A SURVEY FOR EXTRASOLAR PLANETS WITH THE METHOD OF TRANSITS

Ioannis Antoniadis

Aristotle University of Thessaloniki
Department of Physics
Section of Astrophysics, Astronomy and Mechanics
DIPLOMA THESIS
· 2009 ·

A SURVEY FOR EXTRASOLAR PLANETS WITH THE METHOD OF TRANSITS

Ioannis Antoniadis

Supervisor: Proffesor John H. Seiradakis

Aristotle University of Thessaloniki
Department of Physics
Section of Astrophysics, Astronomy and Mechanics

In loving memory of my grandfather,
Moschos Keramitsoglou

Abstract

This Diploma Thesis is divided in five chapters.

The first chapter is a brief overview of the history and evolution of extrasolar planetary science and the methods that have been used so far for the detection and characterization of extrasolar planets.

In the second chapter, I describe the aspects of transit modeling. A new method for calculating the Transit Detection Probability and locating the promising spots for observations on the celestial plane is also introduced.

The third chapter of this dissertation is devoted to the actual design of the transit survey. In order to make a system sufficient to detect transits of Hot Jupiter planets, several tests have been made with different CCD cameras and telescopes. The selection of target fields is also described.

In the fourth chapter, a detailed overview of the data reduction pipeline that have been developed is presented. Because of the tiny variations we are trying to detect, a very precise correction of systematic effects is needed. For this purpose algorithms like the Trend Filtering Algorithm *(Kovàks et al, 2005)* and Sys-Rem *(Tamuz et al 2005)* are implemented and tested together with self developed methods. After deriving the light curves, the Box *Fitting Least Squares* algorithm *(Kovàks 2004)* and the self developed *Fast Decetion Algorithm* are used in search of possible transit-like futures inside the light curves . Algorithms for the detection of periodical signals are also utilized in search of variable stars in the field. The chapter also contains the results of the tests of the pipeline's performance on various kinds of data.

The last chapter contains the actual observations and the results of the analysis.

Contents

List of Figures

List of Tables

Acknowledgments

There are several people without whom you would not be reading this Thesis.

First of all I would like to thank my Father, Mother and Grandmother for their psychological and financial help. Without them it would have been impossible to study at first place.

Of course I own a lot to my Supervisor Professor John H. Seiradakis (Aristotle University of Thessaloniki-AUTH) for his guidance, his vital ideas on this project and for giving the chance to an undergraduate student to work with professional equipment on real research projects. Most of all I would like to thank him for his kindness and his tolerance on my mistakes and for always being there when he was needed.

In terms of technical knowledge and support, the help of Evangelos Tsorlinis (AUTH) and George Dimou (National Observatory of Athens-NOA) was crucial. Special thanks to them.

Also, many thanks to the stuff of the Department of Forestry and Natural Environment (DFNE) at Mt. Holomon, Chalkidiki, and especially to its Director, Mr Panourias George and its curator, Papanikolaou Christos for their trust and hospitality.

Maybe a simple acknowledgment is not enough for my friend and former fellow student Dimitris Mislis (Hamburg Observatory). I feel deep gratitute for him, first of all for getting me involved with extrasolar planets, for the countless hours of discussions, for checking my pipeline and most of all for giving me courage to continue.

To my friends, John Nestoras, Lilia Tremou, Stelios Pyrzas, Vassilis Karamanavis and all the others (you know who you are) thank you for supporting me.

To my girlfriend Emmanouela Athanasiou for not saying anything about spending most of my time in front of a computer and not with her, I give my love.

Last, but not least , I would like to thank my grandfather for sharing his wisdom with me. This Diploma Thesis is devoted to him. May he rest in peace.

Nomenclature

This Thesis employs a number of standard astronomical terms, symbols and units. For the continuity of the text, the most frequent are summarised bellow:

- $\oplus$ = Earth, $\odot$ = Sun, jup = Jupiter
- M = Mass, ESP = Extrasolar Planet
- $M_{\odot} = 2.0 \times 10^{30} kg$
- $M_{jup} = 9.5 \times 10^{-4} M_{\odot}$
- $M_{\oplus} = 3 \times 10^{-3} M_{jup}$
- 1 AU = 1 astronomical unit (mean Sun-Earth distance) $\simeq 1.5 \times 10^{11} m$

With α we will symbolize the semi-major axis of an orbit and with P the period. Stellar distances will be given in parsec (pc), defined as the distance at wich 1 AU subtents an angle of 1 second of arc. $1 pc \simeq 3.1 \times 10^{26} m \simeq 3.25 ly$

Star Names will be given with the standard nomenclature adopted by the astronomical community:

- **Constellation-based names**: β Andromedae or β And , is the second brightest star of the constellation Andromedae. γ Virginis or γ Vir is the third brightest star of Virginis. 70 Taurus is the 70th brightest star of Taurus, etc.
- **Spectral classification**: The Sun is of spectral type G2V. Cool stars (late type stars, G, K, M) are also refered as red dwarfs and have smaller masses.
- **Standard catalogue name**: *HD 112345* is the 112345th star of the Henry Drapper catalogue etc
- **Celestial Coordinates**: e.g $PSR1233 + 11$, $PSR2212 - 12$ etc
- **Planetary companions**: *WASP 12a* is the 12th star found to have a planet by the WASP survey. *WASP 12b* is the planetary companion. If a second planet would be discovered in the system it will be named *WASP 12c*, etc

By the term *Hot Jupiter* we mean all the gaseous planets with $M > M_{jup}$ and α of a few AU

Two possibilities exist: Either we are alone in the Universe or we are not. Both are equally terrifying!

Arthure Clarke

1 The star and its planetary system

1.1 A brief history of planetary searches

Even though the concept of our Earth as being the center of the universe was bounded to peoples' believes until the times of *Kepler and Galileo*, the actual idea of planets orbiting other stars is very old. Many philosophers had cosmological ideas highly advanced for the ancient world. *Simplicius of Cilicia* (490 AD -560 AD) in his *Physica Auscultatio* is revealing that *Anaxagoras* (500 BC) believed that *'The worlds that exist are infinite'*. *Metrodoros of Chios* (400 BC) was saying that *'A single ear of corn in a large field is as strange as a single world in infinite space'*. In the scripts of *Epicuros* (341 BC - 270 BC) we read that *'The Universe is infinite. We are not the center of Cosmos but only one of the infinite worlds that exist'*. We find similar ideas to the writings of several other Ionian Philosophers like Hypolitos, Leukipos and Dimokritos. At the other side of the world and many centuries later, *Teng Mu* (13th century A.D), a Chinese scholar of the Sung Dynasty adopted a position similar to that of Ancient Greeks. He was claiming that *'Empty space is like a kingdom, and earth and sky are no more than a single individual person in that kingdom. Upon one tree are many fruits, and in one kingdom there are many people. How unreasonable it would be to suppose that, besides the earth and the sky which we can see, there are no other skies and no other earths.'*.

These discerning ideas, were deeply challenged by luminaries such as Aristotle and been forgotten for centuries.

In terms of real scientific formalism as we mean it today, the first theory of planetary formation originated with *Pierre Simon de Laplace* (1784) and *Immanuel Kant* (1798). In their theories they introduced for the first time the idea that our planetary system formed from a gaseous nebula that has collapsed under its own gravity into a compact flat disk rotating around the Sun. Planets have formed by the split of this protoplanetary disk into rings similar to those of Saturn. This theory was a real breakthrough for the scientific standards of the time and it survived undoubted for more than a century. The first to question Kant's and Laplace's model was *Moulton* (1900) who suggested that conservation of angular momentum should result to a star at the center of the disk rotating far more rapidly than we observe at the Sun. He proposed that planets may have formed from gravitational instabilities introduced by a passing star. These first ideas continued to evolve through time (*Weizsacker 1944; Edgeworth 1949; Safronov & Smith 1969; Weidenschilling 1976; Cameron 1978; Goldreich & Tremaine 1980*) to lead to what we believe today.

After years of pure theory the first observational evidence of a protoplanetary disk came from the Satellite IRAS which discovered a dust disk around the star Vega at the constellation of Lyra (*Harvey, Wilking and Joy 1984*) [23] . But where are the planets? Suggestions of planets orbiting stars different than the Sun are older. Some of the first claims came from observations of the double eclipsing system *70 Ophiuchi*. Orbital anomalies that were observed (*Capt W.S Jacob 1855- East India Company's Mandras Observatory & Thomas J.J See 1890 -University of Chicago*) were attributed to a dark planetary body orbiting the system with a period of 36 years. However a paper published by *Forest Ray Moulton* proved that a three-body system with the observable characteristics of *70 Ophiuchi* would be highly unstable.

During the 1950s and 1960s, *Peter van de Kamp* (Swarthmore College) published a series of papers in which he postulated that he had found a planetary system orbiting *Barnard's Star*. However, the use of different and old equipment that he utilized for his astrometric observations, make these claims hard to believe and this planetary system has not yet been confirmed.

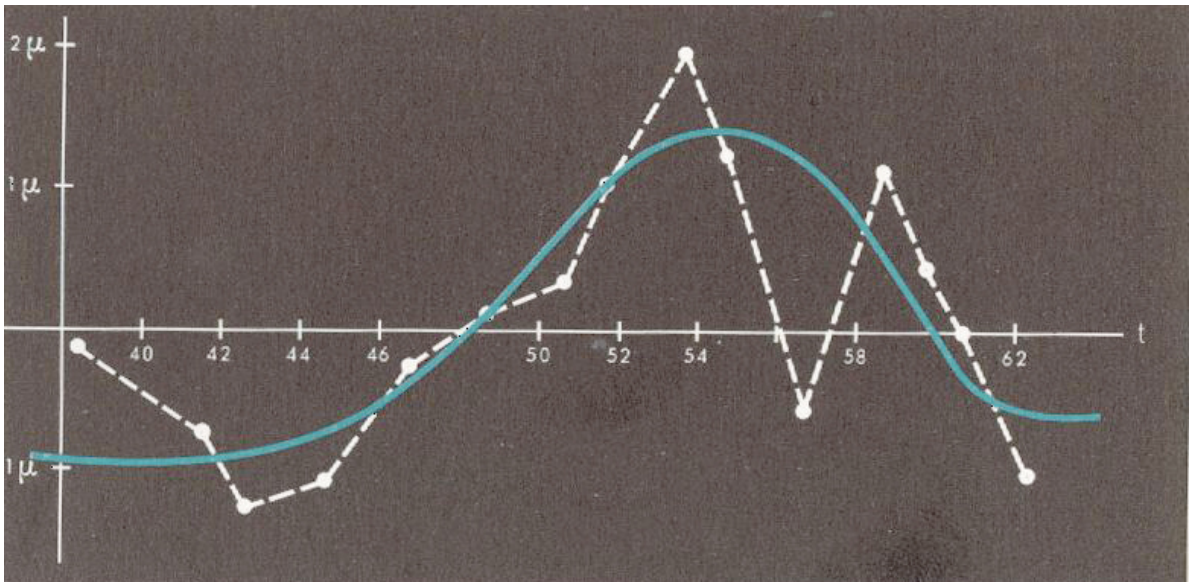

Figure 1.1: de Kamp's observations and a theoretical curve of a planet with $M = 1.6M_J$, $a = 4.42AU$ and $e = 0.77$. Ollivier (2008). School on Corot Astrophysics. Observatoìre de Parìs

The first unquestioned discovery of an extrasolar planet came from the Radio astronomy community in 1992. *Aleksander Wolszczan* and *Dale Frail* using the pulsar timing technique discovered two planets of mass $M = 2.8M_\oplus$ and $M = 3.4M_\oplus$ orbiting the milisecond pulsar $PSR1257 + 12$ at orbital distances of 0.47 and 0.36 AU (*1992 Nature.355..145W*) [44] . That discovery opened a new era in Observational Astronomy. It became pretty obvious that if planets can survive the violent environment of a neutron star they should surround also main sequence stars.

The confirmation of this hypothesis came in October 6, 1995 from *Michel Mayor & Didier Queloz* [30] (Geneva University). In a historical report published by *Nature* they announced their findings of radial velocity measurements on the star *51 Pegasi*. Their observations revealed a planet of a Mass $0.47/sini$ times the mass of Jupiter orbiting the star at a distance of a few AU. Later it was predicted that a planet of this size must be a gaseous giant and not a supersized rocky planet (*Burrows et al* 2000)**?** .

Announcements of other groups using the same technique followed that first discovery. Most important we should mention the first discovery of a planet orbiting a multiple star system (*Cochran et al* 1997) and the first announcement of a multiple planetary system orbiting the star υ Andromedae (*Butler et al* 1998).

In 1999 came the first discovery of a transit event of an extrasolar planet. The transit was observed at the star HD209458. Previous radial velocity observations showed that a Hot Jupiter planet is orbiting the star with a period of 3.5 days (*Marcy et al* 2000). Once the orbital parameters were established it was obvious that HD209458 was a promising candidate for giving a transit event. *Henry et al* (2000) [24] and *Charbonneau et al* (2000) [14] published in the same volume both the radial velocity results and two full transits of the planet. *Charbonneau et al* (2000) performed also a theoretical fit to the data and for the first time derived the true mass of the planet (0.63 Jupiter Masses with a radius of 1.27 R_{jup}).

Since the discovery of HD209458 the list of transiting planets is rapidly growing[1]. Astronomers realized that they had a very powerful tool in their hands that could allow them to fully characterize the orbital parameters of the transiting planet (*see chapter 2 for details*) .This fact sparked off a "rush" of observational effort and several groups have been created to search for transits. Among them the most successful, till now, are the SuperWASP team *(Super Wide Angle Search for Planets, Cameron et al)*, with 15 confirmed candidates, the HAT network (Hungarian-Made Automated Telescope,*Bakos et al* with 10 confirmed candidates and the TrES survey (TransAtlantic Exoplanet Survey, *Serio et al*) with 4 confirmed candidates.
CoRoT space mission (ESA, launched in 2007) and KEPLER mission (NASA, to be launched in 2009) are expected to increase the number of transiting extrasolar planets to at least one thousand.

The effect of gravitational microlensing, namely the amplification of a light source by the curvature of space caused by an intervening object, first considered by *Einstein & Link* (1936) gave the first planet-mass candidate in 2004. The OGLE team (*Trimble et al.* 2003) discovered the microlensing effects caused by the planet OGLE235-MOA53b (*Trimble V. & Aschwanden M 2005*). That planet has a mass of 1.5 M_{jup} or 2.5 M_{jup}, depending on the lens model and a semi-major axis of $\sim$ 3 $\pm$0.2 AU. Up to date 7 more planets have been discovered with this method, among them an Earth-like planet (OGLE-05-390L) with a mass of 5.8 $\pm 0.1 M_{\oplus}$, semi-major axis of $\sim$ 2.6 AU and orbital period of $\sim$ 3800 days.

The most obvious way to detect a planet of course is to take a direct image of its surface. Although this has not yet been achieved, images of planets seen as point sources have been taken with interferometric and coronographic techniques. 6 planets have been detected

[1] 51 candidates by the end of October 2008

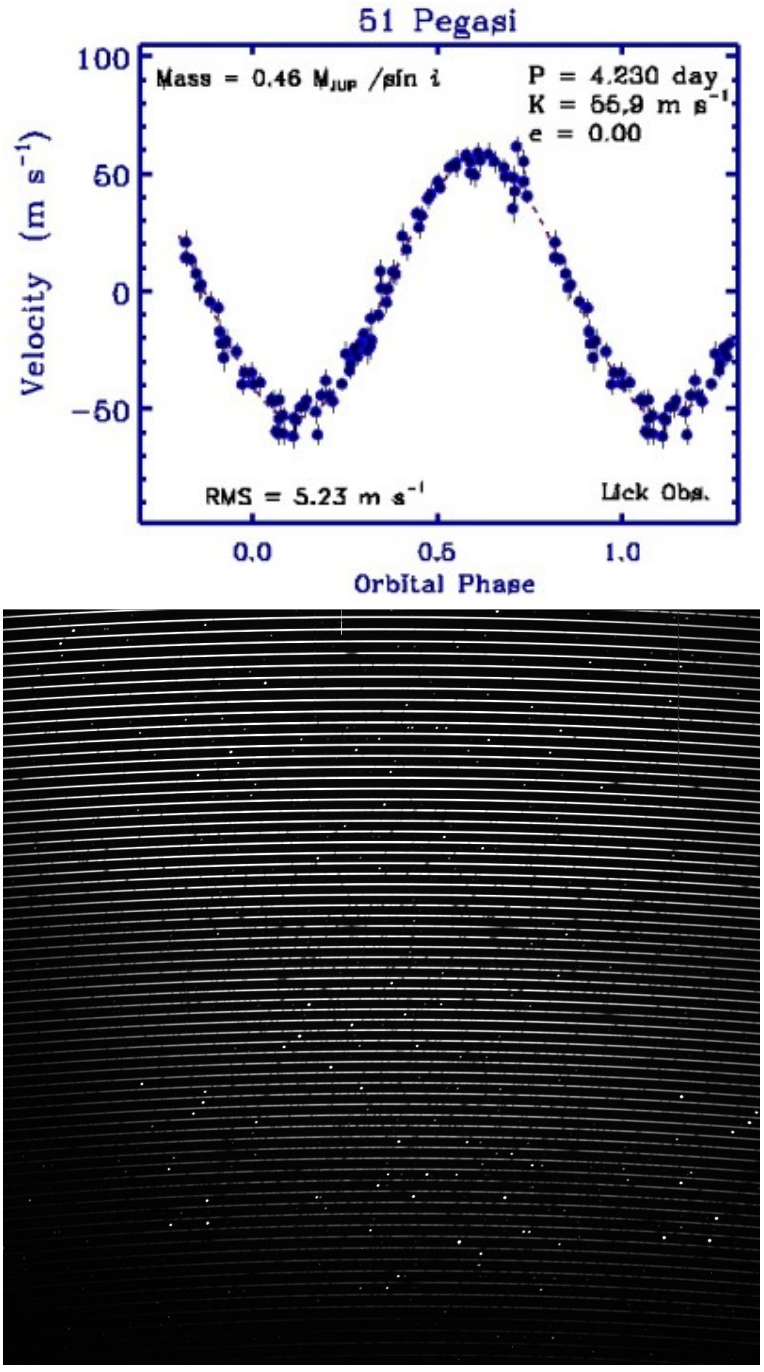

Figure 1.2: Radial velocity plot of 51 Pegasi and an image of it's spectra obtained with the ELODIE instrument

directly till now and prospects for this technique are highly promising.

Up to date *(January 2009)* 313 extrasolar planets have been discovered, of which 295 with radial velocity, 52 with transits, 8 with microlensing and 5 with other methods. Among them we also find 34 multiple systems(*Extrasolar Encyclopedia, Jean Schneider*).

New methods like *timing of periodical systems* and infrared monitoring of the phases of planets orbiting their parent star, are under high consideration over the last few years. These methods might help us to detect more earth-mass planets in orbits similar to that of the Earth. In any case, a deep understanding of planetary systems, including our own, will only be achieved by observing and characterizing a wide range of planetary systems orbiting different stellar environments.

1.2 Detection Methods

In this section, methods implemented for the detection and characterization of extrasolar planets (from now on ESPs) are summarized. But why there are more than one method at the first place? Why can't we just point our telescopes at an ESP and take an image of its surface? To answer this question we have to bare in mind the following: ESPs are small, faint objects placed at vast distances from us. A typical Hot Jupiter orbiting a star 10 pc away, has an angular separation of $\backsim$ 0.005 seconds of an arc for an observer in our Solar System. If we consider that the minimum angle that a telescope can resolve is given by [41] :

$$\vartheta_{min} = 1.22\frac{\lambda}{D}(arcsec) \tag{1.1}$$

where λ the wavelength of observations (in nm), and D the diameter of the telescope (in m), we realize that we need an instrument of $D \backsim 150m$ just to be able to separate the planet from its star at optical wavelengths! Moreover, atmospheric turbulence, due to high atmospheric layers, limits resolution even more, to $\geq$ 0.2arcseconds[2]. A more detailed comparison of angular separation, together with Star/Planet Contrast and Flux emitted from planets is given in Table 1.1.

PLANET	Separation $@10pc$	Separation $@50pc$	Contrast $@0.5\mu m$	Contrast $@10\mu m$	Flux $@0.3-2\mu m$ $(phs^{-1}m^{-2})$	Flux $@6-20\mu m$ $(phs^{-1}m^{-2})$
Earth	0.1	0.02	$5\ 10^9$	10^3	0.3	10
Jupiter	0.52	0.1	$4\ 10^8$	$5\ 10^7$	3.8	2
Saturn	0.95	0.2	2×10^9	5×10^8	0.75	0.2
Neptune	3.0	0.6	10^{11}	10^{10}	0.015	0.01
Hot Jupiter	0.005	0.001	4	10^3	38000	100000

Table 1.1: Characteristic quantities of known planets, as seen by a hypothetical distant observer. For comparisson, the same numbers for a Hot Jupiter are also given

From the facts described above it is obvious that direct imaging of ESPs is extremely difficult and requires sophisticated instrumentation. Efforts pointing toward this direction so far are concentrating in finding ways to reduce the brightness of the star and/or connecting instruments with interferometric techniques in order to increase the angular resolution of observations. With current telescope technology, images of planets have been taken only under exceptional circumstances. A planet has to be far separated from its parent star, large in size and relatively hot, so that its thermal radiation can be observed at the infrared regime.

Planets, however interfere in several ways with their stars. These perturbations can be detected by indirect observational methods.

[2]The technology of *Adaptive optics* allows us to minimize these effects, however "seeing-free" measurements can be made only from space

We can divide these techniques into two major groups:

- **Methods for the detection of the dynamical perturbation of the parent star, caused by the planet:**

  It is known, from celestial mechanics that two masses interacting by gravity are moving around their common center of mass. If m_* is the mass of the star and m_p the mass of the planet, the center of mass is given by:

  $$R_{mc} = \frac{m_* R_* + m_p R_p}{m_* + m_p} \tag{1.2}$$

  and the radius of stellar orbit is

  $$r_* = \frac{m_*}{m_* + m_p} S\,P \tag{1.3}$$

  where SP is the stellar planet distance. Methods resorted to observe this stellar motion are:

  1. Astrometry
  2. Radial velocity technique
  3. Pulsar timing
  4. Eclipsing binary minima timing

- **Methods for the detection of photometric variations caused by the orbiting planet:** ESPs may cause variations in star's light in many ways. The most common scenario is that a planet can dim the light of the star by hiding a fraction of its surface as it orbits. Moreover, variations in planet's temperature can be detected at infrared wavelengths and in particular cases, the reemitted stellar light from circumstellar disks can indicate the presence of small objects, at the same spectral regime. Photometry in narrow spectral lines is also implemented for the detection and characterization of planetary atmospheres. In addition, atmospheric molecules could change the polarization of stellar light. Observations of this phenomenon are favored by the fact that polarimetry is not limited by the stability of earth's atmosphere. Gravitational interference of two aligned objects (*microlensing effect*) may also reveal the existence of a planet at the foreground object.

  Summarized the methods of this group are the following:

  1. Transit method
  2. Orbital phase reflected light variations
  3. Circumstellar disks
  4. Narrow band photometry
  5. Polarimetry
  6. Gravitation microlensing

A more detailed review of all detection methods is followed in the next sections.

1.2.1 Astrometry

This method consists of measuring the changes of the coordinates of a star produced by the motion (1.3). The angular amplitude of stellar motion as it is projected on the plane of the sky is given by:

$$d\vartheta = \frac{\alpha_*}{D} = \frac{m_p}{m_*}\frac{\alpha_p}{D} \tag{1.4}$$

where α_* is the semi-major axis of the star's orbit, α_p the semi-major axis of the planet, D the Earth-star distance and m_*, m_p, the stellar and planetary masses. From the above relation it is seen that $d\theta$ is proportional to both the orbital radius and the mass of the planet and inversely proportional to the star's distance. Therefore this method is more sensitive in detecting massive planets with large orbits.

In general this "wobble" is of the order of milli arc-second or less and hence very difficult to be detected. Table 1.2 gives characteristic values of $d\vartheta$ for Solar System Planets as seen by observers at various distances and figure 1.3 [9] shows the expected changes in a star's position through time for a hypothetical massive planet.

PLANET	$\alpha_*(r_\odot)$	$d\vartheta@5pc$(mas)	$d\vartheta@10pc$(mas)
Earth	6.5×10^{-4}	6×10^{-4}	$5\,3 \times 10^{-4}$
Jupiter	1.07	1	0.5
Saturn	0.59	0.55	0.28
Hot Jupiter	0.01	0.093	0.046

Table 1.2: Astrometric signature of Solar system planets and a Hot Jupiter. In most cases $d\vartheta$ is smaller than the detection limits of ground based telescopes

Although this method is the oldest, it has not yet yelded confirmed candidates. Early time announcements (see *section 1.1*) are nowdays considered to be false alarms. However, in 2002 the Hubble space telescope successfully detected an already confirmed planet at the *Gliese* planetary system (*Benedict et al* 2002). Future space missions like GAIA (*ESA*) will reach the required astrometric precession and they are expected to give a large number of planetary candidates.

1.2.2 Radial velocity

The goal in this method, is to detect the motion (1.3), this time by measuring variations in stellar velocity. These changes in radial velocity can be measured by the *relativistic doppler effect* (DE). Stellar spectral lines appear either *blue-shifted*, if the star is moving towards us, or *redshifted*, if it grows away (Figure 1.4).

Displacement in spectral lines and velocity are connected via the formula:

$$\frac{\Delta\lambda}{\lambda} = \frac{u}{c}(u \ll c) \tag{1.5}$$

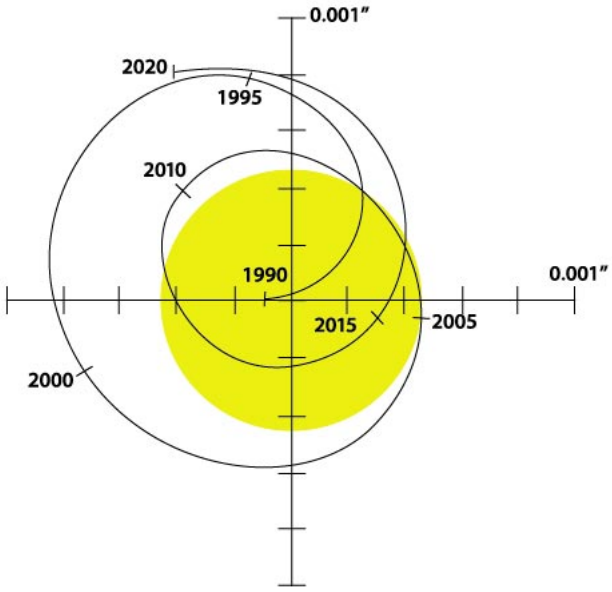

Figure 1.3: Astrometric motion due to the presence of a planet[enersec.org]

The radial velocity amplitude V , for a star of mass m_* orbited by a companion of mass m_p is (*Cumming et al* 1999) [16] :

$$V = \left(\frac{2\pi G}{P}\right)^{\frac{1}{3}} \frac{m_p sini}{\left(m_p + m_*\right)^{\frac{2}{3}}} \frac{1}{(1-e^2)^{\frac{1}{2}}} \tag{1.6}$$

where i is the inclination of the orbit ($i = 0^o$ for face on and $i = 90^o$ for edge on), e the eccentricity and P the orbital period wich is connected with the semi-major axis α by Kepler's third law:

$$P = \left(\frac{\alpha}{1AU}\right)^{\frac{3}{2}} \left(\frac{m_*}{m_\odot}\right)^{-\frac{1}{2}} years \tag{1.7}$$

$m_p sini$ means that the measured amplitude is resulted only by the projected movement along the line of sight. That is why the term *radial velocity* is used. Moreover, radial velocity method can only deliver $m_p sini$ rather than m_p and therefore provide only a minimum value for the planets mass. All values for V are of the order of ms^{-1} as shown in Table 1.3 [9]

Achieving the accuracy, for detecting velocities of that amplitude require high-resolution spectroscopy over a wide range of spectral orders. Usually, with modern instruments, the $\Delta\lambda$ to be measured is of the order of $1/10th$ of a pixel or less and thus, a high signal-to-noise ratio is required. This means that only the brightest stars, down to the magnitude of $\sim$ 13 can be searched with current instrumentation. Another limiting factor is that the usual data analysis procedure requires a big number of well characterized spectral lines. That excludes

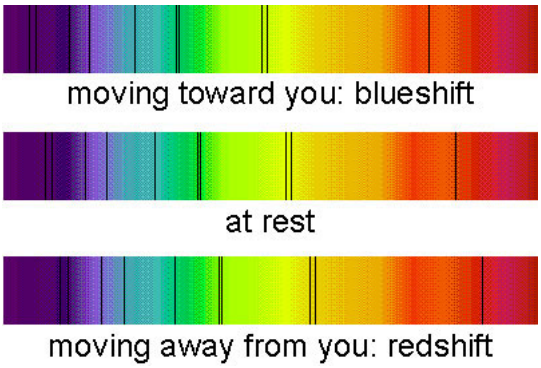

Figure 1.4: Relativistic doppler effect [snaplbl.gov]

PLANET	$m_p(m_\oplus)$	$\alpha_p(AU)$	$\alpha_*(r_\odot)$	T (days)	$V(ms^{-1})$
Jupiter	317.83	5.2	1.07	43326	125
Saturn	95.15	9.54	6.59	10759.2	28
Uranus	14.54	19.18	0.18	30685.4	0.3
51 Peg b	130	0.05	0.004	4.23	50.2
Earth	1	1	0.5×10^{-4}	365.25	0.09

Table 1.3: Comparison of Radial Velocity amplitudes V, for Solar System Planets and 51 Peg b

almost outright fast rotating stars with broadened lines and early-type stars. Moreover, stellar seismic activity (like P-mode oscillations), earths radial velocity, atmospheric conditions and telescope guiding errors among other factors, introduce biases to the observations.

Despite the described difficulties, this method is the most successful so far. Most of the ESP's known today (more than 313) have been detected using radial velocity techniques.

1.2.3 Pulsar timing

Pulsars are highly magnetized, fast rotating neutron stars which emit narrow beams of radio-waves parallel to their magnetic axis. The observational fingerprint of pulsars is an intensive periodic radio-beam seen due to rotational and magnetic axis misalignment. Most of the known pulsars have periods of $\sim 1sec$, but there is a special group with periods of the order of milliseconds. These millisecond pulsars are believed to be refueled old neutron stars that have been spun-up by mass and angular momentum transfer from a stellar companion. Their frequency is very accurate and only spin down at a rate of $\sim 10^{-19} ss^{-1}$ (*Bailes* 1996).

If a planetary companion is present, then the motion (1.3) will cause a periodic increase/decrease in the distance light have to travel. This will lead in timing residuals in the arriving of pulsar beams. The amplitude of these residuals was calculated by *Wolszczan*

(1992) [44] to be:

$$\tau_p = 1.2\left(\frac{m_p}{m_\oplus}\right)\left(\frac{P}{1year}\right)^{\frac{2}{3}} ms \tag{1.8}$$

Pulsar timing is a very powerful method that can detect companions even down to the size of small asteroids. However, the limitation is that it can only be applied in the case of millisecond pulsars and thus it can not deliver planets that have formed in the same way as Sun-like Stars.

1.2.4 Eclipsing binaries minima timing

This method resembles the pulsar timing technique. In this case, the periodic phenomenon to be measured is the variation of light caused by the rotation of a double eclipsing system. Eclipsing minima are considered to be very stable, hence timing of their arrival may indicate the presence of a third-body in the system.

If we account a fully detached[3] binary system, with a linear ephemeris $t_{min} = t_{0min} + P * n$ then the motion of both stars around the Binary-planet barycenter will produce periodic deviations at the arriving time of the light as described in 1.2.3. In an *Observed vs Calculated*(O-C) diagram this will appear as a sinusoidal component. The amplitude of this feature will be proportional to both the planet-binary distance and the mass of the planet. This method is also sensitive to small planets, however due to slow time-scales of binary eclipses, a very-long term monitoring of such systems is required.

1.2.5 Transits

Transits are easy to understand. Under a suitable geometrical alignment, a planet intervening the optical plane of the observer and the star will cause a drop to the amount of light coming from the star (Figure 1.5).

If we consider a totally dark planet then the luminosity drop will be:

$$\frac{\Delta F}{F_*} = \left(\frac{R_p}{R_*}\right)^2 \tag{1.9}$$

ΔF is usually small, less than 2%, yet distinctly detectable with current instrumentation. The transit method is the most suitable for survey work, since a large number of stars can be imaged and measured simultaneously. Given the infrequency of proper alignment geometry, the latter also increases the probability of transit detection. Regardless the casualness of the phenomenon, both the observational and theoretical demands are extremely challenging. Due to small light variances that need to be measured a very careful data reduction need to be followed. In addition, deriving the physical characteristics from the light-curve is complex and strongly dependent on the astrophysical modeling of the host star. A more detailed analysis on transits theory will be followed in Chapter 2.

[3] Both stars are smaller than their Roche lobes and no mass transfer occurs

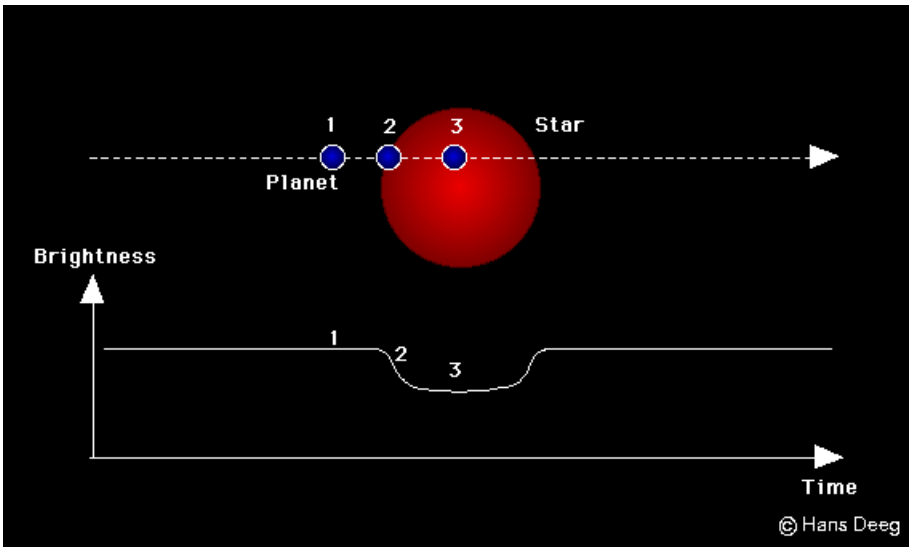

Figure 1.5: Schematics of a transiting planet.

1.2.6 Gravitational microlensing

Gravitational microlensing is the amplification of the intensity of light from a distant source due to the gravitational field of an intervening object. Necessary condition, is the exact alignment of the amplified object, the object that acts as a lens and the observer. In Literature (see *Perryman* 2000), lensing phenomena are parametrized in terms of the *Einstein ring radius*:

$$R_E = \left[\frac{4GM_L}{c^2}\frac{(D_S - D_L)\,D_L}{D_S}\right]^{\frac{1}{2}} = 8.1\left(\frac{M_L}{M_\odot}\right)^{\frac{1}{2}}\left(\frac{D_S}{8kpc}\right)^{\frac{1}{2}}\left[(1-d)\,d\right]^{\frac{1}{2}} AU \tag{1.10}$$

where M_L is the mass of the lensing object, D_L and D_S are the distances to the lens and to the source and $d = \frac{D_L}{D_S}$. The Einstein angle, namely the angle that the observer sees, E_R, is given by:

$$\vartheta_E = \frac{R_E}{D_L} = \left(\frac{M_L}{M_\odot}\right)^{\frac{1}{2}}\left(\frac{D_L}{8kpc}\right)^{-\frac{1}{2}}(1-d)^{\frac{1}{2}}\, mas \tag{1.11}$$

A time evolving model is:

$$A\,(t) = \frac{u^2\,(t)+2}{u\,(t)\left[u^2\,(t)+4\right]^{\frac{1}{2}}} \tag{1.12}$$

where $u\,(t)$ is the projected distance between the lens and the object in units of Einstein radius.

Even though gravitational lensing phenomena are complex to understand, they provide astronomers a powerful tool to study distant, typically unresolved sources. Fine structure within the photometric light-curves of such events, allow the monitoring of lens kinematics, stellar atmospheres and, of course, the detection of planetary systems. Concerning the latter, the detection probability, even for low-mass planets, is relatively high[4], especially when the

[4] $\sim$ 17% for a Jupiter-like system and 3% for a Saturn-like system

planet orbits inside the 'lensing zone' between 0.6-1.6 R_E (*Wambsganss* 1997). However, a big disadvantage of this method is that, lensing events are rather unlikely to occur more than once. Hence all the required data must been taken at the time. Figure 1.5 shows the photometric characteristics of a microlensing event due to a planet and Figure 1.6 explains the geometry of gravitational microlensing.

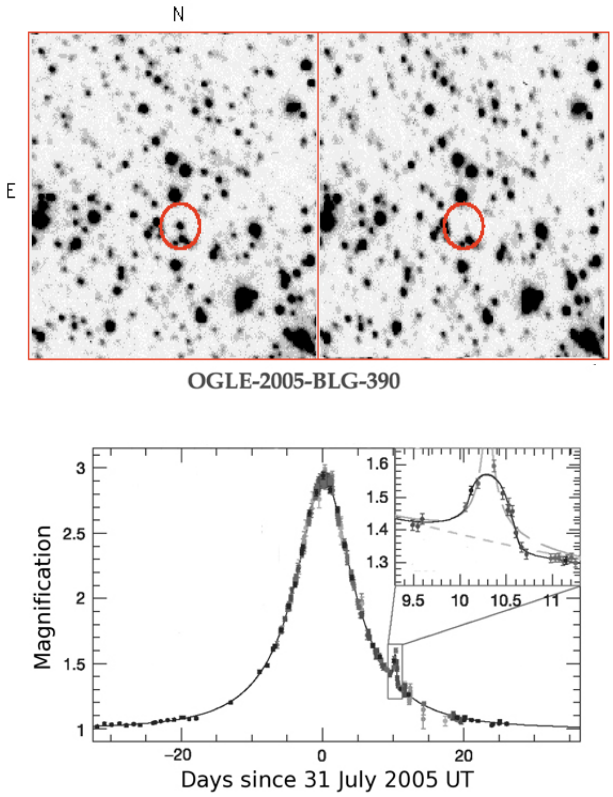

Figure 1.6: Microlensing event due to the presence of a planet.The technique itself requires no light at all coming from the lensing object. That puts strong constrains to the modeling of the light-curve, since that case is almost never identified

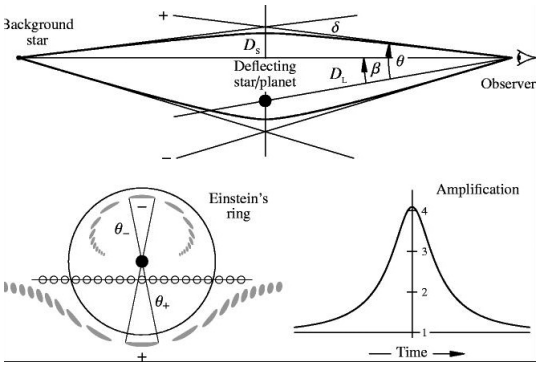

Figure 1.7: Geometrical explanation of gravitational microlensing

1.2.7 Orbital phase reflected light variations

If a giant planet is orbiting close to its host star, then an observer at $i \neq 0$ will see a different amount of reflected star's light from the planet's surface through different phases of the orbit. These reflected light variations are similar to the phases of the Moon. The amplitude of this effect is more or less the same to the transit depth of an Earth-like planet. This method may contribute a large number of planets in the future, especially with space telescopes like COROT and KEPLER.

1.2.8 Circumstellar disks

Many stars are surrounded by dusty circumstellar material. These debris disks can be detected through their infrared radiation, produced by absorbed and reemitted star light. Circumstellar disks are believed to be the product of collisional effects between comets and asteroids. That is because radiation pressure from the star should normally desolve such a disk within a small period of time. Hence debris must be produced by small bodies collisions in order to fuel the disk and keep it stable. Moreover some disks have features like central cavities, strong indicators of full sized planets which have "cleaned out" the dust inside their orbits.

1.2.9 Transit spectroscopy

Transit spectroscopy is not used for the discovery of new planets, nevertheless it is worth mentioning because it provides useful information on the formation and evolution of extrasolar planets.

Observations of that kind are based on the following concept: If an extrasolar planet has an atmosphere, then the profile of the transit light-curve would vary with wavelength, since the size of the planet would appear larger at wavelengths where its atmosphere is strongly absorbing light. The first such study, was performed by *Charbonneau et al.* (2002), who used HST to search for Sodium absorption lines near 600 nm. The yields of this survey showed that absorbtion was smaller than predicted, a fact that was attributed to clouds at the upper

limits of Hot Jupiter atmospheres. After that first discovery, *Brown et al.*(2002), *Richardson et al.*(2003) and *Deming et al.*(2005) extended the survey to the infrared in search of H_2O, CH_4, and CO absorption lines.

An other interesting phenomenon that was observed is that of the *evaporation of Hot Jupiters*.It was early suggested that planets like HD 209458b, orbiting close to their parent stars, may get hot enough leading to evaporation of their external atmospheric layers. *Vidal-Madjar et al.*(2003) and *Vidal-Madjar et al.*(2004) reported such observations of hydrogen, oxygen and carbon escaping atoms, on HD 209458 b with HST.

1.2.10 Direct imaging

Despite the difficulties of direct imaging, described in section 1.2, 11 planet mass bodies[5] have been imaged directly so far. Although most of the discoveries are controversial and the candidates' masses might exceed what could be classified as planet. Most of the candidates, orbit about nearby stars at high distances. In this subsection we will look at brief all the methods that have been used -successfully or not- so far for directly detecting reflected starlight or thermal emission of exoplanets.

Nulling interferometry

The combination of the images of two telescopes, separated by a variable and very accurately known and controlled baseline D, can be used to eliminate the light from the parent star, making the offset planet easier to be seen. At this case constructive interference will occur at a number of angles θ with $sin\theta = (2n+1)\frac{\lambda}{2D}$. By changing the baseline D, a number of angles θ can be surveyed for weak planetary emission or reflected starlight.

Doppler spectral separation

An orbiting extrasolar planet, might weakly reflect the starlight of its host star towards the observer. That light would be doppler sifted differently than the rest analyzed stellar spectrum, by a factor depending on the instantaneous orbital velocity of the planet. This information could be used to distinguish the reflected light and characterize the planet.

Coronographic imaging

In this technique the light of the host star is eliminated, using a mask in the focal plane. It has been successfully used in the past for studies of the Sun's atmosphere and corona and it is only suitable for space missions.

[5] 9 planetary systems, 11 planets and 1 multiple planet system up to the end of 2008

1.3 Characteristics of detected planetary systems

Fifteen years after the discovery of the first extra solar planet, more than 300 planetary systems are known, most of them discovered by radial velocity measurements. The sample is large enough to allow us to consider some statistics of the physical parameters of the detected systems. In order to do this we have to consider the selection effects caused by the observational methods. Moreover, most of the planets masses are known with an uncertainty of $m_p sini$. However, as it was shown by *Jorissen et al.*(2001) [26], for random distributions the distribution of $m_p sini$ is almost the same with the distribution of $m_p sini$.

1.3.1 Mass distribution

The radial velocity method is more sensitive to massive planets and therefore, a dominance of such planets might be expected. However, most of the planets discovered so far, have masses smaller than 5 M_{jup}.

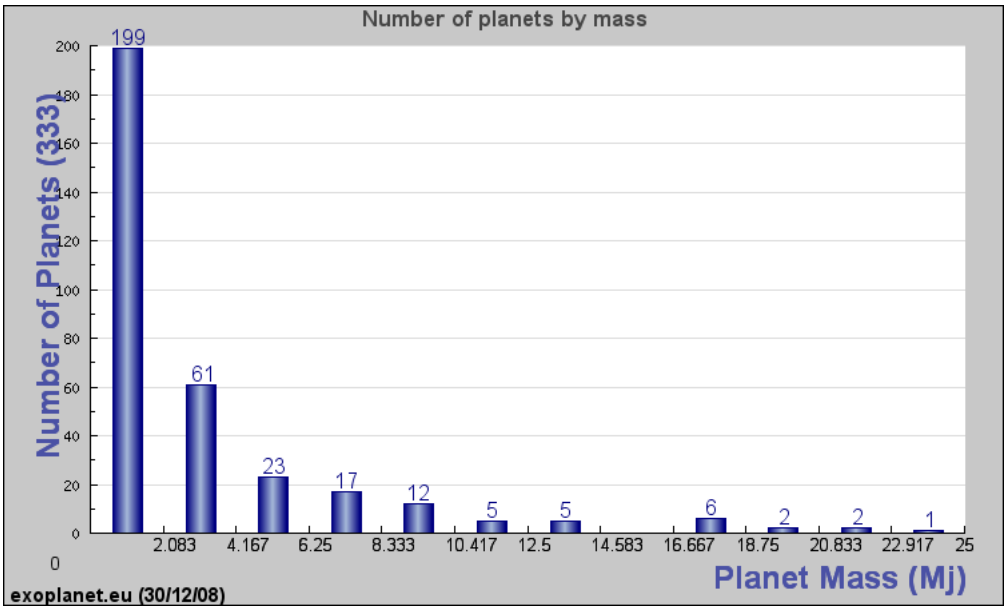

Figure 1.8: Mass population histogram of discovered planets. The distribution clearly follows a power law, despite the selection effects of the detection methods.

As it can be seen in Figure 1.8, the population rises rapidly for small masses. It is also clear that there is an unexpected absence of planets with masses that are close or exceed the deuterium burning limit ($\backsim$ 14 M_{jup}) of Brown Dwarfs. This absence is known as the *Brown Dwarf Desert*. This might lead us to the assumption that massive planets and Brown Dwarfs are formed in a different way than small planets, leading to larger orbital radii. This is also supported by the resent direct imaging findings.

1.3.2 Orbital radii and orbital periods distribution

The orbital radii and the orbital period distribution are presented in Figure 1.9 and Figure 1.10.

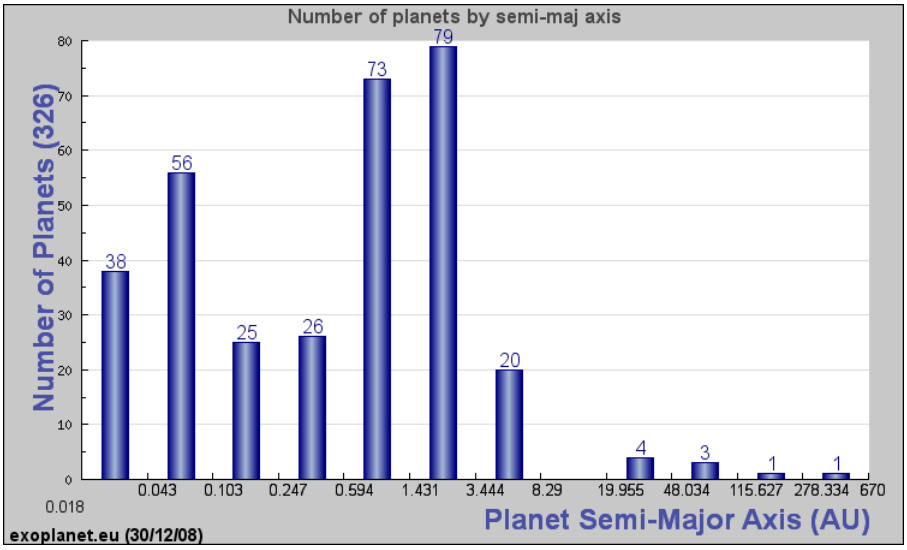

Figure 1.9: Orbital radius distribution

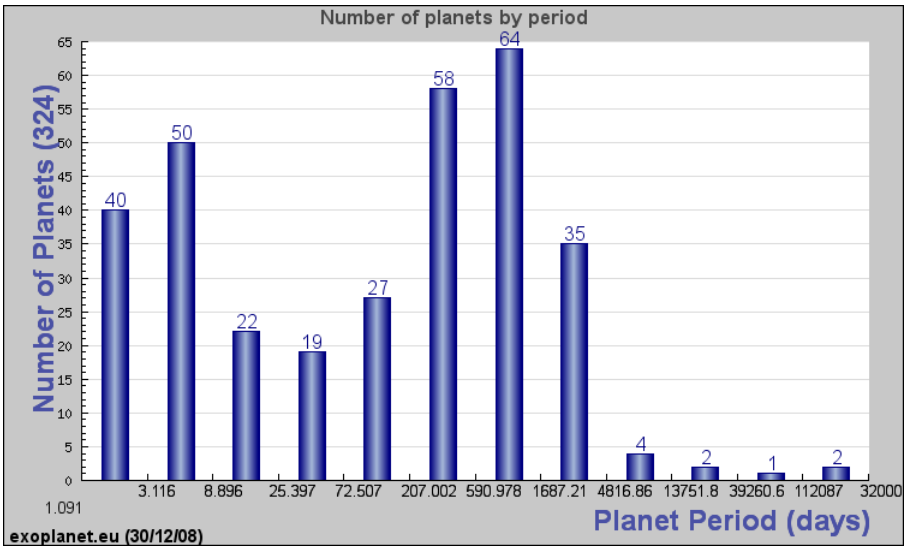

Figure 1.10: Orbital period distribution

Most of the planets discovered so far are orbiting at distances $\leq$ 1.5 AU. Although that, might not correspond to a real distribution since planets with large semi major axis, have larger orbital periods and most of the planet surveys run for less than 10 years. The period histogram appears to have a bimodal distribution, with peaks at $\backsim$ 3 days and $\backsim$ 500 days.

1.3.3 Correlation diagramms

Mass and orbital radius relation

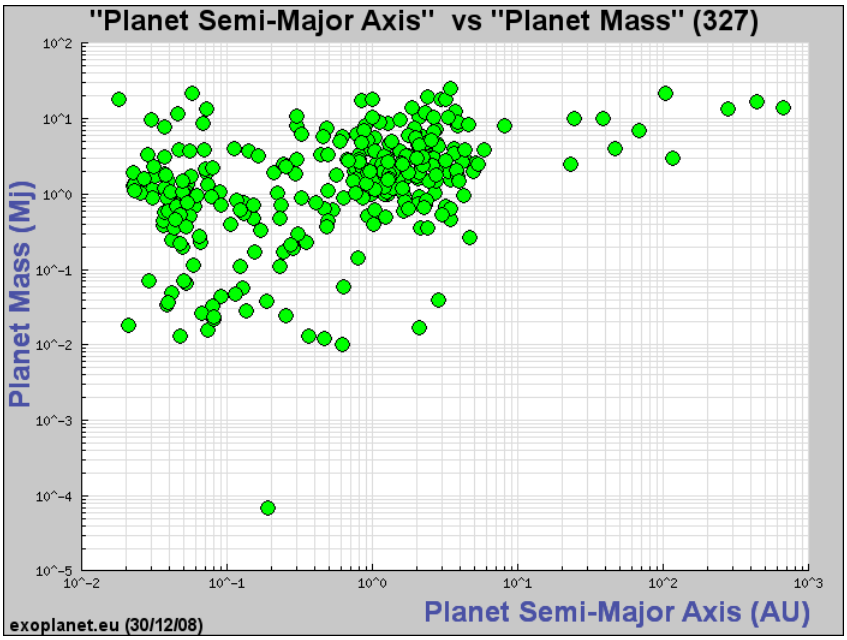

Figure 1.11: Correlation between planetary mass and orbital radius

Figure 1.11 shows the $m_p sini$ of known planets plotted against the semi-major axis of their orbits. It is clear that there is an absence of large planets orbiting close to their host stars. That is not due to selection effects since current methods should easily detect such planets. The above characteristic might be connected with the formation and evolution of planets and their atmospheres. It appears that large planets that are forming close to their stars, are quickly evaporating and loose their atmospheres as suggested by *Vidal-Madjar et al.* (2003) [43].

Eccentricities

In contrast to the above, it appears that there is no correlation between the eccentricities and the orbital radii of the detected planetary systems (Figure 1.12). There is an almost uniform spread of planets between eccentricities 0 and 0.9. That suggests that there is no specific mechanism that distributes eccentricities during the formation process. That includes also multiple planetary systems, that are indistinguishable from single planets.

Metallicity of parent stars

The metallicity of stars is represented by [Fe/H] ratio, which is defined as the abundance of iron in a star to that found in the Sun, expressed on a logarithmic scale. The [Fe/H] ratio of stars with detected planetary companions can be seen in the histogram of Figure 1.13.

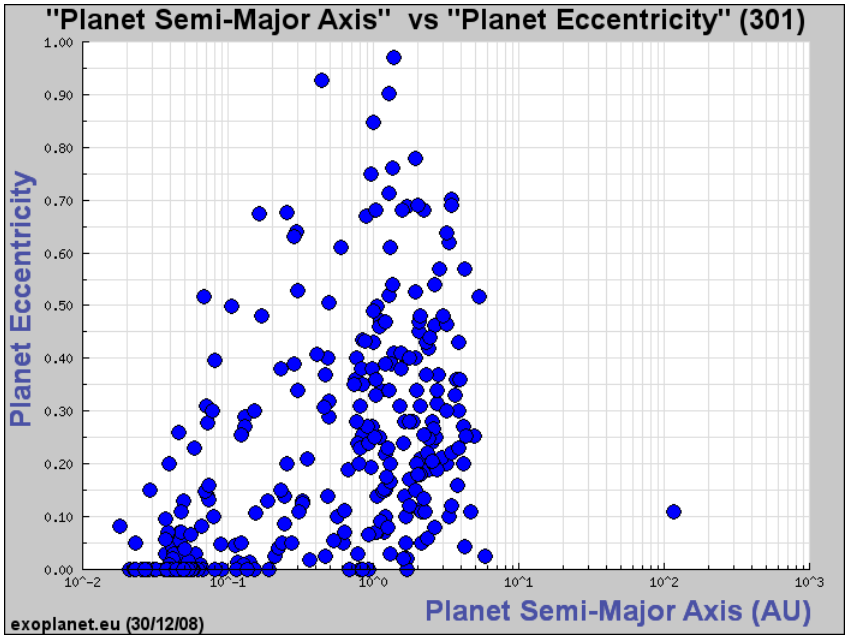

Figure 1.12: The eccentricity distribution of detected planetary systems.

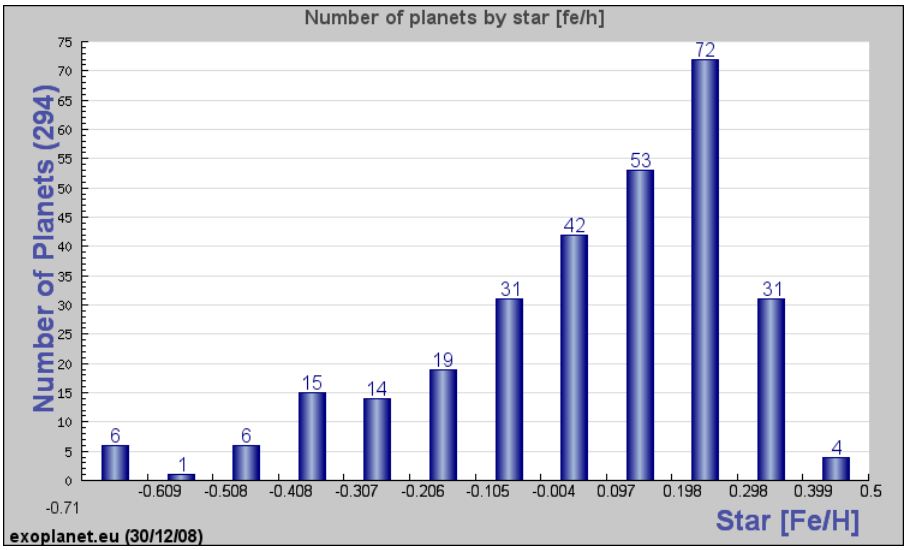

Figure 1.13: The [Fe/H] distribution of Stars with detected planetary companions.

It is obvious from Figure 1.13 that, most of the stars are distributed around $\backsim$ 0.2 [Fe/H], which means that most of the planets discovered so far have been found to orbit 2_{nd} generation stars with metallicity comparable to that of the Sun. This might be due to selection effects, as discussed in subsection 1.2.1. but it could also meen that the presence of dust in protostellar discs plays a significant role to the formation of Planets.

We live in a changing Universe, and few things are changing faster than our conception of it.

Timothy Ferris, "The Whole Shebang"

2 Planetary transits

2.1 Introduction

This chapter is devoted to the theoretical analysis concerning extrasolar transits.

As described in the previous chapter, the light-curve of a transit has a very simple, yet rich in information morphology. A fit of a transit model, combined with spectral data, can deliver all the essential parameters about the planet and its orbit. Moreover, in some cases, astrophysical information about the host star may also be extracted. The need of a precise model was obvious after the first high signal-to-noise measurement of the planet HD209458 (*Brown et al. 2001*) with the Hubble Space Telescope. The excellent quality of this light-curve made obvious, that a correct treatment of the limb darkening effect must have a key -role to transit modeling.

In the framework of this diploma thesis, I chose to use the modeling approach developed by *Gimenez (2006)* [22] . Unlike other models, the latter is a set of analytical equations, valid for any law of limb darkening and any type of transit. The cases of eccentric orbits, third light, or non-zero relative luminosity of the planet can be easily taken into account. Moreover these equations are easy to treat computationally, a fact that is of high importance for a low-bugget transit survey such as the one developed at the University of Thessaloniki.

An algorithm for the prediction of the yields of photometric planetary searches is also introduced. The probability of observing a transit of a planetary system is, of course, linked with the geometry of the system in perspective to the earth, but it is also connected with the equipment employed for the survey as well as with the quality of the observing site and the total time of the observations. The method developed here, is based on the implementation of several empirical laws, and allows the graphical interpretation of the probability as a function of celestial coordinates. This can be very useful for both the selection of targets for new Transit Surveys and the testing of the yields of already running surveys.

2.2 Transits lightcurves: A simple approach

As we discussed in 1.2.5, a transit occurs (Figures 2.1 and 2.2) when a planet is intervenes the line of sight between the observer and the host star. Under the assumption of a non-emitting, circular orbiting planet, and a uniformly emitting star (eg. without limb darkening), one can

describe the transit event through three basic equations (*S. Seager et. al.* 2003) [39]:

$$\Delta F = \frac{F_{notransit} - F_{transit}}{F_{notransit}} = \left(\frac{R_p}{R_*}\right)^2 \tag{2.1}$$

$$\frac{sin\left(\frac{t_f \pi}{P}\right)}{sin\left(\frac{t_T \pi}{P}\right)} = \frac{\left\{\left[1 - \left(\frac{R_p}{R_*}\right)\right]^2 - \left[\left(\frac{\alpha}{R_*}\right) cosi\right]^2\right\}^{1/2}}{\left\{\left[1 + \left(\frac{R_p}{R_*}\right)\right]^2 - \left[\left(\frac{\alpha}{R_*}\right) cosi\right]^2\right\}^{1/2}} \tag{2.2}$$

$$t_T = \frac{P}{\pi} arcsin\left(\frac{R_*}{\alpha}\left\{\frac{\left[1 + \left(\frac{R_p}{R_*}\right)\right]^2 - \left[\left(\frac{\alpha}{R_*}\right) cosi\right]^2}{1 - cos^2 i}\right\}^{1/2}\right) \tag{2.3}$$

where F is the total observed flux, t_f the duration of the flat part of the light-curve, and t_T the total transit duration.

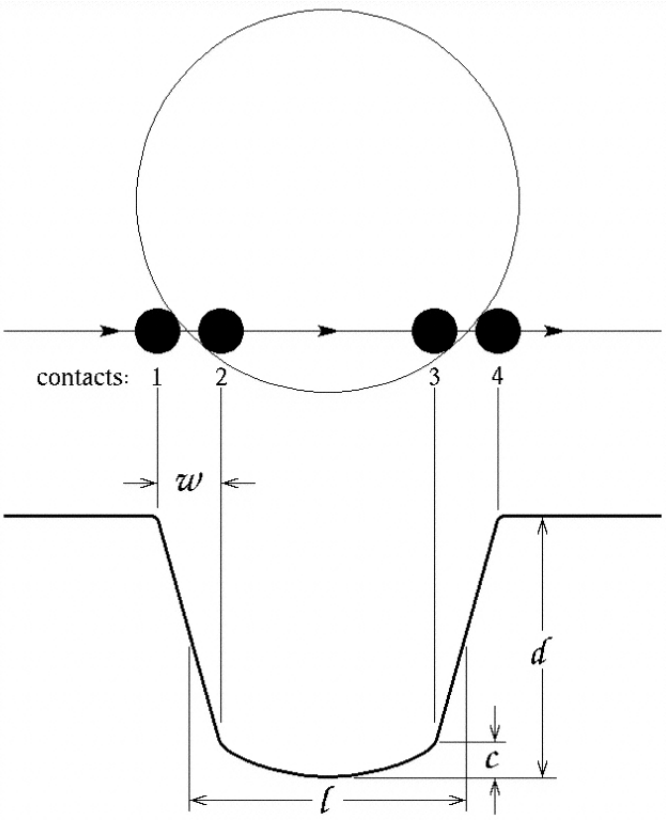

Figure 2.1: The geometry of a transit

The first equation describes the luminosity drop due to the planetary transit, while the second gives the transit's shape, parameterized by the duration between the second and third

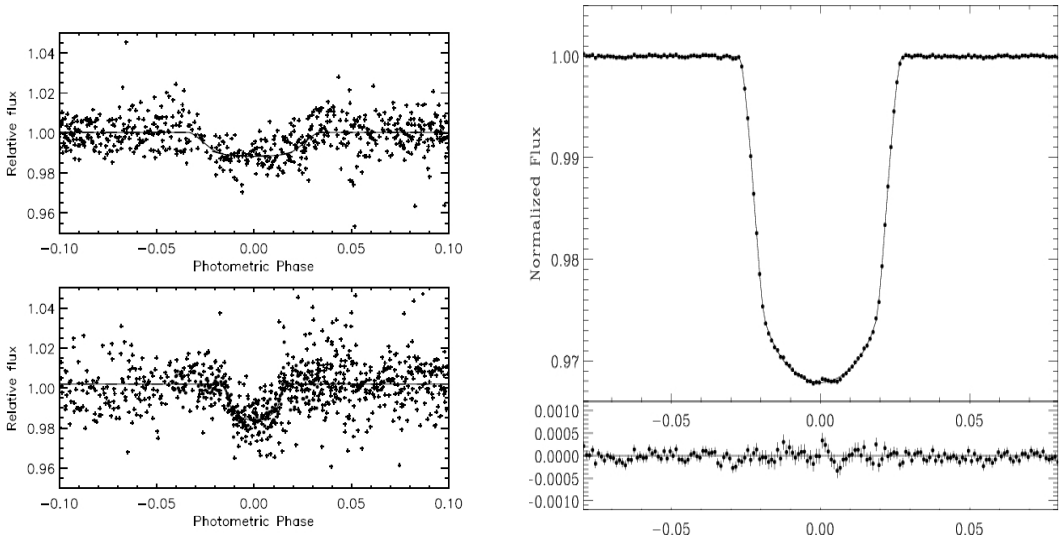

Figure 2.2: Combarison of three real transit lightcurves. *Left*: The discovery ligtcurves of two wasp planets as they were monitored by SuperWASP. Due to low S/N a simple modeling aproach can be employed. *Right*: A lightcurve of a transit obtained from space with CoRoT. The effects of limb darkening are more than clear.

contact (t_f) to the total transit duration (t_T). Finally, equation 2.3 gives the total transit duration.

The above equations can be combined with Kepler's third law $P^2 = \frac{4\pi^2\alpha^3}{G(M_*+M_P}$, and a stellar mass-radius relation, $R_* = kM_*^x$ to derive information about the physical parameters of the planet and the host star. In particular, four combinations of parameters are derived directly from the observable parameters as follows:

Equation 2.1 gives directly the planet-to-star ratio:

$$\frac{R_p}{R_*} = \sqrt{\Delta F} \tag{2.4}$$

Equation 2.2 combined with the above relation, derives the impact parameter b, defined as the projected distance between the planet and star centers during mid-transit in units of R_* (Fig 2.1):

$$b = \frac{\alpha}{R_*}cosi = \left\{\frac{(1-\sqrt{\Delta F})^2 - [sin^2(t_f\pi/P)/sin^2(t_T\pi/P)](1+\sqrt{\Delta F})^2}{1-[sin^2(t_f\pi/P)/sin^2(t_T\pi/P)]}\right\}^{1/2} \tag{2.5}$$

The rearrangement of 2.3 gives the ratio α/R_*:

$$\frac{\alpha}{R_*} = \left\{\frac{(1+\sqrt{\Delta F})^2 - b^2[1-sin^2(t_T\pi/P)]}{sin^2(t_T\pi/P)}\right\}^{1/2} \tag{2.6}$$

Finally, the mean stellar density ρ_* is derived from the above relation and Kepler's law under the assumption of $M_p \ll M_*$:

$$\rho_* = \frac{M_*}{R_*^3} = \left(\frac{4\pi^2}{P^2 G}\right)\left\{\frac{(1+\sqrt{\Delta F})^2 - b^2[1-sin^2(t_T\pi/P)]}{sin^2(t_T\pi/P)}\right\}^{3/2} \tag{2.7}$$

From the above equations and the stellar mass-radius function we can solve for five physical parameters:
The stellar mass,

$$\frac{M_*}{M_\odot} = \left(k^3\frac{\rho_*}{\rho_\odot}\right)^{1/(1-3x)} \tag{2.8}$$

the stellar radius,

$$\frac{R_*}{R_\odot} = \left(k^{1/x}\frac{\rho_*}{\rho_\odot}\right)^{x/(1-3x)} \tag{2.9}$$

the orbital radius,

$$\alpha = \left(\frac{P^2 G M_*}{4\pi^2}\right)^{1/3} \tag{2.10}$$

the orbital inclination,

$$i = cos^{-1}\left(b\frac{R_*}{\alpha}\right) \tag{2.11}$$

and the planetary radius,

$$\frac{R_p}{R_\odot} = \left(k^{1/x}\frac{\rho_*}{\rho_\odot}\right)^{x/(1-3x)}\sqrt{\Delta F} \tag{2.12}$$

Applications and limitations Once a planetary candidate is found, one may apply the above analysis to the light-curve, for a quick estimation of the stellar and planetary parameters. This is proven to be very useful, especially for a planetary survey like ours. The best candidates may easily be selected, for further follow-up photometric and spectroscopic observations with larger telescopes. However, in practice, this approach is limited by several factors. First of all, the orbital eccentricity is not treated at all. This is not the case for most of planetary systems discovered to date (Fig 1.12), hence an application of this model may lead to completely false estimation of parameters. Additionally, stellar limb darkening plays a significant role to the shape of the transit, mostly by curving the "flat" part of the light-curve. Finally ,when it comes to fitting, errors may occur, due to the limited photometric precision and time sampling.

2.3 A precise analytical model

Gimenez (2006) [22] derived an analytical formula for the computation of light-curves, that is valid for almost any case of transit, including eccentric orbits, third light coming from a neighbor star, any degree of limb darkening and non-zero relative luminosity of the planet. An additional advantage is that the analytical nature of the equations makes them easier to treat

computationally, compared to other widely used models. That, serves best our requirements, since, we only have access to regular PCs, and therefore in a case of a possible discovery and follow-up observations we will need a robust algorithm to derive the physical parameters.

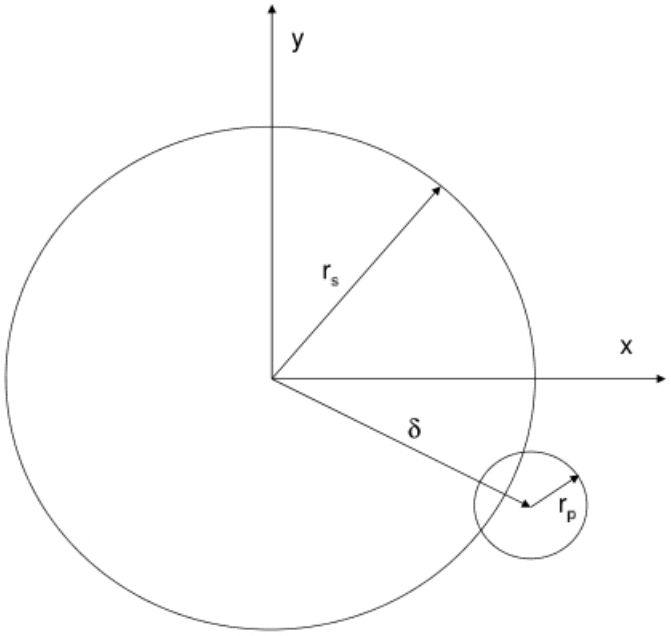

Figure 2.3: The geometry of the *Gimenez* model

2.3.1 The equations

Considering a star-planet system, properly alighted geometrically, to give a transit, we can parameterize the total luminosity L, as seen from earth as $l(t) = L_* + L_P - \alpha(t)L_*$, where L_* and L_p are the luminosities of the star and the planet respectively and a(t) represents the relative drop to luminosity due to the partial occultation of the star by the planet. If we use the orbital phase $\vartheta = (2\pi/P)(t - t_0)$ instead of t, and we set $L_* + L_p = 1$ the above equation becomes:

$$l(\theta) = 1 - \alpha(\theta)L_* \tag{2.13}$$

The condition for the validity of this equation is that the projections of both the star and the planet to the celestial sphere are circular discs moving in front of each other. The main problem consists of finding an analytical relation for $a(\theta)$ by taking into account a non stable luminosity distribution over the surface of the star.

2.3.2 Limb darkening laws

By the term *limb darkening* we generally refer to the dimming of the stellar brightness as we move from the center of the star to the edge. Limb darkening depends on stellar temperature, surface gravity, metal abundance and atmospheric micro-turbulence velocities (*Claret 2000*) [15] . The simplest way to model the luminosity distribution is to adopt a linear law of the form:

$$I(\mu) = I(1)[1 - u(1 - \mu)] \tag{2.14}$$

where $\mu = cos\gamma$ with γ being the angle between the line of sight and the surface normal, $I(1)$ the intensity of light at the center of the star and u the so-called coefficient of linear limb darkening. The latter, is usually fixed to a value near 0.6, for Sun-like stars.

One may adopt a quadratic law, in order to avoid possible misleadings to the estimation of physical parameters, especially the impact parameter and inclination:

$$I(\mu) = I(1)[1 - u_a(1 - \mu) - u_b(1 - \mu)^2] \tag{2.15}$$

In his model, *Gimenez* (2006) adopted a more general law that is given by

$$I(\mu) = I(1)\left[1 - \sum_{n=1}^{N} u_n(1 - \mu^n)\right] \tag{2.16}$$

The degree of the polynomial can be fixed to a value suitable for the precision needed.

2.3.3 The $a(\theta)$ function

The problem of finding an analytical model for a transit light-curve, is simplified in computing the term:

$$\alpha = \int_s I(\mu) cos\gamma d\sigma \tag{2.17}$$

where $d\sigma$ is the surface element and the integral is extended over the hole surface S of the star. It can be shown that the above, with the use of equation 2.16 becomes

$$\alpha = \sum_{0}^{N} C_n \alpha_n \tag{2.18}$$

where the constants C_n are given by

$$C_0 = \frac{1 - \sum_{n=0}^{N} u_n}{1 - \sum_{n=1}^{N} \frac{nu_n}{n+2}}; C_n = \frac{u_n}{1 - \sum_{n=1}^{N} \frac{nu_n}{n+2}} \tag{2.19}$$

The limits of integration can be calculated with the use of the parameters δ, R_*, R_p, i, with δ defined as the projection of center-to-center distance of the Planet and the Star, in the vertical to the line of sight plane. In the case of eccentric orbits δ is:

$$\delta^2 = \left[\frac{(1-e^2)}{1-esin(\theta-\omega)}\right]^2 (1-cos^2\theta sin^2 i) \tag{2.20}$$

with ω the position of the periastron.

The a_n functions

First of all, if we take a closer look to the geometry of the problem (Figure 2.2), we see that for $\delta > R_* + R_p$ no transit occurs, and $a = 0$. Ingress and egress phases occur when $R_* + R_p > \delta > R_* - R_p$, and finally $\delta < R_* - R_p$ describes the main phase of the transit.
Gimenez (2005) evaluated the terms a_n by making use of eclipsing binary modeling tools introduced by *Kopal* (1977). The idea is to treat the problem of the fractional loss of light, as a cross-correlation of two apertures. That links the problem, to the diffraction patterns of two apertures as described by physical optics and allows the use of well established mathematical tools like diffraction integrals.

With the use of this approach, it is found that the a_n functions can be expressed, using a Hankel transform, as

$$a_n(k,h) = 2^\nu \Gamma(\nu) \int_0^\infty (x/k)^{-\nu} J_\nu(x/k) J_1(x) J_0(hx/k) dx \tag{2.21}$$

where J_n are Bessel functions of the first kind with real arguments, $h = \delta/R_*$ and $\nu = (n+2)/2$. After some calculations we finally derive:

$$a_n(b,c) = \frac{b^2(1-c^2)^{\nu+1}}{\nu\Gamma(\nu+1)} \sum_{j=0}^{\infty} (-1)^j (2j+\nu+2) \frac{\Gamma(\nu+j+1}{\Gamma(j+2)} \times \Big\{G_j(\nu+2;\nu+1;1-b)\Big\}^2 G_j(\nu+2,1;c^2) \tag{2.22}$$

with $b = R_p/(R_* + R_p) = k/(1+k)$, $c = \delta/(R_* + R_p = b\delta/R_p)$ and G_n the Jacobi polynomials. Equation 2.21 is valid for any degree of limb darkening, eccentricity, and type of eclipse. Moreover, it requires no more than a few FORTRAN lines and a few seconds to be calculated, even up to a degree of Jacobi polynomial as high as $j = 300$. The precision of the model is obviously a function of the polynomial degree as well as the degree adopted for the limb darkening law. It can be proved though that Equation 2.22 converges fast, and hence a good geometric precision for ground based observations can be achieved for $j \leq 20$ (*Gimenez* 2006).

2.3.4 Fitting the transit parameters

The inverse problem, namely fitting the model to a real error-dominated light-curve is somewhat, more complicated and also requires more computing time. Parameters that need to be considered for the fitting are: The relative radius of the star R_*, the ratio of the radii $k = R_p/R_*$ (Figure 2.4), the inclination i and the limb darkening coefficients u_n (Figures 2.5 and 2.6). If

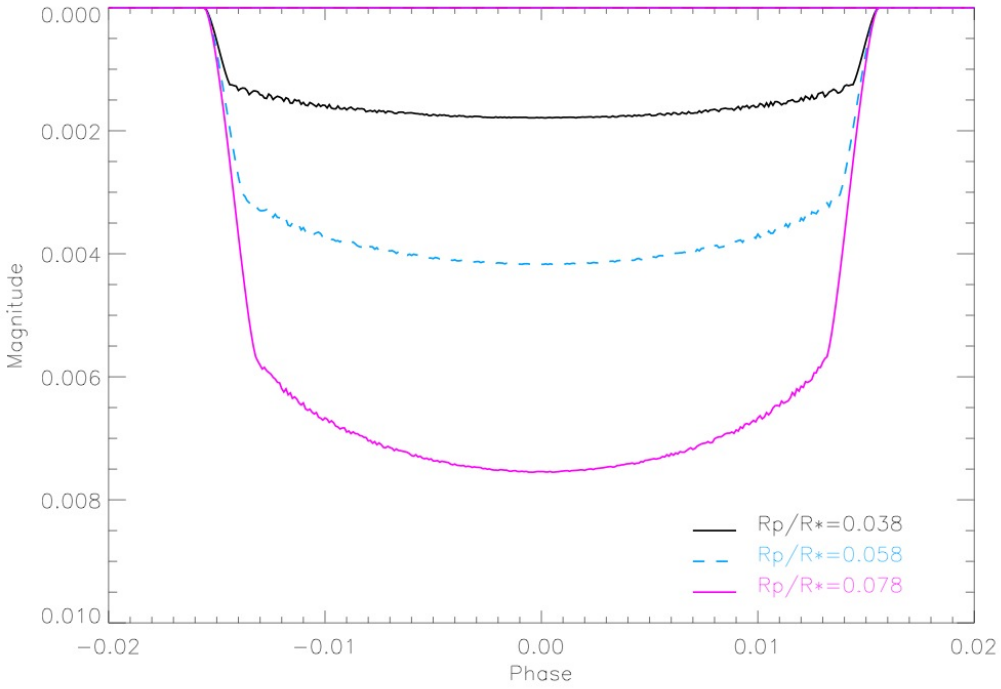

Figure 2.4: Changes of the transit profile is respect to R_p/R_*.

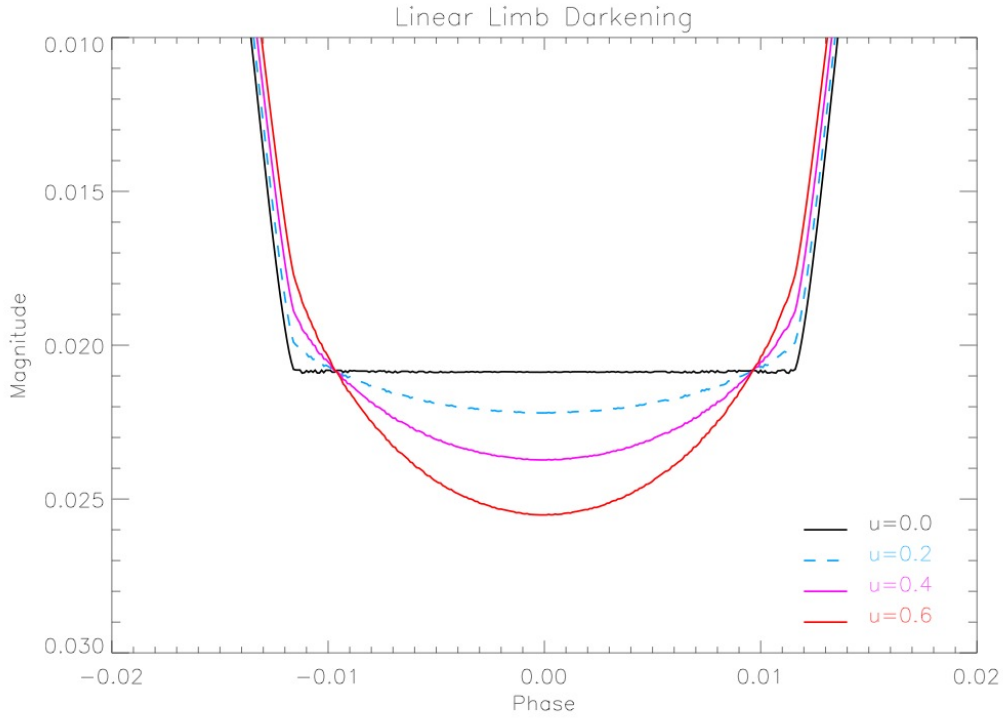

Figure 2.5: Changes of the transit profile in respect to the linear limb darkening coefficient. The flat part of light-curve depends strongly on the coefficient u while some changes are also noticed at the ingress and egress parts of the curve.

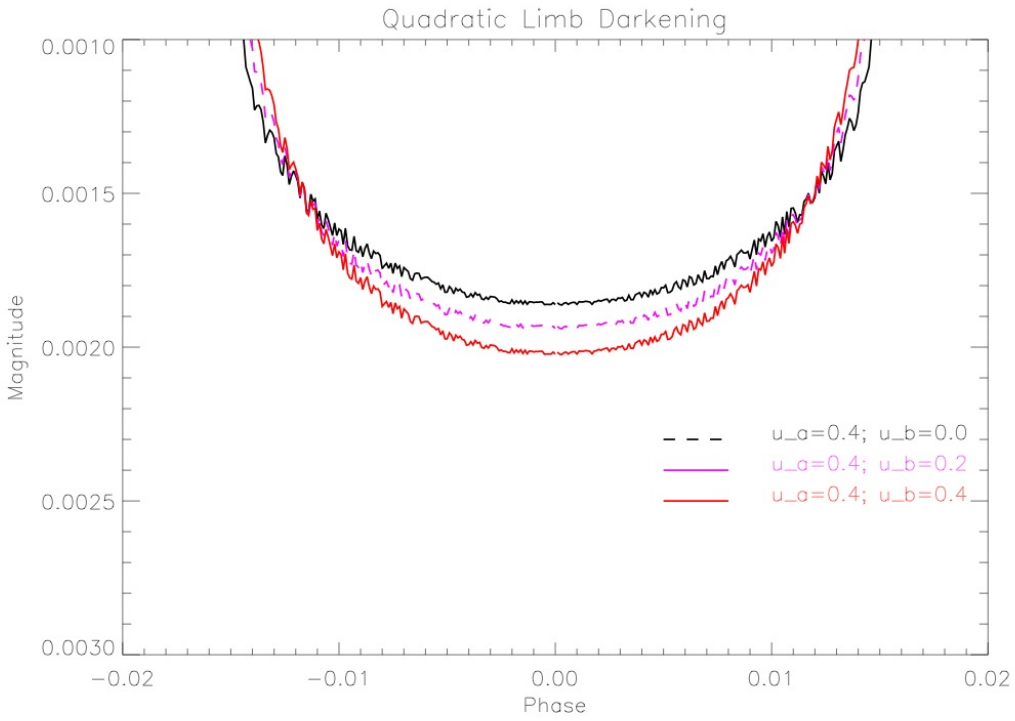

Figure 2.6: Changes of the transit profile in respect to the quadratic limb darkening coefficients. u_b mainly effects the flat part of the light-curve. Changes at the ingress and egress faces are negligible

we adopt a quadratic limb darkening law we only need, u_a, u_b and the total number of parameters reduces to five. Following the *Gimenez* approximation, we can adopt the phase of the ingress θ_1 instead of R_* by applying the transformation:

$$R_* = \frac{(1 - cos^2\theta_1 sin^2 i)^{1/2}}{1 + k} \tag{2.23}$$

It is also better to use $u_+ = u_a + u_b$ and $u_- = u_a - u_b$ in order to avoid correlations between the two limb-darkening coefficients. For the purposes of this dissertation, I use fixed values of u_+ and u_- from model atmospheres. The best parameter solution $l(\theta, \theta_1, k, i, u_+, u_-)$ is found through the *Levenberg - Marquard* method [18] by minimizing χ^2 between the observed and the calculated light-curve. We can reduce the computing time by adjusting properly the upper and lower limits to the space domain of each parameter, for example we can set $arccos(R_* + R_p) \leq i \leq 90^o$ for the inclination and a proper value for θ_1 which can be estimated straight from the light-curve.

2.4 Transit detection probability

The possibility of detecting a transit is a very important aspect for both astrophysical and observational reasons. Calculations for the derivation of transit detection probabilities have already been done by many others (Rosenblatt 1971 [38] ; Borucki and Summers 1984 [11] ; Rachel Street 2001 [42] ; Beatty and Gaudi 2008 [20]). *Beatty and Gaudi*, gave an approximation for the transit detection probability based on an analytical Galactic stellar density model(*Bahcall and Soneira 1980*). In brief, they evaluated the transit detection probability by making use of the formulae:

$$\frac{d^6 N_{det}}{dR_P dp dM dr dl db} = \rho_*(r, l, b) r^2 cosb \frac{dn}{dM} \frac{df(R_p, p)}{dR_P dp} \times P_{det}(M, r, R_p, p) \tag{2.24}$$

where $P_{det}(M, r, R_P, p)$ is the probability of a given star of mass M and distance r, to host a planet of radius R_P and period p.

However, such an approach does not treat for stellar density inhomogeneities inside the Galactic plane, which play a significant role in locating promising spots for transit surveys on the celestial plane. In order to take into account these inhomogeneities, I calculate the transit detection probabilities of hot Jupiters, by using observational data available through the Tycho catalogue. At the basis of this analysis, is the segmentation of the celestial plane into a mosaic of virtual field of views. Such a treatment allows the visualization of the probability function with probability sky maps.

The Tycho catalogue

The Tycho catalogue includes observations of roughly 1 million stars taken with the HIPPARCOS satellite between 1989 and 1993. For most of Tycho entries information about the position (α, δ), the color index $(B - V)$, the visible magnitude m_v and stellar parallax are given

for most of the stars with $m_v \leq 11.5$. The catalogue is almost complete for stars with $m_v \leq 12$ but we also find complete entries for fainter objects.

2.4.1 Astrophysical and observational parameters

For the calculation of the transit probability, the Tycho catalogue is segmented into $x^0 \times y^o$ virtual field of views. x and y can be adjusted to correspond to the field of view of a specific instrumental setup. In a first approach a field of $1^o \times 1^o$ is adopted. Each field is then correlated to the coordinates (α, δ) of the center of the field. The number of stars at each individual field is n. As a first step, the sky-projected mean separation $S(\alpha, \delta)$ of all stars in one field (Figure 2.7) is calculated by:

$$S(\alpha, \delta) = \sqrt{\frac{3600}{n}} \tag{2.25}$$

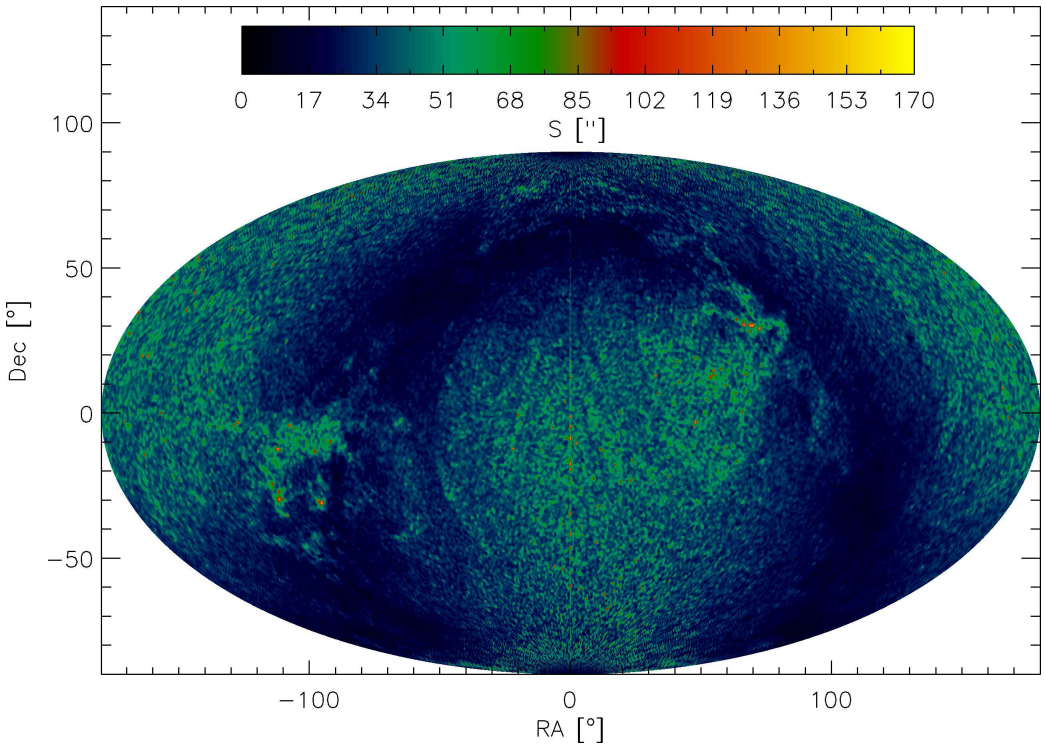

Figure 2.7: Separation $S(\alpha, \delta)$ of the stars in the Tycho catalog. The size of each field of view is 1×1^o

If l_{pix} is the linear pixel size of a given CCD detector, in units of μm, F the focal length of the instrument in mm and n_{pix} the total number of pixels per dimension, the number of the stars that can be resolved within a field of view is given by:

$$N = 206265 \frac{l_{pix} n_{pix}}{FS(\alpha, \delta)_{arcsec}} \tag{2.26}$$

if $S(\alpha, \delta) \geqslant 206265 \frac{l_{pix} n_{pix}}{FS(\alpha, \delta)}$ (*McLean* 1997) [32].

α	STELLAR MASS REGIME
3.05 ± 0.14	$M_* \leq 0.5 M_\odot$
4.76 ± 0.01	$0.5 M_\odot \leq M_* \leq 1.5 M_\odot$
3.68 ± 0.05	$1.5 M_\odot \leq M_*$

Table 2.1: Empirical values for α in the *mass − luminosity* relation

Stellar parameters

An important step for further analysis, is to deliver the astrophysical parameters T_{eff}, M_v and R_*. Those parameters are not part of the Tycho entries so they need to be calculated indirectly. First we can deliver the effective temperature T_{eff} of the star, from the color-index (B-V) by

$$T_{eff} = 10^{14.5501-(B-V)]/3.684} K \tag{2.27}$$

This equation is valid for main-sequence stars of the spectral regime $O3 - M8$ with $T_{eff} \leq 9100K$ (Reed 1998) [37]. We can derive the absolute visible magnitude M_v from m_v and d as

$$M_v = m_v - 5^m log\left(\frac{d}{10pc}\right) \tag{2.28}$$

where the intergalactic extinction, is neglected. The stellar radius R_* and the stellar mass M_* can now be calculated via:

$$\frac{R_*}{R_\odot} = \left[\left(\frac{5770K}{T_{eff}}\right)^4 10^{(4.83-M_v)/2.5}\right]^{1/2} \tag{2.29}$$

and

$$M_* = (4\pi R_*^2 \sigma_{SB} T_{eff}^4)^{1/\alpha} \tag{2.30}$$

where the coefficient α of the $Mass - Luminosity$ relation $L \propto M^\alpha$, depends on the mass of the star (*Cester et al. 1983*) [13]. Values for α near the mass regime of the Sun, are listed in Table 2.1.

Now, we have all the essential information to derive the stellar metallicity $[Fe/H]_*$ by

$$[Fe/H]_* = \frac{1}{411}\left(\frac{T_{eff}}{K} - 8423 + 4736(B-V) - 1106(B-V)^2\right) \tag{2.31}$$

as given in *Santos et al.* (2004). This relation is valid for stars with $0.51 \lesssim B - V \lesssim 1.33$, $4495K \lesssim T_{eff} \lesssim 6339K$, $-0.7 \lesssim [Fe/H]_* \lesssim 0.43$ and $log(g) \geqslant 4$. This limitation, combined with the fact, that such an approach is only valid for MS stars, reduces our sample from 1031992 to roughly 400000 objects.

2.4.2 Transit detection probabilities

We now can use the stellar parameters, to calculate the probability of a given star to host an extrasolar planet. *Fischer and Valenti* (2005) [19] , derived the empirical relationship:

$$P_{\exists planet}([Fe/H]_*) = 0.03 \,.\, 10^{2[Fe/H]_*} \tag{2.32}$$

based on Doppler measurements on 850 stars , sensitive enough to detect semi-amplitudes of $V \gtrsim 30ms^{-1}$ and orbital periods shorter than $4yrs$. Since we are only interested in typical surveys with observing sessions of $\lesssim 50d$, the latter does not put limitations to our analysis.

In the next step, the probability of a given planet to actually give a transit has to be considered. By theoretical reflections about arbitrary inclinations of the orbital plane with respect to the observer's line of sight and including Kepler's third law, *Gilliland et al* (2000) [21] found the geometric transit probability to be

$$P_{geo} = 23.8\left(\frac{M_*}{M_\odot}\right)^{-1/3}\left(\frac{R_*}{R_\odot}\right)\left(\frac{P}{d}\right)^{-2/3} \tag{2.33}$$

where P is the orbital period. A more detailed expression , that also includes eccentricity and the angle of periastron is given by *Seagroves et al* (2003) [40] . Note that P_{geo} also depends implicitly on the semi-major axis α via $P = P(\alpha)$.

By combining equations 2.32 and 2.33 we can deliver the total probability of a transit to occur on a given star via

$$P_{tot} = P_{\exists planet} \times P_{geo} \tag{2.34}$$

At this point we can already produce a visualization of the probability function 2.34, by calculating the total probability for each field to show at least one transit via

$$P^{field}_{\exists transit} = 1 - P^{field}_{\nexists transit} = 1 - \prod_{i=1}^{n}(1 - P_{tot}) \tag{2.35}$$

Plots for $P = 3.63d$ (the mean orbital period of all the known transiting planets), $5, 10, 25, 50$ are shown in Figure 2.9. A stepsize of 0.5^o between two continuous fields was chosen, in order to smooth the visualization. The maximum transit probability for the first case (Figure 2.8), which is the most interesting observationally, is 96.75% at $(\alpha, \delta) = (21.4h, -20^o)$.

Observational restrictions

Even if a transit occurs, it is not sure if it will be observed by a specific instrument. In order to estimate the detection probability, we have to consider for factors that put limitations to our observations. A detailed analysis on that has been undertaken by *Beautty and Gaudi* (2008) who treat for instrumentation, observing conditions (seeing, scintillation, brightness of the sky) and integration time.

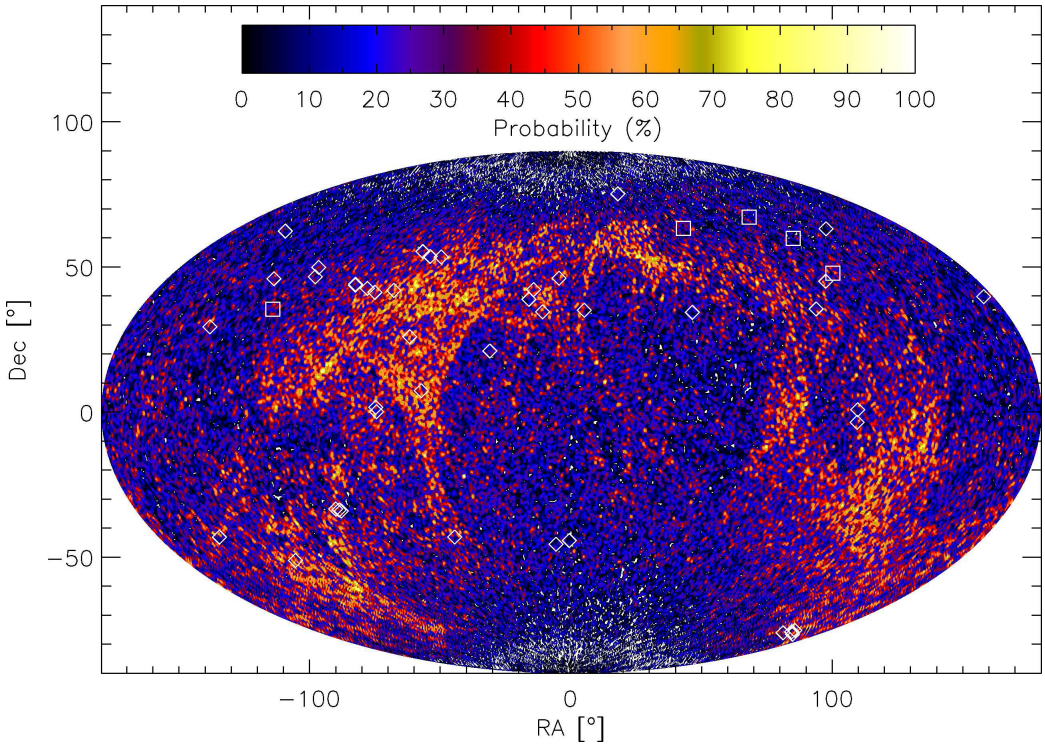

Figure 2.8: Sky map of the transit probability $P^{field}_{\exists transit}$ .The size of each individual probability field is 1×1^o while the step-size between the fields is 0.5^o. The orbital period that this map corresponds is $P = 3.63d$ which is the mean orbital period of all transiting planets. The positions of these 57 planets are indicated with circles.

In a first approach, it is assumed that these factors do not influence our calculations. Specifically, because the Tycho catalogue is mainly limited to stars with $m_v \leq 11.5$, it is assumed that all stars that can be resolved within a specific field, can be monitored for transits. Moreover, the window function, namely the amount of finite observing nights that are required to observe a transit of unknown phase, only depends on the period. Hence, it can be maximized separately for a given period and be ignored for the rest of the calculations.

The only observational factor that is considered to be significant is the number of stars that can be resolved by a specific instrumental setup. By combining equations 2.26 and 2.35 a general equation for the transit detection probability of a specific instrument in a certain field of view is derived:

$$P^{field}_{\exists transit} = 1 - P^{field}_{\nexists transit} = 1 - \prod_{i=1}^{N}(1 - P_{tot}) \tag{2.36}$$

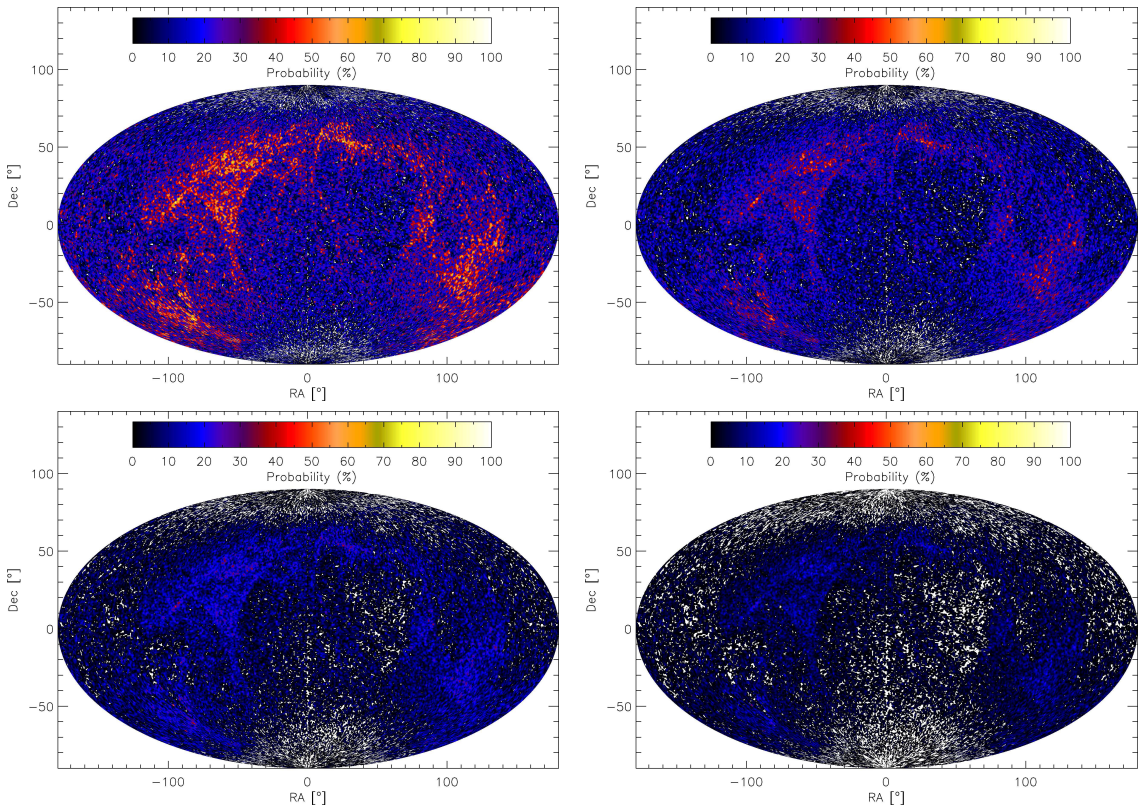

Figure 2.9: Same as before but for $P = 5d$, $10d$, $20d$ and $50d$ respectively

2.4.3 Tests of the algorithm

In order to test the algorithm, I chouse to compare two well established surveys: The Berlin Exoplanets Search Telescope (BEST) (*Rauer et al.* 2004) [36] and the XO project (*McCulough et al.* 2005) [31] . Both surveys share some technical properties on the one hand (eg. they are set at similar latitudes) but differ in important instrumental aspects on the other hand. This allows for a comparison of the efficiency of the two surveys.

BEST has a magnitude range of $8 \leq m_v \leq 14$ (*Rauer et al.* 2004) while XO gazes at stars with $9 \leq m_v \leq 12$ (*McCullough et al.* 2005). That means that some of the Tycho objects are not observed by the XO project but retained in the BEST sample. This number is found to be 154715 or 14.99% of our overall sample.

What segregates the two projects is the differend CCD resolution and field of view. BEST has a small field of 3.1^o with a resolution of $5.5"/px$ while XO is able to monitor a larger field of 7.2^o but at a lower resolution of $25.4"/px$.

Results and conclusions

Probability sky maps for BEST and XO are shown in Figure 2.10.

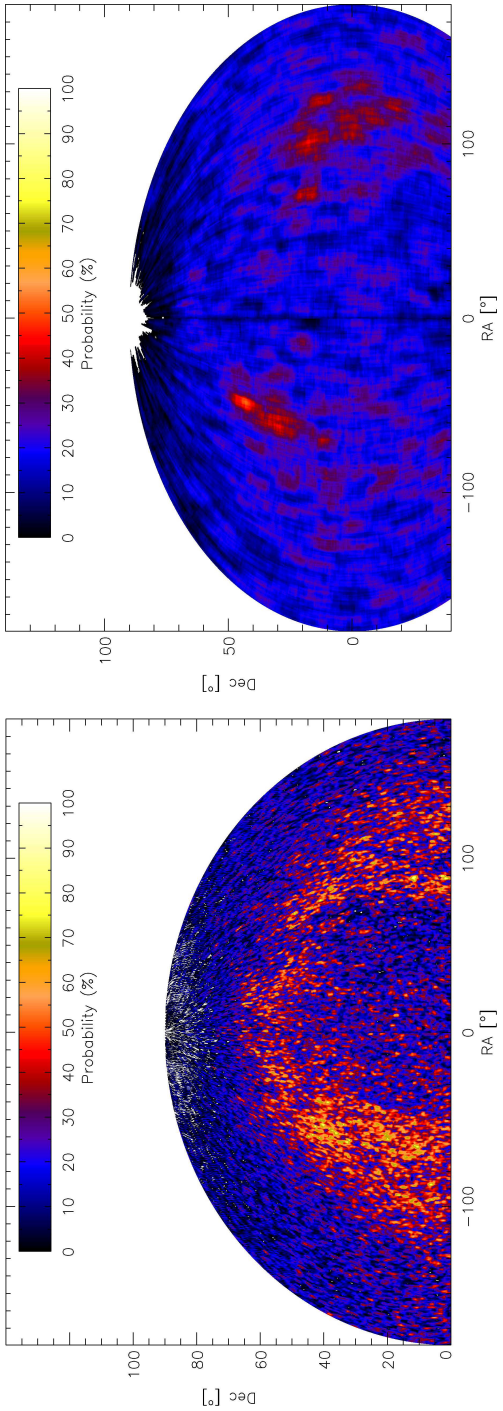

Figure 2.10: Transit detection probability for two different transit surveys. *Left*: The Berlin Exoplanets Search Telescope (BEST). and *Right*: The XO project. Due to the smaller angular separation, more stars merge to unresolved light sources and transits detection becomes more challenging.

The virtual field of view for each map was fixed to the field of view of each individual project. A step size of 0.5^o was chosen between two continuous fields for both surveys. The resulting maps show that the transit probability functions of the two surveys follow the trace of the galactic plane, as expected, due to the higher stellar density of these areas. However the limited resolution of the XO survey makes a large number of stars in its images to merge into indistinguishable objects. That results to a significantly lower detection probability compared to the BEST survey. Namely a high angular resolution instrument yields better results comparable to a wide field monitoring with low resolution.

An other interesting result is that several promising fields with detection probabilities $\gtrsim$ 50% of the BEST map does not appear on the XO map, and vise versa. That leeds us to the conclusion that every individual instrumental setup has its own "hot spots" on the celestial sphere. The latter can be used to coordinate efforts of all established and future surveys, in order to increase the rate of detection.

A time will come when men will stretch out their eyes. They should see planets like our Earth.

Christopher Wren

3

Designing a new planetary survey

3.1 Introduction

The first thoughts for the initiation of a planetary survey at the Aristotle University of Thessaloniki (AUTH) began about 3 years ago, when a new large-format CCD detector was purchased. Some experience on the field had been gained before that, due to the partitipation of AUTH students to the preliminary tests of the WASP0 instrument (*Mislis et al, 2006*) [33]. However, the observatory did not own at the time suitable equipment and the experience in data reduction techniques was limited. After the purchase of the instrument, some initial tests had began, together with an attempt to develop a reduction pipeline for high precision photometry.

In this chapter, I describe the instrumentation, as well as the observing site and the design of the observing strategy.

3.2 The Holomon astronomical station.

The Holomon Astronomical Station, is located on the slopes of Mt. Holomon at Chalkidiki, about 100 km away from the city of Thessaloniki. The facilities are situated inside the *University Forest of Taxiarchis* which is managed by the *Department of Forestry and Natural Environment* (DFNE). Despite it's vicinity to alarge city (Thessaloniki), Mt. Holomon suffers very limited light pollution and has extraordinary stable atmospheric conditions. Moreover the DFNE accommodation facilities are excellent and make the place ideal for safe and comfort night observations. After a few years of observations with portable small telescopes, the Laboratory of Astronomy of the University of Thessaloniki, in collaboration with DFNE decided to build a small station (*Figure* 3.1), able for hosting $2-3$ small telescopes.

After the station was completed, an extensive long-term monitoring of the astronomical seeing with the DIMM method [34] was initiated. The results confirmed the initial impression for excellent atmospheric conditions. The seeing is usually bellow $1 arcsec$ with an average value of $0.86''$ obtained over the course of 5 years. (*Nestoras et al.* 2006) [34].

From the same measurements, the average value of the scintillation index was found to be 0.049 mag and the average value of the isoplanar angle 45.32 arcsec. Plots for four different nights as well as a histogram of the overall values are shown in figures $3.2-3.5$. Some general

Figure 3.1: The Holomon observatory

information about the site are given in Table 3.1.

Longitude	$23^o\ 30'\ 19.6''E$
Latitude	$+40^o\ 25'\ 58.4''$
Altitude	800 m
Temperature	-15^o to 35^o C
Annual rainfall	750 mm
mean seeing	0.86 arcsec
mean scintillation	0.049 mag
mean isoplanar angle	45.32 arcsec

Table 3.1: General information about Mt. Holomon

The favorable atmospheric conditions as well as the low light pollution, make the Holomon Astronomical Station perfect for high-precision photometry and thus, for hosting an extrasolar planetary survey.

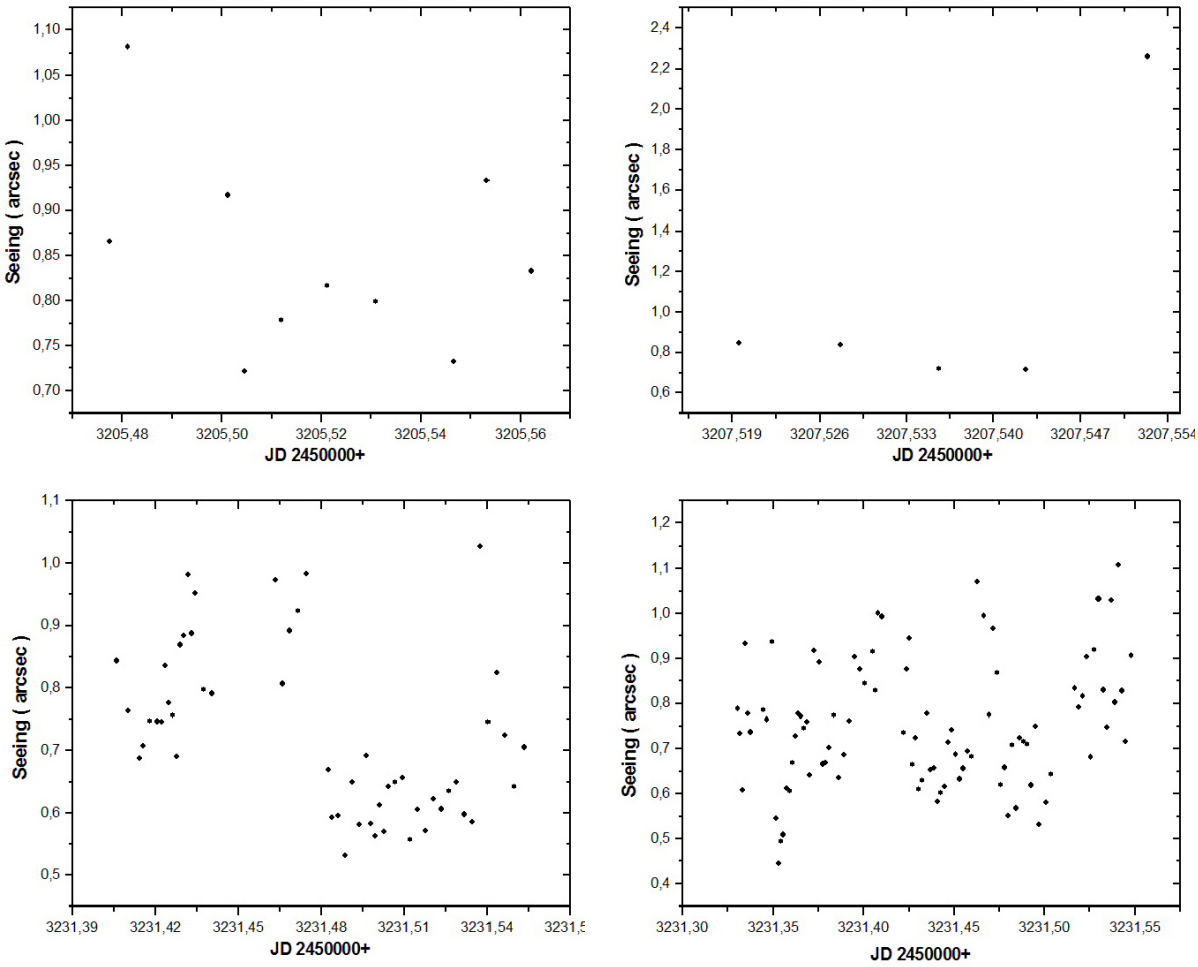

Figure 3.2: Seeing conditions at Mt. Holomon. The long term monitoring with the DIMM method confirmed the extraordinary good atmospheric conditions of the area. Plots of 4 individual nights are shown here.

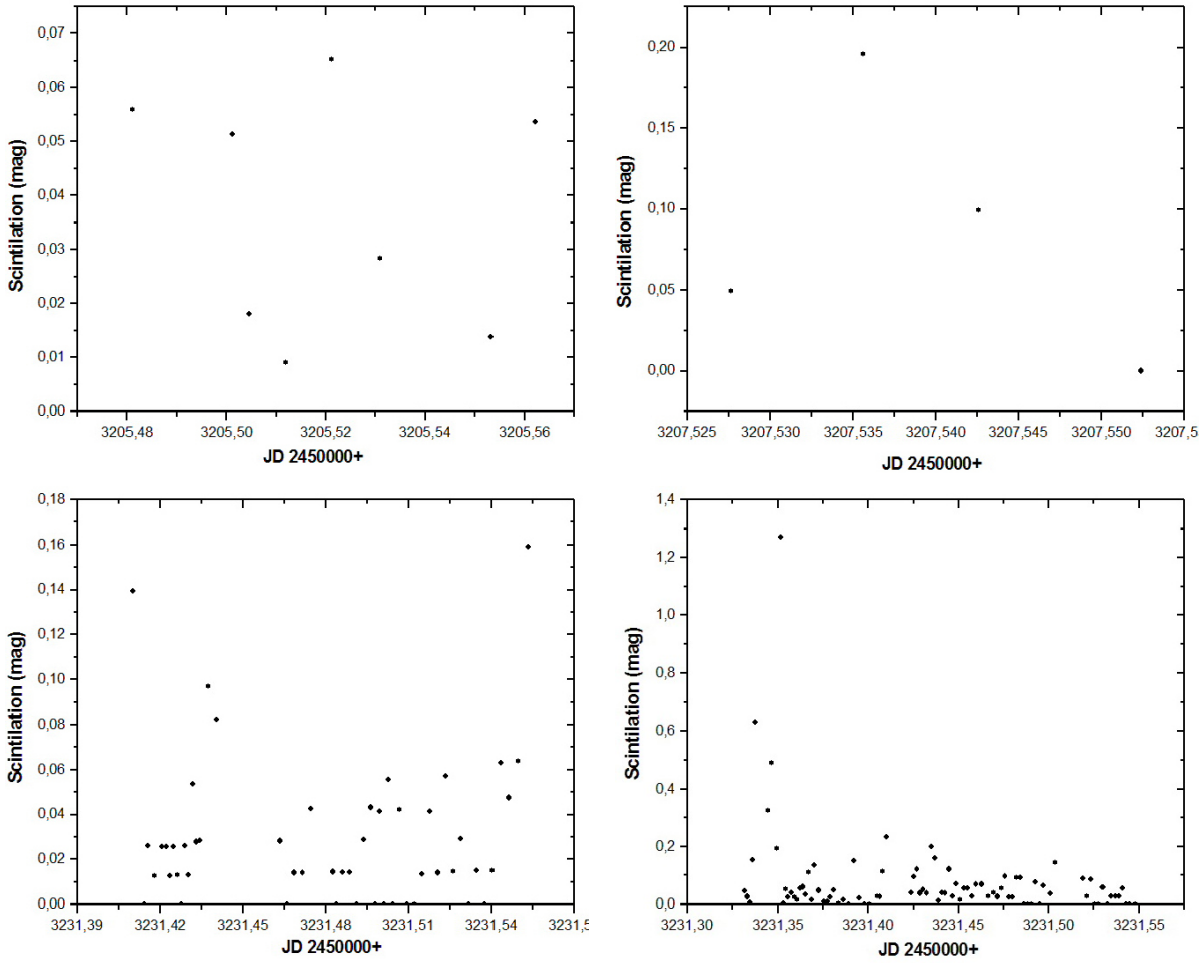

Figure 3.3: Diagram of scintillation index for the same nights as in figure 3.2

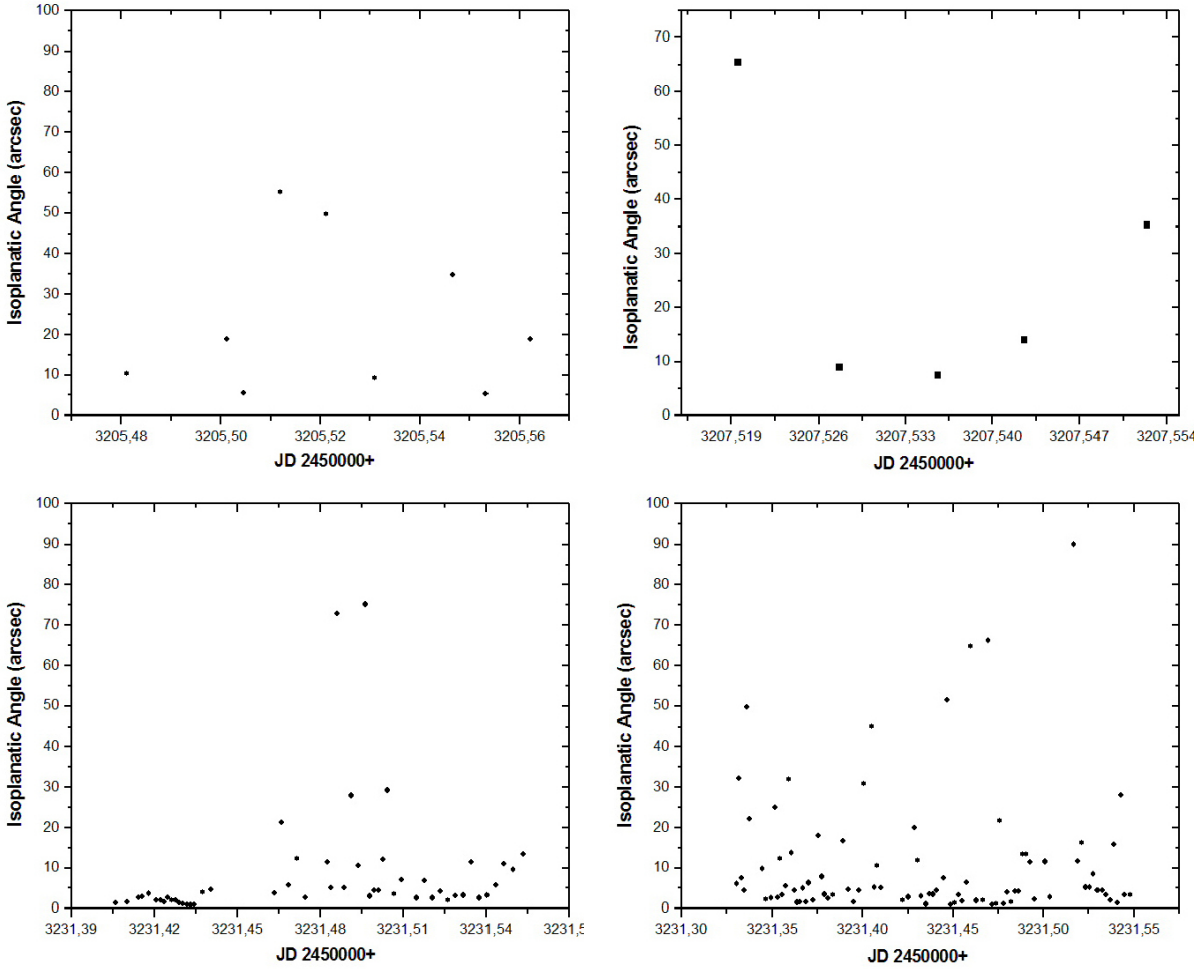

Figure 3.4: Diagram of Isoplanatic Angle for the same nights as in figure 3.2

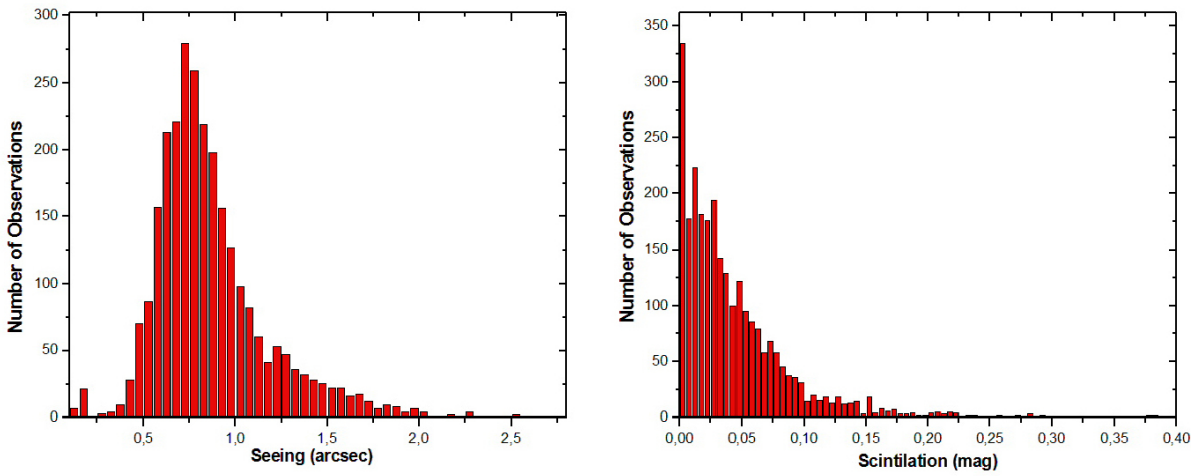

Figure 3.5: Histogram of the overall seeing and scintillation measurements

3.3 Instrumentation

Holomon Astronomical Station houses a wide variety of astronomical instruments. The main telescopes of the Station are a Meade LX-200 ($D = 250mm$, $F = 2000mm$, $f/8$) and a Celestron C-11 XLT ($D = 280mm$, $F = 2800mm$, $f/10$) with an addition of a few smaller supporting telescopes. EQ-6 Skyscan Pro mounts, with full autoguiding and computer connection capabilities. The main camera is a large format Fingerlakes PL6303E. The station is also equipped with a DCF-77 unit for accurate timing.

The selection of the instruments to be uses for the extrasolar planetary survey, was based mainly on the analysis described in Chapter 2, although, some practical issues, such as the weight tolerance of the mount were also considered. The finally selected combination consists of the Celestron C-11 and the Fingerlakes CCD camera with an addition of an $f/10 - 6.3$ focal reducer. A secondary small IMAGING SOURCE DMK CCD camera, mounted on a Konus telescope ($D = 114mm$,$F = 500mm$, $f/4.3$) is used for auto-guiding. Camera and mount operations are controlled via a Windows operating system (O/S) Workstation and the MAXIM DL program.

3.3.1 The CCD detector: Fingerlakes PL6303E

The Fingerlakes ProLine PL6303E camera houses a KODAK KAF 6303 CCD chip . The latter is a front illuminated sensor with an array of 3072×2048, 9 μm pixels. Its full well capacity is 100000 e^- and its peak quantum efficiency reaches 68% (Figure 3.6). The detailed specifications of the CCD sensor and the camera are shown in Tables 3.2 and 3.3.

Interface	USB 2.0
Digital Resolution	16-bit
Maximum Download Speed	8 MHz
Typical System Noise	9 e^-
Temperature Stability	0.1^o C
CCD Grade	2

Table 3.2: Camera Specifications

Array Size	3072 × 2048 pixels
CCD Type	Front illuminated
Coating	None
Sensor Diagonal	33.3 mm
Linear Full Well	100000 e^-
Anti Blooming	None
Peak Quantum Efficiency	68%

Table 3.3: Sensor Specifications

Laboratory tests

Before tcommencing the observations, the camera was tested for its linearity at the Laboratory of Astronomy. For the test, a uniformly illuminated field was created with the use of a white canvas and two defocused projection devices. The camera was mounted on the Celestron telescope and pointed to the canvas. Images of the uniform field were then acquired with various exposure times ranging from 1 to 12 seconds. In order to minimize possible systematics, caused by the variation of the field's luminosity, four sub-frames where taken at each exposure time. The images were also obtained with a random "exposure-time" order. After the acquisition, the frames where de-biased and dark subtracted. Each set of sub-frames was then merged into one single image with median stacking. The mean counts value of all pixels was then calculated and plotted versus exposure time [25] . The result is presented in Figure 3.7.

The results show a high degree of linearity over the full range of the camera's capacity. The least squares line is given by the equation

$$y(ADU) = (4383.89 \pm 11.34) * t(sec) + 364.758 \pm 83.5 \tag{3.1}$$

Usually the linearity is parameterized by the factor a which is defined as the slope of the line $ADUvsRate_{ADU}$, with $rate_{ADU} = \frac{ADU}{t_{exp}}$. The least square line is:

$$y(ADU) = -0.0029x + 4545.6 \tag{3.2}$$

A plot of equation 3.2 is given in Figure 3.8. We find that $\alpha = 0.0029$, which corresponds to a nearly perfect linear response.

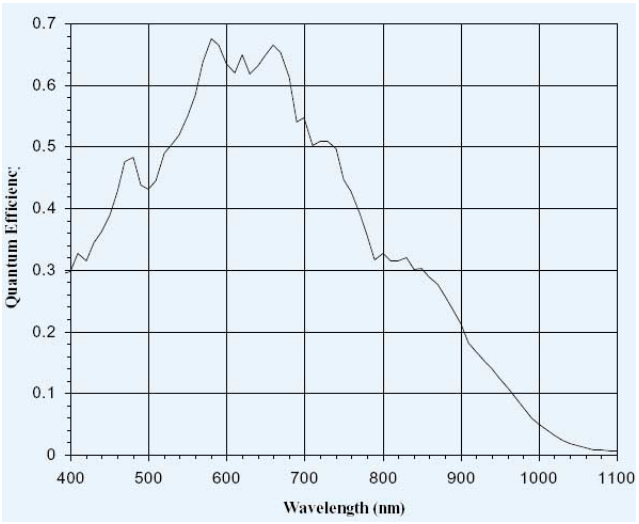

Figure 3.6: Quandum efficiency of KAF 6303 CCD chip

A problem that was noticed during the acquisition of the frames but was washed out with the reduction, is a low-level *shutter effect* at short exposures. However the problem is small and it is not expected to influence the observations.

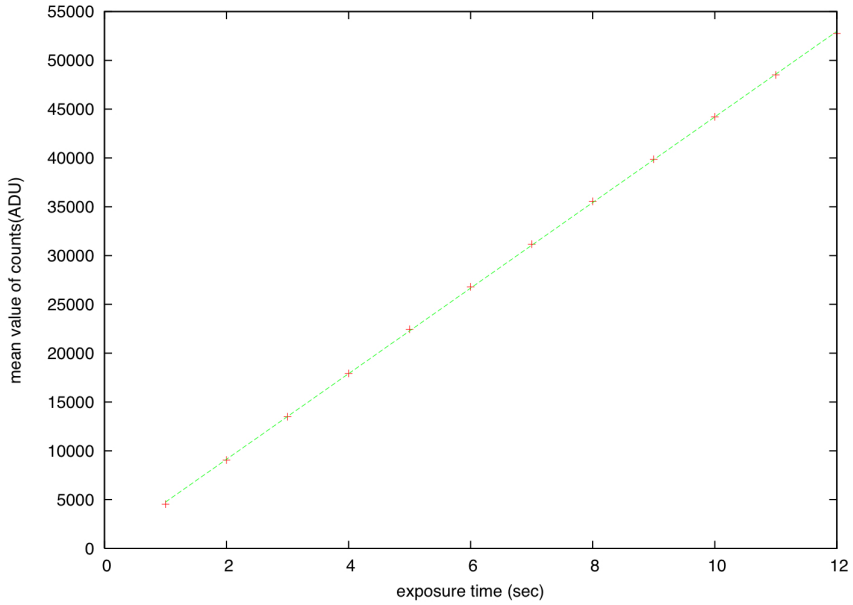

Figure 3.7: Linearity of KAF 6303 CCD chip

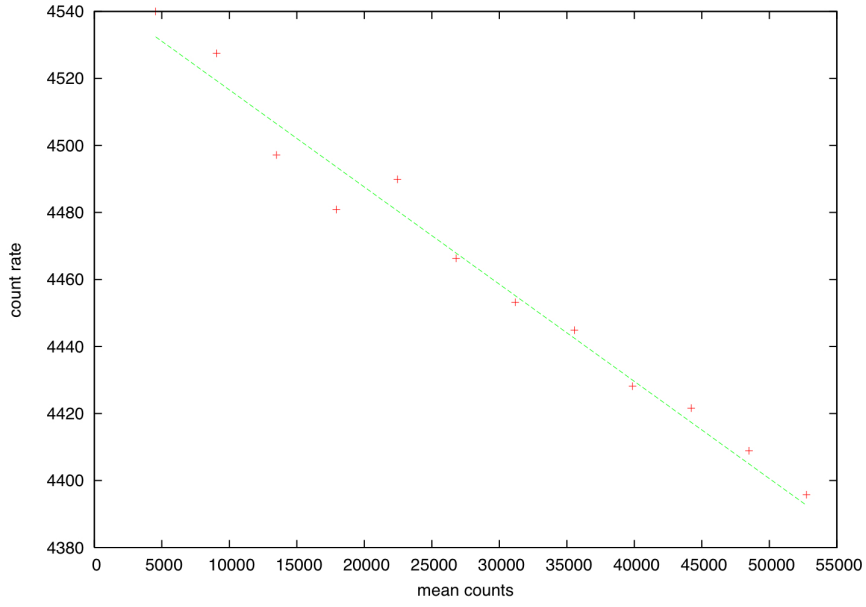

Figure 3.8: Linearity of KAF 6303 CCD chip

3.3.2 The telescope: Celestron C-11

The Celestron C-11, is a Schmidt-Cassegrain telescope with $D = 279.4$ mm, $f/10.02$ and $F = 2800$ mm. The model is popular among the amateur community because of its good quality of optics and relatively low price. Some basic characteristics of the instrument are given in Table 3.4.

Optical Design	Schmidt-Cassegrain
Aperture	279.4 mm (11 in)
Focal Length	2800 mm (110.24 in)
Focal Ratio	10.02
Focal Lenght with reducer	1760.22 mm
Focal Ratio with reducer	6.3
Field flattener	*yes*

Table 3.4: Telescope specifications

Tests of the telescope-camera system

Initially, the camera was adapted straight to the optical tube, without the interference of any other optical element. However, the first light images revealed a significant vignetting problem which is believed to be caused by the small exit pupil of the telescope comparing to the camera's diameter. In order to solve the problem , an $f/10$ to $f/6.3$ focal reducer was inserted between the chip and the optical tube. The latter enlarges the exit pupil of the telescope and it is also designed to work as a field-flattener. After this, the vignetting problem was reduced considerably and we were able to remove it during the data reduction. Some additional advantages include the enlargement of the field of view and the decrease of the exposure time by a factor of 3. A table with the basic parameters fot the telescope-camera system is given below:

Field of view	33.8×22.5 arcmin
Field of view with reducer	53.8×35.8 arcmin
Pixel scale (1×1 bin)	0.66 arcsec
Pixel scale with reducer (1×1 bin)	1.05 arcsec
Pixel scale (2×2 bin)	1.32 arcsec
Pixel scale with reducer (2×2 bin)	2.10 arcsec

Table 3.5: Telescope-camera system specifications

The camera was chosen to be operated at a 2×2 bining mode, in order to decrease the exposure and download time and the size of the images. That of course increases the pixel scale factor from 1.05 to 2.10 leading to a decrease of the number of stars that can be resolved

within a given field (see *Chapter 2*). However, 1×1 images are consuming $\backsim$ 4 times more computing time during reduction and thus, the 2×2 binning seems to be a good compromise. Moreover, the achived pixel scale reduces the size of the PSFs and hence the images are less oversampled, leading to better photometric precision.

Image quality

A visual examination of the frames, obtained with the Celestron telescope, indicated no significant coma error or vignetting. The shape of the PSF of the stars remains round at the edges of the frame.

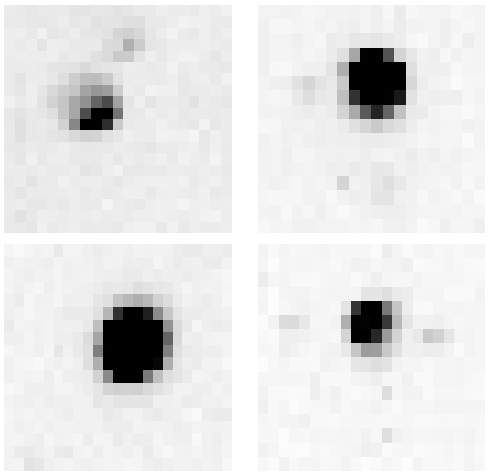

Figure 3.9: PSFs located at the edges of the CCD field.

A more careful examination with image reduction software exposed some small PSF variability over the surface of the CCD chip. A linear correlation was found between the FWHM of the PSFs and the distance from the center of the image. A least squared fitting gave

$$FWHM = (0.0016 \pm 0.00011)r + 0.749 \pm 0.061 \tag{3.3}$$

which is the analytical expression for the PSF size as a function of the distance from the center of the image. The slope of the line is very small (0.0016) which means that the variance is not expected to influence the precision of the photometry.

3.3.3 The mount and the auto-guiding system

The Celestron C-11 telescope is mounted on a SYNTA EQ-6 german equatorial mount, which can be controlled through a regular Windows O/S Workstation. The mount also supports auto-guiding via the standard SBIG ST-4 protocol. The pointing accuracy is $\backsim$ 20 arcmin and the periodic error of the gears is smooth with an amplitude of about $\backsim$ 15 arcsec.

For the guiding a Konus newtonian reflector ($D = 114mm$, $F = 500mm$, $f/4.3$) is used with the addition of an IMAGING SOURCE DMC CCD detector. Characteristics of the auto-guiding systems are given in Table 3.6.

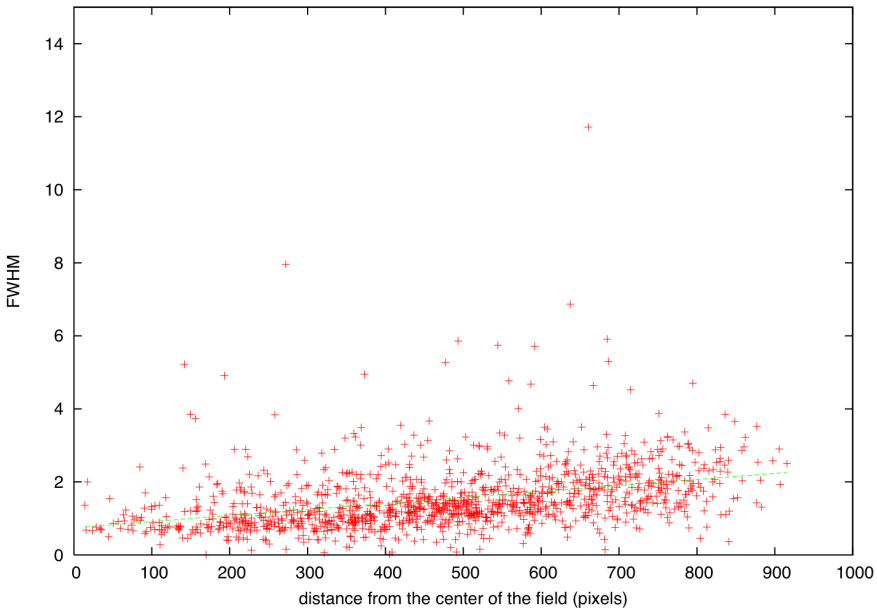

Figure 3.10: PSF size as a function of the distance from the center of the field

All pointing and guiding operations of the telescope are controlled via the commercial astronomy program MAXIM-DL. The auto-guiding algorithm allows for sub-pixel accuracy that reaches $\backsim 1/15$ of a pixel. This corresponds to 0.15 arc seconds or $\backsim 1/14$ of a pixel accuracy for the main instrument.

3.4 Observational parameters

Optical filters

Most of the well established transit projects survey the sky using either R or I filters in order to minimize trends caused by air-mass extinction. However, due to the relatively low quantum efficiency of the Fingerlakes CCD detector at these spectral regimes, we decided to use no filter for the observations and correct for air-mass during data reduction (see *Chapter 4*).

Exposure time

Through a series of tests we desided to use an exposure time ranging from 35 sec to 3 minutes with the exact value depending on the specific star field to be monitored. The magnitude range obtained is between $8 \leq m_v \leq 14$ (see Chapter 5).

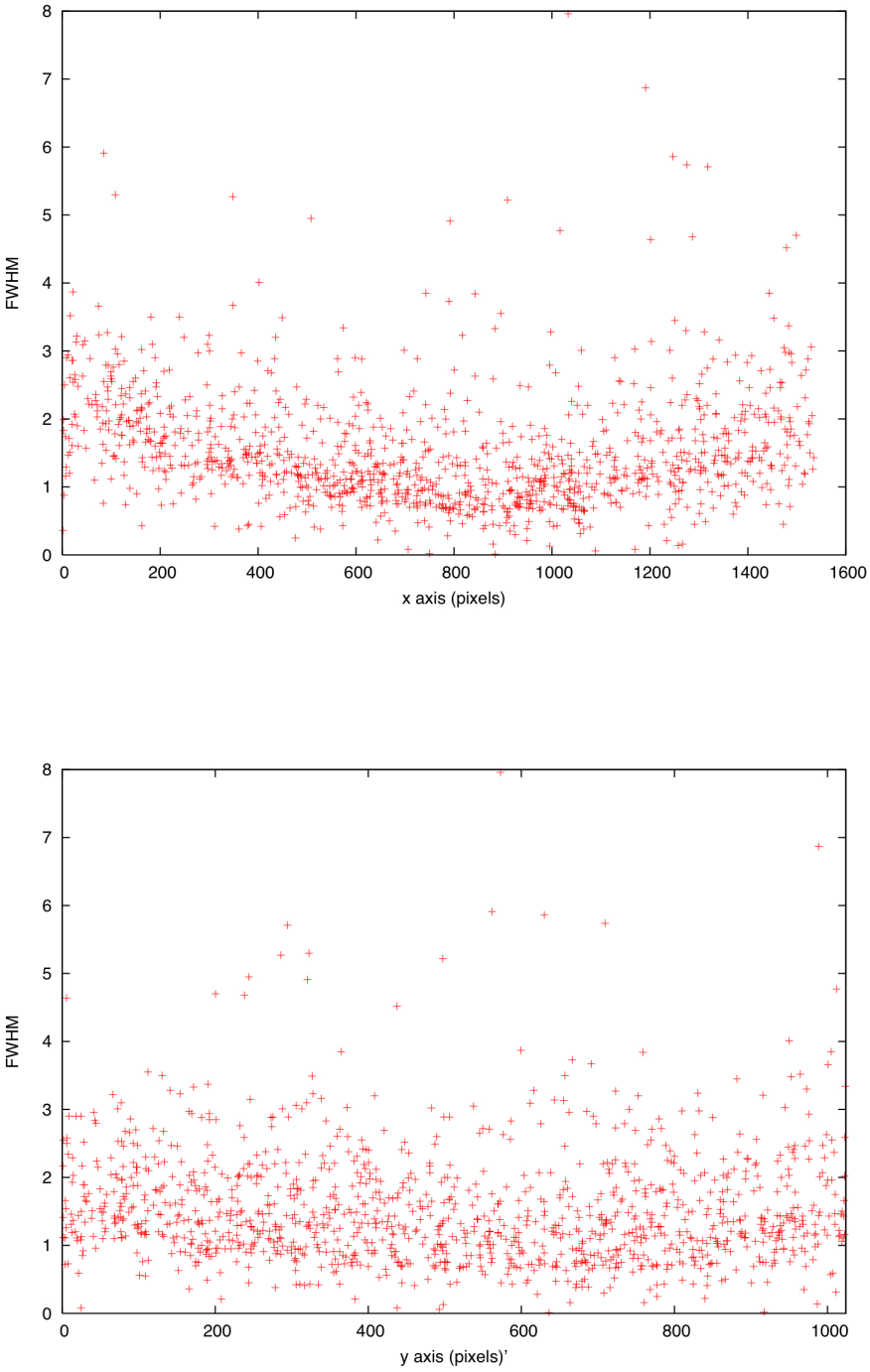

Figure 3.11: PSF variation over the x and y axis.

Pixel size	$5.6\mu m$
Focal Lenght	$500mm$
Pixel scale (1×1 bin)	2.31 arcsec
Pixel scale with reducer (1×1 bin)	1.05 arcsec
theoretical tracking accuracy	0.15 arcsec

Table 3.6: The auto-guiding system specifications

Target selection

The selection of the target fields was based on the analysis described in *Chapter 2*. A transit probability sky map was created with the use of the specific properties of our instruments (fFgure 3.12). With the use of the probability map, we located five favorable locations for the five runs of the survey. Specifications of the fields are given in the table below while visibility diagrams for each target are given in Figures 3.13-3.17

Field	Probability[%]	α	δ	Season
1	71.99	$22^h\ 53^m\ 00$”	$+44^o\ 30^m\ 00$”	Summer
2	67.24	$06^h\ 27^m\ 36$”	$+58^o\ 30^m\ 00$”	Winter
3	73.27	$17^h\ 42^m\ 00$”	$+20^o\ 30^m\ 00$”	Summer
4	80.95	$00^h\ 07^m\ 58$”	$+33^o\ 30^m\ 00$”	Autumn
5	61.15	$10^h\ 49^m\ 50$”	$+37^o\ 30^m\ 00$”	Spring

Table 3.7: Target fields of the Holomon planetary survey

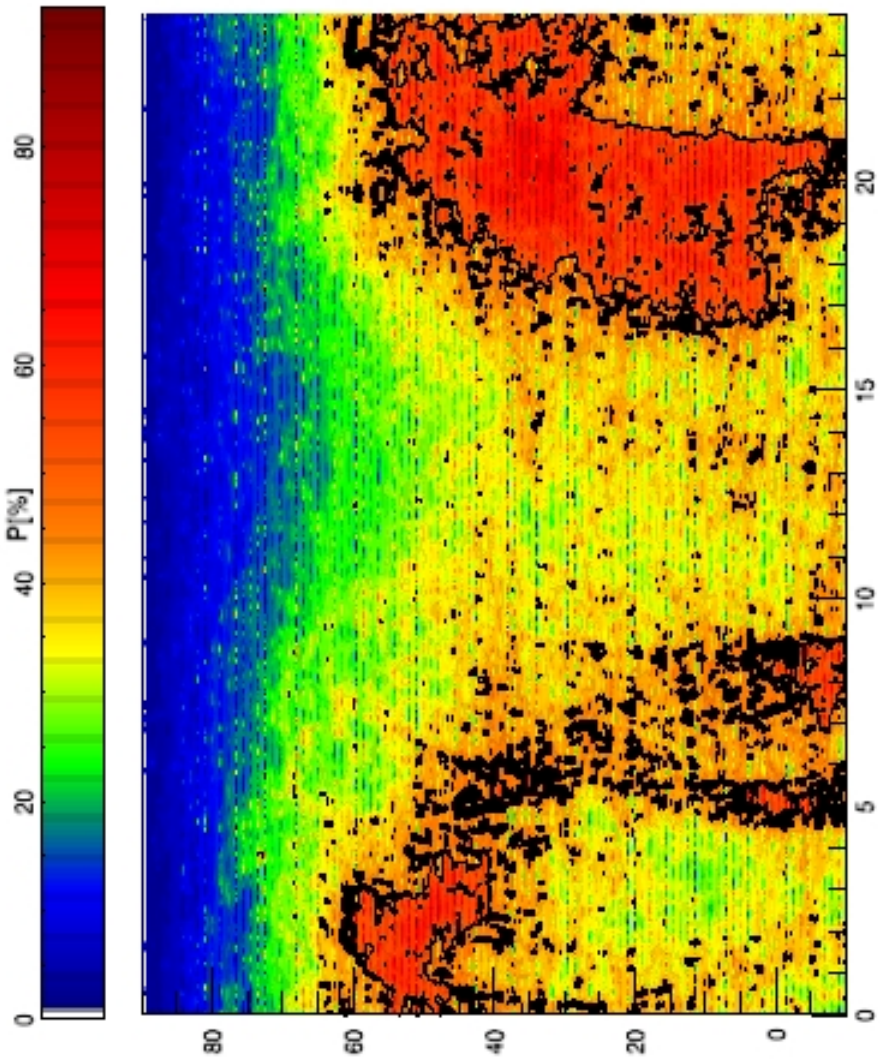

Figure 3.12: Probability map for the Holomon planetary survey. The interpretation was changed according to Chapter 2, and level lines were added in order to make the image sharper.

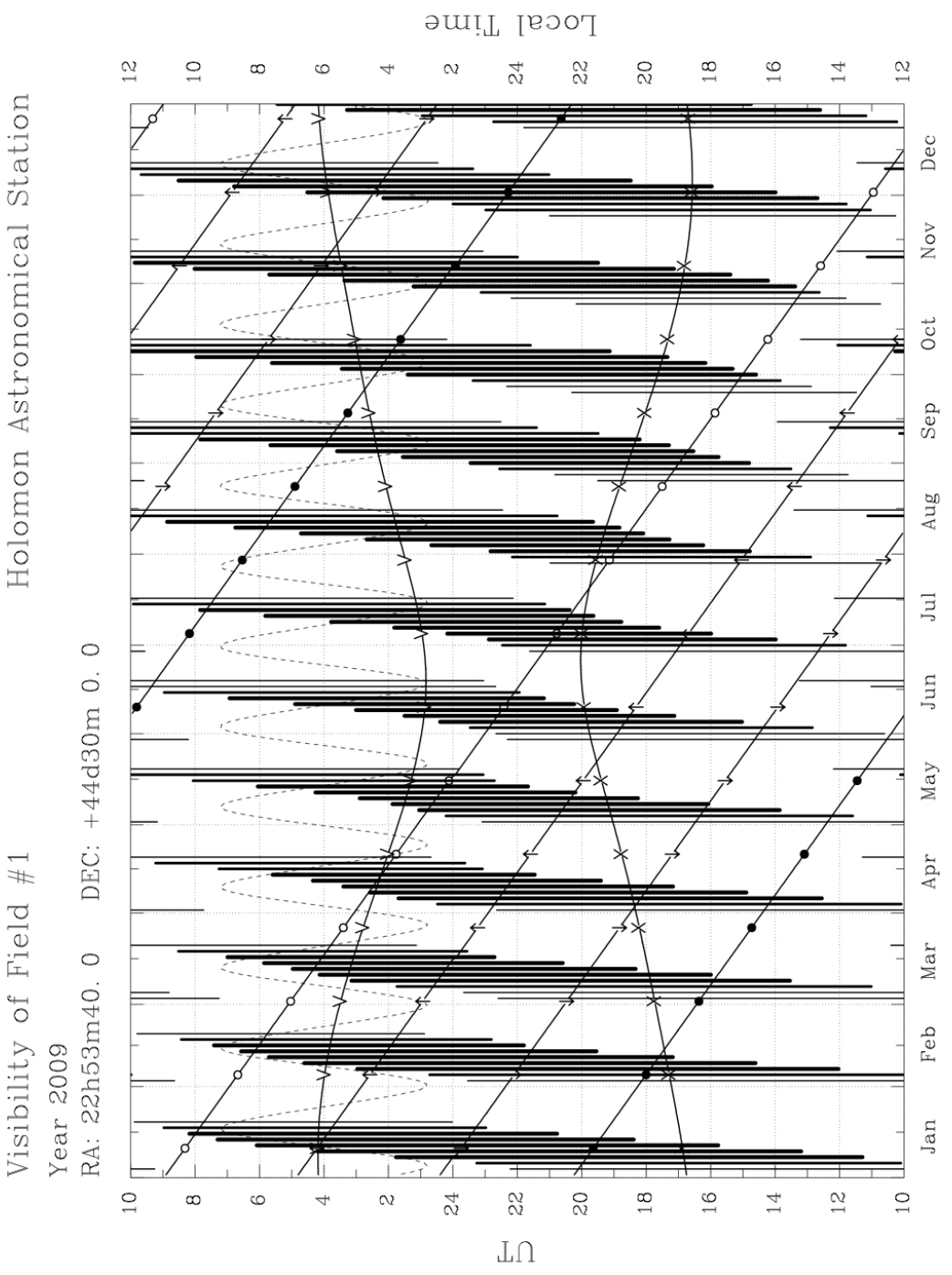

Figure 3.13: Visibility plot for Field 1

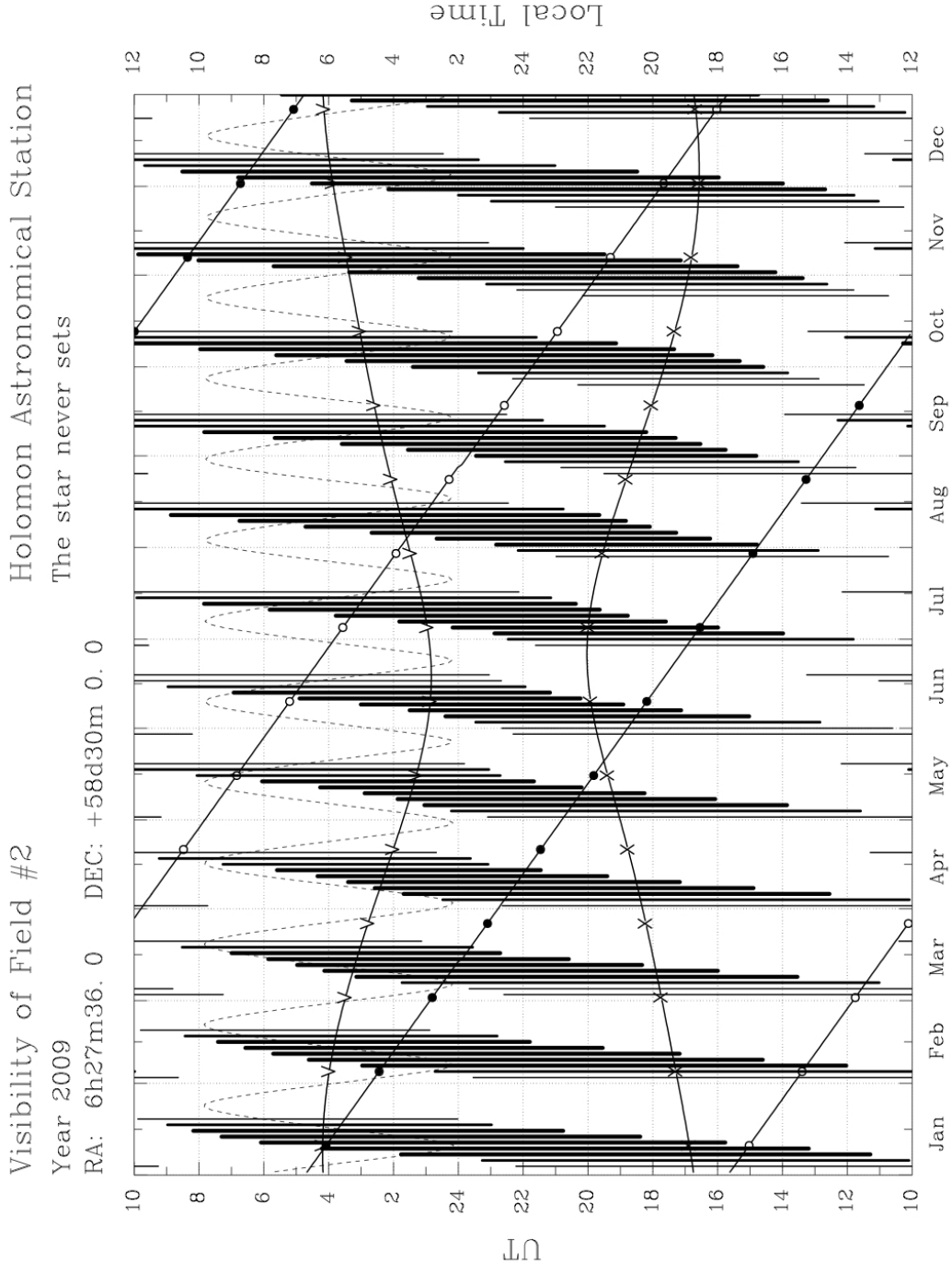

Figure 3.14: Visibility plot for Field 2.

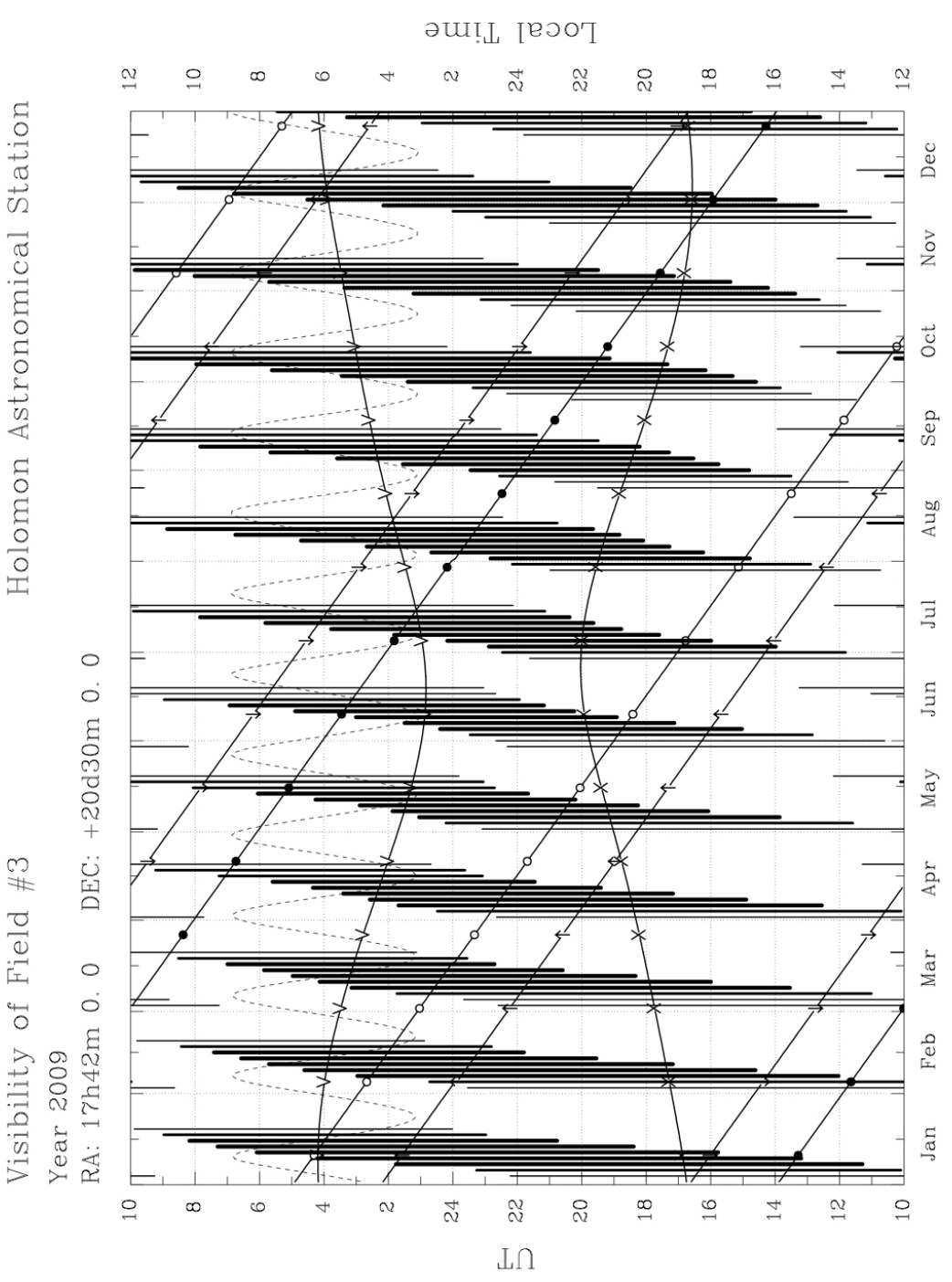

Figure 3.15: Visibility plot for Field 3.

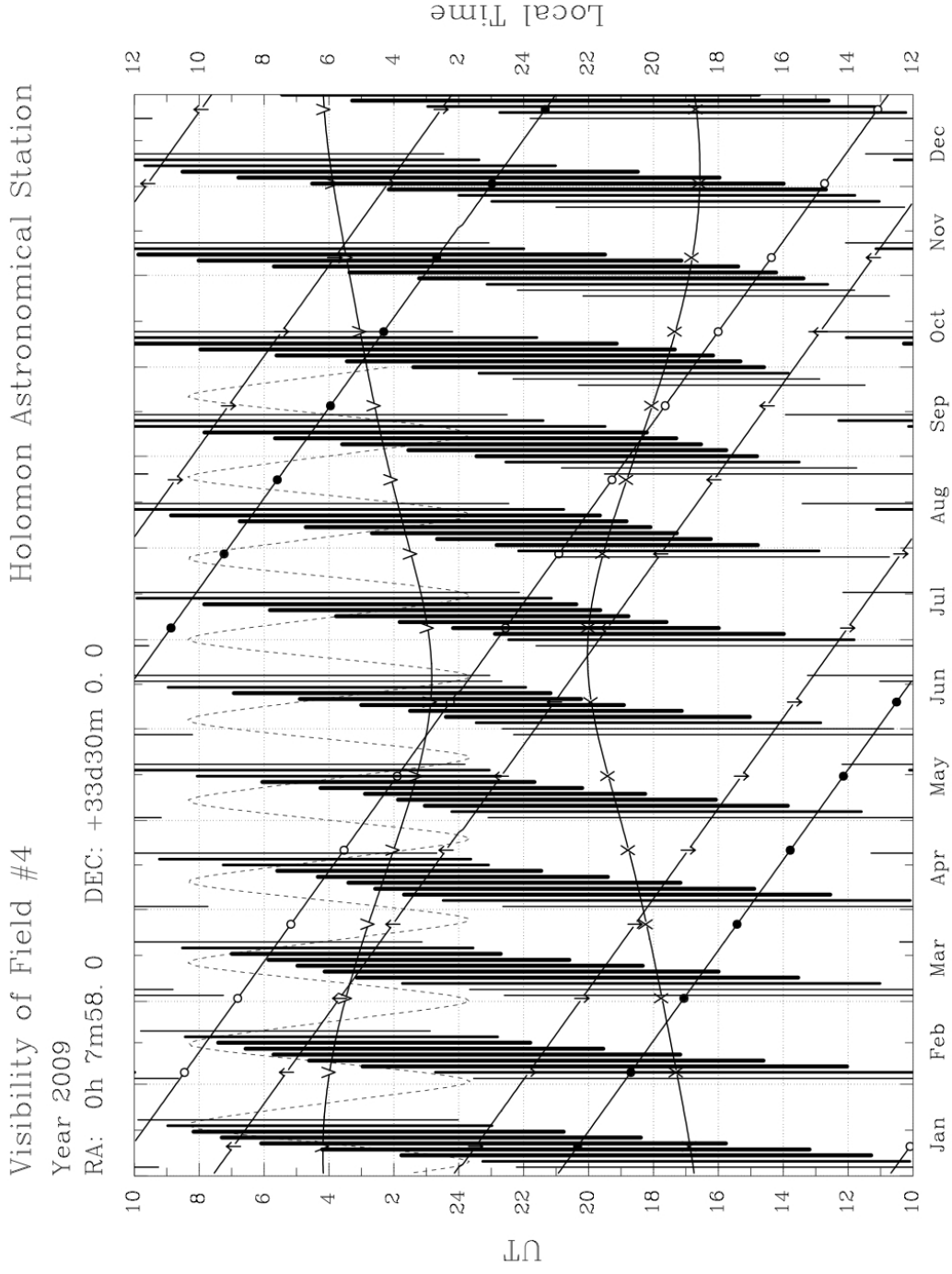

Figure 3.16: Visibility plot for Field 5.

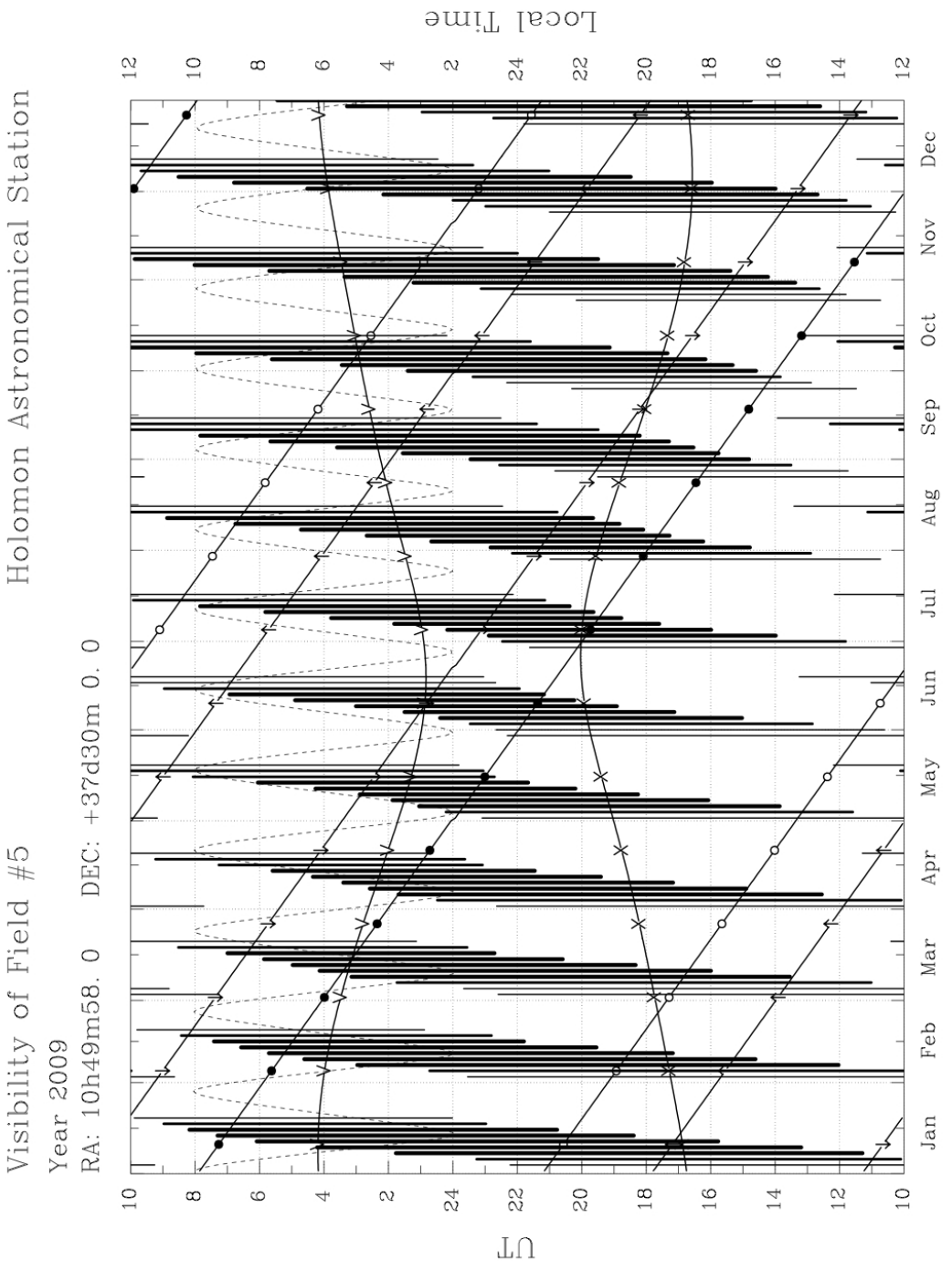

Figure 3.17: Visibility plot for Field 5.

Computer science is no more about computers than astronomy is about telescopes.
Edsger Dijkstra, Turing Award in 1972

4 The ThReT reduction pipeline

4.1 Introduction and overview of the pipeline

Perhaps the most important part of an extrasolar planetary survey is the data analysis that will derive the photometric light-curves. In the framework of this Diploma Thesis I have developed the *"Thessaloniki Research for Transits"* (ThReT) pipeline. The ThReT pipeline is able to perform precision photometry on dense star fields. Although it was build to match the needs of the Holomon Transits Survey, its performance was tested successfully on data-sets acquired with other optical instruments such as the Hamburg 1.2 m SC telescope.

The pipeline consists of a series of c-shell scripts and fortran routines. It uses standard STARLINK libraries for image reduction and the DAOPHOT facilities to perform photometry. The pipeline was developed before the final instrumental setup was decided, and thus it is designed to be able to perform both aperture and PSF photometry for the use with a wide variety of instruments. After the light-curves are accummulated, I use *Ensemble Photometry*, as well as the *Sys-Rem* (*Tamuz et al.*) and *TFA* (*Kovacs et. al.*) algorithms for systematic errors removal. In the final part of the analysis, the light-curves are scaned for variable stars using statistical analysis tools, as well as for transit-shape signals using the *Box-Fitting Least Squares* algorithm and a practical self developed *Fast Detection Algorithms* that can easily find single-night events.

Here I chose to interpret the pipeline script-per-script with an addition of essential comments on theory where needed. The code is available from the author by request.

4.2 Preliminary reduction

4.2.1 thret.csh

The pipeline initiates with the call of *thret.csh* script. The first job of thret.csh is to initialize STARLINK and call all the necessary libraries (FIGARO, KAPPA, SEXTRACTOR, PISSA, DAOPHOT). Then, the script, opens a user-interactive screen where the user is prompt to give some informations to the pipeline. The answers are stored into variables for later use. Specifically, the program asks for:

- Directory name that contains the data (*eg "/home/Desktop/data"*)

- RA coordinate of the center of the field(*eg "20 32 04"*)
- DEC coordinate of the center of the field(*eg "+04 22 05"*)
- Main name of the Bias Frames (*eg "bias-" or "Bias" or "bias"*)[1]
- Extension of the Bias Frames (*eg "fit" or "FIT" or "FITS"*)
- Main name of the Dark Frames (*eg "dark" or "DARK"*)
- Extension of the Dark Frames
- Main name of the Flat Frames (*eg "Flat-" or "FLAT" or "Flat"*)
- Extension of the Flat Frames
- Main name of the Science Frames (*eg "Frame-" or "ngc129-" or "frame"*)
- Extension of the Science Frames
- Value for linearity correction (*constant α as defined in Equation 3.2. It is set to zero if no information about linearity excists*

After storing these parameters, the pipeline is ready to start the reduction of frames. *thret.csh* script, enters the directory that contains the data and calls the next script.

4.2.2 convert.csh

convert.csh simply converts all *.FIT* images to the *.sdf* standard STARLINK format. The initial images are kept unaltered and the pipeline from now on changes only the *.sdf* frames. *convert.csh* finally calls the next script for image reduction.

4.2.3 reduce.csh

reduce.csh is the image reduction script and its operation is to correct the science frames from noise and prepare them for photometry.

bias subtraction

Initially, the script calls the fortran routine *biascheck.for* and checks if the bias frames suffer from any kind of anomaly (eg from cosmic rays), and if they do it rejects them. Then, the remaining bias frames are merged into one master-Bias frame by median averaging and subtracted from dark, flat and science frames.

[1]The pipeline only accepts files of the specific form: *text*[*number*].*extention*, eg. *Bias* – 024.*fit* or *frame*345.*fit*

dark subtraction

After the bias subtraction, dark frames are normalized to $t_{exp} = 1$ and merged into one single master-Dark frame by average combining. The master-Dark frame is scaled to match the exposure of flat and science frames respectively and then subtracted from both.

linearity correction and flat frames

After the dark noise subtraction, both flat frames and science frames are corrected for CCD linearity by:

$$M = C(1 + C \cdot \alpha) \tag{4.1}$$

where M is the measured intensity of a pixel, C is the corrected value and α is the linearity coefficient.

flat fielding

At the final step, the flat frames are normalized to the count value of 1 and median combined to make the master-flat frame. All science images are then divided by the master-flat frame. After the image reduction is complete, the pipeline stops, in order to allow the user to prepare the necessary files for the photometry.

4.3 Photometry

Photometry is performed by *DAOPHOT II: The Next Generation*, which was written by *Peter Stetson* at the Dominion Astrophysical Observatory. Before calling the script that will perform stellar photometry at all images, the user must select a reference image and modify some DAOPHOT parameters inside the *allstar.csh* script. Explicitly, these parameters (as described in DAOPHOT manual) are:

- READ NOISE : The readout noise of the CCD detector in count units
- GAIN : The gain factor of the CCD detector
- LOW GOOD DATUM: The level, in standard deviations below the mean sky value, that a pixel will be considered defective
- HIGH GOOD DATUM: The level in Counts above which a pixel value is considered defective
- FWHM : A representative value for the mean FWHM of the PSFs in the frame
- THERESHOLD : The level in standard deviations, above the mean sky value, that a signal will be considered as a detection

- FITTING RADIUS: This value defines the circular area around a star, within which the PSF model will be fitted
- PSF RADIUS: The radius, in pixels, of the circle within which the point-spread function is to be defined
- VARIABLE PSF: This variable defines the degree of complexity with which the PSF function is to be modeled. The choices are : VARIABLE PSF=0, for a constant model, VARIABLE PSF=1, for a point-spread function which varies linearly with position in the frame, VARIABLE PSF = 2 for a quadratic variation and VARIABLE PSF = -1 for an analytic PSF model.This value is usually set to 0 when there is no information about the PSF variability or equal to 2 for the Holomon Transit Survey frames (*see Figure 3.11*).
- ANALYTIC MODEL PSF: This defines the analytic function to be used for the PSF estimation. The user has to choose between 6 functions:
  1. A Gaussian function, with axis aligned with the x and y directions in the image
  2. A Lorenz function, having three free parameters, FWHM at x and y axis and a position angle for the major axis of the ellipse
  3. A Moffat function, $MF = \frac{1}{(1+z^2)^\beta}$ with $z = x^2/a_x^2 + y^2/a_y^2 + a_{xy}xy$ and $\beta = 1.5$.
  4. A Moffat function but with $\beta = 2.5$
  5. A Penny function, which is the sum of a Gaussian and a Lorenz function, having four free parameters, the FWHM in x and y; the fractional amplitude of the Gaussian function at the peak of the stellar profile and the position angle of the tilted elliptical Gaussian.
  6. A Penny function with one more free parameter, the position angle of the Lorenz function
- The radius of the inner, outer and sky circles for the aperture photometry. DAOPHOT allows for 11 inner circles (A1-A9, AA, AB and AC), one inner sky circle (IS) and one outter sky circle (OS)

After modifying the script, the user must run manually DAOPHOT once on the reference image, in order to determine the PSF function. A typical run of DAOPHOT is as follows:

```
5.53 <-Readout noise
1.28 <-Gain
OPTIONS
KEYBOARD  INPUT
FWHM=2.0
TH=4.0
LOW=4
```

```
FI=6.0
PSF=8.0
MS=50000
HI=65000
VA=+2
ANALYTIC = +2 <- This corresponds to the Lorenz function and it is
                      preferred  because of possible  guiding errors that
                      make the point spread  function elliptical and inclined
ATTACH reference
FIND
1,1
ref.po
Y

PHOTOMETRY
photo.opt
A1 =0.5
A2 =1.5
A3 =2
A4 =3
A5 =4
A6 =4.5
A7 =5
A8 =6
A9 =7
AA =8
AB =8.5
IS =9.5
OS =14
OK
reference.po
reference.ap <- The file where the aperture photometry results are stored

PICK  <- This routine selects the stars to be used for the PSF estimation
reference.ap
50 <- Number of the desired stars to be used for modeling.
           This number is a function of the expected variance to be
            detected. It is usually set between 30 and 99
           reference.lst

PSF <- This routine calculates the PSF
reference.ap
reference.lst
```

```
ref.psf <- The file where the PSF is stored
EXIT
```

4.3.1 allstar.csh

After changing the DAOPHOT parameters for each frame and determining the PSF function, the pipeline continues with the call of *allstar.csh* script. Initially the script calculates the mean $FWHM$ of the stars in each frame and set this value to the FWHM parameter of DAOPHOT. The $FITTING\ RADIUS$ is also adjusted acoording to that value as $FR = 1.2 \cdot FWHM$. Then aperture and psf photometry is performed to all stars at each frame. The program outputs two different files for each frame: [$data - name$].ap, which contains the results of the aperture photometry and [$data - name$].als, which contains the results of the PSF photometry. The format of the two files is given bellow:

$*.ap$ files:

```
 NL   NX   NY  LOWBAD HIGHBAD  THRESH     AP1  PH/ADU  RNOISE
  2 1536 1024   144.1 32766.5   57.02    0.50    1.28    5.53

187  408.390  101.870   20.446   18.805   18.164
17.924   17.809   17.684   17.494   17.390   17.420
196.134 12.34  0.13   0.235    0.134    0.121    0.156
 0.181    0.201    0.205    0.223    0.249

188  605.500  102.820   20.300   18.155   17.383
 16.344   15.508   14.656   14.119   99.999   99.999
 200.140 12.87  0.01   0.216    0.081    0.065
 0.041    0.025    0.015    0.011    9.999    9.999
```

$*.als$ files:

```
 NL   NX   NY  LOWBAD HIGHBAD  THRESH     AP1  PH/ADU  RNOISE    FRAD
 1 1536 1024  5965.4 65000.0  315.55    0.50    1.28    5.53    6.00

61  594.867   16.261   19.644    0.093 6229.525  4.  0.81  0.818
174  692.705   43.623  19.534   0.110 6230.427  4.  1.03  1.471
```

4.4 Interpretation

The next step, after performing photometry, is of course to derive the light-curves of each star.

4.4.1 timer.csh

timer.csh script calculates the Julian date each image was captured at, as well as the air-mass of the field at the time of the exposure. The first thing to do, is to extract the date and the time of the exposure from each image header. An example of a header file is given below. Date and time information are placed in line 9:

```
SIMPLE  =                    T
BITPIX  =                   16 /8 unsigned int, 16 & 32 int, -32 & -64 real
NAXIS   =                    2 /number of axes
NAXIS1  =                 1536 /fastest changing axis
NAXIS2  =                 1024 /next to fastest changing axis
BSCALE  =   1.0000000000000000 /physical = BZERO + BSCALE*array_value
BZERO   =   32768.000000000000 /physical = BZERO + BSCALE*array_value
INSTRUME= 'Finger Lakes' /      instrument or camera used
DATE-OBS= '2008-07-11T23:37:17' /YYYY-MM-DDThh:mm:ss observation start, UT
EXPTIME =   75.000000000000000 /Exposure time in seconds
EXPOSURE=   75.000000000000000 /Exposure time in seconds
SET-TEMP=  -35.000000000000000 /CCD temperature setpoint in C
CCD-TEMP=  -35.257951796054840 /CCD temperature at start of exposure in C
XPIXSZ  =   18.000000000000000 /Pixel Width in microns (after binning)
YPIXSZ  =   18.000000000000000 /Pixel Height in microns (after binning)
XBINNING=                    2 /Binning factor in width
YBINNING=                    2 /Binning factor in height
XORGSUBF=                    0 /Subframe X position in binned pixels
YORGSUBF=                    0 /Subframe Y position in binned pixels
IMAGETYP= 'Light Frame' /       Type of image
FOCALLEN=  0.00000000000000000 /Focal length of telescope in mm
APTDIA  =  0.00000000000000000 /Aperture diameter of telescope in mm
APTAREA =  0.00000000000000000 /Aperture area of telescope in mm^2
SWCREATE= 'MaxIm DL Version 4.51' /Name of software that created the image
SBSTDVER= 'SBFITSEXT Version 1.0' /Version of SBFITSEXT standard in effect
```

After extracting date and time, *timer.csh* calls the fortran routine *time.for*, which calculates the Julian date via:

$$JD = 367 \cdot YEAR - int(((YEAR + int((MONTH + 9)/12))) \cdot 1.75)$$
$$+int((MONTH) \cdot 30.55555) + DAY + 1721013.5 + (UT/24) \quad (4.2)$$

and airmass via[2] :

$$air = a - 0.001867(a-1) - 0.002875(a-1)^2 - 0.0008083(a-1)^3 \quad (4.3)$$

[2]Note that for the calculation of air-mass, the variables *longinute* and *latitude* need to be changed within the fortran script, in order to correspond to the actual place of observations.

where

$$a = \frac{1}{sin(la)sin(dec) + cos(la)cos(dec)cos(hourangle)};$$

$$hourangle = t^{sidereal}_{Local} - RA;$$

$$t^{sidereal}_{Local} = t^{sidereal}_{grenwitch} - longnitude + 1.00274UT$$

and

$$t^{sidereal}_{grenwitch} = 6.69738 + 0.0675[LD - (UT/24) - 2451545]$$

Julian date and air-mass values are stored to the files [*data−name*].*jd* and [*data−name*].*air* respectively.

4.4.2 ref.csh

The next script modifies the DAOPHOT output files and prepares them for object matching (see 4.4.3). First, the user is promed to chose between aperture and PSF photometry. In case of aperture photometry, the user also needs to decide between one of the eleven apertures $A1 - AC$ (daophot output file includes all sky appertures). Otherwise, the script continues to the next operation.

The next job of *ref.csh* script, is to cut DAOPHOT output files, to make them suitable for the *match* − 0.8 routine. The new lists are stored under the name [*data* − *name*].*list*. An example of *ref.csh* outputs is given below:

```
1 453.886 3.461 208740125
2 1383.766 24.498 183360034
3 775.071 22.495 205240099
4 963.270 24.549 203870104
5 730.425 26.589 204370118
```

The first column is the id of each object, while the second and third columns are its *x* and *y* cordinates on the CCD frame. The last column includes information about both the instrumental magnitude of the star and the error of the measurement. The codification is as follows:

The first two digits represent the integer part of the magnitude while digits 3 − 5 are its first three decimal. The next four numbers are the decimal digits of the error.

This unorthodox interpretation was chosen at this point because of a bug of the external routine *match* − 0.8. The latter only accepts lists with 4 columns and this seems to be the only way to include error information. The original form of *magnitude* and *error* is restored at a later point.

Finally, the script choses the file with the maximum number of entries as reference list.

4.4.3 matcher.csh

Matcher.csh uses the external routine *match* − 0.8 (*Valdes et. al.* 1995), to match the entries of $*$.*list* files with those of the reference list. The algorithm allows for arbitrary translation,

rotating and scaling , and it require N^6 operations, where N is the number of entries at each file. $Match - 0.8$ produces two kids of output files: The one, $[data - name].mtA$, contains all the matched stars and the other, $[data - name].unA$, all the unmatched stars. The first list consists of 3 columns, just like the input file, but the ID number of each entry is changed to that of the same object at the reference list.

4.4.4 curves.csh

This script generates the raw light-curves by cross-correlating the $*.mtA$ output files of the $match - 0.8$ program. First a plain file for each individual star of the reference list is generated under the name $ID.curve$, where ID is the identity number of each star (first column). Then the script loops over the $*.mtA$ files, and whenever a star is found to exist, it is passed to the relative $ID.curve$ file, together with the opposite julian date and air-mass. The 4^{th} column of the $*.mtA$ file is also split into the original *mag* and *error* columns.

As a last step, the pipeline rejects all the light-curves that contain 60% less points than the total number of exposures.

4.5 Post reduction

4.5.1 Statistics

[3] A key step for all subsequent analysis is the calculation of basic statistical values of the data sample. We first derive the error-weighted mean magnitude of each light-curve by

$$M = \frac{\sum_{i=1}^{n} \frac{m_i}{\sigma_i^2}}{\sum_{i=1}^{n} \sigma_i^2} \tag{4.4}$$

where m_i is the i^{th} exposure and n is the total number of exposures. Then, the *standard deviation* and *Root Mean Square* values are calculated via

$$\sigma = \sqrt{\frac{1}{n} \sum (m_i - M)^2} \tag{4.5}$$

and

$$rms_{log} = \log\left(\sum \left(\frac{m_i - M}{\sigma} \right)^2 \right) \tag{4.6}$$

For reasons that will become obvious below, we also derive the *inner pulsation index* and the *pulsation factor* by

$$\sigma_{in} = \sqrt{\frac{1}{2(n-1)} \sum_{i=1}^{n-1} (m_i - m_{i+1})^2} \tag{4.7}$$

[3] Until now, we have adopted a script-per-script interpretation of the pipeline. From this point, we will give more emphasis on theory. All scripts and fortran routines are given in Appentix A

and

$$P = \frac{\sigma}{\sigma_{in}} \tag{4.8}$$

The inner pulsation index is an indicator of the coherence of the light-curve, while the pulsation factor illustrates the variability of a light-curve. At this point, the curves are suffering from trends (red noise) and therefore a high value of P might not reflect an endogenous variability of the star. However, equation 4.8 can be used after the noise removal procedure, as a fast indicator of stellar variability. The calculated statistics are stored into the file *sigma.txt* together with the *id* name of each light-curve. As a first step to the post reduction procedure, the missing points of each light-curve are replaced with the mean value (see equation 4.4), of error 4.5, so that every light-curve of the sample has the same number of points.

An example diagram of M vs σ (*sigma* plot refered as hereafter), is given at Figure (4.1). As it can be seen, the σ value of the light-curves increases as the signal drops (magnitude rises). This trend reflects the poison noise (white noise) which increases with $\backsim n$, where n the number of photons captured from the CCD detector. The diagram also contains information about the systematic noise that dominates the data. A non-coherent plot, with high deviated points, illustrates a relatively high level of trends. On the other hand, divergent points of a tenacious curve, are usually identifiers of variable stars.

The sigma plot is a very important diagram that is used by the *ThReT* pipeline for further investigation of the data (see sections 4.5.4, 4.5.4 and 4.5.7)

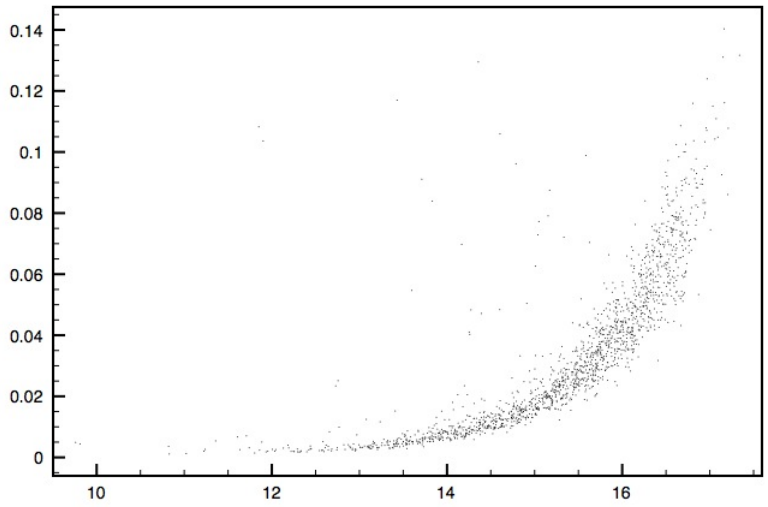

Figure 4.1: An example of sigma vs magnitude plot

4.5.2 Red Noise and its possible sources

By the term *Red Noise* (or trends), we generally refer to the systematic effects (namely, the changes in signal that are not caused by an endogenous variability), that dominate the raw

light-curves. Red noise might be caused by a range of different factors, such as uncorrected instrumental effects or a change of observing conditions. It may also be intrinsic to the data, or originate from an imperfect data reduction.

The most common red-noise generator, is the change of conditions during the observations. Passing clouds, changes of the wind's velocity as well as changes of atmospheric turbulence, are causing deeps and rises to the signal. Even under perfect, stable atmospheric conditions, the target crosses atmospheric layers of different thickness, due to the earth's rotation. The latter means that the incoming photons are submitted to different atmospheric extinction throughout the night. Moreover, since extinction depends on wavelength, it will differ between stars of different color.

The instrumental sources of red noise include guiding errors, optical errors of the instrument as well as pixel-to-pixel and inter-pixel variations.

In general each individual star is dominated by different trends. These trends are a function of the color of the star and the occupying area on the CCD chip. Although all trent reduction algorithms that will be discussed below take as an assumption that at least two or more lightcurves are biased by the same red-noise.

The efficiency of signal detection is highly correlated with the noise on the light-curve and therefore a successful trend minimization on the lightcurves of our survey becomes a necessity. Below we employ and test some of the most commonly used de-trending algorithms.

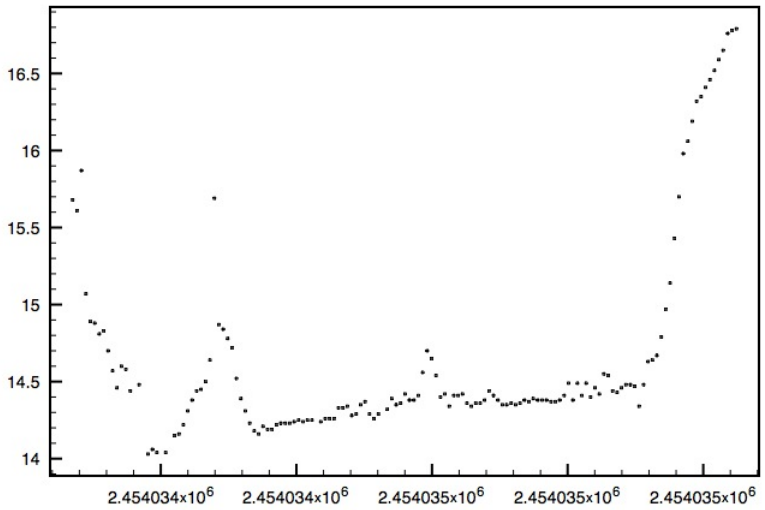

Figure 4.2: An example of an error dominated light-curve

4.5.3 Differential photometry

The oldest, easiest, and most widely used method to remove systematics from a light-curve is *differential photometry*. In differential photometry we assume that the observed variation

on a light-curve of a known constant star, is reflecting the red noise. The lightcurves are also assumed to be affected by the same trends. The red noise subtraction is then simply given by

$$m_0(s) = m(e, s) - em(e) \tag{4.9}$$

where $m_{e,s}$ is the raw magnitude of the star at e exposure, to be corrected and $em(e)$ is the magnitude of the comparison star.

Despite its simplicity, differential photometry works well on small uncrowded fields. When it comes to large scale photometric data, this method becomes inefficient for the reasons explained above. An improvement of differential photometry is *ensemble photometry* (ref), where multiple comparisons are used for the determination of trends.

4.5.4 Ensemble photometry

This method employs ss comparison stars rather than one. If e is the number of exposure, ee the total number of exposures and s is the comparison star, then the corrected magnitude $m_0(s)$, is given by

$$m(e, s) = m_0(s) + em(e) \tag{4.10}$$

as in differential photometry. We now want to minimize the deviation of the signal of all constant stars from their mean magnitude value. The quantity to minimize is given by

$$\beta = \sum_{e=1}^{ee} \sum_{s=1}^{ss} [m(e, s) - m_0(s) - em(e)]^2 w(e, s) \tag{4.11}$$

where $w(e, s)$ is the weighting factor of $m(e, s)$. The number of free parameters to solve for are $ee + ss$, and can be expressed as:

$$\begin{array}{rcccccccc}
\frac{\partial\beta}{\partial em(1)} & = & \sum_{s=1}^{ss} w(1,s)em(1) & - & \sum_{s=1}^{ss} w(1,s)m(1,s) & + & \sum_{s=1}^{ss} w(1,s)m_0(s) & = & 0 \\
\vdots & & & & & & & & \\
\frac{\partial\beta}{\partial em(ee)} & = & \sum_{s=1}^{ss} w(ee,s)em(ee) & - & \sum_{s=1}^{ss} w(ee,s)m(ee,s) & + & \sum_{s=1}^{ss} w(ee,s)m_0(s) & = & 0 \\
\frac{\partial\beta}{\partial m_0(1)} & = & \sum_{e=1}^{ee} w(e,1)m_0(1) & - & \sum_{e=1}^{ee} w(e,1)m(e,1) & + & \sum_{e=1}^{ee} w(e,1)em(e) & = & 0 \\
\vdots & & & & & & & & \\
\frac{\partial\beta}{\partial m_0(ss)} & = & \sum_{e=1}^{ee} w(e,ss)m_0(ss) & - & \sum_{e=1}^{ee} w(e,ss)m(e,ss) & + & \sum_{e=1}^{ee} w(e,ss)em(e) & = & 0
\end{array} \tag{4.12}$$

As an example , the equations for a hypothetical system of $ee = 2$ and $ss = 3$ is given below:

$$\begin{vmatrix} \sum(1,s) & 0 & w(1,1) & w(1,2) & w(1,3) \\ 0 & \sum w(2,s) & w(2,1) & w(2,2) & w(2,3) \\ w(1,1) & w(2,1) & \sum w(e,1) & 0 & 0 \\ w(1,2) & w(2,2) & 0 & \sum w(e,2) & 0 \\ w(1,3) & w(2,3) & 0 & 0 & \sum w(e,3) \end{vmatrix} \begin{vmatrix} em(1) \\ em(2) \\ m_0(1) \\ m_0(2) \\ m_0(3) \end{vmatrix} = \begin{vmatrix} \sum w(1,s)m(1,s) \\ \sum w(2,s)m(2,s) \\ \sum w(e,1)m(e,1) \\ \sum w(e,2)m(e,2) \\ \sum w(e,3)m(e,3) \end{vmatrix} \tag{4.13}$$

After solving this system, the lightcurves are corrected via 4.9.

Weighting factor

At the original paper, *R.K Honeycutt* calculates the weighting factor as

$$w(e,s) = w_1(e)w_2(s)w_3(e,s)w_4(e,s) \tag{4.14}$$

where w_1, w_2 and w_3 are set to either 1 or 0 depending on the image and star. If a specific exposure is going to be excluded then $w_1 = 1$ and if a star is going to be excluded then $w_2 = 0$. If only the particular star s on image e is to be excluded then $w + 3 = 0$. Finally w_4 is calculated as $1/\sigma_1^2[m(e,s)$, where σ is the error as it was extracted by DAOPHOT. ThReR pipeline, though does not need to account for $w_1 - w_3$ because all the inappropriate exposures and stars have already been subtracted. Hence the overall weight of an exposure is simply $w(e,s) = w_4(e,s)$.

Implementation

Th.Re.T pipeline selects the comparison stars (template base from now on), based on the *sigma plot* (*see 4.5.1*) First the user is promed to select the desired number of comparison stars. Then the pipeline selects the ss brightest stars with the minimum deviation.

The set of normal equations are usually dependend because the instrumental magnitude scale has an arbitrary zero point. Hence only $ee + ss - 1$ equations are independent. To solve this problem we set $em(1) = 0$ and the equation to be minimized becomes

$$\beta = \sum_{e=2}^{ee} \sum_{s=1}^{ss} [m(e,s) - m_0(s) - em(e)]^2 w(e,s) + \sum_{s=1}^{ss} [m(e,s) - m_0(s)]^2 w(e,s) \tag{4.15}$$

The new system is then solved with the *LU* decomposition method (*see Numerical Recipes for Fortran sec 2.3*. The lightcurves are then corrected and the mean magnitude is calculated via

$$m_0(s) = \frac{\sum_{e=1}^{ee} [m(e,s) - em(e)]w(e,s)}{\sum_{e=1}^{ee} w(e,s)} \tag{4.16}$$

The error for each light-curve point is given by

$$\sigma^2[M(e,s)] = \sigma^2[m(e,s)] + \sigma^2_{mean}[em(e)] \tag{4.17}$$

with

$$\sigma^2_{mean}[em(e)] = \sigma^2[em(e)]/ss_{eff}(e) \tag{4.18}$$

Results of the ensemble photometry routine are given in section 4.7 and chapter 5.

Limitations and problems

As we discussed in section 4.5.2 trends may differ from star to star and hence the main assumption of ensemble photometry can be wrong. The problem becomes more intense when the field of view of the instrument covers a relatively wide area on of the celestial plane. As an attempt to solve this problem I tried to create subsets of stars that are dominated by similar trends and perform ensemble photometry at each group separately. For grouping I calculated the *correlation factor*

$$c_{xy} = \frac{n\sum x(e)y(e) - \sum x(e)y(e)}{\sqrt{n\sum x(e)^2 - (\sum x(e))^2}\sqrt{n\sum y(e)^2 - (\sum y(e)^2)}} \tag{4.19}$$

of all possible pairs $x(i), y(i)$ of lightcurves. Then, for each individual star x, the lightcurves y with $c_{xy} \geqslant 0.5$ are chosen as a template basis and the correction series $em(e)$ is evaluated for each one of these groups.

This method works well and solves most of the problems of ensemble photometry. However, an obvious limitation is the huge amount of computing time that is required. A more intelligent approach on this technique is introduced by *Dae-Won Kim et al.* (to be published), with the use of a hierarchical clustering algorithm to sort the data into groups.

4.5.5 The Trend Filtering Algorithm

A more successful noise reduction method has been developed by *Kovacs et al (2008)* [27] . The trend filtering algorithm is also based on the observation that in a large photometric database many stars will be affected by the same systematics. A representative sample of all possible systematics is then chosen as a template bases and a filter $\{F(i); i = 1, 2....N\}$ is constructed by the linear combination of all template lightcurves $\{X(j); j = 1, 2.....N\}$:

$$F(i) = \sum_{j=1}^{M} c_j X_j(i) \tag{4.20}$$

The coefficients $c(j)$ are calculated via the minimization of the parameter

$$D = \sum_{i=1}^{N} [Y(i) - A(i) - F(i)]^2 \tag{4.21}$$

where $\{Y(i); i = 1, 2....N\}$ is the target time series and $A(i)$ the current best estimate of the detrended light curve. In our application, since we have no a priori knowledge of the signal and

its period, $A(i)$ is assumed to be constant and equal to the mean value (namely equal to 0) of the target series. If the signal is known then $A(i)$ can be set properly and TFA can be used in an iterative way for signal reconstruction [27]. An estimate of the variance of noise of the filtered data can be obtained via

$$\sigma^2 = D/(N - M) \tag{4.22}$$

Implementation

For the implementation of the Trend Filtering Algorithm , I followed the steps that are described at the original paper fron *Kovacs et al (2008)*. Specifically, the main steps are the following:

1. *Select M template time series from the database, distributed uniformly in the field*. For the selection of the template time series, several techniques where tested. Specifically, the user of ThReT pipeline can choose as template stars:

   - The M stars with the lower σ value from the $sigma - plot$ diagram,
   - The M brightest stars,
   - The M stars with the lower pulsation factor,
   - The M brightest stars with the lower pulsation factor

   Each of these selection methods, was found to lead at similar results.

2. *Compute zero-average template time series*

3. *Compute normal matrix from the above template time series*:

$$g_{j,k} = \sum_{i=1}^{N} X_j(i) X_k(i) \tag{4.23}$$

4. *Compute the inverse of* $g_{j,k}$. The inverse matrix $G_{j,k}$ is computed by ThReT by the LU decomposition method (*see Numerical Recipes for Fortran and Appendix A*)

5. *Fore each light-curve, compute scalar products of the target and template time series*

$$h_j = \sum_{i=1}^{N} Y(i) X_j(i) \tag{4.24}$$

6. *Compute solution for* c_j

$$c_j = \sum_{k=1}^{M} G_{j,k} h_k \tag{4.25}$$

   and finally,

7. *Apply correction to the target time series*

$$Y_c(i) = Y(i) - \sum_{k=1}^{M} c_j X_j(i) \tag{4.26}$$

Advantages and disadvantages of TFA

TFA was proven to be robust and perform well on a wide range of data-sets (see section 4.7 and Chapter 5). However, is some cases, the linear combination *4.20* of trends might be similar to a real signal and hence the last may also be removed from the algorithm. The problem becomes more intense when signal and trends have similar periods and when the number of template stars is large, leading to an oversambled fitting. An additional limitation is that all template time series are left out and cannot be corrected. That means that the most "well measured" stars can not be further investigated.

4.5.6 The Sys-Rem algorithm

Sys-Rem algorithm was initially developed by *O. Tamuz et al* (2008) to correct for the color-dependent atmospheric extinction. However the generalization of this algorithm is able to remove linear systematic effects as well as several other unknown effects.

Color extinction

Given a set of M lightcurves, with each of them containing N number of exposures, the effective extinction coefficient c_i can be found via the minimization of the expression

$$S_i^2 = \sum_j \frac{(r_{ij} - c_i a_j)^2}{\sigma_{ij}^2} \tag{4.27}$$

where $\left\{a_j; j = 1, 2.....N\right\}$ is the air-mass at which the j^{th} exposure was taken and r_{ij} is the residual of each observation, namely the zero averaged instrumental stellar magnitude.
If the air-mass is known then

$$c_i = \frac{\sum_j \frac{r_{ij} aj}{\sigma_{ij}^2}}{\sum_j \frac{a_j^2}{\sigma_{ij}^2}} \tag{4.28}$$

Now, since atmospheric effects, such as clouds or low transparency atmospheric layers may also effect our observations, we can search for the optimum air-mass coefficient that describes our data. The "effective air-mass" can thus be found by the minimization of

$$S_j^2 = \sum_i \frac{(r_{ij} - c_i a_j)^2}{\sigma_{ij}^2} \tag{4.29}$$

with $\{c_i\}$, the previously calculated coefficients. The effective air-mass will now be

$$a_j^{(1)} = \frac{\sum_i \frac{r_{ij}ci}{\sigma_{ij}^2}}{\sum_j \frac{c_j^2}{\sigma_{ij}^2}} \tag{4.30}$$

This process can be continued iteratively until our data are fully corrected.

Generalization

The above method is actually a search for the best two sets of $\{c_i; i = 1, 2....M\}$ and $\left\{a_j; j = 1, 2....M\right\}$ that minimize the global expression

$$S^2 = \sum_{ij} \frac{(r_{ij} - c_i a_j)^2}{\sigma_{ij}^2} \tag{4.31}$$

Therefore the resulted a_j and c_i may not be the true air-mass extinction but represent well all linear systematic effects. If our data set includes a few different systematics, with different c_i and a_j then we can generalize the algorithm to also include these cases.

For doing that, we correct all residuals r_i after the first iteration by

$$^{(1)}r_{ij} -^{(1)} c_i^{(1)} a_j \tag{4.32}$$

and proceed as described, with the corrected $\{r_i\}$ sample

More specifically, we search for the best c_i, a_j that minimize

$$S^2 = \sum_{ij} \frac{(^{(1)}r_{ij} -^{(2)} c_i^{(2)} a_j)^2}{\sigma_{ij}^2} \tag{4.33}$$

and proceed repeatedly until no significant linear effects are found.

Results of the *Sys-Rem* algorithm are presented in Section 4.7

4.6 Searching for variable stars and Transit signals

ThReT pipeline uses one or more of the above mentioned algorithms in order to de-bias the raw lightcurves. After the successful removal of trends, the next and final step is to search for variability in the sample and especially for transit signals.

4.6.1 Searching for variable stars

In order to search for variable stars ThReT pipeline recalculates the $pulsationfactor$ of all de-biased lightcurves. Then the stars with the highest pulsation are visually investigated. This method obviously is not the optimum especially when it comes to low amplitude variability. However, since the prime aim of this work is to find extra-solar transits no other method was tried.

4.6.2 The search for transits-The BLS algorithm

Transit signals are difficult to be discovered by traditional algorithms due to the low amplitude variability and the relatively short transit duration in respect to the period. For example a finite sum of Fourier sinusoidal components (*Discrete Fourier Transformation - DFT* fails to model such a signal properly, due to the leakage of the signal power at many higher harmonics. A solution to this problem is the so-called Phase Dispersion Minimization (PDM). PDM folds the data and searches for the best period that yields the smoothest time-series (in terms of χ^2).

In order to search for transit shape signals in the lightcurves , ThReR pipeline makes use of a variation of PDM that is called *Box fitting least squares algorithm* (*Tamuz et al.* [28] 2002)

The algorithm searches for periodic signals that alter between two discrete levels, with much lest time being spend at the lower level. A FORTRAN routine for BLS is available from the Authors via the page *http://www.konkoly.hu/staff/kovacs.html*.

Implementation

Let us assume a strictly periodic signal with period P_0 that takes only two discrete values, H and L. The time spent in L (the low level) is qP_0 with q being a small number $(0.01-0.05)$. For each light-curve, BLS searches for the best box-shaped model by estimating five parameters: P_0, q, L, H and t_0 (the time of the mid-transit). If the time -series is zero-averaged then $H = -Lq/(1-q)$ and the number of free parameters is reduced to four.

The time series $\{x_i; i = 1, 2.....n\}$ is the addition of signal and Gaussian noise σ_i . The latter is presented as a weighting factor $w_i = \sigma_i^{-2}[\sum_{j=1}^{n} \sigma_j^{-2}]^{-1}]$. For any given trial period P, BLS folds the data and fits the step function $L([i_1, i_2; H[1, i_1), (i_2, n]$. The time spend at the lowest level is parameterized by $r = \sum_{i=i_1}^{i_2} w_i$. The fitting is performed via the minimization of the expression

$$D = \sum_{i=1}^{i_1-1} w_i(x_i - H)^2 + \sum_{i=i_2+1}^{n} w_i(x_i - H)^2 + \sum_{i=i_1}^{i_2} w_i(x_i - L)^2 \tag{4.34}$$

The minimization of D yields the weighted arithmetic averages over the respective regimes

$$L = \frac{s}{r}, H = -\frac{s}{1-r} \tag{4.35}$$

where

$$s = \sum_{i=i_1}^{i_2} w_i x_i \tag{4.36}$$

The average squared deviation of the fit is then given by

$$D = \sum_{i=1}^{n} w_i x_i^2 - \frac{s^2}{r(1-r)} \tag{4.37}$$

After the last expression is evaluated, BLS repeats the computation with other $(i_1, i_2$ and finally finds the absolute minimum of D for any given period. Note that the first term of Eq.

4.37 does not depend on the trial period and therefor the second term can be used intepentantly, to characterize the quality of the fitting. The Box-Fitting Least Squares frequency spectrum is therefore defined at any given period as

$$SR = MAX\left\{\left[\frac{s^2(i_1, i_2)}{r(i_1, i_2)[1 - r(i_1, i_2)]}\right]^{1/2}\right\} \tag{4.38}$$

To characterize the signal detection efficiency BLS uses the expression

$$SDE = \frac{SR_{peak} - \langle SR\rangle}{sd(SR)} \tag{4.39}$$

where SR_{peak} is the SR at the highest peak, $\langle SR\rangle$ is the arithmetic average and $sd(SR)$ is the standard deviation of SR over the frequency band tested.

4.6.3 A fast detection algorithm

BLS is a fast and efficient method to search for planetary transits in a large database of lightcurves. However, it requires more than 2 observed transits to work and hence, it is not convinient for single-night observations.

Due to the fact, that our survey did not yet run for a considerable amount of time (see Chapter 5), we wanted to create an algorithm to extract these events. For that purpose, the Fast Detection Algorithm (FDA) was created.

FDA is very simple practical, yet roburst method, to extract transit-like futures from a large database of lightcurves . In is based on the empirical observation that the *depth of a transit expected to be detected by our instrument ranges from* $dmag = 0.007$ *to* $dmag = 0.03$. Taking advance of this observation, FDA works as follows:

1. It first applies a median filter to each light-curve $\{Y(i); i = 1, 2...n\}$, by averaging every m points, where $3 \leq m \leq 10$, regarding to the amount of points collected at a single-night observation.

2. Then the algorithm calculates the value $g(j) = Y(j) - Y(i)$ for each point of the binned light-curve

3. If $g(j)$ and $g(j + 1)$ are between the range $0.007 \leq dmag \leq 0.03$ then an *indicator index* k is increased by 1. If the next point does not lye in that range then k is reset to 0.

4. Finally the algorithm is reapplied for negative values of *dmag* to include the case of a transit at the beginning of the observation run.

5. If k exceeds a predefined value (Which is set in regards to the exposure time, to correspond to a transit length from $30mins$ to $4hours$) then the detection is considered successful.

This simple method of course has many false detections but *it is sure that it will exclude all lightcurves that do not have a transit feature*. The remaining light-curve can then be examined visually in order to see if the star is a true transit candidate or not.

4.7 Tests of ThReT pipeline

Here we present the test of the performance of ThReT pipeline at two different types of data. The first data-set is one night of photometric data of the globular cluster $M92$ obtained with the Hamburg 1.2 m RC telescope wile the second data-set is a set of photometric images of the open cluster NGC129, obtained with the low-cost CCD camera SBIG ST-7 and a fast Nikor $f/2.8$ photographic tele-lens. The data from M92 where de-trended with all three methods while NGC129 was reduced only with the TFA algorithm.

4.7.1 NGC129 open cluster photometry

This set of data was acquired during a night with pour atmospheric conditions and foul moon. The data consist of 195 images obtained at $10/11/2006$ from the Holomon Astronomical Station. The exposure time was $60sec$ and no filter was used. The images where bias subtracted, dark subtracted and flat fielded with ThReT pipeline. No correction for linearity was applied due to the lack of information about the linearity factor.

The field of NGC 129 is dense and many stars were merged due to the small angular resolution of the instrument ($\backsim 10arcsec/pixel$) (Figure 4.3). The pipeline detected 2565 stars at the reference image. The number of lightcurves with measurements within 60% of the images or more was 2129.

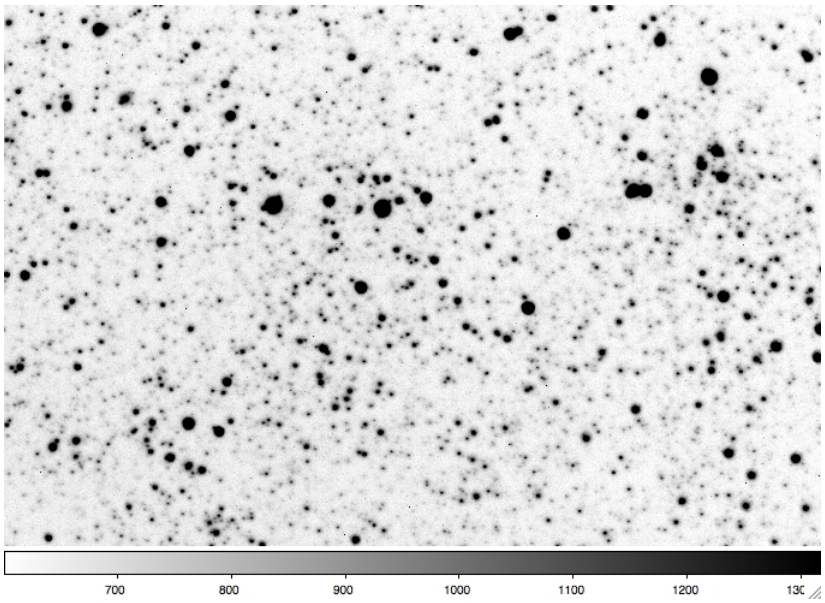

Figure 4.3: The NGC129 Open cluster at Cassiopeia as captured with ST-7 camera

After the preliminary reduction, PSF photometry was performed in order to extract the instrumental magnitudes. The PSFs were integrated with a *Lorenz* function within a radius of 6 pixels.

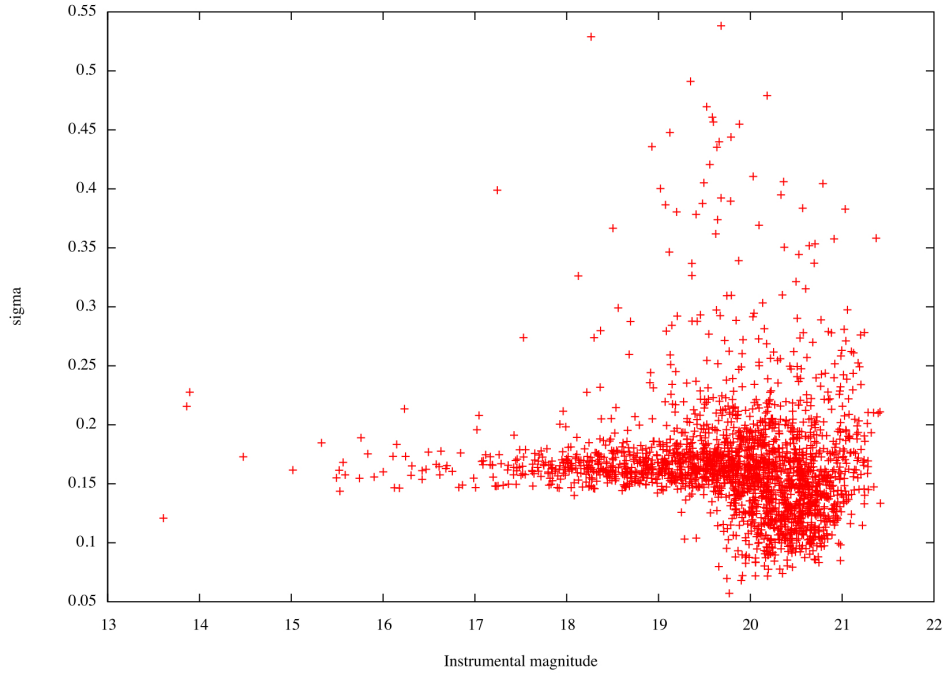

Figure 4.4: Raw Sigma vs Mag Plot of NGC129 data. We see that the plot does not follow the trace of poison statistics. That is the result of cloud-passing during the night. The bright stars at the frame dimmed, leading to higher σ values, while the faint stars disappeared leading to no points at all.

The σ plot of the instrumental magnitudes extracted, can be seen at Figure 4.4. The strange shape of the raw *sigma* plot means that the lightcurves are dominated by non-linear trends. Indeed, after a visual examination of the images, it appears that a thin cloud passed in front of the field, causing the signal of all stars to drop. The brighter stars where not completely diminished ,but their signal dropped, causing a deep at their lightcurves. On the other hand the dim stars, where disappeared and no point at their lightcurves appear during the cloud pass. The deep at the bright stars' lightcurves caused a rise at the value of standard deviation, while the standard deviation of dim stars was not affected. That explains the shape of the *sigma* plot.

The lightcurves where, first corrected with ensemble photometry, with the use of the 20 brighter stars as a template base. The statistical values where then recalculated for the ensembled curves. The "poisson shape" of the sigma-plot was reestablished, meaning that the cloud trend was successfully corrected(Figure 4.5).

The ensembled lightcurves where then corrected with the TFA algorithm, with the use of 80 template stars. The results of TFA are presented below. Figure 4.6 shows the *sigma* plot of the corrected lightcurves, plotted together with the ensembled data for comparison. Figure 4.7 is the $\sigma(TFA) - \sigma_{BeforeTFA} vs \sigma_{BeforeTFA}$ plot and Figure 4.8 is a light-curve before and after the TFA correction. As it can be seen, the trend filtering algorithm removed the trends

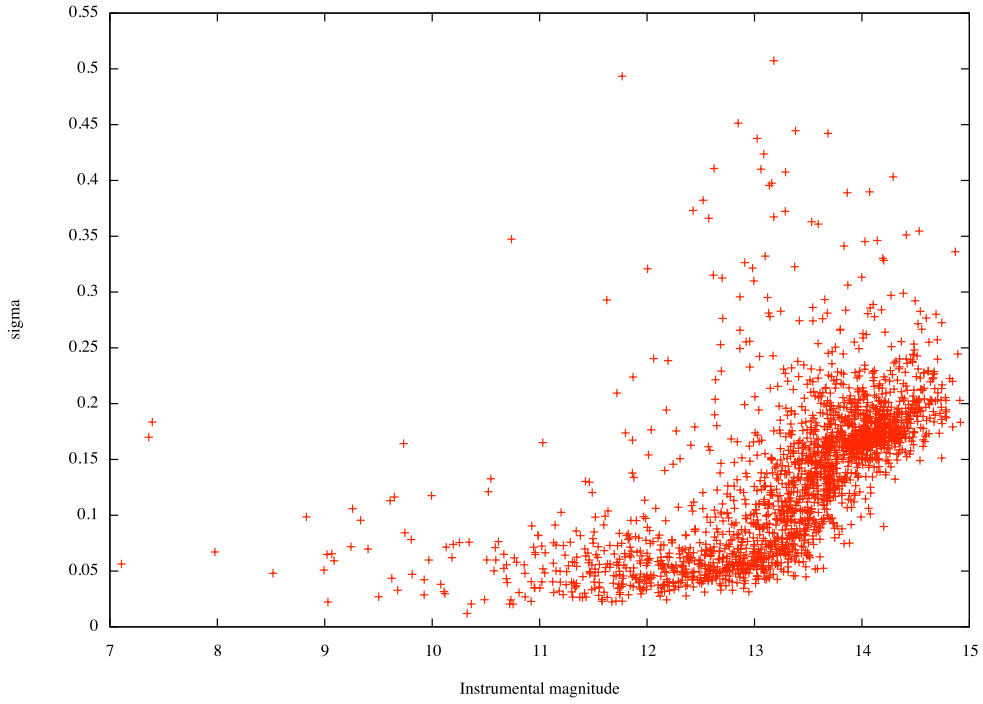

Figure 4.5: Sigma vs Mag plot of ensembled lightcurves. The Poisson trend is reestablished.

efficiently. Although, even after the systematic errors removal, the lightcurves do not have the required accuracy for transit hunting[4]

After the reduction, the pulsation index was used for variable star hunting. Figure 4.9 shows the pulsation factor plotted against the magnitude. The value $p = 1.5$ was chosen as a variability limit, and all stars with pulsation factor larger than this value were visually examined. A promising variable star candidate was found (Figure 4.10) at the coordinates $x = 356$,$y = 440$ of the reference image.

The variability seems to be instrict since all neighbor stars are stable and area of the CCD chip were the candidate was found does not have bad or hot pixels. Although, more data are required in order to further investigate the candidate and determine it's type, if the star is a true variable star.

4.7.2 M92 globular cluster photometry

This field was observed with the Hamburg 1.2m telescope, in order to test the pipeline for forthcoming observations of the Palomar-10 globular cluster with the Aristarchos 2.3 m telescope (NOA, Greece). The observations where made by *Dimitris Mislis* (Hamburg Observatory PhD student). An image of M92 as captured by the telescope is depicted below.

The data where reduced with the standard procedure described above. As it is seen in

[4]We take as a threshold for transit hunting the empirical value of $\sigma \leq 0.01$

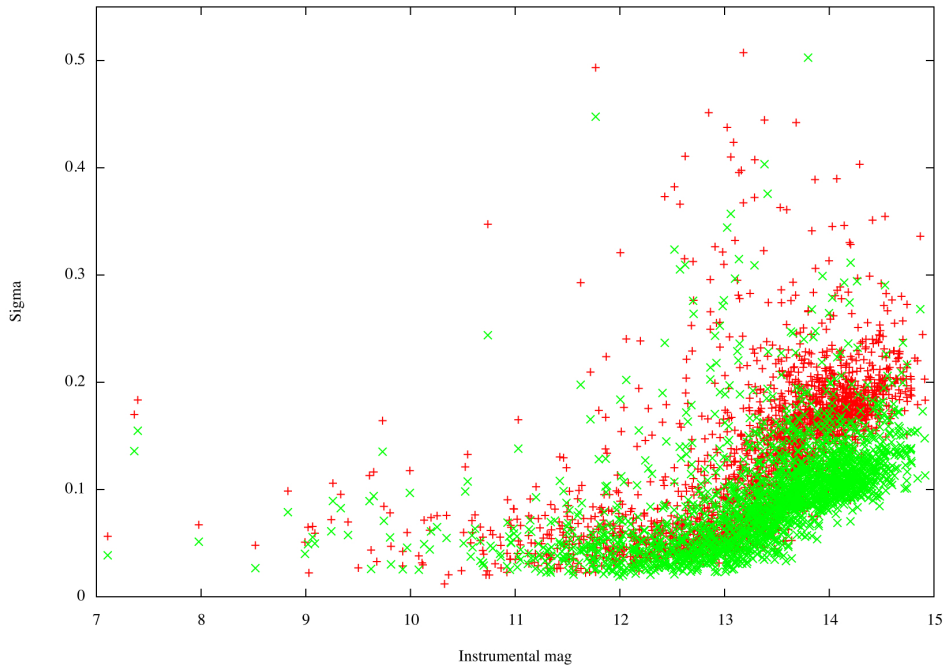

Figure 4.6: Sigma vs Mag plot, after the application of TFA

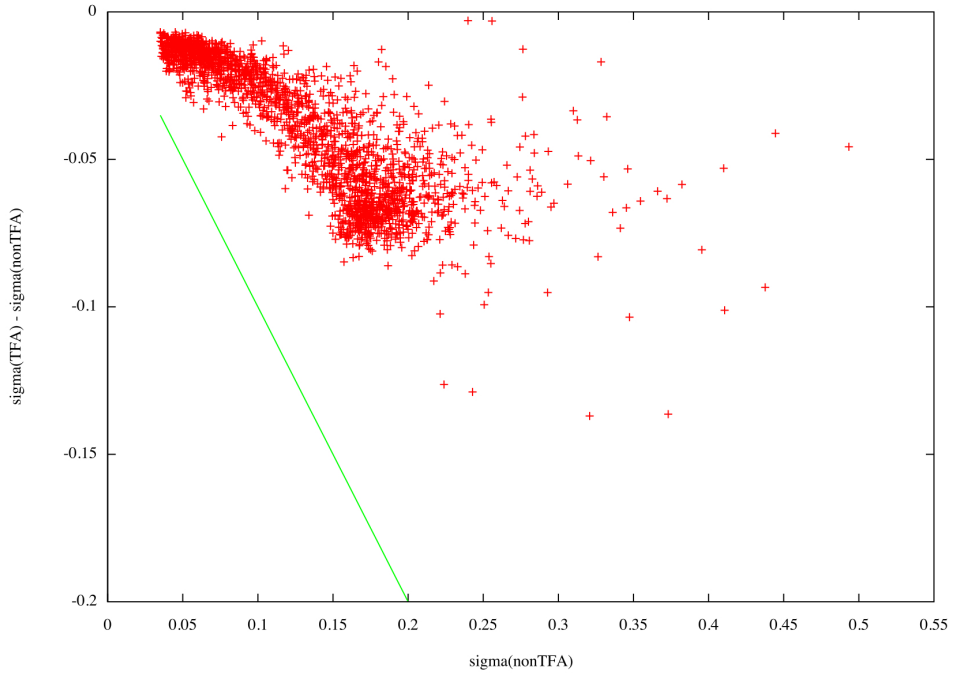

Figure 4.7: $d\sigma$ vs σ_{raw} plot

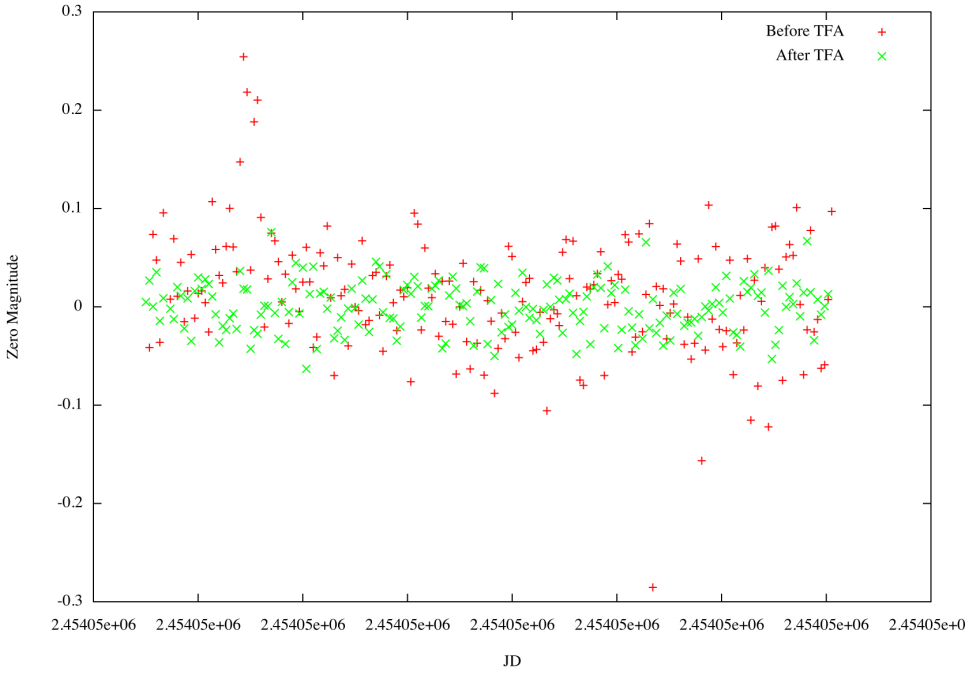

Figure 4.8: A light-curve before and after reduction with TFA

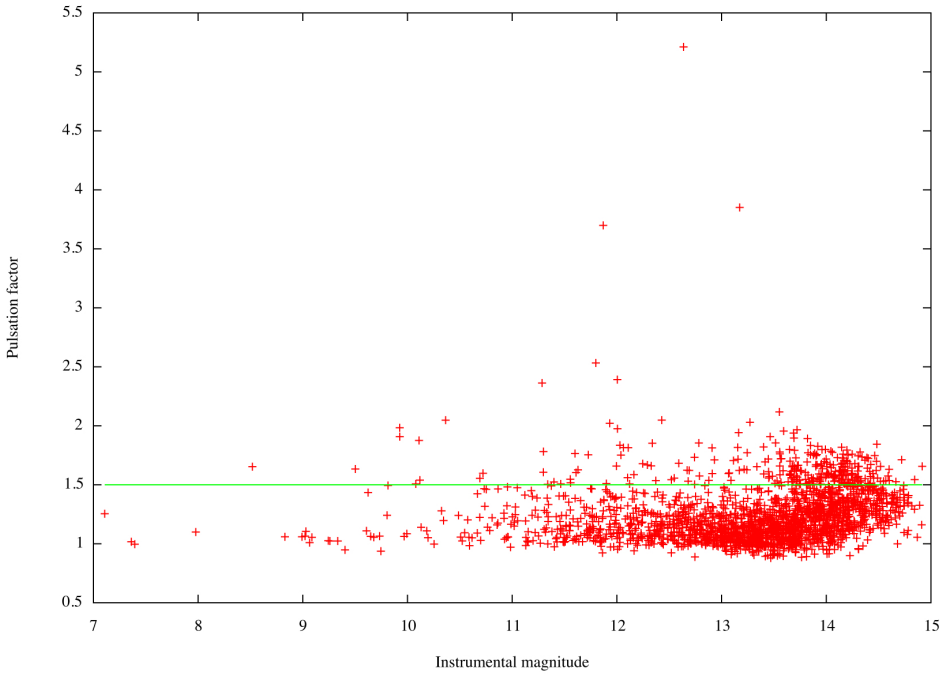

Figure 4.9: Pulsation factor for NGC129. The line $p = 1.5$ was chosen as a variability limit

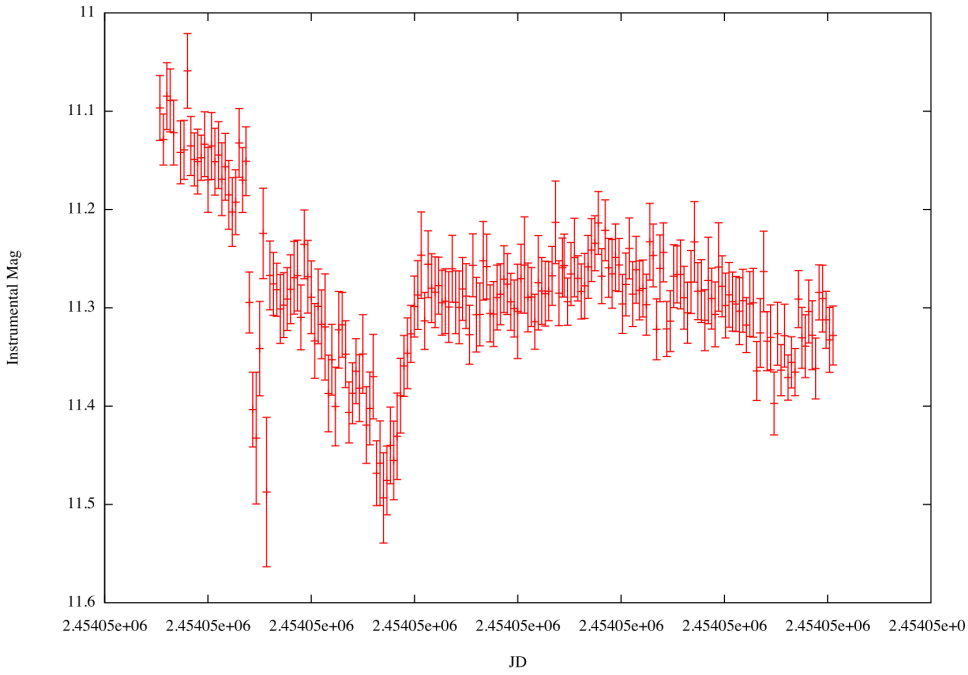

Figure 4.10: Variable star candidate at NGC129

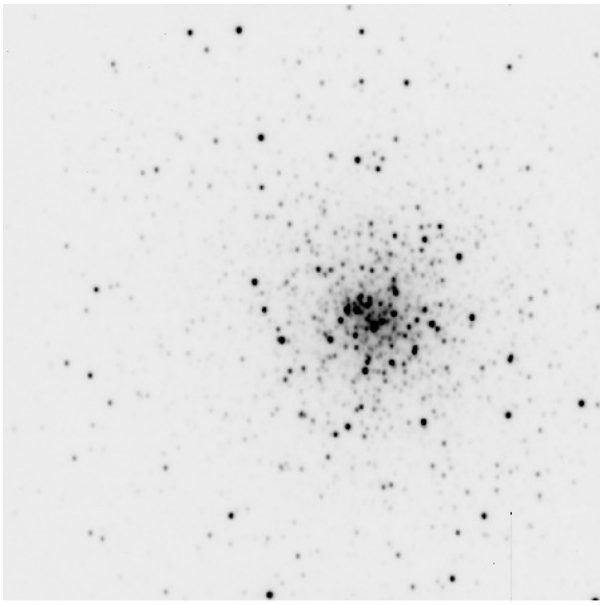

Figure 4.11: M92 Globular cluster as captured by the Hamburg 1.2m telescope

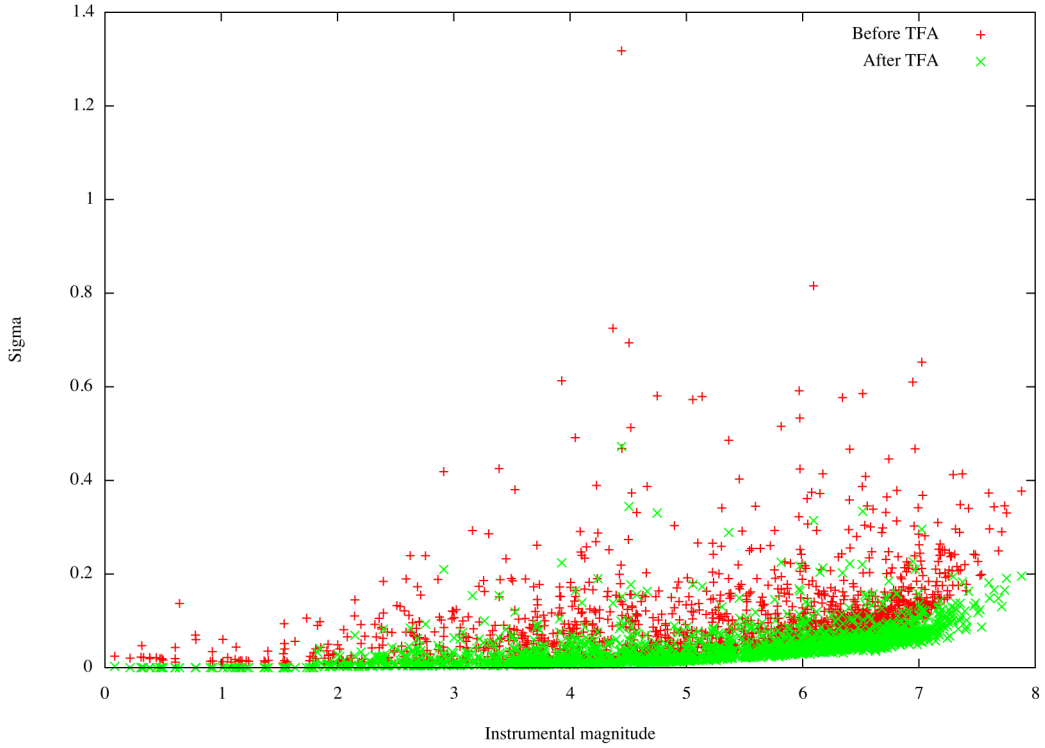

Figure 4.12: Sigma plot, before and after the TFA application

Figure 4.11 the density of the stars in the field is huge and therefore PSF photometry must be applied in order to extract the instrumental magnitudes.
After the derivation of the light-curves the data where de-trend with ensemble photometry, TFA and Sys-Rem algorithm. I used 20 comparison stars for the ensemble photometry and 70 templates for the Trend filtering algorithm. The results are depicted at Figures 4.12 – 4.16.

Variable star candidates

With the use of the *pulsation factor*, we found 5 new *intraday* variable star candidates. The line $p = 2$ was chosen as a threshold for variability. All of the suspected variables are presented below.

We didn't observe a full-phase for none of these variables and thus, their types where not identified. In our future plans is to re-observe the cluster in order to indicate the type of variability for these candidates and detect more longer-period variables. The lightcurves were also searched for transit events with the FDA, without any results.

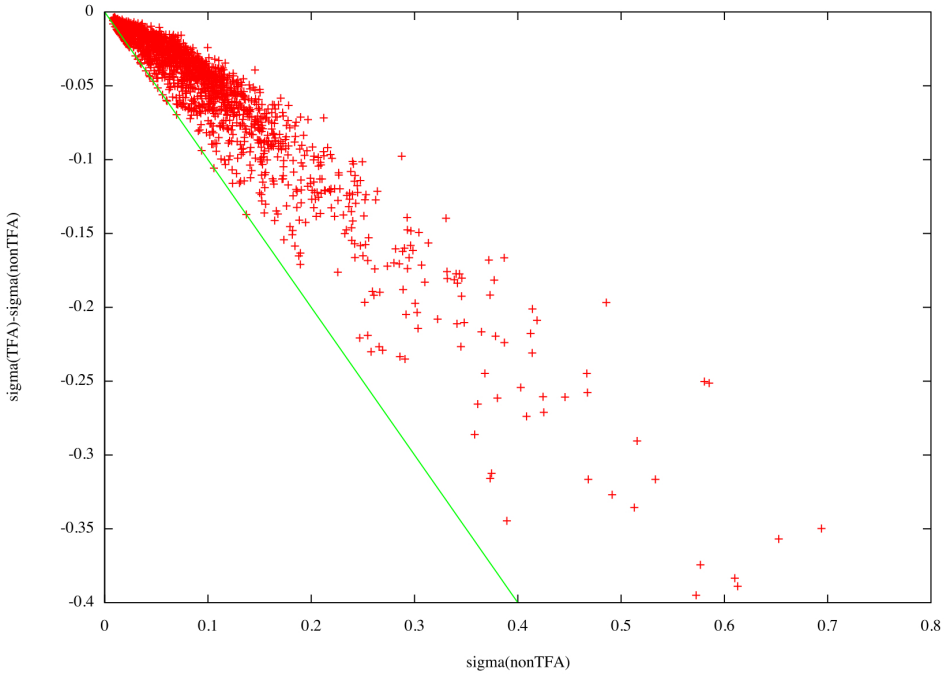

Figure 4.13: $d\sigma$ vs σ_{raw} plot

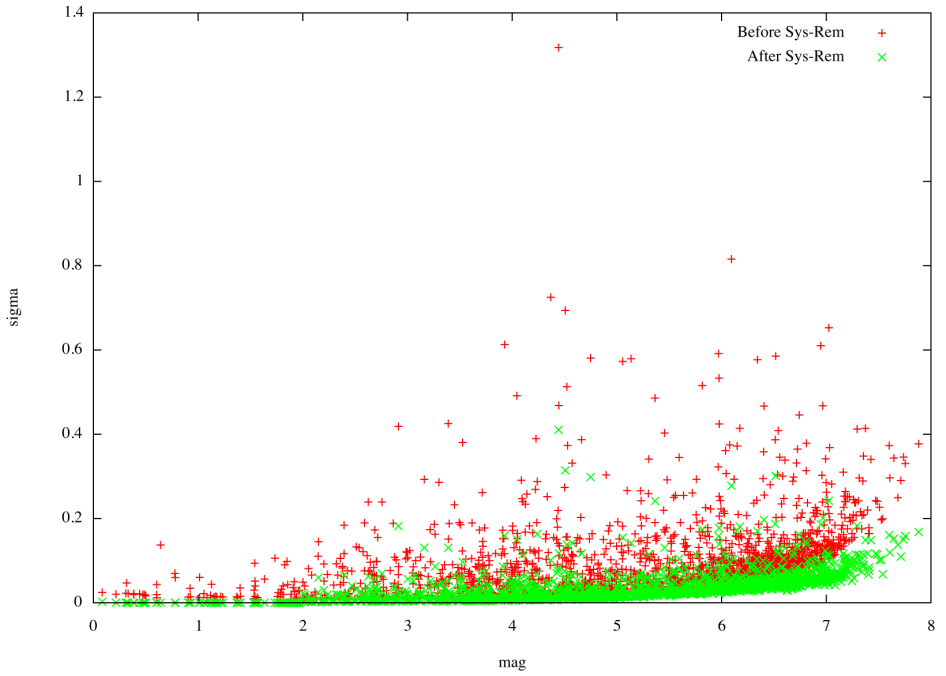

Figure 4.14: An example of sigma vs magnitude plot

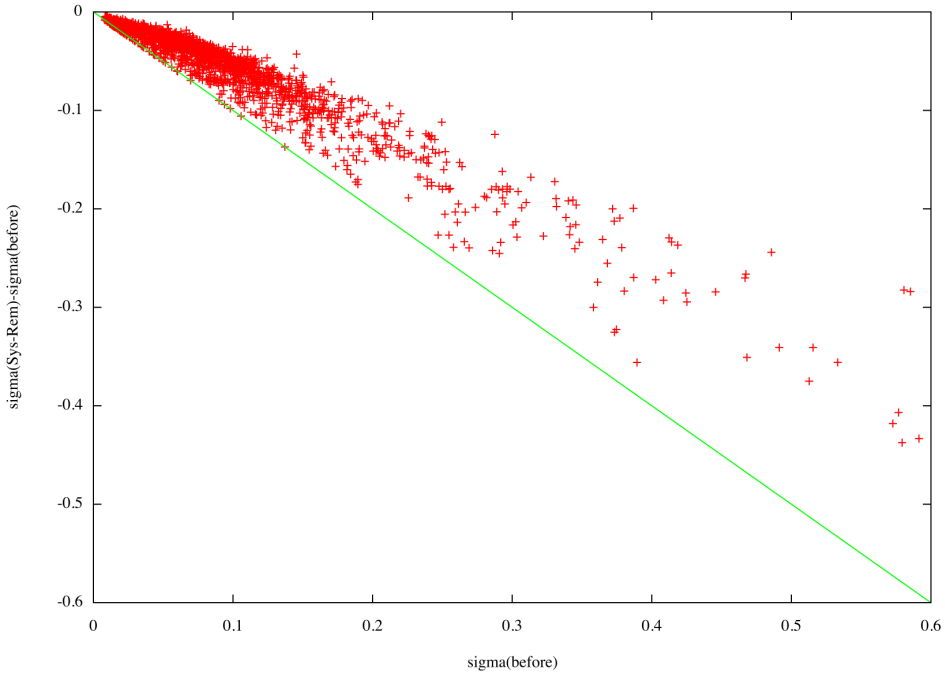

Figure 4.15: $d\sigma$ vs σ_{raw} plot

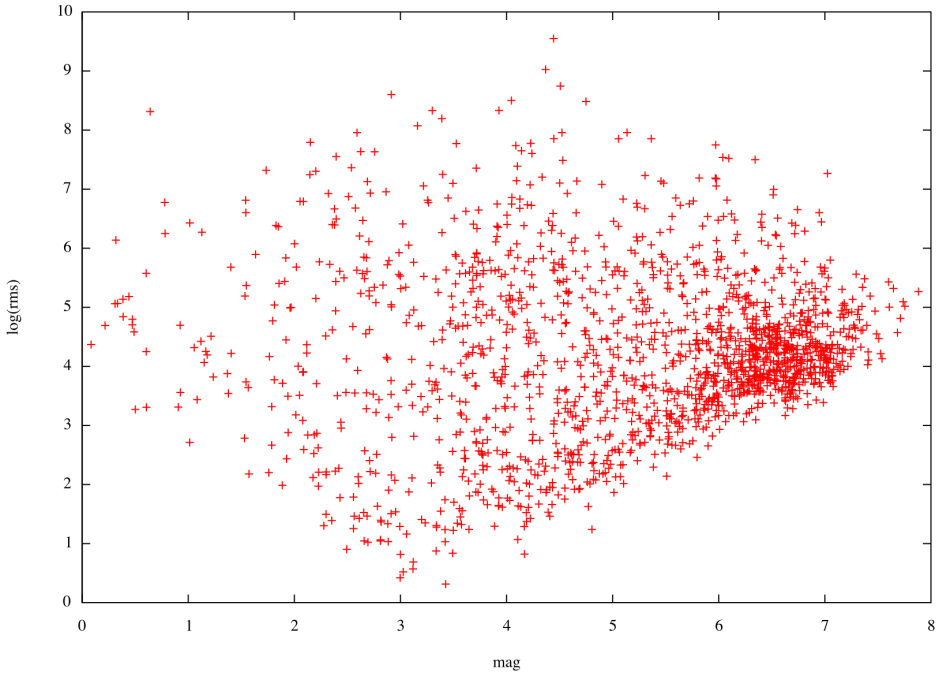

Figure 4.16: log(rms) plot of the M92 data. It turns out that DAOPHOT, may had overestimated the errors and thus the plot is not coherent.

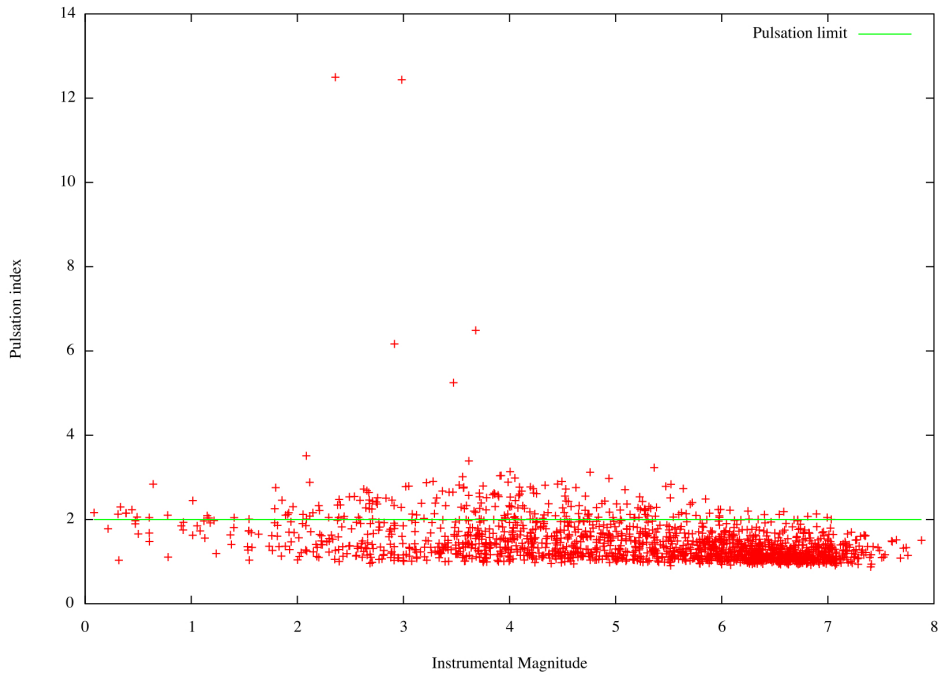

Figure 4.17: Pulsation index for M92.

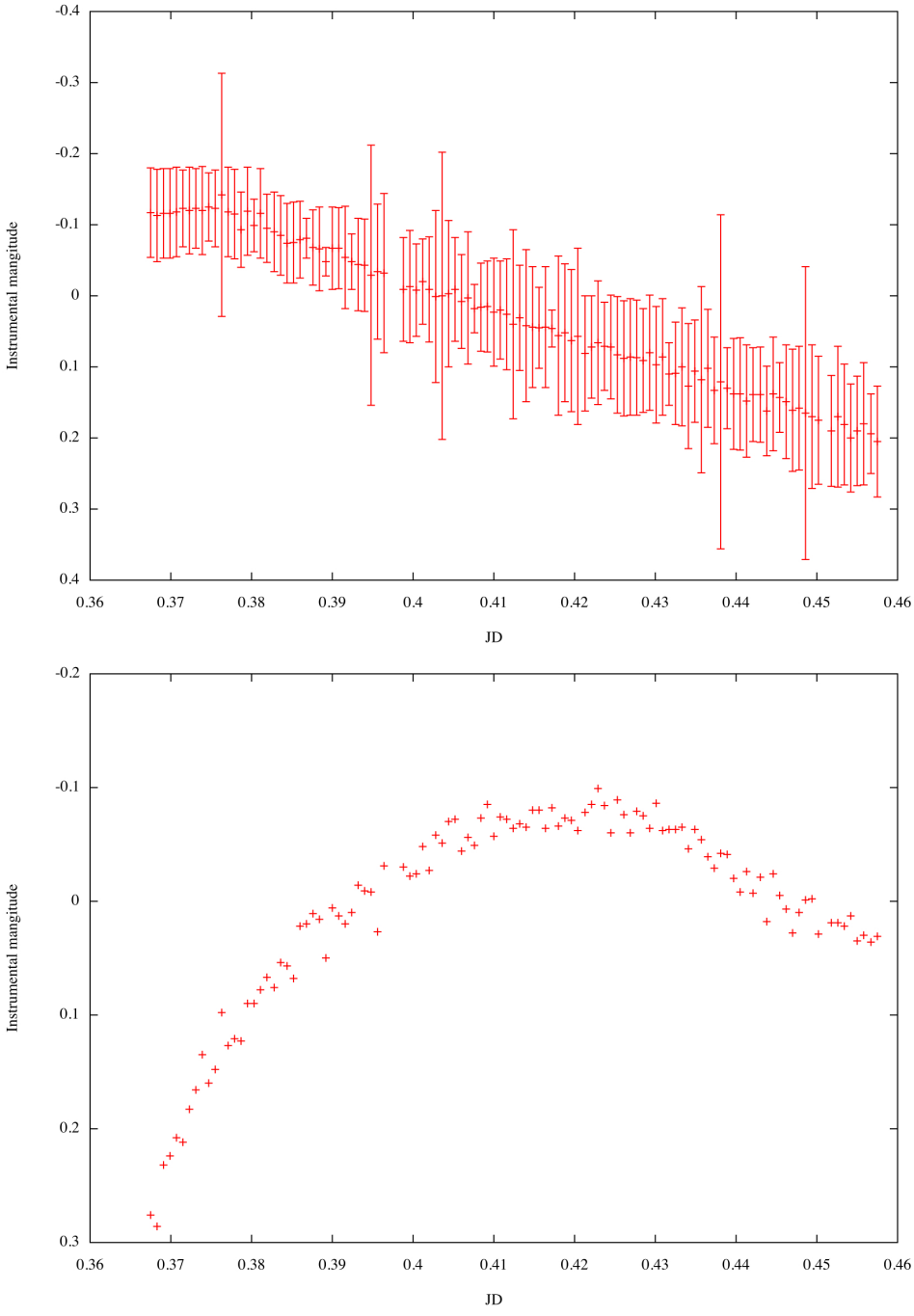

Figure 4.18: Variable star candidates in the field of M92

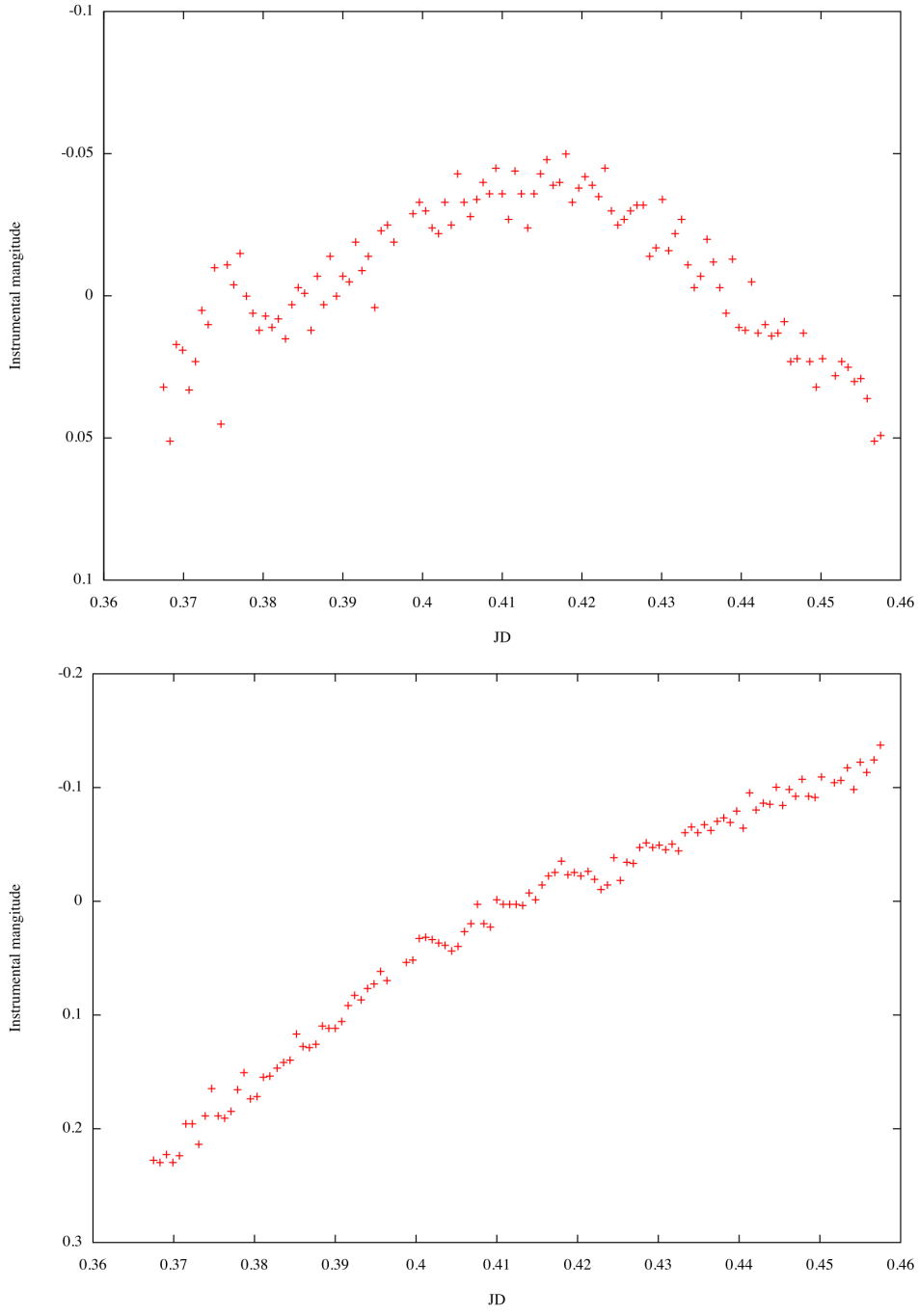

Figure 4.19: Variable star candidates in the field of M92

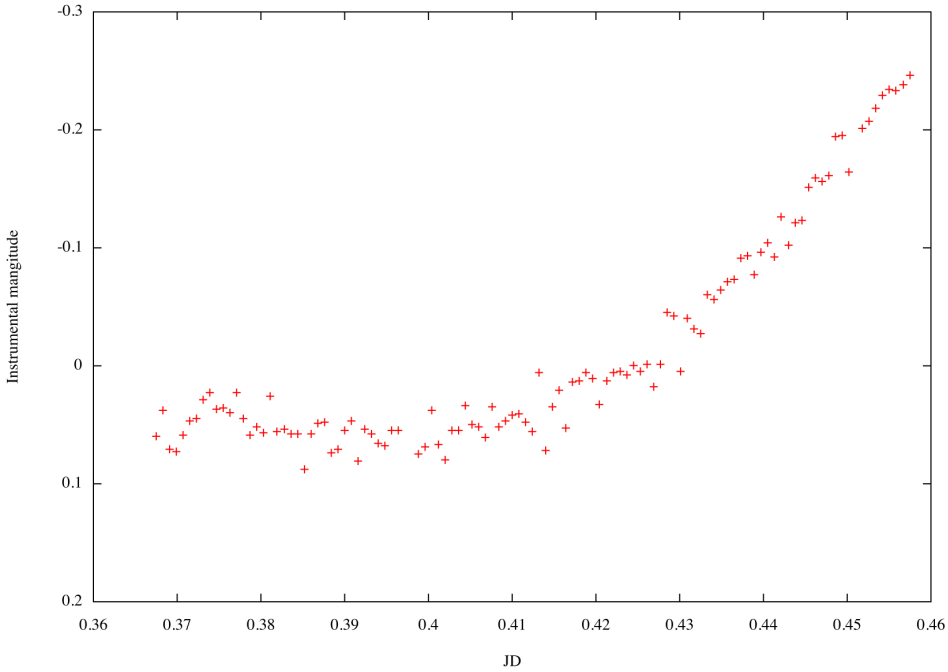

Figure 4.20: Variable star candidate in the field of M92

Lost in the milky way, smile at the empty sky and wait for the moment, when a million chances may all collide.

The Lightning Seeds, "The Life of Reilly"

5 Observations and Results

5.1 Introduction

In this final Chapter of this thesis, I describe the observations and the results from the analysis of the first run of the Holomon Transit Survey. The observations were initiated at July 20 2008 at Mt. Holomon and lasted 7 days. Some more days before July 20^th where also spend in order to fix problems of the Telescope and the guiding system. The field observed was the first field of Table 3.7 ($\alpha = 22^h53^m00''$, $\delta = +44^o30^m00''$) with an overall probability for a transit detection of $p = 71.99\%$. From now on we will refer to that field as *ExoField-1*. The main purpose of these observations where to test our Instrumentation and the ThReT pipeline in real-life conditions and not to make potential discoveries of Variable stars and transiting planets. Unfortunately, after the initiation of the observations, the weather was not on our side and we only managed to get 1 night of worth data. We plan to re-observe ExoField-1 in the future for a more considerable amount of time. In the rest of the Chapter an overview of the observations is given together with the results from the analysis of the 24/07/2008 night's data.

5.2 Overview of the observations of 24/07/2008

We observed *ExoField-1*, with the Celestron C11 Telescope and the Finger-lakes PL6303E camera at the night of 24 07 2008. During the night the weather was clear and stable. The moon was at the last quarter. We used an exposure time of 120 sec and a time interval of 20 sec between the exposures. No filter was used for the observations. The Point Spread Function (PSF) of the stars had a mean FWHM of $\backsim$ 3 pixels. We obtained a total of 113 exposures, 40 bias frames 1 dark frame of 15 min exposure time and 3 flat frames. The mean value of the last was about $\backsim$ 45000 ADU. The observations were once stopped, during the night due to a cloud pass. This time was used to obtain some of the calibration frames.

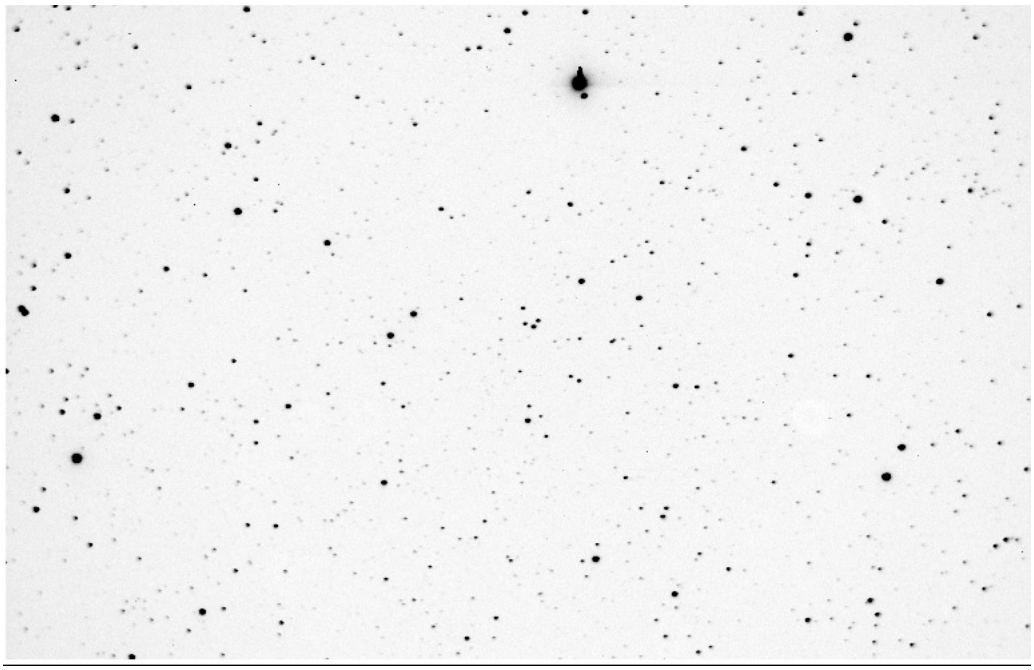

Figure 5.1: The first test field as captured by our instrument.

5.3 Analysis

The images where bias and dark subtracted and corrected for non-linear response and pixel-to-pixel variations (flat fielding) with the ThReT pipeline. Because the mean angular separation of the PSFs in the image is relatively large, we used conventional aperture photometry. The results do not significantly depend on the radii of the apertures. The lightcurves presented here were all extracted with an inner aperture of 5.5 pixels, and inner and outer sky radius of 9 and 14 pixels respectively. A threshold of 5σ was used as an object detection limit.

With these parameters , DAOPHOT detected 1354 objects at the reference frame, of which 1154 existed in more than 60% of the exposures.

The lightcurves were then post-reduced with ensemble photometry using 20 comparison stars and the *Sys-Rem* algorithm using 4 iterations. The results of the noise removal procedure are depicted below. Figure 5.1 shows a light-curve at each point of iteration of the Sys-Rem algorithm. Figure 5.2 is the $d\sigma$ vs σ_{before} plot and Figure 5.3 is the *Sigma-plot* of the final lightcurves. 754 of the stars have been measured with precision better than 1% as it can be seen in Figure 5.4.

5.4 Results

After the noise reduction. the lightcurves were searched for variables with the implementation of the *Pulsation factor*. The variability threshold was arbitrary chosen to be $p = 1.3$. FDA

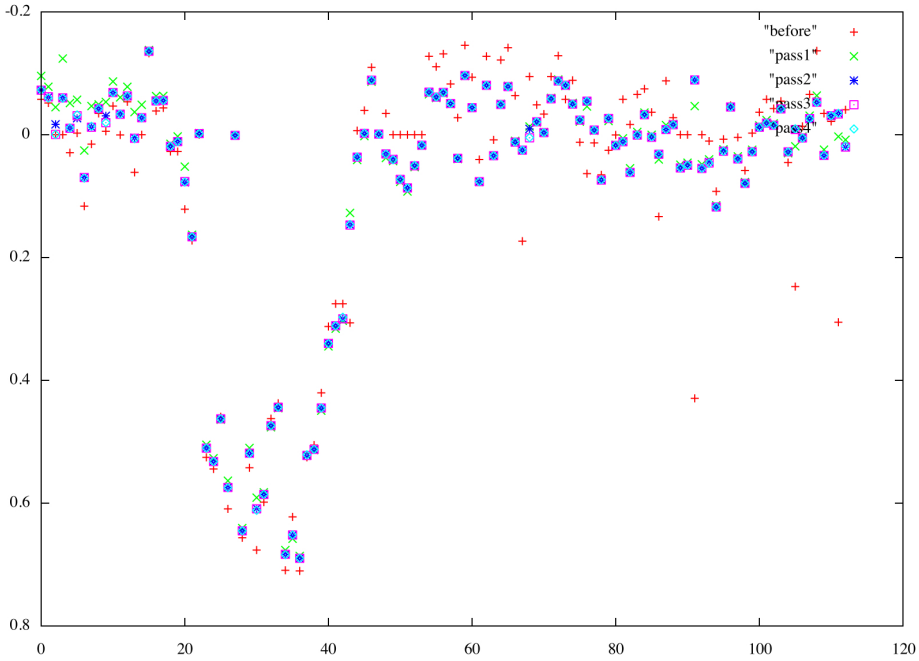

Figure 5.2: Example of a light-curve at each loop of the iteration of the Sys-Rem algorithm

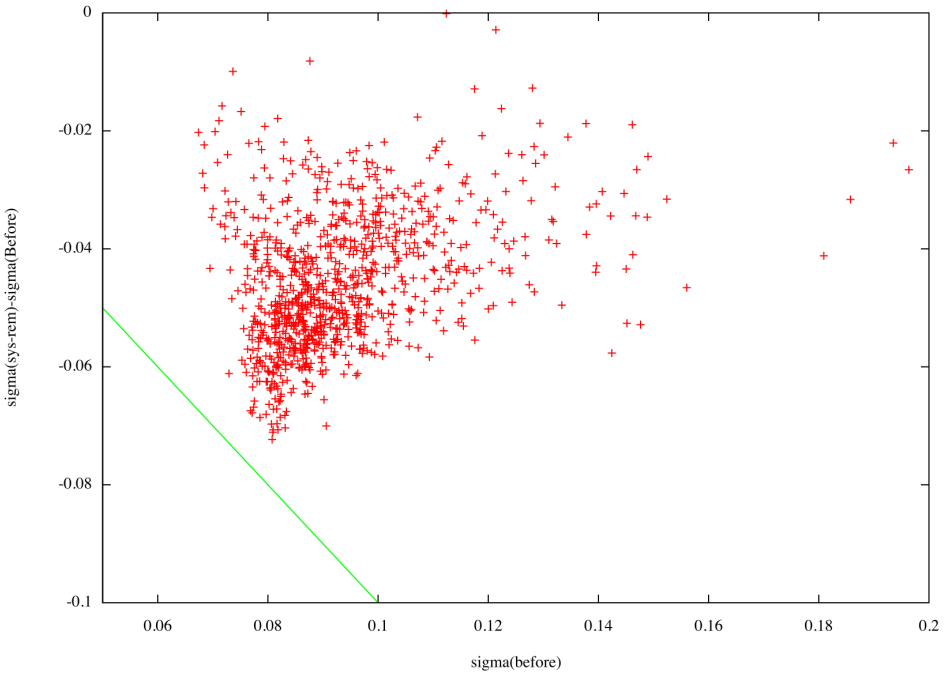

Figure 5.3: .
$d\sigma$ vs σ_{before} of the ExoField-1 data

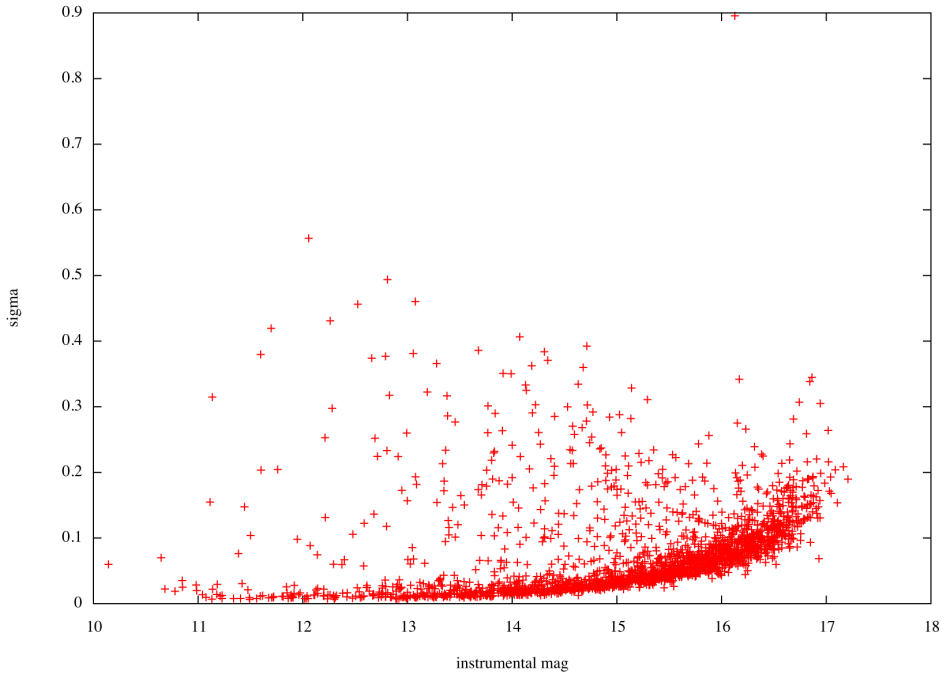

Figure 5.4: Sigma plot of the final lightcurves

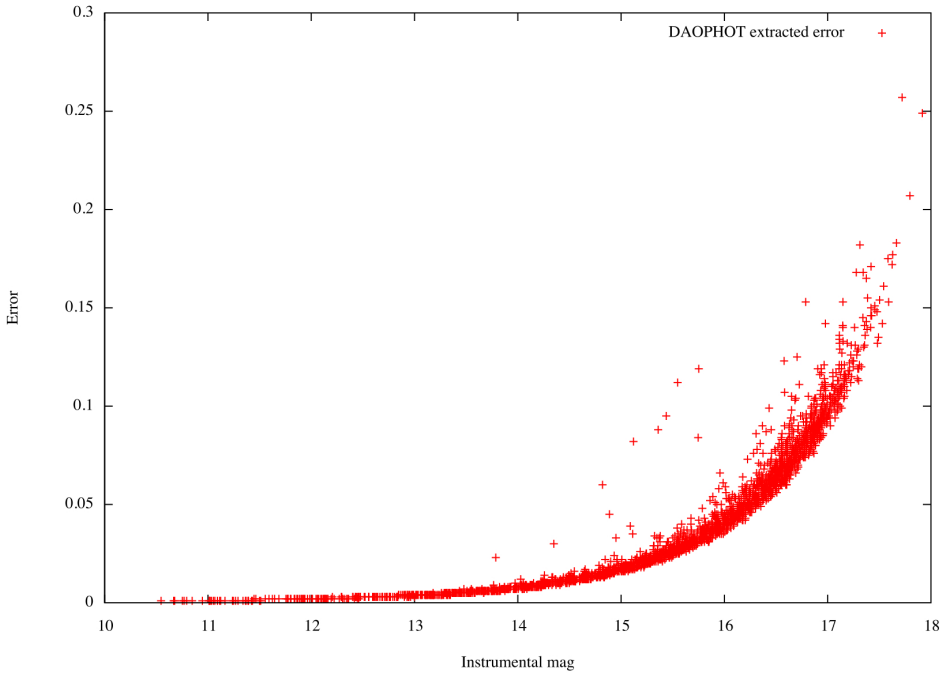

Figure 5.5: Daophot extracted error, for the reference frame of 24/07/2008. As it can be seen the precision achieved allows for transit hunting at a large fraction of the stars.

ID	α	δ	R_{mag}	B_{mag}
USNO-A2.0 1275-17970800	22^h 52^m 58”	$+44^o$ 56^m 14.32”	14.6	16.3
USNO-A2.0 1350-17532765	22^h 54^m 36”	$+45^o$ 13^m 40.03”	13.9	15.0
USNO-A2.0 1275-17982205	22^h 54^m 36”	$+45^o$ 13^m 40.03”	11.9	13.1
USNO-A2.0 1275-17972957	22^h 53^m 06”	$+44^o$ 58^m 48.91”	14.69	13.4

Table 5.1: The variable candidates of ExoField-1

was also applied in search of transit events. Astrometry was also performed at the reference image, with the starlink GAIA facility, in order to find the coordinates and the characteristics of the detected variables, if any. The results are presented bellow

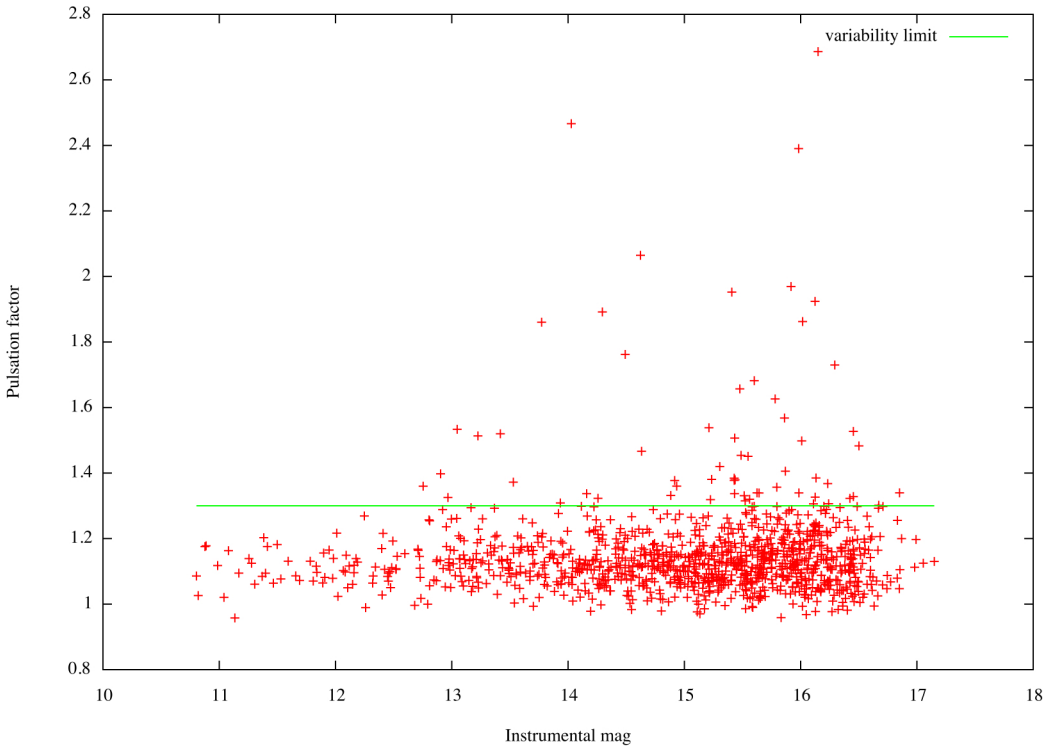

Figure 5.6: Pulsation factor for ExoField-1. The line $p = 1.3$ was chosen as a variability threshold.

5.4.1 Variable star candidates

We found several variable light curves within our data-base. Four of them are promising variable star candidates. The lightcurves of these stars are presented below (Table 5.1 and Figures 5.7-5.10)

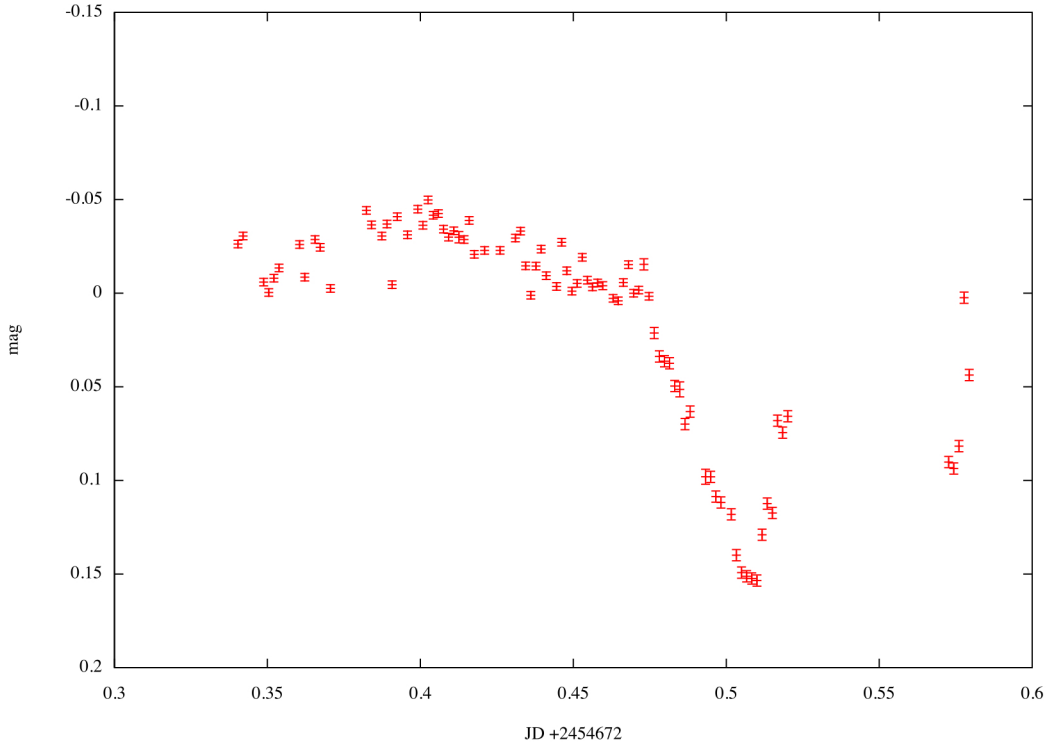

Figure 5.7: Light curve of USNO-A2.0 1275-17972957

None of these stars is listed as a variable in the SIMBAD database. We did not observe a full phase at any light curve and thus we are unable to determine the type of the variability.

5.4.2 Transit Candidates

We found one promising transit candidate with the use of FDA. The characteristics of the star are listed in table 5.2 and the time-series is depicted at Figure 5.11

ID	α	δ	R_{mag}	B_{mag}	Spectral Type
USNO-A2.0 1275-17995584	$22^h\ 54^m\ 40$”	$+44^o\ 52^m\ 29.00$”	13.1	14.2	K0

Table 5.2: Characteristics of the ”Host” star

Unfortunately, we did not observe a full transit, and thus we can not model the event properly. Moreover the variability is not yet confirmed to be instrict. The ”host” star is of unknown classification and hence we do not know if it belongs to the Main-Sequence or not. However, the star is of spectral type K0 and hence if it is actually a main-sequence star it is very possible to be a true extrasolar planet host star. We shall monitor the star in the future and if a transit is re-observed then we will perform follow-up spectroscopic and photometric observations with a bigger instrument in order to examine the nature of the variability.

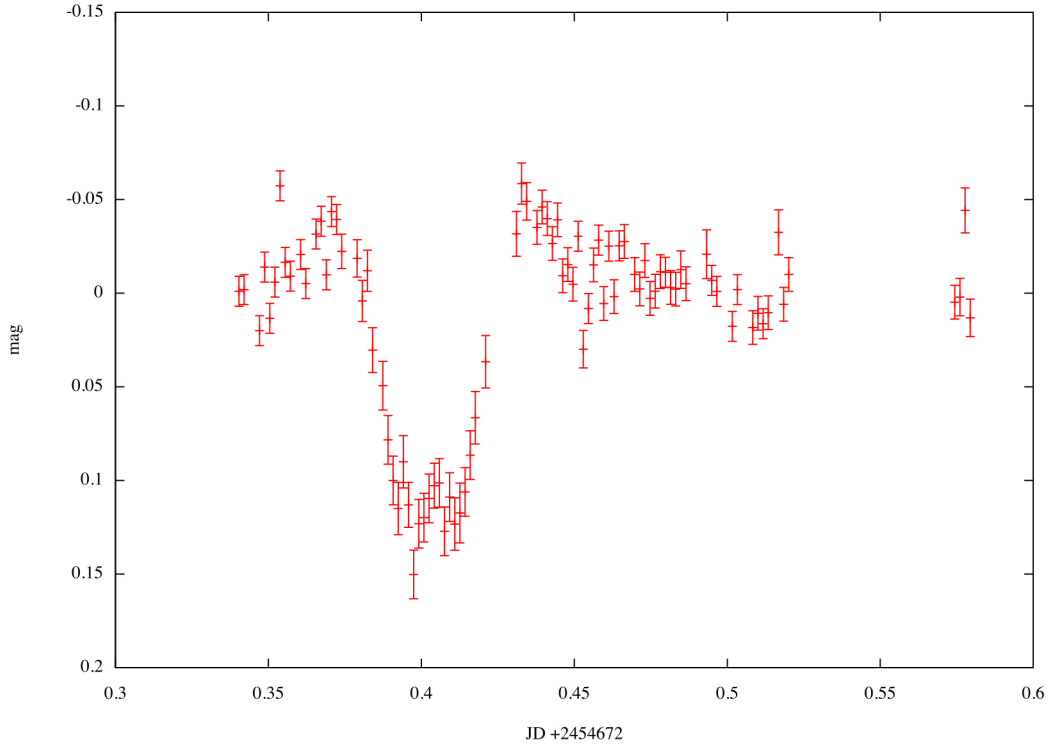

Figure 5.8: Light curve of USNO-A2.0 1275-17970800

Modeling of the transit

As it was mentioned above, we did not observe neither a full phase or a full transit and thus there is little information that we can derive from the light-curve. Although some information can be extracted under the hypothesis that the star belongs to the main sequence.

The star is listed as a $K0$ star and thus, its effective temperature is $T_{eff} = 5000K$ and its approximate radius is $R_* = 0.85R_\odot$ (*http://www.enchantedlearning.com*). With the use of the simple geometrical model:

$$M(t) = b + a\sin(kt - t_0) \tag{5.1}$$

we found that the drop of the brightness is approximately $\Delta F = 0.012936$ and the duration of the transit is $d = 1.29h$. The values of the fitted parameters are listed in Table 5.3 and the model together with the light curve is depicted at Figure 5.12

	b	a	k	t_0
value	1.00058	0.01293	29.1259	13.3414
error (%)	0.108	7.2	7.6	6.89

Table 5.3: Values of the fitted parameters

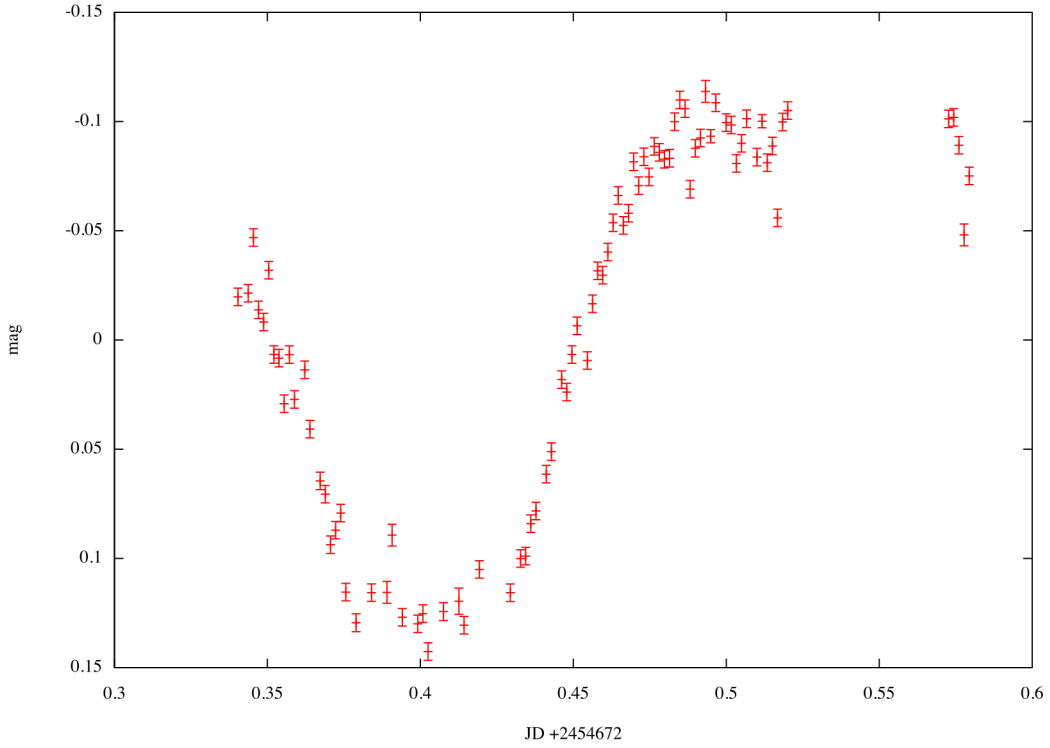

Figure 5.9: Light curve of USNO-A2.0 1350-17532765

Using Equation 2.4 and the extracted values for a and b we find the radius of the planet to be:

$$R_p = \sqrt{\Delta F} R_* = 0.093 R_\odot = 67605.29 Km = 0.94 R_{jup} \tag{5.2}$$

We see that such a drop to the luminosity of a K0 MS star, can be caused by an occulting body 0.94 times the size of Jupiter. The above analysis makes the star USNO-A2.0 1275-17995584 a promising planetary host.

5.5 Conclusions

The goal of this Diploma thesis was the design of an extra-solar planetary survey with the use of the instrumentation offered by the Aristotle University of Thessaloniki's Observatory. To sum up the results:

1. We have created a model for transiting planets with the use of *Gimenez* (2006) [22] equations
2. We have developed a new method for calculating the Transit Detection Probability and locating the promising spots for observation on the celestial plane
3. We performed laboratory tests and observations in order to select and test the equipment to be used in our survey

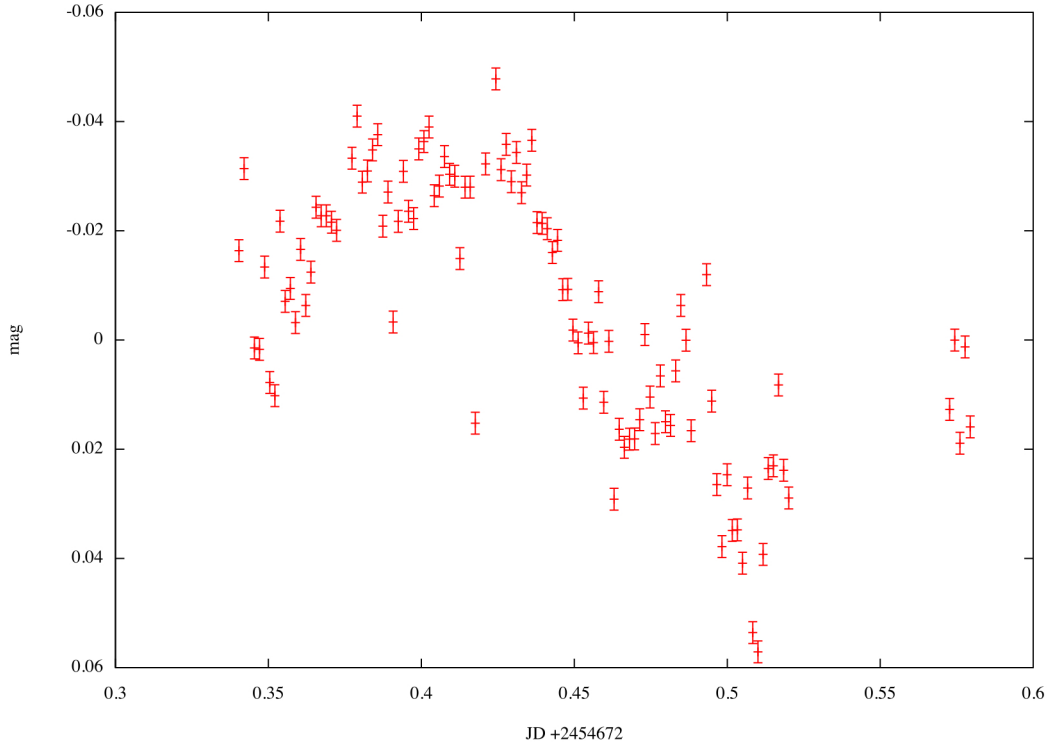

Figure 5.10: Light curve of USNO-A2.0 1275-17982205

4. We developed a Pipeline that is able to reduce the data of multi-star, multi- frame photometric surveys

5. We tested the pipeline in various data -sets

6. We located 10 variable candidates in the field of NGC 129, M92 and ExoField-1 and one transit candidate in the field of ExoField-1

The above results show that we would expect a large number of discoveries with the initiation of the actual observations. Regarding to the preliminary results described here, the achieved photometric precision is high enough to allow us the mili-magnitude monitoring of stars up to the magnitude of $\backsim$ 14. With this performance, the *Holomon Transit Survey* is able to gain a significant role among other planetary surveys in the future. The potential discoveries of extra solar planets would help us to expand the knowledge for these exciting objects, so resent to the cognizance of the Humanity but yet, so important for answering its age-old questions .

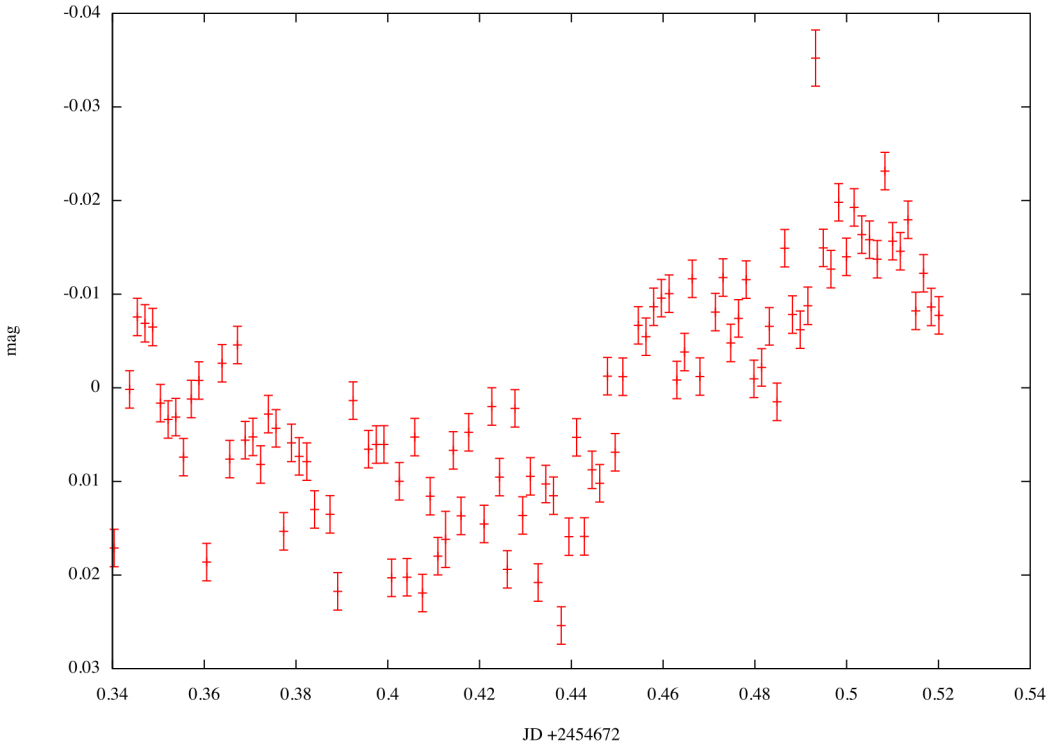

Figure 5.11: The first test field as captured by our instrument.

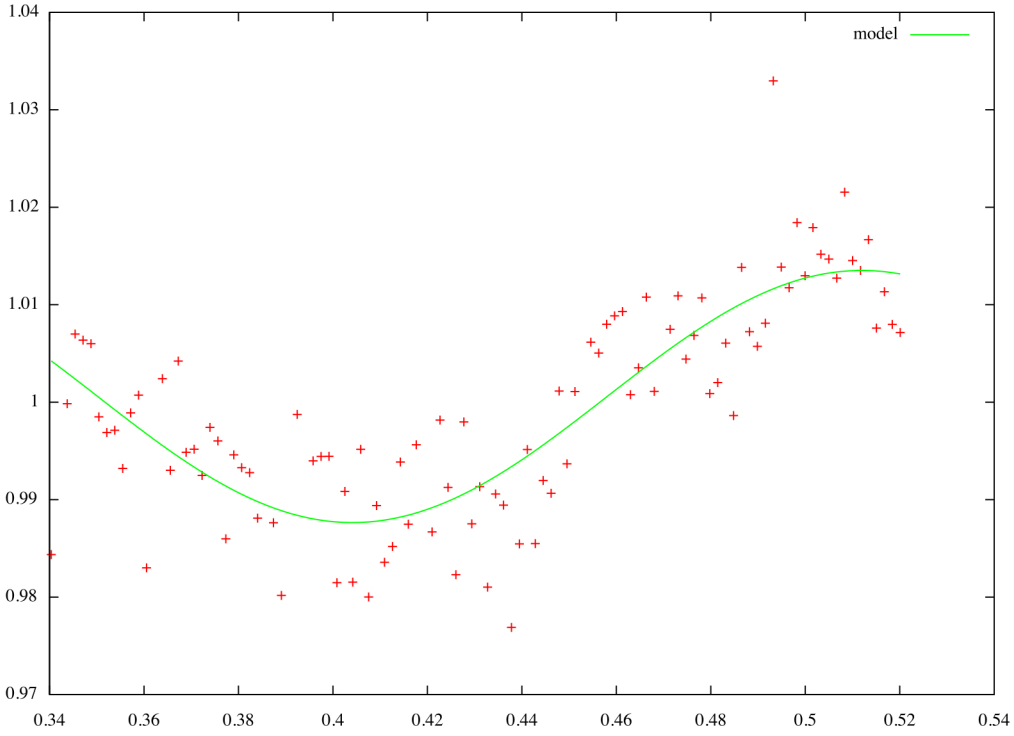

Figure 5.12: Model of the transit candidate

A
DATA REDUCTION PIPELINE

THRET.CSH

```
#! /bin/csh
#last change 01/11/08 J.A ----> Changes in "sigma.for" and "sigma.csh"

clear

#initialize starlink
source /star/etc/cshrc
source /star/etc/login
figaro>junk
pisa>junk
kappa>junk
extractor >junk
daostart>junk
rm -f junk

# welcome screen
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
```

```
echo
echo
echo
echo " ###################################################################"
echo " #                                                                 #"
echo " #                           T  H  Re   T                          #"
echo " #                                                                 #"
echo " #                Software for large image reduction               #"
echo " #                             and                                 #"
echo " #                          photometry                             #"
echo " #                                                                 #"
echo " #                                                                 #"
echo " #                                                                 #"
echo " #                           AUTHOR                                #"
echo " #                       John Antoniadis                           #"
echo " #                            2 0 0 8                              #"
echo " #                            V 0.5                                #"
echo " ###################################################################"
echo
echo
echo
echo
echo

#initialize variables
set home=‘pwd‘
setenv ADAM_DIR $home/adam
echo "PLEASE ENDER THE FULL PATH OF THE DIRECTORY THAT CONTAINS YOUR DATA >"
set directory=$<
echo "PLEASE ENDER THE RA COORDINATES OF THE TARGET(hh mm ss)             >"
set ra="22 53 40"
echo "PLEASE ENDER THE DEC COORDINATES OF THE TARGET(deg mm ss)           >"
set dec="44 44 55"
echo "PLEASE ENDER THE MAIN NAME OF THE BIAS FRAMES                       >"
set bias=$<
echo "EXTENTION ?                                                        >"
set bext=$<
echo "PLEASE ENDER THE MAIN NAME OF THE DARK FRAMES                       >"
set dark=$<
echo "EXTENTION ?                                                        >"
set dext=$<
echo "PLEASE ENDER THE MAIN NAME OF THE FLAT FRAMES                       >"
set flat=$<
```

```
echo "EXTENTION ?                                                    >"
set fext=$<
echo "PLEASE ENDER THE MAIN NAME OF THE DATA FRAMES                  >"
set data=$<
echo "EXTENTION ?                                                    >"
set ext=$<
echo "PLEASE ENDER X DIMENTION OF     FRAMES                        >"
set maxx=$<
echo "PLEASE ENDER Y DIMENTION OF     FRAMES                        >"
set maxy=$<
@ maxx = $maxx - 10
@ maxy = $maxy - 10
first:
echo "DAYLIGHT SAVINGS ? (y/n)                                       >"
set sav=$<
switch($sav)
case y:
goto sec
breaksw
case n:
goto sec
breaksw
default
echo "UNRECOGNIZED ANSWER PLEASE TRY AGAIN                           >"
goto first
breaksw
endsw
sec:
set minx=1
set miny=1
@ minx = $minx + 10
@ miny = $miny + 10
cd REDUCTION
source convert.csh
```

CONVERT.CSH

```
#! /bin/csh
cd $directory

#make sure that all past lists have been deleted
```

```
if (-e bias.list)then
rm -f bias.list
endif
if (-e dark.list)then
rm -f dark.list
endif
if (-e flat.list)then
rm -f flat.list
endif
if (-e data.list)then
rm -f data.list
endif

if (-e sflat.list)then
rm -f sflat.list
endif


#convert images to starlink format
clear
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo "            CONVERTING IMAGES TO STARLINK FORMAT PLEASE WAIT"
cd $directory
echo "            BIAS FRAMES.................................    "

foreach image($bias*.$bext)
set im=$image:r
rdfits $image $im \\
end

clear
echo
echo
echo
```

```
echo
echo
echo
echo
echo
echo
echo
echo "            CONVERTING IMAGES TO STARLINK FORMAT PLEASE WAIT"
echo "            BIAS FRAMES.................................DONE"

clear
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo "            CONVERTING IMAGES TO STARLINK FORMAT PLEASE WAIT"
echo "            DARK FRAMES.................................    "
foreach image($dark*.$dext)
set im=$image:r
rdfits $image $im \\
echo $im >>dark.list
end

clear
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo "            CONVERTING IMAGES TO STARLINK FORMAT PLEASE WAIT"
echo "            DARK FRAMES.................................DONE"
```

```
clear
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo "            CONVERTING IMAGES TO STARLINK FORMAT PLEASE WAIT"
echo "            FLAT FRAMES.................................    "
foreach image($flat*.$fext)
set im=$image:r
rdfits $image $im \\
echo $im >>flat.list
end
clear
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo "            CONVERTING IMAGES TO STARLINK FORMAT PLEASE WAIT"
echo "            FLAT FRAMES.................................DONE"

clear
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
```

```
echo "            CONVERTING IMAGES TO STARLINK FORMAT PLEASE WAIT"
echo "            DATA FRAMES.................................    "
foreach image($data*.$ext)
set im=$image:r
rdfits $image $im \\
echo $im >>data.list
echo $im >>sflat.list
end
clear
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo "            CONVERTING IMAGES TO STARLINK FORMAT PLEASE WAIT"
echo "            DATA FRAMES.................................DONE"


#run biascheck.csh
cd $home/REDUCTION
source biascheck.csh
```

BIASCHECK.CSH

```
#! /bin/csh

#bias quallity check script

cd $directory


#this section puts the Mean values of each bias frame
#onto a .list file
foreach file($bias*.sdf)
set image=$file:r
set mean=‘istat $image xstart=min ystart=min xend=max yend=max|grep Mean‘
```

```
echo $mean[3] >>biascheck.list
end
unset mean

#summarise all mean values and set it into a variable
awk '{sum +=$1} END{print sum}' biascheck.list >sum
set sum=‘cat sum‘

#obtain the number of frames and set it into a variable
awk 'END{print NR}' biascheck.list >total
set total=‘cat total‘

#calculate the total mean value of all frames
set meanv=‘calc $sum/$total prec=_double‘

#bias quality check subrutine
set mi=‘calc $meanv-3‘  #change this according to the needed precision
set ma=‘calc $meanv+3‘
set min=‘calc "nint($mi)"‘
set max=‘calc "nint($ma)"‘
foreach file($bias*.sdf)
set f=$file:r
set mean=‘istat $f xstart=min ystart=min xend=max yend=max|grep Mean‘
set x=‘calc "nint($mean[3])"‘
@ a= ( $min < = $x && $x < = $max )
if ($a == 1)then
echo $f >>bias.list
endif
end

#cleanup
rm -f total
rm -f biascheck.list

#continue with the next script
cd $home/REDUCTION
source reduce.csh
```

REDUCE.CSH

```
#! /bin/csh
```

```
#bias,dark and flat reduction script
cd $directory

clear
#mastebias
medsky files=bias.list scaled=no output=masterbias
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo "                      CREATING MASTER BIAS.......DONE"

#masterdark
foreach dark($dark*.sdf)
set frame=$dark:r
isub $frame masterbias $frame
set exp=‘fitskeys $frame|grep EX‘
icdiv $frame $exp[2] $frame
end
medsky files=dark.list scaled=no output=masterdark
echo "                      CREATING MASTER DARK........DONE"

#masteflat
foreach flat($flat*.sdf)
set frame=$flat:r
isub $frame masterbias $frame
set exp=‘fitskeys $frame|grep EX‘
icmult masterdark $exp[2] junk
isub $frame junk $frame
rm -f junk.sdf
end
medsky files=flat.list scaled=yes output=masterflat
set mean=‘istat masterflat xstart=min ystart=min xend=max yend=max|grep Mean‘
icdiv masterflat $mean[3] masterflat
echo "                      CREATING MASTER FLAT........DONE"

#reduce the data
```

```
clear
foreach d($data*.sdf)
set frame=$d:r
isub $frame masterbias $frame
set exp=‘fitskeys $frame|grep EX‘
icmult masterdark $exp[2] junk
isub $frame junk $frame
rm -f junk.sdf
idiv $frame masterflat $frame
surfit $frame junk
set mean=‘istat junk xstart=min ystart=min xend=max yend=max|grep Mean‘
icdiv junk $mean[3] junk
idiv $frame junk $frame
echo "                        REDUCTING FRAME $frame........DONE"
rm junk.sdf
end


#clear the mess
rm -f $bias*.sdf
rm -f $dark*.sdf
rm -f $flat*.sdf

#continue with next script
cd $home/PHOTOMETRY
source allstar.csh


#!/bin/csh
cd $directory
#extractor >junk
rm -f junk
#  Loop over the files we’re interested in.
foreach file($data*.sdf)

#  Get filename without the .sdf extension.
set fileroot=$file:r

#  Write filename to screen.
         echo "***"
         echo "*** Processing file $file ***"
         echo "***"
```

```
extract $fileroot config=/star/bin/extractor/config/default.sex
set test=‘cat test.cat‘
set arithmitis=‘awk ’{ sum += $1}; END {print sum }’ test.cat‘
set paranomastis=‘echo $#test‘
set fwhm=‘calc $arithmitis/$paranomastis prec=_double‘
set a1=‘calc "0.05*$fwhm" prec=_double‘
set a2=‘calc "0.08*$fwhm" prec=_double‘
set a3=‘calc "0.2*$fwhm" prec=_double‘
set a4=‘calc "0.5*$fwhm" prec=_double‘
set a5=‘calc "0.8*$fwhm" prec=_double‘
set a6=‘calc "1*$fwhm" prec=_double‘
set a7=‘calc "1.5*$fwhm" prec=_double‘
set a8=‘calc "2.1*$fwhm" prec=_double‘
set a9=‘calc "3*$fwhm" prec=_double‘
set is=‘calc "4*$fwhm" prec=_double‘
set os=‘calc "4.7*$fwhm" prec=_double‘
#  Run DAOPHOT with the appropriate commands.
daophot <<__DAOPHOT-END__
5.53
1.28
OPTIONS
KEYBOARD  INPUT
FW=$a6
TH=4.0
LOW=4
FI=$a8
PSF=$a9
MS=50000
HI=65000
WA=-2.0
VA=-1
EXTRA=9.00

ATTACH $fileroot
FIND
1,1
$fileroot.po
Y

PHOTOMETRY
photo.opt
A1 =0.5
A2 =1.5
```

```
A3 =2
A4 =3
A5 =4
A6 =4.5
A7 =5
A8 =6
A9 =7
AA =8
AB =8.5
IS =9.5
OS =14
OK
$fileroot.po
$fileroot.ap
EXIT
__DAOPHOT-END__
#PICK
#$fileroot.ap
#30
#$fileroot.lst

#PSF
#$fileroot.ap
#$fileroot.lst
#$fileroot.psf
#EXIT
#__DAOPHOT-END__

allstar <<EOF
IS=9.5
OS=14
CE=6.00
CR=2.50
FITTING RADIUS=7.5
MS=40000.00

$fileroot
ref.psf
$fileroot.ap
$fileroot.als
$fileroot's'
EOF
```

```
end
#  Get rid of scratch file that DAOPHOT has written.
rm $fileroot'jnk'.sdf
rm *.po
#rm *.ap
rm *.st
#rm *.psf
rm *.grp
rm *.nei
rm *.cmb
rm *s.*
rm *jnk.*

#  End of loop.
#end

cd $home/CURVES
clear
rm *.list
source timer.csh
```

TIMER.CSH

```
#!bin/csh

cd $directory

# fortran script timer.for compiling
cp $home/CURVES/timer.for .
ifort timer.for
echo $ra >ra.txt
echo $dec >dec.txt
#loop over the sdf files
foreach file($data*.sdf)
set root=$file:r

#grep all necesary information from the fits header
set time=‘fitskeys $root|grep DATE-OBS‘
echo $time[2] >ut.txt
set h=‘awk '{print substr($0,12,2)}' ut.txt‘
set m=‘awk '{print substr($0,15,2)}' ut.txt‘
set s=‘awk '{print substr($0,18,2)}' ut.txt‘
```

```
echo $h $m $s >ut.txt
set date=‘fitskeys $root|grep DATE-OBS‘
echo $date[2]>date.txt
set year=‘awk ’{print substr($0,1,4)}’ date.txt‘
set month=‘awk ’{print substr($0,6,2)}’ date.txt‘
set day=‘awk ’{print substr($0,9,2)}’ date.txt‘

echo $year $month $day >date.txt

./a.out
set jd=‘cat jdates.txt‘
set air=‘cat airmass.txt‘
#set jda=‘awk ’{print $1}’ jdates.txt‘
#set jdb=‘awk ’{print $2}’ jdates.txt‘
#set jdf=‘calc $jda+$jdb prec=_double‘
echo $jd >$root.jd
echo $air>$root.air
end
rm -f a.out
rm -f jdates.txt
rm -f date.txt
rm -f ut.txt
cd $home/CURVES
source ref.csh
```

TIMER.FOR

```
      real*8 rahour, decdeg, hour, min ,sec, airmass
      real*8 lat, lon, la, d, a
      real*8 ramin, rasec, decmin, decsec, ra, dec, exp, de
      real*8 hourangle, lst, gmst, jd, jdo, ut
      integer day, month, year

      !open files
      open (1, file='ra.txt')
      read (1,*) rahour, ramin, rasec
      open (2, file='dec.txt')
      read (2,*) decdeg, decmin, decsec
      open (3, file='ut.txt')
      read (3,*) hour, min, sec
```

```
open (4,file='date.txt')
read (4,*) year, month, day
open (5,file='jdates.txt')
open (6, file='airmass.txt')
!Constants
lat =  40.43277        !latitude Lat/lon ....
lon = -23.50527        !longintute ( - ) means east ( + ) means west

!Find Coordinates in decemical form
ra = rahour + (ramin / 60) + (rasec / 3600)
if ( 0.gt.decdeg ) then
decdeg =  abs(decdeg)
dec = -(decdeg + (decmin / 60) + (decsec/3600))
else
dec = decdeg + (decmin / 60) + (decsec/3600)
end if

!Compute UT time - Universal Time
ut = hour +( min/60) + (sec/3600)

!Compute Julian Date
jd = 367*year - int((( year + int((month + 9) / 12 ))) * 1.75)
jd = jd + int((month) * 30.55555 ) + day + 1721013.5 + (ut / 24)

!Compute GMST - Greenwich Mean Sidereal Time
jdo = jd - (ut/24)
d = jdo - 2451545
gmst = 6.69738 + (0.0657 * d)        !for H = 0

!Compute LST - Local Sidereal Time
lst = (gmst - (lon*0.06666666666666667)) + (1.00274 * ut) !The *0.6
is to convert lon to hours

!Compute Hour Angle
hourangle = lst - ra
hourangle = hourangle * 15

!Compute Air Mass
la = lat * 0.01745329          !Convert to radians !same for la
de = dec * 0.01745329          !Convert to radians !de is a new variable fo
leaving dec value without change
hourangle = hourangle * 0.01745329
airmass = (sin(la)*sin(de))+(cos(la)*cos(de)*cos(hourangle))
```

```
       airmass = 1 / airmass

       !Compute Correction for low altitudes
       a = airmass-0.0018167*(airmass-1)-0.002875*(airmass-1)**2
       airmass = a - 0.0008083 * (airmass - 1 )**3

       write (5,*) jd
       write (6,*) airmass
       end
```

REF.CSH

```
#! bin/csh

cd $directory
cp $home/CURVES/ref.for .
rm *.list
clear
echo
echo
echo
echo
echo  "   ************** Selecting referance Please wait **************"

if (-e junk)then
rm -f junk
endif

if (-e a.out)then
rm -f a.out
endif

rm *.lst
cp $home/match .



#this is for apperture photometry
#foreach file(*.ap)
#set r=$file:r
#awk '{if (NR>3&&NF!=0) print $1,$2,$3,$12*100}' $file >file
```

```
#set n=‘awk 'END{print NR}' file‘
#set id=‘awk '{print $1}' file‘
#set x=‘awk '{print $2}' file‘
#set y=‘awk '{print $3}' file‘
#set mag=‘awk '{print $4}' file‘
#set i=1
#@ f = $n / 2
#while( $i <= $f)
#@ k = 2 * $i - 1
#echo $id[$k] $x[$k] $y[$k] $mag[$k] >> $r.list
#@ i = $i + 1
#end
#end

#foreach file(*.list)
#awk '{if ($4!=9999.9) print $0}' $file >junk
#mv junk $file
#end



#this is for psf photometry
foreach file($data*.als)
set d=$file:r
awk '{if (NR>3&&NF!=0) print $1, $2, $3, $4*10000 substr($5,3,4)}' $file >$d.list
rm -f file
end

foreach file(*.list)
set k=‘awk 'END{print NR}' $file‘
set root=$file:r
set i=1
set x=‘awk '{print $2}' $file‘
set y=‘awk '{print $3}' $file‘
set mag=‘awk '{print $4}' $file‘
set jd=‘cat $root.jd‘
#set er=‘awk '{print $5}' $file‘
while ( $i <= $k )
echo $i $x[$i] $y[$i] $mag[$i] $jd >> junk
@ i = $i + 1
end
cat junk >$file
rm junk
```

```
end

#foreach file(*.list)
#awk '{if ($5<0.05) print $1, $2, $3, $4, $5}' $file >junk
#cat junk >$file
#end
#rm -f junk

foreach file(*.list)
awk 'END{print NR}' $file >>junk
end
set n=`awk 'END{print NR}' junk`
echo $n>n
set num=`cat junk`
echo $num>num
rm -f junk
ifort ref.for
./a.out
rm -f num
rm -f n
set max=`cat mtBres`

foreach file(*.list)
set num=`awk 'END{print NR}' $file`
if ($max == $num)then
echo $file >>posref
endif
end

set refall=`head -1 posref`
set ref=$refall:r
rm -f posref
rm -f mtBres
cat $ref.list > referance
cd $home/CURVES
source matcher.csh
```

ref.for

```
      dimension A(10000)
       integer n ,x
       open (1,file='n')
```

```
      read (1,*) n
      open (2,file='num')
      read (2,*) (A(x),x=1,n)
      open (3,file='mtBres')
      k=maxval(A)
      write (3,*) k
      end
```

matcher.csh

```
#!bin/csh

cd $directory
rm *.lst
cp $home/CURVES/match .
rm *.als


foreach file($data*.list)
clear
echo
echo
echo
echo
set f=$file:r
./match $file 1 2 3 referance 1 2 3 outfile=$f id1=0 id2=0 cubic
 trirad=0.0001 matchrad=2 nobj=70
echo "Doing file $file"
end
cd $home/CURVES
source curver.csh
```

CURVER.CSH

```
#!bin/csh

cd $directory
cat jd.txt > jd
rm *.txt
```

```
foreach file(*.mtA)
set froot=$file:r
awk '{print $1}' $froot.mtB>id
set id=‘cat id‘
awk 'END{print NR}' $froot.mtB>n
set n=‘cat n‘
echo "******** Ploting stars in file $froot ******** "
set jd=‘cat $froot.jd‘
set air=‘cat $froot.air‘
set i=1
awk '{print $4}' $file > mag
awk '{print substr($1,1,5)}' mag >junk
set mag = ‘awk '{print $1/1000}' junk‘
awk '{print substr($1,6,4)}' mag >junk
set er = ‘awk '{print $1/1000} ' junk‘
while ( $i <= $n)
#set mm=‘cat $froot.list‘
echo $jd  $mag[$i] $er[$i] $air >> $id[$i].curve
@ i = $i + 1
end
end


#foreach file(*.txt)
#set f=$file:r
#awk '{printf("%s",$1,$2);getline<"jd";print $1}' $file > junk
#awk '$3=$1,$4=$2 {print $4,$3}' junk >$f.curve
#rm -f $file
#end
#rm junk
rm *.txt
#rm *.mtA
#rm *.mtB
#rm *.jd
#rm *.list
#rm *.sdf
rm *.unB
rm id
rm sum
rm *.for
rm mag
```

```
cd $home/Analysis

source sigma.csh
```

SIGMA.CSH

```
cd $directory
cp $home/Analysis/sigma.for .
rm *.txt
ls -l $data*.mtA >junk

set k=‘awk 'END{print NR}' junk‘

foreach file(*.curve)
clear
echo
echo
echo
echo
  set b=‘awk 'END{print NR}' $file‘

  if ($b > 48)then
    echo "CHECKING NUMBER OF POINTS ON FILE $file.....OK"
  else
    echo "CHECKING NUMBER OF POINTS ON FILE $file......REMOVED"
    rm $file
  endif
end

if (-e junk)then
rm junk
endif

if (-f time)then
rm time
endif

if(-f air)then
rm air
endif

foreach file(*.jd)
```

```
set rr=$file:r
cat $file >>time
cat $rr.air >>air
end
set time=‘cat time‘
set air=‘cat air‘
set tnr=‘awk ’END{print NR}’ time‘

ifort sigma.for
foreach file(*.curve)
echo
echo
echo
set n=‘awk ’END{print NR}’ $file‘
echo $n > n
echo $n > q
awk ’{print $2}’ $file > mag
awk ’{print $3}’ $file >error
./a.out
set mean=‘cat mean.txt‘
set sigma=‘cat sigma.txt‘
set inner=‘cat inner.txt‘
set pulse=‘cat pulse.txt‘
set rms=‘cat rms.txt‘
set lrms=‘cat logrms.txt‘
set i=1
  while ($i <= $tnr)
    set b=‘cat $file|grep $time[$i]‘
    set nb=$#b
      if ($nb == 4)then
       echo $b >> c$file
      else
       echo $time[$i] $mean $sigma $air[$i] >>c$file
      endif
  @ i = $i + 1
  end
echo $mean $sigma $inner $pulse $rms $lrms $file >> sigmaplot.txt
echo "         FILE $file UNDER ANALYSIS"
echo "         CALCULATED MEAN : $mean"
echo "         CALCULATED SIGMA: $sigma"
echo "         CALCULATED PULSATION: $pulse"
end
```

```
rm junk
cd $home/Analysis/
```

SIGMA.FOR

```
        real*8 s, mean, sum, inner, sum2
        real*8 k, ssum, fac, dsum, pulse
        real*8 calc, inv, mag, siner, sum3
        real*8 rms, lrms
        integer n, x, i, j ,m
        real*8 A(10000), E(10000)

! Open and Read Input
        open (1, file='n')
        read (1,*) n
        open (11, file='q')
        read (11,*) q
        open (2, file='mag')
        read (2,*) (A(x),x=1,n)
        open (3, file='error')
        read (3,*) (E(x),x=1,n)
        inv= 1/q
        i=1

! Calculate mean magnitute (x square fitting)
        sum2=0
        sum3=0
        do i=1,n
         sum2 = sum2 + 1/(E(i)**2)
         sum3 = sum3 + A(i)/(E(i)**2)
        enddo
        mean = sum3/sum2

!Calculate sigma deviation
        do i=1,n
         k= A(i) - mean
         ssum = ssum + k**2
        enddo
        calc = ssum * inv
        sigms = calc - mean**2
        sigma = sqrt(calc)
```

```
!Calculate inner sigma scattering

        fac = 1/(2*(q-1))
        m = n - 1
        dsum = 0
        do i=1,m
         dsum = dsum + (A(i)-A(i+1))**2
        enddo
        siner = dsum*fac
        inner = sqrt(siner)

!Calculate Pulsation factor
        pulse = sigma/inner

!Calculate rms
        rms = 0
        do i=1,n
         rms = rms + ((A(i)-mean)/(E(i)))**2
        enddo

!Calculate log(rms)

        lrms = log(rms)

!Write results to file
        open (3,file='mean.txt')
        open (4,file='sigma.txt')
        open (5,file='inner.txt')
        open (6,file='pulse.txt')
        open (7,file='rms.txt')
        open (8,file='logrms.txt')
        write (3,*) mean
        write (4,*) sigma
        write (5,*) inner
        write (6,*) pulse
        write (7,*) rms
        write (8,*) lrms
        end
```

TFA.CSH

```
#!/bin/csh
#This is a csh script for applying Trend Filtering Algorithm(TFA) to data
#For Holomon survey data use only
#Before use, make shure that all lightcurves are of the same number of points.
#For more information on the TFA algorithm see "Kovacs et al. A&A (2004)"
#Author: John Antoniadis
#Last Change 15/05/08 @ 11:49pm from John Antoniadis


#Initialize
clear
echo "Initializing......"
set tfa=‘pwd‘                                #set the current directory as Home

cd $directory                                #Go to data directory


if (-f sigmaplot.ens)then
  cat sigmaplot.txt > sigmabefore            #backup MAGvsSD
  diagram for later use
endif


mkdir curves                                 #make directory "curves"

cd curves

cp $directory/c*.ens .                      #Copy all lightcurves from parent
directory to "curves"

mkdir base                                    #Create essential subdirectories
"base" and "targets"
mkdir targets


#STEP 1: Zero average all curves
echo "STEP 1: Zero average all curves"
cp $tfa/zero.for .                           #copy fortran code "zero.for"
 from home directory
```

```
ifort zero.for                              #compile the code.Output "a.out"

foreach curve(*.ens)                      #Begin loop over the lightcurves

  awk 'END{print NR}' $curve >n             #use AWK to print number
  of points for the
   current curve
  set n=‘cat n‘                             #Set this number to variable "n"
  awk 'END{print NR}' $curve >q             #Do the same, but this time,
  the number is going to be used
  awk '{print $1}' $curve >jd               #as a real number
  awk '{print $2}' $curve >mag              #use AWK to print Julian date and
   Magnitute
  o files "jd" and "mag"
  awk '{print $3}' $curve >err

  ./a.out                                   #run 'a.out'. Product output is "zero".
  See comments
  in "zero.for" for
  set zero=‘cat zero‘                       #further information
  set jd=‘cat jd‘
  set err=‘cat err‘
  rm -f $curve
set i=1
     while ($i <= $n)                       #Begin while loop


                                            #replace lightcurve with its
                                             zero averaged
echo $jd[$i] $zero[$i] $err[$i] >>$curve
@ i = $i + 1
     end                                    #end of while loop
end                                         #end of loop
rm *.out
echo "DONE"

#cp $directory/sigmaplot2.txt .
#mv cc621.curve tt
#sort sigmaplot2.txt >junk2
#head -50 junk2 >junk
#rm junk2
#set curve=‘awk '{print $3}' junk‘
#set n=$#curve
```

```
#set i=1
#while ($i <= $n)
#cp $curve[$i] base
#@ i = $i + 1
#end
#mv tt cc621.curve
#cp *.curve targets
#rm junk

#STEP 2: Select templates for algorithm's base (comparison stars)
echo "STEP 2: Select templates for algorithm's base (comparison stars)"
cp $tfa/templates.for .                  #copy fortran code "templates.for" from
 home directory

ifort templates.for                        #compile it to "a.out"

foreach file(*.ens)                      #begin loop over the files
  set n=‘awk 'END{print NR}' $file‘     #set number of lightcurve's
  points to
  variable "n" and
   echo $n > n                           #then to files "n" and "q" to
    be used as an integer
    and
   echo $n > q                           #as a real number
   awk '{print $2}' $file > mag          #copy magnitute row to file "mag".
   This is gona be
   used from fortran
   ./a.out                               #Run 'a.out'. This produces the
    output "sigma.txt"


echo >base.junk
echo >base2.junk
awk '{if ($3<=2.00) print $0}' sigma.txt > base.junk
awk '{if ($2<=0.15) print $0}' base.junk >base2.junk
set junk=‘cat base2.junk‘
set logic=‘echo $#junk‘

if ($logic == 3)then
  mv $directory/curves/$file $directory/curves/base
else
  mv $directory/curves/$file $directory/curves/targets
endif
```

```
rm base.junk
end
rm *.out

echo "DONE"



#STEP 3: Create Base matrix
echo "STEP 3: Create Base matrix"
cd base                                  #move to directory "base"

foreach file(c*.ens)                    #Begin loop over the curves

awk 'END{print NR}' $file >n             #print number of points
 to file "m"
set mag=‘awk '{print $2}' $file‘         #set magnitute row to
 variable "mag"
echo $mag >> matrix                      #print "mag" as a line to
 file "matrix"

end                                      #end of loop


                                         #file "matrix" now containing
                                         magnintute points from
                                          each base curve
                                         #as a line. It has n lines, where n
                                         is the number of
                                         lightcurves used for base
                                         #and m rows, where m is the
                                         number of points
                                         from each curve

echo "DONE"


#STEP 4: Create Normal Matrix
echo "STEP 4: Create Normal Matrix"
cat matrix >matj
cat matrix >matk
cp $tfa/matrixn.for .
```

```
awk 'END{print NR}' matrix > m
ifort matrixn.for
./a.out
set base=‘pwd‘
echo "DONE"

#STEP 5: Inverse Normal Matrix
echo "STEP 5: Inverse Normal Matrix"
cp $tfa/inverse2.for .
rm *.out
cat m >m2
ifort inverse2.for
./a.out
rm a.out
echo "DONE"


#starting point of correcting loop
#STEP 6: Compute scalar products of the target and
template time series
echo "STEP 6: Compute scalar products of the target
 and template time series"
echo "&STEP 7:Compute the solution for {Ci}"
cd ..
cd targets
cp $base/matrix .                   #Base matrix
cp $base/n .                        #Number of rows
cp $base/m .                        #Number of lines
cp $tfa/scalar.for .
cp $tfa/Csolution.for .
cp $tfa/correct.for .
cp $directory/time .
cp $base/G .                        #Inverse matrix
foreach curve(*.ens)
  awk '{print $2}' $curve >mag
  awk '{print $3}' $curve >err
  set cur=$curve:r
  ifort scalar.for
  ./a.out

#STEP 7:Compute the solution for {Ci}
  awk '{print $2}' $curve > y
  ifort Csolution.for
```

```
  ./a.out
  ifort correct.for
  ./a.out
  set n=‘cat n‘
  set jd=‘cat time‘
  set i=1
  set mag=‘cat magf‘
  while ($i <= $n)
   echo $jd[$i] $mag[$i] >> cor$curve
   @ i = $i + 1
  end
echo "Reconstucting lightcurve  $cur"
end
echo "DONE"

#Compute Statistics
cp $home/Analysis/sigma2.for .
foreach file(corc*.ens)
set n=‘awk 'END{print NR}' $file‘
echo $n > n
echo $n > q
awk '{print $2}' $file > mag
ifort sigma2.for
./a.out
set mean=‘cat mean.txt‘
set sigma=‘cat sigma.txt‘
set pulse=‘cat pulse.txt‘
set inner=‘cat inner.txt‘
echo $mean $sigma $inner $pulse $file >> sigmaplot.txt
echo "          file $file under analysis"
echo "          Calculated mean : $mean"
echo "        Calculated deviation: $sigma"
echo "        Calculated pulsation: $pulse"
end

rm junk
cd $tfa
source sbsa.csh
```

ZERO.FOR

```
real q, sum, inv, mean, esum
integer n, i, x
dimension a(10000), b(10000), e(10000)

open (1,file='n')
read (1,*) n
open (2,file='mag')
read (2,*) (a(x),x=1,n)
open (3,file='q')
read (3,*) q
open (5,file='err')
read (5,*) (e(x), x=1,n)
inv=1/q
do i=1,n
if (a(i).ne.0)then
   sum = sum + a(i)/(e(i)**2)
   esum = esum + 1/(e(i)**2)
else
q=q-1
endif
enddo
inv=1/q
mean=sum/esum
do i=1,n
if (a(i).ne.0)then
   b(i) = a(i) - mean
else
b(i) = a(i)
endif
enddo
open (4,file='zero')
write(4,*) (b(x),x=1,n)
end
```

TEMPLATES.FOR

```
 double precision s, mean, sum, q
 double precision k, ssum, fac
 double precision calc, inv, mag, dsum
```

```
        double precision siner, inner, pulse
        integer n, x, i, j, m
        dimension A(10000)

        open (1, file='n')
        read (1,*) n

        open (11, file='q')
        read (11,*) q

        open (2, file='mag')
        read (2,*) (A(x),x=1,n)

        inv= 1/q

        do i=1,n
          sum= sum + A(i)
        enddo

        mean= inv * sum

        do i=1,n
          k= A(i) - mean
          ssum = ssum + k**2
        enddo

        calc = ssum * inv

        sigma = sqrt(calc)


        fac = 1/(2*(q-1))
        m = n - 1
        dsum = 0
        do i=1,m
         dsum = dsum + (A(i)-A(i+1))**2
        enddo
        siner = dsum*fac
        inner = sqrt(siner)

!Calculate Pulsation factor
        pulse = sigma/inner
```

```
open (4,file='sigma.txt')
write (4,*) mean, sigma, pulse

end
```

MATRIXN.FOR

```
dimension g(10000,10000), x(10000,10000)
integer j, k, i, q, n, m
open (1,file='n')
read (1,*) n            ! rows!
open (2,file='m')
read (2,*) m            !lines
open (4,file='matrix')
read(4,*) ((x(i,q),q=1,n),i=1,m)
do j=1,m
     do k=1,m
        do i=1,n
           g(j,k)=g(j,k)+x(j,i)*x(k,i)
        enddo
     enddo
enddo
open (5,file='a')
  do j=1,m
    write(5,*) (g(j,k),k=1,m)
  enddo
end
```

INVERSE.FOR

```
SUBROUTINE ludcmp(a,n,np,indx,d)
INTEGER n,np,indx(n),NMAX
REAL d,a(np,np),TINY
PARAMETER (NMAX=500,TINY=1.0e-20)
INTEGER i,imax,j,k
REAL aamax,dum,sum,vv(NMAX)
d=1.
do i=1,n
  aamax=0
```

```
  do j=1,n
   if (abs(a(i,j)).gt.aamax) aamax=abs(a(i,j))
  enddo
  if (aamax.eq.0.) pause 'singular matrix in ludcmp'
  vv(i)=1./aamax
enddo
do j=1,n
  do i=1,j-1
    sum=a(i,j)
    do k=1,i-1
      sum=sum-a(i,k)*a(k,j)
    enddo
    a(i,j)=sum
  enddo
  aamax=0.
  do i=j,n
     sum=a(i,j)
     do k=1,j-1
       sum=sum-a(i,k)*a(k,j)
     enddo
     a(i,j)=sum
     dum=vv(i)*abs(sum)
     if (dum.ge.aamax) then
        imax=i
        aamax=dum
     endif
  enddo
  if (j.ne.imax)then
      do k=1,n
        dum=a(imax,k)
        a(imax,k)=a(j,k)
        a(j,k)=dum
      enddo
      d=-d
      vv(imax)=vv(j)
  endif
  indx(j)=imax
  if(a(j,j).eq.0.)a(j,j)=TINY
  if(j.ne.n)then
   dum=1./a(j,j)
   do i=j+1,n
     a(i,j)=a(i,j)*dum
   enddo
```

```
  endif
enddo
return
END

SUBROUTINE lubksb(a,n,np,indx,b)
INTEGER n,np,indx(n)
REAL a(np,np),b(n)
INTEGER i,ii,j,ll
REAL sum
ii=0
do i=1,n
 ll=indx(i)
 sum=b(ll)
 b(ll)=b(i)
  if (ii.ne.0)then
    do j=ii,i-1
      sum=sum-a(i,j)*b(j)
    enddo
  else if (sum.ne.0.) then
   ii=i
  endif
 b(i)=sum
enddo
do i=n,1,-1
  sum=b(i)
   do j=i+1,n
    sum=sum-a(i,j)*b(j)
   enddo
b(i)=sum/a(i,i)
enddo
return
END


integer np, indx(1000), i, j
integer n, q
real a(1000,1000), y(1000,1000)
open (1,file='m')
read(1,*) np
open (5,file='m2')
read(5,*) n
open (2,file='a')
```

```
      read(2,*) ((a(i,q),q=1,n), i=1,n)

      do i=1,n
         do j=1,n
           y(i,j)=0.
         enddo
         y(i,i)=1.
      enddo
      call ludcmp(a,n,np,indx,d)
      do j=1,n
        call lubksb(a,n,np,indx,y(1,j))
      enddo
      open (4,file='G')
      do np=1,n
         write(4,*) (y(np,i), i=1,n)
      enddo
      end
```

SCALAR.FOR

```
  integer n, j ,k, r
      real h(10000), matrix(10000,10000), y(10000)
      real err(10000), mean, esum
      open (1,file='m')                !Number of lines
      read (1,*) m
      open (2,file='n')                !Number of rows
      read (2,*) n
      open (3,file='matrix')
      read(3,*) ((matrix(k,j),j=1,n),k=1,m)
      open (4,file='mag')
      read(4,*) (y(r), r=1,n)
      open (6,file='err')
      read(6,*) (err(r), r=1,n)
      open (5,file='curve.h')

      mean = 0.0
      do i=1,n
      mean = mean + (y(i)/(err(i))**2)
      esum = esum + 1/(err(i)**2)
      enddo
      mean = mean/esum
```

```
do j=1,m
    do i=1,n
       h(j)= h(j) + (y(i)-mean)*matrix(j,i)
   enddo
enddo
write (5,*) (h(r), r=1,m)
end
```

CSOLUTION.FOR

```
 dimension c(10000), G(10000,10000)
 dimension h(10000)
 integer m, j, k

 open (1, file='m')
 read (1,*) m
 open (2, file='G')
 read (2,*) ((G(j,k),k=1,m),j=1,m)
 open (6,file='curve.h')
 read (6,*) (h(k),k=1,m)

 do j=1,m
   c(j)=0
 enddo
 do j=1,m
   do k=1,m
       c(j)=c(j)+G(j,k)*h(k)
   enddo
 enddo


 open(4,file='c')
 write(4,*) (c(j),j=1,m)
 end
```

CORRECT.FOR

```
dimension y(10000), c(10000)
dimension x(10000,10000)
integer n, m, i, j
```

```
      open(1,file='n')
      read(1,*) n
      open(2,file='m')
      read(2,*) m
      open(3,file='y')
      read(3,*) (y(i), i=1,n)
      open(4,file='c')
      read(4,*) (c(j), j=1,m)
      open(5,file='matrix')
      read(5,*) ((x(i,j),j=1,n),i=1,m)
      do i=1,n
         do j=1,m
             y(i)= y(i) -c(j)*x(j,i)
         enddo
      enddo
      open(6,file='magf')
      write(6,*) (y(i),i=1,n)
      end
```

AIRERROR.CSH

```
#!/bin/csh
clear
cd $directory
cp $home/Analysis/air.for .
foreach curve(cc*.curve)
set r=$curve:r
awk '{print $2}' $curve > mag$r.zr
awk '{print $3}' $curve >mag$r.error
end
awk '{print $4}' $curve >airmass.txt
clear
echo
echo
echo


cp airmass.txt air
set arr = `cat air`
```

```
# ----------------- SET STEP OF ALGORITHM

set no = 1
set fin = 4

# ----------------- STARTING MAIN SCRIPT
while ( $no <= $fin )


set arr = ‘cat air‘

# ----------------- F-FACTOR
rm -f cf

foreach i (mag*.zr)
end
 awk 'END{print NR}' $i > n
 set end = ‘cat n‘
 set start = 1

 while ( $start <= $end)
   rm -f cfactor
   foreach i ( mag*.zr)
     set f = $i:r
     echo
     echo
     echo
     echo '                FILE : '$f' IN POINT '$start' FROM '$end' POINTS TOTAL
     cp $i mag
     cp $f.error err
     set mm = ‘cat mag‘
     set ee = ‘cat err‘
     echo $mm[$start]' '$ee[$start]' '$arr[$start] > tmp
     set filt = $ee[$start]

     if ($filt == '0.0') then
     else
     awk '{print $1*$3/($2**2), ($3**2)/($2**2)}' tmp >> cfactor
     endif
     clear
   end
   awk '{print $1}' cfactor > arth
   awk '{print $2}' cfactor > parn
```

```
   awk 'END{print NR}' arth > n
   awk '{ s += $1} END {print s} ' arth > we
   awk '{ s += $1} END {print s} ' parn >> we
   set rt = `cat we`
   echo $rt[1]' '$rt[2] > tte
   awk '{print $1/$2}' tte >> cf
   @ start = $start + 1
 end

# ----------------- A-FACTOR

rm -f air_new
set rg = `cat cf`
foreach i (mag*.zr)
set r = $i:r
echo 'CALCULATE AIRij FACTOR - FILE : '$r
awk '{print $1}' $i > temp1
awk '{print $1}' $r.error > temp2
awk 'END{print NR}' temp2 > n
set end = `cat n`
set start = 1

set rt = `cat temp2`
set rmm = `cat temp1`
rm -f new

while ( $start <= $end )
set rr = $rt[$start]
if ($rr == '0.0' ) then
else
echo $rmm[$start]' '$rt[$start]' '$rg[$start] >> new
endif

@ start = $start + 1
end


awk '{print $1*$3/($2**2)}' new > arth
awk '{print ($3**2)/($2**2)}' new > parn
awk '{ s += $1} END {print s} ' arth > we
awk '{ s += $1} END {print s} ' parn >> we

set rt = `cat we`
```

```
echo $rt[1]' '$rt[2]' '$i > tte
awk '{print $1/$2, $3}' tte >> air_new


end


# -------------------- LIGHTCURVE CORRECTION


awk '{print $1}' air_new > aij
awk '{print $2}' air_new > filename
set af = `cat aij`
set fln = `cat filename`
awk 'END{print NR}' filename > n
set end = `cat n`
set start = 1

ifort air.for

while ( $start <= $end )
echo 'REDUCTION PASS No : '$no' FILE idd = '$start' FROM '$end' TOTAL'
cp $fln[$start] temp
awk 'END{print NR}' temp > n
echo $af[$start] > flag
./a.out
mv fort.9 $fln[$start]
@ start = $start + 1
end
awk '{print $1}' air_new > air
@ no = $no + 1
end
```

AIR.FOR

```
      real mag fl
      integer x, y, z, q
      dimension a(1)
      dimension b(100000)
      dimension c(100000)
      dimension d(1)
```

```
open (1,file='n')
read (1,*) (a(x), x=1, 1)

n = a(1)

open (2,file='temp')
read (2,*) (b(y), y=1, n)

open (3,file='cf')
read (3,*) (c(z), z=1, n)

open (4,file='flag')
read (4,*) (d(q), q=1, 1)

fl = d(1)

do i=1, n
mag = b(i) - c(i)*fl
write (9,*) b(i)-c(i)*fl
end do
end
```

ENSEMBLE.CSH

```
#!/bin/csh
#Script for ensemble photometry

cd $directory
cp $home/Analysis/ensemble.for .
cp $home/Analysis/ensemblec.for .


#Remove previus files if any
if (-e magnitutes.mat)then
rm magnitutes.mat
endif

if (-e errors.mat)then
rm errors.mat
endif
```

```
clear
echo "ENDER DESIRED NUMBER OF COMPARISON STARS(LESS THAN 300):"
set stars=$<
#Select templates
clear
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo "SELECTING TEMPLATES..............."
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
sort sigmaplot.txt >junk
head -300 junk > junk2
awk '{print $1,$7}' junk2 >junk3
sort junk3 >junk4
head -$stars junk4 >comparisons
set compar=`awk '{print $2}' comparisons`
rm junk*
clear
echo
echo
echo
echo
echo
echo
echo
echo
echo
```

```
echo
echo ".......SELECTING TEMPLATES.....DONE"
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
#Create matrixes for instrumental magnitutes and weighting factors
set i=1
while ( $i <= $stars )
awk '{print $2}' c$compar[$i] >comp$i.mag
awk '{print $3}' c$compar[$i] >comp$i.err
@ i = $i + 1
end

foreach comp(comp*.mag)
set root=$comp:r
set mag=‘cat $comp‘
set err=‘cat $root.err‘
echo $mag >>magnitutes.mat
echo $err >>errors.mat
end
set ee=$#mag
echo $ee >ee
awk 'END{print NR}' magnitutes.mat >ss    #<----Debuged up to here

#Run fortran script and find the solution em(i)
ifort ensemble.for >junk
rm -f junk
./a.out
clear
echo
echo
echo
echo
echo
echo
echo
```

```
echo
echo
echo
echo ".............SOLUTION FOUND................."
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo

#Correct all lightcurves
ifort ensemblec.for >junk
foreach file(c*.curve)
set time=‘awk ’{print $1}’ $file‘
awk ’{print $2}’ $file >mag
set err=‘awk ’{print $3}’ $file‘
set air=‘awk ’{print $4}’ $file‘
set n=‘awk ’END{print NR}’ $file‘
echo $n >n
./a.out
set ensemble=‘cat cor‘
set root=$file:r
set i=2
while( $i<= $n )
echo $time[$i] $ensemble[$i] $err[$i] $air[$i] >>$root.ens
@ i = $i + 1
end
clear
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
```

```
echo "............. CORRECTING LIGHTCURVE.............$file"
echo
echo
echo
echo
echo
echo
echo
echo
echo
echo
end
cd $home/Analysis
source sigmaens.csh
```

ENSEMBLE.FOR

```
      subroutine ludcmp(a,n,np,indx,d)
      integer n, np, indx(100000), NMAX
      real*8 d, a(100000,100000), TINY
      parameter (NMAX=50000, TINY=1.0e-20)
      integer i, imax, j, k
      real*8 aamax, dum, sum, vv(100000)
      d=1.
      do i=1,n
      aamax=0.
      do j=1,n
       if (abs(a(i,j)).gt.aamax) aamax=abs(a(i,j))
      enddo
      if (aamax.eq.0) pause 'singular matrix'
      vv(i)=1./aamax
     enddo
     do j=1,n
      do i=1,j-1
      sum=a(i,j)
      do k=1,i-1
        sum = sum - a(i,k)*a(k,j)
      enddo
      a(i,j)=sum
     enddo
```

```
 aamax=0.
 do i=j,n
   sum=a(i,j)
   do k=1,j-1
    sum=sum-a(i,k)*a(k,j)
   enddo
   a(i,j)=sum
   dum=vv(i)*abs(sum)
   if (dum.ge.aamax) then
      imax=i
      aamax=dum
   endif
 enddo
 if (j.ne.imax)then
   do k=1,n
     dum=a(imax,k)
     a(imax,k)=a(j,k)
     a(j,k)=dum
   enddo
   d=-d
   vv(imax)=vv(j)
 endif
indx(j)=imax
if(a(j,j).eq.0.)a(j,j)=TINY
if(j.ne.n)then
  dum=1./a(j,j)
  do i=j+1,n
    a(i,j)=a(i,j)*dum
  enddo
endif
enddo
return
end

subroutine lubksb(a,n,np,indx,b)
integer n, np, indx(100000)
real*8 a(100000,100000), b(100000)
integer i, ii, j, ll
real*8 sum
ii=0
do i=1,n
ll=indx(i)
sum=b(ll)
```

```
b(ll)=b(i)
if (ii.ne.0)then
do j=ii,i-1
   sum=sum-a(i,j)*b(j)
enddo
else if (sum.ne.0) then
   ii=i
endif
b(i)=sum
enddo
do i=n,1,-1
 sum=b(i)
  do j=i+1,n
   sum=sum-a(i,j)*b(j)
  enddo
  b(i)=sum/a(i,i)
 enddo
 return
 end


integer e, s, ee, ss, np, NMAX, ee2
integer n, i, j, t, me, ms, ww
real*8 m(100000,100000), er(100000,100000)
real*8 w(100000,100000), a(100000,100000), TINY
real*8 b2(100000),em(100000),er2(100000,100000)
real*8 b(100000), m2(100000,100000)
real*8 d, aamax, dum, sum, vv(100000)
integer indx(100000), imax, k, ii, ll
open(1,file='ss')
read(1,*) ss
open(2,file='ee')
read(2,*) ee
open(3,file='magnitutes.mat')
read(3,*) ((m2(i,j),j=1,ee),i=1,ss)
open(4,file='errors.mat')
read(4,*) ((er2(i,j),j=1,ee),i=1,ss)

t = ee + ss - 1

do j=1,ee-1
   e = j + 1
  do s=1,ss
```

```
        m(j,s)=m2(s,e)
      enddo
    enddo

    do j=1,ee-1
       e = j + 1
       do s=1,ss
          er(j,s)=er2(s,e)
       enddo
    enddo

     ee = ee - 1

   !Create statistical weights
    do e=1,ee
      do s=1,ss
       w(e,s)=1/(er(e,s))**2
      enddo
    enddo




    !Build matrix A

    !First quarter of matrix
    do e=1,ee
      do s=1,ss
        a(e,e) = a(e,e) + w(e,s)
      enddo
    enddo

    !Second quarter of matrix
     do e=1,ee
       do s=1,ss
         me=ee+s
         a(e,me)=w(e,s)
       enddo
     enddo


   !Third quarter of matrix
      do e=1,ee
```

```
     do s=1,ss
      ms= ee + s
      a(ms,e)=w(e,s)
     enddo
   enddo

!Fourth quarter of matrix
    do s=1,ss
      me=ee+s
      do e=1,ee
       a(me,me) = a(me,me) + w(e,s)
      enddo
    enddo

 !Build matrix b
     do i=1,ee
       do s=1,ss
         b(i) = b(i) + w(i,s)*m(i,s)
       enddo
     enddo
    do i=1,ss
      j=ee+i
        do e=1,ee
         b(j) = b(j) + w(e,i)*m(e,i)
        enddo
       enddo
 !solve equation ax=b
 np = ee + ss
 n = np

 call ludcmp(a, n, np, indx, d)
 call lubksb(a, n, np, indx, b)

 open (8,file='s1')
 write (8,*) ss, ee
!write results to file
 open(5,file='em.mag')
 write(5,*) (b(i),i=1,n)
 open(6,file='g')
 do i=1,n
 write(6,*) (a(i,j),j=1,n)
 enddo
 end
```

SIGMAENS.CSH

```
cd $directory
cp $home/Analysis/sigma.for .



if (-e junk)then
rm junk
endif

if (-f time)then
rm time
endif

if(-f air)then
rm air
endif

foreach file(*.jd)
set rr=$file:r
cat $file >>time
cat $rr.air >>air
end
set time=‘cat time‘
set air=‘cat air‘
set tnr=‘awk 'END{print NR}' time‘

ifort sigma.for
foreach file(*.ens)
echo
echo
echo
set n=‘awk 'END{print NR}' $file‘
echo $n > n
echo $n > q
awk '{print $2}' $file > mag
awk '{print $3}' $file >error
./a.out
set mean=‘cat mean.txt‘
set sigma=‘cat sigma.txt‘
set inner=‘cat inner.txt‘
```

```
set pulse=‘cat pulse.txt‘
set rms=‘cat rms.txt‘
set lrms=‘cat logrms.txt‘

echo $mean $sigma $inner $pulse $rms $lrms $file >> sigmaplot.ens
echo "          FILE $file UNDER ANALYSIS"
echo "          CALCULATED MEAN : $mean"
echo "          CALCULATED SIGMA: $sigma"
echo "          CALCULATED PULSATION: $pulse"
end

rm junk
```

CORRELATION.CSH

```
#!/bin/csh
set g=‘pwd‘
cd $directory
cp $g/correl.f .
cp $g/correl2.f .
clear
echo
echo
echo
set jd=‘cat time‘

if (-e temp2)then
rm temp2
endif

#Select Comparisons
echo
echo
echo ’                   SELECTING COMPARISONS’
echo ’                        BE PACIENT’
cat sigmaplot.txt >temp
head -35 temp >temp2
sort -gr sigmaplot.txt >temp
head -10 temp >>temp2
set comp1=‘awk ’{print $3}’ temp2‘
set n=$#comp1
```

```
set i=1
while ($i <= $n)
set root=c$comp1[$i]:r
awk '{print $2}' c$comp1[$i] >$root.c
mv c$comp1[$i] swap$root
@ i = $i + 1
end

rm temp
rm temp2

foreach i (c*.curve)
set k=$i:r
awk '{print $2}' $i > $k.lc
end

foreach i (*.lc)
mv $i temp
set g=`cat temp`
echo $#g > n
set ff = $i:r
rm -f tmp
foreach h (*.c)
set root=$h:r
cp $h temp2
clear
echo
echo
echo
echo '                              CORRELATION SEARCH '
echo '                            FILE '$ff' TO FILE '$root
ifort correl.f
./a.out
set cor = `cat fort.9`
echo $cor' '$h >> tmp
end

sort -gr tmp > coo
awk '$1 !~ /Inf/ {print $0}' coo|awk '$1 !~ /inf/ {print $0}'|;
awk '$1 !~ /NaN/ {print $0}' >junk
mv junk coo
set r = `cat coo`
cp $r[2] zero
```

```
cp $r[4] zero2
echo $r[1] > ccc
rm -f tmp new coo
ifort correl2.f
./a.out
mv fort.9 $ff.l
mv temp $i
end

foreach i(*.l)
set v=$i:r
set mag=‘cat $i‘
set time=‘cat time‘
set n=$#time
set k=1
#
while ($k <= $n)
    echo $time[$k] $mag[$k] >> c$v.curve
@ k = $k + 1
end
end

rm *.lc
```

CORREL2.FOR

```
      real*8 mean, cor, l, d, mean2
      real*8 mag(10000)
      integer a, y, z, i, v
      real*8 b(10000)
      real*8 c(10000)
      real*8 x(10000)
      real*8 cc(10000)
      integer n


      open (1,file='n')
      read (1,*) a

      n = a

      open (2,file='temp')
```

```
read (2,*) (b(y), y=1,n)
open (3,file='zero')
read (3,*) (c(z), z=1,n)
open (4,file='ccc')
read (4,*) d
open (5,file='zero2')
read (5,*) (x(v), v=1,n)
cor = d

mean = 0.0
mean2= 0.0
l = 0.0
do i=1, n

mean = mean + c(i)
mean2 = mean2 + x(i)
l = l + 1
enddo

mean = mean/l
mean2 = mean2/l

do i=1, n
cc(i) = c(i) - mean + x(i) - mean2
cc(i) = cc(i)/2
end do

do i=1,n
mag(i) = b(i) - cc(i)
end do
write (9,*) (mag(i),i=1,n)
end
```

EEBLS.FOR

```
integer n, j, nb, nf, in1, in2, i, zzz
real*8 t(10000)
real*8 x(10000)
real*8 u(10000)
real*8 v(10000)
```

```
       real*8 p(100000)
       real*8 fmin, df, qmi, qma, bper, bpow, fff
       real*8  depth, qtran


       nb=400
       qmi=0.03
       qma=0.98
       df=0.001
       nf=5000

       open(1,file='n')
       read(1,*) n
       open(2,file='t')
       read(2,*) (t(i), i=1,n)
       open(3,file='x')
       read(3,*) (x(i), i=1,n)
       open(4,file='pspec')
       open(5,file='results')

       fmin = 1 / (t(n)-t(1)) +0.000001
       call eebls(n,t,x,u,v,nf,fmin,df,nb,qmi,qma,p,bper,bpow,depth,
     &qtran,in1,in2)
       write(4,*) (p(i),i=1,5000)
       write(5,*) bper, bpow, depth, qtran, in1, in2
       end
c
c
      subroutine eebls(n,t,x,u,v,nf,fmin,df,nb,qmi,qma,
     &p,bper,bpow,depth,qtran,in1,in2)


c
c--------------------------------------------------------------------------
c     >>>>>>>>>>>> This routine computes BLS spectrum <<<<<<<<<<<<<
c
c         [ see Kovacs, Zucker & Mazeh 2002, A&A, Vol. 391, 369 ]
c
c     This is the slightly modified version of the original BLS routine
c     by considering Edge Effect (EE) as suggested by
c     Peter R. McCullough [ pmcc@stsci.edu ].
c
c     This modification was motivated by considering the cases when
```

```
c     the low state (the transit event) happened to be devided between
c     the first and last bins. In these rare cases the original BLS
c     yields lower detection efficiency because of the lower number of
c     data points in the bin(s) covering the low state.
c
c     For further comments/tests see  www.konkoly.hu/staff/kovacs.html
c-----------------------------------------------------------------------
c
c     Input parameters:
c     ~~~~~~~~~~~~~~~~~
c
c     n    = number of data points
c     t    = array {t(i)}, containing the time values of the time series
c     x    = array {x(i)}, containing the data values of the time series
c     u    = temporal/work/dummy array, must be dimensioned in the
c            calling program in the same way as  {t(i)}
c     v    = the same as  {u(i)}
c     nf   = number of frequency points in which the spectrum is computed
c     fmin = minimum frequency (MUST be > 0)
c     df   = frequency step
c     nb   = number of bins in the folded time series at any test period
c     qmi  = minimum fractional transit length to be tested
c     qma  = maximum fractional transit length to be tested
c
c     Output parameters:
c     ~~~~~~~~~~~~~~~~~~
c
c     p    = array {p(i)}, containing the values of the BLS spectrum
c            at the i-th frequency value -- the frequency values are
c            computed as  f = fmin + (i-1)*df
c     bper = period at the highest peak in the frequency spectrum
c     bpow = value of {p(i)} at the highest peak
c     depth= depth of the transit at   *bper*
c     qtran= fractional transit length  [ T_transit/bper ]
c     in1  = bin index at the start of the transit [ 0 < in1 < nb+1 ]
c     in2  = bin index at the end   of the transit [ 0 < in2 < nb+1 ]
c
c
c     Remarks:
c     ~~~~~~~~
c
c     -- *fmin* MUST be greater than  *1/total time span*
c     -- *nb*   MUST be lower than  *nbmax*
```

```
c     -- Dimensions of arrays {y(i)} and {ibi(i)} MUST be greater than
c        or equal to  *nbmax*.
c     -- The lowest number of points allowed in a single bin is equal
c        to   MAX(minbin,qmi*N),  where   *qmi*  is the minimum transit
c        length/trial period,   *N*  is the total number of data points,
c        *minbin*  is the preset minimum number of the data points per
c        bin.
c
c=======================================================================
c
      implicit real*8 (a-h,o-z)
c
      real*8 t(100000),x(100000),u(100000),v(100000),p(100000)
      real*8 y(20000),ibi(20000)
c
      minbin = 5
      nbmax  = 2000
      if(nb.gt.nbmax) write(*,*) ' NB > NBMAX !!'
      if(nb.gt.nbmax) stop
      tot=t(n)-t(1)
      if(fmin.lt.1.0d0/tot) write(*,*) ' fmin < 1/T !!'
      if(fmin.lt.1.0d0/tot) stop
c-----------------------------------------------------------------------
c
      rn=dfloat(n)
      kmi=idint(qmi*dfloat(nb))
      if(kmi.lt.1) kmi=1
      kma=idint(qma*dfloat(nb))+1
      kkmi=idint(rn*qmi)
      if(kkmi.lt.minbin) kkmi=minbin
      bpow=0.0d0
c
c     The following variables are defined for the extension
c     of arrays  ibi()  and  y()  [ see below ]
c
      nb1   = nb+1
      nbkma = nb+kma
c
c================================
c     Set temporal time series
c================================
c
      s=0.0d0
```

```
      t1=t(1)
      do 103 i=1,n
      u(i)=t(i)-t1
      s=s+x(i)
 103  continue
      s=s/rn
      do 109 i=1,n
      v(i)=x(i)-s
 109  continue
c
c*****************************
c     Start period search     *
c*****************************
c
      do 100 jf=1,nf
      f0=fmin+df*dfloat(jf-1)
      p0=1.0d0/f0
c
c=======================================================
c     Compute folded time series with  *p0*  period
c=======================================================
c
      do 101 j=1,nb
        y(j) = 0.0d0
      ibi(j) = 0
 101  continue
c
      do 102 i=1,n
      ph     = u(i)*f0
      ph     = ph-idint(ph)
      j      = 1 + idint(nb*ph)
      ibi(j) = ibi(j) + 1
        y(j) =   y(j) + v(i)
 102  continue
c
c-----------------------------------------------
c     Extend the arrays  ibi()  and  y() beyond
c     nb   by  wrapping
c
      do 104 j=nb1,nbkma
      jnb    = j-nb
      ibi(j) = ibi(jnb)
        y(j) =   y(jnb)
```

```
 104  continue
c-----------------------------------------------
c
c===============================================
c     Compute BLS statistics for this period
c===============================================
c
      power=0.0d0
c
      do 1 i=1,nb
      s     = 0.0d0
      k     = 0
      kk    = 0
      nb2   = i+kma
      do 2 j=i,nb2
      k     = k+1
      kk    = kk+ibi(j)
      s     = s+y(j)
      if(k.lt.kmi) go to 2
      if(kk.lt.kkmi) go to 2
      rn1   = dfloat(kk)
      pow   = s*s/(rn1*(rn-rn1))
      if(pow.lt.power) go to 2
      power = pow
      jn1   = i
      jn2   = j
      rn3   = rn1
      s3    = s
 2    continue
 1    continue
c
      power= dsqrt(power)
      p(jf) = power
c
      if(power.lt.bpow) go to 100
      bpow  =  power
      in1   =  jn1
      in2   =  jn2
      qtran =  rn3/rn
      depth = -s3*rn/(rn3*(rn-rn3))
      bper  =  p0
c
 100  continue
```

```
c
c     Edge correction of transit end index
c
      if(in2.gt.nb) in2 = in2-nb
c
      !return
      end
c
c
```

Bibliography

1. http://archive.eso.org/dss/dss.

2. http://en.wikipedia.org/wiki/Extrasolar_planet.

3. http://exoplanet.eu/.

4. http://simbad.u-strasbg.fr/simbad/.

5. http://webviz.u-strasbg.fr/viz-bin/VizieR-2?-kw.cat=1239.

6. http://www.extrasolar-planets.com/.

7. http://www.extrasolar.net/.

8. *EXOPLANETS:Detection,Formation,Properties,Habitability*. Springer, Praxis Publishing, 2006.

9. *The Star and its planetary system in the wake of CoRoT Advances* (2008).

10. ASPC, Ed. *Extreme Solar Systems* (2007), vol. 398 of *Conference series*.

11. Borucki, W., and Summers, A. The photometric method of detecting other planetary systems. *Icarus* (April 1984), pp 121–134.

12. Budding, E., and Demircan, O. *Introduction to Astronomical Photometry*. Cambridge University Press, 2007.

13. Cester, B., Ferluga, S., and Boehm, C. The empirical mass-luminosity relation. *Astrophysics and Space Science* (November 1983), pp 125–140.

14. Charbonneau, D., Brown, T., Latham, D., and Mayor, M. Detection of planetary transits across a sun-like star. *The Astrophysical Journal* (2000), pp 44–48.

15. Claret, A. A new non-linear limb-darkening law for lte stellar atmosphere models.calculations for -5.0 ¡=log[m/h]¡=+1,2000k¡=teff¡=50000k at several surface gravities. *Astronomy and Astrophysics 363* (November 2000), pp.1081–1190.

16. Cumming, A., Marcy, G., and Butler, R. The lick planet search: Detectability and mass thresholds. *The Astrophysical Journal* (1999), pp 890–915.

17. Dvorak, R., Ed. *Extrasolar Planets:Formation Detection and Dynamics*. WILEY-VCH, 2006.

18. et al., P. *NUMERICAL RECIPIES IN FORTRAN 90: THE ART OF SCIENTIFIC COMPUTING*. Cambridge University Press, 1992.

19. Fischer, D., and Valenti, J. The planet-metallicity correlation. *The Astrophysical Journal* (April 2005), pp 1102–1117.

20. Gaudi, T. B. B. Predicting the yield of photometric of photometric surveys for transiting extrasolar planets. *The Astrophysical Journal* (November 2008), pp 1302–1330.

21. Gilliland, R., Brown, T., Guhathakurta, P., Sarajedini, A., Milone, E., Albrow, M., Baliber, N., Bruntt, H., Burrows, A., Charbonneau, D., Choi, P., Cochran, W., Edmonds, P., Frandsen, S., Howell, H., Lin, D., Marcy, G., Mayor, M., Naef, D., Sigurdsson, S., Stagg, C., and Vandenberg, D. A lack of planets in the 47 tucanae from a hubble space telescope search. *The Astrophysical Journal* (December 2000), pp 47–51.

22. Gimenez, A. Equations for the analysis of the lightcurves of extrasolar planetary transits. *Astronomy and Astrophysics 450* (January 2006), pp.1231–1237.

23. Harvey, P., Wilking, B., and Joy, M. On the far-infrared excess of vega. *Nature* (February 1984), pp 441–442.

24. Henry, G., Marcy, G., Butler, R., Paul, R., and Vogt, S. A transiting 51 peg-like planet. *The Astrophysical Journal* (2000), pp 41–44.

25. Howell, S. B. *Handbook of CCD Astronomy*. Cambridge University Press, 2006.

26. Jorissen, A., Mayor, M., and Udry, S. The distribution of exoplanet masses. *Astronomy and Astrophysics*, pp 992–998.

27. Kovacs, G., Bakos, G., and Robert, N. A trend filtering algorithm for wide–field variability surveys. *Monthly Notices of the Royal Astronomical Society 3* (January 2005), pp.557–567.

28. Kovacs, G., Zucker, S., and Mazeh, T. A box-fitting algorithm in the search of periodic transits. *Astronomy and Astrophysics* (April 2002), 369–377.

29. Mandel, K., and Agol, E. Analytic light curves for planetary transit searches. *The Astrophysical Journal 580*, 2 (December 2002), pp.171–175.

30. Mayor, M., and Queloz, D. A jupiter-mass companion to a solar-type star. *Nature* (November 1995), pp 335.

31. McCullough, P., Stys, J., Valenti, J., Fleming, S., Janes, K., and Heasley, J. The xo project:searching for transiting extrasolar planet candidates. *The Publications of the Astronomical Society of the Pacific* (2005).

32. McLean, I. *Electronic imaging in astronomy:detectors and instrumentation*. UTQL, 1997.

33. Mislis, D., Pyrzas, S., Harlaftis, E., Street, R., Horne, K., Pollacco, D., and Seiradakis, J. Wasp0: A wide field ccd search for extra solar planets. instrumentation, data analysis and preliminary results. *RECENT ADVANCES IN ASTRONOMY AND ASTROPHYSICS* (August 2006).

34. Nestoras, J., Mislis, D., Pyrzas, S., Tremou, E., Seiradakis, J., and Avgoloupis, S. Seeing measurements from mt. holomon. In *RECENT ADVANCES IN ASTRONOMY AND ASTROPHYSICS* (2006).

35. Percy, J. R. *Understanding Variable Stars*. Cambridge University Press, 2007.

36. Rauer, H., Voss, H., Erikson, A., Hatzes, A., Eisloffel, J., and Guenther, E. Recent results from the berlin exoplanet search telescope. In *Second Eddington Workshop:Stellar structure and habitable planet finding*.

37. Reed, B. Ubv beta database for cas-hamburg northern and southern luminous stars. *Astrophyscal Journal Supplement* (1998).

38. Rosenblatt, F. A two-clolor photometric method for detection of extra solar planetary systems. *Icarus*, pp 71.

39. Seager, A., and Mallen-Ornelas, G. A unique solution of planet and star parameters from an extrasolar planet transit light curve. *The Astrophysical Journal 585* (March 2004), pp.1038 –1055.

40. Seagroves, S., Harker, J., Laughlin, G., Lacy, J., and Castellano, T. Detection of intermediate-period transiting planets with a network of small telescopes. *The Publications of the Astronomical Society of the Pacific* (December 2003), pp 1355–1362.

41. Seiradakis, J., and Augoloupis, S. *Observational Astronomy*. AUTH, 1986.

42. Street, R. PhD thesis.

43. Vidal-Madjar, A., Lecavelier, A., Desert, J., Ballester, G., Hebrard, G., and Mayor, M. An extended upper atmosphere around the extrasolar planet hd209458b. *Nature* (March 2003), pp 143–146.

44. Wolszczan, A., and Frail, D. a planetary system around the millisecond pulsar psr1257+12. *Nature* (January 1992), pp 145–147.