An Interference-Resistant Search for Extraterrestrial Microwave Beacons

A thesis presented by

Darren Laney Leigh

to

The Division of Engineering and Applied Sciences

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the subject of

Applied Physics

Harvard University

Cambridge, Massachusetts

June 1998

An Interference-Resistant Search for

Extraterrestrial Microwave Beacons

by

Darren Laney Leigh

Submitted to The Division of Engineering and Applied Sciences
on May 21, 1998, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The Billion-channel ExtraTerrestrial Assay (*BETA*) is a radioastronomical search for microwave beacons from intelligent civilizations. It searches the 1400-1720 MHz "waterhole" region with 0.5 Hz resolution for narrow-band carriers. *BETA* incorporates several systems for terrestrial radio frequency interference mitigation: a terrestrial "veto" feed, two sky feeds to detect sidereal motion, and adaptive filtering to reduce intermittent interference. The search has surveyed the entire sky from $+60^\circ$ to -30° declination twice and is starting a third. During this time it has sifted through $\sim 10^{16}$ frequency bins, followed $\sim 10^9$ candidate features and archived 3500 of these which passed preliminary tests. No candidate has repeated or otherwise presents the assumed characteristics of an extraterrestrial intelligent origin.

Thesis Supervisor: Paul Horowitz
Title: Professor of Physics

The author can be contacted at `dlleigh@alum.mit.edu`.

Acknowledgments

BETA turned out to be a much larger project than we expected when we first started out, so we got a lot of help. A multitude of people provided support, ideas, research, blood, sweat and tears.

We received generous financial support from The Planetary Society and its members, the Bosack/Kruger Charitable Foundation, the Shulsky Foundation, Dr. John Kraus and NASA. Donations of equipment and parts were received from Micron Technology Inc., Hewlett-Packard Company, John Fluke Manufacturing Company Inc., Advanced Micro Devices Inc. and Intel Corporation.

The architecture of *BETA*'s spectrometer core evolved from ideas from the Berkeley SERENDIP program. Special thanks go to Dan Wertheimer for all of his help.

Our warmest thanks to the wonderful John Kraus for supplying the antenna calibration technique and absorber material described in Appendix A, on top of financial support.

We also received valuable help and encouragement from Michael Davis and John Hagen of the Arecibo Observatory, Martin Gimersky of the University of Victoria, BC, William "Crash" Yerazunis of Mitsubishi Electric Research Laboratories, David Staelin of MIT and Kevin Duesman of Micron Technology.

Massive support was provided by Joe Caruso, Robert Stefanik and Joe Zajac of the Oak Ridge Observatory. These are guys who climb up the dish during bad storms to pin it in place and otherwise protect it, handle student tours and field random visits from the press. I'm not sure which task is more dangerous. Our thanks, guys.

Many students worked on the design and construction of the project and we would especially like to point out the efforts of Greg Galperin (who performed a lot of research into the numerical properties of the FFT), the brilliant and super-helpful Derrick Bass (who worked on so many things it's hard to keep track), Neil Hendin and Nick "Phi" Schectman. A herd of others helped with the building stage: Suhail Shah, Eric Wey, Dylan Manna, James Higbie, George Marcus, Jonathan Wolff and Paul Eremenko.

Special thanks to my cohorts and fellow graduate students Jonathan "Jono" Weintroub, Charles "Chip" Coldwell and Ian "Max" Avruch. (How come they all have nicknames and I don't?) I couldn't have done it without you. The following people helped keep me sane: Tom Hayes, Carol Davis, J. D. Paul, Steve Rowley and others.

I'd like to thank my parents for raising me in an environment where learning and questioning were encouraged, where creativity and a love for education were fostered and where high expectations for the future were the norm.

I'd especially like to thank my committee, Professors Patrick Thaddeus, Victor Jones, Costas Papaliolios and Paul Horowitz for their patience, help and good advice. Extra special thanks go to Cos for always being around, always being super-friendly, always being knowledgeable and always being... Cos.

Now, what can I say about Paul Horowitz, my advisor, mentor and friend? Paul is God. There's no getting around that fact. He's brilliant, funny, creative, fun to be around, etc. I have learned *so* much. It's been a pleasure working for him. A lot of my friends have had evil advisors. They regale me with stories about how badly they are being mistreated by these professors, but they always end up telling me how much they're jealous of mine.

And if anyone reading this thinks that I've put this in to butter Paul up before my defense, think again. I've rigged LaTeX to say "**Stuff to be inserted later.**" in the version the committee gets to read. This stuff is only going into the version I turn into the department.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This dissertation discusses a system that is capable of detecting extraterrestrial microwave beacons in the presence of terrestrial interference. We will not deal with the probabilities of life and intelligence arising, or any of the other factors in the Drake Equation [14], nor will we treat the subject of non-electromagnetic signaling. All of these have been dealt with extensively elsewhere [43, 45, among many others...]. We will only discuss the process of initiating radio contact across interstellar distances and equipment and methods for accomplishing this.

Life arose quickly on the early earth, almost as soon as it could have. Among the 10^{11} sun-like stars in the galaxy there are undoubtedly planetary systems where life could exist. It is entirely plausible that many sites in our galaxy harbor intelligent life. Apart from these speculations, it is a *fact* that microwaves are altogether adequate for interstellar communication (as Purcell observed in 1960 [43]). From our current knowledge, they may even be an optimal method. Of course, there could be more advanced signaling techniques of which we are currently unaware, but the known methods can be used to do experiments now.

The experiment described herein, dubbed the Billion-channel Extraterrestrial Assay or *BETA*, was designed as a successor to the previous experiment called *META* (for Mega-channel ExtraTerrestrial Assay). Both have operated at the 26 meter radiotelescope at Agassiz Station (Oak Ridge Observatory) in Harvard, Massachusetts. *META* searched approximately 400 kHz of spectrum (in three different reference

frames) with 1/20 Hz resolution near the neutral hydrogen line at 1420 MHz. Technological advances in the decade since *META* was designed allowed us to construct a much more ambitious search: *BETA* has hundreds of times the frequency coverage as well as real-time analysis and follow-up capabilities. It is also far more complicated. Where *META* had one general purpose computer, *BETA* has 25 (with all of the inherent inter-processor communication issues). Where the *META* spectrometer generated a data stream of about 800 kBytes/sec which could be analyzed in software, *BETA*'s generates over 250 MBytes/sec and requires dedicated analysis hardware. Since its frequency resolution is 0.5 Hz, *BETA* is only about 1/10 as sensitive as *META*. The wider frequency bins also make it impossible to use the previous "doppler-chirp" method for detecting radio frequency interference; the new methods are a little more cumbersome, but adequate.

Our overall goal is to establish communication with an extraterrestrial civilization. This will involve sending information between star systems which are many light years apart. Relativity dictates that information cannot be transferred faster than the speed of light so any communication will take many years to proceed. Add this to the time scales required for biological evolution and we can see that interstellar communication is a waiting game that can be won only by the extraordinarily patient. It is our hope that most of the patience has already been displayed by more advanced civilizations so that we can communicate with them relatively quickly. We are arguably the youngest technological civilization anywhere (our radio communication capabilities arose only a few decades ago and are still developing) so this is not an unreasonable hope.

Chapter 2

Philosophy and Objectives

2.1 Interstellar Communication

2.1.1 Link Budget

First we will show that interstellar communication is not only possible, but easy to do with current technology. Figure 2.1 and equations 2.1 through 2.3 show a simple link

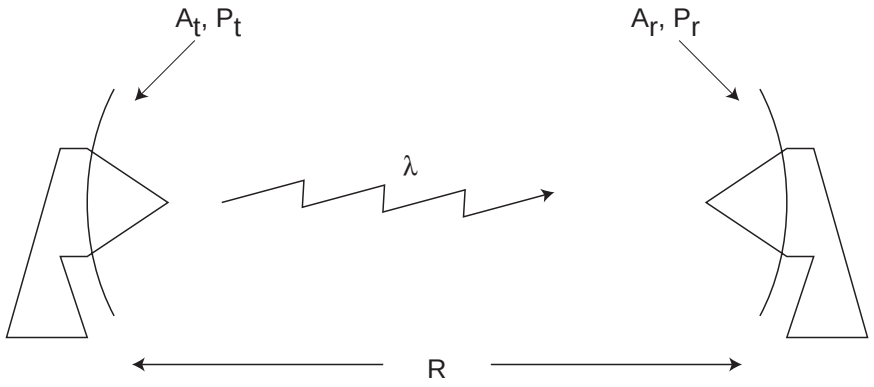


Figure 2.1: Interstellar communication link budget.

budget calculation of the power received at a distant antenna given various transmitter and system parameters. The variable P_t is power transmitted, P_r is the amount of that power collected by the receiving antenna, A_r and A_t are the effective areas of the receiving and transmitting antennas, respectively, λ is the wavelength and R is

the distance between the transmitter and receiver.

An isotropic transmitter of power P_t will produce a power flux of $\frac{P_t}{4\pi R^2}$ at a distance R. Non-isotropic transmitters have *directivity*, a function of direction, defined as the ratio of the power flux in a given direction to what the flux would be if the transmitter were isotropic.[1] Such a transmitter will therefore produce a flux of $\frac{P_t D_t}{4\pi R^2}$ where D_t is the directivity in the direction of interest. If a receiving device with an effective area of A_r is placed a distance R from the transmitter (which has directivity D_t in the direction of the receiver), then the amount of power it receives will be

$$P_r = P_t \frac{D_t A_r}{4\pi R^2} \tag{2.1}$$

Antennas have identical transmitting and receiving properties[2] with the relationship between their directivity and effective area [33] given by $D = \frac{4\pi}{\lambda^2} A$. Therefore a transmitter will have $A_t = \frac{\lambda^2}{4\pi} D_t$ and equation 2.1 becomes

$$P_r = P_t \frac{A_t A_r}{\lambda^2 R^2} \tag{2.2}$$

P_r will have to compete with any noise received or generated by the receiving system. If we are operating in the Rayleigh-Jeans region, the noise power generated in the receiving system is $P_n = kT_N B$ where k is the Boltzmann constant, T_N is the noise temperature of the receiving system and B is the receiver bandwidth (see Appendix A for details). If we integrate over a period of time τ and assume a bandwidth $B = 1/\tau$, then the ratio of received signal energy to noise energy (signal-to-noise ratio or SNR) is $\frac{P_r \tau}{kT_N}$. For a simple on-off keying modulation scheme (1 bit/τ – the transmitter is either on or off) and coherent detection over τ, the bit error rate (BER) of the received data stream will be $\mathrm{BER} = e^{-\mathrm{SNR}}$ (see Appendix D for information on noise statistics). An SNR of 4.6 yields a BER of 1%.[3] We can

[1] Colloquially this is called *gain*, but gain should formally include the losses of the antenna and feed system. See [32] for details.

[2] This is really only true of reciprocal systems, but that includes practically everything in use.

[3] A data link with a BER this high is quite usable when implemented with forward error correction (FEC).

combine these relationships into one formula giving the maximum distance a signal can be received given the transmitter and receiver characteristics, the wavelength, τ and the target BER:

$$R = \sqrt{\frac{P_t A_t A_r \tau}{-\ln(\text{BER}) \cdot k T_N \lambda^2}} \tag{2.3}$$

Substituting in some reasonable values

$$\begin{aligned} P_t &= 10^6 \text{ Watts} \\ A_t = A_r &= \pi \times 10^4 \text{ m}^2 \text{ (200 meter dishes)} \\ \lambda &= 3 \text{ cm (X-band)} \\ \tau &= 1 \text{ second} \\ \text{BER} &= 10^{-2} \\ T_N &= 10\text{K} \end{aligned}$$

yields a distance of $R = 4 \times 10^{19}$ meters or over 4000 light years. If electricity costs 10¢ per kWH, then we can send 36 bits for a dollar. With compression one word can easily be represented with 36 bits. This means that with a system as described above, we can send interstellar telegrams a distance of 4000 light years for an energy cost of about \$1/word. There are about 2×10^7 sun-like stars within that distance. Communication between star systems is not only possible, it's cheap!

2.1.2 The Interstellar Medium

In order to be received here on Earth, an extraterrestrial signal must cross many light years of space. The intervening distance is not empty and the material present in the interstellar medium (ISM) will affect the signal in several different ways.

Absorption

Some of the signal will be absorbed. The amount of absorption will depend on the frequency of the signal, the distance that it has traveled and what part of the galaxy it passes through. Broadband absorption in the ISM is chiefly due to dust grains. When the wavelength of the signal becomes larger than the grain size ($\sim 10\,\mu$m) then

there is very little absorption. Shorter wavelengths can be attenuated severely.

Over narrower bandwidths, there can be substantial absorption due to particular spectral lines. At extremely low frequencies, $\lesssim$ the plasma frequency ($\sim$ 1 kHz for the ISM), there will also be absorption (and reflection, refraction, etc.) due to plasma effects.

From the earth's surface, reception will be strongly affected by the ionosphere (plasma frequency $\sim$ 10 MHz) and molecular lines (especially H_2O at 22 GHz and O_2 at 60 GHz).

Dispersion

Because of the presence of free electrons, the interstellar medium is dispersive; signals of different frequencies travel at different speeds. We have much information about this from pulsar studies. The group velocity is given by

$$v_g = c\sqrt{1 - \left(\frac{\nu_0}{\nu}\right)^2} \tag{2.4}$$

where ν_0 is the plasma frequency. We can derive an approximation for the difference in arrival times of two simultaneously transmitted signals of different frequency, ν_1 and ν_2, assuming uniformity of the interstellar medium and that $\nu_0 \ll \nu_1, \nu_2$. Since $\sqrt{1+x} \approx 1 + \frac{1}{2}x$ for $x \ll 1$, then

$$v_g \approx c\left[1 - \frac{1}{2}\left(\frac{\nu_0}{\nu}\right)^2\right] \tag{2.5}$$

If the signals are transmitted at time $t = 0$, then their arrival times will be

$$t_1 = \frac{L}{v_1} \text{ and } t_2 = \frac{L}{v_2} \tag{2.6}$$

$$\Delta t = L\left(\frac{1}{v_2} - \frac{1}{v_1}\right) = L\left(\frac{v_1 - v_2}{v_1 v_2}\right) \approx L\frac{\Delta v}{c^2} \tag{2.7}$$

and Equation 2.5 gives us

$$\Delta t \approx \frac{L}{2c}\left[\left(\frac{\nu_0}{\nu_2}\right)^2 - \left(\frac{\nu_0}{\nu_1}\right)^2\right] \tag{2.8}$$

where L is the distance the signals have traveled. Dispersion will affect a modulated (finite bandwidth) signal by smearing it out in time. If we assume a narrow-band signal, i.e. $\nu_2 = \nu_1 + \Delta\nu$ where $\Delta\nu \ll \nu_1$, then equation 2.8 becomes

$$\Delta t \approx \frac{L\nu_0^2\Delta\nu}{c\nu_1^3} \tag{2.9}$$

A modulated signal will suffer distortion when $\Delta t \approx 1/\Delta\nu$ or

$$\frac{\Delta\nu}{\nu_1} \approx \sqrt{\frac{c\nu_1}{L\nu_0^2}} \tag{2.10}$$

For example, if $\nu_1 = 1$ GHz, $\nu_0 = 1.6$ kHz and $L = 1000$ light years, then $\Delta\nu \approx$ 100 kHz. Wider bandwidth modulation schemes will have problems at interstellar distances.

Faraday Rotation

Faraday rotation [44] is related to dispersion. Because of free electrons, a magnetic field parallel to the direction of propagation will cause one circular polarization to travel faster than the other. A linearly polarized signal can be decomposed into two equal-amplitude circularly polarized signals, one right-handed, the other left-handed. The direction of linear polarization depends on the phase difference between the two circularly polarized signals. If one of those travels faster than the other, the phase difference between them will change, causing the direction of linear polarization to rotate. The result is that the direction of polarization of any linearly polarized signal will change significantly over interstellar distance. Even worse, the earth's ionosphere and strong magnetic field will cause a further, time-variable change for ground-based observers. [31]

Scintillation

Scintillation [52] is Rayleigh-fading of a signal caused by multi-path transmission. The scattered parts of a signal recombine with different phases (due to the difference in distances they have traveled) and interfere constructively or destructively. Cordes and Lazio [8] describe several relevant effects that this will have:

1. Diffractive intensity scintillations causing signal variations in both time and frequency, "analogous to the twinkling of the stars". Nearby transmitters ($<$ 100 parsecs at 1 GHz) will show weak scintillation with much less than 100% variation in intensity. Farther sources ($>$ 500 parsecs at 1 GHz) will show strong scintillation with 100% intensity variations. The characteristic time scale for these variations "is of order minutes only for the most distant regions of the galaxy. For most directions, it is typically hours." [9]

2. Spectral broadening due to the scattering. For most lines of sight through the galaxy this will be $\lesssim 0.1$ Hz at 1 GHz, though some paths will have as much as 5 Hz. [7]

3. Temporal broadening due to the differential arrival times of the scattered signal.

These effects will cause the distortion of modulation schemes but, more importantly, large intensity fluctuations can make a signal disappear entirely. On the other hand, constructive interference can occasionally cause a signal normally too weak to be detected to rise above threshold. Unfortunately, the statistics make these amplifications rare, and scintillation harms us on the average. [9]

2.1.3 Beacons or Leakage?

There are two types of signals which could be received by a SETI system: intentionally transmitted *beacons* and unintentional *leakage* radiation. The design of our system and the search strategy will depend on which type we are looking for.

A beacon is a signal which has been designed and radiated for the express purpose of initiating interstellar communication. It is meant to attract attention and may

carry no information other than its very existence. The signal will be designed to appear artificial, to be easy for a search to detect and to be able to cross the interstellar medium with minimal corruption.

Leakage radiation includes signals that are designed and radiated for use by the transmitting civilization and are not specifically intended to be received by others. Some examples from earth are television carriers (EIRP[4] $\sim 10^7$ watts) and the Arecibo S-band radar (EIRP $\sim 10^{13}$ watts). One advantage of leakage is that the transmitting civilization need not do anything special for the receiver to take notice. A major disadvantage, however, is that leakage from an advanced civilization may be weak, intermittent or non-existent. Even if a leakage signal were constant and strong, we might have difficulty distinguishing it from natural noise. As a signal is transmitted more efficiently, redundancy is removed and it appears increasingly noise-like. It is unlikely that an advanced civilization would transmit strong signals inefficiently.

For these reasons we have designed our SETI system to receive beacons. The following sections discuss the design of a beacon and appropriate search strategies.

2.1.4 The Spectrum of Interest

What region of the spectrum shall we search? An interstellar beacon would be transmitted at a frequency with little interfering background noise, where there is little absorption and where a receiving civilization is likely to look.

Figure 2.2 shows the characteristics of various parts of the spectrum. Below about 1 GHz, synchrotron radiation from the galaxy contributes significant noise excess. The quantum nature of EM radiation produces shot noise of effective temperature $T_{\text{eff}} = \dfrac{h\nu}{k}$ which starts to become significant above about 100 GHz. In between, the sky's background noise temperature is dominated by the unavoidable 3° K cosmic microwave background (CMB). Receivers on the earth's surface will be hampered by the atmosphere's extremely absorptive molecular lines above about 50 GHz.

Cocconi and Morrison [5] did the first study of the best part of the EM spectrum

[4]Effective Isotropic Radiated Power = $P_t D_t$.

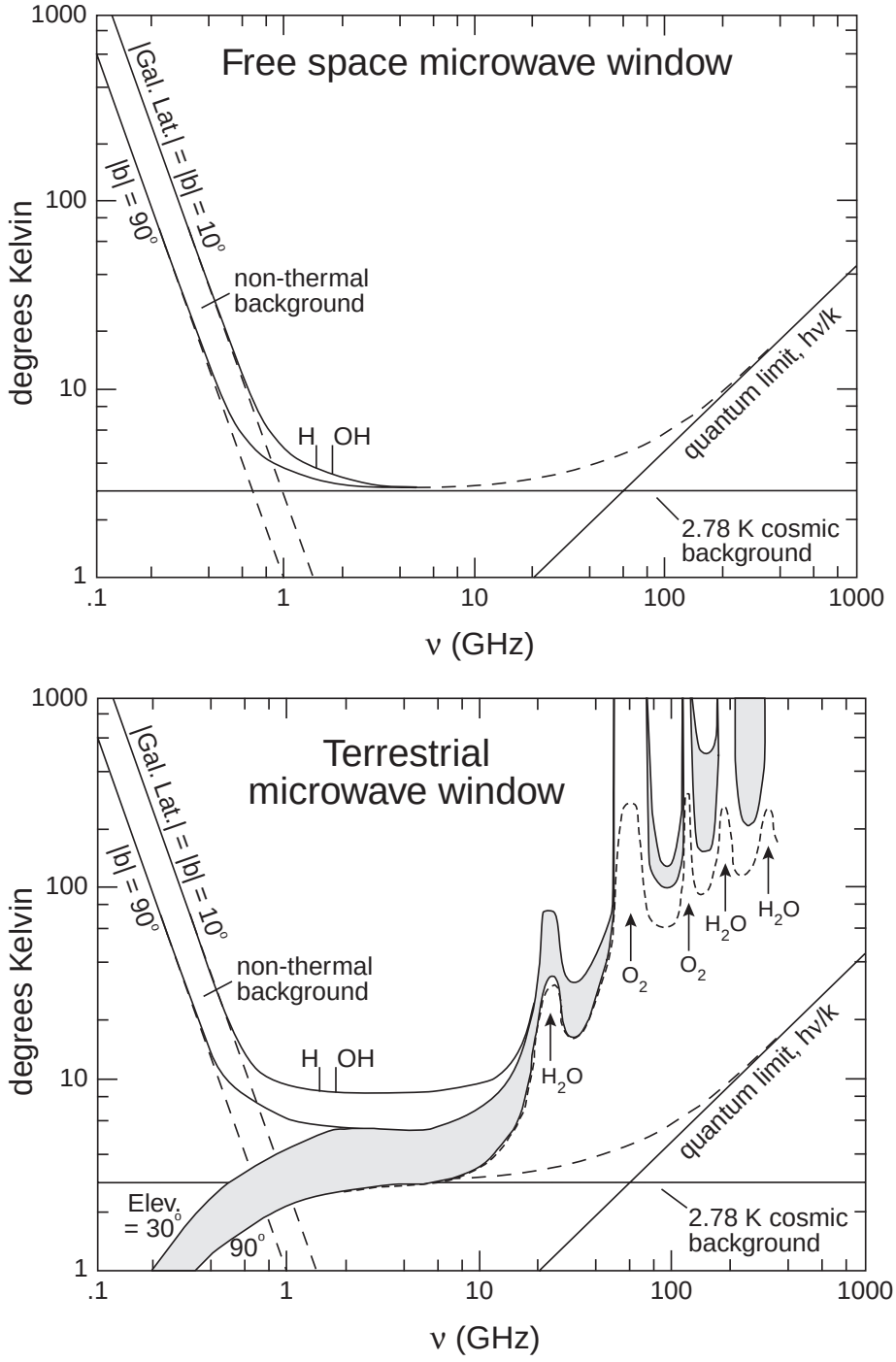


Figure 2.2: Free space and terrestrial microwave windows with the H and OH "magic" frequencies marked. Note the atmosphere's strongly absorptive molecular lines. These diagrams are adapted from [38].

for interstellar communication. They concluded that an excellent location is the vicinity of the neutral hydrogen line at 1420 MHz. The sky is quiet there, absorption is negligible and the hydrogen line is a "magic" frequency.

Since there is so much spectrum to search, it would be a good idea for a beacon to transmit at a frequency that can be easily guessed. Prominent spectral lines are a good choice since they would be known to both civilizations and a beacon might be detected serendipitously during a regular astronomical survey. Since hydrogen is the most abundant element in the universe and surveys of this line are necessary for mapping the structure of the galaxy, it seems to be a natural place to look. Other magic frequencies have been suggested: harmonics of the hydrogen frequency, fundamental constants (π and e) times it, other spectral lines, etc.

The "Water Hole"

More generally, we think it is a good idea to search the entire "water hole". This is the 300 MHz piece of spectrum between the magic frequencies of the neutral hydrogen line and the hydroxyl lines (1612, 1665, 1667 and 1720 MHz). H and OH make water, the stuff of life (as we know it) so it is also biologically relevant. More poetically, in the wild, animals of different species gather around water holes, so this part of the spectrum might be a good place for alien civilizations to meet. [38]

2.1.5 Signal Characteristics and Guesses

Since we have no *a priori* way of knowing what an extraterrestrial signal would look like, we have to make educated guesses based on our knowledge of physics and signal processing. It is reasonable to assume that the transmitting civilization would design its signal to be easily received, efficient and as simple as possible.[5]

[5]This is the Search for ExtraTerrestrial *Intelligence*. We're never going to find stupid aliens.

Waveforms and Modulation

To be confident of the validity of a received signal, there must be a reasonably high signal-to-noise ratio (SNR). While the receiving system can be designed to reduce noise levels, a properly designed signal can contribute greatly to high SNR.

The amount of information C we can transfer over a noisy link is bounded by the Shannon limit [48] which states that

$$C \text{ (bits/second)} \leq B \log_2 (1 + P_r/P_n) = B \log_2 \left(1 + \frac{P_r}{kTB}\right) \tag{2.11}$$

C is called the *channel capacity*. It is obvious from equation 2.11 that we can increase the data rate of a signal by increasing P_r or decreasing T. The effect of changing B is less obvious, but can be seen by differentiating:

$$\frac{\partial C}{\partial B} = \frac{1}{\ln 2} \left[\ln \left(1 + \frac{P_r}{kTB}\right) - \frac{\frac{P_r}{kTB}}{1 + \frac{P_r}{kTB}} \right] \tag{2.12}$$

This has roots only at $\pm\infty$ and is positive for $B > 0$ so C increases monotonically with B. We can maximize C by increasing B until $\frac{P_r}{kTB} \ll 1$ and

$$C_{\max} \approx \frac{1}{\ln 2} B \frac{P_r}{kTB} = \frac{P_r}{kT \ln 2} \tag{2.13}$$

so the maximum data rate no longer depends strongly on the bandwidth. Using a large bandwidth and low SNR is a common technique in spread spectrum [12] and other high data rate systems. One problem with these techniques is complexity: they often have many parameters and degrees of freedom. For example, a receiver of spread spectrum signals needs to know the approximate frequency of the transmission, the spreading code and the code phase. In SETI these modulation parameters would not be known ahead of time and the system would have to search for them. A properly designed beacon signal should require the receiver to search as small a parameter space as possible.

A beacon need not transmit any information other than its existence which means

that, theoretically, the SNR can be arbitrarily low. However, if we need to verify the existence of a beacon within a finite time τ, the amount of information transmitted must be 1 bit/τ, i.e. the beacon either is or is not transmitting during τ. Over the period τ the receiver has intercepted an energy of W_r from the beacon which will compete with an energy of W_n from the system noise. The bandwidth B equals $\frac{1}{\tau}$ and so equation 2.11 becomes

$$C = \frac{1}{\tau}\log_2\left(1 + \frac{W_r/\tau}{W_n/\tau}\right) = \frac{1}{\tau}\log_2\left(1 + \frac{P_r\tau/\tau}{(kT\frac{1}{\tau})\tau/\tau}\right) \tag{2.14}$$

With $C = 1$ bit/$\tau = \frac{1}{\tau}$ we find

$$C = \frac{1}{\tau} = \frac{1}{\tau}\log_2\left(1 + \frac{P_r\tau}{kT}\right) \tag{2.15}$$

and

$$\frac{P_r}{kT} = \frac{1}{\tau} = C = \ln 2 \cdot C_{\max} \tag{2.16}$$

so this kind of strategy can still operate at about 69% of the theoretical maximum channel capacity.

A simple beacon could be a continuously transmitted, unmodulated sinusoid. In this case a Discrete Fourier Transform (DFT) would be used on the sampled time-domain signals to concentrate the received signal energy into a few bins while distributing the the noise energy equally across all bins. The bin width B is $\frac{1}{\tau}$ where τ is the duration of the input sample series.

There are an infinite number of transforms which have this signal-concentrating, noise-distributing characteristic. Which transform is used will depend on the nature of the expected signal. The Karhunen-Loève (KL) transform expands a given input signal into the set of magnitudes of its projections onto a set of orthonormal basis vectors. [40] An input signal that looks a lot like one of the basis vectors will have a very large projection onto that vector and small projections onto the other vectors. On the average, noise will look like none of the basis vectors and so its projections onto them will be small and more-or-less equal. The Fourier transform is a special

case of the KL transform where the basis vectors are sinusoids. Some research has been done in the use of the general KL-transform for SETI. [37]

With an infinite number of possibilities, what type of signal should we listen for? Something simple with only a small parameter space to search. Two signal types come readily to mind:

1. **Pulses** The basis vectors are short pulses with different time delays. One of the beauties of this is that input signals are already in the transform domain and so very little computation needs to be done. On the downside, dispersion smears out radio frequency pulses making it harder to distinguish them from noise. There are natural radio sources which emit pulses (pulsars) which could make distinguishing an artificial pulse source confusing. On the other hand, a pulsed beacon might be picked up serendipitously during a pulsar search. Pulse searches at both radio [53, 26] and optical [49, 28] frequencies have been performed.

2. **Sinusoids** The basis vectors are sinusoids with different frequencies and phases. Since a constant phase carries no information, we can remove it by taking the square magnitude in the frequency domain. This leaves us with a single, linear parameter space to search. The natural radio sources with the narrowest bandwidth are microwave masers (bandwidth $\sim 10^3$ Hz). Since an artificial radio transmitter can easily generate sinusoids of bandwidth much less than 1 Hz, extreme narrowness in frequency is an excellent indicator that the signal is non-natural.[6] The ISM treats narrow sinusoids kindly: dispersion has negligible effect and scintillation will normally cause less than 1 Hz of broadening. The large intensity fluctuations caused by scintillation, however, can be a problem.

In addition to their good propagation characteristics, sinusoids seem to be a "natural" signal type. They are seen in spectral lines, orbital motion, pendula and other ele-

[6]There is excellent reason to believe that very narrow-band natural radio sources do not exist. Any such source strong enough to be received over interstellar distances has to be either very large, or very hot. If it is large, differential motion within the source will widen the bandwidth due to the doppler effect. If it is hot, thermal motion will also cause doppler broadening.

mentary physics, because they are solutions to simple second-order, linear differential equations. Pulses, chirps and pseudo-random spread spectrum are encountered less frequently in nature.

Polarization

Because of Faraday rotation, it is likely that a beacon will be transmitted using circular polarization. Since there is no way to know ahead of time which handedness, left or right, the senders will pick,[7] a SETI receiver needs to monitor both polarizations. One way to do this is to double the amount of equipment used and monitor both polarizations simultaneously. Another way is to alternate between them, spending half of the observing time on each. This keeps the equipment cost down but decreases the time resolution of the search by a factor of two. A third option is to receive with a linearly polarized feed. This will pick up both circular polarizations without affecting equipment cost or time resolution, but the received signal strength will be a factor of two lower. Once a signal is detected, the receiver can then be configured to the proper circular polarization for higher signal strength.

One interesting method that the transmitting civilization could use is on-off keying via polarization modulation: constant power output with right-hand circular for mark and left-hand for space. Such a beacon would

- be clearly recognized as artificial,
- always be transmitting so it wouldn't be missed even during the "off" part of slow on-off keying,
- be readily received regardless of the polarization of the receiver, and
- carry information about the distance to the transmitter and the intervening interstellar medium due to the differential arrival times of the two polarizations. This would be useful in planning modulation schemes for the return signal.

[7] Though perhaps there is some galactic standard of which we are unaware.

2.2 Search Strategies

2.2.1 Directed or All Sky?

Search strategies are usually divided into two categories: *directed* and *all sky*. A directed search concentrates on individual targets (usually stars), tracking them over a period of time. This has several advantages: Radio telescopes with very high resolution can be used since the target is unresolved and has a known position. Integration techniques can be used to gain higher signal-to-noise ratios over time. The doppler shifts (at the reception site, relative to an inertial frame) are well known and the search can completely compensate for them. On the down side, only a small portion of the sky can be searched so if the civilization is not within the set of targets, it will be missed.

An all sky search can observe the entire sky visible from the observatory, so it does not share its designers' target prejudices. However, if one wants to search the sky within a reasonable amount of time, this kind of strategy cannot use integration techniques or high resolution telescopes. Since the target position is unknown, it will not be possible to completely compensate for doppler shifts and the search will have to use compromise techniques.

2.2.2 Optimizing Search Parameters

Certain parameters of the search system, such as the beamwidth and integration time, affect the sensitivity and amount of sky coverage. We would like to see quantitatively what the effects are. The directivity D of a dish antenna will be approximately equal to $\frac{4\pi}{\Omega}$ where Ω is the solid angle (in steradians) subtended by the beam. Since an antenna's directivity is related to its effective area by $D = \frac{4\pi}{\lambda^2} A_{\text{eff}}$, then $A_{\text{eff}} = \frac{\lambda^2}{\Omega}$ and Equation 2.1 becomes

$$P_r = \frac{P_t A_t}{\Omega_r R^2} \tag{2.17}$$

A signal may be detected if the energy received from it is significant compared to the noise energy competing with it, say $E_r = \gamma E_n$. With coherent detection methods

$B = 1/\tau$ and $E_n = kT_N \frac{1}{\tau}\tau = kT_N$ which is constant for the system, so the probability of correctly detecting a signal is a function of $E_r = P_r\tau$ only. Therefore,

$$E_r = \gamma k T_N = \frac{P_t A_t \tau}{\Omega_r R^2} \tag{2.18}$$

and

$$R = \sqrt{\frac{P_t A_t \tau}{\Omega_r \gamma k T_N}} \tag{2.19}$$

If $n_\star$ is the number density of stars in the region being scanned, then the total number of stars in the detectable range of the search is

$$N_\star = \frac{1}{3} n_\star \Omega_r R^3 \tag{2.20}$$

$$= \frac{1}{3} n_\star \Omega_r^{-1/2} \tau^{3/2} \left(\frac{P_t A_t}{\gamma k T_N} \right)^{3/2} \tag{2.21}$$

which gives the surprising result (as shown by Drake [15]) that we actually "see" more stars with a larger antenna (even though the solid angle being observed is smaller) because the added sensitivity lets us look deeper. This suggests searching with the largest possible antenna, and therefore the smallest possible solid angle. Because of the 3/2 exponent, larger values of τ (longer integrations) and smaller values of γ and T_N (lower thresholds and system noise temperature) are also very important.

What if we want to observe the entire sky? The number of observations required is $\frac{4\pi}{\Omega_r}$ and the time necessary to complete them will be $\tau\frac{4\pi}{\Omega_r}$. If we need to finish in a specific time Y, then $\tau = Y\frac{\Omega_r}{4\pi}$ which makes

$$N_\star = \frac{1}{3} n_\star \Omega_r \left(\frac{P_t A_t Y}{4\pi \gamma k T_N} \right)^{3/2} \tag{2.22}$$

This would seem to suggest that for an all-sky search, we should observe using the largest possible solid angle. However, our analysis assumed that we can coherently integrate a signal for as long as we wish. This is not possible because the signal itself will have a non-zero bandwidth B_s ($\sim 10^{-2}$ Hz in the interstellar medium), so the

largest useful τ is $1/B_s$. We could still do non-coherent averaging, co-adding power spectra, which yields an improvement in SNR proportional to $\tau^{1/2}$ instead of τ. This also has limitations: changing noise statistics and gain fluctuations in the receiver make very long integrations tricky. A strategy that observes the most candidates will therefore include:

- Lowering the system noise temperature as much as possible,
- Setting the detection thresholds as low as the analysis speed permits (or conversely, including as much fast analysis hardware as possible),
- Choosing the longest possible integration time (a function of the signal's inherent bandwidth and system stability parameters),
- And observing with the largest, most sensitive telescope that can be obtained.

2.2.3 Interference Rejection

A good candidate signal needs to have two qualities: it must be artificial and it must not come from earth.[8] A signal with a bandwidth less than 1 Hz is undoubtedly artificial so a good strategy is to perform a Fourier transform and then examine the narrow features. Proving that a signal did not come from earth is trickier. Earth has many transmitters in the frequency range of interest so there is a serious interference problem. We need to filter all of the received narrow-band features and pass only those with some quality that an interstellar signal would uniquely possess.

One thing we do know about an interstellar signal is that it would come from a fixed location on the celestial sphere.[9] All SETI projects take advantage of this characteristic in one way or another to filter out terrestrial interference.

Project *META* [24], *BETA*'s predecessor, used the doppler chirp (changing doppler shift) caused by the earth's rotation to distinguish terrestrial signals from interstellar

[8] Or from an earth-built spacecraft.

[9] If alien signal-emitting craft exhibit obvious proper motions, then interstellar transportation is much easier than we have conjectured and SETI will fail anyway.

ones. Non-terrestrial signals near the zenith show a $\sim 10^{-1}$ Hz/sec doppler shift (at L-band) due to the observatory accelerating towards the earth's rotation axis. Terrestrial signals are in the same frame of reference as the observatory and therefore do not show such a shift. *META* swept its local oscillator at precisely the right rate to cancel any doppler shift for non-terrestrial signals while adding that shift to terrestrial ones. With a frequency bin width of 1/20 Hz and an integration time of 20 seconds, terrestrial signals were spread across ≈ 60 bins, while non-terrestrial ones would remain only a few bins wide. Despite this clever method, 74 candidate signals passed the test. Half of these were later proved to be terrestrial in origin, but there was not enough information to verify or eliminate the remaining half.

The Big Ear [30] project at Ohio State University used the time history of signals to classify them as terrestrial or non-terrestrial. It used two horns, one pointed slightly east and the other slightly west, and subtracted their signals, causing a sidereal point source to trace out a characteristic double lobe pattern. The "WOW" signal was found by this search, but has never repeated despite a large number of re-observations.

Project SERENDIP [2] at the University of California at Berkeley uses interference databases, frequency drift analysis, antenna beam shape and other methods to reject interference. [4]

The SETI Institute's Project Phoenix [51] also uses the doppler shift due to the earth's rotation to filter out terrestrial interference. A second radio telescope and receiver (called a FUDD for "follow-up detection device" [11]) are placed a substantial distance ($\sim 10^2$ km) from the main radio telescope. This has two advantages:

1. Very local interference will be received by only one of the telescopes and can thus be rejected.

2. An interstellar signal will be received by both telescopes at *different* frequencies, because the earth's rotation moves the telescopes at different speeds along the line-of-sight. Interfering signals are very unlikely to mimic this behavior.

Search Criteria

Due to past experience we wanted our search to be more robust than previous searches. It should be designed to provide sufficient information to either prove or disprove the source of a signal. It should allow immediate, automatic follow-up so that short-lived signals would not remain unproven mysteries. We decided to use a combination of beam shape and doppler shift to filter out terrestrial signals. Two beams point at the sky, one slightly east and the other slightly west. Any signal coming from a fixed sidereal position will trace out the characteristic pattern of the beam lobes, first in the east beam and then in the west. There was no need to use a swept LO as in *META* since *BETA*'s frequency bins are 10 times as wide. Over the course of a 2 second integration, the 0.14(max) Hz/sec doppler chirp will smear a narrow carrier over 0.28 Hz. This is smaller than the system's 0.5 Hz frequency resolution so it does not present a problem. The LO frequency does need to change between integrations, however, since a sidereal signal will remain in the beam system for about six minutes. Fixed frequency terrestrial signals will not show the same doppler characteristic so this adds a measure of interference protection as well.

A third low-gain, azimuthally omnidirectional antenna looks predominately at the horizon to pick up any interfering terrestrial signals. The spectra from all three beams (east, west and terrestrial) are computed simultaneously and synchronously so that any frequency comparisons can be quite accurate. *BETA* was designed to follow the time history of these spectra over the course of several minutes (enough time for a sidereal source to transit both sky beams) and compare them to the expected behavior of an extraterrestrial signal. If the comparison is good, the antenna can be moved to allow the source to pass through the beams several more times. We call this "leapfrogging".

Because of this immediate re-observation capability, the data analysis must be done in real-time. Offline analysis or any significant delay might allow a source to disappear before it can be observed again. This means that the analysis algorithms must be implemented chiefly in hardware, and that they must be reasonably simple.

Chapter 3

Architecture and Implementation

3.1 Overall Architecture

Figure 3.1 shows a block diagram of the entire system. In a nutshell: L-band microwaves are collected by the dish in dual feedhorns, amplified, downconverted to an intermediate frequency (IF), split into bands, downconverted to 20 parallel basebands and digitized. The digital time-domain data is sent to the spectrometer which performs a fast Fourier transform (FFT). This frequency domain data is then forwarded to the feature recognition (FR/FC) hardware which looks for "spikes": narrow features in frequency that are much stronger than the background noise level. Certain spikes are designated as "hits" and their time history is tracked. The histories are sent to the Unix workstation where a battery of tests are performed on them. Those passing the tests are archived; those failing are discarded.

The system has a 40 MHz instantaneous bandwidth frequency coverage in each of the dual sky feeds and terrestrial feed. It covers the entire 1400-1720 MHz "waterhole" in eight 2-second hops.

This chapter discusses these operations in depth, as well as some other systems which contribute to the maintenance and general well being of *BETA*. Table 3.1 lists the specifications of the system.

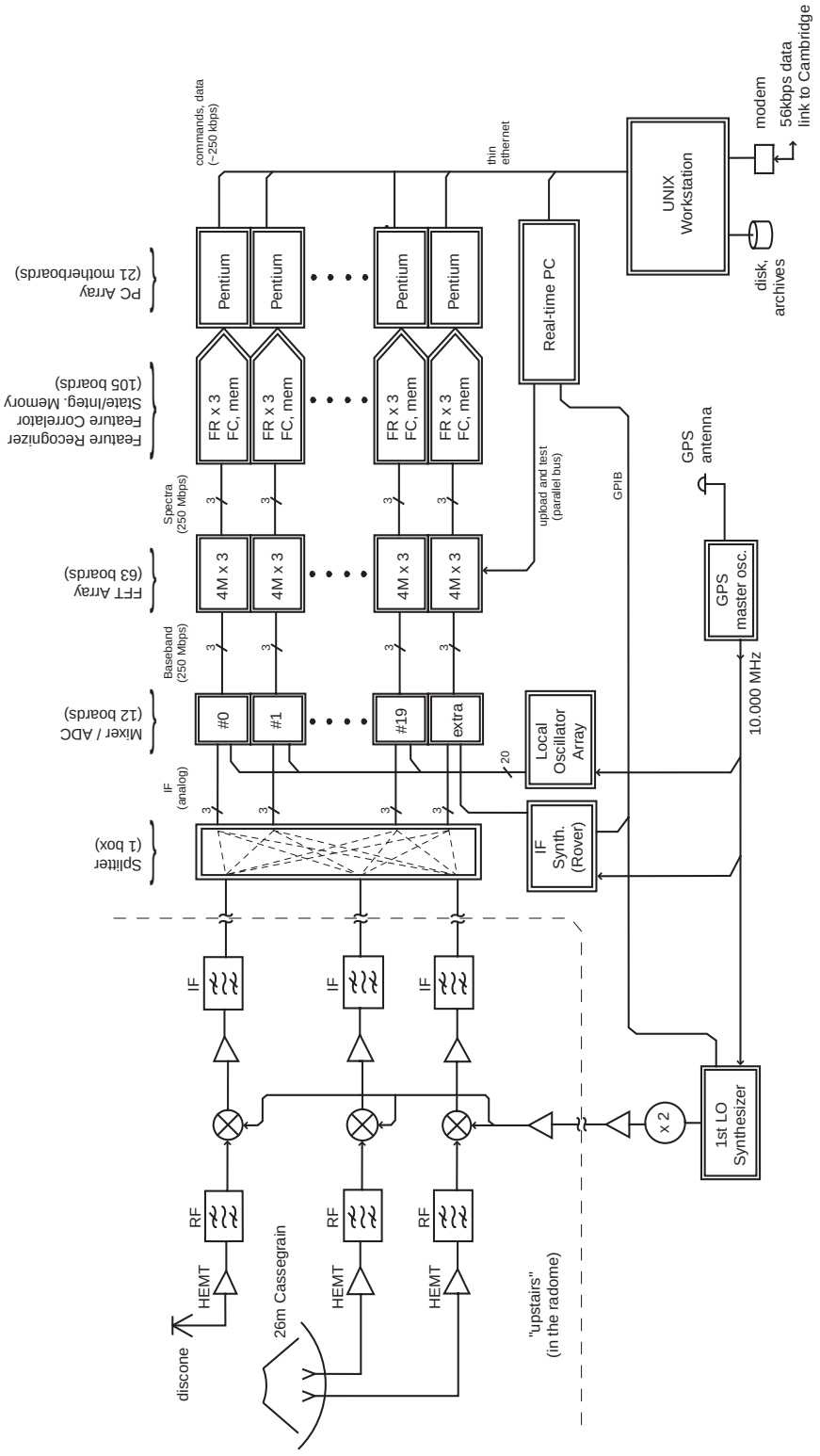


Figure 3.1: Block diagram of the Billion-channel ExtraTerrestrial Assay

Antenna: 26 meter Cassegrain, equatorial mount, fully steerable.

Observatory: Agassiz Station (Oak Ridge Observatory). 71.5583° west longitude, 42.5036° north latitude, 183 meters altitude.

Feed System: 21-cm dual pyramidal horns, linearly polarized.

Amplifiers: 1.4–1.7 GHz low-noise HEMT: $T_n = 30$K.

Spectrometer Core: FFT based on 40 MHz Austek A41102 integrated circuits. 63 FFT boards, 2^{22} channels per board, 251,658,240 channels (+ 12,582,912 spares), 2 second integration, 0.5 Hz per channel ($B\tau = 1$), 40 MHz total instantaneous bandwidth in each of three feeds, doppler acceleration compensated LO.

Sky Coverage: declination $-30°$ to $+60°$ (70% of the sky), 0.5° per day, ~ 1 year to complete entire sky.

Frequency Coverage: 1400–1720 MHz (the "waterhole") in 8 hops (2 seconds per hop) of 40 MHz each.

Sensitivity: $T_r \approx 85$K, 1.5×10^{-22} W/m^2 (correct linear polarization) and 3.0×10^{-22} W/m^2 (circular polarization) for $15P_0$ detection of celestial carrier in bandpass.

Signal Analysis Hardware: 21 60-MHz Pentium microcomputers with 4 special-purpose ISA boards each, Sun Sparc 20 workstation.

Interference Rejection Techniques: 2 sky beams for sidereal position verification, 1 horizon antenna for terrestrial veto, manual and adaptive frequency filtering. Immediate followup of candidates.

Miscellaneous 56 kbps link to Cambridge, full remote control of experiment and antenna. Uninterruptible power supply.

Table 3.1: Specifications and details of *BETA*.

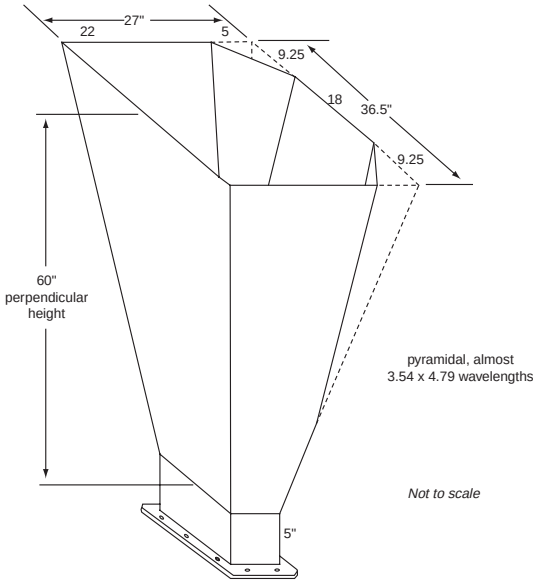


Figure 3.2: Physical measurements of the horn antennas.

3.2 RF Hardware

3.2.1 Antenna System

The business end of *BETA* is the 26 meter, equatorial mount, fully steerable, Cassegrain radio telescope at Agassiz Station in Harvard, Massachusetts. At L-band it has an effective area of about 250 square meters and a beam width of about one-half degree. The dish is fed by two linearly polarized feed horns, set up so that one horn's beam points slightly to the east and the other's points slightly to the west of the symmetry axis. The axis of polarization is parallel to lines of constant declination. The feed horns are constructed of $1/8''$ welded aluminum sheet and are pyramidal horn antennas. In order to fit both horns into the circular radome, we designed them with "lopped-off" corners. This doesn't seem to affect their performance significantly.[1]

The radio telescope optics were designed for a single feed horn, directly on-axis. Moving a feed off-axis is called "squint"; it decreases overall gain, widens the main lobe and increases the side lobes. Figure 3.4 shows the beam shapes of the two horns,

[1]During the design stage we were wondering how well this trick would work and so asked the eminent Ed Purcell. His response? A big shrug and "What the hell."

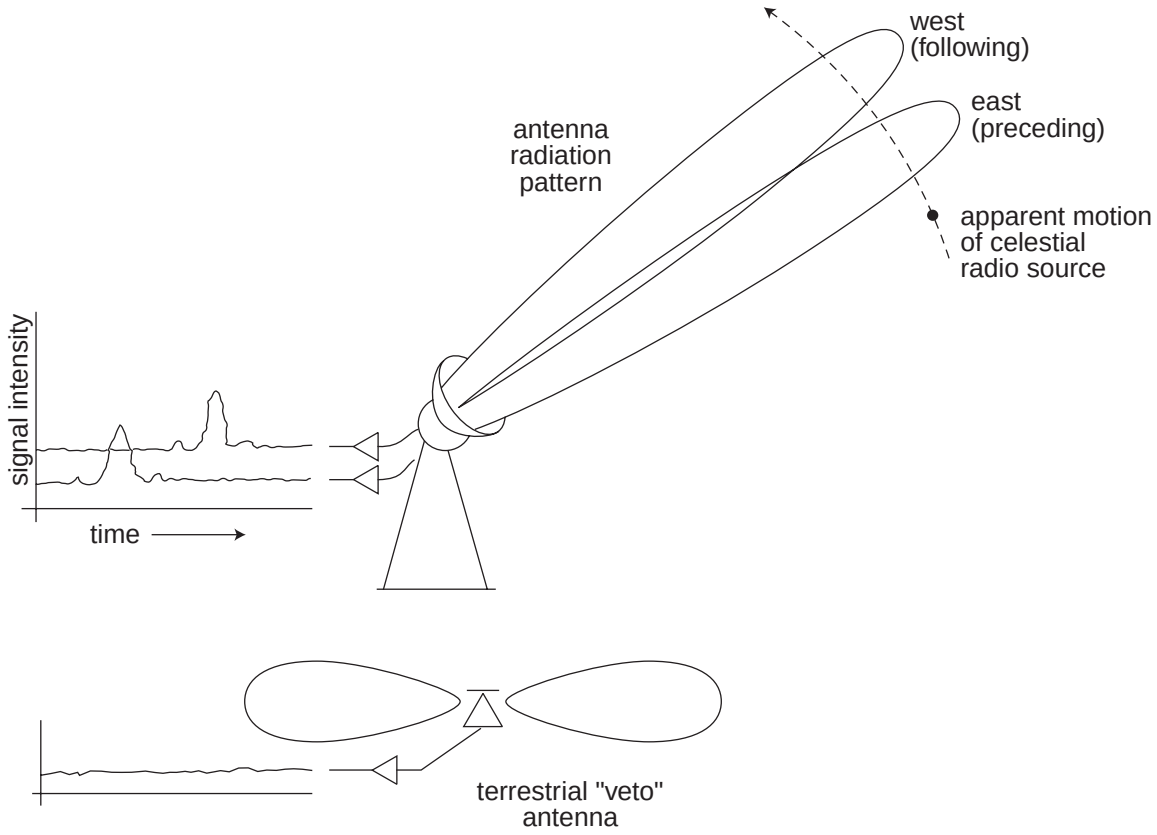


Figure 3.3: *BETA*'s three beam system for interference rejection.

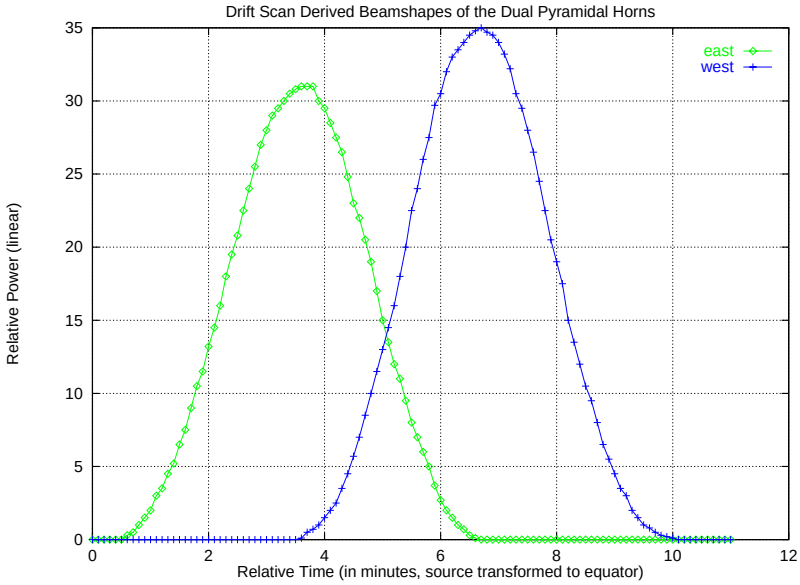


Figure 3.4: Drift scan derived beam shapes of the dual pyramidal horns

derived from a drift scan of Cygnus A (declination = 40.6°). The data has been transformed to make it appear as if the source were on the equator, i.e one minute of time equals fifteen minutes of arc. Despite the "squint" of the two beams, the drift scan shows that the lobe patterns are good and that the side lobes are at least not horrible.

A third feed is used for RFI discrimination. It looks at the horizon and any signals detected there can be vetoed due to their presumed terrestrial origin. This antenna is required to be broad-band, azimuthally omni-directional and a good match to 50Ω. The observatory is very close to two cellular phone base stations. To avoid desensitization problems from these powerful transmitters a further requirement is that the antenna be *poorly* matched below 900 MHz. A discone design can meet these requirements and has the additional advantages of being easy to construct and forgiving of errors. Figure 3.5 shows the voltage standing wave ratio at the antenna feed point. This was calculated from return loss measurements made using a spectrum analyzer with tracking generator and a directional coupler. The VSWR is below 1.7 : 1 for all frequencies in the range of interest and below 1.5 : 1 for most of them. Notice that the VSWR climbs dramatically at low frequencies and is well above 5 : 1 in the

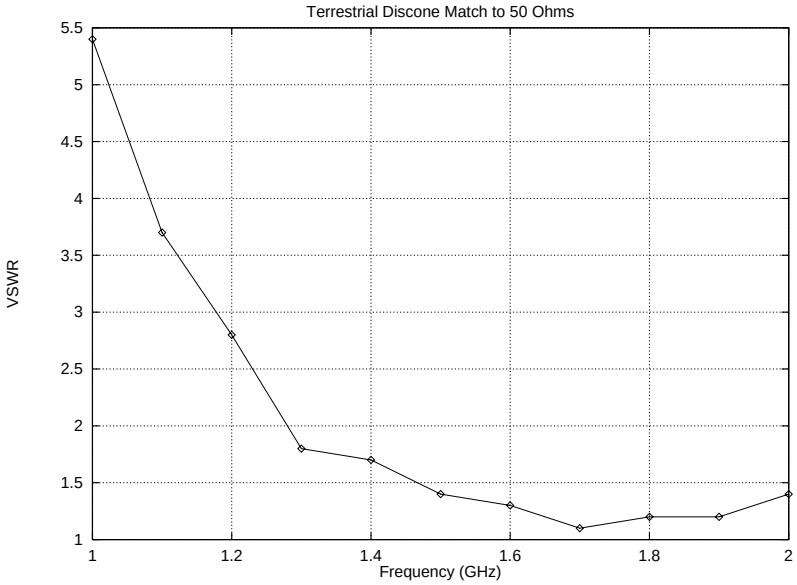


Figure 3.5: The terrestrial discone antenna's match to 50Ω.

cellular phone bands. Figure 3.6 gives a diagram and physical measurements of the discone antenna (and was adapted from a diagram in the *The ARRL Antenna Book* [19]). The antenna was constructed from eight mil brass shim stock and an SMA connector. Optimal parameters for a discone [20] are (typically):

$$S = 0.3\,C_{MIN} \qquad D = 0.7\,C_{MAX} \qquad L_S/C_{MIN} > 22 \qquad \phi = 60^\circ$$

The discone has a high-pass characteristic. Overall scaling is governed by the lowest frequency we desire the antenna to receive. The slant height of the cone, L_S, is approximately equal to $1/4\,\lambda$ at this cutoff frequency. Our design uses a cutoff frequency of 1 GHz and so has $L_S = 7.5$ cm, $L_V = 6.5$ cm, $C_{MAX} = 7.5$ cm and $D = 5.25$ cm. The high frequency response is dictated by S and C_{MIN}. A discone can have a decade of useful bandwidth, but we need less than an octave for our purposes so S and C_{MIN} were not critical. They were dictated by construction techniques and the materials used.

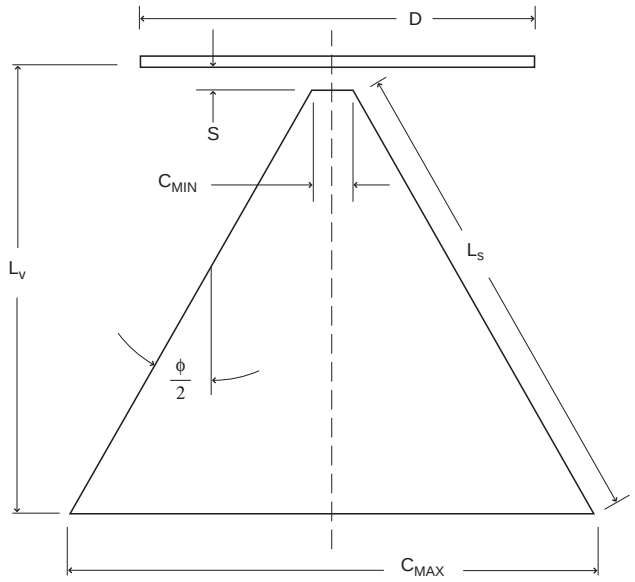


Figure 3.6: Physical measurements of the terrestrial discone antenna.

3.2.2 Telescope Position Control

The radio telescope was originally built in the 1950's and the control panel shows it.[2] Angle information is sent directly from the telescope to the position dials via synchros. The movement motors are controlled with switches. While it is relatively easy to do manually, we needed to be able to position the telescope precisely and safely under computer control (by the telescope control computer, designated TC). We do this by having the computer "read the dials" and "flip the switches" for us.

Reading the Dials

We were able to purchase synchro-to-digital converter cards (ILC Data Device Corporation Model SDC-36016) which allow a computer to read the dials. These are ISA bus cards that plug directly into a PC and enable it to indicate a synchro's angular position to arc-minute precision. The control panel has five dials: three for hour angle (hours, minutes and seconds) and two for declination (degrees and minutes). Since the antenna beam width is only about half a degree we did not use the hour

[2]It looks like something out of a WWII era submarine.

angle/seconds dial. The dial with the larger units is read to get the coarse position of that axis, and the dial with the smaller units is used to make the position precise.

Mechanical backlash in the declination gearing forced us treat that axis somewhat differently. If the telescope is moved northward (or back and forth), the declination/degrees dial may be more than one degree off. The work-around for this problem is to move the telescope southward several degrees before reading the dials. The declination/minutes dial does not suffer from this problem so the computer takes a hybrid approach. When the computer starts up, it moves the telescope southward for ten degrees and then reads the declination/degrees dial. It then remembers this coarse position and uses it (ignoring the declination/degrees dial) until the next power-cycle. The computer then updates the antenna's declination position by monitoring the declination/minutes dial and integrating it over any entire revolutions. The telescope control computer responds quickly and reliably enough so that there is no chance of missing an entire revolution and thus being off by a degree.

The hour angle/hours dial does not have this backlash problem: since each hour is 15 degrees, the dial is always accurate to a fraction of one hour.

Pushing the Buttons

The telescope control computer moves the telescope by pushing the normal control buttons and switches like a human would do. We rewired the control panel to have two modes, *computer* and *manual*, chosen by a switch on the side of the panel. In *manual* mode it works normally. In *computer* mode, the manual switches are disabled and the computer takes over their function via optically-isolated, solid-state relays.

The telescope is very massive and doesn't stop immediately when a switch is released. We calibrated the overshoot and compensate for it in software by releasing the motion switches early. This technique is accurate to 2 or 3 minutes of arc, which is perfectly adequate give the 30+ minute wide beams.

Fail-safe Considerations

The Agassiz Station radio telescope is the one irreplaceable part of *BETA*. Under computer control there is a chance that a system error will move the telescope into the ground, severely damaging it. While there are limit switches to prevent this, they were installed when the telescope was originally built and only 60 feet in diameter. The telescope was upgraded to 84 feet in the early 1970's so now certain combinations of right-ascension and declination will allow the dish to hit the ground.

Even if the limit switches were sufficient for all such combinations, the risks of unattended equipment operation are severe; it requires careful planning to minimize them. The major issues we came up with for this fail-safe design were:

1. The telescope fails safe by stopping, which can be achieved by cutting power to the telescope motors. The cost of stopping the telescope for a false alarm is minimal, while not stopping the telescope for a true problem can be catastrophic.

2. Never trust software. Subtle program bugs, strange machine states or crashed computers can cause unexpected behavior.[3] *BETA* was designed to run continuously for many years so even unlikely circumstances can occur.

3. Use defense in depth. Many different layers of safety system are less likely to fail simultaneously than one single one.

4. Permit the equipment to have only as much capability as necessary when running unattended.

5. Simple systems are less likely to fail than complicated ones.

We came up with a set of safety systems that has done well so far. The first layer is in software: the telescope control computer software will not permit the user to move the telescope beyond certain pre-programmed limits. Any requests to go beyond the

[3]For example, we gathered several months worth of bad data because of a compiler error. *Our* C++ code was correct, but the generated object code was not.

limits will be truncated to the limits. If the software finds the telescope beyond any limit, it will allow motion back into the safe region but not farther into the unsafe one.

Since we don't trust software, the next layer is a hardware watchdog to make sure the software is still running correctly. What happens if the computer commands the telescope to move and then crashes before it can tell it to stop? The watchdog hardware requires that the software signal it periodically to keep operating. If the watchdog does not receive a signal within a preset amount of time ($\approx 1/10$ second), it shuts down the power to the telescope motors. If a bug in the computer software prevents the proper software (with the watchdog signal) from running, or if the computer crashes or if the plug gets kicked out of the wall, the telescope will stop moving immediately.[4] The software signals the watchdog hardware by writing a number to a port. To keep a simple looping bug from continuously writing the same number to the port and defeating the system, the watchdog's "crypto-enable" feature requires a different value be written every time. While just about any algorithmically predictable function would have worked, we used an increment function because it was easy to implement and debug.

If the software and watchdog fail, a set of mercury tilt switches will keep the telescope from moving too far. Under computer control, the telescope needs to move only between $+80^\circ$ and -40° declination and $\pm 1.5^h$ hour-angle, so we used the tilt switches to constrain it to this region. The switches are set up so that if the telescope moves beyond the proper range, the switch opens and power to the motors in that direction ceases. Again, the telescope may still be moved back into the safe range. The tilt switches have no effect in *manual* mode so the telescope's full range of motion is still available.

Finally, if all else fails, the original limit switches will probably be sufficient to prevent damage.

[4]Important watchdog tip: Never signal watchdog hardware from an interrupt routine! If your program has a bug or crashes partly, the interrupt may continue being serviced, merrily keeping the watchdog hardware from doing its job.

Software Services

The telescope control computer (TC) is a PC running a single-threaded, real-time application. It writes pertinent information to the screen and accepts debugging commands from its keyboard, but the main user interface is through a daemon program on the Unix workstation, via the network. This daemon (called *tserv*) is the only program in the system which talks directly to TC. It acts as a broker for other clients, allowing limited access to the telescope control capabilities and preventing conflicting requests. It also has several security features to prevent unauthorized access.[5]

Two other daemons talk to *tserv*: *resyncher* makes sure that the telescope control computer has run its anti-backlash protocol upon rebooting. *frogger* provides a simplified programmatic interface for complete sequences of motion commands.

3.2.3 Down-conversion and Digitization

BETA's RF frontend is a double conversion, superheterodyne design. See Figure 3.7 for the details. Signals from the low-noise amplifiers are mixed down to the first intermediate-frequency (IF) of 40-80 MHz by circuitry in the antenna radome. The first local oscillator (LO), located down in the control room, provides the tuning for this; its frequency specifies which 40 MHz portion of RF will be converted to the IF. Since the tuning needs to change quickly every two seconds (to perform the 8 frequency hops) we use a "direct" synthesizer (a PTS-1000), which has no phase-locked loops with their associated settling times of $\sim$ 100 msec. Instead it uses many PIN diode switches and mixers to synthesize the selected frequency, with a switching/settling time of about 10 microseconds.[6] Since this synthesizer only goes up to 1 GHz, we program it to put out half of the necessary frequency and use a doubler in the mixer circuitry.

[5]These will not be discussed here. Yes, that may be partly "security through obscurity" but we've decided to be professionally paranoid.

[6]Since we sample at about 2 MHz, the first $\sim$ 20 time-domain samples of each spectrum will be corrupted because of the switching transient. The effect will be negligible since this is less than $1/200,000$ of the data and these data points are highly attenuated by the windowing function anyway.

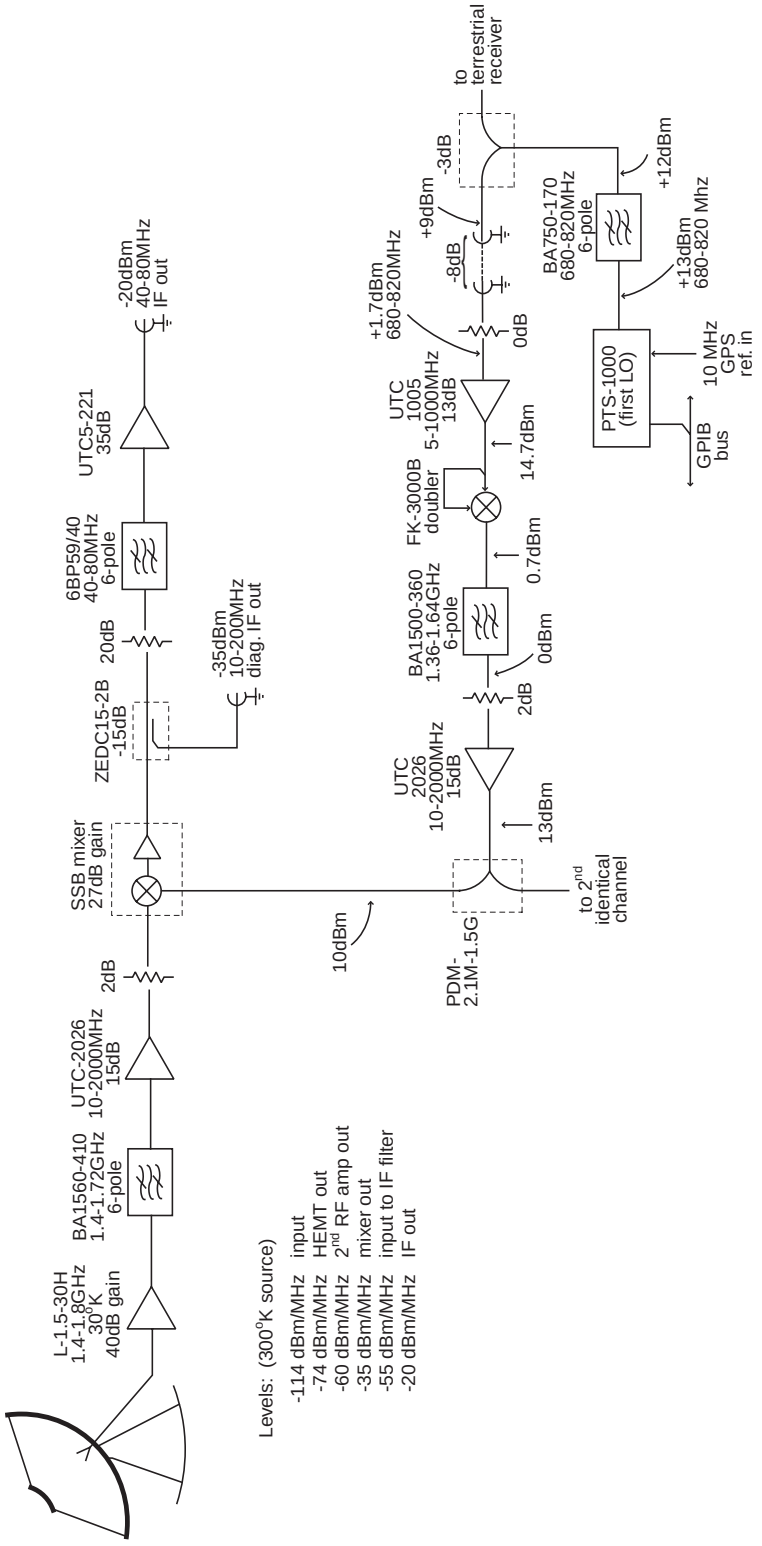


Figure 3.7: Block diagram of the RF front end.

The mixer is of the image-reject variety, which uses quadrature phase shifts to produce only one of the RF sidebands at the IF. (For good explanations of what is called the "phasing method" of generating single sideband see [47] and [23].)

The second conversion mixes the 40 MHz bandwidth IF signal down to baseband. This signal is too wide to be handled by the FFT hardware so we split it into 2 MHz sub-bands and downconvert and digitize each of these separately. We built a local oscillator array which generates 20 separate LO signals from 40 to 78 MHz in 2 MHz steps. They are generated with phase-locked loops from the 10 MHz system reference. We feed the IF signal and LO array signals into a set of 20 mixer/digitizer boards. Each of these downconverts a 2 MHz portion of the IF signal to two baseband signals (in-phase (I) and quadrature (Q)) using quadrature mixers. The I and Q signals are separately digitized using fast 8-bit flash A/D converters and then fed into the real (in-phase) and imaginary (quadrature) inputs of the FFT boards.

3.2.4 Frequency Control and Doppler Compensation

Performing narrow-band spectroscopy at L-band requires some care. In order for the feature-recognition hardware to work correctly the system must make sure that the frequency of a received beacon is mapped near the same FFT bin over the course of several minutes. To provide the frequency stability for this, all oscillators and clocks in the system (up to and including the FFT hardware) are slaved to a single disciplined master clock. This is a Trak Microwave Model 8812 GPS station clock which uses signals from the Global Positioning System satellites to provide accurate time and frequency information. Table 3.2 shows the specified accuracy and stability of the unit installed in the system. Since the FFT bin width is 0.5 Hz (which is broadened roughly a factor of 2 by the Hanning window we used. See Appendix G for details.), we only need a stability of about 3×10^{-10} so the Trak unit is perfectly adequate for our needs.

A much larger problem is Doppler effect due to motion of the earth. We are assuming that the transmitting civilization will correct for any motion on its end so that the signal will appear to have a constant frequency in an inertial frame. We do

Accuracy (while tracking)	2×10^{-10}
1 second Allan variance	3×10^{-12}
10 second Allan variance	5×10^{-12}
100 second Allan variance	1×10^{-11}

Table 3.2: Specified accuracy and stability of the Model 8812 GPS station clock with low noise oscillator option "B6".

Earth motion component	r (m)	ω (rad/s)	$r\omega$ (m/s)	$r\omega^2$ (m/s^2)
Earth rotating about its axis (at 42° N latitude)	4.7×10^6	7.3×10^{-5}	3.4×10^2	2.5×10^{-2}
Earth-moon barycenter orbiting the sun	1.5×10^{11}	2.0×10^{-7}	3.0×10^4	6.0×10^{-4}
Earth orbiting the earth-moon barycenter	4.7×10^6	2.7×10^{-6}	1.2×10^1	3.3×10^{-5}

Table 3.3: Significant earth motions

need to correct for motion on *this* end, however. Table 3.3 shows the three motion components large enough to affect the system. Included in the table are r, the distance from the rotation axis in meters, ω, the angular velocity in radians per second and the magnitudes of the velocity and acceleration components.

Compensating for Velocity

To completely compensate for local velocities, effectively putting the receiver in an inertial frame, we need to accommodate velocity components as small as about $f_{\text{resolution}}/f_{\max} \cdot c \approx 10^{-1}$ m/s. While we can easily calculate these quantities and tune the local oscillator accordingly, we can only do it for one particular direction at a time. If the antenna beam were small or if we knew exactly where the signal was coming from, this would be adequate. However, the main beam system is over 1.5 degrees wide and, since we are doing a drift scan, the signal could be anywhere inside that region. The worst case would be for a source in the plane of the earth's orbit, perpendicular to the earth's motion vector $\vec{v}_\oplus$. As the beam system moves across the source, the angle between its center and $\vec{v}_\oplus$ changes from 90.75° to 89.25°, which

necessitates a change in velocity of $v_\oplus\,(\cos(89.25°) - \cos(90.75°)) \approx 785$ m/sec. This would entail a frequency change of over 4 kHz, despite the fact that the angle between the source and $\vec{v}_\oplus$ has changed negligibly (due only to the change in direction of the earth over those few minutes) and so only a tiny frequency change is needed.

Compensating for Acceleration

META's narrow bandwidth of 400 kHz required it to hop around in frequency to compensate for the doppler shift of three specific reference frames. *BETA*'s frequency coverage (320 MHz in 8 hops of 40 MHz each) is wide enough that we don't have to have to worry about reasonable reference frames. Since we don't need to correct for any specific frame, we can just allow for acceleration and effectively correct for the frame that the observatory is in when a signal first appears.

The real-time PC calculates an ephemeris of the earth, sun and moon and figures out how much acceleration to allow for at any specific moment. These accelerations are reflected in the system's first LO frequency such that a signal coming from a fixed reference frame will always appear at the same frequency in the system. The corrections are not perfect, but over the length of an observation (the drift time through the beam system, about 6 minutes) the error is small.

Doppler Skew

A problem with these doppler shift compensation techniques is that they are correct for one frequency only. The shifts are calculated for the middle of each 40 MHz sub-band, so they are somewhat off at the band edges, 20 MHz on either side. For example, if we were compensating the 1680-1720 MHz sub-band for the earth's velocity v around the sun then all the frequencies there would be shifted by $\frac{v}{c} \cdot f_c \approx 170$kHz. However, the upper edge should have been shifted by $\frac{v}{c} \cdot f_u \approx 172$kHz, so there is 2 kHz of skew.

The acceleration compensation problem is similar. Compensating for the observatory's (maximum) acceleration of $a \approx 0.026$ m/sec, we see that the shift will be $\frac{a}{c} \cdot f_c \approx 0.1473$ Hz/sec. The upper edge's shift should be $\frac{a}{c} \cdot f_u \approx 0.1491$ Hz/sec, so

there is ≈ 0.00173 Hz/sec of skew. Will this amount be a problem? Over a 10 minute observation (longer than the sidereal drift time through the beam system) a signal at the band edge will drift by about 1 Hz. This is nearly BETA's frequency resolution if you include window broadening. Also, the system preserves a swath of ~ 10 Hz while accumulating a "slot", so this amount of drift does not present a problem for us.

A doppler shift can be removed *exactly*, without skew, in the time domain by directly sampling a signal at intervals which are corrected to the transmitter's frame. These intervals will be non-uniform in the accelerating receiver frame (which has a velocity $v(t)$ with respect to the inertial frame). For example, a signal transmitted with frequency F will be received with frequency $F' = F(1 - v(t)/c)$, so sampling non-uniformly with $t' = \dfrac{t}{1 - v(t)/c}$ will produce data with no doppler shift or skew.

However, direct sampling is difficult at high frequencies so most receivers use superheterodyne techniques to lower them. The sampling method mentioned above will not work if the frequency has been shifted, as in a superheterodyne system, so a better method is needed. Lou Scheffer [46] came up with an interesting scheme for exact doppler shift removal over wide bandwidths in superheterodyne receivers. Instead of the usual local oscillator correction, $F'_{\mathrm{LO}} = F_{\mathrm{LO}} - F_c(v(t)/c)$, we correct the LO for its own frequency $F'_{\mathrm{LO}} = F_{\mathrm{LO}}(1 - v(t)/c)$. This makes the shift proportional to the IF frequency:

$$F(1 - v(t)/c) - F_{\mathrm{LO}}(1 - v(t)/c) = (F - F_{\mathrm{LO}})(1 - v(t)/c) = F_{\mathrm{IF}}(1 - v(t)/c)$$

so we can sample non-uniformly at the IF using $t' = \dfrac{t}{1 - v(t)/c}$ completely eliminating any skew.

3.3 FFT Hardware

The heart of *BETA* is the 250-million channel fast Fourier transform spectrometer. It is composed of 63 individual boards, each of which can compute a 2^{22}-point complex FFT every two seconds. These boards are based on the Austek A41102 Frequency

Domain Processor. This integrated circuit can continuously compute 256-point complex FFTs in 102.4 microseconds each. A million of these ICs could compute our 250-million channel spectrum, but this would be overly complicated and wasteful. Since we don't need the spectra every 100μsecs (2 seconds will do), we can use fewer processors to compute the transform more slowly. Appendix C describes how a long DFT can be computed as a series of smaller ones. Figure 3.8 is a "bottle" diagram which shows schematically how this is done by our FFT boards and Figure 3.9 shows a block diagram of the actual circuitry. First the 4M-point FFT board multiplies the time-domain data by a Hanning window (see Appendix G for details). Then the board computes the 2^{22}-point transform as a $2^{14} \times 2^8$ 2-dimensional transform with corner turns and complex twiddle factor multiplications. A corner turn involves using a piece of memory to bit reverse the order of a sequence of data. The bit reverse issue is inherent in the fast Fourier transform algorithm and is handled by writing data into the memory in one order and reading it out in the other. The 2^{14}-point transform itself is implemented as $2^7 \times 2^7$ points. The entire 4M-point transform uses three pipelined Austek chips, one handling each of the smaller transforms. Multipliers built into the chips are used for the window function and twiddle factors (whose coefficients are stored in ROMs).

Real (in-phase) and imaginary (quadrature) time-domain data (digitized as signed, 8-bit integers) enters the 4M-board in normal order, is processed, and exits six seconds later (due to pipeline delays) as 16-bit spectral magnitudes in normal order. The transform is computed with 20-bit integer arithmetic, which turns out to be quite adequate for such a large transform as long as right shifts are performed at the butterflies (see Appendix F). The final magnitude computation, $M = \sqrt{I^2 + Q^2}$, is performed quickly with a look-up table, because doing it by brute force would have been expensive. However, in order to keep the table small enough to implement, we had to devise a non-uniform scalar quantization technique which would compress the I and Q values with as small a fractional error as possible. This technique, which works quite well, is described in Appendix E.

Managing such a large array of fast hardware requires some housekeeping tasks.

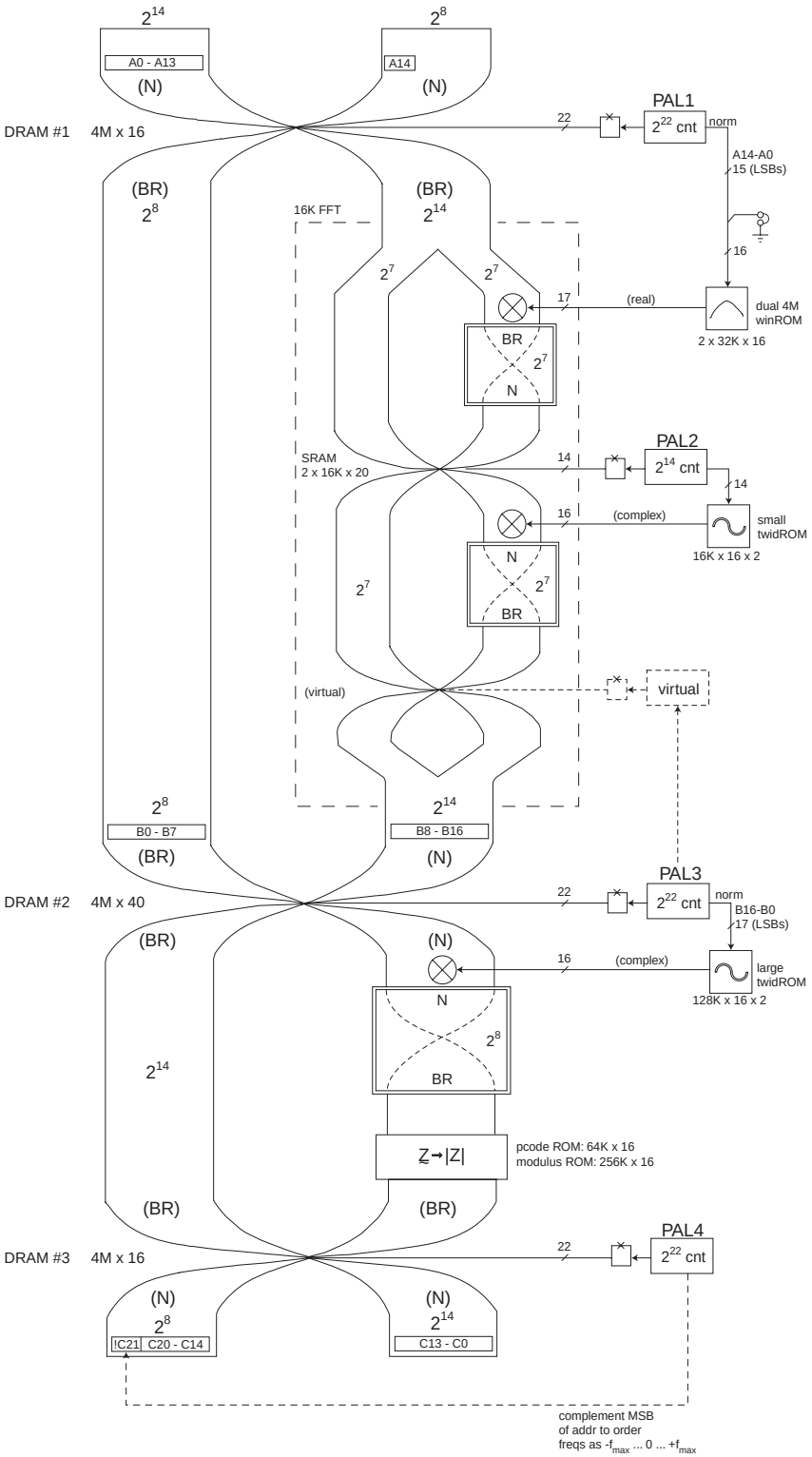


Figure 3.8: FFT "bottle" diagram showing how the large FFT is computed using smaller FFTs and twiddle factors.

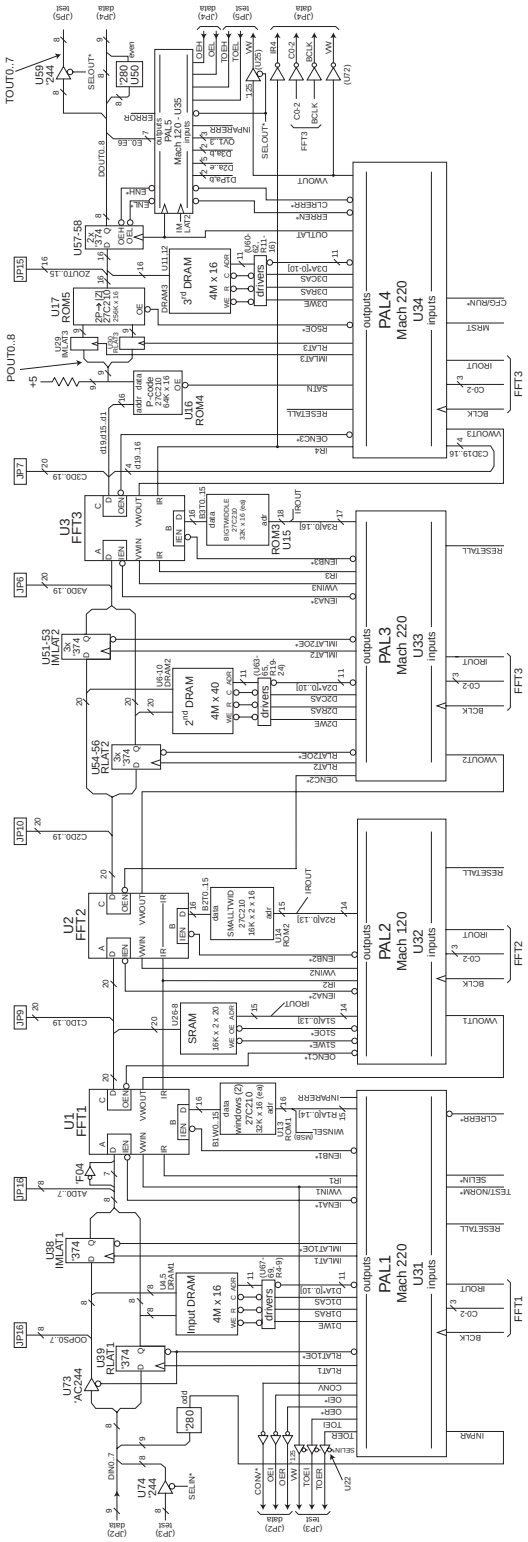


Figure 3.9: Block diagram of the 4M-point FFT computation board.

All 63 boards must run in lock-step from a set of fast clocks, and all require downloaded parameters from the main system. It would also be very useful to be able to detect bad boards that are generating incorrect spectra. Figure 3.10 shows how this is implemented in *BETA*. Clocks and and synchronization signals are distributed with fast ECL circuitry. A differential bus connects the FFT array with the real-time control computer (described in section 3.5.3), allowing the Austek chip control registers to be set and changed at run time.

Detecting incorrect spectra is accomplished by feeding pairs of FFT boards the same input vector and comparing their outputs. Any discrepancy, no matter how small, indicates a flaw in one of the boards. In order to identify which of the boards is bad, we have divided the boards into three groups, A, B and C, and then run two tests on each triplet: A vs. B and A vs. C. The results of these tests (described in Table 3.4) indicate which board has the problem. Since board problems tend to be independent and uncommon, if both tests fail, then A is probably the sole cause. However, this is not guaranteed.

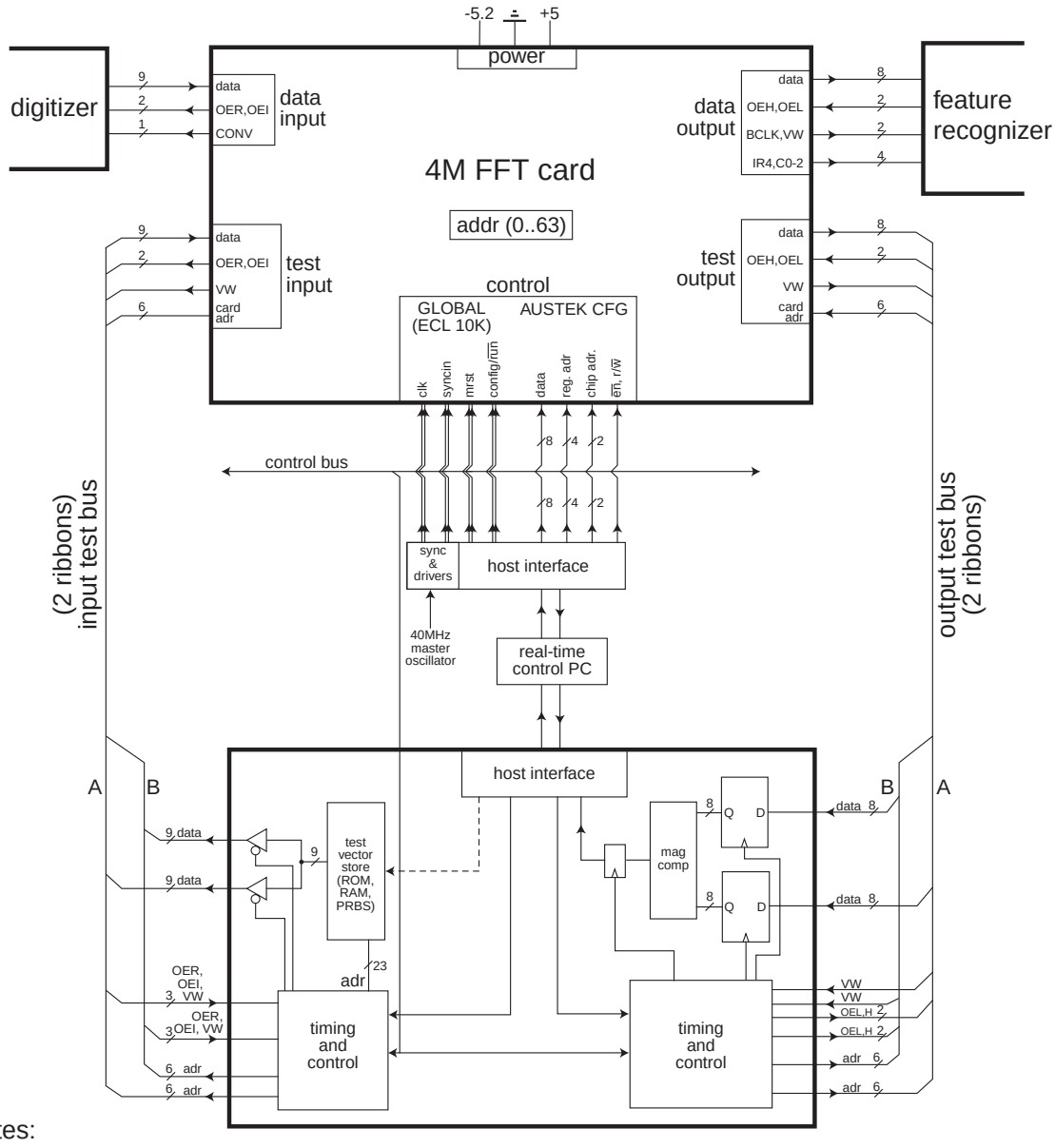
A vs. B	A vs. C	Interpretation
passed	passed	All three boards are good.
passed	failed	A and B are good. C is bad.
failed	passed	A and C are good. B is bad.
failed	failed	B and C are good. A is bad.

Table 3.4: Interpreting the results of FFT triplet tests.

3.4 Feature Recognition Hardware

The spectrometer output data rate is $\approx$ 250MBytes/sec or about a CD-ROM every two seconds.[7] It is currently infeasible to store this amount of data in order to process it at a later time, so we must analyze it all "on the fly". Most of the data must be thrown away. *BETA*'s data analysis algorithms attempt to prove that a piece of data

[7]Which amounts to a terabyte/hour, or almost 10^{16} bytes/year!



Notes:

1. 40 MHz clk, & SYNCHIN* "on-the-fly" distributed via ECL 10124/10125 bussed pairs; SYNCIN* carefully timed.
2. Austek upload from host; non-existant chip number is a register (20V8), to hold test/norm*, clear, etc. states.
3. Austek upload & extra state bits set in "config"; transition to "run" lets PALs begin counting.
4. All FFT bds are in lockstep; control-bus driver cktry "knows" C-states & BCLK timing.
5. IN "test" mode, normal input/output disabled. Identical data is applied to a chosen pair (1 on A, 1 on B) of FFT bds, & the outputs compared (after latency -- can be made "seamless").

Figure 3.10: Control and verification of the FFT array. Master distribution of clock and synchronization signals puts the array in lock-step. The addressable test ports are used for routine verification of FFT board function, demanding identical spectra for identical inputs.

is of terrestrial origin; if so, it is discarded. Data that survives the tests is archived and analyzed further.

Figure 3.11 is a block diagram of the system's feature recognition hardware. Data from one of the 4M FFT boards is received by a Feature Recognizer (FR) board, which computes a running average and compares each data point to that average using several thresholds. The data points, average and threshold results are then reported to a Feature Correlator (FC) board which handles "slots", "notches" and communication with the host pentium PC. Each FC receives the input from three FR boards: one carrying signals from the east beam, one from the west and one from the terrestrial antenna. The FC compares the results from the three data streams and takes various actions depending on its programming. It can then hand the results to the host PC which does further processing. The PCs in the pentium array handle communication between the FCs and the Unix workstation. They also implement the adaptive "mini-notch" processing (described in section 3.5.1), track the time history of slots, handle much of the frequency hopping work and perform other tasks necessary to keep the system working in synchronization. The time history of interesting data points is then forwarded on to the workstation for more analysis and archiving.

3.4.1 The Feature Recognizer Hardware

The primary function of the FR boards is to tag spikes in the frequency data. They do this by comparing each data point to a threshold-scaled version of the running average. Points above the threshold are marked as "hits" and scheduled for later processing; the rest are (usually) ignored. A block diagram of the FR is in Figure 3.12. The FR computes a 4K-point running baseline by summing the squares of the magnitude data points received from the FFT. This is done in an efficient, pipelined manner by adding $x^2_{t-2\mathrm{K}}$ to a running sum and subtracting $x^2_{t+2\mathrm{K}}$ from it. Since the length of the average is a power of two, dividing to get the mean can be done trivially with a shift operation. The data point x_t at the center of this (boxcar averaged) "moving" baseline is forwarded along with the baseline value.

One complication forced us to modify the FR boards after the system was fielded:

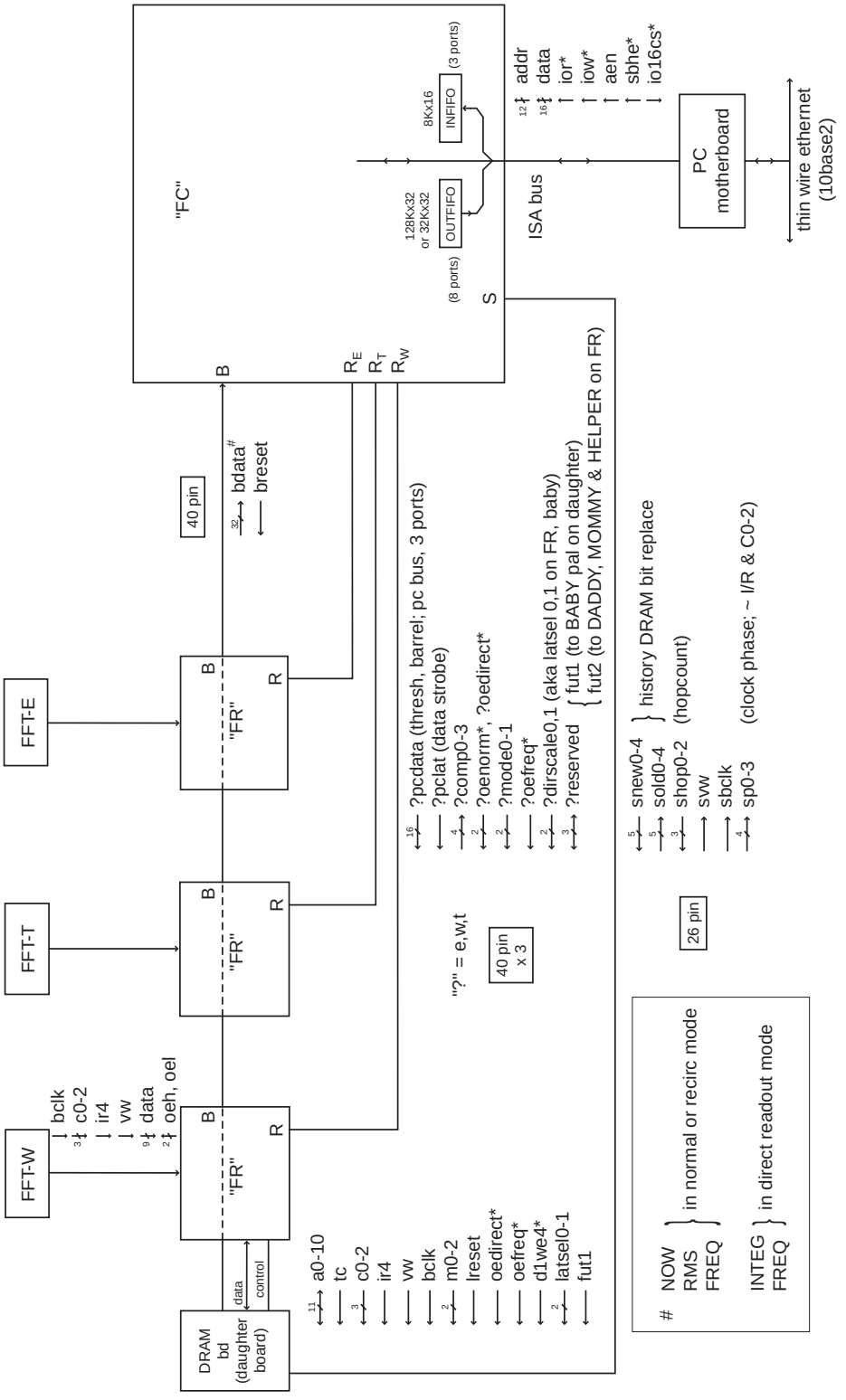


Figure 3.11: Block diagram of the backend hardware showing the connection between the feature recognizers and the feature correlator.

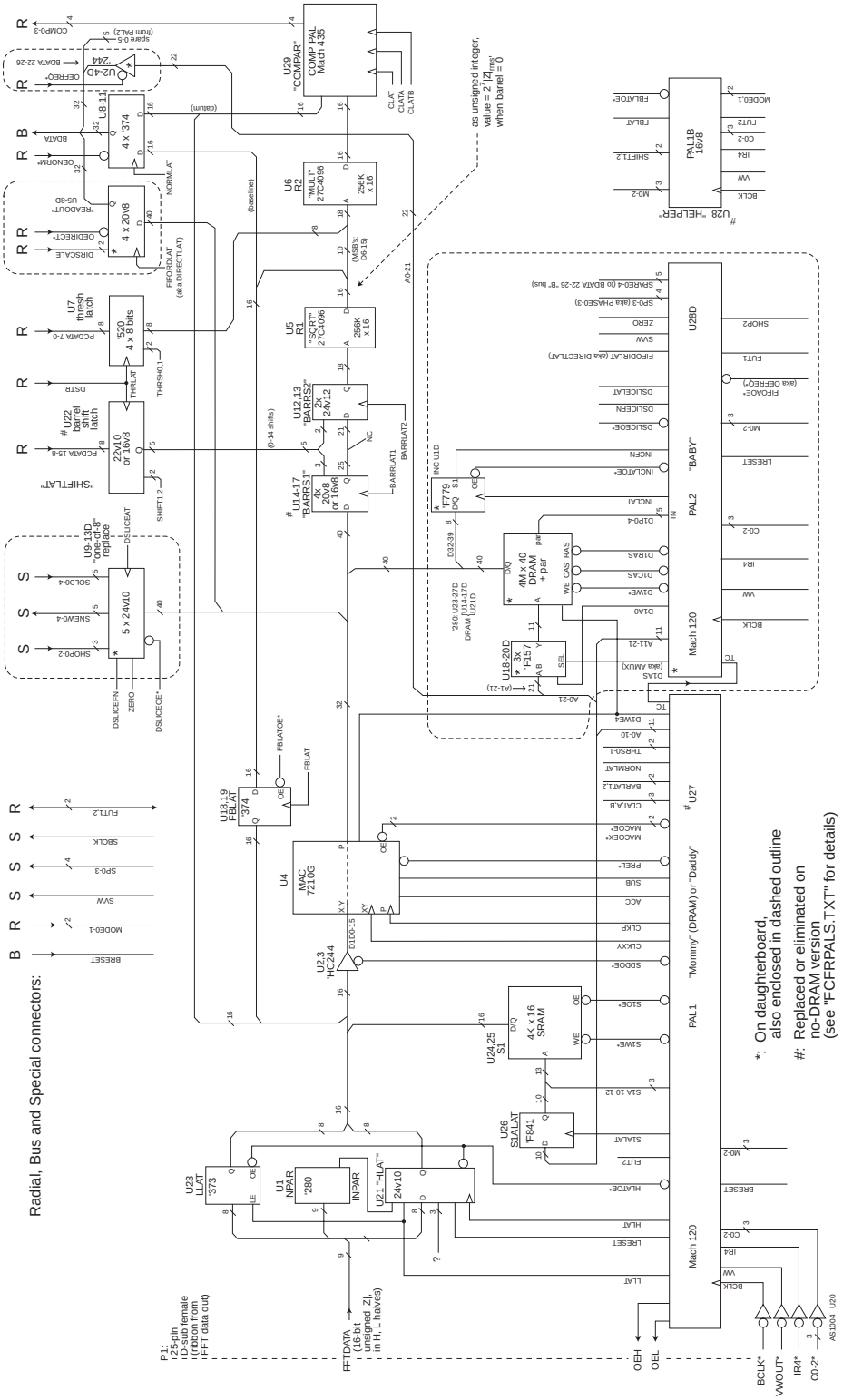


Figure 3.12: Block diagram of the feature recognizer.

because we were summing the squares of the data points, a single large RFI spike would dominate the sum, substantially raising the baseline. This abnormally large baseline would make the system far less sensitive for the surrounding 4K points. To remedy this situation, we modified the FR boards so that only the low-order 8 bits of each data point contribute to the baseline; the high-order bits are ignored. This means that strong signals do not count for as much in the baseline, while the lower level "baseline noise" signals pass through unchanged.

After the power sum is computed, it is barrel shifted, converted back to magnitude, and multiplied by four programmed threshold values.[8] Both the amount of the barrel shift and the value of of the thresholds are programmable on the fly by the host PC. Every element of the magnitude spectrum is compared with all four baseline-times-threshold values.

The largest threshold represents a signal that is well above the noise baseline, while lower thresholds can be used to track signals of smaller intensity without necessarily forwarding them to the host PC, as will be discussed in the next section. Large signals are usually reported as "hits".

The FRs also have an integration mode, in which power spectra can be summed and stored. This integrated power can then be read out directly to the host, via the FCs. Because of the slow PC bus, an entire spectrum requires many seconds to fetch, and pieces of it would be lost. The hardware therefore allows the PC to request only certain portions of the spectrum (using the "slot" mechanism described below). In addition, the integrated spectrum can be recirculated through the FR signal-recognition circuitry, just as an ordinary spectrum, allowing the same threshold analysis to be performed on the integrated spectrum. The ability to read out the full spectrum can be useful in a sidereal tracking reobservation of many minutes' duration. Only one of each group of three FRs, the "mommy", contains the RAM required to perform integrations. The RAM is contain on a daughterboard called "baby". The

[8] Actually the multiplication table is stored in ROM, so arbitrary two variable functions can be programmed. The square root function is also encoded in a ROM, and can therefore be an arbitrary one variable function.

non-RAM versions of the FR are called "daddy".

3.4.2 The Feature Correlator Hardware

The feature correlator forwards interesting frequencies to the host PC. A block diagram is included as Figure 3.13. This data includes the baselines and signal magnitudes for each of the three horns as well as the calculated threshold results and the bin number. The PC has the ability to set certain frequencies to be automatically forwarded, indicated by a start and stop frequency. These are called slots. When a good hit is detected, the PC may begin a slot, meaning that for a given length of time (presumably while the source is still in one of the beams) the PC requests that all the data for a frequency range around the hit frequency be automatically reported to it. The opposite of a slot is a "notch", which specifies that all the data from a given range of frequencies should be ignored, no matter how strong the hits there might be. Notches are important for eliminating fixed-frequency RFI. The bandwidth of the FC/PC interface is limited and an excess number of hits due to interference could cause the loss of potentially interesting data.

This limited bandwidth exposes a significant flaw in the FC design. All of the data from the FC to the host PC, including hits, slots and synchronization information, travels through the same finite sized FIFO. If the FIFO overflows due to a large number of hits, the slot and synchronization information will probably be lost. Losing synchronization is serious enough to warrant restarting the affected computer. We designed the FC circuitry before we were aware of the severe RFI environment that we would be working in, and so did not realize the full extent of the problem. After some experience with running the system, we noticed that many of the computers were restarting frequently. Putting fixed-frequency notches on the worst areas slowed the problem down but did not eliminate it. We then made some simple hardware modifications to the FC to allow it to tag hits and slots. Without the tagging, the host PC had to figure out which was which by checking to see whether it had asked for a slot at that frequency previously. While this scheme worked, it took too long. Tagging the output data allowed the PC to rapidly discard excess hits with very little

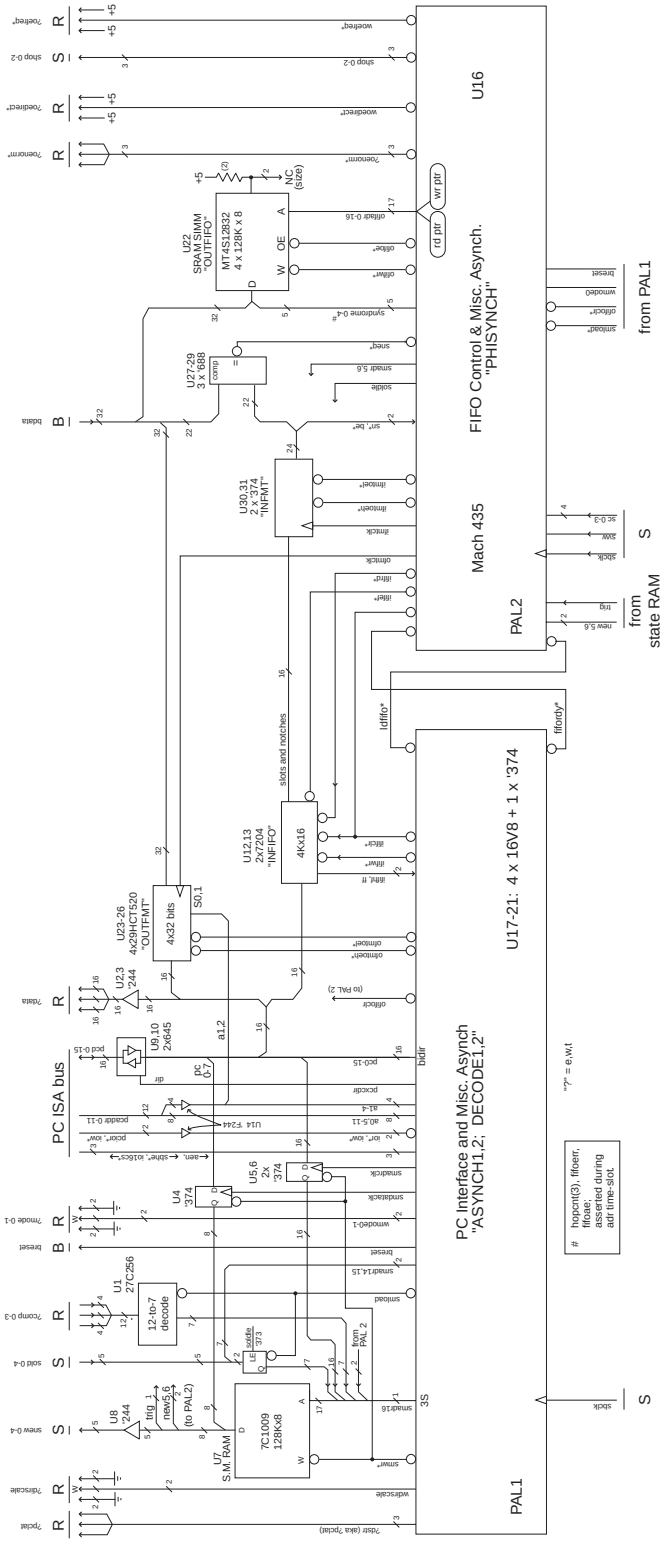


Figure 3.13: Block diagram of the feature correlator.

computational overhead.

This modification works reasonably well but does not eliminate the problem; the PCs covering regions with severe, but periodic RFI will restart several times per day. The restart is automatic, but some observing time at those frequencies is lost. A better solution would be to limit, in hardware, the number of hits that can be generated per spectrum so that infrequent bursts of RFI will cause as little damage as possible. Designing such capability into the hardware is simple, but it was infeasible to modify the existing FC boards to such an extent. A second revision would have such capabilities. The more recent SERENDIP search hardware has this capability. [54]

The FC State Machine

If a frequency is neither a slot or a notch, a programmable state machine (SM) determines whether it is interesting enough to forward to the host PC anyway. The SM is implemented with static RAM whose contents are configurable from the PC. Such reprogramming is normally only done during PC initialization since it can take a substantial amount of time. However, the SM can actually contain eight separate sub-programs between which the PC can instantly select via a configuration register. This is useful for implementing several modes of operation. For example, SM program number 0 is designed to produce no hits and is used during system warmup.

The SM receives a 7-bit quantity describing how many thresholds were exceeded by the signals in the three beams (4 thresholds $\longrightarrow$ 5 states per FR, three of which gives $5^3 = 125$ total states). It also receives a 5-bit quantity ("time based state") that it stored in the FR RAM the last time a particular frequency was visited (16 seconds previously), and a 2-bit quantity ("the frequency based state") that it sent to itself from the previous frequency bin (0.5 microseconds previously). Based on this information, the SM generates 5 bits of new time based state to be stored and retrieved the next time that frequency appears, 2 bits of frequency based state to send to the next frequency bin, and a bit specifying whether the data of this frequency should be forwarded to the PC.

The SM allows great flexibility in making intelligent decisions at speeds with which

the PC cannot compete. For example, the frequency based state can be used to detect a signal that happens to lie between two bins, so that it does not exceed the highest threshold, but exceeds the second highest threshold for two consecutive frequency bins. The time based state can be used to monitor a signal that is crawling into the edge of a beam, but has not yet arrived at the center and so has not exceeded a high threshold, but has repeatedly exceeded lower thresholds. The PC can be sent this frequency bin with the understanding that it may wish to follow it so that if it does turn out to eventually exceed a threshold, less data will have been lost.

Since the SM sees the comparison results from all three horns, it can perform rejection of terrestrial interference and of signals that are simultaneously high in both the east and west lobes, which may indicate RFI that has not been detected in the terrestrial (discone) antenna.

We have developed an intuitive language for generating SM programs in an easy to understand format. Called DSML[9], the language allows an SM programmer to specify simple rules for only the states and cases in which he or she is interested. The DSML interpreter handles the remaining cases and all of the ugly details of the state machine itself.

A DSML program is a sequence of rules which the interpreter tries to match in order. Each rule is of the form:

```
[ EastThresh; WestThresh; TerrThresh; TimeBasedState;
  FreqBasedState; ProgramNumber ]
-> [ NewTimeBasedState; NewFreqBasedState; hit? ] ;
```

The first part in brackets specifies the inputs which will trigger this rule. The programmer can use specific values, ranges of values or a period which means "any value". The second part in brackets (after the `->` arrow) specifies the outputs for this rule. The programmer can again use specific values or simple functions. The syntax is similar to C/C++. Below is a simple DSML program which we will explain line-by-line.

[9]For **D**errick's **S**tate **M**achine **L**anguage, pronounced *"dismal"*.

```
// simple DSML program
[.; .; .; .; .; 0] -> [0; 0; 0] ;
[0-4; 0-4; 0-4; .; .; .] -> [0; 0; ((e==4) && (t==0)) ] ;
[ .; .; .;   .; .; . ] -> [0; 0; 0];   // catch all
```

The first line is a comment. The second indicates that, no matter what the thresholds or previous state bits are, if the program number is zero, then all of the outputs should be zero, i.e. clear the state bits and generate *no* hits. This line will match every possible state with a program number of zero. The second line indicates that for all of the thresholds in the range 0 through 4, ignoring the state bits and in any program number except zero (which was matched earlier), then trigger a hit if the east value exceeds threshold 4 *and* the terrestrial value exceeds none of the thresholds. Simply, trigger a hit if there is a strong east signal and no strong terrestrial signal. This implements the terrestrial veto function. The last line is a catch all, which matches any remaining rules and specifies their output as zero (no hits, clear the state bits).

More complicated programs can be written to perform other tasks such as simple high-speed signal analysis or various kinds of testing.

3.4.3 Pentium Array Computers

The FR and FC boards plug into the ISA bus of an array of 60 MHz Pentium PCs. Each is diskless and equipped with 32MBytes of DRAM, an Ethernet card, a monochrome video card (for debugging purposes) and a "wart" board.

The Wart Board

The wart board performs various control and monitoring tasks for the Pentium PCs. All the wart boards in the system are bused together with the real-time control computer, which can address them individually. The boards allow the PCs to be rebooted either individually or all at once and to have their internal temperatures and power supply voltages read. They also perform a transfer function for the Ethernet card LEDs: wires from these go to the wart board which amplifies the signals and

sends them to unused LED indicators on the computer cases. The yellow "TURBO" light blinks when an Ethernet card receives a packet and the red "HDD" light blinks when the card transmits a packet. This is very useful for debugging purposes, and visitors love looking at the rippling array of colored lights.

3.5 Feature Recognition and System Software

3.5.1 Pentium Array Software

Each FC talks with one Pentium-based PC motherboard. The PC tells the FC which slots and notches it should track, collects the results and forwards them to the workstation. The PCs also perform the adaptive notching function (see below) which keeps the RFI problem manageable. The PC software is event-driven and handles I/O from the Ethernet card, the FC and various PC timing components.

Because of the real-time nature of these tasks and the amount of data they need to manipulate, we could not use any of the higher-level operating systems available at the time. The PC software runs under extended DOS[10] which means there is no operating system to get in the way and no arbitrary 64K/640K memory barriers. The software was written in C++[11] and some assembler (for the interrupt routines).

Networking

Each PC communicates with the Unix workstation over an Ethernet LAN. Ethernet uses a "carrier-sense multiple access/collision detection" (CSMA/CD) protocol for sharing the bandwidth between its users. If two computers try to transmit simultaneously, there is a "collision" and both retry the operation a random amount of time later. With 21 PCs running synchronized programs, all needing to talk to one host computer over the same network cable, there will be major collision problems. This

[10]Using the Pharlap 32-bit DOS extender.
[11]Using the Metaware 32-bit C/C++ compiler for extended DOS.

is one reason we did not use TCP/IP protocols on top of Ethernet.[12] We instead designed our own simple protocol and implemented it directly with raw Ethernet packets. Our original intention was to have the Unix workstation mediate all transfers so that there would be no collisions at all. This turned out to be too cumbersome and slow for our application and had to be abandoned. Another problem was the small input buffers of the network hardware which dropped packets if they arrived too quickly. Most protocols (TCP, etc.) use a "sliding window" protocol [6] to avoid this, but we did not want to implement something so complicated. Our protocol has two layers:

- The lower layer transfers large blocks of data (up to 64KB) between two computers, hiding the Ethernet details, correcting for dropped and out-of-order packets, etc. This layer also serializes the blocks and ensures that they all arrive, in order. The Unix side code transmits packets to the PCs in round-robin fashion, with a guaranteed minimum delay between successive packets to the same PC, to allow for the small buffer sizes. This is just as efficient as sliding windows when sending messages to a large number of PCs at once. When the PCs report their data to the workstation, they do so after a PC-dependent delay. This keeps them from clogging the the network by trying to write to it at once.

- The higher layer uses the blocks of the lower layer to transfer specific, formatted data. The formats are hierarchically structured and extensible.

All data is transfered over the network in big-endian byte order. This order was chosen for efficiency since there is more processing power in the (little-endian) Pentium array than in the (big-endian) Sun Sparc workstation.

[12] Another reason is that we did not have any PC routines for such that we trusted in a real-time environment.

Adaptive Notching

Radio frequency interference is still the largest problem that we face. While *META* operated entirely within a protected radioastronomy band (and *still* had interference problems), *BETA*'s wider frequency coverage requires that we share spectrum with "local" transmitters. The band from 1427-1720 MHz is highly occupied [29] and is used in satellite and aeronautical applications as well as by the military. It also includes strong harmonics from UHF television transmitters and cellular telephone base stations.

Interference from ground-based sources is reasonably straightforward to deal with. It has no intrinsic doppler shift and can be vetoed by the terrestrial antenna. Interference from satellites is a much more difficult problem. In our band of interest, a satellite in low earth orbit (LEO) can present a doppler shift that changes by tens of kilohertz in minutes. Some orbiting transmitters are powerful enough to be received by small hand-held receivers (e.g. the Global Positioning System) and so can be easily detected by the terrestrial feed and the many side-lobes of the main antenna. Interference also appears and disappears as the satellites move above and below the local horizon.

One aggressive solution is to use wide, fixed notches to mask out the offending frequency bands. Some problem bands have thick, more-or-less constant interference and this is the most satisfactory solution. However, most of the interference is intermittent and scattered so using fixed notches would remove a significant portion of the system's frequency coverage. A better solution is to use an adaptive notching technique which will mask out smaller frequency bands, and do so only when interference is actually present. We can take advantage of the fact that interference is correlated both in time and frequency:

- If there is interference at a certain frequency at a certain time, there will probably be interference at that frequency shortly before and after that time.
- If there is interference at a certain frequency at a certain time, there will probably be interference at nearby frequencies at that time.

We exploit these correlations by breaking the FFT spectrum into small blocks, counting the number of hits we get in each block, and following the evolution of this number over time. We chose to break each 2^{22}-point (2 MHz) spectrum into 2^{12}-point (2 kHz) blocks because this is typically the amount of drift that a crystal controlled L-band oscillator (stability of $\sim 10^{-6}$) will present. A counter (initially zero) is associated with each block. If the number of hits in a block is greater than zero, its counter is incremented by 5. If there are no hits, the counter is decremented by 1. If the counter reaches 25, then that block is declared to be "notched" and no slots are generated in its frequency range. If the counter reaches zero, then the block becomes "un-notched" and hit generation is allowed again. Interference must be continuously present for 80 seconds to trigger the mini-notch, and then must continuously disappear for at least 400 to un-notch it. The counters have a maximum value beyond which they are not allowed to go; this allows a block to un-notch relatively quickly after a long period of interference.

Since the mini-notches are meant to deal with interference (which will not show the sidereal doppler chirp), the current doppler offset (provided by the real-time control computer) is subtracted from the each hit frequency before it is processed.

Tests in our RFI environment have demonstrated that this straightforward adaptive "mini-notching" scheme eliminates more than 99% of the interference while blocking less than 1% of the spectrum overall. Regions of heavy interference still need to be explicitly notched, however, as they generate too many hits and overload the ISA bus of the motherboards. For example, we have permanently notched the GPS and Glonass frequencies. Motorola's new *Iridium* system is now starting to be deployed; while we haven't seen too much interference from it yet, we expect that we soon will.

Collection and Submission of "Slots"

If a hit survives the mini-notches, the PC can turn it into a "slot". This means that about a 10Hz region of frequency around the hit will be followed, i.e. *all* of the data (regardless of hit-status, notch-status, amplitude, etc.) will be stored and analyzed. This region is followed for about six minutes, the time it would take for the source

to transit both beams, and then all of the data is forwarded to the workstation for further handling.

Won't the adaptive notching scheme eliminate extraterrestrial signals too? We have been asked this question often enough that we feel it deserves explanation here. The scheme will only ignore hits from regions that contain recurrent interference. If an extraterrestrial signal happens to be at one of those frequencies then, yes, it will be eliminated too. This is why it is important that less than 1% of the spectrum is adaptively notched at any time.

If an extraterrestrial signal enters the beam, it will suddenly appear as a hit and the software will immediately generate a slot for it. Slotted pieces of spectrum do not affect the adaptive notching mechanism so no mini-notch will be generated.

Then how do mini-notches work if slots keep getting generated? First, when the system is initially started, no slots are generated during a warm-up period of about ten minutes[13] so the adaptive notching mechanism has a chance to get established and "settle". Even after this, only a small number of slots may be generated per spectrum (usually two) and, since interference tends to come in groups, nearby spikes will trigger the notch.

3.5.2 Unix Workstation Software

Completed slots are sent to the workstation for analysis and archiving.

Data Analysis Algorithms

Unlike thermal noise, whose statistics we know, radio frequency interference is much harder to analyze. It comes from a variety of different sources with varied, non-stationary statistics and so cannot be handled with the standard techniques. It is possible that, with a long period of observation and data taking, we could come up with some sort of statistical model for the RFI at our location, but it is unlikely that

[13] If The Signal comes during this time, we're sunk.

this would be useful. Since the RFI situation changes on timescales from seconds to decades, the statistics would become obsolete as soon as the data was taken.

Since we assumed several characteristics of an extraterrestrial signal when we designed the search, we can compare the data to a model of these characteristics and see how close it is. We devised a number of tests to do this. None of the tests is sufficient by itself, since they tend to have high false-alarm rates, but the results of several tests can do a pretty good job. Please don't be upset that the tests have cute names; they are just mnemonics that we came up with.

PaulH Test This is a simple first-cut test that any signal with our assumed characteristics must pass. It only examines the first four data points (starting with the hit that triggered the slot) and the four data points which are delayed from those by the beam separation time. It ignores everything else. The test checks that the initial four points have greater amplitude in east than in west and terrestrial and that the subsequent points have greater amplitude in west than east or terrestrial. Any signal that does not meet these criteria doesn't fit our assumption of a sidereally stationary point source.[14] The figure of merit is proportional to the ratios of east to west and vice versa. A higher figure is not necessarily better than a lower one (it *is* indicative of strong signals), but a very low figure means that the signal has failed the test.

Derrick1 Test This test fits a Gaussian to the east and west horn data and a straight line to the terrestrial horn data. The figure of merit in this test is the inverse of chi-squared normalized to the level of the signals.

Derrick2 Test This test computes a correlation between the east and west horn data and the expected beam shape. One problem with this is that the expected beam shape is somewhat ambiguous because the source does not necessarily pass dead center through the beam. To get this figure of merit we find the area under the

[14] A sidereal source with large amplitude fluctuations or mixed with RFI could fail this test, but we'd be unlikely to recognize it no matter what tests we performed.

largest peaks and normalize by the overall signal size. We also decrease the figure for high terrestrial signal levels.

DarrenD Test We often get slot data that is "spiky": large valued bins stand alone or alternate with small ones. We'd prefer to see data that has some "shape" to it; that has some continuity and looks something like the beams. This (temporally) differential test is sensitive to signals with "shape" and insensitive to ones that oscillate. The difference between two successive data points is compared to the previous difference. If their signs are the same, the magnitude of the current difference is added to the accumulating figure of merit. Otherwise it is subtracted. At the end, the figure of merit is normalized to the signal levels. The east and west figures (which should have "shape") are added and twice the terrestrial figure (which should be random and therefore *not* have "shape") is subtracted.

Staelin Test Prof. David Staelin of MIT suggested a unique differential test that is (more-or-less) immune to amplitude fluctuations of the incoming signal. A plot of the function $\frac{(E-W)}{(E+W)}$ versus time will show a characteristic pattern that is (in the noiseless case) a function only of the antenna's gain pattern. The signal amplitudes are irrelevant since the function is normalized to them. If noise is added to the signals (and T_N will be), then the pattern is a function of it too. If the signals are much stronger than the noise then the problem is negligible, otherwise it must be accounted for.

KS Test This is a Kolmogorov-Smirnov test to see how close the signals' statistics are to exponentially distributed noise. This test is not very useful since RFI is more of a problem than thermal noise, and since individual slots have too few data points to make the comparison valid. We currently don't give it much weight.

Eyeball Test The human mind excels at pattern matching, so we archive and look at all data that passes the minimal PaulH test. This cannot be done in real time

and so the other tests need to be sufficiently accurate to properly trigger the followup modes.

Leapfrogging and Followup Modes

If the candidate signal passes a suitable battery of tests, the system does even more automatic followup by moving the antenna west (leapfrogging) and "inviting the signal to perform an encore". After a particularly good slot, software will place the system in *followup mode*: the Pentium computer that produced the slot begins to generate special *followup slots* or *fslots*. These consist of all of the frequency bins within about a kilohertz of the original hit frequency. The antenna hour-angle is then moved west in 15 minute increments every 15 minutes for an hour and a half. This will give 6 complete, independent re-observations of the source transiting the beam system and 2700 independent re-observations of the hit frequency. All of the fslots are archived for later analysis. That ends up being a lot of data so we wrote a special real-time, three-dimensional data visualization program for fslot archives. A user can quickly scroll through frequency and time, instantly rotating the data to view it at arbitrary angles.

After followup mode has completed, the antenna is moved back to 0^h hour-angle and the daily run is re-started at the same declination. This is for two reasons:

1. The daily run that triggered the followup has not finished, and cannot be properly continued since 90 minutes of data were missed when the antenna was pointed west.

2. Re-doing that declination gives us a chance to have another look at the suspect point in the sky.

On the average we will lose about twelve hours of observing time by re-starting the daily run after a followup occurs. To keep from spending too much time re-observing we have set the followup threshold to trigger no more often than about once a week.

Data Archiving

All data that passes the PaulH test, all fslots and all relevant timing information are archived. We have written viewing software (with a nice graphical user interface) to browse the data. We are currently working on using an actual database program for archiving data which will give us easier programmatic access to the data: we will be able to search and sort it according to various criteria.

3.5.3 System Synchronization

The real-time control computer (designated RT) is in charge of keeping the entire system synchronized, both internally and with the outside world. It is a PC running a single-threaded, real-time program which talks to the GPS station clock and LO synthesizer via a GPIB link, and to the rest of the system over the ethernet.

RT controls the two fundamental "heartbeats" of *BETA*:

- the 2 second period of a single FFT integration, and
- the 16 second period of 8 frequency hops, each of which require one short heartbeat period.

Every two seconds RT reads the correct time, calculates an ephemeris (to compensate for doppler acceleration) and pre-programs the LO synthesizer with the compensated frequency value. It then waits for the two-second heartbeat signal from the FFT array, which triggers the LO synthesizer to "instantly" (within 10 microseconds) change frequencies. RT then broadcasts a packet on the ethernet with the current time, frequency hop number and ephemeris data. The rest of the system uses this for synchronization purposes. RT controls how many frequency hops occur, what frequencies they are and when they occur. Its monitor shows a real-time display of the current time, ephemeris, frequency and hop information.

RT also has some secondary functions including power control for the FFT and Pentium racks, voltage monitoring for all power supplies in the FFT rack and temperature and voltage monitoring for the Pentium array PCs.

3.5.4 A Day in the Life of *BETA*

The master control program, *mcp*, is a daemon running on the Unix workstation which handles the scheduling and all major operations of *BETA*. It controls the power to various devices, starts and stops other programs and keeps the operation of many disparate subsystems synchronized. *mcp* handles a day's run by cleaning up from the previous day and setting up for the current:

1. The workstation data handling daemons are terminated.

2. The data collected from the previous day is cleaned up and moved into the permanent archive.

3. Power to the real-time PC and the FFT and Pentium racks is shut off. Doing a major shutdown once a day helps alleviate the problems of "long-term" bugs (resource leaks, etc.) that only present themselves after extended running time.

4. The antenna is moved to the current day's position, usually half a degree (the antenna beamwidth) from the previous day's position.

5. The real-time PC is powered back on and boots over the network. It then downloads system and daily parameters, turns on the FFT and Pentium racks, programs the LO synthesizers and reads the time from the GPS station clock. The PC then begins its synchronization duties, controlling *BETA*'s two second and sixteen second heartbeat cycles.

6. The Pentium array computers boot over the network.

7. The workstation data handling daemons, *etherd* and *back*, are started. They query the Pentium PCs and download parameters and state machine code to them. The PCs start a 10 minute warm-up period, after which the whole system will start taking data.

8. *mcp* then sleeps for 24.5 hours, or until some unusual condition wakes it up. Although a sidereal day of 23 hours, 56 minutes will cover an entire declination

sweep of the sky, it takes a while for the system to do the daily maintenance of the previous 7 steps, so there is some added slop.

After *mcp* wakes up from its daily nap, the cycle starts over again.

3.6 Miscellaneous Housekeeping Systems

Diskless Booting

There are about twenty-five general purpose computers in the *BETA* system[15] and they all need to boot and read program files from some mass storage device. If they each had their own local disk drives, maintenance and software updates would be a nightmare. Because of this problem we made almost all of the system's computers diskless: they boot and get program files from a server computer over the network. The only computer with disks at the observatory is the Unix workstation; the pentium array, the telescope control computer and the real-time computer all boot over the network.

We wrote a special boot-ROM for the ethernet cards that we use. When the diskless computers power up, the boot-ROM appears as a BIOS extension and remaps the floppy-disk service interrupts. When the computer tries to access a floppy disk sector, instead of going to a physical drive, the request is forwarded over the network to the server machine (the Unix workstation). There, a daemon program (called *bootpc*) retrieves the sector from a disk image on a local drive and returns it over the net. For all intents and purposes, the image on the server looks like a 1.44 MByte floppy disk to the diskless computers. Server disk images are maintained with the free `mtools` MS-DOS file system manipulation package.

The diskless computers boot faster over the network than they would from a floppy. Updating software is now easy; only one file needs to be changed and this can even be done over the network. Practically all software development for *BETA* can be done remotely, eliminating many trips to the observatory.

[15] Far more when you include special purpose computers and microcontrollers.

Uninterruptible Power Supplies

The observatory is located in a rural area and the quality of the electrical service is, well, less than ideal. Since the system can take tens of minutes to completely restart after a power failure, even a quick "blink" can cause substantial data loss. To minimize power problems we installed a system of three APC Matrix 3000[16] uninterruptible power supplies (UPSs) in the observatory control room. These can power the system for approximately thirty minutes after the outside power goes down. A daemon program (called *apcserv*) monitors the status of the UPSs and controls the power to the rest of the system, allowing *BETA* to operate smoothly through short blackouts and to shutdown and come up cleanly after longer ones.

Weather Monitoring

Adverse weather conditions at the observatory sometimes force us to park the antenna at certain positions in order to protect it from damage. To monitor weather conditions at the site, we installed a small Peete Brothers weather station. It tracks the wind speed and direction, recent rainfall and the temperature and forwards this information to the Unix workstation. A daemon called *wserv* serves this information to other interested programs.

Currently the weather station's temperature probe is mounted inside the control room and *wserv* has been modified to broadcast alerts or shut down the system if the temperature goes beyond pre-set boundaries.

Digital Network Link

In order to allow remote development, administration and data downloading, we built a computer network link from the observatory to our lab in Cambridge. Early in the project we investigated the possibility of using a dedicated microwave link. Although we checked terrain maps and surveyed several tall buildings, we eventually

[16]Special thanks to American Power Conversion Corporation for their helpful tech support and for waiving the fee for their UPS serial protocol. This saved us much reverse-engineering effort.

decided that this would be too difficult. Instead, we had Nynex install a 56 kbps DDS (Digital Data Service) line. This is a dedicated 4-wire synchronous data link which works constantly without dialing. On both ends of the DDS line we installed Motorola MR56 CSU/DSU units which allow us to send 57.6 kbps asynchronous data along the line. For all intents and purposes the Motorola units make things look like one long RS-232 serial line. A Unix workstation on each end talks to the serial line with a high-speed, buffered serial card. The workstations use PPP protocol and their built-in routing features to put the observatory computers on the Internet.

Chapter 4

Results and Conclusions

BETA is the world's first "all-waterhole" sky survey. Since it went on-line we have surveyed the sky (from $+60^\circ$ to -30° declination) twice and have begun a third run. During this time the system has examined $\sim 10^{16}$ frequency bins, tracked $\sim 10^9$ candidates and archived over 3500 of these which passed preliminary tests. None of these candidates has the characteristics that we expect of an extraterrestrial signal. A diagram showing this winnowing procedure and the amount of data passed by each step is in Figure 4.1.

4.1 Radio Frequency Interference and its Suppression

When we designed *BETA*, a much more powerful successor to *META*, we knew that its wide frequency coverage would increase the amount of interference received. We were not prepared, however, for the huge number of interfering signals we saw. One of the first things we noticed at *BETA*'s "first light" was a large set of spikes from the Global Positioning System satellites at 1575 MHz.[1] Next to be identified were the second harmonic of cellular telephone signals, the third harmonic of TV channel 27,

[1]These were readily identifiable from their frequency, $\left(\frac{\sin x}{x}\right)^2$ envelope and fine structure characteristic of direct-sequence spread spectrum modulation.

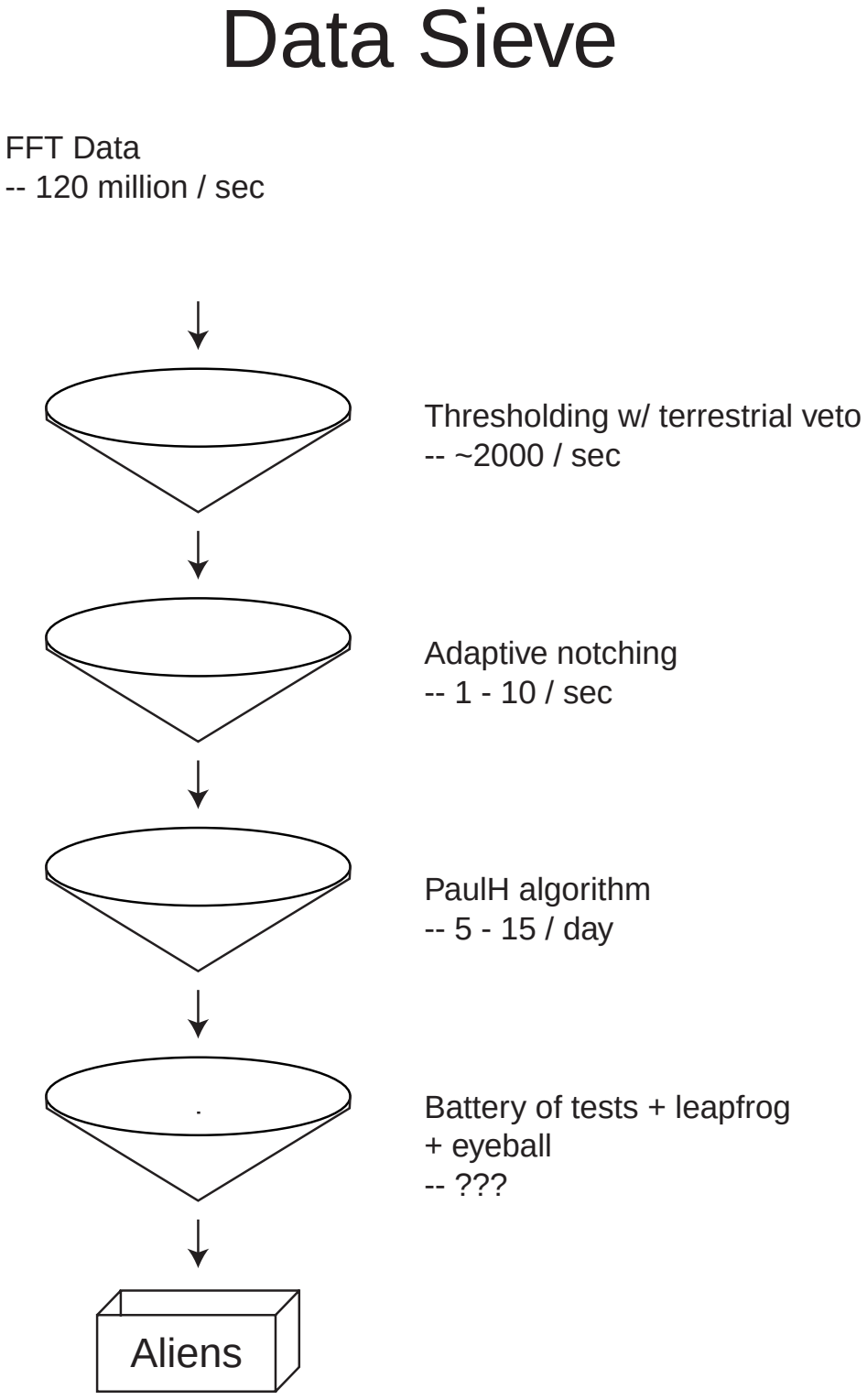


Figure 4.1: *BETA* data sieve, showing the amount of data passed by each stage.

the Russian Glonass satellites and the SARSAT search-and-rescue system. There was also a lot of unidentified clutter near 1500 MHz and several other frequency regions.

BETA was originally designed with two interference rejection systems: the terrestrial "veto" feed and the dual sky beams. The terrestrial feed was meant to process a high volume of data, but the dual sky beams were not (since useful comparisons between the beams require a delay and therefore storage of data). The combination of excessive RFI and a terrestrial veto that didn't work as well as expected forced us to make additions and changes that upset some of our designed-in features. The modifications do make the system work reasonably well; they are described below.

Terrestrial Veto Results

The terrestrial feed provides a fast way to veto strong interfering signals early in the recognition process. The vetoing is performed in hardware by the Feature Correlator state machine using a simple algorithm: if the terrestrial signal strength for a frequency bin is above a certain threshold, no hits will be generated for that bin regardless of the strength in the other two channels.

This strategy works well for strong carriers, but it is not perfect. Some satellite signals may appear in the sky horns but be too weak for the terrestrial horn thresholds. Also, although RFI signal strengths are approximately the same for all of the feeds (the terrestrial discone and most of the telescope sidelobes have gains of around 0dBi), the terrestrial feed has more thermal noise because it looks mostly at the ground, raising the noise floor and thus desensitizing it to RFI. It is clear from looking at examples of interference that signals from the terrestrial channel are proportional to those in the sky channels. Figure 4.2 is an example of this: interference from a global positioning system satellite showing banding. The terrestrial signal level (on the bottom) is obviously proportional to the others, yet it did not get eliminated by the terrestrial veto. However, this requires averaging a group of nearby frequency bins, since the terrestrial signal level in an individual bin may be indistinguishable from thermal noise even when the sky horn levels are quite strong. Despite this, the terrestrial channel is useful in the signal recognition algorithms. Because we

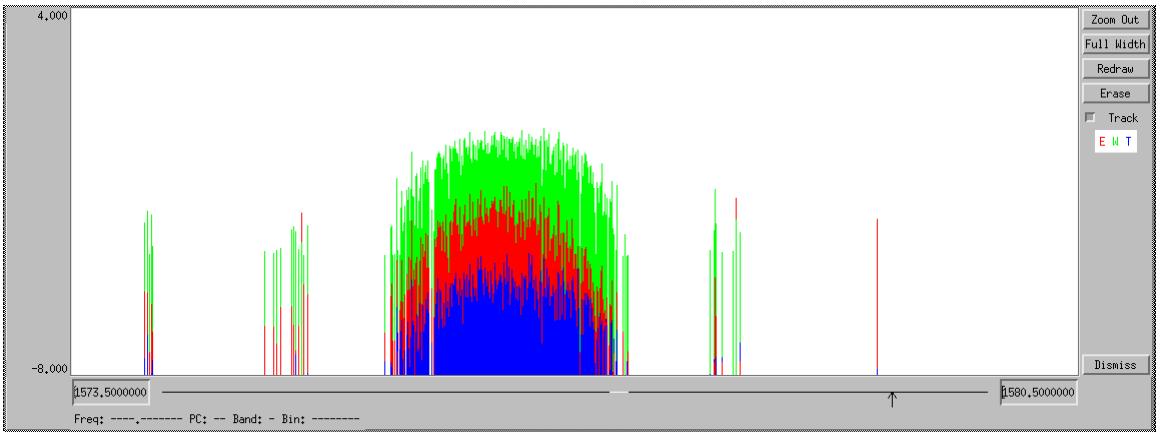


Figure 4.2: Interference from a GPS satellite showing banding. The terrestrial signal (on the bottom) is clearly proportional to the two sky beam signals, and yet it was not eliminated by the terrestrial veto.

have a time history of a band of frequency bins, we can analyze correlations between them and verify the presence of even weak terrestrial interference. Figure 4.3 has an example of this.

State Machine Results

We had high hopes for the state machine. The prospect of fast, very low threshold detection was extremely appealing during the design phase. The idea was to track every frequency bin's progress in hardware, looking for the east to west characteristic of a sidereal source. In an interference-free environment (or if the terrestrial veto were better for single events) this would have worked very well; since thermal noise is uncorrelated between spectra it is highly unlikely that such a pattern would occur from chance, even at low thresholds. The problem is that interference *is* correlated between spectra. At the higher thresholds at which slots are generated, we find $\sim$ 10 events per day that pass an east-west test. Lower thresholds produce even more. Since the state machine is built around a piece of memory, it can only have a small number of inputs[2] (used for time and frequency based state, threshold results,

[2]Because the inputs are just the address lines for the memory, and memory size increases exponentially with the number of address lines.

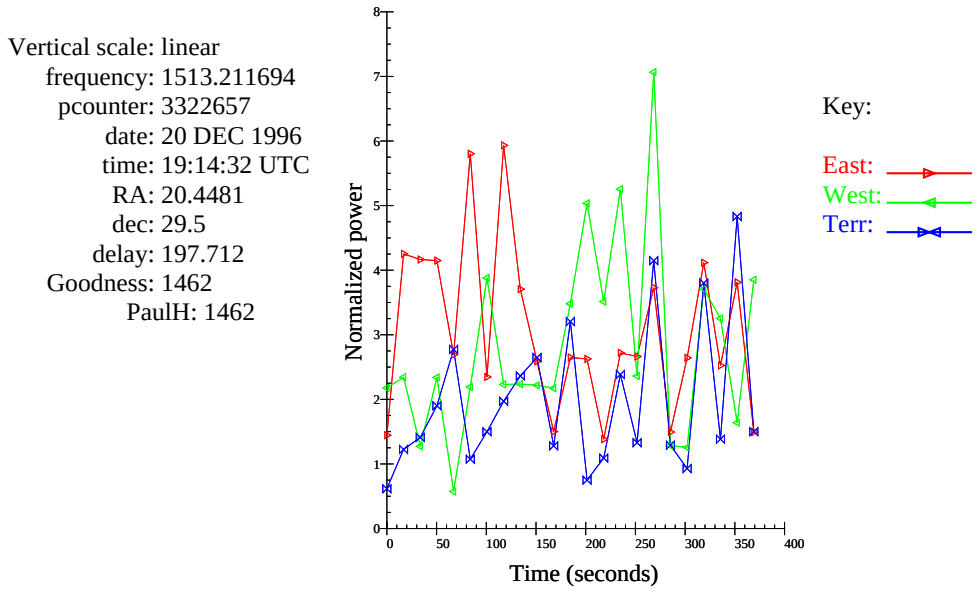


Figure 4.3: A low level terrestrial signal showing correlation with the east beam.

etc.) which means that it can only distinguish between a small number of patterns. Random interference will frequently generate any of the patterns if the thresholds are low enough. This low-level detection scheme still might be useful except for one thing: since a state machine "hit" is not triggered until the event has already occurred (and finished), no data can be stored and there is nothing to analyze further.

The state machine can still be used to trigger regular hits and then proceed with the slot generation and following mechanism. It is also very useful for debugging and running various tests on the equipment. It was, however, tricky to design and debug, so we would probably not implement something this general in a future system.

Adaptive Filtering Results

The adaptive filtering (sometimes called "mini-notching") scheme was designed after we had some experience with the RFI environment that *BETA* contends with. Before we implemented it, the interference was so bad that the system was nearly useless; now the situation is satisfactory. The scheme consistently eliminates more than 99%

of the interference while masking less than 1% of the available spectrum.

In order to test the adaptive filtering properly, we incrementally lowered the thresholds while watching the hits that were produced. Every time we decreased the threshold, new RFI suddenly appeared and then gradually went away. Since thermal events appear at different frequencies in each spectrum, the non-interference, thermal background got thicker each time and did not go away. Because the mini-notch thresholds we use (see Section 3.5.1) ramp the counters up five times as fast as they ramp down, we should theoretically be able to use the scheme with thresholds so low that each 2 kHz block gets a thermal hit every five spectra. This corresponds to a threshold of $8.5P_0$, or about 800 hits per spectrum, which is far more data than the rest of the system can handle.

It would not be difficult to implement the adaptive notching scheme in hardware. This would enable the system to handle large quantities of interference without overloading any of its data paths (i.e. the ISA bus). In hindsight, we probably should have done this instead of developing the generic state machine.

Results of the Analysis Algorithms

The PaulH algorithm provides a first cut test that any candidate signal that meets our assumptions must pass. It tests for the basic east-to-west/not-terrestrial characteristic that a sidereal source should display. Figure 4.4 shows a histogram of the frequencies of every candidate that has passed this test. The candidates are scattered over the entire bandwidth, with "spikes" at certain frequency bands which have intermittent interference. If the interference at those locations were constant, the adaptive filtering would have eliminated it. Out of $\sim 10^9$ slots that were tracked during *BETA*'s run, 3500 passed the PaulH test. This means that at the $15P_0$ threshold level in our interference environment, a part in 10^5 of all slot events have at least a minimal east-to-west/not-terrestrial character. Figure 4.5 shows a fake slot, meant to look somewhat like a real signal. The best slots we get in reality are more like Figure 4.6 with the "signal" barely out of the noise. Note that the terrestrial signal tends to have a value < 1. Thermal noise would not have such statistics and this is a sign of

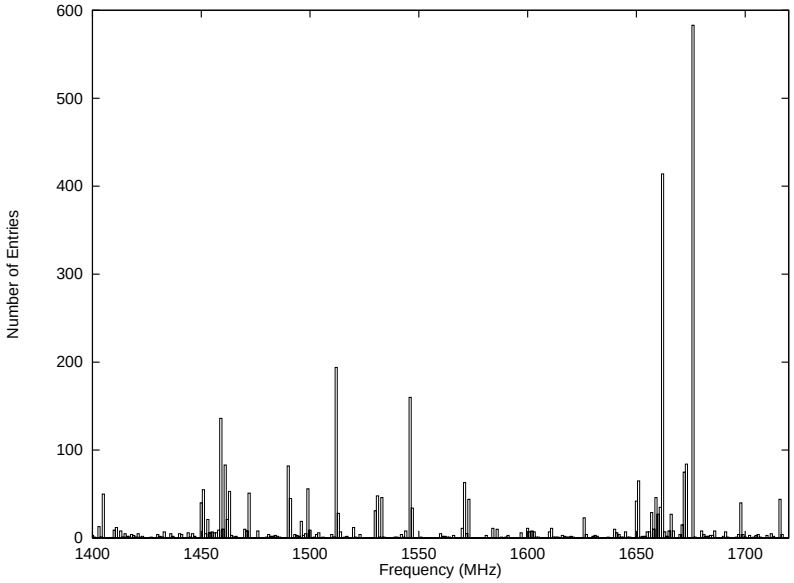


Figure 4.4: Histogram of archived candidates by frequency.

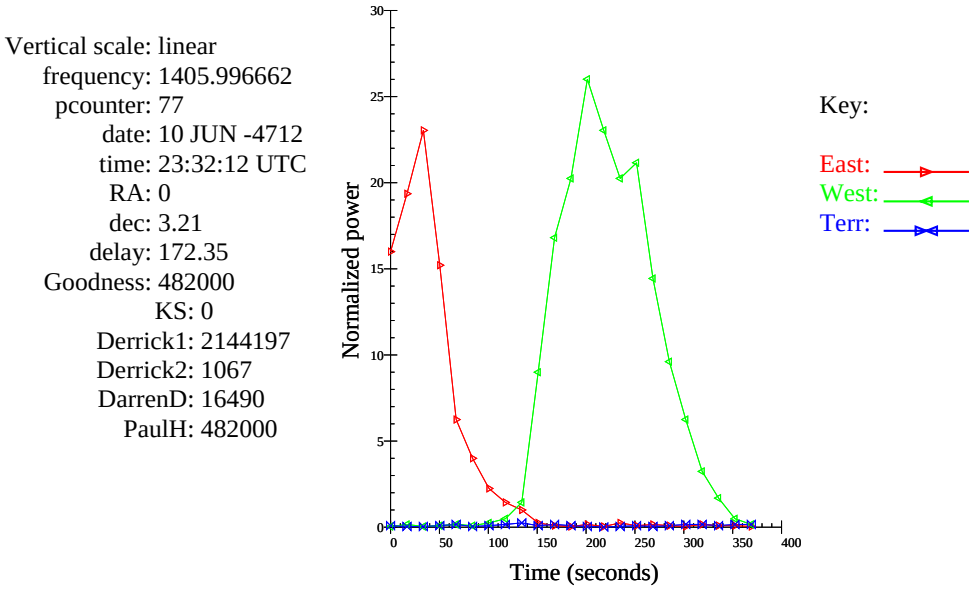


Figure 4.5: Fake data meant to look like a successful slot.

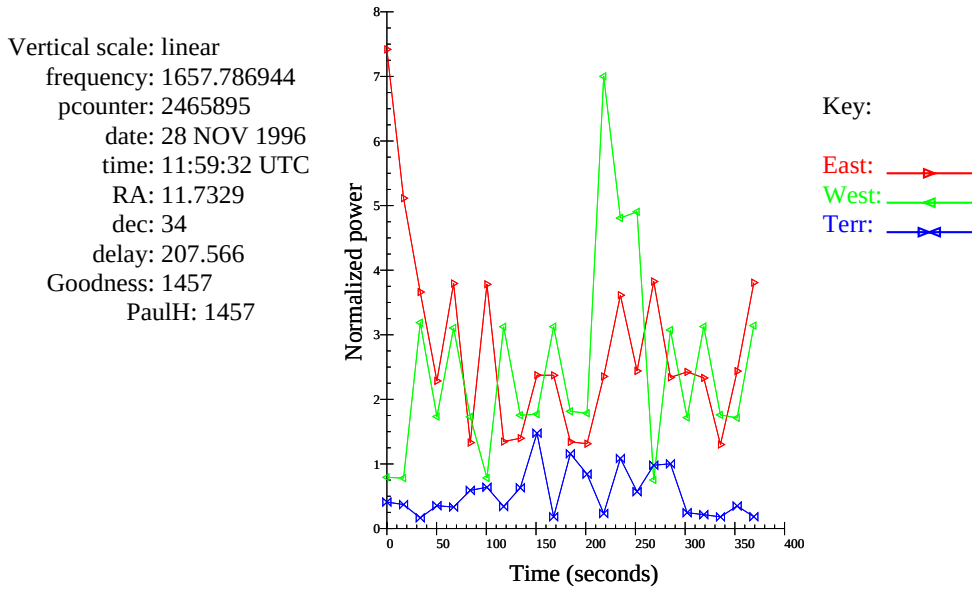


Figure 4.6: A slot which is typical of the better ones we have received. Note the depressed terrestrial signal.

nearby terrestrial interference raising the baseline and thus depressing the normalized power values.

Leapfrog Results

The "leapfrog" followup mode (described in Section 3.5.2) is tricky to trigger properly. If it happens too often, then a lot of observing time is lost. If it doesn't happen often enough, it's nearly useless. We think a false alarm rate of about once/week is a good compromise between the two. The only way to know when to trigger the followup is from the results of the various analysis algorithms. Figure 4.7 shows some an example of an automatic followup. The box shown can be rotated by the user in real-time, making it easy to pick out small details in the large, 3-D followup data set. The axis running NW-SE represents frequency (and the particular band can be changed by moving the slider on the bottom), the NE-SW axis represents time and the vertical axis represents amplitude. The six divisions correspond to six different re-looks of

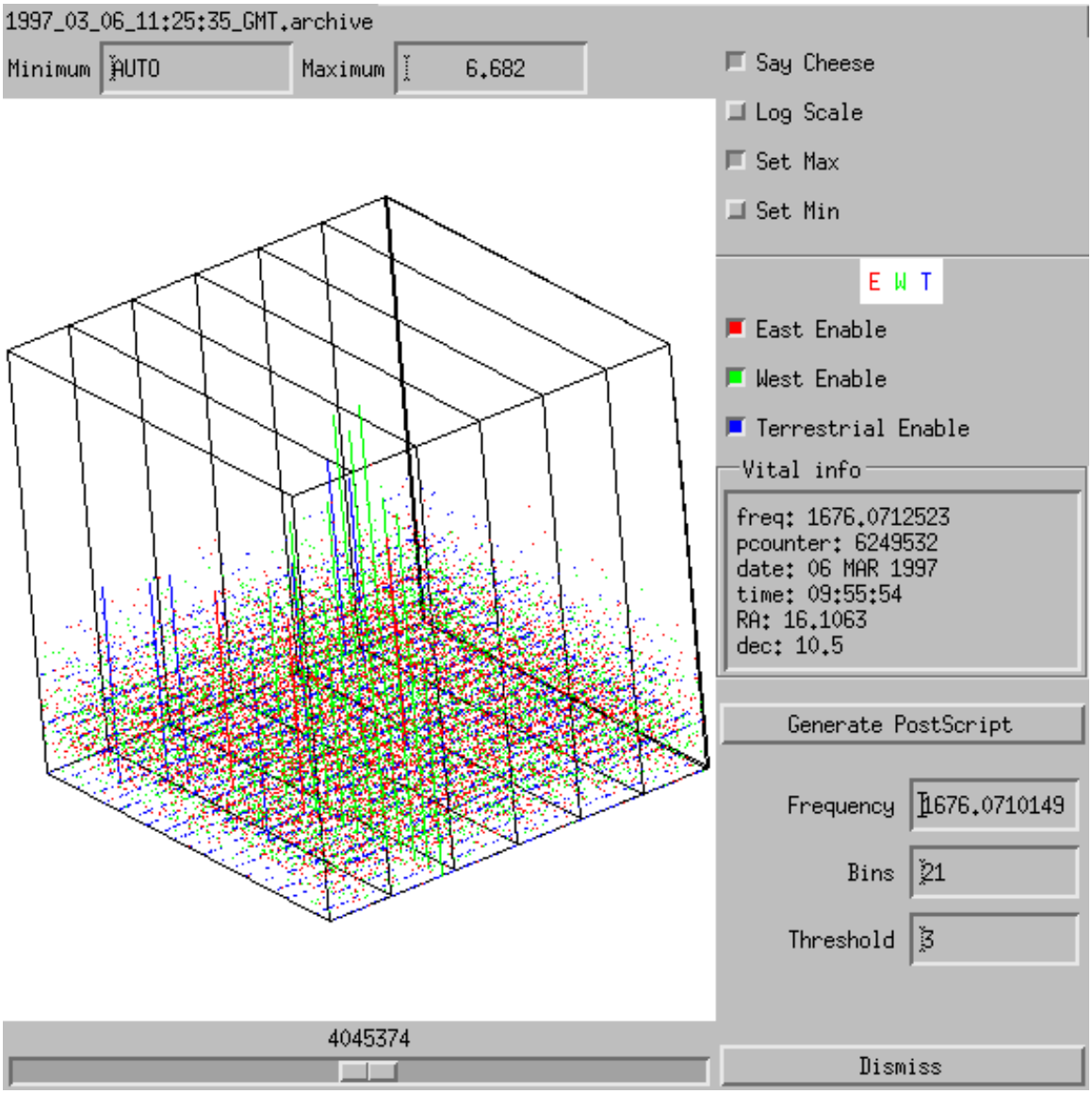


Figure 4.7: Sample "leapfrog" followup results. (It looks great in color and rotating, but black-and-white and static doesn't do it justice.)

the same sky position, each 15 minutes apart. A real signal would show a double lobe pattern, similar to Figure 3.4, in each of the six divisions. This pattern would would be narrow in the other axis, presumably occupying only a few frequency bins. What we see in the example shown here is just the opposite: the pattern is spread out along the frequency axis and is narrow along the time axis. This is characteristic of transitory RFI.

4.2 Some Conclusions about the Prevalence of Transmitting Civilizations

The current negative results of the *BETA* search allow us to set some limits on the prevalence of transmitting civilizations, with certain qualifications. From the system's sensitivity parameters we can derive a relationship between a transmitter's EIRP and the maximum distance at which we could detect it. Since EIRP $= P_t D_t$, then $P_t A_t = \dfrac{(\text{EIRP})\lambda^2}{4\pi}$ and equation 2.3 becomes

$$R = \sqrt{\frac{(\text{EIRP})A_r\tau}{8\pi\gamma k T_N}} \tag{4.1}$$

The extra factor of 2 in the denominator is due to the fact that we are receiving with linear polarization, but the signal was presumably transmitted with circular polarization. With our values of $A_r = 239$ m^2, $\tau = 1/2$ second, $\gamma = 15$ and $T_N = 85$K, we obtain

$$R = C\sqrt{(\text{EIRP})} \tag{4.2}$$

with $C = 1.6 \times 10^{10}$ m$/\sqrt{\text{W}} = 1.7 \times 10^{-6}$ ly$/\sqrt{\text{W}}$. Figure 4.8 shows a log-log plot of this relationship.

Like the *META* analysis in [24], we will consider three types of transmitting "super-civilizations", similar to the Kardashev[3] definitions [27]. *Type 0* civilizations

[3]Kardashev's Type I is the Type 0 here. He did not specify civilizations that use their entire planetary insolation.

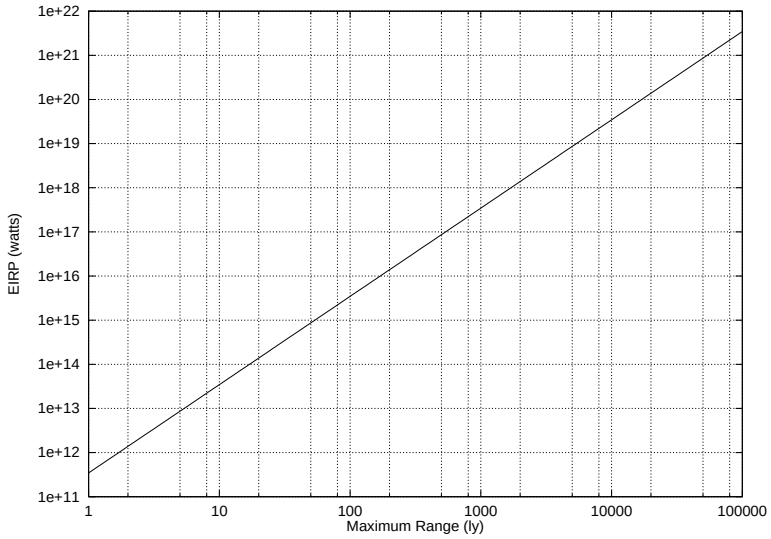


Figure 4.8: Transmitter EIRP vs. maximum range of the search.

are similar to ours and have total power resources of about 10^{13} watts. *Type I* civilizations have power available equal to the solar insolation on earth: about 10^{17} watts. *Type II* civilizations harness the entire power of their star: about 10^{26} watts. We will also consider three types of beacons: isotropic ($D_t = 1$), galactic-isotropic[4] ($D_t \approx 50$) and directed beams with sufficient gains to be received here from anywhere in the galaxy ($R < 80,000$ ly). If these civilizations (lavishly) use a significant part of their power to broadcast a SETI beacon within *BETA*'s limitations (frequency range, doppler-compensated to an inertial frame, within our declination range, not conflicting with RFI, etc.), we can set the following limits on the number N of such civilizations in our galaxy (where ϵ is the earth-incident duty cycle of transmission, i.e. the fraction of time spent transmitting in our direction):

For Type 0 civilizations, $N\epsilon < 1$ out to a distance of 5.4 ly for an isotropic

[4]Since the galaxy is disk-shaped, an isotropic beacon is wasteful. It is more efficient to transmit with higher gain in the plane of the galaxy and with less toward the poles. If the galaxy is considered a disk with radius $R_\mathrm{g} \approx 50,000$ ly and thickness $d_\mathrm{g} \approx 2000$ ly then, from the center of the galaxy, the rim subtends an angle of about $d_\mathrm{g}/R_\mathrm{g} \approx 0.04$ radians in galactic latitude (θ_g). Therefore, the solid angle of the rim as seen from the galactic center is $\Omega_\mathrm{rim} \approx (2\pi R_\mathrm{g} d_\mathrm{g})/(4\pi R_\mathrm{g}^2) = 0.02$ steradians. A galactic-rim beacon would therefore have a gain of about 28dB. To illuminate the nearer disk-boundaries properly (their distance from the center is $d_\mathrm{g}/2\sin\theta_\mathrm{g}$) we need $D_t \propto \sqrt{R}$. Integrating D_t over 4π steradians we find that over ten times as much power is needed to illuminate the nearer regions than the rim and that $D_{t(\max)} \approx 17$dB.

beacon. $N\epsilon < 1$ to a distance of 38 ly ($\approx$ 100 sun-like stars) for the main lobe of a galactic-isotropic beacon. $N\epsilon < 1$ for the entire galaxy ($\approx 10^{11}$ sun-like stars) for directed beacons with transmitting gain $D_t \approx 83$dB. An 83dB antenna would be about 1000 meters in diameter at $\lambda = 21$ cm, and would have a beamwidth of about one arc-minute.

For Type I civilizations, $N\epsilon < 1$ out to a distance of 540 ly ($\approx 2 \times 10^5$ sun-like stars) for an isotropic beacon. $N\epsilon < 1$ to a distance of 3800 ly ($\approx 3 \times 10^7$ sun-like stars) for the main lobe of a galactic-isotropic beacon. $N\epsilon < 1$ for the entire galaxy ($\approx 10^{11}$ sun-like stars) for directed beacons with transmitting gain $D_t \approx 43$dB. A 43dB gain antenna at 21 cm has a diameter of 10 meters and a beamwidth of 1.2°.

For Type II civilizations, $N\epsilon < 1$ out to a distance of 1.7×10^7 ly ($\approx 10^{11}$ sun-like stars in our galaxy and $\sim 10^{12}$ in neighboring galaxies) for an isotropic beacon. At this power level, directed beacons could be detected at cosmological distances, which is not very useful for SETI considering the time-scales of biological evolution.

4.3 Suggestions for Future Searches

BETA is not a particularly sensitive search. The 26 meter telescope we are using is quite small for a research instrument. While we were constrained by budgets and the facilities available to us, Equation 2.21 suggests that future searches should endeavor to use the largest aperture possible. Our short 0.5 second integration time, while necessitated by the desired frequency coverage and equipment cost, is insufficient for interstellar signals with intrinsic bandwidths of $\sim 10^{-2}$ Hz.

Decreasing the beam size and increasing the integration time will make a meridian transit (drift scan) sky search, like ours, impossible. The beam(s) will have to move in hour angle, tracking a particular position on the sky. A modern way to solve these problems is with aperture synthesis. Signals from a phased array of antennas can be digitally combined to synthesize several simultaneous sky beams. These beams can be electronically steered and otherwise controlled. Several different users can even be supported at once. While the cost of heavy hardware and construction remains the

same or increases, the cost of electronics and computer processing power continues to plummet – it makes sense to design projects that take advantage of this. Array antenna systems like the above are already in the planning stages, e.g. Ohio State's Argus [13] and The Square Kilometer Array Interferometer [3] in the Netherlands.

Any future SETI searches should have substantial RFI-proofing designed in from the beginning. To begin with, the radiotelescope should be sited in a low-interference environment.[5] The search equipment should be designed to handle interfering signals that are many times the power and number of the expected thermal noise. If possible, the interference should be reduced or eliminated as far "up the data stream" as possible: analog filtering in the front-end, hardware adaptive filtering, etc. These RFI reduction methods should be *data independent*, i.e. they should be able to operate despite the amount of data they receive. This is to keep strong bursts of interference (such as a satellite passing through the main beam) from completely disrupting operation. The worst thing that should happen is for data to be lost only during the period and near the frequencies of the RFI.

4.4 Final Thoughts

The *BETA* system was designed to provide sufficient information to either prove or disprove the extraterrestrial provenance of a signal. We were not disappointed in this regard – the system detected *no* signals that remain mysterious. None of the archived candidates has the characteristics that we expect of an extraterrestrial signal. This is due to one of the following three reasons: either our assumptions about signal characteristics were incorrect, or we were unlucky or no signals were within our capability to receive (as discussed about in 4.2).

[5] SETI researchers have even picked out a spot on the far side of the moon for this purpose: Saha Crater near the lunar equator. [22]

Incorrect Assumptions?

Because no extraterrestrial signals have ever been received, we could only make educated guesses about their characteristics. We may have guessed wrong. The "waterhole" band of frequencies may not be the preferred place to transmit a beacon. Perhaps communicative civilizations are expected to be near-space faring, and signals are transmitted outside of the atmospheric window. Perhaps the optical regime is considered better for beacon use. Transmitting an unmodulated beacon does seem rather wasteful; it may be that the signals are modulated and so are not recognized by our carrier detection scheme. Perhaps basis vectors other than sinusoids are used. *BETA* assumes that any signals will be doppler corrected to an inertial frame, but a different civilization may consider the doppler changes to contain interesting information about their planet's motion and might therefore leave them in. The duty cycle of our search is very low: we spend only about 10^{-5} of our search time on any particular piece of the sky, so we will certainly miss any low-duty-cycle signals.

We could second-guess ourselves forever; the only way to really find the answer is to keep looking through the search space and exploring a wide range of options.

Appendix A

Antenna System Calibration

The antenna temperature T_A is defined as the fictional black-body temperature with which the antenna would have to be surrounded in order to receive the same power density (power per unit frequency) that it does when pointed at some source. The system noise temperature T_N is defined as the excess temperature added by the receiving system regardless of where the antenna is pointed. The system temperature T_S is defined as $T_A + T_N$. Since we are operating in the Rayleigh-Jeans region where power is proportional to temperature, the intermediate-frequency (IF) output power P_{IF} is proportional to T_S and therefore an affine transformation of T_A.

We calibrated the dual-feed antenna system using "hot load/cold load" techniques. Figure A.1 shows various measurements made on May 31, 1993. Note that the old GaAs low-noise amplifiers were still installed at that time. T_N was calculated by comparing P_{IF} when the antenna was pointed at cold sky ($P_{\mathrm{IF}} = 26.0\mu\mathrm{W}$, $T_A = 0$[1]) to that when then antenna area was covered with absorbing material ($P_{\mathrm{IF}} = 93.8\mu\mathrm{W}$, $T_A = 291°\mathrm{K}$ – the ambient temperature that day). The ratios of T_S to P_{IF} will be equal, i.e.

$$\frac{T_N}{26.0} = \frac{T_N + 291°\mathrm{K}}{93.8} \tag{A.1}$$

[1] T_A is not really zero; there are actually contributions from the 3°K cosmic microwave background, atmospheric absorption, side-lobes hitting the warm ground and loss in the antenna itself. We're going to roll these into T_N which won't be a problem unless we point the antenna at something that doesn't include all of them. We promise not to.

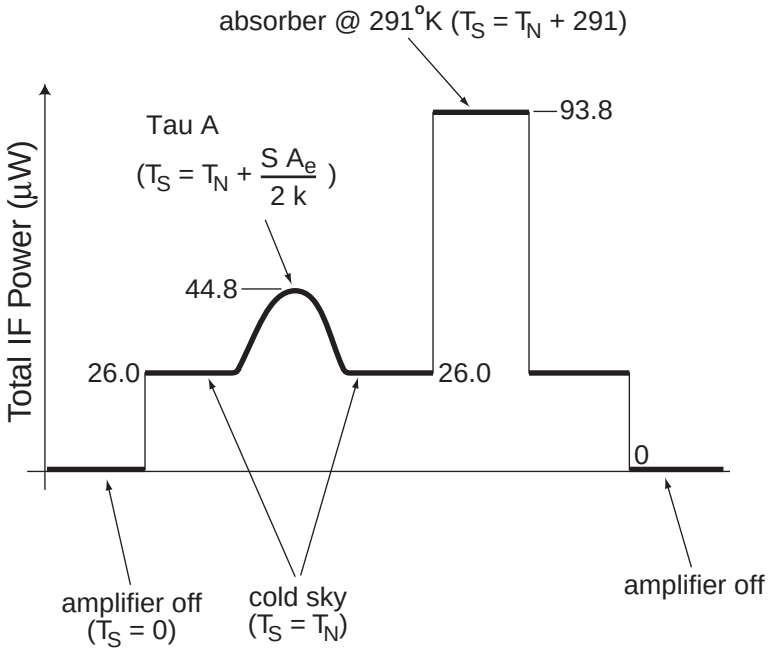


Figure A.1: Calibration of the system noise temperature (T_N) and effective area (A_e) of each feed of the dual-feed 26-meter Cassegrain. Bolometric total-power measurements were taken at IF, using cold sky, Taurus A and absorber material at ambient temperature. The Tau A flux value is computed from Green [18].

This yields $T_N \approx 112°\text{K}$.

We calculated the antenna's effective area with the same technique, except using an astronomical source instead of the absorber material. The source was Taurus A, the Crab nebula, which is a bright radio source with

$$S_{\text{TauA}} \text{ at } 21\text{cm} = 936 \text{ Jy} = 9.36 \times 10^{-24} \text{ W/Hz/m}^2$$

With the antenna pointed there, $T_A = \dfrac{S_{\text{TauA}} A_e}{2k}$ where A_e the is effective area and k is Boltzman's constant (1.38×10^{-23} joules/K). There is a 2 in the denominator because S_{TauA} is the total power but, since the antenna is sensitive to only one polarization, we receive half. Plugging the values into

$$\frac{T_N}{26.0} = \frac{T_N + S_{\text{TauA}} A_e / 2k}{44.8} \tag{A.2}$$

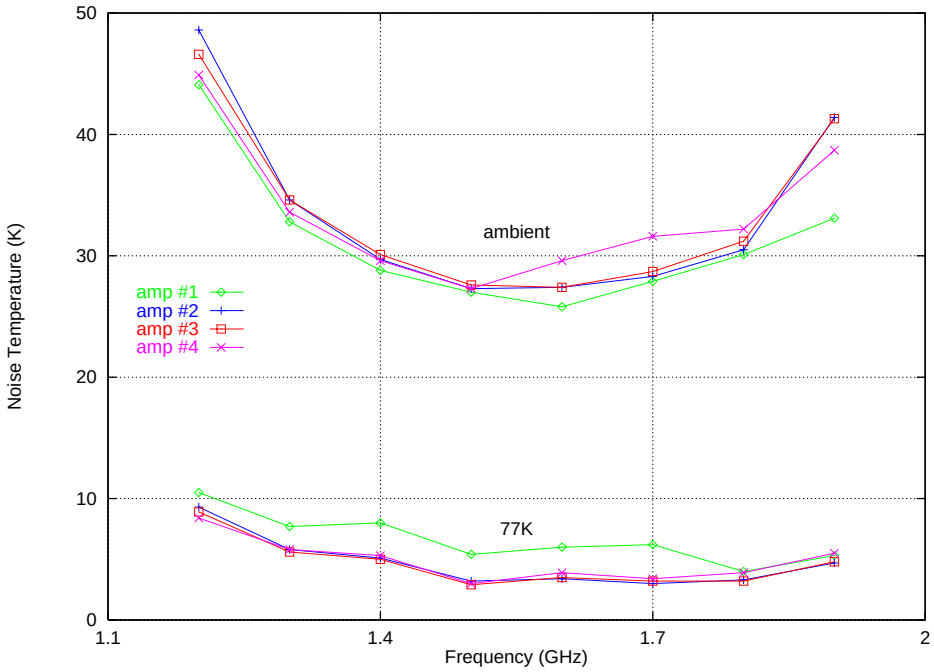


Figure A.2: Noise temperature vs. frequency of the 4 L-band HEMT amplifiers from Berkshire Technologies. Measurements were made for the amplifiers both at room temperature and at 77K. Performance over the "waterhole" is about 30K and 5K, respectively.

we get $A_e \approx 239$ m^2. The antenna's efficiency ϵ is the ratio of its effective area to its physical area. For our 84 foot diameter dish, $\epsilon = \dfrac{A_e}{\pi R^2} \approx 0.46$. Note that after the 84 foot upgrade in 1970, the antenna had an efficiency of 0.516 at 21 cm with the optimal horn for that wavelength [1]. The new horns are only 0.5dB below that.

The *sensitivity* of the system, measured in kelvins/jansky is the incremental increase in T_S per jansky increase of the source. It is equal to $\dfrac{A_e \cdot (1\text{Jy})}{2k}$ which for our system is 87 mK/Jy.

The previous system's low-noise amplifiers have since been replaced by even lower-noise HEMT amplifiers from Berkshire Technologies of Oakland, California. Figure A.2 shows the measured noise temperatures of the four amplifiers (which we built from kits) at both room temperature and 77K. The old amplifiers had a noise temperature of about 55K, so new ones have decreased T_N to about 87K. The effective area and sensitivity are, of course, unchanged.

Appendix B

Beam Forming

BETA is designed to reject terrestrial interference (the number one problem in SETI) by exploiting the property that a genuine extraterrestrial signal must be both point-like and exhibit sidereal motion. Thus we built a two-horn receiver, with stationary beam lobes oriented east-west, along with a third "terrestrial" low-gain antenna. Although we originally envisioned a pair of beams separated by several beam widths, we decided in favor of some degree of overlap,[1] such that the hand-off from one beam lobe to the other keeps the source in sight continuously. To implement this, we looked at two schemes (Figure B.1), namely

- A phased array consisting of 10 hexagonally-packed feedhorns and three low-noise preamps (with the central cluster of four horns passively combined, then buffered and phased with each of the passively combined outer sets of three horns; each horn could be either linear or dual-circular polarization), and
- A simpler arrangement of two pyramidal (linear polarization) horns, aligned along their E-plane axes.

The 10-horn arrangement exhibits a remarkable azimuthal symmetry of beam pattern (Figure B.2), but requires a major (lengthy and expensive) construction effort; by

[1]Suggested by Prof. David Staelin of the Research Laboratory of Electronics at MIT. Note that because of energy conservation it is not possible to have (efficient) beams overlap closer than their 3dB points. If it were, in the overlap region each beam could receive more than half of the energy incident on the dish from that direction.

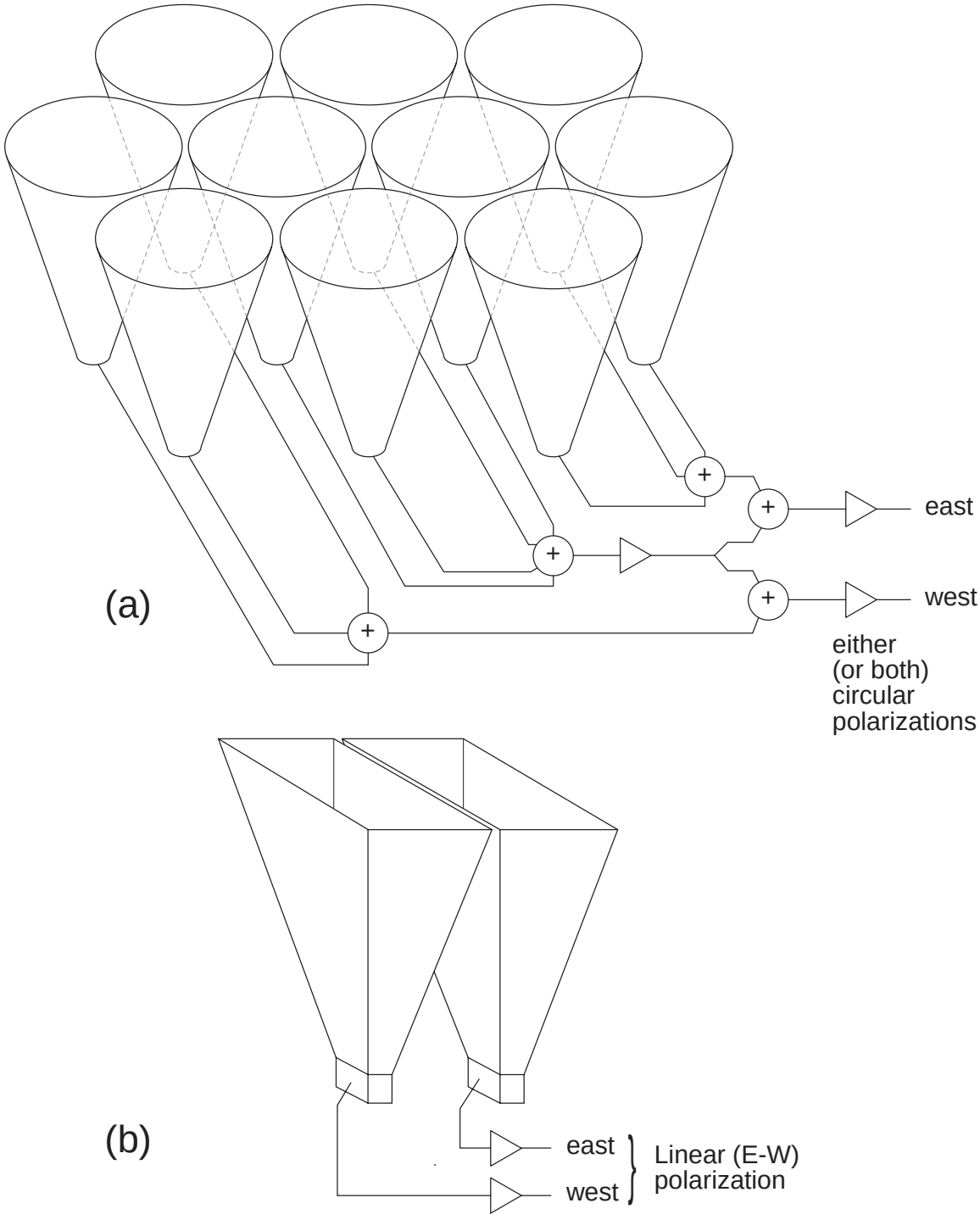


Figure B.1: Feed alternatives for generating dual-beams from a single parabolic antenna. At (a) is a phased array of dual-polarization circular horns using passive combiners and low-noise amplifiers (only one polarization is shown). At (b) is a pair of pyramidal, linearly-polarized horns, stacked along the E-plane.

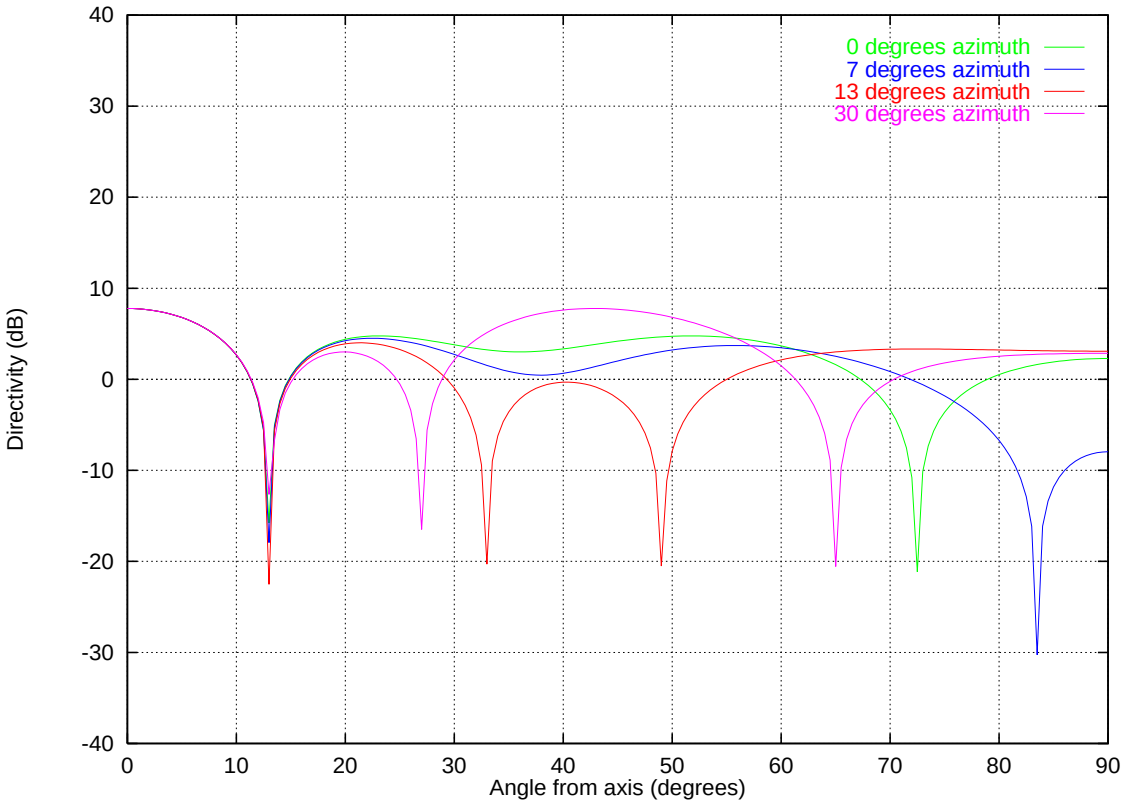


Figure B.2: Antenna gain as a function of zenith angle (for four choices of azimuth), for a hexagonal phased subarray of non-interacting point radiators, with lattice-plane spacing of 1.5 wavelengths. For an array of finite sized conical feeds, as in Figure B.1, this plot must be multiplied by the single-horn diffraction pattern, which largely suppresses the off-axis grating lobes. Note that the use of dual circular polarization feeds, with separate combining networks, allows one to construct a dual-beam, dual-polarization focal plane array. Though the pattern shown is for a hexagonal array, the result for a heptagonal array is nearly identical. The close matching of main-lobe patterns is maintained for all azimuth angles; the particular choices here were meant to be "random".

contrast, the pyramidal pair is sensitive to only one linear polarization, but is easy to build and try out.

With either scheme we were told that near-field aperture interactions would cause major distortions of the far-field pattern, though no one was able to quantify the effect. To satisfy our curiosity we made some laboratory "test-range" measurements with a pair of pyramidal X-band (3 cm) horns, driven (via a magic-T hybrid) both alternately and simultaneously with Gunn oscillator sources, while scanning the far-field pattern with a small dipole connected to a spectrum analyzer. We were unable to see any interaction effects at the measurement accuracy (approximately 1dB).

While favoring the simplicity of the stacked pyramidal horns, we were concerned about two additional issues:

1. Does the far-field pattern have reasonable azimuthal symmetry, and

2. Is there adequate beam overlap with separate horns (which is guaranteed with the interleaved 7-horn array)?

To answer these questions we asked Martin Gimersky[2] to perform diffraction calculations, starting with a pyramidal horn design whose H-plane to E-plane dimensions are in the ratio of 1.35 (this ratio produces equal -3dB beam widths, owing to cosine taper in the H-plane field amplitude, combined with uniform field amplitude along the E-plane). We assumed no feedhorn interaction, and simply calculated the far-field pattern, using the parameters of our 26-meter antenna (full-width illumination angle of 18.5 degrees). Figure B.3 shows the far-field beam intensity, from a single displaced horn, for three choices of horn aperture. In each case the horn center has been offset along the E-plane, relative to the Cassegrain axis, by half the horn aperture (i.e., stacked horns). A rule of thumb to achieve maximum efficiency in Cassegrain design is to taper the illumination to approximately -10dB at the reflector edge. That corresponds to Figure B.3c, producing beam overlap at the -6dB point; it also results in feedhorns that do not fit in the radome!

[2]Of the University of Victoria, B.C. See also [16].

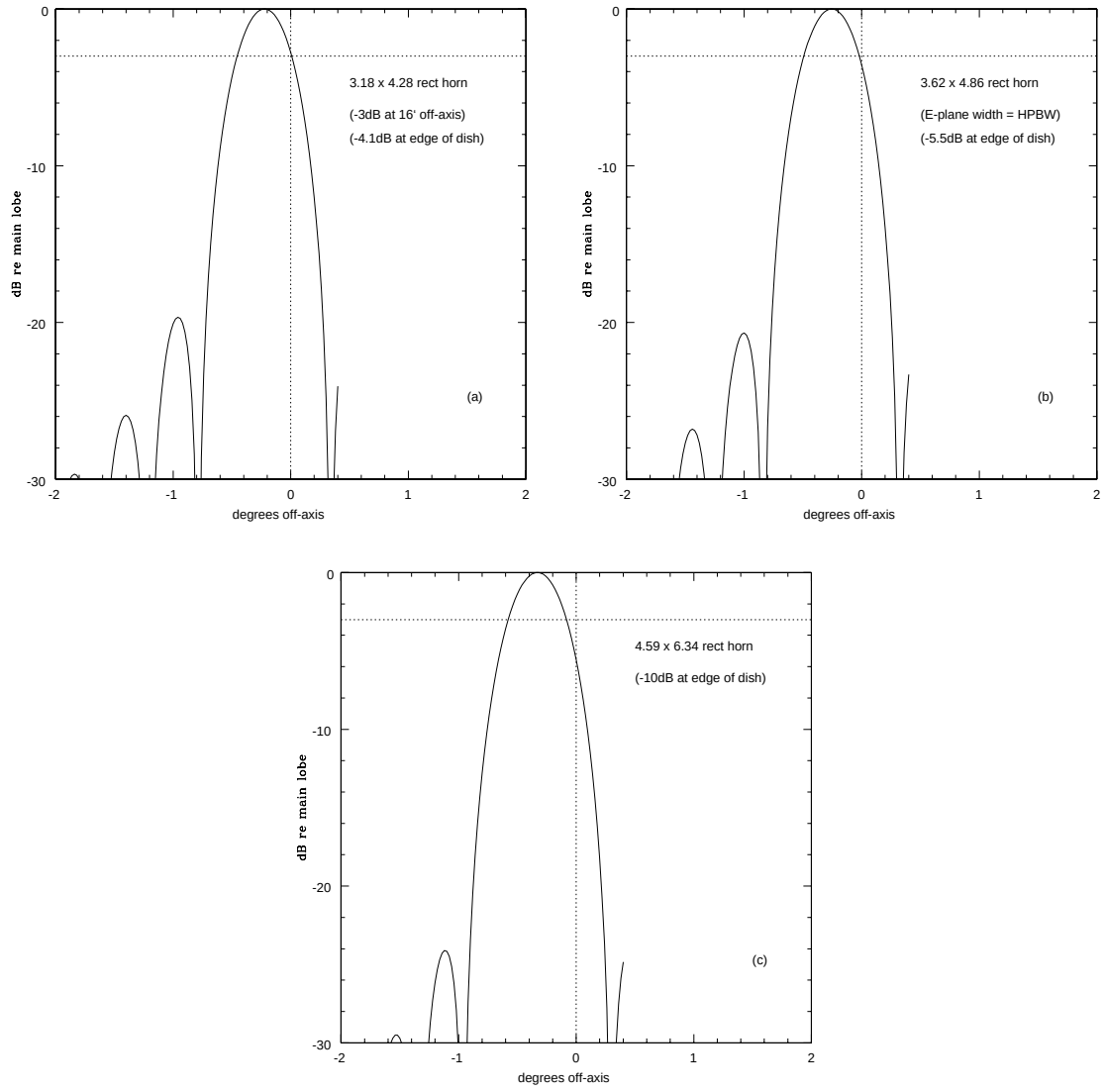


Figure B.3: Far-field antenna pattern for a 26-meter Cassegrain illuminated by a single pyramidal horn that is displaced along its E-plane by half its aperture. The three cases plotted progress to larger apertures, specified in wavelengths, with edge tapers of -4.1dB, -5.5dB, and -10dB, respectively. Each graph is centered on the H-plane and plotted versus angle in the E-plane, with the vertical dashed line indicating the antenna axis, and the horizontal dashed line indicating -3dB relative to maximum gain. The horns are assumed non-interacting.

We finally settled on the design of Figure B.3b which we estimated to have a paraboloid efficiency (i.e., spillover efficiency times aperture efficiency) just 0.6dB less than the ideal; its taper at the edge of the dish is -5.5dB, compared with the conventional -10dB, resulting in somewhat increased sidelobe amplitude. It fits in the radome, if the outside corners are cut diagonally (Figure 3.2). The feedhorn length was chosen to produce wavefront curvature of about 0.2 by 0.3 wavelengths (E-plane by H-plane), resulting in finished feedhorns that fit in the radome with about 1 inch to spare. They are constructed of 1/8 inch aluminum sheet (6061-T6), heliarc welded and joined to a WR-650 waveguide section with flange.

We mounted the horns, and performed drift scans of astronomical point sources (Sgr A, Cyg A, Tau A). Figure 3.4 shows such a scan. The beam shape and overlap are ideal. However, the observed signal strength was lower than expected because, for mechanical reasons, the horns are mounted a little closer to the telescope's secondary reflector than optimum.

Appendix C

Computing Large Discrete Fourier Transforms

The Discrete Fourier Transform (DFT) $X(k)$ of a series of numbers $x(j)$ is calculated with

$$X(k) = \sum_{j=0}^{N-1} x(j)\, W_N^{jk}, \text{ for } 0 \leq k \leq N-1 \tag{C.1}$$

where $W_N = e^{-i(2\pi/N)}$.

It is possible to compute a large DFT as a series of smaller ones. If the length of the transform N is the product of two integers N_R and N_C, then we can represent the input time series values as a two-dimensional array with N_R rows and N_C columns. We can define

$$j \equiv (cN_R + r) \quad \text{and} \quad k \equiv (\kappa + \rho N_C) \tag{C.2}$$

$$\text{with} \;\; c, \kappa \in [0, N_C] \;\; \text{and} \;\; r, \rho \in [0, N_R]$$

which means that the time-domain data is written in column-by-column and the frequency domain data is read out row-by-row. The W_N^{jk} term of equation C.1 then becomes

$$W_N^{jk} = W_N^{(cN_R+r)(\kappa+\rho N_C)} = W_N^{\rho c N_R N_C}\, W_N^{\rho r N_C}\, W_N^{\kappa c N_R}\, W_N^{\kappa r} \tag{C.3}$$

Since $N_R N_C = N$, $W_N^{\rho c N_R N_C} = e^{-i2\pi\kappa r} = 1$ because κr is an integer which makes the phase angle an integral multiple of 2π. Similarly, $W_N^{\rho r N_C} = W_{N_R}^{\rho r}$ and $W_N^{\kappa c N_R} = W_{N_C}^{\kappa c}$.

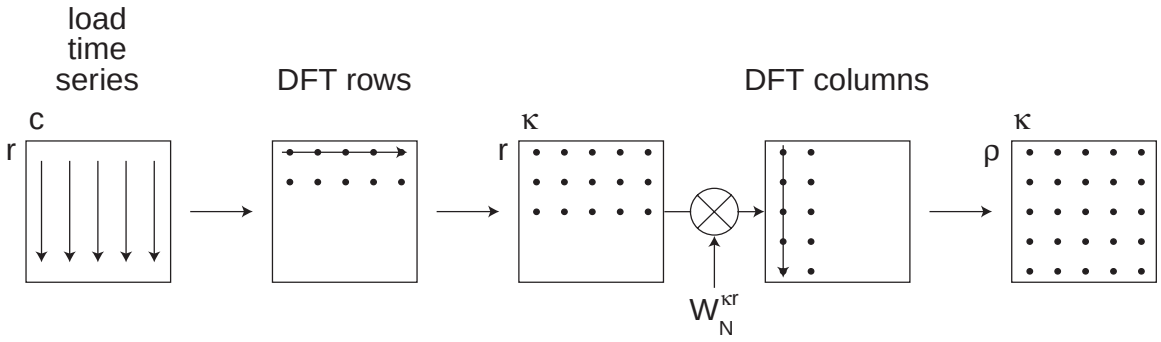


Figure C.1: Computing a large DFT via several smaller ones.

Equation C.1 then becomes

$$X(\kappa,\rho) = \sum_{r=0}^{N_R-1} \left\{ \sum_{c=0}^{N_C-1} x(r,c)\, W_{N_C}^{\kappa c} \right\} W_N^{\kappa r}\, W_{N_R}^{\rho r} \tag{C.4}$$

where we have made X and x two-dimensional according to the mapping of equation C.2. Notice that the term in curly-braces is just the N_C-point DFT of the r^{th}-row. If we call this term $G(r,\kappa)$ then equation C.4 becomes

$$X(\kappa,\rho) = \sum_{r=0}^{N_R-1} G(r,\kappa)\, W_N^{\kappa r}\, W_{N_R}^{\rho r} \tag{C.5}$$

We can define $G'(r,\kappa) \equiv G(r,\kappa)\, W_N^{\kappa r}$ where $W_N^{\kappa r}$ is called a "twiddle factor". We then get

$$X(\kappa,\rho) = \sum_{r=0}^{N_R-1} G'(r,\kappa)\, W_{N_R}^{\rho r} \tag{C.6}$$

which is the N_R-point DFT of the κ^{th}-column of the "twiddled" one-dimensional row DFTs. In summary, as pictured in Figure C.1, we can break the large DFT into smaller ones by the following process:

- Fill the two-dimensional array with the time-domain values, column-by-column.
- Take the DFT of each row.
- Multiply the entire array by the twiddle factors $W_N^{\kappa r}$.

- Take the DFT of each column.
- Read the frequency domain values out of the array, row-by-row.

If you are using a Fast Fourier Transform (FFT) to compute the DFT, the the usual bit-reverse issues apply and will have to be taken care of.

A regular FFT has a time complexity of $O(N \log N)$. Using the above process with small FFTs will yield a time complexity of

$$\begin{aligned} & O(\underbrace{N_C N_R \log N_R}_{\text{DFT rows}} + \underbrace{N_R N_C \log N_C}_{\text{DFT columns}} + \underbrace{N_R N_C}_{\text{twiddles}}) \\ = \; & O(N_C N_R[\log N_R + \log N_C + 1]) \\ = \; & O(N \log N) \end{aligned}$$

which is the same. More accurately, computing a regular FFT requires $\frac{N}{2} \log_2 N$ complex multiplications and $N \log_2 N$ complex additions. The above process requires

$$\begin{aligned} & N_C \frac{N_R}{2} \log_2 N_R + N_R \frac{N_C}{2} \log_2 N_C + N_R N_C \\ = \; & \tfrac{1}{2} N_C N_R (\log_2 N_R + \log_2 N_C + 2) \\ = \; & \tfrac{1}{2} N (\log_2 N + 2) \end{aligned}$$

complex multiplications and thus $N \log_2 N$ complex additions. The only extra work is the N twiddle factor multiplications. If $\log_2 N \gg 2$, that extra work is negligible.

Do we really need the twiddle factors?

Is it really necessary to perform the twiddle factor step? Without it, the above process is exactly the two-dimensional DFT of the array. Eliminating the twiddle factors is equivalent to multiplying the data by a phase factor of $W_N^{-\kappa r}$ before doing the column DFTs. Note that for a given column κ, this is a *linear* phase factor. The Fourier transform of a function multiplied by a linear phase factor is just a shifted version of the transform of the original function, i.e.

$$W_N^{-nm} x(n) \longleftrightarrow X(k+m)$$

so the columns will be shifted slightly. The maximum shift will occur for column $\kappa_{\max} = N_C - 1$ where the shift (due to $W_N^{-r(N_C-1)} = W_{N_R}^{-r(N_C-1)/N_C}$) will be $\frac{N_C - 1}{N_C}$ or about one bin.

Won't this shift be harmful? It can be. The shift will move spectral energy to adjacent rows generating unwanted spurs. Some of these can be quite large but, if you can live with them, you don't need the twiddles. The non-twiddled transform is still linear so the spurs do not interact and are always at predictable places.

BETA's predecessor, *META*, did not use twiddle factors and this permitted it to compute large DFTs via an efficient data flow technique. In a two-dimensional DFT it is irrelevant whether the rows or columns are transformed first. *META* fed the time domain data through a 144-point DFT and then used the output of each bin as a new "time" series, on which it performed a 64K-point DFT. This was equivalent to filling an array row-by-row, transforming the rows, then transforming the columns and finally reading the output column-by-column. This made it easy to build a multiprocessor supercomputer to do the job as little inter-processor communication was required. Special software was used to take care of the spurs.

Since implementing twiddles is not difficult in a modern system, we recommend that future designers always do so.

Appendix D

Discrete Fourier Transforms of Gaussian White Noise

A zero-mean Gaussian random variable x with standard deviation σ has the probability density function (PDF)

$$p_x(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-X^2/2\sigma^2} \tag{D.1}$$

The Discrete Fourier Transform (DFT) $X(k)$ of a complex[1] series of numbers $x(j)$ can be calculated with

$$X(k) = \sum_{j=0}^{N-1} x(j) W_N^{jk}, \text{ for } 0 \leq k \leq N-1 \tag{D.2}$$

where $W_N = e^{-i(2\pi/N)}$ (the N^{th} root of unity). Notice that the DFT is linear; each value of the output series is just a linear combination of the input values. A linear combination of independent Gaussian random variables is also a Gaussian random variable. If the input "time" series is zero-mean stationary Gaussian white noise then each sample will be independent and have the same Gaussian PDF. The output "frequency" series values will also be independent and Gaussian. Gaussian white

[1] Well, they're not *really* complex. Since the transform is linear we use real numbers to represent the "in-phase" signal and imaginary numbers to represent the "quadrature" signal.

noise transforms to Gaussian white noise.

We're interested in the square modulus of each frequency bin. Since each bin is a vector of two independent values (real and imaginary), their joint PDF is

$$p_{x,y}(X,Y) = p_x(X)p_y(Y) = \frac{1}{\sigma^2 2\pi} e^{-(X^2+Y^2)/2\sigma^2} \tag{D.3}$$

which we can convert to polar coordinates with $x = r\cos\phi$ and $y = r\sin\phi$ to get

$$p_{x,y}(X,Y) = \frac{1}{\sigma^2 2\pi} e^{-R^2/2\sigma^2} \text{ for } 0 \le R < \infty,\ 0 \le \Phi < 2\pi \tag{D.4}$$

We can use probability masses [41] via the the probability distribution function

$$P_x(X) = \int_{-\infty}^{X} p_x(X')\, dX' \tag{D.5}$$

to get

$$\begin{aligned} P_{r,\phi}(R,\Phi) &= \int_0^{2\pi}\int_0^{R} \frac{1}{\sigma^2 2\pi} e^{-R^2/2\sigma^2}\, R'\, dR'\, d\Phi' & \text{(D.6)} \\ &= P_r(R) = 1 - e^{-R^2/2\sigma^2} & \text{(D.7)} \end{aligned}$$

Since $p_x(X) = \frac{d}{dX}P_x(X)$ we can calculate the PDF of r, the modulus of the signal strength at a given frequency. It is a Rayleigh distribution.

$$p_r(R) = \frac{d}{dR}P_r(R) = \frac{R}{\sigma^2} e^{-R^2/2\sigma^2} \tag{D.8}$$

We can again use probability masses to find the PDF of the square modulus of the signal strength. If $w = r^2$, then

$$P_w(W) = P_r(\sqrt{W}) = 1 - e^{-W/2\sigma^2} \tag{D.9}$$

which yields an exponential PDF for the power.

$$p_w(W) = \frac{d}{dW}P_w(W) = \frac{1}{2\sigma^2} e^{-W/2\sigma^2} \tag{D.10}$$

n	$\Pr(w > n\lambda_w)$	n	$\Pr(w > n\lambda_w)$	n	$\Pr(w > n\lambda_w)$
1	3.68×10^{-1}	8	3.35×10^{-4}	15	3.06×10^{-7}
2	1.35×10^{-1}	9	1.23×10^{-4}	16	1.13×10^{-7}
3	4.98×10^{-2}	10	4.54×10^{-5}	17	4.14×10^{-8}
4	1.83×10^{-2}	11	1.67×10^{-5}	18	1.52×10^{-8}
5	6.74×10^{-3}	12	6.14×10^{-6}	19	5.60×10^{-9}
6	2.48×10^{-3}	13	2.26×10^{-6}	20	2.06×10^{-9}
7	9.12×10^{-4}	14	8.32×10^{-7}	21	7.58×10^{-10}

Table D.1: Probabilities of thermal events.

Time scale	Probability	Number of λ_w
Once per hour	6.62×10^{-12}	25.7
Once per day	2.76×10^{-13}	28.9
Once per week	3.94×10^{-14}	30.9
Once per month	8×10^{-15}	32.4
Once per year	7.55×10^{-16}	34.8
Once per decade	7.55×10^{-17}	37.1
Once per century	7.55×10^{-18}	39.4

Table D.2: Time scales of rare thermal events.

The exponential PDF has a mean $\overline{w} = 2\sigma^2$ and a standard deviation $\lambda_w = 2\sigma^2$. We can measure the probability of an unlikely thermal event in units of the standard deviation. Table D.1 shows the probability that a thermal event will exceed a certain number of standard deviations, i.e. $\Pr(w > n\lambda_w) = 1 - P_w(n\lambda_w) = e^{-n}$.

Applying these statistics to *BETA*, we can see how often thermal events will cause false alarms. Each spectrometer board has 2^{22} frequency bins whose contents are independent (modulo some broadening due to the window function). On the average we can expect each board to receive one $> 15\lambda_w$ event, $200 > 10\lambda_w$ events and 1400 $> 8\lambda_w$ events per spectrum. With 20 spectrometer boards monitoring the east beam (where hits will be generated), producing one spectrum every two seconds, we should see rare, strong thermal events on the time scales shown in Table D.2.

Figure D.1 shows the statistics of some genuine modulus data from one of the spectrometer boards whose input is the normal (Gaussian, white) system noise. No-

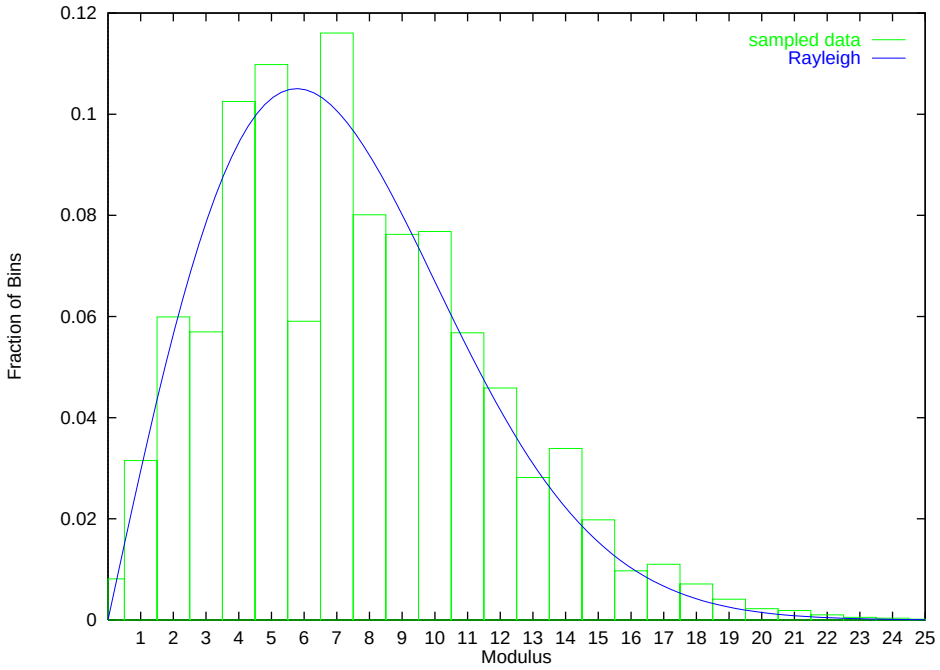


Figure D.1: Histogram of spectrometer magnitude output compared with a Rayleigh function.

tice that it is very similar to a Rayleigh distribution, as predicted by equation D.8. There are some significant differences, however, which are caused by the use of integer arithmetic. This is analyzed in detail in Appendix E.

Appendix E

Easy Lookup-Table Computation using Scalar Quantization

One method that we have used for fast computation in *BETA* is the lookup-table. While this can be extremely fast, it is unwieldy (or impossible) for a large number of input bits. In this particular problem we used a unique non-uniform scalar quantization technique [25] to lower the number of input bits sufficiently for table lookup.

The problem was to devise an inexpensive ROM-based scheme to compress a pair of 20-bit signed Fourier amplitudes (I and Q) into a single 16-bit unsigned Fourier magnitude, i.e. $\widetilde{M} \approx \sqrt{I^2 + Q^2}$. We also needed to minimize the worst-case fractional error $(M - \widetilde{M})/2$. Our method was as follows:

1. Begin with "saturation logic" to truncate 20-bit signed I and Q to 16-bit signed integers (a "sign de-extend"), then strip the sign to yield 15-bit unsigned integers.

2. Now compress the 15-bit integers (n) to 9-bit representations (p), minimizing the worst-case fractional error upon inversion (i.e. $(\widetilde{n} - n)/n$).

3. Combine the I and Q p-representations to generate an 18-bit address into a

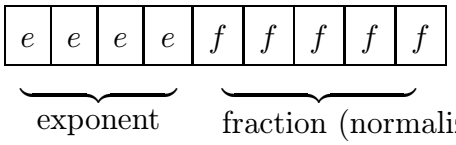
256K×16 ROM lookup table, containing the 16-bit magnitudes, namely

$$M_{16}(2^9 \cdot p(I) + p(Q)) \tag{E.1}$$

The heart of the problem was to find the best way to accomplish step 2 – compressing a 15-bit unsigned integer n to a 9-bit representation p, with least (worst-case) fractional error. We investigated three possibilities:

1. Floating point

We would represent p as an exponent and normalized mantissa,

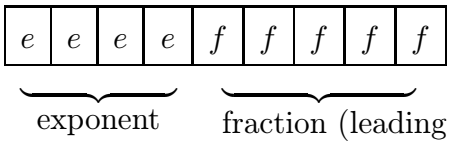
p: 

i.e. $n = 2^e \cdot (fffff)_2$.

This method has a worst-case step size (upon inversion) of $2^{-4} \approx 6.3\%$.

1a. Floating Point with "hidden" bit

The representation is the same as above, but without storing the the MSB of fractional part since we can assume that it is 1 (similar to IEEE floating point).

p: 

i.e. $n = 2^e \cdot (1fffff)_2$.

This method has a worst-case step size of $2^{-5} \approx 3.1\%$.

Note that floating point is wasteful because only 10 of 16 exponents are used. It also has unequal step sizes, ranging over a factor of 2.

2. Logarithmic

In this method p is a logarithm representation of n, i.e. $p = \mathrm{rnd}[\alpha \ln(n+1)]$. We want α so that

$$n = 0, 1, \ldots, 32\mathrm{K} - 1 \quad \longrightarrow \quad p = 0, 1, \ldots, 511$$

so we choose $\alpha = 49.15$:

$$p = \mathrm{rnd}[49.15 \ln(n+1)] \longleftrightarrow n = \exp(p/49.15) - 1$$

then $n(p+1)/n(p) \to 1.021$ and the worst-case step size is 2.1%. This is better than floating point with "hidden bit", but it is still wasteful because of missing codes:

n	0	1	2	4	8	16	32	64	128	256	512	1K	...	8K	16K	32K
p	0	34	54	79	108	139	172	205	239	273	307	34	...	443	477	511

missing, wasted codes

3. Hybrid

This method is a hybrid scheme which is linear until the step quantization equals the fractional error of the logarithmic method, then it continues logarithmically, i.e.

$$p = \begin{cases} n & n < n_0 \\ \mathrm{rnd}[\alpha \ln n - \beta] & n \geq n_0 \end{cases}$$

with n_0 chosen to minimize worst-case step size (and α, β chosen to make the mapping continuous and have the proper range and domain). Figure E.1 shows this graphically. For 15-bit n and 9-bit p, $n_0 = 70$, $\alpha = 71.7$ and $\beta = 234.6$ and we get...

n	0	1	2	...	70	71	...	100	...	128	...	1K	...	8K	16K	32K
p	0	1	2	...	70	71	...	96	...	113	...	262	...	411	461	511

and the worst case step size is 1.4%.

For each of these lossy compression methods, the worst-case *error* is half of the

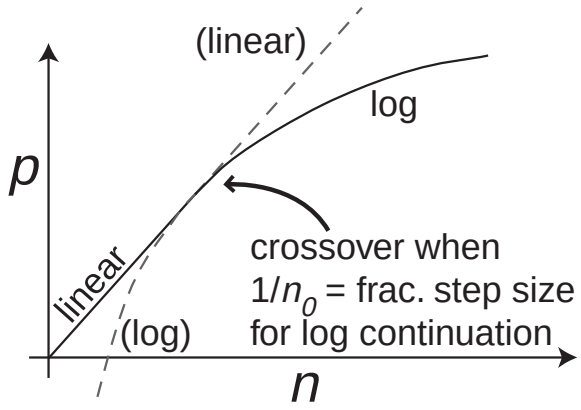


Figure E.1:

worst-case step size. The conventional floating point format (with hidden bit) produces 1.6% error, a pure logarithmic format produces 1% error and the hybrid logarithmic scheme produces 0.7% error. We believe the latter to be nearly optimal for compression of our integer data. All three methods are easily implemented in a small ROM lookup table. We chose method three and dubbed it *pcode*.

Numerical Considerations

Of course, computing using integer arithmetic and compressed values will degrade the results. Figure E.2 shows a histogram of the values in the *p2mag* ROM of the FFT boards. This is the ROM which computes the amplitude of a frequency bin value from the pcode-ed I and Q components (as described in step 3 at the beginning of this appendix). Notice that the higher values appear as a series of separated spikes. This is caused by the pcode compression: above the log/linear threshold, many I and Q values map to a single pcode value. The spikes correspond to the moduli of those few pcode values. Without the compression technique, the spikes would be spread out evenly and the histogram would be quite smooth.

Figure E.3 shows a close up view of the same histogram, which gives a better idea of its overall shape. Note that it is very similar to a Rayleigh distribution. It would *be* a Rayleigh distribution if the inputs to the ROM were weighted by a Gaussian

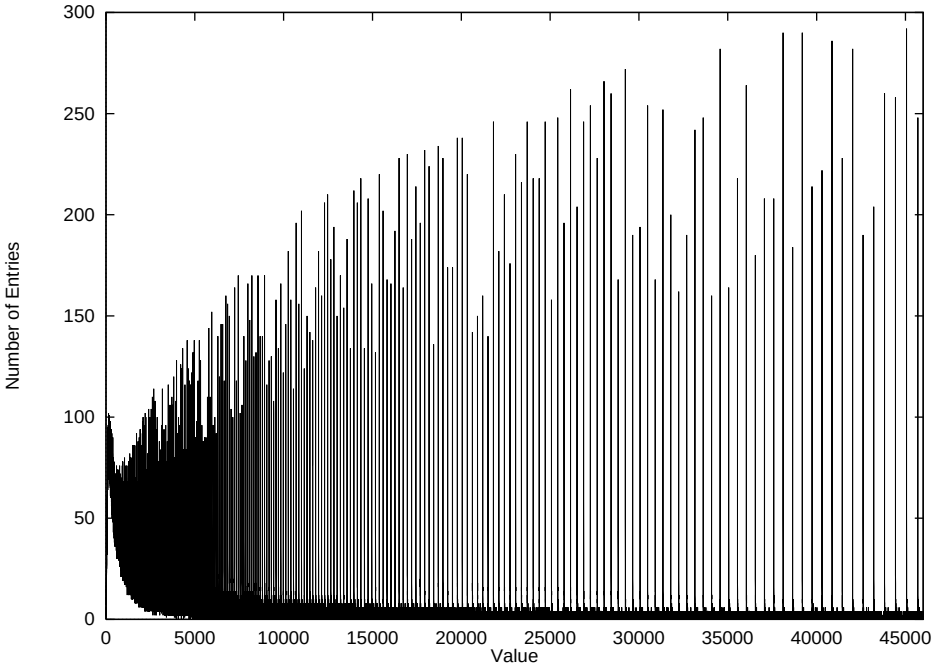


Figure E.2: Histogram of the values in the *p2mag* ROM of the FFT boards.

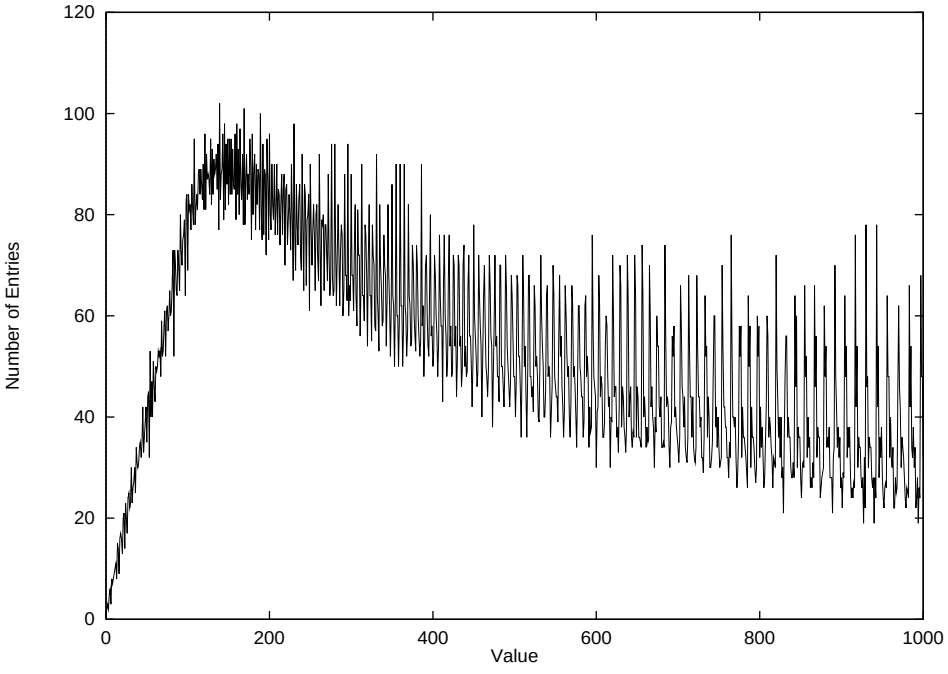


Figure E.3: Close up view of the same histogram, showing the Rayleigh-like overall shape.

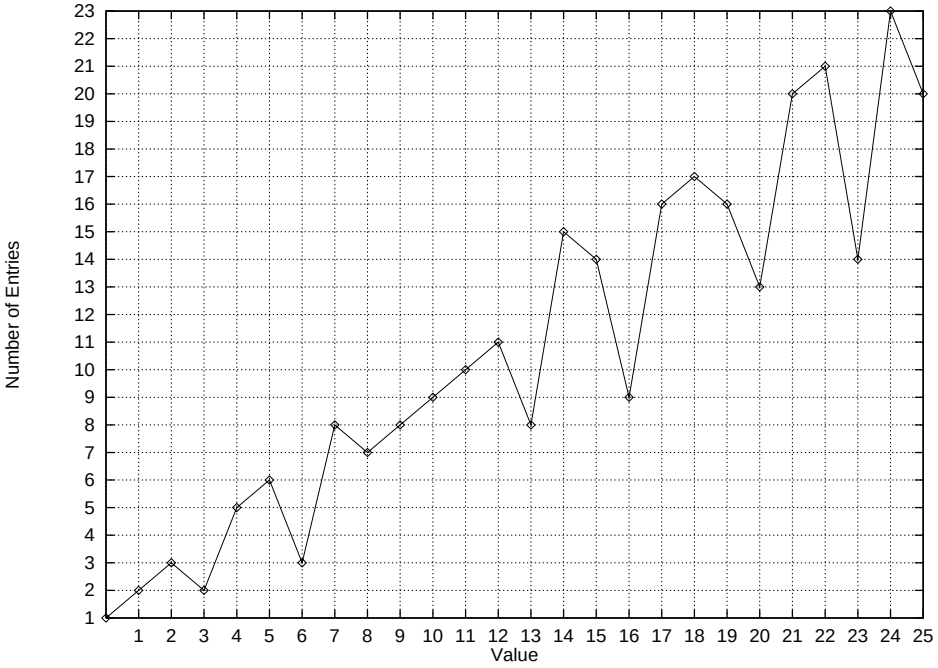


Figure E.4: Extreme close up view of the same histogram, showing the detail at the very beginning.

distribution. But they are not since the ROM is only responsible for mapping values; the Gaussian weighting of the inputs is a property of the noise itself.

We noticed an example of degradation in Figure D.1 where the histogram produced some values which were significantly different from the calculated Rayleigh distribution. Since the values involved were all less than the pcode linear/log threshold, the compressed values were identical to the original ones, so the pcode cannot be held responsible for this. Figure E.4 shows a close up view of the very beginning of the *p2mag* ROM histogram. Notice the "jaggy" look it has. If you compare the places where Figure D.1 deviates significantly from the Rayleigh distribution (values 3, 6, 8, 13, 16, ...), you will notice that Figure E.4 has strong "dips" there. These are caused by the use of integer arithmetic in computing the moduli of integer-valued vectors. Figure E.5 shows the truncated-to-integer moduli of the vectors whose components are the x, y values along the axes. Figure E.6 shows a histogram of these truncated values. Notice that it also has a jaggy appearance, but does not look exactly like the

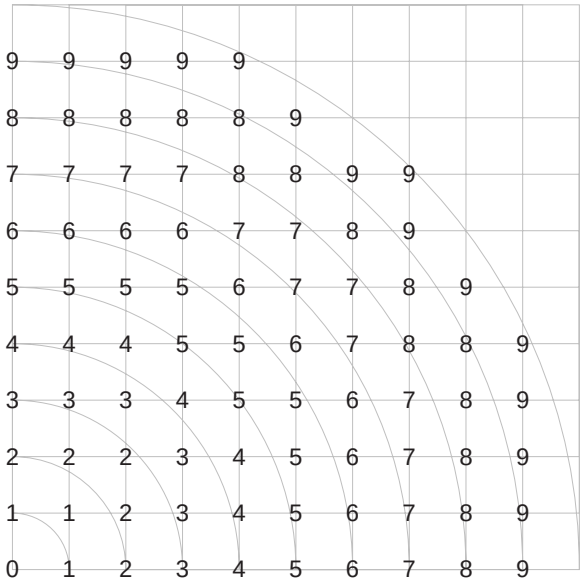


Figure E.5: The locations of the truncated-to-integer modulus values in the first quadrant.

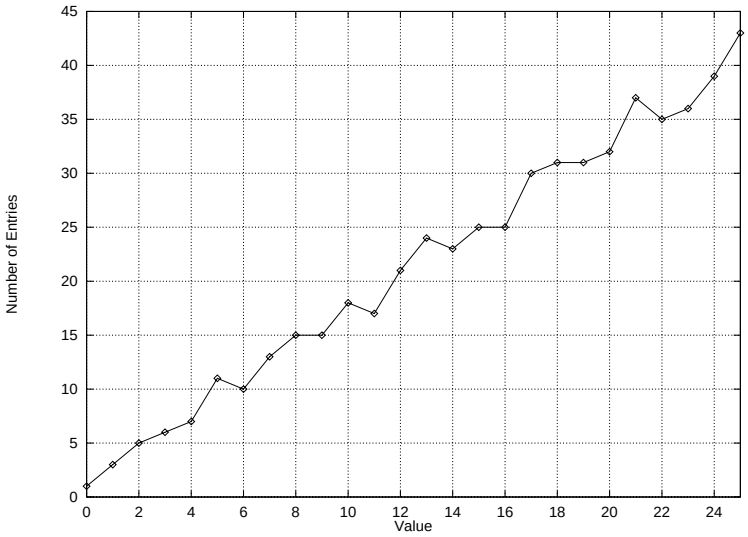


Figure E.6: Values of the truncated-to-integer modulus function. Notice the resemblance to the beginning of Figure E.4.

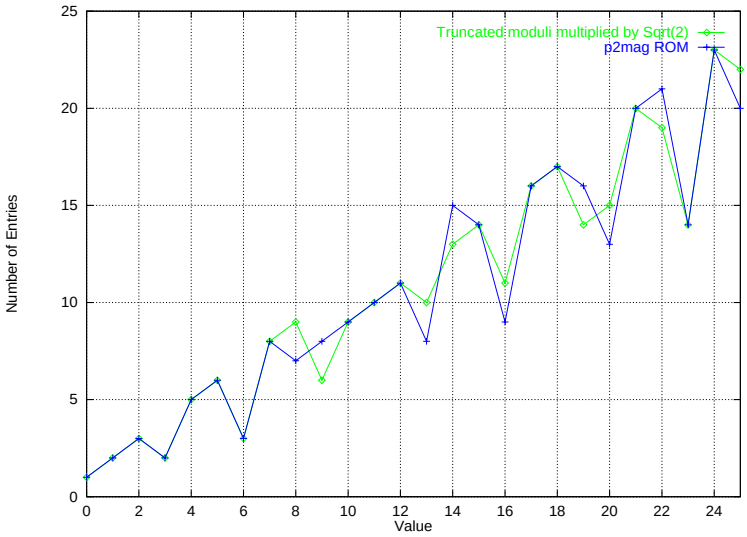


Figure E.7: Comparison of the *p2mag* ROM and scaled, truncated modulus histograms.

ROM histogram. For various reasons, the ROM histogram multiplies its values by $\sqrt{2}$, and when we do this and plot the histograms together, we get Figure E.7. These are identical at the very beginning and very close otherwise. The discrepancies are due to minor scaling and round-off issues in the *p2mag* ROM.

Appendix F

FFT Simulation Results[1]

We devoted considerable time to simulating our FFT architecture, using a simulator ("FDPSIM") supplied by Austek. In particular,

- We needed to verify that the 3-chip architecture is numerically correct.
- We wished to determine the effect of word size quantization and roundoff on the spectral dynamic range.
- We wished to determine the effects of finite twiddle factor word size ("depth") and quantization ("width"), since the use of full-size twiddle factor ROMs (e.g., 20 bits by 1M points) would raise costs considerably, and
- We wished to verify by actual numerical simulation that a weak sinusoidal signal embedded in wideband noise, in the presence of additional strong sinusoidal signals, could be reliably and accurately detected by the FFT system we intended to build.

We now describe the results of these simulations in detail, because there appears to be considerable confusion (and not a little folklore and mythology) in the signal processing community on precisely this issue. We hope the reader will be as surprised and enlightened by these results as we ourselves were.

[1]This section includes work by Paul Horowitz, Greg Galperin and Derrick Bass.

Austek's simulator was never intended for such large transforms, and, as supplied, it took nearly a day to complete a single 4 megapoint transform on a Sun SPARC-2 (it took many days on a '486-type PC). We began by verifying the numerical correctness of the 3-chip architecture on scaled-down transforms, then proceeded to modify some of the simulator's modules to speed up performance. The improved code performs a 4 million point simulation in 2 hours.

The first simulations verified that a suite of sinusoidal waves, chosen with frequencies relatively prime but with each sine "on-bin" (i.e., an integer multiple of the lowest FFT frequency, $1/T_{\rm fft}$), and covering a range of amplitudes, was properly resolved by the FFT, when using "perfect" (double precision floating point) arithmetic and twiddle factors.

We then explored the effects of finite word length in the FFT computation, in particular the 16-bit and 20-bit integer options that can be set by the initial command register load of the A41102. The results can be summarized as follows: With all "scales" enabled (i.e., with a 1-bit right shift of data following each FFT butterfly, required to prevent word growth for coherent frequencies present in the initial time series) the effect of finite word length and arithmetic precision is to introduce a "numeric noise" into the spectrum, consisting of an average of 1 LSB fluctuations in the final spectral amplitudes (and a *peak* fluctuation of 1.4 LSB, i.e., 1 LSB in each of the real and imaginary components). Stated this way, the result is independent of word length. It may seem surprising that a 2^{22}-point integer computation of the Fourier Transform introduces so little roundoff error; but the effect of the successive scale-by-2's is to keep pushing the roundoff error off the right end of the word.

The numeric noise is, of course, to be compared with any periodic signal present in the digitized input. A single tone, present as a full-scale on-bin sinusoid in the digitized time series, produces a full-scale output. Thus a 16-bit "all-scale" integer transform has a dynamic range of 2^{15} in amplitude (90dB); as we will soon see, however, there are several effects that can introduce spurious responses in the spectrum from a single sinusoidal input. These "spurs" can be important in SETI, because a single interfering signal may produce a set of false responses in addition to the obvious

large peak. Some of these effects are finite input quantization (i.e., shorter than the computation word size), spectral "leakage" (signal "off-bin"), and truncated twiddle ROM (both in word size and argument step size). We discuss these in the following paragraphs.

The effect of omitting some of the right-shift scales is interesting: Most obviously, one introduces a risk of numeric overflow – a full-scale input sine causes overflow if any scales are omitted, a half-scale input causes overflow if more than one scale is omitted, etc. I.e., the spectral amplitudes grow by a factor of 2 for each omitted scale. This seems obvious, and in fact one might easily conclude that roundoff error grows the same way. However, the situation is more complicated – it turns out, as revealed by our simulations, that the *peak* numeric noise grows as expected (a factor of 2, or one bit, in amplitude for each missing scale), but the *average* numeric noise amplitude grows only as the square root of the number of omitted scales (1/2 bit per missing scale). Thus one can squeeze some extra average dynamic range out of an integer FFT by omitting some scales, at the risk of numeric overflow (if a large signal is present); but note that the dynamic range relative to the *peak* numeric noise is not improved.

We studied the effects of ROM truncation next. To set the stage, note that a "full-sized" complex twiddle ROM (4M $\times$ 20 bits, say) would require 40 4-megabit ROMs, at that time priced at about $30 each, thus approximately doubling the parts cost! Of course, one need not store both sine and cosine (factor of 2 savings), and one need store only a quarter-sine table (another factor of 4); that puts the ROM cost at about $150. Even at that price the ROMs are a significant portion of the board cost, so it is worth asking how wide and deep the ROM needs to be.

We ran a set of simulations, and learned the following:

1. The spectral amplitude of an on-bin sinusoidal signal is very little affected by rather extreme ROM wordsize truncation; in particular, 8-bit ROM amplitudes affect spectral amplitudes by less than 1%.

2. ROM "width" (number of table entries) can also be reduced substantially, with

almost negligible effect upon signal amplitude, but with production of spurs that are absent when using a full-sized ROM.

3. If ROM width is to be reduced for a non-square corner turn, truncate the larger address first.

Further explanation of 2 and 3: Our 4M-point transform is implemented as 128×128×256, with a "small" (16K) twiddle factor multiplication following the first corner turn, and a "large" (4M) twiddle factor multiplication following the second corner turn (there are, in addition, a pair of 4M corner turns, without twiddle multiplication, at both ends of the overall FFT). The small twiddle factor ROM is cheap, and no truncation is needed there. The second ROM is the issue. It is a 16K×256 corner turn (the initial 128×128 transform pair is exactly equivalent to a single 16K transform), requiring 14 and 8 address bits, respectively (22 address bits for a full-size ROM: 4M coefficients). We found that one can use a 16-bit (amplitude) ROM with 8 bits by 8 bits of address (256×256, or 64K complex coefficients, a factor of 64 less than a full-sized ROM) with no loss of signal amplitude, but with production of spurs whose peak amplitude is -54dBc (dB relative to the "carrier," i.e., the sinusoidal signal). That peak spur occurs at $f_0 \pm 16$K bins, with additional spurs at multiples of 16K bins offset, dropping at 6dB per 16K bins of offset. The peak spur amplitude depends on ROM width truncation, dropping 6dB per additional address bit used (e.g., a 512×256 addressed ROM – 512K complex coefficients – has a peak spur amplitude of -60dBc, again at ± 16K bins offset from the carrier). These results are independent of the FFT computation word size, i.e., identical for 16-bit and 20-bit integer FFTs.

There are several possibilities for calculating the precise coefficients in a truncated ROM; for example, should each entry be the *average* of the multiple "true" coefficients for which that entry substitutes? or perhaps it should be simply the exact coefficient corresponding to the (smaller) FFT for which the ROM is full-sized. We explored this question, trying what we called a "mean ROM" and an "expanded ROM," respectively; we also considered a "median ROM" and a "topographic center ROM". The

result of simulation showed that it hardly matters, but where there is a difference the expanded ROM is better. For example, "spurs" of a pure dc input signal are identically zero for the truncated ROM constructed as an expanded ROM, whereas for a mean ROM they are the same size as the carrier spurs (which we might call "ac spurs") described in the previous paragraph, e.g., −54dBdc for a 4M-point FFT using a ROM containing 64K complex coefficients. (These spur amplitudes are for ROMs of 16-bit precision, by the way, whether doing a 16-bit or 20-bit FFT computation. Using instead a truncated ROM of perfect precision has no effect on spur level, which is caused entirely by the ROM's truncated "width.") This perfect suppression of "dc spurs" when using an expanded ROM is less than meets the eye, by the way: when the input data is multiplied by a window function (to reduce spectral "leakage"), as must be done in the real system (see below), the dc spurs reappear, at the canonical level specified in the previous paragraph.

Our next set of simulations involved the addition of uncorrelated Gaussian noise to a discrete array of pure sinusoids, in order to determine how much precision (word size) is needed to ensure that the spectrum of input noise dominates over "numeric noise". This is clearly word-size dependent, since the amplitude of numeric noise in an all-scale FFT equals the LSB (independent of word size), whereas the amplitude of the spectrum of random noise approximately equals the input amplitude reduced by a half bit per butterfly. This estimate suggests that 16 bits is marginal in an all-scale FFT, because even if the input noise level is set to the full-scale amplitude of 2^{15} (a radical approach, allowing no signal headroom), it will emerge in the spectrum at an amplitude of 2^4 after the 22 butterflies of a 4 megapoint FFT; that is +24dB relative to roundoff ("dBr"). The corresponding figure for a 20-bit word size is an output noise amplitude of 2^8 (+48dBr). Of course, one cannot set the input noise amplitude to full scale without severe clipping, owing to the high crest factor of white noise; thus these figures should be reduced by a factor of at least $\approx$ 10dB.

The purpose of the noise simulations was to quantify these estimates of the dynamic range, in the output spectrum, of input noise (call it "antenna noise") over roundoff noise ("numeric noise"). We carried out many simulations, with the follow-

ing result: If the input noise amplitude is set so that approximately 1 sample in 4 million saturates at full scale (call this "full-scale noise"), then the RMS noise amplitude in the spectrum that results is approximately 65 (for 20-bit integer arithmetic, with unwindowed input), or 4 (for 16-bit arithmetic). Thus for full-scale noise, the antenna noise in the spectrum has an amplitude +36dBr for a 20-bit computation, +12dBr for a 16-bit computation. In practice one would probably set the input amplitude some 6dB or so below full-scale noise, implying that a 4M-point all-scale FFT must be done at 20-bit precision (or better). The only way to survive with a 16-bit transform is to omit some intermediate scales, a perfectly reasonable (though less conservative) approach.

At this point we decided to use a 20-bit word size, and performed all further simulations at that precision. We next experimented with the "dithering" effects of combining wideband noise with input data, the sum being quantized to 4 or 8 bits. For 8-bit data quantization, dithering increases the dynamic range to $\approx$ 90dB (from the 48dB of an undithered 8-bit quantization); however the quantization is a nonlinearity that produces harmonics at the ≈ -48dBc level (in addition to the -60dBc spurs caused by a 128K truncated ROM). For 4-bit input quantization the harmonic spurs are far worse, approximately -20dBc, even though dithering continues to provide a wide dynamic range; 4-bit quantization thus appears an unwise choice for SETI.

Appendix G

Good Window Hunting[1]

We experimented with various window functions to find ones with the best characteristics for our application. A "window" is jargon for a multiplicative function applied to the input time series for the purpose of reducing sidelobes and leakage: If an FFT is applied to an unwindowed input time series, the finite data length corresponds to multiplication of a continuing time series by a rectangular function (of length equal to the transformed data frame), thus producing in the frequency domain (by the convolution theorem) the convolution of the proper sampled spectrum with a sinc (that's shorthand for $\sin(x)/x$) function, the transform of a rectangle. For an "on-bin" signal (i.e., a sinusoid whose period is an integral submultiple of the transformed time series length) all off-signal bins lie at zeros of the sinc function, producing an accurate spectrum with no sidelobes or leakage; but that is a rare case, and in general one sees sidelobes and signal leakage corrupting the spectrum. The usual cure is to use a multiplicative window function, of unit amplitude at the center of the time series and generally tapering to zero at the ends of the time series (in optics the 2-dimensional analog is known as "apodizing"). The simplest example is the triangle (also called "Bartlett"), but there are literally dozens of contenders for "best window function," named after the famous (and not-so-famous), such as Hanning, Blackman, Dolph-Chebyshev, etc.; for an excellent review see the article by Harris [21]. In general, one

[1]This section includes work by Paul Horowitz, Greg Galperin and Derrick Bass.

trades off improved sidelobe rejection for a broader central response in the frequency domain.

The rectangular window (i.e., no window at all) is a disaster, with peak sidelobe of -13dBc, and slow falloff of sidelobe with offset from the spectral peak (-6dB/octave). At the other extreme, the Blackman-Harris "minimum 4-sample" window has peak sidelobe level of -92dBc, bought at the expense of a factor of ≈ 2 decrease in spectral resolution (i.e., the response to a pure sinusoid is a peak that spans perhaps 4 or 5 frequency channels before it has fallen off by 30dB). We wished to look at windows because

- We needed one, and wanted to choose rationally, and
- We wanted to see if windowing had side effects on the parameters already simulated (e.g., peak spur level, average noise level, headroom, etc.).

The results are approximately as expected: Windows have negligible effect on spur levels relative to signal amplitudes (because both are similarly affected by the window), etc., and they have the predicted effect on resolution. The average noise level, for "full-scale" input noise, is reduced by about 3dB or so, owing to the reduction of average signal level by the window; this makes the choice of 20-bit arithmetic mandatory, if an all-scale transform is used. Finally, signal amplitudes are reduced by a few dB, relative to numeric noise; this is unimportant, because the system is designed so that antenna noise dominates numeric noise (by some 20dB or more).

The major effect of windowing is to reduce leakage and sidelobes. We tested three windows, namely Hanning (von Hann: a cosine-squared), Blackman-Harris, and triangular (Bartlett), in comparison with a uniform (no-window) window. Of these, the Blackman-Harris has the lowest peak sidelobe level (-92dBc, falling 6dB per octave offset from the carrier), while the Hanning has only a modest peak sidelobe level (-32dBc) combined however with very rapid falloff away from the carrier (-18dB/octave); the advantage of the Hanning, of course, is a narrower main lobe (1.6 times the width of the uniform window, versus 2.1 for the Blackman-Harris). The choice is not absolute, but depends very much on the nature of the signals and

interference expected. For example, if interference is often modulated with audio bandwidths (a few kilohertz), the Blackman-Harris's precipitous drop to −70dBc is of no benefit, and its broader central lobe (its response to a narrowband ETI beacon) thus makes it a poorer window. What is needed in this case is a window that confines spectral "splatter" to a handful of contiguous channels, which the Hanning's rapid falloff adequately achieves; thus for this application it is a superior window to the Blackman-Harris because of its superior resolution and sensitivity.

On the other hand, if one is dealing often with interfering *carriers*, the Blackman-Harris is the better window, since it keeps the signal within just a few channels before it falls below the antenna noise continuum. Although the choice is not critical, we believe that experience with the system will dictate which window is better. Thus we designed the hardware to permit run-time selection – we loaded both windows into the (small) window ROM, selected via a downloaded segment address. Our simulations of windows showed, incidentally, that the window ROM can be truncated enormously with no observable effect: the "full" 4 million coefficients can be replaced by an 8K×16 expanded ROM (512 times smaller). Thus a single 27C1024 (64K×16), which cost less than ten dollars at the time of construction, can hold 8 window functions.

Based on the simulations just described, we chose the following parameters for the 4M-point FFT: 8-bit data quantization (in the mixer-filter-digitizer module), an 8K×16 expanded window ROM, 20-bit integer arithmetic with all scalers (or all−1) enabled, a full (16K×16) small twiddle ROM, and a $512 \times 256 \times 16$ expanded large twiddle ROM. Figure G.1 summarizes the behavior of the FFT with regard to signals, noise, roundoff, and spurs, and Figures G.2 and G.3 demonstrate the output data from a pair of simulation runs: The "signal" consists of wideband antenna noise to which has been added three large sine waves (amplitude 0.1, at channels 1M, 2M+0.25, and 3M+0.5) and, nearby, three weak sine waves (amplitude 0.001, at channels 1M−8, 2M+20.25, 3M+20.5). In both cases the data has been quantized to 8 bits, and transformed with 20-bit arithmetic in an all-scale FFT using a 512×256×16 expanded large twiddle ROM. The output table prints complex pairs, 4 to a line, beginning at the labeled channel number. Note that only certain selected regions of interest have

been printed (about 0.01% of the full 4M complex output channels). In Figure G.2 we have used a uniform window: the on-bin signal at 1M is cleanly resolved (one bin, all real), allowing clear detection of the nearby signal (at 1M−8); but the off-bin signals (at 2M+0.25 and 3M+0.5) are broadened by "spectral leakage" to more than the 68 contiguous channels shown, burying the nearby weak signals. In Figure G.3 we have used the same data and transform parameters, but with a Blackman-Harris window (truncated to 16K×16). Now the on-bin signal at 1M has been broadened to a half dozen channels, somewhat degrading its detection, but the weak signal at 1M−8 is still cleanly resolved. More importantly, the off-bin signals at 2M+0.25 and 3M+0.5 are of comparable width, dropping below the antenna noise level at ±3 bins – the nearby weak signals (at 2M+20.25, 3M+20.5) now show clearly! Given that only a small fraction of real-world signals are on-bin, the wisdom of windowing should be apparent.

In these tables, the small peak near 4M is a harmonic spur of the strong signal at 2M, at a level a few times the average antenna noise level. The spectral regions offset by 16K from the large peaks have been listed because the worst ROM-truncation artifact occurs there; even with these relatively strong signals the spur does not rise above antenna noise (it is −60dBc, corresponding to an amplitude of about 15 in Figure G.3).

Figure G.4 shows the actual performance of the 4M spectrometer board and its windows during a two-tone test. We've put a strong carrier exactly midway between bins ("mid-bin") and a much weaker carrier (40dB down) just 10 Hz away. At this resolution that's 21 bins away, again mid-bin. In Figures G.4a and G.4b we've done the FFT unwindowed (euphemistically called a "rectangular" window), displaying the serious spectral "leakage" resulting from the convolution of a *sinc* function in the spectral domain (the FFT of a rectangular impulse). Because the weak sine combines coherently with the spectral leakage tail of the stronger signal, the rotation of their relative phases produces a non-stationary combined amplitude, illustrated in Figure G.4a (best relative phase) and G.4b (worst relative phase – nearly complete cancellation of the weaker signal). In the latter, the weak sine wave is lost in the

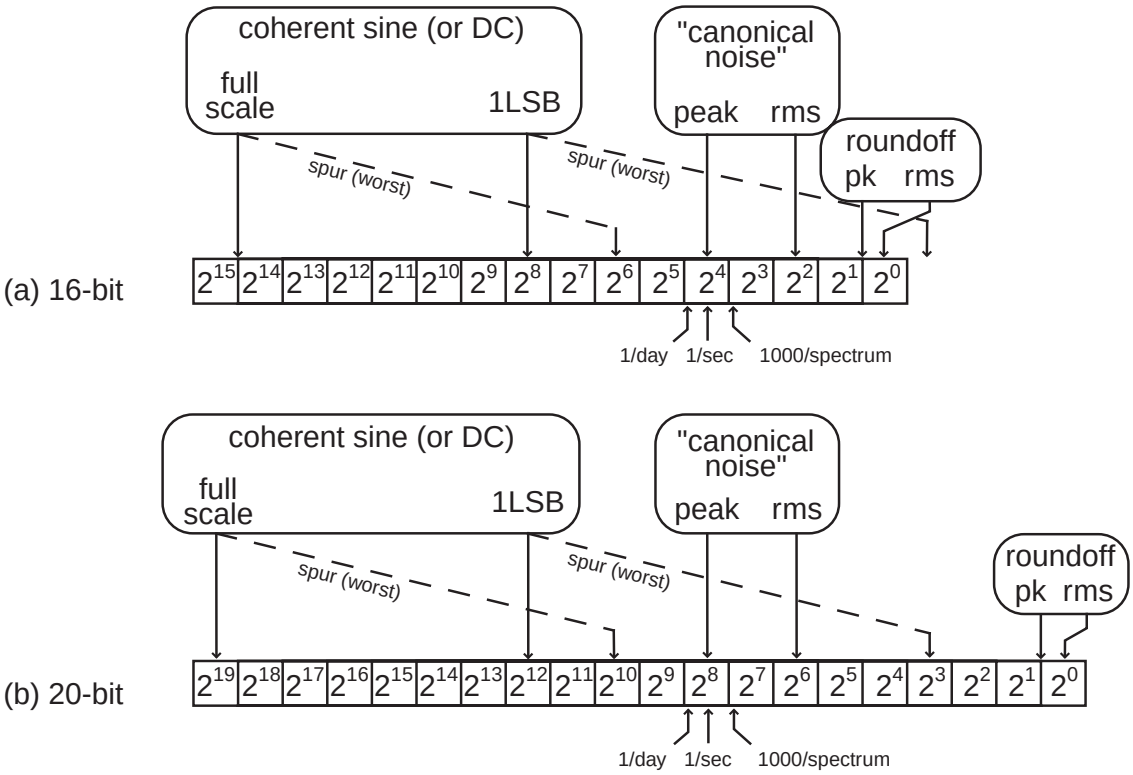


Figure G.1: 4-million-point integer FFT behavior with regard to coherent signals, noise, roundoff, and spurs. 8-bit I and Q input amplitudes are assumed, left-justified in the FFT's word, with one righthand bit shift per butterfly, and a $256 \times 256 \times 16$ "expanded ROM" for the large twiddle factor. (a) 16-bit integer arithmetic; (b) 20-bit integer arithmetic. The diagrams show the location in the output spectral amplitudes that the indicated inputs emerge. For example, a full-scale input sine wave, whose period is commensurate with the transform window, produces a full-scale output peak in the corresponding frequency bin. "Canonical noise" is input ("antenna noise") Gaussian white noise of amplitude such that approximately one sample in 4 million would overflow full scale (and is forced to saturate at full scale); its modulus has a mode of approximately 20% of full scale. "Spurs" are spurious spectral responses to genuine sinusoidal components in the input time series, caused primarily by the twiddle ROM truncation; each doubling of ROM size reduces them by 6dB; our design uses ROMs twice as large as assumed here, hence produces worst-case spurs that are shifted one bit to the right of the positions shown. Omission of bit shifts between butterflies affects all signals and spurs linearly; however, the effect on roundoff noise is different – although the *peak* roundoff noise grows linearly with omission of bit shifts, the *rms* roundoff noise grows only as the square root. The three arrows below the boxes point to the thresholds that produce the indicated "hit" rate, assuming random noise, and a 4-million point FFT every 2 seconds.

```
1000000   -41     8    -6     8    25   -23   -33   3000000   -3   32724     3  -32754   -11  -10003   -32  -6533
2000000     5    42     2     7   -37   -29   -10   3000004   22   -4647    25   -3650    15   -2967     8   -2476
3000000   -17   -25    21   -11    15   -31    42   3000008    7   -2179   -18   -1931     7   -1732   -18   -1556
4000000   -30    -8    12   -17    31   -11    14   3000012  -19   -1426   -12   -1304     3   -1168    37   -1073
5000000   -10    13   -11     3    57   -24    23   3000016   -9   -1014   -31    -900    19    -879    -7    -784
6000000    17   -19     1    36    -2   -24    -4   3000020   17    -508   -18   -1141    -9    -879    11    -826
7000000    -5     3     9    -7   -12   -19    46   3000024  -12    -750    10    -686    19    -629     7    -633
8000000    14   -17   -28    -7   -23   -19    46   3000028   -9    -627    13    -667   -12    -558    41    -550
9000000    -1    16   -20    11    20    14     1   3000032   22    -527   -18    -518   -13    -515   -38    -452
9999992   503   -31    27    13    -3    10   -11   3000036   -3    -447    23    -467     3    -474    -5    -457
9999996    11   -28   -14    17   -40    18    19   3016372   25      -3    22      29     1     -11    17      -2
10000000 51422   -20     -4   -25    18    28   -17   3016376   31      10    -9      27   -74     -20    45      -9
10000004   -23     5    -4    24     4     9    -3   3016380    2      -3    15     -13    29     -19    26     -21
10000008   -14    -1   -15    24     3    -9   -17   3016384  -28      26    21      43   -11       9    -37     -11
10163372   -18     3    -7    -4     8    16   -35   3016388  -67       1    -8     -71    23      -7    25      -5
10163376   -10     5    41    22     9     0    -3   3016392   39       8   -18      -9   -15      -9    -5      13
10163380   -12     1    26    27   -18    10    15   3016396   -6      23   -13      -1    20     -27    15      -5
10163384     1    -1   -21   -55   -38    15    21   3100000  -19      11    19       5   -17     -40    -7     -13
10163388     1    -9    41    23     5   -16   -31   3200000   -8     -35   -11     -24    18       3    28      37
10163392    15    14    11    18    16    -6    -9   3300000   -4     -11   -11      73    18       1   -11     -23
10163396    -5    -7   -49     6     6   -53    -7   3400000   -1      -1    -1     -14    35      -1   -19       7
10100000   -36     3    -7   -15    30   -22    -4   3500000  -11      23    31      27    45     -33   -78     -13
10200000   -22   -43   -37   -43    35    17   -15   3600000   19       5    33      13    45      28    39      22
10300000    12    -1   -12   -21    35    11   -19   3700000   19       9     6      -3   -31      42    39      19
10400000     3    -1   -40   -21    29   -31    -5   3800000  -15      21   -37     -27    -5      -6    10      -7
10500000    -5   -26    -9   -32     0    -1     3   3900000  -17      -1   -47      -7   -61      -7    -1      -7
10600000    17    29    17    60    13     6    -3   3999992  -19      -5    48      21   -27      31   -31     -42
10700000    -4    55   -47   -23    -5   -13   -33   3999996  -33     -27   -15      68   -22       0   -37      19
10800000     9   -37   -21    35   -11   -49    21   4000000  -15      57   138     -197    39     -37   -15      -4
10900000    37     1    30     4    -3    38    20   4000004  -111      6   -3       21     1      19   -53     -11
10999972   270   277   283   286   330   305   323   4000008   -3       7     5     -10    30      16     3     -30
10999976   333   321   311   395   370   337   394   4016372   -6      17    52      32    18      11    -2      22
10999980   423   425   451   439   441   451   469   4016376  -66      19    -9      -1    10     -17     0      -6
10999984   509   540   507   525   552   633   615   4016380   -7      42    37      -2    12      23    14     -15
10999988   700   733   741   783   816   791   932   4016384   18     -41    37       1     7      17     8      25
10999992   979   986  1107  1103  1319  1284  1538   4016388   -4     -38   -12      27   -34      23    -3      -39
10999996  1871  1920  2531  2574  3607  3663  6543   4016392   15      27     1      33    16     -21    24       6
20000000 32825 32779-10941-10937 -4646 -4711 -2955   4016396    5      13    26       6    23      19    -3      -21
20000004 -2170 -2165 -1743 -1697 -1433 -1434 -1173   4100000   33       2   -23     -20    35      17    43       17
20000008 -1040 -1075  -903  -920  -817  -899  -711
20000012  -731  -631  -639  -619  -582  -555  -486   window: Uniform, 16K x 16, expanded to 4M
20000016  -538  -487  -475  -459  -443  -413  -348
20000020  -107   -53  -525  -493  -421  -427  -373   data: noise (0.5 pk amp) + 3 sines (each 0.1 amp) @ 1M, 2M+0.25, 3M+0.5
20000024  -355  -391  -325  -367  -348  -329  -298         + 3 nearby weak sines (each 0.001 amp) @ 1M-8, 2M+20+0.25, 3M+20+0.5
20000028  -312  -289  -241  -259  -242  -251  -294           the sum quantized to 8 bits (i.e., each weak sine is 1/8 LSB amplitude)
20000032  -295  -259  -259  -282  -293  -229  -255
20000036  -231  -227  -218  -257  -189  -211  -187   fft: 20-bit, all scale
20163372    28   -25    33   -24     7   -29    43
20163376   -11   -18    41    -2   -24    -8   -37   twiddle ROM: 128K x 16, expanded to 4M
20163380     0     1   -24   -45    -6    47   -23
20163384   -26    -3     1    58     3    -1   -24   plot: complex amplitude
20163388     9    31   -23   -82   -17   -25   -19
20163392     7   -34   -25    48    19    37    -8
20163396    -4   -13     3   -12    -2   -11   -15
21000000     8   -33   -24    -8   -23    33    37
21200000    21    -3   -23   -18   -39    13    12
22000000   -19    41   -41   -24   -25    21   -20
23000000   -12   -16     2    27     3     5   -19
24000000   -82    45   -63   -37   -26    21    33
25000000     5    -7    -7    13    15   -26    24
26000000     9   -25   -32   -23   -48   -25   -17
27000000   -71   -19     2     2     6   -10    38
28000000     5     9    33     2    -9     8    27
29999972     7   609   -25   593    34   633    17
29999976   -12   637     8   675   -16   747   -11
29999980    10   836   -17   848    39   905     6
29999984   -31   984    -2  1055    -9  1093    17
29999988    37  1332    12  1434    -7  1557    -9
29999992   -20  1950     2  2165   -30  2504   -19
29999996   -40  3673   -16  4699   -13  6572   -63
```

Figure G.2: Result of 4M complex integer FFT, with three strong sines and three weak sines (both "on-bin" and "off-bin") embedded in strong noise (parameters listed on figure), using no windowing. The on-bin strong sine (at 1M) is cleanly resolved, permitting detection of the nearby weak signal (at 1M−8); but the off-bin sines (at 2M+0.25 and 3M+0.5) are unacceptably broadened, by spectral "leakage", burying the weak sines located 20 channels above.

```
window: Blackman-Harris, 16K x 16, expanded to 4M

data: noise (0.5 pk amp) + 3 sines (each 0.1 amp) @ 1M, 2M+0.25, 3M+0.5
      + 3 nearby weak sines (each 0.001 amp) @ 1M-8, 2M+20+0.25, 3M+20+0.5
        the sum quantized to 8 bits (i.e., each weak sine is 1/8 LSB amplitude)

fft: 20-bit, all scale

twiddle ROM: 128K x 16, expanded to 4M

plot: complex amplitude
```

Figure G.3: Result of 4M complex integer FFT, with three strong sines and three weak sines (both "on-bin" and "off-bin") embedded in strong noise (parameters listed on figure), using a Blackman-Harris window (implemented as a 16K×16 "expanded ROM"). The on-bin strong sine (at 1M) is now somewhat broadened, but not so much as to obscure the nearby weak signal (at 1M−8); now, however, the off-bin sines (at 2M+0.25 and 3M+0.5) are kept reasonably narrow, permitting clear detection of the weak sines located 20 channels above.

leakage wings.

In Figure G.4c we've used the Hanning (VH) window, and in Figure G.4d the Blackman-Harris (BH) window. The BH window is a severe window, with peak sidelobes of −92dBc, bought at the expense of a fairly broad main lobe (a pure sinusoid typically becomes 4 to 5 channels wide in the frequency domain). The VH window has a peak sidelobe level of −32dBc, but then falls at 18dB/octave; its main lobe is typically 3 channels wide. These characteristics are apparent in Figures G.4c,d: Note the narrow top portion of the main peak, when using the VH window, but broader width of the base (this is a logarithmic plot), compared with the BH window. Both satisfactorily separate the weaker signal from the sidelobes of the stronger; the comparison with the rectangular window is stunning.

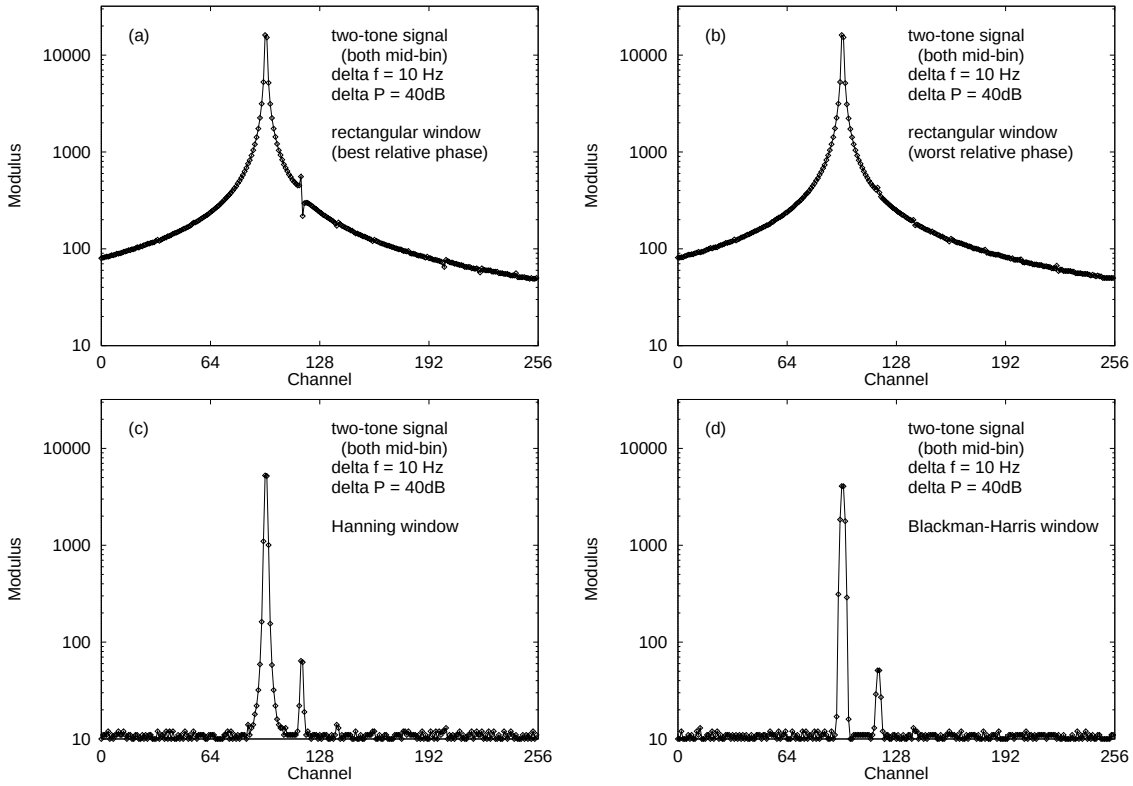


Figure G.4: Two-tone signal detection, separated 10 Hz in frequency and 40dB in power; both tones are placed between bins, the most difficult case. The vertical axis is logarithmic in modulus; the spectra in (c) and (d) have 10 added to the output, to suppress the visual raggedness that otherwise results from logarithmic plots of small numbers. Subfigures (a) and (b) use a rectangular window function (no windowing) and demonstrate the best and worst phasing of the weaker signal. Subfigures (c) and (d) use Hanning and Blackman-Harris windows, respectively. The reduction of spectral "leakage" is stunning.

Glossary

A/D Analog-to-Digital converter

ARRL American Radio Relay League

BER Bit Error Rate

BETA Billion-channel ExtraTerrestrial Assay – the subject of this document.

BIOS Basic Input Output System – a tiny, built-in operating system available to microcomputers when they boot.

CMB Cosmic Microwave Background

dB decibels

Dec Declination

DFT Discrete Fourier Transform

DOS Disk Operating System – a simple microcomputer operating system. Usually refers to a Microsoft or compatible product which runs on an Intel platform.

DRAM Dynamic RAM

DSML Derrick's State Machine Language – a language created for writing state machine programs for *BETA*. (described in section 3.4.2)

ECL Emitter-Coupled Logic – a fast bipolar logic family

EIRP Effective Isotropic Radiated Power – a transmitter's power output multiplied by its gain.

EM Electro-Magnetic

Ethernet A local area network standard.

FC Feature Correlator (described in section 3.4.2)

FFT Fast Fourier Transform. An $O(N \log N)$ algorithm for computing the DFT.

FIFO First In First Out – a queue-like memory device.

FR Feature Recognizer (described in section 3.4.1)

FUDD Follow-Up Detection Device

Glonass The Russian version of GPS.

GPIB General Purpose Interface Bus – an standard for communication between various test equipment. Also known as HPIB and IEEE-488.

GPS Global Positioning System

HA Hour Angle

HEMT High Electron Mobility Transistor

I in-phase

I/O Input/Output

IC Integrated Circuit

IF Intermediate Frequency

ISA Industry Standard Architecture – the 16-bit PC bus standard for expansion cards.

ISM Interstellar Medium

kWH Kilowatt-Hours

KL Karhunen-Loève (transform of expansion)

L-band The frequency range from 1–2 GHz.

LAN Local Area Network

LED Light Emitting Diode

LEO Low Earth Orbit

LO Local Oscillator

LSB Least Significant Bit

ly light-years

META Mega-channel ExtraTerrestrial Assay – *BETA*'s predecessor

MSB Most Significant Bit

PC Personal Computer, used here to refer to a consumer-class, Intel based micro-computer.

pc parsecs

PDF Probability Density Function

PIN diode A positive-intrinsic-negative material diode good for radio frequency switching applications.

PLL Phase-Locked Loop

PPP Point-to-Point Protocol – an internet standard for serial links.

Q quadrature

RA Right Ascension

RAM Random Access Memory – read-write memory

RF Radio Frequency

RFI Radio Frequency Interference

RMS Root Mean Squared.

ROM Read Only Memory

RT Real-Time Control Computer (described in section 3.5.3)

S-band The frequency range from 2–4 GHz.

SARSAT A Search and Rescue Satellite Service

SERENDIP Search for ExtraTerrestrial Radio Emissions from Nearby Developed Intelligence Populations – A set of SETI projects at the University of California at Berkeley.

SETI The Search for ExtraTerrestrial Intelligence

SM State Machine (described in section 3.4.2)

SMA A coaxial connector standard useful at microwave frequencies.

SNR Signal to Noise Ratio

TC Telescope Control Computer (described in section 3.2.2)

TCP/IP An internet protocol standard for reliable byte-streams.

UHF Ultra High Frequency – the frequency range from 300–3000 MHz.

Unix A multi-user, multi-tasking operating system for workstation computers.

UPS Uninterruptible Power Supply

VSWR Voltage Standing Wave Ratio

X-band The frequency range around surrounding 10 GHz.

Division of Labor

The committee has asked me to describe which parts of the project that I, the author, worked on personally. It's probably easier to describe the parts that I did *not* work on: I had little to do with the design of the RF downconversion and digitization components including the local oscillator box and the mixer digitizer boards. I did not work on the 4M FFT board and was only peripherally involved in the simulations, though I did do some analysis of the noise properties after-the-fact. Practically every other piece of the system has my fingerprints on it.

I *was* involved with the design of the dual feedhorns, though not principly. I was completely responsible for the design and testing of the terrestrial discone feed. I had complete responsibility for everything downstream of the FFT boards, including all software in the system. I was part of the design team for the FR/FC boards, and I designed, spec-ed and built the Pentium array with my own hands, including the homemade Ethernet boot ROM and all of the software. While I did not write all of the Unix-side software, I did help design it and supervised the development. I also modified certain parts of it extensively. I designed and built the RT and TC computers and their software almost single-handedly. I explored options for the Internet link and eventually designed and built the one we are currently using from the DDS phone line up. I designed and implemented all subsystems for synchronization and automatic control of *BETA* (though certain bits of hardware were designed and built by our talented team of undergrads).

I was involved in the system design from the start and provided a lot of the input for the basic architecture and functionality of *BETA*.

Bibliography

[1] Original blueprint in our possession, 1970. Plan number C84-081 for the HCO/SAO Radio Astronomy 84' Antenna, Gain-Feedhorn Combinations.

[2] Stuart Bowyer, Dan Werthimer, Charles Donnelly, Jeff Cobb, David Ng, and Michael Lampton. Twenty years of SERENDIP, the Berkeley SETI effort: Past results and future plans. In Cosmovici et al. [10], pages 667–676.

[3] R. Braun. The concept of the square kilometer array interferometer. In N. Jackson and R. J. Davis, editors, *High Sensitivity Radio Astronomy Meeting*, pages 260–268. Jodrell Bank, January 1996, Cambridge University Press, 1997.

[4] Jeff Cobb, Charles Donnelly, Stuart Bowyer, Dan Werthimer, and Michael Lampton. The SERENDIP IV interference rejection and signal detection system. In Cosmovici et al. [10], pages 677–682.

[5] Giuseppe Cocconi and Philip Morrison. Searching for interstellar communications. *Nature*, (184):844, 1959.

[6] Douglas E. Comer. *Internetworking with TCP/IP*, volume 1, pages 175–177. Prentice-Hall, Englewood Cliffs, New Jersey, second edition, 1991.

[7] James M. Cordes and T. Joseph Lazio. Interstellar scattering effects on the detection of narrow-band signals. *Astrophysical Journal*, (376):123–134, 1991.

[8] James M. Cordes and T. Joseph Lazio. Interstellar scintillation and SETI. In G. S. Shostack, editor, *Proceedings of the Third Decennial US-USSR Conference on SETI*, page 143, 1993.

[9] James M. Cordes, T. Joseph Lazio, and Carl Sagan. Scintillation-induced intermittency in SETI. *Astrophysical Journal*, (487):782, 1997.

[10] Cristiano Batalli Cosmovici, Stuart Bowyer, and Dan Werthimer, editors. *Astronomical and Biochemical Origins and the Search for Life in the Universe*, Proceedings of the 5^{th} International Conference on Bioastronomy, IAU Colloquium Number 161, Capri, Italy, 1996. Editrice Compositore, Bologna - Italy.

[11] D. Kent Cullers and Richard P. Stauduhar. Follow-up detection in Project Phoenix. In Cosmovici et al. [10], pages 645–651.

[12] Robert C. Dixon. *Spread Spectrum Systems.* John Wiley and Sons, New York, NY, second edition, 1984.

[13] Robert S. Dixon. The Ohio SETI program and the Argus telescope. In Cosmovici et al. [10], pages 623–632.

[14] Frank Drake. The Drake equation was first presented at the 1961 Green Bank Conference. Good explanations of this are in [39] page 26 and `http://www.seti.org/drake-eq.html`.

[15] Frank D. Drake. Estimates of the relative probability of success of the SETI search program. In Frank Drake, John H. Wolfe, and Charles L. Seeger, editors, *1983 Report of the SETI Science Working Group*, appendix IV. 1983.

[16] Martin Gimersky and Paul Horowitz. Design of an L-band dual feedhorn. *Radioengineering*, 3(2):1–4, 1994.

[17] Donald Goldsmith, editor. *The Quest for Extraterrestrial Life: A Book of Readings.* University Science Books, Mill Valley, California, 1980.

[18] D. A. Green. A Catalog of Galactic Supernova Remnants (1996 August version). Catalog, Mullard Radio Astronomy Observatory, Cambridge, United Kingdom, 1996. (Available on the World-Wide-Web at `http://www.mroa.cam.ac.uk/surveys/snrs`).

[19] Gerald Hall, editor. *The ARRL Antenna Book*. The American Radio Relay League, Newington, CT, fifteenth edition, 1988.

[20] Gerald Hall, editor. *The ARRL Antenna Book*, pages 9-7 – 9-12. In Hall [19], fifteenth edition, 1988.

[21] Frederic J. Harris. On the use of windows for harmonic analysis with the Discrete Fourier Transform. *Proceeding of the IEEE*, 66(1):51–83, 1978.

[22] J. Heidmann. Saha crater: A candidate for a SETI lunar base. *Acta Astronautica*, 32:471, 1994.

[23] Paul Horowitz and Winfield Hill. *The Art of Electronics*, pages 896–898. Cambridge University Press, Cambridge, England, second edition, 1989.

[24] Paul Horowitz and Carl Sagan. Five years of project META: An all-sky radio search for extraterrestrial signals. *Astrophysical Journal*, (415):218–235, 1993.

[25] N. S. Jayant and Peter Noll. *Digital Coding of Waveforms*. Prentice-Hall, Englewood Cliffs, New Jersey, 1984.

[26] Kardashev, Gindilis, Popov, Soglasnov, Spangenberg, Steinberg, et al. Pulse search at 371, 408, 458 and 535 MHz. Search site: Caucasus, Pamir and Kamchatka regions and the Mars 7 spacecraft. Search period: 1972–1974. In [50].

[27] N. S. Kardashev. *Soviet Astron. – AJ*, 8:217, 1964.

[28] Kingsley. Columbus Optical SETI Observatory optical pulse search at 0.55μ. Search periods: 1995–. In [50].

[29] Bennett Z. Kobb. *Spectrum Guide: Radio Frequency Allocations in the United States, 30 MHz – 300 GHz*. New Signals Press, Inc., Falls Church, VA, second edition, 1995.

[30] John Kraus. *Big Ear Two: Listening for Other Worlds*. Cygnus-Quasar Books, Powell, Ohio, 1994.

[31] John D. Kraus. *Radio Astronomy*, pages 5-13 – 5-17. Cygnus-Quasar Books, Powell, Ohio, second edition, 1986.

[32] John D. Kraus. *Antennas*, page 27. In [34], second edition, 1988.

[33] John D. Kraus. *Antennas*, pages 46–47. In [34], second edition, 1988.

[34] John D. Kraus. *Antennas*. McGraw-Hill, new York, NY, second edition, 1988.

[35] Darren Leigh. *An Interference-Resistant Search for Extraterrestrial Microwave Beacons*. Ph.D. dissertation, Harvard University, June 1998.

[36] Darren Leigh and Paul Horowitz. Millions and billions: The META and BETA searches at Harvard. In Cosmovici et al. [10], pages 601–610.

[37] Claudio Maccone. *Telecommunications, KLT and Relativity*, volume 1. IPI Press, Colorado Springs, Colorado, 1994.

[38] Bernard M. Oliver. Rationale for the water hole. *Acta Austronautica*, 6:71–79, 1979.

[39] Bernard M. Oliver and John Billingham. Project Cyclops: A design study of a system for detecting extraterrestrial intelligent life. Technical Report CR114445, NASA, 1973.

[40] Athanasios Papoulis. *Probability, Random Variables and Stochastic Processes*, pages 412–416. In [42], third edition, 1991.

[41] Athanasios Papoulis. *Probability, Random Variables and Stochastic Processes*, pages 86–. In [42], third edition, 1991.

[42] Athanasios Papoulis. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, New York, NY, third edition, 1991.

[43] Edward Purcell. Radioastronomy and communication through space. Report BNL-658, U. S. Atomic Energy Commission, 1961. Reprinted in [17], pages 188–196.

[44] George P. Rybicki and Alan P. Lightman. *Radiative Processes in Astrophysics*, pages 229–231. John Wiley and Sons, 1979.

[45] Carl Sagan, editor. *Communication with Extraterrestrial Intelligence (CETI).* MIT Press, Cambridge, Massachusetts, 1973.

[46] Louis Scheffer. Exact doppler shift removal over wide bandwidths. Unpublished manuscript. Mr. Scheffer can be contacted at `lou@alumnae.caltech.edu`, April 1997.

[47] Robert Schetgen, KU7G, editor. *The ARRL Handbook*, pages 12.3–12.6. The American Radio Relay League, Newington, CT, seventy-second edition, 1995.

[48] C. E. Shannon. Communication in the presence of noise. In *Proceedings of the IRE*, pages 10–21, 1949.

[49] Shvartsman et al. "Mania" optical pulse searches at 5500Å. Search periods: 1973–1974, 1978–. In [50].

[50] Jill Tarter. Summary of SETI Observing Programs, October 1996.

[51] Jill C. Tarter. Results from Project Phoenix: Looking up from down under. In Cosmovici et al. [10], pages 633–643.

[52] A. Richard Thompson, James M. Moran, and George W. Swenson Jr. *Interferometry and Synthesis in Radio Astronomy*, pages 449–455. Krieger Publishing Company, Malabar, Florida, 1994.

[53] Troitskii, Bondar, and Starodubtsev. Pulse search at 1863, 927 and 600 MHz. Search site: Gorkii, Crimea, Murmansk and Primorskij regions. Search period: 1969–1983. In [50].

[54] Dan Wertheimer. Personal communication.

Colophon

This document was composed with the Gnu Emacs text editor on a Sun Unix workstation using the LaTeX 2_ε document preparation system. Original diagrams were created with Adobe Illustrator and saved as Adobe Illustrator 7.0 `.ai` files (basically encapsulated Postscript). Legacy diagrams available only on paper were scanned with an HP Scanjet IIc, converted to line art with Adobe Streamline and then cleaned and finished up with Illustrator. Most of the images were scanned from photographic prints and cleaned up with Adobe Photoshop. Plots were chiefly made using Gnuplot version 3.5 although several legacy Mongo plots were redone with SM (Super Mongo). The Adobe software ran on PCs under Microsoft Windows 95. All other software ran on the Unix workstation.

All figures, images and plots were saved as encapsulated Postscript files and merged with the LaTeX using epsfig version 1.20 and Tom Rokicki's dvips version 5.66a.

The archive copy was printed from the dvips output on Crane's thesis paper (100% cotton fiber, acid free) using an HP Laserjet 4M Plus at 600 dots per inch. Electronic copies are available either as the dvips Postscript output or as Adobe PDF converted from the Postscript using Adobe Distiller.