Physics 403

Spectral Analysis

Segev BenZvi

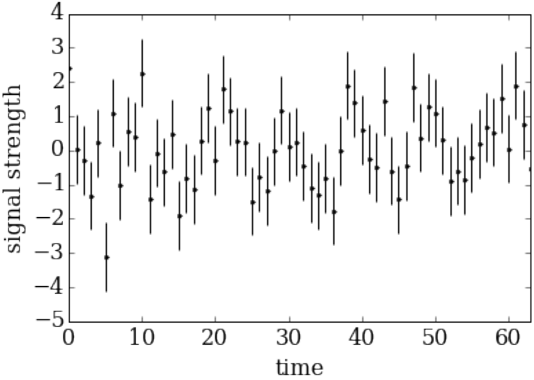
Department of Physics and Astronomy
University of Rochester

Table of Contents

Searching for Structure in a Time Series

Suppose we measure some quantity as a function of time, like the flux of particles in a detector, and we want to estimate f assuming it obeys

$$y(t) = A\cos(2\pi f t + \varphi) + \text{Gaussian noise (mean}= 0, \sigma = 1)$$



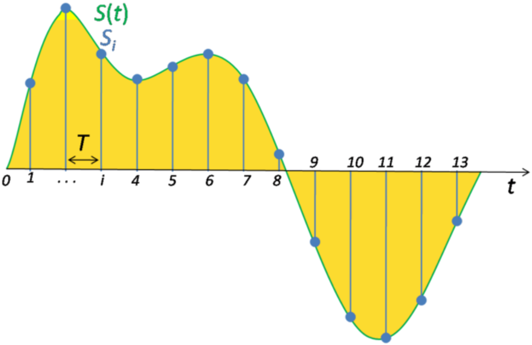
This falls under the domain of spectral analysis

Analysis in the Frequency Domain

- There are many ways one can try to analyze the data in the time domain using maximum likelihood and Bayesian techniques
- Today, we'll talk about solutions in the frequency domain
- This means we need to review the basics of Fourier analysis, because the study of signals in the frequency domain is done using the Fourier transform of the data
- But first, we have to also review some basic concepts from the processing of digital (i.e., sampled) signals:
  1. Analog to digital conversion
  2. Nyquist sampling theorem

Signal Sampling (Digitization)

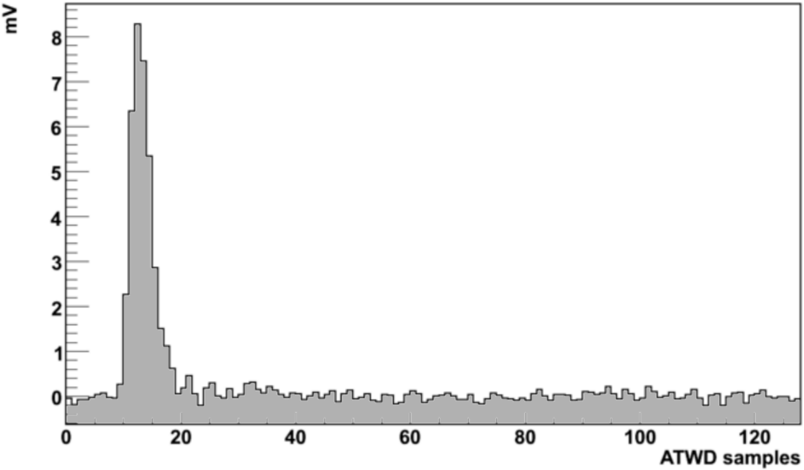
Sampling is the compression of a continuous (analog) signal $S(t)$ into a discrete (digital) signal S_i.



If the signal is sampled at intervals of width T, we say the sampling rate is $f_s = 1/T$.

Example: Analog to Digital Conversion

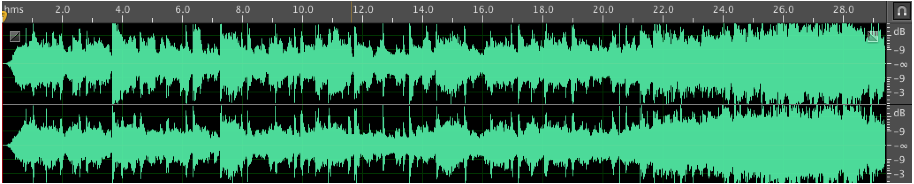
Example of a digitized waveform produced when a single photon triggers a digital optical module, or DOM, in the IceCube detector [1]:



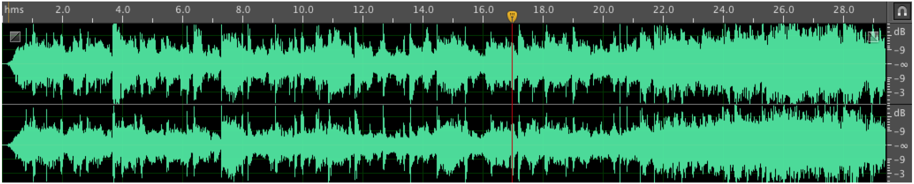
Effect of Sampling Rate

The cochlea in yours ears are frequency analyzers.

30 seconds of song sampled at 44 kHz:

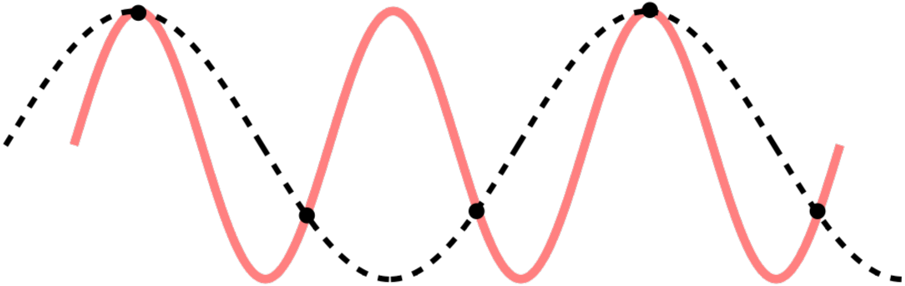


30 seconds of song sampled at 6 kHz:



Nyquist-Shannon Sampling Theorem

If a signal's highest frequency $f < f_s/2$, where f_s is the sampling rate, the signal can be reconstructed perfectly [2, 3].



If $f \geq f_s/2$ the signal can exhibit aliasing, in which several different functions can be reconstructed from the same set of samples. The peak frequency that can be reconstructed is called the Nyquist frequency:

$$f_{\text{Nyquist}} = f_s/2 = 1/2T$$

Table of Contents

Harmonic Analysis

- Fourier's theorem: a periodic function $h(t)$ can be expanded in terms of an infinite sum of sines and cosines, which form an orthogonal basis on $t \in [-\tau/2, \tau/2]$. Defining $f_0 = 1/\tau$, we have

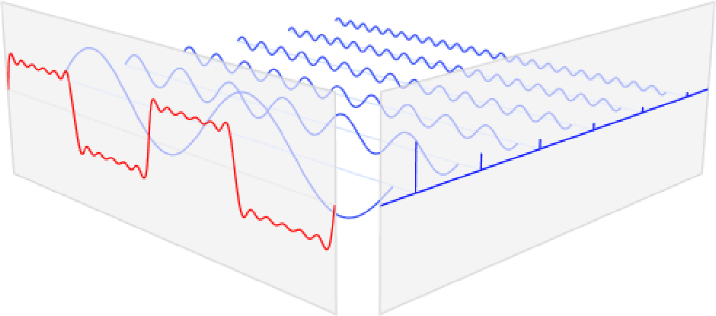
$$h(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(2\pi n f_0 t\right) + \sum_{n=1}^{\infty} b_n \sin\left(2\pi n f_0 t\right)$$

- The terms a_n and b_n are called the Fourier coefficients of $h(t)$ and can be picked out by calculating the inner product of $h(t)$ with the basis functions:

$$\begin{aligned}
a_0 &= \frac{2}{\tau}\int_{-\tau/2}^{\tau/2} h(t)\, dt \\
a_n &= \frac{2}{\tau}\int_{-\tau/2}^{\tau/2} h(t)\, \cos\left(2\pi n f_0 t\right) dt \\
b_n &= \frac{2}{\tau}\int_{-\tau/2}^{\tau/2} h(t)\, \sin\left(2\pi n f_0 t\right) dt
\end{aligned}$$

Fourier's Theorem

Red: a periodic function (square wave) approximated by the first six terms in the Fourier series



The series of lines on the right indicate the power spectral density (PSD) of the function. We will spend the next few slides explaining what the PSD is and how to interpret it

Fourier Transform

- For more convenient notation, if we allow the Fourier coefficients to be complex-valued we can write the much simpler expression

$$h(t) = \sum_{n=-\infty}^{\infty} H_n\, e^{i2\pi n f_0 t}$$

where

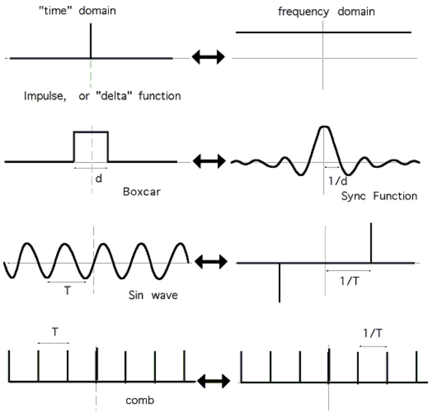
$$H_n = \frac{1}{\tau}\int_{-\tau/2}^{\tau/2} h(t)\, e^{-i2\pi n f_0 t}\, dt$$

- In the limit as $\tau \to \infty$, the separation between Fourier components $f_0 = 1/\tau \to 0$ and the H_n become a continous function $H(f)$ known as the Fourier transform (FT):

$$H(f) = \int_{-\infty}^{\infty} h(t)\, e^{-i2\pi f t}\, dt$$

The FT decomposes $h(t)$ into the frequencies that contribute to it

Time-Frequency Relationship



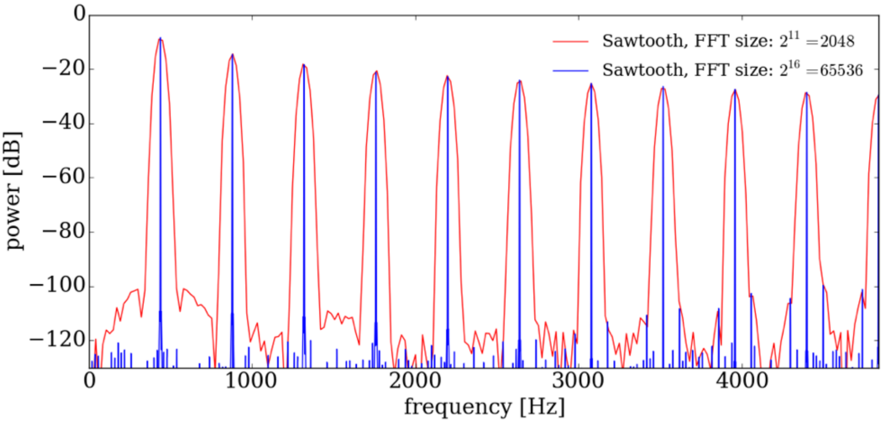
- When $h(t)$ is more concentrated, $H(f)$ becomes spread out, and vice-versa
- Gabor limit: uncertainty relation in time and frequency analysis. Follows because t and f are Fourier pairs
- For a measure of bandwidth Δf and a measure of time duration Δt (e.g., variances),

$$\Delta t \Delta f \geq 1$$

- Proof: use definition of variance with the Cauchy-Schwartz Inequality

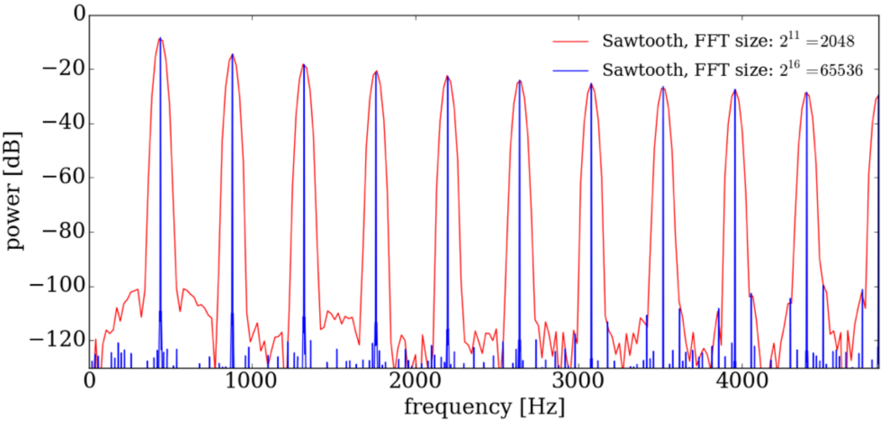
Example: Effect of Sample Length

Increasing the length of a sampled waveform from $N = 2048$ to $N = 65536$ samples improves the localization of the frequency peaks. Why?



Example: Effect of Sample Length

Increasing the length of a sampled waveform from $N = 2048$ to $N = 65536$ samples improves the localization of the frequency peaks. Why?



Small $N \implies$ waveform localized in time, worse resolution in frequency

Discrete Fourier Transform

- Suppose that $h(t)$ is sampled in N intervals lasting T seconds each, so that the function is given by N equally-spaced samples $h_k = h(kT)$ for $k = 0, 1, \ldots, N-1$
- We calculate $H(f)$ at the discrete frequencies

$$f_n = \frac{n}{NT}, \qquad n = -\frac{N}{2}, \ldots, \frac{N}{2}$$

where we obtain useful information only when $|f| < f_{\text{Nyquist}} = 1/(2T)$

- The Discrete Fourier Transform (DFT) is

$$\begin{aligned} H(f) &= \int_{-\infty}^{\infty} h(t)\, e^{-i2\pi ft}\, dt \\ &\approx \sum_{k=0}^{N-1} h(kT)\, e^{-i2\pi f_n kT}\, T = T \sum_{k=0}^{N-1} h_k\, e^{-i2\pi nk/N} = TH_n \\ \therefore H_n &= \sum_{k=0}^{N-1} h_k\, e^{-i2\pi nk/N} \end{aligned}$$

Discrete Power Spectral Density

- The power spectrum, or power spectral density (PSD), is the power per unit cycle of h
- Energy is defined by Parseval's theorem:

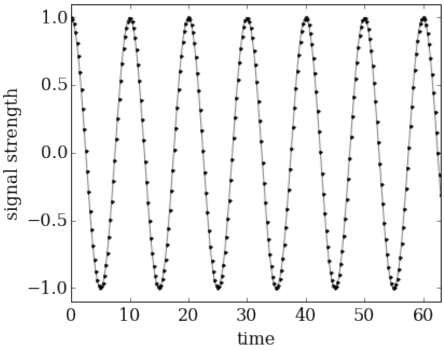
$$\begin{aligned}\text{Energy} &= \int_{-\infty}^{\infty} h^2(t)\,dt = \int_{-\infty}^{\infty} |H(f)|^2\,df \\ &= \sum_{k=0}^{N-1} h_k^2 T = \sum_{n=0}^{N-1} |H(f_n)|^2 \Delta f\end{aligned}$$

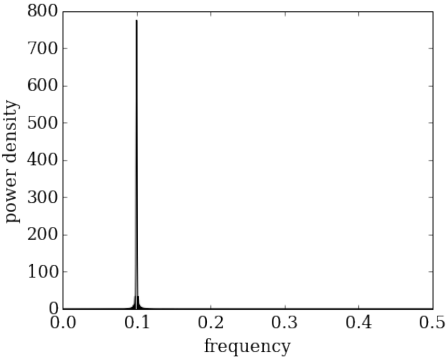
  Power is energy/(waveform duration), where duration is NT
- For a sampled waveform, the PSD $\propto$ (Fourier coefficient)^2

$$P(f_n) = \begin{cases} T/N\,|H_0|^2, & n = 0 \\ T/N\,\left[|H_n|^2 - |H_{N-n}|^2\right], & n = 1, 2, \ldots, (N/2-1) \\ T/N\,|H_{N/2}|^2, & f_{N/2} = \text{Nyquist frequency}\end{cases}$$

Periodogram

The PSD is how we typically investigate the number and relative strength of frequency constributions to a signal:

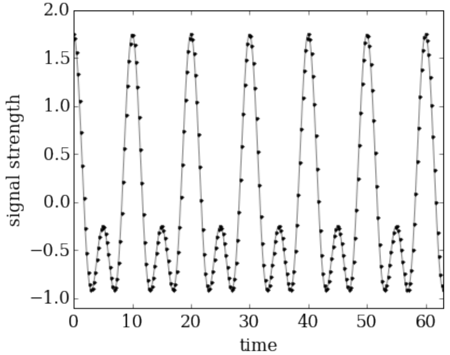


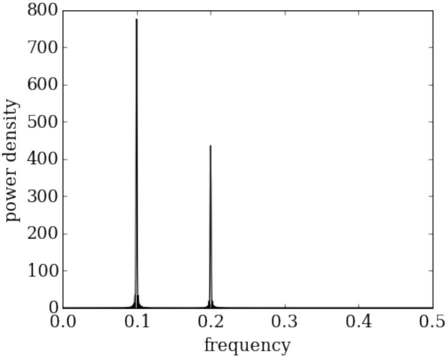


In f-domain problems we call plots of the PSD periodograms. It is just proportional to the square modulus of the (discrete) Fourier transform

Periodogram: Two Frequencies

The PSD is how we typically investigate the number and relative strength of frequency constributions to a signal:

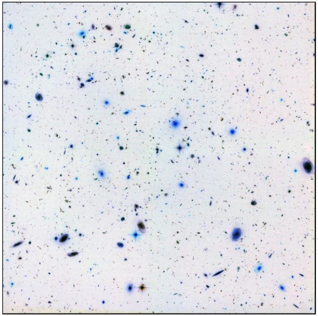




It is just proportional to the square modulus of the (discrete) Fourier transform

Example: Power Spectrum of Galaxies

The power spectrum doesn't just have to involve t and f; it can be calculated for any Fourier pair such as position $\boldsymbol{x}$ and wavenumber $\boldsymbol{k}$



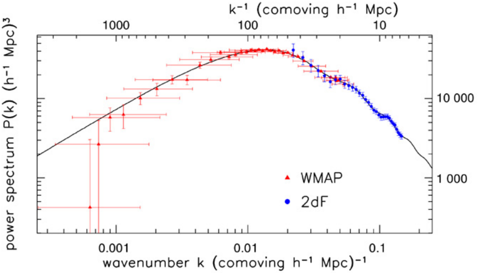
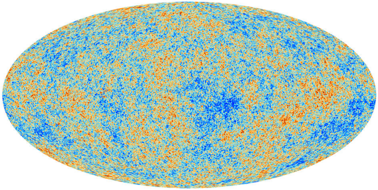


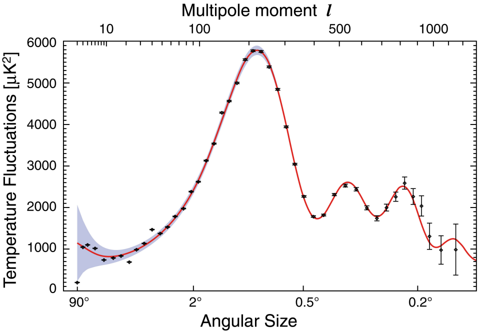
Fig 8.17 (A. Sanchez) 'Galaxies in the Universe' Sparke/Gallagher CUP 2007

How far away from each galaxy are other galaxies? The power spectrum tells us. We care because it is sensitive to the ratio of dark matter in the universe (23%) to normal matter (4%)

Example: Angular TT Power Spectrum of the CMB

Another Fourier pair: angular separation θ between hot and cold spots in the CMB and multipole ℓ



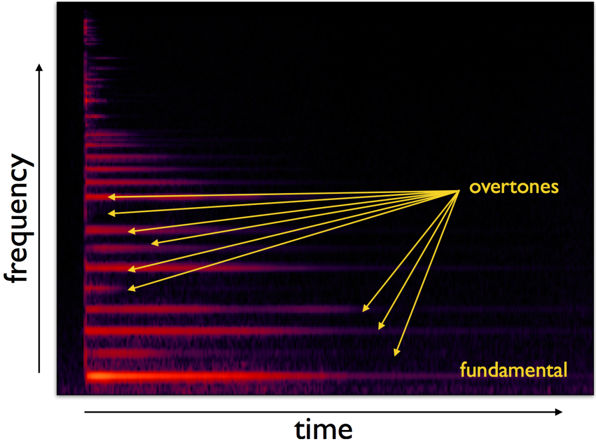


Note: the temperature-temperature (TT) power spectrum comes from

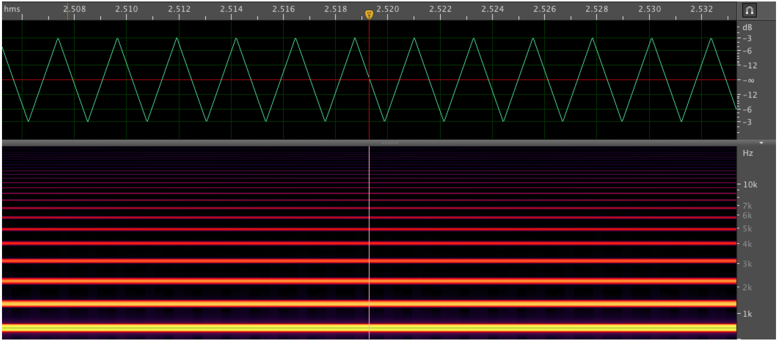
$$\mathrm{TT} = \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\theta, \varphi), \qquad C_{\ell}^{TT} \propto \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2$$

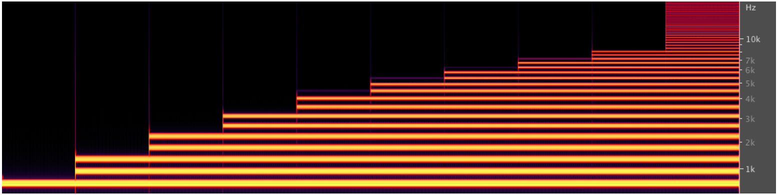
Spectrogram

It is possible for the PSD to evolve in time; we show this with a spectrogram. Below: spectrogram of a plucked harp:



Triangle Wave Built Up from Fourier Terms

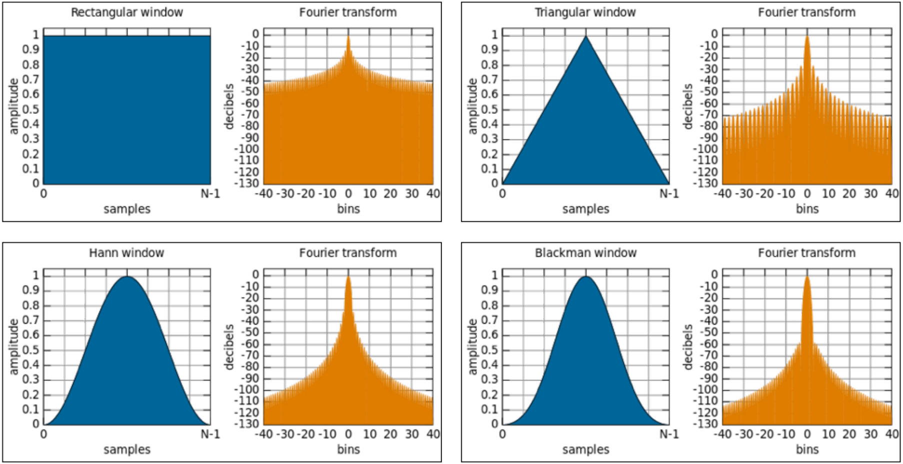




Need for Apodization/Windowing

- When we deal with real data, we don't have infinitely long time samples of periodic functions
- As a result, your FFT will contain artifacts (sidebands) because we've "cut off" the ends of the function in the time domain
- Windowing or apodization is used to reduce the creation of artificial structures in the frequency domain due to our cutoffs in the time domain
- **Concept**: we convolve our signal with a function that drops slowly and smoothly to zero at the edges of the sampling window
- **Result**: the FFT reflects the underlying periodic signal with much less contamination from the cutoffs

Common Window Functions



Example: FFT of a Sampled Square Wave

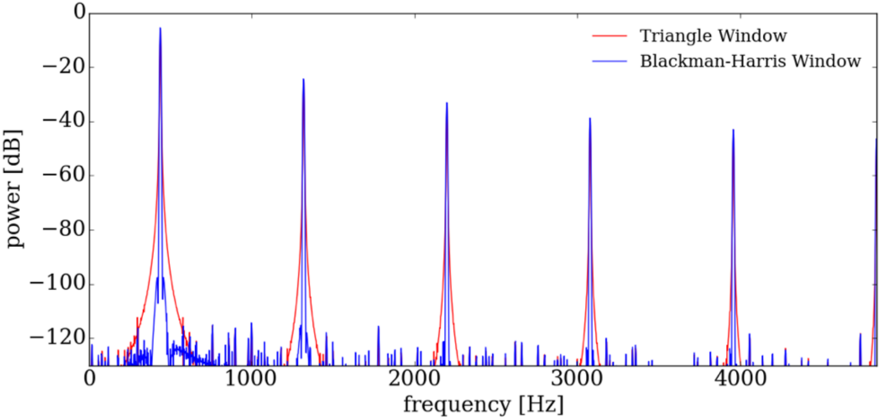


Table of Contents

Bayesian Insight into the Periodogram

- Let's get back to the example from the start of the presentation: given a time series with noise, we want to test if the data are sinusoidal with frequency f, where the model is:

$$y(t) = A\cos\left(2\pi ft + \varphi\right)$$

- Given data D and Gaussian noise of known size σ, we solve

$$\begin{aligned} p(f|D,\sigma,I) &\propto p(D|f,\sigma,I)\,p(f|I) \\ &= \int dA \int d\varphi\, p(D|A,f,\varphi,\sigma,I)\,p(A|I)\,p(f|I)\,p(\varphi|I) \end{aligned}$$

- The uncertainties are Gaussian, so the likelihood is

$$p(D|A,f,\varphi,\sigma,I) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\frac{(d_i - y(t_i))^2}{\sigma^2}\right]$$

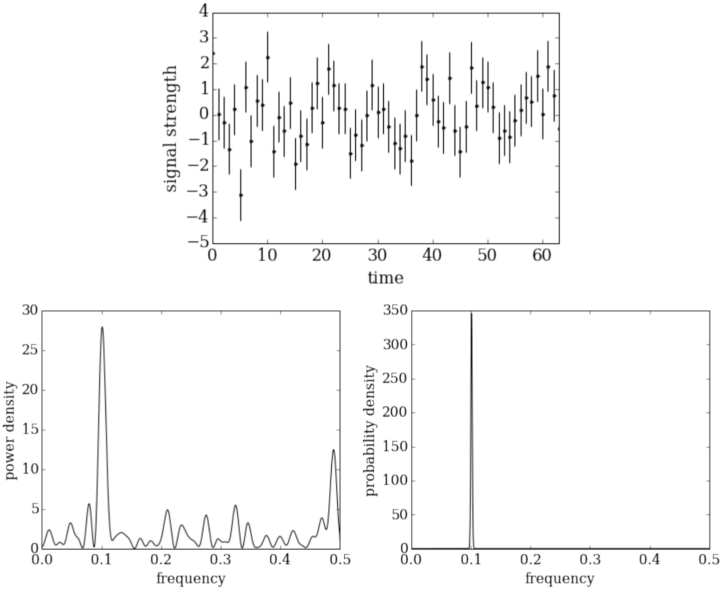
Bayesian Insight into the Periodogram

- Choose a uniform prior for amplitude A and a unifor prior for the phase φ
- Marginalizing the unwanted parameters A and φ gives

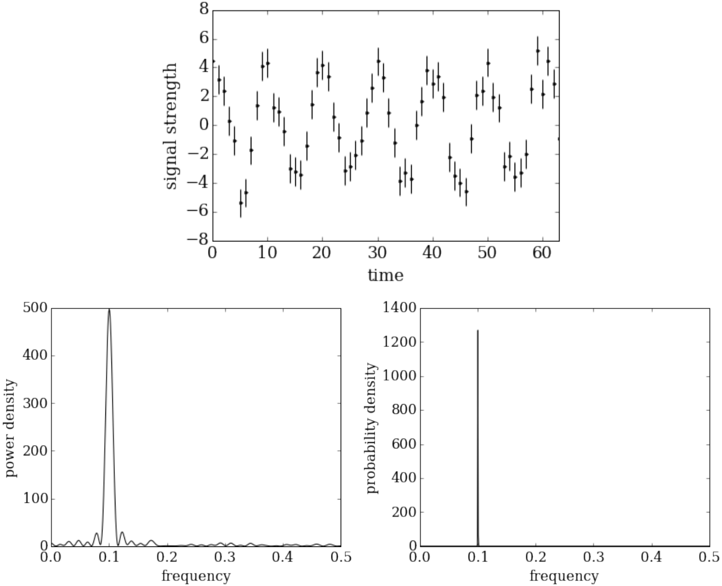
$$p(f_n|D,\sigma,I) \propto \exp\left[\frac{C(f_n)}{\sigma^2}\right],$$

  where $C(f_n) = |H_n|^2/N \propto \text{PSD}$ and σ^2 is the known variance of the noise [4]
- The Bayesian analysis shows that the DFT is the optimal estimator of f if N is large, DC offsets have been removed, there are no lower frequencies, the data contain just one frequency, A and φ are constant, and the noise is Gaussian [4]
- **Bonus**: the expression naturally attenuates noise features in the base of the PSD without requiring any kind of smoothing

Periodogram with Low Signal/Noise Ratio



Periodogram with High Signal/Noise Ratio



Bayesian Periodogram with Unknown Variance

- Suppose that we actually don't know the variance of the data. In this case it also becomes a nuisance parameter that needs to be marginalized:

$$p(f|D,I) = \int dA \int d\varphi \int d\sigma \, p(D|A,f,\varphi,\sigma,I) \\ p(A|I)\, p(f|I)\, p(\sigma|I)\, p(\varphi|I)$$

- Since σ is a scale parameter we use $p(\sigma|I) \propto 1/\sigma$, giving

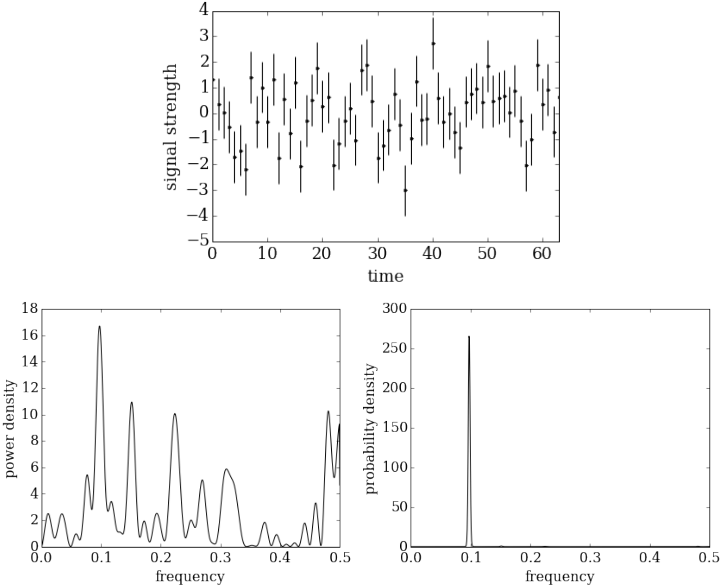
$$p(f_n|D,I) \propto \left[1 - \frac{2C(f_n)}{N\overline{d^2}}\right]^{\frac{2-N}{2}},$$

which looks like a Student t distribution. Note that

$$\overline{d^2} = \frac{1}{N}\sum_{i=0}^{N} d_i^2$$

is the mean square average of the data values.

Periodogram with Unknown Variance in Data



Nonuniform Sampling

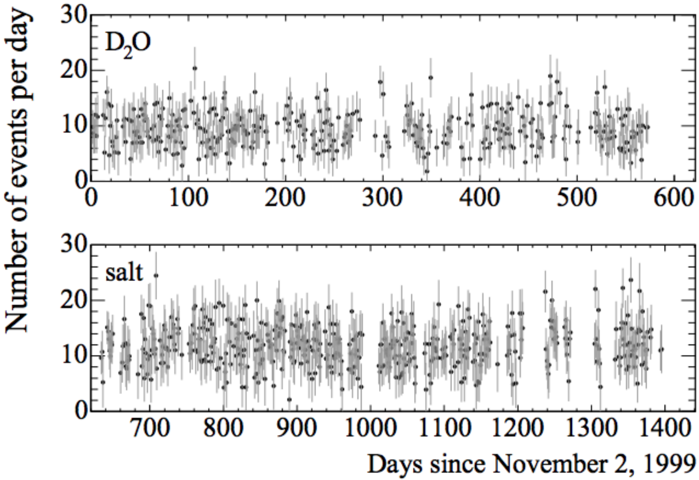
- An extremely common problem in time-domain analysis is nonuniform sampling caused by downtime
- New approach: fit a new model to the data of the form

$$y(t_i) = A\cos\left(2\pi f t_i - \theta\right)Z(t_i) + B\sin\left(2\pi f t_i - \theta\right)Z(t_i)$$

- A and B are the amplitudes of the sine and cosine functions (equivalent to one amplitue plus phase in a cosine function)
- $Z(t)$ is a weighting function that accounts for missing data or any other effect of importance
- θ is defined to make the sine and cosine functions orthogonal on discretely sampled times

Example: Searching for Periodicities in ^{8}B Solar ν Flux

Looking for periodicity in solar ν flux in SNO detector, D_2O and salt mode [5]. Note the gaps in the data:



Classical Solution: Lomb-Scargle Periodogram

The classical solution to non-uniform sampling is called a Lomb-Scargle periodogram:

$$\overline{h^2} = \frac{R(f)^2}{C(f)} + \frac{I(f)^2}{S(f)}$$

where

$$R(f) = \sum_{i=1}^{N} d(t_i)\ \cos\left(2\pi f t_i - \theta\right) Z(t_i)$$

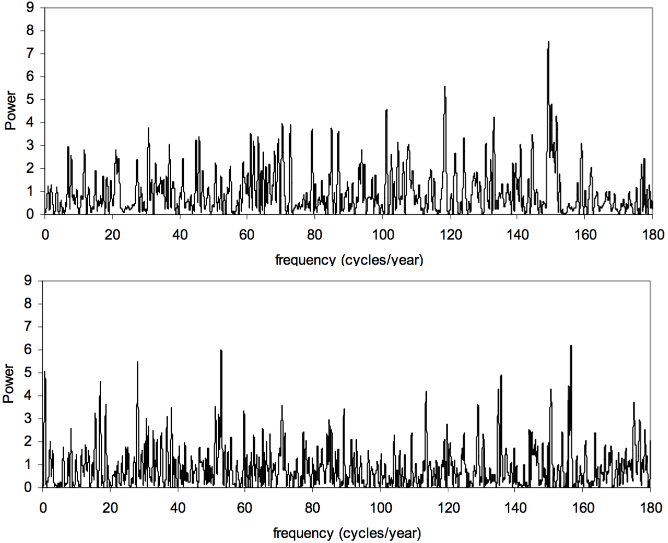
$$I(f) = \sum_{i=1}^{N} d(t_i)\ \sin\left(2\pi f t_i - \theta\right) Z(t_i)$$

$$C(f) = \sum_{i=1}^{N} \cos^2\left(2\pi f t_i - \theta\right) Z(t_i)^2$$

$$S(f) = \sum_{i=1}^{N} \sin^2\left(2\pi f t_i - \theta\right) Z(t_i)^2$$

Example: ^{8}B Solar ν Flux

SNO Lomb-Scargle periodograms for D_2O (top) and salt (bottom) [6]:



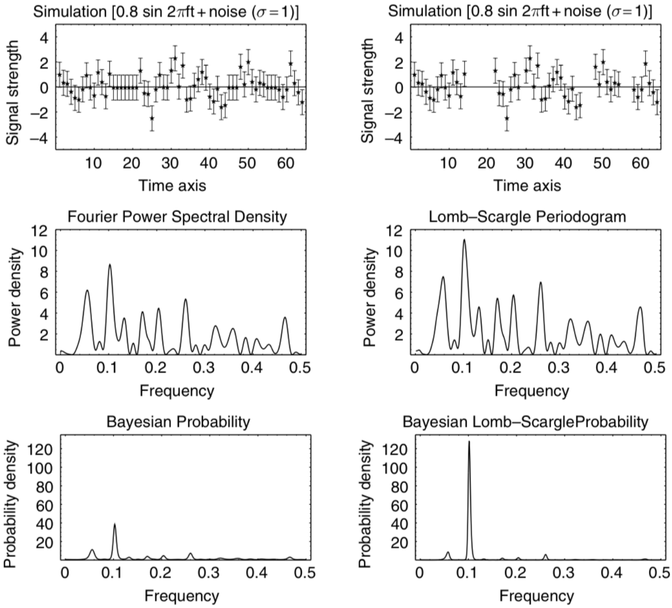
Bayesian Lomb-Scargle Calculation

- Assume independent uniform priors for A and B
- Assume a Jeffreys prior for the noise variance σ. Hence, any variation not described by the model is assumed to be noise
- Putting it all together:

$$p(f_n|D,I) \propto \frac{1}{\sqrt{C(f_n)S(f_n)}} \left[N\overline{d^2} - \overline{h^2}\right]^{\frac{2-N}{2}}$$

- Like the periodogram with uniform sampling, $p(f_n|D,I)$ involves a nonlinear processing of the Lomb-Scargle periodogram
- Spurious base features in the periodogram are attenuated

Bayesian Lomb-Scargle Calculation



Summary

- For a sampled signal, the (Schuster) periodogram given by the PSD of the signal is an easy way of picking out the frequency components in a signal
- The periodogram can be derived from first principles in a Bayesian analysis by marginalizing the amplitude and phase of a periodic signal
- The Bayesian periodogram goes like $p(f|D,I) \propto \exp\left[C(f_n)/\sigma^2\right]$, resulting in the natural attenuation of ripples below the main peak
- If one carries out a Bayesian analysis on a nonuniformly sampled signal, a version of the Lomb-Scargle periodogram pops out
- Note: this is not the only way to search for periodicity in an analysis, but it is probably the most popular

References I

[1] R. Abbasi et al. "The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping". In: *Nucl.Instrum.Meth.* A601 (2009), pp. 294–316. arXiv: `0810.4930 [physics.ins-det]`.

[2] H. Nyquist. "Certain topics in telegraph transmission theory". In: *Trans. AIEE* 47 (1928), pp. 617–644.

[3] Claude E. Shannon. "A Mathematical Theory of Communication". In: *Bell Sys. Tech. J.* 27 (1948), pp. 379–423.

[4] L. Bretthorst. *Bayesian Spectrum Analysis and Parameter Estimation*. New York: Springer-Verlag, 1988.

[5] B. Aharmim et al. "A Search for periodicities in the B-8 solar neutrino flux measured by the Sudbury neutrino observatory". In: *Phys.Rev.* D72 (2005), p. 052010. arXiv: `hep-ex/0507079 [hep-ex]`.

References II

[6] Gioacchino Ranucci and Marco Rovere. "Periodogram and likelihood periodicity search in the SNO solar neutrino data". In: *Phys.Rev.* D75 (2007), p. 013010. arXiv: `hep-ph/0605212 [hep-ph]`.