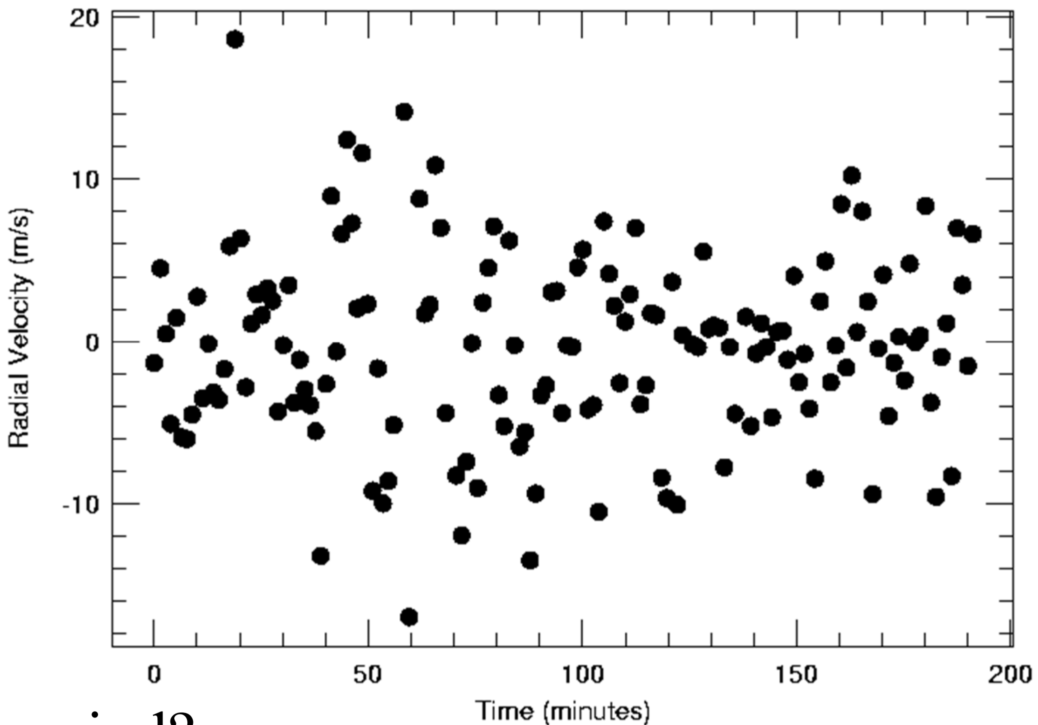
Searching for Periodic Signals in Time Series Data

1. Least Squares Sine Fitting
2. Discrete Fourier Transform
3. Lomb-Scargle Periodogram
4. Pre-whitening of Data
5. Other techniques
   - Phase Dispersion Minimization
   - String Length
   - Wavelets

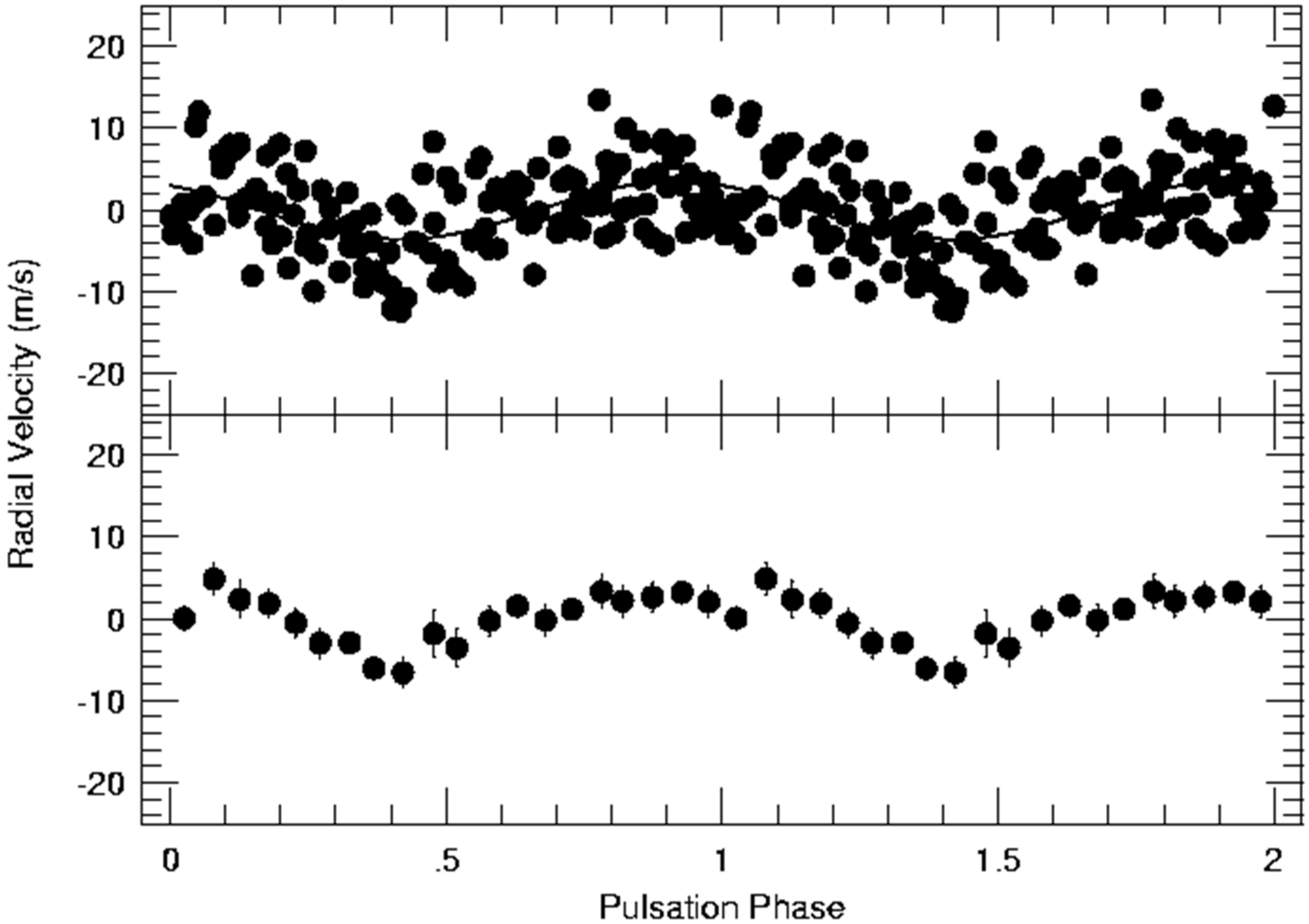
Period Analysis

How do you know if you have a periodic signal in your data?



What is the period?

Try 16.3 minutes:



1. Least-squares Sine fitting

Fit a sine wave of the form:

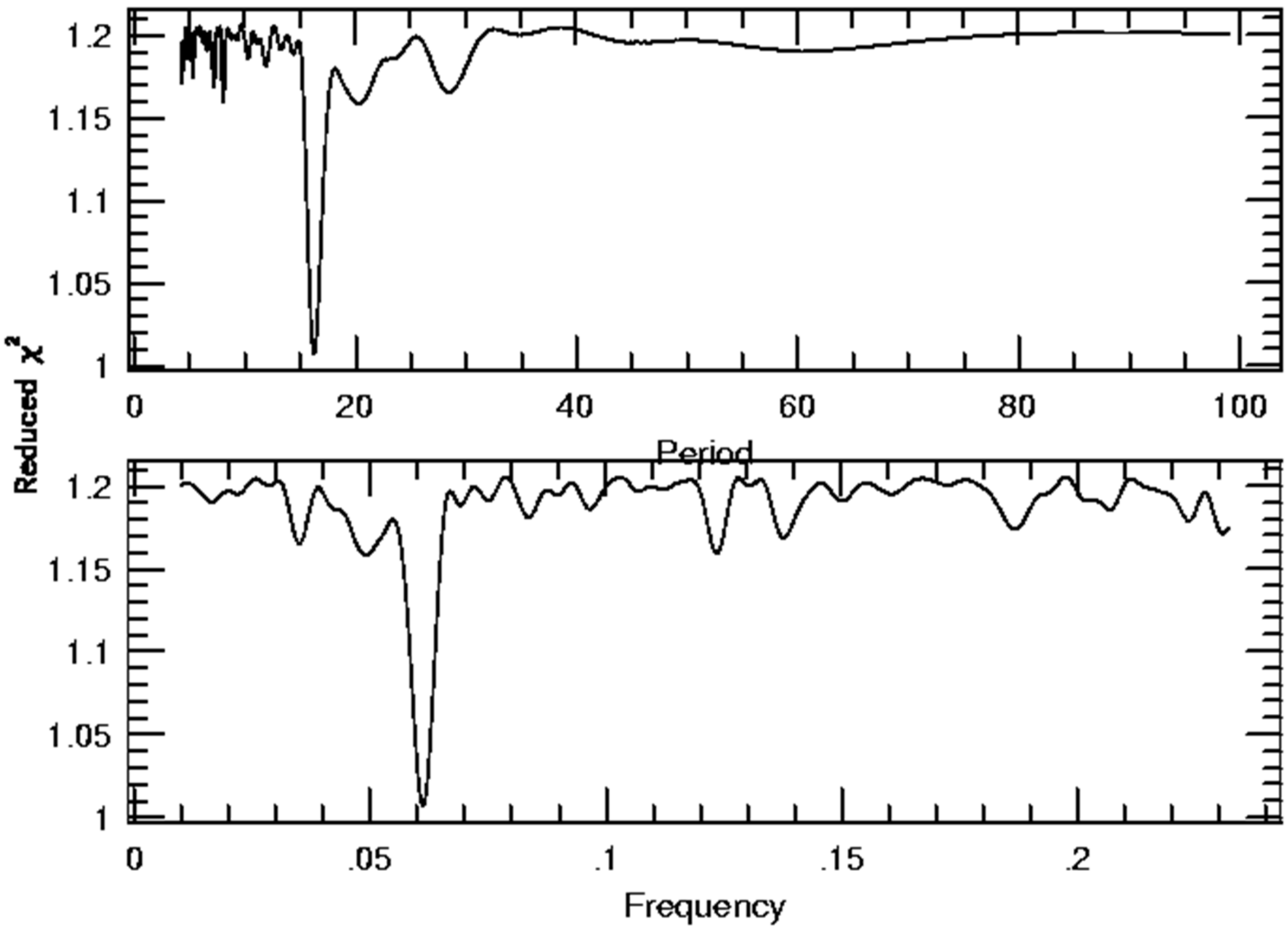
$V(t) = A\cdot\sin(\omega t + \phi) + \text{Constant}$

Where $\omega = 2\pi/P$, ϕ = phase shift

Best fit minimizes the χ^2:

$\chi^2 = \Sigma\, (d_i - g_i)^2/N$

d_i = data, g_i = fit



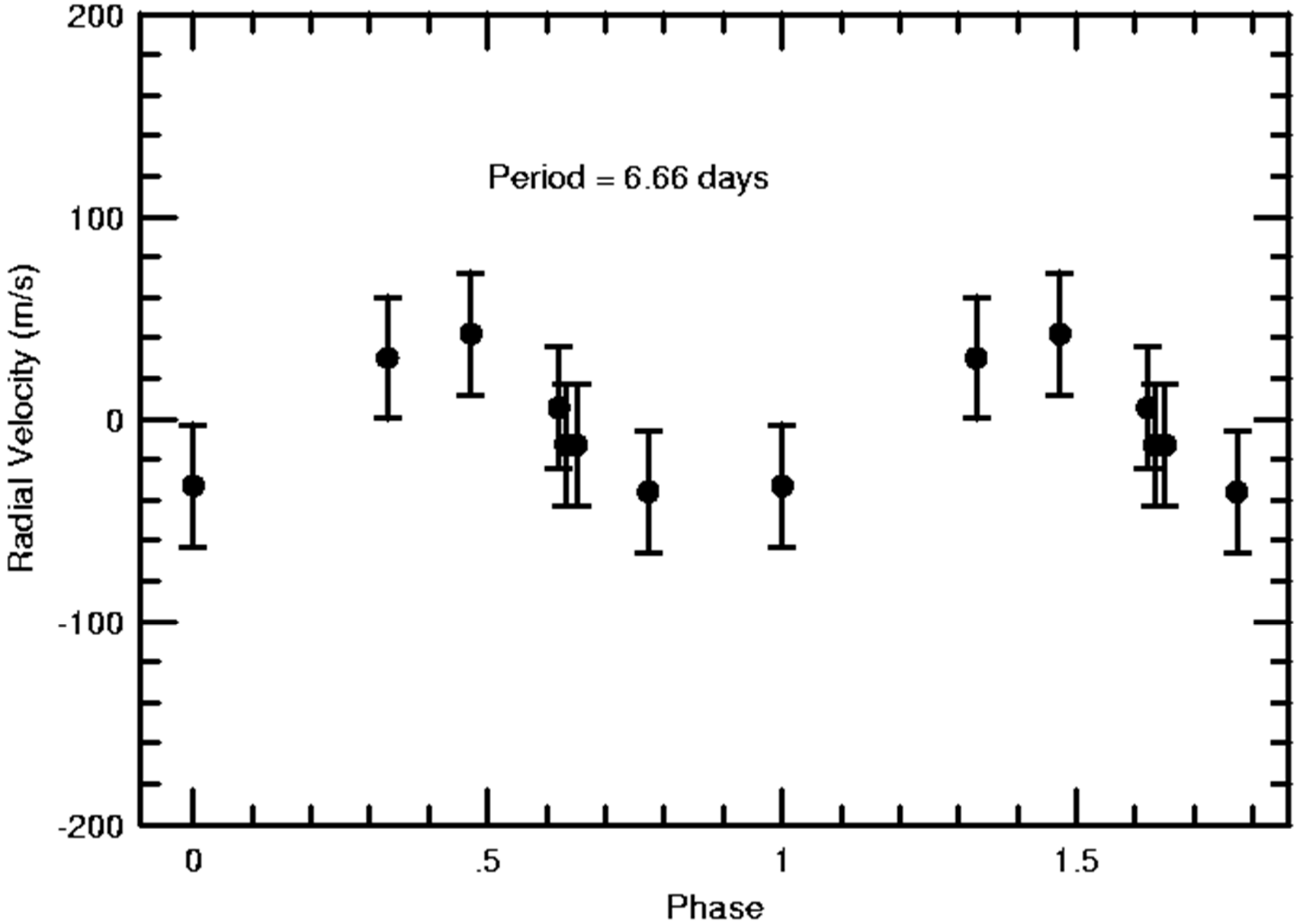
1.2
1.15
1.1
1.05
1
0
20
40
60
80
100
Period
Reduced χ²
1.2
1.15
1.1
1.05
1
0
.05
.1
.15
.2
Frequency

Advantages of Least Sqaures sine fitting:

- Good for finding periods in relatively sparse data

Disadvantages of Least Sqaures sine fitting:

- Signal may not always be a sine wave (e.g. eccentric orbits)
- No assessement of false alarm probability (more later)
- Don‘t always trust your results



This is fake data of pure random noise with a σ = 30 m/s. Lesson: poorly sampled noise almost always can give you a period, but it it not signficant

2. The Discrete Fourier Transform

Any function can be fit as a sum of sine and cosines

$$FT(\omega) = \sum_{j=1}^{N_0} X_j\,(t)\, e^{-i\omega t}$$

Recall $e^{i\omega t} = \cos \omega t + i \sin\omega t$

X(t) is the time series

Power: $$P_x(\omega) = \frac{1}{N_0} \,|\, FT_X(\omega)|^2$$

N_0 = number of points

$$P_x(\omega) = \frac{1}{N_0}\left[\left(\sum X_j \cos \omega t_j\right)^2 + \left(\sum X_j \sin \omega t_j\right)^2\right]$$

A DFT gives you as a function of frequency the amplitude (power) of each sine wave that is in the data

The continous form of the Fourier transform:

$$F(s) = \int f(x)\, e^{-ixs}\, dx$$

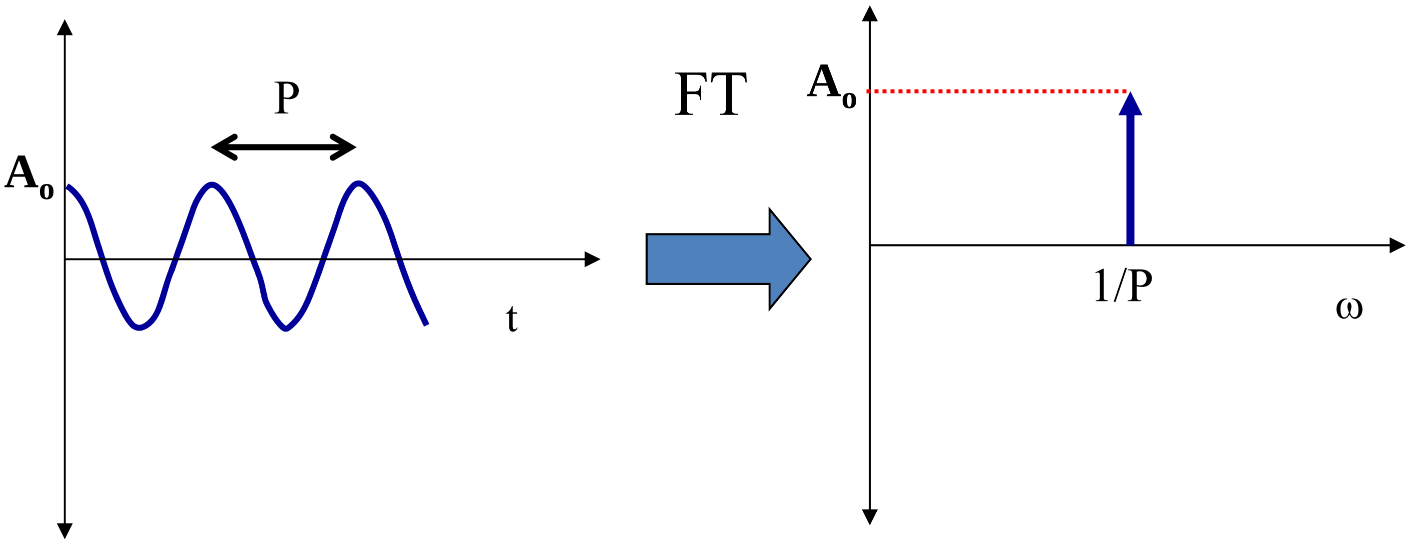
$$f(x) = 1/2\pi \int F(s)\, e^{ixs}\, ds$$

$$e^{ixs} = cos(xs) + i\, sin\, (xs)$$

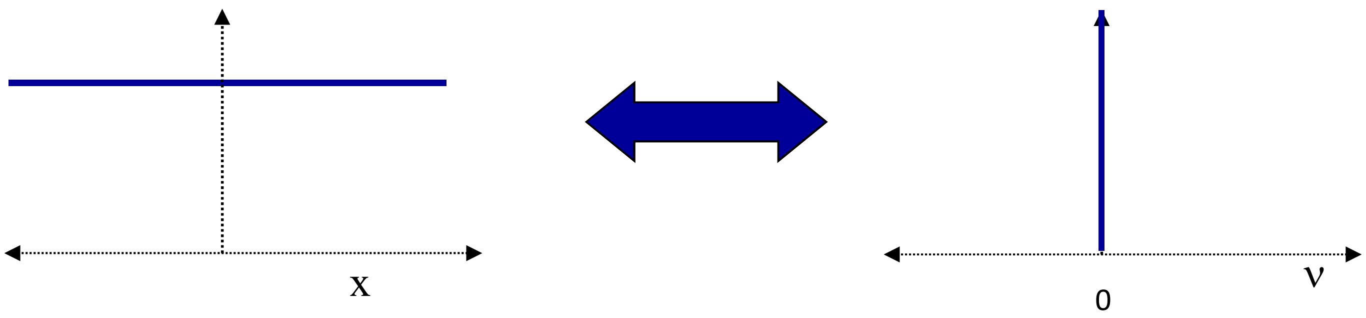
This is only done on paper, in the real world (computers) you always use a discrete Fourier transform (DFT)

The Fourier transform tells you the amplitude of sine (cosine) components to a data (time, pixel, x,y, etc) string

Goal: Find what structure (peaks) are real, and what are artifacts of sampling, or due to the presence of noise



A pure sine wave is a delta function in Fourier space

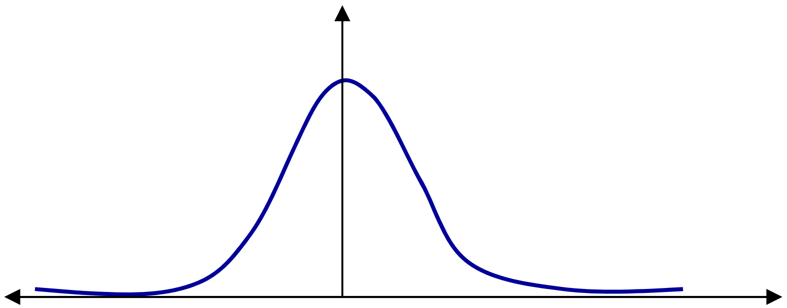


A constant value is a delta function with zero frequency in Fourier space:
Always subtract off „dc“ level.

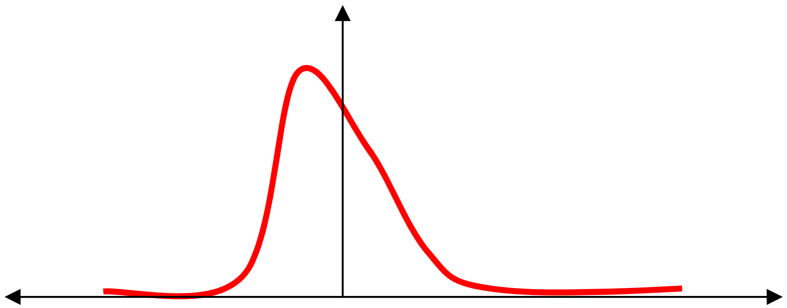
Useful concept: Convolution

$$\int f(u)\phi(x-u)du = f * \phi$$

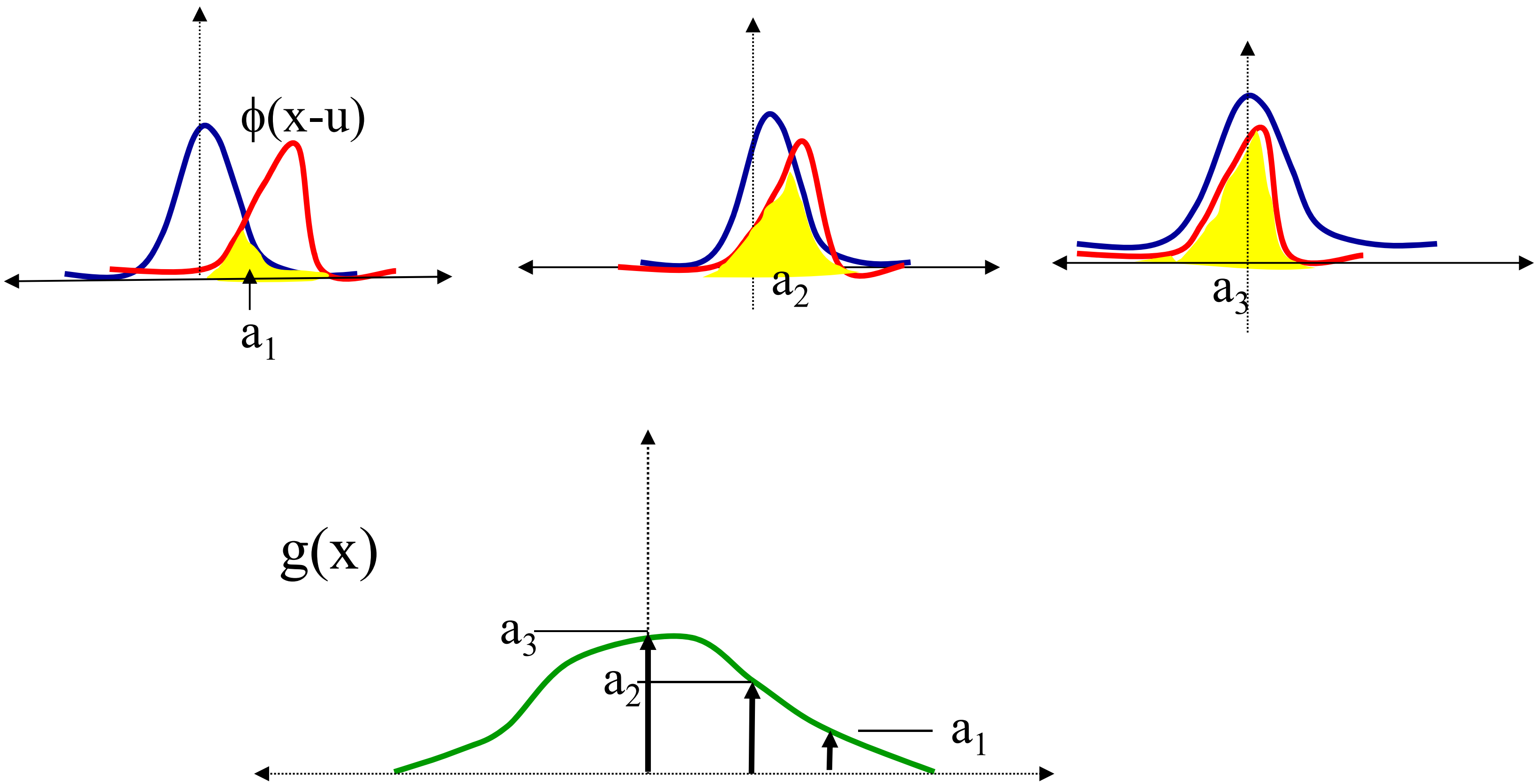
f(x):



$\phi(x)$:



Useful concept: Convolution



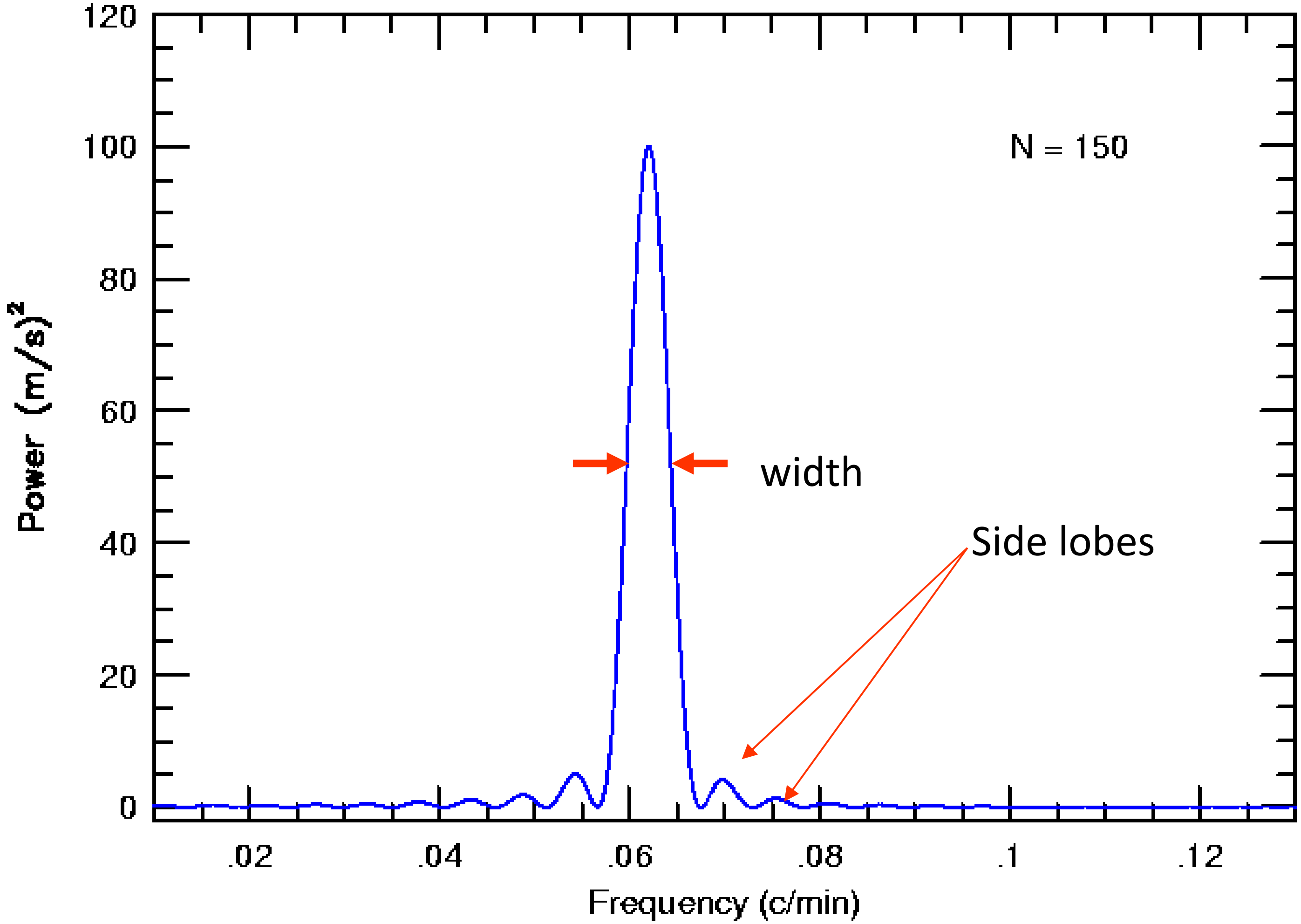
Convolution is a smoothing function

Convolution

In Fourier space the convolution is just the product of the two transforms:

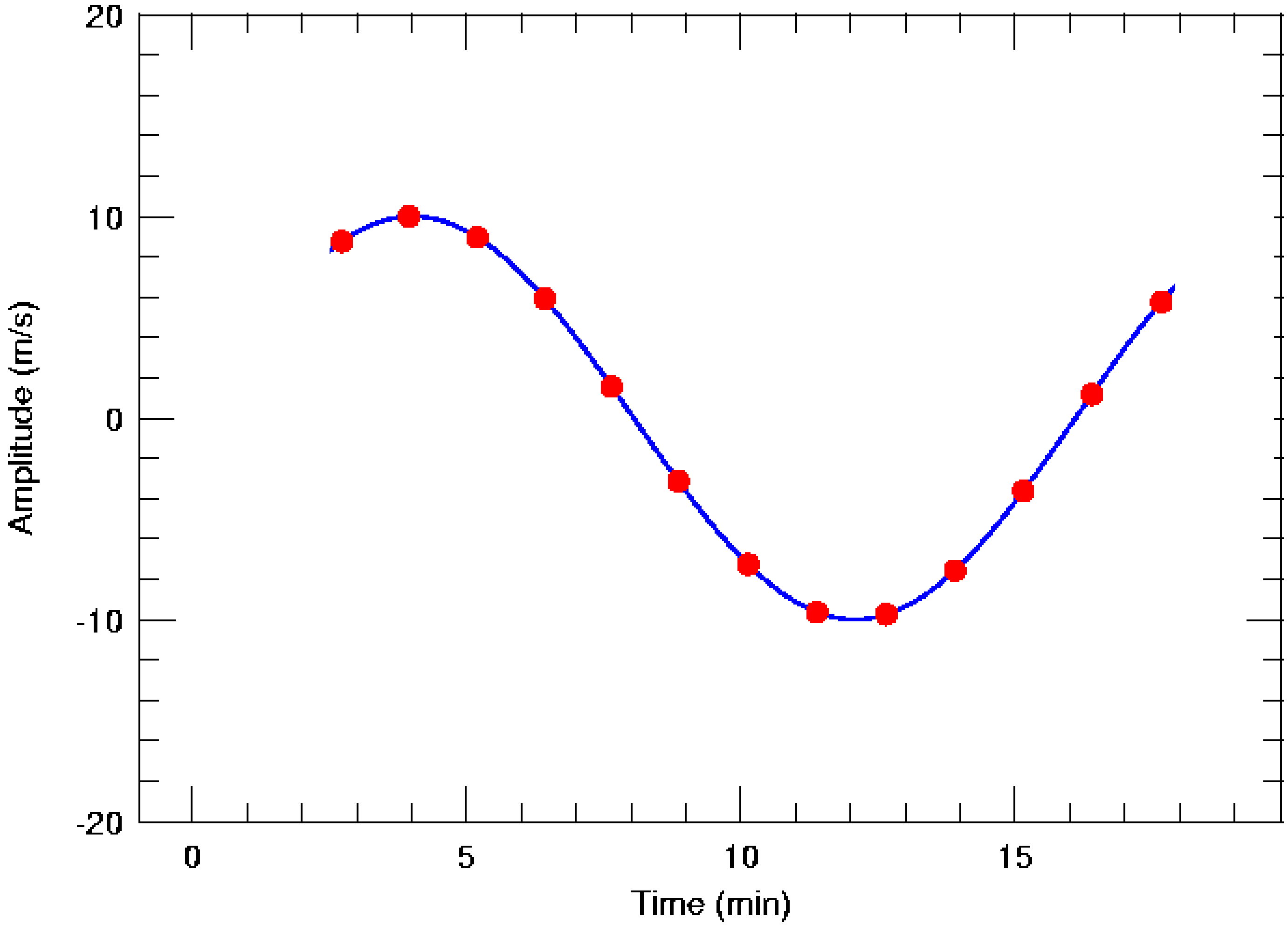
Normal Space	Fourier Space
$f * g$	$F \cdot G$

DFT of a pure sine wave:



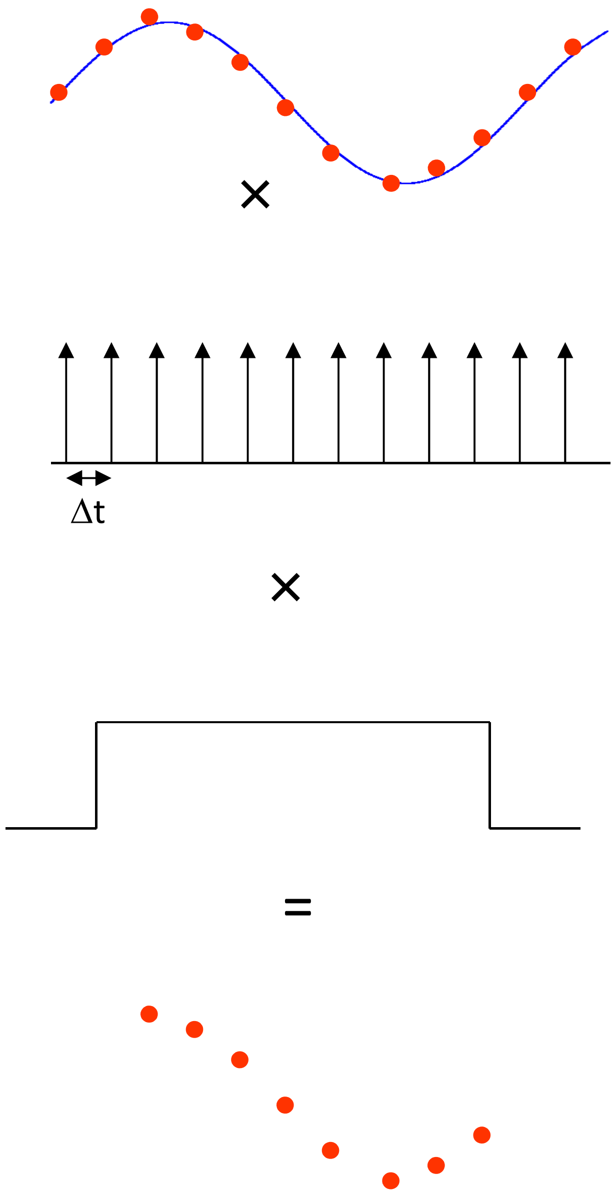
So why isn‘t it a δ-function?

It would be if we measured the blue line out to infinity:

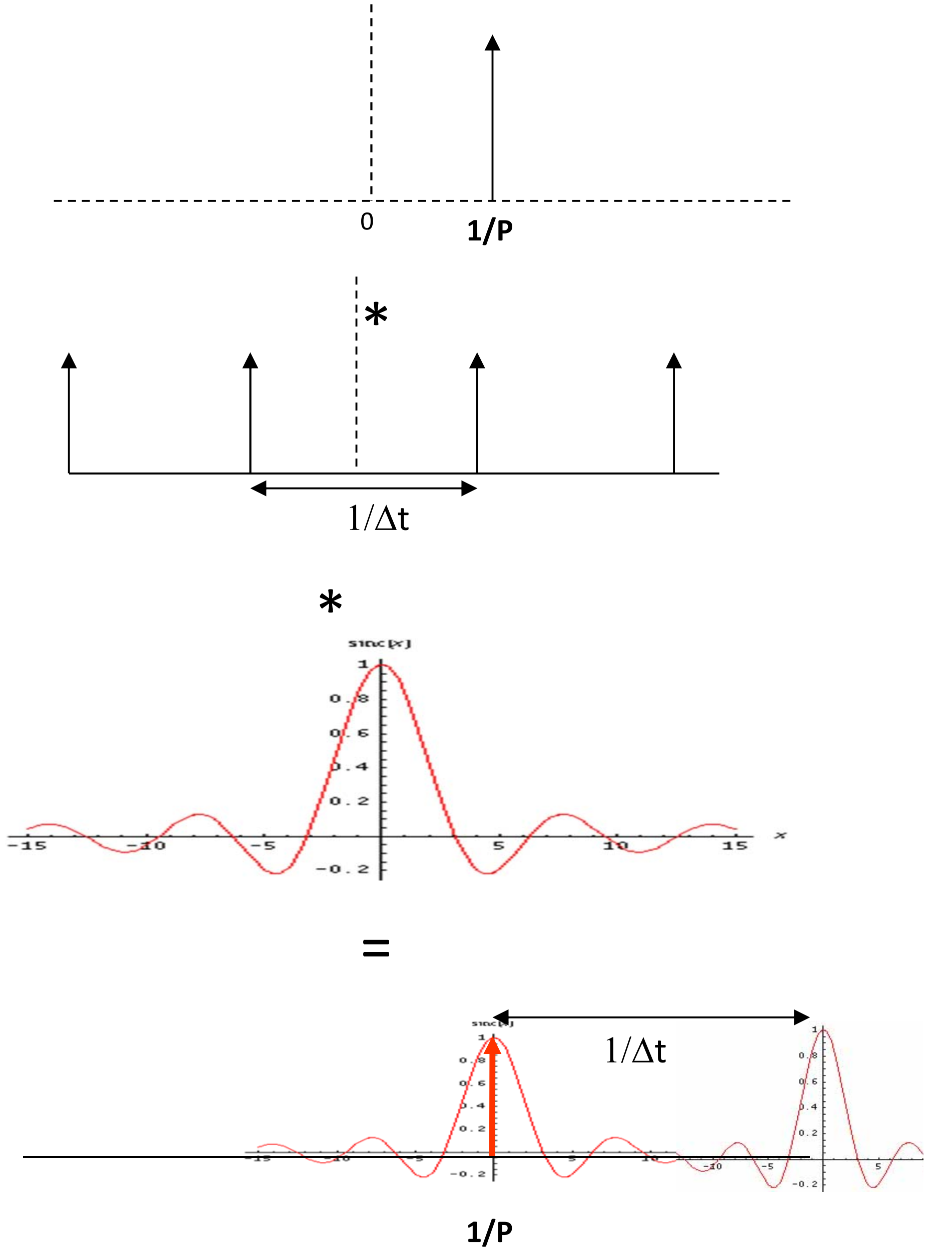


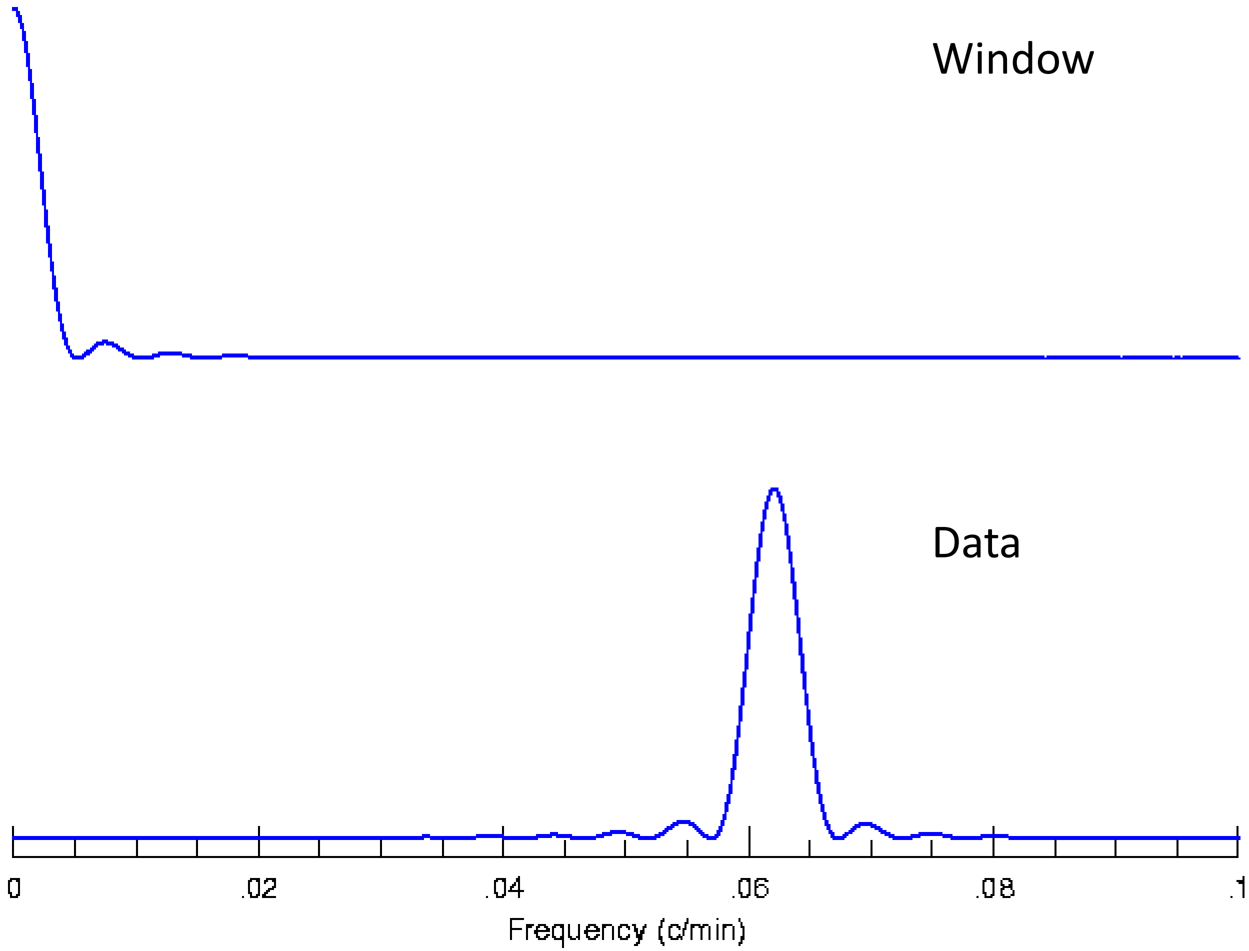
But we measure the red points. Our sampling degrades the delta function and introduces sidelobes

In time space



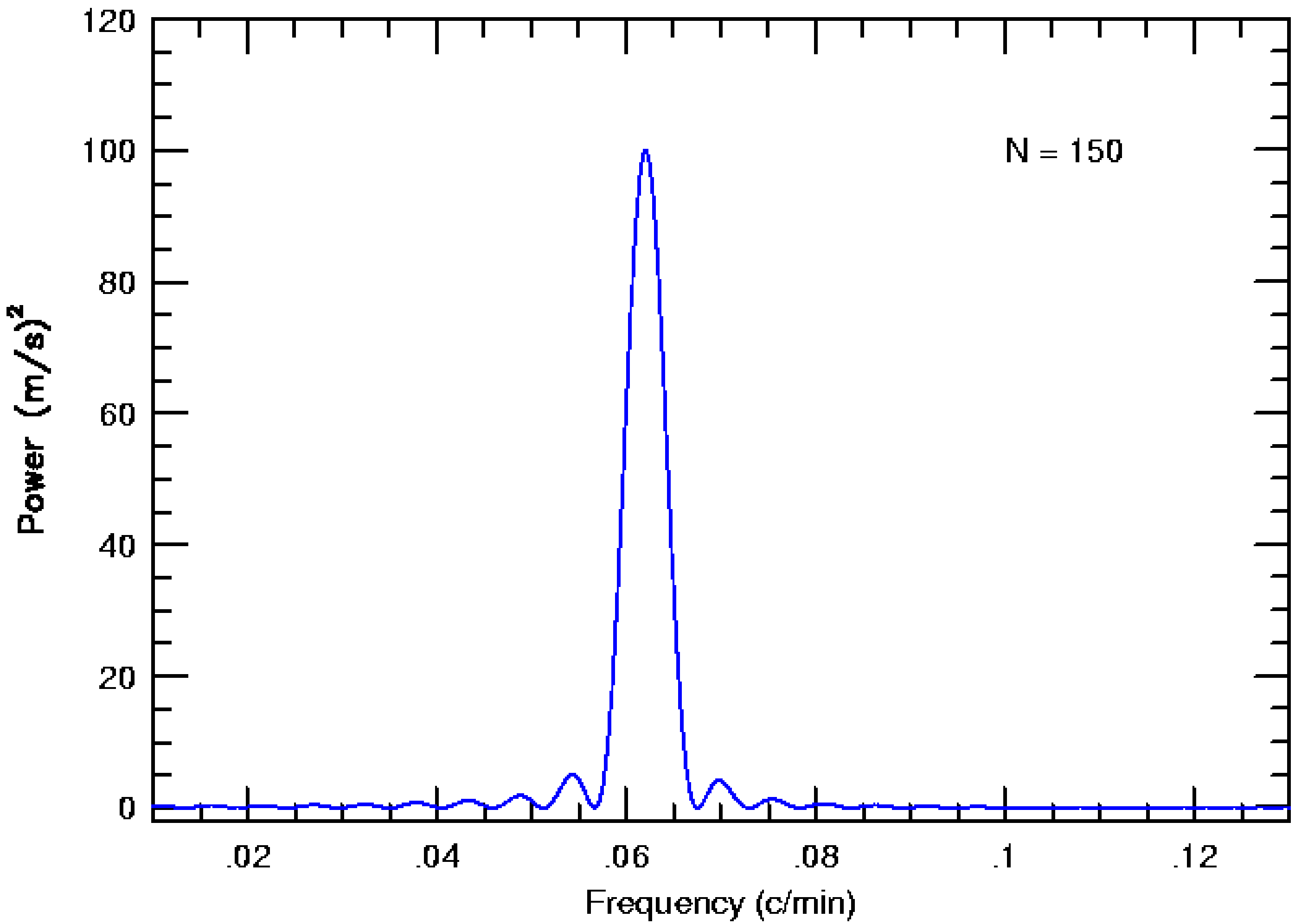
In Fourier Space





16 min period sampled regularly for 3 hours

The longer the data window, the narrower is the width of the sinc function window:



Error in the period (frequency) of a peak in DFT:

$$\delta\nu = \frac{3\pi\sigma}{2\, N^{1/2}\, T\, A}$$

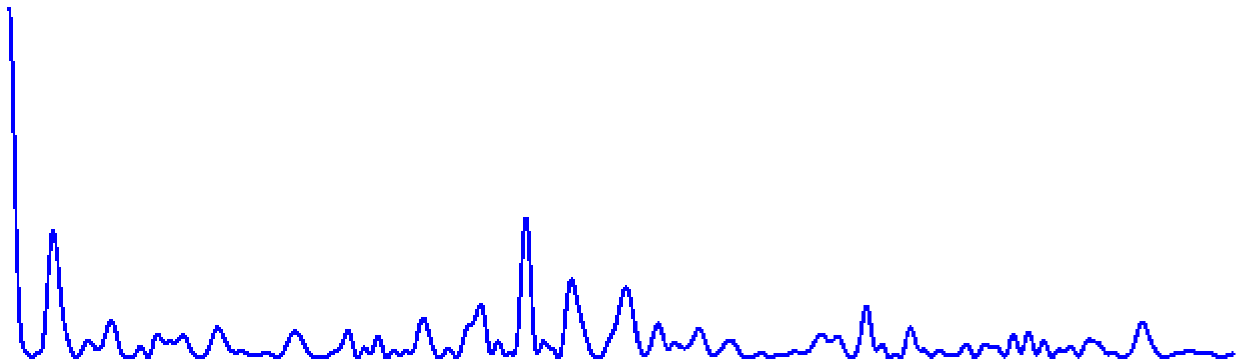
σ = error of measurement

T = time span of your observations

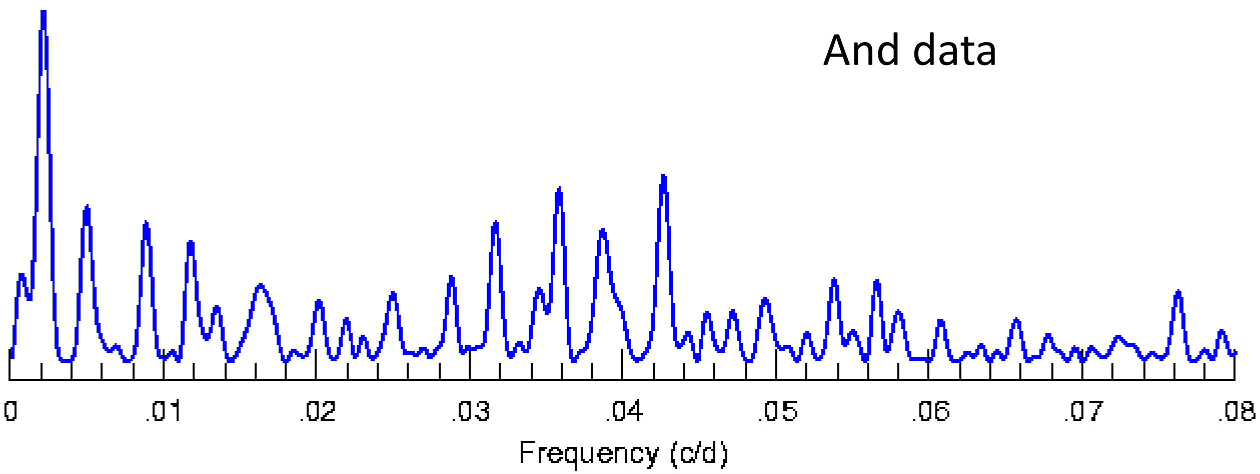
A = amplitude of your signal

N = number of data points

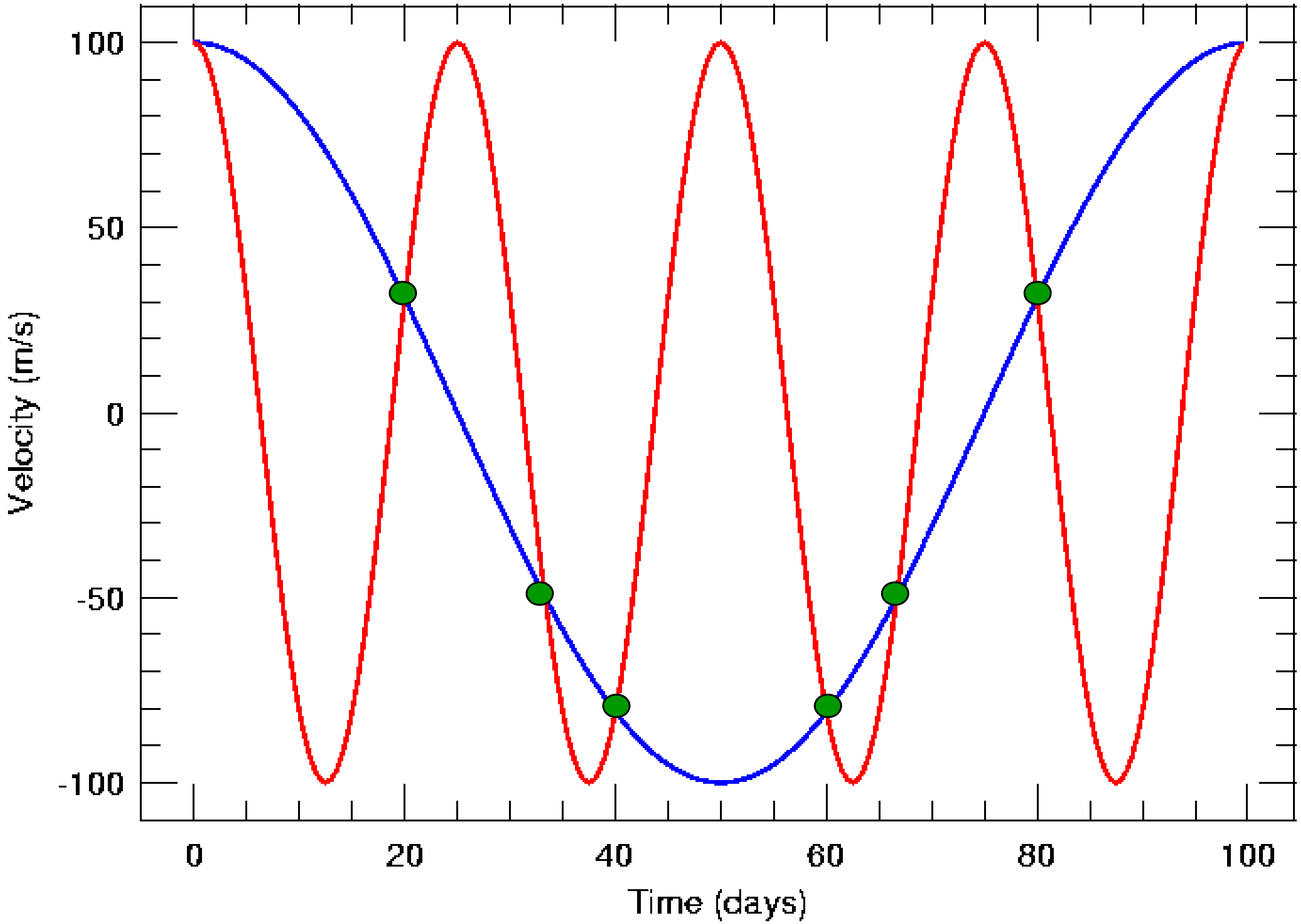
A more realistic window



And data



Alias periods:



Undersampled periods appearing as another period

Alias periods:

$$P_{\mathrm{false}}^{-1} = P_{\mathrm{alias}}^{-1} + P_{\mathrm{true}}^{-1}$$

Common Alias Periods:

$$P_{\mathrm{false}}^{-1} = (1\ day)^{-1} + P_{\mathrm{true}}^{-1}$$ day

$$P_{\mathrm{false}}^{-1} = (29.53\ d)^{-1} + P_{\mathrm{true}}^{-1}$$ month

$$P_{\mathrm{false}}^{-1} = (365.25\ d)^{-1} + P_{\mathrm{true}}^{-1}$$ year

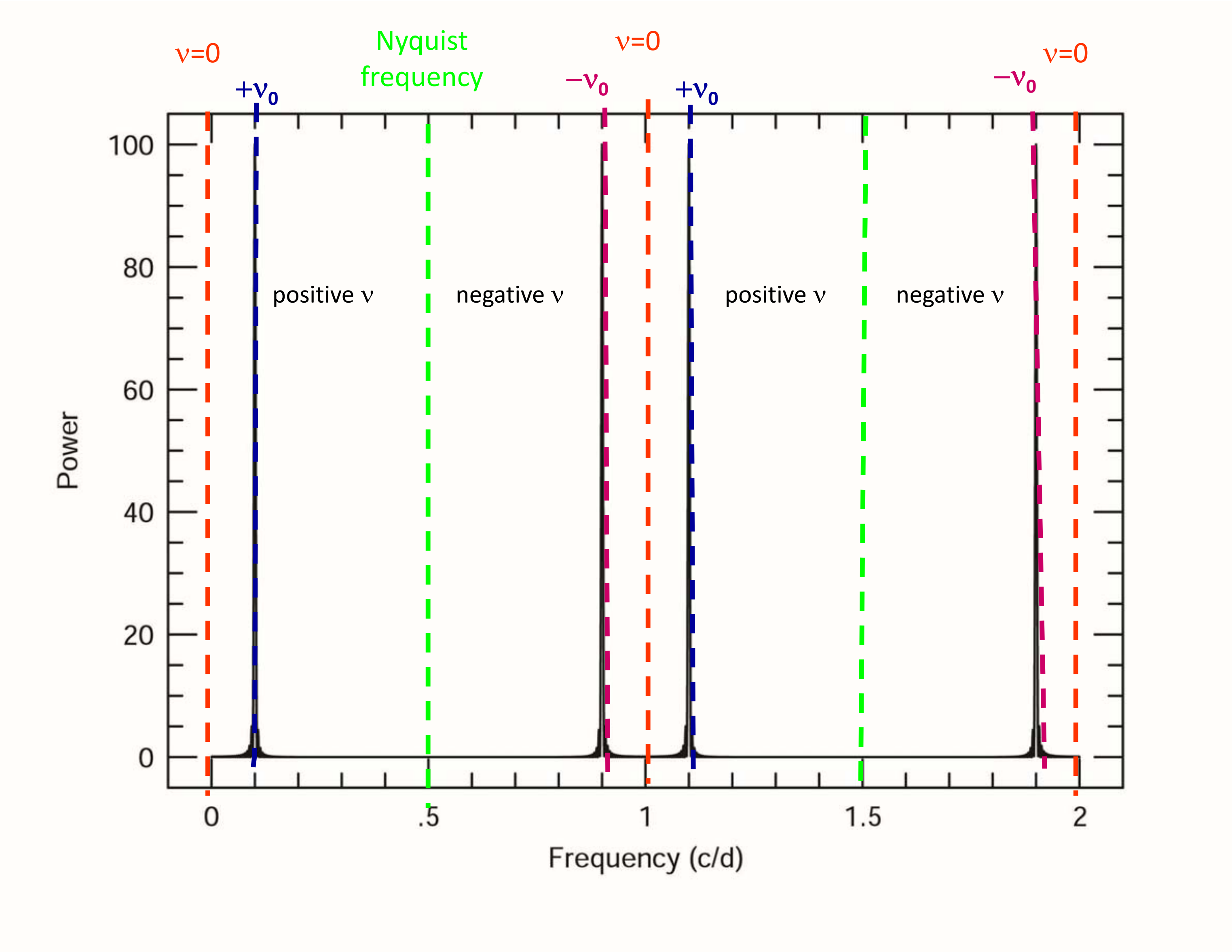
Nyquist Frequency

If T is your sampling rate which corresponds to a frequency of f_s, then signals with frequencies up to $f_s/2$ can be unambiguously reconstructed. This is the Nyquist frequency, N:

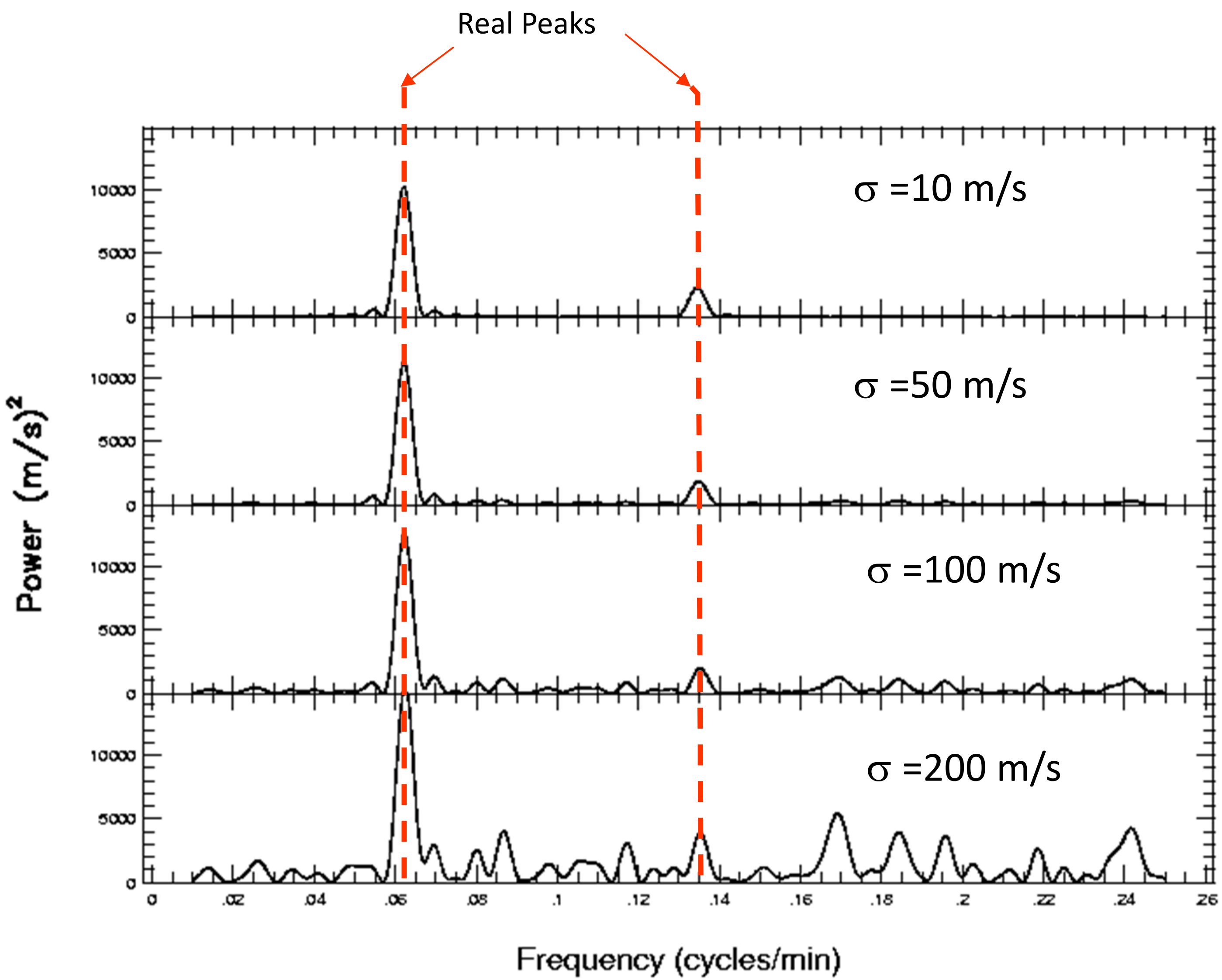
$$N < f_s/2$$

e.g. Suppose you observe a variable star once per night. Then the highest frequency you can determine in your data is 0.5 c/d = 2 days

When you do a DFT on a sine wave with a period = 10, sampling = 1: ν_0=0.1, $1/\Delta t$ = 1



The effects of noise:



2 sine waves ampltudes of 100 and 50 m/s. Noise added at different levels

3. Lomb-Scargle Periodograms

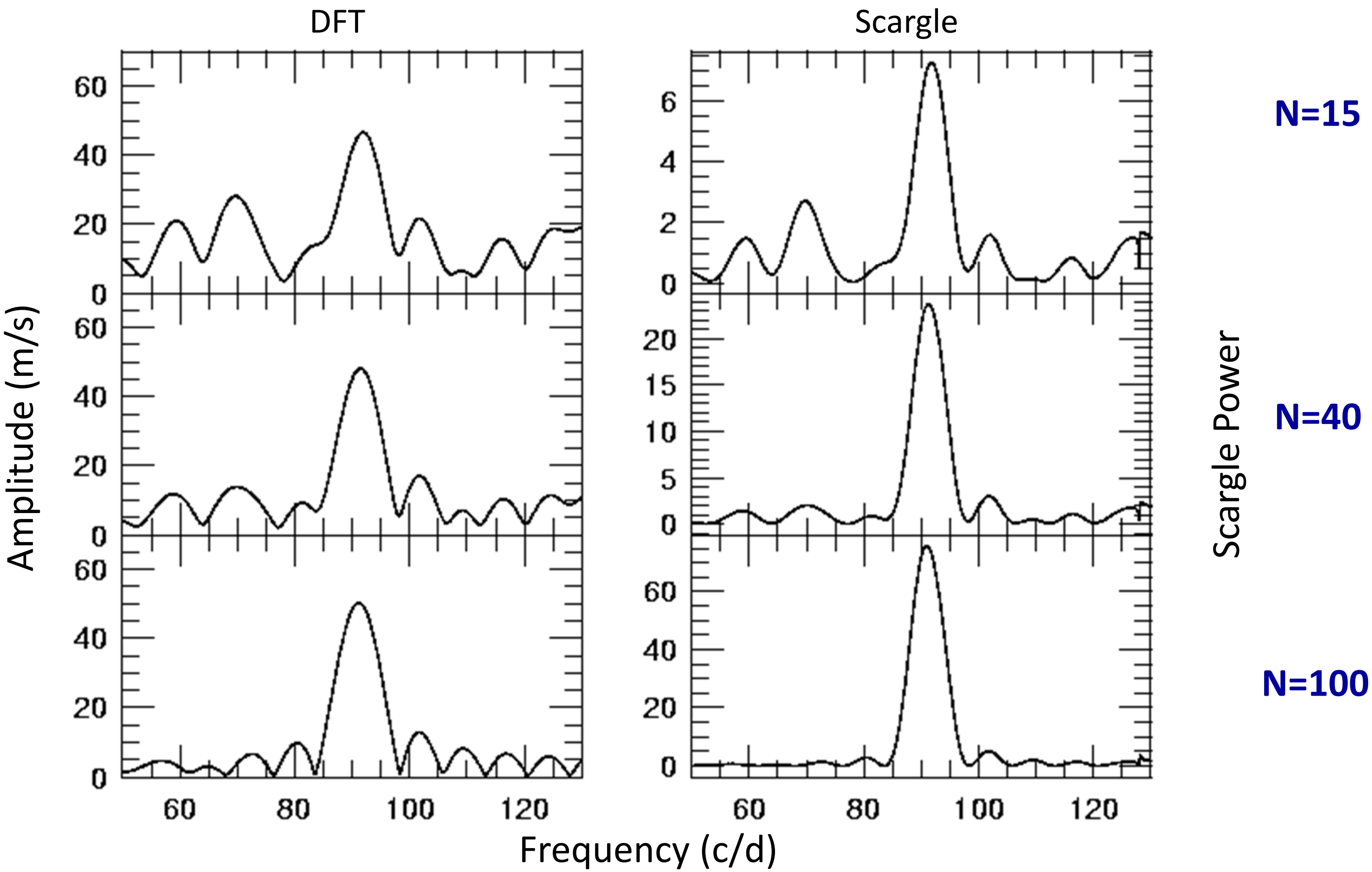
$$P_x(\omega) = \frac{1}{2}\frac{\left[\sum_j X_j \cos\omega(t_j-\tau)\right]^2}{\sum_j X_j \cos^2\omega(t_j-\tau)} + \frac{1}{2}\frac{\left[\sum_j X_j \sin\omega(t_j-\tau)\right]^2}{\sum X_j \sin^2\omega(t_j-\tau)}$$

$$\tan(2\omega\tau) = (\sum_j \sin 2\omega t_j)/(\sum_j \cos 2\omega t_j)$$

Power is a measure of the statistical significance of that frequency (period):

Scargle, Astrophysical Journal, 263, 835, 1982

Power: DFT versus Scargle



DFTs give you the amplitude of a periodic signal in the data. This does not change with more data. The Lomb-Scargle power gives you the statistical significance of a period. The more data you have the more significant the detection is, thus the higher power with more data

False Alarm Probability (FAP)

The FAP is the probability that random noise will produce a peak with Lomb-Scargle Power the same as your observed peak in a certain frequency range

Unknown period:

$$FAP \approx 1 - (1 - e^{-P})^N$$

Where P = Scargle Power

N = number of independent frequencies in the frequency range of interest

Known period:

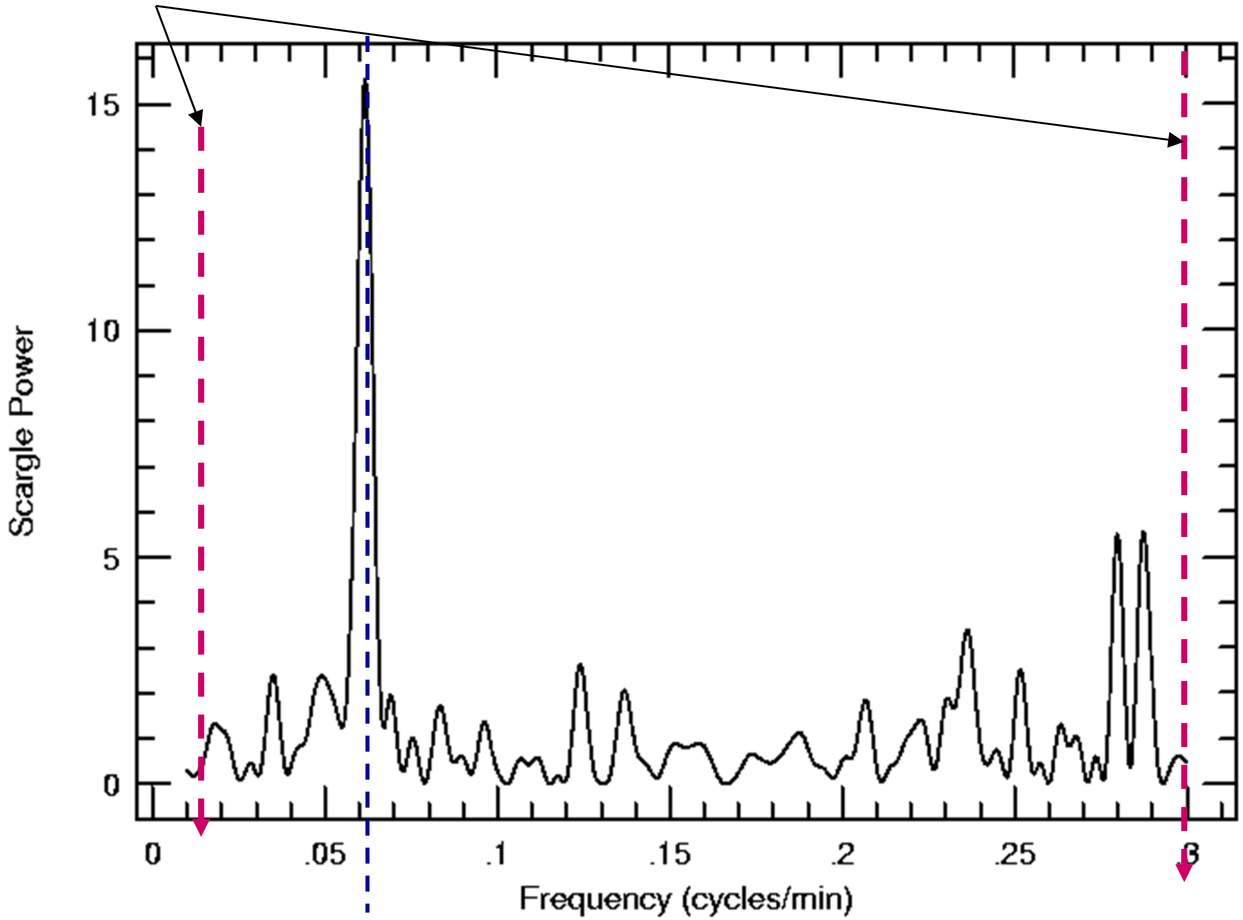
$$FAP \approx e^{-P}$$

In this case you have only one independent frequency

Scargle Power (significance) is increased by lower level of noise and/or more data points

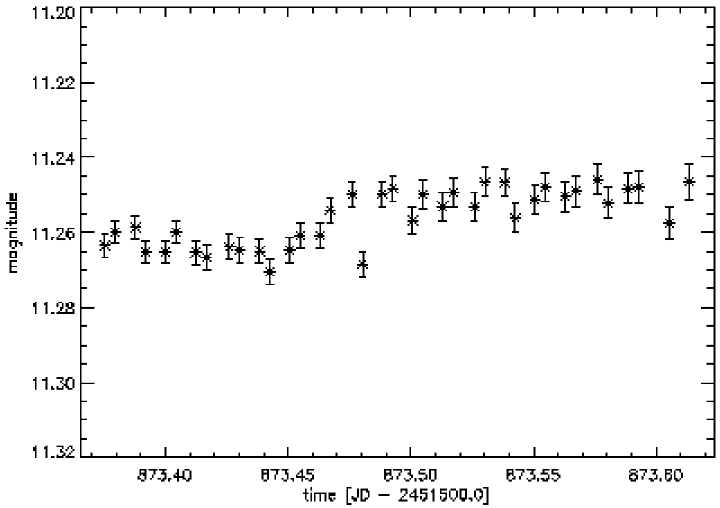
False Alarm Probability (FAP)

The probability that noise can produce the highest peak over a range ≈ $1 - (1 - e^{-P})^N$



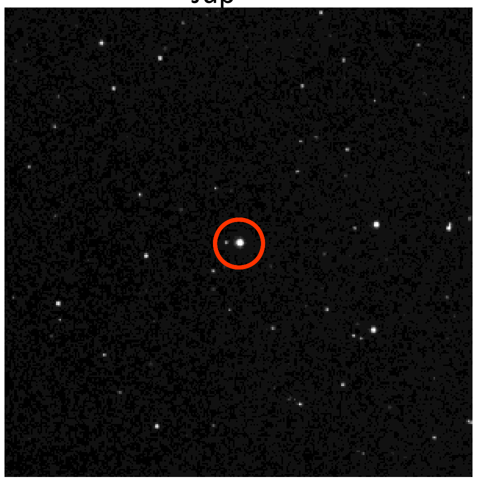
The probability that noise can produce the this peak exactly at this frequency = e^{-P}

Why is the FAP Impotant?
Example: A transit candidate from BEST



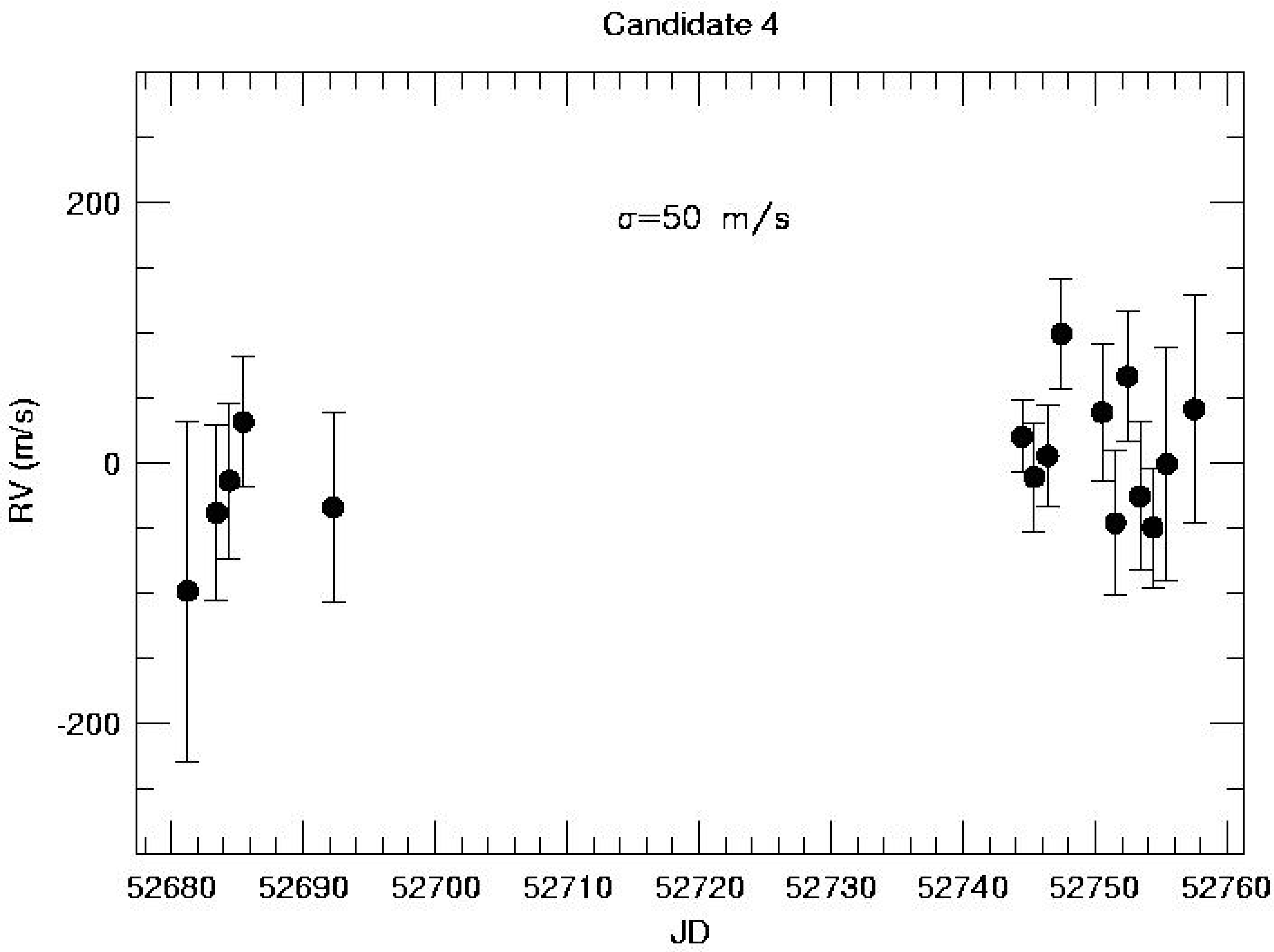
Depth [%]		1.2
Duration [h]		?
Orbital period [d]		
Semi mayor axis[AU]		
Number of detections		1
Target field No.		8
Host star	K...(?)	
Magnitude(B.E.S.T.)	11.25	
Radius[R_{sun}]		0.65-0.85
Radius of planet ? [R_{Jup}]		0.71-0.89?

To confirm you need radial velocity measurements, but you do not have a period...

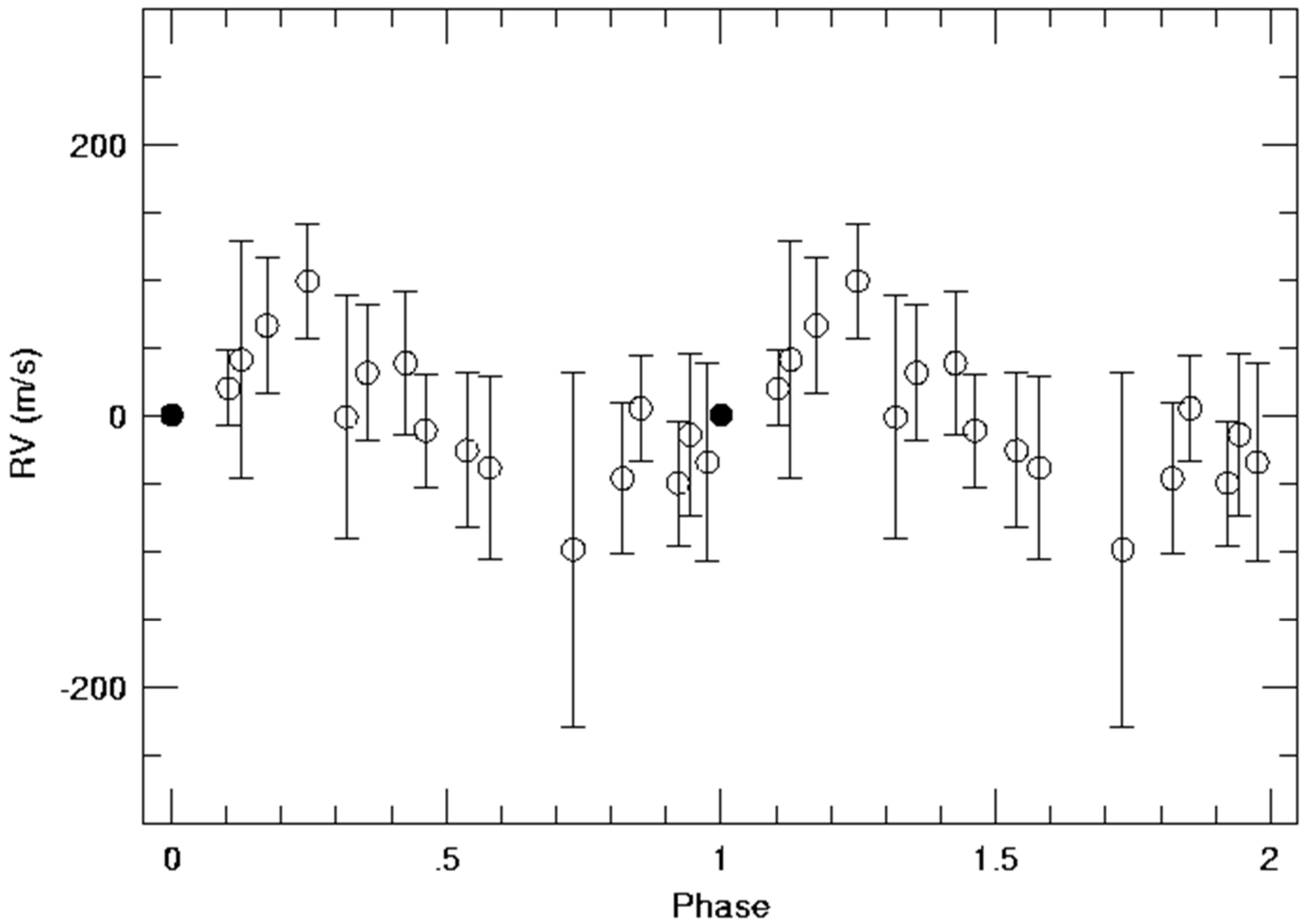


DSS1/POSSI

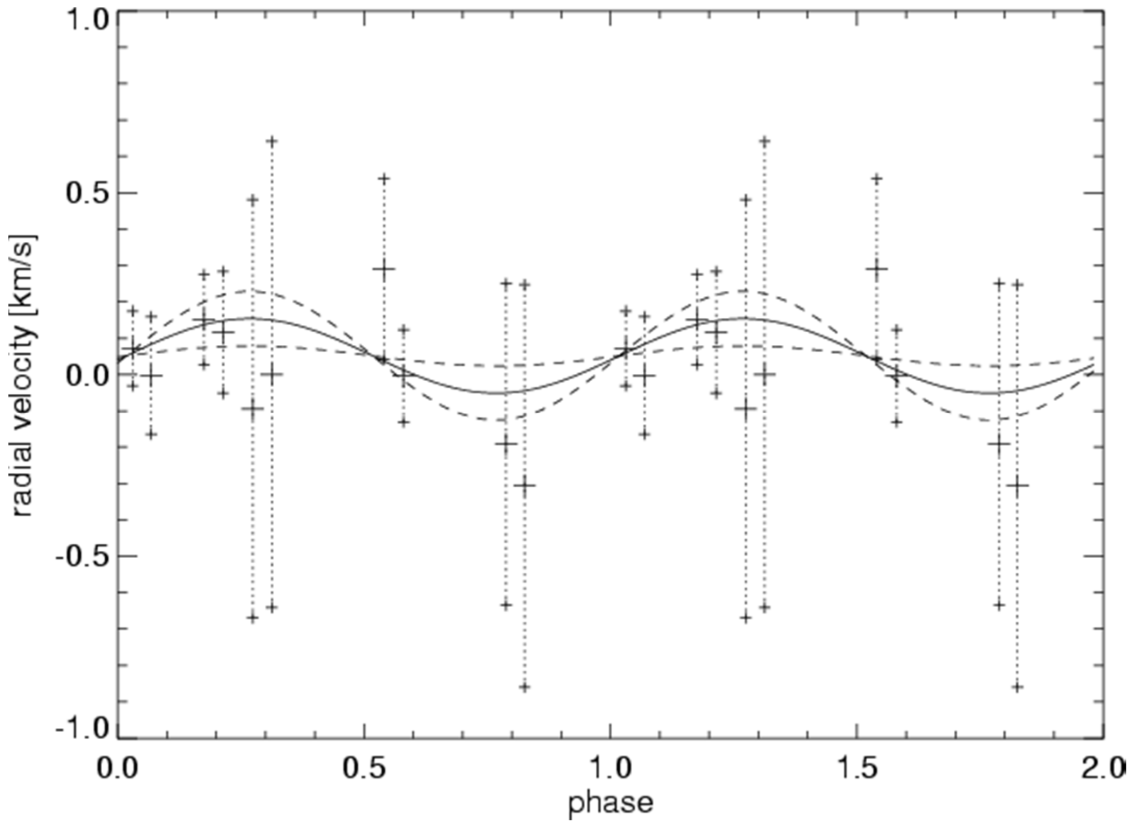
16 one-hour observations made with the 2m coude echelle



Least Square sine fit yields of 2.69 days

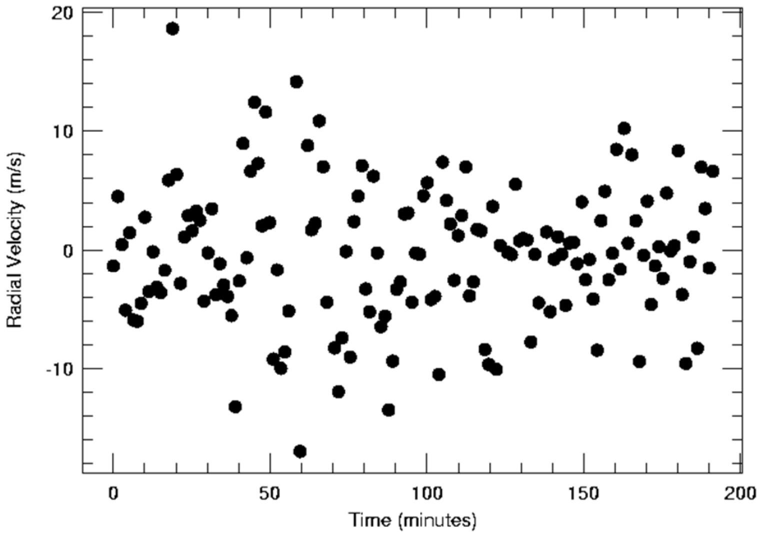


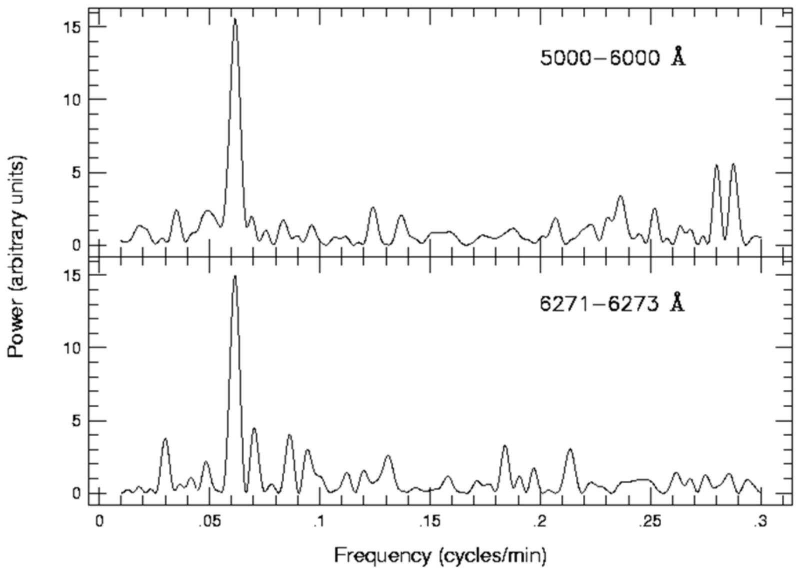
Published (Dreizler et al. 2003) Radial Velocity Curve of the transiting planet OGLE 3



Wrong Phase (by 180 degrees) for a transiting planet!

Discovery of a rapidly oscillating Ap star with 16.3 min period





FAP ≈ 10^{-5}

β CrB

Small FAP does not always mean a real signal

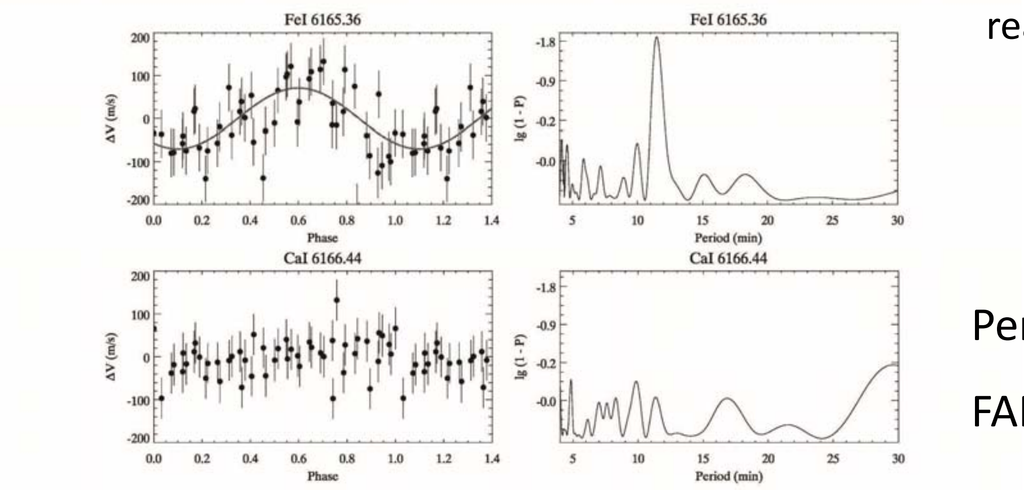
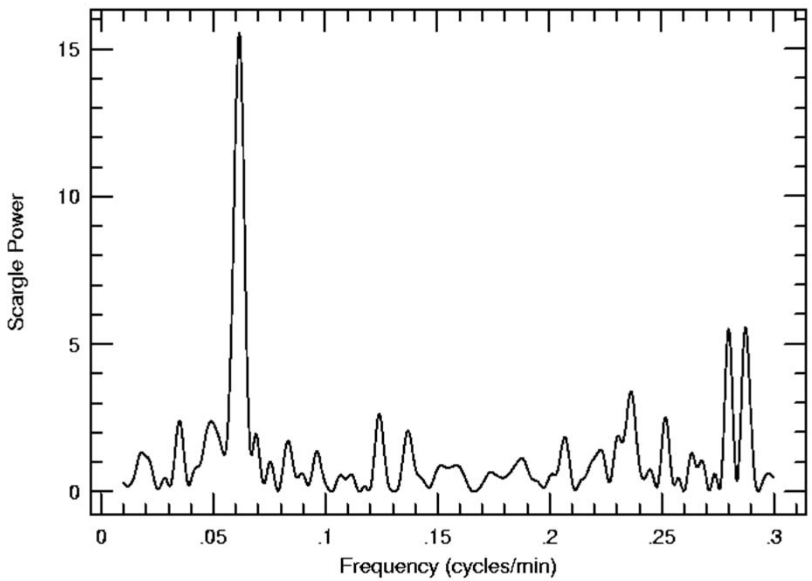


Figure 4. Radial velocity variations of the Fe I 6165.36 Å and Ca I 6166.44 Å lines in the spectrum of β CrB. Left panels show individual RV measurements (symbols) phased with the period of 11.5 min. The best cosine fit to the RVs of the Fe I line is shown by the solid curve. Right panels show periodograms for

Period = 11.5 min

FAP = 0.015



Period = 16.3 min

FAP = 10^{-5}

Lesson: Do not believe any FAP < 0.01

My limit: < 0.001

Better to miss a real period than to declare a false one

Determining FAP: To use the Scargle formula you need the number of independent frequencies.

How do you get the number of independent Frequencies?

First Approximation: Use the number of data points N_0

Horne & Baliunas (1986, Astrophysical Journal, 302, 757):

$N_i = -6.362 + 1.193\, N_0 + 0.00098 N_0^2$ = number of independent frequencies

Use Scargle FAP only as an estimate. A more valid determination of the FAP requires Monte Carlo Simulations:

Method 1:

1. Create random noise at the same level as your data
2. Sample the random noise in the same manner as your data
3. Calculate Scargle periodogram of noise and determine highest peak in frequency range of interest
4. Repeat 1.000-100.000 times = N_{total}
5. Add the number of noise periodograms with power greater than your data = N_{noise}
6. FAP = N_{noise}/N_{total}

Assumes Gaussian noise. What if your noise is not Gaussian, or has some unknown characteristics?

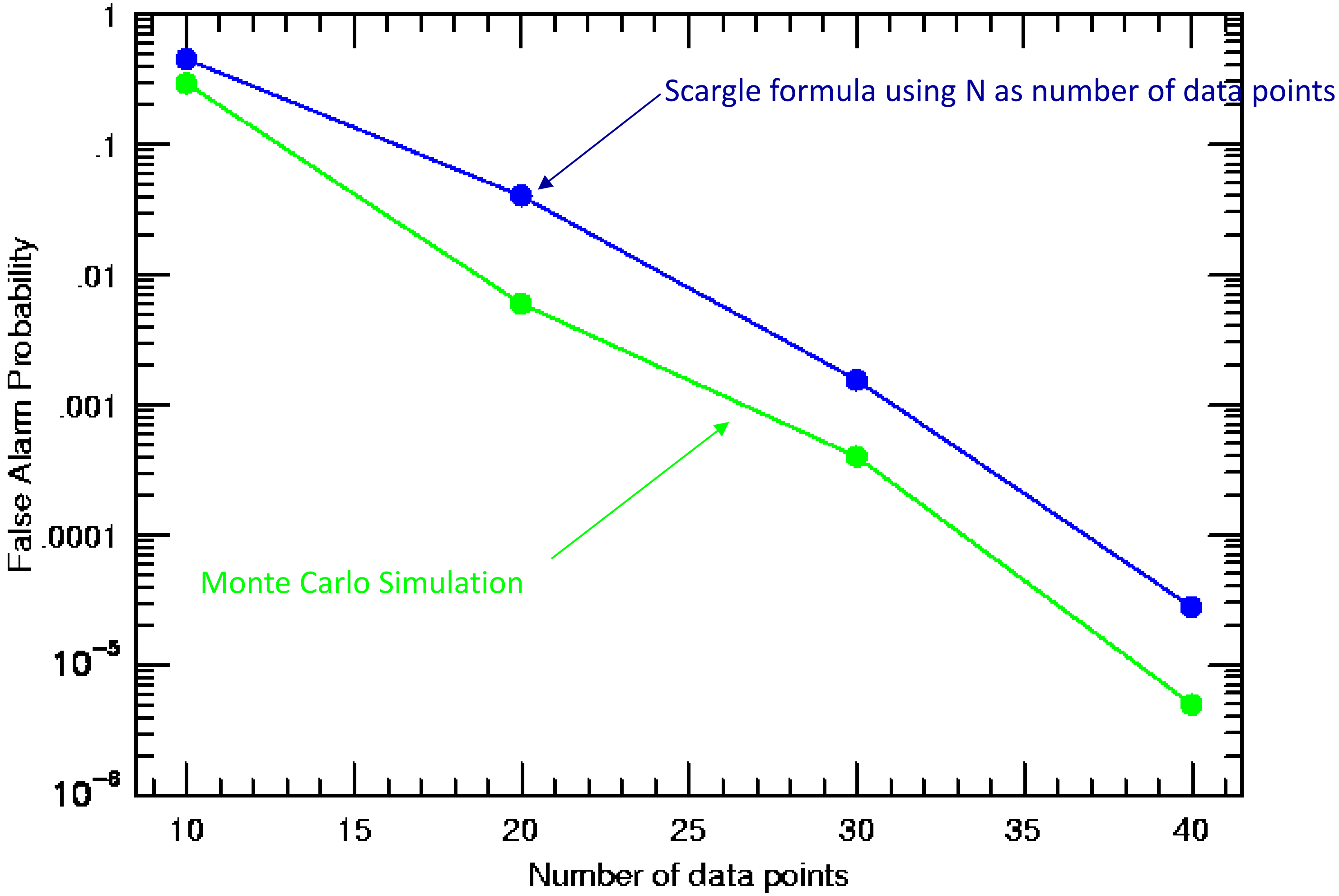
Use Scargle FAP only as an estimate. A more valid determination of the FAP requires Monte Carlo Simulations:

Method 2:

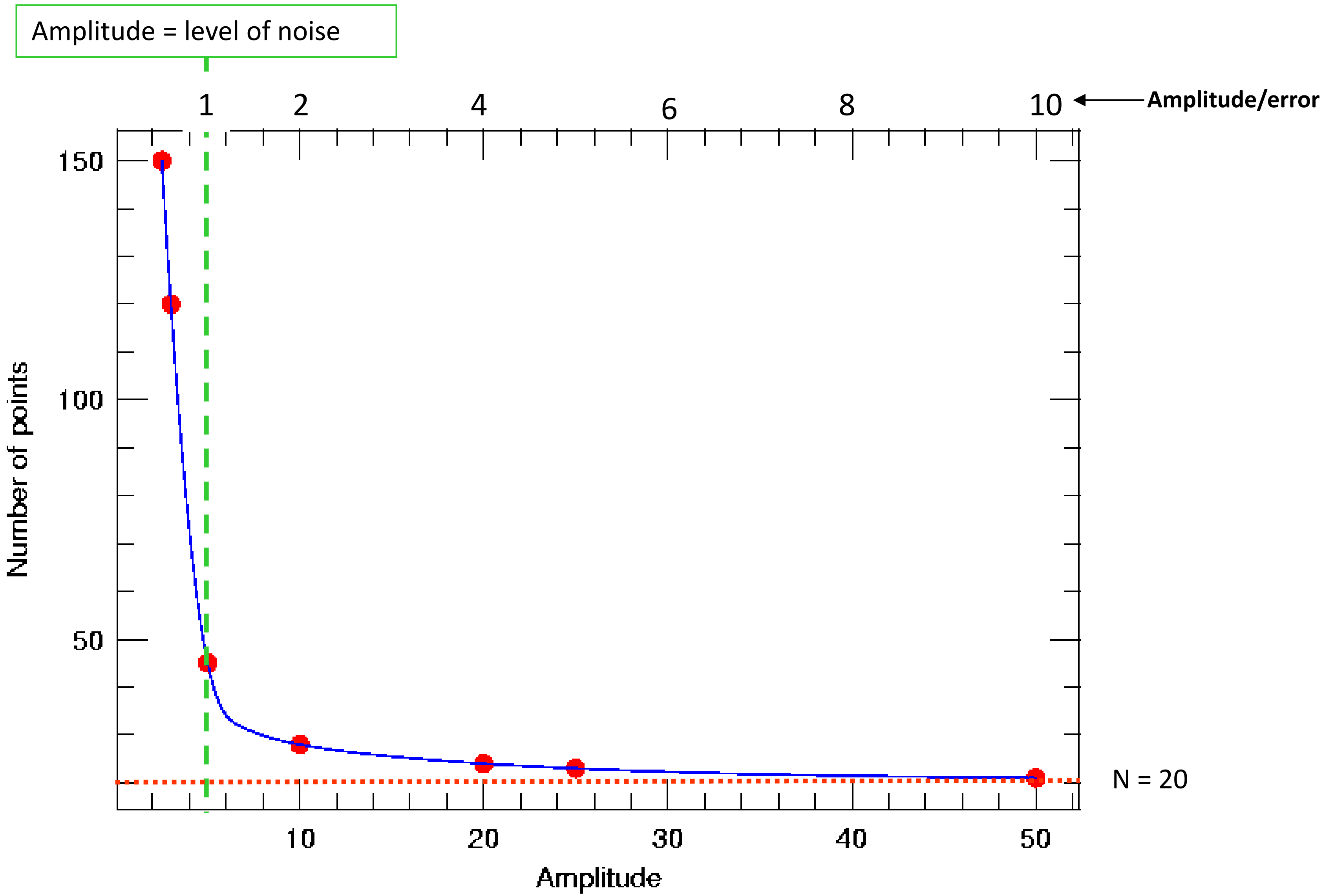
1. Randomly shuffle the measured values (velocity, light, etc) keeping the times of your observations fixed
2. Calculate Scargle periodogram of random data and determine highest peak in frequency range of interest
3. Reshuffle your data 1.000-100.000 times = N_{total}
4. Add the number of „random“ periodograms with power greater than your data = N_{noise}
5. FAP = N_{noise}/N_{total}

Advantage: Uses the actual noise characteristics of your data

FAP comparisons

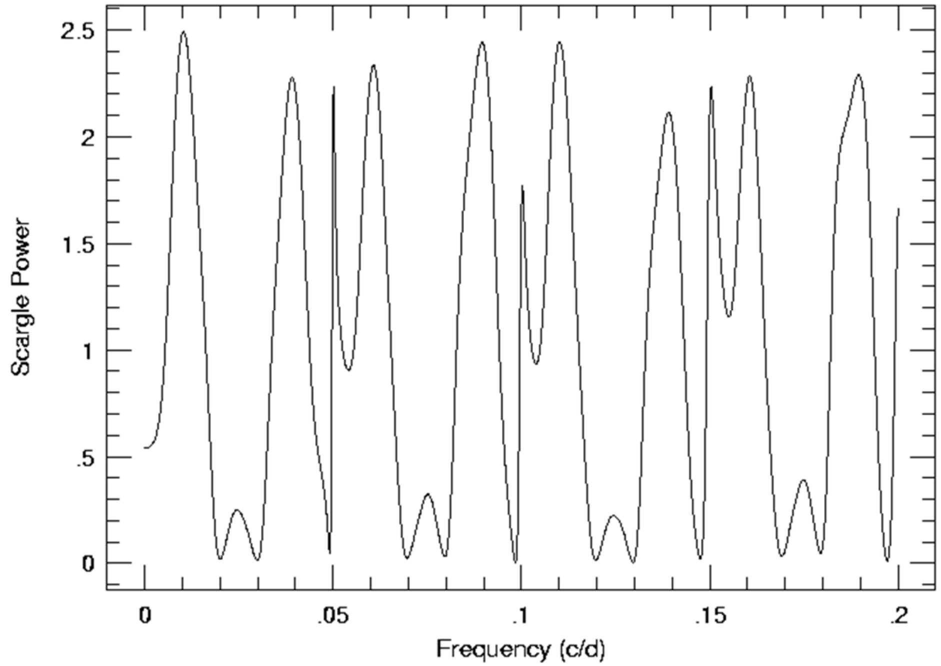


Using formula and number of data points as your independent frequencies may overestimate FAP, but each case is different.



Number of measurements needed to detect a signal of a certain amplitude. The FAP of the detection is 0.001. The noise level is σ = 5 m/s. Basically, the larger the measurement error the more measurements you need to detect a signal.

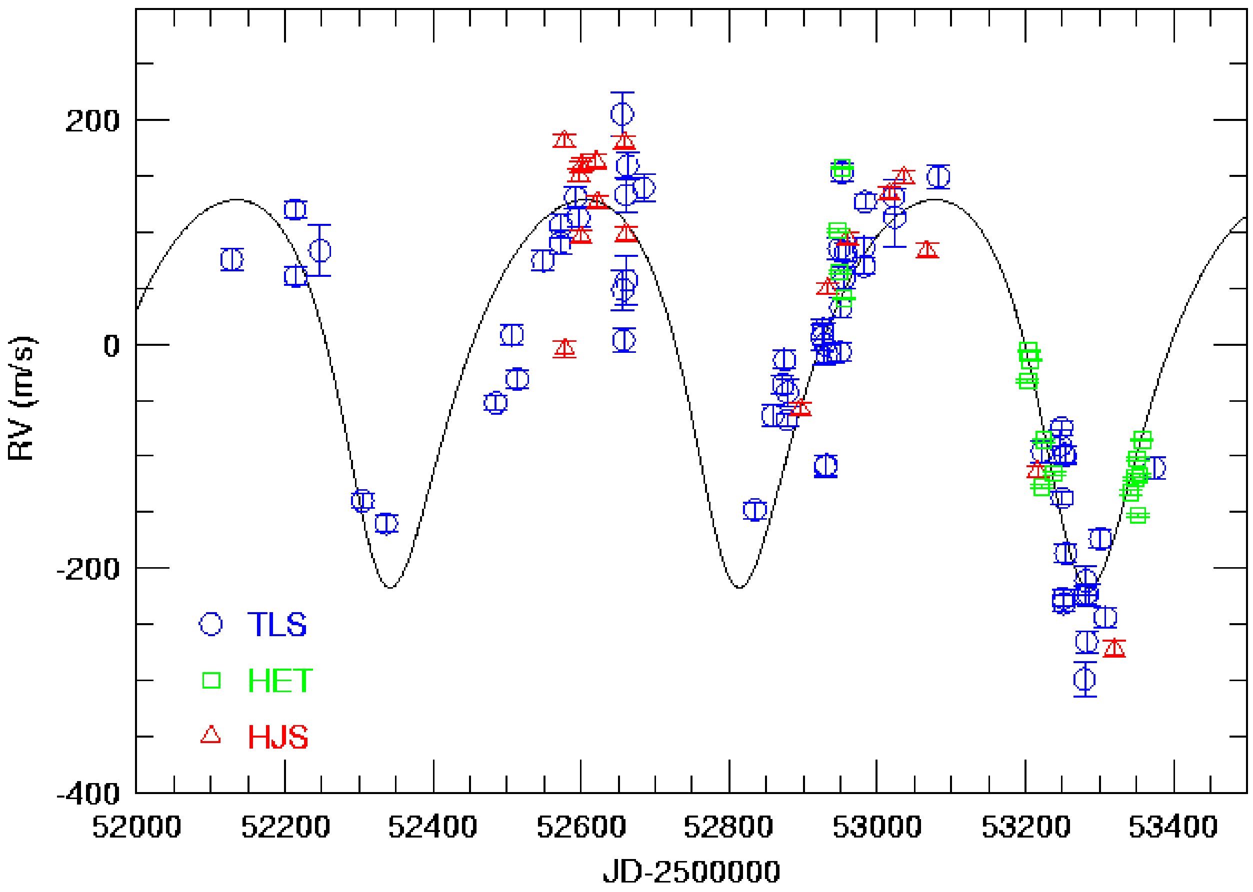
Lomb-Scargle Periodogram of 6 data points of a sine wave:

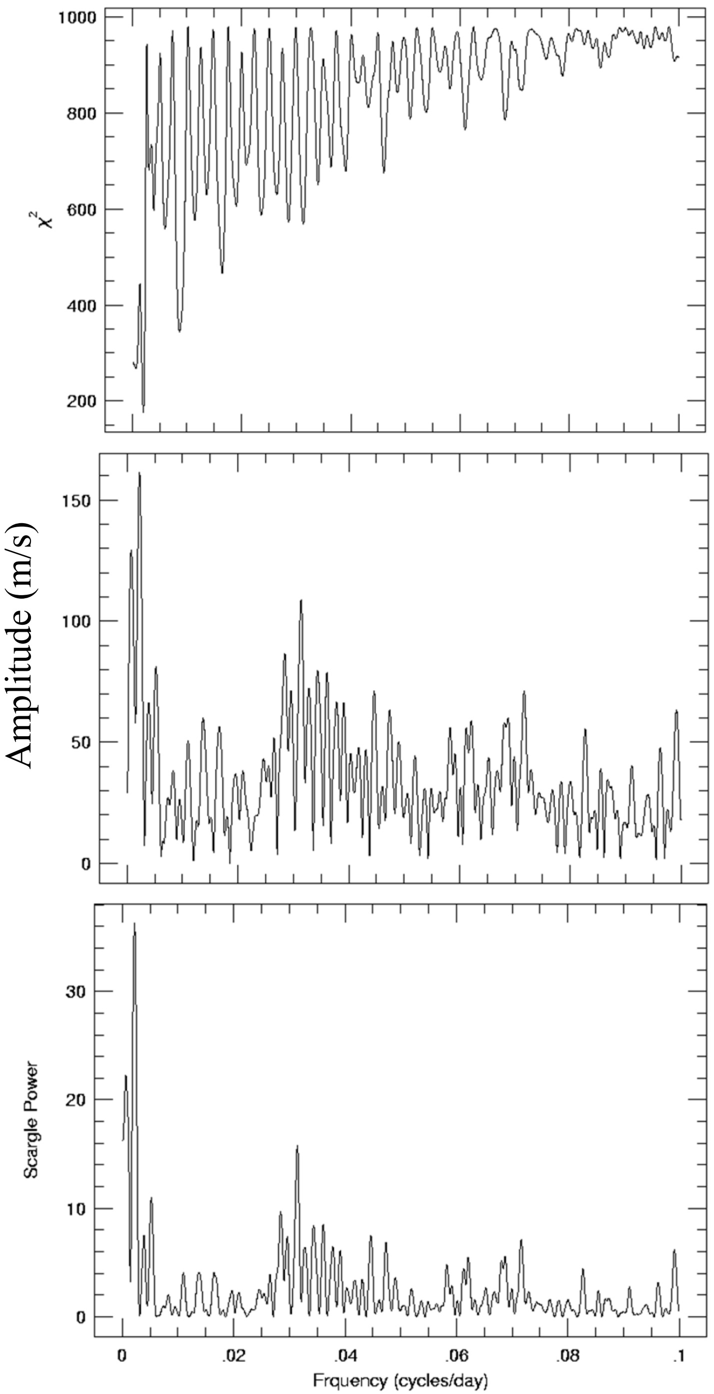


Lots of alias periods and false alarm probability (chance that it is due to noise) is 40%!

For small number of data points do not use Scargle, sine fitting is best. But be cautious!!

Comparison of the 3 Period finding techniques





Least squares sine fitting: The best fit period (frequency) has the lowest χ^2

Discrete Fourier Transform: Gives the power of each frequency that is present in the data. Power is in $(m/s)^2$ or (m/s) for amplitude

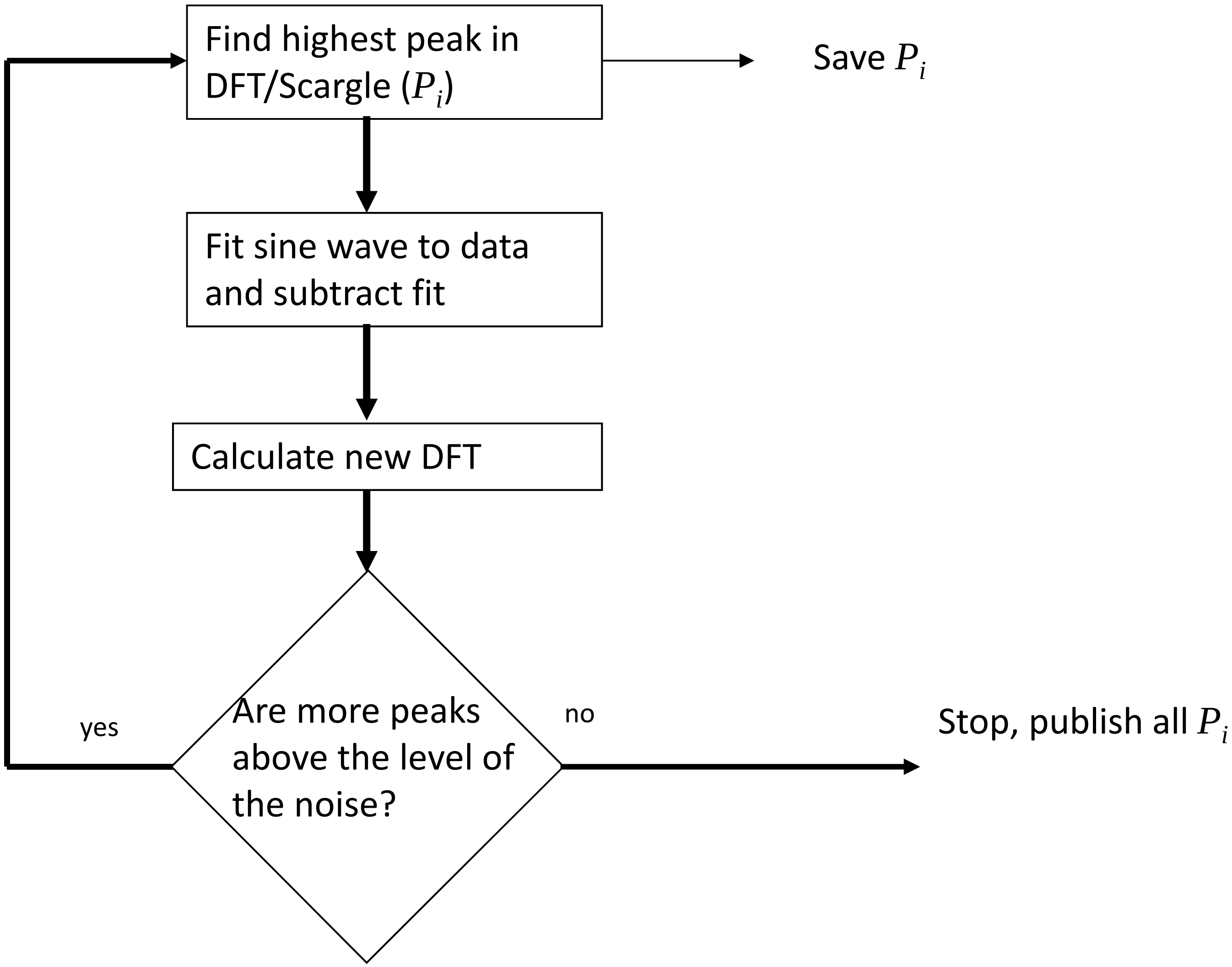
Lomb-Scargle Periodogram: Gives the power of each frequency that is present in the data. Power is a measure of statistical signficance

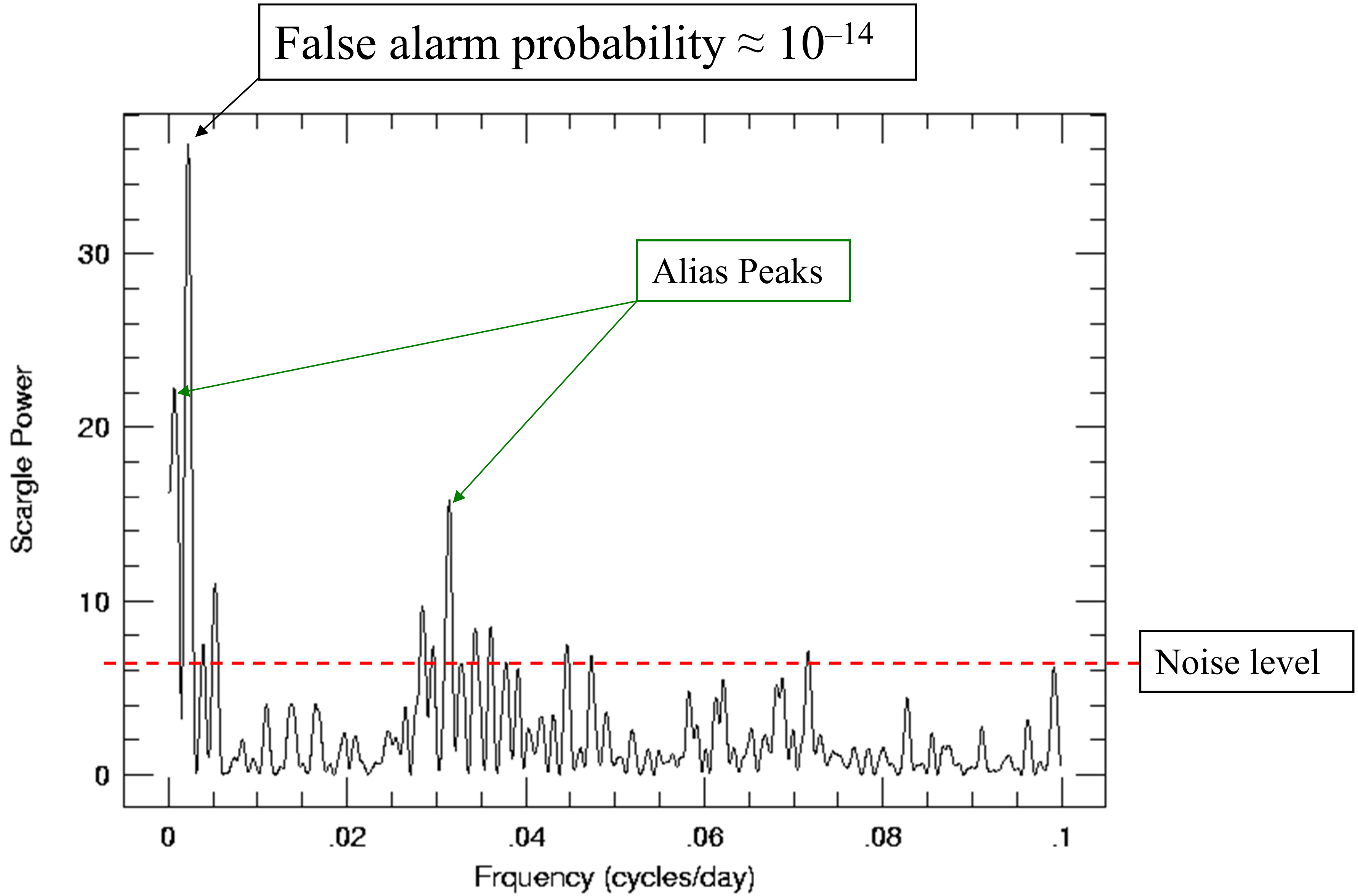
4. Finding Multiple Periods in Data: Pre-whitening

What if you have multiple periods in your data? How do you find these and make sure that these are not due to alias effects of your sampling window.

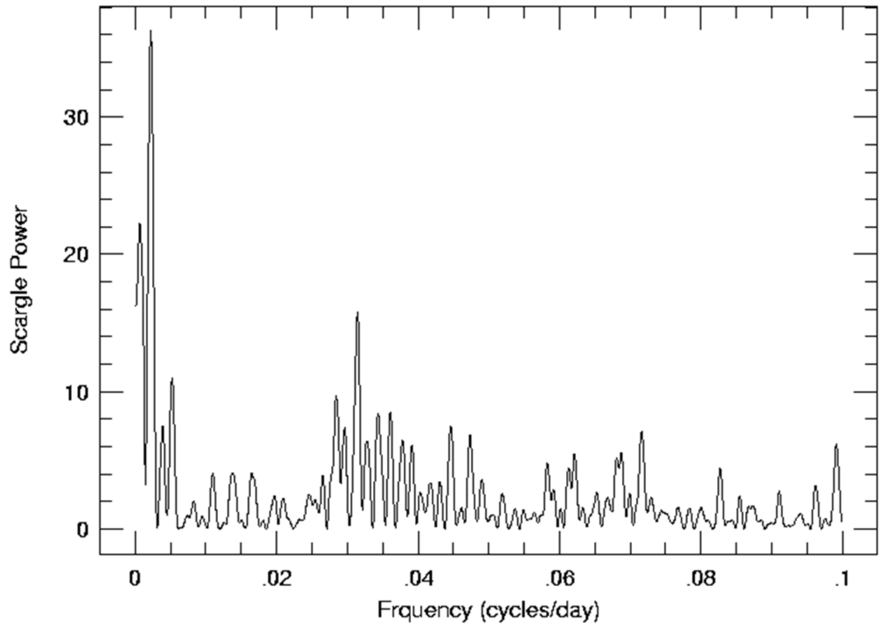
Standard procedure: ***Pre-whitening.*** Sequentially remove periods from the data until you reach the level of the noise

Prewhitening flow diagram:

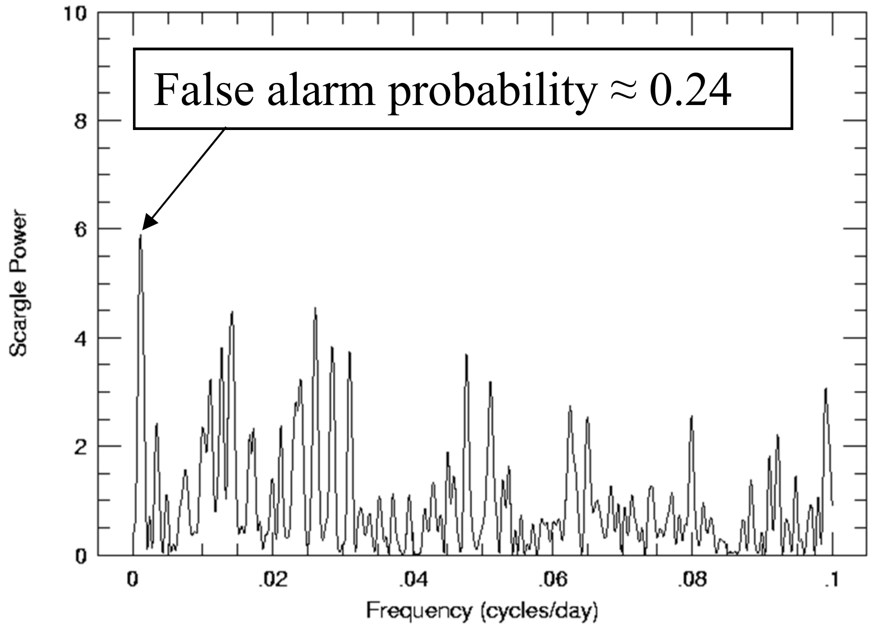




False alarm probability ≈ 10^{-14}
Alias Peaks
Noise level
Scargle Power
30
20
10
0
.02
.04
.06
.08
.1
Frquency (cycles/day)



Raw data



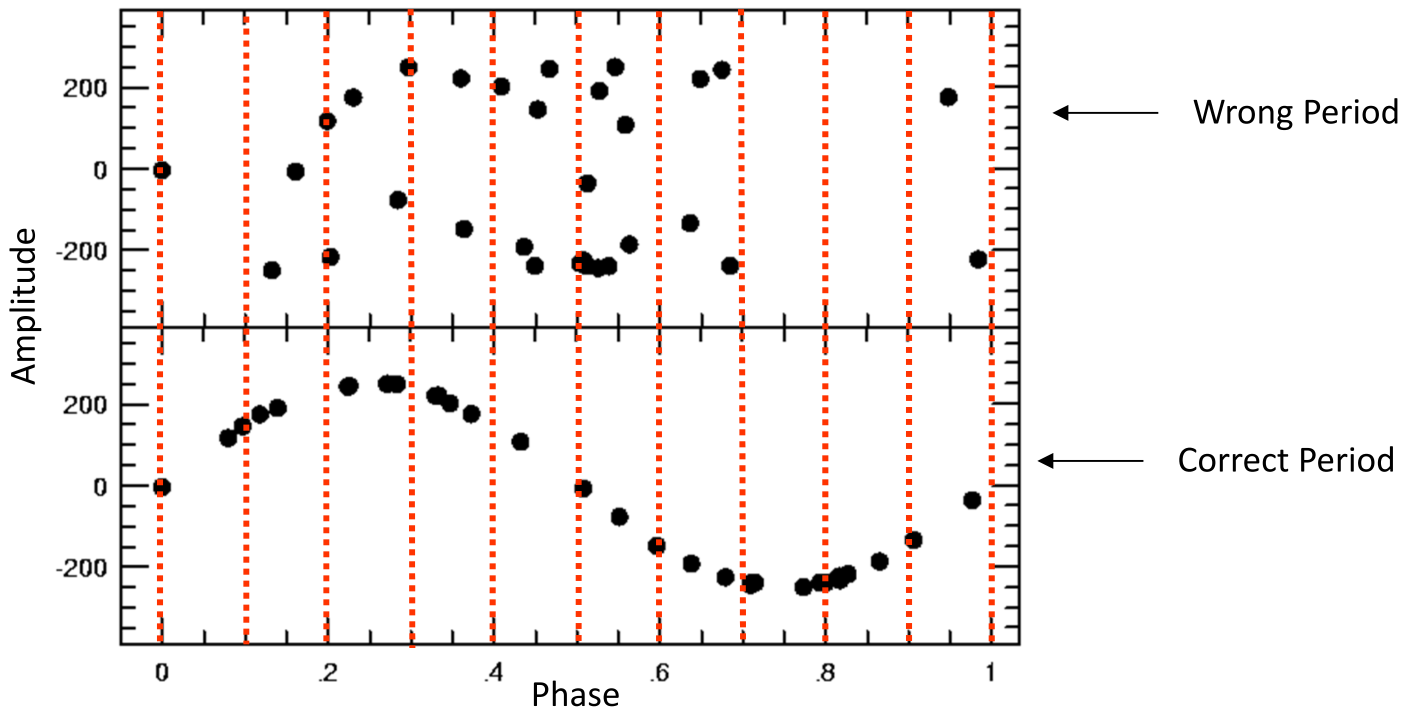
After removal of dominant period

Useful program for pre-whitening of time series data:

http://www.univie.ac.at/tops/Period04/

- Program picks highest peak, but this may be an alias
- Peaks may be due to noise. A FAP analysis will tell you this

5. Other Techniques: Phase Dispersion Minimization



Choose a period and phase the data. Divide phased data into *M* bins and compute the standard deviation in each bin. If σ^2 is the variance of the time series data and s^2 the total variance of the *M* bin samples, the correct period has a minimum value of Θ :

$$\Theta = s^2/\sigma^2$$

See Stellingwerf, Astromomical Journal, 224, 953, 1978

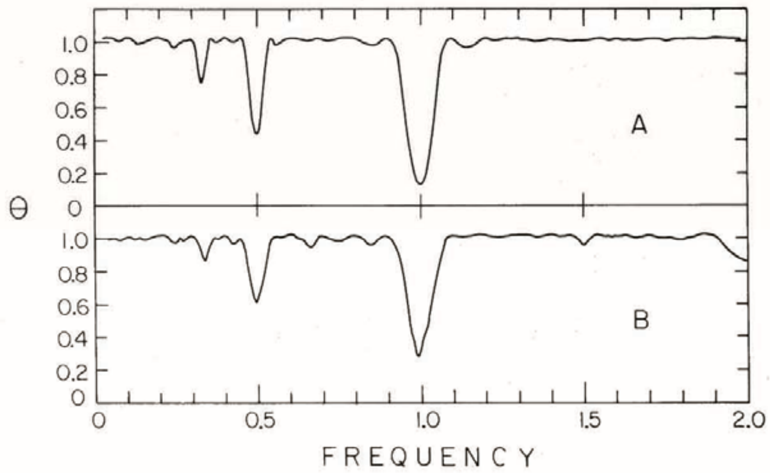


FIG. 1.—The Θ statistic versus frequency for the two test cases described in the text. (*a*) Sine-wave transform; (*b*) sawtooth function transform.

PDM is suited to cases in which only a few observations are made of a limited period of time, especially if the light curves are highly non-sinsuisoidal

PDM

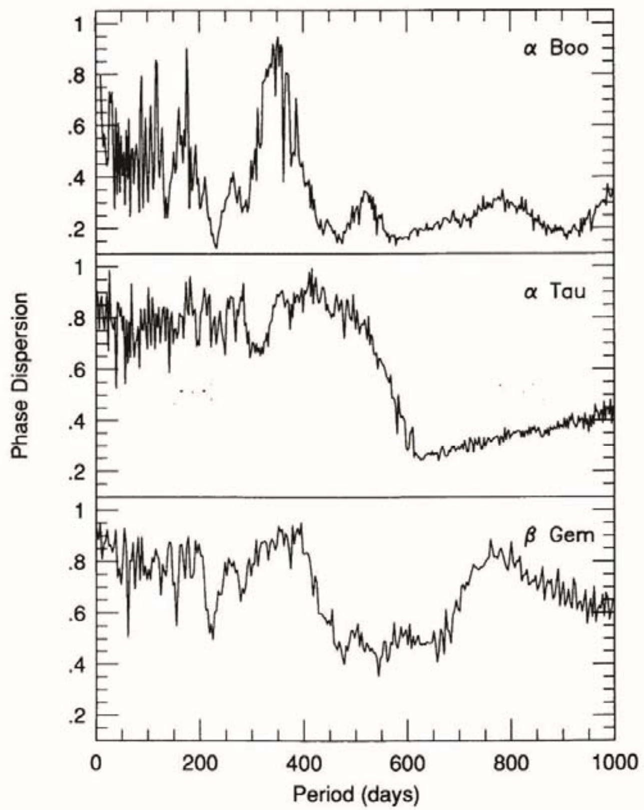


FIG. 3.—The phase dispersion minimization θ-diagram for α Boo (*top*), α Tau (*middle*), and β Gem (*bottom*). Minima represent possible periods in the data sets.

Scargle

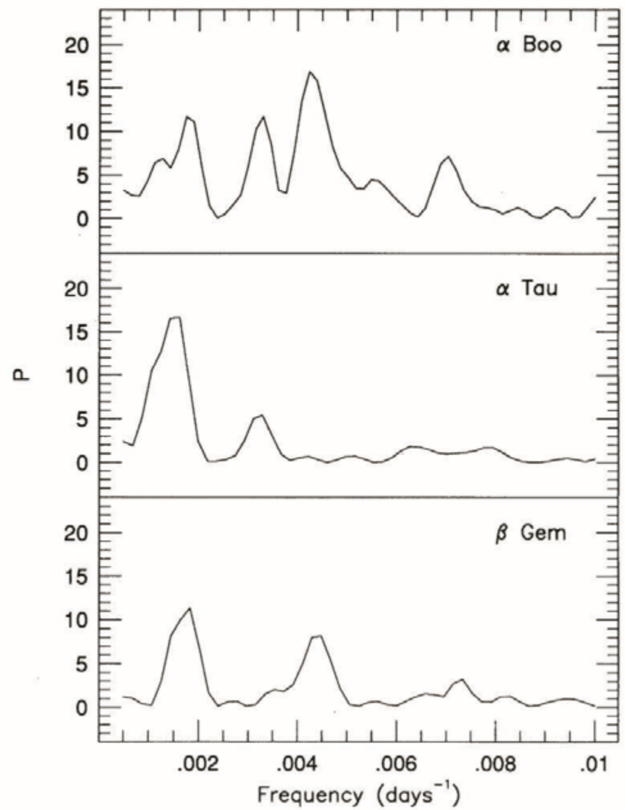
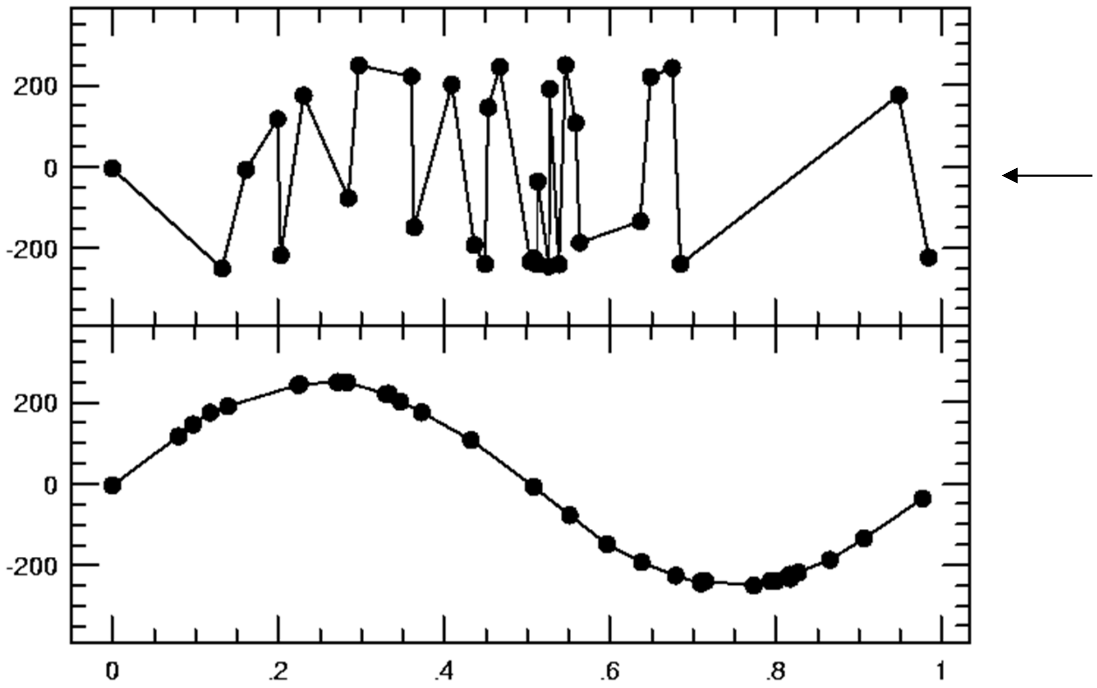


FIG. 2.—Periodograms in the frequency range $0 < \nu < 0.01$ days^{-1} for α Boo (*top*), α Tau (*middle*), and β Gem (*bottom*).

In most cases PDM gives the same answer as DFT, Scargle periodograms. With enough data it should give the same answer

5. Other Techniques: String Length Method



Wrong period longer string length

Phase the data to a test period and minimize the distance between adjacent points

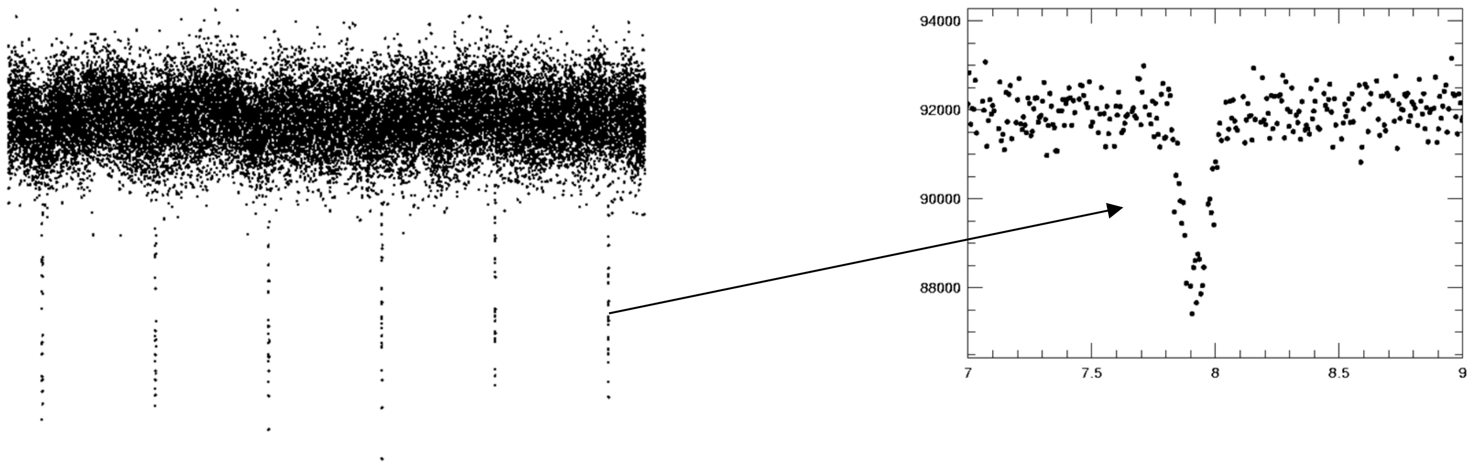
Lafler & Kinman, Astrophysical Journal Supplement, 11, 216, 1965

Dworetsky, Monthly Notices Astronomical Society, 203, 917, 1983

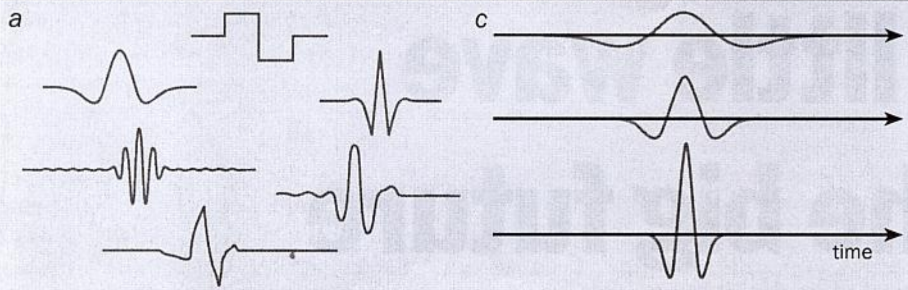
5. Other Techniques: Wavelet Analysis

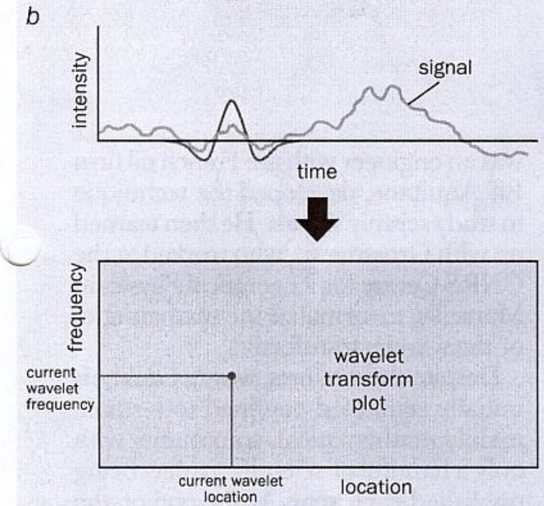
This technique is ideal for finding signals that are aperiodic, noisy, intermittent, or transient.

Recent interest has been in transit detection in light curves



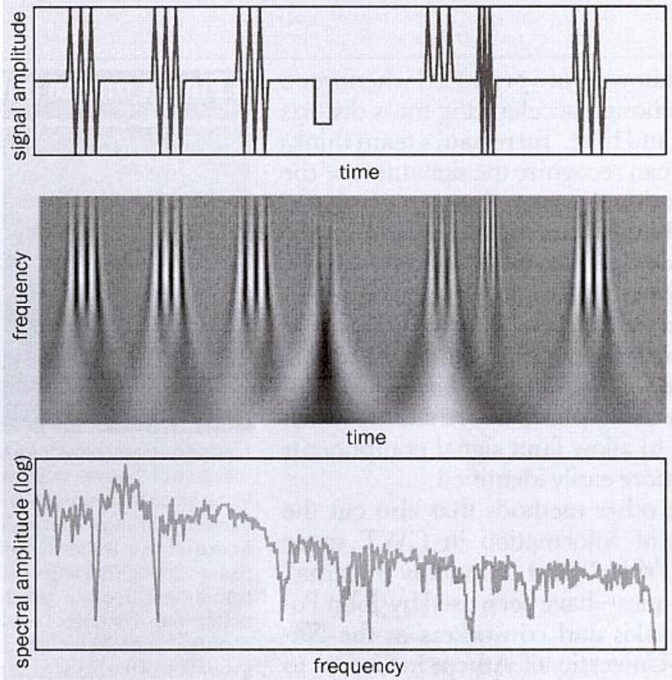
1 The secrets of the little wave





(*a*) Wavelets are used to transform a signal into another representation that presents the information in a more useful form. A few common examples of wavelets are shown here, including the square "Haar" wavelet. (*b*) In the wavelet-analysis technique the wavelet function (red) is "convoluted" with the signal (green). If the wavelet matches the signal well, a large number is obtained. If the wavelet matches the signal poorly, a low number is obtained. The transform is computed at different times in the signal, with wavelets of different frequency used on each occasion. The transform value for each position and frequency of the wavelet function are usually plotted on a 2D plane. (*c*) Stretching and squeezing the wavelet during the transform process changes its frequency make up. The height of the wavelet also changes to conserve its signal energy.

2 Forget Fourier



Wavelet transforms are good at picking out features in a signal that occur only intermittently. This advantage can be seen with a synthetic signal (top), which contains a number of isolated features. It has been transformed using a wavelet that is in the form of a "Mexican hat". The four identical wavegroups all have the same morphology in the wavelet-transform plot (middle), while the remaining features each have their own unique appearance. Although the correlation between features in the signal and those in the transform can be seen visually in this example, statistical techniques have to be used for the messier signals that are more likely to be encountered in real applications. In contrast, the conventional Fourier transform of the original signal (below) provides no useful information about the obvious features in the original signal.

„You have to be careful that you do not fool yourself, and unfortunately, you are the easiest person in the world to fool"

Richard Feynmann