UNIVERSITY of CALIFORNIA
SANTA CRUZ

CREATION OF SYNTHETIC RADIAL VELOCITIES TO FOLLOW UP TESS OBSERVATIONS

A thesis submitted in partial satisfaction of the
requirements for the degree of

BACHELOR OF SCIENCE

in

ASTROPHYSICS

by

Barbara Giusti

January 2012

The thesis of Barbara Giusti is approved by:

Professor Gregory Laughlin
Advisor

Professor Adriane Steinacker
Theses Coordinator

Professor Michael Dine
Chair, Department of Physics

Abstract

CREATION OF SYNTHETIC RADIAL VELOCITIES TO FOLLOW UP TESS OBSERVATIONS

by

Barbara Giusti

The use of synthetic radial velocity data sets and Monte Carlo simulations is a valuable exercise to analyze and predict the time frame necessary for a ground-based telescope to observe and verify the data collected by a space-based telescope. The proposed Transiting Exoplanet Survey Satellite Mission will collect and record transit photometry (and transit timing variations) of stars spanning F5 to M5 spectral type. Stars harboring candidate planets will then be examined, using Doppler velocimetry, by ground-based telescopes to verify and confirm the existence of the planetary systems. The minimal time interval and ideal location for the follow-up observations will be established according to a pre-specified 10% estimation error on the planetary masses. The synthetic data utilized for this task will be generated, analyzed, and then compared to the known input values using the Systemic Console software package. Although the Systemic Console is currently used to analyze existing RV data sets, it can also be employed to create comparable new ones. The script, built using a Java script code (BeanShell), assigns coordinates to the observed stellar targets (Right Ascension, RA and Declination, DEC) and to the available observing locations on the ground (Latitude, LAT and Longitude, LON). When the program is initialized, it registers the current time and runs the clock forward at a 15-minute increment. Constraints, such as chances of site availability, favorable weather, full moon, and altitude are an essential

component of the script. The code can also ingest orbital parameters derived from the experimental data and then generate new sets of synthetic radial velocities. Following this, the new data will be loaded into a Levenberg-Marquardt optimization algorithm and the output is a new, best-fit planetary system. The thus-produced *Original.fit* and *Minimized.fit* files are compared by looking at the corresponding planetary masses. If the planetary masses of the synthetic and observed systems are to within 10% of each other, then no more data will be collected. If the two examined planetary masses differ by more than 10%, then more RV data will be generated to decrease the error. The final result, which is a module within a larger project that is currently underway, will compare four variables: 1) a projected budget, 2) within this budget, the number of ground-based instruments from which we can observe, 3) the best location for observation, and 4) the minimum number of observation needed to corroborate the observational data received form TESS. It is our hope that this methodology will be employed to substantially increase the prospects of success for the TESS Mission.

To

my mother, Rosella, for letting me free to choose the life I wanted and

my husband, Kurt, for believing in me and boosting my confidence level

during the hardest times of my academic career.

Acknowledgements

I want to thank my technical advisor, Dr. Gregory Laughlin for his immense patience and expertise on the subject: he helped me set a higher standard in my career goals. Stefano Meschiari, the brain behind the Systemic Console, always available, has also helped me tremendously in understanding the logic and syntax of the code. Finally, I would like to thank my thesis advisor, Dr. Adriane Steinacker, for her scrupulous eye and constructive criticisms without which it would have been hard to write a coherent and organized thesis.

Contents

1

Introduction

The first Jovian extra-solar planet, 51 Pegasi b, was discovered in 1995 by two Swiss astronomers, Michel Mayor and Didier Queloz. In the intervening years, over 600 candidates have been considered to qualify as exoplanets (http://exoplanet.eu). We have discovered an average of seven exoplanets every month since 1995 and the number is increasing rapidly as technology advances. This has made it necessary to find ways to corroborate the planetary discoveries with the least possible amount of available data and within a definable budget. The use of synthetic radial velocity data sets, within the context of a lerge-scale Monte-Carlo simulation, is a valuable exercise in analyzing and predicting the time frame necessary for a ground-based telescope to observe and verify the data collected by an orbiting one. TESS (Transiting Exoplanet Survey Satellite) is a joint project between MIT, Caltech, UCSC, NASA Ames Research Center, and the Harvard-Smithsonian Center for Astrophysics. The projected mission operation time is two years. The project just recently passed NASA's phase-A study and it is projected to launch in late 2016 [Ricker (2011)]. The TESS mission will be an improved successor to the currently active KEPLER mission.

It will corroborate KEPLER's findings and use new and more sensitive data gathering techniques with the purpose of confirming more than 1,000 extra-solar planets. KEPLER was launched on March 6, 2009 [Borucki et al. (2009)] to quantify the frequency of Earth-size planets around Sun-like stars. About 1/3 of the 1200 transiting planet candidates detected in the first four months of the Kepler mission were found to be members of multiple candidate systems [Lissauer et al. (2011)]. The satellite was equipped with a 95 cm aperture telescope and 42 CCD cameras acquiring images with a total resolution of 95 million pixels [Borucki et al. (2009)]. As a comparison, the Hubble Space Telescope (HST) has an optical aperture that is about two and a half times greater than the aperture of the KEPLER telescope, but is only equipped with 8 CCD cameras. KEPLER is a better device to analyze large areas of the sky. KEPLER is currently surveying and monitoring about 150,000 stars. It's field of view is limited to a 105 square degree area of the Cygnus-Lyra region, just above the galactic plane. This area provides a rich field of stars in our vicinity.

Two systems are primarily used to identify the presence of a planetary system around a star. 1) The detection of periodic drops in brightness of a star (or transit photometry) which can indicate the passage of a potential planet in front of its parent object as shown in Figure 1.1. 2) The detection of the stars' Doppler radial velocities (RV), method accredited with 90% of planetary discoveries, as represented in Figure 1.2. The analysis of detected RVs is one of the most widely effective techniques in the study of orbital parameters such as planetary masses, orbital eccentricities, and orbital periods [Meschiari et al. (2010)].

While KEPLER was able to indicate the existence of more than 1,235 planetary system candidates, the lack of an effective verification system creates uncertainty and doubt within the scientific community. Ground-based telescope observations have revealed an

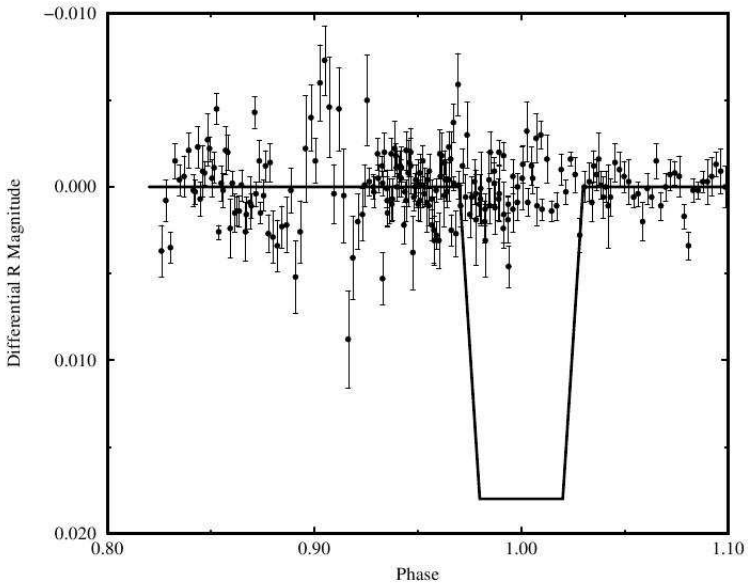


Figure 1.1: Photometric data belonging to the star HD187123 which shows a prediction of planetary transit. The curve shown is based on the presence of one planet believed to have a period of 3.097 days and a semi-amplitude of 72 m/s. [Robb R.M. et al. (1999)]

average of ten false positives to one possible detected planetary system to be a base-line expectation for purely photometric surveys [Torres et al. (2010)]. The exact ratio of false positives for a space-based telescope such as KEPLER has not yet been determined with accuracy, but we can analyze its data and understand its main pitfalls.

In particular, the Kepler mission is expected to yield transiting Earth-mass planets in the Habitable Zone (HZ) as part of its mission objectives, through continuous and simultaneous photometric sampling of more than 100,000 main-sequence stars. However, this class of objects will likely represent a small percentage of the detections (given the constraints of the mission design), and a much larger number of Neptune-mass and giant planets will be likely detected [Meschiari & Laughlin (2010)]. There are three major false positives common to KEPLER's analysis. 1) The presence of a binary star system. When

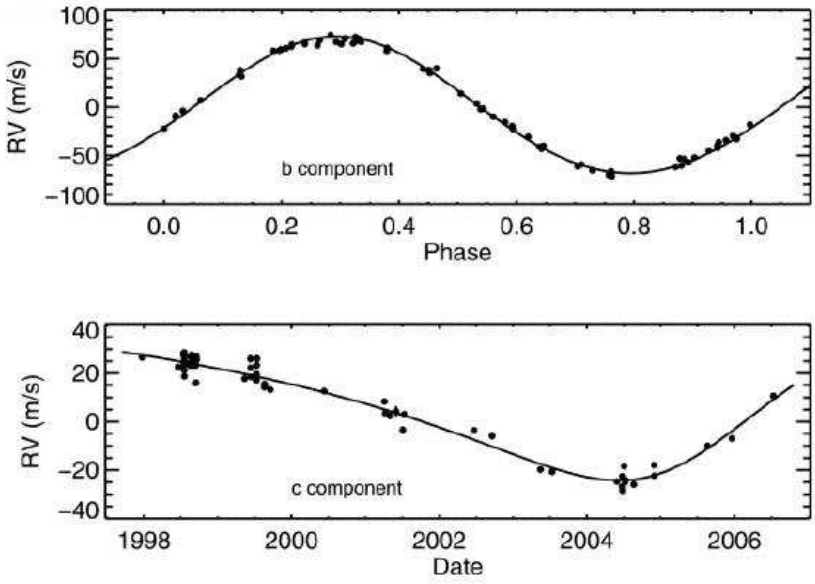


Figure 1.2: Radial velocity data for the star HD187123. This data was collected at the Keck Observatory and shows the inner planet (top) with a period of 3.1 days and a lower limit on the mass of 0.5 $M_{\mathbb{Q}}$. The bottom graph shows an outer planet with a period of 3810 days and a lower limit on the mass of 2 $M_{\mathbb{Q}}$ [Wright et al. (2007)].

two stars orbit each other they can mimic the photometric and radial velocity data collected for a planetary system they will give rise to a photometric detection similar to that of a planetary system. 2) Planets with a radius and mass similar to Earth are difficult to confirm using RV readings alone. As the stars magnitude increases and the amount of light gathered decreases, the detection of RVs becomes even more difficult if not impossible. A planet with similar terrestrial characteristics is unlikely to cause a detectable change in motion of the parent star. 3) The geometry of the KEPLER telescope aperture is such that the field of view can include more than one star. This effect increases the possibility of blending a collection of photometric data gathered from more than one subject [Torres et al. (2010)].

This phenomenon occurs when the photometric data collected from a specific candidate is contaminated by the presence in the foreground or background of a binary star and/or a combination of star-planet systems. For instance, in the case of a binary star situated in the background of the host star of a candidate planet, the signature light of the target reduces the depth of the binary system eclipsing event, resulting in a transit-like signal. In another case, the candidate's light is diluted by the presence of other stars in the telescope's aperture. Even at high resolution, there have been images counting up to four stars within fifteen arc seconds of the target [Torres et al. (2010)]. These combinations of star and planets can cause "blending", and they must be systematically ruled out before reaching any conclusion. This means that while the ground-breaking KEPLER mission collected thousands volumes of photometric data, it is ultimately not able to differentiate between potential planetary systems and possible anomalies. To minimize follow-through on false positives, and to attain a more effective use of research fundings, we need the right tools to make more accurate guesses on the existence of planetary systems around observed stars.

TESS will examine over 2 million stars with an efficient all-sky coverage 400 times greater than KEPLER's (wide), and with magnitudes between 4.5 and 12 (shallow). The project is thus termed an all-sky, wide-shallow survey [Ricker et al. (2009)]. The TESS budget is going to have a greater consideration for follow up using ground-based observations and its parameters are going to be established according to trials and predictions using synthetic data matched to actual data collected by the space telescope. Thus, the ground-based follow-up will independently corroborate the data collected by the orbiting telescope and use it to establish the presence of extra-solar planets.

For this purpose, we use a computer algorithm that creates thousands of simulated

sets of RVs that will subsequently be matched with the existing sets. The final achievement is to minimize the amount of synthetic values the computer program creates and see how they fit the model with a minimal standard deviation and, through this method, the ground-based observing time and location for optimal detection can be set within 10% of error. The program was built using a Java script code (Beanshell), that serves as an API to a complex software package, the Systemic Console, which will be examined more in detail in Chapter 4. The Systemic Console is currently used with the purpose of analyzing Doppler radial velocities and transit timing observations [Meschiari et al. (2009)] and fitting the data to possible orbiting planets.

The program takes into account various weather and astronomical constraints and creates multiple sets of RV for up to one thousand stars per data set. The synthetic data will then be matched to the actual data in order to predict and verify planetary motion. The final result will compare four variables: 1) a projected budget, 2) within this budget, the number of ground-based instruments that could be built, 3) optimal observing time frames, 4) matched expectation values. The following research thesis will be dedicated to the explanation and description of this method. Chapter 2, will explain what methods can be used to collect information about exoplanetary systems, and how RV detections occur and are represented by mathematical expressions. Chapter 3 will discuss the Keplerian method of integration, uncertainties, and false positives. Chapter 4 will show how the Systemic Console works, how the files are divided, and how it can be used to obtain orbital elements' `best fit`. Chapter 5 explains and shows how the code built in the Systemic Console was manipulated to aid the TESS project. Finally, Chapter 6 presents the current results and discusses the next phases of this ongoing research.

2

Theoretical Understanding

The core accretion model is currently the most widely accepted theory of planet formation. The accretion starts partially from a pressure-supported sub-Keplerian disk containing gas and dust. The dust particles collide, grow, and form rock-ice planetary cores. When a core becomes massive enough with gas still present in the disk, it gravitationally gathers gas and rapidly increases in mass. Through hydrodynamical, n-body simulations, and appropriately parametrized damping, the phenomenon described above can be explained by noting that short period planets form farther out and migrate inwards to their final semi-major axis [Wright et al. (2009)]. Planetary formation has been at the heart of astronomical research for the past 250 years, and it has been recognized that there is an important constraint to apply when considering planetary formation: *co-planarity* [Lissauer et al. (2011)]. Our current understanding leads us to believe that growth within an accreting disk generally yields circular orbits and low relative inclinations. The Solar System is an example of such a configuration: eccentricity and inclination reach their maximum with Mercury that holds $e_{☿} = 0.21$ and Earth's inclination (if Pluto is not considered)

$i_{\oplus} = 7.2°$ with respect to the Sun's equator.

2.1 Planetary inclinations

The inclination distribution of exoplanets is a fundamental aspect of planetary system dynamics. Yet neither transit observations, nor any other technique, have directly measured the true mutual inclination between planets observed in multiple systems, other than in a few in unusually fortuitous circumstances. Indirect constraints, however, can be obtained for stable systems with multiple transiting planets, considered the best examples of mutual inclinations around main sequence stars. For example, the lack of transit timing variations (TTV) can be used to constrain mutual inclinations. Occasionally, it will be possible to measure mutual inclinations from exoplanet mutual events such as measurements of the Rossiter-McLaughlin effect of planets in the same system [Lissauer et al. (2011)]. The Rossiter-McLaughlin effect determines the sky-projected angle between the stellar rotation axis and the planetary orbital axis. Furthermore, two or more planets initially on coplanar circular orbits (e.g. similar inclination) with sufficient orbital separation can never develop crossing orbits. These exoplanets are called "Hill Stable" [Lissauer et al. (2011)].

$$\Delta_{orb} = \frac{a_o - a_i}{R_H} > 2\sqrt{3} \approx 3.46, \tag{2.1}$$

where a_i is the semi-major axis of the i^{th} planet and R_H is the mutual Hill sphere radius, the distance within which a planet feels the gravitational force of another body.

2.2 Detection of exoplanets

The stellar mass, $M_\star$, and a exhaustive sample of its radial velocities can help determine the presence of a planetary system. However, the strength of the radial velocity signal usually diminishes with increasing periodicities. This phenomenon is due to the smaller gravitational force that a distant planet exerts on its star. Furthermore, a planet with orbital eccentricity, e, and longitude of periastron (the orbital angle of a planet when it is the closest to its host star), ω, produces transits visible from the Earth with a probability given by [Charbonneau et al. (2007)]

$$P_{tr} = 0.0045 \left(\frac{1AU}{a} \right) \left(\frac{R_\star + R_p l}{R_\odot} \right) \left[\frac{1 + e \cos\left(\frac{\pi}{2} - \omega\right)}{1 - e^2} \right]. \tag{2.2}$$

By noting that P_{tr} is inversely proportional to the semi-major axis a we can deduce that the greater the planet's orbital distance from its host star, the harder it will be to detect it.

If the star is bright enough, a reasonable estimate of its radius, $R_\star$, can be calculated and the planetary radius, R_{pl}, can be obtained by measuring the fraction of the parent star light that is occulted. By definition, transiting planets have their orbits oriented so that the Earth lies in, or close to, their orbital plane, but this is not a very common circumstance and we therefore have to rely on other techniques to characterize most of the planetary systems. In the case of a single planet orbiting its host star, (multiple planet systems are harder to identify and describe), if the planetary transit is observed and we can accurately measure the orbital inclination, i, we can evaluate the planetary mass M_{pl} directly from the limiting mass $M_{pl} \sin i$. If the orbital inclination is not known, by assuming that the stellar rotation is aligned with the orbital plane, we can derive $\sin i$ by combining the observed radial velocity (RV), $v \sin i$, with the equatorial velocity, or angular motion of

the star about its axis V_{equ} [Mayor & Queloz (2005)]. The faster the star rotates on its own axis, the harder it is to disentangle the two sets of velocities belonging to the system.

$$V_{equ} = 2\pi R/P(\upsilon \sin i), \tag{2.3}$$

where $\upsilon \sin i = V_{equ} \sin i$.

The first system known to include multiple planets was a triple one, ν_{And}, discovered by Butler et al. in 1999. In 2007, at least 1/4 of known planetary systems showed evidence of multiple companions [Wright et al. (2007)]. By March 2009, 14% of known stars hosting planets within 200 pc ($\approx$ 653 ly = 6.17×10^{15} km) were known to have multiple-planet systems, and another 14% showed significant evidence of multiplicity observed through the long-term radial velocity data monitoring [Wright et al. (2009)]. Needless to say, multiple planet systems that have been observed and described only in recent years, and yet very common in the observable universe, require much more complex algorithms to disentangle system parameters from the radial velocity signature. Planet-planet interactions can significantly alter the system dynamics and, often, should not be underestimated. These dynamics can cause detectable variations in the orbits of the planets and a set of constant Keplerian orbital elements is no longer sufficient to model the observed radial velocities. The resulting system models are non-linear and cannot be described analytically (through closed-form functions). A proper fit must be derived with the use of a *n-body* code (a *Newtonian* or a *dynamical* fit). Characterizing such arrangements in the shortest possible time, and the efficient search of the χ^2 space has developed into a real art. The final goal is to find a global minimum. Long-term stability is another constraint that must be taken into account when considering multi-planet fits and elements such as *inclination* (see Section 2.1), and eccentricities can seriously constraint the development of a stable

planetary system. Fortunately, in many cases, planet-planet interactions are sufficiently weak that they can be ignored, and the resulting radial velocity signal is simply the linear superposition of multiple Keplerian radial velocity curves (a *Keplerian* or a *kinematic* fit). Figure 2.1 helps summarize the characteristics of planetary systems discovered before and

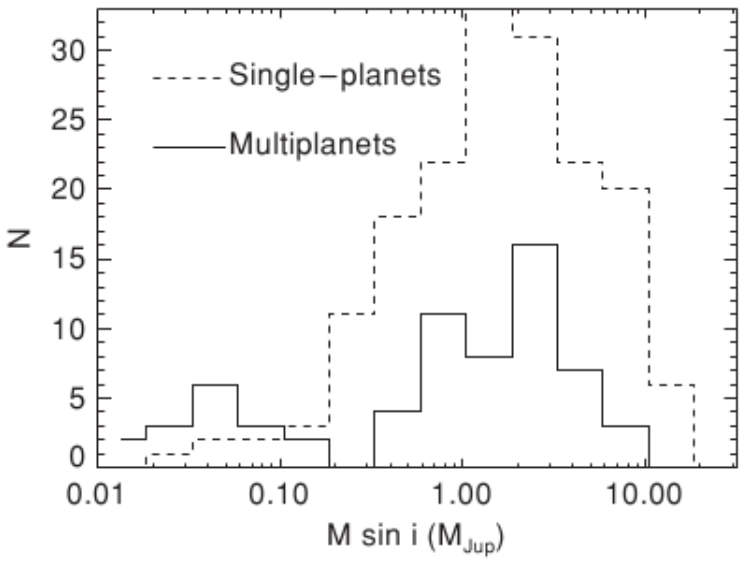


Figure 2.1: Distribution of exoplanets according to $M \sin i$ with an upper limit of 13 $M_{\mathcal{Q}}$ for known multiple-planet systems (solid) and apparently single systems (dashed) [Wright et al. (2009)].

during the KEPLER mission. 1) Multi-planet systems exhibit an apparent overabundance of planets with $0.01\ M_{\mathcal{Q}} \leq M_{pl} \sin i \leq 0.2\ M_{\mathcal{Q}}$ (as a point of reference $M_{\oplus} = 0.0315\ M_{\mathcal{Q}}$). 2) The number of known systems hosting planets with $M_{pl} < 1\ M_{\mathcal{Q}}$ is much less than that of systems with $M_{pl} > 1\ M_{\mathcal{Q}}$. 3) It is easier to find a small planet in a multiple-planet system, than it is to find it in a single-planet system. 4) Interactions between giant planets are expected to excite eccentricities, whereas Figure 2.1 exhibits multi-planet systems with lower orbital eccentricities. This last effect may be partially explained by the orbital stability constraint which favors low eccentricity orbits. When a multi-planet system is

characterized by high eccentricities, it makes the system more unstable and prone, through time, to planetary collisions and expulsion of planets from the system. Thus, some single-planet systems may exhibit high eccentricities as a result of ejections of former members from the system. However, it is important to remember that this finding may be amplified by a selection effect and small planet discoveries are still limited by instrumentation. A greater population of Earth size planets might be found in the next future as technology rapidly advances.

2.3 Methods implemented to discover exoplanets

Some of the most successful methods that have been developed and implemented for the detection of extra-solar planets are listed below.

i) **Transit method** (implemented with Spitzer Space Telescope (2005), Corot(2006) Kepler(2009), and (assuming successfull launch) TESS(2016))

As shown in Figure 1.1, the observed brightness of the star decreases by a small amount when a planet transits in front of it and in or about our line of sight. This photometric method is particularly effective when employed by orbiting telescopes, and it can help determine the radius of a planet. The light curves derived from space-based observations such as KEPLER and the future TESS mission, can be used to measure the radii and orbital periods of planets to high accuracy. The ratios of planetary radii to stellar radii are also well measured, except for low signal to noise transits (e.g., planets much smaller than their stars, those systems for which few transits are observed, and/or planets orbiting faint or variable stars). This technique, combined with the radial velocity method as a follow up, can give insights even on

planetary atmospheric compositions. When a transit occurs, the stellar photons pass through the upper atmosphere of the planet and, by studying the high-resolution stellar spectrum carefully, one can detect major constituents of the planetary atmosphere. Additionally, the secondary eclipse (which occurs when the star eclipses the planet) allows for direct measurement of planetary radiation (light from the planet can be detected through thermal emission), giving estimates about the planet's temperature. The transit method has three major disadvantages. First, if the planetary orbit is not within the Earth's line of sight, the transit cannot be observed and the probability of detection is given by the previously presented Equation 2.2. Roughly 10% of planets with orbits of order $3-10$ days have such alignment, and the fraction decreases for planets with larger orbits. A solution to this problem will be given by TESS, programmed to scan the observable sky and to find extra-solar planets at a rate that exceeds that of the radial velocity method. The second issue is related to planetary sizes. For a given planet transiting its host star, the likelihood of detection is a decreasing function of the orbital period, $t_{transit} \propto 1/P^{2/3}$ [Lissauer et al. (2011)]. Third, the high rate of false positives can exactly mimic planetary systems. This last point will be further discussed in Section 2.5. Finally, it is worth emphasize that only for transiting planets direct estimates of planetary masses from their observed radii can be obtained [Charbonneau et al. (2007)].

ii) **Transit timing variation method (TTV)**

If a planet was previously detected by the transit method, then variations in the periodicity and duration of the transit can provide an additional tool for the detection of a multiple-planet system. This method can be used to infer the orbital elements of

the perturbing planet, or, at least place limits on the presence of additional planets. TTVs are caused by gravitational perturbations exerted by additional planets, which cause deviations from Keplerian orbits with predictable periodicity. The successful analysis of TTVs often occurs in cases where the planets are in mean-motion resonance as in the case of Jupiter's moons as described later in Section 3.1. These dynamics can be particularly difficult to identify and characterize solely from radial velocity curves because the `best fit` orbital parameters exhibit degeneracies (multiple solutions).

iii) **Gravitational microlensing**

The luminosity of a star is influenced by the relativistic effect of the planets mass. When two stellar bodies are almost exactly aligned, the gravitational field of a foreground star magnifies the light of a background star and, if the former comprises a planetary system, then that system's own gravitational field can make a detectable contribution to the lensing effect. Gravitational microlensing presents two major disadvantages. First, lensing events are very brief, and second, an ideal alignment is highly improbable as the system composed of two observed stars and the Earth are constantly moving with respect to each other. This method is mainly used to monitor stars towards the center of the galaxy, as this region of the sky provides the largest number of background stars.

iv) **Astrometry (HST)**

This technique uses precise measurements of a star's position in the sky to observe how the position changes over time due to the gravitational influence of an orbiting planetary system. Astrometry dates back to the mathematician William Herschel. In the late 18th century, he claimed that an unseen companion was affecting the position

of a star he had cataloged as 70 Ophiuchi, located in the Ophiuchus constellation. Precisely, he claimed the discovery of the binary system in August 1779, after 1700 observations [Herschel & Dreyer (1912)]. In the 19th Century, this technique was still implemented visually with hand-written records, but the method has evolved since then through the use of photographic plates. In 2002, the Hubble Space Telescope succeeded in using astrometry to characterize a previously discovered planet around the star Gliese 876 [1]. One advantage of the astrometric method is its sensitivity to planets with large orbits. However, long periods require long observation intervals, ranging form years to decades.

v) **Eclipsing binary minima timing**

When an eclipsing binary (a 2-star system) is in our line of sight (LOS), the times of minimum light (occurring when the brightest star is obscured by its companion), constitute a time stamp on the system, much like the pulses from a pulsar (characterized as a dip in brightness). If a planetary system is present and set in a circumbinary orbit around both stars, it will be offset around a binary star-planet center of mass. The periodicity of this offset is indicative of the presence of a planetary system.

vi) **Polarimetry**

The light emitted by a star is unpolarized, however, when photons hit the atmospheric molecules of a planet, the light waves can become polarized. Polarimeters can be set to a very high sensitivity and can sift through unpolarized light coming from seemingly single stars.

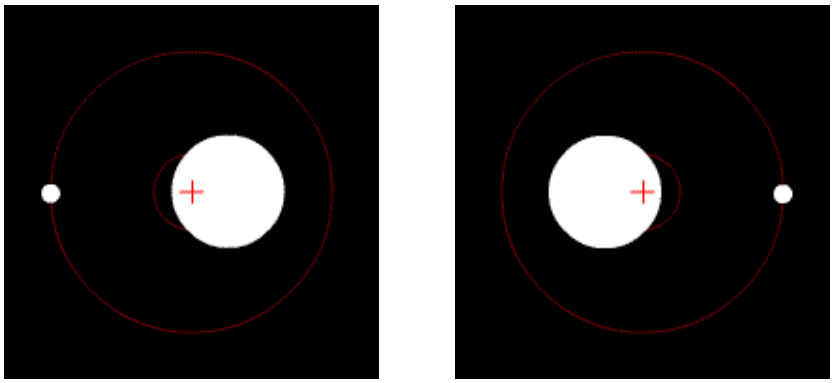
vii) **Doppler Spectroscopy**

[1] The actual measurements contained serious errors.

This method is the most effective at claiming planetary discoveries, and it is responsible for the characterization of more than 90% of all currently known planets [Meschiari et al. (2010)]. Because it is also very important to this thesis, it is discussed in considerable detail in the following section.

2.4 Doppler Spectroscopy: an efficient method to confirm the presence of exoplanets

The Sun orbits the center of our galaxy at a speed of approximately 250 km/s. While most of the stars in the solar neighborhood are moving in roughly the same manner, some are moving more quickly than the Sun, while others are moving more slowly. The Alpha Centauri system, for example, is headed toward us with a radial velocity of $V_r = -21.6$ km/s. The average difference in orbital velocity between neighboring stars is about 20 km/s. Part of this velocity will be in the transverse direction, while the rest is along the radial line connecting our solar system to the star. A radial velocity (RV) measurement is the component of the velocity of the star measured along the LOS from the Earth to the star where most of this radial velocity stems from the natural motion of the star (e.g. angular rotation and translation in space with respect to our solar system). In addition to the random motion that a given star has with respect to the Solar system, there is also a small superimposed component of motion that is generated as the star wobbles back and forth in response to any planets that are in orbit around it (see Figure 2.2). By carefully noting the Doppler shift of the stellar lines, it is sometimes possible to measure the speed of the star in our LOS with a precision down to 1 m/s. For the case of a single planet in



(a) **(b)**

Figure 2.2: The star moves about its center of mass to compensate for the planetary orbit. If the Earth was located West of the system (left of the page), a positive RV would be collected in (a) and a negative RV would be observed in (b).

a circular orbit, the system can be visualized by imagining that the star and the planet are attached to the opposite ends of a rigid rod. The point of balance can be reached by choosing the correct coordinate of the point much closer to the heavy star than to the less massive planet. For example, if the star is ten times more massive than the planet, then the point of balance would lie 10 times closer to the star than it would to the planet. Thus, the star delineates a circle that is 1/10 times smaller, and which executes in the same amount of time that the planet takes to complete one circular revolution. If the RVs collected vary with time, corresponding to movement of the star along our LOS, it could mean that one or more planets orbit the target. As of 2011, spectrometers can detect velocity variations down to less than 1 m/s. The HARPS (High Accuracy Radial Velocity Planet Searcher) instrument, connected to a 3.6 m telescope located at La Silla Observatory (Chile), and HIRES (High Resolution Echelle Spectrometer), another spectrometer which is a vital component of the Keck telescopes complex in Hawaii, are examples of excellent

instrumentation devoted to such detections. This method requires high signal-to-noise ratios in order to achieve high precision, and is primarily used for stars up to 160 light-years from Earth [Charbonneau et al. (2007)]. Doppler spectroscopy is readily suited to the detection of massive planets close to their hosts, but the detection of such bodies orbiting further out requires many months, if not years of observation. Planets with orbits highly inclined to our LOS are more difficult to detect because the change in velocity of the incoming star's photons is minimal.

Back in the 1950s, the Russian-American astronomer, Otto Struve, foresaw the future use of powerful spectrographs to detect distant planets. He predicted that small Doppler shifts of light emitted by a star, due to the constant and continuous change of its radial velocity, could be detected by sensitive instrumentation. These motions would be identified on a spectrographic image as small shifts towards longer wavelengths (red color) or shorter wavelengths (blue color) of the star's emitted photons. During his lifetime, technological limitations kept the detection of radial velocities in the 1000 m/s range or greater. This precision was practically useless for the detection of extra-solar planets [http://en.wikipedia.org/wiki/Otto_Struve].

2.4.1 How Doppler shift is measured

Until the beginning of the 1990s, RV measurement techniques rarely exceeded a precision of $200 - 500$ m/s. The wavelength reference (WR) was taken at a different time from observation and often transversed a different light path than that of the stellar spectrum. One solution proposed was to use RVs collected from standard stars. This system avoided the problem of different light paths for the reference and stellar spectra, but the

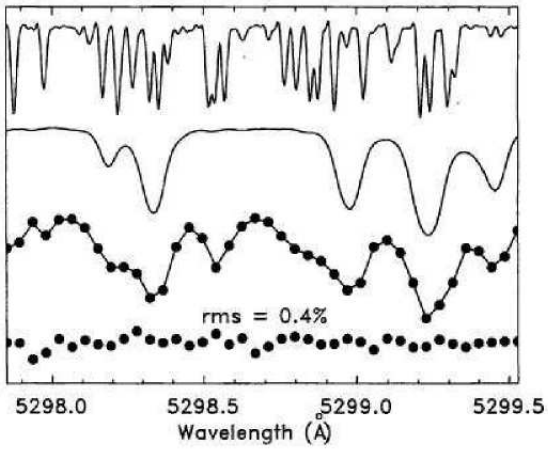


Figure 2.3: First (top): the template iodine cell spectrum. Second: the template of a stellar spectrum. Third: the points represent observation of the target through the I_2 absorption while the solid line is a model of the observation. Fourth (bottom):10 times the difference between the model and the observation corresponding to a 0.4% rms between model and observation [Butler et al. (1996)].

sample observation was still made at a different time and there was always the danger that the standard star was subject to radial velocity jitter, arising either from stellar variability, a binary companion, or other confounding factors. However, a new solution first considered in 1973 by Griffin & Griffin at the McDonald Observatory and then successfully implemented with high precision in 1994 by Steve Vogt at the Keck HIRES Spectrometer. The new technique consisted of minimizing the instrumental errors by superimposing a WR on top of the stellar observation. This task is accomplished by passing the incoming light through an absorbing gas prior to its entrance into the spectrograph. The gas produces its own set of absorption lines against which velocity shifts of the stellar spectrum can be measured. Instrumental shifts affect both the WR and stellar spectrum equally and a high degree of precision can be achieved. The precision of this method is limited by pressure and temperature changes of the earth's atmosphere as well as by Doppler shifts of the element's

lines due to winds. These errors can be eliminated if the observer has some control over the absorbing gas. A main improvement has been applied by enclosing a gas in a cell that can be temperature and pressure regulated. The use of such a method was first applied to astronomical research by Campbell and Walker in 1979. They chose hydrogen fluoride (HF) as the absorbing gas. They demonstrated, through more than a decade of use, that the HF cell could measure relative radial velocities with a long term precision of about 13 m/s. Although the HF cell proved capable of achieving the precision needed for detecting Jovian-sized planets around Solar-type stars, there were a number of disadvantages in using such a device [Kürster et al. (1994)].

1) The path length of the cell that is required to produce reasonably deep HF absorption lines is about 1 m and it can be a problem if space in front of the spectrograph slit is limited.

2) HF has significant pressure shifts and must be regulated to a rather high temperature of 100℃.

3) It is a highly reactive chemical, and prolonged exposure of the cell to HF will destroy the container walls. Consequently, the absorption cell must be made of inert material and has to be refilled before each observing run.

4) At last, HF gas is fatal to humans who have been exposed to it.

Gaseous molecular iodine, I_2, is a benign alternative to HF and the advantages of using this substance in an absorption cell are numerous.

1) The vapor pressure is high enough (≈ 0.5 torr) to produce significant absorption at room temperature in a cell $10 - 20$ cm in length, instead of the 1 m long HF cell.

2) I_2's pressure shifts are much smaller than HF's and this results in a very stable reference spectrum.

3) Iodine gas does not react with glass so that the construction of the cell is relatively easy and can be done by any glassblower.

4) A fixed number of I_2 molecules are permanently sealed in the cell for its entire lifetime, so there is no need to refill it prior to each observation.

5) Its spectral region is rich of extremely narrow lines starting at about 4800 Å and ends near 6000 Å.

An iodine absorption cell consists of a sealed glass bottle with a small amount of I_2 crystals that is placed into position directly in front of the slit. When heated to a temperature above 35℃, the iodine sublimes and the gas produces an absorption spectrum on the beam from the telescope as it enters the spectrometer. The absorption spectrum yields a very stable zero-velocity reference spectrum superimposed on the spectrum of the object being observed as shown in Figure 2.3. The Iodine cell is sealed and temperature controlled to 50 ± 0.1 ℃ such that the column density of Iodine remains constant. A block of the spectrum containing the Iodine region is divided into $\approx$ 700 slices of 2Å each [Meschiari et al. (2010)]. The final measured velocity is the weighted mean of the velocities of the individual slices. All radial velocities have been corrected to the Solar system barycenter, but are not related to any absolute radial velocity system and are therefore "relative" radial velocities. Although this system has been able to detect RVs with a precision of 1 m/s, through measurements obtained at Keck (Hawaii), AAT (Australia), and La Silla (Chile) telescopes, systematic errors such as detector imperfections, optical aberrations, effects of under-sampling the

Iodine lines, photon statistics, and stellar jitter must be taken into consideration.

2.4.2 Mathematical expression of radial velocity data

The Systemic Console allows for a choice of various modeling schemes. In particular, if the N planets do not experience significant dynamical interactions during the time in which a set of RVs are collected, the RV variation of the star can be represented by a sum of N Keplerian orbits, each described by its planets' orbital elements [Meschiari et al. (2009)]. The RV Curve will be given by

$$v_{RV}(t) = \sum_{N} K_N[\cos(\theta_N + \omega_N) + e_N \cos\omega_N] = \sum_{N} v_{RV_N}(t) + C_{Tel}, \tag{2.4}$$

where the one singly collected RV for a star that includes a perturbing planet on a Keplerian orbit is given by

$$v_{RV_N}(t) = K_N[\cos(\theta_N(t) + \omega_N) + e\cos\omega_N]. \tag{2.5}$$

The orbital parameters presented in Equation 2.4 and 2.5 are K, the RV semi-amplitude, θ, the true anomaly (the angular parameter formed by the periapsis, the star's and the planet's positions, or also the angular parameter that defines the position of a body moving along a Keplerian orbit), ω, the longitude of periastron, and e, the eccentricity. The presence of an unknown constant velocity offset parameter, C_{Tel}, in Equation 2.4, is due to the fact that RV observations can be sampled from telescopes at different locations and different times. However, even though no information about planetary existence can be provided by the latter parameter [Ford (2005)], the Systemic Console allows sources with different zero point velocity offsets to be combined and facilitate a better fitting. Other essential orbital parameters that need to be assessed are the period, P, the mass, m, and the mean anomaly,

MA. When the modeling RV curve equation is in place, the orbital parameters of the i^{th} planet can be derived.

2.5 False positives

The detection of exoplanets through photometric transits, as explained in Section 2.3, suffers from a very high rate of biases and false positives. The two most notable observational biases are 1) that the probability of transiting decreases with increasing orbital period and 2) for a given size planet transiting a given star, the duty cycle (the fraction of the orbit spent in transit) is a decreasing function of orbital period, because the fraction of time that a planet spends in transit is proportional to $P^{2/3}$ [Lissauer et al. (2011)]. Other factors which have which have strong biasing effect on the detectability of transiting planets, or alternately, which can generate false positives are:

i) **Eclipses of planets with smaller orbits**

Photometric transit searches are strongly biased in favor of planets with small orbits. Such objects have a greater probability of going through periodic eclipses.

ii) **Transits of planets with smaller orbits**

Larger orbits imply longer orbital periods and fewer chances for transits to occur and therefore small orbits are more readily detected by transit surveys. However, multi-site surveys, such as the second phase of the TESS project, can monitor a given field for several months and have more chances to frequently achieve a positive detection nearing 100% for periods up to 6 days.

iii) **Stroboscopic effect**

A stroboscopic effect can afflict single-site surveys, favoring orbital periods near integer numbers of days and may account for the tendency of the longer period transiting planet periods to clump near 3 and 3.5 days [Charbonneau et al. (2007)].

iv) **False signals**

Most transit surveys, including the current KEPLER mission, are characterized by a high rate of candidate systems that display light curves precisely mimicking the sought-after data. Some of this collected data results from setting an overly-permissive detection threshold. In these cases, the algorithm used will select even events that result purely from photometric noise [Charbonneau et al. (2007)].

v) **Blending**

If the field of view is crowded, erroneous photometry can result when light signals from nearby stars leaks in the photometric aperture.

vi) **Eclipsing multiple star systems**

The overwhelmingly majority of false positives are expected to come from eclipsing binary or multiple star systems. These eclipses are generated when fainter stars are in the same system of the target or stars that are the faint background of foreground objects within the camera aperture.

vii) **Grazing eclipses**

Grazing eclipses in binary systems (two star's disks graze each other without ever overlapping) can result in transit-like signals with depths and durations that resemble planetary ones.

3

Keplerian Integration and Uncertainty Estimates

3.1 Keplerian method of integration

The gravitational force that the star exerts on the planets is not the only effect to consider. The planets also interact with one another, providing perturbations. In particular, if the planets repeatedly encounter one another at nearly the same point in their orbits, their mutual gravitational pull will tend to add up over time. A classical example of gravitational resonance occurs between Jupiter's moons Io, Europa, and Ganymede. Io orbits Jupiter twice as frequently as Europa, and Europa orbits Jupiter twice as frequently as Ganymede. This 4:2:1 orbital ratio is maintained by the mutual gravitational tugs exerted by the moons. If we assume, however, that the planetary motion consists of independent elliptical orbits, the model RV curves are computed very quickly by solving Kepler's Equation. The Keplerian parameters, responsible for the planet's ID, can be extrapolated from

the RV data. There are $5N+1$ parameters that need to be fit: 1) period, P_i, 2) eccentricity, e_i, 3) mean anomaly, M_i, 4) inclination of the planetary orbit, i_i, 5) semi-amplitude of the RV signal, K_i, and the stellar mass, $m_\star$. The mass of a planet, m_i, can be determined from the RV semi-amplitude the following way:

$$K_i^3 = \frac{2\pi G}{P_i(1-e_i^2)^{3/2}} \frac{m_i^3 \sin^3 i_i}{(m_\star + m_i)^2}, \tag{3.1}$$

[Wright & Howard (2009)] which becomes

$$K_i = \left(\frac{2\pi G}{P_i}\right)^{1/3} \frac{m_i \sin i_i}{(m_\star + m_i)^{2/3}} \frac{1}{\sqrt{1-e_i^2}}. \tag{3.2}$$

The planet's mean anomaly, M_i, is a Keplerian parameter that needs to be known before implementing Equation 3.2. The planet's true anomaly, $\theta_i(t)$, given by Equation 3.3 and present in Equations 2.4 and 2.5, is related to the planet's eccentric anomaly $E_i(t)$ as a function of time.

$$\tan\left[\frac{\theta_i}{2}\right] = \sqrt{\frac{1+e_i}{1-e_i}} \tan\left[\frac{E_i}{2}\right]. \tag{3.3}$$

In turn, the eccentric anomaly is related to the planet's mean anomaly M_i through Kepler's equation [Wright & Howard (2009)]

$$M_i(t) = E_i(t) - e_i \sin E_i(t) = 2\pi(t - t_{p,i})P_i, \tag{3.4}$$

where t is the time of RV detection, t_p is the time of periastron passage of the planet i, and the other parameters present have been specified in previous equations. The true and eccentric anomalies can be visualized and better understood by looking at Figure 3.1.

When revisiting Equations 3.2, 3.3, and 3.4 there are three parameters left that cannot be derived analytically, but need to be fitted at best with many trials: the period, P_i, the eccentricity, e_i, and the time of periastron passage, $t_{p,i}$. The use of an algorithm

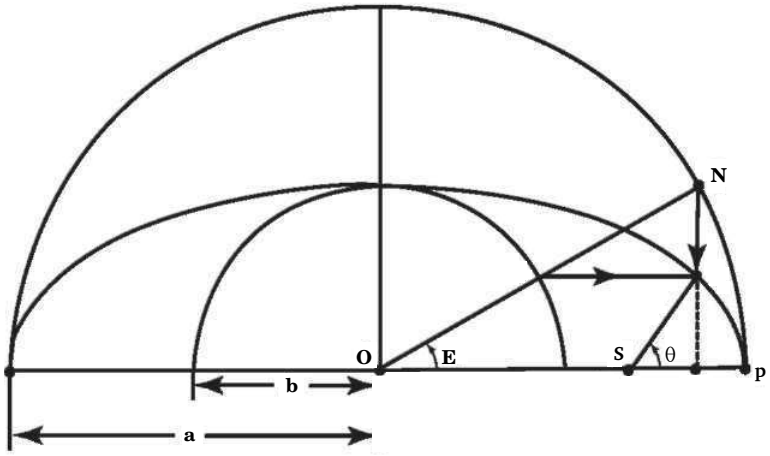


Figure 3.1: The eccentric anomaly is identified by **E** and the true anomaly by θ. **S** symbolizes the position of the star, while **N** is the planet, and **p** is the location of periapsis

such as the Levenberg-Marquardt or the Markov Chain Monte Carlo fitting routines, both available in the Systemic Console and later discussed in Section 3.2, can be used for this purpose.

3.2 Uncertainties

In the quest for ever-smaller planets, proper error analysis is key. This happens frequently at the threshold of detection, where signal-to-noise ratios are low. Indeed, once the `best fit` parameters have been identified, the estimation of their uncertainties is of equal or greater importance to the measurement of the orbital parameters themselves as evidenced by the recent controversies surrounding Gliese581 g . The Systemic Console currently offers several tools for error estimation, including a Levenberg-Marquardt (LM) minimizer, a Bootstrap re-sampler, and a Markov-chain Monte-Carlo (MCMC) estimator.

The Bootstrap Monte-Carlo works well when little is known about the underlying process, or the nature of the measurement errors [Press et al. (2007)]. It initializes by

using the collected n $RV_{original}$ data points and generated orbital elements. Successively, it generates any number of new synthetic data sets branching out from the original one and containing m (where $m < n$) original data points. Because of the replacements, the outcome is not the original data set, but a new set with a random fraction of the original points, typically 37% ($\approx 1/e$), replaced by newly generated ones. These new synthetic sets are subjected to the same estimation procedure of the original one and generate new orbital parameters, ELS_{synth}. The synthetic orbital parameters will be distributed around $ELS_{original}$, just as $ELS_{original}$ is part of the distribution around ELS_{true}, the true orbital parameters. A considerable number of bootstrap iterations can generate an estimate of uncertainties based on the obtained standard deviations in the orbital parameters of all the present planets. To aid in the analysis, the Systemic Console provides scatter plots of the correlations between the uncertainties of the orbital parameters.

The Markov-chain Monte-Carlo is an alternative technique for estimating uncertainties and it works in a way that is qualitatively different. The goal of the MCMC method is to generate a chain of states (chain part), or sequence of sets of parameters ELS_n, that are sampled from a desired probability distribution $f(x)$, provided an initial set of parameters ELS_0. The Monte Carlo aspect of MCMC simulation refers to randomness in the generation of each subsequent state. The Markov property specifies that the probability distribution for the determination of ELS_n depends on ELS_{n-1} but not previous states such as $ELS_{n-2}, ELS_{n-3}, ...$ [Press et al. (2007)]. Bayesian methods, often implemented using MCMC, provide a powerful way of estimating the parameters of a model and their degree of uncertainty. For application to RV measurements and parameters determination, the observational errors are believed to be very nearly Gaussian with accurately

estimated variances. Thus, if the data are generated by the model specified by ELS, then the probability of drawing the observed values, $Pr(RV_{original}|ELS)$ (a prior distribution), is roughly proportional to $e^{-\chi^2(ELS/2)}$. If we choose a uniform previous model ELS with $Pr(ELS) \sim 1$, then the probability of obtaining specific parameters given the observed RVs, $Pr(ELS|RV_{original})$ (a posterior distribution), is also roughly proportional to $e^{-\chi^2(ELS/2)}$ [Ford (2005)]. MCMC is important for long-period planets and multiple-planet systems, where there are multiple free, unknown parameters and whose values can be drastically replaced to still obtain similarly good fits. The RV observations of such systems can result in large valleys in the $\chi^2(ELS)$ space that permit a broad range of parameter values. While this type of $\chi^2(ELS)$ space presents difficulties for local minimization routines, the MCMC method is able to jump between these local minima and accurately calculate the posterior probability distribution for model parameters, $Pr(ELS|RV_{original})$. In summary, MCMC has the following advantages compared to other methods:

1. The results can be interpreted by examining the posterior distribution for a given prior and the set of $RV_{original}$.

2. In some cases the resulting distribution can be efficiently updated to account for additional RV observations, or alternative priors, provided they do not greatly alter the posterior distribution.

3. Calculating the next step in a Markov-chain is much faster than performing an additional minimization with re-sampled data.

4. Computationally, MCMC is more efficient than other techniques, particularly for high-dimensional parameter spaces such as multi-planet systems.

The Simulated Annealing (SA) or global minimization is a technique that is suitable for optimization problems, especially ones where a desired global extremum is hidden among many poorer, local extrema, which is often the case for the Doppler RV problem. In practice, it is a more dynamic version of methods such as LM minimizer, and its goal is to reduce the risk of the solution being trapped in a local minimum [Ford (2005)]. At each iteration, random perturbations are applied. These perturbations are initially large and are then gradually reduced. To understand this methodology one can picture the process of crystallization. The material is cooled so slowly that atoms within have time to geometrically align themselves and create a flawless crystalline lattice. By contrast, when a material is "quenched", that is, rapidly cooled, there is ample opportunity for defects to remain in the final structure. The same idea pertains to SA and, provided that the perturber is reduced sufficiently slowly, this method can convert a local minimization algorithm into a global one.

Refitting to synthetic data is another method used to estimate the uncertainty in orbital parameters fitted to the observed data. The same fitting technique is repeatedly applied to many sets of simulated data. Each set of simulated data is meant to represent a possible set of measurement values. If the same fitting procedure is applied to the actual data and each simulated data set, then one can obtain the distribution of *best-fit* parameter values. This method will be implemented and later explained in Chapter 5 to demonstrate the nature of the scripts built during this research.

3.2.1 Least Square Fitting (χ^2)

How does one know whether a particular planetary model provides a good fit to the radial velocity data? In general, the better the job that the model curve does of running through the data, the better the fit. Data, however, are generally not exact, but rather subject to measurement errors such as noise and/or stellar jitter. Observational data never exactly fit the model that is being used, even when that model is correct. For this reason, it is necessary to assess whether or not the model is appropriate, or in other words, to test the *goodness of fit* against some useful statistical standard. To be useful, a fitting procedure should provide three important components: parameters, error estimates on the parameters (or a way to sample from their probability distribution), and a statistical measure of the *goodness of fit*. When the third item suggests that the model is an unlikely match to the data, then the first two items are meaningless. In the Systemic Console, the *goodness of fit* is quantified by two numerical measures. 1) The root mean square (RMS) of the fit corresponds to the square root of the average (mean) of the squares of the velocity differences between the actual radial velocity measurements and the model curve. 2) The Reduced Chi Squared (χ^2) is a related statistical measure that gives stronger weight to points that have smaller error bars. In general, the best fit will have a $\chi^2_{reduced} \approx 1$. The error bars on the RV data points correspond to estimates of the uncertainties introduced at the telescope by the measurement process and, importantly, they do not include another source of error known as the *stellar jitter*. The latter is caused by the star's photosphere convective eddies, spots, plages, granulation and stellar oscillations. These phenomena alter the shape of the stellar spectral lines, injecting random velocity shifts (RVs) that may mask or mimic planetary signals. For stars similar to the Sun (G spectral type), the stellar jitter

generally has a value in the neighborhood of $3-5$ m/s [Wright et al. (2007)].

Now, suppose there is a set of n, RV data points, $v_{obs}(t_i|ELS_k)$ where t_i is the time of the i^{th} observation with explicit dependence on the m orbital parameters, ELS_k, and where $i = 1, ..., n$ and $k = 1, ..., m$. The purpose is to establish if these data points are well described by a modeling function $y = f(v_{model})$. Assuming that the uncertainties from single observations are expected to closely follow a normal distribution, the latter is given by

$$y = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(v_{model} - \mu_i)^2}{2\sigma_i^2}\right], \tag{3.5}$$

where μ is the *expectation* value or location of the peak, σ is the standard deviation, and σ^2 is the *variance*, a measure of the width of the distribution (see Figure 3.2). To evaluate

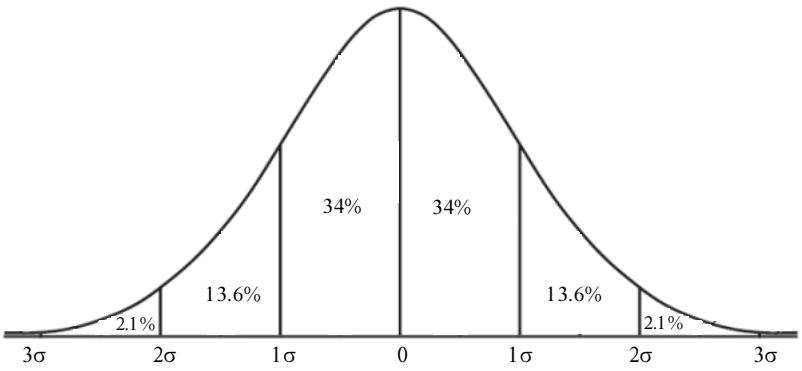


Figure 3.2: In a normal distribution, 1σ accounts for $\approx 68\%$ of the RV data set, while 2σ for $\approx 95\%$, and 3σ for $\approx 99.7\%$. A small σ indicates that the data points tend to be very close to the mean, whereas a large σ indicates that the data points are spread out over a large range of values.

the *goodness of fit* for our given model, y, we can calculate the χ^2_{RV} statistic:

$$\chi^2_{RV}(\upsilon) = \sum_{i=1}^{n} \frac{[\upsilon_{model,i} - \upsilon_{obs}(t_i|ELS_k)]^2}{\sigma_i^2}, \tag{3.6}$$

where σ_i is the observational uncertainty of the i^{th} radial velocity observation. In particular, the Systemic Console implements a minimization of a fit through the *Reduced Chi Squared Statistic* [Meschiari et al. (2009)] as follows:

$$\chi^2_{RV_{Systemic}}(\upsilon) = \frac{1}{n_{RV} - m_{ELS}} \sum_{i=1}^{n} \frac{[\upsilon_{model,i} - \upsilon_{obs}(t_i|ELS_k)]^2}{\sigma_i^2}. \tag{3.7}$$

In this case, $m = 5$: period, mass, mean anomaly, eccentricity, and longitude of periastron. Parameters minimization, is obtained by finding the minimum, χ^2_{min}. The new orbital elements are identified as the *best fit model*, but it is worth remembering that finding the set of parameters that minimizes χ^2 can be mathematically exhausting, particularly for multiple-planet systems. For this task, computer algorithms can save much time and eliminate random errors.

3.2.2 Levenberg-Marquardt (LM) method for local minimization

The Levenberg-Marquardt, or local minimization method is an iterative minimization algorithm (a process repeated with the aim of approaching the desired result) and it has become a standard non-linear least-squares routine. LM is useful for finding the best fitting parameters of a non linear model to a set of data for which plausible starting guesses, n_0, have already been established [Press et al. (2007)]. Once an initial, reasonable estimate of the orbital parameters is available, it can be used to refine them by minimizing the χ^2. Unfortunately, even though the χ^2 space can have many local minima, the LM algorithm will identify only a single local minimum [Ford (2005)]. According to *"Numerical Recipes*

in C" by Press et al. (2007), given an initial guess for the set of fitted parameters n_0, the recipe is as follows:

- Compute the $\chi^2_{RV}(\upsilon)$ as shown in Equation 3.7
- Pick a small value for λ (e.g. $\lambda = 0.001$).
- Solve the linear equation

$$\sum_{l=1}^{m} \alpha'_{kl} \delta ELS_l = \beta_k. \tag{3.8}$$

- If $\chi^2_{RV}(\upsilon + \delta\upsilon) \geq \chi^2_{RV}(\upsilon)$ then λ should be increased by a factor of 10 and Equation 3.8 resolved.
- If $\chi^2_{RV}(\upsilon + \delta\upsilon) \leq \chi^2_{RV}(\upsilon)$, then λ should be decreased by a factor of 10 and Equation 3.8 resolved.

Equation 3.8 is the key formula of the Levenberg-Marquardt method and the calculations to derive it unfold as follows.

If the original fit is ELS_k, where $k = 0, ..., 5$ identifies the orbital parameters to minimize, then its $\chi^2_{RV}(\upsilon)$ can be used as a starting point. The gradient of χ^2 with respect to ELS is vector that indicates the direction of steepest descent and is given by

$$\nabla\chi^2_{RV}(\upsilon) = \frac{\partial\chi^2}{\partial ELS} = -\frac{1}{2}\sum_{i=1}^{n} \frac{[\upsilon_{model,i} - \upsilon_{obs}(t_i|ELS)]}{\sigma_i^2} \frac{\partial\upsilon_{obs}(t_i|ELS)}{\partial ELS_k} = -2\beta_k, \tag{3.9}$$

where $(t_i|ELS)$ is the i^{th} observation obtained at time t and if $\nabla\chi^2_{RV}(\upsilon) = 0$ then $\chi^2_{RV}(\upsilon)$ is already at its minimum. The partial derivative with respect to ELS of Equation 3.9 can also be thought of as the curvature of the function and is given by

$$\frac{\partial^2\chi^2}{\partial ELS_k \partial ELS_l} = 2\sum_{i=1}^{n} \frac{1}{\sigma^2} \left[\frac{\partial\upsilon_{obs}(t_i|ELS)}{\partial ELS_k} \frac{\partial\upsilon_{obs}(t_i|ELS)}{\partial ELS_l} - C \right] = 2\alpha_{kl}, \tag{3.10}$$

where l is the next iteration and generates a new, improved set of parameters ELS_l. In Equation 3.10, the term C can be ignored as it is the second derivative of $\partial^2\chi^2$ multiplied by a random $\pm$ measurement error of each point which cancels out when summed over i. Therefore we are left with

$$2\alpha_{kl} \approxeq 2\sum_{i=1}^{n} \frac{1}{\sigma^2}\left[\frac{\partial v_{obs}(t_i|ELS)}{\partial ELS_k}\frac{\partial v_{obs}(t_i|ELS)}{\partial ELS_l}\right], \tag{3.11}$$

and

$$\alpha_{kl} = \frac{1}{2}\frac{\partial\chi^2}{\partial ELS_k \partial ELS_l} \quad , \quad \beta_k = -\frac{1}{2}\frac{\partial\chi^2}{\partial ELS_k} \quad \text{and} \quad \beta_l = -\frac{1}{2}\frac{\partial\chi^2}{\partial ELS_l}. \tag{3.12}$$

Sufficiently close to the minimum, we expect the $\chi^2_{RV}(\upsilon)$ function to be well approximated by a quadratic form. However, this representation might be a poor local approximation to the shape of the function that we are trying to minimize. In this case, the suggested method is to take a step down the gradient

$$ELS_l = ELS_k - constant \times \nabla\chi^2(ELS_k), \tag{3.13}$$

and based on Equations 3.9, 3.11 3.12, and 3.13 we have

$$\delta ELS_l = constant \times \beta_l, \tag{3.14}$$

where the constant is found by looking at the dimensions of δELS_l and β_l and therefore $constant = 1/\lambda\alpha_{ll}$ where λ is a non-dimensional fudge factor that can be $\gg 1$, ≈ 1, or $\ll 1$. Putting it all together Equation 3.14 becomes

$$\delta ELS_l = \frac{1}{\lambda\alpha_{ll}}\beta_l \quad \text{or} \quad \delta ELS_l \lambda\alpha_{ll} = \beta_l. \tag{3.15}$$

Once more, the equation we need to solve for exactly displayed as in Equation 3.8 is

$$\sum_{l=1}^{m} \alpha'_{kl}\delta ELS_l = \beta_k, \tag{3.16}$$

and can be solved with the help of Marquardt's insight by recognizing that

$$\text{if} \quad k = l \rightarrow \alpha'_{kl} = \alpha_{ll}(1+\lambda) \quad \text{and if} \quad k \neq l \rightarrow \alpha'_{kl} = \alpha_{kl}. \tag{3.17}$$

While there is a degree of ambiguity in the choice of the weighing factor λ, this is not a concern in the vicinity of a solution, where the contribution from RVs is approximately equal for $\lambda = 1$. Far from the solution, $\chi^2_{RV}(\upsilon)$ can become very large due to possible transits. However, this effect can be tamed by fitting one planet at a time and reducing the initial $\chi^2_{RV}(\upsilon)$ [Meschiari & Laughlin (2010)].

In conclusion, iteration can stop after few occurrences of χ^2 decreasing by a negligible amount (e.g. $> 10^{-3}$). If χ^2 increases more than trivially it means that λ has not yet adjusted itself optimally. As noted before, when fitting for parameters that characterize a non-linear model, the χ^2 space might not have only a single minimum. The LM method offers no insight into finding the global minimum. It is just a downhill (gradient-based) search and it should be regarded as the resource for achieving a final optimization of the parameters, preceded by more global methods, such as MCMC or SA, for manipulating the system into the correct basin of convergence.

4

Synthetic Software Overview

Two main problems occur when analyzing extra-solar planetary data: 1) error estimates within the non-stationary, significantly varying signal-to-noise RV data sets can be underestimated, and 2) there is no continuous, up-to-date library of known extra-solar planetary fits associated with a strictly uniform analysis. The Systemic Console collaboration been established to address this problem. The Console is an all-in-one software package built with the purpose of analyzing and combining multi-parameter fitting of Doppler RV and transit timing observations [Meschiari et al. (2009)]. The Systemic Console is freely downloadable at `http://www.ucolick.org/~smeschia/SystemicConsole`. The Console includes a Graphical User Interface (GUI) and a library of routines which are very useful to write scripts in a non-interactive mode. An extensive and well-explained list of Javadocs is provided as well as a link to useful tutorials on how to use the GUI. The program is written in the Java programming language, simplifying cross-platform portability.

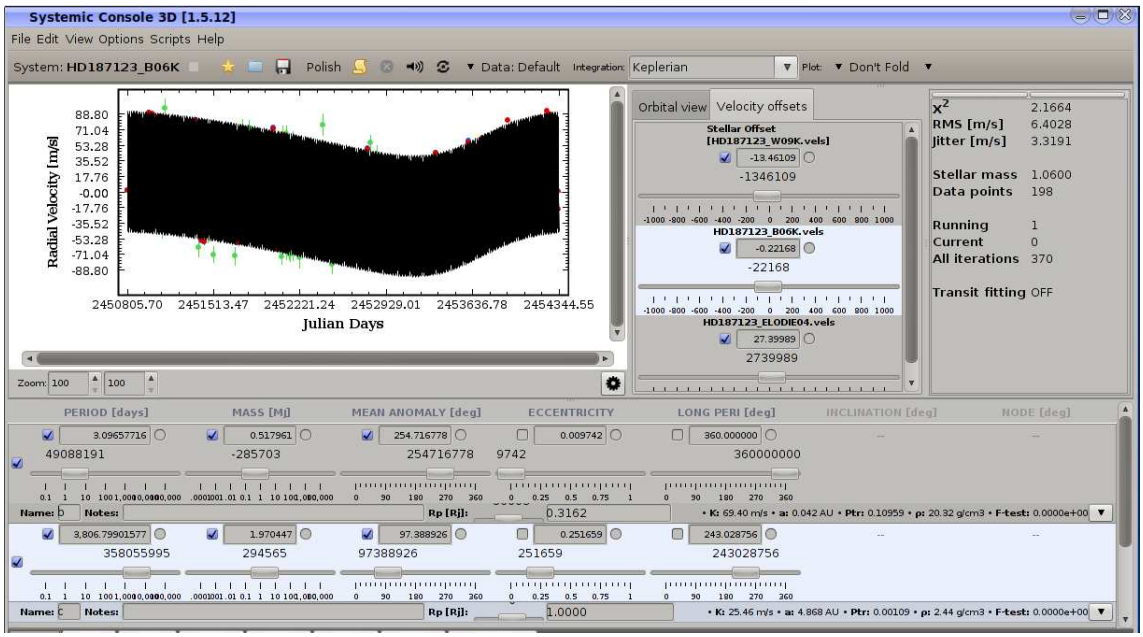


Figure 4.1: The Console GUI displaying the star HD187123's loaded RVs. The sinusoidal wave delineates and covers almost perfectly the RV data, suggesting that the fit is good. On the right hand side, the χ^2 and *RMS* are displayed. On the bottom, the two sets of planetary parameters are listed.

4.1 GUI of the Systemic software (Graphical User Interface)

The GUI, shown in Figure 4.1, can display

1. RV curves and transits which can be folded to a chosen period (on the upper left),
2. orbital parameters of the chosen planetary architecture (on the bottom),
3. orbital figures of each of the chosen planets (not displayed in the above figure),
4. stellar velocity zero-point offsets (on the upper center),
5. summary statistics associated with error analysis such as *RMS* and χ^2 (on the upper right),

6. a Lomb-Scargle (LS) periodogram which can also display the periodicities of the residual RVs after a primary selection (see Figure 4.2),

7. minimization schemes such as, but not limited to Levenberg-Marquardt (LM), or Simulated Annealing (SA) (these are controlled with the *Polish* button located on the left of the upper row and in the *scripts* tab respectively),

8. integration routines such as, but not limited to Keplerian, Runge-Kutta, or Bulirsch-Stoer N-body integration (control buttons located on the right side of the upper row the *integration* drop-down menu),

9. error estimations algorithms including Bootstrap, Resampling, and Markov Chain Monte Carlo (MCMC) (under the *scripts* tab).

The Console features many other useful tools for data analysis. One of them allows the Console to perform long-term integrations that can help decide whether a planetary system is stable or not. This utility is accessed via *Orbital Evolution & Stability* under the *View* menu. A variety of integration schemes are available, and the results can be plotted. The integration is performed and displayed in real time. As a rule of thumb, if the orbital elements of the planets are varying only slowly, over a period of time in the order of 10^3 years, then the system is recognized as stable.

Another important feature of Console is the possibility to phase fold the RV data to a chosen period. This task can be executed by either manually entering a selected number of days through the GUI and then selecting the *Fold to period* button from the drop-down menu or by selecting the most probable period displayed in the *Periodogram* window as shown in Figure 4.2. The tallest peaks in the periodogram are used as guesses for the

fitting algorithm where the dominant planet (possibly identified by the tallest peak) can be fit for alone. Its signal is then subtracted from the data, and additional planets (or peaks) can be searched for among the residuals. This process can then be repeated until all of the planets have been identified, and then a full, multi planet fit on the original data starting at the values found for the individual planets. This is a graphical interpretation of

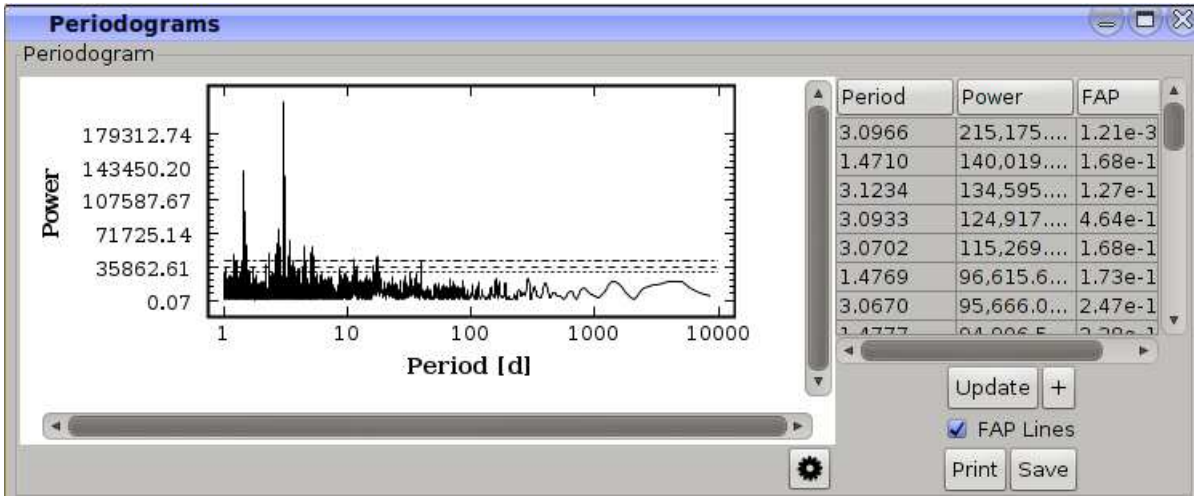


Figure 4.2: Lomb-Scargle Periodogram of the HD187123 RV set shown in Figure 4.3 displaying from top to bottom the most probable periodicities. The relative Powers and analytical False Alarm Probabilities (FAP) at levels 10^1 (long dashed), 10^2 (short dashed), and 10^3 (dotted) are overlaid.

a Fourier method called the Lomb-Scargle (LS) periodogram. It is an algorithm for time series analysis of unevenly spaced data [Meschiari et al. (2009)] and one of the means used to identify the possible periodicities of prospective planets.

4.1.1 Introduction to Fourier technique

The choice of a Fourier technique to analyze the RV data is an appropriate one because it consists of observations that vary with time. A data set, composed of n discrete observations, taken at times t_i, are necessarily limited in length. Both data length and

data spacing have important limiting effects on the accuracy with which Fourier analysis can be performed. The Fourier technique adopted in the Systemic Console applies is well-suited to data that has been unevenly sampled. A deterministic model assumes a basic predictability and repeatability, which is the signature character of classical, non-quantum physical processes. A discrete versions of the Fourier transform where the integration (or summation) limits are finite rather than infinite is appropriate for discrete data. A discrete Fourier transform is defined as

$$F(\upsilon) = \sum_{i=1}^{n} f(t_i) e^{i2\pi t_i}, \tag{4.1}$$

where $f(t_i)$ is a pure cosine wave of frequency v_0 and the Fourier transforms $F(v)$ has amplitudes that are significantly different from zero only in the immediate vicinity of $v = \pm v_0$. A Fourier analysis is able to detect the presence of a frequency in the data and, with some care in the normalization, to determine its amplitude. In the case of a multiply periodic function, such as in the case of the LS periodogram, with frequencies $v_0, v_1, v_2, ...$,etc., the transform will be large in the vicinity of $v = \pm v_1, \pm v_2, ...$,etc., and, ideally, the analysis can detect the presence of each of these frequencies independently, and determine their amplitudes. In practice, this ideal cannot quite be realized because of the finite data length and the discrete sampling. The interference between frequencies is the limiting factor to a straight forward implementation. Two are two types of such interference: 1) interference from nearby frequencies, which is described by a *spectral window*, $\delta(v)$, and is primarily a product of the finite length of the data, and 2) interference from distance frequencies, which is called *aliasing*, and is a product of the data spacing. $\delta(v)$ is a function only of v and the

times of observation t_i [1]

$$\delta(v) = \sum_{i=1}^{n} e^{i2\pi v t_i}. \tag{4.2}$$

In practical computations it is convenient to use instead a corresponding spectral window

$$\gamma(v) = \frac{\delta(v)}{n}, \tag{4.3}$$

because this variable returns a normalized spectral window, $\gamma(0) = 1$. Thus, if $F(v)$ is a series of delta functions, corresponding to a multiplicity of periodic functions, then $\gamma(v)$ will consist of a series of spectral windows centered on the various frequencies present. The signature of the data distribution is all contained in the spectral window, $\gamma(v)$. A plot of the amplitude of $\gamma(v)$ versus frequency shows 1) a reasonably well defined central peak at $v = 0$ and 2) some subsidiary peaks corresponding to peculiarities in the data spacing. For example, Figure 4.2 shows the effect of having uneven data spacing. The spectral window obtained for the given series of observations shows the presence of a sharp peak at frequency $v_0 = 1/3.0966$ days. If another moderate frequency was present in the data at v_a (e.g. $v_a = 1/3810$ days), an interference between the two frequencies would arise in the data spacing. Thus, there should be subsidiary peaks (*aliases*) in the transform at frequencies $v = v_a \pm v_0$. However, in this specific case, the weak signal obtained by the period $P_c = 3810$ days does not produce any relevant subsidiary peaks.

4.1.2 A practical example: system HD187123

The year 1998 brought the announcement of a $M \sin i = 0.5\ M_{♃}$ planet in a three-day orbit around HD 187123 [Butler et al. (1998)]. This star is constant in brightness [Robb R.M. et al. (1999)] and is a close solar analog: $M_\star = 1.1\ M_\odot$ and $T_{eff} = 5810K \approx$

[1] if the data points were to be continuous, then $\delta(v)$ would be a Dirac-Delta function

$T_{eff,\odot}$. In 2007, a second outer companion was announced with orbital period > 10 yr [Wright et al. (2007)]. Assuming that there is no linear trend or detectable third planet in the system, the data constrain HD187123c to have $P = 10.4 \pm 1.2\ yr$, $e = 0.25 \pm 0.03$, and $M \sin i = 2.0 \pm 0.3\ M_{\mathrm{Jup}}$. The system can be analyzed using the tools built in the Systemic Console and utilizing both the older data from 1998 and the latest collected RVs presented in 2007. This planetary system can be loaded and analyzed through the GUI. First, the RV data is selected and loaded from the *datafiles* folder, using the ★ button located at the upper left corner. The RV data will be displayed as shown in Figure 4.3. Specifically in this case, the set chosen has been modified to allow the introduction of more up-to-date RV data since the recent observations, made at the Keck telescope, have enabled the discovery of the second planet. The most recent RVs can be found at

http://iopscience.iop.org/0004-637X/693/2/1084/apj_693_2_1084.tables.html

[Wright et al. (2007)], and are shown in Figure 4.3 together with the older collected data.

The easily observable peaks of its LS periodogram, (Figure 4.2), correspond each to a period with a given Power and False Alarm Probability (FAP). If the period of 3.0966 days, accompanied by the highest Power and lowest FAP (1.21×10^{-31}) is selected, a new planetary system is created. The RV set can now be folded to this new period and is shown in Figure 4.4. At this point, the GUI can further be further employed to reshape the sinusoidal function created and exactly trace and overlap the set of folded RVs. If a slight eccentricity is added to the orbital parameters, and the mass is set $M_b \approx 0.5\ M_{\mathrm{Jup}}$ the sinusoidal model, begins to match the delineated curve formed by the collected RVs. By adding another planet with a mass of $M_c \approx 2.0\ M_{\mathrm{Jup}}$, period of $P_c = 3810$ days, and eccentricity of $e_c = 0.25$, the fit seems to improve drastically. The RV data, however, suggest that the time of the first

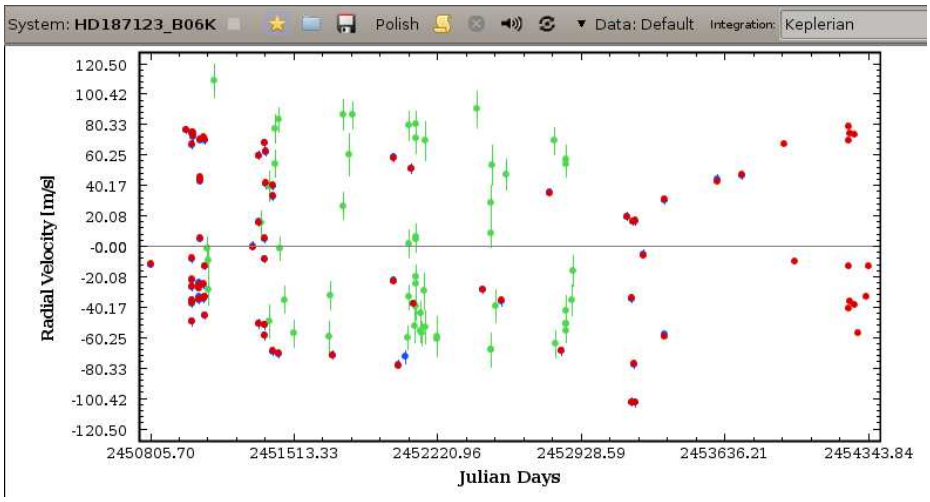


Figure 4.3: 65 RVs collected by the Keck/HIRES telescope spanning 8 years (blue), 57 RVs collected by the ESO/Elodie telescope spanning 5 years (green), and 76 new RVs collected up until August 2007

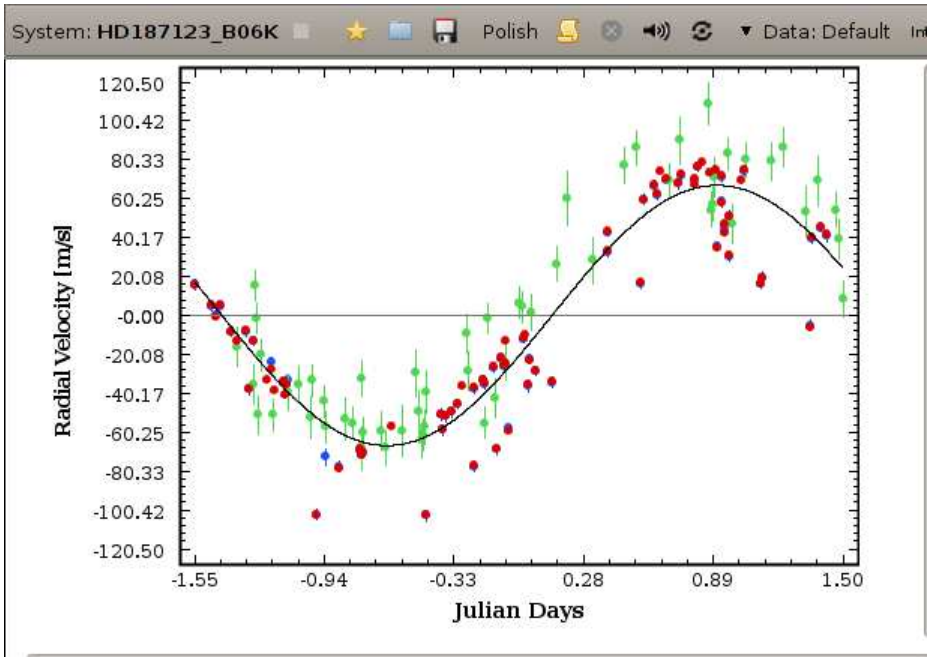


Figure 4.4: The same set of RVs shown in Figure 4.3 has been folded to a period of 3.09 days and clearly shows a sinusoidal shape.

measurement did not occur at the maximum of the velocity cycle. An improvement can be obtained by adjusting the *mean anomaly* slider in the GUI (see Figure 4.5). This feature slides the modeling curve to the left or to the right without changing its shape.

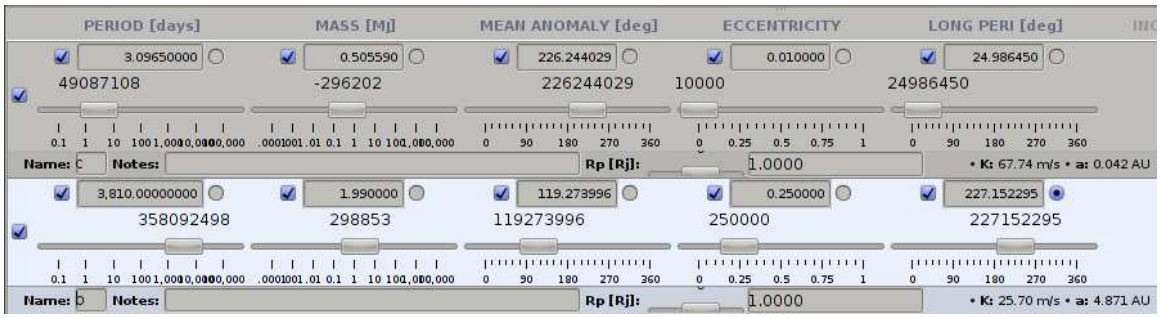


Figure 4.5: The bars shown appear when one or more planets are added to the given RV set after selecting the most probable period from the periodogram

4.2 Location and organization of the code

The Console can be used in two different modes: through the GUI as explained in Section 4.1 or through scripts that can also be available in the GUI, but do not strictly depend on it. The latter is the most effective mode for developers and for "no frills", straight forward implementation. At the same web address, a complete guide to the *Javadocs* needed to understand the already existing scripts and to write new ones can be found. After the Systemic Console is downloaded, the compressed folder will contain two options: a folder for MAC OS and a folder for WINDOWS/LINUX OS. Each of these folders contain:

1. A *Console.jar* file through which the scripts and the GUI must be loaded.

2. A *datafiles* folder where all RV data is stored in *.vels* and *.sys* files.

3. A *lib* folder containing all the pertinent *java* libraries.

4. A *fits* folder where all the *.fit* files will be eventually collected and saved.

5. A *altnames.txt* file containing an extensive catalog of approximately 1500 candidate stars that could host a planetary system.

6. A *scripts* folder containing some of the scripts used by the Systemic Console, but not all as the full code is only available to the developers involved.

The scripts are key to understanding the code. The first one, *derivefit.conscript*, can be called directly from the GUI (as all other scripts) when RV data is loaded. It will examine and use the most prevalent peaks in the LS periodogram and residuals and, after evaluating their RMS, will use these values as starting guesses for the planetary masses. For each planet found, it will automatically perform a Keplerian integration, then activate a Levenberg-Marquardt (LM) minimization routine (explained in Section 3.2.2), and finally update the GUI as it moves forward in the iterations. The other script available in this folder is *genSystems2.conscript*, which was written to generate synthetic datasets. After a set of orbital elements is specified, the code generates a file containing an array of synthetic RVs (*.vels* file), a file containing an array of observed transits (*.tds* file), and a file containing a summary of star's mass, RVs and transits (*.sys* file). Finally, when running one of these or any other Console scripts through a terminal shell, the command must be entered where the *Console.jar* file is: java -jar Console.jar [location of the script]/[script].conscript.

4.2.1 .vels, .sys, and .fit files

The folder *datafiles*, contains one or more sets of RV data, collected in the past 25 years, for each of the stars listed in the *altnames.txt* file. These files, identified by the characteristic "*.vels*" suffix, are named after the star and the telescope that collected the

RVs and contain an array composed of three columns. Each line lists the Julian date of the measurement, the RV measurement, and the σ value (or standard deviation from the mean) respectively. The "*.vels*" files are then grouped together in a "*.sys*" file that list the star's mass (compared to solar masses, $M_{\odot}$) and the *.vels* files available. The "*fits*" folder is empty at first, but as new planetary systems are created, it will get filled with "*.fit*" files. These files list the orbital parameters of the guessed planetary system and gives other specifications in the following order:

1 The first Julian Date of observation.

2 Identification of the "*.vels*" and "*.sys*" files used and to which folder they belong.

3 The relative RV offset as briefly explained in Section 2.4.2.

4 The planets' periods, masses, mean anomalies, eccentricities, longitudes of periastron, inclinations, nodes, and radii.

Table 4.1 is an excerpt of a *.vels* file displayed RVs of the star HD187123 and collected at Keck I, Mauna Kea, Hawaii.

Julian Date (JD)	$\boldsymbol{RV_{Orig}}$(m/s)	$\boldsymbol{\sigma_{Orig}}$
2450805.7017	-12.30	2.00
2451009.9419	-27.00	2.20
2451011.8744	-8.50	2.20
2451012.8400	-49.90	2.70
2451013.0751	-22.40	2.10
2451013.9150	72.30	2.90
2451042.9614	-24.10	2.40
2451043.9607	-35.70	2.50
2451050.7308	42.40	1.40
2451051.7294	44.80	1.40
2451068.8311	-25.30	2.30
2451070.8914	-34.80	2.40
2451072.8169	70.00	2.40
...	...	...

Table 4.1: Sample of the 60 detected RVs at the Keck Observatory, starting on December 23, 1997 and ending on December 19. By astronomical convention, motion away from the Earth is defined as a positive radial velocity.

5

The Code

The purpose of the code developed and presented in this paper is to reproduce and analyze synthetic RV data points that constitute a model for what might prospectively be collected with ground-based telescopes. Because the TESS mission is still in the proposal phase, the first part of the code is dedicated to generating data conforming to the distribution of planets that we believe TESS can detect. The second part of the simulation framework was built to output and analyze multiple RV data sets, from different ground locations, and to reproduce data that replicates the expectations from realistic observing conditions. Finally, the pipeline finds the best orbital parameter fit for a given system through a LM minimization. The scripts described below were built using the Java script, Beanshell, an embeddable Java source interpreter. They begin by initializing a new Kernel context, the bridge between one or more application and the actual data processing done at the hardware level.

5.1 Content of the TESS folder

The TESS folder contains various *.conscript* and *.txt* files, and two Java classes. The *.conscript* files are executable, while the *.txt* files are arrays, data types that contain a series of elements of the same type placed in contiguous memory locations and that can be individually referenced. Specifically, the folder includes:

- **Mass_Generator.conscript**. It generates stellar masses ranging from F5 to M5 spectral type masses and are characterized by $0.45\ M_{\odot} \leq M_{planet} \leq 1.4\ M_{\odot}$.
- **Synth_Masses.txt**. It contains the star mass array generated by *Mass_generator.conscript*.
- **Location_generator.conscript**. It generates a set of right ascensions, RA, and declinations, DEC (location of the target in the sky), and includes a rejection method that allows an even distribution of points in the sky and eliminates crowding at the ecliptic poles.
- **Synth_RA_DEC.txt**. It is an array that contains 2500 synthetic star locations generated by *Location_generator.conscript*.
- **Period_Radius_Generator.conscript**. An investigation is currently ongoing regarding whether outer planets are larger than inner planets on average. The script creates periods and radii of exoplanets according to the Lissauer J. (2011) article which states "When considering the entire multi-candidate population, there is a slight but significant preference for outer planets to be larger [Lissauer et al. (2011)]." Random planetary periods are generated in the range of $2\ days \leq P_{planet} \leq 500\ days$ [Ricker (2011)] and planetary radii consistent with the generated periods, $0.1\ R_{\oplus} \leq R_{planet} \leq 12\ R_{\oplus}$, are created.

- **Synth_Period_Radii.txt**. It contains an array of planetary periods and radii created by *Period_Radius_Generator.conscript*.

- **OriginalFitGenerator.conscript**. This code randomly picks a number of planets between 1 and 3 and randomly assigns them a period and a radius drawn from the *Synth_Period_Radii.txt* file. The script also generates planetary masses related to the selected radii and periods according to the power law relationship (Equation 5.1) between Earth and Saturn as shown in Figure 5.1 [Lissauer et al. (2011)]

$$\frac{M_{planet}}{M_{\oplus}} = \left(\frac{R_{planet}}{R_{\oplus}}\right)^{2.06}. \tag{5.1}$$

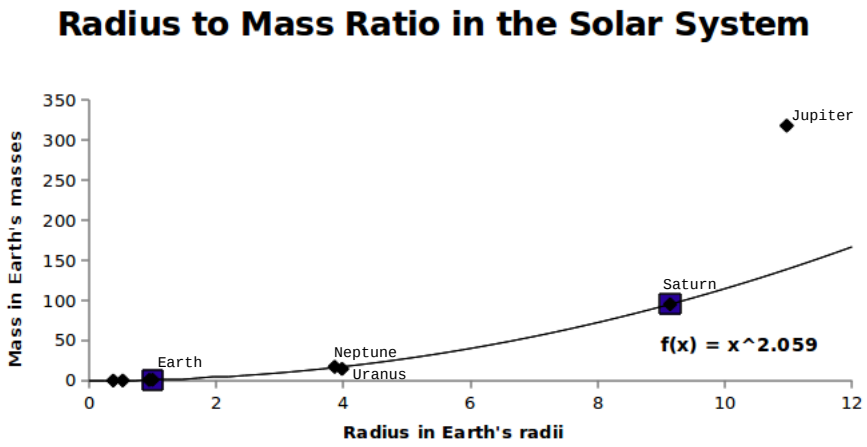


Figure 5.1: The power-law was obtained by fitting to Earth and Saturn. It slightly overestimates the mass of Uranus and slightly underestimates the mass of Neptune. Jupiter, whose equation of state is dominated by degeneracy pressure, is far away from the fit line.

- **PlanetaryParameters.txt**. It contains an array with the planetary parameters picked by the *OriginalFitGenerator.conscript*

- **[system name]_Original.fit**. It contains the previously selected orbital elements, but in a format that is compatible with the Systemic Console scripting language.

- **LAT_LONG_3scopes.txt**. It is an array containing three observing sites, but more can be loaded by simply extending the array with more ground-based locations.

- **RV_Generator_from_FitFile.conscript**. It generates a number of RVs for each loaded ground-based observing site and includes a set rejection methods to simulate at best true observing conditions. A flowchart of this script is shown in Figure 5.2.

  - When the script is initialized, it examines all days within a forward time interval of one year (more observing time can be easily added). The program knows about the current date and time by calling the *Calendar* class available in the *java.util* method. The date is then converted to a Julian date (JD) [Duffett-Smith (1988)] which, given the need to keep track of leap years and Universal Time, provides a much easier way to work with time calculations. The JD is the interval of time in days and fractions of a day since January 1, 4713 BC, Greenwich noon, [http://en.wikipedia.org/wiki/Julian_day], and it can contain up to five decimal places depending on the precision of the recorded time.

  - After the current JD is loaded, the clock is run forward in fifteen minutes increments. This specific time step was chosen because it simulates at best the time gap between ground-based telescope observations: the fainter the object in the sky, the longer the exposure time required to gather enough light to detect a Doppler's shift of the photons' frequency.

  - The script then encounters the first constraint. A science team cannot make

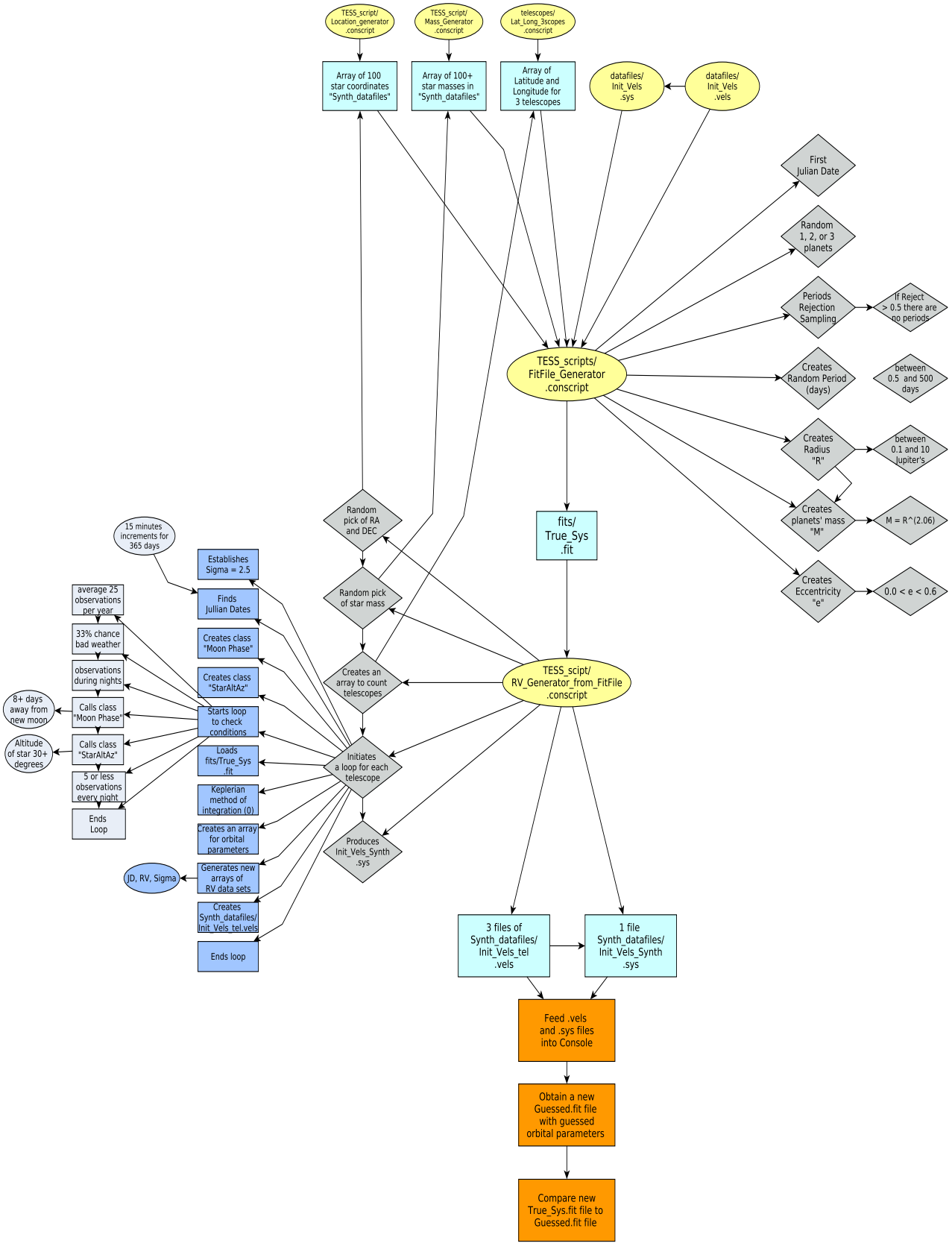


Figure 5.2: Flowchart describing the sequential steps taken by the *RV_Generator_from_FitFile.conscript* to create sets of synthetic RVs

ground-based observations each day of the year, therefore a random float number (with six decimal places) anywhere between 0 and 1 is generated and if the value of the float number is less than 0.9845, the script stops and moves on to the next observing time, fifteen minutes later. This constraint represents a to 1.55% chance of available telescope time or an average of 25 observations every year. The yearly frequency of observations is the same as it would be for a science team at the Keck I telescope in Mauna Kea, Hawaii [http://www2.keck.hawaii.edu/observing/schedule/index.php].

- The next constraint samples the distribution of atmospheric conditions to see if they are favorable to observe. The code performs the same random number draw as it did in the previous constraint, and it assumes a 66.0% chance of favorable weather.
- The third constraint is implemented to differentiate between night and day. Knowing that a Julian date starts at noon, the code is instructed to consider only time increments between 10:00 pm and 4:00 am. A more refined version of the code will be able to include Sun's rise and set times and will adjust the time available to observe with respect to sky darkness.
- **MoonPhase** Java class called by the *RV_Generator_from_FitFile.conscript* to calculate the phase of the Moon [Duffett-Smith (1988)] and is examined in detail in Section 5.1.1.
- **AltAz** Java class called by the *RV_Generator_from_FitFile.conscript* to calculate the altitude of the chosen star in the sky. It initializes by calculating the local sidereal time (LST) and then proceeds to find the hour angle (HA), altitude

(ALT), and azimuth (AZ) (see Section 5.1.2).

- **derivefit_levenberg.conscript** loads the newly created RV data and the *[system name]_Original.fit* file and runs the Levenberg-Marquardt minimization method to find the *best fit*. The code also displays the improved $\chi^2_{min,RV}(\upsilon)$ and *RMS*.

- **[system name]minimized.fit** is generated by the *derivefit_levenberg.conscript* and describes the improved orbital elements.

The flowchart displayed in Figure 5.3 shows how the full code proceeds through the scripts to generate the "best fit".

5.1.1 Full moon constraint

As stated above, the time available to make observations and collect data on a telescope is restricted to an average of 25 days per year. The optimal observing times occur when light pollution and atmospheric moisture are at minimum, but in these circumstances, the priority for observations is given to teams who collect data from very faint objects such as galaxies. Because the TESS research is focused on fairly bright stars (F5-M5 spectral type) such as the Sun (a G2-type star), sufficiently accurate observations can be made relatively close to a full moon cycle. However, the script will constraint the program from making observations too close to new moon, the darkest days, such that if $-8\ days \leq t_{\bullet} \leq +8\ days$, the code will move on to the next JD. The calculations used to correctly predict the moon phase must take into account the ecliptic orbit of the Earth with respect to the Sun, the mean anomaly of the Sun ($M_{\odot}$) which refers to the average motion of the Sun in a circular orbit, its ecliptic longitude ($\lambda_{\odot}$), the Moon's true orbital longitude ($\lambda_{☾}$) and its mean anomaly ($M_{☾}$), and five smaller corrections including, but not limited to evection, annual

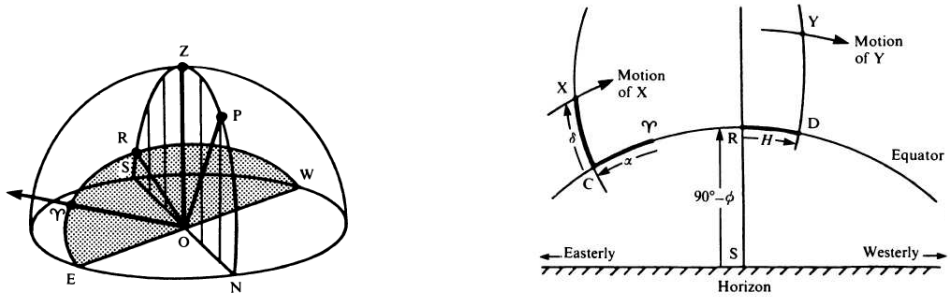
Figure 5.3: Flowchart describing the sequential steps taken by the code to create the *best fit* for a given planetary system.

equation, and variation [Duffett-Smith (1988)]. Finally, the correct moon phase can be calculated using Equation 5.2 where μ is the Moon's phase spanning from 0 (new moon, ●) to 1 (full moon, ○)

$$\mu = \frac{1}{2}[1 - \cos(\lambda_{☾} - \lambda_{\odot})]. \tag{5.2}$$

5.1.2 Star's altitude constraint

The last constraint that needs to be considered is the altitude and azimuth of the target star. The best location to observe a celestial body is at the zenith (90 ° angle) and the worst, or rather impossible time, is when the target is located on the horizon (0 ° angle). Telescopes such as the Keck HIRES (Hawaii) and the VLT (Chile) can make quality observations well below the right angle of altitude. The constraint in the code was set to limit observations of celestial bodies at $> 30°$ above the horizon. The altitude of a star is



(a) **(b)**

Figure 5.4: Equatorial coordinates. a) On the celestial sphere where the observer, O, is on the northern hemisphere. b) As seen from the ground by the observer, O, looking up in the sky.

found by using the latitudinal and longitudinal coordinates of the observing site and the RA and DEC of the target celestial body as displayed in Equation (5.3).

$$\sin(ALT) = [\sin(DEC) \times \sin(LAT)] + [\cos(DEC) \times \cos(LAT) \times \cos(HA)] \tag{5.3}$$

The coordinate system used to calculate this last constraint can be better understood by looking at Figure 5.4. In Figure 5.4a , the observer, O, is at the center of the hamisphere, the semicircle E♈RW is the equatorial plane and the line OP is the axis of rotation. At P all stars seem to describe circles as they move across the sky. In Figure 5.4b the equatorial coordinates are represented as seen from the ground in the northern hemisphere and the curve C♈RD represents the equator and the star is represented by X. XC represent the Declination (DEC) describing how far up in the sky the star is located from the equator. C♈ represents the Right Ascension (RA) describing the westward movement of the star. RA is a fixed point in the sky because the star moves along the equator with the rest of the bodies. H describes the Hour Angle (HA) and it is a measure of how far the star Y has traveled along the equator from the southern point R (time since the star crossed the meridian).

6

Results: The code output

The code first geneartes synthetic RVs. Since the observational errors are believed to be of Gaussian nature and each radial velocity measurement has a corresponding uncertainty estimate, it is possible to construct simulated data sets by adding Gaussian random values to the actual data points. Each synthetic data set is meant to represent a possible set of measurement values [Ford (2005)]. If the same fitting procedure is applied to the actual and synthetic data set, then it is possible to obtain the distribution of *best fit* orbital parameters. In principle, it is prudent to apply a global minimization routine to each data set such as MCMC or SA, but in practice the heavy computational load suggests that a local minimization routine such as the Levenberg-Marquardt is more appropriate. Even then, the computational requirements can be challenging. For instance, for a system characterized by four free parameters, the code must recalculate the χ^2 8 times for each planet and then once more after including all present planets. The synthetic *.vels* files include JDs, RVs and a $\sigma = 2.5$ m/s common to all data sets. Similar in format to the previously collected RV set explained in Section 4.2.1, Table 6.1 is an excerpt of a synthetically created *.vels*

file and is similar to Table 4.1. Then, minimization routines can be implemented on these synthetic RVs to find the *best fits*.

Julian Date (JD)	$\boldsymbol{RV_{Synth}}$(m/s)	$\boldsymbol{\sigma_{Synth}}$
2455796.4838	81.17	2.50
2455821.4629	64.87	2.50
2455836.4838	90.57	2.50
2456061.5879	-13.98	2.50
2456076.4733	-55.18	2.50
2456091.5254	-27.05	2.50
2456109.5150	51.75	2.50
2456115.4213	77.85	2.50
2456116.5046	-43.35	2.50
...	...	...

Table 6.1: Sample of synthetically created RVs for the system HD187123 based on the Keck I telescope coordinates. By astronomical convention, motion away from the Earth is defined as a positive radial velocity.

6.1 Example continued: HD 187123 and its planetary system

Continuing with the example presented in Section 4.1.2, it is possible to derive self-consistent fits and Monte-Carlo dynamical analyses of the orbital parameters of the HD187123 system [Meschiari et al. (2009)]. Supposing that the planets HD187123b and HD187123c had been discovered by a previous survey and that their periods, P_b and P_c, and radii, R_b and R_c, were given with great accuracy from the RVs and transits detection, the planets' masses, m_b and m_c are calculated using Equation 5.1. The two periods and masses are loaded into a *[system name]Original.fit* file which can then be loaded into the executable file *derivefit_levenberg.conscript*. The latter minimizes the masses, mean anomalies (M_b and M_c), eccentricities (e_b and e_c), and longitudes of periastron (ω_b and ω_c). The reduced χ^2_{RV}

Planet	$\boldsymbol{P}$(days)	$\boldsymbol{m(m_{\Uparrow})}$	$\boldsymbol{M(^\circ)}$	$\boldsymbol{e}$	$\boldsymbol{\omega(^\circ)}$	**Uncertainties**
$Minimized_b$	3.10	0.50	334.96	0.02	279.57	$\chi^2_{RV,b\_c} = 26.98$
$Public_b$	3.10	0.52	N/A	0.01	25.00	
$Minimized_c$	3,810.00	1.60	121.50	0.79	229.33	$RMS_{b\_c} = 13.29\ m/s$
$Public_c$	3,810.00	1.99	N/A	0.25	250.00	

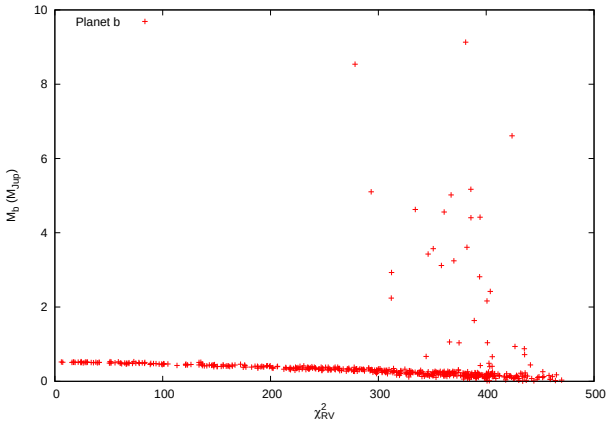
Table 6.2: Orbital elements of the star HD187123 including uncertainties after the execution of the LM minimizing method. The periods were left unchanged because the LM routine is not applied to periodicities as TESS will give their accurate values.

generated from the planetary elements of HD187123b solely, and without applying a minimization routine is fairly high, $\chi^2_{RV,b} \approx 1988.0$. The least square fit obtained with both planets present, $\chi^2_{RV,b\_c} \approx 1978.0$, still describes a rather poor fit (considering that for a perfect fit $\chi^2_{reduced} \approx 1$), but is lower than $\chi^2_{RV,b}$ and therefore it can be considered as a step towards a possible solution. When the LM minimizing routine is executed, better results are achieved for the least square fit, given that the four orbital elements introduced at first are acceptable guesses. The system obtained in Table 6.2 exhibits a definitely improved Chi Squared ($\chi^2_{RV,b\_c} = \chi^2_{Minimized_b} + \chi^2_{Minimized_c}$), but some parameters differ from the ones publicly available ($Public_b$ and $Public_c$) [http://www.exoplanet.eu/star.php?st=HD+187123]. For instance, the minimized planetary mass of planet c differ by $\approx 20\%$ and the longitude of periastron of planet b (ω_b) is $\approx 83\%$, far from the theoretical value. This differences can be symptomatic of 1) insufficient RV data and/or 2) possible undiscovered bugs in the code, particularly in the LM minimizing script and/or 3) the planetary system is too difficult to recognize because it consists of 3 or more planets and it exhibits high eccentricities, and/or 4) the signal-to-noise ratio produced is high.

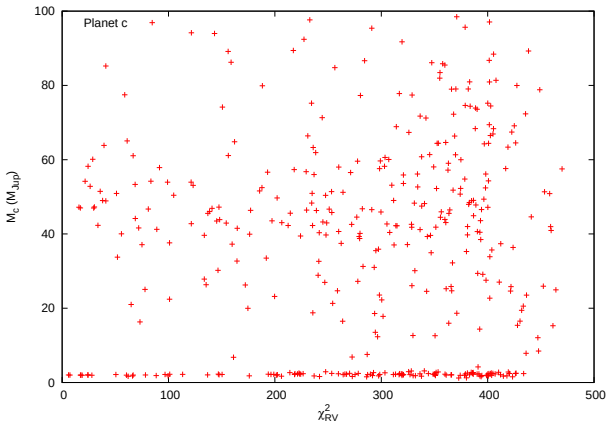
Another way to explore the χ^2_{RV} space and derive the orbital parameters' *best fit* is by generating many synthetic RV data sets (as shown in Table 6.1) that are similar in values

and number to the original ones. In the case of the system HD187123, 500 ($i = 0, ..., 499$) synthetic *.sys* files, each containing 3 RV sets and an average of 70 RVs per set, are created during a 10+ year time span. The *derivefit_levenberg.conscript* examines these 500 systems one at a time by minimizing the original planetary system according to the new RV sets. The script produces 500 $\chi^2_{RV,i}$ and corresponding minimized elements. When the planetary masses versus $\chi^2_{RV,i}$ are plotted as shown in Figure 6.1, it is possible to fit them to a normal or Gaussian distribution as shown in Figure 6.2.

In the case of planet b (Figure 6.1a), the amplitude obtained is $A_b = 0.468 \pm 0.064\ m_{\mathbb{4}}$, or $\pm 13.7\%$ (Figure 6.2a), fully contains the published solution $m_b \approx 0.52\ m_{\mathbb{4}}$ [http://www.exoplanet.eu/star.php?st=HD+187123], and gives a $\chi^2_{reduced} \approx 1$. The plot for planet c (Figure 6.1b) shows a very random distribution of masses versus least square fits. However, at the bottom of the plot, there is an area with a noticeable higher concentration of points. When zooming onto this area and, after completing the Gaussian fit (Figure 6.2b), the amplitude obtained is $A_c = 2.199 \pm 0.035\ m_{\mathbb{4}}$ which is $\sim 10\%$ away from the published value $Public_c = 1.99\ m_{\mathbb{4}}$. It is, although, a stretch to exclude $\sim 75\%$ of the minimized masses and this poor fit is indicative of insufficient RV data to confirm and characterize the existence of a second planetary mass. At this point, it is worthwhile to re-run the code with an increased number of generated RVs. With an average of 112 observations over a 10+ years time span, the results obtained have greatly improved (Figure 6.3). After 85 iterations and no data manipulation, the fit converged to an amplitude $A = 2.11 \pm 0.014\ m_{\mathbb{4}}$. This value is $\approx 7\%$ away from the theoretical value, $m_c = 1.99\ m_{\mathbb{4}}$, and confirms decisively the presence of planet c. However, the achieved result, so close to the theoretical value, is indicating that less RV data is necessary for confirmation. This next minimization step

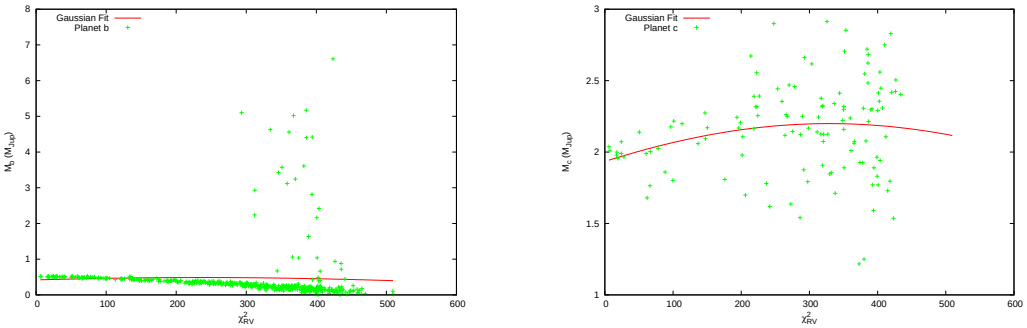


(a) Plot of m_b versus χ^2_{RV} accounting for all data points.



(b) Plot of m_c versus χ^2_{RV} accounting for all data points.

Figure 6.1: The two planetary masses, m_b and m_c are respectively plotted versus the produced χ^2_{RV}



(a) **(b)**

Figure 6.2: Gaussian distribution fitted to planet *b* (a), and planet *c*.

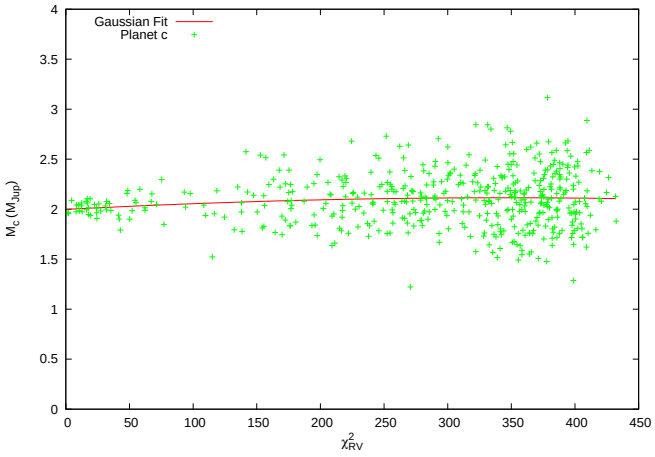


Figure 6.3: 500 Synthetic RV data sets, each containing an average of 112 data points, were minimized and then fitted to a Gaussian distribution.

will be implemented in a future version of the code: a small, synthetic RV data set will be generated first and then the number of RVs will be increased one at a time in order to achieve a result with 90% confidence.

6.2 Code implementation and confirmation of planets discovered by TESS

TESS will measure with great precision planetary periods and radii and because of the accuracy of the period measurements, this particular parameter will not need to be minimized. Masses, however, are the key parameters for dynamical studies, and will need to be minimized as they will be approximated by the power law equation obtained in Section 5.1. Equation 5.1 is derived by sampling the planets' masses in the Solar system as this is the only model known with precise measurements and describes the best fit considering Earth and Saturn solely. With the understanding that this fit is not optimal and does not reflect many of the characteristics of the observed extra-solar systems (e.g. hot Jupiters), it is however a concrete starting point for a planetary mass theory.

The first step consists in implementing the *OriginalFitGenerator.conscript* to obtain an *Original.fit* as shown in Figure 6.4a. This file includes information about RV datafiles, *Init_Vels.vels* and *Init_Vels.sys* (currently empty), stellar mass, orbital periodicities obtained by TESS, and the corresponding planetary masses obtained by using the planetary radii and Equation 5.1. Next, the *RV_Generator_from_FitFile.conscript* generates new RV data to replace the *Init_Vels.vels* file and the *Init_Vels.sys* file, taking into account Keplerian orbits and a standard deviation of $\sigma = 2.5$ m/s. In a normal distribution this standard deviation corresponds to 68% of the observations constrained within 2.5 m/s from the mean and $\approx$ 95% of the RVs constrained within 5 m/s from the mean. The new RV data sets (if the observing sites are more than one) are loaded onto the next script together with the *Original.fit* file. With the exception of periodicities, the orbital parameters

```
# Initial Epoch: 2455888.6069
#start

Parent "datafiles/Init_Vels.sys"
ID 831
Components 2
PrimaryRVSet "Init_Vels.vels"
OverallRVOffset 0.0
RelativeRVOffset {
  0.0
}
"0" {
 Period 131.40
 Mass 0.09
 MeanAnomaly 0.00
 Eccentricity 0.00
 LongOfPeriastron 0.00
 Inclination 0.0
 Node 0.0
 Radius 0.0
}
"1" {
 Period 13.82
 Mass 0.03
 MeanAnomaly 0.00
 Eccentricity 0.00
 LongOfPeriastron 0.00
 Inclination 0.0
 Node 0.0
 Radius 0.0
}

Properties {
}

Notes {

}

#end
```

(a)

```
# InitialEpoch: 2455896.523542
#start

Parent  "datafiles/System831.sys"
ID  831
Components  2
PrimaryRVSet  "System831_0.vels"
OverallRVOffset  0.026499667111960577
RelativeRVOffsets  {
  "System831_1.vels"  -0.13742123708587836
  "System831_2.vels"  -0.30438982130568176
  "Trend"  0.0
}

"1"  {
  Period  131.4
  Mass  0.08227021633218997
  MeanAnomaly  1.8762887352476951
  Eccentricity  0.0
  LongOfPericenter  0.0
  Inclination  0.0
  Node  0.0
  Radius  0.0
}

"0"  {
  Period  13.82
  Mass  0.02811874649848838
  MeanAnomaly  347.85876086107373
  Eccentricity  0.0
  LongOfPericenter  0.0
  Inclination  0.0
  Node  0.0
  Radius  0.0
}

Notes {
null
null
}


#end
```

(b)

Figure 6.4: 831 planetary system's *.fit* file outputs as loaded from the TESS file (a) and then loaded again after the LM minimization routine (b)

are minimized through the LM routine, and a new *Minimized.fit* file gets created (Figure 6.4b)[1]. The latter still displays the same number of planets as the original system, but the LM routine has found a best fit based on the new RV data. The comparison of the original planetary masses and the minimized ones, can help determine how a *best fit* is dependent on the number n of collected RVs.

As an example, the synthetic system 831 exhibits 2 planets with masses $m_{b,init} = 0.09\ m_{\updownarrow}$ and $m_{c,init} = 0.03\ m_{\updownarrow}$. After loading an average of 126 RVs per telescope, the

[1]The absence of eccentricities in the minimized system is a common feature to all double planet systems that the code produces and the nature of this outcome still needs to be investigated. One hypothesis could be that small planetary systems governed by Keplerian orbits favor circular fits.

LM routine minimizes the given fit and produces a new fit containing two planets with masses $m_{b,min} = 0.082\ m_{\oplus}$ and $m_{c,min} = 0.028\ m_{\oplus}$. Since $m_{b,min} = 91.1\%\ m_{b,init}$ and $m_{c,min} = 93\%\ m_{c,init}$ are sufficiently close to the original values, the orbital parameters predicted by TESS can be confirmed. Thus, the minimized fit gave insights regarding RV offsets and planetary mean anomalies and it established the number, n, of RVs that were collected by a specified ground-based telescope to corroborate TESS' guesses.

```
# Initial Epoch: 2455889.6069
#start

Parent "datafiles/Init_Vels.sys"
ID 935
Components 3
PrimaryRVSet "Init_Vels.vels"
OverallRVOffset 0.0
RelativeRVOffset {
  0.0
}
"0" {
 Period 191.60
 Mass 0.28
 MeanAnomaly 0.00
 Eccentricity 0.00
 LongOfPeriastron 0.00
 Inclination 0.0
 Node 0.0
 Radius 0.0
}
"1" {
 Period 58.19
 Mass 0.05
 MeanAnomaly 0.00
 Eccentricity 0.00
 LongOfPeriastron 0.00
 Inclination 0.0
 Node 0.0
 Radius 0.0
}
"2" {
 Period 259.90
 Mass 0.16
 MeanAnomaly 0.00
 Eccentricity 0.00
 LongOfPeriastron 0.00
 Inclination 0.0
 Node 0.0
 Radius 0.0
}

Properties {
}

Notes {

}

#end
```

(a)

```
# InitialEpoch: 2455896.440208
#start

Parent  "datafiles/System935.sys"
ID  935
Components  3
PrimaryRVSet  "System935_0.vels"
OverallRVOffset  -0.11444453372555057
RelativeRVOffsets  {
  "System935_1.vels"  -0.4704367088956301
  "System935_2.vels"  0.3225857596157749
  "Trend"  0.0
}

"1"  {
  Period  191.6
  Mass  0.3211197724335909
  MeanAnomaly  357.7410893960517
  Eccentricity  0.021025934180228908
  LongOfPericenter  0.04320392197227427
  Inclination  0.0
  Node  0.0
  Radius  0.0
}

"0"  {
  Period  58.19
  Mass  0.06453420058917264
  MeanAnomaly  1.4317312290183812
  Eccentricity  0.0811487717574434
  LongOfPericenter  0.024041772455819132
  Inclination  0.0
  Node  0.0
  Radius  0.0
}

"2"  {
  Period  259.9
  Mass  0.18337400060819467
  MeanAnomaly  353.6925246090658
  Eccentricity  1.6007034014123953E-5
  LongOfPericenter  0.12343372414931082
  Inclination  0.0
  Node  0.0
  Radius  0.0
}

Notes {

}

#end
```

(b)

```
# InitialEpoch: 2455897.429792
#start

Parent  "datafiles/935Synth.sys"
ID  935
Components  3
PrimaryRVSet  "935_tel0.vels"
OverallRVOffset  -0.4724482490675283
RelativeRVOffsets  {
  "935_tel1.vels"  0.5074485058476147
  "935_tel2.vels"  1.024581270909027
  "Trend"  0.0
}

"1"  {
  Period  191.6
  Mass  0.3080349123105966
  MeanAnomaly  3.3940419653512137
  Eccentricity  0.013787048472365972
  LongOfPericenter  0.21241119645038756
  Inclination  0.0
  Node  0.0
  Radius  0.0
}

"0"  {
  Period  58.19
  Mass  0.04783391326205045
  MeanAnomaly  354.8018902082518
  Eccentricity  1.3474294721570526E-4
  LongOfPericenter  0.08778667875365231
  Inclination  0.0
  Node  0.0
  Radius  0.0
}

"2"  {
  Period  259.9
  Mass  0.18952502095107213
  MeanAnomaly  4.309485113066635
  Eccentricity  0.039058449425547065
  LongOfPericenter  0.04602794905125343
  Inclination  0.0
  Node  0.0
  Radius  0.0
}

Notes {

}
```

(c)

Figure 6.5: 935 planetary system's *.fit* file outputs as loaded from the TESS file (a), loaded again after the LM minimization routine (b), and loaded one final time after LM routine was performed on an increased number of RV data

Another example is given by the system 935, displaying 3 orbiting planets, as

System 935	$m_b(m_\varphi$	$m_c(m_\varphi$	$m_d(m_\varphi$	n
Original	0.28	0.05	0.16	0
Minimized$_1$	0.32	0.065	0.18	263
(min/orig)$_1$	87.5%	77%	89%	
Minimized$_2$	0.30	0.048	0.19	420
(min/orig)$_2$	93%	104%	84%	

Table 6.3

shown in Figure 6.5. Figure 6.5a displays its found orbital masses and periods just as Figure 6.4a did for system 831. However, this time, when the new *Minimized.fit* file is obtained, the planetary masses are not within 10% of the originals. Table 6.3 shows the orbital values obtained according the number of RVs collected.

When looking at the mass ratios of the System 935, the first time the orbital elements are minimized, the new masses fall within 77-89% of the original masses. According to the TESS mission standards, this difference is too large to decisively confirm the presence of the identified planets. We find, however, that, by increasing the average number of collected RVs per telescope by 60% (from $n_{ave} = 88$ to $n_{ave} = 140$), better mass estimates are obtained. During the second trial, the first two planetary masses of System 935 are confirmed, while the last one still needs more RV data to ensure confirmation. As a side note, although the implemented LM routine considers planets governed by Keplerian orbits, a system composed of three or more perturbing bodies cannot be described by such physical process. The mutual gravitational force of the objects creates different dynamics, often with catastrophic effects that the current LM minimizer does not take into account.

In summary, according to TESS established standards, if a minimized planetary mass is well within 10% of the original one, as in the case of System 831, a lower number of ground-based, follow up observations could be performed without putting its confirmation

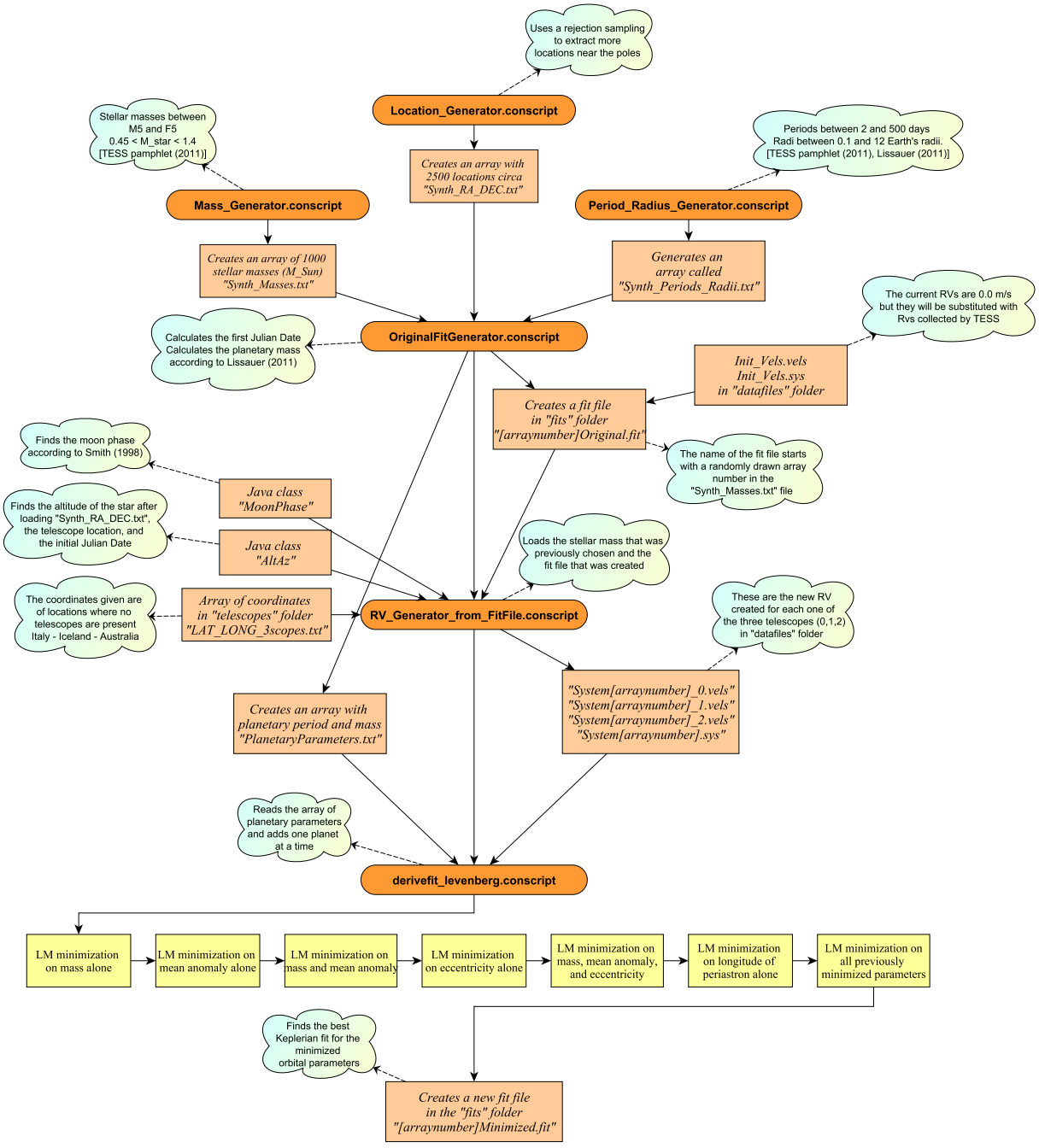


Figure 6.6: This flowchart is a map of the sequential steps needed to minimize RV data and analyze and confirm the planetary systems observed by TESS

at risk. Otherwise, if a minimized planetary mass is beyond 10% of the observed one, as in the case of System 935, more observations must be carried out. Through a balanced

plan, more corroborating efforts can be spent on the latter case (and less on the former) to confirm the presence (or absence) of a specific planetary system.

The flowchart displayed in Figure 6.6 summarizes the multiple steps necessary to obtain improved and minimized planetary systems. To notice in Figure 6.5c and 6.5b, is the appearance of eccentricities in the *[system name]Minimized.fit* files. The actual dynamical stability of planetary systems depends greatly on the eccentricities and mutual inclination, none of which can be measured well from transits and RV data alone. The most stable planetary systems are those with circular coplanar orbits (which minimizes the momentum deficit) [Lissauer et al. (2011)]. Thus, it is important to investigate possible high eccentricities generated by the code because the dynamical stability of the system can be forever jeopardized.

6.3 Minimization of the time frame and estimate of best locations

There is an unavoidable truth to the saying "Theres no such a thing as a free lunch." A data set that includes hundreds if not thousands of RVs would permit near-perfect measurements of orbital parameters. Three problems arise, however. First, the Time Allocation Committees do not favor long and drawn out data collection time spans. Second, RV detections are expensive in terms of real dollars, telescope time, and third, the competition between the various, planet-hunting teams is intense, and improved "guessing" techniques are highly sought after. If the given initial orbital parameters are good guesses compared to the actual present ones, then it seems that the *Refitting to Synthetic data*

method might be one of the most efficient at finding out how many radial velocities should be collected, via ground observations, before confirming a planetary system.

Another factor that can greatly improve the efficiency of follow up observations is the selection of observing sites. Given that an optimal location should comprise all the characteristics of already existing sites such as altitude and minimal light pollution, there could be the possibility of erecting new telescopes. A newly built site could cost less than $6,000,000 including hardware and labor, fairly inexpensive in comparison to the expenses incurred to build HIRES at Keck: without including labor, the hardware was estimated to cost $4,000,000. With a hypothetical $600,000,000 budget, the TESS mission could easily consider new locations. This option would ease the competition for observing time and shorten the follow up time frame.

7

Conclusions

TESS will try to take a step towards solving a major outstanding question in planet formation: how common are planetary schemes with architectures similar to the Solar system? In the known planetary systems within 200 pc, the average $m \sin i \approx 1.6\ m_{\oplus}$, and the average orbital distance from the host star is 0.9 AU. Given all the observational biases and technical constraints, these values suggest that the Solar system may well prove to be typical. However, the Solar system is currently considered atypical due to the nearly circular orbits of its planets (which also give it long term stability) versus the average exoplanetary eccentricity, $e_{ave} = 0.25$. An important improvement to planetary architecture theory, will be brought by a more efficient planetary mass theory, consistently fitting the observed data. Furthermore, the code demonstrates that while detections of short period planets can be rapidly corroborated, planets with long orbital periods will require observations spanning years if not decades to obtain reasonable results. It is also more difficult to obtain precise orbital elements for planets with large orbital periods because, while planets with short periodicities can be routinely observed for multiple cycles, long period candidates can often

be observed only for one period or less. Fitting multiple planet systems requires many more free parameters, so that observational data may not precisely constrain all the orbital elements or even distinguish between degeneracies (multiple possible orbital solutions).

Bibliography

[Borucki et al. (2009)] Borucky, W. J. et al., 2009, **KEPLER: Search for Earth-Size Planets in the Habitable Zone**. IAU Symposium, 253, 289-99.

[(Borucki et al. (2010)] Borucki, W. J. et al., 2010, **Kepler Planet-Detection Mission: Introduction and First Results**, Bulletin of the American Astronomical Society, 42, 1052.

[Brown & Latham (2008)] Brown, T. M., and Latham, D. W., 2008, **Expected Planet and False Positive Detection Rates for the Transiting Exoplanet Survey Satellite**, arXiv:0812.1305v1.

[Butler et al. (1996)] Butler, R.P. et al., 1996, **Attaining Doppler precision of 3 ms^{-1}**, Astronomical Society of the Pacific, 108, 500-509.

[Butler et al. (1998)] Butler, R.P. et al., 1998, **A Planet with a 3.1 Day Period around a Solar Twin**, Astronomical Society of the Pacific, 110, 1389-1393

[Charbonneau et al. (2007)] Charbonneau, D. et al., 2007, **When Extrasolar Planets Transit Their Parent Stars**, Protostars and Planets V, 701-16.

[Cumming et al. (2008)] Cumming, A. et al., 2008, **The Keck Planet Search: De-**

tectability and the Minimum Mass and Orbital Period Distribution of Extrasolar Planets, ASP, 120, 531-554.

[Duffett-Smith (1988)] Duffet-Smith, P., 1988, **Practical Astronomy with Your Calculator, 3rd ed.**, Cambridge University Press, 86-88, 142-144, 147.

[Ford (2005)] Ford, E. B., 2005, **Quantifying the uncertainty in the orbits of extrasolar planets**, AJ, 129, 1706-1717.

[Herschel & Dreyer (1912)] Herschel, W., and Dreyer, J.L.E., 1912, **The Scientific Papers of Sir William Herschel**, London, Royal Society and the Royal astronomical Society, 58-90.

[Kürster et al. (1994)] Kürster, M. et al., 1994, **A radial velocity search for extra-solar planets using an iodine gas absorption cell at the CAT+CES**, The Messenger 76, 51-55.

[Latham (2009)] Latham, D.W., 2009, **Towards Earth Like Planets**, ASP Conference Series, 420, 297-302.

[Lissauer et al. (2011)] Lissauer, J.J. et al., 2011, **Architecture and Dynamics of Kepler's Candidate Multiple Transiting Planet Systems**, arXiv:1102.0543v4.

[Lyons (1991)] Lyons, L., 1999 **A Practical Guide to Data Analysis for Physical Science Students**, Cambridge University Press, 58-61, 89-91.

[Mayor & Queloz (2005)] Mayor, M., and Queloz, D., 1995, **A Jupiter-mass companion to a solar-type star**, Nature, 378, 355-359.

[Meschiari et al. (2009)] Meschiari, S. et al., 2009, **Systemic: a testbed for characterizing the detection of extrasolar planets. I. The systemic console package**, arXiv:0907.1675v1.

[Meschiari et al. (2010)] Meschiari, S., et al., 2010, **The Lick-Carnegie Survey: Four New Exoplanet Candidates**, arXiv:1011.4068v2.

[Meschiari & Laughlin (2010)] Meschiari, S., and Laughlin G., 2010, **Systemic: a testbed for characterizing the detection of extrasolar planets. II. Numerical approaches to the transit timing inverse problem**, arXiv:1005.5396v1.

[Pal (2009)] Pal, A., 2009, **An analytical solution for Kepler's problem**, MNRAS, 396, 1737-42.

[Press et al. (2007)] Press, W.H. et al., 2007, **Numerical Recipes, the art of scientific computing, Third Edition**, Cambridge University Press, 549-555, 824-836.

[Ricker (2011)] Ricker, G. R., 2011, **TESS, Transit Exoplanet Survey Satellite. Explorer Proposal**, Cambridge, Massachussets: MIT.

[Ricker et al. (2009)] Ricker, G. R. et al., 2009, **Transiting Exoplanet Survey Satellite (TESS)**, American Astronomical Society Meeting Abstracts, 214, 306.

[Robb R.M. et al. (1999)] Robb, R. M. et al., 1999, **PHOTOMETRY OF THE HD 187123 FIELD**, IBVS 4820, 1.

[Torres et al. (2010)] Torres, G. et al., 2010, **Modeling Kepler transit light curves as false positives: rejection of blend scenarios for KOI-377, and strong evidence for a super Earth-size planet in a multiple system**, arXiv:1008.4393v2.

[Wright et al. (2007)] Wright, J.T. et al., 2007, **Four New Exoplanets and Hints of Additional Substellar Companions to Exoplanet Host Stars**, APJ 657, 533-545.

[Wright & Howard (2009)] Wright, J.T. & A.W. Howard., 2009, **Efficient Fitting of Multiplanet Keplerian Models to Radial and Astrometry Data**, APJ Supplement Series 182, 205-15.

[Wright et al. (2009)] Wright, J.T. et al., 2009, **Ten new and updated multiplanet systems and a survey of exoplanetary systems**, APJ 693, 1084-1099.

[http://en.wikipedia.org/wiki/Julian_day] **Julian Day**, Wikipedia, Wikimedia Foundation Inc., Web, 2011, available at http://en.wikipedia.org/wiki/Julian_day.

[http://en.wikipedia.org/wiki/Otto_Struve] **Otto Struve**, Wikipedia, Wikimedia Foundation Inc., Web, 2011, available at http://en.wikipedia.org/wiki/Otto_Struve.

[http://www.exoplanet.eu/star.php?st=HD+187123] **Star: HD 187123**, Extrasolar Planets Encyclopedia, Jean Schneider, Web, 2011, avaiable at http://www.exoplanet.eu/star.php?st=HD+187123.

[http://www2.keck.hawaii.edu/observing/schedule/index.php] **Telescope Schedule**, W. M. Keck Observatory, Web, 2011, available at http://www2.keck.hawaii.edu/observing/schedule/index.php.