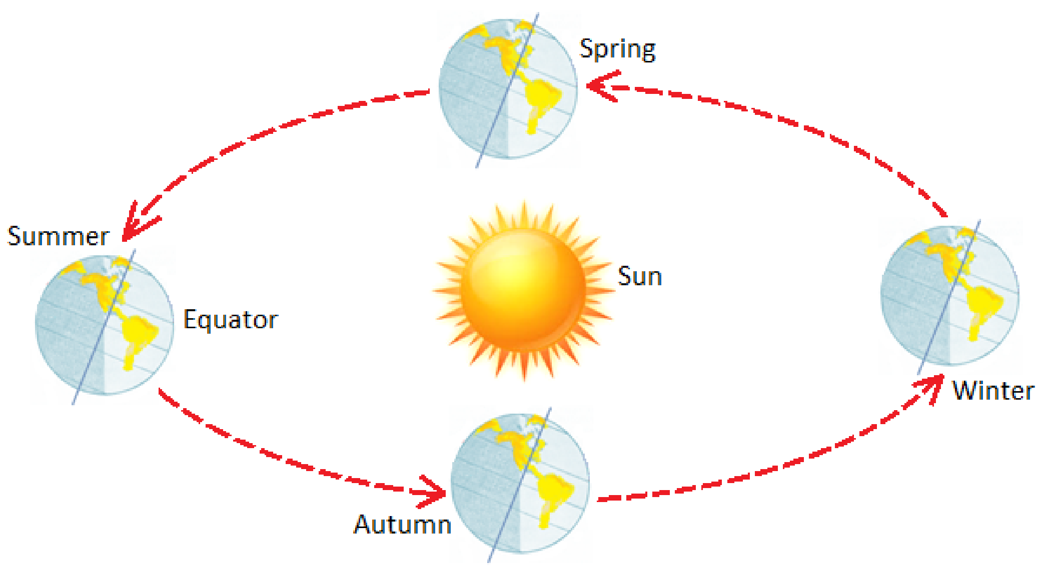
Mathematical description of variability

1. Motivation

As I mentioned already several time before, the astronomical sources do very. We also talked about the characteristic timescales and how they help to determine the properties of the radiating source. As the measurement of the distance is inherently difficult, the measurement of the variability can frequently help.Looking at variability we can determine for example the period of the binary system, and with its help, the masses of the stars. We talked how variability allows to measure the size of the radiating region and the efficiency of accretion. The data at our disposal in increasingly more complex, and we need tools appropriate for determination of the requested information, and the tools have to be adjusted to the character of the variability of a given source and requested information.

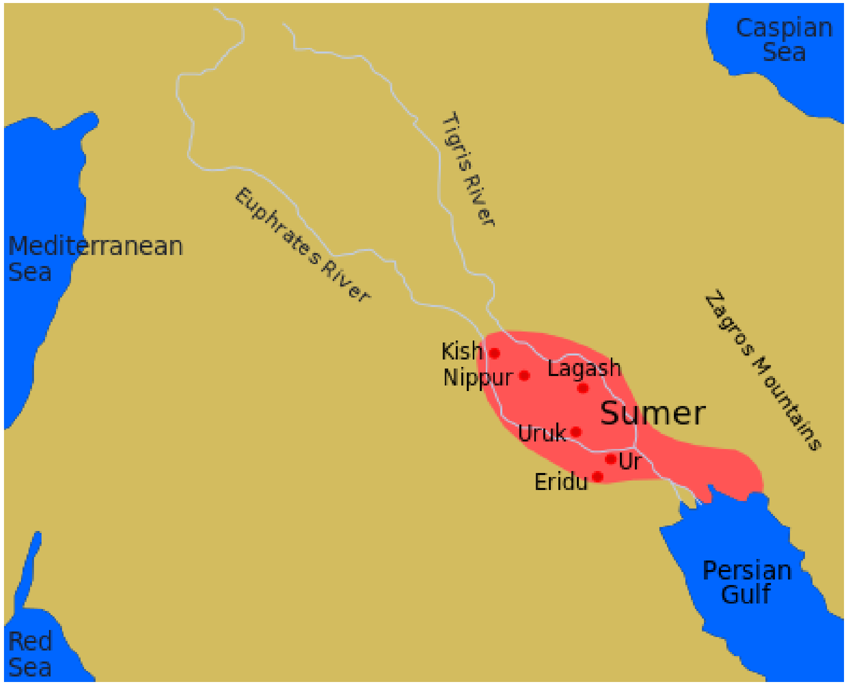
2. Periodic variability

This is the oldest of the patterns of interest to astronomy. Days/nights, then months, seasons and years are all related to the periodic aspects of the motion of Earth with respect to Sun.



2. Periodic variability

This aspect of the changes of the light is of crucial importance for our life, and precise predictions of seasons were of crucial importance for people when they started to develop agriculture. Studying periodic changes has lead to the development of astronomy as applied science thousands of years ago.

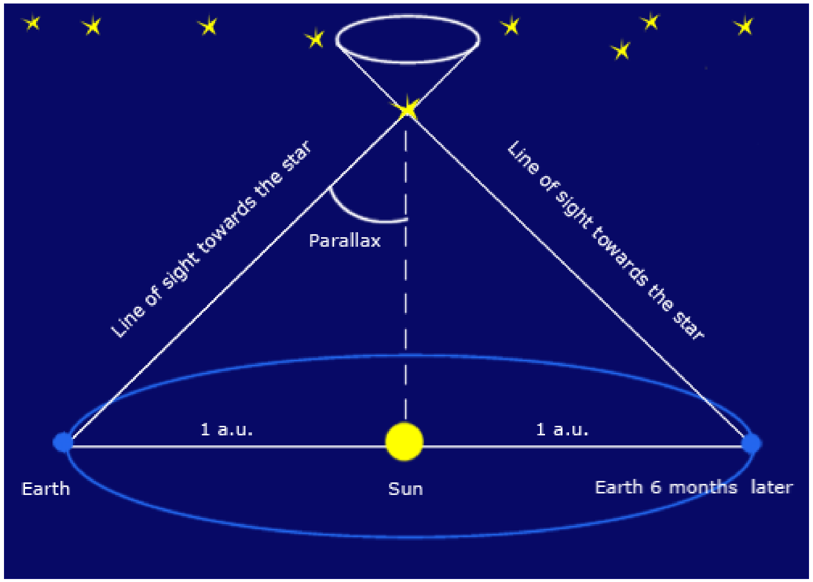


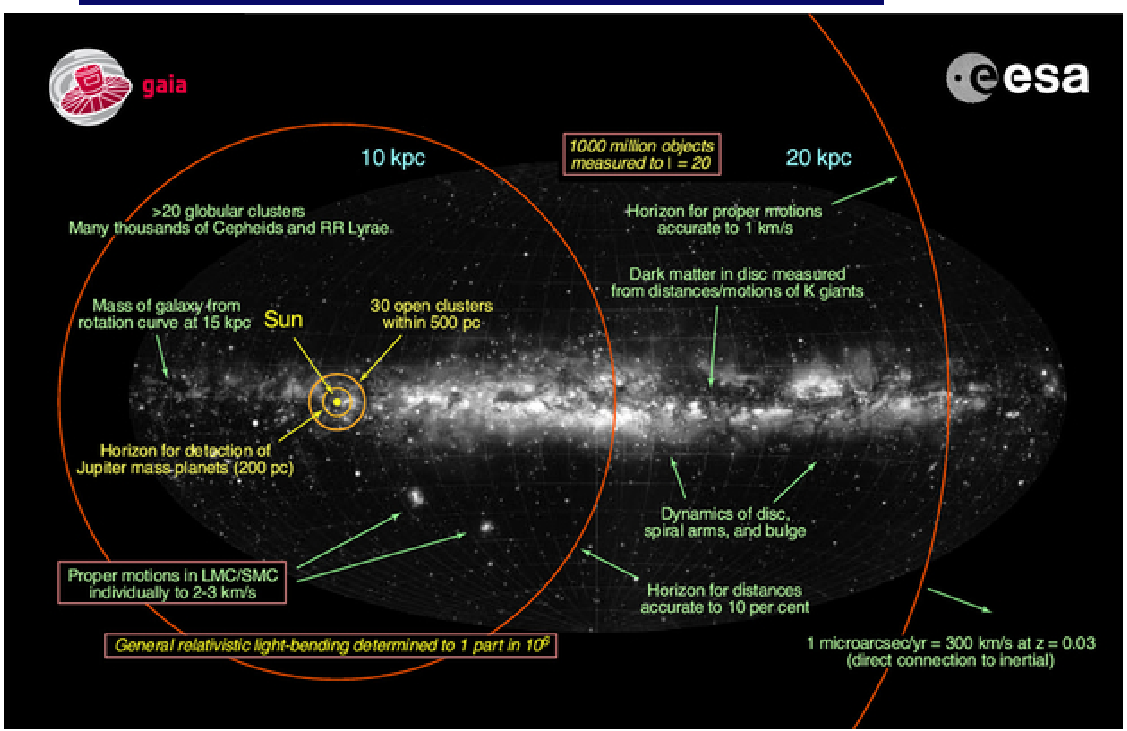
CALENDAR, Sumer culture, around 3000 BC

https://www.sutori.com/item/this-is-a-ancient-calendar-sumerian-s-used-to-tell-the-date-and-time-of-year-ju

2. Periodic variability

Even nowadays studying periodic signals is a very important branch of astronomy. Most of the direct distance measurements in astronomy are bsed on periodic motion (paralax).



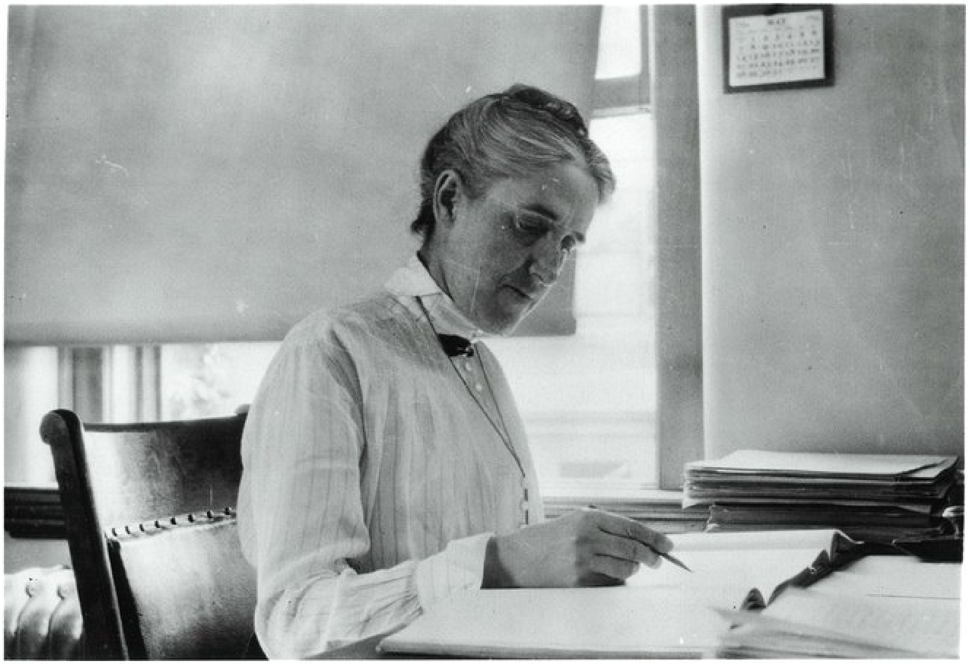


- Now we know the distance Earth-Sun with high precision, including ellipticity etc.
- The distance to nearby stars are thus measured as paralax due to the Earth yearly travel around the Sun
- 1 parsec is a basic distance unit in astronomy

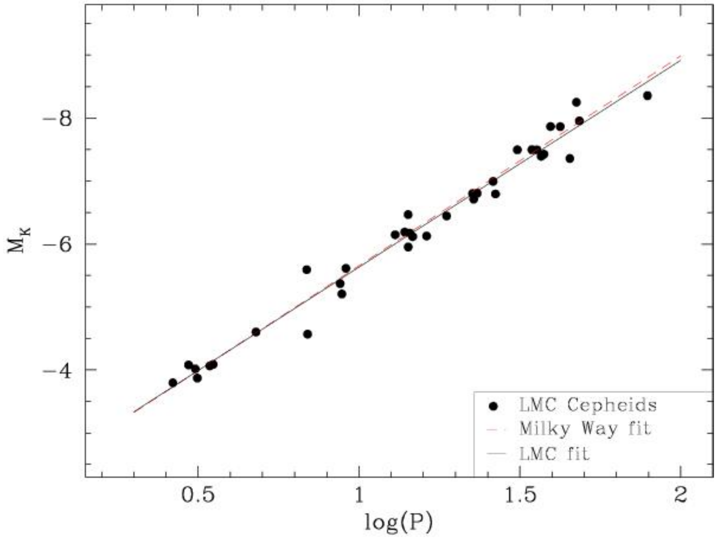
- High accuracy has been occasionally reaches with HST
- Currently with Gaia mission 10 per cent accuracy of paralaxy measurements (for bright stars) up to 10 kpc

2. Periodic variability

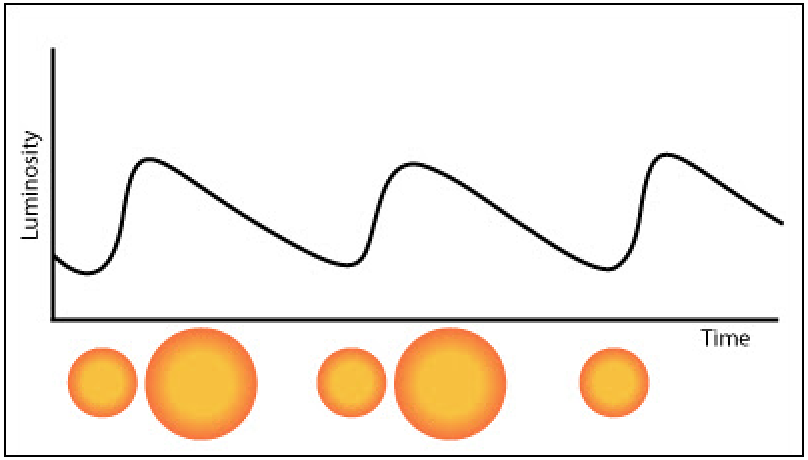
- With paralax method we now directly reach Cepheid stars
- Currently 15 Cepheids have HST paralaxes
- This allows to use Leavitt law (period-luminosity relation)
- From Cepheids we nowadays go directly to SN Ia
- Currently 19 galaxies which host SN Ia and well studies Cepheids (Riess et al. 2016)
- These SN Ia can now be used to callibrate distance-luminosity relation for SN Ia



Henrietta Swan Leavitt (1868-1921)



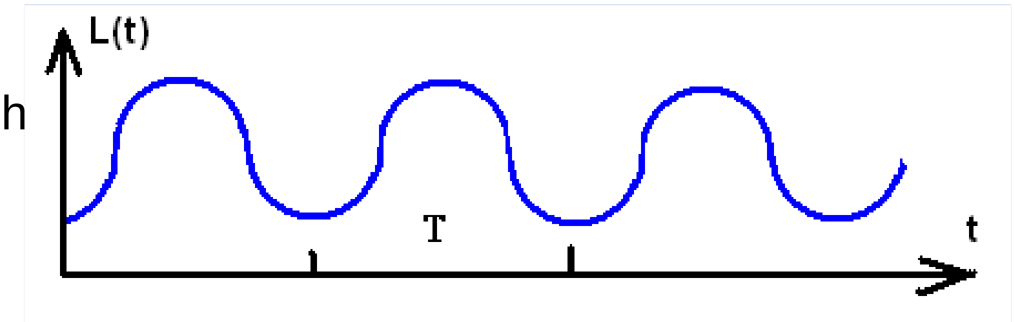
Modern plot of Period-luminosity relation for Milky Way and LMC Cepheids (Storm...Pietrzynski… 2011)

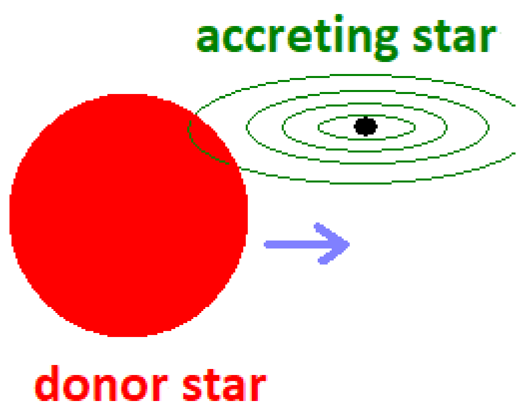


Pulsations of Cepheid stars

2. Periodic variability

Mathematically, the simpleast case of a periodic lightcurve is a single sinusoid. Such case happens in nature (approximately) in accreting binaries, when we look at the radial velocity curve.



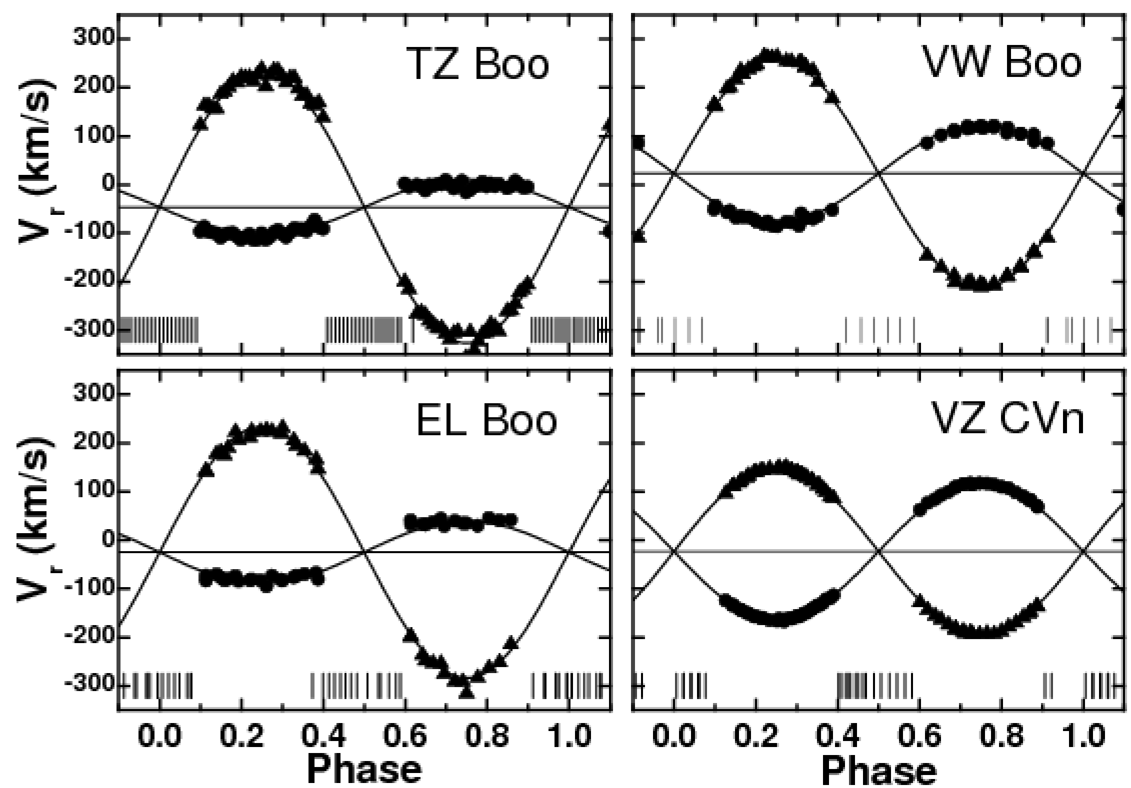


Mathematically, this function is given by

$$F(t) = A\sin(\omega t + \phi)$$

and the period is given by

$$T = \frac{2\pi}{\omega}$$



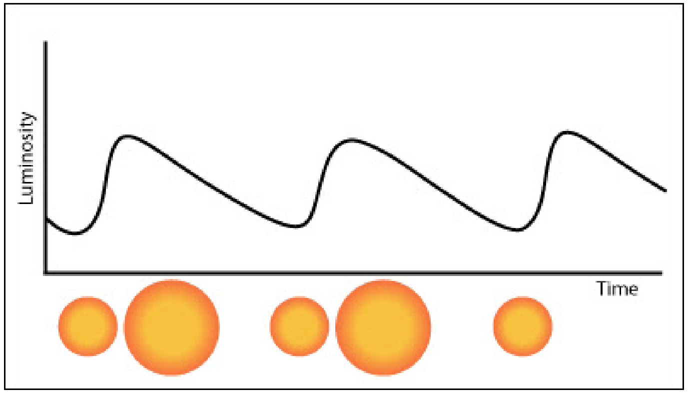
Pribulla et al. (2009), radial velocities for selected close binary contact systems (white dwarfs).

2.1 Periodic variability: Fourier series

In general, periodic functions do not have to look like a single sinusoid. Nevertheless, each periodic function having the property

$$F(t+T)=F(t) \qquad \text{for every value of T}$$

can be expressed as a sum of sinusoids



Pulsations of Cepheid stars

$$F(t)=\sum_{n=1}^{n=+\infty} A_n \sin(\omega n t+\phi_n)$$

with appropriate normalizations and phases. Here $T=\dfrac{2\pi}{\omega}$

The coefficients of the decomposition can be easily determined if we know the function F(t), but the elegant expression can be achieved starting from different notation:

$$F(t)=\sum_{n=1}^{n=+\infty} [a_n \cos(\omega n t)+b_n \sin(\omega n t)]$$

We then get the coefficients by integration:

$$a_n=\frac{2}{T}\int_{-T/2}^{T/2} F(t)\sin(\omega n t)\,dt$$

$$b_n=\frac{2}{T}\int_{-T/2}^{T/2} F(t)\sin(\omega n t)\,dt$$

Note that the operation in Fourier series between time and frequency is not symmetric.
This is because time is a continuous variable while the required frequencies are not.
This is because the function is strictly periodic.

2.1 Periodic variability: Fourier series

Instead of real functions sin and cos, it is convenient to use the complex notation

exp(ix) = cos x + i sin x

And in this case the Fourier series can be expressed in a more compact form as

$$F(t) = \sum_{n=-\infty}^{n=+\infty} c_n \exp(i\omega n t) \qquad c_n = \frac{2}{T}\int_{-T/2}^{T/2} F(t)\exp(i\omega n t)\,dt$$

3. Aperiodic variability: Fourier transform

If the considered function is not strictly periodic, then the new notation is particularly useful, and the symmetry between the frequency domain and the time domain becomes obvious

$$\hat{F}(\omega) = \int_{-\infty}^{\infty} F(t)\exp(-i\omega t)\,dt \qquad F(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{F}(\omega)\exp(i\omega t)\,dt$$

Sometimes the normalization is also adjusted to complete the symmetry, with the factor $\frac{1}{\sqrt{(2\pi)}}$ in front in both cases.

Any (regular) function can be Fourier-tranformed back and through, the information is not lost, both F(t) and $\hat{F}(\omega)$ contain the full information.

3. Aperiodic variability: Fourier transform

Conditions for the use of the Fourier transform: $\int_{-\infty}^{\infty} |L(t)|\, dt < \infty$

This means that the function must be 'localized' in time.

$$\hat{F}(\omega) = \int_{-\infty}^{\infty} F(t) \exp(-i\omega t)\, dt \qquad F(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{F}(\omega) \exp(i\omega t)\, dt$$

Fourier transform of a real function is a complex function (contains amplitude and phase). Frequently we are not that much interested in phase but we want to now the dominant periodicities, so we are interested in the power spectrum density (PSD)

$$P(\omega) = |\hat{F}(\omega)|^2$$

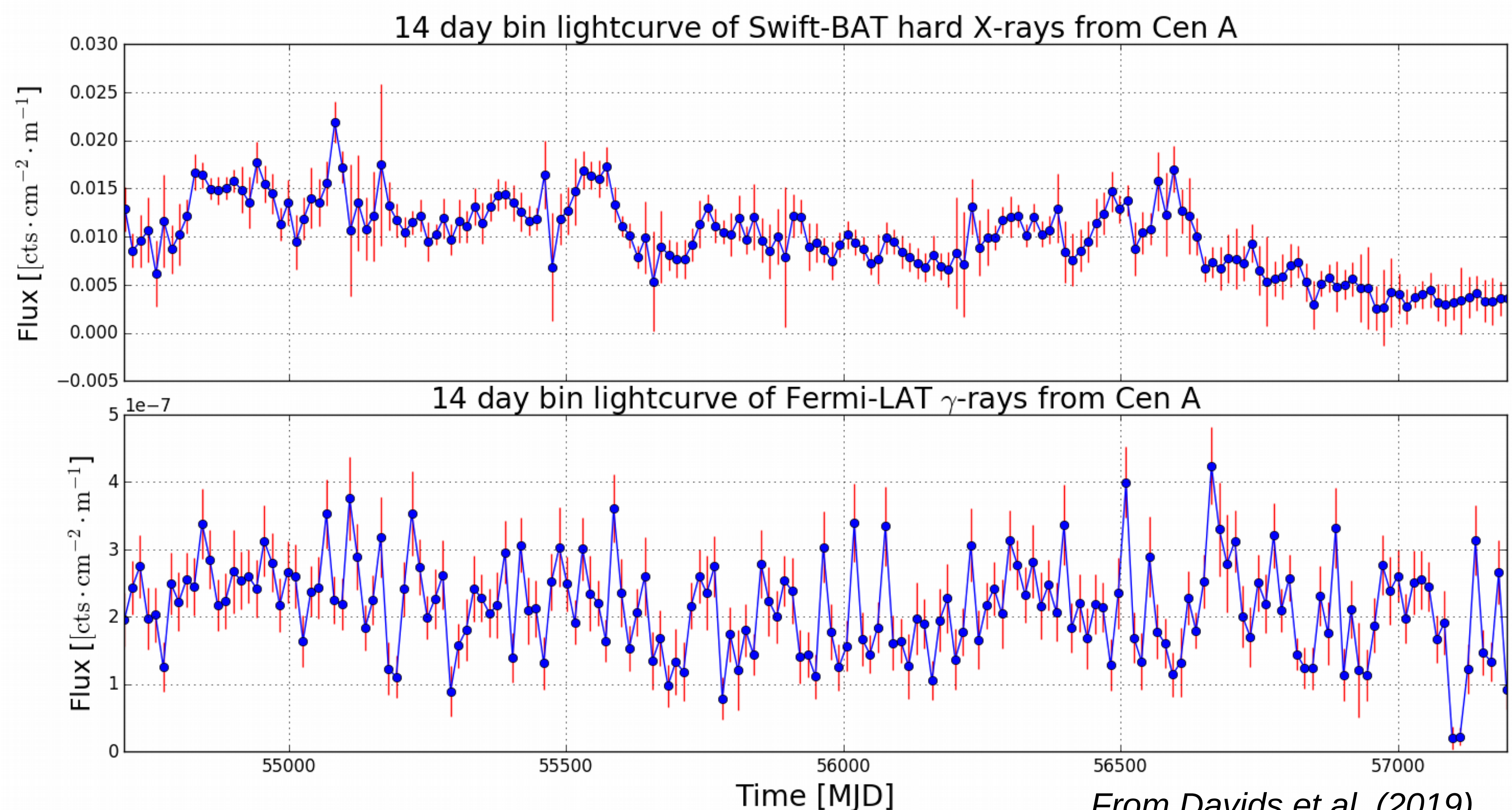
If we try to calculate the PSD for a single sinusoidal function, we will get

$$P(\omega) \propto \delta(\omega - \omega_0)$$

4. Discreet time-limited signals

The problems start when we confront the theory with reality. We do not have continuous lightcurves covering time in an arbitrarily dense way. In practise, at best we have a finite lightcurve with a constant step (e.g. in X-ray data), or we have a finite lightcurve with non-uniform time coverage.

4. Discreet time-limited signals



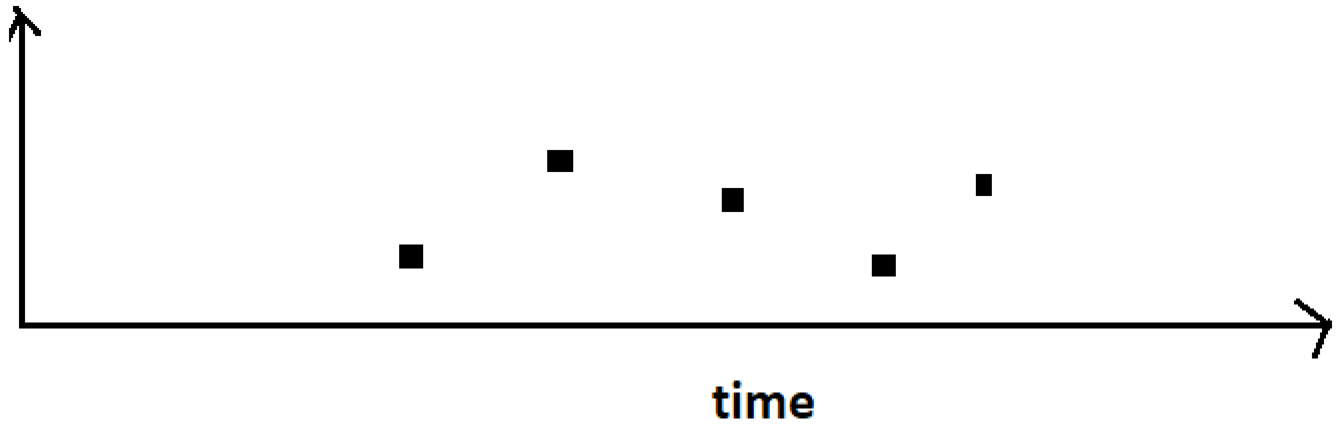
From Davids et al. (2019)

The source of the problems is our limited knowledge about a function: we know its behavior only between t_{start} and t_{end}, and we do not know how it behaves ***BETWEEN*** the points.

If we now do some transformation from the time to the frequency domain and back, the uniqueness of the procedure is not guaranteed.

4. Discreet time-limited signals

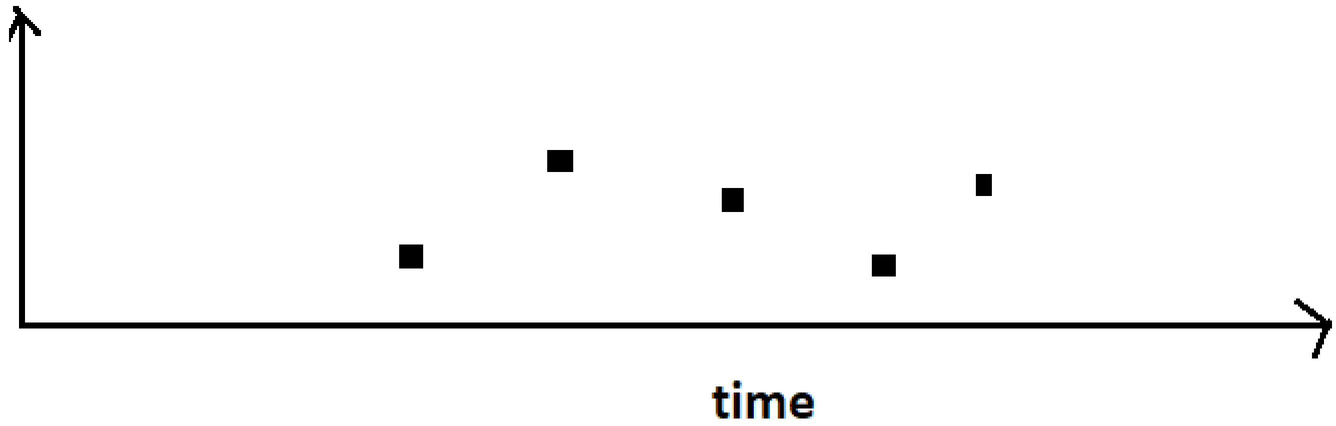
For example, if this is your lightcurve:



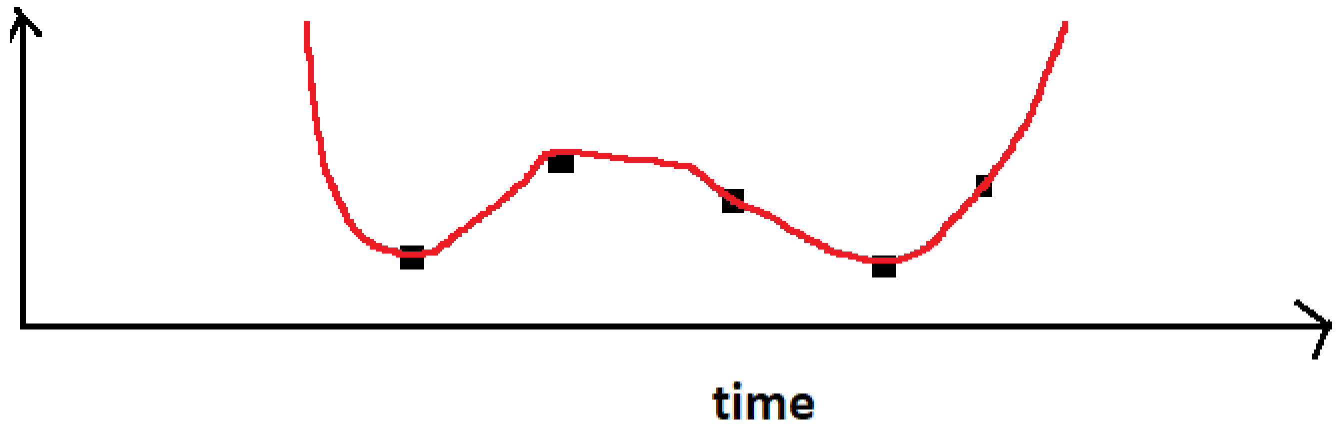
then after going to frequency space through Fourier transfor and back you might get

4. Discreet time-limited signals

For example, if this is your lightcurve:

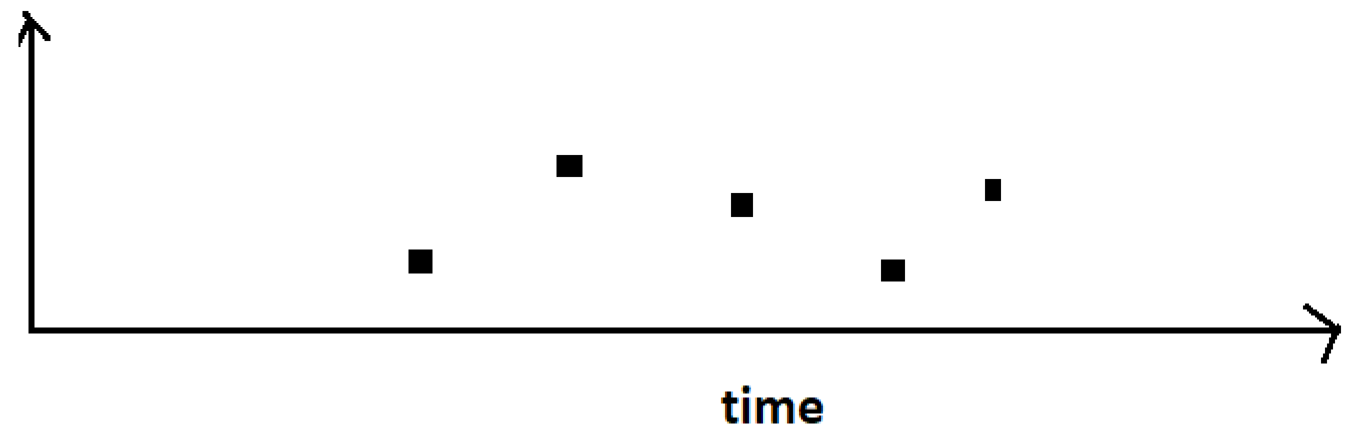


then after going to frequency space through Fourier transform and back you might get a new continuous lightcurve like that

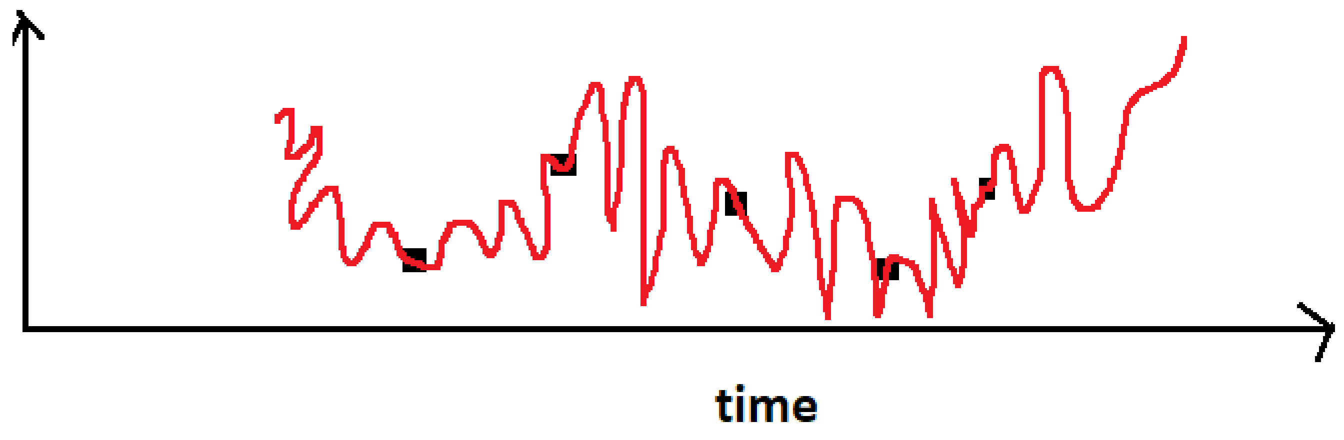


4. Discreet time-limited signals

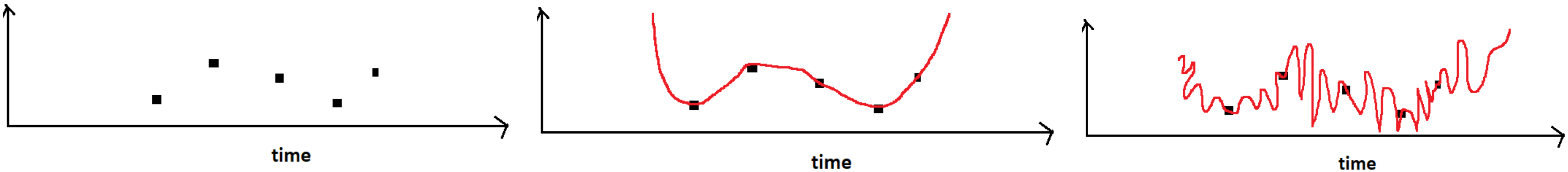
For example, if this is your lightcurve:



then after going to frequency space through Fourier transform and back you might get a new continuous lightcurve even like that



4. Discreet time-limited signals



There are ways to prevent basic problems like that, but they are never perfect, and this is a complex issue in signal processing. Aliases, window functions, power leaking are just a few of terms in this context.

Experts in time-domain analysis at CAMK:
Alex Schwarzenberg-Czerny
Alex Markowitz

In this lecture I will mostly qualitatively touch issues which are particularly important in the context of accreting sources. I will always assume that the signal is given in a limited number of time points, covering a limited time span.

5. Looking for a period

If the curve is likely periodic (brigthness of a binary star) we most frequently use two options

- Periodogram
- Epoch folding

5.1 Periodogram is constructed in the following way:

$$P(\omega)=\frac{1}{N}\left|\sum_{j=1}^{N} F(t_j)\exp(-i\,\omega\,t_j)\right|^2$$

Here we do not request that the separation between the points is the same, we just take whatever we have. For the frequency range, we should not go to frequencies smaller than

$T_{total}=\frac{2\pi}{\omega_{min}}$ Here T_{total} is the total duration of the lightcurve. Also the probes frequencies should not be higher than the smallest time separation between the two time points

$$min(\Delta t_{j,j+1})=\frac{2\pi}{\omega_{max}}$$

Most frequently, instean of a cyclic frequency ω we use the frequency f defined as

$$f=\frac{\omega}{2\pi}$$

And then the periodogram is defined as

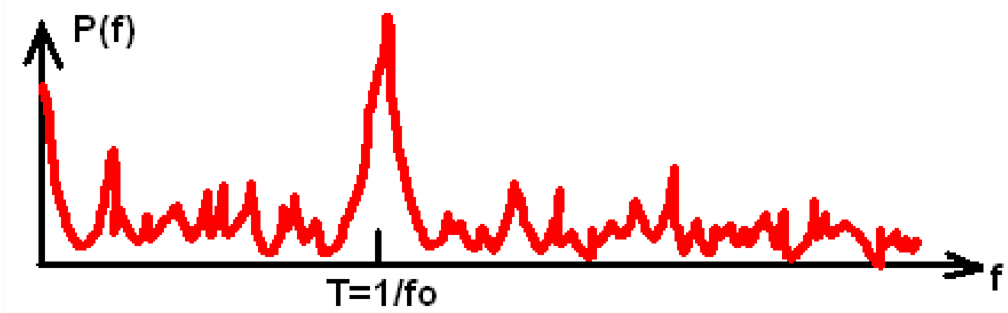
$$P(f)=\frac{1}{N}\left|\sum_{j=1}^{N} F(t_j)\exp(-2\pi\,i\,\omega\,t_j)\right|^2$$

Be carefull and always check, whether the authors use ω or f, the difference is by a factor 6!

5.1 Looking for a period - periodogram

You get something like that:

$$P(f)=\frac{1}{N}\left|\sum_{j=1}^{N} F(t_j)\exp(-2\pi i\omega t_j)\right|^2$$



And the high peak implies the searched period T.
However, note that on one hand, it is not obvious which frequency grid to use, on the other hand, you cannot get more numbers than you actually measure (N). This makes the estimate of the peak determination accuracy quite complicated.

In practise, instead of a standard periodogram we mostly use its modified version, known as Lomb–Scargle periodogram, or or Least-squares spectral analysis. The difference is in optimization for the phase at each frequency first, and then computing the coefficients.

$$A_n\sin(\omega n t+\phi_n)$$

Computing Lomb-Scarge periodogram is equivalent to the fitting of a sinusoid to all the tested points using χ^2 method.

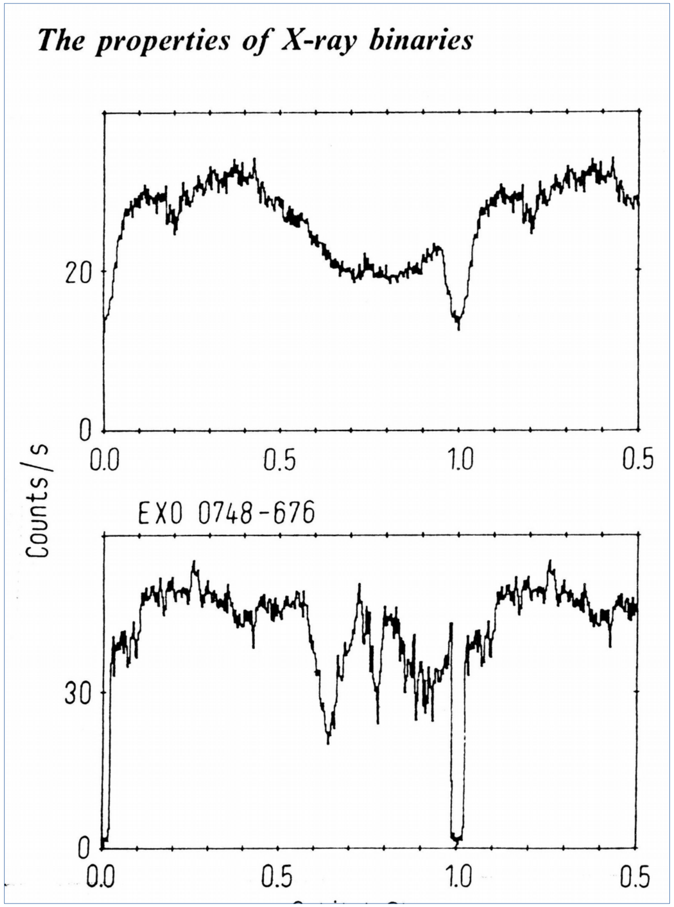
There are new developments for computing the power spectra from uneven data, e.g. orthogonal matching pursuit, which are more general than just derivation of the power spectrum.

5.2 Looking for a period – epoch folding

A simple and independent method is just to assume we know the period, and we fold the curve with the dopted period. If we see a nice pattern, then we have found a solution. Instead on relying on a personal judgements, we rather use some quantitative criterium to decide which value is the best. For example, we can calculate the mean folded curve in some time bins, and then calculate the total dispersion

$$\sigma^2 = \frac{\sum_{i,j} (x_{ij} - \bar{x}_j)^2}{N-1}$$

This approach is known as ***Phase Dispertion Minimization***. We thus choose the period T as the one corresponding to its lowest value. We show the folded curve as the result.

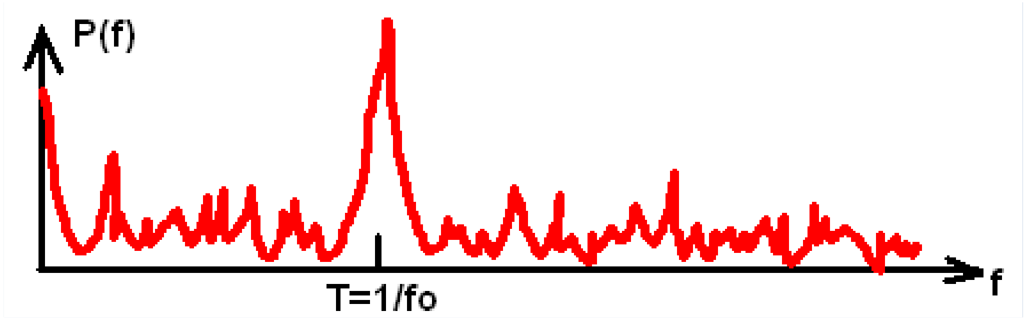


Even if we use a periodogram, it is frequently interesting to show the folded lightcurve using the selected period since such a lightcurve is less noisy and shows better the periodic pattern.

Two examples of the colded lightcurves in 1-10 keV range for the sources X1822-371 and X0748-676 (Parmar et al. 1986), showing 1.5 cycle. Those are low mass X-ray binaries, containing neutron stars.

5.3 Looking for a period – equally spaced data

If points are equally separated, then you should perform discreet Fourier transform, and that one is reversible, and symmetric between the time and frequency domain. Here the frequencies start formally from 0 and continue down to 1. The formula uses Δt = 1 and implicitly assumes that the function repeats with the period T_{max} = NΔt .



$$\hat{F}_k = \frac{1}{N} \sum_{n=0}^{N-1} F_n \exp(-2\pi i k n / N) \qquad F_k = \frac{1}{N} \sum_{n=0}^{N-1} \hat{F}_n \exp(2\pi i k n / N)$$

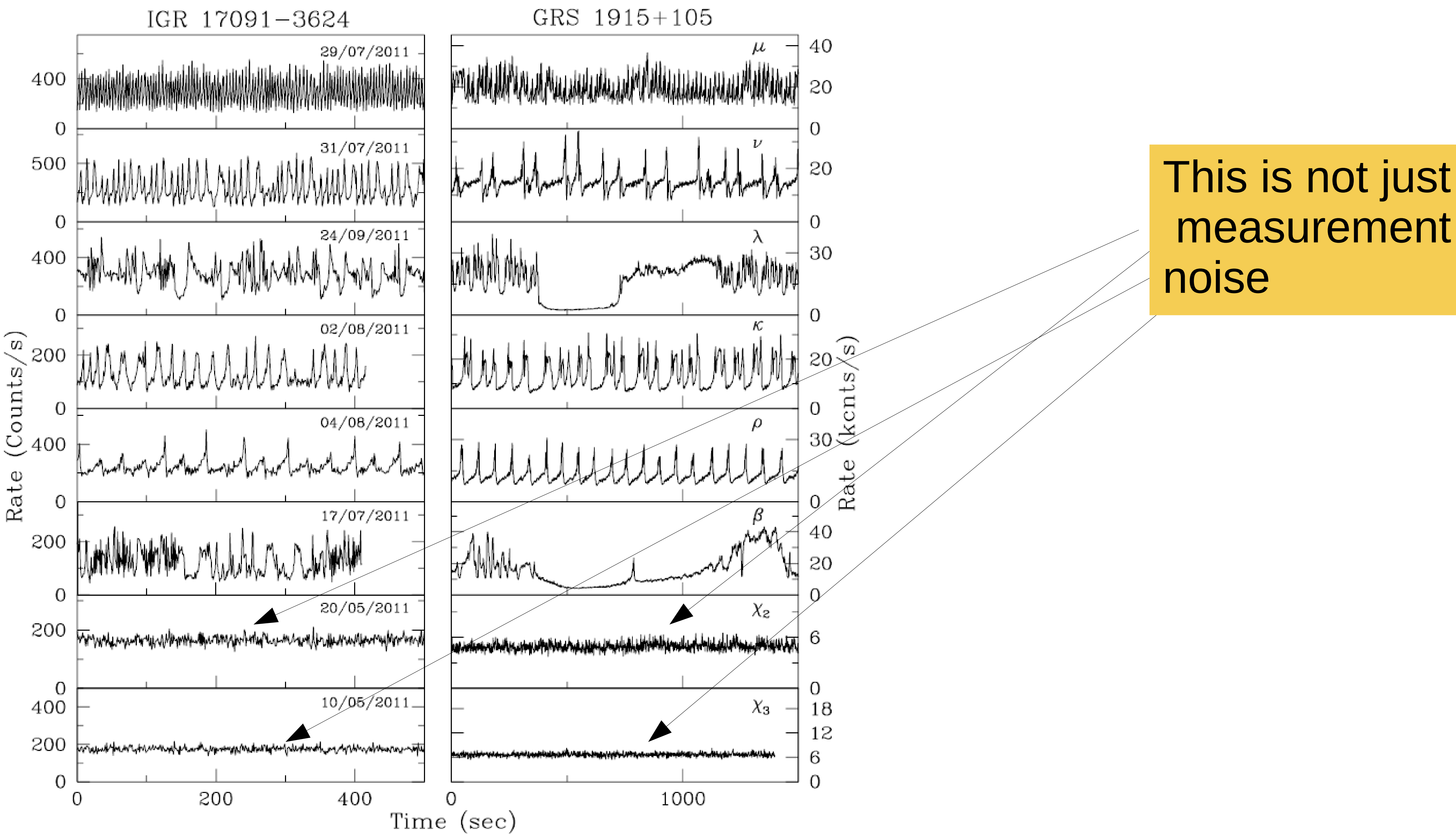
Additionally, if the number of observed points can be expressed as 2^M, when M is a natural number, then computations are performed much faster using the **Fast Fourier Transform (FFT)** algorithm.

If we use the formula above directly, you have to perform N x N operations. FFT allows to do the same by an inteligent trick which required only N log(N) operations. For very large values of N like thousands or millions, this saves enormous amount of time.

If the data is not of 2^M form but equally spaced and long, it is usually better to loose even up to almost 50 % of the data but to apply FFT.

6 Stochasic variability

During lecture 9 we discussed the issues of outbursts due to accretion disk instabilities. Now we will concentrate on the periods when there is no outbursts.



6 Stochasic variability

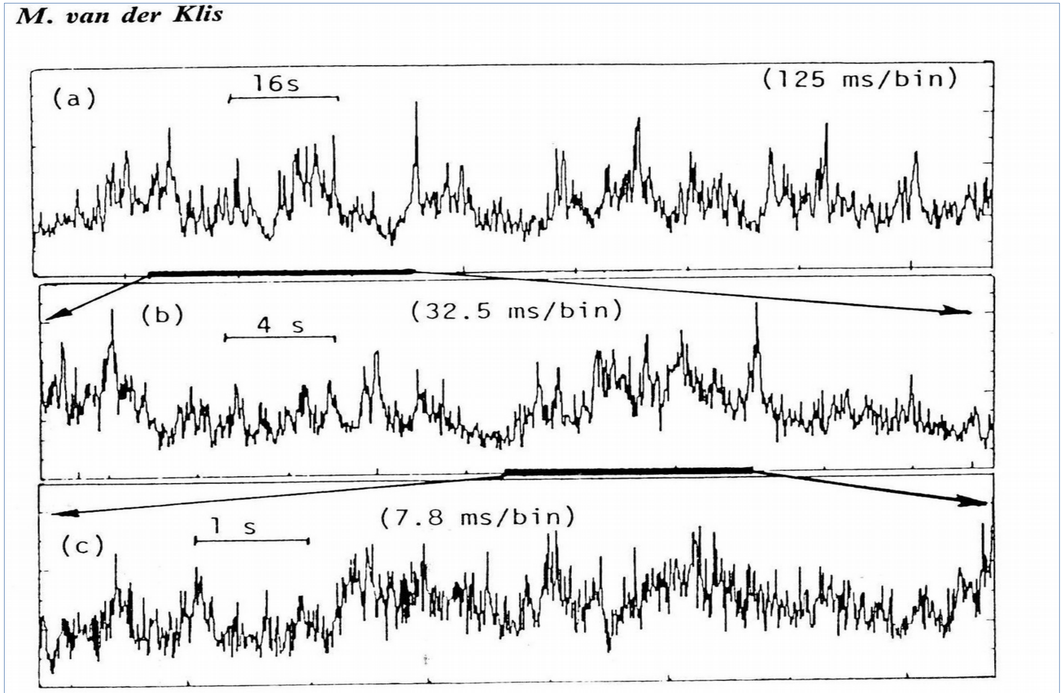


Fig. 6.8. Light curves of Cyg X-1 in the LS in various time resolutions (Makishima 1988).

We do not see any clear periodicity, but the variability is also an important part of the source properties, so we measure the power spectrum density of the source.
But we change the way how we normalize the power spectrum.
Standard approach:

$$P(f)=\frac{1}{N}\left|\sum_{j=1}^{N} F(t_j)\exp(-2\pi i\omega t_j)\right|^2$$

If the function is roughly periodic, then this normalization is ok. But if we have a Brownian motion, like in the case of photon scattering, the departure from the starting value rises as sqrt(N).

Normalization used for stochastic lightcurves

$$P(f)=\frac{T}{N}\left|\sum_{j=1}^{N} F(t_j)\exp(-2\pi i\omega t_j)\right|^2$$

That is we multiply the PSD by the duration of the observation. This give the result which less depends on the choice of the original lightcurve duration.

If we additionally divide p(f) by the mean value $(\bar{F})^2$

we get a quantity that has units of Hz, and in addition it is related to the normalized variance calculated for a fixed length T of the data

6 Stochasic variability

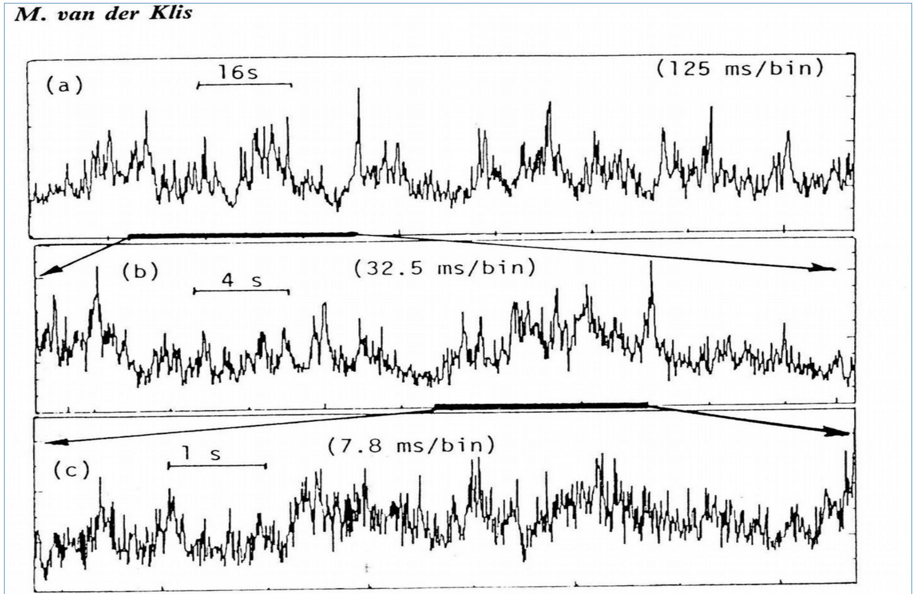


Fig. 6.8. Light curves of Cyg X-1 in the LS in various time resolutions (Makishima 1988).

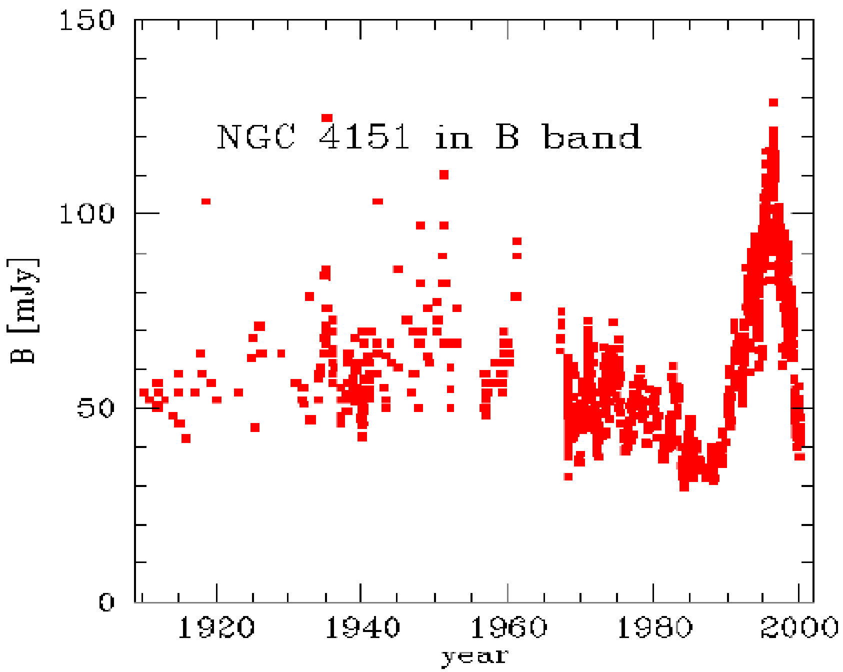
If we additionally divide p(f) by the mean value $\bar{F}^2$

we get a quantity that has units of Hz, and in addition it is related to the normalized variance calculated for a fixed length T of the data

$$P(f)_{norm} = \frac{T}{N\,\bar{F}^2} \left| \sum_{j=1}^{N} F(t_j) \exp(-2\pi i \omega t_j) \right|^2$$

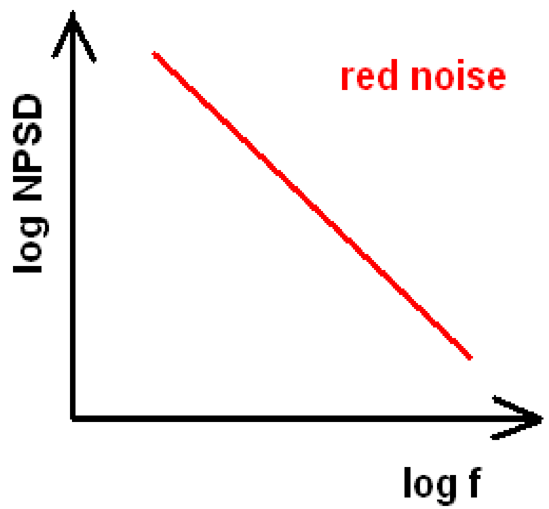
$$\sigma^2_{norm} = \sum_{1/T}^{\infty} \frac{(F_k - \bar{F})^2}{N\,\bar{F}^2} \qquad \sigma^2_{norm}(T) = \int_{1/T}^{\infty} P_{norm}(f)\,df$$

Those expressions are explicitly dependent on the duration of the observational data. This is the important property of the stochastic variability, which shows up in the lightcurves of accreting sources.



100 years of optical data from NGC 4151

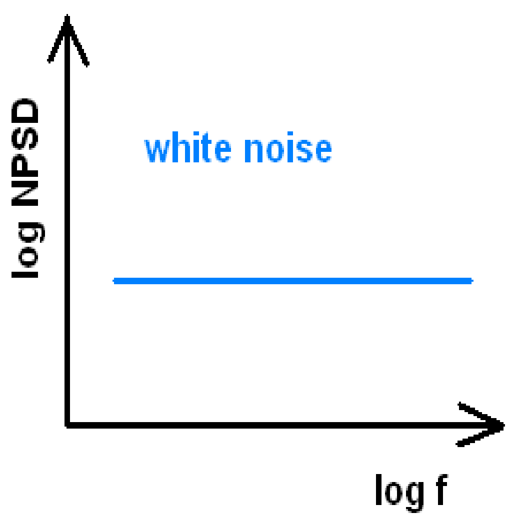
6 Stochasic variability: basic examples of normalized power spectrum densities



If the plot looks like a power law, in astronomy we name that a red noise. Sometimes the term ‘red noise’ is reserved to slope = 1 only.

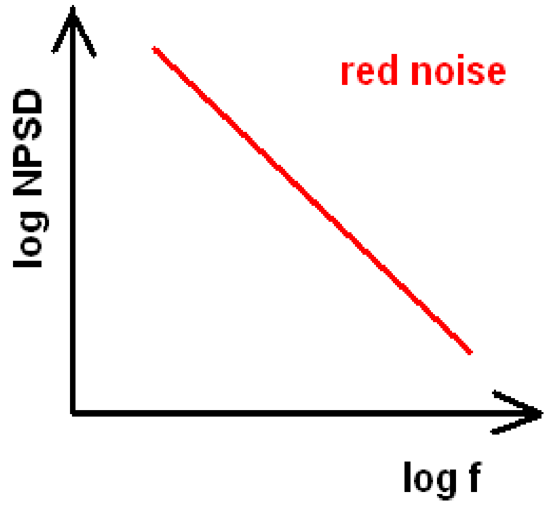
Here we see that the value of the Normalized excess variance rises with the length of the curve.

$$\sigma^2_{norm}(T) = \int_{1/T}^{\infty} P_{norm}(f)\,df$$



This is what happens if we have just measurement errors, and no real signal in the data.

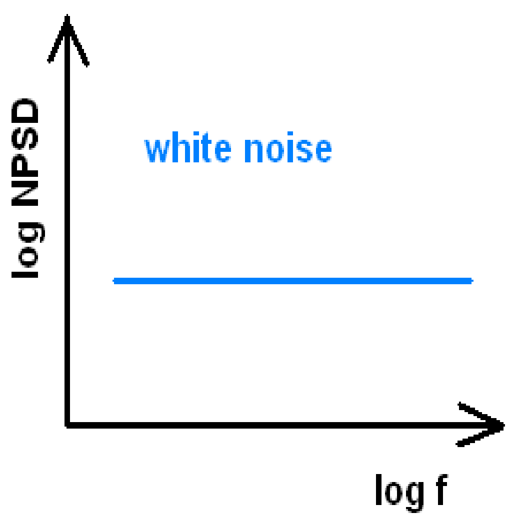
6 Stochasic variability: basic examples of normalized power spectrum densities



If the plot looks like a power law, in astronomy we name that a red noise. Sometimes the term ‘red noise’ is reserved to slope = 1 only.

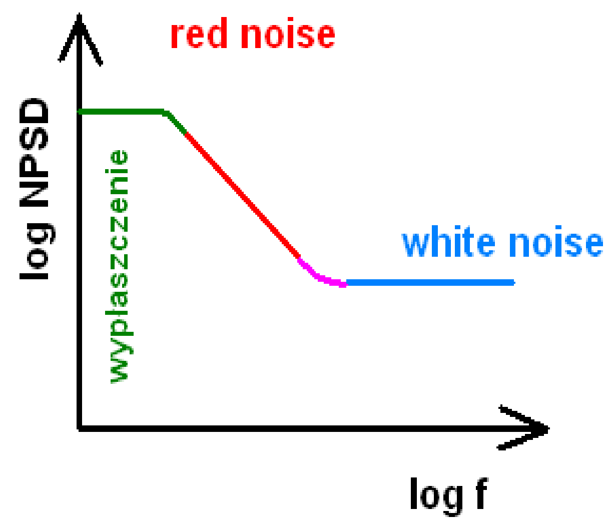
Here we see that the value of the Normalized excess variance rises with the length of the curve.

$$\sigma^2_{norm}(T) = \int_{1/T}^{\infty} P_{norm}(f)\, df$$

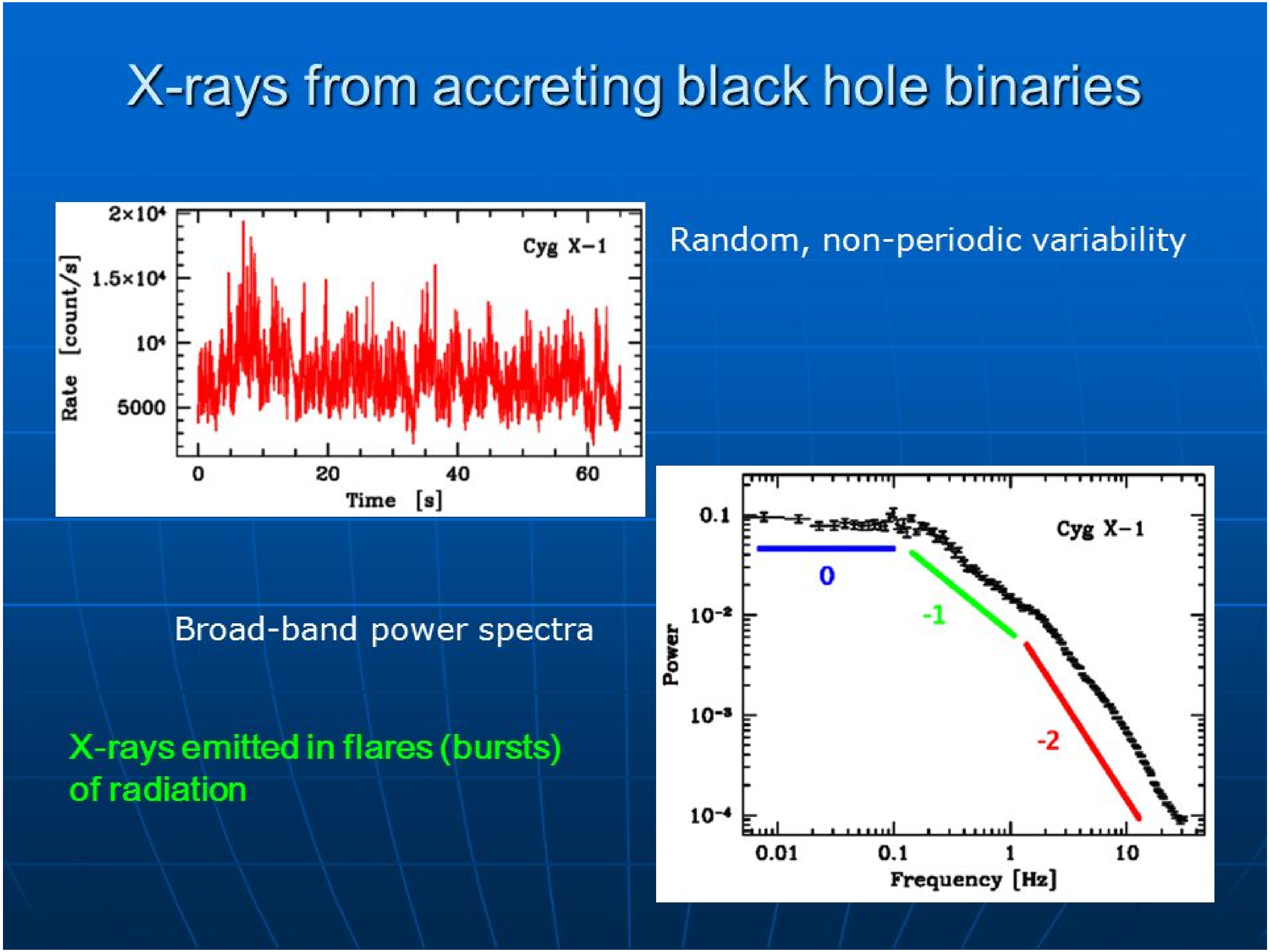


Thus in the real situation, if we do not subtract the measurement error from the measured P(f) we see something like that at the high frequency tail.

At the low frequency tail also usually some flattening appears when we reach the longest timescales physically present in the system.

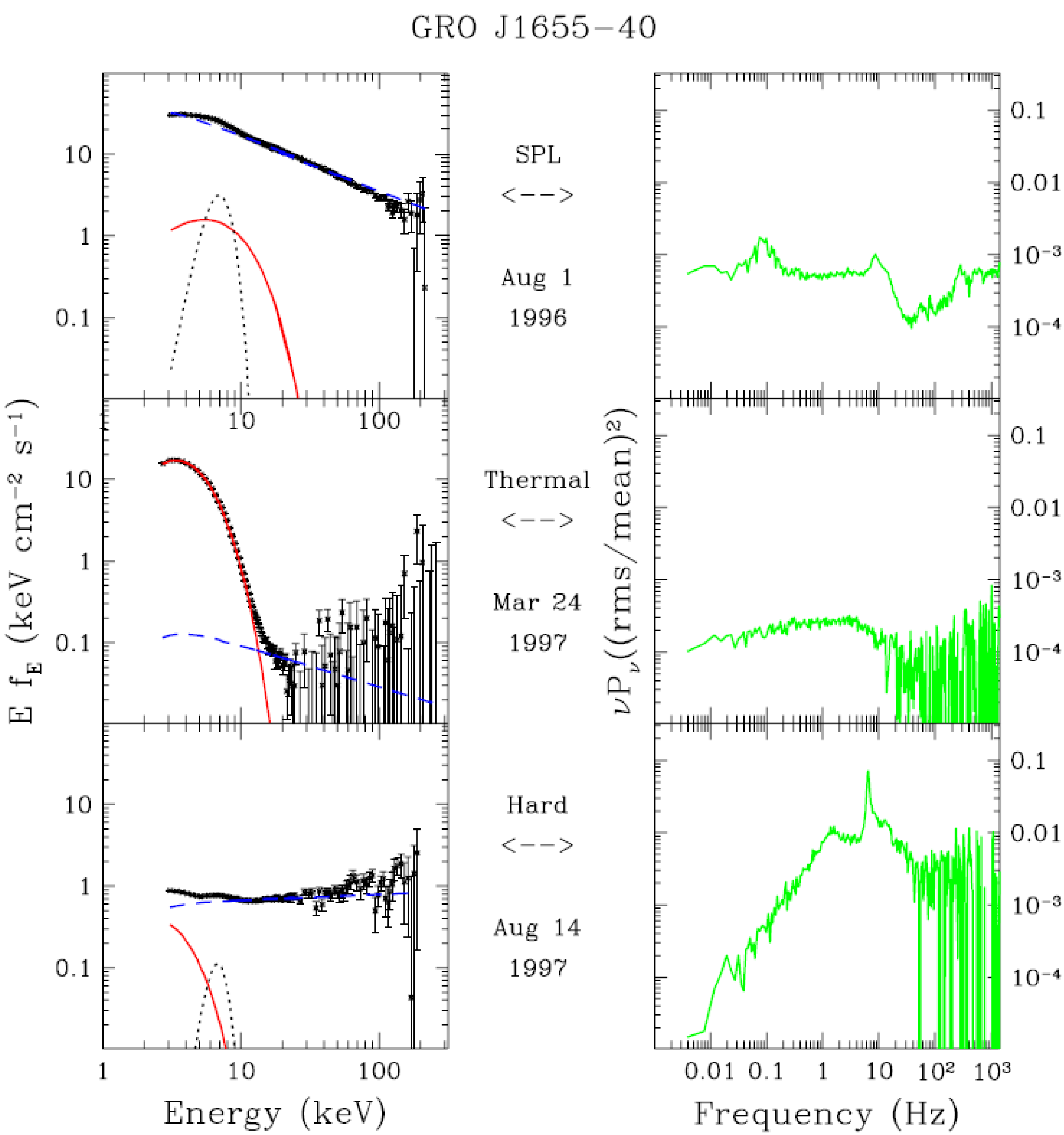


6 Stochasic variability: real examples



From Życki & Sobolewska

6 Stochasic variability: real examples



In galactic sources we see that the variability properties changes as the state of the source changes.

Important note: frequently a power spectrum is also plotted in such way as to show the maximum power, so we plot

log (f P(f)) vs log(f)

This changes the slopes by 1, i.e. flat slope corresponds to -1

Important note: power spectra show sometimes narrow peaks – quasi-periodic oscillations.

From Remillard & McClintock (2006)

6 Stochasic variability: real examples

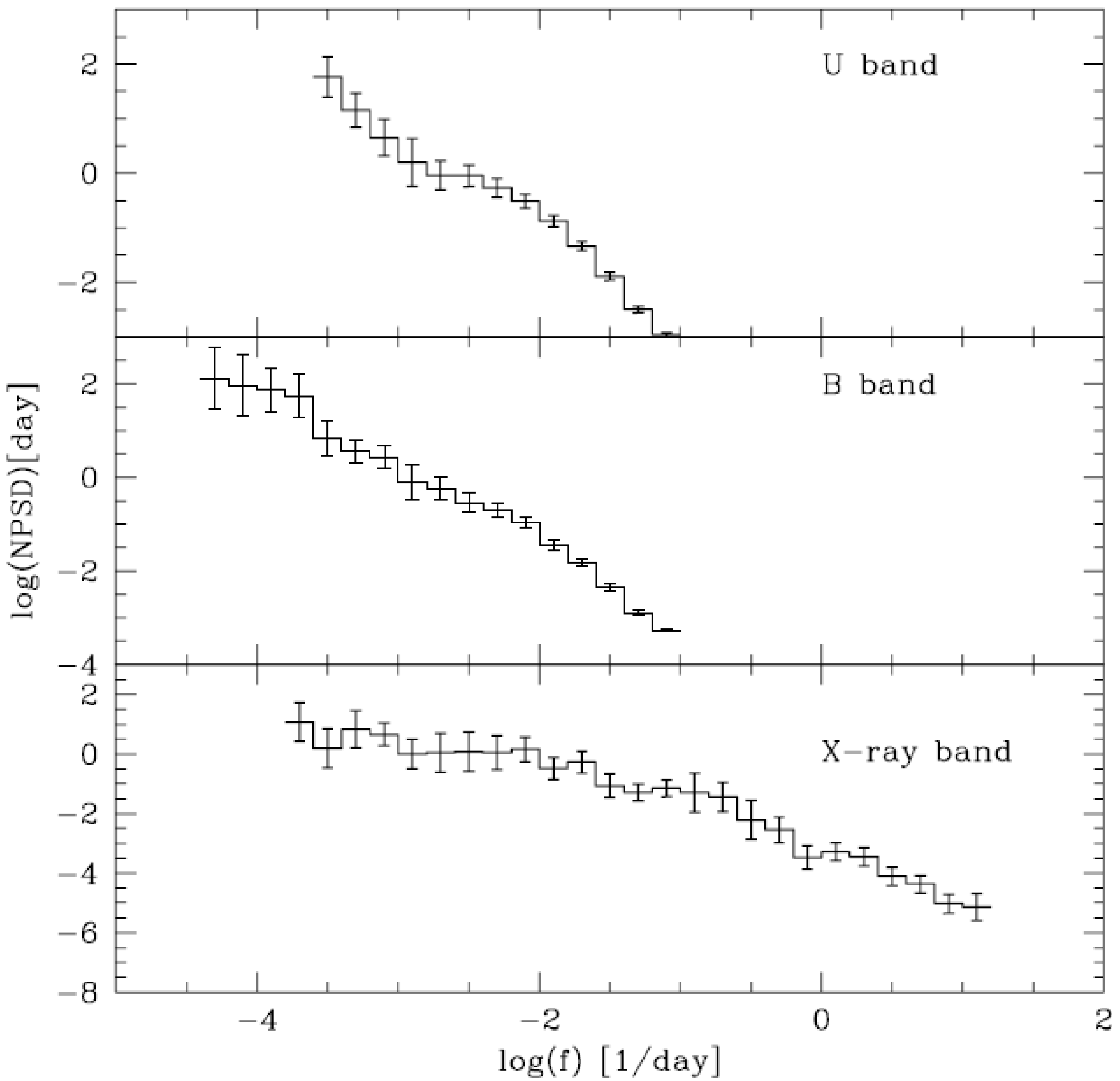


Figure 3. The normalized power spectrum density of NGC 4151. Upper panel, *U* band (1968–2000); middle panel, *B* band (1910–2000); lower panel, X-ray band (1974–2000). Marked errors are direct observational errors, as described in Appendix A.

Basically, AGN and galactic sources she the same variability properties but timescales are clearly longer.

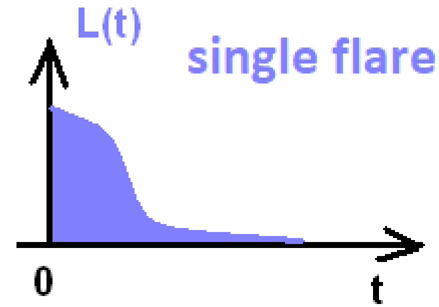
Optical and X-ray variability of NGC 4151 (Czerny et al. (2003)

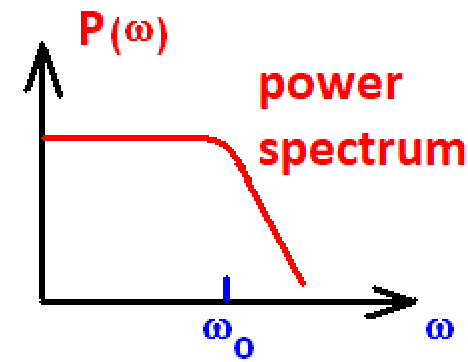
7 Modelling the broad band power spectrum

Two simple example which can be calculated analytically and show how the complex braod band power spectrum can actually form.

Case 1

$$F(t) = A \exp(-\omega_0 t) \qquad t > 0$$





$$\hat{F}(\omega) = A \int_0^{\infty} \exp(-\omega_0 t) \exp(i \omega t) dt = \frac{A}{i\omega - \omega_0}$$

$$P(\omega) = \frac{A^2}{\omega_0^2 + \omega^2}$$

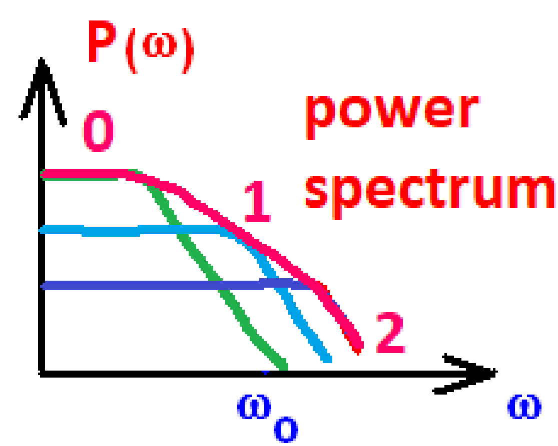
A single flare starting at some moment and then decaying exponetially gives the power spectrum of the slope -2 at high frequencies.

If we have more flares, but starting randomly, this will not change the shape of the power spectrum, just the normalization

$$P(\omega) = \lambda \frac{A^2}{\omega_0^2 + \omega^2}$$

where λ is the number of flares per unit time.

If we have a range of decay timescales in the ransom shorts, we can produce the slope -1 with proper normalization of the contribution from various flares.

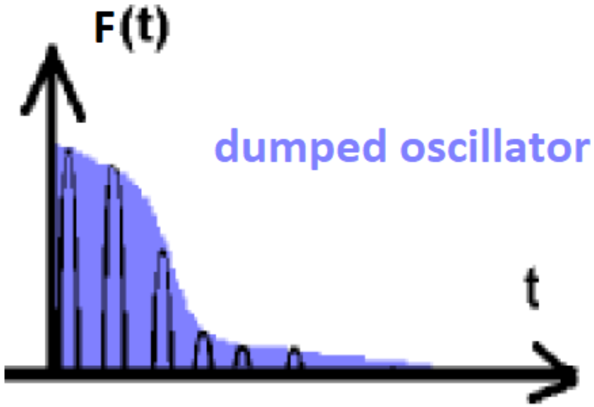


7 Modelling the broad band power spectrum

Case 2

We now assume the dumped oscillator

$$F(t) = A\sin(\Omega t + \phi)\exp(-\omega_0 t) \qquad t > 0$$

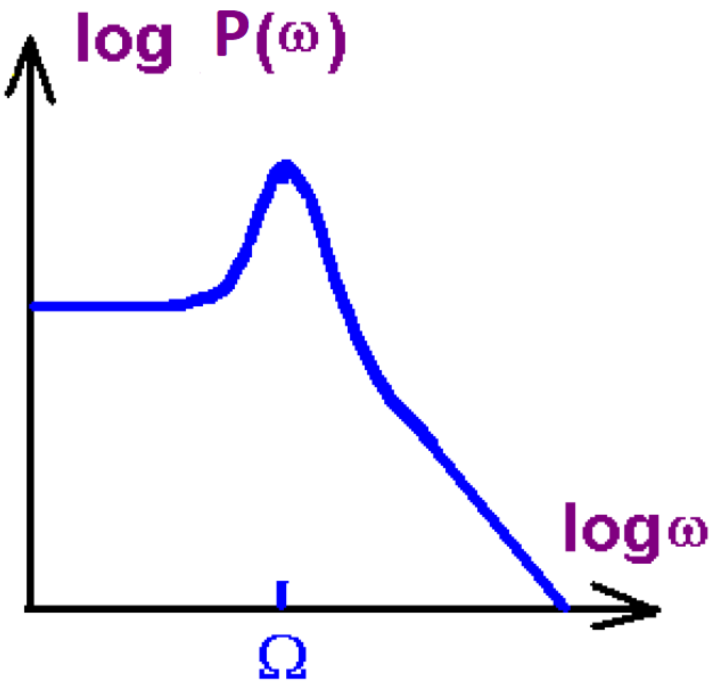


The power spectrum from such shape is a Lorentzian

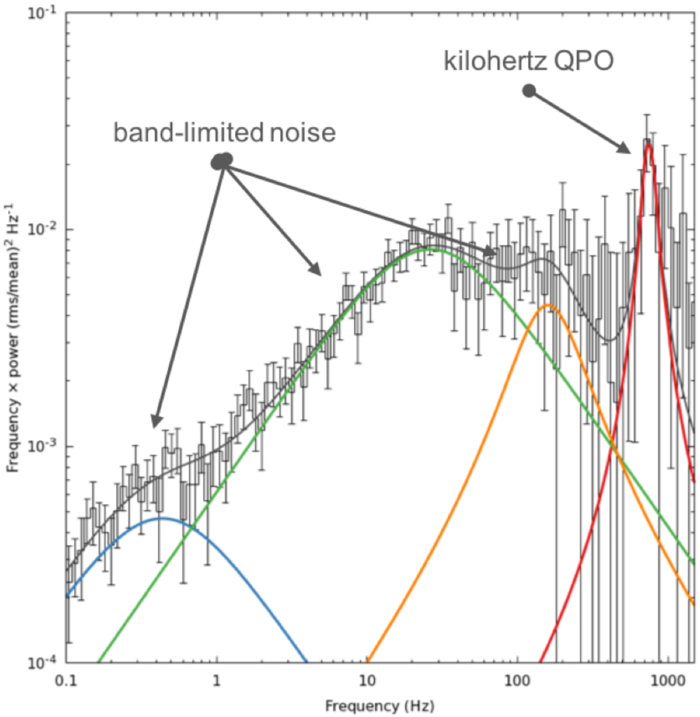
$$\frac{2N^2 Q\Omega}{\Omega^2 + Q^2(\omega - \Omega)^2}$$

Where $Q = \frac{\Omega}{2\omega_o}$ and $N = \frac{A}{Q^{1/2}}$

The parameter Q describes the prominence of the peak, and 1/Q actually gives the relative peak width. It is frequently used to characterize the Quasi-Periodic Oscillation component in the power spectrum.



7 Quasi-periodic oscillations and their relation with the theory of accretion disks



NICER observation of a kilohertz QPO in 4U 0614+09, Bult et al. (2018)

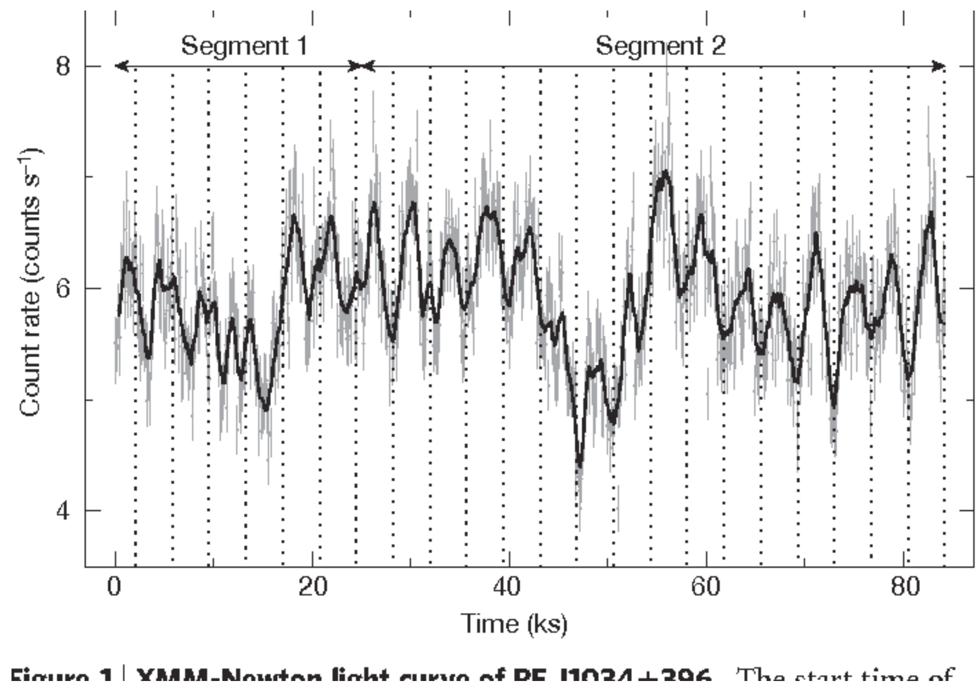
The connection between the observed QPO frequencies are not as clear as we would like.

The kilohertz QPO are usually considered to be related to ISCO. The observational relation allows the eventually to use for black hole mass measurement

$$\nu_0 = 931 \left(\frac{M}{M_\odot} \right)^{-1} \text{[Hz]},$$

(from Remillard & McClintock (2006).

However, they come frequently in pairs (suggests resonance) but what is worse, they are actually observed in the high energy part of the X-ray emission, thus they are not likely related to the cold Keplerian accretion disk but instead they come from the hot plasma.



Strong QPO detected in AGN RE J1034+396 (Gierliński et al. 2008)

8. Other methods to analyse the lightcurve

Calculating the power spectrum is not the only thing we can do. There are many other techniques:

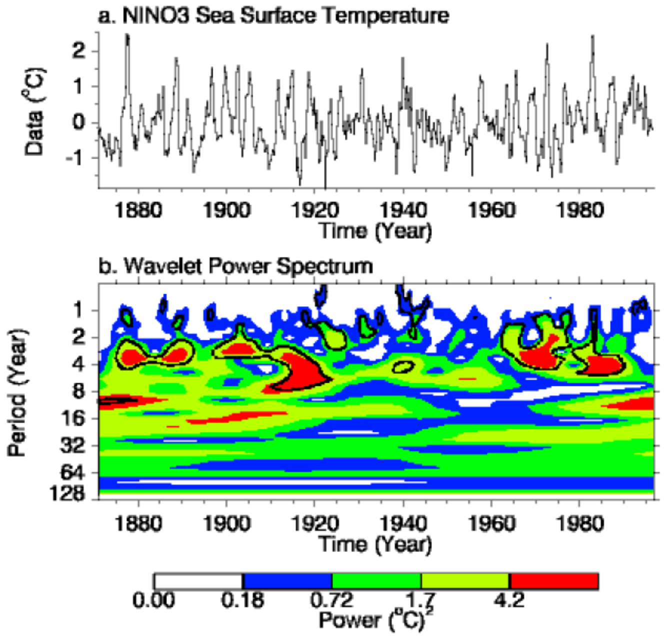
- Running variance (for analysis of stronlgly non-stationary behaviour)
- Time-dependent power spectrum
- Wavelet analysis (should be also good for QPO)
- Autocorrelation function (measures the characteristic timescale of correlations)

$$C(\Delta T) = \frac{1}{D^2}\frac{1}{T}\int_{-T}^{+T} L(t)\,L(t-\Delta T)\,dt$$

- Structure function (alternative to power spectrum)

$$SF(\Delta T) = \frac{1}{D^2}\frac{1}{T}\int_{-T}^{+T} [L(t)-L(t-\Delta T)]^2\,dt$$

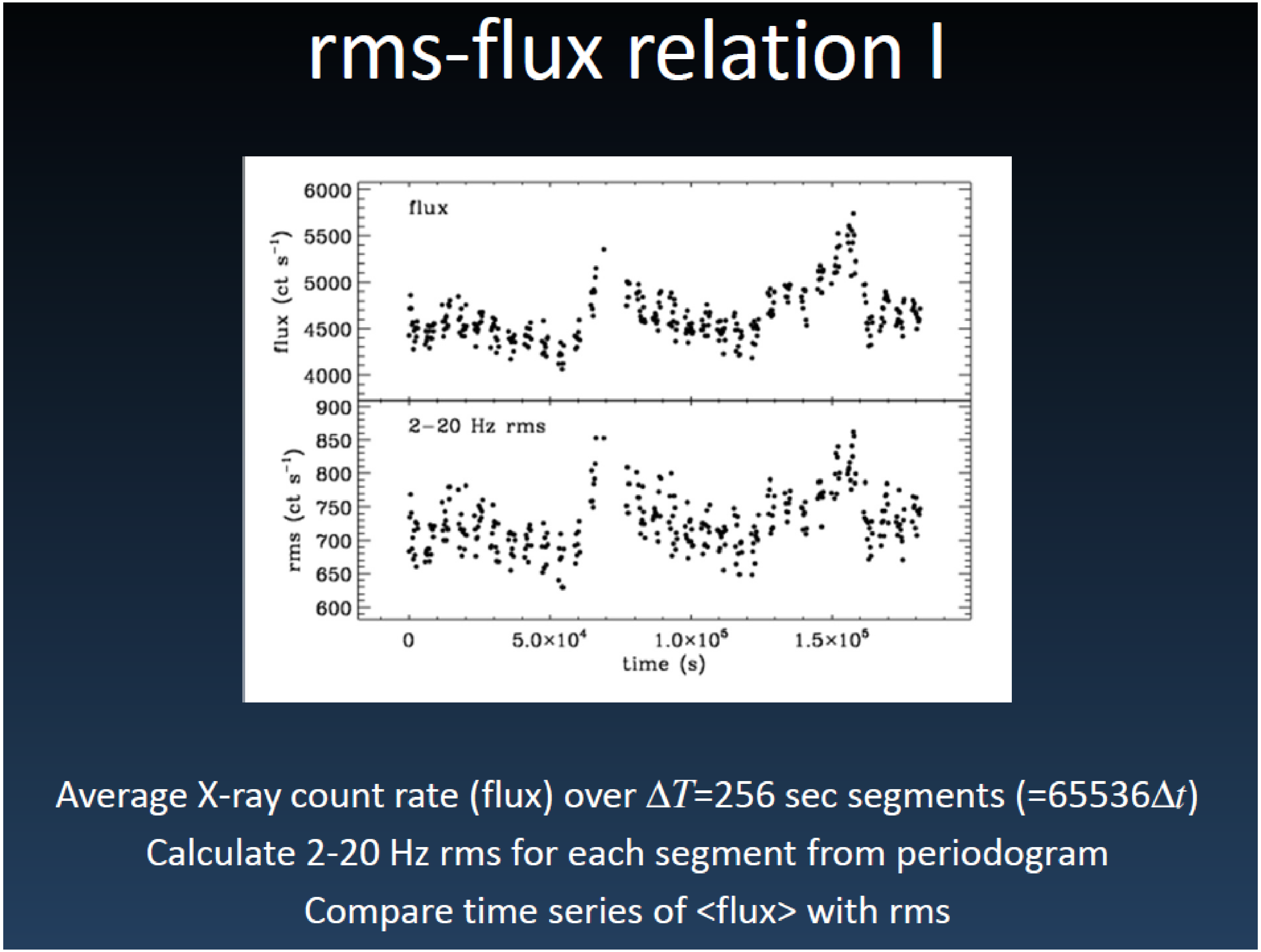
- Fractal analysis (searches for deterministic chaos)
- Non-linear prediction method (allows for example to see a difference between the stochastic signal and the deterministic chaos)
- Matching pursuit methods (more general than wavelet analysis)
- Rms – flux relation
- …..



Here D is the dispersion

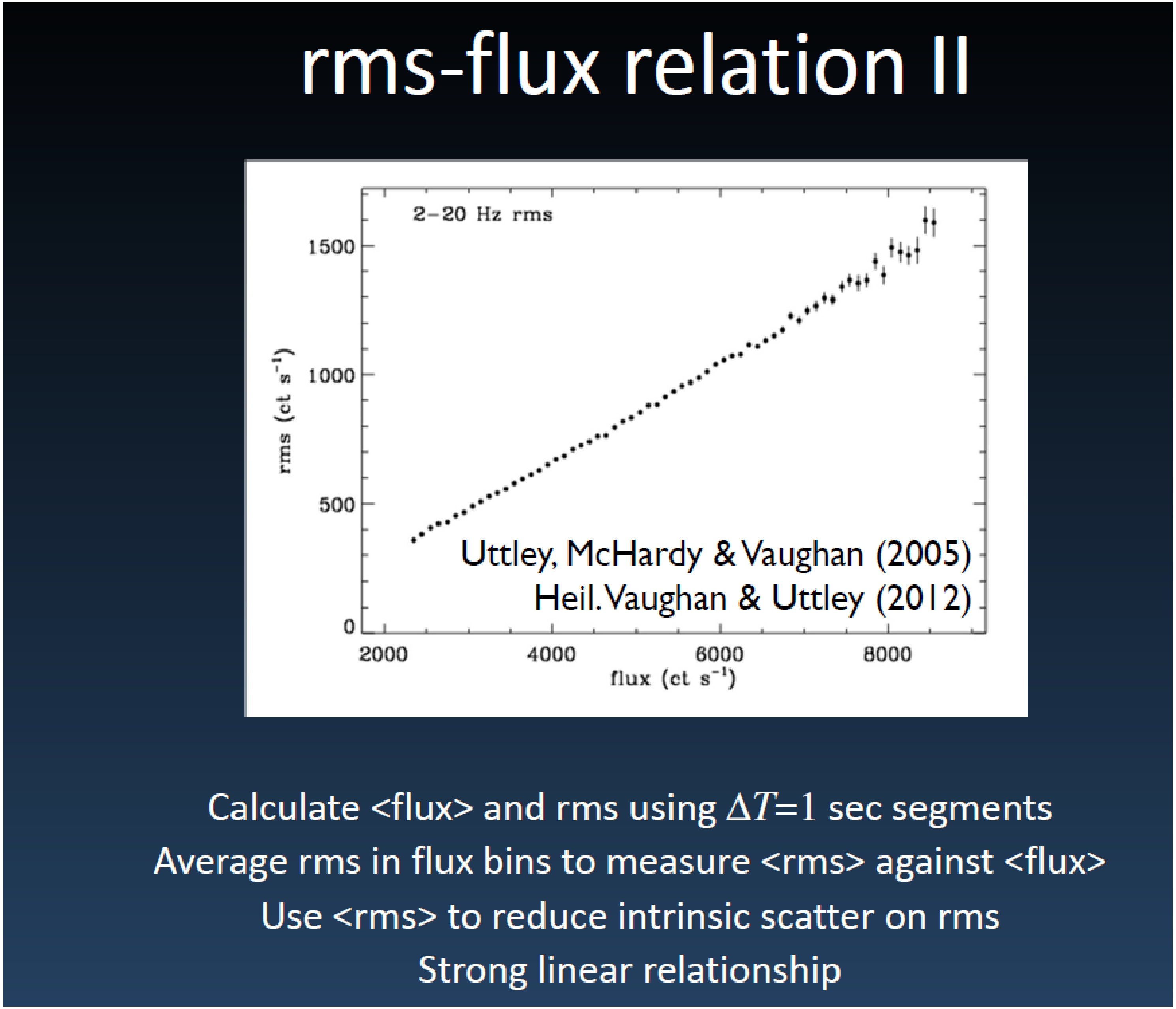
This is an interesting aspect of the accreting sources variability, most studies in X-rays for binary stars

Digression: rms-flux relation



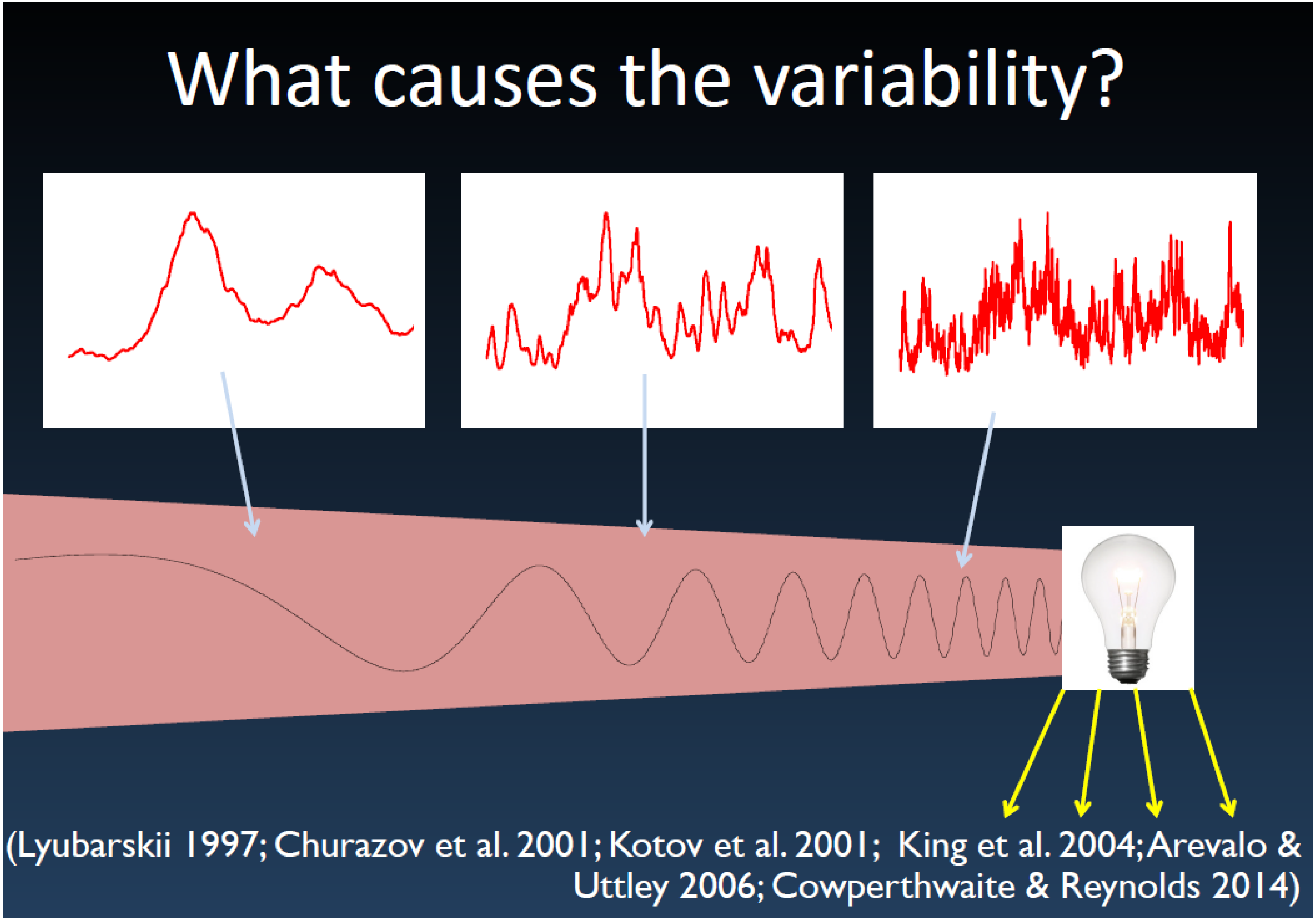
From presentation by Uttley

Digression: rms-flux relation



From presentation by Uttley

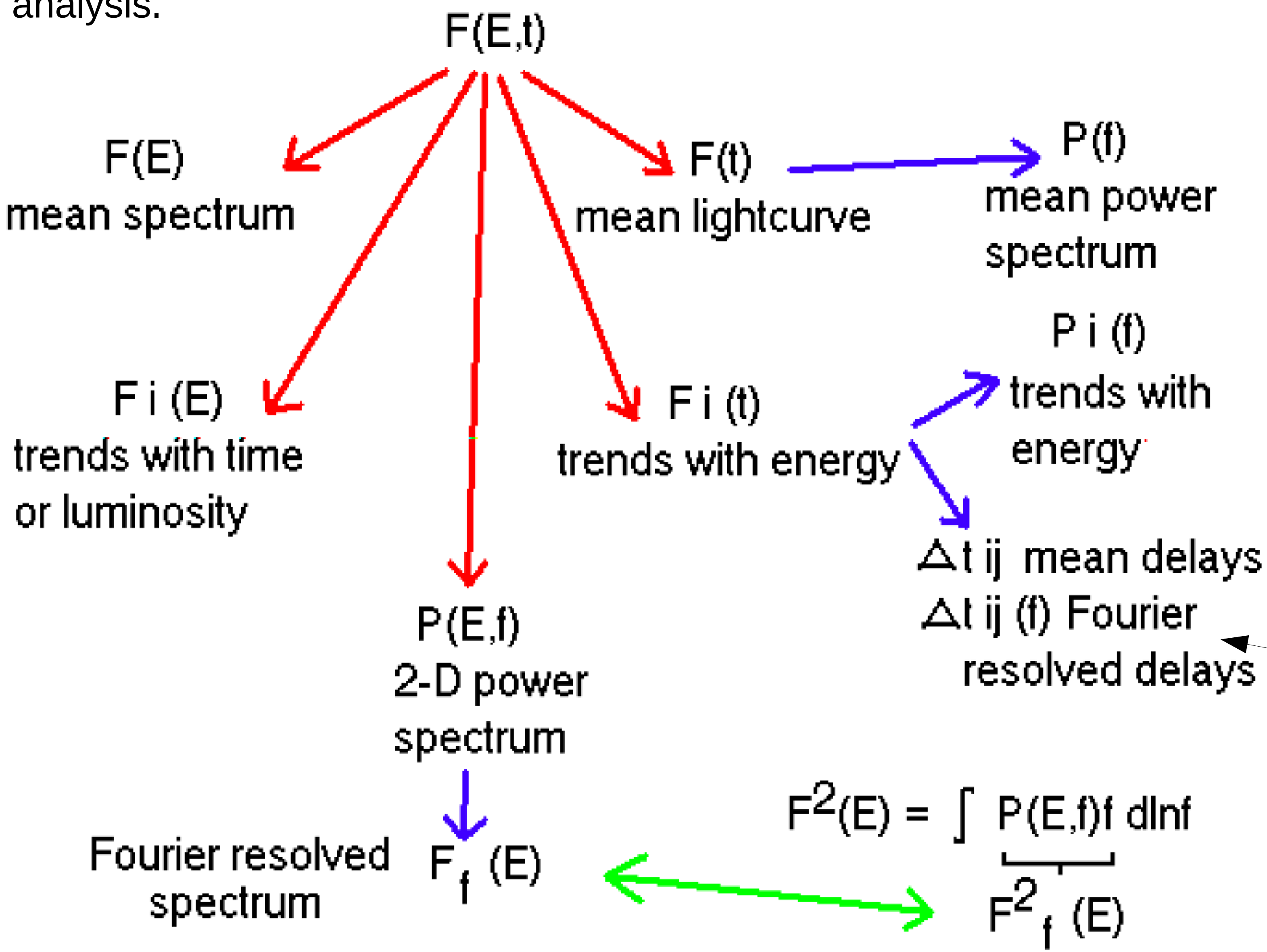
Digression: rms-flux relation



From presentation by Uttley

9. Time-energy spectral analysis as an example of two-dimensional analysis

In general, id we measure the photon arrival time and energy, we can perform more advanced 2-D analysis.



Ask Barbara de Marco

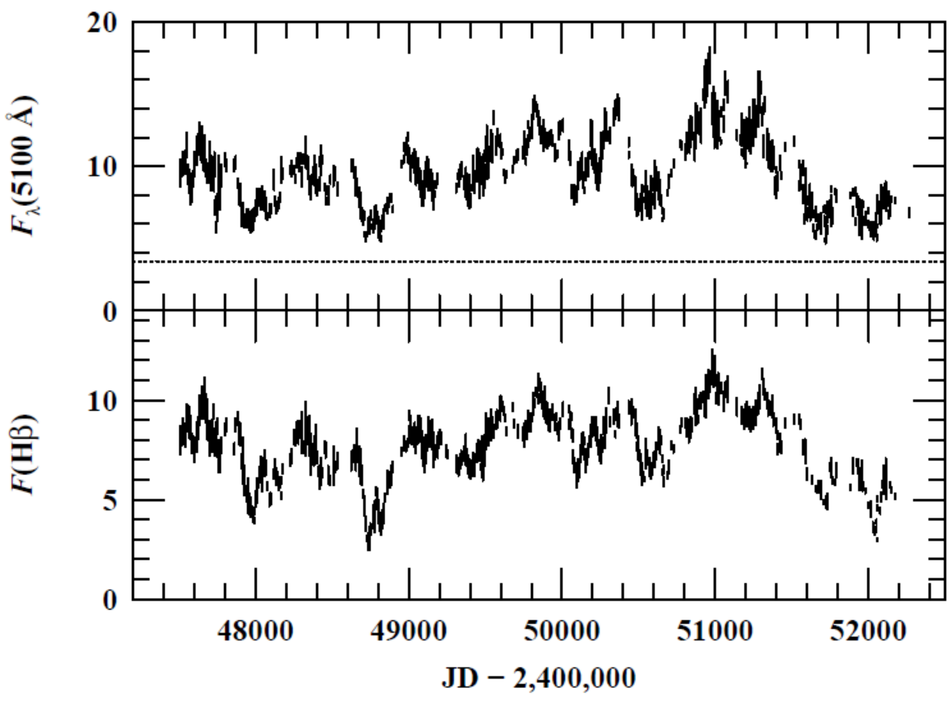
10. Mean time delays

If we have two lightcuves in two energy band, we can be interested in determining the time delays between them.

Frequently used technique: Interpolated Cross Correlation Function

$$CC(\Delta T) \propto \frac{1}{T}\int_{-T}^{+T} F(t)\,G(t-\Delta T)\,dt$$

Example 1: time delay measurement of Hbeta with respect to the continuum



13 years of monitoring of NGC 5548, Peterson et al. (2002)

Results:

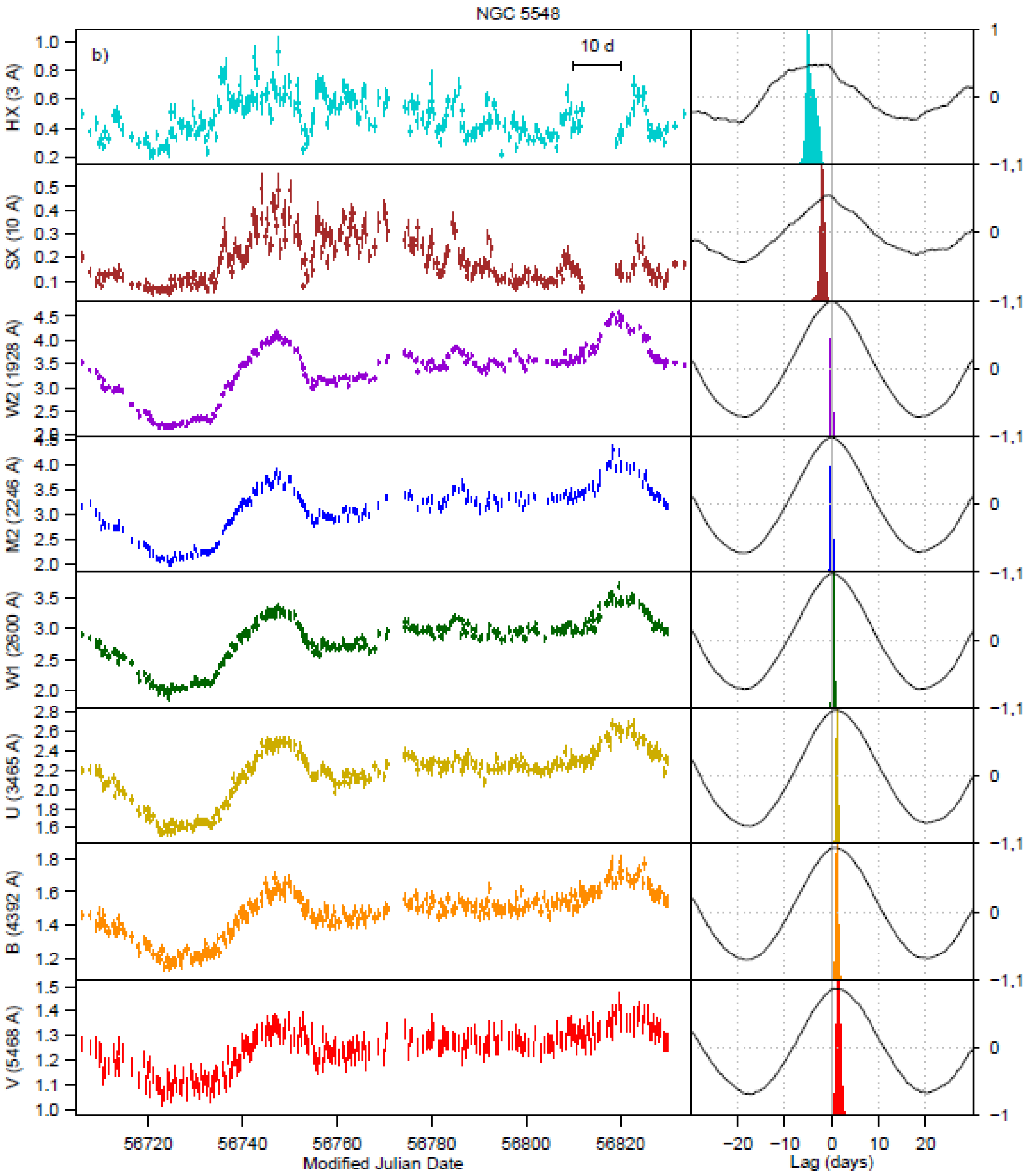
- black hole mass measurement from the Keplerian motion,

$$v^2 = \frac{GM}{R}$$

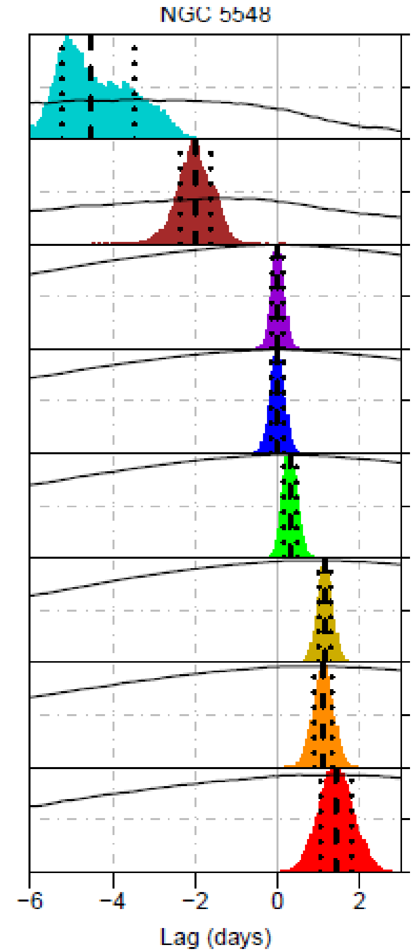
$M = 6 \pm 2 \times 10^7\ M_\odot$ from Wandel et al. 1999

- proof that the size of the broad line region reacts to the changes of the continuum since in the individual years the time delay changed, with the minimum of 6 days and the maxium of 26 days (Peterson et al. 2002)

10. Mean time delays



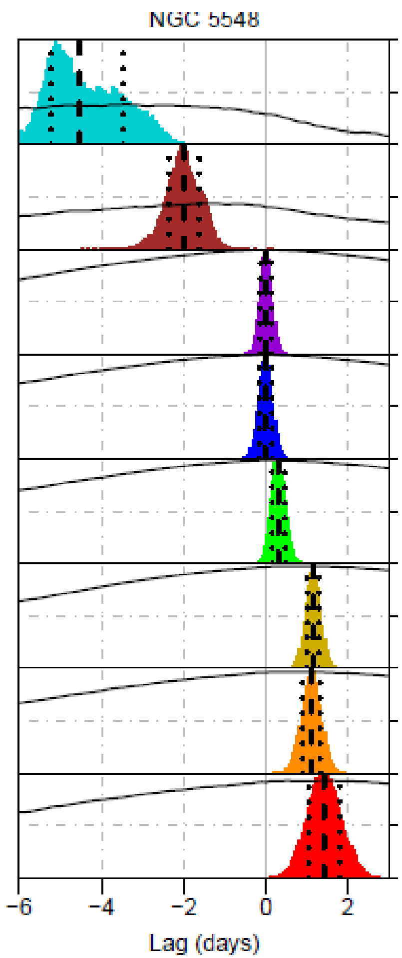
Example 2: time delay from multi-band dense monitoring of NGC 5548 focused on understanding of the continuum emission (Edelson et al. 2019).



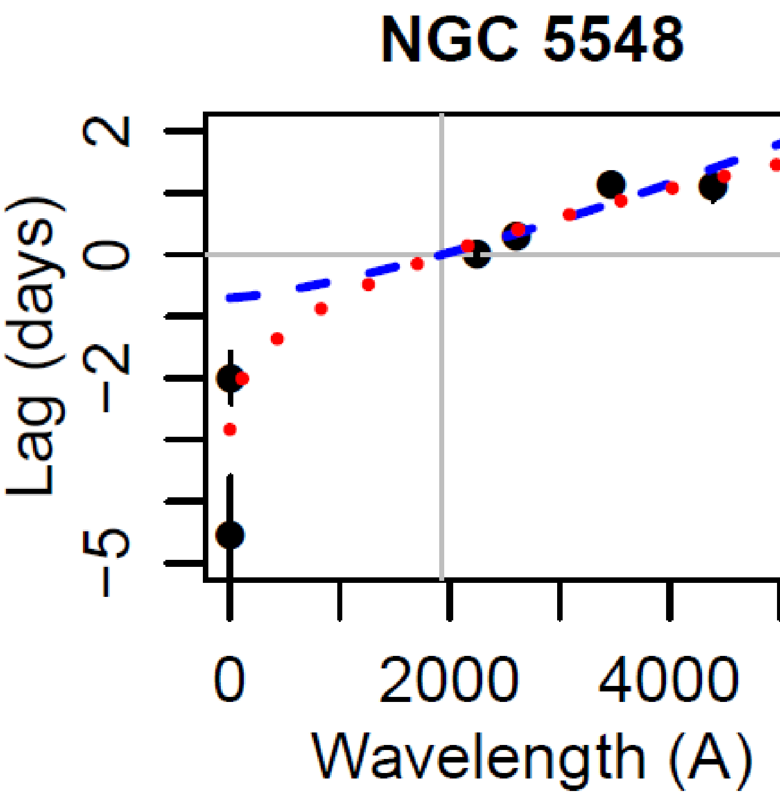
Expanded version of time delay plots

10. Mean time delays

Example 2: time delay from multi-band dense monitoring of NGC 5548 focused on understanding of the continuum emission (Edelson et al. 2019).



Expanded version of time delay plots



Here those time delays are plotted against the expectation from a standard accretion disk irradiated by X-rays (blue dashed line)

$$\tau \propto \lambda^{4/3}$$ (Cackett et al. 2007)

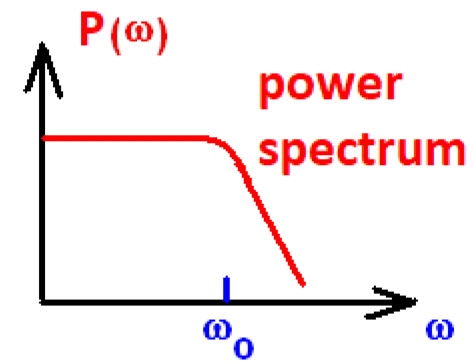
The first two points (hard X-rays and soft X-rays) do not follow the extected pattern.

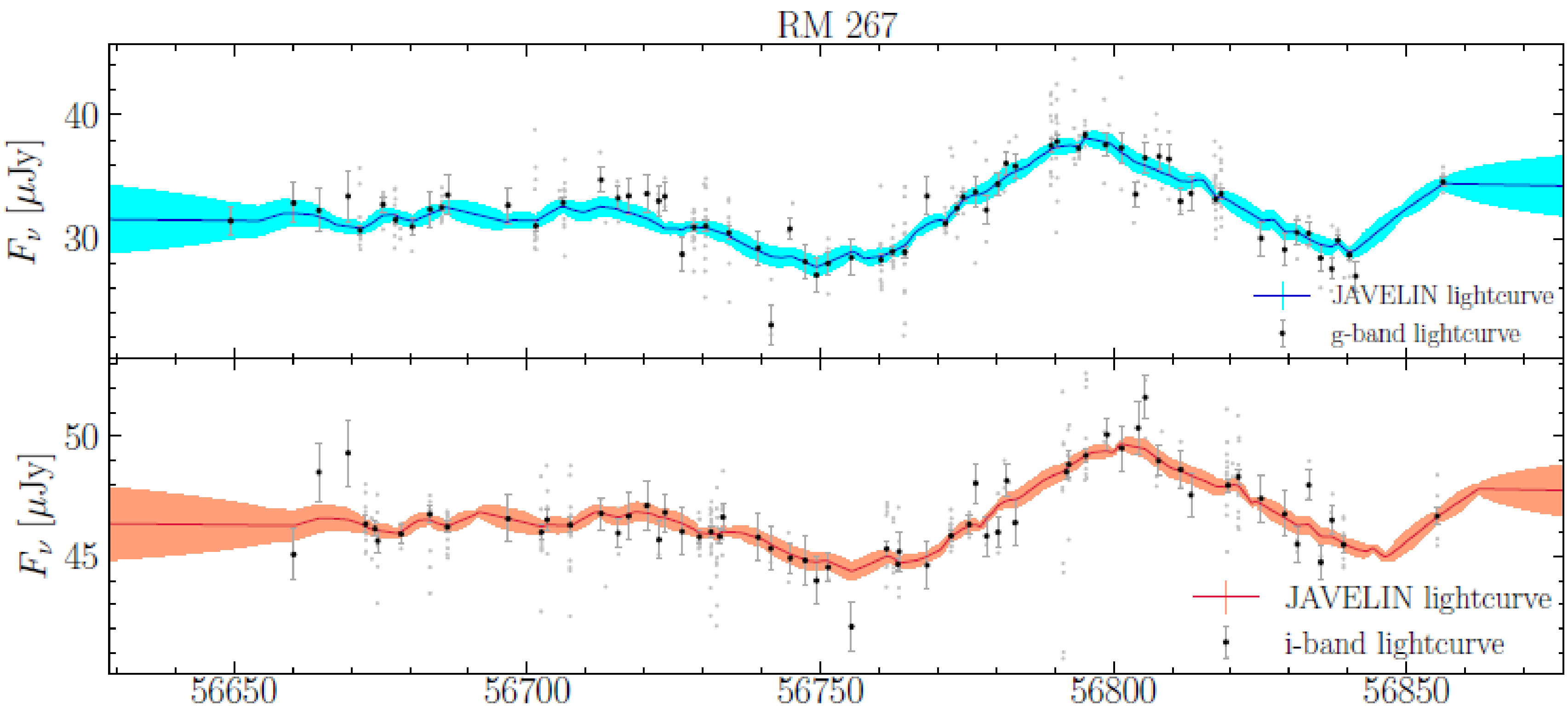
Such data are excellent starting point to study the complex issue of geometry of the central region.

10. Mean time delays – other delay measurement techniques

The time delay can be measured not only using ICCF (Intepolated Cross Correlation Function). For example, in our recent paper about a quasar CTS C30.10 (z = 0.90052) we used

- Discreet Correlation Function (DCF)
- Z-transformed Discreet Correlation Function (ZDCF)
- Javelin (based on data interpolation using Dumped Random Walk)
- von Neumann estimator
- Shifting and X^2 fitting





Example of the lightcurve with Javelin representation from Homayouni et al. (2019)

11. Variability as a tool to study accretion objects

- It is a powerfull tool
- But usually we study only one aspect of the source, with one technique
- Global approach to monitoring and to modelling is necessary to make the best of it and finally to disentangle the issue of hot phase geometry and cold disk/hot phase transition

HOMEWORK

- Why in lecture 9 we used the expression $t_{dyn} = \frac{1}{\Omega_K}$ instead of $T = \frac{2\pi}{\Omega_K}$?
- Calculate the power spectrum from the shape

$$F(t) = A \sin(\Omega t + \phi) \exp(-\omega_0 t) \qquad t > 0$$