

Analysing Data from Five Stars

Jonas Telle
(Dated: Fall 2024)

We analyse daily data spanning about 30 years of $H\alpha$ -lines and relative fluxes from five stars for which the mass is known. We find their peculiar velocity and plot a centered velocity curve and light curve for each star, and fit the velocity curves with stochastic gradient descent. We find that stars 1, 2, 3 and 5 have planets, and that stars 1, 2 and 5 have inclination close to 90 degrees. Their planets have an estimated mass of respectively $3.4M_J$, $5.0M_J$ and $5.4M_J$, while the planet of star 3 has a mass of at least $0.71M_J \approx 225$ Earth masses. These values are surprisingly large, and it is unclear whether the planet classification is appropriate. Finally, we reconstruct the observed absorption line at $t = 0$ for star 5, which has a known temperature of 6000 K, and use our fitted velocity model to predict an observation to be made about 4.5 years after our final data-point.

I. INTRODUCTION

In analysing distant stars, we are interested in the star mass, temperature and its motion relative to our solar system. If the star also has orbiting planets, we would like to estimate their masses and radii. Knowing the common ranges on these metrics provide reference for future observations and help explain the origins of our own solar system. Furthermore, these ranges could cast light on deep laws of our universe. Indeed, certain physical insights, such as the existence of dark matter and our theories of gravity, are sooner gained from observing the distant than by constructing experiments on our tiny home planet. In the search of extraterrestrial life, one might expect Earth-like planets to be more likely to support life, making distant planet systems highly intriguing.

One way through which we obtain data about distant star systems is by measuring the wavelengths of electromagnetic radiation from the stars. In the current analysis, we have daily measurements of the radiation from five different stars over a time period of between 24 and 36.5 years, depending on the star. Capable scientists have extracted from these measurements the wavelength of the $H\alpha$ -line and the normalised flux of light from each star each day (Figure 1), and also estimated the mass of the stars (Table III). In the next section, we will explain what this means and how this data can be used to model each star's motion and determine which stars have planets.

In figure 1, we see that our data has noise due to effects such as atmospheric instability. However, we assume that systematic errors in the observations, such as the effect of the motion of the Earth, have been corrected. To get an accurate model fitting the noisy data, we implement a stochastic gradient descent algorithm. Finally, we use the fitted model to predict future observations for star 5, for which the temperature is known to be 6000 K, which is about the temperature of the Sun [9].

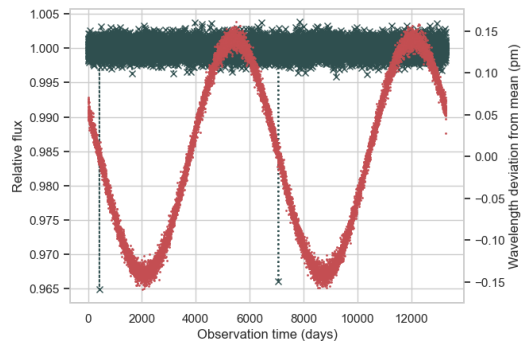


Figure 1. Our data for star 5. In red with axis on the right, the measured $H\alpha$ wavelength, plotted as the deviation from the mean, which is the shifted $H\alpha_5 \approx 657,314.93\text{pm}$ due to the peculiar velocity of star 5. Note that the original data is not centered in this way. In dark blue with axis on the left, the flux relative to the mean unobstructed flux of star 5, which is the format of the original data. We have about one measurement per day spanning almost 36 and a half years for this star.

II. THEORY

A star is to good approximation a black body, which is an object that absorbs all electromagnetic radiation it receives. Thus, the continuum flux $F_{cont}(\lambda)$, where λ is a given wavelength, follows Planck's law of radiation. In [5], this is given in terms of wavelength as

$$F_{cont}(\lambda; T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/kT\lambda} - 1}, \quad (1)$$

where h is Planck's constant, c is the speed of light, k is the Boltzmann constant and T is the star temperature. Note that this is a smooth function with a single extremum, as described by Wien's displacement law [5]. However, observed spectra will typically have certain lines missing, indicating that some particular wavelengths have been absorbed on their way to us. Since the energy levels of atoms are quantised, only specific wavelengths can be absorbed to excite an electron to a higher orbital. In particular, hydrogen atoms in the stars' atmosphere will absorb away the $H\alpha$ -line - which we will take to be at $\lambda_0 = 656.3\text{nm}$ - from our observations (figure 2).

Interestingly, the placement of the absorbed line is not constant in our observations, as is clear in figure 1. This lets us find the radial velocity of the star. Due to the Doppler effect, the light from a star moving towards us is blue-shifted, while the light from a receding star is red-shifted. Solving the Doppler-formula in [4] for the radial velocity v_r , we get

$$v_r = c \frac{\lambda - \lambda_0}{\lambda_0}, \quad (2)$$

where λ is the observed wavelength and λ_0 is the expected wavelength from a reference frame stationary with respect to the light source. In our case, λ_0 comes from the $H\alpha$ -line, since the star atmosphere containing hydrogen atoms is moving with the star, and λ is taken from the center line of an observed absorption line (the dotted black line in the Gaussian spectrum of figure 2). Thus, given the placement of the observed absorption lines and knowing that the line close to 656.3nm indeed comes from the hydrogen atoms in the star atmosphere, one can calculate the star velocity to good precision. Plotting this velocity against time, we obtain the *radial velocity curve* for that star (figure 5).

There is a nuance here, for while the star atmosphere is certainly stationary with respect to our star, the hydrogen atoms in that atmosphere are in motion. Thus, the radiation from the star will be Doppler-shifted with respect to individual atoms, and thus the absorption line will be spread out (figure 2). Still, since there are so many particles, we can use statistical physics to get a very accurate idea of how the absorption line will appear. Due to the high temperature, we can safely disregard the forces between these atoms and use the ideal gas approximation. Thus the particles' speeds - that is, the euclidean norm of their velocities - follow the Maxwell-Boltzmann distribution in [3], and since no velocity component is favoured, the direction of movement is uniformly distributed. Since λ is a linear function of v_r (see equation 2), we also expect the line profile to be Gaussian (figure 2). A proper derivation can be found in [5], from where we get the following function of λ , which approximates the observed spectrum:

$$F(\lambda) = F_{cont}(\lambda) + (F_{min} - F_{cont}(\lambda))e^{-\frac{\lambda - \lambda_0}{2\sigma^2}}. \quad (3)$$

Here, λ_0 is the mean absorbed wavelength, that is, the center-line of the absorption line - this is the measurement provided in our data - and F_{min} denotes the minimal flux, or the flux at $\lambda = \lambda_0$. The standard deviation $\sigma = \sigma_\lambda$ is obtained as follows: In [3], we find the standard deviation of the Maxwell-Boltzmann distribution, and we get the variance by squaring this:

$$V(v_r) = \frac{kT}{u},$$

where u is the mass of a hydrogen atom. Substituting v_r with equation 2 and using standard theory of variances

we get

$$V\left(\frac{c}{\lambda_0}\lambda - c\right) = \frac{c^2}{\lambda_0^2}V(\lambda) = \frac{kT}{u}$$

and reordering and taking the square root gives

$$\sigma_\lambda = \frac{\lambda_0}{c} \sqrt{\frac{kT}{u}}. \quad (4)$$

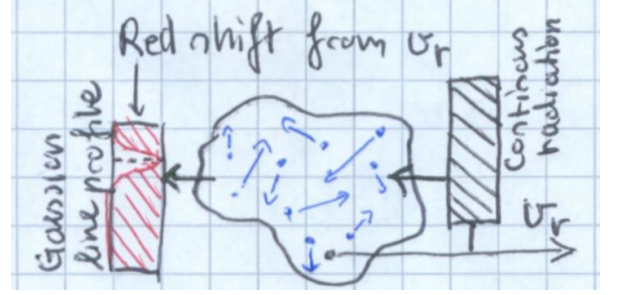


Figure 2. Illustration of the changes the radiation from a receding star undergoes before being observed by us. We have drawn a portion of the star atmosphere, with mean velocity v_r , receiving continuous radiation from the star moving with the same velocity away from us. Within this atmosphere, the hydrogen atoms move according to the Maxwell-Boltzmann distribution in the ideal gas approximation, and absorb only those wavelengths which from their point of view are Doppler-shifted precisely to match the energy required to excite an electron. This radiation is re-emitted in random directions, and we thus observe a dark line in the outgoing spectrum. Due to the normally distributed particle velocities, and since equation 2 gives a linear relationship between v_r and λ , the line profile is Gaussian. Note that we can still read off the mean of the absorption line at the minimal value for the flux, which is why our data only contains a single wavelength for each observation. Finally, the radiation is red-shifted since the radial velocity v_r is positive; the star is receding.

Using orbital mechanics, we can use the radial velocity curve to determine whether a star is orbiting in a non-perpendicular plane with respect to our line of sight, and in such a case find an upper bound on the planet mass. The angle between our line of sight and the normal direction of the orbital plane is called the inclination i . If $i = 0$, there is no radial motion even if the star is orbiting, and in this case we cannot determine whether there is a planet or not; in general, it is hard to prove that something does *not* exist. However, if the star's motion is periodic, it indicates an orbit, and thus some other object must be orbiting the star to provide a centripetal force. In [4], the following formula is derived from the Newtonian version of Kepler's third law:

$$m_p \sin i = M^{2/3} v_r^* \left(\frac{P}{2\pi G} \right)^{1/3}. \quad (5)$$

That is, from the star mass M , the radial star velocity v_r^* and the period P , using the gravitational constant G ,

we find the mass of the planet m_p times the sine of the inclination. This provides a lower bound on the mass, since $|\sin i| \leq 1$.

If we knew the inclination i , we would also know the planet mass. This is generally difficult, but in the case of $i \approx \pi/2$ we would expect the planet to regularly pass in between us and the star. Thus, the total flux from the star should diminish at regular intervals when the planet eclipses the star. Using the data of figure 1, we can plot a *light curve* for the star (figure 5) and check for such eclipses. With frequent measurements, one could determine the radius of a planet with such a curve, counting the time from partial eclipse until the planet is fully inside the star disc, where the flux stabilises until the planet begins to move out of the disc. However, this is unlikely to be possible with only one measurement per day since the planet is likely to fully enter the disc in one such time step (figure 6).

III. METHOD

We have data from several telescopes around the world in five files, one for each star, with three columns containing observation times in days, observed wavelength of the $H\alpha$ in nanometers and the flux relative to the maximum flux of that star (the dataset can be found at [10]). For each star, we read this data and make velocity and light curves as described in detail in the previous section. Importantly, we center our velocity curve by subtracting the mean value of the velocities. This mean value is a good estimate for the peculiar velocity of the star system - that is, the (assumed constant) radial velocity of the system center of mass - which can be found for each system in table III. Do note that, when estimating the peculiar velocity in this way, one should ensure that the mean is taken over a whole number of periods, else the result is skewed. In our case, as is clear from the velocity curves in figure 5, the data spanned about two periods for all the orbiting stars, so we did not need to take this into special consideration.

Since stars 1, 2, 3 and 5 exhibit periodic motion (figure 5), we have explained in Theory that these stars have planets. None of the velocity curves is a clear superposition of different sine waves, so we assume that one planet of each star is much more massive than the others and disregard other potential planets. By equation 5, we require the maximal radial velocity and the period of each star to obtain a lower bound on the mass on its planet. For an initial estimate, we can read off these values by eye from the velocity curve as illustrated in figure 3. Since we assume that systematic errors have been corrected, we know that the noise has mean value zero. Thus, we can imagine a line going through the middle of the noise in our velocity curves, like the light blue lines in the velocity curves of figure 5. The maximal radial velocity is the amplitude of the sine-wave, and thus - since our data is centered - the maximal value of this imagined line, while

the period is the time it takes for the star to make one full revolution, returning to its initial position in the system center of mass reference frame. For simplicity, we choose two wave-tops and estimate the distance along the x-axis between these, giving the period in years.

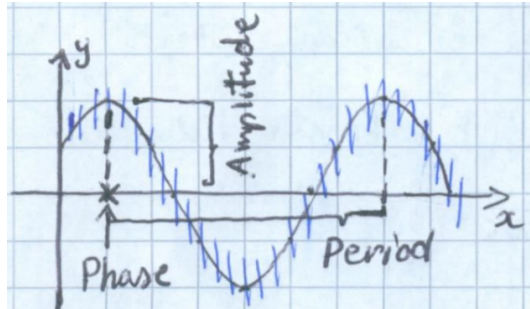


Figure 3. How to read off the amplitude, period and phase from a cosine-wave such as our model (equation 6). The blue lines denotes noisy data. The black line is a by-eye fit of this data, from which we read off the values as drawn. A periodic velocity curve such as this shows that the star is orbiting, indicating that it has a planet. Since it is a simple cosine, the orbit is due to a single planet; if there are others, they are much lighter or orbiting at inclination 0 degrees.

However, such by-eye measurements are not precise, and since we assume that the velocity curves come from a single planet, we should be able to make a reasonable fit with a simple cosine. Note that the data is centered such that we do not have any added constant. We want a model

$$v_r^{model}(t; v_r, P, t_0) = v_r^{max} \cos\left(\frac{2\pi}{P}(t - t_0)\right), \quad (6)$$

where v_r^{max} is the amplitude, P is the period and t_0 is the phase. The phase is only needed to ensure a good fit, as we do not need it for our calculations, however, it could be read off the velocity curve as the time when our function attains its first maximum, since cosine is maximised at 0 (figure 3).

We will use stochastic gradient descent (SGD) to fit our model. The method is described in detail in [6], but the main idea of gradient descent is to define a cost function of the parameters we want to optimize and iteratively calculate the gradient of this function in order to update our parameters in the negative gradient direction. This way, we hopefully find a minimum of the cost function (figure 4). The details, including the 'stochastic' part, will be explained shortly.

For simplicity when calculating the gradient, we let $A = v_r^{max}$, $B = 2\pi/P$ and $C = -Bt_0$ to get

$$v_r^{model}(t; A, B, C) = A \cos(Bt + C).$$

We use residual sum of squares as our cost function

$$RSS(\mathbf{v}_r^{pred}) = \|\mathbf{v}_r^{data} - \mathbf{v}_r^{pred}\|_2^2$$

where \mathbf{v}_r^{pred} denotes the predicted values by our model each day and $\|\cdot\|_2^2$ is the squared l_2 -norm. In other words,

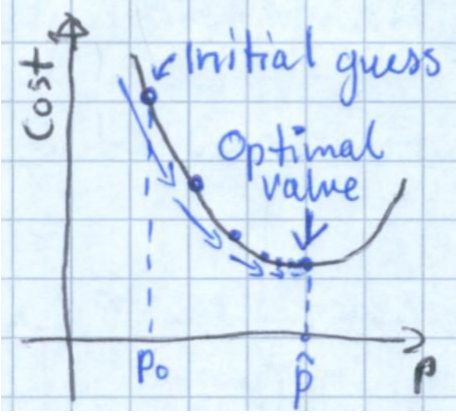


Figure 4. The main idea of gradient descent, here illustrated with a one dimensional case for a convex problem. We show the cost as a function of a parameter p ; in reality, we optimise three variables, but the multi-dimensional case is conceptually the same. Given an initial guess, we move in the opposite of the gradient direction; notice that the norm of the gradient affects our step size, so that we take very small steps once we are close to the optimal value \hat{p} . In SGD, we run through the whole dataset multiple times. For each such run, we do multiple training steps, using only a small portion of our data to calculate the gradient for each step, but we update our parameters in just the same way.

we are doing a least squares optimisation. Letting

$$e_t = v_r^{data}(t) - v_r^{model}(t) \quad (7)$$

for each time t for which we have data, the derivatives with respect to each parameter are

$$\frac{\partial RSS}{\partial A} = -2 \sum_t e_t \cos(Bt + C) \quad (8)$$

$$\frac{\partial RSS}{\partial B} = 2 \sum_t e_t A t \sin(Bt + C) \quad (9)$$

$$\frac{\partial RSS}{\partial C} = 2 \sum_t e_t A \sin(Bt + C) \quad (10)$$

summing over all t . Finally, we update each parameter $p \in \{A, B, C\}$ with $p_{new} = p - \delta \frac{\partial RSS}{\partial p}$ for each iteration, where δ is the learning rate, which is a hyperparameter which we choose before fitting (table I). We use the maximal centered radial velocity as our initial amplitude A_0 , and the other initial values are somewhat arbitrarily chosen to be $B_0 = 0.51/yr$ and $C_0 = 0yr$. Most of our stars indeed seem to have a phase of about 0, but this C_0 -value also worked well for star 5, which has a final $C_5 = 1.177yr$. These values seemed reasonable, and we got a good fit without tinkering with them. In hindsight, the B_0 -value was a little lucky, since all final B -values falls within the range $[0.345yr, 0.524yr]$. The parameters for the most challenging case of star 5 are displayed in figure 8, though the B -values must be converted to units of $1/yr$, as this is the unit used in the current discussion

and the training of the model. Our complete training journal with parameter values can be found in the supplementary material at [11].

Initially, we used standard gradient descent. While this worked well in testing for small datasets ($n \approx 100$), it seemed to break down when using larger datasets; especially, it did not work well on our dataset of $n \approx 10,000$. In particular, there seems to be a strong pull towards the global mean of 0. Thus, we require a very low learning rate to avoid the amplitude crashing to zero when the initial period and phase are off. With a lot of data-points, the sum of the residuals will be very large, and by extension $|\frac{\partial RSS}{\partial p}|$ (equations 7-10), so that even with very small learning rates, the descent gets way off track. Since we always initially overshoot with A_0 by definition, this value falls off fast and the model lands at the global mean of 0.

With inspiration from [6] and the fact that the model works well on small datasets, we turn to SGD. Here, we shuffle our data and batch it into subsets of m elements. Then we update our parameters on each batch. We repeat this for M epochs, where each epoch is one full run through our data. Both the batch size m and the number of epochs M are hyperparameters which we tune before training. Notice that this method avoids the issue with large datasets, since we make mini-batches of our data and train on these. Furthermore, it introduces randomness into our model, so that we can retry if we find our model initially gets off track. This could be useful, since we only tune our parameters for a single star, hoping that the algorithm is then robust enough to fit all other stars. This was indeed the case, and it seems to be relatively stable, so that the retry mechanism is almost never activated.

SGD hyperparameters	
M	1,000
m	64
δ_A	$5 \cdot 10^{-7}$
δ_B	$5 \cdot 10^{-8}$
δ_C	$5 \cdot 10^{-7}$

Table I. Final SGD hyperparameters. M is the number of epochs, m is the mini-batch size and δ_p is the learning rate for a parameter p . Notice that we, perhaps unconventionally, use a smaller learning rate for the second parameter B relating to the period. This took ages to tune, but once tuned on the first star it fitted the others very well without modification.

The final SGD hyperparameters can be found in table I. Note particularly that we use a smaller learning rate for the period-dependent variable B . The rationale is that we want the amplitude and phase to settle in before messing too much with the period, but the hyperparameters are mostly due to trial and error, as is often the case in machine learning. It is a custom to use powers of two for the mini-batch-size due to parallel processing [6], so 64 was a natural choice since we already noticed better performance on smaller datasets ($n \leq 100$). In figure 8,

we see that 1000 epochs seems sufficient for the parameters to settle. Note that the parameters are only tuned for the first star, but it still seems to fit the other four - which then act as a test-dataset - more or less perfectly.

After fitting the model for each star (see the light blue lines in figure 5), we can more precisely calculate the maximal radial velocity $v_r = A$ and the period $P = 2\pi/B$ from the model parameters A and B , giving a more accurate estimate of the bound on the planet mass with equation 5. The by-eye values and the model values can be compared with table II, where we also provide the signed relative difference between the lower bound mass estimates using the model estimate as the true value.

With such a simple velocity curve, one can do reasonably well by eye, but if we had multiple planets orbiting each star, giving a superposition of sine waves, it should be greatly advantageous to fit the data numerically, since it is non-trivial to pick out the different sine waves by eye. Still, the model parameters are better than by-eye measurements; their precision will be further discussed in the Discussion. We thus use the model estimates for the star period and maximal radial velocity in all final estimates presented in the Results section.

Star	Amplitude (m/s)		Period (yr)		Diff.
(i)	By-eye	Model	By-eye	Model	(%)
1	35	34.41	13	13.42	0.6
2	65	64.48	15	14.71	1.5
3	10	9.74	12	12.00	2.7
5	65	64.25	18	18.19	0.8

Table II. Comparison of by-eye values and the model including the signed relative difference in the calculated lower bound on the mass using the by-eye versus the model numbers.

We now plot the relative flux - that is, the third column of our data - against the observation time. In the previous section, we explained how we can use these curves to find which systems have approximately a 90 degree inclination; clearly, this is systems 1, 2 and 5 (figure 5). Thus, the lower mass bounds on these planets are good estimates for the actual planet mass. Unfortunately, we cannot estimate the planets' radii with only one measurement per day, as anticipated in the previous section. This is clear from the eclipses of figure 6, where we zoom in on a time-span of nine days about the minimal value of the light curve for each star with an eclipsing planet. Indeed, the planet fully enters the curve in between consecutive measurements.

Finally, we may use our fitted model to attempt to predict future observation data for a star with a known temperature. Star 5 is known to have a temperature of 6000 K [4]. We want to predict not only the placement of the $H\alpha$ -line, but also how the absorption line appears on the spectrum. We have explained how this can be done in the previous section, and will now simply use equation 3 to plot the flux as a function of wavelength.

Importantly, we make a simplification with constant

$F_{cont} = 1$ instead of using Planck's law of radiation; when zooming in on a small range of wavelengths just around the $H\alpha$ -line as shifted by the peculiar velocity, the continuum flux will not change too much in such a step, so that this approximation is perhaps tolerable. Furthermore, we know that the lowest relative flux at the absorption line is $F_{min} = 0.7$ [4].

We first test our approach against the first measurement on star 5. This star has a peculiar velocity of about $464 \frac{km}{s}$, and with rearranging equation 2 we get that the $H\alpha$ -line from this star will typically be about $657.3nm$. Thus, we plot a range from $656nm$ to $658nm$ to also include the stationary $H\alpha$ -line. We then make a prediction for a time one fourth of a period - corresponding to 4.55 years - after our last measurement. Both plots can be found in figure 7. Since the peculiar velocity dominates the centered radial velocity, these plots look very similar, but even such small differences can be measured accurately; in fact, the plot of the difference between these plots show that they are actually noticeably different, falling on different sides of the shifted $H\alpha$ -line (figure 7).

IV. RESULTS

We provide the system peculiar velocities in table III. While we cannot guarantee that any star does *not* have a planet, as it could be either very light or be orbiting at inclination 0 degrees, we know that stars 1, 2, 3 and 5 have planets. For stars 1, 2 and 5, the planet eclipses its mother star, showing that the inclination is close to 90 degrees. For the corresponding planets, we have a good estimate of the actual planet mass. For the planet of star 3, we have only a lower bound on the mass. These values are provided in table III.

Star	Mass	Peculiar velocity	Planet mass
(i)	(M_\odot)	(km/s)	(M_\oplus)
1	1.29	477	1,083
2	0.86	-388	1,597
3	0.86	230	≥ 225
4	5.35	-312	-
5	0.86	464	1,708

Table III. All but star 4 are orbiting in a plane visible from our line of sight. Among the rest, only star 3 deviates sufficiently from a 90 degree inclination to avoid being eclipsed; for its planet, we have only a lower bound for the mass.

Our peculiar velocities are on the order of $100 \frac{km}{s}$. In [1], Carlberg et al. state that the median velocity dispersion - which is a measure of the relative movement of galaxies within a group - for certain galaxy groups is $\sigma_1 \approx 200 \frac{km}{s}$, which demonstrates that such velocities are not unprecedented. Our planets' masses are however shockingly large, though Schlafman writes in [7] that objects up to ten times the mass of Jupiter (that is, more than 3,000 Earth masses!) have been found orbit-

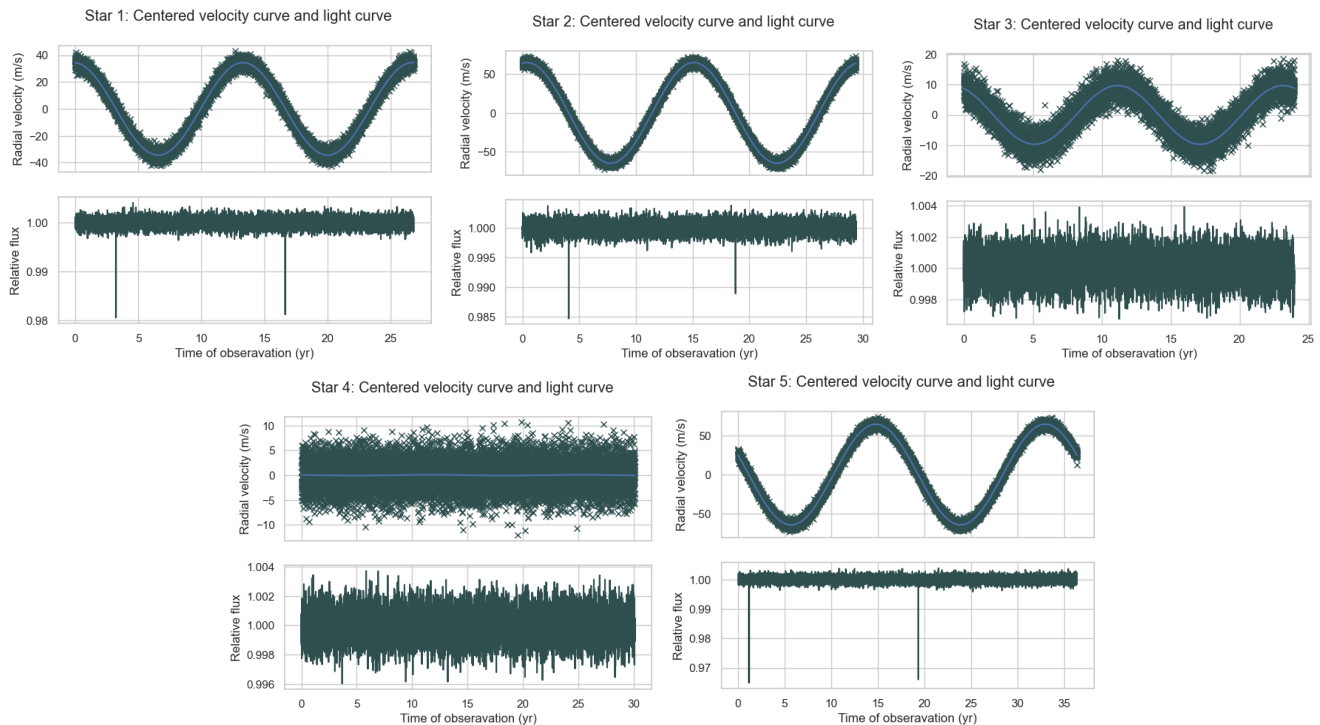


Figure 5. *Centered velocity curves and light curves for each star. The light blue line is the fitted velocity model. We see clearly that stars 1, 2, 3 and 5 are orbiting due to periodic motion. Furthermore, stars 1, 2 and 5 clearly have eclipsing planets, with clear spikes beyond the noise in their light curves. While it might appear that star 2 has two distinct eclipses, the spikes are due only to a single data-point (figure 6), and the difference is not larger than that it could well be due to noise. The simple sine-wave of the star 2 velocity curve indicates at least that if there are two planets, the second is very light. Assuming the apparent difference is only the effect of noise seems the most likely explanation without further observations.*

ing nearby stars. Still, this finding warrants particular attention, and will be so granted in the next section.

Finally, we plot our reconstruction of the first observed absorption line for star 5, the predicted absorption line at a time $P/4 \approx 4.5\text{yr}$ after the final measurement for this star, and the difference between these two spectra, in figure 7. For reference, we plot the stationary $H\alpha$ -line and the doppler-shifted $H\alpha$ -line as a result from the peculiar velocity. Since the star moves away from us, we expect a red-shift, and indeed the shifted $H\alpha$ -line has a longer wavelength than the stationary line. Since the peculiar velocity dominates the centered radial velocity, we expect that the reconstructed and predicted absorption lines should fall very close to the shifted line, as is the case. At last, we compare the plot of the difference between the spectra to the velocity curve. At $t = 0$, the radial velocity is slightly larger than the peculiar velocity. Extrapolating one quarter of a period beyond the final measurement, we strongly expect the velocity to be slightly smaller than the peculiar velocity. Thus, our reconstructed spectrum should be shifted to the left of the predicted spectrum, so that our difference plot should start with a peak and end with a trough, splitting the doppler-shifted $H\alpha$ -line just about down the middle. This is precisely what we see.

V. DISCUSSION

Our most surprising result is the planet masses. For reference, Jupiter has a mass of about $M_J = 318$ Earth masses [8], and our mass estimates for planets 1, 2 and 5 are respectively $3.4M_J$, $5.0M_J$ and $5.4M_J$, while we only have a lower bound on the mass of planet 3. As mentioned, [7] makes clear that celestial bodies of nearly twice our highest mass have been found orbiting nearby stars. In fact, Schlaufman explains that the *IAU Working Group on Extrasolar Planets* have used the “definition of a planet as an object with a true mass below the limiting mass for thermonuclear fusion of deuterium that orbits a star or stellar remnant” [7], thus using a bound of about $13M_J$ which our planets fall well within. He then goes on to propose a conservative reduction to $10M_J$.

However, he shows that nearly all unanimously classified giant planets have a mass below $4M_J$. Thus, if our mass estimates are correct, they fall into a grey-zone and might be difficult to classify. Either way, it is very surprising that all three planets for which we have a mass estimate fall in this rare range, suggesting that further critical analysis, or at least further observation of star systems 1, 2 and 5, is warranted. As such, our dataset is available at [10]. While unlikely, there could also be a selection bias in our data, in that stars with very large

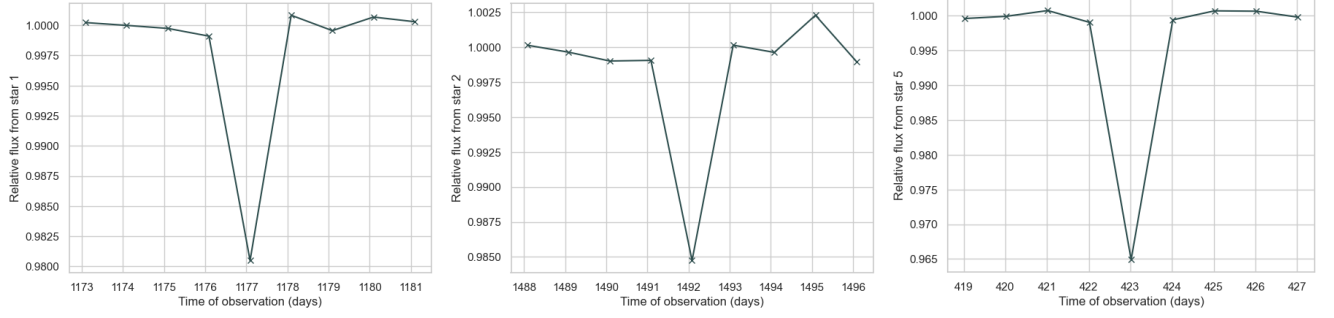


Figure 6. Close up of eclipses for stars 1, 2 and 5. We plot 9 days including the day of the eclipse. Notice that in all cases, the planet completely enters the star disc between consecutive observations, so that we cannot estimate the planet radius.

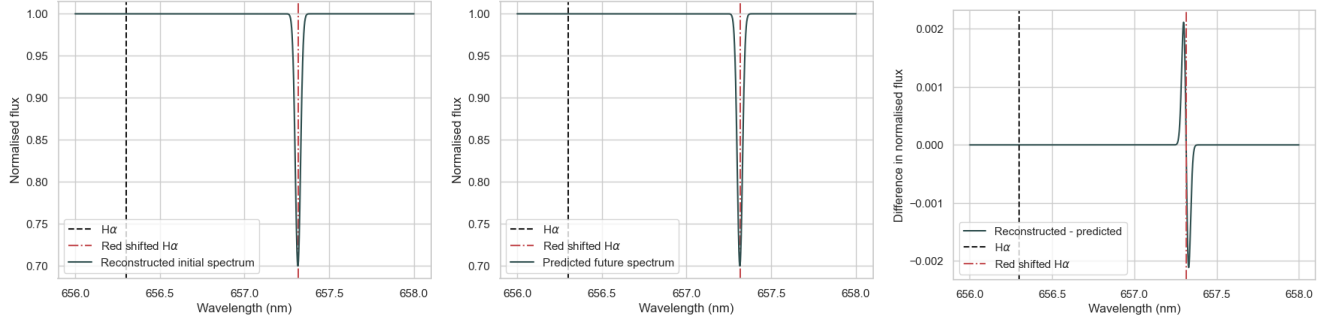


Figure 7. Simulated absorption spectra as could be observed from star 5, at $T = 6,000K$. From left to right: Reconstruction of the first observation, then a prediction for an observation to be made 4.55 years after our final data-point, and finally the difference between these spectra. The black dotted line denotes the stationary $H\alpha$ -line, while the red dotted line denotes the red-shifted $H\alpha$ -line due to the peculiar velocity of star 5. Note that we use a constant $F_{cont}(\lambda) = 1$ instead of Planck's law of radiation for simplicity. $F_{min} = 0.7$ is the minimal observed flux at the absorption line.

motions had a higher chance to be chosen as points of interest.

One could reconsider our assumption that there is only one planet, but our light curves do not give any indication of this; the difference in the spikes in the light curve of star 2 in figure 5 is not so large that it cannot be due to noise, keeping in mind that each dip only corresponds to a single data-point as seen in figure 6. Most importantly, the star orbits are clearly simple sine-waves, discrediting the notion that another planet not eclipsing the star - that is, outside the 90 degree inclination orbit - could be contributing anything beyond the size of the noise to the star motion.

We also considered the extremely unlikely situation that the star has two planets placed such that the other is in the L2-Lagrange point of the first. Then, they would orbit together, causing a single sine-wave yet distributing the large mass over two bodies. However, according to NASA in [2], L2 is unstable, so we take this to be impossible at our time scales without providing any further analysis. However, even if this was possible, it seems very doubt-able that mass could be even close to evenly distributed between such objects. Taking all these points together, the assumption of one dominant planet does seem justified.

The final SGD-algorithm seems stable and fits all stars without problems. Since the parameters seem to stabilise before our training concludes, it is likely that the obtained parameters are close to the optimal parameters (figure 8). Rudimentary testing suggests that the variances in the estimates for the maximal radial velocity and the period are no more than 0.01, and we take this together with the visual proof of good fit to be sufficient evidence that these values are correct. A more rigorous analysis of the variance using bootstrap was thus not performed, being computationally intensive.

However, if one were to fit other star velocity curves with much different, though still periodic, profiles, we suggest improving the initial guesses. This can be done by marking the observation times for about 100 of the largest measurements and grouping these by distance (this can easily be done by an unsupervised algorithm, since the distances within wave-tops will be much smaller than between tops). Taking then the average of each group gives an estimate for each peak. The first peak observation is the phase and the distance between consecutive peaks is the period. Using these estimates as initial values in our SGD should ensure more or less complete stability with all sine-waves, but requires that the data spans at least two wave-crests as in our data.

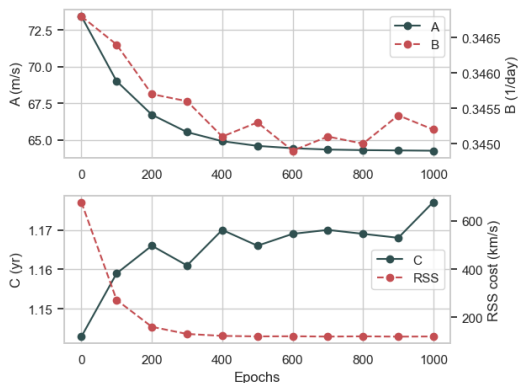


Figure 8. *Parameter evolution during SGD training for star 5. The dark blue lines have y-axes on the left, while the red dotted lines have axes on the right (see labels). Notice that the cost stabilizes well before we end training, as does the amplitude, which gives consistent results across re-runs. B and C are slightly more unstable, but only on the order of 0.01 in the units of table II. Note that the units on the axes in this plot have been adjusted for readability, else we would need more digits to distinguish the values.*

VI. CONCLUSION

We have analysed daily data of $H\alpha$ -lines and relative fluxes from five stars, for which the mass is known, spanning about 30 years. We find their peculiar velocity and plot a centered velocity curve and light curve for each star, fitting the velocity curves with a SGD algorithm. We have determined that stars 1, 2, 3 and 5 are orbiting, indicating that these stars have planets. For stars 1, 2 and 5, the light curves indicate that the inclination of this orbit is close to 90 degrees, and we estimate these planets' masses to be respectively $3.4M_J$, $5.0M_J$ and $5.4M_J$. The planet of star 3 has a mass of at least $0.71M_J \approx 225$ Earth masses. Clearly, none of these planets are similar to Earth. Finally, we reconstruct the absorption line as would have been observed at the initial measurement of star 5, which has a temperature of 6000 K, and make a prediction on how this line will look in about 4.5 years using our SGD-fitted velocity model for this star. Future research should be done to verify our surprisingly large planet mass estimates, and if found correct, decide whether to classify them as planets.

-
- [1] Carlberg, R. G., et al. (2001): *Galaxy Groups at Intermediate Redshift*. The Astrophysical Journal, 552(2), 427–444.
 - [2] Cornish, N. J. (2023): *What is a lagrangian point?* from NASA.
 - [3] Hansen, F. K. (2017): Lecture note 1A for the course AST2000 at UiO.
 - [4] Hansen, F. K. (2017): Lecture note 1C for the course AST2000 at UiO.
 - [5] Hansen, F. K. (2017): Lecture note 1D for the course AST2000 at UiO.
 - [6] Hjorth-Jensen, M. (2021): *Applied Data Analysis and Machine Learning: Ch. 7.7. Stochastic Gradient Descent*.
 - [7] Schlafman, K. C. (2018): *Evidence of an upper bound on the masses of planets and its implications for giant planet formation*. The Astrophysical Journal, 853(1), 37-50.
 - [8] Williams, D. R. (2024): *Jupiter Fact Sheet* from NASA.
 - [9] Williams, D. R. (2024): *Sun Fact Sheet* from NASA.
 - [10] Dataset (Seed 42).
 - [11] The SGD training journal can be found in 'example-output1C4.txt'.