Multi-Periodic Variability in Low Mass X-ray Binaries

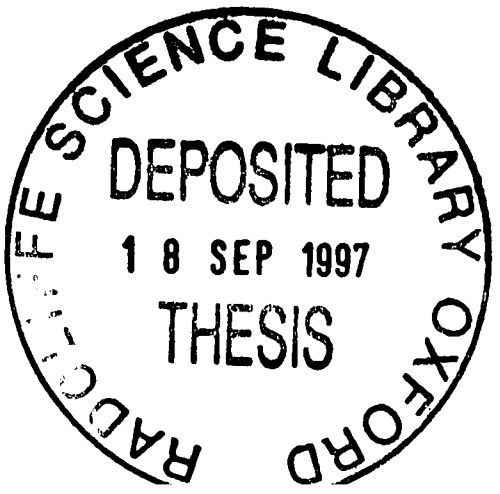
by

Andrew Martin

Exeter College, Oxford

Michaelmas Term 1995

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy



Declaration

I declare that this thesis is my own work and that no part of it has been accepted, or is currently being submitted, for any degree or diploma or certificate or other qualification in this University or elsewhere.

The I-band photometry of J0422+32 in Chapter 4 was published in the *Monthly Notices of the Royal Astronomical Society* (MNRAS) **274**, 1995; as Martin A.C., Charles P.A., Wagner R.M., Casares J., Henden A.A., Pavlenko E.P. A paper based on the analysis of the six hour periodicity in V404 Cyg in Chapter 3 has also been submitted to MNRAS; as Pavlenko E.P., Martin A.C., Casares J., Charles P.A., Ketsaris N.A (1996). Also, the photometry of V2107 Oph described in Chapter 5 was published in MNRAS **274**, 1995; as Martin A.C., Casares J., Charles P.A., van der Hooft F., van Paradijs, J.

All of the observations, data reduction and subsequent analysis were undertaken by the author, with the exception of Chapter 6. These data were combined from various sources (see references in Chapter 6), and analyzed using the new Bayesian techniques outlined in Chapter 2.

Multi-Periodic Variability in Low Mass X-ray Binaries

Andrew Martin

Exeter College, Oxford

Michaelmas Term 1995

A thesis submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Abstract

An introduction is given to the fields of low mass X-ray binaries (LMXBs), soft X-ray transients (SXTs) and related classes of objects. Aspects such as evolution, variability and populations are reviewed.

The methods of time series analysis are listed and the new method of Bayesian spectrum analysis and parameter estimation is outlined. In particular the methods required to analyze multi-periodic variability in LMXB light curves are described, as used in subsequent chapters. Artificial datasets are used to examine the robustness of the Bayesian frequency analysis methods. An *ad hoc* method of deconvolving the effects of data sampling is then detailed.

The optical variability of V404 Cyg is analyzed using the methods of the preceding chapter. Orbital and short period variations are examined in several wavebands and the implications for this black-hole candidate discussed. Models for the short period variation are compared to the observed behaviour, to locate the origin of this effect. Modelling of ellipsoidal variability in V404 Cyg, J0422+32 and several faint X-ray transients allows the determination of the inclination for these systems. Other important physical parameters are also considered and compared to previous values cited in the literature.

The globular cluster LMXB AC211 is then examined to try to shed some light on the complex orbital variability observed. Several photometric and spectroscopic datasets are combined to provide an overall, relative probability for the two possible periods.

The distributions of neutron star and black-hole masses are analyzed to provide constraints on models of the neutron star equation of state. Inclinations obtained from ellipsoidal analysis and the excess noise seen in SXT light curves, are discussed.

Acknowledgements

First of all I wish to thank my supervisor, Phil Charles, for his help and guidance over the past three years, and for his careful reading of this thesis. Mention must also go to the various telescope operators and support staff at the South African Astronomical Observatory, European Southern Observatory, Anglo-Australian Observatory and the Isaac Newton Group of telescopes on La Palma, operated by the Royal Greenwich Observatory. In particular Emilios, who showed me the ropes during the observing run at the JKT in La Palma. Thanks also to Mark Wagner, at the Lowell observatory in Arizona, for his hospitality and generosity during my stay at Flagstaff. I am also indebted to the friendliness displayed by the many people I met around the world during the past three years.

The collaboration with the TRIFFID group at University College, Galway provided the basis for the work on AC211; their image sharpening research provides an essential service to ground-based astronomy. Larry Bretthorst gave me advice and insights into Bayesian analysis, which was the basis for Chapter 2, and was always quick to reply to my questions.

Thanks to everyone all the regulars who played in the departmental football team: Bart, Brad, Dave, John, Phil L., Steve_R and anyone else I have forgotten, for providing me with a good excuse for leaving early on Fridays. Thanks to Ian and Simon for printing out the crosswords and generally chatting. Also, thanks Jorge for the help and background on V404 Cyg and other areas, and Tariq with whom I had many useful and informative discussions. Christian helped with the significance testing, discussed time series analysis with me and provided me with many ideas that are included in this thesis. Mention should also be made of my other office mates over the last two years Andy B. and Anthony L. for occasionally laughing at my jokes. I should also like to thank everyone else in the Astrophysics Department here at Oxford for creating a warm and friendly atmosphere during my time here. My friends outside the department: Gavin, Steve, Dave, Sharad and Julian I thank for all the fun I have had over the last three years.

I acknowledge the support of a Particle Physics and Astronomy Research Council (PPARC) studentship during the course of this thesis. I am also grateful to my parents for further financial assistance and worrying about me all the time. Finally, thank you to Juliet, for the love and encouragement given to me over the last year.

"I saw Eternity the other night
Like a great Ring of pure and endless light,
All calm as it was bright;
And round beneath it, Time, in hours, days, years,
Driven by the spheres,
Like a vast shadow moved, in which the world
And all her train were hurled."

Henry Vaughan, "The World", 1-7.

Contents

1 Introduction **1**

1.1 Low Mass X-ray Binaries 1

1.1.1 Relation to HMXBs 5

1.1.2 Optical and X-ray Variability of the Persistently Bright LMXBs 5

1.1.3 Populations 9

1.1.4 Formation and Evolution of LMXBs 14

1.2 Soft X-ray Transients 16

1.2.1 Outburst 16

1.2.2 Quiescence 22

1.3 Summary 27

2 Bayesian Spectrum Analysis **32**

2.1 Methods of Frequency Analysis 32

2.1.1 Discrete Fourier Transform 32

2.1.2 Lomb-Scargle Periodogram 34

2.1.3 Phase Dispersion Minimization 34

2.1.4 CLEAN 35

2.1.5 Other methods 36

2.2 The Bayesian Method of Periodogram Analysis 36

2.2.1 General Outline 36

2.2.2 A Simple Bayesian Model 37

2.2.3 Examples for the Single Sinusoid Model 40

2.2.4 Examples for the Two Frequency Model 44

2.2.5 The General Model 47

2.2.6 Multiple Frequencies 52

2.2.7 Model Comparison 53

2.2.8 Parameter Estimation and Resolving Power 54
2.2.9 An Ellipsoidal Model 55
2.2.10 Example: Uniformly Sampled Ellipsoidal Dataset 58
2.3 The Problem of Aliasing 60
2.3.1 Introduction 60
2.3.2 Best Fit Periodogram Algorithm 64
2.3.3 Estimating the Signal Amplitude 65
2.3.4 Example – Unevenly Sample Data 66
2.3.5 Violating the Assumptions 67
2.3.6 Summary 70

3 Orbital and Quasi-Periodic Optical Variations in the Black-Hole X-ray Binary V404 Cyg 72
3.1 Summary – Optical Photometry 72
3.1.1 Introduction 73
3.1.2 Observations 74
3.1.3 Ellipsoidal Modelling of the R-band Light Curve 74
3.1.4 Colour Dependent Variations in V404 Cyg 78
3.1.5 Bayesian Frequency Analysis 85
3.2 The 6-hour Quasi-Periodicity 90
3.2.1 Changes in the six hour period. 90
3.2.2 Modelling the Observed Periodogram 92
3.3 I-band Photometry 93
3.3.1 Introduction 93
3.3.2 Period Analysis 95
3.3.3 The Amplitude of the Six Hour Modulation 95
3.3.4 Random Flickering in the Light Curve 98
3.4 Discussion – The Origin of the Six Hour Period 98
3.4.1 The Spectroscopic Phase of the 6 Hour Period 103

4 Outburst, Decline and Quiescent Photometry of J0422+32 107
4.1 Near Quiescent Photometry 107
4.1.1 Summary 107
4.1.2 Introduction 107

4.1.3 Observations . . . 109
4.1.4 Results . . . 111
4.1.5 Significance Tests . . . 112
4.1.6 Discussion – An Ellipsoidal Effect? . . . 120
4.1.7 Relation to the 5.1 Hour Period . . . 122
4.1.8 Nature of the Secondary Star . . . 123
4.1.9 Comparison with other X-ray Binaries . . . 124
4.2 Outburst photometry . . . 124
4.2.1 Introduction . . . 124
4.3 Quiescent photometry . . . 129
4.3.1 Observations . . . 129
4.3.2 Quiescent variability . . . 130
4.3.3 Intensity Dependent Changes in the Modulation of J0422+32 . . . 135
4.3.4 Analysis of the 2.55 (5.1) Hour Period . . . 136
4.3.5 The Origin of the 16.18 hour Period . . . 136
4.4 Discussion . . . 138

5 Optical Photometry of Three Faint Soft X-ray Transients (V2107 Oph, BW Cir, KY TrA). 142

5.1 V2107 Oph . . . 142
5.1.1 Summary . . . 142
5.1.2 Introduction . . . 142
5.1.3 Observations . . . 143
5.1.4 Orbital Period Search . . . 144
5.1.5 Ellipsoidal Light Curve and System Parameters . . . 149
5.1.6 Conclusions . . . 155
5.2 BW Cir (Nova Cir 1987, GS 1354–64) . . . 155
5.2.1 Introduction . . . 155
5.2.2 Observations . . . 156
5.2.3 Period Analysis of the Light Curve . . . 159
5.2.4 The Inclination of BW Cir . . . 161
5.3 KY TrA (TrA X-1) . . . 163
5.3.1 Introduction . . . 163
5.3.2 Observations . . . 164

5.3.3 The Reddening and Nature of the Secondary Star 164

6 AC211 – an LMXB in the Core of M15 **171**

6.1 Introduction 171

6.2 TRIFFID Photometry 172

6.2.1 Introduction 172

6.2.2 Observations 173

6.2.3 Period Analysis 176

6.3 U-band Photometry 181

6.3.1 Period Analysis 182

6.4 Spectroscopy 184

6.5 Discussion 192

7 Conclusions **196**

Chapter 1

Introduction

1.1 Low Mass X-ray Binaries

The low mass X-ray binaries (LMXBs), are semi-detached systems consisting of a late-type star and an accreting compact object (either a neutron star or black hole). The compact object is surrounded by an accretion disc fed by mass loss from the secondary star ($\sim 10^{-8}$ – 10^{-10} $M_\odot$/yr; McClintock, 1986). Due to the energy released as this matter falls into the potential well of the compact object these systems are extremely bright in X-rays (typically $L_X \sim 10^{36}$ – 10^{38} ergs s^{-1}), with X-ray to optical luminosity ratios of ~ 100. This ratio is useful as a measure of the amount of optical radiation produced by reprocessing by the disc of X-rays from the central source. Also, there is a wide range of variability due to accretion processes, geometrical effects and X-ray heating. The separation of the two components is given by the equation below, derived from Kepler's third law and also from the assumption that the secondary fills its Roche lobe:-

$$a\,(R_\odot) \;=\; 2.12(M_1 + M_2)^{1/3}$$

If typical system parameters of 1.4 $M_\odot$ for the mass of a neutron star compact object and 1 $M_\odot$ for the secondary are assumed, the separation obtained is 3 $R_\odot$, i.e. LMXBs are extremely compact. The large X-ray emission is produced by the accreted matter falling into the large potential well of the compact object. Neutron star radii are typically of the order

of 10 km, and a black hole's event horizon is given by $2GM/c^2$, which, for a stellar mass black hole is $\sim$ 5 km. Thus, black holes convert accreted matter into energy more efficiently (approximately 40 % of the accreted rest mass is converted for rotating black holes and 15% for neutron stars). The orbital period of these systems is typically a few hours, although a few periods of the order of tens of days have been measured.

The rotation of the two binary components, combined with the rotation of the system gives rise to the Roche potential (see figure 1.1). The final stable surface for each star is called its Roche lobe, and any matter that flows over one of these lobes will be lost from the binary or accreted by the other companion. The form of the Roche potential means that it is much easier for matter to flow from one star to the other across the L_1 point, where the two Roche lobes touch. Assuming that the material is pushed out of the secondary by pressure forces, then the radial velocity of the gas stream with respect to the primary will be small ($<$ 10 km s^{-1}), and the transverse velocity of the stream will be given by its angular momentum at that point ($\sim$ 100 km s^{-1} for a typical LMXB). The material from the gas stream will free fall from this point on a ballistic trajectory until it collides with the outer edge of the accretion disc at the bright spot. Initially the gas stream would form a disc at a radius with the same angular momentum as the material flowing over the L_1 point (Frank, King and Raine, 1992). This *circularization radius* is equal to:-

$$R_{circ}/a = (1+q)\ (0.500 - 0.227\ log_{10}[q])^4$$

Where q is the mass ratio, which is equal to M_2/M_1. As the disc evolves, material flows inwards and angular momentum is transported outwards, meaning that the true disc radius can extend beyond the circularization radius. The exact size of the disc will vary with time, depending on the viscosity, accretion rate, temperature and other physical parameters. The radius of the disc is important in the disc outburst mechanisms, which will be discussed later in this chapter.

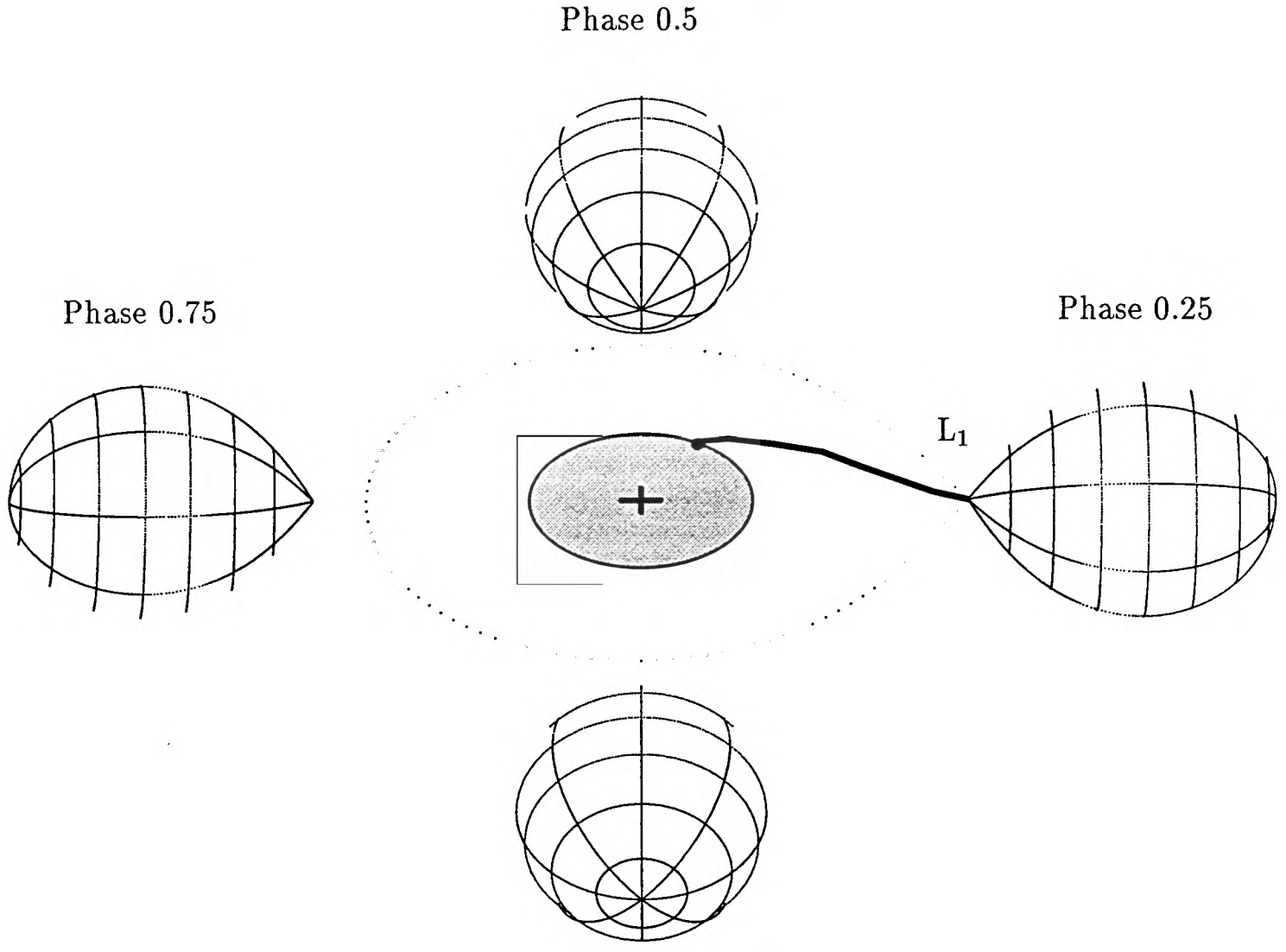


Figure 1.1: 3-Dimensional plot of a typical LMXB or SXT system with a mass ratio $q = 5$ and an inclination of 60°. The rotation of the secondary, which fills its Roche lobe, is shown at several phases and the primary Roche lobe is displayed as a dotted curve. The disc radius was taken to be equal to the circularization radius, with the location of the primary indicated by a cross at its centre. For clarity the expected flaring of the disc ($< 20°$) is not represented.

There are several mechanisms for driving the mass transfer in close binaries. For short period systems (P_{orb} < 3 hours) gravitational radiation is strong enough to drain energy from the binary and to drive the two components together. Longer period systems may be driven by magnetic braking of the secondary, if it is convective, causing the binary separation to decrease, or evolution may cause the secondary to outgrow its Roche lobe. This final mechanism will occur on the nuclear timescale (determined by hydrogen shell burning), and so it is unlikely that this can be the cause of any short term changes (as postulated by the mass transfer burst model, see section 1.2.1).

The mass ratio in LMXBs must be smaller than $\sim$ 1 for mass transfer to be stable (this figure is not entirely certain as some mass and therefore angular momentum may be lost from the binary during mass transfer). If it is larger than this value, then the secondary Roche lobe shrinks as the secondary loses mass and a runaway process ensues, unless the star can reduce its radius quickly enough. This will only stop once the mass ratio decreases below the above limit. Thus, all of the systems that are studied here have low mass companions.

The soft X-ray transients (SXTs) are a sub-class of the LMXBs which remain in a dormant, quiescent phase for most of the time, but produce a huge X-ray outburst every 10–50 years becoming several orders of magnitude brighter. The object then returns to its quiescent state on a timescale of several months (in quiescence, SXTs typically have an X-ray luminosity in the range 10^{32} – 10^{34} ergs s^{-1} , see van Paradijs et al., 1987; Verbunt et al., 1994). The SXTs are important because in quiescence most of the light comes from the secondary star, and so it is possible to determine the orbital period and model the (ellipsoidal) brightness variations of the system with little influence of contaminating light from the accretion disc. Combined with a radial velocity curve for the secondary star it becomes possible in principle to completely solve for the binary stellar parameters (see e.g. Charles, 1995).

1.1.1 Relation to HMXBs

The high mass X-ray binaries are a related class of objects which have a companion star with a mass > 10 $M_\odot$ orbiting a compact primary (some of which are thought to be black holes, e.g. Cyg X-1). Two broad groups are distinguished: the 'standard' systems, with giant or super-giant companions which nearly fill their Roche lobes and the Be-X-ray binaries, which have rapidly rotating B-emission stars of luminosity class III-IV. The 'standard' systems such as SMC X-1 exhibit ellipsoidal light variations and X-ray eclipses, and observations indicate that they come from O-type progenitors. The Be-star companions are deep within their Roche lobes, indicated by the long periods of these systems (> 15 days) and the absence of ellipsoidal variations and X-ray eclipses. These systems are often driven by wind-fed accretion, or by means of a ring or disc of material.

1.1.2 Optical and X-ray Variability of the Persistently Bright LMXBs

Type I bursts

LMXBs undergo a large range of variability on all timescales, in both X-ray and optical bands. Type I bursts (see the review by Lewin et al., 1988) occur as a sudden rise in X-rays, taking place over a few seconds, up to a level several times above the persistent X-ray flux, this then decays back to its original level after a few minutes. Blackbody spectral analysis of the bursts show that the apparent radii stays roughly constant at ~ 10 km, a typical neutron star radius. The bursts have longer tails in the lower energy bands, which indicates cooling after the onset of the burst. This implies that accretion onto the surface of the neutron star is the cause and that the burst arises from a thermonuclear flash on the surface, of the accreted material. This model is now well established and fits the observed data on type I bursts extremely well, so that occurrence of type I bursts is used as an unambiguous indicator that the compact object in a system is a neutron star. Unfortunately, no such simple observational signature exists to prove that a compact object is a black hole.

Type II bursts

Type II bursts have been seen in only one source, the rapid burster MXB 1730−335, which bursts every few minutes, each burst lasting a few seconds (see Hoffman et. al., 1978). The blackbody models of these bursts show a roughly constant temperature with decreasing radius, and are thought to be due to an accretion instability. The behaviour of these bursts goes through three phases. Firstly, sharply peaked bursts occur with longer and longer durations, these become trapezoidally shaped, and then finally their duration decreases again. The height of the trapezoidal bursts is inversely proportional to their length, indicating a maximum possible energy for each burst, perhaps due to the maximum size of some sort of reservoir of material supplying the energy for the bursts. The time averaged energy of the bursts is constant, indicating that the mechanism for the bursts is fed by a constant accretion flow. The height of the shorter bursts is proportional to the time to the next burst, which could be related to the way the reservoir is filled. A burst of a certain size empties the reservoir by an amount proportional to its size and, if the reservoir is fed by a constant accretion flow, the time to the next burst will be the time required to fill up the reservoir again. Other modes of behaviour have also been recently discovered, such as small quasi-periodic bursts preceding a type II burst (Dotani et al., 1990), and it has also been observed that the decay part of type II bursts show regular structure (Tawara et al., 1985). These phenomena have yet to be explained, but indicate that accretion processes in the LMXBs are not as uniform as simple models suggest.

Quasi-Periodic Oscillations

The phenomenon of quasi-periodic oscillations (QPO) has been seen in several sources, where an almost periodic variation of 1–10 % in the X-ray flux occurs at a frequency of 5–50 Hz (see the review of van der Klis, 1989). The X-ray power spectra of these sources show a broad peak superimposed on a general increase in power toward low frequencies, which is known as low frequency noise (LFN). These seem to correlate in intensity with QPO. Power spectra diagrams are widely used in QPO studies to analyze the frequency shifts in the various

components associated with QPO.

Sources which undergo QPO can be separated into two classes: the Z sources and the atoll sources. This distinction arises from the shape these sources trace out in a diagram of X-ray hardness versus intensity (or sometimes in X-ray colour-colour diagrams). The Z sources, such as Cyg X-2 and Sco X-1, have three branches of QPO. The horizontal branch, the normal branch and the flaring branch. The frequency of QPO on the horizontal branch are intensity dependent, unlike the normal branch where the frequency remains low throughout. Also, LFN appears on the horizontal but not the normal branch. The atoll sources show 'island'-like structures in their X-ray colour-colour diagrams and are generally fainter than the Z sources.

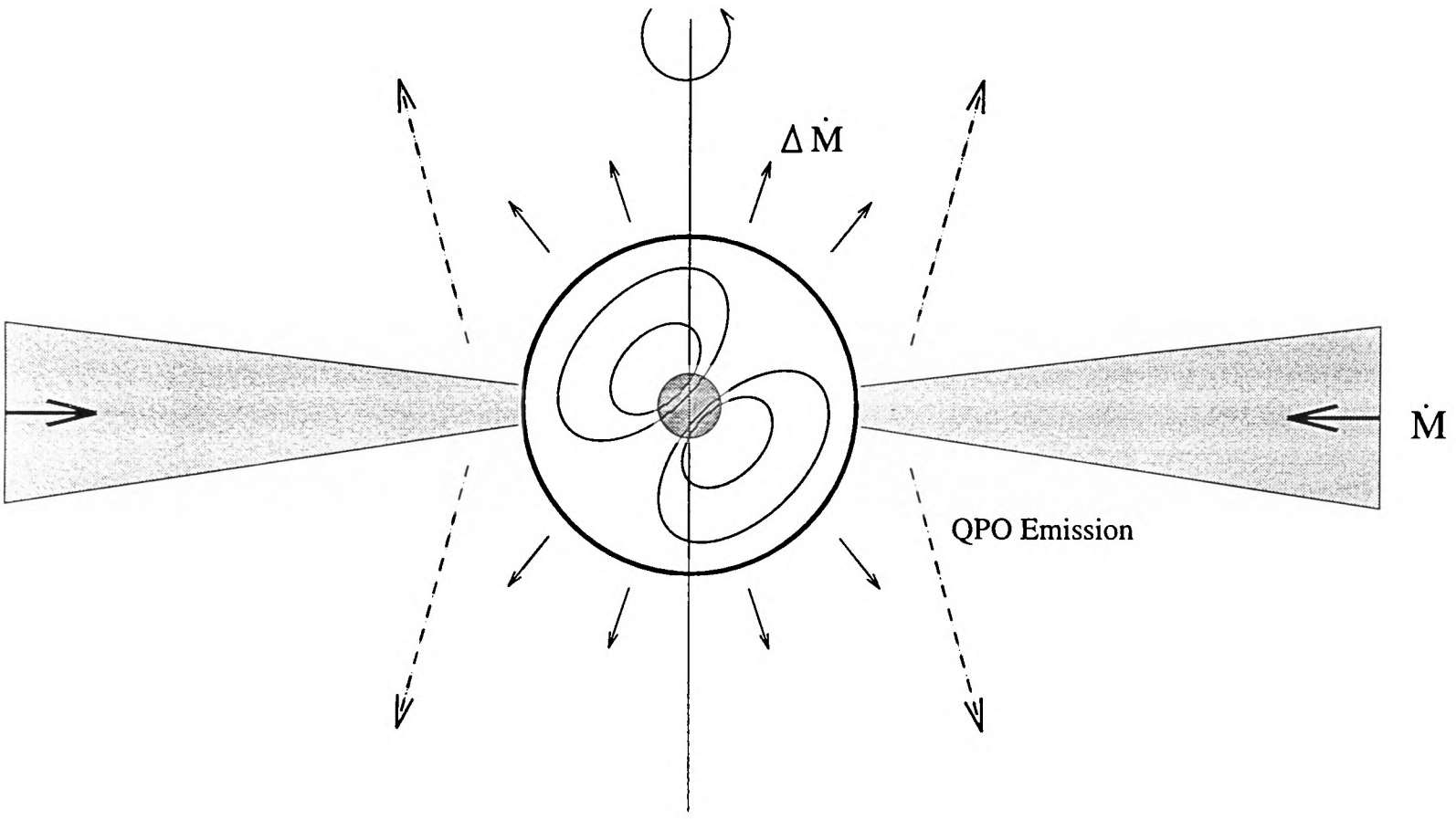


Figure 1.2: Beat frequency model of QPO emission for LMXB systems as described in the main text. For the typical QPO frequencies observed ($\sim$ 10 Hz) the neutron star rotation frequency that is predicted is $\approx$ 100 Hz (spin period 10 ms).

One model proposed to explain this modulation is the beat frequency model, which models the interaction of the magnetosphere and the inner accretion disc, as shown in figure 1.2. The frequency of the QPO is equal to the beat frequency of the rotation of the neutron star and the Keplerian frequency of the disc at the emission region, $f_{QPO} = \mid f_k - f_* \mid$. The emission region is thought to be near to the Alfvén radius, R_A, where the disc begins to follow the magnetic field lines of the neutron star.

Period Distribution

The period distribution of LMXBs tails off at long and short periods and also possibly has a gap analagous to that found in cataclysmic variables (CVs), see White (1986). Recently, several CVs have been found inside the gap (e.g. Shafter et al., 1995), but the gap may still be significant. The gap in the LMXB period distribution is either due to pure coincidence that no systems have periods within the gap, the superposition of two distributions or mass transfer switching off in the gap. In the CV distribution some objects near the gap have epochs of very low mass transfer, which could be an indication that mass transfer is beginning to switch off in those systems. But statistical studies of CVs with different mass transfer rates show that low mass transfer systems do not preferentially clump near to the gap. The most favoured explanation of the gap for CVs, which probably applies to LMXBs too, is that at a period $\sim$ 3 hours the secondary becomes fully convective (at a secondary mass of 0.3 $M_\odot$ and a spectral type of $\sim$ M4). This could then reduce the effect of magnetic braking, which would mean that mass transfer stops until shorter periods, when gravitational radiation becomes efficient and shrinks the Roche lobe onto the surface of the secondary ($P_{orb} \approx 2.3$ hours).

The main difference between the period distributions of the CVs and LMXBs appears to be the lack of LMXBs with periods between 80 mins and 2 hours. At a period of two hours gravitational radiation has become effective and shrunk the Roche lobe to the secondary radius thereby re-initializing mass transfer. Further mass loss from the secondary causes its central temperature and pressure to drop such that the core becomes degenerate and the star expands leading to an increase in binary period. This is expected to occur when the

mass transfer timescale is of the order of the Kelvin-Helmholtz timescale, which happens at a period of 81 minutes (Paczyński, 1981) and a secondary mass of about 0.12 $M_{\odot}$ when it can no longer remain on the main sequence. Shorter period systems can be explained by a common envelope spiral-in phase when the neutron star is left with the core of the evolved secondary as a companion star. In this way the secondary may become a helium burning star ($P_{orb} \sim 10$ mins) or a white dwarf ($P_{orb} \sim 1$ min). The rarity of this type of system suggests that such an evolutionary scenario is unlikely or short-lived.

The long period tail of the LMXB distribution could be partly due to selection effects (see Watson 1985). For example, EXOSAT observations typically lasted for around 6–12 hours and so were biased toward the discovery of shorter period systems. Also, the fact that mass transfer becomes unstable when q is larger than 1 sets an approximate upper limit on the period of the system, as the secondary cannot be much more massive than the primary for stable mass transfer to occur. In wider binary systems, the secondary Roche lobe is large and so the secondary would have to be evolved in order to fill it and for mass transfer to begin. For late-type stars, around K spectral type, the expected timescale for main sequence turn-off is 10^{10} years (in the absence of any external effects such as forced co-rotation and X-ray heating). Thus, many secondary stars will not have evolved into sub-giants as yet, and so long period binaries will be scarce.

1.1.3 Populations

The galactic distribution of LMXBs is concentrated towards the galactic centre and has a fairly wide latitude distribution ($< |b^{II}| > = 9.2°$; van Paradijs, 1983). This is consistent with these sources being related to older (population II and old disc population) objects. The number of sources found near the centre of our galaxy has given rise to the classification of these objects as bulge sources, many of the other LMXBs are found in globular clusters. The bulge sources were probably formed by evolution of one or more components in the binary, whereas the higher densities found in globular clusters means that tidal capture is a viable mechanism.

The bulge sources are not seen to pulsate, which combined with the large number that undergo type I bursts, seems to imply that they possess only weak magnetic fields. This means that the accretion disc will extend all the way down to the surface of the compact object, and the X-ray emission will come from material in the disc spiraling inward. The bursting bulge sources appear to have luminosities of $\sim 0.1\ L_{Edd}$, whereas those that do not burst appear to have much higher luminosities. This suggests that high accretion rates tend to suppress type I bursts. Globular cluster sources have an average maximum luminosity that is equal to galactic LMXBs and most of them emit X-ray bursts. For a further description of the globular cluster LMXBs and their formation see section 1.1.3.

A further classification that can be applied to LMXBs is whether or not they undergo type I bursts. As has already been described, these are due to thermonuclear burning on the surface of the compact object and so are considered to be an unmistakable signature of a neutron star primary. The physical difference between sources that burst and those that do not, is likely to be the mass transfer rate from the secondary.

The average properties of LMXBs such as colour index, the ratio of X-ray to optical luminosity and their absolute visual magnitude all seem to lie in a small range of values (see van Paradijs, 1981a; van Paradijs and McClintock, 1994). The (B−V) and (U−B) colours, corrected for interstellar reddening, are equal to 0.03 ± 0.25^m and -0.95 ± 0.25^m, respectively. These colours are similar to those of a flat energy distribution (F_ν = constant), which for the Johnson calibration of the UBV system are equal to: B−V = 0.2^m and U−B = -1.0^m. Taking into account the scatter induced by the error on the reddening corrections applied, which are probably only accurate to $\approx 0.1^m$, this implies that the colours of LMXBs are tightly clustered around the above mean values.

The difference between the optical B-band magnitude and X-ray magnitude m_x (= -2.5 $\log_{10} F_X$ (μJy)) was defined as a convenient measure of the relative strengths of optical and X-ray emission strengths in LMXBs by van Paradijs (1981a). This was found to be equal to B $- m_x = 21.5 \pm 1.1^m$ (excluding 3 systems which showed evidence of an eclipse of the optically

emitting region or extreme X-ray heating). Part of the spread of this distribution is again due to the error on the interstellar extinction in B (0.4^m), and also due to random orientations of LMXBs in space, which can account for a further spread of $\pm$ 0.7^m. This suggests that the scatter is mostly due to observational effects and that the actual LMXB distribution is narrow. The average ratio of X-ray (2–11 keV) to optical luminosities (3000–7000 Å), corresponding to the above magnitude difference, was found to be equal to $\sim$ 350.

To estimate the average absolute visual magnitude of LMXBs, several assumptions were made. The first was that the peak X-ray luminosity is the same for all of the sources included. Only the X-ray bursters were included in the sample as the peak luminosity of a given burst is believed to be Eddington limited. This assumption is supported by the luminosities estimated for globular cluster sources with independent distance estimates (e.g. using colour-magnitude diagrams and reddening studies). These vary by a small enough amount to make the standard meaningful (van Paradijs, 1981b). This information was then used to obtain distance estimates for optically identified burst sources. The average value of the absolute visual magnitude was thus $\overline{M_v} = 1.2 \pm 1.0$. The main variation in the distribution of M_v was due to the inclusion of Aql X-1 ($M_v = -1.1$), an SXT which undergoes 'mini' outbursts on a timescale of 1 year making it difficult to estimate the mean optical flux of this object.

All LMXBs and CVs appear to have late-type secondaries, including those that are sub-giants. This is coincidental with the fact that the rotation speed of single stars drops sharply beyond the spectral type F5, due to later type stars having efficient, magnetically locked winds which remove angular momentum from the star. This effect can therefore be explained if only those systems with later spectral types manage to shed angular momentum rapidly enough to reach their Roche lobe on a timescale less than the current age of the universe ($< 10^{10}$ years). Another possibility is that A to F stars undergo unstable mass transfer and become either a G or K star, which are expected to transfer mass in a stable manner. An evolutionary scenario, first suggested by van den Heuvel (1983), begins with a binary having an extreme mass ratio. The higher mass star evolves more quickly, engulfing its companion, and may eventually undergo a supernova explosion. This leads to the formation of a neutron

star – main sequence binary and, after subsequent angular momentum loss, the low mass star will fill its Roche lobe causing the onset of mass transfer, thus creating an LMXB.

LMXBs in Globular Clusters

The formation of LMXBs in globular clusters can provide insights into the evolution and structure of clusters (Grindlay, 1987), and conversely the physics of the host cluster can reveal a lot about the formation and evolution of LMXBs. Due to equipartition of energy it is possible to state that statistically, more massive stars will lie nearer to the centre of a globular cluster. Thus, the positions of all the cluster sources known were calculated relative to the dynamical centres and, using the core radii, the typical binary mass of globular cluster LMXBs was estimated. This has shown that the neutron star's companion in an LMXB is indeed a low mass star (Grindlay, 1984), as the total masses were found to be in the range 0.9–1.9 $M_\odot$ (with 90% confidence). For a neutron star primary with $M_1 \approx 1.5$ $M_\odot$, this implies that the secondary is a low mass object. This is consistent with the fact that the stars in globular clusters are $\geq 10^{10}$ yrs old, and so must have masses ≤ 0.8 $M_\odot$ in order not to have already evolved off the main sequence.

Compact binaries can be formed by at least two methods in globular clusters – three-body interactions or tidal capture (see Cohn, 1987). In the tidal capture mechanism a bulge is raised on the captured star which leads to heating and dissipation of energy and therefore enhances the probability of capture. Three-body interactions can occur either through the simultaneous collision of three stars (the probability of which is small), or when a neutron star interacts with an existing binary. The former case either leads to a triple forming or a binary: the latter results in either the neutron star being exchanged for a binary member, or in the neutron star being ejected. As tidal capture is only a two-body process it is more important, unless there exist very high stellar densities where three body interactions have a higher probability (e.g. in cluster core collapse).

The maximum closest approach distance found for tidal capture to occur was $\sim 3\ R_\odot$ (Fabian et al., 1975), this was recently modified but only by a small amount. The cross-section depends linearly on the closest approach distance and so at least one third of these encounters will be direct collisions which may not produce a binary system.

The tidal capture mechanism operates by depositing the kinetic energy that each star possesses into non-radial oscillations of the other. If enough energy is absorbed the two stars become bound and subsequent periastron passages cause the newly formed binary orbit to circularize. The modes that are excited depend on the structure of the envelope of the star, but not explicitly its mass or radius. It has been shown that for two particular stars the cross-section $\sigma(v)$ or impact parameter R_0 can be expressed as a function of the minimum closest approach distance and their relative velocity (see the equations below). For a more detailed description of this proof see Press and Teukolsky (1977).

$$\sigma(v) = \pi R_0(v)^2$$

$$R_0(v) = (2GM_T R_{min}/v^2)^{1/2}$$

In direct collisions between a main sequence star and a compact object it has been suggested that the main sequence star will be totally disrupted, forming a disc around the compact object. In collisions between a giant star and a compact star it is thought that the degenerate core of the giant and the compact star will orbit each other in the common envelope of the giant (these are known as Thorne-Zytkow objects, see Thorne and Zytkow, 1977). Tidal friction will eventually remove the envelope and reduce the orbital separation until an ultra-compact binary is formed.

The formation of compact binaries and their subsequent interaction with other cluster stars is thought to be a mechanism that can halt core collapse in globular clusters. Thus, it is possible that LMXBs are a signature of post core collapse globular clusters. It has been suggested that many X-ray burst sources were formed by tidal capture in globular clusters

which have since been disrupted (Grindlay, 1984) either through interactions with the galactic gravitational field or giant molecular clouds. Cowley et al. (1987) found that bulge X-ray sources have a similar distribution and velocity dispersion to the bulge globular clusters, which supports this idea. But this cannot explain how all of the LMXBs throughout the galaxy were formed.

1.1.4 Formation and Evolution of LMXBs

The formation of LMXBs in globular clusters was described above, but for the rest of the LMXB population different formation scenarios must be investigated (Verbunt, 1993). Outside of globular clusters the density of stars is much lower and so the probability of a neutron star interacting with other stars to form a binary is small. Binaries consisting of non-degenerate stars are common and could lead to the formation of LMXBs when one of the binary stars undergoes a supernova (Paczyński, 1971). The very fact that LMXBs are seen at all, albeit in small numbers, shows that a binary can sometimes survive the supernova. In HMXBs it is the low mass component which undergoes supernova and so it is fairly easy to see why the binary is not disrupted, but in LMXBs it is harder to see how the binary can survive the explosion.

To see how the orbit of a binary will be affected by the supernova a simple estimate can be derived for the parameters of the resulting orbit. It is assumed that the orbit prior to the supernova was circular and that the only change that occurs after the explosion is in the mass of the exploding star. Thus the periastron distance between the two stars is equal to the previous binary separation, and the velocity at this point is equal to the velocity of the pre-explosion orbit. The eccentricity e can then be derived as well as the velocity v_s of the centre of gravity after the supernova, giving the following two equations:-

$$e = \frac{\Delta M}{M_1 + M_2 - \Delta M}$$

$$v_s = \frac{M_2 v_2 - (M_1 - \Delta M) v_1}{M_1 + M_2 - \Delta M} = e\, v_1$$

Where v_1 and v_2 are the velocities of the two components with masses M_1 and M_2 before the supernova explosion. Although the derivation assumes that the supernova does not produce a change in velocity in the star concerned (supernova kicks appear to be ~ 100 kms^{-1} from radio pulsar studies) it does give an upper limit to the likelihood of binary formation by this process. The above equations show that a binary may be disrupted ($e > 1$) if more than half the binary mass is lost, i.e. if the centre of mass velocity exceeds that of one of the binary components. The higher mass star in a binary evolves faster and is thus expected to explode first, disrupting the binary, so some process to reduce its mass must occur if the binary is to survive. If conservation of mass and angular momentum are not assumed for the binary, the envelope of the higher mass star can be ejected from the binary in the spiral-in scenario. Alternatively, accretion-induced collapse of a white dwarf could result in a relatively quiet supernova. So, both the supernova mechanism and neutron star capture scenarios have problems associated with the formation of LMXBs. Study of the wide variety of LMXBs found in the galaxy and in globular clusters should provide clues as to how these objects form and evolve.

Stripped Giant model

The stripped giant model was devised to explain systems which appear to contain under-dense secondaries, such as V404 Cyg which has $P_{orb} = 6.47$ days implying a secondary density of $\bar{\rho} = 0.0045$ g cm^{-3} (see King, 1993). In such a binary, the secondary has evolved to become a sub-giant and is overflowing its Roche lobe such that its outer layers are stripped away. The luminosity of the stripped giant is provided by a hydrogen burning shell around the core, the outer envelope is extended and has effectively no influence. The effective temperatures of such stars are similar ($T_{eff} < 3000$ K) and so the star's radius is proportional to $L^{0.5}$, which in turn is dependent only on the core mass. Thus, the mass loss from the outer envelope has no effect on the evolution of the star at this point. Webbink et al. (1983) fitted the luminosity and radius as a function of the core mass which, for a given binary period and assuming the secondary fills its Roche lobe, allows a set of solutions for M_2 and M_c to be derived.

$$(M_c/0.25)^{7.65} M_2^{-0.5} = 0.392$$

All that is required for a given system is to place limits on M_2, so that all of the parameters of the stripped giant can be calculated. Clearly, the minimum total mass is equal to the core mass, so that $M_2 \geq M_c$, and the secondary would not have left the main sequence if M_c/M_2 were less than the Schonberg-Chandrasekhar limit ≈ 0.17, which limits the total mass to $M_2 \leq 5.88\ M_c$. The high power of M_c in the above equation means that it is not very sensitive to the value of M_2, although the luminosity and radius have a sensitive dependence on the core mass. For this type of evolution it is expected that the orbital period will be larger than one day, implying a small enough secondary density. Below this value the secondary will either be a zero age main sequence star or a partly evolved main sequence secondary.

1.2 Soft X-ray Transients

1.2.1 Outburst

As described earlier, the SXTs are systems which remain in a dormant, quiescent phase for much of the time, and only go into outburst every 10 – 50 years. They may increase in brightness by as much as 9 magnitudes (in the case of J0422+32), accompanied by a large increase in the amount of X-ray luminosity produced (from a quiescent level in the range 10^{32} – 10^{34} erg s^{-1} up to 10^{38} erg s^{-1}, close to the Eddington limit). During the outburst these systems are very similar in behaviour to the persistently bright LMXBs and their overall behaviour is similar in nature to the dwarf nova CV systems. Also, it is known that the increase in brightness is due to accretion onto the compact object, causing X-ray production and the optical changes are caused by X-ray heating and gas flow through the disc. Unfortunately, the cause of this change in the level of accretion is not completely understood, although there are several models which attempt to explain the behaviour of the transients.

The optical spectra of SXTs during outburst resemble that of the persistently bright LMXBs, showing HeII λ 4686 emission and emission from the Bowen blend around 4640 Å, indicative of strong ionization in the system (Casares et al., 1995). Narrow emission has also been detected in the HeII λ 4686 and Balmer lines, which has been associated with X-ray heating of the inner face of the secondary (Harlaftis et al., 1994). These features are due to the high level of accretion and X-ray luminosity during outburst.

The X-ray spectra of SXTs during outburst have a soft component, and also a hard X-ray tail which extends up to several hundred keV (the photon index for the 20 – 200 keV range is between 1.5 to 2.0). The soft component (probably arising in the accretion disc) positively correlates with accretion rate and in most of these systems there is a significant absence of the 500 eV – 2 keV soft blackbody radiation seen in many LMXBs (Tanaka and Lewin, 1995). The cause of this emission is thought to be due to the thermalization, in the dense atmosphere of the neutron star, of hard X-rays arising in the boundary layer. Thus, the absence of these soft X-rays is seen as evidence that the compact object has no hard surface and could therefore be a black hole. Also, the absence of an atmosphere in black-hole SXTs means that there is no known mechanism for the production of soft X-ray photons, and thus the temperature in the inner disc can become very high ($\sim$ 30,000 K: Endal, 1976). Therefore a hard X-ray tail is created up to tens or hundreds of keV and the large-scale creation of electron-positron pairs also becomes possible. Electron-positron annihilation lines have already been seen in SXTs (Nova Muscae: Sunyaev et al., 1992).

'Mini' Outbursts

Several of the X-ray transients have now had a series of 'mini' outbursts after their main outburst (e.g. J0422+32 – Shrader et al., 1994; J1655−40 – Bailyn et al., 1995). These outbursts repeat over a timescale of months and rise by several magnitudes, before decaying after a period of around a week. The optical outburst is normally accompanied by a sharp rise in the X-ray flux. After a period of over a year, these outbursts cease and the object returns to quiescence. It is thought that recent, more intensive monitoring of these objects

has meant that more of these mini outbursts have been discovered, and that previous X-ray transients might have displayed similar behaviour.

Several models have been proposed to explain this behaviour, and to try to relate it to the processes causing the main outburst. Generally, the models require that X-ray illumination drives an increase in the mass transfer in some way. Chen et al. (1993) invoked X-ray heating of the outer disc to explain the first mini outburst and then X-ray heating of the secondary to produce the second (this model was published before tertiary outbursts were observed). Augusteijn et al. (1993) explain these outbursts as repeated 'echoes' of the main burst. The X-ray flux of the main outburst heats up the secondary, causing increased mass transfer, which in turn increases the X-ray flux in a repeated cycle. The delay between X-ray heating and the resumption of mass transfer sets a timescale between the outbursts and should cause them to constantly decrease in size with time, before eventually stopping altogether. Unfortunately, the mini outbursts all appear to have approximately the same increase in flux, and the predicted outburst times are not in complete agreement with the observations. However, these sort of feedback mechanisms seem promising as explanations of these recurrent outbursts.

Outburst Models

There are three basic models for the X-ray transients, which have been adapted from the models which are used to explain SU UMa outbursts: the thermal instability model, mass transfer model and the tidal instability models. More recent work (Ichikawa et al., 1994) has tried to link combinations of these models to try to explain the differences between the main outburst of SXTs and subsequent 'mini' outbursts that have now been seen in J0422+32 (Nova Persei 1992), GRS 1009−45 (Nova Velorum 1993) and J1655−40 (Nova Scorpius 1994).

The disc instability occurs because there are two possible temperature solutions for certain surface densities: one solution is a low viscosity cool state and the other is a high viscosity hot state (see figure 1.3). This is caused by the strong variations in opacity around 10^4 K where

the ionization of hydrogen occurs. During quiescence matter is transferred into the disc from the secondary, but does not accrete significantly onto the primary because the disc has a low viscosity. When the surface density reaches a critical value Σ_1 at some point in the disc an instability is triggered and the disc moves into the hot, high opacity state and because of the increased viscosity, mass accretes rapidly onto the primary. As the surface density decreases below a second critical surface density (Σ_2), the disc goes back into the cool, low opacity state. The overall $T - \Sigma$ relation is an elongated 'S' shape (see figure 1.3), and so this cyclical relaxation oscillation can explain the outbursts in dwarf novae and SXTs.

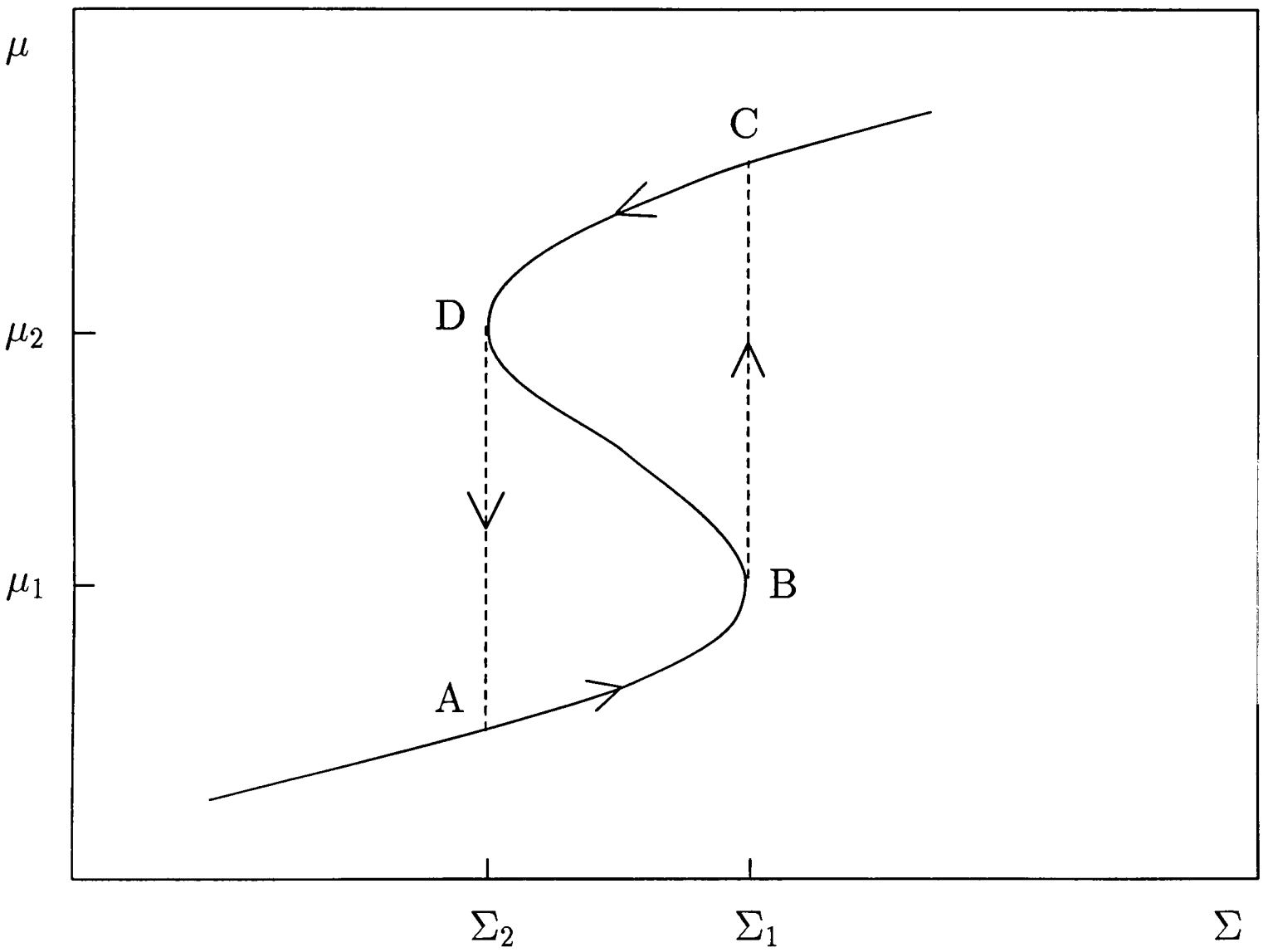


Figure 1.3: The viscosity (μ) versus surface density (Σ) diagram for the disc instability model of CVs and SXTs. Limit cycle behaviour (A $\rightarrow$ B $\rightarrow$ C $\rightarrow$ D $\rightarrow$ A) occurs because the viscosity is not a monotonically increasing function of surface density, due to the ionization of hydrogen within the disc.

Various numerical models have been computed (Mineshige and Osaki, 1985; Mineshige and Wheeler, 1989; Cannizzo et al., 1982) for the thermal instability, but they all require some prescription to calculate the viscosity in the disc. They use the standard α-disc assumptions, which parameterize the uncertainty about the viscosity in one quantity. Next, all that is needed is to define the value of α and any dependence on physical variables. It was found that constant α models only gave small amplitude outbursts ($\sim 2^m$) and so it was found to be necessary to have a larger value for α in the hot state than in the cool state equal to α_{hot} and α_{cool} (Ichikawa et al., 1994). Thus, thermal instability models also give us information on the physical properties of viscosity in the disc, but this also means that the exact form of the model is not known.

Tidal Instabilities and 'Superoutbursts'

The tidal instability was discovered by Whitehurst (1988) and was found to be most prevalent in high mass ratio systems, which makes SXTs ideal candidates for this effect. This instability is caused by the tidal interaction of particles orbiting in the disc with the secondary star. The simulation of Whitehurst considered particles in the disc undergoing a 3:1 resonance with the secondary, i.e. rotating three times for each orbital period. But, the 3:2 resonance can also occur for $q > 3$ and so this may also be important for some SXTs. The perturbation of these particles causes the accretion disc to become axisymmetric and rotate with respect to the inertial frame of reference. The period of this rotation is given approximately by :

$$\frac{\Delta P}{P_{orb}} = \frac{3q\ r_t^{3/2}}{4\sqrt{(1+q)a^3}}$$

Where $\Delta P = P_s - P_{orb}$ is the superhump period excess, r_t is the tidal radius beyond which the disc becomes unstable to perturbations from the secondary, a is the binary separation and q is the mass ratio. This expression assumes that the particles are isolated and that the tidal effect of the secondary is restricted to the radial direction only. However, simulations gave similar values as the above equation, showing that these assumptions are reasonably valid. The tidal instability requires a disc of a certain size and so either needs a burst of mass

transfer from the secondary or steady accretion over a long period of time to increase the radius of the disc to beyond the critical radius. Once the instability has set in, it becomes very efficient at removing angular momentum from the disc and can clear most of the mass from the disc.

In SXTs the rotation of the elliptical disc is thought to produce periodic X-ray heating of the secondary during the outburst, at a period which is slightly larger than the orbital one. This is known as the superhump period, from the humps which are seen in SU UMa light curves during super-outburst. For mass ratios $q > 4$ the difference between these two periods is at most 5.7 %, which should be resolvable using light curves spread over several days. This provides a further way of measuring the mass ratio of the binary, but the superhump effect has not been seen simultaneously with the orbital modulation of an SXT.

The mass transfer burst model (MTB) assumes that there is a period of temporary enhancement of mass transfer onto the disc to explain SXT outbursts (Osaki, 1985). The burst is postulated as being due to expansion of the secondary on a timescale of years over the critical Roche surface, allowing mass to flow across the L_1 point. This could be caused by X-ray heating from the primary, but calculations have shown that the timescale for the radius of the secondary to change is several orders of magnitude too large for this model to explain the outbursts of SXTs. However, most MTB models do not attempt to model the exact cause of the burst of mass transfer, but simply try to examine the effect it has on the accretion disc and the resulting flux variations.

The MTB model was applied to dwarf novae, but predicted that at the onset of the outburst the disc should undergo a transient shrinkage and very little shrinkage should occur in the disc radius in quiescence, contrary to observations (Ichikawa and Osaki, 1992). Thus, for dwarf novae the disc instability model appears to be the most viable. But this model alone cannot account for the 'mini' outbursts which have also been seen in SXTs after the initial outburst. These could be explained by a combination of the tidal and thermal instabilities as suggested by Ichikawa et al. (1994). The thermal instability causes the disc to become

brighter and expand radially, until it reaches a critical radius, after which the tidal instability sets in clearing most of the matter from the disc. The tidal instability would then cause a gradual disc shrinkage, which can explain the slow exponential decays of SXTs.

1.2.2 Quiescence

It typically takes SXTs many months to decay into quiescence after the initial outburst, and some take over a year after repeated 'mini' outbursts (e.g. J0422+32). The decay is exponential with typical e-folding timescales of $\sim$ 30 – 200 days, but the e-folding time has been seen to vary over the duration of the decay. Quiescence lasts for 10 – 50 years in SXTs (apart from Aql X-1 which goes into repeated 'mini' outbursts on a timescale of 1 year, see Kaluzienski et al., 1977). The SXTs have typical X-ray luminosities of 10^{32} – 10^{34} erg s^{-1} during quiescence and the optical light mainly comes from the secondary, with a small percentage coming from the disc (rising toward the blue). This means that spectroscopy of the secondary star can be achieved with little contamination, allowing its radial velocity and rotational broadening to be measured.

The spectra of SXTs show double peaked Hα and Hβ emission lines which presumably comes from the disc, and various absorption lines from the secondary. Lithium absorption (LiI 6707.8 Å) in particular has been detected in several systems (e.g. V404 Cyg, A0620−00, Cen X-4; Martín et al., 1994) which is not currently understood, as it should be rapidly depleted in the secondary over a timescale of 10^6 years. This suggests either that these systems have some mechanism for the production of lithium such as spallation, or that lithium depletion has not yet progressed very far. No HeII 4686 emission is usually seen in quiescent SXT spectra and little HeI, indicating that there is little ionizing radiation coming from the primary. This is due to the fact that there is little accretion during quiescence, but Doppler tomography (A0620−00: Marsh et al., 1994) has shown that mass transfer still continues.

Ellipsoidal studies

During quiescence the accretion disc in SXTs becomes very faint and it is possible to observe the ellipsoidal modulation in the secondary star. This modulation is caused by the changing aspect of the distorted secondary with respect to us, combined with limb and gravitational darkening (see figure 1.1). These effects cause a double modulation per orbital cycle with minima occurring at phases 0 and 0.5. The phase 0.5 minimum (when the inner face of the secondary is towards us) is the deepest due to the larger gravity darkening in that direction, and so there should be differing minima in the light curve when it is folded on the (correct) orbital period. Thus, the ellipsoidal light curve is in effect the sum of two sinusoids with periods P_{orb} and $P_{orb}/2$ ($P_{orb}/2$ has the larger amplitude), and so the most common way to detect it is to take the discrete fourier transform of the data and multiply the period of the largest peak by two to obtain the orbital period.

The gravity darkening effect is due to the fact that the emergent flux of total radiation over the surface of a tidally distorted star in radiative equilibrium varies proportionally to the local gravity. Since equipotential surfaces are isothermal in the star, the temperature gradient over such a surface is proportional to Ψ, which is simply the value of the local gravity. Therefore, the outward flow of heat over the equipotential surface should be proportional to the local gravity. The dependence of the effective temperature (and also the luminosity) is parameterized by the gravity darkening exponent β (see below). This is used because transfer of heat by convective currents beneath the photosphere means that the flow of heat may no longer be exactly proportional to the temperature gradient. This exponent allows this factor to be taken into account (a value of 0.08 for β is used for convective stars). Gravity darkening acts in addition to the limb darkening which is caused by the fact that radiation viewed normally to the emitting surface originates from a greater depth within the star and thus from a region with a greater temperature. The apparent surface brightness is dependent on the angle of foreshortening (between the surface normal and the line of sight).

Modelling the ellipsoidal effect in SXTs allows the inclination of the system to be determined, but the mass ratio cannot be measured accurately with this technique. The ellipsoidal model requires several other parameters to be known: the secondary's effective temperature T_{eff}, the limb darkening coefficient u, the gravity darkening exponent β (normally taken to be equal to = 0.08 for convective stellar atmospheres, see Lucy, 1967), and effective filter wavelength λ_{eff}, which leaves the mass ratio q and inclination i as free parameters. The effective temperature of the secondary in an SXT system can be obtained by modelling the energy distribution with a blackbody curve, or by estimating its spectral type by cross-correlation with template stars. The limb darkening coefficient, u, relates to the linear limb-darkening law of Al-Naimiy (1977), which is given by:-

$$I(\mu) = I(1)\{1 - u + u\cos\gamma\}$$

Where $\mu = \cos\gamma$, and γ is the angle of foreshortening of the emergent radiation. This law is a parameterization of the limb-darkening and is obtained empirically from stellar model atmospheres of convective stars; the value of u that is used is essentially a mean value taken over the entire stellar disc. This assumption has a negligible effect as the actual value remains within 10% of the mean value of u over the surface of the star. The gravitational darkening exponent is related to the effective temperature by the following expression: $T_{eff} \propto g^{\beta}$. The value of β was determined by a model of a convective stellar envelope, and was found to be consistent with values predicted by observational studies (Sarna, 1989).

The model used in this work is the same as that used to fit the infra-red light curves of Cen X-4 and V404 Cyg (Shahbaz et al., 1993 and Shahbaz et al., 1994), and is similar to that used by Tjemkes et al. (1986). The secondary star is split into surface elements, and for each element the limb darkening, gravity darkening and projected surface area onto the line of sight are calculated. This is performed over the whole of the stellar surface, using the known T_{eff} and filter wavelength to calculate the brightness distribution. This procedure is repeated for the whole binary orbit to generate the light curve, which is then compared to the observed data by means of the chi-squared statistic.

Radial velocity curves and rotational broadening

During quiescence it is possible to observe the spectrum of the secondary with little disc contamination, and this means that it is possible to measure its radial velocity curve. This can be done by cross-correlation of each spectrum with a template star of the correct spectral type, and then the radial velocity curve of these measurements allows the mass function of the system to be determined. This is given by:

$$f(M) = \frac{(M_1\ sini)^3}{(M_1 + M_2)^2} M_\odot = \frac{K_2^3\ P_{orb}}{2\pi G}$$

The mass function gives the lower limit to the mass of the compact object M_1 (when $i = 90°$ and the secondary mass, $M_2 = 0$), and so can be a useful indicator of the nature of the primary. The mass function is useful because it can be obtained from measurable quantities, i.e the secondary's radial velocity K_2 and the orbital period P_{orb}. If the mass function is significantly above 3 $M_\odot$, then this is strong evidence that the compact object is a black hole, as most models of the neutron star equation of state predict that the highest stable configuration for neutron stars occurs at this order of mass. SXTs are ideal for this sort of study because the mass ratios are usually large and so M_2 is small, which means that the denominator in the above equation will be as small as possible and the mass function will be large. In HMXB systems where the secondaries are much more massive, the above indicator is less useful as the high secondary mass will make the denominator large and the values for $f(M)$ will tend to be small.

Once the mass function has been measured all that is needed to obtain the system masses, is the mass ratio q. This can be measured by studying the Doppler broadening of the secondary star as it rotates around the primary. One way to accomplish this is to measure the broadening in metallic absorption lines in the stellar spectrum and then compare this against a template spectrum from a star with the same spectral type as the secondary. The excess broadening can then be measured by cross-correlation of the spectra with a Gaussian and assuming that the secondary fills its Roche lobe (this is a safe assumption as all SXTs are seen to be undergoing

mass transfer) the rotation velocity can be estimated. The rotational broadening can then be related to the mass ratio with the further assumptions that the secondary is tidally locked with the binary orbit and spherically symmetric (see Wade and Horne, 1988):-

$$v_{rot} \sin i = (K_1 + K_2) \frac{R_{L_2}}{a} \sim 0.462 K_2 q^{1/3}(1+q)^{2/3}$$

To obtain the value of q requires the inclination and radial velocity to be input into the above equation. Once these parameters are known, an estimate of the mass of the primary can be made and tested against the values for the highest mass rotating neutron star models (≈ 3.8 $M_\odot$). Alternatively, a firm upper limit can be made by assuming that general relativity is the correct theory of gravity; causality is satisfied (i.e. the speed of sound is less than c), and the condition of 'microscopic stability' preventing the spontaneous collapse of matter (see Rhoades and Ruffini, 1974). This yields a firm upper limit of ≈ 3.2 $M_\odot$, on the maximum mass of a neutron star.

Another method for estimating q requires that the primary's velocity about the system centre of mass be measured. As $q = M_1/M_2 = K_2/K_1$, this means that if K_2 is known K_1 can give the value of q directly. This can be done by use of a technique known as the diagnostic diagram, which represents a series of double Gaussian cross-correlations of an emission line in a series of spectra. The centre of the two Gaussians, for a given width and separation, traces out a sine wave with time, the amplitude of this sine wave is a measure of the rotational velocity of the material in the disc at a given radius about the centre of mass (see Schneider and Young, 1980). As the Gaussian separation is increased towards the wings of the line, this corresponds to a higher velocity and therefore comes from material closer to the primary, until at a very large separation the velocity tends to a constant value which is the velocity of the primary. Normally, the limit on the separation that can be used in this method is the decreasing signal-to-noise as you progress toward the wings of the line.

1.3 Summary

LMXBs can provide insight into the mechanisms of stellar evolution. By studying LMXBs, the physical parameters of the two components can be discovered, such as masses, radii and so on. In the case of the secondary this provides clues as to how it evolves over time in the binary and by looking at many different systems an overall picture can be formed of the nature of LMXBs. The mass of the primary can constrain the physics of black holes and neutron stars and give a useful guide to theoretical models of these objects. But, to use LMXBs to provide such information requires that the binary parameters are well known. This means that it is necessary to obtain orbital periods, and to understand the other types of variability that occur simultaneously in the light curves and radial velocities of LMXBs. The remainder of this work is an attempt to examine optical variability in LMXB light curves by using new multi-frequency analysis techniques and modelling the periodicities obtained.

References

Al-Naimiy H.M., 1978, Ap. & Space Sci., **53**, 181.

Augusteijn T., Kuulkers E. and Shaham J., 1993, Astr. and Astrophys., **279**, L13.

Bailyn C.D., Orosz J.A., McClintock J.E. and Remillard R.A., 1995b, IAU Circ., No. 6173.

Cannizzo J.K., Ghosh P. and Wheeler J.C., 1982, Ap. J., **260**, L83.

Casares J., Marsh T.R., Charles P.A., Martin A.C., Martín E.L., Harlaftis E.T., Pavlenko E.P. and Wagner R.M., 1995, MNRAS, **274**, 565.

Charles P.A., 1995, IAU Symp. No. 165, ed. E.P.J. van den Heuvel, in press.

Chen W., Livio M., and Gehrels N., 1993, Ap. J., **408**, L5.

Cohn H., 1987, in IAU Symp. No. 126, *Globular Cluster Systems in Galaxies*, eds. J.E. Grindlay and A.G.D. Philip, Reidel, Dordrecht, pg. 379.

Cowley A.P., Hutchings J.B., Crampton D. and Hartwick F.D.A., 1987, **320**, 296.

Dotani T., Inoue H., Murakami T., Nagase F., Tanaka Y., Tsuru T., Makishima K., Ohashi T. and Corbet R. H. D., 1990, Nature, **347**, 534.

Endal A.S., Devinney E.J. and Sofia S., 1976, Ap. Lett., **17**, 131.

Fabian A., Pringle J. and Rees M., 1975, MNRAS, **172**, 15P.

Frank J., King A.R. and Raine D.J., 1992, *Accretion Power in Astrophysics*, 2nd ed., Cambridge University Press, pg. 56.

Grindlay, J.E., 1984, Adv. Space Res., **3**, No. 10, 19.

Grindlay J.E., 1987, in IAU Symp. No. 125 *Origin and Evolution of Neutron Stars*, eds. Helfand D. and Huang J., Dordrecht: Reidel.

Harlaftis E., Jones D.H.P., Charles P.A. and Martin A.C., 1994, in Holt S.S., Day C., eds., *The Evolution of X-Ray Binaries*, AIP, New York, 91.

Hoffman J.A., Marshall H.L. and Lewin W.H.G., 1978, Nature, **271**, 630-633.

Ichikawa S., Mineshige S. and Kato T., 1994, Ap. J., **435**, 748.

Ichikawa S. and Osaki Y., 1992, PASJ, **44**, 15.

Kaluzienski L.J., Holt S., Boldt E.A. and Serlemitsos P.J., 1977, Nature, **265**, 606.

King A.R., 1993, MNRAS, **260**, L5.

Lewin W.H.G. and van Paradijs J., 1988, Sp. Sci. Rev., **46**, 273.

Lucy L.B., 1967, Z. Astrophys., **65**, 89.

McClintock J.E. and Remillard R., 1986, Ap. J.,**308**, 110.

Marsh T.R., Robinson E.L., Wood J.H., 1994, MNRAS, **266**, 137.

Martín E.L., Rebolo R., Casares J. and Charles P.A., 1994, Ap. J., **435**, 791.

Mineshige S. and Osaki Y., 1985, PASJ, **37**, 1.

Mineshige S. and Wheeler J.C., 1989, Ap. J., **343**, 241.

Osaki Y., 1985, Astr. Astrophys., **144**, 369.

Paczyński B., 1971, Ann. Rev. Astr. Ap., **9**, 183.

Paczyński B., 1981, Acta. Astr., **31**, 1.

Press W.H and Teukolsky S.A., 1977, Ap. J., **213**, 183.

Rhoades C.E. and Ruffini R., 1974, Phys. Rev. Lett., **32**, 324.

Sarna M.J., 1989, Astron. Astrophys., **224**, 98.

Schneider D.P. and Young P.J., 1980, Ap. J., **238**, 946.

Shafter A.W., Misselt K.A., Szkody P. and Politano M., 1995, MNRAS, **448**, L33.

Shahbaz T., Naylor T. and Charles P.A., 1993, MNRAS, **265**, 655.

Shahbaz T., Ringwald F.A., Bunn J.C., Naylor T., Charles P.A. and Casares J., 1994, MNRAS, **271**, L10.

Shrader C.R., Wagner R.M., Hjellming R.M., Han X.H., and Starrfield S.G., 1994, Ap. J.,

434, 698.

Sunyaev R. et al., 1992, Ap. J., **389**, L75.

Tanaka Y. and Lewin W.H.G., 1995, in *X-ray Binaries*, eds. W.H.G. Lewin, J. van Paradijs and E.P.J. van den Heuvel, Cambridge University Press, pg. 126.

Tawara Y., Kawai N., Tanaka Y., Inoue H., Kunieda H. and Ogawara Y., Nature, **318**, 545.

Thorne K.S., Zytkow A.N, 1977, Ap. J., **212**, 832.

Tjemkes S.A., Zuiderwijk E.J. and Van Paradijs J., 1986, Astr. Astrophys., **154**, 77.

van den Heuvel E.P.J., 1983, *Accretion-driven stellar X-ray sources*, eds. W.H.G. Lewin and E.P.J. van den Heuvel, Cambridge University Press, pg. 303.

van der Klis M., 1989, Ann. Rev. Astr. Ap., **27**, 517-553.

van Paradijs J., 1981a, Astron. Astrophys., **103**, 140.

van Paradijs J., 1981b, Astron. Astrophys., **101**, 174.

van Paradijs J., 1983, in *Accretion-driven stellar X-ray sources*, eds. W.H.G. Lewin and E.P.J. van den Heuvel, Cambridge University Press, 189.

van Paradijs J. and McClintock J.E., 1994, Astr. Astrophys., **290**, 133.

van Paradijs J., Verbunt F., Shafer R.A. and Arnaud K.A., 1987, Astr. Ap., **182**, 47.

Verbunt F., 1993, Ann. Rev. Astr. Ap., **31**, 93.

Verbunt F., Belloni T., Johnston H.M., van der Klis M. and Lewin W.H.G., 1994, Astr. Ap., **285**, 903.

Wade R.A. and Horne K., 1988, Ap. J., **324**, 411.

Watson M.G. et. el., 1985, IAU. Circ. No. 4051.

Webbink R.F., Rappaport S. and Savonije G.J., 1983, Ap. J., **270**, 678.

White N.E., 1986, in *The Evolution of Galactic X-ray Binaries*, ed. Truemper J., Lewin

W.H.G. & Brinkman W., p. 227, Dordrecht: Reidel.

Whitehurst R., 1988, MNRAS, **232**, 35.

Chapter 2

Bayesian Spectrum Analysis

2.1 Methods of Frequency Analysis

The first statistical method used for frequency analysis was least squares, initially published by Legendre (1806) and Gauss (1963 reprint). A function is defined to model the data (normally a sinusoid), and then the sum of the residuals of the data minus the model (χ^2) are minimized by varying the model parameters. This allows one to obtain the frequency, amplitude and phase of the best fit. The value of χ^2 versus frequency gives the chi-squared periodogram.

At the turn of this century, Sir Arthur Schuster devised the periodogram as an *ad hoc* method of frequency analysis (Schuster, 1905). It was defined as the sum of the data projected onto sine and cosine functions, over a series of discrete times (see below). So, if the data contain a frequency at ω_0, the periodogram calculated at a frequency $\omega = \omega_0$ yields a peak. This is now more commonly known as the Discrete Fourier Transform, and any general method of summing over the data to produce a frequency spectrum is called a periodogram.

2.1.1 Discrete Fourier Transform

The Discrete Fourier Transform (as described above) was adapted from the fourier transform by replacing the integral with a summation in order to estimate the power at a given

frequency for discretely sampled data. The power at a frequency of ω is defined as the squared magnitude of the discrete fourier transform of the data:

$$C(\omega) = \frac{1}{N}[R(\omega)^2 + I(\omega)^2] = \frac{1}{N}|\sum_{j=1}^{N} d_j e^{-i\omega t_j}|^2 \tag{2.1}$$

where,

$$R(\omega) = \sum_{j=1}^{N} d_i \cos(\omega t_i)$$

$$I(\omega) = \sum_{j=1}^{N} d_i \sin(\omega t_i),$$

and the data consist of the set of values:

$$D = \{d_1, d_2, ..., d_N\}$$

C(ω) is just the sum of the real and imaginary parts of the DFT; d_i are the data taken at times t_i and N is the total number of data points. The way this statistic works as a frequency estimator is that if the data contain a sinusoid at a frequency ω_0 then, near $\omega = \omega_0$, the terms d_j and $e^{-i\omega t_j}$ in equation 2.1 are in phase and the summation becomes very large. Thus, at any frequency where there is power in the data, a peak will appear in the periodogram. The error associated with this peak is approximately equal to its full width half maximum. The error can be more accurately obtained by fitting a Gaussian to the peak – the error is then equal to the standard deviation given by the width of this Gaussian.

The DFT serves as a useful frequency estimator under a variety of conditions. But it breaks down if N is small, the frequency varies in amplitude or phase, or there is more than one frequency present in the data. Added to this is the problem of aliasing which is caused by the occurrence of regularly spaced gaps in the data. This can cause spurious peaks to appear in the periodogram and make it difficult to detect the true frequency of any modulation. The problems associated with the DFT have led to several related methods to be developed, some of which are described below.

2.1.2 Lomb-Scargle Periodogram

The DFT is the correct frequency estimator for evenly spaced data, in that it contains all of the frequency information present in the data (it is a sufficient statistic). But if the data are not evenly spaced then the statistical behaviour of the periodogram is not well known, and for data whose sampling is pathological, the DFT will not yield the correct frequency present in the data. In order to correct this defect the Lomb-Scargle periodogram was defined (Scargle, 1982) as a modified DFT, generalized to the case of uneven spacing. For the Lomb-Scargle periodogram the statistical behaviour is simple, and the method is equivalent to least squares fitting of sinusoids to the data (Lomb, 1976). As this statistic is merely an extension of the DFT, it reduces to the same form as the DFT in the case of even sampling.

Astronomical time series are often unevenly sampled due to the diurnal cycle, and so the Lomb-Scargle periodogram is often the best method to use. Throughout this work it was used instead of the DFT on all of the astronomical datasets, as they are typically unevenly sampled.

2.1.3 Phase Dispersion Minimization

Phase dispersion minimization (PDM; Stellingwerf, 1978) is a method which is useful for the detection of non-sinusoidal modulations in data. The algorithm works as follows: at a given frequency, the data is folded and then divided into n bins. The dispersion in each of these bins is calculated, each bin is moved along by a fraction f of the bin width and the process is repeated $1/f$ times. This produces the PDM Θ statistic, the minimum of which represents the strongest signal. This method is not very well understood statistically, but does have the advantage that it does not assume that the signal in the data is necessarily sinusoidal.

2.1.4 CLEAN

The CLEAN algorithm (Roberts et al., 1987) is a further extension of the DFT. It attempts to remove the effects of aliasing due to the data sampling by deconvolving the window function from the data, and is particularly useful for data with multiple frequencies. The fourier transform of the sampled signal D(ν) is defined as a convolution of the spectrum of the true signal F(ν), with the fourier transform of the sampling function W(ν):

$$D(\nu) = F(\nu) \otimes W(\nu)$$

D(ν) and W(ν) are known as the dirty spectrum and the spectral window function, respectively. The aim of CLEAN is to deconvolve D(ν) and thus to find F(ν), the true (CLEANed) spectrum of the data. The algorithm works by finding the amplitude and frequency of the first spectral component and removing it from the data. However, only a fraction g (called the loop gain) is removed in order to improve the stability of the algorithm. This process is repeated for a pre-set number of iterations until a set of spectral components is obtained. These spectral components are then convolved with the frequency resolution of the data ($\sim 1/T$), which is known as the clean beam B(ν). This is then added to the residual spectrum, left after the removal of all the spectral components from the data, to obtain the CLEANed spectrum (for a good example of the performance of the CLEAN algorithm, see figure 2 of O'Dell and Collier Cameron, 1993, who used it to detect the rotational periods of G–K dwarfs). The algorithm performs well at removing the effects of the data sampling of a dataset, but the method assumes the data contain zero noise, which can cause the algorithm to fail. Noisy datasets, which are unevenly spaced, often have spurious alias peaks which are higher than the true signal peaks (Scargle, 1982). Because the CLEAN algorithm does not take noise into account it will simply assume that the highest peak is the true signal peak, and will then deconvolve the periodogram using the wrong periodicity. The resulting CLEANed spectrum may then contain several peaks, none of which correspond to the true signal frequency.

2.1.5 Other methods

There are various other methods in use besides the ones outlined above, the most common of which is the chi-squared periodogram. As mentioned previously, this is mathematically equivalent to the DFT and is essentially an algorithm to fit sinusoids to the data and calculate the chi-squared for the fit versus frequency.

The 'string' method is similar to PDM in that it folds the data over a set of trial frequencies. It then calculates the distance between consecutive points, and computes the sum of these lengths of 'string' over all points in the dataset. This is not a very practical method, as the string length is not a well defined statistic and can vary erratically.

Maximum entropy is useful in many different fields of research and has also been used in periodogram analysis in order to remove the implicit assumption underlying most methods, i.e that the data contain only a single sinusoid. Maximum entropy seeks the maximum entropy periodogram and uses an effectively unlimited number of sinusoids to model the data, using the maximization of entropy as the main constraint. Unfortunately, this method is also not robust to the effects of aliasing and noise. Hence, a Bayesian periodogram was developed to allow multiple frequency models of the data to be defined.

2.2 The Bayesian Method of Periodogram Analysis

2.2.1 General Outline

Bayesian probability theory is a simple but very effective tool in many different areas of science, although it is only now becoming widely used in astrophysics. Using Bayesian methods to analyze data it is possible to include any prior information about the data, sometimes producing orders of magnitude improvement on the results obtained. It also allows models of the data to be explicitly defined and then compared using probability analysis.

The origin of Bayesian statistics is Bayes' theorem. This can be derived from the law of commutativity of probabilities. In other words the statements 'both H and D are true' and 'both D and H are true', are equivalent. This leads to the following theorem:

$$P(H, D|I) = P(D|I)P(H|D, I) = P(H|I)P(D|H, I)$$

which gives Bayes' theorem

$$P(H|D, I) = \frac{P(H|I)P(D|H, I)}{P(D|I)}$$

In this notation, $P(H|D, I)$ means the probability of H given D and I, where D represents the data, H is the hypothesis to be tested and I is any prior information that is known. In Bayesian theory $P(H|D, I)$ is the probability that the hypothesis or model for the data is the correct one and $P(H|I)$ is the prior probability of the hypothesis, before any data have been accumulated. $P(D|H, I)$ is known as the direct probability of the data, which, for a given hypothesis is just the sampling distribution of the data. $P(D|I)$ is a normalization constant and does not affect relative probabilities.

It is necessary to calculate the posterior probability from the data given the direct probability of the data, and any prior information that is known about the parameters of the model. These prior probability distributions are known as priors and indicate the current state of knowledge about the parameters. For example, if the amplitude of the modulation is known to within an order of magnitude, a Gaussian prior could be used for the amplitude with a width equal to the uncertainty. But, to be as conservative as possible uninformative priors are often used, and so no prior information is assumed. In this section no prior information is assumed and just a straight forward Bayesian model is used for the data.

2.2.2 A Simple Bayesian Model

In order to analyze a set of data it is necessary to define a model and then to estimate the parameters for that model (for a detailed discussion of Bayesian spectrum analysis and parameter estimation see Bretthorst, 1988). For time series analysis the most useful model

is that of a sinusoid, as most modulations are smooth, and any small departures from this assumption do not have a large effect. This is because most non-harmonic signals still have a large component at the fundamental frequency and so periodogram analysis will still find the correct frequency but with less accuracy than before. The most simple model therefore is that of a single sinusoid:

$$f(t_i) = B_1 \cos(\omega t_i) + B_2 \sin(\omega t_i)$$

which models the data,

$$d_i \equiv y(t_i) = f(t_i) + e_i$$

In this equation d_i is the ith data point, $f(t_i)$ is the model (a single stationary sinusoid) and e_i is any noise associated with the data. Each frequency consists of sine and cosine terms in order to define the model function in terms of amplitudes only, which simplifies the mathematics, and so for a given model there are m frequencies consisting of $r = 2m$ sine and cosine terms. Next, the distribution of the noise is required. Firstly, it is assumed that the noise has zero mean and variance σ^2, i.e. the data must have had the mean subtracted from it. This leads to the following conditions on the noise distribution:

$$\int_{-\infty}^{\infty} ep(e)de = 0$$

$$\int_{-\infty}^{\infty} e^2 p(e)de = \sigma^2$$

Now, the distribution which assumes only these two pieces of information (and also therefore maximizes the entropy) is a Gaussian distribution:

$$p(e) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{e^2}{2\sigma^2}\right)$$

From this the direct probability of the data (d_i) can be obtained, given the parameters of the model. The parameters are: the amplitudes of the model B_1 and B_2, which define the phase and amplitude of the sinusoid; the frequency, noise and initial information I:

$$P(D|B_1, B_2, \omega, \sigma, I) \propto \prod_{i=1}^{N} \sigma^{-1} \exp\{-\frac{1}{2\sigma^2}[d_i - f(t_i)]^2\}$$

Now the single sinusoid model is substituted in for $f(t_i)$ and expanded. The result is multiplied by a uniform prior and then integrated with respect to B_1 and B_2 to remove the dependence on phase and amplitude of the result. This gives the posterior probability density:

$$P(\omega|D, \sigma, I) \propto \exp\left(\frac{C(\omega)}{\sigma^2}\right)$$

where,

$$C(\omega) = \frac{1}{N}\left[R(\omega)^2 + I(\omega)^2\right]$$

$C(\omega)$ is the definition of the discrete fourier transform as in equation 2.1. The probability of a given frequency representing the true model of the data is therefore essentially the exponential of the DFT in the single sinusoid case (if no prior information is assumed). But, the noise level may not be known and so in that case the probability is multiplied by a prior for the noise $(1/\sigma)$, and integrated as before to obtain the posterior probability density for data with an unknown noise level:

$$P(\omega|D, \sigma, I) \propto \left[1 - \frac{2C(\omega)}{N\overline{d^2}}\right]^{\frac{2-N}{2}}$$

with $\overline{d^2}$ given by:

$$\overline{d^2} = \frac{1}{N}\sum_{i=1}^{N} d_i^2$$

Normally, though, for astronomical time series studies in this field the random noise can be estimated from the data. Then any systematic errors present can be taken account in interpreting the resulting periodogram. Also, it should be noted that the proportionality sign in the above equation arises because certain unnecessary normalizing factors have been dropped. The probability can be normalized to one numerically, but this is not always done to save computing time. Usually, the only quantities of importance are relative probabilities. For a more detailed discussion of the above analysis see Bretthorst (1988).

2.2.3 Examples for the Single Sinusoid Model

Example 1: Uniformly Sampled Data.

To compare the behaviour of the Bayesian method with DFT related periodograms an artificial dataset was constructed, and both the 1 sinusoid Bayesian model and a Lomb-Scargle periodogram analysis were then applied. The data were created with a frequency of $f = 4.0$ cycles/day, unit amplitude, Gaussian noise with $\sigma_N = 1.0$ and generated over time intervals that were randomized with σ_t equal to one tenth their spacing. A total of 250 ($=N$) points were generated:

$$d_j = \cos[2\pi(4t_j)] + e_j,$$

$$t_j = 1, 2, ..., N; \ \pm \ \sigma_t = 1/10$$

Thus, the signal-to-noise level (S/N) is 1, given the above sinusoidal amplitude and noise level. The artificial dataset, Lomb-Scargle and Bayesian periodograms are shown in figure 2.1. The frequency estimates given by the two methods are as follows: $f_{DFT} = 3.93 \pm 0.48$ and $f_{Bayes} = 4.014 \pm 0.02$. The accuracy for the DFT estimate was taken to be its full width half-maximum and that for the Bayes periodogram was calculated using equation (2.2) (see section 2.2.8). The parameters derived from the Bayesian model are: amplitude $= 1.15 \pm 0.08$, $\sigma_N^2 = 0.91 \pm 0.08$, and posterior probability P(f)= $10^{29.7}$.

$$f_{est} = \hat{f} \pm \frac{\sigma}{2\pi \hat{B} T} \sqrt{48/N^3} \tag{2.2}$$

The DFT finds the correct frequency with an accuracy similar to the Bayesian method, but it overestimates the error on the frequency. This is because the DFT has been used indiscriminately without defining what model is being used. If the data had not been just singly periodic the DFT error would have been more reasonable. The Bayesian method calculates the error by including the knowledge that the data contain only one sinusoid,

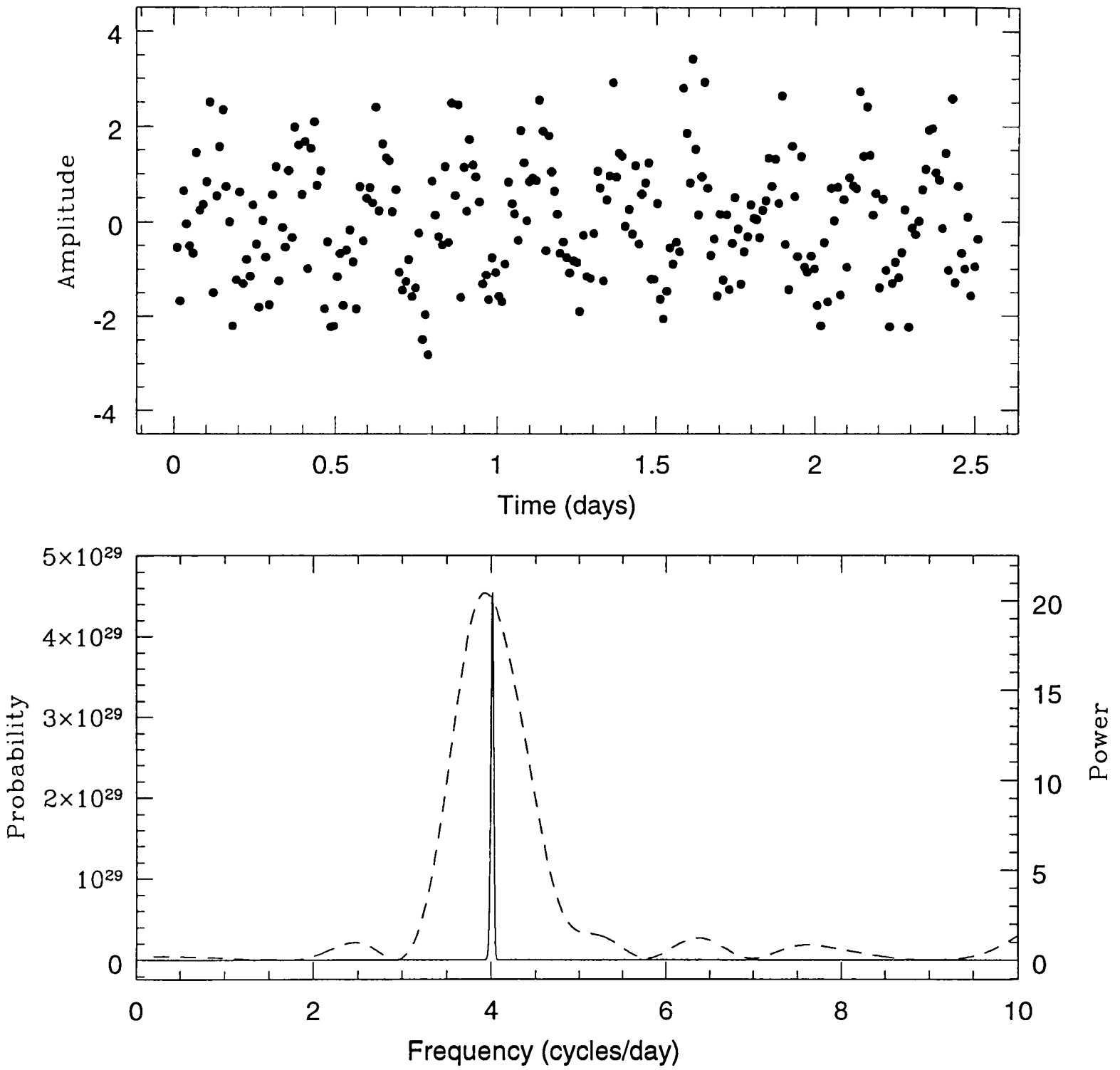


Figure 2.1: The Lomb-Scargle periodogram and Bayesian posterior probability density for example 1. Top: the dataset consisting of 250 points and bottom: the DFT and Bayes periodograms superimposed. The Bayes periodogram can be seen to be much more sharply peaked and so gives a higher accuracy for the frequency estimate.

whereas the DFT does not. This illustrates the fact that the Bayesian method makes all the information that is included explicit, so that the meaning of any result can be defined exactly.

Example 2: Non-uniformly Sampled Data

Example 2 compares the two methods under more realistic conditions, such as are met in astronomy. This dataset consisted of 100 points spread over five days with Gaussian noise at the same level as the previous data, and a sinusoid of amplitude 1.0 and frequency equal to 4.0 cycles/day. To simulate the missing observations due to the diurnal cycle, the data were calculated in blocks of 8 hours with 16 hour gaps. This mimics the aliasing problems often met in astronomical time series, where gaps in the data cause spurious peaks to appear in the corresponding periodograms. Each data point was separated by 0.02 days (28.8 minutes), with a random variation in the timings of 0.001 days (σ_t = 1.44 minutes).

The data from this simulation and the two periodograms are shown in figure 2.2. The data contain large gaps which cause the periodic structure in the DFT around the 4 cycles/day peak. Because the Bayesian periodogram is essentially an exponentiated DFT in this case, the sidelobe structure in the posterior probability density is suppressed. The two periodograms give the following frequencies: $f_{DFT} = 4.005 \pm 0.09$ and $f_{Bayes} = 4.012 \pm 0.02$. From the Bayesian model the calculated parameters are: amplitude = 1.004, $\sigma_N^2 = 1.16 \pm 0.17$, and P(f) = $10^{7.8}$. The DFT does give a frequency closer to the actual value for these data than the Bayesian estimate, but on average both methods would do equally well. The only difference is that the DFT again overestimates the error on the frequency estimate, whereas the Bayesian method gives a more reliable estimate. It assumes explicitly, the additional information that the data contain only one periodicity.

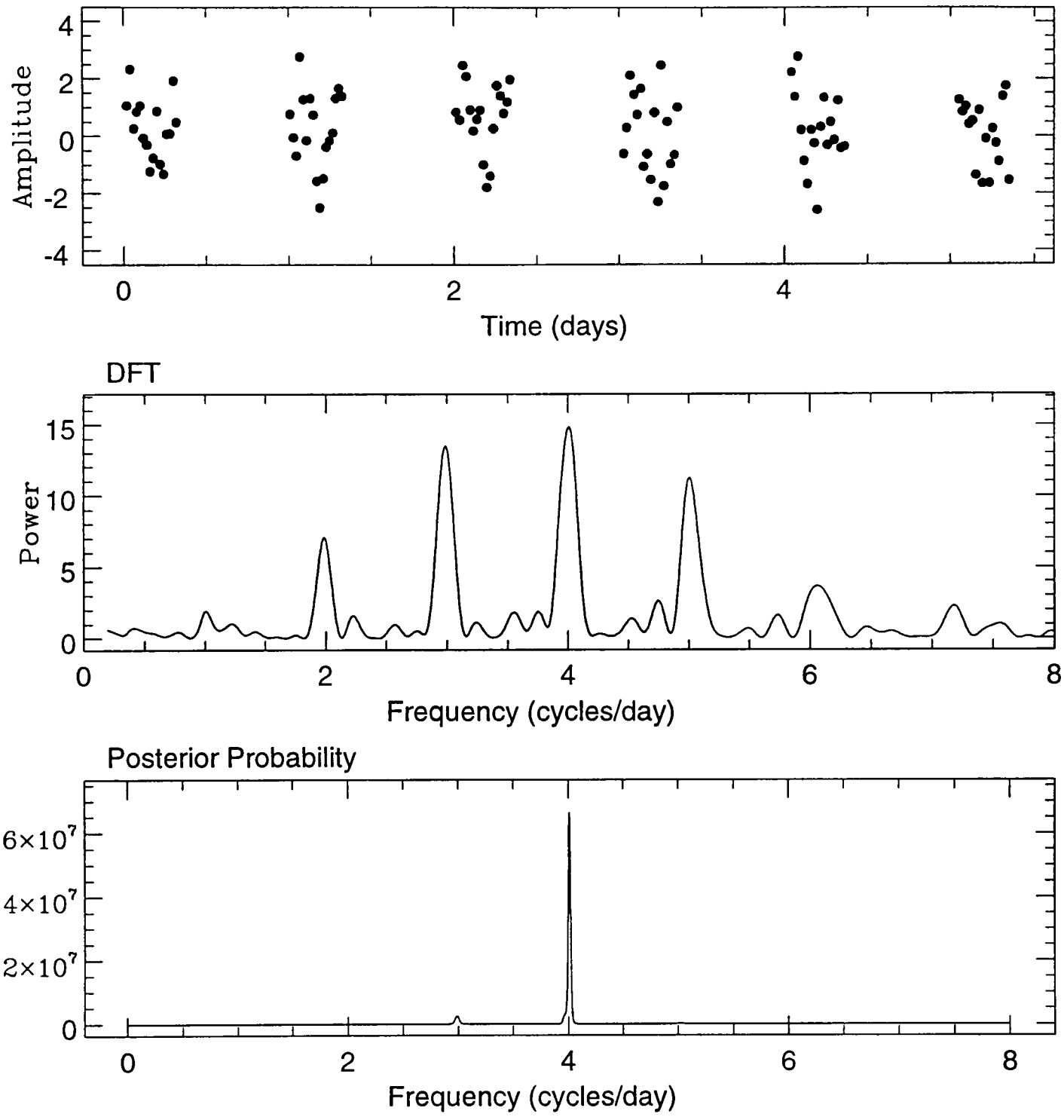


Figure 2.2: The Lomb-Scargle periodogram and Bayesian posterior probability density for example 2. This dataset simulates real astronomical time series by imposing a 1/3 duty cycle on the data points. The resulting alias pattern in the DFT causes it to be unclear which peak is correct in the DFT, however, the Bayesian probability plot indicates that the peak at 4.0 cycles/day is the most probable by more than one order of magnitude.

2.2.4 Examples for the Two Frequency Model

Example 3: Uniformly Sampled Data

For the case where there is only a single stationary sinusoid present in the data, the DFT does not perform much worse than the Bayesian method, and even when there are two frequencies the DFT can still detect both of them if they are well separated. But in the case of two close frequencies it will fail because of the interference between the two frequencies. The dataset generated next (example 3) is identical to example 1, except that the data consists of two frequencies at $f = 0.048$ and 0.049, both of unit amplitude (and N=500).

As can be seen in figure 2.3, the DFT cannot resolve these two frequencies. It has a broad peak half way between the frequencies 0.048 and 0.049 which reaches a maximum at a frequency of $f_{DFT} = 0.0485 \pm 0.0011$. However, the two frequencies are easily resolved by the Bayesian model, which estimates the two frequencies to be: $f_1 = 0.0477 \pm 0.0004$ and $f_2 = 0.0488 \pm 0.0004$. Figure 2.4 shows a surface plot of the posterior probability in two dimensions, with f_1 and f_2 along the two axes, and probability displayed vertically. The plot is symmetrical about the $f_1 = f_2$ line and so the model gives two peaks for the above frequencies ($f_1 = 0.048$, $f_2 = 0.049$ and $f_2 = 0.048$, $f_1 = 0.049$). From the plot it can be seen that the two frequencies are well resolved by this method. The Bayesian periodogram gives an accuracy (for this example) which is an order of magnitude better than the DFT. The assumption underlying the DFT, that the data contain only one sinusoid, means that it can no longer be used to estimate frequencies accurately when more than one periodicity is present. To confirm that the data really do contain two frequencies, Bayesian analysis can be used (see model comparison section) to calculate the probability that either the 1 or 2 sinusoidal model is correct. This gives the relative probability that the data contain 2 sinusoids: P(2 frequencies)/P(1 frequency) = 10^{42}. This means that there is a virtually zero probability that the data contain only one sinusoid.

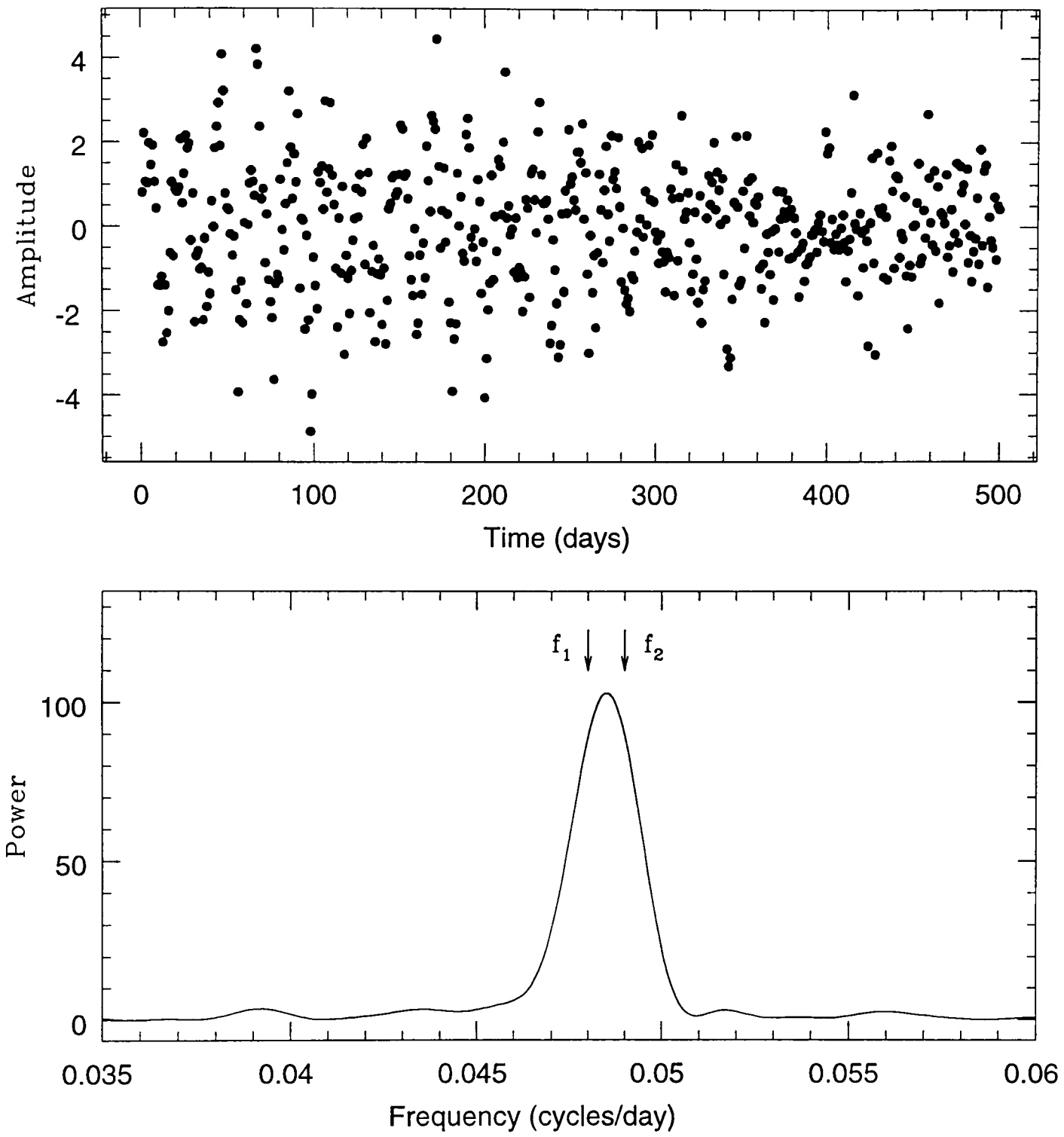


Figure 2.3: A plot of the data for the two sinusoid case (example 3), with its associated DFT. The DFT does not resolve the two frequencies present in the data (f_1 and f_2), instead it gives a broad peak roughly half way between them. The error given by the full width half maxima of the DFT is much larger than the separation between the f_1 and f_2, indicating that it cannot resolve the two frequencies in this case.

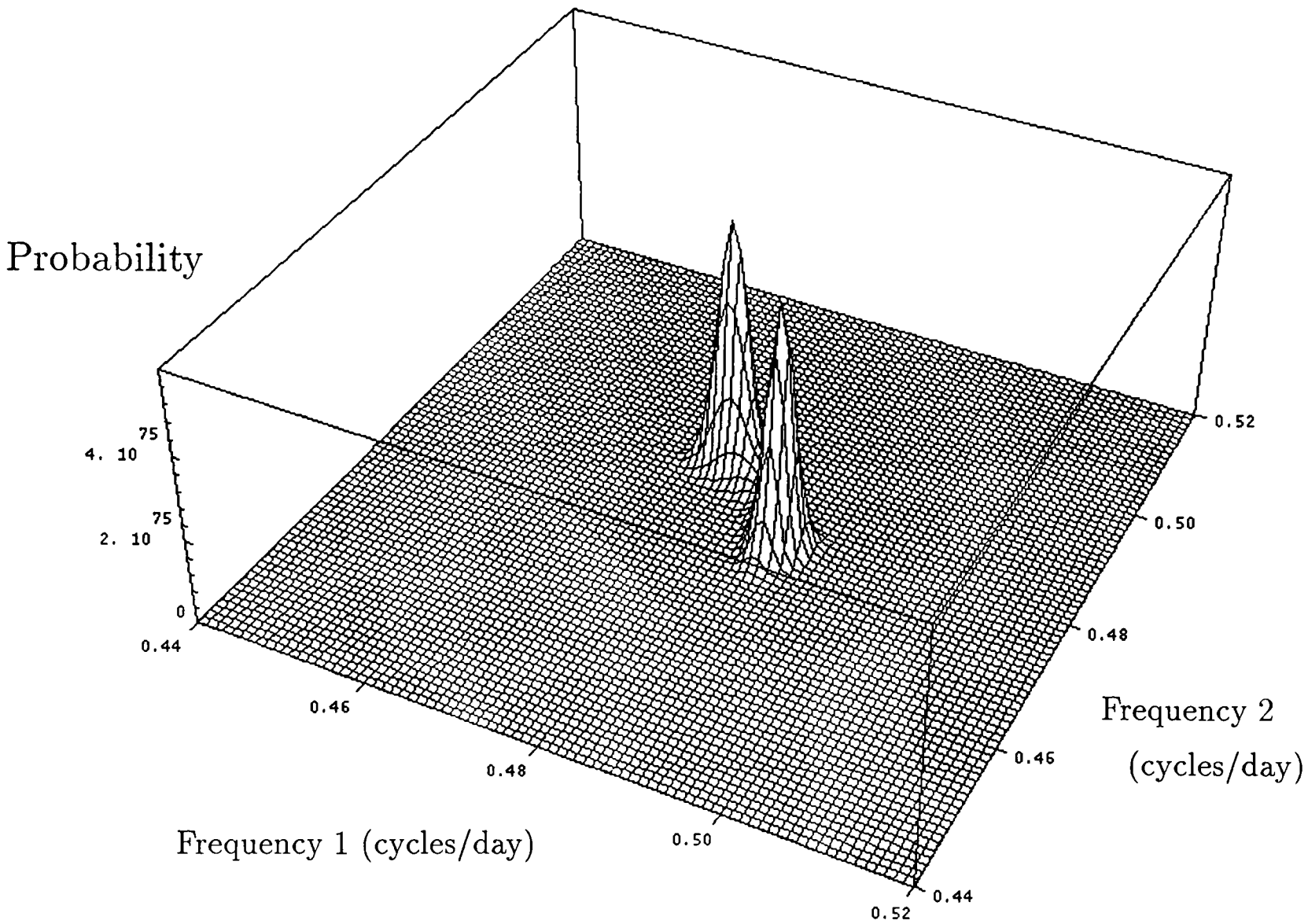


Figure 2.4: A surface plot of the posterior probability against the two frequencies f_1 and f_2 (example 3). The two peaks represent the same solution reflected in the line of symmetry $f_1 = f_2$, where $f_1 = 0.0477 \pm 0.0004$ and $f_2 = 0.0488 \pm 0.0004$. The probability reaches a maximum value of $10^{75.8}$ (the errors on the frequencies correspond to a region enclosing 99% of the probability).

Example 4: Nonuniformly Sampled Data

Now, a more realistic 2 frequency dataset is considered (example 4). The signal for this example consists of two unitary amplitude sinusoids at frequencies of 4.0 cycles/day and 4.5 cycles/day, with noise ($\sigma = 1.0$), a 1/3 duty cycle and a total of 82 points.

As can be seen from figure 2.5 the DFT becomes very distorted in this example. The data sampling creates a set of aliases for each of the two frequencies (offset by 0.5 cycles/day), which overlap as the frequencies are close. But using the 2 frequency Bayesian model (see figure 2.6) results in two clear peaks at the correct frequencies. The resulting frequencies from the Bayesian method are: $f_1 = 4.01 \pm 0.03$ and f_2 4.49 ± 0.03, with a relative probability of $10^{15.4}$ (the errors were taken from the area enclosing 99% of the probability). For comparison the next highest peak occurs at a probability of $10^{14.52}$. Again, the probability of the two frequency model can be compared to the single frequency one, which gives P(2 frequencies)/P(1 frequency) = $10^{5.05}$. This result is less clear cut than the previous result due to the poor sampling of the data, but still shows the two frequency model to be more likely.

In order to illustrate the structure of aliasing in 2 dimensions the scale of figure 2.6 was compressed to P = 10^{14}. Two peaks appear at the coordinates f_1 (f_2) = 3.5 cycles/day and f_2 (f_1) = 4.0 cycles/day. These are alias peaks which are analogous to the spurious peaks seen in DFT periodograms. Generally the alias patterns for the 2-D periodogram are seen as rows of equally spaced peaks oriented parallel to the f_1 and f_2 axes, with a separation equal to the frequency of gaps in the data. Thus, the 4.0/4.5 cycles day peak is aliased onto the 4.0/3.5 cycles/day peak. More complex gap structures will result in more confused alias patterns.

2.2.5 The General Model

All of the above analysis was based on two specific models, that of a single or double sinusoidal modulation. It is possible to generalize the analysis to any model function. The model can be defined more generally to be the sum of a set of m functions G with amplitudes B_j, which for time series analysis will normally be a series of sinusoids of different frequencies.

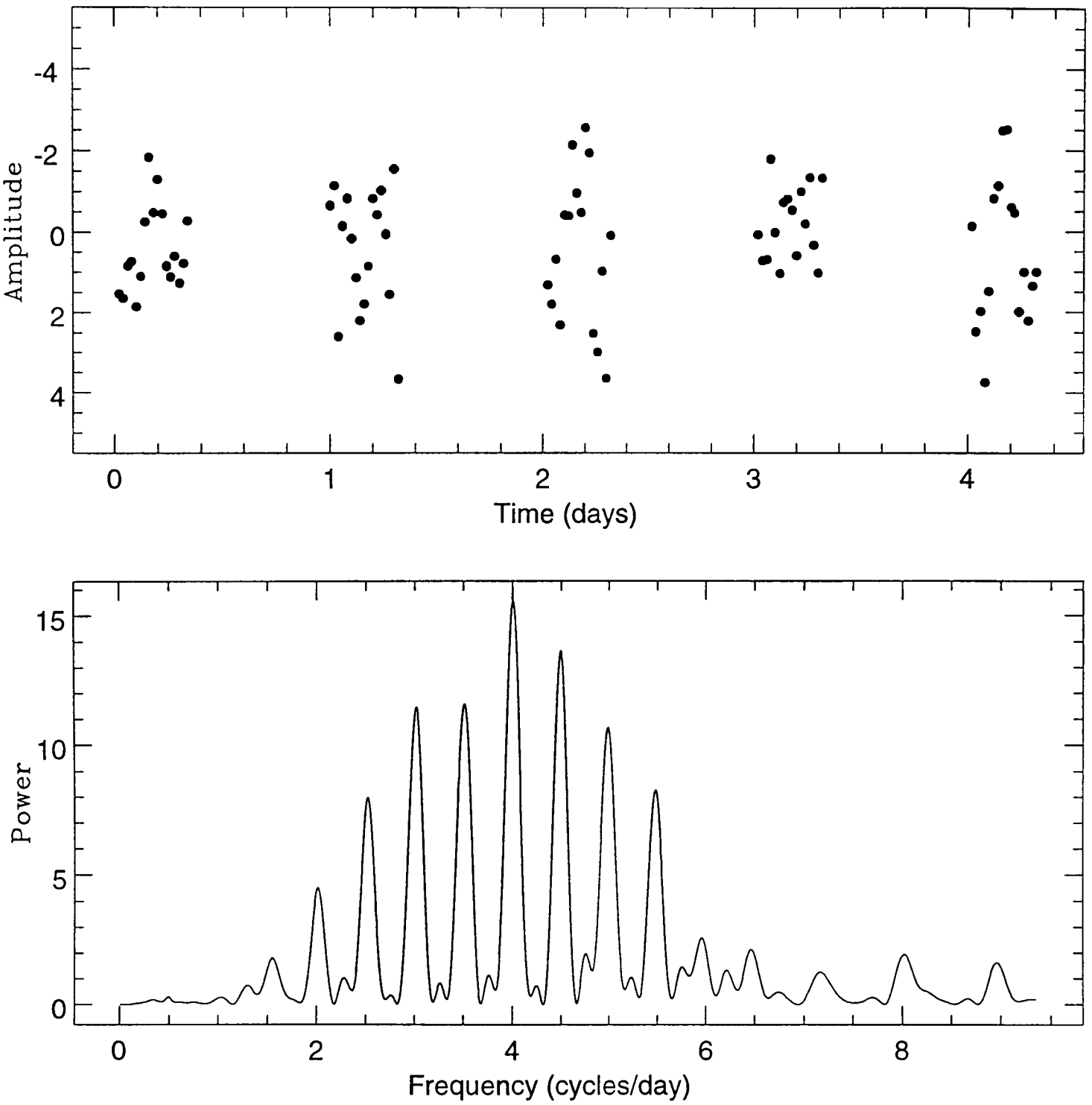


Figure 2.5: The DFT for the 2 frequency case with a 1/3 duty cycle and a S/N ratio of 1. The data do not display any obvious modulation and the resulting DFT does not give a clear indication as to whether there are 1 or 2 frequencies present. The aliasing structure makes it impossible to decide which are the real frequencies, and which are spurious.

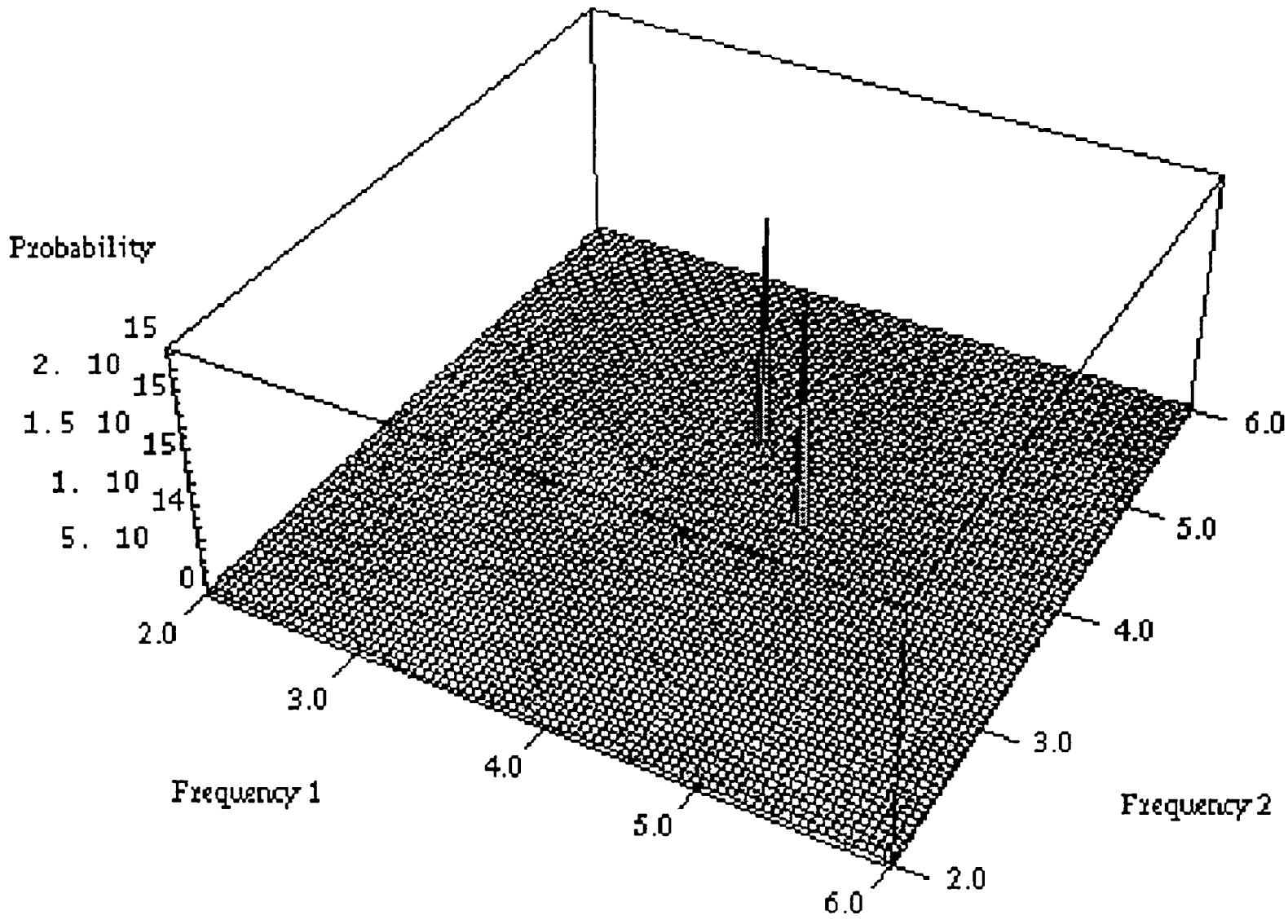


Figure 2.6: The 2 frequency Bayesian surface plot for example 4. The axis of symmetry $f_1 = f_2$ results in two sharp peaks at 4.0/4.5 cycles/day. Even though the data are poorly sampled, the alias peaks have a probability at least one order of magnitude less than the two maxima shown.

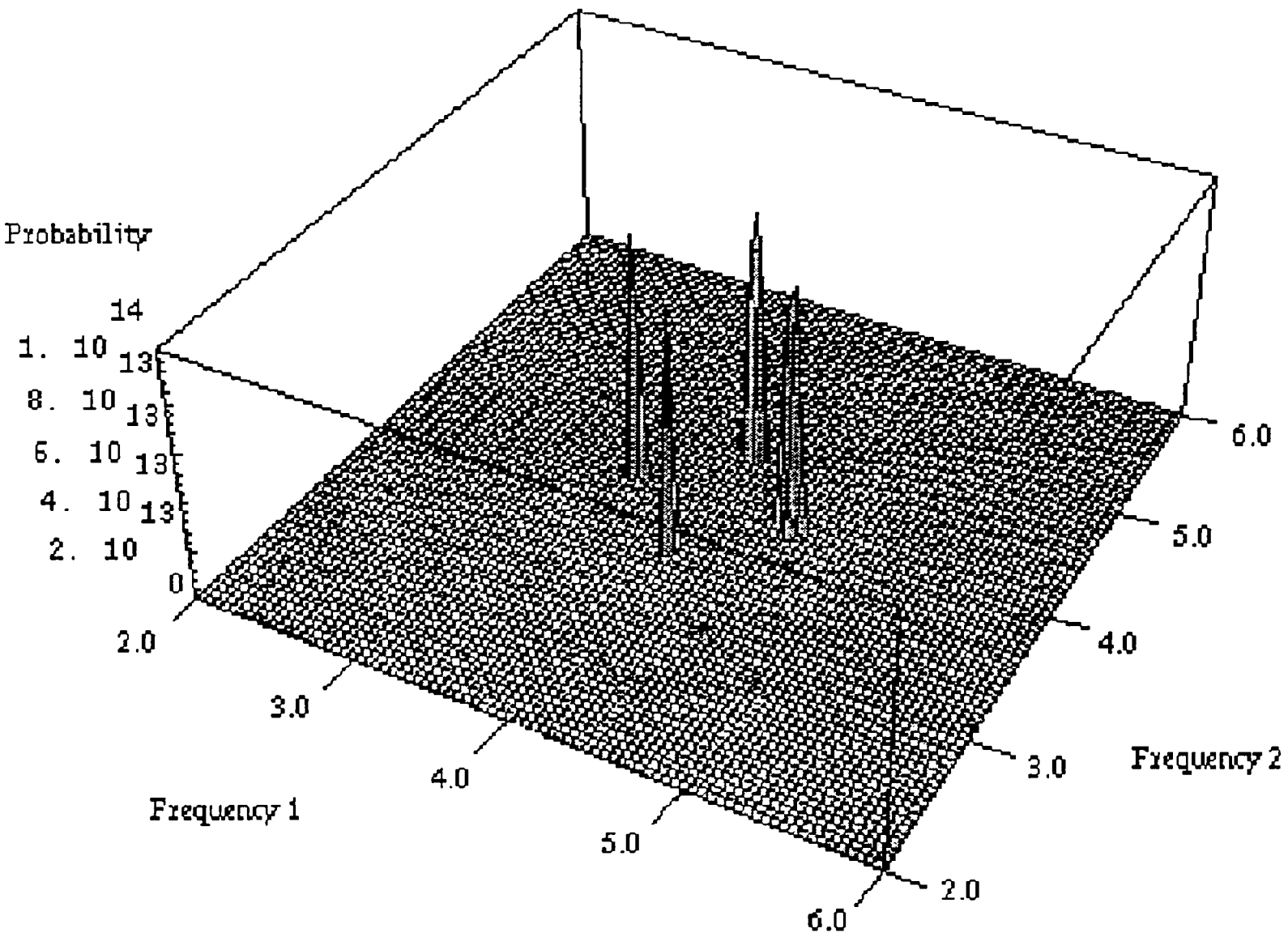


Figure 2.7: The 2 frequency Bayesian surface plot, for example 4, with the probability axis compressed. The axis of symmetry $f_1 = f_2$ results in two large peaks at the frequencies 4.0/4.5 cycles/day, as well as two alias peaks at 3.5/4.0 cycles/day.

$$f(t) = \sum_{j=1}^{m} B_j G_j(t, \{\omega\})$$

This gives the posterior probability for a set of m functions:

$$P(\{\omega\}|D, I) \propto \left[1 - \frac{m\overline{h^2}}{N\overline{d^2}}\right]^{(m-N)/2}$$

This is the posterior probability for m functions, ω represents the parameters of interest, which need not necessarily be frequencies. As before, N is the total number of data points, and $\overline{d^2}$ and $\overline{h^2}$ are given by the following summations:

$$\overline{d^2} = \frac{1}{N}\sum_{i=1}^{N} d_i^2$$

$$\overline{h^2} = \frac{1}{m}\sum_{j=1}^{m} h_j^2$$

and,

$$h_j = \sum_{i=1}^{N} d_i H_j(t_i)$$

$$H_j(t_i) = \frac{1}{\sqrt{\lambda_j}}\sum_{j=1}^{m} e_{jk} G_k(t)$$

Where λ_j and e_{jk} are the eigenvalues and eigenvectors of the matrix g_{jk}, which is simply the matrix of the model functions:

$$g_{jk} = \sum_{i=1}^{N} G_j(t_i) G_k(t_i)$$

Essentially this definition is exactly the same as the definition for the discrete fourier transform, except generalized to a set of m functions. In the case where m=1 the method produces the DFT of the data (exponentiated). For m=2 the model produces the 2 frequency periodogram of the data, and so on.

2.2.6 Multiple Frequencies

The general model can be used to estimate frequencies and the other subsidiary parameters for the multiple frequency case. Firstly, the r model functions are defined to be sinusoids with frequencies ω_r.

$$f(t) = \sum_{j=1}^{r} B_j \cos \omega_j t + \sum_{j=1}^{r} B_{r+j} \sin \omega_j t$$

Now it is necessary to determine a method to search for all the frequencies present in the dataset. The amount of time taken to search the frequency parameter space increases rapidly with the number of frequencies, and so for a complex model the time required to search for the most probable solution becomes unfeasibly large. Therefore, the method used is to test each model in turn. Initially, the data are tested with the single sinusoid model, and the most probable frequency is then found (within the Nyquist range: $1/2T$ up to $1/2\Delta T$). This is then repeated for the two sinusoidal model. To discriminate between these two models the above model selection process is used, and if the relative probability for the two models is greater than 1 the two sinusoid model is accepted. This continues to higher and higher complexity models, until the data reject one. Alternatively, one can increase the complexity of the model until the residual signal left in the data, after the model of the data has been subtracted, falls below a level $\sim 3\sigma$ (these two procedures are mathematically identical). The model comparison equation contains certain 'Occam' factors which automatically penalize the more complex models and so the above procedure will find the simplest model consistent with the data. Another method used by Bretthorst (1988) marginalizes over the frequencies of the model by integrating them out of the probability. For a model with m frequencies it would be convenient to display them all on a single periodogram, and so this method, in effect, estimates the m frequencies and then calculates the power spectral density which can be displayed versus one ω parameter. This means that a multiple sinusoid model can be displayed on a simple 1-D periodogram plot. For the data here, there are only a few periodicities present in the data and so the model selection method is used.

2.2.7 Model Comparison

In order to analyze the data using different models, it is important to be able to compare these models to see which fit the data better. In terms of probabilities what is needed is the probability of model f_i relative to model f_j. The probability of model f_j is given by:

$$P(f_j|D,I) = \frac{P(f_j|I)P(D|f_j,I)}{P(D|I)}$$

the relative probabilities of f_j and f_i are,

$$\frac{P(f_i|D,I)}{P(f_j|D,I)} = \frac{P(f_i|I)}{P(f_j|I)}\frac{P(D|f_i,I)}{P(D|f_j,I)}$$

Thus the factor $P(D|I)$ can be ignored in this calculation as it is a normalization constant, and cancels out when two different models are compared. The direct probability $P(D|f_j,I)$ can be calculated as before and the prior information about the model parameters $P(f_j|I)$, is taken to be very uninformative. For a detailed description of model comparison methods using Bayesian statistics, see Bretthorst (1990). Integrating over the nuisance parameters gives:

$$P(D|f_j,I) = \frac{\Gamma(m/2)}{2\log(R_\delta)}\left[\frac{m\overline{h^2}(\hat{\omega})}{2}\right]^{-m/2}\frac{\Gamma(r/2)}{2\log(R_\gamma)}\left[\frac{r\overline{\omega^2}}{2}\right]^{-r/2}\nu_1{}^{-1/2}\ldots\nu_r{}^{-1/2}$$
$$\times\ \frac{\Gamma([N-m-r]/2)}{2\log(R_\sigma)}\left[\frac{N\overline{d^2}-m\overline{h^2}(\{\hat{\omega}\})}{2}\right]^{(m+r-N)/2}$$

For this analysis, m is the number of sinusoids used in a given model ($r = 2m$); $\overline{h^2}(\hat{\omega})$ is the value of $\overline{h^2}(\omega)$ taken at the peak of the probability distribution and $\overline{\omega^2} = 1/m\sum_{j=1}^{m}\omega_j$. The function Γ is the gamma function, and ν_r are the eigenvalues of the matrix given by:

$$b_{jk} = -\frac{m}{2}\frac{\partial^2\overline{h^2}}{\partial\omega_j\partial\omega_k}$$

R_δ, R_γ and R_σ are constants which give a measure of the uncertainty in the amplitude of the signal, its frequency and the noise present in the data respectively. These cancel out when two models are compared and so can be ignored for model comparison methods. For a proof of this form of the model comparison equation see the Appendix.

What is normally needed for time series analysis is to compare the jth model with the $(j+1)$th model, to determine the probability that there is one more sinusoid present in the data. The above equation gives the probability of the two models and so it is the ratio of the two, i.e. $P(D|f_{j+1}, I)/P(D|f_j, I)$ that is calculated. This gives the probability (known as the odds ratio) that the $(j+1)$th model is correct compared to the jth model. Thus, successively more complex models can be tested against the data. The above equation contains several factors to penalize more complex models, and so will only predict that a model is correct if it fits the data significantly better than more simple models.

In example 3, the odds ratio of the two frequency model compared to the one frequency model ($P(D|f_2, I)/P(D|f_1, I)$) was 10^{37}, if the noise is assumed to be unknown, and 10^{42} if the correct noise level is put into the calculation ($\sigma_N = 1.0$). This means there is one chance in 10^{42} that the data are only singly periodic. To test the robustness of the model comparison method, the posterior odds ratio for the three frequency model relative to the two frequency model can be calculated. This gives a relative probability of 10^{-4}, and so the Bayesian method predicts that there are only two frequencies present in the data. It rejects the one and three frequency models, and has therefore successfully characterized the data.

2.2.8 Parameter Estimation and Resolving Power

Once the most probable set of frequencies have been found it is possible to estimate all of the other parameters of the model. The following equations give the estimates of these parameters, but it must be stressed that there is an implicit dependence on the frequency in these equations. The parameters must be evaluated at the peak of the probability distribution, and so all values of $\overline{h^2}$ are taken at the maximum of the periodogram. Firstly, the estimate of the noise variance is given by:

$$\sigma^2 = \frac{N}{N-m-2}\left[\overline{d^2} - \frac{m\overline{h^2}}{N}\right](1 \pm e)$$

where,

$$e = \sqrt{2/(N-m-4)}$$

In the single sinusoidal case the model is the sum of a sine and cosine, to define the phase and simplify its removal from the analysis. Thus, there are two individual amplitudes $\hat{B}_1$ and $\hat{B}_2$ for each frequency, meaning that the total amplitude $\hat{B}$ of each sinusoid is:

$$\hat{B}^2 = \hat{B_1}^2 + \hat{B_2}^2$$

the amplitudes B_k are equal to:

$$E(B_k|\{\omega\}, \sigma, D, I) = B_k = \sum_{j=1}^{m} \frac{h_j e_{jk}}{\sqrt{\lambda_j}}$$

The definitions of h_j, e_{jk}, λ_j were described in the previous section. From the above equations it is possible to calculate all of the amplitudes in the model, as well as the phase which is simply $\arctan(\hat{B}_1/\hat{B}_2)$.

It is also possible to calculate the accuracy of the frequency estimate obtained. This is derived from the curvature of the (assumed) Gaussian peak (see Jaynes, 1987):

$$\delta f = \frac{\sigma}{2\pi \hat{B} T}\sqrt{48/N}$$

This shows that the accuracy of the frequency estimate depends on the signal-to-noise ratio of the data and even more strongly on the number of data points. The above equation differs from the Rayleigh limit of $1/2T$, which is only strictly a rule of thumb and does not apply under certain conditions depending on the values of N and $\hat{B}$. This equation can be used to estimate the error on the frequency, but it implicitly assumes that the correct model has been used on the data. A more conservative estimate is therefore used here: the width of the peak enclosing 99% of the probability is taken to define the error on the frequency in all of the analysis that follows.

2.2.9 An Ellipsoidal Model

A type of modulation commonly found in SXT systems is the ellipsoidal modulation. This is caused by the varying aspect of the secondary star, combined with limb and gravity darkening.

The resulting light curve shape is basically sinusoidal, but with minima of alternating sizes. The usual way to analyze such data is just to use a straightforward DFT, under the assumption that the difference in the minima is small. Normally the difference in the minima is less than $\sim 0.05^m$ and the size of the overall modulation is $\sim 0.2^m$. But, using Bayesian analysis allows any prior information that is known about the data to be incorporated and so improves the estimates of frequency and other parameters.

The ellipsoidal modulation is equivalent to a sinusoid of frequency $2f$ which lags 90° in phase, relative to a sinusoid at a frequency f (see figure 2.8). It is also usual for the difference in minima to be small, which means that the amplitude of the sinusoid at the frequency f will be smaller than that at $2f$ (the ratio is typically in the range 5–10). This information can be put into the 2 frequency model to constrain the frequency and phase of the 2 sinusoids relative to each other ($f_2 = 2f_1$ and $\phi_2 = \phi_1 - 90°$). The resulting model for the data is therefore:

$$f_1(t_i) = A\ \sin(2\pi f_1 t_i + \phi_1) + B\ \sin(2\pi[2f_1]t_i + \phi_1 - \pi/2)$$

No information about the size of the amplitudes at the two frequencies has been assumed to make the method as general as possible, only the relative phases and frequencies are constrained. The maximum amplitude of ellipsoidal modulation that is physically possible is $\approx 0.3^m$, this could be included in the model to improve the parameter estimates, but would only make a small difference to the result. The common method of taking the DFT and halving the estimate of the period obtained is equivalent to assuming that the minima are of equal depths (A=0). To estimate the frequency of the ellipsoidal modulation from this model, the posterior probability has to be calculated:

$$P(f|\sigma, D, I) \propto \exp\left(\frac{-\overline{d^2} + \overline{h^2}}{2\sigma^2}\right)$$

or, if the level of noise is unknown:

$$P(f|D, I) \propto \left(1 - \frac{4\overline{h^2}}{N\overline{d^2}}\right)^{(4-N)/2}$$

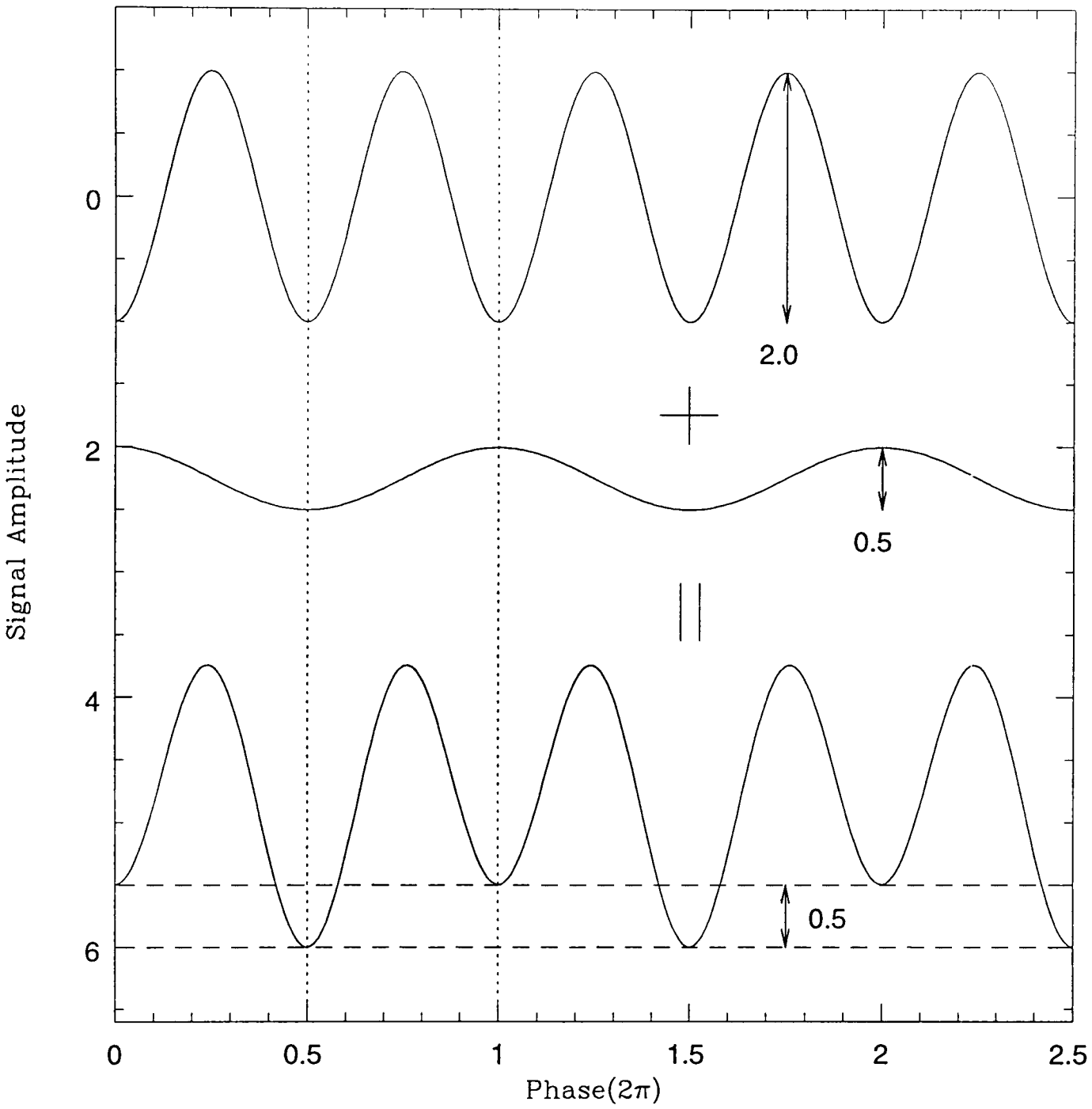


Figure 2.8: The ellipsoidal modulation (bottom) is equivalent to the sum of two sinusoids at frequencies $2f$ (top) and f (middle), which are offset by $1/4$ in phase relative to each other. The amplitude of the sinusoid with frequency $2f$ governs the overall amplitude of the modulation and the amplitude of the sinusoid with frequency f determines the difference in the depths of the minima.

which is essentially the same equation as for the 2 frequency model. $\overline{d^2}$ was defined earlier and $\overline{h^2}$ is given by:

$$\overline{h^2} = ds^{(1)^2} + dc^{(1)^2} + ds^{(2)^2} + dc^{(2)^2}$$

The terms $ds^{(1)}$ are the projections of the data onto the model functions, and are given by:

$$ds^{(1)} = \sum_{i=1}^{N} d_i \sin[2\pi f_1 t_i] \qquad dc^{(1)} = \sum_{i=1}^{N} d_i \cos[2\pi f_1 t_i]$$

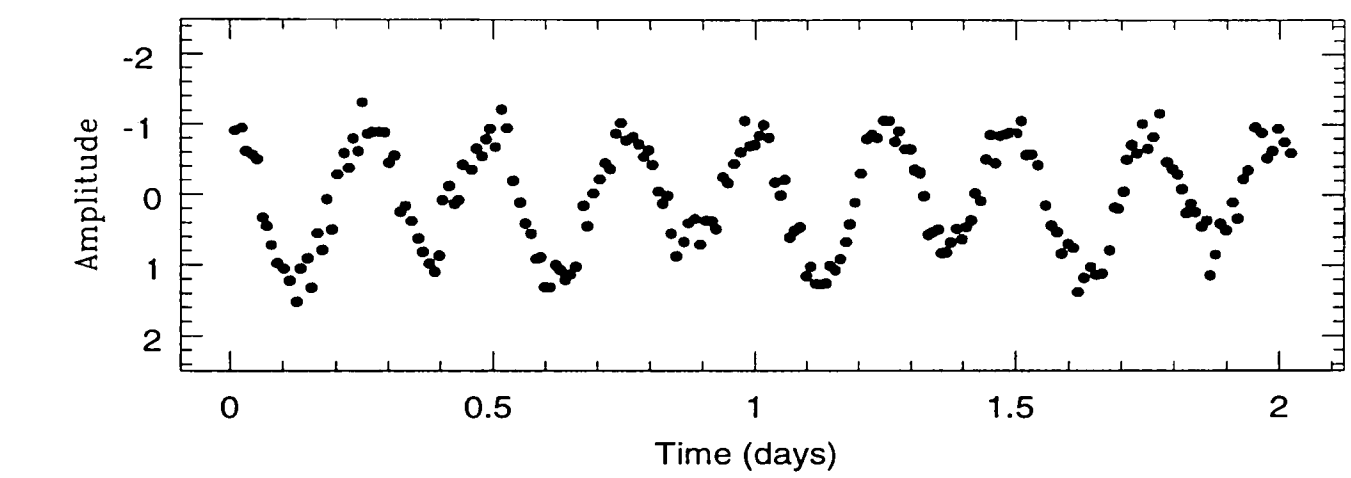
$$ds^{(2)} = \sum_{i=1}^{N} d_i \sin[2\pi (f_2 t_i) - \pi/2] \qquad dc^{(2)} = \sum_{i=1}^{N} d_i \cos[2\pi (f_2 t_i) - \pi/2]$$

These equations are the same as for the simple two dimensional case, except that they are projected against slightly modified sine functions. The above equations can be modified by the transformation of variables that was used for the general model equation, to allow for the effect of uneven spacing, which slightly alters the ellipsoidal periodogram. This method can now be compared with the DFT, remembering that the frequency f is the one of interest, rather than $2f$, as the orbital period of binaries that show an ellipsoidal modulation is equal to $1/f$.

2.2.10 Example: Uniformly Sampled Ellipsoidal Dataset

A dataset was generated with 200 points, evenly spaced ($\Delta T = 0.01$ days), except for a random 'jitter' of 0.002 to recreate the uneveness of typical astronomical time series sampling, and a noise level of $\sigma_N = 0.2$. To simulate the ellipsoidal modulation, the data were constructed as the sum of two sinusoids with amplitudes 1.0 and 0.3, and frequencies 4.00 cycles/day and 2.0 cycles/day, respectively. The simulated data and posterior probability versus frequency, for the ellipsoidal model outlined above, and the periodogram for the DFT are shown in figure 2.9.

The data for this example can be seen by eye to have alternating size minima and an overall period of 2 cycles/day. The DFT produces a large peak at 4.01 ± 0.4 cycles/day and cannot



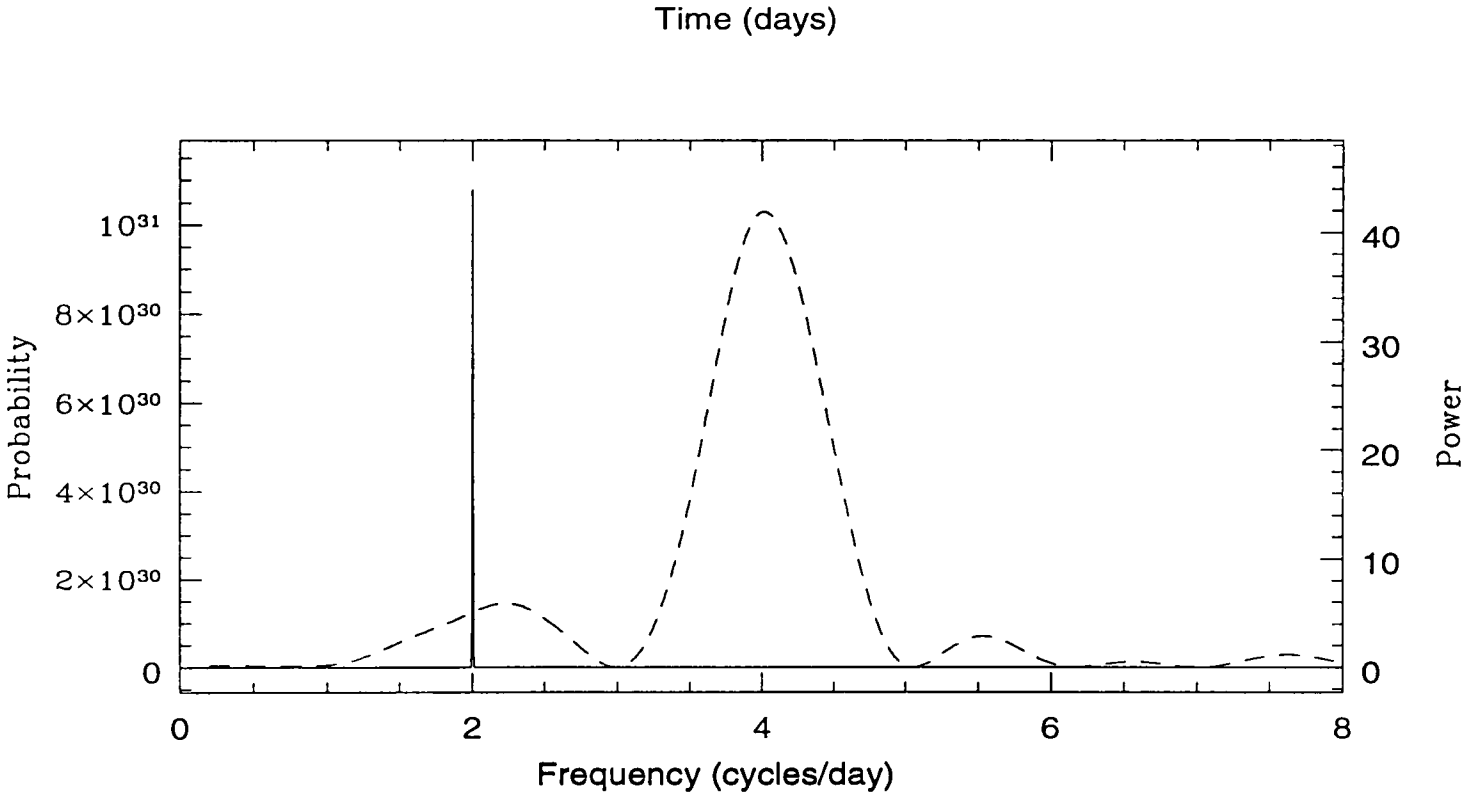


Figure 2.9: Data, posterior probability functions for the ellipsoidal model (solid line) and the DFT of the data (dashed line). The DFT does not produce a peak at the orbital frequency, but the ellipsoidal model predicts that the frequency at 2 cycles/day is more than 10^{32} more likely to be correct than any other frequency.

determine whether the data consist of one or two frequencies. The ellipsoidal model produces a much clearer result, though, with a sharp maximum at f = 2.001 ± 0.007 cycles/day. The next highest peak at 4 cycles/day is a factor 10^{32} smaller than the main peak, which indicates that it is much less probable that this is the correct period.

A second dataset was generated to illustrate the effects of aliasing on this problem. The data sampling convolved with the two frequencies present in the ellipsoidal modulation will cause the resulting periodogram to become very confused. Using the new model, however, should reduce this problem and increase the accuracy of frequency estimation. This dataset was exactly the same as in the previous example, except that a 1/3 duty cycle was imposed (equivalent to an 8 hour observing run each 24 hours). The data and model periodograms for these data are shown in figure 2.10. The data are less obviously due to an ellipsoidal modulation and the DFT has become very confused due to the interaction of the aliases and the ellipsoidal variation. Despite this, the ellipsoidal Bayesian model is still free of aliases and has a sharp peak at 2.002 ± 0.003. The DFT still gives 4 cycles/day as the most likely frequency, and has become very unclear with several alias peaks.

2.3 The Problem of Aliasing

2.3.1 Introduction

The previous section dealt with a problem that affects most methods of frequency analysis, i.e. the underlying assumption that the data are singly periodic. But, the situation is made worse by the problem of aliasing, which is caused by unevenly spaced data. Examples 2 and 4 in section 2.2 showed how regularly spaced gaps can cause the resulting periodograms to become confused and difficult to interpret. In figure 2.11 an example of aliasing is shown. The true frequency of the data is 2.4 cycles/day which is sampled once each day, producing aliases at 2.4 ± 1 cycles/day. If the sampling becomes more random, however, these aliases can be eliminated. The constraint of nightly observing means that this is not possible and so aliasing is often a problem. Even though each run may last 8 hours out of each 24, this is still not enough to provide effective sampling.

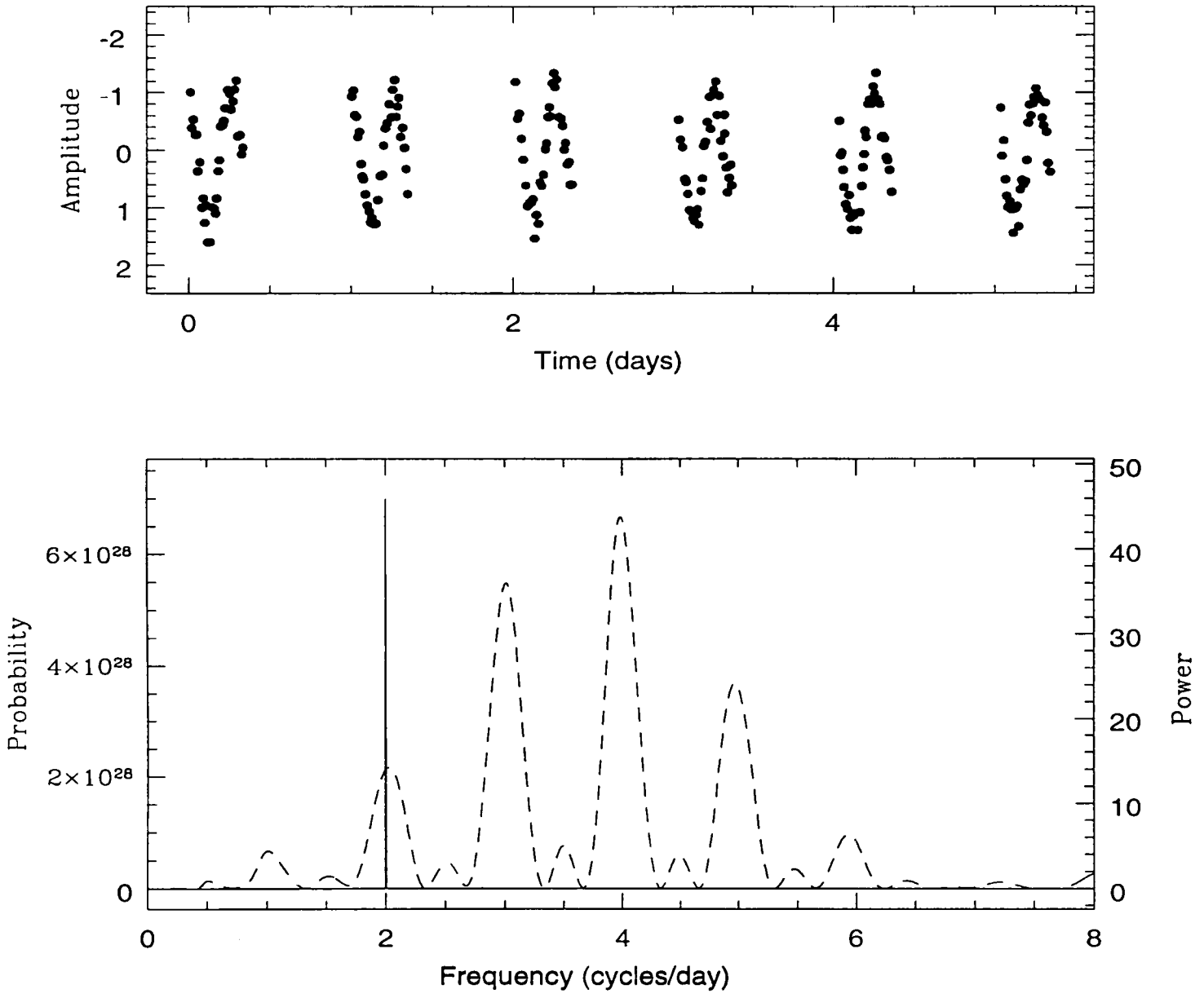


Figure 2.10: Second example – the data (top) and posterior probability periodogram for the ellipsoidal model (solid line) and DFT (dashed line). The next highest peak is a factor $10^{17.8}$ below the peak at 2 cycles/day.

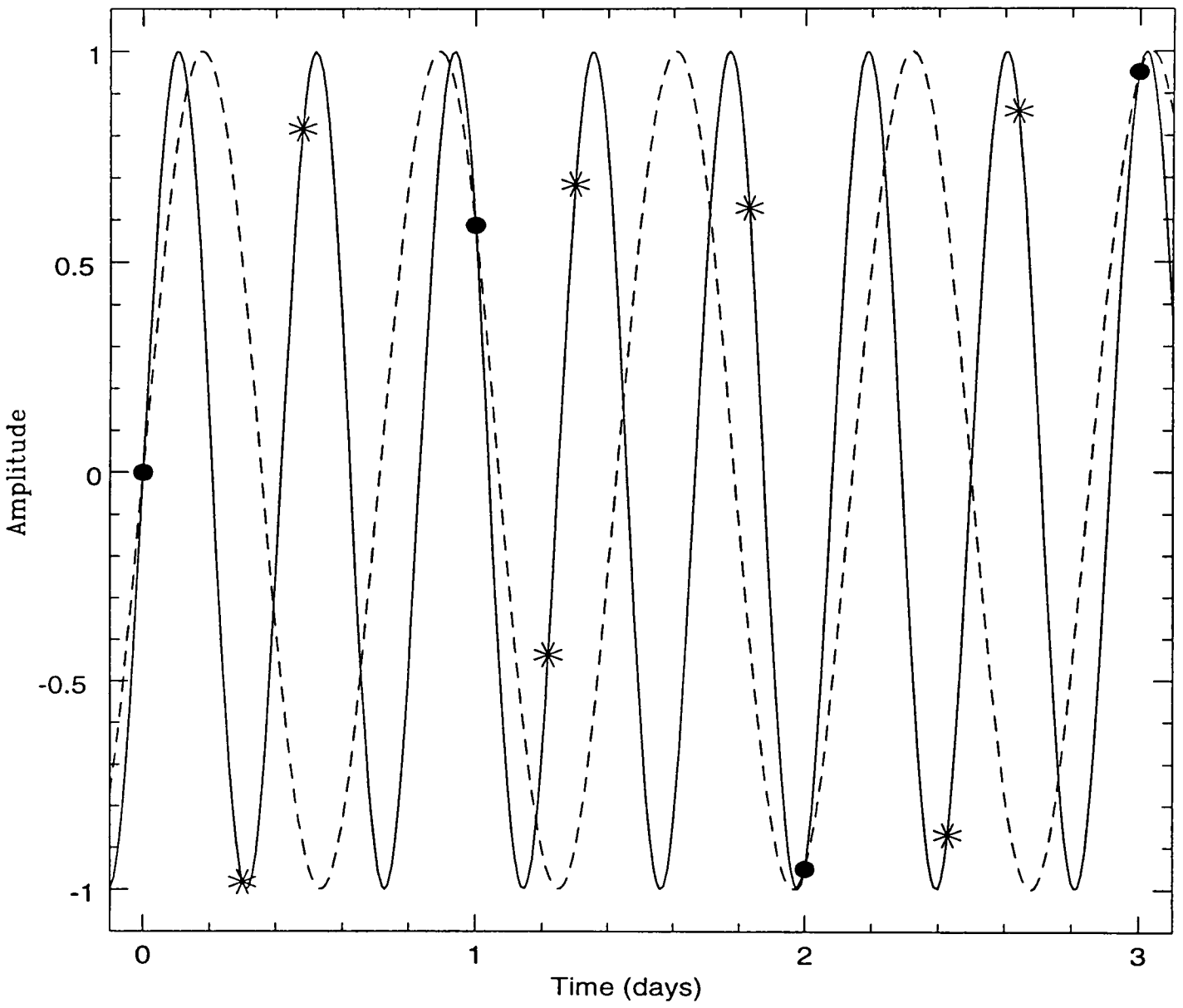


Figure 2.11: Aliasing is caused by regularly space gaps in time series datasets. Here a frequency at 2.4 cycles/day (solid line) is sampled once each day (filled circles) as in most ground based astronomical observations. This means that it is impossible to distinguish between this frequency and its aliases, one of which is shown – 1.4 cycles/day (dashed line). If the observing is more random and covers more phases (asterisk), then it becomes possible to rule out the aliases as spurious.

A common method of analyzing the effect of the data sampling on the resulting periodogram is to look at the window function. All of the data values are set to one, whilst keeping the same timings. Then, a periodogram of this 'windowed' data produces the window function. It is therefore a measure of the uneveness of the data sampling. Usually the window function has a maximum near zero frequency, which indicates that a lot of power will stay near the signal frequency. Then, with increasing frequency a set of equally spaced peaks appear, separated by 1 cycle/day, which indicate that a signal will leak power onto a series of one day aliases.

Another way of testing the behaviour of the periodogram is to input Gaussian, white noise with the same temporal sampling as the data, and then take its discrete fourier transform. As white noise contains equal power at all frequencies, this indicates which frequencies will leak power into sidelobes, and so this can therefore be a useful tool in understanding the data.

A method to alleviate the aliasing problem is described which takes into account the sampling of the data. Previously, the CLEAN algorithm was used to correctly identify several frequencies in unevenly spaced datasets. This works well under a variety of conditions, but unfortunately includes a questionable assumption – that the highest peak in the periodogram is the correct frequency. As described in section 2.1.4, CLEAN repeatedly subtracts the frequency which has the largest power in the periodogram. This works well when there is no noise present in the data, but low signal-to-noise levels can cause alias peaks to become larger than true signal peaks (Scargle, 1982). This causes CLEAN to subtract the wrong periodicities and so distorts the resulting spectrum.

A related method was devised by Gray and Desikachary (1973). Instead of just considering the central frequency of a peak in the periodogram and subtracting, they used the window function of the data, and shifted it until it corresponded to the pattern seen in the observed periodogram. They then subtracted this from the periodogram and repeated the process for several frequencies. This method also suffers from the assumption that the largest peak in the periodogram is correct and so can still fail in the presence of noisy data.

2.3.2 Best Fit Periodogram Algorithm

To alleviate the problem of the effects of noise and data sampling a different method has been adopted, which only assumes that the modulation of the data is sinusoidal. This is a fairly robust assumption, as a wide variety of modulations have most of the their power at the fundamental frequency and so can be approximated in this way. Firstly, the DFT of the original data is computed over the frequency range of interest. After this, an artificial dataset is constructed which has the same data sampling as the original measured data, consisting of a sinusoid at a set frequency and noise at a specified level. The DFT of this artificial dataset is computed and compared to the DFT of the original data by means of a chi-squared fit parameter. This procedure is repeated over a range of frequencies to give a periodogram of the fit.

When the fit is closest the parameter of the fit has a maximum, indicating which frequency best reproduces the observed periodogram. What this algorithm does in effect is to find the dataset that best reproduces the observed alias pattern, and so it takes into account the convolution of noise and signal with the data sampling.

The fit parameter is a measure of the goodness of fit of the artificial periodogram compared to the observed periodogram. The goodness of fit is given by the following chi-squared statistic:

$$\chi^2_{fit}(f) = 1/N \sum_{i=1}^{N} \frac{(P_i^{obs}(f) - P_i^{art}(f))^2}{P_i^{obs}(f)}$$

The chi-squared fit is summed over N points, and P^{obs} and P^{art} are the power levels of the observed and artificial periodograms, respectively. To compare this method with other periodograms such as the DFT the inverse of this statistic was used, so that the most probable frequency occurs at the maxima of the resulting $1/\chi^2$ periodogram.

The noise in the DFT periodogram has an (normal) ogive shaped cumulative distribution, which can be calculated exactly for a given dataset by using a Monte-Carlo simulation. This

distribution could be convolved with the above statistic to account for the effects of noise. This would effectively be a weight to bias the fit toward higher peaks, which are less likely to be due to noise. Unfortunately, noise also tends to shift power away from peaks in the power spectrum, but this effect is dependent on the data sampling in a complex way. These and other complications mean that it is best just to use the simple statistic given above, assigning all parts of the periodogram equal weight.

To average out the $1/\chi^2$ periodogram the statistic was calculated 5 times for each frequency. Each peak in the artificial periodogram will have a Gaussian distribution centred about some 'mean' power level and so this increases the accuracy of the estimate of this mean by approximately $\sqrt{5}$. This value was high enough to ensure that the DFT fitting routine gave consistent, accurate results. The reason that noise must be included at all in the simulation is because it has the effect of altering the size of alias peaks, tending to shift power away from the true signal (see Scargle, 1982), and altering the shape of the periodogram.

2.3.3 Estimating the Signal Amplitude

The above method requires that an estimate of the signal amplitude and noise variance be input. Normally, the noise variance can be estimated from the photometric errors of the data, but the signal amplitude needs to be estimated by some other means.

It is possible to estimate the signal amplitude in the following way. The total variance σ_y^2 of the data is calculated, and knowing the photometric error (variance $= \sigma_N^2$) the signal variance (σ_s^2) is given by:

$$\sigma_s^2 = (\sigma_y^2 - \sigma_N^2)$$

The variance of a sinusoid of unit semi-amplitude is equal to 0.5, and so given σ_s it is possible to calculate the amplitude of the modulation in the data, independent of its frequency.

$$Semi - amplitude = \sqrt{2}\,\sigma_s = \sqrt{2(\sigma_y^2 - \sigma_N^2)}$$

The error in this amplitude depends on the error in σ_s^2, which in turn depends upon the error on σ_y^2. This is equal to:

$$error(\sigma_y^2) = \sqrt{\frac{\mu_4 - (\mu_2)^2}{N}}$$

Where μ_4 is the fourth moment of the data, i.e. the sum of the fourth power of the deviations from the mean divided by N, and μ_2 is the second moment, equal to σ_y^2. The fractional error can then be propagated through the analysis to obtain the error on the amplitude.

2.3.4 Example – Unevenly Sample Data

To test the behaviour of the $1/\chi^2$ statistic, a dataset similar to those in section 2.2 was generated. A total of 250 points were generated with a sinusoidal frequency of 4 cycles/day (6 hours), with 2/3 of the duty cycle removed (leaving 84 points). The signal semi-amplitude of 1.0 and noise level of 0.2 gives a S/N ratio of 5 for this example. The DFT fitting algorithm requires the amplitude and noise variance for the simulated datasets to be input and so the actual values were used. The estimate of the semi-amplitude using the above procedure gives a value of 1.02 ± 0.05 showing that it does accurately predict the true amplitude along with its associated error.

Figure 2.12 shows the periodogram of this artificial dataset (top), and has a typical aliasing pattern centred around the correct frequency of 4 cycles/day. Below this is the best fit periodogram, which is almost identical and shows all of the gross features of the original periodogram. At the bottom of figure 2.12 is the $1/\chi^2$ statistic. This peaks very strongly at 4.00 ± 0.01 cycles/day and has very little alias structure. The reason for the large error compared to the discrepancy of the result is that the full width half maximum was used to estimate the error. In the section on Bayesian analysis this was seen to be an overestimate of the error in many cases, and is only used in this example to enable comparison with the estimates given by the DFT.

The peak of the $1/\chi^2$ statistic has smaller full-width half maximum than the DFT. This indicates that the error on the frequency is smaller (the DFT gives: $f = 3.998 \pm 0.09$ cycles/day). This is because the whole periodogram is used to obtain the frequency, not just the area around the largest peak.

All of the alias structure in the above periodograms is caused by the 4 cycles/day modulation convolved with the sampling of the data (plus any noise that is present). It is therefore reproducible and this is used to obtain a clearer and more accurate frequency estimate from the data. A related method was used by Gray and Desikachary (1973), which subtracted the alias pattern of a signal from the periodogram and therefore also increased on the normal accuracy obtainable. But, as with CLEAN, this method suffers from the assumption that the highest peak is the correct one.

2.3.5 Violating the Assumptions

To test the robustness of the method, the same data were used, but the amplitude and noise were input into the algorithm as 10.0 and 0.2, respectively. This amplitude is a factor 10 larger than the true amplitude, and so this will test the effect of uncertainty in the size of the modulation on the algorithm. Figure 2.13 shows the results of this fit. The fitting routine produces as good a fit as before, and the plot of the $1/\chi^2$ statistic still shows the aliases to be very well suppressed. Overestimating the size of the modulation in the data has little effect on the results obtained. The frequency obtained in this test was f = 3.998 ± 0.01.

The main assumption of this algorithm is that the data can be represented by a single sine wave. For most types of modulation this is reasonable as they have their largest amplitude at the fundamental frequency. To test this, the behaviour of the statistic was investigated for various input functions, such as a triangle, square wave etc. These results are shown in table 2.1. The $1/\chi^2$ statistic produces slightly better estimates of the frequency than the DFT and there is considerable suppression of aliasing (the ratio R of the two highest peaks in the periodogram is smaller).

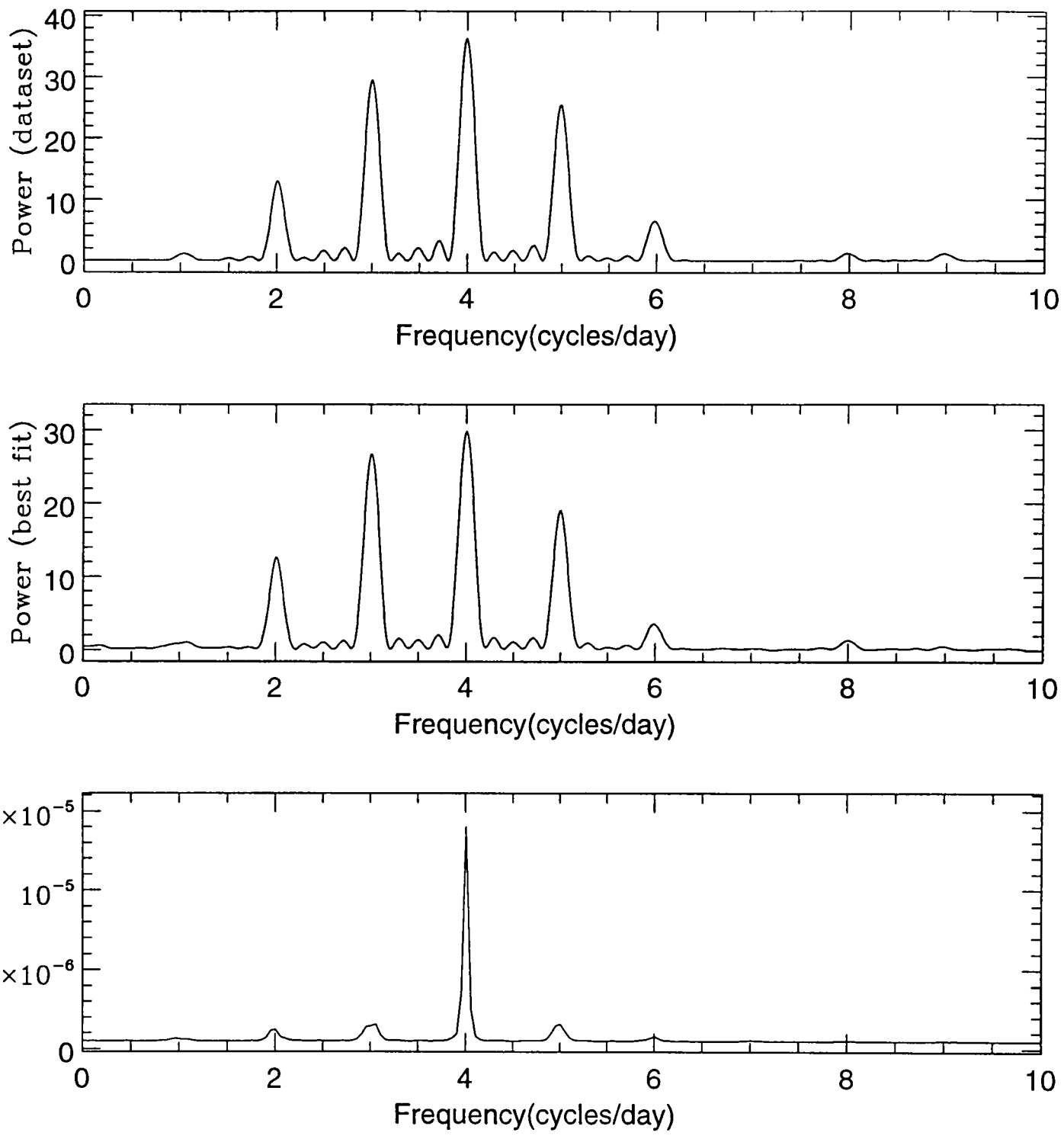


Figure 2.12: Top – the DFT periodogram for the dataset described in the text, with a modulation at 4.0 cycles/day. The usual 1 day alias structure is seen. Middle – the best fit periodogram as measured by the $1/\chi^2$ statistic, which is almost identical to the observed periodogram. Bottom – the $1/\chi^2$ statistic versus frequency. This peaks very strongly at 4.0 cycles/day and has very little alias structure.

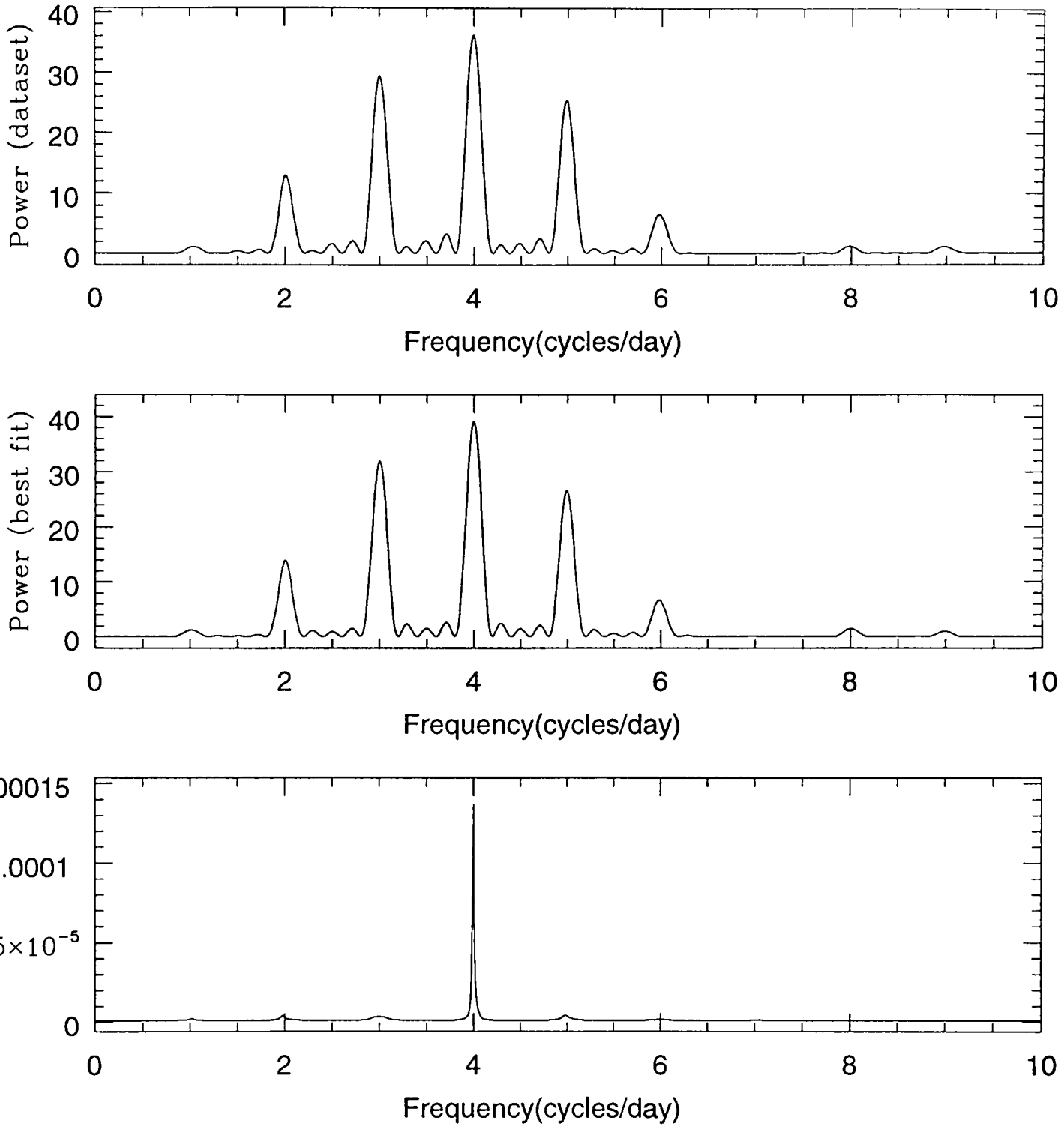


Figure 2.13: The second DFT fitting example, with the amplitude overestimated by a factor 10. Top – the DFT periodogram for the original dataset with a modulation at 4.0 cycles/day. Middle – The best fit periodogram as measured by the $1/\chi^2$ statistic. Bottom – the $1/\chi^2$ statistic versus frequency. This gives the correct frequency as before, and suppresses the aliases.

Table 2.1: Results of $1/\chi^2$ fitting for several periodic functions. The frequency estimates for the DFT and $1/\chi^2$ statistics are listed as well as the ratios, R_{DFT} and R_{1/χ^2}, of the two highest peaks in the periodograms.

Wave Type	Duty Cycle	Frequency in Original DFT	R_{DFT}	$1/\chi^2$ Frequency	R_{1/χ^2}
Sine	1	2.00 ± 0.19	0.059	1.99 ± 0.02	0.002
	1/3	1.99 ± 0.17	0.677	2.01 ± 0.03	0.089
Triangle	1	1.99 ± 0.17	0.073	1.99 ± 0.02	0.002
	1/3	1.99 ± 0.18	0.859	2.02 ± 0.02	0.24
Square	1	1.99 ± 0.18	0.116	1.99 ± 0.01	0.032
	1/3	1.97 ± 0.21	0.913	2.00 ± 0.03	0.13
Sawtooth	1	1.96 ± 0.17	0.272	2.01 ± 0.03	0.169
	1/3	1.94 ± 0.18	0.958	1.97 ± 0.05	0.568
Pulse	1	2.01 ± 0.17	0.630	2.04 ± 0.15	0.394
	1/3	2.08 ± 0.15	0.874	2.08 ± 0.06	0.459

2.3.6 Summary

The Bayesian method outlined in the previous section was basically designed to cope with multiply periodic data, but even this method can become unusable in the presence of strong aliasing. A combination of both methods has been used in this work: the Bayesian method was used for data that contained more than one sinusoid, and to check that aliasing did not cause this method to fail. The DFT fitting routine above was used to remove the effect of poor sampling. Including the $1/\chi^2$ statistic, or some derivative of it, into the overall framework of Bayesian theory would combine the robustness of this statistic, with the flexibility of Bayesian probabilities.

References

Bretthorst G.L., 1990, J. Magn. Reson., **88**, 552.

Bretthorst G.L.,*Lecture Notes on Statistics: Bayesian Spectrum Analysis and Parameter Estimation*, vol. 48, Springer-Verlag, New York, 1988.

Gauss K.F., (1963 reprint),*Theory of the Motion of the Heavenly Bodies Moving About the Sunb in Conic Sections*, Dover Publications Inc., New York.

Gray D.F. and Desikachary K., 1973, Ap.J., **181**, 523.

Jaynes E.T., 1987, 'Bayesian Spectrum and Chirp Analysis', in *Maximum Entropy and Bayesian Spectral Analysis and Estimation Problems*, ed. D.Reidel, Dordrecht-Holland, 1-37.

Legendre A.M., (1806),*Nouvelles Méthods pour la Détermination des Orbits des Cométes*,Paris.

Lomb N.R., 1976, Ap. Space Sci., **39**, 447.

O'Dell M.A., Collier Cameron A., 1993, MNRAS, **262**, 521.

Roberts D.H., Lehar J., Dreher J.W, 1987, Astron.J., **93**, 968.

Scargle J.D., 1982, Ap.J., **263**, 835.

Schuster A., 1905,*The Periodogram and its Optical Analogy*, Proceedings of the Royal Society of London, **77**, 136.

Stellingwerf R.F., 1978, Ap.J., **224**, 953.

Chapter 3

Orbital and Quasi-Periodic Optical Variations in the Black-Hole X-ray Binary V404 Cyg

3.1 Summary – Optical Photometry

R-band CCD photometry of V404 Cyg, the optical counterpart of the soft X-ray transient GS2023+338, was obtained on La Palma in July 1992, over a period of two weeks. The R-band light curve showed a clear 0.3^m ellipsoidal modulation on the established 6.5 day orbital period, but with substantial (up to 0.3^m) superposed variations occurring during each night on a timescale of 6 hours. In addition to this B,V and I-band colours were obtained once each night to investigate the colour dependence of the orbital and short period variability.

Modelling was performed on both the mean orbital light curve and its lower envelope (which presumably shows the secondary star with minimum contamination from the accretion disc). With a mass ratio already known from optical spectroscopy (Casares & Charles 1994), the best fit to the R-band light curve gave an inclination $i = 56° \pm 6°$. The model required no hot spot contribution. The weighted mean inclination, obtained from the complete set of colour photometry was $57.7° \pm 2.0°$, if it assumed no disc contribution, and $59.8° \pm 2.0°$ if the disc veiling was taken into account.

The shorter timescale of variability is consistent with the 6 hour quasi-periodicities that have been seen in outburst and during the early quiescent period (Casares and Charles, 1992; Casares et al., 1993). The nature of this period is unknown.

3.1.1 Introduction

V404 Cyg exhibited optical outbursts in 1938 and 1989 (see Charles et al. 1989), with possible additional outbursts reported in 1956 and 1979 (Richter 1989). During the 1989 outburst, V404 Cyg was identified as the optical counterpart of the X-ray transient GS2023+338, discovered by the Ginga satellite in late May 1989 (Makino et al., 1989). The mass function of V404 Cyg was subsequently found to be $6.08 \pm 0.06 M_{\odot}$ (Casares & Charles, 1994) which establishes it as the strongest black-hole candidate yet found.

V404 Cyg also displayed spectral and photometric variability on several time scales. High-speed photometry during the outburst showed evidence for a 10-min modulation (Wagner et al., 1990) and quasi-periodic oscillations (QPO) on time-scales of 3–10 min (Gotthelf et al., 1991). However, Casares et al. (1992) determined the actual orbital period to be 6.5 days from the radial velocity curve of the K0 secondary. Their conclusion that it is the binary period was confirmed by subsequent photometric observations (in I: Wagner et al., 1992; in white light: Casares et al., 1993). Furthermore, Shahbaz et al. (1994) obtained a K-band ellipsoidal light curve from which they infer a compact object mass of 12^{+3}_{-2} $M_{\odot}$. Nevertheless, there remain many puzzling features about the V404 Cyg system, such as the origin of the short period variability which has now been seen by many observers at different sites and over several years, during both outburst and quiescence.

It was possibly first detected by Haswell & Shafter (1990) as a 5.0 hour modulation in the radial velocity of the Na D_1 λ5896 absorption line. Leibowitz et al., (1991) and Udalsky & Kaluzny (1991) both reported R-band photometric modulations near 3 hours. High-resolution Hα spectroscopy in quiescence showed evidence for a 0.24 day periodicity (Casares and Charles 1992; Casares et al., 1993), but at the same time photometry in 1991 indicated that the

0.24 day modulation was very much weaker than in 1990 (Casares et al., 1993). Casares and Charles (1992) also then detected a spectroscopic modulation (*a posteriori*) in the Hβ and HeII emission lines during their outburst observations. Here, far more extensive CCD photometry of V404 Cyg is reported, undertaken in 1992 to further investigate the nature of this unusual modulation, which is clearly not directly related to the orbital period.

3.1.2 Observations

A total of 16 days of continuous R-band photometry were obtained, with the GEC CCD camera (EEV7) on the 1.0 m Jacobus Kapteyn Telescope on La Palma from June 27 to July 12, 1992. The seeing was typically 1–1.5 arcseconds, with little or no cirrus during the run. A total of 1083 frames of V404 Cyg were taken, which were de-biased and flat-fielded using standard procedures. The typical integration time was 200 s and the duration of the observations was approximately 6 hours per night. Straightforward aperture photometry was performed with a circular aperture of 11 pixels radius (3.6 arcseconds) around both V404 Cyg and its close contaminating star (separation 1.4 arcseconds). Several stars of comparable brightness to V404 Cyg were used as check stars, and all were referred to a nearby bright reference star. A correction was then made for the contribution of the contaminating star (measured at times of very good seeing using DAOPHOT) to obtain the true relative brightness of V404 Cyg itself. This was done by converting the individual magnitudes of V404 Cyg to fluxes and then subtracting the contribution of the companion star to this flux (the reason for this is that as the brightness of V404 Cyg varies relative to the companion the correction needed is constant in terms of flux, but not magnitude). The flux was then converted back into magnitudes and the process repeated for all of the data points.

3.1.3 Ellipsoidal Modelling of the R-band Light Curve

The R-band light curve of V404 Cyg relative to the reference star is shown in figure 3.1. An overall variation of 0.3^m can clearly be seen, as well as a short term variation within each night. For comparison the random scatter of the reference star relative to a nearby check star is $\sigma = 0.017^m$ (both of which are of a comparable brightness to V404 Cyg).

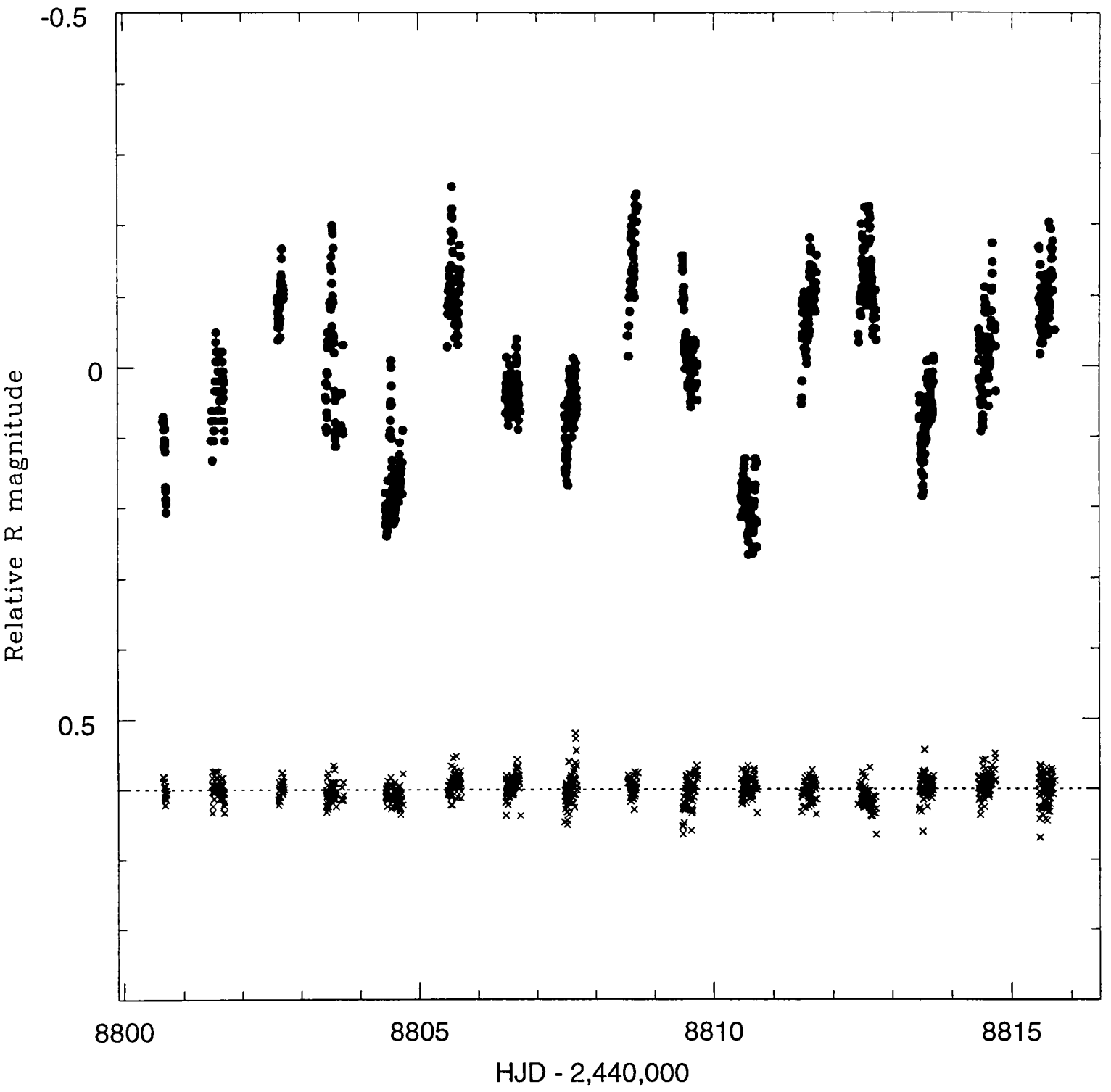


Figure 3.1: The overall R-band light curve of V404 Cyg relative to the reference star (top) and the reference star relative to the check star. The ellipsoidal variation of V404 Cyg ($\sim 0.3^m$) can clearly be seen, whereas the reference star only exhibits a random scatter of $\sigma = 0.02^m$.

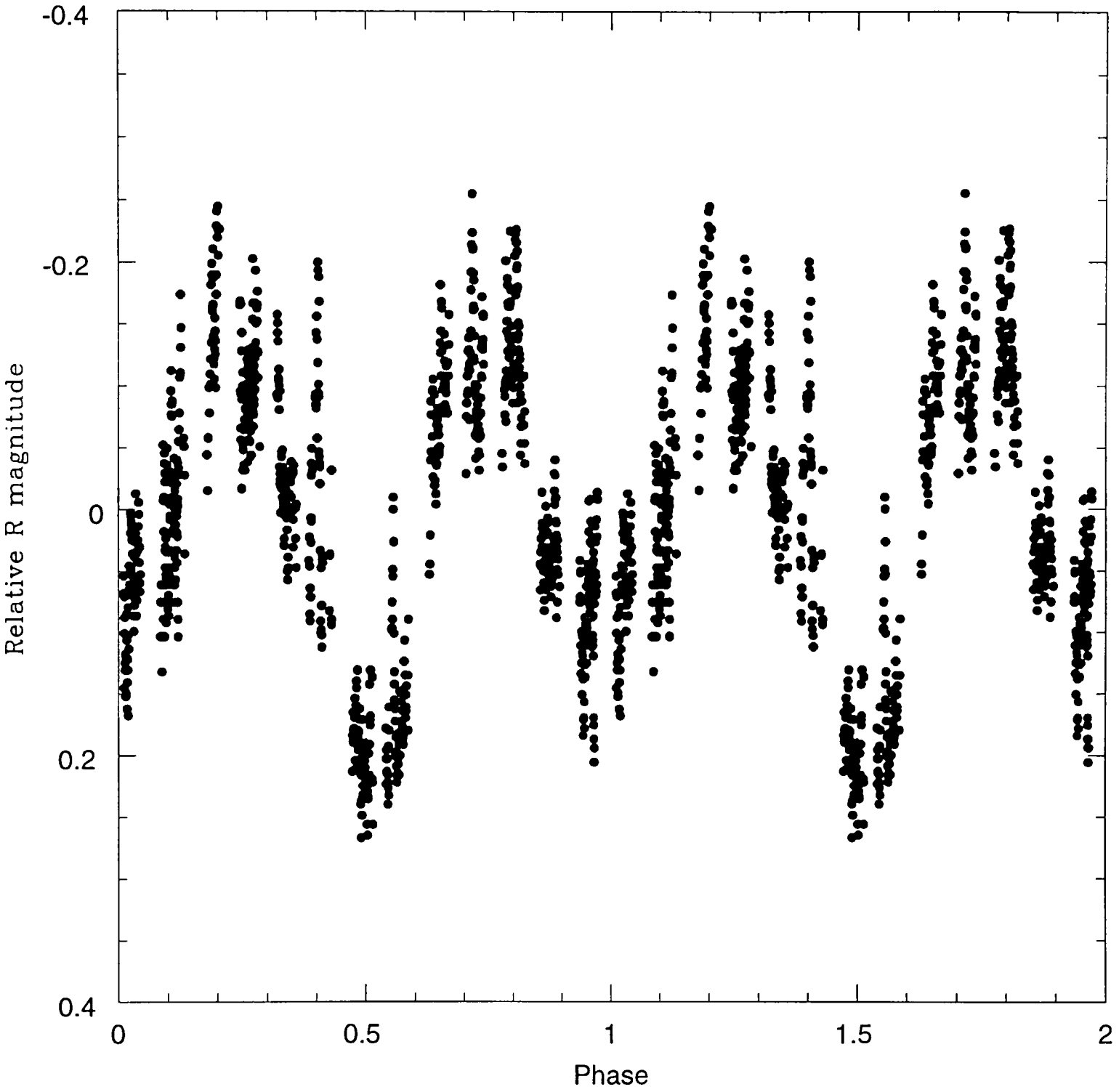


Figure 3.2: The relative R-band light curve of V404 Cyg folded on the 6.4714 day spectroscopic period. The ellipsoidal modulation can clearly be seen, as well as short period variations within each night. The light curve consists of a total of 1083 measurements.

The orbital ephemeris of V404 Cyg is known from spectroscopy (P_{orb} = 6.4714 days, T_0 = 2,448,813.873; Casares and Charles, 1994) and so the light curve was folded using these values (see figure 3.2). The ellipsoidal variation can clearly be seen in the light curve, although there is a significant variation in the brightness of V404 Cyg within each night. The data was divided into 16 bins to enable an ellipsoidal fit to be made to the light curve. Two sets of binned data were computed: one using the mean of the light curve each night and another derived from an estimate of the 'lower envelope' of the data. To estimate the lower envelope the mean of the five lowest points from each night was computed, and then the average phase of these points was used to place them on the folded light curve. This *ad hoc* method was used to try to remove the short period variations that are seen in the data, assuming that they are caused by a variation of the flux above the lower envelope values.

An ellipsoidal model fit to the binned data was calculated (for a description of the ellipsoidal effect see Chapter 1). The fit was performed using the following parameters: T_{eff} = 4400 K, limb-darkening coefficient u = 0.716, gravity darkening exponent β = 0.08, and effective R-band wavelength λ = 6460 Å, leaving only the inclination as a free parameter. The effective temperature was taken from Casares et al. (1993), and was estimated by fitting the energy distribution of V404 Cyg between 3600 Å and 22,000 Å. The limb-darkening coefficient appropriate for the R-band was taken from Al-Naimiy (1978), assuming the secondary star has a convective envelope. The gravity darkening derived by Lucy (1967), also assumes that the secondary has a convective envelope. This value is reasonable as observations have generally determined β to be within 0.02 of the value used here (see Sarna, 1989, and references therein). The variation necessary for each of these three main parameters to cause an error of 1° was computed giving: ΔT_{eff} = 500 K, Δu = 0.15 and $\Delta\beta$ = 0.025. Thus, the inclination is not sensitive to errors in the effective temperature and limb-darkening; the accepted value of β is unlikely to be in error by an amount large enough to affect the inclinations obtained using this model.

Using the known mass ratio of $q = 16.7^{+1.5}_{-1.1}$ (Casares and Charles, 1994) the fit to the mean light curve gave an inclination of $i = 56° \pm 6°$, and for the lower envelope a value of

Table 3.1: Colours for V404 Cyg and two nearby reference stars, C4 and C5.

Filter	C4	C5	V404 Cyg	V404 Cyg – Dereddened ($A_v = 4.0$)
B	20.396(0.02)	21.568(0.06)	20.657(0.04)	15.367
V	18.009(0.006)	18.993(0.013)	18.534(0.012)	14.543
R	16.080(0.003)	16.711(0.004)	16.656(0.004)	13.662
I	15.157(0.003)	15.723(0.004)	15.749(0.006)	13.814

$i = 57° \pm 5°$. The errors on the inclination were obtained by finding the minimum reduced chi-squared, after which the 1-σ error is given by $\chi^2_\nu + 1.0$ (Lampton, Margon and Bowyer, 1976). The reduced chi-squared values for the two fits are $\chi^2_\nu = 4.676$ and 2.176, respectively. This implies that the lower envelope does remove some of the contamination of the light curve caused by short period variations. However, the inclination differs little between the two datasets, and because of the rather arbitrary definition of the 'lower envelope', the mean light curve was used in all subsequent analysis. The light curve model required no hot spot contribution (this was parameterized as a Gaussian variation in amplitude centred at phase 0.9). If the hot spot was included in the model, the fitting routine assigned it a negligible amplitude $\sim 10^{-3}$ magnitudes.

3.1.4 Colour Dependent Variations in V404 Cyg

Colours of V404 Cyg were obtained simultaneously with the R-band data to measure the variation in the data as a function of the observed colour. A series of B,V,R and I-band frames were exposed once each night, giving a total of 16 measurements. These were calibrated using the standards 114 548, 107 684, 110 340 and 106 1024 from Landolt (1983). The extinction corrections were applied using the extinction curve derived in the ING La Palma documentation, which was calculated using the method outlined by Hayes and Latham (1975), assuming an aerosol-free atmosphere. A colour equation was derived for each colour to test for any colour effects in the calibration, but it was found that this did not improve the consistency of the calibrated magnitudes and so it was not used.

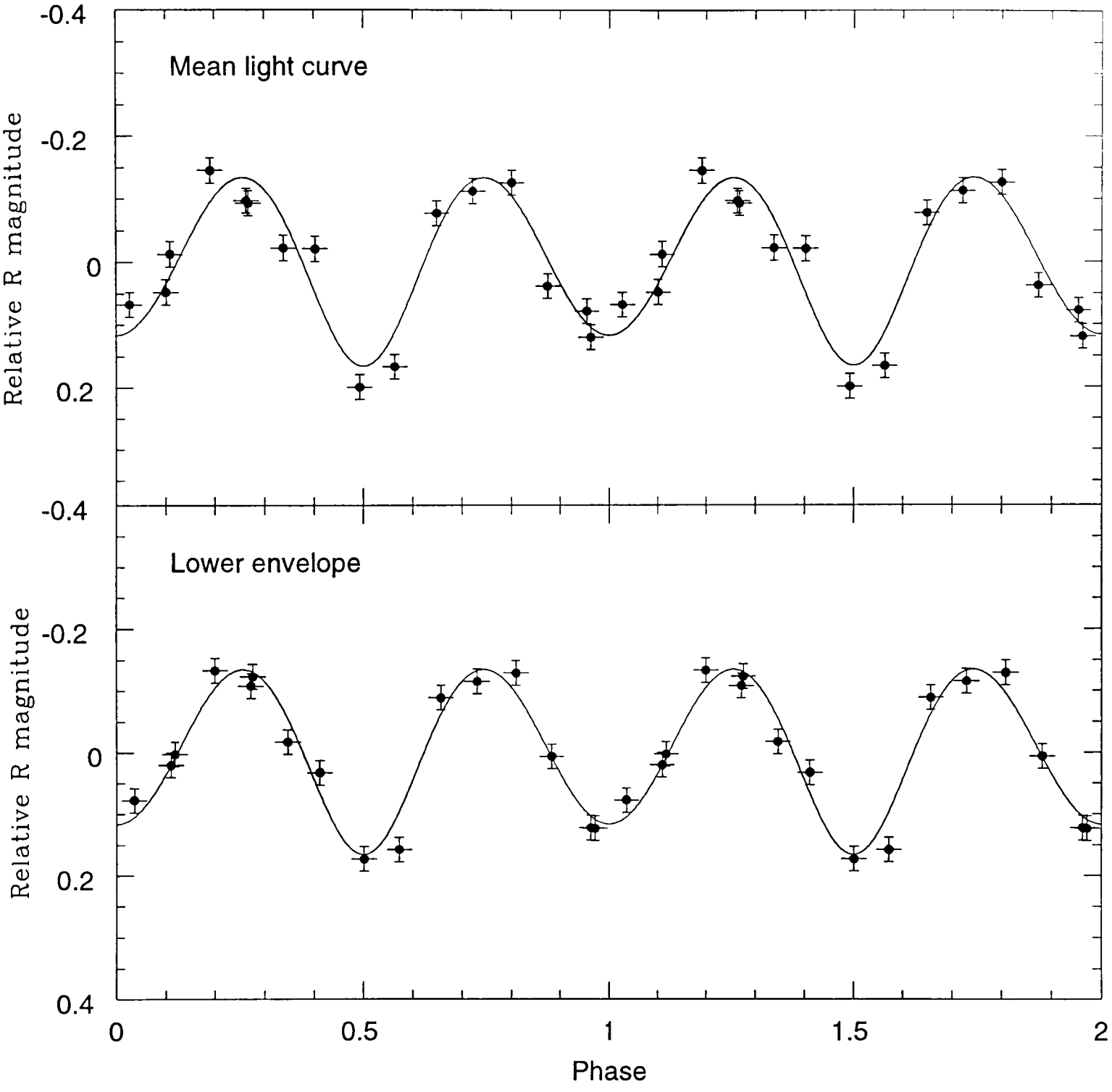


Figure 3.3: The ellipsoidal fits to the mean light curve and lower envelope in the R-band. Both sets of data give similar inclinations ($i = 56° \pm 6°$ and $57° \pm 5°$ respectively), but the lower envelope has a much lower chi-squared ($\chi^2_\nu = 2.176$, compared to 4.676 for the mean light curve). This implies that the lower envelope has less contamination from the source of the 6 hour modulation.

To perform the calibration, it was necessary to determine which of the nights were photometric, as some of this photometry was taken in borderline transparent conditions. This was done by checking the consistency of the photometry against that of Casares et al. (1993) and Udalski and Kaluzny (1991), as well as ensuring that the calibration did not alter from night to night by more than was expected due to the statistical errors. The results for this calibration are shown in table 3.1, along with the dereddened magnitude of V404 Cyg which was obtained using the reddening $A_v = 4.0$ (Casares et al., 1993). These average magnitudes were calculated by taking the weighted mean of all of the photometric nights to produce the least biased estimate of the true magnitudes.

The B,V,R and I-band light curves folded on the orbital ephemeris are shown in figure 3.4. The ellipsoidal fit for each colour is plotted as a dotted line (the inclinations for the fits are: $57° \pm 25°$, $59° \pm 10°$, $56° \pm 5°$ and $58° \pm 6°$ respectively). The parameters for the fits were the same as the earlier R-band model, except for the wavelengths of the filters used and the limb-darkening (which is wavelength dependent). The values used were: $u_B = 1.000$, $\lambda_B = 4400$ Å; $u_V = 0.835$, $\lambda_V = 5470$ Å; $u_I = 0.5765$, $\lambda_I = 8300$ Å. These ellipsoidal fits and the extended I-band fit ($i = 65° \pm 7°$; see section 3.3) are displayed versus wavelength in figure 3.5. The central dashed line is the inclination obtained from the K-band infra-red photometry of Shahbaz et al. (1994; $i = 56° \pm 4°$), and the lower and upper lines indicate the 1-σ errors for this value. In the presence of an accretion disc it is expected that the inclination obtained from ellipsoidal fits to photometry should asymptotically increase toward the true value with increasing wavelength. Thus, the values plotted in figure 3.5 should tend toward the dashed line, from blue to red. There is little evidence of this, indicating that the disc contribution is small, in agreement with the result of Casares et al. (1993), who placed an upper limit of 10% on the disc contribution in the R-band.

The weighted mean inclination calculated from the complete set of colour photometry is $57.7° \pm 2.0°$, if there is 0% disc contribution (the chi-squared contours for the entire set of data are shown in figure 3.6, along with the X-ray eclipse region). However, Casares et al. (1993) found that there is a small but significant amount of veiling due to the disc, and that

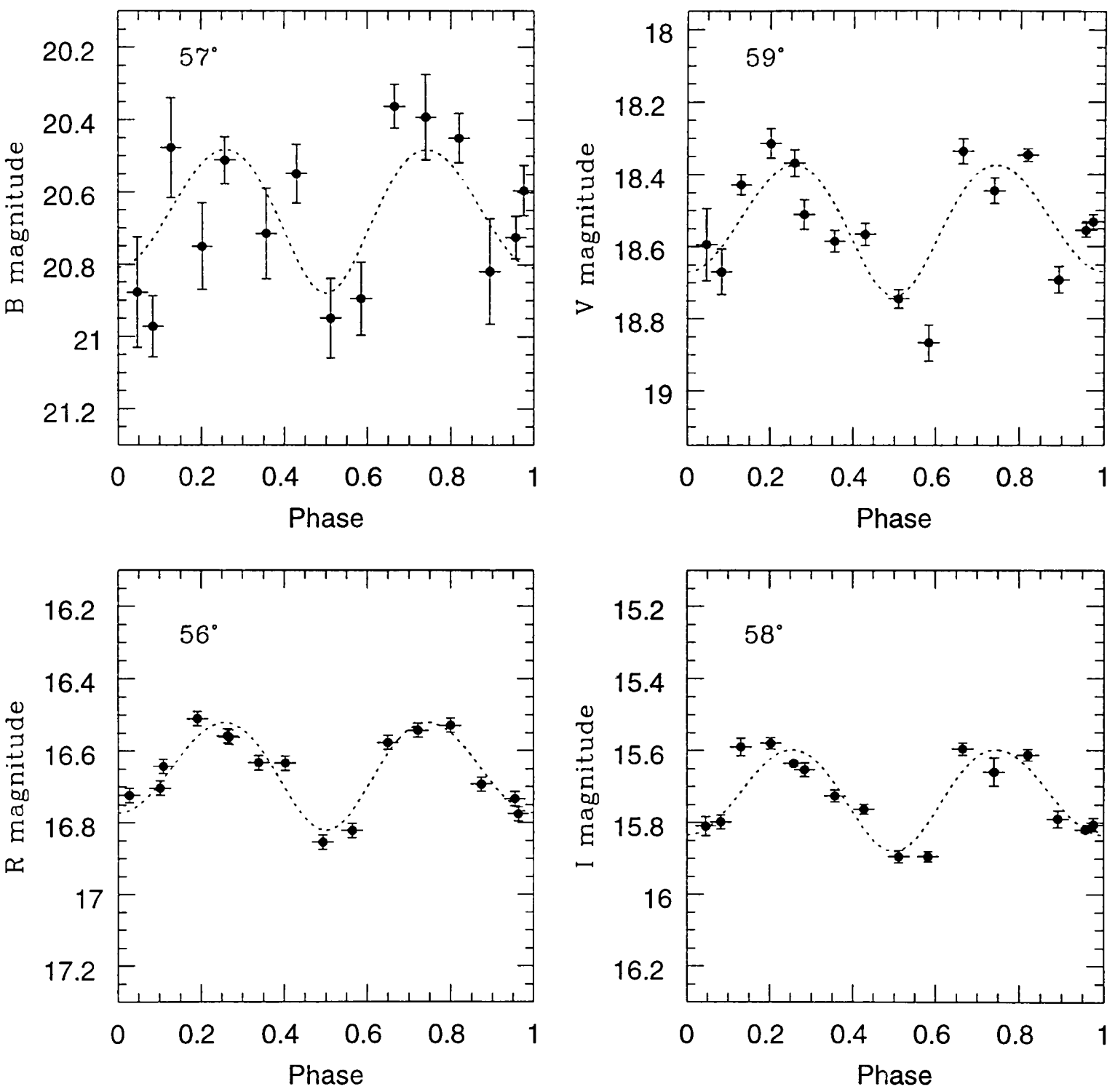


Figure 3.4: The B, V, R and I-band ellipsoidal light curves, folded on the orbital ephemeris. The model fit to each curve is shown as a dashed line (the best fit inclinations are shown at the top left of each panel). It can be seen that the data cannot be wholly explained just by the ellipsoidal model.

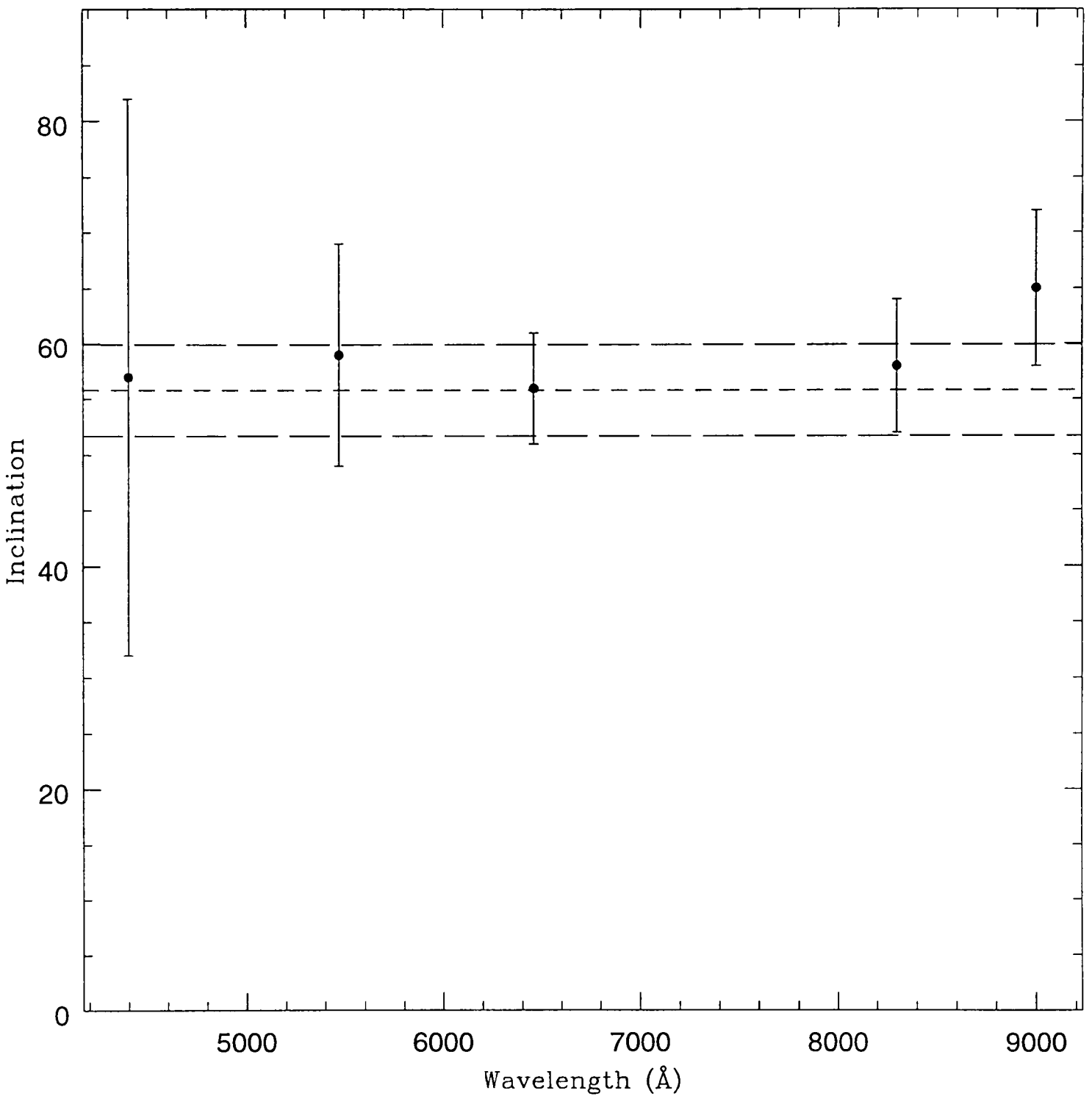


Figure 3.5: The resulting inclinations from the ellipsoidal fits to the B,V,R,I and extended I-band data (see section 3.3). The dashed lines are the best fit of Shahbaz et al. (1994), and its associated errors. There are no obvious trends with wavelength which is consistent with a small disc contribution in V404 Cyg.

it increases toward the blue. They calculated the fractional veilings for the B,V and R-bands to be: $r_B = 0.36$, $r_V = 0.18$, $r_R = 0.10$, where r is defined as the ratio of the excess emission relative to the secondary star flux. These values are consistent with figure 3.5, which would have inclination constant as a function of wavelength (within the errors) if the fractional veiling was taken into account. The veiling causes the above, overall value for the inclination to be an underestimate of the true value, and so the veiling must be taken into account to allow for this systematic error. The veiling distribution (power law $\alpha = -3.4 \pm 0.9$) was used to extrapolate the veiling factor up to the I, extended I, and K-bands, which have λ_{eff} = 8300, 9000 Å and 2.205 μm, respectively. Thus, the values calculated for these wavebands are: $r_I = 0.04$, $r_{longI} = 0.03$, $r_K = 0.0015$. Finally, the mean inclination from the entire set of photometry having taken into account the veiling is equal to $59.8° \pm 2.0°$, which is consistent with previous results of other authors, taking into account the errors.

The only other bias on this estimate is caused by the short period modulation of the light curve. Over a large set of data this bias should decrease as the 6 hour modulation averages out over the six day period, but it could have a small systematic effect on the above result as the B,V and I datasets are relatively small. To estimate the effect of this bias, an upper limit can be obtained from the error on the mean of the six hour variation, given the number of data points n. This error is equal to the semi-amplitude of the six hour variation ($\sim 0.025^m$, see section 3.1.5) divided by $\sqrt{2n}$. For the I-band data ($n = 16$) this is equal to 0.0044^m and for the R-band data ($n = 1083$), the error is 5.4 $\times 10^{-4}$ magnitudes. Therefore, this bias is negligible compared to other uncertainties in the analysis.

It is clear that the data varies significantly more than would be expected purely from the orbital modulation. To quantify this the variance of the data was calculated for each of the four colours, as well as the expected variance of the sum of the photometric noise and orbital variations. This is shown in table 3.2. The variances due to noise and the ellipsoidal modulation increase toward the blue, as does the data itself. But the residual variance in the data, left after subtraction of the noise and ellipsoidal variances, still increases toward the blue. This suggests that there is another source of variability in V404 Cyg that is bluer than

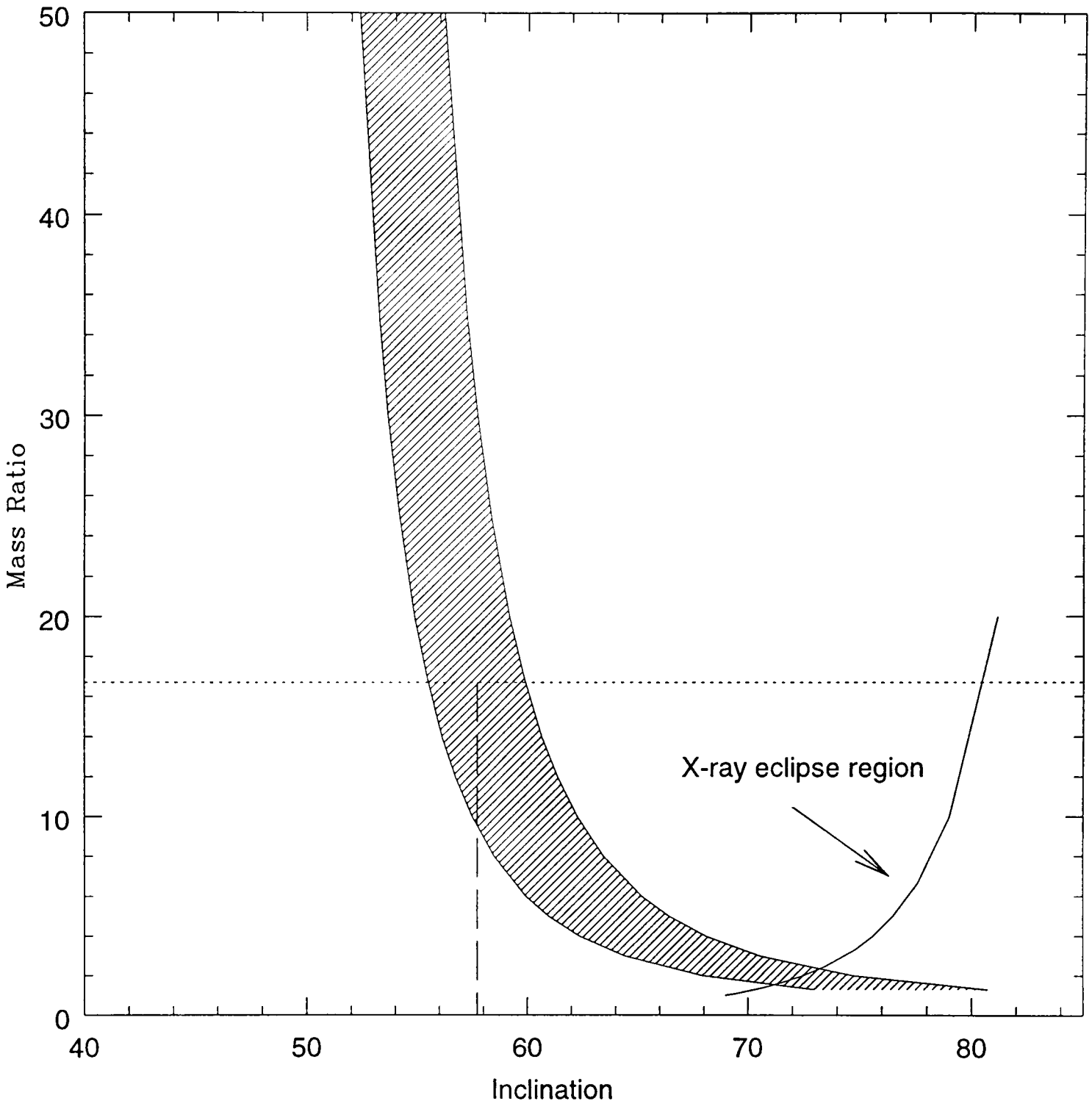


Figure 3.6: The chi-squared contours for the entire set of photometry, with a best fit of 57.7° ± 2.0° marked along the dashed line, corresponding to the mass ratio of V404 Cyg, q = 16.7 (indicated by the dotted line). This solution assumes no disc contamination, whereas the best fit taking into account the disc veiling ($\approx$ 10% in the R band) gives an inclination of 59.8° ± 2.0°. The X-ray eclipse region is also shown.

Table 3.2: Variances for the B,V,R, I-band photometry.

Filter	Variances			
	Ellipsoidal (1)	Noise (2)	Data (3)	Difference 3−(2+1)
B	0.01644	0.00990	0.03998	0.01364
V	0.01377	0.00146	0.02355	0.00832
R	0.00961	0.00040	0.01271	0.00270
I	0.00853	0.00034	0.01108	0.00221

the secondary star. This variability is due to the short period variations ($\sim$ 6 hours), which will be discussed in more depth in section 3.2, combined with intrinsic scatter in the light curve. Both of these effects appear to arise in the disc.

3.1.5 Bayesian Frequency Analysis

To calculate the probability that the short period variations are real, a Bayesian analysis of the data was performed. The probability that the data consist of one, two or three sinusoids can be measured using a model comparison method, similar to that used by Bretthorst (1988). The equation used is basically a discrete fourier transform with multiple frequencies. For a description of the method see Chapter 2.

To perform the analysis the noise level used was $\sigma = 0.016^m$ (obtained from the photometric errors in APPHOT in IRAF. The first comparison was made between a single sinusoid model at 3 days and a model including the variation in the minima over the 6 day orbital period, to test if the data show an ellipsoidal modulation. For this we get the relative probability of the two frequency model compared to the single frequency: $P(2/1) = 10^{784}$; a very large probability that the light curve is ellipsoidal. The two dimensional Bayesian periodogram is shown in figure 3.7. This has a large peak at the periods of the orbital, ellipsoidal modulation, 3.13 ± 0.19 and 5.26 ± 1.4 days. The reason for the large error associated with the 6 day component of the modulation is that it is closer to the resolution limit of the data than the 3 day period, and it also has a much smaller amplitude which decreases the accuracy of the estimate of this component.

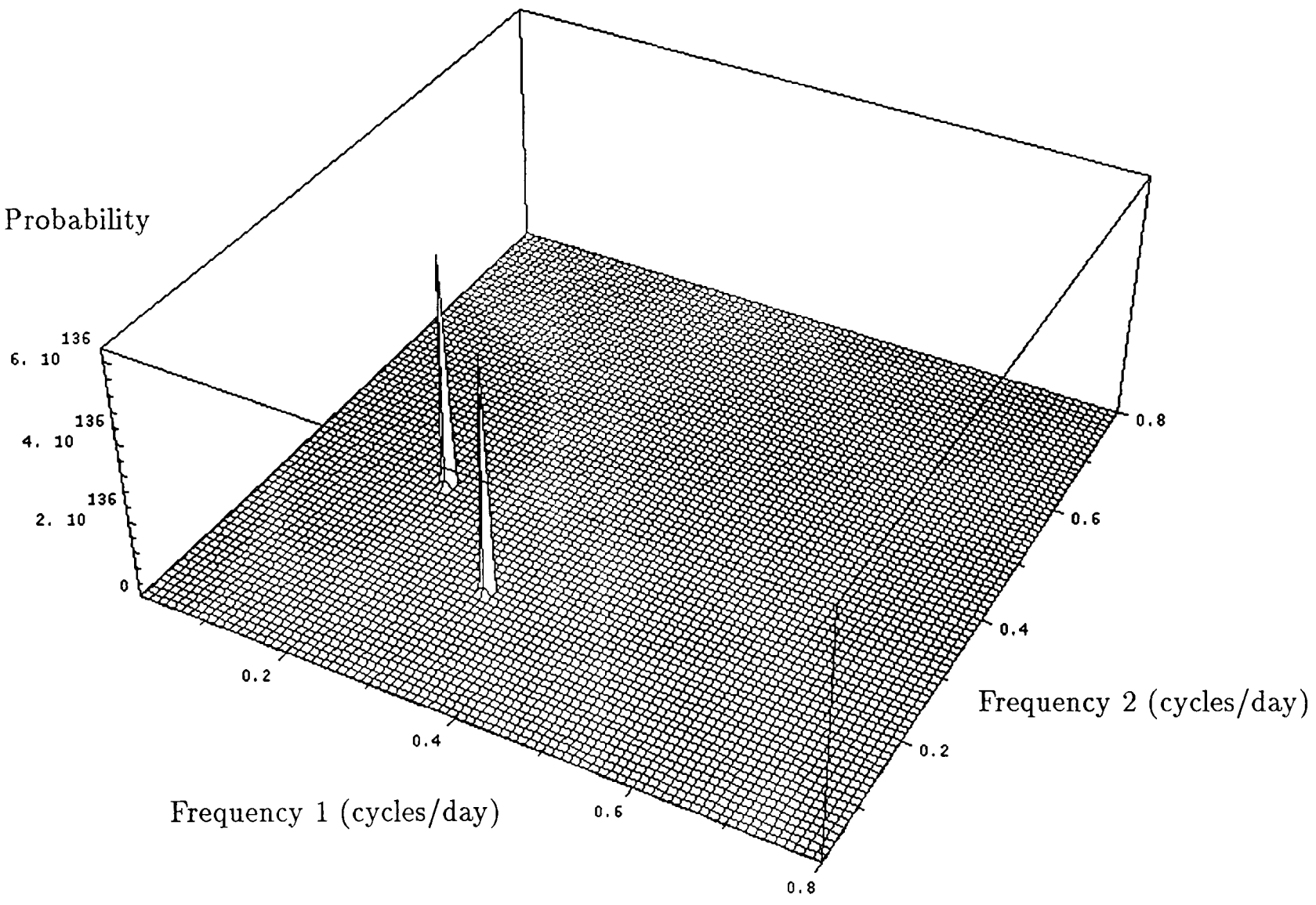


Figure 3.7: The two dimensional periodogram of V404 Cyg showing the ellipsoidal modulation. The two peaks represent the periods at 6 and 3 days, caused by the ellipsoidal effect, reflected in the $f_1 = f_2$ axis.

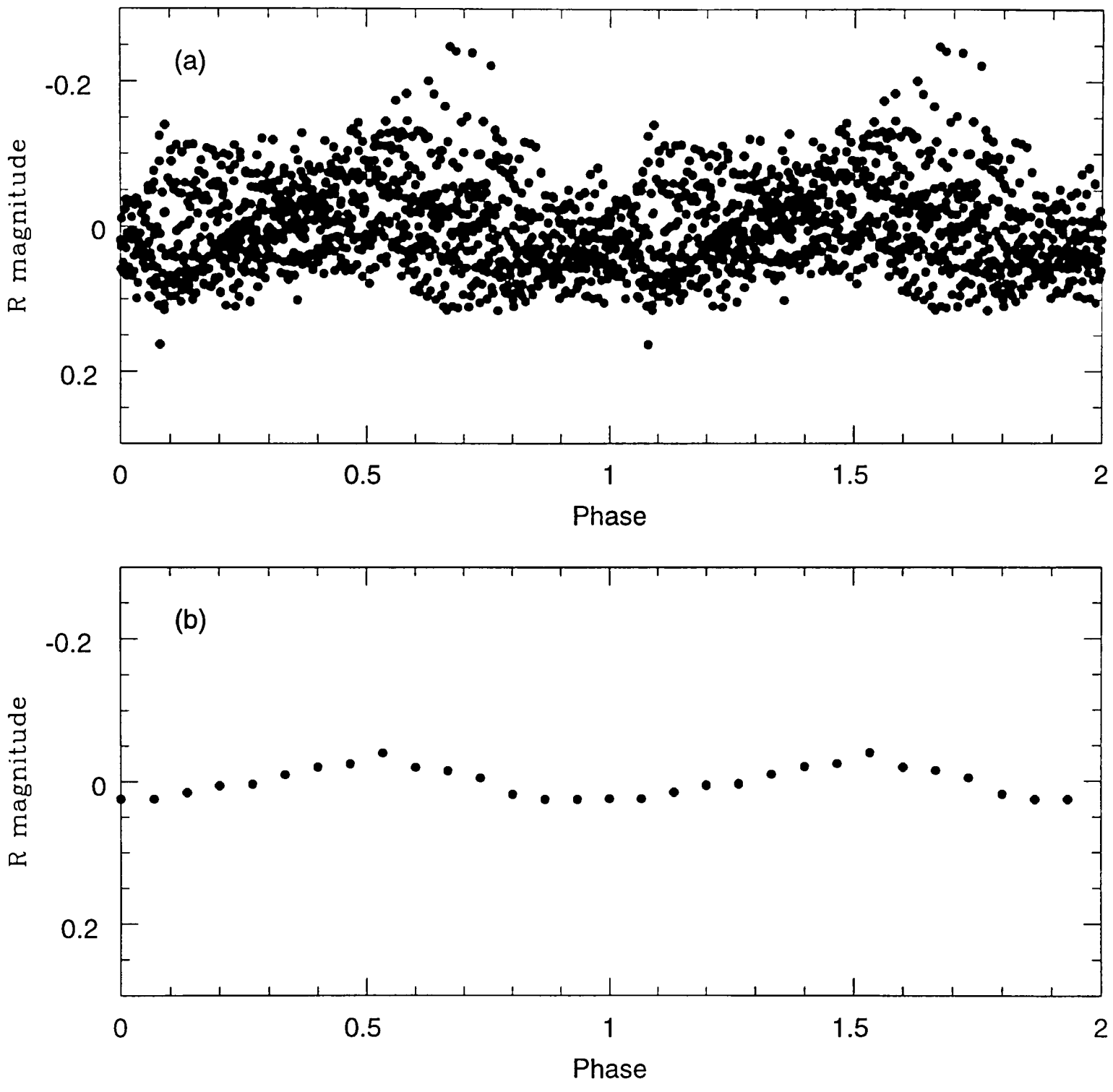


Figure 3.8: The ellipsoidal-subtracted data folded on the 6.22 hour period (top) and divided into 15 phase bins (bottom). The data exhibit a non-sinusoidal modulation $\sim 0.05^m$ superposed on a random variation which is larger than would be expected from purely statistical noise.

To analyze the short term variations, a periodogram of the three frequency model was calculated, as is shown in figure 3.9. The log of posterior probability is plotted versus frequency 3 (frequencies 1 and 2 were set to the orbital values found previously). The peak at 6.22 $\pm$ 0.01 hours (3.852 $\pm$ 0.005 cycles/day) has the largest probability by a factor 170, compared to the peak at 8.377 hours (2.865 cycles/day). The main subsidiary peaks appear at the 1-day aliases of the 6.22 hour period. Now it is possible to test the probability that this six hour modulation is real. For 6.22 hrs the probability of the three frequency model compared to the two frequency is $P(3/2) = 10^{228}$. This indicates that this modulation is very significant relative to the noise level of the data.

To test whether there are any other periodicities in V404 Cyg, the probability of a four frequency model can be compared to the three frequency model. This gives a peak probability of $P(4/3) \sim 10^{21}$. The probability that the four frequency model is more likely than the three frequency is not insignificant. This is probably indicative of the fact that the scatter in the light curve is not purely due to statistical noise. The Bayesian estimate of the residual noise level taking into account the first three periodicities (two orbital and the 6 hour modulation) is $0.053^m \pm 0.005^m$, whereas the noise due to photometric errors is 0.016^m. This implies that there is a variation in the light curve due to flickering of 0.051^m ($= (0.053^{m2} - 0.016^{m2})^{0.5}$). If the total noise of 0.053^m is put into the model comparison equation as the true noise level of the data, then the probability for the four frequency model becomes: $P(4/3) = 10^{-2.0}$. Finally, the probability that there are five frequencies compared to four gives: $P(5/4) = 10^{-1055}$. There are no more detectable periodicities in the data.

So, the Bayesian analysis predicts that there are three frequencies present in the data, as well as noise which is significantly above that predicted by the measuremental errors. Before the nature of the 6 hour variation is examined it is necessary to remove the first two frequencies (i.e, the ellipsoidal variation). To do this an estimate of the amplitudes can be made using the 3 frequency Bayesian model. The full amplitudes for the 6 and 3 day periods are $0.086^m \pm 0.006^m$ and $0.28^m \pm 0.006^m$, respectively (the amplitude of the 6 hour period was estimated to be $0.056^m \pm 0.005^m$). These values agree quite well with those obtained by

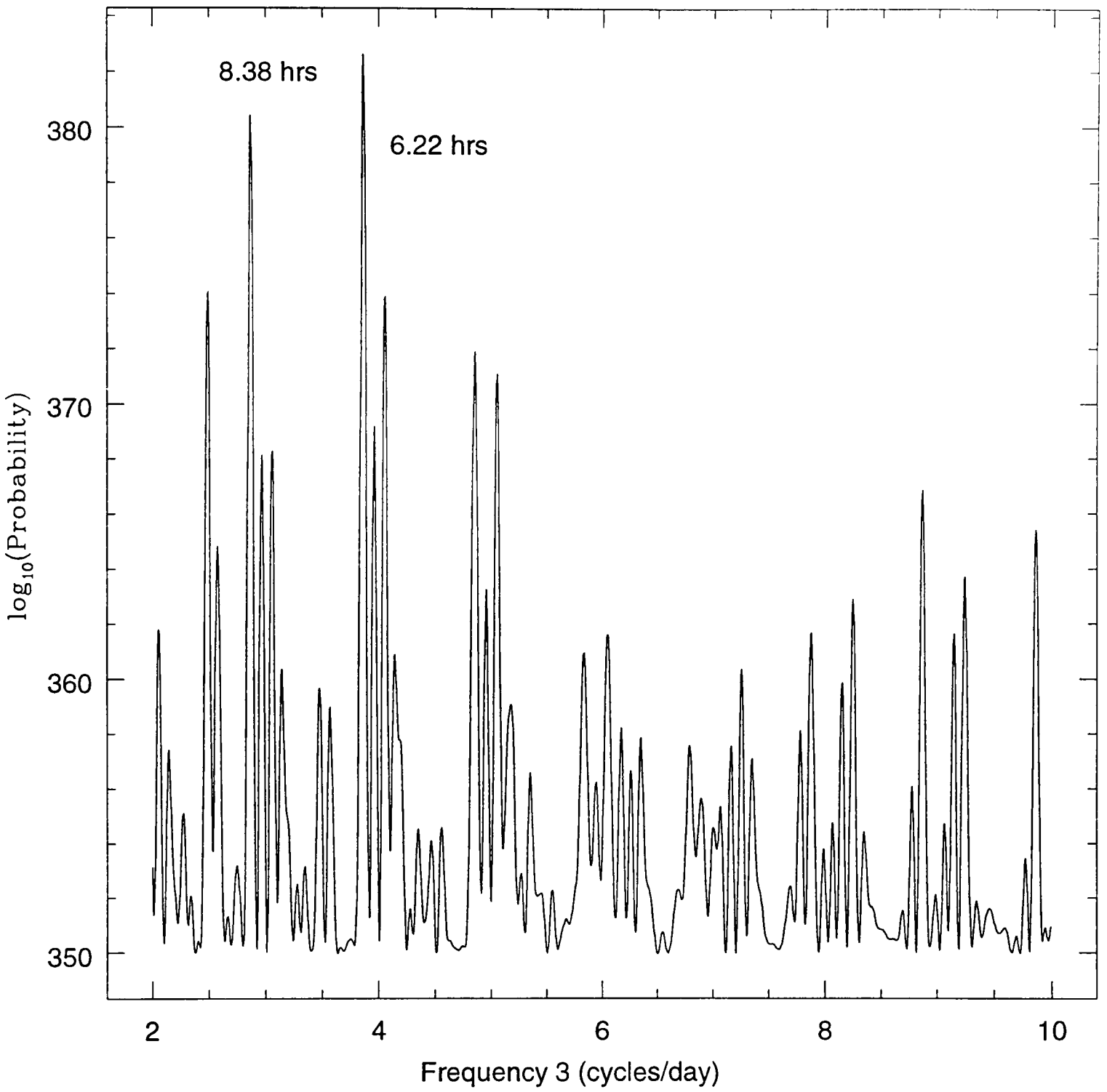


Figure 3.9: The 3 dimensional Bayesian periodogram of the R-band dataset. The axes are the log of probability versus frequency 3 (frequencies 1 and 2 were set to the values found for the 2-D periodogram). The 6.22 hour peak is largest with a relative probability of 10^{382}, a factor of 170 larger than the nearby alias peak at 8.38 hours.

fitting the observed light curve (this gave: 6 days – $0.076^m \pm 0.004^m$; 3 days – $0.33^m \pm 0.02^m$). The ellipsoidal modulation was then subtracted from the data using the known zero phase by using the model fit, as this is a physical rather than a statistical fit to the data.

The ellipsoidal-subtracted data folded on the 6.22 hour period are shown in the top panel of figure 3.8. The ephemeris used to fold these data was obtained from a least squares sinusoidal fit to the data. This gave: $T_0 = 2448800.441 \pm 0.001$, $P_{6hr} = 6.218 \pm 0.008$ hours (3.860 ± 0.005 cycles/day) and amplitude = 0.056^m. The data exhibit a modulation $\sim 0.05^m$ superposed on random scatter which is larger than would be expected from purely statistical noise for an object with the brightness of V404 Cyg. In the bottom panel the data are divided into 15 bins to average out this noise (the formal error on the points is 0.002^m given that the original mean error was 0.02^m, and that there are on average 72 points per bin). It is clear from the binned data that the modulation is not exactly sinusoidal (the reduced chi-squared for a sine fit to the data was found to be 5.04).

3.2 The 6-hour Quasi-Periodicity

3.2.1 Changes in the six hour period.

To investigate the nature of the six hour modulation, without the use of specific models for the data, a least squares analysis was undertaken. The ellipsoidal-subtracted data were divided into groups of four nights, nights 1 to 4, nights 2 to 5 and so on. For each section of data. a least squares sinusoidal fit was performed with three free parameters: phase, frequency, amplitude, with the frequency constrained to be in the region of the six hour modulation. In this way 13 measurements of each were made as is shown in figure 3.10. From this it can be seen that all three parameters vary during the observations. The frequency of the six hour modulation decreases steadily during the run (figure 3.10(b)), and from this the expected phase can be calculated relative to the overall phase zero of the entire dataset ($T_0 = 8800.441 \pm 0.001$). The solid line (figure 3.10(a)) represents this phase difference calculated for each corresponding frequency, which is consistent with the observed phase. This demonstrates that the same modulation is being measured over the whole run. The amplitude also decreases

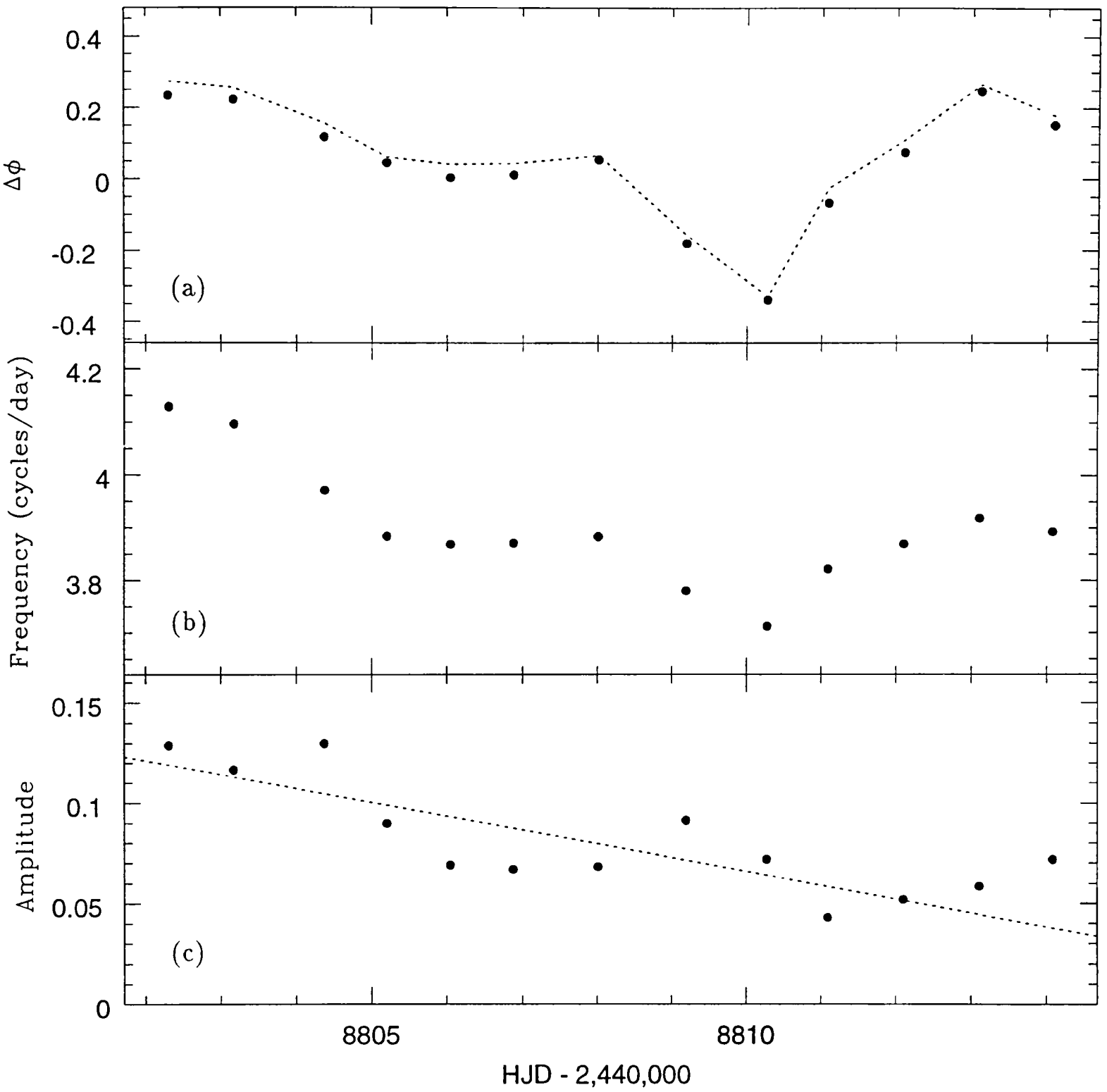


Figure 3.10: The variation in the fitted parameters, derived as described in the main text. The frequency appears to generally decrease, whereas the phase exhibits more complex behaviour. The phase derived from the frequency change is plotted (dotted line) alongside the measured phase in the top panel. This shows that the fits to the 6 hour variation are consistent throughout the dataset. The lower panel of amplitude versus time shows a large decrease in the amplitude of the modulation of at least a factor of two.

during the run from 0.14^m to 0.05^m. The solid line represents a straight line fit to the amplitude.

These changes in the six hour modulation confirm its quasi-periodic nature. The fact that the changes do not occur on the orbital period indicates that the source of this periodicity is likely to be the disc. The secondary star and hot spot region should vary in absolute brightness over the orbital period, which would cause the six hour variation to undergo an amplitude modulation over the orbital cycle. The decreasing frequency suggests that the differing short periods that have been seen in V404 Cyg at 3, 5 and 6 hours could all arise from the same source. Also, the decrease in amplitude seen of over a factor two concurs with evidence that this periodicity has been seen to be weaker during some observations.

3.2.2 Modelling the Observed Periodogram

The 3-D Bayesian periodogram of the data appears to show the 6 hour period, but with some sidelobe structure. The three peaks that appear near six hours could be due either to several frequencies being present near six hours, or just a single period with changing amplitude. It could also be due in some way to interference between the periodicities present in the data and the uneven data sampling. So, to try to mimic the observed periodogram, a model was created to parameterize the structure seen in the 3-D periodogram (figure 3.9) in a simple way. It was found that the structure could be successfully described by a model involving a single period, but with variable amplitude. An artificial dataset was generated using the following formula:

$$f(t_i) \;=\; \sin(2\pi t_i/P_A)\; A\; \sin(2\pi t_i/P_* + \phi)\;; \quad i = 1, N$$

Where t_i represents the N (=1083) observation times, P_A is the period on which the amplitude changes and P_* is the central period at 6 hours. The model was then used to create artificial Lomb-Scargle periodograms, whilst varying the values for each parameter, and the best fit model was evaluated using the least squares method. The Lomb-Scargle periodogram was used here as it has well understood statistical properties in the presence

of noise and irregular sampling, which is not the case for the Bayesian periodogram. The value of ϕ was obtained by this procedure (the fit gave a value of 0.256) and A was set to the Bayesian estimate of 0.056^m. The values obtained from the fitting routine for P_A and P_* were 9.0 days and 0.252 days. The Lomb-Scargle periodogram calculated for the best fit model (see figure 3.11) shows three component peaks around 6 hours. For comparison the Lomb-Scargle periodogram of the dataset (with the ellipsoidal modulation subtracted) is also plotted in figure 3.11. P_* coincides with the period of the central peak and the amplitude varies with the beat period of the central peak and its two neighbours, $P_A \sim 9.0$ days.

The 4.5 day/9.0 day beat period could be simply an alias caused by the irregular sampling of the data, or an artefact left due to uncertainty in the estimation of the ellipsoidal modulation. It appears to have no relation to the orbital period and so the nature of this period, if real, is not clear. It is also not possible with these data to distinguish whether the structure seen around six hours is due to a variable amplitude, variable frequency (chirping), or the existence of several frequencies around this period. The Bayesian analysis in section 3.1.5, however, clearly showed that the six hour period itself is most definitely real. But it appears that it is not a stable period but has quasi-periodic properties, which seems to rule out the possibility that it is due to an orbital effect.

3.3 I-band Photometry

3.3.1 Introduction

Time-resolved differential CCD photometry of V404 Cyg was obtained with the 1-m Ritchey-Chrétien reflector of the United States Naval Observatory Flagstaff Station (USNOFS) from October 26 to November 7, 1993. Of the 13 nights allocated, 2 were lost due to cloud cover or other poor weather conditions. The USNOFS Texas Instruments 800 x 800 CCD camera was used in combination with an extended I-band filter (see Wagner et al., 1992 for a description of the extended I-band filter) to image a 5.7 x 5.7 arcmin region of the sky at a scale of 0.43 arcsec/pixel. The seeing during our run typically averaged 2 arcseconds. Several bias and twilight sky flat field frames were also obtained each night. These bias and twilight

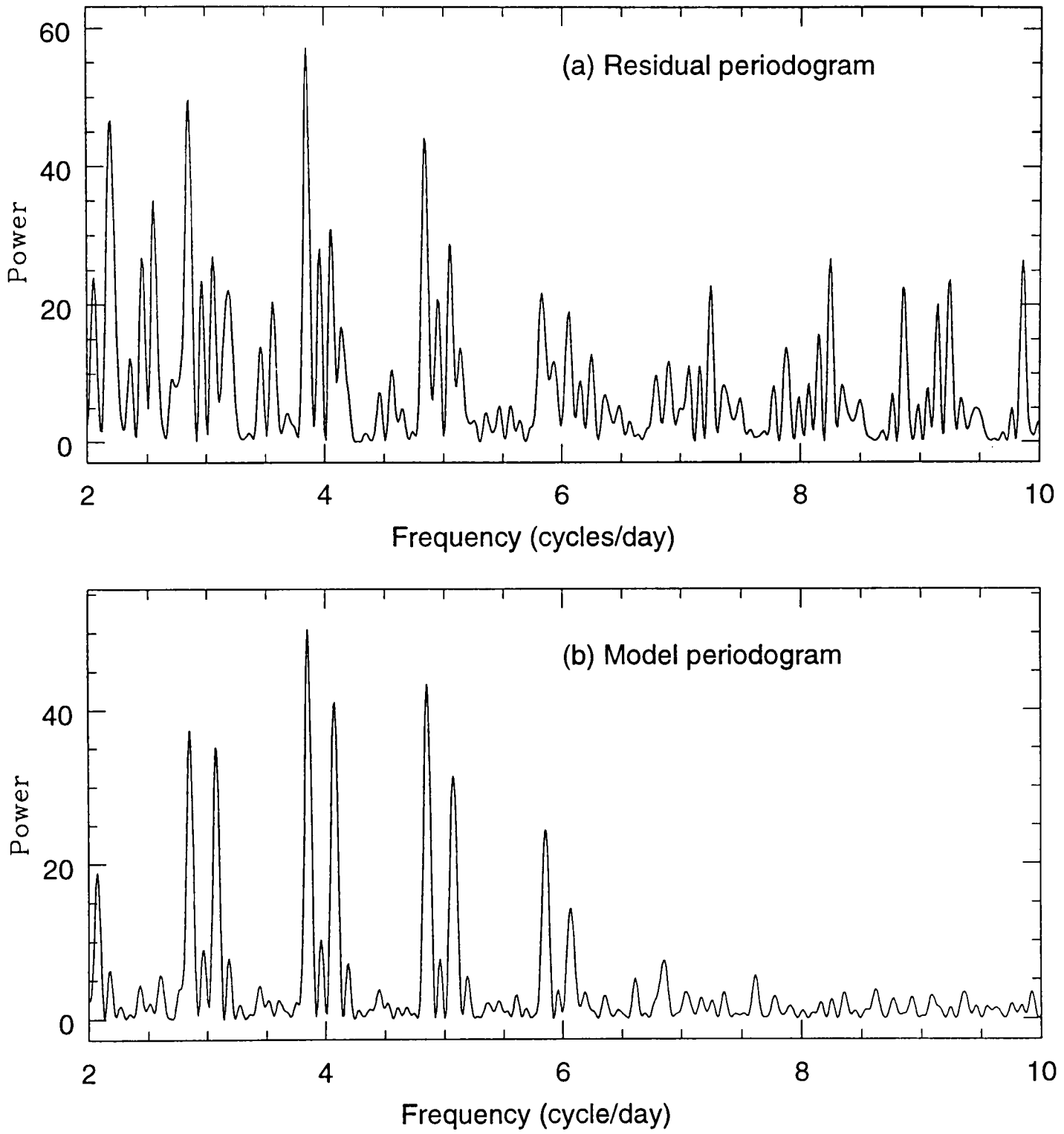


Figure 3.11: The residual periodogram (top) and best fit model periodogram for the R-band dataset (the residual periodogram was left after subtraction of the ellipsoidal modulation, as estimated by the Bayesian frequency analysis). The best fit parameters are as described in the text. The structure around six hours is reproduce fairly well, but there is evidence for a harmonic around ~ 8.0 cycles/day.

flat fields were separately median combined within IRAF using IMCOMBINE. The image frames of V404 Cyg were then de-biased and flat fielded with CCDPROC, using these master biases and twilight flats from each individual night.

The I-band light curve of the 1993 dataset folded on the orbital ephemeris of V404 Cyg is shown in the top panel of figure 3.12. The ellipsoidal fit to the light curve calculated as before (but with $\lambda_{eff} = 9000$, and $u = 0.55$) is superimposed. The fit gave an inclination of $65° \pm 7°$ for the mass ratio of 17, which is just consistent with the previous fits (see figure 3.5).

3.3.2 Period Analysis

A period analysis was performed on the I-band data similar to that used previously on the R-band light curve. A 2-D Bayesian periodogram was calculated to characterize the ellipsoidal variation of the data, which is shown in figure 3.13, note the reflection in the $f_1 = f_2$ line. A peak at the ellipsoidal periods of 6.47 days and 6.47/2 days (within the errors) is apparent in the periodogram, at least a factor ten larger than any other peaks in the periodogram. The peak is more elongated in the 6 day direction because of the larger error associated with this period. This is caused by its small amplitude and proximity to the low frequency resolution limit ($1/24.0$ days^{-1}) of the dataset. The Bayesian estimate of these two frequencies is 0.31 ± 0.01 and 0.11 ± 0.06 days^{-1}, which agrees with the corresponding ellipsoidal periods within the errors.

3.3.3 The Amplitude of the Six Hour Modulation

Next, the six hour period can be examined in more detail. The 3-D periodogram yields 4.01 ± 0.015 cycles/day (5.99 ± 0.02 hours) as the most probable value for the period in the six hour region. The amplitudes for the 3-D periodogram are: 3 day – 0.284 ± 0.008^m; 6 day – 0.051 ± 0.008^m; 6 hour – 0.0894 ± 0.008^m. The first two amplitudes are in agreement with the ellipsoidal variation seen in the top panel of figure 3.12 (3d – 0.292^m; 6d – 0.06^m). The estimate of the variation of the six hour period is also accurate as is shown in the lower panel of figure 3.12. This shows the residuals of the data, after the ellipsoidal effect has

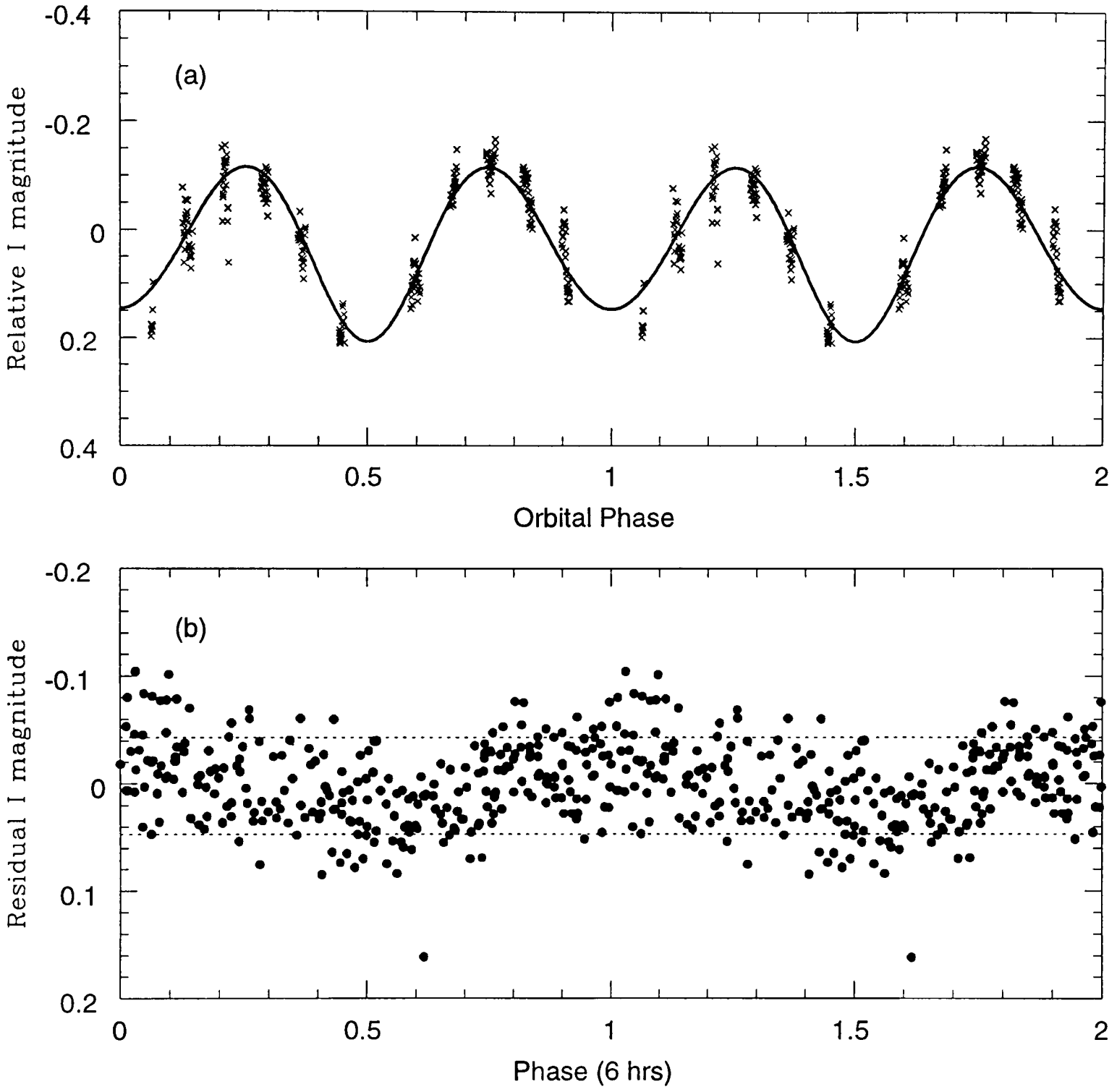


Figure 3.12: Top – the I-band light curve folded on the orbital ephemeris. The best fit to this curve is superimposed, corresponding to an inclination of $65° \pm 7°$. Bottom – the residuals, after subtraction of this fit, folded on the most probable period of 5.76 hours (0.240 days) (using the T_0 obtained for the R-band data). The variation is smaller and not as well defined as in previous datasets. The Bayesian amplitude estimate of 0.045 ± 0.002^m (dotted lines) is consistent with the data.

been removed, folded on the 5.76 hour period (the best fit period to the residuals). This period differs slightly from the Bayesian 3-D estimate because it was obtained from the 1-D periodogram of the residuals, which were obtained by subtracting the ellipsoidal fit, not the Bayesian estimate of the ellipsoidal modulation. It is better to use a physical model of the light curve than to simply remove the orbital modulation using frequency analysis. The amplitude of the light curve agrees with the Bayesian amplitude estimate, which is shown as a dotted line. This amplitude is slightly smaller than was found for the 1992 R-band data (0.056^m), but the difference is negligible given the errors on the estimate. This implies that the six hour modulation did not significantly decrease in size from 1992 to 1993. The phase of the six hour modulation cannot be directly compared to the ealier R-band data as the data show slightly different periods near six hours. However, a calculation of the relative phase was attempted by folding the residuals on the 1992 ephemeris for the six hour period, but this did not give a clear modulation and so the phases could not be reliably matched.

The fact that the six hour modulation did not change significantly from 1992 to 1993 implies that once V404 Cyg reached quiescence the six hour period did not change in strength. Although the values for the amplitude found from these data ($\sim 0.05^m$) are much smaller than found previously. In 1990, just one year after outburst, a variation of $\sim 0.2^m$ was seen in several visible bands as well as the infra-red (Casares et al., 1993). In 1991, the six hour modulation decreased to $\sim 0.1^m$, which implies that there was a further decrease in amplitude up to the section 3.1 observations in 1992 ($A_{6hr} = 0.056^m$), followed by no significant change in the interval 1992–1993. All of this is consistent with the six hour modulation being located in the disc, which would have been much brighter during outburst, with a higher temperature than in quiescence. Also in 1993, K-band photometry was taken of V404 Cyg by Shahbaz et al. (1994), which showed no short term variability with an upper limit of 0.03^m, showing that in quiescence this period has little or no infra-red component. This means that the inclination obtained by Shahbaz et al. (1994) from the ellipsoidal modulation, should be free of any systematic bias due to this effect.

3.3.4 Random Flickering in the Light Curve

The Bayesian estimate of the noise level of the data for the 3-D periodogram was $\sigma_N = 0.035^m$, whereas the photometric error is equal to $\sigma_{ph} = 0.0075^m$, this means that there is another source of noise with $\sigma = 0.034^m$. Thus, both the R-band and I-band data show evidence of flickering over and above that due to purely statistical noise. The I-band is less contaminated by this effect than the R-band (which had $\sigma = 0.051^m$) which indicates that it is a blue effect in agreement with the results of table 3.2. It is therefore likely to be due to flickering in the light from the accretion disc. Unfortunately, it is not possible to directly compare these data with the earlier I-band data, as they were taken using filters with different effective wavelengths (8300 Å and 9000 Å). But the 1-σ variation due to flickering for the two datasets was equal to 0.047^m and 0.033^m, for the I-band and extended I-band, respectively, and so the evidence is consistent with there being little change in the level of flickering from 1992 to 1993.

3.4 Discussion – The Origin of the Six Hour Period

The modulation in excess of the orbital variation was found to increase towards the blue in section 3.1.4 which would imply that it arises in the disc. Assuming that the gas motion in the disc is Keplerian the distance from the compact object, for a 6.22 hour orbit, is equal to $0.12a$ (where a is the binary separation). This is well within the inner Lagrangian point ($0.78a$), and is also smaller than the estimated disc (circularization) radius $R_{circ} = 0.39a$, for a mass ratio of 16.7. This is calculated by assuming the gas flowing across the L_1 point has little or no velocity and then evolves into a circular orbit at the disc's outer edge. Furthermore, it is known from Casares et al. (1993), that almost all of the Hα emission arises from the disc, and thus the detection of the 6 hour periodicity in the Hα emission lines is quite significant. The calculated velocity for a 6 hour Keplerian orbit in V404 Cyg is 740 kms^{-1}, which projected onto the inclination of 59.8°, gives an observed velocity of 640 kms^{-1}. This is fully consistent with the estimate of the velocity for the six hour period, obtained from the change in FWHM of Hα (Casares et al., 1993), of $\sim$ 635 kms^{-1}. This evidence implies that the source of the

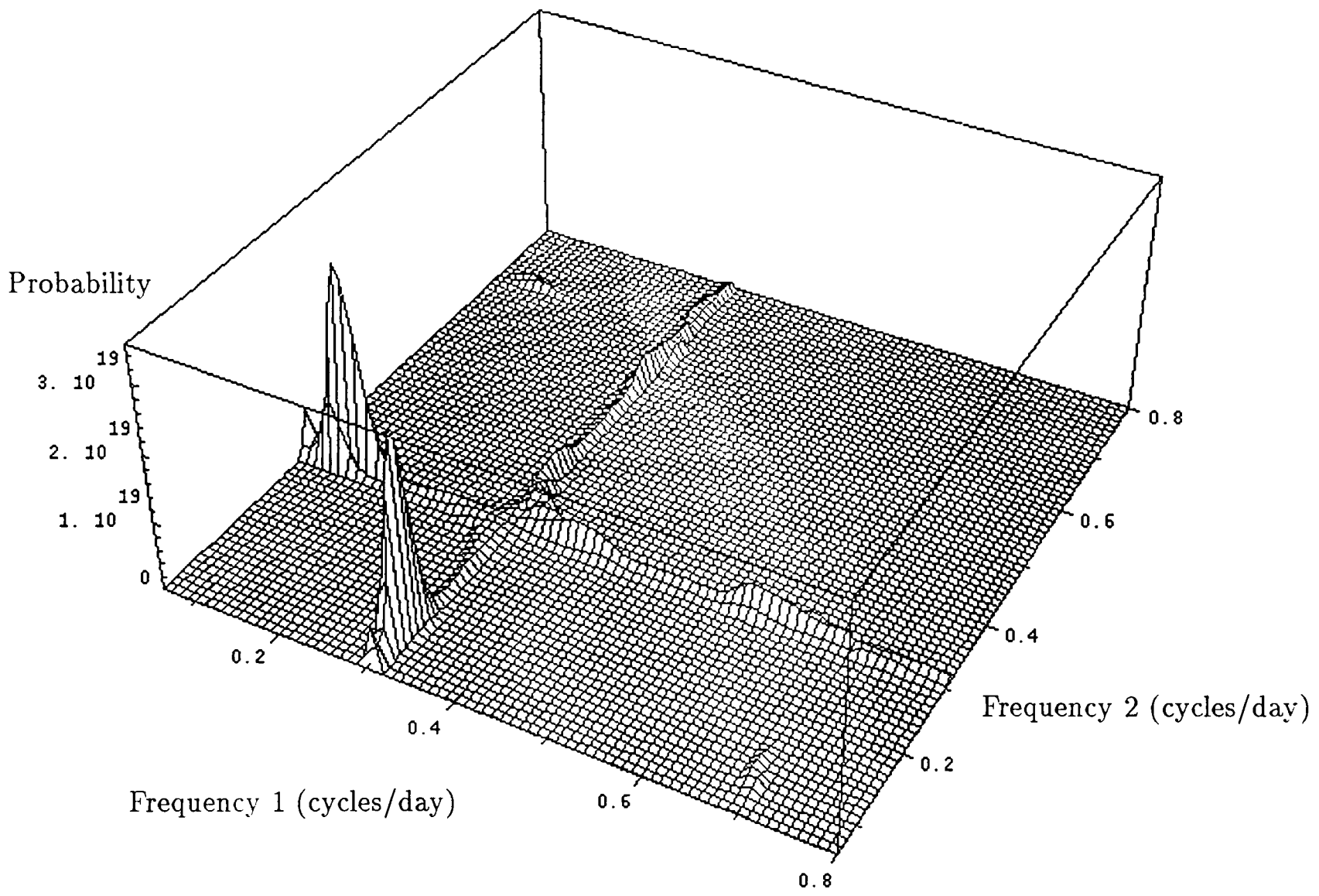


Figure 3.13: The two dimensional periodogram of the 1993, I-band dataset (taken from the low frequency resolution limit ≈ 0.05 cycles/day). There are two peaks representing the most probable frequencies: 0.11 ± 0.06 and 0.31 ± 0.01 days^{-1}, i.e. P_{orb} and $P_{orb}/2$, caused by the ellipsoidal modulation of the secondary star in V404 Cyg. The ellipsoidal periodogram of Chapter 2, taking into account both periods simultaneously, gives a more accurate value for the orbital period of 0.155 ± 0.002 days^{-1} (6.45 ± 0.08 days).

6 hour periodicity is localized in the disc and could be connected with some non-stationary accretion process.

If it is assumed that there is some instability in the disc which gives rise to the short period modulation, then all that needs to be known is how the variability arises physically. To investigate this, several different models were used to try to match the observations in the R-band which clearly exhibited the 6 hour variation. A blob or spot of matter rotating in a circular orbit around a black hole will change in brightness due to several effects, such as gravitational lensing, Doppler beaming and occultation by the disc. For V404 Cyg, where the inclination of the disc to the line of sight is $\approx 60°$, occultation by the disc does not appear to be possible unless the disc flares by more than 30°, which is unlikely as pressure forces cannot maintain the vertical height for such a disc (Meyer and Meyer-Hofmeister, 1982). The other two effects are gravitational lensing by the black hole and beaming due to the motion of the spot. Lensing should be a small effect, but to test this the light curve computed using the model of Bao (1992) was compared to a model which only included beaming. This gave curves with total amplitudes that differed by less than 1% (thus general relativity is not important at this radius, and so it is not possible to distinguish between Schwarzchild and Kerr metrics). This means that the effect of lensing is a factor of 100 weaker than beaming and so it was left out of the model (for higher inclinations the effect of lensing increases and so would need to be included), leaving Doppler beaming as the main cause of the modulation. The model equation used to generate the light curves is shown below (Zhang and Bao, 1991). The intensity of light received by the observer is proportional to the fourth power of the redshift factor (Cunningham and Bardeen, 1973), and so:-

$$I_{obs} = g^4 I_{em} \tag{3.1}$$

where,

$$g = \frac{\sqrt{1 - 2M/r_s}}{\sqrt{1 - 2M/r_s - v_{orb}^2}}(\sqrt{1 - 2M/r_s} - v_{orb} \sin i \sin \phi_s) \tag{3.2}$$

In the above equation $v_{orb} = \Omega_s r_s$, r_s is the orbital radius for the spot, Ω_s is the orbital frequency, M is the black-hole mass, i is the inclination and ϕ_s is the position of the spot along its orbit. The following parameters were put into the model for V404 Cyg: $i = 59.8° \pm 2.0°$, $P = 6.22 \pm 0.01$ hours and $M = 12^{+3}_{-2} M_\odot$. Using equation (3.1) the light curve of magnitude versus orbital phase for the spot was generated as is shown in figure 3.14 (labelled e=0), with the R-band data folded on 6.22 hours also shown for comparison (the formal error on each point is 0.002^m, not taking into account any possible systematic errors caused by the subtraction of the ellipsoidal modulation). This model gave an amplitude of 0.02^m, which is not large enough to match the observed short period variation (amplitude = 0.056^m).

The amplitude of the variation can be increased if elliptical orbits are considered. To generate these curves, the orientation of ellipse with the largest amplitude was considered only, i.e. that with the semi-major axis perpendicular to the line of sight of the observer. The curves for the eccentricities 0.75, 0.9 and 0.99 are shown in figure 3.14. Only the highest eccentricity curve can reproduce the variation seen, and that is without including the 'veiling' effect due to the remaining disc flux and the secondary star flux. To create the amplitude of variation seen, for the $e = 0.99$ case requires that the spot is 0.59 times as bright as the secondary (spectroscopy indicates that the secondary is not veiled by this large an extent). Also, the elliptical orbits modelled are highly eccentric and have little observational evidence to support them, and so the circular case seems more physically reasonable. Unfortunately, the circular model cannot reproduce the data: if the disc is faint compared to the secondary ($<$ 10%) the amplitude of the modulation is 0.002^m. In outburst the disc would have been much brighter, but then the model would only be able to produce a change of $\sim 0.02^m$, compared to the variations seen of $\sim 0.2^m$ [1]. Thus, it appears that if the short period modulation in V404 Cyg is due to a rotating spot in the accretion disc, then it must vary intrinsically in some way.

[1] Although this model does not appear to be able to mimic the six hour variation, it could be used to model QPOs seen in this and other systems. This would require an inhomogeneous inner disc comprising of a population of rotating overdense regions, which would produce variability on a range of timescales.

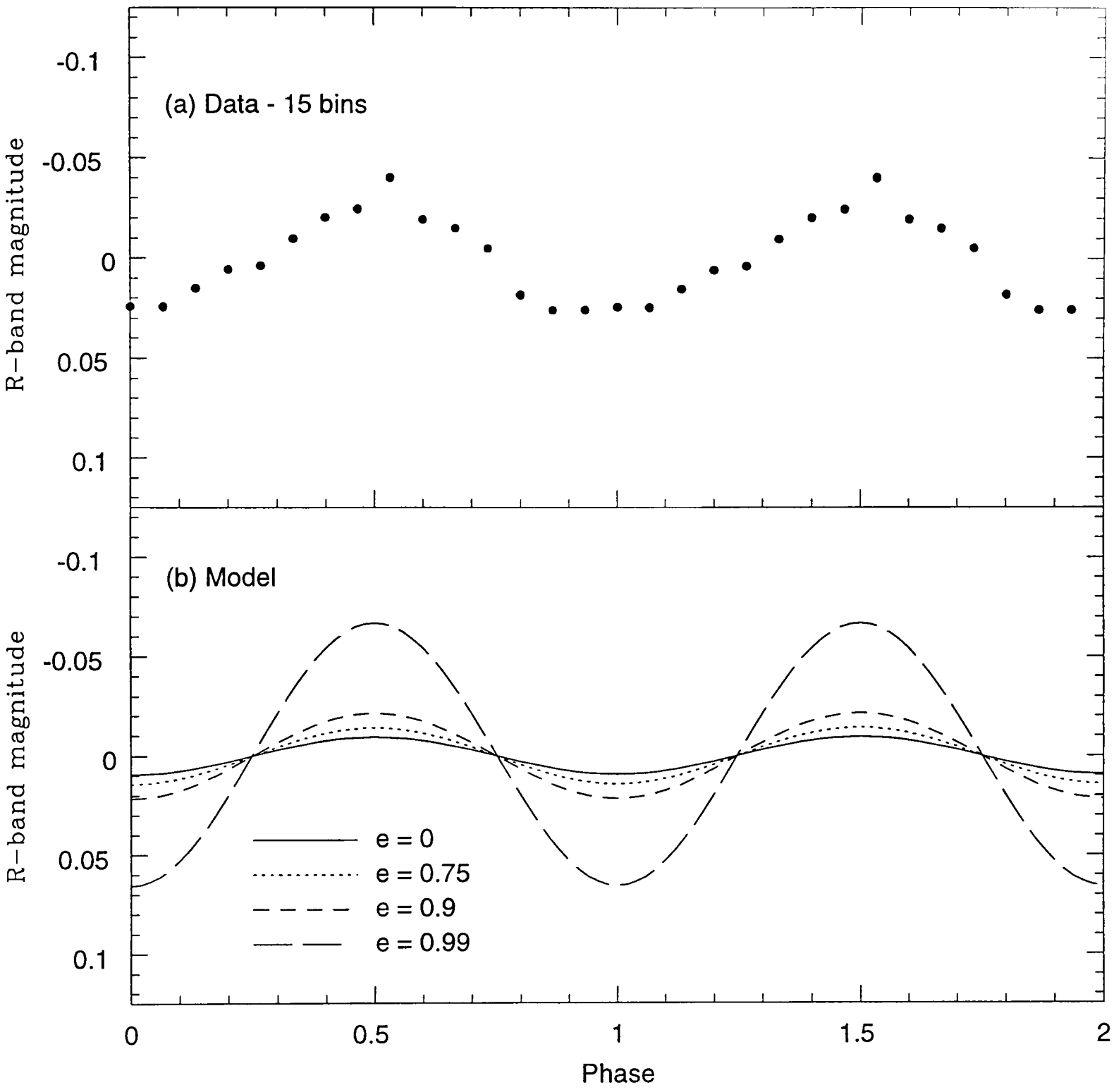


Figure 3.14: Top – the data folded on the 6.22 hour period and divided into 15 bins. The formal error on each point is equal to 0.002^m. Bottom – the model light curves calculated as described in the main text. Only the eccentricity of 0.95 can reproduce the data, and then only if the spot is of the same order of brightness as the secondary star.

3.4.1 The Spectroscopic Phase of the 6 Hour Period

The phase of the six hour modulation was computed from the spectroscopy of Casares et al. (1993), which was taken partly simultaneously with the 1992 R-band photometry presented here. This enables the phase difference between the spectroscopy and photometry to be calculated accurately, despite the quasi-periodic nature of this variability (the two T_0 values are less than a day apart). The phase difference between the spectroscopic phase zero (taken at the blue to red crossing point of the spectra) and the maxima of the photometry is 0.504 ± 0.007. This means that the maximum in the photometric flux occurs at the red to blue crossing point, i.e. when the emitting region is behind the compact object. This could mean that the modulation is caused by the inner, X-ray heated face of some vertical structure in the disc. The gas stream as it impacts the outer disc edge may have a larger vertical extent than the disc itself (Lubow and Shu, 1976) and so could be X-ray heated, but the location of the 6 hour periodicity places it closer to the inner disc (assuming Keplerian velocities in the disc). Lubow (1989) proposed that the gas stream could overflow the disc and would impact the inner disc at a single point, close to the closest approach of the stream with the disc. The difference in velocities between the the gas stream and the disc could then result in strong shocks and turbulence, causing large variations in flux. The radial distance of the impact point was tabulated for different mass ratios, and so using table 1 of Lubow (1989), and the mass ratio of $16.7^{+1.5}_{-1.1}$ of V404 Cyg, the impact point of the gas stream on the inner disc occurs at a radius $0.18a \pm 0.004a$ from the compact object. The position of the 6.22 hour orbit assuming Keplerian rotation in the disc is $0.12a$, with a 0.2% error and so all that can be deduced from this calculation is that it is in approximately the same region as this source of instability.

The nature of the six hour modulation is still puzzling and several aspects of its behaviour appear to be contradictory. The long term stability over several years of the modulation suggests that it is an orbital feature, but the quasi-periodic changes and radial velocities imply that it is an instability located in the disc. The analysis above shows that the variability of the periodicity cannot be due to beaming or lensing, even during the outburst of V404 Cyg.

This leaves the possibility of X-ray heating or some other intrinsic form of modulation in the flux at this period. The phase of the maximum of the modulation implies that it is the X-ray heated, inner face of some raised structure in the disc. More extensive monitoring of this phenomenon is needed to provide information on its secular variability and dependence on wavelength. This may shed light on the process by which the variation of flux occurs.

References

Al-Naimiy H.M., 1978, Ap. & Space Sci., **53**, 181.

Bao G., 1992, Astron. and Astrophys., **257**, 594.

Bretthorst G.L., in *Lecture Notes on Statistics: Bayesian Spectrum Analysis and Parameter Estimation*, vol. 48, Springer-Verlag, New York, 1988.

Casares J. and Charles P.A., 1992, MNRAS, **255**, 7.

Casares J., Charles P.A. and Naylor T., 1992, Nature, **355**, 614.

Casares J., Charles P.A., Naylor T., Pavlenko E.P., 1993, MNRAS, **265**, 834.

Casares J. and Charles P.A., 1994, MNRAS, **271**, L5.

Charles P.A et al., 1989, in Hunt J., Battrick B., eds, 23rd ESLAB Symp. on *Two-Topics in X-ray Astronomy*, ESA Publ. Division. SP 296, p. 103.

Charles P.A., Kidger M.R., Pavlenko E.P., Prokofieva V.V. & Callanan P.J., 1991, MNRAS, **249**, 567.

Cunningham C.T. and Bardeen J.M., 1973, Ap. J., **183**, 237.

Gotthelf E., Patterson J., Stover R.J., 1991, Ap. J., **374**, 340.

Haswell C.A. and Shafter A.W., 1990, IAU Circ., No. 5074.

Lampton M., Margon B., Bowyer S., 1976, Ap. J., **208**, 177.

Landolt A.U., 1983, A.J., **88**, 439.

Leibowitz E.M. et al. , 1991, MNRAS, **250**, 385.

Lubow S.H., 1989, Ap. J., **340**, 1064.

Lubow S.H. and Shu F.H., 1976, Ap. J. Lett., **207**, L53.

Lucy L.B., 1967, Z. Astrophys., **65**, 89.

Makino F. et al., 1989, IAU Circ., No. 4782.

Meyer F. and Meyer-Hofmeister E., 1982, Astr. Astrophys., **34**, 106.

Richter G.A., 1989, Inf. Bull. Var. Stars, No. 3362.

Sarna M.J., 1989, Astr. Astrophys., **224**, 98.

Shahbaz T., Ringwald F.A., Bunn J.C., Naylor T., Charles P.A., Casares J., 1994, MNRAS. **271**, L10.

Udalsky A., and Kaluzny J., 1991, PASP, **103**, 198.

Wagner R.M. et al., 1990, in *Proc. 11th North American Workshop on CVs and LMXRBs*, Cambridge Univ. Press, 29.

Wagner R.M., Kreidl T.J., Howell S.B. and Starrfield S.G., 1992, Ap. J., **401**, 97.

Zhang X.-H., Bao G., 1991, Astron. Astrophys., **246**, 21.

Chapter 4

Outburst, Decline and Quiescent Photometry of J0422+32

4.1 Near Quiescent Photometry

4.1.1 Summary

I-band photometry of the soft X-ray transient (SXT) J0422+32 was taken during a low brightness state over 2 weeks in October 1993. The object had a mean V magnitude of 19.9, approximately 2 magnitudes above its subsequent, presumably quiescent, level of V $\sim$ 22.4. These data revealed a periodicity at 16.18 hours (significance level $>$ 99%), but there was no evidence for the previously reported 5.1 hour period. It is not possible to account for the physical origin of this 16 hour modulation, but it may be related to the brightness state of J0422+32. It is shown that it cannot be an ellipsoidal effect as the secondary star would be much brighter than the current minimum. The 5.1 hour period could be orbital, with a very late-type ($\sim$M0–M5) secondary which is too faint to be observable if the distance is $\geq$1 kpc.

4.1.2 Introduction

J0422+32 has been extensively studied since its main outburst and discovery in 1992 by the Compton Gamma Ray Observatory (GRO) satellite (Paciesas et al., 1992), when it reached a hard X-ray brightness of 2.9 Crab (40–150 keV). The presence of a high energy tail extending to 600 keV (Harmon et al., 1992; Cameron et al., 1992; Sunyaev et al., 1992) suggested that

J0422+32 was a black-hole candidate by comparison with other such objects (see Tanaka 1991 for a review). The optical counterpart, Nova Persei 1992 (V518 Per), was discovered by Castro-Tirado et al. (1992) at a V magnitude of 12.5. But examination of the Palomar Sky Survey plates showed no object down to approximately 20th magnitude (Mueller, 1992), thereby giving an upper limit to the quiescent brightness of J0422+32. Photometry published by Kato et al. (1992), taken during the decline from the main outburst, show a 5.18 hour modulation at a level of 0.14 magnitudes. They also observe dips with a period of 5.0906 hours which they claim to be the orbital period, and if these dips are due to an eclipse, they imply a high inclination, making this an important system (known eclipsing SXTs are rare at this time). These dips caused them to reinterpret the longer period to be due to superhumps which have been seen in SU UMa systems during their 'superoutbursts' (Vogt, 1980). Using the empirical relation between the mass ratio and the orbital and superhump periods derived for SU UMa systems (Mineshige et al., 1992), they calculate a minimum compact object mass for J0422+32 of $3M_{\odot}$, and have therefore suggested that J0422+32 is a black-hole candidate. Photometry was obtained in October and November of 1992 when J0422 was bright (V $\sim$ 14), and these also show short period variations of the order of five hours (Harlaftis et al. 1994). Similar periodicities were found in outburst by Chevalier et al., (1993b and 1994a) and Shrader et al. (1994).

During the main outburst in 1992, J0422+32 exhibited a slow decline 'plateau' which was followed after some weeks by a much faster decline toward quiescence. The following August in 1993, J0422+32 had a 'mini' outburst with a rise time of a few days, and the whole event lasting for several weeks. J0422+32 then returned to a low, near quiescent state (V $\sim$ 19th magnitude). This pattern was repeated in December of the same year, although the outburst was of a shorter duration, and again in the following January. This behaviour is unprecedented in the class of SXTs. Then, after these events, on January 30, 1994, J0422+32 was observed to have declined to V = 20.67 (Zhao et al., 1994a), which was thought at the time to be its true quiescent level. However, subsequent observations taken in September, 1994 (Zhao et al., 1994b), showed that it had dimmed further to V = 22.39 $\pm$ 0.27, which is

now accepted as the quiescent level. For the overall V-band light curve of J0422+32 see figure 4.1. The open triangles are the 1993 USNO and 1994 NOT V-band magnitudes; the filled squares are the points taken from Shrader et al. (1994); the open squares are those published in IAU Circulars and the crosses indicate photo-electric magnitudes from the Ukraine. The plot shows that after an initially slow decline, J0422+32 went into a faster decay. The linear decay in figure 4.1 indicates an exponential decay in flux.

All the photometry reported prior to 1995 had been taken when J0422+32 was in a bright state (outburst and decline). When SXTs are bright, the light output of the system is dominated by the disc and any variations in the disc brightness, due to the hot spot, or gas flows through the disc, dominate any contribution from the secondary star which is much fainter (unless there is significant X-ray heating). The modulation searched for here is the ellipsoidal effect which is due to the changing aspect of the distorted secondary star with respect to us (see Tjemkes et al., 1986). To detect this modulation, photometry was obtained once the object had faded substantially from its outburst level.

4.1.3 Observations

Time-resolved differential CCD photometry of J0422+32 was obtained with the 1.0 m Ritchey-Chrétien reflector of the United States Naval Observatory Flagstaff Station (USNOFS) from October 26 to November 7, 1993. Of the 13 nights dedicated to this program, 5 were lost due to cloud cover or the proximity of the moon in the sky to J0422+32. The USNOFS Texas Instruments 800 x 800 CCD camera was used in combination with Johnson V-band and extended I-band filters (see Wagner et al. 1992 for a description of the extended I-band filter) to image a 5.7 x 5.7 arcmin region of the sky at a scale of 0.43 arcsec/pixel. The seeing during the run typically averaged 2 arcseconds. Several bias and twilight sky flat field frames were also obtained each night. These bias and twilight flat fields were separately median combined within IRAF using IMCOMBINE. The image frames of J0422+32 were then de-biased and flat fielded with CCDPROC, using these master biases and twilight flats from each individual night.

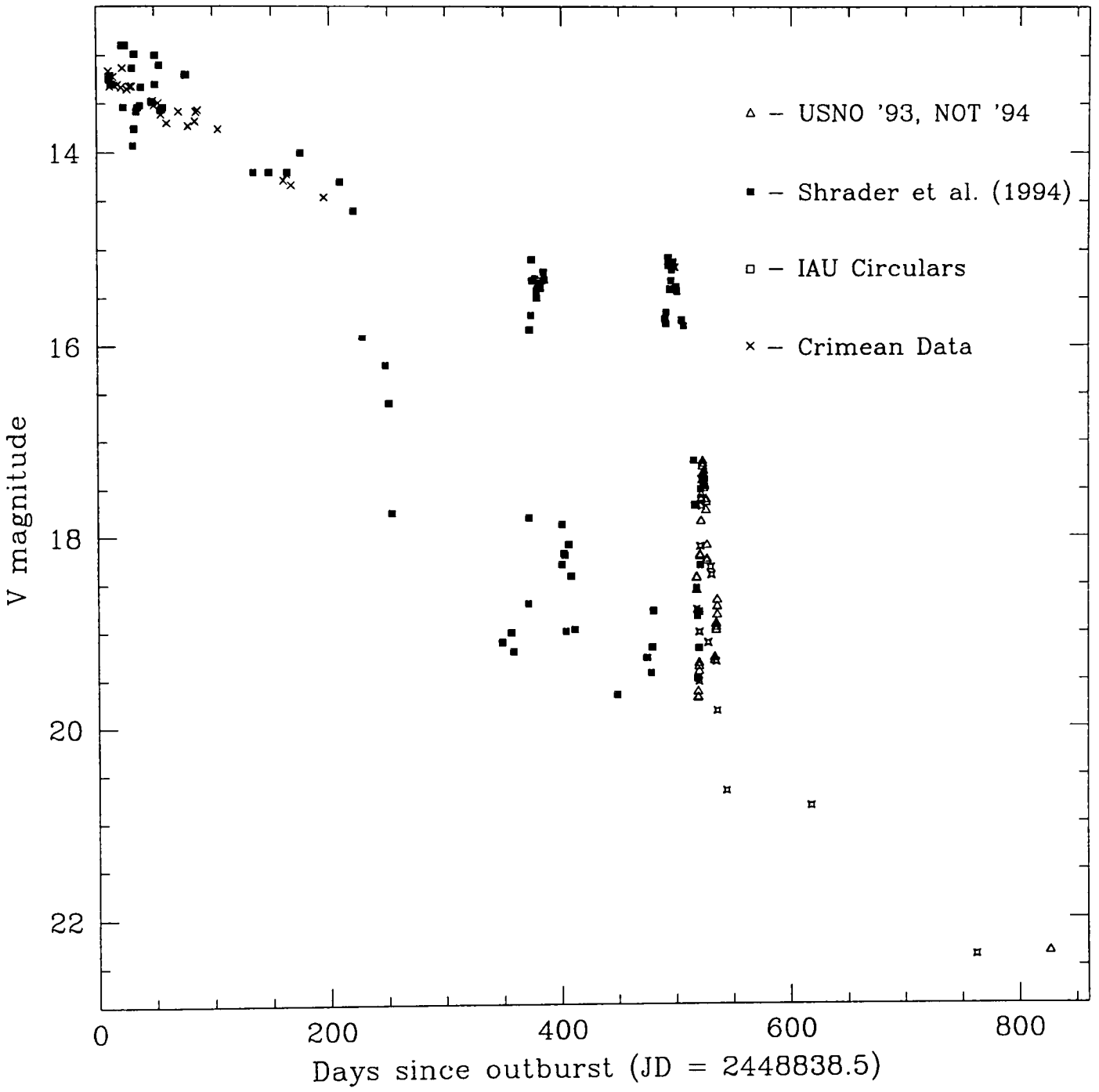


Figure 4.1: The overall light curve of J0422+32 to date. The horizontal axis is the number of days since the initial outburst, which was taken to be JD = 2448838.5 and the magnitudes are in either the V or visual bands. On this plot the decay after outburst is linear which indicates that the flux decayed exponentially.

On October 26 and November 7 several V-band CCD frames of J0422+32 were taken to compare with the overall V-band light curve, from which it was found that V = 19.6 and 20.3 respectively. Subsequent monitoring of the overall V-band light curve showed that the photometry was obtained in the middle of a low brightness state of J0422+32 and between two 'mini' outbursts which occurred in 1993 August and December (Filippenko and Matheson, 1993; Zhao et al., 1993). If it is assumed that at V = 22.4 the light is dominated by the secondary, the contribution of the disc and secondary were comparable during these observations. This suggests that it should have been possible to detect the ellipsoidal modulation of the secondary, given the high inclination implied by the observation of absorption dips.

In addition to the V-band photometry that was obtained on the first and last nights, time-resolved differential I-band photometry was obtained in an effort to detect the ellipsoidal light variations of the late-type secondary star. The integration time was 900s with 30s of deadtime between frames for readout and storage. The average duration of the observations for each night was 7.2 hours (maximum: 8.2 hours, minimum: 6.4 hours) and thus cover more than a single 5.1 hour cycle per night. The brightness of J0422+32 with respect to two reference stars was measured by two-dimensional digital aperture photometry. The photometric accuracy of the two reference stars' relative magnitudes, was approximately 0.015^m (1-sigma) and showed no systematic trends.

4.1.4 Results

Power Spectrum Analysis

The overall light curve of J0422+32 shows clear night to night variations of ~ 0.2 magnitudes as can be seen in figure 4.2, with little short term modulation. This behaviour looks markedly different to that observed during outburst, where no long term modulation was found and non-sinusoidal 5 hour variations were apparent. A Lomb-Scargle (Lomb, 1976; Scargle, 1982) periodogram was used to search for any periodicities present in the data, this periodogram is shown in figure 4.3. Large peaks at 16.18 ± 0.05 hours (1.48 cycles/day) and 50.11 ± 0.50 hours (0.48 cycles/day) are seen, but there is surprisingly little power at the 5.1

hour period (4.71 cycles/day) which was present in outburst. The window function shows the usual peak at 24 hours, caused by the gaps in the data sampling.

Light curves folded on the 16.18 and 50.11 hour periods are shown in figures 4.4 and 4.5. A least squares sine fit was applied to the data folded on 16 hours and gave an amplitude of 0.28^m. It is clear from the clustering of the data points that the 50 hour peak is the 24 hour alias of the 16 hour peak. The strength of the 50 hour peak is due to power leakage from the 16 hour periodicity, which is exacerbated by noise in the dataset. It is quite easy for the noise to cause alias frequencies to exhibit more power than their corresponding signal frequencies (Scargle, 1982). To test this an artificial dataset was created with a sinusoid of amplitude 0.28^m at the 16.18 hour period and a noise level with $\sigma = 0.10^m$ which was obtained from the IRAF error on J0422+32, calculated by APPHOT. It should be emphasized that the σ of 0.10^m for J0422+32 represents a combination of statistical noise and intrinsic random variability ($\sigma = 0.07^m$ for a star of the brightness of J0422+32). The periodogram for the simulated data is very similar to the observed one, with comparable strength peaks at 16 and 50 hours. The observed data were also analyzed for any evidence of a 5.1 hour periodicity. As is shown in figure 4.6, a fold of the data at this period shows only a scatter of $\sim 0.10^m$, with no obvious modulation.

The data show a small downward slope at the beginning of the run. So to detrend this a straight line fit was made and subtracted from the data, and a new Scargle periodogram calculated. Some of the aliasing between the two large 16 and 50 hour peaks disappears, but the 16 and 50 hour peaks are still approximately the same size. The reason for this is that detrending with a linear fit has roughly the same effect as removing a large period sine wave from the data. This means that aliases that are close together in frequency space are removed.

4.1.5 Significance Tests

The significance of the periodicity in the data was calculated by modelling the data using a Monte-Carlo simulation, to model the effects of Gaussian noise. The noise σ of the data

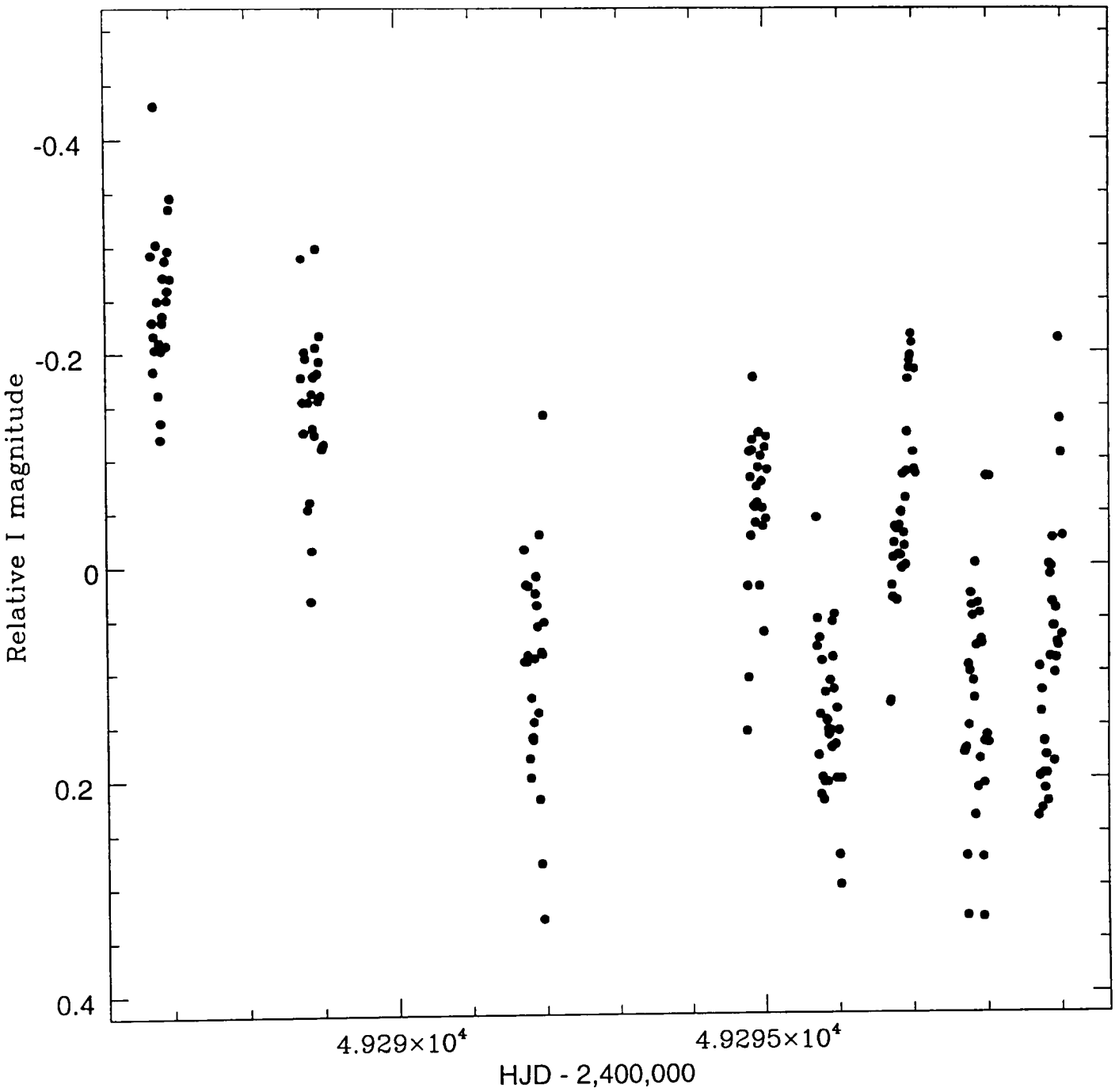


Figure 4.2: Overall I-band light curve of J0422+32 taken on the 1.0 m telescope at the US Naval Observatory from 26/10/93 to 07/11/93 (8 nights), nightly variations can be seen as well as a possible downward trend at the beginning of the observing run.

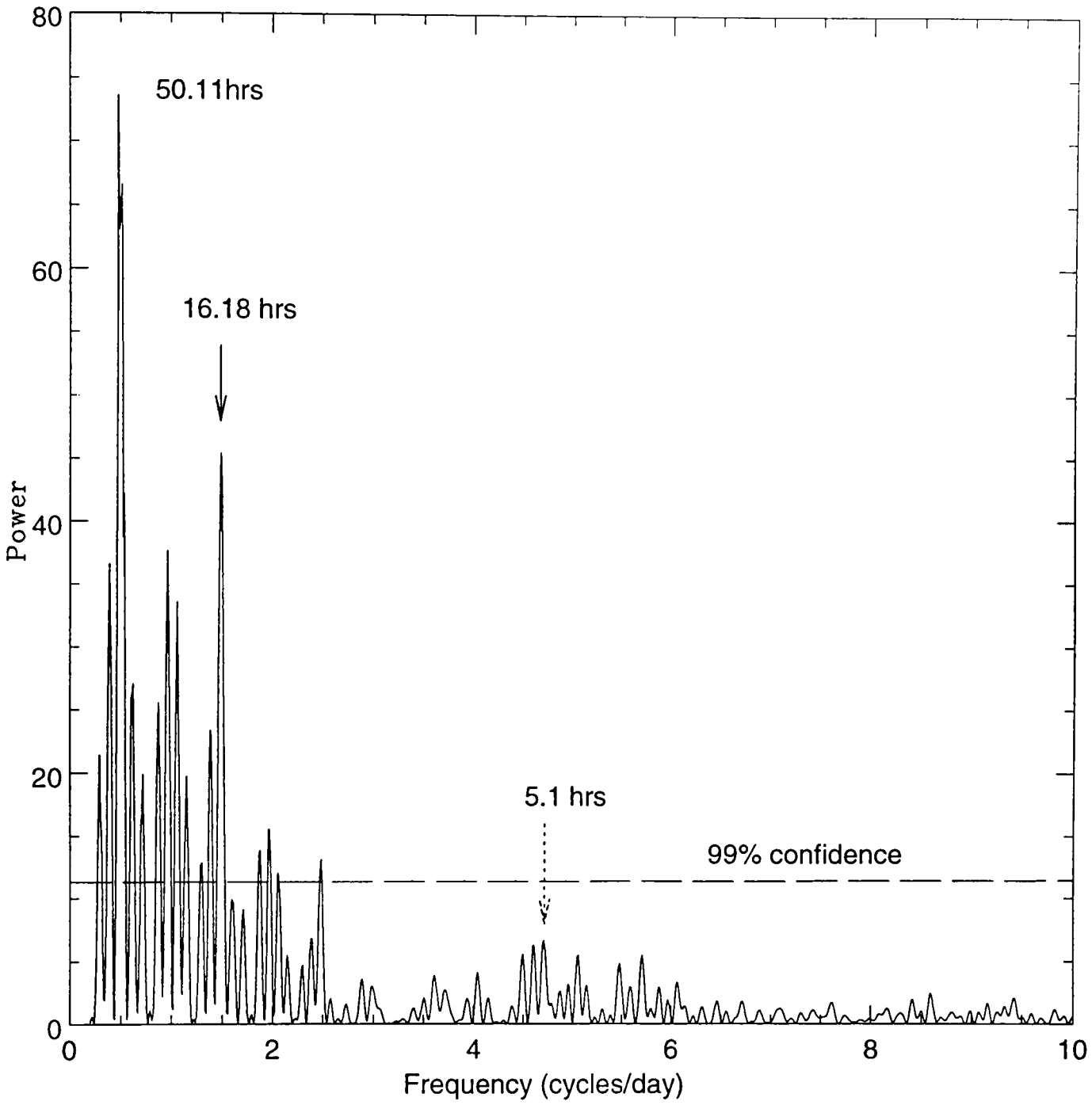


Figure 4.3: Scargle periodogram for entire dataset, showing 16.18 and 50.1 hour peaks. The 99% confidence level as derived in the text is also shown.

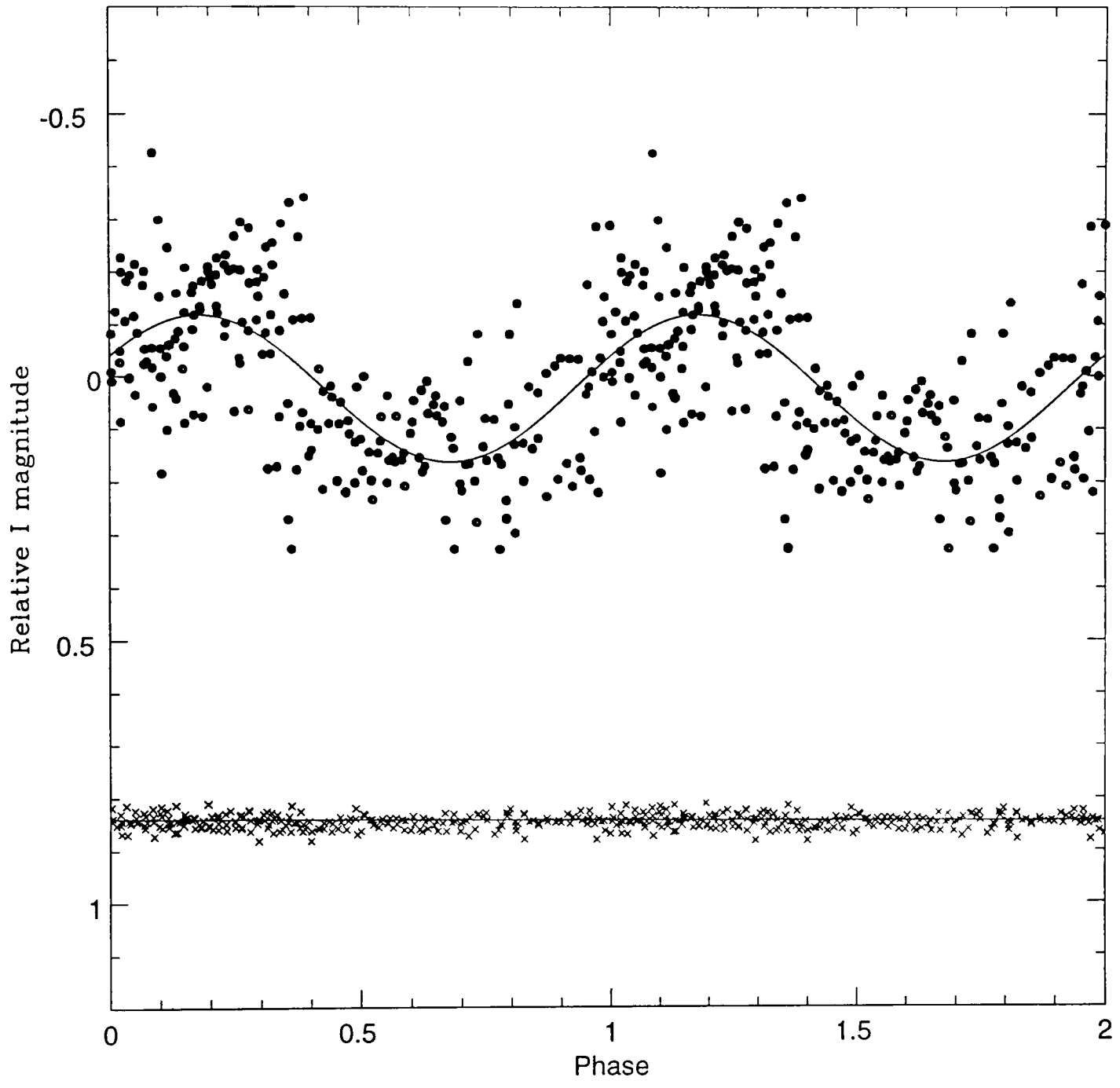


Figure 4.4: Data folded on the 16.18 hour periodicity, (top) the light curve of J0422+32 relative to a nearby comparison star and (bottom), the comparison relative to the check star. The sinusoid plotted is the least squares fit to the data (amplitude = 0.28^m, and a straight line at the mean level of the relative magnitude of the reference stars is also shown (sigma = 0.015^m).

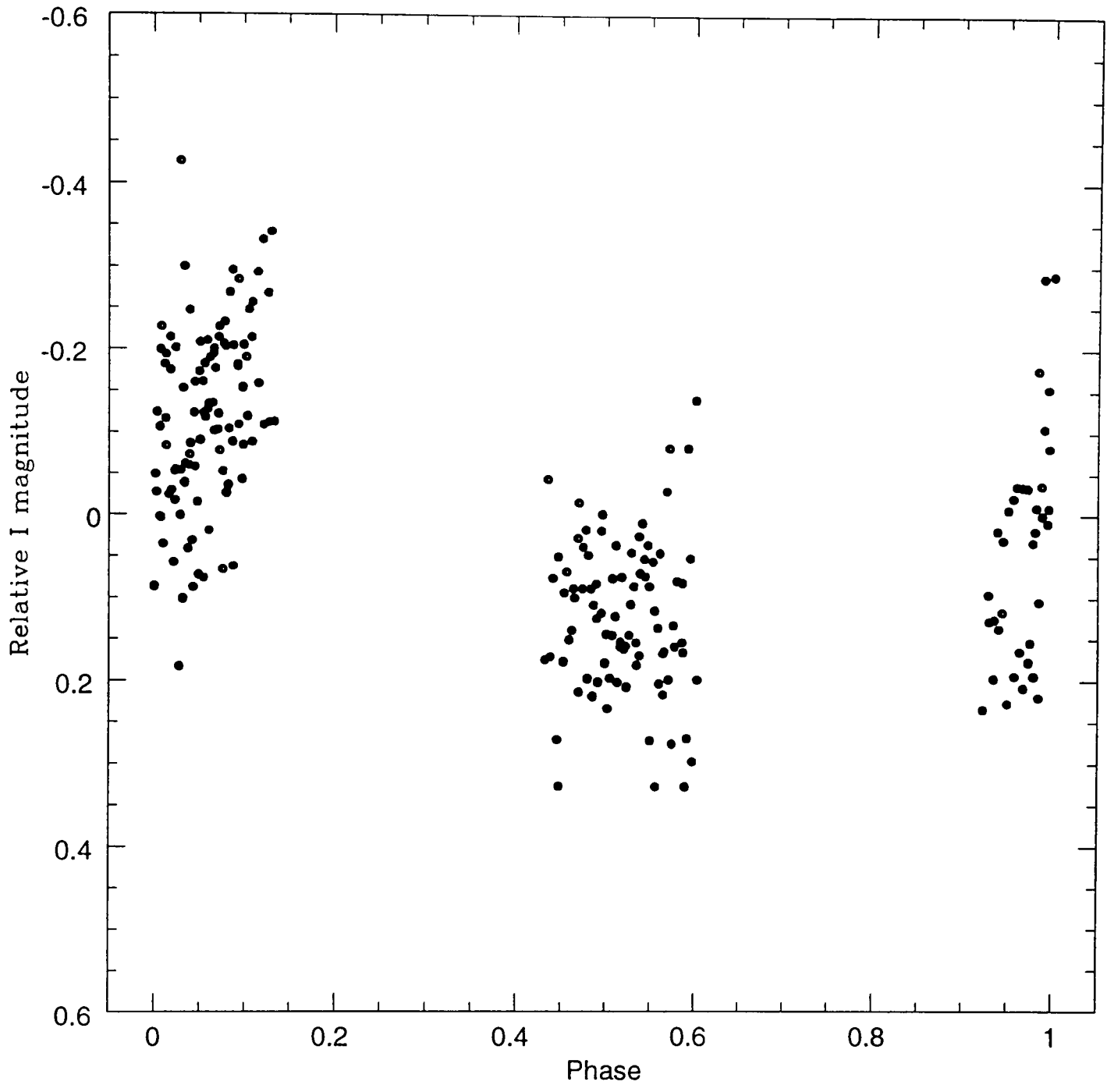


Figure 4.5: Data folded on 50.11 hour periodicity, a close 24 hour alias of 16.18 hours. The large amount of power seen at this period is caused by power leakage from the 16.18 hour peak, and also because of the clustering in the folded light curve, as it is a multiple of the 24 hour observation interval.

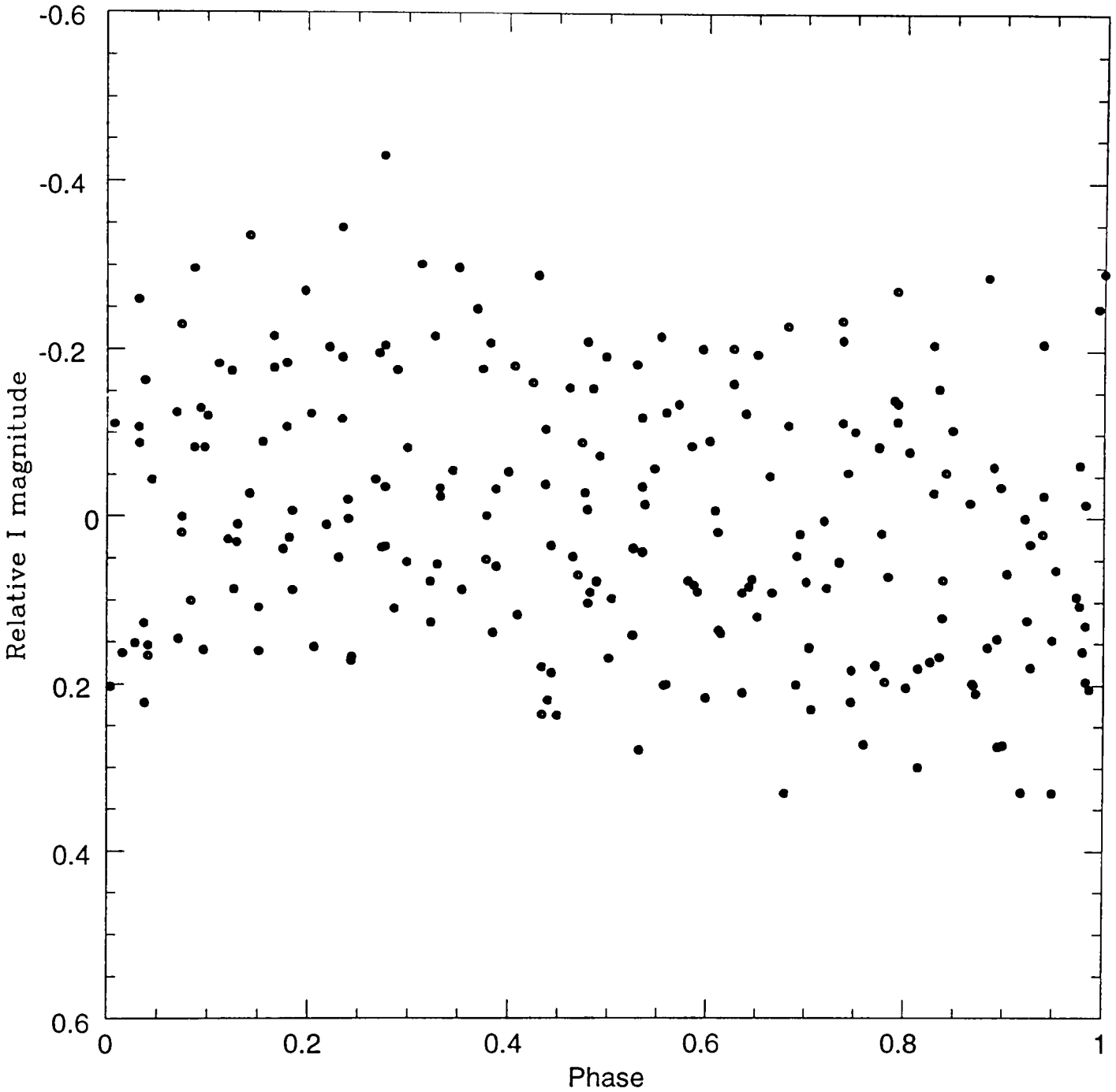


Figure 4.6: Data folded on the 5.1 hour previously reported periodicity (Kato et al., 1992). The scatter on the data is 0.10^m.

about its mean was calculated as 0.156 magnitudes, and 10,000 Gaussian noise sets created using this value, with the same sampling as in the observed dataset. A Lomb-Scargle periodogram was calculated for each dataset over the Nyquist interval. The maximum peaks in the periodograms of these data were calculated over this frequency range, and a cumulative distribution function curve obtained. This CDF (see figure 4.7) gives us the distribution of peaks that would be expected if all of the variation in the data was caused by (random, normally distributed) noise. The 99% significance level is obtained from figure 4.7, at the point where the left curve crosses $p = p_0 = 0.01$; this is the probability of a noise fluctuation exceeding the power z_0.

The 99% significance power level obtained in this way occurs at a power level of 11.47, but the 16 hour peak reaches a power of 45.44, and so is > 99% significant. No peaks around five hours approach the 99% significance level. To quantify the apparent absence of the 5.1 hour period, another Monte-Carlo simulation was run with Gaussian noise plus a signal at this period. The noise level used was $\sigma = 0.10^m$, the calculated error on the magnitude of J0422+32, plus the intrinsic variability. This was obtained from the aperture photometry error estimation in APPHOT, and the full amplitude of the sinusoidal signal used was 0.14 magnitudes. This amplitude was seen by Kato et al. (1993) in the V-band, and so gives us an estimate of the I-band variation. Figure 4.7 shows the noise and noise plus signal CDFs for the data. The power level Z (= $P_X(\omega)$ for a set of data $X(t_i)$, i = 1,2,... N_0, is defined as $P_X(\omega) = 1/N_0|FT_X(\omega)|^2$, where $FT_X(\omega)$ is the discrete Fourier transform of the data.

The 99% ($p_0 = 0.01$) level of the Gaussian noise CDF gives $z_0 = 11.47$, the 99% confidence power level. For the noise plus signal case the probability of non-detection can then be calculated for the 5.1 hour period, where the CDF (of noise plus signal) crosses z_0. This was calculated from the CDF to occur at p = 0.9932, see figure 4.7. Thus the probability that the signal at 5.1 hours would not have been detected, with the above quoted signal-to-noise ratio, at the 99% confidence level is 0.7 %. It was also calculated that the 5.1 hour modulation could have been detected down to a level of 0.02^m (estimated from the amplitude required to obtain a power level of 11.47).

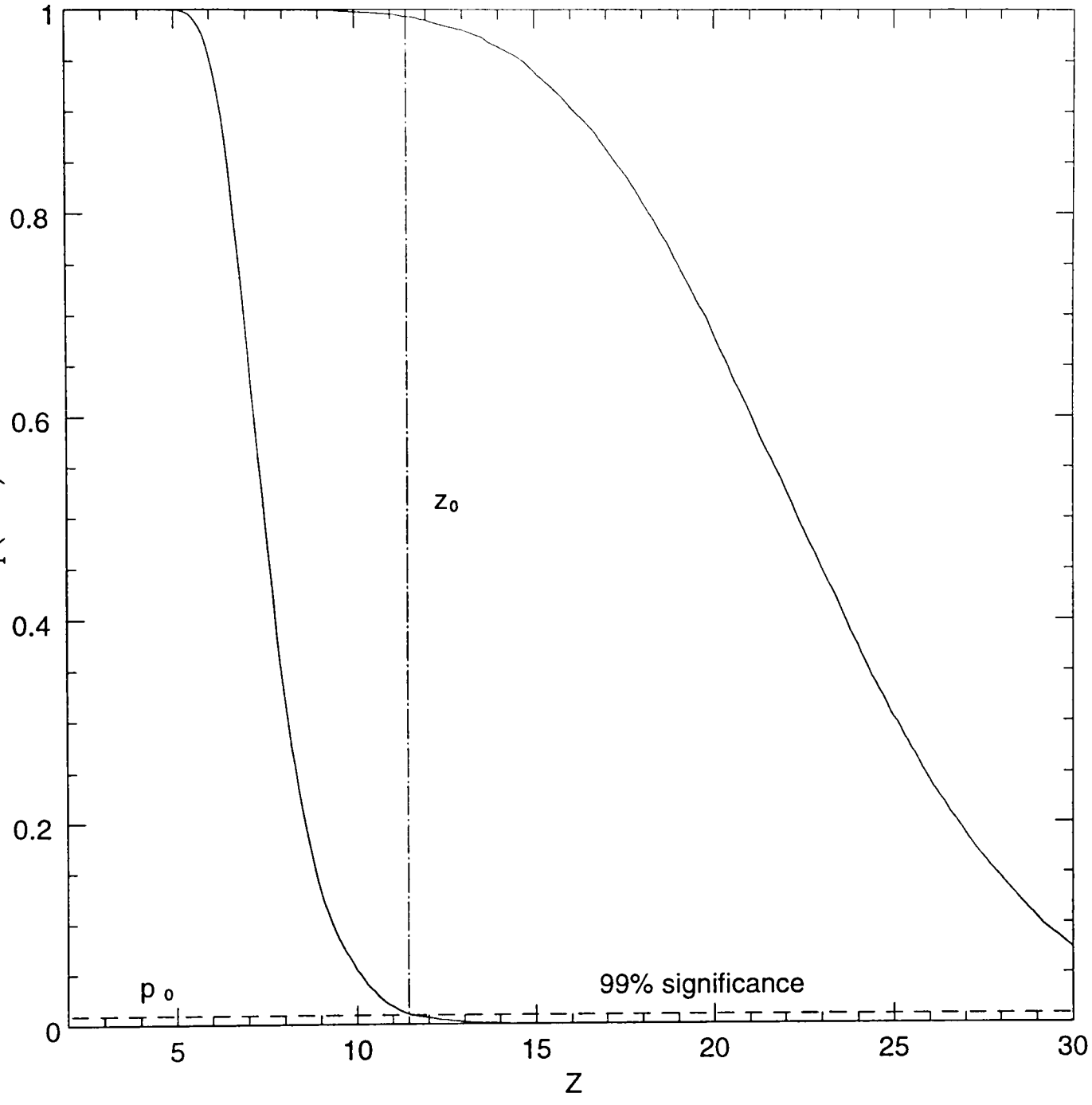


Figure 4.7: Cumulative distribution curves for Monte-Carlo simulations with Gaussian noise (left curve) and noise plus signal at 5.1 hours (right hand curve), with the power Z plotted versus the cumulative probability p(z>Z). The line z_0 indicates the power required to reach 99% significance, the intersection of this line with the signal curve gives the non-detection probability of the 5.1 hour period (0.007).

4.1.6 Discussion – An Ellipsoidal Effect?

If the 16.18 hour periodicity were related to the secondary and not to the disc, then the orbital period would be 32.4 hours (in the absence of X-ray heating), as the ellipsoidal modulation of the secondary will produce two minima per cycle. These ellipsoidal modulations occur because of the changing aspect of the distorted secondary with respect to us, combined with limb and gravitational darkening (see Tjemkes et al., 1986). These effects cause a double modulation per orbital cycle with minima occurring at phases 0 and 0.5. The phase 0.5 minimum (when the inner face of the secondary is towards us) is the deepest due to the larger gravity darkening in that direction, and so there should be differing minima in the light curve when it is folded on the (correct) orbital period. The folded light curve (32.4 hours) is shown in figure 4.8. It features minima that differ by Δm = 0.06^m, but also different maxima, with Δm = 0.05^m.

If this light curve is the true ellipsoidal modulation of the system, the differing maxima can be accounted for in several ways. One explanation involves the secondary having a slightly eccentric orbit, as has been seen in the high mass X-ray binary system Vela X-1 (Tjemkes et al., 1986). There, an eccentricity of 0.092 (Rappaport et al., 1976) leads to a difference in the ellipsoidal maxima of $\Delta V = 0.015^m$. If periastron occurs at phase 0.25, and it is assumed that the secondary star fills its Roche lobe at this phase, then the deformation of the secondary is maximal. At phase 0.75 (apastron) it under-fills the critical lobe and so is less deformed causing the maxima to be unequal. However, this effect produces a small difference in the maxima, unless the eccentricity is much larger than above, and so would be very difficult to measure in any of the currently known SXTs as they are extremely faint.

The differing maxima have also been seen in the soft X-ray transients A0620−00 and GS2000+25 (Haswell et al., 1993; Chevalier et al., 1993a), but at a much higher level. The maxima in A0620−00 differed by $\sim 0.10^m$ in the U, B, V, and R-band light curves, the maximum at phase 0.25 was the larger, which might be explained as above with periastron occurring near this phase. However, Haswell et al. (1993) model it as a phase varying

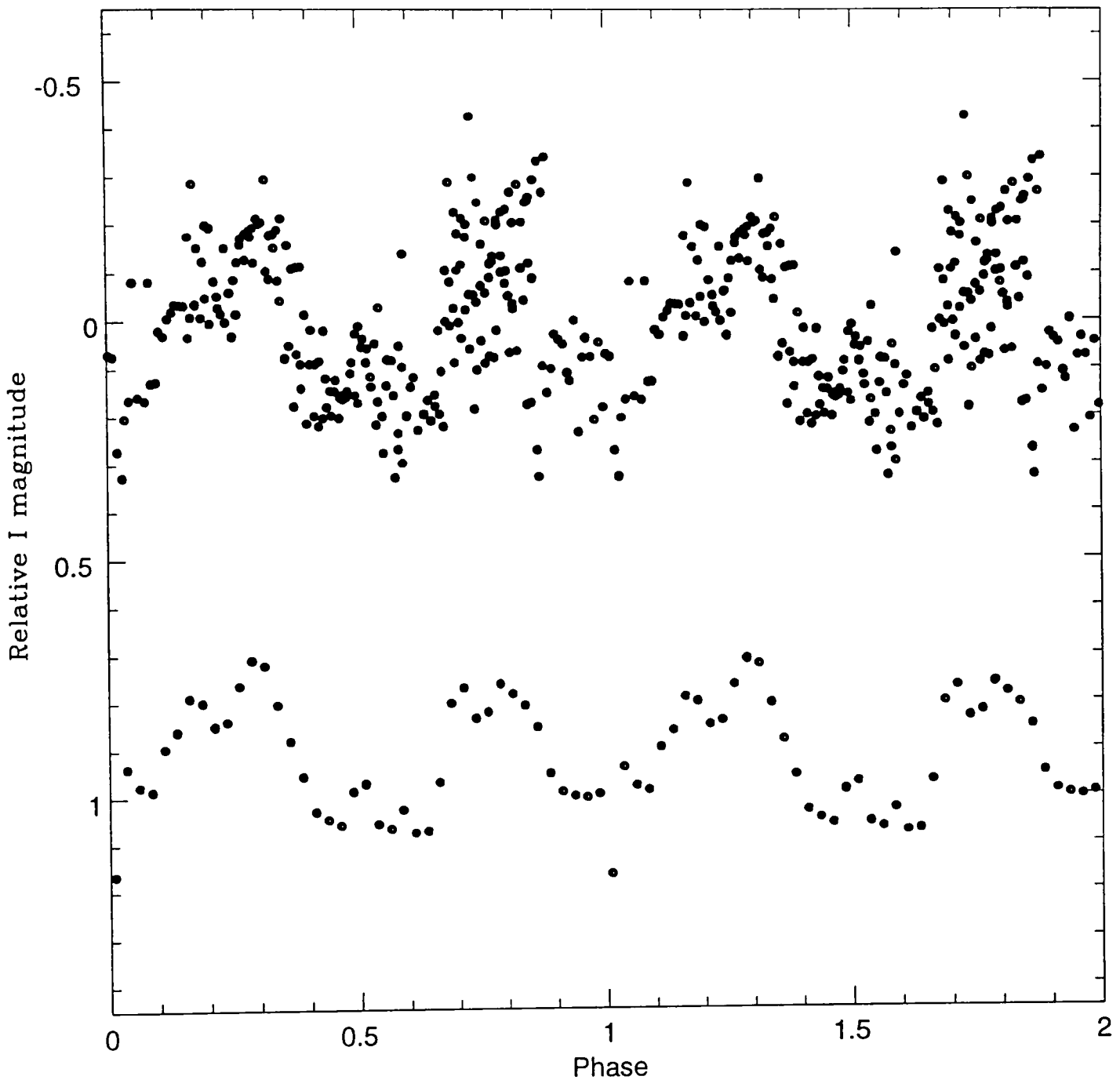


Figure 4.8: Data folded on 32.4 hour ellipsoidal modulation, the top curve shows the raw folded data and the bottom curve shows the data sectioned into 40 bins 0.025 apart in phase. The light curves show minima differing by 0.06 magnitudes, but also unequal maxima (Δmax $\approx 0.05^m$) which could be accounted for by a slightly eccentric orbit.

component in the disc which peaks at phase 0.25, without invoking a detailed model for this phase variation. Other observations show the larger maximum at phase 0.75. In GS2000+25 the ellipsoidal light curves folded on the orbital period (8.258 hours) have the larger maxima and minima alternating on successive nights. This was explained as a 10 hour distortion wave superposed on the main period. It is possible, though, that the alternating maxima are due to starspots migrating across the surface of the secondary.

4.1.7 Relation to the 5.1 Hour Period

Unfortunately, no spectra were taken simultaneously with these observations, the next spectroscopic run occurring during the following 'mini' outbursts. Casares et al. (1995a) observed a 5.1 hour S-wave component in their HeII λ4686 spectra with a radial velocity of 755 kms^{-1}, but no long period variations. The component they see is probably due to the hot spot, whereas the radial velocity variation in their spectra and that of Filippenko et al. (1995) has a semi-amplitude of $\approx$ 380 km s^{-1}, which comes from the secondary star.

The 16 or 32 hour modulation seen here appears to be unrelated to the five hour variation seen in outburst. More importantly, this long period modulation only seems to occur when the object is faint ($\leq$ 18th magnitude), and the shorter period only when it is bright.

The fact that the five hour period only occurs in the bright state suggests that it is related to the disc or vicinity of the compact object, and if it is the orbital period, may be caused by the bright spot on the disc or X-ray heating of the secondary. If this is true then the source of the 16 hour modulation is very puzzling. The observations took place with J0422+32 at a mean V of 19.9, approximately 2 magnitudes above quiescence (assuming that V = 22.4 is the true quiescent level); the disc was therefore $\sim$ 9 times the secondary's brightness, and so for the 0.28 magnitudes variation to be entirely due to the secondary, it would have to vary by $\sim$ 140% (although in the I-band the decrease will be smaller than the V-band, see later). It is not possible to produce this large a variation with ellipsoidal models of the secondary. The periodicity could be caused by a residual variation in the disc, but it is still surprising that no

modulation near the well established 5.1 hour period was detected. This could be explained by the quiescent magnitude of V = 22.4, which means that the light was still dominated by residual flux from the disc, during the run of observations. The secondary is therefore an extremely faint low mass object, which would explain the difficulty to date in finding late-type features in the spectra that have been obtained (Bonnet-Bidaud and Mouchet, 1994).

4.1.8 Nature of the Secondary Star

One can combine the observed periods with the current quiescent brightness to set limits on the nature of the secondary star and its distance. Assuming only that the secondary star fills its Roche lobe (a necessary condition for mass transfer) it is possible to follow the argument of King (1993) as was applied to V404 Cyg. The mean density of such a star is given by $\rho_2 = 110 P_{hr}^{-2}$ g cm^{-3}, which leads to values of 4.24 and 0.11 g cm^{-3} for the 5.1 and 32.4 hour periodicities, respectively. Clearly the longer period requires an evolved secondary, as late-type main sequence stars have densities in the range 1.4 (G2) to 10 (M5) g cm^{-3} (Allen, 1973).

Now, for the 32.4 hour period, King's (1993) equation (4) becomes: $0.084\, M_2^{0.5} = (M_c/0.25)^{7.65}$. This implies a minimum mass for the secondary star (when $M_2 = M_c$, the core mass) of $0.16 M_\odot$ (i.e the Schonberg-Chandrasekhar limit), and hence a luminosity of $0.9 L_\odot$. Using an A_V of 0.72 and a distance to J0422+32 of 2.4 kpc, obtained by Shrader et al. (1994), with a bolometric correction of -0.6, the derived m_v is 18.09. This can therefore be ruled out by the current much fainter m_v of 22.4.

Also, if the 5.1 hour dip period is orbital, then the observed mean density would imply a spectral type in the range M0–M5. This would be consistent with the current minimum m_v for distances $\sim$ 1 kpc, which requires $M_V \sim 9.7$. The information on late-type binaries and

M_V values from Allen (1973) is collected together in table 1.

Table 1. Physical quantities for late-type secondaries (Allen, 1973).

Spectral type	Mean density g cm^{-3}	M_2 ($M_\odot$)	R_2 ($R_\odot$)	M_V	$V-R$	$V-I$
K0	1.8	0.78	0.85	+5.9	0.74	1.4
K5	2.4	0.69	0.74	+7.3		
M0	2.5	0.47	0.63	+9.0	1.1	2.2
M5	10	0.21	0.32	+11.8		

4.1.9 Comparison with other X-ray Binaries

The recurring outburst behaviour of J0422+32 is very unusual. It is reminiscent of the neutron star transient Aql X-1 (4U1908+00) which has repeated 'mini' outbursts on a time-scale of the order of one year (Charles et al., 1980). Also, the X-ray binary system GX 339−4 (thought to be a black-hole candidate on the basis of its rapid X-ray variability) shows high, low and off states (Markert et al., 1973) which could be analagous to the outburst, near quiescence, and quiescence in J0422+32. Between the outbursts, J0422+32 did not return to quiescence, the lowest it ever reached was V $\approx$ 19, which would seem to be a low state. Only after the last 'mini' outburst did J0422+32 finally reach its current quiescent state, $\sim$ 2 years after the initial outburst event. GX 339−4 also shows spectral changes that are similar to the soft X-ray transient black-hole candidates. These systems are fundamentally very similar – the accretion processes occurring should be independent of the type of compact object, and so it is expected that there should be comparable modes of behaviour in each. J0422+32 would seem to be a system whose nature is intermediate between the X-ray transients, and those which show more than one mode of accretion.

4.2 Outburst photometry

4.2.1 Introduction

The outburst observations were taken at the USNO in Flagstaff using the USNOFS Texas Instruments 800 x 800 CCD camera and extended I-band filter. All other technical details of

the observations and subsequent reduction are identical to that in section 4.1.3. The outburst photometry was obtained during 1994 January 3–25. Conditions were good, allowing a total of 610 frames to be taken over 15 nights, out of a possible 22. At the start of the run, J0422+32 was declining from a previous 'mini' outburst, after which it went into a second outburst, reaching a peak I-band magnitude of $\sim$ 17. Subsequently, J0422+32 decayed over several days down to a magnitude of $\sim$ 19 (for the overall light curve see figure 4.9). To study the orbital modulations in this system it was therefore necessary to remove these secular variations from the overall light curve.

The changes in the light curve can be seen to divide into roughly 3 parts: an initial decline of 2 nights duration; the mini-outburst, lasting 8 nights; a final decline over the last 5 nights of the run. So, to remove these long-term variations, the data were split into three sections and polynomials were fitted to the data. The F-statistic was used to determine what order of polynomial was necessary to fit the data. This is a measure of the significance of the fit (strictly speaking, the F statistic used here is the significance that the last term in the polynomial is different from zero). However, as the data appear in blocks separated by a day, not each point is independent when looking at long term trends and so only the first and last points of each night were taken into account when evaluating the significance. The form of the F-statistic used here is given by:

$$F(2/1) = \frac{(res^2_{(1)} - res^2_{(2)})/(\nu_1 - \nu_2)}{res^2_{(2)}/\nu_2} \tag{4.1}$$

The terms $res^2_{(1)}$ and $res^2_{(2)}$ are the sums of the squared residuals for polynomials of order 1 and 2, respectively; ν_1 and ν_2 are the degrees of freedom for the two polynomials. The value of F(2/1) will be high if a linear fit to the data is much better than constant, F(3/2) will be large if a parabola fits better than a line, and so on. The values of F for increasing orders of polynomial applied to the 3 sections of data are shown in table 4.1.

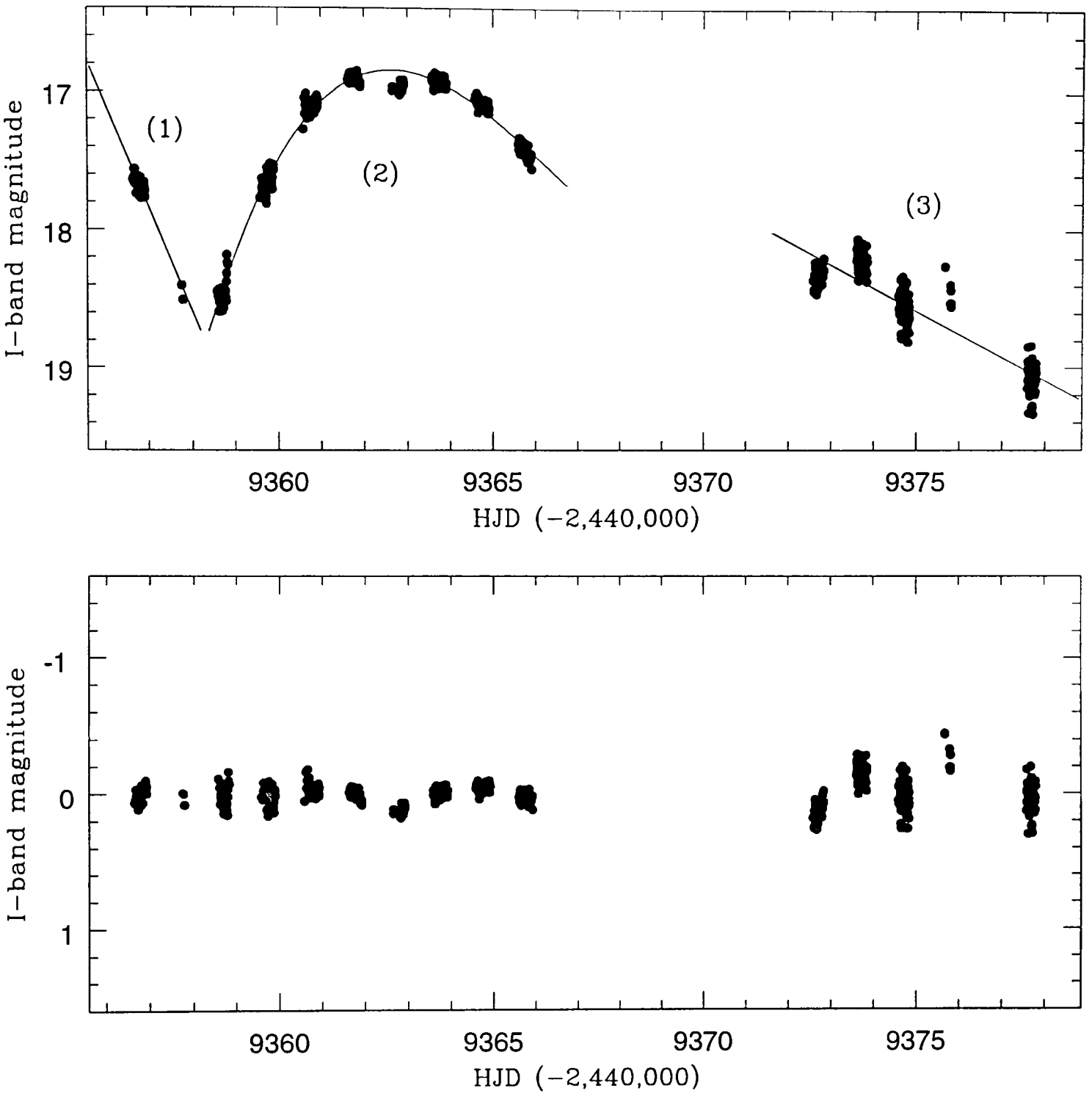


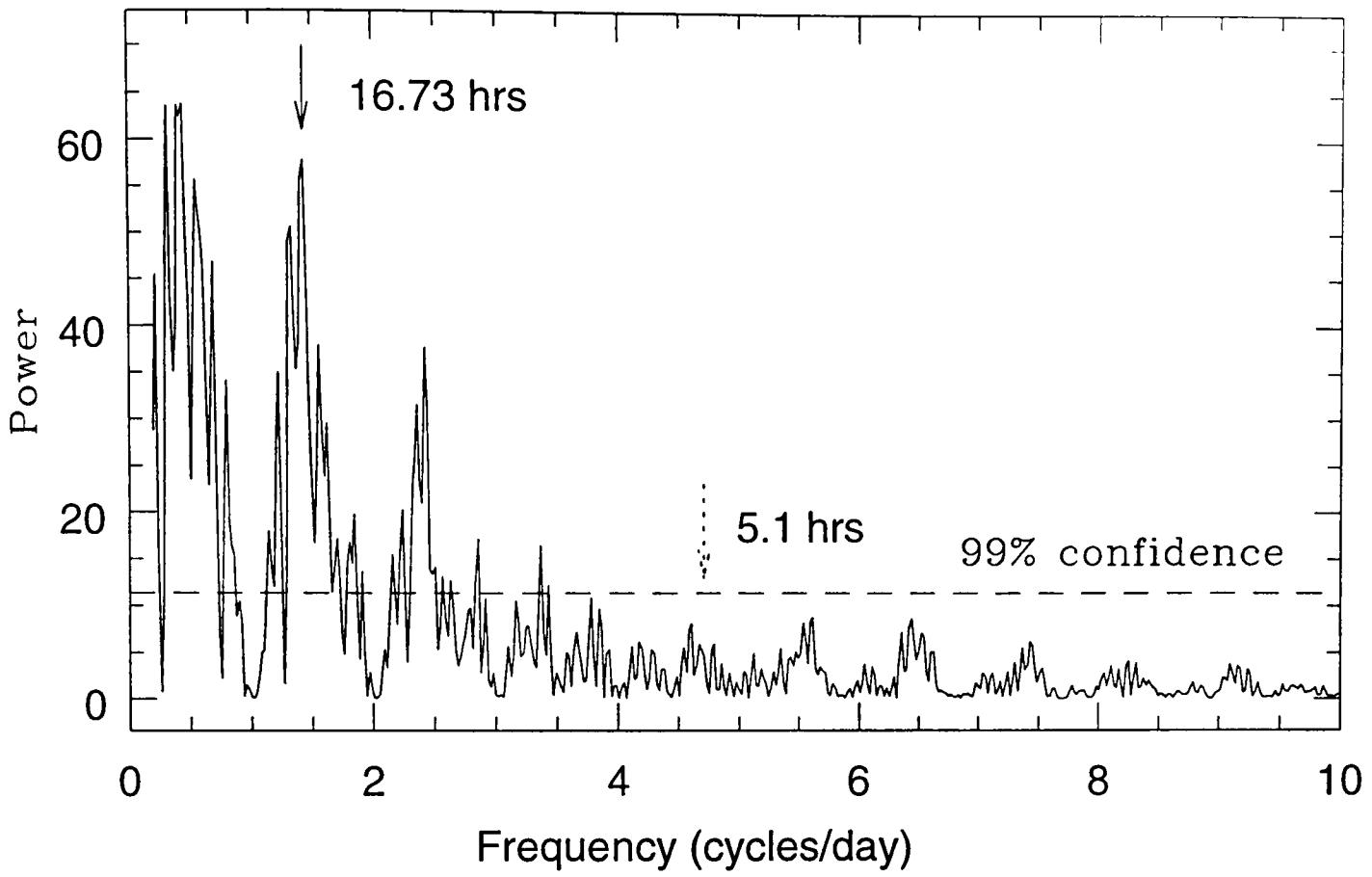
Figure 4.9: The outburst light curve taken during the January 1994 outburst, in the I-band. The polynomial fits to the 3 sections of data are plotted over the top of the data points (solid curves), and in the bottom panel the residuals are plotted, after subtraction of the fits.

Table 4.1: Significance tests for the polynomial fits to the data.

F	Section 1	Section 2	Section 3
(Linear/Constant)	28.48	5.14	31.77
(Parabolic/Linear)	0.31	249.68	2.59
(Cubic/Parabolic)	–	13.40	–
(Quartic/Cubic)	–	1.30	–

Values in the table higher than the corresponding critical value taken from F-statistic tables, indicate that the fit is significantly better than the previous one with 99% confidence. Thus, a linear fit is adequate for sections 1 and 3 of the data, but a cubic must be used for section 2 (e.g. the final value for section 2 gives a confidence level which falls below the 99% confidence level of 3.54, which means that the quartic does not fit the data significantly better than a cubic). The table quantifies the fitting procedure that would normally be performed by eye. These polynomials were then subtracted from the data, to give the bottom panel of figure 4.9. The residuals appear to be free from changes in the light curve due to the 'mini' outburst, but appear to vary in a periodic manner. To search for periodicities these data were analyzed using the Lomb-Scargle periodogram.

The Lomb-Scargle periodogram of the residuals of the above dataset is shown in figure 4.10, as well as the data folded on the peak at 16.73 $\pm$ 0.3. This peak is close to 16.18 hours, and has $>$ 99.9 % confidence. There is little power at the orbital period of 5.1 hours. The discrepancy between the frequency of the two 16 hour periods could be due to the fact that these data were taken in outburst and had to be detrended. This can cause quite large shifts in the estimated frequency, which could easily be large enough to explain this difference. The amplitude of the 16.73 hour periodicity seen here (0.12^m) is much less than that seen previously (0.28^m). Although, as can be seen on the bottom panel of figure 4.10, the modulation seemed to increase in amplitude as J0422+32 decayed back towards 19th magnitude (in the I-band). The 16 hour period has not been seen in quiescence, which implies that the size of this variation becomes larger with increasing disc brightness. The fact that this period was seen around 19th magnitude in the I-band, at which level J0422+32 stayed constant for at least six months, could be suggestive of an intermediate source state. This is comparable to GX 339−4 which exhibits at least three states (Grabelsky et al., 1995). In the *high* or *soft* state, the soft X-ray flux is at its maximum, and the hard tail is weak. The *low* or *hard* state has a much lower soft X-ray flux and a more luminous X-ray tail extending up to several hundred keV. Finally the *off* state is characterized by undetectable X-ray flux in the 2–10 keV range. These three states are analagous to outburst, inter-outburst and quiescent modes, respectively which have



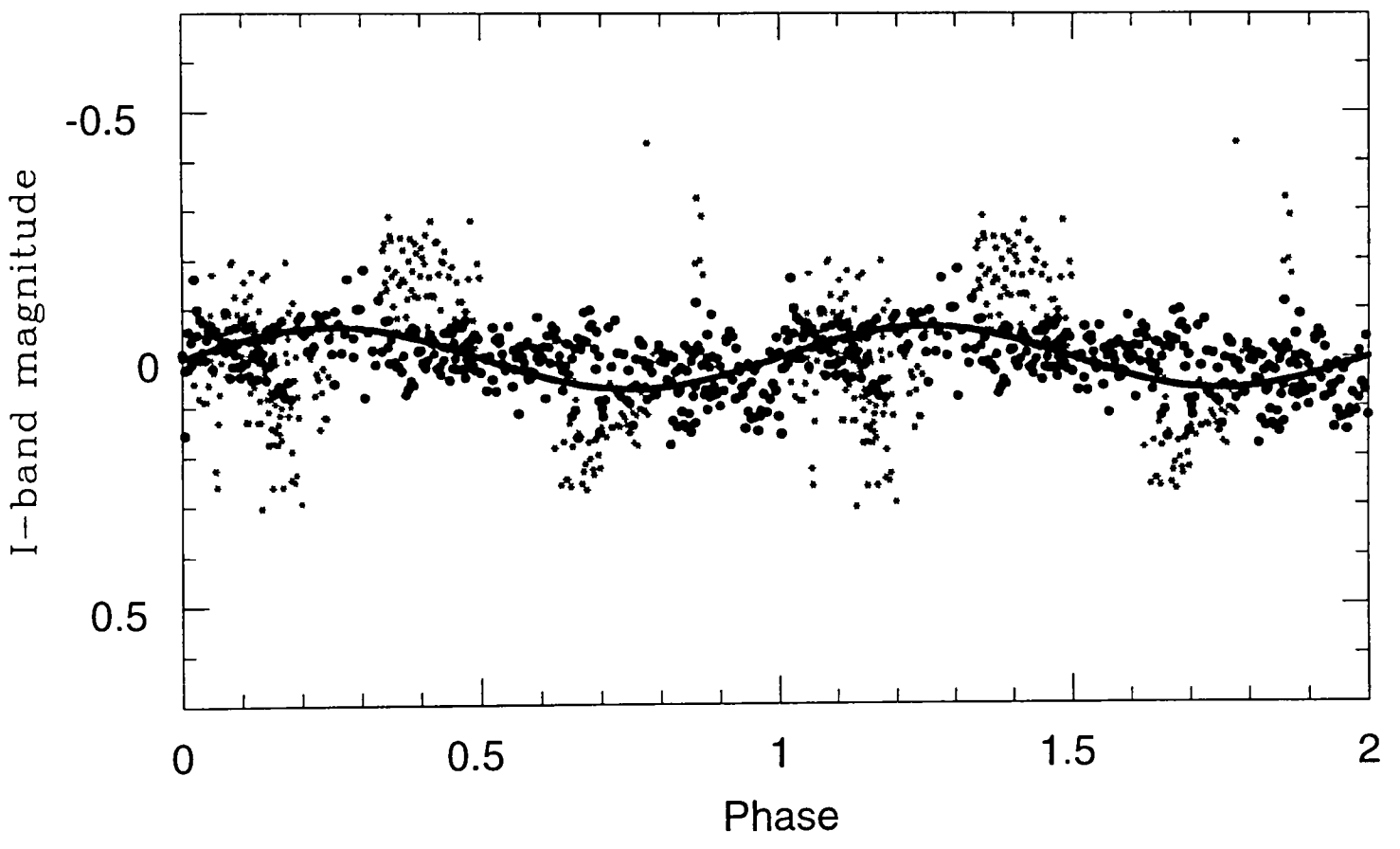


Figure 4.10: The Lomb-Scargle periodogram of the detrended outburst data. There is a significant amount of power at the period 16.73 ± 0.3 hours, but no detectable peak at 5.1 hours (the 99% confidence level is shown as a dashed line). The bottom panel shows the light curve folded on 16.73 hours, with a sinusoidal fit to the data (solid curve). The first 10 nights of data fit the sinusoid well (filled circles), but the last 5 nights show considerably more scatter (asterisks).

been observed in J0422+32, and could indicate that the global changes in accretion behaviour in these two black-hole candidates are similar.

4.3 Quiescent photometry

4.3.1 Observations

J0422+32 had dimmed to a magnitude of 22.39 ± 0.27 by September, 1994 (Zhao et al., 1994b), which was the faintest state reached so far, and so it seemed likely that J0422+32 had reached quiescence. Therefore, I-band photometry was obtained to study the secondary star light curve with little disc contamination, on the NOT in 1994 (November 8) and on the JKT in 1994 (December 22–29). Four of the eight nights allotted on the JKT were lost due to cloud, but the conditions were good on the remaining nights (1–1.5 arcsec seeing). The Tek 4 CCD was used, giving a 5.6 x 5.6 arcmin^2 field of view and a scale of 0.33 arcsec pixel^{-1}. The KPNO broad band V, R and I filters were used and a total of 43 frames were taken. Exposure times were varied between 15 and 30 min, depending on the seeing conditions. The magnitude of J0422+32 was measured relative to two nearby reference stars using the aperture photometry routine APPHOT within IRAF. These were then inter-compared to check the consistency of the photometry (r.m.s. scatter = 0.007). The standard field PG0231 (Landolt 1992) was used to calibrate the V, R and I magnitudes.

The B,V,R and I colours were obtained on the second night, during photometric conditions, as well as B,V,R and I colours for the standard star PG 0231 (Landolt, 1992). These were taken consecutively to reduce systematic errors due to changes in the seeing or weather conditions, the airmasses were also similar, which reduces the extinction correction needed. The extinction corrections were obtained from the ING La Palma documentation, which was calculated using the method outline by Hayes and Latham (1975), assuming an aerosol-free atmosphere.

Unfortunately, J0422+32 was so faint in the B-band that an exposure time of 900s was not sufficient to render it visible on the CCD image taken, and so only V, R and I magnitudes could be calculated. The magnitudes obtained are shown in table 4.2. The reddening to

Table 4.2: Calibrated and dereddened magnitudes for J0422+32 ($E_{B-V} = 0.3$).

	V	R	I
Calibrated	22.24 ± 0.14	20.97 ± 0.1	20.22 ± 0.07
Dereddened	21.31 ± 0.14	20.27 ± 0.1	19.77 ± 0.07

J0422+32 was taken to be E(B−V) = 0.3 (from Bonnet-Bidaud and Mouchet, 1995) and the interstellar extinction curve of Savage and Mathis (1979) was used to obtain the dereddened magnitudes.

Thus the dereddened colours are: V−R = 1.04 ± 0.24; R−I = 0.50 ± 0.17. The dereddened V−R colour implies a spectral type in the range K2–M0 (Popper, 1980), using the 1-σ error and the fact that the secondary is a main sequence star (the 5.1 hour period gives a main sequence density if the secondary fills its Roche lobe). Various blackbodies were convolved with the filters used to obtain the temperature and spectral type corresponding to the R−I colour. This gave a spectral type K1 which seems somewhat early, but could indicate that the reddening used here is inaccurate. Also the error associated with the R−I colour is quite large and gives spectral types in the range G5 – K7. Thus, the values are just consistent with the range of possible spectral types (K7V–M0V) published by Chevalier and Ilovaisky (1995), which were estimated from V and I-band photometry taken during varying brightness states of J0422+32, and the value of M2V from Casares et al. (1995b). All of the above magnitudes assume the object is quiescent and therefore that there is no contribution to the flux from the accretion disc, which has a much higher colour-temperature than the secondary star. This will cause the V−R and R−I colours to become smaller and this will cause the estimate of the spectral type to be too early. The above estimates for the spectral type should therefore be treated as lower limits.

4.3.2 Quiescent variability

The Lomb-Scargle periodogram of the entire dataset is shown in figure 4.11. None of the peaks on the periodogram are significant relative to the noise (the 90% confidence level is

6.904 on the power axis), although there is a peak near the 16.18 hour period as well as a broad set of peaks near 2.55 hours. To try to further clarify the nature of the variability of the data, a set of periodograms for each separate night was calculated in figure 4.12. There is a peak in the periodogram of each night's data near the 2.55 hour period (9.41 cycles/day – labelled by a dashed line), but the position of this peak varies from night to night. The reason for this variation is not clear but could be due to some long period modulation, or the limited size of this dataset.

The folded light curves for the NOT and JKT datasets are shown in the top and bottom panels of figure 4.13, respectively. The mean magnitudes of J0422+32 in December were found to be: V = 22.24 ± 0.14, R = 20.97 ± 0.1, I = 20.22 ± 0.07. These values suggest that J0422+32 has reached its quiescent state, and so the I-band light curve should be dominated by the secondary. The data were folded on the photometric period (5.0944 ± 0.0017 hours) of Chevalier and Ilovaisky (1994b), and the T_0 (2,449,715.345 ± 0.004) from Casares et al. (1995b); then divided into 14 bins to give the above light curves. The errors were calculated from the aperture photometry routine APPHOT, taking into account the number of points in each bin.

The contribution of the secondary star to the flux of J0422+32 was found to be 52 percent, in the $\lambda\lambda$6700–8000 region, by Casares et al. (1995b) from fitting M2V templates to the spectra of the source. As the secondary contribution rises toward the red, this can be used as a lower limit for the secondary star contribution in the I-band photometry (λ 8300). The light curves should therefore be dominated by the ellipsoidal modulation of the secondary, and so ellipsoidal fits were performed on the folded light curves using the above estimate for the secondary flux. The other parameters used were: T_{eff} = 3600 K, limb darkening coefficient u = 0.61, gravity darkening β = 0.08, λ = 8300 Å (see Tjemkes et al., 1986). As the mass ratio is uncertain the fits were performed over a range of values, with the following results: q = 1, i = 49° ± 6 °; q = 5, i = 43° ± 6 °; q = 10, i = 41° ± 6 °. These results are combined from both the NOT and JKT datasets. The fit for a mass ratio of 5 is plotted on top of the 2 sets of folded data in figure 4.13, and can be seen to be consistent with the variation seen

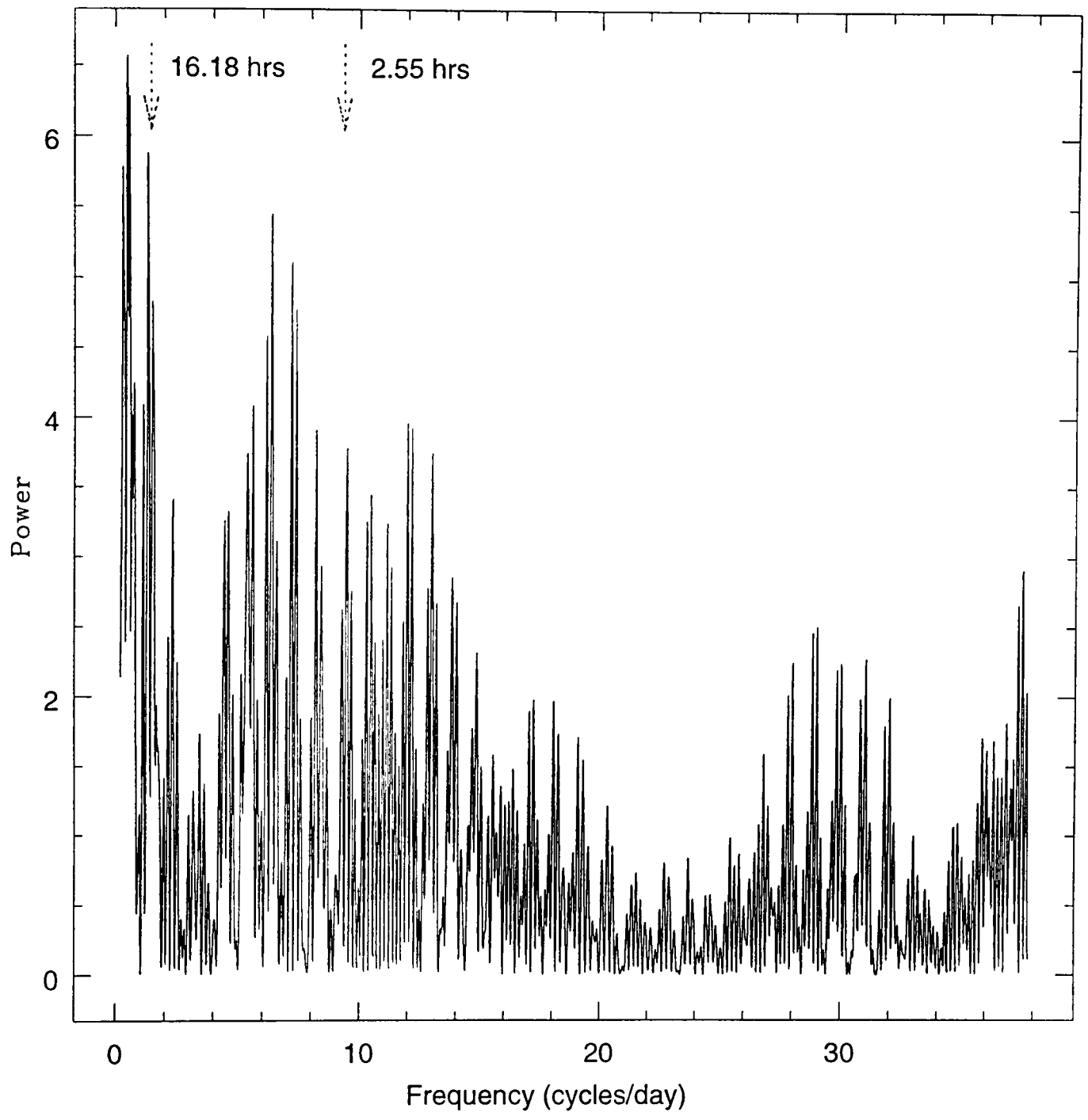


Figure 4.11: Lomb-Scargle periodogram of the quiescent photometry of J0422+32 (22–29/12/94). No peaks stand out significantly from the noise, the 90% confidence level occurs at a power level of 6.904. The orbital period (5.1 hours) found by Kato et al. (1992) is labelled by a dashed line, along with the 16.18 hour period which was found in the near quiescent photometry of section 4.3. One peak does appear near the longer period, but with a low significance and the shorter period has no obvious single peak in this periodogram.

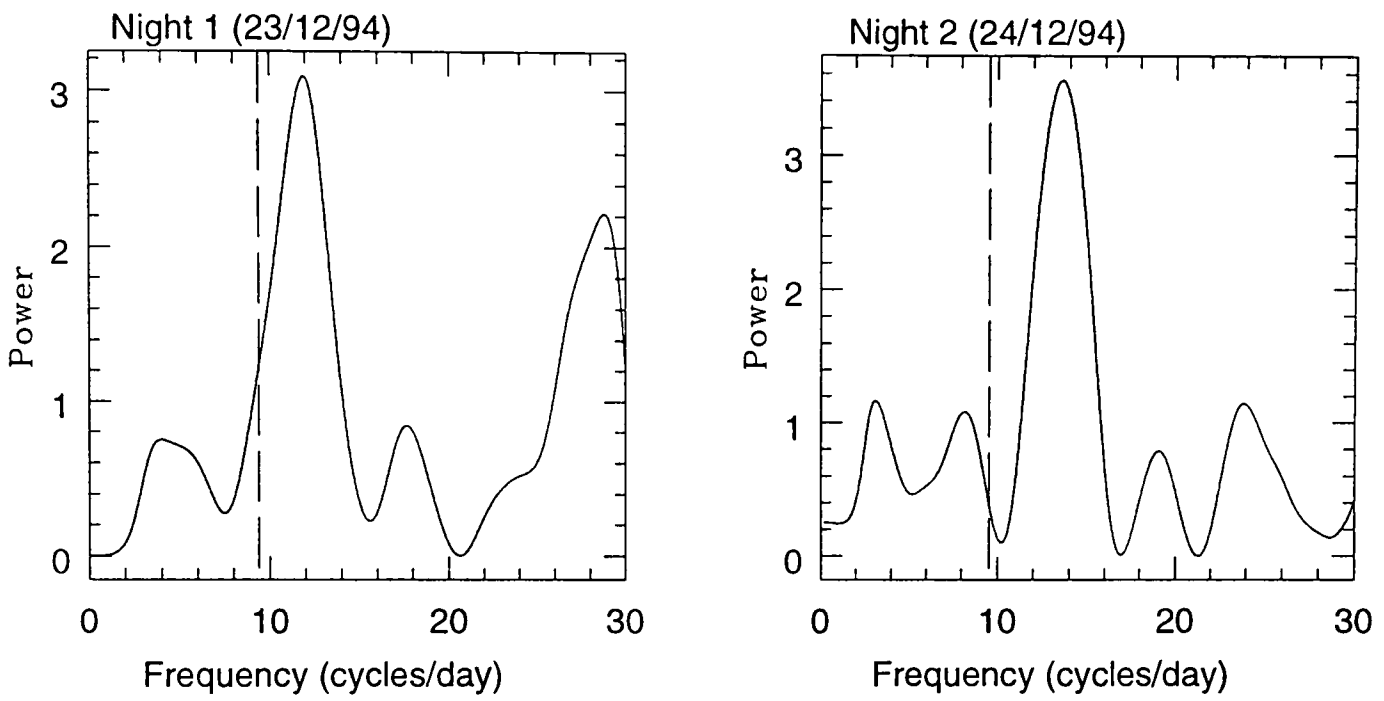


Figure 4.12: Separate Lomb-Scargle periodograms for the 4 nights (23,24,28,29) of the December 1994 run. Each shows a clear peak at $\sim$ 10 cycles/day (2.4 hours), near 2.55 hours (labelled with a dashed line). Unfortunately, each periodogram gives a different estimate of the frequency, making the overall result difficult to interpret.

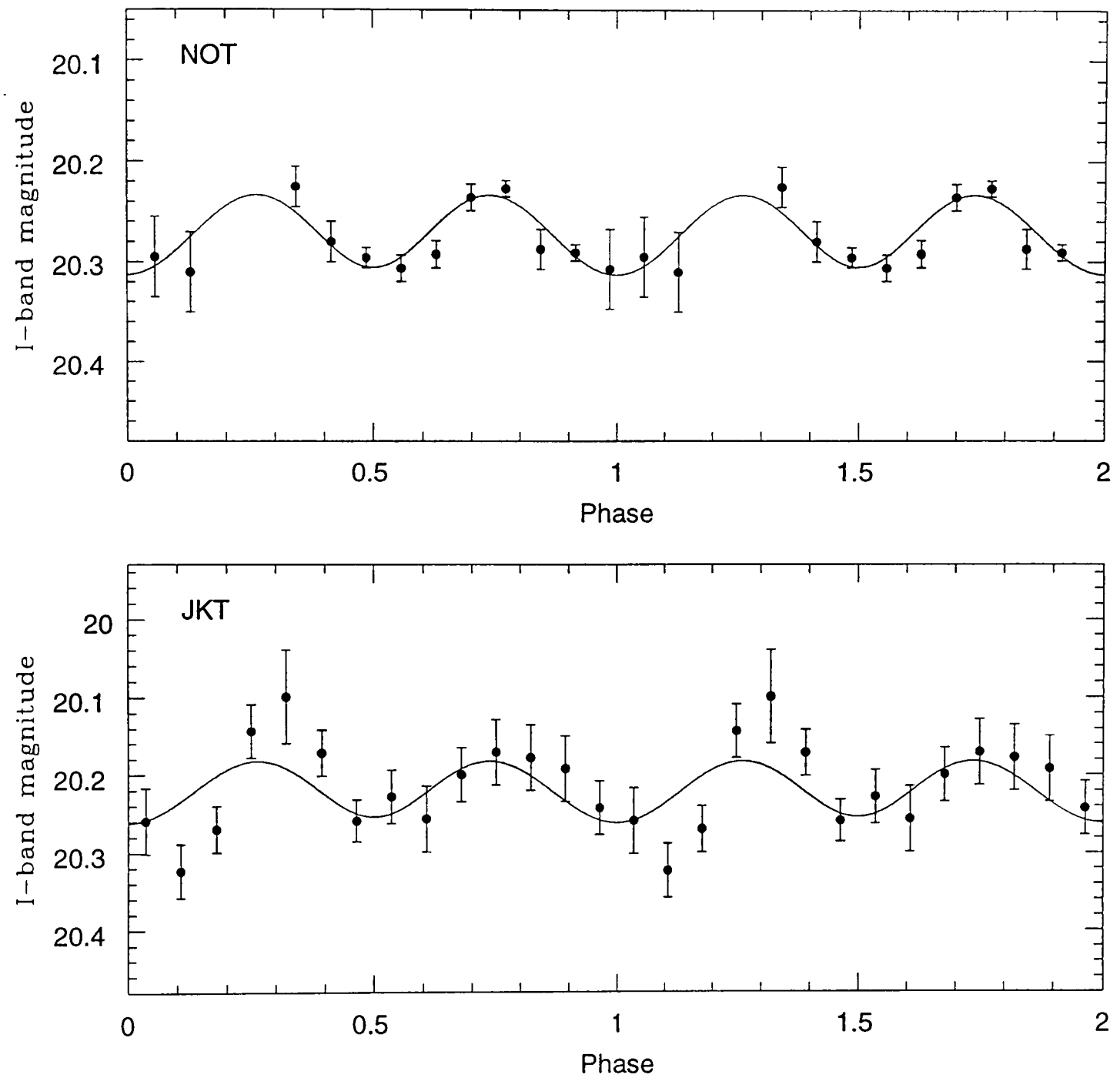


Figure 4.13: The I-band folded light curves for the NOT and JKT datasets, and the ellipsoidal fit at q = 5, i = 43°, with a secondary contribution of 52%. This inclination combined with the mass function of 1.2 $M_\odot$ (Casares et al., 1995b), implies a compact mass in the range M_x=4.8–5.3 $M_\odot$.

in the data [1]. The inclination obtained appears to be low, and is not consistent with the orbital dips seen by Kato et al. (1995), which would require an inclination $> 70°$. If the mass ratio is high in J0422+32 (Filippenko et al., 1995 estimated a mass ratio of ~ 10), then the inclinations above, combined with the mass function of 1.2 $M_\odot$ (Casares et al., 1995b), give a compact mass in the range $M_x = 4.8 - 5.3$. This is above the maximum mass for a rotating neutron star (≈ 3.8 $M_\odot$), and is much larger than the neutron star masses measured to date, which all lie near 1.5 $M_\odot$. Therefore J0422+32 is a strong black-hole candidate.

4.3.3 Intensity Dependent Changes in the Modulation of J0422+32

The periodograms of the datasets from October, 1993 (section 4.1.4) and December, 1994 (section 4.3.2), do not appear to be consistent with each other. But the fact that they were obtained during different brightness states of J0422+32 can be used to explain the differences seen. In particular, it is possible to constrain the size of the variability of the 5.1 hour and 16.18 hour modulations, and thus to deduce some information about their possible origins.

What is required is an estimate of the size of modulation of the 2.55/5.1 hour and 16.18 hour periods, from both the 1993 and 1994 datasets. The 2.55 hour modulation has been seen in quiescence (Casares et al., 1995b), from both photometric and spectroscopic data, and so would seem to be due to the secondary, which implies that the orbital period is 5.1 hours. The size of the 2.55 hour modulation as the brightness of J0422+32 changes can be examined in the different brightness states of J0422+32, to verify that it consistent with the secondary.

A way of determining the amplitude from the periodogram is to ascertain the power - amplitude relationship, by using Monte-Carlo simulations to create artificial periodograms. To do this, the data sampling of the 1993 and 1994 datasets were taken, after which, noise and sinusoids of varying amplitudes were added. This enabled a curve of power versus amplitude to be computed separately for each dataset. It was found that altering the frequency of the sinusoid changed the power/amplitude curve, and so two separate curves for each dataset were

[1] The two datasets were not combined into the same light curve because they have different mean magnitudes. Different bins consisting of different numbers of points from each dataset would then have a systematic bias which could affect the final result.

computed at the 2.55 hour and 16.18 hour periods. These curves were used in the following analysis (see figure 4.14).

4.3.4 Analysis of the 2.55 (5.1) Hour Period

It is possible to verify that the 2.55 hour period is due to the secondary by using the two periodograms obtained for J0422+32. The 1994 data (taken in quiescence) give a power level of ≈ 3.5 at a period of 2.55 hours, which implies a semi amplitude of 0.036^m. Assuming that this is now the true quiescent level, the 1993 data were taken when J0422+32 was 1.36^m brighter (in the I-band) and so the 'veiling' due to this increased brightness would mean that the ellipsoidal modulation decreases to S.A = 0.008^m. This corresponds to a power level of 4.646 for the 1993 dataset, for this size of modulation. This is well below the 99% confidence level for the 1993 data and so one would not have expected to see the secondary modulation.

4.3.5 The Origin of the 16.18 hour Period

The above analysis can also be applied to the 16.18 hour periodicity, the source of which is currently unknown. Using the semi-amplitude (0.14^m) found in the, high state photometry, it is possible to calculate the amount of power expected in the quiescent periodogram. If the modulation were due to the secondary then in quiescence, where it is expected that the disc contribution is small, the light curve should vary with a semi-amplitude of 0.52^m. This corresponds to a power level of z = 18.49, which is well above the 99% confidence level for the 1994 dataset and so this can be ruled out with a non-detection probability of 10^{-4} (calculated from 10,000 Monte-Carlo simulations of the data).

The semi-amplitude from the 1993 data (0.14^m) can be used again to test if the source of the modulation could be due to the disc. To get an idea of the power expected in the periodogram in quiescence it is assumed that the disc now contributes a small amount of flux $\sim$ 30%. This means that, in quiescence, there should be a variation of $\approx 0.055^m$, which gives an expected power level of $z \approx 6$. This is below the 99% confidence level and so is not detectable with these data. Conversely, this means that because the modulation was not

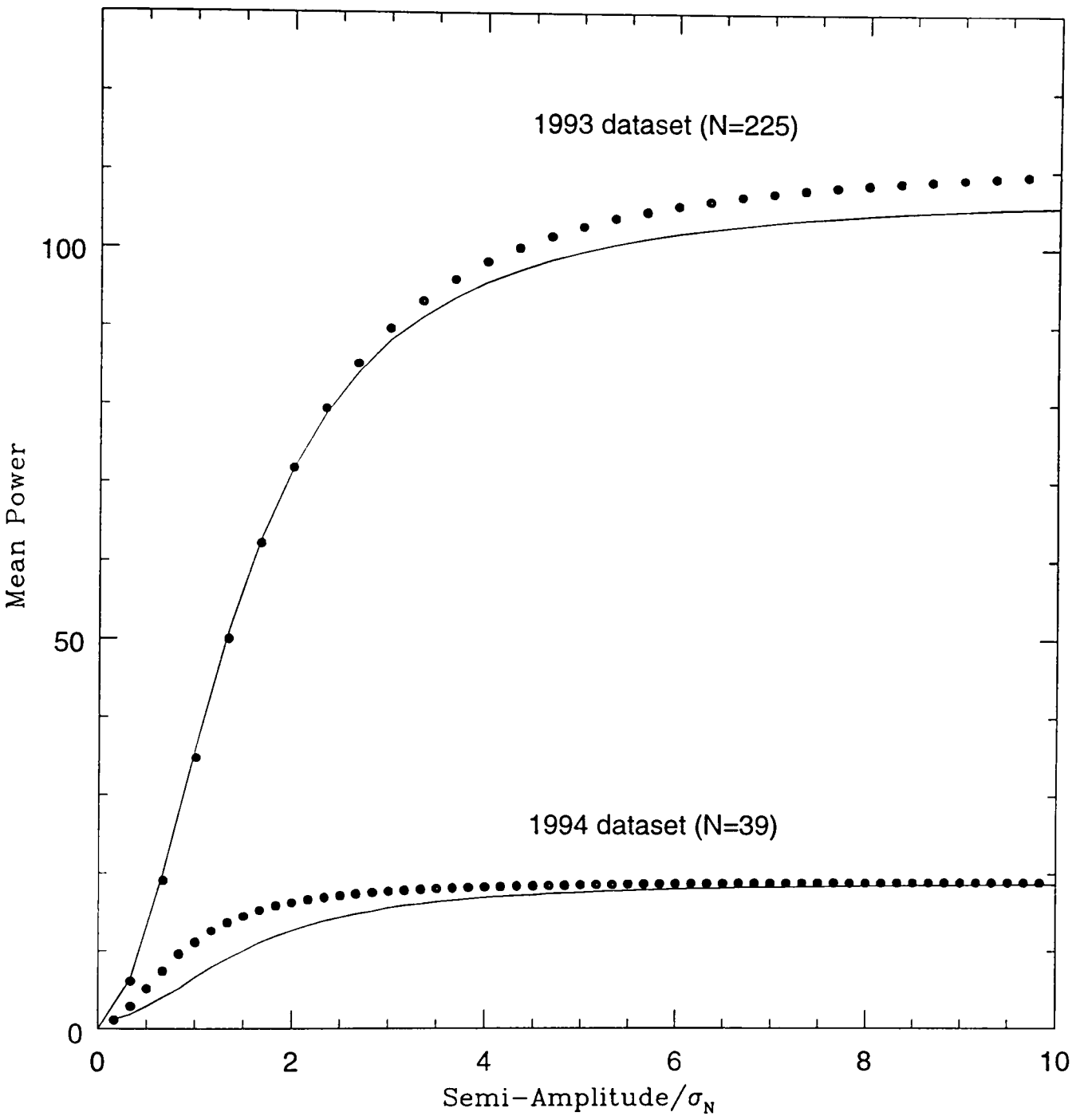


Figure 4.14: The power versus amplitude curves for the 1993 and 1994 datasets, at the periods 2.55 hours (dotted curve) and 16.18 hours (solid curve). The different periods can be seen to give slightly different curves, and so the correct period must be used in estimating the amplitude from the power spectrum. Note that the x-axis scale is semi-amplitude divided by the noise level and that the curves asymptotically approach N/2 when the amplitude is large.

detected in quiescence, the disc must contribute little to the flux.

4.4 Discussion

The above analysis implies that the 16.18 hour period is from a source other than the secondary, possibly the disc. Using the above arguments it is not possible to discover the origin of the modulation, all that can be said is that it increases in brightness with the disc. Unfortunately, no spectra were taken with the 1993 photometry and so no information on the radial velocity behaviour of this modulation is available. No other observations have yielded this result, although some observers have found that they needed to subtract long term trends from their light curves to remove low frequency power (Callanan et al., 1995).

The radial velocity curve of J0422+32 yielded a mass function of 1.2 $M_\odot$ (Casares et al., 1995b), which, combined with the inclination obtained from the ellipsoidal light curve fit (q = 5, i = 43°) gives a compact mass in the range M_x=4.8–5.3 $M_\odot$. This implies that J0422+32 is a strong black-hole candidate.

References

Allen, C.W., 1973, *Astrophysical Quantities*, Athlone Press, London.

Bonnet-Bidaud, J.M. and Mouchet, M., 1994, IAU Circ., 5980.

Bonnet-Bidaud, J.M. and Mouchet, M., 1995, Astr. Astrophys., **293**, L69.

Callanan P.J., Garcia M.R., McClintock J.E., Zhao P., Remillard R.A.,
Bailyn C.D., Orosz J.A., Harman B.A. and Paciesas W.S., Ap. J., 1995, **441**, 786.

Cameron R.A., Grove J.E., Kroeger R.A., Johnson W.N. and Kurfess J.D.,
1992, IAU Circ., 5587.

Casares J., Marsh T.R., Charles P.A., Martin A.C., Martin E.G., Wagner R.M., Harlaftis E.T. and Pavlenko E.P., 1995a, MNRAS, **274**, 565.

Casares J., Martin A.C., Charles P.A., Mart'in E.L., Rebolo R., Harlaftis E.T and Castro-Tirado A.J., 1995b, submitted.

Castro-Tirado, A.J. et al., 1992, IAU Circ., No. 5588.

Charles P.A. et al., Ap. J., 1980, **237**, 154.

Chevalier C. and Ilovaisky S.A., 1993a, Astr. Astrophys., **269**, 301.

Chevalier C. and Ilovaisky S.A., 1993b, IAU Circ., no. 5692.

Chevalier C. and Ilovaisky S.A., 1994a, IAU Circ., no. 5928.

Chevalier C. and Ilovaisky S.A., 1994b, IAU Circ., no. 5974.

Chevalier C. and Ilovaisky S.A., 1995, Astr. and Astrophys., **297**, 103.

Filippenko A.V. and Matheson T., 1993, IAU Circ., no. 5842.

Grabelsky D.A., Matz S.M, Purcell W.R., Ulmer M.P., Grove J.E., Johnson W.N., Kinzer R.L., Kurfess J.D. and Strickman M.S., 1995, Ap. J., **441**, 800.

Harlaftis E.T., Jones D.H.P., Charles P.A. and Martin A., 1994, in Proc. of the 4th Annual Astrophysics Conference in Maryland,*The Evolution of X-ray Binaries*, ed. American

Institute of Physics, pg. 91.

Harmon B.A., Wilson R.B., Fishman G.J., Meegan C.A., Paciesas W.S., Briggs M.S, Finger M.H., Kroeger R., and Grove, E., 1992, IAU Circ., 5584.

Haswell C.A., Robinson E.L., Horne K., Stiening R.F., Abbott T.M.C., 1993, Ap. J., **411**, 802.

Kato T., Mineshige S., Hirata R., 1995, PASJ, **47**, 31.

King, A.R., 1993, MNRAS, **260**, L5.

Landolt A.U., 1992, Astron. J., **104**, 340.

Lomb N. R., 1976, Ap. Space Sci., **39**, 447.

Markert T.H., Canizares C.R., Clark G.W., Lewin W.H.G., Schnopper H.W., and Sprott G.F, 1973, Ap. J. Lett., 184, L67.

Mineshige S., Hirose M., Osaki Y., 1992, Publ. Astron. Soc. Japan, 44, L15.

Mueller J., 1992, IAU Circ., 5597.

Paciesas, W.S. et al., 1992, IAU Circ., no. 5580.

Popper D.M., 1980, Ann. Rev. Astr. Ap., **18**, 115.

Rappaport S., Joss P.C. and McClintock J.E., 1976, Ap. J. Lett., **206**, L103.

Savage B.D. Mathis J.S., 1979, Ann. Rev. Astron. Astrophys., **17**, 73.

Scargle J.D., 1982, Ap. J., **263**, 835.

Shahbaz T., Sansom A.E., Naylor T. and Dotani T., 1994, MNRAS, **268**, 763.

Shrader C.R., Wagner R.M., Hjellming R.M., Han X.H., and Starrfield S.G., 1994, Ap. J., **434**, 698.

Sunyaev R., Churazov E., Gilfanov M., Novikov B., Goldwurm A., Paul J.,
Mandrou P. and Techine P., 1992, IAU Circ., 5593.

Tanaka Y., 1991, Proc. Workshop on Iron Line Diagnostics in X-ray Sources,

Lect. Notes. Phys., **385**, 98-110.

Tjemkes S.A., Zuiderwijk E.J. and Van Paradijs J., 1986, Astr. Astrophys., **154**, 77.

Vogt N., 1980, Astron. Astrophys., **88**, 66.

Wagner R.M., Kreidl T.J., Howell S.B. and Starrfield S.G., 1992,
Ap. J. Lett., **401**, L97.

Zhao P., Callanan P., Garcia M., and McClintock J., 1993, IAU Circ., No. 5901.

Zhao P., Callanan P., Garcia M., and McClintock J., 1994a, IAU Circ., No. 5929.

Zhao P., Callanan P., Garcia M., and McClintock J., 1994b, IAU Circ., No. 6072.

Chapter 5

Optical Photometry of Three Faint Soft X-ray Transients (V2107 Oph, BW Cir, KY TrA).

5.1 V2107 Oph

5.1.1 Summary

The detection of an orbital, ellipsoidal modulation at a period of 16.8 hours for V2107 Oph (Nova Ophiuchi 1977, H1705−25), from R-band photometry, is reported. From this period and the observed magnitude, constraints on the distance are calculated (2.0 kpc – 8.4 kpc), which is consistent with that obtained previously from X-ray observations (3 kpc). By comparison with other SXTs it is suggested that $q \geq 5$, and combining this with an ellipsoidal fit to the data gives an orbital inclination in the range 48° – 51°.

5.1.2 Introduction

The X-ray transient H1705−25 was discovered independently with the HEAO-1 scanning modulation collimator (MC) (Griffiths et al., 1977) and the Ariel 5 all sky monitor (ASM) (Kaluzienski and Holt, 1977). Subsequently, the coincident optical nova (N Oph 1977, V2107 Oph) was identified on plates taken at the Anglo-Australian Telescope and UK Schmidt Telescope (Longmore et al., 1977; Griffiths et al., 1978); it reached an estimated magnitude

of B = 16.5 ± 0.5. Examination of the blue and red POSS plates revealed a possible pre-nova star at the plate limit (i.e. at B ≈ 21 and R ≈ 21). The range of B magnitudes observed appears to be smaller than for a typical transient event which normally have $\Delta B \sim 6^m$. This could be due to the fact that the photometry was taken 2 days after the initial outburst was detected in the X-ray, or that the reddening reduced the observed magnitude when the object was brighter. An optical spectrum of V2107 Oph taken during outburst showed the presence of HeII 4686 Å emission but no detectable Hα (Griffiths et al., 1978).

The X-ray light curve of the outburst (Watson et al., 1978) showed a short rise time to maximum (~ 1.7 days), a secondary maximum that occurred ~ 2 days after the initial peak and an irregular decline over a period of months. This behaviour and the soft X-ray spectrum led to the classification of V2107 Oph as a soft X-ray transient (Griffiths et al., 1978). The X-ray data were analyzed for any possible time variability by period folding, which did not give any clear result. Subsequently, a hard X-ray component was found by Wilson and Rothschild (1983); the X-ray spectrum consists of a thin thermal bremsstrahlung component (kT = 1.8 ± 0.1 keV) and a hard power-law component (photon index 2.19 ± 0.06) . The X-ray properties of V2107 Oph are similar to those of other SXTs (A0620−00: Cooke et al., 1984; GS 2023+338: Sunyaev et al., 1991a; GS 1124−684: Sunyaev et al., 1991b) for which a dynamical mass estimate (McClintock & Remillard, 1986; Remillard et al., 1992; Casares & Charles, 1994) showed the compact star to be a strong black-hole candidate (for a recent X-ray review of these systems, see Tanaka & Lewin 1995). V2107 Oph is therefore to be considered a good black-hole candidate.

5.1.3 Observations

Six nights of photometry were obtained on the Danish 1.54m at ESO from May 1 – 6, 1994. Five of the nights had seeing mostly around 1 arcsecond; a large part of the remaining night was lost due to high cirrus. The Tektronix TK1024M CCD was used with a Gunn R-band filter (see the ESO operating manual) giving a 4.78 x 4.78 arcmin^2 field of view, with a scale of 0.28 arcseconds/pixel. A total of 87 CCD frames of V2107 Oph were obtained, and

the exposure times varied between 15 and 30 minutes depending on the seeing conditions. Standard fields were observed at the start and end of each run of images.

The brightness of V2107 Oph was measured with respect to two reference stars, which were themselves inter-compared using DAOPHOT (Stetson, 1987). Because of the crowded nature of the field and the faintness of V2107 Oph in quiescence, a new finding chart is presented in Fig. 5.1. Calibration was performed using two Landolt standard fields (Landolt, 1992), Markarian A (3 stars) and PG0918 (1 star), during two separate photometric nights that had 1 arcsecond seeing or better. Observation of these standard fields from night to night showed that the photometric calibration was stable to within 0.005 magnitudes. These standards were observed at different airmasses, and the magnitude of V2107 Oph was calibrated, to allow for the extinction, close to the airmass of one of the standard observations. The reference stars are labelled 'R' and 'C' on Fig. 5.1 and have R magnitudes of 20.13 $\pm$ 0.05 and 20.49 $\pm$ 0.05 respectively. The light curves of V2107 Oph and reference star C are shown in Fig. 5.2.

5.1.4 Orbital Period Search

Before trying to detect a modulation in the data, a search was made for possible systematic errors caused by the conditions during the run. To do this a histogram was made of the magnitude differences between the two reference stars. This had a basically Gaussian shape as expected, but with a tail to one side consisting of a few points. These data points correspond to images taken during periods of cirrus or very poor seeing (> 2 arcseconds). These outliers were removed by performing a Gaussian fit to the histogram, all points deviating by more than 3σ were excluded from the analysis, as they were systematically biased due to poor weather conditions. A total of 9 images were removed by this procedure.

A Lomb-Scargle periodogram (Lomb, 1976; Scargle, 1982) of these data is shown in Fig. 5.3, and has its highest peak at a period of 8.4 $\pm$ 0.05 hours. The 99% confidence level was calculated from a Monte-Carlo simulation, which calculates the maximum power level in 10,000 sets of Gaussian noise with mean and variance equal to that of V2107 Oph. The

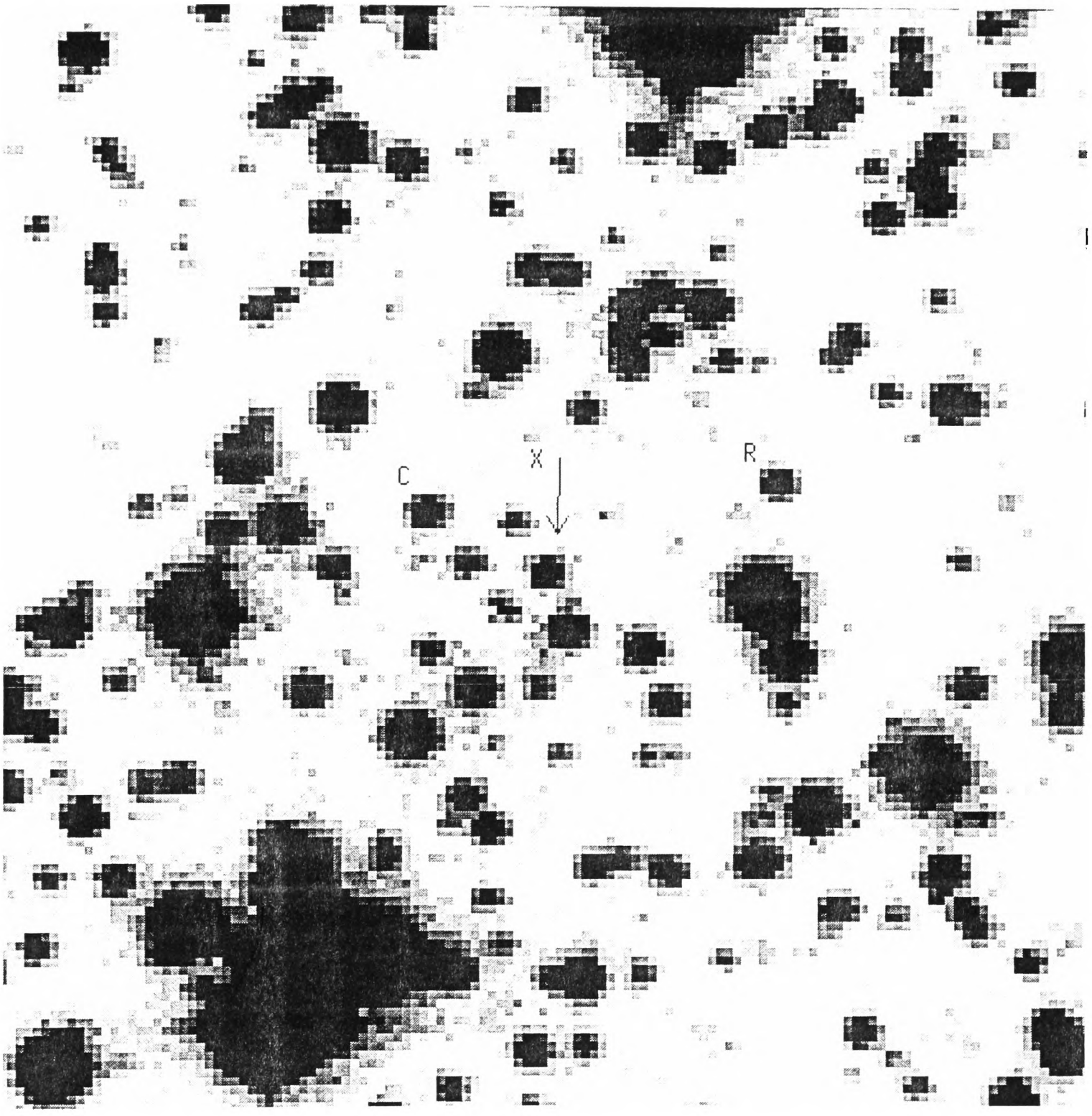


Figure 5.1: Improved finding chart (taken 1/5/94) for V2107 Oph (H1705−25), arrowed and labelled X. North is up, and east is to the left. Also shown are the two reference stars 'R' and 'C' used in the data reduction. The field is 1.19 x 1.19 arcmin^2.

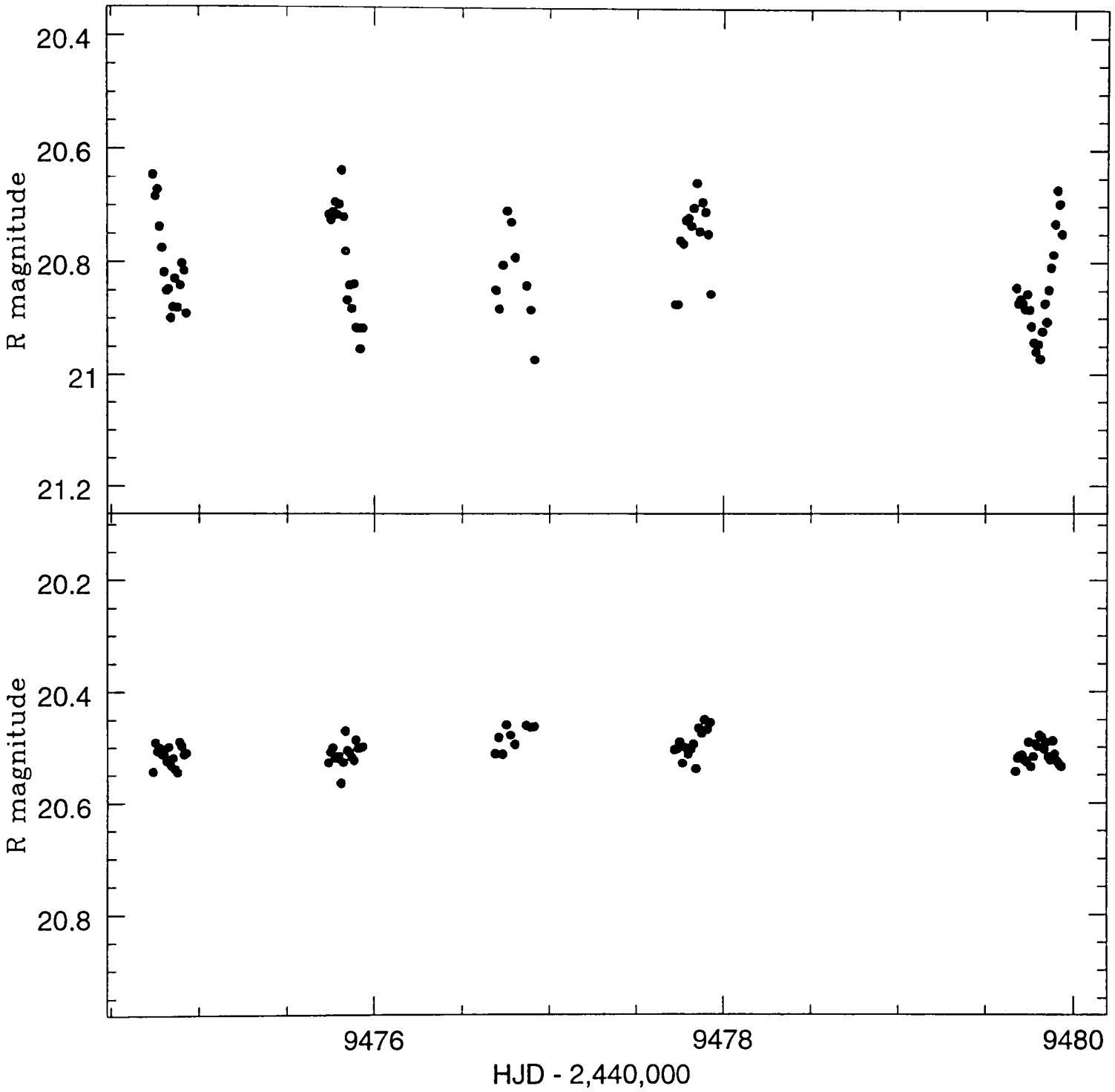


Figure 5.2: Top – R-band photometry of V2107 Oph in May, 1994 at ESO (6 nights, 78 points), with poor quality images removed as described in the text; bottom – reference star C relative to reference star R ($\sigma = 0.04$).

power level at which only 1% of the artificial datasets showed a peak due to noise (calculated between the Nyquist frequency and the resolution limit $1/2T$, where T is the total length of observations) gives us the 99% confidence limit (at a power level of 9.707), as shown in Fig. 5.3.

The periodogram is not as clear as one would hope, with significant aliasing, and so to reduce this problem the $1/\chi^2$ statistic periodogram was calculated for the data. The algorithm for this technique was outlined in Chapter 2. It involves creating artificial datasets with a sinusoid, noise and the same data sampling as the observed data. The closeness of fit between the real and artificial periodograms is then calculated over a range of frequencies giving $1/\chi^2$(f). To use this method it is necessary to know the noise level of the data, and also to estimate the amplitude of the modulation of the data. The mean of the errors on the photometry of V2107 Oph was found to be 0.039^m, and so then all that is required is some way of estimating the amplitude of the variation in the data, without assuming a known period.

To estimate the amplitude, the method described in Chapter 2 is used. The variance due to the signal is equal to the variance of the data (0.0079 ± 0.0008) minus the noise variance (0.039^2), which gives 0.0063 ± 0.0007 (i.e. the noise is summed in quadrature). The semi-amplitude of the modulation of the data is then $\sqrt{2 \times 0.0063} = 0.11 \pm 0.01^m$ and the full amplitude is twice this value (0.22 ± 0.02^m). This is very close to the value found from the ellipsoidal fit to the data (0.21^m) see section 5.1.5. The amplitude of the signal modulation and noise variance is then put into the DFT fitting routine and the best fit periodogram found. The resulting plot is shown in figure 5.4. The first periodogram is of the original dataset and is identical to figure 5.3, below this is the best fit periodogram with the largest value of $1/\chi^2$, and at the bottom is the periodogram produced by the program at a period of 6.18 hours (the second highest peak in the periodogram). The period at 8.33 hours gives the best fit to the observed periodogram, and it can be seen that the artificial periodogram reproduces all of the gross features of the observed one, whereas the 6.18 hour periodogram is significantly different. This is quantified in figure 5.5 (top) which shows $1/\chi^2$ against frequency. There is

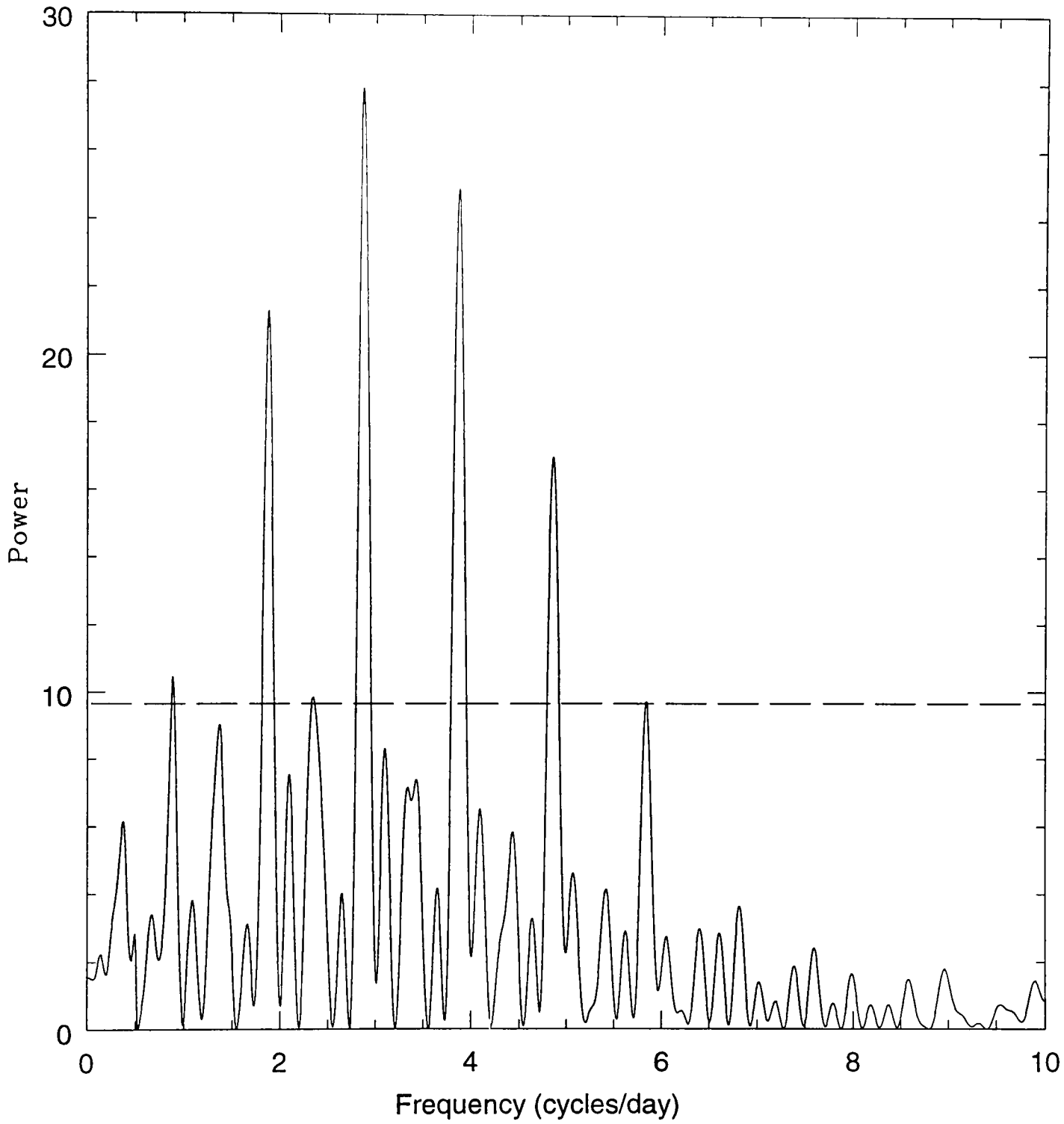


Figure 5.3: Lomb-Scargle periodogram of the dataset, after the removal of the outliers. A Monte-Carlo simulation with 10,000 noise sets provides the 99% confidence level as shown.

a large peak at 8.33 hours (2.88 cycles/day) and little alias structure – the 6.18 hour period is shown to be a poor fit. The most likely period is 8.33 ± 0.03 (this has a smaller error than the DFT result as the whole periodogram is used to obtain this estimate of the frequency). Also in figure 5.5 (bottom), are the residuals of the 8.33 hour periodogram subtracted from the dataset periodogram. The residuals have no significant features left – they lie well below the 99% confidence level (shown as a dashed line). This indicates that all of the signal has been subtracted, and all that is left is the noise. Some of the residuals for 6.18 hour period lie outside these limits, and so indicate the presence of something other than noise with 99% probability, i.e. the period cannot completely explain the variation of the data.

The suppression of the alias structure in the $1/\chi^2$ periodogram and the similarity of the artificial 8.33 hour periodogram to the observed periodogram indicates that it is the most likely period, and that the nearby side-lobe structure is an artefact of the data sampling. The power level of the 8.33 hour peak in the Scargle periodogram is 99.98% significant, and the symmetrical alias pattern also suggests that it is the correct periodicity. For this to be the actual orbital period would require the modulation to be due to X-ray heating of the secondary or a hot spot/hump in the accretion disc. These are very unlikely to occur in SXTs in quiescence, as the observed X-ray flux (Verbunt et al., 1994) is at least four orders of magnitude below the outburst level, and the mass transfer rates are much lower than in conventional LMXBs. It is assumed, therefore, by analogy with the results obtained for other SXTs (McClintock & Remillard, 1986, 1990; Chevalier et al., 1989; Shahbaz et al., 1993, 1994a) that the light curve is of the double-waved ellipsoidal type; hence the implied orbital period of V2107 Oph is 16.8 ± 0.1 hours. This interpretation is further supported by the difference in the two minima (see next section), and the Bayesian ellipsoidal periodogram which gave the same orbital period.

5.1.5 Ellipsoidal Light Curve and System Parameters

Ellipsoidal modelling is, in principle, a very powerful diagnostic for determining the inclination in non-eclipsing binaries. This is demonstrated by Shahbaz et al. (1993) in their

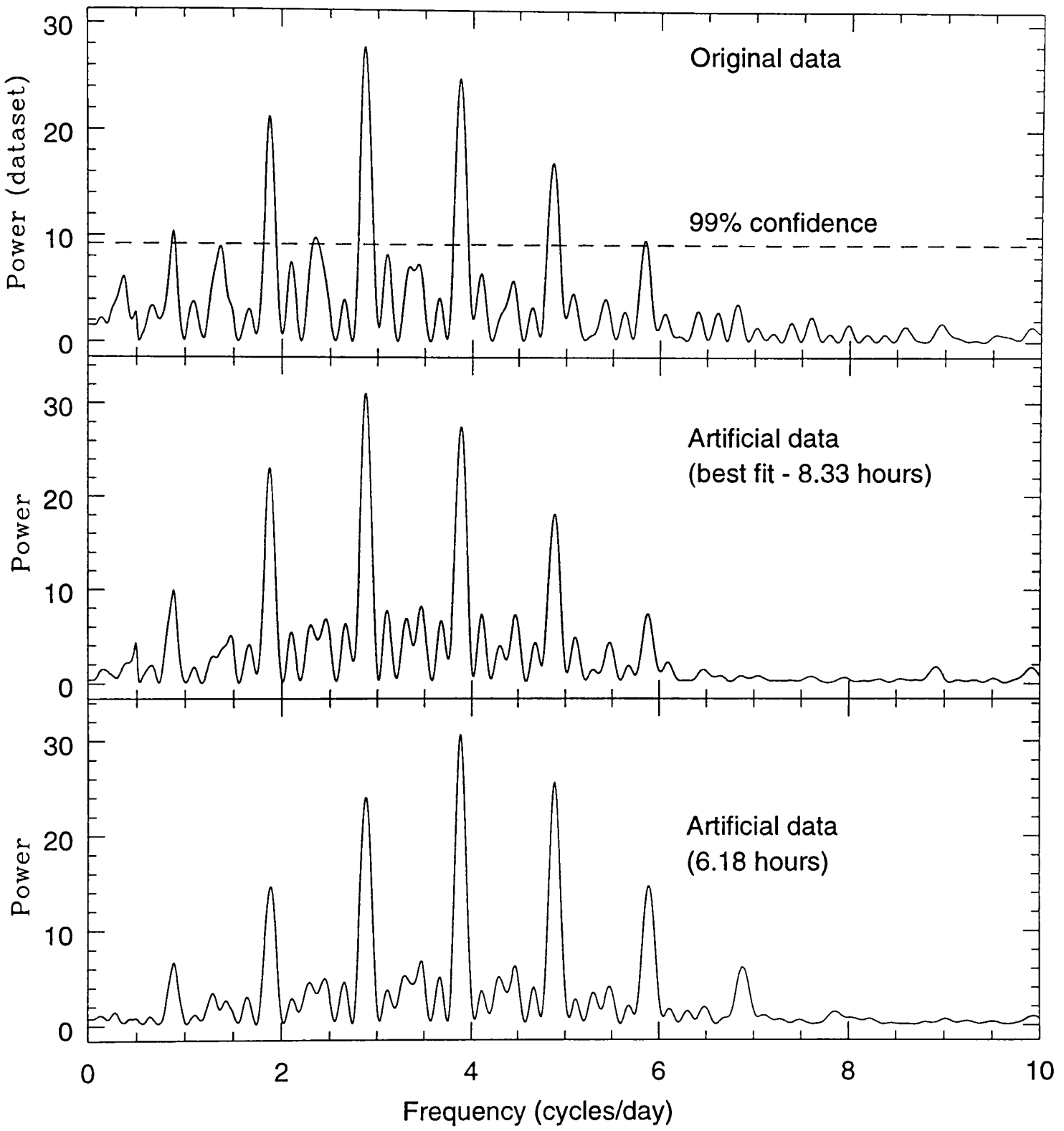


Figure 5.4: Top – the original periodogram of the data; middle – the best fit periodogram; bottom – the artificial periodogram at the 6.18 hour period. The original and best fit periodograms are almost identical and show that all of the alias structure is reproducible, whereas the 6.18 hour periodogram is markedly different.

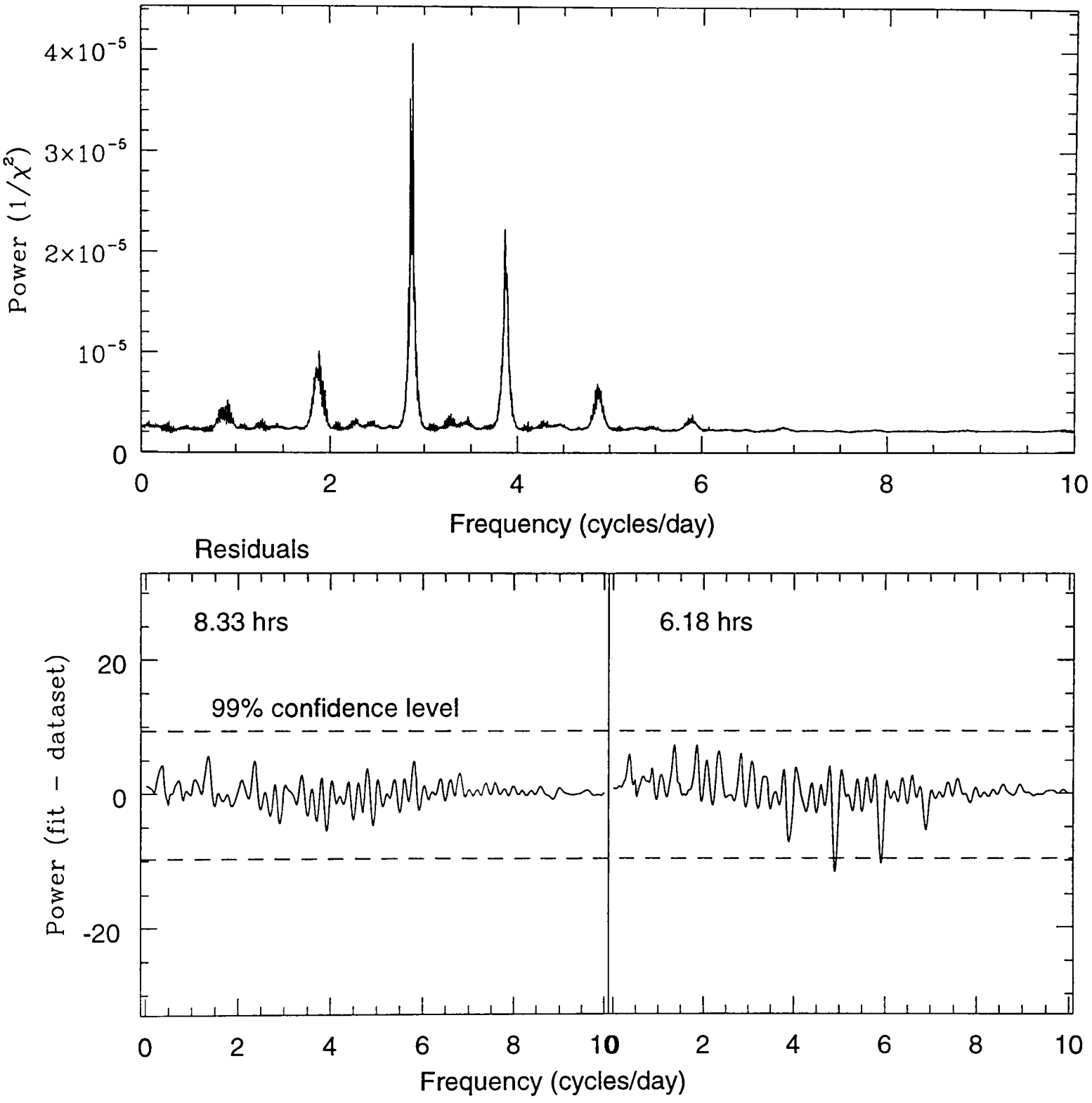


Figure 5.5: Top – the $1/\chi^2$ statistic versus frequency. The 8.33 hour peak is most prominent and the alias pattern is greatly reduced. Bottom – the residuals from the fits at the two most probable frequencies. The 8.33 hour residuals all lie within the 99% confidence range, whereas the 6.18 hour exceeds this power level twice. This indicates that there are systematic effects left in the periodogram, after the 6.18 hour period has been subtracted, which are not due to noise.

studies of the SXT Cen X-4. For a given inclination angle i, the amplitude of this modulation changes little for $q > 5$ and therefore depends entirely on i, allowing the inclination of V2107 Oph to be obtained directly, assuming only that it is a high mass ratio system. However, using the calibrated photometric magnitudes measured for V2107 Oph, combined with estimates of interstellar absorption, it is possible to obtain a range of possible secondary masses and find a lower limit for the mass ratio, and further constrain the inclination.

The assumption underlying this type of analysis is that all brightness modulation is caused by the tidal and rotational distortion of the secondary star, i.e. that the contribution of the disk to the system brightness is negligible. Recently, Haswell (1995) has presented evidence that small brightness variations of an accretion disk, at a period slightly larger than the orbital period (i.e., the analogue of superhumps seen in superoutbursts of SU UMa type cataclysmic variables, see Ichikawa et al., 1993), can affect the detailed shape of the double-waved optical light curves of quiescent SXTs. They can produce unequal maxima, and reversals of the relative depths of the minima. According to Haswell the importance of these superhump variations decreases with the time elapsed since the SXT outburst. Since the outburst of V2107 Oph occurred almost two decades ago, and there is no evidence for a non-ellipsoidal distortion of the light curve (see Fig. 5.6) the perturbing effects of an accretion disk are likely to be small; however, the possibility of a small systematic effect on the results cannot be fully excluded. Future observations in the infra-red should be undertaken in order to reveal such an effect.

An ellipsoidal model fit to the data, sectioned into 20 phase bins, was calculated. The ellipsoidal effect is due to the changing aspect of the distorted secondary with respect to the observer, combined with limb and gravitational darkening. These effects cause a double modulation per orbital cycle with minima occurring at phases 0 and 0.5. The phase 0.5 minimum (when the inner face of the secondary is towards us) is the deeper due to the larger gravity darkening near the inner Lagrangian point. The model fit was performed using the following parameters: $T_{eff} = 4500$ K, limb darkening coefficient $u = 0.716$, gravity darkening exponent $\beta = 0.08$, and effective R-band wavelength $\lambda = 6460$ Å (see Tjemkes et al., 1986),

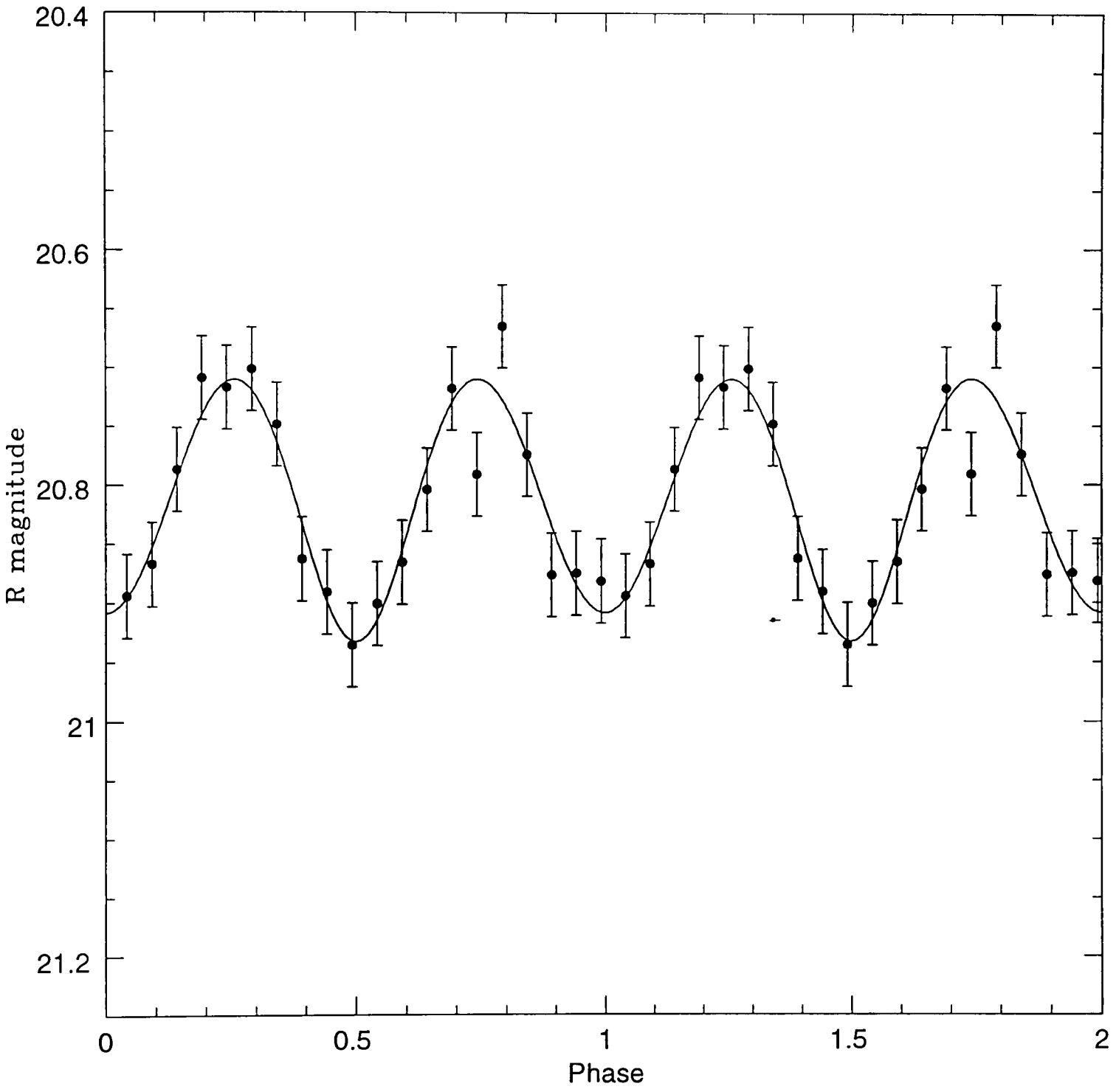


Figure 5.6: Folded light curve of V2107 Oph on the 16.8 hour, ellipsoidal period, with 20 phase bins. A variation of 0.2 magnitudes is evident and a possible difference in the minima of $\sim 0.03^m$. The ellipsoidal fit to the light curve was calculated using $q = 5$ and i $= 51°$, χ^2_ν $= 0.83$.

leaving the mass ratio and inclination as free parameters. By comparison with other SXTs it is possible to assume $q \geq 5$, and therefore obtain an inclination in the range $48° - 51°$ making no assumptions about the nature of the compact object. The ellipsoidal fit for $q = 5$ and $i = 51°$ is shown in Fig. 5.6.

Using the results from the photometry some of the system parameters of V2107 Oph can also be estimated. From the orbital period of 16.8 ± 0.1 hours the mean density of the secondary can be calculated, assuming that it fills its Roche lobe. This gives $\bar{\rho} = 0.35$ g cm^{-3}, which is an order of magnitude less than for a typical main sequence star, implying that the secondary is evolved, probably a sub-giant (see King, 1993). The stripped giant model (see Chapter 1) is used in order to constrain the mass of the secondary in V2107 Oph, and to calculate its lowest possible mass assuming the giant envelope has been completely stripped, using the orbital period of 16.8 hours. This gives a lower limit of $M_2 = 0.15\ M_\odot$. Also, by analogy with other SXTs which are known to have low mass secondaries (V404 Cyg: $M_2 = 0.6\ M_\odot$, Shahbaz et al., 1994a; A0620−00: $M_2 = 0.21\ M_\odot \sin^3 i$, Marsh et al., 1994, which corresponds to $0.2 - 0.7\ M_\odot$, Shahbaz et al. 1994b) it is possible to set a conservative upper limit of $1.0\ M_\odot$.

The distance can be estimated from these two limits by using the R-band magnitude of 20.79. The V−R colours for the spectral types K0 and M0 were used (Popper, 1980) in order to obtain the bolometric magnitude, which yields the upper and lower limits, respectively; after which the bolometric correction and reddening are applied. The reddening value $A_v \approx 1.4$ (from Griffiths et al., 1978), was inferred from low energy X-ray absorption. The resulting distance range , 2.0 kpc to 8.4 kpc, concurs with the value ~ 3 kpc estimated by Griffiths et al. (1978), which assumed that the peak X-ray luminosity of V2107 Oph was of the same order as A0620−00.

5.1.6 Conclusions

The similarity of the X-ray behaviour of this system during outburst to the other soft X-ray transient black-hole candidates and absence of any type I X-ray bursts (caused by thermonuclear burning on the surface of a neutron star) suggests that V2107 Oph is a promising black-hole candidate. Assuming a black-hole primary with minimum mass $M_1 = 3\ M_\odot$ and using the upper limit for the secondary mass, the mass ratio obtained is $q \geq 3$. Unfortunately it is not possible to derive a mass function for the system as yet, as there is no currently available radial velocity curve for the secondary. This will be the subject of future studies.

5.2 BW Cir (Nova Cir 1987, GS 1354–64)

5.2.1 Introduction

BW Cir was first discovered by the X-ray satellite *Ginga* (Makino and *Ginga* Team, 1987), and may be related to the outburst source MX 1353–64 (Markert et al., 1977). It is also consistent with the position of Cen X-2, which went into outburst in 1966 (Francey et al., 1971; Cominsky et al., 1978), and was the first X-ray source to be described as a transient (see Chodil et al., 1968). The optical counterpart of the X-ray source was later identified by Pedersen et al. (1987) on a 60 minute blue plate as BW Cir, and on three subsequent nights they observed tenth magnitude variations on a timescale of 8 hours. Subsequently, Ilovaisky et al. (1987) reported a possible period around $\sim$ 46 hours, with an amplitude of 0.3–0.4 magnitudes in V, at a mean V-band magnitude of 17.5. They also saw a secular decrease of 0.04 magnitudes per day, which, along with the brightness of the object, indicates that BW Cir was declining back towards quiescence after the outburst.

A finding chart for BW Cir was published by Kitamoto et al. (1990), along with the overall X-ray light curve. The maximum flux was 2.9 $\times 10^{-9}$ erg s^{-1} cm^{-2} in the 1–10 keV band, with a decay on a timescale of 66 days. They also took a 10 minute low dispersion (13 Å) spectrum which showed strong HeII λ4686 emission, indicative of the presence of ionizing radiation, which is typical of LMXBs and SXT systems in outburst.

During observations taken at the AAT in 1993 BW Cir was found to have a magnitude of ~ 21, indicating that BW Cir had declined back into quiescence. Thus, it was possible to observe the object with little or no disc contamination and obtain an ellipsoidal light curve from the variations in light output of the secondary.

5.2.2 Observations

The photometry of BW Cir were taken during the same observing run as V2107 Oph, and so all the details of the observations are essentially the same. Unfortunately, part of the first night was lost due to technical difficulties and so only 5 nights of photometry were obtained giving a total of 55 CCD frames. The exposure time was again varied between 15 and 30 minutes (dependent on the seeing), to acquire photometry with enough signal-to-noise to give an accuracy of a few percent.

The finding chart of Kitamoto et al. (1990) was taken during outburst with an exposure time of 5 minutes, when the V magnitude of BW Cir was ≈ 16.9, and so a new, deeper (1800s exposure) finding chart for BW Cir in quiescence is shown in figure 5.7. BW Cir is labelled 'X' and the two reference stars (the reference and check stars), are labelled 'R' and 'C'. The field around BW Cir is very crowded – several of the point spread functions of nearby stars overlap with BW Cir, and so it was necessary to perform the photometry using DAOPHOT (Stetson, 1987) to take account of this crowding. The two reference stars were chosen from several stars because they showed the least photometric variation ($\sigma = 0.0068$), and both are in clearer areas of the field.

Calibration of the photometry was performed using the standard star PG0918 (Landolt, 1992), during sub-arcsecond photometric conditions. The mean magnitude of BW Cir, from this calibration is 20.34 ± 0.03. The two reference stars 'R' and 'C' have magnitudes 19.35 $\pm$ 0.02 and 19.36 ± 0.02, respectively. The R-band light curve is shown in figure 5.8. The photometry of BW Cir is plotted above that of reference star R (r.m.s. scatter $\sigma = 0.022$). The variation in BW Cir is clearly much larger than this ($\sigma = 0.127$).

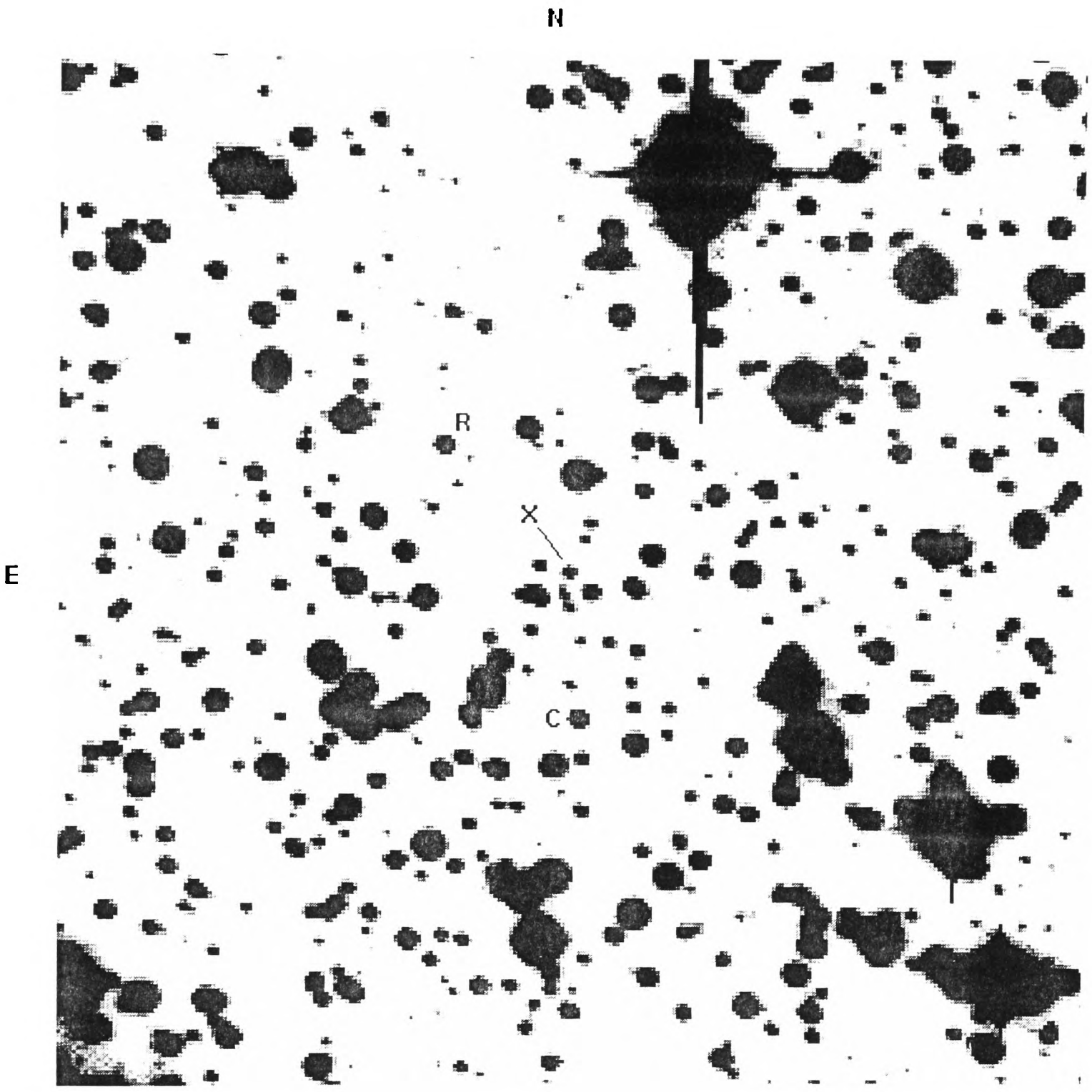


Figure 5.7: Improved finding chart (taken 1994, May 4th) for BW Cir, labelled X. North is up, and east is to the left. Also shown are the two reference stars 'R' and 'C' used in the data reduction. The total field of view is 1.19 x 1.19 arcmin^2.

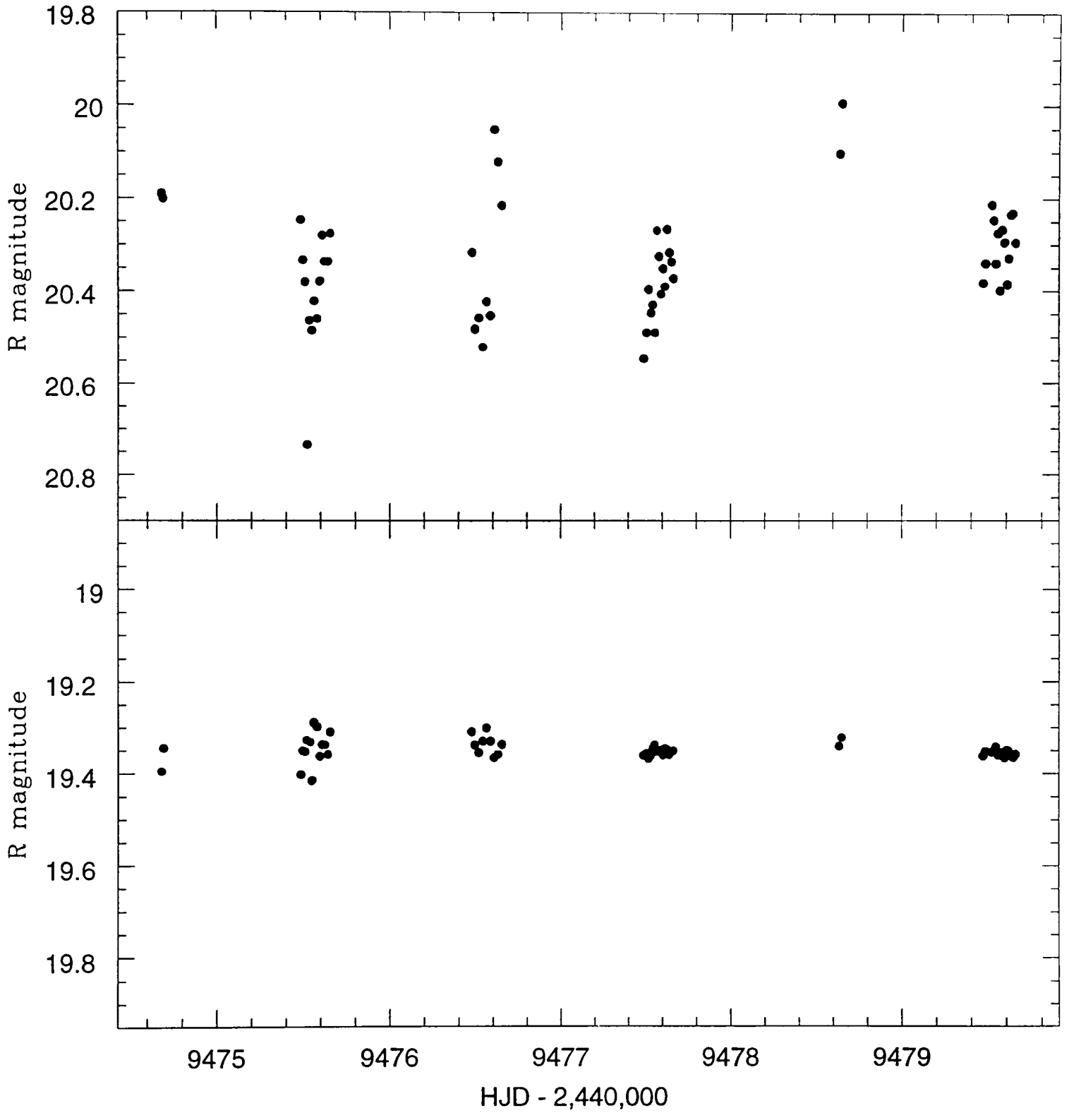


Figure 5.8: Top – R-band photometry of BW Cir (6 nights, 55 points); bottom – reference star C relative to reference star R ($\sigma = 0.022$).

5.2.3 Period Analysis of the Light Curve

Figure 5.9 shows the Lomb-Scargle periodogram (top) for BW Cir. The large gaps in the data confuse the periodogram, causing a set of regularly spaced peaks separated by 1 cycle per day. The largest peak in the periodogram occurs at a period of 7.82 $\pm$ 0.20 hours (3.07 $\pm$ 0.08 cycles/day), but the nearby aliases are too large to allow this to be associated with the orbital period. Also, the 99% confidence limit is quite high relative to the size of the peaks, but still indicates that there is a signal present in the data with a high degree of confidence (99.9 %).

As for V2107 Oph the sinusoidal amplitude required to produce the variance of the data can be estimated. Using the mean error of BW Cir ($\sigma_N = 0.037^m$), combined with the observed scatter ($\sigma_s = 0.12 \pm 0.01^m$), the full amplitude is 0.31 ± 0.03^m ($= 2\sqrt{2(0.12^2 - 0.037^2)}$). These values were put into the DFT fitting routine, to attempt to clarify the periodogram. The results for the $1/\chi^2$ statistic are shown in figure 5.9. The top panel shows the original periodogram (the two highest peaks occur at 7.82 and 5.91 hours); the middle panel shows the best fit periodogram which occurred at a period of 23.76 $\pm$ 0.23 hours and the bottom panel shows the $1/\chi^2$ statistic versus frequency. The $1/\chi^2$ statistic is clearer than the original periodogram, but still has some aliasing, and the proximity of the peaks to exact multiples of 1 cycle per day makes the result less convincing, as well as the disagreement between the original periodogram and the $1/\chi^2$ statistic. The final periodogram looks reminiscent of the window function of data with 1 day spacings, and so this could indicate that it is not possible to determine the orbital period of BW Cir from these data.

Finally the Bayesian ellipsoidal periodogram was applied to the data to try to determine the orbital period. This is shown in figure 5.10 along with the data folded on this period. The periodogram has a very sharp peak at 1.54 $\pm$ 0.02 cycles/day (15.6 $\pm$ 0.2 hours) and the peak probability is a factor $10^{1.6}$ larger than the next highest peak. However, the folded data do not show the expected ellipsoidal variation with differing minima, but instead have maxima that are 0.08^m apart (from the amplitude estimates below). Given the appearance of the light

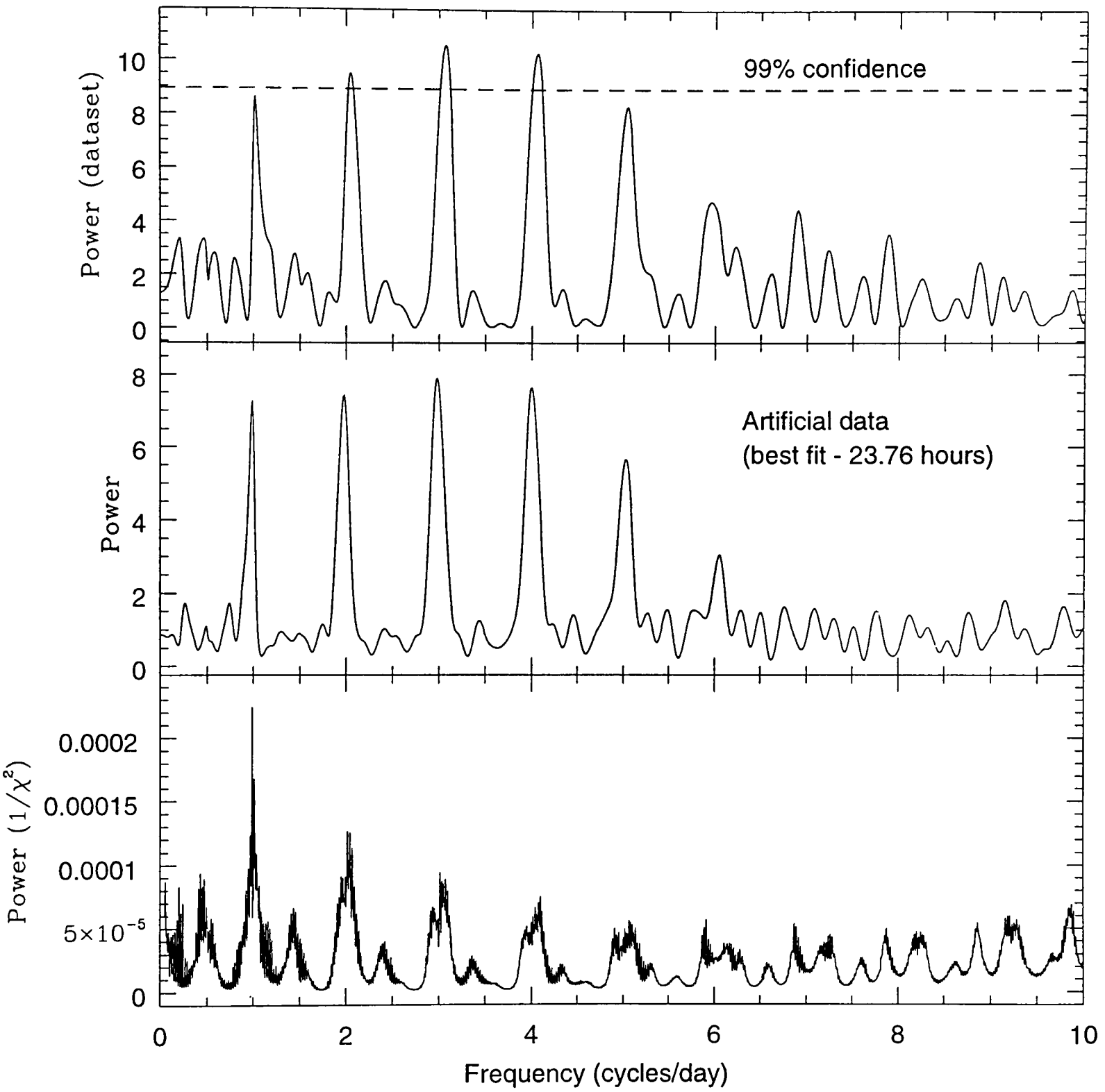


Figure 5.9: Top – Lomb-Scargle periodogram of BW Cir from 0.01 to 10.0 cycles per day; middle – the best fit periodogram (at 23.76 hours); bottom – the $1/\chi^2$ statistic versus frequency.

curve and the disagreement between this period and the earlier result from the $1/\chi^2$ statistic, this period is very tentative. The amplitudes obtained from the model are $0.13^m \pm 0.02^m$ and $0.08^m \pm 0.02^m$, and the variance is estimated to be 0.0083 ± 0.0017. This is again larger than would be expected purely from the statistical variance ($\sigma_N^2 = 0.037^2 = 0.001$). The excess variation is $\sigma = 0.083^m$, which is comparable to the levels found in the light curves of V404 Cyg and J0422+32.

5.2.4 The Inclination of BW Cir

The amplitude of the modulation, which was derived independently of the frequency, can be used to place a limit on the inclination of BW Cir. The amplitude of the ellipsoidal modulation depends on the inclination and mass ratio of the system, as well as other subsidiary parameters such as secondary temperature, and limb and gravity darkening coefficients. And so the amplitude can be used to obtain the inclination as a function of the unknown mass ratio. The parameters used here are the same as for V2107 Oph except for T_{eff} which was taken to be 5000 K (which changes the limb darkening to $u = 0.609$), a typical value for an early K spectral type main sequence star. The period appears to be reasonably short (15.6 hours or a 1 day alias of this period) and so the secondary should not have evolved very much given the mean density of the secondary, and all SXTs discovered so far have had companions with late spectral types between late G and early M. Even so, any error on this temperature has only a relatively small effect on the inclinations calculated below – a change of 500 K alters the inclination by 3°.

The amplitude of the ellipsoidal modulation is defined here to be the difference in magnitude between the maximum at phase 0 and the average of the two minima. This is equal to the amplitude found from the earlier, ellipsoidal analysis which was equal to $0.26^m \pm 0.046^m$. For a mass ratio of $q = 1$, this value cannot be reached by the model even for an inclination of 90°, but $q = 2$ can produce this large an amplitude. Thus, $q \geq 2$ is implied by the data, which is not surprising since all SXTs to date have large mass ratios. The inclinations for $q = 5$ and $q = 10$ are $i = 65° \pm 9°$ and $i = 61° \pm 7°$. For much larger values of q the inclination

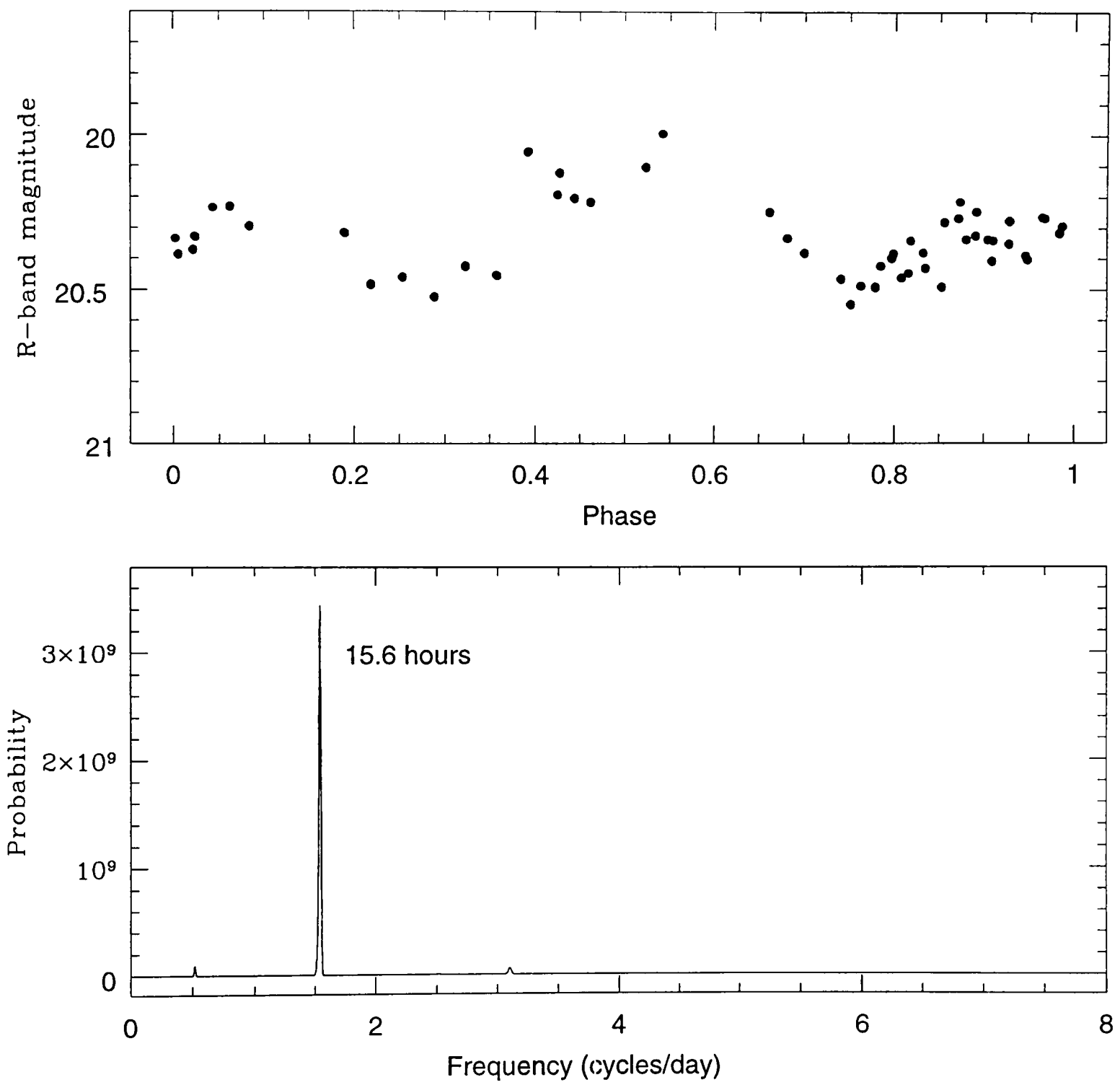


Figure 5.10: Top – Folded light curve of BW Cir (each point has an associated error of 0.037^m); bottom – the ellipsoidal periodogram of the data. The folded light curve does not look typically ellipsoidal and so the identification of the orbital period remains uncertain. The ellipsoidal periodogram has a sharp peak at 15.6 hours, which occurs at twice the value of the period found from the DFT (7.82 hours), as would be expected for this type of modulation.

changes by a very small amount, as the amplitude becomes virtually independent of q, and so the inclination of 61° can be used as a lower limit within its errors.

5.3 KY TrA (TrA X-1)

5.3.1 Introduction

TrA X-1 (A1524−61) was discovered by the *Ariel V* satellite (Pounds, 1974). Subsequent observations revealed that the X-ray emission of TrA X-1 reached a peak on November 22, 1974, which was followed by a second, larger peak 12 days later. TrA X-1 then declined with an e-folding time of 2 months. TrA X-1 was then optically identified with an optically faint nova (hereafter KY TrA) by Murdin et al. (1977), after which, using the similarity of the optical and X-ray behaviour to other X-ray transients such as A0620−00, they estimated the distance to be at least 3 kpc. More recently van Paradijs and Verbunt (1984) calculated a distance of 4.4 kpc, assuming $M_V(\mathrm{max}) = 1.0$ and E(B−V) = 0.7 (see section 5.3.3 for a discussion of the reddening to KY TrA). The value of $\overline{M_V}(\mathrm{max}) = 1.0 \pm$ was found to be an approximate mean value for all of the LMXBs (see van Paradijs, 1981). This is thought to be a consequence of the small range of X-ray luminosities found in LMXBs ($\sim 10^{36} - 10^{38}$ ergs s^{-1}).

KY TrA was initially considered a good black-hole candidate on the basis of its soft X-ray spectrum, although no hard X-ray tail had been found. However, several neutron star systems (e.g. Cir X-1; Tennant et al., 1988) were then observed to have similar soft X-ray properties to KY TrA, and so this source was removed from the list of black-hole candidates (Tanaka, 1989). Subsequently, the SIGMA telescope observed KY TrA in the range 41 – 152 keV on 1990 August 27 and 1991 January 21. On the first occasion it detected KY TrA and found a hard X-ray tail extending up to 90 keV, with a spectral index of 1.8 (± 0.7). The second observation did not detect KY TrA and so it is unclear whether the hard X-ray flux was due to normal emission or a different source state – if the former were the case, KY TrA would seem to be a more promising black-hole candidate. The probability that KY TrA was mis-identified in either of these observations is small – the nearest source to KY TrA is Cir X-1, which lies

5° away. The uncertainty in the position associated with the X-ray emission of the source is 5 arcminutes, at a confidence level of 90%, and so it seems likely that the X-ray emission originated from KY TrA. By chance, a few days prior to the first SIGMA observations (1990 August 17 until August 19), the ROSAT/PSPC also observed the region of the sky containing KY TrA (Barret et al., 1995). The unabsorbed X-ray luminosity of KY TrA was estimated to be $\sim 10^{36}$ erg s^{-1} (for a distance of 4.4 kpc). The fact that this possible 'mini' outburst was not detected by the X-ray all sky monitors at that time (namely GINGA and WATCH) could be due to the low intensity of the outburst. A second scan six months later revealed no X-ray emission with an upper limit of $\approx 10^{33}$ erg s^{-1} on the quiescent X-ray luminosity of KY TrA.

5.3.2 Observations

Five CCD frames of KY TrA were taken on the Danish 1.54m at ESO on two separate nights (May 3 and 5, 1994), during the V2107 Oph and BW Cir observing run. Three R-band and two I-band exposures were made in order to obtain colour information. All other technical details are the same as for the V2107 Oph photometry.

The previous finding chart for KY TrA was taken during its initial outburst and so an improved finding chart is shown in figure 5.11. KY TrA (labelled 'X' on the finding chart) lies in a crowded field, and so DAOPHOT (Stetson, 1987) was used to compute the magnitudes of KY TrA and two nearby reference stars. These reference stars ('R' and 'C') were inter-compared to measure their constancy. To calibrate the photometry the standard fields PG1323 and PG1633 were used, consisting of 1 and 3 standards respectively. The mean R-band and I-band magnitudes and R–I colour of KY TrA are shown in table 5.1.

5.3.3 The Reddening and Nature of the Secondary Star

Murdin et al. (1977) estimated the reddening to KY TrA by comparison with other nearby stars in the field, such as Circinus X-1, which is approximately 5° away and has a reddening of $E_{B-V} \sim$ 0.3 – 0.5 mag (Bok, Bok and Miller, 1972). A nearby G star (Murdin et al., 1977)

Table 5.1: Calibrated and dereddened magnitudes and colours for KY TrA ($E_{B-V} = 0.7$).

	R	I	R–I
Calibrated	22.37 ± 0.1	21.12 ± 0.1	1.25 ± 0.2
Dereddened	20.75 ± 0.1	20.07 ± 0.1	0.68 ± 0.2

has a reddening of $E_{B-V} = 0.2$, and from several other stars in the field of KY TrA which are all redder than this G star, they estimate a lower limit for the reddening to KY TrA of $E_{B-V} \geq 0.5$. Following this van Paradijs and Verbunt (1984) assumed a reddening of $E_{B-V} = 0.7$, to obtain a distance of 4.4 kpc. Unfortunately, no colours are available for KY TrA during its outburst and so no estimate for the reddening can be made from the assumption that a flat disc spectrum dominates during the outburst.

Using this estimate of the reddening, it is possible to deredden the above R and I magnitudes of KY TrA. To do this the interstellar reddening law of Savage and Mathis (1979) was used to calculate A_R and A_I given the above value of E_{B-V}. They quote $A_R = 2.32E_{B-V}$ and $A_I = 1.50E_{B-V}$, which gives $A_R = -1.62$ and $A_I = -1.05$. These two values are then used to give the dereddened magnitudes in table 5.1. Now, using the dereddened R–I colour, the spectral type of the secondary can be estimated assuming negligible disc contribution. This is reasonable given that the error on the spectral type due to photometric inaccuracies should be much larger than any systematic errors caused by this assumption. The temperature of the blackbody that corresponded to this R–I colour is 4100 K. This implies a spectral type of K7 for a main sequence star, with a possible range of K1–M4 given the error on the colour.

A rough estimate of the bolometric magnitude can also be calculated to indicate whether the secondary star is on the main sequence, or if it is evolved. Using the estimate of the reddening to obtain $A_v = 2.2$ and the distance, gives $M_v = 6.1$. For a spectral type around K7 the bolometric correction is -1.2, which gives the bolometric magnitude $M_{bol} = 4.9$. This is far too faint for the secondary star to be a giant ($M_{bol} \sim -1.0$, for a K-type star), and is closer to the typical values for a main sequence late-type stars ($M_{bol} \sim 6.0$). The

main uncertainties on this result are due to the distance, spectral type and reddening values. But, taking the possible ranges for these parameters into account gives an uncertainty in the bolometric magnitude of $\pm$ 2.0, which still does not easily admit the possibility that the secondary in KY TrA a giant star. However, the secondary star could still be partly evolved, i.e a sub-giant, similar in nature to V404 Cyg (Casares et al., 1993). Measuring the orbital period for KY TrA will be difficult due to its faintness, but this would give the mean density of the secondary with the assumption that it fills its Roche lobe, and so would better constrain its luminosity class. Radial velocity measurements of these faint SXTs will provide mass functions and should help to determine the nature of the compact object in these systems.

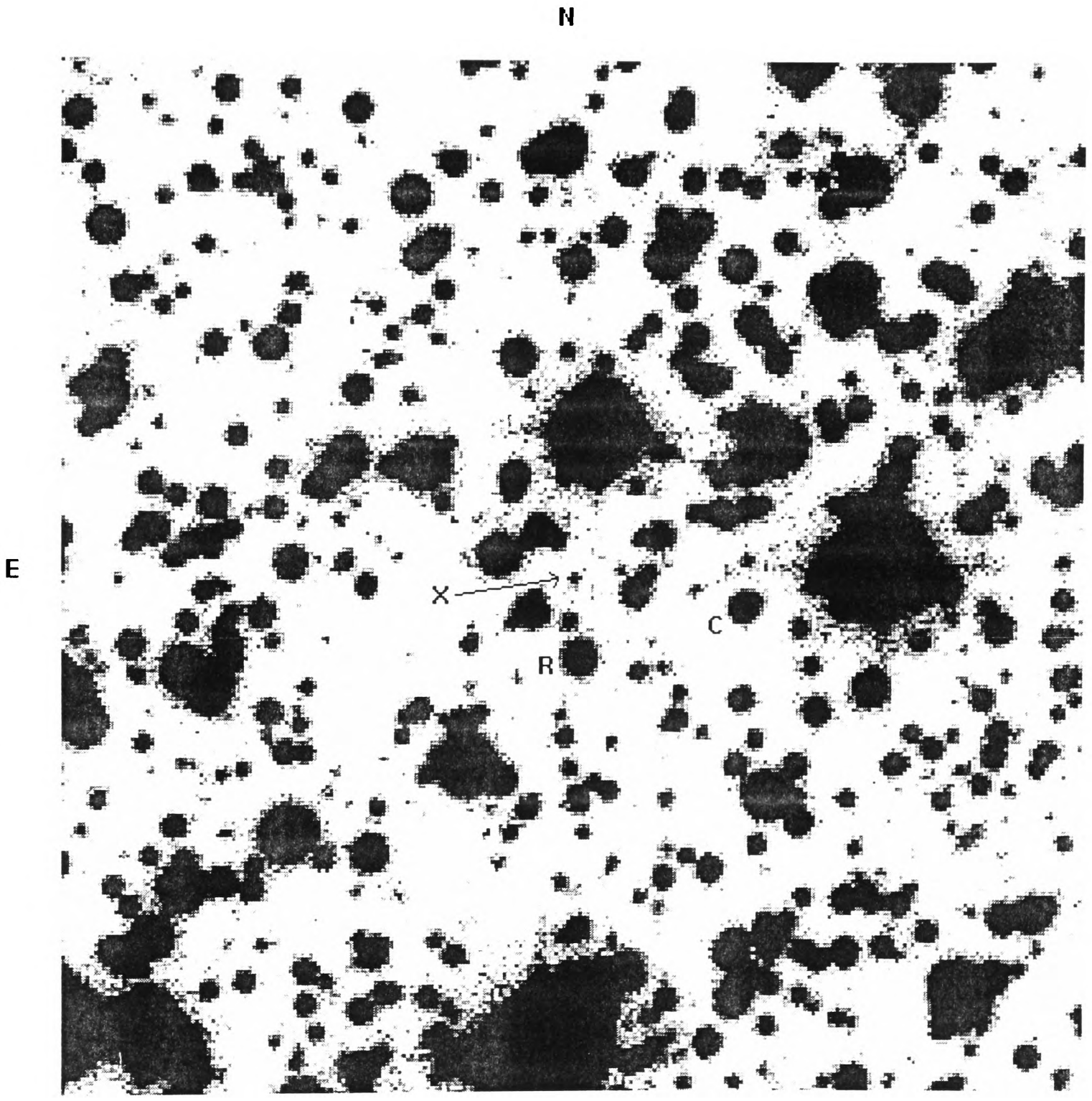


Figure 5.11: Improved finding chart for KY TrA (1800s exposure taken 1994, 5 May), labelled X. North is at the top, and east is to the left. Also shown are the two reference stars 'R' and 'C' used in the reduction. The total field of view is 1.19 x 1.19 arcmin^2.

References

Barret D., Motch C., Pietsch W. and Voges W., 1995, Astr. Astrophys., **296**, 459.

Bok B.J., Bok P.F. and Miller E.W., 1972, Astron. J., **77**, 733.

Casares J., Charles P.A., Naylor T. and Pavlenko E.P., 1993, MNRAS, **265**, 834.

Casares, J., Charles, P.A., 1994, MNRAS, **271**, L5.

Charles P.A., 1995, IAU Symp. No. 165, ed. E.P.J. van den Heuvel, in press.

Chevalier C., Ilovaisky S.A., van Paradijs J., Pedersen H. and van der Klis M., 1989, Astr. Ap.,**210**, 114.

Chodil G., Mark H., Rodrigues R. and Swift C.D., 1968, Ap. J. Lett., **152**, L45.

Cominsky L., Jones C., Forman W. and Tananbaum H., 1978, Ap. J., **224**, 46.

Cooke B.A., Levine A.M., Primini F.A. and Lewin W.H.G., 1984, Ap. J. 285, 258.

Francey R.J., 1971, Nature Phys. Sci., **229**, 228.

Griffiths R.E., Bradt H., Doxsey R., Friedman H., Gursky H., Johnston M., Longmore A., Malin D.F., Murdin P., Schwartz D.A. and Schwarz J., 1978, Ap. J., **221**, L63-L67.

Griffiths R., Johnston J., Doxsey R., Fabbiano G., Schwartz D. and Schwarz J., 1977, IAU Circ., 3110.

Haswell C., 1995, IAU Symp. No. 165, ed. E.P.J van den Heuvel, in press.

Ichikawa S., Hirose M. and Osaki Y., 1993, PASJ, **45**, 243.

Ilovaisky S.A., Pedersen H. and van der Klis M., 1987, IAU Circ., No. 4362.

Kaluzienski L.J. and Holt S.S., 1977, IAU Circ., 3104.

King A.R., 1993, MNRAS, **260**, L5.

Kitamoto S., Tsunemi H., Pedersen H., Ilovaisky S.A. and van der Klis M, 1990, Ap. J., **361**, 590.

Landolt A.U., 1992, Astron. J., **104**, 340.

Lomb N. R., 1976, Ap. Space Sci., **39**, 447.

Longmore A., Cannon R., Murdin P. and Malin D., 1977, IAU Circ., 3110.

Makino F., and *Ginga* Team, 1987, IAU Circ., No. 4342.

Markert T.H., Canizares C.R., Clark G.W., Hearn D.R., Li F.K., Sprott G.F. and Winkler P.F., 1977, Ap. J., **218**, 801.

Marsh T.R., Robinson E.L., Wood J.H., 1994, MNRAS, **266**, 137.

McClintock J.E., 1986, in *The Physics of Accretion onto Compact Objects* (eds. Mason K.O., Watson M.G. and White N.E.) 211 (Springer, Berlin)

McClintock J.E. and Remillard R., 1986, Ap. J.,**308**, 110.

McClintock J.E. and Remillard R., 1990, Ap. J.,**350**, 386.

Murdin P., Griffiths R.E., Pounds K.A., Watson M.G. and Longmore A.J., 1977, MNRAS, **178**, 27P.

Pedersen H., Ilovaisky S.A. and van der Klis, 1987, IAU Circ., No. 4357.

Popper D.M., 1980, Ann. Rev. Astr. Ap., **18**, 115.

Pounds K.A., 1974, IAU Circ., No. 2729.

Remillard R., McClintock J.E., Bailyn C.D, 1992, Ap. J.,**399**, L145.

Savage B.D. and Mathis J.S., 1979, Ann. Rev. Astr. Ap., **17**, 73.

Scargle J.D., 1982, Ap. J., **263**, 835.

Shahbaz T., Naylor T. and Charles P.A., 1993, MNRAS, **265**, 655.

Shahbaz T., Ringwald F.A., Bunn J.C., Naylor T., Charles P.A. and Casares J., 1994a, MNRAS, **271**, L10.

Shahbaz T., Naylor T. and Charles P.A., 1994b, MNRAS, **268**, 756.

Stetson P.B., 1987, PASP, **99**, 191.

Sunyaev, R.A. et al. 1991a, SvAL, **17**, 123.

Sunyaev R.A., Jourdain E. and Laurent P., 1991b, IAU Circ., 5176.

Tanaka Y., 1989, 23rd ESLAB Symposium Bologna, eds. J. Hunt and B Battrick, 3.

Tanaka, Y. and Lewin W.H.G., 1995, in *X-ray Binaries*, eds. W.H.G. Lewin, J. van Paradijs and E.P.J. van den Heuvel, Cambridge University Press, pg. 126.

Tennant A.F., 1988, Adv. Space Res., **8**, (2)397.

Tjemkes S.A., Zuiderwijk E.J. and van Paradijs J., 1986, Astr. Astrophys., **154**, 77.

van Paradijs J., Verbunt F., Shafer R.A. and Arnaud K.A., 1987, Astr. Astrophys., **182**, 47.

Verbunt F., Belloni T., Johnston H.M., van der Klis M. and Lewin W.H.G., 1994, Astr. Ap., **285**, 903.

Watson M.G., Ricketts M.J. and Griffiths R.E., 1978, Ap. J., **221**, L69.

Wilson C.K. and Rothschild R.E., 1983, Ap. J., **274**, 717.

Chapter 6

AC211 – an LMXB in the Core of M15

6.1 Introduction

M15 is one of the prototype post-core-collapse globular clusters, with a bright central cusp containing a high density of stars. Very close to the cusp is X2127+119, a bright LMXB 2 arcseconds from the core centre which was optically identified by Auriere et al. (1984) to be AC211, a 15th magnitude variable star. This was confirmed by Charles et al. (1986), by the discovery of Hα and HeII emission lines from the proposed optical counterpart. The period of the object was first identified by Naylor et al. (1986) to be 9.1 hours, then Hertz (1987) found an X-ray period of 8.66 hours and Ilovaisky et al. (1987) obtained an optical period of 8.5 hours, all of which agreed within their error limits. AC211 has $L_X/Lopt \sim 20$, which is typical of accretion disc corona systems (see White and Holt, 1982) where the X-ray source is hidden from view, but the discovery of an X-ray burst (Dotani et al., 1990), which reached a peak luminosity of 10^{38} erg s^{-1}, has shown that the neutron star is seen directly and that the normal X-ray luminosity is the total X-ray flux. This makes the optical brightness and large variability ($\sim 1.5^m$ in the U-band) even harder to understand. The high mass transfer rate required for AC211 has been explained using a model in which a neutron star captures a main sequence star just as it evolves onto the giant branch (Bailyn and Grindlay, 1987).

Subsequently, Ilovaisky et al. (1993) claimed that the period is in fact 17.1 hours, twice what was previously thought. If this period is correct the mass losing star would have to be a sub-giant in order to fill its Roche lobe. They derived this period partly from episodes when the source exhibited no modulation and remained at an intermediate brightness level. These recur with the 17.1 hour period and fall between phases 0.3 and 0.7. They also state that on the night of 30 June 1988, the object appeared to be abnormally faint with a V-band magnitude $\sim$ 18, whereas the minimum flux level of the source usually occurs at V $\approx$ 16.4. This could be due to a different source state and would give AC211 an L_X/L_{opt}, at that time, closer to a value typical of normal X-ray binaries, although this could not be confirmed from contemporaneous X-ray observations.

To search for further evidence in support of either periodicity, various observations of AC211 were investigated including photometry from 1988, 1992, 1994 and spectroscopy over the period 1986 – 1988. The Bayesian methodology outlined in Chapter 2 was used to examine the evidence for these two periods and also to determine if any other effects were present in the data.

6.2 TRIFFID Photometry

6.2.1 Introduction

Accurate photometry of AC211 is difficult due to the exceptionally crowded field and its proximity to the cusp of M15, and so requires very good seeing conditions. The new photometric data were taken in June 23 – 26 1992, and July 16 – 18 1994 on the 4.2 m WHT in La Palma using TRIFFID (TRansputer Instrument For Fast Image Deconvolution). TRIFFID is an image sharpening instrument using two high speed detectors for on-line data capture, with image sharpening taking place post-exposure. Separate sub-pupils are imaged onto each detector, image sharpening is then performed on each pupil separately before summation into a final image. These measurements were taken under excellent seeing conditions, and showed evidence for large-scale variability. No unmodulated state was seen during the observations.

The technique of image sharpening used by TRIFFID is described in detail by Redfern et al. (1993). Briefly, the instantaneous atmospheric effects causing random image motion, within each sub-pupil image, were determined by tracking the motion of a composite 'reference star'. The 'reference star' profile was produced by adding the relative photon centroid positions from the brightest stars within the field (up to 19 for this study) integrated in 50 millisecond segments. The image segments are then summed according to the motion of the 'reference star' to attempt to remove the effects of atmospheric turbulence on the final images.

The observation aperture is split into sub-pupils whose diameters are chosen to match the seeing, to enable the correct level of spatial over-sampling to be obtained. The optimum falls in the range $3 - 5\ r_0$, where r_0 is the Fried parameter, which can be considered to be the diameter of a telescope which has a diffraction limit equal to the available seeing (for a more detailed description see Beckers, 1993). During the 1992 observations r_0 was found to be in the range $15 - 20$ cm. These sub-apertures were summed after the initial sharpening process to obtain the final images.

6.2.2 Observations

A series of B-band images of AC211 were taken over the total of seven nights, the data were sharpened, and the best images obtained were selected. This resulted in a total of 46 sharpened images which were used in the following analysis. Some of the discarded images were unusable, as there were no well resolved stars, and the quality of others was variable. This is partly due to the fact that the sharpening technique works better for images taken during good seeing ($<$ 1 arcsecond), because of the inability to produce a good quality composite 'reference star' in the crowded field of M15 in bad seeing. To quantify the quality of the images the FWHM, for several test stars, was measured using IMEXAMINE in IRAF. The range of FWHM in the sharpened images varied from 5 (0.37 arcseconds) to 7 pixels (0.52 arcseconds).

To demonstrate the effectiveness of image sharpening an image of AC211 taken on July 17, 1994 with TRIFFID is shown in figure 6.1, along with a U-band image of the field taken with the Hubble Space Telescope (HST) on September 19, 1990. The HST image was exposed for a total of 800 s using the Planetary Camera (WFPC 1), and the F336W (U-band) filter. The image scale is 0.044 arcseconds/pixel. The HST image easily resolves AC211 from the contaminating light of the core of M15, two arcseconds from the source. The TRIFFID image also distinguishes AC211 with a resolution of 0.4 arcseconds, compared to 0.1 arcseconds for the HST observation. As can be seen from figure 6.1, sub-arcsecond seeing is essential to enable photometry of AC211 to be obtained.

Straightforward aperture photometry was performed on the images using the program PHOT in IRAF, because of the large light gradients and crowded nature of the field of AC211. The high sampling of the TRIFFID measurements (0.074 arcseconds/pixel) and excellent effective seeing (0.3 arcseconds), mean that it is feasible to use the centroid centreing routines in IRAF. The magnitudes of two, dim reference stars (AC220 and AC199), were measured relative to a brighter star (AC101), all of which are close to AC211. None of these are listed as variable in the AC catalogue of M15 (Aurière and Cordoni, 1981). AC211 was then compared to the bright reference star to generate the light curves of the two datasets. The radius of the circular aperture used was adjusted so as to obtain the least variation in the reference star's brightness, the aim being to minimize the noise in the light curve of AC211. Thus a circular aperture with a radius of 8 pixels (0.59 arcseconds) was used. The dominant source of error in the measurements is in the background determination, which is complicated by the close proximity of the core of M15 to AC211. When AC211 was bright it had an error comparable to that of AC207 ($\approx 0.1^m$) which is in a similarly crowded part of the field, and when it declined to its faintest level a larger value of 0.25^m was estimated. AC207 was not used as a reference star in this analysis as it was seen to be double in some of the frames, with a stellar separation of 0.3 arcseconds.

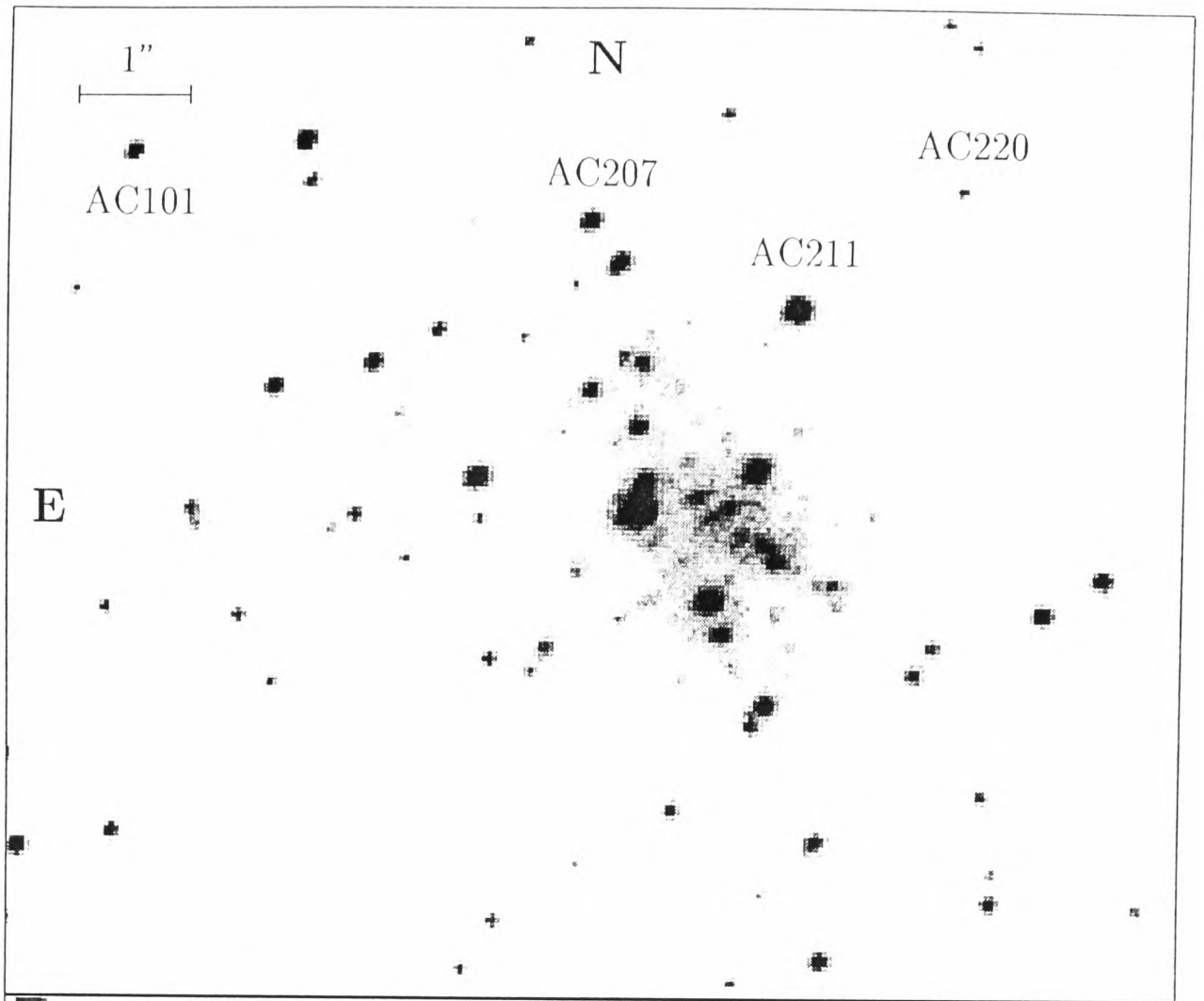


Figure 6.1: Top – An 800 s U-band exposure of the core of M15 taken with the Planetary Camera (WFPC 1) on board the Hubble Space Telescope, showing the reference stars AC207, AC220, AC199 and AC101 as well as AC211 itself. The field of view is 10.0 × 8.3 arcsec^2. Bottom – a 600 s ground-based TRIFFID image of M15. The field of view of this image is 13.3 × 11.1 arcsec^2. North is at the top and east to the left for both images.

6.2.3 Period Analysis

To search for any periodicities in the two TRIFFID datasets, the Bayesian periodogram described in Chapter 2 was used. The frequency range from 1 to 10 cycles was searched with a frequency step of 0.005 cycles/day, which is a little larger than the Nyquist limit for these datasets. The two TRIFFID datasets consist of only 21 and 26 points, which, combined with the large noise level present in the data means that it is important to test the significance of any signal detection. This was done by means of a series of 10,000 Monte-Carlo simulations of the data, using the same sample times and (assumed Gaussian) noise level of the data. This number of simulations means that the error on the significance is very small. A run of 10 Monte-Carlo simulations with 10,000 sets each were calculated for the 1992 data to empirically estimate the error on the 99% confidence level, yielding an error of 0.1 (the confidence level itself was 2.70, in units of $\log_{10}$(Probability)). This is a 4% error, which should be comparable to the error on the other significance levels obtained throughout this chapter. All of the signal detections in this chapter were well above the 99% confidence level and so the error on the significance can be neglected in the analysis that follows. It should be noted that the Monte-Carlo simulation used the photometric noise level of the data to produce the artificial noise sets and not the total variance, which means that the hypothesis being tested is whether there is a signal above the noise level of the data, but not if this signal is periodic.

The behaviour of the Bayesian statistic in the presence of noise has not been studied at present and so there is no analytical method of obtaining significance levels. Thus, the Monte-Carlo method was used as it does not rely on any knowledge of the characteristics of the Bayesian periodogram. Another possibility would have been to use a randomization test, where the time tags of the data points are randomized and the highest peak in the periodogram of the resultant data is found. This ensures that no assumptions about the distribution of the noise are made, and so there is no effect on the significance level. However, this is performed when there is reason to believe that some non-random effect exists in the data, whereas here the photometric noise should have little or no systematic components. Also, the randomization test uses the overall variation of the data, and so does not allow testing of whether there are

random effects above that expected by the measurement errors. To ensure the validity of this method, the randomization test was compared to the Monte-Carlo method for the TRIFFID data and it was found to be consistent. Therefore Monte-Carlo simulations were used to test significances in the following analysis.

The two periodograms for the data are shown in figure 6.2. The 1992 periodogram has a sharp peak at 16.6 $\pm$ 0.3 hours (1.45 $\pm$ 0.03 cycles/day) [1], well above any other peaks and also larger than the 99% confidence level (shown as a dashed line). The folded light curve in the next panel shows variability of 0.5^m in the light curve, but does not cover all phases and is not obviously sinusoidal. The 1994 TRIFFID data does not produce one clear peak in the periodogram. The three highest peaks in order of likelihood are: 9.1 $\pm$ 0.2 hours (2.65 cycles/day), 4.79 $\pm$ 0.07 hours (5.01 cycles/day) and 4.06 $\pm$ 0.07 hours (5.92 cycles/day). Only the 9.1 hour peak is close to a previously reported period, and there are no peaks near the 17.1 hour period, but this may be due to the short duration of each night's observations. The 1994 data folded on the 9.1 hour period are also shown in figure 6.2. They have an amplitude of 0.84^m, but again not all phases are sampled and so it must be concluded that the 17.1 hour variability could not have been unambiguously detected with these datasets.

The two TRIFFID datasets above were then combined to improve the above result. Initially, the two datasets were mean subtracted and the periodogram of the two datasets combined, was calculated. This is shown in figure 6.3. All of the main peaks seen in the two datasets are 99% significant, but it is not possible to determine which are produced by aliasing and the method of combining the data, and which are real. The 1992 and 1994 datasets had 1-σ variations which were significantly different (0.23 ± 0.01^m and 0.30 ± 0.02^m), which could be due to the existence of 'modulated' and 'unmodulated' states in AC211, as suggested by Ilovaisky et al. (1993). These differing states may be due to some masking effect in AC211, possibly by the disc blocking the light from the secondary star between phases 0.35 and 0.65. As the disc probably varies in a non-periodic manner, this would explain why

[1] All of the errors on the periods in this chapter were taken to be the region enclosing 99% of the probability in the peak of the Bayesian periodogram.

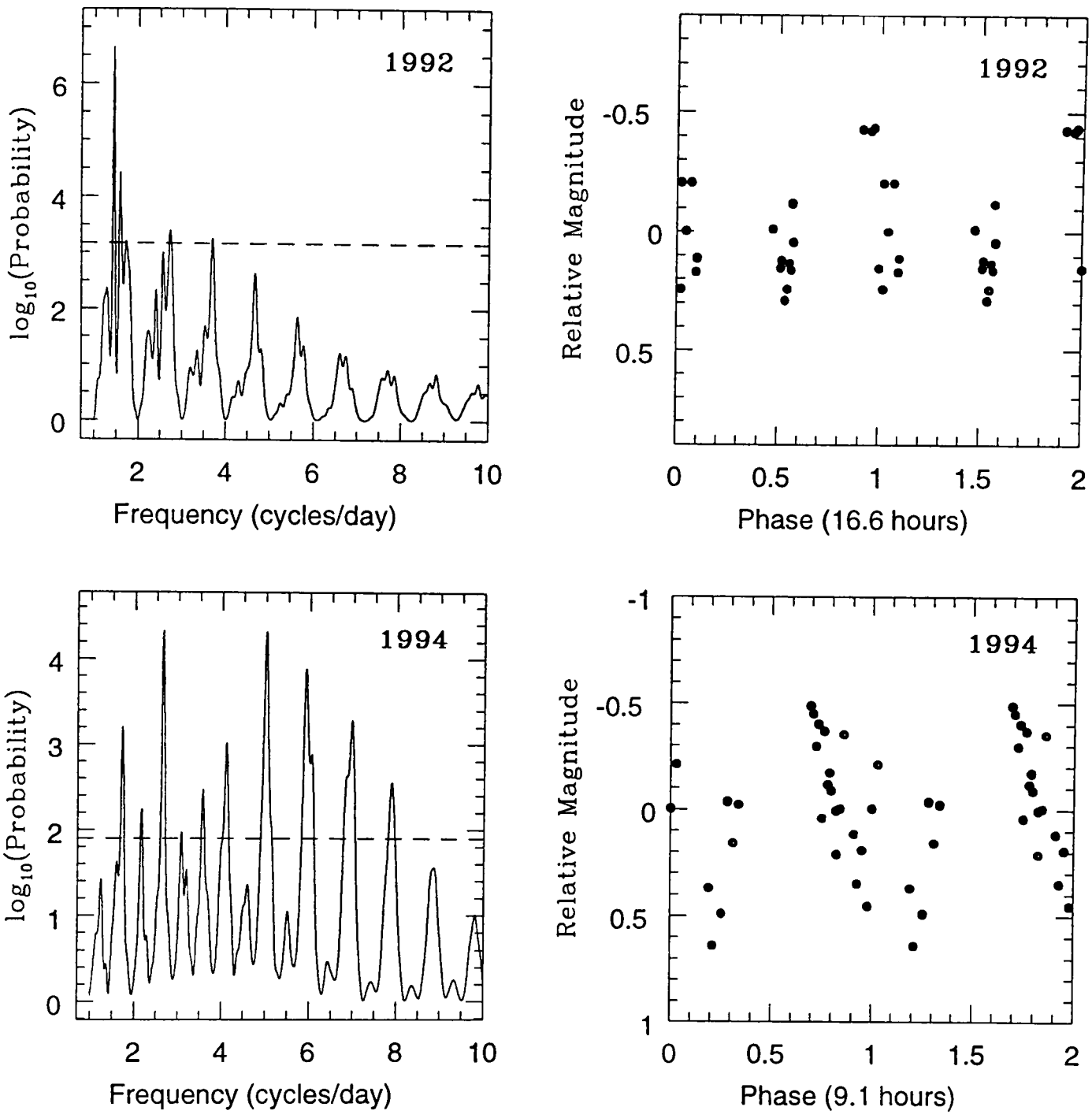


Figure 6.2: The periodograms for the 1992 and 1994 TRIFFID photometry. The most probable period for the 1992 data is 16.6 hours, but the 1994 data does not yield one clear peak, although of the two reported periodicities the peak near 9.1 hours is more likely than 17.1 hours. The 1992 data folded on 16.6 hour period and the 1994 data folded on the 9.1 hours are shown in the right-hand panels.

the eclipse feature seen at zero phase (when the secondary is in front of the disc) is the only reproducible feature seen in the light curve. This complicating factor could cause the multiple aliasing seen in figure 6.3(a), as the modulation in AC211 switches on and off, so this will produce power in the periodogram at several frequencies. Therefore, the two periodograms were simply convolved to remove the effect of any changes in the state of AC211 during the observations (see figure 6.3(b)).

Because the periodograms are on a probability scale, they can be multiplied to obtain the overall probabilities at each frequency. This joint analysis is more efficient than simply averaging the data, producing a $\sqrt{n}$ improvement under a wider variety of conditions as it makes fewer assumptions. Here it is only assumed that the three datasets exhibit the same periodic modulation. The result appears to be clearer, with two peaks at 14.0 ± 0.2 hours (1.71 cycles/day) and 8.92 ± 0.07 hours (2.69 cycles/day) reaching well above the 99% confidence limit. These two peaks are both 1-day aliases of each other and so it is likely that only one is the frequency of the real signal. Thus, the period at ≈ 8.9 hours is more likely given the above data, and not the 17.1 hour period.

Only the 1-dimensional periodogram was used in this analysis as there is not enough data to produce a significant result for the 2-D periodogram. From the model comparison section of Chapter 2 it was shown that for the 2-D model to be accepted as correct then the projection of the model onto the data (h_2) must be $\geq 3\ \sigma$, where $h_2 \approx N\hat{A}_2^2/2$. Thus, for a noise level of 0.2^m and an amplitude of 0.2^m, at least 30 data points are required to detect a second periodicity. A 2-D periodogram was computed for both datasets and no periodicity was found to be significant. For example the odds ratio of the two period model versus the single period model for the 1994 data was equal to $10^{0.85}$, which indicates that the two period model is not a significantly better fit to the data. This is consistent with the above rule of thumb that $\approx$ 30 points or more are necessary to significantly detect a second period in these data.

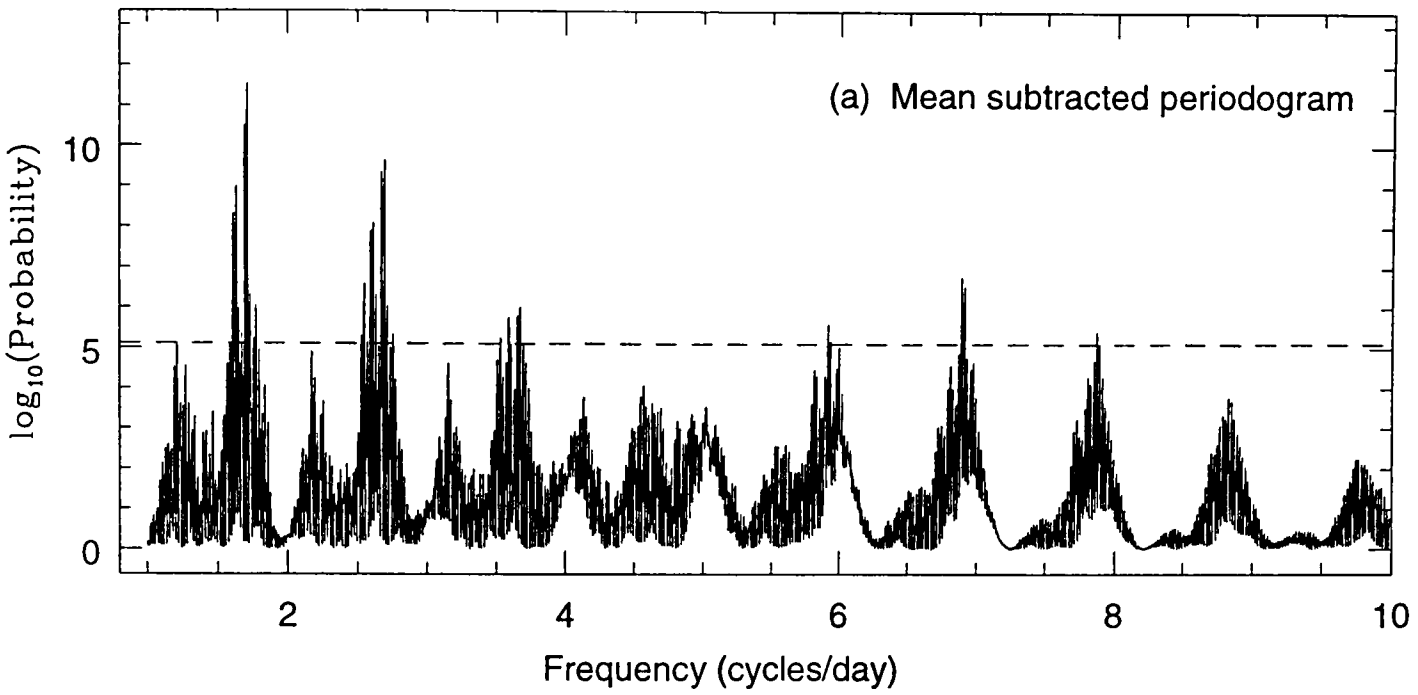


Figure 6.3: The combined periodograms for the TRIFFID datasets: a) the mean subtracted data and b) the convolved periodogram, obtained by multiplying the individual probabilities at each frequency. They are grossly similar with two peaks at 14.0 hours (1.71 cycles/day) and 8.9 hours (2.69 cycles/day). However, the convolved periodogram has its largest peak at 16.6 hours (1.45 cycles/day). The dashed lines indicate the 99% confidence lines derived as in the main text.

6.3 U-band Photometry

The U-band dataset was obtained on the WHT at one of the Nasmyth foci on the WHT in 1988 August 4 – 7, using a blue-sensitive GEC CCD chip. The exposure times were 100 s and the seeing varied from 0.7 to 1.0 arcseconds throughout the run. There was no image rotator, so all the images were rotated, and co-added to give a total integration time of 300 s/image. A total of 45 frames were taken over four nights, and the observing duration was 4 – 5 hours each night.

The data were reduced using STARMAN (Penny and Dickens, 1986), which was developed to cope with photometry in crowded fields and with large light gradients. The program DAOPHOT in IRAF, which was primarily developed for use in uniformly crowded fields, could not cope with large gradient in the light from the core of M15, separated by only 2 arcseconds from AC211. STARMAN uses a Lorentzian profile with wide, Gaussian wings to fit the stars in each image, as well as a linearly sloping background component. Initially, a set of isolated stars is used to define the profile for an image. These stars are then fitted, and the parameters for the best fitting profile is used to fit to the remaining stars by varying only the position, height and background level. After finding and fitting all of the visible stars, they are subtracted to show up any faint background stars which were not initially resolvable. These stars are then also removed and the process is repeated until there are no stars left in the subtracted image.

To generate the light curve, AC211 was measured relative to the nearby star AC207, and the calibrated magnitude was calculated from the U-band magnitude of AC207 (14.97^m) taken from Ilovaisky et al. (1993). The errors on the photometry were difficult to determine as they are dominated by the determination of the background light from the cluster core. But, by comparison with the error on AC207, errors of 0.1^m and 0.25^m were found to be appropriate when AC211 was bright and faint, respectively (the amplitude of the light curve in the U-band is $\sim 1.5^m$).

6.3.1 Period Analysis

The U-band light curve shows strong evidence of variability, the U-band magnitude varies by about 1.5^m, with a mean level of 15.78 ± 0.3^m. The period analysis was performed, as for the TRIFFID data, over a range of 1–10 cycles per day, yielding several peaks (see figure 6.4). The 99% confidence level was 2.2 ($\log_{10}$(Probability)) and so at least two peaks in the periodogram are significant, although only one is likely to be real as they are all separated by a 1-day alias. The two highest peaks are: 13.0 ± 0.4 hours (1.85 cycles/day) and 8.5 ± 0.2 hours (2.84 cycles/day). The semi-amplitudes for these two periods are 0.60 ± 0.06^m and 0.58 ± 0.07^m, which are consistent within their respective errors. The residual noise level of the data (1-σ) is equal to 0.43 ± 0.05^m or 0.45 ± 0.05, respectively for the two periods. This is higher than is suggested by the photometry, which implied a 1-σ error of approximately 0.2^m. This discrepancy may be due to additional systematics uncertainties in the photometric measurement, or could indicate that there is an additional source of random flickering in the light curve similar to that observed in the SXTs.

The U-band data folded on the two most likely periods are shown at the top of figure 6.4, but the light curves do not appear to be clearly sinusoidal as there is a jump in the magnitude level once per cycle. This effect is normally associated with aliasing in the periodogram, but if these peaks are aliases there are no other peaks in the periodogram which could be the true signal. The light curve was also folded on 28.2 hours (0.7 cycles/day), which is also an alias of these peaks, but the clustering effect was even more pronounced with a phase coverage of only 0.58, and the scatter on the folded light curve was very large. Therefore, this cannot be the true period.

The total number of measurements in the U-band data is 45, and so it is feasible to search for multiple periodicities in the data. A 2-frequency, Bayesian periodogram was computed, part of which is shown in figure 6.5 ranging from 1 to 8 cycles/day. Two peaks symmetrical about the $f_1 = f_2$ axis are evident, corresponding to frequencies of 9.6 ± 0.4 hours (2.51 cycles/day) and 12.7 ± 0.6 hours (1.90 cycles/day). The probability of this model relative to

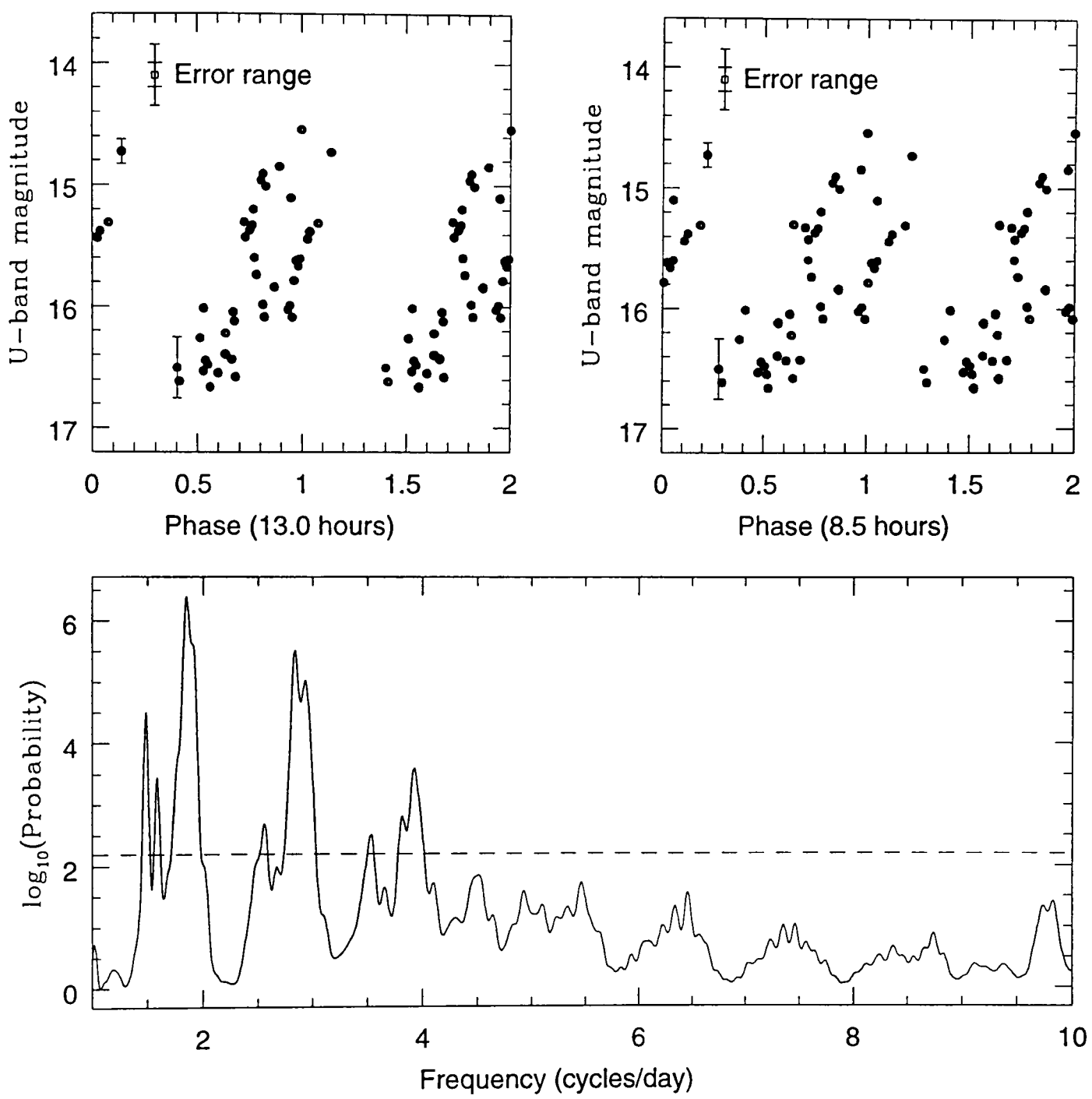


Figure 6.4: The Bayesian periodogram of the U-band photometry from 1–10 cycles/day. Significant peaks are seen at 13.0 hours (1.85 cycles/day) and 8.5 hours (2.84 cycles/day). A peak does appear at 16.1 (1.49 cycles/day), which may indicate the existence of a 17.1 hour periodicity in these data. The 99% confidence level is shown as a dashed line.

the 1 frequency model is $10^{1.4}$, which is only marginally in favour of the 2 frequency model.

6.4 Spectroscopy

Spectroscopic observations of AC211 were taken in 1986 August 1 – 3; 1987 August 16 – 19; 1988, August 4 – 7 at the 2.5 m INT on La Palma. The IPCS and RGO Intermediate Dispersion Spectrograph were used, and the resolution of the spectra obtained was 1.5 Å over a spectral range of $\lambda\lambda$ 3950 – 4975. The spatial resolution was such that, along the slit there were 0.4 arcseconds/pixel, oversampling the seeing and enabling AC211 to be distinguished from other stars in the crowded field of M15. The seeing varied from 0.6 – 1.0 arcseconds.

The spectra were flat-fielded and then wavelength calibrated using a series of Cu-Ar arcs fitted with 5th order polynomials. The drift in the wavelength of a particular channel with time was less than $\pm$ 0.2 Å. A further check on the wavelength calibration was made using the planetary nebula K648, in M15. This was observed and the Balmer lines were fitted with a single Gaussian to measure the radial velocity of K648. The scatter of these values implied an error of 0.3 Å or less. These errors are similar for AC211, but are smaller than the error incurred in the line fitting. Further details of the observations and reduction are given in (Naylor, 1988).

The signal-to-noise was not large enough in the 600 s single measurements, and so the spectra were binned such that the length of time from the beginning of the first spectrum to the end of the final one was 1 hour. The HeI λ 4471 position was measured using a single Gaussian least squares fit to each profile.

Periodograms of the data from each of the three years (1986 – 1988) were computed separately (figure 6.6). The 99% confidence regions were calculated as before using the mean errors for each dataset, σ_{86} = 23.7 km s^{-1}, σ_{87} = 10.3 s^{-1} and σ_{88} = 23.5 km s^{-1}. The largest peak in the 1986 periodogram occurs at a period of 9.4 $\pm$ 0.2 hours (2.56 cycles/day), and there is no peak near the period of 17.1 hours found by Ilovaisky et al. (1993). Both of

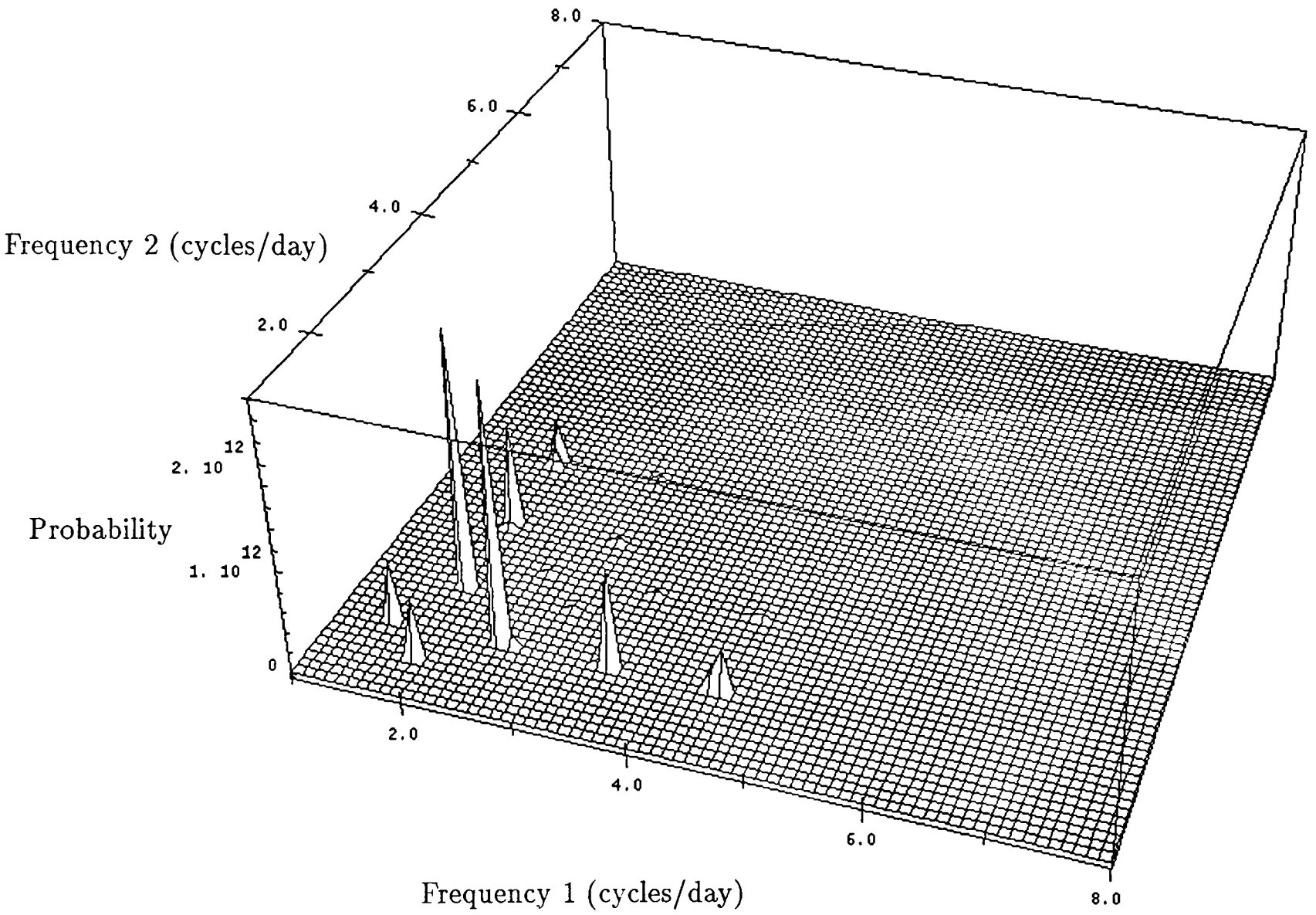


Figure 6.5: The 2-D periodogram of the U-band data. Two peaks appear corresponding to the frequencies 9.6 ± 0.4 hours (2.51 cycles/day) and 12.7 ± 0.6 hours (1.90 cycles/day). Note the probability scale, which indicates that there are several peaks with a similar probability of being correct.

these periods are indicated by a vertical, dotted line. The 1987 has three peaks which are only marginally above the 99% confidence limit, at 6.92 hours (3.47 cycles/day), 6.10 hours (3.94 cycles/day) and 5.44 hours (4.42 cycles/day). It was not possible to calculate the errors on these values from the area enclosing 99% of the local probability, because the peaks only reached a level of $10^{1.2}$ above the underlying probability level. However, the width of these peaks indicates that the precise values are very uncertain and given the 99% confidence level no useful conclusion can be made. The 1988 radial velocities yielded a periodogram with a large peak near to 17.1 hours, although the peak is divided into two aliases at 17.9 $\pm$ 0.6 hours (1.34 cycles/day) and 19.5 $\pm$ 0.6 hours (1.23 cycles/day). There are no other significant peaks in the periodogram.

To combine the results for these spectra without causing aliasing due to the large separation of each dataset it is possible to convolve the periodograms into one, as was done previously on the TRIFFID data. The convolved spectroscopic periodogram is shown in figure 6.7. There are two significant peaks near the 17.1 hour and 8.55 hour periods (19.4 $\pm$ 0.6 hours and 9.4 $\pm$ 0.2 hours). But these peaks have subsidiary aliases next to them at, 17.8 $\pm$ 0.6 and 8.8 $\pm$ 0.2 hours, both above the 99% confidence level. They are separated from the main peaks by $1/5.56$ days^{-1}, which is equal to the mean resolution of the 1986 – 1988 datasets ($= 1/2\Delta T$ $= 1/5.6$ days^{-1}), implying that two of these peaks are produced by aliasing.

The 99% confidence region for the convolved periodogram was obtained by taking the 0.215 ($= \sqrt[3]{0.01}$) significance points in each of the three periodograms and summing them to obtain the 99% confidence level of the combined periodogram (the values for each of the three years were: 0.912 + 0.786 + 1.499 = 3.197). Thus, the final confidence level is the combined probability that noise could have reached the given level only 1% of the time.

Only the 1986 and 1988 datasets gave significant detections of a periodicity in AC211 (at 8.55 hours and 17.1 hours, respectively), and these folded light curves are plotted in figure 6.8. The mean gamma velocities for the two years are -271.3 km s^{-1} and -252.2 km s^{-1}, and the semi-amplitudes of the two modulations are 32.8 $\pm$ 8.0 km s^{-1} and 18.73 $\pm$ 4.7 km

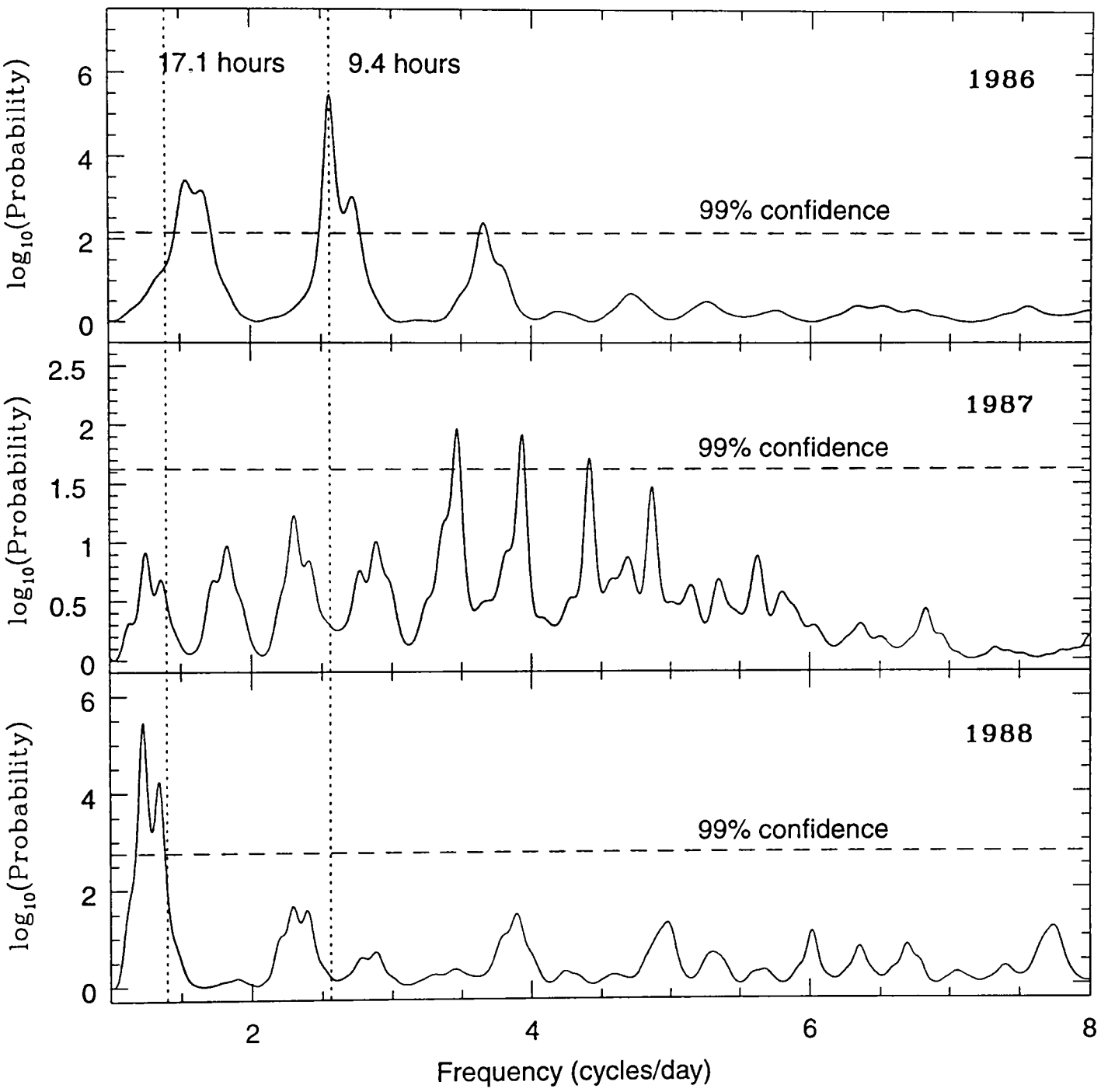


Figure 6.6: The Bayesian periodograms for the 1986 – 1988 HeI λ 4471 spectroscopy. Only the 1986 and 1988 datasets give significant detections of a signal near the expected periods. The 1986 data favours the shorter period, whereas the 1988 data finds the period near 17.1 hours to be more probable (the two periods are indicated by a vertical, dotted line). The 99% confidence limits, derived as before, are shown as dashed lines.

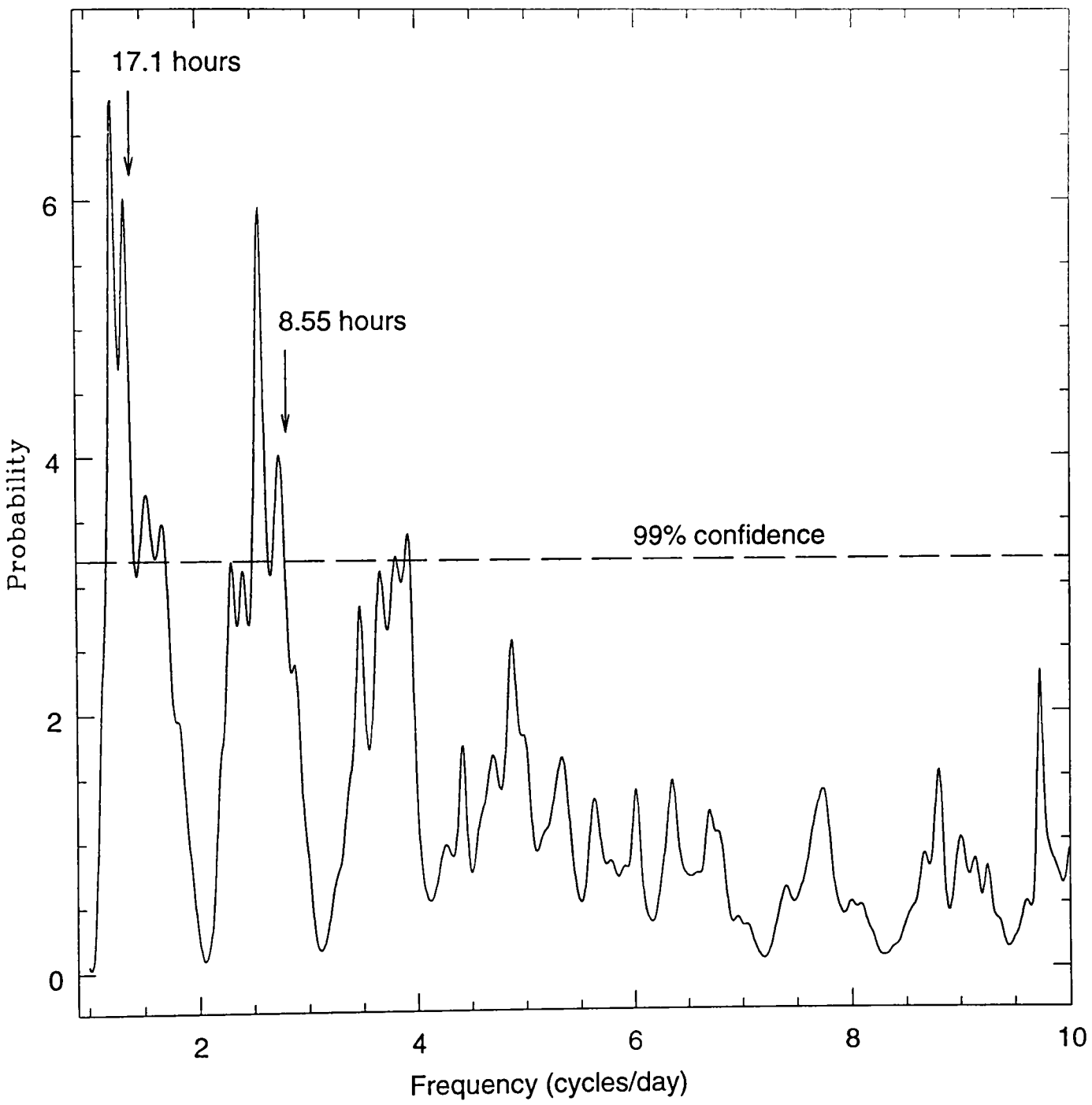


Figure 6.7: The convolved spectroscopic periodogram of $\log_{10}$(Probability) versus frequency. Two peaks near 17.1 and 8.5 hours are well above the 99% confidence limit, derived as in the main text. The 17.1 hour peak is more probable, but only by a relatively small factor (~ 10).

s^{-1}. The gamma velocity is very large with respect to the other cluster members (-105 km s^{-1}, see Peterson, Olszewski and Aaronson, 1986), implying that AC211 is not gravitationally bound to M15. However, Bailyn et al. (1989) suggested that the apparent gamma velocity is due to outflow near the L_2 point (see section 6.5). The radial velocity is quite small, and is thought to arise from orbital changes in the absorption of an envelope surrounding AC211.

The 1988 spectroscopy and the 1988 U-band data were taken simultaneously and so it is possible to calculate the relative phasing without incurring a large error because of errors in the period assumed. This gives a relative phase between the point of maximum light and the red to blue crossing point of the spectroscopy of 0.051 ± 0.11 (for the 17.1 hour period). This implies that they occur at the same time within the random error. This is consistent with the model of AC211 proposed by Bailyn et al. (1989), which stated that the HeI λ 4471 variations are caused by outflow near the L_2 point. If this model is correct the phase zero above would correspond to the point where the secondary is behind the compact object, and so the system is at maximum light. However, the unusual radial velocity curve predicted by the model is not observed.

Finally, the entire set of TRIFFID, U-band and spectroscopic data were convolved in a similar way as before, into a single overall periodogram. This requires the implicit assumptions that the photometric and spectroscopic periods are the same and that they are stationary (unchanging with time). This periodogram is shown in figure 6.9. The highest peak occurs at 9.4 ± 0.2 hours (2.56 cycles/day) close to the 9.1 hour period found in earlier studies. Near 17.1 hours there are a series of peaks with a large probability, almost as high as the 9.4 hour peak. The fact that this periodogram was constructed by combining datasets which only had observing durations of around $4 - 5$ hours may have biased against the detection of the 17.1 hour period, possibly causing this effect. The second highest of this group of peaks corresponds to 16.4 ± 0.3 hours (1.46 cycles/day).

The light curve observed by Ilovaisky et al. (1993) folded on his period of 17.1 hours is very complex and not even approximately sinusoidal. This means that there would be a lot of

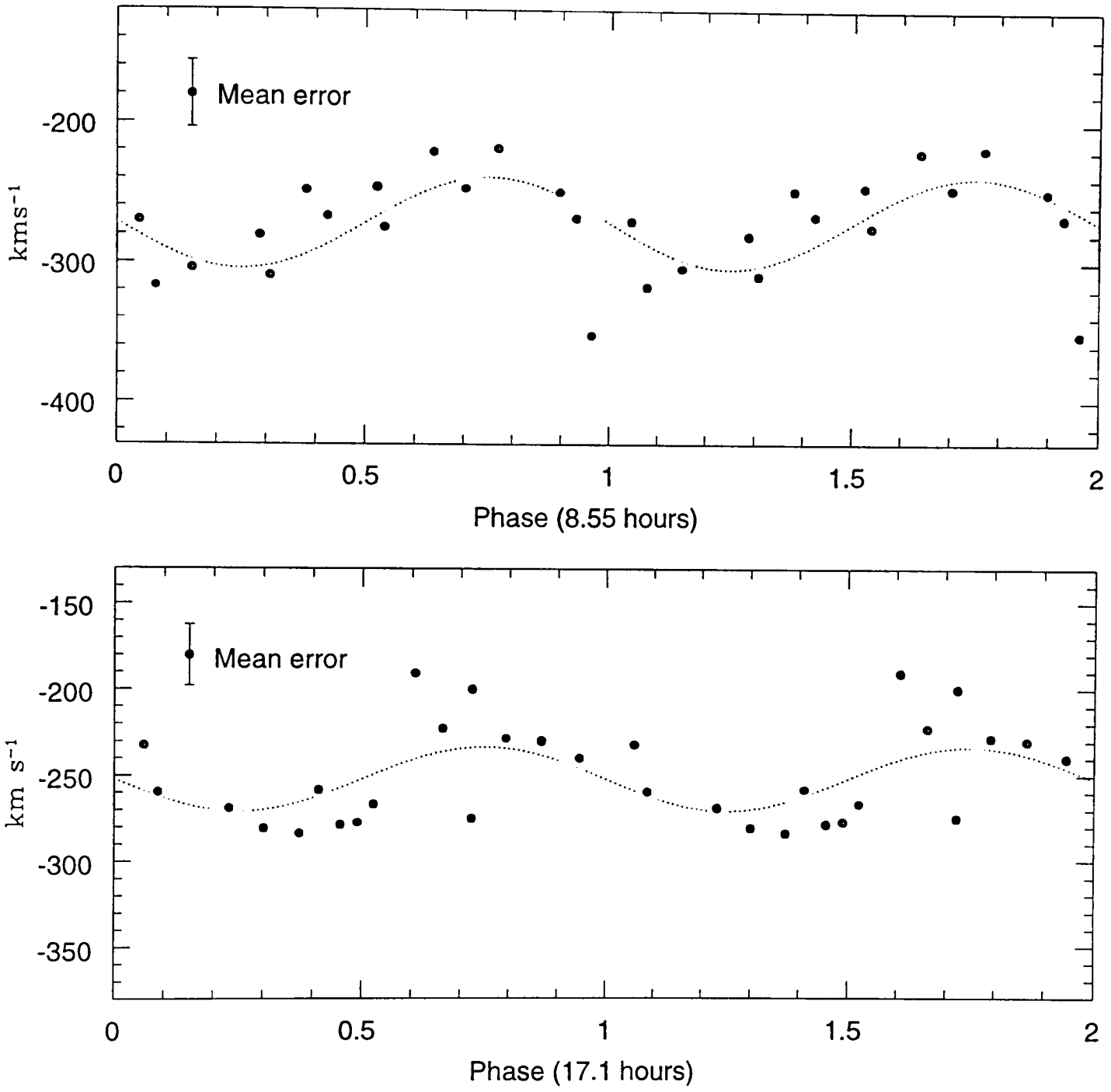


Figure 6.8: The 1986 and 1988 HeI λ 4471 spectroscopy folded on the two most probable periods for these data (8.55 hours and 17.1 hours). The gamma velocity is substantially different from the cluster velocity and the semi-amplitude of the modulation is relatively small.

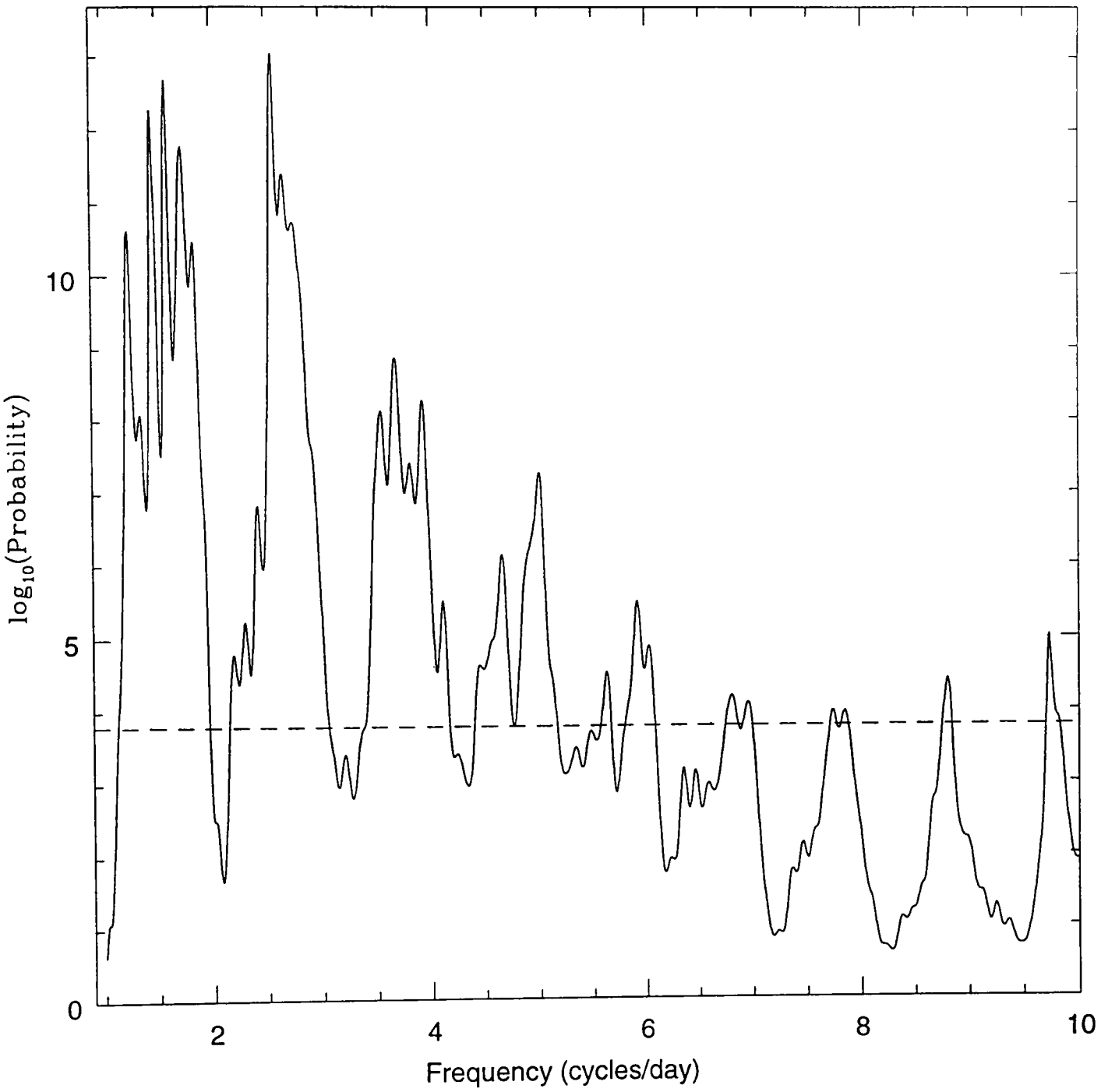


Figure 6.9: The convolved periodogram of the entire set of data on AC211. Peaks at 9.4 hours and around 17.1 hours appear well above the 99% confidence level of 3.79 (=$\log_{10}$(Probability)) calculated from the individual confidence levels assuming they are all independent. If this is not assumed, a more conservative 99% confidence level of 8.88 is obtained. The two highest peaks are still well above this level and so there is a significant probability that either the 8.55 or 17.1 hour periods are present in the data.

power at the harmonics of the 17.1 hour period, 2.85, 3.85 cycles/day and so on. This could explain the confusion that exists between the 8.55 hour and 17.1 hour periods in the data presented here. This is analagous to that seen in the detection of SXT periods, which have two components at frequencies f_{orb} and $2f_{orb}$. Because of the complex nature of the light curve, a multi-periodic analysis of the data would require many components and there is not enough data to justify this and obtain a significantly high probability of the model. The two peaks in the periodogram are well above the confidence limit and so it can be said without doubt, that there exists a signal significantly above the noise level of the respect sources of data.

The 99% confidence line was obtained by summing the 0.464 ($= \sqrt[6]{0.01}$) significance levels from each dataset, which gave 3.79. This level appears somewhat low given the number of peaks which lie above it, and so a better, more conservative estimate can be made by taking all of the 99% confidence limits and summing them to obtain the overall limit of 8.88. This appears to be more consistent with the periodogram. The limit obtained by the latter method guards against any systematic effects which could cause spurious peaks in all of the datasets in the same way, as the datasets are not considered to be independent.

6.5 Discussion

The large amplitude of the light curves seen in AC211 implies that the inclination is high, but the observation of a type I X-ray burst (Dotani et al., 1990) means that the X-ray emitting source can be seen directly, limiting the inclination to be less than $\approx 80°$ (if the disc flares by more than 10°, which is very likely, see Meyer and Meyer-Hofmeister, 1982). Also, the low X-ray to optical ratio ($L_X/Lopt \sim 20$) implies that the optical luminosity is high, since it seems unlikely that the X-ray luminosity has been underestimated as we see the source directly. Models which attempt to explain this and other features of AC211, invoke an ionized accretion disc corona (Bailyn et al., 1989), which scatters X-rays back onto the disc causing it to become hot (22,000 K). In this model the HeI λ 4471 absorption is caused by cooler outer gas. The hot inner gas is strongly suggested by the blue colour of AC211 (U−B

$= -1.2$) and the strong HeII emission. They also suggested that the HeI λ 4471 absorption could originate in an outflow near the L_2 point. Their model reproduced the blue-shifted gamma velocity of the line with respect to the rest of the cluster ($\approx$ 150 km s^{-1}), and the semi-amplitude of the radial velocity variation ($\approx$ 60 km s^{-1}) was also fairly close to the observed value.

The morphology of the optical light curve observed by Ilovaisky et al. (1993) is very complex. A reproducible minimum at phase zero allowed the determination of the 17.1 hour period, the depth was 1 magnitude over 0.3 in phase and may be related to a partial eclipse of the disc by the companion. Between phases 0.2 – 0.7 the light curve appeared to be variable, sometimes with a secondary minimum, at other times AC211 stayed in a low, unmodulated state. The non-sinusoidal nature of the light curve may explain the confusion between the 8.5 and 17.1 hour periods. Many other datasets prior to the Ilovaisky et al. result found evidence for variability at the shorter period, and to date the determination of the period of AC211 remains inconclusive. The data presented here show evidence for both periods in the Bayesian periodograms of the spectroscopy and photometry and so confirmation of the correct period of AC211 remains elusive.

Ilovaisky et al. (1987) used the 8.5 hour period, combined the relationship of Paczynski (1971) and Kepler's third law to obtain a radius of 0.9 $R_\odot$, for a mass less than 0.8 $M_\odot$. The mass is constrained to be less than 0.8 $M_\odot$ by the fact that the stars in globular clusters are older than 10^{10} yrs, and so must have a mass less than this value not to have evolved off the main sequence. Also, the 8.5 hour period implies a mean, stellar density $\overline{\rho} = 1.52$ g cm^{-3} ($= 110/P_{hr}^2$). These parameters are close to those for a G main sequence star, near to the turn-off. However, the 17.1 hour period implies that the density of the companion star in AC211 is small ($\overline{\rho} = 0.38$ g cm^{-3}). For a G spectral type star or later the density is at least 1.35 (G0), and so this longer period implies that the secondary star is an evolved sub-giant. However, the stripped giant model of is not applicable to AC211 as the orbital period is certainly less than 1 day, and therefore not underdense. The nature of the secondary star in AC211 depends critically on the orbital period, and only when this is known can progress be

made in understanding this unusual system.

AC211 is still not well understood at this time. Without the orbital period it is difficult to estimate any of the physical parameters of the system, and models to explain the large amplitude of the light curve, optical brightness and other unusual features of AC211 remain speculative. The close proximity of the core of M15 makes it a difficult object to study, with background light from the cluster scattering into the observations. HST is almost ideally suited to studying AC211, as it could obtain photometry in the ultra-violet where AC211 is brightest, with the required resolution from the core, but the probable 17.1 hour period means that it is not feasible to monitor the object for the amount of time that is necessary. It will require the development of techniques like TRIFFID and adaptive optics to produce the spatial resolution required, combined with a long enough time-base to detect the orbital period unambiguously.

References

Aurière M. and Cordoni J.-P., 1981, Astron. Astrophys. Suppl. Ser., **46**, 347.

Aurière M., Le Fèvre O. and Terzan A., 1984, Astr. Astrophys., **138**, 415.

Bailyn C.D. and Grindlay J.E., 1987, Ap. J., **316**, L25.

Bailyn C.D., Garcia M.R. and Grindlay J.E., 1989, Ap. J., **344**, 786.

Beckers J.M., 1993, Ann. Rev. Astr. Ap., **31**, 13.

Charles P.A., Jones D.C and Naylor T., 1986, Nature, **323**, 417.

Dotani T., Inoue H., Murakami T., Nagase F., Tanaka Y., Tsuru T., Makishima K., Ohashi T. and Corbet R. H. D., 1990, Nature, **347**, 534.

Hertz P., 1987, Ap. J., **315**, L119.

Ilovaisky S.A., Aurière M., Chevalier C., Koch-Miramond L., Cordoni J.-P. and Angebault L.P., 1987, Astr. Astrophys., **179**, L1.

Ilovaisky S.A., Aurière M., Koch-Miramond L., Chevalier C., Cordoni J.-P. and Crowe R.A., 1993, Astr. Astrophys., **270**, 139.

Meyer F. and Meyer-Hofmeister E., 1982, Astr. Astrophys., **34**, 106.

Naylor T., Charles P.A., Drew J.E. and Hassall B.J.M., 1988, MNRAS, **233**, 285.

Naylor T., Charles P.A., Callanan P.J. and Redfern R.M., 1986, IAU. Circ., No. 4263.

Paczynski B., 1971, Ann. Rev. Astr. Ap., **9**, 183.

Penny A.J., Dickens R.J., 1986, MNRAS, **220**, 845.

Redfern R.M., Devaney M.N., O'Kane P., Ballesteros-Ramirez R., Gomez-Renasco R. and Rosa F., 1989, MNRAS, **238**, 791.

White N.E. and Holt S.S., 1982, Ap. J., **257**, 318.

Chapter 7

Conclusions

The most important step in understanding any binary system is to obtain its orbital period. Subsequently, knowledge of the remaining binary parameters enable the system to be characterized and allow further investigations to shed light on the specifics of each system. Thus, only when a system is known to have a black-hole primary is it possible to begin to understand the underlying physics of the observational effects seen in that system. This means that it is important to understand what selection effects apply to the various binary parameters, as well as attempting to discover if any physical processes cause such selection effects, such as the period gap in the CV distribution.

The Mass Distribution of Compact Objects

From the studies in this work and elsewhere it has become evident that the masses of LMXBs divide into two distributions: the high mass black-hole candidates and the lower mass neutron star systems (see Chapter 1). Many of the lower mass systems are known to have neutron star compact objects because of type I X-ray bursts, which occur on the surface of the star, or because of pulsations. The low mass objects all seem to cluster around a mass of 1.4 $M_{\odot}$, the canonical mass of a neutron star, whereas theoretical models allow masses up to 3.8 $M_{\odot}$, for a rapidly rotating neutron star. What is the significance of this clustering in the neutron star mass distribution, and what can be said about the range of black-hole

masses?

The neutron star masses measured to date (see table 7.1) are tightly clustered. The mean of these masses is 1.35 $M_\odot$ and the 1-σ width of the distribution is 0.22 $M_\odot$. Much of the scatter in the neutron star masses can be explained by the error on the mass determinations in each case, which may mean that the actual distribution is even more tightly clustered about 1.4 $M_\odot$. This is in contrast to the black-hole candidate systems, (see table 7.2) which have a mean mass of 7.1 $M_\odot$ and 1-σ of 2.7 $M_\odot$. This does not mean that the black-hole mass distribution is normal, it is simply a measure of the larger range of masses that the black-hole candidates take. These two distributions were compared using a Kolmogorov-Smirnov test, which rejected the hypothesis that these two distributions are the same at the 99% level. From the histogram of masses (figure 7.1), the neutron star masses seem to be clustered with an apparent Gaussian distribution near 1.4 $M_\odot$, whereas the black-hole masses show no obvious departure from a uniform distribution.

The neutron star and black-hole distributions also have a possible gap dividing them, between 2 $M_\odot$ and 3 $M_\odot$. This is most striking if it is considered that many neutron star equations of state predict stable configurations up to 3 $M_\odot$ (Salgado et al., 1994), and so the existence of this gap may provide clues as to which equations of state are most physically realistic. If the neutron star distribution is assumed to be uniform, from 1 – 3 $M_\odot$, then the probability of a given system not falling in the gap is equal to 0.5. Given the fact that there are 18 systems this gives a probability that no neutron stars would fall in this gap of 3.8×10^{-6}. The probability that any gap would occur with a width of 1 $M_\odot$ is 3.8×10^{-5}, for the 0.1 $M_\odot$ bin width used here. Thus, the clustering of the neutron star masses and absence of any in the gap is significant. The same calculation for the 8 black-hole masses, taking the same gap and a range of masses from 2 – 12 $M_\odot$, gives a probability of 0.43 that the gap exists. This means that it is not possible to rule out the fact that the black-hole distribution extends down to a mass of $\sim$ 2 $M_\odot$. However, this may be due to the fact that there are only 8 measured masses to date and so the significance of this calculation is very uncertain.

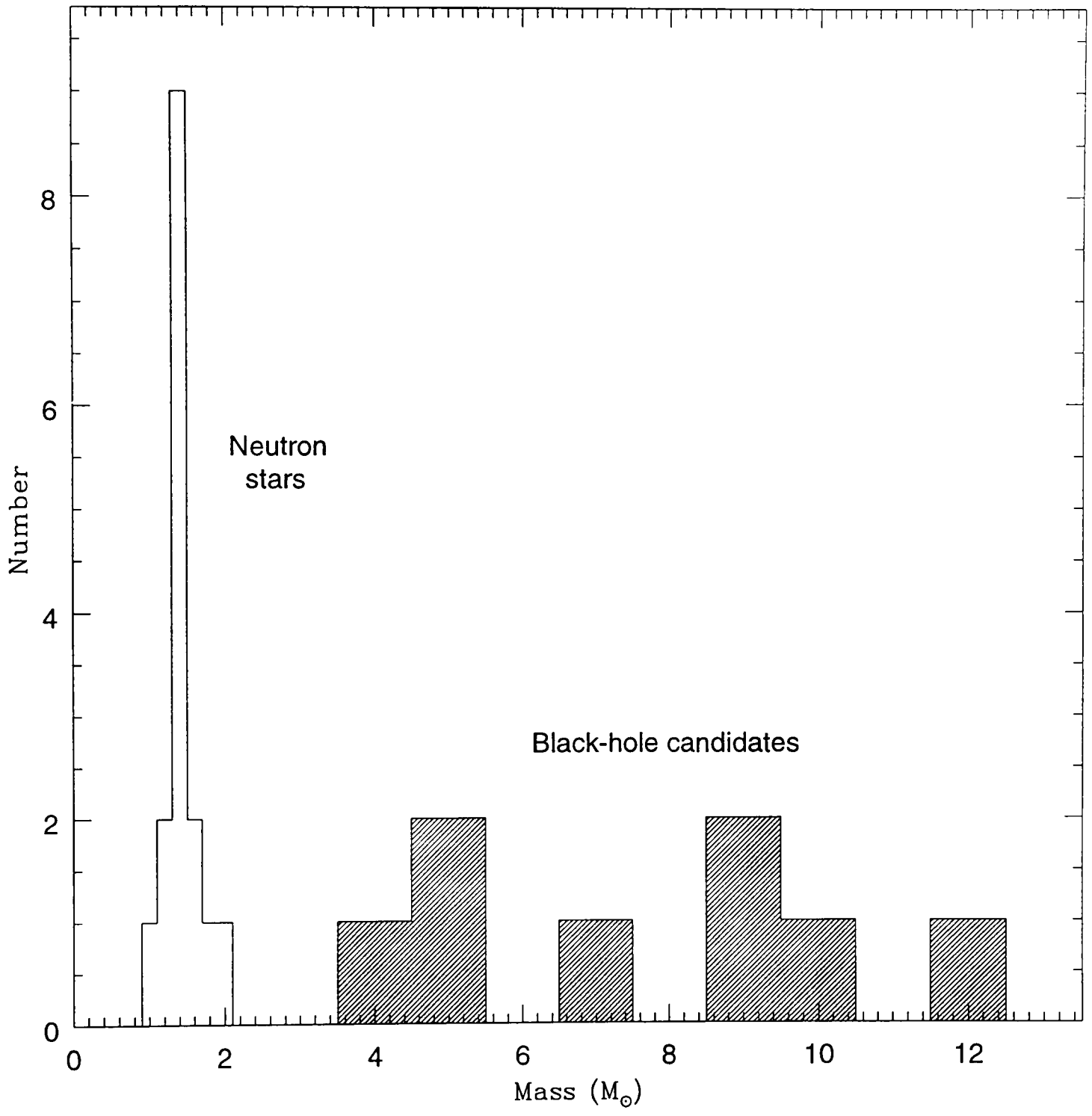


Figure 7.1: The black-hole (shaded) and neutron star (unshaded) mass distributions. The 18 neutron star masses were divided into bins of width 0.2 $M_\odot$ and the 8 black-hole masses were placed in bins of width 1.0 $M_\odot$. The neutron star masses cluster near 1.4 $M_\odot$, whereas the black-hole candidates have a large spread of masses up to 12 $M_\odot$.

Table 7.1: Neutron star and pulsar mass determinations.

Object (P – pulsar, C – companion)	Mass ($M_\odot$)	Error ($M_\odot$)
Cen X-3	1.06	0.56
Cen X-4	0.5 – 2.1	–
LMC X-4	1.38	0.25
SMC X-1	1.06	0.3
Vela X-1	1.77	0.21
Her X-1	0.98	0.12
4U 1538−52	1.3	0.2
4U 1700−37	1.8	0.4
PSR 1913+16 (P)	1.442	0.003
PSR 1913+16 (C)	1.386	0.003
PSR 1534+12 (P)	1.32	0.03
PSR 1534+12 (C)	1.36	0.03
PSR 1855+09 (P)	1.27	0.23
PSR 1802−07 (P)	1.4	0.4
PSR 2303+46 (P)	1.16	0.28
PSR 2303+46 (C)	1.37	0.24
PSR 2127+11c (P)	1.34	0.2
PSR 2127+11c (C)	1.37	0.2

See van Paradijs and McClintock J.E., 1995;
Nagase F., 1989;
Thorsett et al., 1993 and references therein.

Table 7.2: Black-hole candidate mass determinations.

Object	Mass ($M_\odot$)	Error ($M_\odot$)	Reference
Cyg X-1	9	$^{+7}_{-2}$	(1)
LMC X-3	~ 9	–	(1)
J0422+32	~ 5	–	(2)
A0620−00	10	$^{+7}_{-2}$	(1)
Nova Mus 91	~ 4	–	(1)
Nova Sco 94	3.8 – 4.6	–	(3)
GS 2000+25	7 – 8	–	(4)
V404 Cyg	12	$^{+3}_{-2}$	(1)

(1) van Paradijs and McClintock J.E., 1995.
(2) Chapter 4.
(3) Bailyn et al., 1995b.
(4) Casares et al., 1995.

Thus, the gap in the neutron star mass distribution above 2 $M_\odot$ is significant, as is the clustering around 1.4 $M_\odot$. This clustering is physically important, as it occurs at the Chandrasekhar mass. If the neutron star forms by accretion-induced collapse of a degenerate stellar core then this is expected to be the subsequent mass. However, the gravitational (observed) mass of a neutron star can be less than 1.4 $M_\odot$, because the gravitational mass is equal to the baryonic mass minus the binding energy (radiated away when the progenitor star initially collapses). The gravitational mass (i.e. the observed mass) is typically 10% to 15% less than the baryonic mass (Burrows and Woosley, 1986), and so there are a range of values allowed near 1.4 $M_\odot$, with a scatter of $\sim$ 0.2 $M_\odot$. It should be noted that the scatter in the neutron star values is partly due to the errors of measurement and partly due to intrinsic variation in the neutron star masses. Also, there is a large range in the observed errors.

The Inclination Distribution of SXTs

The inclinations of 6 SXT systems are now known from ellipsoidal studies in this work and elsewhere. A list of these values is shown in table 7.3, as well as the relevant list of references. The inclinations measured to date all lie in the range 30° – 60°. Thus, there appears to be a selection effect biasing against the discovery of extreme inclination systems. It must be noted that this does not imply that these systems are less common, necessarily, only that no inclinations have been measured for these systems. However, a possible eclipsing system has recently been discovered (Nova Scorpius; Bailyn et al., 1995a), which should therefore have an inclination larger than 70°, although this has not yet been measured independently. Also, GS2000+251 is thought to have an inclination larger than 67.5° (Chevalier and Ilovaisky, 1993), from the amplitude of the ellipsoidal, optical modulation. But, it must be less than 80°, because of the absence of X-ray eclipses during the 1988 outburst. Therefore, the absence of low inclination systems is more significant than the paucity of high inclination SXTs.

To explain the absence of measured low inclination binaries, consider the orientation in 3-D space of the orbital plane (see figure 7.2). Assuming that SXTs form with a random orientation in 3-D space, it is possible to calculate the distribution of inclinations (with respect to our line

Table 7.3: The measured inclinations of SXTs to date.

Object	Inclination (most probable)	Reference
A0620−00	37°	(1)
Cen X-4	43°	(2)
V404 Cyg	58°	(3)
J0422+32	45°	(3)
Nova Oph 1977	50°	(3)
BW Cir	$\approx 63°$	(3)

(1) Shahbaz et al. (1994).
(2) Shahbaz et al. (1993).
(3) Chapters 3,4 and 5.

of sight) that would be expected. The orientation of the binary plane is randomly distributed and so a sphere enclosing the binary should have an equal number of binaries oriented with the orbital axis pointing towards each unit surface element of that sphere. To calculate the inclination distribution it is necessary to integrate around the sphere to calculate the area in each inclination range that we observe. This integral is shown below:-

$$\begin{aligned} A &= \int_{i_1}^{i_2} 4\pi r^2 \sin\phi \, d\phi \\ &= 4\pi r^2 \left[\cos(i_1) - \cos(i_2)\right] \end{aligned} \tag{7.1}$$

The inclination distribution obtained in this way, biases toward high inclination systems, as they have more surface area per inclination range than lower inclination systems. This is consistent with the fact that few of these are observed, because they are less likely to occur. For example, 1.5% of systems should lie between 0° – 10° and 17% should have inclinations between 80° – 90°.

The fact that, at present, few high inclination systems have measured inclinations could be due to some other effect. The area projected onto the line of sight by the disc in SXTs is proportional to cos$i$. Thus, in outburst but less so in quiescence, the discs of high inclination will appear fainter to the observer (in the X-ray as well as optical bands). Also, X-rays will

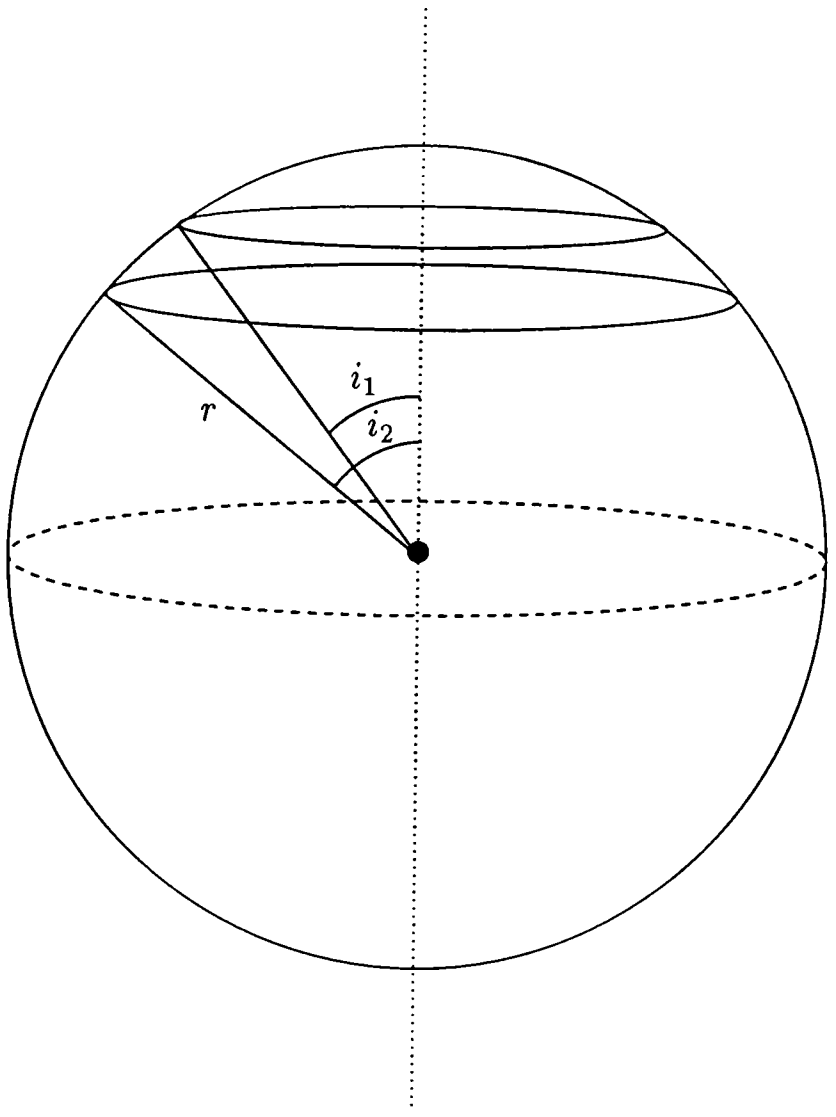


Figure 7.2: The orientation of each binary's orbital axis points toward a given position on the sphere enclosing the binary. To obtain the number of binaries with inclinations, with respect to the line of sight (dotted line), between i_1 and i_2 one simply integrates over the surface area on the sphere in the inclination range as shown.

Table 7.4: The estimated and observed inclination distribution of known SXTs to date.

Inclination Range (degrees)	Predicted Fraction of SXTs	Number observed
0 – 10	0.030	0
10 – 20	0.086	0
20 – 30	0.133	0
30 – 40	0.163	1
40 – 50	0.173	3
50 – 60	0.164	1
60 – 70	0.133	1
70 – 80	0.088	0
80 – 90	0.030	0

be scattered more from the line of sight at higher inclinations making the initial detection of high inclination SXTs more difficult. Disc flaring may also absorb most of the X-rays at high inclinations (Milgrom, 1978). These observational cutoffs to the observation of SXTs will have a complex effect on the distribution of inclinations obtained, but as an approximate guide this factor is taken to be proportional to cos$i$. Combining this with the angular distribution above gives the proportion of SXTs in each inclination range (the total is scaled to 1) shown in table 7.4.

The predicted distribution roughly agrees with that observed, showing that these selection biases are occurring in the measured inclinations (although a possible eclipsing system, Nova Scorpius, has recently been discovered, see Bailyn et al., 1995b). The distribution is perhaps even more sharply centred around the middle range of inclinations than was estimated from the simple analysis described previously. This could be due to additional selection effects such as the fact that extremely low inclination systems below 30° will only exhibit a very small orbital modulation ($< 0.10^m$), which means that period detection would be harder for such systems. Also, the CV inclination distribution, which has many more measured values (Shahbaz, private communication), has few low inclination systems in agreement with the above discussion, but does not have a cutoff until 80° or more. This may indicate that the high inclination systems are not as scarce as argued above.

Table 7.5: The noise levels (in magnitudes) of several SXT light curves.

Object	Noise (1-σ)	Photometric Noise	Residual Noise
V404 Cyg (R)	0.053	0.016	0.051
V404 Cyg (Extended I)	0.035	0.008	0.034
J0422+32	0.10	0.07	0.071
Nova Oph 77	0.049	0.039	0.029

The inclination and mass distributions yield insights into the physics underlying these objects and the limitations of the methods of observing them (e.g sensitivity limits, and constraints of finite observing time). The ellipsoidal studies presented here also enable other types of variability in SXTs to be examined. For those systems with known periods it is possible to measure the noise variance of the data after removal of the orbital variation. This noise level can then be compared to the expected, photometric noise. This is shown in table 7.5, along with the residual noise level, which is equal to the noise variance minus the photometric variance.

Table 7.5 shows that, consistently, there is a residual level of noise left in the light curve which cannot be accounted for by known variations or measurement errors. This is presumably due to random flickering of the accretion disc, and possibly other hot components in the binary. Evidence for this was obtained in Chapter 3, where the BVRI light curves of V404 Cyg showed an increase in the random noise toward the blue, implying that this effect arises in the disc.

The level of noise observed is a few percent and so could be due to the low frequency analogue of quasi-periodic oscillations (QPOs). These occur in many LMXBs in X-ray light curves at frequencies between 0.1 Hz and 100 Hz, but some systems show no cutoff below frequencies of $\sim 10^{-3}$ Hz (Langmeier et al., 1990). This is thought to be due to variations in the accretion rate, and so if the X-ray output of the compact object is varying, then the optical flux of the disc should also vary. The observations taken here had integration times of ~ 600 s and so QPO variations at 10^{-3} Hz could have caused erratic variations in the light curve, as was observed.

References

Bailyn C.D. et al., 1995a, Nature, **374**, 701.

Bailyn C.D., Orosz J.A., McClintock J.E. and Remillard R.A., 1995b, IAU Circ., No. 6173.

Burrows A. and Woosley S.E., 1986, Ap. J., **308**, 680.

Casares J., Charles P.A. and Marsh T.R., 1995, MNRAS, in press.

Chevalier C. and Ilovaisky S.A., 1993, Astr. Astrophys., **269**, 301.

Langmeier A., Hasinger G. and Trümper J., 1990, Astr. Astrophys., **228**, 89.

Milgrom M., 1978, Astr. Astrophys., **67**, L25.

Nagase F., 1989, PASJ, **41**, 1.

Salgado M, Bonnazola S., Gourgoulhon E. and Haensel P., 1994, Astron. Astrophys. Suppl. Ser., **108**, 455.

Shahbaz T., Naylor T. and Charles P.A., 1993, MNRAS, **265**, 655.

Shahbaz T., Naylor T. and Charles P.A., 1994, MNRAS, **268**, 756.

Thorsett S.E., Arzoumanian Z., McKinnon M.M. and Taylor J.H., 1993, Ap. J., **405**, L29.

van Paradijs J. and McClintock J.E., 1995, in *X-ray Binaries*, Eds. Lewin W.H.G, van Paradijs J. and van den Heuvel E.P.J., Cambridge University Press, 101.

Appendix

Model Comparison

In order to analyze the data using different models, it is important to be able to compare these models to see which fit the data better. In terms of probabilities what is needed is the probability of model f_j relative to model f_i (for a detailed description of model comparison methods using Bayesian statistics see Bretthorst, 1990). The probability of model f_j is given by:

$$P(f_j|D,I) = \frac{P(f_j|I)P(D|f_j,I)}{P(D|I)}$$

The relative probabilities of f_j and f_i are,

$$\frac{P(f_j|D,I)}{P(f_i|D,I)} = \frac{P(f_j|I)}{P(f_i|I)}\frac{P(D|f_j,I)}{P(D|f_i,I)}$$

Thus the factor $P(D|I)$ can be ignored in this calculation as it is a normalization constant, and cancels out when two different models are compared. It is assumed that there is no prior information about the model parameters (so that $P(f_j|I)$=1). Thus, all that needs to be calculated is the ratio of $P(D|f_j,I)$ to $P(D|f_i,I)$. To calculate this it is assumed that the noise level is known (this can be marginalized from the calculation later). The probability can be separated as follows:

$$P(D|\sigma,f_j,I) = \int P(A,\omega|I)P(D|A,\omega,\sigma,f_j,I)dAd\omega \qquad \text{(A.1)}$$

The most conservative assumption about the noise is that it is Gaussian and this gives the direct probability:

$$P(D|A,\omega,\sigma,f_j,I) \;=\; (2\pi\sigma^2)^{-N/2}\;\; exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N}[d_i - f_j(t_i)]^2\right) \qquad \text{(A.2)}$$

The data d_i are modelled by the functions $f_j(t_i)$ over a set of N points. The jth model consists of a series of $r = j$ sinusoids with amplitudes A:

$$f_j(t_i) = \sum_{l=1}^{r} A_l sin(\omega_l t_i)$$

The function is a sine for odd l, and cosine for even l:

$$f_j(t_i) = \sum_{l=1}^{r} A_l cos(\omega_l t_i)$$

The model is substituted into equation (A.2) to give:

$$P(D|A,\omega,\sigma,f_j,I) = (2\pi\sigma^2)^{(-N/2)} exp\left(\frac{-NQ}{2\sigma^2}\right)$$

where,

$$Q = \overline{d^2} - 2/N\sum_{l=1}^{r} A_l h_l + 1/N\sum_{l=1}^{r} A_l^2$$

and,

$$\overline{d^2} = 1/N\sum_{i=1}^{N} d(t_i)^2$$

$$h_l = \sum_{i=1}^{N} d_i H_l(t_i)$$

Again the function H_l is a sine for odd l and cosine for l. The other factor in the integral of equation (A.1) can be obtained by assuming a conservative Gaussian distribution for our state of knowledge about the amplitude A and frequencies ω:

$$P(A,\omega|I) = P(A|I)P(\omega|I)$$

$$P(A,\omega|I) \;=\; (2\pi\delta^2)^{-r/2} exp\left(-\sum_{l=1}^{r}\frac{A_l^2}{2\delta^2}\right)\;\; (2\pi\gamma^2)^{-r/2} exp\left(-\sum_{l=1}^{r}\frac{\omega_l^2}{2\gamma^2}\right) \qquad \text{(A.3)}$$

The parameters γ and δ determine the uncertainty in the values of the frequencies and amplitudes, respectively. These can be removed by integration in the next step. Substituting equations (A.2) and (A.3) into the integral for the direct probability (equation (A.1), and then integrating over the frequencies and amplitudes gives:

$$P(D|f_j,\sigma,I) = \frac{\Gamma(m/2)}{2log(R_\delta)}\left[\frac{m\overline{h^2}(\hat{\omega})}{2}\right]^{-m/2}\frac{\Gamma(r/2)}{2log(R_\gamma)}\left[\frac{r\overline{\omega^2}}{2}\right]^{-r/2}$$

$$\times \quad {\nu_1}^{-1/2}\ldots{\nu_r}^{-1/2}(2\pi\sigma^2)^{\frac{N-m-r}{2}}$$

$$\times \quad exp\left[\frac{N\overline{d^2} - m\overline{h^2}(\{\hat{\omega}\})}{2}\right] \tag{A.4}$$

This assumes that the level of noise in the data is known. This can normally be estimated from photometric errors, and so this form of the equation was used here. For this analysis, m is the number of sinusoids or frequencies for a given model (r=$2m$); $\overline{h^2}(\hat{\omega})$ is the value of $\overline{h^2}(\omega)$ $(= 1/m\sum_{l=1}^{m} h_l^2)$ taken at the peak of the probability distribution and $\overline{\omega^2} = 1/m\sum_{l=1}^{m}\omega_l$. The function Γ is the gamma function, and ν_r are the eigenvalues of the matrix given by:

$$b_{kl} = -\frac{m}{2}\frac{\partial^2\overline{h^2}}{\partial\omega_k\partial\omega_l}$$

R_δ, R_γ and R_ω are constants which give a measure of the uncertainty in the amplitude of the signal, its frequency and the noise present in the data respectively. These cancel out when two models are compared and so can be ignored for model comparison methods.

What is normally needed for time series analysis is to compare the jth model with the $(j+1)$th model, to determine the probability that there is one more sinusoid present in the data. The above equation gives the probability of the two models and so it is the ratio of the two, i.e. $P(D|f_{j+1},I)/P(D|f_j,I)$ that is calculated. This gives the probability (known as the odds ratio) that the $(j+1)$th model is correct compared to the jth model. Thus, successively more complex models can be tested against the data. The above equation contains several factors to penalize more complex models, and so will only predict that a model is correct if it fits the data significantly better than more simple models.

References

Bretthorst G.L., 1990, J. Magn. Reson., **88**, 552.