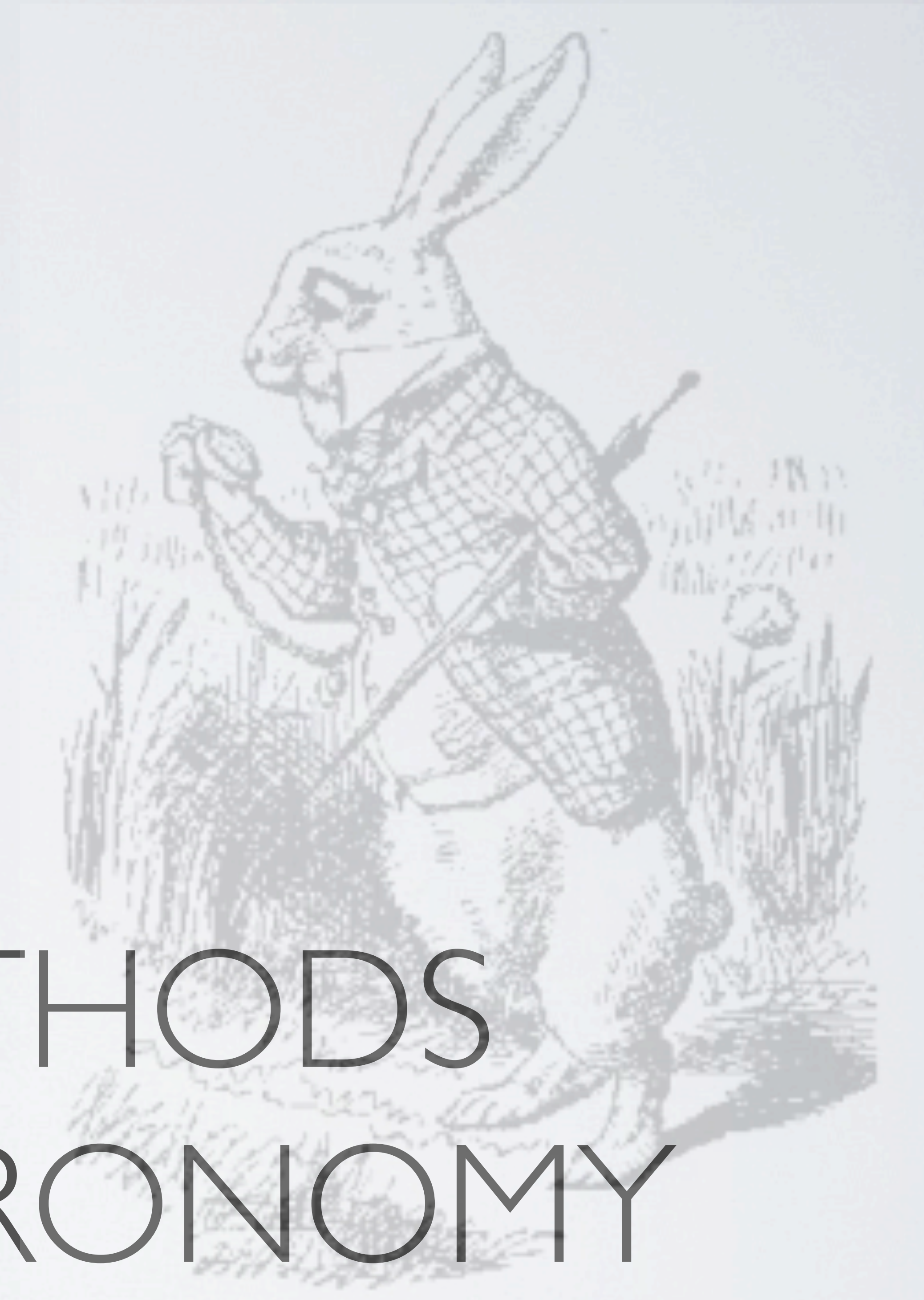


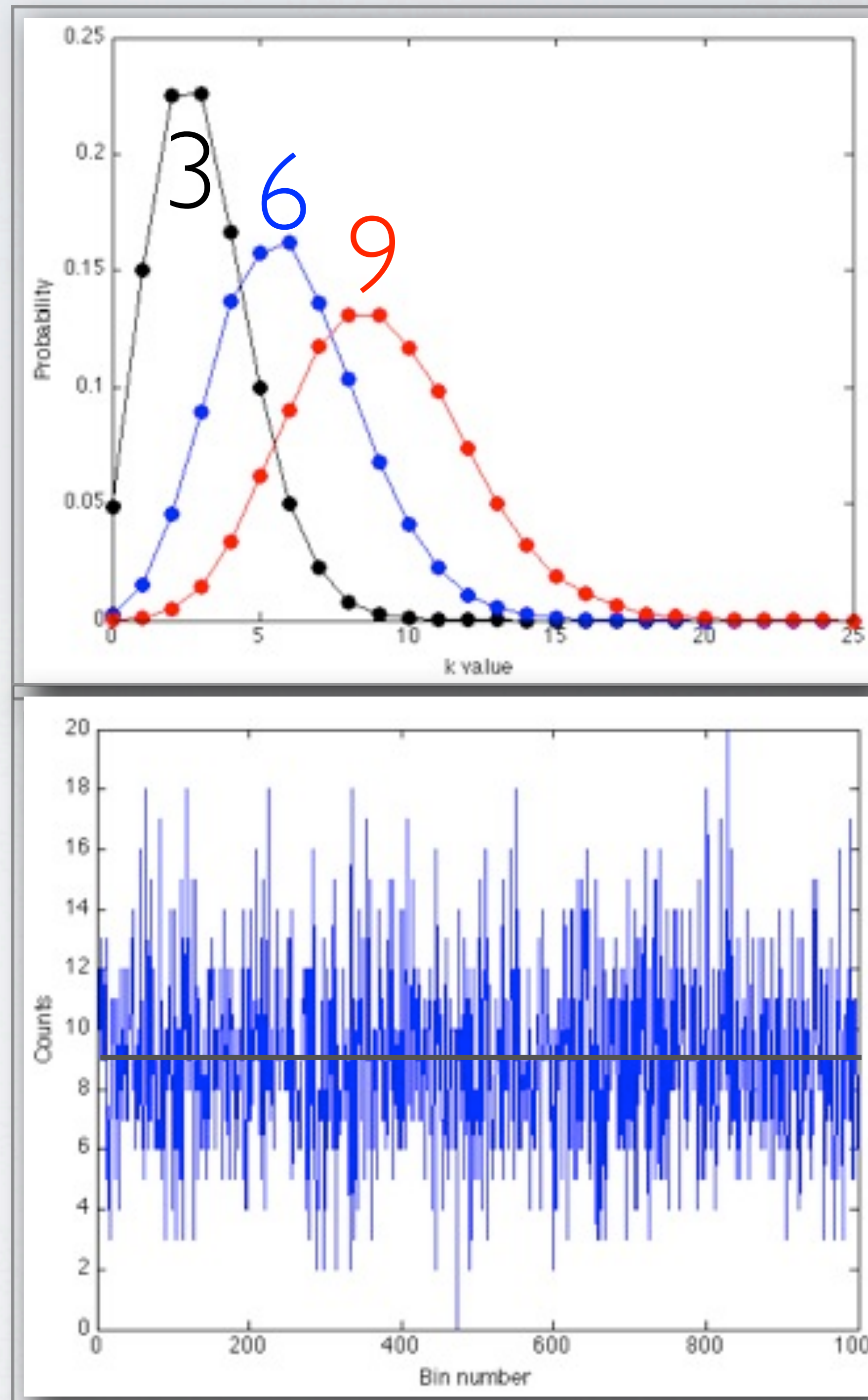
TIMING METHODS IN X-RAY ASTRONOMY

Tomaso Belloni (INAF - Osservatorio Astronomico di Brera)

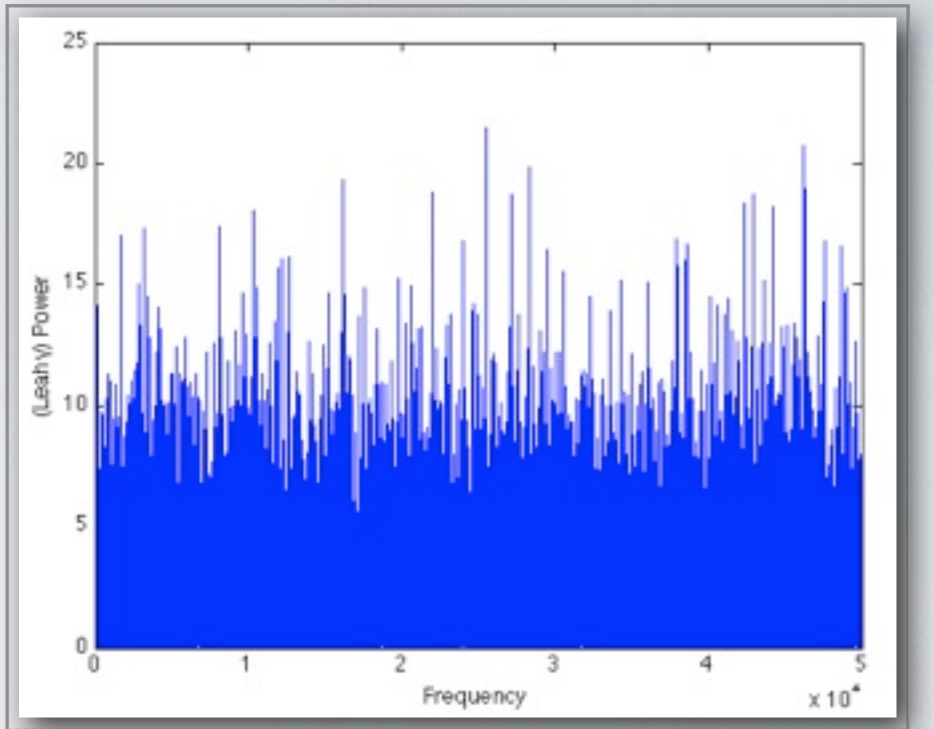
POISSON NOISE EFFECTS

$$p(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

- Counting detector
- Counting noise
- Background negligible
- Independent arrival times
- Exponential waiting time between photons



POWER SPECTRUM NORMALIZATION



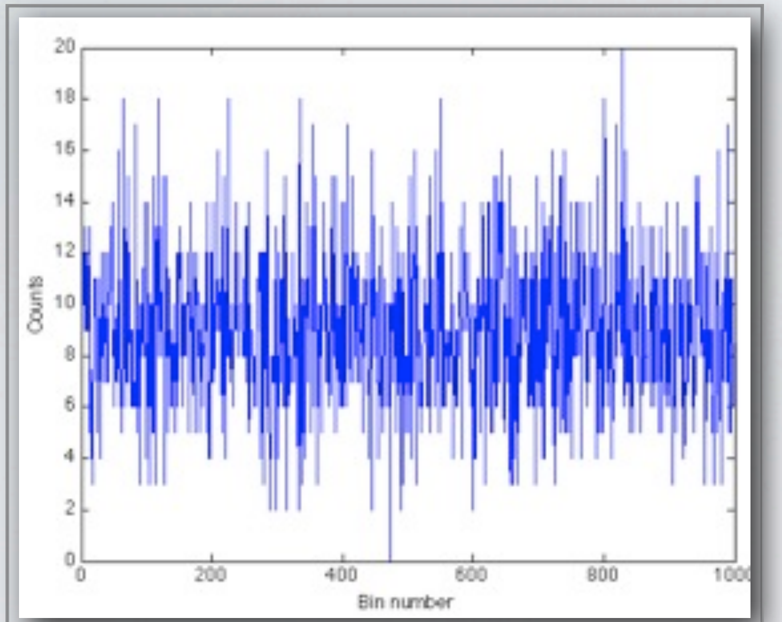
- With this choice, noise power a χ^2 with 2 d.o.f.

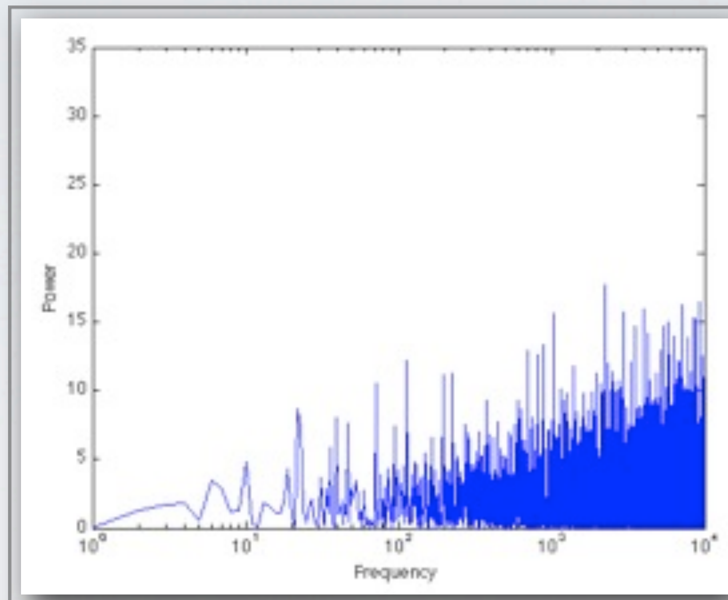
$$P = \frac{2}{N_{phot}} |a|^2$$

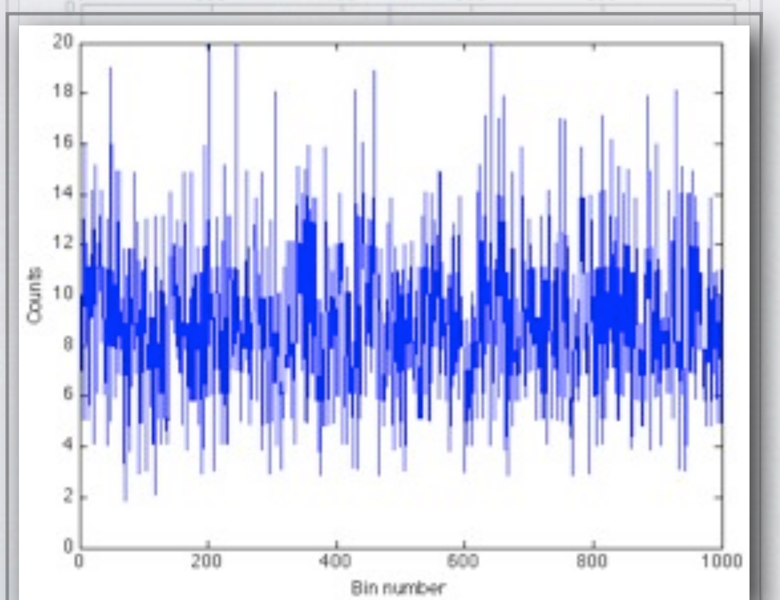
Leahy Norm.

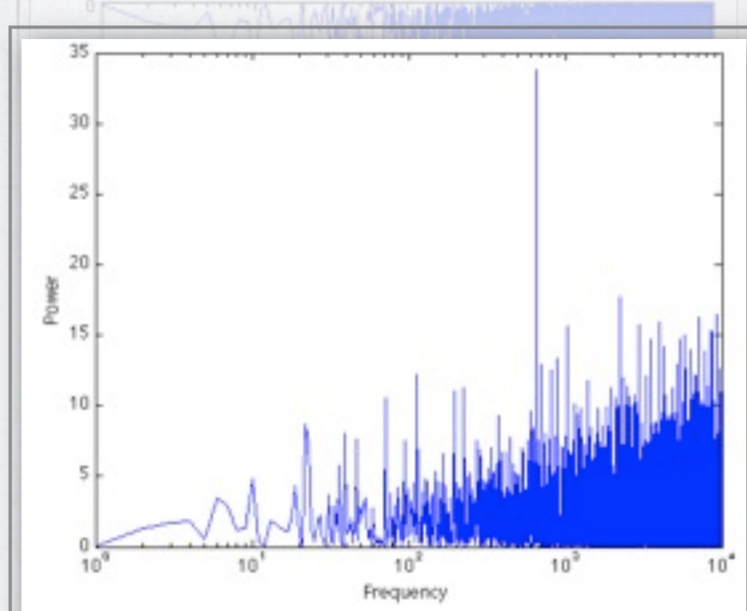
- Most noises do
- Average power is 2. I can calculate statistics
- Noise & signal independent: $P_{total} = P_{signal} + P_{noise}$
- Not always so... (count rate!)
- More complex: deadtime

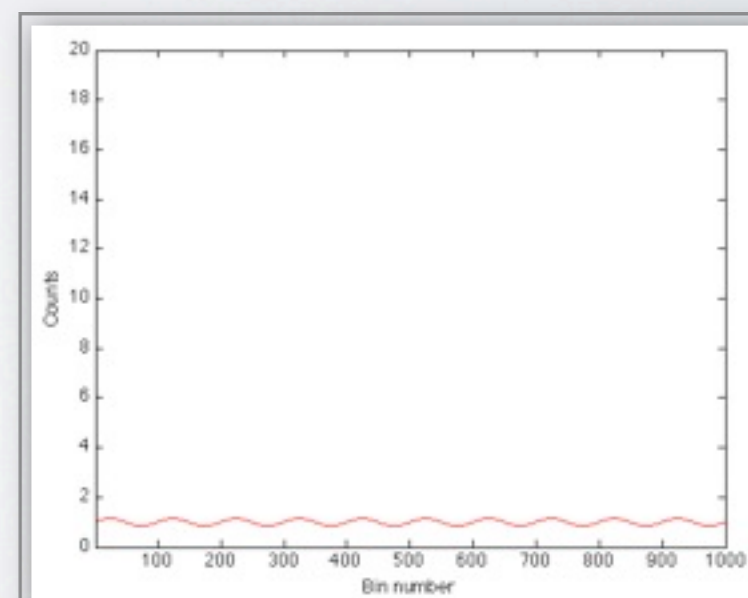
POWER OF POWER SPECTRUM





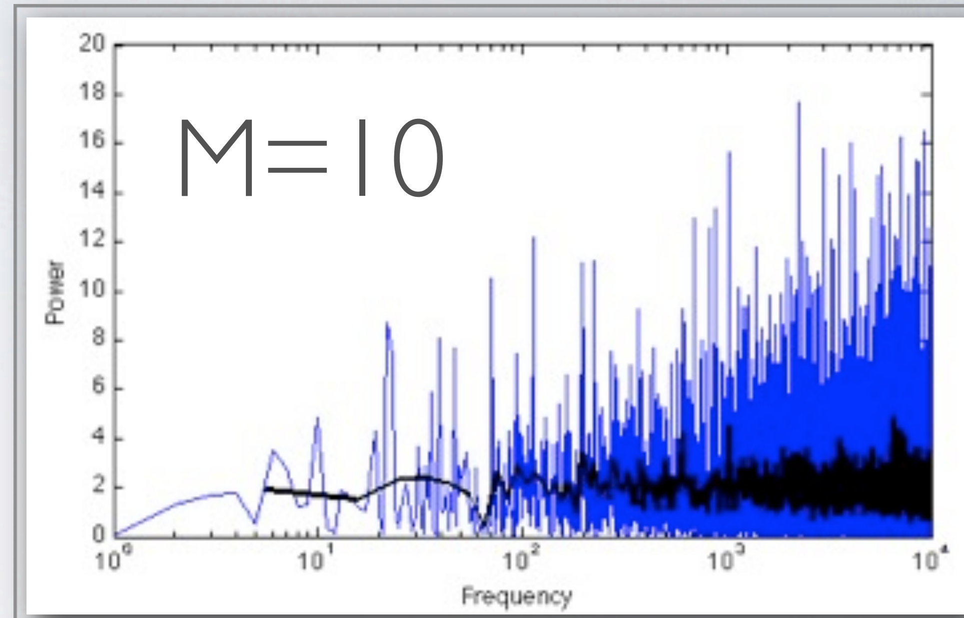






NOISY NOISE

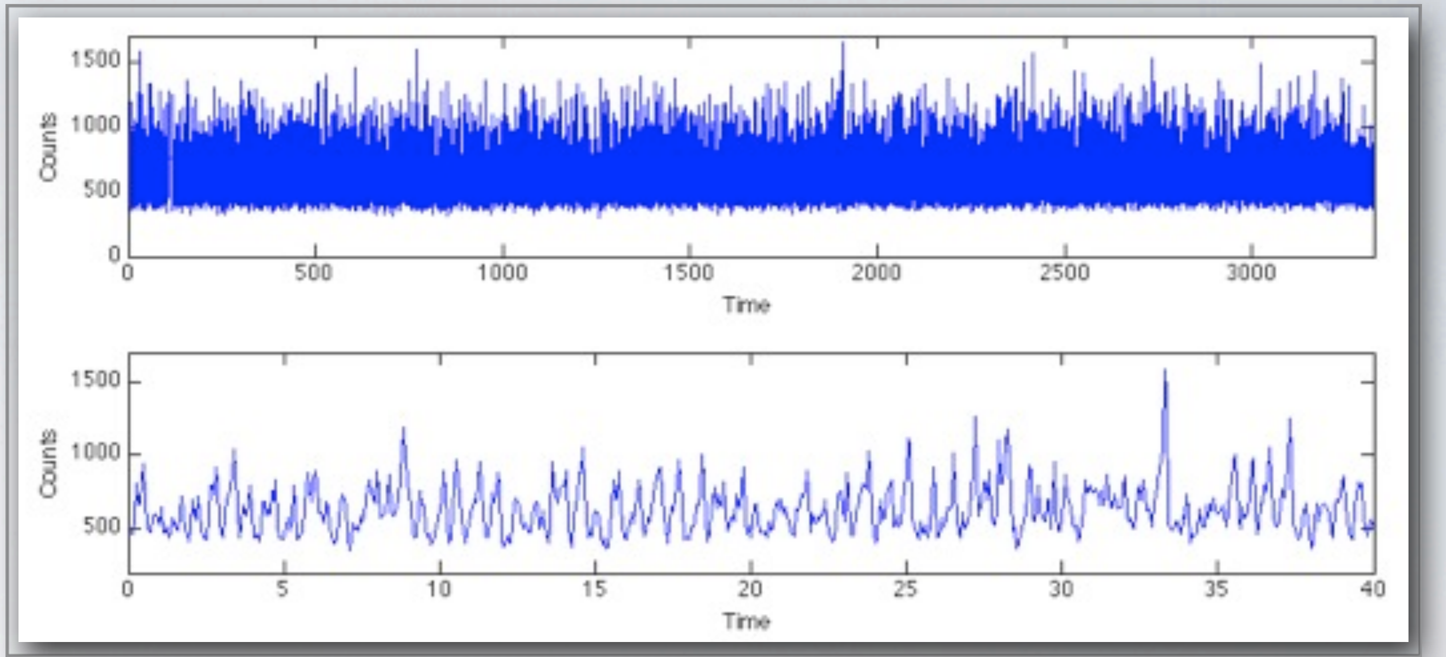
- Power spectrum of noise is very noisy! $\sigma_{P_j} = < P_j > = 2$
- Increasing length or Δt not useful
- Two ways out:
  - a) Frequency rebinning by M
  - b) Time slicing by W and averaging powers

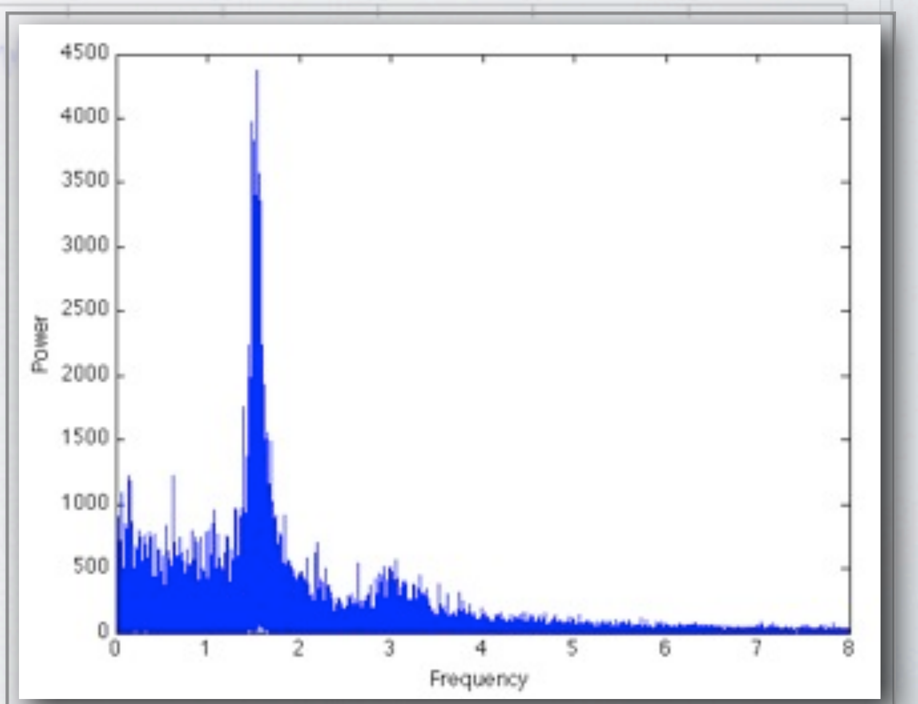


2 with 2MW dof
2MW distribution scaled by MW
Mean: 2
Standard dev: $\frac{2}{\sqrt{MW}}$

FULL POWER SPECTRUM

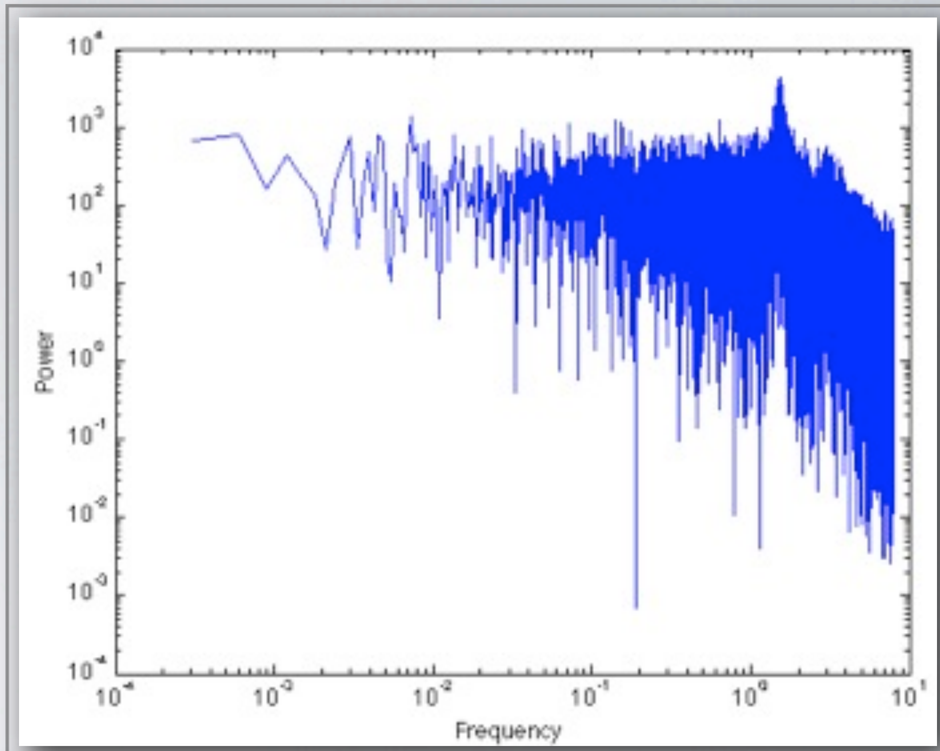
- RXTE light curve
- t = 1/16 seconds
- T = 3325 seconds
- Something can be seen by eye in the light curve
- Full power spectrum
- High-power signal, no coherent peak

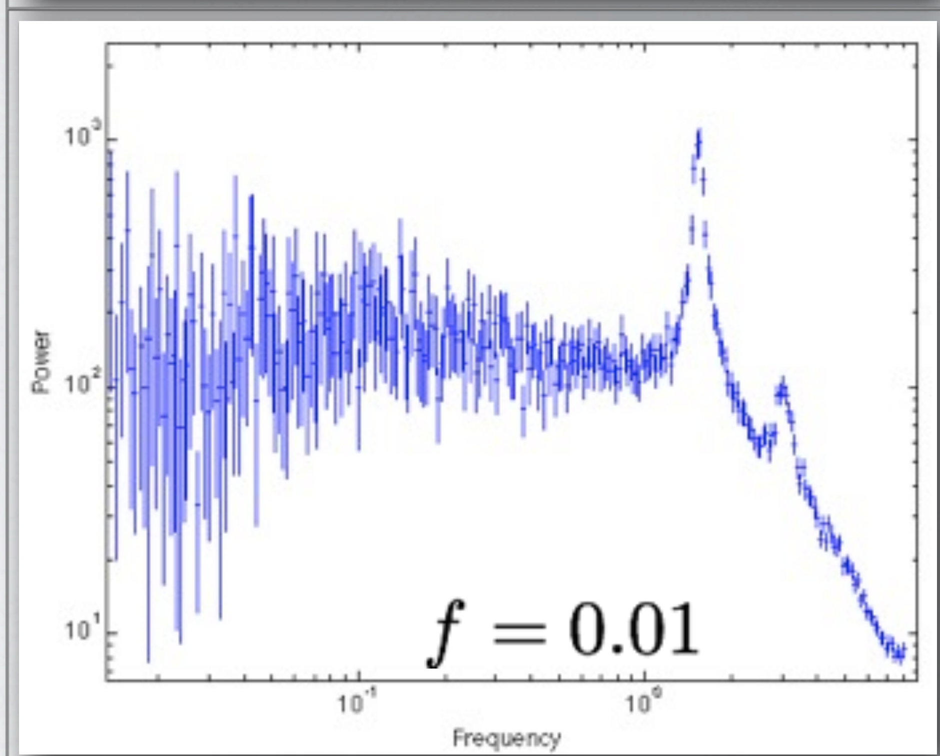




LOG SPACE AND REBINNING

- Log-log plot more appropriate for all frequencies
- Errors are 100%
- Frequency rebinning (M)
- Log-rebinning: $\Delta\nu_j = \Delta\nu_{j-1} * (1 + f)$
- Error bars, better shape
- Poisson level below scale



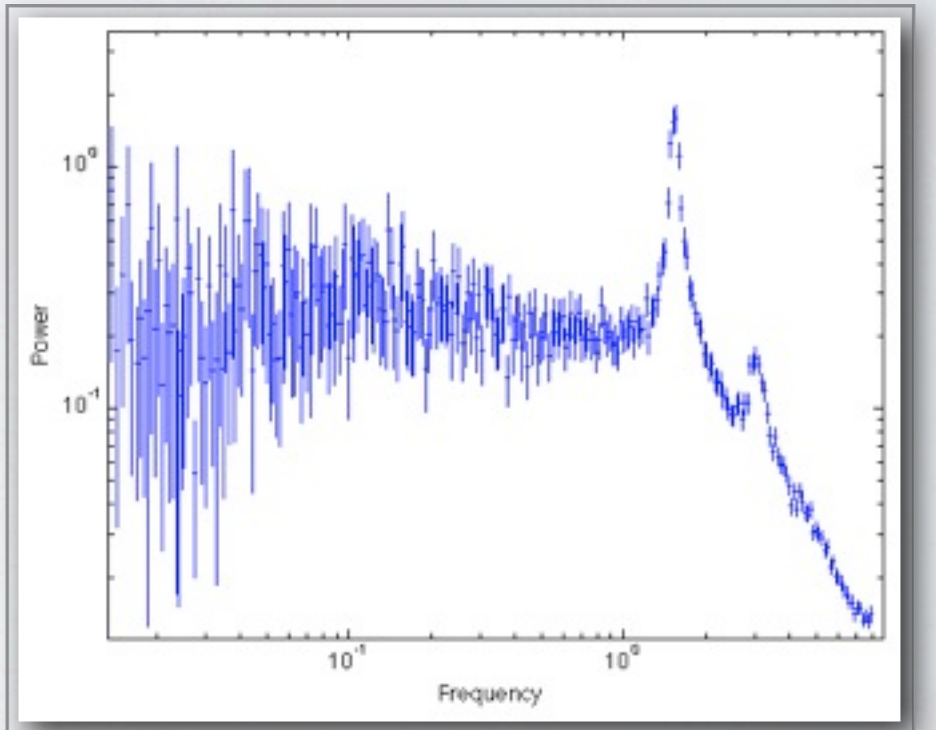


NORMALIZATION

- Leahy normalization very useful for statistics
- Power ∝ square intensity
- Remove it by dividing by square intensity: rms (Belloni) normalization
- Caveat: from Leahy to rms^2
- Meaning: squared rms per decade
- Root of integral gives fractional rms

$$P_{rms} = \frac{P}{C^2}$$

$$P_{rms} = \frac{P_{Leahy}}{C}$$

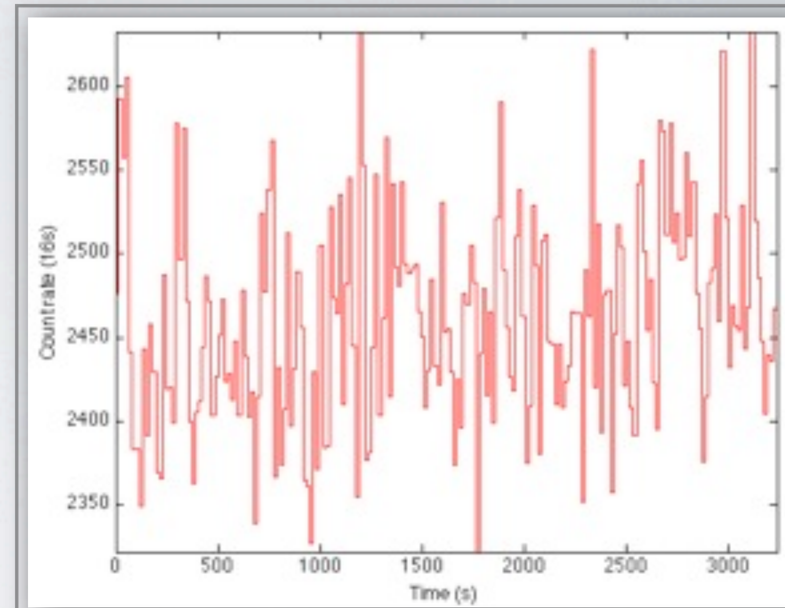


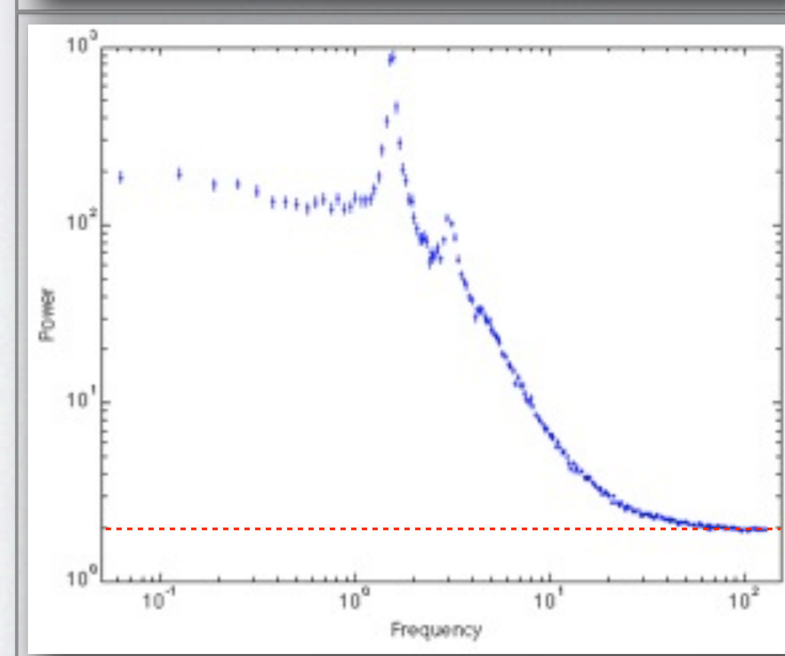
A NOTE ABOUT REBINNING

- Coherent peak: narrow power distribution - least rebinning - the longer the observation span, the better
- Broad peak: broad power distribution - rebinning helps - length of observation not crucial
- Very important for maximizing sensitivity

W: WELCH POWER SPECTRUM

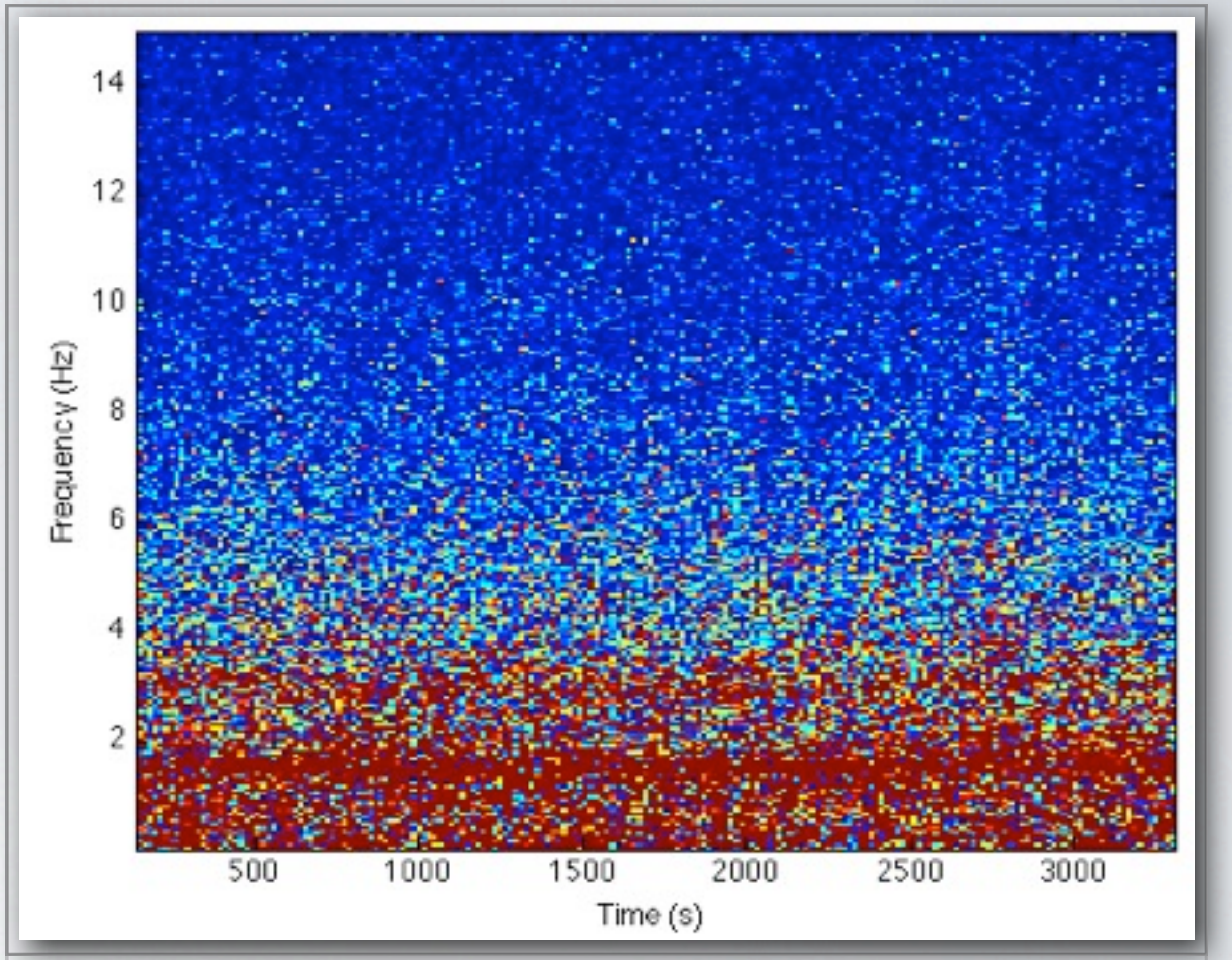
- If signal stationary
- Slice the signal
- Power spectrum of slices
- Add the W slices
- Sliding slices are also possible (statistics?)
- Windowing is also possible

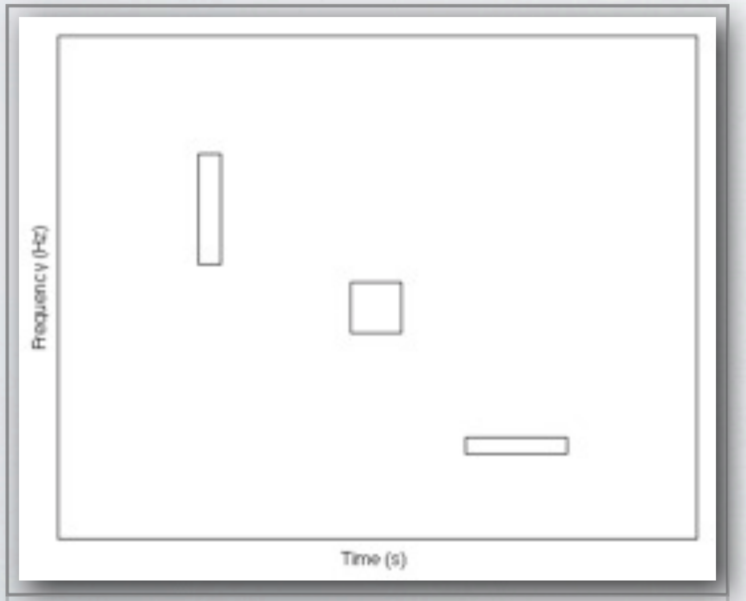




TIME-FREQUENCY ANALYSIS

- If signal is not stationary
- No average of power spectra
- Image: time-frequency-power
- Uncertainty principle





THE UNCERTAINTY PRINCIPLE

$$T^2 = \sigma_t^2 = \int (t - <t>)^2 |s(t)|^2 dt$$

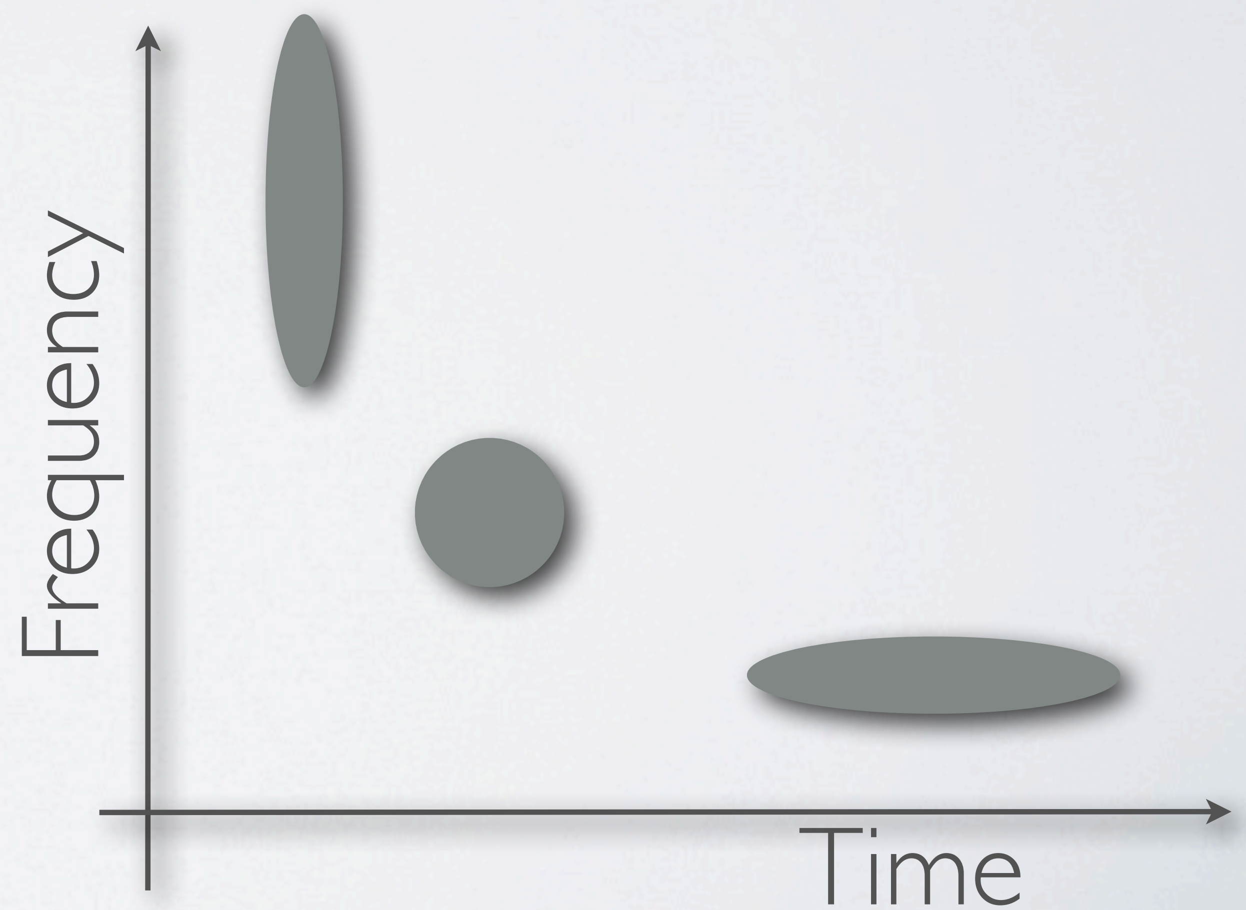
$$B^2 = \sigma_\omega^2 = \int (\omega - <\omega>)^2 |S(\omega)|^2 d\omega$$

$$TB \geq \frac{1}{2}$$

σ_t Duration

σ_ω Bandwidth

- You cannot beat it
- It's a big limitation



THE EASY WAY OUT

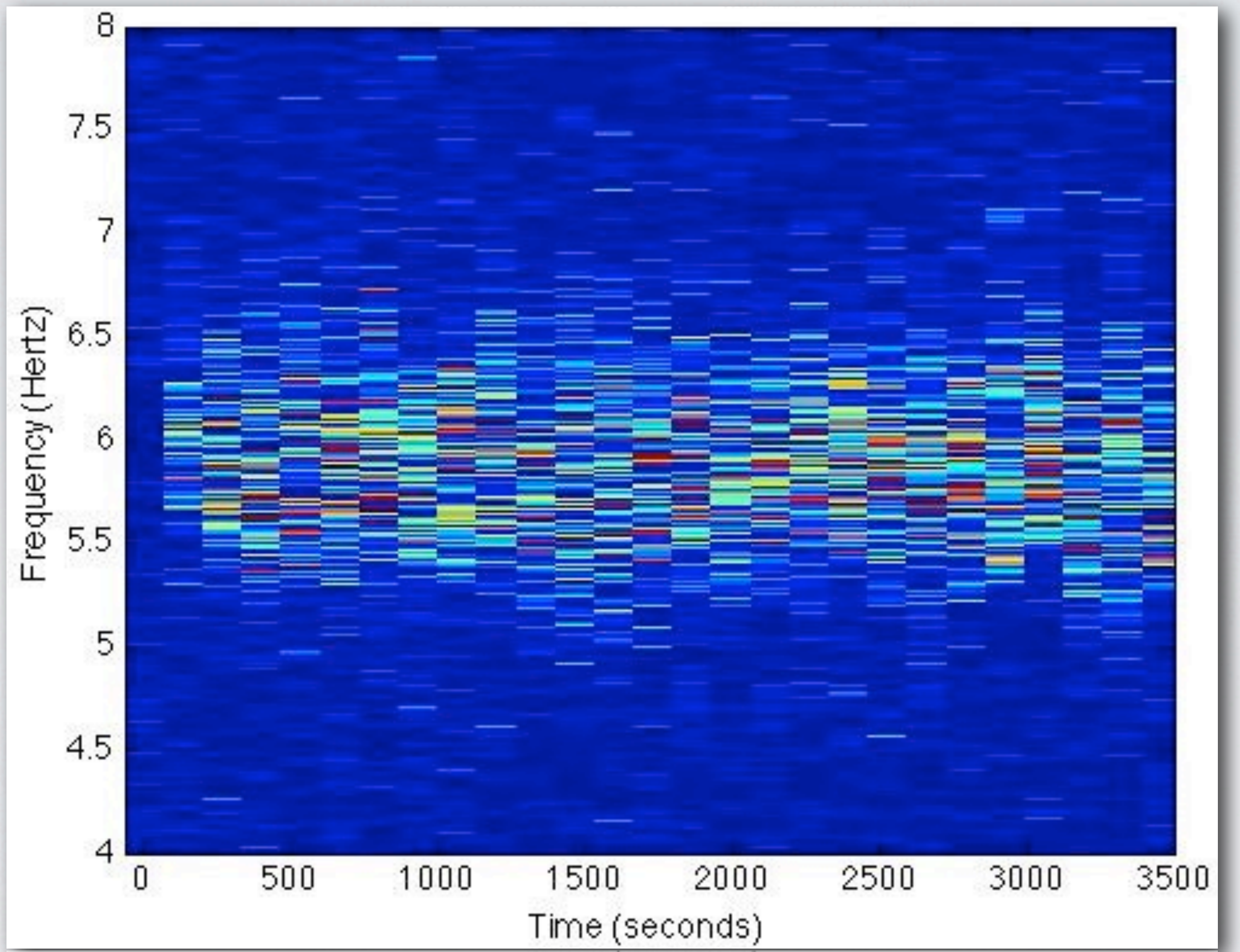
- Spectrogram (from short-term Fourier Transform)
- Sliding window to select time (window can be chosen)
- Obtain a time-frequency image

$$s_t(\tau) = s(\tau)h(\tau - t)$$

$$P(t, \omega) = \left| \frac{1}{\sqrt{(2\pi)}} \int e^{(-i\omega\tau)} s(\tau)h(\tau - t)d\tau \right|^2$$

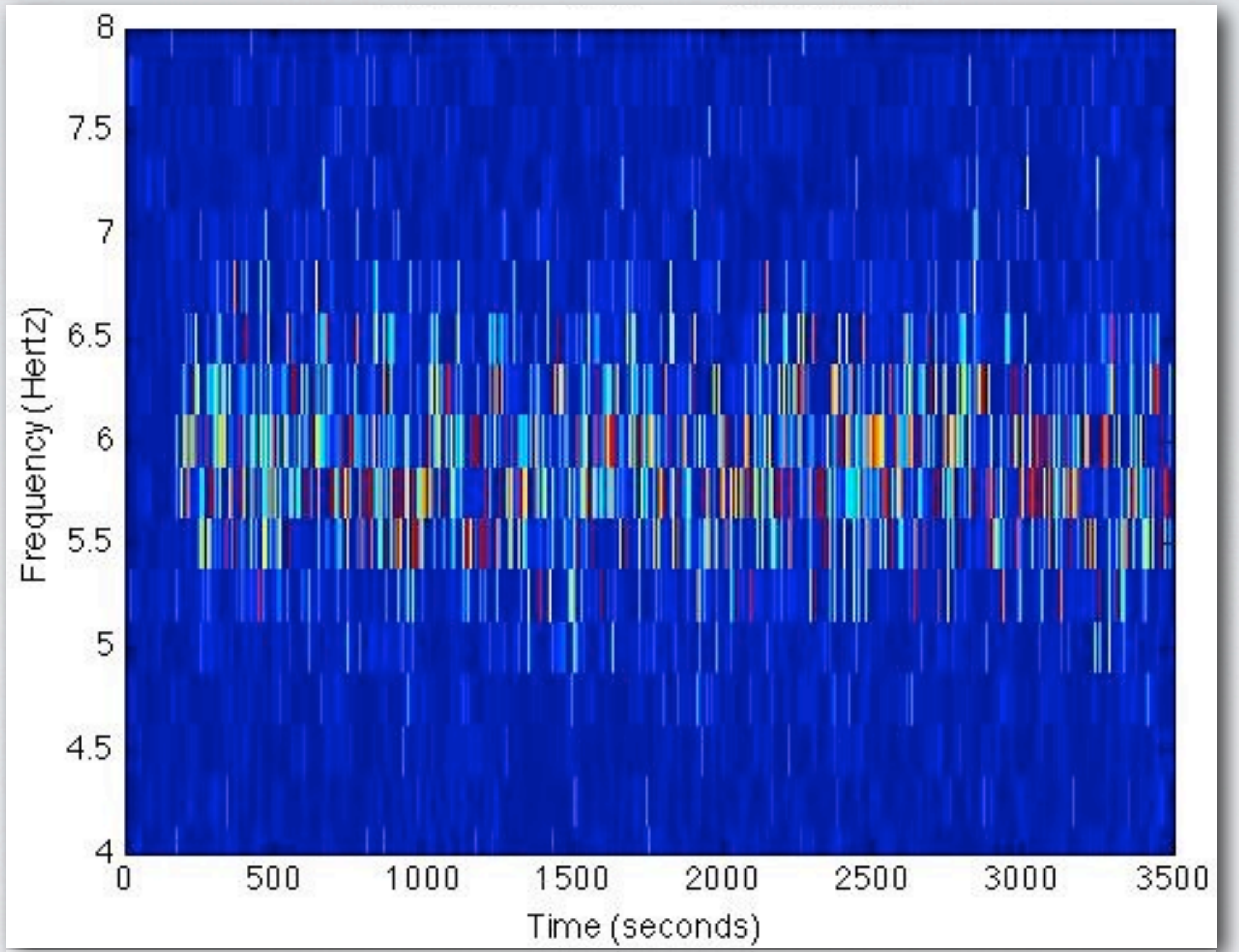
AN EXAMPLE

- Quasi-Periodic Oscillation T = 128 s



AN EXAMPLE

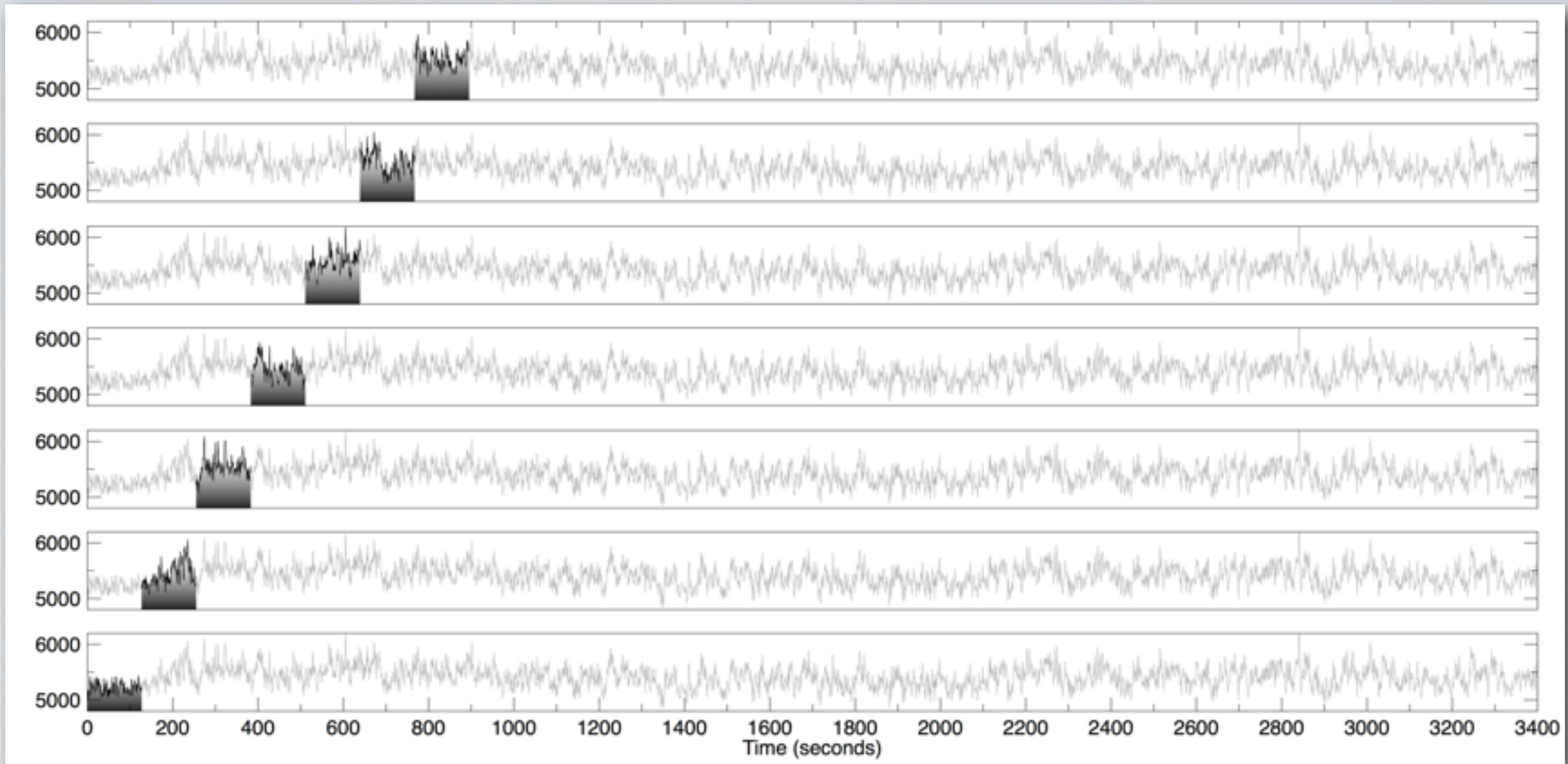
- Quasi-Periodic Oscillation T = 4 s



NON-OVERLAPPING

- Sliding window to select time $s_t(\tau) = s(\tau)h(\tau - t)$

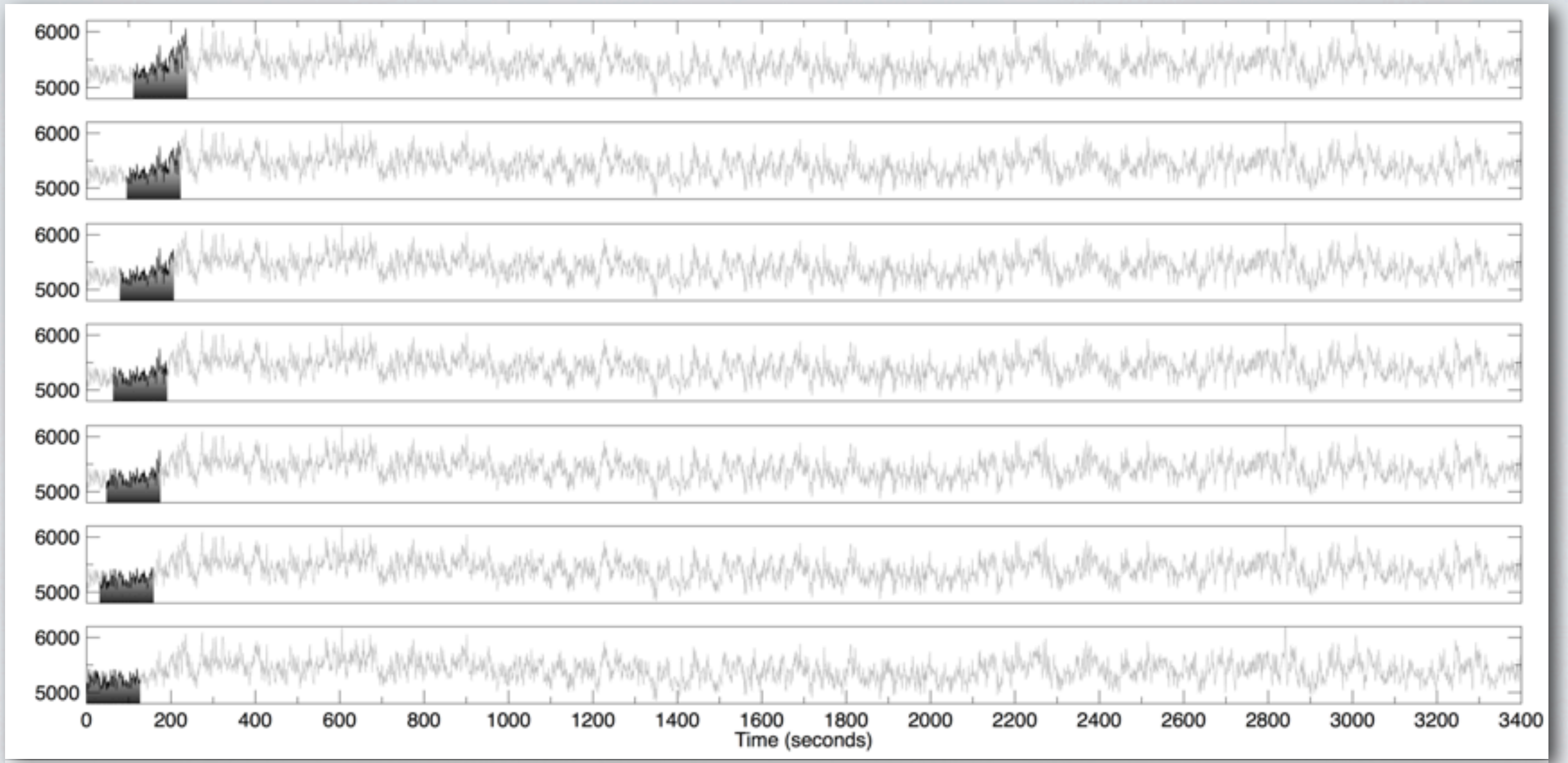
$$t = \tau$$



NON-OVERLAPPING

- Sliding window to select time $s_t(\tau) = s(\tau)h(\tau - t)$

$t < \tau$

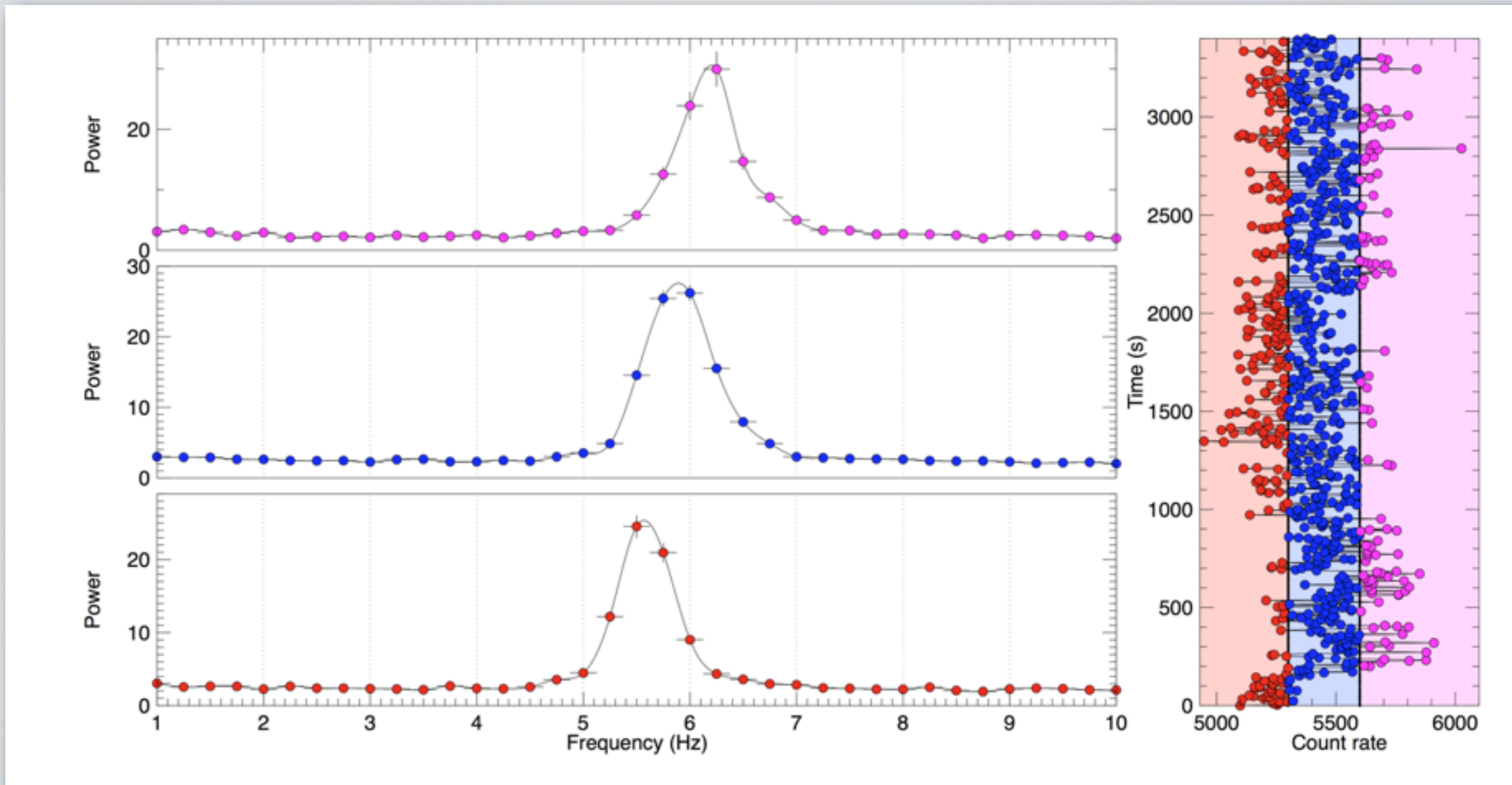


SHIFT 'N' ADD TECHNIQUE

- Used for twin high-frequency peaks
- You see one, not the other
- The one you see moves
- Correct for the movement, align the spectra in an additive way
- More complex: multiplicative technique (tricky to implement)

LINEAR SHIFT AND ADD

- Good to recover features at a constant distance in ν



INSTRUMENTAL DEAD TIME

- After a photon, dead time
- Introduces correlations between photons (no Poisson!)
- It must be as small as possible and well-known and modeled
- Two types of dead time:
    - Paralyzable — Every incident event causes a dead time t_d even if it's not detected
    - Non-paralyzable — Only a detected event causes a dead time t_d

PARALYZABLE DEAD TIME

- If incident rate r_{in} is very high, no detected counts at all!
- Detected rate: $r_0 = r_{in} e^{-r_{in} t_d}$ $\lim_{r_{in} \to +\infty} r_0 = 0$
- In RXTE/PCA, for binning time $t_b \geq t_d$

$$< P_j >= 2 \times \left[1 - 2 r_0 t_d \left(1 - \frac{t_d}{2 t_b}\right)\right] - 2 \frac{N-1}{N} r_0 t_d \left(\frac{t_d}{t_b}\right) \cos\left(\frac{2\pi j}{N}\right)$$

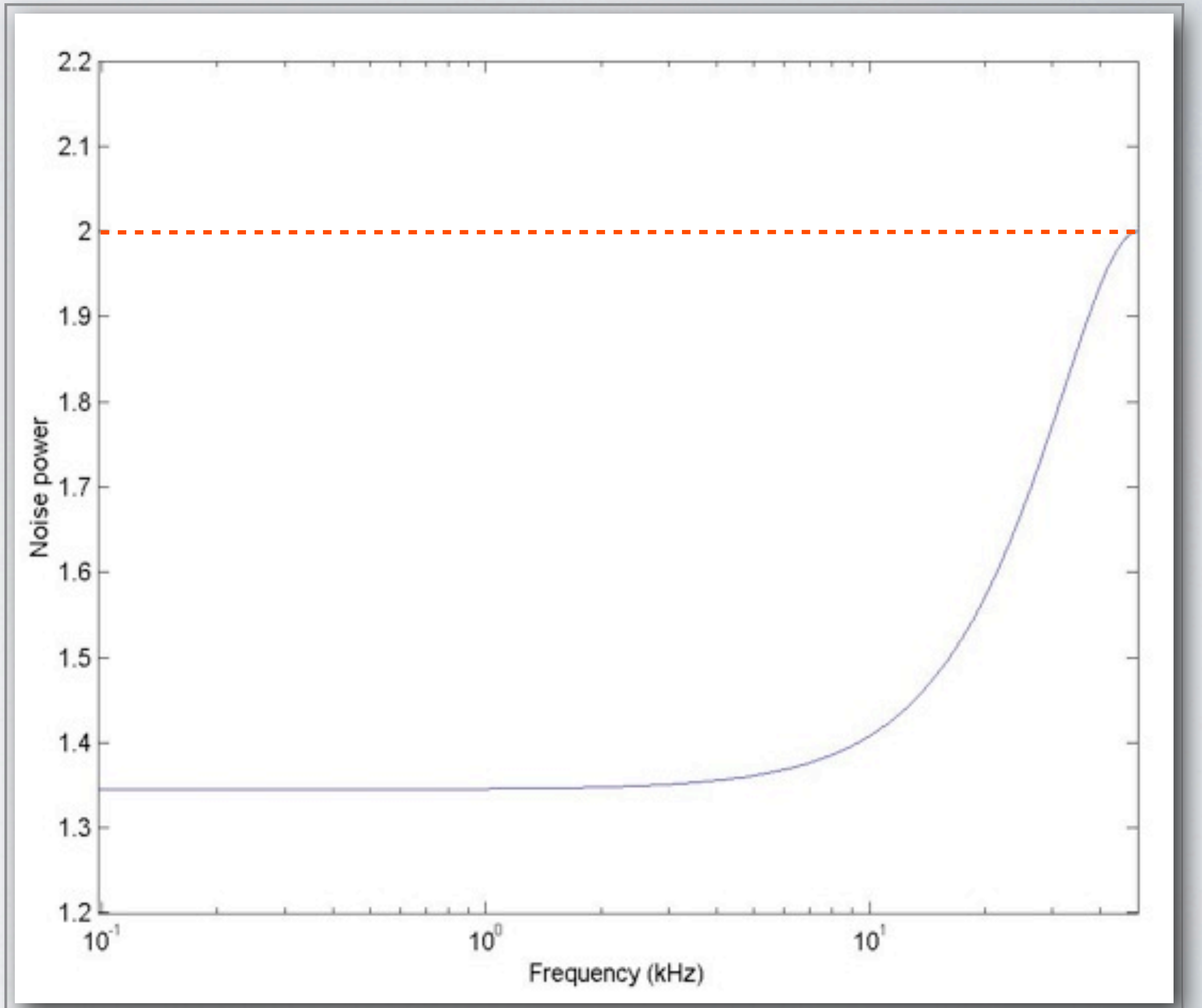
PARALYZABLE DEAD TIME

r_{in} = 20 kcts/s

r_0 = 16.385 kcts/s

$t_d = t_b$ = 10 μs

N = 1024

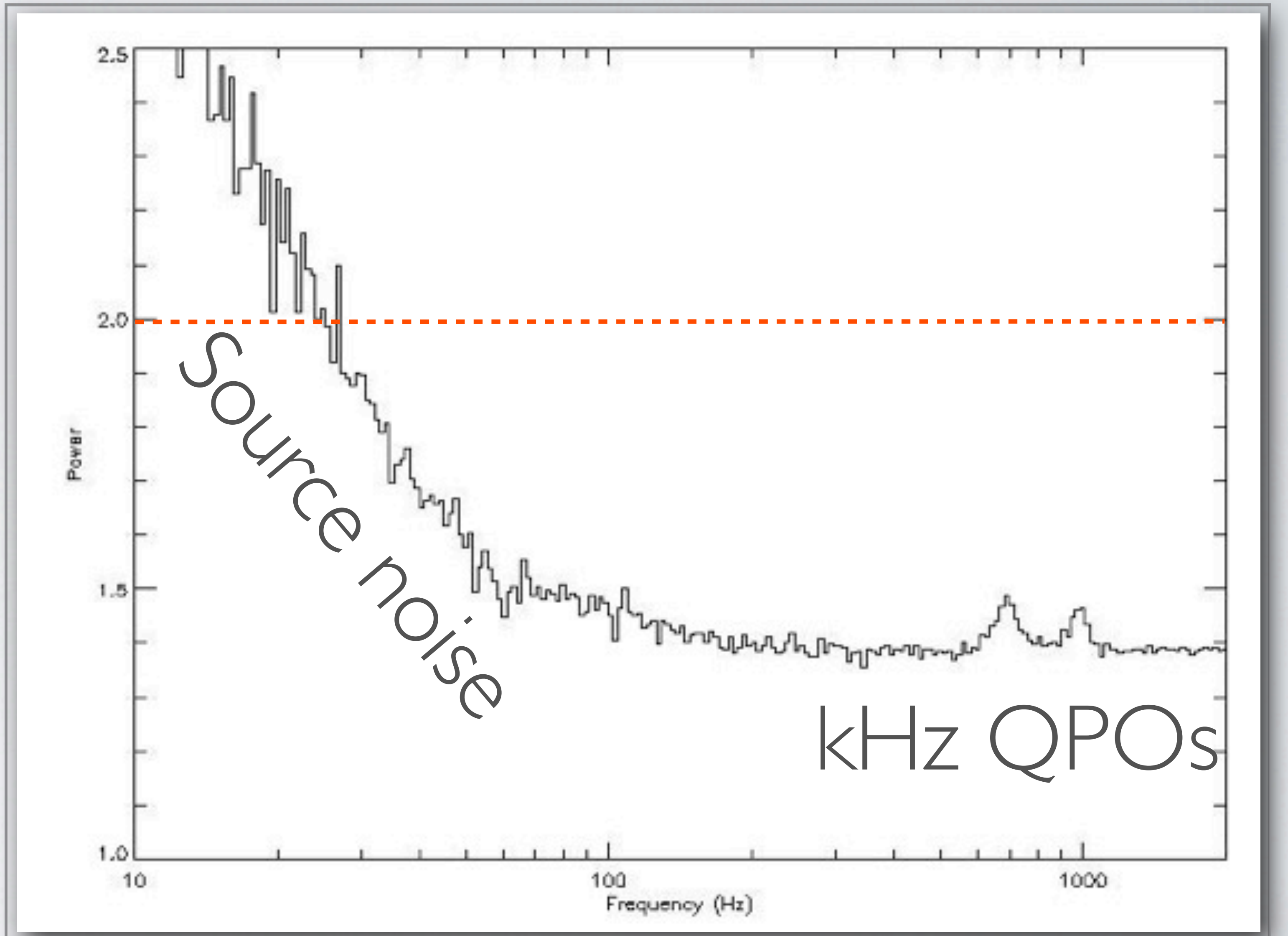


NON-PARALYZABLE DEAD TIME

- If incident rate r_{in} is very high, one count every t_d
- Detected rate: $r_0 = \frac{r_{in}}{1 + r_{in} t_d}$ $\lim_{r_{in} \to +\infty} r_0 = t_d^{-1}$
- Formula is even more complicated, result is similar
    - Depression of noise level @ low frequencies (correlation)
    - Peak @ t_d (quasi-periodicity)

PARALYZABLE DEAD TIME: SCO X-1

$r_0 = 10^5$ cts/s
$t_d = 10$ μs



FITTING POWER SPECTRA

- Fit with typical minimization (χ^2)
- Rebinning is important for χ^2
- Error estimation vs. significance
- Limit in power an NOT rms
- Coherent peaks: distribution of powers and number of trials

NUMBER OF TRIALS

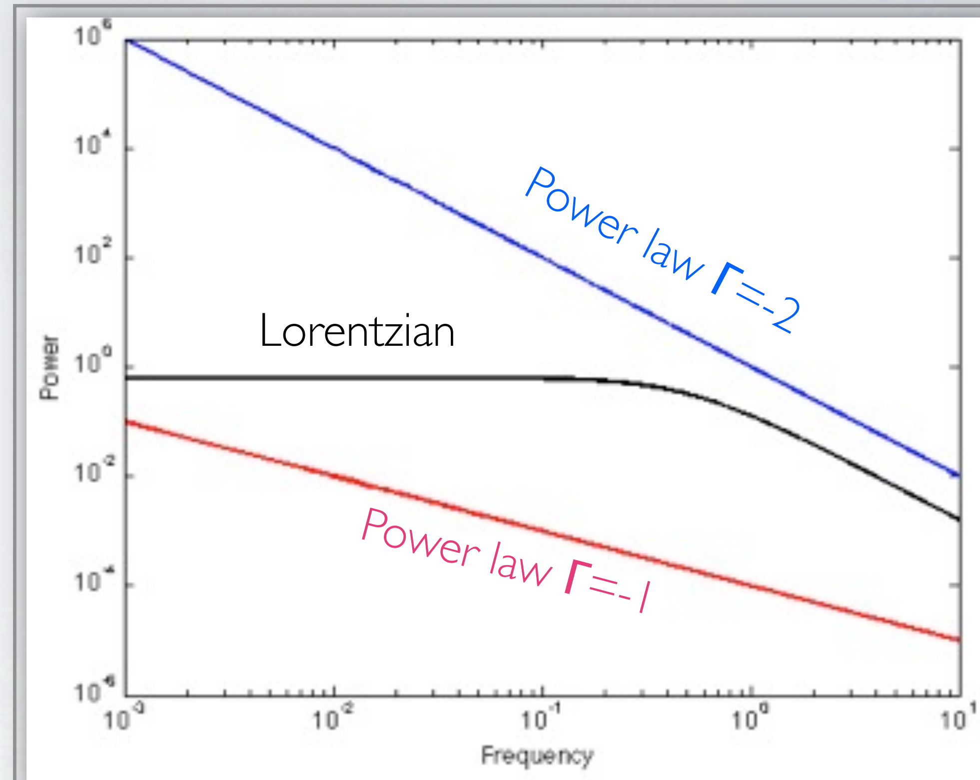
- Important statistical concept
- Should be done correctly, but if P is small can be approximated

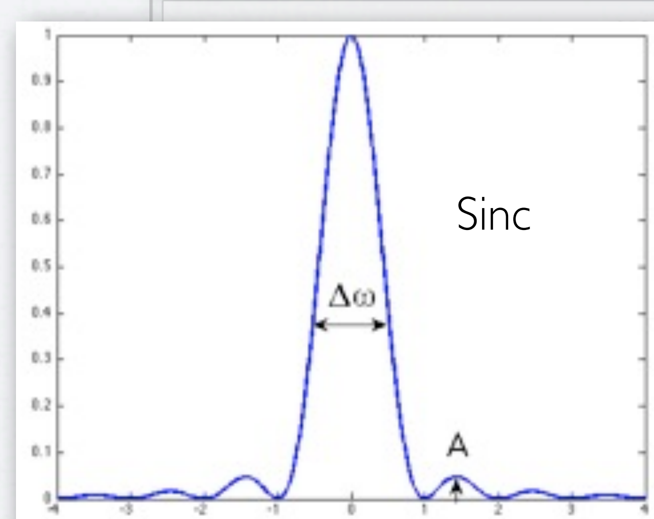
$$\tilde{P}_{chance} = P_{chance} \times N_{trials}$$

- IMPORTANT: how to estimate N_{trials}
- For Power Spectra: number of *independent* frequencies

CONTINUUM COMPONENTS

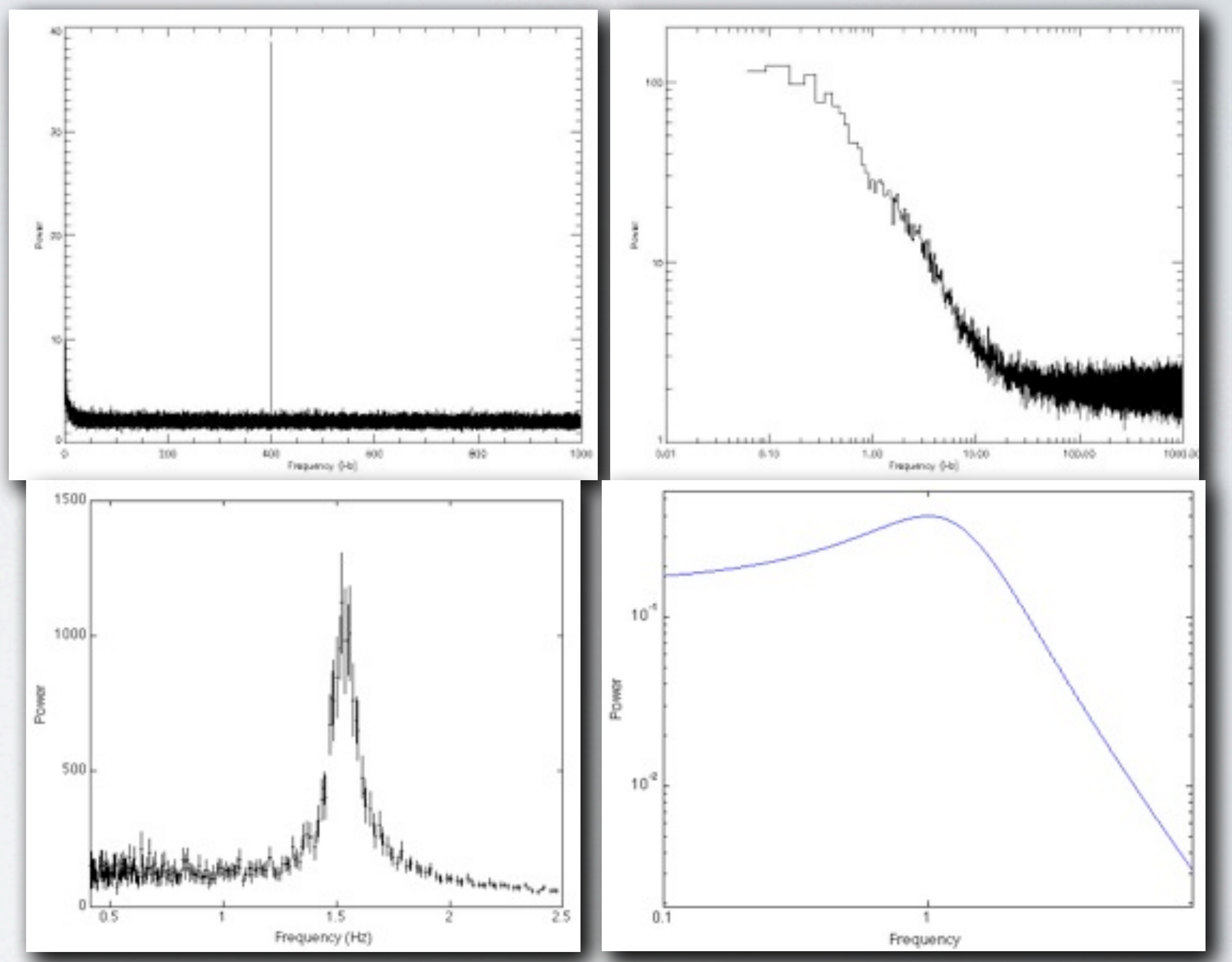
- Very important for accreting sources
- Slope is limited by the window
- Window overflow
- Γ=-2 is the steepest value
- If an issue (pulsar noise):
  exotic methods





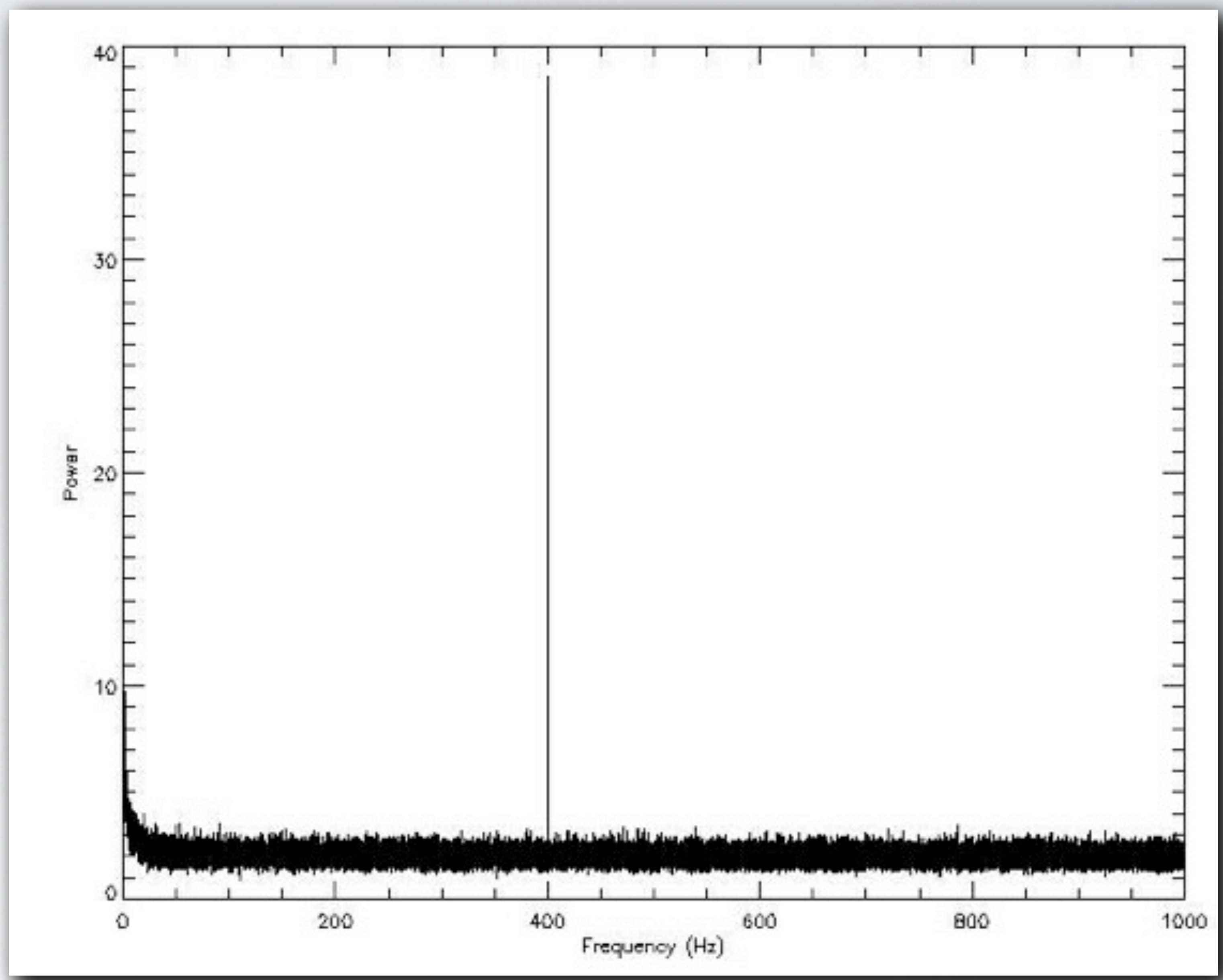
MAIN TYPES OF SIGNALS

- Coherent pulsation
- Broad-band noise
- Broad peak (QPO)
- "Peaked-noise"



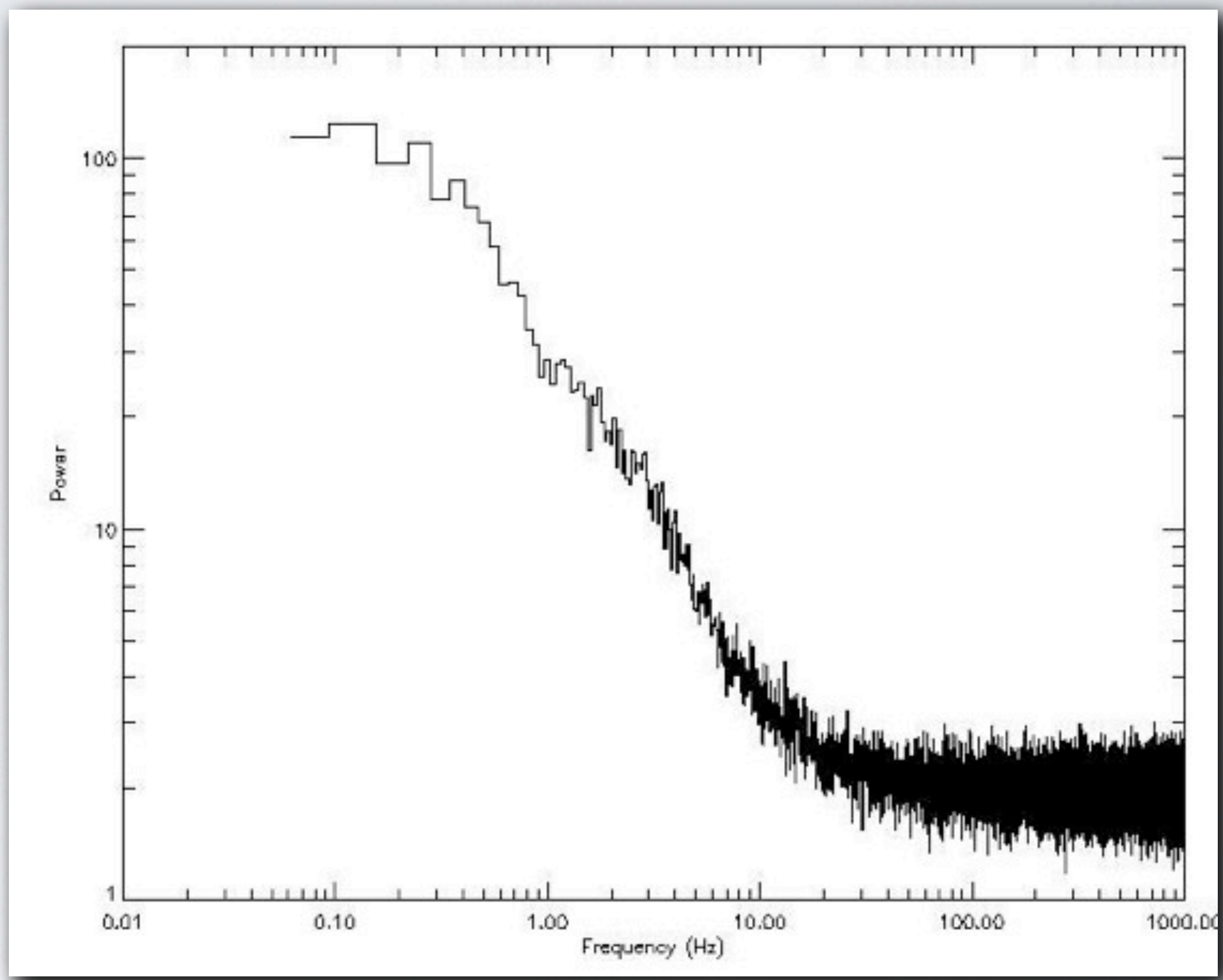
MAIN TYPES OF SIGNALS

- Coherent pulsation
- Broad-band noise
- Broad peak (QPO)
- "Peaked-noise"



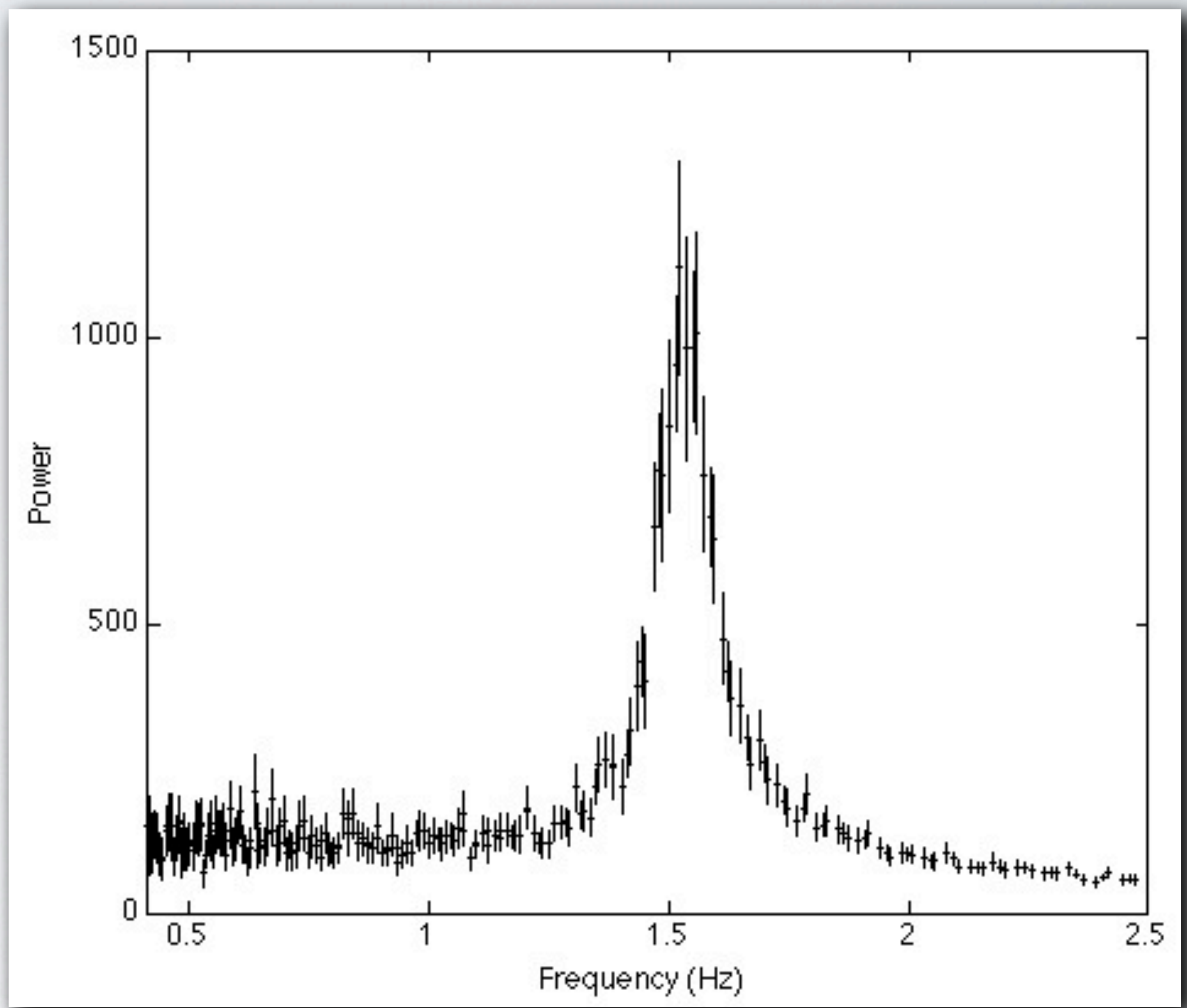
MAIN TYPES OF SIGNALS

- Coherent pulsation
- Broad-band noise
- Broad peak (QPO)
- "Peaked-noise"



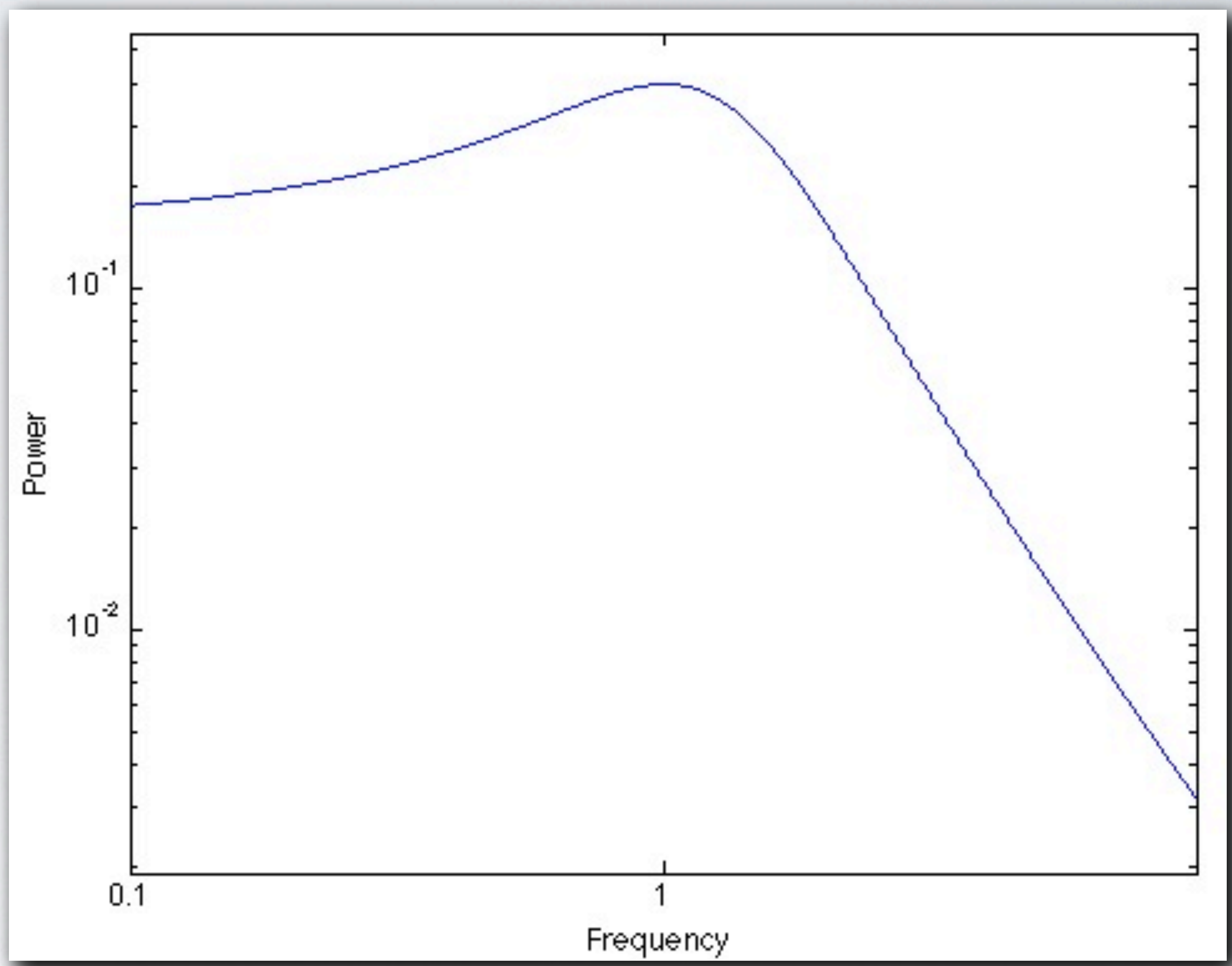
MAIN TYPES OF SIGNALS

- Coherent pulsation
- Broad-band noise
- Broad peak (QPO)
- "Peaked-noise"



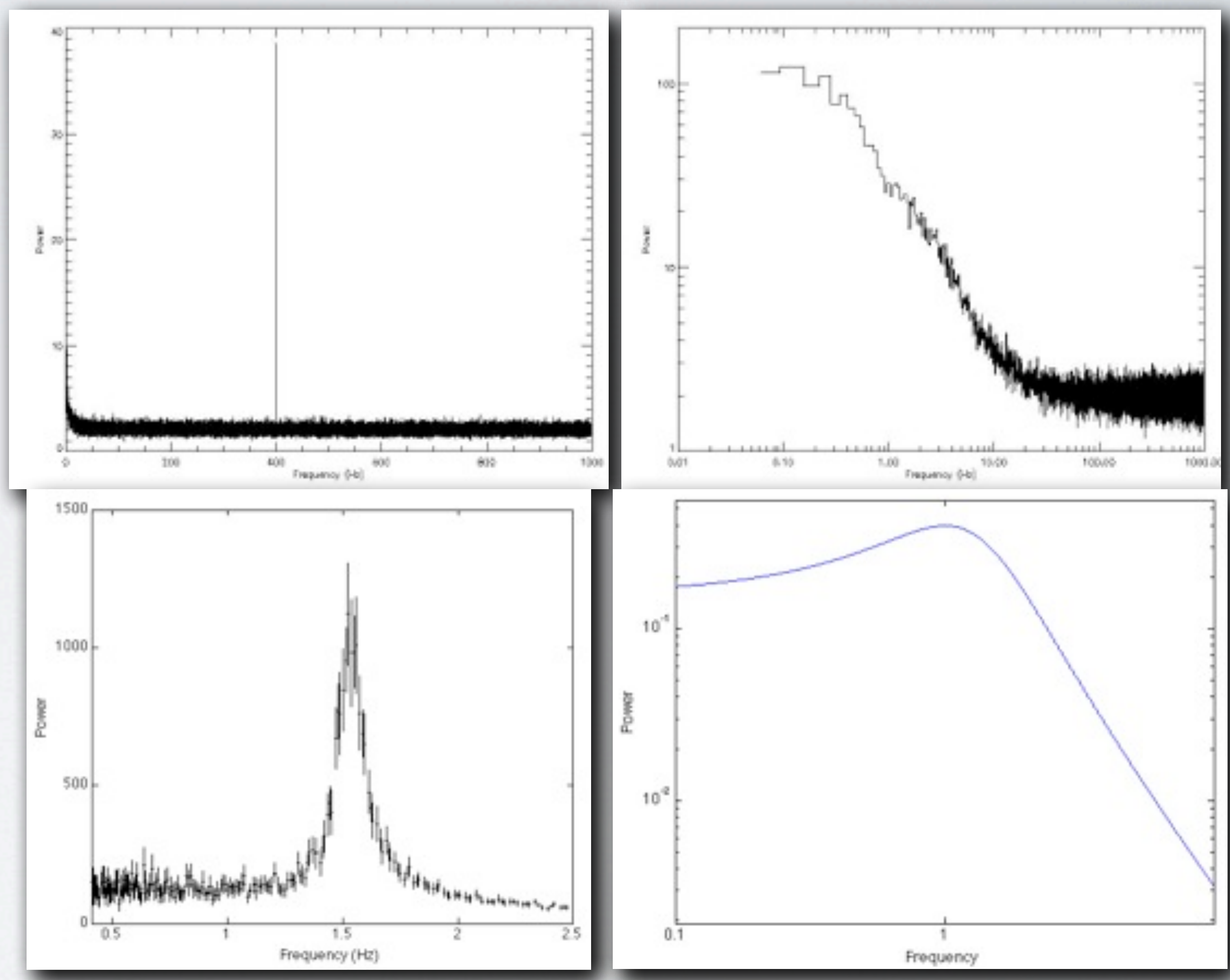
MAIN TYPES OF SIGNALS

- Coherent pulsation
- Broad-band noise
- Broad peak (QPO)
- "Peaked-noise"

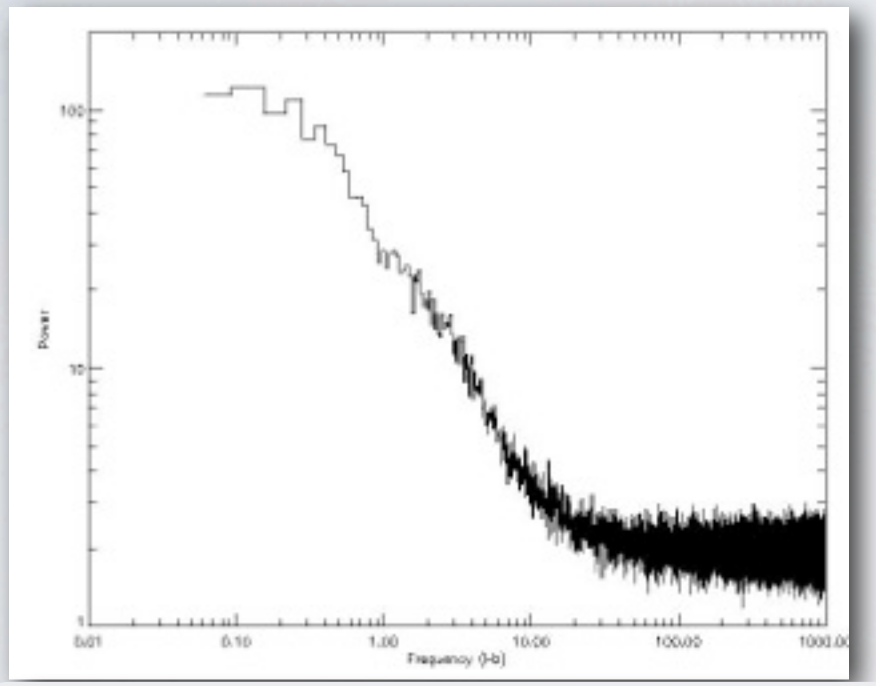


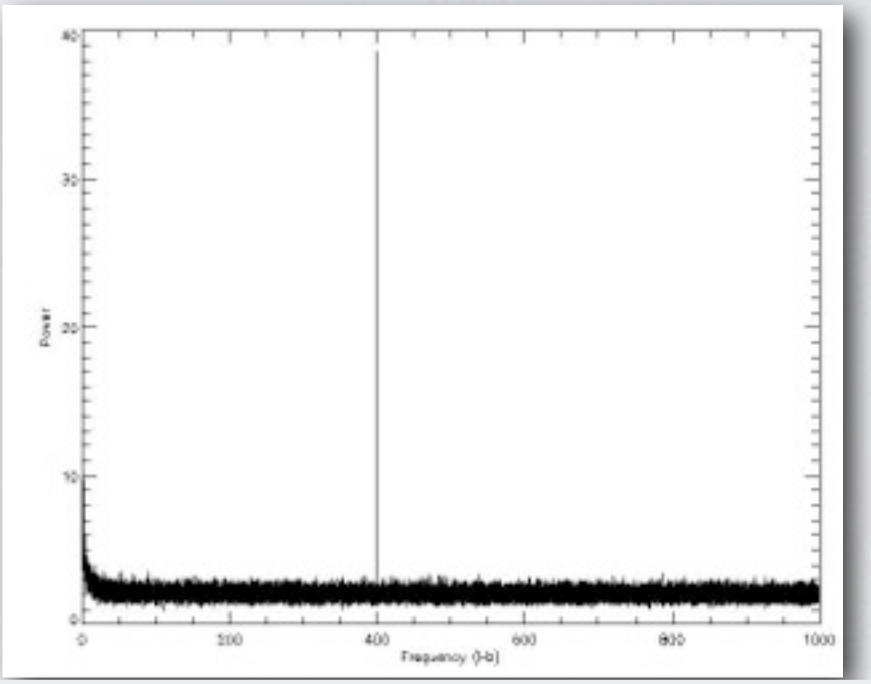
MAIN TYPES OF SIGNALS

- Coherent pulsation
- Broad-band noise
- Broad peak (QPO)
- "Peaked-noise"

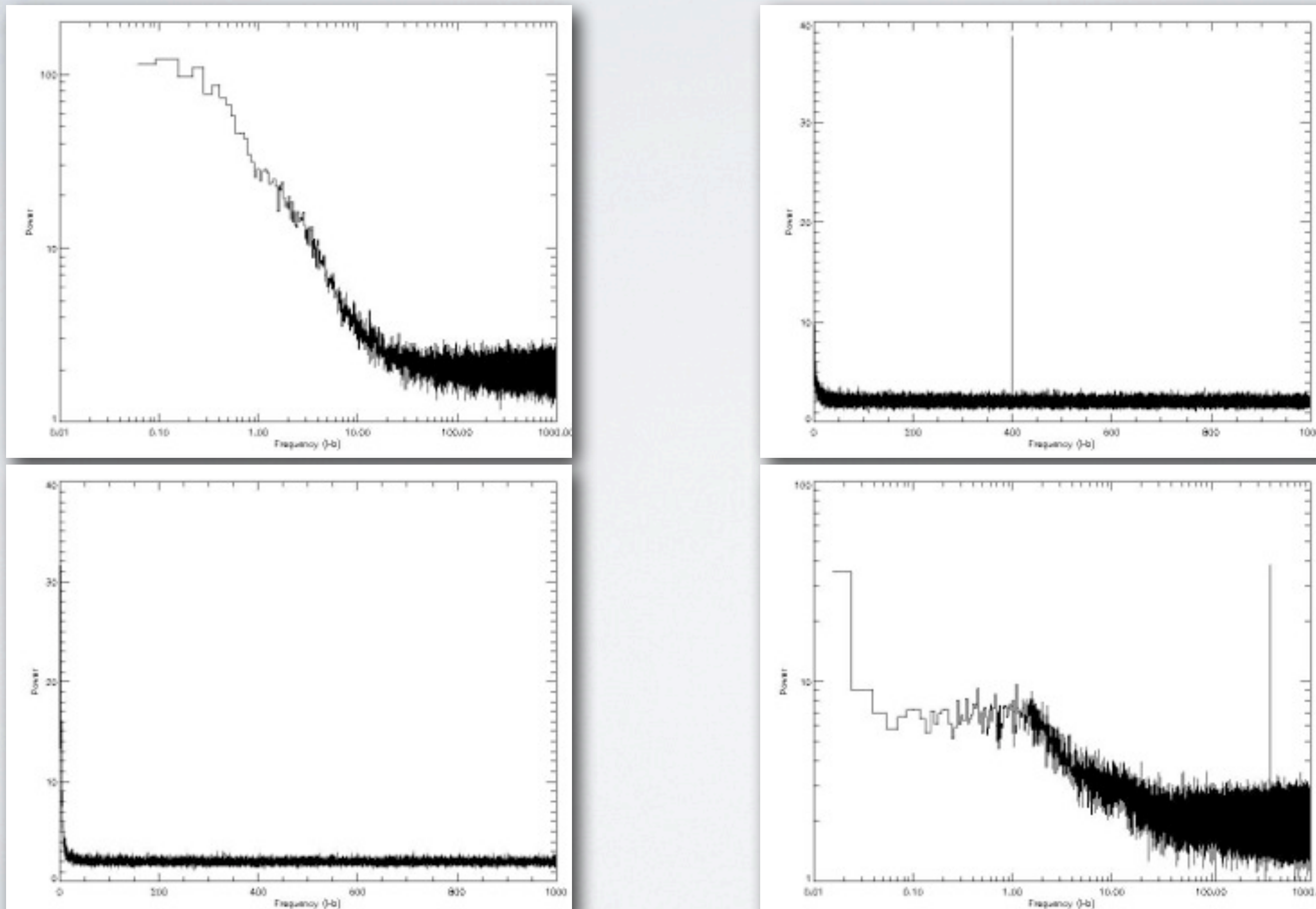


A WORD ON REPRESENTATION





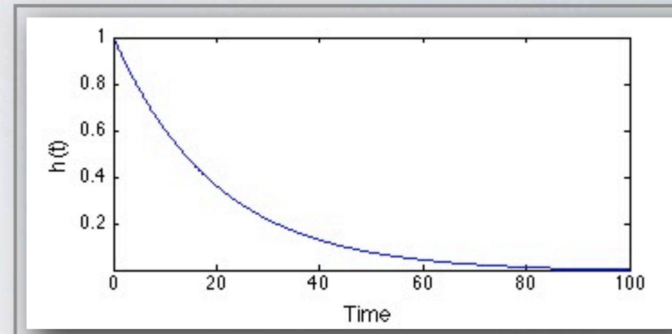
A WORD ON REPRESENTATION

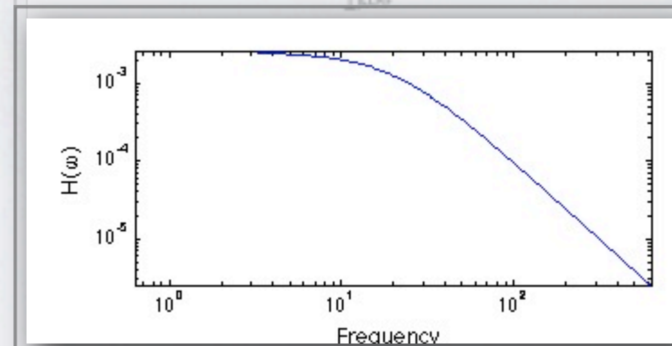


THE LORENTZIAN (ZERO-CENTERED)

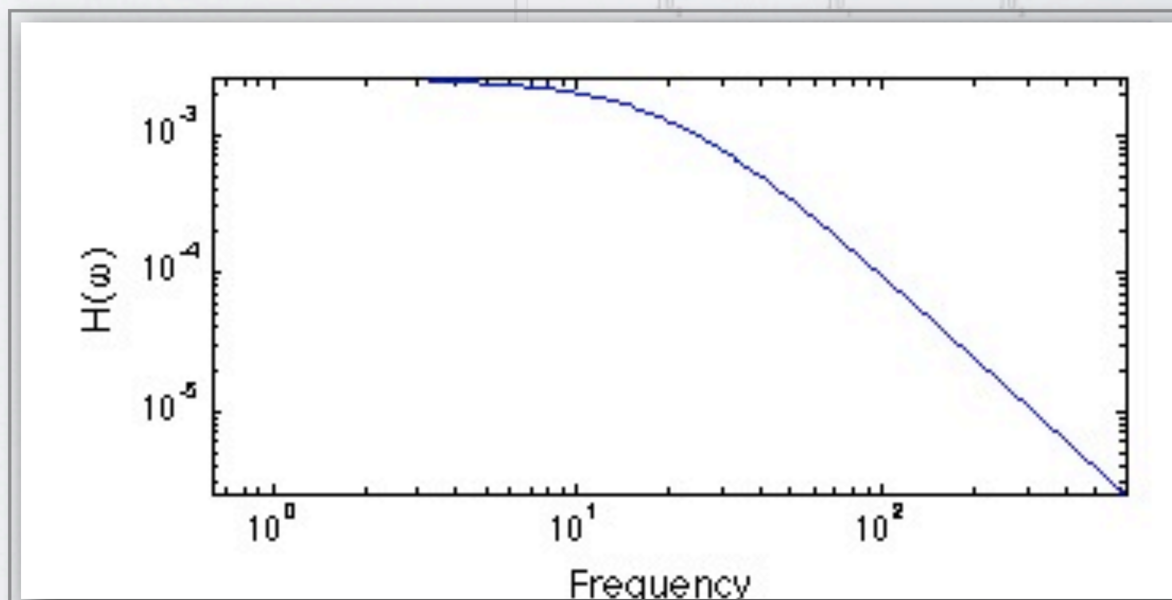
- Power spectrum of a one-sided exponential

$$L(\nu; N, \Delta) = \frac{\Delta}{2\pi} \frac{1}{\nu^2 + (\frac{\Delta}{2})^2}$$





- Good for modeling broad-band noise components (flat-top)

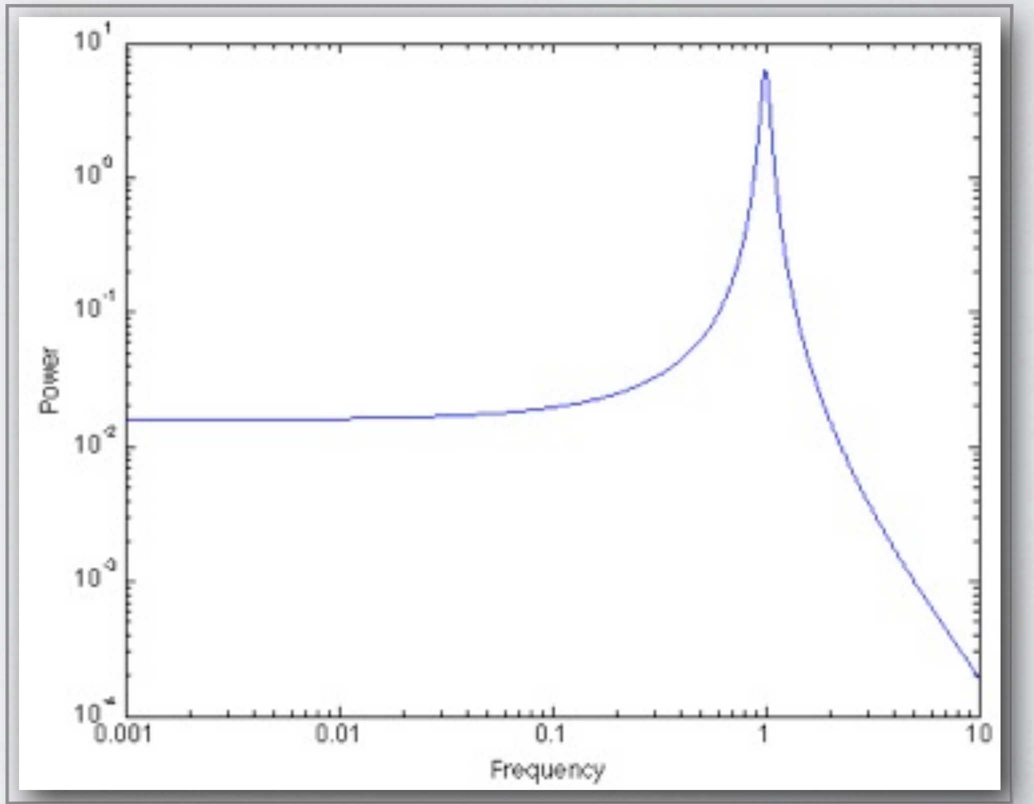


THE LORENTZIAN

- Centroid of Lorentzian not at zero

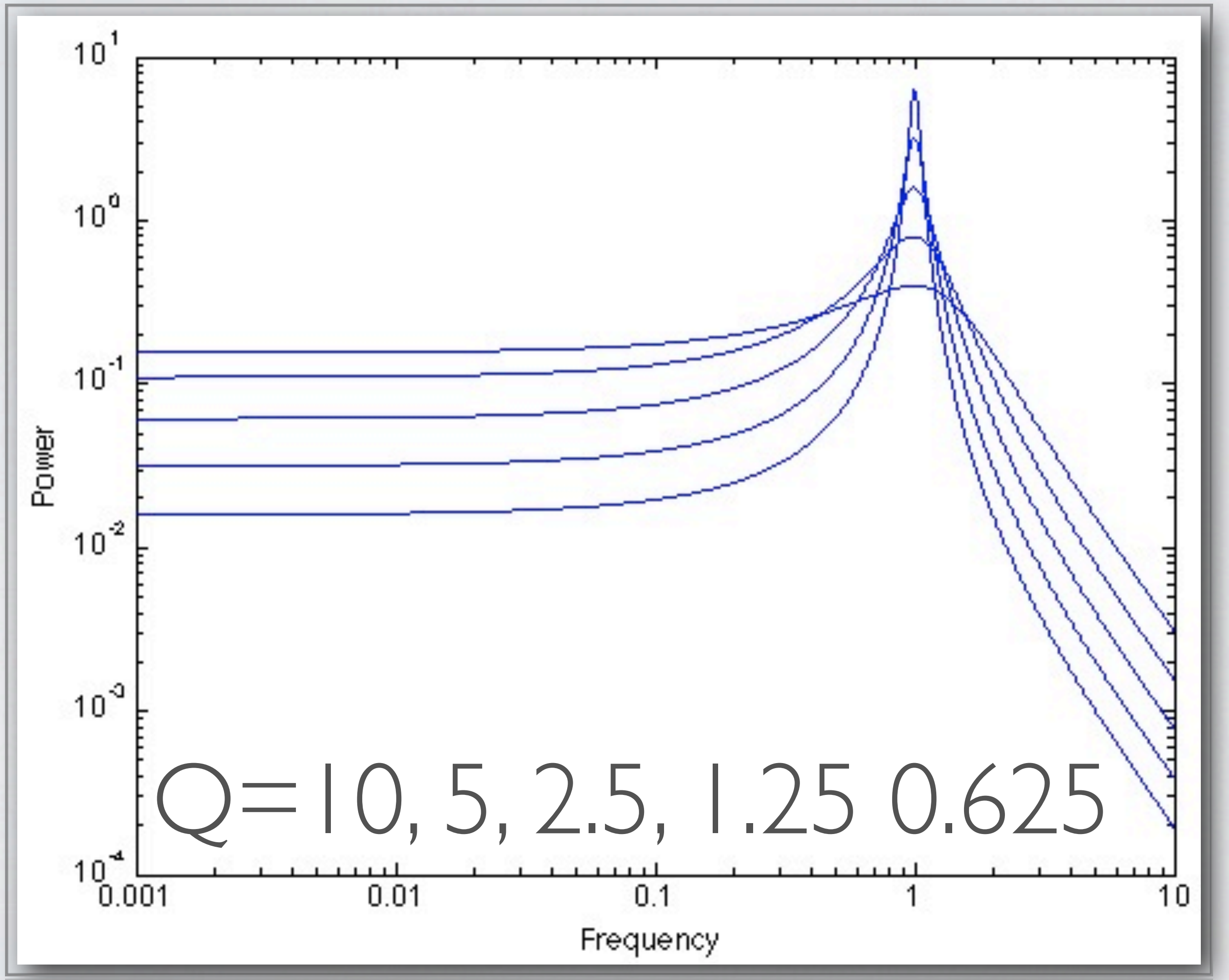
$$L(\nu; N, \nu_0, \Delta) = \frac{\Delta}{2\pi} \frac{1}{(\nu - \nu_0)^2 + (\frac{\Delta}{2})^2}$$

- Good for modeling Quasi-Periodic Peaks



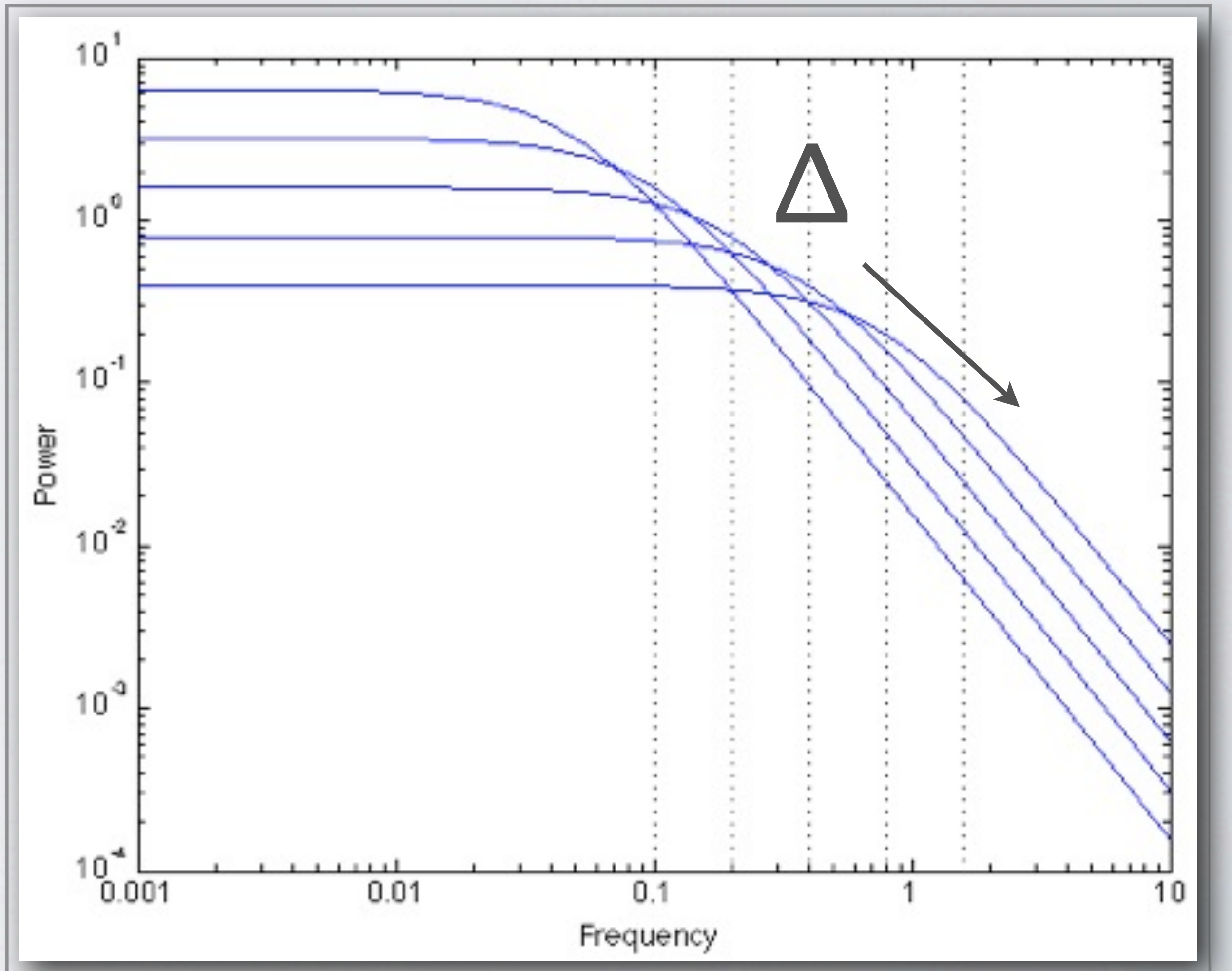
THE QUALITY FACTOR Q

- To quantify the coherence of a component $Q = \frac{\nu_0}{\Delta}$



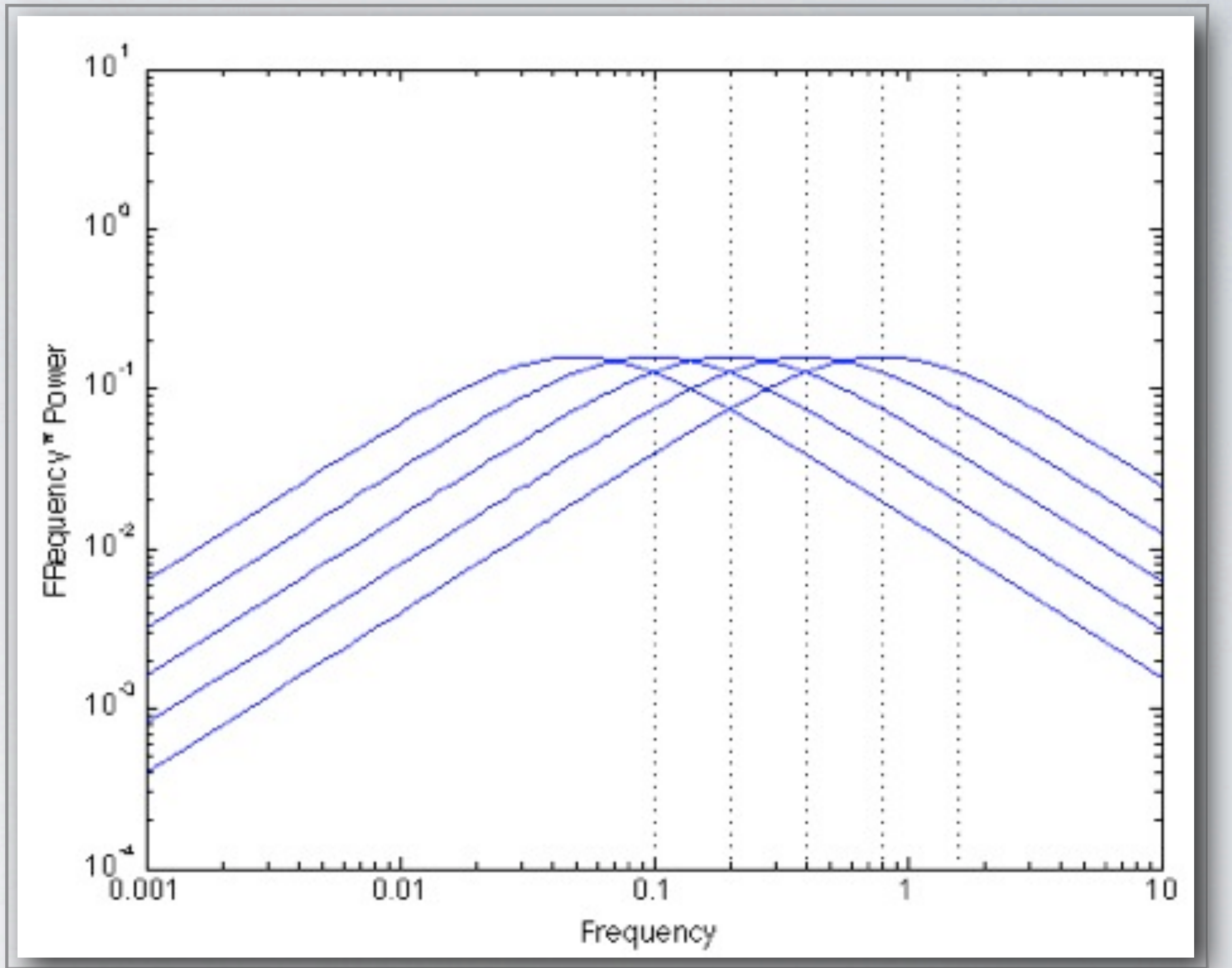
Q=0: THE PEAK WITHOUT QUALITY

- Here $\nu_0 = 0$, equal N
- Notice position of the break
- Factor of two higher



BETTER REPRESENTATION

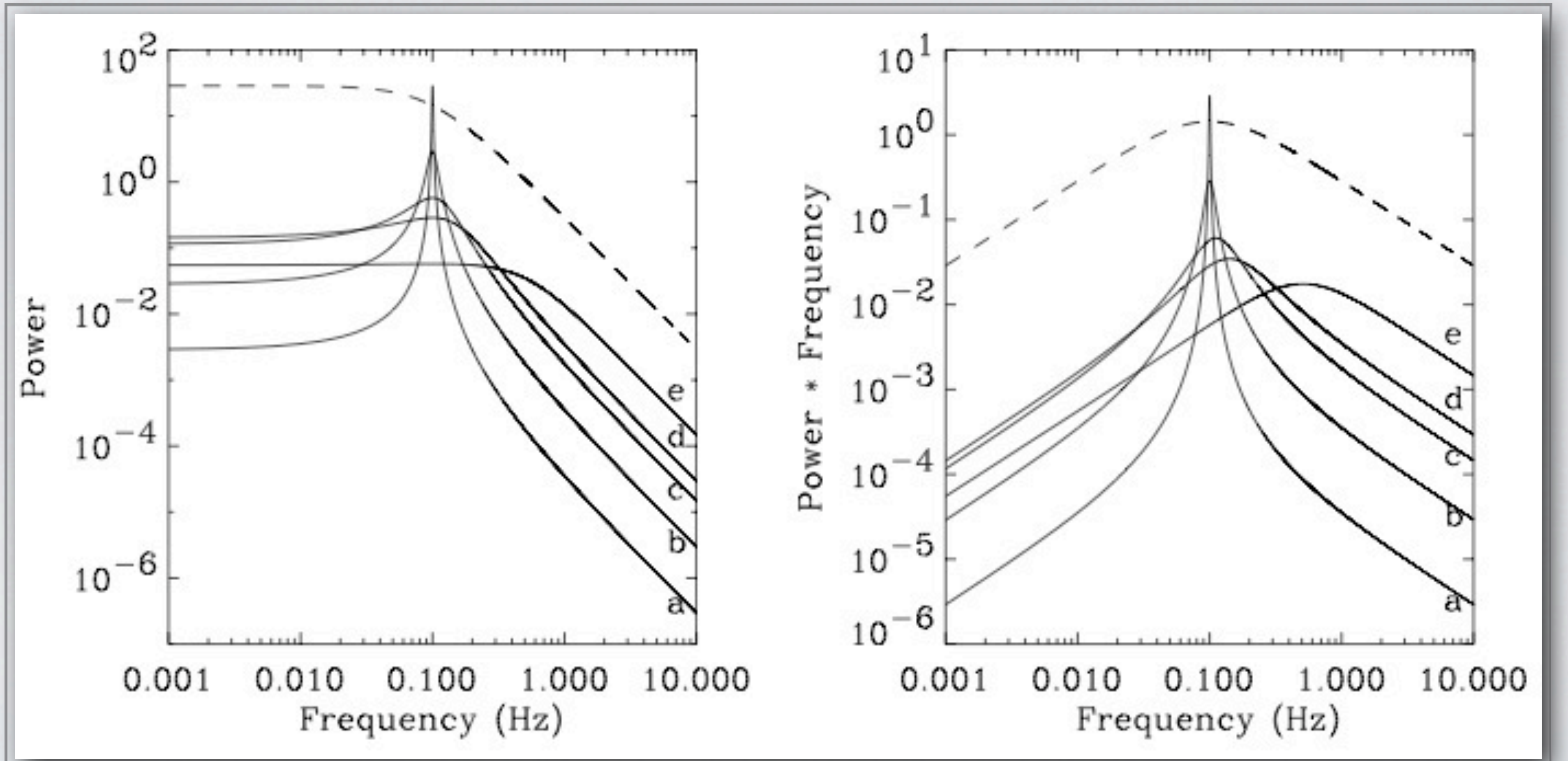
- In νP_ν the effect is the same
- Better value is $\Delta/2$
- But... how do I treat things homogeneously and how do I treat peaked noise?



CHARACTERISTIC FREQUENCY

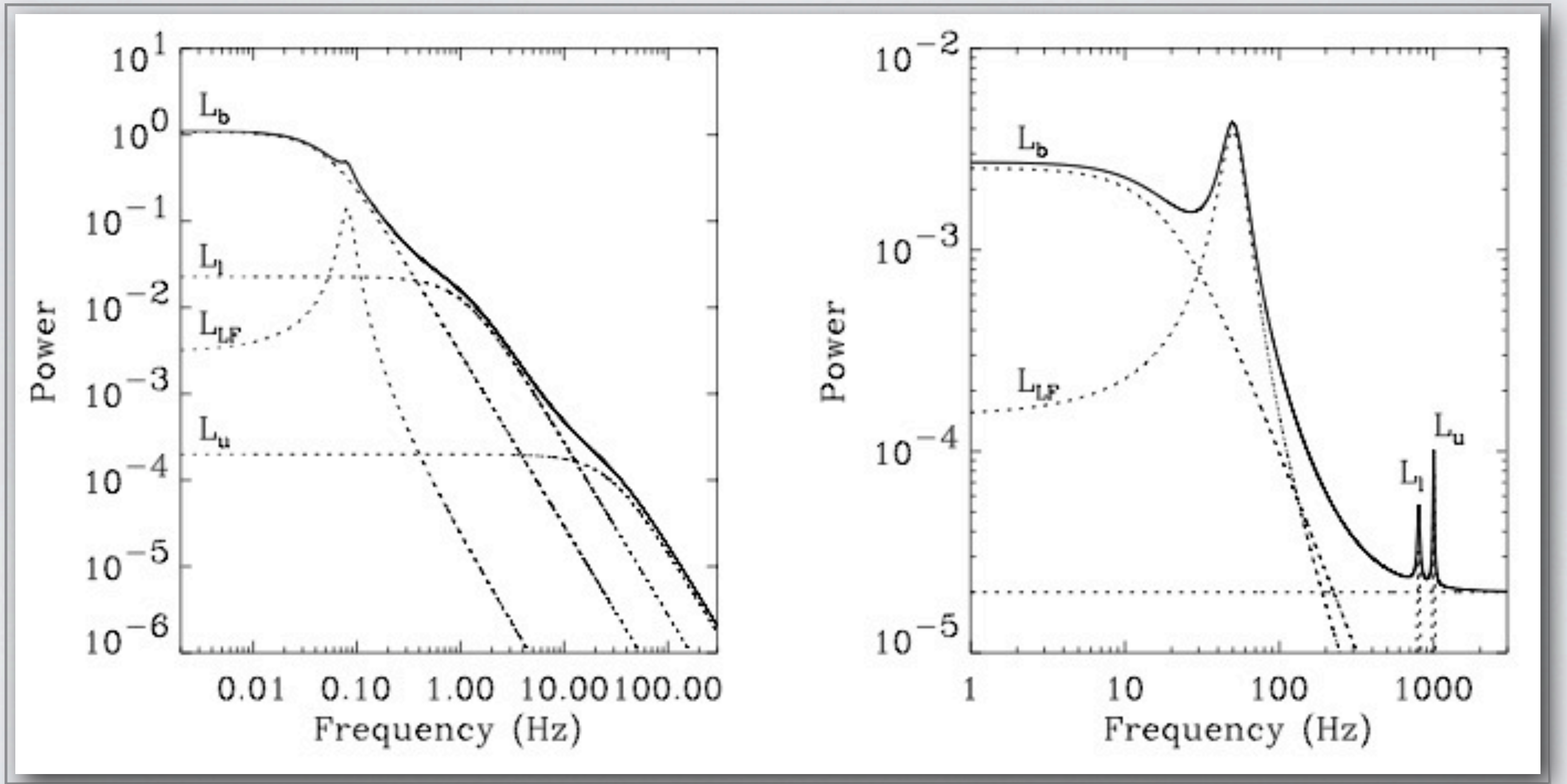
- We can use the peak in νP_ν

$$\nu_{max} = \sqrt{\nu_0^2 + \frac{\Delta}{2}^2}$$



LORENTZIAN DECOMPOSITION

- With these tools we can fit power spectra



NO PHYSICAL BACKING (YET)

- Power spectrum of a damped oscillator
- Also called Cauchy distribution
- Even if it looks like a Lorentzian, it might not be a Lorentzian

DEALING WITH GAPS

- Some solutions are obvious:
  - Welch method (skip gaps)
  - Zero padding (or local average)
- Other methods are available: Lomb-Scargle
  - Good for general uneven sampling
  - Equivalent to linear least-square fit to sin+cos
  - Statistically robust

LOMB-SCARGLE PERIODOGRAM

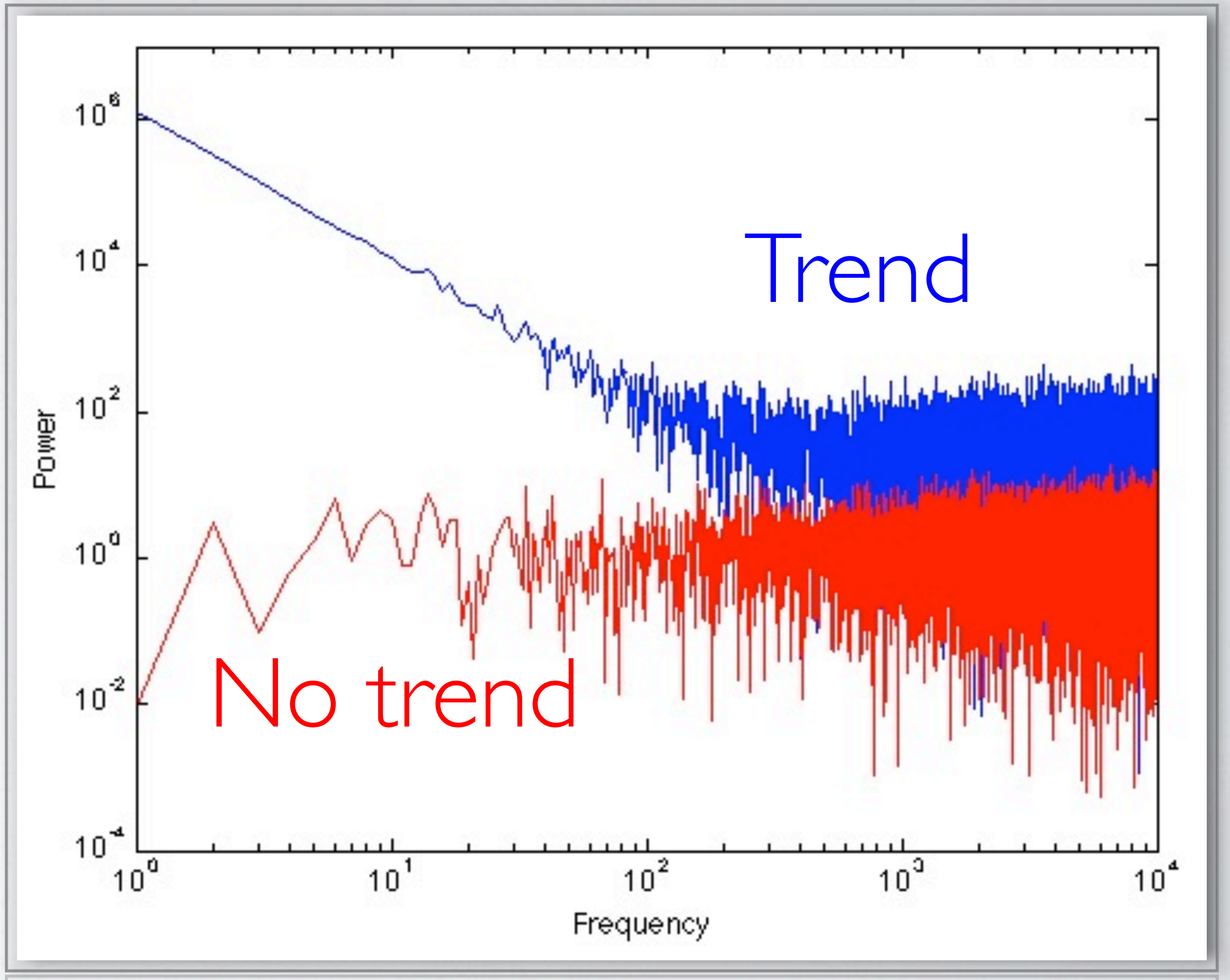
- h_j sampled at t_j

$$P_N(\omega) = \frac{1}{2\sigma^2}\left\{\frac{\left[\sum_j (h_j - \bar{h})\cos\omega(t_j-\tau)\right]^2}{\sum_j \cos^2\omega(t_j-\tau)} + \frac{\left[\sum_j (h_j - \bar{h})\sin\omega(t_j-\tau)\right]^2}{\sum_j \sin^2\omega(t_j-\tau)}\right\}$$

- where: $\tan(2\omega\tau) = \frac{\sum_j \sin 2\omega t_j}{\sum_j \cos 2\omega t_j}$ ensures shift independence
- Powerful method: it can go beyond "Nyquist"

BEWARE OF TRENDS!

- A trend is a modification to the window
- Must be de-trended
- Same about possible drop outs

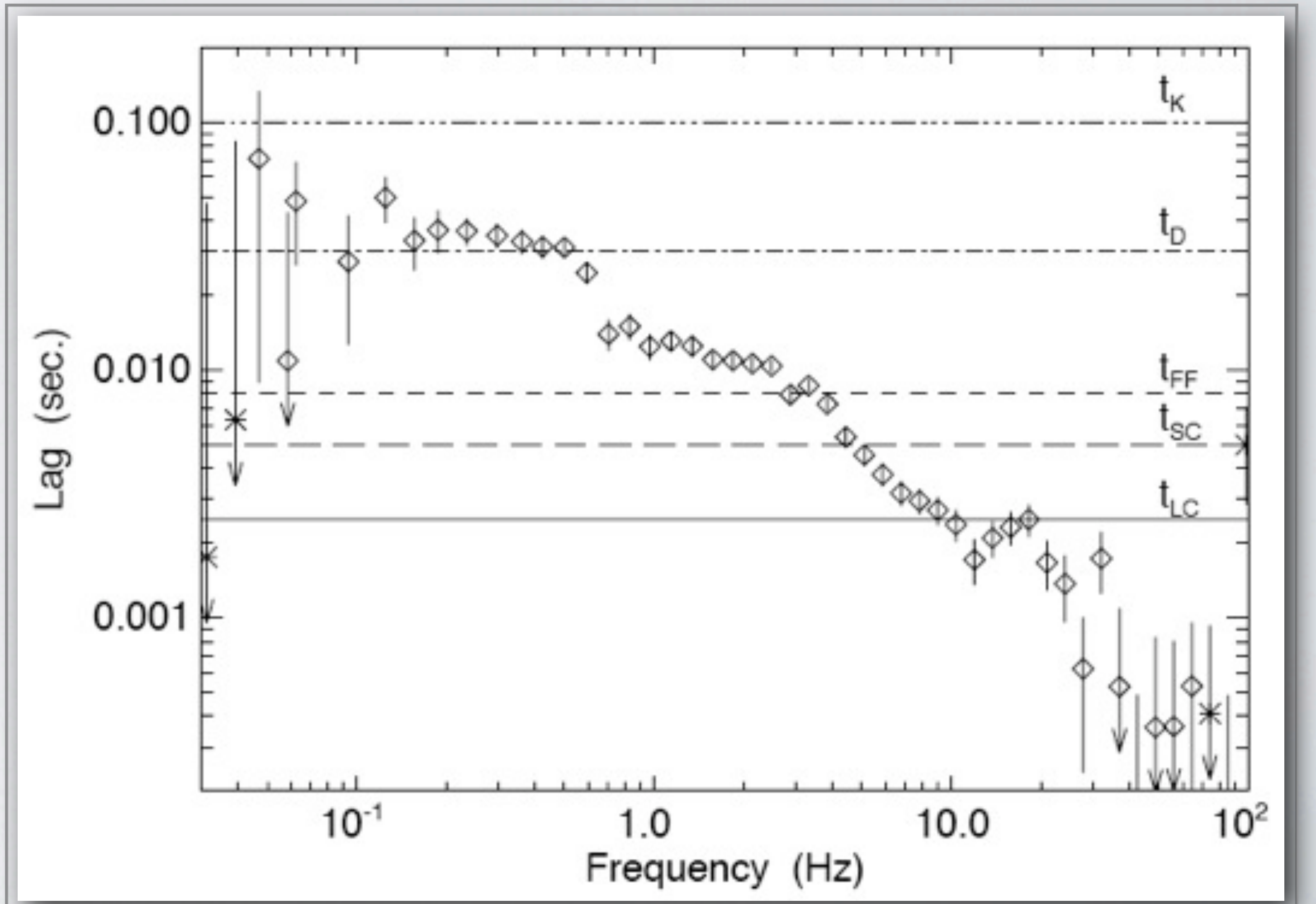


CROSS-SPECTRUM

- Power spectrum: amplitudes of the FFT
- We throw away the phases
- If we take *two* time series *f(t)* & *g(t)*, the phases make more sense
- Cross-spectrum: $C_{f,g}(\nu) = F_f^*(\nu) \times F_g(\nu)$
- If *f*=*g*, it becomes the power spectrum
- What is it useful for?

PHASE/TIME LAGS

- The phases give us the phase delay between the two time series
- Not easy to interpret, can be linked to physical models
- Time lags: phase φ/ν
- Additional technical details (not shown)



AUTO/CROSS-CORRELATION

- The power spectrum is the FT of the autocorrelation

$$Corr(g,g) = \int_{-\infty}^{\infty} g(t+\tau)g(\tau)d\tau \Longleftrightarrow |G(f)|^2$$

- Autocorrelation is real and even, power spectrum is real and even

- The cross spectrum is the FT of the crosscorrelation

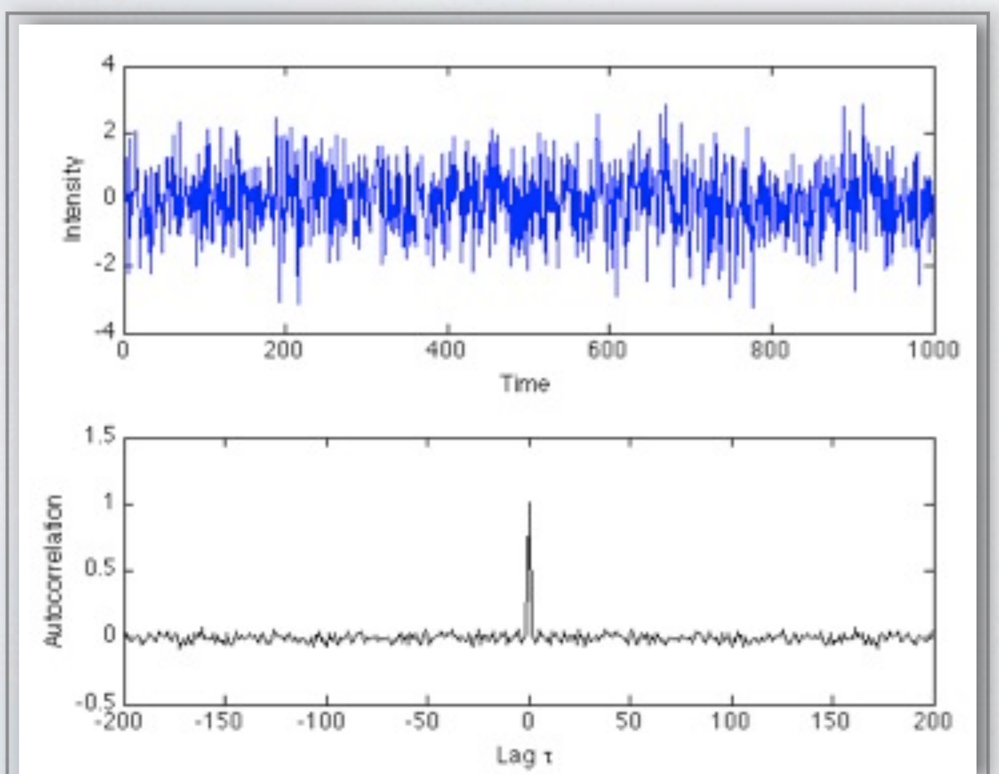
- Power- an cross-spectrum contain more information (if you can afford them because of statistics)

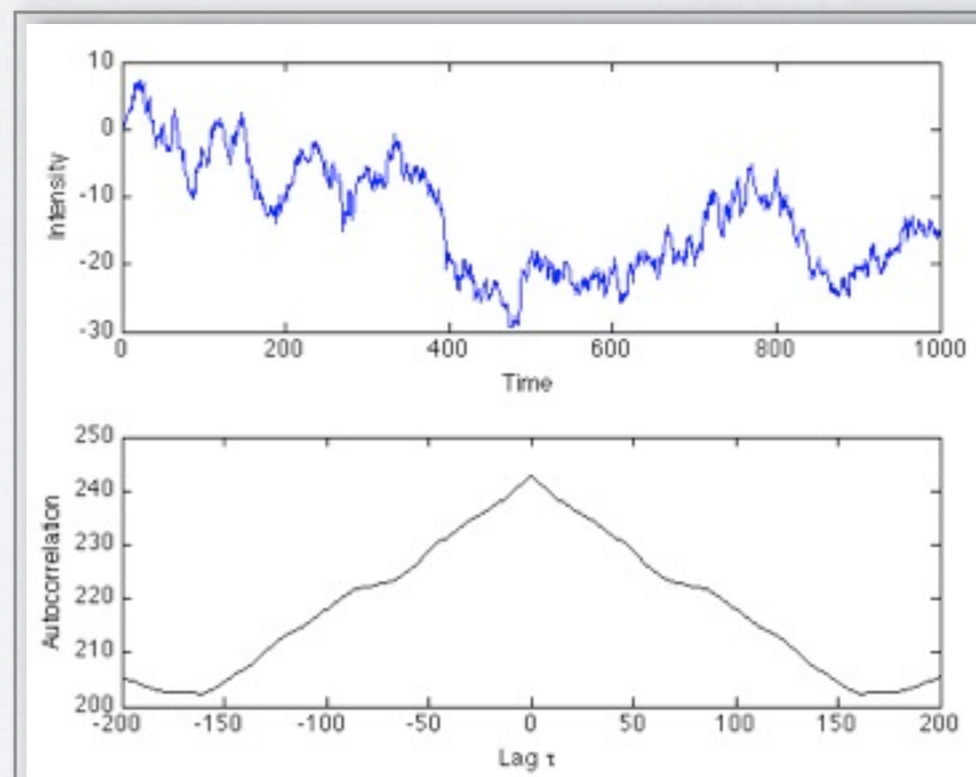
AUTOCORRELATION

- Uncorrelated noise: ACF is zero everywhere but at **τ**=*0* [variance]

$$Corr(g,g) = \int_{-\infty}^{\infty} g(t+\tau)g(\tau)d\tau$$

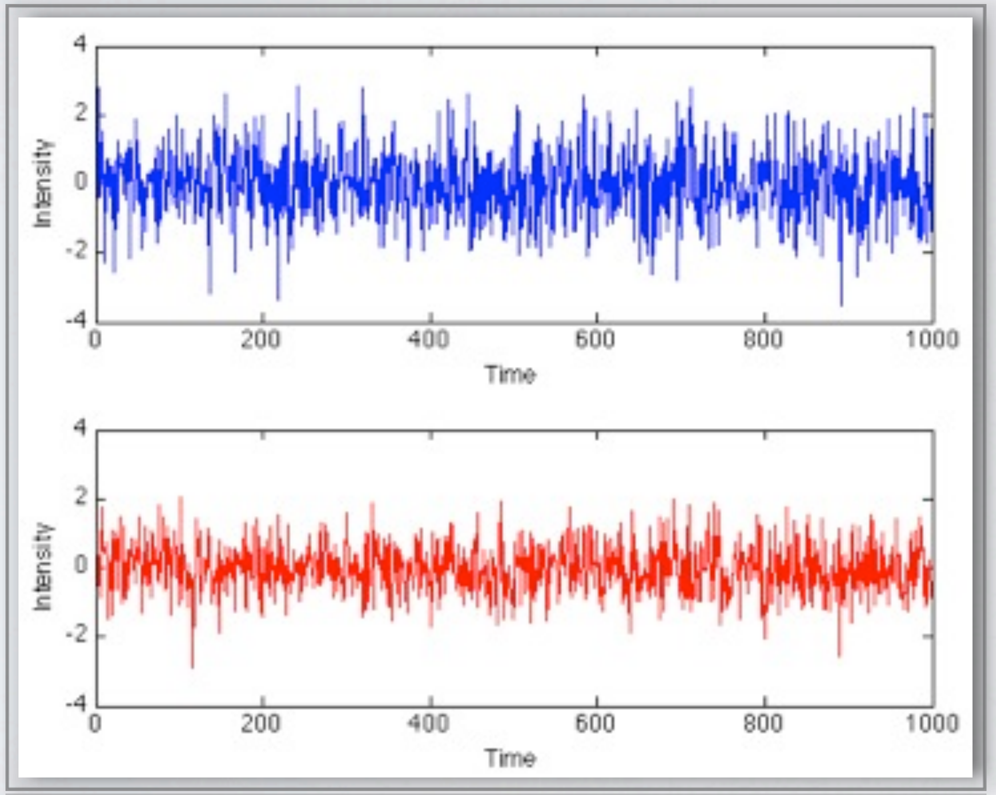
- Biased ACF: dividing by *N*
- Unbiased ACF: dividing by *N*-|*m*|

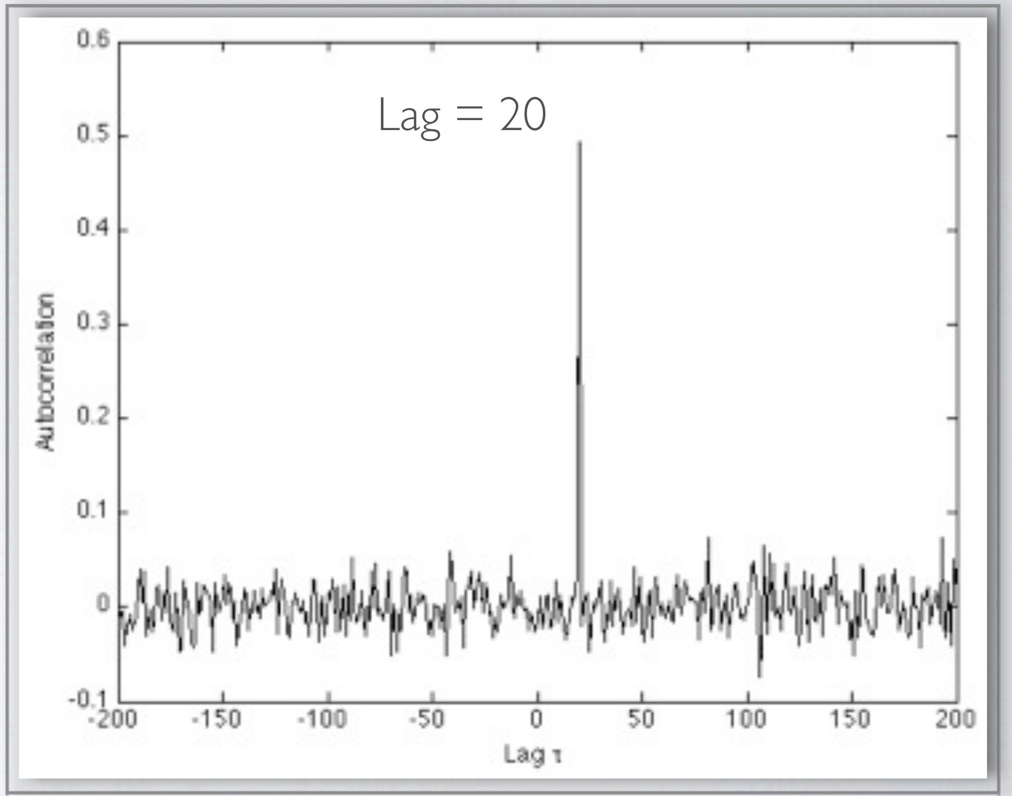




CROSSCORRELATION

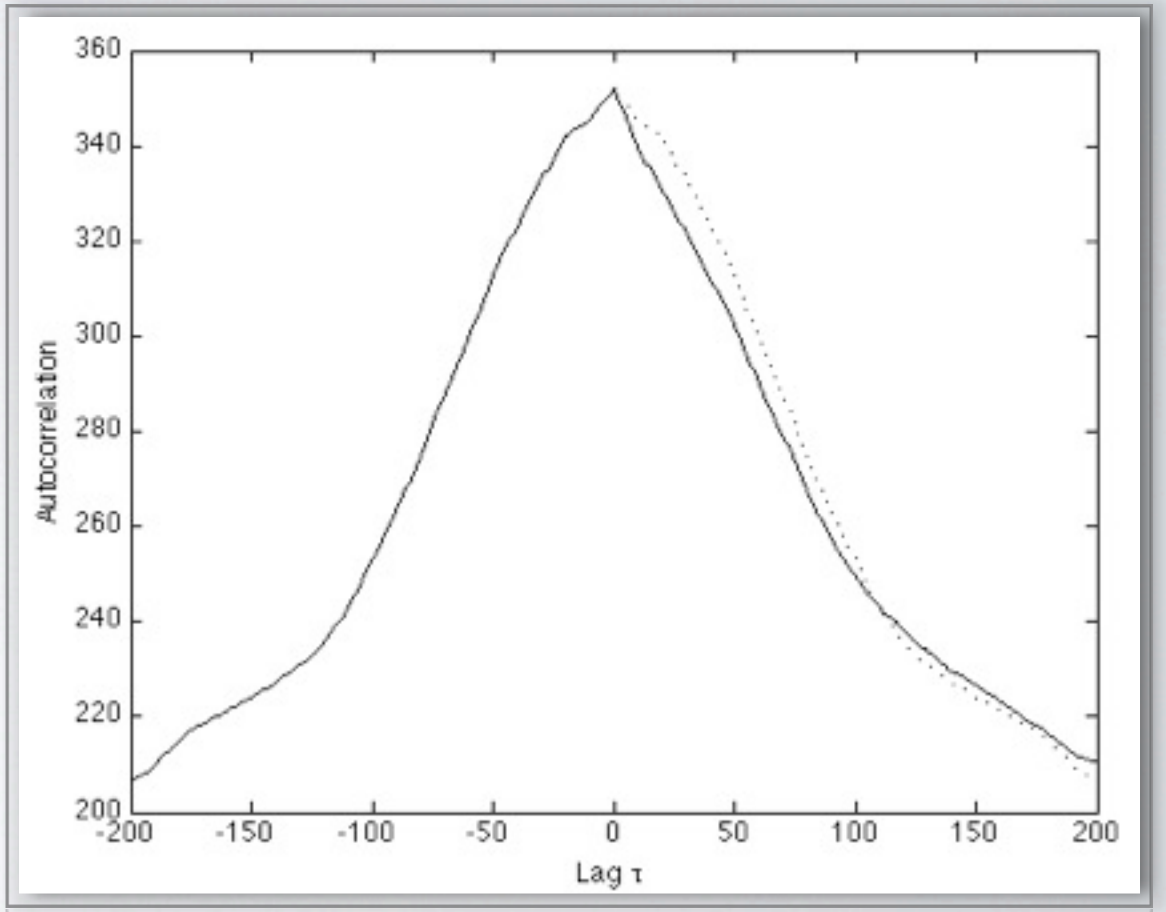
- Uncorrelated series: CCF is zero everywhere
- Simple shift: peak somewhere





WHEN DO I HAVE A LAG?

- *"The CCF peaks at 0, therefore there is no measurable lag"*
- NO!
- CCF is a superposition of sinusoids of different periods
- Any asymmetry implies a lag



COHERENT SIGNALS: BARYCENTRIC CORR.

- The Earth moves and rotates, the satellite also moves
- This has an effect on the period (doppler modulation)..
- .. and on the absolute phase
- Times are corrected to the barycenter of the solar system
- Standard routines and ephemeris
- Not relevant for aperiodic signals

PERIOD FOLDING I: χ^2 TEST

- Photon arrival times t_j
- For trial period produce phases $\phi_j = \mathrm{Frac}\left(\frac{t_j}{P}\right)$
- Put photon in appropriate phase bin
- Test vs. constancy (χ^2)
- If time bins and not times, easy to generalize
- Problem: binning and statistics (few photons?)

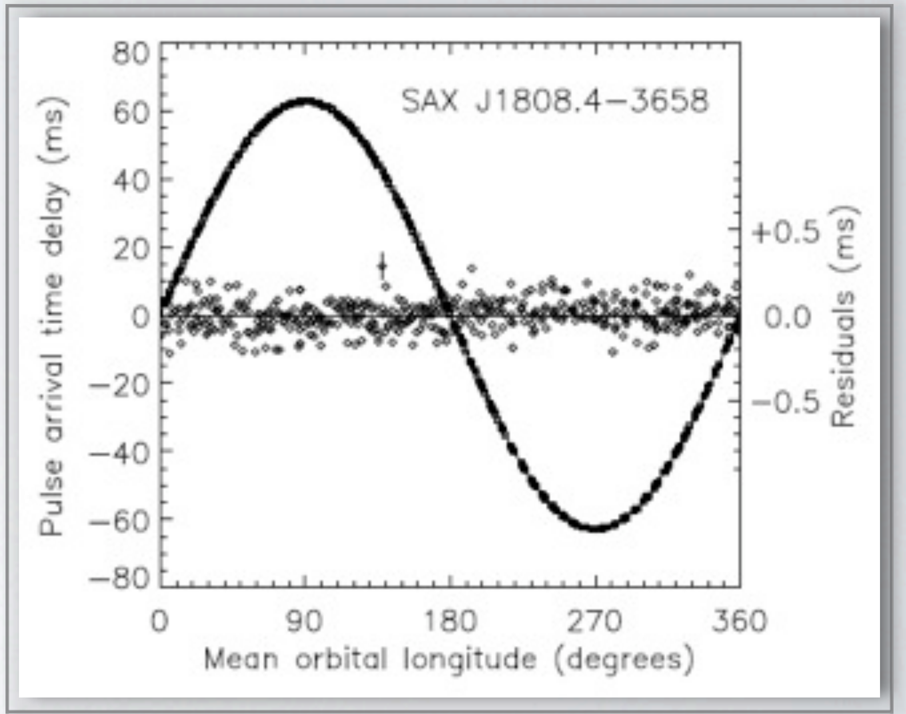
PERIOD FOLDING II: Z^2 TEST

- Photon arrival times t_j
- For trial period produce phases $\phi_j = \mathrm{Frac}\left(\frac{t_j}{P}\right)$
- Compute $$Z_n^2 = \frac{2}{N}\sum_{k=1}^{n}\left[\left(\sum_{j=1}^{N}\cos k\phi_j\right)^2 + \left(\sum_{j=1}^{N}\sin k\phi_j\right)^2\right]$$

  where *n* is the desired number of harmonics
- Z is distributed as a χ^2 with *2n* d.o.f.
- Good for small number of photons [Rayleigh test]

COMPLICATIONS

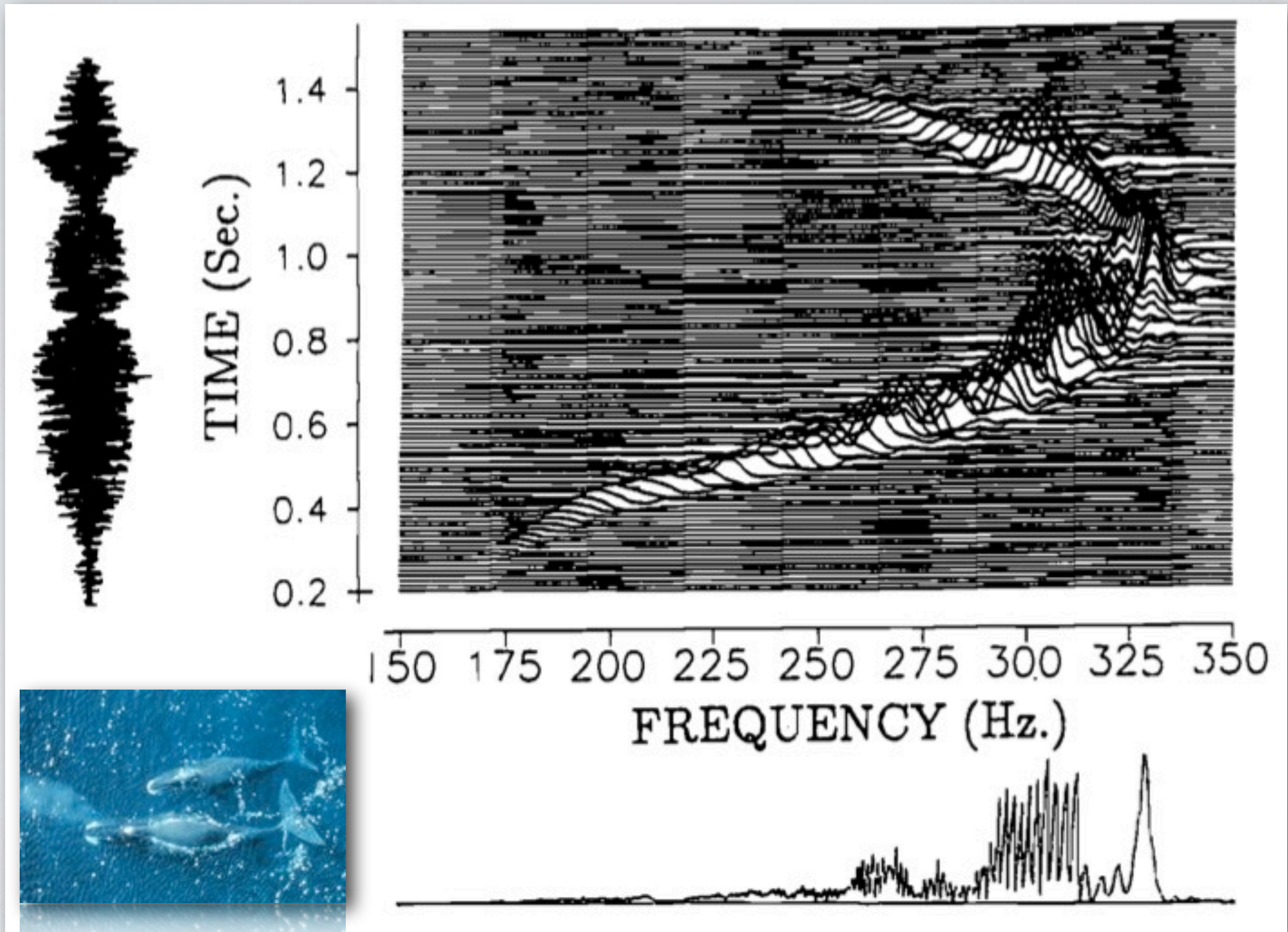
- There can be a significant period derivative
- If your pulsar is in a binary system, there is Doppler effect



- Easy to lose a pulsation
- Power spectrum smeared, folding as well
- Must factorize possible orbit in the solution
- Many free parameters

MORE ON TIME-FREQUENCY

Bowhead whale



ALTERNATIVE TECHNIQUES

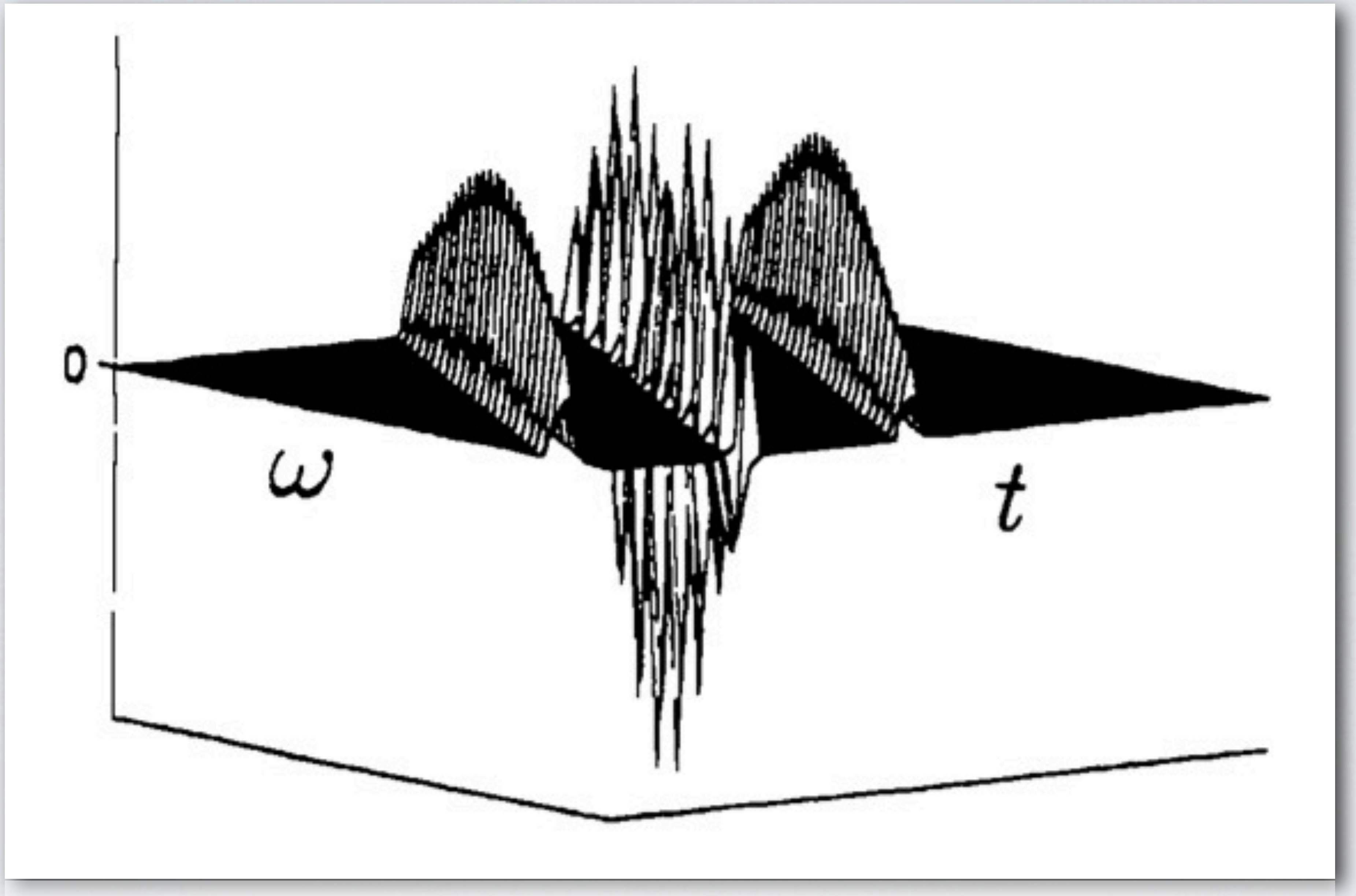
- The Wigner Distribution

$$W(t,\omega) = \frac{1}{2\pi}\int s^*(t - \tfrac{1}{2}\tau)s(t + \tfrac{1}{2}\tau)e^{-i\tau\omega}d\tau$$

- Signal in the past by the signal in the future!
- Problem: only for Gaussian chirps W is everywhere positive
- You can beat the uncertainty principle, but at the cost..
- ... of generatin additional monstruosities

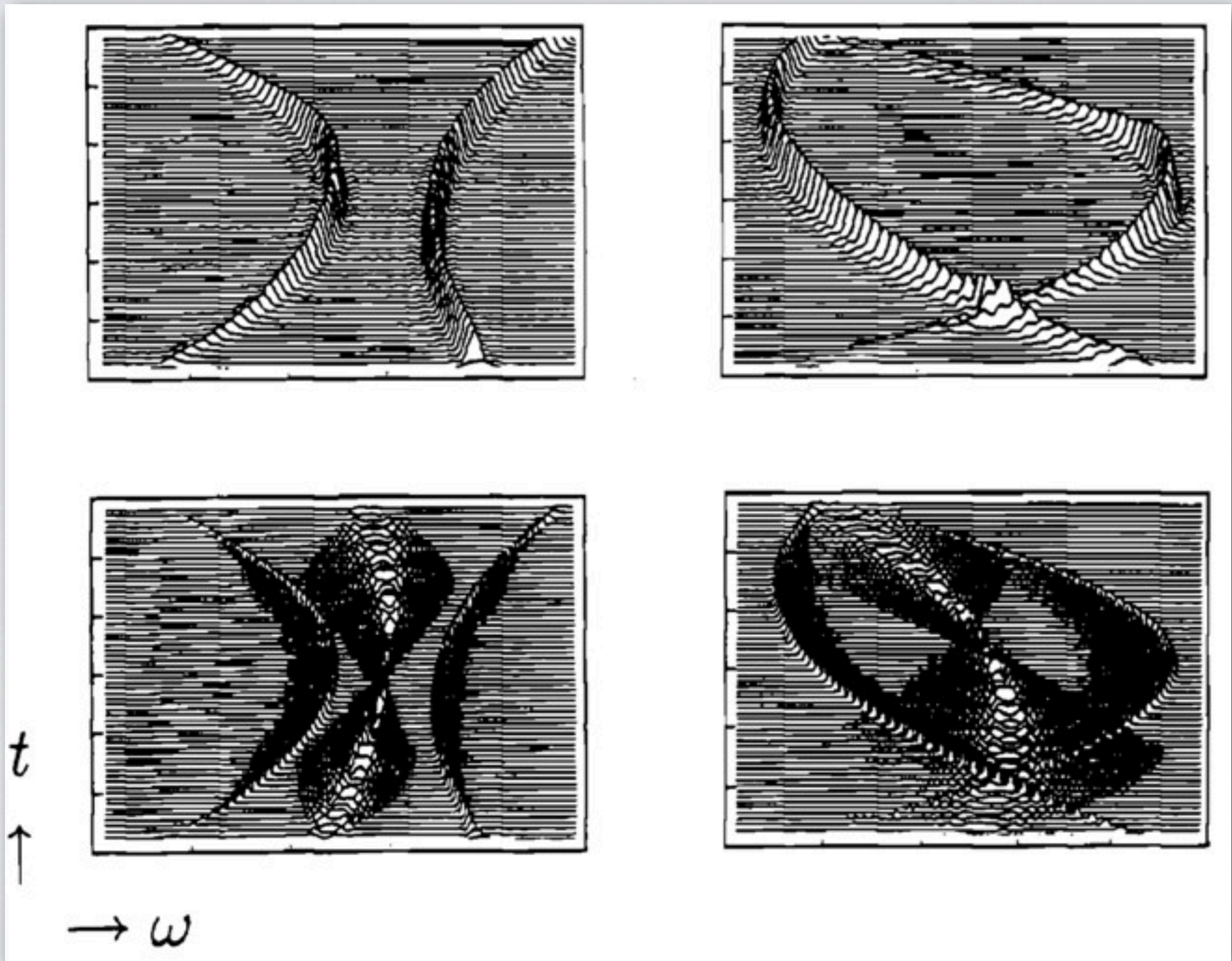
THE WIGNER DITRIBUTION

Signal: sum of two chirps



THE WIGNER DITRIBUTION

Comparison with the spectrogram



COHEN'S KERNEL

ALL time-frequency representations come from:

$$C(t,\omega) = \frac{1}{4\pi^2}\int\int\int s^*(u-\tfrac{1}{2}\tau)s(u+\tfrac{1}{2}\tau)\phi(\theta,\tau)e^{-i\theta t-i\tau\omega+i\theta u}du\,d\tau\,d\theta$$

Where $\varphi(\theta,\tau)$ is the kernel

Changing the kernel you change the representation

The properties of the representation depend on the proprties of the kernel

Name	Kernel: $\phi(\theta,\tau)$	Distribution: $C(t,\omega)$
General class (Cohen[125])	$\phi(\theta,\tau)$	$\frac{1}{4\pi^2}\iiint e^{-j\theta t-j\tau\omega+j\theta u}\,\phi(\theta,\tau)\, s^*(u-\tfrac{1}{2}\tau)\,s(u+\tfrac{1}{2}\tau)\,du\,d\tau\,d\theta$
Wigner[584]	1	$\frac{1}{2\pi}\int e^{-j\tau\omega} s^*(t-\tfrac{1}{2}\tau)\,s(t+\tfrac{1}{2}\tau)\,d\tau$
Margenau-Hill[358]	$\cos\tfrac{1}{2}\theta\tau$	$\mathrm{Re}\,\frac{1}{\sqrt{2\pi}}\,s(t)S^*(\omega)\,e^{-jt\omega}$
Kirkwood[305] Rihaczek[484]	$e^{j\theta\tau/2}$	$\frac{1}{\sqrt{2\pi}}\,s(t)S^*(\omega)\,e^{-jt\omega}$
Born-Jordan[1] (Cohen[125])	$\frac{\sin\frac{1}{2}\theta\tau}{\frac{1}{2}\theta\tau}$	$\frac{1}{2\pi}\int\frac{1}{\lvert\tau\rvert}e^{-j\tau\omega}\int_{t-\lvert\tau\rvert/2}^{t+\lvert\tau\rvert/2} s^*(u-\tfrac{1}{2}\tau)\,s(u+\tfrac{1}{2}\tau)\,du\,d\tau$
Page[419]	$e^{j\theta\lvert\tau\rvert}$	$\frac{\partial}{\partial t}\left\lvert\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} s(t')\,e^{-j\omega t'}dt'\right\rvert^2$
Choi-Williams[117]	$e^{-\theta^2\tau^2/\sigma}$	$\frac{1}{4\pi^{3/2}}\iint\frac{1}{\sqrt{\tau^2/\sigma}}\,e^{-\sigma(u-t)^2/\tau^2-j\tau\omega}\, s^*(u-\tfrac{1}{2}\tau)\,s(u+\tfrac{1}{2}\tau)\,du\,d\tau$
Spectrogram	$\int h^*(u-\tfrac{1}{2}\tau)\,e^{-j\theta u}\,h(u+\tfrac{1}{2}\tau)\,du$	$\left\lvert\frac{1}{\sqrt{2\pi}}\int e^{-j\omega\tau}\,s(\tau)h(\tau-t)\,d\tau\right\rvert^2$
Zhao-Atlas-Marks[626]	$g(\tau)\,\lvert\tau\rvert\,\frac{\sin a\theta\tau}{a\theta\tau}$	$\frac{1}{4\pi a}\int g(\tau)\,e^{-j\tau\omega}\int_{t-\lvert\tau\rvert a}^{t+\lvert\tau\rvert a} s^*(u-\tfrac{1}{2}\tau)\,s(u+\tfrac{1}{2}\tau)\,du\,d\tau$
Positive[2] (Cohen, Posch, Zaparovanny[127, 128])	see Chapter 14	$\lvert S(\omega)\rvert^2\,\lvert s(t)\rvert^2\,\Omega(u,v)$

A WORD ON WAVELETS

- The resolution element in time-frequency cannot be made smaller than the minimum

- However, there is no reason why it should be of the same shape

- You can adapt it to what you need where you need it