Flickering Analysis of CH Cygni Using Kepler Data

Thomas Holden Dingus
East Tennessee State University

Follow this and additional works at: https://dc.etsu.edu/honors
Part of the Stars, Interstellar Medium and the Galaxy Commons

Recommended Citation
A Flickering Analysis of CH Cygni Using Kepler Data

THOMAS HOLDEN DINGUS

East Tennessee State University

dingust@goldmail.etsu.edu

April 2016

A Thesis submitted in Partial fulfillment for
The Midway Honors Program
in the
Department of Physics and Astronomy
College of Arts and Sciences

Faculty Readers

__________________________
Dept. of Physics: Richard Ignace

__________________________
Dept. of Physics: Gary Henson

__________________________
Dept. of Mathematics: Debra Knisley
Abstract

Utilizing data from the Kepler Mission, we analyze a flickering phenomenon in the symbiotic variable star CH Cygni. We perform a spline interpolation of an averaged lightcurve and subtract the spline to acquire residual data. This allows us to analyze the deviations that are not caused by the Red Giant’s semi-regular periodic variations. We then histogram the residuals and perform moment calculations for variance, skewness, and kurtosis for the purpose of determining the nature of the flickering. Our analysis has shown that we see a much smaller scale flickering than observed in the previous literature. Our flickering scale is on the scale of fractions of a percent of the luminosity. Also, from our analysis, we are very confident that the flickering is a product of the accretion disc of the White Dwarf.

1 Introduction

In this paper we analyze a symbiotic variable star named CH Cygni, in the constellation Cygnus. The data we use is acquired from the Kepler Mission Satellite. In the following sections we will summarize the background on symbiotic variables, CH Cygni itself, and the Kepler Mission. From there we will discuss the procedure of how we handle and analyze the data obtained, look into the code that we used, and discuss how we made sure we did not introduce any uncontrolled variables inside our procedure. At the end of this paper we discuss the results of our study.

1.1. Symbiotic Variables

Symbiotic variable stars get their name from the composite spectra that they exhibit. They were defined as a new subset of variable stars by Annie J. Cannon because of their peculiar spectra (Mikołajewska 2003). Usually symbiotic variables consist of a Red Giant and a hot White Dwarf or main-sequence star in a binary pair with a planetary nebula surrounding them (Kenyon and Webbink 1984). Red Giants are stars that have achieved the temperatures necessary for their cores to begin fusion of helium atoms. White Dwarfs are stars that have reached the end of their life; they are highly compressed stars. The White Dwarfs are stellar cores that have blown off their outer layers, leaving only electron degenerate matter behind. These stars are so dense that they have about the mass of the sun, but the size of the earth.

The two very different stars can create spectra that are quite distinct from typical spectra due to the fact they they are in a binary pair. These spectra
Figure 1: The above is an artist’s depiction of a Symbiotic Variable Star. Specifically RS Ophiuchi. The large red star is a Red Giant, and the bright white object is a White Dwarf. Artist: David Hardy

consist of features such as: TiO bands, neutral metal bands, and a red continuum (associated with Red Giants), HI Balmer lines, a blue continuum, and strong lines from highly ionized species like HeI, HeII, OIII, and other lines usually associated with planetary nebulae (Mikołajewska 2003). Below we are going to discuss the anatomy of a Symbiotic Variable and how the accretion disk forms.

In a symbiotic variable, the Red Giant and White Dwarf are interacting. As shown in Figure 1, the White Dwarf is capturing the wind from the Red Giant. The figure shows the White Dwarf gravitationally pulling matter from the Red Giant. The interactions between the White Dwarf and the Red Giant can cause an accretion disc to form around the White Dwarf (Sokoloski 2003). This happens due to the rotation of the binary. The rotating binary causes the matter to overshoot the star as it is pulled to the White Dwarf, forcing the matter to spin around the star, creating an accretion disk. White Dwarf’s have fairly large gravitational fields allowing them to pull material off of the other star. The mechanism for this is called Roche Lobe Overflow. This means that sometimes stars get big enough that their ability to hold on to their outermost layers is diminished, they grow past their Roche Limit. When this happens the matter can be pulled off of the star onto its binary partner. Sometimes this accretion gives rise to an area of ionized gas creating a symbiotic nebula (Shagatova, Skopal, and Cariková 2016). In some symbiotics an optical flickering in the accretion disk can happen (Sokoloski 2003).
What do we mean by this flickering? In our system there are long-period variations from the Red Giant, but we also see very short fluctuations in the brightness on the scale of minutes. Using the Kepler data we hope to learn more about these short period variations. While we have variations on the scale of weeks and days and even months, the variations we are looking for are on the scale of minutes. To observe the flickering we need a way to smooth out the long variations while letting us examine the short term variations. From flickering we can learn about the accretion physics going on behind this Symbiotic Variable.

1.2. CH Cygni

CH Cygni is a symbiotic variable star in the constellation of Cygnus. The location of the star is just under the western wing of the Northern Cross in the night sky. In Figure 2 the constellation of Cygnus is shown to be the Swan constellation. Looking towards the right wing tip there is a circle with a crosshair that shows where CH Cygni is in the constellation. It was only recently discovered to be a symbiotic variable when it started behaving differently. Before then it was thought to be a Red Semiregular variable. CH Cygni showed a spectrum of M6-M7 III, a Red Giant spectra. This changed with the addition of a hot blue continuum that had not yet been witnessed. When this happened CH Cygni, grew to its brightest magnitude yet recorded and started fluctuating quickly. It was identified as a symbiotic system by A.J. Deutsch in 1963 when he noticed the change
in its spectra (Faraggiana and Hack 1969). These two spectra point to CH Cygni being composed of multiple stars. There is some debate as to whether CH Cygni is a double or triple star system. In the double system the stars that compose CH Cygni are an M6-M7 III Giant and a White Dwarf (Slovak and Africano 1978), while the triple system adds a G-K dwarf in a currently unknown location (Hinkle et al. 1993). Both systems are equally viable at this point in time. There was no evidence for the hot component of CH Cygni prior to 1963 (Mikołajewski, Mikołajewska, and Khudiakova 1990). Several papers have found CH Cygni’s lightcurve was variable on very short time scales (Deutsch et al. 1974, Slovak and Africano 1978, Mikołajewski et al. 1990). This shows that the flickering is connected with the blue-violet continuum (Mikołajewski et al. 1990). It is still unclear exactly what is causing this flickering and a few different models have been presented such as the magnetic rotator model (Mikołajewski et al. 1990). There is also evidence for jets in the system (Crocker et al. 2001, Galloway and Sokoloski 2004, Belczyński et al. 2000). Crocker et. al discuss data obtained by the Very Large Array for CH Cygni, which shows Radio Jets from the star system. Galloway and Sokoloski discuss Chandra data for CH Cygni that shows an X-ray jet in the star system. Finally, Belczyński et. al discuss an optical jet along with the radio jet in their catalog.

A number of periods have been found thanks to the use of data obtained from the American Association of Variable Star Observers’ (AAVSO). AAVSO data is collected by variable star observers all over the globe dating back to the founding of the organization in 1911. This allows a high density of observations from the inception of the organization. Hinkel et al. (2008), assembled several papers on the parameters of CH Cygni. Most

1https://www.aavso.org/lcg
papers in this compilation find a long orbital period of 15.6 years, and some papers report another orbital period of 2.1 years. Most tend to agree on the pulsation period of the Red Giant being around 100 days, with another pulsation period at 770 days (Hinkle, Fekel, and Joyce 2009). The catalogue "A Catalogue of Symbiotic Stars" by Belczyński et. al assembled data for several symbiotic variable stars and Table 1 shows these parameters for CH Cygni (Belczyński et al. 2000). This paper based the ephemeris on a U-band minimum. The ephemeris is used to approximate when an eclipse occurs. In the ephemeris the variable E represents how many cycles have occurred since the reference date which is the one that the quantity multiplied by E is added to. To examine the flickering effect we need to figure out a way to remove these long period variations from the data. This is a process we will discuss later in this paper.

Table 1: Stellar Parameters of CH Cygni

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photometric Period</td>
<td>100 days</td>
</tr>
<tr>
<td>Radial Velocity Curve Period</td>
<td>756 days</td>
</tr>
<tr>
<td>Ephemeris</td>
<td>JD 2446275 + 5700 × E</td>
</tr>
<tr>
<td>Right Ascention</td>
<td>19h 24m 33.0s</td>
</tr>
<tr>
<td>Declination</td>
<td>50° 14′ 29.1″</td>
</tr>
</tbody>
</table>

Figure 3 shows the AAVSO Lightcurve. Lightcurves are one way astronomers study variable stars. They are plots of Brightness/Intensity on the y-axis using magnitudes with the x-axis measuring time. In Figure 3 the left end of the graph is when CH Cygni was in a dormant period. At this point, the star was fairly well behaved and thought to be only a semi-regular variable star. The variations are quite regular, coming up to about the same value then dimming to about the same value again. In the dormant period these variations also happen in about the same time periods. As we move to the right of the lightcurve CH Cygni has an outburst and it becomes obvious that it was a symbiotic variable; as time goes on the star enters an erratic period. We can see this by looking at how irregular and spaced out the variations are. There is a very large spike, reaching a brightness it had never reached before, followed by a dimming that had never been seen before.
Figure 4: Here the Kepler Field of View is pictured. The field encompasses most of the western wing tip of Cygnus, Making CH Cygni well in the view of the Telescope. Each square represents two CCD cameras in the Kepler Array. Photo taken from the NASA Documentation on Kepler.

1.3. Kepler Mission

The Kepler Mission began in 2009 to study exo-planets (Koch et al. 2010). Using the Kepler Space Telescope to observe a section of sky near the constellation Cygnus, the mission obtained huge amounts of data. Figure 4 shows the Kepler Field of View. The section of sky was determined because of the photometer. The photometer chosen had a 55° sun avoidance angle. Taking this angle into account this leaves two sections of sky; from there we look for the area with the most stars above V-magnitude 14, which leads us to the area near Cygnus (Koch et al. 2010). To look for exo-planets the CCDs on the satellite measure incident light from stars in the field of view and look for regular dips in the intensity (Koch et al. 2010). This shows that the light from the star is changing, if that change is very regular, and can be predicted as such, it is possible that the source of those dips is an exo-planet. Since this telescope observed a large portion of the sky, we can use it as a tool to learn more about stars in general, not just exo-planets (Koch et al. 2010). In terms of light variations, Variable Stars are not that different from stars with exo-planets. One such variable star that was studied by Kepler was CH Cygni. We can use this telescope to study this star.

Variable stars can have exo-planets. In such a case astronomers try to remove the regular longer period variations of the star so they can identify the planet transit. To do this they create a model for the variations and subtract from the model the long-period variations in order to get what are called residuals, the variations that are left over. The ways astronomers model lightcurves include: fourier transforms, spline interpolations, etc. These functions "flatten" the lightcurve as will be evident later in this paper. The
Figure 5: Here we can see the Kepler Telescope itself. The long cylinder is the main barrel of the telescope, where the light enters to be measured by the camera. The blue plate near the back is the solar panel that powers the entire telescope. Since it is solar powered, it needs to keep the solar panel pointing towards the sun. To do this the entire telescope rotates to keep the panel aimed at the sun once every quarter, causing ‘jumps’ in the data.

remaining residuals are phenomena like an exo-planet transit, or flickering variability. This is basically the same procedure we use in this study.

Kepler takes two kinds of data sets: long cadences and short cadences. The long cadences are thirty second exposures taken every thirty minutes over the course of three months, while the short cadences are thirty second exposures taken every two minutes over the course of a single month (Koch et al. 2010). A list of the data sets we use is in Appendix 1. Figure 5 is a diagram detailing the Kepler Telescope instrumentation and its appearance. The structure of the Kepler Space Telescope causes it to do some unintuitive things during its mission that affect the data sets. Every quarter (about three months) the satellite needs to roll to ensure its solar panels are pointed at the sun (Koch et al. 2010). This motion can create some offsets in the data. This is because when the telescope rolls, the star ends up on a different set of pixels on the CCD or even a completely different CCD. Only the long-cadence data suffers this rolling. For the purposes of this study, only the short-cadence data will be used. There are still some problems that affect the short-cadence data set. Every so often there are gaps in the data where the satellite stopped taking data for mechanical reasons. We remedy this by clipping the lightcurves around these gaps.

The error associated with the data is much less than the actual signal. The instrument
error associated with most of the Kepler data for CH Cygni with is on the level of $10^3 \text{ e}^{-s}$, while our signal is of the order of $10^7 \text{ e}^{-s}$. This means that the signal to noise ratio is of the order of $10^4$, this puts the percent error at 0.01%. The telescope is programmed to tell when the data points are not as good as they should be. The telescope’s programming assigns a quality flag to every data point. The best quality data gets assigned a flag of 0. Other flags signify things that make data non-quality, phenomena such as: cosmic ray hits, loss of fine pointing of the telescope, and others.

2 Methods

To get the Kepler data into a useful form we need to condition it. This means we need to run some analysis to remove data that we know is incorrect. The Kepler data as mentioned above has zeroes (where the telescope did not take data) and discontinuities in it. To fix these we used only the short-cadence data. Short-cadences have thirty second exposures taken every two minutes for a month. Using these data sets we do not have to deal with the quarterly roll. However, there are still gaps in the data where the instrument did not take any exposures. To do this procedure we utilize a computer code written in python to do most of the work and to make the process quick and simple. The procedure to prepare and use our data is as follows:

- Included an ‘if’ statement in code to get rid of gaps.
- ‘Binned’ averaging
- Cubic spline interpolation
- Subtract spline from actual data to get residuals
- Histogram residuals
- Moment statistics on residuals

In the next few subsections we will discuss exactly what each of these steps entails.

2.1 Binned Average

The Kepler data set is raw when we acquire it. We need to pre-process it to make it usable. We remove the zero values from the gaps in the data and any values that are flagged as non-quality. To do this we include an ‘If’ statement in our code. ‘If’ Statements allow you to include and use boolean operators. Since we are taking data points from something that is not zero, the intensity of a star, we know we can pull these values out. We include this line before we run the operations in our python code:

```python
for index in range(data):
    if data[index] > 0:
```
Figure 6: Here we can see the eight different data sets arranged by date. This is the lightcurve we use from the Kepler Data without stitching. We can see the discontinuities especially towards the right side of the graph inbetween the red and black sections. One can see that they do not quite meet up. The y-axis shows the intensity measured in electrons per second and the x-axis shows the Kepler Mission Date.

This segment of code creates a list of indices for the data and since each data point now has an index we can call the data point by that index. The code then runs through all those indices to see if the data values are greater than zero. Then it will perform any needed subsequent operations on only that portion of the data which satisfies what we wish to perform on the data set: binned averaging, cubic spline interpolation, subtraction from data, and histogram the data.

Figure 6 shows the lightcurve we obtained from the Kepler Data. You can see the erratic nature of the light curve from how much it changes over time. These changes need to be taken out in order to look at the small variations. To take out the long period variations we need to average the curve. If we do not average the lightcurve, our spline procedure will match our data exactly. If our model matches the data then we will not obtain residuals that will help us locate the flickering. The type of averaging we use here uses a binned average. This entails taking the first few points and evaluating the average. That average value is our new first data point. We repeat this process for the whole data set. The formula for this is shown in Equation (1), where \( \{D_i\} \) represents the original data points, and N is the number of points binned.
Flickering Analysis • Spring 2016

Figure 7: This figure shows an example of a binned average with data of the form \( y = x \). The blue circles represent the original data, while the red squares represent the binned data. Here the data points are binned on 10 points.

\[
\overline{D_j} = \frac{\sum_{i=1}^{i+N} D_i}{N}
\]  

We created a code to run through the thousands of data points in the files. The binning we used were ten, twenty, and thirty. So if we were binning our data on ten points and our first ten points were 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10, then our first data point for our binned average would be 5.5. This is shown for a set of numbers from 1 to 60 in the figure below. The unbinned data set is shown as the blue dots, while the binned data is shown as red triangles. The data points shown are binned by ten.

This procedure gives us an averaged lightcurve to generate our spline interpolation from. This step is crucial because if we did not average the lightcurve we would be modeling the whole lightcurve and not the long period variations we are trying to remove. Figure 8 shows the differences, or lack thereof, of the binned data set and the non-binned data. The two data sets are different, but they still look amazingly similar to each other.

As you can see when we view the whole data set, there is almost no way to differentiate between the two just from the graphs at this scale. Thanks to the high resolution of Kepler
data, the binned data still has a very large number of points despite using binning of $N = 10, 20, 30$. The raw data sets have approximately 50,000 data points, while for $N = 10$ the binned data has 5,000 data points. The raw data has a span of about thirty days, meaning each point reflects about a minute (51.8 seconds). When we bin ten points each point reflects about 10 minutes (518 seconds). While there is no visible different between the data sets on this scale, there are differences that appear as we zoom in on each data point. We need those differences to be able to model the long term variations in the lightcurve well.

2.2. Cubic Spline Interpolation

While there are many ways to model a curve, we have chosen to perform a cubic spline interpolation on our data set to model the lightcurve. We do this to remove the long-term variations of the star so that we can study the short-term variations that we call "flickering". We use the cubic spline almost like a smoothing function. To perform a cubic spline the following equation is satisfied.

$$S_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i$$  \hspace{1cm} (2)

Equation (2) describes a cubic function between two points $x_{i+1}$ and $x_i$. We set the
Figure 9: Above is a set of data connected with a cubic spline and with a linear interpolation function. The cubic spline connects the points smoothly. This is similar to what a zoomed in view of our data with a spline would look like. The dashed line shows the cubic spline, the dots represent the original data, and the green line shows the linear function of the data.

spline equal our data binned at the points \( \{x_j\} \), and to make sure the transitions are smooth, we require the following to be true,

\[
S'_i(x_i) = S'_{i-1}(x_i) ; \quad S''_i(x_i) = S''_{i-1}(x_i) ; \quad S'_0(x_0) = f'(x_0) ; \quad S''_{n-1}(x_n) = f''(x_n)
\]

These are conditions on the first and second derivatives of the Spline function. We force the spline to be continuous by making the first and second derivatives to be equal to the previous first and second derivatives. We also make the derivatives of the end points to be equal to the derivatives of the data at the end points.

For just one data set this would take a great length of time to do by hand. Luckily, there is a python module written to do the interpolation for us. We used the Scipy and Numpy modules to help us with the interpolation. Figure 9 shows an example of how a cubic spline works. This graphic is from the Scipy documentation on their cubic spline code.

We now employ this process on our data to generate our spline interpolation. We use our binned data to create our spline. If we used our raw data we would model the data in such a way that takes out the flickering as well as the long period variations. Using the binned data we can subtract the spline from the raw data to examine the left over variations. So we use our binned data to create our spline and use the spline to interpolate
Figure 10: This figure shows the process we go through with the data, from raw data, to binned data, to the spline, and the final plot is the spline shown over the raw data.
Figure 11: Here we zoom in on the fourth graph in Figure 9 so we can take a closer look at how the spline models the data. The spline goes through the majority of the raw data points while going through every binned data point. Subtracting the spline from the raw data leaves some residual data.

the rest of the data points to get the same amount we have in our raw data. In Figure 10 we show the entire process of working with our data, from raw data to spline. This shows that our spline does in fact model our data fairly well, but remember we do not want our spline to perfectly match our data or when we subtract it from the raw data we will be left with nothing to examine.

Now we zoom on the fourth plot in Figure 10 to come to Figure 11. Here you can see just how the data points are threaded by the spline. We use our binned data to create an average lightcurve, our spline interpolation. Our spline has values at every t-value between the first and last data point on the binned data set. Think of the light curve as a mountain or a hill covered in trees, our aim is to flatten this mountain so we can look at the heights of the trees covering it. To look at the flickering we flatten the hills and mountains of the overall lightcurve to look at what is left over, our residuals.
Figure 12: Here we see the residuals of the data subtracted from the spline. This allows us to look at the small quick variations in the data without seeing the larger slower variations. This is also another step to allow us to run statistics on the data.

2.3. Residuals

From the picture in Figure 11, it seems like it would be a difficult process to level out all those peaks and bumps in the data, but with our spline routine, this is as simple as a subtraction. To get the residuals of the data we need to subtract the spline from the data. Since the spline is a function, we input the same date values we have into the spline function to get the spline values for the data set. Now we take those values and subtract them from the corresponding intensity values. This leaves us with values that range typically between -10,000 and 10,000. Since the instrument error for Kepler is only about 800 we know these are well above the error. Our residuals show us the variations above the noise.

Figure 12 is a plot of our residuals. The two thin lines near the zero line show the average instrument error of the Kepler data. Most of our data points are well above that line. This tells us that there is some sort of source noise associated with our star. That is to say that this noise is caused by our source, CH Cygni, not the instrument. From our residuals we continue on to create a more useful plot of the residuals, the histograms.
2.4. Histograms

Having show the source signal in the residuals is well in excess of the measurement uncertainty we move on to plot histograms from those residuals. To do this we bin values, but this time we create 100 bins across the min and max of the residual value. Those bins are our x-values, with our y-values being the incidence of the residual value falling inside that bin. If we have 10 incidences of values between 0 and 1 then the y value for the bin of 0 to 1 is 10. If we have 50 incidences of values between 2 to 3 then the y value for the bin corresponding to values between 2 to 3 is 50. This process continues for however many bins are chosen. These plots help us to see the distribution of the source noise we are detecting. Figure 13 is a plot of one of our histograms. The histograms we plot look like gaussian distributions. From here we can perform the moment statistics on the data sets. The moment statistics we will be working with will be described in subsection 5.

2.5. Moment Statistics

From our histograms we acquire the moment statistics. Once we run these functions on our data we can learn a lot about how our histograms are shaped. Moments are defined by Equation (4), where $s$ is the order of the moment, and $n$ is the number of points in the data set.
Flickering Analysis • Spring 2016

\[ M_s = \frac{N}{N} \sum_{i=1}^{N} x_i^s = \frac{\sum_{i=1}^{N} x_i^s}{N} \quad (4) \]

The moments we are interested in are listed by order as follows:

- **Residual Average**: \( \bar{x} \), Mean of all of the residual values.
- **Standard Deviation**: \( \sigma \), This is a measure of how spread out the numbers are.
- **Skew**: \( \delta \), This is a measure of how asymmetric the distribution is.
- **Kurtosis**: \( \kappa \), This is a measure how how long the tails of the distribution are.

These are the moments we use to help understand a bit more about the flickering we are observing. These help us to study different aspects of the histograms.

To perform the average we use here simply add up the values and divide by the number of values. Equation (5) describes the equation for this.

\[ \bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i = \quad (s = 1) \quad (5) \]

The next moment we are interested in is the Standard Deviation, which is actually the root of the second order moment, its equation is listed in Equation (6). This value tells us two times the width at the half-maximum of the graph peak. This also tells us the standard deviation of the data set, and if the deviation is above the instrument error we know that the flickering shown there is real.

\[ \sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} x_i^2} \quad (s = 2) \quad (6) \]

Now we look at the third order moment, this is called Skewness. This value is a measure of how symmetric the distribution is about the mean. This value shows where the bulk of the data is located at. Figure 14 is taken from the paper "Measuring Skewness: a Forgotten Statistic?" by David Doane (Doane and Seward 2011). In this figure you can see how the two different types of skewness relate to a normal distribution in between them. The equation that defines skewness is shown in Equation (7).

\[ \delta = \frac{1}{N} \sum_{i=1}^{N} x_i^3 \quad (s = 3) \quad (7) \]

From here we look at the final moment we consider. This is the fourth order moment, and it is called the kurtosis. This is a measure of the ‘tailed-ness’ of the distribution relative to a Gaussian distribution. It measures if the tails are drawn out, or cut off. Equation (8) shows this moment. Figure 15 is taken from a paper by Lawrence DeCarlo titled "On the Meaning and Use of Kurtosis" and shows what exactly is meant by
Figure 14: Here we have a figure that displays the skewness of a distribution. One side looks more stretched out than the other. (Doane and Seward 2011)

Figure 15: Here we have a figure that displays the difference between a positive and negative kurtosis in relation to a normal distribution. In this figure the notation $\beta_2 - 3$ is used for the value of kurtosis. This makes the distribution on the left a positive kurtosis and the distribution on the right a negative kurtosis. (Decarlo 1997)

'tailed-ness' (Decarlo 1997). The positive kurtosis values have much longer tails on their distributions than normal distributions. The opposite is true for the negative kurtosis values.

$$\kappa = \frac{1}{N} \sum_{i=1}^{N} x_i^4 , \quad (s = 4) \quad (8)$$

We ran these statistics on each data segment binned on 10. From those we obtained the following values as given in Table 2. From Table 2 we can see that the standard deviation of each data set is at least 4-5 times larger than the 800-1000 we get from the instrumentation error, $\sigma_{err}$. From this analysis we see that $\frac{\sigma}{\sigma_{err}} \gtrsim 4$.

### 2.6. Simulation Analysis

After the statistical analysis we need to ensure that we account for any error that we introduce with our procedure. A stream of numbers was generated to simulate our data. Gaussian noise was added to simulate flickering. With these codes we can control free parameters such as: how many points are in a data set, and the measure of the noise
Table 2: Statistical Moments of Residual Data

<table>
<thead>
<tr>
<th>Points</th>
<th>Average $r$</th>
<th>Standard Deviation $\sigma_r$</th>
<th>Skew $\delta$</th>
<th>Kurtosis $\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>29634</td>
<td>12.0</td>
<td>3972.1</td>
<td>0.070</td>
<td>0.308</td>
</tr>
<tr>
<td>23609</td>
<td>4.5</td>
<td>4110.6</td>
<td>0.122</td>
<td>0.050</td>
</tr>
<tr>
<td>16930</td>
<td>106.2</td>
<td>11100.3</td>
<td>0.334</td>
<td>3.815</td>
</tr>
<tr>
<td>19830</td>
<td>29.2</td>
<td>5457.5</td>
<td>0.246</td>
<td>1.664</td>
</tr>
<tr>
<td>35679</td>
<td>13.1</td>
<td>4144.3</td>
<td>0.135</td>
<td>0.111</td>
</tr>
<tr>
<td>26760</td>
<td>18.3</td>
<td>4210.7</td>
<td>0.135</td>
<td>0.277</td>
</tr>
<tr>
<td>31332</td>
<td>21.3</td>
<td>5501.5</td>
<td>0.020</td>
<td>1.053</td>
</tr>
<tr>
<td>17622</td>
<td>16.3</td>
<td>4522.5</td>
<td>0.049</td>
<td>0.354</td>
</tr>
<tr>
<td>21519</td>
<td>39.7</td>
<td>4728.8</td>
<td>-0.021</td>
<td>0.951</td>
</tr>
<tr>
<td>19310</td>
<td>39.4</td>
<td>5659.5</td>
<td>0.194</td>
<td>2.013</td>
</tr>
</tbody>
</table>

added to it. We then run our procedure on it. Figure 16 shows the results for one of these data sets while Figure 17 shows the histogram for the same data set. This procedure allows us to determine what changes when we change the inputs for the data; this lets us examine how these changes reflect in the real data.

Now that we have our histograms we run the moment statistics again. We compute the statistics sets of data with make ups according to Table 3. This brings us to the results at Table 4 for the different data sets we created. We also compute the percent error between the input standard deviation and the inferred standard deviation. This allows us to find out how accurate the data our procedure yields is.

2.7. Results

From Table 4 we can see that our percent error ranges from 5% to 0.5%. We see a flickering in our data at around the level of 0.01% of the intensity of the lightcurve. This is a smaller level of flickering viewed than previous studies. In previous studies they see a flickering level between 7.6% to 2.29% (Mikołajewski, Mikołajewska, and Khudiakova 1990), or as low as approximately 1.1% (Slovak and Africano 1978).

The paper “Symbiotic Binaries III. Flickering Variability of CH Cygni: Magnetic rotator Model” by Mikolajewska reports that the average V magnitude for CH Cygni is 6.55 and they also reported the largest flickering in the U-filter being approximately 0.5 magnitudes, with the smallest flickering being 0.15 in the R filter. Since Kepler is in the visible spectrum we use the V magnitude average and can see that the flickering would be between 0.5 and 0.15 on time scales as small as 100 seconds making the percentage between 7.6% and 2.29% (Mikołajewski, Mikołajewska, and Khudiakova 1990). The paper “Flickering of the Symbiotic Variable CH Cygni During Outburst” by Slovak reports a V
magnitude of 6.85 for CH Cygni while it has a flickering in the V-filter of between 0.08 magnitudes and 0.10 magnitudes on a time scale of about 5 minutes (Slovak and Africano 1978).

This makes the percent of flickering here between 1.46% and 1.17%. The lowest of these percentages are approximately two orders of magnitudes larger than the flickering level that we see from our data. These percentages were taken as a percentage of magnitude. We feel that we may be looking at a couple of different levels of flickering. Our procedure may even be taking out larger scale flickering. But we are finding this flickering on similar timescales as the rapid flickering mentioned in other papers. There are several different possibilities for this phenomena that we will discuss in our conclusions.

3 CONCLUSIONS

From this procedure, applying the techniques of binning, spline interpolation, and moment statistics, we feel reasonably certain that we see a flickering on a smaller level than previously observed. The previous observations were on the scale of 1-3 percent. The flickering we have revealed is a factor of 10 smaller than that. We currently do not know the physics behind this flickering, but with further research we will be able to apply what we have learned in this study in order to to better understand the system. We are also reasonably certain that the flickering results from either the White Dwarf or the Accretion
Disc surrounding it. This is due to the observations of the CH Cygni system when the Red Giant eclipses the White Dwarf (Sokoloski and Kenyon 2003 & Hinkle, Fekel, and Joyce 2009). During these eclipses we see a cessation of flickering in the system. When there is no eclipse, the blue component dominates the spectra.

Are we introducing a false signal into our data? We discussed measures we took to make sure we have not. Earlier in the paper it was mentioned that we used three different binnings to make sure that we still see signal in each data set. In the Appendices we include the data for each binning set.

For further research we might repeat our procedure with the other symbiotic binary in the Kepler Field StHa 169. By running our procedure on this set of data we can get an even better understanding of the errors our data introduces and how this flickering phenomena presents itself.

As for further research for CH Cygni we can run additional analysis on the data we have analyzed so far. Power Series analysis and much more can be done to further analyze the flickering effect.

What is causing the flickering? With the accretion disk we have processes of turbulence and convection. This star is radiating energy; this process is a form of cooling known as radiative cooling, and it is known to have instabilities. These arguments are typically made in terms of Poisson noise. Poisson noise is described by equation 9. In the equation there are three variables $S, \sigma$, and $N$. Here $S$ is the signal, $\sigma$ is the noise, and finally $N$ is
the number of random events that are causing the noise. This would mean should we achieve a flickering of 1%, our signal to noise or $\frac{S}{\sigma}$ is 100. This would make our $N$ 100² or 10,000. Meaning that for that level of flickering we would expect to see 10,000 random events causing them. Our flickering that we have discovered is at a 0.01% making our signal to noise 10,000 which makes our $N = 100,000,000$.

$$\frac{S}{\sigma} \approx \sqrt{N}$$  \hspace{1cm} (9)

Suppose that our $N$ in this instance were bright patches on the accretion disk. Since the disk is already glowing, there would be the bright patches and less bright patches, but when taken as a whole on average on the disk they would negate each other. At any given instance there will be more of one than the other. That is our reasoning behind the flickering.

Following this vein, if we could find out the area of our accretion disk we could divide that area by our $N$ we would be able to get a characteristic area of the bright patches.
### Table 4: Statistical Moments of Residual Simulation Data

<table>
<thead>
<tr>
<th>Points</th>
<th>Average $r_{sim}$ ($10^{-3}$)</th>
<th>Standard Deviation $\sigma_{inferred}$ ($10^{-3}$)</th>
<th>Skew $\delta$</th>
<th>Kurtosis $\kappa$</th>
<th>% Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>9989</td>
<td>-0.36</td>
<td>9.6</td>
<td>0.012</td>
<td>-0.065</td>
<td>-4.3</td>
</tr>
<tr>
<td>9979</td>
<td>-0.35</td>
<td>9.8</td>
<td>0.007</td>
<td>-0.036</td>
<td>-2.0</td>
</tr>
<tr>
<td>9959</td>
<td>-0.36</td>
<td>9.9</td>
<td>0.011</td>
<td>-0.027</td>
<td>-1.1</td>
</tr>
<tr>
<td>9989</td>
<td>-0.12</td>
<td>2.9</td>
<td>-0.016</td>
<td>0.042</td>
<td>-3.7</td>
</tr>
<tr>
<td>9979</td>
<td>-0.12</td>
<td>3.0</td>
<td>-0.004</td>
<td>0.012</td>
<td>-1.3</td>
</tr>
<tr>
<td>9959</td>
<td>-0.12</td>
<td>3.0</td>
<td>-0.006</td>
<td>0.013</td>
<td>-1.3</td>
</tr>
<tr>
<td>9989</td>
<td>-1.12</td>
<td>29.0</td>
<td>0.038</td>
<td>-0.095</td>
<td>-3.3</td>
</tr>
<tr>
<td>9979</td>
<td>-1.07</td>
<td>29.6</td>
<td>0.035</td>
<td>-0.097</td>
<td>-1.2</td>
</tr>
<tr>
<td>9959</td>
<td>-1.15</td>
<td>29.9</td>
<td>0.031</td>
<td>-0.092</td>
<td>-0.47</td>
</tr>
<tr>
<td>989</td>
<td>-0.30</td>
<td>9.5</td>
<td>0.016</td>
<td>-0.065</td>
<td>-5.3</td>
</tr>
<tr>
<td>979</td>
<td>-0.17</td>
<td>9.7</td>
<td>0.016</td>
<td>-0.051</td>
<td>-2.9</td>
</tr>
<tr>
<td>959</td>
<td>-0.08</td>
<td>9.8</td>
<td>0.015</td>
<td>-0.030</td>
<td>-1.5</td>
</tr>
<tr>
<td>99988</td>
<td>-0.42</td>
<td>9.6</td>
<td>-0.003</td>
<td>-0.013</td>
<td>-4.3</td>
</tr>
<tr>
<td>99979</td>
<td>-0.42</td>
<td>9.8</td>
<td>-0.003</td>
<td>-0.005</td>
<td>-2.2</td>
</tr>
<tr>
<td>99957</td>
<td>-0.41</td>
<td>9.9</td>
<td>-0.003</td>
<td>-0.002</td>
<td>-1.5</td>
</tr>
</tbody>
</table>

Further research should try to determine the area of the disk. However, CH Cygni has an eclipse, meaning that the disk must be either edge of or nearly so. This would make the bright patches to be in the edge of the disk. This means we would only need the diameter of the accretion disk. Again, further research is needed to find these values.

With our research we have aimed to get a better grasp on the astrophysics behind the flickering phenomenon of CH Cygni. We have discovered that there may be multiple levels behind this phenomenon using the procedure we have detailed here. Further research will be needed to determine the cause of this phenomenon.
REFERENCES


Acknowledgements

In the production of this Thesis there have been many hands at work. Some right in the thick of it and others on the sidelines cheering on. Here I would like to give credit to those people. I would like to thank Dr. Gary Henson and Dr. Richard Ignace for their guidance and expertise in creating this project. I would also like to thank Dr. Joy Wachs for her moral support through my academic career at East Tennessee State University. Similarly, I would love to thank my family for their support all my life.

Now I would like to thank some of the organizations that made this possible. I would like to thank East Tennessee State University’s Midway Honors Scholar Program for accepting me and recognizing my ability as a student. Also, I want to thank the Columbus Phipps Foundation and the Buchanan County First Presbyterian Church for their financial support with my education. We acknowledge with thanks the variable star observations from the AAVSO International Database contributed by observers worldwide and used in this research. Some of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NNX09AF08G and by other grants and contracts. This paper includes data collected by the Kepler mission. Funding for the Kepler mission is provided by the NASA Science Mission directorate. This research has made use of NASA’s Astrophysics Data System Bibliographic Services.

Without these people and organizations my research and education would not be what it has become. I whole-heartedly thank you all.
4 Appendices

4.1. Appendix: Kepler Files

Here is a list of the data files we use for this study. These can be accessed from the Kepler archive in the Mikulski Archive for Space Telescopes (MAST). MAST is an archive that is used to store space telescope data. Here you can find data for Kepler, Hubble, GALEX, and so many more.

Table 5: List of Kepler Files used in this Study

<table>
<thead>
<tr>
<th>Kepler Mission Date Start</th>
<th>File</th>
</tr>
</thead>
<tbody>
<tr>
<td>443</td>
<td>kplr100002053-2010111051353_slc.fits</td>
</tr>
<tr>
<td>476</td>
<td>kplr100002055-201014023957_slc.fits</td>
</tr>
<tr>
<td>504</td>
<td>kplr100002057-2010174090439_slc.fits</td>
</tr>
<tr>
<td>539</td>
<td>kplr100002732-2010203174610_slc.fits</td>
</tr>
<tr>
<td>567</td>
<td>kplr100002736-2010234115140_slc.fits</td>
</tr>
<tr>
<td>808</td>
<td>kplr100002053-2011116030358_slc.fits</td>
</tr>
<tr>
<td>845</td>
<td>kplr100002053-2011145075126_slc.fits</td>
</tr>
<tr>
<td>874</td>
<td>kplr100002053-2011177032512_slc.fits</td>
</tr>
<tr>
<td>906</td>
<td>kplr100002732-2011208035123_slc.fits</td>
</tr>
<tr>
<td>937</td>
<td>kplr100002732-2011240104155_slc.fits</td>
</tr>
<tr>
<td>970</td>
<td>kplr100002732-2011271113734_slc.fits</td>
</tr>
</tbody>
</table>

4.2. Appendix: Graphs

In this Appendix we show the graphs for every data set we have created. The Data Sets will be shown in order by binning and then by date. The first three sets will be interpolations, followed by three sets of Residuals, followed then by three sets of Histograms.

---

2https://archive.stsci.edu/
Figure 18: This shows the Binned Data and the Spline Data for the whole data set with a binning of 10. The y-axis is an intensity is an electron count measured in electrons per second. While the x-axis is the Kepler Mission Date measured in days.
Figure 19: This shows the Binned Data and the Spline Data for the whole data set with a binning of 20. The y-axis is an intensity is an electron count measured in electrons per second. While the x-axis is the Kepler Mission Date measured in days.
Figure 20: This shows the Binned Data and the Spline Data for the whole data set with a binning of 30. The y-axis is an intensity is an electron count measured in electrons per second. While the x-axis is the Kepler Mission Date measured in days.
Figure 21: Here the residuals for Figure 18 are shown. The y-axis shows the Residual value, while the x-axis shows the Kepler Mission date measured in days.
Figure 22: Here the residuals for Figure 19 are shown. The y-axis shows the Residual value, while the x-axis shows the Kepler Mission date measured in days.
Figure 23: Here the residuals for Figure 20 are shown. The y-axis shows the Residual value, while the x-axis shows the Kepler Mission date measured in days.
Figure 24: Here the histograms for Figure 18 are shown. The y-axis shows the incidence count, while the x-axis shows the Kepler Mission date measured in days.
Figure 25: Here the histograms for Figure 19 are shown. The y-axis shows the incidence count, while the x-axis shows the Kepler Mission date measured in days.
Figure 26: Here the histograms for Figure 20 are shown. The y-axis shows the incidence count, while the x-axis shows the Kepler Mission date measured in days.