

SEARCH FOR EXOPLANETS

submitted to the University of Delhi for PH-EL 564

Author

Yogita Kumari(21026762068)

M.Sc Physics (F)

Under the guidance of

Prof. HP Singh

Prof. TR Seshadri



Department of Physics and Astrophysics
University of Delhi-110007

Acknowledgement

Prof.HP Singh has graciously supported us by providing valuable guidance on how to solve problems. He gave us a perspective to look at things . We would like to express our sincere gratitude to him. We also want to acknowledge the outstanding lectures by Prof.TR Seshadri which help us build understanding concept of transit and radial velocity. Additionally, we would like to extend our thanks to our fellow M.Sc students for their invaluable discussions. Their support and encouragement have been instrumental in shaping our understanding of the field of exoplanet research.

Abstract

Search for exoplanets is manifestation of human desire to find life like us beyond earth. This project is all about methods of finding exoplanets using different method. We have studied Kepler photometric data for stars Kepler-8 ,Kepler-7,Kepler-6,Kepler-5 and WASP-28 with magnitude of order 12-14 and make conclusions regarding planets orbiting them. All of these planets are hot jupiters.Kepler-8b ,Kepler-7b,Kepler-6b and Kepler-5b are first 4 planets to be discovered by kepler mission and WASP28-b is first to be discovered by k2 mission. We have studied radial velocity data for HD 31253 ,HD 218566,HD 177830, KIC 6922244 and KIC 11853905.

Contents

1 Exoplanets	4
1.1 Introduction	4
1.1.1 Type	4
1.1.2 Method of Detection	5
1.2 Time Series	9
1.3 Periodogram	11
1.3.1 Lomb Scargle Method	12
1.3.2 Box Least Squares (BLS) Periodogram	14
1.3.3 BLS algorithm in lightkurve	15
1.4 Barycentric Time	16
1.5 Long cadence and short cadence	16
1.6 Linear Least Square Fitting	17
2 Transit Method	18
2.1 Sources of Data:	18
2.2 Lightkurve	23
2.3 Characterizing the transit:	25
2.3.1 Kepler-8b	25
2.3.2 Kepler-7b	27
2.3.3 Kepler-6b	27
2.3.4 Kepler-5b	30
2.3.5 WASP28-b	30
2.4 Results	35
3 Radial Velocity Method	36
3.1 Introduction	36
3.1.1 Lomb-Scargle method	36
3.1.2 False Alarm method	37
3.1.3 Period04	38
3.2 HD 31253	39
3.3 HD 218566	41
3.4 KIC 11853905	43
3.5 KIC 6922244	45
3.6 HD 177830	47
Summary and Conclusion	50
A Radial Velocity Data	51
Bibliography	57

Chapter 1

Exoplanets

1.1 Introduction

The very first question that arises in our mind is to what are exoplanets. We know that there exist billions of galaxies in the observable universe and each single one has billions of stars in it. Mankind has been intrigued by the possibilities of existence of planetary systems other than our own. This zest has provided us the database today we have on these extrasolar planets known as exoplanets i.e. exoplanets are planets which are outside our solar system. Generally these planets are also orbiting stars but certain free-floating exoplanets called as *Rogue Planets* also exists which orbit the galactic center. Most of the exoplanets are discovered in small region of our Milky Way galaxy by missions like Kepler, K2, TESS etc. Until date around 5241 exoplanets have been confirmed so far with the very first one being 51 Pegasi b in 1995 and the latest one LHS 475b in 2023 by James Webb Telescope.

There exists a wide range of diversity in exoplanets. Their orbital properties vary widely. Many have very elliptical orbits compared to the largest eccentricities in our own solar system. Based on estimated masses many are around that of Jupiter($0.3\text{-}3M_j$). Some are close in orbit to their host star than Mercury and some are located far out at distances of 100 au or more. These exoplanets are being discovered around a wide variety of stellar types. Host stars with varying internal structure and composition such as ranging from main sequence stars like Sun to very low mass, to white dwarfs. Exoplanets atmospheres are also been probed through secondary eclipse photometry and spectroscopy. There even exists multi-planetary systems with more than one planet orbitting the star.

Since these exoplanets are accompanied by host stars and shined by the reflected starlight, they will be very much fainter than their host stars and direct detection is extraordinarily demanding. In this work we will be discussing the theory regarding different types of exoplanets, host stars and all the detection methods formulated till now with special emphasis on the transit method and radial velocity. We will discuss how different parameters of an exoplanet can be found geometrically and through analysis of the light curves from transit method and spectroscopic data through radial velocity measurements. We will further provide an insight into exoplanets atmosphere and habitability.

1.1.1 Type

These planets exists in different sizes and constituents such as from gas giants like Jupiter even bigger in size to small and rocky planets just like Earth and Mars. These are generally categorised into four types :-

1. Gas Giants : Giant planets composed mainly of gas. They are generally of size of Saturn or Jupiters. They include *Hot Jupiters* which are loosely defined as jupiter mass planets with

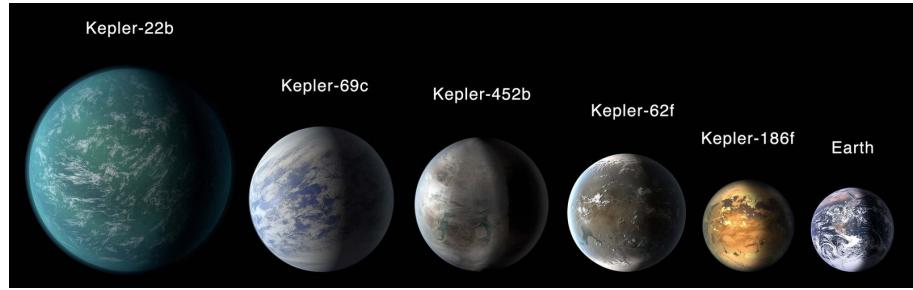


Figure 1.1: Some exoplanets compared to Earth(www.gps.caltech.edu)

$a \leq 0.1$ au or $P=3\text{-}9$ days ,planets in close orbit around their stars hence temperature very high. There are also very hot jupiters($P \leq 3$ days),ultra-short-period hot jupiters($P \leq 1$ day),and warm jupiters($P \geq 10$ days).Example is the very first exoplanet discovered 51Peg b.

2. Super-Earth : Rocky planets larger than Earth,generally in mass range of 2-10 times mass of Earth.They might or might not have atmospheres.Kepler has shown that super-Earths are amongst the most common planet type around nearby Sun-like stars. There are two other class also i.e.'lava planets' with surface mostly covered by molten lava and 'ocean planet' with significant mass fraction in H_2O . For e.g. Kepler-186 f.
3. Neptune-like : Gaseous Planets around the size of Neptune and Uranus with atmosphere mainly composing of hydrogen and helium.For e.g. OGLE-2005-BLG-390L b.
4. Terrestrial : Planets mostly made up of silicate rocks and metals and of mass order of one to few times mass of Earth.They have a solid surface and may have basic type of structure like comprising a central metallic coreSome might have possibility of oceans and atmospheres and perhaps signs of habitability. For e.g. TRAPPIST-1 e,two Earth sized planets of Kepler-20,two in the Kepler-70 system.

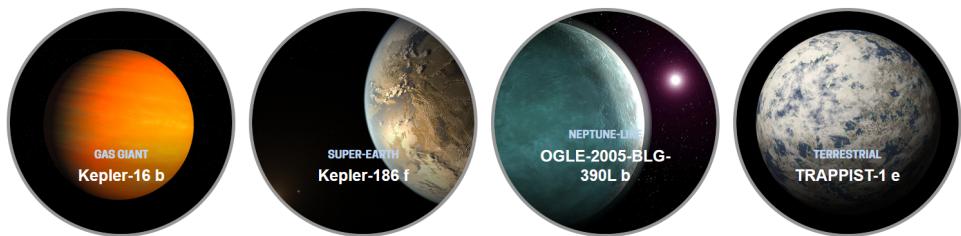


Figure 1.2: Visualisation of some examples of different types of exoplanets (exoplanet.nasa.gov)

1.1.2 Method of Detection

To detect these exoplanets 5 methods are used which are discussed further.

1. **Radial Velocity:** We know that there exists a gravitational interaction between the planets and the stars which we can see in the figure below. Since the mass of star is way much more than that of star we see the planets revolving around the star not the other way around.But star also experience the gravitational effect of planet which also cause star to wobble. But how do we see the wobble? It is done using the Doppler shift.We know that light emitted from the star is electromagnetic waves so the waves as object

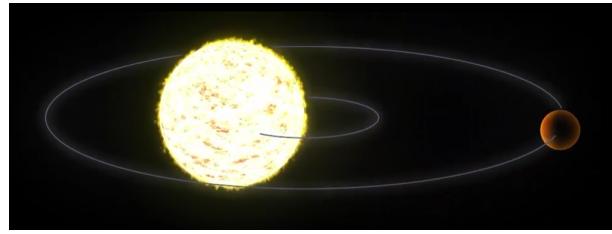


Figure 1.3: Wobbling of star(exoplanets.nasa.gov)

moves towards or away from the observer tends to bunch up or stretch out. So the visible light crunch together ,they look more blue in color and when the visible light stretch out ,they make an object look more reddish. This change in color is called as 'redshift'. The radial velocity method is one of the most successful ways to find exoplanets with two notable observatories where this work happens are the Keck Telescopes in Hawaii and the La Silla Observatory in Chile. The wobbling of star is its motion around the star-planet barycentre (centre of mass). To get further into this method first we will go through some basics prerequisites.

Kepler Laws: Kepler's three laws of planetary motion which are

- (a) Motion of a planet around the Sun is traces an ellipse with Sun at its one of foci.
- (b) The areal velocity of the planet is a constant i.e. the line joining a planet and Sun sweeps out equal areas in equal intervals of time.

$$\frac{dA}{dt} = \frac{J}{2m} \quad (1.1)$$

where A is the area swept,J is the total angular momentum of planet with mass m.

- (c) The square of the orbital period of a planet is directly proportional to the cube of length of semi-major axis.

$$P^2 = \frac{4\pi^2 a^3}{GM} \quad (1.2)$$

where P is the Orbital time period,a is the semi-major axis,G is Universal gravitational constant and M is the mass of Sun.

Another important things to understand are the various angles in the orbit plane which are referred as 'anomalies'. *True anomaly* $v(t)$ is the angle between the direction of pericentre and the current position of the body measured from the barycentric focus of the ellipse. It is the angle normally used to characterise an observational orbit. *Eccentric anomaly* $E(t)$ is a corresponding angle which is referred to the auxiliary circle of the ellipse. The true and eccentric anomalies are geometrically related. *Mean anomaly* $M(t)$ is an angle related to a fictitious mean motion around the orbit, used in calculating the true anomaly. The mean anomaly at time $t-t_p$ after pericentre passage is then defined as

$$M(t) = \frac{2\pi(t - t_p)}{P} \quad (1.3)$$

The relation between these quantities is

$$M(t) = E(t) - e \sin E(t) \quad (1.4)$$

$$\cos v(t) = \frac{\cos E(t) - e}{1 - e \cos E(t)} \quad (1.5)$$

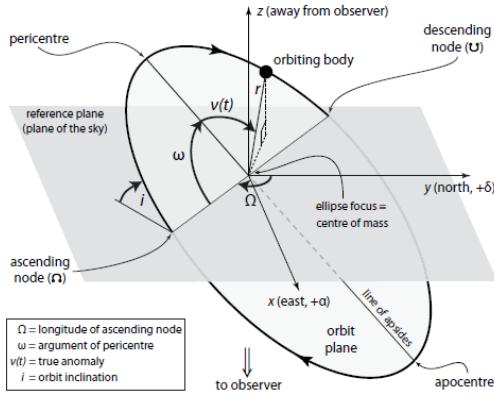


Figure 1.4: depiction of all discussed quantities(The Exoplanet Handbook)

The three angles (i, Ω, ω) represent the projection of the true orbit into the observed (apparent) orbit. All these quantities can be visualised in the diagram below.

Radial velocity semi-amplitude: We obtain this quantity after applying the third kepler law and applying some geometric identities. For a circular orbit with $M_p \ll M_*$

$$K = 28.4 \left(\frac{P}{1\text{yr}} \right)^{-\frac{1}{3}} \left(\frac{M_p \sin i}{M_j} \right) \left(\frac{M_*}{M_\Theta} \right)^{-\frac{2}{3}} \text{ms}^{-1} \quad (1.6)$$

Here M_Θ and M_j denotes the mass of Sun and mass of Jupiter respectively. M_* can be determined from spectral and luminosity class of the star. And from the measurement of K we can determine the $M_p \sin i$.

2. **Transit Method:** In this method when the planet passes directly between the observer and the star it is orbiting ,it blocks some of the star light for a time period .During that time the flux received from the star is dipped,so we see a drop in brightness in the associated light curve(a plot of flux with time for a star) of the star.So this change may be tiny but enough to confirm the presence of an exoplanet around a star. The duration for



Figure 1.5: Transit method(exoplanets.nasa.gov)

which the drop in flux happens and the amount can help us to find out the size of planet ,its orbit radius and moreover the blocked starlight passing through the atmosphere of the star can help us to find out its composition. We can obtain the following parameters from the light curve and stellar parameters provided in database.

- **Planet's Radius (R_p):** It can be measured by measuring the amount of dip in the light curve.

$$\frac{\Delta f}{f} \approx \left(\frac{R_p}{R_*} \right)^2 \quad (1.7)$$

- **Planet's orbital period (P):** It can be simply determined by measuring the time passed between successive dips in the light curve.
- **Orbital semi-major axis (a):** It is simple to measure for transiting planets with circular orbits. We have to simply equate the gravitational force between the star and planet to the centrifugal force on the planet. Then we can find velocity of planet in terms of semi major axis 'a' and orbital period P. It will give us the expression which is nothing but the expression of Kepler's third law.

$$a = \left(\frac{GM_*}{4\pi^2} \right)^{\frac{1}{3}} P^{\frac{2}{3}} \quad (1.8)$$

- **Transit impact parameter (bR_*):** It is a measure of closeness to the centre of stellar disc the planet passes at the midpoint of transit. It is measured in units of R_* . Its expression is found to be

$$bR_* = \left[(R_* + R_p)^2 - a^2 \left(\sin \frac{\pi T}{P} \right)^2 \right]^{1/2} \quad (1.9)$$

Here T is the duration of transit.

- **Inclination angle (i):** It is the angle between the orbital plane of planet and plane of elliptic. From geometry it is

$$i = \arccos \left(\frac{bR_*}{a} \right) \quad (1.10)$$

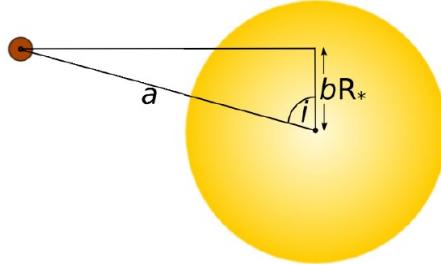


Figure 1.6: Transit parameter and inclination angle (Richard P. Nelson, Guillem Anglada-Escudé and Gavin Coleman)

3. **Direct Imaging:** The method is quite hard since the exoplanets are millions of times dimmer than the star they orbit and very far away. So the imaging is done generally using two methods to block the light of the star. One, called coronography, uses a device inside a telescope to block light from a star before it reaches the telescope's detector. Coronagraphs are built as internal add-ons to telescopes, and are now being used to directly image exoplanets from ground-based observatories. Another method is to use a 'starshade', a device that's positioned to block light from a star before it even enters a telescope. For a space-based telescope looking for exoplanets, a starshade would be a separate spacecraft, designed to position itself at just the right distance and angle to block starlight from the star astronomers were observing. (Reference:- exoplanet.nasa.gov)

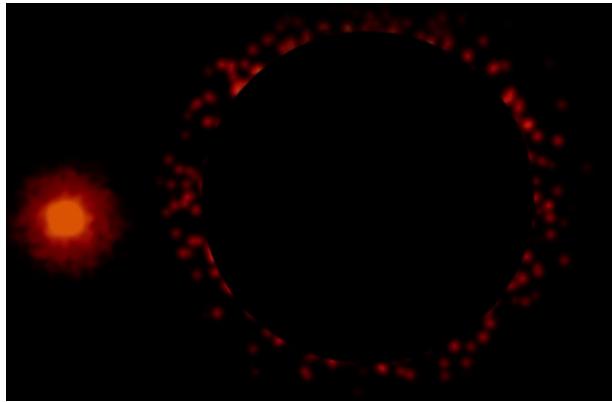


Figure 1.7: Direct Imaging by blocking light of star(exoplanets.nasa.gov)

4. **Gravitational Microlensing:** It happens when the gravity of a star or a planet focuses the light of another distant star in a way that makes it temporarily seem brighter. In the figure below, we can see the rays of light from the more distant star bend around the exoplanet and then the exoplanet's star. The graph on the left indicates the changing brightness of the distant star as its light is lensed and focused onto the observer. The star starts to get brighter, then there's a brief blip of brightness from the lensing action of the planet. The light levels fall after the planet is lensed but they continue to increase because of the continued lensing action of the star. Once the lensing star moves out of the optimum position, the brightness of the more distant star fades away.

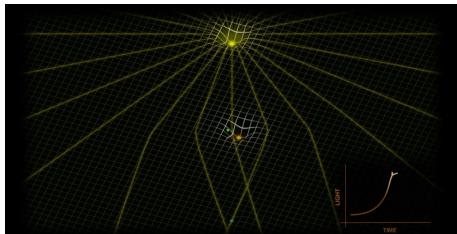


Figure 1.8: Gravitational Microlensing(exoplanets.nasa.gov)

5. **Astrometry:** It is also related to wobbling of star as discussed earlier. But in this method this wobbling is not measured using the Doppler shift but by changes in the apparent position of star with respect to another stars in the sky. Astrometry, as this method is called, is still amazingly hard to do. Stars wobble such a minute distance that it's very difficult to accurately detect the wobble from planets, especially small ones the size of Earth. In order to track the movement of these stars, scientists take a series of images of a star and some of the other stars that are near it in the sky. In each picture, they compare the distances between these reference stars and the star they're checking for exoplanets. If the target star has moved in relation to the other stars, astronomers can analyze that movement for signs of exoplanets. Astrometry requires extremely precise optics, and is especially hard to do from the Earth's surface because our atmosphere distorts and bends light.

1.2 Time Series

- **Time Series :** Time series can be define as collection of data taken at different point of time where all observations are taken independent and identically distributed. There

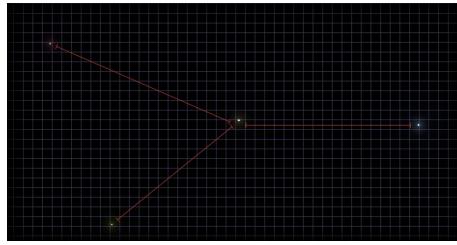


Figure 1.9: change in star's relative position(exoplanets.nasa.gov)

can be two approaches to analyse the time series- (1) Time Domain Approach and (2) Frequency Domain Approach. Lets try to explore both approaches-

1. Time Domain Approach : It works with preassumption that there is dependence of present value on past values. In this approach focus is on modelling of future value in terms of present and past values. It is used in forcasting.
2. Frequency Domain Approach : It works with pressumption that time series can be expressed as linear combination of periodic variations . These variation are usually caused by biological, physical or envirnoment phenomena.This approach is used when we are interested in knowing periodicities in series. It is also known as spectral analysis.

- **Exploratory Data Analysis(EDA) :**

It is a term coined by John W. Tukey. It is just used to describe informal way to have a look at data first to see what it looks like and then employing techniques .

First thing that can be done is to see whether time is stationary or non stationary. Stationary time series implies that series properties do not change with time while non stationary series properties change with time. Time series with trends or seasonality are non stationary. It is usually easier to deal with stationary series so autocovariace and mean need to be stationary at least.

Let try to see how can we work with trend stationary model which has stationary behavior around a trend. This can be given as follows-

$$x_t = \mu_t + y_t \quad (1.11)$$

where x_t are observations , μ_t is trend and y_t is stationary process. Strong trend can obscure the behavior of the stationary process so first step in exploratory analysis is to remove the trend . We need to first find the trend then subtract it from observations. Residual is given as follows-

$$\hat{y}_t = x_t - \hat{\mu}_t$$

This process is known as **detrending**.

Differencing can also be used to remove trend.In case if we have trend like equation (1) then differencing will give stationary process.

$$\begin{aligned} x_t - x_{t-1} &= \mu_t + y_t - (\mu_{t-1} + y_{t-1}) \\ &= \delta + w_t + y_t - y_{t-1} \end{aligned} \quad (1.12)$$

here $z_t = y_t - y_{t-1}$ is stationary since we are taking stationary behavior around trend. When we calculate covariance of z_t that will be independent of time . Another case could

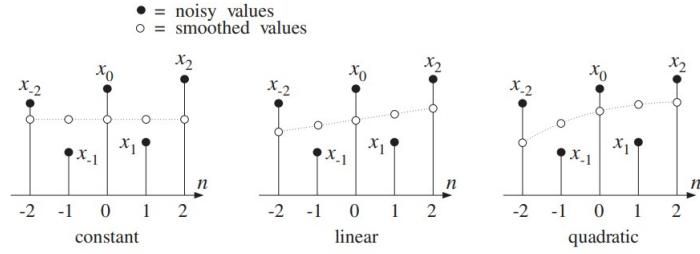


Figure 1.10: Figure adopted from Introduction to Signal Processing by SJ Orfanidis showing data smoothing via polynomial best fit

be fixed trend,i.e, $\mu = \beta_1 + \beta_2 t$ then differencing will give stationarity.

$$\begin{aligned}\nabla x_t &= x_t - x_{t-1} \\ &= \mu_t + y_t - (\mu_{t-1} + y_{t-1}) \\ &= \beta_2 + y_t - y_{t-1}\end{aligned}\tag{1.13}$$

Savitzky Golay Filter :It is way of smoothing of data to remove noise via taking small range of data points and doing polynomial fit using least square fit method.This can be seen in Figure 1.10.

- Fourier Transformation : Any stationary time series can be thought of as superposition of sine and cosine oscillating at various frequencies.

If data is taken at discrete time interval then to determine cycle we need at least two point so highest frequency of interest is 0.5 cycle per point.Folding frequency is define as highest frequency that can be seen in discrete sampling.here highest frequency implies that frequency which appear less frequently.

To know the important frequencies of time series we need to use a technique called periodogram. Time series can be expressed as follows

$$x_t = \sigma[\beta_1\left(\frac{j}{n}\right)\cos(2\pi\omega_j t) + \beta_2\left(\frac{j}{n}\right)\sin(2\pi\omega_j t)]$$

where $\beta_1\left(\frac{j}{n}\right)$ and $\beta_2\left(\frac{j}{n}\right)$ are regression parameters. We need to estimate these parameters. It can be done using method called Fast Fourier Transform. in that we need to detemine regression parameter-

$$P\left(\frac{j}{n}\right) = \beta_1^2\left(\frac{j}{n}\right) + \beta_2^2\left(\frac{j}{n}\right)$$

1.3 Periodogram

Periodogram analysis is a widely used method for analyzing time series data in various fields such as finance, economics, astronomy and engineering. It provides a graphical representation of the spectral density of a signal, which is useful for identifying the underlying frequencies and periodicities in the data.

One of the most common uses of the periodogram in astronomy is the detection of periodic signals in stellar light curves. For example, starspots, which are regions of a star's surface that are cooler and thus appear darker than the surrounding area, can cause periodic fluctuations in a star's brightness as they rotate in and out of view. By analyzing the periodicity of these brightness fluctuations using a periodogram, astronomers can determine the rotation period of

the star and study the distribution and behaviour of starspots on its surface.

Another important application of the periodogram in astronomy is the search for exoplanets. As we know, when an exoplanet transits, or passes in front of its host star, it causes a periodic dip in the star's brightness. By analyzing the periodicity of these brightness dips using a periodogram, the orbital period of the exoplanet and other properties such as its size and distance from its host star can be determined.

There are several algorithms used in periodogram analysis to analyze time-series data. Here are some of the most common ones:

a) Fast Fourier Transform (FFT):

This is a widely used algorithm for calculating the discrete Fourier transform of a time-series data set. It is computationally efficient and is especially useful when analyzing regularly spaced data.

b) Welch's method:

This algorithm is a modification of the FFT method that involves dividing the data into overlapping segments and calculating the periodogram for each segment. The resulting periodograms are then averaged to reduce noise and improve the estimate of the power spectral density.

c) Maximum Entropy Method (MEM):

This algorithm is a non-parametric method that uses a Bayesian approach to estimate the power spectral density. It is particularly useful for analyzing short or unevenly sampled time-series data.

d) Lomb-Scargle method:

This algorithm is based on fitting sinusoidal functions to the data at each frequency and calculating the power at each frequency. It is especially useful when analyzing unevenly sampled data or data with missing values.

e) Wavelet Transform:

This algorithm decomposes the time-series data into a set of wavelet functions of different scales and calculates the power spectral density at each scale. It is particularly useful for analyzing non-stationary time-series data.

f) Box Least Squares method (BLS):

This algorithm is used to search for the periodic dips in the brightness of a star caused by the passage of a planet in front of it. It works by dividing the time-series data into a series of boxes and fitting a model to each box to search for periodic signals. The algorithm then searches for a periodic signal that best matches the data by minimizing the sum of the squared residuals.

Overall, the choice of algorithm will depend on the characteristics of the time-series data being analyzed, such as the sampling rate, length of the time series, and presence of noise or missing data. It is important to carefully consider the strengths and weaknesses of each algorithm before choosing one for a particular analysis.

1.3.1 Lomb Scargle Method

The Lomb-Scargle algorithm is a method for estimating the power spectral density (PSD) of a signal from unevenly sampled time-series data. It is commonly used in signal processing and astronomy to analyze time-series data with irregular time intervals between measurements.

The algorithm was developed by N.R. Lomb and J.W. Scargle in 1976 and is a modification of the classical Fourier periodogram method. The Lomb-Scargle algorithm is particularly useful when dealing with data that has irregular sampling, where the traditional Fourier analysis

techniques may not be applicable.

The Lomb-Scargle periodogram is a measure of the power spectral density of a signal at a particular frequency. It can be computed for a discrete set of frequencies and provides a measure of the relative strength of the signal at each frequency. The periodogram is a normalized measure of the power spectral density, which is the squared magnitude of the Fourier transform of the signal. The Lomb-Scargle algorithm involves fitting a sinusoidal function to the data at each frequency, and then computing a measure of the goodness of fit. This measure is then used to compute the periodogram at each frequency. It provides a measure of the relative strength of the signal at each frequency and can be used to identify periodicities or other patterns in the data.

The Lomb-Scargle algorithm has several advantages over other methods for analyzing time-series data, including its ability to handle unevenly spaced data and its sensitivity to signals with low power or low signal-to-noise ratio. It is a powerful and widely used tool for analyzing time-series data in a variety of fields, including astronomy, signal processing, and finance.

The Nyquist Limit

The Nyquist limit theorem states that in order to accurately reconstruct a continuous-time signal from its sampled version, the sampling rate must be at least twice the highest frequency component of the signal. In other words, the Nyquist limit sets a lower bound on the sampling rate required to avoid aliasing, which occurs when high-frequency components in the original signal are misrepresented in the sampled version.

It can be also stated as when a band-limited signal, which , which contain the frequency component of the signal and whose fourier transform is zero outside the $\pm B$ range, must be sampled with a rate of atleast $f_{Ny} = 2B$.

Use of Lomb-Scargle to search for an exoplanet

The Lomb-Scargle algorithm can be used to search for exoplanets by analyzing the variations in the brightness of a star over time. When a planet orbits a star, it causes a periodic variation in the star's brightness, which can be detected through time-series photometric observations. It can be used to identify periodic signals in the brightness variations of a star. The periodogram will show a peak at the period of the planet's orbit, which can be used to determine the planet's period and therefore its distance from the star. The amplitude of the peak in the periodogram is related to the size of the planet, and the shape of the peak can reveal information about the planet's orbit and any additional planets in the system.

To use the Lomb-Scargle algorithm for exoplanet detection, astronomers typically collect high-precision photometric data of a star over an extended period of time, typically several weeks to months. They then process the data using the Lomb-Scargle algorithm to search for periodic signals in the star's brightness variations. The algorithm Is sensitive to small variations in the brightness of the star, making it a powerful tool for detecting small exoplanets, such as Earth-sized planets, that produce only subtle variations in the star's brightness.

Use of Lomb -Scargle in Lightkurve to search for an exoplanet

Lightkurve is a Python package that provides tools for analyzing time-series data from space-based telescopes, including Kepler, K2, and TESS. One of the key features of Lightkurve is its

ability to perform Lomb-Scargle periodogram analysis on light curves to search for exoplanets.

To use Lightkurve to search for exoplanets, we first need to download the time-series photometric data from the telescope archive. Once we have the data, we can use Lightkurve to perform a series of steps to search for exoplanets:

- Import the Lightkurve package and load the time-series data using the ‘`lightkurve.open()`’ method.
- Detrend the data to remove any instrumental or systematic effects using the ‘`flatten()`’ or ‘`remove_outliers()`’ method.
- Use the ‘`to_periodogram()`’ method to compute the Lomb-Scargle periodogram of the detrended light curve. This generates a periodogram that shows the power spectral density as a function of frequency.
- Use the ‘`periodogram.plot()`’ method to visualize the periodogram and identify any peaks that may correspond to the orbital period of an exoplanet.
- If a peak is identified, use the ‘`fold()`’ method to fold the data at the period of the peak and visualize the resulting folded light curve.

Analyze the folded light curve to determine if it exhibits the expected features of an exoplanet transit, such as a periodic dip in the brightness of the star.

By following these steps, we can use Lightkurve and the Lomb-Scargle algorithm to search for exoplanets in time-series photometric data from space-based telescopes. Lightkurve provides a user-friendly interface for performing these analyses and allows astronomers to quickly and easily search for exoplanets in large data sets.

1.3.2 Box Least Squares (BLS) Periodogram

The Box Least Squares (BLS) algorithm is a commonly used method for detecting periodic signals in time-series data, such as those obtained from astronomical observations or transit searches for exoplanets.

In the context of periodograms, the BLS algorithm involves dividing the time-series data into a series of evenly spaced bins, each with a certain width or duration. The algorithm then searches for periodic signals within each bin by fitting a model to the data using a least squares approach. Specifically, the BLS algorithm models the data as a series of transit-like events, each of which has a fixed duration and depth, and is separated by a certain period.

To search for periodic signals, the algorithm varies the period of the transit-like events and evaluates the goodness-of-fit of the resulting models using a least squares approach. The algorithm then identifies the period that produces the best fit to the data, and reports the corresponding transit depth and duration as well as the signal-to-noise ratio (SNR) of the detection.

Box Least Square (BLS) algorithm to search for an exoplanet

The Box Least Squares (BLS) algorithm is a method used to search for exoplanets by analyzing the periodic transit signals they produce as they orbit their host star. The BLS algorithm is used to identify periodic changes in the brightness of a star that may indicate the presence of an exoplanet passing in front of it.

The BLS algorithm works by dividing the observed light curve of a star into many small segments, and for each segment, it calculates a statistical measure of how well a box-shaped transit signal fits the data. The width and depth of the box-shaped signal are varied over a range of values to identify the best fit. The resulting periodogram shows peaks at the periods where the box-shaped signal best matches the data.

To search for an exoplanet using the BLS algorithm, astronomers typically obtain a series of high-precision observations of the target star using a ground-based or space-based telescope. The observations are then processed using the BLS algorithm to search for periodic transit signals that may indicate the presence of an exoplanet. If a candidate signal is identified, further observations and analysis are typically performed to confirm the presence of an exoplanet and to determine its properties, such as its size, mass, and orbit.

1.3.3 BLS algorithm in lightkurve

To use the BLS algorithm in Lightkurve, the first step is to download and preprocess the time-series data from the Kepler or TESS mission using the Lightkurve data access and target pixel file tools. Once the data is loaded, the BLS algorithm can be applied using the ‘Lightkurve.periodogram’ module.

The ‘Lightkurve.periodogram’ module allows the user to specify the range of periods to search for, as well as the number of bins in the periodogram. The module then computes the BLS periodogram and returns a power spectrum that indicates the strength of periodic signals at different periods.

The BLS algorithm in Lightkurve can be used to search for exoplanet transits by looking for periodic signals in the power spectrum that match the expected shape and duration of a planetary transit. Once a potential transit is identified, further analysis can be performed using statistical methods such as the transit fitting algorithm in Lightkurve to determine the properties of the exoplanet and its orbit.

BLS algorithm to determine the orbital properties of the exoplanet

In the case of transit photometry observations, the BLS algorithm provides the period, duration and depth of the transits, which can be used to estimate the orbital period, radius and inclination of the exoplanet, assuming a model for the transit geometry and the properties of the host star.

In the case of radial velocity observations, the BLS algorithm provides the period of the radial velocity variation, which can be used to estimate the orbital period of the exoplanet. The amplitude of the radial velocity variation provides a measure of the exoplanet’s mass, assuming a model for the orbital geometry and the properties of the host star.

1.4 Barycentric Time

Barycentric time, also known as Barycentric Julian Date (BJD), is a system of timekeeping that takes into account the motion of the Solar System. It is a standard for measuring time used in astronomy, particularly in the study of exoplanets.

The barycenter is the center of mass of two or more celestial bodies, in this case, the center of mass of the Earth and the Sun. Barycentric time is the time as measured from the barycenter of the Solar System, and is used to correct for the motion of the Earth in observations of celestial objects.

In the study of exoplanets, barycentric time is important because the motion of the Earth around the Sun can cause a periodic shift in the arrival time of light from a distant star. This shift can be mistaken for a variation in the orbital period of the exoplanet, and can lead to incorrect estimates of the exoplanet's properties. By converting observations to barycentric time, astronomers can correct for this effect and obtain more accurate estimates of the exoplanet's orbit and other properties.

Barycentric time is typically expressed as a Julian Date (JD) or Modified Julian Date (MJD), which are systems of timekeeping that count the number of days since a starting point. The starting point for JD is January 1, 4713 BCE, while the starting point for MJD is November 17, 1858 CE.

1.5 Long cadence and short cadence

Long cadence and short cadence are two different observing modes used in NASA's Kepler and TESS missions to search for exoplanets.

Long cadence refers to the standard observing mode in which Kepler and TESS observe a target star once every 30 minutes (for Kepler) or once every 2 minutes (for TESS). During each observation, the spacecraft measures the brightness of the star, looking for the small dips in brightness that can occur when a planet passes in front of the star, blocking a small fraction of its light. Long cadence observations are used to detect larger exoplanets with orbital periods of weeks to months.

Short cadence, on the other hand, is an observing mode in which Kepler and TESS observe a target star once every few seconds or minutes. This high-cadence observing mode is used to detect smaller exoplanets with orbital periods of hours to days. By taking many observations in quick succession, Kepler and TESS can detect the small, rapid dips in brightness that occur when a small exoplanet passes in front of its host star.

In short, long cadence is used to search for larger exoplanets with longer orbital periods, while short cadence is used to search for smaller exoplanets with shorter orbital periods. By using both observing modes, astronomers search for a wide range of exoplanets and gain a more complete understanding of the exoplanet population in our galaxy.

To analyze data obtained in long cadence mode using Lightkurve, we can use the ‘`lightkurve.search_lightcurve`’ function to search the Kepler or TESS archive for the target star and download the long cadence data. Once we have the data, we can use functions such as ‘`remove_outliers()`’ and ‘`flatten()`’ to clean and preprocess the data, and then use functions such as ‘`periodogram()`’ and ‘`box_least_squares()`’ to search for periodic signals and transit signals indicative of exoplanets.

To analyze data obtained in short cadence mode using Lightkurve, we can use the ‘`lightkurve.search_targetpixelfile()`’ function to search the Kepler or TESS archive for the target star and download the short cadence data. Once we have the data, we can use functions such as ‘`to_lightcurve()`’ to convert the target pixel data to a light curve, and then use the same functions as for long cadence data to analyze the light curve.

1.6 Linear Least Square Fitting

Linear least squares fitting is a statistical method used to determine the best-fit line through a set of data points. In this method, the goal is to find a linear equation of the form $y=mx+b$ that minimizes the sum of the squared residuals between the observed data points and the predicted values of the dependent variable (y) based on the independent variable (x).

To perform a linear least square fit, we calculate the slope (m) and the y-intercept (b) of the best-fit line using the following formulas:

$$m = \frac{nx\Sigma(x + y) - \sum x \sum y}{n\Sigma(x^2) - (\Sigma x)^2} \quad (1.14)$$

$$b = \frac{\sum y - m \sum x}{n} \quad (1.15)$$

where n is the number of values we have in the data set.

These equations give us the values of m and b that minimize the sum of squared differences between the observed y values and the predicted y values based on the linear equation. These values can then be used to plot the best-fit line and to make predictions or to interpolate between data points.

The accuracy of the linear least square fit can be assessed by calculating the coefficient of determination (R^2), which measures the proportion of variance in the dependent variable that is explained by independent variable.

The value of R^2 ranges between 0 and 1, with a value of 1 indicating a perfect fit of the data to the regression line. A value of 0 indicates that the regression line does not explain any of the variance in the data.

The formula for calculating R^2 is:

$$R^2 = 1 - \frac{RSS}{TSS}$$

Where RSS is sum of squares of residuals, which is the difference between the observed value of y and the predicted value of y based on the regression line, squared and summed over all data points and TSS is total sum of squares, which is the squares of differences between the observed values of y and the mean of y, summed over all data points.

Chapter 2

Transit Method

2.1 Sources of Data:

Kepler Mission: Kepler Space telescope,moving in earth-trailing heliocentric orbit with period of approx 372.5 days, was design to find earth-sized planets in habitable zone via observing certain region of milky way galaxy centered at R.A.= $19^h22^m40^s$ and Dec= $+44^\circ30'00''$ between 2009 and 2013. Kepler telescope has 0.95 m aperture and 115.6° field of view(As shown in Figure 2.1).Kepler data is being organised into 17 quarters(93 days period each).Since kepler telescope was being rolled at 90° quarterly for solar panels to remain oriented towards sun and radiator pointed towards deep space such that stars in its field of view remains same but now target falls on different CCD channel.LC targets were obsevered at least for a quarter and SC targets at least for a month.

K2 Mission:

This mission utilize spacecraft of kepler mission which came to its end 2013 after failure of two reaction wheels.Precision positioning was being done using thrusters and remaining two reaction wheels. Total obsevation period lasted 83 days ,designated as campaign. Field of view and balancing is shown in Figure 2.2.This mission ended in 2018.

There are two components of kepler flight system-

1. **Spacecraft:**Spacecreft provides power, communications (through High Gain Antenna(HGA) ,Low Gain Antenna(LGA),Kepler Control Boxes(KCBs)) and attitude control (Via Star Trackers).
2. **Photometer:**Photometry is measurement of brightness(flux) of star. The device that is being used for very purpose is Photometer as shown in figure 2.3.It has wavelength range is of 420-900nm.This range is resultant of all elements of optics(Schmidt corrector,primary mirror,field flattener lenses, CCD quantum efficiency etc) combined and chosen to avoid time-variable spectral lines found in spectra of solar-like stars.It consists of following parts-
 - **Telescope:**Wide-field Schmidt Telescope ,with schmidt corrector of diameter 0.95 m at circular entrance and primary mirror of 1.4m diameter, is being used.It has 16.1° field of view. Schmidt corrector consists of aspheric lens which is used to remove spherical aberration via introducing equal and opposite spherical aberration to primary mirror. f/ratio for it is 1.473 on axis with an effective focal length of 1399.20 mm.
 - **Focal Plane:**Focal plane consist of back illuminated and anti-reflection coated charge coupled devices(CCDs).CCD is composed of silicon which convert incoming photons into electrons(photoelectrons).There are 21 science CCD modules with 2

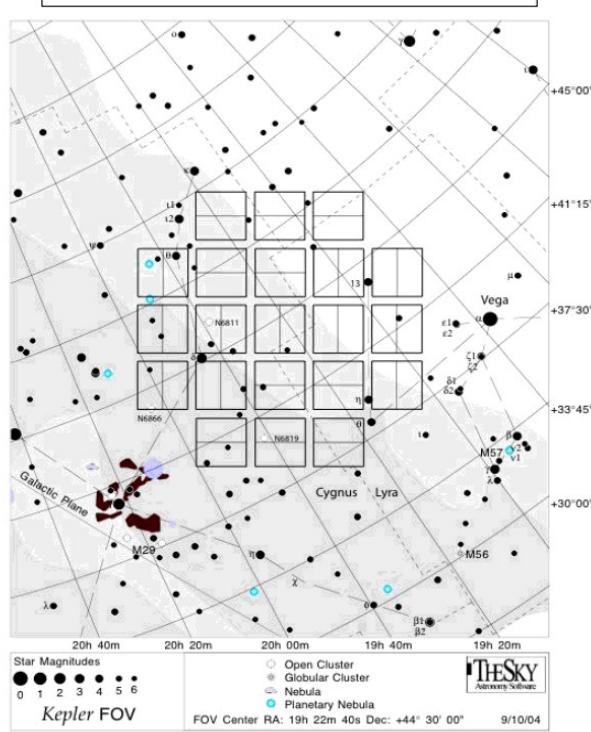


Figure 2.1: Figure adopted from Mast Kepler Archive Manual Kepler Field of View Assembled and Edited for the MAST by Susan E. Mullally showing Kepler Field of View

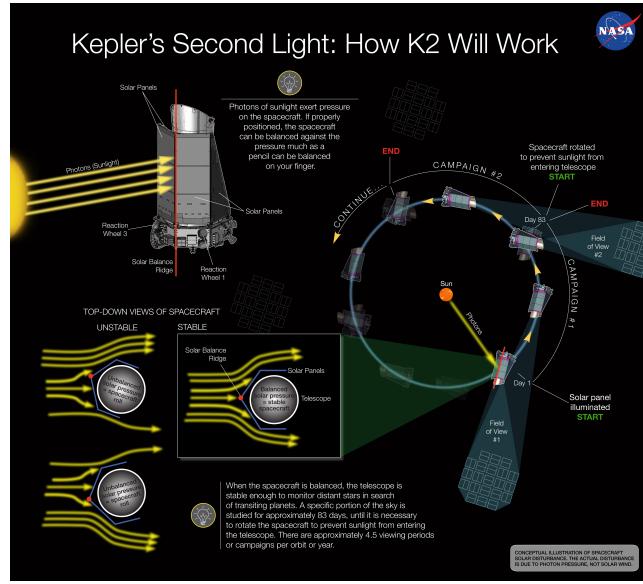


Figure 2.2: Figure adopted from Astrobites showing balancing of kepler spacecraeft against and solar pressure and new field of view

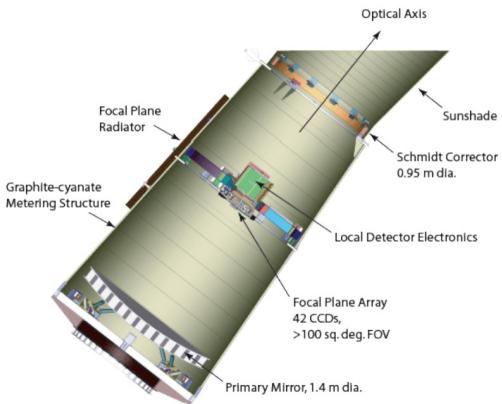


Figure 2.3: Figure adopted from Kepler Instrument Handbook, 22 April 2016 showing Kepler Photometer

CCDs in each and 4 FGS(Fine Guide Sensor) CCD modules. Each science module contains 2 CCDs of size 50x25 mm each with total 2200x1024 pixels and 2 outputs. So, total pixels observing sky are 94.6 million pixels. There are also masked and virtual pixels on board which are used for collateral data(Collateral data is being utilized to calibrate the pixels). Value of pixel corresponds to intensity of photons striking at point. CCDs have no shutter .

- **Field Flattener Lenses**
- **Radiator**
- **Local Detector Electronics:** Function of LDE is to digitize output of all science pixels.

Building an understanding of data:

1. How photometric data is being taken:

Steps that are involved in taking photometric data to downlinking it are as follows:

- (a) Star light falling is converted into photoelectrons using CCDs which is converted into digital signal by local detector electronics . So, Output of CCDs are metric of numbers (one per pixel) . Number is related to amount of light fallen on pixel. Time interval between two reads is 6.538753 seconds (known as integration time), consist of 6.019803 seconds for exposure and 0.51895 seconds for readout.
- (b) Output of each pixel at end of a cadence is sent to Science Data Accumulator(SDA) where it is coadded.
- (c) Since all coadded pixels could not be downlinked so Pixel of interest for target are being chosen using Target definition tables. There are three targets:
 - Photometric Target: Long Cadence(downlinked Quarterly) and Short Cadence data(Downlinked monthly).32 pixels were read out for a target on any average.270 integrations were done for long cadence data (29.424 minutes bins) and 9 for short cadence data (1 minute bins).
 - Reference Target: These pixels(96000) are used to get information regarding health of Focal Plane.(downlinked every 4 days)
 - Background Target: These pixels are utilized for removing zodiacal light and unresolved stars from data. They are being collected each time LC is collected. These are downlinked monthly

- (d) Coadded pixel are being communicated to KCBs . Coadded pixel of interest are stored in Solid State Recorder(SSR).
- (e) Using Deep Space Network(DSN) data were downloaded from spacecraft.
2. **What kind of data available to us:** Using Deep Space Network(DSN) data were downloaded from spacecraft.Data and telemetry packets are received by Mission Operations Center(MOC) at LASP which binned the files by data type.This data is sent to Data management Center(DMC) at STScl.Here data was decompressed,sorted by cadence and pixel type , and converted to FITS(Flexible Image Transport System) files.This is raw (uncalibrated) data which was sent to Science Operation Center(SOC) at NASA where detailed calibration is done and returned to DMC.Data products(Combination of data set and data processing description) are delivered to MAST(Mikulski Archive for Space Telescopes) which is made available to users as FITS files.Data available can be categorized as follows(we have mention data which is being utilized by us)-

- (a) **Light Curves:** Light Curves are flux time series.Flux is calculated in unit electron per second.Information available in Lightcurve FITS is as follows:
- i. SAP(Simple Aperature Photometry) : Flux in optimal aperture pixels. It is output of PA module in SOC pipeline.
 - ii. Error in SAP (SAP_FLUX_ERR) is also provided.It is sum of minimal error calculation(shot noise + read noise) and offset term to account for extra error from full propagation of errors.
 - iii. SAP_BKG:Flux for background
 - iv. SAP_BKG_ERR
 - v. PDCSAP(Presearch Data Conditioning Simple Aperature Photometry):After PDC module has been applied to PA Light Curve,flux in optimal aperture. It is basically flux corrected for systematics.
 - vi. PDCSAP_FLUX_ERR: 1-sigma error in PDC
 - vii. SAP_QUALITY :give information regarding quality of data.

(b) **Target Pixel Files(TPFs):**

TPFs are cutout images of each observed target which is known as target mask. Raw and Calibrated pixels are available as time series of images in binary FITS table.These are made available for users via MAST archive. Information available in FITS file is as follows-

- Raw counts
- Flux after calibration
- 1σ flux uncertainties of each pixels
- Calibrated Sky Background
- 1σ flux uncertainties of sky background
- Cosmic Ray incidences
- Quality Flag asscoiated with every cadence: have information about quality of photometric measurement. This can be used by user to make a choice regarding use of data.
- Information about pipeline mask
- motion of target over the time (using motion of a set of reference stars across the detector)

Dealing with TPFs: Target mask(TPF associated with a target) contains optimal aperture and halo around it.Optimal aperture is used for calculating flux while halo is used to mostly for calibration purposed.

- **Optimal Aperture:**

Optimal Aperture refers to pixels in target mask which are going to summed over to get flux (with background subtracted). Choosing of these pixel is known as aperture photometry.

Choice should be made in such a way that artifacts can be mitigated(which usually requires more pixels result into more sky background) whiling not increasing signal to noise ratio(SNR).

- Moment of centroids can be calculated using it

While the calibration astronomical signals and systematics are not removed so resultant of Aperture photometry (Calibrated light curve) do contains these which need of processed further

- (c) **Cotrending Basis Vectors(CBVs):**They represents most common trends present over each CCD channel . These are arranged by order of their relative amplitude. These can be utilized to remove systematic trends . These are created while creating PDC time series. They are being fitted to light curves using linear least square fit to find best fit which can remove systematic effect while reducing noise . For kepler and K2 data single scale CBVs are available ,in which all systematic trends are combined in single set of basic vector.

16 CBVs are provided each quarter to MAST .It is over to user how many of them need to be fitted since every inclusion results into addition of noise also.

3. To understand data there is need to be familiar with following concepts:

- (a) **Pixel Response Function(PRF):**Since transmit signals for Earth-sized exoplanets orbiting Solar-like stars are of order 100 ppm so there is need for high photometric precision . For high photometric precision, there is need for detailed knowledge of how kepler pixels respond to starlight during nominal observations which is provided by Pixel Response Function(PRF). PRF is super resolution representation of interaction of light coming from stars with pixels. It contains information regarding point spread function(PSF) modulation by pointing jitter and other systematic effect during observation.PSF should neither be most compact since it can degrade photometry nor be too broad since it can increase background noise. It provides us with continuous representation which is used to estimate star flux value in a pixel.
- (b) **Part Per million:** It is unit which is used to indicate fraction change in signal if we normalize it to 10^6
- (c) **Combined Differential Photometric Precision(CDPP):**It is measure of how effortlessly one can determine the transit signal from the data. In lightkurve library , which is been used by us to calculate cdpp uses “sgCDPP proxy algorithm” discussed in [2]. It measures the scatter that remains after long term trend is removed using Savitzky-Golay method .

4. Data Characteristics:

- (a) **Cadence :**Readout pixels are being coadded for specific number of integrations. This set of coadded pixels is called a cadence. It is absolute unique number which is enumerated with cadence interval numbers (CIN),which is assigned even if no cadences are being collected.

(b) **Raw Pixels :**

- Bias Level
- dark current:almost zero
- Smear: Since readout is shutterless,star illuminate CCDs during this period. This is being calculated using masked and virtual smear pixels
- Gain:Non Linear Transfer function
- Undershoot
- Flat Field

(c) **Pointing Error:** It is being introduced due to movement of telescope line of sight over time.Attitude Determination and Control System (ADCS) consists of Star trackers , FGS (40 stars) , reaction wheels and thrusters.

(d) **Differential Velocity Aberration:** It is due to change in angle between velocity vector of kepler telescope and line of sight. It causes image motion of 0.6 pixels over about 90 day period.

(e) **Read Noise:** It is electronic noise due to amplifiers on CCDs that is being introduced everytime while CCDs are readout.

(f) **Shot Noise:** This noise arise due to fluctuation in current that powers LDE.

(g) **Reasons for poor photometric cadence:**

- Monthly data downlinking :Telescope change its orientation every month to downlink data ,last for one day.It introduces data gap.Thermal transient can increase flux when telescope return to regular position. It can be identified in SAP as downward slope.
- Safe Mode: Telescope shut down its operation for some time due to some unpredicted event. It introduce gap in data.
- Coarse pointing:cadence with suspected loss of fine pointing
- no fine point : Cadence due to unexpected event
- loss of fine pointing:for these cadences flux values are replaced by NaNs.
- cosmic rays:results into sudden pixel sensitivity drop
- Argabrightening : It is caused by illumination of some portion of focal plane due to unexpected events like falling of debris on focal plane. Effect of it last for few minutes.
- Attitude Tweaks:for Q0-Q2,kepler's attitude is adjusted every few days so that star don't move over pixel . From Q-3 onwards ,change in FGS is done so no tweaks were required
- Reaction wheel event:zero crossings and momentum desaturation. These cadences are assigned NaNs values.
- Coronal mass ejection:manual exclusion of cadence is requires

(h) **K2 Mission specific errors:** Source motion on pixel due to solar pressure and thrust firing.

2.2 Lightkurve

Lightkurve is a python package which is designed to handle kepler data. It provides with different correctors that are being used to remove errors in tpf or lightcurves.

1. **flatten:** It removes long term trend using Savitzky -Golay(SG) filter with parameters - window length (number of coefficients) ,polyorder (order of polynomial). By reducing window length and polyorder,we can get shorter segements to apply SG filter.
2. **Pixel Level Decorrelation (PLD Corrector):** K2 Mission use kepler telescope with two reaction wheels by pointing is being done using solar wind pressure and periodic thrust firing Still it does not have precision in pointing of that of kepler which complicates the transit search so systematic trend due to point error need to removed . One way is using PLD corrector.It operates on background corrected pixels time series to remove trends. Working of it can be explained in following steps:
 - Intensity of each pixel of detector is being calculated for defined optimal aperture for target star.
 - Normalized intensity with total flux in aperture is used as basis vector for PLD model . Since normalization remove common trend in all pixels so astronomical signal is being removed and pld corrector work on variation in flux over pixels.
 - PLD model is expressed as follows:

$$m_i = \sum_l a_l \frac{p_{il}}{\sum_k p_{ik}} + \sum_l \sum_m b_{lm} \frac{p_{il} p_{im}}{\left(\sum_k p_{ik}\right)^2} + \sum_l \sum_m \sum_n c_{lmn} \frac{p_{il} p_{im} p_{in}}{+} \dots \quad (2.1)$$

where m_i is noise model at time t_i , p_{il} refer to flux in l^{th} pixel at time t_i , a_l , b_{lm} , c_{lmn} is PLD coefficients ,summation is being done on all pixels in aperture.

- Principle Component Analysis (PCA) [7] is being performed to reduce dimensionality . These reduced vectors are now basis for PLD model
- To obtain coefficents mimimization of χ^2 (eq2.2) is being done with respect to coefficents.

$$\chi^2 = \sum_i \frac{(y_i - m_i)^2}{\sigma_i^2} \quad (2.2)$$

where σ_i is standard error of SAP flux and y_i is SAP flux given by

$$y_i = \sum_k p_{ik} \quad (2.3)$$

- Now detrending is being performed by subtracting this noise to signal
3. **CBV Corrector:**CBVs are fitted to remove instrumental trend.
 4. **Self Flat Fielding Correction(SFF Corrector):**This method is being used to remove photometric variablity due to motion of spacecraft. Following are steps that involved-
 - With the assumption that short period variablity is dominated by 6 hr spacecraft pointing jitter, lightcurve is divided into small parts with breaking point every 1.5 days
 - Via iterative fitting basic spline (piecewise polynomial function) to these segement low frequency variability is estimated

- High frequency trend is estimated by dividing aperture photometric time series by final B-spline fit.
- Remaining variability is combination of astrophysical signal and spacecraft's pointing jittering. These are being separated by removing noise component that correlates with position of star on detector.
- Arc length which have info about position of star is being divided into 15 bins and mean was calculated within each bin with clipping of 3σ outliers (to avoid astrophysical signal in correction).
- linear interpolation between mean is being done and this thing is applied to aperture photometry to remove low frequency trend.

2.3 Characterizing the transit:

We have chosen 4 kepler object and 1 k2 object for analysis of light curve.

2.3.1 Kepler-8b

TPF associated with kepler-8b (KIC 6922244) for quarter 16 is obtained from MAST using lightkurve library in python. First we downloaded it without any target mask to have a look at all cadence present even with quality issues and NAN values for flux. To find reason for poor cadences we went through Kepler Data Release 25 Notes Q0–Q17 which describe about all events that take place in a particular quarter. Reason for few quality flags particular to Q16 are mentioned in Table :2.1 Instead of using tpfs without removing any of quality flag

Table 2.1: Quality Flags for quarter 16

Quality flag	Explanation
Coarse Points	temporary increase in friction of reaction wheel 4 resulting into degradation of pointing stability
Safe Mode	increase in amount of torque required in one of remaining three wheels which lead to kepler to go to safe mode for period of 11.3 days
Exclude	Solar flares

we used default quality mask which removes cadences with severe quality issues (e.g. coarse points). Now this tpf file is being used to file optimal aperture. We have used threshold mask with threshold 2σ to choose optimal aperture (Figure:2.4). Basically threshold mask choose pixels with median flux value higher than threshold times standard deviation above overall median. Standard deviation is estimated by multiplying median absolute deviation (MAD) with 1.4826. The value of threshold is being decided based on cdpp value. Mask with lowest cdpp value is being chosen. Now with this mask lightcurve is created which is shown in Figure:2.5. Data gaps are due to safe mode, monthly downlinking of data. We observed an increasing trend after data gaps this was due to attitude change for downlinking which led to thermal transient and some time is consumed to come into thermal equilibrium. Now this lightcurve still has systematic and astrophysical noise associated with it which are corrected in following steps:

1. To apply correction first we have prepared transit mask which is going to mask cadences to involve any correction. For that we normalized light curve and applied BLS to find

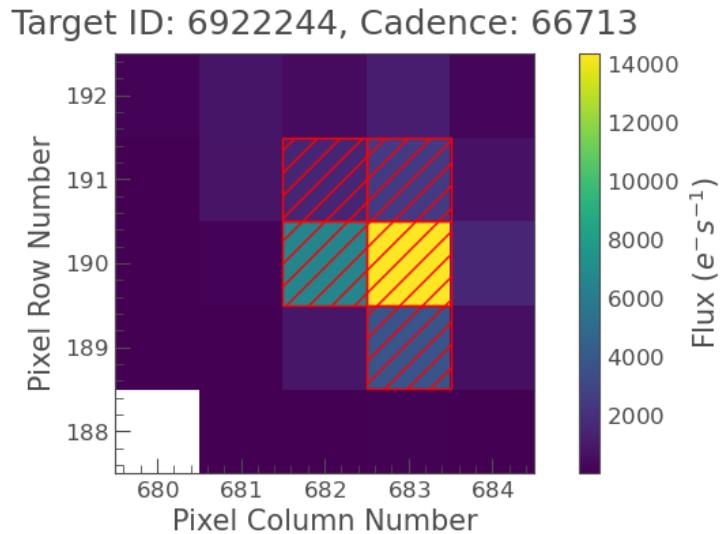


Figure 2.4: TPF with optimal aperture using threshold mask of 2σ threshold

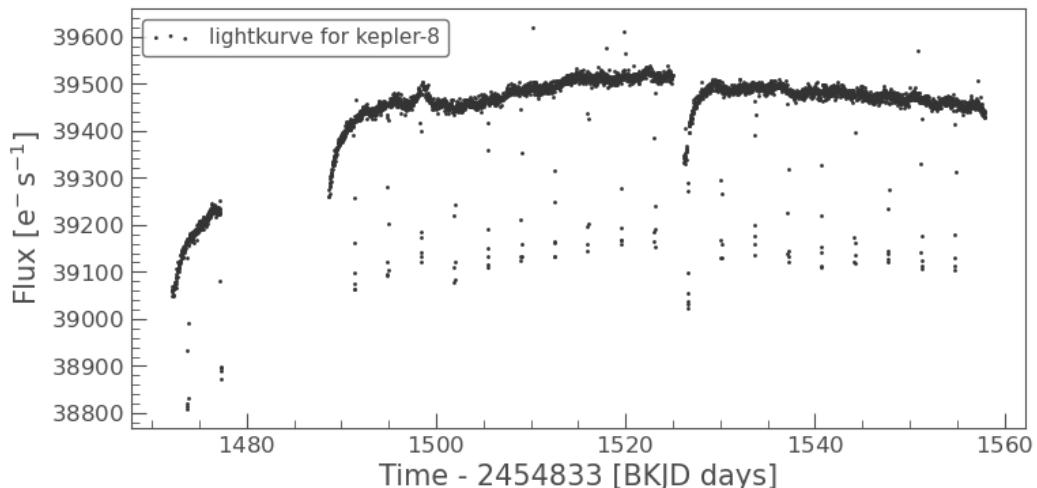


Figure 2.5: Uncorrected Light curve for kepler-8

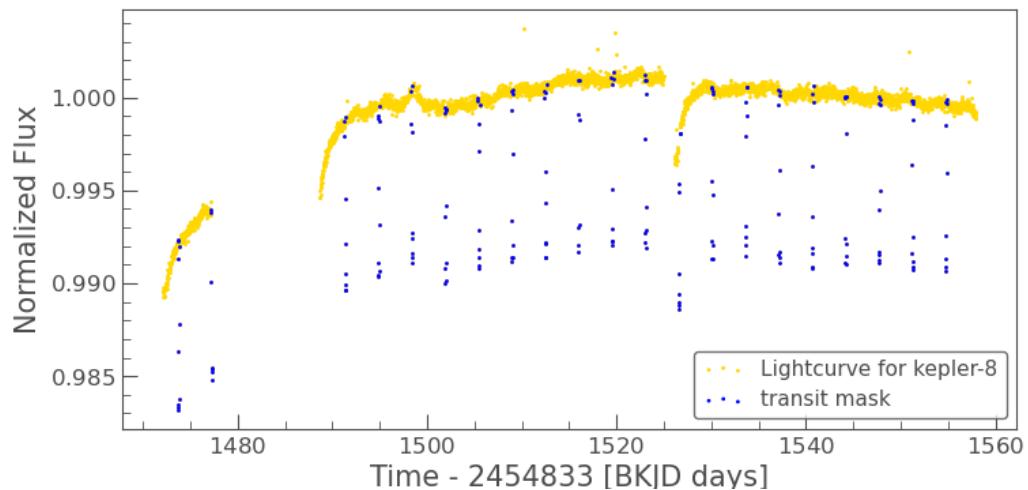


Figure 2.6: Transit mask for Light curve for kepler-8

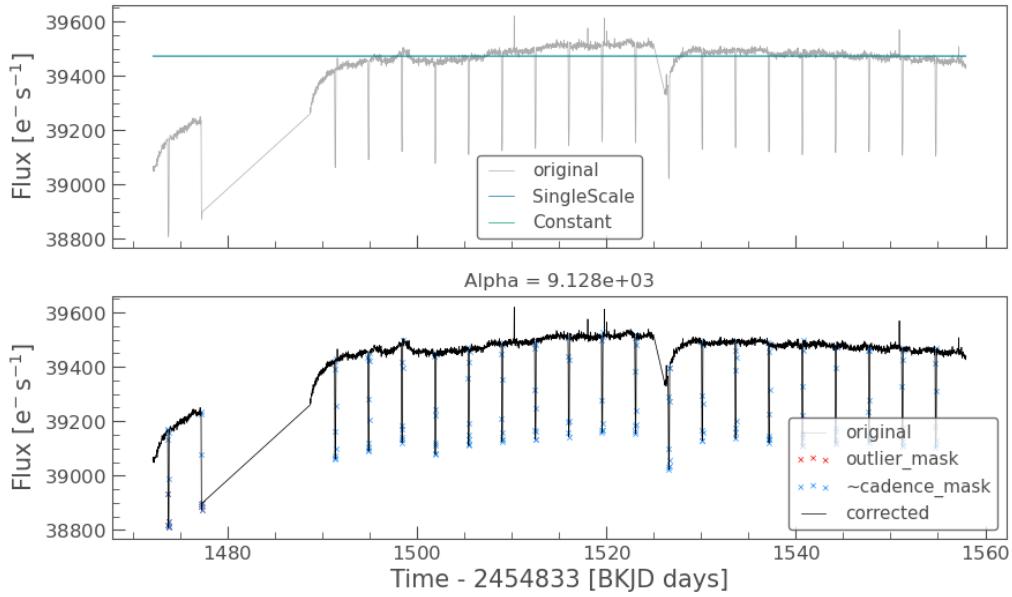


Figure 2.7: CBV correction to lightcurve of kepler-8

period,duration and depth of transit so that we can choose cadences to be masked .Transit mask is shown in Figure: 2.6

2. CBV correction was also applied. First 9 CBVs with transit mask are applied to uncorrected light curve so that systematic trend can be removed.But this come to no avail since via `cbv.diagnose()` we found no difference so we decided to continue with our original uncorrected light curve.(see Figure:2.7)
3. Now flatten command was used again on uncorrected light curve with number of coefficients require to fit are chosen so that CDPP is low and detrended lightcurve was obtained (Figure:2.8).This could be due to unavailability of multi scale CBVs, which are given different band pass.
4. Phase folding for this corrected light curve is done and BLS model is implemented.(as shown in Figure:2.9)
5. With help of transit mask all transit were removed and again BLS was applied but no significant dip was observed this time.(Figure:??)

2.3.2 Kepler-7b

for kepler-7 we have used data of Quarter one.There was no specific event related to Q1 only.Same process is being followed as for kepler-8b. Figure:2.11 represents uncorrected lightcurve obtained from aperture photometry and corrected light curve (free of all systematics). Transit cadences were removed and again bls was applied to know whehter there is another transit like deep is there.

2.3.3 Kepler-6b

We have taken data for Q9 to analysis light curve of Kepler-6b.In this quarter LDE become out od sync but it has not affect long cadence data.This time , TPF with hard mask is used

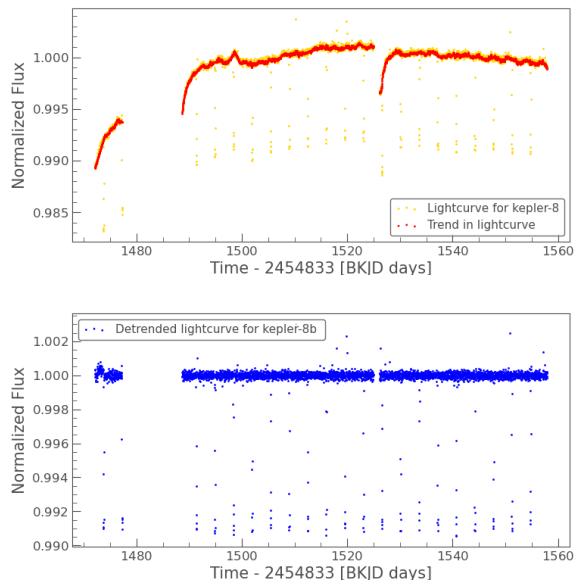


Figure 2.8: Detrended Light curve for kepler-8

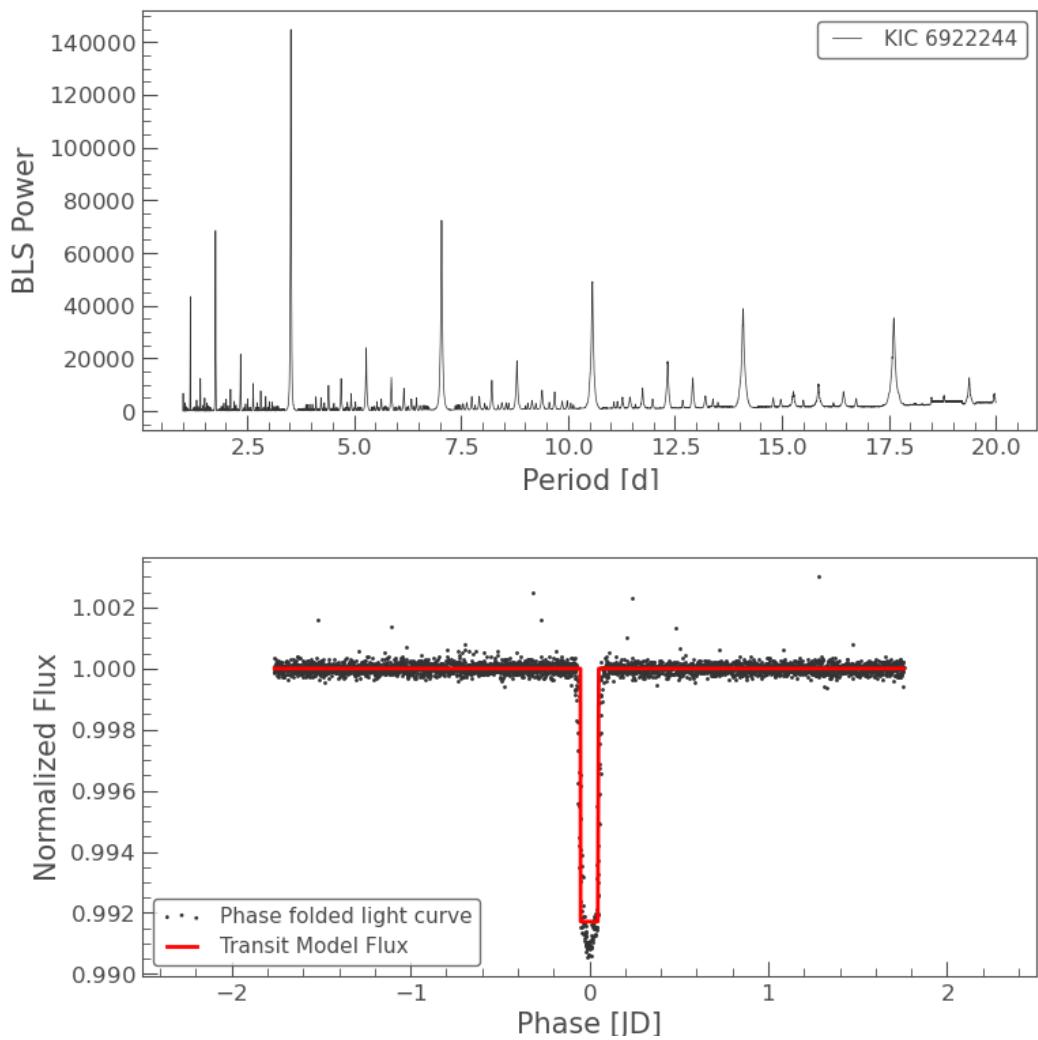


Figure 2.9: Phase folded light curve and BLS model for Kepler-8b

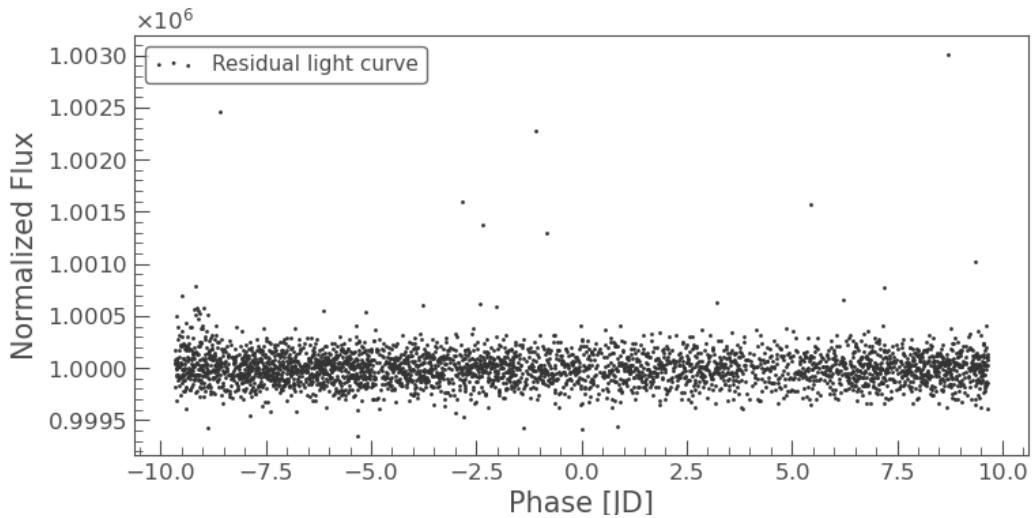


Figure 2.10: Residual light curve

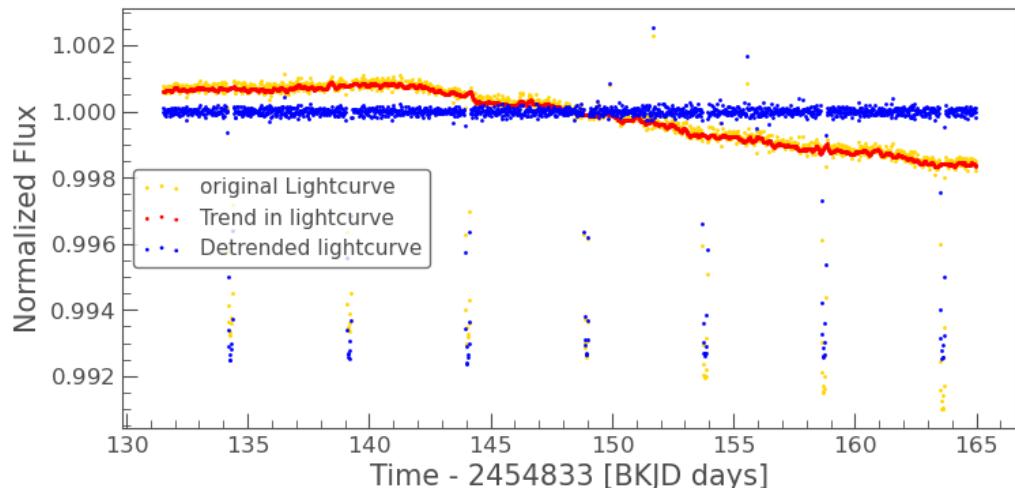


Figure 2.11: Kepler-7 lightcurve

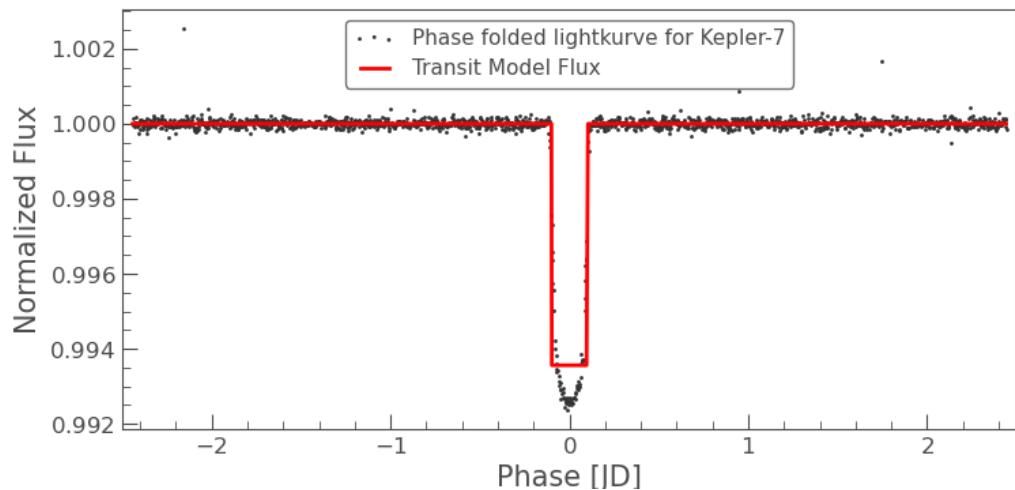


Figure 2.12: Phase folded light curve and BLS model for Kepler-7b transit

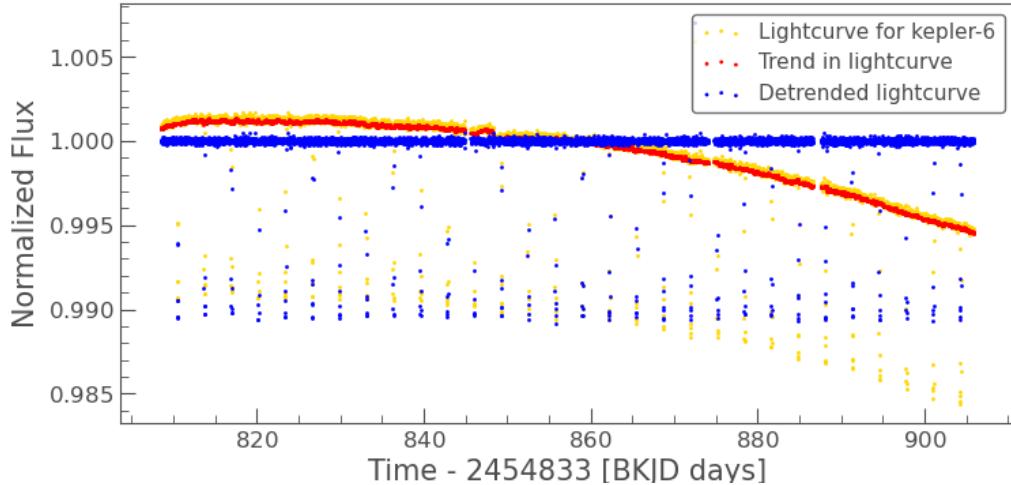


Figure 2.13: Kepler-6 lightcurve

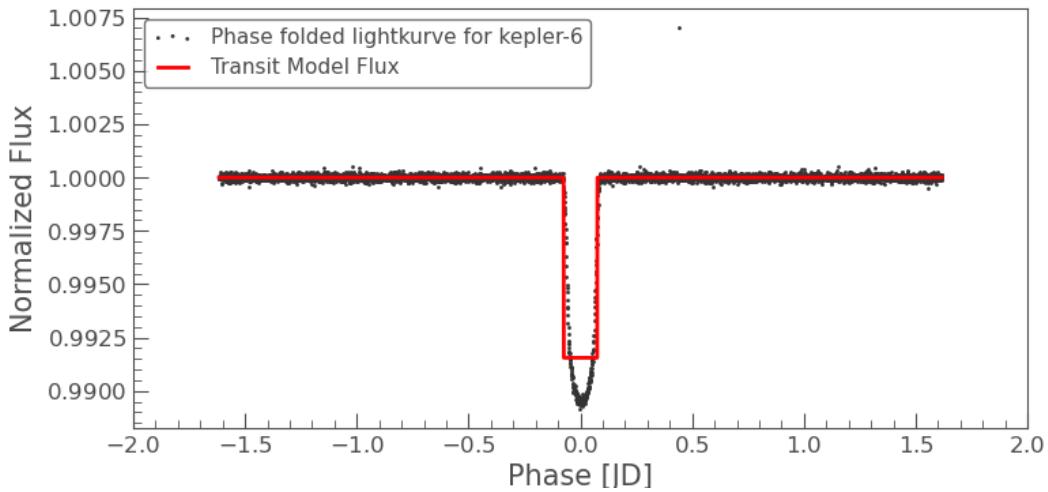


Figure 2.14: Phase folded light curve and BLS model for Kepler-6b transit

since there were lot of cosmic event present . From TPF raw light curve was obtained using threshold mask with threshold 3σ and then same process as in section:2.3.1 When we removed transit cadences and applied BLS all periods were giving same power so no secondary transit was observed.

2.3.4 Kepler-5b

For Kepler-5b data for quarter-7 was taken .No specific event was associated with it. Same steps were required as in section:2.3.1. Figure:2.15 and 2.16 is showing final results.

2.3.5 WASP28-b

For WASP28 data was collected by K2 mission. Since this mission same spacecraft but with only two reaction wheels with balacing as explained in section:?? .This causes degradation in pointing. So addition correctors were required to tackle the issue.We have taken data of campaign 12 for getting light curve.Specific events associated with it are shown in Table:2.5

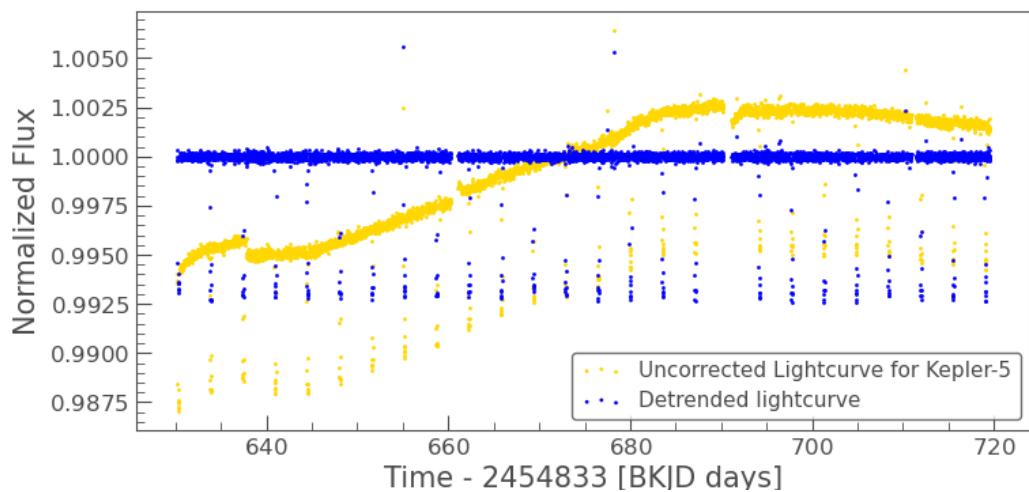


Figure 2.15: Kepler-5 lightcurve

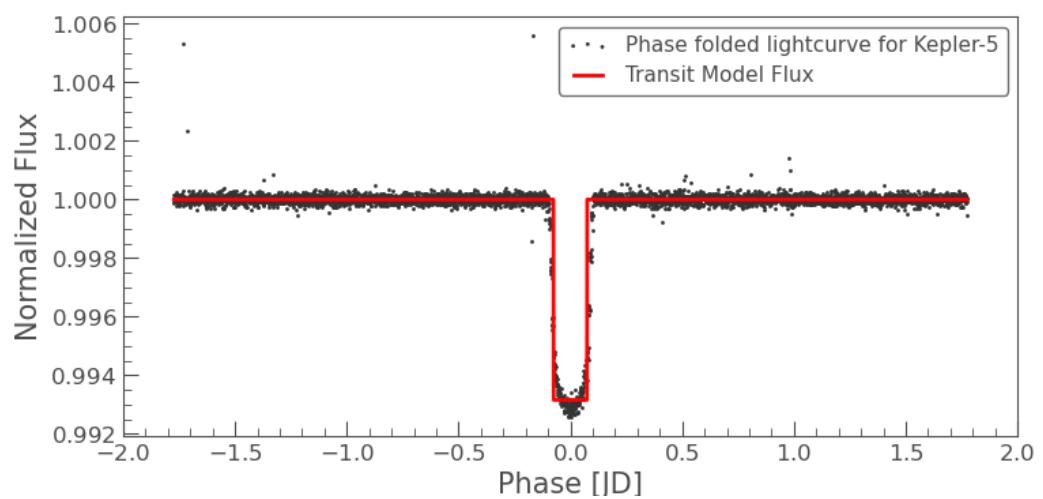


Figure 2.16: Phase folded light curve and BLS model for Kepler-5b transit

Table 2.2: Events related to campaign 12

Event	Explanation
Safe Mode	Flight software reset (total time loss was 5.25 days),engineering data(thrust firing) was not being downlinked during safe mode recovery

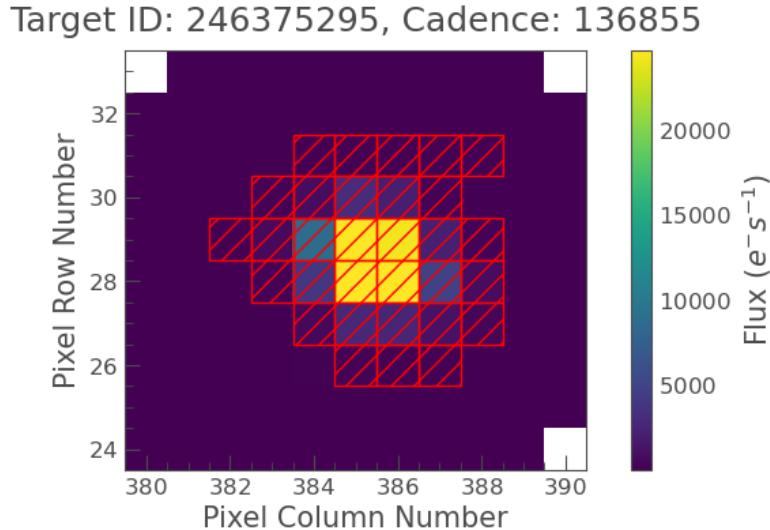


Figure 2.17: Target pixel file with optimal aperture chosen using threshold mask with $\sigma = 3$ for WASP28

From TPF to uncorrected lightcurve was obtained in same way as in kepler objects . Transit mask is also prepared in same way(shown in Figure:2.18) .

For further analysis following steps were being followed:

1. PLD corrector was applied to tpf which remove trend due to spacecraft motion. (2.20)
2. Now CBV correction is applied to obtained lightcurve but again it did not things much
3. Flattening command was applied which made cdpp lower
4. Phase folding was done using bls .Results are shown in Figure:2.21

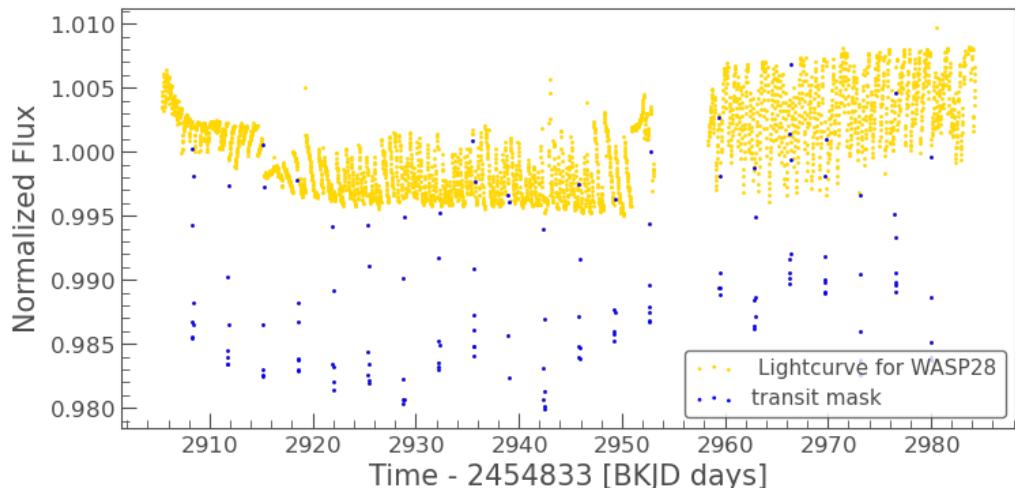


Figure 2.18: Transit mask for WASP28

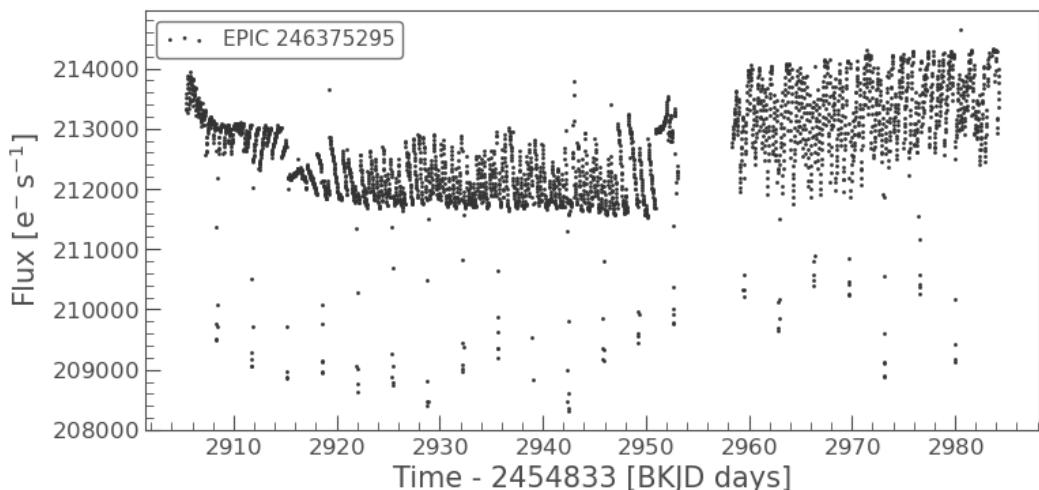


Figure 2.19: Uncorrected lightcurve for WASP28

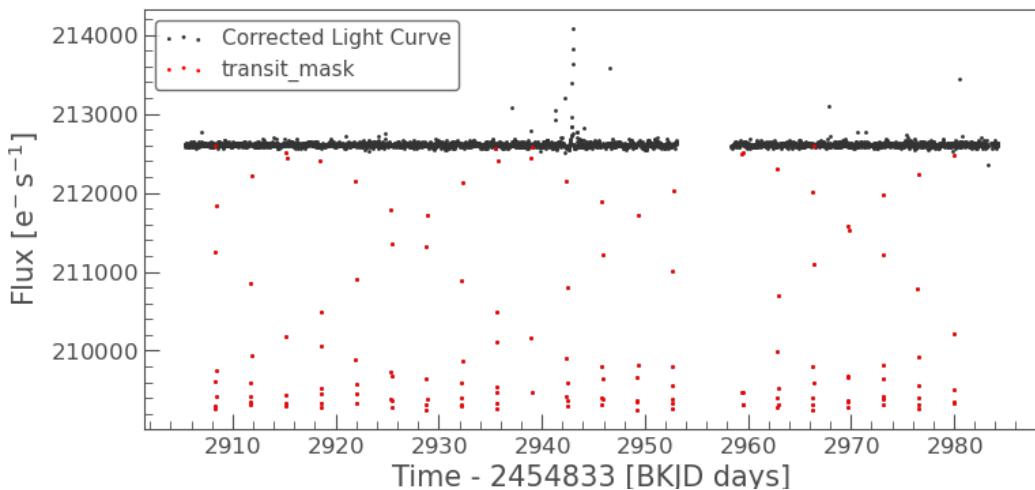


Figure 2.20: PLD corrector

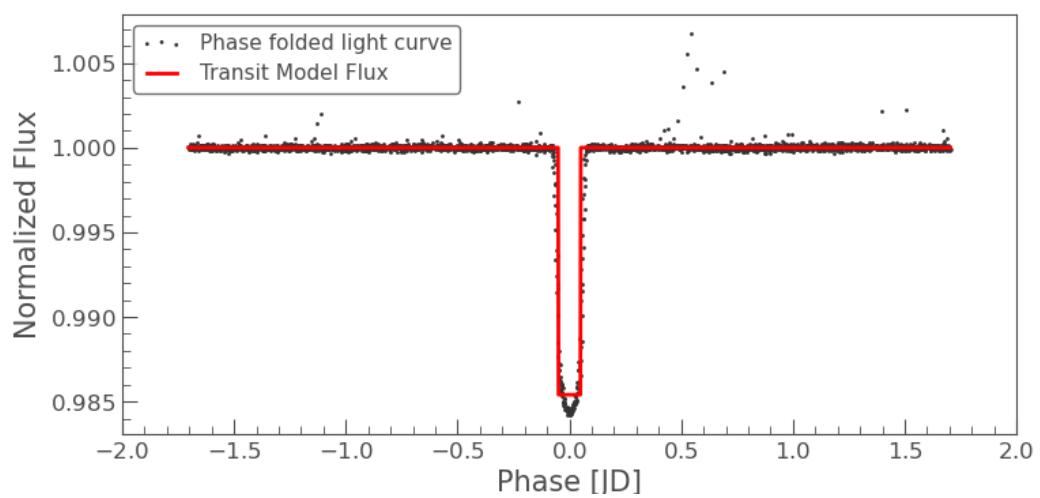


Figure 2.21: Phase folded light curve for WASP28

2.4 Results

BLS modeling is used to find parameters of transit- period , depth and duration of transit. From there we found semi-major axis ,radius,impact parameter and inclination angle of planet using formulas given in chapter 1 :

Table 2.3: Targets used for transit analysis

Planet	Target ID	Mass(in M_{sun})	Radius (in R_{sun})
Kepler-8b	KIC 6922244	1.13	1.45
Kepler-7b	KIC 5780885	1.35	1.84
Kepler-6b	KIC 10874614	1.21	1.39
Kepler-5b	KIC 8191672	1.347	1.793
WASP-28b	EPIC 246375295	0.993	1.083

Table 2.4: Parameter of transit

Planet	Period of Transit(in days)	Duration of Transit(in days)	Depth of transit
Kepler-8b	3.5216	0.1	0.0083
Kepler-7b	4.8859	0.2	0.0064
Kepler-6b	3.2346	0.15	0.0084
Kepler-5b	3.5482	0.15	0.0067
WASP-28b	3.4094	0.1	0.0146

Table 2.5: Planet Characteristics

Planet	Orbital Period(in days)	Radius(in $R_{jupiter}$)	Semi-major axis(in AU)	Inclination Angle (in degree)
Kepler-8b	3.5216	1.285	0.047	82.65
Kepler-7b	4.8859	1.437	0.062	85.72
Kepler-6b	3.2346	1.243	0.045	86.93
Kepler-5b	3.5482	1.428	0.050	83.05
WASP-28b	3.4094	1.273	0.044	84.92

Chapter 3

Radial Velocity Method

3.1 Introduction

Radial velocity (RV) measurements of stars can be obtained using an iodine cell. The iodine cell is a small glass cell filled with iodine vapor, which is placed in front of the telescope's spectrograph. When the light from a star passes through the iodine cell, the iodine absorbs specific wavelengths of the star's light, creating a complex set of absorption lines known as the iodine spectrum.

By comparing the iodine spectrum to a calibration spectrum taken without the iodine cell, we can determine the precise wavelength calibration of the spectrograph. This allows them to measure the precise Doppler shift of the star's light caused by its motion towards or away from Earth. The Doppler shift is caused by the star's motion towards or away from us, which changes the wavelength of the star's light due to the Doppler effect.

By monitoring the Doppler shift of the star's light over time, astronomers can track the star's motion and determine its RV. RV measurements are essential for detecting and characterizing exoplanets, as the gravitational pull of an orbiting planet causes the star to move in a small but measurable orbit around the center of mass of the system. By analyzing the RV data, we can infer the presence, mass, and orbital parameters of the planet(s) around the star.

3.1.1 Lomb-Scargle method

This part of the project contains code and images related to the analysis of the radial velocity data of certain stars. The code is written in Python (Jupyter notebook) and the plots are generated using Matplotlib. First of all we can discuss a little about use of lomb scargle periodogram in this data. Lomb-Scargle is a well-known method used in astrophysics for the analysis of unevenly sampled time-series data. It is particularly useful in the field of radial velocity studies of exoplanets, where the goal is to determine the orbital parameters of a planet orbiting around a star by measuring the star's radial velocity over time.

In the case of radial velocity studies, the Lomb-Scargle periodogram is constructed by computing the power spectrum of the residuals after fitting the RV measurements with a sinusoidal model. The periodogram is a plot of the power (or signal-to-noise ratio) as a function of frequency or period. It can reveal the presence of periodic signals in the data, which can be attributed to the orbital motion of the planet around the star.

The Lomb-Scargle method is particularly useful for analyzing unevenly sampled data, as it takes into account the nonuniform spacing of the data points. It works by fitting a sine wave to the data at each trial frequency and then computing the power of the residuals as a measure of the goodness-of-fit. The Lomb-Scargle periodogram can be used to determine the most likely period of the signal in the data, as well as the amplitude and phase of the

sinusoidal fit. The Astropy library provides a Python implementation of this algorithm. The mathematical formulation of the Lomb-Scargle periodogram is given by:

$$P(f) = \frac{1}{2\sigma^2} \left[\frac{\left(\sum_{i=1}^N y_i \cos(2\pi f(t_i - \tau)) \right)^2}{\sum_{i=1}^N \cos^2(2\pi f(t_i - \tau))} + \frac{\left(\sum_{i=1}^N y_i \sin(2\pi f(t_i - \tau)) \right)^2}{\sum_{i=1}^N \sin^2(2\pi f(t_i - \tau))} \right] \quad (3.1)$$

where $P(f)$ is the Lomb-Scargle periodogram, y_i and t_i are the observed values and times of the data points, respectively, σ is the standard deviation of the data, f is the frequency at which the periodogram is calculated and τ is a time offset.

The normalization used in the Lomb-Scargle periodogram is designed to ensure that the periodogram is unbiased, which means that the expected value of the periodogram at any frequency is equal to the true power of the signal at that frequency. The normalization is given by:

$$N = \frac{2}{\sigma^2} \left[\sum_{i=1}^N (y_i - \bar{y})^2 - \frac{\left(\sum_{i=1}^N (y_i - \bar{y}) \cos(2\pi f(t_i - \tau)) \right)^2}{\sum_{i=1}^N \cos^2(2\pi f(t_i - \tau))} - \frac{\left(\sum_{i=1}^N (y_i - \bar{y}) \sin(2\pi f(t_i - \tau)) \right)^2}{\sum_{i=1}^N \sin^2(2\pi f(t_i - \tau))} \right] \quad (3.2)$$

where \bar{y} is the mean value of the data. The Lomb-Scargle periodogram is then normalized as

$$P_{norm}(f) = \frac{P(f)}{N} \quad (3.3)$$

The normalization factor N ensures that the power in the periodogram is correctly scaled in relation to the noise in the data. It is important to use this normalization when interpreting the Lomb-Scargle periodogram, as it allows for meaningful comparison of the power at different frequencies. These methods can be used to obtain more accurate estimates of the orbital parameters of the planet, as well as their uncertainties.

In summary, the Lomb-Scargle method is a powerful tool for the analysis of radial velocity data in the search for exoplanets. It allows astronomers to detect the presence of periodic signals in unevenly sampled data, and to estimate the orbital parameters of the planet with high accuracy. By combining the Lomb-Scargle method with other techniques for fitting radial velocity data, astronomers can obtain a more complete understanding of the dynamics of exoplanetary systems.

3.1.2 False Alarm method

The normalized Lomb-Scargle periodogram is a widely used method for detecting periodic signals in time series data. However, when using this method, it is important to consider the possibility of false alarms, which occur when the periodogram identifies a spurious signal as significant.

An approach to estimating the significance of peaks in the periodogram is to use the bootstrap method. The basic idea behind the bootstrap method is to create many simulated datasets by randomly resampling the original data while preserving the statistical properties of the original data. Each of these simulated datasets is then analyzed using the same method as the original data, and the resulting periodograms are used to estimate the distribution of peak heights that would be expected by chance.

To perform a bootstrap analysis for the normalized Lomb-Scargle periodogram, one would follow these steps:

1. Create many simulated datasets by randomly resampling the original data.

2. Compute the normalized Lomb-Scargle periodogram for each simulated dataset.
3. For each frequency in the periodogram, calculate the fraction of simulated periodograms that have a peak height greater than or equal to the peak height in the original periodogram.
4. Use the resulting distribution of fractions to estimate the false alarm probability for each peak in the original periodogram.
5. By using the bootstrap method, one can estimate the probability of false alarms in the normalized Lomb-Scargle periodogram, which can help to avoid the interpretation of spurious signals as being significant.

3.1.3 Period04

Period04 is a software package for analyzing and modeling time-series data, particularly in the context of astronomical observations. One of the key features of Period04 is its ability to fit sinusoidal functions to the data, allowing for the detection and characterization of periodic signals. The fitting formula used in Period04 is based on a least-squares approach, and is given by:

$$y(t) = A \sin(\omega t + \phi) + c \quad (3.4)$$

where $y(t)$ is the value of the time-series data at time t , A is the amplitude of the sinusoidal function, ω is the angular frequency, ϕ is the phase offset, and c is a constant offset. The parameters A , ω , ϕ , and c are determined by minimizing the sum of squared residuals between the model and the data, using a nonlinear least-squares algorithm.

In addition to fitting sinusoidal functions to the data, Period04 also provides tools for analyzing the resulting periodograms, which show the power spectrum of the data as a function of frequency. The periodogram is calculated using the Lomb-Scargle method, which is a widely used technique for detecting periodic signals in time-series data.

To further characterize the periodic signals in the data, Period04 also provides tools for performing a phase analysis. This involves dividing the data into a set of phase bins, and calculating the mean value of the data in each bin. The resulting phase plot shows how the value of the data varies as a function of phase, allowing for the identification of any phase-dependent trends or anomalies. In Period04, Monte Carlo simulation is used to measure the uncertainty associated with the fitted parameters of the sinusoidal function. This is done by generating a large number of synthetic time-series data sets that are consistent with the noise properties of the original data and then fitting sinusoidal functions to each of these data sets using the same algorithm as was used for the original data.

Synthetic data sets are generated by randomly perturbing the original data points according to their error bars. This preserves the statistical properties of the original data, while allowing for the creation of multiple data sets that are consistent with the observed noise. For each synthetic data set, a sinusoidal function is fitted using the same algorithm as was used for the original data, and the values of the resulting parameters are recorded.

After a large number of synthetic data sets have been generated and analyzed, the distribution of parameter values can be used to estimate the uncertainty associated with the fitted parameters. This is typically done by calculating the standard deviation of the parameter values across the simulated data sets, and using this as a measure of the uncertainty.

Monte Carlo simulation is a powerful technique for estimating uncertainties, as it allows for the creation of multiple synthetic data sets that are consistent with the observed noise properties. This approach is particularly useful when the noise properties of the data are complex or difficult to model analytically, as is often the case in astronomical observations. By using Monte Carlo simulation to estimate uncertainties, Period04 provides a robust framework for the analysis and modeling of time-series data.

Overall, Period04 is a powerful tool for analyzing and modeling time-series data, particularly in the context of astronomical observations. Its fitting formula and associated tools provide a robust framework for detecting and characterizing periodic signals, and its flexible user interface allows for easy customization and visualization of the results.

3.2 HD 31253

HD 31253 is a star of magnitude $V = 7.133$ of spectral class F8. In this section we have used the radial velocity data listed in the appendix to search for planetary companion around the star. First, we plotted the observation data with corresponding error bars (Figure 1) and then using the Lomb Scargle method discussed above, we plotted the following normalized periodogram with indicated false alarm levels of 5%, 1% and 0.1% calculated using the bootstrap method for the above data (Figure 2). We can see that it consists of a single dominant peak of period 465.166 days whose probability of being spurious is $5.26e-08$. This period is further corroborated from period04 which also gave a peak in the power spectrum at period of 465.166 ± 3.179 days.

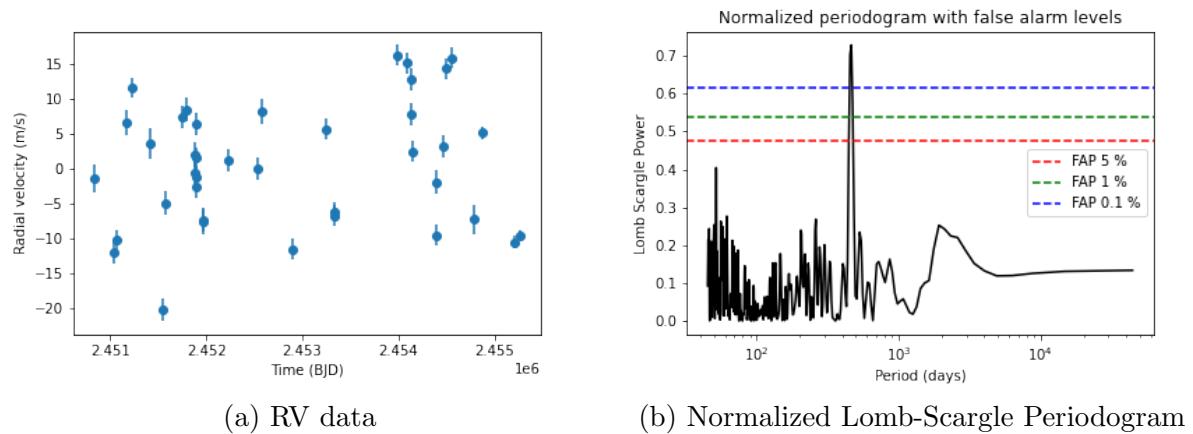


Figure 3.1: Radial velocity data and normalized Lomb-Scargle periodogram

Then further using a sinusoid fit by defining a function and using the curve fit of SciPy we evaluated the best-fit period (P), eccentricity (e), semi-amplitude (K), minimum mass ($M \sin i$), and semi-major axis (a) along with their errors. The mass of the star here is taken to be $1.23 M_{\odot}$. The value of eccentricity is measured from the maximum and minimum radial velocity of the best fit. The errors in minimum mass limit and semi major axis are measured using the error propagation method. Here is the best fit to RV observations. and also best fit to phase folded.

The table:3.1 lists the parameters found using the best fit to RV data along with uncertainties and reference values are also given. The reference values are obtained from Meschiari_2011_ApJ_727_117. The errors in period04 are calculated using Monte Carlo simulations.

Now after finding the parameters we can look further into the residuals to search for any additional planetary companion to the star. Following plots are for residual and again Lomb Scargle periodogram to the residuals. The peak in the residual can be ignored by the fact that its false alarm probability is around 1.8 % and is well below the safe FAP level of 0.1%. There are no further indications of any additional planets due to absence of any strong peak.

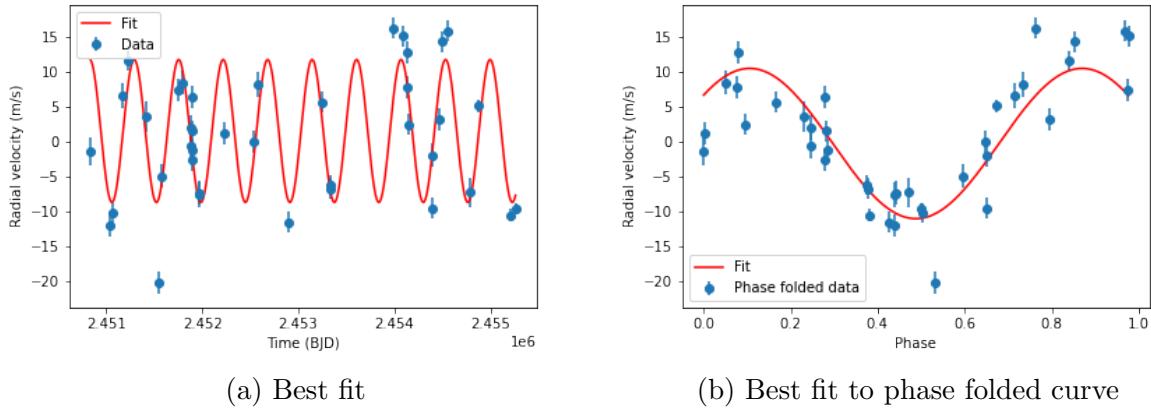


Figure 3.2: Least square fitting to RV data and associated phase folded curve

Table 3.1: Parameters and derived quantities for HD 31253b

Parameters	Calculated (LS)	Calculated (P04)	Reference
Epoch	2450838.75189	2450838.75189	2450838.75189
P (days)	462.678 ± 3.320	465.166 ± 3.179	466 ± 3
K (m/s)	10.192 ± 1.301	10.100 ± 1.063	12 ± 2
$M \sin i (M_J)$	0.446 ± 0.057	0.443 ± 0.047	0.50 ± 0.07
a (AU)	1.253 ± 0.006	1.257 ± 0.006	1.260 ± 0.006
e	0.159	-	0.3 ± 0.2
χ^2	11.41	25.20	8.87
Phase	-165.241 ± 239.051	3.492 ± 0.118	-
Offset (m/s)	1.555 ± 0.881	1.555 ± 0.849	-

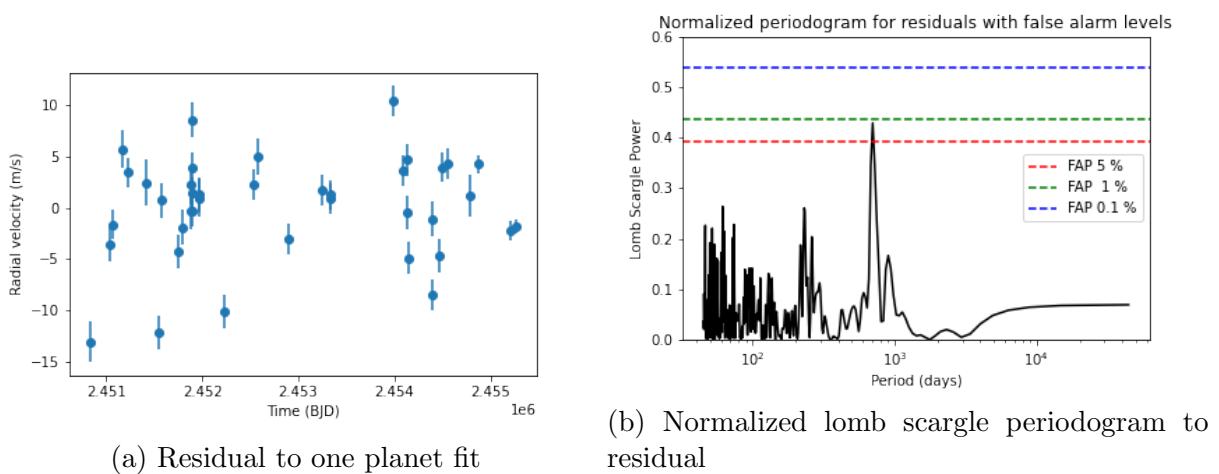


Figure 3.3: Residual and its LS periodogram

3.3 HD 218566

HD 218566 is a star of $V=8.628$ magnitude of spectral class K3V. In this section we have used the radial velocity data listed in the appendix to search for planetary companion around the star. First of all we plotted the observation data with corresponding error bars (Figure 1) and then using the above discussed Lomb Scargle method we plotted the following normalized periodogram with indicated false alarm levels of 5%, 1% and 0.1% calculated using bootstrap method for the above data (Figure 2). We can see that it consists of a single dominant peak whose probability of being spurious is $5.11e-10$. This period is further corroborated from period04 which also gave a peak in power spectrum at period of 225.559 ± 0.54 days.

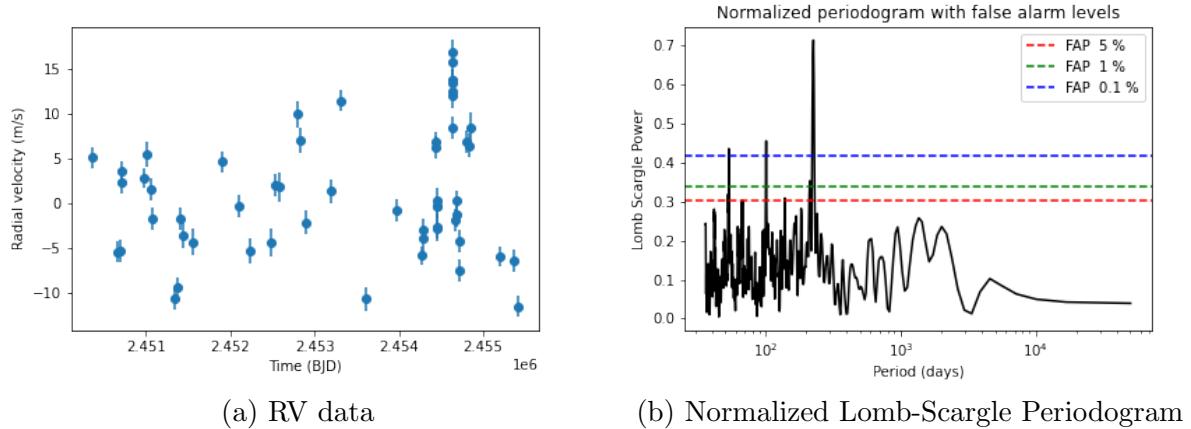


Figure 3.4: Radial velocity data and normalized Lomb-Scargle periodogram

Then further using a sinusoid fit by defining a function and using the curve fit of SciPy we evaluated the best-fit period (P), eccentricity (e), semi-amplitude (K), minimum mass ($M \sin i$), and semi-major axis (a) along with their errors. The mass of star here is taken to be $0.85 M_{\odot}$. The value of eccentricity is measured from the maximum and minimum radial velocity of the best fit. The errors in minimum mass and semi-major axis are measured using the error propagation method. Here is the best fit to RV observations and also the best fit for phase folded data.

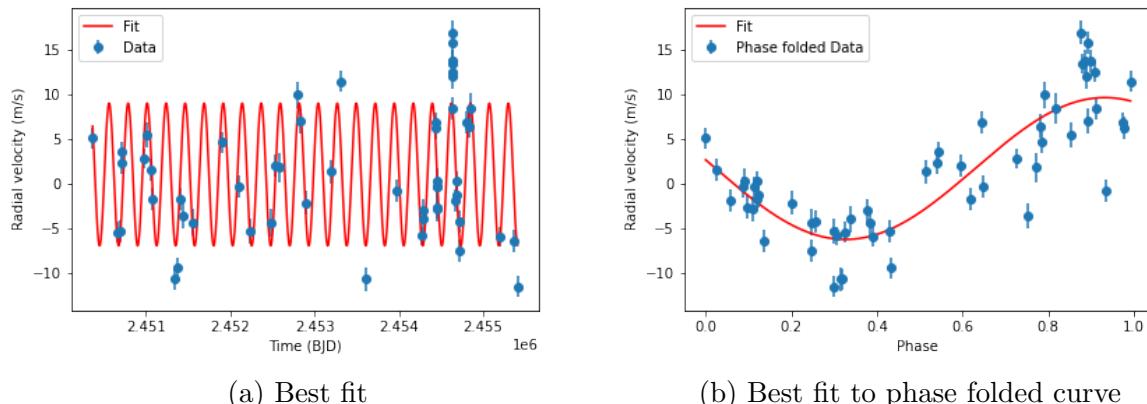


Figure 3.5: Least square fitting to RV data and associated phase folded curve

Following table lists the values of parameters found out using the best fit to RV data along with uncertainties and reference values are also given. The errors in period04 are calculated using Monte carlo simulations.

Table 3.2: Parameters and derived quantities for HD 218566b

Parameters	Calculated (LS)	Calculated (P04)	Reference
Epoch	2450366.85498	2450366.85498	2450366.85498
P (days)	226.065 ± 0.561	225.559 ± 0.540	225.7 ± 0.4
K (m/s)	7.996 ± 0.990	7.996 ± 0.794	8.3 ± 0.7
$M \sin i (M_J)$	0.208 ± 0.026	0.215 ± 0.021	0.21 ± 0.02
a (AU)	0.675 ± 0.001	0.686 ± 0.001	0.6873 ± 0.0008
e	0.128	-	0.3 ± 0.1
χ^2	9.82	14.82	8.41
Phase	-123.970 ± 169.904	5.534 ± 0.096	-
Offset (m/s)	1.029 ± 0.605	1.029 ± 0.522	-

Now after finding the parameters we can look further into the residuals to search for any additional planetary companion to the star. Following plots are for residual and again Lomb Scargle periodogram to the residuals. The peak in the residual can be ignored by the fact that

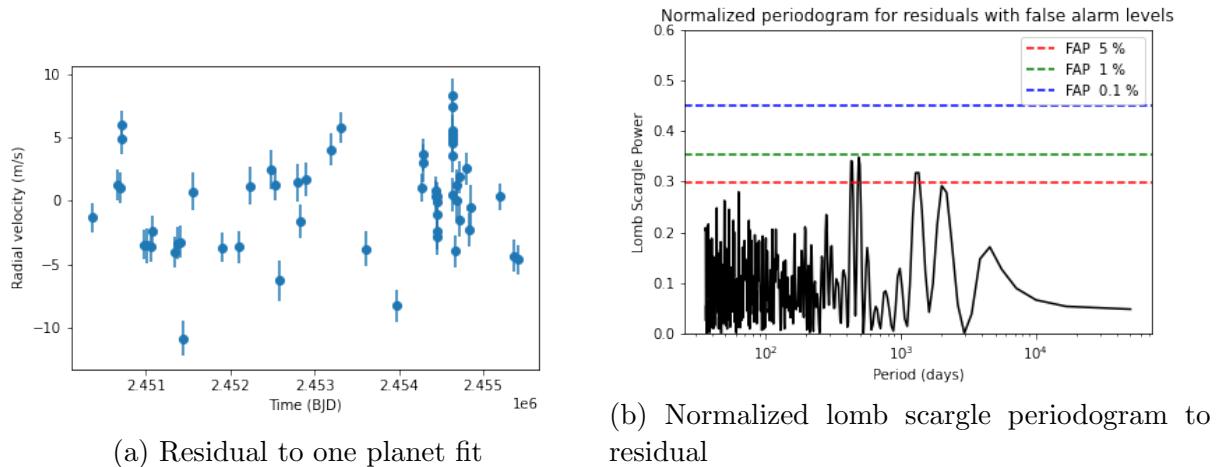


Figure 3.6: Residual and its LS periodogram

its false alarm probability is around 0.7 % and is well below the safe FAP level of 0.1%. There are no further indications of any additional planets. There are no further indication of any additional planets due to absence of any strong peak.

3.4 KIC 11853905

The planetary companion to this star was discovered by the Kepler mission in 2009. It is a star of $V=12.7$ magnitude and of spectral type G0 with mass of $1.22 M_{\odot}$. Although we have already discussed about this in transit method but here we are also providing a radial velocity analysis also. Below are the RV observations and normalized lomb scargle periodogram. We can see a dominant peak of false alarm probability 0.04 % of period 3.186 days.

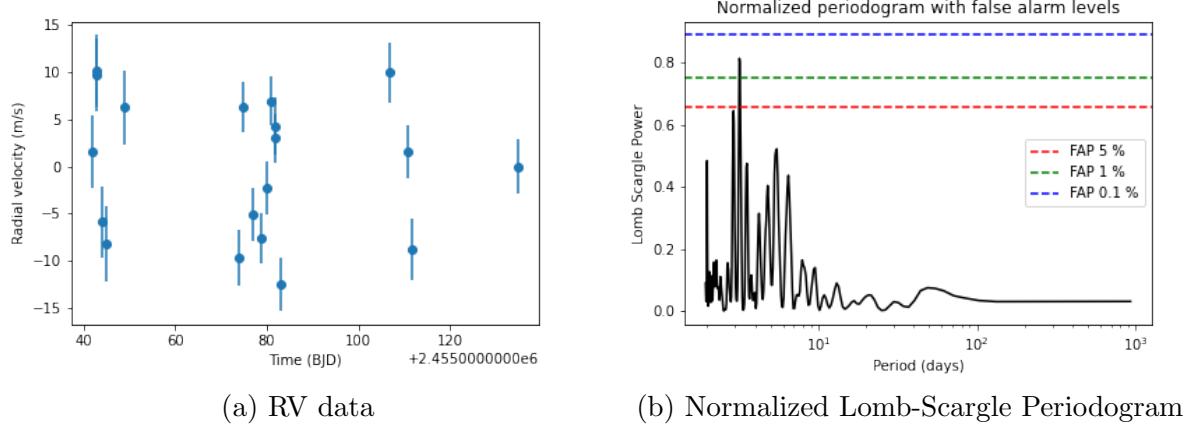


Figure 3.7: Radial velocity data and normalized Lomb-Scargle periodogram

Then we will repeat the process done for previous single planet stars to find out the best fit and associated parameters as shown in the below figures.

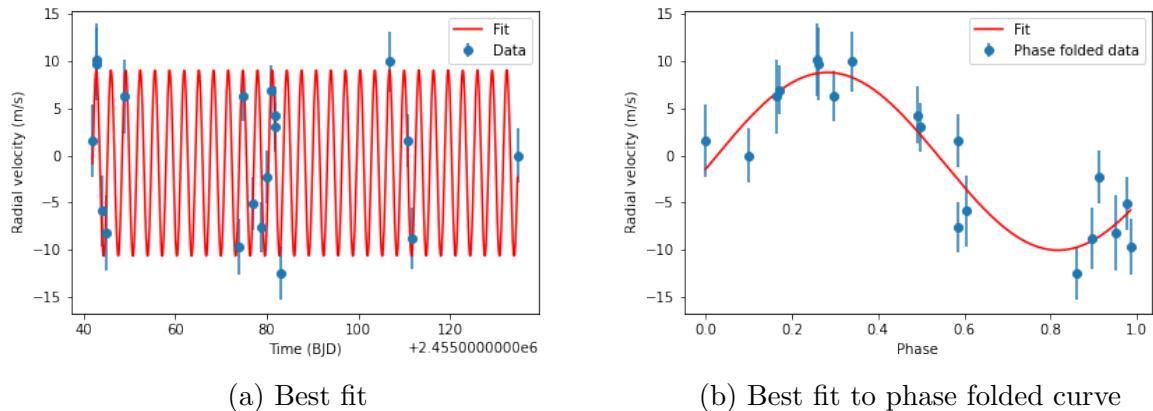


Figure 3.8: Least square fitting to RV data and associated phase folded curve

Below is the table where we have compared the parameters calculated by us from the best fit to phase folded plot and reference parameters are taken from exoplanet.eu. Now after it we

Table 3.3: Parameters and derived quantities for Kepler-4b

Parameters	Calculated (LS)	Reference
Epoch	2455041.990999	-
P (days)	3.200 ± 0.006	3.21346 ± 0.00022
K (m/s)	9.428 ± 1.211	9.3 ± 1.1
$M \sin i (M_J)$	0.078 ± 0.010	0.077 ± 0.012
a (AU)	0.045 ± 0.000	0.0456 ± 0.0009
e	0.059	0.00
χ^2	1.13	-
Phase	3.230 ± 0.225	-
Offset (m/s)	-0.636 ± 0.742	-

have shown the residual plot and lomb scargle to residual to check for further presence of any dominant peak which could indicate the presence of any additional companion and we found none. Thus we conclude that this star has only one companion.

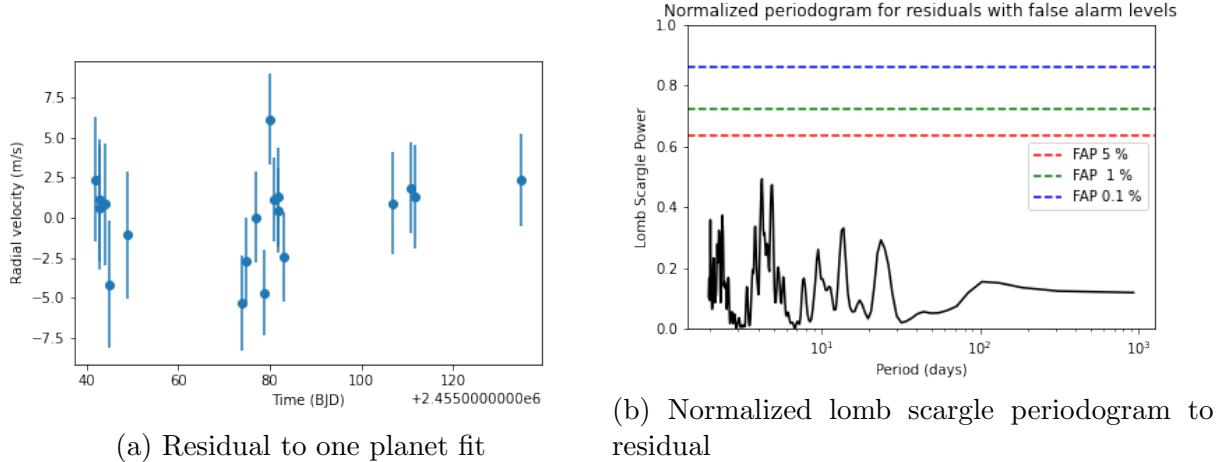


Figure 3.9: Residual and its LS periodogram

3.5 KIC 6922244

The planetary companion to this star was discovered by Kepler Mission in 2010. It is a star of $V=13.9$ magnitude and of spectral type F5V with mass of $1.213 M_{\odot}$. Although we have already discussed about this in transit method but here we are also providing a radial velocity analysis also. Below are the RV observations and normalized Lomb Scargle periodogram. We can see that here due to less number of points ,there is no dominant peak visible but from transit we know there is a planetary companion of period 3.52 days and we are getting that peak having power greater than all in the plot below with a false alarm probability of 14.8 %

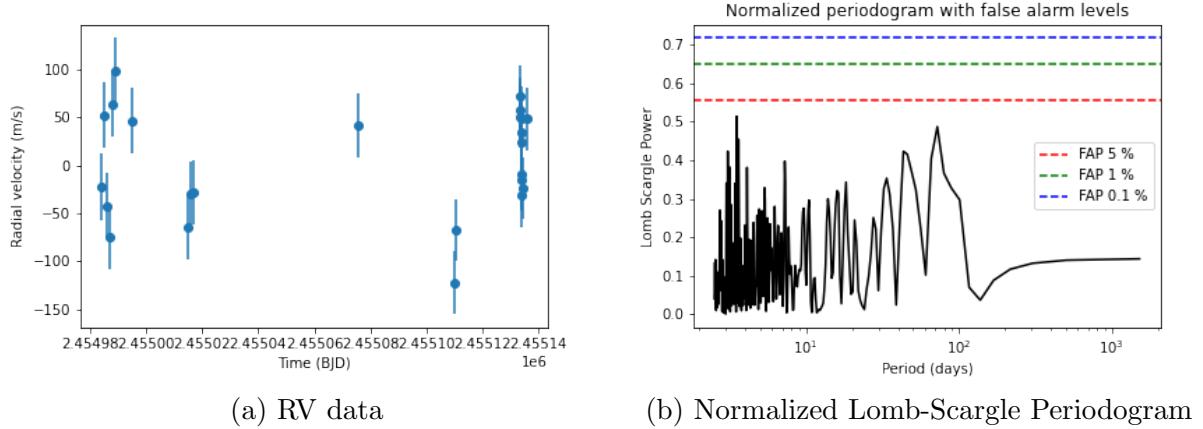


Figure 3.10: Radial velocity data and normalized Lomb-Scargle periodogram

Then we will repeat the process done for previous single planet stars to find out the best fit and associated parameters as shown in the below figures.

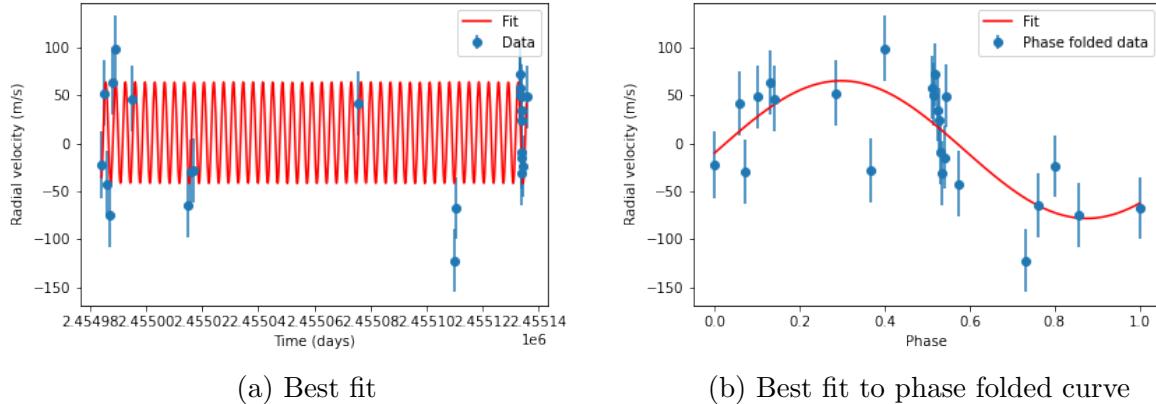


Figure 3.11: Least square fitting to RV data and associated phase folded curve

Below is the table where we have compared the parameters calculated by us from the best fit to phase folded data and reference parameters taken from exoplanet.eu and wikipidea. Now

Table 3.4: Parameters and derived quantities for Kepler-8b

Parameters	Calculated (LS)	Reference
Epoch	2454984.040029	-
P (days)	3.507 ± 0.009	$3.522 \pm e-07$
K (m/s)	71.696 ± 14.558	68.4 ± 12
M sin i (M_J)	0.616 ± 0.125	0.59 ± 0.12
a (AU)	0.048 ± 0.000	0.0474 ± 0.0019
e	0.237	0.00
χ^2	1.47	-
Phase	-0.049 ± 0.328	-
Offset (m/s)	-6.482 ± 12.272	-

after it we have shown the residual plot and lomb scargle to residual to check for further presence of any dominant peak which could indicate the presence of any additional companion and we found none. Thus we conclude that this star has only one companion.

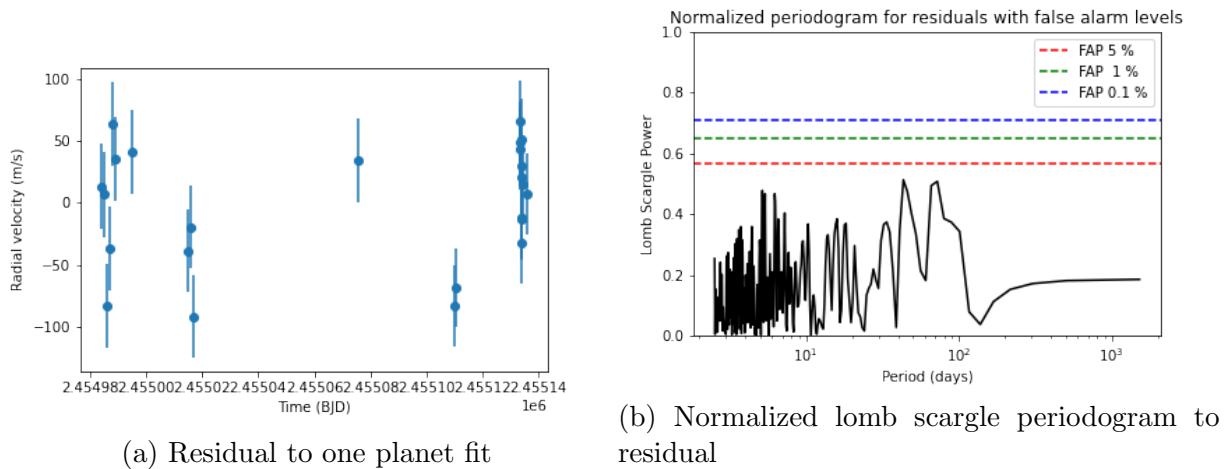


Figure 3.12: Residual and its LS periodogram

3.6 HD 177830

Now we will move to stars which are reported to have two planetary companions. We will again approach to these in a similar manner as to one-planet systems. HD 177830 is a V=7.177 magnitude star of spectral type K0IV. Below there are plots of RV observation and corresponding LS periodogram indicating a dominant peak at period of 404.098 days with FAP nearly zero. This period was also found from period04 which also gave a peak in power spectrum at period of 408.928 ± 4.384 days.

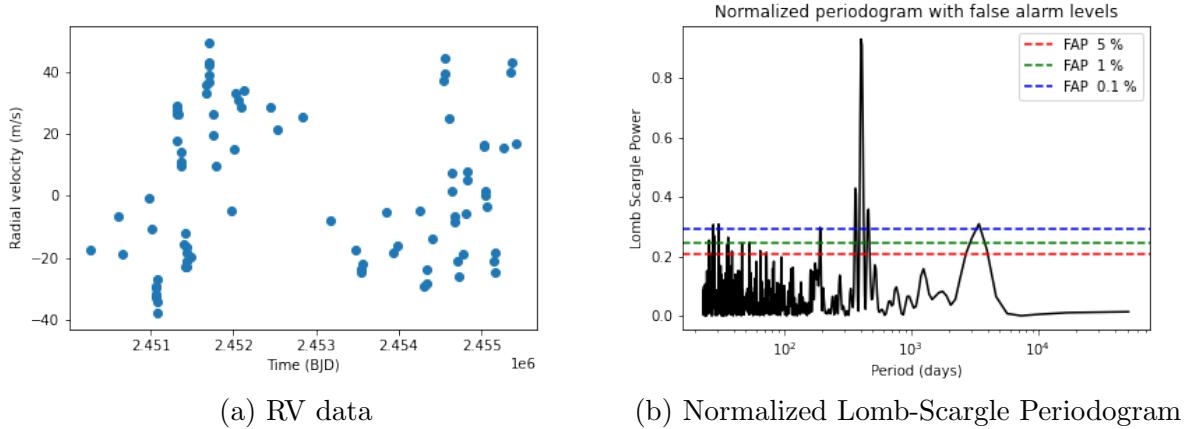


Figure 3.13: Radial velocity data and normalized Lomb-Scargle periodogram

Now we found out the one planet best fit to the data using curve fit and listed the best fit parameters along with uncertainties in the values.

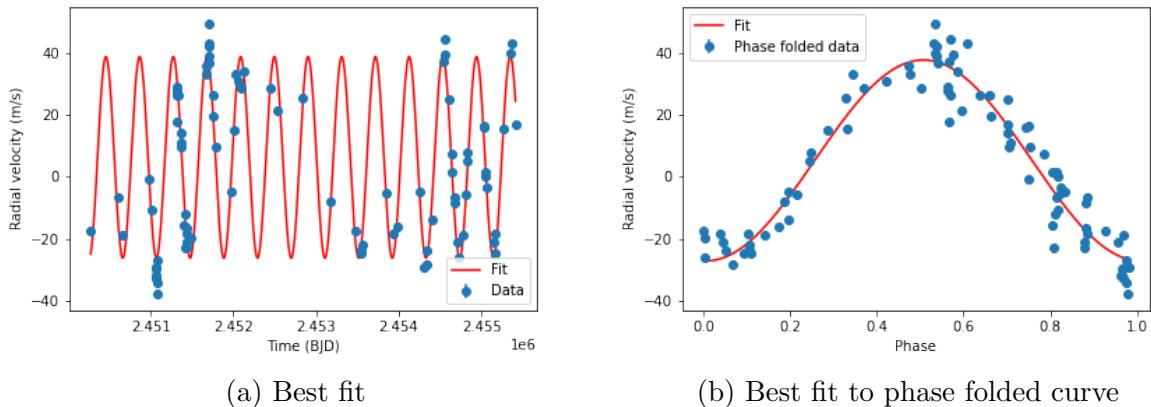


Figure 3.14: Least square fitting to RV data and associated phase folded curve

The table below lists the parameters from best fit and reference one then we will look for residuals and check for any indication of additional planets by again calculating power spectrum for residuals to one planet fit. The errors in period04 are calculated using Monte carlo simulations.

Table 3.5: Parameters and derived quantities for 177830b

Parameters	Calculated (LS)	Calculated (P04)	Reference
Epoch	2450276.02388	2450276.02388	2450276.02388
P (days)	406.980 ± 0.450	408.928 ± 4.384	406.6 ± 0.4
K (m/s)	32.553 ± 0.875	32.348 ± 0.766	31.6 ± 0.6
M sin i (M_J)	1.544 ± 0.041	1.536 ± 0.037	1.49 ± 0.03
a (AU)	1.223 ± 0.001	1.227 ± 0.009	1.2218 ± 0.0008
e	0.018	-	0.009 ± 0.004
χ^2	27.19	-	27.53
Phase	92.135 ± 41.867	5.520 ± 0.027	-
Offset (m/s)	6.310 ± 0.632	6.309 ± 0.633	-

Figure 1.15 consists one planet residual and their LS periodogram which again indicate another planetary companion with FAP very low and period of nearly 110 days. Period04 also indicated an additional companion with period of 110.909 ± 0.409 days. We again repeated

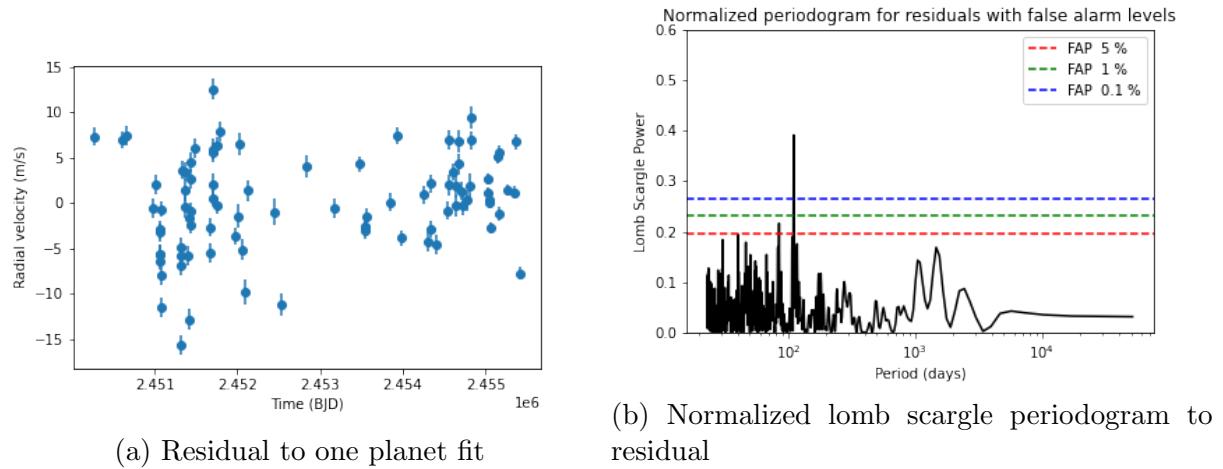


Figure 3.15: Residual and its LS periodogram

same curve fitting process to achieve best fit to residuals to one planet fit which we have shown in figure 1.16 (χ^2 of 16.28). The table below shows the parameters of second planet along with uncertainties and further we have plotted the residual to two planet fit and coressponding LS periodogram as shown in figure 1.17. The errors in period04 are calculated using Monte carlo simulations. It is quite clear that this star does not have any other additional planetary companion with no dominant peaks in periodogram to two planet residuals.

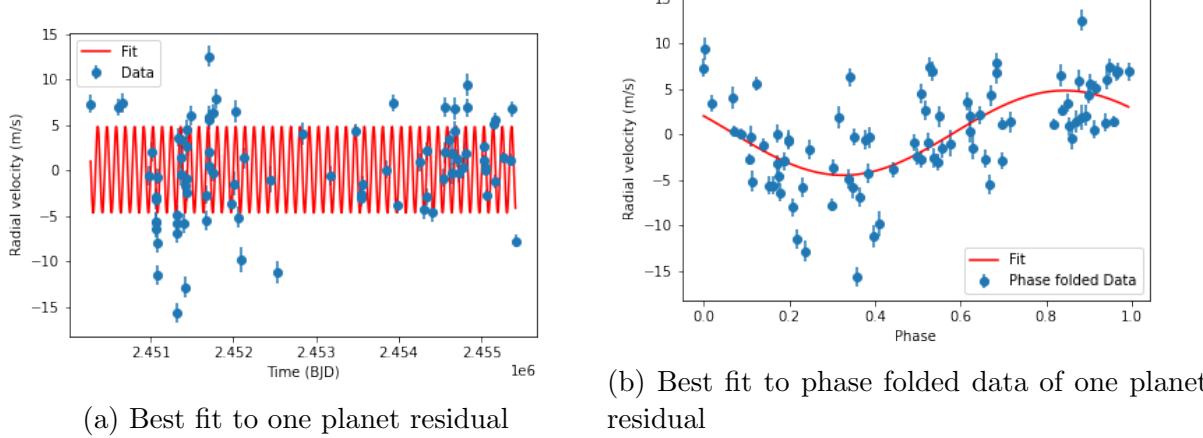


Figure 3.16: Least square fitting to RV data and associated phase folded curve

Table 3.6: Parameters and derived quantities for 177830c

Parameters	Calculated (LS)	Reference
Epoch	2450276.02388	2450276.02388
P (days)	111.072 ± 0.159	110.9 ± 0.1
K (m/s)	4.751 ± 1.133	5.1 ± 0.8
$M \sin i (M_J)$	1.514 ± 0.027	1.49 ± 0.03
a (AU)	0.146 ± 0.238	0.15 ± 0.02
e	0.016	0.3 ± 0.1
χ^2	16.28	15.31
Phase	1351.814 ± 198.874	-
Offset (m/s)	0.078 ± 0.445	-

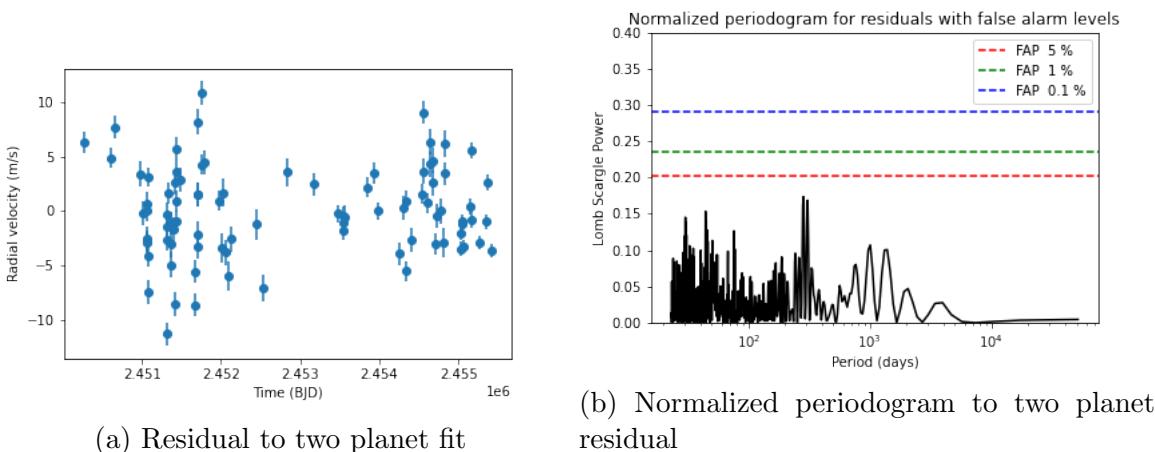


Figure 3.17: Residual and its LS periodogram

Summary and Conclusion

In this project, we have explored two methods to find exoplanets. Transit is being studied using Kepler data. Since Kepler mission was designed to detect earth-size exoplanets, highly precise photometry was required so there was need to apply different correction before making any conclusion. Light curve are obtained from target pixel file and characterisation of the transit using BLS method was being done to determine various planet parameter for 4 Kepler objects and 1 k2 object. This method gave us insight various aspects associated with photometric data. The radial velocity analysis of the selected stars provides valuable insights into their motion and characteristics. By measuring the changes in the star's spectral lines, we were able to determine the velocity of the star relative to our observation point. This information allows us to study the star's orbit, mass, and even the presence of planets around it. The results of our analysis indicate that some of the stars exhibited periodic variations in their radial velocities, suggesting the presence of planets in orbit around them. This project highlights the importance of radial velocity analysis in the study of stars and their planetary systems. Further research in this field may lead to new discoveries and a better understanding of the universe.

Appendix A

Radial Velocity Data

Below are the tables with columns as Barycentric Julian date, radial velocity and uncertainty in radial velocity for stars HD 31253, HD 218566, KIC 11853905, KIC 6922244 and HD 177830. These data sets are obtained from reference paper Meschiari_2011_ApJ_727_117 which was further taken from HIRES at Keck Observatory.

Table A.2: Radial Velocity Measurements of HD 218566

BJD	RV (m/s)	Uncertainty in RV (m/s)
2450366.85498	5.12	1.16
2450666.08942	-5.42	1.21
2450690.03236	-5.31	1.24
2450714.98919	2.30	1.24
2450715.94463	3.55	1.16
2450983.10118	2.79	1.13
2451012.03640	5.50	1.37
2451050.94400	1.57	1.21
2451071.96025	-1.72	1.19
2451343.03657	-10.72	1.22
2451369.03547	-9.47	1.24
2451410.98964	-1.71	1.28
2451440.89265	-3.58	1.36
2451552.72759	-4.34	1.46
2451900.73713	4.62	1.15
2452096.07187	-0.28	1.29
2452242.72507	-5.34	1.44
2452488.04912	-4.37	1.55
2452535.87165	2.10	1.25
2452575.73785	1.82	1.61
2452806.11771	9.98	1.49
2452828.95884	7.04	1.34
2452898.94265	-2.18	1.32
2453195.98598	1.34	1.31
2453303.88246	11.48	1.21
2453603.06616	-10.68	1.35
2453969.05624	-0.75	1.28
2454279.10762	-5.84	1.06
2454286.09348	-3.99	1.48

Continued on next page

Table A.2 – *Continued from previous page*

BJD	RV (m/s)	Uncertainty in RV (m/s)
2454295.07876	-3.03	1.28
2454429.76038	6.83	1.17
2454430.74580	6.25	1.25
2454455.76108	-0.38	1.23
2454456.75051	0.33	1.31
2454457.73398	-2.62	1.36
2454460.74730	-2.79	1.42
2454461.77113	-0.27	1.12
2454634.09288	16.93	1.31
2454635.06583	13.45	1.13
2454636.05577	13.70	1.28
2454637.11095	12.04	1.39
2454638.06397	15.78	1.21
2454639.07829	13.78	1.24
2454641.10707	12.59	1.16
2454642.12403	8.46	1.31
2454674.93117	-1.84	1.27
2454688.94686	0.27	1.21
2454689.96737	-1.23	1.44
2454717.93868	-7.51	1.30
2454719.95623	-4.29	1.27
2454807.79837	6.87	1.26
2454838.77427	6.45	1.32
2454846.72884	8.46	1.78
2455202.76214	-5.91	1.07
2455371.00493	-6.44	1.25
2455408.11285	-11.53	1.17

Table A.4: Radial Velocity Measurements of HD 177830

BJD	RV (m/s)	Uncertainty in RV (m/s)
2450276.02388	-17.61	0.97
2450605.04340	-6.49	0.97
2450666.88551	-18.82	1.03
2450982.93951	-0.89	1.11
2451009.93206	-10.55	1.10
2451068.81716	-31.85	0.99
2451069.85005	-31.82	1.07
2451070.89530	-29.51	0.97
2451071.83118	-32.61	1.05
2451072.82007	-29.18	1.04
2451073.81795	-26.90	0.90
2451074.80785	-34.21	0.99
2451075.89814	-37.65	1.10
2451311.10969	29.04	1.12
2451312.10751	27.79	1.16
2451313.10582	17.70	1.07

Continued on next page

Table A.4 – *Continued from previous page*

BJD	RV (m/s)	Uncertainty in RV (m/s)
2451314.12858	26.21	1.09
2451341.95441	26.60	0.99
2451367.91448	14.22	1.06
2451368.90654	9.80	1.20
2451369.91809	11.16	1.26
2451409.84682	-15.43	1.13
2451410.80182	-22.85	1.10
2451411.79910	-12.17	1.10
2451438.74187	-21.18	0.98
2451439.76021	-22.98	0.93
2451440.86979	-16.38	1.08
2451441.72339	-18.55	1.07
2451488.72274	-19.79	1.06
2451679.05640	36.15	1.03
2451680.09782	33.30	1.04
2451703.07726	43.09	1.14
2451704.01943	49.50	1.15
2451705.06457	38.87	1.17
2451706.03088	42.20	1.07
2451707.09076	37.01	1.02
2451754.88242	26.59	1.07
2451755.95107	19.59	0.91
2451792.76348	9.67	1.09
2451972.15641	-4.89	0.90
2452008.12955	15.20	1.35
2452031.09883	33.18	1.27
2452061.99221	30.93	1.20
2452094.87498	28.88	1.38
2452128.88145	34.21	1.19
2452445.97578	28.67	1.42
2452536.83488	21.32	1.23
2452832.84731	25.64	1.29
2453180.04734	-7.97	1.08
2453479.09162	-17.49	0.82
2453546.91590	-24.74	0.88
2453550.00554	-23.82	1.07
2453551.08720	-23.90	0.89
2453552.03093	-22.15	1.01
2453842.11990	-5.14	0.95
2453927.89270	-18.56	0.93
2453982.89128	-16.11	0.82
2454250.06746	-4.81	1.04
2454309.06516	-29.23	1.10
2454337.79715	-23.69	1.00
2454343.76189	-28.09	0.92
2454396.73487	-13.82	1.07
2454546.12536	37.14	0.93

Continued on next page

Table A.4 – *Continued from previous page*

BJD	RV (m/s)	Uncertainty in RV (m/s)
2454547.13854	44.74	1.06
2454549.12824	39.55	1.16
2454601.10229	24.93	0.96
2454634.02218	7.46	1.18
2454641.86106	1.34	1.12
2454673.06538	-8.41	1.16
2454674.85466	-6.67	1.20
2454702.93200	-21.20	1.03
2454721.81080	-25.87	1.14
2454778.68230	-18.61	1.19
2454807.68779	-5.63	1.37
2454819.68690	5.04	1.06
2454820.70300	8.04	1.26
2455022.06175	16.10	0.65
2455024.07505	16.66	0.56
2455049.98881	1.48	0.57
2455051.99502	0.20	0.58
2455053.99931	-3.63	0.59
2455143.76106	-21.14	0.78
2455166.72651	-18.15	0.77
2455168.71841	-24.60	0.76
2455259.15406	15.67	0.61
2455341.09201	39.94	0.71
2455370.90765	42.98	0.79
2455408.07014	16.76	0.65

Table A.1: Radial Velocity Measurements of HD 31253

BJD	RV (m/s)	Uncertainty in RV (m/s)
2450838.75189	-1.44	1.97
2451043.12399	-11.98	1.61
2451073.03263	-10.12	1.45
2451170.90537	6.52	1.79
2451228.78516	11.64	1.41
2451411.13343	3.54	2.21
2451550.87137	-20.13	1.59
2451581.85884	-4.86	1.71
2451757.13309	7.45	1.61
2451793.11793	8.50	1.76
2451883.00052	2.05	1.80
2451884.08282	-0.63	1.78
2451898.01005	6.44	1.65
2451898.99796	-2.59	1.57
2451899.99476	1.53	1.48
2451901.00753	-1.13	1.49
2451973.74609	-7.57	1.83
2451974.76077	-7.26	1.67
2452235.85335	1.28	1.62
2452536.09260	0.00	1.53
2452575.99470	8.16	1.77
2452898.09851	-11.51	1.48
2453241.13429	5.68	1.53
2453338.87214	-6.23	1.39
2453339.99226	-6.70	1.50
2453984.09331	16.24	1.44
2454084.05142	15.15	1.52
2454130.84528	7.82	1.62
2454131.74964	12.85	1.61
2454138.73395	2.50	1.53
2454396.90266	-9.46	1.50
2454398.01742	-1.91	1.70
2454464.89612	3.27	1.61
2454491.84891	14.30	1.44
2454545.76603	15.83	1.50
2454778.98495	-7.23	2.03
2454872.85868	5.12	0.87
2455202.81016	-10.48	0.95
2455257.82898	-9.44	0.72

Table A.3: Radial Velocity Measurements of KIC 11853905

BJD	RV (m/s)	Uncertainty in RV (m/s)
2455041.990999	1.54	3.86
2455042.818474	10.14	3.79
2455042.826449	9.67	3.88
2455043.921701	-5.89	3.79
2455045.022184	-8.21	3.94
2455048.886762	6.23	3.95
2455073.807980	-9.66	2.95
2455074.796690	6.30	2.70
2455076.962912	-5.12	2.83
2455078.903107	-7.59	2.70
2455079.948260	-2.26	2.82
2455080.770697	6.96	2.62
2455081.790097	4.29	3.09
2455081.812747	3.01	2.60
2455082.965408	-12.55	2.79
2455106.794786	9.93	3.17
2455110.765944	1.56	2.84
2455111.759862	-8.81	3.22
2455134.714618	-0.01	2.86

Table A.5: Radial Velocity Measurements of KIC 6922244

BJD	RV (m/s)	Uncertainty in RV (m/s)
2454984.040029	-22.15	34.5
2454985.042856	52.45	34.5
2454986.065428	-42.39	34.0
2454987.060244	-74.37	33.6
2454988.026564	63.04	33.7
2454988.972744	98.87	33.8
2454995.103097	46.58	34.2
2455014.891326	-64.61	33.7
2455015.975326	-29.97	33.4
2455017.019858	-28.28	33.3
2455075.788330	41.65	33.9
2455109.849781	-122.00	32.8
2455110.790309	-67.42	31.9
2455133.717887	58.11	32.6
2455133.732343	51.13	32.5
2455133.747053	71.85	32.7
2455133.761161	34.51	32.8
2455133.775640	24.38	32.6
2455133.790026	-9.49	33.2
2455133.804527	-31.04	33.2
2455133.821205	-14.32	32.6
2455133.838751	49.58	32.8
2455134.739907	-23.48	31.9
2455135.789679	48.50	32.6

Bibliography

- [1] Introduction to Signal Processing by S.J. Orfanidis (SGF)
- [2] KEPLER MISSION STELLAR AND INSTRUMENT NOISE PROPERTIES REVISITED
Gilliland and et al
- [3] <https://keplergo.github.io/KeplerScienceWebsite/pipeline.html>
- [4] <https://exoplanets.nasa.gov/k2/>
- [5] The Kepler Pixel Response Function by ST Bryson
- [6] Everest :Pixel Level Decorrelation of k2 light curves by Rodrigo Luger and et.al.
- [7] <https://builtin.com/data-science/step-step-explanation-principal-component-analysis>
- [8] Kepler data characteristic handbook
- [9] Demystifying Kepler data by K.Kinemuchi and et.al.
- [10] MAST Kepler archive manual 2020
- [11] <https://docs.lightkurve.org/tutorials/>
- [12] Kepler instrument handbook
- [13] Photometry of the δ scuti star HD 40372 by S.Deb,S.K.Tiwari,H.P.Singh,T.R.Seshadri,U.S.Chaubey
- [14] A technique for extracting highly precise photometry for 2-wheeled kepler mission by
Vanderburg and et.al.
- [15] Jacob T. VanderPlas. Understanding the Lomb–Scargle periodogram. Publications of the
Astronomical Society of the Pacific, 130(985):054501, 2018. Received 2017 August 26; revised
2017 December 6; accepted 2017 December 7; published 2018 May.
- [16] Michael Perryman. The Exoplanet Handbook. Cambridge University Press, Cambridge,
UK, 2nd edition, 2018.
- [17] Stefano Meschiari. Mcmc methods for exoplanet orbital inference. The Astrophysical Journal,
727(2):117, 2011.
- [18] Richard P. Nelson, Guillem Anglada-Escudé, and Gavin Coleman. Planet hunting with
python. Astronomy and Computing, 28:100291, 2019.
- [19] R. V. Buluev. Assessing the statistical significance of periodogram peaks. Monthly Notices
of the Royal Astronomical Society, 385(3):1279–1286, 2008. Accepted 2007 November 2.
Received 2007 June 13.

- [20] Understanding the Lomb-Scargle Periodogram by Jacob T. VanderPlas
- [21] <https://docs.astropy.org/en/stable/timeseries/lombscargle.html>
- [22] <https://docs.astropy.org/en/stable/timeseries/bls.html>
- [23] <http://neurocline.github.io/dev/2016/02/22/barycentric-coordinate-time.html>
- [24] Time Series Analysis and its Applications-Robert H.Shumway and David S. Stuffer
- [25] <https://otexts.com/fpp2/stationarity.html>
- [26] <https://www.ibm.com/in-en/topics/exploratory-data-analysis>
- [27] Time Series Analysis by James D. Hamilton
- [28] <https://online.stat.psu.edu/>
- [29] Modern Statistical Methods for Astronomy by Erin D. Feigelson and G.Josgesh Babu
- [30] <https://www.researchgate.net/figure/The-comparison-between-Kepler-short-cadence-top-panel-and-long-cadence-bottom-panelfig1338989591>
- [31] <https://www.freecodecamp.org/news/the-least-squares-regression-method-explained/>