Variable Stars (53932: 5pts)

- Homepage of this course *http://www.helsinki.fi/ jetsu/vars/vars.html*
- General Catalogue of Variable Stars (GCVS) *http://www.sai.msu.su/groups/cluster/gcvs/gcvs/*
- GCVS contains over 43,000 stars
- An endless list of variable types *http://www.sai.msu.su/gcvs/gcvs/iii/vartype.txt*
- Every star is variable in some or several time scales
- Sun is a variable star
- We analyse unpublished standard Johnson V photometry of a chromospherically active star
- The name of the analysed object is **HD 8357**
- Aim: To publish the results in a refereed Journal
- First author = Best student
- Second author = Lecturer
- Other authors = Other passing students in alphabetical order
- IDL=Interactive Data Language (**Exercise 2** $\Rightarrow\sim$ 3 weeks)
- `ssh -X -lusername heaven.astro.helsinki.fi`

$\boxed{\text{Test now}}$

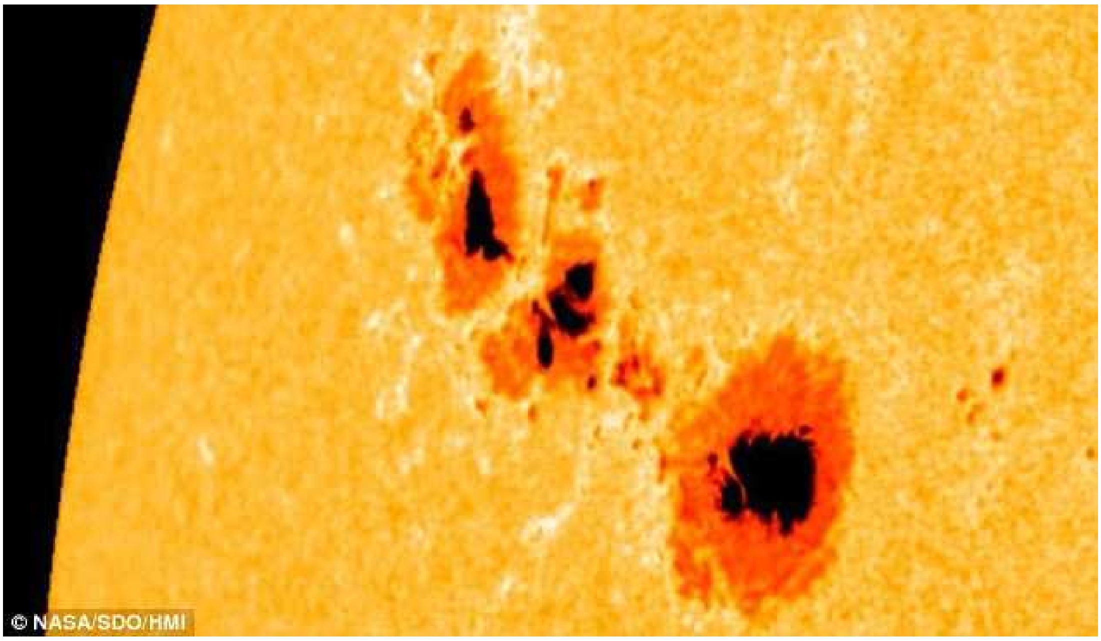
Sketch of stages of analysis

This sketch can change

1. Signatures of solar and stellar magnetic activity
2. Introduction: What is known about **HD 8357**?
3. Processing data (Segments, Datasets, Outliers ...)
4. Figure 1: All data, Table 1: Information of data
5. Light curve modelling: parameters
6. Activity cycle(s) – Power spectrum
7. Differential rotation – Period cahnges
8. Active longitudes – Solution of minima – Kuiper test
9. Time scale of change
10. Scientific writing
11. Scientific references
12. Comparison to Sun, and other active stars

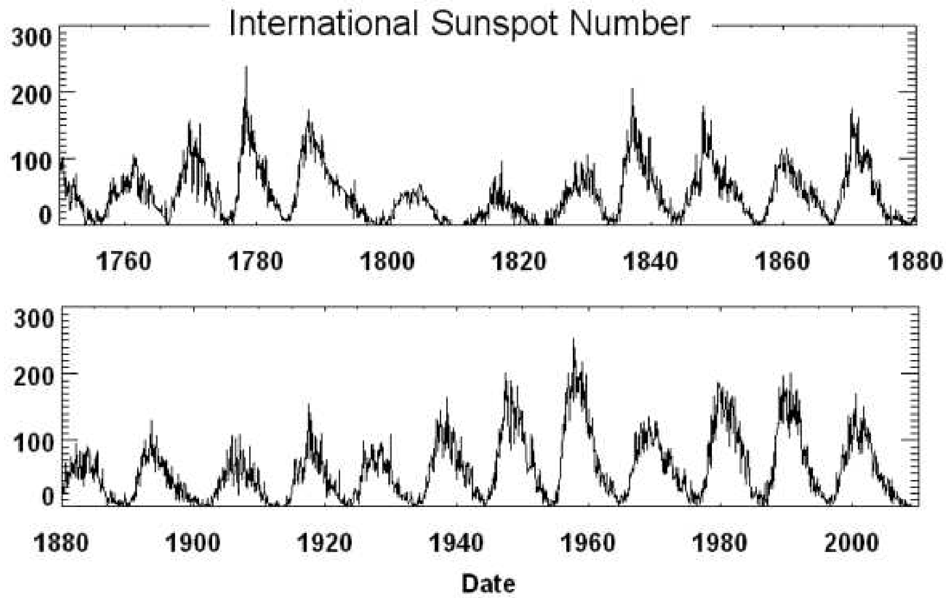
Sunspots and starspots

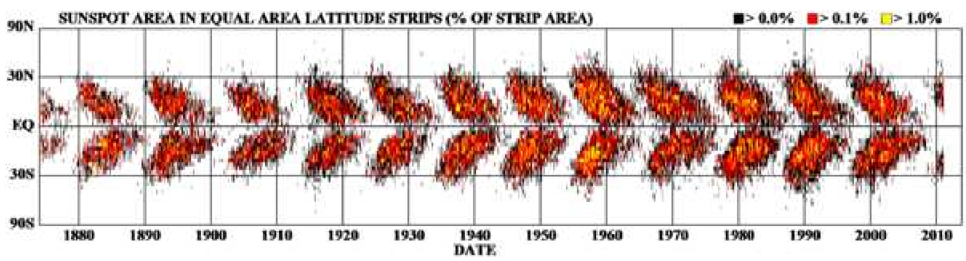
- Convection tranfers energy to surface surface
- Convection & differential rotation $\Rightarrow$ magnetic field
- Solar surface magnetic field strongest in sunspots
- Magnetic field weakens convection
- Solar photosphere about 6000 K
- Sunspots about 3000-4500 K $\Rightarrow$ dimmer $\Rightarrow$ darker
- Snspots: Luminosity changes of about 0.1%
- Starspots: Largest luminosity changes $\sim$ 30%
- **HD 8357**: What can we expect?



Sunspot and starspot cycles

- Sunspot cycle of 11 years
- Irregular periods from 9 to 14 years
- Cycle mean and amplitude also irregular
- No sunspots, e.g. Maunder minimum: 1645 – 1715
- Butterfly diagram $\Rightarrow$ Regular latitude changes during sunspot cycle
- **HD 8357**: What can we expect?

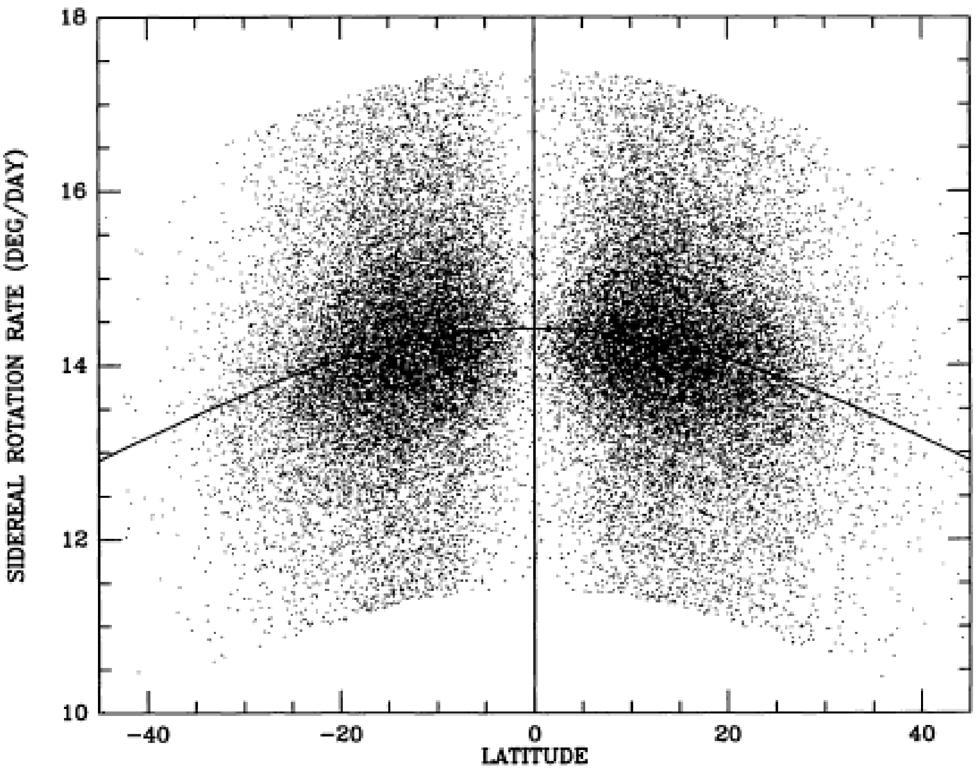




Solar and stellar differential rotation

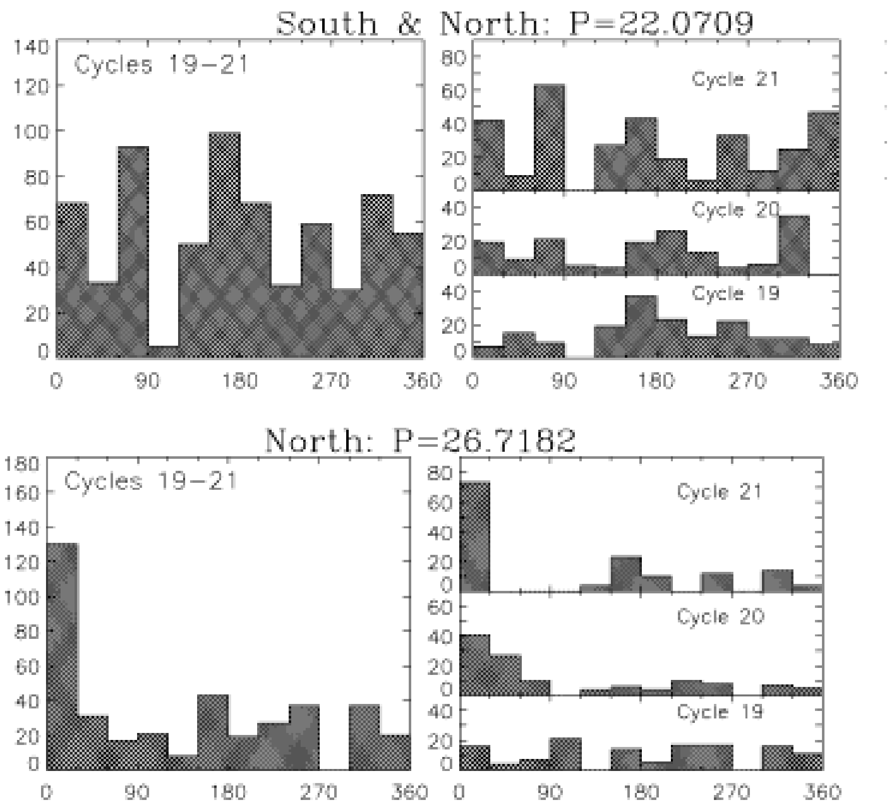
- Solar surface does not rotate like a rigid body
- Approximation of solar surface differential rotation

  $P(\Theta) = P_{\rm eq}/[1 - k\sin^2\theta]$
- $P_{\rm eq}$ = rotation period at the solar equator
- Θ = latitude
- $k \approx 0.2$ = constant
- $P_{\rm eq} \approx 25^{\rm d}$ (sidereal) $\approx 27^{\rm d}$ (synodic)
- k depends on tracers: sunspots, magnetograms, ...
- Figure below (from Howard 1994, PASPC 68, 1)
- Periof of any arbitrary sunspot is highly unpredictable
- **HD 8357**: What can we expect?

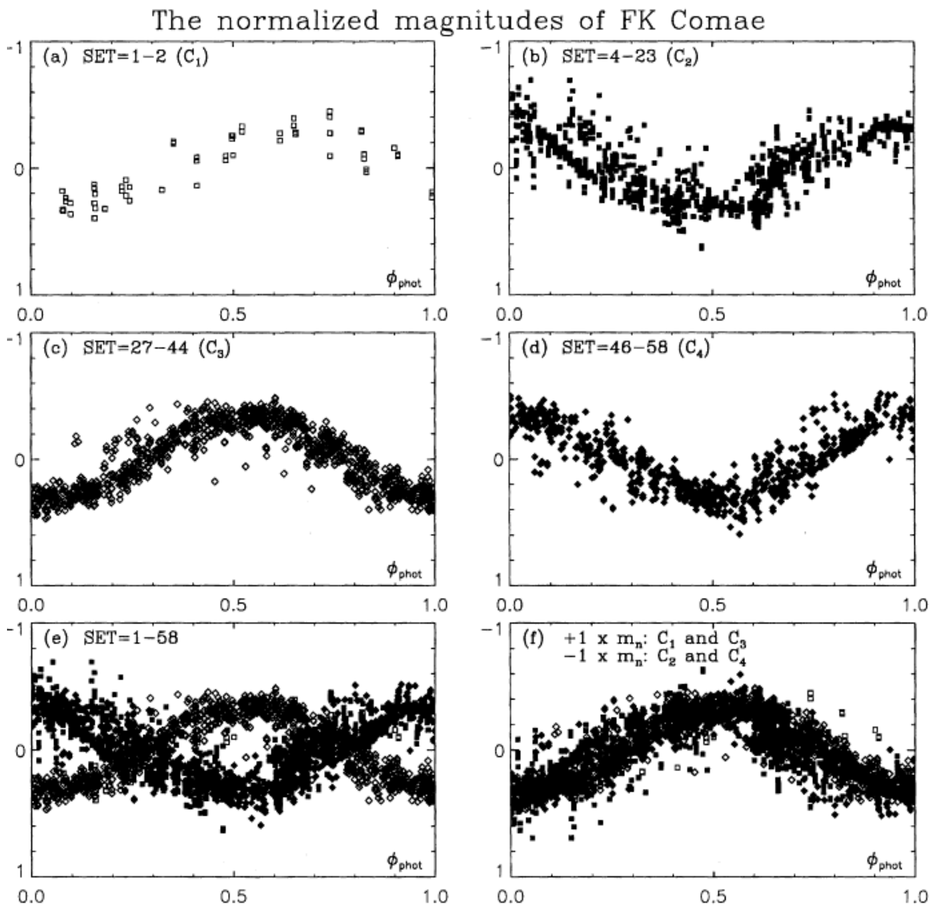


Solar and stellar active longitudes

- Sunspots: form and fade during one or two rotations
- Distibution longitudinally uniform
- Active regions: form and fade during 4 – 12 months
- Solar flares 1956-1985 (Jetsu et al. 1997, A&A 318, 293)
- Active longitudes rotating with 22.1 days
- Active zone in northern hemisphere 26.7 days



- Starspots: form and fade during X=? rotations
- Distribution longitudinally nonuniform
- Cause large rotational modulations of brightness
- Light curves of FK Comae (Jetsu et al. 1993, A&A 278, 449)
- Two stable active longitudes separated by 180 degrees
- "Flip-flop" phenomenon – quarter of a century
- **Some** rapidly rotating giants and subgiants: Yes
- **Most** rapidly rotating main sequence stars: No
- **HD 8357**: What can we expect?



Photometry

- Flux $[F]$ measured in a limited wavelength interval
- Magnitudes: $\mathrm{m} = -2.5\lg[F/F_0]$
- m increases as F decreases
- $0.^{\mathrm{m}}1 \approx 10\%$, $0.^{\mathrm{m}}01 \approx 1\%$ or $0.^{\mathrm{m}}001 \approx 0.1\%$
- Instrumental system photometry transformed to standard system photometry
- For example, Johnson UBVRI, Cousins $\mathrm{UBV(RI)_C}$, Strömgren uvby, ...
- Automatic Photoelectric Telescopes (APTs)
  http://schwab.tsuniv.edu/papers/html/precisepep/precisepep.html
- Differential photometry: Sky (S), Comparison star (C), Check star (K) and Variable (Va)
- APT measurement sequence of 10^{s} integrations
  $K - S - C - Va - C - Va - C - Va - C - S - K$
- $\Delta V_{K-C} = -2.5\lg[F_K/F_C] \approx$ constant $\approx \pm 0.^{\mathrm{m}}01 \Rightarrow$
  $\Delta V_{VA-C} = -2.5\lg[F_{VA}/F_C]$ variability is real

Your manuscipt

- You will learn how to find the references and other information of **HD 8357** from `ADS` and `Simbad` databases. Links to these databases are given in the course homepage.
- **Exercises 0A and 0B**

  All references are divided between all participants. Everyone reads about the same number of references.

  A: Copy files `model.tex` and `model.bib`, from the course homepage. Collect all your references into `model.bib`.

  B: Make a note of the pages (or possibly tables or figures), where **HD 8357** mentioned in your publications. Describe with a few sentences, why **HD 8357** has been included into the publication. Collect this information into `model.tex`

  Everyone reports during the next lectures, **what** is relevant about our **HD 8357** in his/her references?

∼ • ∼ Exercise ends here. ∼ • ∼

- **Preliminary** work before Exercise 1: Your lecturer compiles all information in all references into files given in the homepage. We will decide which references are relevant and which ones are not. Use these files in writing your introduction.

- **Exercise 1.**

  Write an **Introduction** to the paper of **HD 8357**. You can use section **Introduction** in Kahanpää et al. 1999, Lyytinen et al. 2002 Kajatkari et al. 2014, and Siltala et al. 2016 as a model. These papers can be found in the course homepage. Your manuscripts will be compared during the next lecture.

∼ • ∼ Exercise ends here. ∼ • ∼

Short LaTeX instructions

- `latex model`
- `bibtex model`
- `latex model`
- `latex model`
- `dvips model -o`
- You can see the resulting file with the commands

  `gv model.ps &`

  or

  `evince model.ps &`
- Note that the symbol `&` allows you to run the above five LaTeX commands at the same terminal.
- **If** LaTeX "gets stuck", do

  `rm model.aux`

  `rm model.bbl`

  `rm model.blg`

  `rm model.dvi`

  `rm model.log`

  `rm model.ps`
- **Never**, do

  `rm model.tex`

  `rm model.bib`
- Additional LaTeX information of the A&A journal:

  `http://ftp.edpsciences.org/pub/aa/readme.html`

Identifying outliers

How to use $\boxed{\text{emacs}}$ & $\boxed{\text{idl}}$ & $\boxed{\text{terminal}}$ efficiently!

- **Exercise 2.** The data format file is (9 first lines only)

```
2447987.8762 99.999  1.804  1.828 99.999 -0.815  0.177
2447988.9494 99.999  1.814  1.845 99.999 -0.798  0.209
2447989.9185 99.999  1.827  1.853 99.999 -0.820  0.180
2447990.7229 99.999  1.813  1.844 99.999 -0.818  0.189
2447990.8327 99.999  1.814  1.848 99.999 -0.799  0.203
2447991.7459 99.999  1.804  1.824 99.999 -0.813  0.180
2447992.7237 99.999  1.777  1.802 99.999 -0.813  0.173
2447992.9580 99.999  1.763  1.789 99.999 -0.816 99.999
2447993.7243 99.999  1.752  1.777 99.999 -0.817  0.182
```

The 1st column is time (t_i) in heliocentric julian days (abbreviation HJD). The 4th column is ΔV_{S-C} and 7th column is ΔV_{K-C}. The dummy value `99.999` means that no observation was obtained. You can copy this file from the homepage. Write an IDL program that

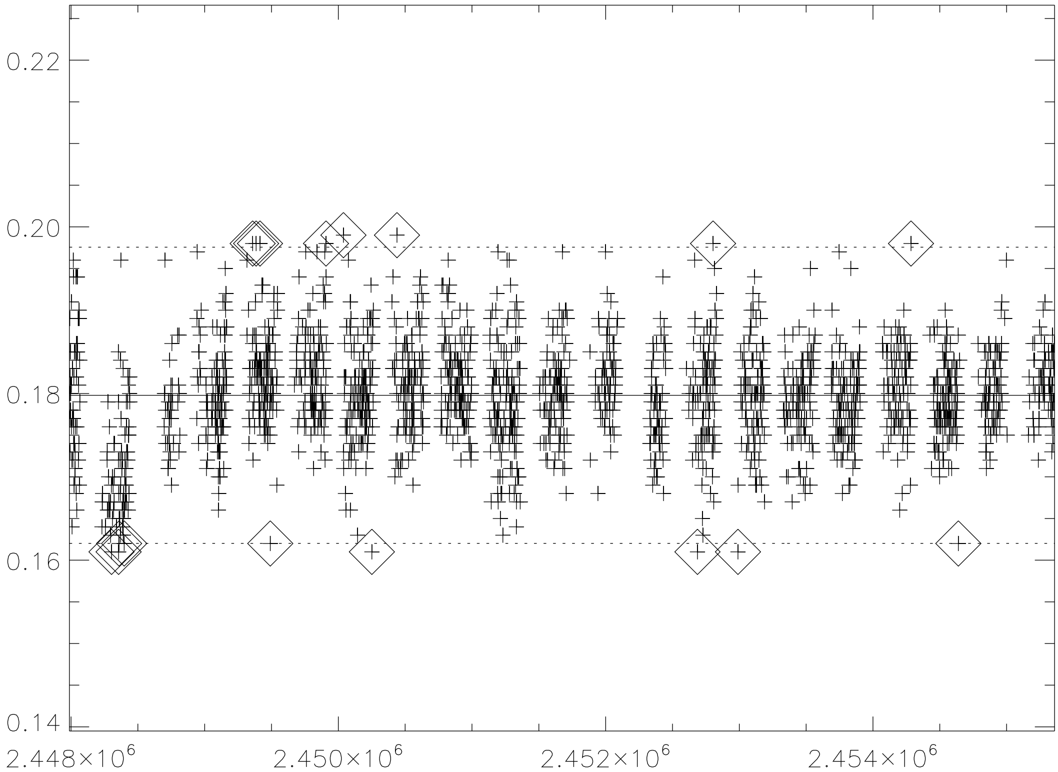
A: Plots the ΔV_{S-C} magnitudes as a function of time (`HJD`). Identify $y_i = \Delta V_{S-C}$ (Variable minus comparison star) that fulfil $|y_i - m| > 3s$, where m is the mean and s is the standard deviation of y_i. Write down those t_i and y_i values, e.g. into your `model.tex`. Overplot the $y = m$ and $y = m \pm 3s$ levels. Use different symbols for the outliers, i.e. the data that fulfils the above criterion.

B: Do the same for the $y_i = \Delta V_{K-C}$ (Check star minus comparison star) data.

$\sim \bullet \sim$ **Exercise ends here.** $\sim \bullet \sim$

One type of a graphical solution for Exercise 2 is shown below.

Note: This result was obtained after removing two times the $y_i = \Delta V_{K-C}$ values that exceed $m \pm 3s$ limit.



Activity cycles

- τ = mean of all observing times t_i of a dataset
- Activity cycles sometimes detected:

  $M(\tau) \equiv$ longitudinally uniform spot distribution

  $A(\tau) \equiv$ longitudinally nonuniform spot distribition

  $M(\tau) + A(\tau)/2 \equiv$ maximum spot coverage

  $M(\tau) - A(\tau)/2 \equiv$ minimum spot coverage
- Scargle (1982: ApJ 263, 835) power spectrum for data $y_i = y(t_i)$ at any arbitrary tested frequency f is

  $$z_{\mathrm{LS}}(f) = \frac{[\sum_{\mathrm{i=1}}^{\mathrm{n}} y'_{\mathrm{i}} \cos 2\pi f(t_{\mathrm{i}} - \delta)]^2}{2\sum_{\mathrm{i=1}}^{\mathrm{n}}[\cos 2\pi f(t_{\mathrm{i}} - \delta)]^2} + \frac{[\sum_{\mathrm{i=1}}^{\mathrm{n}} y'_{\mathrm{i}} \sin 2\pi f(t_{\mathrm{i}} - \delta)]^2}{2\sum_{\mathrm{i=1}}^{\mathrm{n}}[\sin 2\pi f(t_{\mathrm{i}} - \delta)]^2},$$

  where $\hat{y}_i = [\sum y_i]/n$, $y'_i = y_i - \hat{y}_i$ and δ is defined by

  $$\tan(4\pi f\delta) = \left(\sum_{\mathrm{i=1}}^{\mathrm{n}} \sin 4\pi f t_{\mathrm{i}}\right)\left(\sum_{\mathrm{i=1}}^{\mathrm{n}} \cos 4\pi f t_{\mathrm{i}}\right)^{-1}.$$
- Note: subtract the mean $\hat{y}_i = [\sum y_i]/n$ from y_i
- Note: this $z_{\mathrm{L}S}(f)$ was programmed in "Practical methods in astronomy" using `python`

- Horne & Baliunas (1986: ApJ 302, 757) defined

  $z_{\mathrm{H}B}(f) = z_{\mathrm{L}S}(f)/\sigma^2,$

  where σ^2 is the variance of the data y_i.
- The probability that $z_{\mathrm{H}B}(f)$ exceeds z_0 at one arbitrary tested frequency f is $P(z_{\mathrm{H}B}(f) > z_0) = \mathrm{e}^{-z_0}$.
- Tested frequencies between f_{min} and f_{max} were

  $\Delta T = t_{\mathrm{max}} - t_{\mathrm{min}} = t_{\mathrm{n}} - t_1$

  $f_{\mathrm{min}} = 1/\Delta T \Rightarrow P_{\mathrm{max}} = \Delta T$

  $f_{\mathrm{max}} = n/(2\Delta T) \Rightarrow P_{\mathrm{min}} = 2\Delta T/n$
- Number of independent tested frequencies was

  $m = -6.362 + 1.193\, n + 0.00098\, n^2$
- False alarm probability for $z_{\mathrm{HB}}(f)$ reaching z_0 was

  $F = 1 - [1 - \mathrm{e}^{-z_0}]^m$
- Probability for the data containing a signal was $1 - F$

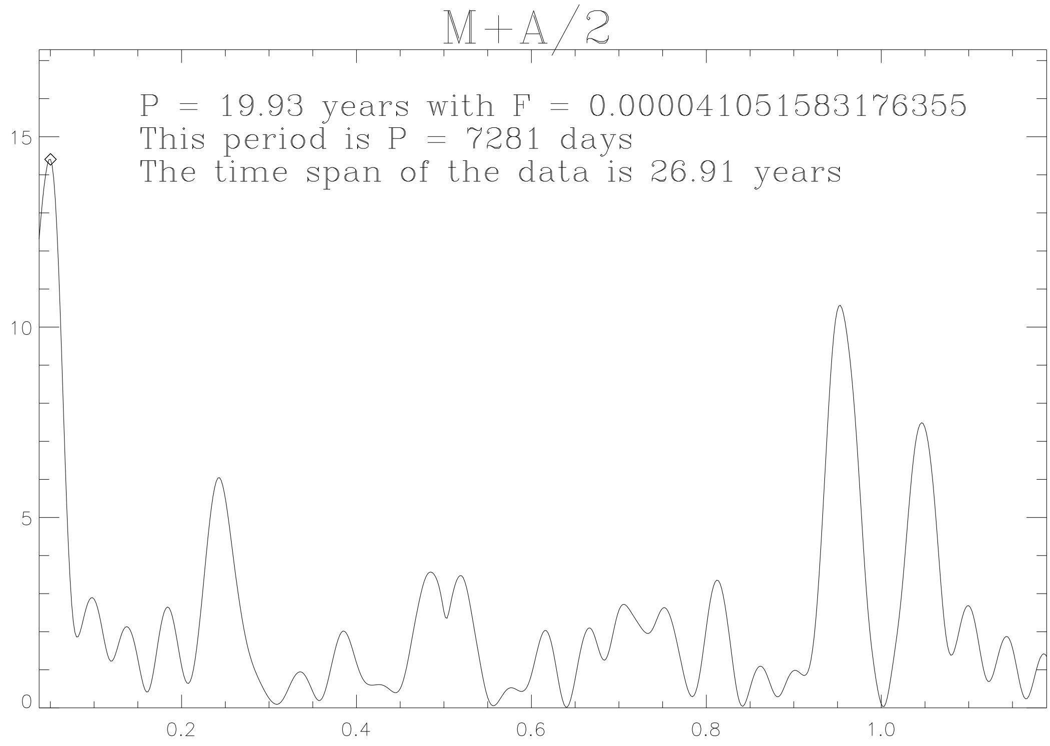
- **Exercise 3.** Compute and plot $z_{\mathrm{HB}}(f)$ for the four estimates $M(\tau)$, $A(\tau)$, $M(\tau)+A(\tau)/2$ and $M(\tau)-A(\tau)/2$ of all independent datasets ($IND=1$). Solve the best period and the false alarm probability F of this period for each of these four estimates. If $K=0$, convert the values $A(\tau)=-1$ to the values $A(\tau)=0$ when calculating $M(\tau)+A(\tau)/2$ or $M(\tau)+A(\tau)/2$. You can copy the file containing the $M(\tau)$ and $A(\tau)$ estimates from the homepage. The file format is

```
Data file: arpsc.16, P_0=       12.380000, GAP=       30.000000

           1     47112.8514           1
  47112.8514     47141.7536  47129.1248             1
          15         1.2819      0.0887
           1         0.0684     -1.0000
      1.2912         0.0200           1
     10.3788         1.4134           1
      0.1834         0.0509           1
  47119.3589         1.4865           1
     -1.0000        -1.0000           1

           1     47112.8514           2
  47173.6609     47203.5844  47191.4821             1
          15         1.2192      0.0318
           1         0.0114     -2.0000
      1.2159         0.0032           0
     11.7449         0.2756           0
      0.0801         0.0078           0
  47179.0753         0.3623           0
     -1.0000        -1.0000           1
```

This format is specified in Lehtinen et al. (2011: Appendix on page 14) and we will also discuss about it during lectures. Test all integer multiples of $f_0=1/(10\Delta T)$ between $f_{\min}=1/\Delta T$ and $f_{\max}=n/(2\Delta T)$.

∼ • ∼ Exercise ends here. ∼ • ∼

One graphical solution for **Exercise 3** is given below, where the τ values have been transformed in years before analysis.



Kuiper method (K-method)

- A nonparametric method by Kuiper (1960), which searches for periodicity in circular data. This method is described in Jetsu & Pelt (1996, A&AS 118, 598)
- "Original" version by Kuiper (1960) did not utilize the information contained by the weights $w_{\mathrm{i}} = \sigma_{\mathrm{t_i}}^{-2}$. This is not required in the next **Exercise 4**.
- **Problem:** Is the distribution of the phases for n time points $t_1, ..., t_{\mathrm{n}}$ periodic with an arbitrary period P?

K–method

- **Step 1**: Calculate the phases $\phi_i = \mathrm{FRAC}[(t_i - t_0)/P]$ of the time points t_i with the tested period P, where $\mathrm{FRAC}[x]$ removes the integer part of x. Note that t_0 can be any arbitary epoch before t_1.
- **Step 2**: Arrange the phases ϕ_1, ..., ϕ_{n} into ascending order (i.e. rank order). Denote these phases by x_1, x_2, ... and x_{n}. For example, the smallest ϕ_{i} value is x_1.
- **Step 3**: Compute the **sample distribution function**

$$\mathbf{F}_{\mathrm{n}}(\phi) = \begin{cases} 0, & \phi < x_1 \\ in^{-1}, & x_{\mathrm{i}} \le \phi < x_{\mathrm{i+1}},\ 1 \le i \le n-1 \\ 1, & \phi \ge x_{\mathrm{n}} \end{cases}$$

Note: $F_{\mathrm{n}}(\phi)$ is a "step" function which has two values at $\phi = x_i$. These two values are $F_{\mathrm{n}}(\phi = x_i) = (i-1)/n$ and i/n.

- **Step 4**: The "null hypothesis" (H_0) is: *"The ϕ_{i} represent a random sample drawn from a distribution, which has a* **cumulative distribution function** *$F(\phi)$".*

Our H_0 is: *The ϕ_{i} distribution over $[0, 1)$ is uniform, i.e. a random distribution, i.e.* $\boxed{\mathbf{F}(\phi) = \phi}$

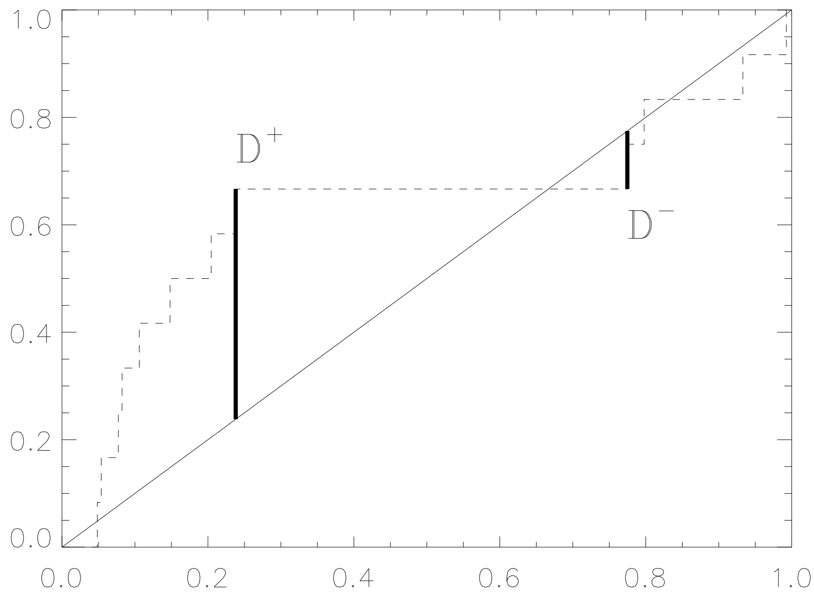
- **Step 5**: Fix the preassigned significance level γ for rejecting H_0. This parameter represents the probability of falsely rejecting H_0 when it is in fact true.
- **Step 6**: Compute the K–method test statistic:

  $V_{\rm n} = D^+ + D^-,$

  where D^+ and D^- denote the maximum values of $F_{\rm n}(\phi)$–$F(\phi)$ and $F(\phi)-F_{\rm n}(\phi)$.
- **Step 7**: Under H_0, the parameter $Q_{\rm K}$ is called the critical level. It is the probability that a particular value of $z = n^{1/2}V_{\rm n}$, or an even larger value, can occur. This critical level is

$$Q_{\rm K} = P(n^{1/2}V_{\rm n} \geq z) = P_{\rm n}(z) = \sum_{\rm k=1}^{\infty} 2\left(4k^2z^2 - 1\right){\rm e}^{-2k^2z^2} - (8z/3)n^{-1/2}\sum_{\rm k=1}^{\infty} k^2\left(4k^2z^2 - 3\right){\rm e}^{-2k^2z^2} + {\rm O}(n^{-1}),$$

  where ${\rm O}(n^{-1}) \approx 0$ when $n \geq 20$.
- **Step 8**: Reject H_0, if and only if $Q_{\rm K} \leq \gamma$
- **Note:** H_0 rejection more probable when $V_{\rm n}$ increases. Draw.



- **Exercise 4.** We continue to analyse the same file as in Exercise 3. Compute the Kuiper test periodogram $V_{\mathrm{n}}(f)$ for the $t_{\mathrm{min},1}$ estimates of independent datasets (IND=1). Test the period interval between $P_{\mathrm{min}} = 10.5$ and $P_{\mathrm{max}} = 14.0$ days. Test all integer multiples of the frequency $f_0 = [(OFAC)(\Delta T)]^{-1}$ between $f_{\mathrm{min}} = P_{\mathrm{max}}^{-1}$ and $f_{\mathrm{max}} = P_{\mathrm{min}}^{-1}$, where $OFAC = 15$ and ΔT is time interval between the first and last analysed $t_{\mathrm{min},1}$ estimate. Plot the periodogram $V_{\mathrm{n}}(f)$, indicate the position of the highest peak and give the value of the best period connected to this peak (typical graphical solution below).

  Some helpful comments

  Ignore **Steps 5, 7** and **8**, because no Q_{K} solution is required in this exercise. In the loop of tested frequencies:

  Step 1: Calculate the phases ϕ_i with the tested $P = 1/f$

  Step 2: Rank ϕ_i and denote with $x_1 \leq x_2 \leq, ..., \leq x_{\mathrm{n}}$

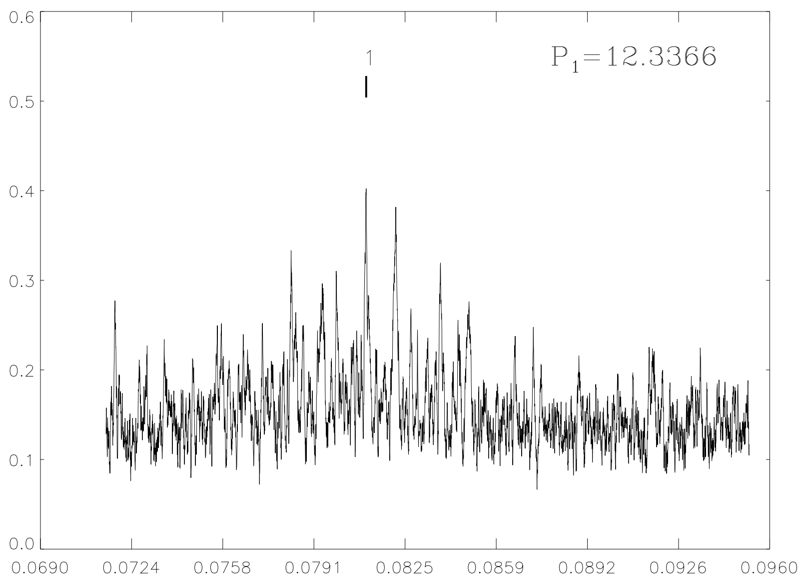
  Step 3: Compute sample distribution function $F_{\mathrm{n}}(\phi)$

  Step 4: Compute cumulative distribution function $F(\phi) = \phi$

  Step 6: Compute $V_{\mathrm{n}} = D^+ + D^-$.

  $\Longrightarrow$ Back to **Step 1** to test the next $P = 1/f$

  One example of a solution is shown in the figure below

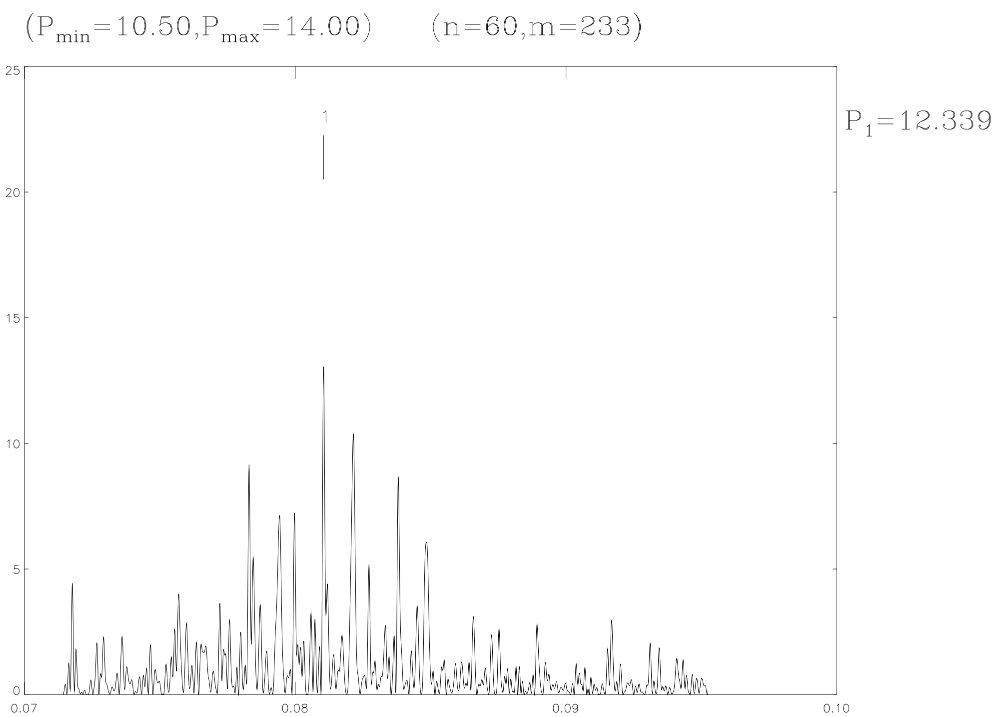


Rayleigh method (R-method)

- Analysing n time points $t_1, t_2, ..., t_n$ with Rayleigh test
- Period is P. Frequency is $f = 1/P$. Zero epoch in time is t_0
- Phases (ϕ is pronounced "phi") are
  $\phi_i = \text{FRAC}[\frac{(t_i - t_0)}{P}] = \text{FRAC}[f(t_i - t_0)]$
- FRAC[x] removes integer part of x, e.g. FRAC[4.12]=0.12
- Angles are $\Theta_i = 360^{\text{o}}\phi_i$ (degrees) or $\Theta_i = 2\pi\phi_i$ (radians)
- The unit vectors are $\bar{r}_i = [\cos\Theta_i, \sin\Theta_i]$.
- Their sum is $\bar{R} = \Sigma_{i=1}^{i=n} \bar{r}_i = \bar{r}_1 + \bar{r}_2 + ... + \bar{r}_n$
- $\bar{r}_1 = \bar{x}_1 + \bar{y}_1$, $\bar{r}_2 = \bar{x}_2 + \bar{y}_2$, ..., $\bar{r}_n = \bar{x}_n + \bar{y}_n$
  $\Rightarrow \bar{R} = \bar{x}_1 + \bar{x}_2 + ...\bar{x}_n + \bar{y}_1 + \bar{y}_2 + ...\bar{y}_n =$
- $\bar{R} = \bar{X} + \bar{Y}$
- $\bar{X} = \Sigma_{i=1}^{i=n} \bar{x}_i = \Sigma_{i=1}^{i=n}[\cos\Theta_i, 0] = [\cos\Theta_1 + \cos\Theta_2 + ... + \cos\Theta_n, 0] = [\Sigma_{i=1}^{i=n} \cos\Theta_i, 0]$
- $\bar{Y} = \Sigma_{i=1}^{i=n} \bar{y}_i = \Sigma_{i=1}^{i=n}[0, \sin\Theta_i] = [0, \sin\Theta_1 + \sin\Theta_2 + ... + \sin\Theta_n] = [0, \Sigma_{i=1}^{i=n} \sin\Theta_i]$
- $|\bar{X}|^2 = (\Sigma_{i=1}^{i=n} \cos\Theta_i)^2 + 0^{\ 2} = (\Sigma_{i=1}^{i=n} \cos\Theta_i)^2$
- $|\bar{Y}|^2 = 0^{\ 2} + (\Sigma_{i=1}^{i=n} \sin\Theta_i)^2 = (\Sigma_{i=1}^{i=n} \sin\Theta_i)^2$
- Pythagoras $\Rightarrow |\bar{R}|^2 = |\bar{X}|^2 + |\bar{Y}|^2$
- Final result $|\bar{R}|^2 = (\Sigma_{i=1}^{i=n} \cos\Theta_i)^2 + (\Sigma_{i=1}^{i=n} \sin\Theta_i)^2$
- Rayleigh test statistic for n phase angles $\Theta_1, \Theta_2, ..., \Theta_n$ is
  $z = \frac{|\bar{R}|^2}{n} = \frac{(\Sigma_{i=1}^{i=n} \cos\Theta_i)^2 + (\Sigma_{i=1}^{i=n} \sin\Theta_i)^2}{n}$
- $m = (f_{max} - f_{min})/f_0$ is the number of independent tested frequecies, where $f_0 = 1/(\Delta T)$ and $\Delta T = t_n - t_1$ is the time span of data
- $Q = 1 - (1 - \text{e}^{-z_0})^m$ is the probability for a $z(f)$ periodogram peak reaching the height of z_0

- **Exercise 5.** Compute the Rayleigh test periodogram $z(f)$ for the $t_{\mathrm{min},1}$ estimates of independent datasets, i.e. IND=1. Test the same frequencies as in Exercise 4. Give the value of the best period connected to the highest $z(f)$ peak. A typical graphical solution is shown below. Completing this assignment gives you 2pts.

  The solution for the best period is nearly, but not exactly, the same as in Exercise 4. The shape of the periodogram also resembles that in the previous exercise. What could be the reason? Answer this question with a few sentences and send your answer to the assistant before the lecture. A correct answer gives you one additional point, i.e. you may get 3pts/2pts. (Hint: see Jetsu & Pelt, 2000, A&A 353, 409: Figure 3).



Period changes

- The figure below shows all independent P_i values of **HD 8357** in the file `arpsc.orig.cl.res`
- The notations in the figure below are

– These independent $P_i \pm \sigma_{P_i}$ values are used to calculate the parameter
$Z = 6\Delta P_{\mathrm{w}}/P_{\mathrm{w}}$,
where $P_{\mathrm{w}} = [\sum_{\mathrm{i=1}}^{\mathrm{n}} w_{\mathrm{i}} P_{\mathrm{i}}]/[\sum_{\mathrm{i=1}}^{\mathrm{n}} w_{\mathrm{i}}]$,
$\Delta P_{\mathrm{w}} = \sqrt{\sum_{\mathrm{i=1}}^{\mathrm{n}} w_{\mathrm{i}}(P_{\mathrm{i}} - P_{\mathrm{w}})^2}/\sqrt{\sum_{\mathrm{i=1}}^{\mathrm{n}} w_{\mathrm{i}}}$ and $w_i = 1/(\sigma_{P_i})^2$.

- **Exercise 6.** Edit a program that produces the figure below. The P_w level is denoted with a continuous line. The $P_w \pm 3\Delta P_w$ levels are denoted with dotted lines. Completing this assignment gives you 2pts.
Read the text connected to Z in Lehtinen et al. 2011 (A&A, 527, A136). Explain with less than 100 words, what could this Z parameter measure in **HD 8357**? A correct answer gives you one additional point, i.e. you may get 3pts/2pts.