



SENIOR THESIS IN MATHEMATICS

Detecting Rotation Periods of
Near-Earth Asteroids: An
investigation of Fourier
Analysis and the Lomb-Scargle
Periodogram

Author:
William Gray

Advisor:
Dr. Johanna Hardin

Submitted to Pomona College in Partial Fulfillment
of the Degree of Bachelor of Arts

December 14 2022

Contents

1	Periodic Signals in Rotating-Asteroid Data	1
1.1	Near Earth Object Observations	1
1.2	Rotating Asteroids	3
1.2.1	Yarkovsky Effect	3
1.2.2	Asteroid Light Curves	4
1.3	Rotation Period Computation Technique	5
2	Fourier Analysis	7
2.1	The Fourier Series	7
2.1.1	The Origin of the Series	7
2.1.2	Statement of the Fourier Series	8
2.1.3	Demonstration of the Fourier Series	8
2.1.4	Fourier Series Coefficients	10
2.2	Fourier Transform	13
2.2.1	Statement of the Fourier Transform	13
2.2.2	Windowing Functions and the Discrete Fourier Transform	14
2.3	Non-Uniform Data Problem	20
2.4	Lomb-Scargle Periodogram	22
2.4.1	Original Equation Statement	23
3	Exploration Through Analysis of Catalog Data	25
3.1	Methodology	25
3.2	Results and Discussion	26
3.3	Conclusion	28
3.4	Acknowledgements	29

Chapter 1

Periodic Signals in Rotating-Asteroid Data

1.1 Near Earth Object Observations

The number of Near Earth Object (NEO) observations has grown exponentially over the past two decades. NEO research gained momentum in 1998, when asteroid 1997 XF11 was observed and projected to hit earth in 2028. Fortunately, further observations of the target contradicted the suspicions, but the event prompted a congressional request given to NASA. The request tasked NASA with cataloging 90% of NEOs larger than 1km by 2008. Since then, NASA and worldwide observers have continuously grown the NEO catalogs [NASA, 2018].

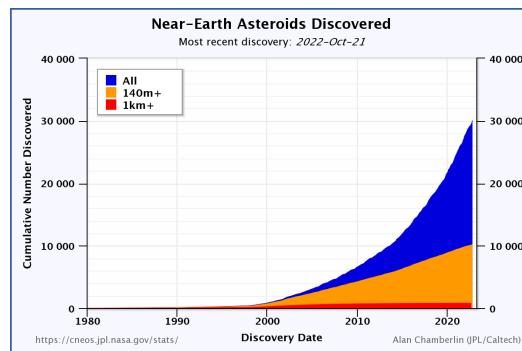


Figure 1.1: Near-Earth Asteroid Discoveries over time. From: Alan Chamberlin (JPL/Caltech), <https://cneos.jpl.nasa.gov/stats/totals.html>

As of October 21st 2022, a total of 30,353 NEOs have been discovered and cataloged by the NASA NEO program [NASA, 2022c]. Pomona College has been a contributor to these observation efforts, collaborating with NASA JPL, observing using Pomona’s 1-meter telescope located on Table Mountain. I have been a member of the observing and research team on this project since my sophomore year under the supervision of Professor Philip Choi in the Pomona College Physics and Astronomy Department. The observing projects were a primary motivation for pursuing my thesis topic.



Figure 1.2: Pomona College’s 1-Meter Telescope on Table Mountain, CA. From: <https://www.pomona.edu/academics/departments/physics-and-astronomy/facilities/table-mountain-observatory>

Just as what prompted the beginning of NEO observations in 1998, one fundamental mission for observing NEOs is for the protection of earth — by cataloging and taking repeated observations of an object, we are able to accurately predict the orbit for the object over a long period of time. Excitingly, this mission was in many ways re-confirmed and validated by NASA’s DART project on September 26th 2022, demonstrating the possibility of human’s altering the orbit of a celestial body, in this case by impacting it with a spacecraft [NASA, 2022b].

As techniques for studying NEOs continue to improve, we are able to understand more about the different movement of these asteroids. One specific movement we care about is the rotation of the asteroids, which will be the primary focus of this thesis.

1.2 Rotating Asteroids

Rotating asteroids are of particular interest in NEO observations currently because their rotation can lead to changes in their orbits. Therefore, if we are to maintain commitment towards knowing the long-term orbit of an object, understanding its rotation effect is critical.

1.2.1 Yarkovsky Effect

One dynamic that complicates our understanding of the long-term orbit of rotating NEOs — particularly small, Rotating Near-Earth Asteroids (NEAs) — is the Yarkovsky effect. The effect describes how the sun’s radiation on a rotating object can change its orbit over time.

To demonstrate the Yarkovsky Effect, consider the illustration in Figure (1.3). Notice in particular the day-time side of an object (pointing towards the sun) and a night-time side of an object (pointing away from the sun). On the day-time side, the object is absorbing radiation, while on the night-time side, the object is emitting radiation. However, for a rotating body, there is a discrepancy between the direction of radiation as the direction of greatest emission is not opposite to that of greatest absorption. As a result, there is a net force on the object that pushes it off of its original orbit. [Vokrouhlický et al., 2015].

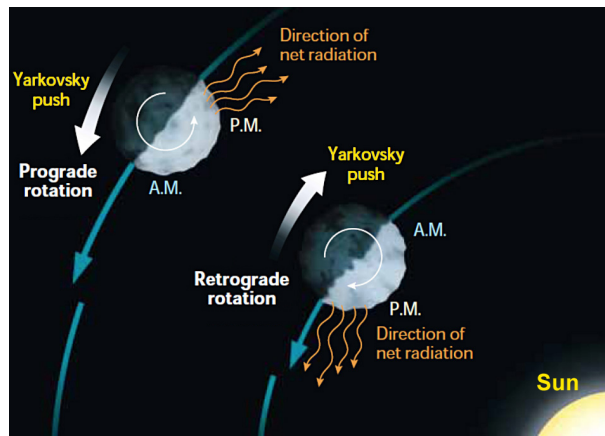


Figure 1.3: Illustration of the Yarkovsky Effect: From <https://skyandtelescope.org/astronomy-news/apophis-pays-a-visit-this-week/>

Because the Yarkovsky effect is primarily due to rotation and radiation absorption/emission, the three primary characteristics of an asteroid that will alter the Yarkovsky effect are its shape, rotation and thermal conductivity. The characteristic we are able to study at Pomona, and from most ground-based observations, is the rotation of the NEAs, and therefore the rotation characteristic will be the primary interest of this thesis.

1.2.2 Asteroid Light Curves

The primary method for determining the rotation of the NEAs is by extracting rotation characteristics from the light curves of the object. A light curve can be generated in many ways, but the method that is used for the Pomona College ground-based telescope is aperture photometry on images. Without including too much detail, this method takes the photon counts for some specified section of the sky (your aperture). There is also background subtraction done with an outer dannulus in order to increase the signal to noise ratio. An image example is shown in Figure 1.4.

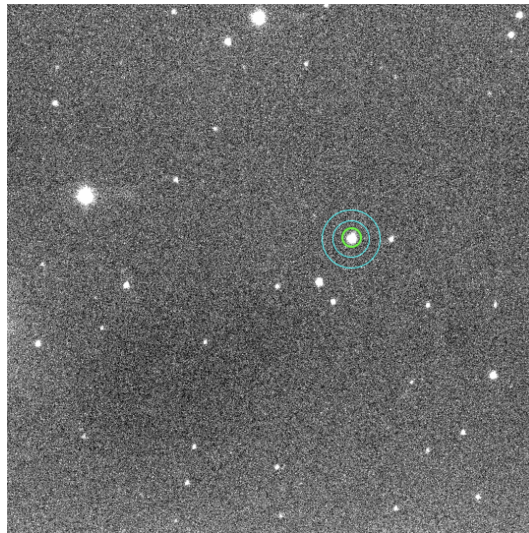


Figure 1.4: Example data image, with example aperture in green. The area between blue lines is the described region that is used for background subtraction. Picture taken by Pomona College’s Table Mountain 1-meter telescope.

Because the majority of asteroids are non-circular, the amount of reflected

light (and therefore photon counts in the aperture) will change based on the surface area visible to earth at any given time. Comparing the two orientations of a cigar-shaped asteroid in Figure 1.5 illustrates this effect, showing orientation 1 having more reflected light and orientation 2 having less reflected light.

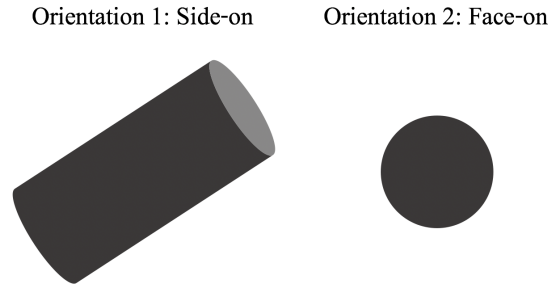


Figure 1.5: Demonstration of potential orientations of cigar-like asteroid

From Figure 1.5, orientation 1, with its increased surface area, will reflect more visible light than orientation 2. Therefore, as the asteroid rotates, the shape exposed to earth will gradually alternate between orientation 1 and orientation 2. Using images of asteroids, we can measure the variation in brightness that is caused by this gradual change of orientation of the asteroid (variation in reflected visual light). The period of these variations will be the same as the rotation period of the asteroid.

1.3 Rotation Period Computation Technique

Once the light curves of an NEA is constructed, extracting the rotation periods of the asteroid is important to understanding the nature of its movement — this is our primary goal.

Because we are handling a consistent rotating body, techniques for studying periodic signals are naturally the first instinct. In particular, the Fourier Transform is a common technique used for extracting periodic signal within data. For evenly sampled, discrete data the Discrete Fourier Transform would be a natural first analysis technique, and is often used for discrete, periodic data.

However, most data, particularly in astronomy, is not evenly sampled. Therefore, new techniques have been developed to overcome the problem of

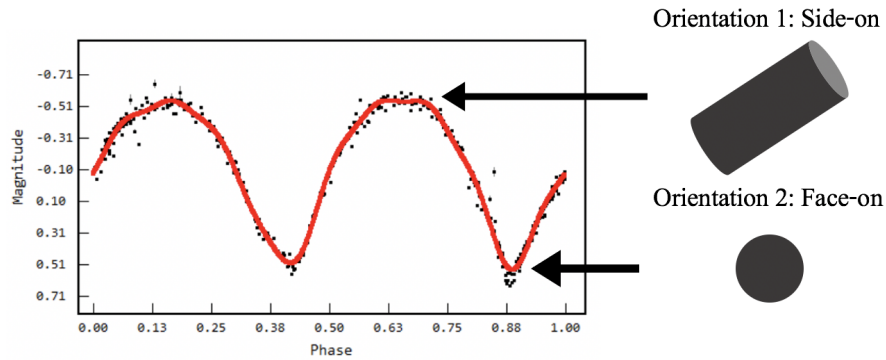


Figure 1.6: Example asteroid light curve and which parts of the curve would correspond to the possible orientations in Figure 1.5. Adapted from: Tycho user manual

uneven sampling. The Lomb-Scargle periodogram is one of these approaches that is often used in Astronomy. In particular, it is cited in the work of [Beniyama et al., 2022], which is a recent paper cataloging rotation period of fast-rotating asteroids. We hope that replicating these results is a step that can be done using our telescope, and therefore understanding their rotation period detection technique is important.

Chapter 2

Fourier Analysis

Fourier Analysis of periodic data is one of the most common techniques for extracting periodic signal within data. The Fourier Transform has its origins in physics and the dissipation of heat and rests on the assumption of that every signal can be estimated with a linear combination of sine and cosine functions — the Fourier Series.

2.1 The Fourier Series

2.1.1 The Origin of the Series

Jean-Baptiste Joseph Fourier was a French mathematician whose primary research focus was the dissipation of heat. His most famous works, "Mémoire sur la propagation de la chaleur dans les corps solides" [Fourier, 1808] and "Théorie analytique de la chaleur" [Fourier and Darboux, 1822], both tackled the question of heat transfer. In particular, Fourier was working towards a solution to the Heat Equation — the second order partial differential equation in Equation (2.1).

$$\frac{\partial u}{\partial t} = \alpha^2 \frac{d^2 u}{dx^2} \tag{2.1}$$

Prior to Fourier, there had been no solution to Equation (2.1). In "Mémoire sur la propagation de la chaleur dans les corps solides" [Fourier, 1808], Fourier explored the idea of a series in which all periodic functions could be expressed as a linear combination of sine and cosine functions. This series allowed for

a decomposition of a complex function, which then had a simpler PDE calculation and contributed to solving the Heat Equation in Equation (2.1) [Debnath, 2012]

This linear combination became known as the Fourier Series, and is the foundation for many modern time-series analysis techniques — particularly the Fourier Transform. The Fourier series is one of the building blocks for the explanation of Fourier Analysis and the Lomb-Scargle periodogram.

2.1.2 Statement of the Fourier Series

The Fourier series suggests that every periodic function can be constructed by a linear combination of sine and cosine functions. Using this assumption, we can take some arbitrary function $f(t)$ and express it as the sum:

$$\begin{aligned} f(t) &= a_0 + a_1 \cos(1\omega_0 t) + b_1 \sin(1\omega_0 t) + a_2 \cos(2\omega_0 t) + b_2 \sin(2\omega_0 t) \dots \\ &= a_0 + \sum_{n=1}^{\infty} (a_n \cos(n\omega_0 t) + b_n \sin(n\omega_0 t)) \end{aligned} \quad (2.2)$$

where a_0 is the mean value of the function, each a_n and b_n are the amplitudes of the underlying cosine and sine functions respectively, and ω_0 is the fundamental frequency. This fundamental frequency is defined as $\omega_0 = \frac{2\pi}{T}$ where T is the duration of the observation window. For example, for some signal along the closed interval of time $[0, 2\pi]$, the fundamental frequency is $\omega_0 = \frac{2\pi}{2\pi-0} = 1$, which can also be thought of as representing a curve that goes through one period in our observation window T [Cheever, 2022].

2.1.3 Demonstration of the Fourier Series

A helpful way to grasp the concept of the Fourier series is with the concept of a musical chord. We will consider some chord played on the piano consisting of four notes. We then take a measurement of the sound produced by the piano and create a graph. The graph may look something like Figure 2.1.

At first glance, to someone who did not know that the signal was produced by a piano, the signal may be difficult to extract information from — much like how data we are interested in studying may be. However, we know that the signal was produced by four different notes — each of which has their own frequency.

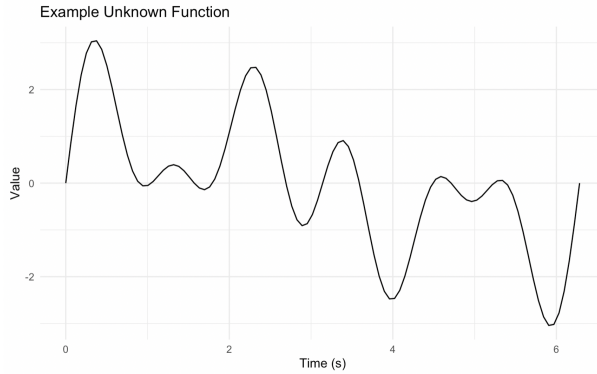


Figure 2.1: Example Chord Function

If we consider each of the notes individually, assuming the signal is sinusoidal and each note is played at the same volume, then we can interpret our chord as the sum of the individual notes.

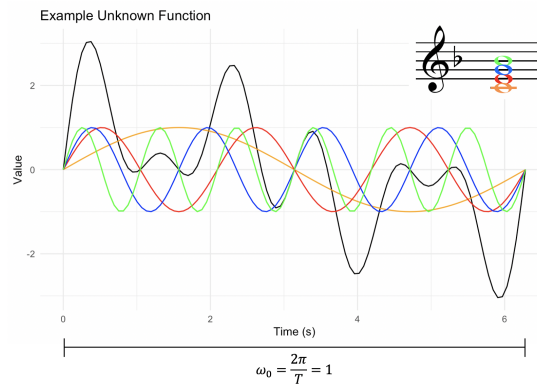


Figure 2.2: Example Chord Function With Associated Notes. Note that the example chord does not actually produce these sign waves in reality — the exact chord is being used here as an analogy.

More precisely, we let one note have frequency ω , another have frequency 3ω , another have frequency 4ω , and the last note have frequency 6ω (Note that with the overall periodic signal being 2π , our fundamental frequency is $\omega = 1$). Therefore, the chord, and the original signal, can be written as a linear combination of the four sine functions for the notes. In other words,

the image in Figure 2.2 is given by Equation (2.3).

$$\begin{aligned} f(t) &= 0 + 1\sin(1\omega_0 t) + 1\sin(3\omega_0 t) + 1\sin(4\omega_0 t) + 1\sin(6\omega_0 t) \\ &= \sin(t) + \sin(3t) + \sin(4t) + \sin(6t) \end{aligned} \quad (2.3)$$

2.1.4 Fourier Series Coefficients

In order to transition from an infinite sum, like that of Equation (2.2), we must consider the coefficients a_n and b_n as $n \rightarrow \infty$.

$$\begin{aligned} f(t) &= a_0 + a_1\cos(1\omega_0 t) + b_1\sin(1\omega_0 t) + a_2\cos(2\omega_0 t) + b_2\sin(2\omega_0 t) \\ &\quad + a_3\cos(3\omega_0 t) + b_3\sin(3\omega_0 t) + a_4\cos(4\omega_0 t) + b_4\sin(4\omega_0 t) \\ &\quad + a_5\cos(5\omega_0 t) + b_5\sin(5\omega_0 t) + a_6\cos(6\omega_0 t) + b_6\sin(6\omega_0 t) \dots \end{aligned} \quad (2.4)$$

Specifically, we consider if a_n and b_n are zero or non-zero. Taking some of these coefficients to be 0 and others positive, we can cancel out certain sine and cosine functions.

$$\begin{aligned} f(t) &= a_0 + \cancel{a_1\cos(1\omega_0 t)} + b_1\sin(1\omega_0 t) + \cancel{a_2\cos(2\omega_0 t)} + \cancel{b_2\sin(2\omega_0 t)} \\ &\quad + \cancel{a_3\cos(3\omega_0 t)} + b_3\sin(3\omega_0 t) + \cancel{a_4\cos(4\omega_0 t)} + b_4\sin(4\omega_0 t) \\ &\quad + \cancel{a_5\cos(5\omega_0 t)} + \cancel{b_5\sin(5\omega_0 t)} + \cancel{a_6\cos(6\omega_0 t)} + b_6\sin(6\omega_0 t) \dots \end{aligned} \quad (2.5)$$

For $a_0 = 1$, $b_1 = b_3 = b_4 = b_6 = 1$ and $\omega_0 = 1$, we notice that the original example function in Equation (2.3) is extracted.

$$f(t) = 0 + \sin(t) + \sin(3t) + \sin(4t) + \sin(6t) \quad (2.6)$$

The calculation of the a_n and b_n coefficients within the function $f(t)$ is critical to understand the underlying frequencies and how they influence the signal. This can be done using the real trigonometric form of the Fourier Series in Equation (2.2), resulting in

$$a_n = \frac{2}{T} \int_T f(t) \cos(n\omega_0 t) dt \quad (2.7)$$

$$b_n = \frac{2}{T} \int_T f(t) \sin(n\omega_0 t) dt \quad (2.8)$$

However, considering the exponential form of the Fourier Series will be helpful in consolidating these coefficients into one term and to generalize to imaginary functions.

Consider the exponential form of the Fourier Series:

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{in\omega_0 t} \quad (2.9)$$

where

$$c_n = \frac{a_n}{2} - i \frac{b_n}{2} \quad (2.10)$$

A proof that the exponential form in Equation (2.9) and the trigonometric form in Equation (2.2) are equivalent is needed. The following proof contains the general structure of a similar proof in [Herman, 2014].

Proof. We can begin with either equation and show equivalency to the other. For this proof, begin with Equation (2.2).

So assume

$$f(t) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(n\omega_0 t) + b_n \sin(n\omega_0 t))$$

Using Euler's Formulas

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}$$

$$\sin(x) = \frac{e^{ix} - e^{-ix}}{2i}$$

We can re-write $f(t)$ as:

$$f(t) = a_0 + \sum_{n=1}^{\infty} a_n \left(\frac{e^{in\omega_0 t} + e^{-in\omega_0 t}}{2} \right) + b_n \left(\frac{e^{in\omega_0 t} - e^{-in\omega_0 t}}{2i} \right)$$

Which can be simplified by grouping like exponential terms

$$f(t) = a_0 + \sum_{n=1}^{\infty} \left(\frac{a_n - ib_n}{2} \right) e^{in\omega_0 t} + \left(\frac{a_n + ib_n}{2} \right) e^{-in\omega_0 t}$$

It should also be noted that the second term can be re-written for $-n$ terms

$$f(t) = a_0 + \sum_{n=1}^{\infty} \left(\frac{a_n - ib_n}{2} \right) e^{in\omega_0 t} + \sum_{n=-\infty}^{-1} \left(\frac{a_{-n} + ib_{-n}}{2} \right) e^{in\omega_0 t}$$

which can further be simplified as

$$f(t) = a_0 + \sum_{n=-\infty}^{\infty} c_n e^{in\omega_0 t}$$

where

$$c_n = \frac{a_n - ib_n}{2}, n = 1, 2, 3, \dots$$

$$c_n = \frac{a_n + ib_n}{2}, n = -1, -2, -3, \dots$$

Since we are focused on finding information about the coefficients, and c_n is a combination of these coefficients, it serves well to solve for c_n .

By plugging in Equations (2.7), (2.8) into the c_n formula in Equation (2.10) c_n can be solved for. Although it will not be elaborated on in this proof, [Herman, 2014] describes how this is the form for c_n for both $n = 1, 2, 3, \dots$ and $n = -1, -2, -3, \dots$

$$c_n = \frac{1}{T} \int_T f(t) e^{-in\omega_0 t} dt \quad (2.11)$$

Where T is the fundamental period and ω_0 is the fundamental frequency.

□

2.2 Fourier Transform

2.2.1 Statement of the Fourier Transform

The Fourier Transform is defined as a scalar multiple of c_n . Specifically $c_n \times T$. We let $F(\omega) = c_n \times T$, and express the Fourier Transform equation in (2.11) as:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad (2.12)$$

where $\omega = n\omega_0$.

We define another function — a periodogram — as:

$$\mathcal{P}(\omega) = \frac{1}{T} |F(\omega)|^2 \quad (2.13)$$

The periodogram allows for us to do away with complex components, and have a positive, real-valued function that peaks at the underlying ω values that appear in the Fourier Transform.

Fourier Transform Example

To work through the process of solving the Fourier Transform, consider the box function in Equation (2.14).

$$f(t) = \begin{cases} 1 & -T \leq t \leq T \\ 0 & |t| > T \end{cases} \quad (2.14)$$

Performing the Fourier Transform on Equation (2.14), we get

$$\begin{aligned} F(\omega) &= \int_{-T}^T e^{-i\omega t} dt \\ &= \frac{-1}{j\omega} (e^{-i\omega T} - e^{-i\omega(-T)}) \\ &= \frac{2\sin(\omega T)}{\omega} \end{aligned} \quad (2.15)$$

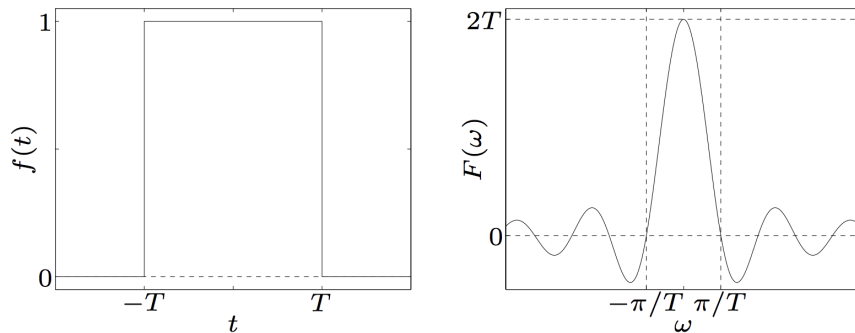


Figure 2.3: $f(t)$ and corresponding Fourier Transform $F(\omega)$. From [Stanford, 2001]

It will be useful in the coming sections to find some common Fourier pairs (functions and the resulting Fourier Transforms). Figure 2.4 shows Fourier Pairs for Sinusoid, Top Hat, and Dirac Comb functions.

2.2.2 Windowing Functions and the Discrete Fourier Transform

Real data is almost never taken continuously, so it is important to find some discrete Fourier Transform that can be performed on discrete data. Windowing functions serve as a way to visualize how this discrete data relates to the underlying signal, and convolution gives an intuition for the Fourier Transform of the Discrete Fourier Transform.

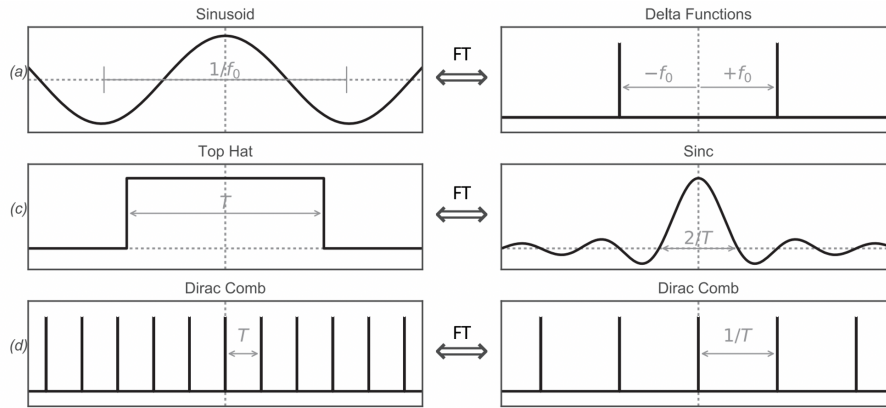


Figure 2.4: Important Fourier Pairs. Modified from: [VanderPlas, 2018]

Windowing Functions

When looking at Figure 2.4, notice that a point-wise product of the three functions would simulate discrete data. Specifically, the top-hat (or box) function would restrict the observation to some discrete time interval, and then the Dirac Comb function will create instantaneous observation times within that time interval.

Functions that are point-wise multiplied from the original signal are called windowing functions, as they create gaps of time in which the signal is observed discretely.

Introducing the Discrete Fourier Transform

A discrete function of interest can be interpreted as a point-wise product of our signal, a box function, and a dirac comb function. A dirac comb function is shown at the bottom left of Figure 2.4 — a sequence of impulses separated by some T . This Dirac Comb function will collapse the Fourier Transform integral into a sum of evenly sampled data.

$$F(\omega) = \sum_{n=-\infty}^{\infty} f(n\Delta t)e^{-i\omega n\Delta t} \quad (2.16)$$

However, because we are taking N samples, we can let $f(n\Delta t) = f_n$ with the following limits.

$$F(\omega) = \sum_{n=0}^N f_n e^{-i\omega n \Delta t} \quad (2.17)$$

The discrete sum in Equation (2.17), however, requires an additional step to become the complete Discrete Fourier Transform. This is because the function does not consider which frequency values ω are of interest, and if any should and can be rejected. The following section on convolution will help focus on particular ω values.

Convolution

A helpful property of the Fourier Transform that can give an interpretation of the Discrete Fourier Transform as well as the frequency (ω) values that we are interested in is the convolution property.

The convolution property states that the Fourier Transform of the point-wise multiplication of two functions is equal to the convolution of the Fourier Transforms of the individual functions. In mathematical form,

$$F\{f \cdot g\} = F\{f\} * F\{g\} \quad (2.18)$$

Before continuing with the definition of convolution, it is useful to keep a few qualitative motivations in mind. Primarily, recall we care about the Fourier Transform of some function that is the point-wise product of other products.

Defining Convolution

Formally defined, the convolution of two functions f and g is given by $*$ and can be formally defined by:

$$[f * g](t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau \quad (2.19)$$

In this case, τ is some pseudo variable. We use this variable to reflect our $g(\tau)$ function across the y axis and slide the $g(\tau)$ function in the τ direction. The convolution of the two functions can therefore be thought of as the original two functions sliding over one another, taking the product, and then integrating that product. Some describe this process conceptually as "using one function to smooth and average the other" [Osgood, 2007].

Figure 2.5 demonstrates an example in which $g(t)$ and $f(t)$ are two square functions. Each function is drawn in terms of the pseudo variable τ and then $g(\tau)$ is reflected across the y axis (equivalent to taking $g(-\tau)$). As t increases, $g(t - \tau)$ approaches the $f(\tau)$ function and then overlaps it. As the overlap occurs, the product is taken and then integrated across all τ .

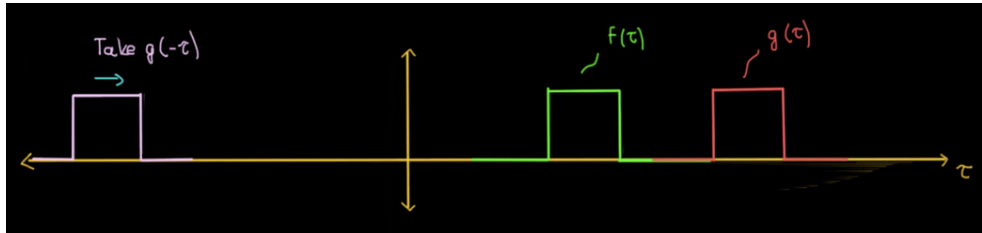


Figure 2.5: Example $g(\tau)$ and $f(\tau)$ functions and the reflected $g(-\tau)$ for box functions. From: <https://www.youtube.com/watch?v=N-zd-T17uiEt=148s>

For the example in Figure 2.5, the resulting integral (and therefore the convolution) is shown in Figure 2.6. The reason this integral is a triangle is because as t increases, and $g(t - \tau)$ begins to overlap $f(\tau)$, the integral is small, but then peaks at the value t in which both boxes are fully over one another. Because there are box functions, the area overlapping with change linearly with t , resulting in a triangle.

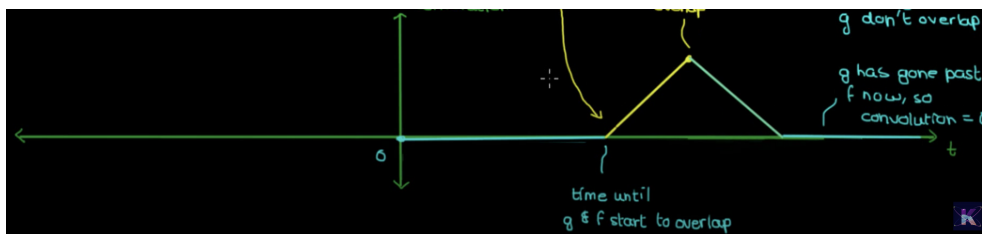


Figure 2.6: Example convolution of functions from Figure 2.5. From: <https://www.youtube.com/watch?v=N-zd-T17uiEt=148s>

Thus, as Fourier Transforms hold under convolution,

$$F\{f \cdot g\} = F\{f\} * F\{g\} \quad (2.20)$$

Convolution Interpretation of Discrete Fourier Transform

Because we now have this convolution property of Fourier Transforms, it is simple to think about what the Fourier Transform will look like for some signal with windowing functions described in Figure 2.4. Performing this convolution, we notice the result will be a sequence of sinc functions peaking at some $\frac{1}{T}$ where T is the length between each observation. In Equation (2.16), the T is Δt . This is not what we want, however, as there is only one frequency peak for the underlying sine function signal in Figure 2.4.

The multiple peaks are actually aliases of the true underlying curve under this convolution example. Because these aliases appear, we know that all signal information can be extracted for some narrower gap of $0 \leq \omega \leq 1/\Delta t$. Although this will not be heavily addressed, these ω limits also give rise to the Nyquist limit, which gives a limit to the sampling frequency that can be picked up in a signal — $\frac{1}{2}$ of the sampling frequency.

Complete Discrete Fourier Transform Equation

With the understanding of Nyquist limiting frequencies, we can re-write the discrete Fourier Transform found in Equation (2.16). We define some N evenly spaced frequencies to run through and define them as $\Delta\omega = 1/N\Delta t$. The resulting function is:

$$F(k\Delta\omega) = F_k = \sum_{n=0}^N f_n e^{\frac{-ikn}{N}} \quad (2.21)$$

Which is the most common form of the Discrete Fourier Transform.

Short Example of Discrete Fourier Transform

To demonstrate an example of the Discrete Fourier Transform, we consider the function in Figure 2.7. In this equation, we find the Fourier Transform over N frequency increments of our original function. Each x_k is the y value of our function at each k^{th} increment.

We can now go through an example for this discrete Fourier transform. For our example graph with $N=8$, we can compute the components for F_k with F_1 as

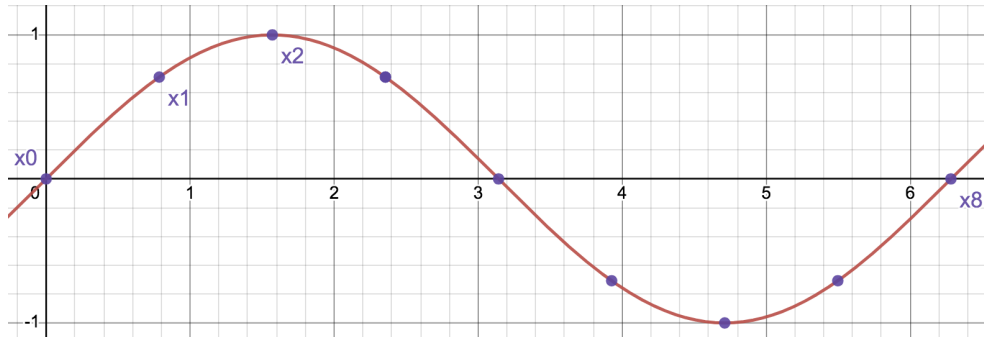


Figure 2.7: Example sine curve, with each x_1, x_2, \dots representing the y values corresponding the n th sample.

$$F_1 = 0e^{-\frac{-i2\pi(1)(0)}{9}} + 0.707e^{-\frac{-i2\pi(1)(1)}{9}} + \dots = 0 \quad (2.22)$$

each computed F_k can then be listed and plotted to form our Fourier Transform. For our example, we compute

$$F_1 = 0, F_2 = 0 - 4i, F_3 = 0, F_4 = 0, \dots \quad (2.23)$$

We find that only our F_2 is non-zero, and is therefore plotted as a component of our Fourier Transform. An example plot representing this F_n component is seen in Figure X.

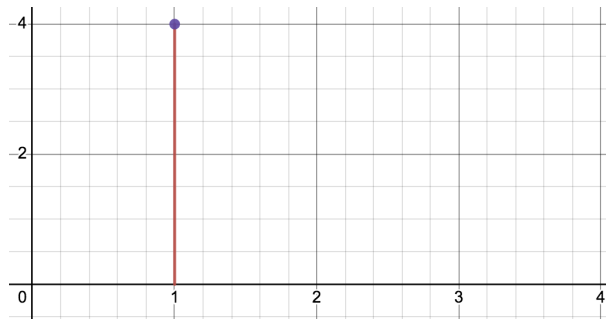


Figure 2.8: Resulting peak of Fourier Transform in Figure 2.7

As suspected, this outcome demonstrates to us that the original function had one singular sinusoidal component of frequency $\omega = 1$ or $T = 2\pi$.

2.3 Non-Uniform Data Problem

Much of the data that is found in real life, and particularly astronomy related fields are not uniformly spaced in time. This could be caused by clouds overhead, gaps of days between observations, etc. Non-uniform observations pose a problem for the Fourier Transform, due to its assumption of uniform N samples over the sample period T .

How the Fourier Transform Handles Unevenly Sampled Data

To illustrate how the Fourier Transform misinterprets non-uniform data, we can consider some simulated data of a sine curve in Figure 2.9. By removing arbitrarily selected values observations, we can simulate an unevenly-sampled data set. Figure 2.10 shows the points that were removed in red.

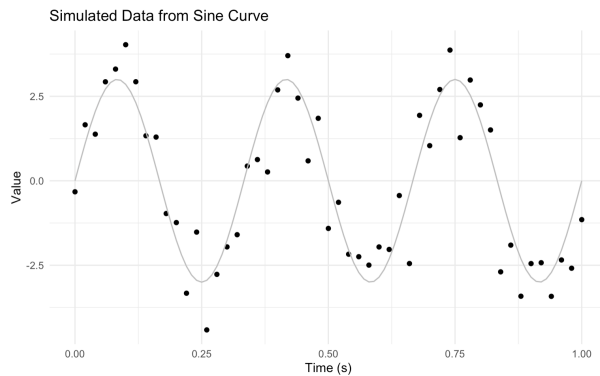


Figure 2.9: Simulated data with random noise from underlying sine function in gray.

Because the Fourier Transform assumes evenly sampled data, it will automatically interpret all the data points as evenly spaced. In fact, when using the periodogram function in R, the only requirement is a vector of values — the time component is not even considered.

More specifically, consider the original, complete data set which had 50 points over a 1 second time period. The Fourier Transform in this case will correctly assume that the 50 points are evenly spaced, assigning a Δt of 0.02 seconds. After removing the 12 red data points, there are now 38 data points available to the Fourier Transform now. So, the Fourier Transform will now wrongly assume even sampling, assigning a Δt of $\frac{1}{38}$ seconds.

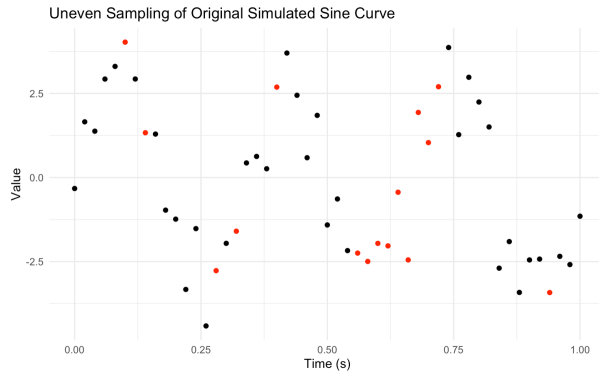


Figure 2.10: Red points indicating removed observations from the data set

It is not immediately clear how this wrong assumption will influence the Fourier Transform, so I have created a plot in Figure 2.11 to help demonstrate the problem. Figure 2.11 shows the remaining, non-removed points in two different orientation. The 38 points in black are evenly distributed along the x-axis with a distance of $\frac{1}{38}$ seconds between each point. The 38 grey points are the non-removed points in their original spacing, with uneven gaps between each point due to the missing data.

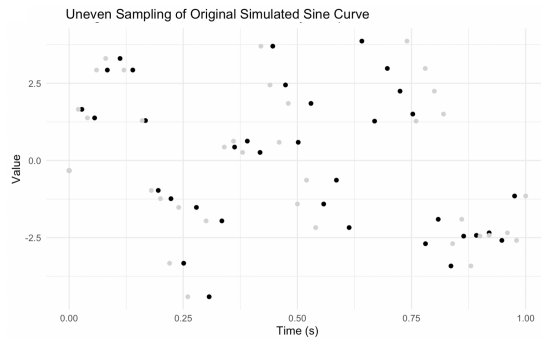


Figure 2.11: "Evenly" sampled data shifted from the unevenly sampled data

If we think about the Fourier Transform as finding the best fitting frequencies for our data, it is clear that if our data is different, then the frequencies found by the Fourier Transform will also be different or at the least smeared and noisier. Under the assumption of even sampled data, the Fourier Transform will interpret the data as different than observed, and therefore match

frequencies to an incorrect data set. Thus, the frequencies will not be correct.

For the purposes of completing this example, Figure 2.12 shows the periodogram result from performing the Discrete Fourier Transform on the unevenly sampled data, confirming the smearing as there are two frequency peaks.

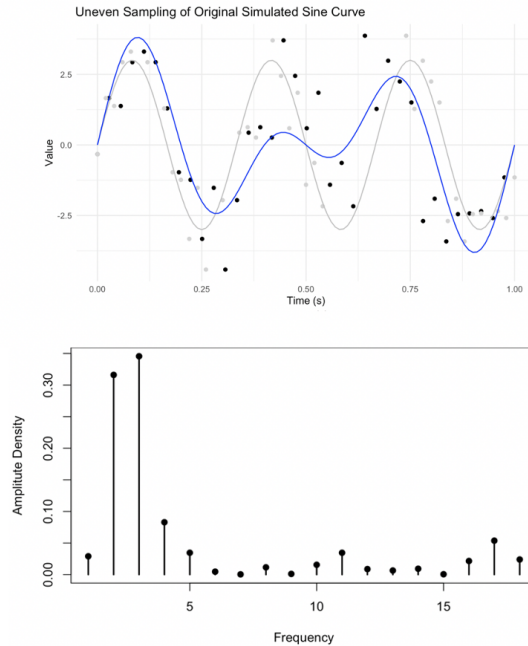


Figure 2.12: Classical periodogram result on unevenly sampled data. The blue line shows an example curve containing both peak frequencies. The gray line is the underlying signal.

2.4 Lomb-Scargle Periodogram

The Lomb-Scargle periodogram was developed by Lomb and Scargle [Lomb, 1976] [Scargle, 1982], with the motivation of handling uneven sampling in astronomy-related data sets. This method has been popular particularly in astronomy-related fields. As mentioned in the introduction, it was also the method used by a group in Japan [Beniyama et al., 2022], which presents results that are of interest to Pomona to replicate.

2.4.1 Original Equation Statement

The Lomb-Scargle periodogram is of the following form. This form considered in particular some time delay τ that allowed for the presence of unevenly sampled data. In other words, τ is responsible for the time-invariant capabilities of the periodogram [Scargle, 1982].

$$P_x(\omega) = \frac{1}{2} \left(\frac{[\sum_j f(t_j) \cos(\omega(t_j - \tau))]^2}{\sum_j \cos^2(\omega(t_j - \tau))} + \frac{[\sum_j f(t_j) \sin(\omega(t_j - \tau))]^2}{\sum_j \sin^2(\omega(t_j - \tau))} \right) \quad (2.24)$$

where τ is given by:

$$\tan(2\omega\tau) = \frac{\sum_j \sin(2\omega t_j)}{\sum_j \cos(2\omega t_j)} \quad (2.25)$$

The derivation of this Lomb-Scargle periodogram is beyond the scope of this paper, but can be found in [Scargle, 1982]. For a qualitative understanding, Scargle has chosen functions for a_n and b_n that create a generalization of the Fourier Transform and therefore the Periodogram for our uneven sampling. We can see the skeleton of a trigonometric form of the Discrete Fourier Transform in the numerators in the fraction — they are the squared discrete Fourier Transform components.

The τ value is considered in greater detail in [Scargle, 1982]. The term is added for the cases in which the underlying signal is phase shifted during our window of observation. Therefore, τ is included to create the time invariant nature of the Lomb-Scargle periodogram. As Scargle describes, τ achieves this when it holds for Equation (2.25) because Equation (2.25) says that if any t_j is shifted by T_0 (in other words, t_j becomes $t_j + T_0$, then τ becomes $\tau + T_0$, effectively giving the periodogram the same form under any shift.

Least-Squares Approach

Although the interpretation of Equation (2.24) is quite difficult to understand, Lomb describes how the resulting periodogram will have peaks corresponding to least squares values from fitting a sinusoid model to the data. In other words, we can take some function at a frequency ω

$$y(t; \omega) = A \cos(\omega(t)) + B \sin(\omega(t)) \quad (2.26)$$

At each frequency, the χ^2 statistic is taken for the model, defined as

$$\chi^2(\omega) = \sum_n (y_n - y(t_n; \omega))^2 \quad (2.27)$$

We can then find the best fit by minimizing the $\chi^2(\omega)$ values. Note that the A and B values can be calculated using sinusoid fitting methods found at [Bloomfield, 2004] for some given frequency.

Calling the minimizing value $\hat{\chi}^2(\omega)$, the periodogram is then defined as

$$P(\omega) = \frac{1}{2}[\hat{\chi}_0^2 - \hat{\chi}^2(\omega)] \quad (2.28)$$

where $\hat{\chi}_0^2$ is the non-varying reference model [VanderPlas, 2018].

Note that if there desire to learn about the equivalence of this least squares approach at that of Equation (2.24), please refer to the Appendix of Scargle [Scargle, 1982] and Lomb [Lomb, 1976].

Chapter 3

Exploration Through Analysis of Catalog Data

To complete my investigation of Fourier Analysis Using the Lomb-Scargle periodogram, the following section begins to construct practices for using the Lomb-Scargle Periodogram on real asteroid data. For this process, I use catalog data from the LCDB [NASA, 2022a]. To calculate the Lomb-Scargle Periodogram, I used the functions within the lomb package in R [Ruf, 1999].

3.1 Methodology

In this analysis, I consider three lightcurves coming from the LCDB. The asteroids of interest are numbered 503293, 508871, and 504025. Each can be searched on NASA's Small-Body Database Lookup website to find referenced, well established periods.

An example light-curve plot for the object 508871 is shown in Figure 3.1. It is important to note that each data point in this plot comes with some uncertainty in the magnitude.

To perform the periodogram analysis, I implemented a similar method to that done in [Beniyama et al., 2022]. For each target, the data were re-sampled 5000 times under the assumption that each data point's magnitude was the mean of a normal distribution with standard deviation of the magnitude error. After taking the 5000 re-samples, the average of all resulting periodogram maximum-peak frequencies was recorded as the period for that target.

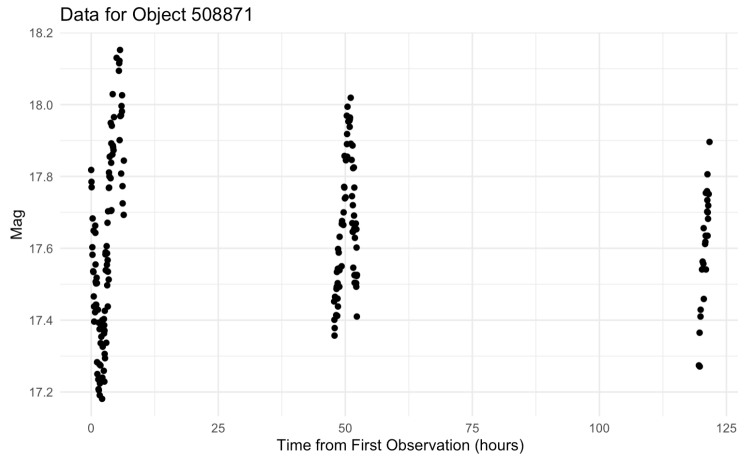


Figure 3.1: Light curve data for target 508871.

3.2 Results and Discussion

Target	LCDB Established Period	My Found Period
503293	8.8 hrs	9.144 hrs
508871	12.998 hrs	6.486 hrs
504025	13.2 hrs	6.445 hrs

Table 3.1: My Periodogram Results and Established, Known periods recorded from the LCDB

Interestingly, the results referenced were often twice that of what I was recording with the Lomb-Scargle method. To investigate, plots of phase-folded light-curves based on the reported periods were made. In the example of 508871 it is evident that there are two peaks within the period. This difference is most likely due to a physical interpretation of period that I am not aware of. If this physical interpretation is particularly necessary, then it would help to consider a second-order periodogram as described in [VanderPlas, 2018]. An example for how this second-order periodogram changes the interpretation is shown in Figure 3.2.

Much like that in Figure 3.2, the single-order interpretation is finding half of the reported period. Without the use of a second-order periodogram,

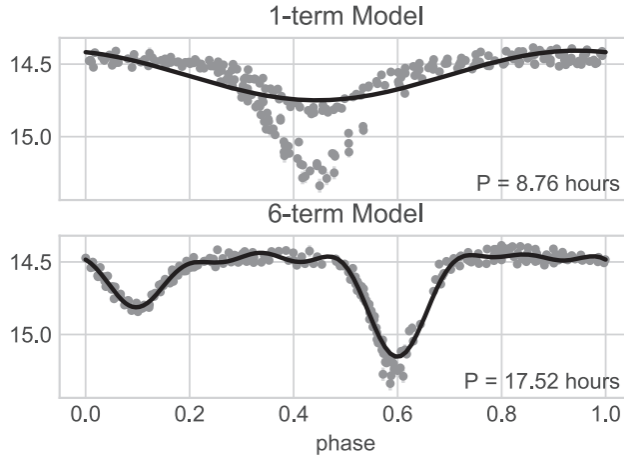


Figure 3.2: Demonstration of multiple-term model doubling resulting period for binary star. Note this is an object not analyzed in my table. From: [VanderPlas, 2018]

it is important to still consider how my results matched with that of half the reported period for those objects in which two peaks were within the designated period. Table 3.2 shows these new results for , which are considerably more promising and demonstrate the validity of the Lomb-Scargle periodogram.

Target	LCDB Period Half (When Two Peaks)	Found Period
503293	8.8 hrs	9.144 hrs
508871	6.5 hrs	6.486 hrs
504025	6.6 hrs	6.445 hrs

Table 3.2: Periodogram Results and Recorded Known Values from LCDB

Interestingly, though, the results are not exactly correct. It also seems that the results reported are better when looking at the phase-folded curves, and noticing them as having a tighter distribution of points. This demonstrates that the Lomb-Scargle periodogram may not be an extremely precise measure of periodicity, but can be a good estimator. In specific, none of my found periods were greater than 3% from the registered LCDB values after halving those with two peaks. It is also possible that the data available may

be the cause of the slight lack of precision.

3.3 Conclusion

The Lomb-Scargle periodogram is a methodology used to extract periodic tendencies from unevenly sampled data. In this thesis, I explored the foundations of periodogram analysis in Fourier Series and Fourier Transforms, and the extension to the Lomb-Scargle periodogram. When applying the Lomb-Scargle periodogram to real datasets, it was clear that the periodogram correctly interpreted periodic tendencies, with results being within 3% of the established periods (or half periods). It did become apparent, however, that to extract exact periods, other methodologies may be necessary, including multi-term Lomb-Scargle periodogram analysis. Further work can be done to explore alternate methods, and/or work towards determining whether the lack of precision in the Lomb-Scargle technique was due to the data I provided the R function rather than the methodology itself.

3.4 Acknowledgements

I want to thank Professor Hardin for all of your support throughout this senior thesis process. I would also like to thank Professor Choi for guiding me through the motivation for this project, and being a key contributor to my desire to be an Astronomy minor and focus my thesis on an Astronomy related topic. Also, thank you to Professor Chandler for your help answering questions regarding the Lomb-Scargle periodogram.

Bibliography

- [Beniyama et al., 2022] Beniyama, J., Sako, S., Ohsawa, R., Takita, S., Kobayashi, N., Okumura, S.-i., Urakawa, S., Yoshikawa, M., Usui, F., Yoshida, F., Doi, M., Niino, Y., Shigeyama, T., Tanaka, M., Tominaga, N., Aoki, T., Arima, N., Arimatsu, K., Kasuga, T., Kondo, S., Mori, Y., Takahashi, H., and Watanabe, J.-i. (2022). Video observations of tiny near-Earth objects with Tomo-e Gozen. *Publications of the Astronomical Society of Japan*, 74(4):877–903.
- [Bloomfield, 2004] Bloomfield, P. (2004). *Fourier analysis of time series: an introduction*. John Wiley & Sons.
- [Cheever, 2022] Cheever, E. (2022). The fourier series. <https://lpsa.swarthmore.edu/Fourier/Series/WhyFS.html>.
- [Debnath, 2012] Debnath, L. (2012). A short biography of joseph fourier and historical development of fourier series and fourier transforms. *International Journal of Mathematical Education in Science and Technology*, 43(5):589–612.
- [Fourier, 1808] Fourier, J.-B. J. (1808). Mémoire sur la propagation de la chaleur dans les corps solides. In *Nouveau Bulletin des Sciences, par la Société Philomathique de Paris*.
- [Fourier and Darboux, 1822] Fourier, J.-B. J. and Darboux, G. (1822). *Théorie analytique de la chaleur*, volume 504. Paris: Didot.
- [Herman, 2014] Herman, R. L. (2014). Introduction to partial differential equations. *UNC Wilmington, Wilmington, NC*.
- [Lomb, 1976] Lomb, N. R. (1976). Least-squares frequency analysis of unequally spaced data. *Astrophysics and space science*, 39(2):447–462.

- [NASA, 2018] NASA (2018). Twenty years of tracking near-earth objects.
- [NASA, 2022a] NASA (2022a). Asteroid light curve database (lcdb). <https://sbn.psi.edu/pds/resource/lc.html>.
- [NASA, 2022b] NASA (2022b). NASA Confirms DART Mission Impact Changed Asteroid’s Motion in Space.
- [NASA, 2022c] NASA (2022c). NEO discovery statistics.
- [Osgood, 2007] Osgood, B. (2007). The fourier transform and its applications. <https://see.stanford.edu/materials/lsoftae261/book-fall-07.pdf>.
- [Ruf, 1999] Ruf, T. (1999). The lomb-scargle periodogram in biological rhythm research: Analysis of incomplete and unequally spaced time-series. *Biological Rhythm Research*, 30:178–201.
- [Scargle, 1982] Scargle, J. D. (1982). Studies in astronomical time series analysis. ii-statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 263:835–853.
- [Stanford, 2001] Stanford (2001). Lecture 11 the fourier transform. <https://web.stanford.edu/class/ee102/lectures/fourtran>.
- [VanderPlas, 2018] VanderPlas, J. T. (2018). Understanding the lomb–scargle periodogram. *The Astrophysical Journal Supplement Series*, 236(1):16.
- [Vokrouhlický et al., 2015] Vokrouhlický, D., Farnocchia, D., D., Chesley, S. R., Pravec, P., Scheirich, P., and Müller, T. G. (2015). The yarkovsky effect for 99942 apophis.