# Spectrometers and Signal Processing Basics

2021 GBO/AO Single Dish Workshop

Ryan Lynch (GBO)

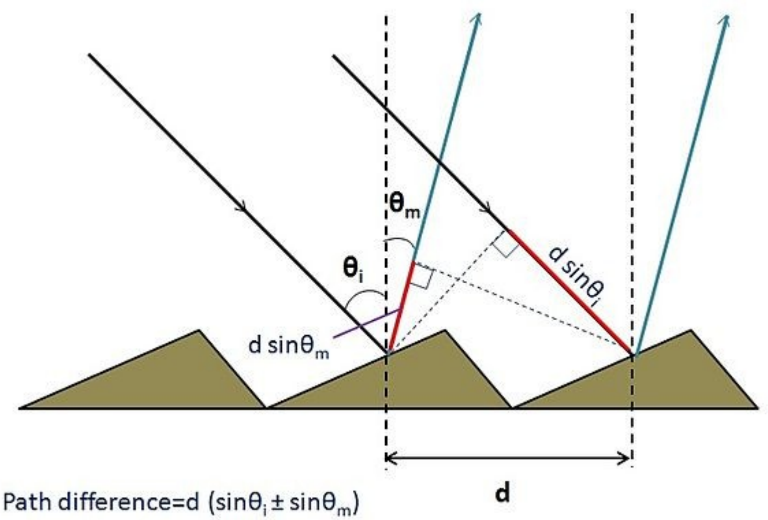
# What is a Spectrometer and How Does it Work?

- A spectrometer measures intensity of electromagnetic radiation at different frequencies / wavelengths
  - In practical applications spectrometers have a finite frequency / wavelength *resolution* and a finite range of frequencies / wavelengths over which they operate
- Most astronomers are introduced to spectrometers at optical wavelengths
  - Use gratings (or prisms) to disperse light, i.e. physically separate different wavelengths
  - Measure intensity of dispersed light using a CCD
    - Position on CCD maps to wavelength
- This is not how radio-frequency spectrometers work
- Why not?

# Spectrometers in Other Contexts

- Physical properties of the spectrometer often scale with wavelength (or it’s inverse)
- Radio-wavelengths span ~10 m – 1 mm
- Consider diffraction angle

$$\theta_m = \sin^{-1}\left(\frac{m\lambda}{d}\right), m = 1, 2, \ldots$$

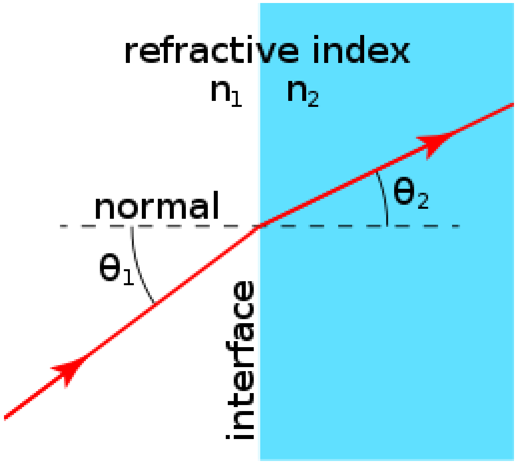



  - A radio-frequency diffraction grating would have to be impractically large

# Spectrometers in Other Contexts

- Physical properties of the spectrometer often scale with wavelength (or it's inverse)
- Or consider a refractive prism

$$\frac{n_1}{n_2} = \frac{\sin(\theta_2)}{\sin(\theta_1)}$$



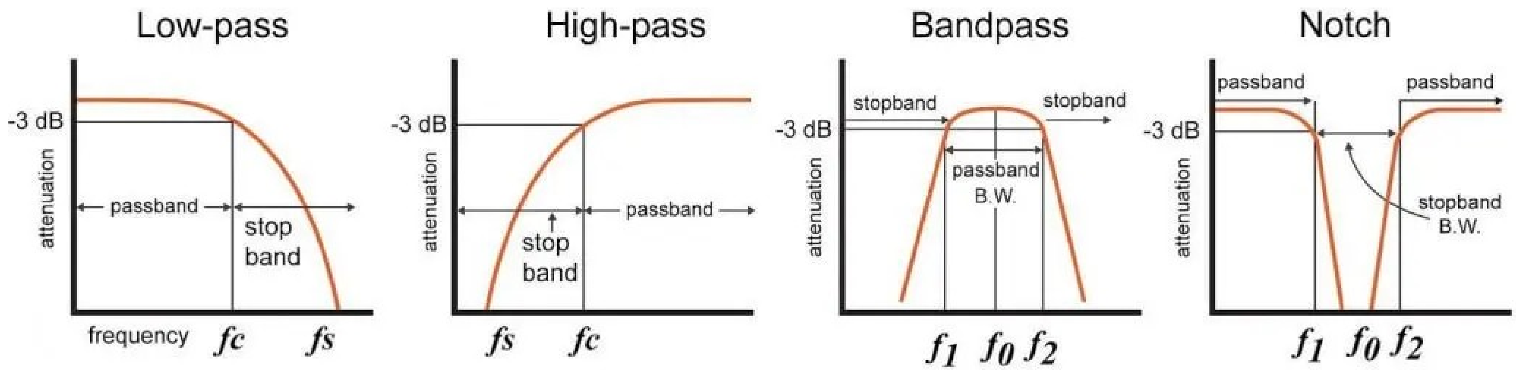

- Index of refraction usually approaches 1 as wavelength increases
- Fractional change in wavelength Δλ/λ is small
- Angle of refraction will be very small and not much separation of different wavelengths, leading to almost no dispersion

# Spectrometers in Other Contexts

- Properties of a radio-CCD are also impractical
- Pixel size has to be > $\lambda$
- Band-gap energy of semiconductor needs to be < radio-photon energy
  - At radio frequencies photons have energies ~ μeV – meV
  - Most semiconductors have band-gap energies ~ 1 eV
  - Thermal noise will swamp signal unless detectors are cooled to extremely low temperatures

# Analog Radio Spectrometer

- Quantum devices like CCDs don’t operate well at radio wavelengths, but analog electronic circuits do
- It is fairly easy to create electronic filters that attenuate power above/below/within certain radio frequencies




- We can envision sending copies of a signal through multiple bandpass filters, each with different frequency cutoffs, and then detecting power that passes through each filter
- A bank of filters → a *filterbank*

# Analog Radio Spectrometer

- Each signal path is known as a *channel*
- Each channel has some some narrow *channel bandwidth* over which it is sensitive
  - Analogous to resolution of an optical spectrometer
- In practical applications, channels are adjacent, with small gaps between them
- Difference between lowest and highest frequency channels defines *total bandwidth*

# Analog Radio Spectrometer

- This setup is intuitively simple and illustrates basic concepts, but is not very practical
- What if you need a narrower channel bandwidth?
  - Need to duplicate the entire filterbank with a different set of filters
- What if you need a large total bandwidth *and* narrow channel bandwidth?
  - Need lots and lots of filters
- Can we do better?
- Yes!  We can create a spectrometer using math…
- But first, what are we actually measuring?

# Radio Telescopes Sample Electric Fields

- EM radiation is a time-varying electromagnetic field
- Radiation incident on a radio receiver causes a change in electric potential, i.e. a change in *voltage*
- For a *monochromatic* wave

$$\vec{V}(x,t) = \underbrace{V_0}_{\text{Amplitude}} \sin\Big(2\pi \underbrace{f}_{\text{Frequency}} t + \frac{2\pi}{\underbrace{\lambda}_{\text{Wavelength}}} \overbrace{\vec{x}}^{\text{Direction}} + \overbrace{\phi_0}^{\text{Initial Phase}}\Big) \qquad c = \lambda \cdot f$$

- Real-world signals have non-zero amplitude at many frequencies (i.e. polychromatic)

# Radio Telescopes Sample Electric Fields

- EM radiation is a time-varying electromagnetic field
- Radiation incident on a radio receiver causes a change in electric potential, i.e. a change in *voltage*
- For a *monochromatic* wave

$$\vec{V}(x,t) = \underbrace{V_0}_{\text{Amplitude}} \sin\left(2\pi \underbrace{f}_{\text{Frequency}} t + \frac{2\pi}{\underbrace{\lambda}_{\text{Wavelength}}} \overbrace{\vec{x}}^{\text{Direction}} + \overbrace{\phi_0}^{\text{Initial Phase}}\right) \qquad c = \lambda \cdot f$$

- Real-world signals have non-zero amplitude at many frequencies (i.e. polychromatic)

# Radio Telescopes Sample Electric Fields

- EM radiation is a time-varying electromagnetic field
- Radiation incident on a radio receiver causes a change in electric potential, i.e. a change in *voltage*
- For a *monochromatic* wave

$$\vec{V}(x,t) = \underbrace{V_0}_{\text{Amplitude}} \sin(2\pi \underbrace{f}_{\text{Frequency}} t + \overbrace{\phi_0}^{\text{Initial Phase}})$$

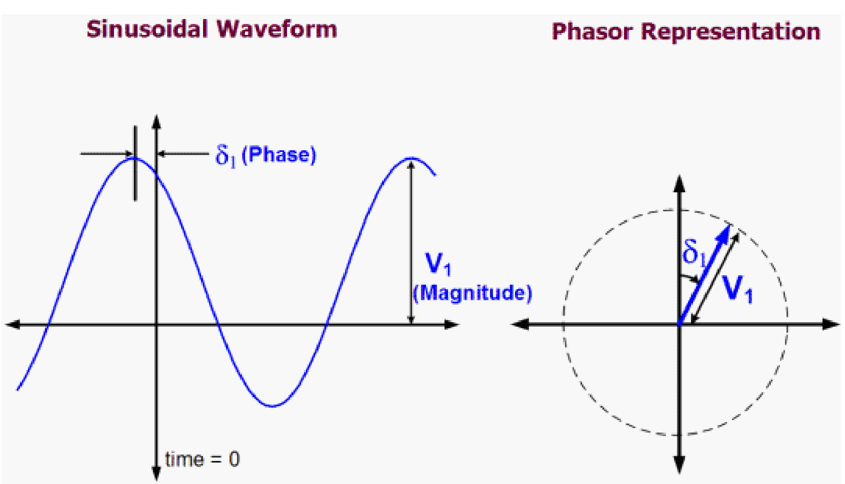
$$c = \lambda \cdot f$$

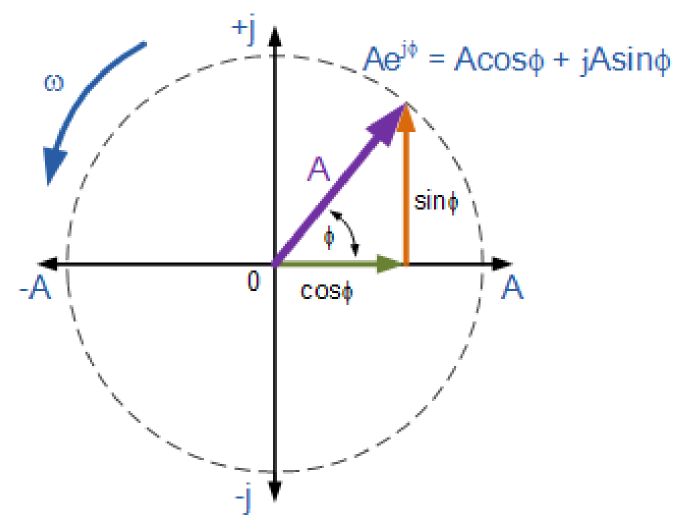
- Real-world signals have non-zero amplitude at many frequencies (i.e. polychromatic)

# Complex Voltage and Power

- Because incoming radiation is described by both an electric field amplitude and phase, it is convenient to represent it as a **phasor**
- This lends itself to using complex numbers to describe the voltage (recall Euler's formula)

$$A e^{i\theta x} = A[\cos(\theta x) + i \sin(\theta x)]$$



# Complex Voltage and Power

- The cosine and sine terms are often referred to as $I(t)$ and $Q(t)$ (i.e. $I/Q$ values)
  - $I$ corresponds to the real part of the complex voltage, and $Q$ to the imaginary part
  - Don’t confuse these with the I and Q of Stokes parameters!

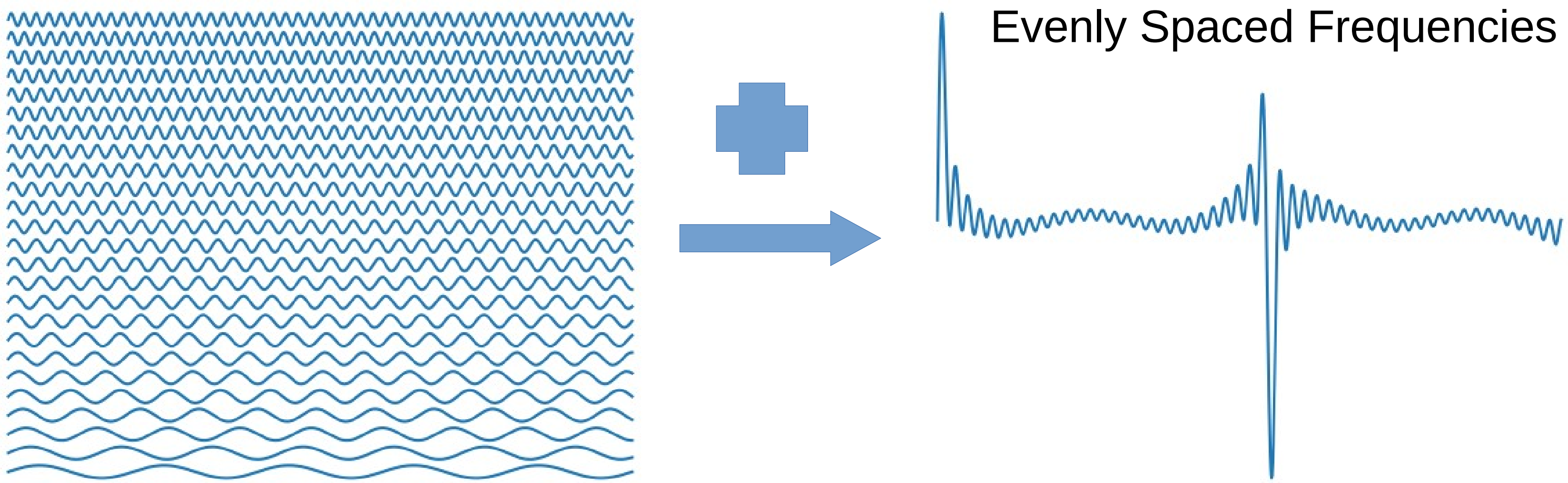
# Complex Voltage and Power

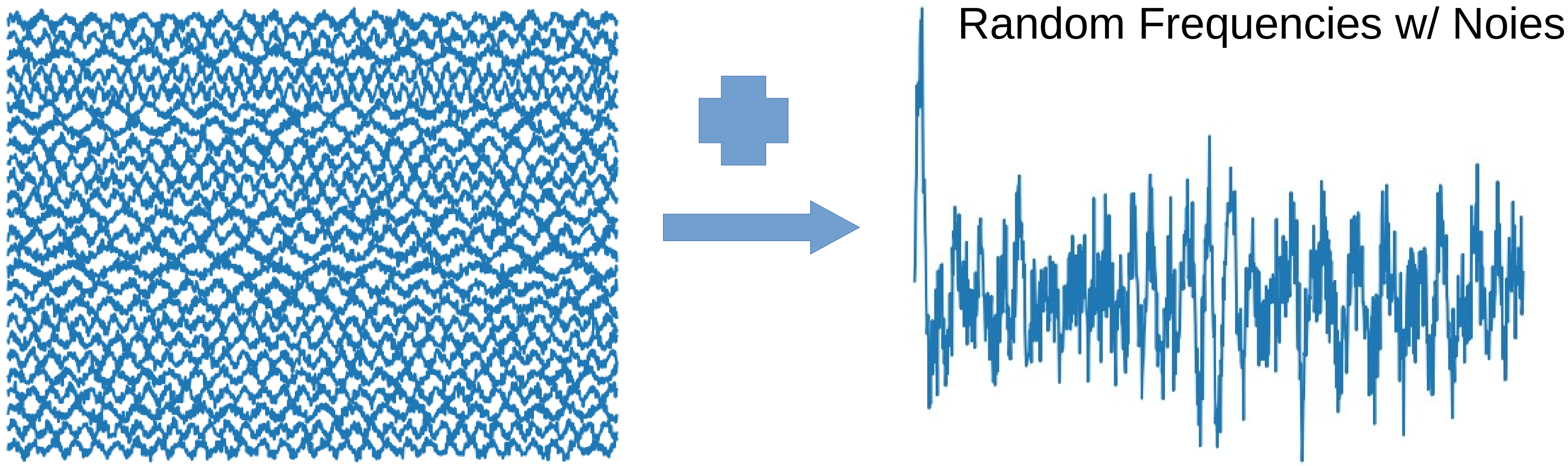
- This allows us to represent the signal with a real and imaginary part
  - Retains full amplitude and phase information so can be used for **coherent** processing
- In the final analysis we are usually interested in the **power** (which has non-zero mean), rather than the amplitude

$$P = \left| A e^{i \theta x} \right|^2$$

- This step is usually referred to as **detection**
  - If we sample two polarization states, we can form Stokes parameters or other polarization products prior to detection
  - Note that we lose phase information at this stage!

# More Realistic Signals

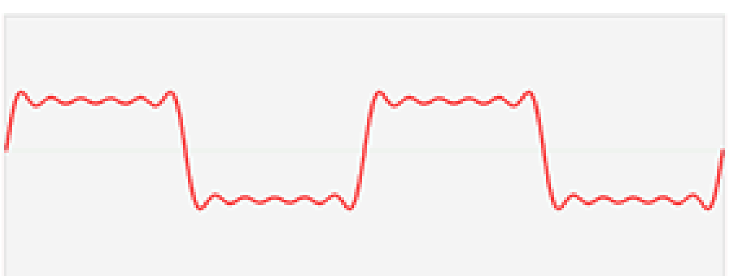


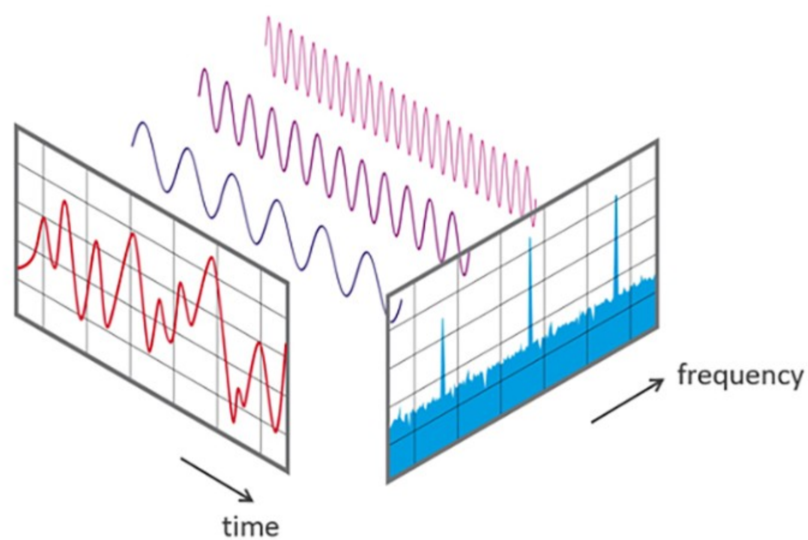



# Fourier Transform Spectrometer

- We measure $V(t)$ but want to measure power vs frequency
- We can harness the Fourier Transform to do this mathematically. Recall that for a continuous signal

$$X(f) = \int_{-\infty}^{+\infty} x(t) e^{-2\pi i f t} dt$$

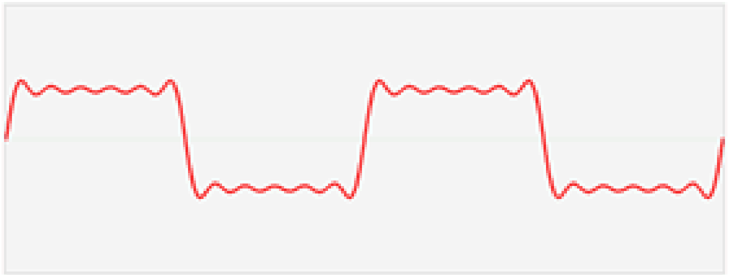


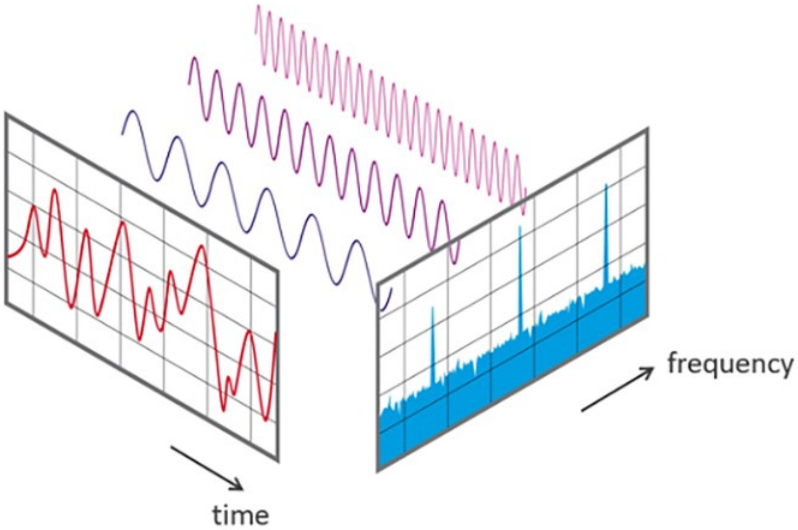


# Fourier Transform Spectrometer

- We measure $V(t)$ but want to measure power vs frequency
- We can harness the Fourier Transform to do this mathematically. For a discrete signal

$$X_k = \sum_{0}^{N-1} x_n e^{-2\pi i k n/N}$$





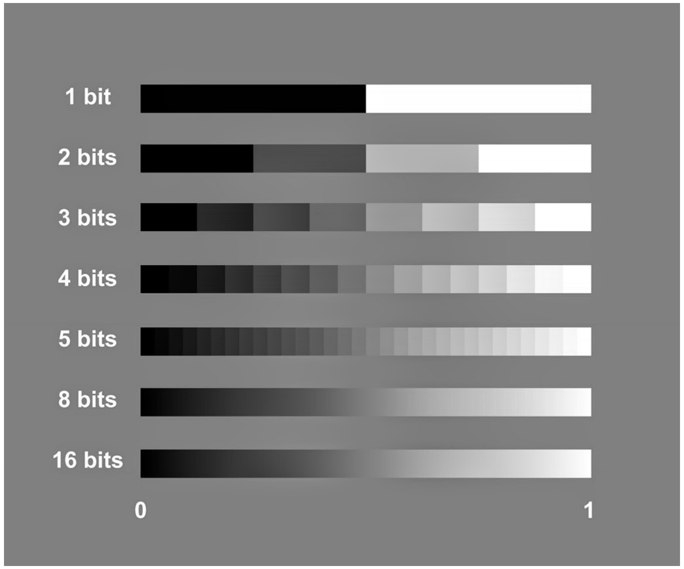
# Fourier Transform Spectrometer

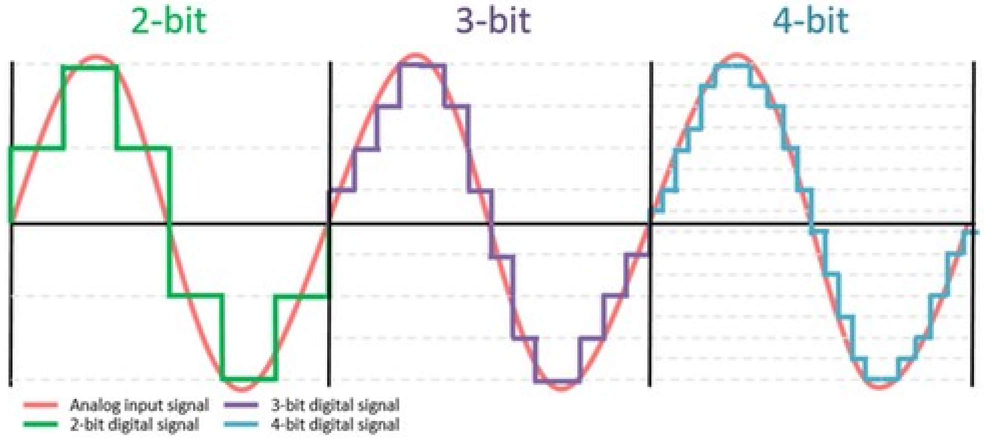
- FT spectrometers are highly flexible and can be implemented on computers *if* we can fully **digitally** sample the voltage time series
- To do this we have to measure amplitude *and* phase
- We also have to sample quickly enough to detect rapidly varying (i.e. high-frequency) signals

# Continuous vs Discrete Signals

- Incoming radiation is a continuous change in electric field over a continuous range of frequencies
- Digital systems work on discrete values that can be represented with some finite number of bits
- 1 bit = $2^1$ values (0,1)
- 2 bit = $2^2$ values (0 – 3)
- 8 bit = $2^8$ = 256 values (0 – 255)

- The number of bits used to sample the signal defines the **dynamic range**
  - Smaller bit depth / resolution provides less granularity (1 bit = high or low)
  - Higher bit depth captures both weak and strong inputs
- This introduces some error, as perfect reconstruction is not possible with a finite number of bits
  - Bit depth chosen to keep quantization errors at or below an acceptable level



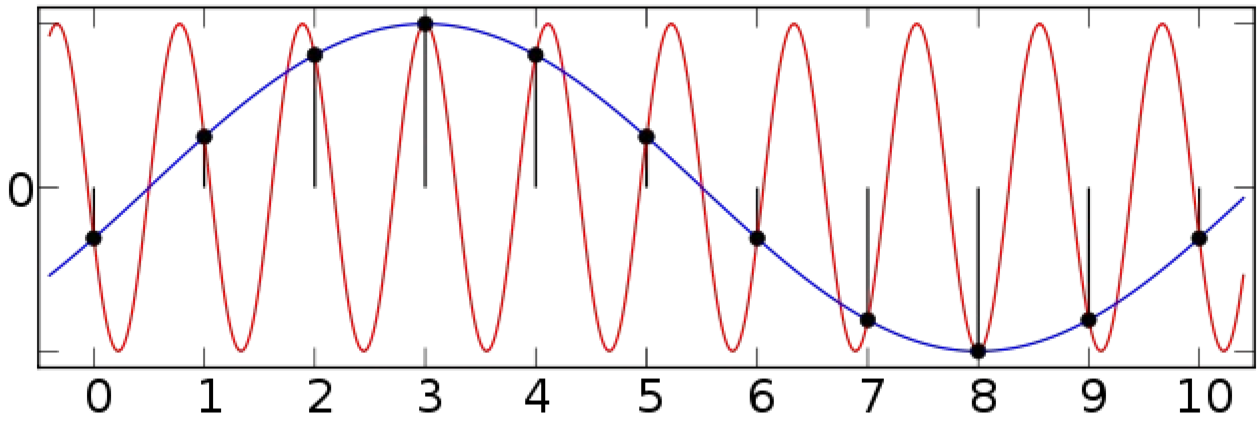



# Analog to Digital Converters

- An **analog-to-digital converter** (ADC) is a device for converting continuous signal to discrete, digital signal
- ADCs are characterized (in part) by the number of bits they use and the maximum *sampling rate*
- The sampling rate determines the *bandwidth* of the sampler

# Nyquist-Shannon Sampling Theorem

- To perfectly reconstruct a time varying signal, we must sample at a critical rate, $f_N$, that is **twice the highest frequency** contained in the signal
  - A signal at a frequency $f > f_N$ will be **aliased** into our sampling band at a lower apparent frequency
- $f_N$ is known as the **Nyquist frequency**



# Nyquist-Shannon Sampling Theorem

- This is not just a time/frequency phenomenon
- Spatial variations can be decomposed into spatial frequencies
  - Sharp features contain higher frequency components
- Nyquist sampling in spatial domain is important in mapping

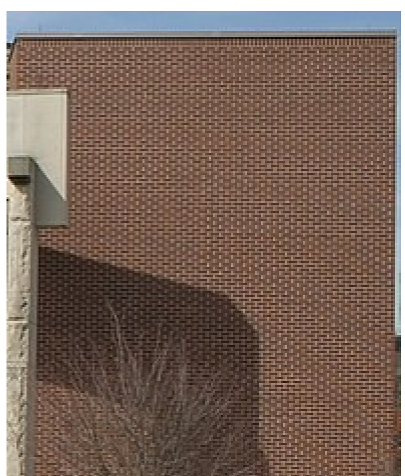
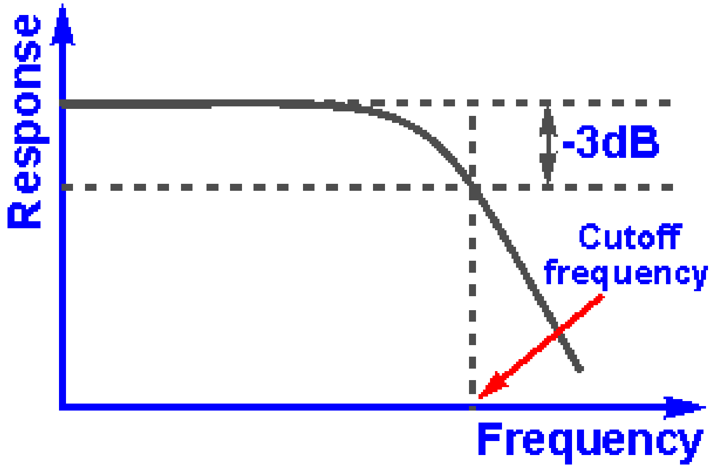


Image credit: Wikipedia

# Sampling Rate and Bandwidth

- To avoid aliasing, we must apply an analog filter to suppress power outside some desired bandwidth *B*
- We then use our ADC to sample at a frequency $f_s = 2 \times B$
  - Example: We want to sample 800 MHz bandwidth
  - Downcovert to **baseband** and apply low-pass filter
  - Sample at 1.6 Gsps
- Remember: filters are not perfectly sharp
  - Filter roll-off needs to start below $f_s/2$ to ensure aliasing is kept below an acceptable level



# Why Don’t Optical Spectrometers Work This Way?

- If a FT spectrometer is so useful, why not use them at shorter wavelengths / higher frequencies?
- Radio frequencies have to advantages for this approach
  - Signals can be *down-converted* from high to low frequency via *mixing*
  - Down-converted signals have frequencies that are low enough to sample with modern electronics
- Optical-frequency mixing is much harder, and electronics can’t sample at the native optical frequencies

# Dynamic Range for Wideband Systems

- Note that ADCs are **total power** devices
  - We have not yet sampled the power contributed at individual frequencies
- As the bandwidth goes up, so to do does the total power contributed by noise, RFI, and signal of interest
- Resolution / bit depth becomes increasingly important for wideband systems
  - Strong signals can push ADCs into non-linear regime

# Types of Spectrometers

- There are different ways of implementing these general principles
- We will talk more about two examples
  - Auto-correlation spectrometer (ACS)
  - Polyphase filterbank (PFB)

# Weiner-Kinchin Theorem

- Relates the power spectrum to the **autocorrelation** of the incoming time series

$$S(f) = \int_{-\infty}^{\infty} r_{xx}(\tau) e^{-2\pi i f \tau}\, d\tau.$$

- $r_{xx}$ is the autocorrelation, defined as

$$r_{xx} = \int_{-\infty}^{\infty} f(u) f^*(u - \tau)\, du$$

  - $\tau$ is known as the **lag**, and * denotes the complex conjugate
- In words, the power spectrum is the Fourier transform of the integral of the input signal multiplied point-wise by a time-delayed version of itself

# Autocorrelation Spectrometer

- An **autocorrelation spectrometer** is highly flexible in terms of total bandwidth and channel bandwidth
  - The sampling interval Δτ and total number of lags *N* completely determine these parameters

$$B = \frac{1}{2\Delta\tau}; \Delta f = 1.2\frac{B}{N}$$

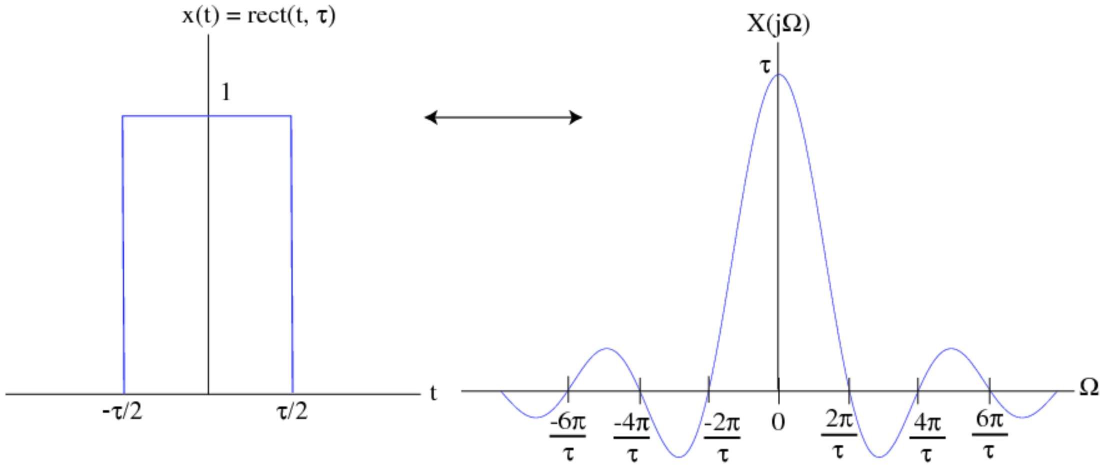
    The factor of 1.2 comes from the **windowing function**, which is simply a hard cutoff at t > Δτ*N* (i.e. *w*(*t*) = 1 for t <= ΔτN, else 0)

  - The observed power spectrum is a convolution of the true spectrum with the Fourier transform of *w*

$$\tilde{S}(f) = S(f) \circ W(f)$$

# Autocorrelation Spectrometer

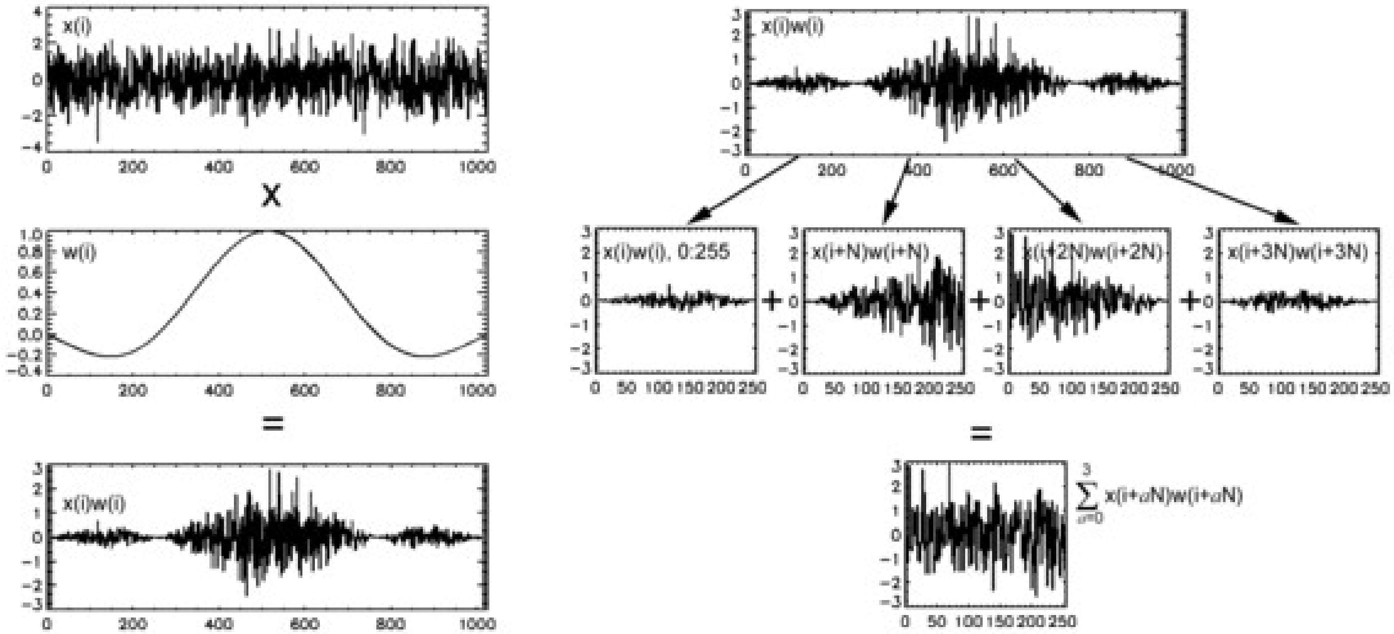
- Because the Fourier transform of a top-hat is a sinc function, the channel shape of an ACS is itself a sinc, defined by it's FWHM
  - This is where the factor of 1.2 comes from
- While an ACS is flexible and easy to implement, this frequency response is undesirable
  - Power can leak into adjacent channels
  - For very strong signals, leakage can impact significant part of band
- Can we do better?
  - Yes!



# Polyphase Filterbank

- In a direct discrete Fourier transform (DFT) we start with a rectangular windowing function (in time) and end with a sinc response (in frequency)
- We prefer to have a rectangular (i.e. flat) response in frequency across a channel
  - Use the Fourier inverse as the time-domain window, i.e. a sinc filter
- In practice, to obtain an $N$-point spectrum, use $M = N$ x $P$ points
  - $P$ is the number of phases in the **polyphase** filterbank, also referred to as the number of taps

# Polyphase Filterbank




- After multiplication by an *M*-point filter, each phase is added to produce an *N*-point input to the DFT
- The DFT can now be taken, the result squared, and then accumulated to produce a power spectrum

Image credit: Dale Gary

# Polyphase Filterbank

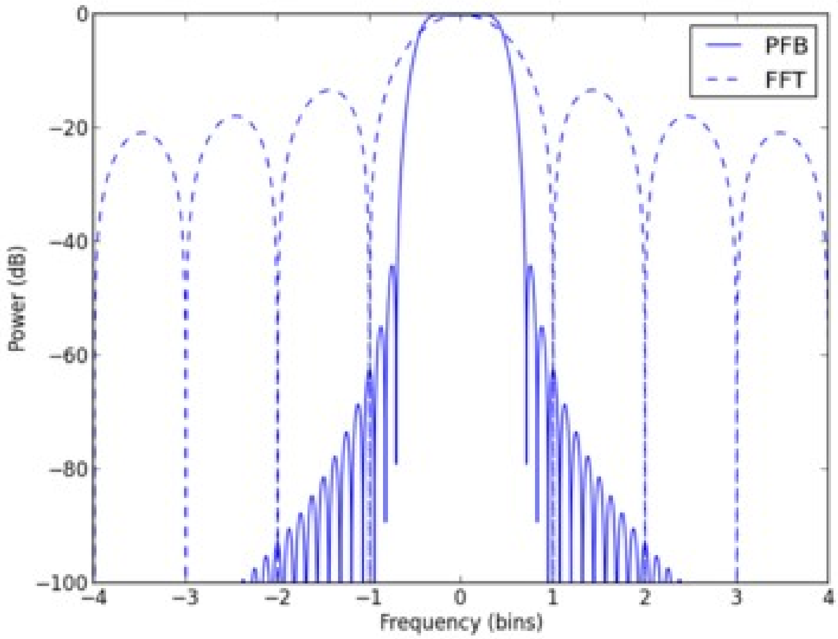



Image credit: Jayanth Chennamangalam

- Caveats
  - In pratice, the sinc window must be truncated so the frequency response is not perfectly flat
  - We typically multiply the sinc window by an finite impulse response (FIR) filter to improve frequency response
  - Using more taps also improves response

- PFB is more computationally intensive (~1.5x) than direct DFT but improved spectral response is usually worth the trade-off

# Astronomical Spectrometers

- Note that the frequency resolution we obtain is determined by the number of points in the FFT
  - The sampling theorem is also relevant here: we need $2N$ time samples for $N$ frequency channels
- This creates an inverse relationship between time and frequency resolution
- In typical spectral line observing, we are more concerned with frequency resolution than time resolution
- In pulsar observing we are usually more concerned with time resolution that frequency resolution

# Astronomical Spectrometers

- The last* step is typically to detect and accumulate power spectra for some integration time
  - The choice of integration time depends on the stability of the instrument and scientific goals
  - Typically use ~0.1 – 10 s for spectral line observing to allow efficient excising of RFI
  - Typically use 10s $\mu$s in pulsar observing to retain sensitivity to fast pulsars

*Additional signal processing often performed in pulsar observing (e.g. dedispersion, folding)

# Polarization Products

- Most receivers sample two polarization states (typically linear [X/Y] or circular [L/R])
- Everything described above must be duplicated for each polarization channel
  - 2x ADCs, 2x spectrometer engines
- The polarization products that one records depends on science goals
  - Typically sufficient to record each channel's self-products independently (e.g. $|X|^2$ and $|Y|^2$ )
- For strongly polarized sources, typically record Stokes parameters or self and cross terms

# Polarization Products

- Stokes parameters allow complete recovery of polarized signal
  - For a linear basis:

    $I = |X|^2 + |Y|^2$ (total intensity)
    $Q = |X|^2 - |Y|^2$
    $U = 2\,\mathrm{Re}(X^* Y)$
    $V = 2\,\mathrm{Im}(X^* Y)$

  - $|V|$ = circular polarization
    $|L| = \sqrt{(Q^2 + U^2)}$ = linear polarization

- We may also record the self and cross terms directly, [i.e. $|A|^2$, $|B|^2$, $\mathrm{Re}(A^* B)$, $\mathrm{Im}(A^* B)$]

# A Note on Complex Voltages

- There are some applications in which it is desirable/necessary to record pre-detection complex voltages
  - Very long baseline interferometry requires phase information for correlation
  - Offline analysis may be needed to form spectra with different resolutions for different applications
- This comes at the expense of very high data rates, requiring lots of storage

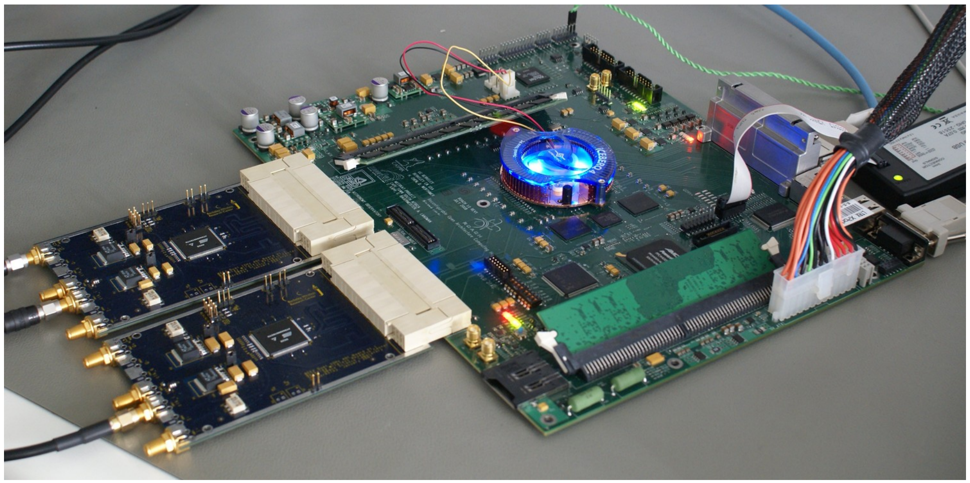
# Hardware for Modern Digital Backends

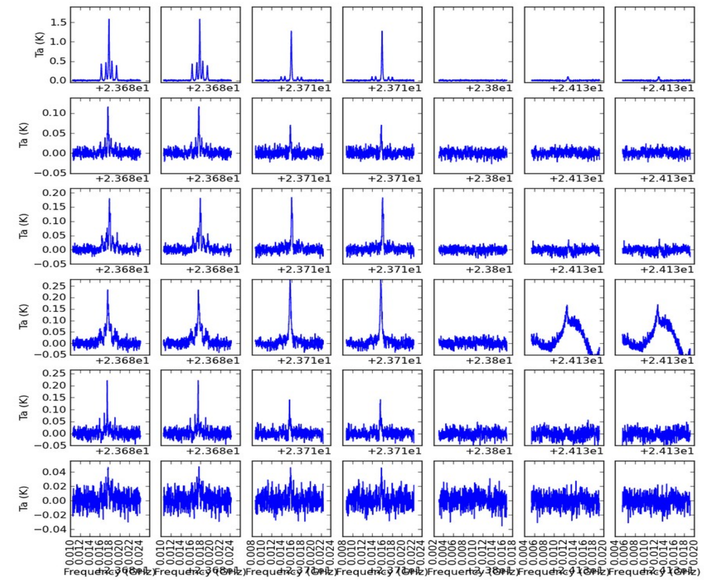
- Modern systems are typically implemented with a combination of field programmable gate arrays (FPGAs) and GPU-equipped high performance computers running specialized digital signal processing software
- GBT currently uses five primary backends
  - Digital continuum receiver
  - Mark VI VLBI baseband recorder
  - VEGAS (spectral line/pulsar observing)
  - JPL Radar Backend
  - Breakthrough Listen (baseband recording for SETI, etc.)

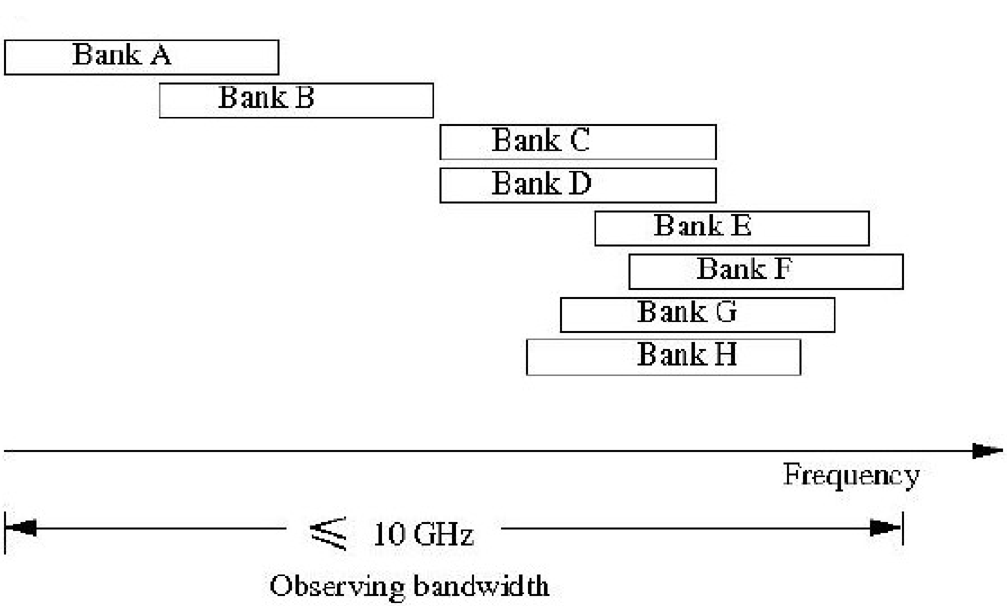
# Hardware for Digital Backends

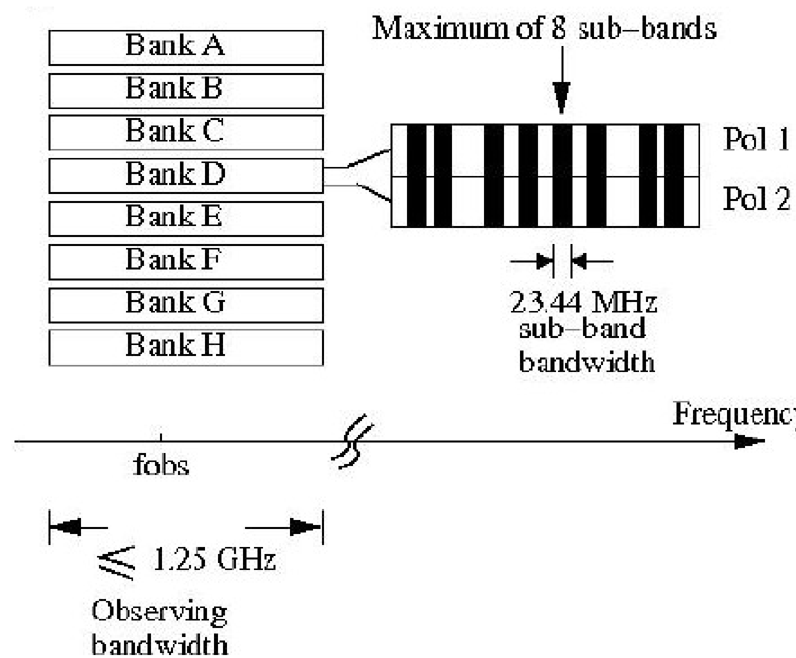
- VEGAS, BTL developed through CASPER (Collaboration for Astronomical Signal Processing and Electronics Research)
- VEGAS uses 8x ROACH2 boards and NVIDIA GPUs
  - Integrated ADCs, FPGAs, 10 gigabit ethernet, serial communication ports, onboard flash memory perform initial conditioning, supply channelized data or I/Q values
  - Additional spectral line / pulsar processing performed on GPUs/CPUs
  - Data stored on beeg-fs distributed filesystem
  - 8 independent spectrometer banks for maximum frequency coverage/flexibility

# Hardware for Digital Backends







# Questions?

GREEN BANK OBSERVATORY

**greenbankobservatory.org**

*The Green Bank Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.*