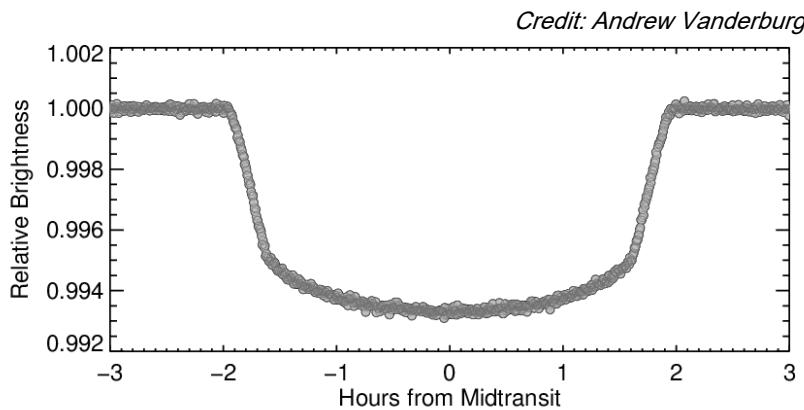
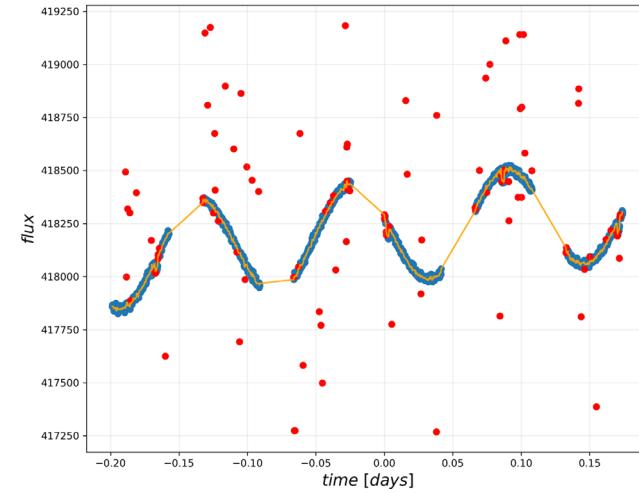


# LAB 3: Transits Jorge Casas, Pablo Drake

In this lab, we were tasked with cleaning and detrending simulated Hubble data of exoplanet transits, of which was mostly handled by an MCMC algorithm to detrend the global oscillations in flux data.



The typical transit light curve is shown, above, with a distinct trapezoidal transit depth blocking out a minuscule magnitude of flux from the light source. The given data in this lab contains four transits of similar shape on smaller transit durations.

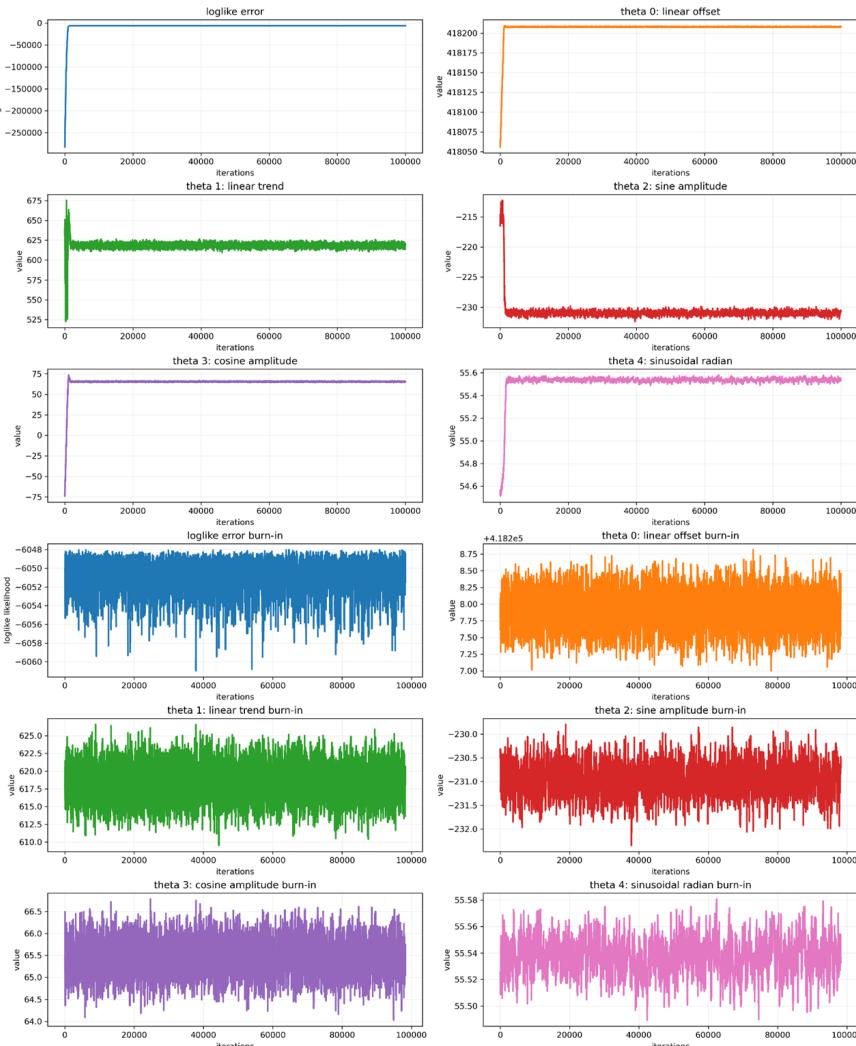
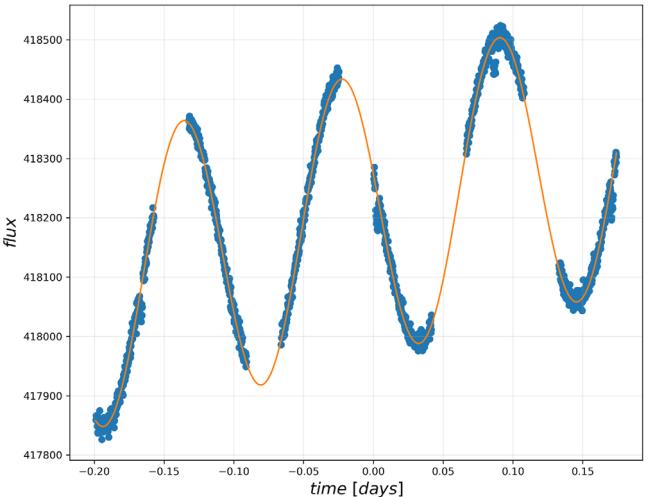


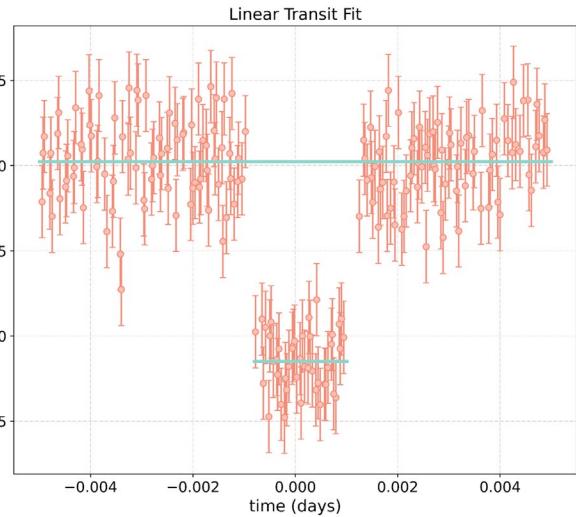
The plot above is a sample of an algorithm using a 5 point moving median to flag and discard flux data larger than  $3\sigma$ , prior to detrending to maximize the log-likelihood error in the MCMC algorithm to detrend the data.

The following model was used in the MCMC algorithm to detrend the data, using a linear term and an offset to detrend the drift in flux and two sinusoidal terms to detrend the oscillation in flux from heating and cooling of sensors.

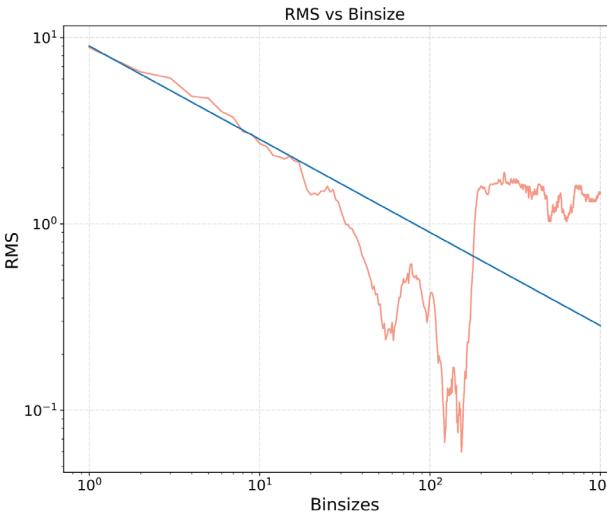
$$\text{flux} = a + bt + c\sin\left(\frac{2\pi t}{P}\right) + d\cos\left(\frac{2\pi t}{P}\right)$$

A periodogram was first used to provide initial parameters for the MCMC algorithm, and the resulting model was plotted below in orange to fit it against the raw flux data.

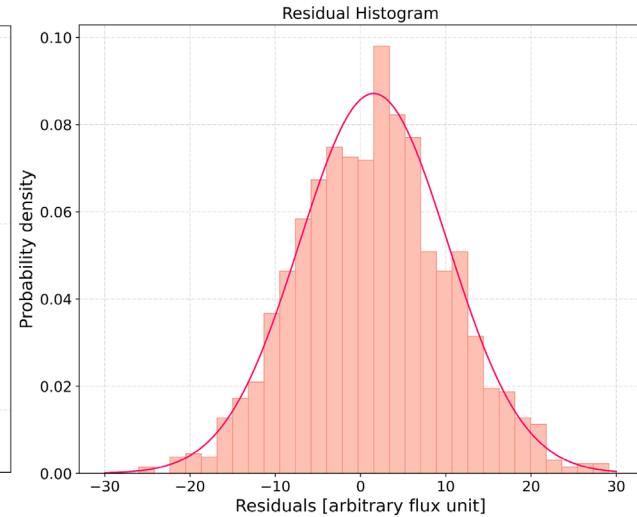




*Folded Transits and Depth*



*RMS vs Binsize scatter plot, along with the  $\text{RMS} \sim \text{binsize}^{1/(0.5)}$  line*



*Histogram of residuals with a normal distribution of (mean, sd) = (1.54, 8.70) superimposed*

The transit depth of  $0.000117 \pm 0.000004$  with respect to the normalized flux with an uncertainty derived from a bootstrapping method using a normal distribution of errors, which was calculated by minimizing the error in a linear regression through each transit and non-transit regions.

For quality of fitment of the model, a  $\chi^2$  value of 1328.17 with 1504 data points was calculated using the model from the global MCMC algorithm with transits included, and a residual data set without transits included to maximize the loglike error in the MCMC without directly changing the global data set before the MCMC.

We furthermore binned the residuals with a range of bin sizes between 1 and 1000, and plotted their corresponding RMS. Lastly, we also plotted the distribution of residuals in a histogram centered about a mean of 1.54 and a standard deviation of 8.70.