Astro 426/526

Fall 2019

Prof. Darcy Barron

Lecture 17: Analysis of Sequential Data

Miscellaneous

- Midterm solutions posted, final grades will be posted very shortly
- Class in P&A this week, in PAIS next week
- Homework from *Practical Statistics* coming soon (probably Monday, after proposal reviews complete)
- Next assignment: proposal reviews

Proposal Reviews

- Proposal review on Wednesday (and probably next Monday)
- 8 proposals to review
- Proposal Review Process
  - You have been assigned two proposals to review.
  - You are the primary reviewer for one proposal, and the secondary reviewer on a second proposal.
  - **Before Wednesday: read your two proposals and submit both a primary review and secondary review on Learn**
  - **Both Due: Wednesday, October 23 at 3:30pm**
  - Come to class prepared to present and lead discussion on your **primary** proposal assignment, with some help from secondary reviewer
  - Be prepared to write up a summary of the class discussion for your **secondary** proposal assignment, with some help from primary reviewer
  - Everyone will vote on score for proposal after class discussion

Can you start a fire with 1 Watt?

- https://www.youtube.com/watch?v=-CIUZBBz0Uc

Sequential Data – 1D Statistics

- Last week: overview of data analysis, some examples of data analysis in Jupyter notebooks
- Today: mostly covering material in *Practical Statistics for Astronomers:* Chapter 9, and Practical Statistics online lectures
- **What are some examples of sequential data?**

Observations with sequential data

- Intensity vs position as a single beam/pixel scans across the sky
- Signal variation along a row of a CCD
- Light curves (intensity vs time on a single pixel)
- Many other measurements vs time (including stock market, population, GDP, etc)
- Typical nomenclature (in astronomy)
  - Scans: spatial domain
  - Timestreams: time domain
  - Spectra – frequency/wavelength domain

What do we want to do with sequential data?

- Trend-finding: can we predict the future behavior?
- Establish a "baseline," so any signal on top of the baseline can be analyzed
- Faint signal detection, when signal and noise are comparable in magnitude
- Filter data to improve signal-to-noise ratio
- Quantify the level of noise
- Search for periodic signals
- Correlation: between antennas, or between spectra

Data transformations

- With these kinds of analyses, the features we are interested in only emerge after **transformation**
  - **Filtering:** remove known noise to find feature
  - **Transform** along known important feature of data
    - For example: Periodicity search, spectral-line correlator

- Expand data into **orthogonal functions**
- **Fourier transform** is just one option

Fourier Analysis

- Extremely common analysis technique, with many reasons for being **physically motivated**
  - Most physical processes at both macro and micro levels involve oscillation and frequency: orbits of galaxies, stars or planets, atomic transitions at particular frequencies, spatial frequencies on the sky as measured by correlated output from pairs of telescopes.
    - We want the frequencies composing data streams; just the amplitudes of these frequency components may be the answer (as in the case of detection of a spectral line).
  - In many physical sciences there is frequent need to measure a single signal from a data series. In measuring a specific attribute of this signal such as redshift, the power of Fourier analysis has long been recognized

Fourier Analysis

- Any continuous function may be represented as the sum of sines and cosines:
  - $f(t) = \int_{-\infty}^{+\infty} F(\omega) e^{-i\omega t} dt$
  - **F** is the phased amplitudes of the sinusoidal components of **f** – the Fourier Transform (FT)

Properties of Fourier Transforms

- The FT of a sine is a delta function in the frequency domain
- The FT of a Gaussian is another Gaussian (convenient!)
- The FT of $f \otimes g$ is F x G
- The FT of $f(t+\tau)$ is (FT of f)($e^{-i\omega\tau}$) (shift theorem)

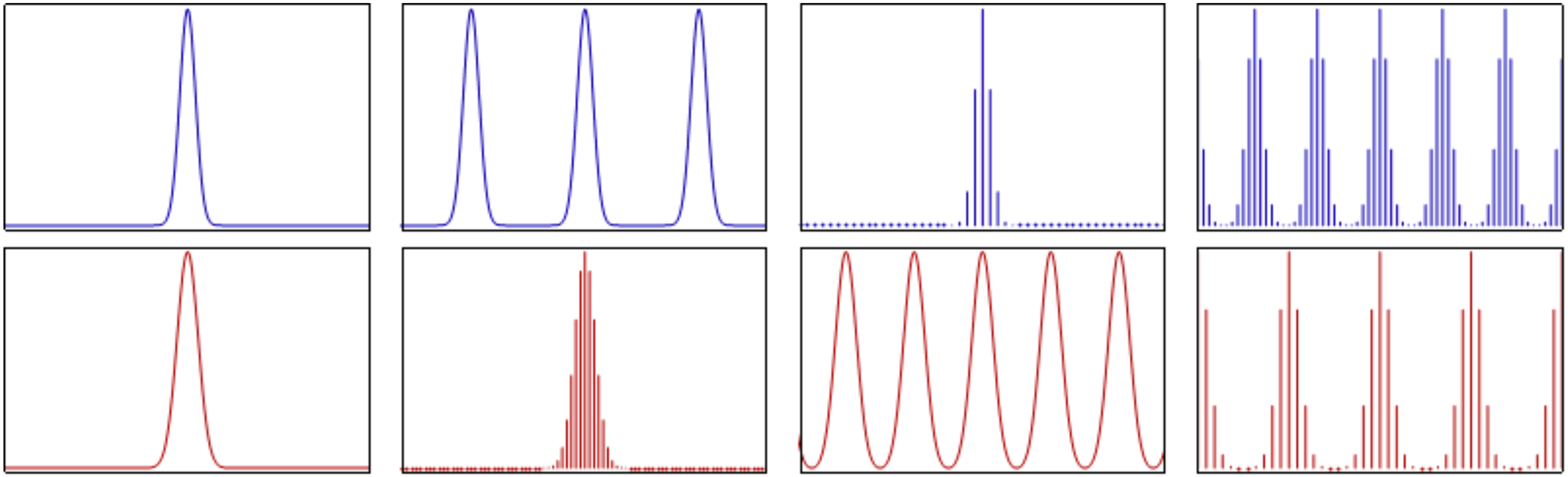
Fourier and Sampling

- Actual data is not continuous and infinite
- Can use the discrete Fourier Transform (DFT)
- With N data points taken at *uniform* interval Δt, DFT results in the continuous function multipled by the 'comb' function
- f'(t) can be represented as:
  - $f'(t) = A_n \sum sin(n\Delta v) + B_n \sum \cos(n\Delta v)$
  - Interval $\Delta v = 2\pi\Delta t$

The Fast Fourier Transform (FFT)

- FFT – Cooley and Tukey 1965
  Does the transform of N points in a time proportional to $N \log N$, rather than the N^2 timing of a brute-force implementation. This is a monumental cpu saver.
- Quirks (see Bracewell, or Numerical Recipes)
  - The typical (?) arrangement of its input / output data
  - normalization
- Critical to most image processing, critical to design of radio telescopes
- Algorithm was apparently known to Gauss – even before Fourier had discovered his series.
- It may be the most used algorithm on the planet. (Used for every .jpg image, for a start)

Discrete Fourier Transform



Relationship between the (continuous) Fourier transform and the discrete Fourier transform. Left column: A continuous function (top) and its Fourier transform (bottom). Center-left column: Periodic summation of the original function (top). Fourier transform (bottom) is zero except at discrete points. The inverse transform is a sum of sinusoids called Fourier series. Center-right column: Original function is discretized (multiplied by a Dirac comb) (top). Its Fourier transform (bottom) is a periodic summation (DTFT) of the original transform. Right column: The DFT (bottom) computes discrete samples of the continuous DTFT. The inverse DFT (top) is a periodic summation of the original samples. The FFT algorithm computes one cycle of the DFT and its inverse is one cycle of the inverse DFT.

https://en.wikipedia.org/wiki/Discrete_Fourier_transform

Five **criteria for successful discrete-sampling**

1. The **Nyquist criterion** or **Nyquist limit** guarantees **no** information at spatial frequencies above ***$\pi/\Delta t$.*** The sampling interval ***Δt*** sets the highest spatial frequency ***$2\pi/\Delta t$*** retained; higher frequencies present in the data are lost.

2. The **Sampling theorem: any** bandwidth-limited function can be specified **exactly** by regularly-sampled values provided that the sample interval does not exceed **a critical length** (approximately half the FWHM resolution), i.e. for an instrumental half-width ***B***, ***$f'(t) \rightarrow f(t)$*** if ***$\Delta t < B/2$***. Any physical system is band-pass limited**, preventing** full recovery of the signal.

3. To avoid any ambiguity - **aliasing** - in the reconstruction of the scan from its DFT, the **sampling interval must be small enough** for the amplitude coefficients of components at frequencies as high as ***$\pi/\Delta t$*** to be effectively zero. Otherwise there's a tangle with the negative tail of the repeating function → **ambiguity.**

4. The lowest frequencies are ***$2\pi/(N\Delta t)$***. Such low-frequency components may be real or instrumental; but to find signal the scan length must exceed the width of single resolved features by **> 10**.

5. The integration time per sample must be **long enough for decent S/N**.

Five criteria for successful discrete-sampling

1. The **Nyquist criterion** or **Nyquist limit** guarantees **no** information at spatial frequencies above ***π/Δt.*** The sampling interval ***Δt*** sets the highest spatial frequency ***2π/Δt*** retained; higher frequencies present in the data are lost.

Can draw out sine waves with sample points to see this effect

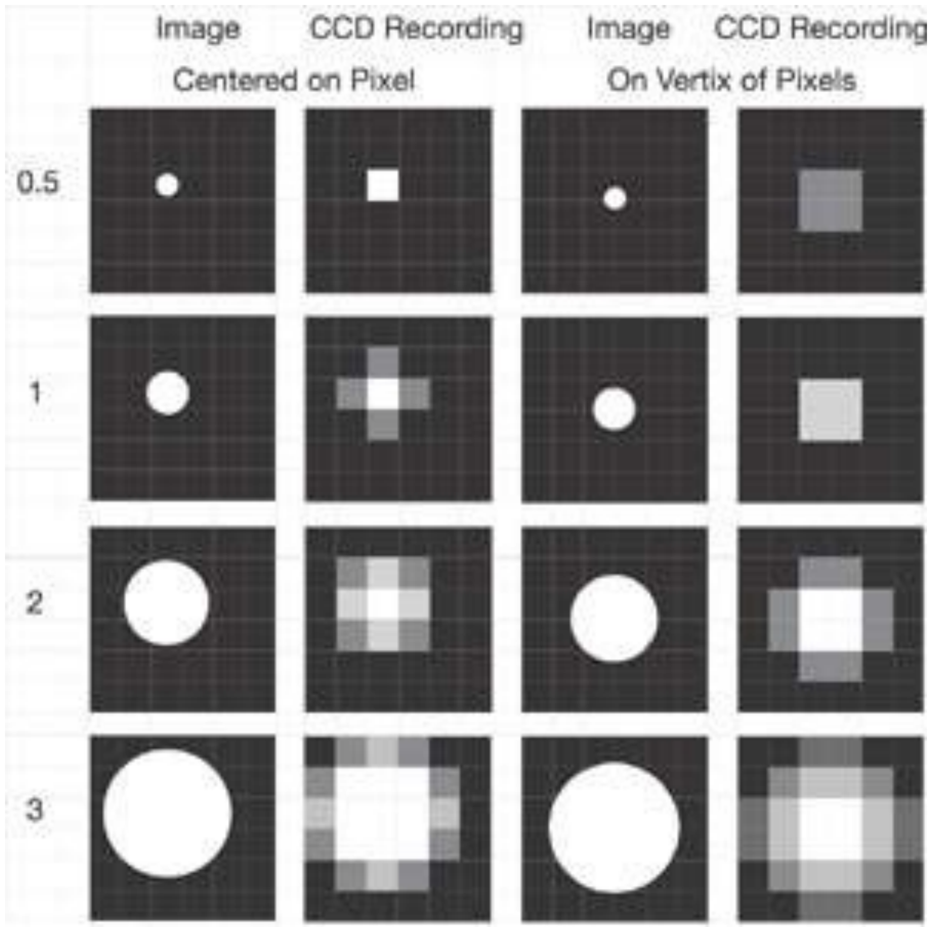
Five criteria for successful discrete-sampling

2. The **Sampling theorem: any** bandwidth-limited function can be specified **exactly** by regularly-sampled values provided that the sample interval does not exceed **a critical length** (approximately half the FWHM resolution)

i.e. for an instrumental half-width ***B***, ***f'(t)*** → ***f(t)*** if ***Δt < B/2***.

Any physical system is band-pass limited, which can prevent full recovery of the signal.

Sampling theorem

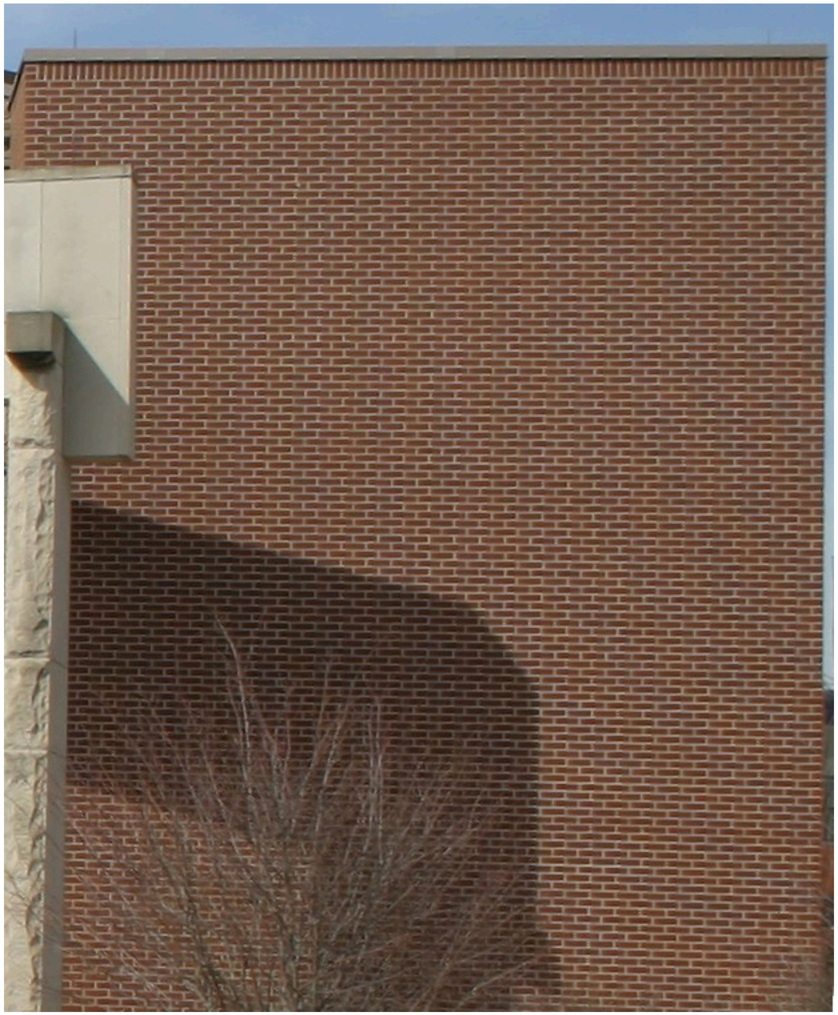


Need ~ three samples per resolvable unit (achieved resolution)

Five criteria for successful discrete-sampling

3. To avoid any ambiguity - **aliasing** - in the reconstruction of the scan from its DFT, the **sampling interval must be small enough** for the amplitude coefficients of components at frequencies as high as $\boldsymbol{\pi/\Delta t}$ to be effectively zero. Otherwise there's a tangle with the negative tail of the repeating function → **ambiguity.**

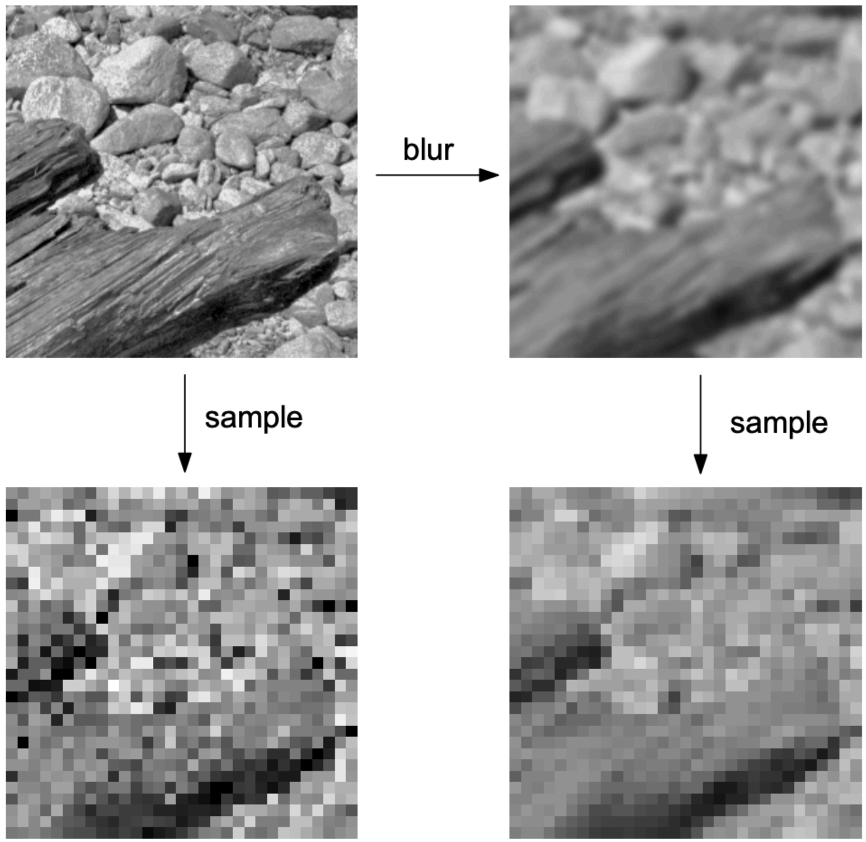
Example of aliasing



https://en.wikipedia.org/wiki/Nyquist%E2%80%93Shannon_sampling_theorem

3.2 Natural images

The avoidance of aliasing is also important in the design of the eye. For example, in the human eye the lens filters out any spatial variations finer than 60 cycles/degree. The Nyquist theorem tells us in this case that the photoreceptor spacing should be at least 120 sample/degree, which is exactly what we have! Consider though what the sampled retinal image might look like if this were not the case:



At upper left is shown a natural image that may typically fall on the retina, and below this is shown the result of subsampling the image without properly blurring beforehand. Compare this to the sampled version after blurring (lower right). It is not difficult to see why nature has gone to the effort it has to match photoreceptor spacing with the point-spread function (spatial-frequency cutoff) of the lens.

- http://www.rctn.org/bruno/npb261/aliasing.pdf

Five criteria for successful discrete-sampling

4. The lowest frequencies are $\boldsymbol{2\pi/(N\Delta t)}$. Such low-frequency components may be real or instrumental; but to find signal the scan length must exceed the width of single resolved features by **> 10**.

Example: long-period exoplanets

Five criteria for successful discrete-sampling

5. The integration time per sample must be **long enough for decent S/N**.

Interactive 2D fourier transforms

- http://www.jezzamon.com/fourier/index.html
- https://betterexplained.com/articles/an-interactive-guide-to-the-fourier-transform/

Example - Redshifts from Cross-Correlation

- http://articles.adsabs.harvard.edu/cgi-bin/nph-iarticle_query?1979AJ.....84.1511T&data_type=PDF_HIGH&whole_paper=YES&type=PRINTER&filetype=.pdf
- A Survey of Galaxy Redshifts. I. Data Reduction Techniques
  - John Tonry and Marc Davis
  - The Astronomical Journal, October 1979

Tonry and Davis 1979

- Galaxy spectrum g(n) with n = A lnλ +B (n is bin number)
- Template spectrum t(n), zero redshift, instrumentally-broadened.
- Set up DFTs: G(k) = Σn g(n) exp(-2πink/N) , and equiv for T(k)
- Then FT for cross-correlation
  - c(n) = g ¤ t(n) is C(k) = (1/Nσgσt) G(k) T*(k)
  - Now set g(n) = α t(n) ¤ b (n – δ)
- The galaxy spectrum is a multiple of the template spectrum convolved with a broadening function shifted by δ. This function accounts for the **velocity dispersion** and the **redshift**, which we seek
- Assume b(n) Gaussian, and likewise for c(n) , centered at δ
- Minimizing X^2(α,δ;b)=Σn [α t¤b(n-δ) -g(n)]^2

  is equivalent to maximizing (1/$\sigma_{t\ x\ b}$)c ¤ b(δ)

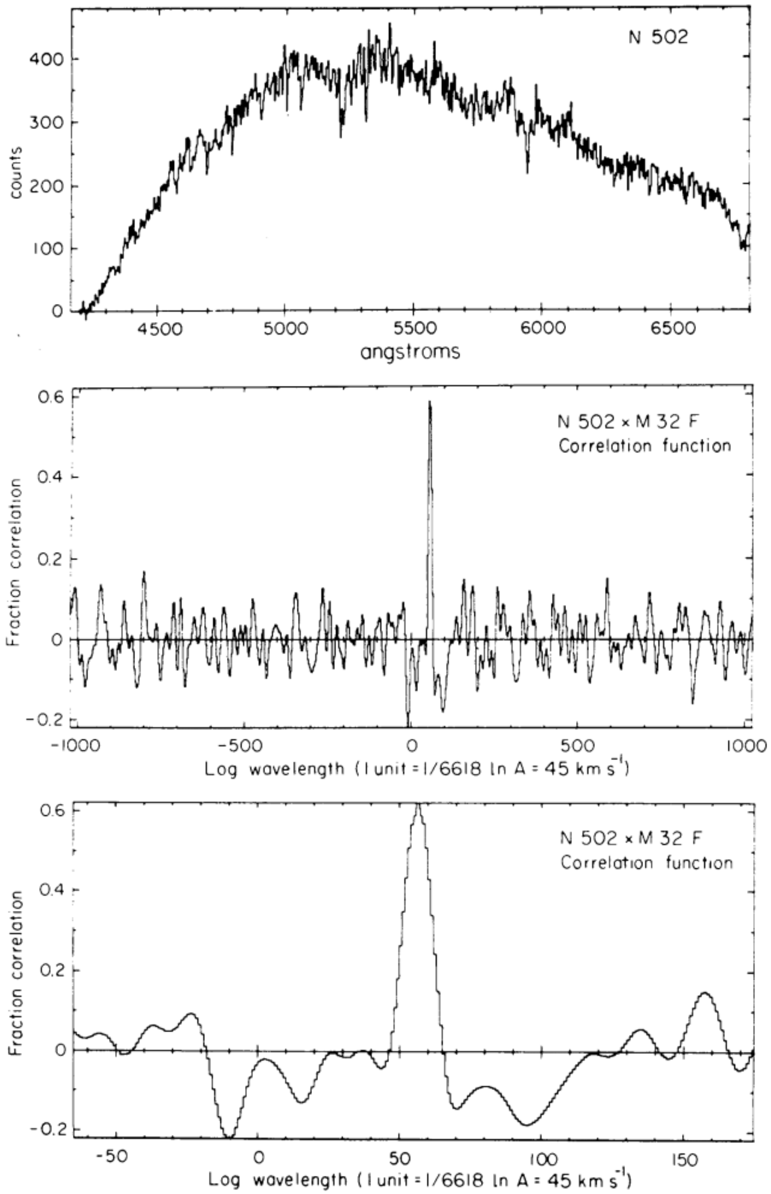


FIG. 10. Same as Fig. 8 for NGC 502.

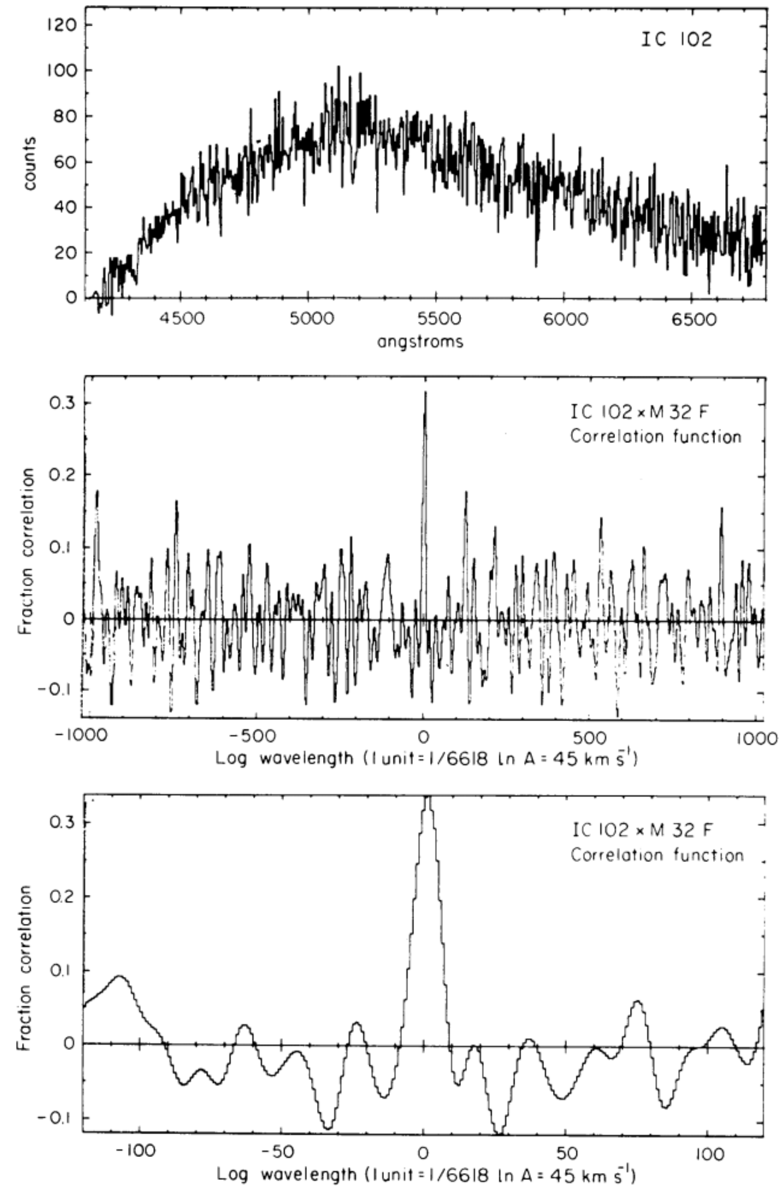


FIG. 11. Same as Fig. 8 for IC102.

Filtering

- Why filter?

Filtering

- To reduce noise
- To compress data

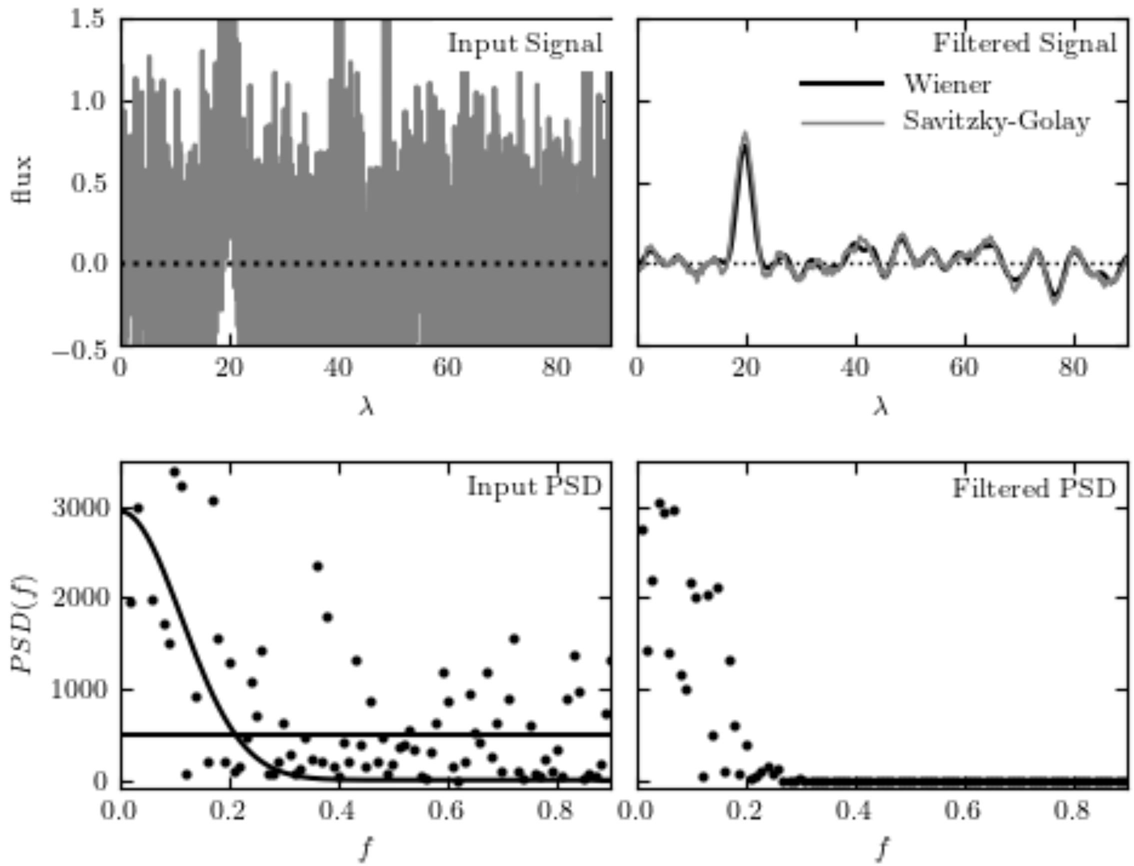
Examples of filtering

- Running average filter
- Low-pass filter
  - Wiener filter (https://reference.wolfram.com/language/ref/WienerFilter.html)
  - Savitsky-Golay filter
- High-pass filter
  - For example: removing a baseline drift

Filtering to reduce noise

- Many sources of noise are “white”
  - Photon noise, shot noise, random noise
  - “White” noise has a flat spectrum extending out to the limit given by the sampling theorem
- The FT of a Gaussian is another Gaussian
  - Other noise or instrumental effects are commonly Gaussian
  - Can taper off the amplitudes of **high frequencies**, where there is little information (low-pass filter)
- Filtering will decrease the noise, **but it must decrease the signal as well**
  - If you do it right, you still “win”

Wiener filter example



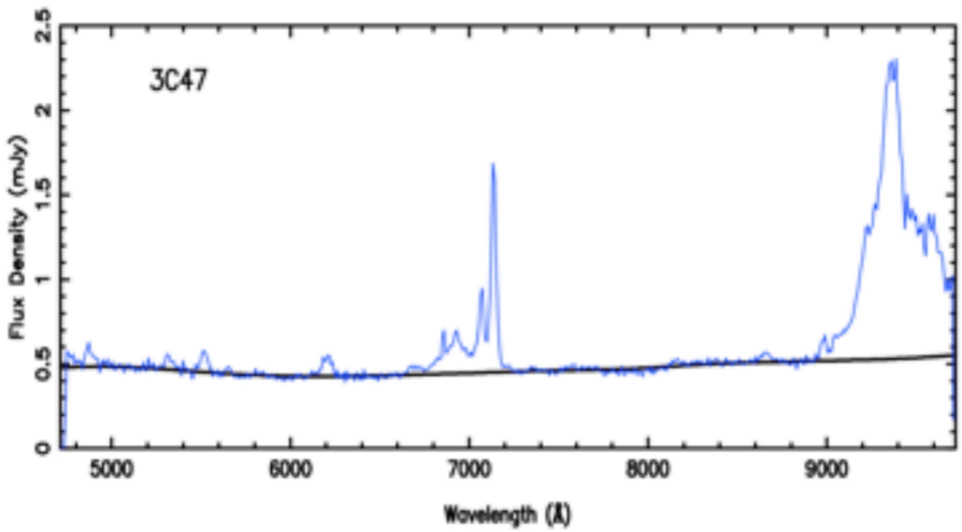
https://www.astroml.org/book_figures/chapter10/fig_wiener_filter.html

High-pass filtering

- Get rid of unwanted low frequencies
  - Fitting baselines
- Have to be careful to preserve signal, can have big signal component at low frequencies

Minimum-Component Baselines: Example

(Above) A spectrum of 3C47 obtained with the Faint Object Spectrograph of the William Herschel Telescope, La Palma. The redshift is 0.345; broad lines of the hydrogen Balmer series can be seen, together with narrow lines of [OIII].



(Below) A spectrum of the A star RZ Cas (Maxted et al. 1994). The continuum obtained with the minimum-component technique is shown as the black line superposed on the original data.