A Phase Measurement Radio Positioning System

for Indoor Use

by

Matthew S. Reynolds

Submitted to the Department of Electrical Engineering and Computer Science

in Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Electrical Engineering and Computer Science

at the Massachusetts Institute of Technology

February 3, 1999

Author____

Department of Electrical Engineering and Computer Science
December 21, 1998

Certified By____

Neil A. Gershenfeld
Thesis Supervisor

Accepted By____

Arthur C. Smith
Chairman, Department Committee on Graduate Theses

A Phase Measurement Radio Positioning System
for Indoor Use
by
Matthew S. Reynolds

Submitted to the
Department of Electrical Engineering and Computer Science

February 3, 1999

In Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Electrical Engineering and Computer Science

ABSTRACT

I present here the detailed design for a low frequency radio navigation system called the Building Positioning System (BPS). This system is designed to work indoors, where the microwave radio signals of the Global Positioning System (GPS) cannot be received. Potential applications for this system range from asset tracking, security, and human-computer interface, to robot navigation and the management of services as diverse as medical care and postal delivery. I first present the issues surrounding its conceptual design and then describe in detail the component level implementation of the prototype BPS system, which I have designed and built but not yet tested as a whole. I also discuss the test procedures which will accompany the deployment of that prototype system in its initial configuration in its test building, as well as further research that is required to make the system scalable and manufacturable.

Thesis Supervisor: Neil Gershenfeld
Title: Assistant Professor of Media Arts and Sciences

Acknowledgements

I wish to thank the many people who have treated me with kindness and generosity over the past 23 years, especially my mother and father, and brother. Their contageous love of knowledge, family, friends, and our world and its people continuously renews my respect for them. They have created in me impossible dreams and unachievable aspirations which will guide me forward forever.

I must also thank my friends of all species, especially Marisa and Papua, for their trust, friendship, love, and understanding.

This document, the thoughts which it contains, and the person I am, are shaped in no small part by Neil Gershenfeld, Joe Paradiso, Susan Bottari, my wonderful labmates, and the many people who toil under, around, and beyond the vision of Nicholas Negroponte and MIT to create what constitutes for me a nearly perfect working, living, and learning environment.

Rehmi Post contributed his time, talents, and wit to the debugging of the CPU module and its software development environment.

This work is supported by the Things That Think Consortium of the MIT Media Lab. The author is the recipient of a Motorola Fellowship and wishes to thank the many Motorolans who have been so exceptionally friendly and generous with their time, knowledge, and resources.

Table of Contents

Chapter 1: Introduction

1-1: Motivation

The location of people and objects relative to their environment is a crucial piece of information for asset tracking, security, and human-computer interface (HCI) applications. These applications may be as simple as tracking the location of a valuable shipping carton or detecting the theft of a laptop computer, or as complex as helping someone to find his or her way around an unfamiliar building. Other applications include the optimization of services such as medical care or postal delivery. Current technologies, such as GPS or differential GPS, can help to solve this problem as long as the people or objects to be tracked are outdoors, where the signals from the 24 orbiting GPS satellites may be received, but there is a latent demand for a similar system that works indoors, where the physics of radio propagation rules out the reception of GPS's weak microwave signals.

This document describes the design and partial implementation of a system which is called the Building Positioning System, or BPS. This system determines the position of a small portable receiving unit by measuring the phase of several radio signals sent from fixed positions in a building. The system uses much lower radio frequencies than GPS, so its signals propagate with relatively little attenuation in a typical home or office building. It requires only four fixed transmitting antennas, placed in the corners of the building, to provide useful coverage over a very wide area inside the building.

At the time the transmitter, receiver, power amplifier, and antennas have been designed and tested individually. A great deal of the real-time signal processing software has been written and tested. The remaining work consists of testing and deploying the prototype system and repairing the inevitable bugs in hardware and software, which will hopefully result in a functional protoype BPS unit.

1-2: LORAN Revisited

This system is similar in principle to the LORAN-A system of radio navigation, which was developed at MIT's Radiation Laboratory in the 1940s. At that time, the complex signal processing techniques required for a spread spectrum system like GPS did not exist; the few computers that did exist were incapable of the real-time computations required to decode GPS signals. Also, the lack of satellite technology prohibited the use of a space-based navigation system. So the engineers of the time turned to a network of very powerful fixed transmitters located strategically along the Atlantic and Pacific coasts of the US. These transmitters operated at the relatively low frequency of 2.6 MHz. Because of the phenomenon of ionospheric reflection, these signals could propagate great distances at night, permitting their reception by ships at sea. The navigation officer of a LORAN-equipped ship would use a special receiver equipped with a cathode-ray oscilloscope to graphically determine the relative phases of pairs of transmitted signals. He would then refer to a printed map upon which he would overlay a transparency plot of the phase measurements he had made. This process would yield a position solution that was generally accurate to within 10Km.

1-3: GPS Technology

Signal processing, computation technology, and space science have advanced considerably over the past 50 years. GPS depends on the use of orbiting space-based atomic clocks, spread spectrum microwave radios, high speed digital signal processing (DSP), and sophisticated mathematics including algorithms that compensate for general relativity. [GPS95] The initial design and deployment of GPS for military use by the Department of Defense in the 1980s has led to the availability of inexpensive commercial GPS receivers in the 1990s. In 1998, a consumer GPS receiver fits in the palm of a hand and costs its manufacturer less than fifty dollars to make. This receiver is capable of reporting the user's position to within 50m RMS accuracy at any point on the Earth's surface.

However, GPS suffers from a fundamental limitation: it cannot be used indoors. The microwave signal from the GPS satellites is extremely weak by the time it arrives at the Earth's surface, and the presence of the leaves of a tree or the roof of a building in the GPS signal path reduce the signal strength to imperceptible levels. Even the use of sophisticated cryogenically cooled receiver electronics cannot recover GPS signals when deep inside a typical reinforced concrete building; the imperceptibly weak GPS signals are overwhelmed by interference from other electronic equipment or more fundamentally by the blackbody radiation of the building itself. The problem is roughly equivalent to trying to see and decode morse code sent by the beam of a flashlight on Earth-- from the surface of Mars.

1-4: How Not To Accomplish Indoor Navigation

Several potential solutions to the problem of indoor navigation have been proposed, ranging in complexity from placing an IR-light emitting beacon equipped with a unique ID in each room, to the use of ultrasonic beacons, to giving the user a computer vision system to identify her place in the indoor world. The first two of these systems have the disadvantage of requiring the deployment of a large number of beacon stations; their setup and maintanance would be prohibitive for more than a few rooms. They also have a problem of high granularity in that one knows only that she is in a particular room, not where she is within that room. The third system, based on computer vision, suffers from the typical problems of computer vision, including great complexity, the need for a clear lines of sight around the room, and slow performance even on very fast computing hardware. Vision based solutions are also generally prohibitive for portable systems for reasons of cost, size, and power consumption.

1-5: Spread-Spectrum Microwave Systems

Recently, several small startup companies have begun to investigate indoor radio navigation, primarily for the location of large portable objects, like hospital beds containing critically ill patients or very expensive packages. They intend to leverage the proliferation of cellular telephones and wireless networking equipment by adapting the components used for those purposes to the problem of navigation. These companies have developed prototype systems which use direct sequence spread spectrum (DSSS) techniques in the 2.4 GHz industrial, scientific and medical (ISM) band set aside by the FCC for such purposes. A good description of how DSSS may be used for radio ranging may be found in [Dix94]. While it is a step in the right direction this approach has several fundamental limitations. First, the 2.4 GHz microwave signal is strongly attenuated by building materials, so a complete set of fixed station units must be replicated in each room to be served by the system. Second, the many fixed units must be connected by a data network of some kind, leading to wiring, antenna mounting, and network management issues. Third, the FCC's bandwidth limitations for the 2.4 GHz band, along with the strong multipath interference caused by even small conductive objects in the room, limit the precision with which one may make DSSS ranging calculations. These companies claim an RMS error of about 3-5 meters. Finally, even though microwave electronics have become much cheaper, easier to use, and less power hungry, these issues still render DSSS a less than adequate solution for many uses, especially those relating to HCI applications.

1-6: Motion Tracking with Low Frequency Magnetics

For HCI applications, specifically the capture of human motion for the purpose of animating computer generated characters, some prior work has been done using the amplitude measurement of very low frequency magnetic fields. These systems consist of a single transmitter, an antenna made from two orthogonal coils, and a large number of small receiving antennas which are placed on the clothing of the person whose motion is to be recorded. These receiving antennas have little onboard circuitry; all of the signal processing is done in a large box containing the transmitter. The attached computer system can then keep track of the motion of as many as 50 points on the person's body.

The person being tracked is therefore wired with a rather large and unwieldy umbilical tether which connects her to a big metal box. The maximum range of this system depends on the size and placement of the transmitting antenna, but typical ranges of 15m are claimed with a measurement precision of 5-7 cm. The system claims to provide positon updates at rates of 30-50 measurements per second. While this performance is quite sufficient for the system's intended use in motion capture from a performer on a stage, it is not sufficient for many other applications in which the performer must be free to move over a wide distance. Also, this system, because all signal processing is done centrally at the transmitter, is not suited to spatially tiling with other such systems to cover a wide operational space.

Chapter 2: Design Goals for BPS

2-1: Overview

The goals of this development effort are to produce a radio navigation system that can cover large areas of space within a building, even though that space may be divided by walls, floors, or ceilings. This system should require a minimum of fixed infrastructure, including antennas, system modules, and interconnecting wiring. The transmitting equipment must be biologically safe and must consume a minimum of bandwidth. As this is an experimental device it need not be FCC compliant, but it should make sensible use of the RF spectrum so that minor redesign of it or the FCC rules may permit its use. The portable receiver should be small, light, reasonably power efficient, and as low in cost as possible. The system should support a virtually unlimited number of such receivers. As the system may be used to track people as well as objects, the system should be capable of being used anonymously, so that the position of a receiver is not known by any central device unless the tagged person releases that information.

2-2: Broadcast Data and Mutual Interference

The system should have a minimal data transfer capability for broadcast messages to the receivers for the purpose of advising the receiver's software or the end user about the health of the system, or for broadcasting a GPS-derived system position reference so that the receiving units may calculate their position in world coordinates as well as system coordinates. Many systems should coexist in adjacent spaces without undue mutual interference. Finally, the position output should be available via a small display on the receiver or by a RS-232 connection to a laptop computer or logging device.

2-3: Quality of Service

The receiving device should report a position solution that has a good availability of part-per-thousand class position solutions. This means that if the coverage area were 50m, the positional accuracy should be 5cm. The positional accuracy should have no functional dependency on the distance from the transmitters to the receiver, but it may vary due to multipath and other artifacts of the radio propagation process.

Chapter 3: Regulatory Issues

3-1: Overview

The designer of a commercial product which uses radio frequencies must take into account such issues as compliance with existing FCC regulations [FCC96] which limit transmitted frequencies, bandwidth, and power levels, as well as more recent regulations which are intended to protect people from excessive exposure to electromagnetic radiation. These issues are among the most nebulous faced by the designers of devices which use radio signals because even the simplest of devices must comply with a large number of regulations. Before it can be sold, each device must be tested to ensure that it is in compliance with all applicable regulations. This process is described in [FCC86]. This is a difficult task, because exactly what the regulations mean is guesswork even for an experienced RF engineer. Fortunately, as the present system is used in a research environment, these concerns may be temporarily ignored for the purpose of demonstrating a working system, as long as the system does not harm existing radio services or cause harm to nearby people.

However, it is important to consider the commercial potential that exists for an indoor radio positioning system. It is the author’s belief that the principles employed in the design of this system are compatible with existing FCC regulations as described in [FCC96]; a similar system could be constructed that may be sold for use under the FCC’s allocations for Industrial, Scientific, and Medical (ISM) devices at 6.78MHz or 13.56MHz. Alternatively, the FCC has set aside spectrum near 2.0MHz for the use of radio navigation devices, although it would be unusual for this spectrum to be used for a nongovernmental purpose.

In the case of the prototype device, the chosen operating frequency of 1.9MHz is within the spectrum allocated to the Amateur Radio Service, and as such this system operates as an experimental beacon device under the author’s Amateur Radio license. Its signal bandwidth of less than 14KHz and its transmitted power of 10W-50W are compatible with the applicable regulations for Amateur Radio stations.

3-2: Certification under Part 15 of the FCC Rules

We will briefly describe how BPS may be redesigned to fit into Part 15.223 of the FCC’s regulations, Title 47, governing unlicensed electronic devices operating in ISM bands. Certification under Part 15 rules occurs after emitted field strength is measured in an FCC approved laboratory’s anechoic chamber at a wide variety of frequencies from 9KHz to 10GHz. All unlicensed devices must emit less than the maximum allowable field strength at every frequency to be certified under Part 15. For these purposes, field strength is the peak electric field intensity measured in microvolts per meter, with a specially calibrated antenna located at a specified distance from the emitting device.

3-3: Field Strength Estimation

At our BPS carrier frequency of 1.9MHz, Part 15.223 specifies a maximum field strength of 100 microvolts per meter, measured at 30 meters distance. In the 13.56MHz ISM band we may emit a field strength of 10000 microvolts per meter at 30 meters. We will estimate the emitted field strength of a BPS transmitter by using an expression relating transmitted power to field strength. This is really only an approximation because it does not take into account such factors as antenna sidelobes, peak versus average field intensity, or power radiated at harmonic frequencies.

$$\frac{P\eta}{4\pi D^2} = \frac{E^2}{120\pi}$$

Eq. 3-1: Field Strength

In this expression, P is the transmitter's output power, η is the antenna efficiency derived in Chapter 7, D is the distance from source to receiver, and E is the strength of the electric field. For those unfamiliar with electromagnetic waves, we note that 120π is the characteristic impedance of free space in ohms. This expression simply restates Ohm's Law. Substituting the BPS parameters discussed in Chapter 7, our system has an estimated field strength of 400 microvolts per meter at 30 meters when operating at 20W output. This system would have to be redesigned to be certified at 1.9MHz but would pass easily at 13.56MHz.

Chapter 4: Signals and Noise

4-1: Overview

The use of a radio navigation system indoors presents several challenges to the designer of that system. Among these are signal attenuation by building materials, people, and objects, self-interference due to multipath propagation effects, and phase distortions caused by the presence of large amounts of metal in the building. These issues require serious consideration because they may affect the range of the transmitted position signals or the accuracy of the position solution.

4-2: Frequency Dependent Signal Attenuation

Signal attenuation is perhaps the most fundamental technical problem to be addressed, because no radio positioning system can operate when the tracking signals are too weak to be received. Any receiver is essentially a device which separates signals from noise, and recovers the information content of those signals. Usually, there are only one or a few sources of the desired signal, but there are many sources of noise in a typical environment. We will first consider consider signal attenuation, and then the expected sources of noise. The goal of our design process is to maximize the signal to noise ratio at the receiver by minimizing both noise and attenuation.

In general, given typical wood or reinforced concrete buildings, attenuation is an increasing function of frequency. This explains why GPS does not work indoors; its microwave carrier signal at 1.5GHz undergoes severe attenuation in the course of passing through even the thinnest roofing materials. Other, commercially available indoor radio positioning systems use the ISM allocation at 2.4-2.45GHz. These frequencies are attenuated even more severely by passage through materials as common as reinforced concrete, people, and plant leaves. It is not uncommon for a 2.4GHz signal to be attenuated by more than 20dB when a single reinforced concrete wall appears in an otherwise free-space signal path. Rain induced attenuations at 1.5GHz can reach as much as 12dB [GPS95], which can render the GPS signals inaudible to receivers with poor sensitivity. In the link budget analysis presented later, we will make several assumptions in an effort to bound path losses at 1.9MHz and determine a lower bound on the system's expected signal to noise ratio in actual use.

4-3: Types of Noise

For our purposes, noise on the BPS channel is of two distinct types. The first type of noise is an uncorrelated, additive white Gaussian noise (AWGN) which simply reduces the signal to noise ratio at the receiver. It does not cause any distortion in the phase of the system's signals as received at the receiving antenna, but it must be rejected by the receiver's signal processing algorithms. The second type of noise is phase noise from the transmitter, receiver, or channel, which will be covered in the next section.

4-4: Sources of Additive White Gaussian Noise

We will first consider sources of AWGN and means of mitigating its effects. Fundamental physical processes such as the motion of thermal electrons in the receiver's passive components, quantum effects in semiconductor devices that consitute the receiver, or the broadband impulses produced by a lightning strike are responsible for the most difficult sources of AWGN to control. It is worthwhile to note that even though quantum noise in transistors, for example, is not Gaussian in profile over a wide bandwidth, engineers will usually approximate its effects as Gaussian in the narrow bandwidths of interest.

4-5: Thermally Dependent Noise Sources

Thermally-dependent noise is the result of many different physical processes. It is manifest as resistor noise due to the motion of thermal electrons, as well as temperature dependent transistor noise which is largely a quantum effect. Regardless of the mechanism of its generation, the best known solution to such noise is to build a receiver that has a superconducting input filter and first amplifier. This is because the noise figure of a receiver with substantial front-end gain is determined principally by the sum of the attenuation of front-end passive components and the contribution of the noise caused by the semiconductor device providing the amplification. We find that the noise power generated in the resistance of the passive components increases with temperature as well as measurement bandwidth:

$$N = \kappa T W$$

Eq. 4-1: Thermal Noise in Resistors

Here, N is the noise power available across any resistive element, independent of its resistance, $\kappa = 1.38 \times 10^{-23} \frac{W}{(°K \cdot Hz)}$ is Boltzmann’s constant, T is the temperature of the passive components in degrees Kelvin, and W is the measurement bandwith in Hertz. The noise contribution from an active device, like a transistor, is called the noise figure of that device and is usually expressed as a ratio of input signal to noise ratio (SNR) to output SNR; it is most convenient to express noise figures in dB at a 290K.

If we wanted to build the most sensitive and selective receiver possible, we would have little choice but to consider the most microscopic effects. Cryogenically cooled receivers attack thermal noise from two fronts. First, superconducting filters have very low resistance and consequently extremely high Q factors. This means that they can pass signals of a very well defined bandwidth with little attenuation while rejecting unwanted signals with extremely high attenuation. This allows very good rejection of out-of-band signals as well as out-of-band noise. Second, cooling the semiconductor devices minimizes the noise that is generated when thermal electrons jump among quantum states. This noise is generally responsible for determining the noise figure of microwave receivers like those used for GPS. Superconducting filter technology, which until recently was used only in radio telescopes and NASA's Deep Space Network, is becoming commercially necessary as the proliferation of wireless devices in crowded urban environments results in ever-increasing levels of interference at cellular telephone base stations. These receivers are extremely sensitive and must be protected at all cost from out-of-band interference.

In our system, however, the dominant contributors to the signal to noise ratio at the position tracking receiver are the more macroscopic sources of noise in the form of atmospheric noise and adjacent-channel interference caused by errant electronic devices. This is because our signal bandwidth is relatively small, and thermally dependent noise sources are completely dominated by externally generated noise.

4-6: Atmospheric Noise

Lightning and other sources of atmospheric noise are problematic mainly at very low frequencies. This is partially because antennas for low frequency signals must be physically very large in order to be effective. Recently it has been determined that the charged portion of the Earth’s atmosphere can act as a very large resonant antenna which forms part of the global circuit for atmospheric electricity. This large antenna can radiate impulse noise from lightning discharges or electrostatic noise from the motion of charged clouds over great distances. Fortunately, the small antennas and the choice of a 1.9MHz carrier frequency for the proposed system minimizes this potential problem because the physically small receiving antennas couple very inefficiently to the electric field portion of the very long wavelength 1.9MHz signal. Their electrostatic shields are designed to admit only the magnetic field component of the radio wave; these shielded loop antennas thus provide immunity in large part to atmospheric noise pickup.

4-7: Noise From Other Electronic Devices

Probably the most significant source of noise in our system is that caused by interference from other electronic devices. This interference ranges from harmonics and subharmonics of computer clock signals to stray emissions from video displays and other electronic equipment. A typical television monitor, for example, has a circuit which deflects the CRT’s electron beam horizontally across its phosphor face. This circuit generates a sawtooth waveform at 18.75KHz and is notorious for the wide spectrum of its harmonics. Other internal oscillators commonly found in monitors and televisions operate at 1.84MHz, 3.58MHz, 14.3MHz, and 40MHz.

Because of the irritating effects of these circuits, and their ubiquity, computer monitor and television manufacturers must now go to great lengths to minimize the radiated harmonics of these signals. This is beginning to clear the spectrum enough to expose the next largest sources of interference. Computer clock signals are often the next major source of interference. In one computer the author recently surveyed with a spectrum analyzer, more than 20 signal peaks of widely varying frequencies and strengths were detected when the computer was examined with a small unshielded loop antenna placed on its plastic case. Any of these signals, if of sufficient power and in close proximity to a BPS receiver, could cause the confusion of the system’s signal processing chain and can therefore result in measurement errors.

4-8: Internally Generated Digital Hash Noise

We know that the Fourier transform of the delta function is 1. This mathematical fact explains many of the RF engineer’s sleepless nights. As digital circuits become faster, the rise and fall times of their signals must necessarily become faster. Fast switching of digital circuits generates broadband noise. The task of immunizing the BPS receiver from the hash noise of high speed digital circuits is made even more difficult because the BPS unit itself contains many of these circuits! For this reason, metal shield partitions form Faraday cages around sensitive receiver circuits in the hope of reducing this source of noise as much as possible. Fortunately, these shield boxes also protect sensitive circuits from externally generated noise as well. Circuit board layouts, miniature coaxial cable, and ferrite beads are all used to prevent noise from entering these little Faraday cages.

This can only help to protect against noise which does not enter the receiver through the antenna or power ports, however. The use of a battery power supply and the use of an electrostatically shielded antenna help to reduce these pathways for noise entry.

4-9: Phase Noise

Phase noise has its origin in three major places. The first source of phase noise is the transmitting chain, especially the system's master oscillator unit. This noise is the result of random processes involved in the operation of the oscillators and amplifiers of the transmitter itself. The second source of phase noise is an environmentally generated position-sensitive phase offset. These offsets occur when the velocity with which a BPS signal travels varies because of objects in the environment. This is unavoidable, as the signals from the four spatially-separated transmitting stations necesssarily take different paths to reach the receiver. The third source of phase noise occurs because we are attempting to measure signal phases with an imperfect receiver, whose own clocks and sampling devices will exhibit phase jitter similar in nature to that caused by the transmitter's oscillators.

4-10: Phase Noise in Oscillators and Frequency Synthesizers

In our system, oscillator induced phase noise is minimized by the use of crystal oscillators and direct digital synthesis techniques. Crystal resonators have Q factors as high as one million and therefore act as their own very narrow bandpass filters. The measured phase noise of the temperature compensated crystal oscillator module used in the BPS transmitter unit is -140dBc at a 10KHz offset. This means that the signal from the oscillator is as pure as can be measured on a sophisticated spectrum analyzer without the use of extraordinary measures; the measured spurious power at the 10KHz offset is very small. If the oscillator output were a 1 Watt signal, the noise 10KHz away would be 10 femtowatts. This oscillator is specified to be stable to within 10ppm per day at room temperature and is not a major contributor to phase measurement errors.

The transmitted signal itself is derived directly from the master crystal oscillator by the use of direct digital synthesis (DDS) techniques. The DDS is a digital device which uses internal logic to perform sine wave lookup from a numerical table stored in the device's on chip ROM. These numbers are then reconstituted at a programmable rate in the device's 10-bit digital to analog converter (DAC) to provide a signal whose frequency and phase may be digitally adjusted by means of a 40-bit tuning word. As the sine wave table is fixed in ROM and the rate of its use is controlled by an extremely low noise crystal oscillator, the DDS's close-in phase noise performance is dominated by the performance of its DAC. Fortunately, the DAC is very good and the aggregate phase noise is roughly equivalent to that of a good LC oscillator; according to Analog Devices in [Ana99], the manufacturer of the DDS IC used in our system, phase noise of -120dBc at 10KHz offset is typical.

4-11: Jitter in Sampling Systems

Phase measurement errors encountered in the receiver unit are caused largely by the jitter of the sample and hold (S/H) unit which is part of the receiver's analog-to-digital converter (ADC). This is due to the fact that all receiver clocks are derived from a crystal oscillator similar in specifications to that of the transmitter. This ADC sampling jitter is specified in [Ana98] to be better than plus or minus 100 picoseconds. Translating this time-domain measurement into a perceived phase error is difficult because of the averaging process that the receiver performs, but the raw 100pS jitter corresponds to approximately 1.62 millidegrees of phase jitter at each sample when the 455KHz receiver intermediate frequency (IF) is digitized. This jitter is the major contributor to non-environmental phase measurement errors.

4-12: Phase Offset Errors of Environmental Origin

The estimation of phase errors caused by environmental effects is troublesome. At microwave frequencies these errors take the form of multipath propagation problems which can at times be extremely severe in areas surrounded by reflective walls or metal objects. The situation is not so clear at low frequencies. To a first-order approximation, physical intuition suggests that metal objects significantly smaller in size than the transmitted signal’s wavelength should present relatively little reflectivity to incident radio waves. However, higher order effects can become very important when extremely precise measurements are made.

Practical experience suggests that these effects do indeed become important for large objects. For example, the US Coast Guard’s published LORAN-C navigational charts include offset errors caused by large metal structures such as the Golden Gate Bridge. According to the Coast Guard in [LOR95], mariners making LORAN-C measurements while located closer than 100m from the bridge should be alert to the possibility of positioning errors and should include the fixed offsets found on the Coast Guard’s navigational charts to account for these errors. This is in some ways encouraging; indoor phase errors of environmental cause should be limited due to the fact that the Golden Gate Bridge is much too large to fit inside a typical building. More research and some experimenting will be needed to determine whether very close proximity to much smaller objects will result in significant position errors in a typical indoor case.

Chapter 5: Positioning by Phase Measurement

5-1: Overview

In this chapter we will examine how poisiton can be determined by signal phase measurement. We will consider this problem in two parts. First, the forward problem is presented. This is the problem of determining what the relative phases of the signals from the BPS fixed stations should be at a receiver located in a particular place relative to those fixed transmitters. The second part is the inverse problem, which is actually what the BPS receiver must do. This is the problem of determining the receiver's position, given a set of signals received from the BPS fixed station. The basic signal processing tools needed for this task are discussed in [Opp89].

5-2: The Forward Problem

To determine the relative phases of signals received at a remote point, we must consider the basic facts of radio wave propagation. For our purposes, waves travel at a fixed speed, the speed of light in a vacuum, $3\times10^{8}\ \frac{m}{S}$. In our system, signals are emitted from BPS transmitter stations S0-S3 which are placed at particular locations within an imaginary cubic building. These signals emit spherical wavefronts which travel until they are received at the BPS receiver R, which is located at some arbitrary place in our building. The system looks like the drawing in Figure 5-1, which omits the internal structure of the building for clarity. We assume that the maximum distance between any transmitter and any possible receiver position is less than one wavelength of the transmitted signal. This is one reason why it is convenient for BPS to operate at low frequencies; at 1.9MHz, the wavelength of the transmitted signal is about 157m, which is larger than the building for which the system is designed.

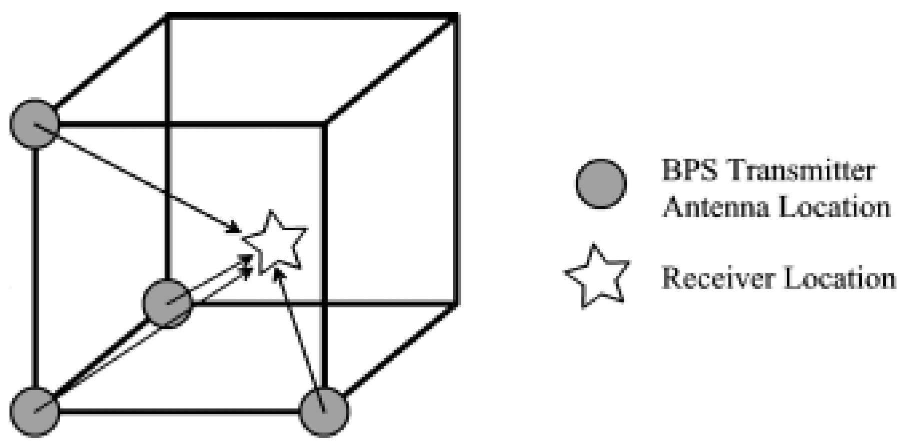


Figure 5-1: Placement of Transmitters and Receiver

Assuming that all transmitting stations emit pure, sinusoidal signals that are of the same frequency f, with zero phase offset, we can use the fact that the distance D_i from each transmitter to the receiver adds a finite transit delay. We will express this delay as a fraction of the signal's wavelength and denote it with the symbol $\phi_i = \frac{2\pi f D_i}{c}$ in Eq. 5-1.

$$\begin{aligned} S_0(t) &= \sin(2\pi f + \phi_0) \\ S_1(t) &= \sin(2\pi f + \phi_1) \\ S_2(t) &= \sin(2\pi f + \phi_2) \\ S_3(t) &= \sin(2\pi f + \phi_3) \end{aligned}$$

Eq. 5-1: Forward Model

This is our simplified forward model for the signals received at the receiver. Graphically, at the receiver's location our transmitted signals would look like those shown below in Figure 5-2.

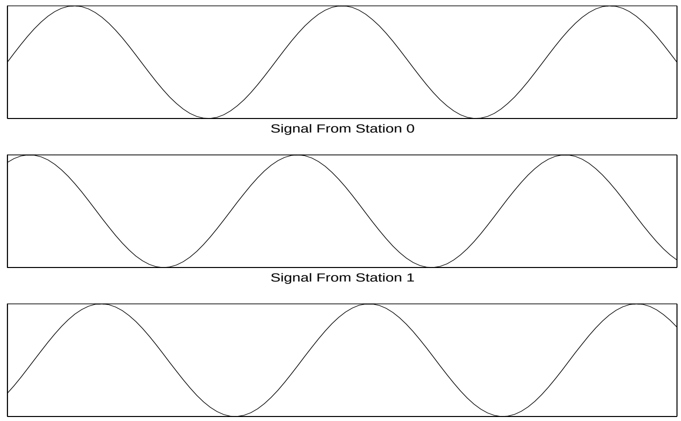


Figure 5-2: Transmitted Signals from Stations S0, S1, and S2

5-3: The Inverse Problem

As long as the transmitted signal's wavelength is longer than the diagonal of our cubic building in Figure 5-1, we notice that the transit time of the signals from each transmitter T_i to the receiver adds a non-zero phase offset ϕ_i to the measured signal. This offset is related to the receiver's distance from the transmitter, the speed of light $c = 3 \times 10^8 \frac{m}{S}$, and to the transmitted frequency f as shown in Eq. 5-2. There is a lurking issue, however; phase is only a meaningful concept when expressed with respect to some reference. Stated another way, our problem is complicated by the fact that there is no notion of absolute phase; we can only measure phase differences between pairs of signals.

$$\frac{c\phi_i}{2\pi f} = D_i$$

Eq. 5-2: Phase-Distance Equation

What the receiver can really measure, then, are phase differences between pairs of stations. To make this useful, though, we need a way to measure the phase difference between two signals. Analog engineers working with continuous-time systems would use a device called a phase detector. A phase detector consists of a mixer (or analog multiplier) followed by a low pass filter (or integrator). This is shown in block diagram form in Figure 5-3.

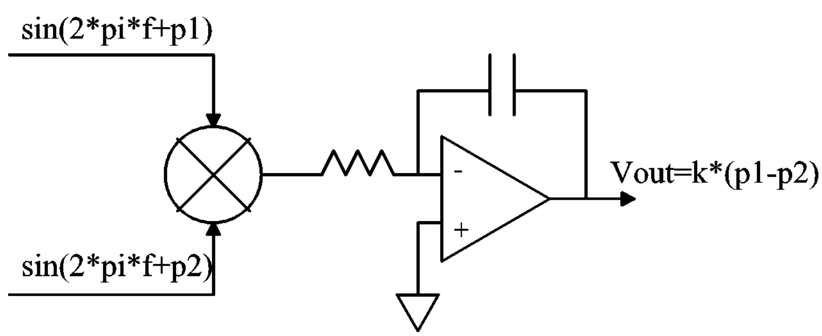


Figure 5-3: Analog phase detector circuit

In this circuit, the output voltage is some constant times the difference in phase between the two input signals, as long as those signals are sinusoids of the same amplitude and frequency. The same technique may be applied to digital signal processing of discrete time signals. That is what the BPS receiver does. Two oversampled sinusoidal signal vectors are normalized. They are multiplied pointwise into a buffer vector, and the mean of that buffer vector is taken. This is the dot product of normalized vectors, divided by the number of samples in the input vectors. The resultant is a value from -1 to 1, representing a phase difference of $-\pi$ to π.

Figure 5-4 illustrates this point by displaying the three pairs of phase differences that arise between any three of the four fixed stations. The source signals for these plots is derived from the same forward model developed in Eq. 5-1. Plotted in the bottom half of Figure 5-4 is the dot product of the two signals in the upper half, along with the mean of that dot product. This shows how the mean of the dot product of two sinusoidal signals of the same frequency indicates the phase difference between them

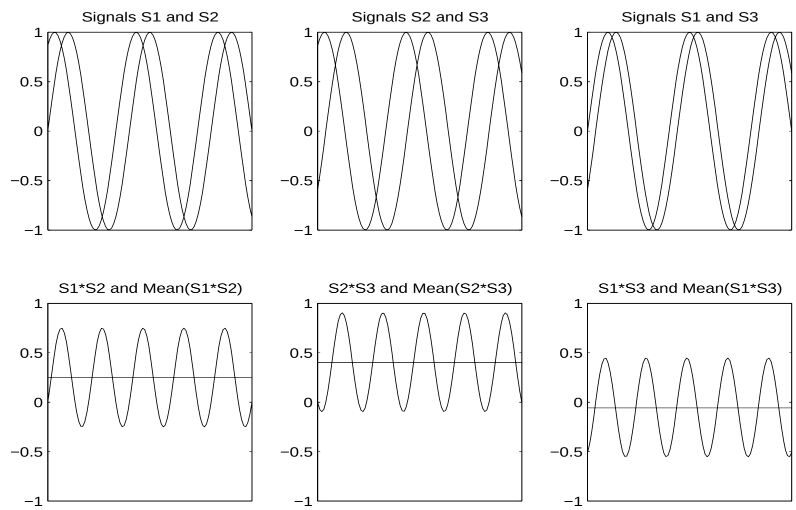


Figure 5-4: Pairwise phase differences for three stations

The last problem is how to take these phase differences and derive the location of the receiver. The issue is that phase differences between stations do not indicate the distance to a particular station. Instead, they represent the difference in distance between two stations, thus leading to hyperbolic curves of constant differential distances. This is why

LORAN operators used maps and overlays; these overlays were graphical calculators for determining the intersection of several hyperbolic curves of the type shown in Figure 5-5.

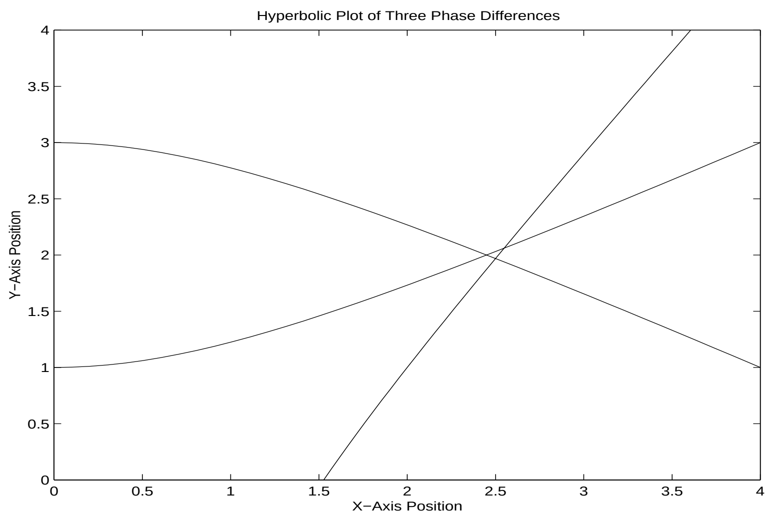


Figure 5-5: Position Solution by Intersecting Hyperboloids

In Figure 5-5 we see the intersection of three hyperbolic curves, derived from the measurement of all three pairs of signals taken from three transmitting stations. While it is only strictly necessary to intersect two hyperboloids to generate a two dimensional position solution, this graph shows why it is best to intersect every available pair. The position solution reported by the receiver should lie somewhere within the small triangular area at the intersection. Because of noise in the measurement, the three curves do not intersect in a single point. This is a reason to provide more than the minimum number of transmitting stations (we provide four stations to determine position in three dimensions). Each additional transmitter adds more information to the position solution process.

When deciding which point to report as the receiver's position, the software should take into account such factors as received signal strength, last known position, and especially the angle formed between each pair of hyperboloids. Uncertainty is minimized when two hyperbolic sections intersect at right angles, because small noise inputs lead to large changes in the point of intersection. In Figure 5-5, the position output from the receiver should not be as close to the intersections with the curve whose Y-intercept is 1 as it is to the other intersection. That is because the middle curve intersects the solution region at a more acute angle than the other two curves.

Mathematically, the receiver should use a Kalman-filtering algorithm as described in [Ger98] to filter the position solutions before they are output to the user. A Kalman filter is a way of formalizing the notion of having an adaptive weighting among many factors, each of which contribute some information about the user's position. Kalman filters are commonly used in all kinds of navigation systems, including missile guidance and ship navigation systems. For example, the Kalman filter "knows" that people do not move discontinuously; if the position solution from the hyperbolic solver suddenly jumps farther away by 20 meters in less time than a person could move 20 meters, the Kalman filter would assign less weight to the current position solution and more weight to the previous position solution, and would report an answer closer to the previous solution.

As usual in cases where noise or Nature intersect measurement needs, solving the inverse problem is much harder than solving the forward problem. There is no doubt that there are better methods than these that would minimize errors as determined by one metric or another. Neil Gershenfeld has proposed that this problem be cast in the form of a nonlinear search problem, and he suggests the application of Powell's method [Ger98] to arrive at a solution with an iterative method. This method could doubtless be made to work, but further research would be necessary to characterize the types of noise that are experienced in an operating BPS system before making a decision about the use of one method of solution or another. I have proposed what I think is the simplest, most straightforward solution and am waiting for the completion of the system to elucidate the shortcomings of this method.

5-4: Indoor Wave Propagation and Phase Errors

All electromagnetic waves, regardless of their frequency or power, consist of an electric field component and a perpendicular magnetic field component. The wave travels in the direction of the Poynting vector, $\vec{P} = \vec{E} \times \vec{B}$, which is in the direction perpendicular to both the electric and magnetic field components. We have all learned that, as is assumed above, an electromagnetic wave "travels at the speed of light", $3 \times 10^8 \frac{m}{s}$. However, when we solve Maxwell's equations for their wave solution, we notice that the velocity of wave propagation is not really a constant. It depends on the material through which the wave is travelling. We usually write the wave propagation velocity in a material as $V = \frac{1}{\sqrt{\mu \cdot \varepsilon}}$, where μ is the magnetic permeability of the material and ε is its dielectric constant. It is a good thing that a vacuum has a small but nonzero dielectric constant ε_0 and permeability μ_0, so that $\frac{1}{\sqrt{(\mu_0 \varepsilon_0)}} = 3 \times 10^8 \frac{m}{s}$. It is also a good thing that air is not very dense; because there is little "stuff" in air, electromagnetic waves travel almost as fast in air as in a vacuum. It is only when the BPS signals travel through large quantities of materials with high permeabilities or high dielectric constants that we have to worry about phase errors. For example, soft iron has a permeability over a thousand times greater than air, and plastics like teflon can have a dielectric constant two to ten times greater than air.

For this reason we might suspect that in a typical building environment, full of metals and plastics, the signals from the BPS transmitters will be travelling slower than expected, leading the system to calculate distances that are longer than they should be and thus report erroneous position information. It is not that simple. One reason is that when a wave that has been travelling in air hits the edge of another material it must abruptly slow down. This leads to reflection at the boundary between the air and the foreign object. Only a fraction of the incident power actually travels through the object; the rest of the power is reflected back and continues to propagate in the air in another direction. Because it is very difficult to predict the paths a wave will travel in a building, further work and a lot of experimentation is needed to make any quantitative statements about indoor propagation. This is especially true at the low frequencies that BPS uses; most of the existing experimental work has been done in the 100MHz-2GHz spectrum used for land mobile radios and cellular telephones. Low frequency propagation in buildings is still largely unexplored.

5-5: A Special Note on the History of Radio Navigation

The underlying mathematical and physical principles underlying this chapter have been well understood for many years. I wish to take this opportunity to acknowledge the pioneering work on radar and LORAN carried out at the MIT Radiation Laboratory in the 1940s. The scientists working at the Rad Lab broke new ground in many areas of engineering and physics while laboring under the enormous pressure of World War II. Reading the volumes of the original Rad Lab series in [Rad48] and [Hal47] has given me a profound appreciation for that work and the people who conducted it. For the authoritative history on the development of radar, which also touches on many of the engineering and social issues surrounding radio navigation, the reader is directed to Robert Buderi's excellent book [Bud97].

Chapter 6: BPS Signal Structure

6-1: Overview

The basic BPS signal structure consists of a set of 10mS frames, each of which is formed from four signal periods and one data period. The four signal periods are each 1.5mS long and consist of unmodulated carrier signal from each of the four primary location service transmitters. The remaining 4mS of each frame consists of 3.9mS of data period, followed by 100uS of silence. The 3.9mS data period consists of digital data transmitted by primary location service transmitter 0, using an FSK modulation at 10.25Kbps. The silent periods serve to allow the receiver sufficient time to recognize the end of the frame and to process the data received during each frame's data period. This is shown graphically in Figure 6-1 below.

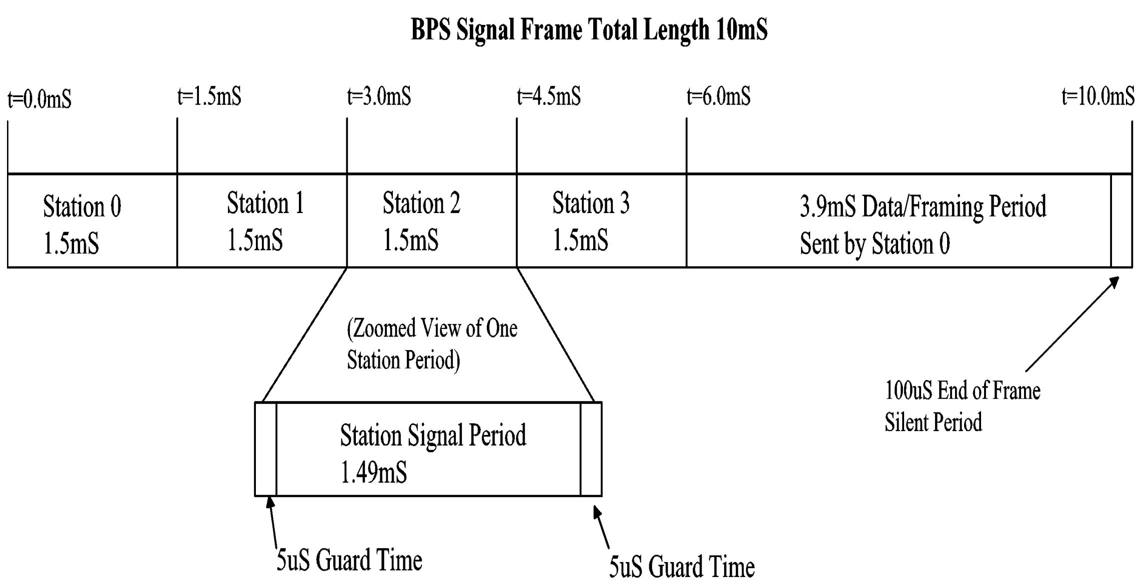


Figure 6-1: BPS Signal Frame Timing

6-2: Frame Alignment Method

The receiver's sampling engine digitizes the receiver's signal strength indicator (RSSI) voltage with the microprocessor's slow ADC and looks for a 100uS period of silence from the receiver IF module signifying the imminent beginning of a signal frame. When the RSSI indicates a silent period of 100uS, the high speed acquisition system begins the acquire period on each of the four transmitter phase intervals. The receiver's DMA system is capable of capturing 6.5mS (65536 samples at 10Msamples/sec) in a single shot measurement, but we capture only the first 6.0mS of the signal frame with the fast ADC. This results in a buffer of 60K 12-bit words, or 90K bytes of raw data. These DMA periods include 5.0uS dead times at the beginning and end of each subframe period, yielding a total of 10.0uS to allow for the transmitter switching and power control systems to settle.

The 6.0mS DMA periods are timed by the timer pattern generator unit in the SH7034's peripheral controller, which requests a highest-priority DMA transfer at the appropriate times. The processor is essentially idle during these periods, but the system bus is saturated by the ADC data filling its SRAM buffer. A simple refinement would be to have correlation or user interface routines running from on-chip RAM as long as they do not require the bus. The 90K byte sample buffer is then passed to the non-real-time remainder of the signal processing system.

6-3: Framing Error Conditions

If two entire frame periods (20mS) pass with no silence interval, the receiver assumes that it is encountering a "Jam" condition, reports the last valid position, and raises the “Jam” indicator on the operator display. It also reports an error code in its serial output stream. This could be the result of interference from enviromental sources, including the close proximity of two BPS systems. If no frames are received at all, or if the RSSI shows too weak an incoming signal, a “No signal” indication is presented to the user. The same indication is also presented if, for some reason, the processor records an unexpected drop in signal levels during the processing of the acquired signal frame. This will not be noticed during the acquire period itself, but is apparent later when the receiver processes each received signal frame.

6-4: Frame Data Interval

The 3.9mS data subframes, each containing data sent at 10.25Kbit/sec, transfer 40 bits of information to the receiver per frame. This occurs at a rate of 100 frames per second to form a 4.0Kbps data downlink channel which is used to communicate the BPS system ID and operating health of the four primary channels. It is sent in a narrow-shift FSK data format in which a logic one is represented by shifting the carrier frequency of Station 0 down by 10.0KHz. This modulation is done in the transmitter unit by loading the transmitter’s direct digital synthesizer with the appropriate new frequency value and restoring it when needed. This signal is demodulated by the simple expedient of digitizing the receiver’s RSSI signal exactly as is done to locate the end-of-frame gaps and looking for the small deviations in RSSI that occur as the received carrier shifts into and out of the 6.0KHz passband of the ceramic IF filter. This is a digital version of slope-detected FM and is a crude but effective way to send small amounts of data to the receiver without requiring additional hardware.

An essential component of the downlink signal is the surveyed location of the transmitter group's origin point. This is sent as an NMEA string by Station 0 as part of the data period. This could be obtained from a long-time-averaged GPS measurement if desired. All computed position solutions are of course made with reference to the group origin point.

6-5: Interframe Signal Processing

The correlation engine is the processor's main task. It consists of a set of sliding multiply-accumulate operations that seek to achieve maximal correlation among the four transmit periods, and records the offset in samples between each pair of transmit periods. The subsample interpolation used to meet the target position resolution goals is based on a simple sine-fitting algorithm that tries to fit a maximal-amplitude sine wave to the given point set. For our four transmit periods we can derive six pairs of phase offsets, which are then used by the position solution task to calculate a position estimate. This process is limited by the bus fetch speed to retrieve the appropriate samples from the main memory. Further discussion appears in Chapter 8 when the receiver unit is discussed.

Chapter 7: Link Parameter Calculation

7-1: Overview

To determine whether a radio frequency system can be made to work in a variety of environments, the system engineer will generally perform a link budget analysis. The objective of the link budget analysis is to account for loss, inefficiency, and noise to ensure an adequate signal to noise ratio at the receiver. This analysis, in the case of BPS, is not as straightforward as a typical satellite link budget due to the poorly understood nature of indoor radio propagation. Until the system is built and signal strengths are measured we can only make educated guesses about the attenuation encountered in a typical building. We allow for a very high link margin to account for all of the inevitably unexpected sources of signal loss in our system. A good discussion of the use of link budget calculations in the context of satellite systems may be found in [Skl88].

7-2: Transmitting Antenna Performance Estimation

A key parameter of the antenna system is the antenna efficiency. Both the transmitting and receiving antennas used in this system are loops whose diameter is a very small fraction of the wavelength of the transmitted signal. In our case, the transmitting antenna is a 5-turn loop of 0.5cm diameter copper refrigeration tubing that is wound around a 0.5m diameter plastic form. Its measured inductance is about 100uH, and its Q factor is about 150. With these parameters we have enough information to make an estimate of the antenna's efficiency. This section follows the standard procedure for the small loop approximation found in [Bal96].

We will first calculate the antenna's radiation resistance as follows, where R_s is the antenna's radiation resistance, A is its area, and λ is the wavelength of the transmitted signal:

$$R_s = 31\text{k}\Omega\left(\frac{A}{\lambda^2}\right)^2$$

Eq. 7-1: Radiation Resistance

In our case, given the transmitting antenna's parameters as described above, the radiation resistance R_s is 1.92×10^{-6} Ohms. This is a small number but is typical of small transmitting loops used at these frequencies. We now have enough information to calculate the antenna's efficiency. In this expression, Q is the antenna quality factor, R_s is its radiation resistance, f_o is the operating frequency, and L_{loop} is the loop's inductance.

$$\eta = \frac{Q \cdot R_s}{2\pi \cdot f_o \cdot L_{loop}}$$

Eq. 7-2: Antenna Efficiency

This yields an efficiency of 2.4×10^{-7}. This is also typical of such a small loop at these frequencies. This efficiency may be written as a loss of 66.1dB. This is the reason why we must have such a high transmitted power to cover such a small area. The effective radiated power of our +43dBm transmitter is therefore 43dBm-66.1dB=-23.1dBm.

7-3: Receiving Antenna Performance Estimation

Performing a similar analysis at the receiving antenna is difficult because the receiving antenna is a very small ferrite core inductor. In this case, the inductance of the loop is increased dramatically by the very high permeability of the ferrite material. Additionally, the effective loop area also increases because of the more permeable medium inside the loop. Because the transmitting loop is less than one wavelength from the receiving loop, and the receiver is therefore in the near field of the transmitting antenna, we will assume that the two loops act more like two inductors loosely coupled like a transformer. It is therefore very difficult to predict this efficiency without very detailed near field modelling of the B-fields surrounding the transmitting antennas. We will therefore assume a loss of 40dB as is typical of small ferrite rod antennas for AM radio use.

7-4: Freespace Path Loss Calculation

Next, we will calculate the path loss of the freespace path between the transmitter and the receiver. The worst case path in the Media Lab's cube-shaped Weisner Building is approximately 50m, measured from the bottom corner of the basement to the top opposite corner of the fourth floor. The Friis free-space path loss equation is used, which is derived from an assumption of spherical spreading of the wave front at both the transmitter and receiver antennas. Here, d is the transmitter-receiver separation in km and f is the frequency in MHz.

$$L_o = 32.5\text{dB} + 20\log d + 20\log f$$

Eq. 7-3: Path Loss

In our system, we note that with d=0.05km and f=1.9MHz, path loss L_o=12.05dB.

Table 7-1: Link Budget Calculation

Origin	Parameter	Value
[Measured]	Transmitted Frequency	1.9MHz
[Measured]	Transmitter Power Output	+43dBm (20W)
[Estimated]	Transmitter Antenna Pattern Gain	+1.7dB (Near-dipole)

Table 7-1: Link Budget Calculation

Origin	Parameter	Value
[Calculated]	Transmitter Antenna Radiation Gain (loss)	-66dB
[Calculated]	Freespace Path Gain (loss) 50m path	-12dB
[Estimated]	Receiver Antenna Pattern Gain	+1.7dB (Near-dipole)
[Calculated]	Receiver Antenna Radiation Gain (loss)	-40dB
[Calculated]	Power at Antenna Port	-71dBm
[Measured]	Receiver Sensitivity	-100dBm for 10dB SNR
[Calculated]	System Link Margin	+28.4dB

7-5: Link Margin and Related Factors

These numbers are hypothetical, but the author believes them to be representative of the actual link parameters of a finished system. The process, in any case, is illustrative of the design process. The design engineer chooses values for variables which are easy to set, such as transmitter power or operating frequency, in order to provide adequate signal to noise ratio at the receiver given his worst-case assumptions about the behavior of an unknown signal propagation path. In our case, the allowable attenuation due to the presence of building materials in the signal path is at most about 28.4dB.

Chapter 8: The BPS DSP-based Receiver Unit

8-1: Overview

The BPS Receiver Unit is based on a digitized IF software radio architecture. As shown in Figure 8-1 below, the signal from the antenna is filtered, preamplified, mixed to a 455KHz intermediate frequency (IF), and is then amplified by a limiting amplifier. This signal is then passed to a high speed ADC which feeds the CPU's direct memory access (DMA) unit with a high speed streaming digitization of the incoming signal. The CPU acquires these signals in real time but processes them at leisure after each signal frame. This arrangement allows the use of a standard Hitachi SH-1 microcontroller instead of a dedicated DSP unit. Of course, this results in a considerable savings in cost and complexity. Another feature of this receiver is its direct digital synthesis (DDS) local oscillator unit, based on an Analog Devices DDS IC. This IC greatly simplifies the task of generating a clean local oscillator signal for the downconversion process.

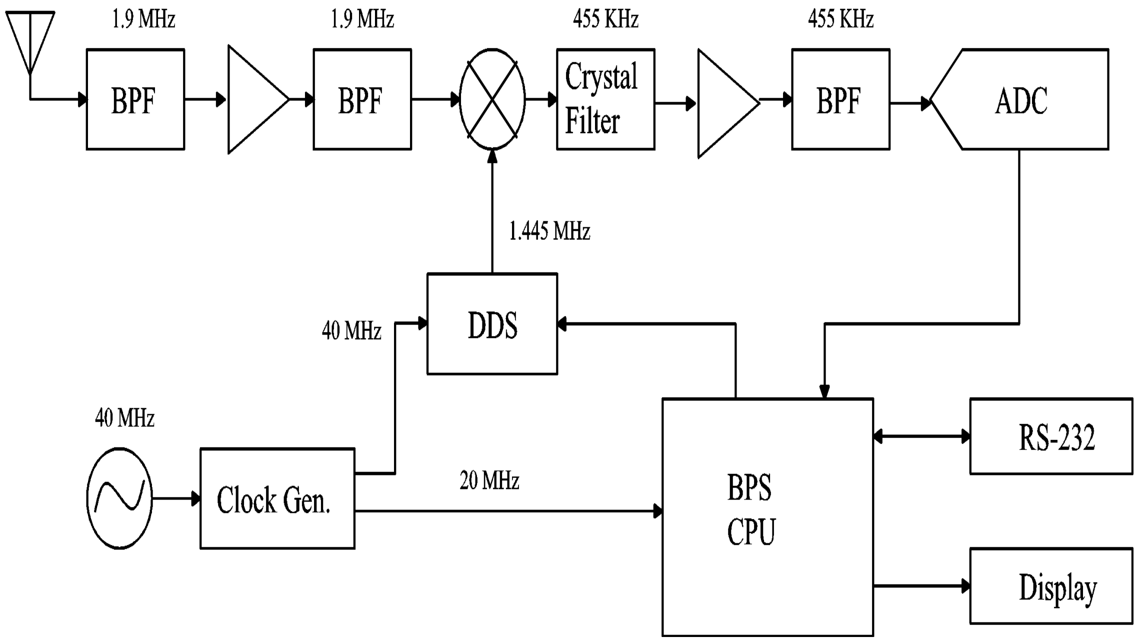


Figure 8-1: BPS Receiver Block Diagram

8-2: Receiver Front End

The receiver front end consists of a low-noise FET RF preamplifier with 5 poles of associated LC coupled-resonator filtering. Input impedance is 50 ohms, while the output of the preamp is matched to the 100 ohm impedance of the subsequent mixer stage; it is transformer coupled to an Analog Devices AD608 mixer-IF IC which provies a Gilbert-cell mixer as well as a limiting amplifier and extremely accurate log-RSSI circuitry. This receiver uses a single conversion design with a 455KHz IF, which is possible due to the relatively high Q input matching and filtering circuits. Signals that are 455KHz above or below the desired signal frequency are attenuated by more than 40dB, with an additional 20dB of selectivity due to the high Q loop antenna.

8-3: Logarithmic Detection and RSSI A/D Conversion

In order to detect the gaps in the transmitted signal which indicate the start and end of each transmitted frame, the CPU's internal 12-bit ADC is used to capture the received signal strength indicator (RSSI) voltage from the AD608 IF device. The RSSI output is logarithmic with respect with to the incoming signal over a range of 60dB and is low pass filtered with a one-pole RC lowpass filter which has a cutoff frequency of 10KHz. A simple filter method is sufficient to remove high-frequency noise due to the much higher 455KHz IF signal which is the unwanted byproduct of the RSSI's logarithmic generation process. This signal is then buffered by a high-speed opamp and is connected to the BPS CPU's internal 12-bit low speed A/D converter for the digital measurement of RSSI. An appropriate compensation function is used to report actual RSSI in dBm units for test purposes. As the RSSI is used as a source of framing information the high cutoff frequency is important to allow the signal acquisition to occur within the 100uS inter-frame interval.

8-4: IF Stages and Filtering

The receiver's 455KHz IF is then filtered by a four-pole ceramic filter with a 3dB bandwidth of 6KHz, and is subsequently limited to a 200mV pk/pk square wave signal by the AD608's internal limiter stage. The limiter stage is of great help in removing amplitude noise contributed by AWGN sources. Shown below is a typical noisy IF signal before, during, and after the limiting process. The sine wave is first limited to a square wave, and then undergoes a sine wave restoration in a subsequent filter.

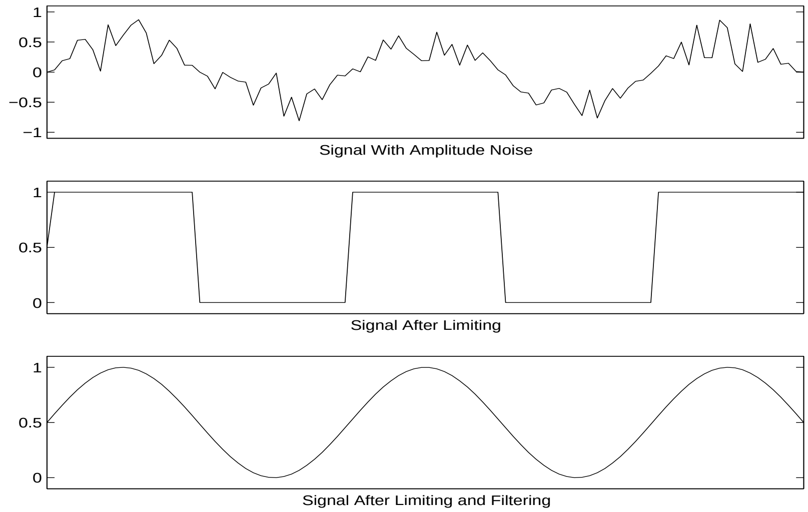


Figure 8-2: Signal with AWGN During Processing

That square wave signal is then applied to a bipolar tuned amplifier stage which operates in a Class C mode. This Class C output stage has a high-Q output tank which then regenerates a sine wave at the same frequency as the IF signal as shown above. This has the advantage of providing a very selective filtering effect without introducing appreciable phase distortion to the differential measurements; as all four transmitters are guaranteed to be transmitting at the exact same frequency, the phase errors through the IF stage are irrelevant. This is one great advantage of a DSP-based receiver system. If an analog receiver were designed, the four individual receiver chains would have to be carefully aligned to preserve phase from signal to signal. Even though the initial receiver stages could be reused for all four receiver chains, the switching process would cause an inevitable noise output which would then have to be filtered later, causing possible phase accuracy problems.

8-5: Signal Conditioning for A/D Conversion

The sine wave RF from the Class-C stage's tank is inherently amplitude limited, and a simple potentiometer adjusts this signal to the proper level for digitization. A Maxim MAX474 high-speed rail-to-rail opamp buffers this signal, and provides a level shift to a 2.048V biased centerline. The resulting signal is digitized at 10Msamples/second by an Analog Devices AD9220 CMOS 12-bit ADC with internal 4.096 bandgap voltage reference; this bandgap reference is also used to generate the 2.048V reference. The AD9220 ADC is a pipelined converter which is always running; the output word changes on falling clock edges but is valid within about 70nS of the rising clock edges. This is fortunate because it makes processor interfacing extremely straightforward. The output bits are buffered by 74VHC logic bus drivers which latch the ADC bits on to the main processor bus during periods of DMA transfer.

8-6: Receiver Local Oscillator

The receiver local oscillator is an Analog Devices AD9850 Direct Digital Synthesizer. This device takes as its reference a 40MHz TCXO module and produces an output signal whose frequency and phase may be specified by a 40-bit control word. This tuning word consists of 32 bits of frequency control information as well as 5 bits of phase control information, which are not used in this design. This 32 bit frequency resolution corresponds to a frequency variation of 0.093Hz (40MHz/2^{32}) at the least significant bit. Its output is filtered by an 3-pole antialiasing LPF and has very good spectral purity (all spurious outputs more than 70dB below the carrier) at the desired 1.445MHz LO frequency. The tuning word is supplied in a serial fashion by the main processor. As a frequency update requires about 10uS, future experimentation with multifrequency operation is quite possible if the input filtering circuitry can successfully be readjusted to cover other possible carrier frequencies.

8-7: Receiver CPU and Interface

The main receiver CPU provides three services. One is a time critical single cycle latency DMA burst transfer of the digitized IF signal to main memory. The second service is the non-real-time analysis of the received signals to decide on the pairwise phase differences and to compute a position solution at a 100Hz rate. The third function is to provide user interface I/O to a small keypad, to a serial port, and to a VFD display unit. The Hitachi SH7034 SH-1 RISC microcontroller is capable of providing all of these services. It is a 32-bit RISC microcontroller clocked at 20MHz; it operates in single cycle mode yielding 20MIPS.

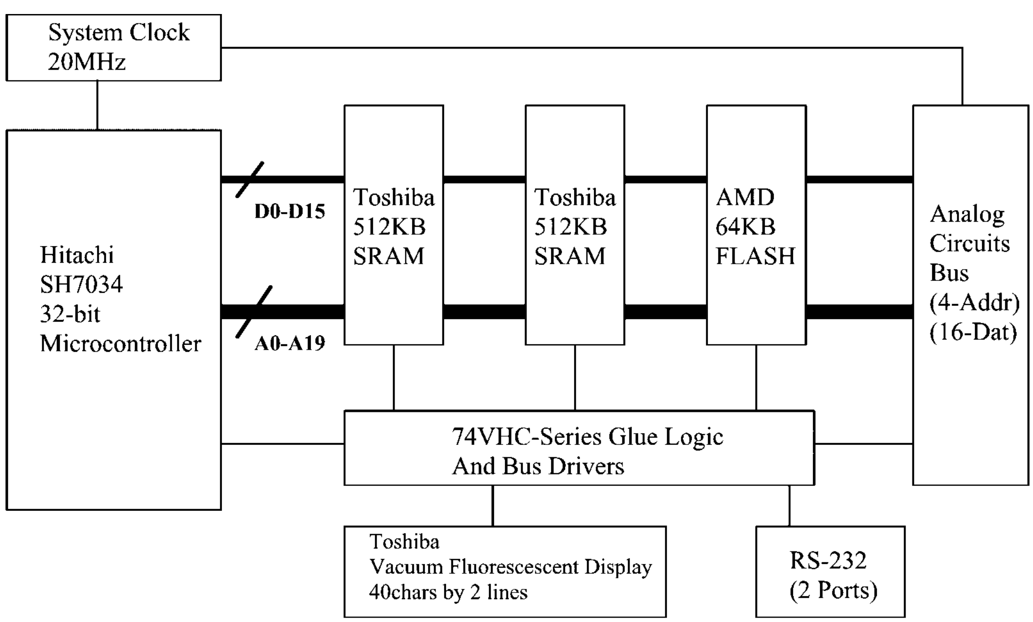


Figure 8-3: The BPS CPU Module

The Hitachi SH7034 is capable of directly addressing over 32Mbytes of storage; the high order address bits are broken out to chip select signals to minimize the need for external glue logic. In our case, as seen above, the system architecture consists of an AMD 1M-bit FLASH ROM from which program code and lookup tables are retrieved, as well as 1Mbyte of 70nS Toshiba static RAM. The static RAM is accessed on a single-cycle write mode so that the processor's DMA controller can provide address sequencing to drive the A/D converter results directly into a specified 64K-word memory block. The 12-bit resultants from the A/D converter subsystem are padded with zeros and stored in 16-bit memory words. From there the required compuations can be done while being interleaved with bursts of transmitted position information from the master clock unit.

8-8: Receiver Real-Time Software

The main processor runs a simple cooperative multitasking OS constructed by the author to run in synchrony with the received frames. This OS is derived from SH-1 monitor code originally written by my colleague Rehmi Post, also of MIT. This cooperative multitasking system is somewhat unique as it results in nondeterministic execution times for user interface (UI) code but extremely precise control of signal acquisition timing, which is necessary for the proper operation of the position computation cycle. The SH7034's internal timer unit is used to set periodic interrupts which then cause the DMA transfer of received signals at the appropriate times. There is a single low level routine which apportions CPU time between signal processing and UI tasks.

8-9: DMA Transfer of Signal Samples

The 60K-word DMA transfers of incoming signal streams consume the CPU's entire bus bandwidth for 6.0mS time slices. The SH7034's DMA controller limits the maximum frame time of our signals because of its 65536-sample maximum transfer word limit. We want all four transmitted signals to fit within a single sampling interval in order to prevent phase jitter differences from detracting from the system's phase accuracy. The A/D conversion clock is derived directly from a divider chain on the receiver unit's single 40MHz clock, thus eliminating CPU or DMA trigger time induced sampling phase jitter. Of course this mandates the use of a single global clock to ensure that the CPU's bus timing expectations are always met; if more than one oscillator were used the A/D clock would slowly slide in phase relative to the CPU clock and thus bus timing would suffer.

During the DMA intervals, the processor can continue to operate against data stored in its onboard static RAM area. This RAM area is used to perform the position solution algorithm. The processor's internal 32-bit data bus is logically disconnected from the external 16-bit bus during DMA periods, so separate operations may occur on the two busses. This makes possible the interleaving of position solution calculations and the gathering of another data set from the A/D converter system.

8-10: Receiver User Interface

The receiver's UI consists of two display options, a small numeric keypad, and two high speed serial ports for connection to an external PC. The first display option is a 4-line 40 character vacuum fluorescent display (VFD) which is readable from any angle, but which can only display symbols from its internal ASCII character set. The second option, which is not yet fully implemented, is a 256x128 full dot matrix monochrome LCD module with a CCFL backlight. Given sufficient development time, the dot matrix display may prove useful for a variety of geographic information service (GIS) applications, such as the display of maps or associated information which is keyed to the computed position solutions. The numeric keypad is used for basic system control when the BPS receiver is used in its standalone mode of operation, and is especially useful for inital test and measurement operations.

The primary BPS receiver I/O is performed through two serial ports. These ports are capable of being operated at speeds of up to 115Kbit/sec. One is allocated to a host-PC connection for the transfer of raw data during testing, and the output of NMEA-like position strings once the initial testing has been completed. The BPS unit presents a command-line interface through a system monitor application based on code initially written by Rehmi Post. The second serial port is intended for use as an interconnect to a secondary peripheral device, such as a conventional GPS receiver, which can provide less position solutions when the handheld unit is outside in view of the GPS constellation. The BPS unit's NMEA-like string will provide a quality of service metric which can be parsed to determine the source and estimated precision of its position solutions.

Chapter 9: BPS Transmitter Design

9-1: Overview

The BPS Transmitter is a key part of the system as a whole. It is responsible for generating a stable carrier signal, maintaining modulation timing, and transferring data packets to the receiver. It must generate power levels as high as +43dBm. A mixture of very delicate low level signal processing, as well as RF power electronics, the transmitter has been an interesting source of design challenges ranging from curing power supply ripple to watching one overstressed ferrite broadband transformer explode and disintegrate into many small pieces due to excessive circulating RF currents.

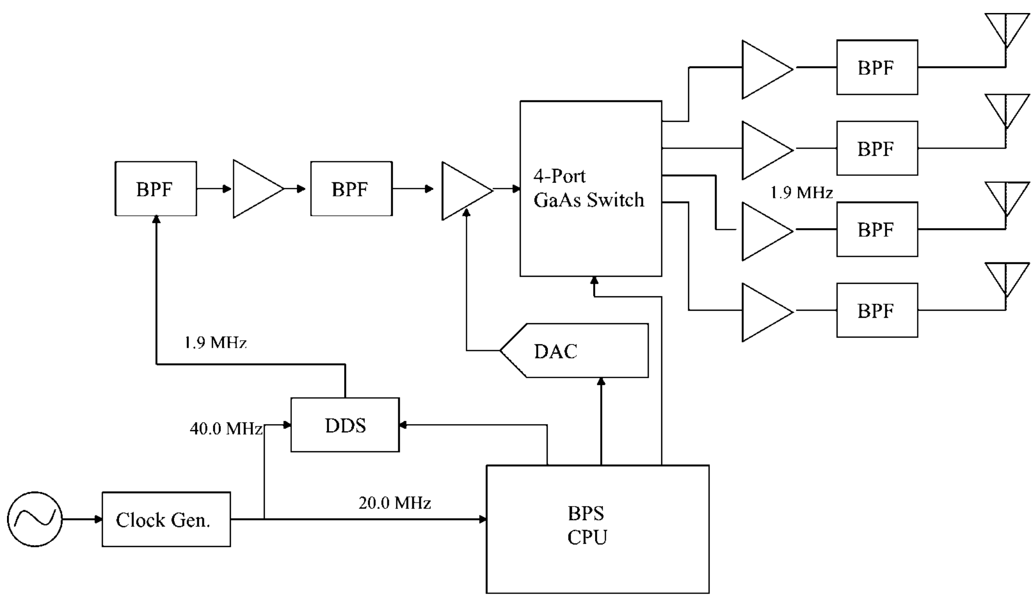


Figure 9-1: BPS Transmitter Block Diagram

9-2: Transmitter Master Clock and Power Amplifier Modules

The BPS Transmitter Unit, whose block diagram appears above, consists of two modules. One, the Master Clock module, is the source of all timing signals in the system. Every clock signal needed for BPS is ultimately derived from a single 40MHz temperature compensated crystal oscillator (TCXO) that is the heart of the Clock module. The Clock accepts serial I/O from a controlling computer and produces four low-level RF signals (+10dBm) which are available at rear-panel-mounted BNC jacks. The Clock module’s rear panel also carries an 8-pin Amplifier Control jack which allows for dynamic control of the power amplifier unit, primarily for bias scheduling of the Class A portions of the amplifier unit.

The other transmitter module is the Power Amplifier module. This module takes as its input the low-level RF signals as well as the Amplifier Control signals. It consists of four submodules, each of which is a complete broadband amplifier module based on two Motorola MRF455 bipolar RF power transistors in a balanced configuration. These modules operate in a Class AB linear mode and deliver a gain of 30dB at 1.9MHz. Efficiency, including bias currents, is about 45% which is roughly typical for such an amplifier.

9-3: Generation of Stable Signals

The BPS Master Clock unit consists of a four-output source of BPS timing signals whose frequency and power output are under microprocessor control. The unit’s master oscillator is a quartz crystal based 40MHz TCXO with a stated short term stability limited by the oscillator’s phase noise. A typical phase noise of -140dBc at 10KHz offset is quite sufficient for our application; this corresponds to a short term phase variation that is negligible compared to propagation based phase distortion. The unit's long term frequency stability is less than 10ppm per day at room temperature range. This, coupled with the Direct Digital Synthesizer (DDS) used to synthesize the BPS carrier frequency, ensures the long term stability of the BPS system. It is possible to update calibration values for the TCXO calibration input based on an external GPS reference signal at 10MHz, as the system includes a phase-synchronous 10MHz output that is ordinarily used as the A/D converter clock in a BPS receive unit. The system operator has access to this calibration output and can easily compare it to a high-accuracy external standard. Alternatively, if a source of extremely stable 40MHz TTL level clock signals is available, it is very easy to inject this external reference and dispense with the onboard TCXO altogether.

9-4: Transmitter Station Switching

The multiport switching system must have extremely high port-to-port isolation. A typical requirement is -70dB from the nominal carrier output. This is readily achieved by a combination of the software programmable gain circuit, which is used to reduce transmit power before the switching operation takes place to minimize switch bleedthrough on transitions, as well as the use of the MA/COM SW-114 GaAs multiport switch. This switch consists of a hybrid module of two dies on a ceramic substrate. One die is the GaAs FET switch element. The other die is a silicon ASIC which is responsible for level translation and break-before-make sequencing. The stated 60dB isolation increases to nearly 70dB at low frequencies where capacitive coupling dominates. The use of matched microstrip technology is necessary to maximize isolation, even at such a low operating frequency, due to the high isolations needed; the GaAs ASIC is rather intolerant of signal mismatches, so good design practices are mandatory.

9-5: Power Amplifier Design

The BPS power amplifier system consists of four identical modules, each capable of supplying up to 100W to a 50 ohm load. They are in practice operated at a significantly lower power output of 40W. As the phase stability requirements for the fixed-station transmitters is rather severe, these amplifier modules are operated in Class AB mode which unfortunately limits efficiency to at most 55%. An efficiency of 41% at 30W output power was measured at 1.9MHz. Phase error over the amplifier's 0.5-10MHz bandwidth must be less than 2 degrees, with the phase error within an arbitrary 10KHz bandwidth less than 0.01 degree. This is achievable with standard broadband amplifier technology as long as the use of high-Q tuned circuits is minimized. The use of ferrite broadband matching transformers is permissible if Q-swamping resistors and balancing capacitors are employed.

Each amplifier consists of a discrete-transistor preamplifier and driver which supplies roughly 15dB of gain. Thus the +10dBm signal from the synthesizer and switch board is amplified to +25dBm which is sufficient to drive the class-AB balanced power amplifier stage (using 2 Motorola MRF455 bipolar RF power transistors) to +45dBm to +50dBm on signal peaks. As duty cycle expected of any individual amplifier is less than 25%, heatsink design may be slightly easier than if a single high power amplifier were switched among many antenna ports. Fortunately the Clock unit supplies a set of bias control signals, which are used to gate the biases of the individual amplifier modules. The use of a power-FET based amplifier was considered but was not practical given the lack of available low-Vdss RF power MOSFETS. Standard power MOSFETs might be applicable if a driver stage capable of driving highly capacitive loads is available.

Output filtering is performed by a set of matched L/C low pass filter elements with resistor-swamped Q minimization, which are designed in the Butterworth lowpass configuration to minimize phase errors in their passbands. As the specified phase error tolerance would be difficult to meet given varying antenna loads, it is imperative that the antennas be carefully matched to the filter's 50-ohm impedance. They must be readjusted if the system operating frequency is changed for some reason.

9-6: Power Supply Issues

The amplifier system requires a clean source of +13.8V at 7A average. A peak current of 15A must be supplied, as the bias circuitry may produce very short high current transients. As the RF output is cycled at a 100Hz rate, it is necessary to eliminate any amplitude ripple which may result in a low frequency whine on the transmitted signal. The use of a large computer-grade electrolytic bypass capacitor may be necessary even with a high frequency switching power supply. It is worthwhile to note that this supply must be free of switching harmonics at 455KHz (the receiver 1st IF) and 1.9MHz (the transmit signal frequency) as the presence of these components even at very low levels may result in the transmission of very slightly amplitude modulated carrier signals. This is extremely undesirable in the high-precision mode of signal phase measurement that BPS employs, as any amplitude modulation that makes it through the limiting amplifier stage results in an apparent phase change at the receiver's A/D converter stage.

9-7: Transmitting Antenna Design

Frequency and environment based mismatching of the antenna system is the single largest cause of fixed phase offset errors. It is important to design the antenna system to minimize phase errors of environmental origin. As the antennas will be operating in the small loop limit, it will not be the movement of effective phase center that will be the chief problem, but rather the matching that will vary depending on the proximity of conducting objects. An antenna housing has been designed in the fashion of a small radome that will serve to keep people and objects from obstructing the antenna system. The radomes were once small, round office waste bins.

We have considered many possible antenna topologies that may be employed in a working BPS system. As transmitted signal wavelengths of 157M are used, it is basically impossible to design a small but efficient antenna. The engineer is then left with the fundamental choice of maximizing system Q to minimize radiation resistance (and thus try to minimize power coupling losses), or minimizing Q to faciliate phase matching among many antennas. We have chosen a compromise antenna Q of 150 which represents a relatively lossy antenna at these frequencies in the belief that we may make up for these losses with a peak amplifier power 10dB larger than would otherwise be necessary. Given the calculated link loss budget of 118dB and our +43dBm power output, receiver signals will be about -70dBm at the antenna which allows a link margin of about 28dB with our receiver's designed MDS of -100dBm. The relatively small noise power at the small receiving antennas employed works greatly in our favor in this case.

9-8: Antenna Mechanical Assembly

The transmitting antenna that is employed is a loop antenna of 50cm diameter fabricated from 0.5cm diameter copper refrigeration tubing. Five turns are used, which results in an inductance of about 100uH. This is made resonant with a 200pF vacuum variable capacitor with a voltage rating of 10kV. Caution must be exercised as peak RF voltages can easily reach over 5kV in normal operation. Voltages may be higher in the case of intentional or deliberate system mismatch. Any bystanders must of course be prevented from approaching closer than 1m to the antenna unit. The 50-ohm impedance of the feed-line is matched by a sliding tap arrangement which can be adjusted to provide a low VSWR to the transmitter.

9-9: Antenna Tuning Methods

The antenna is tuned with a network analyzer to be resonant at the appropriate operating frequency, and the transmitter's digital power control is set to provide a 5W test signal to aid in the final matching and tuning steps. The sliding tap is then adjusted along with the vacuum variable capacitor (these adjustments will interact) to maximize measured RSSI at a strategically placed test location. Amplifier power is then increased until the required RSSI is achieved at all service points.

Chapter 10: Concluding Thoughts

This system is not yet complete. Over the past several months, I have learned a lot about the design and implementation of radio navigation devices. I've also gained a lot of experience with complex hardware and software systems. This is the first project I have worked on that is simply too complicated to design and build without constantly documenting every hardware and software design and fabrication choice as it is made. I've often run into what seems to be an arbitrary decision I made a few weeks before, only to find that decision to be the source of an error costing days of debugging time. In the process, I have found problems ranging from badly documented integrated circuits, to a bug or two in the silicon of the SH-1 microcontroller, to one particularly insidious microscopic solder bridge that was located beneath a 100-pin surface mount package.

On the other hand, even the most painful experience has been valuable in its own way. The writing of this document was no exception, as I have struggled to find a style and format that renders a complicated system understandable to those who have not worked on it nearly constantly for months at a time. I hope that the reader can use my written description to find design errors that I have not yet found; if not, he or she may simply find the way to spell some word that eludes my weary fingers.

I must again thank the many people who have helped me as I have grown from a 10-year-old kid building radios in my basement to an MIT alumnus who wishes he had more time to spend building radios in his basement.

Happily, this is the very last sentence of my thesis.

List of References

[Ana98] "AD9221/9223/9220". Data Sheet Revision A. Analog Devices, Inc.,1998.

[Ana99] "AD9850". Data Sheet Revision C. Analog Devices, Inc., 1999.

[Bal96] Balanis, Constantine I. Antenna Theory: Analysis and Design. John Wiley and Sons, 1996.

[Bud97] Buderi, Robert. The Invention that Changed the World. Simon and Schuster, 1997.

[Dix94] Dixon, Robert. Spread Spectrum Systems With Commercial Applications. John Wiley and Sons, 1994.

[Fcc86] "FCC methods of measurements of radio noise emissions from industrial, scientific, and medical equipment". FCC/OST MP-5. Federal Communications Commission Office of Science and Technology. 1986.

[Fcc96] "Understanding the FCC regulations for low power, non licensed transmitters". OET Bulletin 63. Federal Communications Commission Office of Engineering and Technology. 1996.

[Ger98] Gershenfeld, Neil. The Nature of Mathematical Modelling. Cambridge University Press, 1998.

[GPS95] "Global Positioning System Standard Positioning Service Signal Specification". 2nd Ed. Air Force Space Command, 1995.

[Hal47] Hall, John Scoville. Radar aids to navigation. Mc Graw Hill Book Co., 1947.

[Hit96] "SuperH RISC Engine SH7032 and SH7034 Hardware Manual". 2nd Ed. Hitachi Micro Systems, Inc., 1996.

[LOR95] "LORAN-C User Handbook". US Coast Guard Navigation Center, 1995.

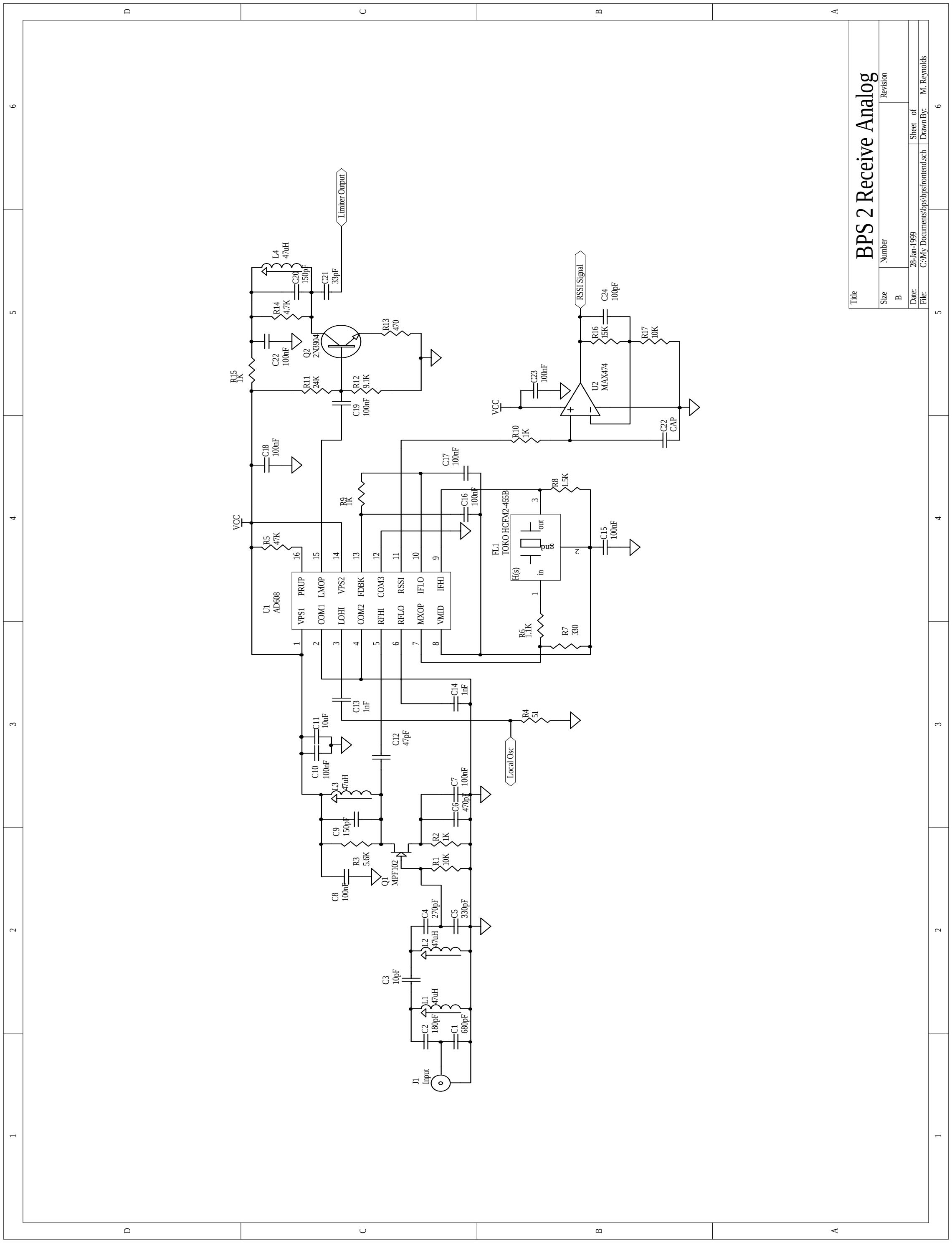
[Opp89] Oppenheim, Alan V. and Ronald Schafer. Discrete Time Signal Processing. Prentice Hall, 1989.

[Rad48] Pierce, John Alvin. LORAN, long range navigation. Mc Graw Hill Book Co., 1948.

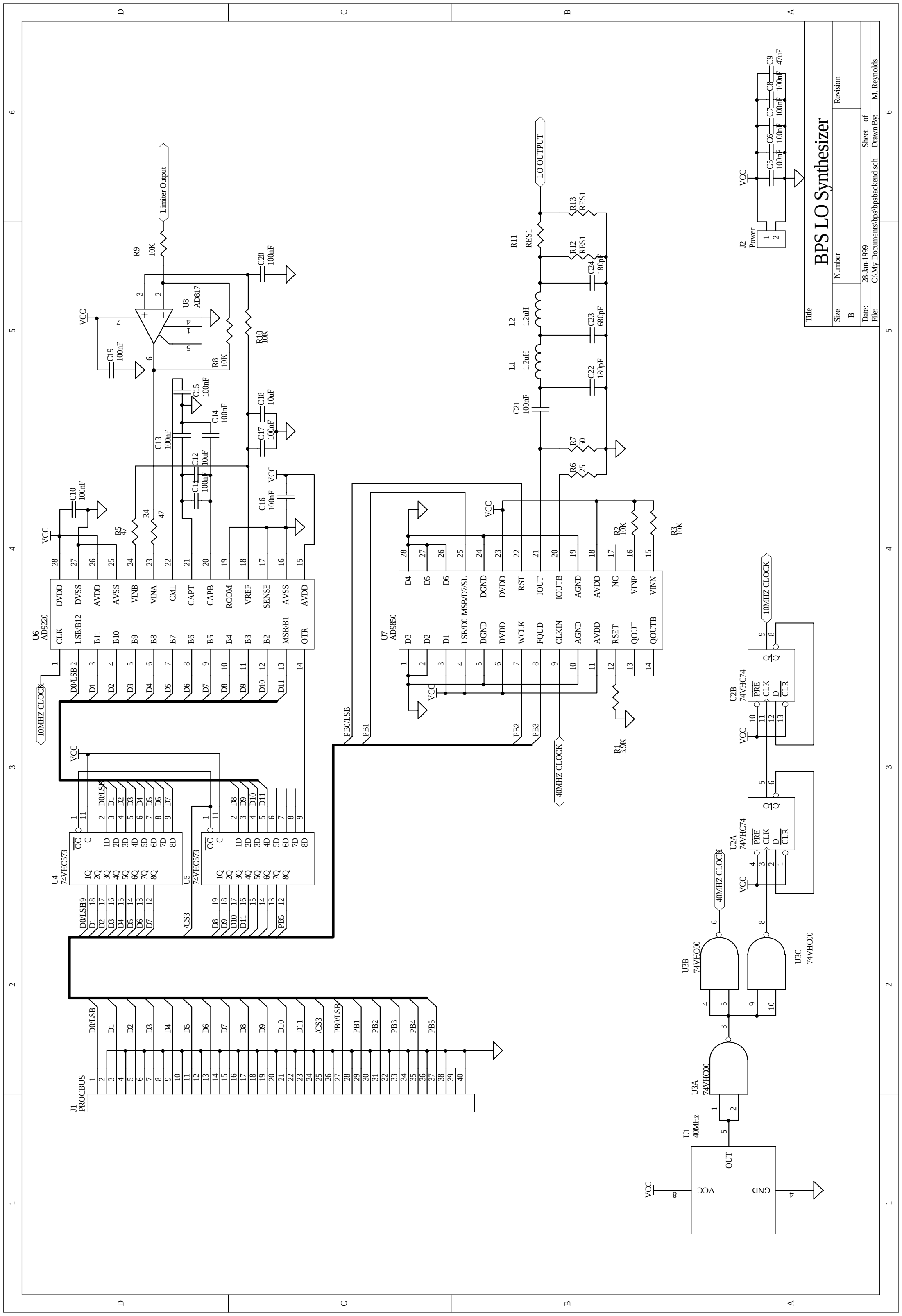
[Skl88] Sklar, Bernard. Digital Communications: Fundamentals and Applications. PTR Prentice Hall, 1988.

Appendix A: System Schematics

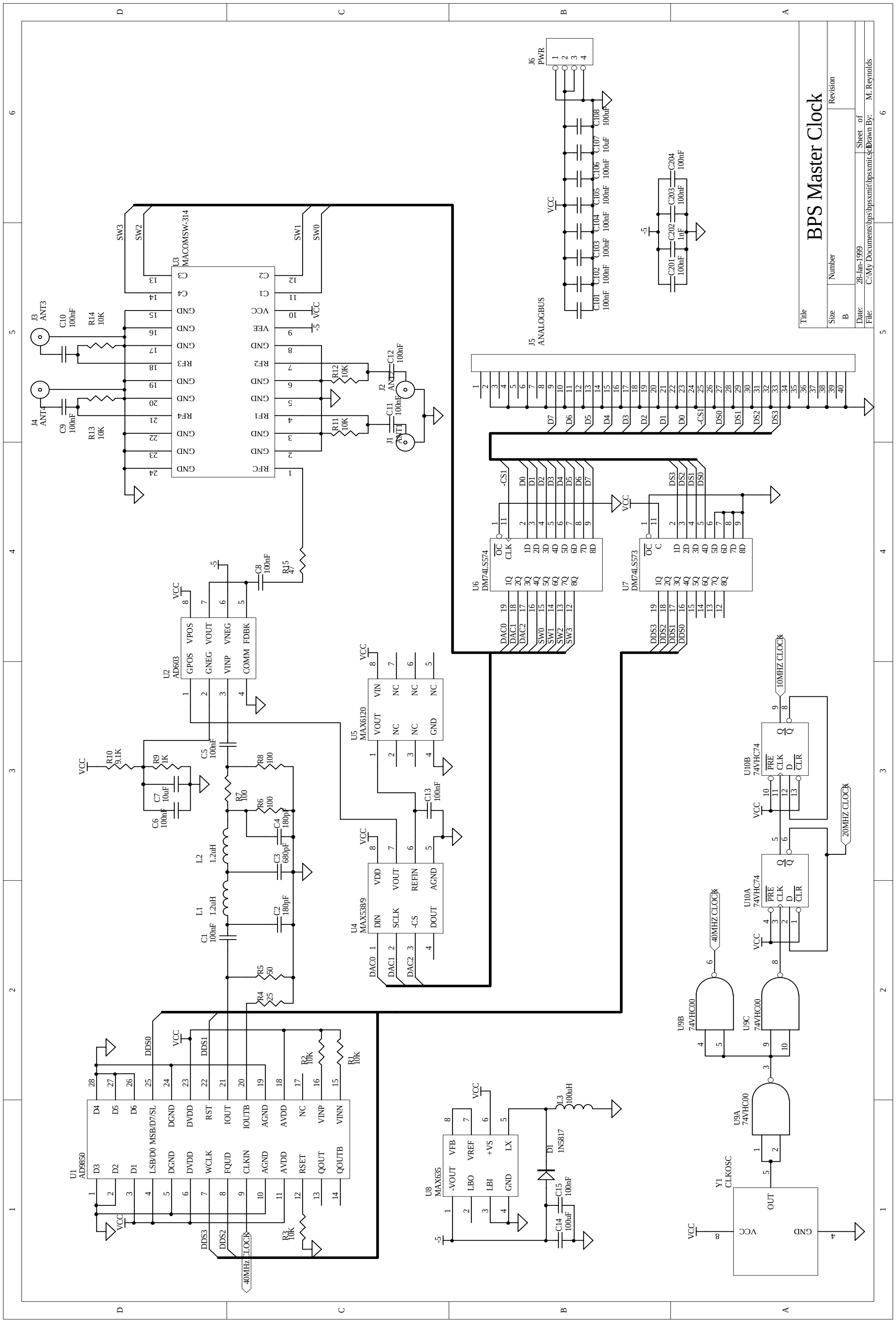
A-1: BPS Receiver Front End



A-2: BPS Receiver LO and ADC



A-3: BPS Transmitter Clock Unit



A-4: BPS CPU Board