

Dissertation

submitted to the

Combined Faculties for the
Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg
Germany

for the degree of
Doctor of Natural Sciences

Put forward by

Dipl.-Phys. Tim Schulze-Hartung

born in Ludwigshafen am Rhein

Date of oral examination: 25 January 2013

Searching and Characterising Exoplanets
using
Astrometry and Doppler Spectroscopy

REFEREES:

PROF. DR. THOMAS HENNING
PROF. DR. ANDREAS QUIRRENBACH

We on Earth have just awakened
to the great oceans of space and time
from which we have emerged.

— Carl Sagan

Zusammenfassung

Für die Suche nach extraterrestrischem Leben wie zur Weiterentwicklung von theoretischen Modellen ist der Nachweis von Exoplaneten von zentraler Bedeutung. Die meisten Planetenkandidaten sind bislang mittels Doppler-Spektroskopie des Zentralsterns entdeckt worden. Die so gemessenen Radialgeschwindigkeiten lassen in Kombination mit Astrometrie Rückschlüsse auf Masse und Charakteristik der Umlaufbahn des Begleitobjekts zu. In der vorliegenden Arbeit werden Modelle zur theoretischen Beschreibung der entsprechenden Observablen hergeleitet. Diese Modelle können unter Annahme eines Fehlermodells mit den Messdaten verglichen werden. Als statistische Methode wird hier der Bayes'sche Ansatz genauer erläutert und weiterverfolgt. Neben der Berücksichtigung von *A-priori*-Wissen erlaubt dieser unter anderem die Ableitung von Wahrscheinlichkeitsdichten der Parameter und den statistisch robusten Nachweis von Exoplaneten. Zur Anwendung der Bayes'schen Methode auf Astrometrie- und Radialgeschwindigkeitsdaten wurde das Computerprogramm BASE stark weiterentwickelt. Nach einer ausführlichen Vorstellung des Tools wird dieses zur Bestimmung der Umlaufbahn des Doppelsterns Mizar A eingesetzt. Hierbei werden frühere Resultate bestätigt und fundierte Aussagen über die Unsicherheiten der Parameter getroffen. In Bezug auf den nahen Stern ϵ Eridani lassen eine Frequenzanalyse und der Bayes-Faktor keine eindeutigen Schlüsse über die Präsenz eines umstrittenen planetaren Begleiters zu. Dies könnte durch stellare Aktivität sowie Eigenschaften der zugrundeliegenden Daten verursacht sein. Eine weitere Untersuchung dieser Effekte erscheint aussichtsreich.

Abstract

The discovery of exoplanets plays a key role both in advancing theoretical models and in the search for extraterrestrial life. Most planet candidates have so far been detected by means of Doppler spectroscopy of their central star. The radial velocities thus measured, in combination with astrometry, allow to draw conclusions on the mass and orbital characteristics of the accompanying object. In this work, models which theoretically describe the corresponding observables are derived. Under the assumption of an error model, these observable models can be compared to the measured data. Here, the Bayesian approach is detailed and pursued as a statistical method. In addition to considering *a priori* knowledge, it allows to derive probability densities over the parameters as well as the statistically robust detection of exoplanets. To apply the Bayesian method to astrometric and radial-velocity data, the computer program BASE has been significantly extended. After presenting the tool in detail, it is employed to determine the orbit of the binary star Mizar A. This leads to a confirmation of earlier results and well-founded statements on the parameter uncertainties. For the nearby star ϵ Eridani, a frequency analysis and the Bayes factor do not support unanimous conclusions about the presence of a controversial planetary companion. This might be caused by stellar activity and properties of the underlying data. Further investigation of these effects seems promising.

Contents

1	The search for exoplanets	13
1.1	Planet formation and statistics	14
1.1.1	From dust to planets	15
1.1.2	Gas giants	15
1.1.3	Statistics and migration	16
1.2	Observational techniques	17
2	Physics and observables	21
2.1	Dynamics and kinematics of planetary systems	21
2.1.1	Stellar motion in the orbital plane	22
2.1.2	Transformation to reference system	23
2.1.3	Relation to the planetary orbit	26
2.1.4	Multiple planets	26
2.2	Observable models	28
2.2.1	Additional observable effects	29
2.2.2	HIPPARCOS intermediate astrometric data	30
2.2.3	Radial velocities	32
2.2.4	Determination of planetary mass	33
2.2.5	Binary systems	34
2.3	Parameters and derived quantities	34
2.4	Errors and noise	38
3	Data analysis	41
3.1	Frequentist inference	41
3.1.1	Likelihood of data	42
3.1.2	Parameter estimation	43
3.1.3	Uncertainty estimation	44
3.1.4	Model selection	44
	Period detection	45
3.2	Bayesian inference	46
3.2.1	Bayes' theorem	46
3.2.2	Encoding prior knowledge	47
3.2.3	Posterior sampling of parameters and derived quantities	48
3.2.4	Marginalisation and density estimation	49
3.2.5	Parameter estimation	50
3.2.6	Uncertainty estimation	50
3.2.7	Model selection	51
3.2.8	Model-uncertainty prediction and observation scheduling	53

4	Informatics and implementation	55
4.1	Requirements	55
4.2	Other software and features of BASE	56
4.3	Modes of operation	58
4.3.1	Normal and binary mode	58
4.3.2	Number-of-planets modes	58
4.3.3	Periodogram mode	58
4.4	Input	58
4.4.1	Invocation and options	59
4.4.2	Data and selection of mode	69
	File format	70
	Data grouping	70
4.4.3	Priors	72
4.5	Program architecture	72
4.5.1	Top-level program flow	73
4.5.2	Posterior sampling by MCMC	74
4.5.3	Improvement of mixing by parallel tempering	76
4.5.4	Assessing convergence by multi-PT	77
4.5.5	Organisation of source code	78
4.6	Specific algorithms	80
4.6.1	Periodogram mode	80
4.6.2	Saving and reading samples	81
4.6.3	Synthetic data	81
5	Bayesian analysis of exoplanet and binary orbits	85
5.1	Introduction	85
5.2	Methods and models	88
5.2.1	Likelihoods and frequentist inference	88
5.2.2	Bayesian inference	90
	Bayes' theorem	90
	Posterior inference	91
5.2.3	Observable models	95
	Stellar motion in the orbital plane	95
	Transformation into the reference system	97
	Relation to the planetary orbit	98
	Observables	98
	Effects of the motion of observer and CM	99
5.3	BASE – Bayesian astrometric and spectroscopic exoplanet detection and characterisation tool	103
5.3.1	Prior knowledge	103
5.3.2	Physical systems and modes of operation	103
5.3.3	Types of data	104
5.3.4	Computational techniques	104
5.4	Target and data	105
5.5	Analysis and results	107
5.5.1	Preparation of RV data	108
5.5.2	Pass A: first constraints on RV parameters	110
5.5.3	Pass B: combining all data	110
5.5.4	Pass C: selecting the frequency f	111

5.5.5	Passes D and E: selecting ω_2 and Ω and refining results	112
5.6	Conclusions	118
5.7	Appendix: Encoding prior knowledge	118
5.8	Appendix: Numerical posterior summaries	119
6	A planet around ϵ Eridani?	123
6.1	Introduction	123
6.2	Previous work	124
6.3	Analysis and results	127
6.3.1	Determination of priors	127
6.3.2	Bayesian periodogram of all data	129
6.3.3	Bayesian periodograms of individual data sets	130
6.3.4	Model selection	134
6.4	Conclusions	134
7	Conclusions	137
	Acknowledgements	151

Chapter 1

The search for exoplanets

Discovering other worlds

The idea that life could exist outside of our Earth has fascinated humankind for countless years. While life has not been unveiled on any of the seven other planets in our Solar System, the investigation has a much wider scope encompassing thousands of stars in the Solar neighbourhood which may host planets. Although this search has been a subject of scientific investigation since the nineteenth century, it is only twenty years ago that the first *extrasolar planet* (or exoplanet) was discovered in orbit around the pulsar PSR B1257+12 (Wolszczan and Frail 1992), followed by the detection of a Jupiter-mass planet orbiting the main-sequence star 51 Pegasi (Mayor and Queloz 1995). Whereas the radiation of a pulsar would probably not allow life as we know it to exist, 51 Pegasi is a host star similar to our Sun. Since that time, more than 800 (partly unconfirmed) extrasolar planet candidates with diverse properties have been unveiled in more than 600 systems, more than 100 of which show signs of multiplicity (Schneider, J. et al. 2011). It has been demonstrated using models for planetary formation and evolution that the diversity in properties of these objects reflect the variety of conditions in which they have formed (Mordasini et al. 2009).

According to the premise that life tends to assume the forms known to us from Earth, particularly that it is based on liquid water, the search increasingly concentrates on the detection of *terrestrial planets*, i.e. rocky, low-mass planets potentially similar to our Earth. Of prime interest are such planets that orbit their parent stars in what is called the *habitable zone*. In this zone – whose scope depends on the stellar properties – planets probably have to reside if most of the H₂O potentially present on their surfaces is to be able to assume the liquid state. Still, atmospheric theory suggests that water surfaces could also be completely frozen in the habitable zone (Boschi et al. 2012).

Thus, characterising the parent stars and orbits of exoplanets only serves to exclude life-hostile environments, while to deliver conclusive evidence for life forms as we know them, it is at least necessary to analyse the planets' atmospheres and find unambiguous *biomarkers* such as O₂ or O₃ (e.g. Seager 2003).

Besides the final question about life, it is interesting how many Earth-like planets exist around the variety of types of stars and how their sizes and masses are distributed. While most previous planet searches have focused on Sun-like stars and mainly found massive *gas giants* similar to Jupiter, which cause the strongest of planetary signals in observational data, stars of different masses and types are increasingly becoming targets of investigation.

Observing low-mass stars is promising because terrestrial planets are easiest to detect around these; the fact that the empirical distribution of planetary minimum mass rises towards lower masses (Butler et al. 2006) might indicate that terrestrial planets are frequent. Furthermore, increasing the sample of planets with well-characterised orbits and properties around stars of types as varied as possible will also help to improve our understanding of planet formation by allowing to better compare predicted planet populations to the actually observed ones (Alibert et al. 2011). Yet in this endeavour, one faces the challenge of detecting and interpreting signals sometimes lower than the level of the disturbing noise which is always present. Moreover, telescope time and thus observational data are always sparse.

In the present work, statistical methods of *detecting* extrasolar planets and *characterising* their orbits in terms of orbital parameters and derived quantities, based on *astrometric* and *radial-velocity* data, are discussed, implemented and applied to such data. Closely related to the orbits of planets are those of binary stars, which are treated analogously.

This work is organised as follows. Current models for planet formation and their observed statistical properties are outlined in section 1.1, while section 1.2 aims to give an overview of today’s observational techniques for exoplanet¹ detection. In chapter 2, models for the signals caused by planetary and binary companions in astrometric and radial-velocity data are derived, followed in chapter 3 by a discussion of data-analysis methods for relating data and models, taking into account noise and prior information on the model parameters. The methods presented are implemented in BASE, a software tool which is described in chapter 4. After validating it in chapter 5 using observational data, BASE helps to assess the presence of putative planets around the star ϵ Eridani in chapter 6. Finally, conclusions are drawn in chapter 7.

1.1 Planet formation and statistics

Planets are believed to be created in a common process together with their parent stars. The formation of these objects takes place in interstellar molecular clouds comprised of gas – predominantly H_2 , but including many other kinds of atoms and molecules – and dust in a variety of species. Due to local density fluctuations, gravitational instabilities arise which may lead to the cloud’s collapse. However, because each cloud has a non-vanishing angular momentum, not all of the mass falls onto the central *protostar*: a protostellar *disk* is also formed in the plane perpendicular to the total angular momentum, possibly accommodating a high fraction of the collapsing cloud’s matter. From solar-mass clouds, a protostellar “embryo” is thought to be created within less than 10^5 yr, while the remainder of mass accretion takes multiples of that time, amounting to about a million years after the initial collapse before the star has reached its final mass and proceeds towards the *main sequence*. At that time, a long process has also begun further out in the disk – the formation of *protoplanetary* bodies, which evolve over tens if not hundreds of millions of years to planets such as those found in our Solar System. The processes of planet formation are still only partly understood, but they may clearly be as manifold as the resulting bodies that have already been detected. Several planet-formation models have been developed, the most widely accepted of which are briefly sketched below. For an in-depth treatment of the matter, the reader is referred to the review article by Papaloizou and Terquem (2006),

¹Although most methods described in section 1.2 apply to any kind of companion to a star, including planets, brown dwarfs, or stars, companions are referred to as “exoplanets” – the least massive case, hence generally the hardest to detect.

as well as to Perryman (2011), who provides a recommendable reference not only in this area, but for contemporary exoplanet science in all its variety.

1.1.1 From dust to planets

The fine dust grains initially present in the protoplanetary disk, with sizes ranging down to below $1\mu\text{m}$, collide and stick together, forming larger bodies between 1 cm and 10 m in diameter. This is accompanied by the dust settling in the disk’s mid-plane. Here, the temporal order of the two processes is just one example of the many aspects of the models still uncertain. Due to their significant interaction with the gas in the disk, the newly formed larger grains experience a force driving them to the centre of the disk on average, causing a time span of only about 100 yr to be available for coagulation. An effect helping to circumvent this difficulty may be *turbulence*, accompanied by local pressure enhancements that facilitate the aggregation of solids.

As the *meter-size barrier* is overcome, *planetesimals* with diameters of order 1 km and more, whose interactions are increasingly determined by gravity, are presumed to form by essentially the same mechanisms as their smaller predecessors, i.e. pairwise collisions and sticking. The increasing role of gravitational interactions then leads to a phase of *runaway growth* between size scales of 10 km and 100 km, where the rate of mass growth increases with time for larger bodies only. This phase is probably followed by an *oligarchic-growth* epoch, in which the larger bodies all grow at a similar rate, dominantly removing material from the influence of their smaller “competitors” and ending up about 1000 km in diameter. Finally, the strong gravitational interactions between these larger bodies become more chaotic, leading to severe disruptions of their Keplerian orbits around the barycentre and allowing for collisions that shape the final planetary system and produce planets of masses similar to that of the Earth.

1.1.2 Gas giants

In contrast to the *rocky planets* discussed above, the existence of *giant planets* such as Jupiter cannot be explained solely by the same mechanisms. Predominantly composed of gaseous matter tens or hundreds of times the mass of the Earth, potentially surrounding a solid core, such objects probably could not be formed from the material found close to the accreting protostar, whereas further out in the disk, the accretion time scales may be too long. Instead, the formation of such *gas giants* is explained today by two competing scenarios: core accretion (which is favoured overall) and gravitational instability.

The *core-accretion* model states that some of the objects produced as rocky planets, with masses of several (tens of) Earth masses, may turn into the cores of newly forming giant planets by gravitationally capturing large amounts of gas and also planetesimals. The developing gaseous envelope further enhances the rate of collection of planetesimals, an increasing fraction of which dissolve in the dense gas. Where conditions are adverse, i.e. the available gas or core mass is insufficient, an *ice giant* such as Uranus and Neptune or no giant planet at all may be formed from the core.

It is suspected that massive cores may only be able to form beyond the so-called *snow line*,² a certain distance from the protostar, where temperatures are so low that various otherwise gaseous species are in the solid phase, thus providing enough matter to form

²For the protoplanetary disk preceding our Solar System, this line has been found to correspond to a distance of 2.7 AU from the centre (e.g. Lecar et al. 2006), which would be intermediate between the orbits of Mars and Jupiter in today’s terms.

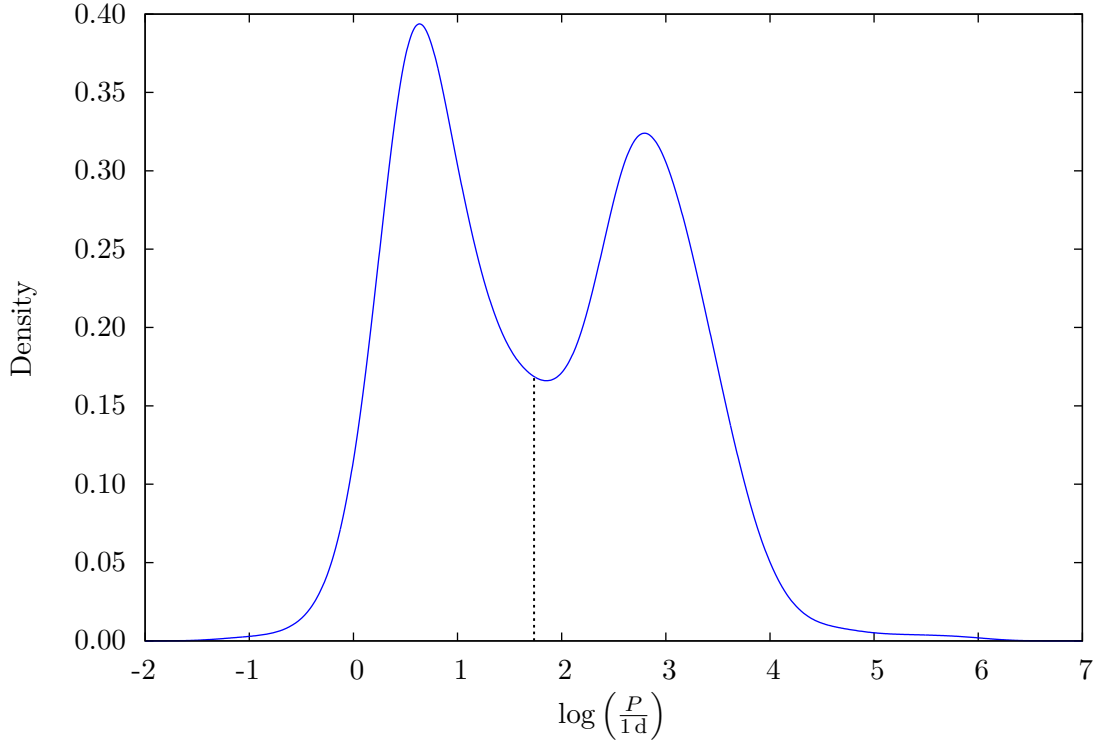


Figure 1.1: Distribution of the periods P listed for 820 planet candidates by Schneider (2012). The dotted line indicates the median period $P = 54.3$ d.

massive cores. There, however, the gas-accretion time scales are longer, possibly conflicting with the onset of dispersion of the gaseous protoplanetary disk by, e.g., *photoevaporation* within about ten million years. While this may let some of the planets end up as ice giants, the formation of gas giants by core accretion can also be modelled over such time scales. One effect thought to accelerate the formation process is the protoplanet’s *migration* (↗ 1.1.3) through parts of the disk.

In the competing model of *gravitational disk instability*, giant planets may be formed by collapsing *fragments* of a dense protoplanetary disk. Taking place over much shorter time scales than the various processes postulated in core accretion, gravitational instability would enable the formation of gas giants well before the disk’s dispersion. Uncertainties of this approach include whether fragmentation can occur at smaller distances from the centre or for less massive protoplanets, and whether it could reproduce the assumed core masses of the giant planets in our Solar System.

While both giant-planet formation scenarios are often viewed as contradictory, it has also been suggested that they may in fact be complementary, with some fraction of the giant planets formed by each of them.

1.1.3 Statistics and migration

The period distribution of the exoplanet candidates detected thus far is clearly bimodal, as is evident from fig. 1.1.³ Its high-period peak is located at $P = 625$ d, corresponding to an

³This graph is based on data compiled at the Extrasolar Planets Encyclopaedia (Schneider 2012). It has been produced using kernel density estimation (↗ 3.2.4), providing a differentiable alternative to the often-used box-shaped histogram.

orbital radius of about 2.6 AU around a solar-mass star (eq. (2.63)), a similar position as that of the suspected snow line in what preceded our Solar System (\nearrow 1.1.2). By contrast, the higher peak is at a period of only $P = 4.31$ d, or a distance of 0.1 AU, corresponding to the so-called *hot Jupiters* unveiled in large numbers: 41% of the putative exoplanets are found to have semi-major axes not exceeding 0.2 AU ($P \lesssim 33$ d around a Sun-like star), while 50% are within 0.5 AU of their host star, equivalent to a period of $P \lesssim 120$ d around the Sun. A fraction of 53% of the candidates within 0.5 AU are at least half as massive as Jupiter.

While the *transit-photometry* technique (\nearrow 1.2) has been yielding an increasing fraction of the candidates, amounting to about a third to date, *Doppler spectroscopy* still accounts for over 50% of the detections. Since each observational technique has its specific biases, the observed distribution of period or other parameters should be treated with some caution. In particular, transit spectroscopy is inclined to detecting lower periods due to the smaller corresponding orbits, half of its detections being associated with periods shorter than 4.2 d, and none exceeding one year. Thus, the transiting planets found cannot contribute significantly to the high-period peak in fig. 1.1. Doppler spectroscopy also has a monotonously increasing bias towards lower-period planets, which cause the stronger signals in measured stellar *radial velocities* when conditions are otherwise unchanged. Nevertheless, about half of its detections are associated with periods exceeding one year. In conclusion, the clear bimodality of the period distribution can be expected to reflect at least to a significant part the physical reality of exoplanet systems.

Although other explanations have been put forward, the observed “pile-up” of hot Jupiters at small orbital radii is often attributed to a presumed *migration* of the giant planets from their birth places further out in the protoplanetary disk (\nearrow 1.1.2) towards the centre. In the vicinity of their host star, they then must be stopped in time by other mechanisms that are still poorly understood. Migration may be caused by the planet interacting with other planets, with planetesimals and/or with the disk’s residual gas, and various hypothetical stopping mechanisms have also been proposed. Although both theories are still a matter of debate, they are perhaps the favoured interpretations to date.

1.2 Observational techniques

Planets outside our own Solar System are regularly discovered at the present time. This is made possible, from an observational point of view, by a variety of different techniques outlined below.

Direct observational methods refer to the *imaging* of exoplanets (e.g. Levine et al. 2009), which reflect the light of their host stars but also emit their own thermal radiation. To overcome the major obstacle of the high brightness contrast between planet and star, techniques such as *coronagraphy* (Lyot 1932; Levine et al. 2009), *angular differential imaging (ADI)* (Marois et al. 2006; Vigan et al. 2010), *spectral differential imaging (SDI)* (Smith 1987; Vigan et al. 2010), and *polarimetric differential imaging* (Kuhn et al. 2001; Adamson et al. 2005) have been invented. Still, imaging has only revealed few detections and orbit determinations so far.

The most productive methods in terms of the number of detected and characterised exoplanets are of an *indirect* nature, observing the effects of the planet on other objects or their radiation.

Of these, *transit photometry and spectroscopy* (e.g. Charbonneau et al. 2000; Seager 2008) are noteworthy because they have helped uncover more than 200 exoplanet candidates,

plus over 2000 still unconfirmed candidates from the Kepler space mission (Koch et al. 2010): small decreases in the apparent visual brightness of a star during the *primary* or *secondary eclipse* point to the existence of a transiting companion, whose spectrum may additionally be inferred by subtracting the target spectra obtained during and outside an eclipse. Such data allow one to determine the ratio of planetary to stellar radius and the orbital inclination as well as the planet’s atmospheric composition and temperature.

Timing methods include measurements of *transit timing variations* (TTV^4) and *transit duration variations* (TDV) (e.g. Holman and Murray 2005; Nascimbeni et al. 2011) of either binaries or stars known to harbour a transiting planet. The method used in the first exoplanet detection (Wolszczan and Frail 1992) is *pulsar timing*, which relies on slight anomalies in the exact timing of the radio emission of a pulsar and is sensitive to planets in the Earth-mass regime.

Microlensing (Mao and Paczynski 1991; Gould 2009), which accounted for about 15 exoplanet candidates, uses the relativistic curvature of spacetime due to the masses of both a lens star and its potential companion, with the latter causing a change in the apparent magnification and thus the observed brightness of a background source.

Perhaps the most well-known technique, and one of those on which this thesis is based, is known as *Doppler spectroscopy* or *radial-velocity* (RV) measurements (e.g. Mayor and Queloz 1995; Lovis and Fischer 2010). With more than 400 exoplanet candidates, it has been most successful in detecting new exoplanets and determining their orbits to date. From a set of high-resolution spectra of the target star, a time series of the line-of-sight velocity component of the star is deduced. These data allow one to determine⁵ the orbit in terms of its geometry and kinematics in the *orbital plane* as well as the minimum planet mass $m_{p,\min} \approx m_p \sin i$. To derive the actual planet mass m_p , the inclination i of the orbit plane with respect to the sky plane needs to be derived with a different method, e.g. astrometry. The RV technique is distance-independent by principle, but signal-to-noise requirements do pose constraints on the maximum distance to a star. Stellar variability sometimes makes this approach difficult because it alters the line shapes and thus mimicks RV variations. The signal in stellar RVs caused by a planet in a circular orbit has a semi-amplitude of approximately

$$K \approx m_p \sin i \sqrt{\frac{G}{m_\star a_{\text{rel}}}}, \quad (1.1)$$

where m_p , m_\star , i , G , and a_{rel} are the masses of planet and host star, the orbital inclination, Newton’s gravitational constant, and the semi-major axis of the planet’s orbit relative to the star, respectively. This approximation holds for $m_p \ll m_\star$, which is true in most cases. It should be noted that the sensitivity of the RV method decreases towards less inclined (more face-on) orbits, which is an example for the selection effects inherent to any planet-detection method.

Finally, *astrometry* (AM; e.g. Gatewood et al. 1980; Sozzetti 2005; Reffert 2009) – on which this work is also based – is the oldest observational technique known in astronomy: a stellar position is measured with reference to a two-dimensional coordinate system attached to the sky plane. The measurements may either be *absolute* (*wide-angle astrometry*), e.g. by using a single interferometer, or *relative* to physical reference stars (*narrow-angle*

⁴For a list of abbreviations used in this article, cf. table 1.1.

⁵Some quantities depend, however, on the knowledge of the stellar mass m_\star , which, using current techniques (e.g. Torres et al. 2010), may be determined up to an uncertainty of about 6% for single (post-)main-sequence stars over $0.6 M_\odot$.

astrometry); in the latter case, the reference may be either given by a binary companion or an unattached star. Alternatively, as is the case in space telescopes such as HIPPARCOS or GAIA, the coordinate system may be defined *globally* by a grid of reference stars spanning all the sky. Astrometry can thus be considered as complementary to Doppler spectroscopy, which measures the kinematics perpendicular to the sky plane, i.e. in the line of sight. In contrast to Doppler spectroscopy, AM allows one to determine the orientation of the orbital plane relative to the sky in terms of its inclination i and the position angle Ω of the *line of nodes*⁶ with respect to the meridian of the target. A planet in circular orbit around its host star displaces the latter on sky with an approximate angular semi-amplitude of

$$\alpha \approx \frac{m_p}{m_\star} \frac{a_{\text{rel}}}{d}, \quad (1.2)$$

where d is the distance between the star and the observer. Again, this approximation holds for $m_p \ll m_\star$.

Imaging astrometry, in its attempt to reach sufficient precision, still faces problems due to various distortion effects. By contrast, *interferometric* astrometry has been used to determine the orbits of previously known exoplanets, mainly with the help of space-borne telescopes such as HIPPARCOS or the Hubble Space Telescope (HST), which presently still excel their Earth-bound competitors (e.g. McArthur et al. 2010). However, instruments like PRIMA (Delplancke et al. 2000; Delplancke 2008; Launhardt et al. 2008) or GRAVITY (Gillessen et al. 2010) at the ESO Very Large Telescope Interferometer are promising to advance ground-based AM even more in the near future.

While planet-induced signals in AM and RVs are both approximately linear in planetary mass m_p , they differ in their dependence on the orbital semi-major axis a_{rel} (eq. (1.1) and (1.2)). Doppler spectroscopy is more sensitive to smaller orbits (or higher orbital frequencies, eq. (2.63)), while AM favours larger orbital separations, viz. longer periods.

Comprehensive reviews of observational methods for exoplanet detection and characterisation can be found in Deeg et al. (2007, chapter 1) and Perryman (2011).

⁶The *line of nodes* is the intersection of the orbital plane with the sky plane.

Table 1.1: Abbreviations used in this work.

Abbreviation	Meaning
ADI	A ngular d ifferential i maging
AM	A strometry
AM _a	A strometry (a ngular positions)
AM _h	A strometry (H IPPARCOS intermediate astrometric data)
API	A pplication p rogramming i nterface
BASE	B ayesian A strometric and S pectroscopic E xoplanet D etection and C haracterisation T ool
BVS	B isector v elocity s pan
CES	C oudé E chelle spectrograph
CFHT	C anada- F rance- H awaii telescope
CM	C entre of m ass
CRC	C yclic r edundancy c heck
CSV	C omma-separated v alues
DFT	D iscrete F ourier t ransform
ESO	E uropean S outhern O bservatory
FAP	F alse- a larm p robability
GOMP	G NU O pen M P
GNU	G NU's N ot U nix
GRAVITY	G eneral R elativity A nalysis via V LT I nterferometry
HARPS	H igh A ccuracy R adial V elocity P lanet S earcher
HIPPARCOS	H igh- p recision p arallax- c ollecting s atellite
HPDI	H ighest p osterior- d ensity i nterval
HST	H ubble S pace T elescope
IRAS	I nfrared A stronomical S atellite
JD	J ulian d ate
LC	L ong C amera
LS	L omb- S cargle
MAP	M aximum a - p osteriori
MCMC	M arkov c hain M onte C arlo
MH	M etropolis- H astings
NASA	N ational A eronautics and S pace A dministration
NLA	N umerical L ebesgue A lgorithm
NPOI	N avy P rototype O ptical I nterferometer
OpenMP	O pen M ulti- P rocessing
PDF	P ortable D ocument F ormat
PRIMA	P hase- R eferenced I maging and M icroarcsecond A strometry
PSR	P otential s cale r eduction
PT	P arallel t empering
RLE	R un- l ength e ncoding
RMS	R oot m ean s quare
RV	R adial v elocity
SB	S pectroscopic b inary
SDI	S pectral d ifferential i maging
SIMBAD	S et of I dentifications, M easurements and B ibliography for A stronomical D ata
SSB	S olar system b arycentre
TDV	T ransit d uration v ariations
TTV	T ransit t iming v ariations
VLC	V ery L ong C amera
VLT	V ery L arge T elescope
VLTI	V ery L arge T elescope I nterferometer
VTA	V olume T esselation A lgorithm

Chapter 2

Physics and observables

Observable effects of Newton's gravity

The motion of two or more bodies in a bound system is responsible for all the deterministic observable effects that are considered in this work. Such systems are either extrasolar planetary systems, where one or more planets and a star orbit each other, or binary systems comprised of two orbiting stars.

In this chapter we derive models for the relevant observables (*observable models*) from orbital kinematics, which are in turn caused by the dynamics of the orbiting bodies. Observable models consist of functions $\mathbf{f}(t; \boldsymbol{\theta})$ of the model parameters $\boldsymbol{\theta}$ and time t which return theoretical values of the observables. These can be compared to measured data by means of the likelihood (section 3.1.1).

In our treatment of the dynamics, we consider only isolated systems, i.e. no external forces are taken into account, and relativistic effects are neglected, implying that all orbits can be considered closed. We begin with the simplest case of a single-planet system, where two bodies of differing masses orbit each other. Since only the star is observable by AM and Doppler spectroscopy, it is the stellar motion which we first describe. Then, because the closely related planetary orbit is of prime interest, we transform the parameters into ones characterising the planet's orbit.

By making several straightforward transformations, the results are then carried over to binary systems, which behave completely analogously for the purposes of this work. Finally, neglecting gravitational interactions between more than two bodies, observable models are derived for multi-planet systems.

An overview of the model parameters used in this work is given in table 2.1, while table 2.2 lists quantities that can be derived from them. For an in-depth treatment of celestial mechanics, the interested reader is referred, e.g., to Moulton (1984).

2.1 Dynamics and kinematics of planetary systems

In the following, we examine the case of an isolated non-relativistic two-body system of star and planet. The stellar motion around the two-body *centre of mass* (CM) is governed by Newton's Law of Gravity,

$$\ddot{\mathbf{r}} = -G\mu \frac{\mathbf{r}}{|\mathbf{r}|^3}, \quad (2.1)$$

where

$$\mu = \frac{m_\star^3}{(m_\star + m_p)^2} \quad (2.2)$$

is the *mass function*, \mathbf{r} is position vector of the star with respect to the CM, m_\star is the stellar mass, and m_p is the planetary mass. We assume that the CM is unaccelerated, viz. no external force acts upon the system, which implies that the reference frame is *inertial*.

2.1.1 Stellar motion in the orbital plane

The general solution corresponds to a motion in a fixed plane – the *orbital plane* – and can be expressed in a polar coordinate system (r, ν) , whose pole coincides with the CM and whose fixed direction is that from the CM to the *periapsis*.¹ Stumpff (1973) found the solution for the radial coordinate to be

$$r = a_\star \frac{1 - e^2}{1 + e \cos \nu} = r(\nu; a_\star, e), \quad (2.3)$$

where a_\star and e are the *semi-major axis* and *orbital eccentricity*, respectively. The angular coordinate $\nu \in [0, 2\pi)$ is known as the *true anomaly*. This equation describes the stationary elliptical *Keplerian orbit* of the star with one *focus* of the ellipse coinciding with the CM.

Since eq. (2.3) implies that r varies in the range $[a_\star(1 - e), a_\star(1 + e)]$ over the course of an orbital revolution, we can define

$$r \equiv a_\star(1 - e \cos E) = r(E; a_\star, e), \quad (2.4)$$

where E is called the *eccentric anomaly*. It follows from eq. (2.3) and (2.4) and a trigonometric half-angle formula that the transformation between true and eccentric anomaly is given by

$$\cos \nu = \frac{\cos E - e}{1 - e \cos E}. \quad (2.5)$$

The time dependence of E is given implicitly by *Kepler's equation*,

$$E - e \sin E = 2\pi f(t - T) = M(t), \quad (2.6)$$

which implies

$$\dot{E} = \frac{2\pi f}{1 - e \cos E}, \quad (2.7)$$

where $f = P^{-1}$ is the *orbital frequency*, P is the *orbital period*, T is the last time the periapsis was passed before the first measurement (known as the *time of periapsis*) and $M(\cdot)$ is the *mean anomaly*, which varies uniformly over the course of an orbit. Kepler's equation is transcendental and can be solved numerically to obtain E for every relevant combination of e and M .²

In the components of position vector $(r, \nu)^\top$ the eccentric anomaly E appears only as an argument to the $\cos(\cdot)$ function (eq. (2.4) and (2.5)). Hence E is equivalent to $E + 2\pi$ and, according to eq. (2.6), M is equivalent to $M + 2\pi$. Consequently, both anomalies E and M can be taken modulo 2π by redefining Kepler's equation as

$$E - e \sin E = 2\pi \cdot \text{mod}(f(t - T), 1) = M(t), \quad (2.8)$$

¹Here, the *periapsis* refers to the stellar position closest to the CM.

²Alternatively, it is possible to solve Kepler's equation using eq. (2.7) and $E|_{t=T} = 0$.

where

$$\text{mod}(x, y) \equiv x - \left\lfloor \frac{x}{y} \right\rfloor y, \quad x, y \in \mathbb{R} \quad (2.9)$$

defines the modulo function, which satisfies for $z \in \mathbb{R}$

$$\text{mod}(zx, zy) = z \text{mod}(x, y). \quad (2.10)$$

Equation (2.8) implies that $E = E(t; f, T)$ is a periodic function of t with period f^{-1} , given f and T .

The time of periapsis T , due to its definition, lies within a range which depends on orbital frequency f . However, T can be transformed into another parameter whose range is simpler to determine. We proceed in a similar approach as Gregory (2005a) by using the alternative χ , defined by

$$\chi \equiv \frac{M(t_r)}{2\pi} = f(t_r - T). \quad (2.11)$$

where t_r is a reference time for the parameters known as the *epoch* (section 2.2). Thus, Kepler's equation becomes

$$E - e \sin E = 2\pi \cdot \text{mod}(\chi + f(t - t_r), 1) = M(t), \quad (2.12)$$

where the *mean anomaly* $M(\cdot)$ varies uniformly over the course of an orbit. By reference to eq. (2.4) and (2.5), it is readily shown that the stellar coordinates are periodic functions of χ with period 1. χ is therefore called a *cyclic* parameter and treated as lying within the range $[0, 1)$.

To express the stellar position in cartesian coordinates, we set up a coordinate system S_1 such that its origin is identical to the CM, its z -axis is perpendicular to the orbital plane and its positive direction chosen such that $\dot{\nu} > 0$ due to the orbital motion, and the vector from the CM to the *periapsis* is orientated in positive x -direction. In S_1 , the stellar barycentric position is given by

$$\mathbf{r}_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = r(\nu; a_\star, e) \begin{pmatrix} \cos \nu \\ \sin \nu \\ 0 \end{pmatrix} \quad (2.13)$$

and, using eq. (2.4) and (2.5),

$$\mathbf{r}_1 = a_\star \begin{pmatrix} \cos E - e \\ \sqrt{1 - e^2} \sin E \\ 0 \end{pmatrix} = \mathbf{r}_1(E; a_\star, e), \quad (2.14)$$

where a_\star is the *semi-major axis*, e is the *eccentricity*, and E is the *eccentric anomaly*.

2.1.2 Transformation to reference system

To derive a model for the stellar barycentric position or velocity, respectively, we transform S_1 to a new coordinate system S_4 by three successive rotations. These are described by Euler angles, termed in our case *argument of the periapsis* ω_\star , *inclination* i , and *position angle of the ascending*³ (or *first*⁴) *node* Ω , and are carried out as follows (fig. 2.1):

³The *ascending node* is the point of intersection of the orbit and the sky plane where the moving object passes away from the observer.

⁴Without RV data, it cannot be determined whether a given node is ascending or descending; then, Ω is defined to be the position angle of the *first* node.

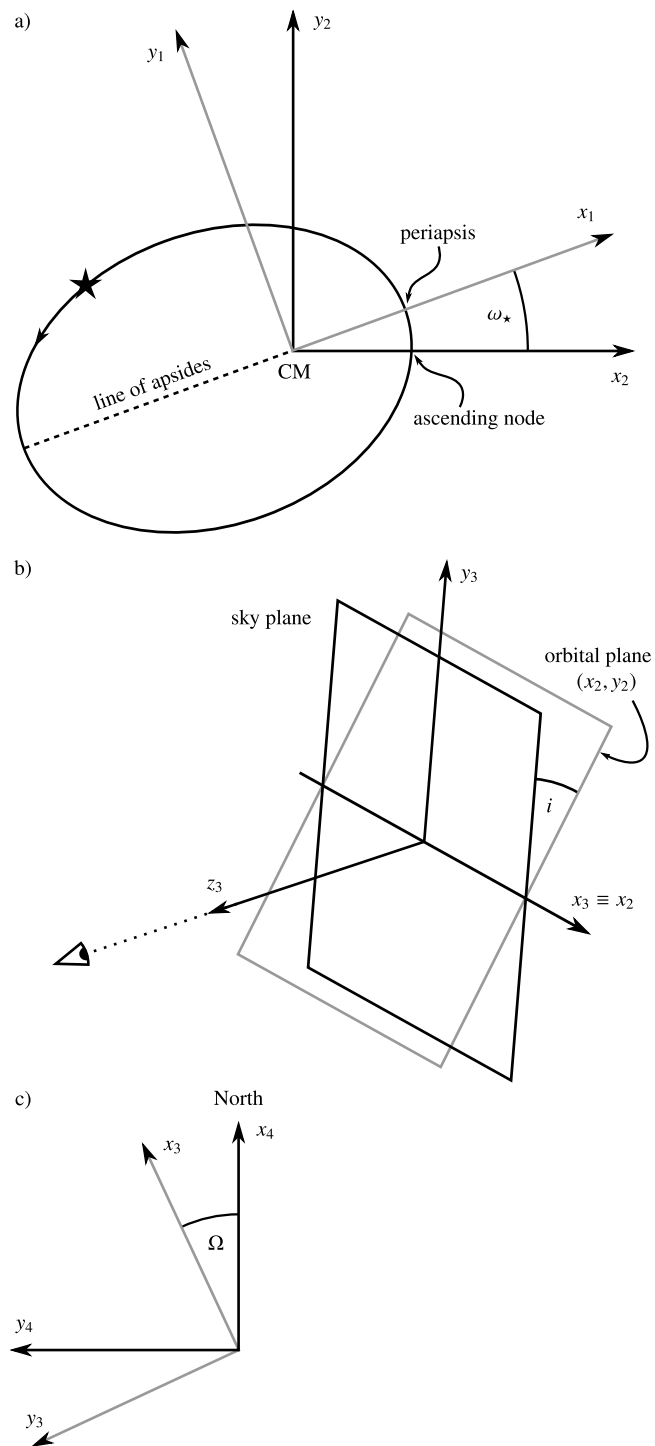


Figure 2.1: Definition of the angles ω_* , i , Ω . a) From S_1 to S_2 , the star and its sense of rotation about the CM are indicated; the dotted line marks the major axis of the orbital ellipse. b) From S_2 to S_3 , the observer and line of sight are indicated. c) From S_3 to S_4 , the positive x_4 -axis points northward along the meridian of the CM.

1. Rotate S_1 about its z_1 -axis by $(-\omega_*)$ such that the ascending node of the stellar orbit lies on the positive x_2 -axis.
2. Rotate S_2 about its x_2 -axis by $(+i)$ such that the new z_3 -axis passes through the observer.⁵
3. Rotate S_3 about its z_3 -axis by $(-\Omega)$ such that the new x_4 -axis is parallel to the *meridian* of the CM and points in a northern direction.

Except for the inclination, the signs of these rotation angles are chosen such that the inverse rotations, leading from the reference system S_4 to the stellar orbit, have positive angles.

Thus, the stellar barycentric position has new coordinates

$$\mathbf{r}_4 = \mathbf{R}_{zzz} \mathbf{r}_1, \quad (2.15)$$

with the passive rotation matrix

$$\mathbf{R}_{zzz} \equiv \begin{pmatrix} A & F & J \\ B & G & K \\ C & H & L \end{pmatrix} \quad (2.16)$$

defining the above rotations; its components are

$$A = \cos \Omega \cos \omega_* - \sin \Omega \cos i \sin \omega_* \quad (2.17)$$

$$B = \sin \Omega \cos \omega_* + \cos \Omega \cos i \sin \omega_* \quad (2.18)$$

$$F = -\cos \Omega \sin \omega_* - \sin \Omega \cos i \cos \omega_* \quad (2.19)$$

$$G = -\sin \Omega \sin \omega_* + \cos \Omega \cos i \cos \omega_* \quad (2.20)$$

$$J = -\sin \Omega \sin i \quad (2.21)$$

$$K = \cos \Omega \sin i \quad (2.22)$$

$$C = -\sin i \sin \omega_* \quad (2.23)$$

$$H = -\sin i \cos \omega_* \quad (2.24)$$

$$L = \cos i. \quad (2.25)$$

$A, B, F,$ and G are known as the *Thiele-Innes constants*, first introduced by Thiele (1883).

By taking the time derivative of eq. (2.15), we obtain the stellar velocity in S_4 ,

$$\mathbf{v}_4 = \dot{\mathbf{r}}_4 = \mathbf{R}_{zzz} \dot{\mathbf{r}}_1 = \dot{E} \mathbf{R}_{zzz} \frac{d\mathbf{r}_1}{dE}, \quad (2.26)$$

where

$$\frac{d\mathbf{r}_1}{dE} = a_* \begin{pmatrix} -\sin E \\ \sqrt{1-e^2} \cos E \\ 0 \end{pmatrix} \quad (2.27)$$

and \dot{E} is given by eq. (2.7).

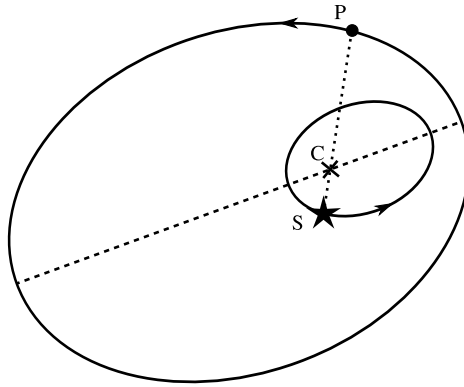


Figure 2.2: The orbits of star S and planet P around the centre of mass C. All three points lie on a common line and the ratio of the lengths of the segments $|CS| : |CP|$ is equal to the mass ratio $m_p : m_*$.

2.1.3 Relation to the planetary orbit

By reference to the above results, the observables of AM and RV are easily derived (section 2.2). They can be parameterised by quantities pertaining to the planetary instead of the stellar orbit based on the following simple relation.

According to the definition of the CM, the line connecting star and planet contains the CM and the ratio of their respective distances from the CM equals the inverse mass ratio,

$$\vec{CS} = -\frac{m_p}{m_*} \vec{CP}, \quad (2.28)$$

where C, S and P stand for CM, star and planet, respectively. This implies invariable relationships between the orbits of the star and the planet as follows. The two bodies orbit their CM with a common orbital frequency f and time of periapsis T . With respect to the corresponding periapsis, they always have the same eccentric anomaly E . Their orbital shapes, viz. eccentricities e , are identical as well, and the orbital semi-major axes relate to each other as

$$a_* = \frac{m_p}{m_*} a_p. \quad (2.29)$$

Additionally, the two bodies share the same sense of orbital revolution, hence those nodes of both orbits which lie on the positive x_2 -axis are ascending. Consequently, the only Euler angle not shared by the stellar and planetary orbits is the argument of periapsis, which differs by π because star and planet are in opposite directions from the CM.

2.1.4 Multiple planets

In the presence of multiple planets, an *n*-body problem with $n > 2$ arises, for which no exact and general solution is known. Instead, several approaches exist which either give an exact solution for special cases or an approximate (general) solution, e.g. by numerical integration or in the form of a truncated Taylor series. Numerical integration can be an important tool especially in systems with interaction between the planets, leading to possibly unstable systems where bodies may collide or be ejected from the system.

By contrast, it is assumed in this work that any interactions between the planets can be neglected, implying that their orbits around the star can be treated separately, as derived

⁵A positive rotation angle is used here to ensure that the node is indeed ascending (fig. 2.1 b).

in the following. Furthermore, only the case where these orbits are coplanar, i.e. all bodies move in the same orbital plane, is considered.

For any number n_p of planets, but particularly when $n_p > 1$, the position of the common CM of all bodies is given by

$$\overrightarrow{OC} = \frac{m_\star \overrightarrow{OS} + \sum_{j=1}^{n_p} m_{p,j} \overrightarrow{OP}_j}{m_{\text{tot}}} \quad (2.30)$$

$$= \frac{m_\star \overrightarrow{OS} + \sum_j m_{p,j} (\overrightarrow{OS} + \overrightarrow{SP}_j)}{m_{\text{tot}}} \quad (2.31)$$

$$= \overrightarrow{OS} + \frac{\sum_j m_{p,j} \overrightarrow{SP}_j}{m_{\text{tot}}}, \quad (2.32)$$

from which the barycentric position of the star is immediately obtained as

$$\overrightarrow{CS} = \sum_j \frac{m_{p,j}}{m_{\text{tot}}} \overrightarrow{P}_j \overrightarrow{S}. \quad (2.33)$$

In the above equations, O is the coordinate origin, P_j is the position of the j th planet and $m_{\text{tot}} \equiv m_\star + \sum_{j=1}^{n_p} m_{p,j}$.

Assuming that the total planetary mass in the system is much lower than the stellar mass,

$$\sum_{j=1}^{n_p} m_{p,j} \ll m_\star, \quad (2.34)$$

eq. (2.33) may be approximated as

$$\overrightarrow{CS} \approx \sum_{j=1}^{n_p} \frac{m_{p,j}}{m_\star + m_{p,j}} \overrightarrow{P}_j \overrightarrow{S}. \quad (2.35)$$

because

$$m_{\text{tot}} \approx m_\star + m_{p,j} \quad \forall j \in \{1, \dots, n_p\}. \quad (2.36)$$

We furthermore define the two-body barycentre C_j of star and j th planet by specialising eq. (2.30) to $n_p = 1$ and thereby obtain

$$\overrightarrow{CS} \approx \sum_{j=1}^{n_p} \overrightarrow{C}_j \overrightarrow{S}. \quad (2.37)$$

Due to eq. (2.37), the stellar barycentric position \mathbf{r}_1 in S_1 can be approximated by the sum of the stellar positions $\mathbf{r}_1^{(j)}$ induced by the individual planets j , viz.

$$\mathbf{r}_1 \approx \sum_{j=1}^{n_p} \mathbf{r}_1^{(j)}. \quad (2.38)$$

As all the bodies orbit in one common plane, the second and third rotations in section 2.1.2 are defined by the same inclination i and position angle of the ascending node Ω for all planets and only the arguments of periapsis ω_j differ. Thus, eq. (2.15) gives the position of the star in S_4 as

$$\mathbf{r}_4 = \sum_{j=1}^{n_p} \mathbf{R}_{zz}(\omega_j, i, \Omega) \mathbf{r}_1, \quad (2.39)$$

with \mathbf{R}_{zz} as defined by eq. (2.16).

2.2 Observable models

In the following, the two types of astrometric data treated in this work are abbreviated respectively as AM_a (angular positions) and AM_h (HIPPARCOS intermediate data), while AM denotes astrometric data in general.

To express the stellar barycentric position as a two-dimensional angular position, we perform a final transformation of S_4 into a spherical coordinate system S_5 with radial, elevation and azimuthal coordinates $(r, \delta, \alpha)^T$. Its origin is identical with the observer, its reference plane coincides with the (y_4, z_4) -plane and its fixed direction is $-z_4$, pointing from the observer to the CM. In S_5 , the radial coordinate of the CM equals a distance d , which relates to the *parallax* ϖ as

$$d = \frac{1 \text{ AU}}{\varpi}. \quad (2.40)$$

The distance is assumed constant in the following and the radial coordinate in S_5 omitted. In the new system, the two-dimensional angular barycentric position of the star is obtained as

$$\mathbf{r}_5 \equiv \begin{pmatrix} \delta_5 \\ \alpha_5 \cos \delta_5 \end{pmatrix} = \frac{1}{d} \begin{pmatrix} x_4 \\ y_4 \end{pmatrix} \quad (2.41)$$

$$= a' \left((\cos E - e) \begin{pmatrix} A \\ B \end{pmatrix} + \sqrt{1 - e^2} \sin E \begin{pmatrix} F \\ G \end{pmatrix} \right) \quad (2.42)$$

with

$$a' \equiv \varpi a_* \cdot 1 \text{ AU}^{-1} \quad (2.43)$$

in S_5 , where the first coordinate $r \equiv d$ has been omitted, δ is called the *declination* and α is called the *right ascension*. The factor $\cos \delta$ in eq. (2.41) stems from the fact that the line element in spherical coordinates, for $r = \text{const}$, is $d\mathbf{r} = r(\mathbf{e}_\delta d\delta + \mathbf{e}_\alpha \cos \delta d\alpha)$, where \mathbf{e}_δ and \mathbf{e}_α are the unit vectors in the local directions of increasing δ and α , respectively. The coordinate $\alpha \cos \delta$ is often abbreviated α_* .

In binary systems, observations can use the binary companion as point of reference and coordinate origin and may therefore be described by eq. (2.42) with only minor modifications (\nearrow 2.2.5). In particular, because $\delta \ll 1$ rad, the factor $\cos \delta \approx 1$ can be omitted and the second coordinate is simply α .

By contrast, single-star planetary systems lack an observable object near the CM in general. Therefore, one or more physically unattached stars or a set of stars need to be referred to when observing the latter type of systems. In the following, we derive the stellar coordinates in an independent coordinate system S_6 assumed to be fixed to an inertial frame

as the previous coordinate systems $S_{1..5}$. We first assume that the coordinates $(\delta, \alpha_*)^\top$ in S_6 of the fixed direction of S_5 change linearly in time, with the rate of change being

$$\boldsymbol{\mu} \equiv \begin{pmatrix} \mu_\delta \\ \mu_{\alpha_*} \end{pmatrix} \quad (2.44)$$

and the values at a given reference time or *epoch* t_r being

$$\mathbf{r}_r \equiv \begin{pmatrix} \delta_r \\ \alpha_r \cos \delta_r \end{pmatrix}. \quad (2.45)$$

Thus, the angular position of the star in S_6 results as

$$\mathbf{r}_6 \equiv \begin{pmatrix} \delta_6 \\ \alpha_6 \cos \delta_6 \end{pmatrix} = \mathbf{r}_5 + \mathbf{r}_r + (t - t_r)\boldsymbol{\mu}. \quad (2.46)$$

In general, the epoch t_r is the instant at which such parameters as δ_r or α_r are defined – parameters that represent the values of time-variable quantities. In particular, if we momentarily ignore the orbital motion by setting $\delta_5 \equiv \alpha_5 \equiv 0$, the position \mathbf{r}_6 at $t = t_r$ is identical to \mathbf{r}_r . In this work, all parameters referring to the following aspects of the observable models are assumed to be constant and therefore not associated with an epoch:

- motion in the orbital plane and orbital shape
- coordinate transformations to systems $S_{2..5}$
- distance of CM
- rates of linear changes of astrometric position and radial velocity.

2.2.1 Additional observable effects

In the following, we list several physical effects which may be important when analysing the types of data considered in this work. However, most of them only affect the analysis of AM data, therefore those also concerning RV data are explicitly pointed out.

- *Light deflection*, a relativistic effect where the masses of objects modify the apparent position of a target, is not considered in this work.
- The following classical effects are related to the finite speed of light and can be treated in combination as *planetary aberration* (Green 1985):
 - *Aberration* designates the change of the apparent target position as seen from Earth due to Earth’s orbital velocity of about $10^{-4}c$, where c is the speed of light.
 - *Light-time correction* takes into account the target’s motion during the time taken for light to reach the observer and affects both AM and RV data.

In this work, planetary aberration is neglected because aberration is irrelevant for AM_a data of binaries and already removed from AM_h data (van Leeuwen 2007, chapter 2.5.3), and because light-time correction cannot be carried out reliably due to the uncertainty in the target’s distance.

- *Annual parallax* is the shift in the apparent position of the target due to Earth's orbital motion around the SSB. For wide-angle or global AM, this effect needs to be accounted for according to the absolute parallaxes ϖ of the relevant bodies, and it is therefore included in the HIPPARCOS model. By contrast, *relative* parallaxes of the observed bodies may be referred to in narrow-angle AM. For AM_a measurements of the relative angular positions of binary components situated (approximately) at the same distance from the observer, annual parallax can be neglected (see e.g. section 5.2.3). In RV data, a correction for the effect of Earth's motion has usually already been applied (section 2.2.3).
- *Perspective secular changes* in certain parameters pertaining to AM and RV data are due to the relative motion of the target with respect to the SSB and, correspondingly, a change in direction towards the target. For a linear motion, the temporal changes in parallax ϖ , proper motion $\boldsymbol{\mu}$ and radial velocity of the CM v are given by (van de Kamp 1977):

$$\dot{\varpi} = -v\varpi^2 \cdot 1 \text{ AU}^{-1} \quad (2.47)$$

$$\dot{\boldsymbol{\mu}} = -2v\boldsymbol{\mu}\varpi \cdot 1 \text{ AU}^{-1} \quad (2.48)$$

$$\dot{v} = \boldsymbol{\mu}^2\varpi^{-1} \cdot 1 \text{ AU}, \quad (2.49)$$

where both $\dot{\varpi}$ and $\dot{\boldsymbol{\mu}}$ change their signs from positive to negative at *perihelion*, whereas $\dot{v} > 0$ always. Assuming values of $\varpi = 0.02''$, $\boldsymbol{\mu} = 0.1'' \text{ yr}^{-1}$, and $v = 20 \text{ km s}^{-1}$ – which are near the median values given in the HIPPARCOS catalogue (Perryman et al. 1997) for stars with $\varpi \geq 0.01''$ and the median of the absolute value of radial velocities given in Chubak et al. (2012, table 3), respectively – the resulting rates of change are

$$\dot{\varpi} = -0.0082 \mu\text{as yr}^{-1} \quad (2.50)$$

$$\dot{\boldsymbol{\mu}} = -0.082 \mu\text{as yr}^{-2} \quad (2.51)$$

$$\dot{v} = +0.011 \text{ m s}^{-1} \text{ yr}^{-1}. \quad (2.52)$$

Expressed as relative rates of change, this gives

$$|\dot{\varpi} \varpi^{-1}| = 4.09 \times 10^{-7} \text{ yr}^{-1} \quad (2.53)$$

$$|\dot{\boldsymbol{\mu}} \boldsymbol{\mu}^{-1}| = 8.18 \times 10^{-7} \text{ yr}^{-1} \quad (2.54)$$

$$|\dot{v} v^{-1}| = 5.75 \times 10^{-8} \text{ yr}^{-1}. \quad (2.55)$$

Due to their typically small values, perspective changes are neglected in this thesis. Nevertheless, a linear acceleration a_v of radial velocity may be allowed (\nearrow 2.2.3), which may include the influence of an outer, unmodelled planet and/or perspective acceleration.

2.2.2 Hipparcos intermediate astrometric data

Instruments aboard the HIPPARCOS satellite measured one-dimensional positions of 118,300 stars along *great circles* on the sky. Observations were taken simultaneously in two fields of view, separated by a basic angle of 58° . The satellite was spinning with a frequency of 11.25 d^{-1} , maintaining a fixed solar aspect angle $\xi = 43^\circ$ between the spin axis and the direction of the Sun in order to improve thermal stability; its spin axis was made to precess

around the direction of the Sun. Due to the motion of the satellite, the stellar images moved over a detector and their transit times, the basic data of the mission, were measured. These were then translated into positions, or *abscissae*, along well-defined *reference great circles* on sky. In this “great-circle reduction”, each abscissa was based on up to five successive rotations of the satellite performed during the time span of ca. 10.7 h, i.e. one orbital revolution. The great-circle reduction included reconstruction of the satellite’s along-scan attitude, or *scan phase*. Afterwards, all abscissae were combined in one common solution (“sphere reconstruction”) and the five AM parameters $\delta, \alpha, \mu_\delta, \mu_{\alpha_*}, \varpi$ and their uncertainties were estimated for each observed star (“astrometric parameter solution”). This led to the original HIPPARCOS catalogue published by Perryman et al. (1997).

In the new reduction by van Leeuwen (2007), the basic data were combined in a different way, including the following aspects:

- The processes involved in the great-circle reduction were decoupled from each other;
- because the published AM parameters could be used as reference values, their errors could be assumed uncorrelated and following a distribution with zero mean and finite variance, which was not previously possible;
- drifts of the basic angle were identified and corrected in 0.7% of the satellite’s orbits;
- in reconstructing the scan phase of the satellite, a different model was used and different weights were assigned to the two fields of view;
- discontinuities in the scan phase and attitude of the satellite were identified and corrected.

While the original procedure involved projecting all the transits of an orbit onto one common reference great circle, the new reduction combined only the data of one field transit each, resulting in up to five abscissae per orbit, whose accuracy excelled that of their predecessors. The definitions of several relevant angles are according to fig. 2.3.

The abscissa residuals are determined as follows. The a-priori reference position \mathbf{r}_r of the target is converted to a reference abscissa $a_{h,r}$ by projecting it onto the scan circle. Then, the difference $\Delta a_h \equiv a_h - a_{h,r}$ between the measured and the reference abscissae gives the abscissa residual.

Following van Leeuwen (2007), the abscissa residuals can be modelled as:

$$\Delta a_h = \sin \psi (\delta(t) - \delta_h(t)) + \cos \psi (\alpha_*(t) - \alpha_{*,h}(t)) + \gamma (\varpi - \varpi_h), \quad (2.56)$$

with model position $(\delta(t), \alpha_*(t))^T \equiv \mathbf{r}_6$ (eq. (2.46)) and reference positions

$$\delta_h(t) = \delta_h + \mu_{\delta,h}(t - t_h) \quad (2.57)$$

$$\alpha_{*,h}(t) = \alpha_{*,h} + \mu_{\alpha_{*,h}}(t - t_h). \quad (2.58)$$

Here, ψ is the scan-orientation angle, $\delta_h, \mu_{\delta,h}, \alpha_{*,h}, \mu_{\alpha_{*,h}}, \varpi_h$ are AM reference parameters (defined with respect to the HIPPARCOS mean and reference epoch $t_h = \text{J1991.25}$), ϖ is the parallax, and γ is the time-dependent *parallax factor* allowing to account for the annual parallax. The values of the reference parameters and the parallax factors for all data are part of the intermediate-data files in van Leeuwen (2007). Thus, the combination of eq. (2.46), (2.56), (2.57) and (2.58) yields the model equation for AM_h data, $\Delta a_h = \Delta a_h(t; \boldsymbol{\theta})$.

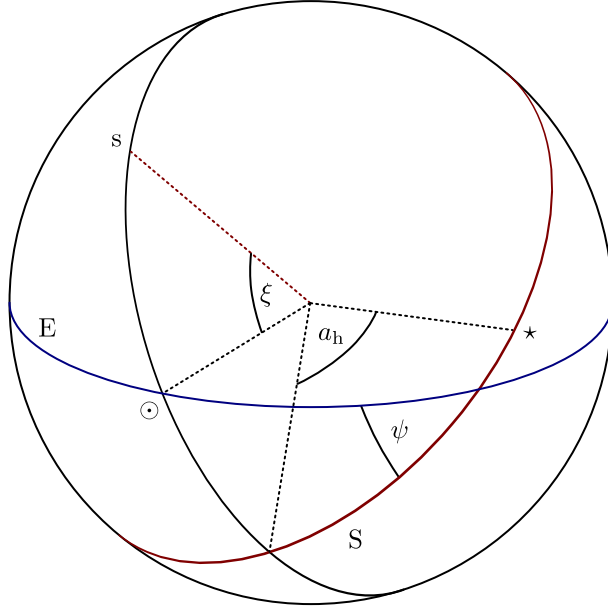


Figure 2.3: Definition of angles for the HIPPARCOS instrument. Shown are the position of the Sun \odot , the celestial equator E , the target star \star , the instantaneous spin axis s and, perpendicular to it, the scan circle S ; labeled angles are the solar aspect angle ξ , the scan-orientation angle ψ and the abscissa a_h .

2.2.3 Radial velocities

In contrast to AM, RV data are usually automatically transformed into an inertial frame resting with respect to the SSB (e.g. Lindegren and Dravins 2003), which allows the Earth's motion to be neglected in this model and treats the observer's rest frame as inertial. The model function for the stellar radial velocity measured by an observer is thus given by

$$v(t; \boldsymbol{\theta}) = -(\mathbf{v}_4)_z + V + (t - t_r)a_v, \quad (2.59)$$

with

$$(\mathbf{v}_4)_z = (\dot{\mathbf{r}}_4)_z = -\sum_{j=1}^{n_p} K_j \frac{\sin E_j \sin \omega_j - \sqrt{1 - e_j^2} \cos E_j \cos \omega_j}{1 - e_j \cos E_j}, \quad (2.60)$$

where V is the reference RV⁶, a_v is a constant RV acceleration which may account for the linear component of a potential perspective secular change of radial velocity (\nearrow 2.2.1) and/or unmodelled low-period, outer planets, $j \in [1, n_p]$ is the planet's index – omitted in the following for brevity – and n_p is the number of planets, which are assumed to be non-interacting (section 2.1.4). Further, the *RV semi-amplitude* K can be expressed as

$$K = 2\pi f a_\star \sin i \quad (2.61)$$

$$= \sqrt[3]{2\pi G f} \frac{m_p \sin i}{(m_p + m_\star)^{\frac{2}{3}}}. \quad (2.62)$$

The last equality holds because of *Kepler's third law*,

$$(a_p + a_\star)^3 f^2 = \frac{G}{4\pi^2} (m_p + m_\star) \quad (2.63)$$

⁶The reference RV consists of the radial velocity of the CM plus an offset due to the specific calibration of the instrument (\nearrow 2.4). An independent parameter V_i is therefore used for each data set (table 2.1).

and the definition of the CM (eq. (2.28)). Owing to eq. (2.62), only one of a_* , K needs to be employed in the AM and RV models; here we adopt K .

It should be noted that a different version of eq. (2.59) involving ν instead of E and the alternative definition $K_{\text{alt}} \equiv \frac{K}{\sqrt{1-e^2}}$ (table 2.2) is often found the literature (e.g. Gregory 2005a).

2.2.4 Determination of planetary mass

In the following, we assume that a given RV semi-amplitude $K > 0$ and frequency f have been estimated from RV data and that the stellar mass m_* is also known exactly, but that the orbital inclination i is unknown. Then, the quantity

$$f_m \equiv \frac{m_p \sin i}{(m_p + m_*)^{\frac{2}{3}}} = \frac{K}{\sqrt[3]{2\pi G f}}, \quad (2.64)$$

whose cube f_m^3 is known as the *mass function*, is also given. From it, the minimum planetary mass $m_{p,\text{min}}$ can be determined as follows. Rewriting eq. (2.64) as

$$m_p \sin i = f_m (m_p + m_*)^{\frac{2}{3}}, \quad (2.65)$$

it is, first, clear that the value of $m_p \sin i$ is not strictly known unless the planetary mass m_p is also known. As both sides of eq. (2.65) depend on m_p , and assuming that a minimum mass $m_{p,\text{min}}$ exists which satisfies this equation, it follows that

$$\min_{m_p} (m_p \sin i) = \min_{m_p} \left(f_m (m_p + m_*)^{\frac{2}{3}} \right) = f_m (m_{p,\text{min}} + m_*)^{\frac{2}{3}}. \quad (2.66)$$

Moreover, rewriting eq. (2.65) as

$$\sin i = \frac{f_m (m_p + m_*)^{\frac{2}{3}}}{m_p}, \quad (2.67)$$

we have

$$\frac{d \sin i}{d m_p} = - \frac{f_m (m_p + 3m_*)}{3m_p^2 \sqrt[3]{m_p + m_*}} < 0. \quad (2.68)$$

With $i \in (0, \pi)$, this implies

$$1 = \max_{m_p} (\sin i) = \sin i |_{m_p = m_{p,\text{min}}}, \quad (2.69)$$

i.e. the minimum mass $m_{p,\text{min}}$ is reached for edge-on orbits, $i = \frac{\pi}{2}$. Insertion of the latter values into eq. (2.65) and comparison with eq. (2.66) implies

$$m_{p,\text{min}} = f_m (m_{p,\text{min}} + m_*)^{\frac{2}{3}} = \min_{m_p} (m_p \sin i), \quad (2.70)$$

which can be solved numerically for $m_{p,\text{min}}$, given f_m and m_* (table 2.2).

Despite the conceptual difference between $m_p \sin i$ and $m_{p,\text{min}}$, it should not go unnoticed that one may write, using $m_* \gg m_p$,

$$m_{p,\text{min}} \approx f_m m_*^{\frac{2}{3}} \approx m_p \sin i, \quad (2.71)$$

an approximation often made implicitly in the literature. The relative error of this approximation is less than 1% even for a 10 M_J planet and thus smaller than the usual uncertainty of several percent in m_\star ($\gtrsim 1.2$), which has been neglected above.

If, by contrast, the inclination i is known, e.g. through the availability of AM data, then the planetary mass m_p itself can be determined. This is achieved by rewriting eq. (2.65) as

$$m_p = \frac{f_m}{\sin i} (m_p + m_\star)^{\frac{2}{3}}, \quad (2.72)$$

which contains only (albeit not exactly) known quantities besides m_p and thus can be solved numerically for m_p (table 2.2).

2.2.5 Binary systems

If the *primary* and *secondary* binary components assume the roles of star and planet, respectively, the above reasoning also yields the observables of a binary system.

For visual binaries, AM_a measurements often refer to the position of the secondary with respect to the primary, implying that proper motion $\boldsymbol{\mu}$ and constant AM offset \mathbf{r}_r can both be ignored. From the definition of the CM, it follows that the orbit of the secondary with respect to the primary is identical with its barycentric orbit but scaled by a factor $(m_1 + m_2)m_1^{-1}$, or with the semi-major axis equaling the sum of the two components' barycentric semi-major axes,

$$a_{\text{rel}} = a_1 + a_2. \quad (2.73)$$

Thus, eq. (2.42) can be used as AM_a model function $\mathbf{r}(t; \boldsymbol{\theta})$ for a binary with a_{rel} replacing a_\star , $\omega_2 + \pi$ replacing ω_\star and

$$a'_{\text{rel}} \equiv \frac{\varpi a_{\text{rel}}}{1 \text{ AU}}. \quad (2.74)$$

Equation (2.59) yields the RV of component i if K is replaced by $(-1)^{i+1}K_i$ and ω by ω_2 . If AM_a and RV data are combined, parameters $a_{1,2}$ can be omitted in favour of $K_{1,2}$ (eq. (2.61)) and a_{rel} is given by the equivalent of eq. (2.62) in combination with eq. (2.73).

2.3 Parameters and derived quantities

For several of the parameters listed in table 2.1, the default prior ranges may be determined from the data. The ranges of the derived quantities listed in table 2.2 follow from corresponding parameter ranges, but need not be explicitly specified as derived quantities are sampled implicitly with the parameters (section 3.2.3). The automatic default ranges $I_{\text{pri},\theta}$ for parameters θ are as follows:

- reference radial velocity V :
 - normal mode:

$$I_{\text{pri},V} = \tilde{\Delta}v_i. \quad (2.75)$$

- binary mode: section of the RV ranges $\tilde{\Delta}v_i$ per RV data set i , with each range extended by one corresponding measurement uncertainty on both bounds, i.e.

$$\tilde{\Delta}v_i = \max_{j,k \in [1, N_D]} (v_j + \sigma_j - (v_k - \sigma_k)), \quad (2.76)$$

where N_D is the number of data, $\{v_j\}$ are the measurements and $\{\sigma_k\}$ are their uncertainties, all pertaining to data set i . Thus,

$$I_{\text{pri},V} = \bigcap_i \tilde{\Delta}v_i. \quad (2.77)$$

- radial-velocity acceleration:

$$I_{\text{pri},a_v} = \left[-\min_i \frac{|\tilde{\Delta}v_i|}{\Delta t_i}, \min_i \frac{|\tilde{\Delta}v_i|}{\Delta t_i} \right], \quad (2.78)$$

where Δt_i is the time span covered by data set i . This means that the RV acceleration is assumed not to exceed the minimum rate of RV change of any data set i determined heuristically according to $|\tilde{\Delta}v_i| \Delta t_i^{-1}$.

- radial-velocity jitter σ_+ :

$$I_{\text{pri},\sigma_+} = [0, 3 \text{ stdev}(\{v_j\})], \quad (2.79)$$

where $\{v_j\}$ is the set of all RVs.

- astrometric jitter τ_+ :

$$I_{\text{pri},\tau_+} = [0, \tau_{+, \text{max}}], \quad (2.80)$$

where

$$\tau_{+, \text{max}} \equiv 3 \cdot \begin{cases} \text{stdev}(\{\Delta a_{h,j}\}), & \text{AM}_h \text{ data} \\ \sqrt{\frac{1}{N_D} \sum_j (\alpha_{*,j}^2 + \delta_j^2)}, & \text{AM}_a \text{ data.} \end{cases} \quad (2.81)$$

The expression used for AM_a data corresponds to three standard deviations of the Euclidean distances of the binary components, calculated with respect to a sample mean identified with zero.

- position angle of the ascending or first node Ω : the prior range of Ω is reduced from the default $[0, 2\pi)$ (*ascending node*) to $[0, \pi)$ if no RV data are provided, in which case Ω is defined as referring to the *first* node (section 2.1.2).
- radial-velocity semi-amplitude K, K_i : while the lower bound $K_{\text{min}} \equiv 0$, the determination of the upper bound varies:
 - normal mode: to determine the upper bound, the following alternatives⁷ have been found: one may either use the physical argument of a maximum allowed “projected mass” $m_p \sin i$,

$$K_{\text{max}} = \sqrt[3]{2\pi G f_{\text{max}}} \frac{(m_p \sin i)_{\text{max}}}{m_*^{\frac{2}{3}}} \quad (2.82)$$

according to eq. (2.62), or the extended RV range $\tilde{\Delta}v$ calculated over all sets,

$$K_{\text{max}} = \tilde{\Delta}v \cdot \frac{\sqrt{1 - e_{\text{min}}^2}}{2}, \quad (2.83)$$

where the latter is based on the fact that for a given K and e , the RV-model range is $2K(1 - e^2)^{-\frac{1}{2}}$.

⁷In BASE, the variant can be chosen with an option (section 4.4.1).

◦ binary mode:

$$K_{i,\max} = \tilde{\Delta}v_i \cdot \frac{\sqrt{1 - e_{\min}^2}}{2}. \quad (2.84)$$

Table 2.1: Parameters of the models in section 2.2.

Symbol ⁸	Designation	Unit	Widest prior support	Cyclic ⁹	Prior ¹⁰	Used with data types				
						AM _a	AM _h	RV	AM _a +RV	AM _h +RV
V	reference RV ¹¹	m s^{-1}	—		U			×	×	×
a_v	RV acceleration	m s^{-1}	—		S			×	×	×
σ_+	magnitude of additional RV noise	m s^{-1}	\mathbb{R}_0^+		M			×	×	×
ϖ	parallax ¹²	arcsec	$(0, 0.77]$		J		×		×	×
α_r	reference right ascension	°	$[0, 360]$	×	U		×			×
δ_r	reference declination	°	$[-90, 90]$		U		×			×
μ_{α^*}	proper motion in $\alpha \cos \delta$	mas yr^{-1}	—		S		×			×
μ_δ	proper motion in δ	mas yr^{-1}	—		S		×			×
τ_+	magnitude of additional AM noise	mas	\mathbb{R}_0^+		M	×	×		×	×
i	inclination	rad	$[0, \pi)$		U	×	×		×	×
Ω	position angle of the ascending (or first) node ¹³	rad	$[0, 2\pi)$	×	U	×	×		×	×
e	eccentricity	1	$[0, 1)$		U	×	×	×	×	×
f	orbital frequency	d^{-1}	$(0, 10]$		J	×	×	×	×	×
χ	mean anomaly at t_r over 2π	1	$[0, 1)$	×	U	×	×	×	×	×
ω, ω_2	argument of periapsis	rad	$[0, 2\pi)$	×	U	×	×	×	×	×
K, K_i	RV semi-amplitude	m s^{-1}	\mathbb{R}_0^+		M			×	×	×
a'	semi-major axis of stellar barycentric orbit over distance ¹⁴	mas	$[10^{-3}, 10^{-5}]$		J	×	×			
a'_{rel}	semi-major axis of orbit of secondary around primary over distance ¹⁴	mas	$[10^{-3}, 10^{-5}]$		J	×				

⁸Parameters printed in boldface pertain to each planet individually in multi-planet mode.

⁹For cyclic parameters θ , the indicated lower and upper bounds are treated as equivalent (section 2.1.1).

¹⁰The default prior type; abbreviations: U (uniform), J (Jeffreys), M (modified Jeffreys), S (signed modified Jeffreys).

¹¹One instance of this parameter, V_i , is employed per RV data set in order to allow for differing offsets (section 2.4).

¹²Prior support includes trigonometric parallax of the nearest star, Proxima Centauri (Perryman et al. 1997).

¹³For details see section 2.1.2.

¹⁴Lower bound corresponds to AM measurement uncertainty of $1 \mu\text{as}$; upper bound according to wide-binary observations by Tolbert (1964).

Table 2.2: Quantities derived from model parameters.

Definition	Designation	Unit	Eq.
$P \equiv f^{-1}$	period	d	—
$T \equiv t_r - \frac{\chi}{f}$	time of periapsis	d	(2.11)
$m_{p,\min} \equiv K(2\pi Gf)^{-\frac{1}{3}}(m_{p,\min} + m_\star)^{\frac{2}{3}}$	minimum planetary mass ¹⁵	M_J	(2.64), (2.70)
$m_p \equiv K(2\pi Gf)^{-\frac{1}{3}}(\sin i)^{-1}(m_p + m_\star)^{\frac{2}{3}}$	planetary mass ¹⁶	M_J	(2.72)
$m_j \equiv \frac{4\pi^2}{G} \frac{K_{3-j}(K_1+K_2)^2}{f(2\pi \sin i)^3}$	binary-component mass	M_\odot	(2.61), (2.63)
$\rho \equiv \frac{m_2}{m_1} = \frac{K_1}{K_2}$	mass ratio of binary components	1	(2.29), (2.61)
$K_{\text{alt}} \equiv \frac{K}{\sqrt{1-e^2}}$	alternative RV semi-amplitude ¹⁷	m s^{-1}	—
$a_\star \equiv \frac{K}{2\pi f \sin i}$	semi-major axis of stellar orbit around CM	AU	(2.61)
$a_\star \sin i \equiv \frac{K}{2\pi f}$	semi-major axis of stellar orbit times sine of inclination	AU	(2.61)
$a_p \equiv \frac{m_\star}{m_p} a_\star$	semi-major axis of planetary orbit around CM	AU	(2.29)
$a_{p,\max} \equiv \frac{m_\star}{m_{p,\min}} a_\star \sin i$	maximum semi-major axis of planetary orbit	AU	(2.29), (2.70)
$a_j \equiv \frac{K_j}{2\pi f \sin i}$	semi-major axis of component's orbit around CM	AU	(2.61)
$a_{\text{rel}} \equiv a_1 + a_2 = \frac{K_1+K_2}{2\pi f \sin i}$	semi-major axis of orbit of secondary around primary	AU	(2.61)
$d \equiv \frac{1 \text{ AU}}{\varpi}$	distance	pc	—

2.4 Errors and noise

Data from realistic, macroscopic measurement processes cannot be described exactly by deterministic observable models, as there is always some amount of *noise* caused by a variety of different and potentially unknown physical effects. Noise contributes a random *error*, or mismatch between the measured and the predicted value. Furthermore, *systematic errors*, i.e. deterministic but *a priori* unknown alterations, may be caused by the chosen observable model being too simplistic and/or overly complex with respect to the different aspects of the true physical process in the target, and/or by a misunderstanding or miscalibration of the instrument. Thus, assuming that an adequate “error-free” observable model is given by function $\mathbf{f}(\cdot; \cdot)$ with parameters $\boldsymbol{\theta}$, the measured value \mathbf{y}_i can be described as

$$\mathbf{y}_i = \mathbf{f}(t; \boldsymbol{\theta}) + \boldsymbol{\epsilon}_{s,i} + \boldsymbol{\epsilon}_i, \quad (2.85)$$

¹⁵The implicit function of minimum planetary mass $m_{p,\min}$ is solved numerically. The planet's minimum mass equals its real mass if the orbit is edge-on, viz. $\sin i = 1$ (section 2.2.4).

¹⁶The implicit function of planetary mass m_p is solved numerically.

¹⁷Refers to the star or any of the binary components, respectively.

where $\epsilon_{s,i}$ is the systematic error and ϵ_i is the random error.

If the functional form of the systematic error $\epsilon_{s,i}$ is known, it can be modelled by either altering the observable model $\mathbf{f}(\cdot; \cdot)$ or by letting it be absorbed into one or several of the model parameters. In this work, only the unknown constant calibration offset affecting most RV data sets is considered; it is accounted for by the reference radial velocity V_i , which is made an independent model parameter for each data set (table 2.1). All other possible systematic errors are assumed to be insignificant ($\epsilon_{s,i} \equiv 0$) and therefore neglected.

Noise, although unpredictable by definition, may obey characterisable distributions given by *noise models*, adjusted by one or several parameters (see also Sozzetti 2005). In the following, we assume that the noise of different data is independent.

According to the *principle of maximum entropy*, given only the mean and variance of a distribution, the normal (Gaussian) distribution has maximum information-theoretic entropy, equivalent to minimum bias or prejudice with respect to the missing information, and should therefore be used in these cases (Kapur 1989).

However, additional noise components may be present whose variance is unknown. We consider two distinct classes of noise-induced error components in measurements of stellar positions or radial velocities:

1. an error $\epsilon_{0,i}$ caused by instrumental effects and photon noise, known as the *internal error*, whose distribution is characterised by the *nominal uncertainty* of the i th datum, derived from an understanding of the physics and statistics of the measurement process;
2. an error $\epsilon_{+,i}$ caused e.g. by atmospheric or stellar effects not modelled otherwise, called *external error* and assumed independent from $\epsilon_{0,i}$.

Throughout, we assume that both error components follow a (uni- or bivariate) normal distribution with zero mean, which can be shown by means of characteristic functions to imply that the total error $\epsilon_i = \epsilon_{0,i} + \epsilon_{+,i}$ is again normally distributed with zero mean.

Sometimes a *non-standardised t-distribution* with (unknown) degrees-of-freedom parameter $\nu \in \mathbb{R}$ is adopted for the total error ϵ_i . This can be derived under the assumption that, for one-dimensional data, the unknown variance of ϵ_i follows an inverse-Gamma distribution¹⁸ and by then *marginalising* (\nearrow 3.2.4) the variance.

By contrast, in this work no marginalisation with respect to the noise variance is applied; instead, the external errors are described by specific parameters, described in the following, that are treated and estimated like model parameters.

For (one-dimensional) RV data, the internal error $\epsilon_{0,i}$ is distributed as $\mathcal{N}(0, \sigma_{0,i}^2)$, the external error $\epsilon_{+,i}$ is distributed as $\mathcal{N}(0, \sigma_+^2)$, where σ_+ is a free parameter, and thus the total error ϵ_i is distributed as $\mathcal{N}(0, \sigma_{0,i}^2 + \sigma_+^2)$.

For (two-dimensional) AM_a data, the internal error $\epsilon_{0,i} \sim \mathcal{N}(0, \mathbf{E}_{0,i})$, where $\mathbf{E}_{0,i}$ is called the AM_a *data covariance matrix*, and the external error $\epsilon_{+,i} \sim \mathcal{N}(0, \mathbf{E}_+)$. Here, the scalar covariance matrix $\mathbf{E}_+ = \text{diag}(\tau_+^2, \tau_+^2)$ characterises the distribution of external noise, assuming that the latter is *independent identically-distributed (i.i.d.)* in the two astrometric coordinates, with τ_+ being a free parameter. Thus, the total error $\epsilon_i \sim \mathcal{N}(0, \mathbf{E}_{0,i} + \mathbf{E}_+)$. The data covariance matrix represents the nominal uncertainty of the two measured astrometric coordinates and can be written using singular-value decomposition as

$$\mathbf{E}_{0,i} = \mathbf{R}(-\phi_i) \begin{pmatrix} a_i^2 & 0 \\ 0 & b_i^2 \end{pmatrix} \mathbf{R}(\phi_i), \quad (2.86)$$

¹⁸This is the *conjugate-prior* distribution of the variance of a Gaussian likelihood, i.e. it is the posterior distribution of the variance if the variance has an inverse-Gamma distribution as prior (\nearrow 3.2).

where $\mathbf{R}(\cdot)$, a_i , b_i and ϕ_i are the 2×2 *passive rotation matrix*, the nominal semi-major and -minor axes of the *uncertainty ellipse* and the position angle of its major axis with respect to North, respectively. An uncertainty ellipse is defined here as the location of all points whose probability density under the noise model is at least $\exp(-\frac{1}{2})$, corresponding to an interval of $\pm 1\sigma_0$ around the datum for RV data.

For (one-dimensional) AM_h data, the internal error $\epsilon_{0,i}$ is distributed as $\mathcal{N}(0, \tau_{0,i}^2)$, whereas the external errors in the underlying astrometric coordinates are distributed as AM_a data. This implies due to eq. (2.56) that the external error in abscissa $\epsilon_{+,i} \sim \mathcal{N}(0, \tau_+^2)$; thus, the total error $\epsilon_i \sim \mathcal{N}(0, \tau_{0,i}^2 + \tau_+^2)$.

As σ_+ and τ_+ are treated as free external-noise similar parameters, they are given for every evaluation of the noise distribution (\nearrow 3.1.1) and not marginalised over.

Chapter 3

Data analysis

Two approaches to inference

Data analysis is a type of *inductive reasoning*, inferring general rules from specific observational data (e.g. Gregory 2005b). These general rules are described by *observable models* which produce theoretical values of the observables as a function of *parameters*. Additionally, models for the errors are set up (section 2.4).

The first steps in the analysis of a given set of data consist in defining the noise model and all the possible concurrent observable models. Then, the following primary tasks of data analysis may be carried out:

1. In *model selection*, the relative probabilities of a set of concurrent models $\{\mathcal{M}_i\}$, chosen *a priori*, are assessed. In this work, only observable models are selected, while noise models are assumed to be known. Specifically, *exoplanet detection* tries to decide the question of whether a certain star is accompanied by a planet or not, based on available data.
2. Additionally, *model assessment* (or *model checking*) may be used to determine whether each of the models under consideration, especially the most probable one, adequately describe the data. In this thesis, the basic assumption is made that the most probable model¹ indeed provides an adequate description of the data, i.e. that potential other effects than those caused by binary or planetary companion(s) play no significant role. Therefore, model assessment is not performed in this thesis.
3. *Parameter estimation* aims to infer the parameters θ of a chosen model. This is specifically referred to as the *characterisation (or determination) of exoplanet orbits* in the present context.
4. The purpose of *uncertainty estimation* is to provide a measure of the uncertainties in the parameters.

3.1 Frequentist inference

The well-established *frequentist* approach to inference is named after the fact that it defines probability as the relative frequency of an event. Measurements are regarded as values of

¹For stars treated as binary, only one possible model is considered in this work (chapter 2).

random variables drawn from an underlying statistical *population* that is characterised by *population parameters*. The kind of population and its parameters are determined by the observable and noise models: e.g. for one-dimensional observables with Gaussian noise, under the condition that no systematic errors are present and the correct model function is $\mathbf{f}(\cdot; \cdot)$ with parameters $\boldsymbol{\theta}$, the population is normally distributed with variance equal to the noise variance and mean given by the observable model $\mathbf{f}(t; \boldsymbol{\theta})$ evaluated at parameters $\boldsymbol{\theta}$ and time t .²

3.1.1 Likelihood of data

Given the observable model (\nearrow 2.2) and the noise model (\nearrow 2.4) with their parameters, hence the population assumed to underlie the measurements, the probability density³ of any datum can be calculated. We start by defining the *residuals* for the relevant data types,

$$\Delta_{\text{AM},i} \equiv \mathbf{r}_i - \mathbf{r}(\boldsymbol{\theta}; t_i) \quad (3.1)$$

$$\Delta_{\text{h},i} \equiv a_{\text{h},i} - a_{\text{h}}(\boldsymbol{\theta}; t_i) \quad (3.2)$$

$$\Delta_{\text{RV},i} \equiv v_i - v(\boldsymbol{\theta}; t_i), \quad (3.3)$$

i.e. the differences between a datum taken at time t_i and the corresponding observable model, respectively for AM_a, AM_h and RV. Furthermore, we define the normalised residuals,

$$\varrho_{\text{AM},i} \equiv s_i \sqrt{\Delta_{\text{AM},i}^{\text{T}} \mathbf{E}_i^{-1} \Delta_{\text{AM},i}} \quad (3.4)$$

$$\varrho_{\text{h},i} \equiv \frac{\Delta_{\text{h},i}}{\tau_i} \quad (3.5)$$

$$\varrho_{\text{RV},i} \equiv \frac{\Delta_{\text{RV},i}}{\sigma_i}, \quad (3.6)$$

where

$$s_i \equiv \text{sgn}((\phi_i - \varphi_i)(\pi - [\phi_i - \varphi_i])) \quad (3.7)$$

gives the sign of the normalised AM_a residual, with φ_i and ϕ_i being the position angles of the residual and of the uncertainty ellipse, respectively (section 2.4). In these equations, total errors are employed as discussed in section 2.4,

$$\mathbf{E}_i \equiv \mathbf{E}_{0,i} + \mathbf{E}_+ \quad (3.8)$$

$$\tau_i^2 \equiv \tau_{0,i}^2 + \tau_+^2 \quad (3.9)$$

$$\sigma_i^2 \equiv \sigma_{0,i}^2 + \sigma_+^2. \quad (3.10)$$

²In general, t_i is the value of an *independent variable*, which may e.g. be temporal or spatial. We assumed the measurement durations to be short in comparison with the characteristic time of orbital motion, given by the orbital period, and thus the observations to take place at points in time t_i which are known exactly.

³Throughout, we use the term *probability density* wherever it refers to a continuous quantity, as opposed to *probability*, i.e. *probability mass function*, for discrete quantities. *Probability distribution*, denoted by $p(\cdot)$, is a generic term used for both cases.

Now the probability density or *likelihood* of an individual datum is dictated by the (Gaussian) noise model

$$\mathcal{L}_{\text{AM},i} \equiv p(\mathbf{r}_i|\boldsymbol{\theta}) = \frac{1}{2\pi\sqrt{\det \mathbf{E}_i}} \exp\left(-\frac{\varrho_{\text{AM},i}^2}{2}\right) \quad (3.11)$$

$$\mathcal{L}_{\text{h},i} \equiv p(a_{\text{h},i}|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\tau_i} \exp\left(-\frac{\varrho_{\text{h},i}^2}{2}\right) \quad (3.12)$$

$$\mathcal{L}_{\text{RV},i} \equiv p(v_i|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{\varrho_{\text{RV},i}^2}{2}\right). \quad (3.13)$$

The joint likelihood of all data of a given type equals the product of the above individual likelihoods, i.e.

$$\mathcal{L}_{\text{AM}} = \prod_{i=1}^{N_{\text{AM}}} \mathcal{L}_{\text{AM},i} = \left((2\pi)^{N_{\text{AM}}} \prod_{i=1}^{N_{\text{AM}}} \sqrt{\det \mathbf{E}_i} \right)^{-1} \exp\left(-\frac{\chi_{\text{AM}}^2}{2}\right) \quad (3.14)$$

$$\mathcal{L}_{\text{h}} = \prod_{i=1}^{N_{\text{h}}} \mathcal{L}_{\text{h},i} = \left((2\pi)^{\frac{N_{\text{h}}}{2}} \prod_{i=1}^{N_{\text{h}}} \tau_i \right)^{-1} \exp\left(-\frac{\chi_{\text{h}}^2}{2}\right) \quad (3.15)$$

$$\mathcal{L}_{\text{RV}} = \prod_{i=1}^{N_{\text{RV}}} \mathcal{L}_{\text{RV},i} = \left((2\pi)^{\frac{N_{\text{RV}}}{2}} \prod_{i=1}^{N_{\text{RV}}} \sigma_i \right)^{-1} \exp\left(-\frac{\chi_{\text{RV}}^2}{2}\right). \quad (3.16)$$

$$(3.17)$$

Here, $N_{\text{AM}}, N_{\text{h}}, N_{\text{RV}}$ are the respective numbers of data and

$$\chi_{\text{AM}}^2 \equiv \sum_{i=1}^{N_{\text{AM}}} \varrho_{\text{AM},i}^2 \quad (3.18)$$

$$\chi_{\text{h}}^2 \equiv \sum_{i=1}^{N_{\text{h}}} \varrho_{\text{h},i}^2 \quad (3.19)$$

$$\chi_{\text{RV}}^2 \equiv \sum_{i=1}^{N_{\text{RV}}} \varrho_{\text{RV},i}^2 \quad (3.20)$$

are called χ^2 *statistics* of the data, providing a measure of the deviation between data and model, known as the *goodness of fit*. If the Gaussian noise model and the observable model are correct, the normalised residuals follow a standard normal distribution $\mathcal{N}(0, 1)$. Then, and if the observable models are linear functions of the parameters, the values of the χ^2 statistics obey the χ_{ν}^2 *distribution*, where $\nu = N_D - k$ are the *degrees of freedom*, with N_D being the number of data and k the number of model parameters. This distribution has an expectation equal to ν and a variance given by 2ν .

In the general case of several types of data being combined, the total likelihood is the product of the corresponding joint likelihoods given in eq. (3.14) – (3.16); e.g. for AM_h and RV data,

$$\mathcal{L} = \mathcal{L}_{\text{h}}\mathcal{L}_{\text{RV}}. \quad (3.21)$$

3.1.2 Parameter estimation

Frequentist parameter estimation is generally equivalent to maximising \mathcal{L} or minimising χ^2 as functions of $\boldsymbol{\theta}$. The resulting *best estimates* of the parameters $\hat{\boldsymbol{\theta}}$ are therefore often called

maximum-likelihood or *least-squares estimates*. For linear models, $\chi^2(\boldsymbol{\theta})$ is a quadratic function and consequently $\hat{\boldsymbol{\theta}}$ can be found unambiguously by matrix inversion. In the more realistic cases of nonlinear models, however, χ^2 may have many local minima, therefore care needs to be taken not to mistake a local minimum for the global one. Several methods exist to this end, including evaluation of $\chi^2(\boldsymbol{\theta})$ on a finite grid, *simulated annealing* or *genetic algorithms* (e.g. Gregory 2005b).

3.1.3 Uncertainty estimation

Under the frequentist framework, parameter uncertainties are usually quoted as *confidence intervals*. Procedures to derive these are designed such that when repeated many times based on different data, a certain fraction of the resulting intervals will contain the true parameters. Popular methods use *bootstrapping* (Efron and Tibshirani 1993) or the *Fischer information matrix*, which is based on a local linearisation of the model (e.g. Ford 2004). However, these methods suffer from specific caveats: the Fischer matrix is only appropriate for a quadratic-shaped χ^2 in the vicinity of the minimum, and bootstrapping, which relies on modified data, may lead to severe misestimation of the parameter uncertainties, especially when these are large (Vogt et al. 2005).

3.1.4 Model selection

Frequentist model selection usually starts by setting up a *null hypothesis* \mathcal{H}_0 stating that the simplest model \mathcal{M}_0 is correct. If \mathcal{M}_0 is linear and the noise is normally distributed, the χ^2 statistic follows the known *reference distribution* χ^2_ν (\nearrow 3.1.1). Under these conditions, the following procedures are often used to select a model.

1. The goodness of fit $\hat{\chi}^2 = \min_{\boldsymbol{\theta}} \chi^2(\boldsymbol{\theta})$ of \mathcal{M}_0 is calculated and the reference distribution is integrated from $\hat{\chi}^2$ to infinity to obtain the *p-value*,

$$p = \text{p}(\chi^2 \geq \hat{\chi}^2 | \mathcal{H}_0). \quad (3.22)$$

If $p < \alpha$ for some previously chosen *significance level* α chosen in advance, say $\alpha = 0.05$, the null hypothesis is rejected as it would imply a better fit with high probability $1 - \alpha$. Rejecting \mathcal{H}_0 is equivalent to the hypothesis that another more complex model is correct. Because it only addresses one model explicitly, this is a type of model selection based on model *assessment*.

2. The relative goodness of fit of \mathcal{M}_0 and an alternative, more complex model \mathcal{M}_1 , defined as correct under hypothesis \mathcal{H}_1 , is determined by calculating χ^2 under both models and applying a certain statistic \mathcal{S} to these values $\hat{\chi}^2_{\nu_0}, \hat{\chi}^2_{\nu_1}$. Then, the measured value $\hat{\mathcal{S}}$ of this statistic is converted into a *p-value* analogously as above, using the reference distribution of \mathcal{S} under \mathcal{H}_0 . Again, if $p > \alpha$, the null hypothesis is rejected in favour of \mathcal{H}_1 . Among such statistics \mathcal{S} are the *F-statistic* and the *likelihood-ratio-test statistic*, whose null-hypothesis reference distributions are known only under certain regularity conditions and if $N_D \rightarrow \infty$ (Protassov et al. 2002).

It should be noted, however, that this type of model selection can only be applied if the models $\mathcal{M}_0, \mathcal{M}_1$ are nested, i.e. \mathcal{M}_1 turns into \mathcal{M}_0 if its additional parameters assume their so-called *null values*.

Period detection

The detection of periodicities in *time-series* of data, e.g. sets of RV data $\{(t_i, v_i)\}$, plays an important role particularly in the present context of discovering and characterising extrasolar planets. In assessing the relative performance of two competing models, it is a special type of frequentist model selection (item 2 above). One of its basic difficulties is to distinguish spurious patterns induced by noise, which mimics periodicities, from a real periodic signal.

An important tool for period detection is the *periodogram*, which proceeds by evaluating a certain statistic of the data at a range of test frequencies f . The *Lomb-Scargle (LS) periodogram* (Lomb 1976; Scargle 1982) is a well-known type of periodograms closely related to the *discrete Fourier-transform* (DFT; Deeming 1975) and briefly sketched in the following. The Lomb-Scargle statistic, or *periodogram power* P_{LS} , at any given frequency f measures the relative goodness of fit of a zero-mean sinusoidal signal of that frequency and a constant signal. In calculating the statistic, the free parameters of both models are according to a linear least-squares fit in which the individual measurement uncertainties are disregarded. The power P_{LS} follows an F -distribution under the null hypothesis \mathcal{H}_0 that only noise and no deterministic signal is present.

From the maximal power $P_{LS,max}$ found at frequency f_{max} , the p -value is calculated analogously to eq. (3.22). Correspondingly, the probability under \mathcal{H}_0 to obtain a power lower than $P_{LS,max}$ at some arbitrary frequency is $1 - p$. If the number of statistically independent frequencies, M , can be determined,⁴ one can calculate the probability that at least one of the frequencies has a peak exceeding the measured $P_{LS,max}$ although no periodic signal exists. This *false-alarm probability (FAP)* then equals $1 - (1 - p)^M$. If the FAP is found to fall below a given significance level α , \mathcal{H}_0 is rejected in favour of \mathcal{H}_1 , postulating a single sinusoidal periodicity of frequency f_{max} .

Other kinds of periodograms have been defined, notably the *floating-mean periodogram* (Cumming et al. 1999), the *Keplerian periodogram* (Cumming 2004), the *SIGSPEC periodogram* (Reegen 2007), and the *generalised Lomb-Scargle periodogram* (Zechmeister and Kürster 2009). While all of them account for a non-zero mean of the signal, the latter two also introduce weighting by considering the measurement uncertainties.

Of the variants mentioned, only the Keplerian and the generalised Lomb-Scargle periodogram are tailored to the signals of eccentric orbits in RV data, while the others correspond to the assumption of circular orbits when applied to RV data.

Besides noise, another problem inherent to period detection in finite time-series of data is *spectral leakage*, i.e. the appearance of spurious peaks in the periodogram due to the finiteness of total observing time (*sidelobes*) and time spacing (*aliases*). In particular, the DFT of a time series is the convolution of the Fourier transform of the underlying signal with a *spectral window* which determines the effects of spectral leakage.⁵

Spectral leakage, also called *aliasing*, often causes strong artefacts and is not accounted for in the reference distributions of most periodogram statistics. However, it affects different techniques to varying degrees (e.g. Reegen 2007). Methods to deal with spectral leakage are discussed, e.g. by Ferraz-Mello (1981); Roberts et al. (1987); Foster (1995); Reegen (2011).

⁴This number may only be possible approximatively or in special cases (e.g. Cumming et al. 1999).

⁵The spectral window is given as the DFT of a constant sampled at the observing times.

3.2 Bayesian inference

Bayesian inference (e.g. Sivia 2006), which has gained popularity in various scientific disciplines during the past few decades, defines probability as the degree of belief in a certain hypothesis \mathcal{H} . While this is sometimes criticised as leading to subjective assignments of probabilities, Bayesian probabilities are not subjective if they are based on all relevant knowledge \mathcal{K} , hence different persons with the same knowledge will assign them the same value (e.g. Sivia 2006). Thus, Bayesian probabilities are conditional on the knowledge \mathcal{K} , and this conditionality should be stated explicitly, as in the following equations.

3.2.1 Bayes' theorem

In the eighteenth century, Thomas Bayes laid the foundation of a new approach to inference with what is now known as Bayes' theorem (Bayes and Price 1763). For the purpose of parameter and uncertainty estimation, the hypothesis \mathcal{H} refers to the values of the model parameters $\boldsymbol{\theta}$, and Bayes' theorem is expressed as

$$\underbrace{p(\boldsymbol{\theta}|D, \mathcal{M}, \mathcal{K})}_{\text{posterior } \mathcal{P}(\boldsymbol{\theta})} = \frac{\overbrace{p(\boldsymbol{\theta}|\mathcal{M}, \mathcal{K})}^{\text{prior } \pi(\boldsymbol{\theta})} \cdot \overbrace{p(D|\boldsymbol{\theta}, \mathcal{M}, \mathcal{K})}^{\text{likelihood } \mathcal{L}(\boldsymbol{\theta})}}{\underbrace{p(D|\mathcal{M}, \mathcal{K})}_{\text{evidence } \mathcal{Z}}} \quad (3.23)$$

where $D \equiv \{(t_i, \mathbf{y}_i)\}$ is the set of pairs of observational times and corresponding data values and \mathcal{M} denotes the particular model assumed. As mentioned above, all probabilities are also conditional on the knowledge \mathcal{K} , including statements on the types of parameters and on parameter space Θ – which we assume to be a subset of \mathbb{R}^k with $k \in \mathbb{N}$ – as well as the noise model.

Using Bayes' theorem, the aim is to determine the *posterior* $p(\boldsymbol{\theta}|D, \mathcal{M}, \mathcal{K}) \equiv \mathcal{P}(\boldsymbol{\theta})$, i.e. the probability distribution of the parameters $\boldsymbol{\theta}$ in light of the data D , given the model \mathcal{M} and prior knowledge \mathcal{K} . The posterior represents all the knowledge about the model parameters based on the data and prior knowledge. The other terms on the right-hand side of the theorem are explained below.

- The term *prior* refers to the probability distribution $p(\boldsymbol{\theta}|\mathcal{M}, \mathcal{K}) \equiv \pi(\boldsymbol{\theta})$ of the parameters $\boldsymbol{\theta}$ given only the model and prior knowledge \mathcal{K} ; it characterises the knowledge about the parameters present before considering the data. For objective choices of priors, based on classes of parameters, see section 3.2.2.
- The *likelihood* $p(D|\boldsymbol{\theta}, \mathcal{M}, \mathcal{K}) \equiv \mathcal{L}(\boldsymbol{\theta})$ is the probability distribution of the data values D , given the observing times, the model and the parameters, where the notation $\mathcal{L}(\boldsymbol{\theta})$ is used to refer to the likelihood as a function of the parameters. It is introduced in the context of frequentist inference in section 3.1.1.
- The *evidence* $p(D|\mathcal{M}, \mathcal{K}) \equiv \mathcal{Z}$ is the probability distribution of the data values D , given the observing times and the model but neglecting the parameter values,

$$p(D|\mathcal{M}, \mathcal{K}) = \int p(D, \boldsymbol{\theta}|\mathcal{M}, \mathcal{K}) d\boldsymbol{\theta} \quad (3.24)$$

$$= \int p(\boldsymbol{\theta}|\mathcal{M}, \mathcal{K}) p(D|\boldsymbol{\theta}, \mathcal{M}, \mathcal{K}) d\boldsymbol{\theta} \quad (3.25)$$

$$= \int \pi(\boldsymbol{\theta}) \mathcal{L}(\boldsymbol{\theta}) d\boldsymbol{\theta}. \quad (3.26)$$

It equals the integral of the product of prior and likelihood over parameter space Θ and plays the role of a normalising constant. In practice, however, it is hard to calculate (section 3.2.7).

It may be instructive here to note that the frequentist approach of maximising the likelihood $p(D|\theta, \mathcal{M}, \mathcal{K})$ is equivalent to maximising the posterior when assuming uniform priors $p(\theta|\mathcal{M}, \mathcal{K})$. This can be seen by inserting $p(\theta|\mathcal{M}, \mathcal{K}) = \text{const}$ into eq. (3.23), which leads to

$$\mathcal{P}(\theta) = p(\theta|D, \mathcal{M}, \mathcal{K}) \propto p(D|\theta, \mathcal{M}, \mathcal{K}) = \mathcal{L}(\theta) \quad (3.27)$$

However, this maximum-likelihood approach ignores the fact that uniform priors are not always the most objective choice (\nearrow 3.2.2) and the posterior cannot be fully characterised just by the position of its maximum. Still, the latter can be used as a posterior summary in the Bayesian framework (section 3.2.5).

3.2.2 Encoding prior knowledge

By means of the prior, Bayesian analysis allows one to incorporate knowledge obtained earlier, e.g. using different data. When no prior knowledge is available for some model parameter, except for its allowed range, maximum prior ignorance about the parameter can be encoded by a prior of one of the following functional forms for the most common classes of *location* and *scale parameters* (Gregory 2005b; Sivia 2006).

- For a *location parameter*, we demand that the prior be invariant against a shift Δ in the parameter, i.e.

$$p(\theta|\mathcal{M}, \mathcal{K}) d\theta = p(\theta + \Delta|\mathcal{M}, \mathcal{K}) d(\theta + \Delta), \quad (3.28)$$

which leads to the *uniform prior*

$$p(\theta|\mathcal{M}, \mathcal{K}) = \frac{\Theta(\theta - a) \Theta(b - \theta)}{b - a}, \quad (3.29)$$

where $\Theta(\cdot)$, a , and b are the Heaviside step function, the lower and the upper prior bounds.

Here, we note that the frequentist approach, lacking an explicit definition of the prior, corresponds to the implicit assumption of a uniform prior for all parameters.

- A positive *scale parameter*, which often spans several decades, is characterised by its invariance against a stretch of the coordinate axis by a factor φ , i.e.

$$p(\theta|\mathcal{M}, \mathcal{K}) d\theta = p(\varphi\theta|\mathcal{M}, \mathcal{K}) d(\varphi\theta), \quad (3.30)$$

which is solved by the *Jeffreys prior*,

$$p(\theta|\mathcal{M}, \mathcal{K}) = \frac{\Theta(\theta - a) \Theta(b - \theta)}{\theta \ln \frac{b}{a}}. \quad (3.31)$$

That a uniform prior would be inappropriate for this parameter is also illustrated by the fact that it would assign higher probabilities to θ lying within a higher decade of $[a, b]$ than in a lower.

- If the lower prior bound of a scale parameter is zero, e.g. for the RV semi-amplitude K , a *modified Jeffreys prior* is used. It has the form

$$p(\theta|\mathcal{M}, \mathcal{K}) = \frac{\Theta(\theta - a) \Theta(b - \theta)}{(\theta + \theta_k) \ln \frac{b+\theta_k}{\theta_k}}, \quad (3.32)$$

where θ_k is the *knee* of the prior. For $\theta \ll \theta_k$, this prior is approximately uniform, while it approaches a Jeffreys prior for $\theta \gg \theta_k$.

- Scale parameters which can have positive or negative sign have a *signed modified Jeffreys* prior which we define as

$$p(\boldsymbol{\theta}|\mathcal{M}, \mathcal{K}) \equiv \frac{1}{2} p_{\text{mJ}}(|\boldsymbol{\theta}| | \mathcal{M}, \mathcal{K}), \quad (3.33)$$

where $p_{\text{mJ}}(\cdot)$ is a modified Jeffreys prior.

3.2.3 Posterior sampling of parameters and derived quantities

In order to obtain the normalisation constant \mathcal{Z} of the posterior from prior $\pi(\boldsymbol{\theta})$ and likelihood $\mathcal{L}(\boldsymbol{\theta})$, one would have to integrate the product $\pi(\boldsymbol{\theta})\mathcal{L}(\boldsymbol{\theta})$ over the whole parameter space (eq. (3.24)). This can be achieved in general by analytic or numerical integration, but neither is feasible in more than about three dimensions, and the former method may carry further disadvantages in being approximative or restricted to simple integrands.

To circumvent this obstacle, the *Markov chain Monte Carlo (MCMC)* method (e.g. Gilks et al. 1996) allows to gather samples distributed as the posterior density,⁶ from these samples, various aspects of the posterior, described below, can then be estimated. Moreover, MCMC allows to explore the whole parameter space without applying a regular grid to all dimensions, which may not be chosen fine enough in more than a few dimensions to adequately sample the posterior. MCMC and several related techniques are described in technical detail in section 4.5.

Not only the parameters $\boldsymbol{\theta}$ are of interest *a posteriori*, but various other relevant quantities can be derived from them (table 2.2). Such derived quantities $\vartheta = \vartheta(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}})$, where $\tilde{\boldsymbol{\theta}}$ are additional parameters such as the epoch t_r or stellar mass m_* , are implicitly sampled with the parameters $\boldsymbol{\theta}$ according to

$$\vartheta^{(j)} = \vartheta(\boldsymbol{\theta}^{(j)}, \tilde{\boldsymbol{\theta}}). \quad (3.34)$$

For example, the time of periapsis T is sampled as

$$T^{(j)} = t_r - \frac{\chi^{(j)}}{f^{(j)}}. \quad (3.35)$$

These derived-quantity samples can be used analogously to parameter samples for posterior inference as described below, where a notation pertaining only to parameters is used for simplicity.

⁶The first M samples belonging to the *burn-in* phase do not follow the posterior distribution and thus need to be excluded (section 4.5.2).

3.2.4 Marginalisation and density estimation

As a density over $k > 2$ dimensions, the posterior $\mathcal{P}(\boldsymbol{\theta})$ cannot be displayed unambiguously in a figure. By reducing the dimensionality of the posterior domain, *marginal posteriors* $\mathcal{P}_i(\cdot)$, i.e. probability densities over each of the parameters θ_i , and *joint marginal posteriors* $\mathcal{P}_{i,j}(\cdot, \cdot)$ over two parameters, can be obtained and plotted. This reduction is known as *marginalisation* and described mathematically by integration over all other parameters,

$$\mathcal{P}_i(\theta_i) \equiv p(\theta_i|D, \mathcal{M}, \mathcal{K}) = \int \mathcal{P}(\boldsymbol{\theta}) d\boldsymbol{\theta}_{\setminus i} \quad (3.36)$$

$$\mathcal{P}_{i,j}(\theta_i, \theta_j) \equiv p(\theta_i, \theta_j|D, \mathcal{M}, \mathcal{K}) = \int \mathcal{P}(\boldsymbol{\theta}) d\boldsymbol{\theta}_{\setminus i,j}, \quad (3.37)$$

where $d\boldsymbol{\theta}_{\setminus i} \equiv \prod_{k \neq i} d\theta_k$ and $d\boldsymbol{\theta}_{\setminus i,j} \equiv \prod_{k \neq i,j} d\theta_k$. In practice, marginal posteriors are estimated from the collected samples $\{\boldsymbol{\theta}^{(j)}\}$ by only considering their i th component and performing a *density estimation* based on these one-dimensional samples. Joint marginal posteriors are derived analogously, based two-dimensional samples of components i and j .

Several *density estimators* exist for deriving a density from a set of samples. One of them – the oldest and probably most popular type, known as the *histogram* – has several drawbacks: its shape depends on the choice of origin and bin width, and when used on two-dimensional data, a contour diagram cannot easily be derived from it.

Generalising the histogram to *kernel density estimation* over one or two dimensions, the samples can be represented more accurately and unequivocally (Silverman 1986). Below, we refer only to the simpler one-dimensional case. There, the kernel estimator can be written as

$$\mathcal{F}(x) \equiv \frac{1}{N\sigma_{\text{ker}}} \sum_{j=M+1}^N K\left(\frac{x - X^{(j)}}{\sigma_{\text{ker}}}\right), \quad (3.38)$$

where x is a scalar variable, M is the *burn-in* length (\nearrow 4.5.2), N is the total chain length (i.e. number of samples), $K(\cdot)$ is the *kernel*, σ_{ker} is the *window width* and $\{X^{(j)}\}$ are the underlying samples. As detailed by Silverman (1986), the efficiency of various kernels in terms of the achievable *mean integrated square error* is very similar, and therefore the choice of kernel can be based on other requirements. Since no differentiability is required for the estimated densities and computational effort plays an important practical role, a triangular kernel,

$$K_{\text{tri}}(x) \equiv \max(1 - |x|, 0), \quad (3.39)$$

was selected for estimating the marginal posteriors and a biweight kernel,

$$K_{\text{bi}}(x) \equiv \frac{15}{16} (1 - x^2)^2, \quad (3.40)$$

for the joint marginal posteriors.

The window width is chosen following the recommendations of Silverman (1986),

$$\sigma_{\text{ker}} \equiv 2.189 \cdot \min\left(\sigma_{\text{samp}}, \frac{r_{\text{iq}}}{1.34}\right) (N - M)^{\frac{1}{5}}, \quad (3.41)$$

where σ_{samp} is the sample standard deviation and r_{iq} is the interquartile range of the samples.

3.2.5 Parameter estimation

To obtain a single most probable estimate of the parameters, the posterior density $\mathcal{P}(\cdot)$ can be summarised by the *posterior mode* $\hat{\theta} \in \Theta$, i.e. the point where the posterior attains its maximum value,

$$\hat{\theta} \equiv \arg \max_{\theta} \mathcal{P}(\theta). \quad (3.42)$$

This point, also known as the *maximum a-posteriori (MAP)* parameter estimate, can be approximated by the MCMC sample with highest posterior density (Gilks et al. 1996), based on the values of $\mathcal{P}(\theta^{(j)})$ already calculated during sampling. This approximation neglects the finite spacing between samples.

Alternatively, the following scalar summaries can be inferred from the samples or, in case of the marginal mode, from the marginal posteriors $\mathcal{P}_i(\theta)$:

- *mean* or *expectation* $\bar{\theta}$,

$$\bar{\theta} \equiv \int_{-\infty}^{\infty} \theta \mathcal{P}_i(\theta) d\theta, \quad (3.43)$$

- *median* $\tilde{\theta}$,

$$\int_{-\infty}^{\tilde{\theta}} \mathcal{P}_i(\theta) d\theta \equiv 0.5, \quad (3.44)$$

- *marginal mode* $\check{\theta}$,

$$\check{\theta} \equiv \arg \max_{\theta} \mathcal{P}_i(\theta). \quad (3.45)$$

3.2.6 Uncertainty estimation

For uncertainty estimation, posterior samples enable *highest posterior-density intervals (HPDIs)* to be estimated. For any given $C \in \mathbb{R}$ with $0 < C < 1$, a HPDI $I_{\text{HPD}} \equiv [a, b]$ is defined as the smallest interval over which the posterior contains a probability C , i.e.

$$\int_a^b \mathcal{P}_i(\theta) d\theta = C, \quad \text{s.t.} \quad b - a = \min. \quad (3.46)$$

In contrast to frequentist confidence intervals, HPDIs are generally not symmetric, meaning that their midpoint does not necessarily correspond to the best estimate. This is because the marginal posteriors may be asymmetric, including any amount of skew.

It should also be noted that HPDIs are not useful with multimodal posteriors because several modes cannot be meaningfully summarised by one interval per dimension, nor by a single best estimate.

To quantify linear dependencies between parameters, the a-posteriori *Pearson correlation coefficient*,

$$r_{\theta_1, \theta_2} \equiv \frac{\text{cov}(\theta_1, \theta_2)}{\sqrt{\text{var}(\theta_1) \text{var}(\theta_2)}} = r_{\theta_2, \theta_1}, \quad (3.47)$$

can be inferred from the samples. There may also be nonlinear correlations between parameters that are not described by the correlation coefficients. Furthermore, one should be aware that for strong linear or non-linear relationships between parameters, uncertainties of single parameters as characterised by HPDIs may not be meaningful.

We stress that the (joint) marginal posteriors can – and should – always be referred to, especially when best estimates and/or HPDIs do not adequately characterise the posterior. The availability of these more informative densities is one of the advantages of a Bayesian approach with posterior sampling.

3.2.7 Model selection

Bayes' theorem (\nearrow 3.2.1) can also be expressed for model selection based on a hypothesis \mathcal{H} stating that a certain model \mathcal{M} is correct. This gives

$$p(\mathcal{M}|D, \mathcal{K}) = \frac{p(\mathcal{M}|\mathcal{K}) \cdot p(D|\mathcal{M}, \mathcal{K})}{p(D|\mathcal{K})}, \quad (3.48)$$

where $p(\mathcal{M}|\mathcal{K})$ and $p(\mathcal{M}|D, \mathcal{K})$ are the model prior and posterior, respectively, $p(D|\mathcal{K})$ is a normalising constant, while $p(D|\mathcal{M}, \mathcal{K})$ is the evidence given equivalently to eq. (3.24) by

$$p(D|\mathcal{M}, \mathcal{K}) = \int p(D, \boldsymbol{\theta}|\mathcal{M}, \mathcal{K}) d\boldsymbol{\theta} \quad (3.49)$$

$$= \int \pi(\boldsymbol{\theta}) \mathcal{L}(\boldsymbol{\theta}) d\boldsymbol{\theta} = \mathcal{Z}. \quad (3.50)$$

When two models $\mathcal{M}_0, \mathcal{M}_1$ are concurrently entertained, their *posterior odds* are therefore given by

$$\frac{p(\mathcal{M}_1|D, \mathcal{K})}{p(\mathcal{M}_0|D, \mathcal{K})} = \underbrace{\frac{p(\mathcal{M}_1|\mathcal{K})}{p(\mathcal{M}_0|\mathcal{K})}}_{\text{prior odds}} \cdot \underbrace{\frac{p(D|\mathcal{M}_1, \mathcal{K})}{p(D|\mathcal{M}_0, \mathcal{K})}}_{\text{Bayes factor}}. \quad (3.51)$$

Absent any knowledge to the contrary, a fair choice for the *prior odds* is to set them equal to 1. In the following, we therefore ignore the prior odds and consider the posterior odds to be given by the *Bayes factor*

$$B_{1,0} \equiv \frac{p(D|\mathcal{M}_1, \mathcal{K})}{p(D|\mathcal{M}_0, \mathcal{K})} = \frac{\mathcal{Z}_1}{\mathcal{Z}_0}, \quad (3.52)$$

where \mathcal{Z}_0 and \mathcal{Z}_1 are the evidences of the two models.

In contrast to frequentist model selection, Bayesian evidences and thus the Bayes factor include an *Ockham's razor*, which imposes a penalty on any model \mathcal{M}_i whose parameter space (i.e. prior support) is larger than warranted by the likelihood, as can be seen from the definition of the evidence, eq. (3.50): the evidence, also known as the *prior-averaged likelihood* or *marginal likelihood*, decreases when the prior supports large areas with low likelihood $\mathcal{L}(\boldsymbol{\theta})$, i.e. the models are “unnecessarily complex”. Additionally, the Bayes factor is a consistent selector, which increasingly supports the correct model in the limit $N_D \rightarrow \infty$ (Weinberg 2012).

A common method of estimating integrals such as that in eq. (3.50) is known as the *Laplace approximation*, where the integrand is approximated as a multivariate normal distribution. However, this only leads to good results if all significant posterior peaks have been found and can be represented accurately by Gaussian distributions, which would be a severe practical limitation for the purposes of this work. By contrast, methods are described in the following to estimate the integral in eq. (3.50) on the basis of samples from the prior or posterior densities.

Assume that a set of samples $\{\boldsymbol{\theta}^{(j)} : j = M + 1, \dots, N\}$ has been drawn, where $\boldsymbol{\theta}^{(j)} \sim g(\boldsymbol{\theta})$ and $g(\cdot)$ is a density. The evidence of a particular model can then be written

$$\mathcal{Z} = \int \frac{\pi(\boldsymbol{\theta}) \mathcal{L}(\boldsymbol{\theta})}{g(\boldsymbol{\theta})} g(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (3.53)$$

and can be approximated by the *consistent estimator* (Gilks et al. 1996, ch. 10)

$$\hat{\mathcal{Z}} = \frac{1}{N-M} \sum_{j=M+1}^N \frac{\pi(\boldsymbol{\theta}^{(j)})\mathcal{L}(\boldsymbol{\theta}^{(j)})}{g(\boldsymbol{\theta}^{(j)})} \quad (3.54)$$

$$\equiv \left\| \frac{\pi\mathcal{L}}{g} \right\|_g. \quad (3.55)$$

Below, the notation $\mathcal{Z} \simeq \hat{\mathcal{Z}}$ denotes the relation between a quantity and its estimator. If the prior $\pi(\cdot)$ is chosen as density g , the estimator results as

$$\hat{\mathcal{Z}} = \|\mathcal{L}\|_\pi. \quad (3.56)$$

Sampling from the prior, however, can be similarly inefficient as pure Monte-Carlo sampling. By contrast, for posterior samples, the density is

$$g(\boldsymbol{\theta}) \equiv \mathcal{P}(\boldsymbol{\theta}) = \frac{\pi(\boldsymbol{\theta})\mathcal{L}(\boldsymbol{\theta})}{\mathcal{Z}}. \quad (3.57)$$

Using

$$\frac{1}{\mathcal{Z}} = \frac{1}{\mathcal{Z}} \int \pi(\boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{\theta} = \int \frac{\pi(\boldsymbol{\theta})}{\mathcal{Z}g(\boldsymbol{\theta})} g(\boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{\theta} \quad (3.58)$$

and eq. (3.57) yields

$$\frac{1}{\mathcal{Z}} = \int \frac{1}{\mathcal{L}(\boldsymbol{\theta})} g(\boldsymbol{\theta}) \, \mathrm{d}\boldsymbol{\theta}. \quad (3.59)$$

An estimator for the evidence based on posterior samples is therefore given by

$$\hat{\mathcal{Z}} = \left(\frac{1}{N-M} \sum_{j=M+1}^N \frac{1}{\mathcal{L}(\boldsymbol{\theta}^{(j)})} \right)^{-1} \quad (3.60)$$

$$\equiv \left\| \frac{1}{\mathcal{L}} \right\|_{\mathcal{P}}^{-1}. \quad (3.61)$$

This *harmonic-mean approximation* of the evidence, however, is prone to disturbance by samples with extremely low likelihood and therefore does not exist in many cases; convergence criteria are given by Wolpert (2002). Weinberg (2012) presents an algorithm based on Lebesgue integrals for calculating eq. (3.60) by ignoring samples contributing a large error. However, this *Numerical Lebesgue Algorithm* is only stable and recommended with informative priors.

Here, another algorithm for estimating the evidence \mathcal{Z} as given by eq. (3.50) is therefore employed. Provided by Weinberg (2012), it uses *balanced kd-trees* for space partitioning. This *Volume Tessellation Algorithm (VTA)* is applied to a subspace $\Theta_s \subset \Theta$ of the parameter space which should be chosen so as to help minimise bias and variance by excluding low-posterior regions, e.g. those containing samples $\boldsymbol{\theta}^{(j)}$ with low values of $\pi(\boldsymbol{\theta}^{(j)})\mathcal{L}(\boldsymbol{\theta}^{(j)})$. The tree is constructed storing samples only in the leaves, not in inner nodes, which simplifies subsequent volume calculations.

Due to eq. (3.23), one has

$$\mathcal{Z} \int_{\Theta_s} \mathcal{P}(\boldsymbol{\theta}) d\boldsymbol{\theta} = \int_{\Theta_s} \pi(\boldsymbol{\theta}) \mathcal{L}(\boldsymbol{\theta}) d\boldsymbol{\theta}, \quad (3.62)$$

and therefore

$$\mathcal{Z} = \frac{\int_{\Theta_s} \pi(\boldsymbol{\theta}) \mathcal{L}(\boldsymbol{\theta}) d\boldsymbol{\theta}}{\int_{\Theta_s} \mathcal{P}(\boldsymbol{\theta}) d\boldsymbol{\theta}}, \quad (3.63)$$

with the denominator simply given as the fraction of posterior samples in Θ_s ,

$$\int_{\Theta_s} \mathcal{P}(\boldsymbol{\theta}) d\boldsymbol{\theta} = \frac{1}{N - M} \sum_{j=M+1}^N \mathbb{1}_{\Theta_s}(\boldsymbol{\theta}^{(j)}), \quad (3.64)$$

where $\mathbb{1}(\cdot)$ is the indicator function. Given a tessellation of Θ_s by n_s sub-volumes ω_i , each containing at most a fixed number c of leaves, the numerator of eq. (3.63) can be estimated by first assigning each ω_i a representative value P_i of $\pi(\cdot) \mathcal{L}(\cdot)$ (the sample median is chosen for P_i in the case of balanced kd-trees) and then approximating the integral as a Riemann sum,

$$\int_{\Theta_s} \pi(\boldsymbol{\theta}) \mathcal{L}(\boldsymbol{\theta}) d\boldsymbol{\theta} \approx \sum_{i=1}^{n_s} \omega_i P_i. \quad (3.65)$$

Thus, one has

$$\mathcal{Z} \simeq (N - M) \frac{\sum_{i=1}^{n_s} \omega_i P_i}{\sum_{j=M+1}^N \mathbb{1}_{\Theta_s}(\boldsymbol{\theta}^{(j)})}. \quad (3.66)$$

In cases tested by Weinberg (2012) with low-informative priors, the VTA excelled both the NLA and the Laplace approximation.

3.2.8 Model-uncertainty prediction and observation scheduling

Posterior samples may also be used to assess the uncertainty in the value of the observable model(s) at any given instant t . In particular, when t is in the future, such an *uncertainty prediction* may help to schedule upcoming observations of the same target with the same technique(s). In this work, uncertainty prediction is applied only to AM_a and radial velocities, as the HIPPARCOS instrument is out of operation.

Given posterior samples $\{\boldsymbol{\theta}^{(j)}\}$, we define uncertainty prediction for a given type of data to consist in the following algorithm:

1. Define the set of times $\{t_i\}$ of interest;
2. for each $t_i \in \{t_i\}$ do:
 - (a) for each $\boldsymbol{\theta}^{(j)} \in \{\boldsymbol{\theta}^{(j)}\}$, calculate the value of the model function $\mathbf{f}_{i,j} \equiv \mathbf{f}(t_i; \boldsymbol{\theta}^{(j)})$;
 - (b) calculate a measure of dispersion ς_i of the values $\{\mathbf{f}_{i,j}\}$, i.e.:
 - the variance or standard deviation, for one-dimensional model functions $f(\cdot; \cdot)$,
 - or

- the semi-major axes a_i and semi-minor axes b_i of the uncertainty ellipse which is defined by the covariance matrix \mathbf{S} of the two vector components of $\{\mathbf{f}_{i,j}\}$, for two-dimensional model functions; equivalently, a_i and b_i are given by the eigenvalues of \mathbf{S} ,

$$a_i = \frac{\text{tr } \mathbf{S}}{2} + \sqrt{\left(\frac{\text{tr } \mathbf{S}}{2}\right)^2 - |\mathbf{S}|} \quad (3.67)$$

$$b_i = \frac{\text{tr } \mathbf{S}}{2} - \sqrt{\left(\frac{\text{tr } \mathbf{S}}{2}\right)^2 - |\mathbf{S}|}. \quad (3.68)$$

The dispersion measures $\{\varsigma_i\}$ can then be plotted over time, followed by locating their extrema. For observation scheduling, re-observation can be recommended at the time t_{\max} of maximum dispersion of the predicted model-function values, i.e. maximum uncertainty in the observable, yielding a maximal constraint of the observed stellar motion.

Alternatively, observation scheduling may be based on the premise that re-observation should yield the maximum gain in information on the parameters. This approach is called *maximum-entropy sampling* (e.g. Loredo 2004; Ford 2008) and employs the (negative) Shannon entropy as a measure for the information on a parameter. In the case of normally distributed predictions $\mathbf{f}(t; \cdot)$ with t given, this is equivalent to uncertainty prediction as described above.

Chapter 4

Informatics and implementation

A tool for Bayesian exoplanet science

This chapter introduces BASE, a **B**ayesian **a**strometric and **s**pectroscopic exoplanet detection and characterisation tool. Its goals are to fulfil two major tasks of exoplanet science, namely the *detection of exoplanets* and the *characterisation of their orbits*, by implementing methods of Bayesian statistics detailed in chapter 3.

BASE has been developed to provide for the first time the possibility of an integrated Bayesian analysis of stellar astrometric and Doppler-spectroscopic measurements with respect to their companions' signals, correctly treating particularly the correlated astrometric measurement uncertainties (↗ 2.4) and allowing one to explore the whole multidimensional parameter space Θ without the need for informative prior constraints. Still, users may readily incorporate prior knowledge, e.g. from previous analyses with other tools, by means of priors on the model parameters. The tool automatically diagnoses convergence of its Markov chain Monte Carlo (MCMC) sampler to the posterior and regularly outputs status information.

For orbit characterisation, BASE performs a complete Bayesian parameter and uncertainty estimation, delivering important results including probability densities and correlations of model parameters and several derived quantities. As opposed to a single best estimate and confidence interval per parameter, this is especially important when the data do not constrain the parameters well, e.g. when only few data have been recorded or the *signal-to-noise* ratio is low (as can be the case for lightweight planets or young host stars). Another important function which BASE has been built to include is Bayesian model selection, performed if the user allows concurrent models with different numbers of planets.

BASE comes in the form of a highly configurable command-line tool, developed in Fortran 2008 and compiled with GFortran (Free Software Foundation 2011a). This chapter details the implementation and modes of using BASE.

4.1 Requirements

BASE has been developed according to the following requirements:

- *Accuracy:*

- the user should be able to reflect their prior knowledge, where present, as accurately as possible in the analysis;
 - calculations should be carried out to an accuracy sufficient for the respective task;¹
 - output should include all significant digits and not be trimmed.
- *Flexibility:*
 - all relevant aspects of how the analysis is carried out should be adjustable by the user;
 - BASE should make sensible assumptions where user demand is missing.
 - *Usability:*
 - the modes of supplying information to BASE and control its behaviour should be as simple and sensible as possible;
 - the user should be able to control the time taken by flexible-duration tasks (such as sampling) and should be provided with an estimate of the remaining runtime for a given task wherever possible;
 - output should be easily comprehensible (including explanations where needed);
 - the user should be warned when BASE acts in an unexpected way;
 - BASE should run automatically, without the need of user supervision nor interaction after the start,² while providing the user with relevant up-to-date status information.

4.2 Other software and features of BASE

Development of BASE has commenced before this work (Schulze-Hartung 2008), leading to an early version with only some of today’s capabilities of the tool. In the course of this dissertation, the program has been significantly extended to include, among other aspects, treatment of AM data, modelling of binaries and multi-planet systems, as well as model selection.

Table 4.1 lists the most essential features of BASE as well as their availability both in BASE before the present work and in three other computer programs introduced recently in the literature. Although other tools for Bayesian inference using MCMC with a more general applicability exist – such as those by Weinberg and Moss (2011) or Foreman-Mackey et al. (2012) – only programs that specialise on exoplanet science, specifically allowing to readily analyse AM_a and RV data, are included in this overview. The table demonstrates that BASE includes considerable new and important functionality.

¹Therefore, most calculations operate on *double-precision* (64-bit) numbers, but *quadruple precision* (128-bit) is employed where essential.

²This is especially important considering the sometimes hour-long durations of BASE runs.

Table 4.1: Essential features of BASE and their presence in other existing tools specialised on exoplanet science. Where fields are blank, no clear indication of the feature being available could be found in the cited article.

Category	BASE feature	Section	BASE before this work	Tuomi et al. (2009)	Gregory (2011)	Anglada-Escudé et al. (2012)
Data	AM _a data	4.4.2		×	×	×
	RV data	4.4.2	×	×	×	×
	AM _a +RV data	4.4.2		×		×
	AM _h (+RV) data	4.4.2				
	User-defined grouping of data	4.4.2				
	Correct treatment of AM _a uncertainties	2.4				
Models	Binary stars (two RV amplitudes)	2.2.5				
	Multiple planets	2.1.4				
	User-defined epoch	4.4.1				
	Special treatment of cyclic parameters	2.1.1			×	
	Additional noise	2.4			×	
Priors	Prior information	4.4.3	×		×	
	Arbitrary prior densities	4.4.3				
Sampling	MCMC	3.2.3	×	×	×	×
	Hit-and-run sampler	4.5.2	×			
	Thinning	4.5.2	×		×	
	Parallel tempering	4.5.3	×		×	
	Automatic convergence diagnostic	4.5.4	×		×	
	Saving and reading samples	4.6.2				
Inference	Derived quantities	3.2.3			×	×
	Marginal posteriors by kernel density estimation	3.2.4	×		×	
	Refinement of kernel window width (for parameter f)	4.3.3				
	User-defined HPDIs, posterior-probability intervals and hypercubes	4.4.1				
	Joint marginal-posterior densities	3.2.4		×		
	Model selection	3.2.7			×	
	Uncertainty prediction	3.2.8				
Other	Generating data	4.6.3	×	×		
	Regular status updates during tasks	4.1	×			
	Gnuplot interface	4.4.1				
	Producing plots with L ^A T _E X formulae	4.4.1				

4.3 Modes of operation

BASE can operate in several kinds of modes, which correspond to the assumption of certain physical systems or to the type of inference to be made. The modes are explained as follows.

4.3.1 Normal and binary mode

BASE is capable of analysing data for two similar types of systems, whose physics have been described in section 2.2. These are systems with:

1. one observable component that may be accompanied by one or more unobserved bodies (generally referred to as planetary systems here for simplicity); this mode of operation is called the *normal mode*;
2. two observable, gravitationally bound components (referred to as binary stars), one of which may serve as reference point for observing the other; this mode of operation is called *binary mode*.

BASE automatically selects normal or binary mode depending on the given data as described in section 4.4.2.

4.3.2 Number-of-planets modes

In normal mode, any combination of up to ten numbers of planets $\{n_{p,i}\}$ can be selected by the user (section 4.4.1). Each number of planets $n_{p,i} \in [0, 9]$ corresponds to a separate n_p mode, including sampling and analysis, carried out by BASE. Each such mode includes the assumption of a corresponding model and parameters which can be configured independently using the `-scope` option (section 4.4.1).

In the case of several n_p modes, a Bayesian model selection is made after the last mode has been finished. Giving the same number of planets several times allows to compare the posterior probability of the same model under different conditions, e.g. with or without additional noise (\nearrow 2.4).

4.3.3 Periodogram mode

Using the option `-periodogram`, BASE can be put in *periodogram mode*, which modifies binary or 1-planet mode by restricting inference to the orbital frequency f . Its marginal posterior is estimated using a refined kernel window width (\nearrow 3.2.4), resulting in narrower peaks and, in general, an altered marginal mode. Posterior summaries calculated from the samples directly are not modified in periodogram mode. The algorithm used for window-width refinement is detailed in section 4.6.1.

4.4 Input

Users can provide different kinds of information to BASE, i.e. a number of configuration options, data, and prior knowledge. These various types of user input are described below.

Table 4.2: Option notation.

Notation	Meaning
<code>[[<arg>]]</code>	Optional argument <code><arg></code>
<code>[[<arg> <arg> ...]]</code>	List of alternative arguments
<code><arg>[[, <arg> ...]]</code>	One or more instances of <code><arg></code> separated by <code>,</code>
<code>...</code>	Repetition of expression from and including the previous <code> </code> or <code>,</code>
<code>[[...]]</code>	Recursive insertion of enclosing optional expression
<code>[]</code>	Square brackets enclosing arguments or intervals, to be put literally

4.4.1 Invocation and options

According to the requirement of running without the need for user interaction, options supplied in starting the program control the behaviour of BASE and allow the user to input information such as the stellar mass or prior knowledge (see section 3.2.1). In the following, options are referred to using a notation explained in table 4.2. Any option can be supplied:

1. in a configuration file³ `base.conf` in the working directory and/or
2. on the command line.

The configuration file is parsed first, followed by the command line, with command-line options overriding those in the configuration file. Besides options, arguments on the command line may be filenames,⁴ where each file must contain a data set in one of the formats described in section 4.4.2. As opposed to filenames, options are preceded by the characters `-` on the command line to distinguish them from file names. No data-file names may be supplied in the configuration file and the `-` prefix is omitted in options. Below, options are always quoted with the prefix for clarity.

Options may be classified according to their syntax as follows:

1. commands:
 - (a) `-option`
 - (b) `-option[<arg>]`
2. assignments:
 - (a) `-option=<rhs>`
 - (b) `-option[<arg>]=<rhs>`

In assignments, depending on the option, `<rhs>` may denote:

1. a single value `<val>`
2. a list of values `<val>[[, <val>|...]]`
3. an interval `[[<lbound>, <ubound>]]`.

Using the options described in the following, BASE may be invoked by:

```
base [[<options>]] <filename> [[<options>|<filename>] [...]]
```

³Only the `-group` option, which refers to a specific given data set, cannot be employed in the configuration file.

⁴By default, filenames are relative to the *data directory* (option `-data-dir`); if the data directory is not set, filenames refer to the working directory or may be absolute paths to a file.

Global and local options. Many options are *scopable*, i.e. it is possible to set their scope either to all n_p modes in a given BASE run (which is the default) or only to a specific mode. Instances of such options are called *global* in the former case and *local* in the latter. A scope referring to a given n_p mode can be started with the `-scope` option, with all succeeding options before the next `-scope` becoming local to that mode. All options given before the first `-scope` option are global. If in a given n_p mode an option is set both globally and locally, the local option takes precedence.

Available options. The following options are available to control the behaviour of BASE:

`-calc-hpdi=<val>[[, <val>[...]]]` Set the probability contents of additional HPDI(s) I_{HPD} (section 3.2.6) to be calculated for all parameters and derived quantities. BASE automatically calculates HPDIs of probability contents 50%, 68.27%, 95%, 95.45%, 99%, and 99.73%.

Arguments: `<val>`: probability content C (section 3.2.6). *Default:* none. *Restrictions:* Not combinable with `-no-inference`. Cannot be supplied more than once.

`-calc-prob[<par>]=[<lbound>, <ubound>]` Calculate the posterior probability of a parameter θ to lie within a given interval I ,

$$p(\theta \in I | D, \mathcal{M}, \mathcal{K}) \simeq \frac{1}{N - M} \sum_{j=M+1}^N \mathbb{1}_I(\theta^{(j)}). \quad (4.1)$$

Arguments: `<par>`: name of θ (table 4.3); `<lbound>`, `<ubound>`: lower and upper bounds of I . *Default:* none. *Restrictions:* Not combinable with `-no-inference`.

`-calc-prob-all-pars=[[<bounds>|,], [[<bounds>|,], ...]` Calculate the posterior probability of parameters θ to lie within the given hypercube $\Theta_{\text{sub}} \subset \Theta$,

$$p(\theta \in \Theta_{\text{sub}} | D, \mathcal{M}, \mathcal{K}) \simeq \frac{1}{N - M} \sum_{j=M+1}^N \mathbb{1}_{\Theta_{\text{sub}}}(\theta^{(j)}). \quad (4.2)$$

Arguments: `<bounds>` ::= `<lbound>`, `<ubound>`: pair of lower and upper bounds of Θ_{sub} in a given dimension; replacing any `<bounds>` by `,` implies that the corresponding parameter is marginalised. *Default:* none. *Restrictions:* Not combinable with `-no-inference`.

`-calc-transit-after=<time>` Set the time after which the first potential transit times $t_{t,1}, t_{t,2}$ of a planet are to be calculated (table 2.2).

Arguments: `<time>`: time [d] after which transit times are to be calculated. *Default:* last observing time. *Restrictions:* Not combinable with `-no-inference`. Cannot be supplied more than once.

`-comment=<text>` Include a comment in the log.

Arguments: `<text>`: comment to be included. *Default:* none. *Restrictions:* Cannot be supplied more than once.

`-data-dir=<dir>` Set data directory.

Arguments: `<dir>`: path to the data directory (relative to working directory, or absolute). *Default:* (empty) (interpreted relative to working directory). *Restrictions:* Cannot be supplied more than once.

`-discard-frac=<val>` Set maximum runtime fraction during which to initially discard samples before starting burn-in (section 4.5.2).

Arguments: `<val>`: fraction of runtime. *Default:* 0.1. *Restrictions:* Not combinable with `-read-samples`. A potential adjustment of tempering parameters occurs after 5 minutes at the latest, immediately followed by the start of burn-in, regardless of this option.

`-discard-psr-scale=<val>` Set the factor ρ_p by which the difference $\hat{\mathcal{R}}^{1/2} - \mathcal{R}_*^{1/2}$ has to decrease for all parameters, i.e.

$$\frac{\hat{\mathcal{R}}_i^{1/2} - \mathcal{R}_*^{1/2}}{\hat{\mathcal{R}}_{\max}^{1/2} - \mathcal{R}_*^{1/2}} \leq \rho_p \quad \forall i \in [1, k] \quad (4.3)$$

before the initial discarding of samples is stopped and burn-in is started (\nearrow 4.5.2), where $\hat{\mathcal{R}}^{1/2}$ is the PSR of a parameter, $\hat{\mathcal{R}}_{\max}^{1/2}$ is the maximum PSR attained initially by any parameter and $\mathcal{R}_*^{1/2}$ is the threshold for convergence (option `-psr-conv`).

Arguments: `<val>`: scale factor ρ_p . *Default:* 0.9. *Restrictions:* Not combinable with `-read-samples`. Cannot be supplied more than once. May be overridden by the maximum runtime before adjusting tempering parameters and/or starting burn-in (option `-discard-frac`).

`-epoch=<time>` Set the epoch (section 2.1.1).

Arguments: `<time>`: epoch t_r . *Default:* centre of observing time span, $\frac{t_1+t_N}{2}$. *Restrictions:* Not combinable with `-read-samples`. Cannot be supplied more than once.

`-generate-data` `[[ang_pos|RV], [[<commands>], [[<# planets>], [<pars>], [[<mass>], [<# data>], <begin>, <end>, <uncertainty>[[, <uncertainty>[...]]]]` Generate AM_a or RV data according to the given arguments (section 4.6.3).

Arguments: `<commands>` ::= `<command>[[, <command>[...]]`: data command(s) (table 4.13); `<# planets>`: number of planets n_p (unless in binary mode); `<pars>`: parameters θ (components separated by `,`); `<mass>`: stellar mass (only for notation in data-file header, omitted in binary mode); `<# data>`: number of data; `<begin>`, `<end>`:

times [d] of first and last observation; $\langle \text{uncertainty} \rangle ::= \llbracket \langle \text{major} \rangle, \langle \text{minor} \rangle | \langle \text{stdev} \rangle \rrbracket$: uncertainty ellipse (a, b) or standard deviation σ , for AM_a and RV data respectively (section 4.6.3). *Restrictions:* Not combinable with `-read-samples`, `-no-inference`. Cannot be supplied more than once. If observing times are specified by the data command `$time`, they must comply with $\langle \# \text{ data} \rangle$ and the number of $\langle \text{uncertainty} \rangle$ both; also in this case, values of $\langle \text{begin} \rangle$ and $\langle \text{end} \rangle$ are ignored.

`-group[⟨thresholds⟩]` Arrange the data from the preceding file in groups according to the given thresholds, then replace each group by a representative synthetic datum (section 4.4.2).

Arguments: $\langle \text{thresholds} \rangle ::= \llbracket \langle \text{time} \rangle, \langle \text{phi} \rangle | \langle \text{time} \rangle, \langle \text{psi} \rangle, \langle \text{gamma} \rangle | \langle \text{time} \rangle \rrbracket$: thresholds for AM_a , AM_h and RV data, respectively, as defined in section 4.4.2. *Restrictions:* Not scopable. Cannot be used in the configuration file.

`-help` Display a help screen.

Restrictions: Not combinable with `-read-samples`, `-save-samples`, `-no-inference`. Not scopable.

`-help-pars` Display information on parameters and derived quantities.

Restrictions: Not combinable with `-read-samples`, `-save-samples`, `-no-inference`. Not scopable.

`-inference-len-frac=⟨val⟩` Set the fraction of the number of links used for inference to the total cold-chain length, $\rho_i \equiv (N - M)N^{-1}$, where N is the total length and M is the burn-in length (section 4.5.2).

Arguments: $\langle \text{val} \rangle$: fraction ρ_i . *Default:* 0.5.

`-inference-len=⟨val⟩|max]` Set the minimal number of links used for inference, or maximise that number, with the latter implying that BASE will not stop sampling before either the maximum chain length is reached or both maximum runtime is reached and burn-in is finished (section 4.5.2).

Arguments: $\langle \text{val} \rangle$: minimum number of links used for inference, $(N - M)_{\min}$. *Default:* $\langle \text{val} \rangle = 0$. *Restrictions:* Not combinable with `-read-samples`. Cannot be supplied more than once. Independent of this option, inference is always based on at least 10^3 samples.

`-K_max-from-vel` Determine prior upper bound of K from RV range and minimum eccentricity as

$$K_{\max} = \tilde{\Delta}v \cdot \frac{\sqrt{1 - e_{\min}^2}}{2}, \quad (4.4)$$

according to eq. (2.60) and (2.76) (section 2.3).

Restrictions: Not combinable with `-read-samples`, `-msini_max`. Cannot be supplied more than once. Ignored if no RV data given.

`-keep-seed` Reuse the random seed from file `.BASE.randomSeed` in the working directory, where a newly generated seed is stored.

Restrictions: Not combinable with `-read-samples`. Cannot be supplied more than once.

`-M=<val>` Set the stellar mass. Needs to be specified by this option or in the `AMa/RV` data file(s) if derived quantities are to be calculated.

Arguments: `<val>`: stellar mass $m_{\star}[M_{\odot}]$. *Default:* none. *Restrictions:* Cannot be supplied more than once.

`-mail-when-finished=<address>` Send an email to the specified address when BASE has finished.

Arguments: `<address>`: email address. *Default:* none. *Restrictions:* Cannot be supplied more than once.

`-max-dur=<val><unit>` Set the maximum runtime to be taken by the Multi-PT procedure. If this option is not given, `-max-len` must be given.

Arguments: `<val>`: value of maximum runtime in the given unit; `<unit>` ::= `[[h|m|s]]`: time unit. *Default:* none. *Restrictions:* Not combinable with `-read-samples`. May be overridden by `-max-len` in case of conflict.

`-max-len=<val>` Set the maximum length of the final cold chain. If this option is not given, `-max-dur` must be given. Takes precedence over `-max-dur` in case of conflict.

Arguments: `<val>`: chain length. *Default:* none. *Restrictions:* Not combinable with `-read-samples`.

`-msini_max=<val>` Set the maximum value of $m_p \sin i$, from which the prior upper bound of K is calculated, according to eq. (2.62), as

$$\sqrt[3]{\frac{2\pi G f_{\max}}{m_{\star}^2}} (m_p \sin i)_{\max}. \quad (4.5)$$

Arguments: `<val>`: value of $(m_p \sin i)_{\max}[M_J]$. *Default:* 13.⁵ *Restrictions:* Not

⁵This corresponds to the Deuterium-burning limit for Solar-metallicity substellar objects.

combinable with `-read-samples`, `-K_max-from-vel`. Cannot be supplied more than once. Ignored if no RV data given.

`-n-periods-range=[⟨lower⟩,⟨upper⟩]` Set assumed range of the number of orbital periods $[n_{\min}, n_{\max}]$ covered by any data file, translated to a prior range on f as

$$f \in \left[\frac{n_{\min}}{\Delta t_{\min}}, \frac{n_{\max}}{\Delta t_{\max}} \right], \quad (4.6)$$

where $\Delta t_{\min}, \Delta t_{\max}$ are the shortest and longest time spans covered by any data set, respectively.

Arguments: $\langle \text{lower} \rangle, \langle \text{upper} \rangle$: lowest and highest numbers of orbital periods. *Default:* none. *Restrictions:* Not combinable with `-read-samples`.

`-n-pl=⟨val⟩[[,⟨val⟩[...]]]` Set number(s) of planets to be assumed, thereby defining a set of n_m different n_p *modes* to be carried out (section 4.3.2).

Arguments: $\langle \text{val} \rangle$: number of planets $n_{p,i}$. *Default:* $n_{p,1} = 1, n_m = 1$. *Restrictions:* Cannot be supplied more than once.

`-n-pt=⟨val⟩` Set number of PT procedures m to be run by Multi-PT (section 4.5.4). Setting $m \equiv 1$ reduces Multi-PT to PT (section 4.5.3).

Arguments: $\langle \text{val} \rangle$: number of PT procedures m . *Default:* 2. *Restrictions:* Not combinable with `-read-samples`. Cannot be supplied more than once.

`-no-deriv` Disable calculation of derived quantities (section 3.2.3).

`-no-files` Disable any file output.

Restrictions: Not combinable with `-save-samples`. Cannot be supplied more than once.

`-no-gnuplot-exec` Don't execute Gnuplot after finishing.

Restrictions: Cannot be supplied more than once.

`-no-inference` Disable inference, i.e. only collect posterior samples.

Restrictions: Cannot be supplied more than once.

`-no-joint-marg-post` Don't produce joint marginal posteriors.

`-no-marg-post` Don't produce marginal posteriors. This also prevents the calculation of marginal-posterior modes.

Restrictions: Not combinable with `-periodogram`. Cannot be supplied more than once.

`-no-plots` Don't produce any files related to plotting. This also prevents the calculation of marginal-posterior modes.

Restrictions: Cannot be supplied more than once.

`-no-ps4pdf-exec` Disable execution of the program `ps4pdf` (Niepraschk and Voß 2001), which serves to convert *PSTricks*-compatible plots produced by Gnuplot to the PDF format if `-pstricks-plots` is used.

Restrictions: Cannot be supplied more than once. Can only be used with `-pstricks-plots`.

`-no-temp-pars-adj` Disable adjustment of tempering parameters. Otherwise, tempering parameters are adjusted so as to optimise the “round-trips” of samples in the space of tempering parameter γ and thus improve the efficiency of PT (section 4.5.3) by a variant of the *feedback-optimized PT* algorithm (Katzgraber et al. 2006). Since the efficiency improvement has not been quantified or closely studied with BASE, using the present option is recommended.

Restrictions: Not combinable with `-read-samples`.

`-only-plot-data` Only plot the given data and exit. If combined with `-group`, data are grouped first (section 4.4.2).

Restrictions: Not combinable with `-read-samples`, `-save-samples`. Cannot be supplied more than once.

`-only-speed-meas` Only perform a measurement of sampling speed and exit.

Restrictions: Not combinable with `-read-samples`, `-save-samples`. Cannot be supplied more than once.

`-out-dir=<output dir>` Set output directory.

Default: `../out` (relative to working directory). *Restrictions:* Cannot be supplied more than once.

`-out-dir-csv=<CSV output dir>` Set output directory for CSV plot files.

Default: `<output dir>/csv`. *Restrictions:* Cannot be supplied more than once.

`-periodogram` Restrict inference to orbital frequency f , producing a marginal posterior with refined kernel window width (section 4.6.1).

Restrictions: Not combinable with `-no-inference`, `-no-marg-post`, `-quick-plots`. Cannot be supplied more than once.

`-predict=[⟨begin⟩,⟨end⟩]` Predict uncertainties of observables corresponding to given data sets in the given time span.

Arguments: $\langle\text{begin}\rangle$, $\langle\text{end}\rangle$: time [d] of beginning and end of prediction time span. *Default:* none. *Restrictions:* Not combinable with `-no-inference`. Cannot be supplied more than once.

`-prior[⟨par⟩]=⟨file⟩` Set the prior probability density of a parameter θ from a CSV file, implying both its prior shape and support; the latter can be additionally clipped using `-range` (section 4.4.3).

Arguments: $\langle\text{par}\rangle$: name of θ (table 4.3); $\langle\text{file}\rangle$: path (relative to working directory, or absolute) to a CSV file containing samples of the prior density (table 4.8). *Default:* none. *Restrictions:* Not combinable with `-read-samples`.

`-psr-conv=⟨val⟩` Set PSR threshold for convergence (section 4.5.4).

Arguments: $\langle\text{val}\rangle$: PSR threshold $\mathcal{R}_*^{1/2}$. *Default:* 1.1. *Restrictions:* Not combinable with `-read-samples`. Cannot be supplied more than once.

`-pstricks-plots` Produce PStricks-compatible plots, i.e. `.tex` files that can be included in a \LaTeX document. Implies `-single-plots`.

Restrictions: Cannot be supplied more than once.

`-pt-n-chains=⟨val⟩` Set number of chains n run by each PT procedure (section 4.5.3). Setting $n \equiv 1$ reduces PT to pure MCMC sampling (section 4.5.2).

Arguments: $\langle\text{val}\rangle$: number of chains n . *Default:* 8. *Restrictions:* Not combinable with `-read-samples`. Cannot be supplied more than once.

`-quick-plots` Produce plots of reduced resolution, e.g. reducing the number of points at which marginal posteriors are sampled from 500 to 100.

Restrictions: Not combinable with `-periodogram`. Cannot be supplied more than once.

`-quiet` Suppress all messages to the standard output.

Restrictions: Cannot be supplied more than once. Overridden by `-help`.

`-range[⟨par⟩]=[[⟨type⟩]:[[⟨lower⟩,⟨upper⟩]` Set the prior range (and type) of a parameter θ .

Arguments: $\langle \text{par} \rangle$: name of θ (table 4.3); $\langle \text{lower} \rangle$: the lower prior bound θ_{\min} ; $\langle \text{upper} \rangle$: the upper prior bound θ_{\max} ; $\langle \text{type} \rangle ::= [\text{u}|\text{j}|\text{m}|\text{s}|\text{n}[\langle \text{n} \rangle]]$: the prior type (uniform, Jeffreys, modified Jeffreys, signed modified Jeffreys or truncated normal, respectively); $\langle \text{n} \rangle$: ratio of $\frac{\theta_{\max}-\theta_{\min}}{2}$ to the standard deviation for a truncated normal prior (whose distribution mean is set to $\frac{\theta_{\min}+\theta_{\max}}{2}$). *Restrictions:* Not combinable with `-read-samples`.

`-read-samples=⟨file⟩` Read MCMC samples from a saved-samples file (\nearrow 4.6.2) instead of drawing them. Only the samples of the requested n -planets mode (`-n-pl`) are used. The original data and potential `-group` option(s) must be retained. Takes precedence over `-save-samples`.

Restrictions: Cannot be supplied more than once.

`-save-cand` Save the sampled candidates (\nearrow 4.5.2) in a CSV file.

Restrictions: Not combinable with `-read-samples`. Cannot be supplied more than once.

`-save-samples` Save the final cold-chain samples (section 4.6.2).

Restrictions: Cannot be supplied more than once. Ignored when used with `-read-samples`.

`-scatter-plots` Produce scatter plots, i.e. plots of pairs of components of the posterior samples.

Restrictions: Not combinable with `-no-inference`.

`-scope=⟨i⟩` Begin a new scope, which makes the following options refer to the indicated n_p mode.

Arguments: $\langle i \rangle$: the position i of the intended n_p mode in the set $\{n_{p,i}\}$ given with the option `-n-pl`.

`-single-plots` Output one graphic file per plot, where possible.

Restrictions: Cannot be supplied more than once.

`-speed-meas-dur=⟨val⟩⟨unit⟩` Set maximum duration of the speed measurement.

Arguments: $\langle \text{val} \rangle$: value of maximum duration in the given unit; $\langle \text{unit} \rangle ::= [\text{h}|\text{m}|\text{s}]$:

time unit. *Restrictions:* Not combinable with `-read-samples`.

`-stdout-line-len=<val>` Set the line length on standard output.

Arguments: `<val>`: line length [characters]. *Restrictions:* Cannot be supplied more than once.

`-swap=<val>` Set mean spacing between swap proposals (section 4.5.3).

Arguments: `<val>`: mean swap spacing n_{swap} . *Default:* 100. *Restrictions:* Not combinable with `-read-samples`.

`-temp-pars=<val>[[, <val>[...]]]` Set the (initial) tempering parameters of the heated chains (section 4.5.3).

Arguments: `<val>`: tempering parameter $\gamma_{(k)}$. *Restrictions:* Not combinable with `-read-samples`. Cannot be supplied more than once.

`-temp-pars-adj-dur=<val><unit>` Set maximum duration of tempering-parameters adjustment.

Restrictions: Not combinable with `-read-samples`, `-no-temp-pars-adj`.

`-thin=[[<val>|auto]]` Set thinning stride q , or have it determined automatically from maximum runtime (`-max-dur`) and maximum chain length (`-max-len`) such that both are approximately reached (section 4.5.2). $q = 1$ implies that thinning is deactivated. When combined with `-read-samples`, the posterior samples are thinned after reading them in.

Arguments: `<val>`: thinning stride q . *Restrictions:* Cannot be supplied more than once. Automatic thinning cannot be performed with `-read-samples`.

`-value[<par>]=<val>` Set parameter θ to a fixed value c , equivalent to a delta-function prior $p(\theta|\mathcal{M}, \mathcal{K}) = \delta(\theta - c)$ (section 4.4.3).

Arguments: `<par>`: name of θ (table 4.3); `<val>`: the parameter value c . *Restrictions:* Not combinable with `-read-samples`.

`-zoom-all-pars=[[<bounds>|,], [[<bounds>|,], ...]` Set the ranges of “zoomed” marginal posteriors and joint marginal posteriors for all parameters and only make zoomed variants of these plots. Marginal posteriors of derived quantities are zoomed accordingly. Omitting any `<bounds>` implies that the prior range will be used for the corresponding parameter.

Arguments: `<bounds> ::= <lower>, <upper>`: pair of lower and upper bounds in a given

Table 4.3: Parameter names in BASE.

Symbol	Name	Symbol	Name	Symbol	Name
V	V	μ_δ	mu_d	i	i
a_v	a_v	τ_+	tau_+	Ω	Omega
σ_+	sigma_+	e	e	K	K
ϖ	pi	f	f	K_1	K_1
α_r	alpha_r	χ	chi	K_2	K_2
δ_r	delta_r	ω	omega	a'_{rel}	a`_rel
μ_{α_*}	mu_acd	ω_2	omega_2	a'	a`

dimension. *Restrictions:* Not combinable with `-no-inference`. Cannot be supplied more than once.

4.4.2 Data and selection of mode

Observational data of the following types can be treated by BASE:

- AM data:
 - AM_a: relative angular position of the secondary with respect to the primary binary component, given in a cartesian or polar coordinate system in a CSV file⁶ with one record per line and each record comprising the fields described in table 4.4 or table 4.5;
 - AM_h: HIPPARCOS intermediate astrometric data (abscissa residuals), given in a binary file as described in van Leeuwen (2007, section G.2.2).
- RV data in a CSV file with one record per line and each record comprising the fields described in table 4.6.

BASE automatically selects binary mode if and only if for all of the given types of data:

- all data of the type refer to the secondary measured with respect to the primary component (implemented for AM_a data) or
- data of both binary components with respect to an external, quasi-inertial reference point are given (RV data).

⁶In the CSV files read by BASE, record fields should be separated by one or more spaces.

Table 4.4: AM_a data (cartesian coordinate system): record fields.

No.	Quantity	Description	Unit
1	t	Time	d
2	$\alpha \cos \delta$	Right ascension	mas
3	δ	Declination	mas
4	a	Semi-major axis of uncertainty ellipse	mas
5	b	Semi-minor axis of uncertainty ellipse	mas
6	ϕ	Position angle of uncertainty ellipse, eastwards from North	°

Table 4.5: AM_a data (polar coordinate system): record fields.

No.	Quantity	Description	Unit
1	t	Time	d
2	ρ	Angular distance	mas
3	θ	Position angle, eastwards from North	$^\circ$
4	a	Semi-major axis of uncertainty ellipse	mas
5	b	Semi-minor axis of uncertainty ellipse	mas
6	ϕ	Position angle of uncertainty ellipse, eastwards from North	$^\circ$

Consequently, analysis is carried out in binary mode if AM_a data and/or RV data of both binary components are given. AM_h data cannot be treated in binary mode.

File format

The first data record is preceded by the file header, in which every non-empty line starts with the comment character `#`. To specify the data type and additional information, any header line may contain a *data command* according to table 4.7; the command must be placed immediately after the comment character.

Data grouping

BASE offers to organise observational data from a given set in subsets containing a number of chronologically successive data and replacing each such *group* by a representative synthetic datum derived from the group members. As defined here, this procedure is different from regular *binning* of the data in that the group boundaries in the time domain are not regularly spaced. Rather, groups may start at arbitrary times, whereas their *maximum* time span is fixed.

Besides time, other aspects of the data are used to determine group membership, depending on the data type. Accordingly, the *grouping parameter(s)* supplied by the user with the `-group` option specify the maximum ranges of:

1. time, Δt (all types)
2. uncertainty-ellipse orientation, $\Delta\phi$ (AM_a data)
3. scan-orientation angle, $\Delta\psi$ (AM_h data)
4. ratio of parallax factor to its group mean, $\frac{\Delta\gamma}{\bar{\gamma}}$ (AM_h data).

In each set, beginning with the earliest datum, groups are created from as many successive data as possible without exceeding any of the relevant grouping parameters. Groups of

Table 4.6: Radial-velocity data: record fields.

No.	Quantity	Description	Unit
1	t	Time	d
2	v	Radial velocity	m s^{-1}
3	σ	Measurement uncertainty	m s^{-1}

Table 4.7: Data commands in CSV files.

Command	Meaning	Remarks
<code>\$name=<string></code>	Set system identifier	
<code>\$M=<mass></code>	Set stellar mass $m_* [M_\odot]$	May be set using <code>-M</code> option
<code>\$type=[[ang_pos RV]]</code>	Data type (AM _a or RV)	Required
<code>\$coord_sys=polar</code>	Polar coordinate system	Only for AM _a data
<code>\$bin_comp=[[1 2]]</code>	Positions given for primary/secondary component	Required for AM _a data: 2
<code>\$ref_pt=companion</code>	Positions measured relative to binary companion	Required for AM _a data

size 1 may also be created where successive data are too widely spaced; in such groups, data are not modified. The resulting group sizes are printed such that it can be reconstructed which original records have been grouped in each set.

The synthetic replacement datum is defined as the weighted group mean, with respect to the following quantities:

- time t (all types)
- observable (RV and AM_h data)
- uncertainty-ellipse orientation ϕ (AM_a data)
- scan-orientation angle ψ (AM_h data)
- parallax factor γ (AM_h data).

For a quantity x , the weighted group mean is given as

$$\bar{x} = \frac{\sum_{i=j}^k w_i x_i}{\sum_{i=j}^k w_i} \quad (4.7)$$

if the group contains the j th through k th datum, where w_i are normalised weights determined, respectively for one- and two-dimensional data, by (\nearrow 2.4)

$$w_i = \begin{cases} \frac{1}{\sigma_i^2} \\ \frac{1}{a_i^2 + b_i^2} \end{cases} \quad (4.8)$$

By contrast, the AM_a replacement observable is derived by first rotating the original coordinate system S_5 around the z_5 -axis by the weighted-mean orientation angle $\bar{\phi}$, i.e. such that all uncertainty-ellipse semi-major axes are approximately aligned with the new x' -axis. Subsequently, the data coordinates $\{x'_i\}$ and $\{y'_i\}$ are independently averaged using weights $\left\{\frac{1}{a_i^2}\right\}$ and $\left\{\frac{1}{b_i^2}\right\}$, respectively. The resulting mean $(\bar{x}', \bar{y}')^\top$ is finally transformed back into S_5 . Its uncertainty ellipse is determined by semi-major and -minor axes

$$\bar{a} \equiv \frac{1}{\sqrt{\sum_{i=j}^k \frac{1}{a_i^2}}} \quad (4.9)$$

$$\bar{b} \equiv \frac{1}{\sqrt{\sum_{i=j}^k \frac{1}{b_i^2}}}, \quad (4.10)$$

Table 4.8: User-defined prior shape: record fields.

No.	Quantity	Description	Unit
1	θ	Abscissa	<i>according to parameter</i>
2	$p(\theta \mathcal{M}, \mathcal{K})$	Prior density	

according to the independent weighted-averaging of $\{x'_i\}$ and $\{y'_i\}$.

This definition disregards the differences in orientations of the individual uncertainty ellipses, but instead treats them as co-aligned. It allows to describe the replacement uncertainty by a bivariate normal distribution just as the original uncertainties and thus keep the noise model unchanged. This is a good approximation if $\Delta\phi$ is chosen small enough. Similarly, the definition for AM_h data ignores the variation in scan-circle orientation and parallax factor, which is a good approximation if the corresponding thresholds are small.

Grouping may be useful by reducing the average noise level for data whose spacing is so close that no significant orbital motion is thought to occur within the corresponding time span. However, owing to its modification of original data, it should be used with caution and inference should not ideally rely on grouped data only.

4.4.3 Priors

Any model parameter can be assigned a prior probability density in one of three forms:

1. a delta-function $\delta(\theta - c)$, equivalent to fixing the parameter at value c (option `-value`);
2. a user-defined shape, i.e. a set of up to 1001 samples $\{(\theta_i, p(\theta_i|\mathcal{M}, \mathcal{K}))\}$ of the prior density given in a CSV file with one record per line and each record comprising the fields described in table 4.8 (option `-prior`);
3. a range (either in combination with a user-defined shape, which is then clipped to the given range, or else using one of the default prior shapes detailed in section 3.2.2 (option `-range`)).

Any parameter for which no such option is present is assigned a default prior shape and range to pose the weakest possible restraints justified by the data, the model, and mathematical/physical considerations.

In case 2 above, the prior density is linearly interpolated between the given samples and normalised accordingly; the prior support is set to the abscissa range of the samples, or to a sub-range specified with the `-range` option. If less than 499 samples are within the prior support or their abscissae are not uniformly spaced, the prior density is re-estimated at 1001 positions by means of kernel density estimation.

4.5 Program architecture

In the following, the architecture of BASE is described in terms of the program flow on the top level, the three integral parts responsible for MCMC sampling, improvement of mixing and detection of convergence, as well as the organisation of the Fortran source code in modules.

4.5.1 Top-level program flow

On the top level, program flow in BASE (excluding some minor actions) can be visualised as in fig. 4.1. In complete analogy, the essential tasks between *Start* and *End* and their dependencies on options are described as follows.

1. *Start*
2. Initialise variables, including several default values which may be modified using options.
3. Parse configuration file and command line (section 4.4.1).
4. If `-help`: display a general help screen;
else if `-help-pars`: display a parameter-related help screen;
else if observational data are supplied, read observational data.
5. If `-group`: group the data as described in section 4.4.2.
6. If `-read-samples`: read header of saved-samples file from an earlier BASE run (section 4.6.2).
7. If `-keep-seed`: initialise the pseudo-random-number generator with the seed stored previously in `.BASE.randomSeed` in the working directory;
else: initialise the pseudo-random-number generator with a seed read from `/dev/urandom`.
8. If `-only-plot-data`: write files for data plots;
else if `-generate-data`: generate synthetic data (section 4.6.3);
else:
 - (a) Set up models and parameters according to user input.
 - (b) Repeat for all n_p modes:
 - i. If `-read-samples`:
 - A. Read posterior samples $\{\boldsymbol{\theta}^{(j)}\}$.
 - B. Check consistency of parameter set in saved-samples file with current model.
 - ii. Set various properties of parameters and derived quantities such as names and bounds.
 - iii. Pre-calculate eccentric anomalies.
 - iv. If `-read-samples`: calculate likelihoods $\{\mathcal{L}(\boldsymbol{\theta}^{(j)})\}$ of samples;
else: execute `multiPT`, including measurement of sampling speed.
 - v. Unless `-only-speed-meas`:
 - A. If `-save-samples`: save posterior samples.
 - B. Unless `-no-inference`:
 - Calculate various summaries and derived quantities.
 - If neither `-no-files` nor `-no-plots`: write certain plot files.
 - If `-calc-prob-all-pars`: calculate posterior probabilities over hypercubes Θ_{sub} .
 - Sort scalar samples of parameters $\{\theta^{(j)}\}$.

- Calculate medians, HPDIs (\nearrow 3.2.6) and marginal posteriors (\nearrow 3.2.4) of parameters.
 - If `-calc-prob`: calculate posterior probabilities over intervals I .
 - Calculate χ^2 .
 - If `-predict`: predict uncertainties of relevant observables (section 3.2.8).
 - If data are supplied and not `-periodogram`: calculate residuals.
 - Print summaries for current n_p mode.
 - If neither `-no-files` nor `-no-plots`: write certain plot files.
- (c) Print summaries for complete BASE run.
9. If neither `-generate-data` nor `-only-speed-meas`: execute Gnuplot.
10. If `-mail-when-finished`: send notification email.
11. *End*

MCMC sampling is carried out by the low-level routine `makeChain`, described in section 4.5.2. `makeChain` itself is called by `prlTempering`, which performs parallel tempering (section 4.5.3) and is in turn called by the top-level routine `multiPT`. The latter is started in step 8(b)iv above and described in section 4.5.4.

4.5.2 Posterior sampling by MCMC

As noted in section 3.2.3, the technique of MCMC helps to estimate the normalised posterior $\mathcal{P}(\boldsymbol{\theta})$ by drawing samples of the parameter vector distributed as the posterior. BASE uses the MCMC variant described by the *Metropolis-Hastings* algorithm (MH; Metropolis et al. 1953; Hastings 1970), which performs a random walk through parameter space, thereby collecting N samples $\{\boldsymbol{\theta}^{(j)} : j = 1, \dots, N\}$.

The distribution of these samples is not initially identical with the posterior but converges to it in the limit of many samples if the chain obeys certain *regularity conditions* (e.g. Roberts 1996). The first $M < N$ *burn-in* samples are still strongly correlated with the starting state $\boldsymbol{\theta}^{(0)}$ and excluded to improve convergence. Because the burn-in length M cannot be determined in advance, BASE considers a fixed *fraction* of samples to belong to the burn-in phase.⁷ Methods for setting the starting state $\boldsymbol{\theta}^{(0)}$ and detecting convergence are described in section 4.5.4.

Posterior sampling by MCMC is performed by the routine `makeChain`. Starting from the current chain link $\boldsymbol{\theta}^{(j+1)} \in \Theta$, the following steps lead to the next link, according to the MH algorithm and the *hit-and-run sampler* (step 1; Boneh and Golan 1979; Smith 1980):

1. Set up a candidate \mathbf{C} :
 - (a) sample a direction, viz. a random unit vector $\mathbf{d} \in \mathbb{R}^k$ from an isotropic density over the k -dimensional unit sphere;
 - (b) sample a (signed) distance r from a uniform distribution over the interval $\{r' \in \mathbb{R} : \boldsymbol{\theta}^{(j)} + r'\mathbf{d} \in \Theta\}$;
 - (c) set candidate $\mathbf{C} \equiv \boldsymbol{\theta}^{(j)} + r\mathbf{d}$;

⁷This fraction can be defined using the `-inference-len-frac` option (section 4.4.1).

2. calculate the acceptance probability

$$\alpha(\boldsymbol{\theta}^{(j)}, \mathbf{C}) \equiv \min \left(1, \frac{\mathcal{P}(\mathbf{C})}{\mathcal{P}(\boldsymbol{\theta}^{(j)})} \right); \quad (4.11)$$

3. draw a random number β from a uniform distribution $\mathcal{U}(0, 1)$ over the interval $[0, 1]$;
 4. if $\beta \leq \alpha$, accept the candidate, i.e. set the next link $\boldsymbol{\theta}^{(j+1)} \equiv \mathbf{C}$; otherwise, $\boldsymbol{\theta}^{(j+1)} \equiv \boldsymbol{\theta}^{(j)}$.

The hit-and-run sampler, compared to alternatives like the Gibbs sampler (Geman and Geman 1984), favours exploring of the whole parameter space Θ without becoming “trapped” in the vicinity of a local posterior maximum (Gilks et al. 1996).

Because only the ratio of two posterior values is used (step 2), the normalising evidence $p(D|\mathcal{M}, \mathcal{K})$ – a constant that is difficult to determine (\nearrow 3.2.1) – is irrelevant in the MH algorithm.

Thinning. Slow *mixing*, i.e. exploration of the parameter space by Markov chains, increases the number of samples needed to meaningfully characterise the posterior. In such cases, constraints of computer memory may prevent enough samples for convergence from being stored. Thus, it may sometimes be useful to only store one in a given number q of samples and discard the others. This simple concept, called *thinning*, can be activated in BASE with the `-thin` option, where the thinning stride $q \in \mathbb{N}^+$ can either be given or determined automatically from maximum runtime and maximum chain length such that both are approximately reached.

4.5.3 Improvement of mixing by parallel tempering

To enhance mixing and decrease the attraction of the chain by local posterior modes, the *parallel tempering* (PT) algorithm (e.g. Gregory 2005b) is employed one level above MH. Its function is to create in parallel n chains of length N , $\{\boldsymbol{\theta}_{(k)}^{(j)} : j = 1, \dots, N; k = 1, \dots, n\}$ each sampled by an independent MH procedure, where the *cold* chain $k = 1$ uses an unmodified likelihood $\mathcal{L}(\cdot)$, while the others, the *heated* chains, use as replacement $\mathcal{L}(\cdot)^{\gamma_{(k)}}$ with the positive *tempering parameter* $\gamma_{(k)} < 1$. After n_{swap} samples of each chain, two chains $k \geq 1$ and $k + 1 \leq n$ are randomly selected and their last links, denoted here by $\boldsymbol{\theta}_{(k)}$ and $\boldsymbol{\theta}_{(k+1)}$, are swapped with probability

$$\alpha_{\text{swap}}(k, k + 1) = \min \left(1, \left(\frac{\mathcal{L}(\boldsymbol{\theta}_{(k+1)})}{\mathcal{L}(\boldsymbol{\theta}_{(k)})} \right)^{\gamma_k} \left(\frac{\mathcal{L}(\boldsymbol{\theta}_{(k)})}{\mathcal{L}(\boldsymbol{\theta}_{(k+1)})} \right)^{\gamma_{k+1}} \right), \quad (4.12)$$

which ensures that the distributions of both chains remain unchanged.

This procedure allows states from the “hotter” chains, which explore parameter space more freely, to “seep through” to the cold chain without compromising its distribution. Conclusions are only drawn from the samples of the cold chain.

In contrast to MCMC sampling, which is sequential in nature, the structure of PT allows one to exploit the multiprocessing facilities provided by modern *symmetric multiprocessor* systems. For this purpose, BASE uses the OpenMP API (OpenMP Architecture Review Board 2008) as implemented for GFortran by the GOMP project (Free Software Foundation 2011b).

The following PT parameters can be adjusted by the user:

- The number of parallel chains n can be set with the option `-pt-n-chains`, with $n \equiv 1$ implying that PT is deactivated.
- To adjust the swapping stride n_{swap} , the option `-swap` can be used.
- Using the option `-temp-pars`, the tempering parameters $\{\gamma_{(k)} : k = 2, \dots, n\}$ can be set, whereas $\gamma_{(1)} \equiv 1$ is fixed. By default, $\gamma_{(n)} \equiv 10^{-3}$, and the tempering parameters are linearly decreasing from $\gamma_{(1)}$ through $\gamma_{(n)}$,

$$\gamma_{(k)} = 1 + \frac{k-1}{n-1} \cdot (10^{-3} - 1). \quad (4.13)$$

4.5.4 Assessing convergence by multi-PT

As a matter of principle, it cannot be proven that a given Markov chain has converged to the posterior. However, convergence may be meaningfully defined as the degree to which the chain does not depend on its initial state $\boldsymbol{\theta}^{(0)}$ any more. This can be determined on the basis of a set of independent chains – in BASE, these are the cold chains of m independent PT procedures – started at different points in parameter space. The procedure of BASE handling these PT procedures is called `multiPT`. The PT starting states should be defined such that for each of their components, they are *overdispersed* with respect to the corresponding marginal posteriors.

Since the marginal posteriors are difficult to obtain before the actual sampling, BASE determines the starting states by repeatedly drawing, for each parameter, a set of m samples from the prior using *rejection sampling*; the repetition is stopped as soon as the sample variance exceeds the corresponding prior variance, which yields a set of starting states overdispersed with respect to the prior. Assuming that the prior variance exceeds the marginal-posterior variance, the overdispersion requirement is met.

Such a test, using the *potential scale reduction (PSR)* or Gelman-Rubin statistic, was proposed by Gelman and Rubin (1992) and was later refined and corrected by Brooks and Gelman (1998). It is repeatedly carried out during sampling and, in the case of a positive result, sampling is stopped before the user-defined maximum runtime and/or number of samples have been reached. It may also sometimes be useful to abort sampling manually, e.g. when convergence does not appear to improve any further, which can be done at any time by creating an empty file `.BASE.finish` in the working directory. If this file is detected by BASE, sampling is finished at the current chain length, the file is deleted and all remaining procedures are carried out as if sampling had finished regularly.

The statistic is calculated with respect to each parameter θ separately, using the post-burn-in samples $\{\theta_l^{(j)} : j = M+1, \dots, N; l = 1, \dots, m\}$ of θ provided by the m independent chains. It compares the actual variances of θ within the chains built up thus far to an estimate of the marginal-posterior (i.e. target) variance of θ , thus estimating how closely the chains have approached convergence.

The PSR $\hat{\mathcal{R}}^{1/2}$ is defined as

$$\hat{\mathcal{R}} \equiv \frac{d+3}{d+1} \frac{\hat{V}}{\bar{W}}, \quad (4.14)$$

where \hat{V} is an estimate of the marginal-posterior variance, d the estimated number of degrees of freedom underlying the calculation of \hat{V} , and \bar{W} the mean within-sequence

variance. The quantities are defined as

$$\hat{V} \equiv \frac{\nu - 1}{\nu} W + \frac{m + 1}{m\nu} B \quad (4.15)$$

$$W \equiv \frac{1}{m(\nu - 1)} \sum_{l=1}^m \sum_{j=M+1}^N (\theta_l^{(j)} - \bar{\theta}_l)^2 \quad (4.16)$$

$$B \equiv \frac{\nu}{m - 1} \sum_{l=1}^m (\bar{\theta}_l - \bar{\theta})^2, \quad (4.17)$$

where B is the between-sequence variance and $\nu \equiv N - M$; a horizontal bar denotes the mean taken over the set of samples obtained by varying the omitted indexes.

Owing to the initial overdispersion of starting states, \hat{V} overestimates the marginal-posterior variance in the beginning and subsequently decreases. Furthermore, while the chains are still exploring new areas of parameter space, W underestimates the marginal-posterior variance and increases. Thus, as convergence to the posterior is accomplished, $\hat{\mathcal{R}}^{1/2} \searrow 1$. Therefore, convergence can be assumed as soon as the PSR has fallen below a threshold $\mathcal{R}_*^{1/2} \geq 1$. If BASE is run with $\mathcal{R}_* \equiv 1$, sampling is continued up to the user-defined length or duration, respectively.

For cyclic parameters (\nearrow 2.1.1), the default lower and upper prior bounds are equivalent, which needs to be taken into account when calculating the PSR. Thus, BASE uses the modified definition of the PSR by Ford (2006) for these parameters.

4.5.5 Organisation of source code

The source code of BASE is organised in a main program, `main`, and 22 modules, each contained in a separate file and comprising data and procedures referring to a specific field of work (table 4.9). Each module can make use of, i.e. depend on, other modules by means of Fortran's `use` statement, giving rise to a dependence hierarchy shown in fig. 4.2.

⁸In BASE, specific constants are assigned to variables to mark them as *undefined*, which simplifies the code in many instances. These values are chosen to lie near the lowest storable value of a given data type.

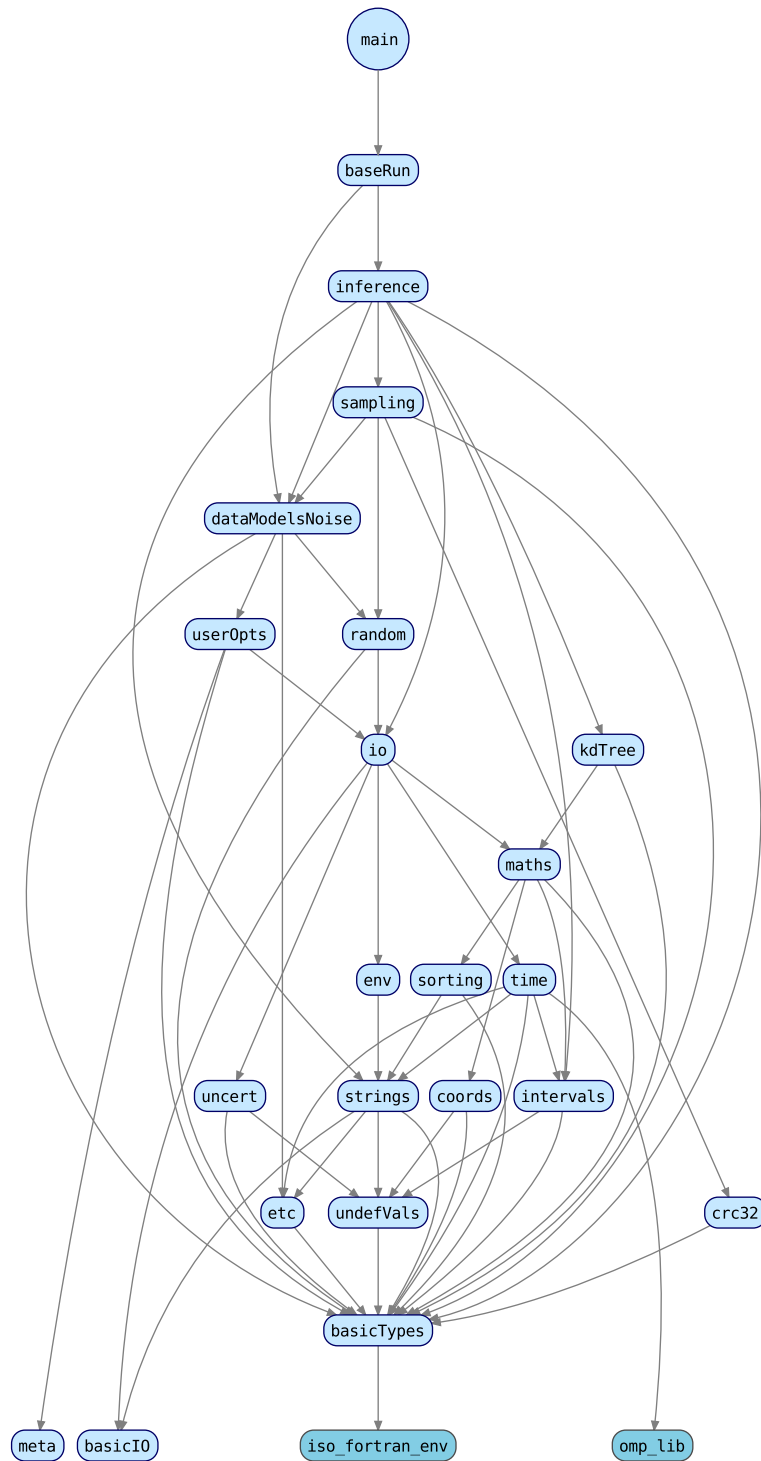


Figure 4.2: Dependency graph of the source modules and main program of BASE. Arrows point from the using module to the used module. The modules `iso_fortran_env` and `omp_lib` are part of GFortran.

Table 4.9: Modules of BASE.

Module	Field of work/tasks
<code>baseRun</code>	Complete runs of BASE
<code>basicIO</code>	Basic in-/output
<code>basicTypes</code>	Basic data types and conversions
<code>coords</code>	Coordinates in \mathbb{R}^2 and \mathbb{R}^3
<code>crc32</code>	Cyclic redundancy check (CRC-32)
<code>dataModelsNoise</code>	Data, models, and noise
<code>env</code>	Computing environment
<code>etc</code>	Further utilities
<code>inference</code>	Posterior inference
<code>intervals</code>	Intervals in \mathbb{R}
<code>io</code>	(Formatted) in-/output
<code>kdTree</code>	kd-trees
<code>maths</code>	Mathematics
<code>meta</code>	Name, author, and version information for BASE
<code>random</code>	Random numbers
<code>sampling</code>	MCMC sampling, PT, Multi-PT, reading and saving samples
<code>sorting</code>	Sorting and other tasks on arrays
<code>strings</code>	Strings
<code>time</code>	Date and time
<code>uncert</code>	Representing quantities with their uncertainties
<code>undefVals</code>	Undefined values of variables ⁸
<code>userOpts</code>	User options

4.6 Specific algorithms

Below, a selection of algorithms implemented in BASE are detailed. These concern inference in periodogram mode, saving and reading posterior samples, and generating synthetic data.

4.6.1 Periodogram mode

By default, the window width for marginal posteriors is based on the MCMC sample standard deviation σ_{samp} (eq. (3.41)). If there are multiple maxima, however, this can lead to artificially broad peaks, which can be particularly problematic for the orbital frequency f (see section 2.1.1), which plays an important role in distinguishing different solutions in orbit-related parameter estimation. Therefore, BASE includes a *periodogram mode*, in which the window width of the marginal posterior of f – a Bayesian analogon to the frequentist periodogram (\nearrow 3.1.4) – is reduced according the following procedure.

1. Initially assume a default window width as given by eq. (3.41);
2. estimate the marginal posterior of f and find its local maximum f_{max} nearest to posterior mode \hat{f} , as well as the local minimum f_{min} nearest to f_{max} ;
3. calculate the marginal-posterior standard deviation σ' over the half-peak between f_{max} and f_{min} ;
4. re-calculate the window width using eq. (3.41), with $\sqrt{2}\sigma'$ replacing σ_{samp} ;

5. repeat step 2, but only consider local minima with ordinates $p(f_{\min}|D, \mathcal{M}, \mathcal{K}) \leq 0.5 \cdot p(f_{\max}|D, \mathcal{M}, \mathcal{K})$ in order not to be misled by weak marginal-posterior fluctuations;
6. repeat steps 3 and 4.

4.6.2 Saving and reading samples

Posterior samples $\{\boldsymbol{\theta}^{(j)}\}$ gathered by BASE can be saved to a file and read in again afterwards, which can be useful when one wishes to make different types of *a posteriori* inference based on the same data, models and prior knowledge. For example, BASE may first be run with the option `-no-plots` in order to obtain textual summaries for a first inspection. If the option `-save-samples` was used, BASE may then be run again with `-read-samples` to re-use the previous posterior samples and produce plots from them by omitting `-no-plots`. To read in samples, the same data files as before must be given in the original order and potential `-group` options also need to be unchanged.

BASE stores posterior samples, along with other relevant information, in *saved-samples files*, which consist of the following parts:

1. *global header* (table 4.10)
2. for each n_p mode carried out or once for binary mode, respectively (table 4.11):
 - (a) *n_p -mode header*
 - (b) *samples*
3. CRC-32 checksum (table 4.12)

After reading the global header (part 1), all available n_p modes (part 2) are read until the requested mode is found, whose samples are then used. While reading parts 1 and 2, a CRC-32 checksum is calculated and compared to the saved checksum (part 3) after reading; in case of a mismatch, an error is cast in order to prevent using a corrupted saved-samples file.

4.6.3 Synthetic data

Based on the observable and noise models introduced in sections 2.2 and 2.4, a set of N_D synthetic data can be easily generated by:

1. Sampling a set of observing times $\{t_i : i = 1, \dots, N_D\}$ as described below;
2. for each time t_i :
 - (a) calculating the model function $\mathbf{f}_i^0 \equiv \mathbf{f}(t_i; \boldsymbol{\theta})$;
 - (b) adding random noise $\epsilon_{0,i}$ to obtain the synthetic datum $\mathbf{f}_i \equiv \mathbf{f}_i^0 + \epsilon_{0,i}$.

⁹The data type of a record, using the abbreviations int (integer) and chr (character), followed by a number specifying the kind, i.e. size in bytes, and/or the number of elements in the array constituting the record (in brackets).

¹⁰(undef) is a special value signifying an undefined variable; see footnote 8.

¹¹The data type of a record, using the abbreviations int (integer) and chr (character), followed by a number specifying the kind, i.e. size in bytes, and/or the number of elements in the array constituting the record (in brackets).

Table 4.10: Saved-samples file: global header.

Length	ID	Content	Remarks	Type ⁹	Count
	<i>A</i>	“BASE posterior samples”		chr(22)	1
	<i>B</i>	BASE revision number		int2	1
	<i>C</i>	—	reserved	chr(128)	1
	<i>D</i>	$length(T)$		int2	1
	<i>E</i>	$length(U)$		int2	1
	<i>F</i>	$length(V)$		int2	1
	<i>G</i>	# configuration-file lines		int2	1
	<i>H</i>	date/time BASE was started		int4×8	1
	<i>I</i>	# data sets		int2	1
const.	<i>J</i>	t_1		real8	1
	<i>K</i>	t_N		real8	1
	<i>L</i>	epoch t_r		real8	1
	<i>M</i>	times in JD scale?		int1	1
	<i>N</i>	# data		int4	1
	<i>O</i>	data CRC		int4	1
	<i>P</i>	# n_p modes	$\langle \text{undef} \rangle^{10}$ in binary mode	int1	1
	<i>Q</i>	# PT procedures in multi-PT		int4	1
	<i>R</i>	# chains in PT		int4	1
	<i>S_i</i>	data type		int1	<i>I</i>
	<i>T</i>	command line		chr(<i>D</i>)	1
var.	<i>U</i>	host-computer name		chr(<i>E</i>)	1
	<i>V</i>	target-system identifier		chr(<i>F</i>)	1
	<i>W_i</i>	$length(X_i)$		int2	<i>G</i>
	<i>X_i</i>	configuration-file line		chr(<i>W_i</i>)	<i>G</i>

Besides N_D and the type of data (AM_a or RV) to be generated, the following information is specified by the user with the option `-generate-data`:

1. the earliest and latest observing times, t_1 and t_N
2. the number of planets n_p and stellar mass m_* (unless binary-mode data is being generated) as well as the parameters θ
3. $n_{\text{unc}} \geq 1$ possible observational uncertainties $\{\sigma_j\}$, for RV data, or $\{(a_j, b_j)\}$, for AM_a data (\nearrow 2.4), from which one is drawn randomly with replacement in step 2b of the generation algorithm

Additionally, the data commands listed in table 4.13 may be given to modify the procedure as follows:

- AM_a data may be generated in a polar rather than cartesian coordinate system by using `$coord_sys=polar`.
- For RV data, `$n_b_obs=2` can be used to switch to binary mode,¹² generating one

¹²According to section 4.4.2, binary mode is automatically selected for AM_a data.

Table 4.11: Saved-samples file: n_p -mode header and samples.

Length	ID	Content	Remarks	Type ¹¹	Count
const.	<i>Y</i>	# planets	$\langle \text{undef} \rangle$ in binary mode	int2	1
	<i>Z</i>	# local parameters		int2	1
	<i>AA</i>	# samples (total)		int8	1
	<i>AB</i>	# burn-in samples		int8	1
	<i>AC</i>	thinning stride		int8	1
	<i>AD</i>	mean PT block length		int8	1
	<i>AE</i>	tempering parameters adjusted?		int1	1
	<i>AF</i>	finished by user request?		int1	1
	<i>AG</i>	compression mode	0: none; 1: RLE	int1	1
var.	<i>AH</i>	tempering parameter		real8	<i>R</i> <i>Q</i>
	<i>AI_i</i>	$length(AJ_i)$		int2	<i>Z</i>
	<i>AJ_i</i>	local parameter name		chr(<i>AI_i</i>)	<i>Z</i>
	<i>AK</i>	prior type		int1	<i>Z</i>
	<i>AL</i>	prior lbound		real8	<i>Z</i>
	<i>AM</i>	prior ubound		real8	<i>Z</i>
	<i>AN</i>	prior knee	$\langle \text{undef} \rangle$ if not applying to prior type	real8	<i>Z</i>
	<i>AO</i>	prior: underlying Gaussian variance		real8	<i>Z</i>
	<i>AP</i>	# user-def. prior points		int4	<i>Z</i>
	<i>AQ</i>	user-def. prior abscissae		real8	<i>AP</i> <i>Z</i>
	<i>AR</i>	user-def. prior ordinates		real8	<i>AP</i> <i>Z</i>
	<i>AS</i>	user-def. prior: log-scale abscissa?		int1	<i>Z</i>
	<i>AT</i>	user-def. prior: inverse delta		real8	<i>Z</i>
	<i>AU</i>	final PSR value		real8	<i>Z</i>
<i>AV</i>	fraction of indep. samples		real8	<i>Z</i>	
<i>AW</i>	posterior mode		real8	<i>Z</i>	
var.	<i>AX</i>	sample	if <i>AG</i> = 0	real8	<i>AA</i> <i>Z</i>
	<i>AY</i>	# sequences	if <i>AG</i> = 1	int8	1
	<i>AZ</i>	sequence size	if <i>AG</i> = 1	int1	<i>AY</i> # free pars.
	<i>BA</i>	value	if <i>AG</i> = 1	real8	

Table 4.12: Saved-samples file: checksum.

Length	ID	Content	Remarks	Type ¹¹	Count
const	<i>BB</i>	check sum	CRC-32	int4	1

Table 4.13: Data commands for generating data.

Command	Meaning
<code>\$coord_sys=polar</code>	Polar coordinate system
<code>\$n_b_obs=2</code>	Two bodies observed (for RV data)
<code>\$unif_time</code>	Apply uniform time sampling
<code>\$window=[⟨period⟩,⟨fraction⟩]</code>	Allow data only in given time window
<code>\$times=[⟨time⟩[,⟨time⟩[...]]]</code>	Explicitly set observing times to generate data at
<code>\$with_ang_pos</code>	When generating RV data, choose set of parameters pertaining to AM _a and RV data
<code>\$with_RV</code>	When generating AM _a data, choose set of parameters pertaining to AM _a and RV data

data set each for the primary and secondary components.

- The times $\{t_i : i = 2, \dots, N_D - 1\}$ are either sampled randomly from $\mathcal{U}(t_1, t_N)$ or, if `$unif_time` is given, distributed such that $\{t_i : i = 1, \dots, N_D\}$ are uniformly spaced.
- When the observing times are randomly sampled, any number of *time windows* $\{(P_j, \rho_j)\}$ can be set up, each defined by a period P_j and a signed fraction ρ_j , implying that for any sampled time t_i the following condition needs to be fulfilled:

$$\frac{\text{mod}(t_i - t_1, P_j)}{P_j} \begin{cases} \leq \rho_j, & \rho_j > 0 \\ \geq 1 + \rho_j, & \rho_j \leq 0 \end{cases} \quad (4.18)$$

- When generating RV data, the set of parameters for the case of both RV and AM_a data can be used, and vice versa when generating AM_a data (commands `$with_ang_pos` and `$with_RV`, respectively).

Chapter 5

Bayesian analysis of exoplanet and binary orbits

Demonstrated using astrometric and radial-velocity data of Mizar A

*This chapter reproduces an article published in *Astronomy&Astrophysics* (Schulze-Hartung et al. 2012) unaltered in content. Therefore, some of the text of previous chapters of this thesis is repeated here.*

Aims. We introduce BASE (Bayesian astrometric and spectroscopic exoplanet detection and characterisation tool), a novel program for the combined or separate Bayesian analysis of astrometric and radial-velocity measurements of potential exoplanet hosts and binary stars. The capabilities of BASE are demonstrated using all publicly available data of the binary Mizar A.

Methods. With the Bayesian approach to data analysis we can incorporate prior knowledge and draw extensive posterior inferences about model parameters and derived quantities. This was implemented in BASE by Markov chain Monte Carlo (MCMC) sampling, using a combination of the Metropolis-Hastings, hit-and-run, and parallel-tempering algorithms to explore the whole parameter space. Nonconvergence to the posterior was tested by means of the Gelman-Rubin statistic (potential scale reduction). The samples were used directly and transformed into marginal densities by means of kernel density estimation, a “smooth” alternative to histograms. We derived the relevant observable models from Newton’s law of gravitation, showing that the motion of Earth and the target can be neglected.

Results. With our methods we can provide more detailed information about the parameters than a frequentist analysis does. Still, a comparison with the Mizar A literature shows that both approaches are compatible within the uncertainties.

Conclusions. We show that the Bayesian approach to inference has been implemented successfully in BASE, a flexible tool for analysing astrometric and radial-velocity data.

5.1 Introduction

The search for extrasolar planets – places where one day humankind might find other forms of life in the Universe – has been a subject of scientific investigation since the nineteenth century, but only became successful in 1992 with the first confirmed discovery of an exoplanet orbiting the pulsar PSR B1257+12 (Wolszczan and Frail 1992). Still, because

it is situated in an environment hostile to life as we know it, this case has been of less relevance to the public than the first detection of a Sun-like planet-host star, 51 Pegasi (Mayor and Queloz 1995). Since that time, more than 700 extrasolar planet candidates have been unveiled in more than 500 systems, more than 90 of which show signs of multiplicity (Schneider 2012).

Closely related to the *detection* of extrasolar planets is the *characterisation* of their orbits. Both these tasks now profit from the existence of a variety of observational techniques, which we briefly sketch in the following. Comprehensive reviews can be found in Perryman (2000) and Deeg et al. (2007).

Direct observational methods refer to the *imaging* of exoplanets (e.g. Levine et al. 2009), which reflect the light of their host stars but also emit their own thermal radiation. To overcome the major obstacle of the high brightness contrast between planet and star, techniques such as *coronagraphy* (Lyot 1932; Levine et al. 2009), *angular differential imaging* (Marois et al. 2006; Vigan et al. 2010), *spectral differential imaging* (Smith 1987; Vigan et al. 2010), and *polarimetric differential imaging* (Kuhn et al. 2001; Adamson et al. 2005) have been invented. Still, imaging has only revealed few detections and orbit determinations so far.

The most productive methods in terms of the number of detected and characterised exoplanets are of an *indirect* nature, observing the effects of the planet on other objects or their radiation.

Of these, *transit photometry* (e.g. Charbonneau et al. 2000; Seager 2008) is noteworthy because it has helped uncover more than 200 exoplanet candidates, plus over 2000 still unconfirmed candidates from the Kepler space mission (Koch et al. 2010): small decreases in the apparent visual brightness of a star during the *primary* or *secondary eclipse* point to the existence of a transiting companion. These data allow one to determine the planet's radius and orbital inclination and may also yield information on the planet's own radiation.

Timing methods include measurements of *transit timing variations (TTV)* and *transit duration variations (TDV)* (e.g. Holman and Murray 2005; Nascimbeni et al. 2011) of either binaries or stars known to harbour a transiting planet. The method used in the first exoplanet detection (Wolszczan and Frail 1992) is *pulsar timing*, which relies on slight anomalies in the exact timing of the radio emission of a pulsar and is sensitive to planets in the Earth-mass regime.

Microlensing (Mao and Paczynski 1991; Gould 2009), which accounted for about 15 exoplanet candidates, uses the relativistic curvature of spacetime due to the masses of both a lens star and its potential companion, with the latter causing a change in the apparent magnification and thus the observed brightness of a background source.

Perhaps the most well-known technique, and one of those on which this article is based, is known as *Doppler spectroscopy* or *radial-velocity (RV)* measurements (e.g. Mayor and Queloz 1995; Lovis and Fischer 2010). With more than 500 exoplanet candidates, it has been most successful in detecting new exoplanets and determining their orbits to date. From a set of high-resolution spectra of the target star, a time series of the line-of-sight velocity component of the star is deduced. These data allow one to determine the orbit in terms of its geometry and kinematics in the *orbital plane* as well as the minimum planet mass $m_{p,\min} \approx m_p \sin i$. To derive the actual planet mass m_p , the inclination i of the orbit plane with respect to the sky plane needs to be derived with a different method, e.g. astrometry. The RV technique is distance-independent by principle, but signal-to-noise requirements do pose constraints on the maximum distance to a star. Stellar variability sometimes makes this approach difficult because it alters the line shapes and thus mimicks

RV variations. The signal in stellar RVs caused by a planet in a circular orbit has a semi-amplitude of approximately

$$K \approx m_p \sin i \sqrt{\frac{G}{m_\star a_{\text{rel}}}}, \quad (5.1)$$

where m_p , m_\star , i , G , and a_{rel} are the masses of planet and host star, the orbital inclination, Newton’s gravitational constant, and the semi-major axis of the planet’s orbit relative to the star, respectively. This approximation holds for $m_p \ll m_\star$, which is true in most cases. It should be noted that the sensitivity of the RV method decreases towards less inclined (more face-on) orbits, which is an example for the selection effects inherent to any planet-detection method.

Finally, *astrometry* (AM; e.g. Gatewood et al. 1980; Sozzetti 2005; Reffert 2009) – on which this work is also based – is the oldest observational technique known in astronomy: a stellar position is measured with reference to a given point and direction on sky. Astrometry can thus be considered as complementary to Doppler spectroscopy, which measures the kinematics perpendicular to the sky plane. In contrast to the latter, AM allows one to determine the orientation of the orbital plane relative to the sky in terms of its inclination i and the position angle Ω of the *line of nodes* with respect to the meridian of the target. A planet in circular orbit around its host star displaces the latter on sky with an approximate angular semi-amplitude of

$$\alpha \approx \frac{m_p a_{\text{rel}}}{m_\star d}, \quad (5.2)$$

where d is the distance between the star and the observer. Again, this approximation holds for $m_p \ll m_\star$.

Imaging astrometry, in its attempt to reach sufficient precision, still faces problems due to various distortion effects. By contrast, *interferometric* astrometry has been used to determine the orbits of previously known exoplanets, mainly with the help of space-borne telescopes such as Hipparcos or the Hubble Space Telescope (HST), which presently still excel their Earth-bound competitors (e.g. McArthur et al. 2010). However, instruments like PRIMA (Delplancke et al. 2000; Delplancke 2008; Launhardt et al. 2008) or GRAVITY (Gillessen et al. 2010) at the ESO Very Large Telescope Interferometer are promising to advance ground-based AM even more in the near future.

While planet-induced signals in AM and RVs are both approximately linear in planetary mass m_p , they differ in their dependence on the orbital semi-major axis a_{rel} (eq. (5.1) and (5.2)). Doppler spectroscopy is more sensitive to smaller orbits (or higher orbital frequencies, eq. (5.45)), while AM favours larger orbital separations, viz. longer periods.

In this article we introduce BASE, a **B**ayesian **a**strometric and **s**pectroscopic **e**xoplanet detection and characterisation tool. Its goals are to fulfil two major tasks of exoplanet science, namely the *detection of exoplanets* and the *characterisation of their orbits*. BASE has been developed to provide for the first time the possibility of an integrated Bayesian analysis of stellar astrometric and Doppler-spectroscopic measurements with respect to their companions’ signals,¹ correctly treating the measurement uncertainties and allowing one to explore the whole parameter space without the need for informative prior constraints. Still, users may readily incorporate prior knowledge, e.g. from previous analyses with other tools, by means of priors on the model parameters. The tool automatically diagnoses

¹Although most methods described in this introduction apply to any kind of companion to a star, we refer here to companions as “exoplanets”, irrespective of whether they are able to sustain hydrogen or deuterium burning.

convergence of its Markov chain Monte Carlo (MCMC) sampler to the posterior and regularly outputs status information. For orbit characterisation, BASE delivers important results such as the probability densities and correlations of model parameters and derived quantities.

Because published high-precision AM observations of potential exoplanet host stars are still sparse, we used data of the well-known binary Mizar A to demonstrate the capabilities of BASE. It is also planned to gain astrophysical insights into exoplanet systems using BASE in the near future.

This article is organised as follows. Section 5.2 provides an overview of the most often-used methods of data analysis, including Bayes' theorem and MCMC as theoretical and implementational foundations of this work, as well as a derivation of the necessary observable models. BASE is described in section 5.3. In section 5.4, the target Mizar A and the data used in this article are discussed. Section 5.5 presents and discusses our analysis of Mizar A. Conclusions are drawn in section 5.6.

5.2 Methods and models

Data analysis is a type of *inductive reasoning* in that it infers general rules from specific observational data (e.g. Gregory 2005b). These general rules are described by *observable models*, simply called *models* in the following, which produce theoretical values of the observables as a function of *parameters*. The primary tasks of data analysis are listed in the following.

1. In *model selection*, the relative probabilities of a set of concurrent models $\{\mathcal{M}_i\}$, chosen a priori, are assessed. Specifically, *exoplanet detection* tries to decide the question of whether a certain star is accompanied by a planet or not, based on available data. Additionally, *model-assessment* techniques can be used to determine whether the most probable model describes the data accurately enough.
2. *Parameter estimation* aims to determine the parameters θ of a chosen model. This is specifically referred to as *exoplanet characterisation* (or orbit determination) in the present context.
3. The purpose of *uncertainty estimation* is to provide a measure of the parameters' uncertainties.

Although model selection is equally important, in what follows we focus entirely on the second and third tasks, viz. parameter and uncertainty estimation. This is because for a known binary system, only one model is appropriate, viz. two bodies orbiting each other. Accordingly, BASE can only perform model selection when it analyses data from stars for which it is a-priori unknown whether a companion exists.

5.2.1 Likelihoods and frequentist inference

The well-established, conventional *frequentist* approach to inference is touched upon only briefly here. Its name stems from the fact that it defines probability as the relative frequency of an event. Measurements are regarded as values of *random variables* drawn from an underlying *population* that is characterised by *population parameters*, e.g. mean and standard deviation in the case of a normal (Gaussian) population. In the following,

we derive the joint probability density of the values of AM and RV data, known as the *likelihood* \mathcal{L} , which plays a central role in frequentist inference.

When combining data of different types, one should generally be aware that potential systematic errors may differ between the data sets, e.g. due to a calibration error in one instrument that renders the data inconsistent with each other. In this case, each data set analysed separately would imply a different result. In other instances, systematic errors in one set do not affect the other data: for example, any constant offset in radial velocity is absorbed into parameter V , which is irrelevant to the analysis of astrometric data. In the following derivation, we assumed that no systematic effects are present that led to inconsistent data.

In the following, we assumed that the error ϵ_i of any datum y_i is statistically independent of those of all other data and consists of two components, each distributed according to a (uni- or bivariate) normal distribution with zero mean:

- a component $\epsilon_{0,i}$ corresponding to a nominal measurement error, whose distribution is characterised by the covariance matrix $\mathbf{E}_{0,i}$ or standard deviation $\sigma_{0,i}$ given with the datum, for AM and RV data respectively, and
- a component $\epsilon_{+,i}$ representing e.g. instrumental, atmospheric or stellar effects not modelled otherwise, whose distribution is characterised by scalar covariance matrix $\mathbf{E}_+ = \text{diag}(\tau_+^2, \tau_+^2)$ – assuming no correlation between the noise in the two AM components – or variance σ_+^2 , respectively, where τ_+ and σ_+ are free noise-model parameters.

The AM *data covariance matrix* $\mathbf{E}_{0,i}$ of datum i , representing the uncertainty of and correlation between the two components measured, can be written using singular-value decomposition as

$$\mathbf{E}_{0,i} = \mathbf{R}(-\phi_i) \begin{pmatrix} a_i^2 & 0 \\ 0 & b_i^2 \end{pmatrix} \mathbf{R}(\phi_i), \quad (5.3)$$

where $\mathbf{R}(\cdot)$, a_i , b_i and ϕ_i are the 2×2 *passive rotation matrix*, the nominal semi-major and -minor axes of the *uncertainty ellipse* and the position angle of its major axis, respectively.

Using characteristic functions, it is readily shown that the sum $\epsilon_i = \epsilon_{0,i} + \epsilon_{+,i}$ of the two independent error components is again normally distributed, with zero mean and covariance matrix $\mathbf{E}_i = \mathbf{E}_{0,i} + \mathbf{E}_+$ or standard deviation $\sigma_i = (\sigma_{0,i}^2 + \sigma_+^2)^{1/2}$, respectively.

The probability density² of the values of N_{AM} two-dimensional AM data $\{\mathbf{r}_i\}$ and N_{RV} RV data $\{v_i\}$, known as the *likelihood* \mathcal{L} , is then given by

$$\mathcal{L} = \mathcal{L}_{\text{AM}} \mathcal{L}_{\text{RV}}, \quad (5.4)$$

where \mathcal{L}_{AM} and \mathcal{L}_{RV} are the likelihoods pertaining to the individual data types,

$$\mathcal{L}_{\text{AM}} = \left((2\pi)^{N_{\text{AM}}} \prod_{i=1}^{N_{\text{AM}}} \sqrt{\det \mathbf{E}_i} \right)^{-1} \exp \left(-\frac{1}{2} \chi_{\text{AM}}^2 \right), \quad (5.5)$$

$$\mathcal{L}_{\text{RV}} = \left((2\pi)^{\frac{N_{\text{RV}}}{2}} \prod_{i=1}^{N_{\text{RV}}} \sigma_i \right)^{-1} \exp \left(-\frac{1}{2} \chi_{\text{RV}}^2 \right). \quad (5.6)$$

²Throughout, we use the term *probability density* wherever it refers to a continuous quantity, as opposed to *probability* for discrete quantities. *Probability distribution*, denoted by $p(\cdot)$, is a generic term used for both cases.

Furthermore, the *sums of squares* χ_{AM}^2 and χ_{RV}^2 are defined by

$$\chi_{\text{AM}}^2 \equiv \sum_{i=1}^{N_{\text{AM}}} (\mathbf{r}_i - \mathbf{r}(\boldsymbol{\theta}; t_i))^{\top} \mathbf{E}_i^{-1} (\mathbf{r}_i - \mathbf{r}(\boldsymbol{\theta}; t_i)), \quad (5.7)$$

$$\chi_{\text{RV}}^2 \equiv \sum_{i=1}^{N_{\text{RV}}} \left(\frac{v_i - v(\boldsymbol{\theta}; t_i)}{\sigma_i} \right)^2, \quad (5.8)$$

where \mathbf{r}_i, v_i are the i th AM and RV datum, $\mathbf{r}(\cdot; \cdot), v(\cdot; \cdot)$ the AM and RV model functions and $\boldsymbol{\theta}$ is the vector of model parameters. The relevant models are derived in section 5.2.3.

Parameter estimation. Frequentist parameter estimation is generally equivalent to maximising \mathcal{L} or minimising χ^2 as functions of $\boldsymbol{\theta}$. The resulting *best estimates* of the parameters $\hat{\boldsymbol{\theta}}$ are therefore often called *maximum-likelihood* or *least-squares estimates*. For linear models, $\chi^2(\boldsymbol{\theta})$ is a quadratic function and consequently $\hat{\boldsymbol{\theta}}$ can be found unambiguously by matrix inversion. In the more realistic cases of nonlinear models, however, χ^2 may have many local minima, therefore care needs to be taken not to mistake a local minimum for the global one. Several methods exist to this end, including evaluation of $\chi^2(\boldsymbol{\theta})$ on a finite grid, *simulated annealing* or *genetic algorithms* (e.g. Gregory 2005b).

Uncertainty estimation. Frequentist parameter uncertainties are usually quoted as *confidence intervals*. Procedures to derive these are designed such that when repeated many times based on different data, a certain fraction of the resulting intervals will contain the true parameters. Popular methods use *bootstrapping* (Efron and Tibshirani 1993) or the *Fischer information matrix*, which is based on a local linearisation of the model (e.g. Ford 2004). However, these methods suffer from specific caveats: the Fischer matrix is only appropriate for a quadratic-shaped χ^2 in the vicinity of the minimum, and bootstrapping, which relies on modified data, may lead to severe misestimation of the parameter uncertainties, especially when these are large (Vogt et al. 2005).

5.2.2 Bayesian inference

Bayesian inference (e.g. Sivia 2006), which has gained popularity in various scientific disciplines during the past few decades, defines probability as the degree of belief in a certain *hypothesis* \mathcal{H} . While this is sometimes criticised as leading to subjective assignments of probabilities, Bayesian probabilities are not subjective if they are based on all relevant knowledge \mathcal{K} , hence different people having the same knowledge will assign them the same value (e.g. Sivia 2006). Thus, Bayesian probabilities are conditional on the knowledge \mathcal{K} , and this conditionality should be stated explicitly, as in the following equations.

Bayes' theorem

In the eighteenth century, Thomas Bayes laid the foundation of a new approach to inference with what is now known as Bayes' theorem (Bayes and Price 1763). For the purpose of parameter and uncertainty estimation, the hypothesis \mathcal{H} refers to the values of model parameters $\boldsymbol{\theta}$, and Bayes' theorem can be expressed as

$$p(\boldsymbol{\theta}|D, \mathcal{M}, \mathcal{K}) = \frac{p(D|\boldsymbol{\theta}, \mathcal{M}, \mathcal{K}) \cdot p(\boldsymbol{\theta}|\mathcal{M}, \mathcal{K})}{p(D|\mathcal{M}, \mathcal{K})}, \quad (5.9)$$

where $p(\cdot)$ is a probability (density). Furthermore, $D \equiv \{(t_i, \mathbf{y}_i)\}$ is the set of pairs of observational times³ and corresponding data values, and \mathcal{M} denotes the particular model assumed. As mentioned above, all probabilities are also conditional on the knowledge \mathcal{K} , including statements on the types of parameters and the parameter space Θ (which we assume to be a subset of \mathbb{R}^k with $k \in \mathbb{N}$) as well as the noise model.

Using Bayes' theorem, the aim is to determine the *posterior* $\mathcal{P}(\boldsymbol{\theta}) \equiv p(\boldsymbol{\theta}|D, \mathcal{M}, \mathcal{K})$, i.e. the probability distribution of the parameters $\boldsymbol{\theta}$ in light of the data D , given the model \mathcal{M} and prior knowledge \mathcal{K} . The other terms, located on the right-hand side of the theorem, are explained below.

- The term *prior* refers to the probability distribution $p(\boldsymbol{\theta}|\mathcal{M}, \mathcal{K})$ of the parameters $\boldsymbol{\theta}$ given only the model and prior knowledge \mathcal{K} ; it characterises the knowledge about the parameters present before considering the data. For objective choices of priors, based on classes of parameters, see section 5.7.
- The *likelihood* $p(D|\boldsymbol{\theta}, \mathcal{M}, \mathcal{K})$ is the probability distribution of the data values D , given the times of observation, the model and the parameters. It is introduced in the context of frequentist inference in section 5.2.1.
- The *evidence* is the probability distribution of the data values D , given the times of observation and the model but neglecting the parameter values,

$$p(D|\mathcal{M}, \mathcal{K}) = \int p(D, \boldsymbol{\theta}|\mathcal{M}, \mathcal{K}) d\boldsymbol{\theta} \quad (5.10)$$

$$= \int p(\boldsymbol{\theta}|\mathcal{M}, \mathcal{K}) p(D|\boldsymbol{\theta}, \mathcal{M}, \mathcal{K}) d\boldsymbol{\theta}. \quad (5.11)$$

It equals the integral of the product of prior and likelihood over the parameter space Θ and plays the role of a normalising constant, which is hard to calculate in practice, however.

It may be instructive here to note that the frequentist approach of maximising the likelihood $p(D|\boldsymbol{\theta}, \mathcal{M}, \mathcal{K})$ is equivalent to maximising the posterior when assuming uniform priors $p(\boldsymbol{\theta}|\mathcal{M}, \mathcal{K})$. This can be seen by inserting $p(\boldsymbol{\theta}|\mathcal{M}, \mathcal{K}) = \text{const}$ into eq. (5.9), which leads to

$$\mathcal{P}(\boldsymbol{\theta}) = p(\boldsymbol{\theta}|D, \mathcal{M}, \mathcal{K}) \propto p(D|\boldsymbol{\theta}, \mathcal{M}, \mathcal{K}). \quad (5.12)$$

However, this maximum-likelihood approach ignores the fact that uniform priors are not always the most objective choice (see section 5.7) and the posterior cannot be fully characterised just by the position of its maximum. Still, the latter can be used as a posterior summary in the Bayesian framework (section 5.2.2).

Posterior inference

Sampling from the posterior. To estimate the normalised posterior $\mathcal{P}(\boldsymbol{\theta})$, i.e. the probability distribution of the parameters $\boldsymbol{\theta}$ in light of the data, N samples of the parameter vector $\{\boldsymbol{\theta}^{(j)} : j = 1, \dots, N\}$ are collected using the *Markov chain Monte Carlo* method (*MCMC*; e.g. Gilks et al. 1996) in the variant described by the *Metropolis-Hastings*

³In general, t_i is the value of an *independent variable*, which may e.g. be temporal or spatial. We assumed the measurement durations to be short in comparison with the characteristic time of orbital motion, given by the orbital period, and thus the observations to take place at points in time t_i which are known exactly.

algorithm (MH; Metropolis et al. 1953; Hastings 1970), which performs a random walk through parameter space. The distribution of these samples – excluding the first $M < N$ *burn-in* samples, which are still strongly correlated with the starting state $\boldsymbol{\theta}^{(0)}$ – converges to the posterior $\mathcal{P}(\cdot)$ in the limit of many samples if the chain obeys certain *regularity conditions* (e.g. Roberts 1996). Methods for setting the starting state $\boldsymbol{\theta}^{(0)}$ and detecting convergence are described in section 5.3.4.

Starting from the current chain link $\boldsymbol{\theta}^{(j+1)} \in \Theta$, the following steps lead to the next link, according to the MH algorithm and the *hit-and-run sampler* (step 1; Boneh and Golan 1979; Smith 1980):

1. Set up a candidate \mathbf{C} :
 - (a) sample a direction, viz. a random unit vector $\mathbf{d} \in \mathbb{R}^k$ from an isotropic density over the k -dimensional unit sphere;
 - (b) sample a (signed) distance r from a uniform distribution over the interval $\{r' \in \mathbb{R} : \boldsymbol{\theta}^{(j)} + r'\mathbf{d} \in \Theta\}$;
 - (c) set candidate $\mathbf{C} \equiv \boldsymbol{\theta}^{(j)} + r\mathbf{d}$;
2. calculate the acceptance probability

$$\alpha(\boldsymbol{\theta}^{(j)}, \mathbf{C}) \equiv \min\left(1, \frac{\mathcal{P}(\mathbf{C})}{\mathcal{P}(\boldsymbol{\theta}^{(j)})}\right); \quad (5.13)$$

3. draw a random number β from a uniform distribution over the interval $[0, 1]$;
4. if $\beta \leq \alpha$, accept the candidate, i.e. set the next link $\boldsymbol{\theta}^{(j+1)} \equiv \mathbf{C}$; otherwise, $\boldsymbol{\theta}^{(j+1)} \equiv \boldsymbol{\theta}^{(j)}$.

The hit-and-run sampler, compared to alternatives like the Gibbs sampler (Geman and Geman 1984), favours exploring of the whole parameter space Θ without becoming “trapped” in the vicinity of a local posterior maximum (Gilks et al. 1996).

Because only the ratio of two posterior values is used (step 2), the normalising evidence $p(D|\mathcal{M}, \mathcal{K})$ – a constant that is difficult to determine, as mentioned above – is irrelevant in the MH algorithm.

Marginalisation and density estimation. Obviously, the posterior mode alone reveals only one particular aspect of the posterior. However, as a density over $k > 2$ dimensions, the posterior cannot be displayed unambiguously in a figure.

To obtain a plottable summary of the posterior, a set of *marginal posteriors* $\mathcal{P}_i(\cdot)$, i.e. probability densities over each of the parameters θ_i , and *joint marginal posteriors* $\mathcal{P}_{i,j}(\cdot, \cdot)$ over two parameters, can be estimated. Theoretically, these densities are derived from the posterior density by *marginalisation*, viz. integration over all other parameters,

$$\mathcal{P}_i(\theta_i) \equiv p(\theta_i|D, \mathcal{M}, \mathcal{K}) = \int \mathcal{P}(\boldsymbol{\theta}) d\boldsymbol{\theta}_{\setminus i} \quad (5.14)$$

$$\mathcal{P}_{i,j}(\theta_i, \theta_j) \equiv p(\theta_i, \theta_j|D, \mathcal{M}, \mathcal{K}) = \int \mathcal{P}(\boldsymbol{\theta}) d\boldsymbol{\theta}_{\setminus i,j}, \quad (5.15)$$

where $d\boldsymbol{\theta}_{\setminus i} \equiv \prod_{k \neq i} d\theta_k$ and $d\boldsymbol{\theta}_{\setminus i,j} \equiv \prod_{k \neq i,j} d\theta_k$. Practically, marginal posteriors are estimated by only considering component i of the collected samples $\{\boldsymbol{\theta}^{(j)}\}$ and performing

a *density estimation* based on them. Joint marginal posteriors are derived analogously, based on components i and j .

Several *density estimators* exist for deriving a density from a set of samples. One of them – the oldest and probably most popular type, known as the histogram – has several drawbacks: its shape depends on the choice of origin and bin width, and when used with two-dimensional data, a contour diagram cannot easily be derived from it. Generalising the histogram to *kernel density estimation* over one or two dimensions, the samples can be represented more accurately and unequivocally (Silverman 1986).

Below, we refer only to the simpler one-dimensional case. There, the kernel estimator can be written as

$$\mathcal{F}(x) \equiv \frac{1}{N\sigma_{\text{ker}}} \sum_{j=1}^N K\left(\frac{x - X^{(j)}}{\sigma_{\text{ker}}}\right), \quad (5.16)$$

where x is a scalar variable, $K(\cdot)$ the *kernel*, σ_{ker} the *window width* and $X^{(j)}$ are the underlying samples. As detailed by Silverman (1986), the efficiency of various kernels in terms of the achievable mean integrated square error is very similar, and therefore the choice of kernel can be based on other requirements. Since no differentiability is required for the estimated densities and computational effort plays an important practical role, a triangular kernel,

$$K_{\text{tri}}(x) \equiv \max(1 - |x|, 0), \quad (5.17)$$

was selected for estimating the marginal posteriors. The window width is chosen following the recommendations of Silverman (1986),

$$\sigma_{\text{ker}} \equiv 2.189 \cdot \min\left(\sigma_{\text{samp}}, \frac{r_{\text{iq}}}{1.34}\right) (N - M)^{\frac{1}{5}}, \quad (5.18)$$

where σ_{samp} , r_{iq} , N and M are the sample standard deviation, the interquartile range of the samples, number of samples and burn-in length, respectively.

Periodogram mode. By default, the window width for marginal posteriors is based on the MCMC sample standard deviation σ_{samp} (eq. (5.18)). If there are multiple maxima, however, this can lead to artificially broad peaks, which can be particularly problematic for the orbital frequency f (see section 5.2.3), which plays an important role in distinguishing different solutions in orbit-related parameter estimation. Therefore, BASE includes a *periodogram mode*, in which the window width of the marginal posterior of f (a Bayesian analogon to the frequentist *periodogram*) is reduced according the following procedure.

1. Initially assume a default window width as given by eq. (5.18);
2. estimate the marginal posterior of f and find its local maximum f_{max} nearest to posterior mode \hat{f} , as well as the local minimum f_{min} nearest to f_{max} ;
3. calculate the marginal-posterior standard deviation σ' over the half-peak between f_{max} and f_{min} ;
4. re-calculate the window width using eq. (5.18), with $\sqrt{2}\sigma'$ replacing σ_{samp} ;
5. repeat step 2, but only consider local minima with ordinates $\text{p}(f_{\text{min}}|D, \mathcal{M}, \mathcal{K}) \leq 0.5 \cdot \text{p}(f_{\text{max}}|D, \mathcal{M}, \mathcal{K})$ in order not to be misled by weak marginal-posterior fluctuations;
6. repeat steps 3 and 4.

Parameter estimation. To obtain a single most probable estimate of the parameters, the posterior density $\mathcal{P}(\cdot)$ can be summarised by the *posterior mode* $\hat{\theta} \in \Theta$, i.e. the point where the posterior assumes its maximum value,

$$\hat{\theta} \equiv \arg \max_{\theta} \mathcal{P}(\theta). \quad (5.19)$$

This point, also known as the *maximum a-posteriori (MAP)* parameter estimate, can be approximated by the MCMC sample with highest posterior density, based on the values of $\mathcal{P}(\theta^{(j)})$ already calculated during sampling. This approximation neglects the finite spacing between samples.

Alternatively, the following scalar summaries can be inferred from the samples or, for the marginal mode, from the marginal posteriors $\mathcal{P}_i(\theta)$:

- *mean* or *expectation* $\bar{\theta}$,

$$\bar{\theta} \equiv \int_{-\infty}^{\infty} \theta \mathcal{P}_i(\theta) \, d\theta, \quad (5.20)$$

- *median* $\tilde{\theta}$,

$$\int_{-\infty}^{\tilde{\theta}} \mathcal{P}_i(\theta) \, d\theta \equiv 0.5, \quad (5.21)$$

- *marginal mode* $\check{\theta}$,

$$\check{\theta} \equiv \arg \max_{\theta} \mathcal{P}_i(\theta). \quad (5.22)$$

Uncertainty estimation. For uncertainty estimation, *highest posterior-density intervals (HPDIs)* can be derived from the posterior samples. For any given $C \in \mathbb{R}$ with $0 < C < 1$, a HPDI $I_{\text{HPD}} \equiv [a, b]$ is defined as the smallest interval over which the posterior contains a probability C ,

$$\int_a^b \mathcal{P}_i(\theta) \, d\theta = C, \quad \text{s.t.} \quad b - a = \min. \quad (5.23)$$

BASE automatically calculates HPDIs of probability contents 50%, 68.27%, 95%, 95.45%, 99%, and 99.73%; others may be added on user request.

In contrast to frequentist confidence intervals, HPDIs are generally not symmetric, meaning that their midpoint does not correspond to the best estimate. This is because the marginal posteriors may be asymmetric, including any amount of skew.

It should also be noted that HPDIs are not useful with multimodal posteriors because several modes cannot be meaningfully summarised by one interval per dimension, nor by a single best estimate.

To quantify linear dependencies between parameters, the a-posteriori Pearson correlation coefficient,

$$r_{\theta_1, \theta_2} \equiv \frac{\text{cov}(\theta_1, \theta_2)}{\sqrt{\text{var}(\theta_1) \text{var}(\theta_2)}} = r_{\theta_2, \theta_1}, \quad (5.24)$$

can be inferred from the samples. There may also be nonlinear correlations between parameters that are not described by the correlation coefficients. One should also be aware that for strong linear or non-linear relationships between parameters, uncertainties of single parameters as characterised by HPDIs may not be meaningful.

We stress that the (joint) marginal posteriors can – and should – always be referred to, especially when best estimates and/or HPDIs do not adequately characterise the posterior. The availability of these more informative densities is one of the advantages of a Bayesian approach with posterior sampling.

5.2.3 Observable models

Independent of the chosen approach to inference – frequentist or Bayesian –, theoretical values of the observables need to be calculated and compared to the data by means of the likelihood. To this end, an observable model is set up for each relevant type of data, i.e. a function $f(\boldsymbol{\theta}; t)$ of the model parameters $\boldsymbol{\theta}$ and time t . An overview of all model parameters used in this work is given in Table 5.1, while table 5.2 lists quantities that can be derived from them.

In this section, we only sketch the derivation of the observable models, beginning with a single-planet system. For an in-depth treatment of celestial mechanics, the interested reader is referred e.g. to Moulton (1984).

Stellar motion in the orbital plane

Newton’s Law of Gravity governs the motion of a non-relativistic two-body system of star and planet, whose *centre of mass (CM)* rests in some *inertial* reference frame. A solution to it is given by both the star and the planet moving in elliptical *Keplerian orbits* with a fixed common *orbital plane* and each with one *focus* coinciding with the CM.

To describe the stellar position, whose variation is observable with astrometry and Doppler spectroscopy, we set up a coordinate system S_1 whose origin is identical to the CM, z -axis perpendicular to the orbital plane and the vector from the CM to the *periapsis* orientated in positive x -direction. In S_1 , the stellar barycentric position is given by

$$\mathbf{r}_1 = a_\star \begin{pmatrix} \cos E - e \\ \sqrt{1 - e^2} \sin E \\ 0 \end{pmatrix} = \mathbf{r}_1(E; a_\star, e), \quad (5.25)$$

where a_\star is the *semi-major axis*, e the *eccentricity* and E the *eccentric anomaly*.

The time-dependent eccentric anomaly is determined implicitly by *Kepler’s equation*,

$$E - e \sin E = 2\pi(\chi + f(t - t_1)) = M(t), \quad (5.26)$$

where $f = P^{-1}$ is the *orbital frequency*, P the *orbital period*, t_1 the time of first measurement and $M(\cdot)$ the *mean anomaly*, which varies uniformly over the course of an orbit. Furthermore, following Gregory (2005a), we use

$$\chi \equiv \frac{M(t_1)}{2\pi} = f(t_1 - T), \quad (5.27)$$

with T standing for the last time the periapsis was passed prior to t_1 (*time of periapsis*). Kepler’s equation is transcendental and needs to be solved numerically to obtain E for every relevant combination of e and M .

BASE performs a one-time pre-calculation of E over an (e, M) -grid, which, because of the monotonicity of E as an (implicit) function of e and M , allows one to reduce the effort of numerically solving eq. (5.26) by providing lower and upper bounds on E .

By reference to eq. (5.25) and (5.26), it is readily shown that the stellar coordinates are periodic functions of χ with period 1. We therefore call χ a *cyclic* parameter and treat it as lying in the range $[0, 1)$.

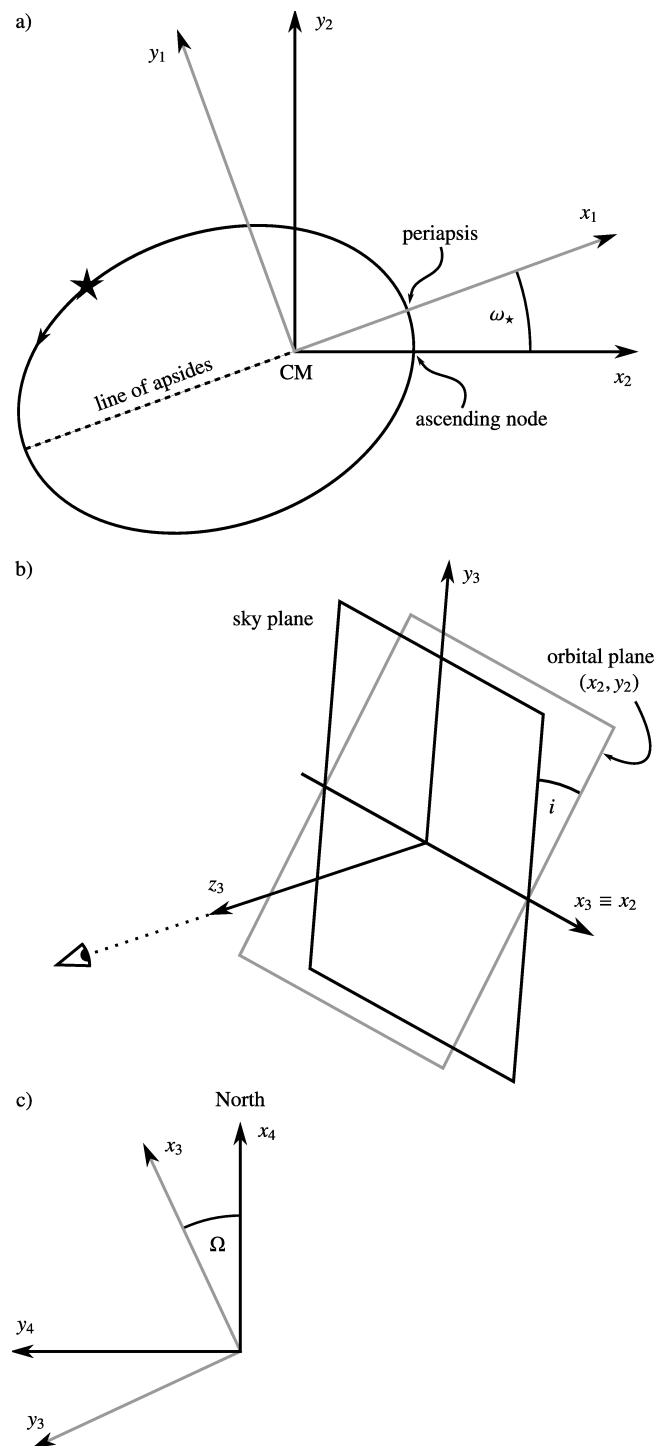


Figure 5.1: Definition of the angles ω_* , i , Ω . a) From S_1 to S_2 , the star and its sense of rotation about the CM are indicated; the dotted line marks the major axis of the orbital ellipse. b) From S_2 to S_3 , the observer and line of sight are indicated. c) From S_3 to S_4 , the positive x_4 -axis points northward along the meridian of the CM.

Transformation into the reference system

To derive the stellar barycentric position as seen from the perspective of an observer, we transform S_1 into a new coordinate system S_4 by three successive rotations. These are described by three Euler angles, termed in our case *argument of the periapsis* ω_* , *inclination* i and *position angle of the ascending node* Ω , and are carried out as follows (fig. 5.1):

1. Rotate S_1 about its z_1 -axis by $(-\omega_*)$ such that the *ascending node*⁴ of the stellar orbit lies on the positive x_2 -axis.
2. Rotate S_2 about its x_2 -axis by $(+i)$ such that the new z_3 -axis passes through the observer.
3. Rotate S_3 about its z_3 -axis by $(-\Omega)$ such that the new x_4 -axis is parallel to the *meridian* of the CM and points in a northern direction.

Thus, the stellar barycentric position has new coordinates

$$\mathbf{r}_4 = \mathbf{R}_{zzz} \mathbf{r}_1, \quad (5.28)$$

with matrix

$$\mathbf{R}_{zzz} \equiv \begin{pmatrix} A & F & J \\ B & G & K \\ C & H & L \end{pmatrix} \quad (5.29)$$

defining the rotations; its components are

$$A = \cos \Omega \cos \omega_* - \sin \Omega \cos i \sin \omega_* \quad (5.30)$$

$$B = \sin \Omega \cos \omega_* + \cos \Omega \cos i \sin \omega_* \quad (5.31)$$

$$F = -\cos \Omega \sin \omega_* - \sin \Omega \cos i \cos \omega_* \quad (5.32)$$

$$G = -\sin \Omega \sin \omega_* + \cos \Omega \cos i \cos \omega_* \quad (5.33)$$

$$J = -\sin \Omega \sin i \quad (5.34)$$

$$K = \cos \Omega \sin i \quad (5.35)$$

$$C = -\sin i \sin \omega_* \quad (5.36)$$

$$H = -\sin i \cos \omega_* \quad (5.37)$$

$$L = \cos i. \quad (5.38)$$

$A, B, F,$ and G are known as the *Thiele-Innes constants*, first introduced by Thiele (1883).

By taking the time derivative of eq. (5.28), we obtain the stellar velocity in S_4 ,

$$\mathbf{v}_4 = \mathbf{R}_{zzz} \frac{d\mathbf{r}_1}{dE} \dot{E} = \frac{2\pi f a_*}{1 - e \cos E} \mathbf{R}_{zzz} \begin{pmatrix} -\sin E \\ \sqrt{1 - e^2} \cos E \\ 0 \end{pmatrix}. \quad (5.39)$$

⁴The *ascending node* is the point of intersection of the orbit and the sky plane where the moving object passes away from the observer. In step 2, a positive rotation angle is used to ensure that the node is indeed ascending (fig. 5.1 b).

Relation to the planetary orbit

By reference to the above results, the observables of AM and RV are easily derived (section 5.2.3). Based on the following simple relation, they can be parameterised by quantities pertaining to the planetary instead of the stellar orbit.

According to the definition of the CM, the line connecting star and planet contains the CM and the ratio of their respective distances from the CM equals the inverse mass ratio,

$$\overrightarrow{CS} = -\frac{m_P}{m_\star} \overrightarrow{CP}, \quad (5.40)$$

where C, S and P stand for CM, star and planet, respectively. This implies a simple relation between the orbits of the star and the planet as follows. The two bodies orbit the CM with a common orbital frequency f and time of periapsis T . With regard to the corresponding periapsis, they always have the same eccentric anomaly E . Their orbital shapes, viz. eccentricities e , are identical as well.

Furthermore, the two bodies share the same sense of orbital revolution, hence those nodes of both orbits which lie on the positive x_2 -axis are ascending. Consequently, the only Euler angle that differs between stellar and planetary orbit is the argument of periapsis, which differs by π because star and planet are in opposite directions from the CM.

Observables

To express the stellar barycentric position as a two-dimensional angular position, we performed a final transformation of S_4 into a spherical coordinate system S_5 with radial, elevation and azimuthal coordinates (r, δ, α) . Its origin is identical with the observer, its reference plane coincides with the (y_4, z_4) -plane and its fixed direction is $-z_4$. In S_5 , the radial coordinate of the CM equals a distance $d = 1 \text{ AU } \varpi^{-1}$, with ϖ being the *parallax*. Stellar coordinates \mathbf{r}_4 therefore correspond to a two-dimensional angular position

$$\mathbf{r}_5 \equiv \begin{pmatrix} \delta \\ \alpha \end{pmatrix} = \frac{1}{d} \begin{pmatrix} x_4 \\ y_4 \end{pmatrix} \quad (5.41)$$

in S_5 , where δ and α are called the *declination* and *right ascension*, respectively. Using eq. (5.25), (5.28) and (5.29), the model function for the angular position of the star with respect to the CM becomes

$$\mathbf{r}(\boldsymbol{\theta}; t) \equiv \mathbf{r}_5 = a' \left((\cos E - e) \begin{pmatrix} A \\ B \end{pmatrix} + \sqrt{1 - e^2} \sin E \begin{pmatrix} F \\ G \end{pmatrix} \right), \quad (5.42)$$

with $a' \equiv \varpi a_\star (1 \text{ AU})^{-1}$. In practice, complications arise because the stellar position is often measured relative to a physically unattached reference star, whose distance and motion differ, and not relative to the unobservable CM. By contrast, for a visual binary, the companion can be used as a reference, yielding the simple model described below.

Other astrometric effects may be caused by the accelerated motion of Earth-bound observers around the *solar system barycentre* (*SSB*). This is discussed for the case of the binary Mizar A in section 5.2.3.

In contrast to astrometry, RV data are usually automatically transformed into an inertial frame resting with respect to the SSB (e.g. Lindegren and Dravins 2003), which allows the Earth's motion to be neglected in this model and treats the observer's rest

frame as being inertial. The model function for the stellar radial velocity measured by an observer is thus given by $(\mathbf{v}_5)_r$, with

$$(\mathbf{v}_5)_r - V = -(\mathbf{v}_4)_z = K \frac{\sin E \sin \omega - \sqrt{1 - e^2} \cos E \cos \omega}{1 - e \cos E}, \quad (5.43)$$

where V is the RV offset, consisting of the radial velocity of the CM plus an offset due to the specific calibration of the instrument – which therefore differ, in general, between RV data sets – and K is the *RV semi-amplitude*, which can be expressed as

$$K = 2\pi f a_\star \sin i = \sqrt[3]{2\pi G f} m_p (m_p + m_\star)^{-\frac{2}{3}} \sin i. \quad (5.44)$$

The last equality holds because of *Kepler's third law*,

$$(a_p + a_\star)^3 f^2 = \frac{G}{4\pi^2} (m_p + m_\star) \quad (5.45)$$

and the definition of the CM (eq. (5.40)). Owing to eq. (5.44), only one of a_\star , K needs to be employed in the AM and RV models; for BASE, K was adopted.

As an aside, in the literature (e.g. Gregory 2005a) often a different version of eq. (5.43) is employed that involves ν instead of E and the alternative definition $K_{\text{alt}} \equiv \frac{K}{\sqrt{1 - e^2}}$ (see table 5.2).

Binary system. If the *primary* and *secondary* binary components assume the roles of star and planet, respectively, the above reasoning also yields the observables of a binary system.

For visual binaries, AM measurements often refer to the position of the secondary with respect to the primary. From the definition of the CM, it follows that the orbit of the secondary with respect to the primary is identical with its barycentric orbit but scaled by a factor $(m_1 + m_2)m_1^{-1}$, or with the semi-major axis equaling the sum of the two components' barycentric semi-major axes,

$$a_{\text{rel}} = a_1 + a_2. \quad (5.46)$$

Thus, the AM model of eq. (5.42) can be used for a binary with a_{rel} replacing a_\star , $\omega_2 + \pi$ replacing ω_\star and

$$a'_{\text{rel}} \equiv \frac{\varpi a_{\text{rel}}}{1 \text{ AU}}. \quad (5.47)$$

Equation (5.43) yields the RV of component i if K is replaced by $(-1)^{i+1}K_i$ and ω by ω_2 . If AM and RV data are combined, BASE uses $K_{1,2}$ instead of $a_{1,2}$ and calculates a_{rel} using the equivalent of eq. (5.44) in combination with eq. (5.46).

Effects of the motion of observer and CM

If we consider the observer, whose position relative to the CM defines the orientation of $S_{2\dots 5}$, to be located on Earth and therefore to be subject to Earth's accelerated motion around the SSB, the angles Ω, i, ω as defined above are not constant but rather functions of time; an additional source of their variation is the (assumedly linear) *proper motion* of the CM. Consequently, these angles are not appropriate constant model parameters even within a timespan of a few months. Furthermore, the systems $S_{2\dots 5}$ defined above are not strictly inertial, which renders the simple coordinate transformations invalid. However, we argue below that the variation in these angles is so weak that these effects are indeed negligible in the context of this work.

Because AM determines instantaneous positions, it is directly influenced by changes in the positions of the observer and the CM. In the following, we assess for Mizar A the greatest possible magnitude of a change in the AM position $\Delta r \equiv |\Delta \mathbf{r}|$ due to angular changes $\Delta\Omega, \Delta i, \Delta\omega > 0$, which are in turn caused by the varying relative position of observer and CM. Finally, we compare Δr to the AM measurement uncertainties.

An upper limit on each of the angular changes $\Delta\Omega, \Delta i, \Delta\omega$ can be derived from the proper motion of the CM and the *annual parallax* from the Earth's motion, as

$$|\Delta\Omega|, |\Delta i|, |\Delta\omega| \leq 2\varpi + \boldsymbol{\mu} \Delta t, \quad (5.48)$$

where the first term corresponds to the annual parallax and $\boldsymbol{\mu}$ is the magnitude of the CM's proper motion. For the second term, we have $\boldsymbol{\mu}\Delta t \approx \sqrt{\mu_{\alpha^*}^2 + \mu_{\delta}^2} \Delta t$.

Using $\sin(x + \xi) \approx \sin x + \xi \cos x$, for $\xi \ll 1$, along with eq. (5.30) – (5.35) and (5.48) as well as the final results for Ω, i, ω and ϖ (using the values $\hat{\theta}$ from table 5.9) and the proper motion from table 5.3 yields the following maximum changes of the Thiele-Innes constants:

$$|\Delta A| \lesssim 8.30 \times 10^{-6} \quad (5.49)$$

$$|\Delta B| \lesssim 8.00 \times 10^{-6} \quad (5.50)$$

$$|\Delta F| \lesssim 7.99 \times 10^{-6} \quad (5.51)$$

$$|\Delta G| \lesssim 4.34 \times 10^{-6}. \quad (5.52)$$

Hence, with $\Delta r \approx \sqrt{|\Delta\delta|^2 + |\Delta(\alpha \cos \delta)|^2}$,

$$|\Delta\delta| \leq a'_{\text{rel}} (2|\Delta A| + |\Delta F|) \quad (5.53)$$

$$|\Delta(\alpha \cos \delta)| \leq a'_{\text{rel}} (2|\Delta B| + |\Delta G|) \quad (5.54)$$

and the final posterior median a'_{rel} (table 5.10), we obtain the maximum change in relative angular position of the two components,

$$\Delta r \lesssim 0.311 \mu\text{as}. \quad (5.55)$$

Comparison with table 5.4 reveals that this is more than two orders of magnitude smaller than the median AM measurement uncertainty, proving that the motions of the Earth and the CM of Mizar A can indeed be neglected.

Table 5.1: Model parameters used by BASE.

Symbol	Designation	Unit	Widest Prior Support	Cyclic ⁶	Prior Type	Data Types ⁵		
						AM	RV	AM+RV
e	eccentricity	1	$[0, 1)$		uniform	×	×	×
f	orbital frequency	d^{-1}	$(0, 1]$		Jeffreys	×	×	×
χ	mean anomaly at t_1 over 2π	1	$[0, 1)$	×	uniform	×	×	×
ω, ω_2	argument of periapsis ⁷	rad	$[0, 2\pi)$	×	uniform	×	×	×
i	inclination	rad	$[0, \pi)$		uniform	×		×
Ω	position angle of the ascending node ⁸	rad	$[0, 2\pi)$	×	uniform	×		×
a'_{rel}	semi-major axis of orbit of secondary around primary over distance	mas	$[10^{-3}, 10^{-5}]$ ⁹		Jeffreys	×		
ϖ	parallax	arcsec	$(0, 0.77]$ ¹⁰		Jeffreys			×
V	RV offset	m s^{-1}	—		uniform		×	×
K, K_i	RV semi-amplitude ¹¹	m s^{-1}	≥ 0		mod. Jeffreys		×	×
τ_+	standard deviation of additional AM noise	mas	≥ 0		mod. Jeffreys	×		×
σ_+	standard deviation of additional RV noise	m s^{-1}	≥ 0		mod. Jeffreys		×	×

⁵The types of data for which each parameter is relevant. Abbreviations: AM (astrometry), RV (radial velocities).

⁶For cyclic parameters θ , the indicated lower and upper bounds are treated as equivalent by BASE (see section 5.2.3).

⁷Normal mode employs ω , while binary mode employs ω_2 .

⁸The widest prior range of Ω reduces to $[0, \pi)$ if no RV data are provided, in which case it cannot be determined whether a given node is ascending or descending; then, Ω is defined to be the position angle of the first node.

⁹Lower bound corresponds to AM measurement uncertainty of $1 \mu\text{as}$; upper bound according to wide-binary observations by Tolbert (1964).

¹⁰Interval includes trigonometric parallax of the nearest star, Proxima Centauri (Perryman et al. 1997).

¹¹Normal mode employs K , while binary mode employs K_1 and K_2 .

Table 5.2: Quantities derived from model parameters.

Definition	Designation	Unit	Equations	Mode ¹²		Data Types ¹³		
				N	B	AM	RV	AM+RV
$P \equiv f^{-1}$	period	d		×	×	×	×	×
$T \equiv t_1 - \frac{\chi}{f}$	time of periapsis	d	(5.27)	×	×	×	×	×
$d \equiv \frac{1 \text{ AU}}{\varpi}$	distance	pc			×			×
$\rho \equiv \frac{m_2}{m_1} = \frac{K_1}{K_2}$	binary mass ratio	1	(5.40), (5.44)		×		×	×
$K_{\text{alt}} \equiv \frac{K}{\sqrt{1-e^2}}$	alternative RV semi-amplitude ¹⁴	m s^{-1}		×	×		×	×
$m_j \equiv \frac{4\pi^2}{G} \frac{K_{3-j}(K_1+K_2)^2}{f(2\pi \sin i)^3}$	component mass	M_{\odot}	(5.44), (5.45)		×			×
$a_j \equiv \frac{K_j}{2\pi f \sin i}$	semi-major axis of component's orbit around CM	AU	(5.44)		×			×
$a_{\text{rel}} \equiv a_1 + a_2 = \frac{K_1+K_2}{2\pi f \sin i}$	semi-major axis of secondary's orbit around primary	AU	(5.44)		×			×
$a_{\star} \sin i \equiv \frac{K}{2\pi f}$	semi-major axis of stellar orbit times sine of inclination	AU	(5.44)	×			×	
$m_{\text{p,min}} \equiv K(2\pi G f)^{-\frac{1}{3}}(m_{\text{p,min}} + m_{\star})^{\frac{2}{3}}$	minimum planetary mass ¹⁵	M_{J}	(5.44)	×			×	
$a_{\text{p,max}} \equiv \frac{m_{\star}}{m_{\text{p,min}}} a_{\star} \sin i$	maximum semi-major axis of planetary orbit	AU	(5.40)	×			×	

¹²The modes in which each derived quantity appears. Abbreviations: N (normal), B (binary).

¹³The types of data for which each derived quantity is relevant. Abbreviations: AM (astrometry), RV (radial velocities).

¹⁴Refers to the star or any of the binary components, respectively.

¹⁵The implicit function of minimum planetary mass $m_{\text{p,min}}$ is solved numerically. The planet's minimum mass equals its real mass if the orbit is edge-on, viz. $\sin i = 1$.

5.3 BASE – Bayesian astrometric and spectroscopic exoplanet detection and characterisation tool

We have developed BASE, a computer program for the combined Bayesian analysis of AM and RV data according to section 5.2, for the following main reasons.

- A statistically well-founded, reliable tool was needed that was able to perform a complete Bayesian parameter and uncertainty estimation, along with model selection (only for planetary systems, not detailed in this article).
- We aimed to combine astrometry and Doppler-spectroscopy analyses.
- A possibility to include knowledge from earlier analyses was needed.
- Finding all relevant solutions across a multidimensional, high-volume parameter space Θ was required. A more detailed knowledge of the parameters than a best estimate and a confidence interval can provide is especially important when the data do not constrain the parameters well, e.g. when only few data have been recorded or the *signal-to-noise* ratio is low (as can be the case for lightweight planets or young host stars).

BASE is a highly configurable command-line tool developed in Fortran 2008 and compiled with GFortran (Free Software Foundation 2011a). Options can be used to control the program’s behaviour and supply information such as the stellar mass or prior knowledge (see section 5.3.1). Any option can be supplied in a configuration file and/or on the command line.

5.3.1 Prior knowledge

Any model parameter can be assigned a prior probability density in one of three forms:

- a fixed value;
- a user-defined shape, i.e. a set $\{(\theta_i, p(\theta_i|\mathcal{M}, \mathcal{K}))\}$;
- a range (either in combination with a specific shape, which is then clipped to the given range, or else using one of the standard prior shapes detailed in section 5.7).

Any parameter for which no such option is present is assigned a default prior shape and range to pose the weakest possible restraints justified by the data, the model, and mathematical/physical considerations.

5.3.2 Physical systems and modes of operation

As mentioned above, BASE is capable of analysing data for two similar types of systems, whose physics have been described in section 5.2.3. These are systems with

1. **one** observable component that may be accompanied by one or more unobserved bodies (generally referred to as planetary systems here for simplicity); in this *normal mode*, the number of companions can be set to a value ranging from zero to nine, or a list of such values, in which case several runs are conducted and the outcome is compared in terms of model selection;
2. **two** observable, gravitationally bound components (referred to as binary stars); this mode of operation is called *binary mode*.

5.3.3 Types of data

Observational data of the following types can be treated by BASE:

- AM data, whose observable, in the case of binary targets, is the relative angular position of the two binary components. Each data record consists of a date, the angular position $(\alpha \cos \delta, \delta)$ or (ρ, θ) in a cartesian or polar coordinate system, and its standard uncertainty ellipse, given by (a, b, ϕ) ;
- RV data, where each data record consists of a date and the observed stellar radial velocity as well as its uncertainty.

5.3.4 Computational techniques

In the following, we describe some computational techniques implemented in BASE that are relevant to the present work.

Improved exploration of parameter space. To enhance the *mixing*, i.e. rapidness of exploration of the parameter space, of Markov chains produced by the Metropolis-Hastings (MH) algorithm (section 5.2.2) and decrease their attraction by local posterior modes, the *parallel tempering* (PT) algorithm (e.g. Gregory 2005b) is employed one level above MH. Its function is to create in parallel n chains of length N , $\{\boldsymbol{\theta}_{(k)}^{(j)} : j = 1, \dots, N; k = 1, \dots, n\}$ each sampled by an independent MH procedure, where the *cold* chain $k = 1$ uses an unmodified likelihood $\mathcal{L}(\cdot)$, while the others, the *heated* chains, use as replacement $\mathcal{L}(\cdot)^{\gamma_{(k)}}$ with the positive *tempering parameter* $\gamma_{(k)} < 1$. After n_{swap} samples of each chain, two chains $k \geq 1$ and $k + 1 \leq n$ are randomly selected and their last links, denoted here by $\boldsymbol{\theta}_{(k)}$ and $\boldsymbol{\theta}_{(k+1)}$, are swapped with probability

$$\alpha_{\text{swap}}(k, k + 1) = \min \left(1, \left(\frac{\mathcal{L}(\boldsymbol{\theta}_{(k+1)})}{\mathcal{L}(\boldsymbol{\theta}_{(k)})} \right)^{\gamma_k} \left(\frac{\mathcal{L}(\boldsymbol{\theta}_{(k)})}{\mathcal{L}(\boldsymbol{\theta}_{(k+1)})} \right)^{\gamma_{k+1}} \right), \quad (5.56)$$

which ensures that the distributions of both chains remain unchanged.

This procedure allows states from the “hotter” chains, which explore parameter space more freely, to “seep through” to the cold chain without compromising its distribution. Conclusions are only drawn from the samples of the cold chain.

In contrast to MCMC sampling, which is sequential in nature, the structure of PT allows one to exploit the multiprocessing facilities provided by many modern computing architectures. For this purpose, BASE uses the OpenMP API (OpenMP Architecture Review Board 2008) as implemented for GFortran by the GOMP project (Free Software Foundation 2011b).

Assessing convergence. As a matter of principle, it cannot be proven that a given Markov chain has converged to the posterior. However, convergence may be meaningfully defined as the degree to which the chain does not depend on its initial state $\boldsymbol{\theta}^{(0)}$ any more. This can be determined on the basis of a set of independent chains – in our case, the cold chains of m independent PT procedures – started at different points in parameter space. These starting states should be defined such that for each of their components, they are *overdispersed* with respect to the corresponding marginal posteriors.

Since the marginal posteriors are difficult to obtain before the actual sampling, BASE determines the starting states by repeatedly drawing, for each parameter, a set of m

samples from the prior using *rejection sampling*; the repetition is stopped as soon as the sample variance exceeds the corresponding prior variance, which yields a set of starting states overdispersed with respect to the prior. Assuming that the prior variance exceeds the marginal-posterior variance, the overdispersion requirement is met.

Such a test, using the *potential scale reduction (PSR)* or Gelman-Rubin statistic, was proposed by Gelman and Rubin (1992) and was later refined and corrected by Brooks and Gelman (1998). It is repeatedly carried out during sampling and, in the case of a positive result, sampling is stopped before the user-defined maximum runtime and/or number of samples have been reached. It may also sometimes be useful to abort sampling manually, which can be done at any time.

The statistic is calculated with respect to each parameter θ separately, using the post-burn-in samples¹⁶ $\{\theta_l^{(j)} : j = M + 1, \dots, N; l = 1, \dots, m\}$ of θ provided by the m independent chains. It compares the actual variances of θ within the chains built up thus far to an estimate of the marginal-posterior (i.e. target) variance of θ , thus estimating how closely the chains have approached convergence.

The PSR $\hat{\mathcal{R}}^{1/2}$ is defined as

$$\hat{\mathcal{R}} \equiv \frac{d + 3}{d + 1} \frac{\hat{V}}{W}, \quad (5.57)$$

where \hat{V} is an estimate of the marginal-posterior variance, d the estimated number of degrees of freedom underlying the calculation of \hat{V} , and W the mean within-sequence variance. The quantities are defined as

$$\hat{V} \equiv \frac{\nu - 1}{\nu} W + \frac{m + 1}{m\nu} B \quad (5.58)$$

$$W \equiv \frac{1}{m(\nu - 1)} \sum_{l=1}^m \sum_{j=M+1}^N (\theta_l^{(j)} - \bar{\theta}_l)^2 \quad (5.59)$$

$$B \equiv \frac{\nu}{m - 1} \sum_{l=1}^m (\bar{\theta}_l - \bar{\theta})^2, \quad (5.60)$$

where B is the between-sequence variance and $\nu \equiv N - M$; a horizontal bar denotes the mean taken over the set of samples obtained by varying the omitted indexes.

Owing to the initial overdispersion of starting states, \hat{V} overestimates the marginal-posterior variance in the beginning and subsequently decreases. Furthermore, while the chains are still exploring new areas of parameter space, W underestimates the marginal-posterior variance and increases. Thus, as convergence to the posterior is accomplished, $\hat{\mathcal{R}}^{1/2} \searrow 1$. Therefore, we can assume convergence as soon as the PSR has fallen below a threshold $\mathcal{R}_*^{1/2} \geq 1$. If BASE is run with $\mathcal{R}_* \equiv 1$, sampling is continued up to the user-defined length or duration, respectively.

For cyclic parameters (section 5.2.3), the default lower and upper prior bounds are equivalent, which needs to be taken into account when calculating the PSR. Thus, BASE uses the modified definition of the PSR by Ford (2006) for these parameters.

5.4 Target and data

¹⁶Because the burn-in length M cannot be determined in advance, BASE considers a fixed but configurable fraction of samples to belong to the burn-in phase at any time.

¹⁷Different values of the parallax have been estimated in this work (table 5.9).

¹⁸Uncertainties are missing in the original publication, but have been estimated in this work (sec-

Table 5.3: Basic physical properties of Mizar A.

Property	Value	Reference
Type	SB II	1
Spectral types	$2 \times A2 V$	2
M_V	2.27 ± 0.07 mag	2
M_{bol}	0.91 ± 0.07 mag	3
L	$33.3 \pm 2.1 L_{\odot}$	3
R	$2.4 \pm 0.1 R_{\odot}$	3
ϖ^{17}	39.4 ± 0.3 mas	3
μ_{α^*}	119.01 ± 1.49 mas yr $^{-1}$	4
μ_{δ}	-25.97 ± 1.65 mas yr $^{-1}$	4

References. (1) Pickering (1890); (2) Hoffleit and Jaschek (1982); (3) Hummel et al. (1998); (4) van Leeuwen (2007)

Mizar A (ζ^1 Ursae Majoris, HD 116656, HR 5054), the first spectroscopic binary discovered, is of double-lined type (SB II; Pickering 1890). Its basic physical properties are summarised in table 5.3. Together with the spectroscopic binary Mizar B, it forms the Mizar quadruple system, seen from Earth as a visual binary with the components separated by about $14.4''$. Mizar is the first double star discovered by a telescope and also the first one to be imaged photographically (Bond 1857).

At an apparent angular separation of about $11.8'$, or 74 ± 39 kAU spatial distance, Mizar is accompanied by Alcor, which has recently turned out to be a spectroscopic binary itself (Mamajek et al. 2010). Mizar and Alcor, also known as the ‘‘Horse and Rider’’, form an easy naked-eye double star, while it is still a matter of debate whether or not they constitute a physically bound sextuplet.

Published data used in this article are displayed in fig. 5.7 – 5.8 along with the model functions determined by BASE and their properties are summarised in table 5.4. Using a Coudé spectrograph at the 1.93-m telescope at Observatoire Haute Provence, Prevot (1961) obtained 17 optical photographic spectra of the light of Mizar A combined with

tion 5.5.1).

¹⁹Uncertainties refer to the semi-minor and -major axes of the standard uncertainty ellipses, respectively (section 5.2.1).

Table 5.4: Published data for Mizar A used in this work.

Data type	Instrument	Observatory	Year	No. records	Median uncertainty	Ref.
Radial velocities ¹⁸	Coudé spectrograph	Haute Provence	1961	2×17	2.050 km s^{-1}	1
Angular positions ¹⁹	Mark III interferometer	Mount Wilson	1995	28	$0.040 \text{ mas} / 0.345 \text{ mas}$	2
Angular positions ¹⁹	NPOI interferometer	Lowell	1998	25	$0.042 \text{ mas} / 0.137 \text{ mas}$	3

References. (1) Prevot (1961); (2) Hummel et al. (1995); (3) Hummel et al. (1998)

that of electric arcs or sparks between iron electrodes. For each of the 13 individual stellar lines identified in the spectra, one intermediate radial-velocity value per binary component was obtained by comparison with a set of reference lines of iron. Finally, a set of 17 pairs of RVs for the two components was calculated as arithmetic means of the corresponding intermediate values. The RV measurement uncertainty is not given by Prevot (1961) and was therefore estimated in the course of the present work, as described in section 5.5.

High-precision AM data of Mizar A were first obtained by Hummel et al. (1995) using the Mark III optical interferometer on Mount Wilson, California (Shao et al. 1988), with baseline lengths between 3 and 31 m. It measured the squared *visibilities* and their uncertainties at positions sampled over the aperture plane due to Earth’s rotation. The visibilities can be modelled as a function of the diameters, magnitude differences, and relative angular positions (ρ, θ) of the binary components, Armstrong et al. (1992). These authors also describe a procedure to derive one angular position for each night of observation from a corresponding set of visibilities, which was adopted by Hummel et al. (1995) to obtain initial estimates of the orbital parameters. These positional data are also relevant for the present work. Hummel et al. (1995) also performed a direct fit to the squared visibilities to derive final estimates of the component diameters and orbital parameters.

Later, a descendant instrument, the Navy Prototype Optical Interferometer (NPOI) at Lowell observatory, Arizona (Armstrong et al. 1998), was used by Hummel et al. (1998) to obtain more accurate results using three siderostats, viz. three baselines at a time. This allowed a better calibration using the *closure phase*,²⁰ which is independent of atmospheric turbulence. Similarly as before, Hummel et al. (1998) separately fitted binary orbits directly to the visibility data and also to the positional angles derived for each night, concluding that the respective results agree well with each other and that those parameters in common with spectroscopic analyses are compatible with Prevot (1961); they also performed a fit to both AM and RV data to obtain their final parameter estimates.

While Hummel et al. (1998) did not include the older, less accurate Mark III data in their analysis, we present a combined treatment of all published data, i.e. the AM positions of Hummel et al. (1995) and Hummel et al. (1998) along with the RV data of Prevot (1961).

5.5 Analysis and results

To illustrate the features of BASE and demonstrate its validity, we used the tool to analyse all published data of Mizar A. This section details the steps taken in and the results of our analysis. Its general goal was, given uninformative prior knowledge, to search the

²⁰The *closure phase* ϕ_{cl} is the phase of the product of three visibilities, each pertaining to a different baseline.

Table 5.5: Analysis passes carried out sequentially.

Pass	Description	Data
A	First constraints on RV parameters	RV
B	Combining all data	All
C	Selecting frequency f	All
D	Selecting ω_2 and Ω	All
E	Refining results	All

parameter space as comprehensively as possible to find and characterise the a posteriori most probable solution together with its uncertainty and other characteristics. For reasons detailed below, once the RV data had been prepared, several runs of BASE (table 5.5) were manually conducted, each using priors derived from the previous pass, with relatively uninformative priors in the first step.

In this analysis, our approach was to regard all nominal measurement uncertainties as accurate by setting parameters characterising additional noise in AM (τ_+) as well as RV (σ_+) data to zero.

Astrometric data alone allow one to constrain neither the RV offset V nor the amplitudes K_1, K_2 but only the sum $K_1 + K_2$ (section 5.2.3). By contrast, these parameters can be constrained using spectroscopic data alone, or AM and RV data both. However, the AM data reduce the relative weight of the spectroscopic data and thus make the determination of these parameters harder. We found in the course of this work that by an iterative approach, starting with a first pass using only RV data, this difficulty could be resolved.

In all passes, BASE was configured to build up eight parallel chains, including the cold chain, using the parallel-tempering technique (detailed in section 5.3.4). 10^8 posterior samples were collected, the first 10% of which were assumed to be burn-in samples. In the final pass, two PT procedures were employed to enable a test of convergence to the posterior, which reduced the number of samples collected with unchanged memory requirements to 5×10^7 . Convergence was not assessed in earlier passes, because convergence was difficult to reach as long as several distinct solutions existed within the prior support.

5.5.1 Preparation of RV data

Assuming appropriate measurement uncertainties is a prerequisite for the proper relative weighting of the different data types when combining them. Because these uncertainties are not quoted by Prevot (1961) for the spectroscopic data, we estimated them according to the following method.

First, we assumed that each RV datum v_i is the sum of the model value $v(\boldsymbol{\theta}_{\text{true}}; t_i)$ and an error e_i , where $\boldsymbol{\theta}_{\text{true}}$ are the true parameter values and the errors are independent and identically Gaussian-distributed with an unknown standard deviation σ . Furthermore, we assumed that the model parameters found in a particular analysis are identical with $\boldsymbol{\theta}_{\text{true}}$. It follows that the uncertainty σ can be estimated as the sample standard deviation of the set of residuals $\{v_i - v(\boldsymbol{\theta}_{\text{true}}; t_i)\}$.

Thus, based on the sample standard deviation of the best-fit residuals of Prevot (1961), viz. 2.13 km s^{-1} , we initially assumed a conservative value of 2.50 km s^{-1} for all data. To quantify the measurement uncertainties based on our own inference, we then conducted a preliminary analysis similar to pass A described in the next subsection, from which we finally inferred $\sigma = 2.05 \text{ km s}^{-1}$. In addition, we took a more correct alternative approach to estimating the RV uncertainties by assuming a relatively low value of $\sigma = 2.00 \text{ km s}^{-1}$ and allowing for higher noise via σ_+ , which led to an estimate $\sqrt{\sigma^2 + \sigma_+^2}$ deviating by less than 1.7% from our previous estimate.

Table 5.6: Pass A: initial prior ranges.

e	f^{21} (d^{-1})	χ	ω_2 (rad)	V^{22} (km s^{-1})	K_1^{23} (km s^{-1})	K_2^{23} (km s^{-1})
0	2.7379×10^{-6}	0	0	-65.25	0	0
1	1	1	2π	56.15	69.50	68.85

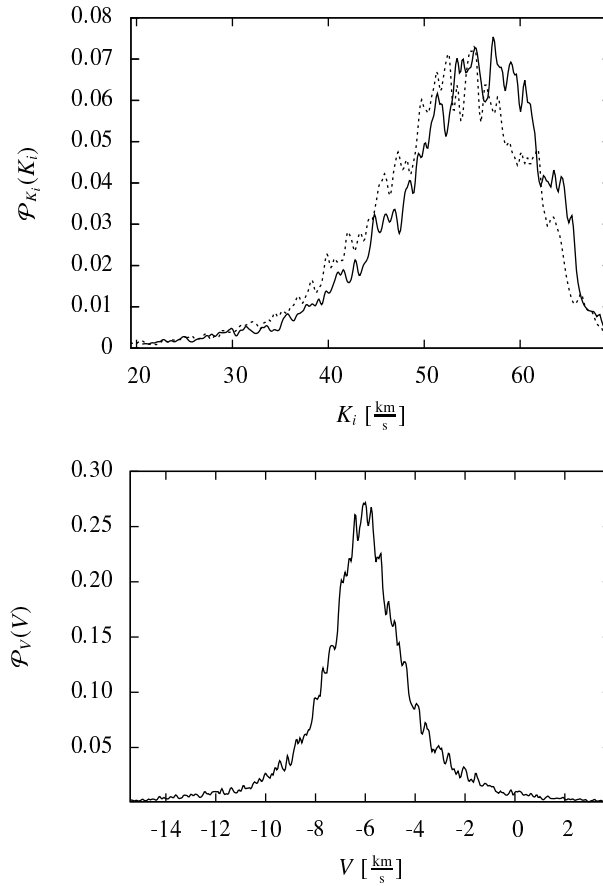


Figure 5.2: Pass A: Marginal posteriors of RV amplitudes K_1 (top, solid line), K_2 (top, dashed line) and offset V (bottom), all plotted over the approximate range of the corresponding 99% HPDI.

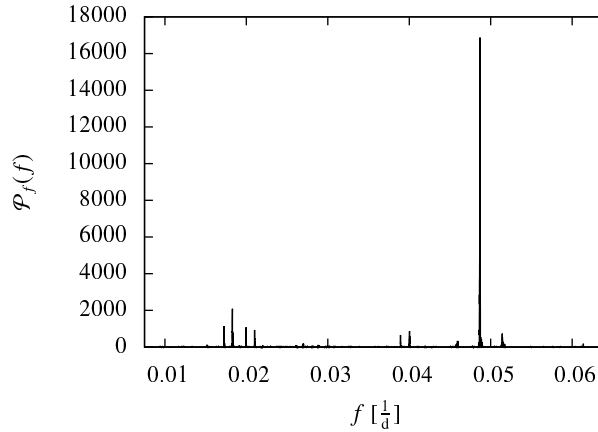


Figure 5.3: Pass B: Marginal posterior of orbital frequency f , plotted over the range of the corresponding 99% HPDI. This is a Bayesian analogon to the frequentist periodogram.

5.5.2 Pass A: first constraints on RV parameters

To facilitate the determination of the RV parameters K_1, K_2 and V , a first pass using only RV data was carried out, as mentioned above. It used uninformative priors (section 5.7 and table 5.1), with bounds listed in table 5.6.

Figure 5.2 shows the resulting marginal posteriors (see section 5.2.2) of the RV offset V and the amplitudes K_1, K_2 , with the abscissae approximately corresponding to the 99% HPDIs. These marginal posteriors represent much tighter constraints on the parameters than the corresponding priors do.

5.5.3 Pass B: combining all data

Providing as new priors the marginal posteriors of all RV parameters $e, f, \chi, \omega_2, V, K_1$ and K_2 from pass A, constrained to the corresponding 99% HPDIs $I_{\text{HPD}, 99\%}$, and again using uninformative priors on the additional parameters (table 5.7), the AM data were added and BASE was run again. Using the resulting posterior samples, BASE was additionally invoked in periodogram mode to refine the kernel window width for the marginal posterior of f as described in section 5.2.2. (This refinement was not used in pass A in order not to constrain the frequency too tightly before adding the AM data, because the combined data are expected to correspond to a marginally differing frequency.)

²¹The prior bounds of f correspond to a period range between 1 d and 1 yr.

²²The bounds of V are given by the section of the RV ranges measured for primary and secondary, with the latter extended by one corresponding measurement uncertainty on both sides.

²³The upper bounds of K_1 and K_2 are given by half the measured RV span of the corresponding component, plus one measurement uncertainty.

²⁴Interval includes trigonometric parallax of nearest star, Proxima Centauri (Perryman et al. 1997).

Table 5.7: Pass B: prior ranges for additional AM parameters.

i	Ω	ϖ ²⁴
(rad)	(rad)	(arcsec)
0	0	0
π	2π	0.77

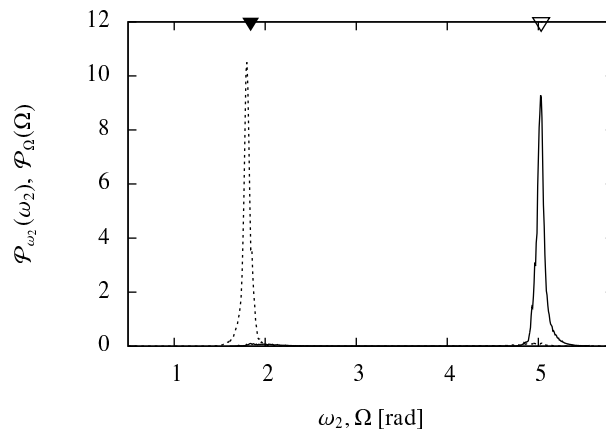


Figure 5.4: Pass C: Marginal posteriors of the argument of periapsis ω_2 (solid line) and position angle of the ascending node Ω (dashed line), plotted over the approximate range of the corresponding 99% HPDIs. Triangles indicate the positions of small local maxima located approximately $\pm\pi$ from the corresponding marginal modes.

The resulting marginal posterior (fig. 5.3) exhibits a very strong mode around 0.04869 d^{-1} , whose height is 8.1 times that of the next lower peak. Over the range of its mode, the marginal posterior contains a total probability of 45.0%. Most other parameters in this stage have very broad and/or multimodal marginal posteriors, hinting at different solutions still probable within the prior support.

5.5.4 Pass C: selecting the frequency f

For the next pass, the prior of f was provided by the marginal posterior of pass B (fig. 5.3), constrained to the range of the marginal mode. For all other parameters, priors were identical with the previous marginal posteriors, constrained to $I_{\text{HPD}, 99\%}$.

With the frequency thus restrained, the pass produced unimodal marginal posteriors in all parameters except for ω_2 and Ω .

The marginal posteriors of the argument of periapsis ω_2 and the position angle of the ascending node Ω exhibited small local maxima located at a distance of -1.02π and $+1.03\pi$ from their marginal modes, respectively (indicated by triangles in fig. 5.4). This can be explained by the fact that the Thiele-Innes constants (eq. (5.30) – (5.33)), which appear in the AM model, contain products of $\sin(\cdot)$ and/or $\cos(\cdot)$ functions with ω_2 and Ω as arguments. These products retain their values when ω_2 and Ω are both shifted by π – with opposite signs, since the marginal modes of these angles lie in different halves of the interval $[0, 2\pi)$. Consequently, the AM model function is invariant with respect to such shifts. (Thus, when analysing only AM data, it cannot be determined which node is ascending, hence Ω is defined only over the interval $[0, \pi)$ and refers to the first node.) Owing to the combination with RV data, which independently constrain ω_2 (eq. (5.43)), the ambiguity is strongly reduced, as illustrated by the very small subpeaks in fig. 5.4 – though it is not completely resolved.

As another result from this pass, the high correlation coefficient of the two angles, $r_{\omega_2, \Omega} = -0.62$, expresses the strong negative linear relationship between them introduced by the possibility of a contrarious change.

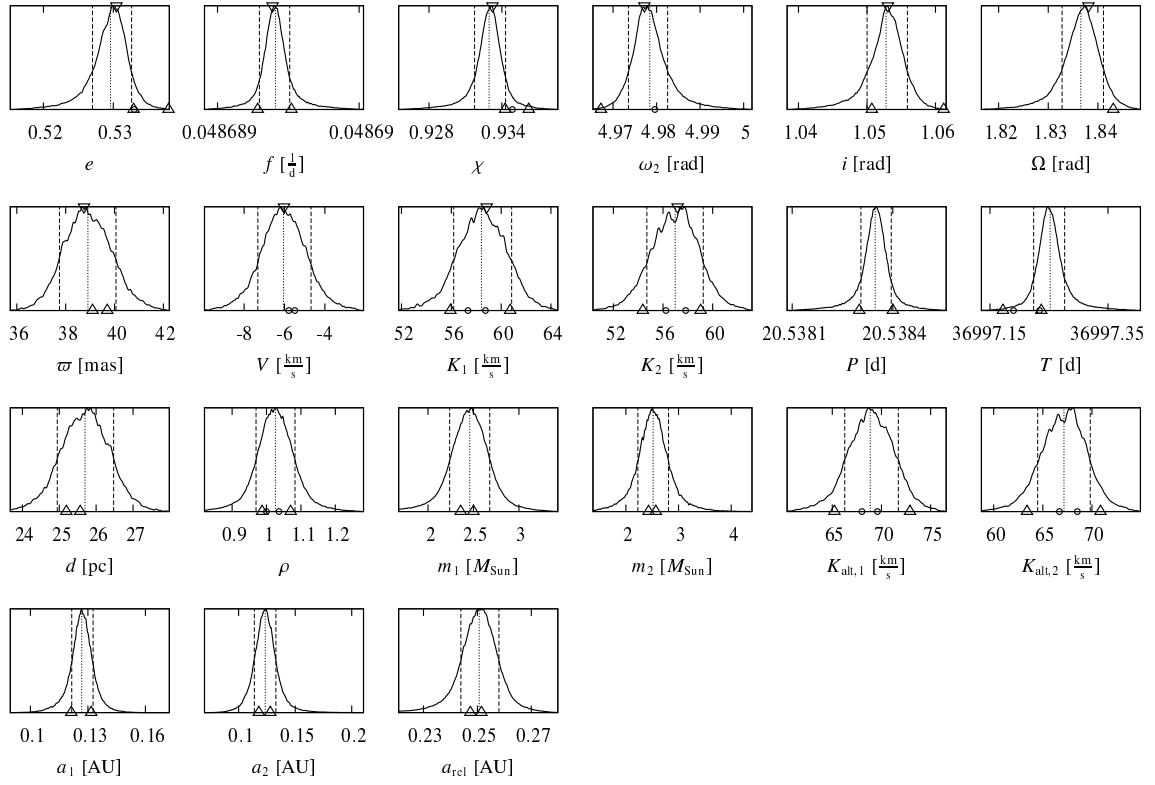


Figure 5.5: Marginal posteriors of parameters e , f , χ , ω_2 , i , Ω , ϖ , V , K_1 , K_2 (see also table 5.1) and derived quantities P , T , d , ρ , $K_{\text{alt},1}$, $K_{\text{alt},2}$, m_1 , m_2 , a_1 , a_2 , a_{rel} (table 5.2) with marginal-posterior medians (dotted line) and 68.27% HPDIs (dashed lines), from the final pass. Abscissae ranges are identical with the corresponding 95% HPDIs. Ordinate values were omitted but follow from normalisation. Open triangles on the upper abscissae indicate the MAP estimates. Open circles on the lower abscissae bound the confidence intervals given by or derived from Prevot (1961), while open triangles refer to the intervals of Hummel et al. (1998). Some of these literature estimates are not plotted because they are outside the abscissa range; for f and χ , no literature uncertainty estimate is available. For derivations and numerical values, see tables 5.9 and 5.10.

5.5.5 Passes D and E: selecting ω_2 and Ω and refining results

In pass D, the ambiguity in ω_2 and Ω was resolved by selecting the range around their marginal modes via the new priors. For all other parameters, priors were again identical to the previous marginal posteriors, constrained to $I_{\text{HPD}, 99\%}$. All resulting marginal posteriors turned out to be unimodal, corresponding to a single solution as opposed to several clearly distinct orbital solutions.

As a final step, pass E was conducted to refine the results by using all previous marginal posteriors, constrained to $I_{\text{HPD}, 99\%}$, as new priors, thus confining the parameter space a priori to the most probable solution.

(Joint) marginal posteriors and correlations. The final marginal posteriors of all model parameters, plotted over the corresponding 95% HPDIs, are shown in fig. 5.5, along with the medians, 68.27% HPDIs (corresponding in probability content to the frequentist 1- σ confidence intervals) and MAP estimates. For comparison, literature estimates and

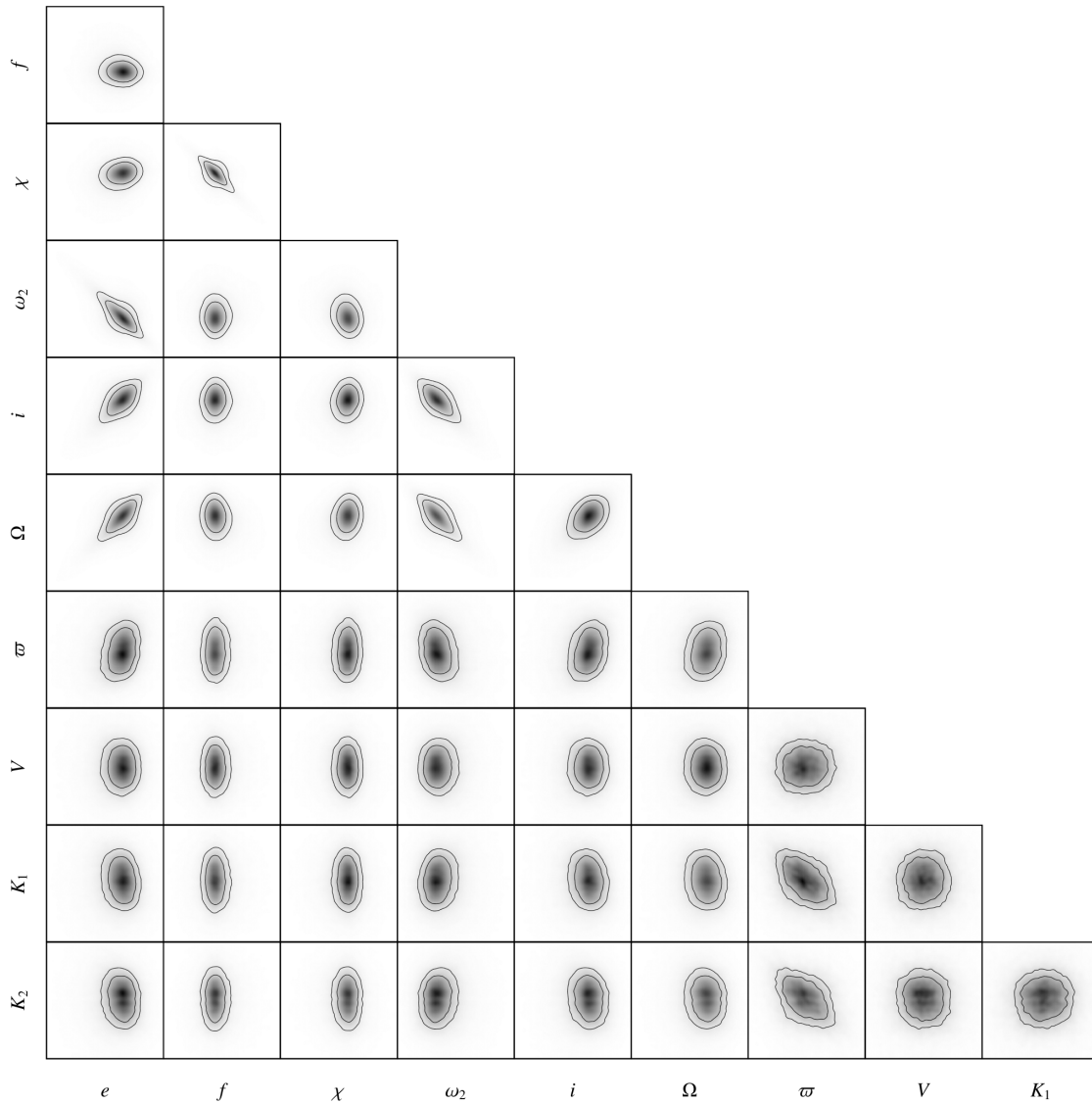


Figure 5.6: Two-parameter joint marginal posterior densities $\mathcal{P}_{i,j}(\cdot, \cdot)$ from pass E. The figure consists of 45 sub-plots, one for each combination of two parameters. Black denotes highest density. The inner and outer contours contain 50% and 64.86% probability, respectively. All plots aligned in one column share the same abscissa, denoted on the bottom; all plots aligned in one row share the same ordinate, denoted on the left. All abscissae and ordinates are displayed over the corresponding 95% HPDIs (table 5.9), such that the total probability content of each plot is between 90% and 95%.

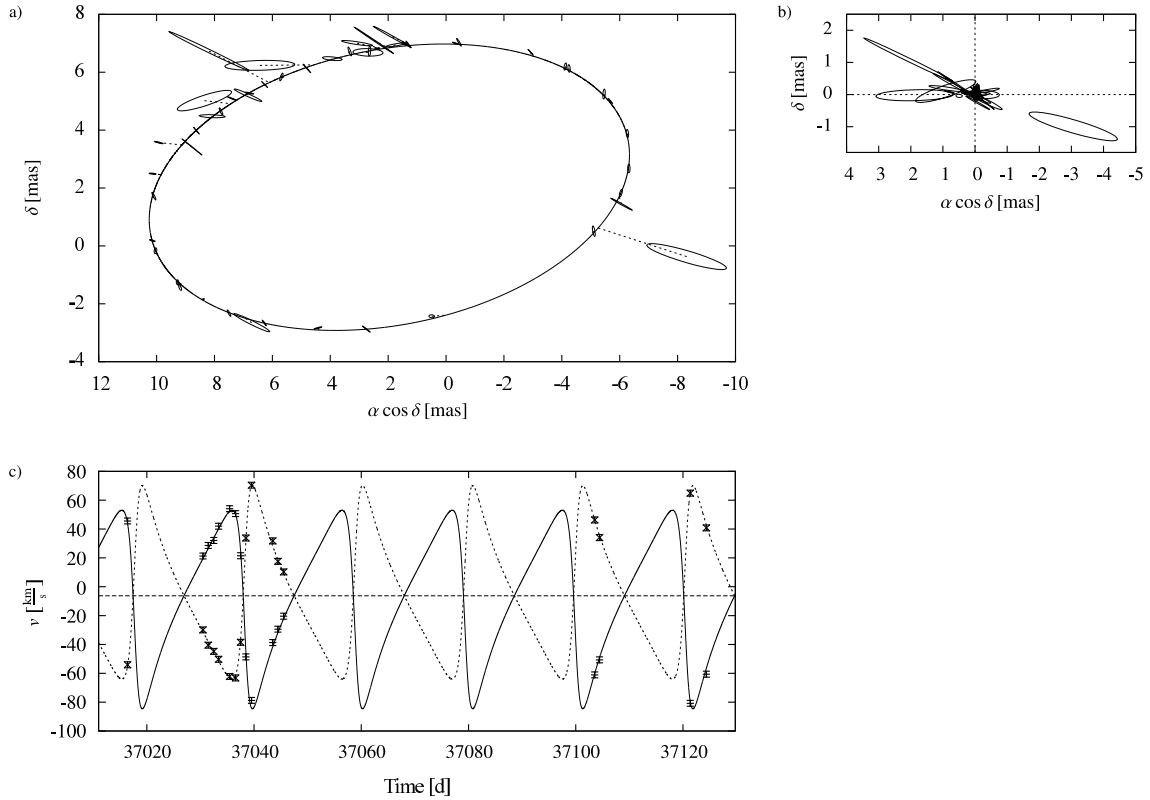


Figure 5.7: AM and RV data and models calculated with the MAP parameter estimates (table 5.9). a) AM data (uncertainty ellipses, solid lines), model (large ellipse, solid line) and residual vectors (dashed lines). b) Residual AM error ellipses. c) RV data and model for primary (normal error bars, solid line) and secondary (hourglass-shaped error bars, dotted line). The horizontal dashed line indicates the RV offset V . RV residuals are presented in fig. 5.8.

confidence intervals are also included, where permitted by the abscissa range.

Linear and nonlinear dependencies between a pair of parameters may be qualitatively judged by means of the joint marginal posteriors (section 5.2.2) shown in fig. 5.6. The inclinations of their equiprobability contours with respect to the coordinate axes are related to the corresponding correlation coefficients. Some joint marginal posteriors exhibit clear deviations from bivariate normal distributions, illustrating that a Gaussian approximation of the likelihood by use of the Fischer matrix (section 5.2.1) would be inappropriate.

Table 5.11 lists the final posterior correlation coefficients. When these attain high absolute values, model-related equations can sometimes serve as an explanation.

For example, the highly negative correlation between f and χ is related to eq. (5.26) in the following way. Given data and estimated model parameters, let t_m be the time midway between the first and last observations, and E_m the eccentric anomaly at t_m . Now increase f by a small amount. This can be balanced by a small decrease of χ (or, equivalently, by increasing the time of periastris T) such that the eccentric anomaly at t_m again equals E_m . This contrarious change of f and χ , as opposed to leaving only one of them altered, will make E deviate less, on average, over the total timespan; i.e. the model will fit the data better. The relation between f and χ so introduced is indicated by their negative correlation coefficient.

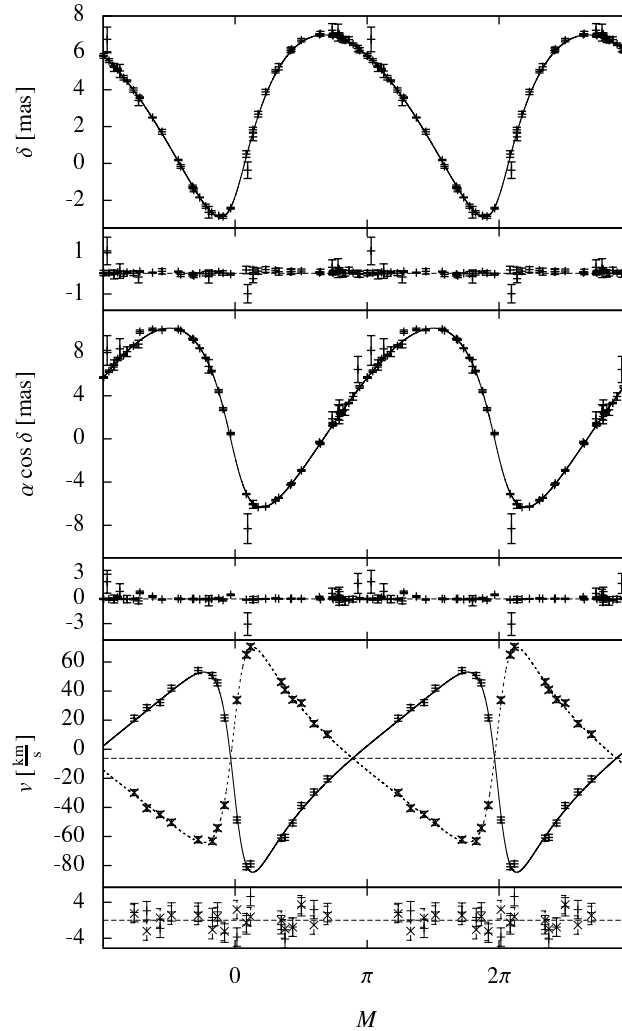


Figure 5.8: All data types: from top to bottom: AM data and model (δ coordinate); AM residuals $\Delta\delta$; AM data and model ($\alpha \cos \delta$ coordinate); AM residuals $\Delta(\alpha \cos \delta)$; RV data and model v (primary: solid line, secondary: dotted line, RV offset V : horizontal dashed line); RV residuals Δv (primary: normal error bars, secondary: dashed error bars, model: dashed line). AM error bars are defined as the sides of the smallest rectangle orientated along the coordinate axes and containing the respective error ellipse. Abscissa values are mean anomalies M (eq. (5.26)), i.e. times *folded* with respect to the MAP estimates of the time of periapsis T and period P .

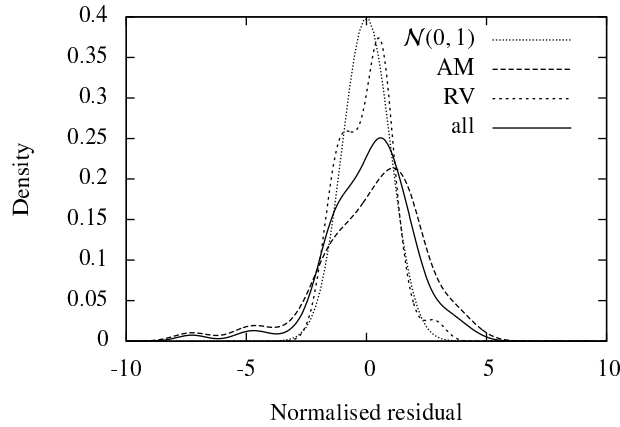


Figure 5.9: Distribution of normalised residuals of AM data (long-dashed line), RV data (short-dashed) and all data (solid), along with the standard normal distribution $\mathcal{N}(0, 1)$ (dotted). For definitions of the normalised residuals, see eq. (5.61) – (5.65).

As another example, the highly negative correlation coefficient ω_2 and Ω can be understood by reference to eq. (5.30) – (5.33), which describes the Thiele-Innes constants of the AM model as follows. In the case of edge-on orbits (inclination $i = 0$), these expressions simplify to sums or differences of $\cos(\cdot)$ and $\sin(\cdot)$ functions, the arguments being $\omega_2 + \Omega$ or $\omega_2 - \Omega$, with both arguments appearing equally often. An additional simplification is observed for face-on orbits ($i = \frac{\pi}{2}$), where the Thiele-Innes constants are $A = G = -\cos(\Omega + \omega_2)$ and $B = -F = -\sin(\Omega + \omega_2)$. Thus, for orbits nearly face-on, the strong appearance of the sum $\Omega + \omega_2$ introduces a negative correlation between the two angles, because their contrarious change can lead to the same value of the model function. For Doppler-spectroscopic data, this is counteracted by the fact that the RV model function does not contain Ω .

Models and residuals. Figure 5.7 presents all data with the models calculated from the MAP estimates listed in table 5.9, as well as residual vectors and error ellipses for astrometry. While this analysis used both the older (Hummel et al. 1995) and newer (Hummel et al. 1998) AM in combination with RV data (Prevot 1961), the AM model is very similar to the one calculated by Hummel et al. (1998, fig. 11) using only the newer AM data because of the overall agreement in parameters.

Figure 5.8 shows all data, models and residuals, separately by coordinate for AM. The abscissa corresponds to the mean anomaly M (eq. (5.26)), i.e. the plots are *folded* with respect to a time of periapsis T and period P corresponding to the posterior mode. Again, due to the similarity in parameters, the folded RV plot in fig. 5.8 is very similar to the corresponding figure in Prevot (1961).

Table 5.8: p -values of the Kolmogorov-Smirnov statistic.

Data type	p -value ²⁵
AM	0.00098
RV	0.75139
All data	0.00331

To assess the distribution of the residuals and compare it to a normal distribution, thus checking the validity of our noise model, we normalised the residuals of both data types as follows. For AM, we defined the normalised residual as a signed version of the Mahalanobis distance (Mahalanobis 1936) between the observed and the modelled values,

$$\varrho_{\text{AM},i} \equiv s_i \sqrt{(\mathbf{r}_i - \mathbf{r}(\boldsymbol{\theta}; t_i))^\top \mathbf{E}_i^{-1} (\mathbf{r}_i - \mathbf{r}(\boldsymbol{\theta}; t_i))}, \quad (5.61)$$

with sign

$$s_i \equiv \text{sgn}((\phi_i - \varphi_i)(\pi - [\phi_i - \varphi_i])), \quad (5.62)$$

where φ_i and ϕ_i are the position angles of the residual and of the uncertainty ellipse (see section 5.2.1), respectively. This definition allows us to write, according to eq. (5.8),

$$\chi_{\text{AM}}^2 = \sum_{i=1}^{N_{\text{AM}}} \varrho_{\text{AM},i}^2 \quad (5.63)$$

For RV data, we define

$$\varrho_{\text{RV},i} \equiv \frac{v_i - v(\boldsymbol{\theta}; t_i)}{\sigma_i}, \quad (5.64)$$

which analogously yields

$$\chi_{\text{RV}}^2 = \sum_{i=1}^{N_{\text{RV}}} \varrho_{\text{RV},i}^2. \quad (5.65)$$

The distributions of normalised residuals of both data types individually as well as of all data, estimated using kernel density estimation (section 5.3.4), are shown in fig. 5.9. Table 5.8 lists the p -values of the *Kolmogorov-Smirnov statistic*, which relates each distribution to the standard normal distribution $\mathcal{N}(0, 1)$.

The p -value equals the probability, under the hypothesis $\mathcal{H}_{\mathcal{N}}$ that the normalised residuals are randomly drawn from $\mathcal{N}(0, 1)$, to observe a distribution of normalised residuals that differs at least as much from $\mathcal{N}(0, 1)$ as is actually the case. This difference is quantified here by the Kolmogorov-Smirnov statistic. Denoting the hypothesis of such an observation by \mathcal{H}_{obs} , the p -value can be expressed as $\text{p}(\mathcal{H}_{\text{obs}}|\mathcal{H}_{\mathcal{N}})$. We note that according to the Bayes theorem, this is not equal to the “inverse” probability $\text{p}(\mathcal{H}_{\mathcal{N}}|\mathcal{H}_{\text{obs}})$ of the residuals coming from the normal distribution, given the observation.

For the AM residuals, as well as for all residuals, a heavy-tailed distribution (fig. 5.9) is observed and the low p -value indicates a minute probability of such an observation under $\mathcal{H}_{\mathcal{N}}$, i.e. the normalised residuals comply poorly with the standard normal distribution $\mathcal{N}(0, 1)$. This reflects the fact that several AM data points are outliers with respect to the observable and noise model (fig. 5.7). We interpret this in terms of systematic effects in the measurements that are not contained in our noise model. In contrast, the normalised RV residuals are more normally-distributed, and there are no such severe outliers (fig. 5.7 and fig. 5.8).

According to the principle of maximum entropy, given only the mean and variance of a distribution, the normal distribution has maximum information-theoretic entropy, equivalent to minimum bias or prejudice with respect to the missing information (e.g. Kapur 1989). Still, it is well-known that this widely used noise model is relatively prone to outliers; Lange et al. (1989) have suggested to replace it by a non-standardised t -distribution, resulting in the down-weighting of outliers. This distribution can be derived from the

²⁵These values are defined with respect to the final residuals and a standard normal distribution.

normal distribution under the assumption that the noise variance is unknown with a certain probability distribution. The t -distribution has an additional unknown degrees-of-freedom parameter $\nu \in \mathbb{R}$; for $\nu = 1$, it resembles the Cauchy-distribution.

In contrast, our approach has been to regard every datum with its standard deviation as accurate, which also implies that we have not discarded any data as outliers. Under this assumption, deviating from the maximum-entropy principle by selecting a different distribution introduces prior “knowledge” that we may not actually have and thus potentially biases the results.

5.6 Conclusions

We have presented BASE, a novel and highly configurable tool for Bayesian parameter and uncertainty estimation with respect to model parameters and additional derived quantities, which can be applied to AM as well as RV data of both exoplanet systems and binary stars. With user-specified or uninformative prior knowledge, it employs a combination of Markov chain Monte Carlo (MCMC) and several other techniques to explore the whole parameter space and collect samples distributed according to the posterior distribution. We presented a new, simple method of refining the window width of one-dimensional kernel density estimation, which is used to derive marginal posterior densities.

We derived the observable models from Newton’s law of gravitation, neglecting the motion of the observer and the target system, which we showed is justified in the case of Mizar A. After sketching how we estimated the RV uncertainties that are missing in the original publication (Prevot 1961), we detailed our analysis of all publicly available AM and RV data of Mizar A. It consists of five consecutive stages and has produced estimates of the values, uncertainties, and correlations of all model parameters and derived quantities, as well as marginal posterior densities over one and two dimensions.

As illustrated in fig. 5.5 and table 5.9, our new results exhibit overall compatibility with previous literature values; this is also the case for the models in fig. 5.7, whose plots differ only slightly from those published earlier. Several outliers in the AM data are visible in the distribution of the corresponding normalised residuals, which deviates significantly more from a standard normal distribution than that of the RV residuals. Nevertheless, it is not necessary to remove outliers for our program to finish successfully.

In the near future, we plan to apply BASE to a potential exoplanet host star. In this study one of the aims will be to determine the existence probability of a planetary companion.

Acknowledgements. We wish to thank David W. Hogg, Sabine Reffert, René Andrae, and Mathias Zechmeister for fruitful discussions, and an anonymous referee for comments that have improved the quality and clarity of this article. This research has made use of the SIMBAD database and VizieR catalogue access tool, operated at CDS, Strasbourg, France, as well as NASA’s Astrophysics Data System.

5.7 Appendix: Encoding prior knowledge

By means of the prior, Bayesian analysis allows one to incorporate knowledge obtained earlier, e.g. using different data. When no prior knowledge is available for some model parameter, except for its allowed range, maximum prior ignorance about the parameter

can be encoded by a prior of one of the following functional forms for the most common classes of *location* and *scale parameters* (Gregory 2005b; Sivia 2006).

- For a *location parameter*, we demand that the prior be invariant against a shift Δ in the parameter, i.e.

$$p(\theta|\mathcal{M}, \mathcal{K}) d\theta = p(\theta + \Delta|\mathcal{M}, \mathcal{K}) d(\theta + \Delta), \quad (5.66)$$

which leads to the *uniform* prior

$$p(\theta|\mathcal{M}, \mathcal{K}) = \frac{\Theta(\theta - a) \Theta(b - \theta)}{b - a}, \quad (5.67)$$

where $\Theta(\cdot)$, a , and b are the Heaviside step function, the lower and the upper prior bounds.

Here, we note that the frequentist approach, lacking an explicit definition of the prior, corresponds to the implicit assumption of a uniform prior for all parameters.

- A positive *scale parameter*, which often spans several decades, is characterised by its invariance against a stretch of the coordinate axis by a factor φ , i.e.

$$p(\theta|\mathcal{M}, \mathcal{K}) d\theta = p(\varphi\theta|\mathcal{M}, \mathcal{K}) d(\varphi\theta), \quad (5.68)$$

which is solved by the *Jeffreys prior*,

$$p(\theta|\mathcal{M}, \mathcal{K}) = \frac{\Theta(\theta - a) \Theta(b - \theta)}{\theta \ln \frac{b}{a}}. \quad (5.69)$$

That a uniform prior would be inappropriate for this parameter is also illustrated by the fact that it would assign higher probabilities to θ lying in a higher decade of $[a, b]$ than in a lower.

- If the lower prior bound of a scale parameter is zero, e.g. for the RV semi-amplitude K , a *modified Jeffreys prior* is used. It has the form

$$p(\theta|\mathcal{M}, \mathcal{K}) = \frac{\Theta(\theta - a) \Theta(b - \theta)}{(\theta + \theta_k) \ln \frac{b + \theta_k}{\theta_k}}, \quad (5.70)$$

where θ_k is the *knee* of the prior. For $\theta \ll \theta_k$, this prior is approximately uniform, while it approaches a Jeffreys prior for $\theta \gg \theta_k$.

5.8 Appendix: Numerical posterior summaries

The following tables list the numerical values of several posterior summaries. Those in tables 5.9 and 5.10 are derived from the marginal posteriors (fig. 5.5), while the correlation coefficients in table 5.11 reflect linear relations between parameters and, in this respect, can be regarded as summaries of the joint marginal posteriors (fig. 5.6).

Compared to the underlying densities, all of these summaries are incomplete. Still, they are useful e.g. for the calculation of model functions or comparison with literature results.

Table 5.9: Model parameters: new and previous results. For definitions of the estimates, see section 5.2.2.

Estimate	e	f (d^{-1})	χ	ω_2 (rad)	i (rad)	Ω (rad)	ϖ (mas)	V (km s^{-1})	K_1 (km s^{-1})	K_2 (km s^{-1})	Reference
$\hat{\theta}$	0.5304	0.0486 89388	0.93322	4.9771	1.0530	1.8381	38.74	-6.02	58.84	57.16	
$\check{\theta}$	0.5299	0.0486 89403	0.93315	4.9779	1.0528	1.8374	38.69	-6.20	58.60	57.39	
$\tilde{\theta}$	0.5295	0.0486 89409	0.93294	4.9784	1.0527	1.8365	38.91	-6.04	58.41	56.97	
$\bar{\theta}$	0.5281	0.0486 89438	0.93244	4.9807	1.0515	1.8346	39.02	-6.02	58.21	56.83	
$I_{\text{HPD}, 50\%}$	0.5282	0.0486 89341	0.93227	4.9751	1.0513	1.8345	38.13	-6.93	57.14	55.47	
	0.5317	0.0486 89467	0.93379	4.9806	1.0549	1.8397	39.66	-5.21	60.21	58.47	
$I_{\text{HPD}, 68.27\%}$	0.5270	0.0486 89295	0.93175	4.9735	1.0500	1.8327	37.74	-7.31	56.18	54.69	
	0.5326	0.0486 89509	0.93430	4.9825	1.0558	1.8411	40.06	-4.68	60.84	59.23	
$I_{\text{HPD}, 95\%}$	0.5152	0.0486 88907	0.92550	4.9653	1.0383	1.8165	35.73	-9.96	51.77	50.35	
	0.5380	0.0486 90031	0.93860	5.0019	1.0615	1.8485	42.24	-2.08	64.54	63.16	
$I_{\text{HPD}, 99\%}$	0.4905	0.0486 88483	0.91472	4.9530	1.0134	1.7790	34.24	-12.40	46.87	45.66	
	0.5451	0.0486 90874	0.94285	5.0423	1.0699	1.8593	45.02	0.22	67.05	66.49	
Lit. estimate ²⁶	0.537	0.0486 8881	0.93487	4.9595	-5.64	58.04	57.03	(1)
Uncertainty ²⁷	0.004	0.0201	0.15	0.70	0.80	
Lit. estimate ²⁸	0.5354	0.0486 89403	0.93524	4.9620	1.0559	1.850	39.4	...	58.33	56.69	(2)
Uncertainty	0.0025	0.0000 00119	0.00097	0.0052	0.0052	0.007	0.3	...	2.40	2.33	

References. (1) Prevot (1961); (2) Hummel et al. (1998)

²⁶For f and χ , literature values and uncertainties are calculated from t_1 and the original parameters P and T according to table 5.2. K_1 and K_2 are derived from $K_{\text{alt},1}$ and $K_{\text{alt},2}$ via e according to table 5.2.

²⁷Uncertainties are missing for P in Prevot (1961) and thus for f and χ .

²⁸For f and χ , see footnote 26. K_1 and K_2 are derived from $P, a'_{\text{rel}}, \varpi, \rho$ and i using eq. (5.40), (5.44) and (5.47).

Table 5.10: Derived quantities: new and previous results. For definitions of the estimates, see section 5.2.2.

Estimate	P (d)	T (d) ²⁹	d (pc)	ρ	$K_{\text{alt},1}$ (km s ⁻¹)	$K_{\text{alt},2}$ (km s ⁻¹)	m_1 (M_{\odot})	m_2 (M_{\odot})	a_1 (AU)	a_2 (AU)	a_{rel} (AU)	Reference
$\check{\theta}$	20.53 8350	36997.247	25.74	1.028	68.70	68.05	2.477	2.500	0.12657	0.12418	0.25074	
$\hat{\theta}$	20.53 8347	36997.251	25.69	1.025	68.85	67.15	2.459	2.517	0.12678	0.12353	0.25068	
$\bar{\theta}$	20.53 8335	36997.262	25.66	1.027	68.56	66.93	2.455	2.521	0.12658	0.12393	0.25017	
$I_{\text{HPD}, 50\%}$	20.53 8322	36997.234	25.18	0.986	67.30	65.39	2.320	2.311	0.12331	0.11742	0.24628	
	20.53 8375	36997.265	26.20	1.061	70.90	68.91	2.609	2.689	0.13051	0.12987	0.25555	
$I_{\text{HPD}, 68.27\%}$	20.53 8305	36997.223	24.94	0.969	66.26	64.48	2.238	2.228	0.12156	0.11385	0.24388	
	20.53 8395	36997.276	26.47	1.082	71.71	69.82	2.678	2.809	0.13269	0.13293	0.25803	
$I_{\text{HPD}, 95\%}$	20.53 8085	36997.135	23.46	0.866	60.95	59.10	1.827	1.775	0.11252	0.10058	0.22977	
	20.53 8558	36997.404	27.72	1.187	75.99	74.19	3.051	3.235	0.14028	0.14681	0.26908	
$I_{\text{HPD}, 99\%}$	20.53 7729	36997.044	21.98	0.751	54.61	53.79	1.475	1.463	0.10452	0.08958	0.21512	
	20.53 8737	36997.622	28.92	1.305	78.41	78.28	3.415	3.641	0.14637	0.16321	0.27939	
Lit. estimate	20.53 860	36997.212	...	1.018	68.80	67.60	(1)
Uncertainty	...	0.022	...	0.018	0.79	0.91	
Lit. estimate	20.53 835	36997.20	25.38	1.029	69.06	67.13	2.43	2.50	0.12652	0.12297	0.24949	(2)
Uncertainty	0.00 005	0.03	0.19	0.041	3.84	3.74	0.07	0.07	0.00519	0.00504	0.00205	

References. (1) Prevot (1961); (2) Hummel et al. (1998)

²⁹ T is given in the *reduced Julian date* scale, i.e. as Julian Date $- 2.4 \times 10^6$ d.

Table 5.11: Pearson correlation coefficients of pairs of parameters.

	e	f	χ	ω_2	i	Ω	ϖ	V	K_1
f	-0.21017
χ	0.35120	-0.43922
ω_2	-0.62560	0.18818	-0.35074
i	0.52716	-0.11723	0.28022	-0.57440
Ω	0.60484	-0.22616	0.34159	-0.62355	0.47121
ϖ	0.02224	0.04643	-0.03317	-0.02066	0.05501	-0.00134
V	-0.01699	0.02057	-0.02176	0.01468	-0.01012	-0.01436	0.00752
K_1	0.11517	-0.07289	0.09702	-0.12194	0.09822	0.11416	-0.33947	0.00868	...
K_2	0.04381	-0.03506	0.04188	-0.04901	0.03520	0.04618	-0.32930	-0.02639	0.02274

Chapter 6

A planet around ϵ Eridani?

Assessing the presence of a companion

6.1 Introduction

For decades, the nearby cool star ϵ Eridani has been suspected to host a planetary system, with the first confirmed hint at a potential companion given based on Doppler spectroscopy by Walker et al. (1995). ϵ Eridani is a low-mass star of $0.82 M_{\odot}$ (Butler et al. 2006) and

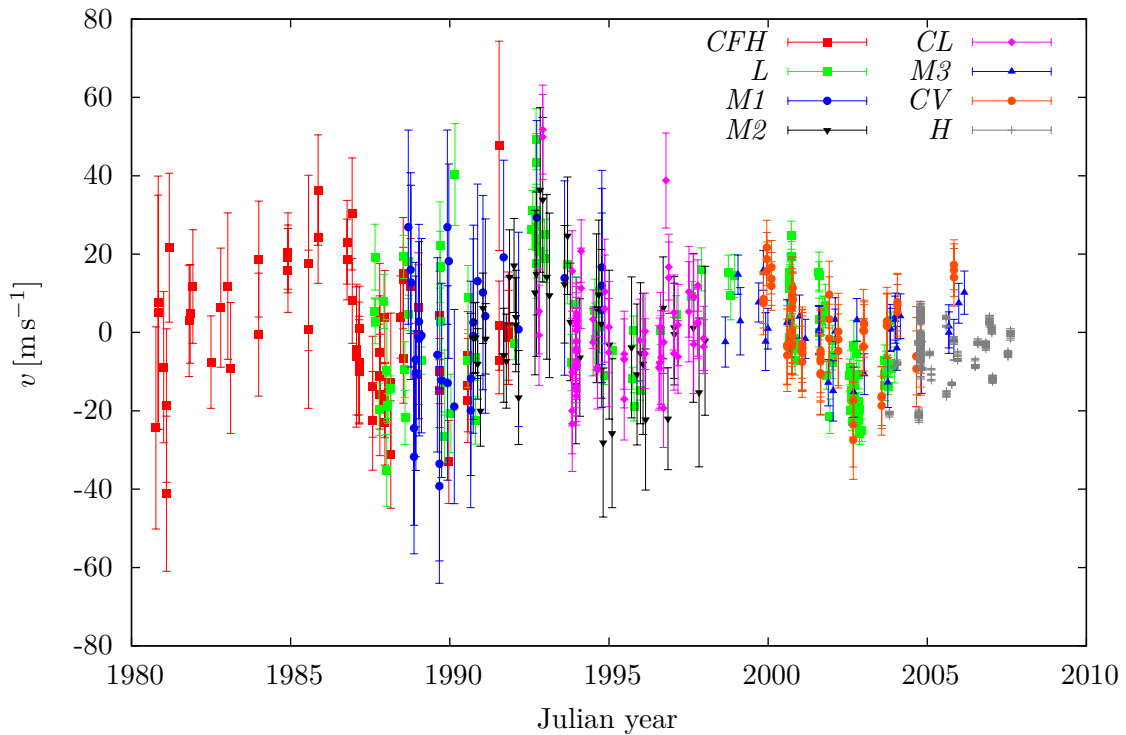


Figure 6.1: The sets of ϵ Eridani radial-velocity data analysed in this chapter. Constants have been added so as to nullify the mean values of each set. Abbreviations are defined in table 6.1.

spectral type K2 V (Gray et al. 2006), situated only about 3.22 pc from the Sun. While the controversial age estimates for ϵ Eridani range from about 100 to 1000 Myr, Janson et al. (2008) denominated 440 Myr, inferred from its rotation rate, the most probable estimate. Its nearness and youth make ϵ Eridani one of the most interesting planet-host candidates and a promising target for imaging searches. To date, however, none of the putative companions to this star have been confirmed by direct imaging.

Analyses of ϵ Eridani’s radial-velocity (RV) time series have so far yielded contradictory outcomes in terms of the frequencies and causes of periodicities. This is partly due to the fact that the star exhibits a high amount of RV *jitter*, i.e. noise caused predominantly by its well-studied strong magnetic activity (e.g. Ruedi et al. 1997) – a common feature of cool stars (Wilson 1978). This activity, perhaps undergoing quasi-periodic cycles, may modulate the photospheric granulation, causing the spectral line profiles and thus the observed radial velocity to vary coincidentally (Dravins 1985). The detection of planetary signals is further hampered by the presence of *star spots* of different temperature on the photosphere, which also modulate the observed radial velocity through changing line profiles with a period corresponding to the stellar rotation.¹

Aspects of ϵ Eridani’s activity during the years 1986 – 1992 were assessed by Gray and Baliunas (1995), who concluded that its strong magnetic activity showed “regular excursions” and hints of an underlying 5-yr cycle in the *S*-index. The Ca II H- and K-line profiles showed rotational modulation with a period of $P_{\text{rot}} = 11.1$ d, varying between 11 and 20 d over the data sets from individual seasons. ϵ Eridani’s luminosity was found to vary by only 1.2%. Walker et al. (1995) could not detect large photometric changes either, making it less likely that stellar *oscillations* are a significant contributor to the intrinsic stellar variability.

Already hinted by a far-infrared excess at 60 μm determined from the IRAS catalogue by Aumann (1985), a dusty ring or debris disk was imaged about 60 AU from ϵ Eridani with an inclination $i_{\text{disk}} \approx 25^\circ$ (Greaves et al. 1998). Its morphology could be modelled by assuming the presence of an outer planetary companion ϵ Eri c with a semi-major axis of 40 AU and a mass of about $m_c = 0.1 M_J$, corresponding to an orbital period of about 280 yr (Quillen and Thorndike 2002).

6.2 Previous work

In the following, the various analyses of AM and RV data of ϵ Eridani found in the literature are discussed. Figure 6.1 shows the RV data sets analysed in this chapter. Some of their properties are listed in table 6.1 along with the abbreviations used below.

Campbell et al. (1988) concluded from RV data spanning about six years that the star is a “probable variable”. They estimated a linear trend and curvature from the RV data, but no perturbation period. In a follow-up article, Walker et al. (1995) employed a generalisation of the Lomb-Scargle periodogram to search for periods exceeding 40 d in RV data² collected over a time span of 11 yr. Periods of $P_1 = 9.88$ yr (with a semi-amplitude $K_{\text{alt}} \approx 14 \text{ m s}^{-1}$) and $P_2 = 56$ d were found but designated only “marginally significant”. The authors determined that P_1 and P_2 were aliases of each other (\nearrow 3.1.4), without being

¹Several observables have been found which may indicate stellar activity. Changes in the symmetry of a stellar line profile can be detected from the *bisector velocity span* (*BVS*) (Toner and Gray 1988), while the *S-index* measures the strength of the Ca II H- and K-lines, thereby allowing to infer variations in the strength of the magnetic field (Schrijver et al. 1989; Baliunas et al. 1995).

²The *CFH* set, see table 6.1.

able to discern the actual periodicity or make a definite detection. However, they derived upper planetary mass limits of $m_p \lesssim 1.8 M_J$ for P_1 and $m_p \lesssim 0.5 M_J$ for P_2 .

The same data were later re-analysed by Nelson and Angel (1998) using least-squares fitting, who found the four periods $11.9 \text{ d} \approx P_{\text{rot}}$, 52.5 d , $7 \text{ yr} < P < 8 \text{ yr}$ and 10 yr , arguing that they were all probably related to stellar rotation. A common plot of RV and BVS time series of ϵ Eridani was presented by McMillan et al. (1996). Due to a visual correlation, the authors suggested that the RV variations were probably caused by granular convection.

Cumming et al. (1999) applied their floating-mean periodogram (\nearrow 3.1.4) for circular orbits to ϵ Eridani RV data covering a time span of 11.2 yr. They inferred a period of $P = 2520 \text{ d}$ with $K_{\text{alt}} = 14.7 \text{ m s}^{-1}$ (the 99th percentile being 20 m s^{-1}), but did not confirm the planetary origin of the signal.

By contrast, based on several RV data sets³ spanning over 19 yr, Hatzes et al. (2000) concluded for the first time that the presence of a planetary companion ϵ Eri b was the simplest and most likely hypothesis to explain the observed variations. The periods derived with two different methods were $P = 2502.1 \text{ d}$ and $P = 2503.5 \text{ d}$, respectively, i.e. $P \approx 6.85 \text{ yr}$, with an RV semi-amplitude $K_{\text{alt}} = 19.0 \pm 1.7 \text{ m s}^{-1}$. A periodogram of the Ca II H and K S -index revealed periods of 20 yr, 3 yr and 3.8 yr, in the order of decreasing periodogram power, while a peak at 6.78 yr, near P , had a lower power and was deemed insignificant. The authors admitted that the variations might still be caused by stellar activity, but noted that since no correlation with the S -index was observed, this would be counter to common understanding. Among other parameters, the orbital solution of Hatzes et al. (2000) was characterised by eccentricity $e = 0.608 \pm 0.041$ and “projected” planetary mass (\nearrow 2.2.4) $m_p \sin i = 0.86 M_J$.

Around the same time, Gatewood (2000) announced the results of a first astrometric orbital analysis of ϵ Eri b. After cancelling proper motion and parallactic displacement from 112 Multichannel Astrometric Photometer (MAP) data, and keeping the RV parameters fixed, they estimated the additional AM parameters, including an angular semi-major axis of the stellar orbit $a' = 1.51 \pm 0.41 \text{ mas}$, a planetary mass $m_b = 1.2 \pm 0.33 M_J$ and orbital inclination $i = 46 \pm 17^\circ$.

Zucker and Mazeh (2001) performed a combined fit of RV data with the original HIPPARCOS intermediate AM data (Perryman et al. 1997) to find an angular semi-major axis of 10.1 mas, with a “peculiarly high” 99th-percentile value of 26.49 mas determined by means of a bootstrapping technique.

Another joint fit to AM and RV data was performed by Benedict et al. (2006), who used AM data from the HST Fine Guidance Sensor 1r obtained over 3 yr along with various RV data sets⁴ covering more than 25 yr. Additionally, MAP AM data were added to improve the determination of proper motion and parallax. The authors found a period of $2502 \pm 10 \text{ d}$, RV semi-amplitude $K_{\text{alt}} = 18.5 \pm 0.2 \text{ m s}^{-1}$ and projected mass $m_p \sin i = 0.78 \pm 0.08$, similar to the results of Hatzes et al. (2000), whereas their best-fit eccentricity $e = 0.702 \pm 0.039$ was somewhat higher. Some of the astrometric parameters of Gatewood (2000) were approximately reproduced by Benedict et al. (2006), including the angular semi-major axis $a' = 1.88 \pm 0.2 \text{ mas}$, mass $m_b = 1.55 \pm 0.24 M_J$, and inclination $i = 30.1 \pm 3.8^\circ \approx i_{\text{disk}}$.

Based on 120 RV data⁵ spanning over 16 yr, Butler et al. (2006) characterised the orbit by a period $P = 2500 \pm 350 \text{ d}$ and RV semi-amplitude $K_{\text{alt}} = 18.6 \pm 2.9 \text{ m s}^{-1}$, both very similar to the results of Hatzes et al. (2000) and Benedict et al. (2006), but a lower

³Probably including subsets of the *CFH*, *L*, *M1*, *M2*, and *CL* sets.

⁴Including set *M3*.

⁵Set *L*.

Table 6.1: Radial-velocity data sets used in this chapter, sorted by first observing time. The median uncertainty $\tilde{\sigma}$ and the root mean square deviation about the mean, ϱ , are defined in section 6.3.2.

Abbrev.	Observatory/ instrument	Timespan [Jyr]	No. records	$\tilde{\sigma}$ [m s ⁻¹]	ϱ [m s ⁻¹]	Reference
<i>CFH</i>	CFHT	1980.8–1991.9	65	13.4	16.48	1
<i>L</i>	Lick	1987.7–2004.0	120	4.5	16.58	2
<i>M1</i>	McDonald	1988.7–1994.8	32	24.8	17.71	3
<i>M2</i>	McDonald	1990.8–1998.1	42	15.0	14.26	4
<i>CL</i>	CES+LC	1992.8–1998.0	66	9.68	13.64	5
<i>M3</i>	McDonald	1998.7–2006.2	33	5.6	7.39	6
<i>CV</i>	CES+VLC	1999.9–2005.9	69	8.29	9.91	7
<i>H</i>	HARPS	2003.8–2007.7	521	0.32	5.73	8

References. (1) Walker et al. (1995); (2) Butler et al. (2006); (3) Benedict et al. (2006); Meschiari et al. (2009); (4) Benedict et al. (2006); Meschiari et al. (2009); (5) Endl et al. (2002); Zechmeister, M. (2011, priv.comm.); (6) Benedict et al. (2006); (7) Zechmeister, M. (2011, priv.comm.); (8) Zechmeister, M. (2012, priv.comm.)

eccentricity $e = 0.25 \pm 0.23$ and higher projected mass $m_p \sin i = 1.06 \pm 0.16 M_J$.

Zechmeister (2010) searched for signatures of circular and eccentric orbits in a combination of three RV data sets⁶ covering nearly 15 yr. No significant periods were detected, with the highest periodogram powers found at 1.49 d (eccentric) and 6.30 d (circular), respectively. The period around 7 yr could not be confirmed.

Reffert and Quirrenbach (2011) performed a fit to AM_h data (\nearrow 2.2.2), keeping the RV parameters fixed at previously published values. The resulting most likely parameters included an inclination $i = 23 \pm 20^\circ$ and companion mass $m_p = 2.4 \pm 1.1 M_J$. However, due to the high uncertainty of the AM parameters, the authors did not consider this a significant detection.

Recently, Anglada-Escudé and Butler (2012) published their analysis of seven RV data sets⁷ spanning over 26 yr. The best-fit results included an eccentricity $e = 0.4$ (the 99% confidence interval being [0.2, 0.68]), period $P = 2651 \pm 36$ d, semi-amplitude $K_{\text{alt}} = 11.8 \pm 1.1 \text{ m s}^{-1}$, and projected mass $m_p \sin i = 0.645 \pm 0.058 M_J$. Since all of these numbers differ significantly from most previous estimates, the authors doubted that the RV variations are of planetary origin but instead referred to activity as a likely cause.

This chapter aims to assess the presence of a planet-induced RV signal in the available AM_h and RV data. This is accomplished by estimating a Bayesian “periodogram” (\nearrow 4.3.3) using BASE. In contrast to the above-mentioned frequentist periodograms, this approach includes well-defined priors on all parameters of the AM_h and RV models (\nearrow 2.2), none of which need to be kept fixed. The “periodogram” is given by the marginal posterior of orbital frequency f . Further, the Bayes factor for the competing zero- and one-planet hypotheses is estimated.

⁶The *CL* and *CV* sets and most of the *H* set.

⁷The data sets *CFH*, *L*, *M1*, *M2*, *M3*, and two other data sets originating from the CES+LC and HARPS instruments.

Table 6.2: Literature values for parameters. Units are given in table 2.1.

Parameter	Best estimate	Uncertainty σ	Reference
ϖ	0.31094	0.00016	1
μ_{α_*}	-975.17	0.21	1
μ_δ	19.49	0.20	1
α_{r}	53.2350902200	3.89×10^{-8}	1
δ_{r}	-9.4583060400	3.06×10^{-8}	1

References. (1) van Leeuwen (2007)

6.3 Analysis and results

The results of a combined Bayesian analysis of HIPPARCOS intermediate astrometric data (AM_h) and RV data of ϵ Eridani are presented in the following. 78 AM_h data (section 2.2.2) spanning Julian years 1990.0 – 1992.6, with a median measurement uncertainty of 610 μas , were analysed along with the RV data listed in table 6.1. The priors of all parameters were chosen as described in the following.

6.3.1 Determination of priors

Section 2.3 introduced the default prior bounds set according to general considerations for most parameters and according to the data for V , a_v , σ_+ , τ_+ , Ω , and K in normal mode. For the RV semi-amplitude K , the upper bound K_{max} was determined from the RV range.⁸ These default priors have not been altered, except for those discussed in the following. The prior ranges used are listed in table 6.3.

Jitter. To determine the maximum allowable jitter, two runs of BASE were conducted:

1. Pass A aimed to determine the amount of RV jitter σ_+ not accounted for in the RV data uncertainties. This pass was based on the assumption that the marginal posterior of σ_+ will be shifted to lower jitter values as the observable model becomes more complex, i.e. that less jitter is “needed” to model the data when a planetary signal is included, compared to a model without planet. Thus, to derive an upper limit $\sigma_{+, \text{max}}$, BASE was run in 0-planets mode with all RV data. The secular perspective acceleration was held fixed at the value of $\dot{v} = 0.07031 \text{ m s}^{-1} \text{ yr}^{-1}$ determined using eq. (2.49) with the literature values of proper motion μ_{α_*} , μ_δ and parallax ϖ listed in table 6.2. The default priors of the other parameters were not changed.

Pass A resulted in a 99% HPDI (\nearrow 3.2.6) upper bound

$$\sigma_{+, \text{max}, 99\%} = 8.365 \frac{\text{m}}{\text{s}}, \quad (6.1)$$

which was kept throughout the following runs as a fixed upper bound, while the lower bound was set to zero to allow arbitrarily small amounts of jitter. The upper bound may be compared with, e.g., the external-noise estimate of $6.6 \text{ m s}^{-1} \approx 0.8 \sigma_{+, \text{max}, 99\%}$ by Anglada-Escudé and Butler (2012).

⁸This is activated in BASE with the `-K_max-from-vel` option.

Table 6.3: Prior ranges for the ϵ Eridani analysis. Non-default ranges are derived in section 6.3.1. Units are given in table 2.1. Parameters printed in boldface only appear in 1-planet mode.

Parameter	Default range	Lower bound	Upper bound
V_1	×	−69.6	48.5
V_2	×	−48.5	56
V_3	×	−60.9	74.4
V_4	×	−47.7	53.8
V_5	×	−23.8	19.8
V_6	×	13675.51	13774.16
V_7	×	13677.28	13743.44
V_8	×	16430.785	16461.749
a_v	×	−5.806	5.806
σ_+		0	8.364
ϖ		0.280	0.342
α_r		53.2350 8998	53.2350 9045
δ_r		−9.4583 0622	−9.4583 0585
μ_{α_*}		−1072.687	−877.653
μ_δ		17.541	21.439
τ_+		0	1.283
i	×	0	π
Ω	×	0	2π
e	×	0	1
f		0.0001	1
χ	×	0	1
ω	×	0	2π
K	×	0	67.65
a'		0.001	4.319

2. In pass B, an analogous procedure was followed for the AM_h jitter τ_+ , independent from the RV data. The prior supports for the five AM parameters were set to relatively wide intervals around the previously published best estimates (table 6.2) in order to reflect the high uncertainty due to the different estimation techniques, whereas the default prior of τ_+ was not altered. For parallax and proper motion, ranges of $\pm 10\%$ around the best estimates were adopted. Because this choice would not be appropriate for location parameters (\nearrow 3.2.2) such as α_r and δ_r , they were assigned ranges of $\pm 6\sigma$. These priors were kept unchanged over all subsequent passes.

Pass B yielded a 99% HPDI upper bound

$$\tau_{+, \max, 99\%} = 1.283 \text{ mas.} \quad (6.2)$$

Orbital frequency. All analyses in 1-planet mode were based on the assumption that at most one period P is contained in the total observing time span $\Delta t = 26.85 \text{ yr}$. Thus, the minimum orbital frequency was set to $f_{\min} \equiv 10^{-4} \text{ d}^{-1} \approx (26.85 \text{ yr})^{-1}$, whereas the upper bound f_{\max} was set to 1 d^{-1} , corresponding to a planetary orbital semi-major axis of 0.034 AU or approximately ten stellar radii of ϵ Eridani.

Angular semi-major axis. When analysing the AM_h data separately, the orbital semi-major axis over distance, i.e. angular semi-major axis a' , appears as a model parameter. For it, an upper bound of $6 \text{ stdev}(\{a_{h,j}\})$ was adopted based on the assumptions that the orbital phases of significant stellar displacement \mathbf{r}_5 (eq. (2.41)) are covered by the data, that the angles between \mathbf{r}_5 and the scan circles are sufficiently small (taken modulo π) to “detect” the displacement, and therefore the dispersion of the abscissa residuals $\{\Delta a_{h,j}\}$ provides an upper limit on a' .

6.3.2 Bayesian periodogram of all data

Under the assumption of a 1-planet model, BASE was started in pass C with all AM_h and RV data and the priors described above to produce a Bayesian “periodogram”, i.e. marginal posterior of orbital frequency f . BASE collected 5×10^6 posterior samples with a thinning stride $q = 10$ to improve convergence with unchanged memory demands (section 4.5.2). The last 75% of the thinned samples were used for inference.⁹

The resulting marginal posterior density is displayed in fig. 6.2. No major peak is present within any of the previously published confidence intervals lying in the zoomed interval of fig. 6.2b. This range has a posterior probability of 5.3×10^{-4} , calculated as the fraction of posterior samples falling into it (see eq. (3.64)). Thus, none of the peaks in this interval can be considered significant. In particular, the highest peak, at $f \approx 3.95 \times 10^{-4} \text{d}^{-1} \approx (2530 \text{d})^{-1}$, covers only 7.4×10^{-5} posterior probability.

No peak with density $p(f|D, \mathcal{M}, \mathcal{K}) > 0.5$ is visible in the frequency range of fig. 6.2c, which has a minute probability of 1.9×10^{-5} and includes the estimates $f = (9.88 \text{yr})^{-1}$ (Walker et al. 1995) and $f = (10 \text{yr})^{-1}$ (Nelson and Angel 1998).

The marginal posterior of f is found to rise for $(1 \text{d} + 17 \text{s})^{-1} = 0.9998 \text{d}^{-1} \leq f \leq 1 \text{d}^{-1}$, i.e. towards the highest frequency searched (fig. 6.2d), with the interval covering a posterior probability of 0.84%. This peak might be explained as an alias as follows. On the one hand, it is evident from fig. 6.1 that any signal potentially present in the RV data likely has a relatively low signal-to-noise ratio, which is also hinted by comparison of the measurement uncertainties with the radial-velocity dispersion: excluding the H data with their very low quoted uncertainties, the median measurement uncertainty is $\tilde{\sigma} = 9 \text{m s}^{-1}$, while the *root mean square (RMS)* deviation of all RVs about the corresponding means is

$$\varrho_{\text{tot}} = \sqrt{\frac{1}{\sum_i N_i} \sum_{i=1}^n \sum_{j=1}^{N_i} \left(v_j^{(i)} - \overline{v^{(i)}} \right)^2} = 14.5 \text{m s}^{-1} = 1.6 \tilde{\sigma}, \quad (6.3)$$

where N_i is the number of data in set i , n is the number of data sets, $v_j^{(i)}$ is the j th RV of set i and $\overline{v^{(i)}}$ is the mean RV of set i . Due to this low dispersion compared with the internal noise level, the data are marginally consistent with a *constant* RV, associated with a frequency $f_0 = 0$. This is especially true since additional jitter σ_+ is allowed for.

On the other hand, due to the timing constraints forced by the day–night cycle, data from Earth-bound telescopes are affected by sampling periods of multiples of one day. For uniformly spaced sampling at frequency f_s , a sinusoidal signal of frequency f can be shown (e.g. Dawson and Fabrycky 2010) to have aliases at frequencies

$$f_a = \pm f \pm n f_s, \quad n \in \mathbb{N}. \quad (6.4)$$

⁹These settings were kept with all passes.

Table 6.4: Summaries for orbital frequency f [d^{-1}] from pass C (all data). Numbers have been rounded to five significant digits unless inappropriate. Definitions of the estimates are given in section 3.2.5.

Estimate	Value
\hat{f}	0.61564
\tilde{f}	0.46730
\bar{f}	0.48043
σ_{mean}	0.12740
\check{f}	0.9999985
σ_{marg}	0.0000010
$I_{\text{HPD}, 50\%}$	0.020628
	0.50148
$I_{\text{HPD}, 68.27\%}$	0.017454
	0.73696
$I_{\text{HPD}, 95\%}$	0.016285
	0.97965
$I_{\text{HPD}, 99\%}$	0.0076181
	0.99999988

Here, with $f = f_0 = 0$ and $f_s = 1 \text{ d}^{-1}$, aliases would therefore be expected at frequencies $f_a = n \text{ d}^{-1}$, including the mentioned peak at 1 d^{-1} . Likewise, the local maximum seen in fig. 6.2a at $f = 0.5 \text{ d}$ might stem from an observation sampling at $f'_s = (2 \text{ d})^{-1}$.

The posterior summaries for f are revealed by table 6.4 to be quite inconclusive. While the median \tilde{f} equals the mean \bar{f} to within its uncertainty σ_{mean} , both MAP value \hat{f} and marginal mode \check{f} are significantly different, with the latter equal to $\check{f} = (1 \text{ d} + (130 \pm 86 \text{ ms}))^{-1}$.

6.3.3 Bayesian periodograms of individual data sets

In order to assess which data set is responsible for which of the many dominant frequencies in the combined marginal posterior (fig. 6.2a), a set of additional passes $D_1 - D_9$ was

Table 6.5: Posterior probabilities for frequencies $f \in I$, with $I \equiv [10^{-4} \text{ d}^{-1}, 5 \times 10^{-4} \text{ d}^{-1}]$, based on individual data sets ordered by increasing observing time span Δt .

Data set	Δt [yr]	$p(f \in I D, \mathcal{M}, \mathcal{K})$
AM_h	2.5	0.0500
H	3.8	5.07×10^{-4}
CL	5.2	0.0131
CV	6.0	0.0596
$M1$	6.1	0.183
$M2$	7.3	0.195
$M3$	7.5	0.287
CFH	11.1	0.344
L	16.3	0.199

conducted, each using only one data set in the order of increasing observing time span (table 6.5).

The resulting marginal posteriors are displayed in fig. 6.3. A striking feature of these densities is the rise of more or less clear peaks around low frequencies $f \lesssim 5 \times 10^{-4} \text{ d}^{-1} \approx (5 \text{ yr})^{-1}$ with increasing time span in certain data sets, i.e. fig. 6.3c – 6.3d (sets *CL* and *CV*) and fig. 6.3f – 6.3i (sets *M2*, *M3*, *CFH*, and *L*). In the latter set of figures, the maximum also rises with an increasing RMS-to-uncertainty ratio¹⁰ $\varrho \tilde{\sigma}^{-1}$ but similar time span Δt (fig. 6.3f – 6.3g) and again becomes more pronounced with a longer time span but similar RMS-to-uncertainty ratio (fig. 6.3g – 6.3h). Finally, for the last data set with its longest time span and second-highest $\varrho \tilde{\sigma}^{-1}$ (fig. 6.3i), a peak at $f = (2668 \text{ d})^{-1}$ containing about 17.4% probability is clearly visible.

Obviously, both Δt and $\varrho \tilde{\sigma}^{-1}$ correlate positively with the height and “sharpness” of the peaks around $f \approx (2500 \text{ d})^{-1}$, a frequency which approximately corresponds to the periods published in some of the previous literature. The role of the time spans is also demonstrated by table 6.5, which lists the posterior probabilities for frequencies $f \in [10^{-4}, 5 \times 10^{-4}]$: a steady rise in probability is observed with the time span increasing for sets *H* through *CFH*, excepting only the shortest and longest data sets. Moreover, in the diagrams of fig. 6.3a, 6.3e, and 6.3f, a general increase in probability towards very low frequencies $f \lesssim 2 \times 10^{-4} \text{ d}^{-1}$, any “specific” peaks being only minor, is also evident. These correspond to the data sets with the lowest $\varrho \tilde{\sigma}^{-1}$. In the latter case, where a sub-peak around $f \approx 3.5 \times 10^{-4} \text{ d}^{-1}$ becomes visible (fig. 6.3f), Δt is the highest of the three and $\varrho \tilde{\sigma}^{-1}$ exceeds that of the second diagram. This positive correlation of the time spans Δt and RMS-to-uncertainty ratios $\varrho \tilde{\sigma}^{-1}$ on the one hand with the probability around $f \approx (2500 \text{ d})^{-1}$ on the other hand may indicate a truth to this signal.

One particular data set, however, is different from all the others (table 6.1) in that it combines the by-far highest RMS-to-uncertainty ratio and number of data with the second-lowest time span: the *H* set. Figure 6.3b and table 6.5 reveal that this set by itself easily has the highest posterior probability for orbital frequencies in excess of $5 \times 10^{-4} \text{ d}^{-1}$, with strong peaks recognisable only for $f \gtrsim 0.01 \text{ d}^{-1}$. A regularity is seen in the spacing of the peaks at these higher frequencies, which may be due to aliasing (\nearrow 6.3.2) caused by the large volume of data taken over a short time span. The latter circumstance may particularly increase the effect of the daily observing windows in comparison to longer windows or true periods. Comparison with fig. 6.3a, i.e. the *AM_h* set with its lower RMS-to-uncertainty ratio $\varrho \tilde{\sigma}^{-1}$ but similarly low time span, may indicate that the high $\varrho \tilde{\sigma}^{-1}$ of the *H* set also plays a role for the strength of higher frequencies.

Consequently, the *H* set may be suspected as a significant cause for the strong presence of peaks at higher frequencies in the combined marginal posterior (fig. 6.2) and the relative weakness of lower frequencies, despite the latter being prominent in several of the individual data sets (fig. 6.3). With its 521 data and very low internal uncertainties – whose median does not exceed 7.1% of that of any other set (table 6.1) – this data set undoubtedly takes the strongest influence on the combined likelihood (eq. (3.21)) and therefore on the marginal posterior of f . This proposition seems to be further corroborated by the fact that the Bayesian evidence, i.e. prior-averaged likelihood, could only be calculated when data grouping was applied, reducing the number of HARPS data to an order of magnitude similar to that of most other data sets (section 6.3.4).

¹⁰The RMS per data set is defined analogously to eq. (6.3). The values of RMS and median uncertainties are listed for each set in table 6.1.

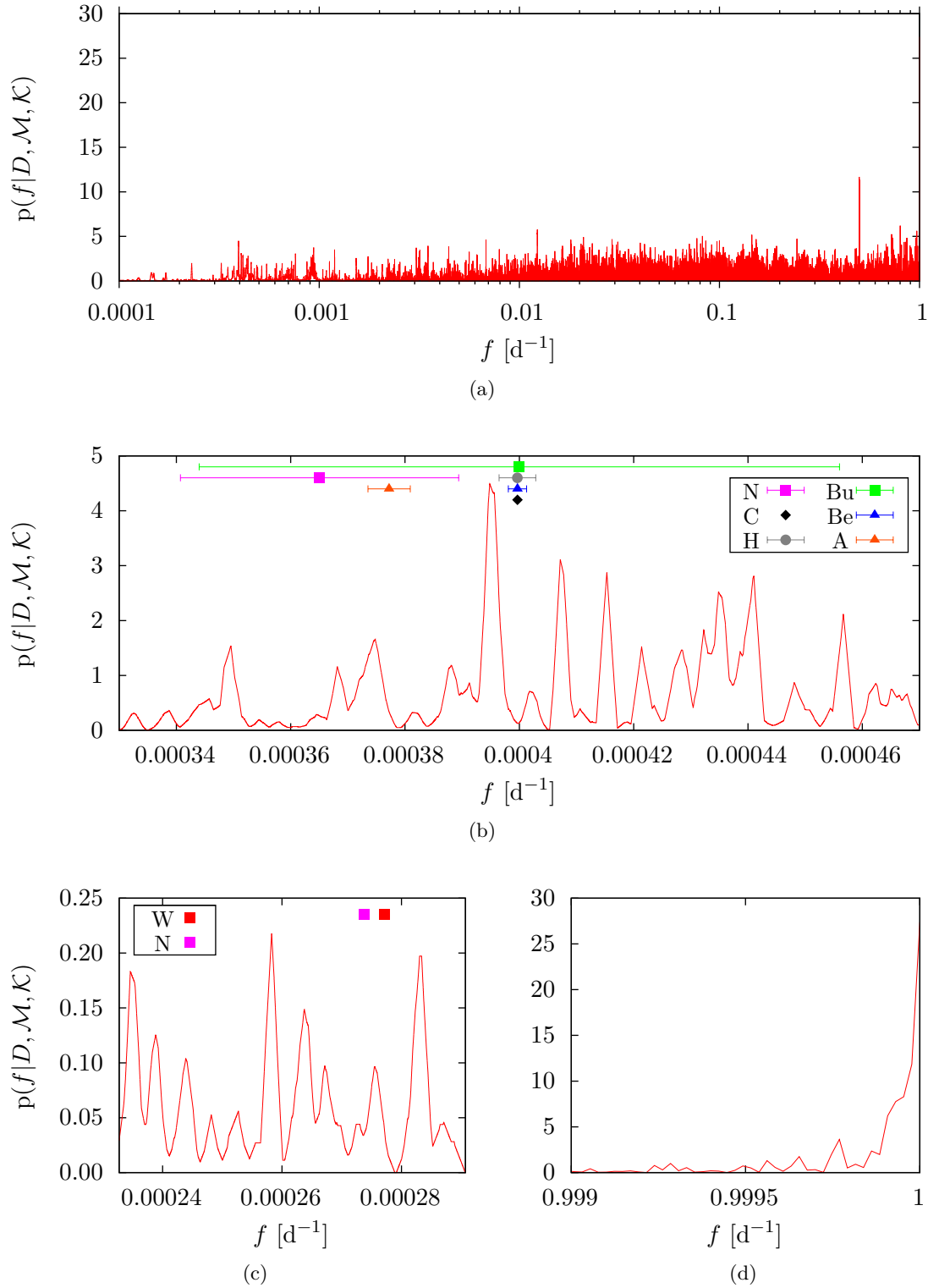


Figure 6.2: Marginal posterior of orbital frequency f under the 1-planet model using all data (pass C). The ordinates are normalised according to linear spacing in a logarithmic abscissa. **(a)** Periodogram over complete f range. **(b)** Blow-up of interval containing most literature estimates, indicated by horizontal error bars. **(c)** Frequency interval free from peaks with $p(f|D, \mathcal{M}, \mathcal{K}) > 0.5$ around two further literature estimates. **(d)** Upper frequency range. Abbreviations: N (Nelson and Angel 1998), C (Cumming et al. 1999), H (Hatzes et al. 2000), Bu (Butler et al. 2006), Be (Benedict et al. 2006), A (Anglada-Escudé and Butler 2012), W (Walker et al. 1995).

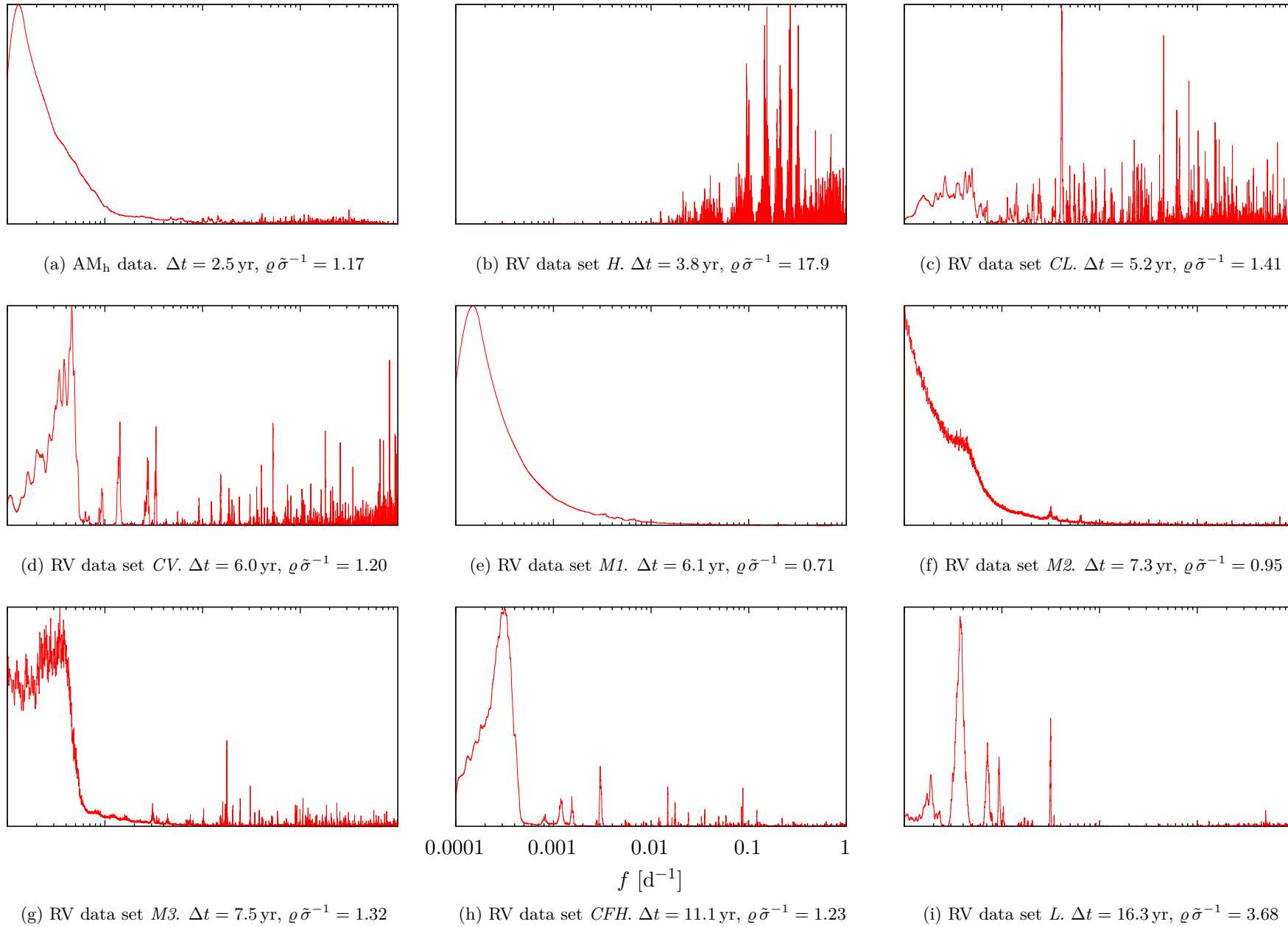


Figure 6.3: Periodograms of each data set taken separately, ordered by increasing observing time span from left to right and top to bottom. Ordinates follow from normalisation and have been omitted for clarity. Also listed are the observing time spans Δt and the ratios $\rho \tilde{\sigma}^{-1}$ of the RMS deviation of the observable from its mean to the median measurement uncertainty.

Table 6.6: Number of original and grouped data for each set.

	AM_h	CFH	L	$M1$	$M2$	CL	$M3$	CV	H
Original	78	65	120	32	42	66	33	69	521
Grouped	42	50	120	32	42	28	29	23	28

6.3.4 Model selection

In order to contrast the above reasoning with a quantitative measure of the probability of a planet’s presence around ϵ Eridani, the Volume Tessellation Algorithm (VTA, ↗ 3.2.7) implemented in BASE was applied to the posterior samples based on all data sets. Using the data unaltered, the evidences $\mathcal{Z}_0, \mathcal{Z}_1$ for the 0- and 1-planet models, respectively, were numerically zero. This may have been caused by the large number of data, especially the over 500 HARPS RVs with very low nominal measurement uncertainties, which in combination imply tiny likelihoods for parameter combinations associated with even small residuals (eq. (3.6), (3.16) and (3.20)).

To circumvent this obstacle, the data were also grouped (↗ 4.4.2) using maximum time spans of 30 min, 2 h, and 5 h respectively. For the L , $M1$, and $M2$ sets, no groups could be formed even to within 5 h. For the CFH , $M3$, and CV sets, the groups remained unchanged regardless of the chosen time span, while for the other two RV sets, the resulting number of data varied within a range of only ± 2 . Therefore, a grouping time span of 2 h was adopted. For the AM_h data, the variations in scan-orientation angle and parallax factor within each group were found to lie below 1° and 1%, respectively. The number of original and modified data in each set are listed in table 6.6.

With grouping, the evidences and Bayes factors given in table 6.7 were calculated. According to Kass and Raftery (1995), Bayes factors $B_{1,0} < 10^{-2}$ may be interpreted as “decisive” evidence for model \mathcal{M}_0 . In the present case, the estimated values $B_{1,0} < 10^{-108}$ are, of course, many orders of magnitude lower and seem to provide definite evidence in favour of the no-planet hypothesis.

Three caveats, however, should not remain unmentioned. First, this inference is based on the specified prior knowledge. While in the frequentist framework, prior knowledge takes the form of flat, implicit priors (↗ 3.2.2), the Bayesian approach makes priors “visible”, causing the necessity – and chance – to explicitly specify them. However, care has been taken in specifying the prior knowledge for the present analysis (section 6.3.1). Second, the values are also based on the available data, their number, time sampling, and other potential imperfections. And third, the data were grouped together in order to be able to perform the calculations. The fact that this reduced the number of HARPS data from 521 to only 28, i.e. the same order of magnitude as for most other sets, illustrates the significant effect of grouping.

6.4 Conclusions

According to the analyses presented, it seems questionable whether the determined Bayes factors should be taken as conclusive evidence against a planet around ϵ Eridani. Besides the modification of data applied in the latter case, the results of sections 6.3.2 and 6.3.3 suggest that the outcome of any analysis based on the present AM_h and RV data may be strongly influenced by the aliasing, large number and small nominal uncertainties of the HARPS data set. As mentioned in section 6.3.3, the fact that the prior-averaged

Table 6.7: Evidences \mathcal{Z}_0 and \mathcal{Z}_1 for the 0- and 1-planet models, respectively, and Bayes factors $B_{1,0}$. Values have been estimated with different numbers of leaves c (section 3.2.7).

c	\mathcal{Z}_0	\mathcal{Z}_1	$B_{1,0}$
4	1.49×10^{-700}	3.59×10^{-826}	2.42×10^{-126}
8	1.23×10^{-698}	2.45×10^{-807}	2.00×10^{-109}
16	2.35×10^{-689}	2.85×10^{-805}	1.21×10^{-116}
32	3.07×10^{-676}	2.85×10^{-805}	9.28×10^{-130}
64	1.36×10^{-676}	6.40×10^{-800}	4.71×10^{-124}

likelihoods, i.e. evidences, could only be calculated with data grouping also hints that an excessive number of low-uncertainty RV data, particularly when strongly affected by aliasing, may interfere with the combined analysis of all data. The fact that a periodicity was clearly detected by Hatzes et al. (2000), Anglada-Escudé and Butler (2012) and others, but not by us, may thus be attributed to the circumstance that their analyses did not make use of a similar number of data comparable with those in the HARPS set.

Further research is therefore necessary to better understand the significance of the effects discussed in this chapter for the analysis of the ϵ Eridani data. Particularly a deepened understanding of ϵ Eridani’s strong and potentially manifold activity seems to be key to answering the question about the reality of a planetary companion. Developing models for the activity-induced effects on the observed data which go beyond the often-made assumption of additional Gaussian stellar jitter may aid any further analyses of astrometric, radial-velocity, or other data for the minute effects of planetary motion. Such an understanding indeed may be helped by large volumes of high-cadence data such as those in the HARPS data set, because they provide a better time resolution of the RV variation than many other data. As detailed analyses of ϵ Eridani’s activity would have been beyond the scope of this work, they remain only to be recommended for future research. Based on such future assessments, supplemented with a better understanding of the effects of the narrow time sampling on the current problem and potentially based on further data, it may be possible to draw definite and statistically reliable conclusions on the presence of its long-suspected planetary companion.

Chapter 7

Conclusions

A short summary of this work

While the search for life outside the Earth still has not been successful, potential places of residence for it are being unveiled in growing numbers. Probably formed in a common process with their host stars, such *exoplanets* are not only of prime interest to those interested in extraterrestrial life. Even the history of our own Solar System, the phases and time scales of its evolution, are still not completely understood, and increasing the sample of well-characterised planetary systems potentially quite different from our own helps to put constraints on the variety of scenarios for the formation and evolution of such systems.

Detecting exoplanets and determining their orbits around their host stars has been made possible by a number of different observational techniques, enabling scientists to obtain observational evidence for the stellar motion caused by its companion. The alternative of directly imaging the planet faces severe problems due to the very high brightness contrast to the nearby host star. For indirectly detected companions, the question of whether they are indeed of planetary nature or rather of higher mass may not be an obvious one to answer, since the stellar motion is governed in all cases by the same physical laws¹ and only aspects of it can be observed. Moreover, noise and other sources of error contribute to the measured value by a generally unknown amount, raising the need for reliable statistical methods and an improved instrumental precision both. This is particularly the case as the focus shifts towards the relatively low-mass rocky planets akin to our Earth.

Two kinds of observational techniques, namely astrometry and Doppler spectroscopy, have been investigated in the present work. We have derived models by which to theoretically describe the values of several corresponding observables: relative astrometric positions of two binary components obtained using interferometry, astrometric *abscissa residuals* pertaining to the host star as measured by instruments aboard now the inoperative HIPPARCOS satellite, and stellar radial velocities routinely inferred from shifts in the stellar spectral lines. These models depend on a set of parameters which ultimately characterise the orbits of the involved objects around each other or the centre of mass.

Given the observed data, the model parameters are adjusted in a statistical data-analysis process so as to make the model approximate the observed values. The unavoidable *residuals*, or mismatch, between data and model are identified with noise in the absence of systematic errors and under the assumption that the model is “correct”. The model parameters

¹Only classical mechanics have been considered in this thesis.

are repeatedly adjusted, each time rating the resulting residuals by reference to a *noise model*, thus judging the appropriateness of the underlying parameters. Finally, the most probable values of the parameters, or a probability density over them, can be estimated, corresponding to the characterisation of the planetary orbit.

In this work, the *Bayesian* approach to inference is detailed primarily. It allows to explicitly specify *prior knowledge* on the model parameters and, in light of the data, to obtain their *posterior* probability density. This density can then be summarised by most probable parameter values and uncertainties. Also, *Bayes factors* can be estimated which include an *Ockham's razor* penalising a too-complex model for its inflated parameter space. Bayes factors therefore provide a statistically sound quantity for selecting the most probable of several models – one of the features of the Bayesian approach missing in the traditional set of *frequentist* techniques. Besides, predictions of the future values of the observable models and their uncertainties can be made, helping to optimally schedule upcoming observations.

The posterior is estimated from parameter samples obtained by the *Markov chain Monte Carlo (MCMC)* technique,² which has been implemented in a computer program called BASE. It allows to explore a multidimensional parameter space, treating astrometric (AM) and radial-velocity (RV) data in a joint analysis. On user request, data may be automatically arranged in groups of chronologically successive measurements, a procedure that may be useful when the time spacing of the data is closer than the lowest orbital period P presumed *a priori*, but should be applied with caution.

When prior knowledge is not explicitly provided by the user, it is supplemented by BASE in terms of relatively uninformative priors. However, prior knowledge may also be specified in several forms, including prior-density samples and fixed parameter values.

BASE can model binary stars as well as (multi-)planet systems. In the latter case, the tool allows to estimate the Bayes factors of competing models, thus assessing the most likely number of planets, including zero. This constitutes a Bayesian type of exoplanet detection. For parameter estimation, BASE estimates *marginal posterior densities* over one or two parameters and also provides a set of numerical summaries for each parameter as well as the correlations between them. Furthermore, in *periodogram mode*, the marginal posterior of the orbital frequency $f = P^{-1}$ is calculated in a refined manner so as to make all of the often very thin peaks visible. BASE incorporates *kernel density estimation*, providing smooth and differentiable density estimates based on a finite number of samples.

In this thesis, we have provided an overview of the modes of using BASE, its functionality as well as its structural components in terms of its program flow, central algorithms, and the organisation of its source code.

We have also applied BASE to publicly available AM and RV data of the well-known binary star Mizar A. We have detailed the estimation of the missing RV uncertainties, followed by the application of BASE in a set of successive passes. After obtaining first constraints on several RV parameters, all data were combined to obtain the most probable range of orbital frequency in periodogram mode, based on a wide prior range. In the next pass, a degeneracy between two other orbital parameters was resolved before obtaining final results for all orbital parameters corresponding to a particular orbit and its uncertainty. The determined parameter values and uncertainties are compatible with previously published results and constitute reliable knowledge on the orbital characteristics of Mizar A.

²Besides MCMC, *parallel tempering* is also implemented in BASE in order to facilitate the exploration of the complete parameter space without becoming “trapped” in small regions of very high probability. Moreover, convergence of the collected samples to the posterior is supervised by the *multi-PT* procedure.

Finally, a well-known putative exoplanet host star has been examined: ϵ Eridani. While first confirmed hints at a planetary companion to this relatively young and nearby star were already given almost twenty years ago, and although ϵ Eridani has been well-studied, the question still has not been answered whether it is in fact part of a planetary system. Direct imaging of the companion has been without success thus far, and due to ϵ Eridani's strong magnetic activity, the abundant RV data from various telescopes have not allowed either to draw definite conclusions on a planet's presence, let alone its orbit.

We have analysed eight separate, publicly available sets of RV data and a set of HIPPARCOS intermediate-astrometry abscissa residuals observed of ϵ Eridani. We have assessed the periodicities present in them, as well as in the combination of all data. The periods found in the literature could not be confirmed with all data sets combined, whereas similar periods were found in *some* of the data sets. From the probabilities of the frequencies found and their correlation with potentially quality-related properties of the data sets, we have concluded that a signal with period $P \approx 2500$ d may be present in reality but hidden due to the strong disturbance caused by the aliasing and dominance of one of the data sets. Due to numerical issues, the important Bayes factor could only be calculated after grouping the data to within time spans of 2 h. Although the determined Bayes factor strongly favours the no-planet hypothesis, it is to be interpreted with caution not only because of the grouping applied, but also since its estimation may have been affected by the data imperfections mentioned above. The latter circumstance could also explain why other authors, particularly including Hatzes et al. (2000) who first denominated the planet hypothesis the most likely of all, did detect a unanimous periodicity – all previous analyses were based on different data.

An improved understanding of ϵ Eridani's strong activity, which may indeed be aided by the availability of closely-spaced data, will help to better characterise the effects of the activity-induced jitter on the observed data. While such an activity analysis was beyond the scope of this work, it remains to be recommended for future research. It appears that guided by a deepened understanding of activity- and sampling-related effects influencing the data, definite conclusions may be drawn on the reality of a perhaps minute planetary signal caused by a companion to ϵ Eridani.

Bibliography

- Adamson, A., C. Aspin, C. Davis, T. Fujiyoshi, and A. Adamson, C. Aspin, C. Davis, & T. Fujiyoshi (Eds.) (2005, December). *Astronomical Polarimetry: Current Status and Future Directions*, Volume 343 of *Astronomical Society of the Pacific Conference Series*.
- Alibert, Y., C. Mordasini, and W. Benz (2011, February). Extrasolar planet population synthesis. III. Formation of planets around stars of different masses. *A&A* 526, A63.
- Anglada-Escudé, G., A. P. Boss, A. J. Weinberger, I. B. Thompson, R. P. Butler, S. S. Vogt, and E. J. Rivera (2012, February). Astrometry and Radial Velocities of the Planet Host M Dwarf GJ 317: New Trigonometric Distance, Metallicity, and Upper Limit to the Mass of GJ 317b. *ApJ* 746, 37.
- Anglada-Escudé, G. and R. P. Butler (2012, June). The HARPS-TERRA Project. I. Description of the Algorithms, Performance, and New Measurements on a Few Remarkable Stars Observed by HARPS. *ApJS* 200, 15.
- Armstrong, J. T., D. Mozurkewich, L. J. Rickard, D. J. Hutter, J. A. Benson, P. F. Bowers, N. M. Elias, II, C. A. Hummel, K. J. Johnston, D. F. Buscher, J. H. Clark, III, L. Ha, L. Ling, N. M. White, and R. S. Simon (1998, March). The Navy Prototype Optical Interferometer. *ApJ* 496, 550–+.
- Armstrong, J. T., D. Mozurkewich, M. Vivekanand, R. S. Simon, C. S. Denison, K. J. Johnston, X. Pan, M. Shao, and M. M. Colavita (1992, July). The orbit of Alpha Equulei measured with long-baseline optical interferometry - Component masses, spectral types, and evolutionary state. *AJ* 104, 241–252.
- Aumann, H. H. (1985, October). IRAS observations of matter around nearby stars. *PASP* 97, 885–891.
- Baliunas, S. L., R. A. Donahue, W. H. Soon, J. H. Horne, J. Frazer, L. Woodard-Eklund, M. Bradford, L. M. Rao, O. C. Wilson, Q. Zhang, W. Bennett, J. Briggs, S. M. Carroll, D. K. Duncan, D. Figueroa, H. H. Lanning, T. Misch, J. Mueller, R. W. Noyes, D. Poppe, A. C. Porter, C. R. Robinson, J. Russell, J. C. Shelton, T. Soyumer, A. H. Vaughan, and J. H. Whitney (1995, January). Chromospheric variations in main-sequence stars. *ApJ* 438, 269–287.
- Bayes, M. and M. Price (1763, January). An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions* 53, 370–418.
- Benedict, G. F., B. E. McArthur, G. Gatewood, E. Nelan, W. D. Cochran, A. Hatzes, M. Endl, R. Wittenmyer, S. L. Baliunas, G. A. H. Walker, S. Yang, M. Kürster, S. Els,

- and D. B. Paulson (2006, November). The Extrasolar Planet ϵ Eridani b: Orbit and Mass. *AJ* 132, 2206–2218.
- Bond, G. (1857, June). Photographical Experiments on the Positions of Stars. *MNRAS* 17, 230–+.
- Boneh, A. and A. Golan (1979). Constraints' redundancy and feasible region boundedness by random feasible point generator (rfpg). In *Third European Congress on Operations Research, EURO III*, Amsterdam.
- Boschi, R., V. Lucarini, and S. Pascale (2012, July). Bistability of the climate around the habitable zone: a thermodynamic investigation. *ArXiv e-prints*.
- Brooks, S. P. and A. Gelman (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7, 434–455.
- Butler, R. P., J. T. Wright, G. W. Marcy, D. A. Fischer, S. S. Vogt, C. G. Tinney, H. R. A. Jones, B. D. Carter, J. A. Johnson, C. McCarthy, and A. J. Penny (2006, July). Catalog of Nearby Exoplanets. *ApJ* 646, 505–522.
- Campbell, B., G. A. H. Walker, and S. Yang (1988, August). A search for substellar companions to solar-type stars. *ApJ* 331, 902–921.
- Charbonneau, D., T. M. Brown, D. W. Latham, and M. Mayor (2000, January). Detection of Planetary Transits Across a Sun-like Star. *ApJ* 529, L45–L48.
- Chubak, C., G. Marcy, D. A. Fischer, A. W. Howard, H. Isaacson, J. A. Johnson, and J. T. Wright (2012, July). Precise Radial Velocities of 2046 Nearby FGKM Stars and 131 Standards. *ArXiv e-prints*.
- Cumming, A. (2004, November). Detectability of extrasolar planets in radial velocity surveys. *MNRAS* 354, 1165–1176.
- Cumming, A., G. W. Marcy, and R. P. Butler (1999, December). The Lick Planet Search: Detectability and Mass Thresholds. *ApJ* 526, 890–915.
- Dawson, R. I. and D. C. Fabrycky (2010, October). Radial Velocity Planets De-aliased: A New, Short Period for Super-Earth 55 Cnc e. *ApJ* 722, 937–953.
- Deeg, H. J., J. A. Belmonte, and A. Aparicio (Eds.) (2007, October). *Extrasolar Planets*. Cambridge University Press.
- Deeming, T. J. (1975, August). Fourier Analysis with Unequally-Spaced Data. *Ap&SS* 36, 137–158.
- Delplancke, F. (2008, June). The PRIMA facility phase-referenced imaging and micro-arcsecond astrometry. *New A Rev.* 52, 199–207.
- Delplancke, F., S. A. Leveque, P. Kervella, A. Glindemann, and L. D’Arcio (2000, July). Phase-referenced imaging and micro-arcsecond astrometry with the VLTI. In P. Léna & A. Quirrenbach (Ed.), *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Volume 4006 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pp. 365–376.

-
- Dravins, D. (1985). Stellar lineshifts induced by photospheric convection. In A. G. D. Philip and D. W. Latham (Eds.), *Stellar Radial Velocities*, pp. 311–320.
- Efron, B. and R. Tibshirani (1993). *An introduction to the bootstrap*. Monographs on statistics and applied probability. Chapman & Hall.
- Endl, M., M. Kürster, S. Els, A. P. Hatzes, W. D. Cochran, K. Dennerl, and S. Döbereiner (2002, September). The planet search program at the ESO Coudé Echelle spectrometer. III. The complete Long Camera survey results. *A&A* 392, 671–690.
- Ferraz-Mello, S. (1981, April). Estimation of Periods from Unequally Spaced Observations. *AJ* 86, 619.
- Ford, E. B. (2004, June). Quantifying the Uncertainty in the Orbits of Extrasolar Planets with Markov Chain Monte Carlo. In S. S. Holt and D. Deming (Eds.), *The Search for Other Worlds*, Volume 713 of *American Institute of Physics Conference Series*, pp. 27–30.
- Ford, E. B. (2006, May). Improving the Efficiency of Markov Chain Monte Carlo for Analyzing the Orbits of Extrasolar Planets. *ApJ* 642, 505–522.
- Ford, E. B. (2008, March). Adaptive Scheduling Algorithms for Planet Searches. *AJ* 135, 1008–1020.
- Foreman-Mackey, D., D. W. Hogg, D. Lang, and J. Goodman (2012, February). emcee: The MCMC Hammer. *ArXiv e-prints*.
- Foster, G. (1995, April). The cleanest Fourier spectrum. *AJ* 109, 1889–1902.
- Free Software Foundation (2011a). GFortran. <http://gcc.gnu.org/fortran/>.
- Free Software Foundation (2011b). GOMP. <http://gcc.gnu.org/projects/gomp/>.
- Gatewood, G. (2000, October). The Actual Mass of the Object Orbiting Epsilon Eridani. In *AAS/Division for Planetary Sciences Meeting Abstracts #32*, Volume 32 of *Bulletin of the American Astronomical Society*, pp. 1051.
- Gatewood, G., L. Breakiron, R. Goebel, S. Kipp, J. Russell, and J. Stein (1980, February). On the astrometric detection of neighboring planetary systems. II. *Icarus* 41, 205–231.
- Gelman, A. and D. B. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* 7(4), pp. 457–472.
- Geman, S. and D. Geman (1984, nov.). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI-6*(6), 721–741.
- Gilks, W. R., S. Richardson, and D. J. Spiegelhalter (1996). *Markov Chain Monte Carlo in Practice* (first ed.). London: Chapman & Hall.
- Gillessen, S., F. Eisenhauer, G. Perrin, W. Brandner, C. Straubmeier, K. Perraut, A. Amorim, M. Schöller, C. Araujo-Hauck, H. Bartko, H. Baumeister, J. Berger, P. Carvas, F. Cassaing, F. Chapron, E. Choquet, Y. Clenet, C. Collin, A. Eckart, P. Fedou, S. Fischer, E. Gendron, R. Genzel, P. Gitton, F. Gonte, A. Gräter, P. Haguenaier, M. Haug, X. Haubois, T. Henning, S. Hippler, R. Hofmann, L. Jocou, S. Kellner,

- P. Kervella, R. Klein, N. Kudryavtseva, S. Lacour, V. Lapeyrere, W. Laun, P. Lena, R. Lenzen, J. Lima, D. Moratschke, D. Moch, T. Moulin, V. Naranjo, U. Neumann, A. Nolot, T. Paumard, O. Pfuhl, S. Rabien, J. Ramos, J. M. Rees, R. Rohloff, D. Rouan, G. Rousset, A. Sevin, M. Thiel, K. Wagner, M. Wiest, S. Yazici, and D. Ziegler (2010, July). GRAVITY: a four-telescope beam combiner instrument for the VLTI. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Volume 7734 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*.
- Gould, A. (2009, March). Recent Developments in Gravitational Microlensing. In K. Z. Stanek (Ed.), *Astronomical Society of the Pacific Conference Series*, Volume 403 of *Astronomical Society of the Pacific Conference Series*, pp. 86–+.
- Gray, D. F. and S. L. Baliunas (1995, March). Magnetic activity variations of epsilon Eridani. *ApJ* 441, 436–442.
- Gray, R. O., C. J. Corbally, R. F. Garrison, M. T. McFadden, E. J. Bubar, C. E. McGahee, A. A. O’Donoghue, and E. R. Knox (2006, July). Contributions to the Nearby Stars (NStars) Project: Spectroscopy of Stars Earlier than M0 within 40 pc-The Southern Sample. *AJ* 132, 161–170.
- Greaves, J. S., W. S. Holland, G. Moriarty-Schieven, T. Jenness, W. R. F. Dent, B. Zuckerman, C. McCarthy, R. A. Webb, H. M. Butner, W. K. Gear, and H. J. Walker (1998, October). A Dust Ring around epsilon Eridani: Analog to the Young Solar System. *ApJ* 506, L133–L137.
- Green, R. (1985). *Spherical Astronomy*. Cambridge University Press.
- Gregory, P. C. (2005a, October). A Bayesian Analysis of Extrasolar Planet Data for HD 73526. *ApJ* 631, 1198–1214.
- Gregory, P. C. (2005b). *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with ‘Mathematica’ Support*. Cambridge: Cambridge University Press.
- Gregory, P. C. (2011, January). Bayesian exoplanet tests of a new method for MCMC sampling in highly correlated model parameter spaces. *MNRAS* 410, 94–110.
- Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika* 57, 97–109.
- Hatzes, A. P., W. D. Cochran, B. McArthur, S. L. Baliunas, G. A. H. Walker, B. Campbell, A. W. Irwin, S. Yang, M. Kürster, M. Endl, S. Els, R. P. Butler, and G. W. Marcy (2000, December). Evidence for a Long-Period Planet Orbiting ϵ Eridani. *ApJ* 544, L145–L148.
- Hoffleit, D. and C. Jaschek (1982). *The Bright Star Catalogue*. Yale University Observatory.
- Holman, M. J. and N. W. Murray (2005, February). The Use of Transit Timing to Detect Terrestrial-Mass Extrasolar Planets. *Science* 307, 1288–1291.
- Hummel, C. A., J. T. Armstrong, D. F. Buscher, D. Mozurkewich, A. Quirrenbach, and M. Vivekanand (1995, July). Orbits of Small Angular Scale Binaries Resolved with the Mark III Interferometer. *AJ* 110, 376–+.

-
- Hummel, C. A., D. Mozurkewich, J. T. Armstrong, A. R. Hajian, N. M. Elias, II, and D. J. Hutter (1998, November). Navy Prototype Optical Interferometer Observations of the Double Stars Mizar A and Matar. *AJ* 116, 2536–2548.
- Janson, M., S. Reffert, W. Brandner, T. Henning, R. Lenzen, and S. Hippler (2008, September). A comprehensive examination of the ϵ Eridani system. Verification of a 4 micron narrow-band high-contrast imaging approach for planet searches. *A&A* 488, 771–780.
- Kapur, J. (1989). *Maximum-Entropy Models in Science and Engineering*. Wiley.
- Kass, R. and A. Raftery (1995). Bayes factors. *Journal of the American Statistical Association* 90, 773–795.
- Katzgraber, H. G., S. Trebst, D. A. Huse, and M. Troyer (2006). Feedback-optimized parallel tempering monte carlo. *Journal of Statistical Mechanics: Theory and Experiment* 2006.
- Koch, D. G., W. J. Borucki, G. Basri, N. M. Batalha, T. M. Brown, D. Caldwell, J. Christensen-Dalsgaard, W. D. Cochran, E. DeVore, E. W. Dunham, T. N. Gautier, III, J. C. Geary, R. L. Gilliland, A. Gould, J. Jenkins, Y. Kondo, D. W. Latham, J. J. Lissauer, G. Marcy, D. Monet, D. Sasselov, A. Boss, D. Brownlee, J. Caldwell, A. K. Dupree, S. B. Howell, H. Kjeldsen, S. Meibom, D. Morrison, T. Owen, H. Reitsema, J. Tarter, S. T. Bryson, J. L. Dotson, P. Gazis, M. R. Haas, J. Kolodziejczak, J. F. Rowe, J. E. Van Cleve, C. Allen, H. Chandrasekaran, B. D. Clarke, J. Li, E. V. Quintana, P. Tenenbaum, J. D. Twicken, and H. Wu (2010, April). Kepler Mission Design, Realized Photometric Performance, and Early Science. *ApJ* 713, L79–L86.
- Kuhn, J. R., D. Potter, and B. Parise (2001, June). Imaging Polarimetric Observations of a New Circumstellar Disk System. *ApJ* 553, L189–L191.
- Lange, K. L., R. J. A. Little, and J. M. G. Taylor (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistical Association* 84(408), pp. 881–896.
- Launhardt, R., D. Queloz, T. Henning, A. Quirrenbach, F. Delplancke, L. Andolfato, H. Baumeister, P. Bizenberger, H. Bleuler, B. Chazelas, F. D erie, L. Di Lieto, T. P. Duc, O. Duvanel, N. M. Elias, II, M. Fluery, R. Geisler, D. Gillet, U. Graser, F. Koch, R. K ohler, C. Maire, D. M egevand, Y. Michellod, J. Moresmau, A. M uller, P. M ullhaupt, V. Naranjo, F. Pepe, S. Reffert, L. Sache, D. S egransan, Y. Salvad e, T. Schulze-Hartung, J. Setiawan, G. Simond, D. Sosnowska, I. Stiliz, B. Tubbs, K. Wagner, L. Weber, P. Weise, and L. Zago (2008, July). The ESPRI project: astrometric exoplanet search with PRIMA. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Volume 7013 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*.
- Lecar, M., M. Podolak, D. Sasselov, and E. Chiang (2006, April). On the Location of the Snow Line in a Protoplanetary Disk. *ApJ* 640, 1115–1118.
- Levine, M., R. Soummer, J. Arenberg, R. Belikov, P. Bierden, A. Boccaletti, R. Brown, A. Burrows, C. Burrows, E. Cady, W. Cash, M. Clampin, C. Cossapakis, I. Crossfield, L. Dewell, R. Egerman, H. Fergussan, J. Ge, A. Give’On, O. Guyon, S. Heap, T. Hyde, B. Jaroux, J. Jasnin, J. Kasting, M. Kenworthy, S. Kilston, A. Klavins, J. Krist, M. Kuchner, B. Lane, C. Lillie, R. Lyon, J. Lloyd, A. Lo, P. J. Lowrance, P. J. Macintosh, S. McCully, M. Marley, C. Marois, G. Matthews, D. Mawet, B. Mazin, G. Mosier,

- C. Noecker, L. Pueyo, B. R. Oppenheimer, N. Pedreiro, M. Postman, A. Roberge, S. Ridgeway, Schneider, J. Schneider, G. Serabyn, S. Shaklan, M. Shao, A. Sivaramakrishnan, D. Spergel, K. Stapelfeldt, M. Tamura, D. Tenerelli, V. Tolls, W. Traub, J. Trauger, R. J. Vanderbei, and J. Wynn (2009). Overview of Technologies for Direct Optical Imaging of Exoplanets. In *astro2010: The Astronomy and Astrophysics Decadal Survey*, Volume 2010 of *Astronomy*, pp. 37.
- Lindegren, L. and D. Dravins (2003, April). The fundamental definition of “radial velocity”. *A&A* 401, 1185–1201.
- Lomb, N. R. (1976, February). Least-squares frequency analysis of unequally spaced data. *Ap&SS* 39, 447–462.
- Loredo, T. J. (2004, April). Bayesian Adaptive Exploration. In G. J. Erickson and Y. Zhai (Eds.), *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Volume 707 of *American Institute of Physics Conference Series*, pp. 330–346.
- Lovis, C. and D. Fischer (2010). *Radial Velocity Techniques for Exoplanets*, pp. 27–53. University of Arizona Press.
- Lyot, B. (1932). Étude de la couronne solaire en dehors des éclipses. Avec 16 figures dans le texte. *ZAp* 5, 73–+.
- Mahalanobis, P. C. (1936, April). On the generalised distance in statistics. In *Proceedings National Institute of Science, India*, Volume 2, pp. 49–55.
- Mamajek, E. E., M. A. Kenworthy, P. M. Hinz, and M. R. Meyer (2010, March). Discovery of a Faint Companion to Alcor Using MMT/AO 5 μm Imaging. *AJ* 139, 919–925.
- Mao, S. and B. Paczynski (1991, June). Gravitational microlensing by double stars and planetary systems. *ApJ* 374, L37–L40.
- Marois, C., D. Lafrenière, R. Doyon, B. Macintosh, and D. Nadeau (2006, April). Angular Differential Imaging: A Powerful High-Contrast Imaging Technique. *ApJ* 641, 556–564.
- Mayor, M. and D. Queloz (1995, November). A Jupiter-Mass Companion to a Solar-Type Star. *Nature* 378, 355–+.
- McArthur, B. E., G. F. Benedict, R. Barnes, E. Martioli, S. Korzennik, E. Nelan, and R. P. Butler (2010, June). New Observational Constraints on the ν Andromedae System with Data from the Hubble Space Telescope and Hobby-Eberly Telescope. *ApJ* 715, 1203–1220.
- McMillan, R. S., T. L. Moore, M. L. Perry, and P. H. Smith (1996, September). Correlation of the radial velocity of Epsilon Eridani with its magnetic cycle. In *Bulletin of the American Astronomical Society*, Volume 28 of *Bulletin of the American Astronomical Society*, pp. 1111.
- Meschiari, S., A. S. Wolf, E. Rivera, G. Laughlin, S. Vogt, and P. Butler (2009, September). Systemic: A Testbed for Characterizing the Detection of Extrasolar Planets. I. The Systemic Console Package. *PASP* 121, 1016–1027.

-
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953, June). Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* *21*, 1087–1092.
- Mordasini, C., Y. Alibert, W. Benz, and D. Naef (2009, July). Extrasolar planet population synthesis. II. Statistical comparison with observations. *A&A* *501*, 1161–1184.
- Moulton, F. (1984). *An Introduction to Celestial Mechanics*. Dover Books on Astronomy. Dover Publications.
- Nascimbeni, V., G. Piotto, L. R. Bedin, and M. Damasso (2011, March). TASTE: The Asiago Search for Transit timing variations of Exoplanets. I. Overview and improved parameters for HAT-P-3b and HAT-P-14b. *A&A* *527*, A85.
- Nelson, A. F. and J. R. P. Angel (1998, June). The Range of Masses and Periods Explored by Radial Velocity Searches for Planetary Companions. *ApJ* *500*, 940.
- Niepraschk, R. and H. Voß (2001). The package ps4pdf: from postscript to pdf. *TUG-boat* *22*(4), 290 – 292.
- OpenMP Architecture Review Board (2008). OpenMP. <http://openmp.org/>.
- Papaloizou, J. C. B. and C. Terquem (2006, January). Planet formation and migration. *Reports on Progress in Physics* *69*, 119–180.
- Perryman, M. (2011, June). *The Exoplanet Handbook*.
- Perryman, M. A. C. (2000, August). Extra-solar planets. *Reports on Progress in Physics* *63*, 1209–1272.
- Perryman, M. A. C., L. Lindegren, J. Kovalevsky, E. Hoeg, U. Bastian, P. L. Bernacca, M. Cr ez e, F. Donati, M. Grenon, F. van Leeuwen, H. van der Marel, F. Mignard, C. A. Murray, R. S. Le Poole, H. Schrijver, C. Turon, F. Arenou, M. Froeschl e, and C. S. Petersen (1997, July). The HIPPARCOS Catalogue. *A&A* *323*, L49–L52.
- Pickering, E. C. (1890, February). On the spectrum of zeta Ursae Majoris. *The Observatory* *13*, 80–81.
- Prevot, L. (1961). Vitesses radiales et  el ements orbitaux de ζ_1 Ursae Majoris. *Journal des Observateurs* *44*, 83–+.
- Protassov, R., D. A. van Dyk, A. Connors, V. L. Kashyap, and A. Siemiginowska (2002, May). Statistics, Handle with Care: Detecting Multiple Model Components with the Likelihood Ratio Test. *ApJ* *571*, 545–559.
- Quillen, A. C. and S. Thorndike (2002, October). Structure in the ϵ Eridani Dusty Disk Caused by Mean Motion Resonances with a 0.3 Eccentricity Planet at Periastron. *ApJ* *578*, L149–L152.
- Reegen, P. (2007, June). SigSpec. I. Frequency- and phase-resolved significance in Fourier space. *A&A* *467*, 1353–1371.
- Reegen, P. (2011, December). SigSpec User’s Manual. *Communications in Asteroseismology* *163*, 3.

- Reffert, S. (2009, November). Astrometric measurement techniques. *New A Rev.* 53, 329–335.
- Reffert, S. and A. Quirrenbach (2011, March). Mass constraints on substellar companion candidates from the re-reduced Hipparcos intermediate astrometric data: nine confirmed planets and two confirmed brown dwarfs. *A&A* 527, A140.
- Roberts, D. H., J. Lehar, and J. W. Dreher (1987, April). Time Series Analysis with Clean - Part One - Derivation of a Spectrum. *AJ* 93, 968.
- Roberts, G. O. (1996). Markov chain concepts related to sampling algorithms. In W. R. Gilks, S. Richardson, and D. J. Spiegelhalter (Eds.), *Markov Chain Monte Carlo in Practice* (first ed.), pp. 45–57. London: Chapman & Hall.
- Rueddi, I., S. K. Solanki, G. Mathys, and S. H. Saar (1997, February). Magnetic field measurements on moderately active cool dwarfs. *A&A* 318, 429–442.
- Scargle, J. D. (1982, December). Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. *ApJ* 263, 835–853.
- Schneider, J. (2012). The Extrasolar Planets Encyclopædia. <http://exoplanet.eu/>. Accessed on November 10, 2012.
- Schneider, J., Dedieu, C., Le Sidaner, P., Savalle, R., and Zolotukhin, I. (2011). Defining and cataloging exoplanets: the exoplanet.eu database. *A&A* 532, A79.
- Schrijver, C. J., J. Cote, C. Zwaan, and S. H. Saar (1989, February). Relations between the photospheric magnetic field and the emission from the outer atmospheres of cool stars. I - The solar CA II K line core emission. *ApJ* 337, 964–976.
- Schulze-Hartung, T. (2008). Bayesian astrometric and spectroscopic exoplanet detection and characterization software. Diploma thesis, University of Heidelberg.
- Schulze-Hartung, T., R. Launhardt, and T. Henning (2012, September). Bayesian analysis of exoplanet and binary orbits. Demonstrated using astrometric and radial-velocity data of Mizar A. *A&A* 545, A79.
- Seager, S. (2003, March). The search for extrasolar Earth-like planets. *Earth and Planetary Science Letters* 208, 113–124.
- Seager, S. (2008, March). Exoplanet Transit Spectroscopy and Photometry. *Space Sci. Rev.* 135, 345–354.
- Shao, M., M. M. Colavita, B. E. Hines, D. H. Staelin, and D. J. Hutter (1988, March). The Mark III stellar interferometer. *A&A* 193, 357–371.
- Silverman, B. (1986). *Density estimation for statistics and data analysis*. Monographs on statistics and applied probability. Chapman and Hall.
- Sivia, D. S. (2006). *Data Analysis—A Bayesian Tutorial* (second ed.). Oxford: Oxford University Press.
- Smith, R. L. (1980). A monte carlo procedure for the generation of feasible solutions to mathematical programming problems. In *Bulletin of the TIMS/ORSA Joint National Meeting*, Washington, DC, pp. 101.

-
- Smith, W. H. (1987, December). Spectral differential imaging detection of planets about nearby stars. *PASP* 99, 1344–1353.
- Sozzetti, A. (2005, October). Astrometric Methods and Instrumentation to Identify and Characterize Extrasolar Planets: A Review. *PASP* 117, 1021–1048.
- Stumpff, K. (1973). *Himmelsmechanik*, Volume 1. VEB Deutscher Verlag der Wissenschaften.
- Thiele, T. N. (1883, January). Neue Methode zur Berechnung von Doppelsternbahnen. *Astronomische Nachrichten* 104, 245–+.
- Tolbert, C. R. (1964, May). A UBV Study of 94 Wide Visual Binaries. *ApJ* 139, 1105–+.
- Toner, C. G. and D. F. Gray (1988, November). The starpatch on the G8 dwarf XI Bootis A. *ApJ* 334, 1008–1020.
- Torres, G., J. Andersen, and A. Giménez (2010, February). Accurate masses and radii of normal stars: modern results and applications. *A&A Rev.* 18, 67–126.
- Tuomi, M., S. Kotiranta, and M. Kaasalainen (2009, February). The complementarity of astrometric and radial velocity exoplanet observations. Determining exoplanet mass with astrometric snapshots. *A&A* 494, 769–774.
- van de Kamp, P. (1977). Perspective secular changes in stellar proper motion, radial velocity and parallax. *Vistas in Astronomy* 21, 289–310.
- van Leeuwen, F. (Ed.) (2007). *Hipparcos, the New Reduction of the Raw Data*, Volume 350 of *Astrophysics and Space Science Library*. Springer Verlag.
- Vigan, A., C. Moutou, M. Langlois, F. Allard, A. Boccaletti, M. Carbillet, D. Mouillet, and I. Smith (2010). Photometric characterization of exoplanets using angular and spectral differential imaging. *Monthly Notices of the Royal Astronomical Society* 407(1), 71–82.
- Vogt, S. S., R. P. Butler, G. W. Marcy, D. A. Fischer, G. W. Henry, G. Laughlin, J. T. Wright, and J. A. Johnson (2005, October). Five New Multicomponent Planetary Systems. *ApJ* 632, 638–658.
- Walker, G. A. H., A. R. Walker, A. W. Irwin, A. M. Larson, S. L. S. Yang, and D. C. Richardson (1995, August). A search for Jupiter-mass companions to nearby stars. *Icarus* 116, 359–375.
- Weinberg, M. D. (2012). Computing the Bayes Factor from a Markov Chain Monte Carlo Simulation of the Posterior Distribution. *Bayesian Analysis* 7(3), 737 – 770.
- Weinberg, M. D. and J. E. B. Moss (2011, August). The umass bayesian inference engine. <http://www.astro.umass.edu/BIE/manual.pdf>.
- Wilson, O. C. (1978, December). Chromospheric variations in main-sequence stars. *ApJ* 226, 379–396.
- Wolpert, R. L. (2002, August). Stable limit laws for marginal probabilities from mcmc streams: Acceleration of convergence. <http://ftp.isds.duke.edu/WorkingPapers/02-22.pdf>.

- Wolszczan, A. and D. A. Frail (1992, January). A planetary system around the millisecond pulsar PSR1257 + 12. *Nature* 355, 145–147.
- Zechmeister, M. (2010, November). *Precision Radial Velocity Surveys for Exoplanets*. Ph.D. thesis.
- Zechmeister, M. and M. Kürster (2009, March). The generalised Lomb-Scargle periodogram. A new formalism for the floating-mean and Keplerian periodograms. *A&A* 496, 577–584.
- Zucker, S. and T. Mazeh (2001, November). Analysis of the Hipparcos Observations of the Extrasolar Planets and the Brown Dwarf Candidates. *ApJ* 562, 549–557.

Acknowledgements

First and foremost, I wish to thank my primary advisor **Prof. Thomas Henning**, who gave me the opportunity to conduct my PhD thesis at the Max Planck Institute for Astronomy (MPIA). This work would not have been possible without his enduring support, his patience, and his scientific guidance.

I owe many thanks to my co-advisor **Dr. Ralf Launhardt** for his constant readiness to provide valuable suggestions and constructive criticism both. His support has been indispensable throughout this work.

Prof. Andreas Quirrenbach is gratefully appreciated for his readiness to act as second co-advisor and referee for my thesis.

The members of my PhD Advisory Committee, **Dr. Eva Schinnerer**, **Dr. Wolfgang Brandner**, **Dr. Coryn Bailer-Jones**, **PD Dr. Hubert Klahr**, and **Dr. Tom Herbst** are acknowledged for providing helpful input on the scope and progress of this work.

I owe special thanks to **Prof. Michael Perryman**, who encouraged me in my work and gave valuable hints on self-management.

I am grateful to **Dr. Mathias Zechmeister** for sharing essential radial-velocity data, and for several fruitful discussions on frequentist data analysis. **Dr. Johnny Setiawan** is thanked for sharing a number of data sets that have been useful in testing BASE.

Dr. Sabine Reffert is appreciated for her advice on the treatment of HIPPARCOS data and other aspects of astrometry.

Prof. David W. Hogg and **Dr. René Andrae** are both thanked for many fruitful and interesting discussions about Bayesian statistics and other aspects of data analysis.

Prof. Edward O. Wiley and **Francisco Rica Romero** are regarded for patiently testing BASE and for many useful hints and ideas.

Dr. Dading Nugroho, **Gabriele Maier**, **Dr. Natalia Kudryavtseva**, and **Dr. René Andrae** are appreciated for sharing their time in our MPIA office.

I thank my parents **Brigitte Schulze-Hartung** and **Klaus Schulze-Hartung** for their patience and support during all this time.

I appreciate my good and reliable friends **Bernhard Wüste** and **Dr. Manfred Bohn**. Manfred is thanked for sharing his L^AT_EX style template.

The deepest gratitude I owe to **Swetlana Stresler** for being there and giving me strong support, tolerance and understanding throughout a very demanding phase of my life.

