Lab #4: Radio Telescopes and Pulsars

Lab report is due Wednesday, Feb 6th, 2019, before 11:59 pm EDT

1. Overview

This handout provides a description of your 4th lab, which explores the real-world use of a radio telescope, both through archival data, and (time & technology permitting), using newly acquired observations. This lab builds on many of the skills learned during the first term, but will require you to handle larger data sets, develop more complex data processing algorithms, and fit models to your data.

> *IMPORTANT: The three labs this term will give you more latitude to explore than the first three, and hence include fewer explicit instructions on how to proceed. If you are lost or need help, attend tutorials, lectures, or office hours, ask your classmates, google around for tips, and by all means ask questions.*

1.1 Schedule

This is a four-week lab, with lectures and sessions the weeks of January 7th, 14th, 21st, and 28th. It begins with analyses on archival data, and you should get started on those asap, even before your lab groups are finalized. You should aim to complete through section 4.1 during the 1st week, section 4.2 during the 2nd and 3rd, and be sure your data looks reasonable for 4.3 by the end of the 3rd lab week. Time permitting, you can tackle 4.4, but no marks will be lost for those who don't have time. As in last term, **do not underestimate the time the analysis and write-up will take!** The lab report should to be submitted electronically, on or before Feb 6th at 11:59pm EDT.

1.2 Goals

Investigate how radio antennas work, and examine one aspect of the time-variable radio sky, pulsars. Characterize and explore a radio system based on measured responses to known sources. Study the behavior of time-domain signals, and generate non-parametric models for use in matched filters.

1.3 Reading assignments

- Signal Processing:
  - Fourier Transforms, Matched Filters
- Radiometers and the Radiometer Equation
  - Essential Radio Astronomy (Condon & Ransom, 2016) Ch 3.1-3.2, 3.5. Available online, http://www.cv.nrao.edu/~sransom/web/Ch3.html
- Pulsars
  - ERA Ch 6, http://www.cv.nrao.edu/~sransom/web/Ch6.html
- Python: arrays, functions, and advanced plotting

1.4 Key Steps

This lab is broken down into several stages.

1. Review the components of a radio telescope, and how they couple to one another.
2. Download the archival data sets from the **Algonquin Radio Observatory** 46m dish.
    a. Calculate various properties (beamwidth, effective aperture, noise temperature and gain) for the telescope.
    b. Inspect the observation of the pulsar B0329+54.
3. Gather data using the ***Small Radio Telescope*** on the roof of McLennan Physical Labs. (The receiver has been struggling with the RFI in Toronto, so there's a risk this will be out of commission. Stay tuned!)
    a. Gather data on a bright source transit: the Sun, Cassiopeia A, or Cygnus A. Calculate beamwidths, dish illumination, and estimate system temperature.
    b. Take your own pulsar data! Can you detect B0329+54?

2. Antennas, feeds, reflectors, and telescopes.

In this lab, you'll be working with data collected from a few different telescopes. Since radio astronomy uses fairly different terms from other branches of the field, it's worth reviewing the main components.

2.1 Radio feeds

As you saw briefly in lab #2, **feeds** are used to couple radio light from free space into a **waveguide**, which can then be fed into a **receiver**. Feeds operate on a basic principle of resonance: incident electric fields push around electrons inside the feed to induce a current in the waveguide, but will only do so efficiently for periodic excitations, i.e. for specific wavelengths or frequencies of incoming light.

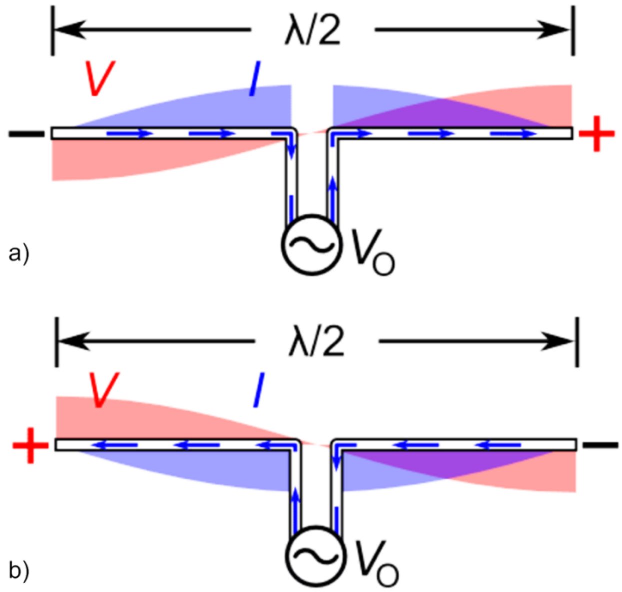

Figure 1 – A resonant dipole feed. Adapted from [Wikipedia](#).

More importantly for the present discussion, they only respond to light incident from certain directions. For example, an ideal dipole is "blind" to light travelling along its primary axis (left-to-right or right-to-left in **Figure 1**), as the electric fields cannot move any electrons in the feed itself. A feed's radiation pattern describes how it would distribute power if used to broadcast, and consequently (by the **antenna reciprocity theorem**), how well it absorbs light from different directions.

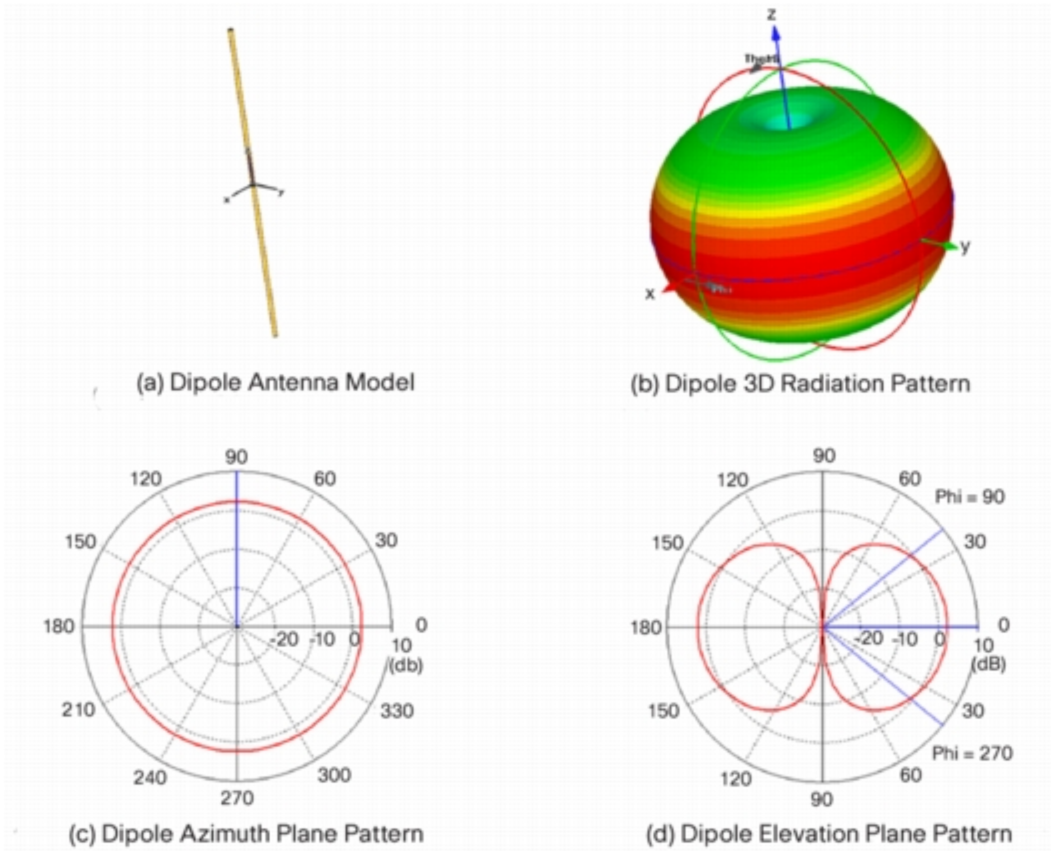

Figure 2 – The radiation pattern of an ideal dipole. Panels (a) and (b) show the dipole and its full 3D pattern. Panel (c) shows a horizontal slice, through the X-Y plane. Panel (d) shows a vertical slice.

The **directivity** or **forward gain** of an antenna describes how focused the emitted / received light is. An isotropic antenna – a theoretical ideal which radiates perfectly evenly in all directions – is often used as a reference point. An ideal dipole can be calculated as ≈ 2.15dBi forward gain, 1.64x more sensitive to the light coming in perpendicular its the main axis.

More complex feeds often show sidelobes and other features in their radiation patterns: increased sensitivity at particular angles, well away from the primary beam.

2.2 *Reflector Dishes*

Observing at radio wavelengths brings with it a host of challenges. In particular, radio observatories typically have enormous **point spread functions** – or, in radio terminology, very broad **beams**. This is driven by the simple diffraction limit,

$$\Delta\theta \approx 1.22\lambda/d$$

These beams should follow something akin to an Airy pattern, though they are frequently approximated as Gaussian in profile. Often, it's easier to work in terms of the Half Power Beamwidth (HPBW), also known as the Full Width Half Max (FWHM),

$$\mathrm{HPBW} \equiv \mathrm{FWHM} \approx 1.028\, \lambda/\mathrm{d}$$

In the Gaussian case,

$$\mathrm{FWHM} \approx 2.355\sigma$$

For a 2.5m reflector, like the one on the top of McLennan Physical Laboratories (see S3, below), observing near the HI 21cm line, that corresponds to a roughly 5° FWHM beam. Anything smaller than that – for example, roughly every astronomical source – will be unresolved, and with its signal correspondingly diluted.

To limit this sensitivity degradation, people have built increasingly large mirrors for their radio telescopes. The largest is now China's FAST telescope, with a 500m-wide aperture, followed by the US Arecibo dish at 300m. Neither of these instruments, however, can be moved: they are built into natural valleys in the landscape. The largest steerable dishes are currently the US's Green Bank 100m dish and Germany's 100m Effelsberg telescope.

The point of these reflectors is simply to increase the **forward gain** of the feed, boosting its sensitivity by focusing in on a smaller region of sky. Indeed, most radio telescopes still operate as **single-pixel** receivers: the mirror focuses onto a single detector, so imaging or more complex mapping of the sky takes place over many sequential observations of different patches of sky. Some facilities have started fitting multi-beam receivers, but for science cases targeting individual objects, these additional pixels looking at blank sky serve little purpose.

Canada's largest single-dish instrument[1] is the Algonquin Radio Observatory's 46m antenna, built just over 50 years ago and shown in **Figure 3**. The 1000-ton dish is mounted on a 2-story concrete foundation, and serves to concentrate radio light just below the small **receiver cab** at the focus. The huge collecting area of ARO yields a vast increase in the forward gain: roughly +60dBi for 10cm radiation, a million times more sensitive than an isotropic feed would be!

[1] Multi-dish arrays can be considerably larger and much more complex, but that's a topic for another lab.

Figure 3 - The ARO 46m dish, a classical parabolic radio antenna.

Illumination Efficiency – Of course, we can only take advantage of the large dish if our feed actually illuminates it! If we were to build a feed with high directivity and place it at the focus, it may only efficiently receive light from a small central portion of the dish. We want to match the feed's radiation pattern to the size and shape of the primary mirror, otherwise we won't be making much use of the big metal reflector that cost so much to build.

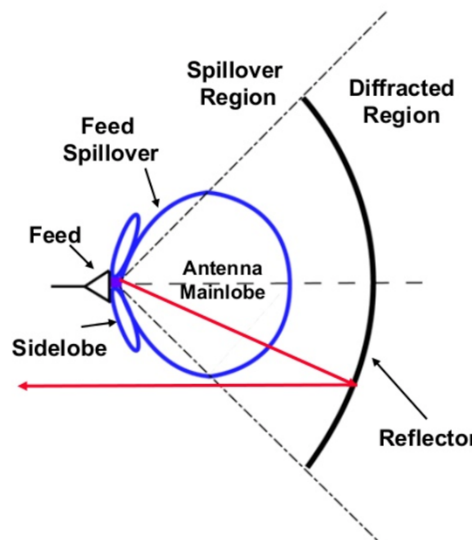

Figure 4 – Coupling a feed to a reflector, and some resulting inefficiencies.

Spillover – It's also possible to go too far the other way, to build a feed with too isotropic a radiation pattern. In this case, the dish is well illuminated, but most of the sensitivity of the feed is missing it entirely. In addition to the decreased efficiency of the reflector (and hence a decrease in forward gain), this means that most of the beam will be looking at the ground, which is probably glowing thermally at 300K – quite bright for a radio astronomer, and it all adds to the noise!

Combined, these two effects typically mean a radio telescope is only 50-70% as efficient as a naïve calculation might suggest.

2.3 *Antennas & Telescopes*

An **antenna** (plural **antennas**; only insects have *antennae*) refers to the full optical system, reflector plus feed. Since the whole operates as a combined unit to collect and capture the light, it is generally described as such, with spec sheets and instrument measurements giving the **antenna gain** and **radiation pattern**: the equivalent of an OIR telescope's Point Spread Function.

A **telescope** may contain one or more antennas, processing light as a coherent unit. In many cases, the telescope, antenna, and dish essentially refer to the same object: good luck avoiding confusion here!

3. Observations

In this lab you will be using both archival and self-acquired observations.

3.1 Archival Data from ARO

To begin, you will be working with archival data taken on the Algonquin 46m dish, using a feed and receiver chain sensitive to 400-800MHz radiation.

You'll have to look up the facility and find its position, in order to translate pointing information into celestial coordinates. You'll want to make use of the excellent `astropy` package, particularly its `time` and `coordinate` modules.

In lab #2 you worked with the full coherent measurement, raw samples of the electric field, and made your own spectra and estimates of antenna temperature. Here, the data has been pre-processed to make it slightly more manageable.

3.1.1 Cassiopeia A

This data set contains a calibration **scan** across Cassiopeia A, one of the brightest radio sources in the sky – you'll want to look it up later. The receiver was left recording while the dish was **slewed** across the source, stepping and pausing at various points along the way. A waterfall plot of the full data set is shown in **Figure 1**.

The response of the receiver will allow you to measure the **primary beam** of the telescope, representing its sensitivity to different parts of the sky. Because most of the data here is hugely redundant, you're only being given a small portion of it, a few of moments at each pointing.

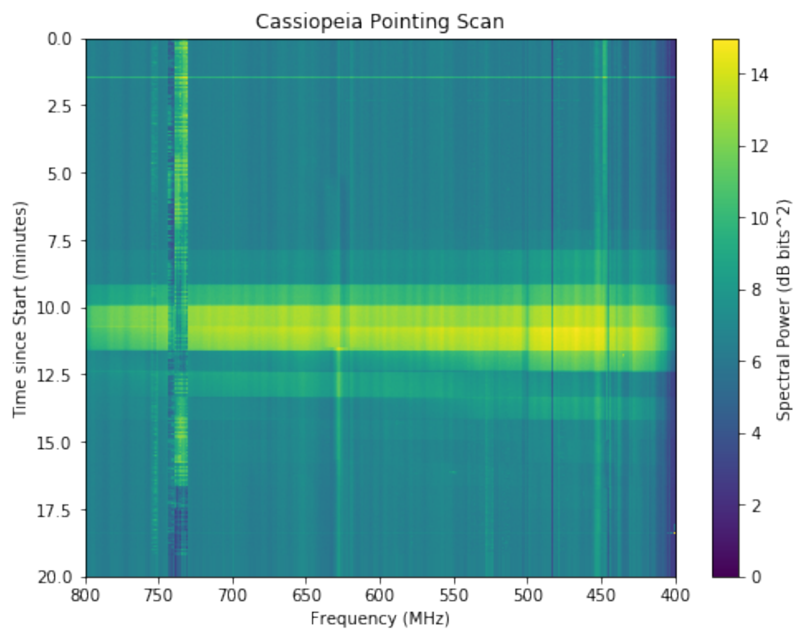

Figure 5 – Waterfall plot of the entire Cassiopeia Calibration scan. The pointing was discretely stepped across the source to sample different portions of the beam. You will only deal with a small subset of the data here.

You should download the data set from the course website, or directly HERE. The data is stored in a binary file with a minimalist structure:

- Start of file
  - 8 Bytes → INT64, nf = number of frequency channels
  - 8 Bytes → INT64, nt = number of time slices
  - 8 Bytes → INT64, ns = size of each time slice (always 1027 samples)
  - 8 Bytes x nf → FLOAT64, lf = list of frequencies
  - ns Time Slices
    - 8B → FLOAT64, timestamp, MJD
    - 8B → FLOAT64, elevation (aka altitude) pointing of the telescope
    - 8B → FLOAT64, azimuth pointing of the telescope
    - 8B x nf → FLOAT64, Stokes I samples for each frequency, in order
- End of File

Note that INT64 is a standard 8B integer; FLOAT64 is a "double precision" 8B floating point.

You'll want to write a function or two to help you read in this file. Use numpy's `fromfile`, `astype`, and `reshape` functions to help! *Hint: This file is small enough that you can probably load the whole thing into memory at once. Load it up as one big array, then index and reshape.*

3.1.2 B0329+54

This data set consists of a long **tracking observation** of the brightest pulsar in the northern hemisphere, B0329+54. You'll want to look up the pulsar by name to find out its properties.

This data has been pre-processed like the Cassiopeia A data: It has been Fourier Transformed to produce 1024 spectral bins (each ≈390kHz wide, all together covering the 400MHz bandwidth), and the variance of each spectral bin has been calculated once every 256 samples ≈ 0.655ms, for each of two orthogonal linear polarizations being measured. Finally, these are summed to yield measurements of the Stokes Intensity.

> *Even with this pre-processing, the data set is large, roughly 5MB/s. You will very likely have to load and process smaller portions of it at a time: remember, loops and functions are your friends! Use them, or you'll likely bog down in processing and trying to plot huge arrays.*

Because of the size of the data set, you're only going to work with the first roughly 90 seconds, which still comes to 500MB even after being reduced to Stokes-I.

> Be glad this is a bright source that doesn't take ages to build up signal on! Pulsar astronomers regularly work with several hours of the data, and often go back to the raw voltages, with data sets stretching easily into the petabytes (millions of gigabytes).

The format is slightly more complex, to allow for the additional detail you'll be needing.

- ➢ Start of file
  - 48 Bytes → Header
    - INT32, sl = slice length
    - INT32, hl = header length
    - INT32, ns = number of samples
    - INT32, dt = data type *(safe to ignore)*
    - FLOAT64, cr = cadence of raw spectra *(safe to ignore)*
    - INT32, nf = number of frequencies
    - INT32, ne = number of elements, 1 *(safe to ignore)*
    - INT32, ss = raw spectra summed per sample
    - UINT32, hidx = handshake IDX (index of start)
    - FLOAT64, hutc = handshake UTC (*Unixtime* of start, in UTC)
  - 4 Bytes x nf x 2 →FLOAT32, list of start and stop frequencies for each of the spectral bins
  - 2^{17} Time Slices of size sl
    - 4B → INT32, index of the time slice
    - 4B → FLOAT32, elevation (aka altitude) pointing of the telescope
    - 4B → FLOAT32, azimuth pointing of the telescope
    - 4B x nf → Stokes I samples for each frequency, in order
- ➢ End of File

A few hints on reading in and working with this file:

- Do it piecemeal! Don't try to read the whole thing at once. It may work, but it'll be painful to manipulate or plot.

- You can use the `stuct` package to `unpack` the header, using the format string `'=iiiidiiiId'`
- Check if the header makes sense before going any farther. Then try reading one slice, using a custom function that reads back slices. Then try reading two. Then add a big loop to read a few thousand.
- Always do sanity checks on your data! Does the frequency make sense? Does the time line up?

3.2 Small Radio Telescope

This portion of the lab will require you to take new data using the dish on top of the McLennan Physical Laboratories. Schedule time along with your group to head up and perform a scan over a bright astronomical source. You can use the Sun (brightest), Cygnus A (fainter, but up in the evenings), or Cassiopeia A (fainter still, but you have already worked the numbers).

> *Note that the dish on the roof has had a rough life. It spent 15 years idle because of the overwhelming RFI environment of downtown Toronto, before we finally got it running in early 2018. After a short but productive spring, a windstorm damaged the mount. We're working on getting it properly refurbished, but you may have to defer to the telescope operator in deciding which source to go after.*

As in the ARO data, a full set of raw voltages is simply too large to save, so we'll have to employ some in-situ signal processing. To record data from the MP telescope, you will need to use a special piece of software to gather data, convert it into a spectrum, and calculate the power across longer timescales.

> *You have the freedom to control these parameters, choosing the number of frequency bands and the power integration timescale, but I suggest leaving these at 64 frequency bins and 50ms integrations.*

DETAILED DESCRIPTION OF FILE FORMAT PENDING. THIS MANUAL WILL BE UPDATED WITH DETAILS IF & WHEN WE GET THINGS WORKING ADEQUATELY!

4. Key Lab Activities

This section outlines the key measurements you should be making throughout the lab, and which should be discussed in your lab report. Remember that the goal is to study and understand the two telescopes (ARO and MP), and to investigate the properties of time-domain radio astronomy through the window of pulsar B0329+54.

4.1 ARO Properties

4.1.1 Beam Width

Look up ARO, find its basic physical properties. Size, focal ratio, latitude / longitude. Look up Cassiopeia A, find its basic radio properties. Is it a point source, or extended? Find a reference for its flux over the 400-800MHz band. *(Hint: I suggest Baars et al, 1977. "The Absolute Spectrum of Cas A," A&A vol 61, though you should beware that it's dimmed roughly 50% in the last 40 years.)*

Inspect the data. Examine the header, see when it was taken, what band it covers, and other critical features of the dataset. Make sure you can read things in and everything makes physical sense. Where was the telescope pointing during the observation? The 46m is an alt-az dish; calculate the telescope coordinates (altitude/elevation and azimuth) and draw a small diagram showing the dish orientation.

Choose a representative frequency sub-band, plot the power as a function of source offset. Using a least-squares fitting routine such as `scipy.optimize.curvefit`, fit a Gaussian and calculate the beamwidth as the Full-Width Half-Max (FWHM). Repeat for all other sub-bands and plot the beamwidth as a function of frequency.

Calculate the diffraction-limited beamwidth for a 46m aperture and overplot. (Don't forget error bars on your measurements!) Do they match? Suggest possible reasons if not.

Plot the beamwidth and aperture efficiency as a function of frequency. Describe the overall shape and any standout features, and explain why they may look as they do.

4.1.2 Receiver Temperature

Choose a representative sub-band. Calculate Cassiopeia A's flux in that band. Given the beam you measured, estimate the effective aperture of the ARO antenna (you can make the assumption that the beam is symmetric to simplify life, though in reality it almost certainly isn't). Recalling the definition of Janskys,

$$1\ \mathrm{Jy} = 10^{-26}\ W\ m^{-2}\ Hz^{-1},$$

and the relation between power and temperature in the Rayleigh-Jeans region of the spectrum,

$$P = kT\Delta\nu,$$

what temperature is Cass A to our telescope, in Kelvin?

Assume the sky around Cass A is 10K. What is the noise temperature of our receiver, in Kelvin? What is the gain of the receiver, in (ADC bits)^2 / K? *(Hint: this is exactly like your calculations in lab #2.)* Repeat the calculation for all frequency sub-bands and plot the noise and gain spectra.

As above, describe the overall shape and identify any peculiar features you see in the data.

4.2 ARO B0329+54

4.2.1 Background

Look up the pulsar B0329+54. The Australia Telescope National Facility (ATNF) provides an excellent and comprehensive online database of pulsar properties that you can search, and there is a rich literature showing many key properties of the pulsar.

Inspect the data. Examine the header, see when it was taken, what band it covers, and other critical features of the dataset. Make sure you can read things in and everything makes physical sense. Where was the telescope pointing during the observation?

Make a **waterfall** plot, showing the Stokes-I as a function of frequency and time, covering at least a few seconds. Can you identify the pulsar? You may need to integrate in time and/or frequency to give the right combination of resolution and sensitivity. *(Hint: the different frequency sub-bands have different responses and noise properties, you can get things to visually pop out much better if you renormalize each sub-band, dividing by its median value.)*

4.2.2 Folding

Often, pulsars are too faint to resolve individual pulses, and we have to stack many of them to make the signal stand out. This is accomplished by **folding** the data stream, on the period of the pulsar. In brief, choose a zero time, then for each subsequent time sample, calculate where in the pulse period it should be. For example, if I have a pulsar with a 0.5s period, samples taken at 0.23s and then at 0.73s (and 1.23s, 1.73s…) are all at the 46% mark within the pulse period. A folded data set shows the signal as a function of pulse phase instead of time.

Fold the data on B0329+54's period and plot a waterfall of the result. You should see a nice bright pulse sitting in it. You might be able to clean it up by doing some basic RFI cleaning: look at the total power in the data as a function of time (i.e. sum across the frequency axis), and remove samples where the power is anomalously high.

Using the calibration found above, what is the typical flux of B0329+54 pulses in this band, measured in Jy?

4.2.3 Dispersion Measure

Consider a single pulse. When it was generated at the source, it was a broadband impulse, with power emitted simultaneously across a broad frequency band. That happened several thousand years ago, depending on the distance of pulsar, and in the intervening millennia, the pulse has been flying toward us through the rarefied interstellar medium.

The ISM is an awfully good vacuum, better than anything we can make on earth, but it's still not quite empty, containing a low density of loose electrons, stripped from interstellar hydrogen. This rarefied plasma very gently disperses light traveling through it, slowing the longer wavelengths by a fraction relative to their higher energy counterparts. By the time the pulse arrives on earth, this becomes visible is a **dispersion**: higher frequencies arrive first, with a gradual sweep in arrival until the lower frequencies trickle in.

Look up the definition of a **dispersion measure** and find the value for this pulsar. Calculate the anticipated time delay across the observing band, and compare to your data. (How would this manifest in non-pulsed signals? Would we ever notice it?) Calculate the dispersive delay for each of the 1024 frequency sub-bands, and **de-disperse** your data, shifting the time in each sub-band to remove the effect. Generate a new waterfall with nice straight pulses, which look more like they would when they were emitted.

4.2.4 Average Pulse Profile

ARO is a big dish, and B0239+54 is a bright pulsar, so you can probably make out individual pulses in this data set. You may notice that they vary a fair bit, with no two quite the same. Folding does a nice job averaging out these variations, but it's common to make a single pulse template, by de-dispersing and summing across your frequency band. Produce a plot of full-band Stokes-I power vs pulse phase, representing your best estimate at what a typical pulsation of power looks like.

What happens if you try to calculate a full-band profile before de-dispersing?

4.2.5 Pulse Variability

How do the pulse amplitudes vary in time? Make a histogram showing the number as a function of flux. (Use the calibration as before to label this plot with physical units, Jy.)

Now that we have a pulse profile, we can look at the pulse variability in a little more detail. Take your raw data set, de-disperse, and sum it across frequency, to yield a single timestream showing the full-band Stokes-I power (the "un-folded" version of what you did above).

Take your template average pulse profile, and use it to maximize your sensitivity to each pulse by convolving the timestream with it. (Why does this help pick out the pulses?) A template convolution is a simple version of a powerful signal processing technique called a **matched filter**, which allows you to quickly search through data in a statistically optimal way, picking out sections that look like some signal of interest.

Using your filtered data, does the pulse move around, or does it always appear in the same place within its period? (Did you get the period right?)

4.3 (Optional) SRT Properties

> This dish is currently being repaired after a windstorm damaged its mount and panels. We had hoped to give you a chance to play with it before lab #6, but things are as of today still broken. Stay tuned for more information on this section, but nothing strenuously time consuming will be required of you here.

4.3.1 Beam Width

The dish on the roof is from the ***Small Radio Telescope (SRT)*** kit, you can look up its properties at https://www.haystack.mit.edu/edu/undergrad/srt/oldsrt.html. Record its size, focal ratio, latitude / longitude. Look up your source, find its basic radio properties. Is it a point source, or extended? Find a reference for its flux at 1420MHz.

As at ARO, choose a representative frequency sub-band, plot the power as a function of source offset. Using a least-squares fitting routine, fit a Gaussian and calculate the beamwidth as the Full-Width Half-Max (FWHM). Repeat for all other sub-bands and plot the beamwidth as a function of frequency. Calculate the diffraction-limited beamwidth for the SRT and overplot. Do they match? Explain any discrepancies, and explain why it looks different from the ARO result.

4.3.2 Receiver Temperature

As at ARO, choose a representative sub-band. Calculate your source's flux in that band. Given the beam you measured, estimate the effective aperture of the MP antenna. Calculate the temperature of your source in our telescope, in Kelvin. Calculate the noise temperature of the receiver, and its gain.