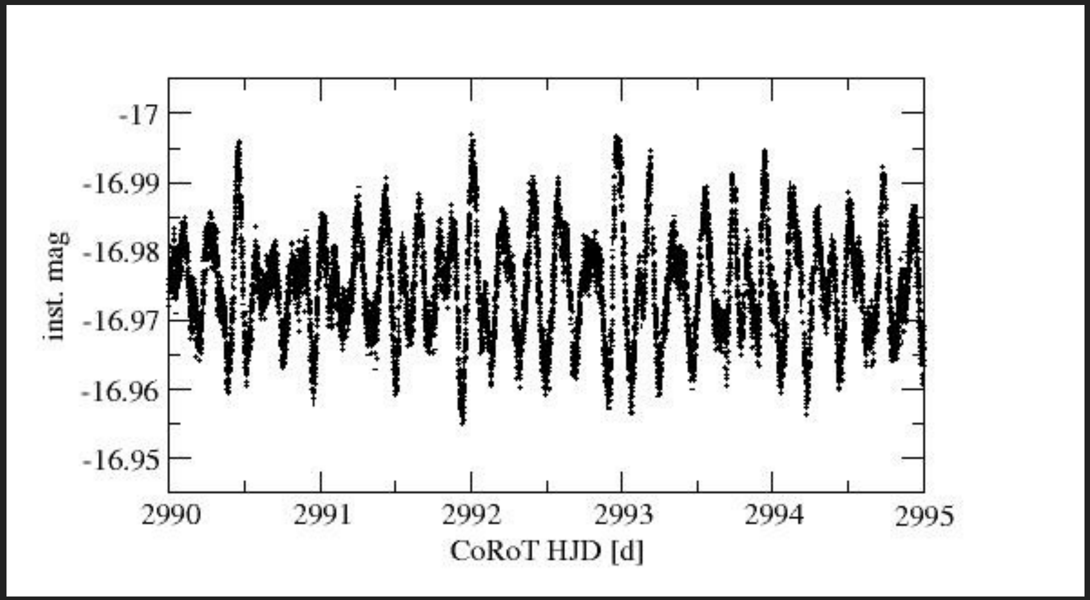
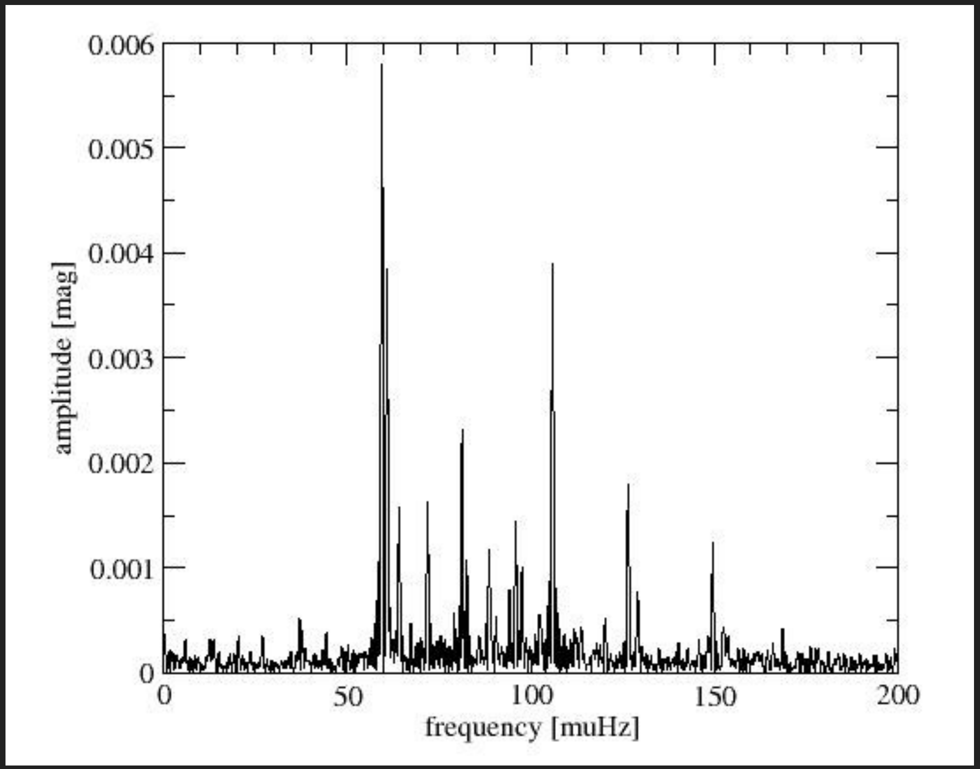


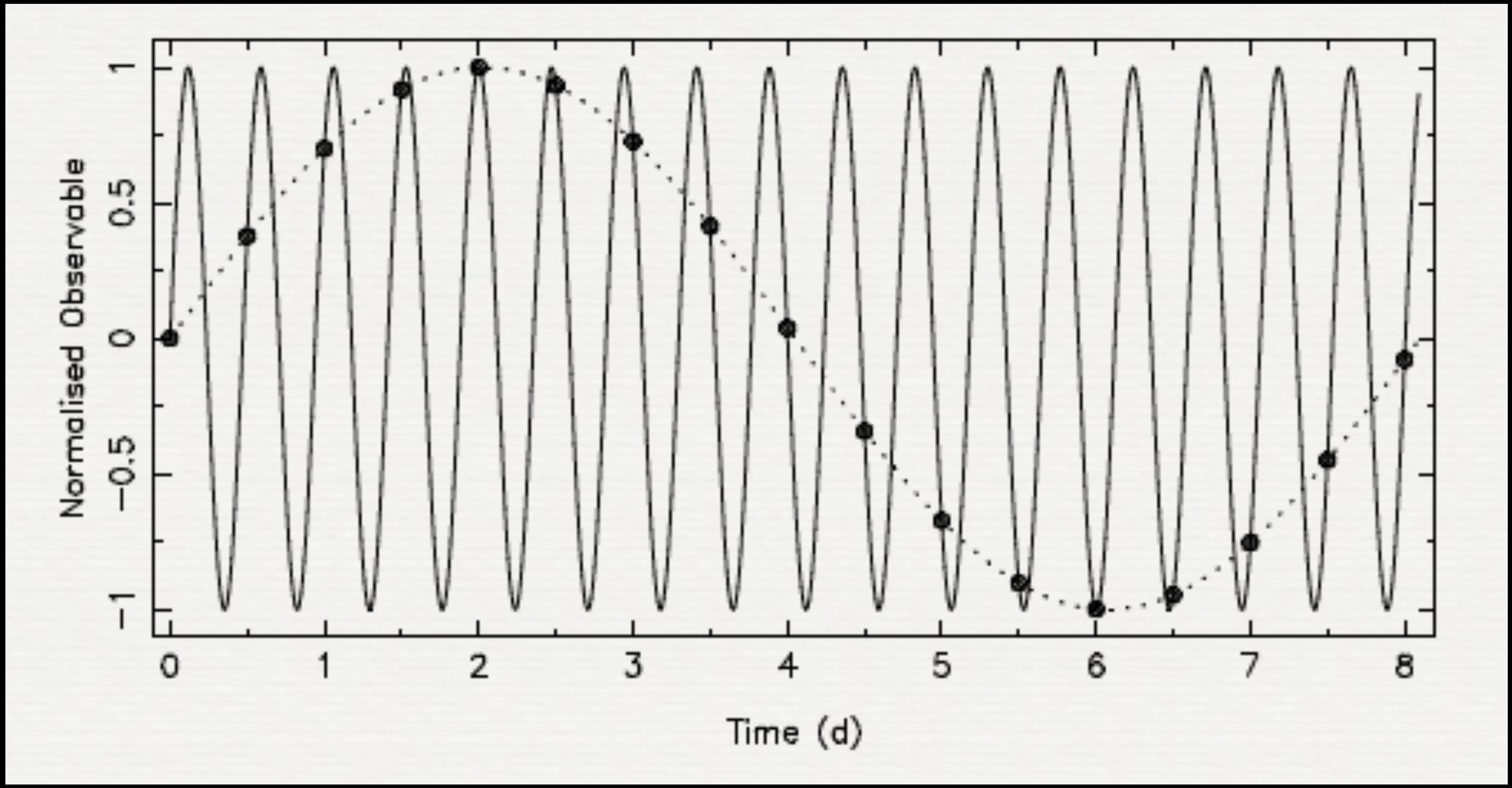


TIME SERIES ANALYSIS

WE ARE DEALING WITH THE TOUGHEST CASES:

TIME SERIES OF UNEQUALLY SPACED
AND
GAPPED ASTRONOMICAL DATA

A PERIODIC SIGNAL

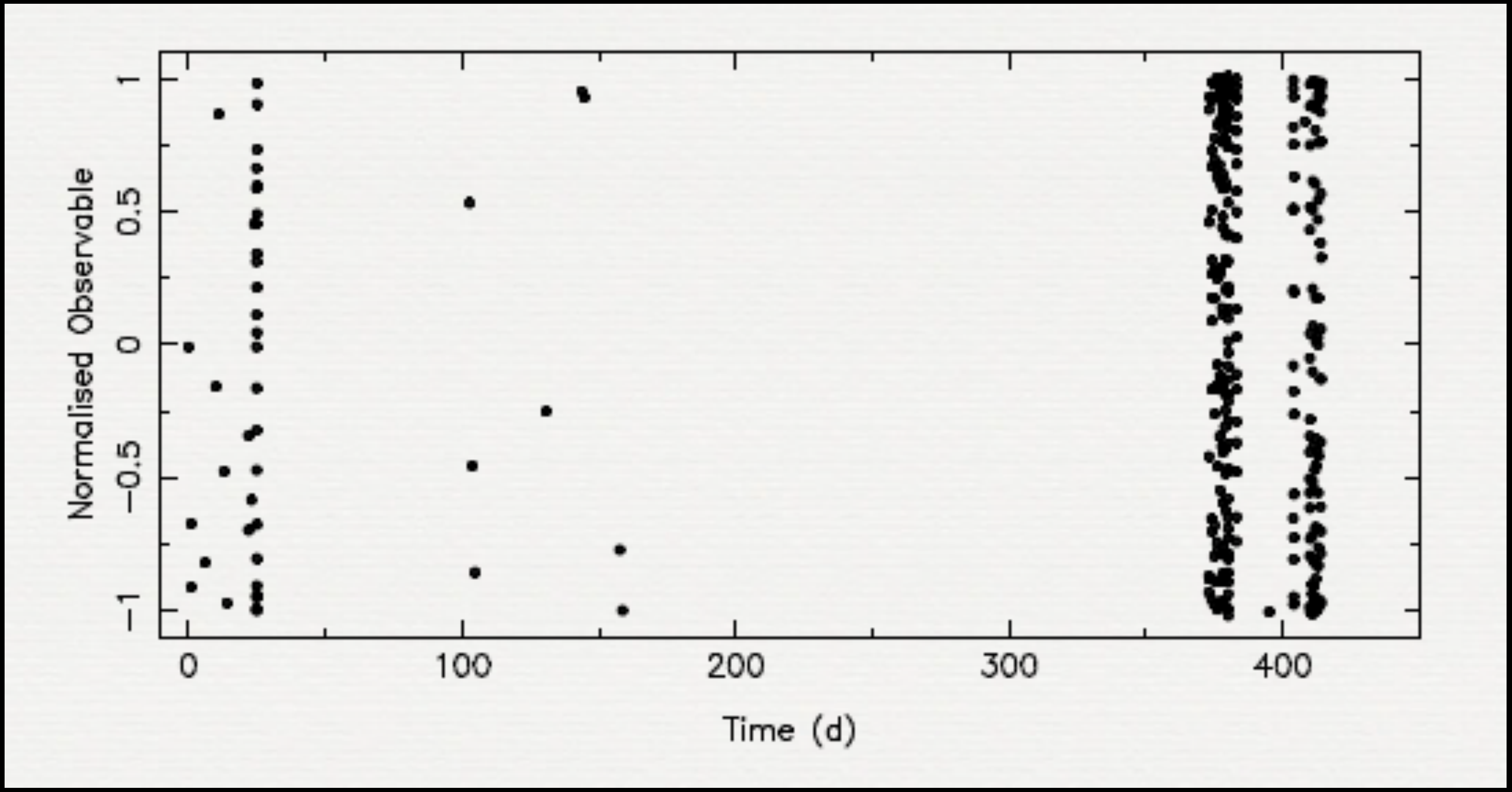


Dots: periodic signal with frequency f = 0.123456789 d^{-1}.

Dotted line: fit for this frequency.

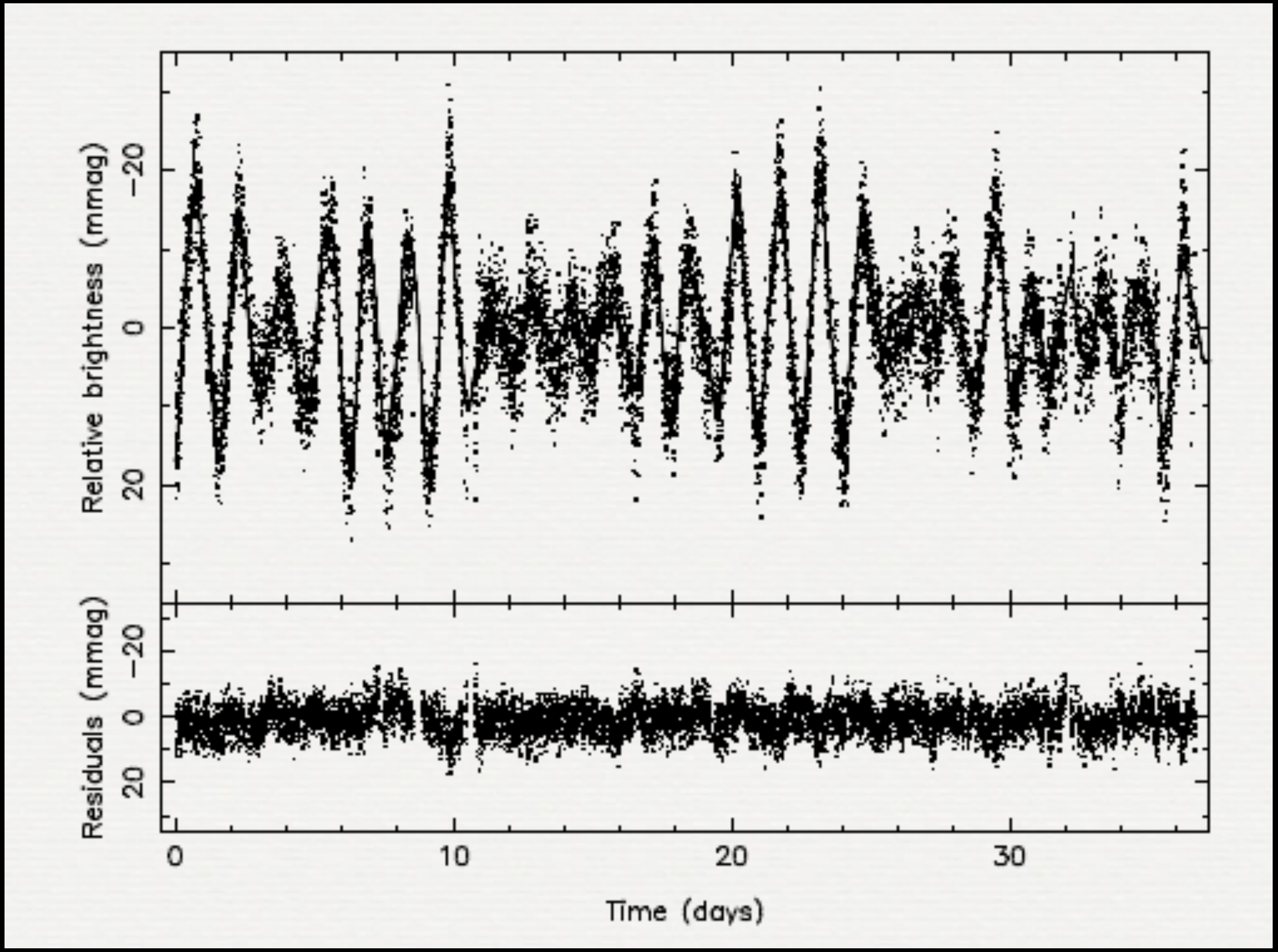
Full line: fit with frequency 2.123456789 d^{-1}.

GAPPED DATA



Simulated gapped data - typical time series for a single-site observing run of a pulsating star

SPACE DATA



Typical time series of a pulsating star as observed from space

FREQUENCY DETERMINATION

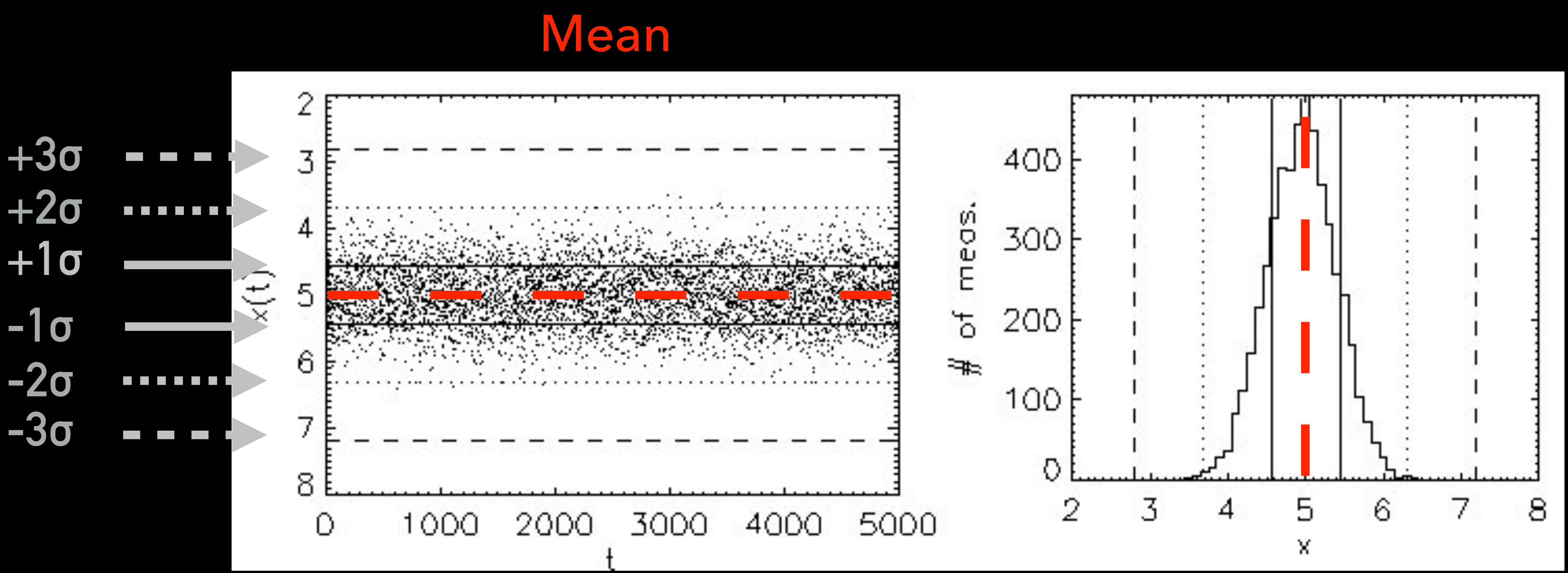
- Aim: Determination of previously unknown periodicities
- We use mathematical and statistical tests for each data set that shows temporal variability. With this we describe the variability itself and through the variability we learn about the physical system (=star) that is responsible for the variability.
  - Mean:

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

  - Variance: measures how far a set of numbers is spread out

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

MEAN & VARIANCE



68% of all data points lie within 1 sigma

99.7% lie within 3 sigma

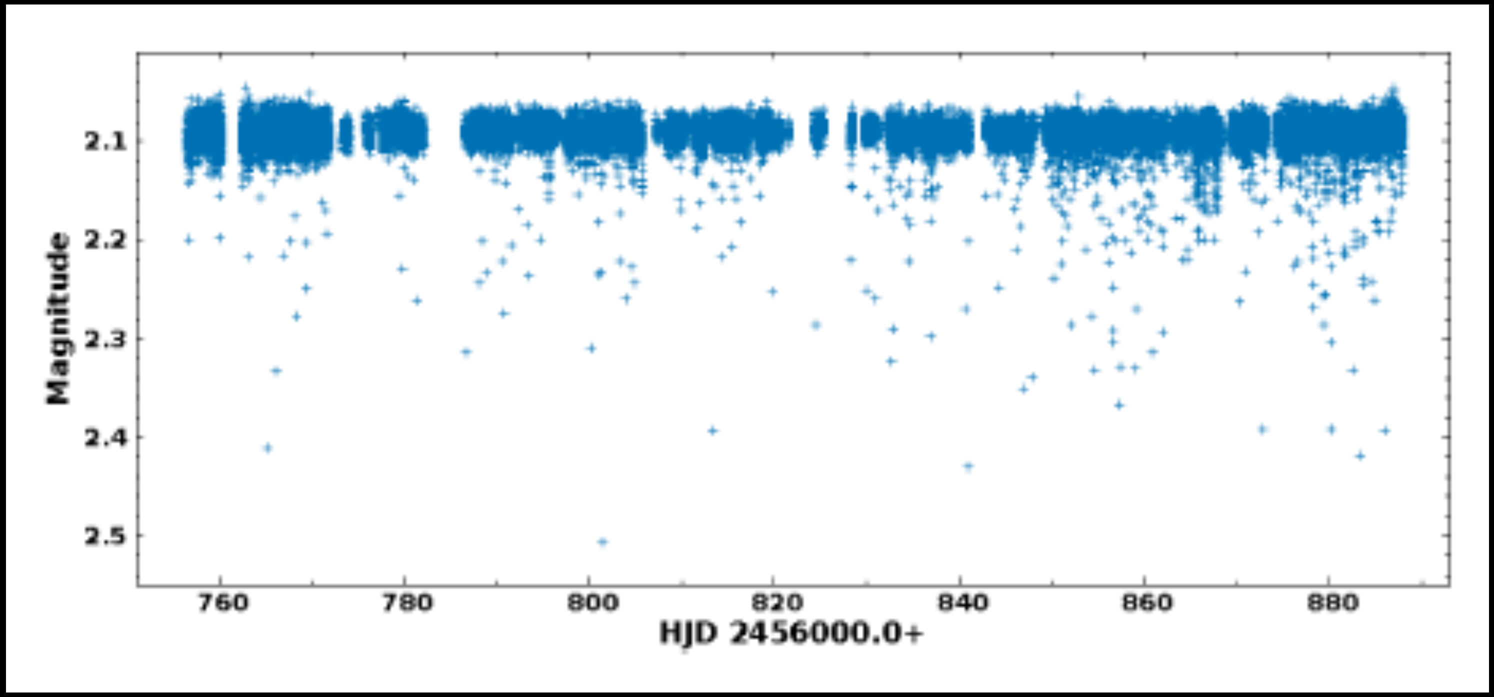
REALITY CHECK

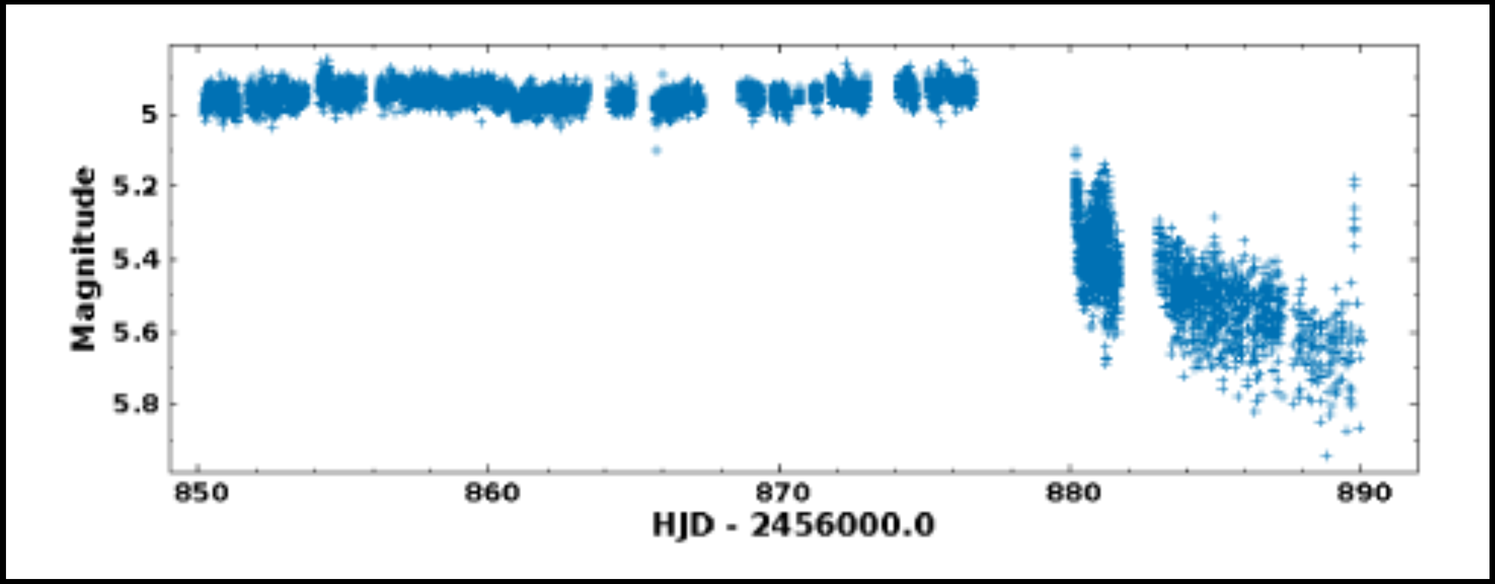
- More complicated in reality… variability of data influences statistics
- Examples:
  - short period variations “disappear” in mean;
  - for large amplitude variables (e.g., Mira) a mean magnitude does not make a lot of sense
  - there might be systematic errors in the data
- **Random errors** are always included in the measurements even if the quality is exceptionally good!
- **Keep checking for systematic errors throughout any analysis!**
- **Message: first you need to understand and characterize your data, then you can start to analyze them!**

UNDERSTANDING YOUR DATA

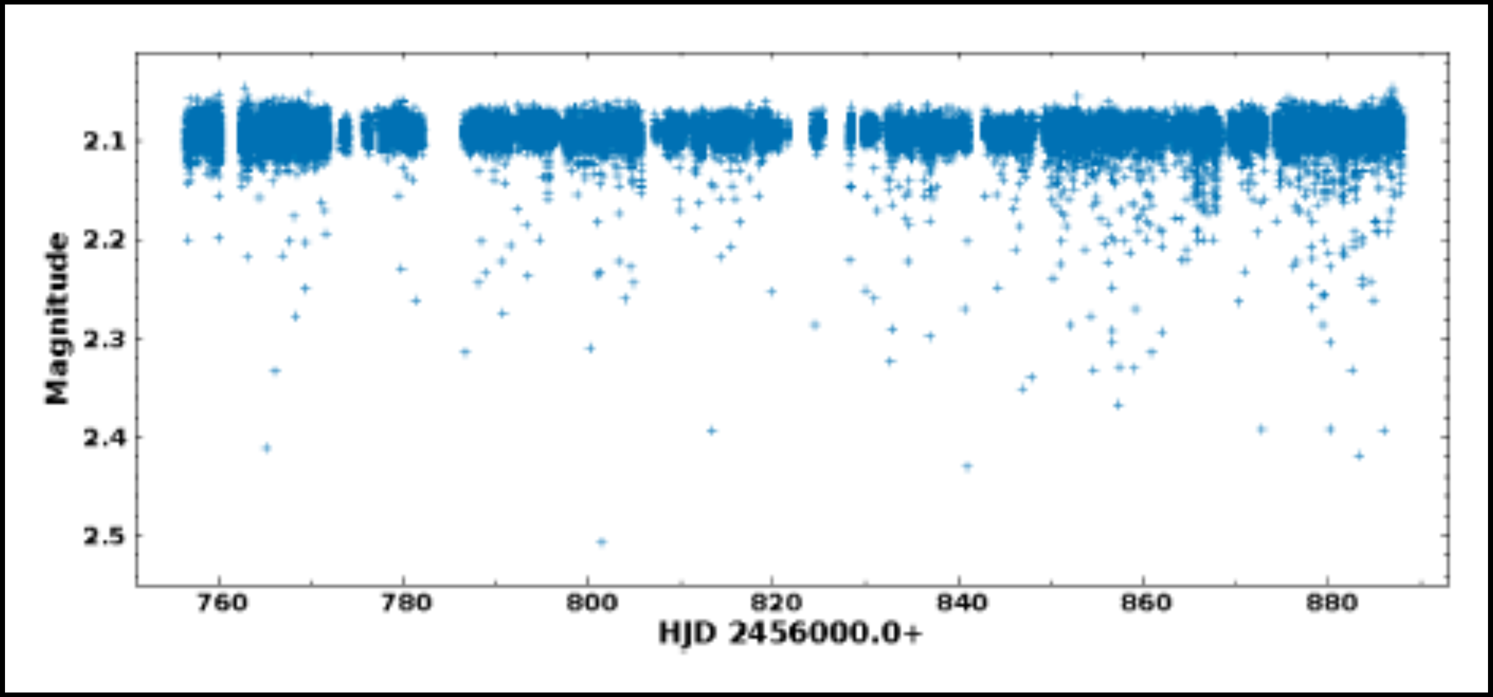
- What is the time sampling?
- Are there gaps in the data? If so, what caused them?
- Is the scatter (variance / sigma) reasonable or too high?
  - If it is very high, are there a lot of outliers? What might have caused outliers? Can we correct them or do we need to remove selected outlier points?
- Are there significant changes in the shape of the light curve?

EXAMPLES OF REAL DATA



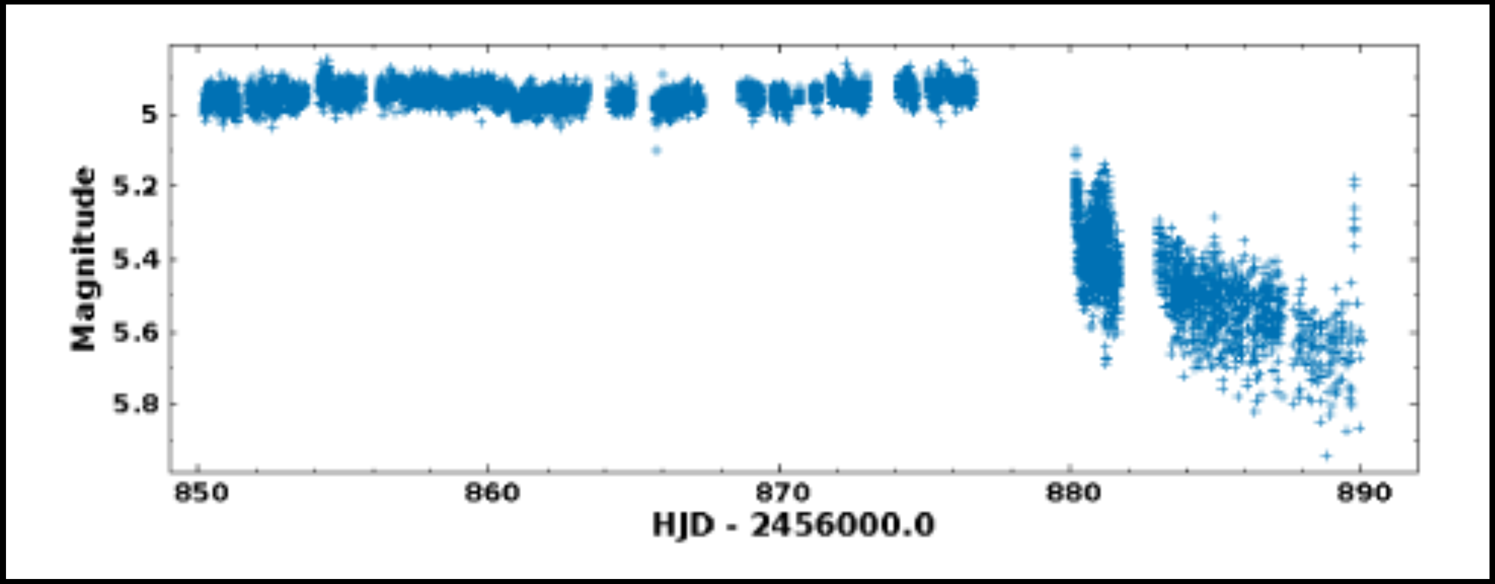


EXAMPLES OF REAL DATA



Gaps

Outliers



Change of "shape"

THE BEST PERIOD

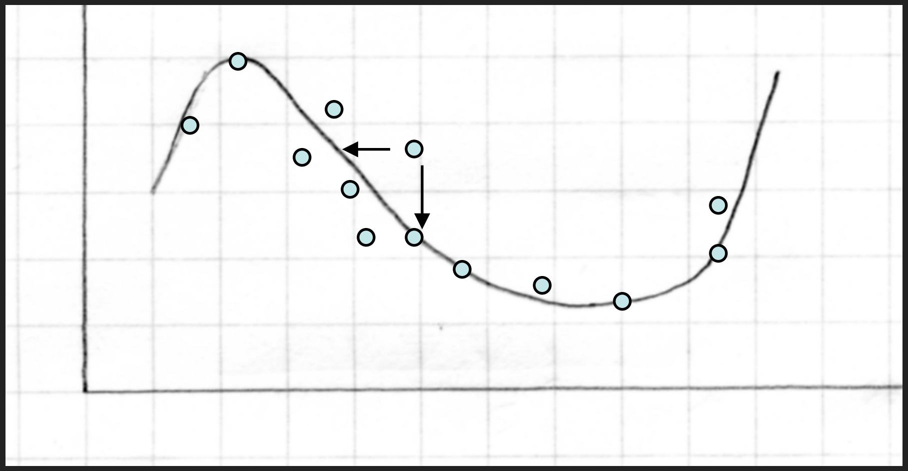
- Consider quantity x observed at t_i: $x_i(t_i)$ with i = 1,…,N
- Phase $\varphi(t_i)$ for cyclic frequency f or period P = 1/f with respect to the reference epoch t_0:

$$\varphi(t_i) = [f(t_i - t_0)] = [t_i - t_0 / P]$$

- Best period
  - minimizes the residuals of the intensity variations ↓
    e.g. through Fourier analysis methods & least-squares fits
  - minimizes the residuals in phase ←
    e.g. through Phase Dispersion Minimization (PDM)

Phase Diagram

Intensity



Phase = Time / Period

FOURIER ANALYSIS

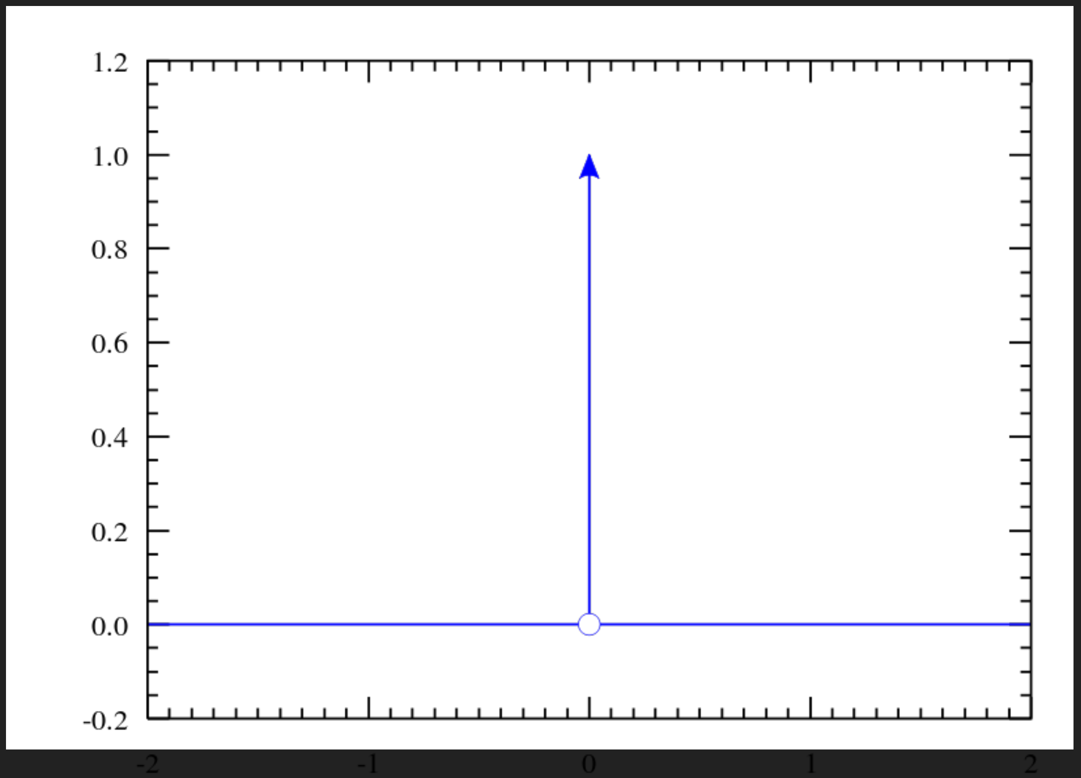
- **Concept:** We fit a series of sine curves with varying frequencies, amplitudes and phases to our data.
- Fourier transform of x(t): $$F(f) \equiv \int_{-\infty}^{+\infty} x(t) \exp(2\pi \mathrm{i}\, f\, t) dt$$
- As in reality we have **measurements at discrete points in time,** we need to replace the integral with a sum → **discrete Fourier Analysis**
- We transform the time series into the **frequency domain** and instead of the time series we investigate their **amplitude spectra** (or power spectra).

FOURIER ANALYSIS

$$x(t) = \sum_{k=1}^{n} A_k \exp(2\pi \mathrm{i}\, f_k t) : \quad F(f) = \sum_{k=1}^{n} A_k \delta(f - f_k)$$

- Fourier transform F(f) of sum of harmonic functions with frequencies $f_1,\ldots,f_n$ and amplitudes $A_1,\ldots,A_n$
- For x(t) = sine with frequency f_1, F(f) ≠ 0 for f = ± f_1
- For x(t) = sum of n harmonic functions with frequencies $f_1,\ldots,f_n$ F(f) = sum of δ-functions ≠ 0 for ± $f_1,\ldots,f_n$
- Real data set: x(t) known for the discrete number of times t_j with j = 1,…, N

DELTA FUNCTION



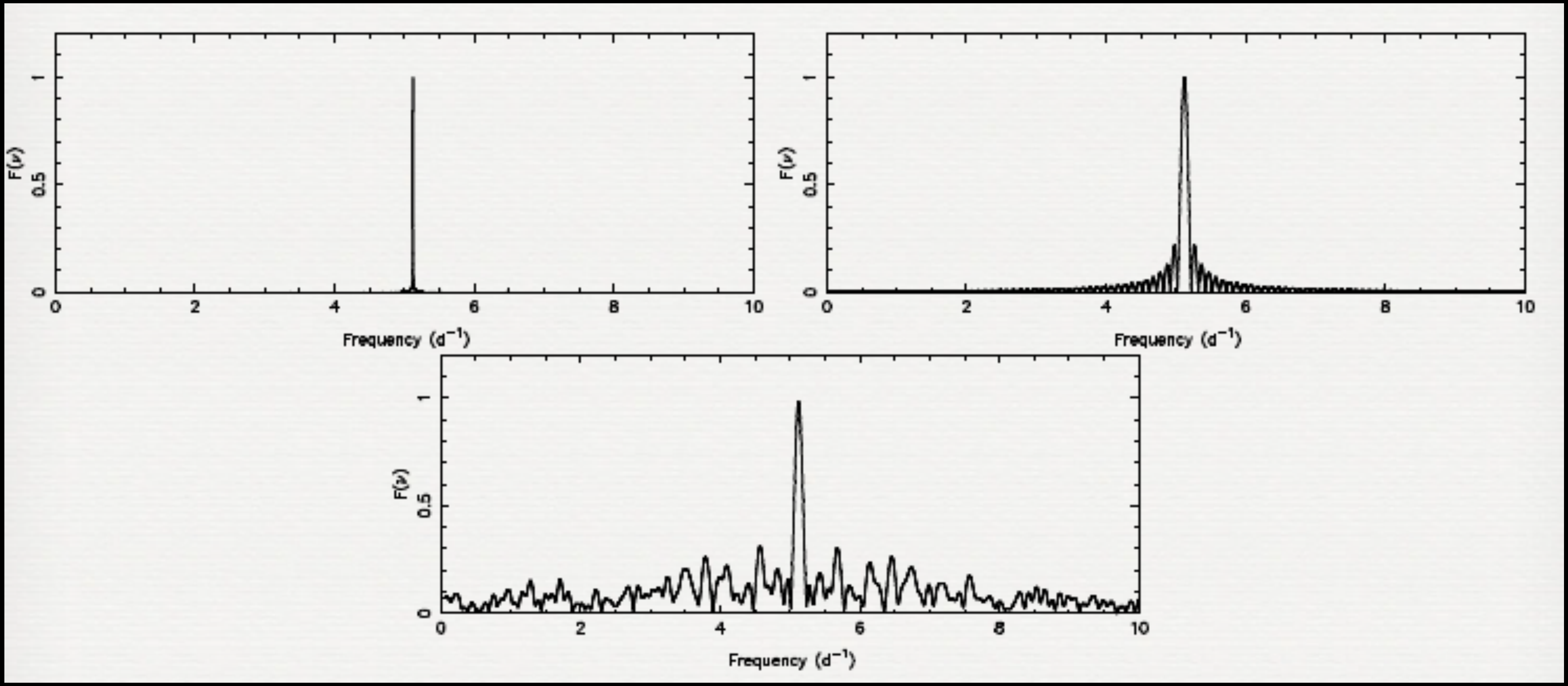
- generalized function or distribution that is 0 everywhere except at 0
- integral over the entire real line is 1
- sometimes pictured as an infinitely high, infinitely thin spike at the origin with total area 1 under the spike
- physically represents the density of an idealized point mass or point charge
- introduced by theoretical physicist Paul Dirac

$$\int_{-\infty}^{+\infty} \delta(\nu)\, d\nu = 1$$

EXAMPLES OF FOURIER TRANSFORMS

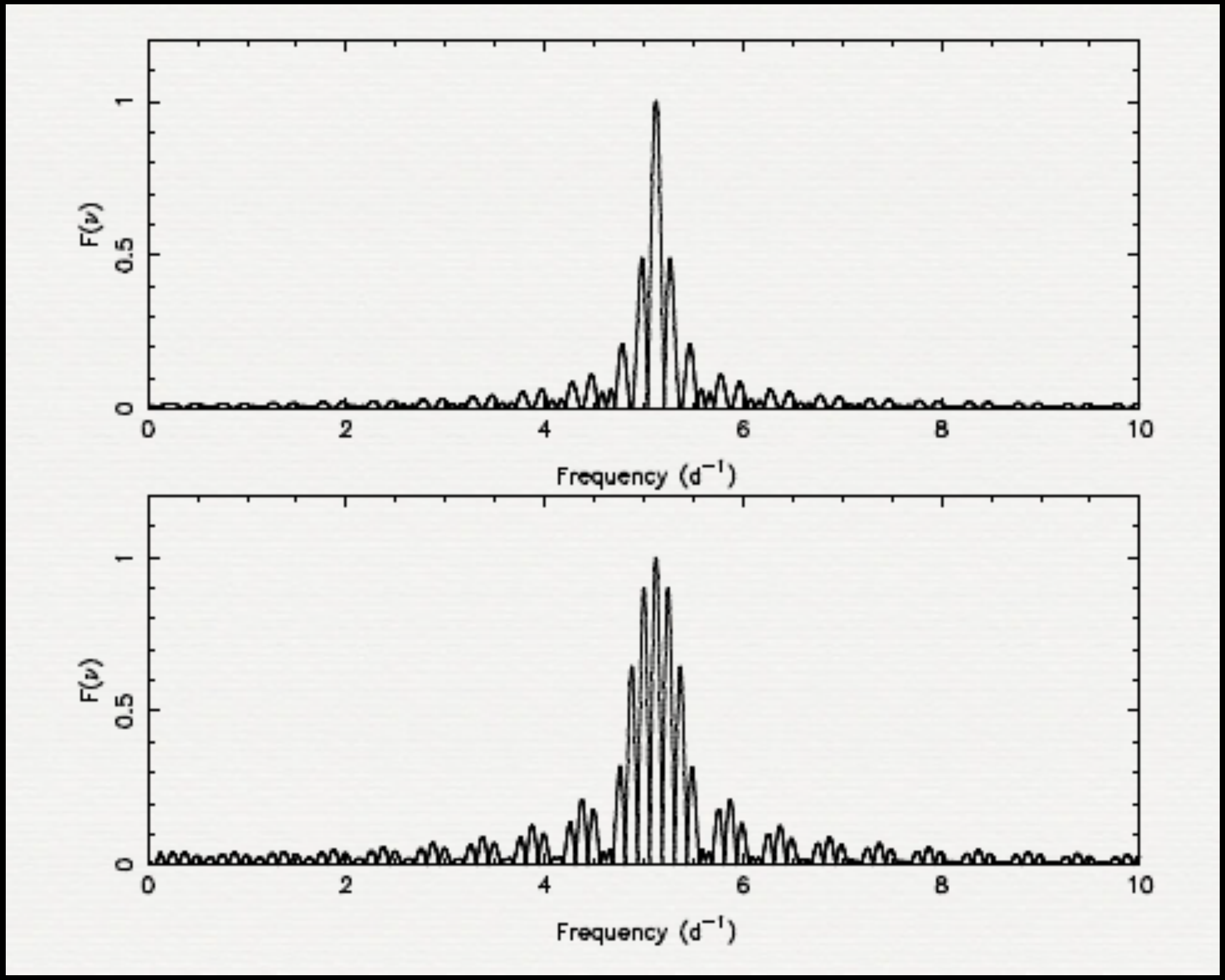
10^6 points over 1000d

10^4 points over 10d



4472 points over 10d

EXAMPLES OF FOURIER TRANSFORMS



Fourier transforms of a noiseless time series of a sine function with frequency 5.123456789 d^{-1} generated for a finite time span of 10 days and containing one large gap from day 4 until day 6 (top) and from day 2 until day 8 (bottom)

WINDOW FUNCTION

- Discrete Fourier transform

$$F_N(f) \equiv \sum_{j=1}^{N} x(t_j)\exp(2\pi\mathrm{i}\, f t_j)$$

- $F_N \neq F$! but connected through the window function

$$w_N(t) \equiv \frac{1}{N}\sum_{j=1}^{N}\delta(t - t_j)$$

- Hence

$$\frac{F_N}{N} = \int_{-\infty}^{+\infty} x(t) w_N(t)\exp(2\pi\mathrm{i}\, f t)dt$$

FOURIER TRANSFORM & WINDOW FUNCTION

- Discrete Fourier transform of window function = = spectral window $W_N(f)$

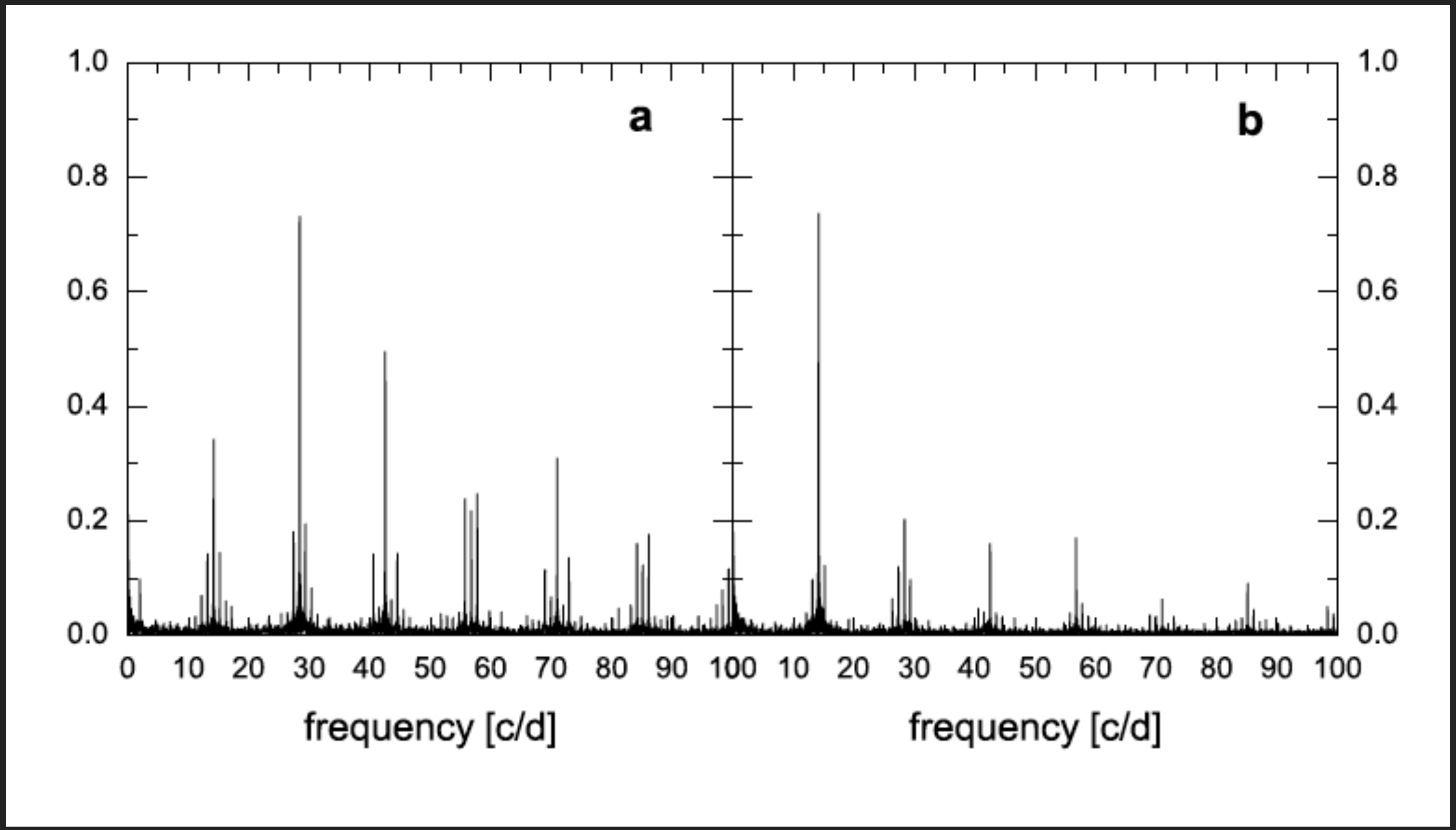
$$W_N(f) = \frac{1}{N}\sum_{j=1}^{N} \exp(2\pi \mathrm{i}\, f t_j)$$

- Discrete Fourier transform = convolution of spectral window and Fourier transform:

$$F_N(f)/N = F(f) * W_N(f)$$

SPECTRAL WINDOW

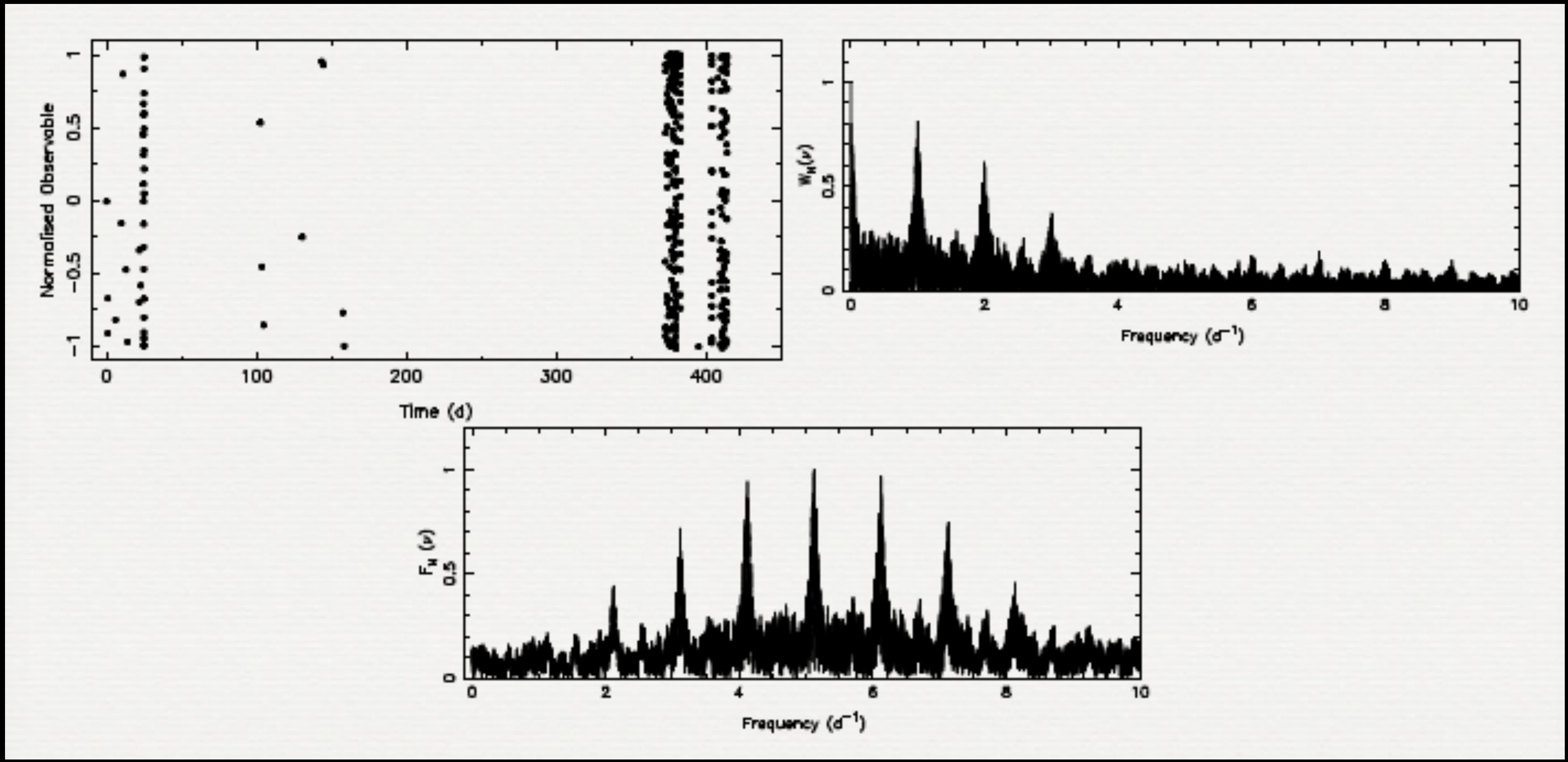
- characteristics of your data
  - instrumental regularities (e.g., day-night cycle, satellite orbit)
  - data set structure (e.g., regular gaps)



ALIAS FREQUENCIES

- Maxima in periodograms at spurious frequencies due to observing times
- Most common alias frequencies for non-equidistant data:
  - $\Delta t \sim 1$ day, 1 year, specific gaps, length of nightly observations, etc.

EXAMPLE



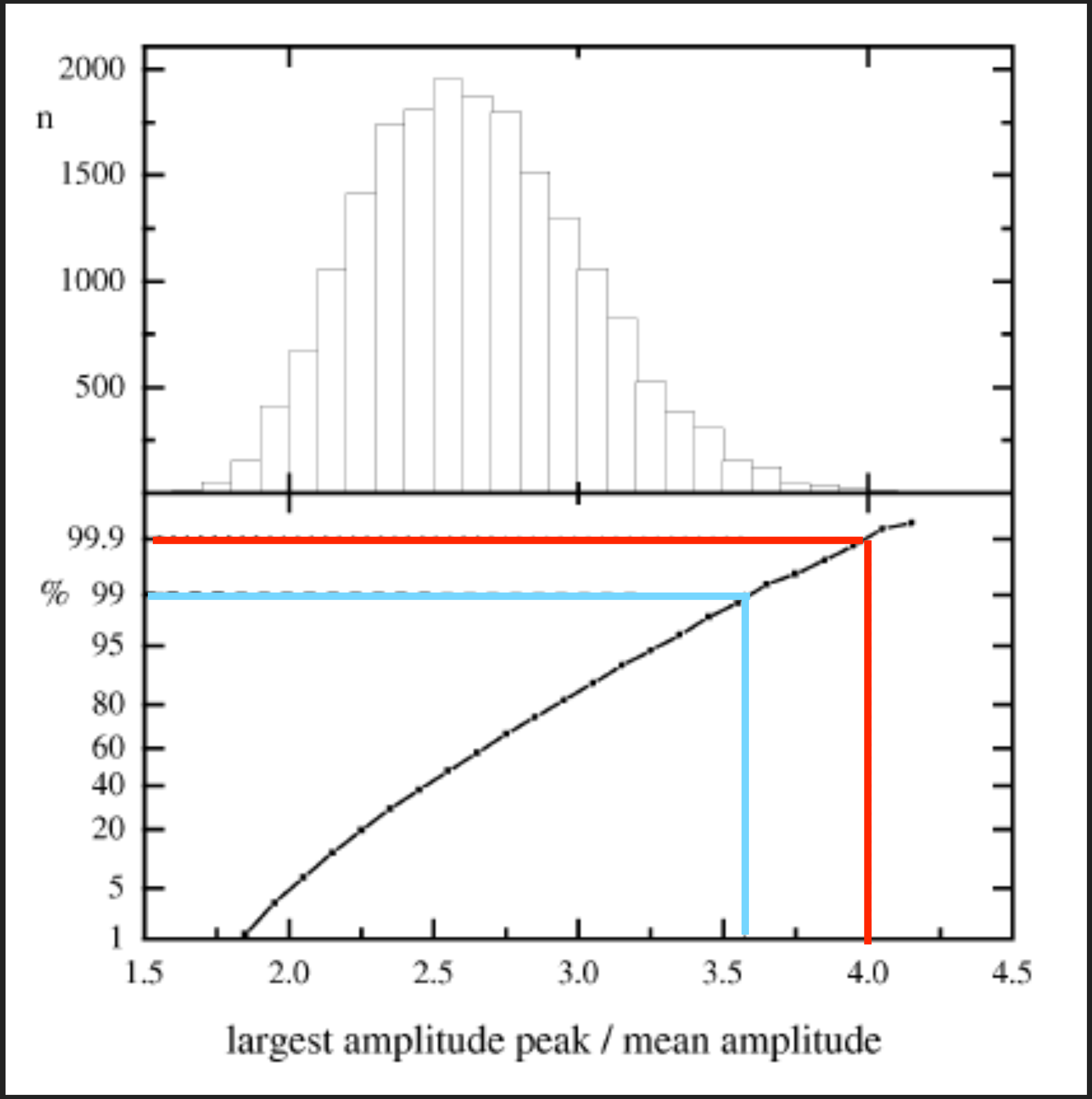
Spectral window (upper right) and DFT (bottom) of a noise-free sinusoid with amplitude 1 at 5.123456789 d^{-1} for the sampling shown in the upper left

ACCEPTANCE CRITERION

- $A > 4\ \sigma_{Fourier}$
  means: 99.9% confidence interval that peak is not caused by noise

  A … amplitude
  $\sigma_{Fourier}$ … noise computed in a test interval around the candidate frequency

  $A > 3.6\ \sigma_{Fourier}$
  99.0% confidence interval that peak is not caused by noise



Kuschnig et al. (1997), Breger et al. (1993)

LEAST SQUARES FIT

- Trying to fit the data with this formula

$$y(t) = a_o + \sum_{i=1}^{i=n} a_i \sin(2\Pi f_i t + \phi_i)$$

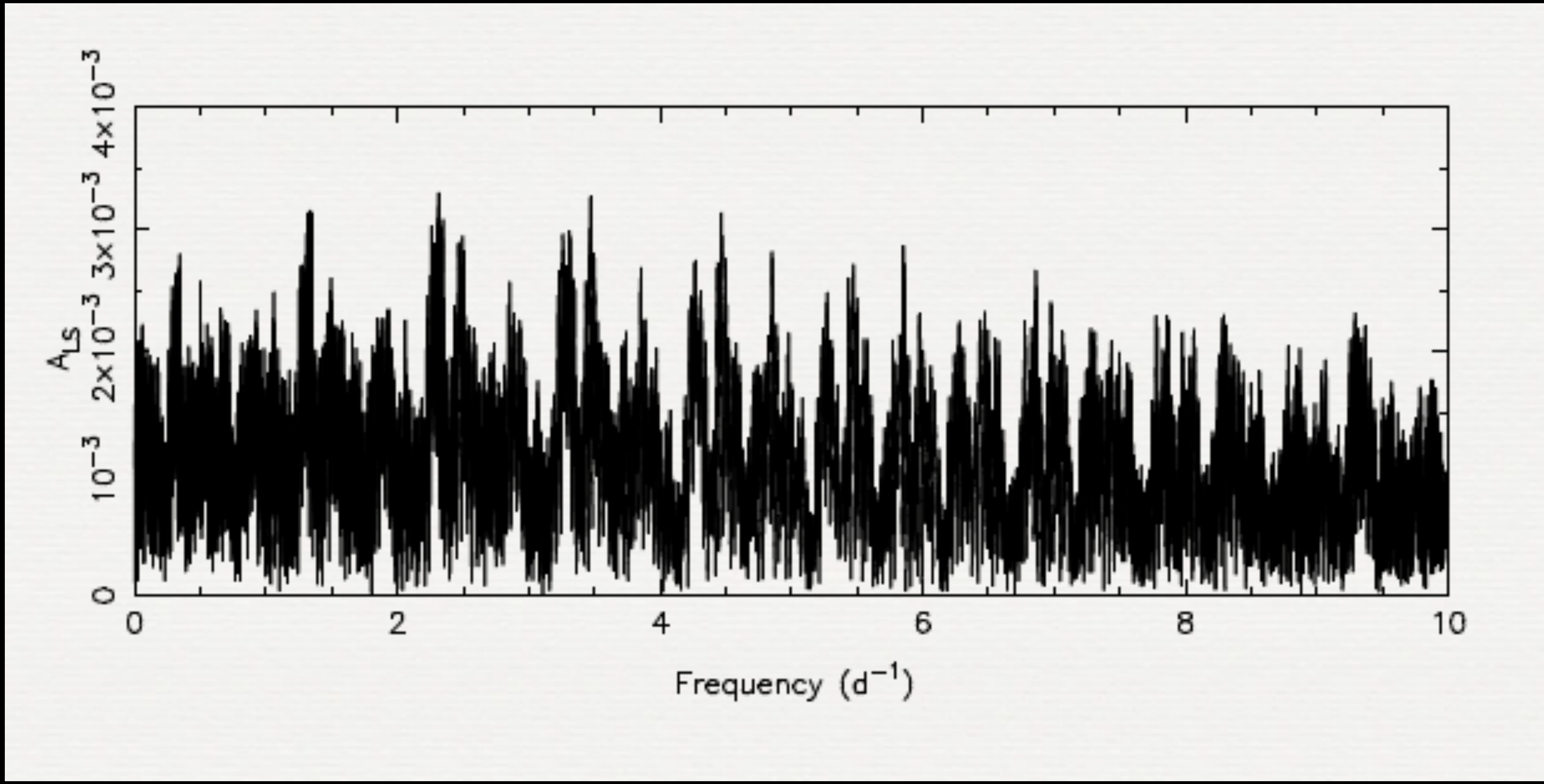
- To find the periodic content in the data & minimize residuals between fit and measurements

with a_0 … zero point (~average magnitude)
a_i … individual amplitudes of the modes (~height of the Fourier spectrum)
f_i … frequency of the i-th mode
Φ_i … phase of the i-th mode

GETTING THE SOLUTION

- Now we fix f: $$x_i(t_i) = A\sin\left[2\pi(ft_i + \psi)\right] + C$$
- Variance reduction $$1 - \frac{\sum_{i=1}^{N}\left\{x_i - \left[A\sin\left(2\pi(ft_i+\psi)\right) + C\right]\right\}^2}{\sum_{i=1}^{N}\left(x_i - \overline{x}\right)^2}$$
- Search for new frequencies in Residuals $R_i(f) \equiv x_i - x_i^c(f)$ with $$x_i^c(f) \equiv A\sin\left[2\pi(ft_i + \psi)\right] + C$$
- and so on until the amplitude $A < 4\sigma_{\mathrm{Fourier}}$

UNTIL YOU REACH THE NOISE LEVEL



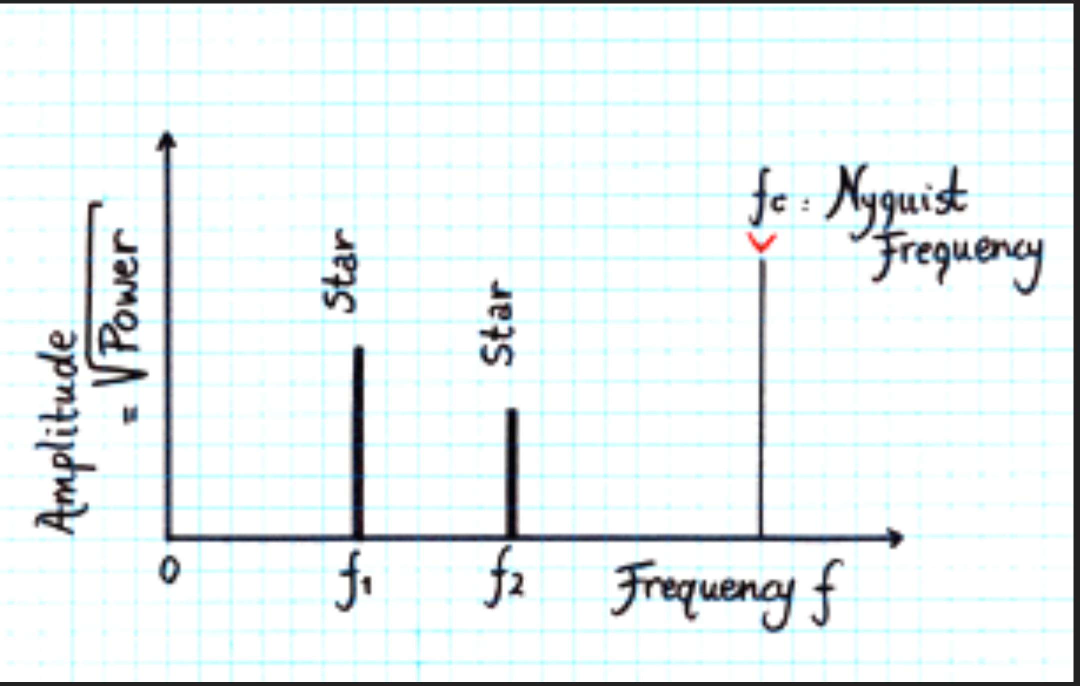
NYQUIST FREQUENCY

- **Nyquist frequency $f_{Nyquist}$ = 1 / 2Δt = N-1 / 2ΔT**

  N … number of measurements that are equally spread over ΔT

  Δt … time separation between two data points (= "time step")
  Δt = ΔT / N-1

Sampling Theorem: *"A signal can be reconstructed from its samples if the original signal has no frequencies above 1/2 the sampling frequency"*



NYQUIST FREQUENCY & ALIASES

- Sampling Theorem means that all frequencies from
  $$-f_{Nyquist} < f < f_{Nyquist}$$
  can be determined
- BUT: real frequencies $f > f_{Nyquist}$ lead to alias frequencies between $-f_{Nyquist}$ and $+f_{Nyquist}$
  e.g., $2 * f_{Nyquist} - f$, $2 * f_{Nyquist} + f$, $3 * f_{Nyquist} - f$ etc.
- Example: roAp star with f = 220 c/d (P = 6.5min)
  $\Delta t = 6$ min $= 1/240$ d $\rightarrow f_{Nyquist} = 1*240 / 2 = 120\ d^{-1}$
  alias frequency at $f_{Nyquist} - f = (240 - 220)\ d^{-1} = 20\ d^{-1}$

ERROR ESTIMATES

- Error for derived frequency assuming white noise and no alias problems

$$\sigma_f = \frac{\sqrt{6}\sigma_R}{\pi\sqrt{N}AT}$$

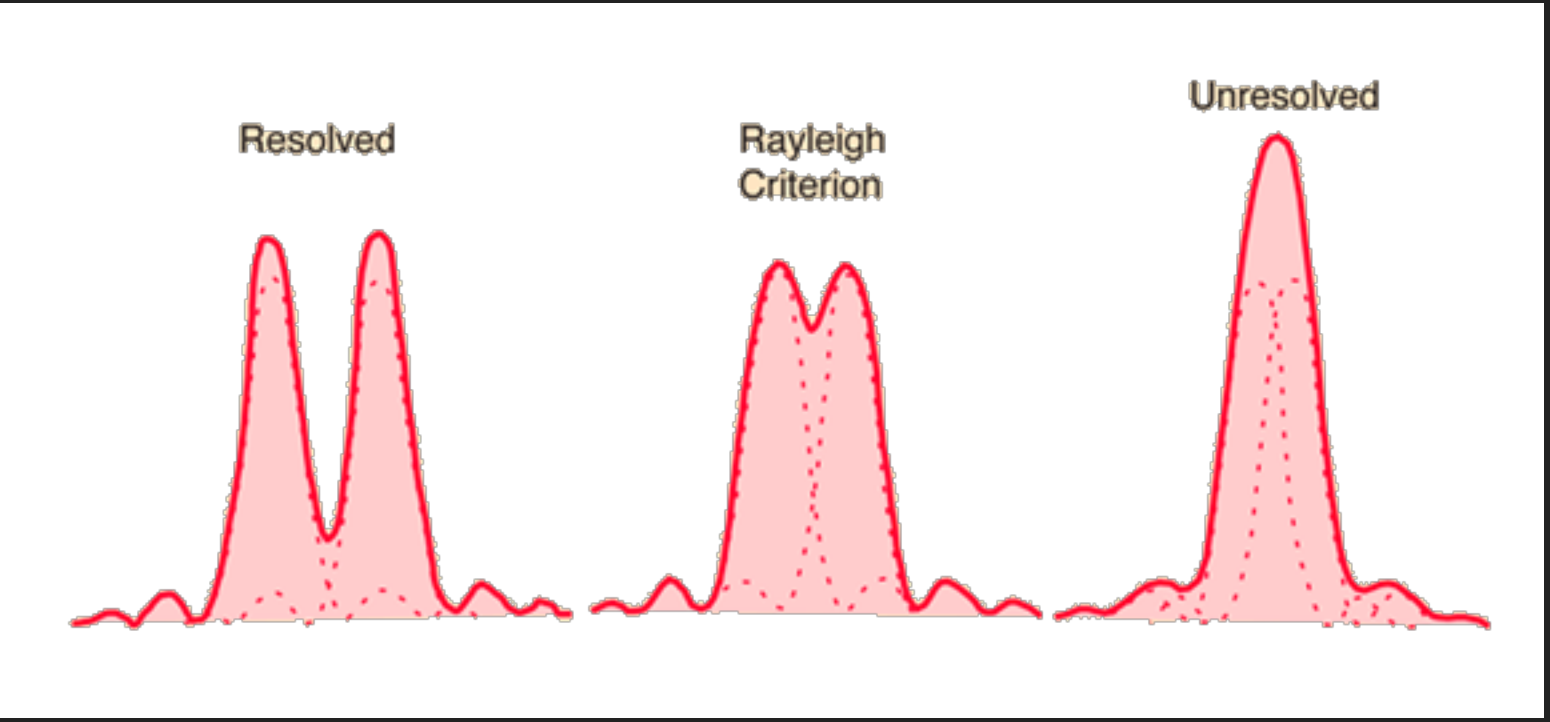
- with σ_R … average error of the data (~standard deviation of noise)
  T … total time span
  N … number of data points

FREQUENCY RESOLUTION

- Resolution of two close peaks, if they are separated by

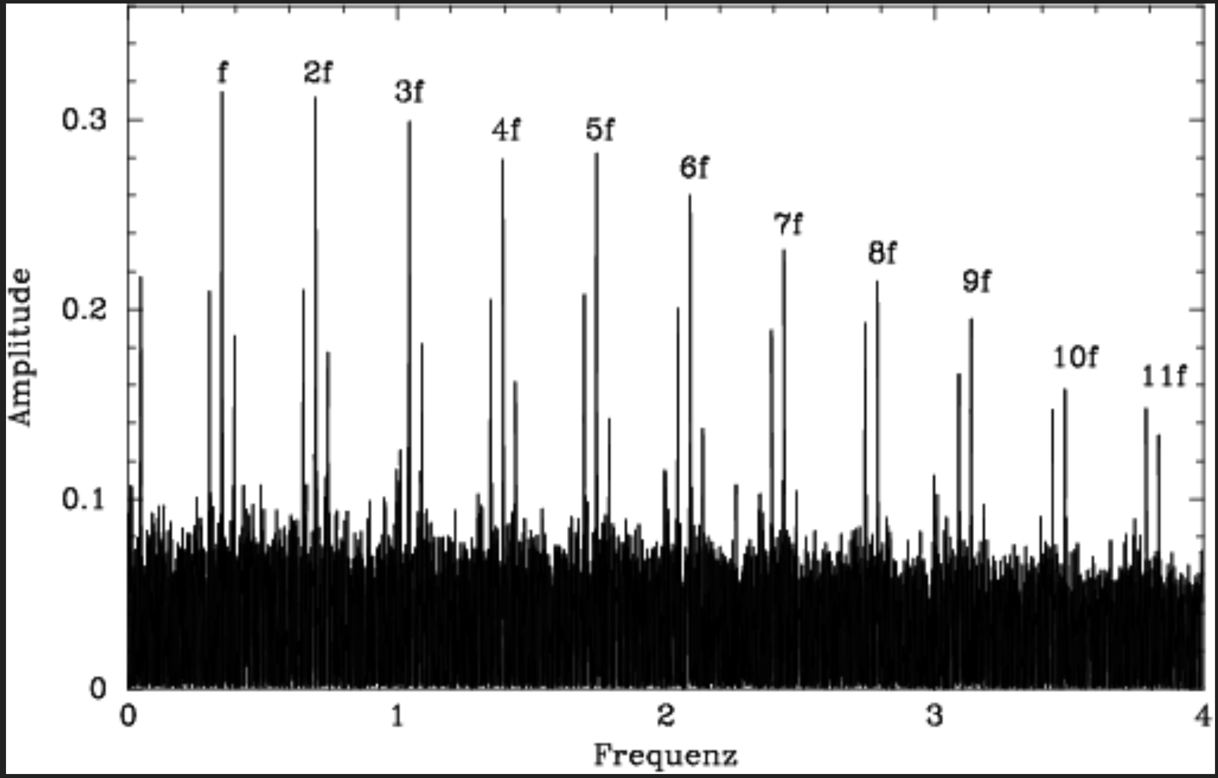
$$f_{res} = 1 / T$$

"Rayleigh criterion"



LIMITATIONS OF FOURIER ANALYSIS

- Advantage: quick, easy, commonly used
- Disadvantages:
  - asymmetric light curves like for RR Lyrae stars
    - formation of “harmonics”, superposition of sine waves needed to reproduce the light curve
  - problematic if periods change with time



OTHER METHOD

- Phase Dispersion Minimization (PDM)
  Stellingwerf 1978, ApJ 224, 953
  - data are compared to a test period and phase diagrams are generated
  - the sum of the data set's variances in certain phase intervals (e.g., length of 0.05*period) is evaluated in comparison to the test period → reaches a minimum at correct period

FOURIER ANALYSIS WITH PERIOD04

- Period04 for download under:
  http://www.lenzpat.at/period04
- You will get your own data sets on the webpage
  http://astro-staff.uibk.ac.at/~zwintz/teaching.html
- **Your task**: In the two weeks in January with no lectures (Wed 11 & 18th January 2017), use Period04 to analyze the light curves, find frequencies, describe the frequency content and try to identify the type of pulsator.
- Please bring your results to the exam on Feb 1, 2017!

INTERPRETING THE FREQUENCIES

PULSATION CONSTANT

- For δ Scuti stars: Pulsation constant, Q
  - to distinguish if the observed period is a radial fundamental or higher overtone mode

$$Q = P\,(\rho/\rho_{Sun})^{1/2}$$

Breger (1979)

$$Q = -6.454 + \log P + 0.5 \log g + 0.1 M_{bol} + \log T_{eff}$$

  - fundamental mode: Q = 0.033 d
    first overtone mode: Q = 0.025 d
    second harmonic: Q = 0.021 d
    third harmonic: Q = 0.017d

higher radial overtone
= lower Q value

ECHELLE DIAGRAM

- First introduced by Grec (1981): based on the fact that the low-degree (p-)modes are essentially equidistant in frequency for a given l.
- The spectrum is cut into pieces of this frequency spacing which are stacked on top of each other. Since the modes are not truly equidistant in frequency, the echelle diagram shows up power as distorted ridges.

PLOTTING ECHELLE DIAGRAMS

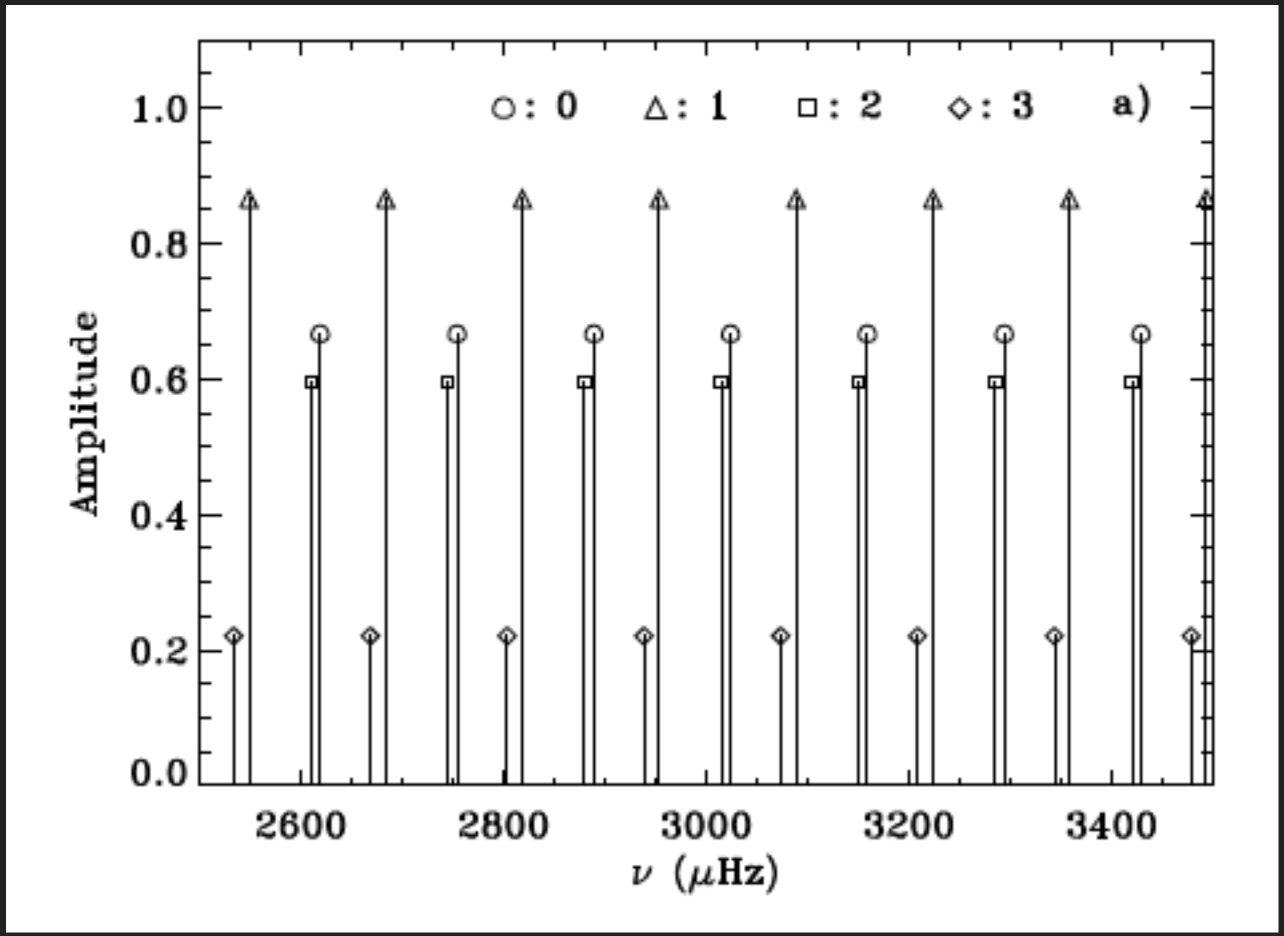
- The frequencies are reduced modulo ν_0 by expressing them as

$$\nu_{nl} = \nu_0 + k\Delta\nu_0 + \tilde{\nu}_{nl} \ ,$$

  where ν_0 is a suitably chosen reference, and k is an integer such that $\tilde{\nu}_{nl}$ is between 0 and $\Delta\nu_0$

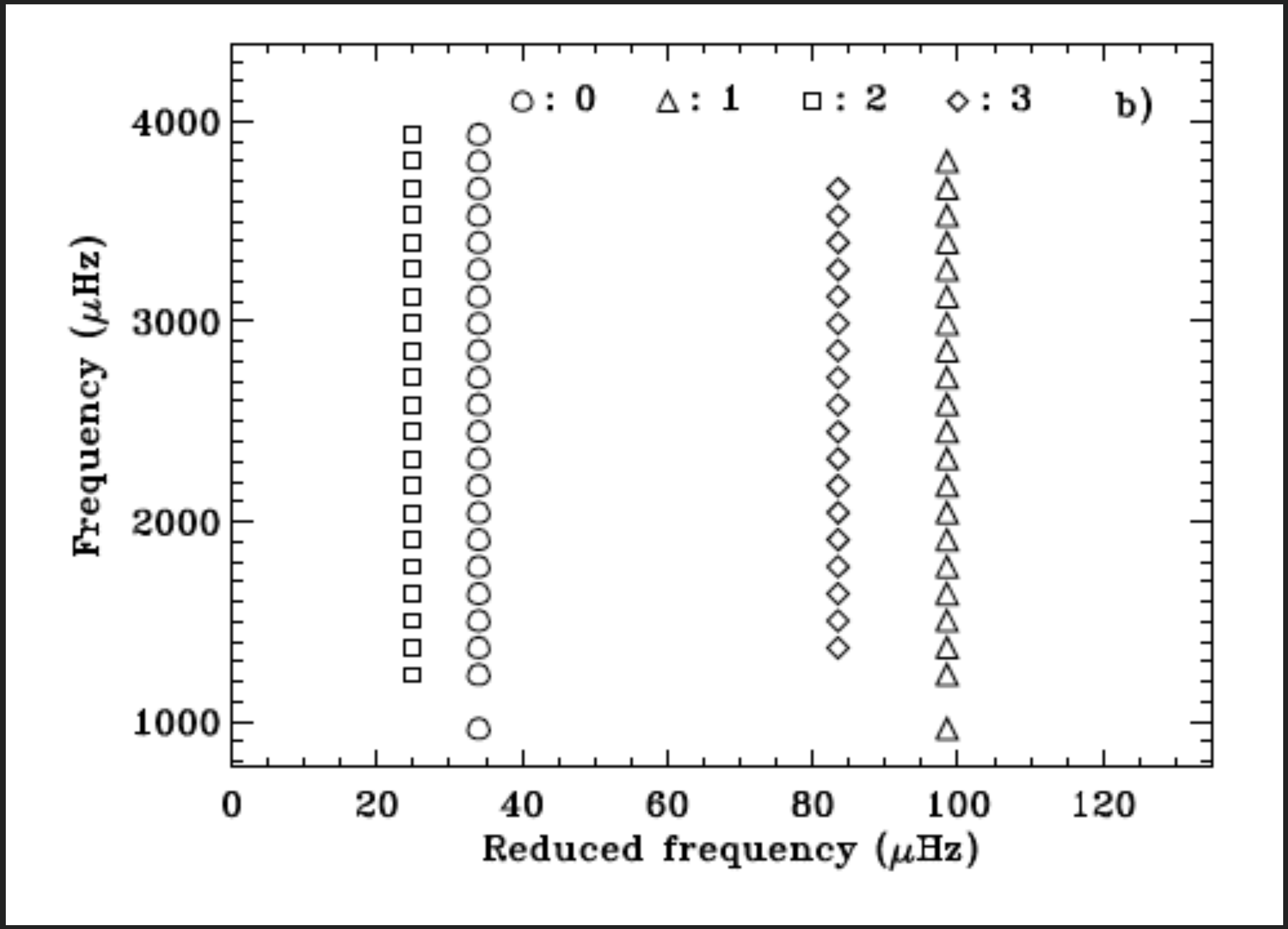
- The echelle diagram is produced by plotting $\tilde{\nu}_{nl}$ on the abscissa and $\nu_0 + k\Delta\nu_0$ on the ordinate.

ECHELLE DIAGRAM PLOTS



Schematic solar frequency spectrum

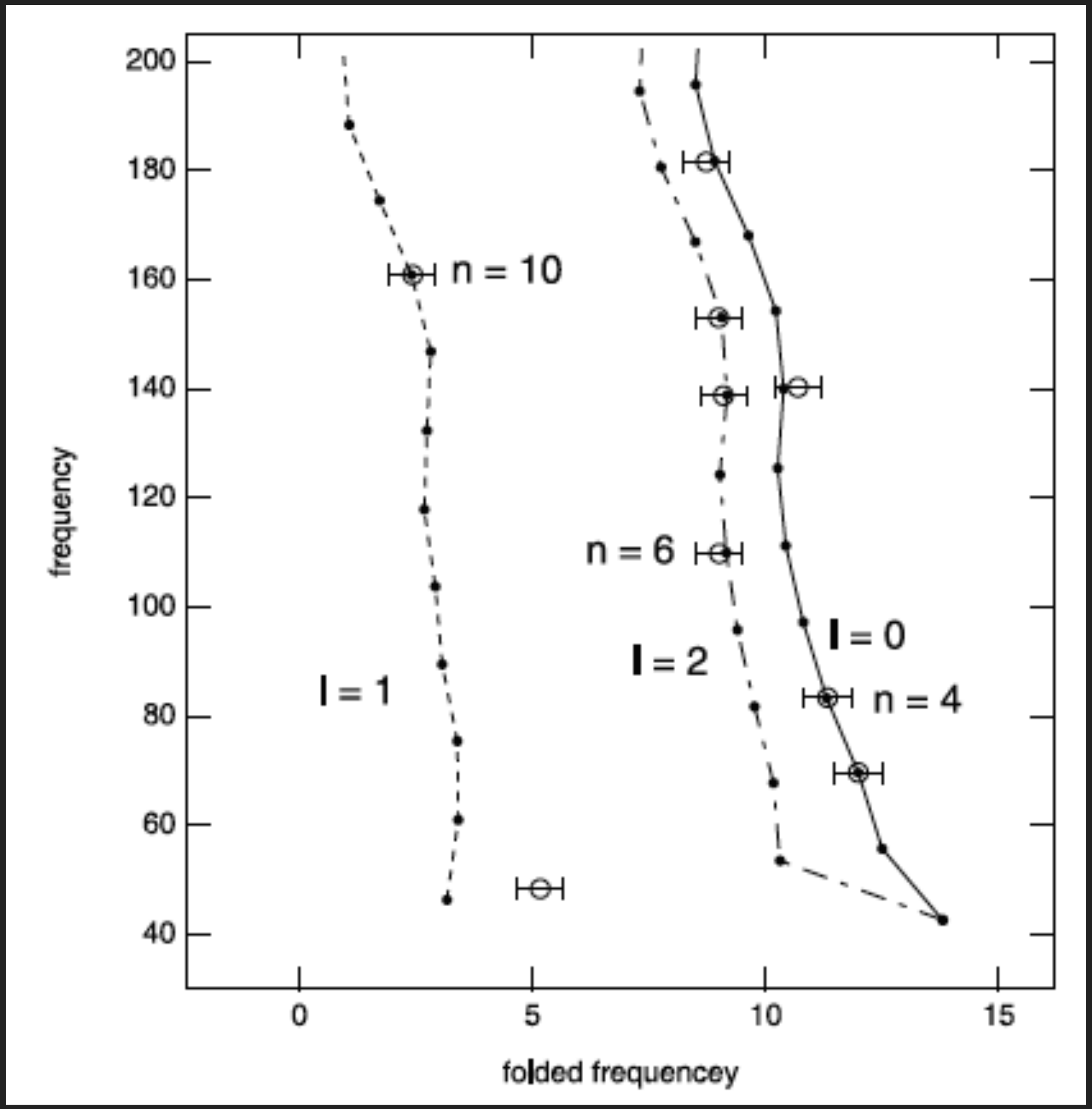
$\Delta\nu_0$ = 135μHz

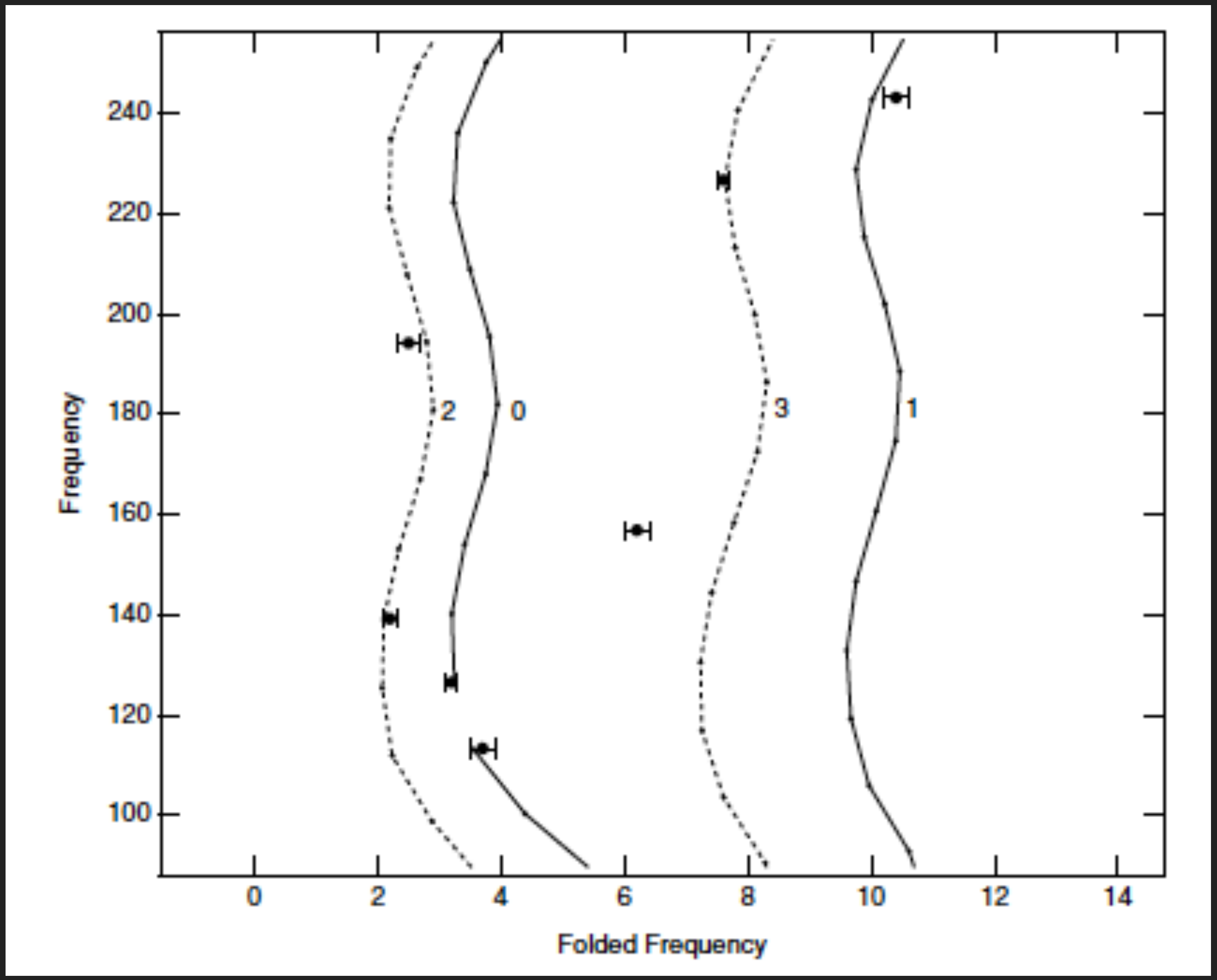


- Graphically: imagine you cut the frequency axis in pieces of $\Delta\nu_0$ and stack them on top of each other

OTHER ECHELLE DIAGRAMS

Pre-main sequence δ Scuti stars





$\Delta\nu_0 = 14.4\ \mu\mathrm{Hz}$

$\Delta\nu_0 = 13.7\ \mu\mathrm{Hz}$

NOT ALWAYS THAT CLEAR…

other effects influence
the pulsations
e.g., rotation,
accretion (in
young stars), …