Report 017392-1-T

SIGNAL PROCESSING EQUIPMENT AND TECHNIQUES FOR USE IN MEASURING OCEAN ACOUSTIC MULTIPATH STRUCTURES

K. Metzger, Jr.

COOLEY ELECTRONICS LABORATORY
Department of Electrical and Computer Engineering
The University of Michigan
Ann Arbor, Michigan 48109

December 1983

Technical Report No. 231

Prepared for
OFFICE OF NAVAL RESEARCH
Department of the Navy
Arlington, Virginia 22217

ACKNOWLEDGEMENTS

The Beamformer Project extended over a period of about 18 months and benefitted from the efforts of a number of individuals. Included below is a list of the participants and their contributions.

Ted Birdsall was the overall project director. In addition to these duties, he generated the sine and cosine tables used in the complex demodulator ROM's and also wrote the initial fast arithmetic device test program.

John Campbell provided support in processing the data acquired during the Tomography experiments (and later experiments) using the TI-980 and the associated color display system.

Linda Conolly taped the A/D converter bay back-plane, layed out, taped, assembled and tested the EPROM memory boards, assembled many printed circuit boards, wrote the production fast arithmetic device test program and wrote the *FORM text formatting program used to produce this document.

Al Davis participated in the development of the DM1200 prototype and checked the artwork for the CPU and the 12K memory printed circuit boards.

Hamed Eshraghian designed and built the isolation amplifier unit, designed and built the EPROM programmer (documented elsewhere) used to program the EPROM's used in the beamformer system and assisted with the ASM48 cross assembler used to program the Intel 8748 microcomputer.

Y. M. Hwang designed the card-cage used in the DM1200 processor boxes, layed out and taped the DM1200 back-plane, adapted the PDP-8/e floppy disc to run on the DM1200 and layed out and taped the communications interface printed

circuit card.

Olive Lockwood handled the placement of orders and kept track of all of the purchasing records. She also served as our resident grammarian.

Al Stoll designed and built the raster scan conversion box used to transfer the ADM-3A+ retrographics graphics display to printed form using a Paper Tiger printer. He also assembled printed circuit boards and made up cables.

Wayne Von Wald was responsible for mechanical assembly of the hardware, hand wrapped several wire-wrap panels on a panic time schedule and prepared final ink drawings of the circuit diagrams. He also prepared several of the figures used in this report.

Betty Willsher prepared the drawings used to describe the equipment to the sponsor and also prepared many of the figures used in this report.

* * *

I have had the good fortune of being able to work for and along with Ted Birdsall for over twenty years. His encouragement, support and friendship are deeply appreciated. I hope that we will have the opportunity to continue working together for many more years.

Walter Munk provided a superb role model in showing how to conduct research on the grand scale. It was largely thru his individual efforts that the Ocean Acoustic Tomography Demonstration Experiment came into being. I appreciate having been included in this effort and look forward to participating in the Pacific Gyre work.

Working with John Spiesberger has also been a rewarding experience. I look forward to working with him on the Pacific heat content experiment and whatever else that the future might hold.

Coralie Parkins and Carol Van Aken entered much of the text of this dissertation into our signal processing computer turned word processor. Both approached the task with enthusiasm and industry. Their efforts are greatly

appreciated.

Betty Willsher and Wayne Von Wald did superb jobs drafting the figures used in this dissertation and provided other much appreciated assistance.

Finally, I would like to thank my wife, Anita, for the patience, support, perseverence and good humor that she showed during the too many years that it took me to get to this point.

The work reported here was supported by the Office of Naval Research Code 486. Dr. J.M. McKisic served as the scientific officer.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF SYMBOLS AND ABBREVIATIONS

AG	Array gain
AM	Amplitude modulation
A_i	Amplitude of line in spectrum of csq(t)
a	Value used to weight sequence bit at transmitter
	Conic angle. Angle between x-axis and the normal to a plane wave
a_i	i-th bit in a binary sequence
a_s	Steered conic angle direction of beam pattern
a_{s0}	Reference steered conic angle direction
a_0	Conic angle smaller than a_s at which the beam pattern takes on a specified value
a_1	Conic angle greater than a_s at which the beam pattern takes on a spefied value
a_{00}	Conic angle smaller than a_{s0} at which the beam pattern takes on a specified value
a_{10}	Conic angle greater than a_{s0} at which the beam pattern takes on a specified value
a(t)	Amplitude modulating waveform
B	Number of bits
BM	Bi-phase modulation
BW	Bandwidth
b	Value used to bias sequence bit at transmitter
b(.)	beam pattern formed using an array of hydrophones

b_c	Specified beam pattern cut value
bw	Beam width
C	Number of bits in digital demodulator multiplier
C(k)	Normalized autocorrelation function of a weighted binary sequence
CM	Complementary phase modulation using 45 degrees
CPM	Complementary phase modulation
CR(k)	Cosine read only memory value
C_p	Estimated carrier power
c	Real part of a residue
	Value used to weight sequence bit at receiver
	Speed of sound in water
c_i	Coeffecient in the characteristic equation of a linear sequence
c_0	Constant used to characterize beam pattern at a specified response level
c_1	Constant used to characterize beam pattern at a specified response level
csq(t)	Complex valued square wave used for demodulation
D	Number of bits needed in digital demodulator accumulator
Dev	Standard deviation
DFT	Discrete Fourier transform
d	Imaginary part of a residue
	Value used to bias sequence bit at receiver
	Distance from plane wave to reference hydrophone
dB	Decibel
d_i	Distance i-th hydrophone to plane wavefront

f_r	Frequency in spectrum of reconstructed waveform
f_s	Sample frequency
f_u	Upper frequency at which a response is down by 3 dB
f_0	Reference frequency used in characterizing design frequency of an array of hydrophones
f_1	Method A correction factor
f_2	Method B correction factor
f_3	Method C correction factor
f_4	Method D correction factor
fract(.)	Fractional part of
H(f)	Filter transfer function
$H_c(f)$	Band-pass filter transfer function
$H_x(f)$	Transfer function of filter in complex demodulator in-phase output channel
$H_y(f)$	Transfer function of filter in complex demodulator quadrature output channel
$H_z(f)$	Baseband representation of band-pass filter transfer function
Hz	Hertz. Unit of frequency
h	Spacing between hydrophones in a uniformly spaced array
h(q)	Discrete replica cross-correlation value
$h_c(q)$	Corrected discrete replica cross-correlation value
Im	Imaginary part of
$J_0(.)$	Zero order Bessel function
j	Square root of -1. Also used as an index

K	White noise spectral level
	Signal amplitude in a demodulate value
K_b	RMS level of beamformer output
k_i	Read only memory address
L	Number of digits in a linear maximal sequence
M	Number of words in a read only memory
	Number of times a sequence digit is sampled
$M_c(f)$	Fourier transform of $m_c(t)$
$M_z(f)$	Fourier transform of $m_z(t)$
$m_c(t)$	Carrier borne band-pass waveform
$m_x(t)$	In-phase component of $m_c(t)$
$m_y(t)$	Quadrature phase component of $m_c(t)$
$m_z(t)$	$m_x(t)+jm_y(t)$
N	Noise level in overlapped spectral replications of sampled noise with spectral level N_o
	Number of hydrophones in an array
	Number of periods used in a circulating sum
N_d	Number of sample values summed in digital complex demodulator filter
N_p	Estimated noise power level
$N(.)$	Spatial noise distribution
NFLF	Non-flatness loss factor
NIF	Noise improvement factor of a hydrophone array
NIF_h	NIF assuming uniform horizontal noise
NIF_{hu}	NIF_h assuming a uniformly spaced array

NIF_s	NIF assuming spherically isotropic noise
NIF_{su}	NIF_s assuming a uniformly spaced array
N_k	Contribution of error energy in k-th spectral line and its replication to output noise level at multiplying dD/A converter output
N_Q	Additional spectral level due to replications of input noise level at demodulator output
N_o	White noise spectral density level
n(t)	Noise waveform
P(k)	Discrete Fourier transform of samples of p(t)
P_p	Estimated correlation peak power
p	Numerator of the ratio of two relatively prime integers
	Pole location in Laplace transform
p(t)	Periodic pulse waveform
Q	Ratio of bandpass center frequency to bandwidth
Q_{FA}	Actual Q of digital demodulator output filter
Q_{FN}	Nominal Q of digital demodulator output filter
Q_S	Signal Q
q	Denominator of the ratio of two relatively prime integers
$R(\tau)$	Autocorrelation function
R(k)	Discrete Fourier transform of samples of r(t)
RMS	Root mean square
Re	Real part of
R_p	Laplace transform residue value
r	Number of bits to scale by

$r(t)$	Received waveform
$r_n(t)$	Noisy reception
$S(f)$	Fourier transform of $s(t)$
$S_{AD}(f)$	Aliased spectrum at A/D converter output
SNRIF	Signal-to-noise ratio improvement factor
SR(k)	Sine read only memory value
$S_b(f)$	Fourier transform of resampled waveform
S_c	Sum of the squares of the alternating sign circulating sum
S_d	Sum of the squares of the demodulates contained in a data set
$S_d(f)$	Discrete Fourier transform of $s(kT_a)$
$S_{dp}(f)$	Spectral contribution of a single pole
$S_{dpp}(f)$	Spectral contribution of a pole pair
S_e	Estimate of noise level obtained using demodulate variance calculation
$S_{k'}$	Sum of replications landing at f_{al}
S_p	Estimated signal power level
$S_r(f)$	Fourier transform of $s_r(t)$
S_s	Sum of the squares of the circulating sum values
S/N	Signal to noise ratio
$s(t)$	Signal waveform
$s_b(t)$	Reconstructed signal waveform
$s_r(t)$	Reconstructed signal waveform
T	Digit duration

T_a	Reciprocal of f_a
T_b	Reciprocal of f_b
t	Continuous time
t_i	Delay time used with i-th hydrophone
V	Input voltage range maximum excursion
$V(f)$	Fourier transform of reference hydrophone output
$V_b(f)$	Fourier transform of beamformer outut
$V_{bq}(f)$	Fourier transform of the demodulate and delay beamformer time delay quantization error
V_c	Noise variance associated with a circulating sum value
V_d	Noise variance associated with a demodulate value
$V_{d,i}(f)$	Fourier transform of i-th demodulated hydrophone output
V_m	Input voltage maximum value at A/D converter input
V_n	Variance of noise waveform sample values
V_{nc}	Noise variance in correlator output value
V_q	Voltage step associated with one A/D quanta
V_r	RMS level to be set at A/D converter input
v	Noise variance
$v(t)$	Output voltage of reference hydrophone
$v_b(t)$	Beamformer output voltage
$v_{d,i(t)}$	Demodulator output voltage associated with i-th hydrophone
v_i	Weighting value

$v_i(t)$	Output voltage of i-th hydrophone
var	Variance
W	Bandwidth
W_i	Bandwidth of waveform to be demodulated
W_s	Signal bandwidth
$W_z(f)$	Spectrum at demodulator input
W(k)	Discrete Fourier transform of w_i
w_i	Weighting value
X(f)	Spectrum at cosine output of demodulator
$X_f(f)$	Filtered demodulator in-phase output
x	Real part of pole location
Y(f)	Spectrum at sine output of demodulator
$Y_f(f)$	Filtered demodulator quadrature output
y	Imaginary part of pole location
Z(f)	Spectrum of complex demodulator output
$\underline{\Phi}(f)$	Fourier transform of $\phi(t)$
ϕ	Elevation angle
ϕ_s	Steered elevation angle
$\phi(t)$	Reconstruction function
θ	Modulation angle used in forming CPM waveform
	Bearing angle
θ_s	Steered bearing angle
$\theta(t)$	Phase modulating waveform

θ_e	Angle related to beamformer delay time quantization error
τ	Delay variable in autocorrelation functions
*	Complex conjugation when used as a superscript

CHAPTER 1

INTRODUCTION

This work describes a set of digital signal processing techniques and associated hardware that were developed at the Cooley Electronics Laboratory (CEL) of The University of Michigan for use in measuring the multipath structure of acoustic propagation channels in the ocean. The design and construction of this hardware, its programming and its use to measure multipaths were done as part of CEL's support of the 1981 Ocean Acoustic Tomography Demonstration Experiment [1].

The 1981 Tomography experiment was a joint effort between the Institute of Geophysical and Planetary Physics of the Scripps Institute of Oceanography, Woods Hole Oceanographic Institution, the Department of Earth and Planetary Sciences of the Massachusetts Institute of Technology, the Cooley Electronics Laboratory of The University of Michigan and the Atlantic Oceanic and Meteorological Laboratory of the National Oceanographic and Atmospheric Administration (NOAA). The Office of Naval Research, the National Science Foundation and NOAA supported this experiment.

The purpose of the Tomography experiment was to use acoustic probe signals to determine the thermal structure of a volume of the ocean which, at the surface, corresponded to a square 300 km per side. This region was monitored over a period of 9 months and changes were observed and tracked. To do this, four self-contained acoustic sources were located along one side of the square and transmitted carrier

borne signals which were modulated using a binary pulse compression code. Located at strategic points within and on the borders of the square were a number of self-contained receivers. The role played by CEL was the monitoring of the transmissions at three distant shore sites. This was done in order to insure that the transmitters were operating properly and to acquire long baseline data with which to help evaluate the feasibility of doing Tomography over distances longer than those provided by the 300 km square.

The available shore sites were located some distance from the Tomography square. Because of this, it was expected that the transmissions would experience significant spreading loss and attenuation. The wavefront at a receiving site is nominally a plane wave. If this wavefront impinges on a collection of hydrophones it produces outputs which differ only by the time shifts due to the time it takes the plane wave to propagate between phones. By properly delaying the individual outputs and summing the delayed versions together the result is an increased sensitivity in the direction of the signal wavefront at the expense of wavefronts arriving from other directions. This process is called beam forming [2]. Since there were multiple hydrophones available at the sites that CEL was going to use there was the possibility of combining their outputs in order to enhance the reception.

Part of CEL's support effort consisted of the design and construction of four portable data acquistion and processing systems. The two main components were a time delay beamformer and a microprocessor based data manager (which was named the DM1200). This dissertation presents the theoretical bases on which design decisions were made, describes the process of beam forming and discusses the use of pulse compression waveforms. It also documents the operation of the resulting hardware and describes how to program it for use in making multipath measurements.

1.1 Background

The speed of sound in the ocean depends primarily on temperature, pressure and salinity. These three influences combine to yield a sound speed variation as a function of depth in deep water as illustrated in Fig. 1.1.

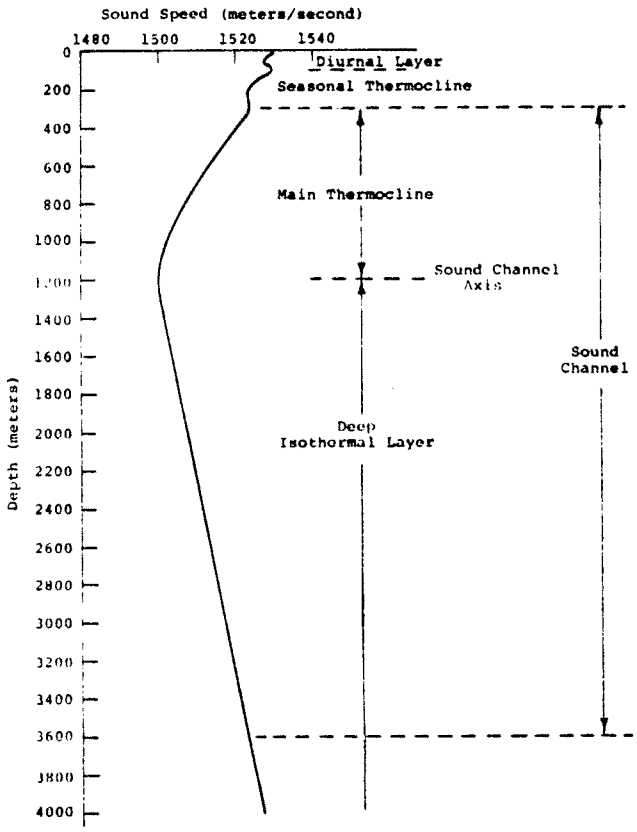


Fig. 1.1. Typical deep water sound speed profile

Near the surface there is a region that extends to about 100 meters which is strongly influenced by daily and seasonal weather. Below this surface region the effects of decreasing temperature and increasing pressure combine to cause the sound speed to decrease as depth is increased. Temperature changes below depths of about 1200 meters are small, the effects of pressure dominate causing the sound speed to increase with increasing depth. The location of

the sound speed minimum is called the sound channel axis. In the ocean a wavefront originating from a non-directional source nominally starts out spherical in shape. As it moves away from the source, portions of the wavefront begin to propagate at different velocities. The direction of propagation of any given point on the wavefront bends in the direction of lower sound speed. Eventually the wavefront tears apart and becomes a collection of non-connected patches. The concept of the ray can be borrowed from geometrical optics in order to aid in keeping track of how these patches propagate.

A ray is the path followed by a vector normal to the wavefront as the wavefront propagates. The movement of individual patches of the original wavefront can be kept track of using rays positioned at the center of each patch.

Given a sound source and a receiver located some distance away only some of these patches arrive at or in the vicinity of the receiver. Because these patches travel over different paths thru regions of varying sound speed their arrival times are also different. For a source and receiver both located on the sound axis the difference between arrival times for the earliest arrival and the latest arrival is on the order of 4.5 seconds per megameter of separation [3]. Figure 1.2 shows three possible paths between a transmitter and receiver located on the sound axis. Although the axial ray propagates over the shortest distance it moves thru water with the lowest value of sound speed and thus arrives last.

By combining multipath measurements with model predictions, it is possible to identify specific arrivals and measure their stability [4]. Arrival time perturbations associated with a given path are due to travel time variations along that path. Using the arrival time fluctuations, it is possible to localize perturbations in the sound speed profile [5]. Since most sound speed variations take place near the surface and are due to

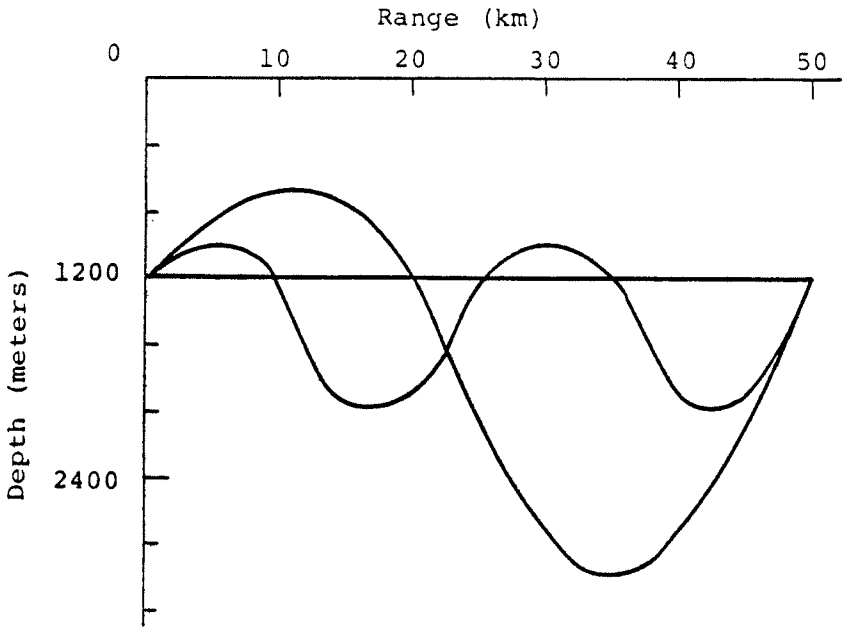


Fig. 1.2. Three possible ray paths for a transmitter and a receiver located on the sound axis

changes in temperature the measured sound speed variations can be related back to temperature differences. The Ocean Acoustic Tomography Demonstration Experiment of 1981 [1] demonstrated the feasibility of using acoustic multipath measurements to monitor changes in the thermal structure of the ocean. It may even be possible to make long term measurements over a large area in order to determine if there is a general long term heating or cooling of the ocean [6].

If the ocean were noise free it would be possible to measure the multipath arrival structure by transmitting a single short acoustic pulse and recording the reception. Similarly, if there were no restrictions on how much power a transmitter could put into the water, a pulse strong enough to overcome the background noise, spreading loss and absorption loss could be transmitted. However, because there is noise in the ocean and transducers are power limited, it is not generally possible to use isolated pulses over ranges greater than a few tens of kilometers. The

basic problem is one of having too low a signal level at the receiver to be useful.

If the isolated pulse is replaced by a periodic pulse and if time synchronization is maintained between the receiver and the transmitter then the signal-to-noise ratio can be built up by averaging together the arrivals due to several periods. The signal-to-noise ratio increases directly with the number of periods included in the average. The separation between pulses should be greater than the expected multipath spread. The limiting factor on the number of periods that can be safely averaged is the ability of the ocean to support coherent receptions over a period of time. How long the ocean remains coherent is a matter of some controversy and measurements are still being made.

The amount of energy per period can be increased thru the use of a pulse compression code which is active during the entire period and which can be processed to give the same response as a periodic pulse [7],[8]. Because the pulse compression code allows almost all of the signal energy to be collapsed back into a single pulse width, the result is an increase in the effective pulse amplitude. For a given required signal-to-noise ratio, fewer periods need to be summed. The cost of using a pulse compression code is an increase in the complexity of the transmitter and receiver. Of the two, the receiver processing is complicated the most but not unmanageably so.

Further enhancement of the reception can be had if more than one hydrophone is available for use. By properly combining the outputs of several hydrophones it is possible to spatially filter the noise field. The process of doing this is termed beam forming and the associated hardware is called a beamformer.

Fig. 1.3 is a plot of a multipath structure monitored over a 2 1/2 hour period in January 1983. This is the starting segment of a measurement that extended over an 86 hour period and shows the start of transmission. The

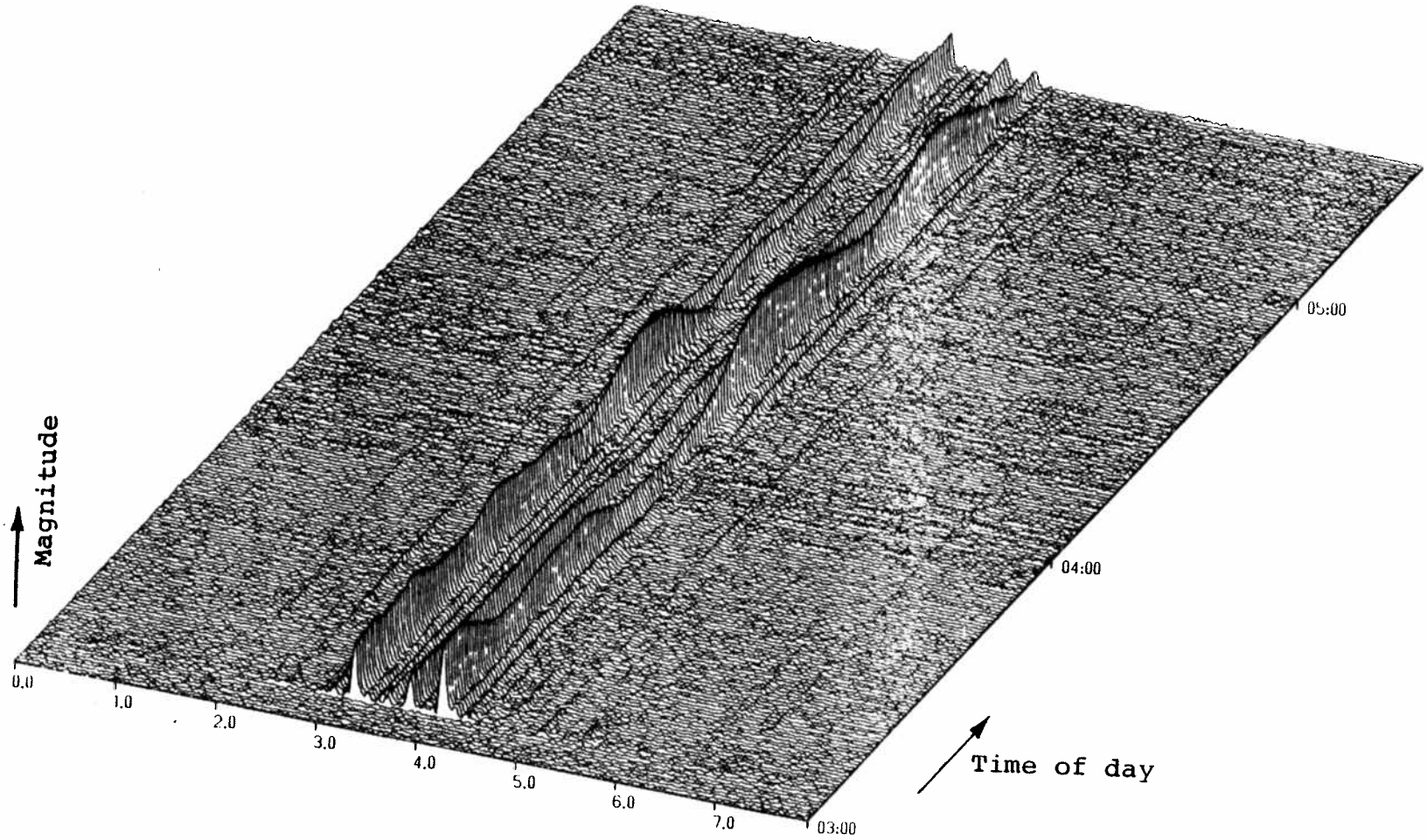


Fig. 1.3. Multipath structure observed over a 2 1/2 hour period in January 1983

received waveform was propagated over a distance of several hundred kilometers. The arrival structure consists of four main paths and several low level paths. The variability of the amplitudes of the main arrivals suggest that they may be due to closely spaced unresolved paths. The data for this plot was acquired using the equipment and techniques described in the following chapters. The plot was generated using the PERSYS plot program [9].

1.2 Overview

Chapter 2 deals with the processes of waveform sampling and reconstruction. Because the CEL beamformer is designed for simultaneous use with receptions at differing non-related carrier frequencies it is not possible to lock the sampling rate to the carrier of any particular reception. However, the signal processing steps applied to a particular reception are generally locked to that reception's carrier frequency. This makes it necessary to convert sample rates at some point in the processing. This is a common problem in general purpose digital signal processing systems and studies concerning sample rate conversion have appeared in the literature [10],[11]. In these studies the process of sample rate decrease by an integer factor of N is handled by first digital filtering to limit the waveform bandwidth and then taking every N-th filter output. The process of sample rate increase by a factor of M is handled by inserting M-1 zeros between the original sample values and then digital filtering. Sample rate conversion by a factor of M/N is then a combination of the two methods. This approach to sample rate conversion is well suited for use in a minicomputer. A hardware implementation would be a relatively large project in its own right.

The simplest approach to sample rate conversion is to use the most recent sample acquired at the old rate whenever a sample is required at the new rate. If the original

sample rate is sufficiently fast, there is the feeling that not too much damage is done. In Chapter 2 the resampling processes is modeled as one of waveform reconstruction followed by a new sampling. The method used to interpolate sample values defines a waveform reconstruction function. This function serves to attenuate the high frequency spectral replications present in the sampled waveform prior to reconstruction. The sampling at the new rate then aliases these attenuated replications. The amount of error introduced by these aliases depends on their magnitudes (controlled by the interpolation rule) and their locations (determined by the two sample rates). Procedures for determining both of these parameters are given. Emphasis is placed on the zero order hold which models the process of taking the most recent sample. This approach to studying the process of resampling appears to be original.

Chapter 3 reviews the process of complex demodulation as a method of shifting a carrier based waveform to baseband while preserving phase information. The use of the complex exponential $\exp\{-j2\pi t\}$ to accomplish this is a classical textbook technique. Borrowing a waveform synthesis method used in frequency synthesizers [12] to generate sample values of the complex exponential, a digital complex demodulator is described and analyses of the effects of time and amplitude quantization are made.

Less general methods of complex demodulation, requiring less sophisticated hardware, are the use of complex valued square waves and "$4f_0$" or quadrature sampling which are also described in Chapter 4. The first use of square waves in a complex demodulator was made by T. G. Birdsall in the early 1960's [13]. When signal processing moved onto the computer this method was replaced by the so-called $4f_0$ method of demodulation by direct sampling [13]. Later workers also published on this technique [14],[15].

Chapter 4 describes the basic properties of linear maximal sequences relevant to their use as the pulse

compression signals in making multipath measurements. Once the received waveform has been demodulated and perhaps had several periods averaged together it is necessary to remove the effects of the sequence induced time spreading. This can be done in a number of ways. Several of these are outlined and their equivalence established. Formulas for the required carrier correction factors are developed. Equations are also presented which allow the signal-to-noise ratio in the processed reception to be related to that of the original reception.

The history of linear maximal sequences is long and varied. Much of the early research was done at CEL [16] and the use of linear maximal sequences has become a part of the local heritage. Other sources of information about these sequences can be found in [17],[18],[19]. Linear maximal sequences appear to have first been used by Stewart and Allen [7] in studying propagation in the ocean. These waveforms were used in the first propagation tests undertaken by CEL in 1963 [13] and have been used in almost every propagation study that CEL has been involved in since then. A number of special purpose signal generators have been designed and built at CEL for use in these experiments [20],[21],[22]. Other researchers have also made use of the technique [23],[24].

Beam forming, the subject of Chapter 5, is a spatial filtering technique where the outputs of several hydrophones are combined in order to discriminate against noise arriving from directions other than the direction of the desired signal. Several methods of implementation are possible. The most traditional method uses time delays to cause the receptions from several hydrophones to be brought into phase with each other. In a continuous time system the output of a time delay beamformer is independent of frequency for arrivals from the steered direction. Quantizing the delay times affects the spatial sensitivity pattern [25] and the beamformer frequency response. The spatial effects of

quantization show up as increased sensitivity (side lobes) in directions other than the one of interest. These effects are generally small and only of concern in a finely tuned system. The CEL beamformer is intended for use with hydrophones of opportunity. It is unlikely that the hydrophone spacings are proper for the frequencies used; the result of which can be the appearance of very large side lobes.

Because the CEL beamformer is to be used with broad band signals the effects of delay time quantization on the beamformer frequency response is of concern. This does not seem to have received consideration in the literature. It is shown that the effects of these quantization errors can be modeled as a low-pass filtering operation. Once a maximum allowable amount of distortion over a frequency band has been settled on the maximum time delay step size can be determined.

Time delay beam forming can also be implemented by delaying the demodulated hydrophone outputs. In the case of narrow-band receptions the required delays can be approximated using phase shifts. Both of the above methods are also described.

Chapter 6 contains a description of the hardware designed and built at CEL for use in the 1981 Tomography experiment. Work on this equipment started in late fall of 1979 and extended into the spring of 1981. Four systems were built each consisting of a signal processing unit and a beamformer. Three of these systems were in field use from March 1981 to October 1981. The fourth system was not placed in the field because of problems at the intended site. Two of the processors were re-configured in the summer of 1982 with a 9-track magnetic tape system and a cartridge tape system being added. This equipment was used in October of 1982 and January of 1983 to acquire data for use in studying the problems associated with transmitting data from a self-contained remote transmitter/receiver to a

shore site. Additional data will be taken in the fall of 1983. If funding is available, one of the beam forming systems will be used to investigate the feasibilty of using multipath arrival time variations over long distances to infer the mean ocean termperature.

Basic to both the signal processor unit and the beamformer is a set of digital logic boards built around the Intersil/Harris IM6100 microcomputer. The IM6100 is an integrated circuit version of the Digital Equipment Corporation PDP-8/e minicomputer. This minicomputer has a 12 bit word size and can address up to 32K words of memory. The main difference between the IM6100 and the PDP-8/e is that the IM6100 runs about one-third as fast. Harris Inc. recently started selling a version that is twice as fast as the original IM6100. Although the Harris unit supports a slightly different bus structure it should be possible to adapt this device for use in the CEL systems.

The IM6100 was chosen because of its compatibility with CEL's exisiting computer systems which include a LINC-8, two PDP-8/e's and a Fabri-Tek MP12. Over the years since the first LINC-8 arrived at CEL in May 1967, a significant amount of software had been developed which could be transferred onto the microcomputer. This includes a program development system written by the author as well as signal processing programs.

The first part of Chapter 6 describes the processor board set and the special devices and interfaces that were included in the processing unit. The second part of Chapter 6 describes the beamformer unit and the special interfaces specific to it.

Chapter 7 describes the signal processing and software techniques which have been developed at CEL for processing multipath receptions. The methods described have evolved over a period of many years. The signal processing techniques are largely due to T. G. Birdsall and the programming procedures are largely due to the author.

CHAPTER 2

SAMPLING, RECONSTRUCTION AND RESAMPLING

Having sampled a waveform at some given rate it is often necessary to generate a new set of samples which correspond to a different sample rate. When the new rate is lower than the original rate the resampling is termed decimation. When the new rate is higher, the process is termed interpolation [11]. The process of resampling is important in systems such as the beamformer described in Chapter 6 where it is necessary to operate portions of the system at different data rates.

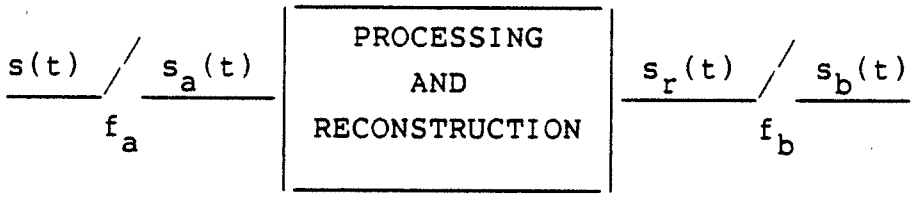


Fig. 2.1. Model for sampling, reconstruction and resampling

The process of sampling, processing and resampling can be modeled as shown in Fig. 2.1. A continuous waveform s(t) is sampled at a uniform rate f_a. The samples are processed in some manner. The processed sample values are then used in conjunction with some waveform reconstruction algorithm to define a new continuous waveform $s_r(t)$. The reconstructed waveform is then resampled at the new rate f_b.

The initial sampling causes a periodic replication of the spectrum of s(t) to take place. The replicas are separated by f_a and add together. If attention is

restricted to the fundamental frequency band $-f_a/2$ to $+f_a/2$ it appears as if the entire spectrum of s(t) has been in some way collapsed into this band. The addition of energy from outside of this band to the energy present in the band prior to sampling is termed aliasing and represents the primary source of distortion caused by sampling. The amount of aliasing can be controlled by shaping the bandwidth of s(t) and by using a high sample rate relative to the bandwidth. In order to change data rates some procedure or algorithm must be used to define sample values at the new sample times in terms of the samples taken at rate f_a. The use of such an algorithm is termed reconstruction. In practice it is not necessary to actually physically generate the corresponding continuous time waveform in order for the analysis contained in this chapter to apply.

The use of a reconstruction algorithm allows the definition of a continuous time function possessing a "normal" spectrum. For the method of reconstruction used in this chapter this spectrum is the product of the replicated sample spectrum and the Fourier transform of the basic reconstruction function. Resampling at rate f_b causes a periodic replication of this spectrum. Given energy at some frequency in the spectrum of s(t) the locations (landing frequencies) and magnitudes of the contributions of this energy in the spectrum of the resampled waveform are readily determined.

Section 2.1 relates the spectrum of a continuous waveform to the discrete spectrum of the sampled version. This material is standard and can be found in texts such as [26]. Section 2.2 builds on this material and shows how the spectra of reconstructed waveforms compare to the original unsampled spectra. Section 2.3 uses these results to determine the effects of the original sampling and the choice of reconstruction function when the resulting waveform is again sampled.

Section 2.4 considers the effects of two error sources

on the spectrum of a sampled waveform. These sources are amplitude quantization and overlap between spectral replications. Amplitude quantization is caused when sample values are represented using a finite number of bits to correspond to a fixed pre-determined voltage range. This noise is commonly represented as being spectrally white with a level related to the quantization step size. The effect of spectral overlap is somewhat more difficult to model and depends on the spectrum of the waveform being sampled. The approach taken in Section 2.4 is to model the unsampled waveform as having been generated by passing stationary white Gaussian noise thru an analog filter. This is referred to as the Wiener noise model [27]. The power spectrum of the sampled waveform can be expressed in terms of the pole locations and the associated residues of the filter transfer function.

The author's contributions are:

determination of the landing frequencies in the spectum of the resampled waveform of the spectral replications caused by the initial sampling

determination of the amplitudes of the re-aliased replications

derivation of the expressions used to determine the power spectrum of the sampled Weiner-model noise processes

This information is of importance to the system designer because it allows the severity of the problems associated with sample rate changes to be quantified. The noise analysis allows the interaction between sample rate, the presampling filter and the amount of noise folded into the signal band to be computed.

2.1 Effects of Sampling on the Spectrum

Given a continuous waveform s(t) its Fourier transform

is defined as

$$S(f) = \int_{-\infty}^{+\infty} s(t)\ \exp\{-j2\pi ft\}dt \qquad (2.1)$$

The associated inverse transform is

$$s(t) = \int_{-\infty}^{+\infty} S(f)\ \exp\{+j2\pi ft\}df \qquad (2.2)$$

Given a sequence of sample values of s(t) at sample rate $f_a=\frac{1}{T_a}$ the discrete Fourier transform is

$$S_d(f) = \sum_{k=-\infty}^{\infty} s(kT_a)\ \exp\{-j2\pi fkT_a\} \qquad (2.3)$$

The associated inverse transform is

$$s(kT_a) = T_a \int_{\frac{-1}{2T_a}}^{\frac{+1}{2T_a}} S_d(f)\ \exp\{j2\pi fkT_a\}df \qquad (2.4)$$

The continuous and discrete spectra can be related. Using Eq. 2.2 letting $t=kT_a$ gives

$$s(kT_a) = \int_{-\infty}^{+\infty} S(f)\ \exp\{j2\pi fkT_a\}df \qquad (2.5)$$

This can be written as

$$s(kT_a) = \sum_{m=-\infty}^{\infty} \int_{\frac{m-1/2}{T_a}}^{\frac{m+1/2}{T_a}} S(f)\ \exp\{j2\pi fkT_a\}df \qquad (2.6)$$

Let $f' = f - \frac{m}{T_a}$, substitute and then drop the primes. This gives

$$s(kT_a) = \sum_{m=-\infty}^{\infty} \int_{\frac{-1}{2T_a}}^{\frac{+1}{2T_a}} S(f + \tfrac{m}{T_a})\ \exp\{j2\pi(f + \tfrac{m}{T_a})kT_a\}df \qquad (2.7)$$

Noting that $\exp\{j2\pi km\} = 1$ and interchanging the order of summation and integration

$$s(kT_a) = \int_{\frac{-1}{2T_a}}^{\frac{+1}{2T_a}} \sum_{m=-\infty}^{\infty} S(f + \tfrac{m}{T_a})\ \exp\{j2\pi fkT_a\}df \qquad (2.8)$$

But this is in the same form as Eq. 2.4. Thus

$$S_d(f) = \frac{1}{T_a} \sum_{m=-\infty}^{\infty} S(f + \tfrac{m}{T_a}) \qquad (2.9)$$

The spectrum of the sampled waveform is formed by shifting versions of the continuous spectrum by amounts that are multiples of the sample rate and then summing the shifted versions together. This process is illustrated in Fig. 2.2 a and b. The resulting spectrum is periodic. The amount of overlap between replications determines the amount

of distortion introduced by sampling. In Fig. 2.2 this overlap is shown in idealized form. In practice, the spectrum of the unsampled waveform falls off at some rate and there may be significant overlap between replicas spaced several sample frequency multiples apart. The amount of overlap between replicas can be reduced by increasing the sample rate and/or by filtering the waveform prior to sampling in order to decrease the amount of energy outside of the frequency band $\frac{-1}{2T_a}$ to $\frac{+1}{2T_a}$.

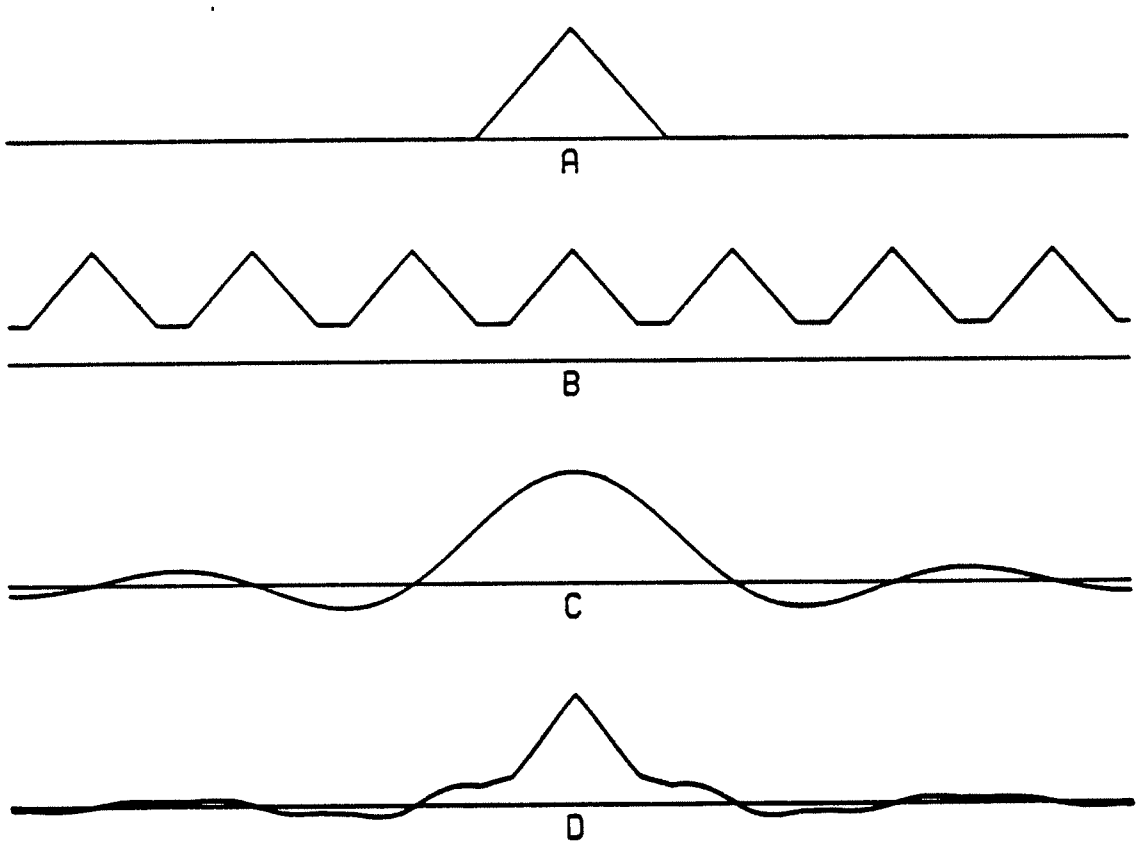


Fig. 2.2. Spectral effects of sampling and reconstruction a) Spectrum of the unsampled waveform b) Replicated spectrum after sampling c) Transfer function of the reconstruction filter d) Spectrum of reconstructed waveform

2.2 Reconstruction

In order to convert a digital sequence into analog form or to be able to change data rates, it is useful to understand the relationship between the spectrum of the sequence and the spectrum of the reconstructed continuous time waveform.

Attention is restricted to simple linear weightings of the form

$$s_r(t) = \sum_{k=-\infty}^{\infty} s(kT_a)\ \phi_k(t) \tag{2.10}$$

The sample values are used to weight a series of continuous time reconstruction functions $\phi_k(t)$. When there is no reason to weight one sample differently than any other sample the $\phi_k(t)$ can be taken to be time translations of the same basic reconstruction function $\phi(t)$. Then

$$\phi_k(t) = \phi(t - kT_a) \tag{2.11}$$

The reconstruction equation becomes

$$s_r(t) = \sum_{k=-\infty}^{\infty} s(kT_a)\ \phi(t - kT_a) \tag{2.12}$$

Taking the Fourier transform gives

$$S_r(f) = \Phi(f) \sum_{k=-\infty}^{\infty} s(kT_a)\ \exp\{-j2\pi fkT_a\} \tag{2.13}$$

$$= \Phi(f)\ S_d(f) \tag{2.14}$$

$$= \frac{1}{T_a}\ \Phi(f) \sum_{k=-\infty}^{\infty} S(f+k/T_a) \tag{2.15}$$

The spectrum of the reconstructed waveform is the

product of the discrete spectrum of the sampled waveform and the Fourier transform of the reconstruction function. Fig. 2.2 b,c and d show how the spectrum of the sampled signal and the transform of the reconstruction function combine to give the spectrum of the reconstructed waveform. Simple hold reconstruction (described below) is used.

Desirable properties of a reconstruction function include:

Adequate suppression of spectral replicas
Negligible effect on the baseband portion of the spectrum
Finite time extent
Ease of implementation

There is no need to actually physically reconstruct a continuous time waveform. The use of a reconstruction function simply provides an algorithm which uniquely defines a waveform value at any time instant at which a new sample is to be generated.

Examples of commonly encountered reconstruction procedures are the sin(x)/x function used in the Shannon sampling theorem, the simple holding of a sample value until a new value is present and linear interpolation.

The reconstruction function used in the Shannon sampling theorem is

$$\phi(t) = \frac{\sin(\pi t/T_a)}{(\pi t/T_a)} \tag{2.16}$$

$$\Phi(f) = T_a \qquad \frac{-1}{2T_a} \leq f \leq \frac{+1}{2T_a} \tag{2.17}$$

$$= 0 \qquad \text{elsewhere}$$

The simple hold reconstruction function amounts to holding the value of a sample until a new value is available.

$$\phi(t) = 1 \qquad \frac{-T_a}{2} \leq t \leq \frac{T_a}{2} \tag{2.18}$$

$$= 0 \qquad \text{elsewhere}$$

$$\underline{\Phi}(f) = \frac{\sin(\pi f T_a)}{\pi f} \tag{2.19}$$

Linear interpolation

$$\phi(t) = \frac{(T_a+t)}{T_a} \qquad -T_a \leq t \leq 0 \tag{2.20}$$

$$= \frac{(T_a-t)}{T_a} \qquad 0 \leq t \leq T_a$$

$$= 0 \qquad \text{elsewhere}$$

$$\underline{\Phi}(f) = \frac{\sin^2(\pi f T_a)}{T_a(\pi f)^2} \tag{2.21}$$

Other possible interpolation functions include the Lagrange interpolation polynomials and the filtering functions used in References [10] and [28]. Reconstruction functions which extend over several samples are usually constrained to pass thru the sample values. It can be shown that the optimum minimum mean square error reconstruction function does pass thru the sample values [28]. This does not mean that there are not good reconstruction functions which do not pass thru the sample values. The simple hold and the linear interpolation functions given above pass thru the sample points and have transforms of the form $\sin(x)/x$ and $(\sin(x)/x)^2$ respectively. The reconstruction function corresponding to the third power of $\sin(x)/x$ is

$$\begin{cases} 0 & t \leq \frac{-3T_a}{2} \\ \frac{1}{2T_a^2}(t+\frac{3T_a}{2})^2 & -\frac{3T_a}{2} \leq t \leq -\frac{T_a}{2} \\ \frac{1}{T_a^2}(\frac{3T_a^2}{4} - t^2) & -\frac{T_a}{2} \leq t \leq \frac{T_a}{2} \\ \frac{1}{2T_a^2}(t-\frac{3T_a}{2})^2 & \frac{T_a}{2} \leq t \leq \frac{3T_a}{2} \\ 0 & \frac{3T_a}{2} < t \end{cases} \tag{2.22}$$

The associated transform is

$$\underline{\Phi}(f) = \frac{\sin^3(\pi f T_a)}{T_a^2(\pi f)^3} \tag{2.23}$$

The waveform generated using this reconstruction function does not pass thru the original sample points. However, this function does heavily attenuate the replica components located at multiples of $f=1/T_a$ while acting as a filter on the baseband components.

In general, one wants a computationally simple reconstruction function which uses only a few sample values, heavily attenuates the replicas and does not filter the baseband energy more than some desired amount. Fig. 2.3 illustrates how various constraints might be placed on the reconstruction function in the frequency domain.

The frequency axis is divided into pass, transition and attenuation zones. In the pass zone the frequency response is constrained to deviate no more than δ_p from some nominal value. In the attenuation zone the frequency reponse is

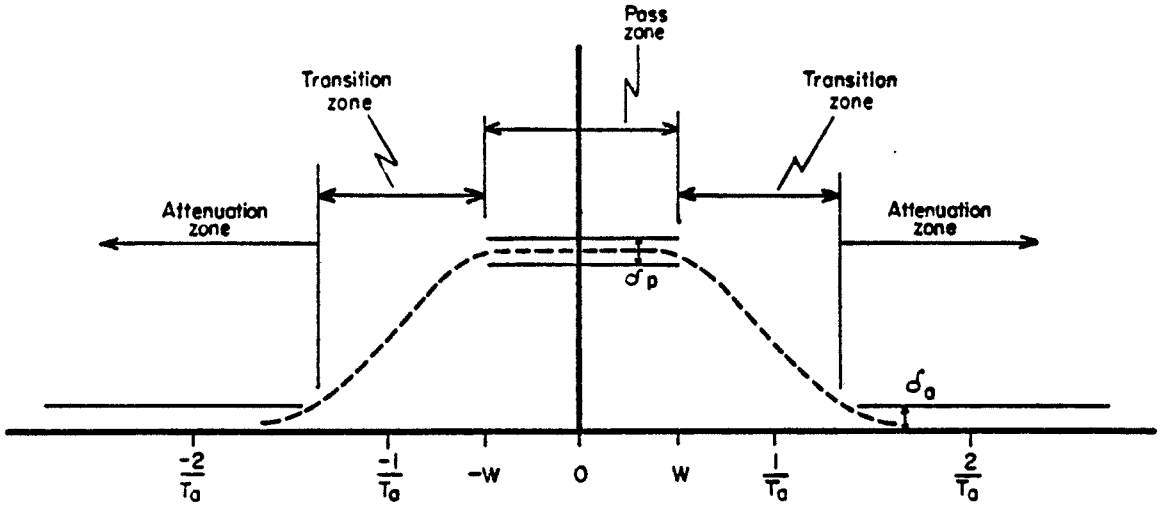


Fig. 2.3. Constraints which might be placed on the Fourier transform of the reconstruction function

constrained to be no larger than δ_a. The frequency response changes from pass to attenuate in the transition zone. The faster that this transition is required to be, the more complicated will be the required interpolation process.

2.3 Resampling

The process of sampling the reconstructed waveform at rate $1/T_b$ causes a new aliasing to occur. The folding frequency is $1/(2T_b)$ and the baseband frequency range is

$$\frac{-1}{2T_b} \leq f \leq \frac{+1}{2T_b} \qquad (2.24)$$

The folding of the components located at multiples of $f=1/T_a$ in the reconstructed waveform causes distortion in the resampled waveform.

Since

$$s_r(t) = \sum_{k=-\infty}^{\infty} s(kT_a)\phi(t-kT_a) \tag{2.25}$$

the sample values of the resampled waveform are

$$s_b(mT_b) = \sum_{k=-\infty}^{\infty} s(kT_a)\phi(mT_b-kT_a) \tag{2.26}$$

The spectrum of the resampled waveform can be expressed in several ways.

$$S_b(f) = \sum_{m=-\infty}^{\infty} \exp(-j2\pi fmT_b) \sum_{k=-\infty}^{\infty} s(kT_a)\phi(mT_b-kT_a) \tag{2.27}$$

$$= \sum_{m=-\infty}^{\infty} \Phi(f+\frac{m}{T_b})S_d(f+\frac{m}{T_b}) \tag{2.28}$$

$$= \sum_{m=-\infty}^{\infty} \Phi(f+\frac{m}{T_b}) \frac{1}{T_a} \sum_{k=-\infty}^{\infty} S(f+\frac{k}{T_a}+\frac{m}{T_b}) \tag{2.29}$$

Attention is restricted to the case of sample rate decrease, $T_b > T_a$.

Given a specific frequency f_r in the spectrum of a reconstructed waveform, the frequency that it maps into in the baseband portion of the resampled spectrum is found by determining the value of m such that

$$\frac{-1}{2T_b} \leq f_r + \frac{m}{T_b} < \frac{1}{2T_b} \tag{2.30}$$

Solving gives

$$m = -[f_rT_b + \frac{1}{2}] \tag{2.31}$$

where [ ] denotes taking the greatest integer of the

enclosed quantity.

The frequency that f_r aliases to is

$$f_{al} = f_r - \frac{1}{T_b}[f_r T_b + \frac{1}{2}] \tag{2.32}$$

Consider the energy in the original unsampled signal at frequency f_c ($|f_c| < \frac{1}{2T_b}$). This frequency gets replicated in the reconstructed waveform to frequencies $\frac{k}{T_a} + f_c$. For a given value of k the associated landing frequency in the spectrum of the resampled waveform is

$$f_{al} = \frac{1}{T_b}\{\frac{kT_b}{T_a} + f_c T_b - [\frac{kT_b}{T_a} + f_c T_b + \frac{1}{2}]\} \tag{2.33}$$

In many applications, such as the system described in Chapter 6, the sample and resample rates are related as the ratio of two integers. Let

$$\frac{T_b}{T_a} = \frac{p}{q} \tag{2.34}$$

Since only the case of sample rate reduction is being considered, $p>q$. It is also assumed that any factors common to both T_a and T_b have been removed so that p and q are relatively prime. Under these conditions

$$f_{al} = \frac{1}{T_b}\{\frac{kp}{q} + f_c T_b - [\frac{kp}{q} + f_c T_b + \frac{1}{2}]\} \tag{2.35}$$

Write

$$k = k' + qr \qquad \begin{matrix} 0 \le k' < q \\ -\infty < r < \infty \end{matrix} \tag{2.36}$$

Substituting gives

$$f_{al} = \frac{1}{T_b}\{\frac{k'p}{q}+f_cT_b-[\frac{k'p}{q}+f_cT_b+\frac{1}{2}]\} \qquad (2.37)$$

This equation is independent of r. There are only q landing frequencies in the baseband portion of the spectrum of the resampled waveform which correspond to a given f_c in the original unsampled waveform and its replications in the reconstructed version. The spacing between landing frequencies is $\frac{1}{qT_b}$.

The amplitude of a line located at frequency $\frac{k}{T_a}+f_c$ in the reconstructed spectrum is

$$\frac{1}{T_a}\Phi(\frac{k}{T_a}+f_c) \qquad (2.38)$$

The sum of the replications landing at f_{al} is then

$$S_{k'} = \sum_{r=-\infty}^{\infty} \frac{1}{T_a}\Phi(\frac{k'}{T_a}+\frac{rq}{T_a}+f_c) \qquad (2.39)$$

Because the simple hold is by far the most commonly used reconstruction function it is worthwhile to evaluate Eq. 2.39 for this case.

$$S_{k'} = \sum_{r=-\infty}^{\infty} \frac{\sin(\pi\{\frac{k'}{T_a}+\frac{rq}{T_a}+f_c\}T_a)}{\pi T_a\{\frac{k'}{T_a}+\frac{rq}{T_a}+f_c\}} \qquad (2.40)$$

$$= \frac{(-1)^{k'}\sin(\pi f_cT_a)}{q}\sum_{r=-\infty}^{\infty}\frac{(-1)^{rq}}{\pi r+\pi(k'+f_cT_a)/q} \qquad (2.41)$$

The series can be summed using Equations 4.3.91 and 4.3.93 from [29].

$$S_{k'} = \frac{(-1)^{k'}\sin(\pi f_c T_a)}{q\tan(\pi(k'+f_c T_a)/q)} \qquad \text{(q even)} \qquad (2.42)$$

$$= \frac{(-1)^{k'}\sin(\pi f_c T_a)}{q\sin(\pi(k'+f_c T_a)/q)} \qquad \text{(q odd)} \qquad (2.43)$$

Substituting $T_a = qT_b/p$ gives

$$S_{k'} = \frac{(-1)^{k'}\sin(\pi\frac{f_c q}{f_b p})}{q\tan(\pi(\frac{k'}{q}+\frac{f_c}{f_b p}))} \qquad \text{(q even)} \qquad (2.44)$$

$$= \frac{(-1)^{k'}\sin(\pi\frac{f_c q}{f_b p})}{q\sin(\pi(\frac{k'}{q}+\frac{f_c}{f_b p}))} \qquad \text{(q odd)} \qquad (2.45)$$

For the Ocean Acoustic Tomography Experiment the values f_a and f_b were 5000 Hz and 64 Hz respectively. The resulting values of p and q were 625 and 8. The amplitudes of the spurious responses for values of k' equal to 1, 2, 3, 5, 6 and 7 are plotted Fig. 2.4 as a function of $f_c T_b$. The value of $S_{k'}$ for k=0 was very close to unity. The value for k'=4 was less than 10^{-5} and was not plotted. The maximum spurious value occurs for $|fT_b| = \frac{1}{2}$ and k' equal to 1 and 7. The amplitude of this spurious signal is down 44 dB in power from any intended signal at the original frequency.

A similar analysis for f_a of 5000 Hz and f_b of 400 Hz gives p=25 and q=2 and a maximum spurious contribution of 4×10^{-3} which is down 48 dB in power.

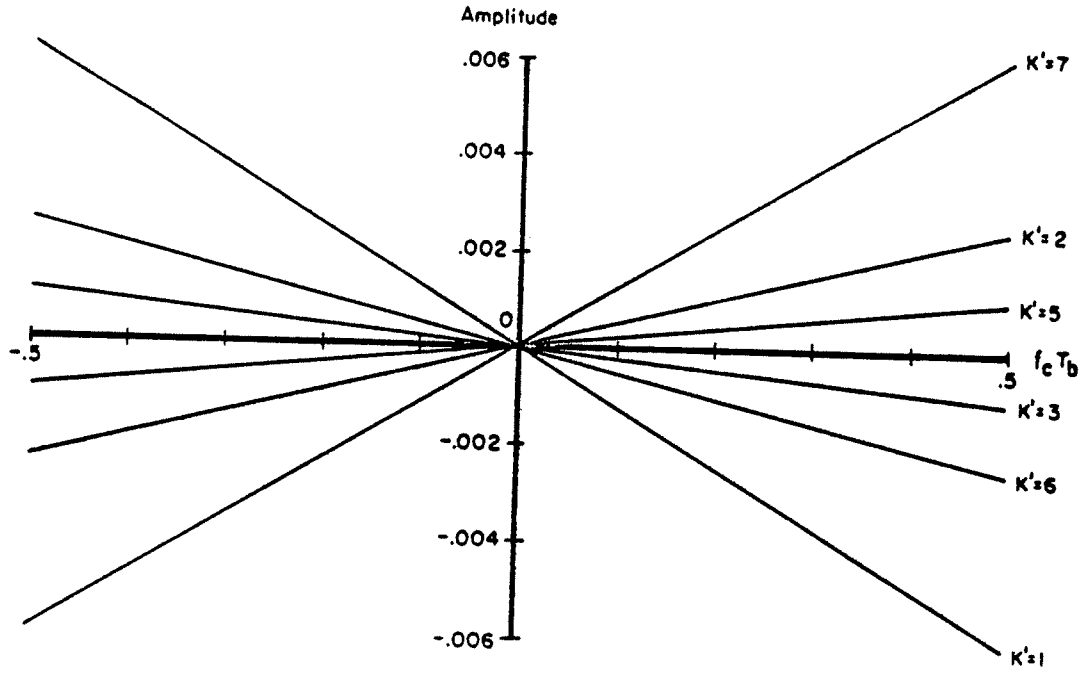


Fig. 2.4. Maximum spurious response levels using Ocean Acoustic Tomography Experiment parameters

In reality, the processes of reconstruction and resampling are equivalent to the application of an interpolation rule. The function of the rule (reconstruction function) is to attenuate the energy contained in the replications centered at multiples of the original sample rate. If the waveform bandwidth is narrow compared to the original sample rate, the demands on the interpolation rule are minimal and even the simple hold function can be quite effective. As the waveform bandwidth becomes larger, the more complex the interpolation rule must become for a given level of performance.

If the two sample rates are related as the ratio of two relatively prime integers then the landing frequencies of the attenuated replicas upon resampling can be determined. Knowing the interpolation rule, the magnitudes of these spurious signal components can be estimated and the expected performace determined.

2.4 Errors Caused by Quantization and Noise Folding

The two major sources of error involved in converting analog signals into digital values are the quantization of sample values into integer values and the spectral aliasing caused by sampling. If a waveform is digitized using a very high sample rate relative to its frequency content then the quantization of the amplitude into discrete values will likely be the main source of error. Similarly, for a waveform sampled at a rate close to its bandwith but using a high resolution A/D converter, the main error source will be due to spectral aliasing.

The noise due to quantizing samples to a given resolution is usually modeled as having a uniform probability density with zero mean, being independent of the waveform and being independent from sample to sample [28]. In practice this is a reasonable model when using A/D converters with 6 or more bits of resolution and the signal spends a small amount of time compared to the inter-sample time in a given quanta step. Given a B bit converter which is used to digitize the voltage range $\pm V$ the variance of the quantization noise is

$$\text{var} = \frac{(\frac{2V}{2^B})^2}{12} = \frac{V^2}{3 \cdot 2^{2B}} \qquad (2.46)$$

Quantization noise can be considered to be an additive background noise with white spectrum in the band $-f_a/2$ to $+f_a/2$ and power spectral density level equal to the variance given in Eq. 2.46 divided by f_a.

The noise introduced by aliasing depends on the character of the waveform being sampled. Contributors to this noise can be random background noise, interfering signals and components of the desired signal. The effects of deterministic waveforms can be analyzed using the

relationships developed in the previous section. The effects of aliasing on the background random noise can be estimated if the power spectrum or autocorrelation function of the noise is known.

The power spectrum of a sampled wide sense stationary random process can be expressed either in terms of its unsampled power spectrum

$$S_d(f) = \frac{1}{T_a} \sum_{k=-\infty}^{\infty} S(f + \frac{k}{T_a}) \qquad (2.47)$$

or in terms of samples of the autocorrelation function $R(\tau)$

$$S_d(f) = \sum_{k=-\infty}^{\infty} R(kT_a)\exp\{-j2\pi fkT_a\} \qquad (2.48)$$

Since the autocorrelation function of a real random process is symmetric, Eq. 2.48 can be written

$$S_d(f) = R(0) + 2 \sum_{k=1}^{\infty} R(kT_a)\cos(2\pi fkT_a) \qquad (2.49)$$

Two types of filters which can be relatively easily evaluated are the sliding integrator and those filters whose transfer functions have a finite number of poles, all in the left half plane. The latter class includes most of those used in the audio to HF frequency ranges.

The impulse response of the sliding integrator integrating over the time interval t to t+T is

$$\frac{1}{T} \qquad 0<t<T$$

$$0 \qquad \text{elsewhere}$$

If this filter is excited by white Gaussian noise the power spectrum at the output is

$$S(f) = \frac{\sin^2(\pi fT)}{(\pi fT)^2} \tag{2.50}$$

The power spectrum of the sampled output can be written in the form

$$S_d(f) = \frac{1}{T_a}\sum_{m=-\infty}^{\infty} \frac{\sin^2(\pi\{f+m/T_a\}T)}{(\pi\{f+m/T_a\}T)^2} \tag{2.51}$$

Since $S_d(f)$ is periodic with period $\frac{1}{Ta}$ attention can be restricted to the frequency range $\frac{-1}{2T_a}\leq f\leq\frac{1}{2T_a}$.

The autocorrelation function of the integrator output is

$$\begin{array}{ll} (1+\frac{t}{T})/T & -T\leq t\leq 0 \\ (1-\frac{t}{T})/T & 0\leq t\leq T \\ 0 & \text{elsewhere} \end{array}$$

Using the autocorrelation function an alternative expression for $S_d(f)$ is

$$S_d(f) = \frac{1}{T}+\frac{2}{T}\sum_{k=1}^{[T/T_a]} (1-\frac{kT_a}{T})\cos(2\pi fkT_a) \tag{2.52}$$

This expression is much easier to use than is Eq. 2.51.

A particularly tractable noise model is the Weiner model where the noise process is generated by passing white Gaussian noise thru a filter with impulse response h(t) and transfer function H(s) [27]. This model is realistic because in many cases the noise present at the A/D converter

input is the result of having passed relatively wide-band noise thru a front end filter. The Laplace transform of the noise autocorrelation function is H(s)H(-s) (assuming that h(t) is real valued). The autocorrelation function is readily found by expanding H(s)H(-s) in a partial fraction expansion and inverting.

Consider the contribution of a single pole in H(s)H(-s) located at s=p (Re(p)<0). Let the associated residue be R_p. The contribution of this pole is

$$S_{dp}(f) = R_p\{1 + 2\sum_{k=1}^{\infty} \exp(pkT_a)\cos(\pi fkT_a)\} \quad (2.53)$$

The cosine can be replaced with its complex exponential representation. Eq. 2.53 then becomes the sum of two convergent geometric series. The summed result can be written

$$S_{dp}(f) = R_p\{1+2\frac{\cos(2\pi fT_a)-\exp(pT_a)}{\exp(pT_a)+\exp(-pT_a)-2\cos(2\pi fT_a)}\} \quad (2.54)$$

Similarly, given a pole pair p and p^* with residues R_p and R_p^* respectively an expression can be obtained describing their contribution to the power spectrum of the sample waveform. This contribution can be written

$$S_{dpp}(f) = 2c + \frac{2\{c*\cos((y+2\pi f)T_a)-d*\sin((y+\pi f)T_a)-c*\exp(xT_a)\}}{\exp(xT_a)+\exp(-xT_a)-2\cos((y+2\pi f)T_a)}$$

$$+ \frac{2\{c*\cos((y-2\pi f)T_a)-d*\sin((y-2\pi f)T_a)-c*\exp(xT_a)\}}{\exp(xT_a)+\exp(-xT_a)-2\cos((y-2\pi f)T_a)} \quad (2.55)$$

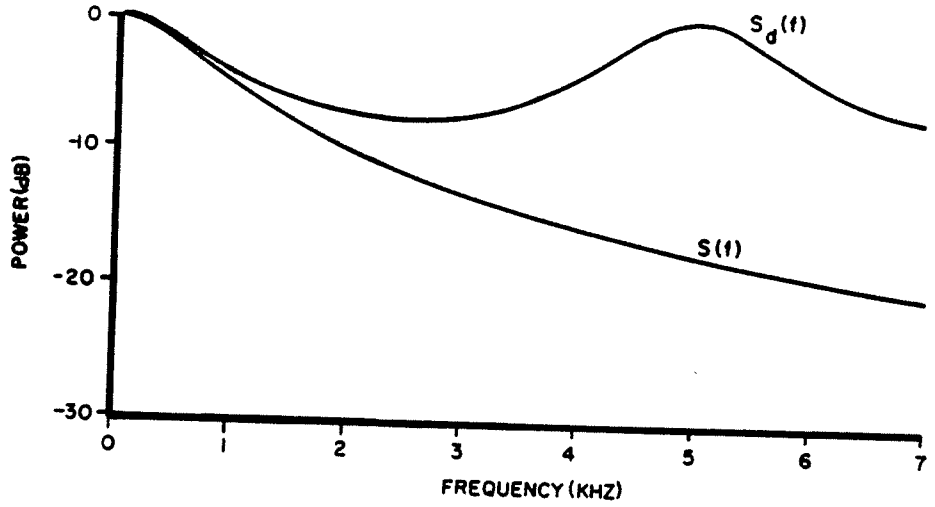
where

$$x = \mathrm{Re}(p) \qquad y = \mathrm{Im}(p)$$

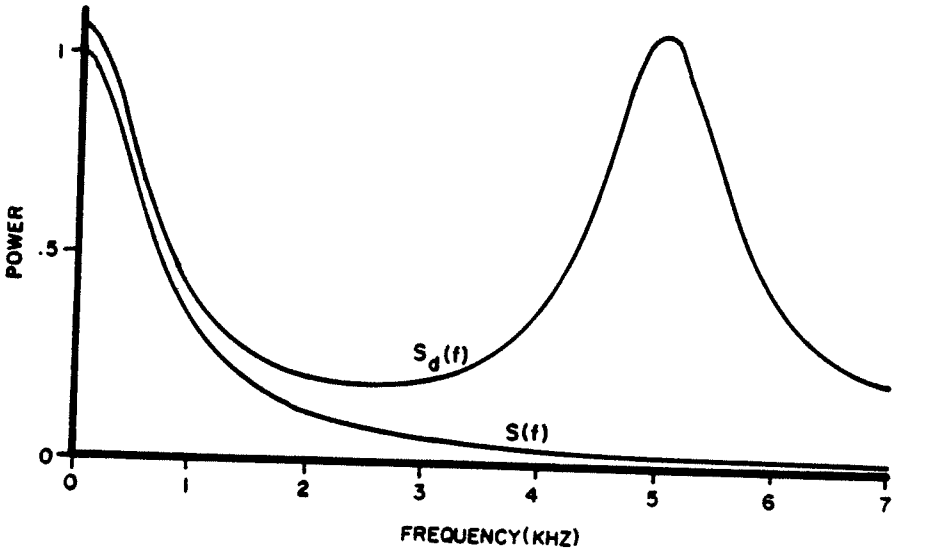
$$c = \mathrm{Re}(\mathrm{Residue}) \quad d = \mathrm{Im}(\mathrm{Residue})$$

Using the expressions for the single pole and for the pole pair contributions to the sampled power spectrum, the result for a transfer function of arbitrary order can be built up.

Fig. 2.5 shows the effects of sampling on the first order Butterworth power spectrum. The 3 dB bandwidth of the unsampled process is 700 Hz and the sample rate is 5000 Hz. Fig. 2.6 shows the effects of sampling for the same parameters but on a second order Butterworth spectrum. Both log and linear (vertical axis) plots are included. The horizontal axis is linear in frequency for all plots. While the log plot is useful for displaying values over a wide dynamic range it also magnifies small or insignificant values. The linear (in power) plot suppresses the tails of the spectra but since the total power in the waveform is directly related to the area under the curve it gives an accurate picture of the relative amounts of error. Assuming that the noise spectrum is shaped by a front end filter prior to the A/D converter, it is clear that for these parameters, the second order filter is a significantly better filter. Also, given that there are no strong spectral lines which need to be suppressed, there will not be a significant improvement in the amount of noise aliased if a higher order filter with the same bandwidth is used.

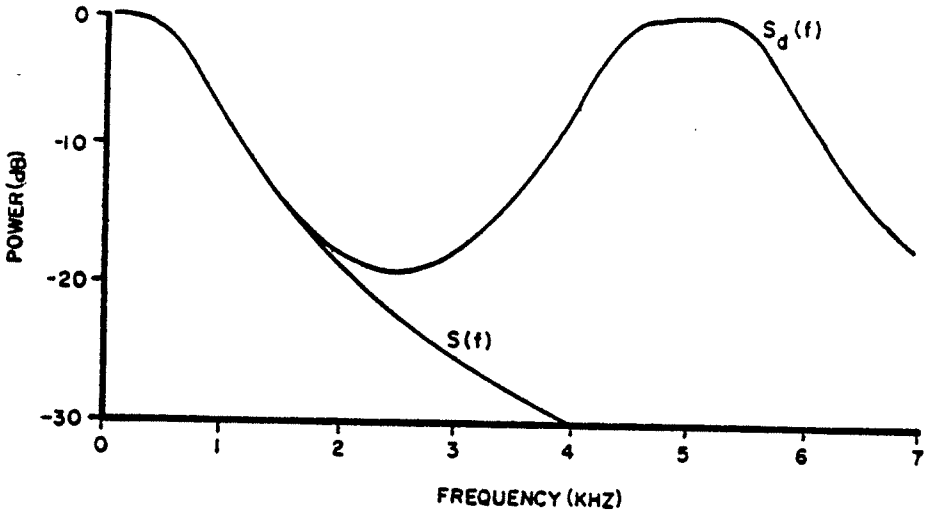


(a)

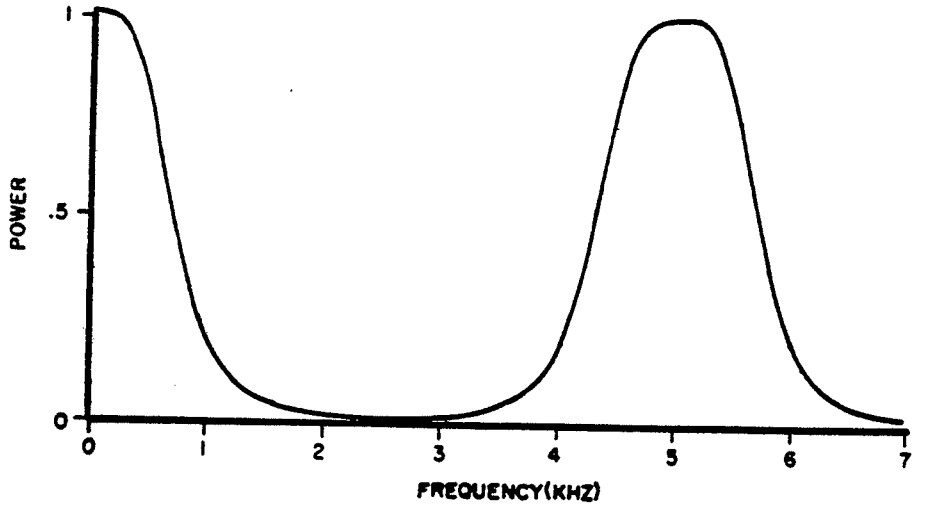


(b)

Fig. 2.5. Aliasing of 1st order Butterworth spectrum, BW=700 Hz, sample rate equal 5000 Hz a) Logarithmic power axis b) Linear power axis



(a)



(b)

Fig. 2.6. Aliasing of 2nd order Butterworth spectrum, BW=700 Hz, sample rate equal 5000 Hz a) Logarithmic power axis b) Linear power axis

CHAPTER 3

COMPLEX DEMODULATION

Given a waveform whose spectrum is either low-pass or band-pass, the process of complex demodulation consists of shifting that spectrum so that the energy contained in a band of frequencies centered at some frequency of interest, f_d, is re-centered at f=0. This is done in such a way that the phase relationships in the unshifted spectrum are preserved. The frequency shifted waveform is then filtered in order to reject the energy outside of the band of interest.

Using Fourier analysis, it is easily shown that multiplying a waveform s(t) whose spectrum is S(f) by the complex exponential $\exp\{-j2\pi f_d t\}$ results in a waveform whose spectrum is $S(f+f_d)$. This property of Fourier theory is sometimes referred to as the Fourier Shifting Theorem [26] and serves as the basis for the design of various forms of complex demodulators. The required complex multiplications can be implemented as pairs of real multiplications using the well known relation

$$\exp\{-j2\pi f_d t\} = \cos(2\pi f_d t) - j\sin(2\pi f_d t) \qquad (3.1)$$

The resulting "complex valued" waveform is handled as a pair of real valued waveforms in the same manner that complex numbers are handled in mathematics.

As defined here, the term "complex rotator" (or rotator) refers to the quantity, $\exp\{-j2\pi f_d t\}$.

Complex demodulators can be implemented in a number of ways. These include:

Using analog waveforms and analog multipliers.

Using multiplying D/A converters and digitally synthesized sine and cosine values with analog inputs and outputs.

A/D converting the input waveform and multiplying by sample values of the complex rotator.

Approximation of the totally analog system using quadrature phased square waves to mimic the action of the sine and cosine waveforms.

Direct sample demodulation (also called quadrature sampling and $4f_0$ sampling).

A completely analog system would use analog multiplier devices to multiply the input waveform by sine and cosine waveforms. The multiplier outputs would then be filtered to eliminate the energy outside the band of interest. The main difficulty with this approach is generating sine and cosine waveforms with precisely controlled amplitude and phase characteristics. Because of the precise control offered by digital techniques, the best way to generate these waveforms is digitally. However, rather than generate analog sine and cosine waveforms from digital versions, it is easier to use a hybrid analog and digital hardware combination.

A hybrid system would use multiplying D/A converters to multiply the analog input by values of the sine and cosine. Multiplying D/A converters act as amplifiers whose gain is digitally set, in this case by digitized values of the sine and cosine. A read only memory is used to hold the sine and cosine values. The ROM addresses can be generated in a number of ways. The simplest way is to use a counter which is clocked at a rate that is a multiple of the desired demodulation frequency. Another method is to use the direct digital frequency synthesis technique used in some frequency synthesizers to generate the required ROM addresses. The

quantizations of amplitude and time in the generation of the sine and cosine values are the main sources of error in the hyrbrid approach.

In applications using digial computers there is often no need for an analog output. An all digital system would A/D convert the input waveform and multiply the resulting samples by samples of the sine and cosine functions. The low-pass output filter would then be digital. The multiplication by the complex rotator values could be performed in the computer or off-loaded into a demodulator front end. In either case the required sine and cosine values would be extracted from a ROM. The addressing of the ROM could be sequential if the input sample rate is a multiple of the desired demodulation frequency or could be controlled using some direct digital frequency synthesis technique. Error sources include the input quantization as well as the quantization of the sine and cosine values. Often the input sample rate and the rate at which demodulator outputs are required at are not nicely related. One is then faced with the problem of sample rate conversion.

If the amount of generality required of the demodulator can be restricted then some simplifications can be made. In particular, if the input waveform is band-pass with center frequency f_d and Q greater than three or four then the following two methods are convenient.

A simplified analog demodulator can be built using quadrature phased square waves to approximate the sine and cosine functions. The required multiplications are then by plus and minus ones. These are easily implemented using buffer amplifiers and analog switches. The switching is controlled using a clock at a rate of four times the desired demodulation frequency. The outputs of the analog switches are low-pass filtered, not only removing the energy in the input waveform outside of the band of interest but also removing the effects of the high frequency components in the

square waves.

If the complex exponential is sampled at a rate of four times per cycle of the demodulation frequency starting at time zero, the resulting sine and cosine samples take on the values zero, plus one and minus one. These are easy values to multiply by in a computer. The properly weighted sample values can then be digitally filtered. One is still faced with the possible problem of sample rate conversion.

Section 3.1 provides a review of band-pass waveforms. The relationship between band-pass waveforms and their equivalent baseband representations is of primary interest. The desired outputs of a complex demodulator are these baseband representations.

Both the hybrid analog and the all digital demodulator configurations described in Sections 3.3 and 3.4 make use of synthesized sine and cosine values. Section 3.2 contains a description of a synthesis technique which can be used to generate the required values. This method is based on the direct digital method used in frequency synthesizers [12]. The effects of time and amplitude quantization errors are analyzed. This method of frequency synthesis is the method used in the multi-channel' complex demodulator described in Chapter 6.

Section 3.3 analyzes the performance of the hybrid analog demodulator. The effects of errors in the synthesis of the sine and cosine values on the demodulator output are evaluated.

The all digital demodulator is described in Section 3.4. This is the demodulator implementation used with the CEL beamforming hardware and so is considered in some detail. Because of the nature of the beamforming process, the input sample rate is fixed and cannot be adjusted to be a multiple of the desired demodulation frequency, f_d. As a consequence, the input sample rate is generally not a multiple of the rate at which samples of the demodulated waveform are desired. At some point in the demodulator, a

conversion must be made between the input sample rate and some multiple of the desired demodulator output rate. Three potential positions at which this conversion can be made are considered. The effects of quantization errors in both the input and the demodulation waveforms are examined. The use of the integrate-and-dump digital filter for the output filtering is described. The effects of sample rate conversion made at the output of this filter are analyzed. Finally, methods of scaling the output values without losing significance or inserting a bias are discussed.

Section 3.5 describes how a simple square wave based complex demodulator can be (and was) implemented. This method of complex demodulation has proven useful for making CW measurements and for tracking doppler shifts caused by moving sources. Historically, this was the first complex demodulation method used at CEL in the early 60's [13].

Section 3.6 considers direct digital demodulation. This method was originally used as a digital equivalent of square wave demodulation. This method was used in the single channel "$4f_0$" type demodulator [13] described in Chapter 6.

3.1 Band-Pass Waveforms

A deterministic waveform or random process is said to be band-pass with center frequency f_c and bandwidth W if its power spectrum contains negligible energy outside of the frequency range $f_c - W/2 < |f| < f_c + W/2$ where $f_c > W/2$. A representative band-pass spectrum is shown in Fig. 3.1.

A common way of generating a band-pass waveform is to take a sinusoidal carrier and modulate its amplitude and/or phase. If the modulation varies slowly compared to the carrier frequency, the resulting waveform is usually considered to be band-pass.

An amplitude and phase modulated cosine wave can be

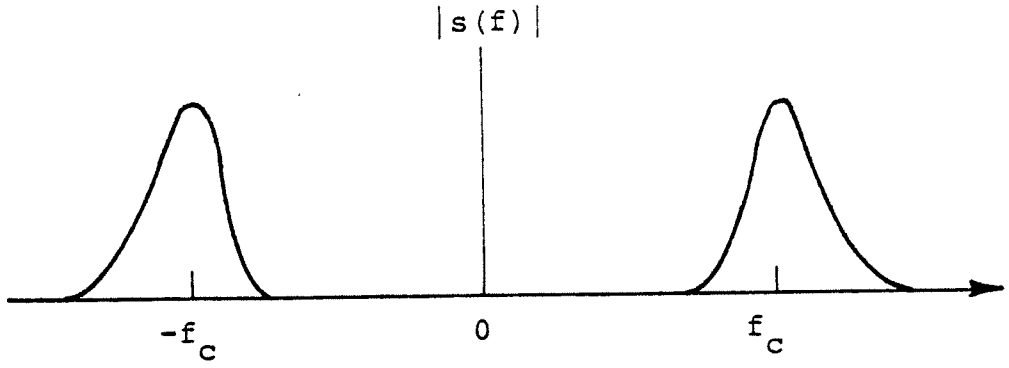


Fig. 3.1. A typical band-pass power spectrum

written in the form

$$m_c(t) = a(t)\cos(2\pi f_c t+\theta(t)) \tag{3.2}$$

where

$a(t)$ = amplitude modulating waveform
$\theta(t)$ = phase modulating waveform
f_c = carrier frequency

Typically $a(t)$ and $\theta(t)$ are real valued waveforms possessing low-pass power spectra.

Applying a standard trigonometric identity for the cosine of the sum of two angles allows $m_c(t)$ to be written as

$$m_c(t) = a(t)\cos(\theta(t))\cos(2\pi f_c t)-a(t)\sin(\theta(t))\sin(2\pi f_c t) \tag{3.3}$$

Define

$$m_x(t) = a(t)\cos(\theta(t))/2 \tag{3.4}$$

$$m_y(t) = a(t)\sin(\theta(t))/2 \tag{3.5}$$

Then $m_c(t)$ becomes

$$m_c(t) = 2m_x(t)\cos(2\pi f_c t) - 2m_y(t)\sin(2\pi f_c t) \qquad (3.6)$$

The waveform $m_x(t)$ is often referred to as the in-phase component of $m_c(t)$ and $m_y(t)$ is termed the quadrature-phase component of $m_c(t)$.

Given $m_x(t)$ and $m_y(t)$, it is possible to recover $a(t)$ and $\theta(t)$ using the relations

$$a(t) = 2\{m_x^2(t)+m_y^2(t)\}^{1/2} \qquad (3.7)$$

$$\theta(t) = \tan^{-1}\{\frac{m_y(t)}{m_x(t)}\} \qquad (3.8)$$

The above quadrature component representation for $m_c(t)$ was obtained assuming an amplitude and phase modulated carrier. The same functional form can be used to represent any stationary, zero mean random process [30] or the impulse response of a band-pass filter [31].

When using the quadrature representation to model a band-pass random process, the exact value chosen for f_c is usually not important and is often selected to simplify analysis. In situations where a carrier borne signal is present, the value of f_c is usually chosen to be the carrier frequency.

The structures of the positive and negative frequency components of $m_c(t)$ are readily determined by substituting the complex exponential representations of the sine and cosine into Eq. 3.3 and grouping terms. Doing this (suppressing the time dependence of $m_x(t)$ and $m_y(t)$ for notational convenience) gives

$$m_c(t) = (m_x+jm_y)\exp\{j2\pi f_c t\}+(m_x-jm_y)\exp\{-j2\pi f_c t\} \qquad (3.9)$$

The form of this equation suggests defining a complex valued baseband waveform, $m_z(t)$ as

$$m_z = m_x + jm_y \tag{3.10}$$

$$= \frac{1}{2}\{a(t)\cos(\theta(t)) + ja(t)\sin(\theta(t))\} \tag{3.11}$$

$$= \frac{1}{2}\, a(t)\exp\{j\theta(t)\} \tag{3.12}$$

The equation for $m_c(t)$ then becomes

$$m_c(t) = m_z\exp\{j2\pi f_c t\} + m_z^*\exp\{-j2\pi f_c t\} \tag{3.13}$$

Using the Fourier transform shifting theorem [30], the spectrum of $m_c(t)$ can be written in terms of the spectrum of $m_z(t)$ as

$$M_c(f) = M_z(f-f_c) + M_z^*(-(f+f_c)) \tag{3.14}$$

The spectrum of $m_c(t)$ is made up of two components. One of these components is generated by taking the complex valued modulating waveform's spectrum and shifting it up in frequency to f_c. The other component is generated by conjugating the modulating waveform's spectrum, flipping it about the origin and then shifting it to $-f_c$. Figure 3.2 illustrates how the spectra are related.

Figure 3.2 is a little deceptive in that it does not show any overlap between the positive and negative components of $M_c(f)$. However, except for a few special cases, there will be contributions from $M_z(f-f_c)$ and from $M_z^*(-(f+f_c))$ at all frequencies. The amount of overlap between these two components depends upon the bandwidth of M_z relative to the value of f_c (signal Q) and the rate at which $M_z(f)$ falls off.

In order to better understand when the interaction between the positive and negative frequency components is significant, consider the spectrum which results when white

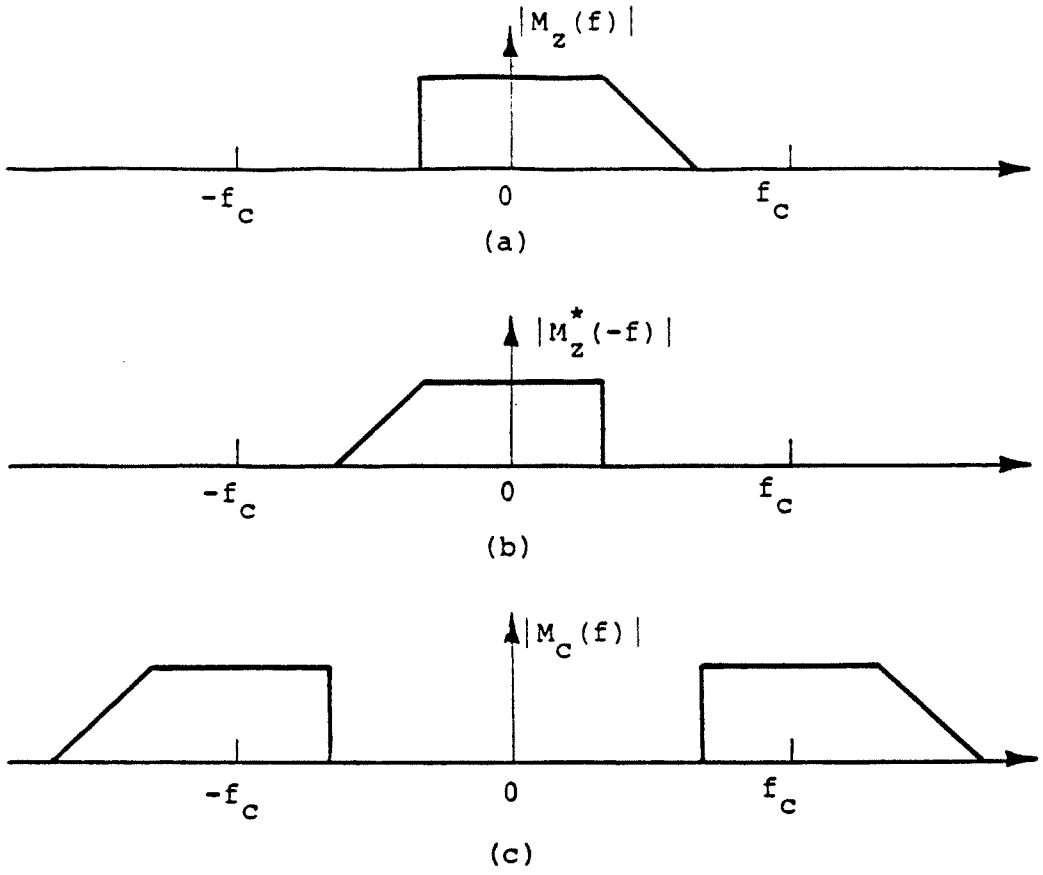


Fig. 3.2. Relationship between $M_z(f)$, $M_z^*(-f)$ and $M_c(f)$
a) $M_z(f)$ b) $M_z^*(-f)$ c) $M_c(f)$

noise is passed through a band-pass filter. The power spectrum at the filter output is

$$S_c(f) = K|H_c(f)|^2 \tag{3.15}$$

where

$H_c(f)$ = band-pass filter transfer function with center frequency f_c

K = white noise spectral level

Expressing $H_c(f)$ in terms of its baseband representation

$$\frac{1}{K}S_c(f) = |H_z(f-f_c)|^2 + |H_z(-f-f_c)|^2 + H_z(f-f_c)H_z(-f-f_c) + H_z^*(f-f_c)H_z^*(-f-f_c) \quad (3.16)$$

The total power measured at the filter output is

$$\int_{-\infty}^{+\infty} S_c(f)df = \int_{-\infty}^{+\infty} |H_c(f)|^2 df \quad (3.17)$$

$$= 2\int_{-\infty}^{+\infty} |H_z(f-f_c)|^2 df + 2\int_{-\infty}^{+\infty} R_e\{H_z(f-f_c)H_z(-f-f_c)\}df \quad (3.18)$$

If there were no interaction between the skirts of $H_z(f-f_c)$ and $H_z(-f-f_c)$, the second term would equal zero and the result obtained using just the baseband model would be exact. The size of the second term reflects the amount of interaction between the positive and negative frequency components of the spectrum. The relative importance of this term can be quantified using the relationship

$$M = \frac{\left|\int_{-\infty}^{+\infty} |H_c(f)|^2 df - 2\int_{-\infty}^{+\infty} |H_z(f-f_c)|^2 df\right|}{\int_{-\infty}^{+\infty} |H_c(f)|^2 df} \times 100 \quad (3.19)$$

The quantity M represents the percent error in the estimated noise power caused by ignoring the interaction between the positive and negative spectral components and

using only the baseband representation.

Given a specific filter type, this expression can be evaluated as a function of the filter Q (center frequency divided by bandwidth). The results of such a computation are plotted in Fig. 3.3. Setting a threshold of a one percent error gives a requirement that the Q of a first order Butterworth filter be 5 or greater and the Q of a second order Butterworth filter be 1.5 or greater.

3.2 Digital Generation of the Complex Exponential

Both the hybrid and the digital demodulators described in the following sections require sample values of the complex exponential $\exp\{-j2\pi f_d t\}$ at some rate f_s. One method of generating these values is to read them out of a read only memory (ROM). The ROM would be programmed to contain M equally spaced values of $\exp\{-j2\pi m/M\}$, $m=0,\ldots,M-1$. At sample time i/f_s, a ROM address k_i would be generated and the corresponding ROM value used in the demodulator. ROM addresses can be generated using the equation

$$k_i = [iMf_d/f_s] \quad \text{modulo } M \tag{3.20}$$

or

$$k_i = [iMf_d/f_s+0.5] \quad \text{modulo } M \tag{3.21}$$

where [ ] represents taking the integer part.

The first equation truncates the address value while the second equation rounds it. Unless $f_s=Mf_d$, the value read from the ROM will not exactly correspond to the desired value. The error due to quantizing the time axis is called time quantization error and is an error source in addition to amplitude quantization.

If M, f_s and the allowed range of values of f_d are properly chosen, the hardware needed to compute the k_i takes on relatively simple form. The value Mf_d/f_s is summed into

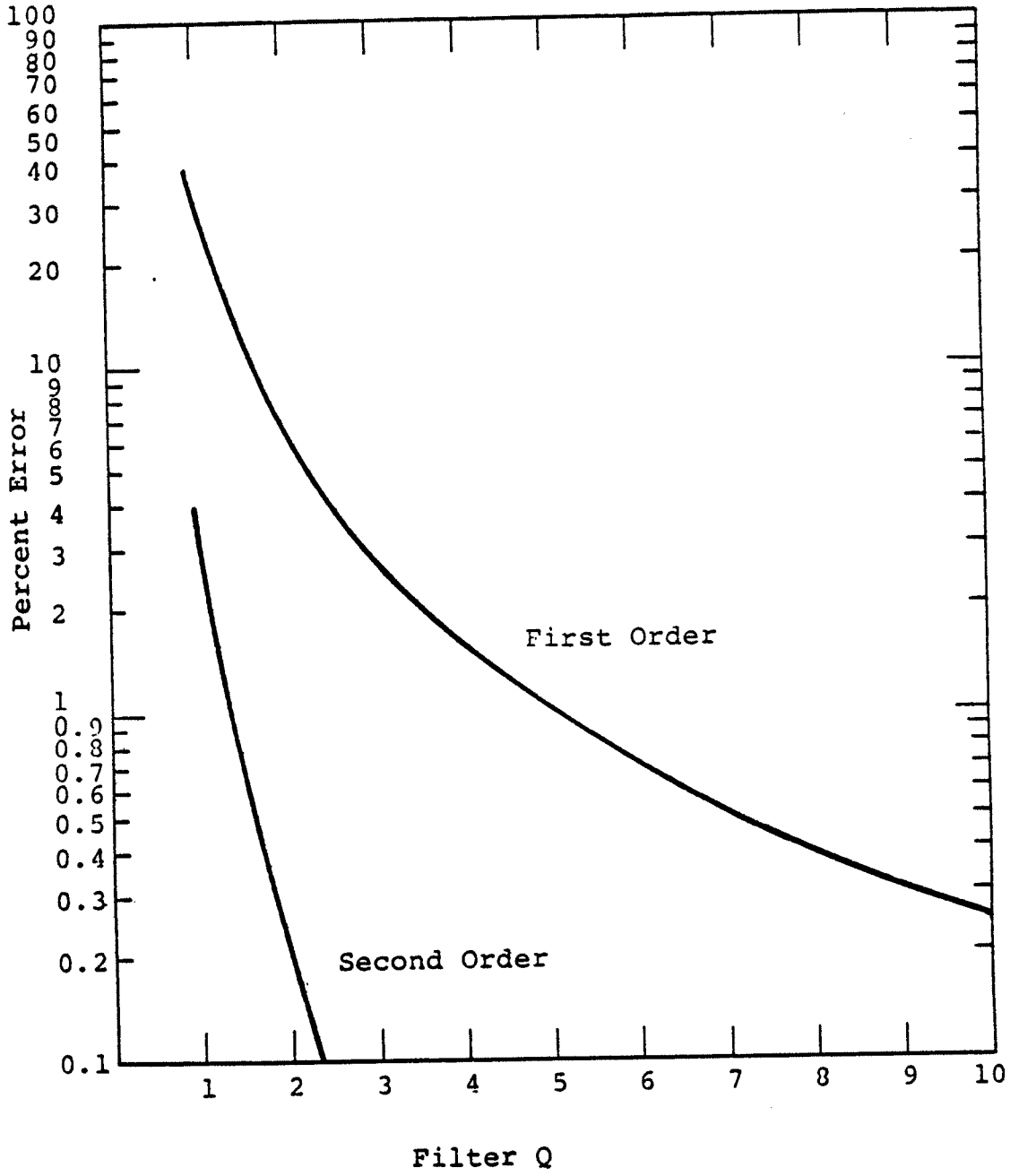


Fig. 3.3. Error in estimated power caused by ignoring cross spectral term a) First order Butterworth filter b) Second order Butterworth filter

a modulo-M accumulator each time the sample clock "tics." The accumulator is then used to address the ROM.

In practice, two ROM's of M words each are used to hold the complex exponential sample values. One ROM is used to contain the real part and the other is used to contain the imaginary part. Assuming B bit words, the values stored in the k-th word of each ROM are

$$CR(k) = [(2^{B-1}-1)|\cos(2\pi k/M)|+0.5]\mathrm{sgn}(\cos(2\pi k/M)) \quad (3.22)$$

$$SR(k) = -[(2^{B-1}-1)|\sin(2\pi k/M)|+0.5]\mathrm{sgn}(\sin(2\pi k/M)) \quad (3.23)$$

where [] represents taking the integer part, || the absolute value and sgn the sign of the enclosed quantity. Since multiplication by $\exp\{-j2\pi f_d t\}$ is being emulated, the contents of the sine ROM need to be negated as shown.

If M is even, the symmetry of the sine and cosine functions insures that the sample values given above sum to zero over one period. If M is odd, there is a small dc component.

At time i/f_s, the desired complex exponential value is

$$(2^{B-1}-1)\exp\{-j2\pi i f_d/f_s\} \quad (3.24)$$

The value read from the ROM is (truncating addresses)

$$(2^{B-1}-1)\exp\{-j2\pi[iMf_d/f_s]/M\} + e_k^a \quad (3.25)$$

where e_k^a is the amplitude quantization error in ROM location $k=[iMf_d/f_s]$. Using the algorithm given above for computing the ROM contents

$$-1/2 \leq e_k^a < 1/2 \quad (3.26)$$

The error in the complex exponential value read from the ROM is then

$$e_i = (2^{B-1}-1)\exp\{-j2\pi i f_d/f_s\}(1- \exp\{j2\pi(iMf_d/f_s-[iMf_d/f_s])/M\}) + e_k^a \quad (3.27)$$

The error made in approximating the sample times is

$$e_i^t = iMf_d/f_s-[iMf_d/f_s] = \text{fract}(iMf_d/f_s) \quad (3.28)$$

If it is possible to set $f_s=Mf_d$, the e_i^t are all zero and the only errors present are due to amplitude quantization. In many cases, however, the value of f_s is fixed by other considerations and it may not be possible to set $f_s=Mf_d$. In order to be able to implement this synthesis method in hardware, it is necessary to place some restrictions on the set of usable f_d values. Typically, f_s and f_d are required to both be an integer multiple of some basic frequency step, say 0.001 Hz. This restriction allows the ratio f_d/f_s to be written as the ratio of two relatively prime integers p and q.

$$f_d/f_s = p/q \quad (3.29)$$

The spectrum of the synthesized complex exponential is periodic with period f_s and contains q lines per replication. If there were no amplitude or time quantization, all of the lines except those corresponding to $-f_d$ and its replications would have zero magnitude. The amplitude of the line at $-f_d$, in this case, would be $2^{B-1}-1$.

Given a specific set of design parameters and a demodulation frequency, the amplitude and time quantization errors are determined. However, in order to estimate the magnitudes of these errors over a wide range of parameter values, these errors will be treated as being random. Additionally, the errors due to amplitude quantization will

be treated as being random and independent of the errors introduced by time quantization. The equations resulting from this assumption yields results that agree well with the results of computer simulations.

First, consider the effects of amplitude quantization.

The real and imaginary parts of e_k^a are modeled as being uniformly distributed between -1/2 and +1/2 quanta and independent of each other. The variance of e_k^a is equal to the sum of the variances of its real and imaginary parts. In terms of ROM quanta squared

$$\text{var}(e_k^a) = 1/12 + 1/12 = 1/6 = 0.17 \qquad (3.30)$$

As a check, the values of $\text{var}(e_k^a)$ were computed over one period of the complex exponential for values of B ranging from 2 thru 16 and values of M equal to powers of 2 ranging from 2^3 thru 2^{11}. For the most part, the computed variance values clustered slightly below the 0.17 value. The largest value found was 0.24 for B=11 and M=32. Even if all of the e_k^a took on the maximum possible value, the resulting variance value would still only be 0.5 quanta squared.

Let E_k^a represent DFT values associated with e_k^a. Parseval's theorem can be used to relate the variance of the quantization errors to the variance of the lines in the associated spectrum.

$$E\{\frac{1}{q}\sum_{k=0}^{q-1}|e_k^a|^2\} = E\{\sum_{k=0}^{q-1}|E_k^a|^2\} = 0.17 \text{ to } 0.5 \qquad (3.31)$$

If all of this energy were contained in a single spectral line, its magnitude would be on the order of 0.41 to 0.71 quanta. If the energy were to split evenly among all of the lines, the magnitudes would be $1/q^{1/2}$ this

amount. The spectral lines could also be modeled as being Chi-square with two degrees of freedom and variance .17/q. Using this model, given q lines, the magnitude of the mode of the distribution of the maximum over q values is

$$0.41(\ln(q)/q)^{1/2} \tag{3.32}$$

Next, consider the effects of time quantization.

The values of f_d and f_s are restricted so that the ratio f_d/f_s can be expressed as the ratio of two relatively prime integers, p and q. The expression for computing ROM addresses is then

$$k_i = [iMp/q] \mod M \tag{3.33}$$

Since M is also an integer, the quantity Mp/q can be expressed as the ratio of two relatively prime integers

$$Mp/q = p'/q' \tag{3.34}$$

The value q' equals q or is the value which remains after all factors in common with M and q have been removed.

The expression for the time quantization error becomes

$$e_i^t = \text{fract}(ip'/q') \tag{3.35}$$

As i is indexed, $q'e_i^t$ takes on the values in the set

$$\{0, 1, 2, \ldots, q'-1\} \tag{3.36}$$

with period q' and ordering as determined by the value p'.

The component of the quantization error in Eq. 3.27 due to time quantization is

$$(2^{B-1}-1)\exp\{-j2\pi ip/q\}(1-\exp\{j2\pi e_i^t/M\}) \qquad (3.37)$$

Taking the DFT yields the portion of the error spectrum due to the effects of time quantization

$$E^t(k) = \frac{2^{B-1}-1}{q}\sum_{i=0}^{q-1}(1-\exp\{j2\pi e_i^t/M\})\exp\{-j2\pi i(p+k)/q\} \qquad (3.38)$$

Replacing k with k'=k+p, substituting and dropping the primes gives

$$E^t(k-p) = \frac{2^{B-1}-1}{q}\sum_{i=0}^{q-1}(1-\exp\{j2\pi e_i^t/M\})\exp\{-j2\pi ik/q\} \qquad (3.39)$$

The e_i^t are periodic with period q' where q' is a factor of q. Let q'=q/r. Since $e_{i+q'}=e_i$, the above sum can be written in the form

$$E^t(k-p) = \frac{2^{B-1}-1}{q'}\sum_{i=0}^{q'-1}(1-\exp\{j2\pi e_i^t/M\})\exp\{-j2\pi ik/rq'\} \qquad (3.40)$$
$$\cdot \frac{1}{r}\sum_{m=0}^{r-1}\exp\{-j2\pi mk/r\}$$

The second summation equals 1 if k is a multiple of r and zero otherwise.

Let k=rk', substitute and drop the primes. The expression for the time quantization error spectrum is then

$$E^t(k-p) = \frac{2^{B-1}-1}{q'}\sum_{i=0}^{q'-1}(1-\exp\{j2\pi e_i^t/M\})\exp\{-j2\pi ik/q'\} \qquad (3.41)$$

The spectrum consists of lines spaced $rf_s/q=f_s/q'$ Hz

apart replicated with period f_s.

For large values of M

$$\exp\{j2\pi e_i^t/M\} \approx 1+j2\pi e_i^t/M \tag{3.42}$$

Substituting into the equation for E^t gives

$$E^t(k-p) = \frac{-j2\pi(2^{B-1}-1)}{q'M} \sum_{i=0}^{q'-1} e_i^t \exp\{-j2\pi ik/q'\} \tag{3.43}$$

Using Parseval's theorem, the energy in the time quantization error spectrum is

$$\sum_{k=0}^{q'-1} |E^t(k-p)|^2 = \left(\frac{2\pi(2^{B-1}-1)}{q'M}\right)^2 \frac{1}{q'} \sum_{i=0}^{q'-1} i^2 \tag{3.44}$$

$$= \frac{2}{3}\left(\frac{\pi(2^{B-1}-1)}{q'M}\right)^2 (q'-1)(2q'-1)$$

The energy associated with the line k=0 causes a slight gain and phase shift in the operation of the demodulator but does not give rise to any spurious lines in the spectrum. It is thus reasonable to remove the contribution of this line from the above sum in order to determine the amount of energy contained in the undesired spectral lines.

$$E^t(-p) = -j\pi(2^{B-1}-1)(q'-1)/(q'M) \tag{3.45}$$

Squaring and subtracting gives

$$\sum_{k=1}^{q'-1} |E^t(k-p)|^2 = \frac{1}{3}\left(\frac{\pi(2^{B-1}-1)}{q'M}\right)^2 (q'^2-1) \tag{3.46}$$

If all of this energy were to reside in a single spectral line, its magnitude would be

$$\frac{\pi}{3^{1/2}} \frac{2^{B-1}-1}{M} \left(\frac{q'^2-1}{q'^2}\right)^{1/2} \approx \frac{\pi}{3^{1/2}} \frac{2^{B-1}-1}{M} = 1.81 \frac{2^{B-1}-1}{M} \quad (3.47)$$

All lines in the time quantization error spectrum must have amplitudes less than or equal to this value.

If the e_i^t values are ordered as shown in Fig. 3.4, the magnitude of the k=1 term is approximately

$$\frac{4}{\pi} \frac{2^{B-1}-1}{M} = 1.27 \frac{2^{B-1}-1}{M} \quad (3.48)$$

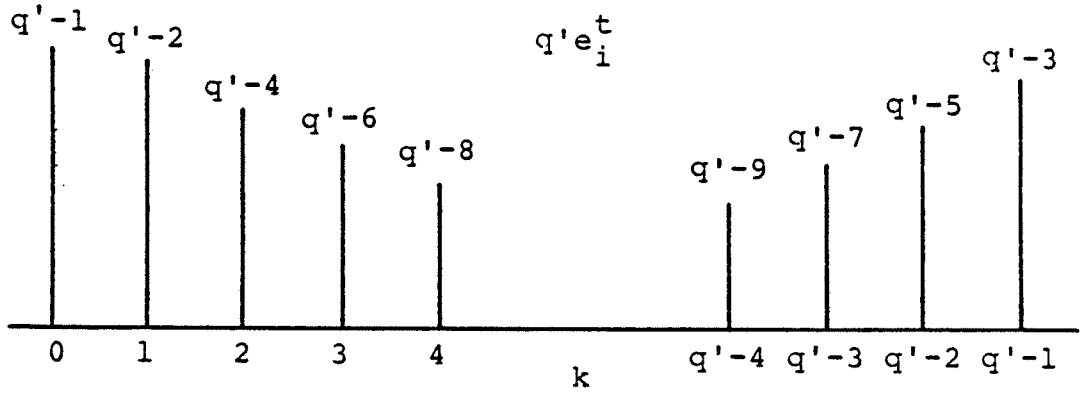


Fig. 3.4. Ordering of e_i^t yielding near absolute maximum amplitude spurious spectral line

This ordering places half of the available error energy in the k=1 line. Although a proof has not been found, it is felt that this is the maximum amount of the error energy which can be placed in a single line. Since the above results are to be used in estimating the magnitude of the maximum non-demodulation frequency line in the demodulator spectrum, the difference between a value of 1.81 and 1.27 is not large enough to merit looking for a tighter bound.

If the ROM addresses are rounded instead of being

truncated, the time jitter error becomes

$$e_i^t = iMf_d/f_s - [iMf_d/f_s + 0.5] \tag{3.49}$$

Again assume that the ratio Mf_d/f_s can be expressed as the ratio of two relatively prime integers p' and q'.

For even values of q', as i is indexed, $q'e_i^t$ takes on the values in the set

$$\{0,\ 1,\ 2,\ldots,\ q'/2-1,\ -q'/2,\ -q'/2-1,\ldots,\ -2,\ -1\} \tag{3.50}$$

For odd values of q', as i is indexed $q'e_i^t$ takes on the values in the set

$$\{0,\ 1,\ 2,\ldots,\ (q'-1)/2,\ -(q'-1)/2,\ldots,\ -2,\ -1\} \tag{3.51}$$

The actual ordering is determined by the value of p'. The k=0 term in the error spectrum for each case is

$$E^t(-p) = \frac{j\pi(2^{B-1}-1)}{q'M} \quad (q' \text{ even}) \tag{3.52}$$

$$= 0 \qquad (q' \text{ odd})$$

Applying Parseval's theorem, the total energy in the error spectrum not including the k=0 line is

$$\sum_{k=1}^{q'-1} |E^t(k-p)|^2 = \frac{1}{3}\left(\frac{\pi(2^{B-1}-1)}{M}\right)^2 \frac{(q'-2)(q'-1)}{q'^2} \quad (q' \text{ even}) \tag{3.53}$$

$$= \frac{1}{3}\left(\frac{\pi(2^{B-1}-1)}{M}\right)^2 \frac{(q'-1)(q'+1)}{q'^2} \quad (q' \text{ odd})$$

The upper limit on the maximum line magnitude is then

approximately

$$\frac{\pi}{3^{1/2}} \frac{2^{B-1}-1}{M} \tag{3.54}$$

The maximum spurious line magnitude is the same as when the ROM addresses are truncated. There is a small difference, however, in the effect on the line at the desired demodulation frequency.

3.3 Hybrid Analog Demodulator

The analog demodulator shown in Fig. 3.5 uses a combination of analog and digital components. Sample values of $\exp\{-j2\pi f_d t\}$ are generated at a rate f_s as described in the previous section. The real and imaginary parts are used to set the gains of separate multiplying D/A converters. Using B-bit converters, the gains are assumed to be set such that a digital input value of $2^{B-1}-1$ corresponds to multiplication by one. The digital values are held constant over the $1/f_s$ second inter-sample period. The waveform to be demodulated is applied to the analog inputs of the multiplying D/A converters. The input gain should be adjusted so that clipping is an infrequent occurance. The weighted analog outputs of the converters are low-pass filtered to remove the higher frequency components introduced by the reconstructed exponential and to reject energy in the input spectrum outside of the band of frequencies that is of interest.

The operation of the digital complex exponential synthesizer in conjunction with the multiplying D/A converter can be modeled as being a reconstruction of the complex exponential using a zero order hold followed by an ideal analog multiplier. The spectrum of the multiplier output corresponds to the convolution of the spectrum of the

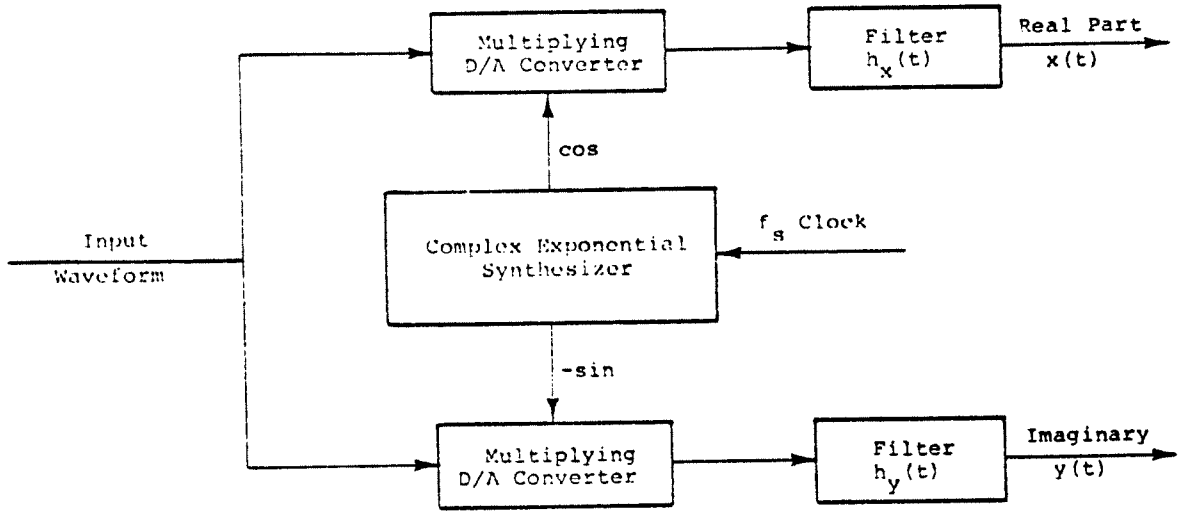


Fig. 3.5. Hybrid analog demodulator structure

input waveform with the spectrum of the reconstructed complex exponential.

The spectrum of the reconstructed complex exponential consists of the periodic line spectrum of the sampled complex exponential with the amplitude of the line at frequency f being weighted by

$$\frac{\sin(\pi f/f_s)}{\pi f/f_s} \tag{3.55}$$

The unweighted periodic line spectrum of the sampled complex exponential consists of shifted replications of the $-f_s$ to 0 portion of the spectrum. The shift amounts are multiples of f_s. In the $-f_s$ to 0 region there are q spectral lines spaced f_s/q apart. The line at $-pf_s/q$ corresponds to $-f_d$, the desired demodulation frequency. The remaining lines are caused by time and amplitude quantization.

Fig. 3.6 illustrates how the spectrum of the

reconstructed complex exponential is structured.

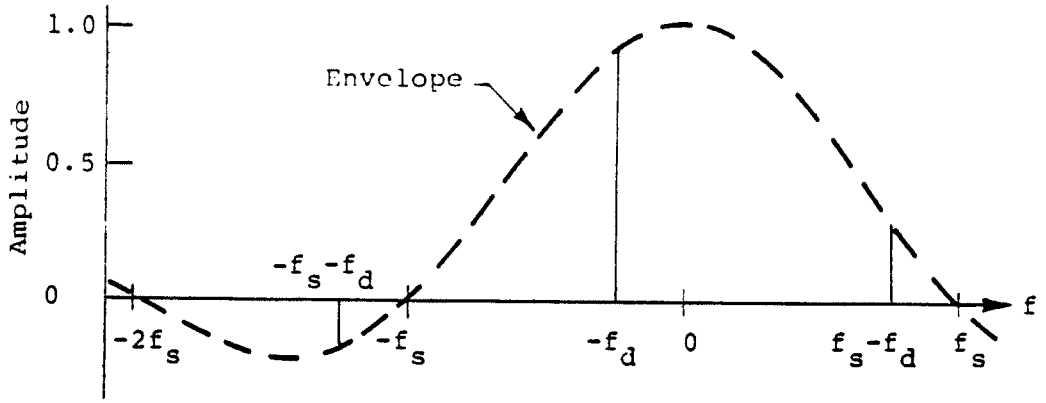


Fig. 3.6. Reconstructed complex exponential spectrum

The spectrum at the output of the multiplying D/A converter, neglecting the contributions of the noise lines, consists of a series of replications of the input waveform's spectrum shifted by $-f_d$ plus multiples of f_s and weighted by

$$\frac{\sin(\pi(f_d \pm mf_s)/f_s)}{\pi(f_d \pm mf_s)/f_s} \tag{3.56}$$

If the bandwidth of the input waveform is W_i, there will be minor spectral overlap of replications if

$$f_s - W_i - f_d > W_i - f_d \tag{3.57}$$

Not unexpectedly, rearranging gives the classical sampling relation

$$f_s > 2W_i \tag{3.58}$$

Given a noisy reception where the signal bandwidth satisfies the above restriction but the noise spectrum does not, the primary effect of the overlap of the replicated spectra is to increase the noise power in the band of interest. Consider the extreme case where the input noise

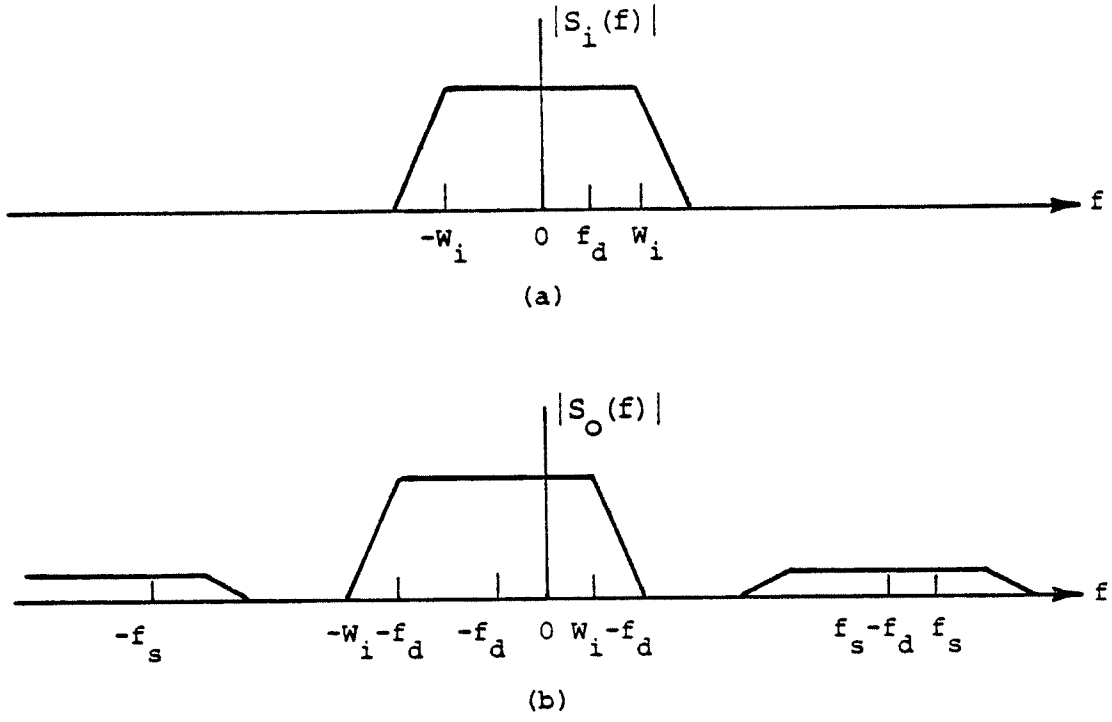


Fig. 3.7. Multiplying D/A Converter input and output spectra a) Input b) Output

spectrum is white noise of power spectral density level N_o. The spectral level of the overlapped replications is then

$$N = N_o \sum_{i=-\infty}^{\infty} \frac{\sin^2(\pi(f_d/f_s+i))}{(\pi(f_d/f_s+i))^2} \tag{3.59}$$

The summation in the above equation can be shown to equal 1 by applying Parseval's theorem to the Fourier series of the exponential $\exp\{-j2\pi f_d t\}$ defined on the interval from $-1/f_s$ to $+1/f_s$.

Thus,

$$N = N_o \tag{3.60}$$

The signal amplitude has been reduced by the amount

$$\frac{\sin(\pi f_d/f_s)}{\pi f_d/f_s} \tag{3.61}$$

The signal to noise ratio in the band of interest is multiplied by a factor of

$$\left(\frac{\sin(\pi f_d/f_s)}{\pi(f_d/f_s)}\right)^2 \tag{3.62}$$

If f_d is restricted to be less than $f_s/4$, the signal to noise ratio with overlapped replications present is no less than 0.81 times the signal to noise at the analog input of the multiplying D/A converter. A factor of 0.81 corresponds to a loss of slightly less than 1 dB.

The effects of the quantization noise lines in the spectrum of the reconstructed complex exponential on the noise power spectral level at the D/A converter output can be determined in a similar manner. The effects of amplitude and time quantization will be treated as being independent of each other.

In the frequency range from $-f_s$ to 0, amplitude quantization contributes energy to q lines spaced f_s/q Hz apart and time quantization contributes energy to q' lines spaced f_s/q' Hz apart. Consider first the effects of the amplitude quantization.

Let e_k^2 be the error energy in the k-th line. Since the k-th line is replicated and weighted by the reconstruction function, the resulting contribution of this line and its replications to the output noise spectral level is

$$N_k = N_o\, e_k^2 \sum_{i=-\infty}^{\infty} \frac{\sin^2(\pi(k/q+i))}{(\pi k/q+i)^2} \tag{3.63}$$

Since the summation in the above equation equals one,

$$N_k = N_o e_k^2 \tag{3.64}$$

The total contribution of the q lines in the $-f_s$ to 0 band and their replications is

$$N_o \sum_{k=0}^{q-1} e_k^2 \tag{3.65}$$

Using the results of the previous section, the value of the above sum is 0.17 quanta squared.

A similar argument holds for the contribution due to the time quantization errors. Thus, the total contribution to the output noise power spectral density (scaling $2^{B-1}-1$ quanta to correspond to unity gain) is

$$N_Q = \frac{N_o}{(2^{B-1}-1)^2} \{0.17 + \frac{((2^{B-1}-1)\pi)^2}{3M^2}\} \tag{3.66}$$

Assuming the use of a three (or more) bit D/A converter, the value of N_Q/N_o is less than 0.1 (-10 dB) for values of M equal to or greater than 7. These values of B and M are much smaller than those that would be used in an actual system.

Independent of how the noise errors distribute spectrally and using the extreme assumption of a white noise input, the above results show that if

$$\begin{aligned} B &\geq 3 \\ M &\geq 7 \\ f_s &\geq 4f_d \end{aligned} \tag{3.67}$$

then there is a negligible increase in the noise level at the multiplying D/A converter output due to spurious lines in the spectrum of the reconstructed complex exponential.

It now remains to estimate the magnitude of the effects on the signal portion of the input spectrum caused by the spurious lines in the spectrum of the synthesized exponential waveform.

Let the analog input to the demodulator be a bandpass waveform with center frequency f_c and bandwidth W_s. Ignoring the effects of the lines due to quantization errors, the spectrum at the demodulator output prior to the filters consists of weighted replications of the input spectrum. This is illustrated in Fig. 3.8 for $f_c=f_d$.

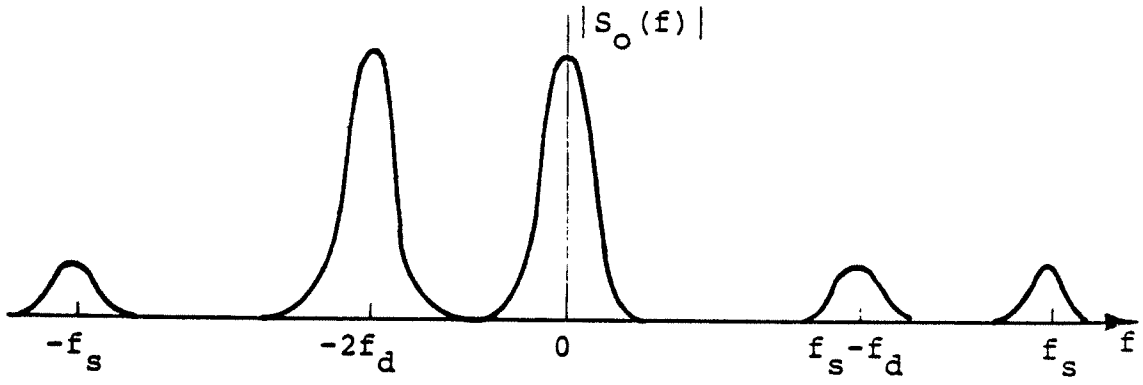


Fig. 3.8. Demodulator output spectrum prior to filtering

Only the spectral component centered about zero frequency is of interest. As long as the lower sideband of the replica centered at f_s-f_d lies at or above $2f_d$, its effects are essentially no worse than those of the sideband energy located at $-2f_d$. That is, if the output filtering effectively suppresses the component at $-2f_d$, then it also effectively eliminates the replicated spectral components. The restriction on f_s and f_d this gives is

$$f_s \geq 4f_d \tag{3.68}$$

The quantization noise in the reconstructed complex exponential values causes replications of the input spectrum, spaced f_s/q Hz apart, to appear in the spectrum at

the output of the multiplying D/A converter.

If all of the quantization error energy were to concentrate into a single spectral line, the worst position for that line would be at $f=+f_d$. This line would interact with the input waveform's lower sideband and shift it to zero frequency. Treating this interference as noise, the resulting signal to noise ratio would be

$$\left(\frac{\sin^2(\pi f_d/f_s)}{(\pi f_d/f_s)^2}\right)\left\{\frac{0.17}{(2^{B-1}-1)^2}+\frac{\pi^2}{3M^2}\right\}^{-1} \qquad (3.69)$$

The contributions to the noise power due to amplitude and time quantization become equal when

$$M = 4.4(2^{B-1}-1) \simeq 2^{B+1} \qquad (3.70)$$

In the CEL system, as configured for the Ocean Acoustic Tomography Demonstration Experiment, parameter values were B=16, M=2000, f_d=224 and f_s=5000. The computed signal to noise ratio is 60.8 dB.

In general, the quantization error energy splits among several lines and only a portion of these lines interact with the input spectrum to cause interfering replications in the band of interest. The resulting noise energy is thus less than the value used in the above signal to noise ratio estimate. Although the above estimate is pessimistic, it is useful in establishing a reasonable lower bound to system performance for specific B, M and f_d range.

A potential trouble spot in an analog system which is not immediately apparent is the need to precisely match the characteristics of the filters used at the outputs of the sine and cosine multiplying D/A converters. Let the spectrum at the input of the demodulator be

$$W_z(f-f_d)+W_z^*(-f-f_d) \tag{3.71}$$

The filtering will be assumed adequate to remove the effects of the spectral replications in the spectrum of the reconstructed sine and cosine waveforms. Similarly, the effects of quantization errors are taken to be negligible.

The spectrum at the output of the cosine multiplier is

$$X(f) = \{W_z(f-2f_d)+W_z(f)+W_z^*(-f)+W_z^*(-f-2f_d)\}/2 \tag{3.72}$$

The spectrum at the output of the sine multiplier is

$$Y(f) = j\{W_z(f-2f_d)-W(z)+W_z^*(-f)-W_z^*(-f-2f_d)\}/2 \tag{3.73}$$

Assuming the filters are adequate to remove the components shifted by $2f_d$, the filtered spectra are

$$X_f(f) = \{W_z(f)+W^*(-f)\}H_x(f)/2 \tag{3.74}$$

$$Y_f(f) = j\{-W_z(f)+W_z^*(-f)\}H_y(f)/2 \tag{3.75}$$

The spectrum of the complex valued waveform formed by the two multiplier outputs is then

$$Z(f) = X(f)+jY(f) \tag{3.76}$$

$$= W_z(f)\{H_x(f)+H_y(f)\}/2 + W_z^*(-f)\{H_x(f)-H_y(f)\}/2 \tag{3.77}$$

If $H_x(f)=H_y(f)=H(f)$, then

$$Z(f) = W_z(f)H(f) \tag{3.78}$$

Any mismatch in the amplitude and phase characteristics

of the two filters allows energy from the $W_z^*(-f)$ component of the spectrum to cause amplitude and phase errors in the demodulated waveform.

3.4 Digital Demodulator

The operation of the digital demodulator is quite similar to that of the hybrid analog demodulator except that the input waveform is digitized prior to being spectrally shifted by $-f_d$ Hz. The structure of the digital demodulator being considered is shown in Fig. 3.9.

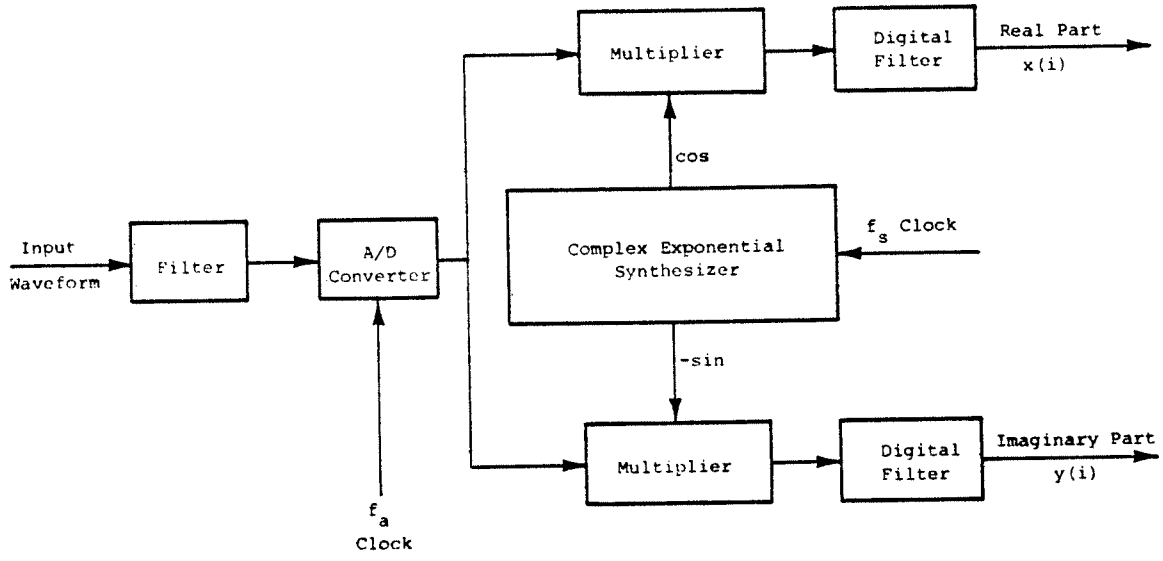


Fig. 3.9. Digital demodulator structure

The input waveform is filtered in order to remove as much noise as is reasonable for the particular application and then sampled at rate f_a. The sample values of the complex rotator $\exp\{-j2\pi f_d t\}$ are digitally synthesized as in the previous section. The input sample values and the complex rotator sample values are digitally multiplied together. The multiplier outputs are then digitally

filtered in order to remove the spectral components outside of the range of interest. The digital filter outputs may themselves be sampled at some rate f_b.

The primary design problem associated with this system is in selecting the sample rates f_a, f_s and f_b.

In a general purpose demodulator design, it may not be possible to relate the three sample rates f_a, f_b and f_s in a fixed manner. For example, the CEL beamformer is intended for use in supplying beam outputs to multiple demodulators operating at independently selected frequencies. Because of this, it is not possible to set the value of f_a based on the requirements of any one particular demodulator. In the digital demodulators considered below, it is assumed that the value of f_a is fixed and cannot be varied. In order to limit the number of degrees of freedom in the demodulator design, the value of f_s will be taken to be related either to f_a or f_b in some fixed manner. It is necessary to convert, at some point within the demodulator, between the f_a based input sample rate and the f_b based output sample rate. The three demodulator configurations shown in Fig. 3.10 illustrate the most natural positions within the demodulator at which sample rate conversion can be made.

The demodulator of Fig. 3.10a samples the input waveform at rate f_a. The spectrum at the A/D converter output is related to the input spectrum S(f) as

$$S_{AD}(f) = f_a \sum_{n=-\infty}^{\infty} S(f+nf_a) \qquad (3.79)$$

The process of resampling the A/D converter output at rate kf_b is accounted for by weighting $S_{AD}(f)$ by

$$\frac{\sin(\pi f/f_a)}{\pi f} \qquad (3.80)$$

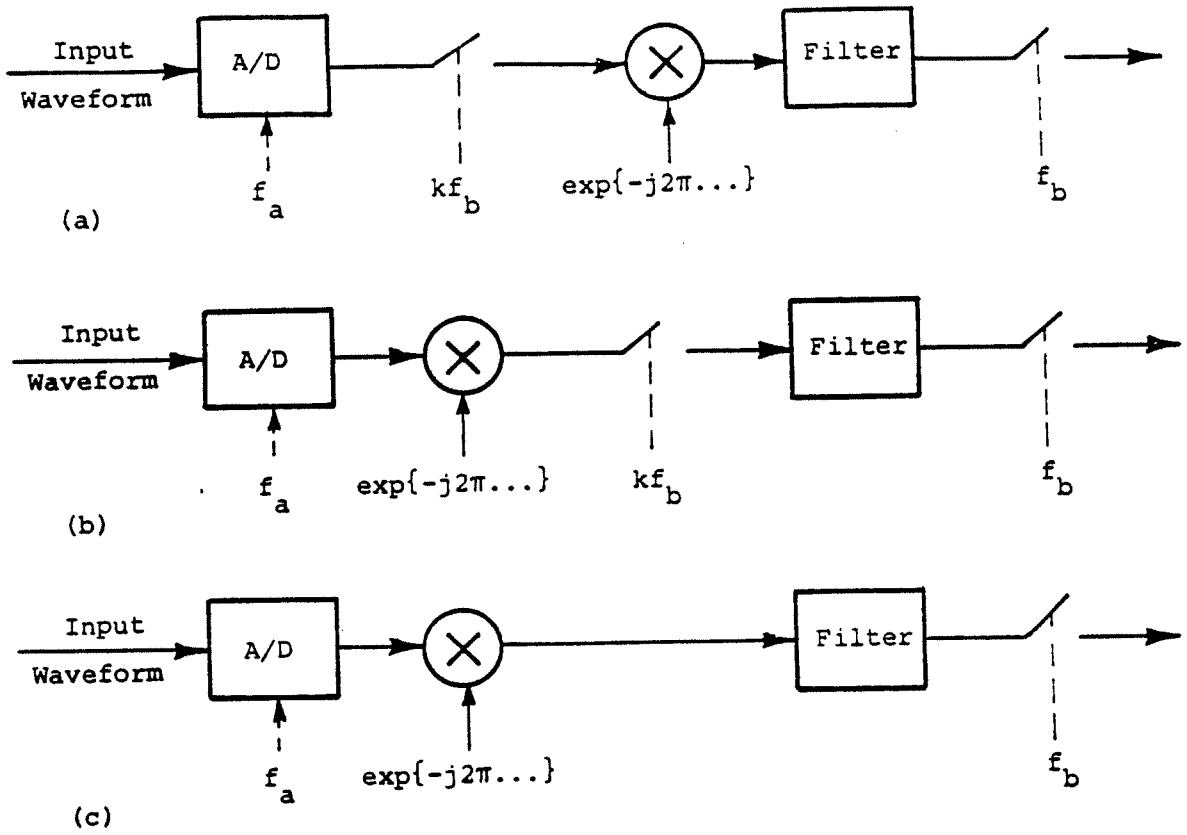


Fig. 3.10. Demodulator implementations with sample rate conversion a) Conversion following the A/D converter b) Conversion following the multiplier c) Conversion at the filter output

and then re-replicating with spacing kf_b and scale factor kf_b. The effects of these three steps are shown in schematic form in Fig. 3.11. In order to keep the figure simple, not all of the effects of the spectral replication are shown in Fig. 3.11c.

The effect of the sin(x)/x weighting is to reduce the magnitude of the replications caused by the initial sampling at f_a. These weighted replications can be shifted into the band of frequencies of interest on resampling at rate kf_b.

The demodulator of Fig. 3.10b samples the input waveform at rate f_a and then shifts the spectrum by $-f_d$ before resampling takes place. The multiplier output is resampled at rate kf_b. Fig. 3.12 shows the effect of these

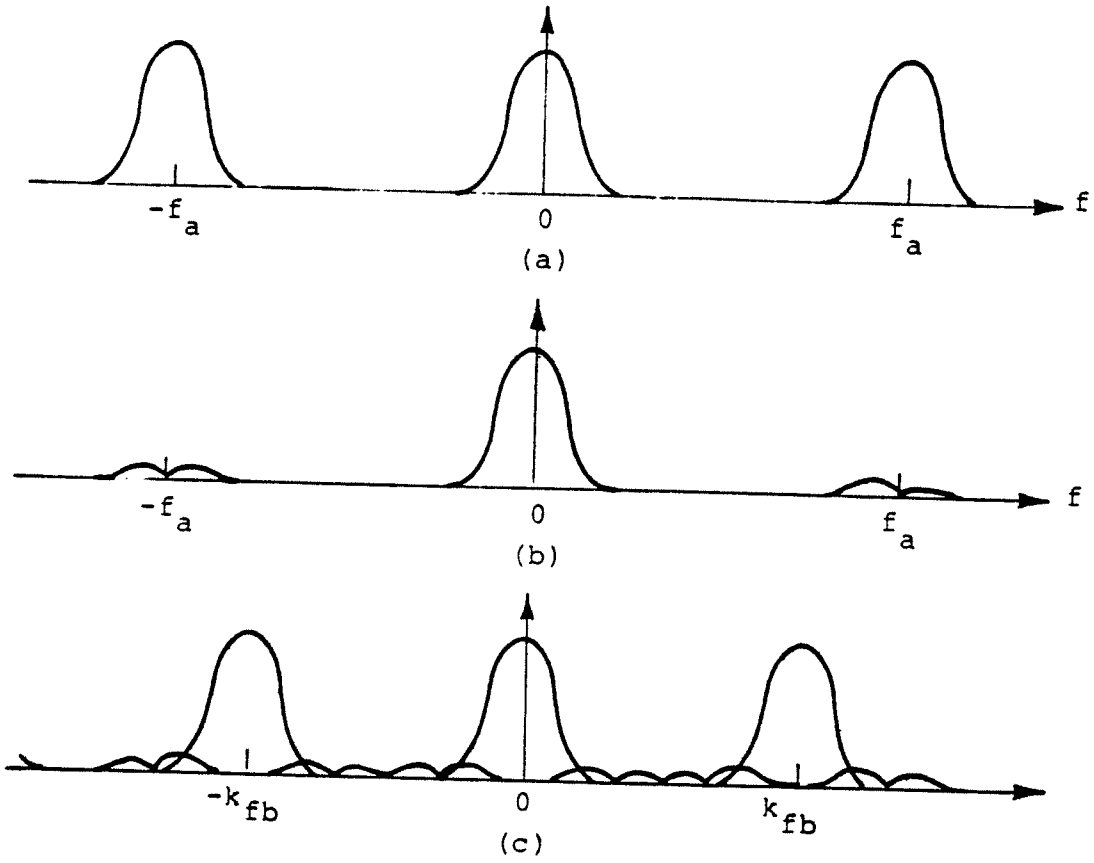


Fig. 3.11. Spectra associated with demodulator configuration A a) Spectrum at the A/D converter output b) Reconstructed spectrum prior to resampling c) Spectrum of the resampled waveform

three operations on the spectrum. Fig. 3.12c does not show all the aliased spectral components.

Performing the frequency shifting prior to resampling positions the spectrum so that the replications of the energy originally at f_d are positioned at multiples of the sample rate f_a. The process of reconstruction using a zero order hold places spectral zeros at these frequencies and provides maximum attenuation around these frequencies. Thus, when the re-replication caused by resampling occurs, the amount of "signal" energy which can be shifted into the band of the remaining processing is significantly reduced.

In particular, if the input waveform has a strong carrier line at frequency f_d, this line is effectively notched out of the replications.

The major improvement of the Fig. 3.10b scheme over that of Fig. 3.10a is the use of the zeros of the reconstruction function to attenuate the replications of the band of frequencies originally centered around f_d in the spectrum of the input waveform. The major drawback to this scheme is that nothing has been done to attenuate the spectral components located at $-f_d$ in the input spectrum.

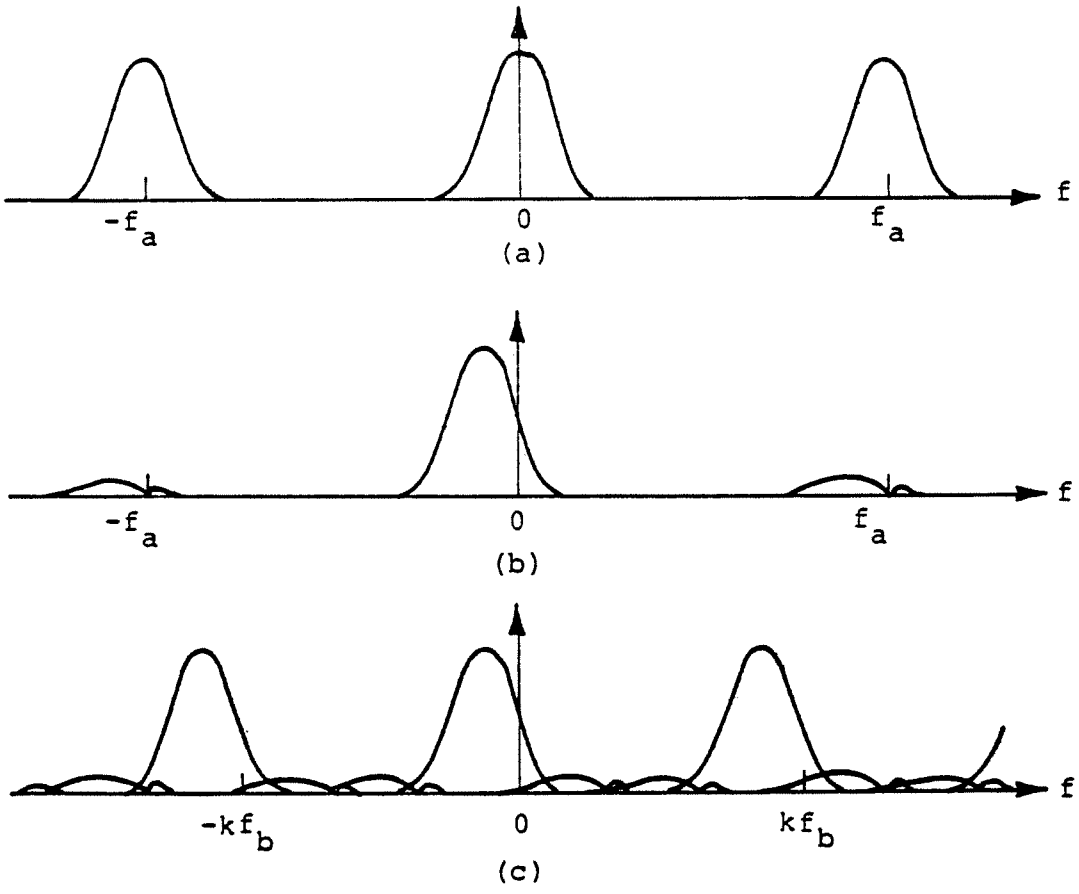


Fig. 3.12. Spectra associated with demodulator configuration B a) Spectrum at the A/D converter output b) Spectrum of reconstructed waveform at the multiplier output c) Spectrum of the resampled waveform

The demodulator of Fig. 3.10c moves the resamping point

so that it follows the output filter. This allows the filter to be used to suppress the energy that was originally located at $-f_d$ and was shifted to $-2f_d$ by the multiplier. This eliminates the problem that plagued the configuration in Fig. 3.10b. In addition, the filtering now reduces the noise bandwidth prior to re-replication decreasing the self noise level of this configuration as compared to the other two configurations. The sample rate is converted at the filter ouputs from f_a to f_b. As in arrangement of Fig. 3.10b, the spectrum is positioned so that the reconstruction function zero crossings heavily attenuate the replications of the band of frequencies originally centered around f_d. Because the demodulator configuration of Fig. 3.10c heavily attenuates these replications, it was the one chosen for implementation.

Having selected the demodulator configuration, the demodulator design consists of choosing sample rates, the number of bits used in the input A/D converter, the number of bits and words used in the complex exponential synthesizer ROM and the choice of the output filter.

3.4.1 Number of Bits in the Input A/D Converter

The number of bits used in the front end A/D converter depends on the expected dynamic range of the input, the expected background noise level and the desired signal power to quantization noise power level.

When the input waveform is essentially noise free, if the input signal sweeps over about 5 or more bits, the quantization noise can be modeled as being independent of the signal, spectrally white and having variance equal to 1/12 quanta squared. Given the average power in the waveform at the A/D input and the desired signal-to-noise ratio after quantization, the range and number of bits in the A/D converter can be determined.

Assume a B bit A/D converter with input range $-V_m$ to

$+V_m$. The voltage step associated with one quanta is

$$V_q = V_m/(2^{B-1}-1) \tag{3.81}$$

In order to accomodate variations in the input signal level, the peak input voltage should be somewhat larger than the expected rms input level. If the input waveform is Gaussian, setting V_m to correspond to 4 standard deviations or more reduces the probability of clipping to 4×10^{-5} or less. Even though the input is not Gaussian, setting V_m to 4 times the expected rms input level V_r is a reasonable procedure, lacking any additional information about the input waveform. Doing this

$$V_m = 4V_r \tag{3.82}$$

The signal power to quantization noise ratio is then

$$S/N = \frac{V_r^2}{V_m^2/(12(2^{B-1}-1)^2)} \tag{3.83}$$

$$= \frac{3(2^{B-1}-1)^2}{4} \tag{3.84}$$

Solving for B in terms of signal-to-noise ratio

$$B = 1 + \log_2\left(\frac{2(S/N)^{1/2}}{3^{1/2}} + 1\right) \tag{3.85}$$

Values of B for various values of S/N are

S/N	B
10 (10 dB)	3.2
10^2 (20 dB)	4.7
10^3 (30 dB)	6.2
10^4 (40 dB)	7.9
10^5 (50 dB)	9.5
10^6 (60 dB)	11.2

Because of power limitations in the source transducers, high signal-to-noise ratio receptions are not commonly encountered when making acoustic propagation measurements over moderate to long distances. Rarely does the signal to noise ratio at the A/D converter input exceed as low a value as 3 dB. Setting the rms noise level to a value corresponding to 3 or 4 bits of A/D converter activity allows the input noise to completely overwhelm the quantization noise. If the average noise level does not vary significantly, an A/D converter with relatively few (~8) bits can be used. If the noise level varies greatly, using a 12 bit converter or an automatic gain control provides additional range.

3.4.2 Quantization Effects

The input waveform is sampled at rate f_a. The spectrum of the sampled waveform consists of replications of the spectrum of the unsampled waveform replicated with spacing f_a. The effects of amplitude quantization are accounted for by adding white noise with spectral level

$$\frac{1}{12 f_a} \text{ quanta}^2/\text{Hz} \tag{3.86}$$

In most applications the quantization noise is much

smaller than the noise contained in the input waveform and can be neglected.

These samples are multiplied by samples of the complex rotator $\exp\{-j2\pi f_d t\}$ which have been formed at rate f_a using the direct digital synthesis method described in Section 3.2. The spectrum of the rotator sample values consists of a line at $-f_d$ nomalized to unit amplitude along with a set of low level lines generated by amplitude quantization and phase jitter. As in the Section 3.3, the amplitude quantization and phase jitter contributions are taken to be independent of each other.

The values of f_d and f_a are expressible as the ratio of two relatively prime integers p and q such that

$$\frac{f_d}{f_a} = \frac{p}{q} \tag{3.87}$$

In Section 3.2 it was shown that the amplitude quantization error energy is distributed among q equally spaced lines. The precise distribution can be computed for a given set of parameters. The worst case condition would be if all of this energy were concentrated in a single line. This line would have magnitude on the order of 0.41 to 0.71 quanta. In the best situation, this energy would be distributed evenly among all q lines. The magnitudes would be $q^{1/2}$ of the worst case level.

It was also shown in Section 3.2 that the phase jitter generated noise was distributed among q' lines where

$$\frac{Mf_a}{f_d} = \frac{Mp}{q} = \frac{p'}{q'} \tag{3.88}$$

The values of p' and q' are relatively prime.

The maximum magnitude of a phase noise generated line was found to be on the order of

$$\frac{\pi}{3^{1/2}} \frac{2^{B-1}-1}{M} \tag{3.89}$$

The spectrum at the multiplier output is the cyclic convolution of the spectrum of the sampled input waveform with the spectrum of the synthesized complex rotator values. Because convolution is a linear operation, the convolution can be expressed as the sum of the convolutions of the sampled input spectrum with the individual lines in the complex rotator spectrum.

The convolution of the spectrum of the sampled input with the line in the complex rotator spectrum at $-f_d$ causes the input spectrum to be rotated by an amount f_d toward the negative frequencies. The energy originally located at f_d is shifted to dc. This is the desired demodulation effect.

As a worst case bound, assume that all of the energy associated with amplitude quantization and phase jitter in the complex rotator samples is contained in a single line. The magnitude squared of this line is (relative to the magnitude of the line at $-f_d$)

$$\frac{1}{(2^{B-1}-1)^2} \{0.17 + \frac{((2^{B-1}-1)\pi)^2}{3M^2}\} \tag{3.90}$$

The parameter values used in the CEL beamforming system were B=16 and M=2000. These values yield a spurious response which is 60.8 dB down from the desired waveform.

3.4.3 The Demodulator Output Filter

The filter at the demodulator output is used to limit the bandwidth of the frequency shifted waveform prior to resampling at rate f_b. The primary purpose of this filter is to reduce the amount of energy replicated into the baseband region without significantly affecting the energy already there. It is expected that the resampled waveform

will be further filtered in the processing which follows the demodulator. This filter must be structured to readily be able to accomodate changes in the number of samples over which it operates (bandwidth) and the resample rate f_b.

One of the most simple filtering operations to analyze is the sliding summation. This operation is a discrete approximation to the sliding integrator described by the equation

$$\int_{t}^{t+T} f(a)da \tag{3.91}$$

The output of the sliding summer is the sum of the N most recent input values. The sum value is held constant between tics of the input clock and is resampled at rate f_b. The hardware needed to implement the sliding summer in its most general form requires N sample values of memory and is moderately complex. If the value of N is restricted to be less than or equal to $[f_a/f_b]$, then no input sample contributes to more than one resampled output value and the summation can easily be implemented using a single word of memory in "integrate and dump" form.

The "integrate and dump" approach sums only the N input values immediately preceeding the f_b clock tic. At the f_b clock tic, the filter value is read and the hardware sum register is reset to zero. If a delay can be tolerated, the f_b clock tic can be used to initiate the summation of N input values. In this case, the resulting output value corresponds to that of a sliding summation over N input values sampled at the f_b clock tics plus N/f_b seconds. This approach to filtering is easily implemented and was the one taken in the design of the CEL complex demodulator used to process the beamformer outputs.

The Kronecker delta impulse response of the sliding summation consists of N unit value outputs spaced $1/f_a$

seconds apart starting at t=0. The associated transfer function is

$$H(f) = \sum_{i=0}^{N-1} \exp\{-j2\pi fi/f_a\} \tag{3.92}$$

$$= \exp\{j\pi f(N-1)/f_a\} \frac{\sin(N\pi f/f_a)}{\sin(\pi f/f_a)} \tag{3.93}$$

With the exception of a phase factor to account for the extra delay, this is also the transfer function of the filter used in the integrate and dump approach (not including the effects of resampling at rate f_b).

The magnitude of the demodulator output filter transfer function is of the form shown in Fig. 3.13.

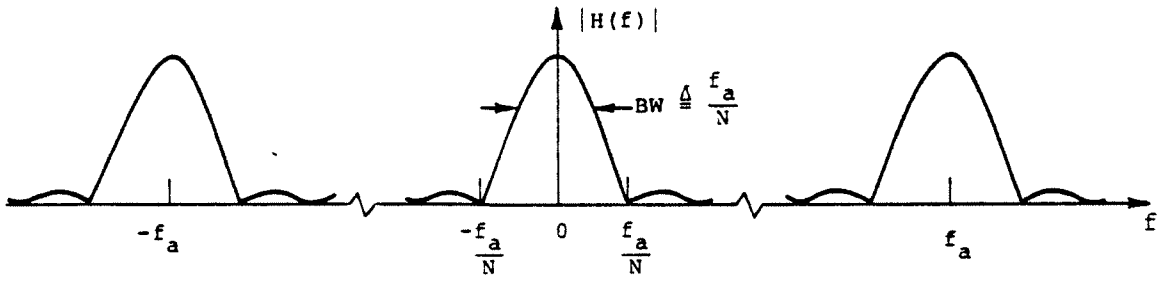


Fig. 3.13. Magnitude of demodulator output filter transfer function

The filter bandwidth is defined to be f_a/N. This value is close in value to the more traditional 3 dB bandwidth but is more easily calculated. Since the primary purpose of the demodulator is to band shift the input waveform, the filter bandwidth is generally set to be somewhat broader than that of the desired input waveform. Typically, the CEL demodulator is configured to use a bandwidth about 4 times that of the waveform of interest.

Because the waveform at the demodulator input is real valued, its spectrum is conjugate symmetric around dc. Any

energy in a band of frequencies centered at f_d is matched by energy centered at $-f_d$. The multiplication of the input waveform by the complex rotator at frequency $-f_d$ causes the energy originally centered at f_d to be recentered at dc (baseband) and the energy originally centered at $-f_d$ to be shifted to $-2f_d$. The rejection of the spectral component at $-2f_d$ can be maximized if the value of N can be chosen such that $-2f_d$ corresponds to one of the zero crossings of the output filter's transfer function.

Let Q_{FN} be the nominal Q desired of output filter

$$Q_{FN} = \frac{f_d}{\text{filter bandwidth}} = \frac{Nf_d}{f_a} \tag{3.94}$$

Solving for N gives

$$N = \frac{Q_{FN} f_a}{f_d} \tag{3.95}$$

Since N must be integer, the above value must be truncated or rounded.

$$N = \left[\frac{Q_{FN} f_a}{f_d}\right] \quad \text{(truncation)} \tag{3.96}$$

or

$$N = \left[\frac{Q_{FN} f_a}{f_d} + .5\right] \quad \text{(rounding)} \tag{3.97}$$

In either case, the actual filter Q

$$Q_{FA} = \frac{Nf_d}{f_a} \tag{3.98}$$

is equal to the nominally desired Q plus some error e.

$$-\frac{f_d}{f_a} < e \leq 0 \qquad \text{(truncation)} \qquad (3.99)$$

$$-\frac{f_d}{2f_a} < e \leq \frac{f_d}{2f_a} \qquad \text{(rounding)} \qquad (3.100)$$

Normalizing the demodulator output filter transfer function to unity at f=0 gives for frequencies within a small range, df, around $-2f_d$

$$|H(-2f_d+df)| = \left|\frac{\sin(N\pi(-2f_d+df)/f_a)}{N\sin(\pi(-2f_d+df)/f_a)}\right| \qquad (3.101)$$

Substituting for N in terms of Q_{FA}

$$|H(-2f_d+df)| = \left|\frac{\sin(Q_{FA}\pi(-2+df/f_d))}{Q_{FA}f_a\sin(\pi(-2f_d+df)/f_a)/f_d}\right| \qquad (3.102)$$

Assume f_d and df are restricted so that

$$(-2f_d+df)/f_a << 1 \qquad (3.103)$$

then, using the small angle approximation for the sine gives

$$|H(-2f_d+df)| \approx \left|\frac{\sin(Q_{FA}\pi(-2+df/f_d))}{Q_{FA}\pi(-2+df/f_d)}\right| \qquad (3.104)$$

Substitute

$$Q_{FA} = Q_{FN}+e \qquad (3.105)$$

$$|H(-2f_d+df)| \approx \left|\frac{\sin(\pi Q_{FN}(-2+df/f_d)+\pi e(-2+df/f_d))}{(Q_{FN}+e)(-2+df/f_d)}\right| \qquad (3.106)$$

If Q_{FN} is restricted to being a multiple of 1/2, the only effect of the $-2\pi Q_{FN}$ term in the numerator is to alter

the sign of the sine. Since absolute values are being taken, this term can thus be dropped. Also, if $f_d << f_a$, then e is small compared to Q_{FN}.

$$|H(-2f_d+df)| \approx \left|\frac{\sin(\pi Q_{FN} df/f_d - 2\pi e)}{\pi Q_{FN}(2-df/f_d)}\right| \qquad (3.107)$$

Given that Q_{FN} is a multiple of 1/2, this expression can be used to estimate the magnitudes of the spectral components around $-2f_d$ at the demodulator filter output.

The above equation describes a transfer function whose magnitude is of the form shown in Fig. 3.14.

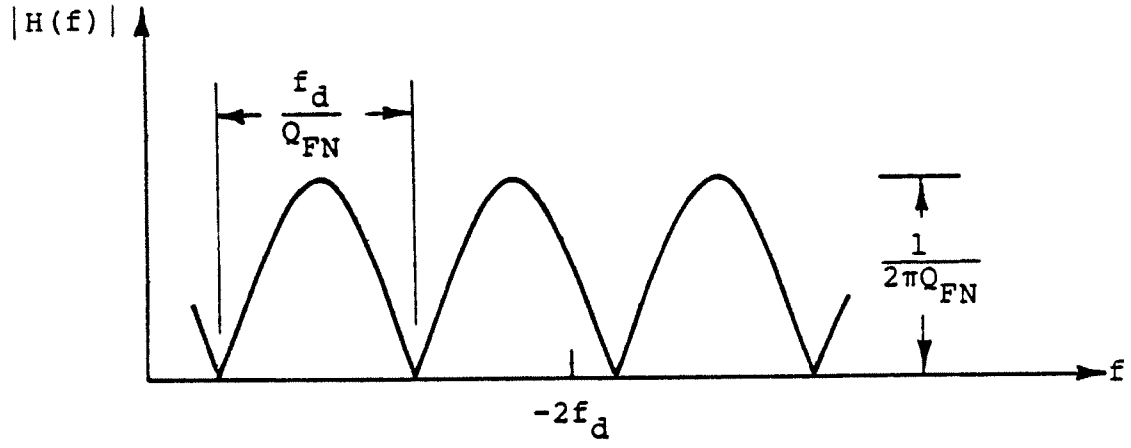


Fig. 3.14. Approximate transfer function magnitude centered around $-2f_d$

The value of e (defined in Eq. 3.105) accounts for the amount that the filter transfer function misses having a zero at $-2f_d$. The filter lobes are essentially sinusoidal in shape with amplitude $1/(2\pi Q_{FN})$ and seperation f_d/Q_{FN} Hz between zeros. A strong carrier at $-2f_d$ would be significantly attenuated.

Let the Q of the signal of interest be some multiple K of the filter Q.

$$Q_S = KQ_{FN} \qquad (3.108)$$

The value of K does not necessarily need to be an integer. The signal Q can also be written in terms of the signal bandwidth W_s as

$$Q_s = KQ_{FN} = f_d/W_s \tag{3.109}$$

The signal spectrum's nominal 3 dB down points are

$$\pm W_s/2 = \pm f_d/(2Q_s) = \pm f_d/(2KQ_{FN}) \tag{3.110}$$

Substituting these values for df gives

$$|H(-2f_d \pm W_s/2)| = \left|\frac{\sin(\pm\pi/(2/K)-2\pi e)}{\pi Q_{FN}(2\pm 1/(2KQ_{FN}))}\right| \tag{3.111}$$

This expression can be used to estimate the amount of attenuation present at the band edges of the input signal spectrum. For example, the 1981 Tomography Experiment parameters

$$\begin{aligned} Q_{FN} &= 3.5 \\ K &= 4 \\ e &= 0.0056 \end{aligned} \tag{3.112}$$

gives

$$|H(-2f_d)| = 0.0016 \qquad \text{(down ~56 dB)} \tag{3.113}$$

$$|H(-2f_d \pm W_s/2)| = 0.0174 \qquad \text{(down ~35 dB)} \tag{3.114}$$

In addition to suppressing the input signal component at $-f_d$, the filter at the demodulator output also acts to improve the signal to noise ratio. If the desired signal is band-pass of bandwidth less than f_a/N to which low-pass noise of bandwidth W_N has been added, then the demodulator filtering improves the signal-to-noise ratio at the output

by a factor of about

$$\frac{NW}{f_a} \tag{3.115}$$

In the CEL demodulator W=700, f_a=5000, giving an improvement factor of

$$\frac{7N}{50} \tag{3.116}$$

3.4.4 Effects of Sampling the Demodulator Output

The spectrum at the filter output prior to resampling at rate f_b consists of the input spectrum replicated with spacing f_a, shifted by $-f_d$ and multiplied by the output filter transfer function. Holding the filter output values constant between tics of the f_a clock effectively causes the replicated, shifted, filtered spectrum to be multiplied by $\sin(\pi f/f_a)/(\pi f)$. Resampling causes a replication of this weighted spectrum to occur with spacing f_b. In Chapter 2 it was shown that if the ratio f_a/f_b could be expressed as the ratio p/q of two relatively prime integers, then there were q frequencies in the range $-f_b/2$ to $+f_b/2$ where the original replications centered at multiples of f_a would be shifted to upon resampling.

The spectrum of the resampled waveform consists of the essentially unmodified "replication" at f=0 plus q spectral "lumps" caused by the f_a replications. The landing frequency upon resampling of the replication originally centered at kf_a is

$$f_k = f_b\left(\frac{kp}{q} - \left[\frac{kp}{q}+\frac{1}{2}\right]\right) \tag{3.117}$$

The magnitudes of the weighted f_a replications vary as 1/k due to the $\sin(\pi f/f_a)/(\pi f)$ weighting.

In order to get a feeling for the amount of "noise" introduced by the presence of the spurious replications, consider the "brick wall" spectrum of bandwidth W shown in Fig. 3.15a. The magnitude of the k-th replication has the form shown in Fig. 3.15b. The maximum magnitude of the k-th replication is approximately

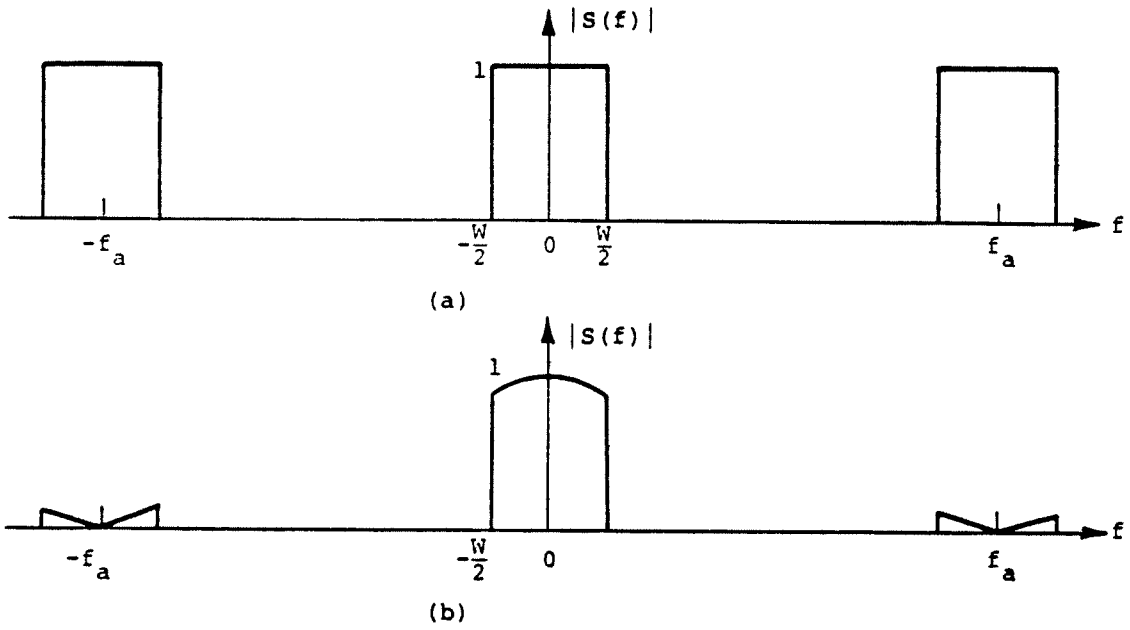


Fig. 3.15. Reconstruction of "brick wall" spectrum a) Spectrum prior to reconstruction b) Spectrum after reconstruction

$$\left|\frac{\sin(\pi(kf_a+W/2)/f_a)}{\pi(kf_a+W/2)/f_a}\right| \qquad (3.118)$$

where the k=0 replication has been normalized to have unit magnitude at f=0.

For $W/2 << f_a$, this is approximately

$$\left|\frac{W}{2kf_a}\right| . \qquad (3.119)$$

Assuming that the replications for $k \neq 0$ are made up of triangular components, the amount of energy contained in the k-th replication is

$$\frac{W^5}{48k^2 f_a^2} \tag{3.120}$$

Taking the view that the contributions from the replications are independent and thus their energies add gives a total "noise" energy of

$$\frac{W^5}{48 f_a^2} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{1}{k^2} \approx \frac{3.3W^5}{48 f_a^2} \tag{3.121}$$

The energy in the main band of interest is W. The resulting signal-to-noise ratio is

$$\frac{14.55\ f_a^2}{W^4} \tag{3.122}$$

The signal-to-noise ratio tends to increase as the square of the frequency of the initial sampling and to decrease as the fourth power of the input waveform bandwidth.

Assuming noise alone and signal alone inputs, the above expressions can be used to estimate the magnitude of the effect of resampling on the noise level and on signal generated self noise. For example, using the Tomography parameter values f_a=5000 Hz, noise bandwidth of 64 Hz and signal bandwidth of 16 Hz, the increase in noise level due to resampling is approximately

$$10\ \log\left(W + \frac{W^5}{14.55 f_a^2}\right) - 10\log(W) = 0.20\ \text{dB} \tag{3.123}$$

and the signal to self noise ratio is about 37.4 dB.

3.4.5 Scaling the Demodulator Output Values

Given a demodulator input consisting of signal plus noise where the rms noise level corresponds to 3 or more A/D converter bits, the effects of amplitude quantization on the signal are overwhelmed by the input noise and can be safely ignored. In this situation, one can quite readily process signals having magnitudes much smaller than the voltage corresponding to a single A/D converter quanta. If there were no input noise present, such signals would fail to cause any variation in the A/D converter output and would be suppressed. Let the demodulator input be a constant amplitude sinusoid at frequency f_d and the noise be "white" with sufficient bandwidth so that successive samples at rate f_a are independent of each other. In the summation, the signal component adds coherently while the noise component adds incoherently. Compared to the input, the signal power in the resulting sum is increased by a factor of N^2 and the noise variance is increased by a factor of N. The net result is an increase in the signal-to-noise ratio by a factor of N.

For a B bit A/D converter, C bit multiplier and a N value sum, if no bits are discarded in the demodulator, the number of bits D needed in forming the output summation should be such that

$$D > B+C-1+\log_2(N) \qquad (3.124)$$

The C bit sine and cosine values used to multiply the B bit A/D converter values consist of a sign bit and a C-1 bit fractional value. Thus, the low C-1 bits in the summation lie to the right of the binary point.

The rms level of the noise has been increased by a

factor of $\log_2(N)/2$. If the filter output is scaled at the same time that it is resampled, it is reasonable to keep about the same number of bits noise driven as there were at the A/D converter. This is accomplished by discarding (using some sort of rounding operation as described below) the low $\frac{1}{2}\log_2(N)+C-1$ bits. The number of bits A needed in the demodulator output samples is

$$A > B + \frac{1}{2}\log_2(N) \qquad (3.125)$$

It may be that even after the filtering in the demodulator output, the signal level is still less than the rms noise level. In this case, discarding additional bits might result in reducing the amount of activity in the low bits of the scaled values to such an extent that it is no longer possible to extract the desired signal in the signal processing steps following demodulation.

Having formed a D bit sum in the demodulator, it is to be converted to A bits upon resampling. If values are being kept in two's complement form, this conversion cannot be properly made by simply arithmetically shifting right and truncating the result. Nor can it be properly made by arithmetically shifting right and rounding the result. Both procedures have the problem that, given two values that sum to zero prior to scaling, the values do not necessarily sum to zero after scaling. The result is that the above two procedures introduce an undesired negative bias into the scaled values. This can be avoided if the following scaling procedure is used:

> Let r equal the number of bit positions that the values are to be scaled by.
>
> To positive numbers add the amount a where $0 \leq a < 2^r$.
> To negative numbers add the amount 2^r-1-a.
>
> Arithmetically shift the resulting values right r bit positions truncating the result.

It will now be shown that this procedure converts values which sum to zero prior to scaling into values which sum to zero after scaling.

Let p be a D bit positive number which is to be scaled by r bits. Its associated negative value in two's complement notation is 2^D-p.

For p the above procedure yields

$$[(p+a)/2^r] \tag{3.126}$$

For -p the above procedure yields

$$[(2^D-p+2^r-1-a)/2^r] = 2^{D-r}+1+[-(p+1+a)/2^r] \tag{3.127}$$

It can be shown that

$$[-(p+1+a)/2^r] = -1-[(p+a)/2^r] \tag{3.128}$$

giving for -p the result

$$2^{D-r}-[(p+a)/2^r] \tag{3.129}$$

which in two's complement notation represents the D-r bit value

$$-[(p+a)/2^r] \tag{3.130}$$

which is the negative of the value obtained for +p.

Various scaling schemes result as a takes on values from 0 through 2^r-1. For a=0, the procedure amounts to reducing the magnitudes of the values by r bits and then truncating. Using a equal to 2^{r-1} or $2^{r-1}-1$ corresponds to reducing the magnitudes of the values by r bits and then

rounding prior to truncating. The two values of r differ only in whether the fractional value of 1/2 rounds up or down. The value of $a=2^r-1$ causes only the value 0 to map into 0 upon scaling.

When scaling values, rounding yields lower computational noise than does truncation, thus values of a equal to 2^{r-1} or $2^{r-1}-1$ are good choices. Of these two values, $2^{r-1}-1$ is recommended since it has a slightly larger number of values which it maps into zero and it is felt that this extra touch of "squelch" is beneficial.

3.5 Square Wave Demodulator

If the waveform at the demodulator input can be suitably filtered, then a simplified version of the analog demodulator of Section 3.3 can be used. In this demodulator, the values of $\exp\{-j2\pi f_d t\}$ are replaced by a complex valued square wave csq(t) which is obtained by retaining only the polarity information contained in the real and imaginary parts of $\exp\{-j2\pi f_d t\}$. The real and imaginary parts of csq(t) take on the values +1 and -1 and are shifted 90° in phase relative to each other. Fig. 3.16. shows the real and imaginary parts of csq(t).

The Fourier series spectrum of csq(t) consists of a series of lines at frequencies $(4i-1)f_d$ with amplitudes

$$A_i = \frac{-4}{\pi(4i-1)} \qquad (3.131)$$

The spectrum of csq(t) is shown in Fig. 3.17a

If the demodulator input possesses a low-pass spectrum of bandwidth $2f_d$, then the resulting spectrum when this input is multiplied by csq(t) is as shown in Fig. 3.17b. There is negligible overlap between the spectral replication centered at $3f_d$ and the energy centered around zero Hz.

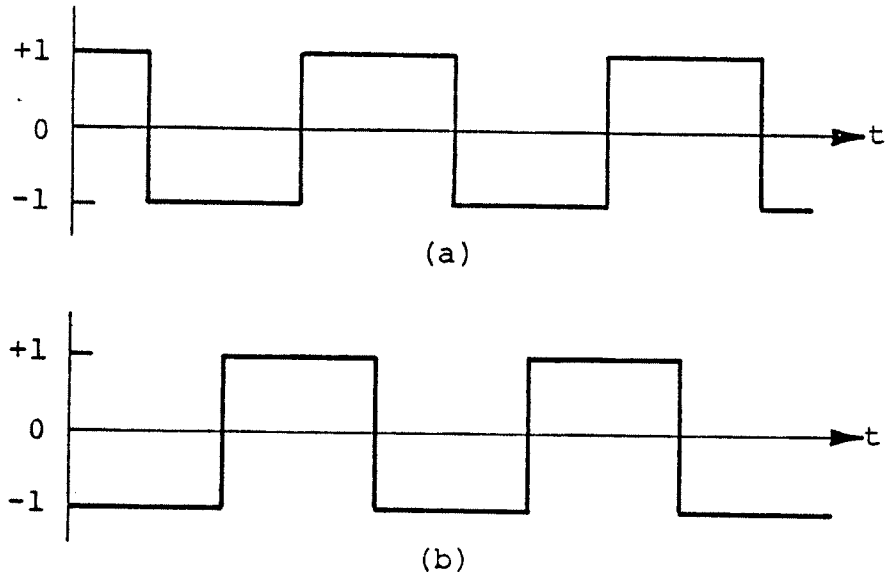


Fig. 3.16. Csq(t) a) Real part b) Imaginary part

Assuming a flat noise like input spectrum and allowing an input waveform bandwidth of as much as $4f_d$ results in a loss of only 1 dB in signal-to-noise ratio in frequency bands of width $2f_d$ or less centered around zero Hz.

Band-pass filtering the demodulator input rather than just simply low-pass filtering it has the advantage of reducing the dynamic range requirements on the components of the demodulator.

The multiplication of the demodulator input by plus and minus ones is easily implemented using operational amplifiers and CMOS switches. Fig. 3.18 contains the schematic diagram of an operational square wave demodulator. This circuit uses the cross-wired flip/flop method devised by T. G. Birdsall in 1963 for generating the required sine and cosine phased square waves. An input clock at $4f_d$ is required in order to band shift by $-f_d$. The output filters are simple single pole RC low pass filters. These were chosen because it is relatively easy to match RC time constants. Output bandwidths of 1, 0.1 and 0.01 Hz are provided.

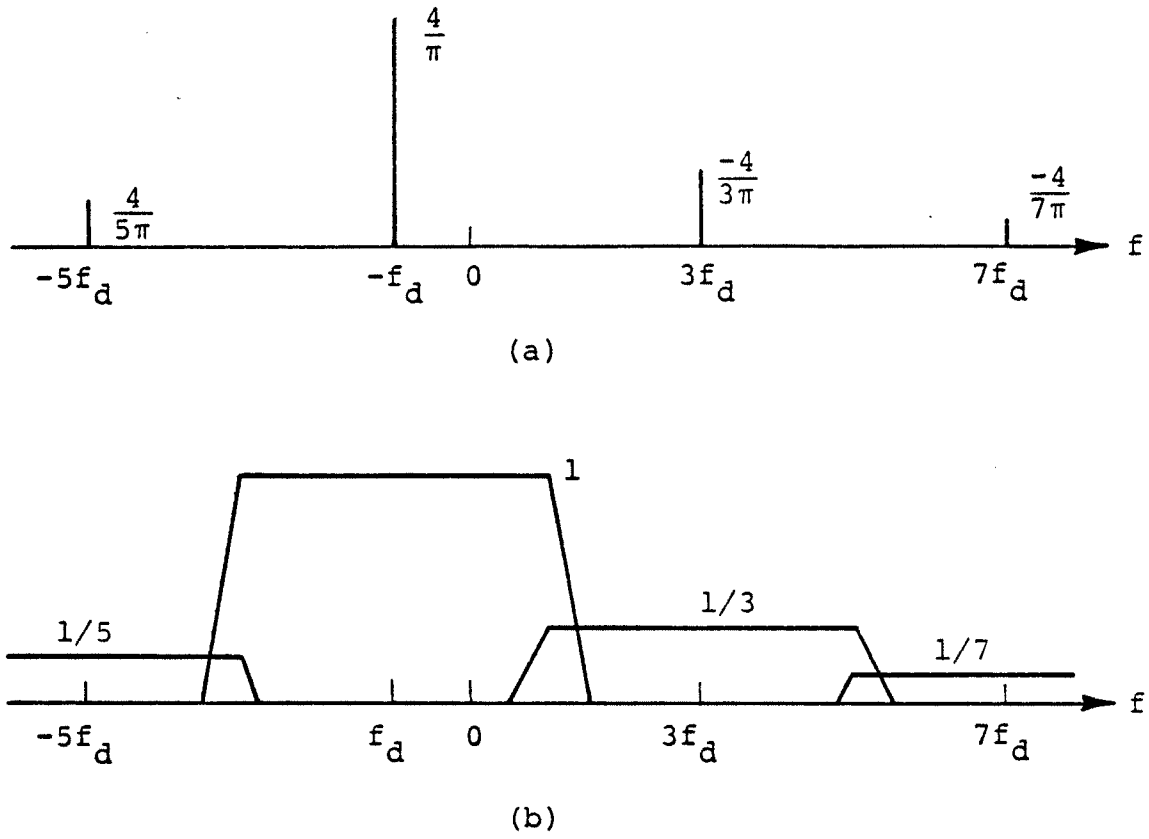


Fig. 3.17. Spectra associated with square wave demodulation a) Spectrum of the complex valued demodulating square wave b) Spectrum prior to output filtering assuming flat low pass input spectrum of bandwidth $2f_d$

3.6 Direct Sample Demodulation

This section describes and analyses a simplified version of the digital demodulator described in Section 3.4. The input waveform is restricted to being band-pass, centered around frequency f_d. Also, the input and output sampling rates are required to be simply related to f_d. The basic idea behind this method is to choose the input sample rate f_a based on f_d so that the only values of the rotator $\exp\{-j2\pi k f_d/f_a\}$ that are used are ± 1 and $\pm j$. These values are easily used to multiply the input data stream by either negating or not negating the sample values and making the

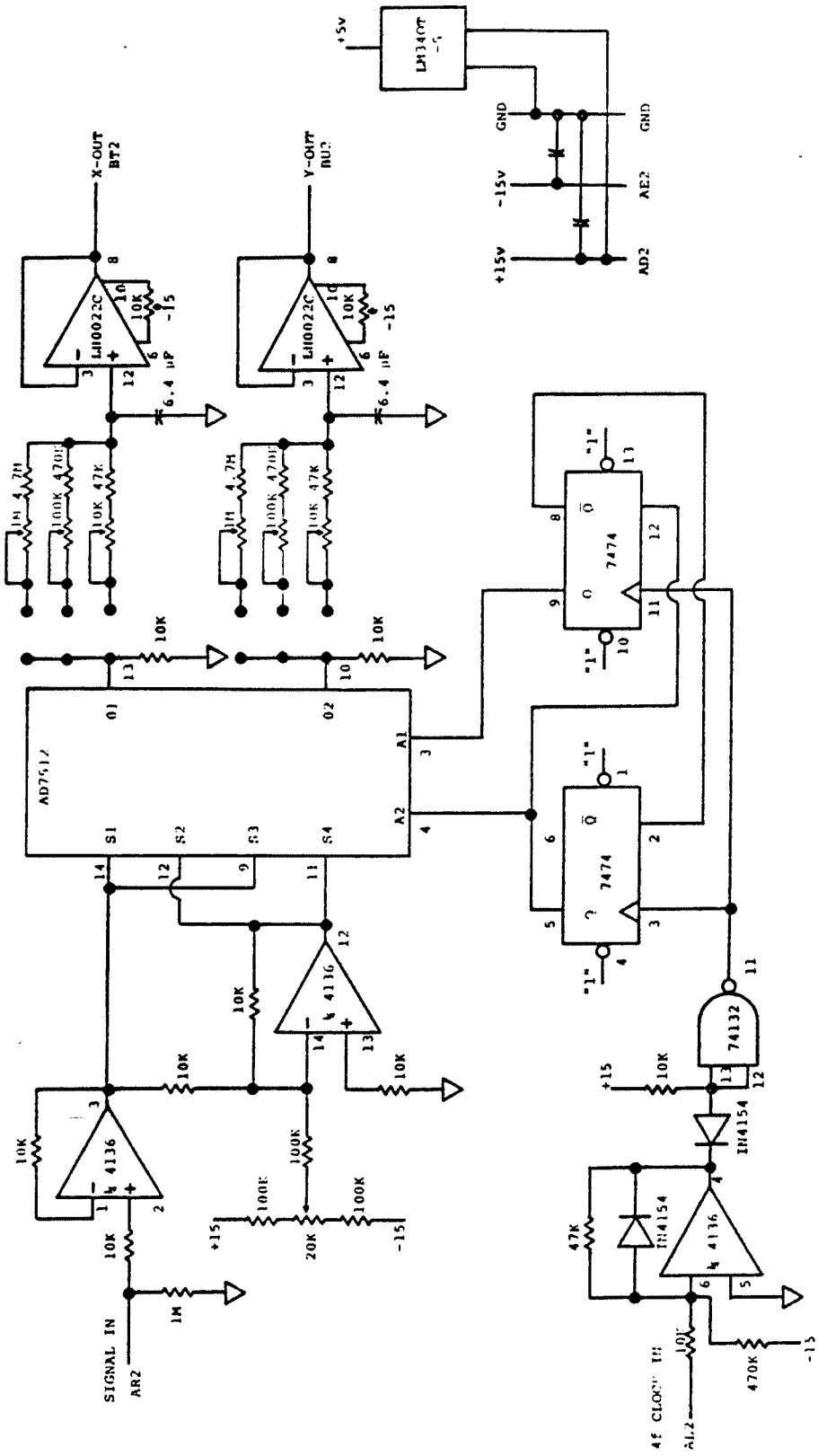


Fig. 3.18. Square wave demodulator circuit diagram

result purely real or imaginary. The filtering operation used to suppress the effects of the input energy at $-f_d$ and possibly any dc offset in the input waveform or the A/D converter is to be a sliding sum of N values. The output sample rate is chosen to be f_a/N, allowing the sampled sliding sum values to be computed using a simple accumulator into which N values are added. When the N-th value is summed, the result is latched into an output register and the accumulator is reset to zero.

The hardware and software needed to implement a direct sampling demodulator is minimal. An A/D converter added to a single chip microcomputer such as the INTEL 8748 can do the job nicely. The price paid for this simplicity is the need for a tuned band-pass front end and the ability to directly relate both the input and output sample rates to the desired demodulation frequency. There are many applications where this is possible.

3.6.1 Sampling a Band-Pass Waveform

Consider a band-pass waveform with center frequency f_d and bandwidth W_B. Sampling this waveform at rate f_a causes spectral replications of the positive frequency component to be centered at frequencies

$$k_1 f_a + f_d \qquad -\infty < k_1 < \infty \tag{3.132}$$

Similarly, the negative frequency component is replicated centered about frequencies

$$k_2 f_a - f_d \qquad -\infty < k_2 < \infty \tag{3.133}$$

If no replication of a positive frequency component is to exactly coincide with a replication of a negative frequency component, then for all k_1 and k_2

$$k_1 f_a + f_d \neq k_2 f_a - f_d \tag{3.134}$$

This requires that

$$2f_d \neq \text{integer} * f_a \tag{3.135}$$

Let

$$f_d = af_a/2 \qquad a>0 \text{ and not integer} \tag{3.136}$$

The positive frequency sideband replications land at $(k_1+a/2)f_a$ and the negative frequency sideband replications land at $(k_2-a/2)f_a$. Let int(a) be the integer part of a and fract(a) be the fractional part. If int(a) is even, the replications closest to the origin are

$$\begin{array}{ll} \text{fract}(a)f_a/2 & \text{positive component} \\ -\text{fract}(a)f_a/2 & \text{negative component} \end{array} \tag{3.137}$$

If int(a) is odd, the replications closest to the origin are

$$\begin{array}{ll} (\text{fract}(a)-1)f_a/2 & \text{positive component} \\ (1-\text{fract}(a))f_a/2 & \text{negative component} \end{array} \tag{3.138}$$

These two cases are pictured in Fig. 3.19.

The separation between the two components is maximum if fract(a)=1/2. In this case, the lowest possible sample rate which gives negligible overlap between the components is $f_a=2W_B$. This limit for band-pass waveforms is termed the band-pass sampling theorem [26].

3.6.2 Frequency Shifting

Consider the digital demodulator shown in Fig. 3.9. If f_a and f_d are related by relatively prime integers p and q

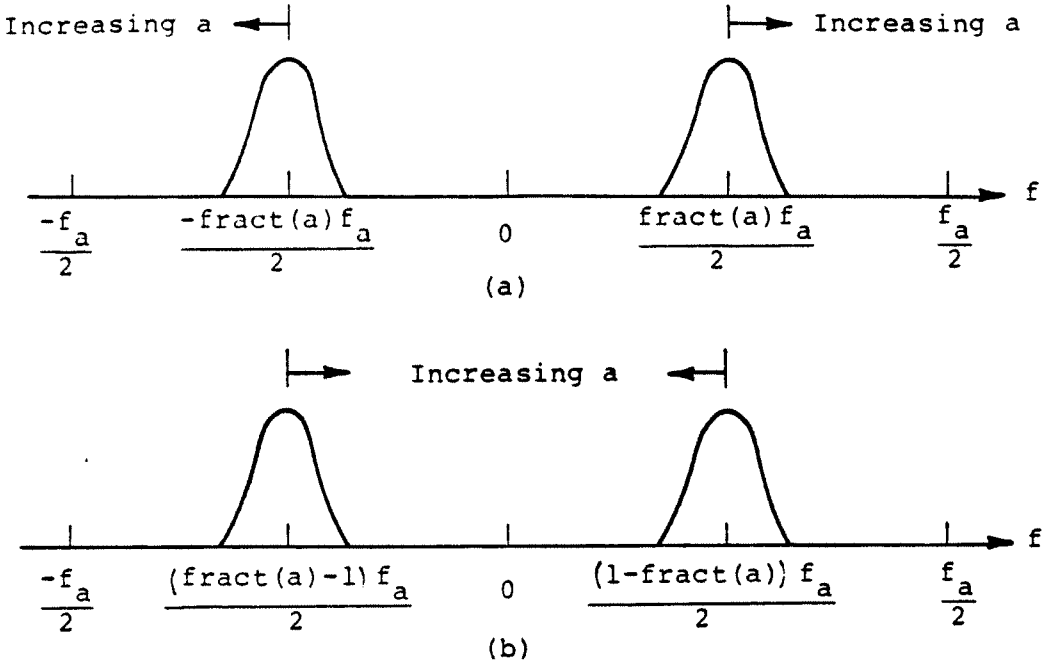


Fig. 3.19. Replicated spectral components in the range from $-f_a/2$ to $+f_a/2$ a) Even int(a) b) Odd int(a)

as

$$f_a = qf_d/p \tag{3.139}$$

then the required values of the complex rotator indexed by i are

$$\exp\{-j2\pi ip/q\} \tag{3.140}$$

If q=4, the values taken on by the rotator are in the set {1,-j,-1,j} with the ordering being determined by the value of p.

Since p and q are relatively prime if q equals 4, this implies that p is odd. Thus,

$$\frac{p}{2} f_a = 2f_d \tag{3.141}$$

giving a=p/2 with fract(a)=1/2. If the input's band-pass components are centered at f_d and $-f_d$ and have bandwidth W_B, the smallest value of f_a that should be used is $2W_B$. Prudence dictates that a value at least 3 times $2W_B$ should be used.

The Q of the input waveform is

$$Q_i = f_d/W_B \tag{3.142}$$

Then, since $f_a>2W_B$ and $f_a=4f_d/p$, the value of p is restricted

$$p<2Q_i \tag{3.143}$$

If a safety factor of 3 is used, the restrictions on p for q=4 are

$$\begin{aligned} &p \text{ is odd} \\ &p < 2Q_i/3 \end{aligned} \tag{3.144}$$

If p=1, the samples of the input waveform are multiplied by the sequence of numbers

$$1, -j, -1, j, 1, -j, -1, j, \ldots \tag{3.145}$$

If p=3, the input samples are multipled by

$$1, j, -1, -j, 1, j, -1, -j, \ldots \tag{3.146}$$

3.6.3 Filtering and Resampling

The purpose of the filter at the demodulator output is to reject the $-2f_d$ component in the frequency shifted waveform and to somewhat limit the noise bandwidth. It is assumed that the demodulator is primarily a frequency shifting device and that any specialized filtering

operations are performed in the processing steps applied to the demodulated sample values.

The easiest filter to implement at the demodulator output is the simple sum and dump filter. This filter is used to sum N suitably weighted input samples. Once N values have been summed, the result is made available to the later processing stages and the sum accumulator set to zero. The resulting output sample rate is then exactly 1/Nth the input sample rate.

The filter transfer function is

$$H(f) = \exp(j\pi(N-1)f/f_a)\frac{\sin(\pi Nf/f_a)}{\sin(\pi f/f_a)} \qquad (3.147)$$

In order to reject the $-2f_d$ component, requires that

$$|H(-2f_d)| = \left|\frac{\sin(2\pi Nf_d/f_a)}{\sin(\pi f/f_a)}\right| = 0 \qquad (3.148)$$

The values of f_d and f_a were restricted in the last section to be such that

$$f_d = pf_a/4 \qquad (3.149)$$

where p is an odd integer. Substituting

$$|H(-2f_d)| = \left|\frac{\sin(\pi Np/2)}{\sin(\pi p/2)}\right| = |\sin(\pi Np/2)| \qquad (3.150)$$

This is equal to zero if Np is even. Since p is odd, the value of N must be even in order to place a zero of the filter transfer function at $-2f_d$.

In many instances, either the input filter and/or the A/D converter used to digitize the input waveform has a dc offset associated with it. After the frequency shifting, any such dc component is shifted to $-f_d$. This line can be

suppressed by selecting N such that $|H(-f_d)|=0$.

$$|H(-f_d)| = \left|\frac{\sin(\pi Np/4)}{\sin(\pi p/4)}\right| = 2^{1/2}|\sin(\pi Np/4)| \quad (3.151)$$

This equals zero if N is a multiple of 4.

The spectrum of the resampled waveform at the filter output can be determined in terms of the spectrum at the filter output assuming a sliding summation filtering operation over the most recent N input samples at rate f_a.

Let S(f) be the spectrum in the range from $-f_a/2$ to $+f_a/2$ at the filter output. The replicated spectrum is

$$f_a \sum_{i=-\infty}^{\infty} S(f+if_a) \quad (3.152)$$

The spectrum upon reconstruction using the zero order hold is

$$\frac{\sin(\pi f/f_a)}{\pi f/f_a} \sum_{i=-\infty}^{\infty} S(f+if_a) \quad (3.153)$$

The spectrum upon resampling at rate f_b is

$$f_b \sum_{k=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \frac{\sin(\pi(f+kf_b)/f_a)}{\pi(f+kf_b)/f_a} S(f+if_a+kf_b) \quad (3.154)$$

As a check on the above equation, f_b will be set to be an integer sub-multiple of f_a. The above equation should simplify to yeild the same spectrum that would have been obtained if the original waveform had been sampled at rate f_b.

Let $f_b=f_a/N$ where N is an integer and interchange the order of summation to get

$$f_b \sum_{i=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \frac{\sin(\pi(f+kf_b)/(Nf_b))}{\pi(f+kf_b)/(Nf_b)} S(f+(iN+k)f_b) \quad (3.155)$$

Let $k'=k+iN$ substitute and drop the primes and again interchange the order of summation to obtain

$$f_b \sum_{k=-\infty}^{\infty} S(f+kf_b)\sin(\pi(f+kf_b)/(Nf_b)) \sum_{i=-\infty}^{\infty} \frac{(-1)^i}{\pi(f+kf_b)/Nf_b-\pi i} \quad (3.156)$$

The sum on i is equal to $1/\sin(\pi(f+kf_b)/Nf_b)$ giving the result for the spectrum upon resampling with $f_b=f_a/K$

$$f_b \sum_{k=-\infty}^{\infty} S(f+kf_b) \quad (3.157)$$

When the ratio of sample rate to resample rate is integer, the spectrum of the resampled waveform can be found by simply replicating with spacing f_b the spectral region from $-f_a/2$ to $+f_a/2$ in the original sampled spectrum.

CHAPTER 4

TIME COMPRESSION USING LINEAR MAXIMAL SEQUENCES

The acoustic transducers used in making propagation measurements in the ocean are electrically and mechanically limited in the amount of peak power that they can put into the water. This in turn limits the amount of energy that a single pulse of a given duration can contain. Because signals propagating in the ocean experience spreading and absorbtion losses, there is a limit to the range at which a single pulse can be detected before it becomes lost in the noise.

One method of extending this range is to transmit a periodic pulse and coherently average the receptions from several periods. The received signal-to-noise ratio increases directly with the number of periods averaged together. The number of periods that can be usefully averaged together is limited by clock drift, mooring motion and the stability of the path structure. The time separation between pulses should be greater than the expected arrival spread.

The energy contained in the transmitted waveform can be further increased thru the use of a periodic pulse compression waveform instead of a single periodic pulse. Each period of the pulse compression waveform consists of a set of L individual pulses which have been amplitude and/or phase modulated. The modulation is carefully chosen so that at the receiver, it is possible to collapse the energy contained in the L pulse positions into a more energetic equivalent single pulse. If this is correctly done, there

is an increase in the signal-to-noise ratio which is proportional to L.

In section 4.1, a time spread waveform is constructed by weighting L shifted versions of a periodic pulse and summing. The shifts used are multiples of the period divided by L. The resulting waveform is uniformly sampled ML times per period. It is found that the DFT of the sampled waveform decomposes into the product of the DFT of the sampled base pulse waveform and the DFT of the weights. This factoring allows the choice of the weights to be studied independently of the chosen pulse waveshape. A quantity, the signal-to-noise ratio improvement factor (SNRIF), is defined which relates the improvement in the signal-to-noise ratio provided by the pulse compression waveform over the single periodic pulse. It is shown that the SNRIF is maximized if the lines in the DFT of the weights have equal magnitudes. By restricting the weight values to a maximum magnitude of one in order to reflect the peak power transducer limitation, the maximum improvement in the signal-to-noise ratio over a single periodic pulse is found to be L. Pulse compression processing can be performed in the time domain as well as in the frequency domain. Using the inverse transform of the reciprocal DFT of the weights, it is shown how the equivalent time operations can be performed. Because the waveform is sampled M times per pulse interval, the time processing can be divided into M independent operations on interleaved data sets. The time compressed time waveform is constructed by merging the individual results. The interleaved data sets used in this approach are equivalent to those obtained by sampling the reception once per digit interval. Because of this, it is reasonable to study the effects of time domain processing on the basis of one sample per pulse interval. This is the approach taken in the remainder of the chapter.

Section 4.1 establishes conditions that must be satisfied in order to achieve pulse compression and provides

a means of evaluating the effectiveness of sets of pulse weightings. Section 4.2 considers the use of binary linear maximal sequences as a basis for selecting weight values. The section starts with a brief introduction of the properties of interest and describes two canonical methods of generation. It is shown how these sequences can be used to amplitude and/or phase modulate a carrier. The autocorrelation function of the resulting baseband representation is derived. An expression for the SNRIF is derived in terms of the autocorrelation function. In order to compare different modulation schemes, the concept of efficiency is introduced. This is the amount of improvement in signal-to-noise ratio obtained by a given modulation as compared to the maximum possible for the same amount of energy per period. Several popular modulation schemes are compared on the basis of their SNRIF and efficiency. Because the length of linear maximal sequences is one less than an integer power of two, it is awkward to pulse compress waveforms based on them using frequency domain processing techniques. Using sequences whose lengths are integer powers of two would allow the use of fast DFT algorithms to perform the pulse compression processing in the frequency domain. One method of generating sequences of the desired length is to add an additional digit to linear maximal sequences. A performance measure termed "non-flatness loss" is introduced and used to evaluate the performance of these augmented sequences. Results from a computational survey of such augmented sequences are included.

Section 4.3 develops algorithms for correlation processing of transmissions that have been time spread using linear maximal sequences. The two level autocorrelation function of these sequences allows the development of correlation algorithms which primarily use add operations.

The material in Section 4.4 consists of odds and ends concerning the use of time spread waveforms. Section 4.4.1

considers the effects of neglecting the correction of the dc line using correlation processing. Section 4.4.2 considers the effects of incorrectly matching the modulation angle direction when using complementary phase modulation. Section 4.4.3 contains a list of correction factors useful for processing CPM based waveforms. Section 4.4.4 lists the number of times that ones and zeros line up when forming linear maximal based cross correlations.

4.1 Time Spreading and Its Removal

The class of time spread waveforms considered here are constructed by summing L weighted shifted versions of a periodic pulse waveform p(t). The weights $\{w_i\}$ are restricted to having magnitudes less than or equal to one in order to account for the transducer peak power limit. The shifts are constrained to be multiples of T seconds with the overall period being LT seconds.

Consider the noisy reception

$$r(t) = \sum_{i=0}^{L-1} w_i p(t-iT) + \text{noise}(t) \qquad (4.1)$$

The noise is modeled as being a stationary random process with white spectrum and sample variance V_n. Samples are taken uniformly at a rate of M samples every T seconds. The resulting sample values are

$$r(n\tfrac{T}{M}) = \sum_{i=0}^{L-1} w_i p(n\tfrac{T}{M} - iT) + \text{noise}(n\tfrac{T}{M}) \qquad (4.2)$$

It is convenient to express all times in terms of the time step index $\frac{T}{M}$ and suppress the actual time step value. This gives

$$r(n) = \sum_{i=0}^{L-1} w_i p(n-iM) + \text{noise}(n) \qquad (4.3)$$

Consider a data set made up of ML samples. It is assumed that the signal is active while the samples are being taken. This is to avoid end effects at signal turn on and turn off. In practice, the time interval to be processed is surrounded by guard zones where the signal is active.

Taking the DFT of ML samples of r(n) gives

$$R(k) = \frac{1}{ML} \sum_{n=0}^{ML-1} \sum_{i=0}^{L-1} w_i p(n-iM) \exp(-j2\pi \frac{nk}{LM}) + \text{noise}_1 \qquad (4.4)$$

As the analysis proceeds, the form of the noise contribution changes. These forms will be indexed at each step. At the end of the analysis, the expected value of the magnitude squared of the final noise term will be evaluated and used to determine the resulting signal-to-noise ratio.

Let $n'=n-iM$, substitute, drop the primes and re-group terms to obtain

$$R(k) = (\frac{1}{ML} \sum_{n=0}^{ML-1} p(n) \exp(-j2\pi \frac{nk}{LM}))(\sum_{i=0}^{L-1} w_i \exp(-j2\pi \frac{ik}{L})) + \text{noise}_1 \qquad (4.5)$$

$$= LP(k)W(k) + \text{noise}_1 \qquad (4.6)$$

The signal portion of the spectrum factors into the product of the transform of the periodic pulse and the transform of the weight values. This factoring allows the effect of the weights to be studied independently of the pulse waveshape.

Divide the DFT'd data by W(k). Remember that the w_i

have to be chosen so that $W(k) \neq 0$ for any k.

$$\frac{R(k)}{W(k)} = LP(k) + \text{noise}_2 \tag{4.7}$$

Take the inverse DFT to obtain sample values of the pulse compressed result

$$DFT^{-1}\left(\frac{R(k)}{W(k)}\right) = Lp(n) + \text{noise}_3 \tag{4.8}$$

The noise contributions have been DFT'd, the resulting DFT divided by W(k) and this result then inverse transformed. Assuming the noise values are independent and identically distributed with sample variance V_n, the variance of the processed noise samples can be shown to be equal to

$$\frac{V_n}{L} \sum_{k=0}^{L-1} \frac{1}{W(k)W^*(k)} \tag{4.9}$$

The signal-to-noise ratio of the pulse compressed sample values is

$$\frac{S}{N}(n) = \frac{L^2 p(n)p^*(n)}{\frac{V_n}{L} \sum_{k=0}^{L-1} \frac{1}{W(k)W^*(k)}} \tag{4.10}$$

Since the signal-to-noise ratio of the sample values of a single pulse is

$$\frac{S}{N}(n) = \frac{p(n)p^*(n)}{V_n} \tag{4.11}$$

the improvement in the signal-to-noise ratio due to the time spreading and collapsing is

$$SNRIF = \frac{L^3}{\sum_{k=0}^{L-1} \frac{1}{W(k)W^*(k)}} \tag{4.12}$$

The signal-to-noise ratio improvement factor is maximized by choosing the w_i to minimize the value of the denominator. Keep in mind that the maximum magnitudes of the w_i were restricted to be less than or equal one. This was done because the transducer was assumed to be peak power limited.

Using Parseval's theorem to relate the energy in the weights and their transform gives

$$\frac{1}{L} \sum_{i=0}^{L-1} w_i w_i^* = \sum_{i=0}^{L-1} W(i)W^*(i) = E \leq 1 \tag{4.13}$$

Subject to this constraint, the {W(k)} which minimize Eq. 4.9 is desired. This is equivalent to minimizing the function

$$f(x_0,\ldots,x_{L-1}) = \sum_{i=0}^{L-1} \frac{1}{x_i^2} \tag{4.14}$$

subject to the constraint

$$g(x_0,\ldots,x_{L-1}) = -E + \sum_{i=0}^{L-1} x_i^2 = 0 \tag{4.15}$$

Using the method of Lagrange multipliers [32] form

$$F(x_0,\ldots,x_{L-1}) = f(.) + \lambda g(.) \tag{4.16}$$

$$= -\lambda E + \sum_{i=0}^{L-1} \left(\frac{1}{x_i^2} + \lambda x_i^2\right)$$

Taking partial derivatives and setting them equal to zero

$$\frac{\partial F}{\partial x_i} = \frac{-2}{x_i^3} + 2\lambda x_i = 0 \tag{4.17}$$

Which gives $\lambda = \frac{1}{x_i^4}$. This can only be true if the x_i^2 are all equal. In this case then

$$W(k)W^*(k) = x_i^2 = \frac{E}{L} \tag{4.18}$$

The signal-to-noise ratio improvement factor is maximized if the spectrum of the $\{w_i\}$ is white. Any deviation from a uniform spectrum reduces performance. The value of the maximized signal-to-noise ratio improvement factor is

$$SNRIF_{max} = L \tag{4.19}$$

The signal-to-noise ratio increases linearly with the number of digits in the pulse compression code.

Given that $W(k)W^*(k) = \frac{1}{L}$, any number of sets of w_i can be found that yield the maximum possible improvement in signal-to-noise ratio. In practice, trade-offs are made between ease of waveform generation, processing and other experimental considerations when choosing the $\{w_i\}$ to be used. The class of waveforms, binary linear maximal sequences, has been utilized with great success and is the topic of the next section.

In the above analysis, pulse compression was achieved by taking the DFT of one period of the reception and dividing by the DFT of the sequence of pulse weights $\{w_i\}$. Pulse compression can also be achieved through the equivalent time operation of convolving the reception with a

stored reference waveform. Depending on how the reference waveform is defined, this process is also called replica cross correlation. Normally, because of the existence of fast DFT techniques which make frequency domain operations more time efficient than the corresponding time domain operations, one seeks to avoid working in the time domain. However, there are conditions under which time domain operations can compete favorably with frequency domain operations. In particular, when the $\{w_i\}$ correspond to the digits in an unmodified binary linear maximal sequence, the correlation of a reception against the stored reference can be implemented by "multiplying" the reception samples by zeros and ones or plus and minus ones. A separate calculation may be needed in order to correct the dc line. In many cases, the size of the program used to perform the correlation is more compact than the code needed to form DFT's and the local reference can be generated as needed.

Earlier in this section, it was shown that in the noise free case the DFT of the sampled reception can be written as the product of the DFT of the samples of the basic pulse waveform, p(t), and the DFT of the weighting sequence used for time spreading.

$$R(k) = LP(k)W(k) \tag{4.20}$$

In order to time compress the reception, a filter whose transfer function is

$$\frac{1}{W(k)} = \frac{W^*(k)}{|W(k)|^2} \tag{4.21}$$

is used. Recall that the w_i are restricted such that none of the W(k) have zero magnitudes.

Let

$$W'(k) = \frac{W(k)}{|W(k)|^2} \tag{4.22}$$

and let w_i' represent the values taken on by the inverse transform of $W'(k)$.

The DFT of the output of the pulse compression filter in the noise free case can be written as

$$H(k) = \frac{1}{L}\sum_{i=0}^{L-1} w_i'^{*}\exp(+j2\pi\frac{ik}{L})\frac{1}{ML}\sum_{n=0}^{ML-1} r(n)\exp(-j2\pi\frac{nk}{LM}) \tag{4.23}$$

$$= \frac{1}{ML^2}\sum_{i=0}^{L-1} w_i'^{*}\sum_{n=0}^{ML-1} r(n)\exp(-j2\pi\frac{(n-iM)k}{LM}) \tag{4.24}$$

Let $n'=n-iM$, substitute, drop the primes and use the fact that the transmitted waveform is periodic

$$H(k) = \frac{1}{ML}\sum_{n=0}^{ML-1}\frac{1}{L}(\sum_{i=0}^{L-1} w_i'^{*}r(n+iM))\exp(-j2\pi\frac{nk}{LM}) \tag{4.25}$$

That is, $H(k)$ is the DFT of the sequence

$$h(n) = \frac{1}{L}\sum_{i=0}^{L-1} w_i'^{*}r(n+iM) \tag{4.26}$$

The indices of the input sample values are stepped by M in the above correlation sum. Since there are ML sample values per data set and the indices are being computed modulo ML, there are in effect M cyclic correlations of length L being formed.

Let $n=qM+m$ when $0\leq q\leq L-1$ and $0\leq m\leq M-1$. Then $h(n)$ can be written as

$$h(qM+m) = \frac{1}{L}\sum_{i=0}^{L-1} w_i'^{*} r((q+i)M+m) \qquad (4.27)$$

This equation verifies that the ML value replica cross correlation can be implemented as M individual replica cross correlations of L values. Define

$$\begin{aligned} h_m(q) &= h(qM+m) \\ r_m(q) &= r(qM+m) \end{aligned} \qquad (4.28)$$

Then

$$h_m(q) = \frac{1}{L}\sum_{i=0}^{L-1} w_i'^{*} r_m(q+i) \qquad \begin{matrix} 0 \le q \le L-1 \\ 0 \le m \le M-1 \end{matrix} \qquad (4.29)$$

The steps involved in using replica cross correlation processing for pulse compression can be implemented as follows:

1. Divide the ML element array of reception samples {r(i)} into M arrays of length L. The m-th array consists of values {r(qM+m)}, $0 \le q \le L-1$, $0 \le m \le M-1$. This division can be accomplished by leaving the data in its original form and properly computing index values in the following operations.

2. Correlate each of the M arrays against the sequence $\{w_i'^{*}\}$.

3. Assemble the ML value array of pulse compressed values {h(n)} from the M correlation arrays by reversing the grouping process used in step 1.

Because no two of the M cross correlations share any of the samples of the reception, the correlations can be performed essentially in place using a single auxiliary array of length L.

If the $\{w_i\}$ used to time spread the transmitted waveform are based on the use of the binary linear maximal sequences described in the next sections, the multiplication in the correlation can be made as simple as multiplying by zero and one. This converts the correlation into simple summations. An adjustment to the dc level may also be required.

Since there are (L+1)/2 ones in a linear maximal sequence, the number of add operations (usually complex valued) needed to accomplish pulse compression is

$$ML(L+1)/2 \qquad (4.30)$$

For the Ocean Acoustic Tomography Experiment, L=127 and M=4. Thus, 32512 complex additions are required per multipath response computation.

4.2 Linear Maximal Sequences

In the previous section, it was shown that a time spread waveform made up of weighted shifted versions of a fixed periodic pulse waveshape could be collapsed into a single more energetic version of that waveshape if the weights were properly chosen. It was found that the spreading/collapsing process is solely a function of the weights and is independent of the chosen pulse waveshape. No consideration was given as to how to choose the weights. This section describes how a class of binary bit patterns called linear maximal sequences can be used to generate such weights.

A binary sequence $\{a_i\}$ is said to be a (binary) linear maximal sequence of degree n if it satisfies a linear

recurrence

$$a_k = \sum_{i=1}^{n} c_i a_{k-i} \pmod 2 \qquad (4.31)$$

and has period $L=2^n-1$ [17].

Linear maximal sequences are also referred to by the names m-sequences, and pseudo-random (or PR) sequences. The properties of these sequences have been studied by a number of investigators [17], [16], [33], [18]. Linear maximal sequences can be generated using an n-stage binary register memory imbedded in a feedback network made up of modulo two adders (exclusive-or gates). If the feedback network is properly structured, clocking the registers causes their contents to cycle thru the states corresponding to the counts 1 thru 2^n-1 (but not in order). The sequences of ones and zeros contained in the individual registers will be shifted versions of each other and will be periodic with period 2^n-1. Many different feedback arrangements can be used to produce a given sequence pattern. Some of the more regular structures have been termed "canonical" and have had their properties tabulated [34]. Regardless of the method of generation, a sequence is uniquely defined by the coefficients of the recurrence relation (law) that it satisfies. Lists of laws corresponding to linear maximal sequences are available for values of n ranging from 2 thru 19 [35],[36]. For a given value of n having a "seed" linear maximal sequence law, it is possible to generate the other linear maximal sequence laws corresponding to that value of n. The number of linear maximal sequences that exist for a given value of n increases rapidly as n increases.

Linear maximal sequences possess many interesting and important properties. The two properties that are exploited in the remainder of this chapter are:

A linear maximal sequence of degree n contains 2^{n-1} ones and $2^{n-1}-1$ zeros.

Let $\{s_i\}$ be the sequence of values obtained by taking a binary linear maximal sequence and replacing the binary ones with the value -1 and binary zeros with the value +1. The unnormalized periodic autocorrelation function of $\{s_i\}$ is

$$C(k) = \sum_{i=0}^{L-1} s_i s_{i-k} = L \quad \text{if } k=0$$
$$= -1 \quad \text{if } k\neq 0 \tag{4.32}$$

Two of the most commonly used shift register generator configurations are shown in Fig. 4.1. These are termed the simple shift register generator (SSRG) and the modular shift register generator (MSRG). The relation between the input and output of the delay stages making up these generators can be described using the shift operator z which is defined

$$u(j) = zu(j-1) \tag{4.33}$$

Using the shift operator, the recursion relations governing the generated sequences are readily determined [36].

SSRG

$$\left(\sum_{i=0}^{n} s_{n-i}^{s} z^i\right) u_0(j) = 0 \tag{4.34}$$

MSRG

$$\left(\sum_{i=0}^{n} s_i^m z^i\right) u_0(j) = 0 \tag{4.35}$$

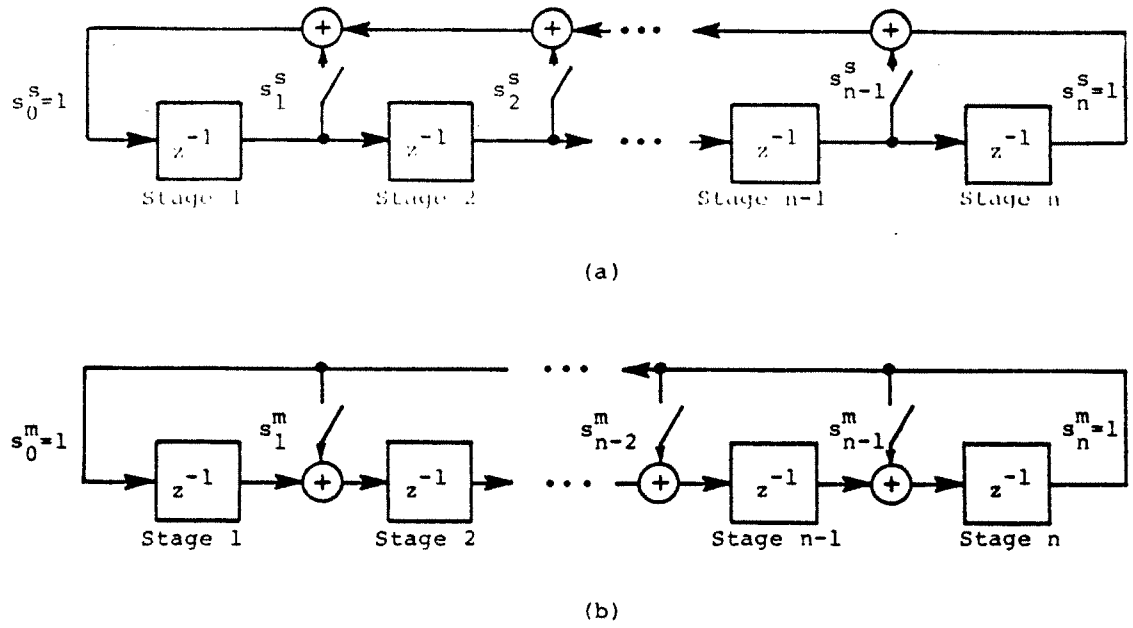


Fig. 4.1. Two canonical shift register generator configurations a) The simple shift register generator b) The modular shift register generator

The sequences generated by the SSRG and the MSRG are described by the same recurrence relation if

$$s_{n-i}^s = s_i^m \tag{4.36}$$

If $s_i^s=s_i^m$ then the sequences generated by these two generator forms are the time reverses of each other.

Using the delay operator the recurrence relationship used to define a linear maximal sequence can be written in the form

$$c_n z^n+c_{n-1}z^{n-1}+\ldots+c_1 z+1 \qquad (\text{mod } 2) \tag{4.37}$$

This equation is also referred to as the characteristic polynomial or law. Only a subset of the 2^{n-2} possible polynomials yield sequences of maximal length. Determining which polynomials characterize linear maximal sequences is a non-trivial process.

It is common practice to refer to a particular sequence using the octal representation of the characteristic polynomial coefficients written as a binary number. For example, the sequence whose coefficients going from c_7 down to c_0 (=1) is 10011101 would be named 235.

Note that the switch settings for the MSRG correspond directly to the coefficients of the characteristic polynomial and that the SSRG switch settings correspond to the coefficients written in reverse order.

Linear maximal sequences can be used to control the amplitude and/or phase modulation of a carrier. Each bit in the sequence is used to select the waveform state during a signaling or digit interval of T seconds. This can be done by clocking a shift register generator at a rate of 1/T Hz and applying the output of one stage to an appropriate modulator. It is not necessary to match the digit duration to the carrier frequency in any particular manner. However, it is very desirable to make the digit duration correspond to an integer number of carrier cycles because this significantly simplifies the equipment and procedures used to process the resulting waveforms.

An amplitude and phase modulated carrier can be written in the form

$$r_c(t) = a(t)\cos(2\pi f_c t+\theta(t)) \tag{4.38}$$

Substituting the complex exponential representation of the cosine gives

$$r_c(t) = \frac{a(t)}{2}\{\exp(j2\pi f_c t+j\theta(t))+\exp(-j2\pi f_c t-j\theta(t))\} \tag{4.39}$$

Using the Fourier shifting theorem, the spectrum of $r_c(t)$ consists of the spectrum of $a(t)\exp(j\theta(t))/2$ shifted from dc to $+f_c$ and the conjugate of this spectrum shifted to $-f_c$.

When the bandwidth of $a(t)\exp(j\theta(t))$ is sufficiently narrow with respect to f_c so that there is negligible overlap between the contributions of the components centered at f_c and $-f_c$, the relationships between components in the baseband representation translate into equivalent relations in the spectrum of the modulated waveform. This narrow-band assumption is made here.

Let $p(t)$ be a possibly complex valued periodic baseband pulse like waveform with period LT seconds. A baseband sequence controlled modulating waveform can be constructed such that

$$r(t) = \sum_{i=0}^{L-1} a_i \exp(j\theta_i) p(t-iT) \qquad (4.40)$$

where the values of a_i and θ_i are selected in some manner by an L digit time spreading sequence. The width of $p(t)$ is not necessarily constrained to equal T.

Since $p(t)$ is periodic, the appropriate spectral representation of $r(t)$ is obtained using Fourier series.

$$R_k = \frac{1}{LT} \int_{-\frac{LT}{2}}^{\frac{LT}{2}} \sum_{i=0}^{L-1} a_i \exp(j\theta_i) p(t-iT) \exp(-j2\pi \frac{kt}{LT}) dt \qquad (4.41)$$

R_k corresponds to the complex amplitude of the spectral line at frequency $\frac{k}{T}$.

Interchanging the operations of integration and summation, substituting $t'=t-iT$, dropping the primes and rearranging terms gives the result

$$R_k = \{\sum_{i=0}^{L-1} a_i \exp(j\theta_i)\exp(-j2\pi\frac{ki}{L})\}\{\frac{1}{LT}\int_{-\frac{LT}{2}}^{\frac{LT}{2}} p(t)\exp(-j2\pi\frac{nt}{T})dt\} \tag{4.42}$$

As expected from the results in the preceding section, the spectrum of r(t) is the product of L times the DFT of the sequence selected parameters and the Fourier series of the basic modulating pulse waveform. This nice result allows the effects of sequence and the pulse shape to be studied separately.

It is convenient to use the representation

$$w_i = a_i\exp(j\theta_i) = as_i+b \tag{4.43}$$

where a and b can be complex and $\{s_i\}$ is the linear maximal sequence based +1, -1 representation introduced earlier in this section.

Because the transducer is considered to be peak power limited, the values of a and b are restricted so that the magnitudes of the resulting w_i are less than or equal to one.

The normalized autocorrelation function of $\{w_i\}$ is

$$\hat{C}(k) = aa^*(\frac{L+1}{L}) + bb^* - \frac{aa^*+ab^*+a^*b}{L} \qquad k=0 \tag{4.44}$$

$$= bb^* - \frac{aa^*+ab^*+a^*b}{L} \qquad k\neq 0$$

The DFT of $\hat{C}(k)$ yields the power spectrum of $\{w_i\}$ which is the magnitude squared of the DFT of $\{w_i\}$.

$$|W(k)|^2 = \frac{(L-1)\hat{C}(\neq 0)+\hat{C}(0)}{L} \qquad k=0 \qquad (4.45)$$

$$= \frac{\hat{C}(0)-\hat{C}(\neq 0)}{L} \qquad k\neq 0$$

The spectrum of the $\{w_i\}$ is flat ("white") with the possible exception of the dc (carrier) line. Thus, the power spectrum of r(t) corresponds to a scaled version of the power spectrum of p(t) with a possibly altered dc value.

Using the results from the previous section, it can be shown that if a linear maximal sequence is used to cause time spreading, the resulting improvement in signal-to-noise ratio at the receiver with respect to the simple pulse waveform p(t) is

$$SNRIF = \frac{L^3}{\sum_{k=0}^{L-1} \frac{1}{|W(k)|^2}} \qquad (4.46)$$

$$= \frac{L\{(L-1)\hat{C}(\neq 0)+\hat{C}(0)\}\{\hat{C}(0)-\hat{C}(\neq 0)\}}{(L-2)\hat{C}(\neq 0)+\hat{C}(0)} \qquad (4.47)$$

This equation must be used with care since it includes any gain factors which might have been built into the selection of the values of a and b. When comparing various modulation schemes, it is customary to set the maximum magnitudes of the associated as_i+b representations equal to unity (peak power limited) or to equate the energy per sequence period.

The efficiency of a given modulation is defined as the ratio of the signal-to-noise ratio obtained with this modulation divided by the maximum possible signal-to-noise ratio for a modulation having the same average energy.

$$\text{efficiency} = \frac{L^2}{\left(\sum_{k=0}^{L-1} |W(k)|^2\right)\left(\sum_{k=0}^{L-1} \frac{1}{|W(k)|^2}\right)} \tag{4.48}$$

Since

$$\sum_{k=0}^{L-1} |W(k)|^2 = \hat{C}(0) \tag{4.49}$$

$$\text{efficiency} = \frac{\text{SNRIF}}{L\hat{C}(0)} \tag{4.50}$$

Two commonly used modulation schemes are on/off (AM) modulation and complementary phase modulation (CPM) (constant amplitude with the phase shifted = $\pm\ \theta$). The signal-to-noise ratio and efficiency for these two methods of modulation will be determined assuming a maximum value of unity for the w_i.

For AM, a=-1/2, b=1/2. This gives

$$\hat{C}(0) = \frac{L+1}{2L} \tag{4.51}$$

$$\hat{C}(\neq 0) = \frac{L+1}{4L} = \hat{C}(0)/2 \tag{4.52}$$

Substituting

$$\text{SNRIF}_{AM} = \frac{(L+1)^2}{4L} \tag{4.53}$$

$$\text{AM efficiency} = \frac{1}{2} + \frac{1}{L} \tag{4.54}$$

For CPM, a=-jsinθ, b=cosθ which gives

$$\hat{C}(0) = 1 \tag{4.55}$$

$$\hat{C}(\neq 0) = \frac{L\cos^2\theta - \sin^2\theta}{L} \tag{4.56}$$

Substituting

$$SNRIF_{CPM} = \frac{(L+1)\sin^2\theta(L^2\cos^2\theta+\sin^2\theta)}{L(L-1)\cos^2\theta+2\sin^2\theta} \tag{4.57}$$

$$\text{CPM efficiency} = \frac{SNRIF_{CPM}}{L} \tag{4.58}$$

If θ is chosen equal to $\tan^{-1}(L^{1/2})$, then the signal-to-noise ratio improvement factor is equal to L and the efficiency is at its maximum value of 1.

Two commonly used CPM modulation schemes use values of θ equal to 45° (named complementary modulation, CM) and 90° (named bi-phase modulation, BM).

For $\theta=45^\circ$, the signal-to-noise ratio improvement factor and efficiency are

$$SNRIF = \frac{(L+1)(L^2+1)}{2(L^2-L+2)} \simeq \frac{L+2}{2} \tag{4.59}$$

$$\text{efficiency} \simeq \frac{1}{2} + \frac{1}{L} \tag{4.60}$$

If the peak power is limited, CM provides a factor of 2 improvement over AM. If the energy per transmission is fixed, then AM is just as efficient as CM.

For $\theta=90^\circ$, the signal-to-noise ratio improvement factor and efficiency are

$$SNRIF = \frac{L+1}{2} \tag{4.61}$$

$$\text{efficiency} = \frac{1}{2} + \frac{1}{2L} \tag{4.62}$$

The time spreading effect of a linear maximal sequence can be removed using operations in either the frequency or the time domain. Because the sequences are periodic, it would appear that frequency domain operations using the DFT are preferred. Unfortunately, there are 2^n-1 digits per sequence period rather than the 2^n which would allow ready use of the fast Fourier transform algorithm. Most of the values of L are factorable. If time spread removal is to be performed in the frequency domain, special transform routines are needed for each value of L. The factors of $L=2^n-1$ for n=3 thru 12 are tabulated below.

Table 4.1. Factors of L for n=3,12

n	L	factors
3	7	none
4	15	3,5
5	31	none
6	63	3,3,7
7	127	none
8	255	3,5,17
9	511	7,73
10	1023	3,11,31
11	2047	23,89
12	4095	3,3,5,7,13

Depending on the circumstances, three viable approaches to collapsing a reception that has been time spread using linear maximal sequences have been found to be:

1. A high speed array processor such as the Floating Point Systems AP-120B can be used to brute force form the DFT via its definition. A 511 point DFT on the AP-120B using the definition takes only 0.35 seconds (or about 0.054 seconds if a special purpose FFT based on the factors 7 and 73 is used). This corresponds to doing a DFT "butterfly" in 1.33 microseconds. This is useful for in lab post processing schemes.

2. Because the time spreading is based on a weighted set of +1 and -1 values, it is possible to develop algorithms for replica cross-correlation which consist mainly of adds with relatively few multiplications. These routines can be made fairly effecient for values of L even as large as 511. The resulting program code can be made very compact and is useful for on-line processing during data acquistion. This approach is described in the next section.

3. Linear maximal sequences can be discarded and sequences of period 2^n can be used. This allows the use of fast Fourier transform algorithms to speed up the frequency domain processing. Unfortunately, it can be shown that there are no binary sequences of even length greater than two which possess a white spectrum [17]. The quest then is to find sequences whose spectra are close to being white. Because the spectra of these sequences is non-white, there will be a loss in SNRIF compared to that of a linear maximal sequence of comparable length. The trade-off here, is between a performance loss and an increase in processing effeciency. If one is already using linear maximal sequences, the most natural class of sequences to consider are linear maximal sequences which have been augmented by adding an additional zero bit. This addition yields a balance of 2^{n-1} ones and 2^{n-1} zeros. The effect of this augmenting varies depending on the sequence involved and the phase position at which the extra bit is added. From a hardware designer's viewpoint, one of the easiest places to insert the extra bit is just after the string of n-1 zeros present in sequences of degree n.

The first approach is useful if the amount of data to be processed is not overwhelming and an array processor is available. The beauty in brute force is that it usually works well and gets the job done.

The second approach has been extensively used in data processing programs written for on-line real-time operation in the field. This approach also has the flavor of being brute force. The cleverness of this method is that the computation can be organized such that the majority of operations are adds and subtracts. Only a relatively few multiplications are required. This approach lends itself nicely to implementation in special purpose correlation hardware. Section 4.3 goes into this approach in some detail.

The third approach is particularly useful in situations where there is some specific need to work in the frequency domain rather than the time domain. Using a time spread waveform to calibrate the frequency response of a signal processing system is one possible application. A trade is made between performance and ease of processing. The cost is often less than 3 dB of processing gain. Searching out good sequences is a major problem. The approach proposed here is to take something that is already good and perturb it slightly, hoping not to disrupt "goodness" too greatly. The result is that a set of sequences have been found that are easy to generate and cost less that 3 dB in performance. Other classes of sequences remain to be investigated.

A modified form of the efficiency equation termed the non-flatness loss factor can be used to evaluate sequences. The efficiency equation relates the performance of a given sequence and modulation to the maximum possible performance possible using the same energy per period. Since the energy contained in the carrier line can be controlled independently of how the remaining signal energy distributes among the other spectral lines, it is reasonable to exclude the dc term from the efficiency equation when evaluating

sequences as opposed to when evaluating methods of modulation.

The non-flatness loss factor is defined as

$$NFLF = \frac{L^2}{\left(\sum_{k=1}^{L-1} |W(k)|^2\right)\left(\sum_{k=1}^{L-1} \frac{1}{|W(k)|^2}\right)} \tag{4.63}$$

Using the representation $w_i = as_i + b$, the DFT of the w sequence can be written in terms of the DFT of the s sequence as

$$\begin{aligned} W(k) &= b + aS(0) \quad k=0 \\ &= aS(k) \quad k\neq 0 \end{aligned} \tag{4.64}$$

Substituting into the above equation for the NFLF gives

$$NFLF = \frac{L^2}{\left(\sum_{k=1}^{L-1} |S(k)|^2\right)\left(\sum_{k=1}^{L-1} \frac{1}{|S(k)|^2}\right)} \tag{4.65}$$

The NFLF is independent of the method of modulation.

If the sequence is balanced (contains an equal number of ones and zeros), then $S(0)=0$ and by Parseval's theorem

$$\sum_{k=1}^{L-1} |S(k)|^2 = 1 \tag{4.66}$$

Thus for balanced binary sequences

$$NFLF_b = \frac{L^2}{\sum_{k=1}^{L-1} \frac{1}{|S(k)|^2}} \tag{4.67}$$

The loss due to adding an additional zero to a linear maximal sequence has been investigated for two cases: 1) Adding the additional zero to make the n-1 digit run of

zeros into an n digit run and 2) Determining the optimal location. The results of this study are tabulated below for n=5 thru 10

Table 4.2. Best Augmented LMS NFLF Values for n=5,10

n	law	n-zero NFLF	law	max NFLF
5	57	0.76	57	0.87
6	147	0.70	103	0.74
7	211	0.64	357	0.77
8	537	0.64	435	0.67
9	1257	0.59	1175	0.67
10	3337	0.54	3117	0.62

The effect on signal-to-noise ratio in dB caused by augmenting is found by taking $10\log_{10}$ of the NFLF number. This does not include the effect of power splitting between the carrier line and the sidebands.

4.3 Time Spread Removal by Cross-Correlation

This section continues the development of correlation processing for time spread removal that was introduced at the end of Section 4.1. The methods outlined here are specific to the use of linear maximal sequences introduced in Section 4.2.

The time spread waveform is a weighted sum of time shifts of a basic pulse waveshape. A total of L pulses with separation T seconds are combined to form the transmitted waveform, r(t).

$$r(t) = \sum_{i=0}^{L-1} w_i p(t-iT) \qquad (4.68)$$

The received waveform is r(t) plus additive noise n(t). This will be written

$$r_n(t) = r(t) + n(t) \qquad (4.69)$$

where any channel gain is included in p(t).

In section 4.1 it was shown that if r(t) was uniformly sampled taking ML samples per period, the time domain processing could be divided into M interleaved cross-correlations of L values of the form

$$h(q) = \sum_{i=0}^{L-1} v_i^* r_n(q+i) \tag{4.70}$$

Because of the number of summations involved it is convenient to use the inner product notation

$$\langle x(i), y(i) \rangle \stackrel{d}{=} \sum_{i=0}^{L-1} x(i)\, y^*(i) \tag{4.71}$$

Using this notation

$$h(q) = \langle r_n(q+i), v_i \rangle \tag{4.72}$$

and

$$r_n(q+i) = \langle p(q+i-j), w_j^* \rangle + n(q+i) \tag{4.73}$$

Substituting the expression for $r_n(q+i)$ into the expression for h(q) gives

$$h(q) = \langle \langle p(q+i-j), w_j^* \rangle, v_i \rangle + \langle n(q+i), v_i \rangle \tag{4.74}$$

Because periodic waveforms and sequences are involved, it can be shown that

$$\langle \langle x(i-j), y^*(j) \rangle, z(i) \rangle = \langle x(i), \langle y^*(j), z^*(i+j) \rangle \tag{4.75}$$

Applying this relation to Eq. 4.74

$$h(q) = \langle p(q+i), \langle w_j^*, v_{i+j}^* \rangle \rangle + \langle n(q+i), v_i \rangle \qquad (4.76)$$

The sequences $\{w_i\}$ and $\{v_i\}$ are restricted here to being linear transformations of the same linear maximal sequence $\{s_i\}$ as defined in Section 4.2. The cross-correlation of $\{w_i\}$ and $\{v_i\}$ then becomes a linear transformation of the autocorrelation of $\{s_i\}$. Since $\{s_i\}$ has a two level autocorrelation, the cross correlation of $\{w_i\}$ and $\{v_i\}$ is also two level. For convenience, the sequence phases used to generate $\{w_i\}$ and $\{v_i\}$ are assumed to be such that the cross-correlation $\langle w_j, v_{i+j} \rangle$ has its peak value for $j=0$ and its off-peak or pedestal level for all $j \neq 0$. Since the off-peak values are identical, the notation $\langle w_j, v_{j+1} \rangle$ will be used to represent this value.

Using the Kronecker delta, $d_k(i)$, $\langle w_j, v_{i+j} \rangle$ can be written as

$$\langle w_j, v_{i+j} \rangle = (\langle w_j, v_j \rangle - \langle w_j, v_{j+1} \rangle) d_k(i) + \langle w_j, v_{j+1} \rangle \qquad (4.77)$$

Substituting this into Eq. 4.76 gives

$$h(q) = p(q)(\langle w_j, v_j \rangle - \langle w_j, v_{j+1} \rangle) + \langle p(q+i), 1 \rangle \langle w_j, v_{j+1} \rangle + \langle n(q+i), v_i \rangle \qquad (4.78)$$

The contribution of the signal to h(q) consists of scaled values of the base pulse waveshape plus a bias term. The bias amounts to a weighted sum of the samples of the pulse waveshape formed over one period and is a constant independent of q. Removing this bias yields the desired time collapsed waveform.

The required bias correction can be made in a number of ways. Some of these are

A. Cross-correlate and correct the result based on the calculated sum of the input samples.

$$h_c(q) = \langle r_n(q+i), v_i \rangle + f_1 \langle r_n(i), 1 \rangle \qquad (4.79)$$

B. Cross-correlate and correct the result based on the calculated sum of the correlation values.

$$h_c(q) = \langle r_n(q+i), v_i \rangle + f_2 \langle \langle r_n(k+i), v_i \rangle, 1 \rangle \qquad (4.80)$$

C. Adjust the mean value of the samples prior to correlation so that no correction is required after correlation.

$$h_c(q) = \langle r_n(q+i) + f_3 \langle r_n(i), 1 \rangle, v_i \rangle \qquad (4.81)$$

D. Demean the samples prior to cross-correlation and add in the proper correction (based on the sum of the original values) after correlation.

$$h_c(q) = \langle r_n(q+i) - \frac{1}{L} \langle r_n(i), 1 \rangle, v_i \rangle + f_4 \langle r_n(i), 1 \rangle \qquad (4.82)$$

E. Choose the values of v_i so that no correction is required.

$$h_c(q) = h(q) = \langle r_n(q+i), v_i \rangle \qquad (4.83)$$

Expanding the expressions for h(q) for each method gives

Method A

$$\begin{aligned} h_c(q) = {} & \langle r(q+i), v_i \rangle + f_1 \langle r(i), 1 \rangle \\ & + \langle n(q+i), v_i \rangle + f_1 \langle n(i), 1 \rangle \end{aligned} \qquad (4.84)$$

Method B

$$h_c(q) = \langle r(q+i), v_i \rangle + f_2 \langle r(i), 1 \rangle \langle 1, v_i \rangle + \langle n(q+i), v_i \rangle + f_2 \langle n(i), 1 \rangle \langle 1, v_i \rangle \quad (4.85)$$

Method C

$$h_c(q) = \langle r(q+i), v_i \rangle + f_3 \langle r(i), 1 \rangle \langle 1, v_i \rangle + \langle n(q+i), v_i \rangle + f_3 \langle n(i), 1 \rangle \langle 1, v_i \rangle \quad (4.86)$$

Method D

$$h_c(q) = \langle r(q+i), v_i \rangle - \frac{1}{L} \langle r(i), 1 \rangle \langle 1, v_i \rangle + f_4 \langle r(i), 1 \rangle \quad (4.87)$$

$$+ \langle n(q+i), v_i \rangle - \frac{1}{L} \langle n(i), 1 \rangle \langle 1, v_i \rangle + f_4 \langle n(i), 1 \rangle$$

Method E

$$h_c(q) = \langle r(q+i), v_i \rangle + \langle n(q+i), v_i \rangle \quad (4.88)$$

Because all of these methods correct the same bias term, the correction factors f_1 thru f_4 are related as follows:

$$f_2 = \frac{f_1}{\langle 1, v_i \rangle} \quad (4.89)$$

$$f_3 = \frac{f_1}{\langle 1, v_i \rangle} \quad (4.90)$$

$$f_4 = f_1 + \frac{1}{L} \langle 1, v_i \rangle \quad (4.91)$$

Substituting these expressions into the correlation equations gives for methods A thru D

$$h_c(q) = \langle r(q+i), v_i \rangle + f_1 \langle r(i), 1 \rangle \tag{4.92}$$

$$+ \langle n(q+i), v_i \rangle + f_1 \langle n(i), 1 \rangle$$

It now remains to determine f_1.

In Eq. 4.78 the values of the signal components of the $h_c(q)$ correspond to scaled values of $p(q)$ with an offset added. This offset is zero if

$$\langle p(q+i), 1 \rangle \langle w_j, v_{j+1} \rangle + f_1 \langle r(i), 1 \rangle = 0 \tag{4.93}$$

The quantity $\langle r(i), 1 \rangle$ can be written

$$\langle r(i), 1 \rangle = \langle \langle p(q+i-j), w_j^* \rangle, 1 \rangle = \langle p(q+i), \langle w_j^*, 1 \rangle \rangle \tag{4.94}$$

$$= \langle p(q+i), 1 \rangle \langle w_j, 1 \rangle \tag{4.95}$$

Using this result and solving for f_1 gives

$$f_1 = - \frac{\langle w_j, v_{j+1} \rangle}{\langle w_j, 1 \rangle} \tag{4.96}$$

In Section 4.2 the representation

$$w_j = as_j + b \tag{4.97}$$

was used to represent the linear maximal sequence based modulation weights. The s_i take on the values +1 and -1 as the underlying sequence takes on the binary zero and one states.

Similarly, the reference weights $\{v_j\}$ will be represented as

$$v_j = cs_j + d \tag{4.98}$$

The values taken on by the autocorrelation function of the $\{s_j\}$ are

$$\langle s_j, s_j \rangle = L \tag{4.99}$$

$$\langle s_j, s_{j+1} \rangle = -1 \tag{4.100}$$

Using these relations and the representations for the w_j and the v_j gives

$$\langle w_j, v_j \rangle = c^*(aL-b) + d^*(bL-a) \tag{4.101}$$

$$\langle w_j, v_{j+1} \rangle = -c^*(a+b) + d^*(bL-a) \tag{4.102}$$

Substituting into the equation for f_1 gives

$$f_1 = \frac{c^*(a+b)-d^*(bL-a)}{(bL-a)} \tag{4.103}$$

Using this result yields the expressions for the other correction factors.

$$f_2 = f_3 = \frac{c^*(a+b)-d^*(bL-a)}{(bL-a)(d^*L-c^*)} \tag{4.104}$$

$$f_4 = \frac{ac^*(L+1)}{L(bL-a)} \tag{4.105}$$

Note that if $\langle w_j, v_{j+1} \rangle = 0$ then no correction is required in Eq. 4.78. This will be when c and d are chosen so that

$$\frac{d^*}{c^*} = \frac{a+b}{bL-a} \tag{4.106}$$

Once the bias correction has been made, the signal component of the correlator output is

$$h_c(q) = p(q)(\langle w_j, v_j\rangle - \langle w_j, v_{j+1}\rangle) = p(q)ac^*(L+1) \quad (4.107)$$

The noise term can be written as

$$\langle n(q+i), v_i\rangle + f_1\langle n(i), 1\rangle = \langle n(q+i), v_i\rangle + f_1\langle n(q+i), 1\rangle \quad (4.108)$$

$$= \langle n(q+i), v_i + f_1^*\rangle \quad (4.109)$$

If the noise samples are independent and have uniform variance V_n, it can be shown, given a deterministic set of x(i) values, that

$$E\{\langle n(i), x(i)\rangle\langle n(i), x(i)\rangle^*\} = V_n\langle x(i), x(i)\rangle \quad (4.110)$$

The variance of the correlator output values is

$$V_{nc} = V_n\langle v_i + f_1^*, v_i + f_1^*\rangle \quad (4.111)$$

$$= V_n\{\langle v_i, v_i\rangle + f_1\langle v_i, 1\rangle + f_1^*\langle 1, v_i\rangle + f_1 f_1^* L\} \quad (4.112)$$

Substituting the appropriate values and arranging terms gives

$$V_{nc} = V_n cc^*(L+1)\{\frac{L-1}{L} + \frac{(L+1)aa^*}{L(bL-a)(b^*L-a^*)}\} \quad (4.113)$$

The resulting signal-to-noise ratio is

$$\frac{S}{N}(q) = \frac{(L+1)aa^*p(q)p^*(q)}{V_n\{\frac{L-1}{L} + \frac{(L+1)aa^*}{L(bL-a)(b^*L-a^*)}\}} \quad (4.114)$$

This depends only on the representation chosen for the weights used in generating the transmitted waveform and does

not depend on the form of the weights chosen for use in the correlation.

If the correlator weights are restricted as indicated in Eq. 4.106, the resulting equations for the signal magnitude and noise variance at the correlator output for Case E are identical to Eq. 4.107 and Eq. 4.113.

Since $\frac{p(q)p^*(q)}{v_n}$ corresponds to the signal-to-noise ratio associated with a sample of the base periodic pulse waveshape the remainder of the above signal-to-noise ratio equation must (and does) correspond to the signal-to-noise ratio improvement factor.

The above development shows that there are few constraints on the choice of the correlator weight constants c and d. If they are not matched to the a and b value, the worst that can happen is that there is a bias to be corrected for. Of the five processing plans presented above, four include such a correction. Two unexamined questions at this point are: 1) how does one select values of c and d and 2) which processing procedure does one use? These two questions will now be addressed, in reverse order.

The basic processing steps in the five methods are:

A. Sum data values
   Correlate
   Adjust the bias based on the sum

B. Correlate
   Sum the correlation values
   Adjust the bias based on the sum

C. Sum the data values
   Adjust the data values based on the sum
   Correlate

D. Sum the data values
   Remove the data mean using the sum
   Correlate
   Adjust the bias based on the sum

E. Correlate (c and d must be properly chosen)

Method A can be used to implement a sliding correlator which produces a new correlation value each time a new input sample becomes available. It can also be used to process a single reception period or the coherent average of several periods. If integer arithmetic is used, the processing can be organized so that no round off errors occur until the final bias correction is made.

Method B is useful when processing a single period or the coherent average of several periods. Basing the bias correction on the sum of the resulting correlation values rules out its use in sliding fashion. The computation of the sum can be incorporated in the correlation process. The processing can be organized so that no round off errors occur until the last stage.

Method C is useful when processing a single period or the coherent average of several periods. The modification of the input data values rules out its use in sliding fashion. Since the modification intended to correct the bias is made prior to correlation, there is a possibilty of round off problems.

Method D is also a non-sliding processing. The modification of the input data values prior to correlation rules out use in a sliding fashion. This method makes one more pass thru the data than do the above three methods and thus requires more program code. There is a chance for a small bias to be added due to round off errors in the demeaning process.

Method E can be used in sliding and non-sliding applications. The values of c and d are chosen based on the input modulation. The correlation is then implemented using the definition. If brute force hardware such as an array processor is available, it is possible to use this method with virtually no cost in time by overlapping the computation with I/O operations.

Methods E and A are closely related. Depending on how the values of c and d are chosen, it is easy to organize

method E code and end up with the same proceedure that would have resulted had method A been used from the start.

The author has experience using methods A, D and E. Methods A and E implement in relatively compact code and can be used in both sliding and non-sliding situations. Method A was used in the Spring of 1983 to sliding process 86 hours of reception on a DM1200. The processing time was about 3 times 86 hours. This method could readily be cast into hardware which would provide a significant increase in speed. Method E has successfully been used on a Floating Point Systems array processor and required significantly less time than the overhead required to move data thru the processing. These two methods are the author's current favorites.

Moving on to the question of how to select the values of c and d.

A particularly nice set of values is c=-1/2 and d=+1/2. Using this set results in the v_i taking on the values zero and one (directly corresponding to zero's and one's in the binary sequence). These values do not require any special hardware for multiplication. The correlation sum becomes a simple sum of (L+1)/2 complex values corresponding to the values selected by the ones in the sequence. The method A correction factor is

$$f_1 = - \frac{b(L+1)}{2(bL-a)} \qquad (4.115)$$

which for large values of L is approximately 1/2.

If an L element buffer is being processed in cyclic fashion, the sum value, S, can be computed once and properly weighted. The result is then added to the correlation sum formed for each q value. For $0 \leq q < L$ the expression for $h_c(q)$ is

$$h_c(q) = - \frac{S\ b(L+1)}{2(bL-a)} + \sum_{i=0}^{L-1} v_i^* r_n(i+q) \qquad (4.116)$$

where the indices of $r_n(\ )$ are computed modulo L.

The total number of complex adds required is $\frac{L}{2}(L+5)$. Only one complex multiply is required.

If the correlation is slid along a continuous input stream, the sum can be updated incrementally. When the correlation is advanced one position, the oldest sample is subtracted out of the sum and the new one added to it. The correlation is then formed using the above equation with the indices not being computed modulo L. For each correlation value produced, a total of $\frac{(L+1)}{2}+3$ complex additions and one complex multiplication are required.

Another attractive selection for c and d is c = -1 and d = 0. The v_i values are then +1 and -1 corresponding to one's and zero's in the binary sequence. The expression for f_1 in this case is

$$f_1 = - \frac{(a+b)}{(bL-a)} \qquad (4.117)$$

which is on the order of 1/L.

This set of c and d values requires about twice the number of additions that was required for c=-1/2, d=+1/2. This set is useful in performing sliding type correlations if one is insecure about maintaining the sliding sum used above. The correlation sum is divided into two sums: one the sum of those values selected by sequence one's, S_1, and the other, the sum of those values selected by sequence zero's, S_0. The correlation value is then

$$h_c(q) = (S_1-S_0) - \frac{(S_1+S_0)(a+b)}{(bL-a)} \qquad (4.118)$$

4.4 Odds and Ends

This section contains odds and ends of information about correlation processing of time spread waveforms based on the use of linear maximal sequences.

It is tempting to save time and reduce program size by not making the bias correction. Section 4.4.1 considers two sets of processing weights that might be used if such a short cut is desired.

One of the many possible problems that can occur when using sequence based waveforms is having the transmitter and receiver use modulation angles which are conjugates of each other. Section 4.4.2 examines the effects of getting the angle direction wrong.

Section 4.4.3 tabulates the values of f_1 thru f_4 for CPM modulation and the two most likely choices of processing parameters c and d.

It is often useful to know the number of times ones and zeros in the local sequence line up with ones and zeros in the received waveform. Section 4.4.4 lists these numbers as a function of L.

4.4.1 Ignoring the Bias Correction

Each arrival contributes to the bias level in proportion to its amplitude. Arrival amplitudes must exceed the bias level before they can be guaranteed to cause peaks instead of valleys. If the bias is not removed, there is a limit to how far down from the largest arrival peak that small arrivals can be reliably detected.

The signal component of Eq. 4.78 is

$$h(q) = p(q)(\langle w_k, v_k\rangle - \langle w_k, v_{k+1}\rangle) + \langle p(q+i), 1\rangle \langle w_k, v_{k+1}\rangle \quad (4.119)$$

Assume that the pulse waveshape is zero outside of the

interval $-\frac{T}{2}<t<+\frac{T}{2}$. Then $p(q)=p(0)$ if $q=0$ and $p(q)=0$ for all other q. With this constraint, $h(q)$ becomes

$$h(q) = p(q)(\langle w_k, v_k\rangle - \langle w_k, v_{k+1}\rangle) + p(0)\langle w_k, v_{k+1}\rangle \quad (4.120)$$

If the bias correction is not made, then

$$h(0) = p(0)\langle w_k, v_k\rangle \quad (4.121)$$

$$h(\neq 0) = p(0)\langle w_k, v_{k+1}\rangle \quad (4.122)$$

Arrivals must be larger in magnitude than the magnitude of the largest arrival multiplied by

$$\frac{|\langle w_k, v_{k+1}\rangle|}{|\langle w_k, v_k\rangle|} \quad (4.123)$$

before they can be detected with any reliability.

Consider a CPM transmission generated using $a=-j\sin\theta$ and $b=\cos\theta$.

Case I

The reception is processed using weight values 0 and 1. This corresponds to $c=-1/2$ and $d=1/2$. Using these values gives

$$h(0) = \tfrac{1}{2}p(0)(L+1)(\cos\theta+j\sin\theta) \quad (4.124)$$

$$h(\neq 0) = \tfrac{1}{2}p(0)(L+1)\cos\theta \quad (4.125)$$

This results in the relation

$$|\text{bias}| = |\cos\theta||\text{peak}| \quad (4.126)$$

Using the Tomography parameter values of $\theta=75^{o}$ and L=127, the bias level is 11.7 dB down from the peak.

Case II

The correlation is performed using weights +1 and -1. This corresponds to values c=-1 and d=0. Using these values gives

$$h(0) = p(0)(\cos\theta+jL\sin\theta) \quad (4.127)$$
$$h(\neq 0) = p(0)(\cos\theta-j\sin\theta) \quad (4.128)$$

This gives the result

$$|\text{bias}| = \frac{|\text{peak}|}{(1+(L^2-1)\sin^2\theta)^{1/2}} \quad (4.129)$$

For the Tomography parameters $\theta=75^{o}$ and L=127, the bias level is down 41.8 dB.

4.4.2 Conjugate Mismatch on CPM Angle

When using conjugate phase modulation there is often a nagging worry about whether the transmitter is advancing phase or retarding phase when the controlling sequence is a one. The net effect of the transmitter and receiver going opposite ways is the presence of a bias in the processed results. As in Section 4.4.1 this bias limits the ability to see small arrivals.

The same assumptions about pulse width that were made in Section 4.4.1 are also made here. The equation describing the uncorrected correlator output is

$$h(q) = p(q)(\langle w_k, v_k\rangle - \langle w_k, v_{k+1}\rangle) + p(0)\langle w_k, v_{k+1}\rangle \quad (4.130)$$

The use of the conjugate angle to correct for the bias term amounts to adding the correction term $-p(0)\langle w_k^*, v_{k+1}\rangle$. Doing this gives

$$h_c(0) = p(0)(\langle w_k, v_k\rangle - \langle w_k^*, v_{k+1}\rangle) \quad (4.131)$$

$$h_c(\neq 0) = p(0)(\langle w_k, v_{k+1}\rangle - \langle w_k^*, v_{k+1}\rangle) \quad (4.132)$$

$$= p(0)\langle (w_k - w_k^*), v_{k+1}\rangle \quad (4.133)$$

For CPM $w_k = \cos\theta - js_k\sin\theta$. Substituting yields

$$h_c(0) = p(0)(-j\sin\theta)(\langle s_k, v_k\rangle - \langle s_k, v_{k+1}\rangle) \quad (4.134)$$

$$h_c(\neq 0) = p(0)(2j\sin\theta)\langle s_k, v_{k+1}\rangle \quad (4.135)$$

Using the representation $v_k = cs_k + d$ gives

$$\langle s_k, v_k\rangle = c^*L - d^* \quad (4.136)$$

$$\langle s_k, v_{k+1}\rangle = -(c^* + d^*) \quad (4.137)$$

The ratio of the magnitude of the bias to the magnitude of the peak value is then

$$\frac{|h_c(\neq 0)|}{|h_c(0)|} = \frac{2|c+d|}{|c|(L+1)} \quad (4.138)$$

If $c=-1/2$ and $d=1/2$ (the v_k take on the values 0 and 1) then for $L=127$ the bias is down 30.1 dB from the peak.

Using $c=-1$ and $d=0$ (the v_k take on the vaues +1 and -1) then for $L=127$ the bias is down 36.1 dB from the peak.

4.4.3 Correction Factor Expressions

Complementary phase modulation is assumed. The processing uses values of 1 (one in the sequece) and 0 (zero in the sequence). The parameter values are:

$$a=-j\sin\theta \quad b=\cos\theta \quad c=-\frac{1}{2} \quad d=\frac{1}{2}$$

The resulting equations for the correction factors f_1 thru f_4 defined in Section 4.3 are:

$$f_1 = -\frac{(L+1)(L-j\tan\theta)}{2(L^2+\tan^2\theta)} \tag{4.139}$$

$$f_2 = f_3 = -\frac{(L-j\tan\theta)}{(L^2+\tan^2\theta)} \tag{4.140}$$

$$f_4 = \frac{(L+1)\tan\theta(\tan\theta+jL)}{2L(L^2+\tan^2\theta)} \tag{4.141}$$

Complementary phase modulation is assumed. The processing uses values of +1 (one in the sequence) and -1 (zero in the sequence). The parameter values are:

$$a=-j\sin\theta \quad b=\cos\theta \quad c=-1 \quad d=0$$

The resulting equations for the correction factors f_1 thru f_4 defined in Section 4.3 are:

$$f_1 = f_2 = f_3 = -\frac{(L-\tan^2\theta)-j(L+1)\tan\theta}{(L^2+\tan^2\theta)} \tag{4.142}$$

$$f_4 = \frac{(L+1)\tan\theta(\tan\theta-jL)}{L(L^2+\tan^2\theta)} \tag{4.143}$$

4.4.4 Number of Times Bit States Match

When working with linear maximal sequences, it is occasionally useful to know the number of times that ones and zeros in the $\{w_i\}$ weights match up with ones and zeros in the $\{v_i\}$ weights. Of course, this is the most useful when the sequences are not perfectly aligned.

The small table given below lists the number of times bits in a sequence match up with bits in a shifted version of that same sequence. This is independent of sequence and holds for all phase shifts except the matching position [17].

Table 4.3. Sequence Bit Match Counts

one	one	$\frac{L+1}{4}$ times
one	zero	$\frac{L+1}{4}$ times
zero	one	$\frac{L+1}{4}$ times
zero	zero	$\frac{L-3}{4}$ times

CHAPTER 5

BEAM FORMING

Consider an array of N hydrophones across which a temporally but not spatially modulated acoustic wavefront is propagating from a fixed direction. Except for the fixed time delays caused by the wavefront propagating across the array, the components in the individual hydrophone outputs due to this wavefront are identical. By properly delaying the hydrophone outputs and summing the results, an electrical waveform is obtained in which the signal component is N times greater than in the output of a single hydrophone. Interfering wavefronts arriving from other directions are not matched to these delays and are not so enhanced. The net result is that the signal-to-noise ratio in the sum is larger than it is in the output of a single hydrophone. This process of enhancing the signal-to-noise ratio is termed beam forming.

This chapter describes and analyzes three methods of implementing digital beamformers for use at frequencies below 500 Hz. These are: 1) the delay and sum method, 2) the demodulate, delay and sum method, and 3) the phase shift method.

The delay and sum beamformer is a direct implementation of the process described above. Because true delay is used and because the speed of sound in the ocean is essentially independent of frequency, the beamformer gain in the direction steered by the chosen delays is independent of frequency. However, the gain off of the steered direction is frequency dependent and arrivals from directions other

than the steered direction are filtered in some manner.

Early delay and sum beamformers used electronic delay lines to obtain the desired delays [37]. The first digital delay and sum beamformer appears to have been the DIMUS (Digital Multibeam Steering) system of Anderson [38]. This early system clipped the individual hydrophone outputs to one bit and used shift registers to implement the required delays.

In a digital delay and sum beamformer, the individual hydrophone outputs are usually sampled using a fixed clock rate. Because of this, sample values may not be available at the exact delays required for beam forming. One approach to approximating samples with the required delays is to use the sample values having the delays closest to the required delays. Another approach is to use some method of interpolation [39]. Errors in the time delay values used to form beams cause grating lobes in the beamformer's spatial sensitivity pattern [25]. The frequency response of the beam forming process is also affected by errors in the delay times.

In designing a delay and sum beamformer, one seeks to use as low a sample rate as possible in order to reduce the delay memory requirements. One also seeks to use as high a sample rate as possible in order to reduce the effects of aliasing and of errors in the delay times. The basic design problem is to balance the trade-offs between these conflicting desires.

If the signal of interest is a band-pass waveform, the individual hydrophone outputs can be demodulated, delayed and then summed. The advantage to this process is that, because of the reduced bandwidth of the demodulated waveform, the accuracy requirements on the delay times are reduced. A beamformer which operates using time delayed demodulated sample values is termed a demodulate, delay and sum beamformer. A description of how such a beamformer can be implemental using interpolation methods is contained in

[40].

If the bandwidth of the demodulated waveform is sufficiently narrow, it is possible to adequately approximate the required delays using simple phase shifts. This approach to beam forming is termed phase shift beam forming and is commonly used in radio frequency applications. A beamformer of this type was built for use in the audio frequency range in the early 1970's and was later described in the literature [41].

In addition to the above three methods, beam forming can also be accomplished in the frequency domain [42]. However, the required hardware is sufficiently complex to have eliminated this approach from consideration by CEL.

The CEL beamformer project was not entered into with an open mind. The intent was to design and build a delay and sum beamformer if this could be done at a reasonable cost. The two driving requirements were: 1) the input filtering had to be lowpass and not depend on the demodulation frequency and 2) the unit had to support simultaneous operation at multiple demodulation frequencies. The time delay beamformer was felt to be the best approach by which these goals could be attained.

The method of delay and sum beam forming is studied in Section 5.1. The coordinate system used in the remainder of this chapter is described first. Next, the concept of the beam pattern is introduced. The beam pattern describes the sensitivity of the beamformer as a function of arrival angle, steering angle and frequency. The beam pattern is useful in determining the effects of pointing errors on the desired reception and the sensitivity to undesired waveforms arriving from directions other than the one in which the beamformer is steered. The beam pattern can be combined with models of the noise field in order to study the effects of various spatial noise distributions on the beamformer performance.

Because it is the most commonly encountered array

configuration, attention is restricted to line arrays. It is shown that, for a fixed line array, the beam sensitivity pattern is a function of a single variable. That variable itself is a function of frequency, the steered direction and the actual arrival direction. Using this property, the behavior of the main lobe in the beam pattern for a given response level can be characterized using only two constants. The values of the two constants depend on the specific hydrophone spacings but are fixed for a given array. These two constants are used to study the effects of pointing errors on the spatial and frequency responses of the beam forming process. Included is an example of the use of these constants in setting up of a fan of beams with a predetermined crossover level bracketing an expected arrival direction.

Two models of the spatial distribution of the noise field are used to evaluate the performance expected of the beam forming process. These are the spherically isotropic noise model and the circularly uniform horizontally arriving noise model. Closed form expressions describing the expected signal-to-noise ratio improvement due to beam forming are derived for these two noise models. The equations developed for the case of spherically isotropic noise are a slight generalization of those included in Urich's book [43] in that the effects of the steering direction are included. The case of circularly uniform horizontally arriving noise does not appear to have been published though it would seem to be an obvious noise model to consider.

The time quantization study contained in Section 5.1.1 was essential to the design of the CEL beamformer. This study was used to determine the effects of time delay quantization errors on the frequency response of the beamformer. It is shown that in the steered direction, the worst case effects of the quantization errors can be modeled as a low-pass filtering of the beamformed reception. The

time step used in the CEL beamformer was determined by requiring a minimum 500 Hz wide 1 dB down low-pass bandwidth due to quantization errors.

The effects of errors in the delay times on the sidelobe structure are relatively small and are of concern only in finely tuned systems. The CEL beamformer is intended for use with collections of hydrophones which were installed for applications other than CEL's and which are made available as a courtesy, on a non-interferring basis. Generally, these hydrophones form line arrays but often have a spacing intended for an operating frequency different from that required by CEL. Often it is necessary to use a given array simultaneously at widely differing frequencies. In these situations the presence of major sidelobes due to spatial undersampling is much more of a concern than the minor perturbations caused by delay time quantization errors. For these reasons the effects of time delay errors on the sidelobe structure are not considered here.

The demodulate, delay and sum beamformer is examined in Section 5.2. As expected, the reduced waveform bandwidth allows use of a coarser time quantization. It is shown that the demodulator outputs have to individually be phase corrected in order to account for the use of the same demodulating waveform in the demodulators. In other respects, this beamformer behaves in the same manner as the delay and sum beamformer.

Section 5.3 contains an anlysis of the phase shift beamformer. This configuration is studied fairly thoroughly because of its potential for use in the future. Not requiring any delays, it is a relatively simple device to build. Its most annoying feature is the deviation of the beam pattern maximum from the steered angle as the signal frequency moves away from the design frequency. This is a handicap not shared by the other two beam forming methods and is a problem when using wide band transmissions.

The analysis of the effects of time quantization errors

on the frequency response of a delay and sum beamformer presented here appears to be original to this work. Also, the realization and exploitation of the single variable nature of the line array beam pattern does not seem to have been recognized in the literature. A third contribution is the comparison of the performances expected using the spherically and horizontally distributed noise distributions.

5.1 Time Delay and Sum Beam Forming

The incoming wavefront at the receiving array is modeled as being a plane wave. One of the receiving hydrophones is selected as a reference and the positions of the remaining phones are measured relative to it. The coordinate system shown in Fig. 5.1 is used.

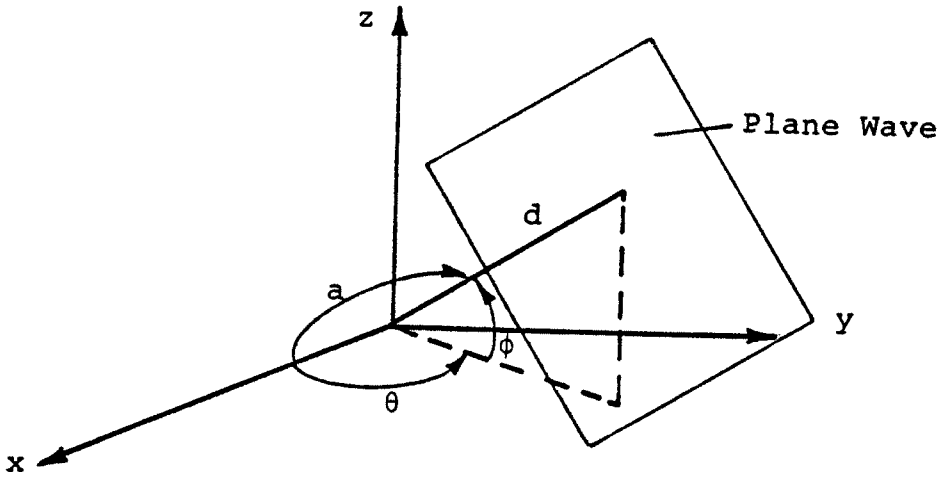


Fig. 5.1. Coordinate system used in beam forming

The angles shown in Fig. 5.1 are defined as

a The conic angle, the angle between the x-axis and the normal to the wavefront (i.e., the incoming ray).

Θ The bearing angle measured in the horizontal between the x-axis and the incoming wavefront. This is called "bearing from endfire."

ϕ The elevation angle of the arriving ray measured from the horizontal upwards. This is also called the vertical arrival angle.

The three angles are related by the expression

$$\cos(a) = \cos(\Theta)\cos(\phi) \tag{5.1}$$

Let d be some distance greater than the distance to the hydrophone furthest from the reference hydrophone. Let (x_i,y_i,z_i) be the coordinates of the i-th hydrophone relative to the reference phone. Given a plane wavefront at distance d from the reference phone in the direction specified by Θ and ϕ, then the distance from this plane wavefront to the point (x_i,y_i,z_i) is given by

$$d_i = d-(x_i\cos(\Theta)\cos(\phi)+y_i\sin(\Theta)\cos(\phi)+z_i\sin(\phi)) \tag{5.2}$$

The propagation delay from the wavefront to the i-th hydrophone is d_i/c where c is the speed of sound in the vicinity of the array.

Given time variations in the strength of the pressure field associated with a noise free arrival from the direction specified by Θ and ϕ, the electrical outputs of all of the hydrophones will be identical except for fixed delays. Let v(t) represent the electrical voltage generated by the reference hydrophone. The voltage generated by the i-th hydrophone is

$$v_i(t) = v(t+(d-d_i)/c) \tag{5.3}$$

A delay and sum beamformer delays the output of the i-th hydrophone by t_i seconds, weights it by w_i and then

sums the results from N phones to obtain the beam output

$$v_b(t) = \sum_{i=0}^{N-1} w_i v(t+(d-d_i)/c-t_i) \qquad (5.4)$$

When the delays are chosen so that

$$t_i = (d-d_i)/c \qquad (5.5)$$

then

$$v_b(t) = v(t)\sum_{i=0}^{N-1} w_i \qquad (5.6)$$

If the w_i are all set to one, then when the beamformer is steered in the direction of an arriving plane wave

$$v_b(t) = Nv(t) \qquad (5.7)$$

If it is assumed that the noise contributions of the individual hydrophone outputs are uncorrelated, the variance of the noise in the sum is N times the variance of a single phone. Since the desired signal has a gain of N, its power has been increased by a factor of N^2 and the signal-to-noise ratio at the beamformer output is N times the input signal-to-noise ratio associated with a single phone. It is this potential processing gain that motivates the use of beam forming.

The actual processing gain achieved using a given array of hydrophones depends on the nature and distribution of noise sources and on the stability of the signal wavefronts at the individual hydrophones.

The Fourier transform of the beamformer output is

$$V_b(f) = V(f)\sum_{i=0}^{N-1} w_i \exp\{-j2\pi f(t_i-(d-d_i)/c)\} \qquad (5.8)$$

The summation term in the above equation describes the sensitivity of the beam forming process at frequency f for a given set of delays $\{t_i\}$ to arrivals from the (θ,ϕ) direction. That is, it represents the spatial sensitivity of the hydrophone array and beamformer combination. Assuming that the delays are chosen to steer the array in the (θ_s,ϕ_s) direction the beam pattern of the combination is defined as [43]

$$b(f,\theta,\phi,\theta_s,\phi_s) = \left| \frac{\sum_{i=0}^{N-1} w_i \exp\{-j2\pi f(t_i-(d-d_i)/c)\}}{\sum_{i=0}^{N-1} w_i} \right|^2 \qquad (5.9)$$

Note that $b(f,\theta,\phi,\theta_s,\phi_s)$ has been defined so that it equals unity when the delays are chosen so that $\theta = \theta_s$ and $\phi = \phi_s$.

Given a spatial noise distribution $N(f,\theta,\phi)$ and steering direction (θ_s,ϕ_s), the gain of the array is defined as [43]

$$AG(f,\theta_s,\phi_s) = 10 \log \frac{\int_{4\pi} N(f,\theta,\phi)d\Omega}{\int_{4\pi} N(f,\theta,\phi)b(f,\theta,\phi,\theta_s,\phi_s)d\Omega} \qquad (5.10)$$

The array gain represents the improvement in noise rejection caused by the use of the array/beamformer combination as compared to a single hydrophone in the same noise field. If the received signal waveform is coherent

across the array and arrives precisely from the (θ_s, ϕ_s) direction, then this quantity also represents the improvement in signal-to-noise ratio obtained by beam forming as compared to a single hydrophone. Perturbations in the beam forming process have a more significant effect on the signal gain then they do on the noise rejection process. For example, steering a five degree wide beam ten degrees off of broadside ($a=90^{\circ}$) for a broadside reception has little effect on the noise level at the beamformer output but devastates the signal. Because of this (and in order to be able to drop the 10 log operation), define the noise improvement factor (NIF) as

$$NIF(f,\theta_s,\phi_s) = \frac{\int_{4\pi} N(f,\theta,\phi)d\Omega}{\int_{4\pi} N(f,\theta,\phi)b(f,\theta,\phi,\theta_s,\phi_s)d\Omega} \qquad (5.11)$$

The structure of the noise field in the ocean is very complex. Potential noise sources include near and distant shipping, acoustic bottom surveys, surface waves and biological activity. Discounting exceptional circumstances such as speed boats passing over the array or whales playing with the hydrophones, the two most significant noise sources are surface waves and distant shipping. Studies made at two sites [44],[45] indicate that at frequencies in the range of 200 to 400 Hz and lower, for relatively calm seas, the major noise source is distant shipping and essentially arrives horizontally. As the surface becomes disturbed by wind driven wave action, the noise intensity tends to become a more uniform function of the vertical arrival angle and at high sea states favors the near vertical angles.

In order to estimate the performance expected from beam forming and to compare beam forming methods, the assumption of a spherically isotropic noise field is often made. In

this case

$$N(f,\theta,\phi) = 1 \tag{5.12}$$

Because the CEL beamformer is intended for use at low frequencies (below 500 Hz), it makes sense to also model both the noise and the signal as arriving horizontally (i.e., $\phi = 0$) and to calculate the performance on this basis.

Assuming spherically isotropic noise and a linear array lying along the x-axis, the expression for the NIF can be written in terms of the conic angle a as

$$NIF_s(f,a_s) = \frac{1}{\frac{1}{2}\int_0^{\pi} b(f,a,a_s)\sin(a)da} \tag{5.13}$$

Similarly, using the same geometry and assuming all receptions are in the horizontal plane with the noise being horizontally isotropic gives

$$NIF_h(f,\theta_s) = \frac{1}{\frac{1}{\pi}\int_0^{\pi} b(f,\theta,\theta_s)d\theta} \tag{5.14}$$

In order to develop a feel for the effects of beam forming, consider a linear array of N hydrophones aligned along the x-axis. Assuming conic arrival angle a and conic steering angle a_s gives values

$$d_i = d - x_i\cos(a) \tag{5.15}$$

$$t_i = x_i\cos(a_s)/c \tag{5.16}$$

Substituting into the expression for the beam pattern gives

$$b(f,a,a_s) = \left| \frac{\sum_{i=0}^{N-1} w_i \exp\{-j2\pi f x_i (\cos(a_s)-\cos(a))/c\}}{\sum_{i=0}^{N-1} w_i} \right|^2 \tag{5.17}$$

The use of the weights $\{w_i\}$ is termed shading. For a given array geometry, the use of the $\{w_i\}$ allows the structure of the beam pattern to be modified. Arrays are often shaded in order to reduce the magnitudes of spurious side lobes in the beam pattern or to aid in steering a null in the direction of a strong interfering source. However, this capability is not without its cost. According to Urick [43], the use of shading with a coherent signal in a uniform incoherent noise field results in an array gain which is lower than that obtained without shading. Because of this, the effects of shading are not considered here and all of the w_i are set equal to unity. Doing this and forming the magnitude squared gives

$$b(f,a,a_s) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} \exp\{-j2\pi f (x_i-x_k)(\cos(a_s)-\cos(a))/c\} \tag{5.18}$$

Exploiting the symmetries in the differences x_i-x_k and noting that the imaginary part sums to zero allows rewriting $b(f,a,a_s)$ as

$$b(f,a,a_s) = \frac{1}{N} + \frac{2}{N^2} \sum_{i=1}^{N-1} \sum_{k=0}^{i-1} \cos(2\pi f (x_i-x_k)(\cos(a_s)-\cos(a))/c) \tag{5.19}$$

Although the above form allows the derivation of closed form expressions for the noise improvement factor, it requires on the order of $N^2/2$ cosine evaluations and generally should not be used to compute the beam pattern.

Given a fixed line array and a constant speed of sound across the array, the dependence of the beam pattern on f, a, and a_s is a particularly simple one since

$$b(f,a,a_s) = b(f(\cos(a_s)-\cos(a)) \qquad (5.20)$$

In essence, the beam pattern is a function of a single variable. That variable itself is a function of three parameters. This feature can be exploited in relating changes in the beam pattern to changes in parameter values.

In the expression for $b(f,a,a_s)$ if $a = a_s$ then $b(f,a_s,a_s) = 1$. This is the maximum value that can be attained since it corresponds to the sum of the maximum possible values of each term. As a moves away from a_s the magnitude of the beam pattern decreases. Holding f fixed, let a_0 be the largest value of a smaller than a_s such that $b(f,a_0,a_s)$ equals some cut level b_c. Similarly, let a_1 be the smallest value of a larger than a_s which causes $b(f,a_1,a_s)$ to equal the same cut level. If this cut level equals 1/2 then the difference between a_1 and a_0 defines the beam width.

Assume that for a specific frequency $f = f_0$ and steering angle $a_s = a_{s0}$, values of $a_0 = a_{00}$ and $a_1 = a_{10}$ are found such that the magnitude of the beam pattern equals b_c. Given any other frequency and steering angle, values of a_0 and a_1 yielding the same values of b_c can be found by solving the equations

$$f(\cos(a_s)-\cos(a_0))/c = f_0(\cos(a_{s0})-\cos(a_{00}))/c \quad (5.21)$$

$$f(\cos(a_s)-\cos(a_1))/c = f_0(\cos(a_{s0})-\cos(a_{10}))/c \quad (5.22)$$

The speed of sound at the array is assumed to be known and constant. If this is not the case, the effects of sound speed variations can be determined by properly perturbing the frequency values.

Define

$$c_0 = \cos(a_{00})-\cos(a_{s0}) \quad (5.23)$$

$$c_1 = \cos(a_{s0})-\cos(a_{10}) \quad (5.24)$$

Then

$$\cos(a_0) = \cos(a_s)+f_0c_0/f \quad (5.25)$$

$$\cos(a_1) = \cos(a_s)-f_0c_1/f \quad (5.26)$$

The values of c_0 and c_1 are determined by the choice of f_0 and the array geometry. They are fixed for a given array.

Consider the Taylor series expansion of cos(a) about the angle a_s. Retaining only the first two terms

$$\cos(a) \approx \cos(a_s)-(a-a_s)\sin(a_s) \quad (5.27)$$

Substituting values of a equal to a_0 and a_1 gives the approximations

$$a_0\sin(a_s) \approx a_s\sin(a_s)-f_0c_0/f \quad (5.28)$$

$$a_1\sin(a_s) \approx a_s\sin(a_s)+f_0c_1/f \quad (5.29)$$

Differencing the two equations gives the following approximation for the beam width as a function of frequency and search angle:

$$bw = a_1 - a_0 \approx \frac{f_0(c_0+c_1)}{f\sin(a_s)} \quad \text{(radians)} \qquad (5.30)$$

Based on the above result, the following observations can be made:

The beam width narrows as frequency is increased.

The beam width is narrowest at broadside ($a_s = \pi/2$) and increases as angles move toward endfire ($a_s = 0$ and π).

As a_s approaches 0 or π, the shape of the beam pattern main lobe changes as shown schematically in Fig. 5.2 [43].

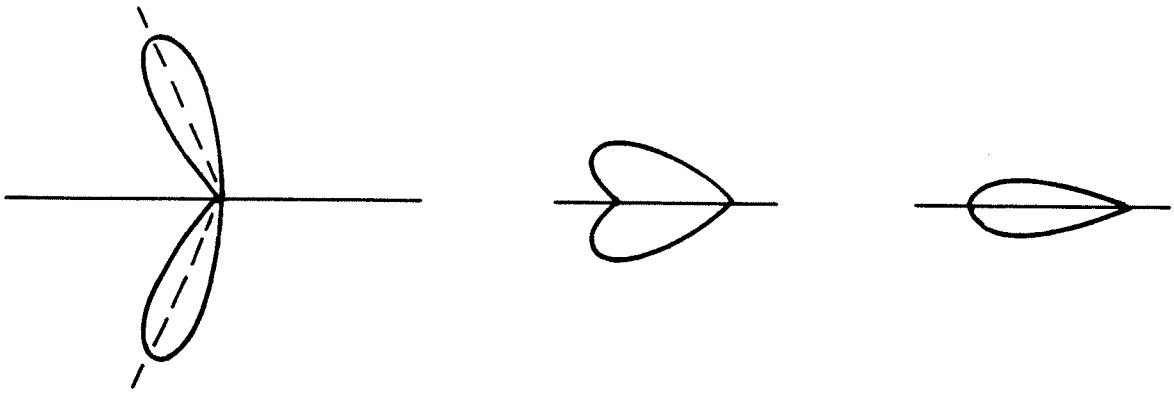


Fig. 5.2. Evolution of the beam pattern main lobe as it is steered toward endfire

The depth of the notch along the x-axis just makes the cut value b_c when

$$a_s = \cos^{-1}(1-f_0c_0/f) \tag{5.31}$$

$$\simeq \left(\frac{2f_0c_0}{f}\right)^{1/2} \text{ radians} \tag{5.32}$$

as the steered direction approaches 0 and

$$a_s = \cos^{-1}(-1+f_0c_1/f) \tag{5.33}$$

$$\simeq \pi-\left(\frac{2f_0c_1}{f}\right)^{1/2} \tag{5.34}$$

as the steered direction approaches π.

The width of the endfire beam in the $a_s = 0$ direction is

$$bw(a_s = 0) = 2\cos^{-1}(1-f_0c_1/f) \tag{5.35}$$

$$\simeq \left(\frac{8f_0c_1}{f}\right)^{1/2} \quad \text{(radians)} \tag{5.36}$$

and for the endfire beam in the $a_s = \pi$ direction it is

$$bw(a_s = \pi) = 2(\pi-\cos^{-1}(-1+f_0c_0/f)) \tag{5.37}$$

$$\simeq \left(\frac{8f_0c_0}{f}\right)^{1/2} \quad \text{(radians)} \tag{5.38}$$

A useful application of the equations for a_0 and a_1 is the setting up of a series of beams either side of the beam in the a_s direction which are down a specified number of dB at the points at which their responses cross. Given f_0, c_0, c_1 and a_s, a fan of "matched" beams either side of the beam centered on a_s have their crossover points, a_i, is defined by the equation

$$\cos(a_i) = \cos(a_s) + f_0 c_0/f - i f_0(c_0+c_1)/f \qquad (5.39)$$

where i can take on negative as well as positive values. Increasing values of i correspond to increasing angles because of the way c_0 and c_1 were defined.

The cosines of the corresponding beam centers a_{si} are given by

$$\cos(a_{si}) = \cos(a_s) - i f_0(c_0+c_1)/f \qquad (5.40)$$

Having developed some insight to the beam forming process, the equations for the noise improvement factors will now be derived. Closed form expressions for the noise improvement factor for the linear array can be obtained for both the spherically isotropic noise and horizontally arriving isotropic noise models.

In the case of spherically isotropic noise, the denominator is obtained by multiplying $b(f,a,a_s)$ by $\sin(a)/2$ and integrating from 0 to π. Doing this and substituting the result into the equation for $NIF_s(f,a_s)$ gives

$$NIF_s(f,a_s) = \qquad (5.41)$$

$$\frac{N}{1+\frac{2}{N}\sum_{i=1}^{N-1}\sum_{k=0}^{i-1}\frac{\cos(2\pi f(x_i-x_k)\cos(a_s)/c)\sin(2\pi f(x_i-x_k)/c)}{2\pi f(x_i-x_k)/c}}$$

Integrating $b(f,\theta,\theta_s)$ from 0 to π and substituting the result into the equation for $NIF_h(f,\theta_s)$ gives

$$NIF_h(f,\theta_s) = \qquad (5.42)$$

$$\frac{N}{1+\frac{2}{N}\sum_{i=1}^{N-1}\sum_{k=0}^{i-1}\cos(2\pi f(x_i-x_k)\cos(\theta_s)/c)J_0(2\pi f(x_i-x_k)/c)}$$

where $J_0(\ )$ is the zero order Bessel function as defined in [29].

A commonly analyzed array configuration uses uniformly spaced hydrophones. If the spacing between adjacent hydrophones is h, then the NIF equations become

$$NIF_s(f,a_s) = \frac{N}{1+\frac{2}{N}\sum_{i=1}^{N-1} \frac{(N-i)\cos(2\pi fih\cos(a_s)/c)\sin(2\pi fih/c)}{2\pi fih/c}} \tag{5.43}$$

and

$$NIF_h(f,\theta_s) = \frac{N}{1+\frac{2}{N}\sum_{i=1}^{N-1} (N-i)\cos(2\pi fih\cos(\theta_s)/c)J_0(2\pi fih/c)} \tag{5.44}$$

The equation describing the uniformly spaced array's beam pattern can be put into simple form after noting that the terms in the summation are terms which form a geometric series. Summing the series gives

$$b(f,a,a_s) = \left|\frac{\sin(\pi Nfh(\cos(a_s)-\cos(a))/c)}{N\sin(\pi fh(\cos(a_s)-\cos(a))/c)}\right|^2 \tag{5.45}$$

The quantity

$$f_0 = c/(2h) \tag{5.46}$$

is the frequency at which the inter-element spacing corresponds to one-half wave length and will be termed the array design frequency. The values of c_0 and c_1 are assumed to be computed for $f=f_0$.

The denominator of the beam pattern equation determines

the number and locations of the main lobes in the beam pattern. Primary lobes occur when the denominator goes to zero. This is always matched by a zero in the numerator with the ratio being equal to one. The numerator determines the width of the lobes.

The locations of the primary beam pattern response lobes are determined by the values of a which cause the argument of the sine function in the denominator to be equal to a multiple of π. That is

$$\frac{\pi f}{2f_0}(\cos(a)-\cos(a_s)) = k\pi \quad (k \text{ integer}) \qquad (5.47)$$

Solving

$$\cos(a) = \frac{2kf_0}{f} + \cos(a_s) \quad (k \text{ integer}) \qquad (5.48)$$

Since $|\cos(a)| \leq 1$, only a limited number of values of k yield lobes. Note that there is always a solution independent of f for k = 0 resulting in a beam response maximum in the steered direction.

Fig. 5.3 shows the locations of the main lobe centers for steered angles 0 degrees, 45 degrees, and 90 degrees as the operating frequency relative to the array design frequency is varied from 0 to 5.

Normally, the beam width for a given steering angle is defined as the angular separation between the smallest angles near a_s for which $b(f,a,a_s) = 1/2$. However, when working with functions of the form $\sin(Nx)/(N\sin(x))$, a particularly convenient definition of the beam width is the difference between x values such that $Nx = \pm\pi/2$. This definition is easy to work with and gives results which do not differ significantly from those obtained using the exact 3dB values.

Using the $x = \pm\pi/(2N)$ approximation gives the

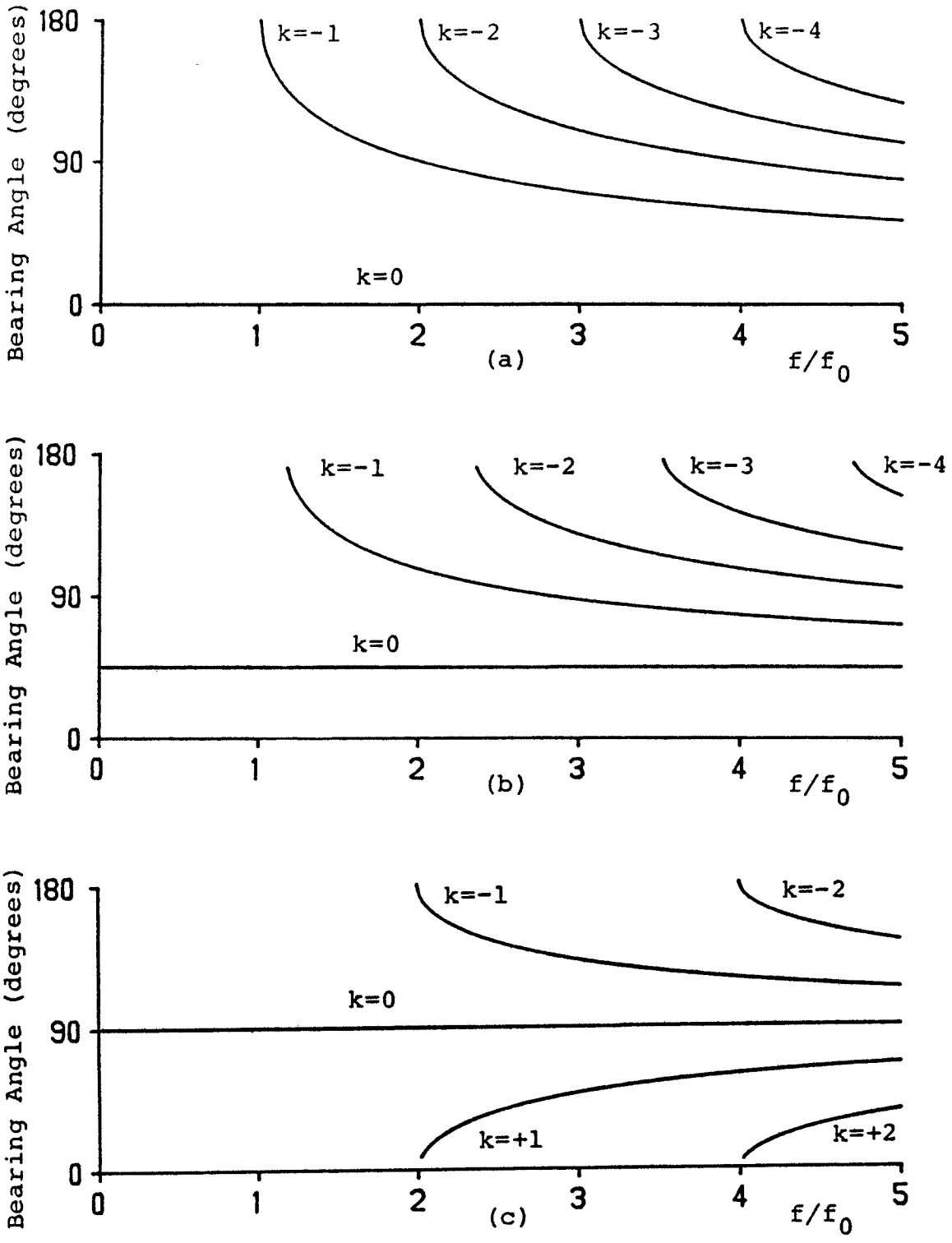


Fig. 5.3. Locations of primary lobes in the uniformly spaced array beam pattern as a function of f/f_0

a) Steering angle 0° b) Steering angle 45°

c) Steering angle 90° degrees

requirement that

$$\frac{Nf}{f_0}(\cos(a_s)-\cos(a)) = \pm 1 \qquad (5.49)$$

The values of a_0 and a_1 defining the beam edges are determined by the equations

$$\cos(a_0) = \cos(a_s)+\frac{f_0}{Nf} \qquad (5.50)$$

$$\cos(a_1) = \cos(a_s)-\frac{f_0}{Nf} \qquad (5.51)$$

Thus, for the linear array

$$c_0 = c_1 = 1/N \qquad (5.52)$$

Substituting this result into the expression for the beam width yields

$$bw = \frac{2f_0}{Nf\sin(a_s)} \quad \text{(radians)} \qquad (5.53)$$

Since the beam width is inversely proportional to the operating frequency, it might be expected that the NIF would also be inversely proportional to frequency. Indeed this is the case in the frequency zones between which spurious major side lobes appear. Fig. 5.4 is a plot of NIF_s for a uniform 20 element array and steered angles ranging from 0 to 90 degrees in 15 degree steps. The number of lines plotted makes the correspondence between lines and steered angles difficult to determine, however, the plot is primarily intended for qualitative use. The general character of the plot does not vary much for values of N equal to 10 or greater. Increasing N primarily smooths out the ripples caused by variations in the smaller sidelobes.

The noise rejection capability improves steadily as the operating frequency goes from zero to the array design frequency, f_0. At f_0, the performance drops for angles near endfire as spurious major lobes begin to appear in the beam pattern. For steering angles near broadside, it takes longer for these side lobes to appear and performance continues to increase until then. For frequencies higher than twice the array design frequency, the performance lies within ± 1.5 dB of 10 log(N) with the bound becoming tighter as f continues to increase.

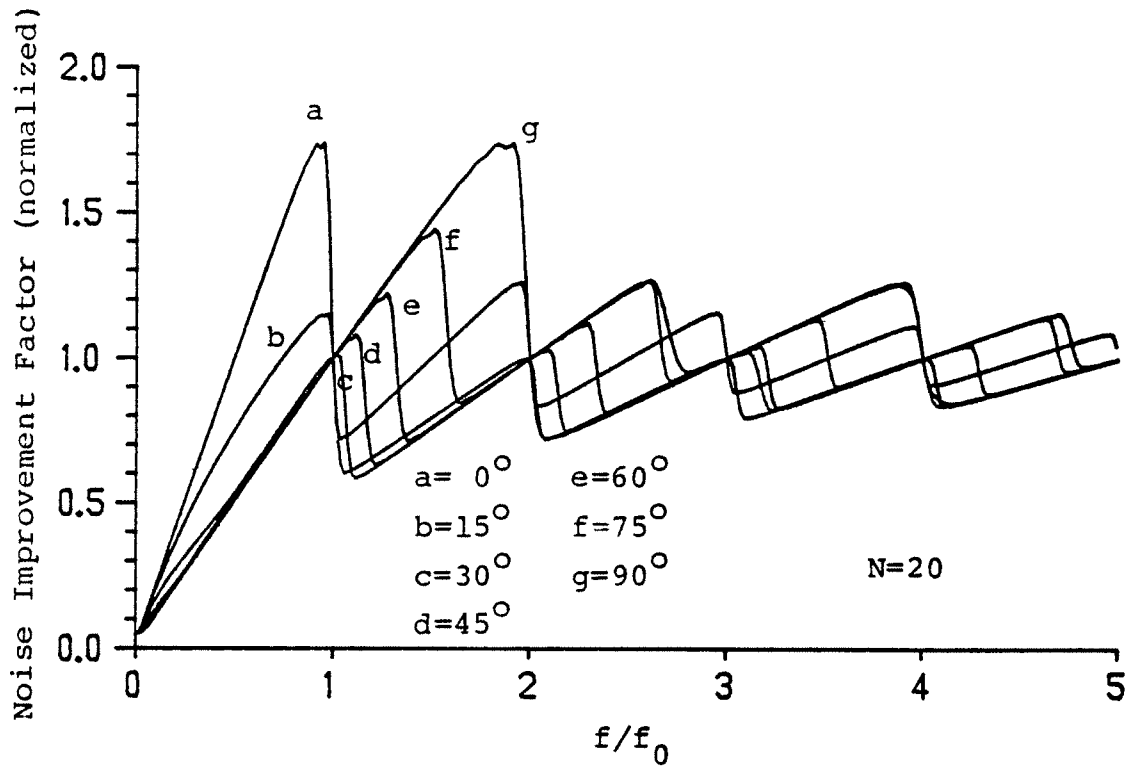


Fig. 5.4. Noise improvement factor for the equally spaced array and spherically uniform noise, N = 20. Steered angles range from 0 degrees thru 90 degrees in 15 degree steps

Using the same set of parameters, the noise improvement factor was computed for the horizontally arriving uniform noise model. The result is plotted in Fig. 5.5. The plots are significantly different from those obtained using the

uniform spherical noise model.

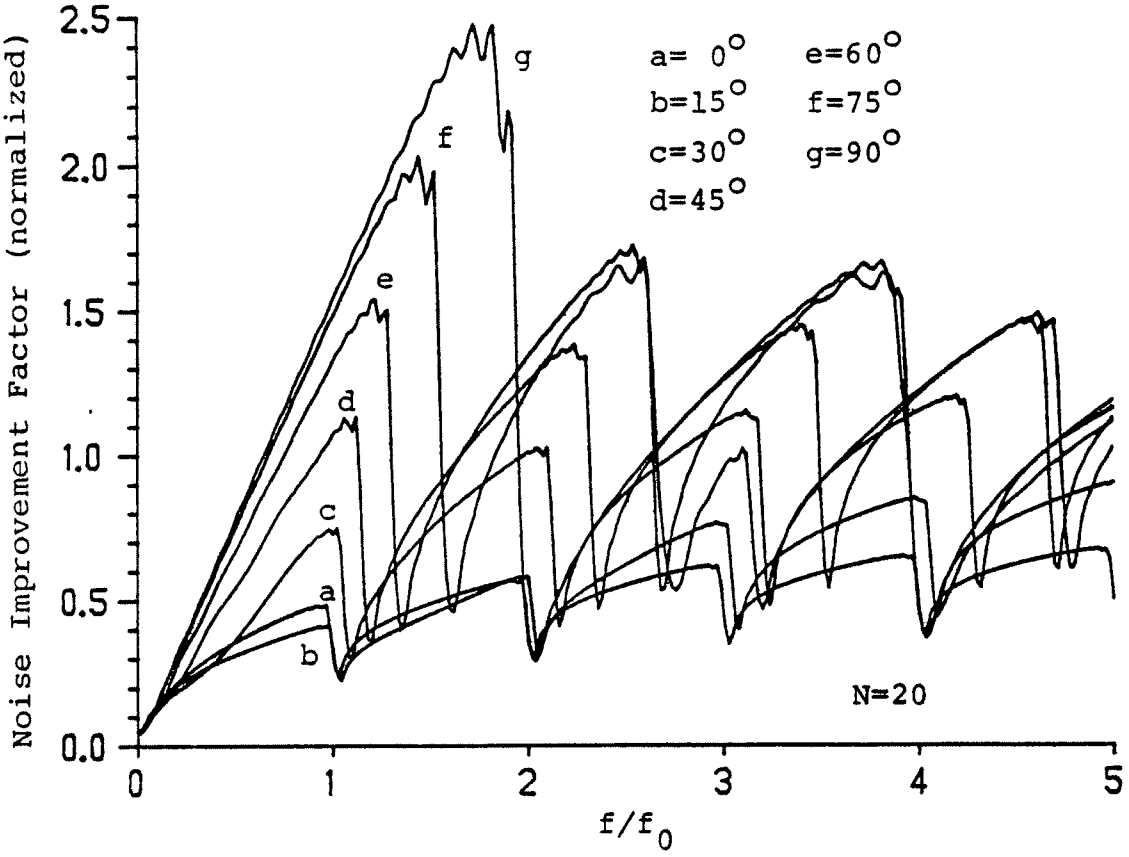


Fig. 5.5. Noise improvement factor for the uniform array and horizontally arriving uniform noise, N = 20. Steered angles range from 0 degrees thru 90 degrees in 15 degree steps

For the spherical model, a beam at steering angle a_s has a nominal beam width of

$$bw = \frac{f_0(c_0+c_1)}{f\sin(a_s)} \tag{5.54}$$

The noise contribution of this beam is equal to the area on the sphere swept out as the beam pattern is rotated through 2π and is approximately

$$2\pi\sin(a_s)\frac{f_0(c_0+c_1)}{f\sin(a_s)} = \frac{2\pi f_0(c_0+c_1)}{f} \tag{5.55}$$

Thus, for the spherical noise model, the amount of noise in a beam output is at most only slightly sensitive to the steered angle. This is borne out by the plots of NIF_s.

In the case of horizontally arriving noise, the amount of noise present in the beam output is roughly proportional to the amount of arc swept out by the primary lobes on the unit circle, i.e., depends directly on the beam width. It should thus be expected that the NIF for the horizontal model is maximum at broadside and varies as $1/\sin(a_s)$. Recall that near end-fire, the nature of the beam pattern changes and the beam width approximation cannot be safely used.

The exceptionally good performance near broadside for the horizontal model is at first disturbing but is to be expected. For broadside or near broadside steering angles, the first set of spurious major lobes does not appear until f/f_0 approaches two. The horizontal NIF can be approximated by dividing the beam width into 2π. Doing this for the linear array

$$NIF_h \approx \frac{Nf}{f_0 \sin(a_s)} \tag{5.56}$$

Near broadside and $f/f_0 = 2$, the value of the noise improvement factor should be near 2 as indicated in the NIF_h plots.

The major difference in the results obtained using the two noise models is that the spherically uniform model predicts noise performance largely independent of steering angle while the horizontal model predicts good performance near broadside falling off toward endfire. The model to be used in practice depends on the expected nature of the noise sources.

5.1.1 Effect of Errors in the Delay Times

There are two potential sources of errors when setting up the delay times for use in a delay and sum beamformer: 1) the actual hydrophone locations may not be known with a high degree of accuracy; 2) when using digital processing, the achievable delay times are quantized. The effects of errors in the delay times are most easily studied in the frequency domain. The Fourier transform of the beamformer output $v_b(t)$ is

$$V_b(f) = V(f)\sum_{i=0}^{N-1} w_i \exp\{-j2\pi f(t_i-(d-d_i)/c)\} \qquad (5.57)$$

Set the weights to unity. Let the error in the t_i value used to set the delay for the i-th hydrophone output be e_i. Then

$$V_b(f) = V(f)\sum_{i=0}^{N-1} \exp\{-j2\pi f e_i\} \qquad (5.58)$$

The main effect of making errors in the values of delay required for a specific steering or look direction is that of filtering the output of an ideally adjusted beamformer. The filter transfer function is

$$E(f) = \sum_{i=0}^{N-1} \exp\{-j2\pi f e_i\} \qquad (5.59)$$

The maximum value of E(f) is equal to N. A bound on the minimum value that E(f) can take on as a function of f provides a lower limit on performance.

Let the e_i values be limited to the range

$$-e \leq e_i \leq e \tag{5.60}$$

and restrict attention to values of f in the range

$$\frac{-1}{4e} < f < \frac{1}{4e} \tag{5.61}$$

The terms in the sum making up E(f) are complex numbers with unit magnitude and angles θ_i restricted to some range depending on f

$$-\pi/2 < -\theta_e \leq \theta_i \leq \theta_e < \pi/2 \tag{5.62}$$

where

$$\theta_e = 2\pi fe \tag{5.63}$$

Assume that for a given value of f, a set of θ_i have been found such that $|E(f)|$ is minimized. Let θ_k be one of these values. E(f) can then be written as

$$E(f) = e^{-j\theta_k} + \sum_{\substack{i=0 \\ i\neq k}}^{N-1} e^{-j\theta_i} \tag{5.64}$$

Let R be the magnitude of the above sum and ϕ the associated angle. Since $|\theta_i| \leq \theta_e$, the value of ϕ is similarly restricted.

The magnitude squared of E(f) is then

$$|E(f)|^2 = 1 + R^2 + 2R\cos(\theta_k - \phi) \tag{5.65}$$

This expression is minimum when θ_k has the opposite sign as ϕ and takes on magnitude θ_e. Since θ_k was arbitrarily selected, this implies that all θ_i have magnitude θ_e.

It now remains to determine the number of positive and negative values of θ_i. Assume p positive values and N-p negative values. Then

$$E(f) = p\exp\{-j\theta_e\} + (N-p)\exp\{j\theta_e\} \qquad (5.66)$$

$$|E(f)|^2 = N^2 + 2p(N-p)(\cos(2\theta_e)-1) \qquad (5.67)$$

Checking the first and second derivatives with respect to p, it is seen that $|E(f)|^2$ is minimized when $p = N/2$ when N is even. If N is odd, then $p = (N\pm1)/2$ does the job.

For even values of N, the minimum value of $|E(f)|^2$ is

$$|E(f)|^2_{min} = N^2\cos^2(\theta_e) \qquad (5.68)$$

For a given look direction and error bound, the magnitude of the error filter transfer function lies between this value and $|E(f)|^2 = N^2$. The frequency at which the worst case error filter response is down by D dB from the ideal is given by

$$f = \frac{1}{2\pi e}\cos^{-1}(10^{-D/20}) \qquad (5.69)$$

For a time delay uncertainty of ±0.1 ms (the time step error in the CEL beamformer assuming rounding), the 1dB down point is 749 Hz,. At 500 Hz the error filter transfer function is down no more than 0.44 dB. For comparison, an error in the delay values of ±0.1 ms corresponds to an uncertainty of ±5.9 inches in the hydrophone location.

The expression for $|E(f)|^2_{min}$ for N odd is

$$|E(f)|^2_{min} = N^2\cos^2(\theta_e) + \sin^2(\theta_e) \qquad (5.70)$$

For f = 500 Hz and e = $\pm$0.1 ms, the value of θ_e is 0.31 radians. The sine squared of 0.31 radians is 0.09 compared to the cosine squared value of 0.91. Thus, for this range of values, it is reasonable to neglect the sine squared term in the above equation and not distinguish between even and odd values of N. The equation derived for even N can be used as an approximation when N is odd.

5.1.2 Effect of Steering Error

Given a specific array, once the values of c_0 and c_1 have been determined, the equations for a_0 and a_1 can be used to study the effects of errors made in setting the steering angle. The equations for a_0 and a_1 are repeated here for convenience.

$$\cos(a_0) = \cos(a_s)+f_0c_0/f \tag{5.71}$$

$$\cos(a_1) = \cos(a_s)-f_0c_1/f \tag{5.72}$$

Having determined a_0 and a_1 the contours of the associated attenuation level can be plotted as a function of the vertical and horizontal arrival angles using the relation

$$\cos(a) = \cos(\theta)\cos(\phi) \tag{5.73}$$

This can be done for a number of frequencies. The expressions for a_0 and a_1 indicate that as frequency is increased, both angles approach a_s. A major effect of making a steering error is to cause the received waveform to be low pass filtered. For broad band receptions, it is wise to use the highest frequency of interest in the reception as the value of f when studying how accurately the arrival angles must be known.

Example

Consider a 10 hydrophone array with uniform spacing between phones corresponding to one-half wavelength at 100 Hz. The expression describing the beam pattern is

$$b(f,a,a_s) = \left|\frac{\sin(\pi f(\cos(a_s)-\cos(a))/20)}{10\sin(\pi f(\cos(a_s)-\cos(a))/200)}\right|^2 \quad (5.74)$$

Setting $f = 100$ and $a_s = \pi/2$ and solving the above equation for the $b(100,a,\pi/2) = 1/2$ points using a hand calculator gives

$$a_{00} = 84.90 \text{ degrees} \quad (5.75)$$

$$a_{10} = 95.10 \text{ degrees} \quad (5.76)$$

The equations for $a_0$and $a_1$then are

$$\cos(a_0) = \cos(a_s)+8.88943/f \quad (5.77)$$

$$\cos(a_1) = \cos(a_s)-8.88943/f \quad (5.78)$$

Assume that it is desired to use this array at 200 Hz, bearing angle 60 degrees and vertical arrival angle 10 degrees

$$\cos(a_s) = \cos(60^\circ)\cos(10^\circ) = 0.492404 \quad (5.79)$$

$$f_0/f = 1/2 \quad (5.80)$$

The corresponding values of $a_0$and a_1 are determined by the equations

$$\cos(a_0) = 0.536851 \quad (5.81)$$
$$\cos(a_1) = 0.447957 \quad (5.82)$$

Using these values and the relation

$$\cos(a) = \cos(\Theta)\cos(\phi) \tag{5.83}$$

the contour of the 3dB down values can easily be calculated and plotted in the Θ,ϕ plane. Fig. 5.6 shows the result of doing this.

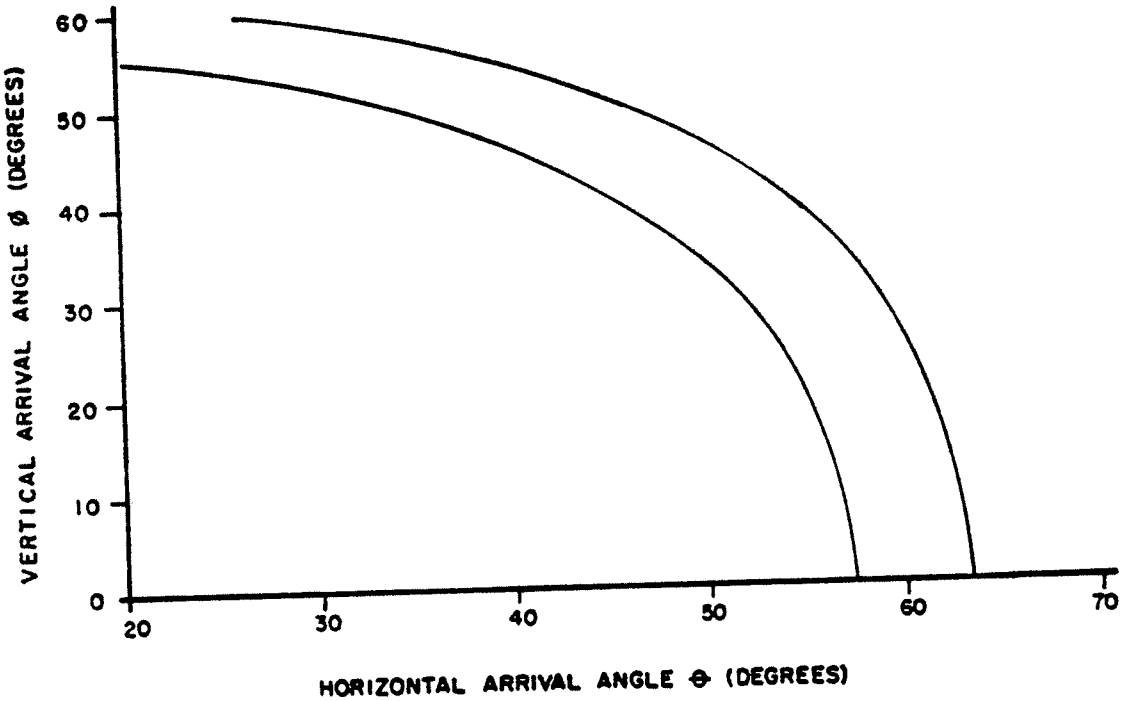


Fig. 5.6. Trajectories of the 3 dB down points

At the selected elevation angle, the 3dB down points correspond to bearing angles of 56.97° and 62.94°, giving a beam width of 5.97° in the horizontal. If the signal arrives from the 60° direction in the horizontal, then values of vertical arrival angle in the range from 0° to 26.37° provide a response which is less then or equal to 3dB down from the maximum.

End of Example

5.2 Demodulate, Delay and Sum Beam Forming

Demodulate, delay and sum beamforming consists of complex demodulating the sampled input data from the frequency band of interest to baseband, delaying the demodulated values and combining the resulting values from several channels as needed for beam forming. Motivating this approach is the observation in Section 5.1 that the allowable delay quantization was determined by the input waveform bandwidth. By demodulating the individual hydrophone waveforms prior to delaying them, the bandwidth can be reduced to its minimum value as determined by experiment design constraints.

The operation of complex demodulation consists of multiplying each hydrophone output by the complex rotator and filtering the result to eliminate the energy outside of the frequency band of interest. The waveform at the output of the i-th demodulator due to a plane wave reception at the array is

$$v_{d,i}(t) = v(t-(d_i-d)/c)\exp\{-j2\pi f_d t\} \circledast h(t) \qquad (5.84)$$

The Fourier transform of $v_{d,i}$ is

$$V_{d,i}(f) = H(f)V(f+f_d)\exp\{-j2\pi(d_i-d)(f+f_d)/c\} \qquad (5.85)$$

The purpose of H(f) is to attenuate the spectrum of the frequency shifted input waveform outside of some band of interest centered at f=0. Let

$$V_d(f) = H(f)V(f+f_d) \qquad (5.86)$$

where $V_d(f)$ has bandwidth W Hz.

The delaying of the waveform $v_{d,i}(t)$ by t_i seconds is accounted for in the frequency domain by multiplying the Fourier transform by $\exp\{-j2\pi f t_i\}$. The Fourier transform of

the sum of N such delayed versions is (again using unit weights)

$$V_b(f) = V_d(f)\sum_{i=0}^{N-1} \exp\{-j2\pi f(t_i+(d_i-d)/c\}\exp\{-j2\pi f_d(d_i-d)/c\} \tag{5.87}$$

The second factor in the above equation results from using the same complex rotator phase in all of the demodulators. This factor is independent of frequency and must be corrected for by multiplying the i-th demodulator output by $\exp\{-j2\pi f_d t_i\}$. Doing this, the transform of the beamformer output becomes

$$V_b(f) = \tag{5.88}$$
$$V_d(f)\sum_{i=0}^{N-1} \exp\{-j2\pi f(t_i+(d_i-d)/c)\}\exp\{-j2\pi f_d(t_i+(d_i-d)/c)\}$$

The beam is steered using $t_i = (2d-d_i)/c$ giving

$$V_b(f) = NV_d(f)\exp\{-j2\pi fd/c\}\exp\{-j2\pi f_d d/c\} \tag{5.89}$$

which corresponds to N times the output of a single hydrophone located at the array reference point, delayed by d/c seconds and phase adjusted.

Next, the effects of delay time quantization will be considered.

Assume that the arrival direction is known exactly but that time values are quantized such that

$$t_i+(d_i-d)/c = d+e_i \tag{5.90}$$

However, it is assumed that the values of the phase factors $\exp\{-j2\pi f_d(t_i+(d-d_i)/c)\}$ are computed exactly. Using the

above t_i values, the expression for the transform of the time quantized beamformer is

$$V_{bq}(f) = V_b(f)\sum_{i=0}^{N-1} \exp\{-j2\pi f e_i\} \qquad (5.91)$$

The effect of quantizing the delays in the demodulate, delay and sum beamformer on the output spectrum is the same as for the simple delay and sum beamformer. The important difference is that the filtering caused by the delay quantization errors affects the base band spectrum rather than the entire input spectrum.

Assume an error criterion of a maximum of D dB fall off in the beamformer response due to the effects of time quantization at the 3 dB points of the demodulator outputs. Then using the result of Section 5.1.1

$$\frac{W}{f_s} = \frac{2}{\pi}\cos^{-1}(10^{-D/20}) \qquad (5.92)$$

If D=1 dB, then the minimum allowed sample rate is

$$f_s = 3.33W \qquad (5.93)$$

In the Tomography experiment the signal bandwidth was 16 Hz. This gives a minimum value of f_s = 53.3 Hz. This corresponds to one demodulator sample value every 4.2 carrier cycles. The planned demodulator sampling rate of one per every 3.5 carrier cycles used in the Tomography experiment would have been adequate for use with this beam forming method.

5.3 Phase Shift Beam Forming

Phase shift beam forming consists of altering the phases of the individual hydrophone outputs and summing the results. This procedure is intended for use with narrow band waveforms which have negligible change in their amplitude and phase while propagating across the receiving array. Since a simple phase shift corresponds to multiplying a complex demodulator output by a unit magnitude complex number, the phase shift beamformer will be modeled by assuming the individual hydrophone outputs are demodulated, multipled by a phase factor and then summed.

Let the values used to phase the demodulator outputs be written as

$$\exp\{j\theta_i\} \tag{5.94}$$

Following the same analysis procedure used in Section 5.2, the Fourier transform of the phase shift beamformer output at baseband is

$$V_b(f) = V_d(f)\exp\{j2\pi(f+f_d)d/c\}\sum_{i=0}^{N-1}\exp\{-j2\pi(f+f_d)d_i/c\}\exp\{j\theta_i\} \tag{5.95}$$

Set $\theta_i = 2\pi f_d(d_i-d)$. The above equation becomes

$$V_b(f) = V_d(f)\sum_{i=0}^{N-1}\exp\{-j2\pi f(d_i-d)/c\} \tag{5.96}$$

Even though the array is steered in the direction of the arriving wavefront, the phase shift beam forming process has the effect of a filter on the corresponding time waveform.

For the linear array

$$d_i = d - x_i \cos(a) \tag{5.97}$$

Steering in the a_s direction requires using phase angles

$$\theta_i = 2\pi f_d (d - x_i \cos(a_s))/c \tag{5.98}$$

The resulting expression for the beam pattern is

$$b(f, f_d, a, a_s) = \left| \frac{1}{N} \sum_{i=0}^{N-1} \exp\{j2\pi x_i (f\cos(a) - f_d(\cos(a_s) - \cos(a))/c\} \right|^2 \tag{5.99}$$

Values of f in the above equation correspond to baseband frequencies. Setting f=0 in the above equation is equivalent to setting $f=f_d$ in the corresponding equation for the delay and sum beamformer. Doing so shows that, for operation at f_d, both systems possess the same beam pattern and thus perform the same. However, the extra term in the exponential of the phase shift beam pattern equation indicates a dependence on frequency that is not present in the delay and sum beamformer.

The beam pattern is at its maximum value if

$$f\cos(a) - f_d(\cos(a_s) - \cos(a)) = 0 \tag{5.100}$$

Given steering angle a_s, the direction of the maximum response is given by the solution of the equation

$$\cos(a) = \frac{\cos(a_s)}{1 + f/f_d} \tag{5.101}$$

When f=0, the direction of the beam pattern maximum is

in the steered direction. As f increases from 0, the direction of the maximum shifts toward broadside. As f decreases from 0, the maximum shifts toward endfire. Unlike for the delay and sum beamformer, the direction of the beam pattern maximum of the phase shift beamformer depends on frequency.

Setting f=0 for a given steering angle a_{s0}, two angles a_{00} and a_{10} can be found bracketing a_{sg} such that b( ) equals some cut value. This is done in the same manner as for the delay and sum beamformer and the values obtained in that case for $f=f_d$ will be the same as those obtained for the phase shift beamformer for f=0.

Define

$$c_0 = \cos(a_{00})-\cos(a_{s0}) \tag{5.102}$$

$$c_1 = \cos(a_{s0})-\cos(a_{10}) \tag{5.103}$$

The values of a yielding the same value of b(.) can be determined using the equations

$$\cos(a_0) = \frac{\cos(a_s)+c_0}{1+f/f_d} \tag{5.104}$$

$$\cos(a_1) = \frac{\cos(a_s)-c_1}{1+f/f_d} \tag{5.105}$$

Using the first two terms in the Taylor series expansion of cos(a) about a_s and assuming that the differences $|a_s-a_0|$ and $|a_s-a_1|$ are small, yields approximations for a_0 and a_1. Differencing these approximations gives the following estimate of angular beam width

$$bw = a_1-a_0 \simeq \frac{f_d(c_0+c_1)}{(f+f_d)\sin(a_s)} \tag{5.106}$$

This approximation should not be used in the vicinity of endfire because of the manner in which the beam pattern changes there.

The beam width increases as the frequency increases. This is the same behavior as for the delay and sum beamformer.

The behavior of the a_0 and a_1 contours as frequency is varied is as shown in Fig. 5.7. The direction of motion changes sign at $a = \pi/2$.

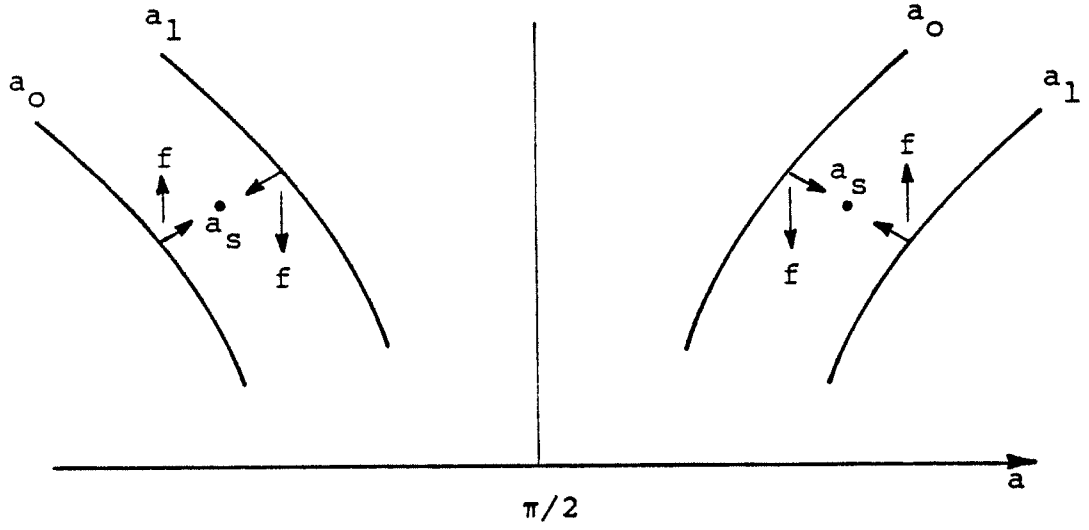


Fig. 5.7. Dependence of a_0 and a_1 on frequency variations

If a_0 and a_1 correspond to the -3dB point on the main beam pattern lobe (assuming that the steered angle is less than $\pi/2$), as f is increased, a_0 moves toward a_s. Let f_u be the upper frequency value such that $a_0 = a_s$. Then

$$f_u/f_d = c_0/\cos(a_s) \tag{5.107}$$

As f is lowered, the a_1 curve moves toward a_s. Let f_l be the low frequency value such that $a_1 = a_s$. Then

$$f_l/f_d = -c_1/\cos(a_s) \tag{5.108}$$

For steering angle a_s the bandwidth is

$$BW = f_u - f_l = \frac{f_a(c_0+c_1)}{|\cos(a_s)|} \tag{5.109}$$

The absolute value sign take care of the case where a_0 and a_1 are greater than $\pi/2$.

If a_s is such that a_0 and a_1 straddle $\pi/2$, then there is only an upper cut frequency. For $a_s<\pi/2$

$$f_u = \frac{f_d c_0}{\cos(a_s)} \tag{5.110}$$

and similarly if $a_s>\pi/2$

$$f_u = \frac{-f_d c_1}{\cos(a_s)} \tag{5.111}$$

If $a_s=\pi/2$ then, not unexpectedly, the bandwidth is infinite. For this angle the delay and sum and the phase shift beamformers are equivalent.

Performing the necessary integrations, the equations for the spherically isotropic and horizontally arriving uniform noise models are found to be

$$NIF_s(f,f_d,a_s) = \tag{5.112}$$

$$\frac{N}{1+\frac{2}{N}\sum_{i=1}^{N-1}\sum_{k=0}^{i-1}\frac{\cos(2\pi f_d(x_i-x_k)\cos(a_s)/c)\sin(2\pi(f+f_d)(x_i-x_k)/c)}{2\pi(f+f_d)(x_i-x_k)/c}}$$

$$NIF_h(f,f_d,\theta_s) = \tag{5.113}$$

$$\frac{N}{1+\frac{2}{N}\sum_{i=1}^{N-1}\sum_{k=0}^{i-1}\cos(2\pi f_d(x_i-x_k)\cos(\theta_s)/c)J_o(2\pi(f+f_d)(x_i-x_k)/c)}$$

Assuming a uniformly spaced array with inter-element spacing h gives

$$NIF_{su}(f,f_d,a_s) = \frac{N}{1+\frac{2}{N}\sum_{i=1}^{N-1}\frac{(N-i)\cos(2\pi f_d ih\cos(a_s)/c)\sin(2\pi(f+f_d)ih/c)}{2\pi(f+f_d)ih/c}} \quad (5.114)$$

$$NIF_{hu}(f,f_d,\theta_s) = \frac{N}{1+\frac{2}{N}\sum_{i=1}^{N-1}(N-i)\cos(2\pi f_d ih\cos(a_s)/c)J_o(2\pi(f+f_d)ih/c)} \quad (5.115)$$

Summing the series for the beam pattern gives for the uniform array

$$b(f,f_d,a,a_s) = \left|\frac{\sin(\pi Nh(f\cos(a)-f_d(\cos(a_s)-\cos(a))/c)}{N\sin(\pi h(f\cos(a)-f_d(\cos(a_s)-\cos(a))/c)}\right|^2 \quad (5.116)$$

Fig. 5.8 consists of plots of NIF_s for a uniformly spaced array of twenty hydrophones and values of f_d/f_0 equal to 1, 1.5, 2 and 4. Steering angles ranging from 0 through 90 degrees in 15 degree steps are included. As for the delay and sum beamformer, these plots are primarily intended to provide qualitative insight into the performance. The frequency range is from $-f_d/2$ to $f_d/2$. For some sets of parameter values, the beam pattern does not have a main response lobe. When this happens, the array performance looks exceptionally good. Unfortunately, not only is the beam forming process not responding to the noise, it is also not responding to the desired waveform. This is what causes the graphs for f_d/f_0 equal to 1 and 2 to go off scale. For

given values of f/f_d, f_d/f_0 and a_s, if an integer k cannot be found such that

$$\cos(a) = \frac{2k+\frac{f_d}{f_0}\cos(a_s)}{\left(\frac{f}{f_0}+\frac{f_d}{f_0}\right)} \qquad (5.117)$$

then there will not be a major lobe in the beam pattern and the noise improvement factor will usually be excessively large.

Fig. 5.9 shows the corresponding results for the same parameter values for horizontally arriving uniform noise.

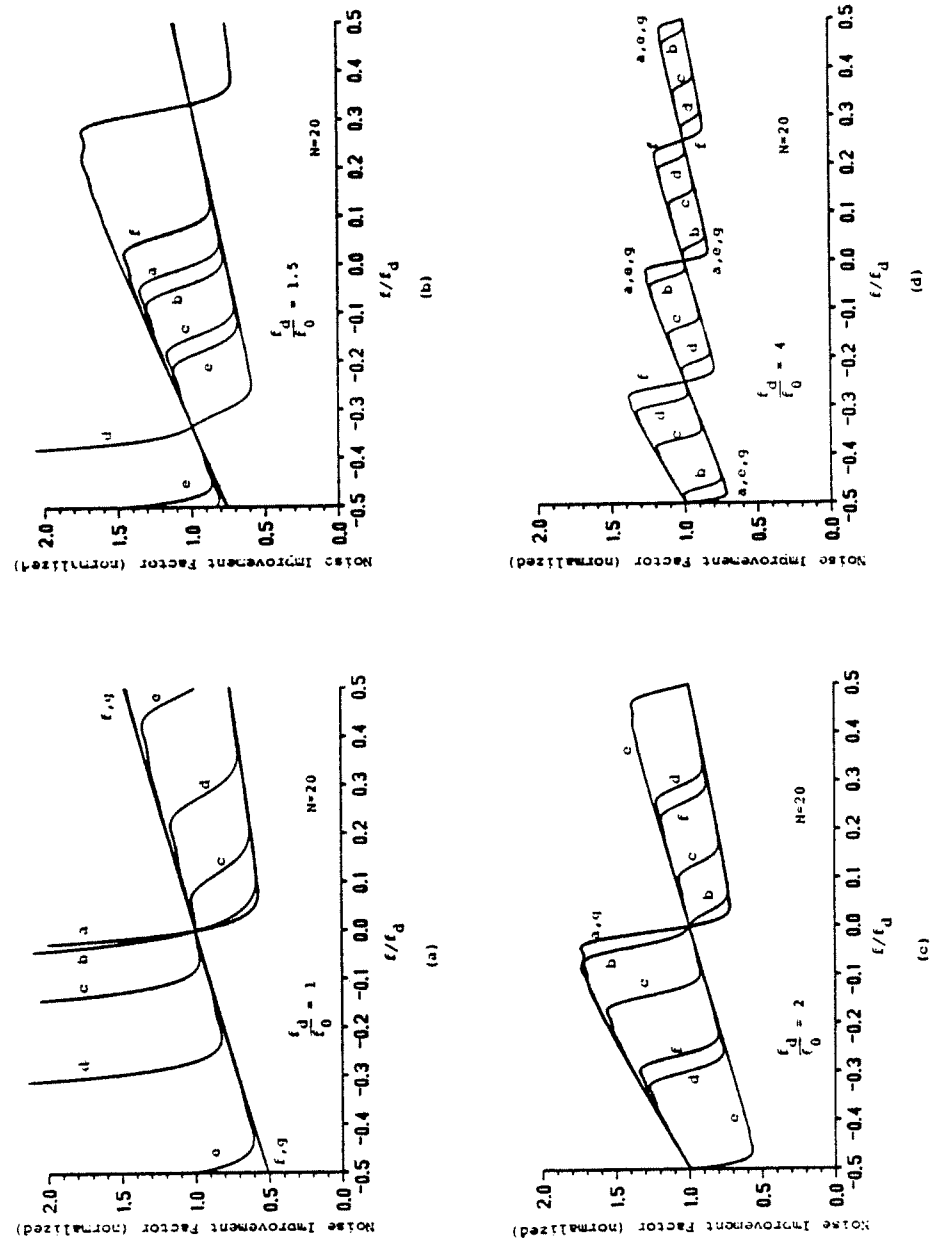


Fig. 5.8. Phase shift beamformer noise improvement factor for the equally spaced array and spherically isotropic noise, N=20. Steering angles range from 0 degrees thru 90 degrees in 15 degree steps as letters go from a thru g a) f_d/f_0=1 b) f_d/f_0=1.5 c) f_d/f_0=2 d) f_c/f_0=4

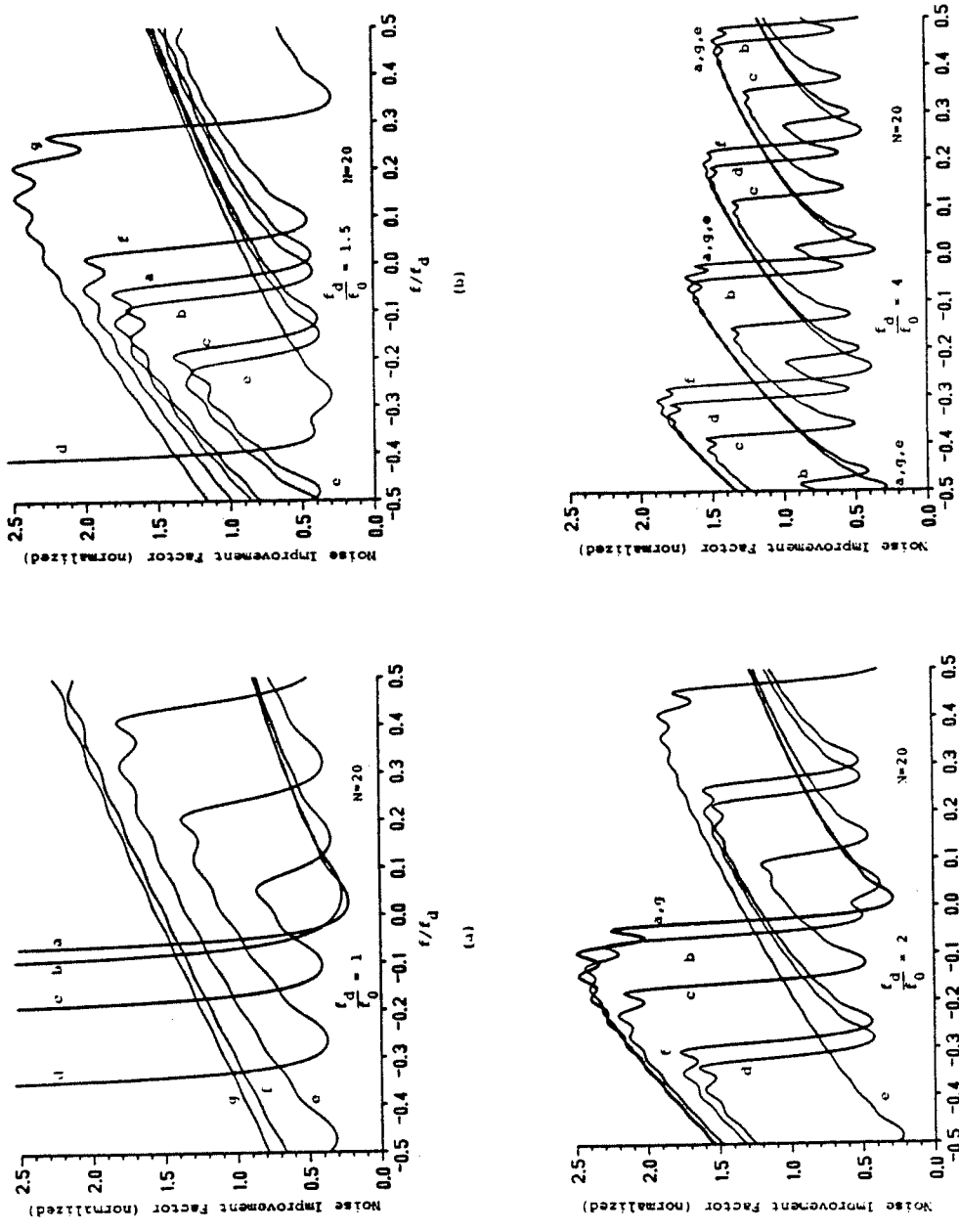


Fig. 5.9. Phase shift beamformer noise improvement factor for the equally spaced array and horrizontally uniform noise, N=20. Steering angles range from 0 degrees thru 90 degrees in 15 degree steps as letters go from a thru g
a) $f_d/f_0=1$ b) $f_d/f_0=1.5$ c) $f_d/f_0=2$ d) $f_d/f_0=4$

CHAPTER 6

BEAMFORMER AND SIGNAL PROCESSING HARDWARE

The data acquisition hardware designed and built at CEL for use in the 1981 Tomography experiment was divided into two major components. These were a beamformer and a signal processor. At the heart of both units is a set of logic cards designed around the Intersil/Harris IM6100 microcomputer. The computer system implemented by this card set was named the DM1200.

The beamformer unit is built in a table top equipment rack about 24 inches tall. Contained in this rack are a set of isolation amplifiers, a card cage containing the A/D converters and delay memories and a card cage containing the beamformer electronics.

The top 3 1/2 inches of space in the beamformer rack contains a set of commercially produced isolation amplifiers. These amplifiers electronically isolate the power and grounds of the source hydrophones and the CEL electronics. They also provide protection to the hydrophone electronics in case of a catastrophic failure in the CEL equipment. The isolation amplifiers are individually mounted on small printed circuit boards that plug into mother boards each of which can accept up to eight amplifier cards. Connection to the amplifier inputs is made via a phone jack panel located on the back of the isolation amplifier chassis. A matching set of phone jacks is available on the front panel to allow access to the signals. Both the front and back panels are non-metallic isolating the phone jacks not only from the CEL ground but also

isolating the grounds of the individual hydrophones.

The outputs of the isolation amplifiers feed individual A/D converter cards. Each card contains a digitally controlled amplifier, track and hold, A/D converter and 2048 words of memory used to contain the 2048 most resent sample values. The selected sample rate was 5 KHz providing a maximum delay of about 0.41 seconds. These cards are mounted in the card cage located just below the isolation amplifiers. Up to 24 cards can be contained in a single cage. It is possible to mount two cages back to back giving a maximum capability of 48 channels in a single beamformer rack. The use of one A/D converter per channel rather than multiplexing several channels onto a single converter may appear to be extravagant. However, there were significant advantages which were felt to justify the extra cost over that of multiplexed A/D converters. This approach allowed the design to be highly modularized. A single printed circuit card contains all of the electronics necessary to produce delayed sample values for a single hydrophone. Channels can be added to a system as needed. A failure on a single board only costs the loss of a single hydrophone output. Using an A/D converter per channel made it easy to simultaneously update all of the delay memories with new sample values taken at the same instant. This decision was also influenced by the availability of relatively low cost integrated A/D converters from Analog Devices.

The A/D delay memory cards plug into a bussed backplane. This backplane bus is connected to a control module contained in the card cage mounted below the one holding the A/D delay cards. The control module controls the timing of the A/D conversions, the updating of the delay memories, and access to these memories by various "users". Timing is derived from an external 5 MHz Rubidium frequency source.

Also located in the cage with the control module are the beamformer electronics and a DM1200 processor system.

The beamformer itself was implemented as a peripheral device to the DM1200.

The DM1200 system is located on the far right of the card cage. A number of card slots were dedicated for use as the DM1200 bus. Card positions near the center of the cage were specifically wired for use by the beamformer electronics and an interface thru which the DM1200 could access the delay memories. In the normal working configuration, the DM1200 system contains a processor board, a read/write memory board, a read only memory board, and two or three communications interfaces. The beamformer and the delay memory access hardware are implemented as special devices on the DM1200 bus. A few card positions are available for plugging in other interfaces.

The beamformer operation starts when the LOAD button is pressed on the DM1200. The processor copies an operating program and set of beam tables out of read only memory and places them into read write memory. The beam tables are then used to configure a control memory located in the beamformer. This control memory contains a list of A/D channels and the associated delays required to form from one to eight beams. The beamformer scans this table every time a new sample is placed into the delay memories. Beam outputs thus become available at a 5 KHz rate. The beam outputs are connected to the signal processing unit using a set of twisted pair cables. When a beam output set becomes available the processing hardware is signalled and serially shifts the resulting values into the demodulating hardware.

Once initialized, the beamformer is not dependent on the DM1200 for any additional support. Beam values are automatically formed and the demodulation hardware signalled. The only time processor intervention is required is when it is desired to change the beams. The control memory is organized so that, if care is taken, this can be done on the fly without affecting the timing.

In addition to setting up the beamformer and changing

beams as needed, the processor can be used to test the system for proper operation. This is usually not done while the beamformer is in active use. By reading the delay memory contents while data is being taken the processor can check the operation of the A/D converters. The sample values can be displayed on an oscilloscope if a D/A converter board is added to the DM1200. Using the delay memory access interface, the processor can stop the acquistion of data and write its own test patterns into the delay memories. It then can verify these patterns to see if the delay memories are working properly. The processor can also read the resulting contents of the beam output registers. Thus, it is possible to verify the proper operation of the system from the A/D converter outputs thru to the resulting beam outputs. The processor can also monitor performance on a less exhaustive basis during normal operation.

The signal processing unit consists of a DM1200 processor to which special devices have been added to demodulate the data and to speed up arithmetic operations. This unit is contained in a small instrument case along with the required power supplies and fans. Connection to the beamformer electronics is made using short cables made up using twisted pairs. A Digital Equipment Corporation LA34 terminal was used as the operator console and to print the data log. A separate small case was used to house a LINCtape drive. The LINCtape system was used to provide random access mass storage needed for program developement and for recording processed results. In 1982, support was added for cartridge tape output and 9-track tape output. The cartridge tape support was primarily in software since the tape system had a serial communication (RS-232) port. The interface for the 9-track tape drives was implemented using a wire-wrapped card cage containing a DM1200 bus with two card positions hard wired for use by the interface.

The demodulators were implemented using serial

multiplier integrated circuits. When a set of beam output samples becomes available, the values are shifted serially from the beam output registers thru the multipliers with the results being accumulated. As the values are being shifted thru the multipliers they are multiplied by sample values of $\exp\{-j2\pi f_d t\}$. The complex exponential values are contained in a read only memory located in the demodulator control electronics. The exponential values are selected using a direct digital frequency synthesizer to generate the required ROM addresses.

Once started, the operaton of the demodulator is automatic and does not require any processor intervention. When the proper number of demodulated beam outputs have been accumulated to form a demodulate the values are copied into a set of output registers (perhaps being scaled in the process) and a flag is set to inform the processor that a set of demodulates has been formed. The processor has a complete demodulation time interval in which to dispose of the demodulates.

Because the IM6100 microcomputer used in the DM1200 is so slow, it was necessary to provide a means by which complex arithmetic and multiplications could be performed rapidly. As much support as could be crammed onto a single wire-wrap panel was provided. Using this device, it is possible to accumulate complex double precision values and form double precision products of the form $R^2=X^2+Y^2$ in a matter of a few microseconds.

The primary device used on the DM1200 system for programming support and data recording is LINCtape. LINCtape drives are present on all CEL systems and offer ready interchange of programs and data. It is also possible to use CEL's Data System Design floppy disk systems on the DM1200.

The design and construction of the hardware described in this chapter benefitted from the contributions of a

number of individuals as noted in the acknowledgment. In addition to directing these efforts, the author's contributions included designing all of the interfaces described in this chapter, designing four printed circuit boards and producing all of the wire-wrapped panels (which were normally machine wrapped).

Sections 6.1 and 6.2 contain descriptions of the various interfaces and devices that are used to make up the beamformer and DM1200 processing unit. The amount of information presented is restricted to that which a programmer would need in order to make effective use of this equipment. Providing more detailed internal descriptions was felt to be beyond the scope of this document. A separate set of manuals is being prepared which will contain final form logic diagrams, board layouts and wire wrap documentation.

Section 6.1 contains the descriptions of the boards which make up the DM1200 and those interfaces not specific to the beamformer. Section 6.2 describes the isolation amplifiers, A/D delay cards, the delay bus controller, the processor delay bus interface and finally the beamformer itself.

6.1 Signal Processing Hardware Descriptions

This section contains descriptions of the modules which normally are used to make up the signal processing unit of the CEL data acquisition system. The first board described is the one containing the IM6100 microcomputer. This board provides the basic computing capabilities contained not only in the signal processing unit but also in the beamformer. Communication between the processor board and the various peripheral interfaces and devices is over a collection of signal lines called a bus. The signals present on this bus are described as part of the processor board description.

Following the processor board description are

descriptions of the other boards which are included in the signal processor. The ordering of the descriptions proceeds from memory to communication to special function and, finally, to mass storage.

The dimensions of the cards used to build up the devices described in this chapter are those used by AUGAT for their R-Series of logic modules. The same connector was also used. This allowed the use of commercially available card cages, connectors and wire-wrap cards.

6.1.1 Processor Board and Bus Description

The processor printed circuit board contains the following:

- Intersil/Harris IM6100 microcomputer
- Intersil IM6102 MEDIC support integrated circuit
- RUN/HALT and bootstrap control logic
- 32 word read only memory bootstrap program
- logic to generate initialize pulse on clear all flags (CAF) instruction
- bus receivers and drivers
- bus termination resistors

The IM6100 microcomputer is a CMOS version of the Digital Equipment Corporation PDP-8/e minicomputer. It uses a 12 bit word size and can directly address 4096 words of memory. The instruction set is described in [46] and will not be described here. Using a 4 MHz clock, the IM6100 runs about one-third as fast as the PDP-8/e. One of the unique additions to the PDP-8/e architecture is the provision for a control panel memory that is separate from the normal user memory. Using this feature it is possible to implement sophisticated program monitoring and debugging aids.

The Intersil IM6102 MEDIC chip comes in a 40 pin package and is specially designed to enhance the capabilities of the IM6100. The features provided by this device are:

- expansion of memory addressability to 32,768 words. This is done in a manner compatible with the PDP-8/e

memory extension hardware.
- real time clock
- dynamic random access memory refresh control signals
- interleaved zero overhead direct memory access (DMA) control

The MEDIC chip was included on the processor board specifically because of its memory extension capability. Neither the dynamic ram nor the zero overhead DMA support features have been used. The real time clock is really an interval timer and has only rarely been used. The accuracy of the interval timer is limited by the accuracy of the crystal used to generate the processor clock. A 0.02 per cent crystal was used on the processor board. It is possible to modify the board to accept an externally supplied 2 MHz clock for use with the real time clock. This was done on one board in order to allow use of a rubidium standard as the time base.

Intersil claims that both the IM6100 and the IM6102 are TTL compatible. However, the data sheet specifies that the lower value allowed for the high logic level is 70% of the supply voltage (5 volts). The minimum guaranteed voltage in the high state for TTL is 2.4 volts. This is about 0.9 volts below the 3.5 needed for the CMOS chips. This was accomodated for by placing pull-up resistors on the CMOS input lines and using the INTEL 8226 for the data lines. The 8226 has a guaranteed minimum of 3.65 volts on the high state. There are no problems with the IM6100 or the IM6102 driving either TTL or LS TTL logic as long as the loading is kept within the specified limits.

Provisions on the processor board for an operator's console are minimal. The provided console support consists of

- a push button used to activate the initial program loading sequence (bootstrap)
- a push button used to toggle the RUN/HALT state
- an indicator to show if the processor is in the RUN or the HALT state
- an indicator to show if the plus five volt power supply

is operating

Connection between the processor board and the front console hardware is made via a 16 wire ribbon cable which plugs onto the top left of the processor board. The pin assignments on this connector are:

Pin 1 Boot push button normally open side
Pin 2 Boot push button normally closed side
Pin 3 Run/halt push button normally open side
Pin 4 Run/halt push button normally closed side
Pin 5 Run indicator LED ground side
Pin 6 Run indicator LED positive side
Pin 7 5 volt power LED positive side
Pin 8 5 volt power
Pins 9 - 16 ground

The bootstrap procedure can be initiated by either pressing the the BOOT push button on the front console or by grounding the BOOT- line on the processor bus. The first step in the booting procedure is the generation of a system wide reset. Next, the bootstrap memory is activated. This memory overlays normal read/write memory and, when active, inhibits instruction fetches from the memory locations that it overlays. These are memory locations 7740 (8) thru 7777 (8). Execution of the bootstrap program by the IM6100 automatically starts at location 7777 (8). Any legal IM6100 instruction can be executed from the boot memory. Normally the bootstrap program consists of a small device handler which is used to read a control program into normal memory and then place it into execution. The boot memory is automatically disabled when the processor fetches an instruction from a memory location not contained within the boot memory address space.

In the PDP-8/e, the clear all flags (CAF) instruction is used to initialize the processor state as well as the states of all peripheral equipment attached to the processor. In the IM6100, the CAF initializes both the processor and the MEDIC. However, no control pulse is generated with which to initialize devices outside of these

two chips. Intersil recommends that each device interface decode the CAF instruction and take the proper action. Rather than have to cope with including the required logic on each device interface, it was decided to decode the CAF instruction on the processor board and generate an initialize pulse.

The signals brought to the processor card edge connector define the structure of the bus used with this card. One of the most important aspects of the design of a bus is its termination. Ideally, each bus line should be terminated at both ends using a load equal to the line's characteristic impedance. Doing this eliminates the possibility of noise generating reflections which is a problem in high speed systems. The characteristic impedance of printed circuit and wire-wrapped buses runs on the order of 100 to 150 ohms. Commonly used termination methods are:

- Series connection of the bus line thru a 120 ohm resistor to a "stiff" power supply. For a TTL compatible system, the supply voltage would be on the order of 3 volts.
- Connection of the bus line to the center of a resistive divider connected between the supply voltage and ground. For a TTL compatible system, a 220/330 divider is often used. The 220 ohm resistor connects to +5 volts and the 330 ohm resistor connects to ground. The common point is tied to the bus line.
- Connection to the bus line using a series connection of a resistor and a capacitor. This configuration reduces the dc load that must be driven but does this by adding extra capacitance to the bus which slows it up. By testing it was found that a 100 pf capacitor in series with 82 ohms made an effective termination.

Each of the above methods has its advantages and its drawbacks. The approach taken in the CEL system was to terminate each line with a 220/330 resistor divider directly on the processor board and use an R/C termination at the far end of the bus. This has worked well with both printed circuit buses and wire-wrapped back planes.

The DM1200 bus is built around the 122 pin connector used by Augat with its R-series of modules. Two versions of

the connector are available. One version accepts 1/8" thick modules and the other accepts 1/16" thick modules. The connector accepting 1/16" modules was chosen for use because it is easier to have printed circuit boards made with this thickness. Of the 122 pins, 4 were used for +5 volts power, 2 for +15 volts, 2 for -15 volts, 56 for ground connections, 48 for signal lines, 6 unassigned and 4 with no connection.

The bus has been implemented both as a printed circuit back plane and as a wire-wrapped back plane. The printed circuit version has provision for 18 modules. The connector side is a ground plane. The signal lines and ground lines are located on the back side. Ground lines are connected to the ground plane via printed-thru holes at each connector position. Each signal line has a ground line on either side of it. The processor card is normally located at one end of the bus and a passive terminator card is located at the other end of the bus. The positioning of other modules is arbitrary. Two wire-wrapped configurations have been implemented to date. The version used in the beamformer is a short bus extending over 9 card positions. It was not necessary to terminate this version at both ends of the bus. The wire-wrap version produced to be used with the 9 track tape system extends over the entire length of the 19 inch wide logic bay. It was necessary to terminate this bus at both ends.

The signals available on the bus are pretty much those provided on the IM6100 and IM6102 for bus use. A notable extension is the provision for an additional 3 bits of address. Using a special controller (not yet designed), it is possible to extend the memory address space up to 256K words. For the most part, signals on the bus are negative logic. A high level corresponds to the logical 0 state and a ground level corresponds to the asserted or logical 1 state. Care was exercised so that the unasserted states on the bus corresponded to the least hazardous function. The list given below identifies the signals present on the bus

and gives a brief description of their functions. For a more detailed description and information about timing, consult the Intersil documentation [46].

Pins 1-2, 5-57, 60-61	ground lines
Pins 62-63, 121-122	+ 5 volts power
Pins 3, 64	+15 volts power
Pins 4, 65	-15 volts power

Pin 8		This line was intended for use in chaining interrupt requests in the manner defined by the Intersil IM6101 Parallel Interface Element (PIE). No PIE units have been used in any of the interfaces built to date and none are anticipated. This line is not in active use.
Pin 66		Not assigned.
Pin 67		Not assigned.
Pin 68		Not assigned.
Pin 69		Part of the interrupt chaining described under Pin 8.
Pin 70		Not assigned.
Pin 71		Not assigned.
Pin 72		Not assigned
Pin 73	DMAEN-	This signal is used to request a transparent DMA cycle from the IM6102 DMA controller. This signal is not used with normal IM6100 DMA transfers.
Pin 74	DMAREQ-	Used to request a normal DMA memory cycle from the IM6100.
Pin 75	CPREQ-	Momentarily grounding this line causes the IM6100 to enter the control panel mode of operation. Memory fetches from normal memory cease and are instead made from control panel memory.
Pin 76	RUNHLT-	Pulsing this pin causes the RUN/HALT flip-

flop in the IM6100 to change state.

Pin 77 RESET- Causes the IM6100 to halt and sets the program counter to location 7777 (8) in memory field 0. Also causes the initialization signal INIT- to go to ground in order to initialize any peripherals that might be present.

Pin 78 SKPINT- During input/output transfer (IOT) instructions this line is used to request a skip of one instruction. During non IOT times this line is used to request a program interrupt. All interfaces must be designed to remove their interrupt requests from this line during the execution of an IOT instruction.

Pin 79 BOOT- Grounding this pin initiates the boot procedure.

Pin 80 IODIS- This line is used to disable device interfaces from responding to IOT instructions addressed to them. This signal was included for possible use in implementing a general I/O monitor in control panel memory. At present, not all device interfaces monitor this line.

Pin 81 MEMDIS- This pin is used to disable normal memory from responding to read or write operations. This line is usually activated by boards containing read only memory in order to prevent bus conflicts when RAM and ROM addresses are overlapped.

Pin 82 EMA2- The least significant bit in the normal 3 bit memory address extension. This line is stable during normal memory cycles and does not need to be latched on memory boards.

Pin 83 EMA1- The middle memory address extension bit. See EMA2-.

Pin 84 EMA0- The most significant memory address extension bit. See EMA2-.

Pin 85 C0- One of three control lines used to determine the type of data transfer to be made during an IOT instruction. The types of transfers made based on the states of the C0-, C1- and C2- control lines are:

C0-	C1-	C2-	Function
H	H	H	data is sent to device
L	H	H	data is sent to device then the AC is cleared
H	L	H	data from device is OR'ed into AC
L	L	H	data from device is loaded into AC
*	H	L	data from device is added to the program counter
*	L	L	data from device is loaded into the program counter

Pin 86 C1- See C0-.

Pin 87 C2- See C0-.

Pin 88 XTC+ Used to control the direction of data transfers on the bus. When low, data is to go to memory or, if an IOT instruction is being excuted, to a device. When high, the transfer direction is to the processor.

Pin 89 INTGNT- This signal is provided for use in extended memory addressing hardware. INTGNT- goes low whenever an interrupt request is granted. The line stays low until the first IOT instruction following the grant is executed.

Pin 90 DATAF- This signal goes low during the execute phase of indirectly addressed AND, TAD, ISZ and DCA instructions. It is intended for use in implementing extended memory addressing hardware.

Pin 91 RUN- This pin is driven low by the processor whenever the processor is in the run state.

Pin 92 DMAGNT- This pin goes low whenever a DMA request

has been granted by the IM6100. This signal is for use in making normal, non-transparent data transfers via DMA.

Pin 93 IFETCH- Indicates that the processor is fetching an instruction from memory during the current memory cycle.

Pin 94 EXT2- The least significant bit in the extra three address bits added to the bus. Although this feature is not presently used, all read/write memory boards check the EXT0- thru EXT2- address line when responding to memory access requests.

Pin 95 EXT1- The mid bit in the extra three memory address bits. See EXT2-.

Pin 96 EXT0- The most significant extra extended memory address bit. See EXT2-.

Pin 97 CPSEL- The control memory select pulse. This is only active when the processor is operating in the control panel mode. Used in conjunction with XTC+ to either write the contents of the data bus into control panel memory or to gate the control memory contents onto the data bus.

Pin 98 MEMSEL- The normal memory select pulse. Used in conjunction with XTC+ to either write the contents of the data bus into memory or to gate the memory contents onto the data bus.

Pin 99 INIT- This pin is used to initialize all interfaces connected to the bus. This line is driven to ground whenever the CAF instruction is executed, whenever the system is booted or whenever the RESET- line is driven to ground.

Pin 100 UP- This signal is used with the transparent DMA and dynamic memory refresh features of the IM6102. At present, no use is made of

this signal.

Pin 101 XTA- This control signal is generated by the IM6100 and indicates input data transfers to the processor. Not used by any interfaces.

Pin 102 LXMAR- Used by memory boards and I/O interfaces to load the contents of the data bus into local address registers. The trailing edge signifies a stable address on the data bus.

Pin 103 DEVSEL- Used to indicate that the current cycle corresponds to the execution of an IOT instruction. The timing on DEVSEL- is the same as for a memory cycle as indicated by MEMSEL-. This signal is used in conjuction with C0-, C1-, C2-, SKPINT- and XTC+ to determine the type of operation to be performed.

Pin 104 SWSEL- This signal is generated whenever operate instructions which access the switch register are executed. Its timing is the same as that of MEMSEL- and DEVSEL-. This level is not presently used.

Pin 105 4MHZ- A 4 MHz square wave. This signal is an inverted version of the 4 MHz clock used by the IM6100. This clock is an excellent source of timing signals for use in DMA interfaces.

Pin 106 LINK- This signal corresponds to the LINK bit contained in the IM6100. It is on the bus in case a more extensive front panel is implemented which would display this bit.

Pins 107 - 118 DX0- thru DX11-

These are the 12 lines which make up the data bus over which all data transfers are made. Depending on the current cycle, these lines may contain memory addresses,

data to be read from or written to memory, data passing between the processor and I/O devices, etc. Line DX0- is the most significant bit and line DX11- is the least significant bit.

6.1.2 Read/Write Memory Board (12K)

The 12K read/write memory board contains 12K words of 12 bits. The memory is organized as three 4096 word memory fields and is composed of 4K by 1 bit static RAMS. The IM6100 cycle time is such that memory chips with 450 ns access times can be used. However, low cost RAM chips were available with 250 ns access times and were used instead. The use of the faster memory chips should prove useful when the processor is upgraded to a faster unit. Their presence was also exploited in the 9 track tape DMA interface.

The low 12 address bits are latched on the board using the LXMAR- control line. The six extended memory address bits are not latched and must remain stable during all memory transfers. The EXT0- thru EXT2- lines are used to assign a board to one of 8 possible "super" banks of 32,768 words. The EMA0- thru EMA2- lines are decoded to allow assignment of any of the three 4K fields to any of the 8 possible field assignments in a given "super" bank.

"Super" bank and field assignments are made using jumper headers. Looking at the memory board from the component side with the connector in the down direction, the "super" bank assignments are made on the center left header position and normal field assignments are made on the center right header.

"Super" bank assignments are made by connecting any one of pins 1 thru 8 as follows:

to pin 15 for "super" bank 0 (normal)
to pin 14 for "super" bank 1
to pin 13 for "super" bank 2

to pin 12 for "super" bank 3
to pin 11 for "super" bank 4
to pin 10 for "super" bank 5
to pin 9 for "super" bank 6
to pin 8 for "super" bank 7

The three 4K fields on the board are named A, B and C. Field A is the field top most on the board, B is the center field and C is the field closest to the connector.

The field assignments within a "super" bank are made using jumpers to connect the decoded field addresses to the enable lines of the individual 4K fields. On the field selection header pin 1 corresponds to memory field A, pin 2 to memory field B and pin 3 to memory field C. Pins 15 thru 8 correspond to field addresses 0 thru 7 respectively.

Connecting pin 4 to pin 5 on the field address header allows memory field C to also respond to control panel read/write operations. This aliases the control panel memory with one field of standard memory. This feature was included in order to simplify the developement of control panel software.

The signals used by this board are:

LXMAR- CPSEL- XTC+ EXT0- EXT1- EXT2-
EMA0- EMA1- EMA2- MEMSEL- DX0- thru DX11-
MEMDIS-

6.1.3 Read Only Memory Board

The read only memory (ROM) board is intended for use as a memory unit which can be used to store executable programs and data. A small program executing out of either the boot ROM or normal read/write memory can be used to transfer information from the ROM board into normal read/write memory. Depending on the type of ROM used, a single ROM board can contain a maximum of either 16K or 32K words.

The ROM memory, once activated, overlays normal memory address space and is addressed in the same manner as normal memory. Whenever an indirect fetch is attempted from a

memory area overlayed by an active ROM card, the ROM card suppresses the the response from normal memory and instead supplies the contents of the corresponding address in its memory. When a write is made to a memory area associated with an active ROM board, the ROM board takes no action allowing the value to be written into normal read/write memory.

Up to 8 ROM boards can be associated with a given device address. Only one of these boards can be active at a given time.

Device address 47 was selected for use by the ROM board.

Instruction Set

RMSL 6470 Skip if there is an active ROM board which overlays memory fields 0 thru 3.

RMSH 6471 Skip if there is an active ROM board which overlays memory fields 4 thru 7.

RMDI 6472 Disable all ROM boards using this device address.

RMEN 6473 Enable all ROM boards associated with this device address.

RMAD 6474 Enable the ROM board associated with this device address as selected by bits 8-11 of the AC. Each ROM board has a select address associated with it which is set on the board via a switch setting or jumper. This IOT loads a three bit select address register from the low three bits of the AC. This address is decoded and compared against the select address set on the board. Only the board set to the proper address will respond to read requests. The select address register on all boards is zeroed when power is first applied and whenever the INIT- bus signal is generated.

6475 Not used.
6476 Not used.
6477 Not used.

If 16K ROMs (type 2716) are used on the ROM board then, when active, the board overlays 16K words of normal address space. Switches located on the board determine whether the low 16K words (addresses 00000 (8)-37777 (8)) or the high 16K words (addresses 40000 (8)-77777 (8)) are overlayed. If 32K ROMs (type 2732) are used on the board, then the entire 32K words of normal address space are overlayed. The "super" bank address bits on the bus are not decoded. The board select address register was included in order to allow addressing beyond 32K words. Presently all boards have been built up using 16K ROMs.

The ROMs are programmed using the CEL EPROM programmer designed and built by H. Eshragian [47]. L. Connolly's DMROM program [48] is used to interface between the PDS operating system [49] and the programmer. Three 2716 errasable programmable read only memories are used per 4K memory field. Fig. 6.1 shows the positions of these ROMs on the board. Each row of three overlays a 4K memory field. The lowest address field is physically the row of ROMs closest to the connector. The closer a row of ROMs is to the top of the board, the higher the address of the corresponding memory field. In a given row the left most ROM contains the most significant 8 bits of the contents of even address memory locations. The middle ROM contains the most significant 8 bits of the contents of odd address memory locations. The right most ROM contains the least signficant 4 bits of the contents of the even address memory locations and the least significant 4 bits of the contents of the odd address memory locations. These ROMs are labeled E, O and M going from left to right.

Fig. 6.1 also shows the positions of three switch groups and one jumper which are used to configure the ROM

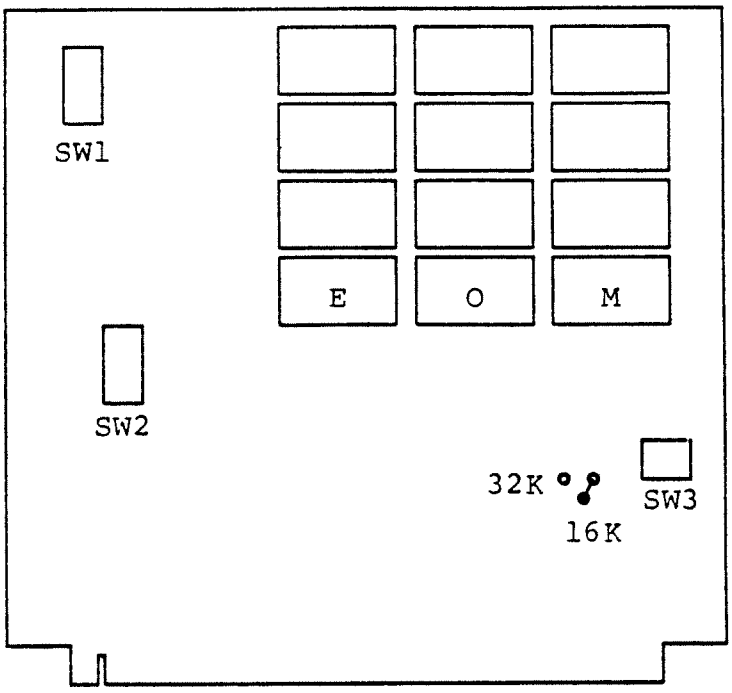


Fig. 6.1. ROM board memory and switch locations

board.

Switch group SW1 consists of 8 switches or jumpers. Positions 1 thru 6 are used to set the six bit device address that the ROM board responds to. Switch position 1 corresponds to the most significant address bit and switch position 6 corresponds to the least significant address bit. A closed switch corresponds to a one in the corresponding bit of the device address. Switch position 7 is not used. Switch position 8 must be closed if the ROM board is to overlay memory fields 0 thru 3 and open if it is to overlay fields 4 thru 7.

Switch group SW2 is used to set the select address that the board responds to. Up to 8 ROM boards can be selected using a given device address and the select register. Only one of the 8 switches should be closed at any given time. Switch position 1 corresponds to select address 0. Select addresses increase directly with switch position. Switch

position 8 corresponds to select address 7.

Switch group SW3 consists of 4 switches. Switch positon 1 is not used. If the board uses 16K ROMs, the position two switch should be closed if the board overlays memory fields 0 thru 3 and the position three switch should be closed if the board overlays memory fields 4 thru 7. If the board uses 32K ROMs, both the position two and three switches should be closed. Switch position 4 should be open if 32K ROMs are used and closed if 16K ROMs are used.

The jumper located near the bottom right corner of the board must be inserted as shown in Fig. 6.1, depending on whether 16K or 32K ROMs are used.

Example

The following is a small program intended for use in the bootstrap ROM. This program assumes the existence of a ROM board and loads the contents of field 0 from the ROM.

```
        *7766
        /
A,      RMEN            /enable the ROM
        DCA 0           /set up addr pointer and loop ctr
        TAD I 0         /get word from ROM
        DCA I 0         /place into r/w memory
        ISZ 0           /adv address and see if done
        JMP .-3         /not done yet
        RMDI            /disable ROM
        JMP I .+1       /start normal execution
        0200            /at location 0200
        JMP A           /boot starts at loc 7777
        $$$$
```

When the accessed address is in the range of 7740 thru 7777, both the boot ROM and the ROM board respond. The contents of the BOOT memory will be copied into normal memory if the ROM memory contains zeros in these locations. Also, the contents of location 0 in the ROM memory is assumed to contain a zero.

The processor used in the beamformer uses a version of the above bootstrap which reads two memory fields out of ROM into read/write memory.

The bus signals used by this board are:

DATAF- MEMSEL- XTC+ EMA0- EMA1- EMA2-
DEVSEL- INIT- LXMAR- SKPINT- DX0- thru DX11-
MEMDIS-

6.1.4 Serial Communication Interface

The serial communication interface supports full duplex bit serial information transfers between the processor and external equipment such as printers and terminals. The instruction set is upward compatible with that used on the PDP-8/e. The features added include the ability to change baud rates under program control and the inclusion of the status register. The interface is contained on a single printed circuit board which plugs into the processor bus. Connection is made to external equipment via a 16 pin cable connector located on the top right corner of the board. Input and output signals are RS-232 compatible. Limited capability for handshake interaction with an external device via RS-232 control lines is provided. The default baud rate is jumper set. The baud rate can be switched between the set value and 9600 baud using a control line on the cable connector. If the control pin is left open, the default baud rate is used. Typically, this is 300 baud. If the control pin is pulled high, the 9600 baud rate is used. Most of CEL's CRT terminals have been wired so that the act of connecting them to the interface switches the rate to 9600 baud. Other baud rates can be selected under program control. Two device addresses are used; one for receiving and one for transmitting. The device addresses are jumper selectable. Device addresses 03 and 04 are always used for standard terminal input and output (in that order). If additional interfaces are present, the address pairs are normally:

66 (output) 67 (input) first additional interface
64 (output) 65 (input) second additional interface

62 (output) 63 (input) third additional interface
60 (output) 61 (input) fourth additional interface

Receiver Instruction Set

The mnemonics given below are those normally associated with the receiver instruction set when the interface is used to connect the operator's console to the processor.

KCF 6030 The received character flag is cleared.

KSF 6031 If the received character flag is not set, this instruction acts as if it were a NOP. If the flag is set, the immediately following memory location is skipped over.

KCC 6032 The received character flag is cleared. The contents of the accumulator are set to 0.

KST 6033 The interface status word replaces the contents of the accumulator. The status word is described below.

KRS 6034 The contents of the received character buffer are OR'ed into the accumulator.

KIE 6035 The contents of the accumulator bits 4 thru 11 are loaded into the interface control register. This register makes up the lower 8 bits of the interface status word described below. The contents of the accumulator are not altered by this instruction.

KRB 6036 The contents of the received character buffer replace the contents of the accumulator. Assuming an eight bit character, the character is placed into bits 4 thru 11 of the accumulator with bits 0 thru 3 of the accumulator being set to 0. Bits 0 thru 3 can instead be selected under program control to represent the error status associated with the received character. This instruction also clears the received character flag.

KTR 6037 If the received character flag is not set, this instruction acts as if it were a NOP. If the flag is set, the contents of the received character buffer replace the contents of the accumulator, the immediately following memory location is skipped over and the flag is cleared.

The status word bits correspond to the state of four EIA compatible input lines and the contents of the interface control register. Bit assignments and interpretation follow:

Bit 0 EIA control input. Cable pin 3. Anticipated use is with a data set as carrier detect, EIA pin 8. A high positive EIA signal yields a "one".

Bit 1 EIA control input. Cable pin 4. This signal is used to switch between the jumper selected default baud rate and 9600 baud. If this pin is left open, the default rate is selected. If this pin is pulled to +5 volts or higher, the rate of 9600 baud is selected. Selecting a baud rate under program control overrides this control action. EIA connector pin 22 is used for this function.

Bit 2 EIA serial input data. Cable pin 1. The serial input data stream. With a terminal, this would connect to EIA pin 2. A negative or zero EIA signal (MARK) yields a "one".

Bit 3 EIA serial output data. Cable pin 5. The serial output stream generated by the interface as monitored at the input of the EIA driver. With a terminal, this would connect to EIA pin 3. A "one" corresponds to a negative (MARK) EIA signal.

The remaining 8 status bits correspond to the bits in the interface control register. Bits 8 thru 10 of this register are set to "zero" and bit 11 is set to "one" whenever the system is reset. This can happen when power is applied, when booting or when a CAF instruction is executed. The interface control register is loaded via the KIE instruction.

Bit 4 Baud rate select bit. If this bit is a "zero", the jumper set default baud rate is used. If a "one", the baud rate is determined by control register bits 8 thru 10.

Bit 5 Error status enable bit. This bit determines whether bits 0 thru 3 of the accumulator are to contain the error status bits when a received character is read into the accumulator. A "zero" in this bit inhibits the reading of the error bits. A "one" in this bit enables the reading of the error bits.

Bit 6 Drives an EIA output level. Corresponds to cable pin 8. Anticipated use is with a data set as the data terminal ready signal, EIA pin 8. A "one" in this bit generates an EIA high level.

Bit 7 Drives an EIA output level. Corresponds to cable pin 7. Anticipated use is with a data set as the request to send signal, EIA pin 4. A "one" in this bit generates an EIA high level.

Bits 8-10 Program control baud rate selector bits. Two sets of baud rates are available depending on the presence or absence of a jumper on the interface board. The primary set of baud rates available (no jumper) and the associated values of bits 8 thru 10 is:

000	9600 baud	100	2400 baud
001	4800 baud	101	300 baud
010	1800 baud	110	150 baud

011	1200 baud	111	110 baud

The alternate set of baud rates (jumper in place) and the associated values of bits 8 thru 10 is:

000	---	100	134.5 baud
001	---	101	200 baud
010	50 baud	110	600 baud
011	75 baud	111	2400 baud

Bit 11 Interrupt enable bit. This bit is set to the "one" state whenever the processor is initialized. When this bit is a "one" and the processor interrupt system is enabled, the setting of either the receiver or transmitter flag will cause a program interrupt. If this bit is a "zero", the setting of either flag will not cause a program interrupt.

The interpretation of the four bits of received character error status which can be placed in accumulator bits 0 thru 3 when a character is read is as follows:

Bit 0 "One" if any of the error bits 1 thru 3 are set. "Zero" otherwise.

Bit 1 Parity error. Active only if the parity option is jumper selected. A "one" indicates that the received character did not have the correct parity.

Bit 2 Framing error. A "one" indicates that an invalid stop bit was contained in the received character frame.

Bit 3 Overrun error. A "one" indicates that a character was received without the previously received character being read. The earlier character is lost.

Transmitter Instruction Set

The mnemonics given below are those normally associated with the transmitter instruction set when the interface is used to connect the operator's console to the processor.

TFL 6040 The transmitter ready flag is set indicating readiness to accept a character for transmission.

TSF 6041 If the transmitter ready flag is not set, this instruction behaves as a NOP. If the flag is set, the immediately following memory location is skipped over.

TCF 6042 The transmitter read flag is cleared.

6043 Not used. Essentially a NOP.

TPC 6044 The contents of the accumulator bits 4 thru 11 are loaded into the transmitter character buffer. The transmitter then converts the 8 bit character into a serial bit stream. The contents of the accumulator are not altered.

TSK 6045 If neither the receiver nor the transmitter flag is set, this instruction behaves as a NOP. If either flag is set and the interrupt enable bit in the interface control register is set, the next memory location is skipped over.

TLS 6046 Combines the operations of the TCF and TPC instructions. The character contained in the accumulator is loaded into the transmitter buffer and the transmitter ready flag is cleared. The contents of the accumulator are not altered.

TTR 6047 If the transmitter ready flag is not set, this instruction behaves as a NOP. If the flag is set, the contents of the accumulator bits 4 thru 11 are sent to the transmitter character buffer for serial transmission, the next memory

location is skipped and the transmitter ready flag is cleared. The contents of the accumulator are not altered.

Jumper Options

The transmitter and receiver device address as well as various options can be selected using jumpers located on the interface board. There are three jumper sockets present. Two of these are located at the top left of the board and the third is located at the bottom right.

Top most left jumper socket

1	Transmitter device address bit	3 (wire="one")
2		4
3		5
4		6
5		7
6		8
7	Receiver device address bit	3 (wire="one")
8		4

Lower top most left jumper socket

1	Receiver device address bit	5
2		6
3		7
4		8
5	Default baud rate select MSB	(wire="zero")
6	Default baud rate select	(wire="zero")
7	Default baud rate select LSB	(wire="zero")
8	Select alternate baud rate set	(wire=alternate)

Lower right side jumper socket

1	If parity enabled, wire=odd parity, otherwise even
2	Character length MSB (wire="zero")
3	Character length LSB (wire="zero")
4	Number of stop bits
5	Wire present enables parity generation
6	Wire disables boot on break detect
7	Wire disables control panel interrupt on break
8	Not used

The character lengths selected by jumpers 2 and 3 do not include the parity bit if one is generated (wire in

position 5). A value of 00=5 bits, 10=6 bits, 01=7 bits and 11=8 bits.

No wire in the stop bit position gives 1.5 stop bits for 5 bit long characters and two for all others. A wire present yields 1 stop bit.

A normally configured header has only jumpers 6 and 7 installed.

The pin assignments on the cable connector are:

Pin 1 Serial EIA signal to interface (RS-232 pin 2 if connected to a terminal).

Pin 2 EIA signal to interface used to allow or inhibit output from the interface. Intended for use a a hardware handshake signal with high speed hard copy devices. If left open, has no effect. If driven to +5, inhibits serial output by preventing transmitter flag from setting. Returning level low again allows normal operation.

Pin 3 EIA signal to interface assigned for use in monitoring the carrier detect output of a modem (RS-232 pin 8).

Pin 4 EIA signal to interface used to set baud rate externally. Open or negative input allows jumper default rate to be used. Positive input forces use of 9600 baud.

Pin 5 EIA serial output from interface (RS-232 pin 3 if connected to a terminal).

Pin 6 Clock generated by interface at 16 time the selected baud rate. Uses RS-232 levels.

Pin 7 EIA control signal generated by interface. Normally used to generate RTS (RS-232 pin 4).

Pin 8 EIA control signal generated by interface. Normally used to generate DTR (RS-232 pin 8).

Pins 9 - 16 Ground.

The bus lines used by this interface are:

CPREQ-	SKPINT-	BOOT-	IODIS-	C0-	C1-
XTC+	INIT-	LXMAR-	DEVSEL-	DX0- thru	DX11-

6.1.5 Fast Arithmetic Device

The Fast Arithmetic Device (FAD) is a hardware addition to the DM1200 designed to speed up the double precision arithmetic operations commonly used in CEL written signal processing software. The FAD is contained on a single wire-wrap board and plugs directly into the processor bus.

The operations supported by the FAD include 24 bit complex addition and subtraction as well as instructions for a variety of multiply and multiply/sum operations. Using the FAD, the time required to multiply two signed 24 bit values is 13 microseconds (not including set-up time) compared to 4.5 milliseconds to do the equivalent operation in software. The FAD does not contain any provision for detecting arithmetic overflows and must be used with some caution.

The FAD is organized as two 24 bit input registers named X and Y and one 48 bit output register named P.

The X and Y registers are loaded from the accumulator using IOT instructions to individually load the most and least significant 12 bit halves. Loading the X register also causes the Y register to be loaded with the same value. Loading the Y register has no effect on the contents of the X register. It is not possible to directly read the contents of these two registers.

The 48 bit P register is used to hold the results of the various arithmetic operations. This register is read into the IM6100 as four 12 bit words named P0 thru P3 going from most significant to least significant quarters. Depending on the operation being performed, the P register is used either as two 24 bit registers or one 48 bit register. When used as two 24 bit registers, the most significant 24 bits (P0 and P1) are named Px and the least significant 24 bits (P2 and P3) are named Py. The P register cannot be directly loaded from the accumulator.

The FAD arithmetic operations are implemented using bit

serial adders and multipliers. While these are not as fast as currently available bit parallel units they are much more compact and require significantly less power. The arithmetic operations require longer than one IOT execution time in order to finish. A small delay must be provided after each of these operations before initiating a new operation or attempting to read the result. The time delay require for the 24 bit add and subtract operations is 6.5 microseconds and for the 48 bit add and the multiply operations is 13 microseconds. This time can be used by the processor to do other things.

Instruction Set

MLXM	6340	The contents of the AC replace the contents of the most significant halves of both the X and the Y registers. The contents of the P register are unchanged. This instruction clears the AC.
MLXL	6341	The contents of the AC replace the contents of the least significant halves of both the X and the Y registers. The contents of the P register are unchanged. This instruction clears the AC.
MLYM	6342	The contents of the AC replace the contents of the most significant half of the Y register. The contents of the P register are unchanged. This instruction clears the AC.
MLYL	6343	The contents of the AC replace the contents of the least significant half of the Y register. The contents of the P register are unchanged. This instruction clears the AC.
MRP0	6344	The contents of word 0 of the P register replace the contents of the AC. No other changes occur.
MRP1	6345	The contents of word 1 of the P register

replace the contents of the AC. No other changes occur.

MRP2 6346 The contents of word 2 of the P register replace the contents of the AC. No other changes occur.

MRP3 6347 The contents of word 3 of the P register replace the contents of the AC. No other changes occur.

The remaining IOT instructions do not affect the contents of the accumulator. Note the additional times beyond the initiating IOT required for the operations to be completed.

MCAD 6350 Complex add. The contents of the X register are summed with the contents of the Px register. The sum replaces the contents of both the X register and the Px register. At the same time, the contents of the Y register and the contents of the Py register are summed with the result replacing both the contents of the Y register and the Py register. This instruction requires an additional 6.5 microseconds.

MCSU 6351 Complex subtract. The contents of the X register are subtracted from the contents of the Px register. The result replaces both the contents of the X register and the Px register. At the same time, the contents of the Y register are subtracted from the contents of the Py register with the result replacing both the contents of the Y register and the Py register. This instruction requires an additional 6.5 microseconds.

MMXY 6352 Multiply x and y. The contents of the X register are multiplied by the contents of the Y register. The 48 bit result replaces the contents of the P register. The contents of

the X register are unchanged. The sign bit associated with the Y register is propagated thru the Y register. The multiply is signed. This instruction requires an additional 13 microseconds.

MX2Y2 6353 Form x squared plus y squared. The contents of P are replaced by the square of the contents of the X register added to the square of the contents of the Y register. The sign bit associated with the X register is propagated into the X register. The sign bit associated with the Y register is propagated into the Y register. This instruction requires an additional 13 microseconds.

MSXY 6354 Add the product of x and y to p. The contents of the X register are multiplied by the contents of the Y register. The resulting 48 bit product is summed with the contents of the P register. The result is is placed into the P register. The contents of the X register are unchanged. The sign bit associated with the Y register is propagated into the Y register. The multiply is signed. This instruction requires an additional 13 microseconds

MSX2Y2 6355 Sum x*x plus y*y into p. The contents of the X and the Y registers are squared and summed with the contents of the P register. The result is placed into the P register. The sign bit associated with the X register is propagated into the X register. The sign bit associated with the Y register is propagated into the Y register. This instruction requires an additional 13 microseconds.

M4PS 6356 Four precision sum. The contents of the X and Y register are treated as if they formed a 48 bit value with the X register holding the most

significant 24 bits and the Y register holding the least significant 24 bits. This value is summed with the 48 bit contents of the P register with the sum being placed into the P register. The most significant 24 bits of the sum also replaces the contents of the X register. The least significant 24 bits of the sum also replaces the contents of the Y register. This instruction requires an addtional 13 microseconds.

MCLP 6357 Clear P. The contents of the P register are set equal to zero. The contents of the X register and the Y register are not affected.

The bus signals used by this board are:

C0- C1- XTC+ INIT- LXMAR- DEVSEL- 4MHZ- DX0- thru DX11-

6.1.6 Single Channel Complex Demodulator

The single channel complex demodulator (SCCD) provides the DM1200 with a single channel A/D converter coupled with a direct digital "$4f_0$" complex demodulator. The inputs to the SCCD consist of a clock at a frequency equal to four times the desired demodulation frequency and an analog waveform to be demodulated. The analog waveform is sampled at the clock rate, weighted by plus and minus ones and summed into the proper X and Y sums as needed to form a complex valued demodulate. Demodulate values are made up of two 24 bit integers. The real part of the demodulate is named X and the imaginary part Y. After the proper number of samples have been summed, the result is transferred to a pair of X and Y output registers, the accumulation registers are zeroed and a new demodulate is started. Once a demodulate has been loaded into the output registers the data ready flag is set. The SCCD can be configured so that the setting of this flag causes a program interrupt in the

processor. The processor has almost the entire time required to form a new demodulate in which to read the contents of the output registers.

The A/D converter used in the SCCD is a 12 bit converter and operates over a minus 5 to plus 5 volt range. The input is ac coupled with an input impedance of 100 Kohms and a low frequency cutoff of 1 Hz. Successive samples over one demodulation cycle are assigned in the order +X, -Y, -X, +Y. Two samples are taken per half cycle of the demodulation frequency. The number of half cycles to be used to form a demodulate is set into the SCCD by the processor as a negative integer value. If a value of -1 is used and if the sign ordering is taken into account, the SCCD can be used as a non-demodulating A/D converter.

The clock input is also ac coupled and must be symmetric about ground. The clock input impedance is 100 Kohms and has a low frequency cutoff of 1 Hz. The minimum input level is 1 volt peak-to-peak and the maximum is 5 volts peak-to-peak.

The SCCD is contained on a single wire-wrap panel onto which a small printed circuit board has been mounted. The printed circuit board contains the A/D converter and the input clock circuit. This method of construction effectively isolates the analog circuit from the digital. The SCCD can be used in any DM1200 system where there is space available.

External connections to the SCCD are made using a 16 pin cable connector located on the analog printed circuit card. Pin assignments are:

Pin 1 Analog input
Pin 2 Analog input ground
Pin 7 Clock input
Pin 8 Clock input ground
Pins 9-16 Ground

There are also three test points and three adjustment pots located on the analog card. Starting at the cable

connector and working to the far end of the card, the test points are:

- Input isolation amplifier output
- Ground
- A/D converter input (track and hold output)

The adjustments starting at the cable connector and working toward the far end of the card are:

- Track and hold offset
- A/D converter offset trim
- A/D converter gain trim

Instruction Set

QDSF 6421 If the data ready flag is set, the immediately following instruction is skipped. If the flag is not set, the instruction is not skipped. If a skip occurs, the data ready flag is cleared.

QDCF 6422 Clears the data ready flag. Bit 11 of the accumulator is also loaded into the interrupt enable register and then the accumulator is cleared. If the interrupt enable register is loaded with a one, the setting of the data ready flag can cause a processor interrupt. If the interrupt enable register is loaded with a zero, the setting of the data ready flag cannot cause an interrupt request. The contents of the interrupt enable register are set to zero whenever the INIT- signal is generated.

QDLH 6423 The contents of the accumulator are loaded into the half cycle counter. The value loaded is equal to minus the number of demodulation frequency half cycles that the demodulates are to be formed over. The accumulator is then cleared.

QDXM 6424 The contents of the most significant 12 bits of the demodulator X output register are loaded

into the accumulator.

QDXL 6425 The contents of the least significant 12 bits of the demodulator X output register are loaded into the accumulator.

QDYM 6426 The contents of the most significant 12 bits of the demodulator Y output register are loaded into the accumulator.

QDYL 6427 The contents of the least significant 12 bits of the demodulator Y output register are loaded into the accumulator.

The bus signals used by this device are:

SKPINT- C0- C1- XTC+ INIT- LXMAR-
DEVSEL- 4MHZ- DX0- thru DX11-

6.1.7 Beamformer Complex Demodulator

The beamformer complex demodulator (BFCD) takes the beamformer output values and shifts them to base band forming complex valued demodulates in the process. The demodulation frequency can be set at any value from 0.001 Hz to 999.999 Hz in 0.001 Hz steps. The integration interval can vary from 0.2 milliseconds to 0.8090 seconds in 0.2 millisecond steps.

The BFCD consists of one wire-wrapped logic panel with an outrigger printed circuit card mounted on it and also from one to eight demodulator printed circuit boards. The wire-wrap panel contains the interface to the DM1200, the timing control logic and the synthesizer used to generate sample values of $\exp\{-j2\pi f_d t\}$. The outrigger printed circuit board holds the ROMs containing the complex exponential samples and generates a bus which is used to provide the demodulator boards with the complex exponential sample values and control signals. Each demodulator board contains the multipliers and accumulators necessary to form demodulates for one beam.

The demodulator boards are connected to the control

board using top edge connectors. This bus uses a 40 wire cable to communicate the complex exponential sample values and a 26 wire cable to communicate control signals. Information is transferred between the demodulator boards and the associated beams via three twisted pairs per demodulator. An additional three twisted pairs are used to transfer timing and control information between the beamformer and the control logic.

Demodulates are formed by multiplying the beamformer output values by sample values of $\exp\{-j2\pi f_d t\}$ and summing. The basic time step between beamformer output values is 0.2 milliseconds (the beamformer forms outputs at a 5 KHz rate). The complex exponential values are amplitude quantized to 16 bits and time quantized to 2000 values. These sample values are stored in ROMs located on the small printed circuit board attached to the wire-wrap panel. No attempt was made at trying to compact the ROM contents. One ROM pair was dedicated to holding one period of the cosine and another pair was dedicated to holding one period of the sine. This resulted in a much simpler device than would have been possible if an attempt had been made to exploit symmetries. The required ROM addresses are generated using a direct digital frequency synthesizer of the form decribed in [12].

The demodulation frequency is set using a 6 digit BCD value. This is loaded into the BFCD using two 12 bit words. The most significant word contains the three BCD digits to the left of the decimal point (integer part) and the least significant word contains the three BCD digits to the right of the decimal point (fractional part).

The number of beamformer output values used to form a demodulate is set by the processor. Using an IOT instruction, this number is loaded into the control logic as a negative integer. The demodulation summations start on the first 5 KHz clock tic that occurs at or just after either the 0 or 180 degree transition of the demodulating complex exponential.

Beamformer output values are a maximum of 20 bits in length. The demodulator sine and cosine values use 16 bits and are normalized so that the binary point lies to the right of bit 0 (the sign bit). The product thus contains a 15 bit fraction which is retained and included in the demodulate formation. Although this appears to be a brute force approach to avoiding scaling problems, there is almost no time or hardware penalty for retaining these bits because of the use of serial multipliers, adders and accumulators. Each of the accumulator registers used to form the demodulates contains 48 bits.

Once a complete demodulate has been formed, it is transferred into a set of output registers. In this transfer, the value is reduced to integer form and scaled down in magnitude by from 0 to 15 bits. This scaling is done carefully in order to avoid introducting any dc bias into the result. The amount scaled is preset by the processor.

The data ready flag is set once the output registers have been loaded with a new demodulate value. The processor has almost an entire demodulate time in which to read the contents of these registers.

Demodulator boards are addressed thru the control logic by loading a board address register and then using "generic" read register IOT instructions. The address that a given board responds to is set using a jumper located near the center of that board. Demodulator board addresses range from 0 thru 7.

There are two 24 bit output registers on each demodulator board. The X register contains the real part of the demodulate and the Y register contains the imaginary part. The contents of these registers are read in 12 bit units.

It is possible to freeze the integration counter and hold it in its initial state. Using this capability along with an externally supplied timing signal, it is possible

for the processor to restart the integration process after an interruption with a timing uncertainty of 0.2 milliseconds or less. Although the integration counter is frozen, the formation of demodulate values continues. Thus, after un-freezing the counter, the first set of demodulate values should be discarded.

Instruction Set

BDSF 6441 Skip the next instruction if the data ready flag is set. If the flag is not set, do not skip the next instruction. If a skip occurs, the flag is then cleared. This instruction always clears the accumulator.

BDCF 6442 Clear the data ready flag. Also loads the contents of accumulator bit 11 into the interrupt enable register and then clears the accumulator. If the interrupt enable register contains a one, the setting of the data ready flag can cause a program interrupt in the processor. If the interrupt enable register contains a zero, the setting of the data ready flag cannot cause a program interrupt. The interrupt enable register is loaded with a zero when ever the INIT- signal is generated.

BDFM 6443 The most significant half of the frequency control register is loaded from the accumulator. The accumulator is then cleared. The value loaded is the integer part of the desired frequency and is represented using three BCD digits.

BDFL 6444 The least significant half of the frequency control register is loaded from the accumulator. The accumulator is then cleared. The value loaded is the fractional part of the desired frequency represented using three BCD digits.

BDLT 6445 The contents of the accumulator are loaded into the integration count register. The accumulator is then cleared. The value in the accumulator should be minus the number of beamformer output values that are to be used in forming a demodulate. Beamformer output values are produced at a 5 KHz rate.

BDLS 6446 The value contained in accumulator bits 8 thru 11 is loaded into the scale register. The accumulator is then cleared. The integerized demodulate values are reduced in magnitude by arithmetically scaling them right by the number of bit positions indicated by the contents of the scale register. The demodulate values can be scaled by from 0 to 15 bit positions.

BDLC 6447 The demodulator address register is loaded from bits 9 thru 11 of the accumulator. The accumulator is then cleared. Only the demodulator board whose jumper selected address corresponds to the loaded address will respond to the read IOT instructions listed below. If the BDLC is executed with a one in accumulator bit 0, the integration count is stopped and the counter set to its starting state and held there. Executing the BDLC with a zero in accumulator bit 0 allows the integration to proceed. The first demodulate formed after un-freezing the integration count should be discarded. The integration counter is un-frozen whenever the bus INIT- signal is generated.

BDXM 6454 The most significant 12 bits of the selected demodulator board X (real part) output register are loaded into the accumulator.

BDXL 6455 The least significant 12 bits of the selected demodulator board X output register are loaded into the accumulator.

BDYM 6456 The most significant 12 bits of the selected demodulator board Y (imaginary part) output register are loaded into the accumulator.

BDYL 6457 The least significant 12 bits of the selected demodulator board Y output register are loaded into the accumulator.

Bus signals used by this board and the associated demodulator boards are:

SKPINT- C0- C1- XTC+ INIT- LXMAR-
DEVSEL- 4MHZ- DX0- thru DX11-

6.1.8 DMA LINCtape System

The LINCtape system designed for the DM1200 reads and writes LINCtapes compatible with CEL's other systems. However, because of the slow speed of the processor, it was necessary to make this system much more automatic in operation than its predecessors. A transfer of data between tape and memory is set up by loading the number of tape blocks to be transferred (the block size is determined by the tape format), the starting block number on tape and the starting memory address into the interface. A single command then causes the transfer to take place. The interface electronics takes care of searching out the proper block, transferring the data using direct memory access, computing the required checksum and verifying it. The processor is notified when the transfer has been completed by the setting of the transfer done flag (TDF). The setting of this flag can be allowed to cause a processor interrupt. After a data transfer, the processor can check for errors by testing the tape error flag (TEF). The time overhead for the DMA data transfers is less than 2% versus the approximately 50% on CEL's older system. It is anticipated that the CEL PDP-8/e LINCtape systems will be upgraded to be compatible with the new electronics.

Device handlers have been written for the PDS programming system and for the CEL real-time monitor BAMMS.

Because the DM1200 lacks a data break multiplexer, the LINCtape system cannot be used simultaneously with the 9-track magnetic tape system.

The DM1200 LINCtape electronics are contained on two wire-wrap panels and one printed circuit board and are separate from the drive electronics. One wire-wrap panel plugs into the processor bus and contains the processor interface. This board generates a data transfer bus which runs to the tape drive. This bus uses two cables, each containing 13 twisted pairs. The data transfer bus connects to the wire-wrapped formatter board which is mounted on the tape drive. Associated with this board is a printed circuit card used to generate various timing signals and to decode the tape format.

Four of the tape drives presently in use on CEL's DM1200 systems are re-cycled dual drives removed from discarded LINC-8 computer systems. One drive is a slightly modified Computer Operations CO-600 dual drive LINCtape and one is a Computer Operations CO-3000 single drive LINCtape. The drives removed from the LINC-8's have had their original electronics replaced using control boards purchased from Computer Operations. The use of these boards made the modified drives compatible with TTL logic levels. The Computer Operations motor control boards were modified so that motion commands tension the tape. The modification to the CO-600 drive consisted of inverting the polarity of the REDY line. The CO-600 does not automatically tension tapes on receipt of a motion command. The CEL formatter board used on the CO-3000 had to be modified in order to accomodate the CO-3000 electronics.

The interaction between the processor and the tape system takes place using four control registers and five flags. The four control registers are:

Block address register (BAR)

This register is loaded with the block number on tape at which the data transfer is to start. This register cannot be read. The contents of the BAR are undefined after a transfer has been made.

Memory address register (MAR)

This register is loaded with the lower 12 bits of the address in memory at which the data transfer is to start. This register cannot be read. The contents of the MAR are undefined after a transfer has been completed.

Block count and field register (BFR)

The low 8 bits are used to hold the number of tape blocks that are to be transferred. From 0 thru 255 blocks can be read or written in one operation. Memory addresses automatically cross field boundaries. The most significant three bits of this register select the memory field in which the transfer is to start. No provision has been made to use the three extended address bits on the processor bus. There was not enough room on the interface board to provide for them. The field bits are incremented whenever the low 12 address bits index from 4095 to 0. The memory field address wraps around from 7 to 0. This register cannot be read. The contents of the BFR are undefined after a transfer has been completed.

Control register (COR)

The bits of the control register define the operation to be performed when the transfer is initiated by a skip using the DLSR IOT instruction. The contents of this register cannot be read. Except for bit 0, the contents of this register are not altered during a data transfer. The contents of this register are set to zero by the DLCL IOT instruction, by the processor stopping and by the bus INIT- signal. Bit assignments are:

Bit 0. If a zero, the tape motion will be in the forward direction when the selected tape drive goes into motion. If a one, the starting direction will be in the backward direction. This bit is the actual direction bit used by the interface to control tape motion. At the end of a successful transfer it is normally a zero.

Bit 1. Referred to as the write bit. If a zero, the direction of the data transfer will be from tape into memory. If a one, the transfer direction will be from memory to tape (the addressed drive must be write enabled).

Bit 2. Referred to as the mark bit. This bit is used to configure the interface and formatter for use in formatting (marking) tapes and in certifying them. If both the mark and write bits are ones then when motion is started the timing, mark and normal data writers are turned on. Timing is generated using a crystal clock. Using the WDF along with the DLWM and DLWD IOTs described below, a tape can be formatted. If the mark bit is a one and the write bit is a zero, the interface and formatter are configured for certifying the format previously recorded on a tape. The WDF along with the DLRM and DLRD IOT instructions are used for this purpose. Whenever this bit is a one no DMA transfers are made between the tape system and memory.

Bit 3. If a zero, the tape drive is stopped after the data transfer has been made. If a one, the drive is left in motion.

Bit 4. If a zero, the transfer is made normally. If a one, the transfer to/from memory is inhibited but the remainder of the operation proceeds normally. This

feature is intended for use in performing a read without transfer just after writing. If there is an error on tape, the values in memory are unaffected and additional attempts can be made to rewrite them.

Bit 5. If a zero, the setting of the block transfer flag cannot cause a processor interrupt. If a one, the setting of this flag can cause an interrupt if the processor interrupt system is enabled.

Bit 6. If a zero, the setting of the data transfer flag cannot cause a processor interrupt. If a one, the setting of this flag can cause an interrupt if the processor interrupt system is enabled.

Bit 7. If a zero, the current CEL checksum convention is used. If a one, the old (pre 1972) checksum convention is used. A jumper on the formatter board must be in place for this feature to work. This capability was included to allow tape interchange with the Institute for Acoustic Research (IAR) which used the old convention on their General Automation systems. IAR has effectively been disbanded and it is not expected that this feature will be required.

Bit 8. Not used.

Bits 9 thru 11. Selects the drive to be used. Drive addresses run from 0 thru 7. Normally, drive 0 is present in all systems using LINCtape.

The four flags used to coordinate the processor activity and the tape system activity are:

Block transfer flag (BTF)

This flag is set whenever the transfer of a single tape block's worth of information has been completed. This flag can be used to synchronize the testing of the error flag on a block by block basis so that if there is an error, the bad block can be identified. This flag is cleared whenever a skip occurs using the DLSB IOT instruction; by a skip on the ready condition; by the DLCL IOT; when the processor halts and by the bus INIT- signal.

Transfer done flag (TDF)

This flag is set whenever the requested number of blocks have been transferred. This flag is cleared whenever a skip occurs using the DLST IOT instruction; by a skip on the ready condition; by the DLCL IOT; when the processor halts and by the bus iNIT- signal.

Error flag (ERF)

The error flag is set if an error is detected in the value of the checksum. This flag is also set whenever the formatter detects an error in the control information read off of the mark track. This makes it possible to detect certain types of errors during tape writes. If on, the writers are turned off whenever this flag becomes set. In a multiple block transfer this flag is not automatically cleared prior to starting to transfer a new block. This flag is cleared by a skip using the DLSE IOT instruction; a skip on the ready condition; the DLCL IOT; when the processor halts and by the bus INIT- signal.

Word done flag (WDF)

During normal data transfers, this flag is used by the interface to request data breaks. When formatting (marking) or certifying LINCtapes, this flag does not cause data breaks but is used instead to synchronize data trasfers between the processor and the tape. When marking a tape (both the mark and the write bits in the

control register are ones), this flag is set whenever the formatter is ready to accept another mark frame and the associated data word. These are loaded using the DLWM and DLWD IOT instructions defined below. When certifying a tape (the mark bit is a one and the write bit is a zero), this flag is set whenever the formatter has assembled a mark frame and associated data word off of tape. The processor reads these using the DLRM and DLRD IOT instructions defined below. Because these programmed data transfers are doubled buffered, once this flag has been set, the processor has approximately 160 microseconds in which to do any required processing. This flag is cleared whenever a skip occurs using the DLSW IOT instruction; by a skip on the ready condition; by a skip on the error flag; by the DLCL IOT; whenever the processor halts; whenever a data break transfer is made and by the bus INIT- signal.

Instruction Set

DLCL 6160 Clears the tape system. This consists of clearing all flags, setting the contents of the control register to zero, stopping motion, turning off the writers and setting the BFR and the MAR contents to zero.

DLSR 6161 Tests the drive status to see if it is compatible with the operation selected by the contents of the control register. Namely, the selected drive must be on-line and, if writing, the drive must be write enabled. If these conditions are met, the immediately following instruction is skipped over and the transfer operation is initiated.

DLWC 6162 The contents of the accumulator are loaded into the control register. The interpretation of the

various bits is described above. The accumulator is cleared by this instruction.

DLWA 6163 The contents of the accumulator are loaded into the memory address register. This is the low 12 bits of the memory address at which the transfer is to start at. The accumulator is then cleared.

DLWB 6164 The contents of the accumulator are loaded into the block address register. The accumulator is then cleared. The block address register is used to hold the starting tape block number at which the transfer is to start.

DLWF 6165 The low 8 bits of the accumulator are loaded into the block counter and the top 3 bits are loaded into the memory address field register. The accumulator is then cleared.

DLST 6166 If the data transfer flag is set, the immediately following instruction is skipped and the data transfer flag is cleared. If the DTF is not set, the immediately following instruction is not skipped.

DLSB 6167 If the block transfer flag is set, the immediately following instruction is skipped and the BTF is cleared. If the BTF is not set, the immediately following instruction is not skipped.

DLSE 6170 If the the data error flag is set, the immediately following instruction is skipped and the DEF is cleared. If the DEF is not set, the immediately following instruction is not skipped.

DLSW 6171 If the word done flag is set, the immediately following instruction is skipped and the WDF is cleared. If the WDF is not set, the immediately following instruction is not skipped.

DLWD 6172 The contents of the accumulator are loaded into

the the tape formatter buffer register. The accumulator is then cleared. This instruction is used when formatting tapes.

DLWM 6173 The contents of the accumulator bits 8 thru 11 are loaded into the formatter mark track buffer register. The accumulator is then cleared. This instruction is used when formatting tapes.

DLRD 6174 The contents of the formatter buffer register are loaded into accumulator. This instruction is used when certifying the tape format.

DLRM 6175 The contents of the mark track register are loaded into accumulator bits 8 thru 11. This instruction is used when certifying the tape format.

Notes:

When searching in the reverse direction, a three block offset is used. The direction of motion is not changed back to the forward direction until the tape has moved three blocks beyond the desired block. This is done in order to allow the tape motion to stabilize after reversing direction before reading or writing a block or set of blocks. This overshoot makes the DM1200 LINCtape system appear to run slower than CEL's other LINCtape systems. They do, slightly. It is felt that this feature adds to the system reliability.

The DMA data transfers are double buffered. Approximately 160 microseconds are available for the processor to repond to a break request. This should allow the LINCtape system to be given the lowest break priority when a data break multiplexer is added.

The bus signals used by the LINCtape interface are:

DMAREQ-	SKPINT-	EMA2-	EMA1-	EMA0-	C0-
C1-	XTC+	RUN-	DMAGNT-	MEMSEL-	INIT-
LXMAR-	DEVSEL-	4MHZ-	DX0- thru DX11-		

6.1.9 9 Track Magnetic Tape System

The DM1200 9 track magnetic tape system consists of from 1 thru 4 Cipher dual mode 800/1600 bpi 9 track tape drives, a formatter and a controller/interface. The formatter is mounted on the master tape drive and was purchased from Cipher with the drive. The controller/interface consists of two wire-wrap panels and was designed at CEL for use with the Cipher hardware. Because of the cabling required to connect the controller/interface to the tape system, the tape system can only be used with a special wire-wrapped bus DM1200 logic bay. The system is controlled using IOT instructions in the DM1200. Data transfers are by direct memory access and occur at a rate of 72,000 bytes per second. The transfer time overhead corresponds to approximately 10% of the available CPU time. Because the DM1200 does not have a DMA multiplexer, it is not possible to simultaneously operate the 9 track tape system and the LINCtape system.

The programmer viewable portion of the controller/interface consists of:

12 bit control register used to set up the desired operation.

12 bit status register used to determine the status of the formatter and the selected drive.

18 bit memory address register used to point to where in memory the data transfer is to start.

18 bit word count register used to control the number of bytes or words to be transferred.

The status, memory address and word count registers can

be read.

Data can be transferred between memory and the 9 track tape system in units of 8 bit bytes or 12 bit words. In 8 bit mode, bytes are packed one per 12 bit word in the low order 8 bits. The top 4 bits in each word are ignored when writing and are set to zero on reading. In 12 bit mode, words are transferred as two 6 bit bytes, each written as an 8 bit byte on tape.

The relation between the bits in a twelve bit word and the data tracks on the 9 track tape are shown below for both 8 bit bytes and 12 bit modes.

Byte transfer mode

```
12 bit word: 00 01 02 03 04 05 06 07 08 09 10 11
tape tracks:             00 01 02 03 04 05 06 07
```

Word transfer mode

```
12 bit word: 00 01 02 03 04 05 06 07 08 09 10 11
tape tracks: 02 03 04 05 06 07                   1st frame
                                 02 03 04 05 06 07 2nd frame
```

Instruction Set

MCLR 6700 Reset the formatter and tape system. Clears all registers and flags. Used to initialize the formatter and to terminate "run-away" conditions.

MSTO 6701 Skips the next instruction if the formatter has completed the last commanded operation and is no longer busy. If bit 9 in the command register is set and the DM1200 interrupt system is enabled, the transition from busy to not busy will cause an interrupt. The interrupt can be cleared by the CAF, MCLR and MSTO instructions or by the skip on the MSGO instruction.

MSTT 6702 This instruction is used to test the state of the DATA BUSY line. A skip occurs if this line is not true. DATA BUSY goes true when the

selected transport has reached operating speed and is about to read or write a block of data. This signal goes false again once the block has been processed and the appropriate post record delay has been completed. This signal is not used to cause an interrupt.

MSFM 6703 Skips the next instruction if the file mark flag is set. This flag is set whenever a file mark is encountered during a non-write operation. This flag is cleared by the MCLR instruction and a skip on the MSGO instruction.

MSET 6704 Skips the next instruction if the EOT marker was sensed during the execution of the most recent command. This flag is cleared by the MCLR instruction or the skip on the MSGO instruction.

MSER 6705 Skips the next instruction if the tape error flag is set. The error flag can be set by a number of error conditions detected by the formatter. During 1600 bpi read operations, soft errors corrected by the formatter do not cause the error flag to be set. However, during writes, corrected soft errors (in the read after write) do cause the error flag to be set. This flag is cleared by the MCLR instruction and by a skip on the MSGO instruction.

MSGO 6706 Skips the next instruction if the command present in the command register can be executed. The indicated operation is initiated on the skip. The following conditions inhibit a skip:

- Illegal command
- Selected drive is off-line or rewinding
- Attempted write on a write protected drive

MSWC 6707 Skips the next instruction if the contents of the word count register are zero.

MLCR 6710 The contents of the command register are replaced by the contents of the AC. The AC is

then cleared. Bit assignments are as follows:

Bits 0-4	Command code
Bit 5	Formatter address (0 for now)
Bits 6-7	Threshold level (0 for now)
Bit 8	Not used.
Bit 9	If bit 9 is a 1 an interrupt is allowed on the setting of the operation done flag.
Bits 10-11	Drive address

The command to be executed is selected by the octal value of bits 0 thru 4. The commands are:

0	Read 8 bit bytes
1	Read 12 bit words in MP12 format
2	Read 8 bit bytes without transfer
3	Read 12 bit words in MP12 format without transfer
4	Write 8 bit bytes
5	Write 12 bit words in MP12 format
6	Write 8 bit bytes in edit mode
7	Write 12 bit words in MP12 format and edit mode
8	Edit backspace
9	Write file mark
10	Erase fixed length (3.5 inches)
11	Erase variable length
12	Space 1 block forward
13	Space 1 block reverse
14	Space 1 file forward
15	space 1 file reverse
16	Rewind
17	Place drive off-line
18-31	Not assigned

MRST 6711 The contents of the tape system status register replace the contents of the AC. The status bits are not affected. The bits are assigned as follows:

0	Formatter busy
1	Data transfer taking place
2	Hard error detected by formatter
3	Soft error detected and corrected by formatter
4	File mark flag
5	Drive ready (tensioned and not loading or rewinding)
6	Drive on-line
7	Drive write protected
8	Tape at BOT
9	Tape encountered EOT during last command

10 not used
11 not used

MLMA 6712 The contents of the AC replace the low 12 bits of the tape system memory address register. The AC is then cleared.

MLWA 6713 Bits 0-5 of the AC replace the contents of the high order six bits of the tape system word count register. Bits 6-11 of the AC replace the contents of the high order six bits of the tape system memory address register. The AC is then cleared.

MLWC 6714 The contents of the AC replace the contents of the low 12 bits of the tape system word count register. The AC is then cleared.

MRWA 6715 The contents of the high 6 bits of the word count register replace the contents of the AC bits 0-5. The contents of the high 6 bits of the memory address register replace the contents of the AC bits 6-11. The contents of the word count and the memory address registers are not altered.

MRWC 6716 The contents of the low 12 bits of the word count register replace the contents of the AC. The word count register is not altered.

MRMA 6717 The contents of the low 12 bits of the memory address register replace the contents of the AC. The memory address register is not altered.

The bus signals used by this interface are:

DMAREQ-	SKPINT-	EMA2-	EMA1-	EMA0-	C0+
C1+	XTC+	DMAGNT-	EXT2-	EXT1-	EXT0-
MEMSEL-	INIT-	LXMAR-	DEVSEL-	4MHZ-	

DX0- thru DX11-

6.2 Beamformer Hardware Descriptions

The descriptions contained in this section are those which specifically concern the beamformer unit. The descriptions follow the same path that the data follows as it passes from the isolation amplifier input thru to the beamformer output.

6.2.1 Electrical Isolation Unit

An electrical isolation unit was included in the beamformer in order to insure that the CEL equipment would not have any adverse effects on the operations of other experimenters who might be sharing the hydrophone outputs. Each input channel is buffered using an Analog Devices type 288K isolation amplifier. Groups of eight amplifiers have their modulation/demodulation clocks supplied by an Analog Devices type 947 driver module. The combination of the 288K and 947 not only provides ground and signal isolation between the hydrophone outputs and the CEL equipment but also between the individual hydrophones.

The electrical isolation unit is mounted in the top 3 1/2 inches of the beamformer cabinet. The type 288K isolation amplifiers are mounted, along with an output buffer amplifier, on small cards approximately 2 inches by 2 inches. Up to eight of these cards can plug onto a single mother board which also contains a type 947 driver module. One 947 can drive up thru eight 288K modules. Up to 6 mother boards can be mounted in the isolation unit chassis.

The input of each 288K amplifier is shunted with a 1 megohm resistor to provide a path for the 25 na (max) bias current in case the signal feed should be ac coupled. The amplifier gain is unity and the frequency response is rolled off with a single pole filter with its 3 dB point at 800 Hz. The amplifier output is buffered using a unity gain amplifier. Each card has provision for attaching a scope

probe or volt meter at the buffer amplifier output. One adjustment is provided for eliminating any dc offset that might be present.

Connections to the isolation amplifier cards are made via phone jacks mounted on the back of the isolation unit. These jacks also feed thru to corresponding jacks on the front panel. The front and rear panels are non-metallic in order to avoid any electrical connection between the sleeves of the individual jacks.

The outputs of all of the isolation amplifier cards mounted on a given mother board go to a 16 pin connector on that board. A cable containing 8 twisted pairs connects the mother board to eight of the A/D-delay cards located in the card cage mounted just below the isolation unit.

6.2.2 Beamformer A/D and Delay Card

Each isolated and buffered hydrophone output is connected to an individual beamformer A/D and delay card. These cards plug into in the card cage mounted directly below the electrical isolation unit. Each card contains a variable gain amplifier, two pole Butterworth low pass filter, track and hold device, twelve bit A/D converter, 2048 words of twelve bit memory and associated interface and control electronics. The card cage contains a back plane into which up to 24 cards can be inserted. Multiple card cages can be connected together to increase the number of channels.

The input signal is ac coupled to an amplifier whose gain can be digitally set at unity (0 dB) or 256 (48 dB). The 48 dB setting is intended for use when connecting directly to hydrophone outputs without benefit of a hydrophone amplifier. When a hydrophone amplifier is present, the unity gain setting is usually sufficient.

The input amplifier is followed by an active two pole Butterworth low pass filter. The 3 dB frequency is 700 Hz

and the gain is approximately 1.6.

The filter is dc coupled to a digitally controlled attenuator. This attenuator is implemented using a CMOS multiplying D/A converter. The gain can digitally be set to a value ranging from 0 to 255/256 in steps of 1/256.

The attenuator output is ac coupled to a fixed gain amplifier whose gain is approximately 128/1.6. The effect of combining this gain with that of the filter and the attenuator gives a digitally controlled gain range of from 0 to 127.5 in steps of 0.5 (ignoring the potential 48 dB gain of the input stage).

The re-amplified signal is fed to a track and hold circuit manufactured by Analog Devices. The track and hold normally tracks the input waveform. Just before the start of A/D conversion, the device is placed into the hold mode and held there while the A/D converter is operating. Once the conversion is complete, the track and hold is placed back into the track mode. The aperature time required to switch from track to hold mode is 125 ns. The output voltage range is minus ten to plus ten volts.

The A/D converter is a twelve bit Analog Devices type AD582J. The converter is configured for a bipolar input ranging from minus five volts to plus five volts. The converter is of the successive approximation type requiring a maximum of 35 microseconds per conversion. Samples are taken at a 5 KHz rate. The most significant bit (used as the sign bit) is a zero for negative inputs and a one for positive inputs. This is the complement of what would normally be used as a sign bit. This must be corrected for in the user interfaces.

The results of the A/D conversions are written into the 2048 words of twelve bit memory contained on each board. Using a 5 KHz sample rate this amount of memory provides delays ranging from 0 to about 0.41 seconds. Space is provided for an additional 2048 words of memory if increased delay times become necessary.

All of the A/D delay cards contained in a beamformer use the same gain setting and are simultaneously commanded to digitize and update the delay memory. The contents of the individual channel delay memories can be accessed working thru the delay bus controller which is mounted in the beamformer card cage.

The address that a card responds to is set using an eight position jumper header which plugs into a socket located at the top rear of the card. Card addresses are set using a binary representation where a wire present is a one and a wire absent is a zero. Only jumper positions two thru seven are used. Position seven corresponds to the least significant bit of the board address and position two corresponds to the most significant bit. Board addresses can range from 0 thru 63. Normally, a system is configured using sequential board addresses starting at zero. The boards are located in the card cage with the lowest address board in position two and with sequentially increasing board positions corresponding to sequentially increasing board addresses. Card position one is usually either left vacant or has a terminator card inserted in it.

There is a ground pin and four test pins along with five trim adjustments on each A/D delay board. These are all accessible from the front of the card cage without having to place the boards on card extenders. Fig. 6.2 shows the locations of these test points and adjustments.

Test point one is connected to the output of the input amplifier stage which can be set to 0 dB or 48 dB gain. Test point two is connected to the input of the amplifier following the digitally controlled attenuator. Test point three is located at the output of this amplifier which is also the input to the track and hold. Test point four is at the track and hold output. This is also the input to the A/D converter.

The front gain adjustment is used to trim the gain of the input amplifier when it is used in it gain of 256 mode.

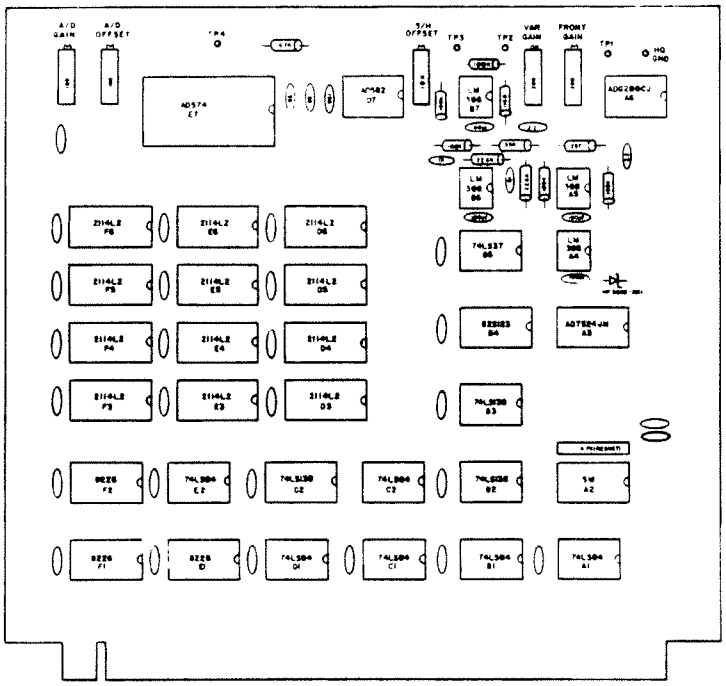


Fig. 6.2. Test point and adjustment locations

To make this adjustment, feed a signal whose amplitude is on the order of 0.01 volts peak into the corresponding isolation amplifier input. With the input amplifier in the unity gain mode, measure the output voltage at test point 1. Switch the amplifier gain to 256 and adjust until the output voltage is 256 times what it was when the gain was unity.

The variable gain adjustment affects the gain of the amplifier following the digitally controlled attenuator. Apply an input on the order of one volt peak at a frequency of about 100 Hz to the associated isolation amplifier input. Set the input amplifier gain to 0 dB. Set the attenuator gain to two. The voltage measured at test point two should be adjusted to match that at test point one.

The next two adjustments affect the dc offset in the digitized output. Normally, the A/D offset adjustment is set mid-range and the dc offset minimized using the track and hold adjustment. The beamformer diagnostic test program

rms test is used to monitor the converted value. Ground the input to the appropriate isolation amplifier. Adjust the offset to minimize the rms value measured by the diagnostic program. It may not be possible to get the value below 1 or 2 A/D counts. There is a background noise level that has to be coped with.

To trim the A/D gain, inject a voltage on the order of 1 or 2 volts peak. Set the digital gain to 2 (this corresponds to a real gain of unity). Measure the rms voltage at the A/D converter input. Compare this to the diagnostic program rms value. The maximum count (2047) on the A/D converter corresponds to 4.9963 volts. Take the computed rms value and multiply it by 4.9963/2047 and compare it against the measured value. Trim as needed.

Two A/D card cages can be chained by cabling position 25 of the unit closest to the delay bus controller to position 1 of the following card cage. The terminator resistors normally located on the back of the first card cage at position 25 should be moved to the back of the second card cage position 25.

6.2.3 Delay Bus Controller

The delay bus controller (DBC) is used to perform the following functions:

- Generate the bus used by the A/D delay boards.
- Generate a bus over which data users can extract sample values from the individual A/D delay memories. Access to these memories is granted on priority basis. All user delay addresses are relative to the most recent sample.
- Generate the master timing from a 5 MHz Rb clock.

The DBC does not control the gains of the A/D delay cards. This is done by the beamformer processor user interface.

Position three of the beamformer card cage is wired

specifically for the DBC. The signals on pins 74 thru 118 are connected to the corresponding connector pins on position one. These pins correspond to the A/D delay bus. A cable which pushes onto the position one connector wire-wrap pins connects these bus signals to the card cage holding the A/D delay boards. Connector pins 5 thru 57 define a user bus over which user interfaces can gain access to the A/D delay memories. As needed, these pins are wired to positions 11, 13, 15 and 17. Positions 11 and 13 are wired for use with beamformer data logic panels, position 15 is wired for use by the beamformer control logic panel and position 17 is wired for use by the beamformer processor user interface.

The DBC causes samples to be taken and the resulting values to be written into the delay memories at a 5 KHz rate. This basic 200 microsecond interval is divided into 250 0.8 microsecond time slices. Eight of these are used to initiate the sampling and to update the delay memories. The remaining 242 time slices are made available to user interfaces.

A user interface requests a time slice by grounding a request line. When the request is granted, the interface must supply a three bit control word, six bit A/D delay board address, twelve bit delay address and (if writing to the delay memory) twelve bits of data. Time slice requests are granted on a priority basis. The beamformer is wired for the highest priority and the processor user interface is wired for the lowest priority. At present, there are no other user interfaces.

The signals provided for user use are:

Pin 5. U10KHZ-. This clock is for the use of the beamformer demodulator in generating the timing used by the frequency synthesizer. It bears a well defined timing relation to the 5 KHz clock used in updating the delay memories. This relation is exploited in the demodulator!

Pins 6 and 7. 5 MHz Clock. Pin 6 is the 5 MHz clock input and pin 67 is the corresponding ground point. This line is terminated on the DBC using a 220/330 terminator between plus five volts and ground. These two pins are connected to a bnc connector on the beamformer front panel. Normally, one of CEL's Rubidium standards is used as the clock source.

Pin 7. U5KHZ-. This is the 5 KHz clock used by the beamformer to initiate the formation of a set of beam outputs. The trailing edge is the active edge.

Pins 8 and 9. UCLK0- and UCLK1-. These waveforms are two square waves with an 800 ns period and shifted 90 degrees with respect to each other. The positive going transition of UCLK0- marks the starts of the 800 ns time slices. This edge should be used to sample the time slice grant response. The positive going transition of UCLK0- marks the time at which the time slice request lines are sampled for access to the next time slice. Signal UCLK1- is included in order to provide timing transitions at times other than the UCLK0- times. Such transitions are often needed in interface designs.

Pins 10 thru 12. Not used.

Pins 13 thru 20. UPR0- thru UPR7-. These are the bi-directional lines over which time slice requests and grants are made. Requests for time slice access should be made when CLK0- is high. This corresponds to the first half of an 800 ns time slice. The response can be sampled using the positive going transition of CLK0- (which indicates the start of a new time slice). The DBC is designed to hold its response well past this transition so that there will not be any race problem. These lines should be driven using open collector devices. Line UPR0- has the highest priority and line UPR7- has the lowest.

Pin 21. UWRINH-. This line is commanded by the processor

user interface. If this line is driven low, all writes by the DBC into the delay memories are inhibited. This feature is used to allow the processor to run memory diagnostics on the delay memories without having to contend with data values changing the memory contents.

Pin 22. Not used.

Pins 23 and 24. UC0- and UC1-. These two lines are used to control the type of transfer to be made between the user interface and the selected A/D delay board. If a single A/D delay board is being addressed (the UENALL- line is not grounded), the effect of these two lines are:

UC0-	UC1-	Action
0	0	A/D conversion is initiated. The data lines are ignored.
0	1	A/D converter output is written to memory. The converter output is also placed on the data lines.
1	0	The contents of the data lines are written into delay memory.
1	1	The contents of the addressed location in delay memory are placed on the data lines.

If the UENALL- line is grounded, action differs in that, because all boards respond, none of them place any values on the data lines, regardless of the values of UC0- and UC1-.

Pin 21. UENALL-. This signal is used to cause all A/D delay boards to respond to the command given in the current time slice. This is typically used to broadcast values when performing memory diagnotics.

Pins 26 thru 31. UAD0- thru UAD5-. These lines are used to address the individual A/D delay boards. The board whose jumper set address corresponds to the address on these lines responds and performs the commanded operation. UAD0- is the most significant board address bit and UAD5- is the least significant.

Pin 32. Not used.

Pin 33. UCLCTR-. Driving this line low causes the DBC address register to be set to zero and held there as long as the line is low. This line is used by the processor user interface in order to suppress the address bias added by the DBC from changing every 0.2 milliseconds when diagnostic programs are being run.

Pins 34 thru 45. UA0- thru UA11-. This is the delay memory address relative to the DBC address register that a user interface wants to access. The value on these lines is added to the value in the DBC address register with the sum being the actual memory address used. DBC memory addresses are advanced by one every 0.2 milliseconds as a new sample becomes available. The delay values supplied by a user device are usually negative.

Pins 46 thru 57. UD0- thru UD11-. These are bi-directional data line used to move data values between the A/D delay boards and the user interfaces. The direction of data flow is controlled by the UC0- and UC1- lines. Line UD0- carries the most significant bit and line UD11- carries the least significant bit.

6.2.4 Processor Delay User Interface

The processor delay user interface (PDUI) provides the beamformer DM1200 processor access to the individual A/D delay card memories. Using this access, the processor can run memory tests on the individual memories and check out the addressing support given by the delay bus controller. When data is being acquired it is also possible, on a non-interfering basis, to monitor, in real time, the outputs of the individual A/D converters.

The processor delay interface is contained on a single wire-wrap panel that plugs into position 17 of the beamformer card cage. This position is wired specifically

for the PDUI and connects to both the DM1200 processor bus and to the delay bus controller.

The PDUI contains four registers and two flags. These are:

Data Register.

This register is used to buffer data words moving between the A/D delay memories and the processor. This register contains twelve bits and can be read or written by the processor. If writing data values to the delay memories, this register must be loaded with the proper value prior to loading the command register. The value is not written until the data done flag becomes set. If reading data values from the delay memory, the desired value is not present until the command done flag becomes set. When reading A/D values, remember that the sign bit must be complemented.

Address Offset Register

The delay bus controller contains an address register whose contents point to the most recently acquired sample value in the delay memories. The PDUI accesses these memories by specifying an offset to this address. The effective address corresponds to the contents of the delay control address register added to the contents of the address offset register. For a delay of zero use a value of zero. For a delay of one use an offset of -1. By setting the appropriate bit in the control register, the contents of the delay bus controller address register can be held at zero. The address offset register is 12 bits in length. Since the A/D delay cards currently only possess 2048 words of memory, all addresses are formed modulo 2048.

Gain Register

The gain settings on all of the A/D delay cards are controlled by this register. This register is loaded from bits 3 thru 11 of the accumulator. Gain values

range from 0 thru 127.5 in steps of 0.5.

Control Register

The control register contents specify the operation to be performed. Loading this register initiates a data transfer request of the form specified by the register contents. Certain bits of this register have a global effect not just local to the granted time slice. In particular, the writing of A/D sample values into the delay memories can be inhibited and the delay bus controller address register can be held in the zero state. Bit assignments are:

Bit 0. A one in this bit inhibits the writing of A/D sample values into the delay memories.

Bit 1. A one in this bit holds the delay bus controller address register in the zero state. Tics of the 5 KHz clock will not increment the address as long as this bit is a one.

Bit 2. Not used.

Bit 3. If a one, the operation to be performed as a result of this IOT should apply to all A/D delay boards.

Bits 4 and 5. These bits correspond to the UC0+ and UC1+ control values. The indicated operations are:

UC0+	UC1+	Action
0	0	Read from the delay memory board addressed by bits 6 thru 11 of the command register. No boards respond if command register bit three is a one.
0	1	Write the contents of the data register into the memory of the A/D delay board

addressed by bits 6 thru 11 of the command register. All boards respond if bit three of the command register is a one.

1	0	Don't use.
1	1	Don't use.

Bits 6 thru 11. The address of the A/D board which is to be involved in the data transfer. Ignored if bit three is a one.

Data done flag. This flag is set whenever the data transfer initiated using the DUWC IOT instruction has been completed. This flag is necessary because the PDUI interface has been wired for the lowest access prioity and it may take some time for the desired transfer to take place. This flag is cleared by the DUWC instruction and by a skip on the DUSD instructin.

Five KHz tic flag. This flag is used to synchronize the operations of the processor with the 5 KHz clock used to time the sampling. This flag is set on the trailing edge of the 5 KHz clock. It is cleared by a skip using the DUST instruction.

Instruction set.

DUWD	6400	The contents of the accumulator are loaded into the interface data register. The accumulator is then cleared.
DURD	6401	The contents of the interface data register replace the contents of the accumulator. The contents of the data register are unchanged.
DUWO	6402	The contents of the accumulator are loaded into the interface address offset register. The accumulator is then cleared.
DURO	6403	The contents of the interface address address register replace the contents of the

accumulator. The contents of the address offset register are unchanged.

DUWG 6404 The contents of the accumulator are loaded into the interface gain control register. The accumulator is then cleared. The interface gain register directly controls the gain settings on the A/D delay cards.

DURG 6405 The contents of the interface gain register replace the contents of the accumulator. The contents of the gain register are unchanged.

DUWC 6406 The contents of the accumulator are loaded into the interface command register. The indicated operation is then initiated. The accumulator is cleared. Bit assignments are described above. The processor user delay interface has been assigned the lowest access priority and it may take some time for the commanded operation to be completed. Use the DUSD instruction to wait for completion.

DURC 6407 The contents of the interface command register replace the contents of the accumulator. The interface command register is unchanged.

DUIN 6410 Initializes the processor user delay interface. This involves setting the contents of all four registers to zero, clearing the two flags and clearing any pending data bus access request.

DUSD 6411 If the command done flag is set, the immediately following instruction is skipped. The command done flag is cleared by the DUWC instruction and by a successful skip using the DUSD instruction.

DUST 6412 Skips the immediately following instruction if the 5 KHz tic flag is set. The flag is cleared if a skip occurs.

The processor bus signals used by this interface are:

SKPINT- C0- C1- XTC+ INIT- LXMAR-
DEVSEL- DX0- thru DX11-

6.2.5 Beamformer

The beamformer is of the classical delay and sum type. A list of A/D delay channels and delay values is used to control the formation of beam output values. This list is maintained in a read/write memory. This memory can be read by the beamformer control electronics and can be read and written by the DM1200 processor mounted in the beamformer card cage. Beam outputs are generated at the same basic 5 KHz rate that input samples are taken at. Once a set of output beam samples has been formed, a pulse noting this is sent to the demodulator control logic in the signal processor unit. The output values are then serially shifted over sets of twisted pairs to the demodulators. When the next 5 KHz clock tic occurs, the process repeats. The beamformer control memory can be read and written by the beamformer processor without affecting the beam forming. It is thus possible to change beams "on the fly."

The beamformer consists of a control board and one or two quad beam boards. The control board is a wire-wrap panel that plugs into position 15 of the beamformer card cage. At present, up to two quad beam boards can be included in a system. These mount in card cage positions 11 (beams 4 thru 7) and 13 (beams 0 thru 3). Positions 11, 13 and 15 are specifically wired for these logic panels. Connection to the demodulators is made using sets of twisted pairs which connect the individual beam outputs to the associated demodulator boards.

Once initialized, the beamformer normally operates independently of the processor. The design of the beamformer is divided into three major parts; processor interface; control logic and beam accumulators.

The processor interface implements the instruction set described at the end of this section. This interface provides the processor with read/write access to the beamformer control memory and read access to the beam

accumulators.

The control logic is built around a 24 bit wide memory of 1024 words. This memory contains lists of A/D delay channels and the associated delays needed to form the desired beam outputs. The processor views this memory as being composed of 2048 words of twelve bits with even addressed locations containing control bits and odd addressed locations containing the associated delay address offsets. The twelve bit control word is organized as

Bit 0	A one in this position terminates the beam list.
Bit 1	A one in this bit causes this entry (except for bit 0) to be ignored.
Bits 2-4	Number of the beam into which the retrieved sample value is to be added.
Bit 5	Not used.
Bits 6-11	A/D delay channel to retrieve a delayed sample from.

In the location following this control word is a twelve bit address offset value. If a delay of n sample times is desired, this value should be the two's complement of n.

The beamformer controller views the memory as being divided into four fields of 256 words 24 bits in length. The beamformer addresses this memory using a two bit field address register and an eight bit control word address register. The eight bit control word address register is set to zero at the beginning of each 5 KHz interval and is incremented by one each time the beamformer extracts a data value from the A/D delay memories.

The field register is only settable by the processor. This register is used to divide the control memory into four banks of 256 words each. This allows four sets of beam tables to be maintained in the control memory with the currently active table being selected by the field register contents. Values loaded into the field register are first buffered until the start of a new 5 KHz interval. At this

time they are loaded into the field register. This allows a glitch free transition between beams to be made. It is suggested that the processor synchronize its loading of the field register buffer with the 5 KHz clock tics.

Once the processor has loaded the beamformer control memory with the desired control and delay information, the operation proceeds as follows:

1. The delay bus controller completes the acquistion of a new sample set and updates the delay memories. The 5 KHz clock tic generated by the delay bus controller signals this event. This pulse is used by the beamformer to zero the beam accumulators and to reset the eight bit control memory address register to zero.
2. Control words (in 24 bit format) are sequentially read from the control memory and are used to initiate data requests from the A/D delay memories. Each 24 bit word specifies a beam number, A/D delay channel and a delay offset.
3. The requested data values are summed into the proper beam accumulator.
4. The eight bit control memory address register is incremented.
5. Steps two thru four continue until a control word with bit 0 set to a one is encountered. After this control word is processed, operation is suspended until the next 5 KHz clock tic.
6. A control pulse is sent to the demodulator control logic. The beam formed values are transmitted serially to the associated demodulator boards. There must be sufficient time allocated for this when setting up the beam table.
7. The control logic waits for the next 5 KHz clock tic. When it occurs, the operation returns to step 1 above.

The processor has the ability to freeze the contents of the beam accumulator registers. This was provided to allow the writing of test programs which can verify system

operation by writing known data values in the A/D delay memories and then checking the resulting beam values to see if they are correct. The use of the freeze command should be synchronized to the beamformer operation using the 5 KHz clock tic. The freeze takes place at the end of the 5 KHz interval that the command was executed in. The beamformer timing proceeds as normal but the beam accumulators are not zeroed and new values are not added into them. The freezing does not affect the clocking of data out of the beam accumulators by the demodulator boards. Because of this, if the beam values are to be checked, it is necessary to disconnect the associated demodulators.

Instruction set.

BFLA 6430 The contents of the accumulator replace the contents of the the beamformer memory address register. Even address memory locations are used to hold beam control words and odd address locations contain delay offsets. Delays are expressed as the two's complement of the desired value. This instruction also clears the freeze state. The accumulator is cleared.

BFRD 6431 Because of how the control memory was organized in order to allow access from both the beamformer control logic and the processor, this instruction is slightly awkward. A buffer register sits between the beamformer control memory and the processor. Executing this instruction copies the contents of that buffer into the accumulator and then copies the contents of the addressed control memory location into the buffer. Thus, if the contents of a specific location are desired, load the address register and execute two successive BFRD instructions. If sequential memory locations

are being read, all that needs to be remembered is that there is a one word delay between the control memory and the processor.

BFWD 6432 The contents of the accumulator are written into the addressed twelve bits of the beamformer control memory. The accumulator is then cleared.

BFRF 6433 The contents of the beamformer field register are read into the accumulator bits eight thru eleven. The field register is four bits long with only the low two bits being used. Bits zero thru seven of the accumulator are set to zero. The contents of the field register are not altered. This instruction also clears the freeze state.

BFWF 6434 The contents of the accumulator bits eight thru eleven are loaded into a buffer register. At the next 5 KHz clock tic, the contents of this buffer register are transferred into the beamformer memory field register. In order to avoid any potential timing problems, this instruction should itself be executed just after a 5 KHz clock tic. The accumulator is cleared.

BFRM 6435 The most significant eight bits of the selected beam output register are read into the accumulator bits four thru 11. The sign bit is not extended into bit positions zero thru three. The contents of the output register are not altered.

BFFB 6436 Freezes beamforming at the end of the current 5 KHz interval and loads the beam address register from accumulator bits nine thru eleven. Scanning of the control memory is not affected. This allow demodulators to maintain their timing even though beam outputs are not being produced. The accumulator is cleared. The freeze state is

cleared by the BFLA and BFRF instructions.

BFRL 6437 The contents of the low twelve bits of the addressed beam output register are loaded into the accumulator. The beam output register contents are not affected.

The processor bus signals used by this device are:

C0- C1- XTC+ INIT- LXMAR-
DEVSEL- DX0- thru DX11-

CHAPTER 7

PROCESSING THE DEMODULATED RECEPTION

This chapter describes how the demodulated reception can be processed in order to remove the time spreading and to estimate the signal-to-noise ratios of various reception parameters.

The received waveform is periodic. The amount of interference caused by additive noise can be reduced by dividing the reception into a number of one period long segments and ensemble averaging. The number of segments that can be usefully combined is limited by frequency drift in the transmitter and receiver clocks, by motion of either the transmitter and/or receiver and by changes in the ocean structure. Section 7.1 describes the filtering effect of forming such an ensemble average and examines the limits imposed by clock drift, mooring motion and ocean stability.

Section 7.2 describes two methods which have been successfully used to estimate the noise level in the processed reception.

Removal of the time spreading caused by the binary sequence used to modulate the transmitted carrier is the topic of Section 7.3. Attention is restricted to the use of linear maximal sequences and correlation processing. The inherent simplicity of the required program organization is demonstrated using program segments written using a Pascal-like notation. Most real-time data acquisition and processing programs are written in assembly language. The transliteration of these code segments into assembly language is a straight forward process complicated only by

restrictions imposed by the architecture of the target computer.

Section 7.4 combines the results of the preceding sections and describes how to calculate the signal-to-noise ratios of various reception parameters.

The processing is often carried out in computers possessing twelve or sixteen bit word sizes using integer rather than floating-point arithmetic to speed up processing. In order to be able to contend with the expected dynamic range of the processed results, one is driven to using double precision. In Section 7.5 the maximum allowable magnitude of the signal component in the demodulated reception is calculated assuming that the intermediate and final values of various parameter calculations are not to overflow double precision.

Before a processing program is placed into production, its proper operation must be verified. Section 7.6 describes the use of three test waveforms for testing. Using the RMS imput level as computed by the beamformer processor, the parameter values to be calculated by the program can be predicted and then compared to the values determined by the program.

7.1 Filtering the Signal

The transmissions used in making multipath measurements are periodic with each period made up of L digit intervals each of duration T seconds. The received waveform is complex demodulated and sampled at a rate of M samples per digit. The demodulated and sampled reception is divided into N segments each one period (LT seconds) in duration. These segments are used to form an ensemble of one period receptions and are summed together. This processing is usually implemented in a computer using a one period long ML value buffer, summing demodulated values into successive buffer locations. The entire buffer is circulated thru once

per period. This form of processing is referred to as forming a circulating sum or average, however, the summed values are rarely, if ever, divided by the number of values summed.

If the portion of the reception due to the signal does not vary from period to period, the expected value of a given position in the circulating sum equals N times the received signal amplitude at that positon. If the noise is zero mean and independent from period to period, the variance of the noise in a given circulating sum positon is increased by a factor of N. The net effect of an N period circulating sum is an improvement in the signal-to-noise ratio by a factor of N.

Assume that the magnitude of the signal dependent portion of the reception does not vary from period to period but that the phase changes at a uniform rate of θ radians per period. The magnitude of the sum of N samples spaced LT seconds apart is the magnitude of a single such sample multiplied by

$$\sum_{k=0}^{N-1} \exp\{jk\theta\} = \exp\{j(N-1)\theta/2\}\frac{\sin(N\theta/2)}{\sin(\theta/2)} \qquad (7.1)$$

Assuming zero mean, independent noise components, the variance of the noise in the sum is N times the variance associated with one sample. The net effect of the summation has been to multiply the single sample signal-to-noise ratio by

$$\frac{\sin^2(N\theta/2)}{N\sin^2(\theta/2)} \qquad (7.2)$$

Given θ, the value of N that maximizes this expression can be determined by taking N to be continuous, differentiating with respect to N, setting the result equal

to zero and solving. Doing this, the desired value of N is found to be the smallest value satisfying the relation

$$N\theta = \tan(N\theta/2) \tag{7.3}$$

Solving gives

$$\begin{aligned} N\theta &= 2.331 \text{ radians} \\ &= 133.6 \text{ degrees} \end{aligned} \tag{7.4}$$

Once the accumulated linear phase drift has reached about 134°, adding additional samples decreases rather than increases the signal-to-noise ratio. For small values of θ, the signal-to-noise ratio initially doubles each time N doubles. As $N\theta$ approaches 134°, the signal-to-noise ratio increases more slowly as a function of N. For example, going from $N\theta = 67^{\circ}$ to $N\theta = 134^{\circ}$ only increases the signal-to-noise ratio by a factor of 1.394 (1.44 dB).

The period of the 1981 Tomography transmission was 7.9375 seconds and the carrier frequency was 224 Hz. If the transmitter and receiver oscillator frequencies were within 1 part of 10^6 of each other and this difference was constant, the phase shift per period would have been 0.64 degrees. This amount of oscillator error would have limited the maximum number of periods which could usefully be averaged to about 208. This corresponds to an integration time of about 27.5 minutes. (The oscillator drifts in the actual experiment were much less than 1 part in 10^6.)

The moorings on which the transmitters were attached experienced a low frequency meander onto which was superimposed a high frequency two cycle per day tidally induced motion. This tidally induced motion was approximately circular. The 1981 Tomography mooring motions due to tides were measured to have excursions typically on

the order of 40 meters. Assuming circular motion with diameter 40 meters and period of twelve hours gives a source velocity of 0.29 cm/sec. As a worst case model, take this velocity to be constant and in a direct line with a fixed receiver. At 224 Hz, assuming the speed of sound to be 1500 meters/sec., the tidally induced mooring motion requires about 860 seconds to induce a 134° phase shift in the reception. The upper limit on integration time due to mooring motion was thus on the order of 14 minutes.

The extent to which the ocean limits the coherent processing is the topic of current research. The results of much of this research have been summarized in a book edited by Flatte' [50]. The primary causes of travel time disturbances which act over time scales of minutes and hours are internal waves. Internal waves are oscillatory displacements of layers of water which differ in salinity and/or temperature from the surrounding water. The limits which these waves place on integration time depend on range, receiver depth, transmitter depth, geographic location and season. Experience has shown that at ranges up to about 500 km, it is generally possible to integrate over intervals of several minutes. At times, the ocean is more stable and supports longer integrations and at other times the ocean is less stable making measurements difficult.

The sample values used in forming a circular sum correspond to N periods of the transmitted waveform. A total of NML sample values are used per data set. If the DFT of these NML values were to be formed, the resulting analysis bins would be 1/(NMLT) Hz apart. Since the transmitted waveform has period MLT seconds, the lines in its spectrum are spaced 1/(MLT) Hz. If there are negligible amplitude and phase variations in the signal portion of the reception, signal energy will only be found in the NML point DFT lines whose indices are multiples of N.

The NML point DFT of the demodulated and sampled

reception is

$$R(k) = \frac{1}{NML}\sum_{i=0}^{NML-1} r(i)\exp\{-j2\pi\frac{ik}{NML}\} \qquad (7.5)$$

Let $i = i'+nML$, $0 \le i' < ML$, $0 \le n < N$. The above sum can be rewritten as (dropping the primes)

$$R(k) = \frac{1}{ML}\sum_{i=0}^{ML-1} \left(\frac{1}{N}\sum_{n=0}^{N-1} r(i+nML)\exp\{-j2\pi\frac{nk}{N}\}\right) \exp\{-j2\pi\frac{ik}{NML}\} \qquad (7.6)$$

When $k = aN$, $0 \le a < ML$

$$R(aN) = \frac{1}{ML}\sum_{i=0}^{ML-1} \left(\frac{1}{N}\sum_{n=0}^{N-1} r(i+nML)\right) \exp\{-j2\pi\frac{ia}{ML}\} \qquad (7.7)$$

The inner term corresponds to the circulating average of N periods of the reception sampled ML times per period. The right hand side of the above equation corresponds to the ML point DFT of the circulating average. The lines in this DFT correspond to the lines whose indices are multiples of N in the NML point DFT of the entire data set.

7.2 Estimating the Noise Level

An estimate of the noise level in the reception is desired in order to be able to evaluate the quality of the multipath measurements. Such a noise measurement can be made in a number of ways. Two of these are:

A circulating sum can be formed where the data values in alternating periods are multiplied by minus one. If the number of periods, N, is even and the signal portion of the reception does

not change significantly in one period, this process very effectively generates signal-free averages of noise samples that can be used to estimate the noise variance and the noise power spectrum, if desired.

The total energy in the reception can be found by summing the squared magnitudes of the demodulate values used to form the N period circulating sum. Parseval's theorem can be used to relate this quantity to the energy in the DFT of the total reception. In the previous section, it was shown that the lines in the DFT of the circulating sum correspond to the signal lines in the DFT of the entire reception. Parseval's theorem can be used to relate the energy in the circulating average to the energy contained in these lines. Subtracting the estimated signal energy from the total energy in the reception yields the energy in the noise. This quantity can then be used to estimate the noise variance.

In the first method, the operations required to form the signal and noise circulating sums can be combined if a single two period buffer is used instead of two single period buffers. (This is T. Birdsall's DUAL method.) Data values are summed into this extended buffer, returning to the buffer beginning every second period. Once N periods of data have been acquired, the front and back halves of the buffer each contain an N/2 period ensemble sum. The N period sum buffer used to filter the signal is formed by ensemble summing the values contained in the two halves of the buffer. The N period alternating sum/difference buffer used to estimate the noise level is formed by differencing the two buffer halves.

The complex valued noise components in the samples

(demodulates) of the demodulated reception are assumed to be zero mean, independent and identically distributed with variance V_d. Write the sample values of the demodulated waveform as

$$r_n(i) = r(i) + n(i) \tag{7.8}$$

where r(i) represents the signal dependent portion of the reception and n(i) the noise portion. Both r(i) and n(i) are complex valued.

The sum of the squares of the values contained in the alternating sign circulating sum can be written

$$S_c = \sum_{i=0}^{ML-1} \left| \sum_{n=0}^{N-1} (-1)^n r_n(i+nML) \right|^2 \tag{7.9}$$

If N is even then

$$S_c = \sum_{i=0}^{ML-1} \left| \sum_{n=0}^{N-1} (-1)^n n(i+nML) \right|^2 \tag{7.10}$$

The expected value of the sum of the squares of the alternating sign noise summation values is

$$E\{S_c\} = MLNV_d \tag{7.11}$$

If the noise is Gaussian, then S_c is Chi-square distributed with 2ML degrees of freedom. The standard deviation of S_c is

$$Dev\{S_c\} = (ML)^{1/2} NV_d \tag{7.12}$$

The estimated variance of the noise in an individual signal sum buffer bin is

$$V_c = \frac{S_c}{ML} \tag{7.13}$$

The expected value of V_c is

$$E\{V_c\} = NV_d \tag{7.14}$$

and the standard deviation is

$$Dev\{V_c\} = \frac{NV_d}{(ML)^{1/2}} \tag{7.15}$$

Next, the noise level estimation method based on Parseval's theorum is analyzed.

The complex valued noise components in the demodulate values are again taken to be zero mean, independent and identically distributed with variance V_d. The sum of the squares of the demodulates contained in N periods of the reception is

$$S_d = \sum_{i=0}^{NML-1} |r_n(i)|^2 = \sum_{i=0}^{NML-1} |r(i)+n(i)|^2 \tag{7.16}$$

The expected value of S_d is

$$E\{S_d\} = NMLV_d + \sum_{i=0}^{NML-1} |r(i)|^2 \tag{7.17}$$

If the signal portion of the reception does not vary significantly over N periods, then

$$\sum_{i=0}^{NML-1} |r(i)|^2 = N \sum_{i=0}^{ML-1} |r(i)|^2 \tag{7.18}$$

giving

$$E\{S_d\} = NMLV_d + N\sum_{i=0}^{NML-1}|r(i)|^2 \qquad (7.19)$$

The sum of the squares of the sum circulating average over N periods of the reception is

$$S_s = \sum_{i=0}^{ML-1}\left|\sum_{n=0}^{N-1} r(i+nML)+n(i+nML)\right|^2 \qquad (7.20)$$

The expected value of S_s is

$$E\{S_s\} = NMLV_d + \sum_{i=0}^{ML-1}\left|\sum_{n=0}^{N-1} r(i+nML)\right|^2 \qquad (7.21)$$

Again assume that the signal dependent portion of the reception changes little over N periods. In this case

$$\left|\sum_{n=0}^{N-1} r(i+nML)\right|^2 = N^2|r(i)|^2 \qquad (7.22)$$

giving

$$E\{S_s\} = NMLV_d + N^2\sum_{i=0}^{ML-1}|r(i)|^2 \qquad (7.23)$$

The signal energy can be subtracted out of the sum of the square of the demodulates using the value estimated by the sum of the squares of the signal enhancing circular sum. Define

$$S_e = S_d - \frac{S_s}{N} \tag{7.24}$$

The expected value of S_e is

$$E\{S_e\} = (N-1)MLV_d \tag{7.25}$$

The estimated variance of the noise in an individual signal circulating sum buffer bin is

$$V_c = \frac{NS_e}{(N-1)ML} = NV_d \tag{7.26}$$

If the noise is Gaussian, then S_e is approximately Chi-square with 2(N-1)ML degrees of freedom. The standard deviation of V_c in this case is

$$Dev\{V_c\} = \frac{NV_d}{((N-1)ML)^{1/2}} \tag{7.27}$$

Both of the above two methods have been successfully used to estimate the noise level in the circularly summed reception. Of the two methods, the alternating circular sum/difference is preferred. The reasons for this preference are:

1. Because alternate periods are summed and subtracted, this approach requires only a negligible change in the signal component of the reception from period to period rather than from beginning to end of N periods.
2. Using a two period buffer to form the circulating sum of the reception allows the noise level to be estimated without additional computational effort during data acquisition.

The primary disadvantage of this method is the need for the dual length buffer. A one period buffer for L = 511, M = 4 requires approximately 8k twelve bit words (using double precision). A 32k DM1200 only has sufficient room for three such buffers if 4k is reserved for program use. This problem will be alleviated once the DM1200 augmented extended memory address hardware is built. For the time being, when processing L = 511, M = 4 based receptions, a buffer of length L is used to form an alternating sum/difference data array using only every fourth reception sample on one input channel. The resulting noise level estimate is somewhat cavalierly used to characterize the noise level on all three channels.

The noise level estimation method based on Parseval's theorem is useful when memory space is at a premium. In those cases where the signal level in a demodulate is not particularly high compared to the noise level, the stability of the signal is not a major problem or at least no more so than it is in trying to estimate the multipath structure. Signal fluctuations cause a leakage of energy from the signal lines into adjacent lines which is not accounted for in the circulating average signal power estimate. The resulting noise level estimate may be slightly high.

Using the DM1200 fast arithmetic device described in Chapter 6, the required squaring of the demodulate magnitudes can be performed quite rapidly but does increase the interrupt time computational load.

7.3 Time Spread Removal

In Chapter 4, it was shown that the time spread placed on the transmitted waveform can be removed using either time or frequency domain operations. The time domain approach is preferred for the following reasons:

1. The lengths of the linear maximal sequences used to

cause the time spreading do not factor "nicely", limiting the computational speedup provided using a fast DFT algorithm. For example, the 1981 Tomography sequence contained 127 digits which is prime. Because fast algorithms split a transform of a given length, L, into a succession of transforms of length equal to the prime factors of L, no speedup is possible when L is prime.

2. The time domain correlations can be implemented mainly using adds and subtracts with relatively few multiplications. Because of this, the times required for correlation processing are comparable to the times required by multiplication intensive DFT's, even when fast algorithms can be used.

3. The memory required for the code to process a reception using correlation is much smaller than that required by the code needed to compute DFT's.

The ML complex valued data array used to form the N period circulating sum contains the data values from which the time spread is to be removed. Because each digit interval has been sampled M times, the correlation processing is organized as M interleaved circular correlations of length L. These correlations cannot be performed in place; an L complex value scratch array is required. The processing organization favored by the author is described below using a Pascal-like notation. The array cs contains the circular sum values and is ML complex values in length. The array sa is a scratch array, L complex values in length. The procedure CORR forms one circular correlation value per call, starting the correlation at the position indicated by i. CORR is described in more detail later.

```
for mi:=0 to M-1 do
  begin
    grandsum=0;
    for i:=0 to L-1 do
      begin
        sa(i):=cs(i*M+mi);
        grandsum:=grandsum+sa(i);
      end;
    f1xgsum:=f1*grandsum;
    for i:=0 to L-1 do cs(i*M+mi)=CORR(sa,i,f1xgsum);
  end;
```

The above code segment is intended to illustrate how the data interleaving is handled. The reason for forming the correlations with the data values first being copied into the scratch array is that this simplifies the addressing calculations in the DM1200 when the equivalent operations are programmed in assembly language. The grand sum value is used in forming all of the correlation values and only needs to be calculated once for each interleave position. The variables grandsum, f_1, and f1xgsum are complex valued.

The desired correlation can be performed in a great number of different ways. The equation describing one possible implementation is

$$cv(i) = f_1*grandsum+\sum_{k=0}^{L-1} sa(k+i)(c^*s(k)+d^*) \qquad (7.28)$$

where cv(i) is the i-th correlation value and $c^*s(k)+d^*$ is the local reference sequence representation as described in Chapter 4. All indices are computed modulo L. In order to make the processing as fast as possible, the local reference is chosen to use values c = -1/2, d = 1/2. The resulting values of $c^*s(i)+d^*$ are then 0 and 1 corresponding to binary

zeros and ones in the underlying sequence. Since f_1*grandsum was precomputed, each value of cv(i) is formed using only (L+3)/2 complex additions. Because there are no "slow" arithmetic operations involved, the amount of time required for address computation and for moving data values and partial sums about often dominates the computation time.

One way to speed up the address computation is to precompute a list of the (L+1)/2 k values corresponding to $c^*s(k)+d^* = 1$ and to use these values in generating the indices of the sa values to be included in cv(i). In those situations where there is insufficient memory to contain such a table, the sequence bits are easily generated as needed. Depending on the hardware resources available, generating the sequence bits usually results in a small to moderate increase in processing time.

Assuming an array sequence named seq which is initialized with the index values corresponding to ones in the base sequence, the computational portion of the correlation procedure can be written as

```
cv(i):=f1xgsum;
for k:=0 to (L-1)/2 do
  begin
    index=i+seq(k);
    if index>=L then index:=index-L;
    cv(i)=cv(i)+sa(index);
  end;
```

Using the mod function this can be compacted to

```
cv(i):=f1xgsum;
for k:=0 to (L-1)/2 do cv(i):=cv(i)+sa(mod(seq(k)+i,L));
```

Which form to use depends on the execution time of the mod function.

7.4 Signal-to-Noise Ratio Relations

Let $r_c(i)$ and $n_c(i)$ be the signal and noise components in the i-th circulating sum bin. The expected value of the total energy in the circulating sum is

$$E\{S_p\} = E\{\sum_{i=0}^{ML-1} |r_c(i)+n_c(i)|^2\} \tag{7.29}$$

$$= \sum_{i=0}^{ML-1} |r_c(i)|^2 + MLV_c \tag{7.30}$$

where as in section 7.2

$$V_c = E\{|n_c(i)|^2\} \tag{7.31}$$

S_p is an estimate of the total signal-plus-noise energy in the circulating summed reception. Define N_p as being ML times the value of V_c which is estimated using one of the methods described in Section 7.2. The signal-to-noise ratio in the circulating sum is then estimated by

$$(\frac{S}{N})_{cs} = \frac{S_p}{N_p} - 1 \tag{7.32}$$

Often the -1 is dropped and the signal-plus-noise-to-noise ratio is accepted as a good approximation to the signal-to-noise ratio.

The energy in the carrier line can be estimated using the magnitude squared of the sum of the circulating sum values. The result is equivalent to evaluating the NML DFT of the demodulated samples for k = 0. The expected value of this sum is

$$E\{C_p\} = E\{|\sum_{i=0}^{ML-1} r_c(i)+n_c(i)|^2\} \tag{7.33}$$

$$= |\sum_{i=0}^{ML-1} r_c(i)|^2+MLV_c \tag{7.34}$$

With N_p equal to ML times the estimated value of V_c, the signal-to-noise ratio of the carrier line is estimated by

$$(\frac{S}{N})_{car} = \frac{C_p}{N_p}-1 \tag{7.35}$$

The -1 is usually dropped and the signal-plus-noise-to-noise ratio is accepted as a good approximation to the signal-to-noise ratio.

Assume a CPM transmission using modulation angle θ. Using Eq. 4.122, the variance of the noise in a correlation value for this case is

$$V_{nc} = \frac{V_c cc^*(L+1)}{L}\{L-1+\frac{L+1}{\frac{L^2}{\tan^2\theta}+1}\} \tag{7.36}$$

The term in braces varies from L-1 to L as θ goes from 0 to the angle that maximizes the signal-to-noise ratio, $\tan^{-1}(L^{1/2})$. For angles in this range and values of L equal to or greater than 63, the following expression is a very good approximation to V_{nc}.

$$V_{nc} \approx V_c cc^* L = \frac{cc^* N_p}{M} \tag{7.37}$$

Let P_p represent the magnitude squared of a correlation value. Given N_p, the estimated signal-to-noise ratio for

this value is

$$\left(\frac{S}{N}\right)_{pk} = \frac{MP_p}{cc^* N_p} - 1 \qquad (7.38)$$

Again, the -1 is usually dropped and the signal-plus-noise ratio is accepted as a good approximation to the signal-to-noise ratio.

For c =-1/2, M = 4, the estimated signal-to-noise ratio is then

$$\left(\frac{S}{N}\right)_{pk} \approx \frac{16P_p}{N_p} \qquad (7.39)$$

Generally, as part of the on-line processing, values of S_p, N_p, P_p, and C_p are computed and printed out in dB on computer's log. The peak correlation value is used to compute P_p. If $10 \log_{10}(\frac{M}{cc^*})$ is added to the computed P_p dB value prior to printing, the peak signal-to-noise ratio in dB can be obtained directly by differencing the printed P_p and N_p values. The same is true for the signal and carrier signal-to-noise ratio except that no adjustments are required.

7.5 Guarding Against Overflow

The magnitudes of the values generated in the course of forming the signal enhancing circular sum and in the removal of the time spreading generally exceed the word size of 12 and 16 bit computers. The most direct solution to this problem is the use of multiple precision arithmetic. In order to keep memory requirements and processing overhead small, one hopes to be able to use no greater than double precision for most calculations. If the scaling of intermediate values is required, the scaling must be done in

such a way as to avoid adding a bias to the scaled values. See Section 3.4.5 for a discussion.

Write the signal component of a demodulate corresponding to a binary one in the time spreading sequence as

$$K(\cos\theta+j\sin\theta) \tag{7.40}$$

and the corresponding signal component corresponding to a binary zero as

$$K(\cos\theta-j\sin\theta) \tag{7.41}$$

K is complex valued.

The maximum signal component value in the circulating sum is

$$NK(\cos\theta\pm j\sin\theta) \tag{7.42}$$

which has magnitude.

$$N|K| \tag{7.43}$$

The value of the grand sum used in the time spread removing correlation is

$$NK(L\cos\theta+j\sin\theta) \tag{7.44}$$

This can be written as

$$LNK\cos\theta(1+j\frac{\tan\theta}{L}) \tag{7.45}$$

For $L\geq 63$ and θ ranging from 0 radius to $\tan^{-1}(L^{1/2})$, ignoring the imaginary part of the above expression results

in less than a 1% error in its magnitude. The approximate grand sum magnitude is then

$$LN|K|\cos\theta \tag{7.46}$$

The maximum correlation sum value prior to using the weighted grand sum value to correct the mean value is

$$(\frac{L+1}{2})NK(\cos\theta+j\sin\theta) \tag{7.47}$$

whose magnitude is

$$(\frac{L+1}{2})N|K| \tag{7.48}$$

From Section 4.4.3, the value of f_1 is used to make the bias correction, assuming a CPM based transmission and local representation parameters c = -1/2, d = 1/2, is

$$f_1 = -\frac{(L+1)(L-\tan\theta)}{2(L^2+\tan^2\theta)} = -\frac{(L+1)}{2L}\left(\frac{1-j\frac{\tan\theta}{L}}{1+\frac{\tan^2\theta}{L^2}}\right) \tag{7.49}$$

$$\simeq -\frac{1}{2} \tag{7.50}$$

The value of the correlation peak after making the bias correction is then approximately

$$j(\frac{L+1}{2})NK\sin\theta \tag{7.51}$$

which has magnitude

$$(\frac{L+1}{2})N|K||\sin\theta| \tag{7.52}$$

For $\theta = 75^\circ$ (the angle used in the Tomography

transmission), $\cos\theta \approx 0.26$, $\sin\theta \approx 0.97$. The largest magnitude encountered in going from the circulating sum to the multipath arrival structure is thus

$$(\frac{L+1}{2})N|K| \tag{7.53}$$

Assuming a 12 bit word size, the largest positive double precision value is $2^{23}-1 \approx 8.4 \times 10^6$. If no scaling is performed, the largest magnitude that the signal component can safely take on is

$$|K| < \frac{8.4\times10^6}{(\frac{L+1}{2})N} \tag{7.54}$$

The Tomography parameter values were L = 127 and N = 22 yielding the restriction

$$|K| < 5958 \tag{7.55}$$

In order to avoid overflow in the calculation of the multipath structure, the signal component in the demodulate values is restricted to being less than about 13 bits in magnitude. Otherwise the circulating sum buffer values must be scaled with the appropriate corrections being added to the computed dB levels.

Overflow can also occur during the signal energy and the carrier energy calculations. First consider the signal power computation.

The signal power is estimated using the sum of the squares of the ML circulating sum bins. The resulting magnitude squared is

$$MLN^2|K|^2 \tag{7.56}$$

Since the circulating sum values are double precision, the accumulation of the squared values should be performed using a four precision accumulator. Doing this places the restriction on |K|

$$|K| < \frac{11.8\times10^{6}}{N(ML)^{1/2}} \tag{7.57}$$

For N = 22, L = 127, M = 4

$$|K| < 23924 \tag{7.58}$$

Thus, for this set of parameter values, the multipath computation is more restrictive on the maximum allowed signal component value than is the S_p power calculation.

The S_p measurement can be used as an aid in setting the demodulator gain. A value of |K| = 5958 corresponds to an S_p value of 129.4 dB. If the input gain is adjusted so that S_p reads below this value, there will not be any overflow problems in the multipath computation. If there are several arrivals, then the total signal energy splits between them and this limit becomes somewhat conservative.

The last calculation in which overflow is a concern is the one for the carrier energy. The carrier line amplitude is estimated by summing the contents of the ML bins in the circulating sum. The resulting complex value is

$$MNK(L\cos\theta+j\sin\theta) \tag{7.59}$$

whose magnitude is approximately

$$MLN|K|\cos\theta \tag{7.60}$$

which for $\theta = 75^{\circ}$ is approximately

$$\frac{MLN|K|}{4} \tag{7.61}$$

If a double precision accumulator is used, the maximum allowable value of |K| is

$$|K| < \frac{33.6\times10^{6}}{MLN} \tag{7.62}$$

which for the Tomography parameter values gives

$$|K| < 3006 \tag{7.63}$$

This is more restrictive than the multipath computation. Rather than further restrict the input magnitude, the carrier line calculation should be made using triple precision.

7.6 Program Testing

Before a program is used to acquire and process data in an experiment, it must be tested to verify that it is working correctly and that there is a reasonable expectation that it is producing valid results. In order to verify the proper operation of programs used to acquire and process multipath data, three test signals are usually used. These are: 1) A noise-free sequence modulated carrier which is phase locked to the A/D conversion clock; 2) A noise-free unmodulated carrier; and 3) A wideband random noise source.

The beamformer is configured to produce "beams" formed using a single input channel. The test signals are individually applied to this channel. The resulting beam output is demodulated and processed. Using the RMS level measurement made by the beamformer's monitor/control processor, the parameter values to be computed by the processing program can be predicted and compared against those values actually produced. If the values don't agree,

either the predictions or the processing program, or both, have to be corrected.

Usually the sequence modulated waveform and the noise waveform are adequate for test purposes. The unmodulated carrier is sometimes useful in tracking down problems already known to exist.

Consider the use of the sequence modulated carrier. The beamformer processor is used to compute the RMS level at the demodulator input in terms of A/D converter quanta. Let K_b be the RMS value measured by the beamformer. The input to the demodulator then consists of samples of

$$2^{1/2}K_b\cos(2\pi f_d t+\theta(t)+\phi) \tag{7.64}$$

where f_d is the carrier frequency, $\theta(t)$ the time spread modulation and ϕ a fixed phase offset.

The positive frequency component in the spectrum is

$$2^{-1/2}K_b\exp\{+j(2\pi f_d t+\theta(t)+\phi)\} \tag{7.65}$$

The negative frequency component is severely attenuated by the complex demodulator and can safely be ignored here.

The beamformer produces samples of the above waveform at a 5 kHz rate. The demodulator multiplies these samples by samples of

$$\frac{2^{15}-1}{2^{15}}\exp\{-j2\pi f_d t\} \tag{7.66}$$

also taken at a 5 kHz rate. A total of N_d products are summed to form a demodulate. The resulting value is approximately

$$2^{-1/2}K_b N_d\exp\{j(\theta_i+\phi)\} \tag{7.67}$$

The θ_i values depend on the modulating sequence. Having θ_i change state during the formation of a demodulate has little effect on the computed power values.

If the demodulator is set to scale by B bits, the resulting demodulate values passed to the processing program are

$$K = 2^{-(B+1/2)} K_b N_d \exp\{j(\theta_i+\phi)\} \tag{7.68}$$

Using this value, predictions of the magnitudes of the S_p, P_k and C_p values can be made using the approximations developed in the previous section for use in checking for the possibility of overflow.

$$S_p \approx 10 \log_{10}(MLN^2|K|^2) \tag{7.69}$$

$$P_k \approx 20 \log_{10}((\tfrac{L+1}{2})N|K\sin\theta|)+10 \log_{10}(M/(cc^*)) \tag{7.70}$$

$$C_p \approx 20 \log_{10}(MLN|K\cos\theta|) \tag{7.71}$$

The expression for P_k assumes that the dB value printed by the program has been adjusted to allow determining the peak signal-to-noise ratio by subtracting the N_p dB value from the P_k dB value.

In addition to the above parameter values, the computed multipath structure should be displayed and examined for anomalies.

The expected results for an unmodulated carrier are

$$K = 2^{-(B+1/2)} K_b N_d \exp\{j\phi\} \tag{7.72}$$

$$S_p \approx 10 \log_{10}(MLN^2|K|^2) \tag{7.73}$$

$$P_k \approx 20 \log_{10}(N|K\tan\theta|/2)+10 \log_{10}(M/(cc^*)) \tag{7.74}$$

$$C_p \approx 20 \log_{10}(MLN|K|) \tag{7.75}$$

The value of P_k is assumed to have been corrected to allow for direct reading of the peak signal-to-noise ratio.

Finally, consider the use of a broadband noise source for program testing. This test waveform is intended to verify the noise level computation. The mean square value (variance) measured by the beamformer monitor computer at the beamformer output is K_b^2. The beamformer input filter is low-pass with a bandwidth of about 550 Hz. The double sided noise power spectral level is approximately

$$\frac{K_b^2}{1100} \tag{7.76}$$

Demodulates are formed using N_d beamformer values which arrive at a 5 KHz rate. The demodulate values are scaled by B bits. The noise bandwidth of the demodulator is approximately $5000/N_d$ Hz and the gain is $2^{-B}N_d$. The variance of the demodulates is then approximately

$$(\frac{K_b^2}{1100})(\frac{5000}{N_d})(2^{-B}N_d)^2 \tag{7.77}$$

The variance of the noise in the circulating sum bins is N times this value

$$V_c \approx \frac{N50K_b^2 2^{-2B}N_d}{11} \tag{7.78}$$

The expected values of the processing parameters are then

$$S_p \approx 10 \log_{10}(MLV_c) \tag{7.79}$$

$$N_p \approx 10 \log_{10}(MLV_c) \tag{7.80}$$

$$C_p \simeq 10 \log_{10}(MLV_c) \tag{7.81}$$

The carrier energy estimate is Chi-square with two degrees of freedom. The value of C_p can be expected to vary significantly compared to the values of S_p and N_p.

CHAPTER 8

SUMMARY

Because of the manner in which the speed of sound in the ocean varies with temperature and pressure there exist a number of paths over which sound can travel between a source and receiver. Using a pulse like signal, these paths can be resolved and identified using the time required for the signal to travel between the source and receiver. Conversely, given a set of measured travel times, it is possible to make statements about the sound speed structure of the portion of the ocean lying between the source and receiver. The Ocean Acoustic Tomography Demonstration Experiment of 1981 demonstrated the feasibility of using travel time measurements as a method of monitoring the thermal structure of a section of the ocean.

The range at which travel time measurements can reliably be made is limited by the background noise level in the ocean and by path stability. This range can be extended by the use of a directional receiver and by placing more energy in the pulse waveform.

If multiple hydrophones are available at the receiver, a method of spatial filtering termed beam forming can be used to enhance the reception. Assuming a plane wave reception, the individual hydrophone outputs are identical except for delay time. By properly delaying the individual outputs, the reception from the signal arrival direction can be enhanced at the expense of the reception of unwanted energy from other directions. In a properly designed system the expected gain in signal-to-noise ratio is on the order

of the number of hydrophone ouputs used in the beam forming process.

The source transducers used to generate the multipath probe transmissions are always limited in the amount of peak power that they can put into the water. Using a linear maximal sequence to phase modulate a carrier the energy contained in one period of the transmitted waveform can be increased by a factor equal to the sequence length, L. Properly processed, the time resolution remains that of a single pulse.

This dissertation describes the design of a time delay beamformer and the associated demodulator/processor system. This equipment is intended for use in acquiring and processing ocean acoustic multipath data. This equipment is unique in its modular organization and its method of demodulation. Emphasis is placed on the development of the theory upon which the design decisions were based. Also included is a description of the use of binary linear maximal sequences to time spread the transmitted energy while retaining the time resolution of a short pulse. Several methods of processing the resulting reception are presented and are shown to be equivalent.

Section 8.1 reviews the relation of the theory developed in this dissertation to the beamformer/demodulator design. Section 8.2 describes the form in which the equipment was implemented. Section 8.3 summarizes the basic ideas behind the use of linear maximal sequences to time spread the transmitted signal energy, and the manner in which the time spreading is removed. Section 8.4 describes possibilities for future work.

The author's primary contributions in this work are:

- Modeling of the effects of beamformer delay time quantization errors as a worst case filter transfer function.

- Comparison of beamformer performance predictions resulting from the use of the spherically isotropic noise and horizontally uniform noise models.
- Recognition that resampling a previously sampled waveform by taking the most recent sample is equivalent to resampling a waveform reconstructed using a zero-order hold. The replicated spectrum caused by the initial sampling is multiplied by a sin(x)/x form weighting prior to the spectral replication caused by the new sample rate. This weighting can be used to minimize the effects of the orginal replications in the spectrum of the resampled waveform.
- Determining the locations and magnitudes of the lines in the spectrum of the resampled waveform due to a single line in the spectrum of the original unsampled waveform.
- Application of a frequency synthesis technique to synthesize sample values of $\exp\{-j2\pi f_d t\}$ for use in a digital complex demodulator.
- Analysis of the effects of amplitude and time quantization errors in the frequency synthesis approach to complex demodulation.
- Design and implementation of the beamformer and signal processor hardware.
- Analysis and comparison of various methods of bias correction that might be used when processing receptions based on the use of linear maximal sequences.
- The method of augmenting a linear maximal sequence by the addition of an extra zero digit.
- The introduction of the non-flatness loss factor to evaluate sequences possessing non-white spectra for use in generating time spread (pulse compression) waveforms.

8.1 The Beamformer and Demodulator

The basic operations involved in beam forming and demodulation are straight forward. The individual hydrophone outputs are delayed and summed in order to form the desired beam output. The beam output is then multiplied by $\exp\{-j2\pi f_d t\}$ and low-pass filtered in order to form the demodulated reception.

The required delays can range from a fraction of a milli-second to times of the order of a tenth of a second or greater. These delays are very difficult to obtain using analog techniques. This leads one to consider digital methods.

In a digital system, the hydrophone outputs are sampled and the sample values are written into a delay memory. Immediately after a set of sample values has been written into memory, a beam output value is formed by reading out of memory the sample values with the required delays and summing. The resulting beam values are multiplied by sample values of $\exp\{-j2\pi f_d t\}$ and low-pass filtered. The errors associated with this approach are due to time and amplitude quantization of the hydrophone outputs and to quantization of the complex exponential values.

If the input waveform sweeps over several A/D converter bits, the amplitude quantization noise can be treated as being spectrally white and independent of the signal. Typically an A/D converter of ten or more bits is used to digitize the hydrophone outputs. The input signal levels are set up so that the background noise level has a standard deviation corresponding to four or more A/D converter bits. In this case, the effects of amplitude quantization can be safely neglected. If the received waveform is so noise free that setting the gain in this manner causes the maximum signal excursion to exceed the A/D converter range, then the gain must be reduced and the quantization noise treated as being the dominant noise source. It is the author's

experience that such noise free signals are only encountered in the laboratory. Indeed, if such noise free waveforms were to be encountered when making an actual propagation measurement it would be an indication that the measurement program was badly designed.

The sample rate used in digitizing the hydrophone outputs must be sufficiently high to allow proper beam formation and to avoid undesired aliasing effects.

In Chapter 5 it is shown that the effects of errors made in the delay values used to form a beam can be viewed as being the same as if the ideal beam output had been low-pass filtered. A worst case transfer function can be hypothesized and used to establish the minimum allowed input sampling rate based on the maximum attenuation level permitted over some frequency range. Requiring no more than a 1 dB fall off over a 700 Hz band requires a minimum sample rate of about 4.7 KHz. This sample rate is almost seven times the specified bandwidth and is sufficiently high to avoid problems due to aliasing. Hardware design considerations resulted in a sample rate of 5 KHz being used.

The carrier frequency of the transmitted waveform is chosen based on the center frequency of the source transducer transfer function. The duration of the digits used in forming the periodic time spread waveform is normally selected to correspond to an integer number of carrier cycles. An integer number of demodulated sample values per digit is required in order to properly remove the effects of the time spreading sequence. Because of this, it is necessary to convert the sample rate at some point in the demodulator from a rate based on that used by the beamformer to a rate related to the carrier frequency of the measurement signal. Generally, it is not possible to specify the value of the carrier frequency in order to simplify the demodulator structure and eliminate the need for sample rate conversion.

In Chapter 3 it is shown that if the multiplication by the sample values of $\exp\{-j2\pi f_d t\}$ and the filtering of the result precedes the sample rate change, then the $\sin(x)/x$ form weighting associated with the resampling effectively attenuates the replications of the input spectrum located at multiples of the beamformer sample rate. In order to exploit this behavior, the demodulator is designed to operate at a rate of 5 KHz and the filter following the multiplier is designed to accommodate the sample rate change at its output.

The demodulator had to be able to operate at any of a large number of carrier frequencies. It was decided that it should be possible to select the demodulation frequency in a range of from 0 Hz to at least 500 Hz in 1 milli-Hz steps. After investigating several alternatives, it was decided to use a table look-up technique based on a method used in direct digital frequency synthesizers in order to generate the required complex exponential samples. A read only memory is used to hold M equally spaced sample values of $\exp\{-j2\pi i/M\}$ where i ranges from 0 thru M-1. Values of $Mf_d/5000$ are added into an accumulator with the sum being formed modulo M. The integer portion of this accumulator is used to address the read only memory. The complex value contained in the addressed location is then used to multiply the current beamformer output value.

The time and amplitude quantization errors associated with this technique are deterministic and periodic. Given values of f_d and M the exact error waveform can readily be computed and its spectrum determined. The spectrum of the multiplier output is the convolution of the spectrum of the sampled beamformer output with the spectrum of the sampled ideal complex exponential plus the spectrum of the error waveform. In Chapter 3 expressions are developed giving the maximum possible spurious line amplitude in terms of the number of bits, B, used to represent the digitized complex exponential values and the number of words, M, in the read

only memory.

The purpose of the filter at the multiplier output is to attenuate the energy outside of a relatively small band of frequencies that, prior to spectral shifting, was centered about the demodulation frequency. This filter has to be easy to reconfigure for use with different output data rates and has to facilitate the change in going from a sample rate based on beamformer requirements to a sample rate based on the signal processing requirements. The filter implemented was the digital equivalent of the analog "integrate and dump" filter. The number of 5 KHz samples in the desired integration period is computed and truncated to an integer value. Starting at a zero crossing of the imaginary part of the synthesized complex exponential, this number of multiplier output values are summed. Once the sum has been formed the resulting value is transferred to the processing computer. The filter then waits for the next zero crossing in order to start the formation of the next demodulate value.

The filter also has the effect of heavily attenuating the spectral component originally centered about $-f_d$ in the input waveform.

A discussion of the proper method to be used in scaling demodulate values is also contained in Chapter 3. Such scaling is often necessary in order to guard against possible arithmetic overflow in later processing steps.

8.2 Hardware Design

The hardware is organized into the following units:

Isolation amplifiers

A/D converter and delay electronics

Beamformer and support processor

Demodulator

Processor

Recording devices

Each hydrophone input is buffered using a commercially available isolation amplifier. This amplifier is used to interrupt ground loops between the hydrophone electronics and the beamformer electronics as well as between individual input signal sources. These amplifiers also protect the hydrophone electronics from the effects of any catastrophic failures in the beamformer.

A major design concern was the organization of the A/D conversion and delay memory electronics. Typically, a multi-channel system such as this would use a track and hold circuit on each input channel in order to acquire simultaneous samples. A single A/D converter with a multiplexer at its input would then digitize from 8 to 16 channels. The resulting values would then be written into a delay memory. This approach requires complicated timing and control electronics. It is also difficult to add and drop channels as needed. The availability of a relatively low cost 12-bit A/D converter allowed the implementation of a more modular structure.

The electronics for a single input channel are contained on a single card. This includes the input low-pass filtering, computer controlled gain, track and hold, A/D converter and the delay memory. These cards are mounted in a card bay which is wired to implement a timing control and data bus. Each card position is associated with an isolation amplifier output.

The A/D-delay bus is generated and controlled by a logic panel contained in a logic bay mounted below the one holding the A/D-delay boards. This bay also contains the beamformer electronics and a microcomputer. The control board generates all of the control signals needed to cause a set of samples to be acquired, A/D converted, and written into the delay memories. The required command signals are

broadcast to all A/D-delay cards which perform these operations in unison. The delay memories are organized as circular buffers of 2048 12-bit words. Samples are written into sequential addresses formed modulo 2048. The control board has a register which is used to keep track of the address containing the most recent sample. Timing is derived from a 5 MHz clock provided by a Rubidium standard.

Devices desiring access to the contents of the delay memories must work thru the control board. Such accesses are made via an access bus also generated by the control board. Up to eight devices can be placed on this access bus. At present only the beamformer and the monitoring computer use this bus. A device signals the control board that it wishes access to the delay memory contents. If there are no other requests of higher priority, an 800 ns time slice is granted. The requesting device then supplies an address relative to the most recent sample, specifies the A/D-delay board to be accessed, and specifies the type of operation desired. A device can read from or write into any given A/D-delay card memory location. The data value is transfered between the device and the selected A/D-delay memory.

The microcomputer used to control the operation of the beamformer and monitor the operation of the delay electronics is based on the INTERSIL IM6100. This is a 12-bit machine that is instruction set compatible with the Digital Equipment Corporation PDP-8/e minicomputer. The processor is mounted on a single printed circuit board along with a hardware bootstrap and the bus generation electronics. Associated with this processor are logic boards containing read/write memory, read only memory and terminal support. The beamformer itself is implemented as a device controlled by the computer. The computer also is supplied with a special interface which gives it access to the A/D-delay bus. This interface is used to allow the processor to check out the operation of the A/D-delay

boards.

The computer control program and the beam information required for a given experiment are contained in a read only memory. When the computer is "booted," the contents of the read only memory are transferred into normal computer memory. The processor then initializes the beamformer by loading the beamformer's control memory with a list of channels, delays and beam numbers. Once initialized, the beamformer operates on its own and interacts with both the delay memories and the demodulator. The processor is then free for use in monitoring the operation of the system.

The beamformer starts a new cycle each time a new sample value has been acquired and written into the delay memories. A list of A/D channels, delay values and beam numbers are read from a memory contained in the beamformer. The required sample values are requested from the delay memories and are summed into the appropriate beam accumulators. Up to eight beams can be formed by a single beamformer. When the list has been exhausted, a flag is set notifying the demodulator that a set of beam outputs is available.

The demodulator is located in a separate logic bay and is implemented as a device contained in a IM6100 based computer. A demodulator is made up of a control/frequency synthesizer wire-wrap panel and from one to eight multiplier/accumulator cards. The processor loads the desired demodulation frequency, the number of beamformer values to be used per demodulate value and the number of bits that the demodulate values are to be scaled by. A direct digital frequency synthesizer is used to address a read only memory containing 2000 samples of one period of $\exp\{-j2\pi t\}$. The addressed complex value is fed to the individual multiplier/accumulator cards. The cards use the complex exponential values to multiply the beamformer output values. Bit serial multipliers are used. The number of products needed to form a set of demodulates is accumulated.

Once a set of demodulate values has been formed, the values are transferred into output buffers. The values are scaled as specified during this transfer. The processor has one full demodulate time in which to extract these values.

A special fast arithmetic device is included in the processor in order to speed up the processing of the demodulated reception. This device is contained on a single wire-wrap panel. The operations supported by this device include those most often encountered in writing signal processing programs. In particular, the device was designed to support various operations involving 24-bit multiplications.

Processed results can be written onto LINCtape, 9-track magnetic tape or 1/4 inch cartridge tape. Special direct memory access interfaces were developed for the LINCtape and 9-track tape systems. The cartridge tape system owned by CEL operates via a standard terminal interface.

8.3 Using Linear Maximal Sequences

A pulse compression waveform is one which has its energy spread over some given time interval but which can be processed to have the time resolution of a much shorter duration waveform containing the same, or nearly the same, energy. Chapter 4 describes how a weighted set of periodic carrier borne pulses can be used to form such a waveform.

A basic periodic pulse waveshape is assumed. The period is divided into L subintervals. A shifted and weighted version of the given waveshape is assigned to the center of each subinterval. It is shown that the spectrum of the resulting waveform is the product of the spectrum of the pulse waveshape and the discrete spectrum of the weighting numbers. If none of the lines in the discrete spectrum have zero magnitude, it is possible to remove the effects of the weighting sequence by dividing the spectrum of the time spread waveform by the spectrum of the weights.

The resulting time waveform is a scaled version of the original pulse waveform. Given a noisy reception based on such a waveform, if the weights have unit magnitude and the associated spectrum is uniform, then the signal-to-noise ratio in the pulse compressed waveform is L times that in an equivalently noisy reception based on a single pulse. If the spectrum of the weights is non-white, the improvement in signal-to-noise ratio will be less than L.

The required processing can also be performed in the time domain by correlating the reception with a suitably chosen set of weights.

Binary linear maximal sequences possess a two level autocorrelation function. The associated spectrum is flat with the exception of the dc line. These sequences can be used to select the weight values used to form a time spread waveform. By suitably choosing the mapping between sequence values and the weight values the amount of energy contained in the dc line can be varied and, if desired, be adjusted to equal that of the surrounding lines.

Because the number of bits contained in one period of a linear maximal sequence is one less than an integer power of two, it is awkward to process waveforms based on these sequences in the frequency domain. One solution to this problem is to augment a linear maximal sequence by adding an additional bit. Adding an additional zero bit results in a sequence containing an equal number of ones and zeros. The cost is a loss in potential signal-to-noise ratio. The results of a survey are presented in Chapter 4 which identify the sequences that provide the best performance when so augmented.

The time domain processing of linear maximal sequence based waveforms can be implemented primarily using add operations with relatively few multiplications being required. One minor problem with this method is that a dc bias is introduced in the result. This bias can be removed using any one of a number of techniques. The four methods

of bias correction felt to be the most reasonable to use are described and shown to yield identical results.

Chapter 7 deals with processing receptions based on the use of linear maximal sequences to time spread the transmitted energy. The signal-to-noise ratio of the processed result can be enhanced by coherently averaging the receptions due to several periods of the transmitted signal. Example program seqments are presented to illustrate the ease with which the time spreading effects of the linear maximal sequence can be removed. Procedures are described for use in estimating the signal-to-noise ratio in the processed reception. The use of laboratory generated signals to verify proper program operation is also described.

8.4 Future Work

The beamformer/demodulator/processor system described in this dissertation was designed and built with the design decisions being based on the theory also presented in this dissertation. The equipment has now been used several times to acquire acoustic propagation data, and it works well. Future efforts should not lie so much in the direction of attempting to refine the theory underlying the hardware but rather in using this hardware to make measurements that will extend our understanding of the propagation of acoustic signals in the ocean.

In the summer and fall of 1983 this equipment will be used to:

Process data acquired as part of a study on the use of linear maximal sequence based signals in a coal mine disaster communication system.

Monitor the installation and operation of an experiment in the Atlantic.

Acquire data for use in designing an acoustic telemetry system for sending data from self contained bouys to shore based receiving sites.

Acquire data in the Pacific to determine the feasibility of using the Hawaii 133 Hz source to make Tomographic measurements of the average heat content of the ocean lying between Hawaii and the west coast of the United States.

In terms of improving hardware capabilities, the following changes and additions would be useful.

Replacement of the INTERSIL IM6100 with the Harris 6120 microcomputer. This would yield about a two times processor speed-up. The two microcomputers are instruction set compatible but possess different bus structures. It should be possible to convert the 6120 bus to one compatible with the existing hardware.

Implement a device to do the summations required for time spread removal. This would be a direct memory device and would allow real time processing of sequence based receptions.

Implement an extended address controller to exploit the address space beyond the basic 32 K presently supported. This could possibly be included on a new processor card using the Harris 6120.

Design and build a new memory card containing 32 K words, at the minimum. The use of CMOS memory integrated circuits would keep power requirements reasonable.

Design and build a communication interface which will support either 2 or 4 terminals.

Add either floppy disk or hard disk storage.

REFERENCES

REFERENCES

1. Ocean Tomography Group: D. Behringer, T. Birdsall, M. Brown, B. Cornuelle, R. Hienmiller, R. Knox, K. Metzger, W. Munk, J. L. Spiesberger, R. Spindel, D. Webb, P. Worcester, and C. Wunch, "A demonstration of ocean acoustic tomography," Nature, Vol. 299, No. 5879, 9 September 1982, pp. 121-125.

2. W. C. Knight, R. G. Pridham and S. M. Kay, "Digital Signal Processing for Sonar," Proceedings of the IEEE, Vol. 69, No. 11, November 1981, pp. 1451-1506.

3. R. J. Urick, Sound Propagation in the Sea, Peninsula Publishing, Los Altos California, 1982.

4. J. L. Spiesberger, R. Spindel, and K. Metzger, "Stability and Identification of Ocean Acoustic Multipaths," The Journal of the Acoustical Society of America, Vol. 67, No. 6, June 1980, pp. 2011-2017.

5. W. H. Munk and C. Wunch, "Ocean acoustic tomography: a scheme for large-scale monitoring," Deep Sea Research, Vol. 26, 1979, pp. 123-161.

6. J. L. Spiesberger, T. G. Birdsall, and K. Metzger, "Acoustic Thermometer Proposal." Submitted to Office of Naval Research, 1983.

7. J. L. Stewart and W. B. Allen, "Pseudorandom Signal Correlation Methods of Underwater Acoustic Research," The Journal of the Acoustical Society of America, Vol. 35, No. 5, May 1963, p. 810.

8. T. G. Birdsall, "MIMI Multipath Measurements," The Journal of the Acoustical Society of America, Vol. 38, No. 5, November 1965, p. 919.

9. G. N. Cederquist, PERSYS: A Collection of FORTRAN Subroutines to Produce Perspective Views of Data Surfaces Defined on Rectangular Grids, Cooley Electronics Laboratory Technical Memorandum No. 113, The University of Michigan, Ann Arbor Michigan, August 1976.

10. R. W. Schafer and L. R. Rabiner, "A Digital Signal Processing Approach to Interpolation," Proceedings of the IEEE, Vol. 61, No. 6, June 1973, pp. 692-702.

11. R. E. Crochiere and L. R. Rabiner, "Interpolation and Decimation of Digital Signals-A Tutorial Review," Proceedings of the IEEE, Vol. 69, No. 3, March 1981, pp. 300-331.

12. V. Manassewitsch, Frequency Synthesizers: Theory and Design, John Wiley & Sons, New York, 1980.

13. T. G. Birdsall and M. Kronengold, "Sound Propagation at 420 CPS, Straits of Florida," Proceedings of the 22nd Navy Symposium on Underwater Acoustics, The University of Texas, Austin Texas, October 1964, pp. 13-22.

14. O. D. Grace and S. P. Pitt, "Quadrature Sampling of High Frequency Waveforms," The Journal of the Acoustical Society of America, Vol. 44, No. 5, November 1968, pp. 1453-1454.

15. O. D. Grace and S. P. Pitt, "Sampling and Interpolation of Bandlimited Signals by Quadrature Methods," The Journal of the Acoustical Society of America, Vol. 48, No. 6, December 1970, pp. 1311-1318.

16. T. G. Birdsall and M. P. Ristenbatt, Introduction to Linear Shift-Register Generated Sequences, Technical Report No. 90, Cooley Electronics Laboratory, The University of Michigan, Ann Arbor Michigan, October 1958.

17. S. W. Golomb, Shift Register Sequences, Holden-Day, San Francisco California, 1967.

18. W. W. Peterson and E. J. Weldon, Jr., Error-Correcting Codes, 2nd Ed., The MIT Press, Cambridge Massachusetts, 1972.

19. F. J. MacWilliams and N. J. A. Sloane, "Pseudo-Random Sequences and Arrays," Proceedings of the IEEE, Vol. 64, No. 12, December 1976, pp. 1715-1729.

20. F. J. Looft, BCSG-74: An Augmented Linear Maximal Shift Register Sequence Generator, Cooley Electronics Laboratory Technical Memorandum No. 110, The University of Michigan, Ann Arbor Michigan, September 1975.

21. K. Metzger and L. W. Von Wald, BCSG-76 Signal Source Documentation, Cooley Electronics Laboratory Technical Memorandum No. 114, The University of Michigan, Ann Arbor Michigan, March 1977.

22. K. Metzger and L. W. Von Wald, The 1979 Towed Source Experiment Signal Generator, Cooley Electronics Laboratory Technical Memorandum No. 116, The University of Michigan, Ann Arbor Michigan, December 1979.

23. H. A. DeFerrari and Lan Nghiem-Phu, "Scattering function measurements for 7-NM propagation range in the Florida Straits," The Journal of the Acoustical Society of America, Vol. 56, No. 1, July 1974, pp. 47-52.

24. R. C. Spindel, "An Underwater Acoustic Pulse Compression System," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-27, No. 6, December 1979, pp. 723-728.

25. G. L. DeMuth, "Frequency Domain Beamforming Techniques," IEEE International Conference on Acoustics, Speech and Signal Processing, Volume 1, 1977, pp. 713-715.

26. A. Papoulis, Signal Analysis, McGraw-Hill, New York, 1977.

27. K. Metzger, Some Aspects of the Weiner Noise Model, Cooley Electronics Laboratory Technical Memorandum No. 111, The University of Michigan, July 1976.

28. S. A. Tretter, Introduction to Discrete-Time Signal Processing, John Wiley & Sons, New York, 1976.

29. M. Abramowitz, and I. Stegun, ed., Handbook of Mathematical Functions with Formulas, Graphs and Tables, National Bureau of Standards, U. S. Government Printing Office, Washington, 1970.

30. A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.

31. A. Papoulis, The Fourier Integral and its Applications, McGraw-Hill, New York, 1962.

32. K. Rektorys, ed., Survey of Applicable Mathematics, The M.I.T. Press, Cambridge, 1969.

33. B. Elspas, "The Theory of Autonomous Linear Sequential Networks," IRE Transactions on Circuit Theory, Vol. CT-6, No. 1, March 1959.

34. C. C. Hoopes and R. N. Randall, Study of Linear Sequence Generators, Cooley Electronics Laboratory Technical Report No. 165, The University of Michigan, June 1966.

35. R. W. Marsh, Table of Irreducible Polynomials over GF(2) through Degree 19, October 24, 1957.

36. K. Metzger, Jr. and R. J. Bouwens, An Ordered Table of Primitive Polynomials Over GF(2) of Degrees 2 Through 19 for Use With Linear Maximal Sequence Generators, Cooley Electronics Laboratory Technical Memorandum No. 107, The University of Michigan, July 1972.

37. C. W. Horton, Sr., Signal Processing of Underwater Acoustic Waves, United States Government Printing Office, Washington, 1969.

38. V. C. Anderson, "Digital Array Phasing," The Journal of the Acoustical Society of America, Vol. 32, No. 2, July 1960.

39. R. G. Pridham and R. A. Mucci, "Digital Interpolation Beamforming for Low-Pass and Bandpass Signals," Proceedings of the IEEE, Vol. 67, No. 6, June 1979, pp. 904-919.

40. R. G. Pridham and R. G. Mucci, "Shifted Sideband Beamformer," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-27, No. 6, December 1979, pp. 713-722.

41. S. P. Pitt, W. T. Adams, and J. K. Vaughn, "Design and implementation of a digital phase shift beamformer," The Journal of the Acoustical Society of America, Vol. 64, No. 3, September 1978, pp. 808-814.

42. D. E. Dudgeon, "Fundamentals of Digital Array Processing," Proceedings of the IEEE, Vol. 65, No. 6, June 1977, pp. 898-904.

43. R. J. Urick, Principles of Underwater Sound, 2nd ed., McGraw-Hill, New York, 1975.

44. G. R. Fox, "Ambient-Noise Directivity Measurements," The Journal of the Acoustical Society of America, Vol. 36, No. 8, August 1964, pp. 1537-1540.

45. E. H. Axelrod, B. A. Schoomer, and W. A. Von Winkle, "Vertical Directionality of Ambient Noise in the Deep Ocean at a Site near Bermuda," The Journal of the Acoustical Society of America, Vol. 37, No. 1, January 1965, pp. 77-83.

46. Intersil IM6100 CMOS 12 Bit Microprocessor, Intersil, Cupertino, 1978.

47. H. Eshraghian, Programmer for Intel's 2716, 2732, 2732A and 2758 EPROMS, Cooley Electronics Laboratory Technical Memorandum No. 117, The University of Michigan, August 1981.

48. L. Conolly, DMROM Program Documentation, Cooley Electronics Laboratory Internal Memorandum, The University of Michigan, 1981.

49. K. Metzger, PDS a Program Development System for the PDP-8 and MP12 Minicomputers 2nd ed., WVM Systems, Ann Arbor, 1977.

50. S. M. Flatte, ed., Sound transmission through a fluctuating ocean, Cambridge University Press, Cambridge, 1979.

DISTRIBUTION LIST

	No. of Copie
Office of Naval Research 800 North Quincy Street Arlington, Virginia 22217 Attn: Code 425 OA Code 425 UA	 1 1
Director Naval Research Laboratory Technical Information Division Washington, D.C. 20375	2
Director Office of Naval Research Branch Office 536 South Clark Street Chicago, Illinois 60605	1
Commander Naval Surface Weapons Center Acoustics Division White Oak Silver Spring, Maryland 20910 Attn: Dr. Zaka Slawsky	 1
Commander Naval Ocean Systems Center San Diego, California 92132 Attn: Dr. Dan Andrews Dr. Henry Aurand	 1 1
Chief Scientist Navy Underwater Sound Reference Division P.O. Box 8337 Orlando, Florida 32806	1
Officer in Charge New London Laboratory Naval Underwater Systems Center New London, Connecticut 06320 Attn: Dr. A. Nuttall Dr. D. M. Viccione	 1 1
Commander Naval Air Development Center Warminster, Pennsylvania 18974 Attn: Unclassified Library	 1
Superintendent Naval Postgraduate School Monterey, California 93940 Attn: Unclassified Library	 1

Commanding Officer
Naval Coastal Systems Laboratory
Panama City, Florida 32401
Attn: Unclassified Library 1

Commanding Officer
Naval Underwater Systems Center
Newport, Rhode Island 02840
Attn: Unclassified Library 1

Superintendent
U.S. Naval Academy
Annapolis, Maryland 21402 1

Commanding Officer
Naval Intelligence Support Center
4301 Suitland Road
Suitland, Maryland 20390
Attn: Dr. Johann Martinek 1
Mr. E. Bissett 1

Officer in Charge
Annapolis Laboratory
Naval Ship Research & Development Center
Annapolis, Maryland 21402 1

Commander
Naval Sea Systems Command
Washington, D. C. 20362
Attn: Code SEA 037 1
Mr. Carey Smith, SEA 06H1 1
Unclassified Library, SEA 03E 1

Commanding Officer
Fleet Numerical Weather Central
Monterey, California 93940 1

Defense Documentation Center
Cameron Station
Alexandria, Virginia 22314 5

Chief of Naval Materials
2211 Jefferson Davis Highway
Crystal Plaza #5
Arlington, Virginia 20360
Attn: Director of Navy Laboratories 1

Commander
Naval Electronic Systems Command
Washington, D.C. 20360
Attn: CDR. A. Miller, NAVELEX 320 1

Commander
David W. Taylor Naval Ship
Research & Development Center
Bethesda, Maryland 20084
Attn: Mr. Craig Olson 1
Unclassified Library 1

Chief of Naval Operations
Department of the Navy
Pentagon Room 4D518
Washington, D.C. 20350
Attn: Capt. A. H. Gilmore 1

Dr. Melvin J. Jacobson
Rensselaer Polytechnic Institute
Troy, New York 12181 1

University of Miami
Rosenstiel School of Marine
and Atmospheric Sciences
4600 Rickenbacker Causeway
Miami, Florida 33149
Attn: Dr. Harry DeFerrari 1
Dr. C. N. K. Mooers 1

General Electric Company
P.O. Box 1088
Schenectady, New York 12301
Attn: Dr. Charles Stutt 1
Dr. Thomas G. Kincaid 1

Dr. Stephen Wolff
John Hopkins University
Baltimore, Maryland 21218 1

Dr. M. Weinstein
Underwater Systems, Inc.
8121 Georgia Avenue
Silver Spring, Maryland 20910 1

Applied Research Laboratories
University of Texas at Austin
P.O. Box 8029
10000 FM Road 1325
Austin, Texas 78712
Attn: Dr. Lloyd Hampton 1

Woods Hole Oceanographic Institution
Woods Hole, Massachusetts 02543
Attn: Dr. R. Spindel 1

Mr. John Bouyoucos
Hydroacoustics, Inc.
321 Northland Avenue
P.O. Box 3813
Rochester, New York 14610 1

Col. Ronald J. Carpinella
F.T.D.
Wright-Patterson AFB
Ohio 45433 1

Atlantic Oceanographic
& Meteorological Laboratories
15 Rickenbacker Causeway
Miami, Florida 33149
Attn: Dr. John Proni 1

Dr. David Middleton
127 East 91st Street
New York, New York 10028 1

Dr. Donald W. Tufts
University of Rhode Island
Kingston, Rhode Island 02881 1

Dr. Loren Nolte
Duke University
Department of Electrical Engineering
Durham, North Carolina 27706 1

Mr. S. W. Autrey
Hughes Aircraft Company
P.O. Box 3310
Fullerton, California 92634 1

Dr. Thomas W. Ellis
Texas Instruments, Inc.
13500 North Central Expressway
Dallas, Texas 75231 1

Mr. Robert Swarts
Applied Physics Laboratory
University of Washington
1013 Northeast Fortieth Street
Seattle, Washington 98195 1

Mr. Beaumont Buck
Polar Research Laboratory
123 Santa Barbara Avenue
Santa Barbara, California 93101 1

Prof. Stuart Schwartz
Princeton University
Department of Electrical Engineering
Brackett Hall, Engineering Quad
Princeton, New Jersey 08540 — 1

Dr. C. V. Kimball
Schlumberger Doll Research Center
P.O. Box 307
Ridgefield, Connecticut 06877
*****Relocated until Summer 1986
EPS
Schlumberger
26 Rue de la Cavee
92140 Clamart, FRANCE — 1

Dr. C. Spofford
Science Applications, Inc.
1710 Goodridge Drive
P.O. Box 1303
McLean, Virginia 22102 — 1

Dr. Walter Munk
IGPP, LaJolla Lab
Scripps Institute of Oceanography
University of California at San Diego
LaJolla, California 92093
Attn: Dr. Walter Munk — 1
Dr. Peter Worcester — 1

Dr. Stanley Adams
Harris Corporation
P.O. Box 37
Melbourne, Florida 32902 — 1

Dr. Ronald L. Spooner
Planning Systems Incorporated
7900 Westpark Drive
McLean, Virginia 22101 — 1

Dr. Victor Anderson
Applied Physics & Information Sciences
Muir College
University of California at San Diego
P.O. Box 109
La Jolla, California 92037 — 1

Dr. William Jobst
U.S. Naval Oceanographic Office
NSTL Station, Code 3443
Bay St. Louis, Mississippi 39522 — 1

Dr. John Spiesberger
Bigelow 208
Woods Hole Oceanographic Institution
Woods Hole, Massachusetts 02543 1

Dr. Gerald Mohnkern
Naval Ocean Systems Center (ARC)
San Diego, California 92152 1

Cooley Electronics Laboratory
The University of Michigan
Ann Arbor, Michigan 48109 25

zeros into an n digit run and 2) Determining the optimal location. The results of this study are tabulated below for n=5 thru 10

Table 4.2. Best Augmented LMS NFLF Values for n=5,10

n	law	n-zero NFLF	law	max NFLF
5	57	0.76	57	0.87
6	147	0.70	103	0.74
7	211	0.64	357	0.77
8	537	0.64	435	0.67
9	1257	0.59	1175	0.67
10	3337	0.54	3117	0.62

The effect on signal-to-noise ratio in dB caused by augmenting is found by taking $10\log_{10}$ of the NFLF number. This does not include the effect of power splitting between the carrier line and the sidebands.

4.3 Time Spread Removal by Cross-Correlation

This section continues the development of correlation processing for time spread removal that was introduced at the end of Section 4.1. The methods outlined here are specific to the use of linear maximal sequences introduced in Section 4.2.

The time spread waveform is a weighted sum of time shifts of a basic pulse waveshape. A total of L pulses with separation T seconds are combined to form the transmitted waveform, r(t).

$$r(t) = \sum_{i=0}^{L-1} w_i p(t-iT) \qquad (4.68)$$

The received waveform is r(t) plus additive noise n(t). This will be written

$$r_n(t) = r(t) + n(t) \qquad (4.69)$$