CoRoT Data and Transit Detection*

Peter Plavchan
On behalf of NStED and NExScI team

* Proper credits to the proper people for figures and so forth...
Outline

* Intro
* Review systematic noise sources
* Removing systematic noise
* Period finding techniques
* CoRoT data
Transit Basics

Primary Eclipse

Measure size of transiting planet, see radiation from star transmitted through the planet's atmosphere

Secondary Eclipse

See thermal radiation from planet disappear and reappear

Learn about:
- Global structure
- Atmospheres
- Eccentricity & Obliquity
Real Data – CoRoT4
CoRoT “raw” data (1st long run)

5/19/09 Peter Plavchan, AY 218

← 35 days →
Colours of noise

white

red

pink
Colours of noise

white

red

pink

5/14/09 Peter Plavchan, Greater IPAC Technology Symposium
Sources of Systematic Noise

- Airmass
- Seeing
- Crowding
- Intra-pixel effects
- Hot Pixels
- Other Detector Effects
- Astrophysical False Positives
Fig. 1.— Secondary eclipse of HD 209458b on UT 2005 Nov. 27, observed at (from top to bottom) 3.6, 4.5, 5.8, and 8.0 μm, binned in 7 minute intervals and normalized to one. The overplotted curves show the best-fit corrections for detector effects (see §2.1 and §2.2).
Fig. 2.— Secondary eclipse of HD 209458b on UT 2005 Nov. 27, observed at (from top to bottom) 3.6, 4.5, 5.8, and 8.0 μm, with best-fit eclipse curves overplotted. Data has been normalized to remove detector effects (see discussion in §2.1 and §2.2), and binned in 7 minute intervals.
Fig. 2.— *Upper Panel:* Baseline-removed aperture photometry of the HD 189733 secondary eclipse. Points are individual 6-second measurements, with error bars suppressed for clarity, but showing the eclipse curve having the best-fit amplitude ($0.551 \pm 0.03\%$) and central phase. *Lower Panel:* Data from the upper panel averaged in bins of width 0.001 in phase ($\sim 3$ minutes), with error bars and the best-fit eclipse curve.
Fig. 1.— Upper Panel: Raw aperture photometry, before background subtraction and baseline correction, for HD189733 versus planetary phase. Note that the secondary eclipse is already visible near phase 0.5 - marked by the vertical line. Lower Panel: Aperture photometry of the comparison star (2MASS20004297+2242342, points), with a polynomial fit (solid line through points, see text). The line below the comparison star shows the background level, which has been increased by an arbitrary factor to place it on the same scale. The background in the HD189733 aperture is about 30% of the total signal, and for the comparison star about 60%.
Astrophysical False Positives
(Other phenomena that mimic the transit signatures)

- Eclipsing binary with grazing orientation
- Small star crossing in front of a large star
- Eclipsing binary diluted by the light of a third star ("blend") → trickiest case
MACHO false positives

5/14/09 Peter Plavchan, Greater IPAC Technology Symposium
How to “De-trend” Light Curves

* **SYS-REM**
  * SYStematic REMoval
  * Reduces to “Principle Component Analysis” for identical photometric uncertainties

* **TFA**
  * Trend Filtering Algorithm
SYS-REM: Minimize $S^2$

$i =$ star #

$J =$ image # / date

$c =$ color dependent “extinction” correction coefficient

$a =$ “air mass”

$r =$ magnitude or relative flux

$\sigma =$ photometry uncertainty

→ Iterative Linear trend fitting and removal

$$S^2 = \sum_{ij} \frac{(r_{ij} - c_i a_j)^2}{\sigma_{ij}^2}.$$
Assume $a_j$ are known, solve for $c_i$

Star no. $i$:

$$S_{i}^2 = \sum_j \frac{(r_{ij} - c_i a_j)^2}{\sigma_{ij}^2}$$

$$c_i = \frac{\sum_j r_{ij} a_j / \sigma_{ij}^2}{\sum_j a_j^2 / \sigma_{ij}^2}$$

Now $c_i$ are known, solve for $a_j$

Image no. $j$:

$$S_{j}^2 = \sum_i \frac{(r_{ij} - c_i a_j)^2}{\sigma_{ij}^2}$$

$$a_j = \frac{\sum_i r_{ij} c_i / \sigma_{ij}^2}{\sum_i c_i^2 / \sigma_{ij}^2}$$
\[ D = \sum_{i=1}^{N} \left( Y(i) - A(i) - F(i) \right)^2. \]  

Here \( \{Y(i); i = 1, 2, \ldots, N\} \) stands for the target time series being filtered, and \( \{A(i); i = 1, 2, \ldots, N\} \) denotes the current best estimate of the detrended light curve, defined in the following way.
discussed later). The filter \( \{ F(i); i = 1, 2, ..., N \} \) is built up from the following linear combination of the template time series \( \{ X_j(i); i = 1, 2, ..., N; j = 1, 2, ..., M \} \)

\[
F(i) = \sum_{j=1}^{M} c_j X_j(i) . \tag{1}
\]
Example: SuperWASP 16h30+28 field
300 stars, 2549 observations spanning 100 days
Detrending algorithms

- **TFA** takes trends from linear combination of randomly selected sub-sample of light curves stars in field to serve as trend “templates”. Can benefit with the period of the science target variability is already known.

- **SYS-REM** fits linear trends with no a priori knowledge of trends or periods.

- Both algorithms are iterative.

- Both algorithms require convergence criteria, or times to stop, and this is somewhat of an art form.

- For SYS-REM, stopping criteria is determined by comparing the ratio of the global dispersion (standard deviation) of photometry before and after the detrending; with a limiting threshold.
Figure 12. Examples of the reconstruction of the light curves of observed intrinsic variables with the aid of iterative TFA filtering. Notation, method of reconstruction and the TFA parameters used are the same as in Fig. 10.
Period Finding

- Brute Force
  - Search through 10,000's periods
  - For each period, “fold” the light curve to that period
    - phase = (date modulo period) / period
  - Calculate some quantity based upon a specific algorithm to evaluate the significance of the “test” period
  - Generate a “periodogram”, and “peaks” in the periodogram may correspond to the correct period
Periodogram Algorithms

- Lomb-Scargle
- Box Least Squares
- Strlen
  - $\text{Min}( \Sigma m_{i+1} - m_i ), m_i$ are ordered by phase after folding
- Analysis of Variance
- Phase Dispersion Minimization
  - Uses phase “bins”
- Plavchan
Why not generate FFT?

- Fast Fourier Transform assumes the observations are evenly spaced in time, with no gaps.
- Real time series observations rarely meet this criteria
- Daylight gets in the way of ground-based efforts

- Lomb-Scargle is effectively a FFT for unevenly sampled data
- For a trial period, fit data to sine wave. Amplitude of sine wave yields significance of the trial period.
Box Least Squares

- Instead of sinusoids, take data folded to trial period and fit to "box-like" transit curve.
Maximize:

\[
\chi^2_{n_0} = \frac{\sum_{i=1}^{n_0} (J_i - J_{\text{mean}})^2}{\sum_{t=1}^{n_0} (J_t - J_{\text{prior}_t})^2}
\]
RA: 279.824077  Dec: 48.900043  Period: 0.1256800 days
Plavchan Peridogram
Summary (Pt 1)

* Published light curves of transiting planets hide the massaging and removal of systematic sources of noise, but fortunately these tools exist.

* Finding a transit signal in a light curve is a brute force extension of a Fourier Transform, with a careful choice/substitution of “basis functions”
Center & Anti-Center
### Results from first 5 runs

<table>
<thead>
<tr>
<th>Run</th>
<th>Stars</th>
<th>Transit signals</th>
<th>Discussed</th>
<th>candidates</th>
<th>F. Up</th>
<th>Planets</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRa01</td>
<td>9921</td>
<td>230</td>
<td>~ 90</td>
<td>51</td>
<td>51</td>
<td>2</td>
</tr>
<tr>
<td>SRc01</td>
<td>7015</td>
<td>256</td>
<td>~ 90</td>
<td>57</td>
<td>57</td>
<td>0</td>
</tr>
<tr>
<td>LRc01</td>
<td>11448</td>
<td>228</td>
<td>~ 50</td>
<td>50</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>LRa01</td>
<td>11448</td>
<td>299</td>
<td>~ 90</td>
<td>45</td>
<td>45</td>
<td>2</td>
</tr>
<tr>
<td>SRa01</td>
<td>5390</td>
<td>155</td>
<td>34</td>
<td>14</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>45,222</strong></td>
<td><strong>1168</strong></td>
<td>~ 354</td>
<td><strong>217</strong></td>
<td><strong>208</strong></td>
<td><strong>6</strong></td>
</tr>
</tbody>
</table>
The "first 4"!

**CoRoT-Exo-1b**
- \( P: 1.5089557 \) d
- \( r: 1.49 \) \( R_J \)
- \( m: 1.03 \) \( M_J \)
- **The star:**
  - G0V
  - \( V = 13.6 \) mag

**CoRoT-Exo-2b**
- \( P: 1.742996 \) d
- \( r: 1.465 \) \( R_J \)
- \( m: 3.31 \) \( M_J \)
- **The star:**
  - K0V
  - \( V = 12.6 \) mag

**CoRoT-Exo-3b**
- \( P: 4.2568 \) d
- \( r: 1.01 \) \( R_J \)
- \( m: 21.66 \) \( M_J \)
- **The star:**
  - G0V
  - \( V = 13.3 \) mag

**CoRoT-Exo-4b**
- \( P: 9.20205 \) d
- \( r: 1.19 \) \( R_J \)
- \( m: 0.72 \) \( M_J \)
- **The star:**
  - F0V
  - \( V = 13.7 \) mag
CoRoTCCDs

5/19/09  Peter Plavchan, AY 218
Field showing Spectral Dispersion

~ 6000 targets/CCD

5/19/09 Peter Plavchan, AY 218
Aperture masks
Pipeline architecture

How CoRoT planet detection works...

- Observations → Data Reduction → Transit Candidate List → Follow-up Observations → Confirmed Planets

- Transit alarm! → Preliminary candidate list (large planets!)

Giant (and even small ones) planets can be detected already in „alarm mode“!
CoRoT “raw” data (1st long run)

5/19/09 Peter Plavchan, AY 218

35 days
Stellar Variability

Object Name: 102693924

WHITEFLUX (ELECTRONS)

DATEHEL (Day)

X axis scaling: add 2000
Y axis scaling: multiply by 10^4

5/19/09  Peter Plavchan, AY 218
Stellar variability + Transit

Transiting planets around variable stars

5/19/09 Peter Plavchan, AY 218
CoRoT7

Raw light curve!

Lightcurve of 144d, demonstrating a rotation period of 22-23d

Cleaned and normalised ➔
150+ transits co-added
Systematic Noise Sources Particular to CoRoT

- Good overview/starting reference:
- Background, dark current, readout
- Jitter aka intra-pixel effects
  - Removed via pointing information from astero channel
- South Atlantic Anomaly, Earth shadow ingress and egress
- Hot pixel events
  - Remove with ~5 day exponential decay
- Linear trends (in flux)
- For given image #, relative magnitude as a function of apparent magnitude shows a “bias”, that is corrected with a quadratic
Frame-to-Frame Residuals Bias

Fitting a parabola to the residuals of one exposure

LRa01 E1 512sec JD 2853.2142
Three runs released to public as of today
- Initial Run (anti center)
- “Long Run” (center)
- “Short Run” (center)
  (included in the release was one corrupted FITS file)

Jitter effects corrected and removed (I think)

SAA photometry flagged, but not removed

Hot pixel effects / discontinuity jumps flagged, but not removed
  - Hard work done for you already

Pay attention to data quality flags

5/19/09  Peter Plavchan, AY 218
Raw Data Format

- FITS binary tables
  - 20 GB binary data (1 long run) → ~100 GB ASCII

- Exo data: (10,000s stars / run)
  - “chromatic” files
    - Stars where the S/N is sufficient (R<15) or science justified to perform aperture photometry on the “spectra” to produce “red”, “green” and “blue” fluxes. No overlap.
  - “monocromatic” files
    - Basically aperture photometry, with the shaped aperture

- Astero data: (typically 10 stars / run)
  - contains raw data, processed, different frames of reference for timing considerations

5/19/09 Peter Plavchan, AY 218
Getting the Data

- From the CoRoT website:
  - Provide IDL routines

- From NStED:
  - [http://nsted.ipac.caltech.edu](http://nsted.ipac.caltech.edu)
  - By contractual agreement, links to CoRoT for actual data download
  - Primarily useful as a target list builder and data visualizer... for now.

- From me in ASCII format?
  - Send me an email with SSH access to a machine with ~100 GB free space / CoRoTrun (plavchan@ipac)
  - I have some Perl code to process FITS files (requires PDL library)

- Nick Law has some Python code
Looking Ahead

- Light curves are currently “cleaned” of hot pixel events and other systematic noise sources on a source-by-source basis.
- There is a need for a generic tool to detrend any light curve, clean up the bad quality data, hot pixel events, allow for stellar variability filtering/fitting.
- 8 teams analyzing CoRoT data. Can you do better?
Useful links

- [http://132.149.11.177/COROT/](http://132.149.11.177/COROT/)  
  CoRoT website

  CoRoT Public Archive, Documentation

  Symposium Talks

- [http://nsted.ipac.caltech.edu](http://nsted.ipac.caltech.edu)  
  Data visualization and access
  - [http://nsted.ipac.caltech.edu/NStED/docs/datasethelp/ETSS_CoRoT.html](http://nsted.ipac.caltech.edu/NStED/docs/datasethelp/ETSS_CoRoT.html)

  Transit detection and characterization talks

- NASA ADS search with “CoRoT” in title words...