Phys 641

Jonathan Sievers (jonathan.sievers@mcgill.ca)

https://github.com/sievers/phys641_2022

TA: Anan Lu (anan.lu@mail.mcgill.ca)

Course Aim

- Meet basic techniques at a variety of wavelengths
- and messengers! (GW, neutrinos ?, CRs ?…)
- Mathematical tools to deal with data extremely important, often skipped over in undergrad
- Major component will be providing practical tools (statistics, Fourier tranforms, linear algebra, model fitting, matched filters…) that are applicable to a wide range of situations.

Mock TAC

- Over course of career, you will almost certainly need to apply for telescope time.
- Everyone will have to write a mock proposal.
- Last week of class, proposals will be reviewed by mock TAC (time allocation committee), consisting of you.
- Primary/secondary reviewers will be assigned for each proposal. Primary leads discussion, secondary writes the report. All people should be ready to comment on any proposal.
- Think about what project you would like to propose, be ready to write science justification (why we should care) and technical justification (how much time you'll need, what telescope modes etc.)

Grade

- 50% problem sets/in-class solutions
- 25% mock time application
- 25% role on TAC - read those proposals!

Some Useful Sources

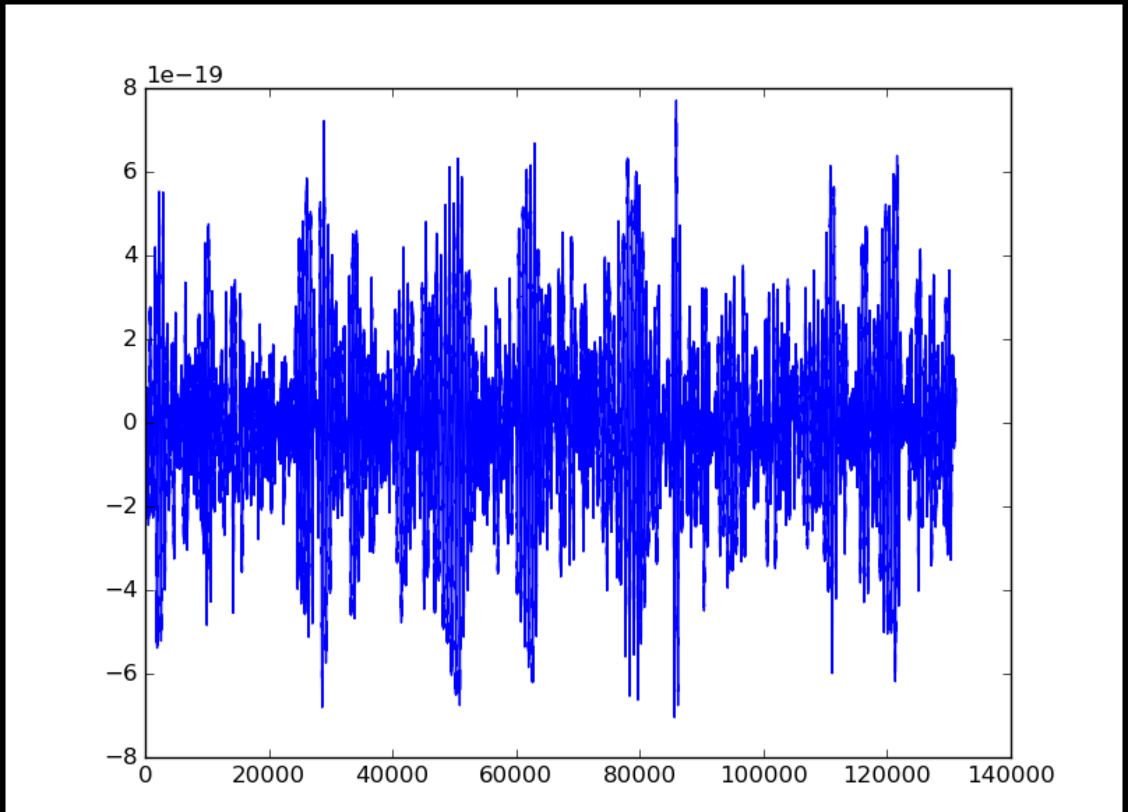
- Bevington (and Robinson) - data analysis
- Thomson Moran & Swenson - bible of radio astronomy
- Saulson - gravitational wave detectors (Adhikari - review 1305.5188)
- Rieke "Measuring the Universe"
- more…

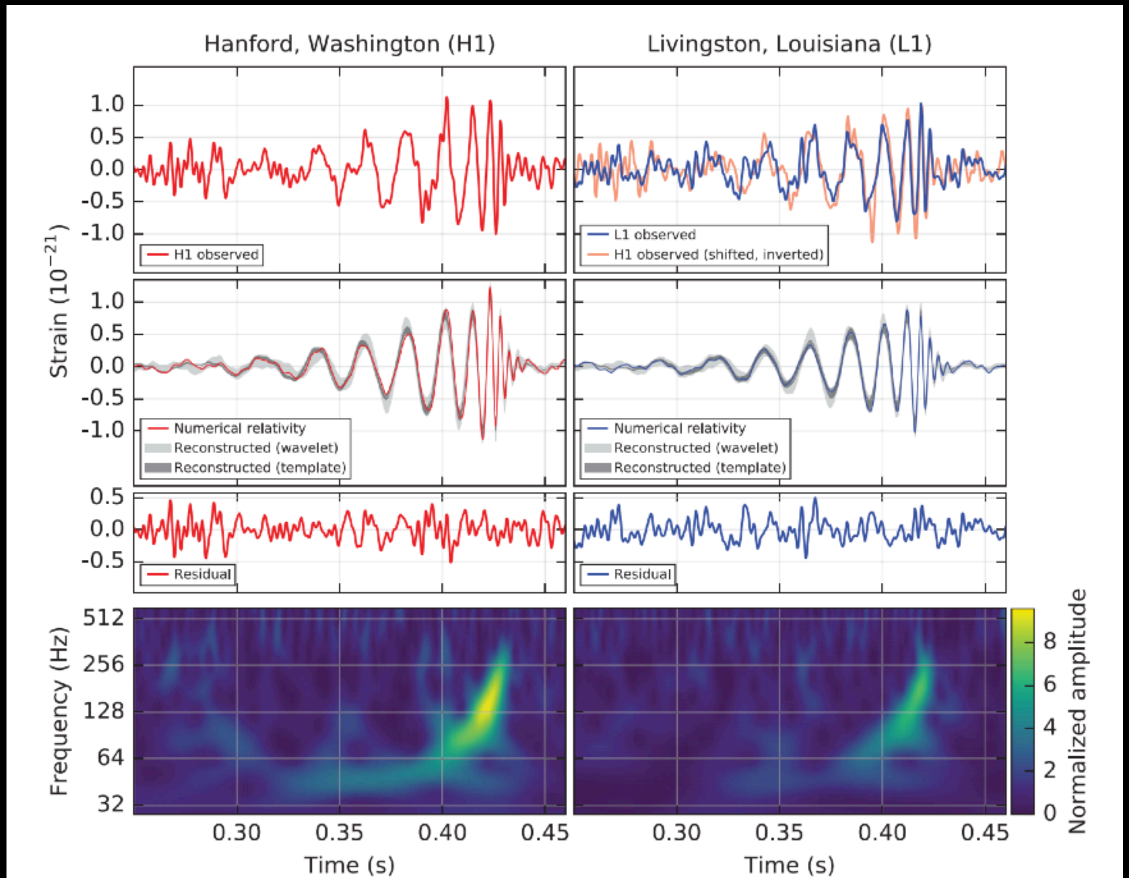
Data Analysis

- “Data analysis is a struggle between you and your computer.” - Alan Weinstein
- Will use python (has everyone used python before?).
- Best computational algorithms may be (quite) different than simplest analytic ones (e.g. fitting polynomials)
- Will point some of these out as we go along.

(Semi-)Near-term Outline

- Go over some basic tools in analysis toolbox.
- Gravitational wave detection a nice, simple case where many of these things come up.
- Basically, how do you go from left to right (and get a Nobel prize in the process)?





To do this…

- If you don't have errors, you haven't done science.
- Look at some basic probability distributions
- Derive χ^2 - everyone's first stop (because of central limit theorem).
- How do we work out χ^2 when the noise has correlations?
- Fourier transforms, Wiener-Khinchin theorem, linear algebra.
- How do we search for signals? Matched filters, which use Fourier transforms, convolution theorem.

Probability Distributions

- Binomial - if I flip a (possibly biased) coin, how many heads/tails might I expect? How many flips to tell if a coin is biased?
- Poisson - limit of binomial. How many photons might I count? Neutrinos? Cosmic rays?
- Gaussian - when you have a hammer, everything looks like a nail…

(this stuff all in Bevington)

Binomial

- If I flip a coin *n* times, what is the probability of getting *m* heads?
- What about if the probability per-flip of getting heads is *p*?
- Answer - (n chose m)$p^m(1\text{-}p)^n$.

Poisson

- If I have some background event rate, what is the probability of detecting k events when I expected λ?
- $e^{-\lambda} \lambda^k/k!$
- Let's derive from binomial…
- How would you calculate probability of getting 10,000 events on a computer?

Gaussian

- basic PDF - $\exp(-0.5(x-\mu)^2/\sigma^2)/\sqrt{(2\pi\sigma^2)}$
- what is the probability of a bunch of (uncorrelated) data points, assuming I know the noise?
  - $\Pi \exp(-0.5(x_i-\mu_i)^2/\sigma_i^2)/\sqrt{(2\pi\sigma_i^2)}$
- If I have two different models for the means, what is the *relative* probability they would have produced the observed data?
- Can you show that a Poisson distribution converges to Gaussian for large λ,k?

χ^2

- Let's take -2ln(PDF)
- That turns into $\sum (x_i-\mu_i)^2/(\sigma_i^2) + \sum \ln(2\pi\sigma_i^2)$
- Second term does not change as we change our model.
- The relative (log-)likelihood between two models is just the first term. We call it that χ^2.
- The *relative* probability of two models giving rise to the observed data is $e^{\delta\chi^2/2}$. where lower χ^2 means higher likelihood.

Linear Algebraing up χ^2

- Usual expression is $\sum(x_i-\mu_i)^2/\sigma_i^2$
- Let N be diagonal matrix with $N_{ii}=\sigma_i^2$.
- Element-wise, $(x-\mu)^T N^{-1}(x-\mu)$ is identically χ^2.
- I can put orthogonal matrices ($V^T=V^{-1}$) in without changing anything: $(x-\mu)^T V^T V N^{-1} V^T V(x-\mu)$.
- In new, rotated coordinates: x->Vx, μ->Vμ, N->VNV^T, χ^2 remains unchanged. Show that expectation of (rotated) x_i noise times x_j noise = (rotated) N_{ij}?

Linear Least-Squares

- Let's say we have a model that depends *linearly* on a set of parameters: $<d>=Am$. What is the set of parameters that most probably gave the observed data?

- $\chi^2=(d-Am)^TN^{-1}(d-Am)$

- differentiate w.r.t. m: $\nabla\chi^2=-2A^TN^{-1}(d-Am)=0$ at minimum

- Group: $A^TN^{-1}Am=ATN^{-1}d$

- Sometimes - not always, we can solve explicitly: $m=(A^TN^{-1}A)^{-1}A^TN^{-1}d$

- These equations cover the vast majority of data analysis questions we'll face. Often they may *look* different, but usually aren't.

Stationary Noise

- For much of what we do, we need to make $N^{-1}x$, but not necessarily write down N^{-1} explicitly.
- For special case of $N^{ij}=f(i-j)$ (i.e. noise statistics are constant with time), we can switch to Fourier space.

Fourier Transforms

- On computer, discrete Fourier transform usually defined as $\sum f(x)\exp(-2\pi ikx/N)$, sum goes from 0 to N-1
- Can write this as a matrix multiply where ij^{th} element is $\exp(-2\pi ikx/N)$
- What is i^{th} column dotted with j^{th} column?
- what is the inverse of this matrix?
- Strongly encourage you to just memorize this. Numerical factors important in real life…

Some Fourier Theorems

- $x \to x+\delta$, $F(k) \to F(k)\exp(-2\pi i\delta k/N)$
- $x \to -x$, $F(k) \to F^*(k)$
- Convolution theorem: $FT(f \otimes g) = FT(f)*FT(g)$