Detection and Removal of Periodic Noise of Kepler/K2 Photometry with Principal Component Analysis
Riley Clarke, Federica Bianco, and John Gizis
Department of Physics and Astronomy, University of Delaware

Abstract
We present a novel method for detrending systematic noise from time series data using Principal Component Analysis (PCA) in Fast Fourier Transforms (FFT). This method is demonstrated on time series data obtained from Campaign 4 of the Kepler K2 mission, as well as two additional objects of interest. Unlike previous detrending techniques that utilize PCA, this method performs the detrending in Fourier space rather than temporal space. The advantage of performing the analysis in frequency space is that the technique is sensitive purely to the periodicity of the unwanted signal and not to its morphological characteristics. This method could improve measurements of low signal-to-noise photometric features by reducing systematics. We also discuss challenges and limitations associated with this technique.

Kepler & Periodic Signals
The Kepler Space Telescope (pictured left) was designed to detect periodic astrophysical signals such as:
• Exoplanet transits
• Photographic variability
• Eclipsing binaries
but the Kepler catalogue includes periodic systematic noise due to scheduled rolling motion every 6 hours. Spectral analysis tools like Fast Fourier Transforms (FFT) can help analyze not only astrophysical periodicity, but systematics as well.

Left: Sample Kepler lightcurve of an M7.5 brown dwarf and the conjugate square of its Discrete Fourier Transform (DFT), i.e. the power spectrum (PS). The high-amplitude spike is an astrophysical signal caused by photographic variability. The smaller spike, indicated by the red arrows, is the systematic Kepler roll frequency.

Kepler & Periodic Signals

Principal Component Analysis
Principal Component Analysis (PCA), is a dimensionality-reduction technique that transforms the data into a new orthonormal coordinate space, the basis vectors of which are the principal components (PCs). The associated eigenvalues of each PC indicate what fraction of a data point, in our case a vector of flux points, is projected along that axis. The eigenvalues represent how much of the variance present in the data is explained by each PC. The components can be sorted by the amount of data variance explained (i.e. the first component is the one that alone explains the largest fraction of the variance in the data).

To detect and characterize systematic periodic noise in a set of synchronous, evenly sampled lightcurves, the following steps are performed:
• The power spectrum of each lightcurve is obtained by taking the conjugate square of the Fourier Transform (Oliphant 2006).
• A PCA decomposition was generated using the power spectra as inputs, a PCA decomposition was generated using the power spectra as inputs, and high-order PCs were inspected.
• The first 5 PCs of the power spectra ensemble are shown below in ascending order from top to bottom. The coefficients of the first PC are uniformly positive and non-zero almost all lightcurves, which means that PC-1, with its prominent peak at 4.09 days, is required to reconstruct every lightcurve in the dataset, connecting PC-1 to a systematic photometric variance.

For left: A synthetic illustrative dataset; note with the principal components of the data are denoted on top of the dataset. Here, the signal is obtained by PCA. The first component captures the majority of the variance of the data. The grey shaded regions indicate the standard deviation along PC-1 and PC-2 (imagination).

Principal Component Analysis

Principal Component Filtering
The systematic periodic signatures in our sample of Kepler lightcurves were removed according to Algorithm 1 and some sample results are shown below.

Left: Results of HYPERs for the first four PCs of the present signal. The left panel shows the results separated on the x-axis, and a crossover in region indicates the results have also been shown on the right.

Right: Upper panel shows the peak-to-peak amplitude of the periodic signal and the left panel shows the maximum amplitude of the residual signal. The bottom panel shows the cumulative variance of the dataset explained by each PC. Upper Right: Excerpt of a lightcurve showing periodic signals. The systematic signal is highlighted in red, and the periodicity is captured by the PC-1.

References
Cosley, J. W. 1987, Mathematica Acta, 93, 33