# Spectral Analysis with Unevenly-Spaced Data

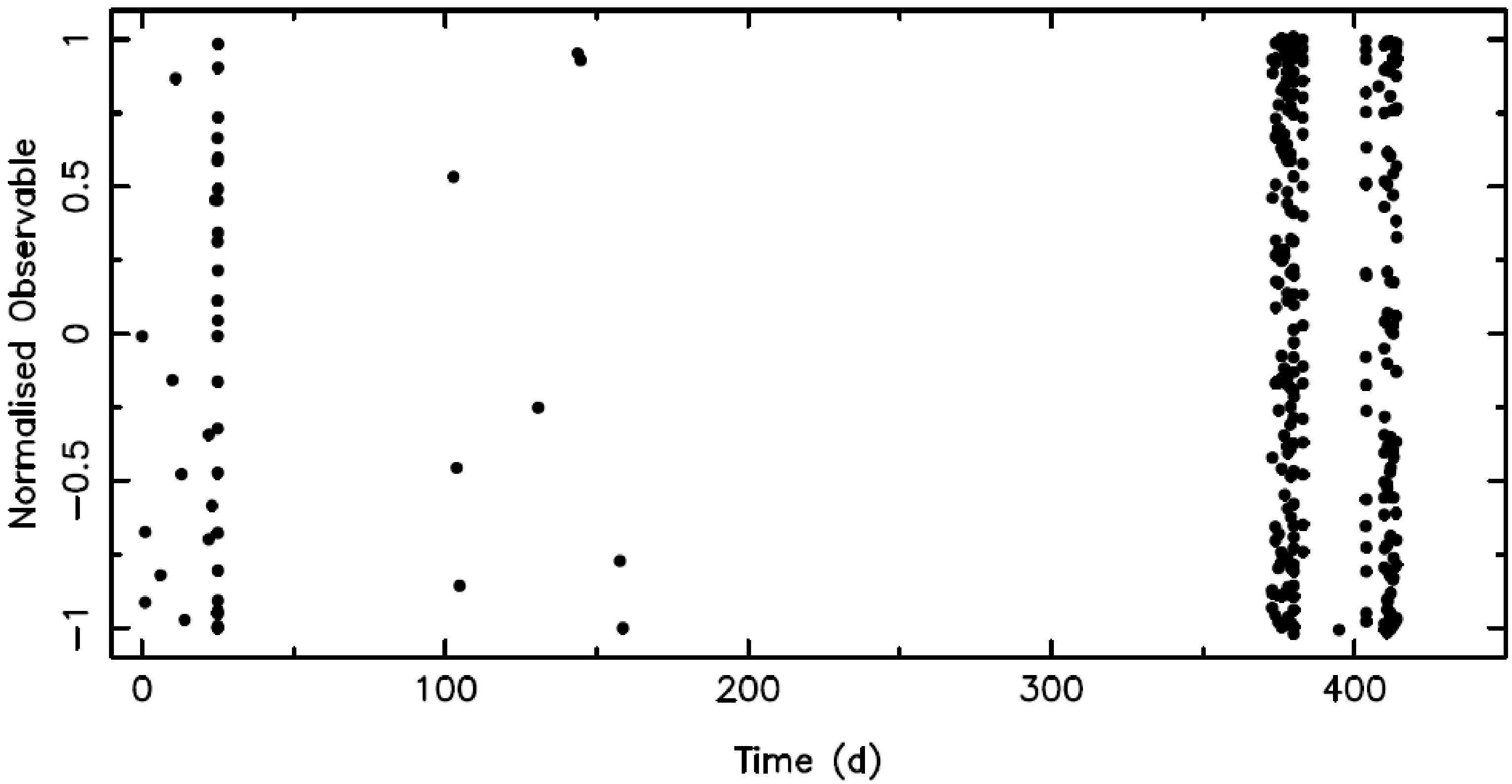


DISCRETE FOURIER TRANSFORM
LOMB-SCARGLE PERIODOGRAM

# Spectral Analysis with Unevenly-Spaced Data



- Gapped data representing a typical time series for a ground-based single-site observational campaign:



# Spectral Analysis with Unevenly-Spaced Data



## FOURIER ANALYSIS WITH UNEQUALLY-SPACED DATA*

T. J. DEEMING

*Dept. of Astronomy, The University of Texas at Austin, Tex., U.S.A.*

(Received 22 March; in revised form 11 November, 1974)

**Abstract.** The general problems of Fourier and spectral analysis are discussed. A discrete Fourier transform $F_N(\nu)$ of a function $f(t)$ is presented which (i) is defined for arbitrary data spacing; (ii) is equal to the convolution of the true Fourier transform of $f(t)$ with a spectral window. It is shown that the 'pathology' of the data spacing, including aliasing and related effects, is all contained in the spectral window, and the properties of the spectral windows are examined for various kinds of data spacing. The results are applicable to power spectrum analysis of stochastic functions as well as to ordinary Fourier analysis of periodic or quasiperiodic functions.

# Discrete Fourier transform



- Time series, $x_k$, $k$=0, … , $N$-1
- **Evenly** spaced data:
  - The discrete Fourier transform decomposes the signal into $N$ sine waves, $a_j$, $j$= -$N$/2+1, … , $N$/2

$$a_j = \sum_{k=0}^{N-1} x_k \, e^{i2\pi jk/N} \qquad j = -\frac{N}{2}+1, \dots, \frac{N}{2}$$

- **Unevenly** spaced data:
  - The discrete Fourier transform decomposes the signal into $M$ sine waves, $a_j$, $j$= 1, … , $M$

$$a_j = \sum_{k=0}^{N-1} x_k \, e^{i2\pi \nu_j t_k} \qquad j = 1, \dots, M$$

$M$ and $v_j$ are now **arbitrary**.

# Fourier Analysis with Unequally-Spaced Data



- PSD is computed as the squared Fourier amplitudes:

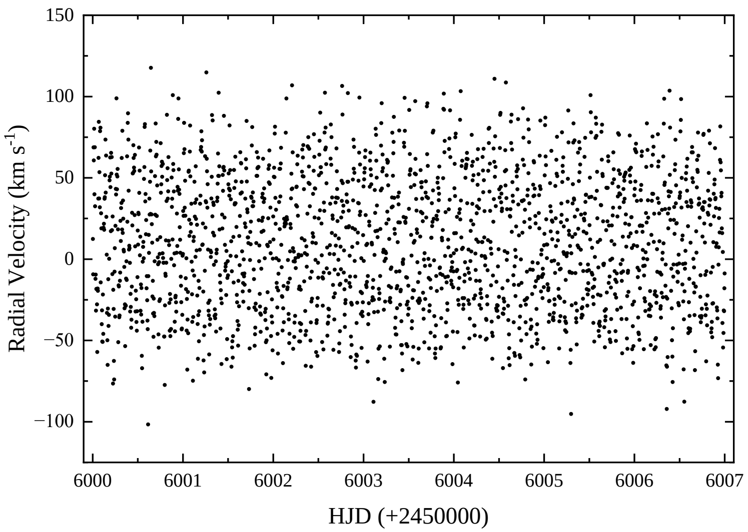
$$P_j = (Normalization)|a_j|^2$$

- **Deeming:**
  **the "pathology" of the data spacing, including aliasing and related effects, is all contained in the spectral window.**

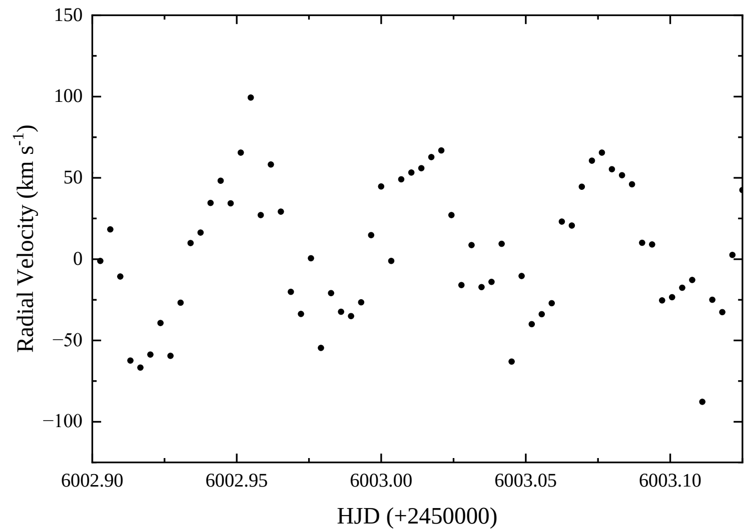
# The variation of spectral window shape



Simulated Radial Velocity data.
Orbital period – 90 min ($\nu$=16.0 day$^{-1}$)

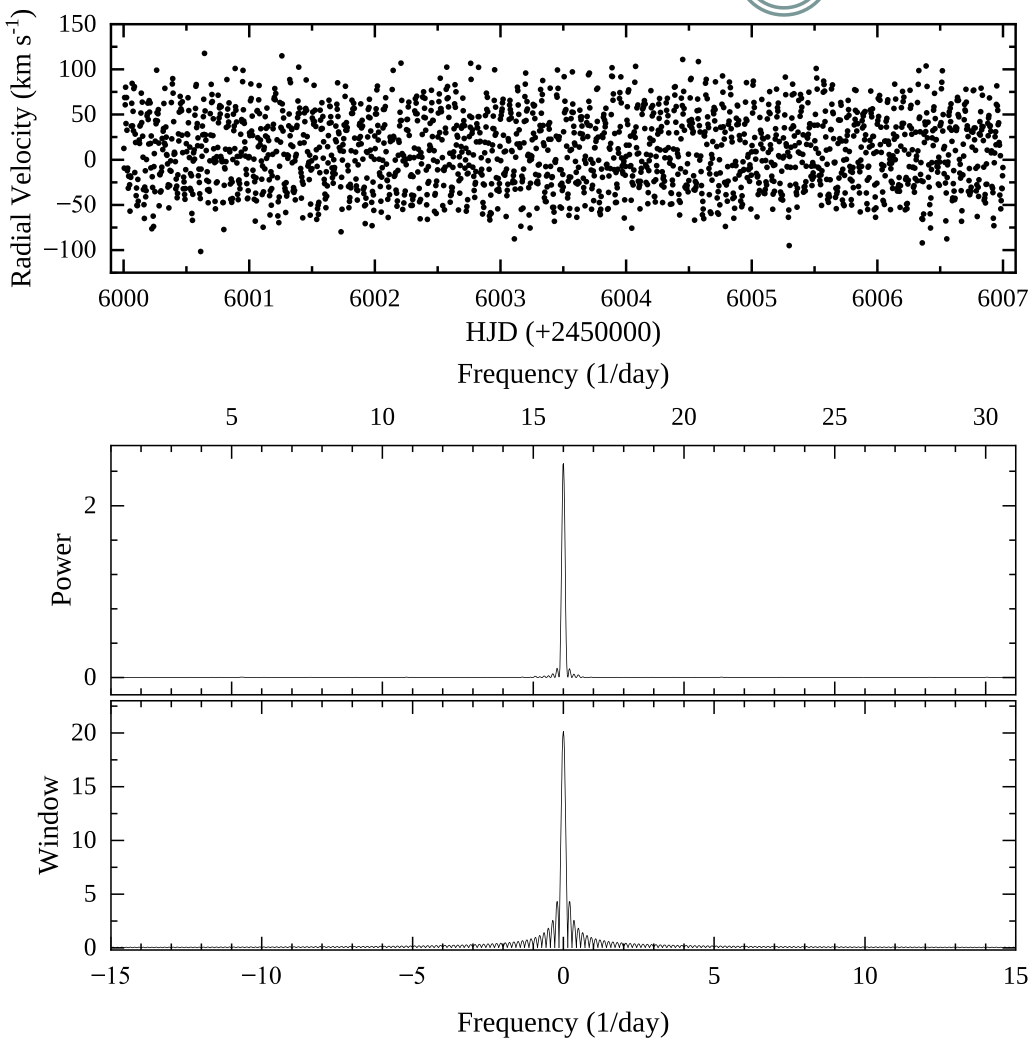



Total duration – 7 nights.
Exposure time – 5 minutes



# The variation of spectral window shape



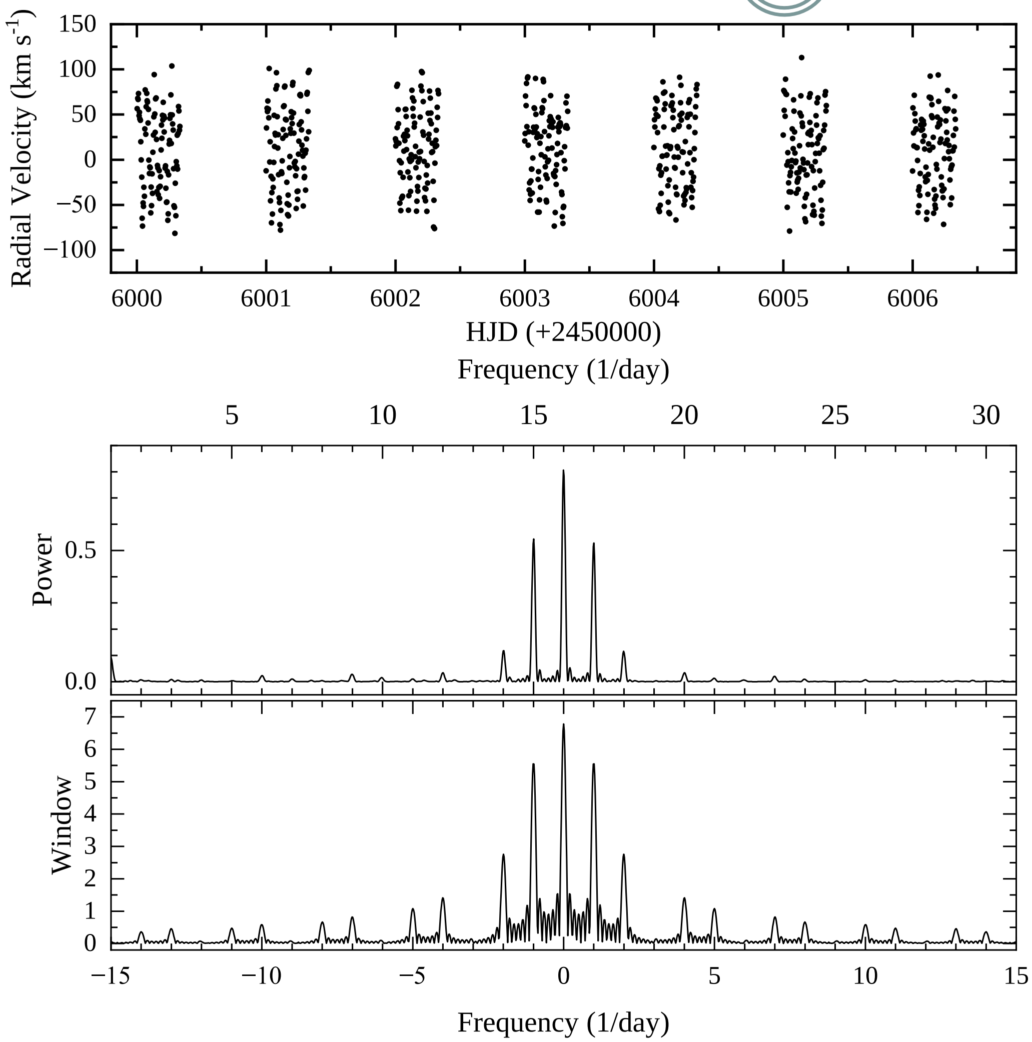



Observations – 7 nights w/o gaps

Exposure times – 5 min.

# The variation of spectral window shape



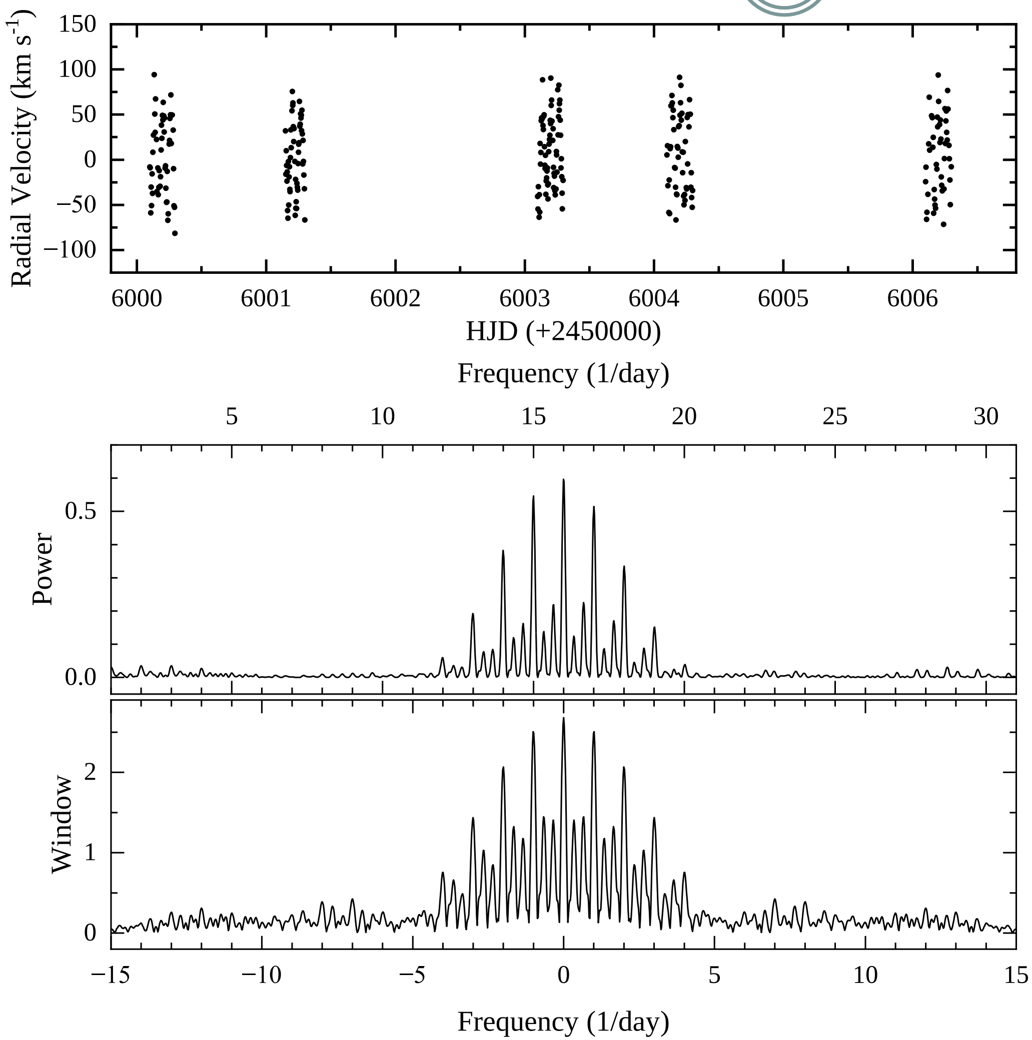



Observations – 8 hours / nights

Exposure times – 5 min.

# The variation of spectral window shape

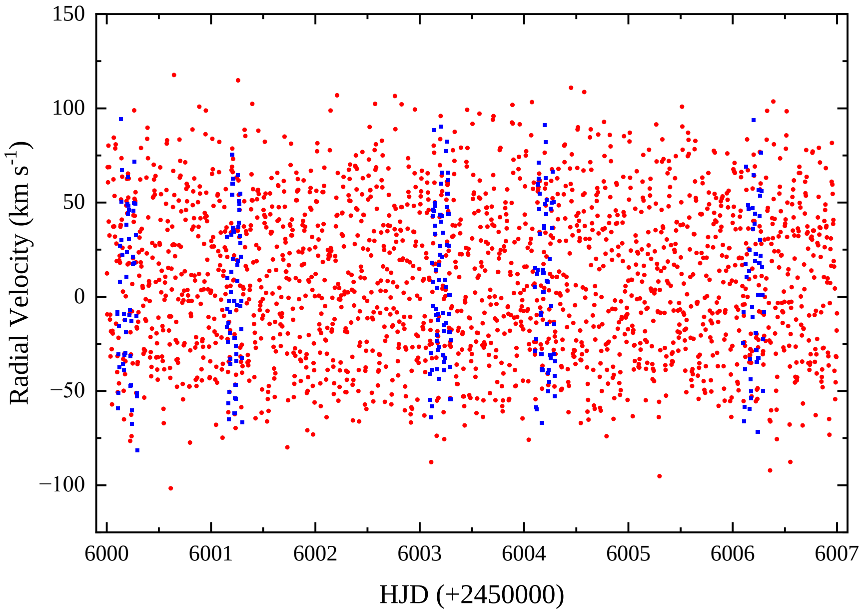


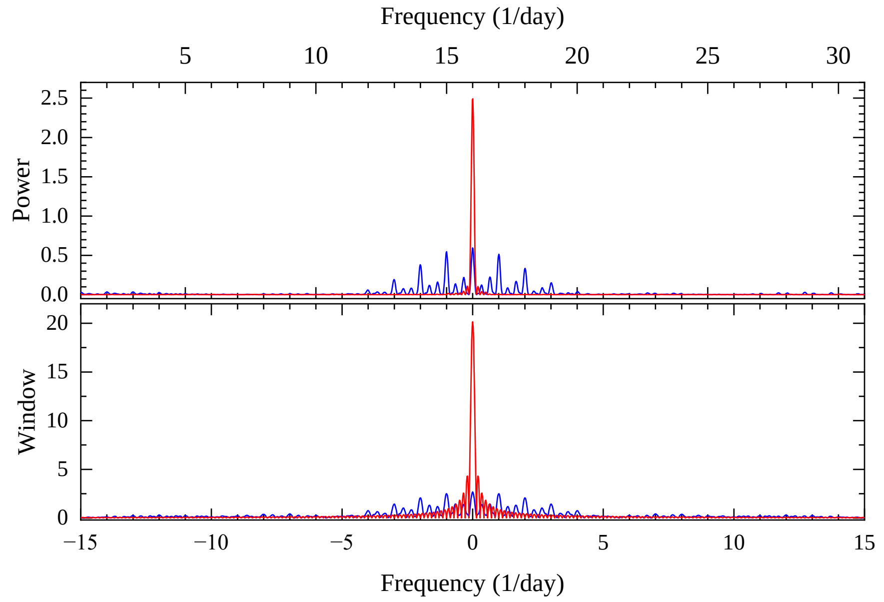



Observations – 5 hours / nights
2 nights excluded
Exposure times – 4 and 5 min.

# The variation of spectral window shape





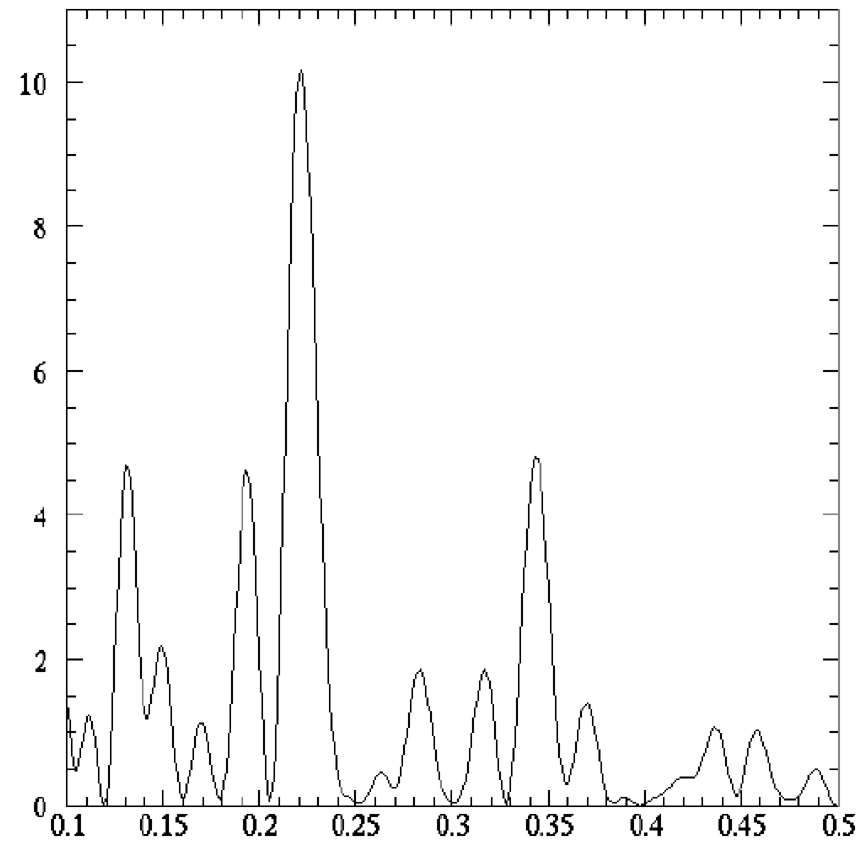



# Fourier Analysis with Unequally-Spaced Data



- How to determine the significance of peaks found in power spectra of unevenly-spaced data?

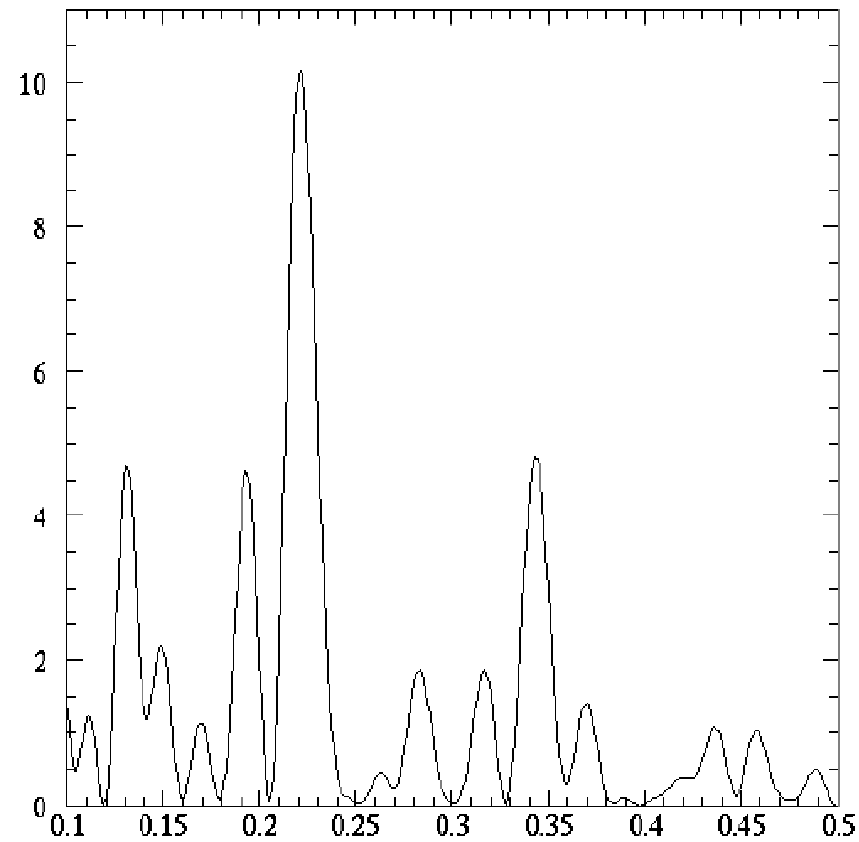
- How many ***independent*** frequencies do we use? In the plot to the right, 1000 values of $\nu_j$ are plotted ($M$=1000), but most frequencies are ***not*** independent!



# Fourier Analysis with Unequally-Spaced Data



- Most frequencies are ***not*** independent!
- Best solution: use Monte Carlo data sets to estimate the probability that the largest peak is bigger than $P_{det}$.
- Second-best rule of thumb: Each peak has a width of $d\omega=2\pi/T$, where T is the length of the entire data set. Here T=365 days, so $d\omega=0.017$. We then estimate that there should be approximately $(0.5-0.1)/0.017 = 23$ independent frequencies over this range.



# Fourier Analysis with Unequally-Spaced Data



- **BE VERY SKEPTICAL OF CLAIMS FOR PERIODICITIES THAT COINCIDE WITH NATURAL FREQUENCIES OF DETECTORS OR OBSERVERS (eg. 1-day, 7-day, 1-year).**

# Fourier Analysis with Unequally-Spaced Data



- The dependence of the PSD on the data length, T, is different for periodic, non-periodic, and stochastic functions:

| | |
|---|---|
| $\lvert a_j \rvert \propto T^0$ | non-periodic |
| $\lvert a_j \rvert \propto T^1$ | periodic |
| $\lvert a_j \rvert \propto T^{1/2}$ | stochastic |

# Lomb-Scargle Periodogram



- Lomb (1976) - Scargle (1982) Periodogram:

$$P_{\mathrm{LS}}(\nu) = \frac{1}{2} \frac{\left\{ \sum_{i=1}^{N} x(t_i) \cos[2\pi\nu(t_i - \tau)] \right\}^2}{\sum_{i=1}^{N} \cos^2[2\pi\nu(t_i - \tau)]} + \frac{\left\{ \sum_{i=1}^{N} x(t_i) \sin[2\pi\nu(t_i - \tau)] \right\}^2}{\sum_{i=1}^{N} \sin^2[2\pi\nu(t_i - \tau)]}.$$

  - Good for general uneven sampling
  - Equivalent to linear least-square fit to sin+cos
  - Statistically robust

# Lomb-Scargle Periodogram



- In this expression, the reference epoch $\tau$ is chosen in such a way that:

$$\sum_{i=1}^{N} \cos[2\pi\nu(t_i - \tau)] \sin[2\pi\nu(t_i - \tau)] = 0,$$

- Or, equivalently

$$\tan(4\pi\nu\tau) = \frac{\sum_{i=1}^{N} \sin(4\pi\nu t_i)}{\sum_{i=1}^{N} \cos(4\pi\nu t_i)}.$$

- It looks complicated, but it's basically the regular periodogram adapted to handle unevenly spaced data. In the limit of equal spacing, it actually reduces to the classical result.

# Lomb-Scargle Periodogram



- One of the reasons to have introduced the Lomb-Scargle periodogram is that its value does not change when all time values $t_i$ are replaced by $(t_i + T)$ because of the definition of $\tau$.

# Properties of the Lomb-Scargle periodogram



- The most important feature of the Lomb-Scargle periodogram is the significance of the power at an individual frequency:

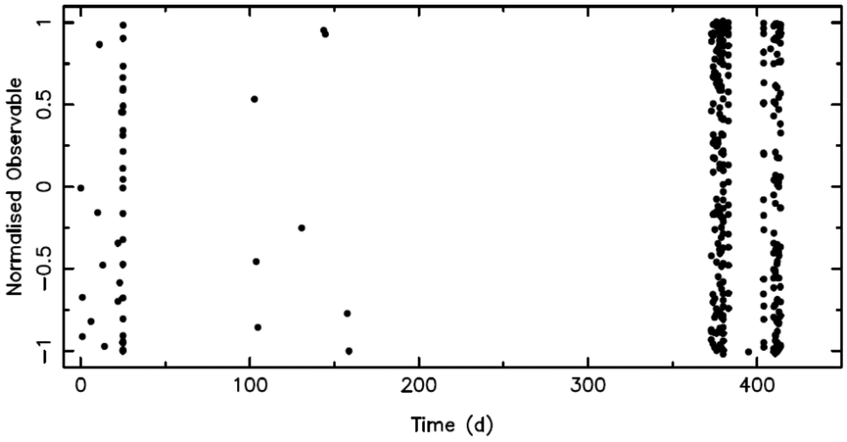
$$\mathrm{Prob}(P(\nu)>P_{det})=\exp(-P_{det})$$

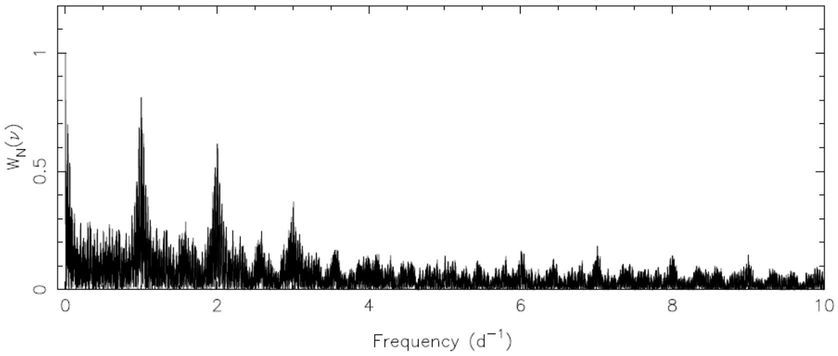
- You still have to worry about the number of independent frequencies you test to account for trials factors

# Spectral Analysis with Unevenly-Spaced Data

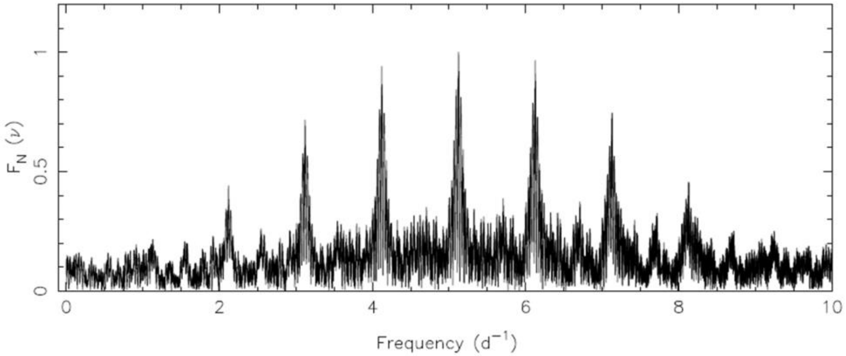


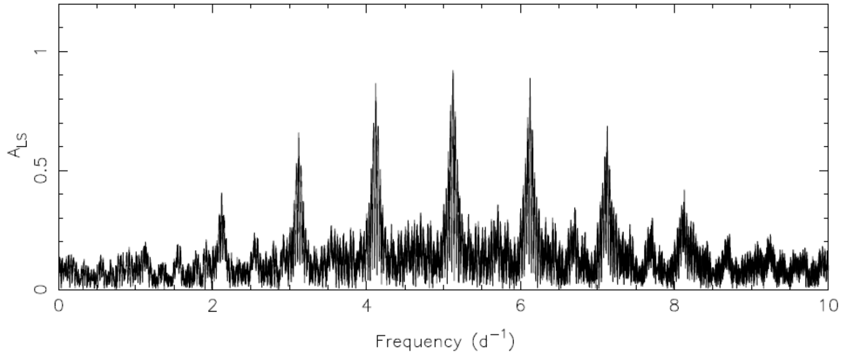
## The gapped data and its Spectral window




## DFT (top) and LS periodogram (bottom)

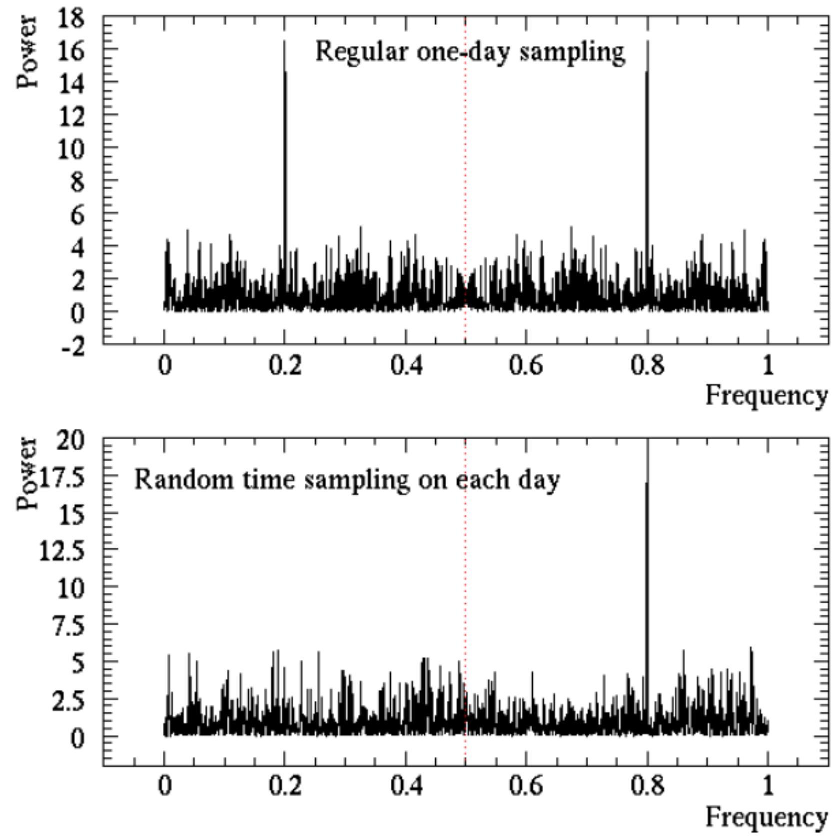


# Advantages of non-uniform sampling



- Consider the following Lomb-Scargle periodograms of an f=0.8 signal.
  Top: sampling exactly once per day at noon
  Bottom: sampling once per day at a random time within the 24-hour period.
- Regular sampling gives strong alias peak at f=0.2: in fact Nyquist frequency is f<0.5, so you'd conclude there was really a signal at f=0.2
- Random sampling gets rid of alias peak! And it gives sensitivity to higher frequencies - since random times can be close to each other, Nyquist cutoff is not a hard limit anymore!



# Combined analysis of power spectra



- In astronomy, it is often necessary to compare the power spectra of two or more time series, e.g.:
    - X-ray binaries are often observed simultaneously in X-rays, UV and optical wavelengths (and γ-rays).
    - the Sun has been observed more or less routinely for many years and in a variety of modes (sunspots, radio, UV, X-ray, irradiance, etc.), so one may need to compare two or more solar data sets.
- One might also wish to estimate the significance of a particular peak that shows up in two or more power spectra.

# Combined analysis of power spectra



- Assuming, that each power spectrum is distributed **exponentially** (e.g., Lomb-Scargle periodogram), Sturrock et al. (2005) proposed three such statistics, that are useful for the combined study of two or more time-series:
  - Combined Power Statistic
  - Minimum Power Statistic
  - Joint Power Statistic

**The paper is on the course web-page**

# Combined analysis of power spectra



- **Combined Power Statistic**
  - If we wish to combine information from $n$ independent power spectra, the combination that would correspond to the chi-square statistic is the sum of the powers, which we write as

$$\boldsymbol{Z} = \boldsymbol{S_1} + \boldsymbol{S_2} + \cdots + \boldsymbol{S_n}.$$

  - The following function of $Z$ ("**combined power statistic**") is distributed exponentially:

$$G_n(Z) = Z - \ln\left(1 + Z + \frac{1}{2}Z^2 + \cdots + \frac{1}{(n-1)!}Z^{n-1}\right).$$

# Combined analysis of power spectra



- **Minimum Power Statistic**
  - We may wish to determine the frequency for which the minimum power among two or more power spectra has the maximum value. Let's consider the following quantity, formed from the independent variables $x_1, x_2, \ldots, x_n$, each of which is distributed exponentially:

$$\boldsymbol{U(x_1, x_2, \ldots, x_n) = \mathrm{Min}(x_1, x_2, \ldots, x_n)}$$

  - It can be shown that the following function of $U$ ("**minimum power statistic**") is distributed exponentially:

$$\boldsymbol{K_n(U) = nU}$$

# Combined analysis of power spectra



- **Joint Power Statistic**

  - Let’s now consider the need to compare spectra from two quite different times series. If one of the time-series has very strong peaks and the other has comparatively weak peaks, then simply adding the powers would not be very revealing, since the sum would be dominated by the stronger spectrum.

  - In this situation, it is more useful to form something resembling a “correlation function" by forming the product of the two powers. It proves convenient to work with the square root of the product (the geometric mean):

$$\boldsymbol{X = (S_1 S_2)^{1/2}}$$

# Combined analysis of power spectra



- **Joint Power Statistic (cont.)**
  - The following function of $X$ is distributed exponentially:

$$J_2 = -\ln\left(2X\, K_1(2X)\right)$$

where $K1$ is the Bessel function of the second kind.

  - A good approximation to $J_2$ is found to be:

$$J_{2A} = \frac{1.943\,X^2}{0.650 + X}.$$

# Combined analysis of power spectra



- **Joint Power Statistic (cont.)**

  - Let's now consider joint power statistics of higher orders, and consider the following geometric mean of $n$ powers:

$$\boldsymbol{X} = (\mathbf{S_1} \ldots \mathbf{S_n})^{1/n}$$

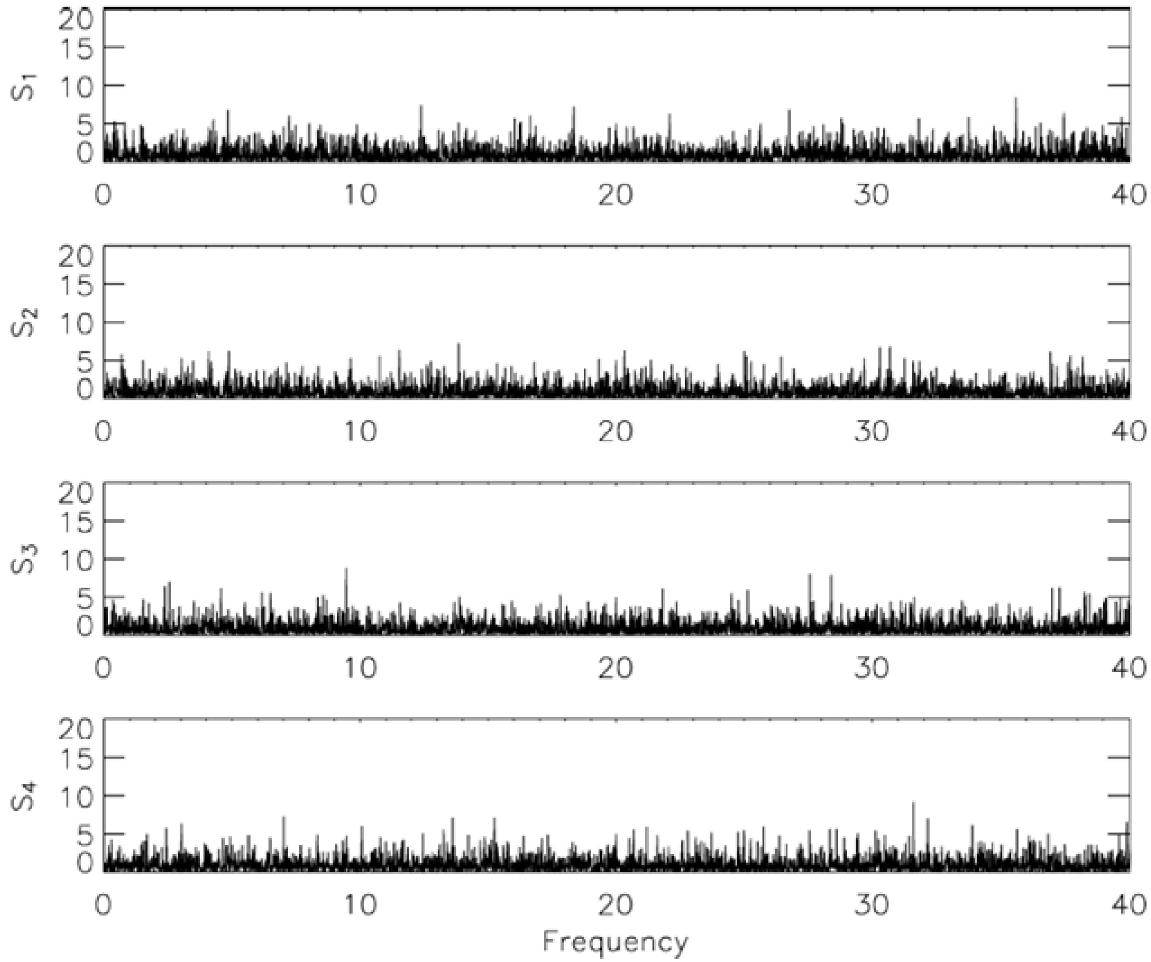
  - There is no useful analytical functions of $X$ that are distributed exponentially, but there are very good approximations:

$$J_{3A} = \frac{2.916X^2}{1.022 + X}, \qquad J_{4A} = \frac{3.881X^2}{1.269 + X}.$$

# Combined analysis of power spectra



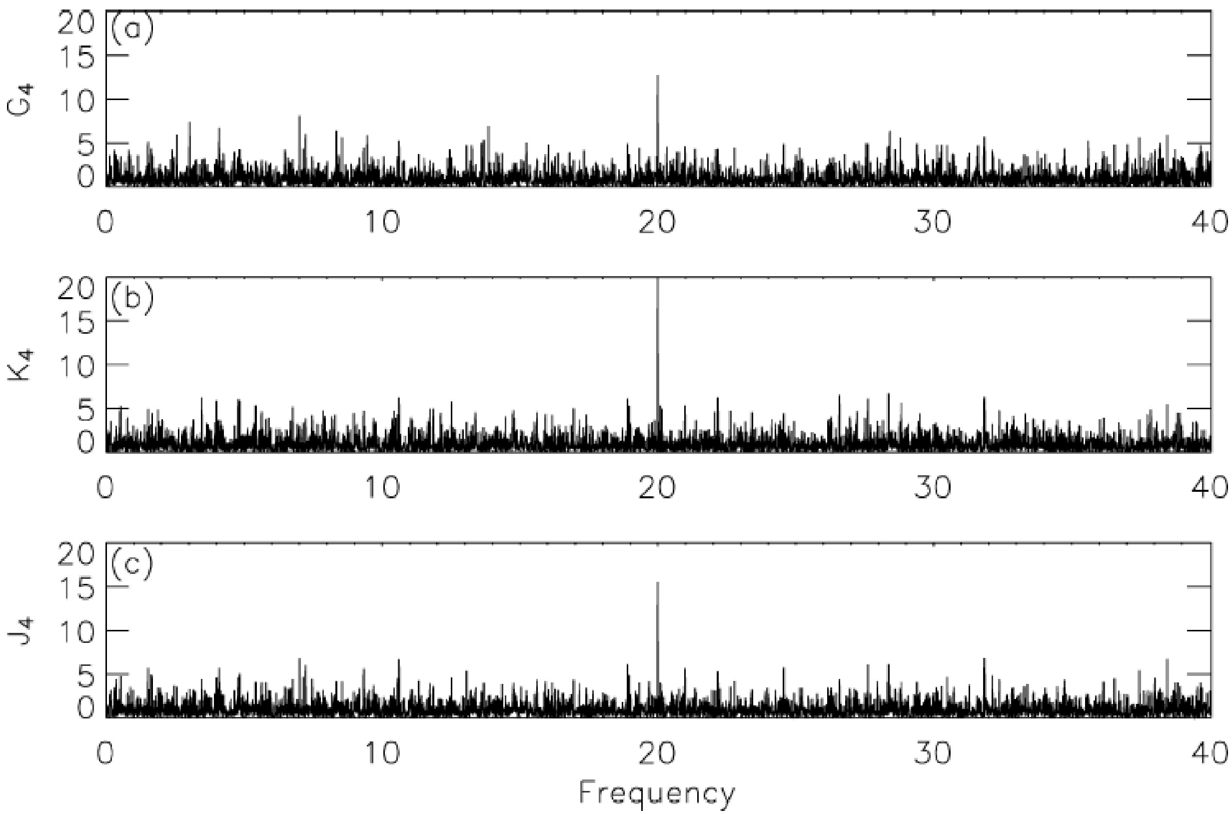
Four synthetic spectra, each with a signal of power 5 at $\nu$ = 20.



# Combined analysis of power spectra



The combined power statistic, minimum power statistic, and joint power statistic, formed from the four synthetic spectra.



# DFT vs Lomb-Scargle



## A critical comparison of the Lomb-Scargle and the classical periodograms

arXiv:1807.01595

Roberto Vio,[1]* P. Andreani,[2]
[1]*Chip Computers Consulting s.r.l., Viale Don L. Sturzo 82, S.Liberale di Marcon, 30020 Venice, Italy*
[2]*ESO, Karl Schwarzschild strasse 2, 85748 Garching, Germany*

Accepted XXX. Received YYY; in original form ZZZ

**ABSTRACT**
The detection of signals hidden in noise is one of the oldest and common problems in astronomy. Various solutions have been proposed in the past such as the parametric approaches based on the least-squares fit of theoretical templates or the non-parametric techniques as the phase-folding method. Most of them, however, are suited only for signals with specific time evolution. For generic signals the spectral approach based on the periodogram is potentially the most effective. In astronomy the main problem in working with the periodogram is that often the sampling of the signals is irregular. This complicates its efficient computation (the fast Fourier transform cannot be directly used) but overall the determination of its statistical characteristics. The Lomb-Scargle periodogram (LSP) provides a solution to this last important issue, but its main drawback is the assumption of a very specific model of the data which is not correct for most of the practical applications. These issues are not always considered in literature with theoretical and practical consequences of no easy solution. Moreover, apart from pathological samplings, it is common believe that the LSP and the classical periodogram (CP) usually provide almost identical results. In general, this is true but here it is shown that there are situations where the LSP is less effective than the CP in the detection of signals in noise. There are no compelling reasons, therefore, to use the LSP instead of the CP which is directly connected to the correlation function of the observed signal with the sinusoidal functions at the various frequencies of interest.

**Key words:** Methods: Statistical – Methods: Data Analysis – Methods: Numerical

# DFT vs Lomb-Scargle



**A critical comparison of the Lomb-Scargle and the classical periodograms** by Vio & Andreani

arXiv:1807.01595

"… apart from pathological samplings, it is common believe that the LSP and the classical periodogram (CP) usually provide almost identical results. **In general, this is true** but here it is shown that there are situations where the LSP is less effective than the CP in the detection of signals in noise. There are no compelling reasons, therefore, to use the LSP instead of the CP …"

However (it is my remark):

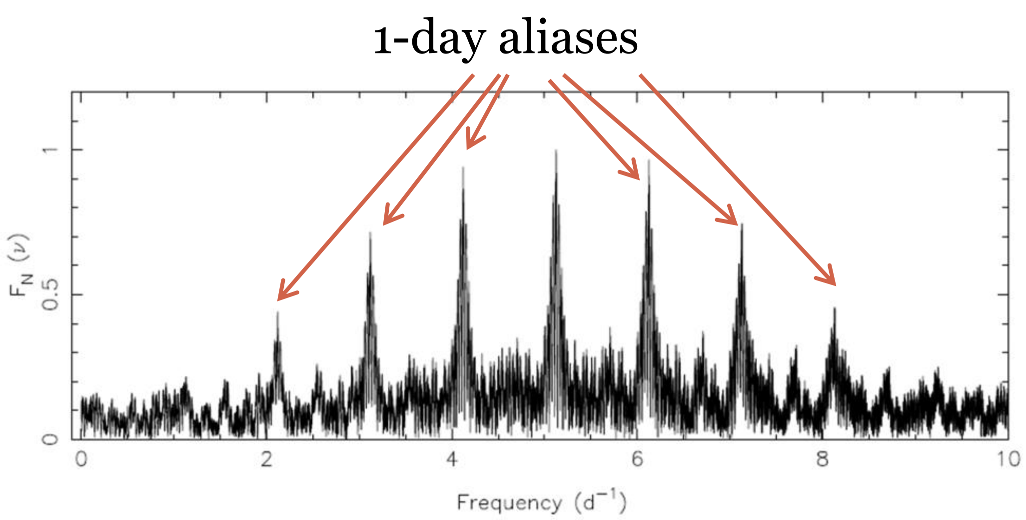
There are situations where the LSP is more effective than the CP.
Thus, there are no compelling reasons to use the DFT instead of the LSP.

**Conclusion: both methods provide almost identical results.**
**However, the LSP is statistically robust.**

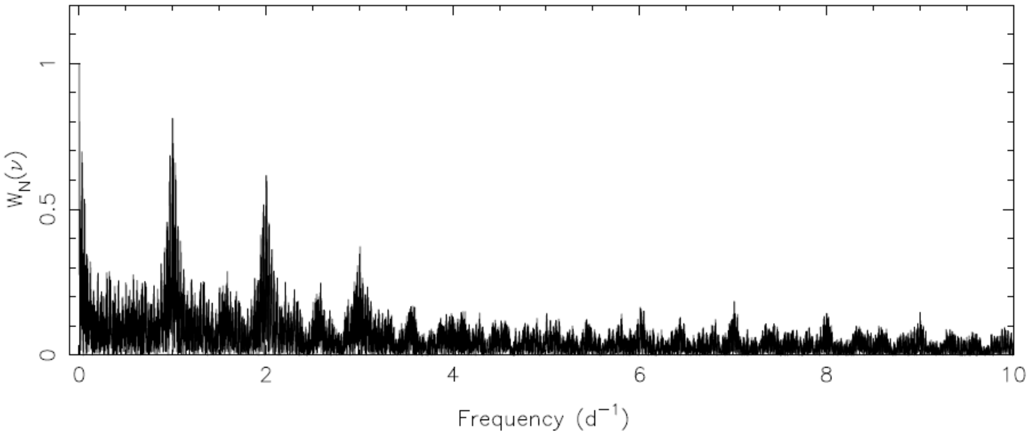
# Dealing with Aliases. CLEAN algorithm.



Power Spectrum:




Spectral Window:



# Dealing with Aliases

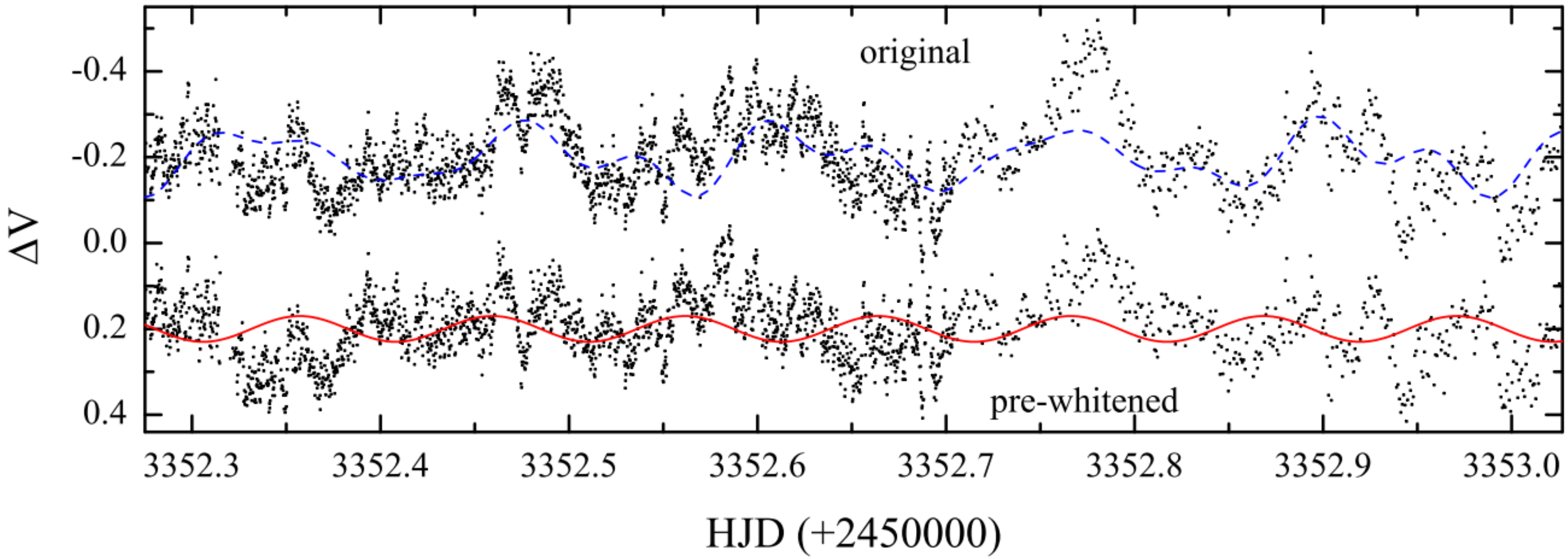


- How to deal with aliases?
  - Pre-whitening
  - Cleaning (Clean algorithm)

# Pre-whitening



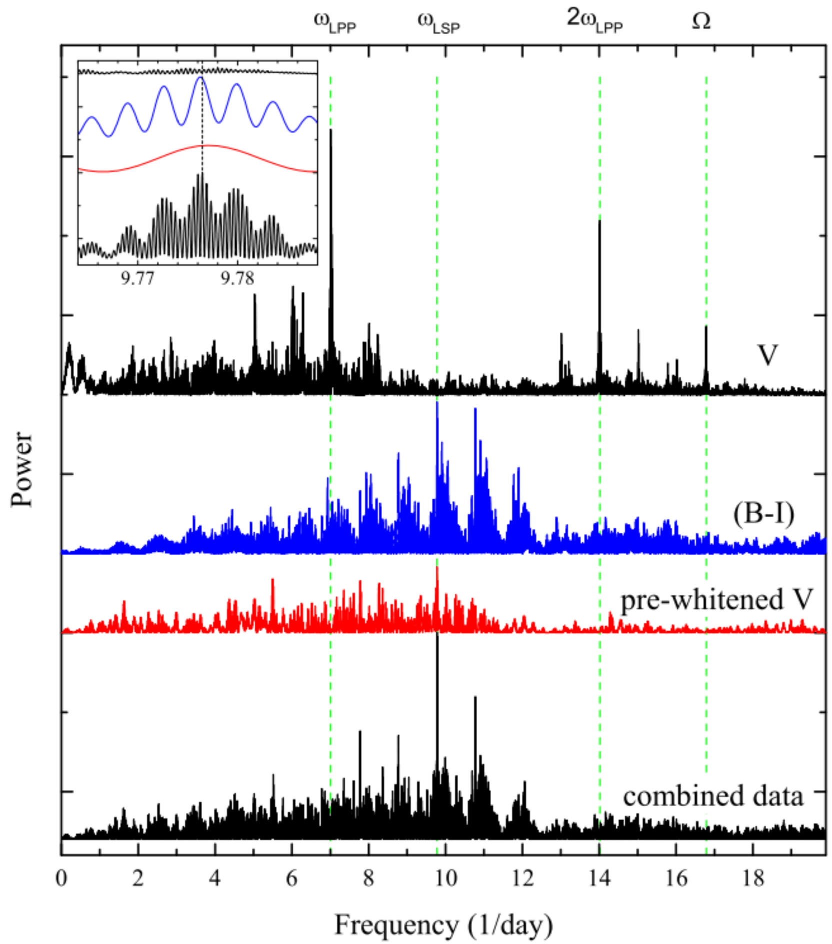
- If the light curve contains more than one periodic modulations, and the signal of interest with an unknown frequency is weak and hidden in noise, then one can try to remove the strongest signal of known frequency from the light curve:
  fit the light curve with a sine-wave (and its harmonics) and subtract it.



# Pre-whitening



- If the light curve contains more than one periodic modulations, and the signal of interest with an unknown frequency is weak and hidden in noise, then one can try to remove the strongest signal of known frequency from the light curve.



# Cleaning



THE ASTRONOMICAL JOURNAL VOLUME 93, NUMBER 4 APRIL 1987

## TIME SERIES ANALYSIS WITH CLEAN. I. DERIVATION OF A SPECTRUM

DAVID H. ROBERTS
Physics Department, Brandeis University, Waltham, Massachusetts 02254

JOSEPH LEHÁR AND JOHN W. DREHER
Physics Department, 26-315, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

*Received 7 October 1986; revised 25 November 1986*

### ABSTRACT

We present a method of time-series spectral analysis which is especially useful for unequally spaced data. Based on a complex, one-dimensional version of the CLEAN deconvolution algorithm widely used in two-dimensional image reconstruction, this technique provides a simple way to understand and remove the artifacts introduced by missing data. We describe the method, give several examples, and point out various analogies with the conventional use of CLEAN.

# Clean Algorithm



- The premise of CLEAN is that our data consist not only of the data amplitudes but also the detailed sampling in time.
- We therefore know that the true spectrum is convolved with a known window function.
- The actual algorithm is based on the fact that any function can be represented as a sum or integral over delta functions.

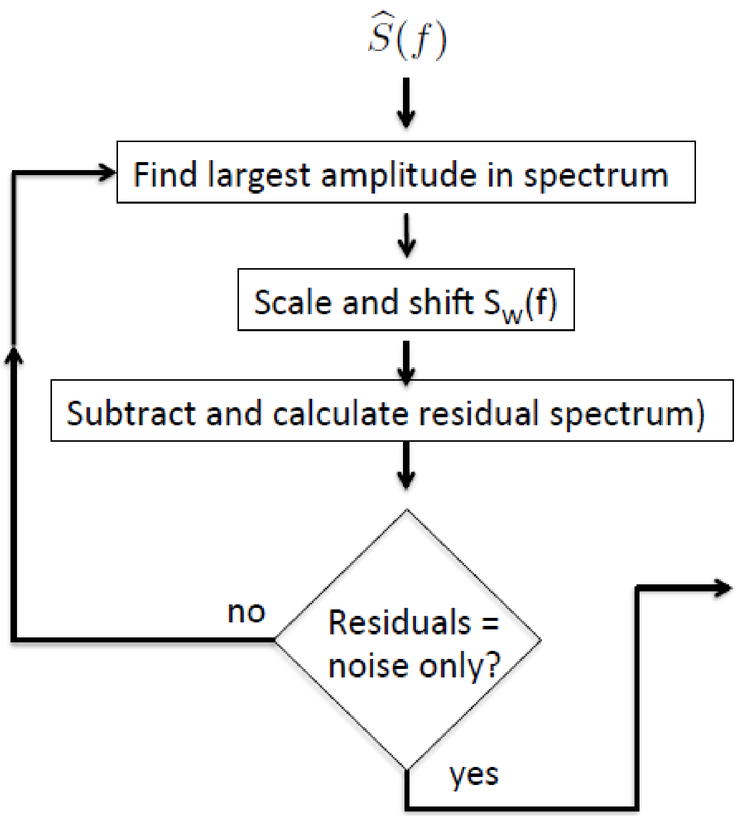
| Spectral Analysis |
| --- |
| $S(f)$ = spectral estimate |
| $S_W(f) = \lvert FT\{W(t)\}\rvert^2$ = spectral window (calculated as the FT of sample times) |
| $S_C(f) = \Sigma\, C_j\, \Delta(f - f_j)$ CLEANed Spectrum Sum over CLEAN components $\Delta(f)$ = restoration function that represents the inherent frequency resolution |

The restoring function is needed to fairly represent the resolution imposed by Fourier transform properties (uncertainty principle)

# Clean Algorithm



Data yield $\widehat{S}(f)$ and $S_W(f) = |\tilde{W}(f)|^2$




$\widehat{S}_i(f_i)_{\max}\ @ f_i \quad i = \text{iteration count}$

$\gamma_i S_W(f - f_i)$

$\widehat{S}_i(f) \rightarrow \widehat{S}_{i+1}(f) = \widehat{S}_i(f) - \gamma_i S_W(f - f_i)$

Collect CLEAN components

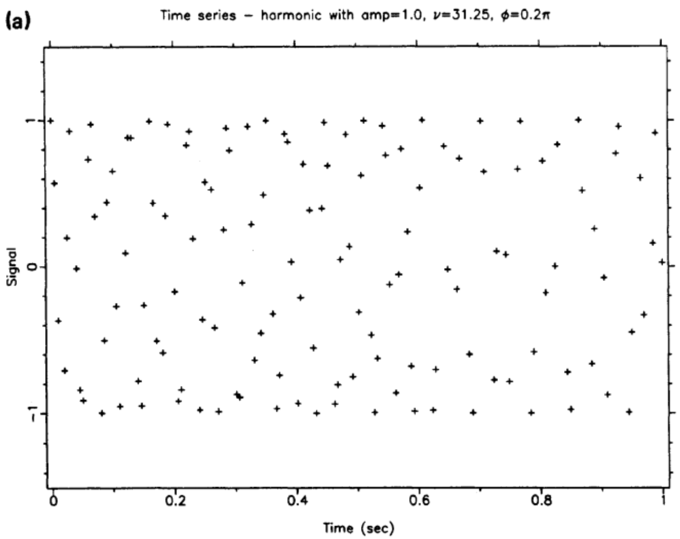
$\{\gamma_i, f_i,\ i = 1, N_c\}$

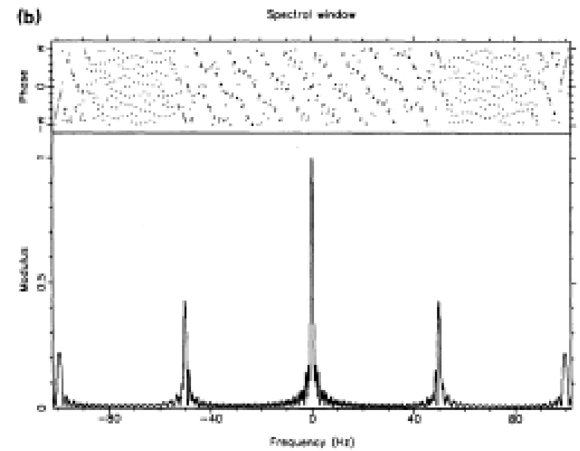
Form CLEAN spectrum

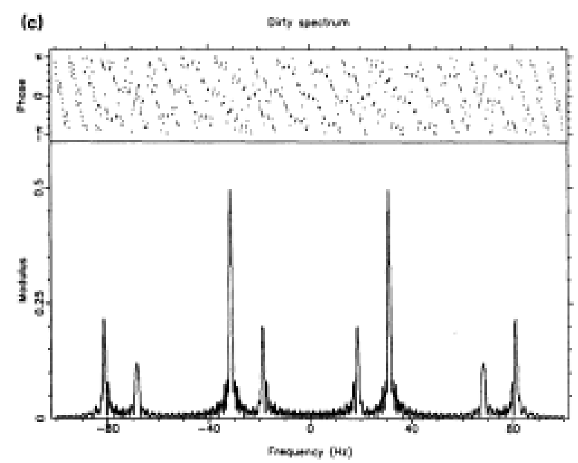
$$S_c(f) = \sum_{i=1}^{N_c} \gamma_i \Delta(f - f_i)$$

# Clean Algorithm






a) Time-Series
b) The window function
c) The dirty spectrum

# Clean Algorithm



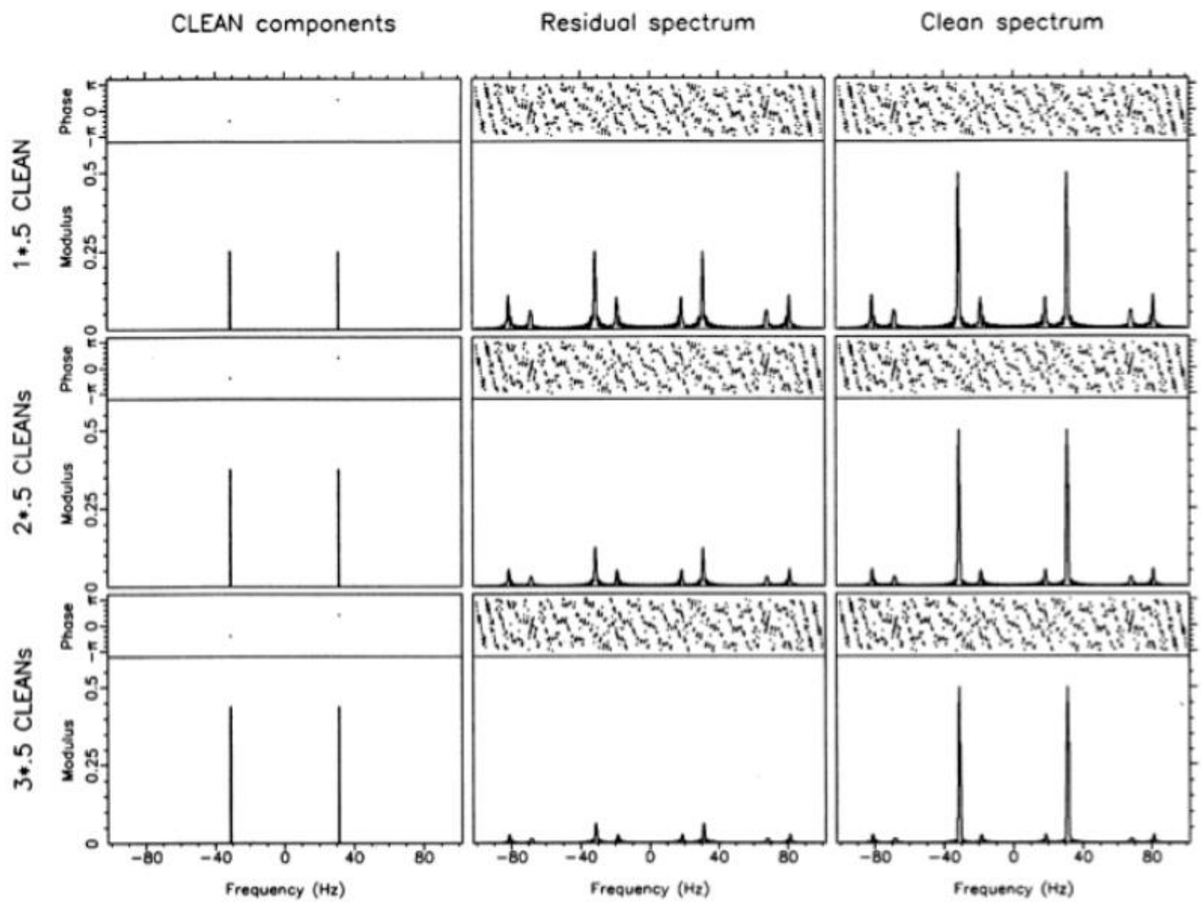



FIG. 2. Analysis of the time series in Fig. 1. (a)–(e). The clean components, residual spectra, and clean spectra after one, two, three, five and one hundred iterations with gain $g = 0.5$, and (f) after one iteration with $g = 1$. Note the change to a logarithmic scale for (d)–(f).

# Clean Algorithm



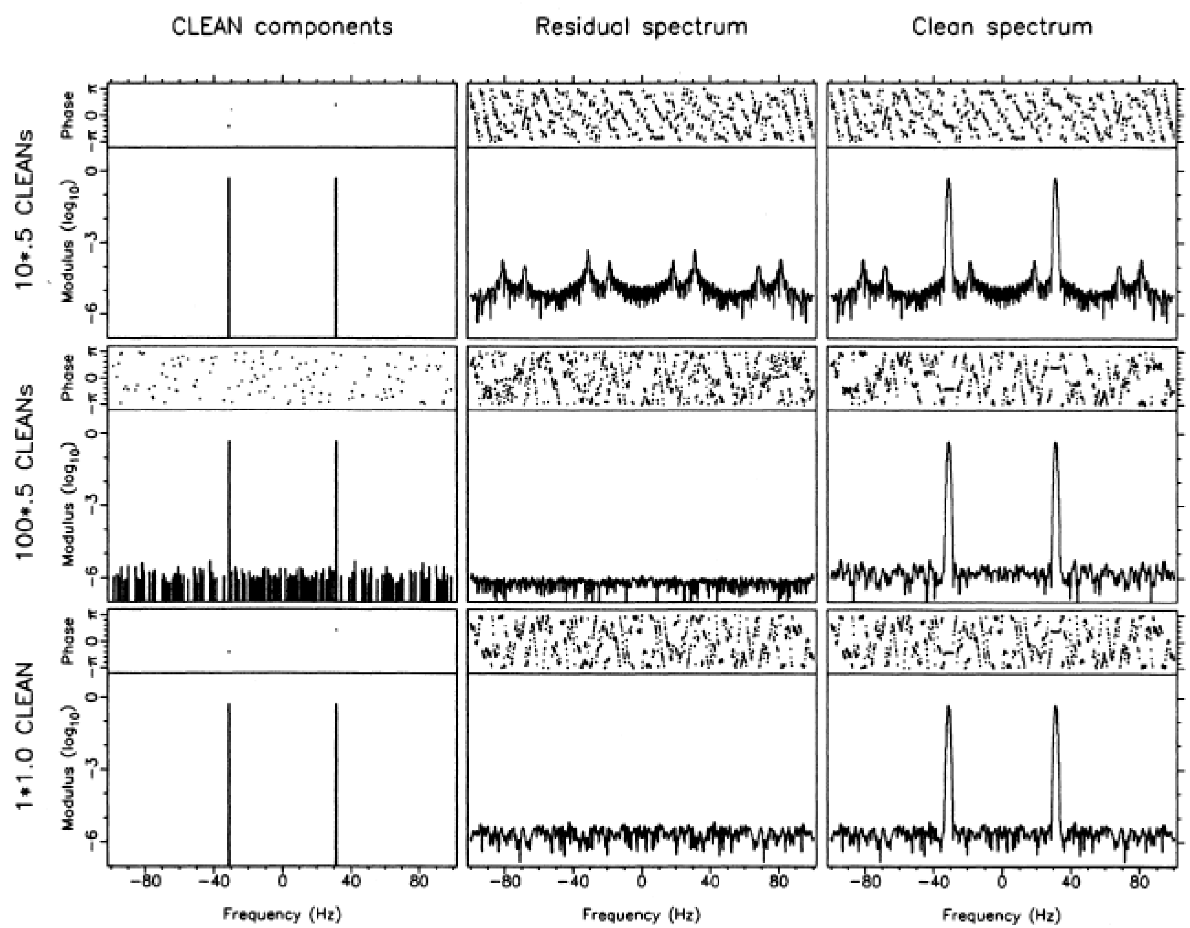



FIG. 2. (continued)