2014-07-01

Introduction to the Development of a Radio Astronomy System at Brigham Young University

Daniel Robert Blakley
Brigham Young University - Provo

Introduction to the Development of a Radio Astronomy System at Brigham Young University

Daniel Robert Blakley

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Victor Migenes, Chair
J. Ward Moody
Karl Warnick

Department of Physics and Astronomy

Brigham Young University

July 2014

ABSTRACT

Introduction to the Development of a Radio Astronomy
System at Brigham Young University

Daniel Robert Blakley
Department of Physics and Astronomy, BYU
Master of Science

The intent of this project was founded upon the need to train students in the techniques of radio astronomy with the purpose of establishing a radio telescope in order to teach the principles and practice of radio astronomy.

This document describes the theory and research necessary to establish the 1st generation radio telescope system within the Department of Physics and Astronomy at Brigham Young University. Included are introductions to: (1) The nature of star forming regions in the spiral arm structure of the galaxy, H I (the hydrogen spin-flip transition) and OH MASERS, (2) The terminology used with the system components and their measurements, (3) The characteristics of the imaging system and its limitations, and (4) Future work and plans. Within the body of this work, I also present an introduction to the purpose and architectural design, of some of the system level functions with associated equipment that constitute the development infrastructure for the 2nd generation radio astronomy system.

This work includes history, some of the fundamental theory behind radio astronomy, significant aspects of the theory behind the system, building of the system, its calibration and characteristics as well as next steps

Keywords: H I, hydrogen spin-flip, radio astronomy, astronomical OH MASER, antenna

ACKNOWLEDGEMENTS

I wish to thank Brigham Young University Physics and Astronomy Department for their support and patience; Dr. Victor Migenes, for his friendship and council over many varied discussions, my committee (Drs. Karl Warnick, and J. Moody) for their encouragement and council; Scott Daniel for his friendship, help, and council in various discussions; I also wish to thank my colleagues, Drs. Steven Simske, (Hewlett-Packard Fellow) and Pankaj Vadgamma, (Queen Mary University of London) for their encouragement, advice, and long-term friendship. I particularly wish to thank Dr. Jeff Macedone (BYU Chemistry Dept.) for his help and his kind sacrifice of time to review the thesis and help with the formatting of the manuscript.

My employers, both in the Department of Physics and Astronomy and at Aribex (Orem, Utah), respectively, for the rich opportunities for learning and for the opportunity to not only provide for my family but the opportunity to design X-Ray generation equipment & high-voltage power supplies, an interesting combination of physics and engineering.

I believe that the LDS Church had great inspiration in the establishment and their support of Brigham Young University. What great insight and vision the early prophets had in the creation of the university, considering the great impact for good that they have made. I wish to thank my wife, who had faith and courage to support her husband on such an endeavor – one that required many sacrifices. I also wish to thank my daughters Kendra and Megan and my sons Aaron and Seth for your encouragement and faith as well as my mother and father for their faith and encouragement.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

This thesis is only intended to describe a limited amount of the physics and radio astronomy principles as applied to imaging star forming regions within our galaxy, since these

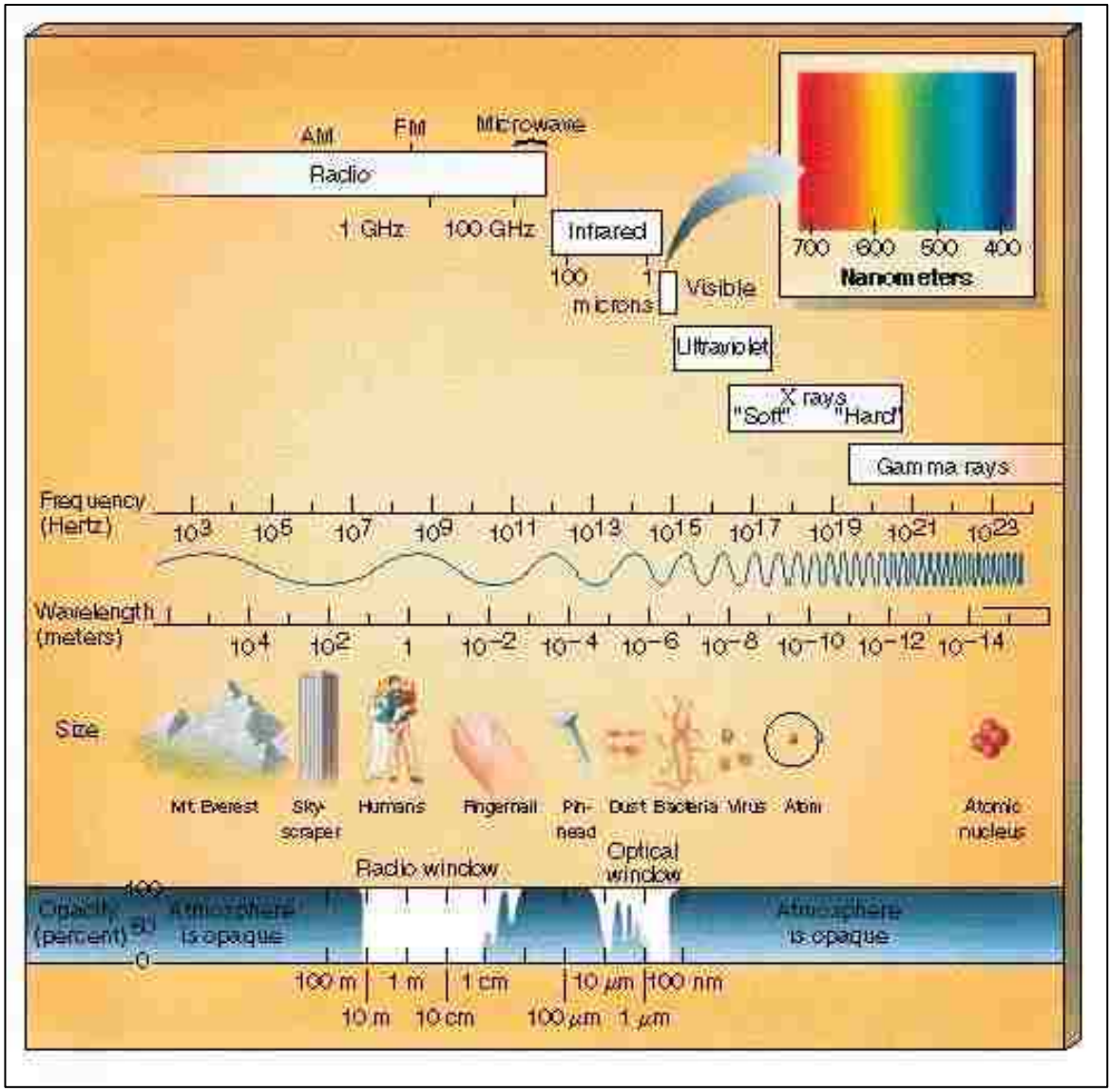

Figure 1-1 Electromagnetic spectrum, and regions of radio spectrum, IR, optical, and other spectrums

are the regions to be studied by this radio astronomy system. In order to study the nature of star forming regions, which are embedded in dense clouds of gas and dust, we first introduce the radio spectrum as it provides a special 'window' to observe and make measurements.

As seen in the electromagnetic spectrum above, one observes that the radio spectrum is much broader than any of the other spectral bands. Historically, it is only from the late 19th century that we have discovered the principles associated with electromagnetic waves, but not until the 20th century were the technology and tools sufficiently developed to study the radio band. One may ask what causes electromagnetic wave radiation. The short answer is that all baryonic matter emits electromagnetic radiation in proportion to their absolute temperature (thermal) or interaction with EM fields (non-thermal). Radiation is the most observable form of the EM spectrum. In the EM spectrum, there are two main types of radiation observed: Continuum and Spectral line. For astrophysical sources, continuum is associated with a spectrum, often originating in blackbody or free-free emissions.

Blackbody radiation is related to its temperature, where the lower the temperature, the lower the energy of radiation given off, and with high temperatures high radiation energies. A blackbody radiator completely absorbs all radiation that impinges upon it, and reflects nothing. Its radiated characteristic spectrum peaks at a wavelength that depends only on the object's temperature. In an astrophysical environment, free-free emission frequently originates within the less dense interstellar gas clouds where atomic collisions are not kinematic, but rather where atoms closely approach one another rather than directly colliding, producing emissions.

Lastly, spectral line implies a monochromatic emission. In the astrophysical environment such monochromatic emissions may be driven thermally, but always tend to have a quantum mechanical constraint or origin, described in more detail within the body of this work. The radio astronomy system described here is able to detect both thermal and non-thermal radiation in both continuum and spectral line modes.

1.1 In the Beginning

It is instructive to first briefly review a bit of history of the universe as we know it now with a few principles relating to the beginning of this universe, especially as it pertains to the contributions amenable to observations using radio astronomy.

Figure 1-2 The Creation, as depicted together with the Christus Statue [1]

Historically, it is generally understood that although there were various theories and beliefs throughout antiquity, up until perhaps 100 years ago, it was generally believed that the universe was stationary or 'always there', as a static and unchanging large environment. However, only within the last 100 years or so has our understanding of the universe and its history developed to a level of understanding that incorporates and agrees with a great number of

observations and measurements. Our current understanding is that our universe began roughly 13.7 billion years ago in what is referred to as the 'Big Bang'. The universe began extremely small, hot, and in a dense state and has expanded since that time [2,3]. As the universe cooled, the energy was converted into various particles including protons, neutrons, and electrons, leading to the production of primarily hydrogen, deuterium, tritium, helium, and lithium. As the universe expanded and cooled, these elements were attracted by gravity to form clouds which later gave birth to stars and galaxies. Heavier elements found in the periodic table up to and including iron were synthesized within stars through the process of fusion. The fusion process creates heavier elements from lighter elements but is limited due to the binding energy curve, which describes the amount of energy required to hold the nucleus together [4]. The binding energy curve falls off after the element iron (element 26). This characteristic of declining binding energy implies that more energy than that available via fusion is required to synthesize elements beyond that of iron. Hence, only elements up to iron can be synthesized in a star [5]. Heavier elements were synthesized during supernovae explosions, occurring at the death of high-mass stars. Early in the universe, stars were different than those that formed later on due to the presence of heavier elements available after the death of stars. We refer to Population III stars as those that formed earlier in the life of the universe. These early stars were composed solely of H and He. Those stars that formed later are classified as Population II and Population I stars. Population II are the older stars and Population I are the newest stars to have formed. It is observed [6] that we generally do not detect heavy elements in Population II stars, although we do observe heavier elements in the younger spectra of Population I stars.

Historically, the idea that the universe was expanding and not simply unchanging or static appears to have had its beginning in the early 1900s, initially with observations of receding

galaxies by Vesto Slipher [7,8] and later, Carl Wilhelm Wirtz [9] who also observed receding nebula. Einstein's [10] General Theory of Relativity originally predicted that gravity would cause the universe to contract, however he added a cosmological constant that would keep the universe static and unchanging, a state that he assumed incorrectly. Later, this constant came into question as evidence of an expanding universe from Hubble and others became evident, causing Einstein to call this assumption his "biggest mistake". In 1927 George LeMaître [11,12] proposed that the universe expanded explosively from an extremely dense and hot state, with this expansion continuing today. LeMaître's original perspective differed from Einstein [13] and later helped Einstein realize his mistake. In 1929 Edwin Hubble [14,15] published a comprehensive study of galaxies that were Doppler (red) shifted, adding further evidence to the idea that the universe is not static, but expanding [16].

Although initially met with much skepticism and scorn, it was found that an expanding universe agreed with a wide variety of observations and measurements. This included a black-body radiation at a temperature of approximately 2.7 degrees Kelvin. Current theory and observation date the big bang to approximately 13.7 billion years ago [17,18]. There has been mounting evidence in the last decade that galaxies are moving away from the Milky Way galaxy at an accelerating rate, as determined from Supernova Type 1A positions. As the universe began, the original numbers of baryonic elements created were rather few, although extremely abundant. These first baryonic elements included hydrogen, deuterium, tritium, helium, and lithium. As stars began to form, heavier elements were formed through the process of fusion within each star. Even today, these lightest elements constitute the highest percentages of the baryonic matter that make up our universe and also serve as a reference for a universal scale of time where other, heavier elements, have been created in stars.

1.2 Our Galaxy and Star Forming Regions

In (Figure 1-3) we see our Milky Way Galaxy, looking down from above, where our sun is one of approximately 200 billion stars. The Milky Way appears to be a typical barred-spiral

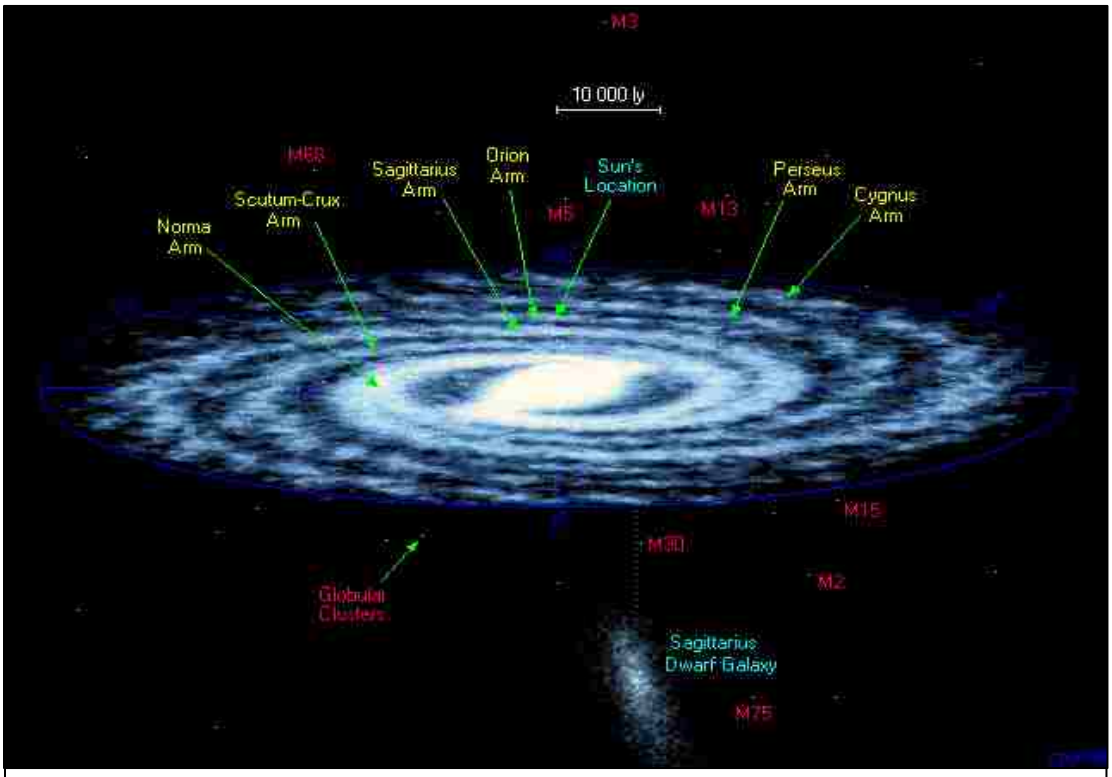

Figure 1-3 Milky Way Galaxy. Note that this is a spiral galaxy, having a "bar" at its center, which is believed to contain a black-hole. Also note the structure which is composed of spiral arms, comprised of HI star forming regions that are able to be imaged by radio astronomy systems such as that which we have constructed. From: www.altasoftheuniverse.com/galaxy.html.

galaxy which is about 100,000 light years in diameter.

From a quick examination of the figure, one may observe that stars lie on the disk of the galaxy with star formation taking place mainly on the spiral arms due to the apparent increase in the amount of light seen along this region. However, there are almost as many stars that lie between the spiral arms as those within the spiral arms. The regions of galaxies along their spiral arms are the major regions of star formation within spiral galaxies and this is where most of the

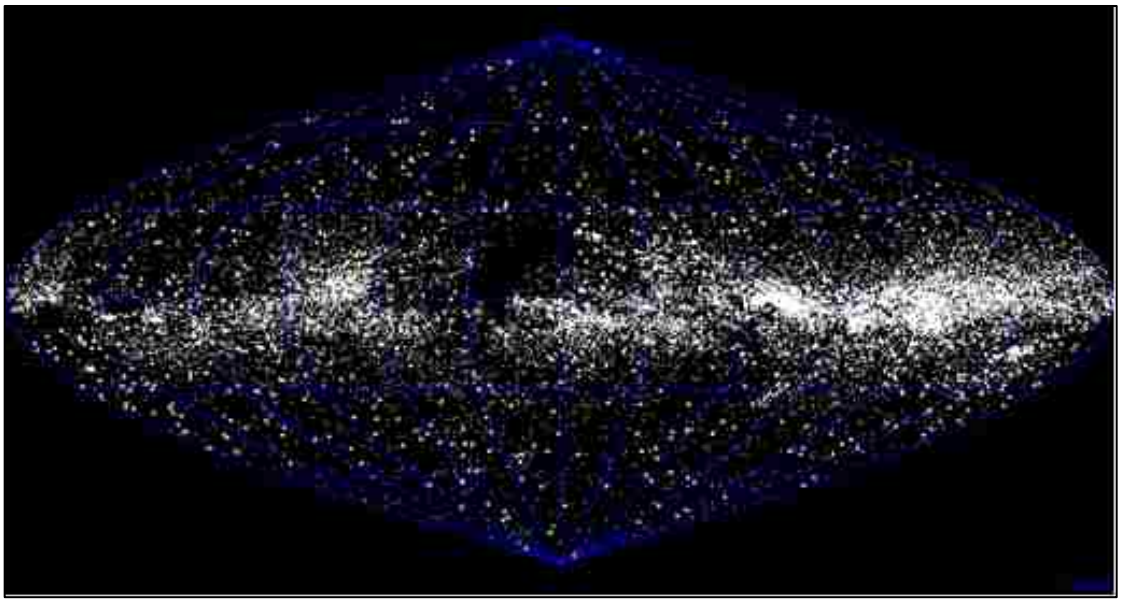

Figure 1-5: This view of our milky Way galaxy, shows a plot of 25000 of the brightest stars. From: "www.atlas.of theuniverse.com/galaxy.html".

major nebulae are found. Although the actual distributions of stars both within and without the

Figure 1-4: The Milky Way: A picture of our galaxy taken in infrared light as part of the 2 Micron All-Sky Survey (2 MASS). The 2 micron infared light can pass through the cloads of dust that fill our galaxy, giving us a clear view of half a billion stars. From: www.altasoftheuniverse.com/galaxy.html.

spiral arms are almost equivalent, the arms of spiral galaxies appear brighter due to the fact that

the brightest stars (those with the highest mass and youngest) tend to be found more frequently in the spiral arms. If we plot the brightest stars in the Milky Way, we find the picture shown here of the 25000 brightest stars, as an inside view of our galaxy (see Figure 1-5). As we examine such distributions of stars within our galaxy, it is particularly illuminating to view using infra-red light (of 2 micron band filter), which has the ability to better see through the obscuring interstellar dust. For example, in the IR imaged picture of Figure 1-4, we see the high concentration of stars along the spiral arms of the galaxy. Additionally, in this photo, we observe the presence of the bulge that lies at the center of our galaxy as well as all other galaxies that we are now aware of.

1.3 Galactic Plane Radio Astronomy Comparisons

In Figure 1 5, we see the galactic plane, where it is instructive to compare this same image as viewed in different wavelengths in order to better understand what one may expect as he/she views objects in radio wavelengths as compared to others wavelengths. So let's begin by comparing a few NASA galactic plane observations. In Figure 1 6 [ ] is shown a composite spectral image of our galactic plane for comparison, when viewed at several different wavelengths. The uppermost spectrum is of the he Radio Continuum (408 MHz) (from Jodrell Bank MkI and MkIA, Bonn 100 meter, and Parkes 64 meter), where most of the emission is from the scattering of free electrons in interstellar plasmas (hot, ionized interstellar gas). The Atomic Hydrogen image (Dickey-Lockman) is of neutral atomic hydrogen, radiating at the 21-cm transition of hydrogen which on a large scale is organized into diffuse clouds of gas and dust that have sizes of up to hundreds of light years. The Molecular Hydrogen (H2) image (Columbia and GISS) is inferred from the intensity of the J = 1-0 line of carbon monoxide (CO), since H2 is

difficult to observe directly. However, the ratio of CO to H2 is fairly constant, so it is used instead to map dense molecular hydrogen gas which tends to be concentrated in the spiral arms in discrete "molecular clouds," which are often sites of star formation.

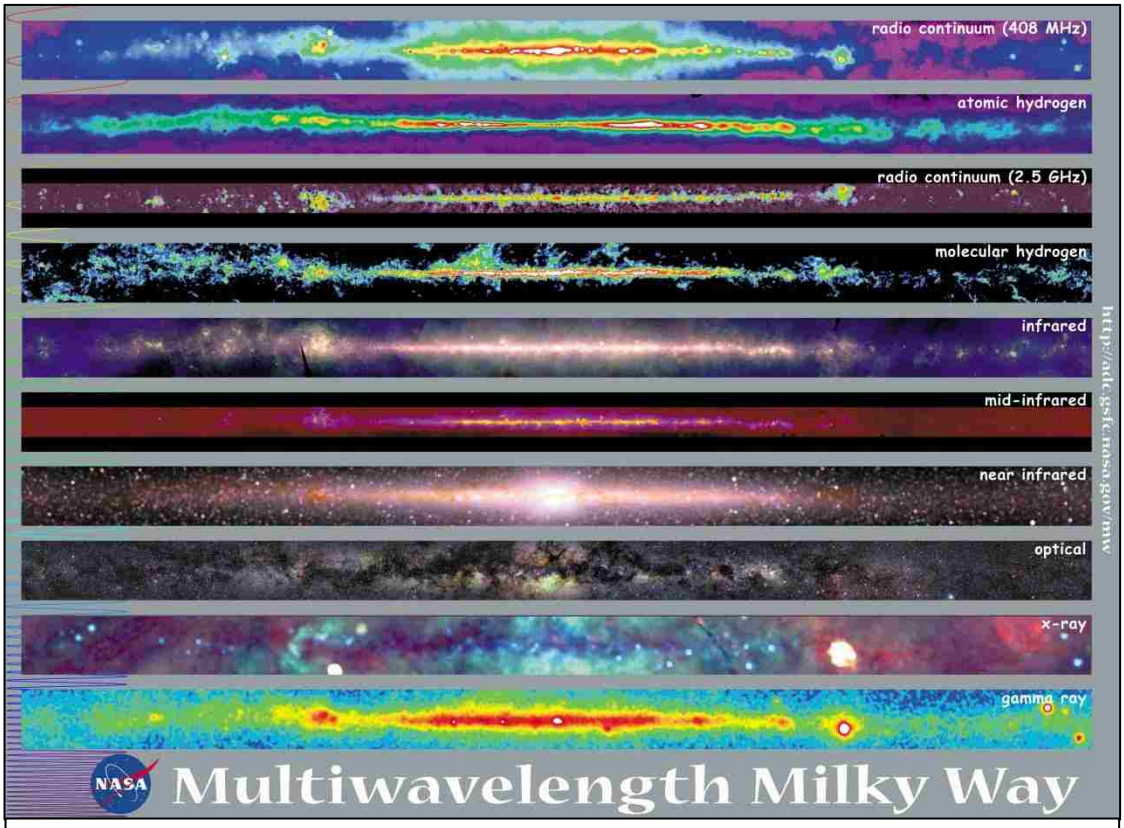

Figure 1-6 Composite images of the galactic plane as viewed using 10 different wavelength bands. These are (from the top): 408 MHz, 1.42 GHz, 2.5 GHz, molecular hydrogen, IR, Mid-IR, Optical, X-Ray, and Gamma-Ray Wavelengths. See: mwmw.gsfc.nasa.gov/mmw_product.html#viewgraph.

These images help us to have context of what we may expect to observe in various wavelength observations, using the galactic plane as a basis by which to contrast [19].

1.4 The Radio Astronomy System, Its Purpose and Limitations

The intent of this Radio-astronomy system is to image both regions of H I and OH MASERS in the star-forming regions of our galaxy (see Figure 1-3 [20]), by use of single dish radio astronomy techniques and with an implicit capability to track galactic objects with large

dwell times in order increase integration time and so improve sensitivity. Although the system described herein is intended to be enhanced through the provided capability, the system's first embodiment consists of a single 4-meter antenna with a triple conversion super-heterodyne narrow-bandwidth receiver, fed by the antenna. The superheterodyne receiver is based upon the use of non-linear “mixers”, in which a local oscillator mixes with the received signal and the resultant sum and difference frequencies are filtered to select one or the other. This is followed, in turn, by a filter in order to improve selectivity of the received signal. A triple order superheterodyne contains three of these circuits in succession, in order to improve selectivity and avoid the interference of adjacent signal noise. At the focal-point of the antenna lies a resonator cavity with two single polarization dipole antennas and two Low Noise Amplifiers (L.N.A.s) for 1.4 and 1.6 GHz frequencies (associated with HI and OH MASERS, respectively).

The system is designed for expanded enhancements in function and capability. Particularly noteworthy among these features is its transformation into an interferometric radio-astronomy array realized through use of an interconnected geographically distributed multi-antenna system with associated multi-receiver interferometric correlator. This capability is realized in its second generation embodiment, for which substantial forethought, design, and system implementation has been provided in this work. Although the full realization of the system's interferometric capabilities lies beyond the scope of this thesis, substantial work was applied to this effort and has resulted in a rich system level infrastructure with which to complete the initial work reported here.

Several aspects of the system were designed for expanded roles for radio astronomy (such as azimuth and elevation controllers, additional antennas, additional frequency coverage, etc.), however there are a few limitations to the basic system as it is envisioned and constructed.

Although the frequencies defined for use are limited by the dipole antennas, their associated resonator cavity (located at the focal point of the antenna), and the LNAs, each of these functional blocks have broad band characteristics and therefore can be used for a much greater bandwidth than their prescribed frequency dictates. It is the first generation receiver that is the primary system constraint, in terms of frequency. This is due primarily to the 1st generation's triple conversion super-heterodyne architecture; specifically it is the narrow bandwidth of the last conversion stage that limits the system to a very narrow bandwidth. Another limitation lies in the front end of the receiver's triple order conversion architecture, where a somewhat narrow band RF filter is located. If it were desired to accomplish wide band radio-astronomy applications, the 2nd generation receiver has none of these described limitations, and is actually capable of imaging bandwidths of 100 MHz and greater. This capability is due to the digitally sampled nature of the CASPER ROACH architecture (described in Section 5), where there is the potential for a bandwidth from DC to the Nyquist sampling limit and potentially beyond.

System gain is, by definition, the sum of all gains and losses within the entire system. The major elements of these include the antenna gain, the low-noise amplifier (found near the focal point of the radio-astronomy antenna). However when considering gain, perhaps just as important is the issue of noise, since it is the ratio of signal to noise that frequently sets the limits of what can be detected as opposed to what cannot. So, by way of limitations, one's attention is drawn to the primary system component that largely determines the overall system gain as well as its noise floor. The component in question is the low-noise-amplifier which is a Peltier-cooled broad-band low-noise amplifier with a peak in its amplification bandwidth centered on the frequency of interest. Its single stage Peltier cooler is constrained to a maximum of 60 degrees below ambient temperatures (theoretically) but often not entirely realized due to other

limitations. However, the Peltier cooler does indeed lower the noise floor (normally by about 3 K from a typical ambient noise temp of 20 - 22 K).

The parabolic dish antennas employed for the system, both on the Eyring Science Center at BYU as well as other established sites, were originally designed for satellite receive operations, primarily in the "L-Band" (1.3 – 1.8 GHz) and "C-Band" (4.2 – 5.8 GHz) frequencies. Although a parabolic antenna's reflecting surface is usually a reflector in terms of its ability to focus electromagnetic energy over a wide area to a single region with gain, as the frequencies increase above some limit, the accuracy of the surface (among other things) becomes a limitation, due to its meshed and non-perfectly uniform surface. Hence, these antennas are wholly not suitable for frequencies above approximately 12 GHz. The exact frequency for each may be readily measured, but has not yet been. It will be noted though that the measurement of antenna gain may be accomplished through the use of two broadband antennas, a spectrum analyzer, and a swept frequency signal generator.

The radio astronomy system, which we have presently constructed, has the capability to detect galactic neutral hydrogen emission and MASER emission from hydroxyl molecules [21,22]. One of the system disadvantages lies in its limitation to extra-galactic imaging. This limitation in imaging is due to very high radiation red-shift caused by the expansion of the universe. Specifically, this limitation is due to the more narrow-bandwidth nature of the antenna resonating elements used and the intermediate frequency bandwidth of the last stage of the 1st generation triple conversion receiver. In order to image extra-galactic sources, one would need to either broaden the bandwidth or alternately shift frequencies from that targeted for imaging to a much lower one. However, it should be noted that the CASPER based system of the 2nd generation system architecture does have both of these capabilities.

Initially, in the 1st generation system, imaging is limited to HI and OH MASER lines, which are conveniently positioned close to one another in terms of frequency and bandwidth. A more specific aspect to the limitation is related to cosmological distance; as HI occurs at approximately 1.4 GHz and OH MASERS occur at approximately 1.6 GHz, imaged regions much more distant will have a higher red-shift associated with them due to the expansion of the universe, resulting in much lower frequencies. In order to image the rest frequency of H I and OH MASER lines, as well as imaging the red-shifted frequencies of more distant galaxies, a receiving system requires a greater bandwidth for imaging. However, for such a system, there are inherent trade-offs that manifest additional limitations particularly when the system is limited by antenna size.

1.5 History of Radio-Astronomy at Brigham Young University

Historically, the Physics and Astronomy department at Brigham Young University has performed optical and I.R. astronomy for many years. The astronomy capability of the department now (as of the publication date of this thesis) includes several optical telescopes on the Provo campus as well as a remote observatory ("ROVER") located in Delta, Utah and also multiple optical telescopes (including a 0.9 m) located at the BYU West Mountain observatory site.

Previous to the work associated with this thesis, there have never been local facilities, or local capabilities for radio-astronomy, within the department. However, since 2004/5 there has been radio-astronomy capability within the BYU Electrical Engineering department in the form of the "Very Small Array (or VSA)". The VSA consists of a series of 4-5 antennas, located on the BYU Clyde building, with an associated analog receiver system intended for laying the

foundation for radio-astronomy phased array research (see Figure 1-7). Although the capabilities of the VSA were intended for a different purpose than what is generally required for astronomy

Figure 1-7. VSA Located on Clyde Building. Thee of the Six Array Antennas are shown here. . Photo courtesy the BYU Electrical Engineering Research Group of Dr. Brian Jeffs and Dr. Karl Warnick.

teaching and research within the Physics and Astronomy Department, the system still has some potential that will be addressed later in this document (see VSA extension).

It was realized that an independent local system would be of much greater benefit to the students of the department, particularly when hands-on teaching and research was an objective.

1.6 H I in Star Forming Regions

Star forming regions are composed of low-temperature, dispersed interstellar gas and

dust. The temperatures of these molecular clouds are typically 20K (slightly above absolute zero) where the gas, although low in density, has a gravitational tendency to clump. The constituent gases in such regions are composed of the original constituent elements from the beginning of the universe, the constituents left over from the death of stars, as well as dust. Since hydrogen is predominant, much of the matter of the universe exists in clouds of neutral hydrogen. As the clouds collapse and the density increases, it kick-starts star formation and the formation of galaxies. As stars evolve and heavy elements and dust are produced, these molecular clouds tend to become opaque to visible light. Yet one may still examine those regions using IR and radio telescopes. Clouds of neutral hydrogen gas may be mapped or imaged by using radio-astronomy within our own Milky-Way Galaxy by making use of a quantum mechanical property of this gas. Due to the cold temperature at which these clouds of gas exist, as collisions occur between atoms, electrons may shift their spins from up to down (or vice versa) relative to the spin of the proton in the atom's nucleus. When the spin of the electron changes spin direction in this manner, a fixed amount of energy is either required or released. From the relationship, $\text{energy} = \text{h}\upsilon,$ where energy is that required for the spin to reverse directions, h is a constant and υ is frequency, one may find the frequency of the emission of photons from such a transition, which turns out to be 1.420 GHz.(or 21 cm) Though the probability of this collision taking place is very small, the amount of neutral hydrogen gas present in many of these interstellar clouds makes this transition detectable. The 21cm oscillatory emission results in a spectral line that is only observed in regions having a significantly great abundance of hydrogen gas present, due to the fact that the electron spin transition that is associated with this spectral line emission has a low probability of occurrence. In order to observe sufficiently high amplitude of the 21cm spectral emission, significant numbers of hydrogen atoms must be present in the H I state for

observation. So, only gas clouds of significant size are amenable to this form of imaging, depending upon the sensitivity of the radio astronomy system used.

Using this property of the electron spin transitions that comprise the gas of the hydrogen clouds in interstellar space, the distribution of H I may be mapped using a radio telescope which is tuned to this frequency (Figure 1-8). Although such images determine the boundaries of star

Figure 1-8: This photo is an image of M17, in the Milky Way Galactic Plane, when imaged at 1.4 GHz (Hydrogen Spin-Flip). Image is by Robert Stephens, of the Alqonquin Radio Observatory. Reference: http://www.radioastronomysupplies.com/show_photo.php?&photo_id=18&photo_title=M17-and-Galactic-Plane-at-1.4-GHz.,Robert-Stephens,-Alqonquin-Radio-Observatory.

forming regions, they may be employed to image and study other related objects, such as the extent of a galaxy, and the structure and morphology of its spiral arms. Examples of imaging H I emission using radio telescopes have included the structure and morphology of spiral galactic arms and the gravitational disruptions between galaxies (such as when two galaxies collide, resulting in interstellar material being pulled out in strands) which were used to determine the direction of movement for the resulting galactic remnants. Although in general, these and other

such studies are possible for H I imaging, our system is not quite capable of such imaging, due primarily to its much smaller antenna size and its use of non-cryogenically-cooled LNAs, both of which limit the signal-to-noise ratio of the system and therefore its overall capability.

Although generally H I regions are not ionized, they do quite effectively absorb photons which are energetic enough to ionize some of the hydrogen gas (which requires an energy of 13.6 eV for each hydrogen atom ionized). H I regions are abundant within our Milky-Way galaxy. In general, along the interface of the transition region between H I and H II, one notes that the H I regions are non-ionized while all H II regions are ionized. At the actual interface of the two regions, referred to as an ionization front, there are collisions where the H I regions collide with expanding ionized gas (such as an H II region). Here, the H II expanding ionized gas glows brighter than it otherwise would within such collision regions. With regard to ionization, the degree of ionization in an H I region is very small at around 10^{-4} (or one atom out of 10,000) and the temperature of an H I region is about 100 K, [23] and it is usually considered as isothermal, except when it is located near to an expanding H II region.[24]. Although both H I and H II may both be expected to be present with great abundance in star forming regions, since H II is not capable of being imaged with our present radio astronomy system, its discussion is not dealt with in much detail as it is beyond the scope of this document.

1.7 The Quantum Mechanical Nature of H I

The 1420 MHz radiation that is emitted from H I (at 20 K) originates from the quantum mechanical transition between two levels of the hydrogen atom in the 1s ground state. This atomic transition comes from the interaction between its electron spin and its nuclear spin. When a transition occurs from this interaction of spins, a small amount of energy is required or

radiated. This energy change is referred to as the hyperfine structure transition. This emission of energy caused by interaction between spins is illustrated in Figure 1-9.

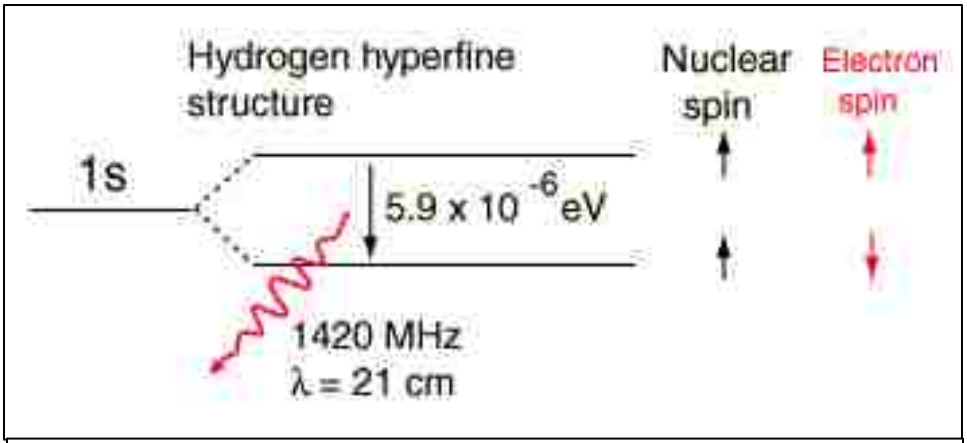

Figure 1-9 The hyperfine structure of H I is caused by the interaction of the hydrogen atom's nuclear spin and its electron spin. The transition is between the alignment of these spins and their anti-alignment, which results in the 21 cm emission line. Reference: hyperphysics.phy-astr.gsu.edu/hbase/quantum/h21.htm

Hydrogen in its lower energy state will absorb 1420 MHz in order to transition to its higher energy state. Similarly, hydrogen in the higher energy state will emit energy at 1420

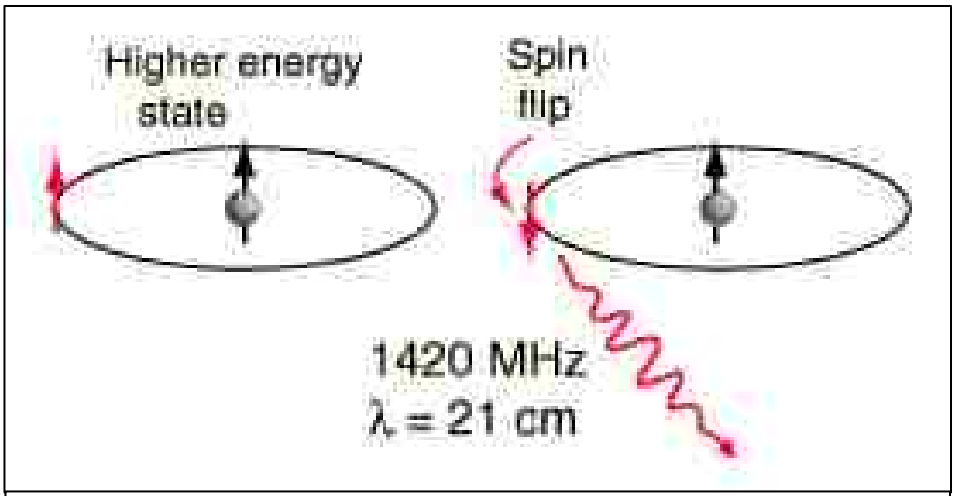

Figure 1-10 Orbital depiction of the spin flip transition of H I. Reference: hyperphysics.phy-astr.gsu.edu/hbase/quantum/h21.htm

MHz, as a photon. This hyperfine transition of hydrogen is roughly 5.8 x $10^{-6}eV$ when compared to the ground state energy of -13.6 eV [25]. These two possible spin states for

hydrogen arise because both the electron spin and the proton spin are ½. Therefore the two possible states for hydrogen are spin parallel and spin antiparallel. The state with the spins parallel is slightly higher in energy (or is less tightly bound) than that of the antiparallel state. In order to better visualize this, we examine the simplified view depicted in Figure 1-10, where we see the transition as a spin-flip of the electron spin with respect to the proton spin. One should be cautioned that this property of "spin" is not literally a classical, spinning charged sphere, but rather quantum mechanical angular momentum is without classical analogy. The figure merely serves as an illustration of its behavior.

Reflecting back upon the physical basis depicted in Figure 1-10, as well as the hyperfine resulting emission structure depicted in Figure 1-9, one may recall that hydrogen is the most abundant element in the interstellar medium (ISM). However hydrogen, when found as a H_2 molecule is symmetric and therefore has no permanent dipole moment and hence does not emit at radio frequencies. Because H I (or neutral hydrogen) atoms are abundant as low-density regions in the ISM, they are detectable as the 21 cm (1420.405751 MHz) hyperfine line due to the two energy levels from the magnetic interaction between the quantized electron and proton spins. [26]

The hyperfine transition frequency of H I, ν_{10}, may be determined as follows:

$$\nu_{10} = \frac{8}{3}\, g_1 \left(\frac{m_e}{m_p}\right) \alpha^2\, (R_m c) \approx 1420.405751\, MHz$$

where $\nu_{10}$is the hyperfine transition frequency of H I, $g_1 \approx 5.58569$ is the nuclear g-factor for a proton, $\alpha \equiv e\text{^}2/(hc) \approx \frac{1}{1.37.036}$, which is known as the dimensionless fine-structure constant, and where $(R_m c)$ is the hydrogen Rydberg frequency [27].

1.8 MASERS

A MASER [28] is based on the principle of stimulated emission, and was first proposed by Albert Einstein in 1917. At the time of his proposal, it was not known that astrophysical MASERs had long ago and that they occur naturally in star forming regions of our universe. A MASER functions similar to a LASER, with a comparison provided by the simplified table below,

Table 1-1, Comparison of MASER and LASER Properties.

	LASER	MASER
Populated Medium	Atoms	Molecules
Emission Wavelengths	Visible (includes IR and UV)	Radio or Microwaves
Quantum Transitions	Major Energy Levels	Rotational Transitions

MASER or LASER emission is enabled to occur in a gas when a number of atoms or molecules have been elevated or "pumped" into an excited energy state, referred to as a population inversion. When these excited atoms or molecules transition or cascade down to a lower energy state, they will emit a single frequency photon. As the photon interacts with other particles in the gas, these other particles are pumped or cascaded as well, emitting other single-frequency photons, a process known as amplification of radiation. As this process occurs, it amplifies radiation at a specific frequency, since its basis of operation lies in the transition of energy levels where the energy of a photon, $E = hv$ (where E is energy, h is Planck's constant, and v represents frequency). To achieve LASER or MASER action, a sufficient number of atoms must be excited into a high energy state within a high density gas or a resonant cavity (where the cavity's resonant frequency corresponds to the transition energy of each atom's decay

energy, as given by the relationship above). As single or multiple energy decays occur (where each are associated with a corresponding photon emission), the whole gas transitions to a lower energy state almost at once. This action results in what is referred to as coherent radiation, which is the fundamental process whereby LASER or MASER emission occurs.

Einstein provided the fundamental theoretical understanding of the process of coherent stimulated emission as the basis by which stimulated radiation occurs. In his work, he treated electromagnetic radiation as quantized packets of energy, each having specific energy. In the case of a large collection of atoms which are full of excess energy, with each potentially ready and able to emit a photon at a random time in a random direction, as energy decay occurs. Then, as a stray photon passes by, the larger group of atoms is stimulated to emit their photons. The emitted photons will mainly travel in the same direction if the gas is in motion along the line of photon travel and have exactly the same frequency as the original photon, resulting in coherent photon emission action. As the small crowd of coherent photons move through the rest of the body of atoms, more and more photons will leave their excited atoms to join in the atomic parade of coherent photons, which in turn causes more photons to join. This results in a cascading or multiplying action of the behavior of those photons present. The term MASER refers to Microwave Amplification by Stimulated Emission of Electromagnetic Radiation. MASER radiation is different than classical electromagnetic radiation in that it is coherent. Due to its spatial coherence, MASER emission (like LASER emission) is able to maintain a high intensity over astrophysical distances since the microwave beam does not disperse as ordinary microwave radiation would.

1.9 OH Masers

Charles Townes was the first individual to build a laboratory MASER. However, long before LASERS were proposed and even long before MASERs, there existed MASERs in space. Astrophysical MASERs often occur naturally in star forming regions. OH MASERS are just one of the several species of naturally occurring astrophysical MASERs [29,30]. Among the known astrophysical MASERs in star forming regions are OH (1,665 MHz, 1,612 MHz, 1,667 MHz, 1,720 MHz), CH, H_2CO (4.829 MHz), H_2O (22.235 GHz), NH_3 (23.870 GHz), CH_3OH (12.178 GHz), HNCNH, SiS (18.155 GHz), HC_3N, SiO (43.122 GHz), HCN (89.087 GHz), and the H I recombination line where such emission arises from large concentrations of molecules located in interstellar clouds and other locations. In star forming regions, OH MASERS are observed as narrow spectral line emissions in the microwave portion of the electromagnetic spectrum,

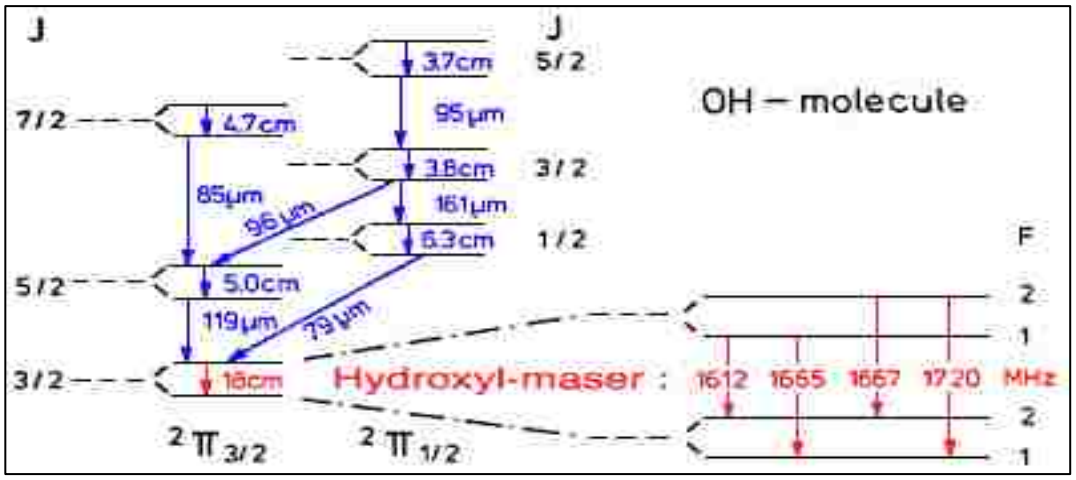

Figure 1-11, Hydroxyl MASER transitions. Fromhttp://laserstars.org/history/hydroxyl.html.

although the OH molecule has been observed to mase at 1612, 1665, 1667, 1720, 4660, 4750, 4765, 6031, 6035, and 13441 MHz as well. The 1720 MHz MASER transition of the OH molecule is known to be associated with supernova remnants that interact with OH molecular

clouds while the 1612 MHz transition is associated with late type stars.

It is important to understand the difference in the functioning of astrophysical MASERs as contrasted with laboratory MASERs, at least for those which have been designed and built to date. Emission from a laboratory MASER occurs as a monochromatic (single) frequency, where its single frequency corresponds to the energy difference between rotational energy levels of its gain medium. A MASER gain medium may be pumped (or excited to a higher energy level) through various means, however usually it employs a set of opposing partial mirrored surfaces, where excitation occurs through multiple reflections through the gain medium. Pumping creates a non-thermal energy population distribution, where a large portion of the molecules have been moved to a higher energy level. In contrast to a laboratory MASER such as was just described, electromagnetic emission which occurs from an astrophysical MASER is due to a single pass through the gain medium as quantum mechanical MASER "pumping", wher e the pumping action refers to the imparting of sufficient energy to the medium containing the molecules (or atoms) to cause them to be raised to a higher or inverted population energy level. In order to establish MASER conditions, the molecules must be excited to the same, higher, energy level. The excitation or pumping of an OH MASER must likewise occur with sufficient energy imparted to enable an overall transition of the medium to a higher rotational energy level, which would correspond to a single frequency of transition.

Of particular importance to a small antenna system, such as our own, are those OH MASER transitions where the MASER action is stronger and more common. Such OH MASER transitions correspond to 1612, 1665, 1667, and 1720 MHz as shown in (Figure 1-11). Massive star forming regions with high energy emissions are best for imaging using smaller antennas such as are had in our existing radio astronomy system. Of these, particularly Orion KL, W3OH,

W75N, and W51 are all excellent candidates for imaging. Some samples of these spectrums are

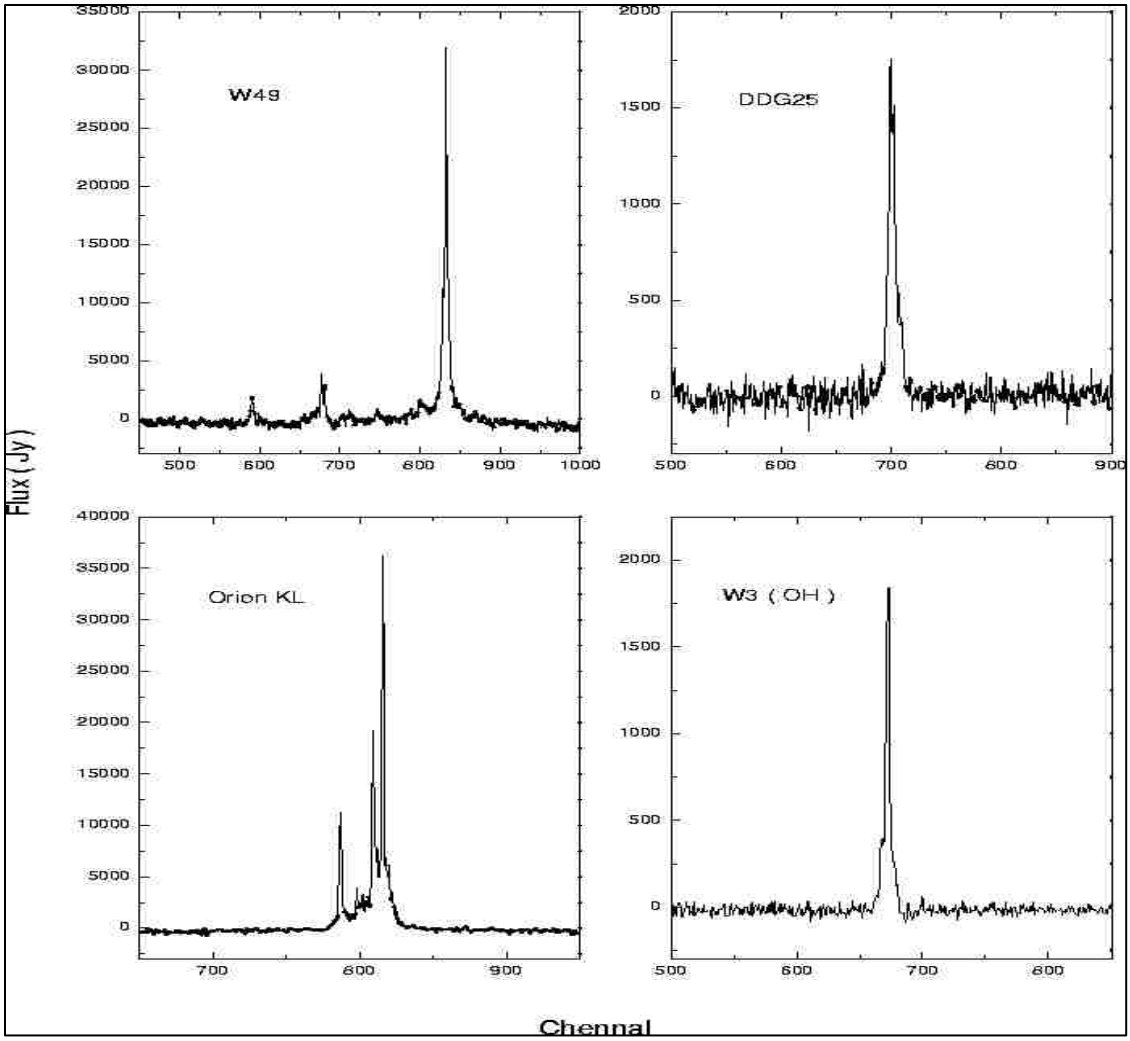

Figure 1-12: Typical OH Emission spectrums , from
http://aas.aanda.org/articles/aas/full/1999/12/ds8480/node4.html

shown in (Figure 1-12).

Chapter 2. Measurements and System Performance Characterization

A radio astronomy system requires many measurements, not only for observations, but for system operation and maintenance. One of the primary tools that was used to measure performance of a radio astronomy system, such as that described here, is the spectrum analyzer [31]. A spectrum analyzer displays signal amplitudes as a function of frequency and thus may be

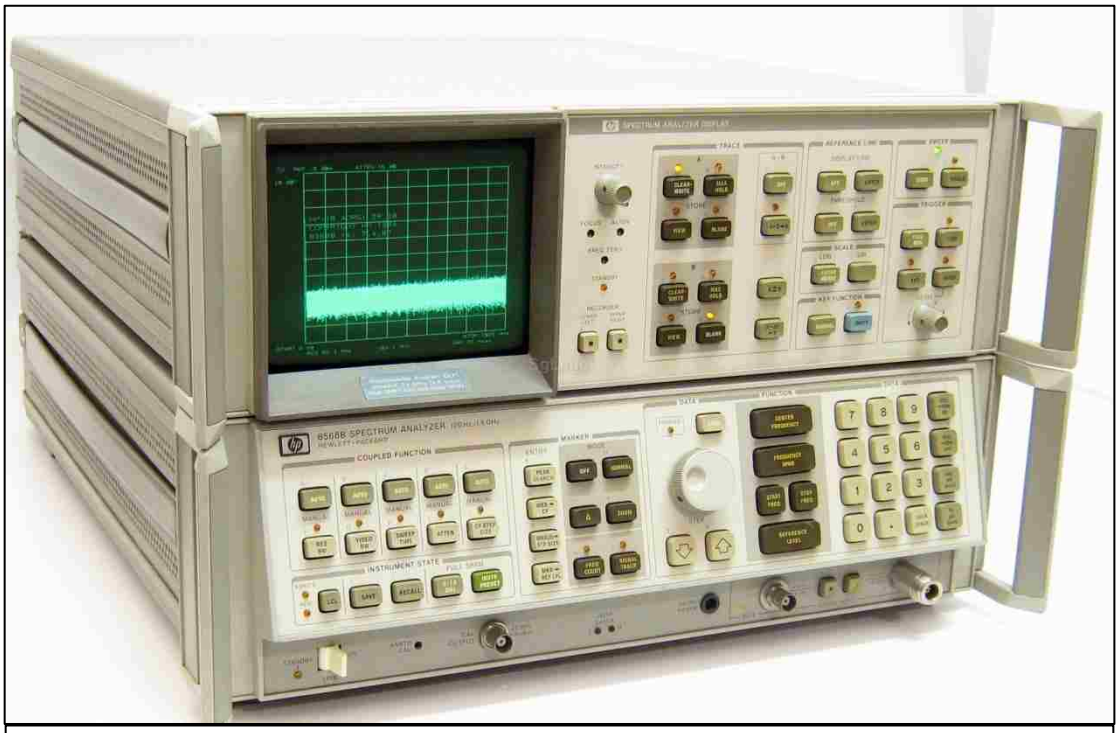

Figure 2-1: HP 8568B Spectrum Analyzer.

used to characterize various components and overall system responses. Furthermore, a spectrum analyzer may be employed for viewing the radio-astronomy spectrum. Due to its capabilities, the spectrum analyzer is an extremely versatile instrument. In our case, we employed the HP 8566 and HP 8568 scalar spectrum analyzers.

These instruments are used to measure several performance characteristics of a system

such as that described here. Among the measurements that may be made with a spectrum analyzer, there are several that are quite pertinent to our system, including:

- Component and system gain, (including antenna and LNA gain).
- Component and system sensitivity.
- System noise floor.
- Antenna tracking variation.
- 3rd Order Inter-modulation Distortion.

Although the first four of these measurements are to be expected, the fifth measurement determines the system response to the presence of local terrestrial interfering signals (such as cell phone tower transmissions) that are capable of undermining the measurements being conducted on a temporary basis.

2.1 The Decibel

The decibel (dB) is a logarithmic unit of measure that indicates the ratio of a physical quantity such as power relative to a specified or implied reference level [32]. A decibel is ten times the logarithm (base-10) of the ratio of two quantities. The decibel is a tenth of a Bel, which is a unit named in honor of Alexander Graham Bell. A decibel is expressed as follows:

$$I(dB) = 10 \log_{10}\left(\frac{I}{I_0}\right)$$

where I is the measured quantity, and I_0 the reference level.

The decibel is used for a wide variety of measurements in science and engineering due to its ability to express both very large and small quantities, especially when their range may vary over a wide magnitude and its ability to carry out multiplication of ratios by simple addition and

subtraction. The gains of amplifiers, attenuation of signals, and signal-to-noise ratios are often expressed in decibels. This practice is due to the fact that a logarithm is more readily able to express a larger dynamic range more easily than a simple linear ratio. There are many examples of dB applications within the field of electronics as applied to radio-astronomy, such as S-Parameter Plots, Bode Plots, and Antenna Radiation Plots, and more.

Often the units of the decibel are accompanied by a suffix, where the suffix refers to reference that is used for comparison when a decibel measurement is made. For instance, dBm indicates a reference of one milliwatt [33,34]. The decibel is often used to express a ratio. Some common ratios used in radio astronomy are illustrated in the following Table 2-1, showing a few power ratios and amplitude ratios as their amplitude ratios as $\sqrt{x}$ and dB equivalents 10 log10 x, where x is the ratio in question.

Table 2-1, Power Ratios and their Decibel Equivalent Representation

dB	30	20	10	6	3	0	-3
Power Ratio	1000	100	10	3.981	2	1	-0.5
Amplitude Ratio	31.62	10	3.162	2	1.41	1	0.708

A decibel (dB) is one tenth of a Bel (B), i.e., 1B = 10dB and represents a ratio between two power quantities of 10:1. The decibel is defined with respect to power, and not amplitude, so when one uses dB to express amplitude or voltage ratios, one must first square the amplitude, or use the factor of 20 instead of 10, as the dB logarithm. Other suffixes frequently used include dBV, (or voltage relative to 1 volt); dBm, (or power relative to 1 milliwatt) and where dBm usually references using a 50 ohm load, (where the associated voltage is 0.224 volts). In antenna measurements, dBi refers a reference of an ideal isotropic antenna radiator, which uniformly

distributes energy in all directions, and dBd refers to the forward gain of a measured antenna compared to a half-wave dipole antenna at the same frequency, where 0 dBd = 2.15 dBi.

2.2 The Jansky

Given the preceding discussion of measurements using the decibel, we are ready to introduce the Jansky. It is named after the earliest recorded radio astronomy pioneer, Karl G. Jansky. The Jansky is a measure of electromagnetic wave intensity given a fixed area with which to make measurements. The Jansky is a unit of flux density or spectral irradiance and is defined as [35]:

$$1 \text{ Jy} = 10^{-26} \text{ W/ m}^2 \text{ Hz}$$

Here, a single Jansky (symbol "Jy") is defined as 10^{-26} Watts per square meter per Hertz. In examining this unit and the reason for its use a little more closely, one notes that 10^{-26} Watts is a very small value, equal to -260 dBW/ m^2 Hz. Formally, the Jansky is:

$$S = \iint_{\text{source}} B(\theta, \phi) \mathrm{d}\Omega$$

where S is the monochromatic flux, B is the spectral radiance over a solid angle and the flux density is integrated over the entire solid angle, dΩ. In use, the flux density is normally integrated over bandwidth as well, as virtually all astronomical sources are not monochromatic, but rather have some finite spectrum associated with them.

One of the useful features of Jansky measurements is that the Jansky simultaneously includes aspects of antenna area and bandwidth measurements into a single quantity with which to compare intensities. In use, because the Jansky combines the these measurements together, it

provides a more concise measure of intensity, although it presumes a constant antenna area which may be true for large radio astronomy installations but is often not the case for research laboratories. Some standard observations using Janskys are provided here as a reference for comparison, which like the above discussion assume a constant antenna area:

Table 2-2, Comparison of Intensities from the Solar System (In Janskys) [Assuming a large antenna area]

Jansky (Jy)	**Source**
110,000,000	Radio-Frequency interference from a GSM phone transmitting 0.5W@1800 MHz at a distance of 1 km (RSSI of -70 dBm)[8]
20,000,000	Sun at 20 MHz (Karl G. Jansky initial discovery published in 1933)
4,000,000	Sun at 10 GHz
1,000,000	Milky Way Galaxy at 20 MHz
2,000	Milky Way Galaxy at 10 GHz
1,000	Sun at 20 MHz

Chapter 3. Antennas & Tracking

An antenna couples electromagnetic energy either to or from the medium to which it propagates into or out of an electronic circuit. There are many varieties of antennas. The simplest antenna, which is often used as a reference to which others are compared, is the so-called “dipole” antenna. By definition, an antenna converts electrical currents into electromagnetic energy and electromagnetic energy to electrical currents. As current from electron movement is driven through an antenna's elements, an oscillating magnetic field is created by the oscillating electric field associated with electron flow. The moving electron current creates an associated oscillating electric field along both driven and parasitic elements. A parasitic element is an element nearby to which energy has been coupled. The time-varying fields radiate into space as a moving transverse electromagnetic wave. During reception, oscillating electric and magnetic fields exert force on the electrons in the antenna elements, creating oscillating currents in the antenna.

Antennas may include both reflective and directive elements that are not connected to the primary receiving element, which can be considered either as parasitic elements or as reflective elements. A reflective element simply reflects the electromagnetic energy to the primary element (or to the parasitic elements, while a parasitic element resonates and then re-radiates energy back to the primary receiving element, functioning electrically “in-phase” so as to increase the intensity of the primary received signal. There are a large potential number of reflective elements, either intended or unintended. Often employed in radio-astronomy, parabolic reflectors and horns direct radio waves into a beam or other desired radiation (re-enforced) pattern and by so doing provide “gain” or amplification of the original signal by concentration of the original energy. It should be noted that both directional and non-directional antennas exist.

Ideally, an antenna which radiates equally well in all directions is defined as an isotropic radiator, although there are no true isotropic antennas. When an antenna's radiation pattern is intended for one or several specific directions, it is referred to as a directional or gain antenna. When referring to antennas, the term gain refers to a comparison with an omnidirectional antenna or isotropic radiator. However, often an omnidirectional antenna refers to non-directional radiation having just one polarization, where other polarizations may have other qualities. Using this relaxed terminology, a common example of an omnidirectional antenna is the vertical antenna, consisting of a metal rod (often, a quarter of a wavelength long), which is vertically polarized. A dipole antenna consists of two such conductors, extending in opposite directions, with a length that is often a half wavelength long.

Antennas more complex than the dipole or vertical designs are usually intended to increase the directivity and hence the antenna gain. Antenna "gain" is then the concentration or directivity of the signal into a particular solid angle of space, as opposed to the spherically uniform radiation of the ideal isotropic radiator. Greater directionality can be had by using beam-forming techniques such as a parabolic reflector. Although a half-wave dipole may be taken as a reference instead of the isotropic radiator, its use is rare. However, in such cases, the gain is given in dBd (decibels over dipole), where 0 dBd = 2.15 dBi.

3.1 Antenna Radiation Patterns

All antennas have a radiation pattern associated with their performance characteristics. For example, Figure 3-1 shows a 2-dimensional polar plot (as horizontal cross section) for a specific Yagi-Uda-antenna. The concentric circular lines shown in this plot mark points having a 3db power level compared to an isotropic radiator that radiates equally in all directions. For

many antennas, their radiation patterns often contain "*lobes*" (or areas of maximum radiation) at various angles, separated by "nulls", at angles where the radiation becomes insignificant or falls

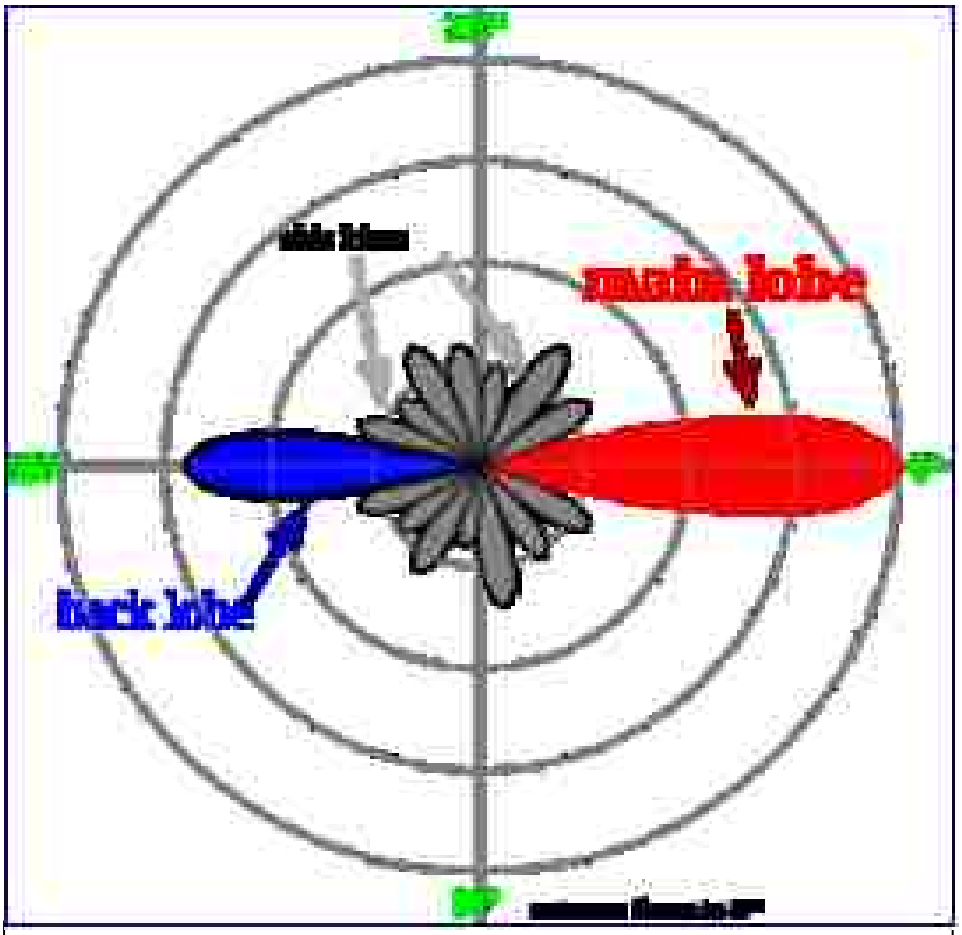

Figure 3-1: Antenna radiation pattern for a Yagi gain antenna showing its main, back, and side-lobes. The concentric circles are each spaced with at 3-dB intervals with reference to an isotropic radiator. Courtesy of "en.wikipedia.org/wiki/Main_lobe".

to zero. This behavior is due to interference, which is caused from the time relationship of electromagnetic waves arriving at distant points in phase with one another versus zero radiation at other angles where each radio wave arrives out of phase with one another. Since, when an arriving wave occurs in phase with the previous wave, the waves add, and when the wave is exactly out of phase, they can cancel. In antennas which are directional, the lobe having a magnitude larger than the others is called the "main lobe", while other lobes are called "side lobes".

3.2 Antenna Beamwidth Measurements

There are many varieties of antennas for different purposes. Of these, parabolic antennas

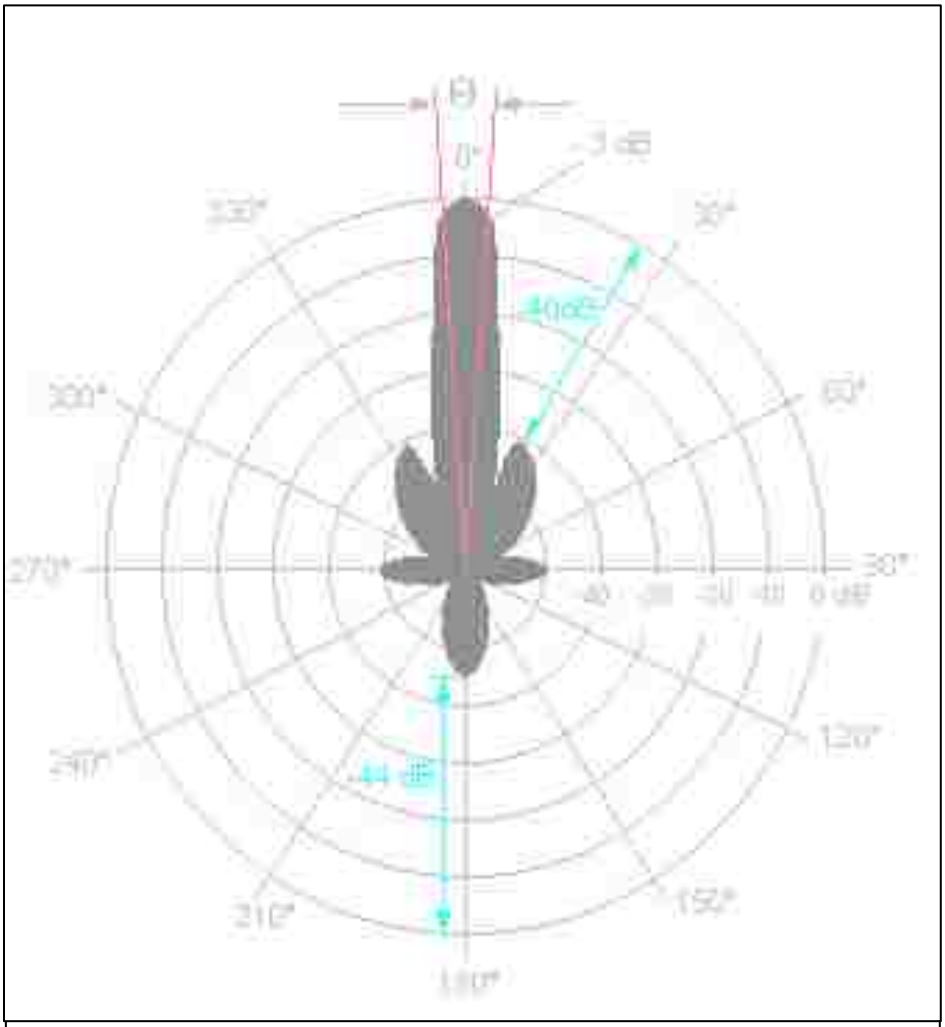

Figure 3-2: Polar Coordinate graph of a dish antenna, showing its main lobe, side lobes, and rear lobes. Here, each circular ring represents 10 dB, with reference to isotropic radiation. Courtesy, http://www.radartutorial.eu/06.antennas/an05.en.html

are typically designed to provide a high gain in a very narrow beam of directionality and are the most commonly used antenna at this time for the purposes of radio astronomy. The angular width of the intended received beam by any high-gain antenna is measured by the "Half-power beam width" (or HPBW) [36]. The HPBW is defined as the angle of separation between the measurement points of the antenna radiation pattern, where the power (with reference to the

antenna central maximum) drops from its peak power to one-half of its maximum value (or -3 dB down from its peak). For parabolic antennas, the HPBW, θ, is given by [37].

$$\theta = \frac{k\lambda}{d}$$

where, λ is the antenna wavelength in meters, d is the diameter of the parabolic radiator in meters, and k is constant which depends on the shape of the reflector and its feed illumination pattern. For a "typical" parabolic antenna k = 70, when θ is in degrees [38, 39].

Figure 3-3: The Arecibo Radio Astronomy 305 meter Observatory. From http://www.geog.nau.edu/

For a typical 2 meter diameter dish antenna operating on C Band (~5 GHz), this formula yields a beam-width of about 2.6°. For the Arecibo antenna at 2.4 GHz the beam-width is 0.028°. It can be seen that parabolic antennas are able to produce very narrow beams, and therefore aiming an antenna with such a narrow beam may potentially become problematic, due to pointing accuracy.

Lastly, there is an inverse relation between antenna gain and its beam width. By

combining the beam width equation with the gain equation, this relationship may be expressed [40, 41] as:

$$Gain \approx \log_{10}\left[\left(\frac{\pi k}{\theta}\right)^2 e_A\right]$$

where, e_A is the aperture efficiency, which is typically between 0.55 and 0.70 .

3.3 Effective area or aperture

The "Effective aperture" of an antenna refers to its effectiveness to convert a passing electromagnetic wave into a signal of electrical energy at the antenna's electrical terminals, which is expressed in terms of an equivalent area. As an example, if an electromagnetic wave passes a fixed location with a flux of 1pW/m^2 and assuming that this antenna has an effective area of 12m^2, then the antenna could ideally deliver 12pW of radio frequency power to the receiver (or 30 microvolts rms at 75 ohms), assuming an e_A of 1. An antenna's effective area averaged over all directions is a function of λ and in this case is equal to $\lambda^2/4\pi$. Here, the "gain" of an antenna is defined by the standard where the average gain, when totaled over all directions, for an antenna with 100% efficiency is equal to 1. Then the effective area, A_{eff}, in terms of the gain, G (in dB) , for a given direction for an antenna is given by:

$$Gain, G = 10\log_{10}\left[\frac{4\,\pi\,A_{\text{eff}}}{\lambda^2}\right] = 10\log_{10}\left[\frac{\pi^2 d^2}{\lambda^2} e_A\right]$$

where, d is the parabolic antenna diameter in meters, and λ is the wavelength in meters.

It should be noted, that not all reflective nor parasitic elements are 100% efficient at reflection nor re-radiation, respectively. For an antenna with an efficiency of less than 100%, both effective area and gain would be less than the ideal described above as each are reduced by

the same amount. In any case as far as efficiency is concerned, the relationship above between gain and effective area applies.

3.4 Electromagnetic Wave Reflections for Conductors vs Dielectrics

When an electromagnetic wave impinges on the surface of an object, two waves result, where one is transmitted through the object and the other is reflected. Depending upon the nature

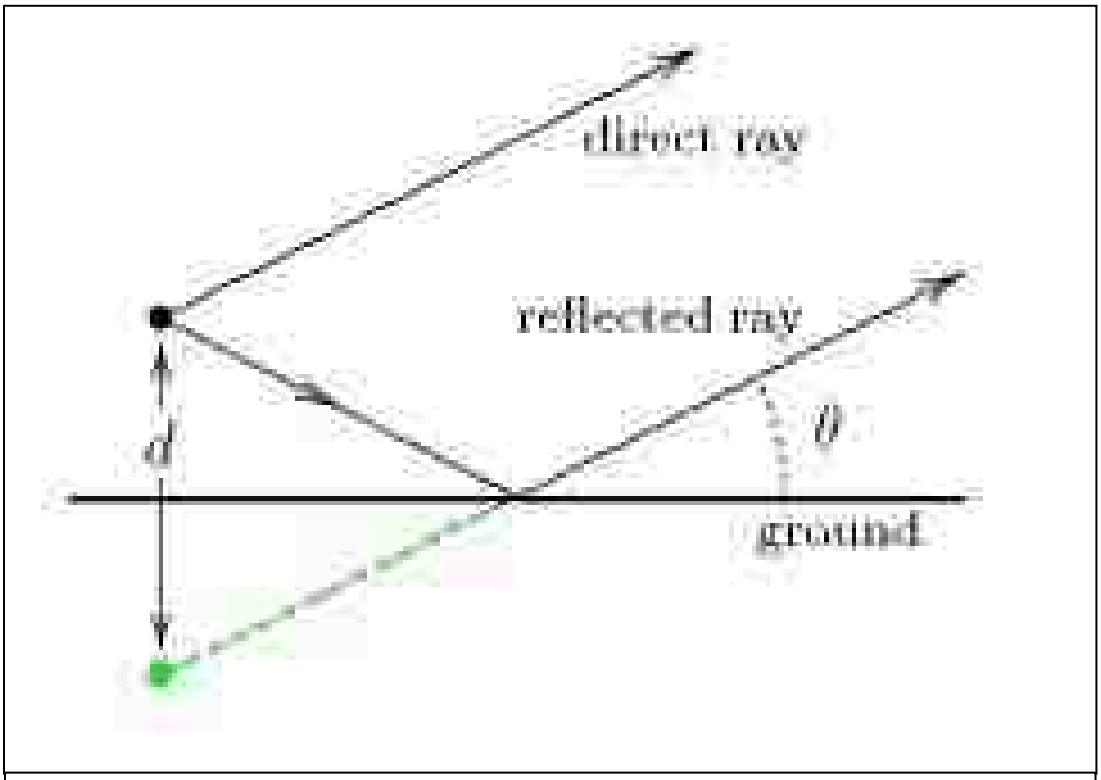

Figure 3-4: This drawing depicts radio frequency wave reflection off of a reflective surface and shows how the reflected rays may be treated as originating from a source behind the reflector.

of the substance of the object, the degree to which the wave reflects or is transmitted becomes evident. When the object is a conductor, the transmitted wave is negligible and the reflected wave has almost the same amplitude as the incident one. When the object is a dielectric, the

fraction of reflected wave depends on the angle of incidence. For a dielectric, when the angle of incidence is small (that is, the wave arrives almost perpendicularly) most of the energy enters the surface with very little reflected. When the angle of incidence is near 90° (grazing incidence) almost all the wave is reflected. An example of this phenomenon is further illustrated in Figure 3-4, where a electromagnetic wave strikes a dielectric surface and is reflected. Here, one notes that as the wave is reflected by the dielectric surface, the reflected wave behaves the same as though another emission is emitted by an "image antenna", located under the reflective surface.

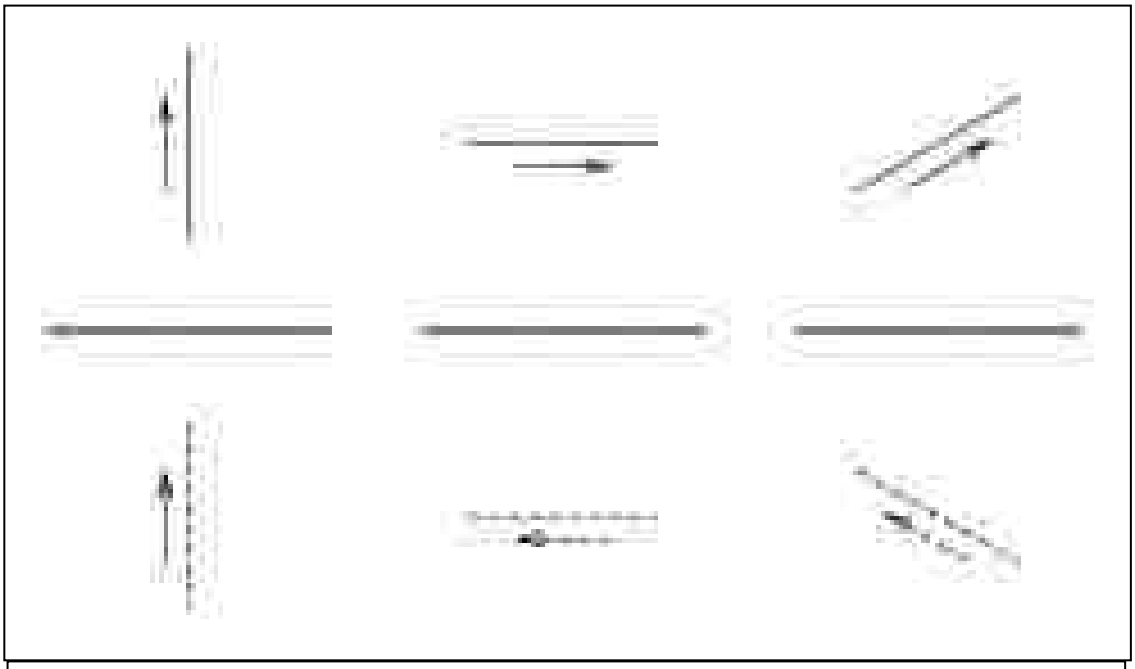

Figure 3-5: This figure illustrates how reflections of electromagnetic waves occur from both dielectric and conductive surfaces with respect to the polarization properties of the incident electromagnetic wave.

Thus this reflection off the dielectric surface functions just the same as though there were a real antenna with an image of the antenna reflected by the dielectric. However, if the surface of the dielectric contains irregularities, the image will appear fuzzy, due to the variation in reflections.

In addition, reflection of an electromagnetic wave depends on the polarization of the incident wave. If the refractive index of the dielectric surface is larger than the refractive index of the air

(where the refractive index of air is 1), the direction of the component of the electric field parallel to the dielectric surface reverses on reflection, equivalent to a phase shift of π radians or 180°. This and other reflections are illustrated in Figure 3-5. Here, the parallel component of the electric field reverses sign while the perpendicular component remains aligned as it was before reflection. This pattern is true for both dielectrics and good electrical conductive surfaces.

As reflections in the ionosphere occur, the angles by which the waves are reflected may vary greatly. Hence, the resulting electromagnetic rays are composed of several reflected rays

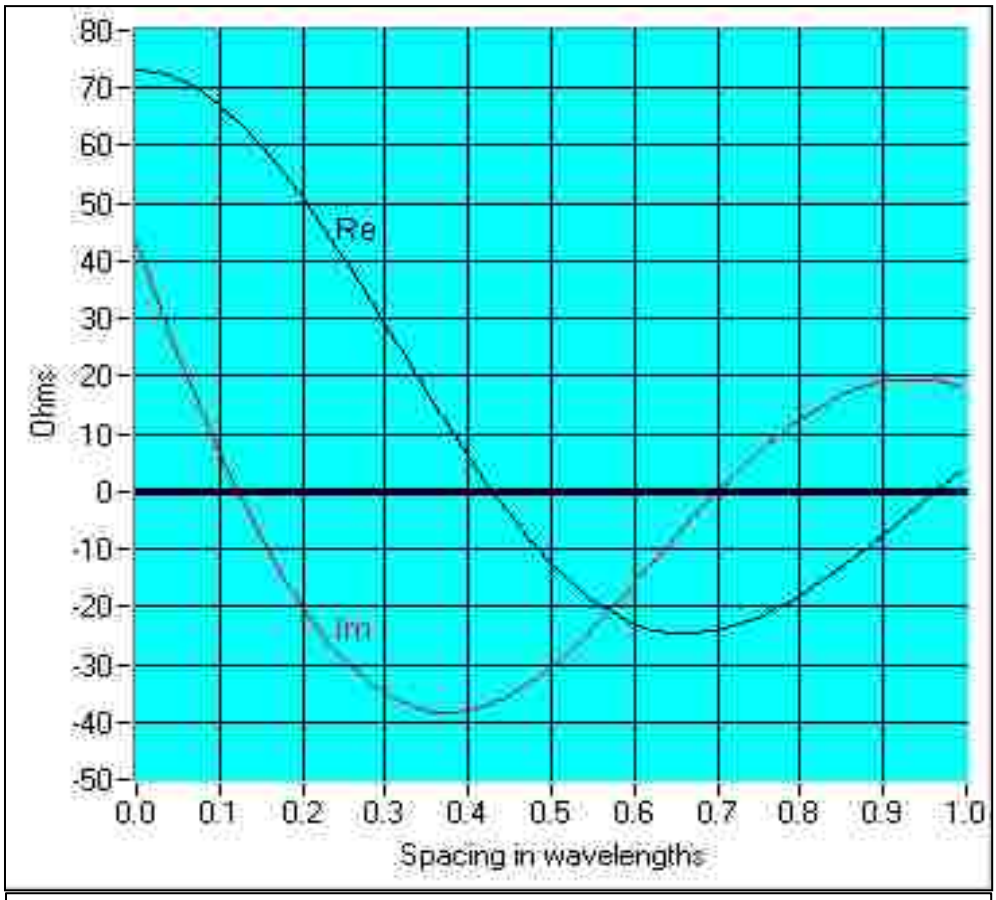

Figure 3-6 This graph depicts mutual impedance between parallel $\frac{\lambda}{2}$dipoles with differing spacing, where Re and Im are resistive and reactive components of the interaction of mutual impedance. From Kai Fong Lee (1984). *Principles of Antenna Theory*. John Wiley and Sons Ltd.

that add with different phase shifts.

3.5 Antenna Interactions

For an antenna having several interacting coupled elements, current circulating in an element induces currents in others. This interaction is illustrated by plotting the mutual impedance Z_{12} between two antenna elements, which behaves somewhat similar to the mutual inductance property of coupled inductors, where changes in one effect changes in the other. Here, the mutual impedance Z_{21} between two antennas is defined as:

$Z_{21} = \frac{v_2}{i_1}$ where i_1 is the current flowing in antenna 1 and v_2 is the voltage that would have to be applied to antenna 2–with antenna 1 removed–to produce the current in the antenna 2 that was produced by antenna 1.

The graph of Figure 3-6 illustrates the interaction of mutual impedance between parallel half wavelength (or λ/2) dipoles. Here, the curves Re and Im represent the resistive and reactive components of impedance.

From the definition above, we can define the general relationship between the currents and voltages in a group of several parasitically coupled antennas. We may write:

$$V_1 = i_1 Z_{11} + i_2 Z_{12} \text{ and } V_2 = i_1 Z_{21} + i_2 Z_{22}$$

where:

- v_i is the voltage applied to the antenna *i*.
- Z_{ii} is the impedance of antenna *i*. and
- Z_{21} is the mutual impedance between antennas i and j.

T his coupling between antenna elements is a result of “Lorentz reciprocity [42], where this mutual interaction refers to the effect of time varying electric currents as sources with the resulting electromagnetic fields described by Maxwell's equations applied to linear media. Lorentz reciprocity states that the relationship between an oscillating current and a resulting electric field is unchanged if one interchanges the points where the current is generated and where the field is measured (and vice versa). When Lorenz reciprocity is applied to parasitic antenna elements that are coupled, then mutual impedance of a first element due to a second is the same as the mutual impedance of the second element due to the first. Another application of Lorentz reciprocity suggests that antennas will work equally well for transmitting as receiving, and that an antenna's transmitting radiation and receiving radiation patterns are the same.

As one applies the mutual impedance relationship above to a set of coupled parasitic antenna elements, the non-active i_i terms are set to zero for parasitic elements, where a parasitic element is defined as non-active elements which reflect, absorb and/or reradiate electromagnetic energy. When two antenna elements are perpendicular in polarization, the coupling of energy between the two can be zero, rendering the mutual impedance between the two antennas also to be zero. Such is the case for “crossed dipoles”, where circular polarization of signals is required or when one wishes to select an opposite polarization.

3.6 Antenna Gain

Antenna gain is a measure of the amplification imparted by a radiator, as imparted through directivity. Measurements of antenna gain, when measured in dB are with respect to an ideal isotropic radiator, where an isotropic radiator is an antenna which radiates equally in all directions. For an antenna to have “gain” or amplification, gain in a preferred direction is had at

the expense of radiation in some other direction.

As an example, suppose we start with a simple theoretical isotropic radiator which has a gain of 1 (or no gain and no loss) and therefore radiates equally in all directions, (with a spherical radiation pattern). Now, suppose we take all the energy from one half of its spherical pattern from the isotropic radiator and cause it to be present in the other half of the spherical radiator, such that one half of the radiator has no radiation at all while the other half has twice as much radiation as the original isotropic antenna radiator had. To find the gain of this example half-spherical antenna radiator, we would use the gain relation:

$$\text{Gain} = 10 \log_{10}\left(\frac{P_{half\ wavelength\ radiator}}{P_{isotropic\ radiator}}\right) = 10 \log_{10}(2/1) = \underline{3.0\ \text{dB}}$$

Thus, we see that an antenna which radiates over half the spherical radiation area as compared to an isotropic radiator has twice as much gain (or 3 dB gain) over one which radiates uniformly in all directions.

3.7 Parabolic Antenna Gain

For radio astronomy applications (as well as others), the gain of a parabolic antenna is the ratio of the power received (or transmitted) by the antenna to that of an isotropic antenna where the gain of a parabolic antenna is[43]:

$$Gain = 10 \log_{10}\left(\frac{4\pi A}{\lambda^2}\, e_A\right) = 10 \log_{10}\left[\left(\frac{\pi d}{\lambda}\right)^2 e_A\right]$$

where A = area of the antenna aperture (in square meters), d is the diameter of the parabolic reflector (in meters), λ is the wavelength desired (in meters), and e_A is the aperture efficiency, where typical parabolic antenna e_A is between 0.55 and 0.70.

The larger the aperture as compared to the wavelength of operation, the higher the antenna gain. Since the gain of the antenna increases with the square of aperture width to wavelength, large parabolic antennas tend to have very high gain. For example, a 25-meter-diameter antenna (often used in radio-telescope arrays) at a wavelength of 21 cm (1.42 GHz), would be expected to have an approximate maximum gain of 140,000 or 50 dBi (above the isotropic reference level).

Aperture efficiency, e_A , accounts for various losses that result in the reduction of gain for an antenna from the maximum that is possible with a given aperture. The major factors reducing the aperture efficiency in parabolic antennas are:.(1) Feed spillover (where radiation of the antenna feed falls outside of the edge of the dish), (2) Feed illumination taper (where the antenna feed is non-uniform across the aperture), (3) Aperture blockage (such as where the antenna feed point is located in front of the beam path), and (4) surface errors in the shape of the reflector.

3.8 The Celestial Coordinate System

In order to view or image various objects in the heavens, it is of great value to be able to identify those objects in a way where they may always be found. When such objects are required to be imaged over an extended period of time, some means of observing while compensating for the rotation of the earth must be found. The celestial coordinate system [44] was devised to enable location of fixed objects in the heavens regardless of the rotation of the earth. The celestial equatorial coordinate system is based upon the idea of an infinite sphere centered on the earth (see Figure 3-7). Like coordinate systems used for identifying geographical locations used on the earth, the celestial coordinate system is used to identify locations of objects observed in

the heavens around the earth. However, since the earth is constantly rotating, unlike the geographical coordinates for locating points of the earth, the celestial coordinate system must be fixed with reference to the objects in the heavens. So the constant speed of rotation of the earth is used to translate between earth centered coordinate and the fixed celestial coordinates of the heavens. Similar to the familiar latitude and longitude of the earth centered coordinate system,

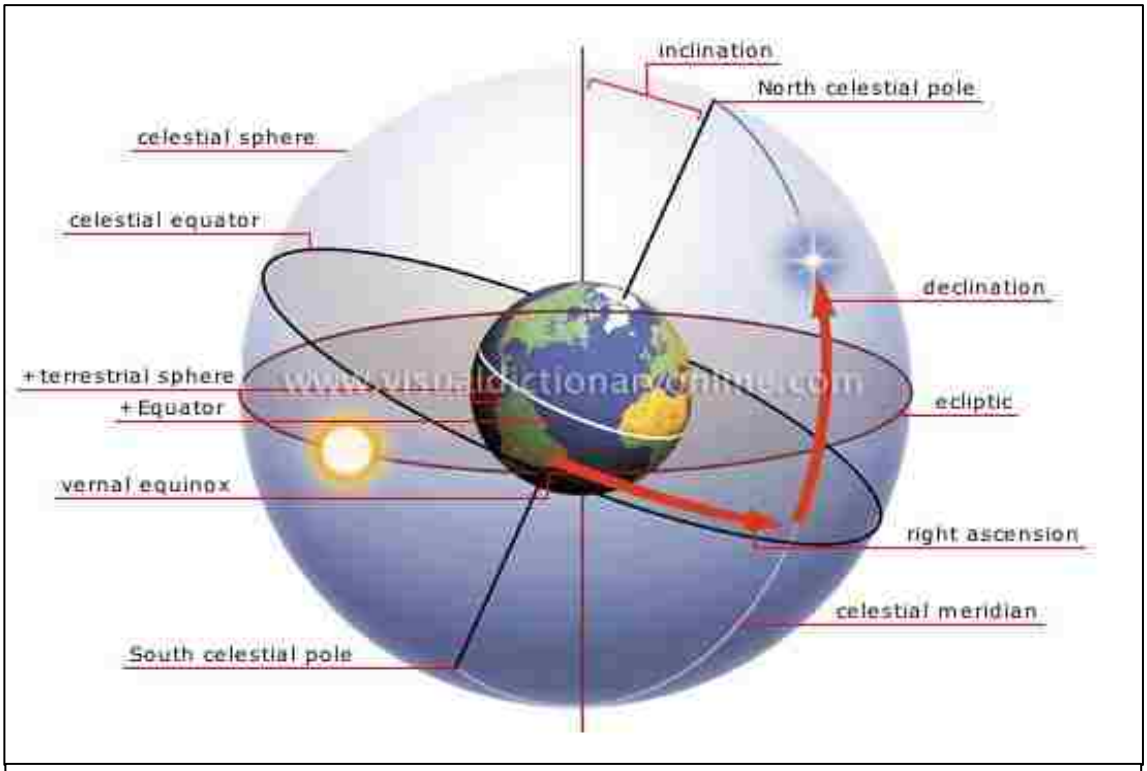

Figure 3-7: The celestial sphere, showing right ascension [RA] and declination [DEC]. From 'visual.merrium-webster.com/astronomy/astronomical-observation/celestial-coordinate-system.php

two coordinates are also employed to identify locations in the celestial coordinate system, as seen in Figure 3-7. Here, one notes that celestial equator is coplanar with the earth's equator and perpendicular to its axis of rotation with the *north celestial pole* directly above the earth's north pole and the *south celestial pole* directly below the earth's south pole. The *declination* is the coordinate of a celestial object located between these two poles and is measured from the celestial equator. Its range of values is from 0° at the celestial equator to +90° at the north

celestial pole and from 0° at celestial equator to -90° at the south celestial pole. The second coordinate in the celestial equatorial coordinate system is *right ascension*. This coordinate functions like longitude, where Greenwich is the zero point for longitude, right ascension also has a zero reference point called the *Vernal Equinox Point*. The Vernal Equinox was chosen as the zero reference point as it is a point in rotation of the earth (of the earth's celestial equatorial plane which intersects with the sun's ecliptic plane or the point at which the earth's rotational plane and the sun's rotational plane intersect. Right ascension is measured in (sidereal) hours, from 0h to 24h east from the Vernal Equinox Point, where east is the direction of increasing right ascension. When considering a RA coordinate, the half-circle with right ascension of zero hours is called the *0 hour circle*. So given a full circle of 360°, a 1 hour right ascension is 15° [45].

3.9 Azimuth and Elevation Antenna Positioning System

The azimuth and elevation antenna positioning system is designed to provide absolute steering of the antenna in any direction possible within its field of view and capabilities. Moreover, for radio-astronomy, the positioning system must provide tracking. Here, tracking refers to compensation for the earth's rotation as well as position movement of the imaged objects in time, with reference to the earth's position. Often in radio-astronomy, a position of an object is given in celestial coordinates, where the celestial coordinates are unchanging in space however do appear to change when viewed from a fixed position on the earth's surface, due to the earth's rotation. These coordinates are always stated with reference to a particular date and time. To first image an object, the given coordinates are first input into a computer controller positioning system, where the on-date RA and DEC are generated from the date and time. The date and time serve as input variables to generate RA and DEC with compensation made for

precision observing as changes in time occur.

In general, an azimuth and elevation positioning system is composed of two primary components. These are the positioning motor with gear box as well as the position controller. The positioning motor and gearbox is composed of: (1) either a stepper or servo motor, (2) gears to increase torque while reducing speed, (3) position sense detection, and (4) limit switches. The position controller is comprised of a power source and control electronics. Although the control electronics may take many forms, frequently it is composed of both a micro-controller with associated interface circuits for the motor positioner capable of both sensing position as well as controlling the heavy currents associated with the positioning motor. If that motor is a DC servo motor, then the interface circuitry contains what amounts to a DC amplifier (where a DC amplifier is roughly equivalent to an audio amplifier, but is directly coupled rather than AC coupled) which is capable of supplying both positive and negative ampere level currents at varying voltages, in order to properly control antenna position. Alternatively, if the positioning motor is a stepper motor, then the interface circuits supply high-current multi-phase stepped voltages appropriate for the stepper motor in question (stepper motors may have two to many phases, so the number of phases and positioning increments associated with each step actuation are dependent upon the stepper in question).

The positioner employed for the 1st generation system is based upon a set of orthogonally mounted positioning stepper motors and gearboxes manufactured by M-Squared which provide both 0-90 degree elevation and over 360 degree rotation. The controller that is associated with this is also manufactured by M-squared. It consists of a single enclosure, having two power supplies (one for azimuth and the other for elevation) as well as two control circuit boards, each having a single micro-controller (manufactured by Microchip) with programming firmware.

Each circuit board contains all interface circuitry necessary to interface with both an external control computer as well as to control both the remote positioning motor/gear-box assemblies. The control box may either be operated manually or under computer control.

Functionally, the controller is very specific and quite limited; its purpose being to control and read back positioning information from the antenna motors/gear box. In itself, the controller is not capable of advanced functions such as tracking an imaged object or compensating for the earth's rotation. However, when coupled with a computer and with appropriate tracking software running, the corrective commands emanated by the computer may be fed to the positioning controller in order to correctly track or compensate for the earth's rotation.

In order to illuminate system operation, it is of value to note the data sequences between the control computer (running the tracking software) and the controller. Here, we have two data streams emanating from the computer in a sequenced fashion. One stream is for elevation and the other is for azimuth. Both data streams are in the data format required by the controller for correct operation. As the earth rotates, the data streams continue to emanate into both ports of the controller as a regular temporal sequence of data streams.

3.10 Manual Operation of Azimuth and Elevation Antenna Positioning System:

In addition to the automated modes of antenna positioning operation described, one may operate the system of antenna positioning in a static or transit mode. Here, the antenna is positioned so as to intercept the source of interest without tracking. Hence, in this mode, one will observe the amplitude of the source grow stronger until it reaches a peak and then grow weaker as the antenna pointing passes by the region of interest, also known as "transit". Although this

mode of operation is not one of the more popular modes of operation, it may be found useful for a variety of reasons, including equipment testing and troubleshooting. Transit mode is natural to use when there is a concern regarding the correct operation of the automated tracking system. Lastly, manual operation may always be employed as a backup means of system operation, when the tracking system is non-operational.

3.11 Positioning Precautions:

By its very nature, a large antenna may be subject to some very significant forces during wind loading. Although the antennas of this system weigh several hundred pounds, due to their large surface areas they may also serve unintentionally as sails during wind loading. Depending upon the maximum wind speed, the forces involved can be quite significant. In the case of a parabolic antenna, large winds ma y be catastrophic, particularly when accompanied by snow, rain or sleet. In order to minimize unbalanced lateral forces, it is recommended that an antenna be "stowed" or parked in a position which minimizes the lateral forces. Normally this is associated with a so called "zenith" position which occurs when the antenna is pointed straight up into the sky and away from the earth. Unfortunately, circumstances may prevent such procedures from being correctly followed on occasion. Figure 3-8 shows the results of one such incident, taken after the occurrence of gale-force winds. The antenna had been left in a position pointed to the south and at a 30 degree angle with respect to the horizon. When the gale-force winds occurred, the upper nearing components were completely stripped out of the mountings by ripping the mounting bolts out of their threaded surfaces. Again, an significant force had obviously been present to cause this amount of damage.

Figure 3-8: Antenna broken and blown off its azimuth/elevation mount as a result of a gale-force windstorm. Photo is courtesy of the author

3.12 Antenna Weight and Counter Weight

Large parabolic antennas are often quite heavy and often present a wind load like a large sail. The antenna on the roof of the Eyring Science Center has a weight of approximately 270 pounds, while the weight of the antenna originally intended to form a baseline extension located on the top of the old KBYU broadcast building in Springville has a weight of approximately 370

pounds. Each of these antennas has a system of counterweights which equalize its static balance.

Each of the antenna's azimuth and elevation motor gearbox assemblies have a pivot point where the antenna rotates for its respective axis. In order to rotate a mass around such a pivot point, there is a torque that must be applied. The required torque for rotation is the (mass)*(radius arm)*(the net gravitational acceleration). The motors and gears which drive the antenna each have a maximum torque at which they are capable of operating. Potential antenna positioning errors and even ultimately failure of the antenna positioning system are potential results associated with exceeding maximum system torque limits.

Static balance is the balance of the antenna with its counterweight when the antenna is not moving. It ignores any wind loading, dynamic loading, or snow loading. Static balance is achieved when the counterweights are equal to the antenna weight, as measured at the pivot point of either the azimuth or the elevation rotors.

Dynamic balance includes changing or moving forces such as wind loading. Wind loading may potentially be significant, yet its duration tends to change with both the direction and magnitude of prevailing wind gusts. Currently the design, as implemented, does not compensate for dynamic loading. This is an inherent limitation of the design. The principle method by which the system is able to overcome dynamic loading is to have more than sufficient torque in order to always achieve necessary torque for movement and control. However, given the large surface area of the antenna dishes (113 square feet for the Provo and Springville antenna dishes while it is 707 square feet for the Spanish Fork antenna dish) as well the potential for high winds it is not always possible to have sufficient torque for antenna positioning under all possible conditions. When high winds are present, it may be necessary to place the antenna in

such an orientation that would minimize the antenna wind loading forces that may be present, for safety reasons. It should also be mentioned that the antenna base assemblies, upon which the antenna systems are mounted, are counter weighted as well in order to prevent any of the antenna bases from movement in the presence of wind.

Other force and balancing considerations for the system include snow loads, which tend to occur in winter when the antenna is stowed (or put to rest in a upright or “bird-bath” position). Hence the stowed position will tend to accumulate the greatest amount of snow as a result of inclement snowy weather. The presence of snow can represent a great deal of weight to the antenna. Since the antenna is statically balanced without snow, the presence of accumulated snow will throw off the static balance and potentially may result in errors and a stall of the positioning system (or its inability to move the antenna). In order to mitigate such a circumstance, the antennas may be rotated sufficient to enable cleaning of their surfaces for snow removal or be pre-fitted with means of snow melting, such as plumber's heat cables (used to prevent pipes from freezing through the use of electrical heating elements.

3.13 Tracking Software

Tracking for a terrestrial based radio-astronomy system provides the means to dwell on a radio-astronomy image or source for an extended period of time through the use of means to compensate for the rotation of the earth. In this system implementation at BYU the tracking software calculates the new antenna pointing position as the earth rotates and sends this position to the azimuth/elevation controller. Since the controller is connected to the azimuth and elevation positioning motors with gearboxes, it is these which finally move the antenna

according to the new positioning coordinates and which compensate for the the earth's rotation

The software used in the first generation system is the "Nova" antenna positioning and tracking software, written by Michael Owen[[46]], and distributed by the American Radio Relay League [ARRL] [[47]]. Some of the relevant features of Nova for radio astronomy include [[48]]:

- Tracks Moon, Sun, planets, cosmic noise sources, cold sky and displays the star background.
- Ability to capture data and save to disk, or send as e-mail.
- Built-in auto-tracking [compensation for earth rotation] for turning azimuth /elevation antennas.
- Keplerian Elements using AMSAT/NASA format, with updating via the Internet.
-

3.14 Building the System

The radio astronomy system required a considerable amount of time to build and

Figure 3-9: Mounting the 4-meter dish upon its azimuth/elevations motor/gearbox assembly.

assemble, as well as the help of others during the process. Our antenna base was built by Wes Liferith, of the Physics machine shop, according to our specifications. We had modeled the antenna base after existing commercial units and performed the calculations to weight it down

during windstorms. In terms of weight, approximately 200 bricks at approximately 21 pounds each (for a little over 2 tons of mounting base weight) were used for the 4 meter dish that is located on the top of the Eyring Science Center. The 4 meter dish, itself, came as a kit in a set of

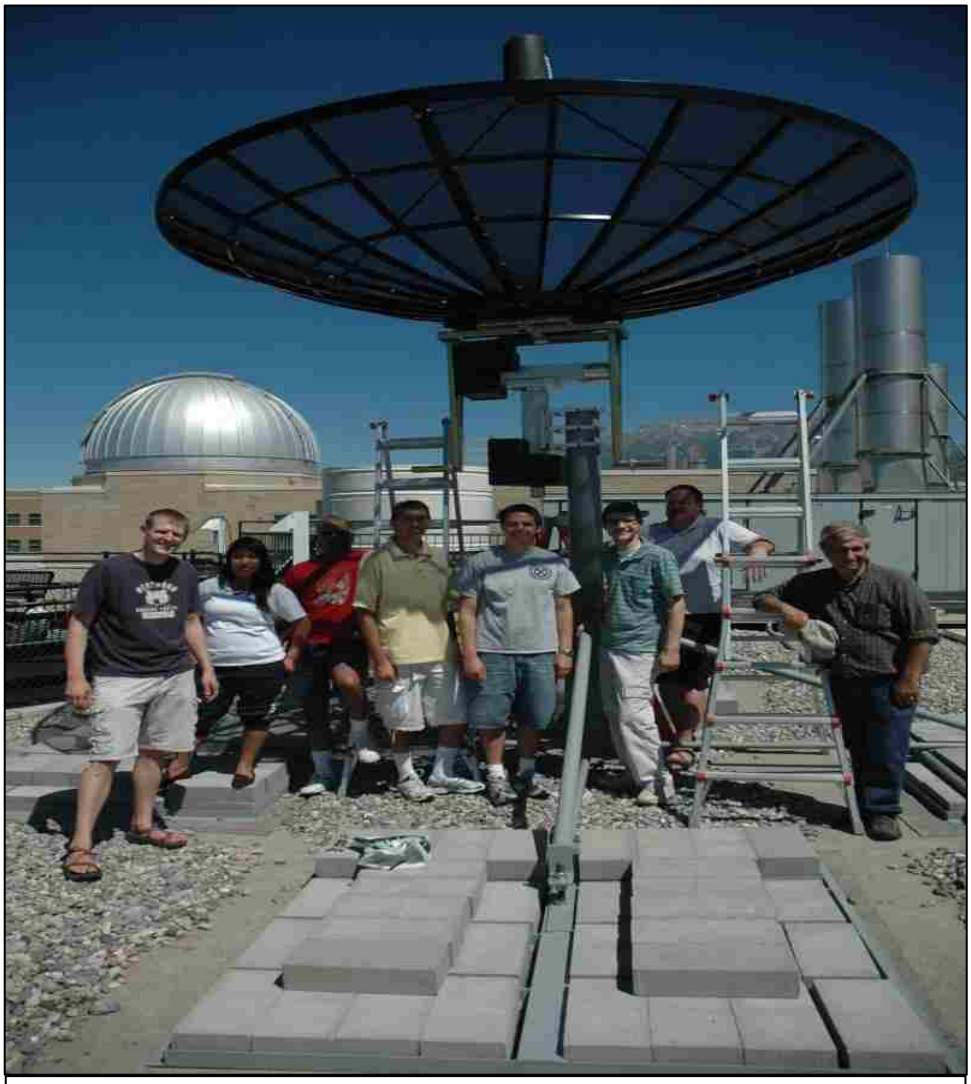

Figure 3-10: Antenna mount complete, on the roof of the Eyring Science Center.

its component parts and was assembled by hand along with assembly of the azimuth and elevation positioning motor and gear assemblies. However, when it can time to place the antenna upon the azimuth and elevation positioning motor and gear assemblies, several students

and a professor worked together to assemble these large pieces as seen in these photos.

3.15 Alignment of Parabolic Reflector with Resonant Cavity

The calibration of the dish antenna itself was primarily that of alignment of the focus of the reflecting dish with the resonator cavity (or what appears to be a 'can') as shown here in (Figure 3-11). The resonator cavity is located at the focal point of the parabolic dish antenna. In general, for amy parabolic dish antenna, its focal spot is found [[49]] from the relation,

$$F = \frac{D^2}{16\,h}$$

where, F = focal length (in meters), D = dish diameter in meters, h = dish height in meters. For the parabolic reflector that we used, D = 4.20 meters, and h = 20 inches = 0.55 meters. Thus we would expect its focal point to be found ideally at 1.82 meters. However, it is noted that this focal point although calculated must eventually be found empirically, as it represents the point of maximum gain for the antenna.

In order to optimize the gain of the antenna, one may first ask what could be expected. From [[40]] we have: $Gain = 10log_{10}\,[e_A\left(\frac{3.14*D}{\lambda}\right)^2]$, where λ is in meters, D = antenna diameter, e_A is antenna efficiency (typically 0.6-0.8). Then, assuming e_A= 0.7, with D=4.0, and λ= c/1.66*10^9 = 0.18m, one would calculate the predicted gain to be: Gain, G = 35.3dB.

3.16 Calibration of Radio Astronomy System

The radio astronomy system must be initially calibrated before reliable measurements may be made. Moreover, it is advisable that either intermittent or preferably periodic system measurements be made at regular intervals of time in order to ensure proper system operation.

Initial calibration consists of three main aspects and then a series of final checks based upon astronomical measurements with comparisons to known documented sources which are checked against observations. The main elements of initial system calibration include: static positioning calibration, dynamic positioning calibration (or tracking), and gain calibration.

Before proper calibration may occur, one may ask: How can one be sure that one is

Figure 3-11 The resonator cavity (or can) located at the focal point of the antenna. Photo courtesy of the author.

imaging what he believes he is imaging? In other words: How does one know he is seeing what he thinks he is pointing to, especially when there can be so many potential sources in the sky? The answer to this dilemma lies in the nature of the sources. Naively, one might expect star forming regions to all be exactly alike. Although there are several aspects of star forming regions that are repeatable from one to the next, not all are exactly alike. This is due primarily to specific component elements which constitute the star forming region in question as well the

astrophysical phenomena that are active in the adjacent areas of those regions. Hence, specific imaged regions have their own distinct spectral characteristics that are evident in amplitude and spectra of different molecular species. One example noted here is the source known as Cassiopeia-A. This star forming region is an excellent source to image for calibration both due to its strong amplitude emission as well as its uniquely recognizable emission spectra, as shown

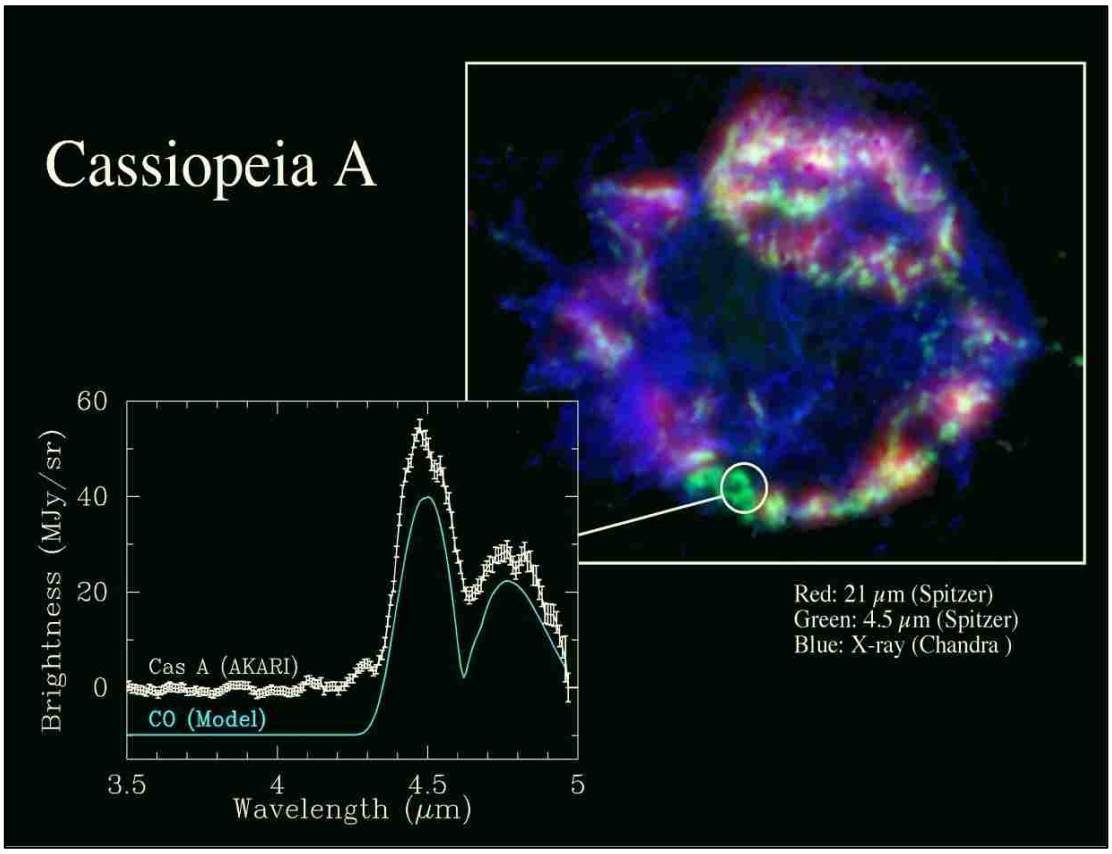

Figure 3-12: Cassiopeia A (Cas A) is a supernova remnant in the constellation Cassiopeia. It is the brightest extrasolar radio source as viewed from earth at frequencies above 1 GHz. As such, it is an excellent source for radio-astronomical calibration. See: www.ir.isas.jaxa.jp - 3513 × 2484

in Figure 3-12 in radio, optical, and x-ray spectrums. Static positioning calibration is achieved using a known source with its position (typically from J2000 data). Given an imaged source such as Cassiopeia-A above, the J 2000 coordinates are employed to seek a rough position of the object. Then the spectrum is displayed on a spectrum analyzer where one may compare the

observed spectra with previously published data. If a match is found, one then manually positions the antenna so as to maximize the spectral amplitude. This point of maximum amplitude is the real position of Cassiopeia-A. This correction in azimuth and elevation should be a first order correction to the J2000 coordinates and one mostly due to the pointing characteristics of the antenna positioner. The positioning controls are then calibrated as per instructions for that position, which causes alignment between published position data corresponding to displayed position.

Dynamic positioning calibration may be considered a check on the ability to maintain static position orientation and control despite the rotation of the earth. When both static positioning and dynamic positioning are maintained with the celestial image, ideally the amplitude of the image should not appreciably vary as the image is maintained within the central portion of the field of view. This is due to the fact that the distance to the object is essentially constant and therefore the amplitude of the signal should be constant as long as tracking is consistently in alignment with the source. In order to accomplish dynamic positioning, the tracking is initiated starting with known J2000 coordinates for an imaged object. Then the amplitude of the object is monitored and used to minimize deviations of observed amplitude during tracking. Once positioned properly for imaging, if the tracking were "perfect" there should be no deviation in amplitude as the earth rotates (assuming the source intensity is constant). Therefore, it is the deviations in amplitude that indicate small tracking errors.

It should be noted that the unique spectra of the image in question is employed as a standard by which one may compare the observed spectra with the expected spectra. Once the system is properly calibrated, it is possible to believe the spectrum of an unknown or poorly studied source.

To calibrate against a matching spectra, one compares the measured value of amplitude with the published value of amplitude. A constant correction factor then enables one to calibrate against the published standard sources which have been studied over time, where the source is known to be stable and to have constant amplitude and therefore be suitable for use in amplitude calibration. We illustrate with an example here for the amplitude of Cassiopeia-A at 1.4 GHz: In Figure 3-12, note particularly the distinctive "Double-Peak" spectrum of Cassiopeia-A.

The published data for Cassiopeia-A is 1600 Jy with the measured raw data for Cassiopeia-A to be 1542 Jy. One then adjusts this measured value to correspond to the published value by use of a correction constant, where Corrected S_{CasA} = 1600 Jy = K*Measured. Therefore K = 1.0379

3.17 System Troubleshooting

Due to system inability to correctly image galactic objects, a process was embarked upon to thoroughly measure the performance of all system level components in order that the source of the error could be correctly identified and corrected. In order to be thorough and sure of one's conclusions a good deal of work was included that is outlined here to be complete. The individual measurements included cable integrity, parabolic focus alignment, and verification of antenna gain.

The 1st generation radio astronomy system employs two 120 foot long cables that relay the output of each LNA to either the SpectraCyber receiver or alternatively to a pair of HP8566B/HP8568B spectrum analyzers for analysis and display. These cables were measured under three conditions: (1) "Open Circuit Termination", where the cables were each connected to an HP1818A Time Domain Reflectometer (TDR) where the far end of the cable was left

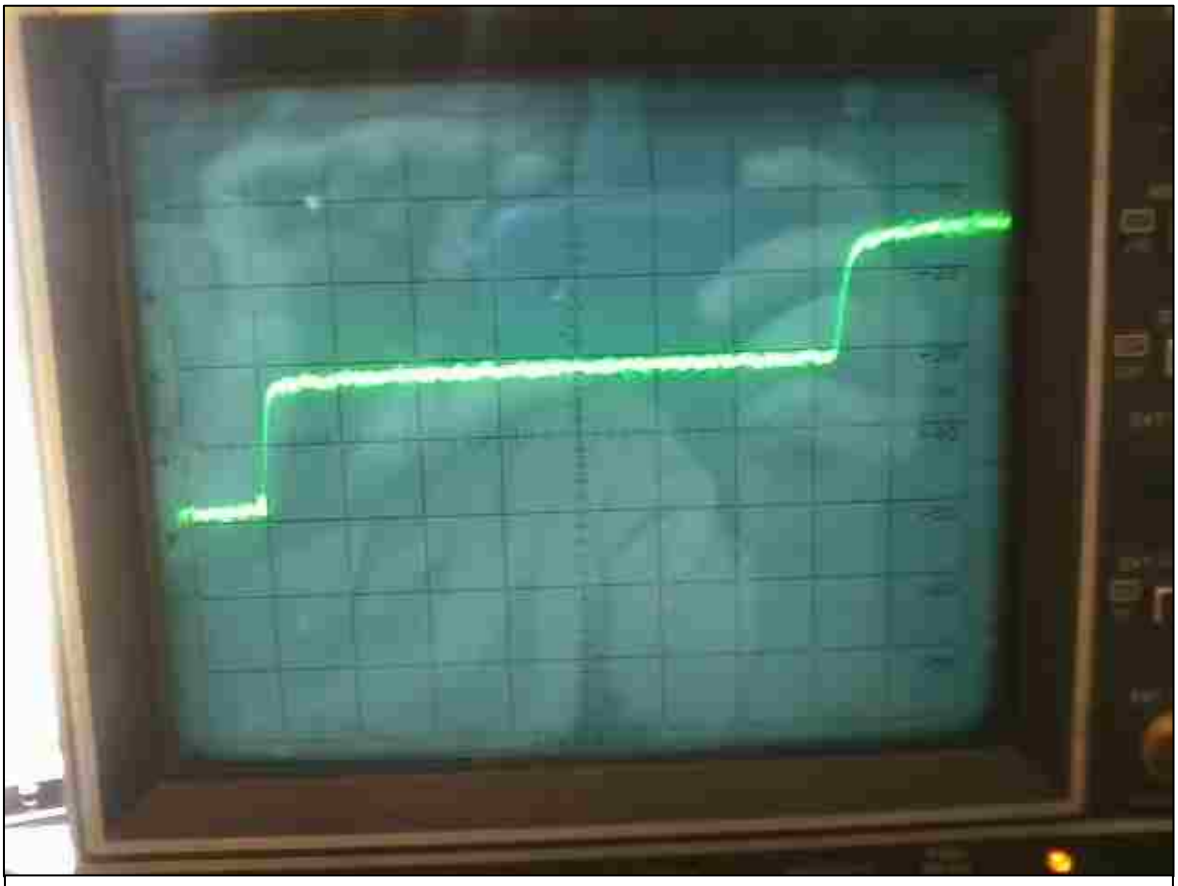

Figure 3-13: This TDR image shows a constant 50 ohm impedance of cable (the center portion) which is open circuit terminated at 120 feet (The upward step toward the right end of the screen). Photo by author.

unconnected, (2) "Short Circuit Termination", where one end of each cable was connected to the HP1818A and the far end of the cable was short circuited, and (3) "50-Ohm Termination", cables. The first of these measurements is shown in Figure 3-13. After making all three sets of measurements upon the two cables, it was found that they are of the expected length (120 feet),

of the expected impedance (50 Ohms), and in all other respects appear to be in proper operating condition, having no apparent impedance discontinuities.

The focus alignment of the parabolic antenna was made using several methods. After calculations and construction and modification of the antenna, the focus was adjusted and checked using a laser pointer, a remote 10 GHz transmitter, and a signal generator for the operating frequencies. Finally, the focus was rechecked using GPS satellite signals. The latter GPS signal was able to be used directly with all electronics of the system in place, including the LNAs. This was possible because of the broad frequency bandwidth of the LNAs, which is 50MHz at its -3 dB points and greater than 100 MHz at its -6 dB points. Since the center frequency of the LNA for OH MASER use is 1667 MHz and the upper frequency of the GPS satellite broadcast is 1575.42 MHz, the frequency difference of 92 MHz was concluded to be well within the -6 dB bandwidth of the OH MASER LNA and thus could be used directly for alignment. A depiction of the GPS satellite orbital constellation around the earth is shown in Figure 3-14, where each GPS satellite circle the globe twice per day at an altitude of 20,200 km and transmits at a power of 25.6 watts. From these figures we may estimate the amount of GPS power from a single satellite as measured on the earth:

The GPS L1 power is listed to be 25.6 Watts. This signal is broadcast from the GPS antenna which has a gain of 13 dBi. This yields a broadcast power of 500 Watts or 27 dBW. Since the GPS satellites orbit at 20,200 km, the path loss of the signal as it propagates to the earth is approximately -182 dB. The signal at the earth is then 27 dBW – 182 dB = -155 dBW. Assuming a 5 dB worst case path loss variation, the signal strength estimated to be available on the earth is -160 dBW = -130 dBm, or 49 mJy (given our 4 m antenna size). Now we add and subtract to find the signal level at the receiver: -130 dBm (for the GPS signal expected upon the

earth) + 35 dB antenna gain (if focused correctly) + 30 dB LNA gain (at its intended operating frequency) – 6 dB (for loss when operating the LNA at 92 MHz away from its intended frequency). These summed together yield a final receiver level of -71 dBm.

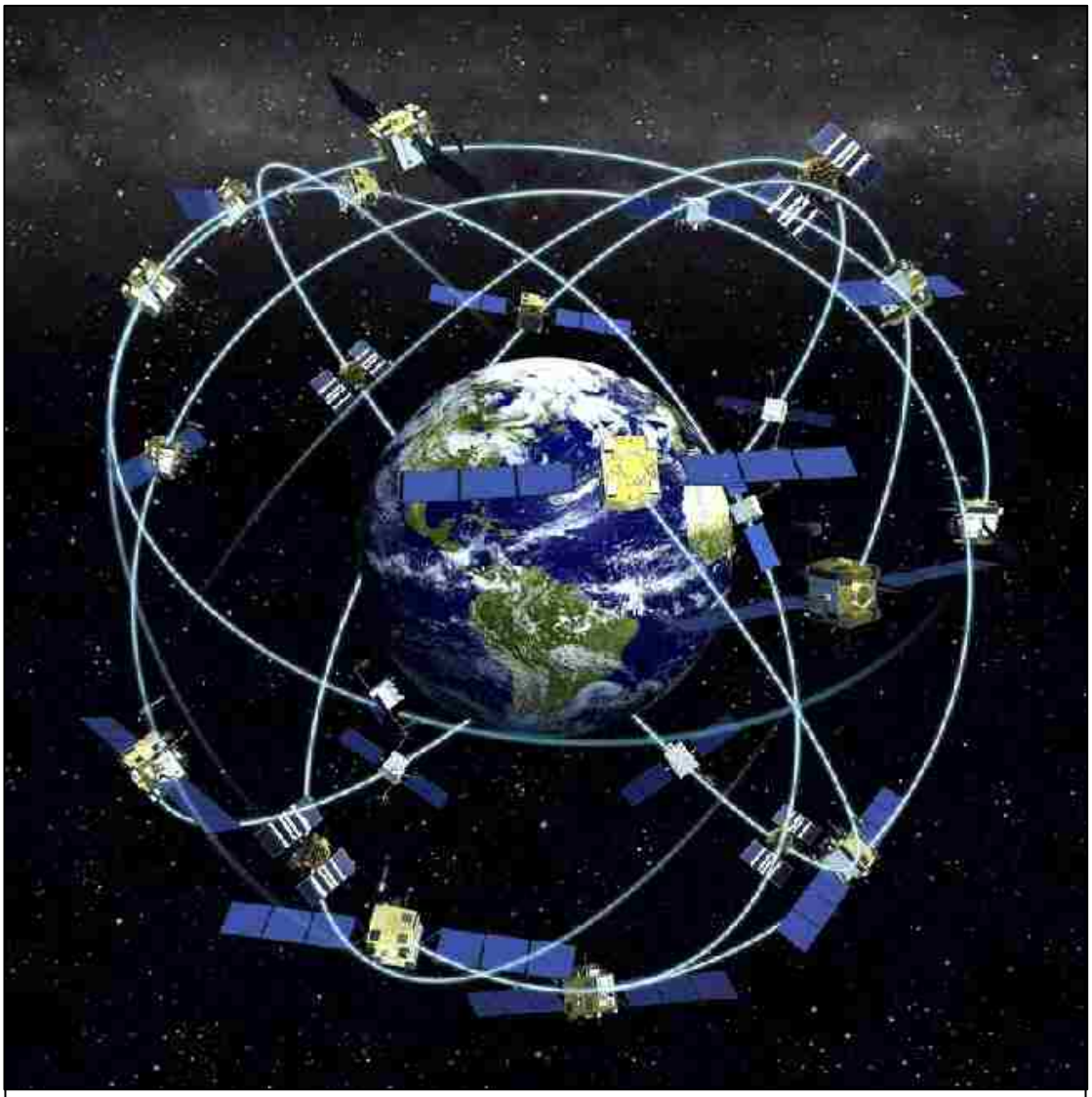

Figure 3-14 GPS Satellite constellation of geocentric orbits. Credit: NASA, and NOAA http://spaceplace.nasa.gov/gps/en

Then to actually measure what was being received in this system, the antenna was pointed at the sky, the spectrum analyzer set for 1557.42 MHz with the span narrowed until the signal was clearly visible. Then the antenna was moved side-to-side and up-to-down to

maximize the signal which indicted we were pointed at a GPS satellite. The measured signal was 72 dB (+/- 1). The fact that the measurement corresponded so closely to the predicted signal strength strongly suggested that the correct focus for the antenna was finally close to being correct.

Although we made many measurements of the antenna, its gain and focus, this single measurement using the GPS signal accomplished both the verification of approximate antenna gain and antenna focus. This is due to the fact that it would not be possible for the predicted value of signal level and the measured value to be so close unless both of these conditions were met. Thus, in one measurement, we have verified focus and antenna gain.

3.18 Operation Verification

There are several measurements and characterizations that are typically performed in order to be assured that all aspects of a radio astronomy system are operating correctly. Moreover, one would like to know what the performance specifications are as well as what percentage error might be expected. For this system at this time, verification is limited to performance verification. This was accomplished by comparison of GPS amplitude measurements (described elsewhere in this document) and transit measurement of solar activity. For the latter, a solar transit was measured using spectrum analyzers in their “Peak Hold” mode (which writes the maximum value found into storage registers. Displayed in these figures are the measurements of a solar transit. All measurements were made using a spectrum analyzer in the Peak Hold mode where integration occurred for several hours.

This proceedure enabled the recording of the radio intensity at this frequency from the sun as it transitioned from an area prior to the focus of the antenna, to one in focus, and then

beyond. Thus, such a recording is used to verify that the antenna is imaging the sun's energy properly. If imaging is occuring correctly, one should see the intensity of the recording begin at a low level and then increase as the sun is in focus and then once again decrease as the sun traverses beyond the focus of the staionary antenna. In Figure 3 15, one observes two waveforms on the spectrum analyzer display. The lower waveform (brighter) is the real-time ,

Figure 3-15 Integrated solar transit at 1296MHz over several hours shows -85dBm amplitude peak with instantaneous level of -95dBm.

instantaneous display after the transit of the sun has occurred. Its value is -95 dBm. The upper trace of the spectrum display (which is dimmer) is a stored display of the recorded peak value of intensity for the spectrum since the stored value occurred from operation the instrument in the 'Peak Hold' mode. One notes that this upper value of -85 dBm, as a peak, has a relatively flat response over its 10 kHz span of spectrum, which seems reasonable as the sun is a broad spectrum emmitter. From this display, one may also conclude the difference in intensity of 10

dBm represents the intensity of the sun as it traversed the antenna focus for this narrow displayed spectrum. The spectrum of Figure 3-16, is for the spectrum analyzer opperated in the Peak Hold mode but with the antenna aimed away from the sun in order to obtain a baseline with witch to compare the previous transit of the sun. The upper trace value of -90 dBm is the value of noise at this frequancy and spectrum width. It is used as a further indicator of solar intensity, before,

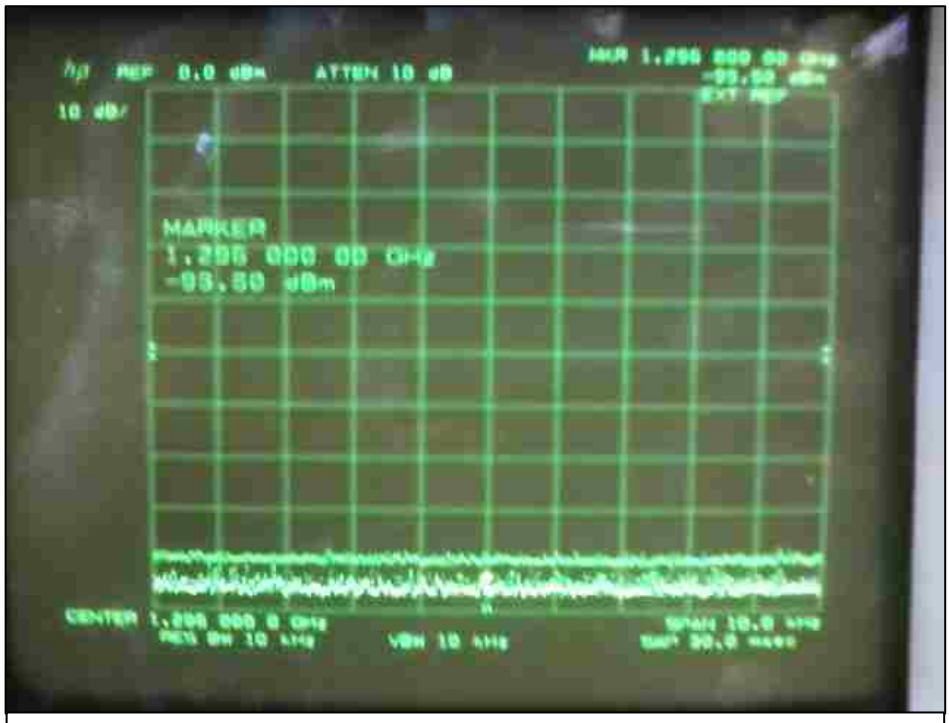

Figure 3-16 Integrated baseline at 1296MHz over hours shows -90dBm.

during, and after transit. On contrasting these spectrums, one may conclude that the fact that one observes a peak intensity during the sun's transit is an indication that the radio astronomy system is operating. However, from this one may not conclude that the system is oppearting at its full capability nor may one totally quantify the system's complete performance characteristics. Further, from these figures, the apparent minimum decernable noise (MDS) becomes apparent as -95 dBm.

Chapter 4. Radio Receiver Systems

4.1 Components of the First Generation System

The first generation radio astronomy system is intended to image HI and OH MASERs in star forming regions within our galaxy. It is based upon use of a singular 4m parabolic antenna (containing two dipole antennas in a resonator assembly at its focal point), two LNAs [Low Noise Amplifiers] (see

Figure 4-3), provisions for Peltier cooling of the LNAs, a triple conversion super-heterodyne receiver, an azimuth and elevation mount and controller for the antenna, and associated computer and software for control and display.

From an instructional standpoint, the purpose of the first generation system is to provide basic hands-on functionality for radio astronomy imaging of galactic star-forming regions, using a single antenna. In addition, it provides a working basis upon which to build a more advanced 2nd generation system which is capable of interferometric measurements.

The 1st generation receiver used is from Radio Astronomy Supplies: The Spectracyber Receiver description with specification as follows: Triple conversion spectrometer covers the rest frequency of hydrogen +/- 2 MHz. Computer controlled display shows both spectral and continuum channels. This design employs - triple conversion (spectrometer) covering the rest frequency of hydrogen +/- 2 MHz using crystal controlled first and second local oscillators and computer controlled synthesized third local oscillator. Receiver noise figure approximately is 1.0 dB. Continuum channel bandwidth is 15 MHz, with final spectral channel bandwidth of 10 KHz.

Figure 4-1: SpectraCyber Receiver for 1420 and 1667 MHz, by Radio Astronomy Supplies.

The standard super-heterodyne design for the standard production model of Spectracyber employs a radio-frequency front-end bandpass filter which precedes all following circuitry. This filter enables the signal to noise sensitivity to be increased by reduction of unwanted wide-band noise that lies out of the range of interest. Following this band-pass filter, is a mixer which is fed by the output of the 1420MHz filter and a 1350MHz oscillator (actual frequency is a sub-multiple of this frequency, but the design uses frequency multipliers to achieve the desired resultant frequency). As the 1420MHz signal is mixed with the 1350 MHz local oscillator, the (non-linear) mixer generates both sum and difference frequencies. The resultant difference frequency is processed through an 70 MHz Intermediate Frequency [IF] amplifier/filter. The output of this first IF filter is presented to another mixer, along with the output from a 70.040 MHz local oscillator. The difference output is then input to a 40 kHz IF amplifier and detector having a sweep-able local oscillator, of almost equivalent frequency as the signal being detected, thus forming a homo-dyne demodulator to result in the signal available at its output.

The model of SpectraCyber that we have employed is slightly modified from the standard production model described above. Since ours is capable of imaging both hydrogen spin-flip (at 1420 MHz) and OH MASERS (at approximately 1.6 MHz), its architecture is quite similar to the standard model, but uses two front end RF stages, one of which is for 1420MHz and the other for 1667 MHz. The output of both RF front ends result in an 70 MHz intermediate frequency stage

Figure 4-2: The Peltier Cooler - used to cool each LNA (by Radio Astronomy Supplies).

with a switch to select which is enabled and which is disabled. The design of the 1667 MHz front end follows very closely to the design described above for the 1420MHz front end, with the exception that the local oscillator frequencies are changed appropriately for the 70 MHz intermediate frequency and instead of just one input filter, the filter is composed of three

Figure 4-3: Radio Astronomy Supplies 1420 MHz LNA, having 33 dB gain at 0.34 dBm noise figure. Design by Tommy Henderson, WD5AGO.

cascaded helical resonators which are tuned to yield the appropriate bandwidth for the desired 1665-1667MHz bandwidth for the OH MASER line spectrums.

The Peltier coolers (see Figure 4-2) provide active heat-pump cooling of each LNA in order to reduce the overall system noise floor by about 3 degrees Kelvin in noise temperature and thereby increase signal sensitivity by increasing the difference between signal to noise. These coolers employ a single-stage Peltier junction array. Typical cooling capability have the potential to achieve up to a temperature difference between ambient and the cooled LNA of up to 74 deg.

4.2 System Reference Function and Gain Stability

Within the overall radio-astronomical system there is a potential for variation of the overall system sensitivity and gain as well as the potential for variation in noise levels. Gain variations often occur corresponding with temperature variations which contribute to overall thermal drift, since all system elements tend to be additive as far as gain and noise is concerned. Gain variations tend to be particularly associated with those system components which are exposed to the weather, such as low-noise amplifiers [LNA], etc. When making astronomical measurements upon different regions, one usually references known measurements performed previously by oneself or others for a reference standard. In order to establish reference intensity, one employs such previous measurements in order to establish correction factors for his/her observations. If there is system gain drift during measurements, these will either tend to cause variations in one's measurements or variations in the correction factors. In either case, such variations are undesirable, leading to measurement errors that may be unacceptable.

One might ask how these errors can be mitigated or at least reduced. The answer is that there is more than one way to address such issues, although there are a few methods in common usage [[50]]. More important is the question of how to address this important issue when faced with both cost and time constraints. The most common methods employ some variation of the so called “Dickie Switch”. In essence, the Dickie Switch alternates the radio-astronomy system's signal between the signal from the antenna and some reference, where the reference is a known and stable standard and the alternation occurs for 50% of the overall time interval. Since half of the signal time is from a known standard, the other half of the time concerning the received signal of interest may be calibrated by this standard. This may occur by using the standard (available half of the time) to establish a consistent output through adjustment of system gain. A

known standard may take one of several forms, such as a 50-ohm termination, a stabilized noise source, or a stabilized signal reference. Finally, the corrected system gain is used to correct for the observed signal amplitude in order to achieve overall gain stability with the overall system element variations.

The method chosen for the first generation of the system does not employ Dickie Switching for automatic gain correction. Instead, it has available reference standards in the form of stabilized signal sources or a stabilized noise source. When observations are made, the input is switched manually to the noise source for standard calibration before or after image amplitude measurements may be relied upon.

4.3 Low Noise Amplifiers

The low-noise amplifiers (or LNA) employed by this system are manufactured by "Radio Astronomy Supplies" [51,52]. The performance of these LNAs include 28dB gain with noise figures of 0.3 dBm, when employed at ambient temperatures. Traditionally, an LNA serves to lower system noise but also may limit system performance other than the antenna. It may be shown that the LNA has the greatest effect on system sensitivity than any other system component [53].

In addition, the system has provisions for active Peltier cooling, to enable the LNA working temperatures to be brought down below ambient and thereby reduce system noise yet further. The design of the LNA amplifier itself is based upon use of a single low-noise amplifier integrated circuit which is biased carefully to maintain its gain across its operating temperature ranges. The same design topology is employed for both LNAs used in H1 (1.4 GHz) and OH

MASER lines (1.6GHz). One may note that although the LNAs are specified for single frequencies, they are in fact broadband, as they have no inherent sharp filters. It may be of value to note that, theoretically, one would expect the use of sharp band-pass filters preceding these LNAs to have the potential to increase system sensitivity due to their decrease in broadband noise before amplification since an amplifier amplifies noise as well as signal, and therefore preceding the amplifier with a filter provides an opportunity to increase signal to noise.

4.4 RF Mixers

Within most radio receivers, mixers are used to non-linearly mix two input signals, often at two different frequencies. The two signals are input to two "ports", with an out or outputs at one or two output ports, When two signals of frequencies f_1 and f_2 are non-linearly mixed, the result of the mixing is that of sum (f_1+f_2) and difference (f_2-f_1) frequencies. The nonlinearities arise from either diode rectifying elements or transistor switching elements employed in the mixer. For instance, when diodes are used as the mixing elements, current and voltage of a diode are related by:

$$I = Is(e^{\frac{qV_D}{nKT}} - 1)$$

This is referred to as the diode equation and describes the current characteristics of a diode, where its charge carriers move by diffusion. Here, I is the diode current, Is is the reverse biased saturation current, V_D is the diode voltage, q is the charge on an electron, n is the ideality factor (and is 1 for silicon), K is Boltzmann's constant, T is the temperature in Kelvin, and $\frac{q}{KT}$ is approximately equal to 39 at room temperature. It is important to note the exponential relationship within the bracket of the diode equation which is nonlinear, since the exponent of

this relation may be expanded as $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.

This may be approximated, for small values of *x* as,

$$\left(e^{\frac{q(V_1 \pm V_2)}{nKT}} - 1\right) \approx \frac{q(V_1 \pm V_2)}{nKT} + \frac{1}{2}\left(e^{\frac{q(V_1 \pm V_2)}{nKT}} - 1\right)^2 + \ldots \approx \frac{q(V_1 V_2)}{nKT}$$

When higher order (x^n , where n>1) terms are filtered out, one sees that mixing of two signals occurs due to the nonlinearities of the active elements (such as diodes or transistors).

One refinement of mixers has been the double-balanced mixer, where common balanced processing of the two input ports inherently enables suppression of the origination inputs, f1 and f2, resulting in only the sum and difference frequencies as the output [54]. Although simple in

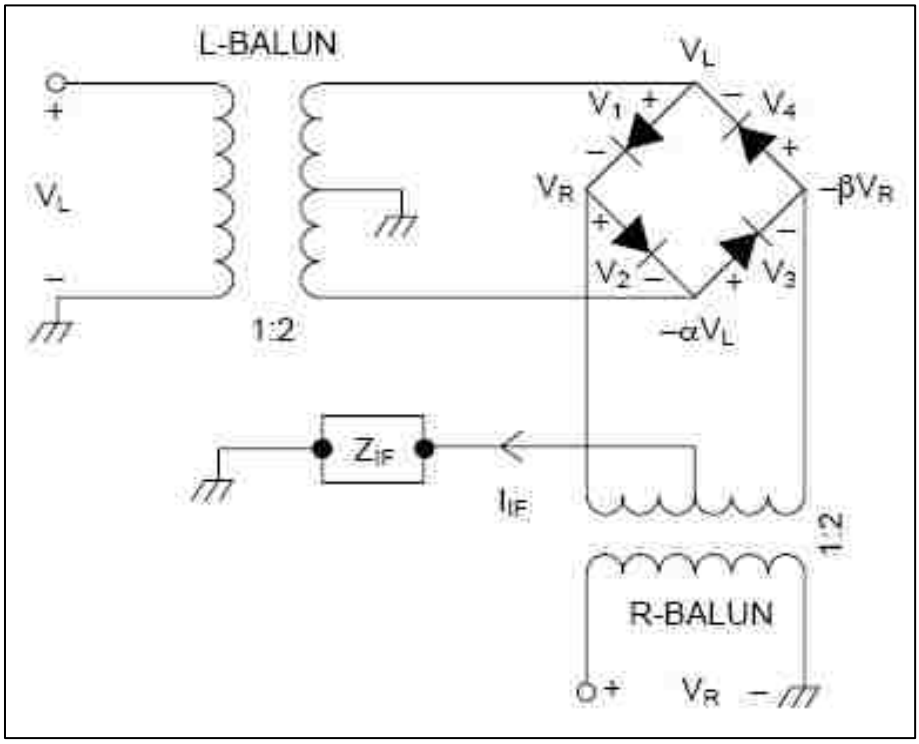

Figure 2.4-4: The Double-Balanced Mixer, implemented as a Shockley diode ring with isolating baluns used to suppress the common input and output

concept, there are several subtleties associated with mixers which are not apparent to the untrained eye. For instance, in Figure 2.4-4 one notes the block "Z_{IF}", implying a constant

impedance load which is presented to the Shockley diode bride. However, providing a constant impedance load that is truly constant over frequency (particularly over a wide range of frequencies) is not always trivial. Providing a non-linear load to the mixer may further exacerbate the resulting sum and difference frequencies due to nonlinearities of the load verses frequency. Such non-linear effects may or may-not be desirable, depending upon the outcome. Another refinement that intentionally emphasizes such nonlinearities is the "Diplexer", which adds a filter to select the frequency to which impedance is applied. A diplexer may provide a separate set of impedances to each sum and difference frequencies as well as the original input port frequencies and is particularly of value when a non-linear filter must be applied to the output, such as a mechanical filter, quartz-crystal resonator, etc. Mixers form the basis of heterodyning and super-heterodyning receivers. Over many years, mixers have been a topic of various refinements, with their application to many common technological appliances, such as cell-phones, televisions, and radios.

4.5 Reciprocal Mixing

In the section above, one learns that a non-linear mixer mixes two signals to result in sum and difference frequencies. Through this process, both the intended and often non-intended aspects of the mixing action take place, with consequential results. Such non-intended results that are referred to here are particularly the noise associated with each source being mixed as illustrated in Figure 4-5 above. Here, we see that a signal often has noise associated with it, which is evident as bands of energy on either side of the carrier shown in the figure. Recalling the non-linear property for mixing, one notes that the mixer does not care which input port has noise and which does not. The mixer simply mixes all components from all its inputs, 'noise and all'. The key idea behind reciprocal mixing is that phase noise of the local oscillator also mixes

with the intended received signal and thus may potentially add noise to it. In a sensitive radio astronomy system, the signals of interest may be 'masked' or hidden by noise originating from the local oscillator, if great care is not applied to first reduce phase noise of the local oscillator. When two signals ("Wanted RF Signal" and the "Local Oscillator") are mixed, the Local Oscillator phase noise also mixes with the wanted signal so that the resulting mixed output contains the intended signal plus phase-noise. Due to reciprocal mixing, the presence of local

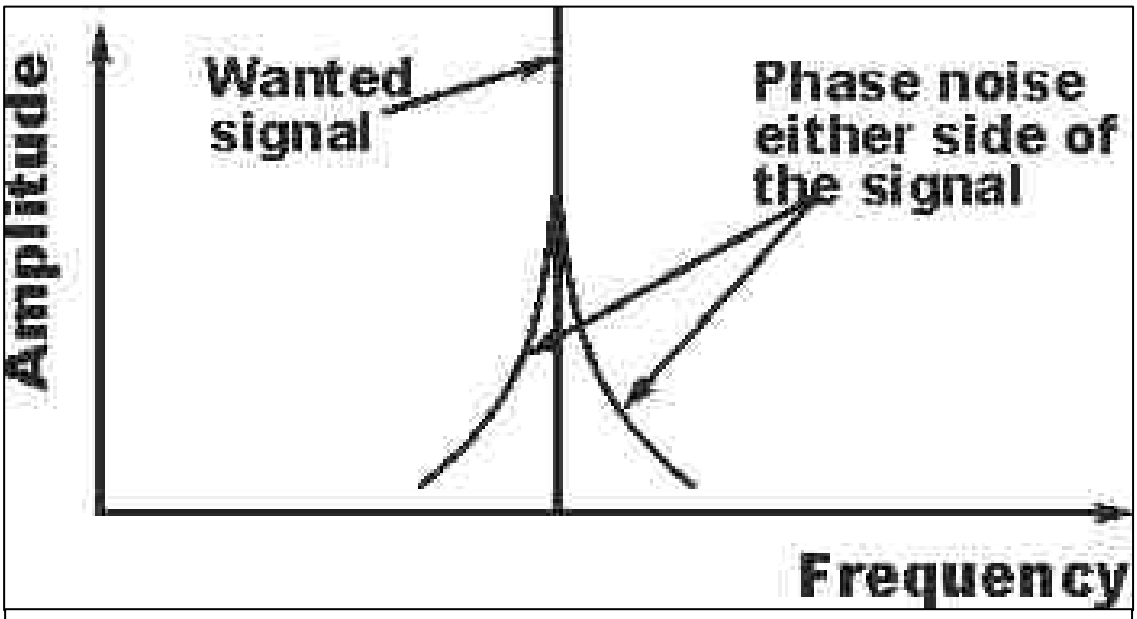

Figure 4-5: Phase noise associated with a signal. From http://www.radio-electronics.com/info/rf-technology-design/phase-noise-jitter/basics-tutorial.php.

oscillator phase noise may has a very distinct and perhaps limiting impact upon receiving properties of the system.

For illustration [[55]], suppose a radio astronomy receiver has a local oscillator having phase noise at 25 kHz offset of -100dBc. What is the receiver sensitivity, (relative to the desired in-band signal), with respect to an interfering signal spaced 25kHz from the carrier, (assuming the IF bandwidth is 15 kHz)? Also, what will be the S/N at the IF from reciprocal mixing if there is an interfering signal of amplitude 50dB above the wanted signal and offset by 25kHz?

The sensitivity relative to the desired signal is: -100dBc/Hz + 10 log10(15kHz) = -58.2dB. Thus an interfering signal with a signal level 50dB above the wanted signal, the noise power produced at the IF will be 50 - 58.2 = -8.2dB relative to the wanted signal, so the S/N is 8.2dB.

The important point intended here is that noise of a local oscillator degrades the sensitivity of a received signal, in the same way that an adjacent interfering received signal degrades the sensitivity of a received signal. Hence, in order to have maximum sensitivity for a receiving system, one must choose or design local oscillators having low-phase noise components, as such concerned care has a very significant effect upon performance.

4.6 Receiver Topologies

Radio astronomy systems are constructed using antennas and electronics, however the electronics may have several options and there are frequently subtle tradeoffs between options which affect performance. So it is prudent for one to at least have a rudimentary introduction to a few such issues.

Technologically speaking, both digital and analog receivers have been built; however, this discussion will focus only on analog receiver technology, while later in this document digital methods are addressed in the discussion of the 2nd generation system. In the analog domain, a receiver is composed of specific functions or blocks. However there are many varieties of how such functions are arranged as well as how they are individually implemented. The functions in an analog receiver are fundamentally composed primarily of amplifiers, filters, mixers, oscillators, and detectors. Various topologies with different arrangements of these functions have been designed and implemented over many years and still continue to evolve.

Common among these topologies are the use of mixers to translate an incoming signal to a lower or higher frequency where selective filtering and finally detection may occur. Common circuit topologies that are based upon mixers in receivers include the heterodyne and the super-heterodyne.

The heterodyne employs a mixer and local oscillator to convert an incoming signal to a higher or lower frequency for selective filtering and detection. The local oscillator may be implemented either in digital technology (such as a phase-locked-loop or digital-direct-synthesis) or analog technology (such as a crystal oscillator, voltage-controlled-oscillator, etc.) or some combination thereof. The super-heterodyne uses a series of heterodyne mixer functions to both achieve greater sensitivity and selectivity, where each mixing and filtering functional block improves noise and undesired image rejection through successively refined filtering.

4.7 1st Gen Block Diagram

The block diagram of the overall 1st generation system is shown in the block diagram of Figure 4-6 which depicts the antenna, LNAs, and SpectraCyber receiver.

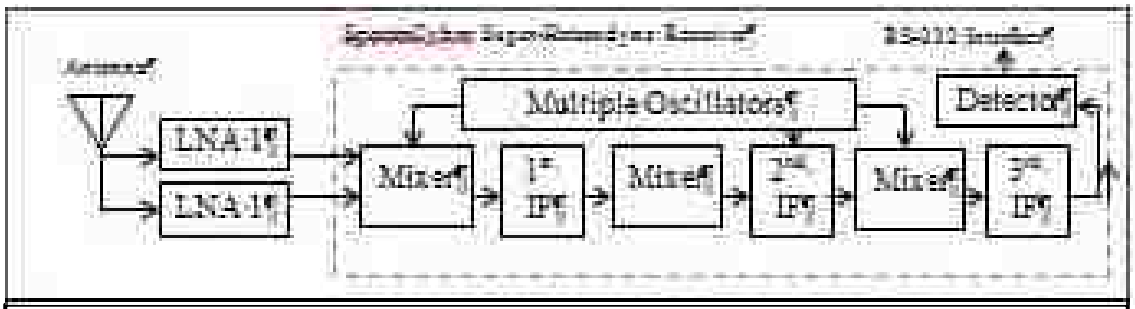

Figure 4-6 This is a block diagram of electronics employed in the 1st generation system.

4.8 The SpectraCyber Radio Astronomy Receiver, a Deeper Look:

Figure 4-7, shows an inside view of the Spectra Cyber receiver, designed and built by Carl Lyster of Radio Astronomy Supplies. In the photo, one notes several circuit boards and associated labels for identification of their respective function. Here, one notes the construction method employed consists of double-sided copper printed circuit board materials, usually associated with good radio-frequency performance as well as lending itself to relatively easy modifications.

Although originally intended for student instruction, due to the relative ease of working

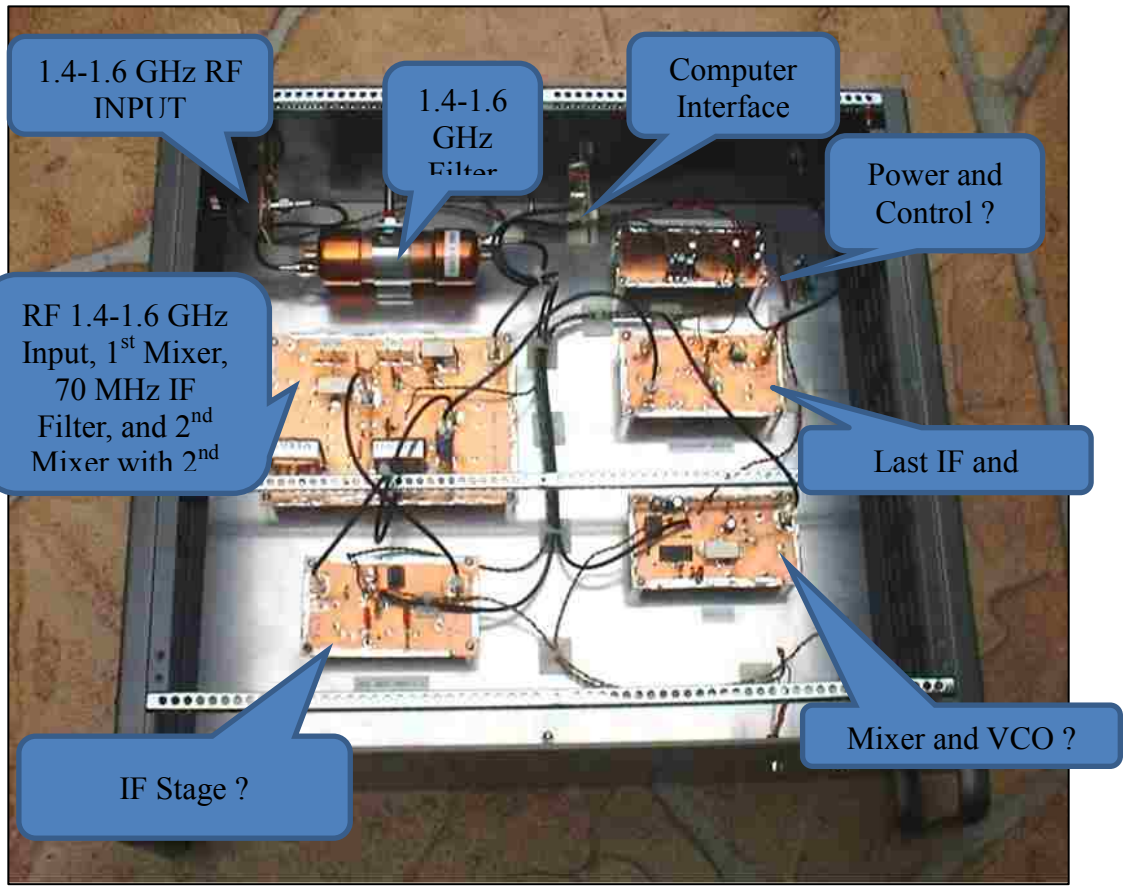

Figure 4-7: Internal view of the SpectraCyber Receiver for H I (1.4 GHz) and OH MASER (1.66 GHz) reception, being of an internal modular construction of the 3rd-order superheterodyne receiver. Photo, courtesy Jeff Lightmann & Carl Lyster, Radio Astronomy Supplies

with such a classical design, there is a potential for a system enhancement through relatively easy modification to the above receiver that would extend its use to advanced observations and associated research in order to enable weak signal imaging, beyond the original design intent of the SpectraCyber above.

I have proposed modification of the SpectraCyber to enable lock-in amplifier use as well as spectral display of observables and have corresponded with the designers of this receiver to check its plausibility and difficulty of its proposal. The designers have verified that my proposed modification is easily accomplished. The proposal consists of modifying the 70 MHz IF in order to insert a 3-dB hybrid RF splitter after the 70 MHz mixer and prior to the 70 MHz filter. The splitter is followed by a 3 dB attenuator to guarantee a matched 50-Ohm impedance for the splitter and more importantly to guarantee a matched relatively consistent load to the mixer. In use, the attenuated 70 MHz output provides a signal input to an external lock-in amplifier having up to approximately 50 MHz bandwidth. This signal may be directly displayed on a spectrum analyzer at its full bandwidth or may be used to display a narrower bandwidth of one's choice. The particular utility of this modification is to extract weak signals from noise, since a lock-in amplifier has the capability to extract signals actually buried under the system noise floor, which promises a significant enhancement over the original design having the potential for enhanced system performance.

4.9 Powering the Radio Astronomy System:

The radio astronomy system consists of both inside-labororatory equipment and that which is outside and thus exposed to the elements. In all cases, the equipment housed inside the

radio-astronomy laboratory is powered through the main power (120 VAC, 60Hz). This equipment also consists of a power supply for powering the outside LNA and associated active Peltier coolers remotely. However, provisions have been made as well for optional battery power for outside equipment. The reason for this battery power is associated with peak system noise performance. Since it is the LNA that dominates the overall system noise and since long lines (over 80 feet of lines) are employed to bring power to the antennas and LNAs, these lines may also serve to contribute to system noise. In this regard, they contribute to both noise pick up as well as coupling of system noise components onto the signals originating at the LNA. Thus in order to explore the limits of low-noise system performance for weaker source imaging, provisions have been made to power the LNAs from batteries which minimize noise pickup sensitivity by totally decoupling the LNA from its power source and associated power lines that are subject to broadband noise pickup.

The battery power is supplied by two 7 Ampere-hour 12 volt gel cells and associated 'Smart Charger' gel cell charger (which is supplied remotely by the internal lab power supply). This gel cell charger has the capability of temperature compensation for charging and is located outside near the antenna. Finally, a remote outside changeover switch is provided for activation/deactivation of this system feature, so one may select LNA power through either the lab supply or the battery supply.

4.10 The Problem with Long (Outside) Remote Power and an Alternative:

Typical of many radio astronomy laboratories, power of equipment is needed for both inside and outside the lab. Outside of the lab power will virtually always be needed for at least

one LNA located at the focal point of the parabolic antenna. In addition, means for cooling (such as use of Peltier junction active cooling) the LNA is often required in order to lower the noise floor and increase effective system sensitivity. Moreover, depending upon the particulars of the laboratory site, long lines leading from the antenna to the laboratory are often required. In the case of our site at the Eyring Science Center, DC power lines of over 80 feet in length connect the LNA (at the focal point of the antenna) to the system receivers located in the laboratory. By nature, both LNA and its associated cooling require current and the long lines have resistance associated with them (resistance per foot) so that there is necessarily a voltage drop between the LNA and the supplied laboratory power source. There are three approaches to address this power loss and voltage drop associated with these long lines, which include:

1. Local remote voltage regulation.
2. Thevenin remote sensing and compensation.
3. Cable/wire voltage drop compensation.

Here is a brief introduction and description of each approach:

In local remote voltage regulation, a voltage regulator is located at the remote outside location. The voltage regulator maintains voltage despite variations in current, etc. but has its own operating voltage loss and must overcome the line/cable loss. Hence, laboratory voltage must be greater than the sum of both in order to supply required power to the LNA and its cooling requirements. Its advantage is its simplicity and its potential to suppress low-frequency noise components while its disadvantage is local heat generation (due to power loss, which adds noise to the system noise floor) and the fact that there is no feedback to the source if the voltage requirements increase or are greater than anticipated. In this latter case, when voltage greater than required is necessary, the regulator “drops out” or falls out of regulation, which can cause

failure of system operation.

In 'Thevenin Remote Sensing and Compensation', the power source supplies power to the load but also uses additional line or cables to remotely sense the voltage at the load (or LNA with cooling). These remote sensed lines have very little current that flow through them and thus may faithfully represent the voltage at the LNA or load to the power source. If the power source is equipped with "Remote Sense", it compensates for the loss in the power supplied through the lines and cables at the remote site. Functioning within the system, such an approach may be very effective. Its disadvantages are few, however if there is ever a loss of remote sense, the compensation is lost and voltage to the LNA can be too low to maintain correct system operation. The main disadvantage of such a situation is that an operator may be totally unaware of such a condition and may wonder why the system does not respond with measurements as it should. Such a condition is likely to be subtle for the non-technically astute operator to easily discern.

The last method to address the problems described above is that of 'Cable/Wire Drop Compensation'. Although a very recent development and presently not one in widespread use, its value is that under relatively constant load conditions found in powering remote LNAs and their associated cooling it has none of the problems described above, but rather holds the potential to be relatively trouble free for remote power use. In operation, this method senses the current required to the load and pre-compensates with a higher or lower voltage in order to maintain a constant voltage at the load. Although a complete product offering for radio astronomy use is unknown to the author at this time, an example of this method is found using a Linear Technology device "LT6110"([56] [57] [58]).

For our 1st generation system, we chose to use the Thevenin remote sensing and

compensation method, as we had available use of a power supply having this capability. In this application, we are wired such that 12V DC power is sent to the remote LNAs and their Peltier coolers with one set of capbles and a 2nd set of caples returns a sense line to the external sense input of the power supply, thus providing for remote sense and control power supply regulation.

Chapter 5. The 2nd Generation System

The first generation system, was based upon a singular radio astronomy system composed of a single parabolic antenna, LNAs, super-heterodyne receiver, and tracking controls for positioning. Unlike the 1st generation system, the 2nd generation system employs interferometric methods to provide imaging from several linked systems for greater sensitivity and resolution. Also, unlike the first generation system, the 2nd generation system uses state-of-the-art digital technology for correlators and zoom spectrometers with much greater capability with a much greater range of imaging capability.

5.1 2nd Gen. Block Diagram

Block diagram of the 2nd generation system, which depicts the antenna, the LNAs, and the CASPER ROACH as its primary system components, as shown in Figure 5-1.

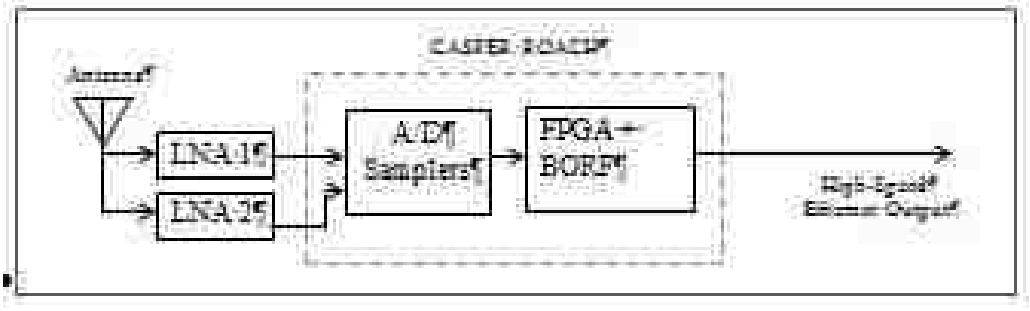

Figure 5-1 Block Diagram of the 2nd generation system

5.2 Baseline Extensions Established and Their Antennas:

A baseline extension is essentially an extension of the original basic radio-astronomy system that has served as the primary objective of this thesis work. Each baseline adds additional receiving capability that adds both sensitivity and accuracy to the original system

through the use of interferometry. This is accomplished by enabling all antennas to slew and track the same object. All signals are then summed per baseline in the frequency domain through the use of correlation methods.

Figure 5-2 Original 4-Meter Antenna donated to Physics and Astronomy Radio Astronomy Research by KBYU Broadcasting. This antenna was later reconditioned, re-purposed, and retrofitted with appropriate HI and OH MASER Imaging LNAs. Additionally a new antenna mount and base was fabricated with tracking hardware installed. Photo, courtesy of Scott Daniel.

A series of baseline extensions were envisioned at the very beginning of this overall radio-astronomy project. Thus, as part of the overall on-going project, work was begun to establish geographically dispersed extended sites. This process began with two parts: identification of the sites and acquiring the necessary permissions for use, and the establishment

of the overall architecture of the system. For the identification of the sites to be considered, several sites were originally included for consideration, with a few of these sites receiving approval before work begun. Original sites have included: West Mountain, Spanish Fork (Figure 5-3: Located at the old BYU Dairy Farm), Springville (Figure 5-4: The Old KBYU Broadcasting

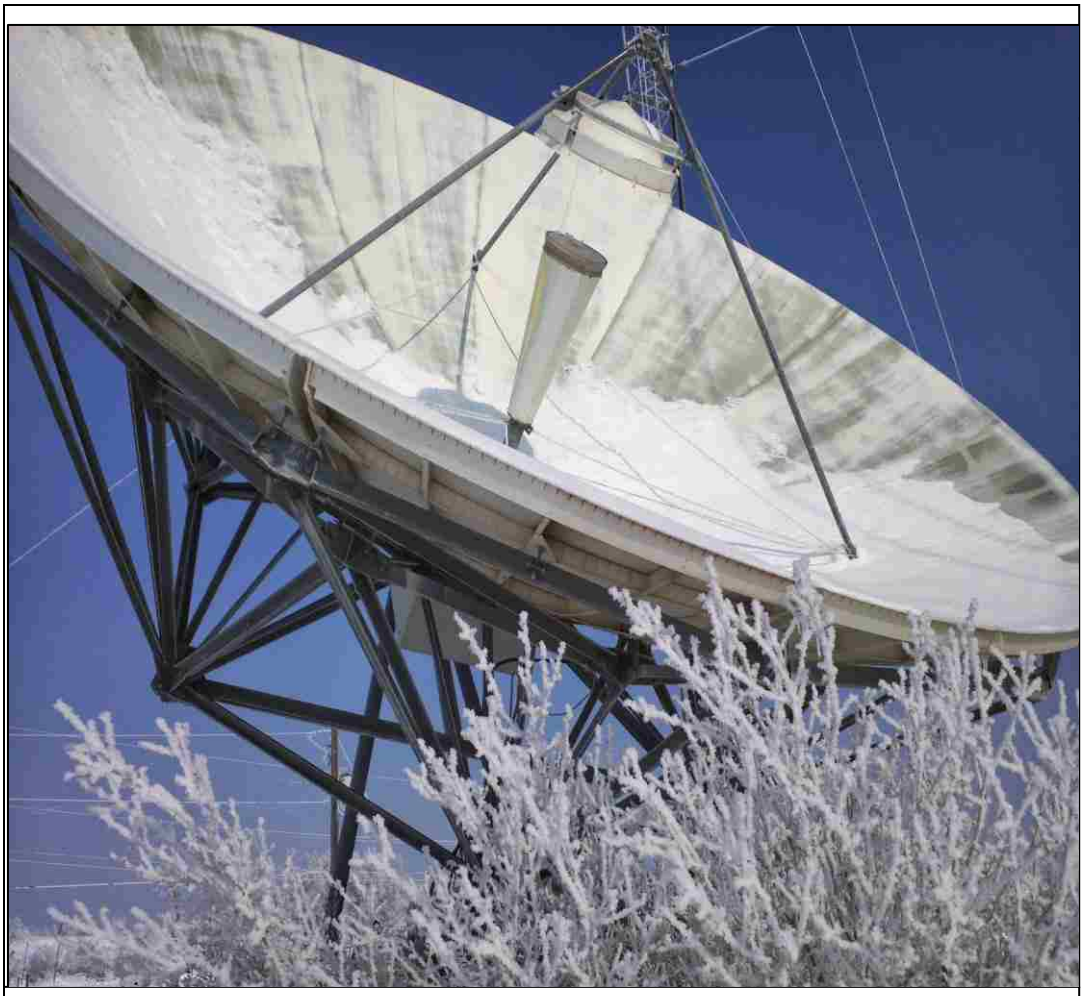

Figure 5-3: The 11 meter antenna located at the BYU Spanish Fork Dairy Farm. Photo, courtesy of Scott Daniel.

Building), The Church Software Development Site, and Mount Vision. Of these potential sites that were under original consideration, permission has already been granted to locate and use the Springville and Spanish Fork sites for baseline extensions. These sites lie in approximately a north-south direction with respect to the original Provo, ESC bldg. Site. For better spatial

resolution of two-dimensional imaging, an east-west extension is also needed. This added dimension may be added either through an extension from one of the sites listed above or by

Figure 5-4: The old KBYU building site, showing the antenna mounting base on the roof in the center of the photo. Photo, courtesy of Scott Daniel.

virtual means via aperture synthesis, which utilizes the rotation of the earth to "synthesize" the needed east-west direction.

5.3 Nyquist Sampling:

The necessity of introducing the Nyquist theorem here is due to the fact that the 2nd generation electronics system uses a digital sampling system to convert the incoming radio frequency signal to a digitally sampled signal. This function, so necessary for advanced correlators and spectrometers, is subject to some constraints, namely the Nyquist limit (unless

one is very careful how one applies this limitation).

When a digitally sampled system is operating intentionally in under-sampling mode (with reference to Nyquist) the system may use aliasing to move signals within some bandwidth x into a lower-frequency portion of the spectrum [59]. Lets take an example [60] for a sample frequency

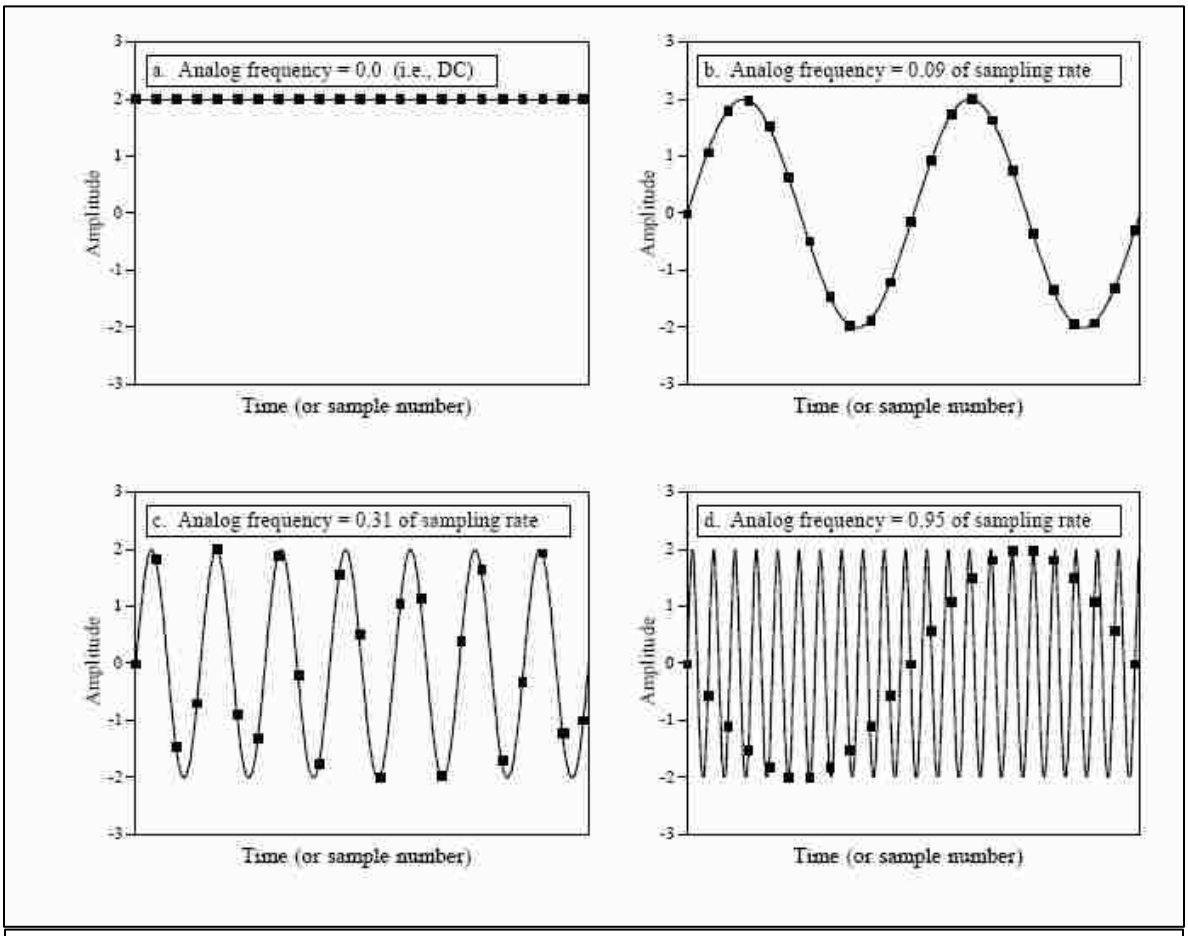

Figure 5-5 In (a), (b), and (c) the sampling of signals is sufficient to guarantee reconstruction of the original signal. However, in (d) the representation of the original signal is aliased (or corrupted) due to the sampling frequency being lower than the Nyquist frequency (one-half of the sampling rate) for the sampled signal. From Lyons, Richard G., "Understanding Digital Signal Processing," 2 ed., Pearson

based on signal bandwidth and the center frequency (f_c) of the bandwidth of interest: Suppose we have a signal at 55MHz with a bandwidth of 8MHz. This produces six upper- and lower-limit sampling frequencies, listed in the table below in Msamples/sec.

In choosing a frequency for sampling, we examine how each upper and lower sampling

frequency affects the aliasing of signals in the bandwidth of 8MHz. The 102-Msamples/sec rate gives an $f_s/2$ value of 51MHz, meaning that aliasing folds the wanted signals along the $f_s/2$ axis to 43-51MHz and aliasing also inverts the order of the frequencies, so those near 59MHz now appear near 43MHz, as shown in the figure below. Alternately, if one were to sample at 59Msamples/sec, the signals would fold along the $f_s/2$ axis at 29.5Hz and appear at 0-8MHz -- also inverted. The diagram below shows results when n equals 4 and 5.

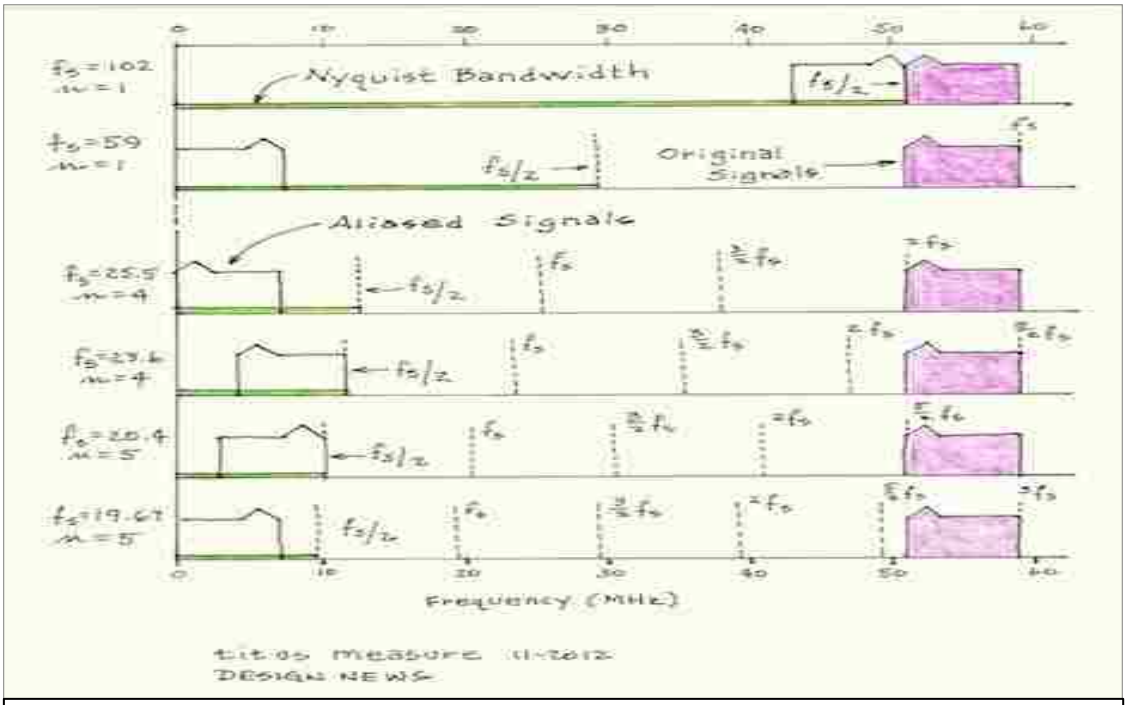

Figure 5-6: Intentional Use of Nyquist Under sampling, where reducing an under sampling rate proportionally reduces the Nyquist bandwidth and folds the resulting signals to lower frequencies between 0 and fs/2 Hz.

After diagramming the aliasing for all six values of n, the following rules emerged. When n equals an odd number, one should use the lower sampling frequency, however one may use signal processing software to invert the frequencies back to their original form. When n equals an even number, one should use the higher sampling frequency, where the signals remain in the proper frequency order.

As n increases, the Nyquist bandwidth, or the 0-to-$f_s/2$ frequency span, decreases and

comes close to the bandwidth of the original signal. For higher values of n, the $f_s/2$ value can encroach on the signal bandwidth and lead to unwanted signal aliasing that will affect measurements.

For this example, one may not create a perfect anti-alias filter that cuts off frequencies below 51MHz and above 59MHz. Some signals and noise outside this range will appear in the aliased signal band. It is advisable to allow space in the Nyquist bandwidth, $f_s/2$, for these signals. In other words, one should try not to choose a sampling frequency that places the Nyquist bandwidth too close to the bandwidth of the signals. As a rule, the bandwidth must not cross an $|f_s/2|$ boundary.

5.4 Frequency and Time References

In order to accomplish the science of radio-astronomy interferometry, one must use multiple antenna systems along with associated electronics, to correlate the signals that are imaged with each antenna. To be able to accomplish this, one needs to correlate them with reference to a given instant in time to which the signals arrive. Assuming the antennas are located in some reasonable proximity to one another, this time difference can be a fraction of a second or a few seconds, depending upon the frequency of observation and the coherence of the signal due to atmospheric variations. This process, which applies a time stamp, needs to be quite accurate, especially when the receive stations are located several miles apart. This is accomplished through use of a time and frequency reference for short term accuracy of at least milliseconds, which is then synchronized to a GPS receiver for long term accuracy. In such long distance interferometric cases such as ours, each location must contain their own time stamp capability. In our system, each site employs its own Rubidium frequency reference, a Stanford

Research Systems FS725 having a low phase noise of –130 dBc/Hz at 10 Hz offset and having a one second Allan variance of less than 2×10^{-11}.

Figure 5-7: Rubidium Time and Frequency Standard, by Stanford Research Systems.

The accuracy and stability of this reference is further stabilized by phase locking the Rubidium reference to a GPS receiver. In our case, this is accomplished through a Trimble Thunderbolt GPS disciplined oscillator, for which GPS-disciplining process is used to enhance accuracy for timescales beyond a minute or so and where the GPS readily overcomes the effects of long-term oscillator aging and frequency drift [61]. The GPS receiver in the disciplined oscillator is capable of simultaneously receiving several GPS satellites in order to generate a greater accuracy average reference value [62]. In general, GPS disciplined oscillators [GPSDO]s are capable of frequency accuracies and stability of one part in 10^9 for even entry-level, low-cost

units but with accuracy up to one part in 10^{13} for the more advanced units such as is used here [63]. As such, GPSDOs comprise one of the highest-accuracy physically-derived reference standards available at this time. As a combined reference GPSDO, although the GPS receiver improves accuracy in the long term, it does nothing for short-term drift and phase noise, where the performance of the Stanford SRS750 rubidium time reference standard "shines" in

Illustration 1: Trimble Thunderbolt Disciplined GPS Receiver. This is used to further improve the long term accuracy of Rubidium clocks. Courtesy of Trimble Engineering. []

performance [64]. We employ the GPS standard as an external 10-MHz reference for our spectrum analyzers, frequency synthesizers, and other applications where it will be multiplied into the GHz neighborhood. This reference's spectral purity is particular value for all measurements made, since it serves as a primary reference where small variations may be multiplied many times over.

5.5 CASPER Usage and Its Range of Applications:

Historically, the electronics associated with radio-astronomy telescopes were designed and built individually and specifically for each site application. In time, there came to be an increasing use of high-technology designs that required a high degree of advanced knowledge and expertise, and was further exaggerated for interferometric radio astronomy or other specialized and special purpose systems. Further, the length of time for development of each system typically required greater than 2 years or more. With such large development times and the associated risk of each, a means of shortening development times was conceived of where the expertise and work of previous development efforts could be re-applied to new sites and needs. Such a method would preserve the understanding gleaned from previous work if its architecture were versatile enough for new applications where it was amenable for re-application by re-programming. This development methodology would hold potential for re-use of major pieces of previously developed efforts and hold potential for shorter development times, easier re-purposing with changing needs, and reduced development costs.

This development model was the basis of CASPER and has been deployed into a wide variety of worldwide sites and systems now using CASPER hardware and software. These designs have included various spectrometers, zoom spectrometers, pulsar imaging engines, beam formers, high-speed data recorders, a wide variety of correlators, and various other dedicated designs.

CASPER [65], stands for Collaboration for Astronomy Signal Processing and Electronics Research. It functions as an international group, having governance and representation from several institutions including: NRAO (Green Bank, NM, et. al.), Swinburne Centre for

Astrophysics and Supercomputing, Australia, Harvard-Smithsonian Center for Astrophysics, NCRA(India), SKA(South Africa), Caltech, and UC Berkeley. CASPER systems are employed at many of the most advanced radio-astronomy systems around the world.

The primary goal of CASPER is to streamline and simplify the design flow of radio astronomy instrumentation by promoting design reuse through the development of platform-independent, open-source hardware and software. In large part, this objective is accomplished by coupling real-time streaming performance of application-specific hardware with design

Figure 5-8: CASPER ROACH Hardware, shown in rack-mount enclosure with high-speed A/Ds.

simplicity of general-purpose software. By providing parameterized, platform-independent "gate ware" libraries that run on reconfigurable, modular hardware building blocks, the complexity of typically encountered systems found in Radio-Astronomy is abstracted away from low-level implementation details in order to enable astronomers, physicists, and engineers to rapidly design and deploy new instruments.

One of the more refined and often used instruments of CASPER is the so called ROACH [66,67], which is named for Reconfigurable Open Architecture Computing Hardware). Its design

is based around a central high speed and high density standalone FPGA (Field Programmable Gate Array) with supporting circuitry to function as a standalone signal processing board. The ROACH was designed by several leading radio-astronomy engineers and physicists from collaborating world-wide radio-astronomy organizations, including KAT (S. Africa), NRAO (Socorro, NM), CASPER (Berkeley), and HKU (MIT).

The ROACH architecture is composed of 3 major and relatively independent subsystems, which operated together where the ROACH board consists primarily of the FPGA subsystem, the PowerPC subsystem, and the monitor/management subsystem.

5.6 .CASPER Firmware Development

The CASPER ROACH has its own self-contained Linux-like operating system, known as BORPH (for "Berkeley Operating system for Re-programmable Hardware") [68]. This system enables one to read and write executable files for the CASPER ROACH hardware. Such files can range from specific functions, such as FFT (Fast Fourier Transform) to complete system functions, such as various types of Correlators.

BORPH is an Operating System designed for FPGA-based reconfigurable computers and is an extended version of Linux that handles FPGA resources as if they were computer files. BORPH operates on the principle whereby a hardware process behaves just like a normal user program except it is a hardware design running on a FPGA. Hardware processes behave like normal software programs. The BORPH kernel provides standard system services, such as file system access, to hardware processes, allowing them to communicate with the rest of the system easily and systematically. Currently, the BORPH operating system is used for the means of manipulating files on the ROACH hardware as well some older designs, such as the BEE.

Under BORPH, hardware and software share the same familiar UNIX interface and the same level of support from the operating system kernel. In BORPH, a hardware process functions the same as a normal UNIX process except that its "program" is an FPGA hardware design instead of a software program. Communications between hardware and software are accomplished through conventional UNIX inter-process communication (IPC) mechanisms, such as shared file, pipe, shared memory, signal, and message passing. Hardware processes have access to system resources as their software counter-parts, such as the general file system, standard input, and standard output.

By maintaining a common hardware and software interface at the kernel level, BORPH provides a system that is language independent for both hardware and software. Software designs can be developed in any language development environment a designer is familiar with. For hardware designs to communicate with the kernel, BORPH defines a standard message passing protocol that resembles a standard operating system software call interface. This standardized protocol enables hardware designs to be developed in any hardware language environment of choice, while reusing common operating system software components.

Although BORPH is based on Linux, it extends the standard Linux kernel to include support for FPGAs. Instead of treating FPGAs as hardware co-processors, BORPH treats FPGAs in the system as normal computational resources. User processes can be either software programs running on processors, or they can be hardware designs running on FPGA's, where a running design on FPGA is referred to as a "hardware process". Here, communicating with a hardware process is no different from communicating with a normal UNIX process. For more details regarding the BORPH operating system see the references [69]. Firmware for the ROACH is developed on a Red Hat Linux Unix Computer and cross-compiled and loaded as

firmware through BORPH. Although it is possible to develop firmware for FPGAs using other operating systems (such as Cent OS), official user support for the required MATLAB and Simulink development software is only offered for Red Hat from Mathworks. Once firmware is debugged and compiled successfully, it is loaded into a ROACH using BORPH, where it is stored non-volatilely and therefore available for reloading with immediate use upon power up.

The firmware used for deployment of CASPER is Graphical User Interface Based, using the Simulink package of the MATLAB software system. It is relatively easy to use, and is based upon graphical interconnection of re-useable functional blocks that are assembled to design a functional system. Each block may range in complexity from simple logical functions such as and/or circuits to oscillators. A block however, may be composed of other hierarchical blocks and thus may become quite complex, depending upon the intended need and the designer's preferences and skills. For instance, a set of commonly used blocks for correlators performs fixed-point fast Fourier transforms with another performing inverse fast Fourier transforms (or FFTs). These FFTs and inverse FFTs are composed of several nested hierarchical blocks in order to finally compute the intended function. The computational capability of such blocks typically far transcends the capability of most computers, and for the example of the FFTs it is capable of 500 million FFTs per second. Normally, this is several orders of magnitude beyond the capability of a conventional personal computer.

Chapter 6. Path Forward and Conclusions

The original intent of the radio astronomy system was to provide hands-on experience for those students wishing to study the regions of star birth in our galaxy, an experience which was rare but now is particularly so due to the use of large and exceedingly expensive nationalized radio telescopes. It was also envisioned that this work would provide the ability to "Give Back" something of substance to the university, to which the author felt indebted for the rare opportunity to attend graduate physics instruction, at a point in life when many are already considering retirement.

In conclusion, I would like to summarize what has been accomplished: We have conceived, built, and established a first generation radio astronomy system for use in imaging strong H I and OH MASER sources within our galaxy. This work has included the building and installation of its 4 meter antenna, the selection and acquisition of its associated LNA and receiver electronics, the design and fabrication of its antenna mast, the installation and calibration of its azimuth and elevation positioning system, the establishment of its computer control, the selection and installation of its 1st rendition of positioning software, and beta tested (working with the software author) a 2nd generation of its positioning software. After a severe gale-force windstorm, we completed repairs and enhancements to the azimuth beating assembly such that it would be strengthened in the event of future storms. Also we calibrated, debugged, and trouble-shot many aspects of its sub-system components for correct operation, including the focal point of the parabolic dish antenna, the calibration of the positioning system, its system gain and more.

Additionally, substantial work has been accomplished in design and establishment of the 2nd generation radio astronomy system. This has included selection and establishment of

multiple locations for interferometric use, donations and conversions of existing satellite antennas for repurposing for radio astronomy purposes, the building of the antenna base for this 2nd antenna, the establishment of the 2nd generation's system architecture, the establishment and operation of the 2nd generation system computational infrastructure consisting of an operation Red Hat Linux computer with appropriate CASPER development software, the establishment of operating rubidium clocks with GPS long-term drift time correction (with the distributions of these accurate clocks to other labs within the physics department), the study and application of CASPER programming methods, yellow blocks, and more).

Although our work included additional efforts, time constraints did not enable us to complete all activities that we had begun.

Concerning the path forward and future efforts, there are two small items that I would recommend, which would include further measurements and refinements of focusing as well as many more measurements of star forming regions with full characterization of its overall system performance. I would also like to better perform calibration constants against known sources. Lastly, for the work on the 2nd generation system that we have begun, there is a substantial amount of work regarding the correlator and zoom spectrometer software development, which is largely comprised of 1st reviewing the programming procedures for CASPER "Yellow-Blocks", particularly for FFT and spectrometer coding, then creating and debugging spectrometer/correlator code, then calibration, and finally comparison of 2nd generation system to the known performance of the 1st generation system.

Appendix A. Established Interferometric Sites

The map below identifies the interferometric sites that were established.

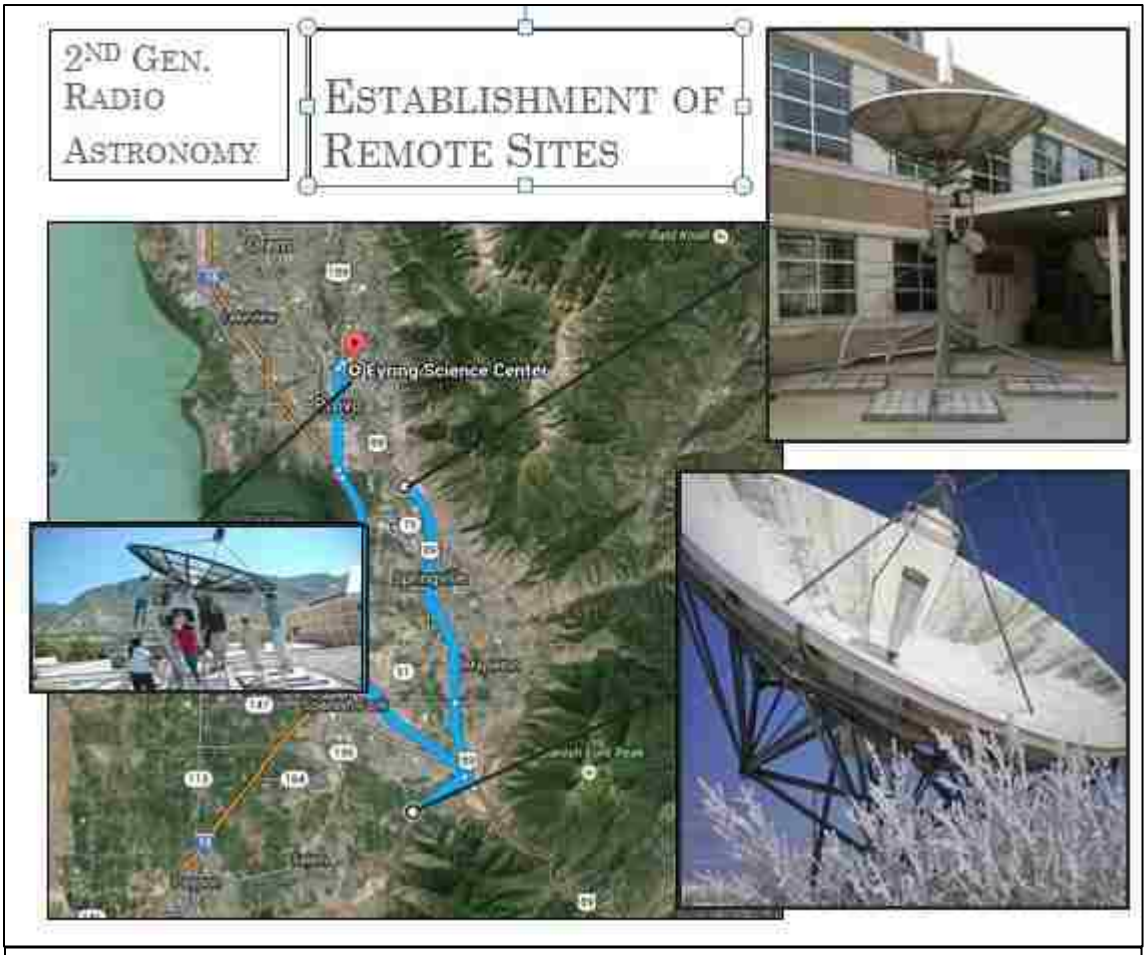

Figure 6-1 Map of interferometric sites established

Appendix B. Advanced Topics

The material in this section relates to aspects of the radio astronomy system that were investigated, used, and/or interacted with at either BYU or Berkeley, but which did not become a part of the final system as implemented in generation-1 embodiments.

Appendix C. Walsh Functions and Applications

Walsh functions, along with Fourier transforms, continue to see much use in the world's leading radio astronomy telescopes. Yet, their background and applications are often generally overlooked and are not frequently taught at the undergraduate or graduate level within the physics or engineering community. Although I was mildly familiar with their use previously, I observed their application into radio-astronomy system development while spending a summer at

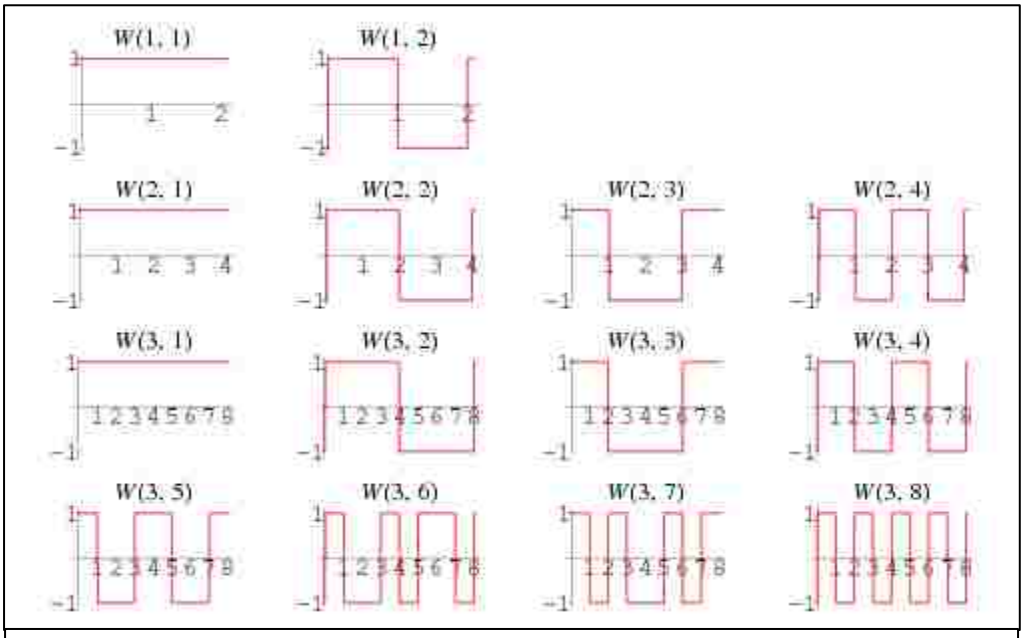

Figure 6-2 Examples of several Walsh functions

UC-Berkeley, where they were required to encode antenna signals for system cross-talk reduction. Their utility is not confined to only radio-astronomy; it is often not employed because it is not well known.

Walsh functions are sets of mutually ortho normal functions. These functions may be used in many applications like that of sine wave functions, such as Fourier transformations. However Walsh functions, unlike sine functions, have only 2 states and thus lend themselves

readily to various digital implementations.

Walsh switching is often used to separate or distinguish one channel from another when there are many antennas situated in relatively close proximity to one another in order to reduce cross-talk and increase performance. In essence, one may switch at both the input and output in synchronization using Walsh functions with frequencies above or below the frequency band of value. On the input, the waveform is summed with a Walsh function having both positive and negative excursions greater than the signal with the net result as zero added or subtracted from the original signal. This summed Walsh signal is subtracted at the final stage, leaving only the original signal with cross-talk often reduced by 40 dB or better. The summation of the Walsh signals with the input signal with back-end subtraction of the same Walsh signal may be readily accomplished.

It is interesting to see some examples of Walsh Switching used at leading radio-astronomy sites worldwide. This is illustrated in the following examples. The GMRT uses Walsh switching for noise calibration since by injecting noise of known power the system temperature can be measured. The noise at any antenna can be switched on and off according to a pre-determined pattern. By synchronously measuring the total power, it is possible to calibrate the system temperature. Additionally, since signals from one antenna could leak into another antenna at various points (referred to as *cross-talk),* this leakage can be minimized by switching the phase of the RF signal for each antenna by an ortho normal pattern with respect to all other antennas. At the correlator the exact reverse phase switching is done for each antenna so that the original phase is recovered just before the cross-correlation is done, greatly reducing the cross-talk at all points between the RF amplifier and the base-band. Typically the ortho normal functions used are Walsh functions, where this scheme is called Walsh switching.

ALMA uses a set of 128 orthogonal Walsh functions [70], both for the sideband separation using 90-degree phase switching at the first local oscillator, and for spurious rejection using 180-degree phase switching at the first local oscillator with demodulation by means of a digital sign change shortly after the digitizer. A set of orthogonal Walsh functions can be defined and ordered in many different ways (PAL, Natural or Paley ordering, WAL, CAL & SAL sequenced ordering) using different phasing definitions (e.g. Harmuth phasing, or "positive phasing".) It was proposed that ALMA standardize on a set of Walsh functions in WAL order with positive phasing since it was found that 180 degree Walsh switching did not work well to eliminate spurious signals in high spectral resolution modes [71]. Also Larry R. D'Addario describes Walsh Methods of implementing fringe tracking, sideband separation, and phase switching in the ALMA telescope [72] where required features make use of direct digital synthesizers for precise phase control of the first and second local oscillators with orthogonal switching waveforms (linear-frequency square waves, rather than Walsh or Rademaker functions).

Appendix D. Tools for Parabolic Antenna Calculations:

There are a few sites having tools for calculations of parabolic antennas as applied to a variety of frequencies. One such site is: "http://www.satcom.co.uk/article.asp?article=22". Here, one will find an on-line calculator for such analysis. It is interesting to examine a few values, as related to the present variety of parabolic dish antennas currently comprising our systems within the physics department:

Table 6-1, Comparison of Antenna Gain vs. Size at Different Frequencies

Location	**Diameter (meters)**	**Gain @1.4 GHz**	**Gain @1.6 GHz**	**Gain @432 MHz**	**Gain @1296 MHz**
ESC-Roof & Springville	4	34.6 dB	35.6 dB	24.2 dB	33.7 dB
Spanish Fork	10	42.6 dB	43.5 dB	32.1 dB	41.7 dB

Note that when applied to EME, the figures above will be double from the values shown, since for moon bounce that assumption is the same antenna gain figures are used for both transmit and receive.

Appendix E. Low Cost Antenna Measurements:

The following link represents one amateur's (Kent Electronics) low-cost kit antenna offerings for relatively high-gain log-period, Yagi, and patch antennas [73] These antennas are suitable for radio-astronomy dish antenna feeds and much more. Standard frequencies available span the spectrum of 300MHz through 15GHz.

Beyond the implicit value and diverse applications for such antennas as referenced above, such wide-band, constant impedance antennas may be useful for large radio-astronomy antenna characterization. For instance, in cases where the suitability of a given parabolic antenna needs to be determined for a particular frequency range and use, its suitability may be determined empirically by its reflection characteristics over frequency. To accomplish this, two such broadband antennas may be employed to qualify the dish antenna's reflective surface over frequency. In practice, one broadband antenna may be placed at the far field of the parabolic antenna and is driven by a programmable swept signal source, while the second broadband antenna is placed at the focal point of the dish antenna with its output connected to a calibrated broadband spectrum analyzer or RF-detector with its output displayed as a voltage level. As the signal generator sweeps a range of frequencies, one then monitors the display for a decrease in signal amplitudes and at end of the received frequency range. Although one may define suitability differently in different cases, a simple definition would occur where a 3-dB drop in power or a 6-dB drop in associated voltage level were detected. Then, for a given antenna, its range of suitable use frequencies occurs between the measured 3-dB power or 6-dB voltage gain points, where antenna gain is lowered to half value. Between such points, the gain remains largely constant, assuming that the signal generator employed uses built-in automatic gain leveling control [ALC] to maintain constant amplitude over frequency.

Appendix F. Low Cost RF Power Measurements

As one constructs a radio-astronomy system, measurements are always necessary to quantify performance as well as to measure the brightness of the objects being imaged. This fact is still evident in the cost-constrained environment of universities and individuals. Although generally, instrumentation for such measurements tends to be costly, there have been several alternative developments, pioneered primarily by radio amateurs. One innovative approach to low cost, high accuracy measurements is illustrated as a poor-man's quantitative measurements [74] where a wide-band radiometer was constructed by use of Linear Technology Log-Linear detectors [75], which are low-cost and accurate over a wide dynamic range.

Appendix G. VSA Extension

As was introduced earlier in the introductory portions of this document, the VSA was originally conceived as a stand-alone system with an intended purpose later oriented towards phased-array antenna design. Over time, research within the electrical engineering department

Figure 6-3: One of the VSA Antennas Located on The Engineering Clyde Building on the BYU Campus, shown with its mounting base, azimuth/elevation controller, counter weights, and base weights. Photo courtesy the BYU Electrical Engineering Research Group of Dr. Brian Jeffs and Dr. Karl Warnick.

has changed to the use of CASPER based ROACH systems. The reader will recall that our 2nd generation radio-astronomy system is designed around a CASPER based distributed interferometric correlator using remote electronics also based upon the CASPER hardware and firmware systems.

Thus since electrical engineering and physics departments have chosen to employ the same type of radio astronomy systems a unique opportunity exists due to the built-in means of high-speed communications between units. It turns out that the CASPER ROACH hardware has

built-in 10/100/1000/10,000 Mbt Ethernet connectivity and moreover, 10GBt Ethernet has been designated as a campus wide inter-building upgrade path by the BYU office of information technology.

Starting approximately three years ago, I initiated an effort to interconnect the Clyde Building with the ESC whereby a 10-GBt Ethernet link would be established in order to share resources. Although the limits to sharing will surely be determined by the imaginations and goals of the users involved, certainly one obvious shared resource could potentially be linking of the Clyde building with antennas and systems in use at the ESC.

Appendix H. Large Antenna Positioner

In order to utilize the Spanish Fork antenna, a positioning system must be designed and built where the az/el-positioning system is not nearly as easy as previous efforts. For this antenna, this aspect becomes its most significant cost. My conversations with various manufacturers of antenna positioners indicates that costs of an az/el-positioner for this antenna may potentially run close to three-quarters of a million dollars, as it must be custom built.

There are perhaps four common varieties of antenna positioning systems that come to mind: These are Fixed, Az/El, Sidereal and Phased Arrays. Given the high cost of a 'Fixed-mount' to Az/El conversion, let us consider the latter two options, one at a time:

The Sidereal positioning system is aligned with the rotation of the earth and is not nearly as versatile as the az/el-positioner. However, the objects being imaged are not moving satellites or fast moving, but instead are at fixed positions in our galaxy, with only slow movement required to compensate for the rotation of the earth. Thus for H I and OH MASERS, a sidereal mount would appear to be adequate.

Phased-Array positioning systems are traditionally expensive and complex. However, our CASPER-based systems could potentially be configured to perform this function with help from those who pioneered this technology, located here at BYU.

Appendix I. Deuterium Spin-Flip

Radio astronomy systems have limitations in their ability to image various objects. Certainly this is true of a small system such as we have built and discussed here. It is illuminating to examine an example of one of our limitations, where we are not able to appropriately image, and why.

Although our present system is designed to image hydrogen spin-flip or H I, one may ask if a similar phenomenon which generates a signal like H I could be imaged in the interstellar medium that arises from the distribution of deuterium rather than hydrogen. This question can be asked, because the changes to the system which are required for imaging deuterium spin-flip are potentially very small and inexpensive.

Like hydrogen, deuterium has been present since the beginning of the universe in relatively fixed proportions. Both were among the original elements of the big-bang. To answer the question regarding the detectability of deuterium, one notes that the difference between H I and D I lies in the presence of a neutron in the nucleus of the deuterium atom. When one does the calculation to find the corresponding 'spin-flip' one finds that the spin-flip frequency of deuterium for these imaging measurements occurs at 327 MHz [76]. However, unlike hydrogen, one finds that the signal for D I is much weaker than H I since deuterium is found in lower abundance (with ratios of D/H less than 4×10^{-5}) [77,78] in the interstellar medium relative to hydrogen. Given this abundance ratio, one would expect that although D I may be used for imaging; it generally will have a much weaker signal. Here one may predict the expected amplitude of D I relative to that found in H I.

Specifically, one may quantitatively relate the expected signal strength to the associated H I line by: $V_{DI\_db} = 20\ log_{10}\left(\frac{V_{DI}}{V_{HI}}\right) = 20\ log_{10}\ (4 \times 10^{-5}) = -88.0\ V_{HI}$

Or, in other words, one would expect that the measurement of deuterium spin-flip would be approximately 88dB smaller in amplitude (at 327MHz) when compared with hydrogen spin-flip (at 1420MHz) in the same region of the interstellar medium. However, one should note that ideally to make this measurement, one would first wish to calibrate the 327 MHz receiver system response with the 1420 MHz receiver system response in order to compensate for variations between the two systems. Such a calibration may be accomplished by means of measurements of the amplitude from a reference oscillator at both frequencies in order to be assured that the amplitudes are equal, as measured by a calibrated spectrum analyzer.

Once calibration is complete, observations could be made. However, since the imaging measurement is so much smaller in intensity than H I, as noted in the above calculation results, one would expect that when given the gain associated with a small antenna, such as is present in this system, it may not be possible to directly image Deuterium. In (Figure 6-4**)** we see the galactic plane imaged at 327 MHz for deuterium spin-flip by the WRST.

Figure 6-4: The galactic Plane imaged at 327 MHz.for Deuterium by the WRST galactic plane survey. See: www.ucalgary.ca/ras/WRSTsurvey.

In this image of D I spin-flip in the galactic plane, one observes that unlike the previous H I imaging of the galactic plane, this distribution of D I appears closer to homogeneous, albeit for the two bright regions closer to the two sides which appear to show a higher concentration. Drawing an implied inference from this image as well as from the significant reduction of

amplitude noted in the calculation of V_{DI_db} , the author concludes that it is highly unlikely that one would be able to successfully image D I spin flip without a significant increase in antenna area as well as significant cooling of the LNA, since the amount of gain sensitivity is so very much beyond that of this system.

Appendix J. Experimental JT65

These sources are provided as an introduction to the field of amateur radio-astronomy. Generally, this area of investigation and experimentation has been associated with efforts by those who hold an innate interest in furthering the science and technology of the subject. Although they are often not funded, from a technological perspective their work is significant and may provide a starting point or even a basis by which to construct systems, follow practice, etc. Some of these may also serve as a technological basis for research, by one purchasing kits or employing their advanced circuits or software. A few of these methods are described briefly here as follows:

An advanced weak signal digital communication method, "JT65", was created by Joe Taylor, a Nobel prize winner in physics. JT65 digital communications mode provides weak signal reception from those communication signals which are deeply "buried" down below noise levels. This quality is exceptional, as most other communication methods require that the signal that is being received be above the noise level. To accomplish this feat, JT65 employs both forward-error correction (using Reed-Salomon encoding) and synchronous demodulation. These techniques and capabilities were originally employed in the advanced and robust communication methods employed by the Voyager deep space satellite which was designed to explore our solar system and beyond [79] since the Voyager was intended to press the limits of our solar system and would be subject to exceedingly high levels of noise with small signal strength. Thus the communication methodology had to be robust sufficiently to enable communication even when noise prevented it.

Historically, Joe Taylor was the winner of the 1993 Nobel Prize in Physics for the discovery of a new type of pulsar. Later, he applied his knowledge to detect the very weak

pulsar emissions to Amateur Radio. Today, thousands of amateur radio individuals around the world use his WSJT (Weak Signal JT), JT65 and WSPR software to propagate and detect radio signals via the ionosphere as well as via EME (Earth-Moon-Earth) without the need for high power and huge antennas. Some have used power as low as 5 Watts to send and receive WSPR signals from all continents, some with signals that are well below the noise level through JT65A unique capability to extract a weak signals hrough the use of Reed-Salomon encoding and synchronous detection.

Reed-Salomon [1,,8081]encoding is a method of "Forward Error Correcting" that has been and continues to flourish in widespread use in many applications. Such applications include magnetic strip bar-code data reading, MP3, CD/DVD data coding, and deep-space communications for space-craft missions. Historically, Reed-Salomon encoding was used in both Voyager-I and II missions for these spacecraft which were designed to travel beyond our solar system. Today, Reed-Solomon codes represent the most frequently used digital error correcting codes in the world [82] as they provide a reliable means of correcting errors in both communications and various devices requiring reliable memory. Their fundamental principles of operation are based on construction of Galois Field Polynomials where a given polynomial may contain corrupted components, yet the original polynomial may be reconstructed within certain limitations through computation of the polynomial according to the Galois field rules. In use, where data has been corrupted, perhaps for the case of scratches on the surface of a DVD or noise-bursts in the case of deep-space communication, the polynomial computations enable full reconstruction of the original data from the corrupted result.

Yet another advanced technology feature found in JT65A communications is the use of "synchronous detection". Synchronous detection uses a local oscillator to mix against the

original signal, resulting in a reduced noise signal when followed by a low-pass filter. Functionally, the combination of local oscillator, followed by a low pass filter, reduces the noise that is amplified and detected.

Appendix K. Other System Applications

Considering that given an investment for an establishment of a small and capable radio astronomy system such as has been described here, one may ask what purposes the system capability may also serve when not actively in use for radio astronomy. Although this is a non-exhaustive, a few purposes are suggested here along with a more detailed example. In its non-scheduled radio astronomy time, the capability for such a system as has been described may be put to many uses other than its original intent. In a educational environment, such uses may serve to encourage interest from a broad variety of students and potentially from many disciplines, which would be enabled by simple changes to the frequencies used and associated electronics.

To begin with, the tracking capability of the antenna coupled with its gain may readily support low-cost tracking and telemetry associated the international space station or Cube Sats, where the latter would typically have low-earth orbits. Changes required to support this function would include LNA frequencies, and potentially addition of a transmitter.

Another alternative use is that of Moon-Bounce. This may have several forms, but in its simplicity is that of transmitting a signal which is broadcast to the moon's surface, reflects, and is bounced back to the earth. Although this may sound rather exotic, various forms of this have been done for years by many, most of whom have had far less capability than this system provides [83]. There are many resources available, a few of which are included in the references [84, 85, 86]. Further, this activity has occurred at the Owens Valley Radio Observatory [87].

As a set of introductory examples of what may consist of moon-bounce capability, let us assume a 350 Watt transmit power with dual sets of Yagi antennas, and in this example we will assume that the radio frequency chosen is 2 meters (or ~144 MHz). Given these constraints, we

may examine a related example of feasibility by assuming that the minimum discernable signal level for our receiver or spectrum analyzer is -95dBm.

Table 6-2, EME Example Summary of Gains and Losses at 2 meters.

EME Feasibility	Gain/Loss
EME Path Loss (at 2meters)	-255 dB
2 YAGI antennas (2 meters @ 16 db each) X 2 [Given two for transmit + two for Receive)	+64 dB
Transmitter Power (2 meters @350W, referenced to 1mW = 0 dBm)	+127 dBm
1 antenna splitter (-3dB)	-3 dB
Calculated receive level	-67 dBm

Summary: This calculated level above is certainly within the current state-of the art technology and would only be enhanced through use of JT65A communications mode. There is much more to this subject matter than what is appropriate to include in this document

Bibliography

---

[1] http://www.christusstatue.com/
[2] http://csep10.phys.utk.edu/astr162/lect/cosmology/age.html
[3] http://wmap.gsfc.nasa.gov/universe/uni_age.html
[4] http://hyperphysics.phy-astr.gsu.edu/hbase/nucene/nucbin.html
[5] http://csep10.phys.utk.edu/astr162/lect/energy/bindingE.html
[6] http://hyperphysics.phy-astr.gsu.edu/hbase/starlog/pop12.html
[7] http://www.roe.ac.uk/~jap/slipher/
[8] arXiv:1108.4864
[9] http://adsabs.harvard.edu/full/1999ASPC..167..237S
[10] http://skyserver.sdss.org/dr1/en/astro/universe/universe.asp
[11] http://www.amnh.org/education/resources/rfl/web/essaybooks/cosmic/p_lemaitre.html
[12] http://www.catholicculture.org/culture/library/view.cfm?recnum=8847
[13] https://www.youtube.com/watch?v=xhLQ_b3bKdI
[14] http://cosmology.carnegiescience.edu/timeline/1929
[15] http://www.spacetelescope.org/about/history/the_man_behind_the_name/
[16] https://www.youtube.com/watch?v=k8fS_W4ZI1A
[17] http://wmap.gsfc.nasa.gov/universe/bb_cosmo_fluct.html
[18] arXiv:0911.1955v2
[19] http://mwmw.gsfc.nasa.gov/mmw_allsky.html
[20] http://www.atlasoftheuniverse.com/galaxy.html
[21] http://astronomy.swin.edu.au/cosmos/N/neutral+hydrogen
[20] Fish; Reid; Argon; Xing-Wu Zheng (2005). "Full-Polarization Observations of OH Masers in Massive Star-Forming Regions: I. Data". *The Astrophysical Journal Supplement Series* **160**: 220–271. arXiv:astro-ph/0505148. Bibcode:2005ApJS..160..220F. doi:10.1086/431669
[23] Spitzer L, Savedoff MP (1950). "The Temperature of Interstellar Matter. III.". *Ap J.* **111**: 593. Bibcode 1950ApJ...111..593S. doi:10.1086/145303
[24] a b Savedoff MP, Greene J (Nov 1955). "Expanding H II region". *Ap J.* **122** (11): 477–87. Bibcode 1955ApJ...122..477S. doi:10.1086/146109
[25]http://www.jca.umbc.edu/~george/html/courses/atomic/h_21cmline.shtml
[26] http://www.cv.nrao.edu/course/astr534/HILine.html
[27] http://www.cv.nrao.edu/course/astr534/HILine.html
[28]http://laserstars.org/history/einstein.html
[29] http://en.wikipedia.org/wiki/Astrophysical_maser
[30] http://ned.ipac.caltech.edu/level5/ESSAYS/Deguchi/deguchi.html
[31] **"**RF Measurements I: Signal Receiving Techniques", CERN, F. Caspers, 16 Jan 2012 , See: http://arxiv.org/abs/1201.3247
[32]http://hyperphysics.phy-astr.gsu.edu/hbase/sound/db.html
[33] Analog Devices : Virtual Design Center : Interactive Design Tools : Utilities : VRMS / dBm / dBu / dBV calculator
[34]"The International System of Units (Si)": http://www.bipm.org/utils/common/pdf/si_brochure_8_en.pdf
[35] http://en.wikipedia.org/wiki/Jansky
[36] http://www.radartutorial.eu/06.antennas/an05.en.html

[37] Straw, R. Dean, Ed. (2000). *The ARRL Antenna Book, 19th Ed.*. USA: American Radio Relay League. pp. 18.14.
[38] "Satellite Systems: Principles and Technologies", by Pattan, Bruno (1993). pp. 267
[39] "Fixed Broadband Wireless System Design", by Anderson, Harry R. (2003), pp. 206–207.
[40] "Satellite Systems Engineering in an IPv6 Environment", by Minoli, Daniel (2009). pp. 78.
[41] http://en.wikipedia.org/wiki/Parabolic_antenna
[42] http://en.wikipedia.org/wiki/Reciprocity_(electromagnetism)
[43] http://en.wikipedia.org/wiki/Parabolic_antenna
[44] http://visual.merriam-webster.com/astronomy/astronomical-observation/celestial-coordinate-system.php
[45] http://astro.unl.edu/naap/motion1/cec_units.html
[46] http://www.nlsa.com/ and http://www.nlsa.com/nfw.html
[47] http://www.arrl.org/shop/Nova-for-Windows
[48] http://www.arrl.org/home
[49] http://www.radio-electronics.com/info/antennas/parabolic/parabolic-reflector-dish-feed-systems.php
[50]"Radiotelescopes 2nd Ed.", Cambridge Monographs on Physics, by W. N. Christensen and J.A. Hogbom. Chapt. 8.2.2 "Noise Fluctuations and Gain Stability", pp. 230-242.
[51] http://www.radioastronomysupplies.com/
[52] http://www.radioastronomysupplies.com/show_detail.php?item_id=11
[53] "RF_Measurements I: Signal Receiving Techniques", CERN, by F. Caspers, See **16** Jan 2012 , See: http://arxiv.org/abs/1201.3247
[54]http://cp.literature.agilent.com/litweb/pdf/genesys200801/elements/system/mixer_double_balanced.htm
[55]http://www.radiolab.com.au/designfile/pnref/rx1.htm
[56] http://www.linear.com/product/LT6110
[57] A PDF specification document may be found at: www.linear.com/docs/43048
[58] A video of its use is found at: www.linear.com/docs/42971
[59] Lyons, Richard G., "Understanding Digital Signal Processing," 2 ed., Pearson Education, Inc., 2004. ISBN: 978-0-13-108989-7.
[60]Jon Titus, , 10/19/2012, Design News, See: http://www.designnews.com/author.asp?section_id=1419&doc_id=251638
[61] http://www.trimble.com/timing/thunderbolt-e.aspx
[62] http://www.ke5fx.com/tbolt.htm
[63] http://tf.nist.gov/general/pdf/2297.pdf
[64] http://www.thinksrs.com/downloads/PDFs/Catalog/FS725c.pdf
[65] https://casper.berkeley.edu/
[66] https://casper.berkeley.edu/wiki/ROACH
[67] https://casper.berkeley.edu/wiki/ROACH_Architecture
[68] http://www.eee.hku.hk/~hso/borph.html
[69] http://www.eee.hku.hk/~hso/Publications/so_codesisss06.pdf
[70]https://science.nrao.edu/facilities/alma/aboutALMA/Technology/ALMA_Memo_Series/alma565/abs565
[71] *https://safe.nrao.edu/wiki/bin/view/Main/ALMALOModulation*
[72] D'Addario, Alma Memo #287, "Fringe Tracking, Sideband Separation, and Phase Switching in the ALMA Telescope"

[73] http://www.wa5vjb.com/links.html
[74] http://www.w1ghz.org/new/portable_powermeter.pdf
[75] http://www.linear.com/parametric/rf_log_detectors
[76] For the derivation of transition frequency, see: "Spin Flip Transition of Hydrogen in Astrophysics", by Crystal Moorman. See: http://www.physics.drexel.edu/~bob/TermPapers/Moorman_QuantumPaper.pdf
[77] "Deuterium Abundance in the Interstellar Gas of the Galactic Anticenter from the 327 MHz Line", The Astrophysical Journal, 630:L41-L44, 2005 September 1,by A.E.E. Rodgers, K. A. Dudevoir, et al. See: http://iopscience.iop.org/1538-4357/630/1/L41/fulltext/19647.text.html
[78]"A High Deuterium Abundance in the Early Universe, *Nature* 385, 137 - 139 (09 January 1997); doi:10.1038/385137a0, by A. Songaila, E. J.Wampler & L. L. Cowie, See : http://www.nature.com/nature/journal/v385/n6612/abs/385137a0.html
[79] http://www.qrz.com/db/K1HTV
[80] S.B. Wicker and V.K. Bhargava, "An Introduction to Reed-Solomon Codes" in Reed-Solomon Codes and Their Applications, eds. S.B. Wicker and V.K. Bhargava. New York: IEEE Press, 1994, pp.1-15
[81]W. Cary Hu man; Vera Pless, "Fundamentals of Error-Correcting Codes", Cambridge University Press; Cambridge, UK 1st Edition, 2010
[82] See: "Introduction to Reed-Solomon Error Correcting Codes", by Yo Sup (Joseito) Moon, Pub. By Harvard University, Department of Mathematics, see: http://isites.harvard.edu/fs/docs/icb.topic982877.files/Moon-%20Introduction%20to%20Reed%20Solomon%20Codes.pdf
[83] http://www.arrl.org/weak-signal-vhf-dx-meteor-scatter-eme-moonbounce
[84] http://www.downeastmicrowave.com/
[85] http://www.advancedreceiver.com/
[86] http://www.minicircuits.com/homepage/homepage.html
[87] http://www.kl7uw.com/emeOVROcalc.htm