MODULATION AND DEMODULATION OF RF SIGNALS BY BASEBAND PROCESSING

By

JORGE A. CRUZ-EMERIC

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL
OF THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1976

DEDICATION

The author proudly dedicates this dissertation to his wife, Sara Hilda Quiñones de Cruz.

ACKNOWLEDGEMENTS

The author wishes to acknowledge his chairman, Professor L. W. Couch, for his advice, sincere cooperation, and encouragement. Thanks are given to other members of his supervisory committee and of the staff of the Department of Electrical Engineering for their comments and suggestions.

Thanks are given to his wife, Sara, for her patience and inspiration. She also had the difficult task of typing the final manuscript.

The author is indebted to the University of Puerto Rico, Mayaguez Campus, and the University of Florida for their financial support during his graduate studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT OF DISSERTATION PRESENTED TO
THE GRADUATE COUNCIL OF THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

MODULATION AND DEMODULATION OF RF SIGNALS BY BASEBAND PROCESSING

By

Jorge A. Cruz-Emeric

August 1976

Chairman: Leon W. Couch, Ph.D.
Major Department: Electrical Engineering

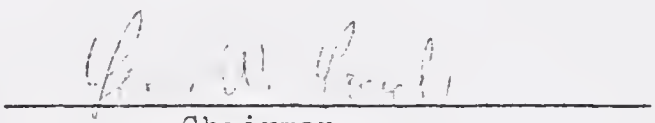
A technique for modulating a carrier by a variety of modulation laws is analyzed. The technique is based on the separate generation of the complex envelope function by a baseband preprocessor. The complex envelope is then used to modulate the carrier. Since the complex envelope function depends exclusively on the modulation law and on the message signal, the baseband preprocessor can be programmed to obtain the desired modulation law. Direct generation of complex valued signals is not physically possible so various alternatives exist to realize the modulation process, but it was found that only three are of practical interest.

All systems considered were a combination of two modulators and two baseband preprocessors. The modulator combinations were an amplitude modulator with a phase modulator, a pair of balanced modulators, and a pair of phase modulators. These configurations were based on the possible representa-

tions of a complex function in terms of real-valued components. Equations were obtained for the complex envelope components and their corresponding autocorrelation functions. The bandwidth required by these components were analyzed to determine what specifications the modulators must meet. Using the digital computer, graphical results were obtained which give the bandwidth for the component as a function of the modulation law parameters.

The modulation laws considered were amplitude modulation, double-sideband linear modulation, single-sideband linear modulation, phase modulation, frequency modulation, compatible single-sideband amplitude modulation, compatible single-sideband phase modulation, and compatible single-sideband frequency modulation.

Also studied was the possibility of demodulating carriers by recovering the complex envelope of the modulated carrier and then extracting the message signal with a baseband post-processor. It was found that only two systems are practical. These were a combination of an envelope detector and a phase detector; and the combination of two homodyne detectors. The latter was found to be better.



Chairman

CHAPTER I
INTRODUCTION

1.1 Problem Statement

Modulation is defined as the systematic alteration of a carrier wave in accordance to the message to be transmitted [1]. The modulation law is the mathematical rule used to alter the carrier properties. The carrier and the modulation law are selected to make better use of the transmission medium [2]. An optimum choice of a carrier waveform and a modulation law for every possible situation does not exist since there are many alternatives.

Each modulation law requires an specific modulator circuit. This is a problem if it is necessary to have a transmitter capable of operating with various modulation laws without drastic modifications. A similar problem exists for the receiver. This dissertation proposes a solution to these problems with the concepts of universal modulation and universal demodulation.

Universal modulation and universal demodulation are defined as techniques used to modulate and demodulate carriers following different modulation laws with a two-stage process that separates the modulation process into a part that depends exclusively on the modulation law and a part that depends only on the carrier waveform. The block diagrams of the proposed universal transmitter and the universal receiver are shown in

Figures 1.1 and 1.2 respectively.

The baseband preprocessor in the transmitter modifies the input message to produce a signal that if applied to the carrier modulator produces a modulated carrier that follows a prescribed modulation law. The carrier modulator is independent of the modulation law. The carrier demodulator and the baseband postprocessor perform the inverse operations at the receiver.

1.2 Historical Background

This section presents a literature review of previous work done by other researchers that is relevant to the research problem.

The concept of complex envelope was popularized by Dugundji [3] and Bedrosian [4] as a technique to analyze modulated carriers. The relationship between the complex envelope and the complex variable theory was studied in detail by Bedrosian [4], Voelcker [5, 6], and Lockhart [7]. The mathematical form of the complex envelope is well known for most types of analog modulation. The idea of combining amplitude and phase modulation was used by some authors [4, 8, 9, 10, 11] to study new forms of modulation.

The properties of the most common types of modulation laws are well known and are available from most communications theory textbooks [12, 13, 14, 15]. The newer types of modulation, like compatible single sideband modulation (CSSB), have been studied in detail by Glorioso and Brazeal [16], Mazo and Salz [17], Kahn and Thomas [18], Couch [19, 20], and others.

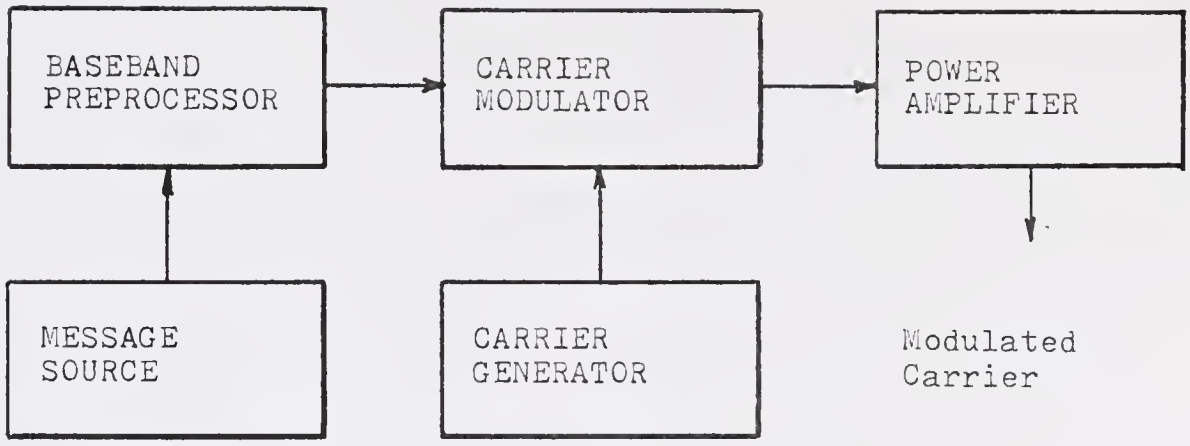


Figure 1.1. The proposed universal transmitter.

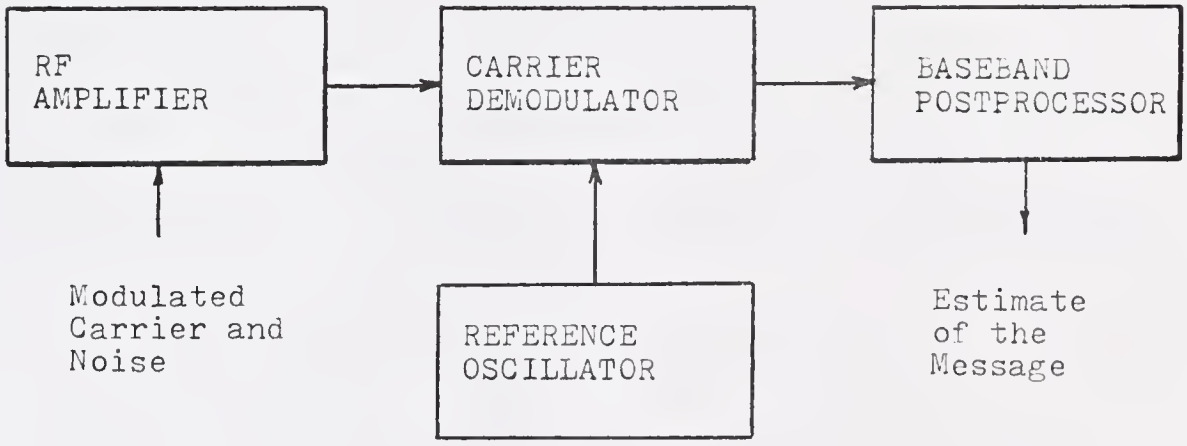


Figure 1.2. The proposed universal receiver.

There is still some controversy among researchers about the virtues and flaws of CSSB. The transmission bandwidths required for distortionless transmission (or with tolerable distortion) for the modulated carrier are well known properties. It is claimed that CSSB modulated carriers can occupy more bandwidth than their double-sided counterparts and that any bandwidth economy depends on the modulation parameters. Also known is the performance of these modulated carriers in the presence of noise. The CSSB modulated carriers have a poorer signal-to-noise ratio when detected with the receiver for which they were designed to be compatible [20].

The idea of designing a modulator capable of following various modulation laws have been proposed by Voelcker [5], Lockhart [7], Couch [19], and Thomas [14]. Voelcker and Lockhart proposed a combination of a phase modulator with an amplitude modulator. Couch proposed a combination of two balanced modulators to obtain different modulated carriers with the same general structure. Meewezen [21] proposed the independent modulation of the sidebands. Cain [22] proposed modifications to a SSB transmitter to obtain CSSB modulated carriers. Leuthold and Thoeny [23] proposed the use of a two dimensional semirecursive filter to obtain some forms of modulation.

The idea of a universal receiver has been suggested by those researchers working on the transmitter problem. Most effort has been channeled toward the development of optimum receivers for specific modulated carriers under specific

restrictions [24-37] rather than to obtain generalized receivers. Nonoptimum receivers are well known and are discussed in many communications theory textbooks.

1.3 Research Areas

Although the idea of universal modulation and demodulation is not new, little attention has been paid to the feasibility of such structures. This dissertation studies the posible structures for the universal transmitter and the physical limitations imposed on the realizations.

To achieve this goal, it is necessary to study the properties of the complex envelope components as related to the different modulation laws, which to the author's knowledge, have not been studied before. This is the subject of Chapter III.

The universal transmitter is covered by Chapter IV. Here various system block-diagrams are proposed as universal modulators and their bandwidth and parameter constraints are studied to determine their feasibility as practical systems.

The receiver case is the subject of Chapter V. The receiver is studied as an extension of the transmitter analysis and general requirements are stated for the receiving problem.

Finally, Chapter VI states the conclusions and highlights the important results obtained throughout the dissertation.

CHAPTER II
MATHEMATICAL PRELIMINARIES

2.1 Introduction

Any cosinusoidal carrier wave is completely described by its amplitude and the absolute phase angle[*]. Consequently, there are only three distinct ways to modulate a carrier: (1) modulate only the amplitude, (2) modulate only the phase angle, and (3) modulate the amplitude and the phase angle simultaneously. The modulated carrier is described by the general expression

$$y(t) = a(t)\cos[\theta(t)], \tag{2-1}$$

where a(t) represents the carrier amplitude, $\theta(t)$ represents the carrier absolute phase angle, and y(t) is the modulated carrier. The function a(t) is also known as the real envelope of y(t) or simply the envelope [4]. The absolute phase angle is

$$\theta(t) = w_o t + p(t), \tag{2-2}$$

where the first term depends exclusively on the carrier frequency in radians/seconds and the time, and the last term

---

[*]Polarization modulation by modulating the transversal components of the radiated RF field will not be considered here.

represents relative phase angle with respect to $\cos(w_o t)$ and includes any other phase-angle variations.

The modulated carrier is a minimum-phase signal if

$$-\pi < p(t) \leq \pi \ , \quad \forall \ t, \tag{2-3}$$

and it is defined as nonminimum-phase otherwise.

2.2 The Complex Envelope

It is a well known fact that a cosinusoidal waveform can be described in terms of a complex exponential using Euler's formula:

$$\exp(jw_o t) = \cos(w_o t) + j\sin(w_o t), \tag{2-4}$$

or

$$\cos(w_o t) = \mathrm{Re}\{\exp(jw_o t)\} . \tag{2-5}$$

Define the complex modulated carrier, $z(t)$, as

$$z(t) = a(t)\exp[jw_o t + jp(t)], \tag{2-6}$$

where $a(t)$ and $p(t)$ are real-valued functions. Comparing Equations (2-4) and (2-5) reveals that the complex modulated carrier is related to the physical modulated carrier by

$$y(t) = \mathrm{Re}\{z(t)\} . \tag{2-7}$$

Equation (2-6) can be rewritten as

$$z(t) = v(t)\exp(jw_o t), \tag{2-8}$$

where

$$v(t) = a(t)\exp(jp(t)). \tag{2-9}$$

The function v(t) is defined as the complex envelope of y(t) [4] and depends exclusively on the message signal and the modulation law.

If v(t) and z(t) are square-integrable functions for $-\infty < t < \infty$, the Fourier transforms V(w) and Z(w) are given by

$$Z(w) = \int_{-\infty}^{\infty} z(t)\exp(-jwt)dt, \tag{2-10}$$

and

$$V(w) = \int_{-\infty}^{\infty} v(t)\exp(-jwt)dt. \tag{2-11}$$

Direct substitution of Equation (2-8) into Equation (2-10) yields

$$Z(w) = V(w - w_o). \tag{2-12}$$

This means that the spectrum of the complex modulated carrier is the result of a linear translation of the spectrum of the complex envelope to the carrier center frequency. The spectrum of z(t) is ussually concentrated in the vicinity of w_o; therefore, the spectrum of v(t) is a baseband function because it is concentrated in the vicinity of zero frequency.

2.3 Analytic Functions

A complex function is classified as Analytic* if the real

*The upper case A is used to denote that the function is analytic in the upper-half t plane as opposed to other regions.

and imaginary parts form a Hilbert pair [4]. This means that the complex funtion s(t) is an Analytic function if

$$\mathrm{Im}[s(t)] = H\{\mathrm{Re}[s(t)]\}, \tag{2-13}$$

where

$$H\{r(t)\} = \hat{r}(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{r(u)}{t-u}\, du \tag{2-14}$$

is the Hilbert transform of the argument [12], and Im{·} and Re{·} are the imaginary and real part operators.

If s(t) is an Analytic function, the following properties are true [4]:

1. The Fourier transform of s(t) vanishes for all negative frequencies.
2. The real and imaginary parts have identical autocorrelation functions.
3. The complex function is completely described by either the real part or the imaginary part.

Any complex modulated carrier, z(t), is not necessarily Analytic. The spectrum of most types of modulated carriers is concentrated near the carrier center frequency and gradually drops to zero as the frequency approaches zero, and remains zero for negative frequencies. Under this condition, z(t) can be considered to be approximately Analytic [4].

The complex envelope is Analytic if and only if V(w) is identical to zero for all negative frequencies. This is satisfied whenever the complex modulated carrier is Analytic and if at the same time Z(w) is zero for all the frequencies

in the interval $[0, w_o]$. This type of modulated carrier spectrum is defined as a single-sided spectrum.

There are single-sided spectrums that do not satisfy Bedrosian's definition of Analytic. A simple example is when s(t) in Equation (2-13) has a nonzero mean value; however, this case can be analyzed because the effect of the nonzero mean value can be considered separately.

2.4 Modulation Techniques

Modulated carriers can be classified in terms of: (1) how the modulation law acts on the spectrum of the message, (2) the distribution of the spectrum of the modulated carrier, and (3) how the effect of modulation is observed in the time domain waveform.

A modulation law is defined as linear modulation if the spectrum of the modulated carrier is a translation of the spectrum of the message and if superposition applies. The modulation law is defined as nonlinear otherwise.

The spectrum of the modulated carrier is classified as double-sided if it is nonzero over at least a finite frequency range on both sides of the carrier frequency. The spectrum is classified as single-sided when all the energy (or power) is concentrated in only one side of the carrier frequency.

In the time domain waveform, modulation of the carrier can be accomplished by the systematic alteration of the carrier amplitude, the carrier phase angle, or both the amplitude and phase simultaneously.

The following are the types of modulation to be considered

in this dissertation. In all cases the message signal is assumed to be a real analog signal.

Amplitude modulation (AM). In this case the message signal is carried exclusively by the real envelope of the carrier. It is a translation of the spectrum of the message and the spectrum of a constant to the carrier center frequency. Due to the constant involved in the modulation process, the modulation law is nonlinear. The AM modulated carrier has a double-sided spectrum.

Linear modulation (LM). It is a translation of the spectrum of the modulating signal to the carrier center frequency [1]. This can be a mixture of amplitude and phase modulation and it can be single-sided or double-sided since both sidebands carry the same information. The single-sided version is known as single-sideband linear modulation (SSB-LM) and the double-sided version is called double-sideband linear modulation (DSB-LM). If the DSB-LM modulated carrier is linearly filtered by an asymmetrical bandpass filter centered around the carrier frequency, the output is classified as vestigial sideband linear modulation (VSB-LM).

Angle modulation (ØM). The information is carried by the phase angle of the carrier [15] and it can be recovered from the zero crossings of the carrier. It is nonlinear. There are two main types, frequency modulation (FM) and phase modulation (PM). In FM the modulating signal is proportional to the derivative of the phase angle, while in PM the modulating signal is proportional to the phase angle itself.

Compatible single sideband (CSSB). The objective of CSSB modulation is to have a modulated carrier with a single-sided spectrum while at the same time retain compatibility at the receiving end with the common types of double-sided demodulators. It is a mixture of amplitude and angle modulation and is nonlinear. There are various alternatives to achieve CSSB [4, 10, 11, 22, 32, 33]. This dissertation restricts its attention to CSSB types of modulations whose complex envelope is an Analytic signal [4]. The three types to be considered are CSSB-AM, CSSB-PM, and CSSB-FM.

2.5 Signal Parameters and Properties

This section defines and discusses the properties and parameters that are of interest in this dissertation.

A deterministic signal is completely described by the time-domain waveform or by its Fourier transform if it exists. If the signal is produced by a random process, there are many possible time-domain waveforms. This situation requires a more general description based on the statistical properties of the signal source. The autocorrelation function of the random process is a satisfactory description in most cases, but it may not be appropiate for nonlinear problems.

The autocorrelation function of the random process $s(t)$ is defined as

$$R_s(t,t+\tau) = E\{s(t)s^*(t+\tau)\} \qquad (2\text{-}15)$$

where $E\{\cdot\}$ is the expectation operator and $s^*(t)$ is the

complex conjugate of s(t). The expectation operator is defined* as

$$E\{s_1 s_2^*\} = \iint_{-\infty}^{\infty} s_1 s_2^* \cdot p(s_1, s_2) ds_1 ds_2, \qquad (2\text{-}16)$$

where the variables s_1 and s_2 represent s(t) and $s(t + \tau)$ respectively. The function $p(s_1, s_2)$ is the joint probability density function of the random variables s_1 and s_2.

The autocorrelation function is related to the power spectral density function by the Fourier transform relationship

$$G_x(w) = \int_{-\infty}^{\infty} R_s(\tau) \exp(jw\tau) d\tau, \qquad (2\text{-}17)$$

where

$$R_s(\tau) = \overline{R_s(t, t + \tau)}^{t}. \qquad (2\text{-}18)$$

The right hand side of Equation (2-18) denotes the time average of $R_s(t, t + \tau)$. Equation (2-17) is known as the Wiener and Kinchine theorem.

2.5.1 Bandwidth

Most signal processing problems involve filters. In some cases, the filter is required by the system realization, in others it is present due to the circuit limitations; therefore, it is necessary to determine the bandwidth required by the signal before the system components are specified. There are various definitions for bandwidth. The most frequently used

---

*Assuming a stationary process.

definitions will follow. All definitions are in terms of Hz.

3 dB down bandwidth (B_d). It is defined for a real valued signal as the frequency where the power spectral density function decreases by 3 dB relative to the maximum level. This is the same as

$$G(B_d) = \frac{1}{2} \max[G(f)] . \tag{2-19}$$

m% power bandwidth (B_p). It is defined for a real-valued process as the frequency band that contains m% of the total power in the signal. The mathematical definition is

$$\int_0^{B_p} G(f)df = \frac{m}{100} \int_0^{\infty} G(f)df. \tag{2-20}$$

Equivalent noise bandwidth (B_n). The equivalent noise is the width of a rectangle with a height of G(0) that contains the same area as the power spectral density curve. This can be written as

$$B_n = \frac{2}{G(0)} \int_{-\infty}^{\infty} G(f)df. \tag{2-21}$$

This equation can be expressed in terms of the autocorrelation function

$$B_n = \frac{R(0)}{G(0)} . \tag{2-22}$$

Second moment bandwidth (B_s). It is defined as

$$B_s^2 = \frac{\int_{-\infty}^{\infty} f^2 G(f)df}{\int_{-\infty}^{\infty} G(f)df} , \tag{2-23}$$

or, in terms of the autocorrelation function

$$B_s^2 = \frac{-R''(0)}{(2\pi)^2 R(0)} , \qquad (2\text{-}24)$$

where $R''(\tau)$ denotes the second derivative of $R(\tau)$ with respect to τ. B_s is also known as the root mean squared (RMS) bandwidth.

Equivalent-filter bandwidth (B_f). The equivalent filter bandwidth is best defined by the block diagram of the procedure used to calculate it rather than by a mathematical expression. This procedure is illustrated by Figure 2.1. There are two signal paths, one through an ideal lowpass filter and the other through a delay line that compensates the delay introduced by the filter. The two signals are applied to the same kind of signal processor and the two outputs are compared to determine the error introduced by the filtering action. The signal processor includes all processing done between the point where the bandwidth is measured to the point where the error is measured. Define the equivalent-filter bandwidth as the width of the passband of the ideal lowpass filter that produces a determinate amount of error. The equation is

$$k = \varepsilon\left\{ f[d[z(t)]] \ , \ f[z(t) \otimes h(t)] \right\}, \qquad (2\text{-}25)$$

where $f[\cdot]$ represents the signal processor algorithm, $\varepsilon[\cdot,\cdot]$ is the error calculation operation, $d[\cdot]$ is the time delay operation, $h(t)$ is the impulse response of the ideal lowpass filter, and k is the value of the measure of error. The ideal lowpass filter impulse response is

$$h(t) = \frac{1}{\pi} \cdot \frac{\sin(2\pi B_f t)}{t} . \qquad (2\text{-}26)$$

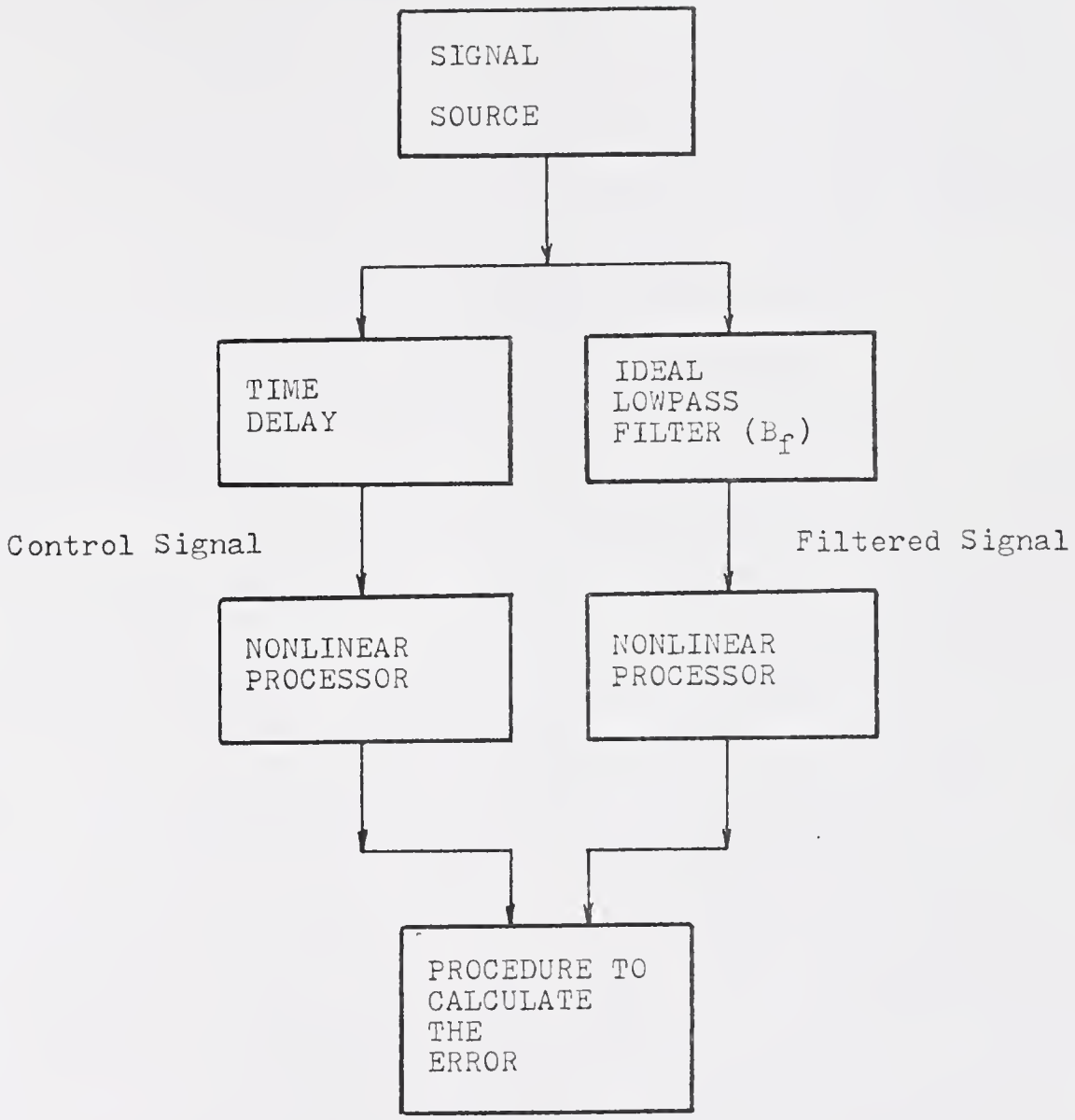


Figure 2.1. Procedure used to calculate the equivalent-filter bandwidth.

The 3 dB bandwidth is often used to specify the useful frequency range of amplifiers and filters. It should not be used to specify the bandwidth of a signal because it does not take into account the signal properties.

The power bandwidth is useful in situations where the evaluation of the power distribution is the prime consideration. It has the disadvantage that Equation (2-20) cannot be solved in terms of B_p except in a few very simple cases; therefore, it must be calculated with a numerical iterative process. It does not specify the actual bandwidth required by a signal unless the power is the only consideration.

The noise equivalent bandwidth and the second moment bandwidth have the advantage of being directly related to the statistical properties of the signal. The disadvantage is that if the signal is subsequently processed by a nonlinear system, the distortion introduced may not be simply related to these bandwidth definitions, although in most cases it is possible to relate the input bandwidth to the output bandwidth. Abramson [34] found that the second moment bandwidth of an arbitrary zero-memory nonlinear transformation of a stationary random process is given by

$$B_{s_s}^2 = \frac{E\left\{g'(x)\cdot x'\right\}}{(2\pi)^2 E\left\{g^2(x)\right\}}, \qquad (2\text{-}27)$$

where $x(t)$ is the input to the nonlinear device, $s = g(x)$ is the nonlinear transformation of the input, $x'(t)$ is the derivative of $x(t)$ with respect to time, and $g'(x)$ is the derivative of $g(x)$ with respect to $x(t)$. For the case where $x(t)$

and $x'(t)$ are statistically independent, Abramson showed that

$$B_{s_s}^2 = \frac{E\left[[g'(x)]^2\right]E\left[x^2(t)\right]}{E\left[g^2(x)\right]} B_{s_x}^2 \qquad (2\text{-}28)$$

where all expectations are with respect to $x(t)$. This relationship shows that B_{s_s} and B_{s_x} may not be simply related. It should be pointed out that the calculation of some of these expectations may be very difficult.

The equivalent-filter bandwidth has the advantage of relating the actual bandwidth requirements of the signal at a point in the system to the criteria used to measure the performance of the system. Therefore, the equivalent-filter bandwidth is the bandwidth required for the actual design at the point where it is measured. It has two disadvantages. The first one is that a closed-form expression is almost impossible to obtain; therefore, it is necessary to simulate the system in order to calculate the bandwidth. This is not a great disadvantage if the reader realizes that all the other definitions may require numerical solutions in complicated cases. The second disadvantage is that the definition is based on a subjective consideration of what is a good performance criterion for a system; therefore, universal acceptance of this definition cannot be expected. However, distortion is precisely what the design engineer will consider as a design criterion; therefore, this definition is acceptable. A similar technique is often used to define the bandwidth of FM signals as a function of distortion [35-37]. Even in the case

of the second moment bandwidth, where an analytical expression may be found, the design engineer cannot use it directly as the design bandwidth without subjective considerations.

2.5.2 Dynamic Range

The voltage (current) dynamic range of an analog signal is defined as the maximum positive voltage (or current) and the maximum negative voltage (or current) that can be found in the signal under consideration. This is important because physical devices and circuits have definite maximum voltage (or current) levels that, if exceeded, may result in improper operation or physical damage to the components.

The numerical dynamic range is defined for digital signals as the maximum (positive or negative) number that can be found in the signal sequence and the minimum number that must be resolved by the computing circuit. These numbers may determine if overflow or underflow will occur and how many bits are necessary to represent the sequence.

2.5.3 Measure of Error

There are two common measures of error. These are the mean-squared error and distortion.

The mean-squared error $\bar{\epsilon}^2$ is defined as [14]

$$E\left[\epsilon^2(t)\right] = E\left\{\left[x(t) - \tilde{x}(t)\right]^2\right\}, \qquad (2\text{-}29)$$

where $x(t)$ is the reference against which $\tilde{x}(t)$ is compared. It involves knowledge of the statistics of the signal so it is a widely accepted theoretical criteria of goodness of a

system. Equation (2-29) can include the effect of noise and other disturbances so the minimization of the mean-squared error is often the objective of optimization problems. The main disadvantage is that it is difficult to measure because any procedure used requires time delay equalization of $\tilde{x}(t)$ and $x(t)$ to guarantee that the measured $E\{\epsilon^2\}$ comes from $x(t)$ and not from the measuring procedure.

The other measure of error is distortion. It is based on the selection of a deterministic test message that is used as the reference against which the distortion is calculated. Distortion is usually specified as harmonic distortion or intermodulation distortion. The <u>total harmonic distortion (THD)</u> is defined for a single tone test signal as

$$\text{THD}(\%) = \sqrt{\frac{\sum_{n=2}^{\infty} P(nf_o)}{P(f_o)}} \times 100\%, \qquad (2\text{-}30)$$

where $P(nf_o)$ is the power of the n-th harmonic of the test signal. The <u>intermodulation distortion (IM)</u> is defined as

$$\text{IM}(\%) = \sqrt{\frac{\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} P(nf_1 + mf_2)}{P(f_1) + P(f_2)}} \times 100\%, \begin{cases} \text{if } n=0, & m\neq 1, \\ \text{if } m=0, & n\neq 1 . \end{cases} \qquad (2\text{-}31)$$

The intermodulation distortion definition takes into account the mixing of the test signal frequency components due to the nonlinearity. It includes the harmonic distortion so it is usually higher than the THD. Since two tones are involved, there are many possible combinations of relative amplitudes and of relative frequency separation. There is no universally

accepted standard. The typical arrangement is a 4:1 amplitude ratio and frequency ratio [38], but there is no technical reason for using it over any other combination. It is advisable to state clearly what is the test signal together with the distortion readings or calculations.

From the engineer point of view, the distortion figures are easier to measure and to understand than the mean-squared error because they are self-normalizing and no phase equalization is required. The disadvantage of the distortion definitions is that they do not take into account the effect of noise and other disturbances.

2.6 Equivalence Between Continuous and Discrete Signals

In some situations it is desirable to replace a continuous-time signal with the sampled or discrete-time version of it. If certain conditions are met, both representations contain the same information.

Let $X_A(w)$ be the Fourier transform of the continuous time signal and $X_D(\exp(jwT))$[*] the discrete Fourier transform (DFT) of the sampled signal. The sampled signal and its DFT are related by

$$x(nT) = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} X_D(\exp(jwT)) \cdot \exp(jnwT)\,dw, \qquad (2\text{-}32)$$

where T is the spacing between samples in the time domain.

---

[*] The notation $X_D(\exp(jwT))$ is a common practice in digital signal processing literature. It is the result of evaluating the z-transform of $x(nT)$ with $z = \exp(jwT)$.

It can be shown [39] that the Fourier transform of the continuous-time signal is related to the DFT of the sampled signal by

$$X_D(\exp(jwT)) = \frac{1}{T} \cdot \sum_{m=-\infty}^{\infty} X_A(w + \frac{2\pi}{T} m), \qquad (2\text{-}33)$$

which is a periodic function in the frequency domain. If the spectrum of the continuous-time signal is absolutely bandlimited to the range $|w| \leq \pi/T$, then

$$X_D(\exp(jwT)) = \frac{1}{T} X_A(w) \qquad (2\text{-}34)$$

provided that $|w| < \pi/T$. If this condition is met, nothing is lost by sampling the continuous-time signal. If the function $X_A(w)$ is not bandlimited to $|w| \leq \pi/T$, $X_D(\exp(jwT))$ will consist of overlapping frequency-shifted spectrums of the continuous signal. This is known as aliasing. The signal can still be sampled, but the sampling rate should be high enough to reduce the effect of aliasing to a tolerable level.

2.7 Summary

This chapter presented the definitions and the concepts that will be used throughout the dissertation. It also discussed the various alternatives to the definition of bandwidth in order to help select the best one for the particular application.

CHAPTER III
THE COMPLEX ENVELOPE

3.1 Introduction

The complex envelope is known to completely describe the modulation process since it contains all the message information. It is reasonable to assume that the generation of the complex envelope is all that is needed to produce a modulated carrier that follows any specified modulation law. This chapter presents a compilation of the complex envelope functions for well known types of modulated carriers and analyzes some properties that have not been considered before in the literature.

Since physical systems cannot handle complex voltages or complex currents, it is necessary to study the representation of the complex envelope in terms of real-valued functions. These components will be used in Chapters IV and V to study the possible structures of the universal transmitter and receiver.

The general properties of the complex envelope are well known and were already presented in Section 2.2. This discussion concentrates on the properties of the components that have received little attention in the literature.

Any complex number can be represented geometrically as a point or a vector in the two-dimensional complex plane defined by the real and imaginary axis. The two most common descriptions of a complex function are in terms of absolute

value and the argument, and in terms of the two quadrature components. These representations are

$$v(t) = a(t)\exp[jp(t)], \tag{3-1}$$

where a(t) is the absolute value and p(t) is the phase angle v(t), and

$$v(t) = i(t) + jq(t), \tag{3-2}$$

where i(t) and q(t) are the quadrature components. In terms of a vector, a(t) is the magnitude or the length of the vector v(t) and p(t) is the phase angle with respect to the real axis. The quadrature components are equivalent to two orthogonal vectors that are the projections of the complex envelope vector along the real and the imaginary axes.

3.2 The Complex Envelope for Different Types of Modulation

The general approach to obtain the equation of the complex envelope for a modulated carrier is to specify certain conditions that the modulated carrier must satisfy. Using available mathematical techniques plus some ingenuity, equations may be found that satisfy all or some of the conditions. These conditions usually require an specific relationship between the message signal and one of the complex envelope components and an specific characteristic of the modulated carrier. For practical purposes, one may like to specify three parameters, namely the magnitude, the phase, and the bandwidth occupancy of the modulated carrier. Unfortunately, there are only two components; therefore, the arbitrary specification of

the three requirements is not possible.

Voelcker [5, 6] developed and Lockhart [7] expanded a theory that serves to study the complex envelope of modulated carriers where the message is a periodic waveform. It is based on the zero-singularity patterns of the z-transform of the complex envelope. By examining and manipulating the location of the zeroes or singularities it is possible to determine if the desired properties can be obtained or if they are not compatible.

This section considers the types of modulation defined in Section 2.4.

3.2.1 Amplitude Modulation (AM)

By definition, amplitude modulation requires the modulated carrier to have a spectrum identical to the translation of the spectrum of the message (except for a carrier line) and the complex envelope to have a magnitude proportional to the message. A complex envelope that satisfies both conditions is

$$v(t) = C[1 + mx(t)], \tag{3-3}$$

provided that

$$mx(t) \geq -1, \tag{3-4}$$

where m is the AM modulation index, x(t) is the message, and C is the unmodulated carrier peak amplitude.

3.2.2 Linear Modulation (LM)

Linear modulation is the linear translation of the spectrum of the message to the carrier center frequency. This requires a complex envelope

$$v(t) = Cx(t). \tag{3-5}$$

The spectrum is double-sided so this is called <u>double-sideband linear modulation (DSB-LM).</u> Comparison of Equation (3-5) with Equations (3-1) and (3-2) reveals that

$$a(t) = C|x(t)|, \tag{3-6}$$

$$p(t) = \begin{cases} 0, & \text{if } x(t) \geq 0 \\ \pi, & \text{if } x(t) < 0, \end{cases} \tag{3-7}$$

$$i(t) = Cx(t), \tag{3-8}$$

and

$$q(t) = 0. \tag{3-9}$$

Equation (3-7) can be rewritten as

$$p(t) = \frac{\pi}{2}\left[1 - \operatorname{sgn}[x(t)]\right], \tag{3-10}$$

where

$$\operatorname{sgn}(u) = \begin{cases} 1, & \text{if } u \leq 0 \\ -1, & \text{if } u > 0, \end{cases} \tag{3-11}$$

The single-sided version is obtained by replacing the message with an equivalent analytic signal that has a spectrum identical to one of the sides of the spectrum of the message. This is the same as requiring the complex envelope to be

$$v(t) = C\left[x(t) + j\hat{x}(t)\right], \tag{3-12}$$

where $\hat{x}(t)$ is the Hilbert transform defined by Equation (2-14). This is known as single-sideband linear modulation (SSB-LM).

The required components are

$$a(t) = C\sqrt{x^2(t) + \hat{x}^2(t)}, \tag{3-13}$$

$$p(t) = \tan^{-1}\left[\frac{\hat{x}(t)}{x(t)}\right], \tag{3-14}$$

$$i(t) = Cx(t), \tag{3-15}$$

and

$$q(t) = C\hat{x}(t). \tag{3-16}$$

3.2.3 Angle Modulation (ØM)

Angle modulation requires that the relative phase angle of the modulated carrier be a linear function of the message. Phase modulation (PM) requires that the relative phase angle, p(t), be directly proportional to the message, while frequency modulation (FM) requires that the message be proportional to the derivative of the phase angle. The required complex envelope is

$$v(t) = C \exp[jr(t)], \tag{3-17}$$

where

$$r(t) = D_p x(t) \text{ for PM}, \tag{3-18}$$

and

$$\frac{dr(t)}{dt} = D_f x(t) \text{ for FM}. \tag{3-19}$$

The constant D_p and D_f are the PM modulation index and the FM frequency deviation constant, respectively

The required complex envelope components are

$$a(t) = C, \tag{3-20}$$

$$p(t) = r(t), \tag{3-21}$$

$$i(t) = C \cos[r(t)], \tag{3-22}$$

and

$$q(t) = C \sin[r(t)]. \tag{3-23}$$

3.2.4 Compatible Single Sideband Modulation (CSSB)

The idea behind compatible single sideband modulation (CSSB) is to obtain a modulated carrier that is compatible with receivers for conventional AM, PM, and FM, while at the same time obtain the bandwidth savings of a single-sided spectrum. This requires the selection of a(t) in the CSSB-AM case, and p(t) in CSSB-PM and CSSB-FM cases. This leaves only one function to specify. This function should result in a single sided spectrum that at the same time is narrower than the spectrum of the conventional case. This may not be possible in all cases.

Several solutions have been proposed for the CSSB-AM signal by Villard [9], Kahn [11], Bedrosian [4], Powers [10], and others. The Villard's system has been shown to be non-bandlimited and not completely single-sided. The Powers' system is absolutely bandlimited and single-sided but it is not

exactly compatible with conventional AM receivers. Kahn's and Bedrosian's systems are very similar, they have single sided spectrums but they have been shown not to be absolutely bandlimited. These systems do show some "effective" bandwidth reduction compared with the double-sided cases but the reduction depends on the message level [7].

Solutions to the CSSB-FM and CSSB-PM cases have been proposed by Bedrosian [4]. Glorioso and Brazeal [16], Mazo and Salz [17], Kahn and Thomas [18], Dubois and Aagaard [40], and Barnard [41] have studied the CSSB-PM and CSSB-FM carrier spectral properties. It is known that CSSB-PM and CSSB-FM require less bandwidth than conventional PM and FM for low modulation levels but that the bandwidth reduction depends on the modulation level.

Bedrosian also showed that if an Analytic periodic signal, s(t), is used to form the complex envelope

$$v(t) = C \exp[s(t)], \tag{3-24}$$

a single-sided complex envelope results. This signal, depends on what characteristics are desired. For CSSB-AM, s(t) is

$$s(t) = \ln[1 + mx(t)] + jH\{\ln[1 + mx(t)]\}, \tag{3-25}$$

provided that Equation (3-4) is satisfied. This results in the following complex envelope components

$$a(t) = C[1 + mx(t)], \tag{3-26}$$

$$p(t) = H\{\ln[1 + mx(t)]\}, \tag{3-27}$$

$$i(t) = C[1 + mx(t)]\cos[p(t)], \tag{3-28}$$

and

$$q(t) = C[1 + mx(t)] \sin[p(t)]. \tag{3-29}$$

The CSSB-PM and CSSB-FM cases require that

$$s(t) = r(t) + j\hat{r}(t), \tag{3-30}$$

where

$$r(t) = D_p x(t) \tag{3-31}$$

for CSSB-PM, and

$$r(t) = D_f \int_{-\infty}^{t} x(u)\,du \tag{3-32}$$

for CSSB-FM. Couch [19] showed that other SSB signals could be obtained by the use of entire functions to obtain analytic complex envelope functions.

The complex envelopes are summarized in Table III-1.

3.3 Autocorrelation Function of the Complex Envelope Components

This section presents the autocorrelation function of the complex envelope and derives the autocorrelation function of the components for different types of modulation laws.

If the autocorrelation function is available, the second moment bandwidth can be calculated using Equation (2-24).

3.3.1 General Expressions for the Autocorrelation Function

The autocorrelation function of the complex envelope can be derived in terms of the a(t) and p(t) functions:

$$R_v(t, t+\tau) = E\left[a(t)a(t+\tau)\exp\left[jp(t) - jp(t+\tau)\right]\right], \tag{3-33}$$

Table III-1. Complex Envelopes

Types of Modulation	Complex Envelope
AM	$C[1 + mx(t)]$, $mx(t) \geqslant -1 \quad \forall\ t$
LM	$Cx(t)$
SSB-LM	$C[x(t) + j\hat{x}(t)]$
PM	$C\ \exp[jD_p x(t)]$
FM	$C\ \exp[jD_f \int_{-\infty}^{t} x(u)du]$
CSSB-AM	$C\ \exp[r(t) + j\hat{r}(t)]$, $r(t) = \ln[1 + mx(t)]$
CSSB-PM	$C\ \exp[jD_p(x(t) + j\hat{x}(t))]$
CSSB-FM	$C\ \exp[jD_f \int_{-\infty}^{t} [x(u) + j\hat{x}(u)]du]$

which can be written as

$$R_v(t,t+\tau) = E\{a(t)a(t+\tau)\cos[p(t) - p(t+\tau)]\} + \\ + jE\{a(t)a(t+\tau)\sin[p(t) - p(t+\tau)]\}. \quad (3\text{-}34)$$

No further simplification is possible unless more is known about the nature of a(t) and p(t). This equation illustrates that it is not possible to separate the effects of the amplitude and phase functions on the modulated carrier complex envelope.

It is possible to define the second moment bandwidth of v(t) in terms of a(t) and p(t). For this purpose, define d(t) as

$$d(t) = v'(t), \quad (3\text{-}35)$$

where v'(t) is the first time-derivative of v(t). Observe that the Fourier transform of d(t) is given by

$$D(w) = jwV(w); \quad (3\text{-}36)$$

therefore, the spectrum of d(t) can be obtained as the output of a filter whose transfer function is

$$H(w) = jw \quad (3\text{-}37)$$

and whose input is v(t). The power spectral density of d(t) can be expressed in terms of the power spectral density of v(t) by

$$G_d(w) = |jw|^2 G_v(w) \quad (3\text{-}38)$$

where $G_d(w)$ and $G_v(w)$ are the power spectral density functions of d(t) and v(t) respectively. Apply the Fourier transform differentiation theorem to obtain the autocorrelation function of d(t) in terms of the autocorrelation function of v(t)

$$R_d(\tau) = -R_v''(\tau). \tag{3-39}$$

Using this equation, Equation (2-24) can be written as

$$B_{s_v}^2 = \frac{E\{|v'(t)|^2\}}{(2\pi)^2 E\{|v(t)|^2\}}. \tag{3-40}$$

Using Equation (3-1), observe that

$$v'(t) = a'(t)\exp[jp(t)] + \\ + j[a(t)p'(t)]\exp[jp(t)] \tag{3-41}$$

so

$$|v'(t)|^2 = [a'(t)]^2 + [a(t)]^2[p'(t)]^2, \tag{3-42}$$

and the second moment bandwidth of v(t) is given by

$$B_{s_v}^2 = \frac{1}{(2\pi)^2} \cdot \\ \cdot \frac{E\{[a'(t)]^2\} + E\{[a(t)]^2[p'(t)]^2\}}{E\{[a(t)]^2\}}, \tag{3-43}$$

or

$$B_{s_v}^2 = B_{s_a}^2 + \frac{1}{(2\pi)^2} \cdot \frac{E\{[a(t)]^2[p'(t)]^2\}}{E\{[a(t)]^2\}}, \tag{3-44}$$

where B_{s_v} is the second moment bandwidth of the complex envelope and B_{s_a} is the second moment bandwidth of the amplitude function. This is similar to a result obtained by Kahn and Thomas [18] and shows that the complex envelope bandwidth is at least as wide as the bandwidth of the a(t) function, and that it depends on both a(t) and p(t).

If the a(t) and the p(t) functions are statistically independent, Equation (3-44) becomes

$$B_{s_v}^2 = B_{s_a}^2 + \frac{1}{(2\pi)^2} E\left\{\left[p'(t)\right]^2\right\} . \qquad (3\text{-}45)$$

Under the same condition, Equation (3-33) becomes

$$R_v(t,t+\tau) = R_a(t,t+\tau)\, E\left\{\exp\left[jp(t) - jp(t+\tau)\right]\right\} . \qquad (3\text{-}46)$$

The autocorrelation function of v(t) can be derived in terms of the quadrature functions,

$$R_v(t,t+\tau) = E\left\{i(t)i(t+\tau)\right\} + E\left\{q(t)q(t+\tau)\right\} + jE\left\{q(t)i(t+\tau)\right\} - jE\left\{i(t)q(t+\tau)\right\}, \qquad (3\text{-}47)$$

or

$$R_v(t,t+\tau) = R_i(t,t+\tau) + R_q(t,t+\tau) + j\left[R_{qi}(t,t+\tau) - R_{iq}(t,t+\tau)\right], \qquad (3\text{-}48)$$

where $R_i(t,t+\tau)$ and $R_q(t,t+\tau)$ are the autocorrelation functions of i(t) and q(t), respectively,

$$R_{qi}(t,t+\tau) = E\left[q(t)i(t+\tau)\right] \tag{3-49}$$

and

$$R_{iq}(t,t+\tau) = E\left[i(t)q(t+\tau)\right]. \tag{3-50}$$

Using Equation (3-39), the second moment bandwidth can be written as

$$B_{s_v}^2 = \frac{+E\left[\left[i'(t)\right]^2 + \left[q'(t)\right]^2\right]}{(2\pi)^2 E\left[i^2(t) + q^2(t)\right]}. \tag{3-51}$$

The autocorrelation functions of i(t) and q(t) are easily related to that of v(t). Observe that i(t) can be written as

$$i(t) = a(t)\cos\left[p(t)\right], \tag{3-52}$$

so

$$i(t) = \frac{1}{2}\left[v(t) + v^*(t)\right]; \tag{3-53}$$

therefore,

$$R_i(t,t+\tau) = \frac{1}{2}\,\mathrm{Re}\left[E\left[v(t)v^*(t+\tau)\right]\right] + \frac{1}{2}\,\mathrm{Re}\left[E\left[v(t)v(t+\tau)\right]\right], \tag{3-54}$$

or

$$R_i(t,t+\tau) = \frac{1}{2}\,\mathrm{Re}\left\{R_v(t,t+\tau) + R_{vv^*}(t,t+\tau)\right\} \quad \text{(3-55)}$$

Since q(t) can be written as

$$q(t) = \frac{1}{2j}\left[v(t) - v^*(t)\right], \quad \text{(3-56)}$$

so following the same procedure,

$$R_q(t,t+\tau) = \frac{1}{2}\,\mathrm{Re}\left\{R_v(t,t+\tau) - R_{vv^*}(t,t+\tau)\right\}. \quad \text{(3-57)}$$

All these equations point out that, in general, knowledge of the autocorrelation function of the complex envelope does not necessarily imply knowledge of the autocorrelation functions of a(t), p(t), i(t), and q(t). The converse is also true. These statements are also valid for the second moment bandwidth.

If a(t) and p(t) are statistically independent, Equations (3-55) and (3-57) can be simplified. Using Equation (3-46), Equation (3-55) simplifies to

$$R_i(t,t+\tau) = \mathrm{Re}\left\{\frac{R_a(t,t+\tau)}{2}\,E\left\{\exp\left[jp(t) - jp(t+\tau)\right]\right\} + \right.$$
$$\left. + \exp\left[jp(t) + jp(t+\tau)\right]\right\}\Big], \quad \text{(3-58)}$$

and Equation (3-57) simplifies to

---

[*] $R_{vv}(t,t+\tau) = R_v(t,t+\tau)$

$$R_q(t,t+\tau) = \mathrm{Re}\left\{\frac{R_a(t,t+\tau)}{2} E\left\{\exp[jp(t) - jp(t+\tau)] + \right.\right.$$

$$\left.\left. - \exp[jp(t) + jp(t+\tau)]\right\}\right]. \qquad (3\text{-}59)$$

3.3.2 Special Case: Analytic Complex Envelope

If the complex envelope is an Analytic signal, by definition

$$q(t) = \hat{i}(t), \qquad (3\text{-}60)$$

$$R_q(t,t+\tau) = R_i(t,t+\tau), \qquad (3\text{-}61)$$

$$R_{iq}(t,t+\tau) = \hat{R}_i(t,t+\tau), \qquad (3\text{-}62)$$

and Equation (3-48) becomes

$$R_v(t,t+\tau) = 2R_i(t,t+\tau) + 2j\hat{R}_i(t,t+\tau), \qquad (3\text{-}63)$$

where $\hat{R}_i(t,t+\tau)$ is the Hilbert transform of $R_i(t,t+\tau)$.

Observe that if Equation (3-60) is valid, then

$$E\{[i'(t)]^2\} = E\{[q'(t)]^2\}; \qquad (3\text{-}64)$$

therefore, Equation (3-52) becomes

$$B_{s_v}^2 = \frac{E\{[i'(t)]^2\}}{(2\pi)^2 E\{i^2(t)\}}. \qquad (3\text{-}65)$$

Remember that i(t) and q(t) have the same autocorrelation functions, so

$$B_{s_v} = B_{s_i} = B_{s_q} \quad . \tag{3-66}$$

A similar simplification of Equation (3-34) is not possible under this condition; however, when v(t) is Analytic,

$$q(t) = a(t)\ \sin p(t)$$

$$= H\{a(t)\ \cos[p(t)]\} \tag{3-67}$$

3.3.3 AM

Assume that the message, x(t), is a zero-mean, unit power, real process with normalized autocorrelation function, $\rho(\tau)$. For the AM case the following autocorrelation functions were obtained using the equations listed on Table III-1[*]

$$R_v(\tau) = C^2[1 + m^2\rho(\tau)], \tag{3-68}$$

$$R_i(\tau) = R_a(\tau) = C^2[1 + m^2\rho(\tau)], \tag{3-69}$$

and

$$R_p(\tau) = R_q(\tau) = 0. \tag{3-70}$$

where it is assumed that m is low enough to prevent frequent violations of Equation (3-4).

---

[*]It should be remembered that these equations assume $|1 + mx(t)| = (1 + mx(t))$.

3.3.4 DSB-LM

The autocorrelation function of the DSB-LM complex envelope is

$$R_v(\tau) = C^2\rho(\tau); \tag{3-71}$$

therefore,

$$R_i(\tau) = C^2\rho(\tau), \tag{3-72}$$

and

$$R_q(\tau) = 0. \tag{3-73}$$

The autocorrelation function for a(t) can be calculated by observing that Equation (3-6) is similar to the output of a full-wave linear rectifier. The autocorrelation function of this system with zero mean gaussian noise (ZMGN) at the input was calculated by Middleton [42]; therefore, adapting that result to this situation

$$R_a(\tau) = \frac{2}{\pi} C^2\left[\rho(\tau)\arcsin\left[\rho(\tau)\right] + \sqrt{1 - \rho^2(\tau)}\right] . \tag{3-74}$$

The autocorrelation function for p(t) is obtained by substituting Equation (3-10) in the definition of autocorrelation function

$$R_p(\tau) = \left[\frac{\pi}{2}\right]^2 E\left\{\left[1 - \mathrm{sgn}(x(t))\right]\left[1 - \mathrm{sgn}(x(t+\tau))\right]\right\} \tag{3-75}$$

If x(t) is a zero mean stationary random process

$$E\left[\operatorname{sgn} x(t)\right] = 0, \tag{3-76}$$

so

$$R_p(\tau) = \left[\frac{\pi}{2}\right]^2 E\left\{1 + \operatorname{sgn}\left[x(t)\right] \cdot \operatorname{sgn}\left[x(t+\tau)\right]\right\}. \tag{3-77}$$

The second term is similar to the autocorrelation function of odd-symmetry limiter. Using published results for this device and ZMGN [42] the autocorrelation of the phase function is

$$R_p(\tau) = \left[\frac{\pi}{2}\right]^2 + \frac{\pi}{2} \arcsin\left[\rho(\tau)\right]. \tag{3-78}$$

3.3.5 SSB-LM

The autocorrelation function of the complex envelope is easily found by comparing Equation (3-12) with Equation (3-65):

$$R_v(\tau) = 2C^2\left[\rho(\tau) + j\hat{\rho}(\tau)\right]. \tag{3-79}$$

and by the definition of an Analytic signal,

$$R_i(\tau) = R_q(\tau) = C^2\rho(\tau). \tag{3-80}$$

The situation for the a(t) and p(t) functions is not as simple because Equations (3-13) and (3-14) involve nonlinearities and memory. Rewrite Equation (3-13) as

$$a(t) = C\sqrt{b(t)}, \tag{3-81}$$

where b(t) is

$$b(t) = x^2(t) + \hat{x}^2(t). \tag{3-82}$$

The new process b(t) is not gaussian. The autocorrelation function of b(t) is

$$R_b(\tau) = E\{b(t)b(t+\tau)\}$$

$$= E\{x^2(t)x^2(t+\tau)\} + E\{x^2(t)\hat{x}^2(t+\tau)\} +$$

$$+ E\{\hat{x}^2(t)x^2(t+\tau)\} + E\{\hat{x}^2(t)\hat{x}^2(t+\tau)\}. \qquad (3\text{-}83)$$

If x(t) is ZMGN, $\hat{x}(t)$ is also ZMGN and

$$E\{x^2(t)x^2(t+\tau)\} = 1 + 2\rho^2(\tau), \qquad (3\text{-}84)$$

$$E\{\hat{x}^2(t)\hat{x}^2(t+\tau)\} = 1 + 2\rho^2(\tau). \qquad (3\text{-}85)$$

$$E\{x^2(t)\hat{x}^2(t+\tau)\} = 1 + 2\hat{\rho}^2(\tau), \qquad (3\text{-}86)$$

and

$$E\{\hat{x}^2(t)x^2(t+\tau)\} = 1 - 2\hat{\rho}^2(\tau). \qquad (3\text{-}87)$$

Equation (3-73) reduces to

$$R_b(\tau) = 4\left[1 + \rho^2(\tau)\right]. \qquad (3\text{-}88)$$

The power spectral density of b(t) is bandlimited for bandlimited x(t), so it is well known that the power spectral density of $\sqrt{b(t)}$ is not absolutely bandlimited [7]. A closed solution for the autocorrelation function of a(t) does not seem possible. Bowen [43] presented a method to calculate the autocorrelation of instantaneous nonlinear devices with non-gaussian noise that yields an infinite series expansion in

terms of the input autocorrelation function. This method can be used if the exact autocorrelation function is known.

An approximation can be obtained for the autocorrelation function of a(t) by observing that x(t) and $\hat{x}(t+\tau)$ are uncorrelated [13] for small τ. This means that a(t) is approximately a two-dimensional chi process. A chi process is defined by

$$r = \left[\sum_{k=1}^{n} u_k^2\right]^{\frac{1}{2}}, \tag{3-89}$$

where u_k represent independent gaussian variables. Miller, et al [44] obtained an expression for the autocorrelation function of a two-dimensional chi process

$$R_r(\tau) = 2R_u(0)\left[1 - \rho_u^2(\tau)\right]^2 \cdot \frac{\Gamma(3/2)}{\Gamma^2(1)} \cdot$$

$$\cdot\ {}_2F_1\left[3/2,\ 3/2;\ 1;\rho_u^2(\tau)\right] \tag{3-90}$$

where $R_u(0)$ is the variance of u_i, $\rho_u(\tau)$ is the normalized autocorrelation function of u, $\Gamma(n)$ is the Gamma function [45] given by

$$\Gamma(n) = \int_0^\infty \exp(-t)t^{(-n)}dt \tag{3-91}$$

and ${}_2F_1(.,\ .;\ .;\ .)$ is a gaussian hypergeometric function given by

$${}_2F_1(a,\ b;\ c;\ z) = \frac{\Gamma(c)}{\Gamma(a)\,\Gamma(c-a)} \int_0^1 t^{a-1}(1-t)^{c-a-1} \cdot$$

$$\cdot\ (1 - zt)^{-b}dt. \tag{3-92}$$

For the case under consideration

$$\Gamma(3/2) = \frac{\sqrt{\pi}}{2}\,, \tag{3-93}$$

$$\Gamma(1) = 1, \tag{3-94}$$

and

$$R_u(0) = C^2 \tag{3-95}$$

so Equation (3-90) becomes

$$R_a(\tau) \simeq \frac{\pi}{2} C^2 [1 - \rho^2(\tau)]^2 \cdot {}_2F_1[3/2,\ 3/2;\ 1; \rho(\tau)] \tag{3-96}$$

The derivation of the autocorrelation function of p(t) is very complicated. The autocorrelation function of the intermediate variable w(t), defined as

$$w(t) = \hat{x}(t)/x(t), \tag{3-97}$$

is calculated in the Appendix A and is listed below:

$$R_w(\tau) = -\hat{\rho}^2(\tau) + \\ - \frac{[1 - \rho^2(\tau) - \hat{\rho}^2(\tau)]}{1 - \rho^2(\tau)} \ln[1 - \rho(\tau)]. \tag{3-98}$$

For a baseband gaussian process bandlimited to W with the autocorrelation function of Equation (3-33) yields

$$R_w(0) = \infty. \tag{3-99}$$

This means that the spectrum is not absolutely bandlimited and requires infinite power. Since w(t) must be applied to the

nonlinear operation $\tan^{-1}(\cdot)$ it is not possible to predict the effect on p(t); however, it is reasonable to expect a very wideband signal.

3.3.6 PM

The autocorrelation function of the PM complex envelope is [13]

$$R_v(\tau) = C^2 \exp\left[-D_p^2[1 - \rho(\tau)]\right] \tag{3-100}$$

provided that x(t) is a stationary ZMGN. Also

$$R_{vv^*}(\tau) = C^2 \exp\left[-D_p^2[1 + \rho(\tau)]\right] . \tag{3-101}$$

The autocorrelation functions of the amplitude and phase functions are

$$R_v(\tau) = C^2 \tag{3-102}$$

and

$$R_p(\tau) = D_p^2 \rho(\tau). \tag{3-103}$$

The autocorrelation function of the quadrature components, i(t) and q(t), can be found by substituting Equations (3-100) and (3-101) into Equation (3-55). The result is

$$R_i(\tau) = \frac{C^2}{2} \exp\left[-D_p^2 (1 - \rho(\tau))\right] + \frac{C^2}{2} \exp\left[-D_p^2(1 + \rho(\tau))\right], \tag{3-104}$$

which can be simplified to

$$R_i(\tau) = C^2 \exp(-D_p^2) \cosh\left[D_p^2 \rho(\tau)\right] . \tag{3-105}$$

Similarly, substituting Equations (3-100) and (3-101) into (3-57) yields

$$R_q(\tau) = C^2 \exp(-D_p^2) \sinh\left[D_p^2 \rho(\tau)\right]. \tag{3-106}$$

3.3.7 FM

If the random process resulting from the integral of x(t) given by Equation (3-19) has zero mean and is stationary, the results derived for PM can be extended to FM and it is only necessary to replace x(t) by $\phi(t)$, where

$$\phi(t) = D_f \int_{-\infty}^{t} x(u)\,du \tag{3-107}$$

with an autocorrelation function of [13]

$$R_\phi(\tau) = \frac{D_f^2}{2} \int_{-\infty}^{\infty} \frac{G_x(w)}{w^2} \exp(jw\tau)\,dw, \tag{3-108}$$

where

$$G_x(w) = \int_{-\infty}^{\infty} \rho(\tau) \exp(-jw\tau)\,d\tau. \tag{3-109}$$

If $\phi(t)$ is ZMGN, the required equations are

$$R_v(\tau) = C^2 \exp\left[-R_\phi(0) + R_\phi(\tau)\right], \tag{3-110}$$

$$R_A(\tau) = C^2, \tag{3-111}$$

$$R_p(\tau) = R_\phi(\tau), \tag{3-112}$$

$$R_i(\tau) = C^2 \exp\left[-R_\phi(0)\right] \cosh(R_\phi(\tau)), \tag{3-113}$$

and

$$R_q(\tau) = C^2 \exp\left[-R_\phi(0)\right] \sinh(R_\phi(\tau)). \tag{3-114}$$

3.3.8 CSSB-AM

A closed form expression for the autocorrelation function of the complex envelope cannot be found in the literature. Preliminary work done by the author reveals that it is necessary to use series expansions that are very difficult to compute and virtually impossible to work with when put into use. That is, no satisfactory result could be obtained for the autocorrelation function of the complex envelope.

The autocorrelation function of the conjugate functions depends on the availability of the result for the complex envelope as shown by Equation (3-65).

The situation for the autocorrelation of the amplitude, a(t), is different. If

$$(1 + mx(t)) \geq 0 \quad \forall t \tag{3-115}$$

the autocorrelation function of a(t) is given by Equation (3-69).

The autocorrelation of the phase function, p(t), is

$$R_p(\tau) = E\left\{H\{\ln[1 + mx(t)]\}\, H\left[\ln[1 + mx(t + \tau)]\right]\right\}. \tag{3-116}$$

It is well known that the autocorrelation function of the argument of the Hilbert transform is the same as the autocorrelation function of the original function [13]; therefore,

$$R_p(\tau) = E\left\{\ln[1 + mx(t)]\ln[1 + mx(t + \tau)]\right\}. \tag{3-117}$$

A closed-form solution of this equation is not possible. If $mx(t) \ll 1$, Equation (3-111) can be approximately solved by

using the first two terms of the series expansion of $\ln(1 + x)$:

$$R_p(\tau) \approx E\left\{\left[mx(t) - \frac{m^2}{2} x^2(t)\right]\left[mx(t+\tau) - \frac{m^2}{2} x^2(t+\tau)\right]\right\} \tag{3-118}$$

which for a ZMGN message signal reduces to

$$R_p(\tau) \approx m^2 \rho(\tau) + \frac{1}{4} m^4(1 + 2\rho^2(\tau)) . \tag{3-119}$$

3.3.9 CSSB-PM

The autocorrelation functions of the CSSB-PM components can be derived in terms of the autocorrelations functions obtained for PM. The required modification is to replace the real-valued process x(t) with the complex valued process s(t) indicated by Equations (3-30) and (3-31). Since x(t) is a gaussian process, $\hat{x}(t)$ is also a gaussian process and the sum of the two variables is also a gaussian process; therefore, s(t) is a complex gaussian process. Miller [46] showed that the same formula used to calculate the moments of a real multivariate gaussian process may be used to calculate the moments of a complex gaussian process. The autocorrelation function of the complex process is

$$R_s(\tau) = 2D_p^2\left[\rho(\tau) + j\,\hat{\rho}(\tau)\right], \tag{3-120}$$

and substituting this expression in Equation (3-100)

$$R_v(\tau) = C^2 \exp\left[-2D_p^2(1 - \rho(\tau) - j\,\hat{\rho}(\tau))\right]. \tag{3-121}$$

The autocorrelation function of a(t) is obtained as

$$R_a(\tau) = E\left\{C^2 \exp\left[-D_p x(t) - D_p x(t+\tau)\right]\right\} \qquad (3\text{-}122)$$

but since [13]

$$E\left\{\exp\left[jc_1x_1 + jc_2x_2\right]\right\} = \exp\left[-\frac{1}{2} R_x(0)(c_1^2 + c_2^2) - R_x(\tau)c_1c_2\right] \qquad (3\text{-}123)$$

where c_1 and c_2 are constants, x_1 and x_2 are ZMGN variables, and $R_x(\tau)$ is the autocorrelation function between x_1 and x_2, Equation (3-122) becomes

$$R_a(\tau) = C^2 \exp\left[\frac{1}{2} D_p^2\left[1 + \rho(\tau)\right]\right]. \qquad (3\text{-}124)$$

The autocorrelation function of p(t) is

$$R_p(\tau) = D_p^2 \rho(\tau). \qquad (3\text{-}125)$$

The autocorrelation functions of i(t) and q(t) are found by expanding Equation (3-123) in terms of its conjugate components and comparing the expansion with Equation (3-57). Note that if v(t) is Analytic, i(t) and q(t) have the same autocorrelation function so

$$R_i(\tau) = R_q(\tau) = \frac{C^2}{2} \exp\left[-2D_p^2\left[1 - \rho(\tau)\right]\right] \cos\left[2D_p^2 \hat{\rho}(\tau)\right]. \qquad (3\text{-}126)$$

3.3.10 CSSB-FM

If the conditions set for in Section 3.3.9 are met, then the equations derived for CSSB-PM are modified to yield:

$$R_v(\tau) = C^2 \exp\left[-2R_\phi(0) + 2R_\phi(\tau) + j2\hat{R}_\phi(\tau)\right], \qquad (3\text{-}127)$$

$$R_a(\tau) = C^2 \exp[2R_\phi(0) + 2R_\phi(\tau)], \tag{3-128}$$

$$R_p(\tau) = R_\phi(\tau) \tag{3-129}$$

and

$$R_i(\tau) = R_q(\tau) = \frac{C^2}{2} \exp\left[-2R_\phi(0) + 2R_\phi(\tau)\right] \cos\left[2R_\phi(\tau)\right]. \tag{3-130}$$

3.4 Second Moment Bandwidth of the Components

Using Equation (2-24), the second moment bandwidth of the components can be calculated.

3.4.1 AM

Using Equations (3-68), (3-69), and (3-70) together with the definition of second moment bandwidth yields

$$B_{s_a}^2 = \frac{-m^2}{1 + m^2\rho(0)} \cdot \rho''(0) = \frac{m^2}{1 + m^2} B_{s_x}^2, \tag{3-131}$$

$$B_{s_i} = B_{s_a}, \tag{3-132}$$

and

$$B_{s_q} = B_{s_p} = \text{not defined}, \tag{3-133}$$

where B_{s_x} is the second moment bandwidth of the message. In practice the carrier term can be neglected and $B_{s_a} = B_{s_x}$.

3.4.2 DSB-LM

The second moment bandwidth of a(t) can be found as follows. First find $R_a''(\tau)$ using Equation (3-74)

$$R_a''(\tau) = \frac{2}{\pi} C^2 \frac{\rho''(\tau)}{\rho(\tau)} \rho(\tau)\arcsin[\rho(\tau)]. \qquad (3\text{-}134)$$

Then

$$R_a(0) = C^2; \qquad (3\text{-}135)$$

therefore,

$$B_{s_a} = B_{s_x} . \qquad (3\text{-}136)$$

The bandwidth of the phase function requires the calculation of the second derivative of $R_p(\tau)$:

$$R_p''(\tau) = \frac{\pi}{2} \frac{\rho''(\tau)[1 - \rho^2(\tau)] + \rho^2(\tau)[\rho'(\tau)]^2}{[1 - \rho^2(\tau)]^{3/2}}, \qquad (3\text{-}137)$$

which can be rewritten as

$$R_p''(\tau) = \frac{\pi}{2} \left[\frac{\rho''(\tau) + \rho(\tau)[R_p'(\tau)]^2}{\sqrt{1 - \rho^2(\tau)}} \right], \qquad (3\text{-}138)$$

where

$$R_p'(\tau) = \frac{\rho'(\tau)}{\sqrt{1 - \rho^2(\tau)}} . \qquad (3\text{-}139)$$

Let $\tau \to 0$, and observe that $R_p'(0) = 0$, so

$$R_p''(0) = \lim_{\tau \to 0} R_p''(\tau) = \infty . \qquad (3\text{-}140)$$

This implies that

$$B_{s_p} = \infty . \qquad (3\text{-}141)$$

The second moment bandwidths of i(t) and q(t) are

$$B_{s_i} = B_{s_x} \tag{3-142}$$

and

$$B_{s_q} = \text{not defined.} \tag{3-143}$$

3.4.3 SSB-LM

Expressions for the second moment bandwidths of a(t) and p(t) cannot be obtained because the autocorrelation functions are not available. Equation (2-27) can be used, but unfortunately the evaluation of the required expectations is not physically tractable.

The second moment bandwidths of the quadrature components are easily found by inspection

$$B_{s_i} = B_{s_x} \tag{3-144}$$

and

$$B_{s_q} = B_{s_{\hat{x}}} = B_{s_x} \tag{3-145}$$

3.4.4 PM

The second moment bandwidths of the amplitude and phase function are easily found

$$B_{s_a} = 0 \tag{3-146}$$

and

$$B_{s_p} = B_{s_x} . \tag{3-147}$$

The bandwidth expressions for the quadrature components are obtained as follows. For the real quadrature component,

$$B_{s_i}^2 = \frac{-D_p^2 \rho''(0)\sinh(D_p^2)}{(2\pi)^2 \cosh(D_p^2)} \tag{3-148}$$

or

$$B_{s_i}^2 = D_p^2 \, B_{s_x} \tanh(D_p^2). \tag{3-149}$$

Similarly,

$$B_{s_q}^2 = D_p^2 \, B_{s_x}^2 \coth(D_p^2). \tag{3-150}$$

This illustrate that the i(t) and q(t) components require different bandwidths.

3.4.5 FM

For the FM case, use the same Equations as for PM, except that D_p^2 is replaced with $R_\phi(0)$,

$$R_\phi(0) = \frac{D_f^2}{2} \int_{-\infty}^{\infty} \frac{G_x(w)}{w^2} \, dw \tag{3-151}$$

where $G_x(w)$ is given by Equation (3-109).

3.4.6 CSSB-AM

Since the expressions for the autocorrelation function of i(t) is not available, the calculation of the second moment bandwidth of i(t) is not possible. Since q(t) also depends on i(t), the bandwidth for this signal cannot be found.

The expression for the bandwidth of a(t) is the same as Equation (3-131) found for AM. The phase function has an

autocorrelation function given by Equation (3-119) from which the second moment bandwidth is obtained

$$B_{s_p}^2 = B_{s_x}^2 \left[\frac{m^2 + m^4}{m^2 + \frac{1}{2}m^4} \right], \text{ if } mx(t) \ll 1. \qquad (3\text{-}152)$$

which illustrates a dependence of the phase function bandwidth on the modulation index.

3.4.7 CSSB-PM

The second moment bandwidth of the amplitude function is found by substituting Equation (3-124) into Equation (2-24). This yields

$$B_{s_a} = \sqrt{\frac{1}{2}}\, D_p B_{s_x} . \qquad (3\text{-}153)$$

The phase function has a bandwidth of

$$B_{s_p} = B_{s_x} . \qquad (3\text{-}154)$$

The bandwidth of the quadrature components is found using Equation (3-126) and the fact that both components have the same autocorrelation function. It is easily shown that

$$R_i''(0) = 2D_p^2 \rho''(0) R_i(0) - \left[2D_p^2 \hat{\rho}'(0)\right]^2 R_i(0); \qquad (3\text{-}155)$$

therefore,

$$B_{s_i}^2 = B_{s_q}^2 = 2D_p^2 B_{s_x} + \left[2D_p^2 \hat{\rho}'(0)\right]^2 . \qquad (3\text{-}156)$$

For the case of a bandlimited ZMGN message with flat spectra, bandlimited to W radians/sec.,

$$\rho(\tau) = \frac{\sin(W\tau)}{W\tau} \tag{3-157}$$

so

$$\hat{\rho}'(0) = \frac{-W}{2} = \frac{-B_{s_x}}{2\sqrt{3}} \tag{3-158}$$

where W is the absolute bandwidth of the message. Substitution of this equation into Equation (3-156) yields

$$B_{s_i}^2 = D_p^2 B_{s_x}^2 \left[2 + \frac{D_p^2}{3}\right]. \tag{3-159}$$

which shows a faster increase in bandwidth as D_p increases compared to the PM case.

3.4.8 CSSB-FM

The equations for CSSB-FM are easily obtained by replacing $R_{\phi}(0)$, given by Equation (3-151), in place of D_p^2 in Equations (3-153) thru (3-159).

3.5 Summary

This chapter presented additional complex envelope properties that are not available in the literature. Mathematical expressions were derived for the second moment bandwidth and the autocorrelation function of the complex envelope components for most of the modulation laws under consideration. The bandwidth equations show how the complex envelope bandwidth is related to the modulation law parameters and the message statistics.

The equations obtained for the bandwidth of CSSB-PM quadrature functions reveals that these functions may require

larger bandwidth as compared to their double-sided counterparts, but any comparison must be done on the basis of a value of D_p and the message statistics.

CHAPTER IV
THE UNIVERSAL TRANSMITTER

4.1 Introduction

This chapter presents the idea that a transmitter can be designed so that it is capable of producing almost any type of modulated carrier. The general structure of the transmitter was already shown in Figure 1.1. First, the complex envelope is generated from the message and the modulation law. A complex carrier is subsequently modulated with this complex envelope, the real part is extracted and amplified. This is an universal transmitter since the baseband preprocessor can be programmed to implement the desired modulation law. The block diagram of this system is shown in Figure 4.1.

The system proposed in the previous paragraph is unrealizable because it requires the generation of complex-valued voltages. A complex function can be described by real-valued functions, so by manipulation of the different representations of the complex envelope specified by the operations illustrated in Figure 4.1, a physically realizable system can be obtained.

In a practical transmitter it is always possible to separate the power amplifier from the actual modulator circuit if a linear amplifier is used. Since amplification is a linear operation, nothing is lost in the analysis if the power amplifier is disregarded. This system without power amplification will be defined as the universal modulator.

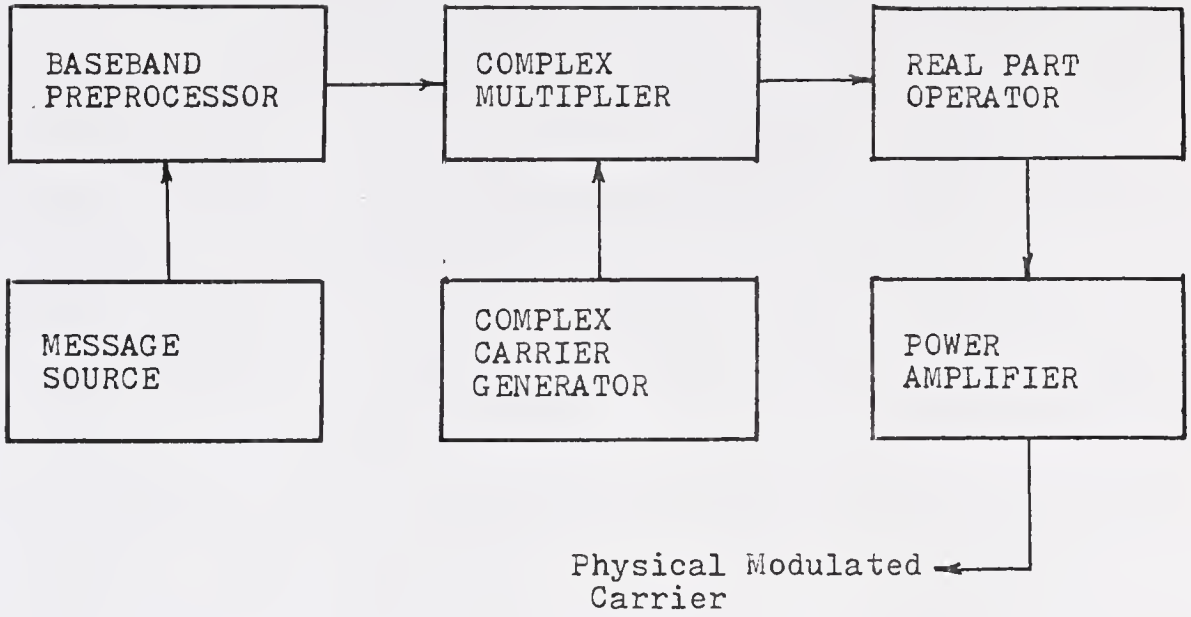


Figure 4.1. Block diagram of the universal transmitter.

4.1.1 The Complex Envelope as a Vector

It is well known that a complex-valued function can be represented as the trajectory of the tip of a vector in the two-dimensional plane described by the real and the imaginary axes. A vector in a two-dimensional plane can be described in terms of its phase angle and its magnitude or in terms of a linear combination of two nonparallel vectors. Four of these posibilities are illustrated in Figure 4.2. The polar representation consists of the magnitude, a(t), and the phase angle, p(t). The quadrature components are two orthogonal vectors that represent the quadrature functions, i(t) and q(t). These are the conventional representations for a complex-valued

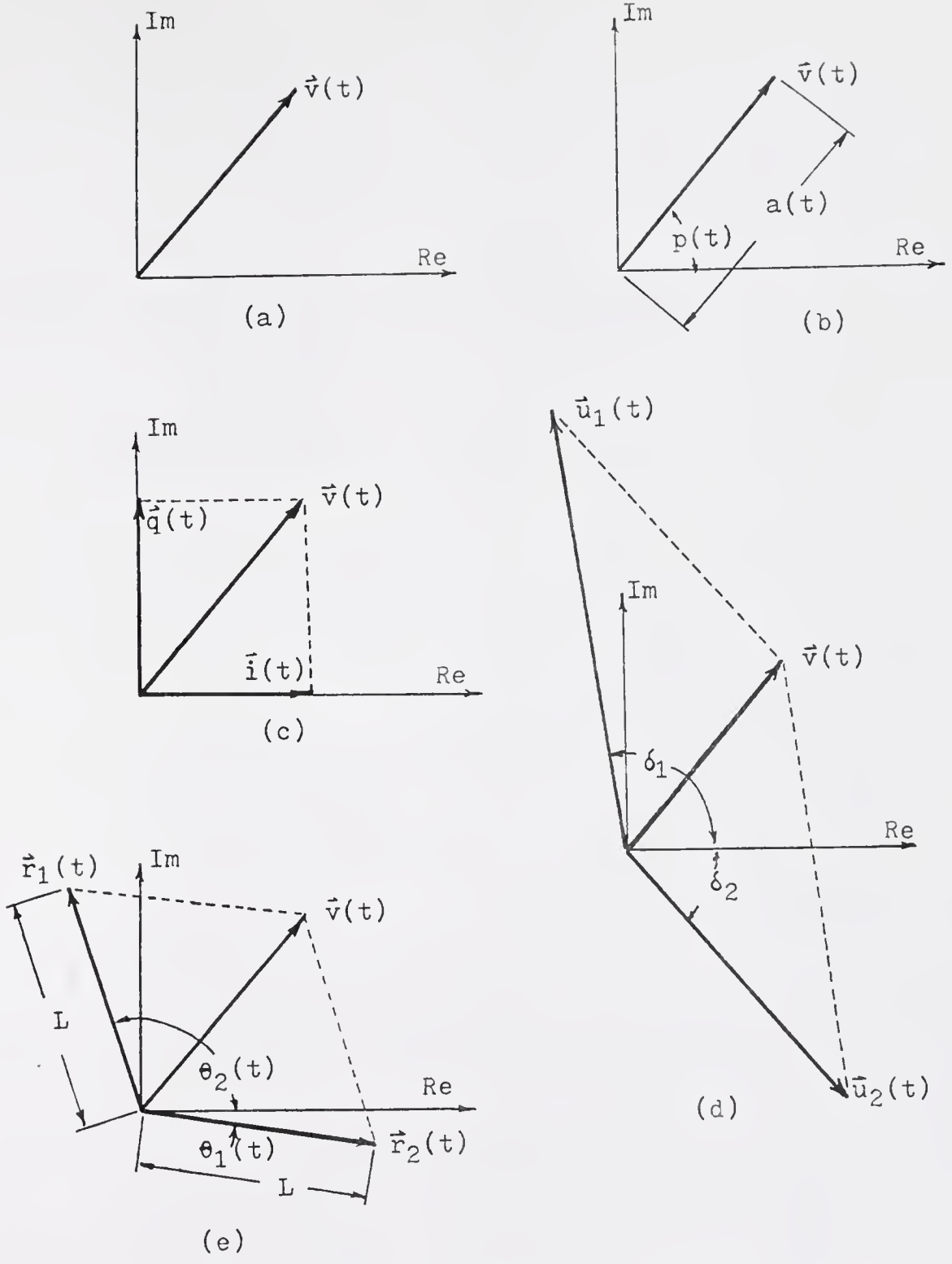


Figure 4.2. Possible vector representations of the complex envelope, (a) the complex envelope vector, (b) polar components, (c) quadrature components, (d) two amplitude-modulated vectors, and (e) two phase-modulated vectors.

function and correspond to the sketches shown in Figure 4.2(b) and 4.2(c).

There are two more representations that are a direct consequence of the vector representation. The linear combination of two nonorthogonal vectors is shown in Figure 4.2(d). Here the magnitudes of the two vectors are adjusted to represent the complex envelope vector while their phase angles remain fixed. In the last case, illustrated by Figure 4.2(e), the length of the vectors remain constant and the phase angle is adjusted. These equivalent representations will generate the same complex envelope, but a certain representations may have economic advantages in terms of circuit realization energy consumption.

4.1.2 Physical Modulators

This section is concerned with the models for the physical modulators. It is necessary to understand the modulator limitations in order to take them into account in the design.

The function of the modulator is to alter the properties of the carrier in accordance to the modulating signal. Ideal modulators do just that, but practical modulator circuits have definite maximum input levels and a limited frequency response. If the maximum input level is exceeded, the circuit may not operate as intended or the circuit is damaged. The effect on the modulated carrier is distortion or no carrier at all. The limited frequency response must be taken into account because if the modulator circuit cannot follow the modulating signal, the signal modulated into the carrier is different from what it is supposed to be.

The limitations on the maximum input level and on the frequency response can be modelled as a saturating amplifier and an ideal lowpass filter respectively. The modulator can be approximated as the cascade combination of these two blocks followed by an ideal modulator. There are two possible arrangements that are not equivalent. The final selection depends on which model describes better the modulator under consideration. It is possible that the maximum input level might be frequency-dependent or that the frequency response might be input-level-dependent. In these cases a new model is necessary.

Signal levels can be attenuated, so the maximum input level problem can be solved in cases where the message does not have such a large dynamic range that circuit noise becomes a problem. The solution to the frequency response problem is not that simple because there is no way of reducing the bandwidth occupied by a signal without changing the signal itself. This means that the modulator model can be simplified to the cascade combination of an ideal lowpass filter and an ideal modulator.

There are two basic modulators. These are the amplitude modulator and the phase modulator. The ideal modulator is defined by the following equation

$$y(t) = \begin{cases} K_a\left[1 + K_b e(t)\right]\cos(w_o t), & \text{if } K_b e(t) \geq -1 \\ 0, & \text{if } K_b e(t) < -1, \end{cases} \quad (4\text{-}1)$$

where $e(t)$ is the input modulating signal, K_a is the carrier peak amplitude in the absence of a modulating signal, K_b is

a modulator constant, w_o is the carrier frequency, and $y(t)$ is the modulated carrier.

The ideal phase modulator is defined by following equation:

$$y(t) = \begin{cases} K_a \cos[w_o t + K_c e(t)], & \text{if } |K_c e(t)| \leq \pi \\ -K_a \cos(w_o t), & \text{if } |K_c e(t)| > \pi, \end{cases} \qquad (4\text{-}2)$$

where $e(t)$, K_a, and $y(t)$ are the same as above; and K_c is a modulator constant. This equation assumes that the modulator is not capable of producing a phase shift of more than π radians. This is in agreement with concept of minimum-phase signal defined in Section 2.1. If this restriction is lifted, the inequalities in Equation (4-2) have to be modified. This kind of modulator is defined as an ideal nonminimum-phase modulator.

Two ideal amplitude modulators can be combined as shown in Figure 4.3 to form an ideal balanced modulator. This is described by the following equation:

$$y(t) = K_a e(t) \cos(w_o t). \qquad (4\text{-}3)$$

If the ideal phase modulator is preceeded by an ideal integrator, an ideal frequency modulator is obtained. It is described by the following equation:

$$y(t) = K_c \cos\left[w_o t + K_c \int_{-\infty}^{t} e(u)\,du\right]. \qquad (4\text{-}4)$$

Note that for most frequency modulators the phase is not restricted to $\pm\pi$.

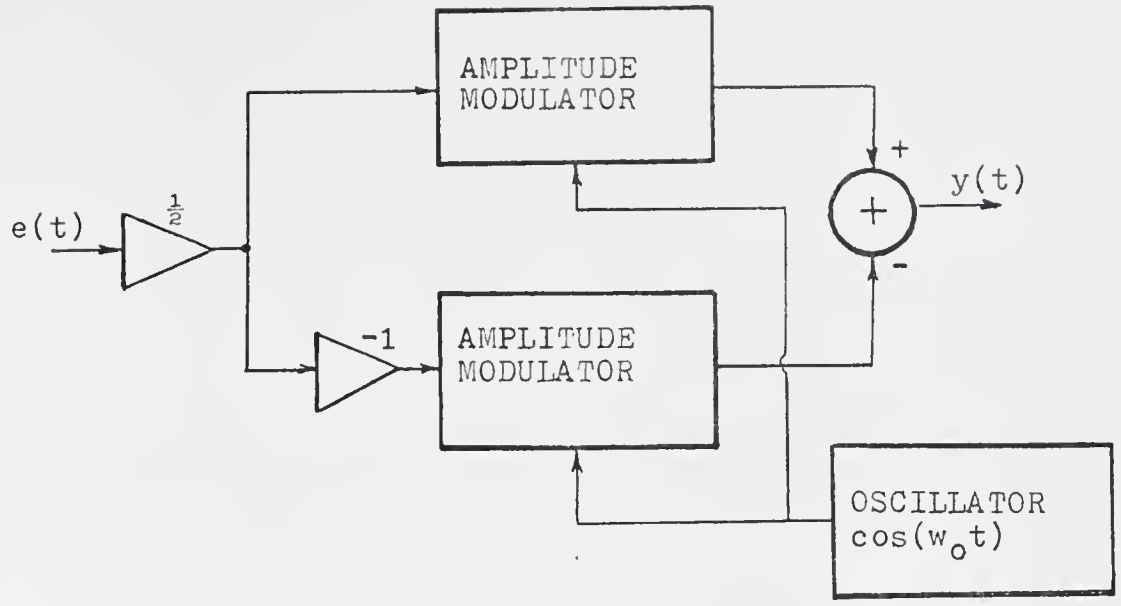


Figure 4.3. The balanced modulator.

A balanced modulator can be used in place of an amplitude modulator if a constant is added to the input. If $e_b(t)$ is the input to the balanced modulator and $e_a(t)$ is the input to the amplitude modulator, they are related by

$$e_b(t) = \left[1 + K_b e_a(t)\right]. \qquad (4\text{-}5)$$

The phase modulator can operate as a frequency modulator, if the minimum-phase condition is lifted. If the modulating signal has a zero mean value, the frequency modulator can replace a phase modulator with a differentiator at the input. If the modulating signal has a nonzero mean value the differentiator will destroy the mean value. Rowe [13] considers in detail the relation between the PM and FM functions and their power spectral density limitations.

4.1.3 The AM/PM Modulator

The first structure to be considered to serve as an uni-

versal modulator is the AM/PM modulator [6, 7]. It consists of a phase modulator followed by an amplitude modulator together with a pair of baseband preprocessors. The baseband preprocessor #1 generates a(t) while the baseband preprocessor #2 generates p(t). To relate a(t) to the input of the amplitude modulator, compare Equation (4-1) with Equation (2-1) and observe that

$$a(t) = K_a\left[1 + K_b e_a(t)\right] , \tag{4-6}$$

so the input to the amplitude modulator, $e_a(t)$, is given by

$$e_a(t) = \frac{1}{K_b}\left[\frac{a(t)}{K_a} - 1\right] . \tag{4-7}$$

To determine the input to the phase modulator, compare Equations (4-2) and (2-1). These equations require that

$$p(t) = K_c e_p(t), \tag{4-8}$$

where $e_p(t)$ is the input to the phase modulator. The required arrangement is shown as a block diagram in Figure 4.4.

Observe that if Equation (4-5) is used, a simpler block diagram results because the amplitude modulator can be replaced by a balanced modulator. If $e_b(t)$ is the input to the balanced modulator, it is easy to see that

$$e_b(t) = \frac{a(t)}{K_a}. \tag{4-9}$$

This arrangement is shown in Figure 4.5 and from now on will be referred to as the AM/PM modulator.

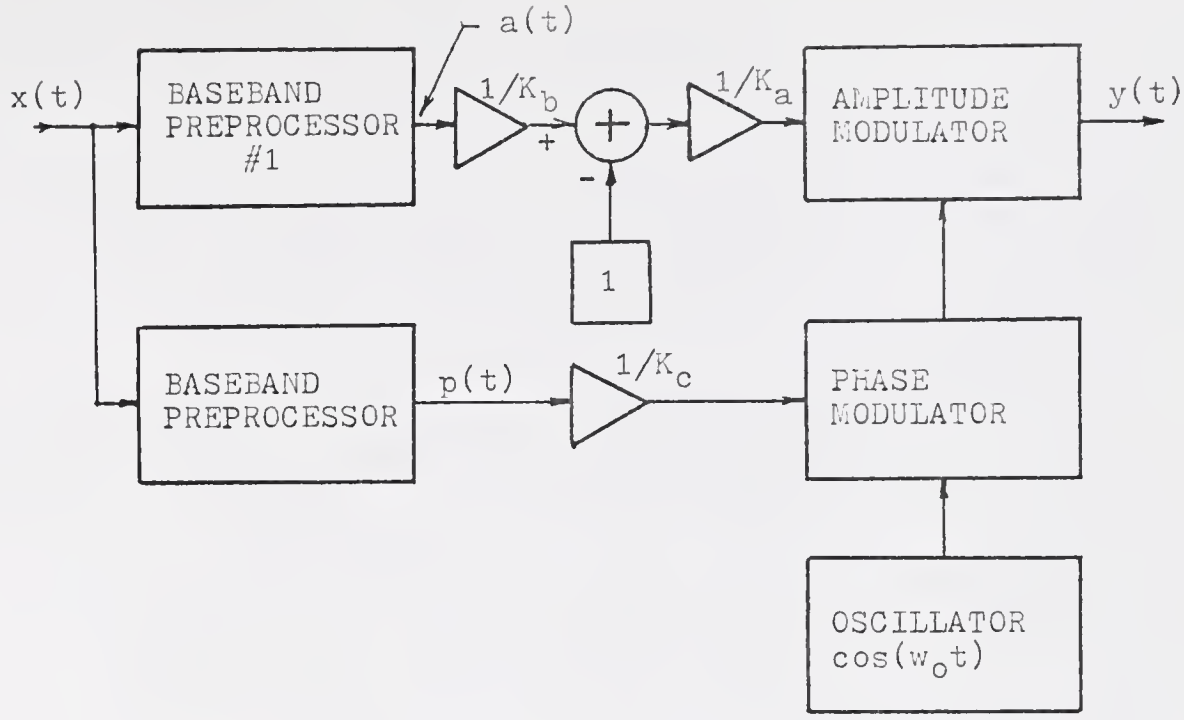


Figure 4.4. The AM/PM modulator.

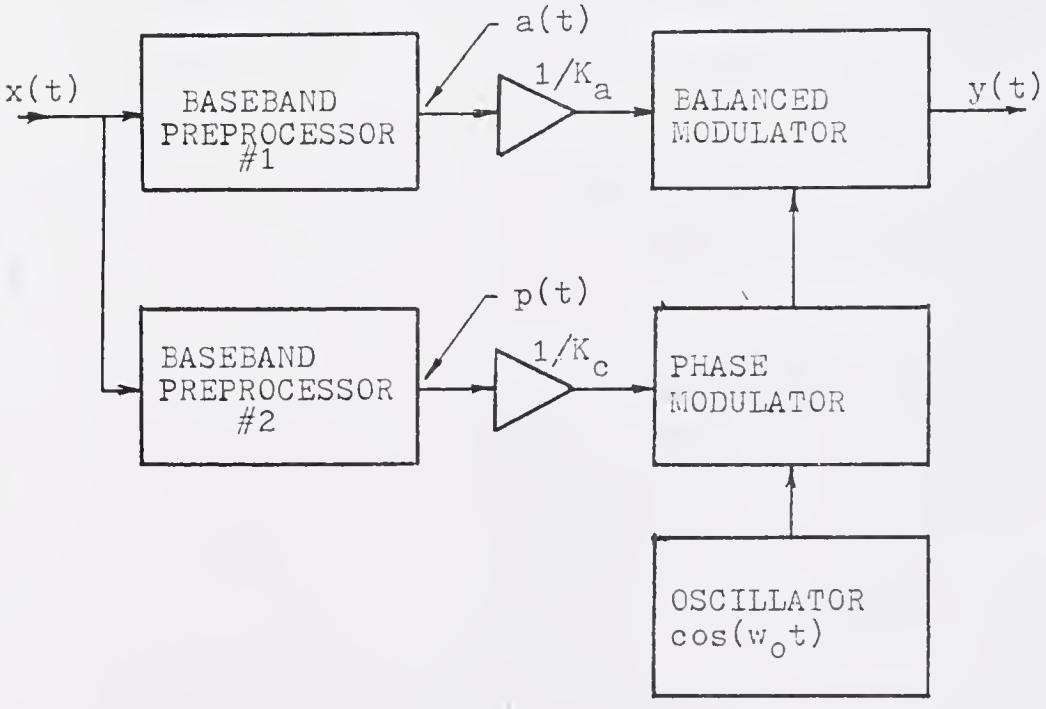


Figure 4.5. The improved AM/PM modulator.

4.1.4 The Quadrature Modulator

The arrangement of two balanced modulators in phase quadrature is a well known method of generatign SSB-LM [47]. The idea has been extended tc other types of modulation by Thomas [14] and others. The block diagram for the quadrature modulator is shown in Figure 4.6. It is desired to generate Equation (2-7) with the complex envelope quadrature components, so

$$y(t) = \mathrm{Re}\left\{[i(t) + jq(t)]\exp(jw_o t)\right\}. \tag{4-10}$$

This can be simplified to

$$y(t) = i(t)\cos(w_o t) - q(t)\sin(w_o t), \tag{4-11}$$

so the baseband preprocessor #3 has to generate i(t) while the baseband preprocessor #4 generates q(t). The i(t) and q(t) functions will be defined as the quadrature components. These quadrature components are related to the inputs of the balanced modulators by

$$e_{b_i}(t) = \frac{i(t)}{K_a}, \tag{4-12}$$

and

$$e_{b_q}(t) = \frac{q(t)}{K_a}. \tag{4-13}$$

4.1.5 The PM/PM Modulator

The idea of combining two phase modulators to generate an AM carrier was proposed by Chireix [48]. This system is

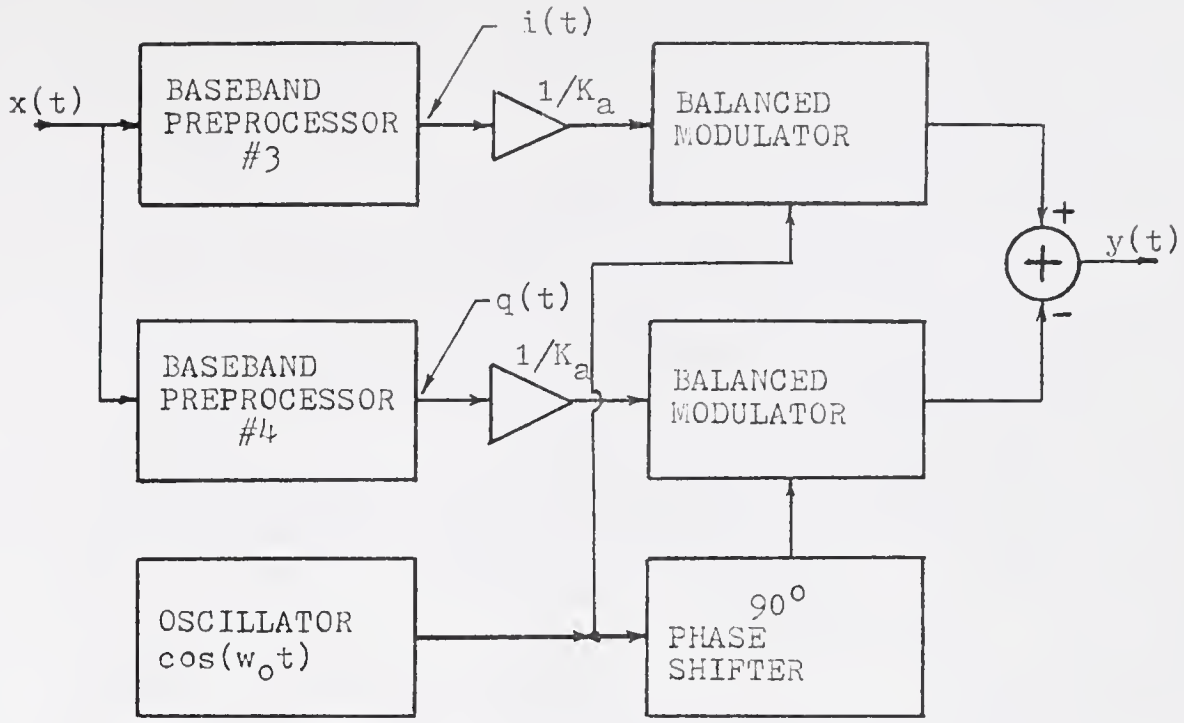


Figure 4.6. The quadrature modulator.

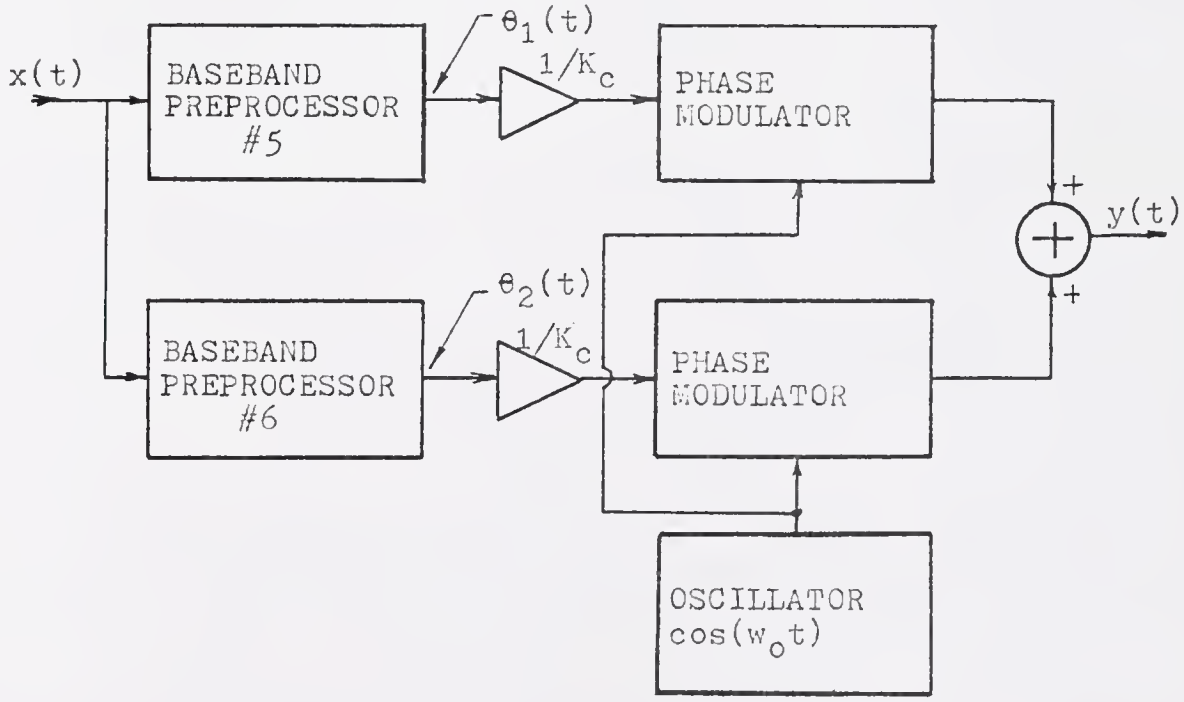


Figure 4.7. The PM/PM modulator.

marketed by RCA under the trademark "Ampliphase". The same idea can be extended to generate other types of modulation. If Figure 4.2(e) is examined, it is seen that the complex envelope can be obtained by adding two constant amplitude vectors with an specific phase arrangement. This is equivalent to the addition of two phase modulated carriers with constant amplitude. The block diagram of the proposed system is shown in Figure 4.7. The details will be worked out later.

4.1.6 Other Configurations

It is possible to obtain the complex envelope by combining two vectors that are not orthogonal as shown in Figure 4.2(d). The required block diagram of a system like this is similar to the block diagram for the quadrature universal modulator, except that the carriers going into the balanced modulators are not in phase quadrature. Observe that in general

$$\vec{v}(t) = \vec{u}_1(t) + \vec{u}_2(t), \tag{4-14}$$

but each vector $\vec{u}_1(t)$ and $\vec{u}_2(t)$ can be written as a linear combination of $\vec{i}(t)$ and $\vec{q}(t)$. Let $\vec{k}_i$ and $\vec{k}_q$ be unit vectors in the real and imaginary axis directions, then

$$\vec{u}_1(t) = a_1\vec{k}_i + a_2\vec{k}_q \tag{4-15}$$

and

$$\vec{u}_2(t) = b_1\vec{k}_i + b_2\vec{k}_q. \tag{4-16}$$

Observe that

$$a_1 = |\vec{u}_1(t)| \cos(\delta_1), \tag{4-17}$$

$$a_2 = |\vec{u}_1(t)| \sin(\delta_1), \tag{4-18}$$

$$b_1 = |\vec{u}_2(t)| \cos(\delta_2), \tag{4-19}$$

and

$$b_2 = |\vec{u}_2(t)| \sin(\delta_2); \tag{4-20}$$

where $|\vec{u}_1(t)|$ and $|\vec{u}_2(t)|$ are obtained using the law of the sines,

$$|\vec{u}_1(t)| = \frac{a(t)\sin[\delta_2 - p(t)]}{\sin(\delta_1 - \delta_2 + \pi)} \tag{4-21}$$

and

$$|\vec{u}_2(t)| = \frac{a(t)\sin[p(t) - \delta_1]}{\sin(\delta_1 - \delta_2 + \pi)}\ ; \tag{4-22}$$

which simplify to

$$|\vec{u}_1(t)| = \frac{-i(t)\sin(\delta_2) + q(t)\cos(\delta_2)}{\sin(\delta_1 - \delta_2)}\ , \tag{4-23}$$

and

$$|\vec{u}_2(t)| = \frac{-q(t)\cos(\delta_1) + i(t)\sin(\delta_1)}{\sin(\delta_1 - \delta_2)}\ . \tag{4-24}$$

Substitute Equations (4-23) and (4-24) into Equations (4-17) thru (4-20) and substitute these into Equations (4-15) and (4-16) to obtain:

$$\vec{u}_1(t) = \frac{-i(t)\sin(\delta_2) + q(t)\cos(\delta_2)}{\sin(\delta_1 - \delta_2)} \cdot \cos(\delta_1)\vec{k} + \sin(\delta_2)\vec{k}_q \tag{4-25}$$

and

$$|\vec{u}_2(t)| = \frac{-q(t)\cos(\delta_1) + i(t)\sin(\delta_1)}{\sin(\delta_1 - \delta_2)} \cdot$$

$$\cdot \left[\sin(\delta_1)\vec{k}_i + \sin(\delta_2)\vec{k}_q\right], \qquad (4\text{-}26)$$

so $\vec{u}_1(t)$ and $\vec{u}_2(t)$ can be expressed in terms of a linear combination of i(t) and q(t); therefore, the analysis is very similar to the quadrature modulator case. This nonorthogonal representation has a serious disadvantage. The magnitude of the vectors $\vec{i}(t)$ and $\vec{q}(t)$ is never greater than the magnitude of $\vec{v}(t)$. The magnitude of $\vec{u}_1(t)$ and $\vec{u}_2(t)$ can be several times larger than the magnitude of $\vec{v}(t)$ as shown by Equations (4-21) and (4-22). This can be avoided with a slight modification of δ_1 and δ_2, for example, 0° and 90°. This results in the quadrature modulator which has been considered before.

Meewezen [21] proposed a system where the modulation process is done by independent modulation of the sidebands. This will be shown to be similar to the quadrature universal modulator. The complex envelope is obtained as the combination of two functions,

$$v(t) = r_u(t) + r_l(t) \qquad (4\text{-}27)$$

where $r_u(t)$ represents upper sideband modulation and $r_l(t)$ represents lower sideband modulation. These two signals are single-sided; therefore, they are Analytic, so each one can be described as

$$r_u(t) = u(t) + j\hat{u}(t) \tag{4-28}$$

and

$$r_l(t) = l(t) - j\hat{l}(t) \tag{4-29}$$

where $u(t)$ and $l(t)$ are real-valued functions not yet determined. Compare these equations with Equation (3-2) and observe that

$$i(t) = u(t) + l(t) \tag{4-30}$$

and

$$q(t) = \hat{u}(t) - \hat{l}(t), \tag{4-31}$$

so

$$u(t) = \frac{i(t) - \hat{q}(t)}{2} \tag{4-32}$$

and

$$l(t) = \frac{i(t) + \hat{q}(t)}{2}. \tag{4-33}$$

This means that the analysis done for the quadrature modulator can be applied to Meewezen's approach to modulation. The block diagram of this system is shown in Figure 4.8. Observe that only $i(t)$ and $q(t)$ have to be obtained, that the system requires two SSB-LM modulators and that a Hilbert transformer is necessary. The complexity of such system is not compensated by any advantage over the other systems already under consideration. For this reason this will not be considered in detail.

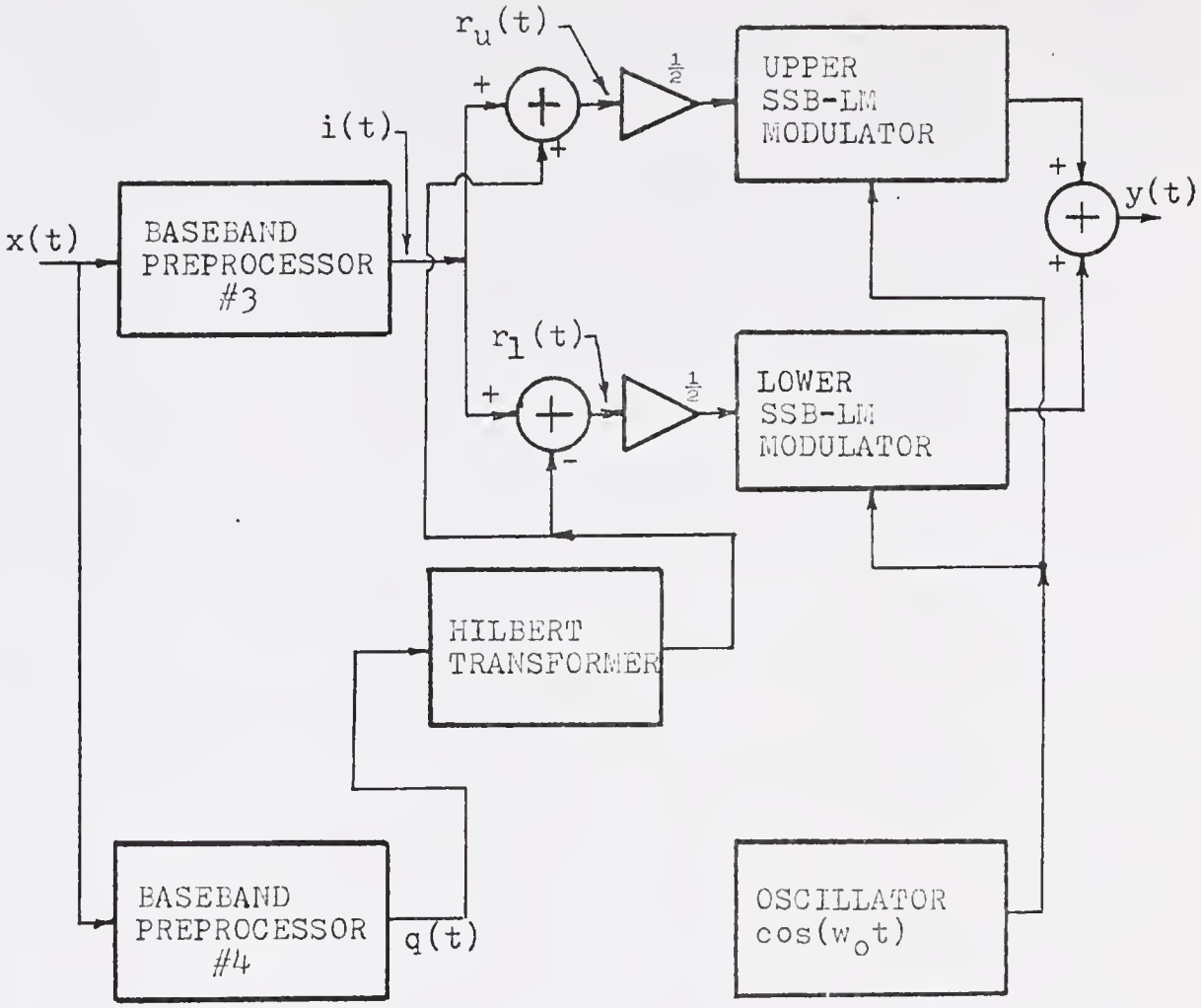


Figure 4.8. Meewezen's independent sideband modulation in terms of the quadrature components.

Cain [22] studied the generation of CSSB modulations with a SSB-LM transmitter. This is similar to the Meewezen's approach if only one sideband is modulated.

4.1.7 Criteria for Comparison and Evaluation of the Modulators

The two main criteria for comparison are the bandwidth required by the signals and their dynamic range. These are the two most important factors to consider in the design and it is important to study their effect for different types of modulation laws.

The point where the signal will be studied is at the output of the baseband preprocessor. These are the reasons to justify that selection: (1) this is the point where all modulation-type-dependent signal-processing ends and where the signals go to the modulation-law-invariant part of the system; (2) this is the point where digital-to-analog (D/A) conversion is most likely to take place; and (3) at this point the signal bandwidths are independent of the properties of the actual circuits that will be used to build the baseband postprocessor or the modulators.

The modulation parameters and the message signal have influence on the bandwidth of the complex envelope component functions and their dynamic range. In addition, the mathematical definition of these components may impose additional constraints that must be considered.

Bandwidth was defined in five different ways in Section 2.5.1. The definition to be used in this discussion is the equivalent-filter bandwidth because it is directly related to

the actual bandwidth required for the system. The measure of error is the intermodulation distortion (IM) at the output of an ideal receiver in the absence of noise. The analysis and calculations will relate the bandwidth of the complex envelope components at the output of the baseband preprocessor to the distortion of the message at the receiver.

The method to be used to calculate the IM distortion is best illustrated by Figure 4.9. There is a control signal that is obtained from the output of the baseband preprocessor, modulated, and demodulated by the receiver. The spectrum of the demodulated control signal is calculated to detect any nonlinearities acting on the message signal. Ideally, the IM distortion should be zero. In practice it is nonzero due to round-off errors and aliasing of sampled signals. The control signal is considered acceptable if the IM distortion is less than 0.0005%. The signal whose bandwidth is to be measured is filtered by an ideal lowpass filter and from there on it follows the same path as the control signal.

It is necessary to perform three different tests because two complex envelope components are involved. For example, the first test consists of filtering a(t) and combining it with the control component, p(t). The second test reverses the situation, that is, a(t) is the control component. The third test involves filtering both a(t) and p(t). The purpose of the test is to determine if a(t) or p(t) requires more bandwidth than the other, and to check what will happen if both components are filtered simultaneously.

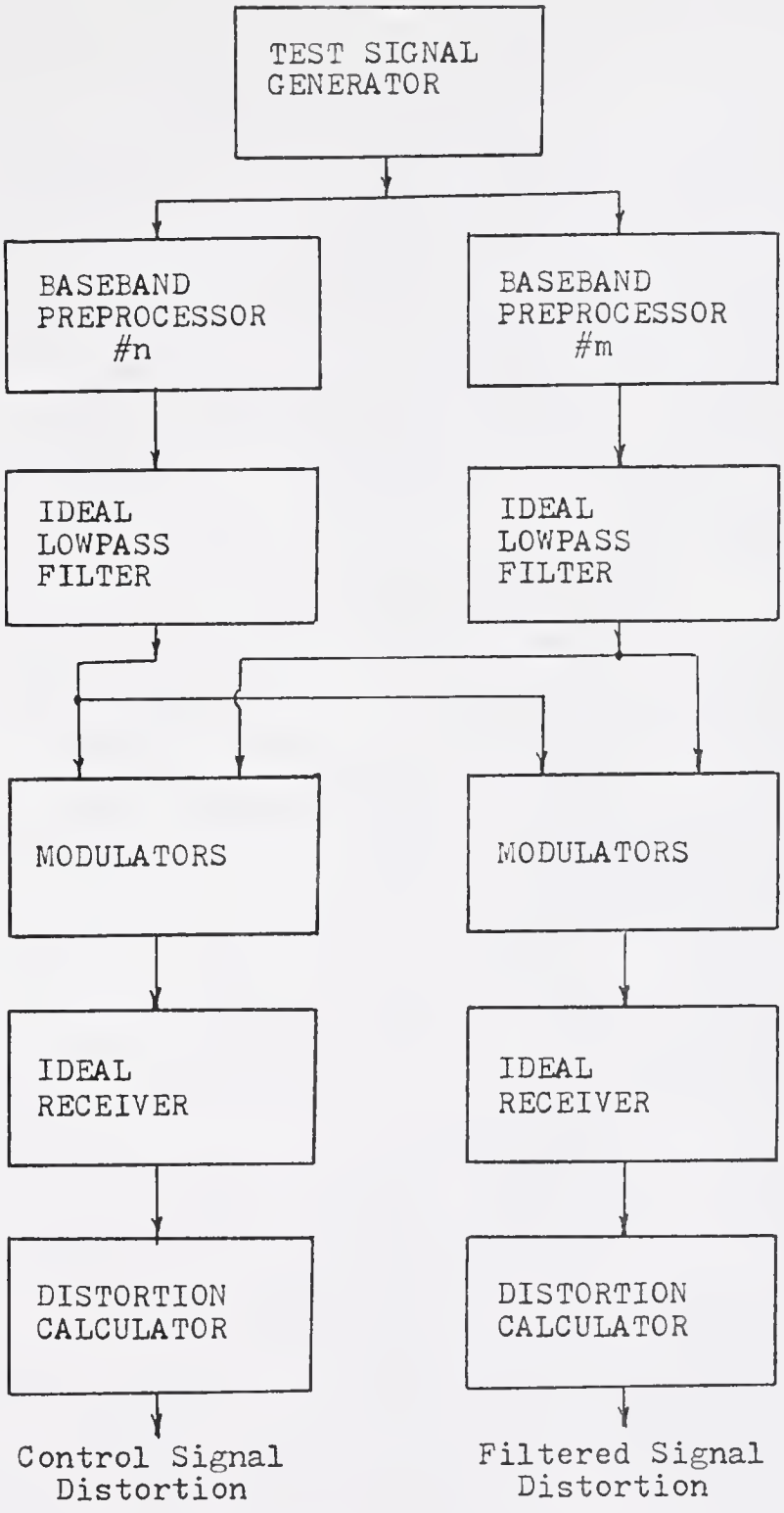


Figure 4.9. Method used to calculate the equivalent-filter bandwidth.

The test signal consists of two tones with equal amplitudes, but at two frequencies, f_1 and f_2, that are not harmonically related. The intermodulation distortion is calculated following Equation (2-31), but instead of including all the frequencies to infinity, it is modified to include all components up to the third harmonic of the first order intermodulation product $(f_1 + f_2)$. The test message signal is shown in Figure 4.10. The frequency scale is easily changed to represent other tones that have the same frequency relationship. Since the signal is deterministic, the peak amplitude is normalized to unity so the modulation parameters take care of the denormalization. For gaussian random signals the theoretical peak amplitude is infinite so it is necessary to normalize the variance and ignore those amplitudes that exceed a given level, for example, five times the standard deviation.

Another consideration in measuring bandwidth is that there are cases, CSSB-AM is an example, where one of the complex envelope components (p(t) in this case) is ignored by the ideal receiver in the noiseless case. The purpose of the

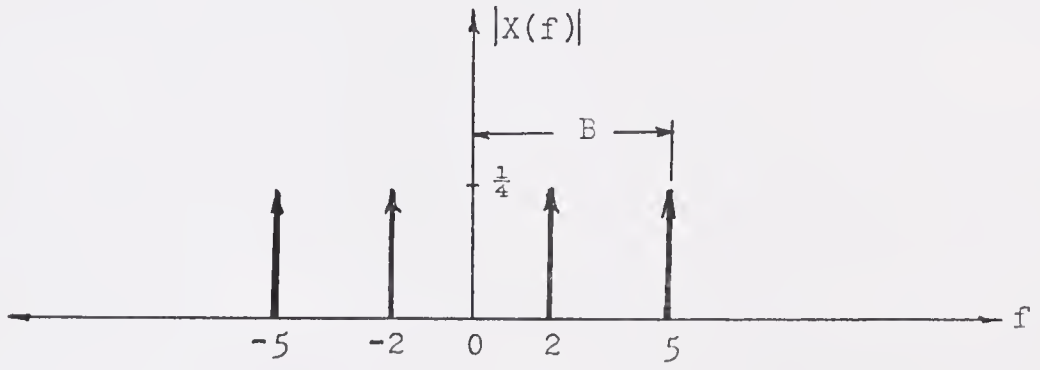


Figure 4.10. Magnitude of the spectrum of the test message signal.

neglected component is to produce a single-sided modulated carrier. The effect of filtering this component is reflected as incomplete sideband suppression and can be measured as the ratio of the power in the unwanted sideband to the total power, expressed as a percentage. This is given by the following equation:

$$S(\%) = \frac{P_u}{P_t} \times 100\% \qquad (4\text{-}34)$$

where P_u is the power present in the suppressed sideband, P_t is the total power, and $S(\%)$ is defined as the sideband suppression factor expressed as a percentage.

The calculations of the intermodulation distortion and the suppression factor were obtained with a computer simulation of the systems shown in Figures 4.6, 4.7, and 4.8. This requires the use of discrete signals to represent continuous signals. This was discussed in Section 2.6. To avoid errors due to aliasing of nonbandlimited signals, the control signal is monitored and the sampling rate is adjusted to reduce the distortion to acceptable levels.

The results obtained in Chapter III for the complex envelope components, their autocorrelation functions, and the second moment bandwidth are used to check the reliability of the calculations. They are not directly applicable because these autocorrelation functions were derived for the special case of a zero mean gaussian random process. However, the same general properties have to apply.

4.2 The AM/PM Modulator

This section analyses the AM/PM modulator whose block diagram is shown in Figure 4.5. This system was described in Section 4.1.3. The mathematical expressions for the transfer characteristics of the baseband preprocessors #1 and #2 are the same as the equations obtained in Section 3.2 for the magnitude and phase of the complex envelope. These functions are summarized on Table IV-1.

4.2.1 Bandwidth Requirements

The bandwidths required for the signals a(t) and p(t) are defined in terms of the equivalent-filter bandwidth. These are calculated following the procedure outlined in Section 4.1.5.

The equivalent-filter bandwidths of the a(t) and p(t) functions for AM, PM, and FM are easily calculated. The only nonlinearity is the addition of a constant in the AM a(t) function. In these cases, the equivalent-filter bandwidth can be defined for zero distortion at the receiver. This is achieved whenever the ideal lowpass filter shown in Figure 4.9 is at least as wide as the message absolute bandwidth B. For the cases under consideration, the equivalent filter bandwidth is B. In practice all signals are considered to be bandlimited [49].

In the DSB-LM and SSB-LM cases, a computer simulation of the system is necessary due to the complexity of the analysis. The computer program used in the simulation is listed in Appendix B. The results of the simulation are presented as

Table IV-1. Magnitude and Phase Functions for the AM/PM Modulator.

Type of Modulation	Baseband Preprocessor Transfer Characteristics.	
	#1 a(t)	#2 p(t)
AM	$C[1 + mx(t)]$	0
LM	$C\|x(t)\|$	$\frac{\pi}{2}[1 - \text{sgn}[x(t)]]$
SSB-LM	$C\sqrt{x^2(t) + \hat{x}^2(t)}$	$\arctan[\hat{x}(t)/x(t)]$
PM	C	$D_p x(t)$
FM	C	$D_f \int_{-\infty}^{t} x(u)du$
CSSB-AM	$C[1 + mx(t)]$	$H\{\ln[1 + mx(t)]\}$
CSSB-PM	$C \exp[-D_p \hat{x}(t)]$	$D_p x(t)$
CSSB-FM	$C \exp\left[-D_f \int_{-\infty}^{t} \hat{x}(u)du\right]$	$D_f \int_{-\infty}^{t} x(u)du$

a graph that indicates the relationship between the calculated intermodulation (IM) distortion and the equivalent filter bandwidth. The plots for the a(t) and p(t) components for DSB-LM and SSB-LM are shown in Figure 4.11. The bandwidth scale is normalized with respect to the message equivalent-filter bandwidth. These plots were obtained for the case where one of the two components is filtered and the other is taken from the experimental control signal. The cases where both a(t) and p(t) are filtered are very close to the plots for the filtered phase alone. These plots indicate that the SSB-LM phase function requires the widest bandwidth of all the modulation types under consideration. The reason for such a large bandwidth is that p(t) exhibits phase jumps in the SSB-LM and DSB-LM cases.

In the CSSB cases, there is usually a component that is ignored by the receiver. For example, CSSB-FM is detected by an FM receiver that ignores any amplitude modulation present in the carrier. This component is necessary to obtain a single-sided modulated carrier; therefore, its bandwidth must be determined. The p(t) function for CSSB-PM and CSSB-FM, and the a(t) function for CSSB-AM are identical to the PM, FM, and AM cases already considered. Instead of trying to measure the distortion in the "unused" component it is better to measure how this component affects the undesirable sideband suppression. The procedure was outlined in Section 4.1. The sideband suppression factors, S(%), for all the single-sided carriers under consideration are presented as a function of

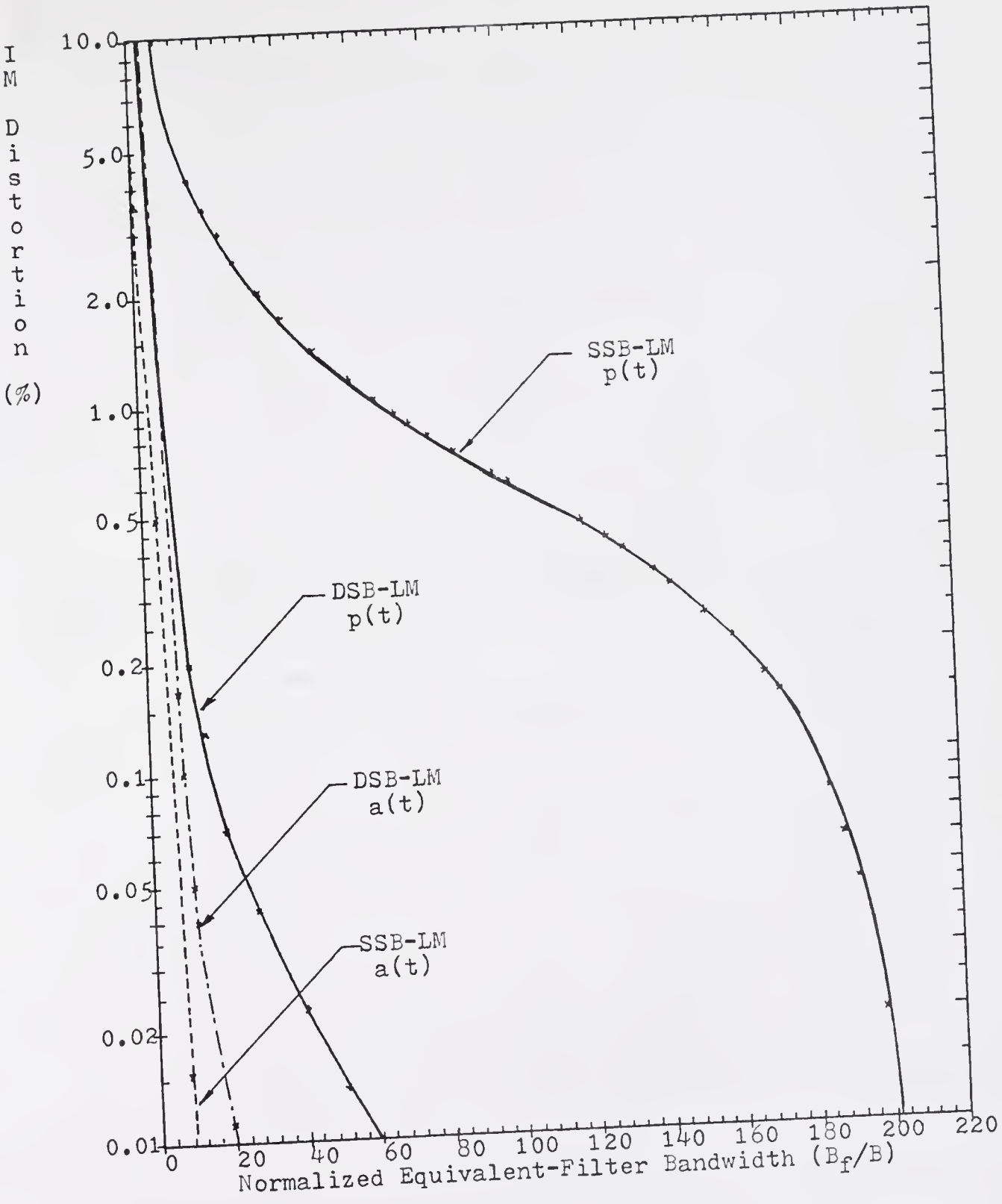


Figure 4.11. IM distortion of the AM/PM modulator a(t) and p(t) functions for DSB-LM and SSB-LM as a function of B_f.

the modulation parameters in Figures 4.12 thru 4.15.

Figure 4.12 presents the S(%) curve for the SSB-LM case. This curve shows that a(t) and p(t) also affect the distribution of power between sidebands. Again, the SSB-LM phase function demands more bandwidth than all the other cases under consideration. This is expected because a large distortion figure was obtained at the output of the receiver indicating that the real part of the complex envelope is distorted; therefore, it is reasonable to expect that the imaginary part is also distorted and the resulting complex envelope is far from being Analytic.

The plots for FM and CSSB-FM are in terms of D, the deviation ratio, defined as [1]

$$D = \frac{f_d}{B} = \frac{D_f}{2\pi B} , \qquad (4\text{-}35)$$

where f_d is the frequency deviation in Hz, B is the message bandwidth, and D_f is the frequency deviation in radians/sec.

The CSSB plots for the sideband suppression factor are presented in Figures 4.13 thru 4.15. These curves show how the equivalent-filter bandwidth normalized with respect to the message bandwidth is related to the modulation constants when the sideband suppression factor is used as a parameter. These modulation constants are the AM modulation index in the CSSB-AM case, the PM modulation index in the CSSB-PM case, and the FM deviation ratio in the CSSB-FM case. These modulation constants were defined in Equation (3-3), (3-31), and (4-35) respectively. All the curves exhibit the same general

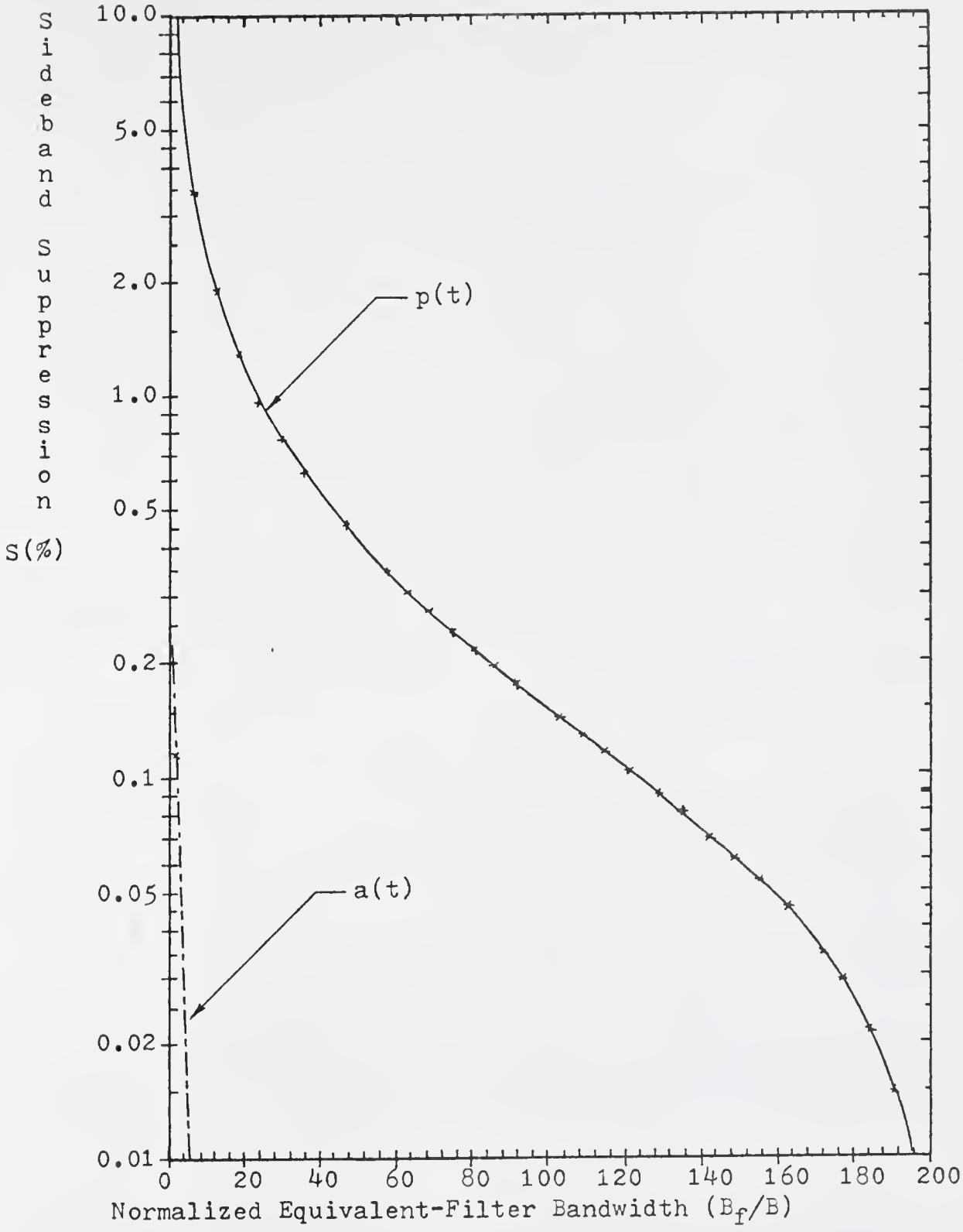


Figure 4.12. Sideband suppression factor (S(%)) of the AM/PM modulator a(t) and p(t) functions for SSB-LM as a function of B_f.

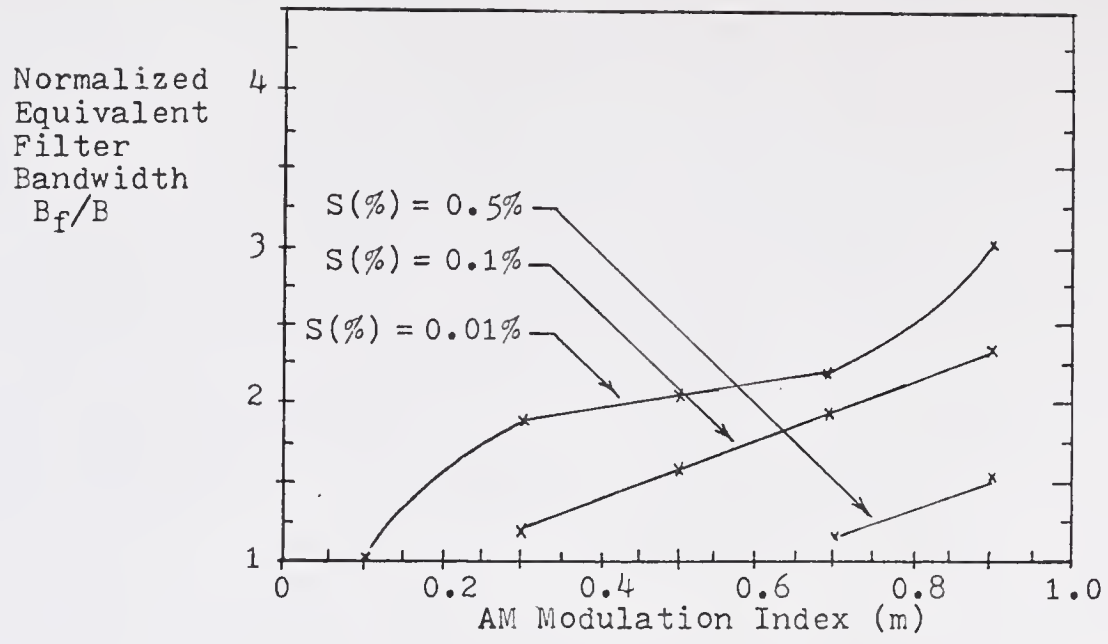


Figure 4.13. Relation between S(%), B_f, and m for the CSSB-AM p(t) function for the AM/PM modulator.

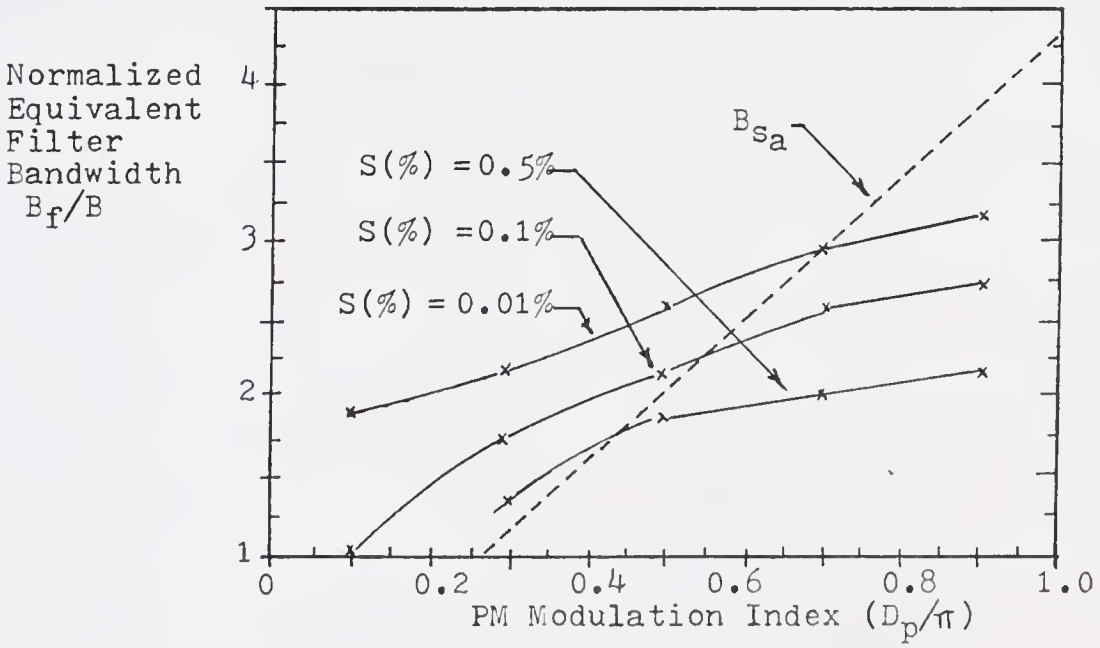


Figure 4.14. Relation between S(%), B_f, and D_p for the CSSB-PM a(t) function for the AM/PM modulator.

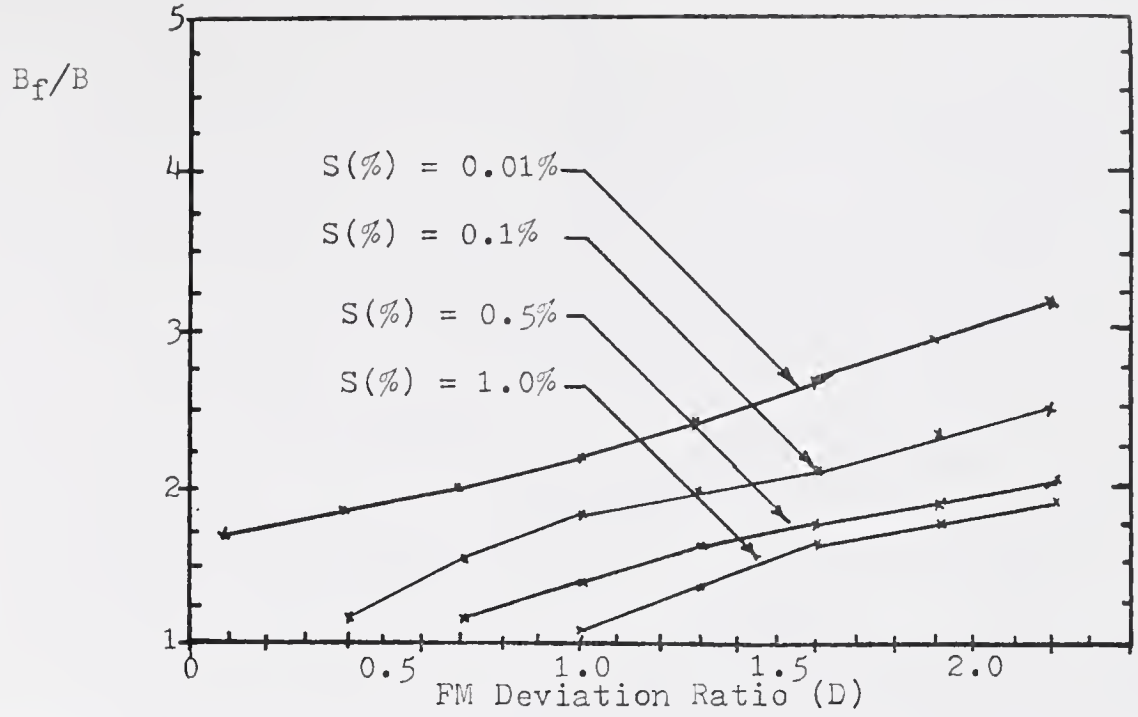


Figure 4.15. Relation between B_f, D, and S(%) for the CSSB-FM a(t) function for the AM/PM modulator.

Table IV-2. Equivalent-Filter Bandwidth Requirements for the a(t) and p(t) Signals in the AM/PM Modulator.

Type of Modulation	Equivalent Filter Bandwidth (Hz.)	
	a(t)	p(t)
AM	B	0
DSB-LM	Figure 4.11	Figure 4.11
SSB-LM	Figure 4.11	Figure 4.11
PM	0	B
FM	0	B
CSSB-AM	B	Figure 4.13
CSSB-PM	Figure 4.14	B
CSSB-FM	Figure 4.15	B

trend. The bandwidth requirements increase with the modulation parameters. In the worst case, an equivalent-filter bandwidth of 2 or 3 times the original message bandwidth is satisfactory.

The equivalent-filter bandwidth requirements of all the cases considered in this section are summarized on Table IV-2.

4.2.2 Dynamic Range

Since physical modulators have definite maximum input levels, it is clear that the modulation parameters may be constrained to a particular range. Let E_b and E_p be the maximum input allowed to the balanced modulator and the phase modulator respectively. Using Equations (4-8) and (4-9) it is clear that the maximum values allowed for a(t) and p(t) are

$$\max[a(t)] \leq K_a E_b \tag{4-36}$$

and

$$\max[p(t)] \leq K_c E_p, \tag{4-37}$$

where $K_c E_p$ is the maximum phase shift produced by the phase modulator.

Assume that the message is normalized to unit peak amplitude. The constraints on the unmodulated carrier amplitude, C, and the modulation parameters can be found by substituting the expressions listed in Table IV-1 for the a(t) and p(t) functions.

Consider the AM a(t), Equation (4-36) requires that

$$\max[C[1 + mx(t)]] \leq K_a E_b \tag{4-38}$$

but since

$$\max[x(t)] = 1, \tag{4-39}$$

this requires that

$$C \leq \frac{K_a E_b}{1 + \max(m)}, \tag{4-40}$$

if m is specified, or

$$\max(m) \leq \frac{K_a E_b - C}{C}, \tag{4-41}$$

when C is specified. Since p(t) is zero, no constraint is placed by the phase modulator.

Take the DSB-LM case, Equation (4-36) requires that

$$\max[C\,|x(t)|] \leq K_c E_b, \tag{4-42}$$

so using Equation (4-39), the carrier peak amplitude, C, is constrained to

$$\max(C) \leq K_c E_b. \tag{4-43}$$

The phase function has only two values, 0 or π; therefore, Equation (4-37) requires that

$$\pi \leq K_c E_p \tag{4-44}$$

so the phase modulator should be capable of producing at least a $\pm 180^{o}$ phase shift.

The SSB-LM case has a new mathematical constraint. The peak value of $\hat{x}(t)$ depends on the waveform of x(t). Squires

and Bedrosian [50] studied the problem of the peak to average power ratio for a deterministic signal and found that it can vary from unity for a sine wave to infinite for a square wave. It is only possible to define

$$\hat{x}_p = \max[\hat{x}(t)], \tag{4-45}$$

where $\hat{x}_p$ is the Hilbert transform peak amplitude. The constraint on the carrier constant, C, can be found in terms of $\hat{x}_p$. Substitute the expression for a(t) for the SSB-LM case in Equation (4-36)

$$\max\left[C\left[x^2(t) + \hat{x}^2(t)\right]^{\frac{1}{2}}\right] \leq K_c E_b, \tag{4-46}$$

using Equations (4-39) and (4-45), Equation (4-46) simplifies to

$$\max(C) \leq \frac{K_a E_b}{\left[1 + \hat{x}_p^2\right]^{\frac{1}{2}}}. \tag{4-47}$$

Since C is related to the average power of the SSB-LM carrier, Equation (4-47) puts a limit on how high this power can be. The phase function is constrained to the $\pm 180^{\circ}$ because $\tan^{-1}(\cdot)$ is constrained to that range; therefore, the phase modulator must satisfy Equation (4-44).

The PM and FM magnitude functions are equal to the constant C. The only restriction placed on C is that

$$\max(C) \leq K_a E_b. \tag{4-48}$$

To obtain the restrictions on the PM modulation index, D_p,

and the FM frequency deviation, D_f, substitute the expressions for p(t) in Equation (4-37)

$$\max[D_p x(t)] \leq K_c E_p \tag{4-49}$$

for PM and

$$\max\left[D_f \int_{-\infty}^{t} x(u)du\right] \leq K_c E_p \tag{4-50}$$

for FM. The restriction placed on D_p is easily found

$$\max(D_p) \leq K_c E_p, \tag{4-51}$$

but for the FM case a situation similar to the Hilbert transform problem arises. There is no way to determine the peak value of the integral of a function without first knowing the function; therefore, the best that can be done is to rewrite Equation (4-50) as

$$\max(D_f) \leq \frac{K_c E_p}{\max\left[\int_{-\infty}^{t} x(u)du\right]}. \tag{4-52}$$

The CSSB-AM a(t) function, and the CSSB-PM and CSSB-FM p(t) functions are identical to the AM a(t) function and the PM and FM p(t) functions; therefore, these cases are already known. The CSSB-AM p(t) restriction is found using Equation (4-37)

$$\max\left[H\{\ln[1 + mx(t)]\}\right] \leq K_c E_p \tag{4-53}$$

which depends on the waveform of x(t).

The CSSB-PM a(t) function requires that

$$\max\left[C \exp[-D_p \hat{x}_p]\right] \leq K_c E_b. \tag{4-54}$$

If

$$\max\left[\hat{x}_p\right] = \max\left[-\hat{x}_p\right], \tag{4-55}$$

then using Equation (4-45) the restrictions on D_p and C are

$$\max(D_p) \leq \frac{1}{\hat{x}_p} \ln\left[\frac{K_a E_b}{C}\right], \tag{4-56}$$

if C is specified, and

$$\max(C) \leq K_a E_b \exp\left[-D_p \hat{x}_p\right], \tag{4-57}$$

if D_p is given.

Observe that Equation (4-51) and (4-56) are restricting the value of D_p. Since both Equations must be satisfied, this means that the smallest of the two values of $\max(D_p)$ should be used.

The CSSB-FM magnitude function is also restricting the value of C and D_f. Substitute the expression for a(t) in Equation (4-36) and observe that

$$\max\left[C \exp\left[-D_f \int_{-\infty}^{t} x(u)du\right]\right] \leq K_a E_b, \tag{4-58}$$

where the value of C is restricted by

$$\max(C) \leq K_a E_b \exp\left[-D_f \max\left[\int_{-\infty}^{t} \hat{x}(u)du\right]\right] \tag{4-59}$$

and

$$\max(D_f) \leq \frac{\ln\left[\frac{K_a E_b}{C}\right]}{\max\left[\int_{-\infty}^{t} \hat{x}(u)du\right]}. \tag{4-60}$$

All these inequalities point out that there are definite

restrictions on the carrier constant, C, and the modulation parameters m, D_p, and D_f due to the modulators limited input range. The constraints can be expressed in terms of the message waveform in those cases where integration is involved.

4.3 The Quadrature Modulator

The quadrature modulator is illustrated in Figure 4.6. The expressions for the transfer characteristics of the baseband preprocessors #3 and #4 are similar to the equations obtained in Chapter III for the quadrature function components of the complex envelope. These equations for the baseband preprocessor are listed in Table IV-3.

4.3.1 Bandwidth Requirements

The bandwidth requirements for the signals i(t) and q(t) are described in terms of the equivalent-filter bandwidth which was defined in Section 2.5.1. The bandwidths are calculated following the procedure outlined in Section 4.1.5.

The equivalent-filter bandwidth of i(t) and q(t) are easily calculated for the AM, DSB-LM, and the SSB-LM cases. The bandwidth can be defined for zero distortion if the ideal lowpass filter shown in Figure 4.9 is at least as wide as the absolute bandwidth of the message, B. In these cases, the equivalent-filter bandwidth is B.

The PM and FM cases requires a computer simulation of the system. The results of the simulation are presented as graphs relating the normalized equivalent-filter bandwidth, the modulation constants, and the intermodulation distortion. These plots are shown in Figures 4.16 and 4.17. Included in these

Table IV-3. Quadrature Component Functions for the Quadrature Modulator.

Type of Modulation	Baseband Preprocessor Transfer Characteristics	
	#3 i(t)	#4 q(t)
AM	$C[1 + mx(t)]$	0
LM	$Cx(t)$	0
SSB-LM	$Cx(t)$	$C\hat{x}(t)$
PM	$C\cos[D_p x(t)]$	$C\sin[D_p x(t)]$
FM	$C\cos[D_f \int x(u)du]$	$C\sin[D_f \int_{-\infty}^{t} x(u)du]$
CSSB-AM	$C[1 + mx(t)] \cdot \cos\{H[\ln[1 + mx(t)]]\}$	$C[1 + mx(t)] \cdot \sin\{H[\ln[1 + mx(t)]]\}$
CSSB-PM	$C\exp[-D_p\hat{x}(t)] \cdot \cos[D_p x(t)]$	$C\exp[-D_p\hat{x}(t)] \cdot \sin[D_p x(t)]$
CSSB-FM	$C\exp[-D_f \int_{-\infty}^{t} x(u)du] \cdot \cos[D_f \int_{-\infty}^{t} x(u)du]$	$C\exp[-D_f \int_{-\infty}^{t} \hat{x}(u)du] \cdot \sin[D_f \int_{-\infty}^{t} x(u)du]$

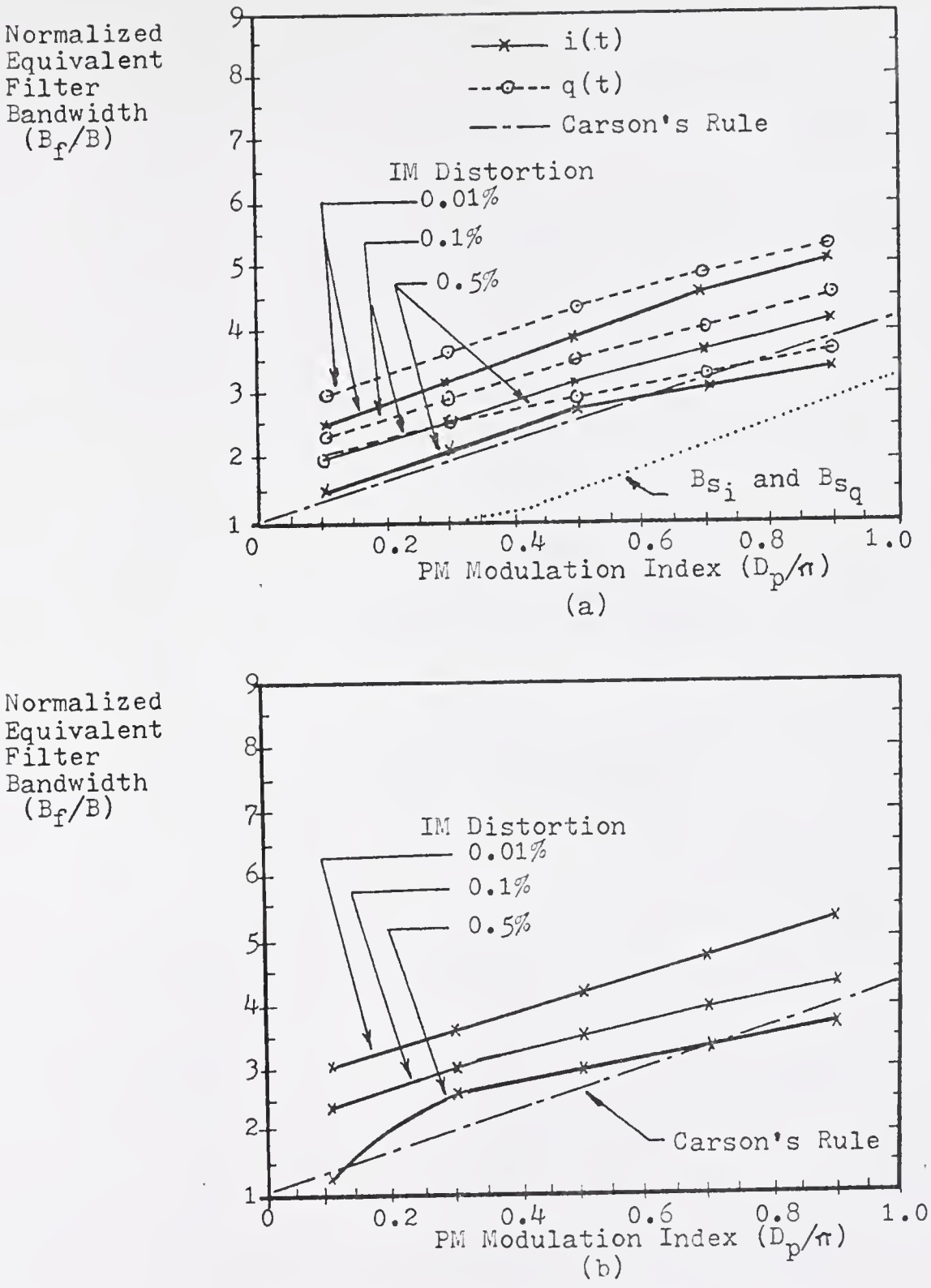
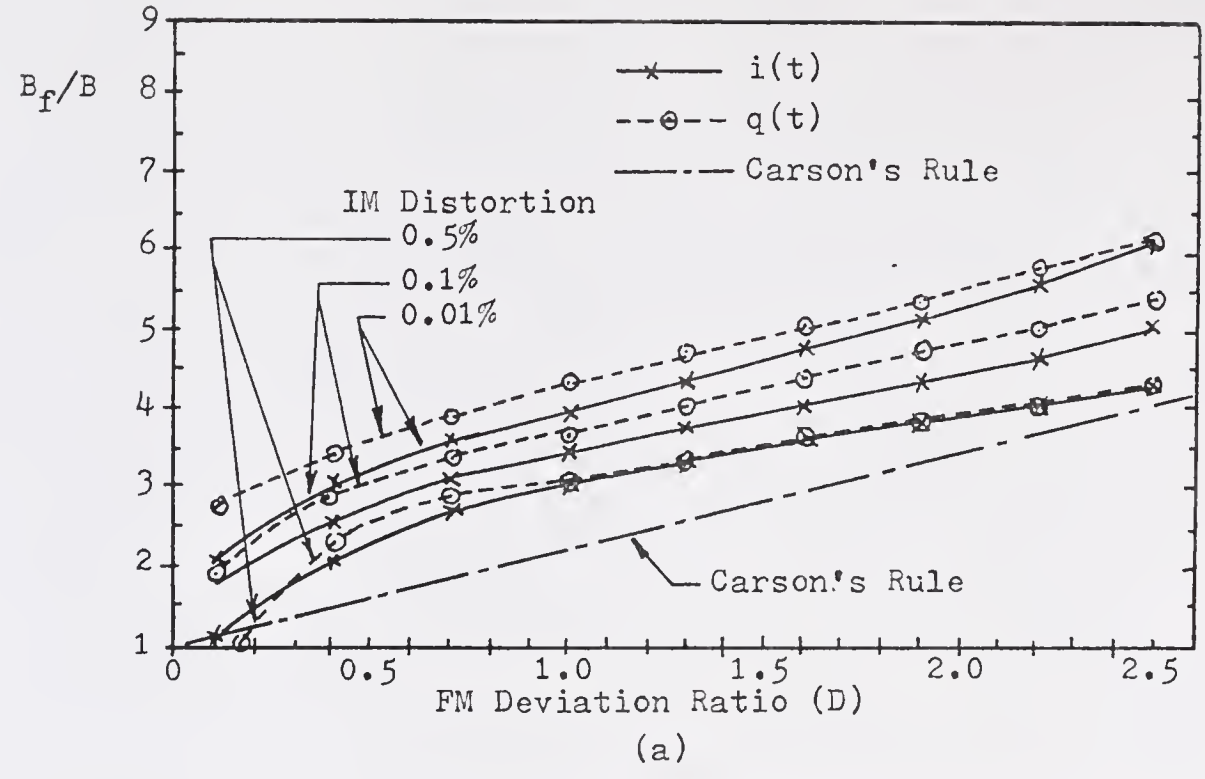


Figure 4. 16. Relation between B_f, D_p, and IM distortion for the PM quadrature components, (a) only one of the components is filtered, and (b) when both components are filtered.



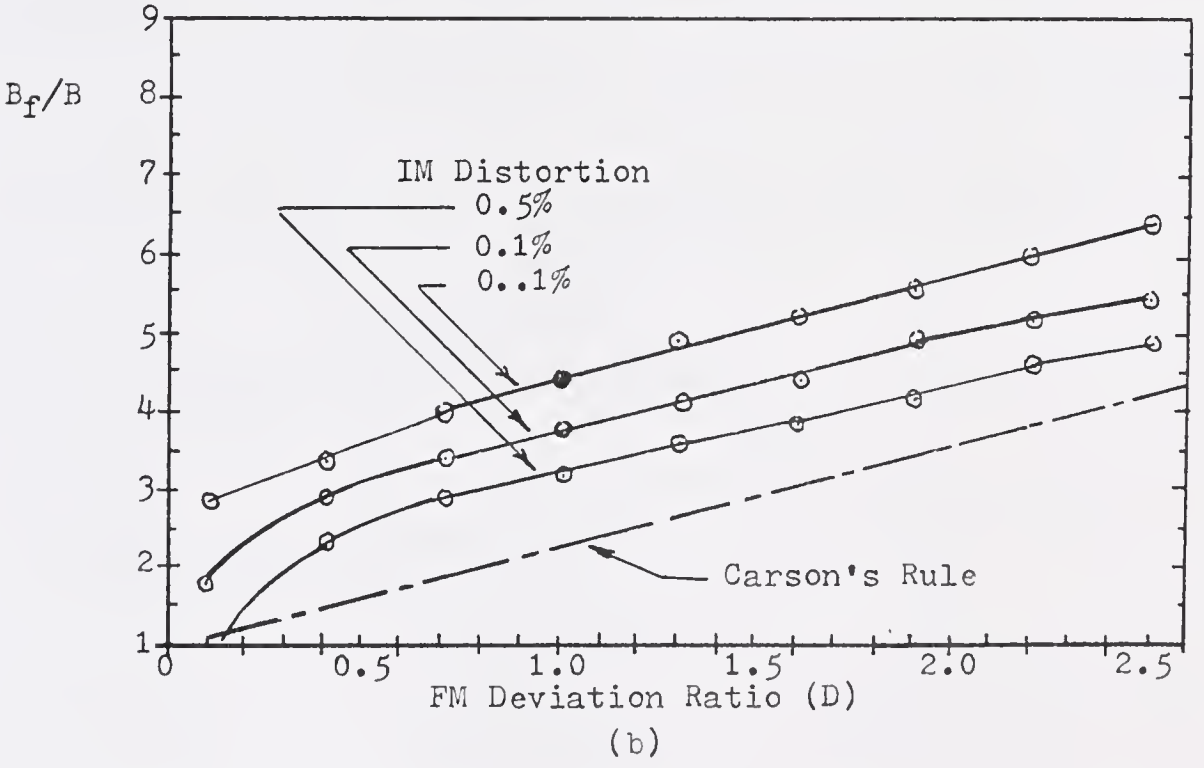


Figure 4.17. Relation between B_f, D, and IM distortion for the FM quadrature components, (a) when only one component is filtered, and (b) when both components are filtered.

figures is the Carson's rule bandwidth B_c [1], defined as

$$B_c = (D_p + 1)B \tag{4-61}$$

for PM and

$$B_c = (D + 1)B \tag{4-62}$$

for FM. From the curves it is seen that the Carson's rule approximation is poor in the two cases. This rule can be modified by adding a constant to every point in the line; therefore, Equation (4-61) is modified as

$$B_c = (D_p + 2)B. \tag{4-63}$$

This equation is in better agreement with the 0.1% curves shown in Figure 4.16(a) and Figure 4.17(b). This was also observed by Carlson [1].

If Equations (3-149) and (3-150) are plotted in the same axes of Figure 4.16(a), a plot is obtained for the normalized second moment bandwidth of the i(t) and q(t) functions for the PM case. Unfortunately the direct application of these equations to design the system will result in a very distorted demodulated message. These equations show the same general trend as the curves based on distortion but they do not agree even after adding an arbitrary constant. This reduces the usefulness of these equations and makes them less attractive for the design of this system.

If the 0.1% IM distortion curve is selected as the design criterion, a set of equations can be found graphically that

relates the approximate bandwidth requirements to the PM modulation index. These Equations are

$$B_{f_i} \approx \left[\frac{2.6}{\pi} (D_p) + 1.8\right] B \qquad \text{(PM)} \qquad (6\text{-}64)$$

and

$$B_{f_q} \approx \left[\frac{2.6}{\pi} (D_p) + 2.2\right] B. \qquad \text{(PM)} \qquad (4\text{-}65)$$

When both i(t) and q(t) are filtered, the equation is

$$B_{f_{iq}} \approx \left[\frac{2.1}{\pi} (D_p) + 2.4\right] B. \qquad \text{(PM)} \qquad (4\text{-}66)$$

The FM case requires

$$B_{f_i} \approx (1.15D + 2.4)B, \qquad \text{(FM)} \qquad (4\text{-}67)$$

and

$$B_{f_q} \approx (1.15D + 2.6)B, \qquad \text{(FM)} \qquad (4\text{-}68)$$

and

$$B_{f_{iq}} \approx (1.23D + 2.5)B. \qquad \text{(FM)} \qquad (4\text{-}69)$$

The CSSB-AM IM distortion curves are shown in Figure 4.18. These curves seem to indicate that the i(t) and q(t) functions are almost bandlimited to twice the message bandwidth because the curves for 0.01%, 0.1%, and 0.5% fall very close of each other. The q(t) function requires less bandwidth than the i(t) function when tha AM modulation index, m, is low. For higher m, the curves for i(t) and q(t) are close. If both i(t)

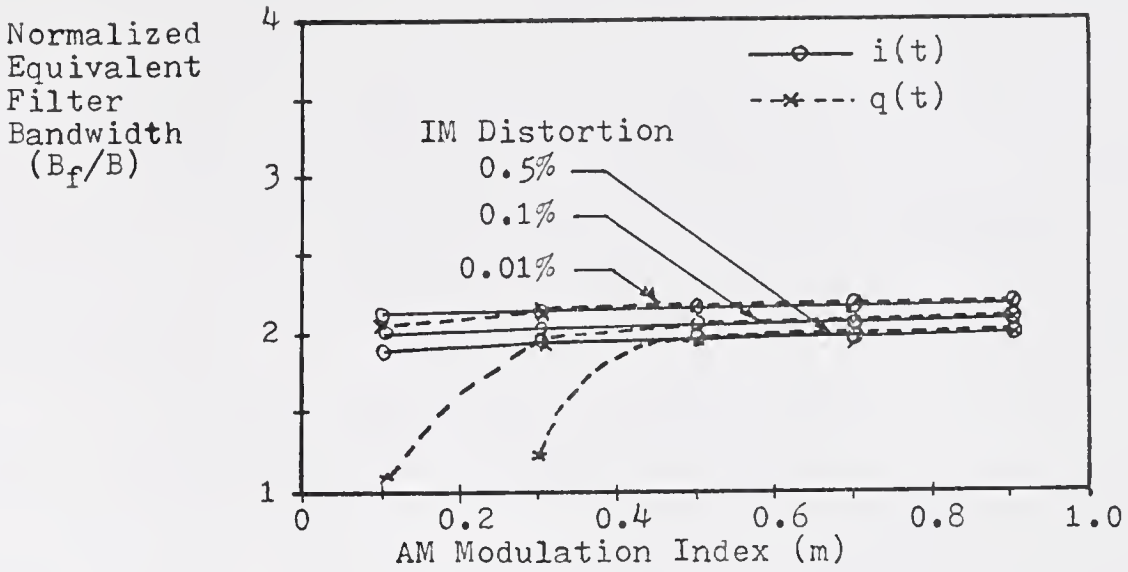
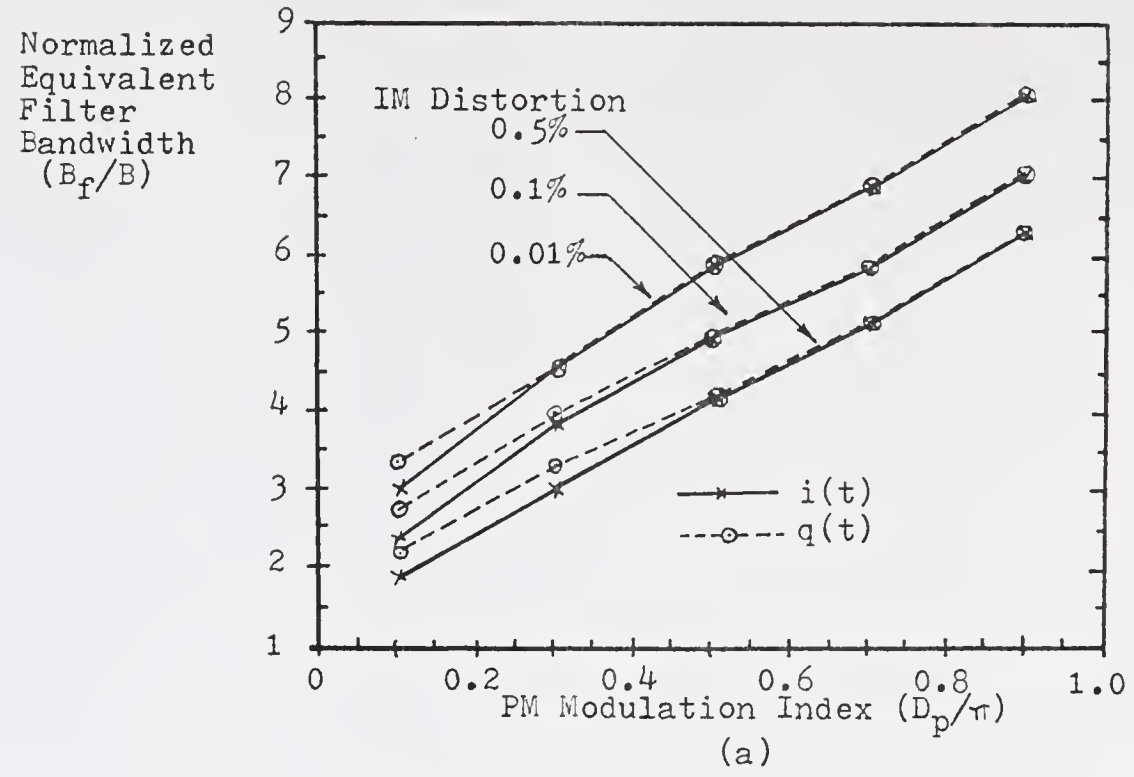


Figure 4.18. Relation between B_f, m, and IM distortion for the CSSB-AM quadrature components (all cases).

and q(t) are filtered, the required bandwidth closely approximates the curves for i(t). Due to the shape of the curves it is very difficult to find an equation for bandwidth, but 2.5B can be used as the design criterion.

The CSSB-PM and CSSB-FM cases are presented in Figures 4.19 and 4.20. If these curves are compared with Figures 4.16 and 4.17 for the PM and FM cases it seems that, individually, the CSSB cases require more bandwidth than their double-sided counterparts. However, if both i(t) and q(t) are filtered, the required bandwidth is less than that required by PM and FM. The reason for this is that i(t) and q(t) are related by the Hilbert transform. When both components are filtered (to the same filter bandwidth), the two components still form a Hilbert pair so it is reasonable to expect some distortion



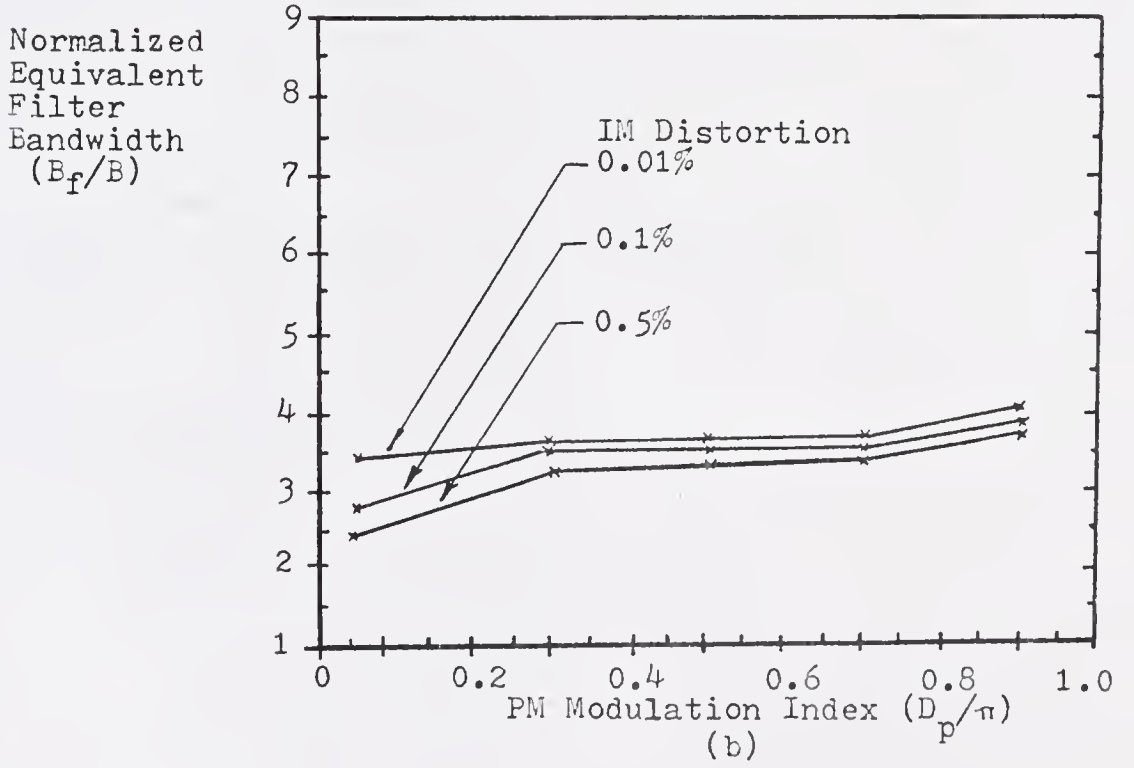


Figure 4.19. Relation between B_f, D , and IM distortion for the CSSB-PM quadrature components (a) when only one component is filtered, and (b) when both components are filtered.

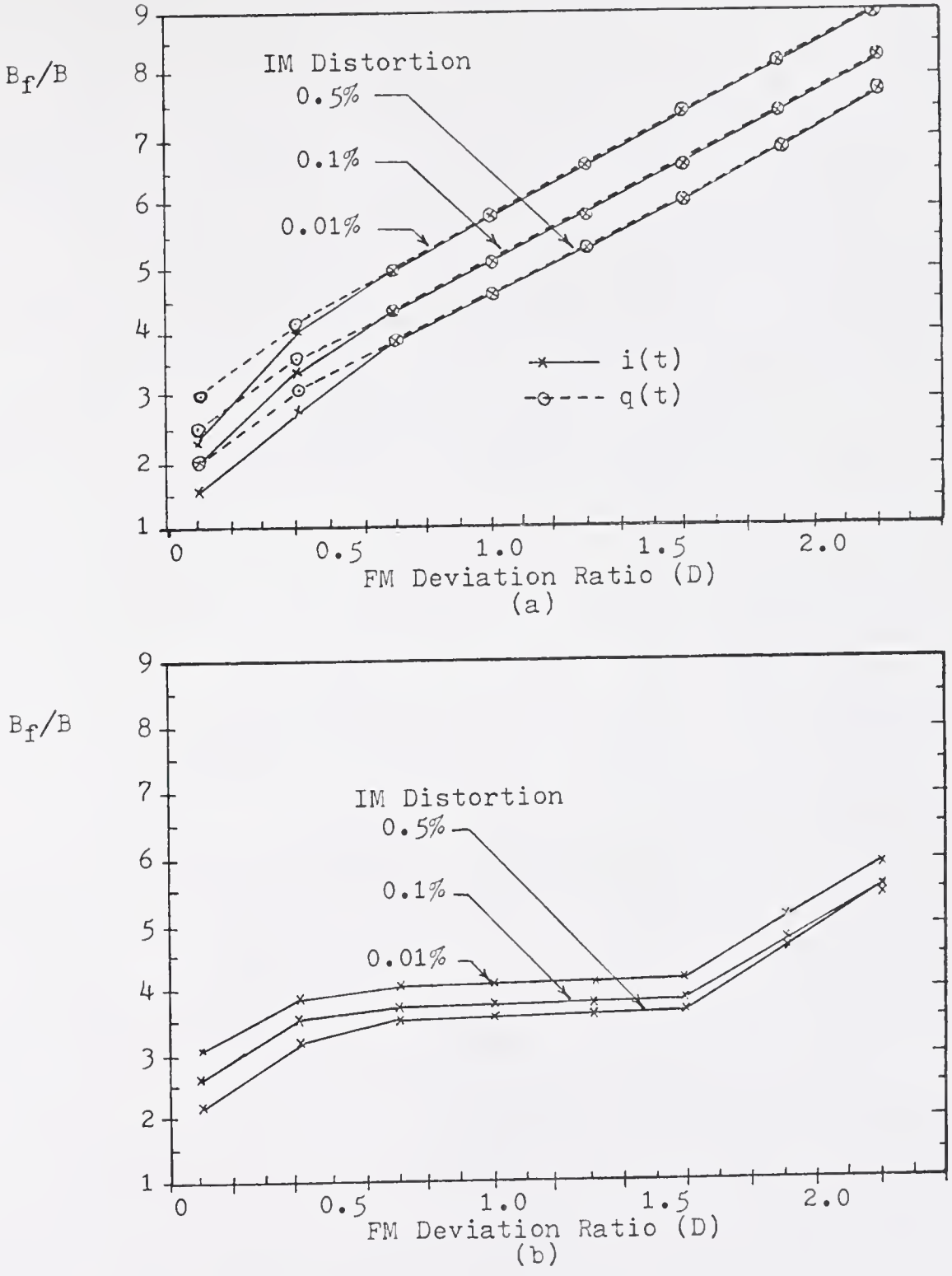


Figure 4.20. Relation between B_f, D, and IM distortion for the CSSB-FM quadrature components, (a) when only one component is filtered and (b) when both components are filtered.

cancellation due to redundant information in both signals.

Equations can be found graphically to describe the relationship between the normalized equivalent-filter bandwidth and the modulation parameters for CSSB-PM and CSSB-FM. These are

$$B_{f_i} \approx B_{f_q} \approx \left[\frac{5.1D_p}{\pi} + 2.4\right]B \qquad (4\text{-}70)$$

for the CSSB-PM case, and

$$B_{f_i} \approx B_{f_q} \approx (2.75D + 2.4)B \qquad (4\text{-}71)$$

for the CSSB-FM case. When both i(t) and q(t) are filtered it is better to refer to Figures 4.19(b) and 4.20(b) rather than to find a simple expression for the bandwidth.

The relationships between the sideband suppression factor, the modulation parameters and the equivalent filter bandwidth are shown in Figures 4.21 to 4.23 for the CSSB-PM and CSSB-FM cases. These curves are for the case where only i(t) or q(t) is filtered. If both signals are filtered the components are distorted but the complex envelope is still an Analytic signal; therefore, the spectrum is single-sided no matter what is the ideal filter. If these curves are compared with those obtained for distortion it is seen that the bandwidth measured in terms of distortion covers very well the sideband suppression cases. It is not necessary to consider the sideband suppression separately.

The equivalent-filter bandwidth requirements are summarized on Table IV-4.

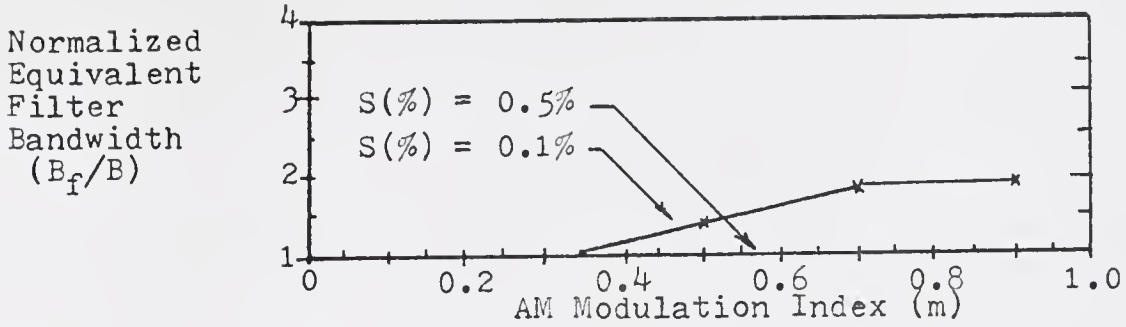


Figure 4.21. Relation between S(%), m, and B_f for the CSSB-AM case where only one of the quadrature components is filtered.

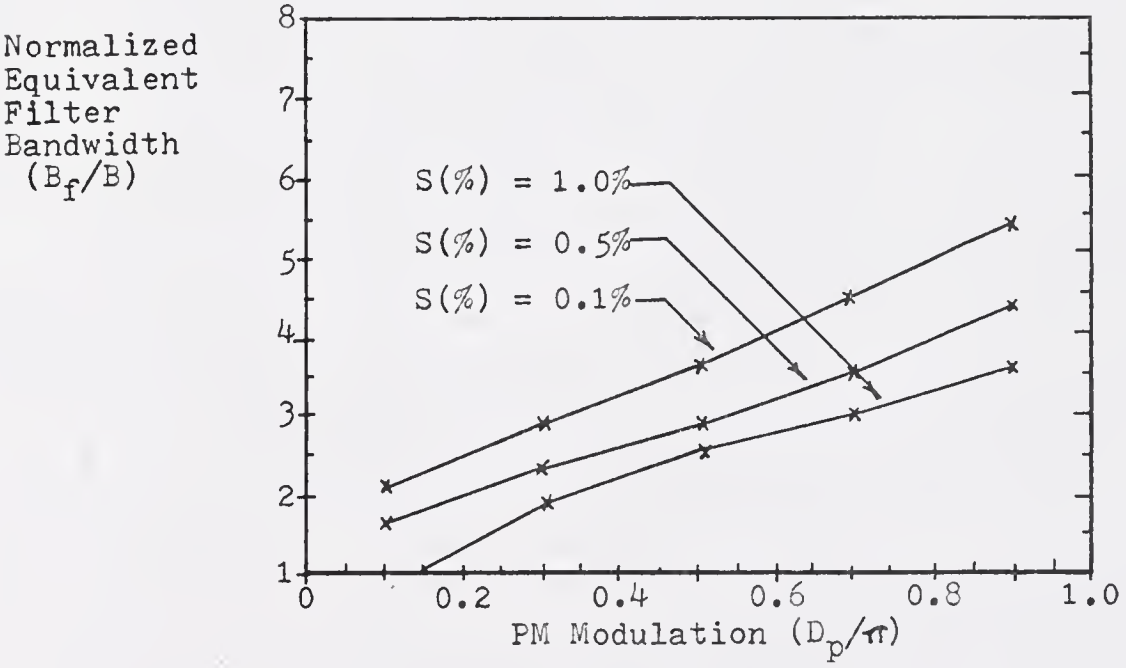


Figure 4.22. Relation between S(%), D_p, and B_f for the CSSB-PM case where only one of the quadrature component is filtered.

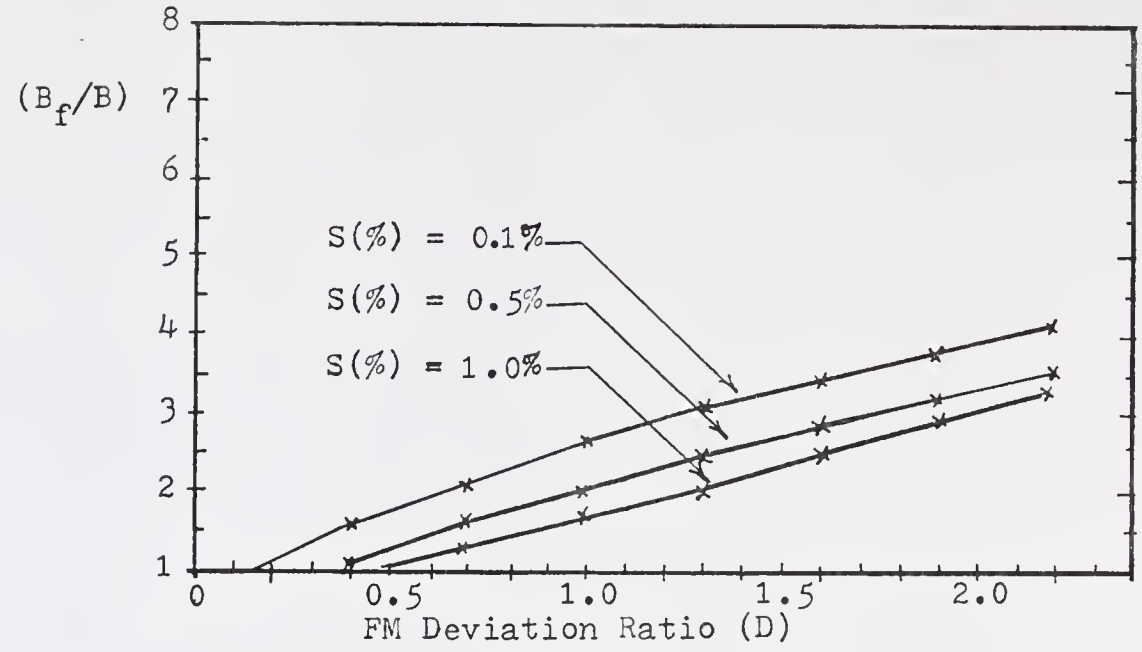


Figure 4.23. Relation between S(%), D, and B_f for the CSSB-FM case where one of the quadrature component is filtered.

4.3.2 Dynamic Range

Let E_b be the maximum input allowed to each balanced modulator. From Equations (4-12) and (4-13) it is clear that

$$\max[i(t)] \leq K_a E_b \tag{4-72}$$

and

$$\max[q(t)] \leq K_a E_b \tag{4-73}$$

where $K_a E_b$ is equivalent to the maximum peak amplitude of the carrier going out of the balanced modulator. The constraints on the modulation parameters can be found by substituting the expressions listed on Table IV-3 for the i(t) and q(t) functions.

Observe that a(t) and i(t) are identical for AM; therefore,

Table IV-4. Equivalent-Filter Bandwidth Requirements for the i(t) and q(t) Signals in the Quadrature Modulator.

Type of Modulation	Equivalent Filter Bandwidth[☆]	
	i(t)	q(t)
AM	B	0
DSB-LM	B	0
SSB-LM	B	B
PM	$\left[\frac{2.6D}{\pi}p + 1.8\right]B$	$\left[\frac{2.6D}{\pi}p + 2.2\right]B$
FM	$(1.15 + 2.4)B$	$(1.15D + 2.6)B$
CSSB-AM	$2.5B$	$2.5B$
CSSB-PM	$\left[\frac{5.1D}{\pi}p + 2.4\right]B$	$\left[\frac{5.1D}{\pi}p + 2.4\right]B$
CSSB-FM	$(2.75D + 2.4)B$	$(2.75D + 2.4)B$

☆ For 0.1% intermodulation distortion at the output of an ideal receiver under zero noise conditions.

Equations (4-40) and (4-41) still apply. Since q(t) is zero, no constraint is imposed by Equation (4-73).

The DSB-LM requires that

$$\max\left[C\,|x(t)|\right] \leqslant K_a E_b. \tag{4-74}$$

Substitute Equation (4-39) and Equation (4-74) simplifies to

$$\max(C) \leqslant K_a E_b. \tag{4-75}$$

This is the same as Equation (4-43) obtained for the DSB-LM magnitude function. Since q(t) is zero, Equation (4-73) does not affect the case under study.

For SSB-LM Equations (4-72) and (4-73) require that

$$\max\left|Cx(t)\right| \leqslant K_a E_b \tag{4-76}$$

for the i(t) function and

$$\max\left|C\hat{x}(t)\right| \leqslant K_a E_b \tag{4-77}$$

for the q(t) function. These inequalities simplify to

$$\max(C) \leqslant K_a E_b \tag{4-78}$$

and

$$\max(C) \leqslant \frac{K_a E_b}{\hat{x}_p} \tag{4-79}$$

where $\hat{x}_p$ was defined by Equation (4-45). Since Equations (4-78) and (4-79) must be satisfied, it is evident that the last inequality must prevail.

The PM and FM cases are very simple. Observe that the

sine and cosine functions are restricted to the range [-1, 1]; therefore, when the expressions for i(t) and q(t) are substituted in Equations (4-72) and (4-73), these inequalities reduce to

$$\max(C) \leqslant K_a E_b \tag{4-80}$$

Observe that unlike the AM/PM modulator case, no restrictions are placed on the PM modulation index or the FM frequency deviation.

Now consider the CSSB cases. Observe on Table IV-3 that the i(t) and q(t) components have the same general form:

$$i(t) = a(t)\cos[p(t)] \tag{4-81}$$

and

$$q(t) = a(t)\sin[p(t)] \tag{4-82}$$

where a(t) and p(t) are given on Table IV-1. Substitute Equations (4-81) and (4-82) into Equations (4-76) and (4-77)

$$\max\left[\left|a(t)\cos[p(t)]\right|\right] \leqslant K_a E_b \tag{4-83}$$

and

$$\max\left[\left|a(t)\sin[p(t)]\right|\right] \leqslant K_a E_b \tag{4-84}$$

so under the worst possible conditions, these inequalities simplify to

$$\max[a(t)] \leqslant K_a E_b. \tag{4-85}$$

This inequality means that i(t) and q(t) constraint the CSSB

modulations in the same way as a(t) did for the AM/PM modulator. Observe that the restrictions imposed by p(t) in the AM/PM modulator are not present for the quadrature modulator.

All these inequalities show that the restrictions imposed by the quadrature modulator are identical to those imposed by the AM/PM modulator except for the PM and FM cases where no restrictions are placed on D_p or D_f and for the SSB-LM carrier constant C where the quadrature modulator allows a higher value of C (Equations (4-47) and (4-79)).

4.4 The PM/PM Modulator

The block diagram of the PM/PM modulator is shown in Figure 4.7. The transfer characteristics between the input message, x(t), and the two phasing functions, $\theta_1(t)$ and $\theta_2(t)$, can be derived from the magnitude and phase of the complex envelope. From Figure 4.2(e) it is seen that

$$v(t) = L \exp[j\theta_1(t)] + L \exp[j\theta_2(t)] \tag{4-86}$$

where L corresponds to the PM carrier peak amplitude. This equation has a magnitude and phase functions given by

$$a(t) = 2L \cos\left[\frac{\theta_1(t) - \theta_2(t)}{2}\right] \tag{4-87}$$

and

$$p(t) = \frac{\theta_1(t) + \theta_2(t)}{2} \, . \tag{4-88}$$

Solving for $\theta_1(t)$ and $\theta_2(t)$ in terms of a(t) and p(t) yields

$$\theta_1(t) = p(t) + \cos^{-1}\left[\frac{a(t)}{2L}\right] \tag{4-89}$$

and

$$\theta_2(t) = p(t) - \cos^{-1}\left[\frac{a(t)}{2L}\right]. \tag{4-90}$$

All these equations are valid provided that

$$a(t) \leq 2L. \tag{4-91}$$

Equations (4-89) and (4-90) can be written as

$$\theta_1(t) = p(t) + \phi(t) \tag{4-92}$$

and

$$\theta_2(t) = p(t) - \phi(t), \tag{4-93}$$

where

$$\phi(t) = \cos^{-1}\left[\frac{a(t)}{2L}\right]. \tag{4-94}$$

The function $p(t)$ is usually limited to the range $[-\pi, \pi]$. Since $a(t)$ is always positive, $\phi(t)$ is limited to

$$0 \leq \phi(t) \leq \pi/2; \tag{4-95}$$

therefore, the phase functions $\theta_1(t)$ and $\theta_2(t)$ are limited to

$$-\pi \leq \theta_1(t) \leq \frac{3\pi}{2} \tag{4-96}$$

and

$$\frac{-3\pi}{2} \leq \theta_2(t) \leq \pi . \tag{4-97}$$

This may be of concern because it may place additional burden on the specifications for the phase modulators. It is evident

that the block diagram of Figure 4.7 may not be a favorable solution.

Observe that Equations (4-92) and (4-93) consist of the same two terms, taken as the sum in one case and as the difference in the other. It is inefficient to generate twice p(t) and $\phi(t)$; therefore, a better solution is to generate $\phi(t)$ and p(t) separately and combine them as necessary. Since the form of Equation (4-92) and (4-93) is independent of the actual $\phi(t)$ or p(t), it is independent of the modulation law and can be made part of the modulation-invariant circuits. This is illustrated in Figure 4.24. This still may not solve the problem of the limited phase shift range of the phase modulators.

A solution that avoids these limitations is to implement directly Equation (4-86) by observing that it can be written as

$$v(t) = L \exp[jp(t)] \cdot [\exp[j\phi(t)] + \exp[-j\phi(t)]] . \quad (4\text{-}98)$$

This suggests the block diagram shown in Figure 4.25 but this system requires three phase modulators. The top two phase modulators form the amplitude modulator configuration known as "Ampliphase".

The equations for $\phi(t)$ are obtained by direct substitution of the equations listed in Table IV-1 into Equations (4-93) and they are listed in Table IV-5. The equations for p(t) are already listed in Table IV-1 but are repreated for completeness.

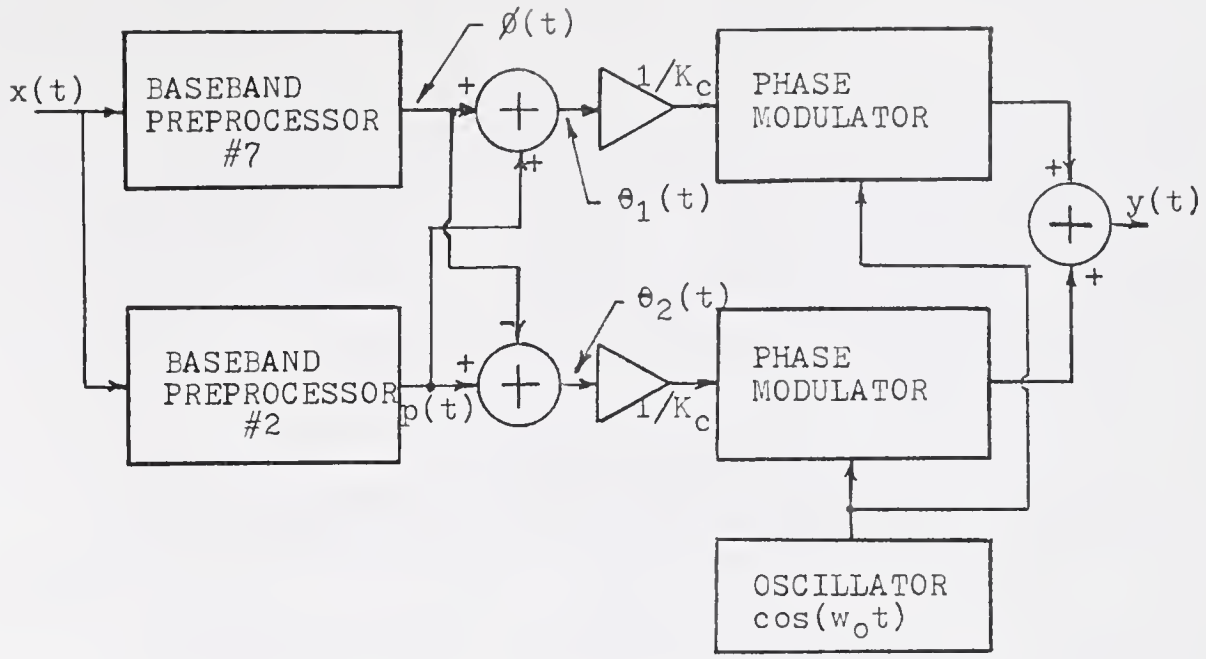


Figure 4.24. Another PM/PM modulator.

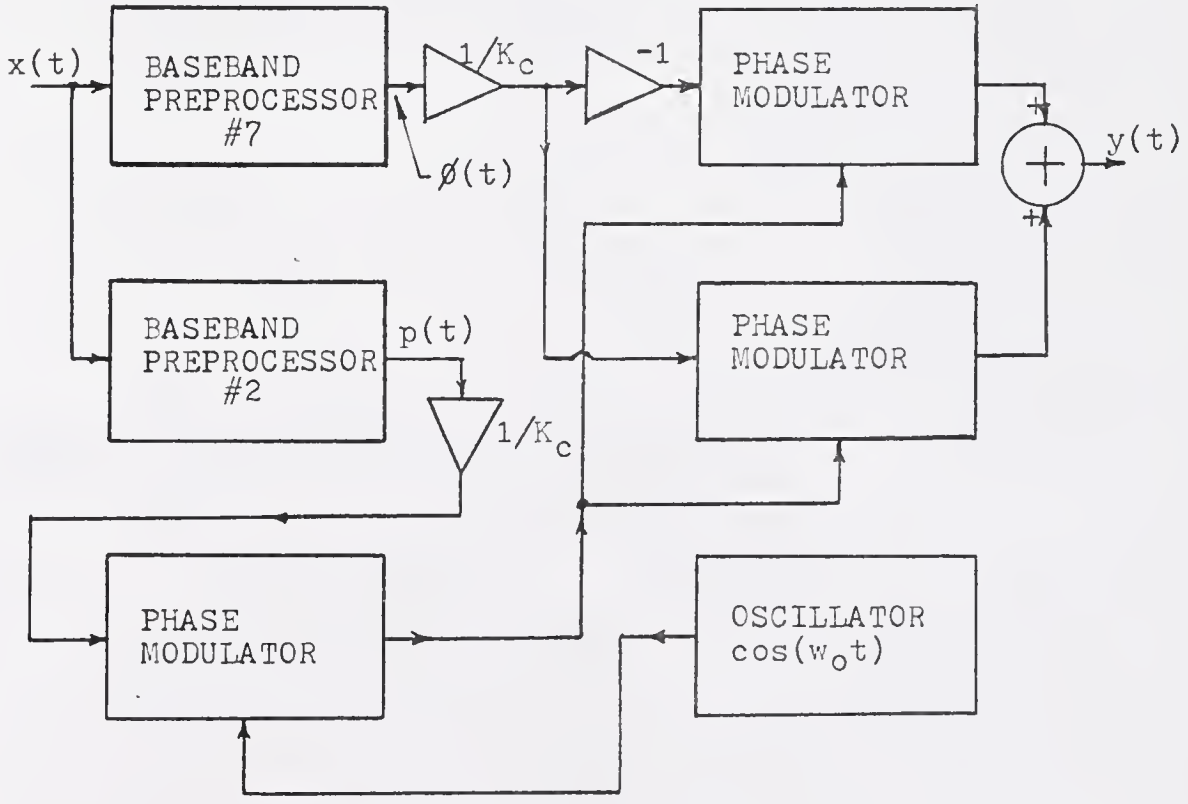


Figure 4.25. The best PM/PM modulator.

Table IV-5. The $\phi(t)$ and $p(t)$ Functions for the PM/PM Modulator.

Type of Modulation	Baseband Preprocessor Transfer Characteristics	
	#7 $\phi(t)$	#2 $p(t)$ ✱
AM	$\cos^{-1}\left[\frac{C[1 + mx(t)]}{2L}\right]$	0
DSB-LM	$\cos^{-1}\left[\frac{C\lvert x(t)\rvert}{2L}\right]$	$\frac{\pi}{2}[1 - \text{sgn}[x(t)]]$
SSB-LM	$\cos^{-1}\left[\frac{C\sqrt{x^2(t) + \hat{x}^2(t)}}{2L}\right]$	$\tan^{-1}[\hat{x}(t)/x(t)]$
PM	$\cos^{-1}\left[\frac{C}{2L}\right]$	$D_p x(t)$
FM	$\cos^{-1}\left[\frac{C}{2L}\right]$	$D_f \int_{-\infty}^{t} x(u)\,du$
CSSB-AM	$\cos^{-1}\left[\frac{C[1 + mx(t)]}{2L}\right]$	$H\{\ln[1 + mx(t)]\}$
CSSB-PM	$\cos^{-1}\left[\frac{C}{2L}\exp[-D_p\hat{x}(t)]\right]$	$D_p x(t)$
CSSB-FM	$\cos^{-1}\left[\frac{C}{2L}\exp\left[-D_f\int_{-\infty}^{t}\hat{x}(u)\,du\right]\right]$	$D_f \int_{-\infty}^{t} x(u)\,du$

✱ Obtained from Table IV-1.

4.4.1 Bandwidth Requirements

In this section it is only necessary to study the bandwidth required for $\phi(t)$ because the p(t) case has already been studied in Section 4.2.1 for the AM/PM modulator. Preliminary work showed that when both $\phi(t)$ and p(t) were filtered, the resulting distortion curve was about the same as the highest distortion produced by the individual components. For this reason, the case where both components are bandlimited is not studied in detail.

There is a parameter that must be watched with care. It is the ratio 2L/C that appears as part of the argument of $\cos^{-1}(\cdot)$. The argument of $\cos^{-1}(\cdot)$ must be in the range[-1,1] or the function is not defined. This can be satisfied if the argument is assigned a value of 1 whenever it is larger than 1, and -1 whenever it exceeds -1, but this causes clipping of the signal in the argument of $\cos^{-1}(\cdot)$ and distortion is observed at the receiver. If the value of the argument never exceeds these limits, the ratio of 2L/C has little effect in the equivalent filter bandwidth. A ratio of 4 was selected for all for all the calculations except for CSSB-FM where a ratio of 16 was used in order to increase the deviation ratio range so that it can be compared with other cases studied before. This ratio of 16 is too high for most applications because the phase modulator will have to be handling 64 times the power of the unmodulated carrier.

Observe on Table IV-5 that AM and CSSB-AM have the same $\phi(t)$ function. This means that they can be considered together.

The relationship between the required equivalent-filter bandwidth and the AM modulation index is shown as a graph in Figure 4.26. A line can be drawn to approximate the required equivalent-filter bandwidth

$$B_{f_\phi} \approx (0.8m + 1.7)B. \tag{4-99}$$

The graph for the CSSB-AM p(t) function was shown in Figure 4.13. The $\phi(t)$ function has no noticeable effect in the single sided property of CSSB-AM.

The plots for the DSB-LM and SSB-LM $\phi(t)$ functions are shown in Figure 4.27. The bandwidth requirements are close to those of a(t) shown in Figure 4.11. This means that $\cos^{-1}(\cdot)$ does not produce much bandwidth expansion. This can be explained by observing that $\cos^{-1}(\cdot)$ can be described by a line segment when the argument is small. This is shown in Figure 4.28.

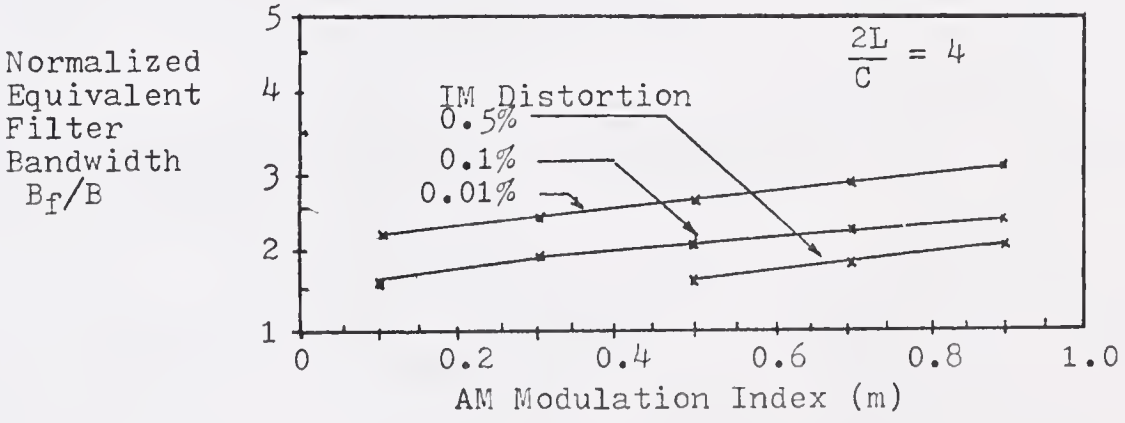


Figure 4.26. Relation between B_f, m, and IM distortion for the AM and CSSB-AM $\phi(t)$ function.

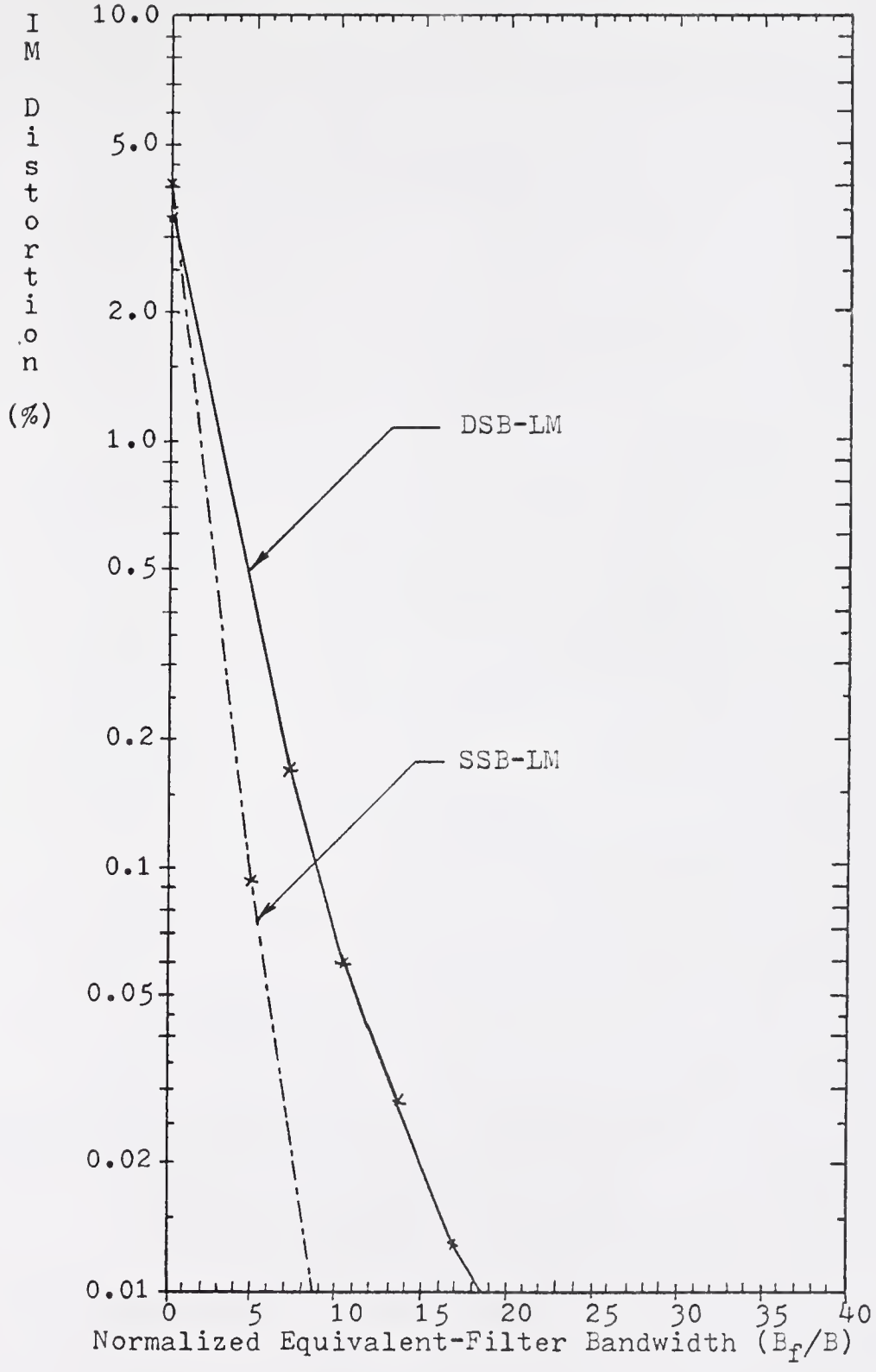


Figure 4.27. IM distortion for the DSB-LM and SSB-LM $\phi(t)$ component as a function of B_f.

The ideal receiver ignores the amplitude of the CSSB-PM and CSSB-FM modulated carriers. If there is no clipping in the argument of $\cos^{-1}(\cdot)$ the $\phi(t)$ function has no effect at all at the receiver but it determines how the power is suppressed in the undesireable sideband. This is illustrated by Figures 4.29 and 4.30 for CSSB-PM and CSSB-FM respectively. For D_p larger than 0.8 and D larger than 1.6 the curves show a fast increase in the slope indicating that clipping is taking place.

The PM and FM $\phi(t)$ functions are constants, so the bandwidth is zero.

The bandwidth requirements are summarized in Table IV-6.

The critical bandwidths obtained for the DSB-LM and the SSB-LM p(t) functions; therefore, from the bandwidth requirements point of view, there is little difference between the AM/PM and the PM/PM modulators.

4.4.2 Dynamic Range

The ratio of 2L/C is important in determining the dynamic range of the modulation parameters in the $\phi(t)$ function. The dynamic range of the parameters in the p(t) function were already discussed with the AM/PM modulator.

The basic restriction imposed in the $\phi(t)$ function is due to the definition of $\cos^{-1}(\cdot)$

$$0 \leq \frac{a(t)}{2L} \leq 1, \qquad (4\text{-}100)$$

where L is the PM carrier peak amplitude. The lower limit is always satisfied since a(t) is always positive, so

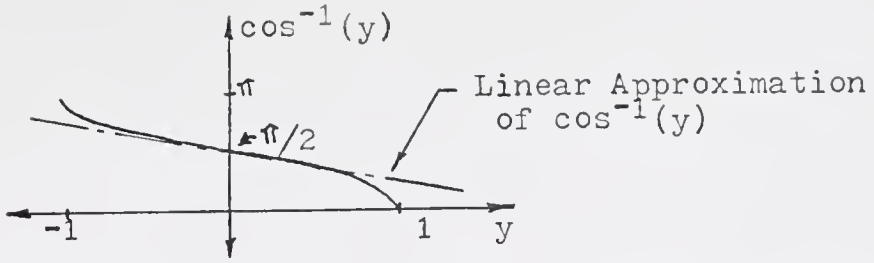


Figure 4.28. The function $\cos^{-1}(y)$ and a linear approximation.

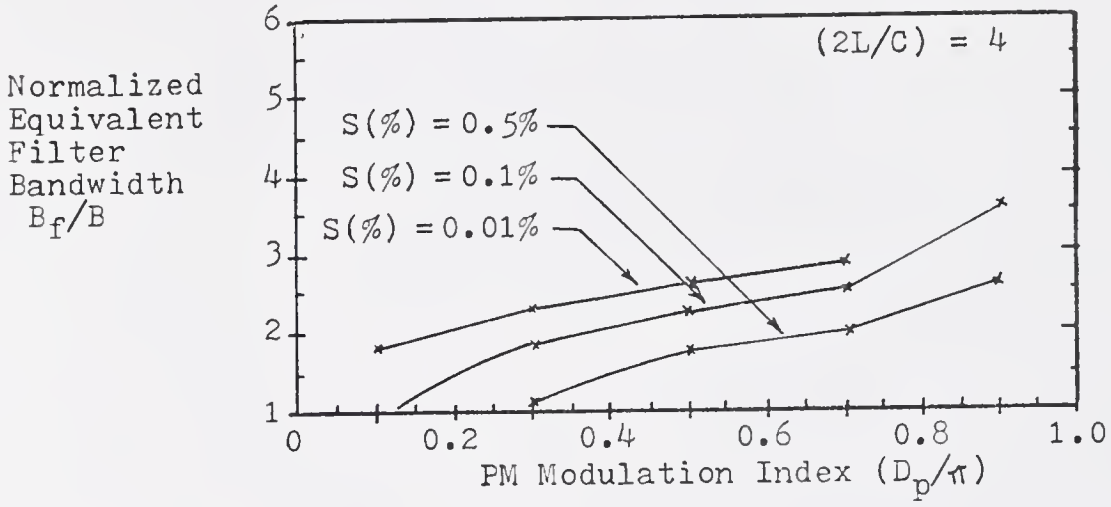


Figure 4.29. Relation between S(%), D_p, and B_f for the CSSB-PM $\phi(t)$ function.

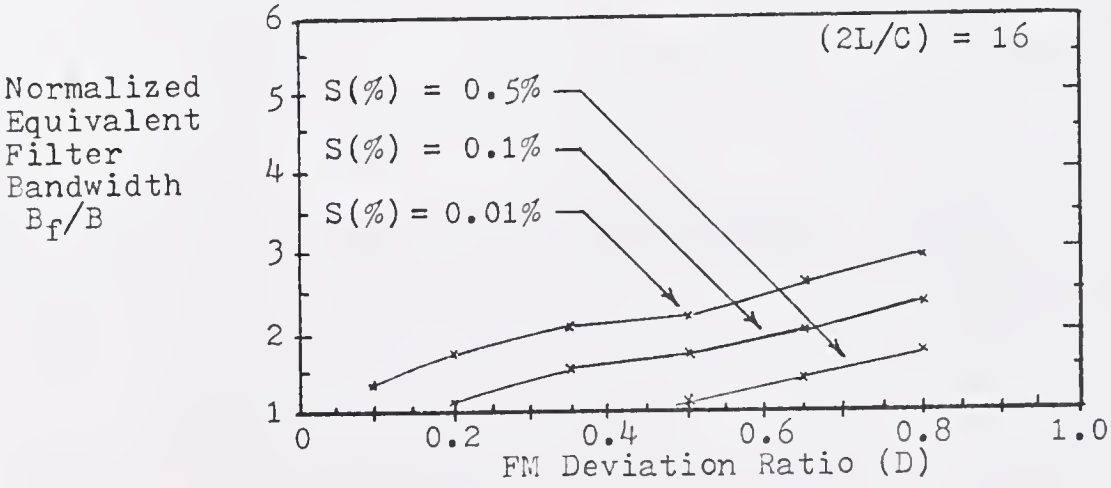


Figure 4.30. Relation between S(%), D, and B_f for the CSSB-FM $\phi(t)$ function.

Table IV-6. Equivalent-Filter Bandwidth Requirements for the $\phi(t)$ and $p(t)$ Signals in the PM/PM Modulator.

Type of Modulation	Equivalent-Filter Bandwidth	
	$\phi(t)$	$p(t)$[★]
AM	$(0.8m + 1.7)B$	0
DSB-LM	Figure 4.27	Figure 4.11
SSB-LM	Figure 4.28	Figure 4.11
PM	0	B
FM	0	B
CSSB-AM	$(0.8m + 1.7)B$	Figure 4.13
CSSB-PM	Figure 4.29	B
CSSB-FM	Figure 4.30	B

★ Obtained from Table IV-2.

$$a(t) \leq 2L. \tag{4-101}$$

This means that the results derived for the AM/PM modulator a(t) function in Section 4.2.2 can be used if $K_a E_b$ is replaced by 2L. In other words, the same results are obtained if

$$L = \frac{K_a E_b}{2} . \tag{4-102}$$

Since $\theta_1(t)$ and $\theta_2(t)$ are the actual inputs to the phase modulators in the system shown in Figure 4.24, the following conditions are set by Equation (4-37)

$$\max\left[|\theta_1(t)|\right] \leq K_a E_b \tag{4-103}$$

and

$$\max\left[|\theta_2(t)|\right] \leq K_c E_p \tag{4-104}$$

Since $\theta_1(t)$ and $\theta_2(t)$ are related to p(t) and $\phi(t)$ by Equations (4-92) and (4-93), Equations (4-103) and (4-104) can be written as

$$-K_c E_p \leq p(t) + \phi(t) \leq K_c E_p \tag{4-105}$$

and

$$-K_c E_p \leq p(t) - \phi(t) \leq K_c E_p \tag{4-106}$$

which cannot be explicitely solved for $\phi(t)$ or p(t). Under the worst case conditions, when

$$\max\left[\phi(t)\right] = \pi/2 \tag{4-107}$$

these inequalities combine as

$$-K_c E_p + \pi/2 \leq p(t) \leq K_c E_p - \pi/2. \qquad (4\text{-}108)$$

This may seriously limit the usefulness of the system shown in Figure 4.24. The system shown in Figure 4.25 does not have this limitation but requires three phase modulators.

4.5 Discussion of Results

This section discusses the results obtained in Sections 4.2 to 4.4 for the three universal modulators under consideration.

The bandwidths were defined in terms of the equivalent filter bandwidth which was defined in Section 2.5.1. The test message was selected to produce IM distortion figures that are higher than those that would be obtained with the usual 4:1 amplitude ratio because it is necessary to make allowances for practical baseband preprocessors where the operations itself introduce more distortion.

Practical baseband preprocessors may have transfer characteristics that are not a perfect reproduction of the mathematical equation. The synthesis of the transfer characteristics may require that approximations of functions and mathematical operations in order to reduce circuit complexity or to handle physically unrealizable operations, like division by zero or the Hilbert transform. These approximations result in additional distortion that has not been considered in the analysis because the distortion is produced by the specific baseband preprocessor realization and is not due to the complex

envelope components characteristics. In this work, all the functions and transformations were assumed to be ideal.

The second moment bandwidth is not used extensively because: (1) it cannot be obtained in closed form for all cases; (2) some of those cases where it can be obtained are restricted to zero mean gaussian noise with specific autocorrelation functions; and (3) it can be misleading. This third reason is very important. Consider the second moment bandwidths calculated for the DSB-LM phase function and for the PM real quadrature component. These are given by Equation (3-136) and Equation (3-149). Compare with Figures 4.11 and 4.16 and observe that to get the bandwidth definitions to agree with 0.1% distortion curve it is necessary to add 3B to the PM case and about 8B for the DSB-LM case. This is inconsistent, the once rigorous second moment bandwidth has been transformed into a partly arbitrary definition that must be checked against distortion anyway. For this reason, the second moment bandwidth is not very useful for nonlinear problems. This bandwidth definition is not completely useless since it serves as a guide of how the different modulation parameters affect the bandwidth and it serves to confirm the simulation results.

The reason why the second moment bandwidth fails to be consistent if compared with with the equivalent-filter bandwidth is in how the different frequency regions of the spectral distribution are used to obtain the final signal. This spectral distribution can be completely altered by nonlinear

processing. The second moment bandwidth definition ignores the fact that the energy above B_s may be as important as the reconstruction of the message. How important this energy is depends on the specifics of each case.

4.5.1 The AM/PM Modulator

The block diagram of the AM/PM modulator was presented in Figure 4.5. The baseband preprocessors transfer characteristics, the a(t) and p(t) functions bandwidth requirements, and the dynamic range constraints were presented in Section 4.2.

The critical modulation laws were the DSB-LM and the SSB-LM. These cases require phase functions with a bandwidth of 16B and 180B respectively, while their magnitude functions require less than 10B and 5B, respectively, for 0.1% IM distortion. The curves for p(t) in Figure 4.11 indicate that the spectrum of the phase functions is wideband which agrees with the discussion in Sections 3.4.2 and 3.4.3. The plot for the SSB-LM indicates that the significant frequency components of p(t) decay rapidly after 200B.

The CSSB cases showed no relationship between the distortion at the output of an ideal receiver and the bandwidth. The ideal receiver uses only one of the two complex envelope components and this component was simply related to the message signal. The "unused" component was necessary to obtain a single-sided carrier, but even for 0.01% of the power in the undesirable sideband the bandwidth requirements were not strict.

The SSB-LM a(t) and p(t) functions also affected the undesireable sideband suppression but the necessary bandwidth was below that required for low distortion reception.

The advantage of the AM/PM modulator over the other two systems is that it is well suited for AM, PM, and FM generation. The main disadvantage is that it is impractical as a SSB-AM modulator. For example, for a message with a 5KHz. bandwidth, the phase modulator must be capable of handling at least a 1MHz. input modulating signal. The baseband preprocessor also has to compute the quotient $\hat{x}(t)/x(t)$ which is troublesome if x(t) is zero or nearly zero.

The AM/PM modulator can be improved. Observe that the balanced modulator allows negative polarity inputs; therefore, instead of generating a(t) for DSB-LM, generate i(t) for DSB-LM. If this is done, p(t) is zero so the DSB-LM problem is quickly solved. The SSB-LM case still remains a problem.

4.5.2 The Quadrature Modulator

The block diagram of the quadrature modulator is shown in Figure 4.6. Section 4.3 discussed the details of the equations for transfer characteristics, and obtained the bandwidth and dynamic range constraints.

In this system there were no critical cases. Although most functions required more bandwidth than the message, none of the cases under study required more than nine times the bandwidth of the message. The largest bandwidths were required for the CSSB-PM and the CSSB-FM components.

When the CSSB components were considered individually, they required more bandwidth than their double-sided counterparts. If both components were filtered with the same kind of filter, the required bandwidth was less or equal to that of the double-sided case. This reduction seems to be due to distortion cancellation caused by built-in redundancy of the quadrature components because they form a Hilbert pair. The second moment bandwidth fails to predict this effect. This kind of bandwidth reduction was not observed in the double sided cases where both components were filtered, the distortion was the same as the highest of the two distortion figures for each individual component.

The single-sided property of CSSB was not affected when both components were filtered to the same bandwidth. The reason for this is that this kind of filtering preserves the Hilbert-pair relationship; therefore, the complex envelope remains Analytic.

The dynamic range constraints on the modulation parameters were found to be the same or more liberal than those required by the other two systems.

The advantages of this system were: (1) the restrictions on the dynamic range were less restrictive than in any of the other two systems under consideration; (2) SSB-LM was easily obtained; and (3) the bandwidths required were never more than two to nine times the bandwidth of the message. The disadvantage of the quadrature modulator was that its baseband preprocessors tended to require more operations than the other two systems.

4.5.3 The PM/PM Modulator

The block diagrams of the PM/PM modulator were shown in Figures 4.24 and 4.26. The last block diagram was preferred because it was less restrictive on the phase modulators. Unfortunately, it required three phase modulators. Both realizations required the same baseband preprocessors. The bandwidth requirements, the equations for the preprocessors, and the dynamic range limitations were discussed in Section 4.4.

The PM/PM modulator and the AM/PM modulator have similar limitations because they share the same p(t) function. The $\phi(t)$ function required slightly more bandwidth than the corresponding a(t) function.

The dynamic range constraints for the PM/PM modulator were found to be the same as those obtained for the AM/PM modulator provided that the maximum input to the balanced modulator, $K_a E_b$, be replaced by twice the amplitude of the carrier going into one of the PM modulators.

The principal advantages of this system over the others was that the use of two phase modulator can be more efficient than an amplitude modulator or a balanced modulator because it allows class C amplifiers [51]. The disadvantage was the same as for the AM/PM modulator plus the fact that the $\phi(t)$ function requires more operations than the a(t) function.

4.5.4 Which One Is Better?

If versatility is the prime consideration, the quadrature modulator seems to be the winner. The bandwidth requirements were not excesive or unreasonable. The dynamic range cons-

traints were always equal to or more liberal than those for the AM/PM and PM/PM modulators. For example, no restrictions are placed on the maximum phase deviation for PM and FM, while the AM/PM and PM/PM modulators require the phase deviation to be restricted to $\pm 180^{o}$. The quadrature modulator was the only system under consideration capable of operating as a SSB-LM modulator without extraordinary bandwidth requirements.

When power efficiency is the main consideration and SSB-LM is not needed, the PM/PM modulator is more atractive because it allows the use of class C power amplifiers.

The AM/PM modulator is better suited for applications where simplicity of the design is the principal concern.

It is clear from this discussion that the final selection depends on the application.

4.6 Physical Realization of the Baseband Preprocessors

There are three basic alternatives to build the baseband preprocessor. These are: (1) continuous time, (2) discrete time continuous-amplitude, and (3) discrete-time discrete-amplitude.

The main objective is to obtain a circuit that will perform satisfactorily the operations required by the entries on Tables IV-1, IV-3, and IV-5. The actual design of the circuits depends on the state-of-the-art of circuit design and it will not be considered in detail.

4.6.1 Continuous-Time

The continuous-time realization involves the use of analog function generators, analog filters, multipliers, and other

operational building blocks. It is necessary to account for, and compensate time-delays and phase shifts introduced by the different signal paths. This is specially important when the Hilbert transform is involved because it is basically a 90^{o} phase-shifter. The baseband preprocessors can be built as separate circuit boards that are simply plugged in a main board that contains the modulators.

The disadvantage of the continuous-time realization is that analog function generators tend to change their characteristics with temperature and aging of components, so calibration may be necessary. Any deviation of the baseband preprocessor transfer characteristics from the ideal equation will result in additional distortion at the receiver or incomplete sideband suppression.

4.6.2 Discrete-Time Continuous-Amplitude

This approach involves sampling the signals and using charge-coupled-devices (CCD) as vaiable delay lines to simulate linear filters and phase shifters [52]. The nonlinear operations must still be handled by analog circuits; therefore, this approach may not be any better than the continuous-time. The advantage is that propagation time-delays are easier to compensate because they occur in discrete steps.

4.6.3 Discrete-Time Discrete-Amplitude

This is basically a digital approach. The continuous time message is converted to a digital sequence, the sequence is processed and the resulting sequence is converted back to a continuous-time signal. Since this process involves sampling,

it is necessary to know the bandwidth of the signals to be converted to digital sequences in order to avoid distortion introduced by aliasing [39, 53]. The equivalent-filter bandwidth for the individual components can be used as a guide to determine the sampling rate since it shows the relationship between forcing a signal to be bandlimited and the resultant distortion.

A proposed realization of an arbitrary baseband preprocessor is shown in Figure 4.31. The input lowpass filter is analog and its function is to limit the bandwidth of the input signal before sampling. The analog-to-digital converter (A/D) changes the bandlimited x(t) to a digital sequence x(nT) where n is the sample number and T is the time between samples. The digital computing network calculates the output sequence r(nT) following the baseband preprocessor transfer characteristic and this sequence is converted back into a continuous time signal by a digital-to-analog converter (D/A). The output lowpass filter is necessary to equalize the transfer function of the D/A converter. The time base coordinates the transfer of data between the different blocks.

The digital computing network is the only part of the system that really depends on the desired transfer characteristic for the baseband preprocessor. Studying the entries of Tables IV-1, IV-4, and IV-7 it is apparent that all the functions share a common computational structure. For example, the a(t) function requires $\hat{x}(t)$, x(t), and then a nonlinear operation involving $\hat{x}(t)$ and x(t). The Hilbert transform is

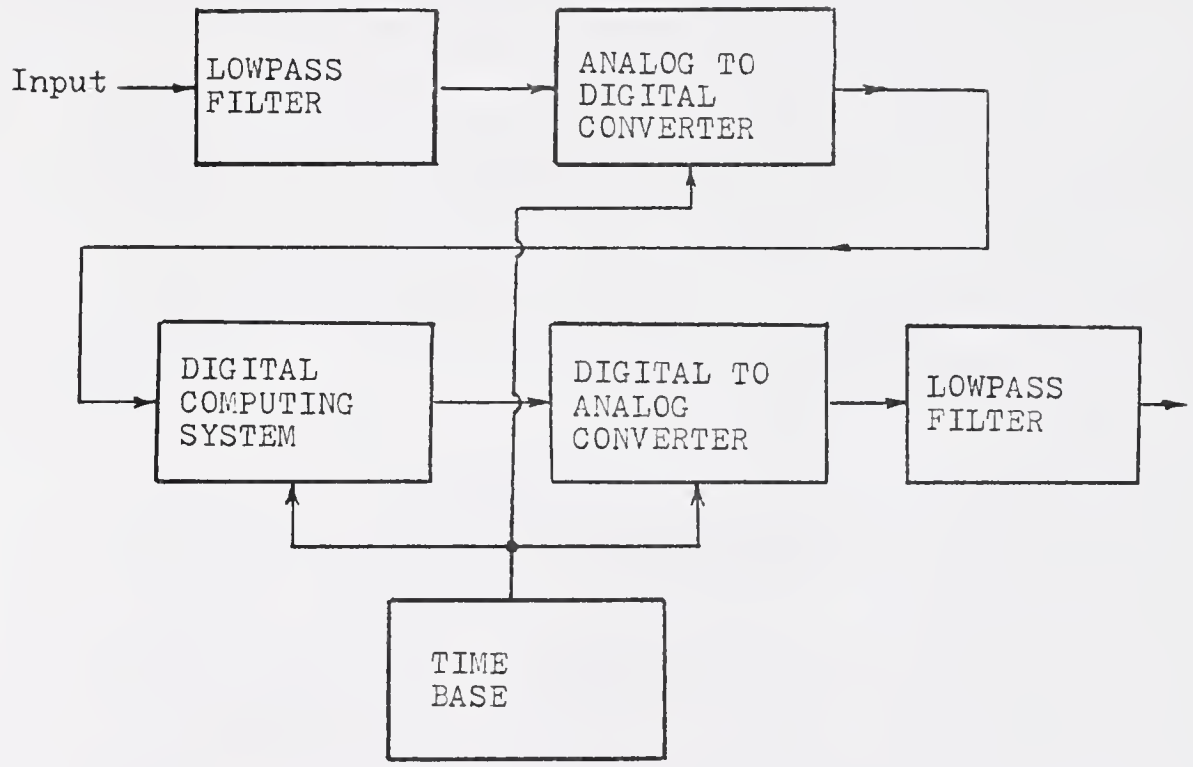


Figure 4.31. Block diagram of a digital baseband preprocessor.

a linear operation and it can be obtained with a digital filter. The message function x(t) must be delayed so that $\hat{x}(t)$ and x(t) have the same phase relationship. The analysis done in Section 4.2.1 revealed that a(t) requires more bandwidth than the input message. This means that the sampling rate must be increased to prevent aliasing when the sequence is converted back to a continuous-time signal [54]. The interpolator is defined as a digital system that increases or decreases the sampling rate. If this analysis is extended to the other functions: p(t), $\phi(t)$, i(t), and q(t), the general block diagram of Figure 4.32 is obtained. This block diagram ignores the possibility of sharing component blocks between baseband preprocessors because that depends on the actual functions to be obtained.

The low bandwidth requirements of the quadrature modulator makes it particularly attractive for the discrete-time realization.

Among the other things that must be taken into account in the discrete-time, discrete-amplitude realization are the number of bits per sample, the numbering system, and the time required to compute an output data point. The time required to compute a data point is important because everything must be done in real-time or at least with a finite time-delay.

The advantages of this realization are: (1) it is easier to obtain most of the required functions because these functions can be computed by numerical methods, (2) it is easy to account for time-delays because of the discrete nature of time;

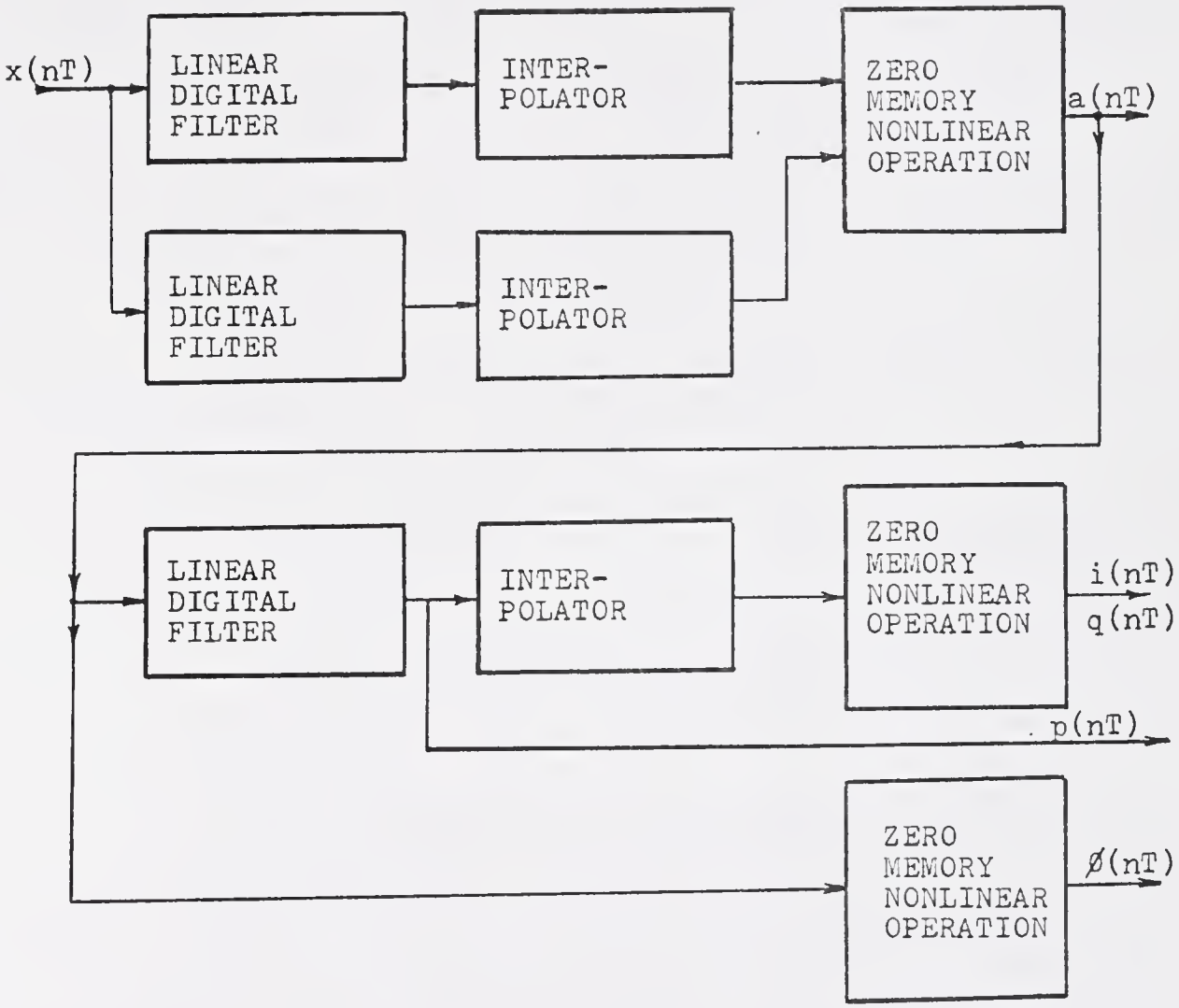


Figure 4.32. Block diagram of the digital computing system for a general baseband preprocessor.

(3) the system constants are unaffected by time or temperature because they are stored numbers; and (4) the baseband preprocessor can be reprogrammed to generate other functions that fit the same general structure without actual circuit changes.

The disadvantages are principally the accuracy and the speed of the computing circuit.

4.7 Summary

This chapter analyzed various alternatives to achieve an universal transmitter. It was shown that only two, the AM/PM and the quadrature transmitters are really different. A third system, the PM/PM transmitter was found to be similar to the AM/PM modulator in most respects.

The transfer characteristics of the baseband processors, the bandwidth requirements for the component signals, and the dynamic range limitations were discussed in detail. The quadrature modulator was found to be the best of the three systems under consideration, but the other two systems have their particular advantages that may make them more attractive in some applications.

Also discussed was the fact that the second moment bandwidth is not the best bandwidth definition for systems where nonlinear operations are involved and where the primary consideration is distortion. This was illustrated with examples that showed inconsistencies between distortion measurements and the second moment bandwidth.

Finally some possible realizations for the baseband pre-

processor were discussed. The discrete-time discrete amplitude realization was found to be particularly attractive because it allows the possibility of obtaining the transfer characteristics of the baseband preprocessor in terms of software instead of hardwired circuits that are difficult to modify.

CHAPTER V
THE UNIVERSAL RECEIVER

5.1 Introduction

The preceeding chapter discussed the possibility of building a transmitter capable of generating different kinds of modulated carriers. It was shown that there are structures that can accomplish this. This chapter is concerned with the possibility of building a receiver capable of demodulating almost any carrier. The general structure of the receiver was shown in Figure 1.2.

The objective is to find block diagram structure that can demodulate a carrier following a two-step process. One part of the process is independent of the modulation law and recovers the complex envelope of the input. The block diagram of such system is shown in Figure 5.1. The input tuned RF amplifier acts as a bandpass filter and is necessary to limit the amount of complex noise, n(t), going into the receiver. The filter must be wide enough to let the complex modulated carrier, z(t), go through without noticeable distortion. The carrier demodulator extracts the complex envelope by multiplying z(t) + n(t) by the complex carrier $\exp(-jw_o t)$. This results in a baseband signal that is used by the baseband postprocessor to estimate the message signal, x(t). The proposed system is physically unrealizable because it involves complex signals, but it can be made physically

realizable if the complex signals are replaced by their real valued components.

For the discussion that follows, it is assumed that the system is studied after the input RF amplifier. The remaining system will be called an universal demodulator.

The demodulator analysis is more complicated due to the presence of various disturbances like noise, fading, Doppler frequency shifts, multipath propagation, interference, etc. The receiver may have to be synchronized with the transmitter which is not always possible. To simplify the analysis the effects of synchronization and noise will be studied separately.

5.1.1 Demodulation as the Inverse of Modulation

The received complex modulated carrier is

$$z_r(t) = kz(t) + n(t), \qquad (5\text{-}1)$$

where $z(t)$ is the transmitted carrier, $n(t)$ is bandlimited noise, and k is a constant that accounts for the attenuation of the modulated carrier between the transmitter and the

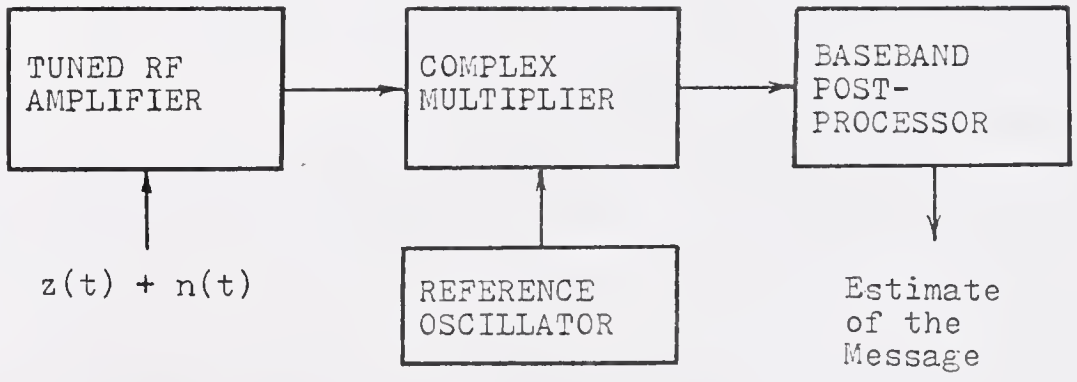


Figure 5.1. The universal receiver.

receiver. Using Equation (2-6), Equation (5-1) can be written as

$$z_r(t) = kv(t)\exp(jw_o t) + n(t). \tag{5-2}$$

The complex noise, $n(t)$, can be expressed in terms of its complex envelope, $N(t)$,

$$n(t) = N(t)\exp(jw_o t). \tag{5-3}$$

The complex envelope of the noise process can be expressed as

$$N(t) = N_o(t)\exp[j\eta(t)] \tag{5-4}$$

where $N_o(t)$ and $\eta(t)$ are random magnitude and phase functions, respectively. The function $N(t)$ can be written in terms of its random quadrature components, $n_i(t)$ and $n_q(t)$, as

$$N(t) = n_i(t) + jn_q(t). \tag{5-5}$$

Substitute Equation (5-3) into Equation(5-2) to obtain

$$z_r(t) = kv(t) + N(t)\exp(jw_o t). \tag{5-6}$$

Now, multiply both sides by $\exp(-jw_o t)$

$$z_r(t)\exp(-jw_o t) = kv(t) + N(t), \tag{5-7}$$

so the received complex envelope is

$$\begin{aligned} v_r(t) &= z_r(t)\exp(-jw_o t) \\ &= kv(t) + N(t). \end{aligned} \tag{5-8}$$

This equation shows that perfect recovery of the transmitted

complex envelope is not possible because k and N(t) are not known. If the noise is zero, then the received and transmitted complex envelopes are linearly related.

The received complex envelope can be expressed in terms of its polar components, $a_r(t)$ and $p_r(t)$, as

$$a_r(t) = \sqrt{k^2a^2(t) + N_o^2(t) + 2ka(t)N_o(t)\cos[\eta(t)]} \tag{5-9}$$

and

$$p_r(t) = p(t) + \tan^{-1}\left[\frac{N_o(t)\sin[\eta(t) - p(t)]}{ka(t) + N_o(t)\cos[\eta(t) - p(t)]}\right]; \tag{5-10}$$

or in terms of the quadrature components, $i_r(t)$ and $q_r(t)$, as

$$i_r(t) = ki(t) + n_i(t) \tag{5-11}$$

and

$$q_r(t) = kq(t) + n_q(t). \tag{5-12}$$

These equations show that all the components are affected by noise. Noise is a random process so it can be described only in terms of statistical parameters.

5.1.2 The Synchronization Problem

The derivation of Equation (5-6) assumed that the receiver had prior knowledge of the transmitter carrier center frequency and phase angle. In practice, this information is not always

known. The reference carrier at the receiver may have a phase angle error, $\varepsilon(t)$, with respect to the incoming carrier. Instead of obtaining Equation (5-7), the following results

$$z_r(t)\exp[-jw_o t + j\varepsilon(t)] = [kv(t) + N(t)]\exp[j\varepsilon(t)] , \tag{5-13}$$

and the received complex envelope is

$$v_r(t) = [kv(t) + N(t)]\exp[j\varepsilon(t)] . \tag{5-14}$$

The noise can be simply replaced by another random process with a complex envelope

$$N_e(t) = N(t)\exp[j\varepsilon(t)] , \tag{5-15}$$

but the received complex envelope still differs from Equation (5-8) because

$$v_r(t) = kv(t)\exp[j\varepsilon(t)] + N_e(t). \tag{5-16}$$

Assume noiseless conditions, that is, $N_e(t)$ is zero, then the received complex envelope components are

$$a_r(t) = ka(t), \tag{5-17}$$

$$p_r(t) = p(t) + \varepsilon(t), \tag{5-18}$$

$$i_r(t) = k[i(t)\cos[\varepsilon(t)] - q(t)\sin[\varepsilon(t)]] \tag{5-19}$$

and

$$q_r(t) = k[i(t)\sin[\varepsilon(t)] + q(t)\cos[\varepsilon(t)]] . \tag{5-20}$$

These equations show that the magnitude function is altered by a constant and that the phase function is distorted by $\varepsilon(t)$. The received quadrature functions are a mixture of the original quadrature functions. This clearly shows the need to reduce $\varepsilon(t)$.

5.1.3 Types of Physical Demodulators

Demodulators can be classified as coherent or incoherent. Coherent demodulators require an external reference in order to demodulate the signal. This kind of demodulator must know the original carrier phase and frequency at the moment it arrives at the receiver.

The incoherent demodulator does not need any reference because it determines only those parameters that can be measured in absolute terms, like frequency and amplitude.

The two basic coherent demodulators are the homodyne detector and the phase detector [1]. The homodyne detector consists of a four-quadrant multiplier followed by a lowpass filter. The block diagram shown in Figure 5.2.

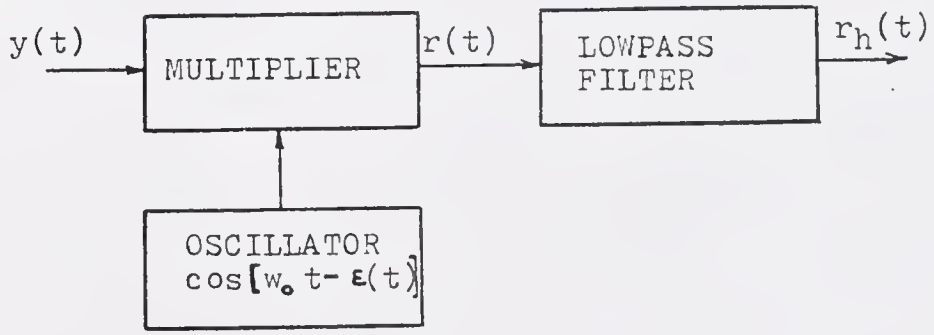


Figure 5.2. The homodyne detector.

Let $\varepsilon(t)$ be the phase angle error between the reference carrier and the modulated carrier center frequency. The output of the multiplier, r(t), is

$$r(t) = \frac{ka(t)}{2} \cos[\varepsilon(t) + p(t)] + \frac{ka(t)}{2} \cos[2w_o t + p(t) - \varepsilon(t)]. \tag{5-21}$$

If there is no overlapping of spectral components between the two terms in the preceeding equation, the second term can be filtered out with an appropiate lowpass filter, so

$$r_h(t) = \frac{ka(t)}{2} \cos[\varepsilon(t) + p(t)]. \tag{5-22}$$

A similar expression can be obtained in terms of the quadrature components

$$r_h(t) = \frac{k}{2} i(t)\cos[\varepsilon(t)] - q(t)\sin[\varepsilon(t)]. \tag{5-23}$$

For the case where $\varepsilon(t) = 0$ but when noise is present, the output of the homodyne detector is [1]

$$r_h(t) = \frac{k}{2} i(t) + \frac{n_i(t)}{2}. \tag{5-24}$$

The phase detector is a phase comparator. Its function is to compare the phase angle of the incoming carrier with an external reference. Its output is proportional to the difference between the two angles. Due to the trigonometric function definitions, the phase comparator cannot distinguish between two phase angles that differ by integer multiples of $\pm 2\pi$.

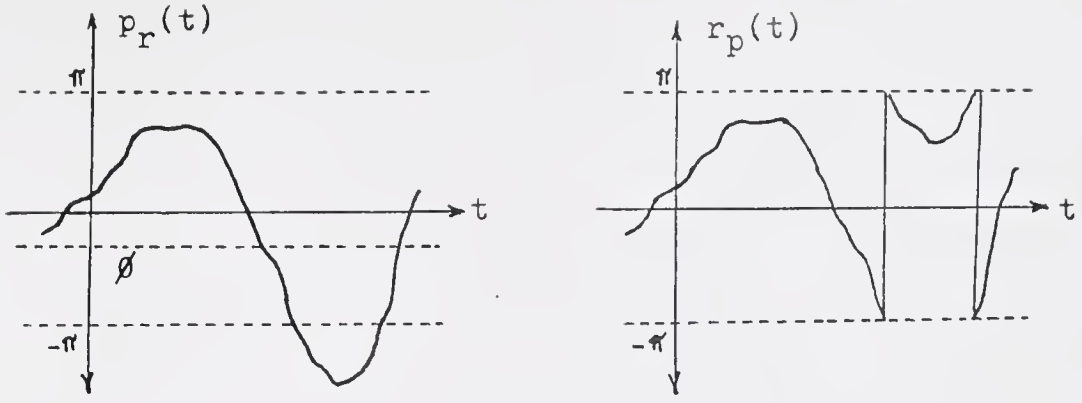


Figure 5.3. Effect of nonzero phase error in the output of the phase detector.

If the phase detector reference has an error $\varepsilon(t)$, the output is (for the noiseless case)

$$r_p(t) = P\{p(t) + \varepsilon(t)\} \tag{5-25}$$

where $P\{\cdot\}$ denotes the operation of extracting the principal value of the argument. Equation (5-25) imposes a particular problem because it seriously limit the dynamic range of $p(t)$. If $\varepsilon(t) = \phi$, a constant, and the input is like the waveform shown in Figure 5.3, the output will show phase discontinuities. This means that the original phase angle will be mutilated. If the phase angle error is due to a frequency difference between the reference and the carrier, the level where discontinuities occur will change periodically with time. It should be mentioned that the phase discontinuities will result in a increase of spectral occupancy of $p_r(t)$.

If noise is present, the output of the phase detector is found by substituting Equation (5-10) into Equation (5-25).

If the noise is small compared to ka(t) most of the time, that is, $ka(t) \gg N_o(t)$, Equation (5-10) is approximately equal to

$$p_r(t) \approx p(t) + \frac{N_o(t)}{ka(t)} \sin[\eta(t) - p(t)], \qquad (5\text{-}26)$$

and the output of the phase detector is obtained using Equation (5-25) with $\varepsilon(t) = 0$

$$r_p(t) \approx P[p_r(t)]. \qquad (5\text{-}27)$$

There are two basic incoherent detectors. These are the envelope detector and the frequency discriminator [1]. The ideal envelope detector recovers the magnitude of the received complex envelope. The ideal frequency discriminator recovers the instantaneous frequency of the complex modulated carrier.

The output of an ideal envelope detector is

$$r_e(t) = ka_r(t); \qquad (5\text{-}28)$$

therefore, it is unaffected by the phase angle. When noise is present, the output is identical to Equation (5-9). If the noise is small compared to ka(t) most of the time, Equation (5-9) can be approximated as[*]

$$r_e(t) \approx ka(t) + N_o(t)\cos[\eta(t)] \qquad (5\text{-}29)$$

or

---

[*] $\sqrt{x^2 + 2xy} \approx x + y$, where y is small with respect to x.

$$r_e(t) \approx ka(t) + n_i(t) \qquad (5\text{-}30)$$

which is similar in form to the result obtained for the homodyne detector in Equation (5-24).

The ideal frequency discriminator is insensitive to the amplitude fluctuations of the carrier. This is achieved with a hard limiter. The output of the discriminator is

$$r_f(t) = \frac{dp_r(t)}{dt} , \qquad (5\text{-}31)$$

provided that no noise is present. When the effect of noise is included and $ka(t) \gg N_o(t)$ most of the time, the output of the frequency discriminator is well known [1, 14] and is given by

$$r_f(t) \approx \frac{dp(t)}{dt} + \frac{1}{kC} \cdot \frac{dn_q(t)}{dt} . \qquad (5\text{-}32)$$

where C is the unmodulated carrier amplitude at the transmitter. This equation assumes that any amplitude fluctuations due to a(t) can be neglected.

The incoherent detectors have a threshold level below which the demodulated signal is severely mutilated by the noise, and above which the noise has little effect on the signal [1].

5.1.4 Coherent Universal Demodulators

The structures for the block diagrams of the coherent universal demodulators are a consequence of the structures for the universal modulators. There are only two feasible universal demodulation systems. These are the AM/PM demodulator and the quadrature demodulator. There is no PM/PM

demodulator analogous to the PM/PM modulator because it involves three parameters. This will be discussed latter.

The AM/PM demodulator is shown in Figure 5.4. It consists of an envelope detector, a phase detector, a baseband postprocessor, and an oscillator. A feedback path is provided from the baseband postprocessor to the oscillator to account for the possibility of reducing the phase angle error on the basis of the information available to the baseband postprocessor. The complex variables and functions are replaced by their real components just like in Chapter IV.

The second system is the quadrature demodulator. It consists of two homodyne detectors, an oscillator, a phase shifter, and a baseband postprocessor. This is shown in Figure 5.5. Again, there is a feedback path from the baseband postprocessor to the oscillator as stated in the previous paragraph.

A system based on two phase detectors is not possible although it is a modulator arrangement. The reason is that the transmitter knows the magnitude of the phase-modulated carrier components and knows what complex envelope has to be generated. As Equation (4-86) indicates, the transmitter knows L, θ_1, and θ_2. The receiver can take only two phase measurements. In other words, the triangle shown in Figure 4.2(e) cannot be constructed with knowledge of only two of its angles and nothing about the length of one of the sides.

There are other structures that can serve as universal demodulator. One of them is based on an extension of a system proposed by Zayezdnyy [55] to demodulate AM. It is based on

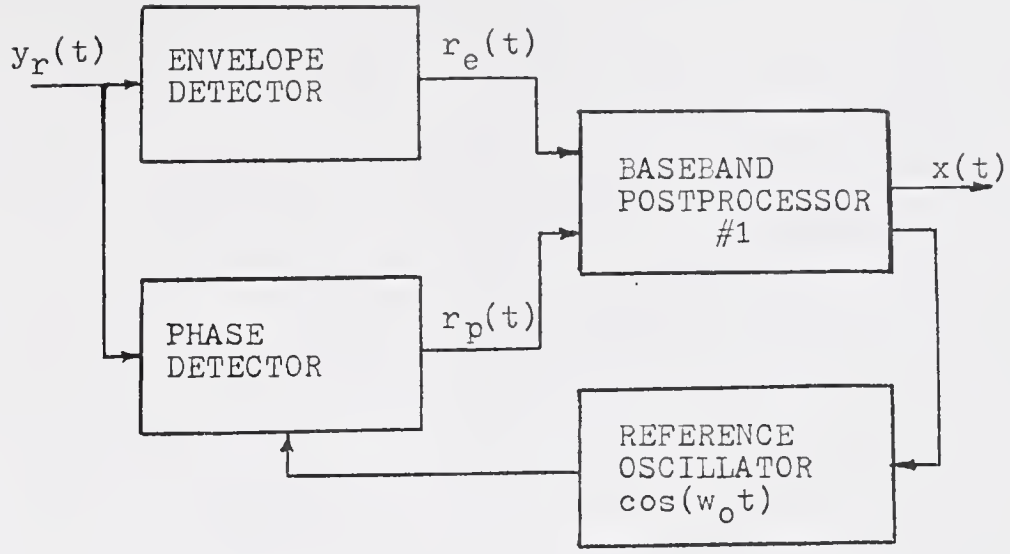


Figure 5.4. The AM/PM demodulator.

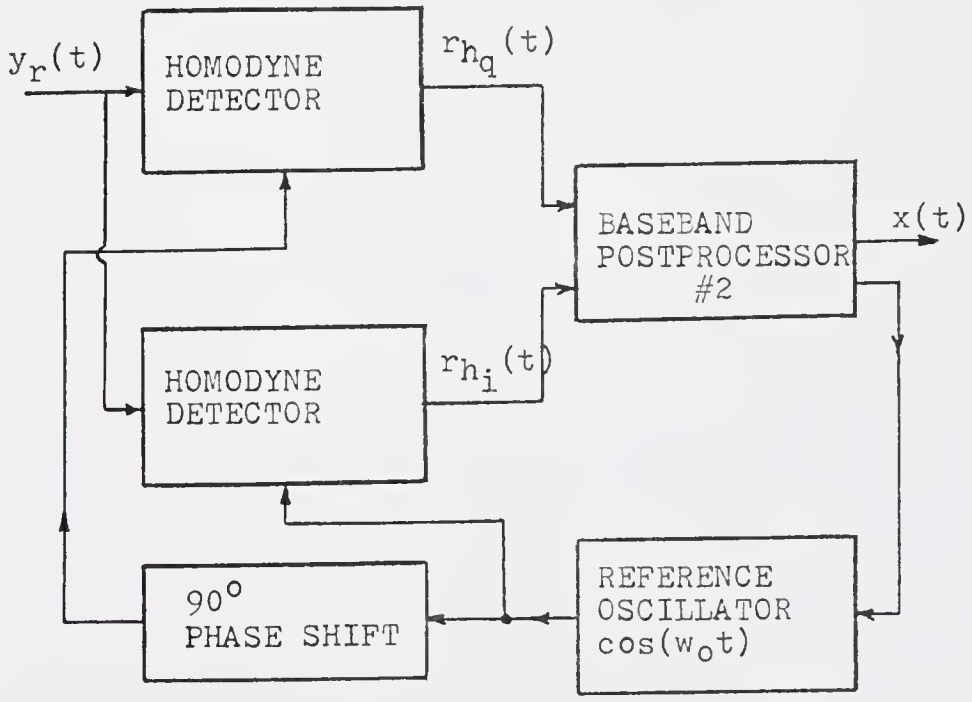


Figure 5.5. The quadrature demodulator.

the generation of an Analytic signal from the real modulated carrier using a Hilbert transformer. From the Analytic signal, the amplitude and phase can be obtained. If Equation (2-1) represents a narrowband modulated carrier, the Hilbert transform is [56]

$$\hat{y}(t) = a(t)\sin[w_o t + p(t)], \tag{5-33}$$

so the Analytic signal is

$$z_a(t) = y(t) + j\hat{y}(t), \tag{5-34}$$

from which the envelope can be recovered as

$$a(t) = \sqrt{y^2(t) + \hat{y}^2(t)}\,. \tag{5-35}$$

The phase angle is

$$p(t) = \tan^{-1}\left[\frac{\hat{y}(t)}{y(t)}\right] - w_o t \tag{5-36}$$

which means that the exact frequency must be known. The complexity of Equation (5-36) does not justify its use.

Another structure that will demodulate various types of modulated carriers is the phase-locked-loop (PLL). The PLL has been analyzed in detail in various communications theory textbooks [57-59] . It will demodulate most double-sided modulated carriers where a carrier term is present. If no carrier term is present then additional modifications are necessary and the system is more complicated and cannot be considered as a universal demodulator.

5.1.5 Incoherent Universal Demodulators

It has been shown in Section 5.1.3 that the complex envelope must be recovered coherently or the components are distorted. It is possible to obtain approximations to the complex envelope components without a phase reference if the relative phase angle is not used directly. The incoherent AM/FM demodulator is shown in Figure 5.6. This system uses an envelope detector to recover a(t) and a balanced frequency discriminator to recover the instantaneous frequency of the complex envelope. This system will be discussed in detail latter.

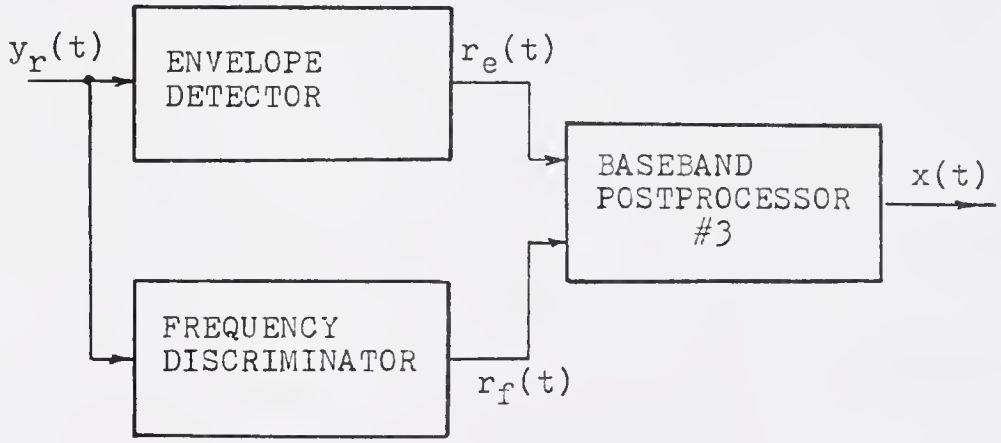


Figure 5.6. The AM/FM demodulator.

5.1.6 Criteria for Comparison and Evaluation of Demodulators

The criteria to compare the different system are the bandwidths required by the signals coming out of the detectors, their dynamic ranges, their noise behavior, and the effect of synchronization on the recovered message.

The bandwidths and the dynamic ranges of the complex envelope components were already studied in Chapter IV. The reader

should refer to Section 4.2 for information on the a(t) and the p(t) functions and to Section 4.3 for the i(t) and the q(t) functions.

Since the noise problem is very complicated in itself, the analysis will be limited to qualitative aspects rather than to analytical work because the main interest of this chapter is to study the systems feasibility and not their optimization. The same applies to the synchronization problem.

5.2 The AM/PM Demodulator

The inputs to the baseband postprocessor #1 and $r_e(t)$ and $r_p(t)$. The necessary equations for these signals are those obtained for the envelope detector and the phase detector.

First consider the noiseless reception with zero phase angle error. There are only four different cases to study. Receivers that demodulate AM, DSB-LM, PM, and FM can demodulate CSSB-AM, SSB-LM, CSSB-PM, and CSSB-FM, respectively. For the AM and CSSB-AM case, substitute the expression for a(t) listed in Table IV-1 in Equation (5-29), with the noise set to zero. This yields

$$r_e(t) = a_r(t) = kC[1 + mx(t)], \tag{5-37}$$

so the message can be recovered as

$$x(t) = \frac{1}{m}\left[\frac{r_e(t)}{kC} - 1\right]. \tag{5-38}$$

The constants m and 1/kC are usually unknown to the receiver. A DC decoupler can be used to obtain an estimate of the message, $x_e(t)$, as

$$x_e(t) = r_e(t) - E\{r_e(t)\} \tag{5-39}$$

where $x_e(t)$ and $x(t)$ differ by an unknown scale factor that does not affect the intelligibility of the recovered message. The phase function is not used.

For the DSB-LM and SSB-LM cases, define

$$f(t) = r_e(t)\cos[p_r(t)]. \tag{5-40}$$

Substitute the equations listed in Table IV-1 for these cases to obtain

$$f(t) = Ck|x(t)|\cos\left[\frac{\pi}{2}\left[1 - \operatorname{sgn}[x(t)]\right]\right] \tag{5-41}$$

for DSB-LM, and

$$f(t) = Ck\sqrt{x^2(t) + \hat{x}^2(t)}\cdot\cos\left[\tan^{-1}\left[\frac{\hat{x}(t)}{x(t)}\right]\right] \tag{5-42}$$

for SSB-LM. Since

$$\cos\left[\frac{\pi}{2}\left[1 - \operatorname{sgn}[x(t)]\right]\right] = \begin{cases} 1, & x(t) \geq 0 \\ -1, & x(t) < 0, \end{cases} \tag{5-43}$$

and

$$\cos\left[\tan^{-1}\left[\frac{\hat{x}(t)}{x(t)}\right]\right] = \frac{x(t)}{\sqrt{x^2(t) + \hat{x}^2(t)}}, \tag{5-44}$$

Equations (5-41) and (5-42) reduce to

$$f(t) = Ckx(t); \tag{5-45}$$

therefore, the estimate of the message is

$$x_e(t) = f(t) = r_e(t)\cos[p_r(t)] = \frac{x(t)}{Ck} \tag{5-46}$$

For the PM, CSSB-PM, FM, and CSSB-FM, ignore the magnitude function. According to the equations listed in Table IV-1, the output of the phase detector is

$$r_p(t) = P[p_r(t)] = P[D_p x(t)] \tag{5-47}$$

for the two PM cases, and

$$r_p(t) = P\left[D_f \int_{-\infty}^{t} x(u)du\right] \tag{5-48}$$

for the two FM cases. The message can be estimated as

$$x_e(t) = r_p(t) \tag{5-49}$$

for the PM and CSSB-PM cases, and

$$x_e(t) = \frac{d}{dt}\left[r_p(t)\right] \tag{5-50}$$

for the FM and CSSB-FM cases. The block diagram of the general baseband postprocessor is shown in Figure 5.7.

Consider the case where perfect synchronization is not possible. From Equations (5-17) and (5-18) it is seen that only the received phase function, $p_r(t)$, is affected by the phase error. This means that if $x(t)$ can be recovered from $r_e(t)$ alone, there is no need to worry about $\varepsilon(t)$. This is true for AM, CSSB-AM, CSSB-PM, and CSSB-FM.

The CSSB-PM and CSSB-FM messages can be recovered as follows. In the CSSB-PM case, Equation (5-17) yields

$$r_e(t) = kC \cdot \exp\left[-D_p\hat{x}(t)\right] \quad ; \tag{5-51}$$

therefore,

$$-D_p\hat{x}(t) = \ln\left[r_e(t)\right] - \ln\left[kC\right]. \tag{5-52}$$

Now take the Hilbert transform of both sides of this equation and observe that the Hilbert transform of a constant is zero, so

$$x_e(t) = D_p x(t) = H\left\{\ln\left[r_e(t)\right]\right\}. \tag{5-53}$$

This is similar to a system proposed by Cain [22] to demodulate SSB carriers incoherently. The disadvantage of this approach is that the message is obtained from the amplitude modulation instead from the phase modulation; therefore, the advantages of angle modulation in the presence of noise

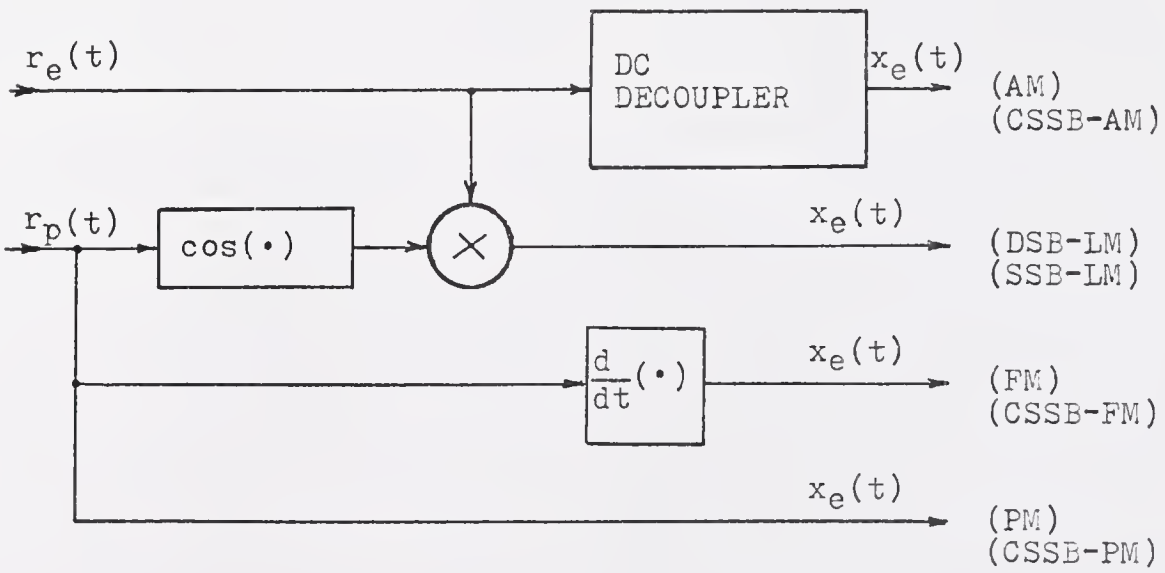


Figure 5.7. Block diagram for the general baseband postprocessor #1 for the noiseless case with perfect synchronization.

are not being utilized. A similar procedure can be followed to obtain the equation to demodulate CSSB-FM

$$x_e(t) = \frac{d}{dt}\left\{ H\left[\ln[r_e(t)]\right]\right\}. \qquad (5\text{-}54)$$

Now it only remains to find out if DSB-LM, SSB-LM, PM, and FM can be demodulated in the presence of a phase error. Consider the DSB-LM case. In this case, the absolute value of the message is linearly related to the magnitude function and the phase function only determines the sign of the message, so the phase function assumes only two values. All that is necessary is to change the sign of $x_e(t)$ whenever a phase jump occurs. This has a flaw. The output of the phase detector is also limited to the range $[-\pi,\pi]$, so the phase jump may be the result of $p_r(t)$ exceeding that range due to $\varepsilon(t)$ rather than due to a change in $p(t)$. The estimated message $x_e(t)$ following this procedure is

$$x_e(t) = r_e(t)\ \mathrm{sgn}[b(t)], \qquad (5\text{-}55)$$

where $b(t)$ is the function that keeps track of the phase jumps. Consider now the direct substitution of Equation (5-25) into Equation (5-46). This yields

$$x_e(t) = kCx(t)\cos[\varepsilon(t)], \qquad (5\text{-}56)$$

This means that the original message is modulated by $\cos[\varepsilon(t)]$. If $\varepsilon(t)$ is a constant, Equation (5-55) predicts an undistorted output because the only phase jumps occuring are due to the carrier phase. If $\varepsilon(t)$ is the result of a frequency error,

Equation (5-55) predicts sudden sign changes whenever $p_r(t)$ exceeds integer multiples of $\pm 2\pi$, while Equation (5-57) predicts a shift of all the frequency components of $x(t)$. Which one is worst cannot be predicted without further study.

The SSB-LM estimate of the message is given by

$$x_e(t) = kC\left[x(t)\cos[\varepsilon(t)] - \hat{x}(t)\sin[\varepsilon(t)]\right] \qquad (5\text{-}57)$$

which is obtained from Equation (5-46) after substituting the corresponding equations for $r_e(t)$ and $p_r(t)$. The result shows crosstalk between the message and its Hilbert transform and the effect of $\varepsilon(t)$ modulating both $x(t)$ and $\hat{x}(t)$. This is a well known result.

It should be noted that if the output of the phase, $r_p(t)$, is used as the argument of a trigonometric function, it is not necessary to worry about principal value limitations. If $r_p(t)$ is used directly, like in PM and FM cases, the principal-value limitation is of concern because it appears as distortion. From the practical point of view, the phase discontinuities are important no matter how $r_p(t)$ is used. These discontinuous points causes $r_p(t)$ to have a wider bandwidth than $p_r(t)$. If the phase-detector output-stage frequency response is not wide enough, distortion will result even if $r_p(t)$ is used as the argument of trigonometric functions.

The PM and FM phase functions are affected by the phase discontinuities discussed in the preceeding paragraph because $p_r(t)$ is supposed to be a linear function of the message. If $\varepsilon(t)$ is small, the principal-value range is not exceeded often

and Equations (5-49) and (5-50) can be rewritten as

$$x_e(t)=p(t) + \varepsilon(t) \tag{5-58}$$

for PM and

$$x_e(t) = \frac{d}{dt}\left[p(t) + \varepsilon(t)\right] \tag{5-59}$$

for FM. If $\varepsilon(t)$ is a constant, the estimate of the PM message differs from the actual message by a scale factor and an unknown constant. This constant can be eliminated if the output is obtained by DC-decoupling the baseband postprocessor. When $\varepsilon(t)$ is a constant, $x_e(t)$ and $x(t)$ differ by a scale factor. If $\varepsilon(t)$ is due to a frequency error, the PM demodulated message is distorted, but the FM message can be recovered by DC-decoupling. This is true assuming that $p_r(t)$ does not exceed $\pm\pi$.

When noise is present, the inputs to the baseband postprocessor are obtained by substituting Equations (5-9) and (5-10) into Equations (5-25) and (5-28). For the small noise case, these equations simplify to

$$r_e(t) = ka(t) + n_i(t), \tag{5-60}$$

and

$$r_p(t) = P\left[p(t) + \frac{n_q(t)}{ka(t)} \cos\left[p(t)\right] - \frac{n_i(t)}{ka(t)} \sin\left[p(t)\right]\right]. \tag{5-61}$$

Note that $n_i(t)$ appears in both equations; therefore, it is possible to obtain an improvement in the output signal to noise ratio by using both $r_e(t)$ and $r_p(t)$ even if only one of

them is required. This is the objective of optimum receivers [24-31], and it is beyond the scope of this dissertation. It is mentioned to point out the possiblility of optimizing the baseband postprocessor.

In overall terms, the limiting case is the SSB-LM phase function because it requires the widest bandwidth. The functions a(t) and p(t) were analyzed in Section 4.2. The dynamic range limitations are the same, except that $K_a E_b$, the maximum input to the balanced modulator, has to be replaced by the maximum output of the envelope detector. Since this output may depend on the strength of the carrier, it is reasonable to assume that the RF gain of the front end of the receiver can be adjusted to avoid overloading the baseband postprocessor; therefore, the actual dynamic range constraints are those imposed by the phase function.

5.3 The Quadrature Demodulator

The quadrature demodulator is easily analyzed for the noiseless case if perfect synchronization exists between the transmitter and the receiver. Let $r_{h_i}(t)$ and $r_{h_q}(t)$ be the outputs of the two homodyne detectors. In Sections 3.2.1 and 3.2.2 it was shown that the i(t) components of the AM, DSB-LM, and SSB-LM cases were linearly related to the message; therefore, the output of the homodyne detector with the cosine reference is

$$r_{h_i}(t) = \frac{i_r(t)}{2} = \frac{kx(t)}{2} \qquad (5\text{-}62)$$

for DSB-LM and SSB-LM cases, and

$$r_{h_i}(t) = \frac{kC}{2}\left[1 + mx(t)\right] \tag{5-63}$$

for AM. It is easily shown that the estimate of the message is

$$x_e(t) = r_{h_i}(t) \tag{5-64}$$

for SSB-LM and DSB-LM cases. If the mean value of $[1 + mx(t)]$ is removed, Equation (5-64) is also valid for AM.

The PM, CSSB-PM, FM, and CSSB-FM can be demodulated if the baseband postprocessor calculates the $r_p(t)$ from $r_{h_i}(t)$ and $r_{h_q}(t)$ as

$$r_p(t) = \tan^{-1}\left[\frac{r_{h_q}(t)}{r_{h_i}(t)}\right], \tag{5-65}$$

so Equations (5-47) and (5-48) can be used to recover the message.

The CSSB-AM case can be demodulated if the baseband postprocessor calculates

$$r_e(t) = \sqrt{r_{h_i}^2(t) + r_{h_q}^2(t)}\ . \tag{5-66}$$

The general block diagram for the ideal case baseband postprocessor is shown in Figure 5.8.

When synchronization between receiver and transmitter is not perfect, the outputs of the homodyne detectors are obtained from Equation (5-23) by observing that one detector uses $\sin[w_o t - \varepsilon(t)]$ and the other $\cos[w_o t - \varepsilon(t)]$; therefore,

$$r_{h_i}(t) = i_r(t) = \frac{k}{2}\left[i(t)\cos[\varepsilon(t)] - q(t)\sin[\varepsilon(t)]\right], \tag{5-67}$$

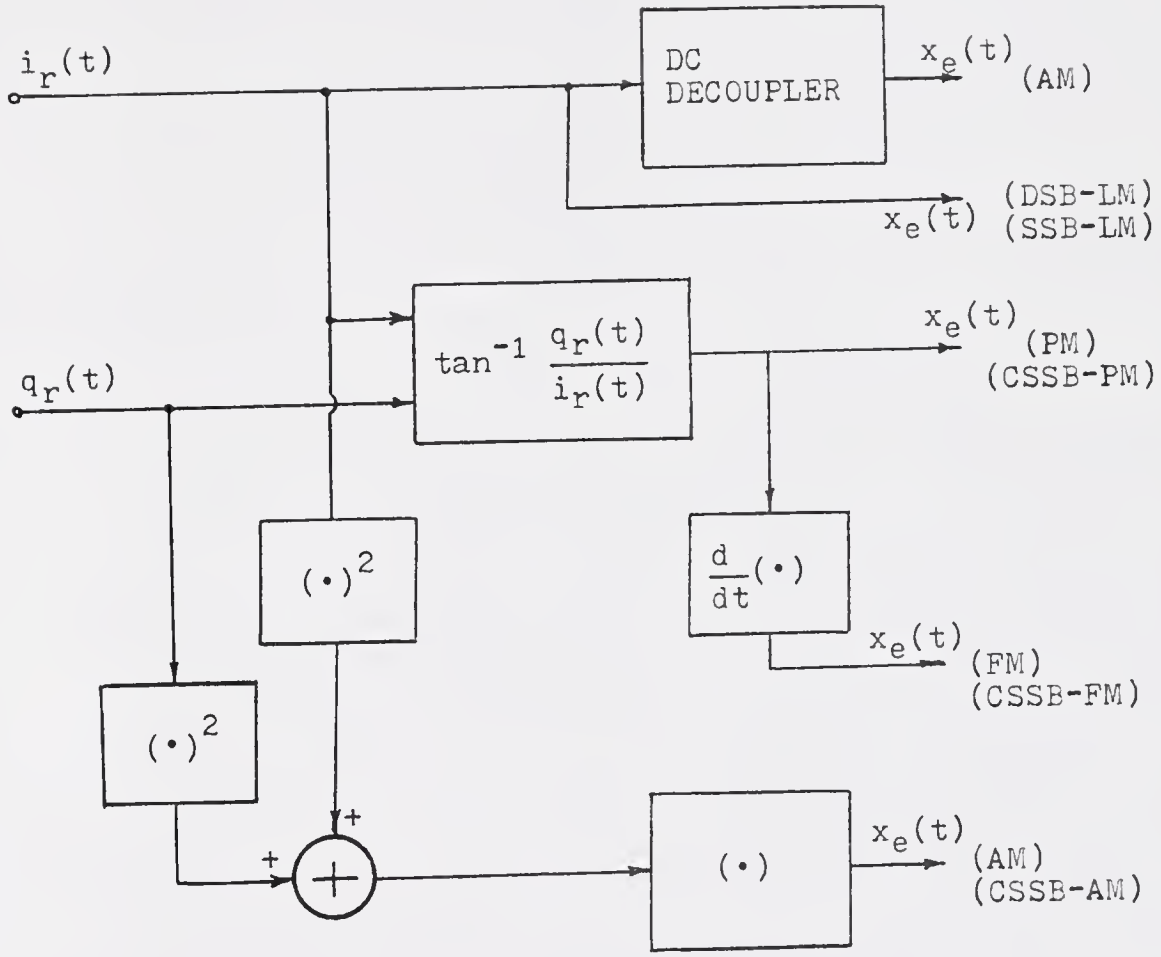


Figure 5.8. Block diagram for the general baseband post-processor #2 for the noiseless case with perfect synchronization.

and

$$r_{h_q}(t) = q_r(t) = \frac{k}{2}\left[i(t)\sin[\varepsilon(t)] + q(t)\cos[\varepsilon(t)]\right] . \quad (5\text{-}68)$$

These equations show crosstalk between i(t) and q(t) that will disappear only if $\varepsilon(t)$ is an integer multiple of $\pm 2\pi$. Since $r_e(t)$ can be obtained from Equation (5-67) and $r_p(t)$ can be obtained from Equation (5-66), those modulated carriers that can be demodulated incoherently from $r_e(t)$ can still be demodulated incoherently from $r_{h_i}(t)$ and $r_{h_q}(t)$.

Although at first glance both systems seem to behave identically, there is a very important difference. The homodyne detector output does not exhibit an effect similar to the phase discontinuities observed in the phase detector. The only thing that comes closer to it is the clipping of the tops of $r_{h_i}(t)$ and $r_{h_q}(t)$ when the detectors are overloaded. This can be solved by reducing the gain of the front-end of the receiver. Only if $r_p(t)$ is computed there is the posibility of phase discontinuities. These discontinuities will occur inside the baseband postprocessor and not outside like in the AM/PM demodulator. Remember that $p_r(t)$ is computed only for the PM, FM, CSSB-PM, and CSSB-FM cases, so the only problems are the dynamic range restrictions on D_p and D_f. The troublesome phase functions were those for DSB-LM and SSB-LM that do not have to be computed here.

It would be desireable to extend the dynamic range of the FM frequency deviation constant in order to detect wideband FM. Remember that in the quadrature modulator that was not

a problem because the i(t) and q(t) functions imposed no restrictions on D_f. At the receiver, the problem is that it cannot distinguish between the angle φ and the angle $\varphi \pm 2\pi$, because the receiver cannot remember the previous phase angle. If memory is added to the sub-system that computes $r_p(t)$ from $r_{h_i}(t)$ and $r_{h_q}(t)$, it is possible to eliminate the phase wrapping by making a correction to the phase whenever a large phase jump is detected. A system like this requires time delay. The correction has to be done after the phase jump is detected and not at the instant it occurs because the jumps cannot be predicted. A situation like this may be best realized with digital circuits because the delays can be accounted for easily.

When noise is present the outputs of the homodyne detectors are

$$r_{h_i}(t) = \frac{ki(t)}{2} + \frac{n_i(t)}{2} \tag{5-69}$$

and

$$r_{h_q}(t) = \frac{kq(t)}{2} + \frac{n_q(t)}{2} \tag{5-70}$$

assuming perfect synchronization. Since $n_i(t)$ and $n_q(t)$ appear separately, it may be possible to reduce the effect of noise with special techniques at the baseband postprocessor.

5.4 The AM/FM Demodulator

This section explores the possibility of using the AM/FM demodulator to recover the message from the carriers being

considered throughout this dissertation. The system is an incoherent demodulator and the block diagram was presented in Figure 5.7. The outputs of the demodulators are $r_e(t)$ and $r_f(t)$ given by Equations (5-28) and (5-31), respectively.

In the noiseless case, the AM, CSSB-AM, FM, and CSSB-FM presents no problem since these modulation laws can be demodulated from the envelope or from the instantaneous frequency. The cases that must be considered are DSB-LM, SSB-LM, PM, and CSSB-PM. Remember that the phase function of DSB-LM is used to change the sign of $|x(t)|$, so it has only two values. The output of the balanced frequency discriminator is the derivative of the phase angle, consequently, an impulse should appear at the output of the discriminator whenever there is a sign change. The baseband postprocessor #3 detects the presence of the impulse and reverses the previous sign. This requires some memory to remember the previous sign.

Since differentiation tends to emphasize the highest frequency components, it is reasonable to expect an increase in the bandwidth of $r_f(t)$ over that required for $r_p(t)$. This is critical for DSB-LM and SSB-LM cases whose phase functions are very wideband signals.

The incoherent demodulation of SSB-LM has been studied by Voelcker [60] and Hanson [61]. Their solution is to add a small unmodulated carrier, c, to the SSB-LM carrier at the transmitter so that $(c + x(t)) > 0$. This restricts the phase angle to $0 < p(t) < \pi$. The SSB-LM can be demodulated by envelope detection and a baseband postprocessor. Note that this

requires modifications at the transmitter so it falls outside of what has been defined as universal demodulation. Incoherent demodulation of SSB-LM might be a possibility. Observe that when SSB-LM is at the input of the discriminator, the output is

$$r_f(t) = \frac{d}{dt}\left[p_r(t)\right] = \frac{d}{dt}\left[\tan^{-1}\left[\frac{\hat{x}(t)}{x(t)}\right]\right] , \qquad (5\text{-}71)$$

and that after simplification

$$r_f(t) = \frac{\hat{x}'(t)x(t) - \hat{x}(t)x'(t)}{x^2(t) + \hat{x}^2(t)} . \qquad (5\text{-}72)$$

Since

$$r_e^2(t) = k^2\left[x^2(t) + \hat{x}^2(t)\right] , \qquad (5\text{-}73)$$

then

$$\hat{x}(t) = \pm \frac{1}{k}\sqrt{r_e^2(t) - k^2x^2(t)} , \qquad (5\text{-}74)$$

and Equation (5-74) reduces to

$$\frac{1}{k^2}\left[r_f(t)r_e(t)\right]^2\left[r_e^2 - k^2x^2(t)\right] = \left[r_e'(t)x(t) - x'(t)r_e(t)\right]^2 \qquad (5\text{-}75)$$

which is a nonlinear differential equation. A solution to Equation (5-75) is very likely to involve an infinite series which may be difficult to implement.

To demodulate PM observe that the message can be recovered from the instantaneous frequency by integration

$$x_e(t) = \frac{2\pi}{D_p}\int_{-\infty}^{t} r_f(t')dt' \qquad (5\text{-}76)$$

so PM can be demodulated provided that x(t) has zero mean value.

This discussion shows that it is possible to demodulate most types of modulated carriers with an incoherent AM/FM demodulator provided that some conditions are met.

5.5 Practical Considerations and Comparison of Demodulators

The demodulator configurations that have been suited in this chapter have been assumed to operate under idealized or controlled circumstances. The discussion has been centered on the qualitative aspects of the demodulator systems rather than in strict mathematical analysis. This section considers several aspects that have been ignored in the previous three sections.

It is well known that most detector circuits involve a lowpass filter at the output. This filter is necessary to limit the noise only to the frequency band where there is a modulating signal. In Sections 4.2.1 and 4.3.1 it was shown that the bandwidth of the complex envelope components change with the modulation law and the modulation parameters. If the lowpass filter bandwidth is kept as wide as the largest bandwidth that is expected, the noise performance will be poorer for those modulated carriers whose components require a narrower bandwidth. A solution is to incorporate the output lowpass filter of the detector as part of the baseband postprocessor. The best results will be obtained when the filter

passband is made as narrow as the signal component can tolerate. The improved baseband postprocessor is shown in Figure 5.9.

Another important consideration is the effect of synchronization. In some cases the phase information is important and it is desirable to synchronize the oscillator with the transmitter carrier. If a carrier term is sent by the receiver, synchronization can be achieved by filtering that carrier component, amplifying it,and using it as the reference. This may require a modification at the transmitter and may be equivalent to a change in the modulation law. The other alternative is to send a known reference signal so that the baseband postprocessor can adjust the local oscillator frequency and phase until the received signal is the same as the reference signal. This reference signal can be sent for a short time at the beginning of a transmission. In voice communication, the listener can adjust the reference oscillator frequency

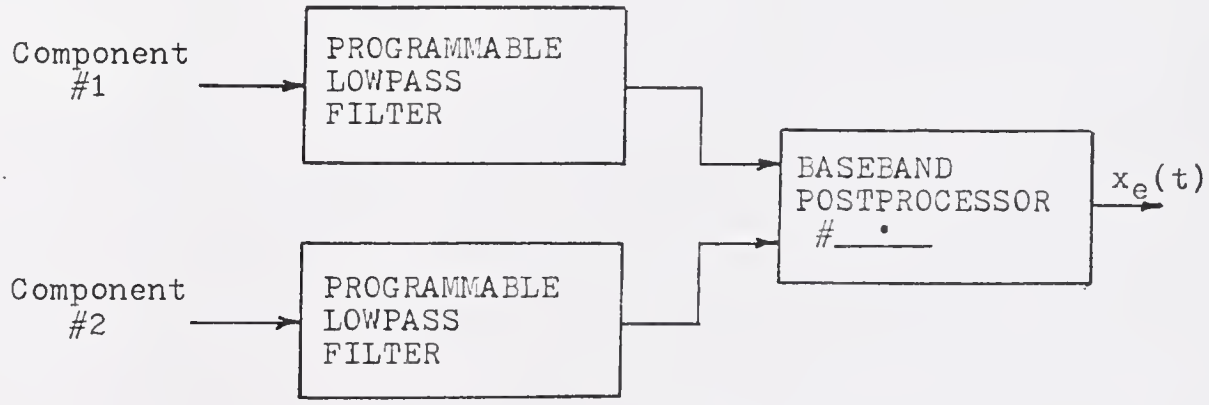


Figure 5.9. Improved baseband postprocessor with programmable lowpass filters.

and phase until the received message is satisfactory.

Finally, it is well known that the envelope detector has a noise threshold level [1]. If the noise power exceeds that level, the recovered message is severely mutilated. This seems to suggest that the envelope detector should be replaced by homodyne detector which has no threshold effect. This is not very restrictive to the designer because the phase reference is already necessary for the phase detector. The output of the homodyne detector is $r_{h_i}(t)$, but $r_e(t)$ can be recovered from

$$r_e(t) = \frac{r_{h_i}(t)}{\cos[r_p(t)]} \cdot \qquad (5\text{-}76)$$

Since this equation is equivalent to an envelope detector, a noise threshold level can be expected to exist. This threshold can be lower than the level of the original envelope detector if some noise is removed from $r_{h_i}(t)$ and $r_p(t)$ before calculating $r_e(t)$ with Equation (5-76). The sub-system discussed here can be described with the block diagram shown in Figure 5.10. The block shown inside the dotted line represent the baseband postprocessor #4.

The quadrature demodulator was the best of the demodulating systems under consideration because it did not have the problem of wideband phase components that the AM/PM demodulator exhibited. The phase detector also suffered from phase discontinuities whenever the phase angle goes from the second to the fourth quadrant. This is very important since it

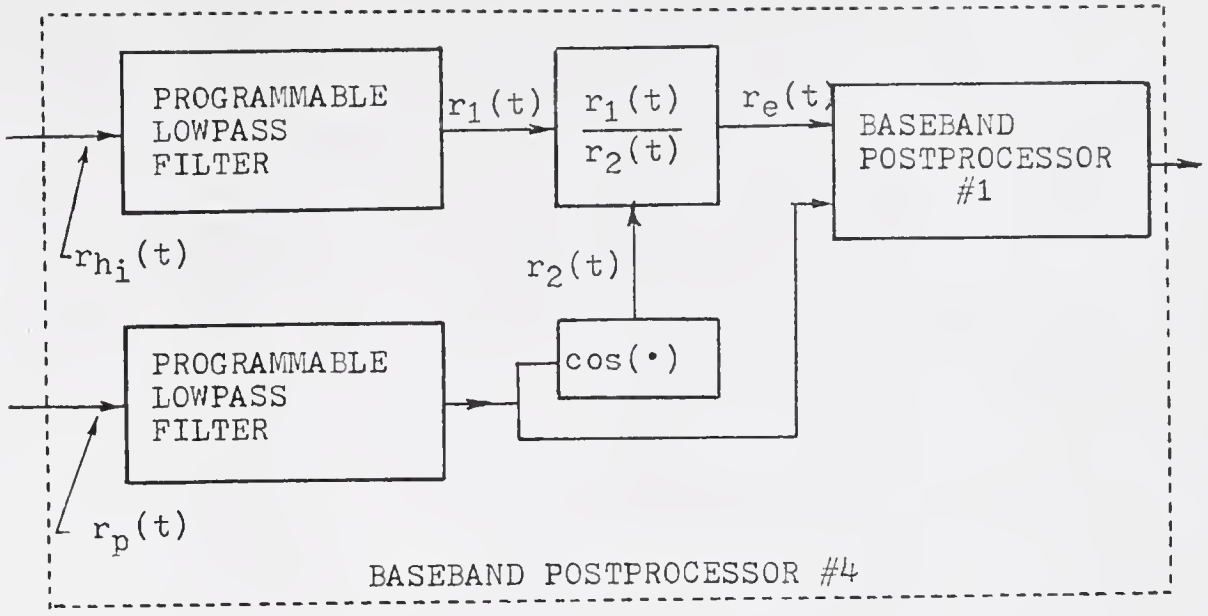


Figure 5.10. Baseband postprocessor #4 for the AM/PM demodulator.

results in wideband signals and the phase detector output bandwidth must be wide enough to allow $r_p(t)$ to jump from one value to another with a very short transition time.

Although the quadrature demodulator seems to require the phase reference at all times, it has been shown that it can demodulate AM, CSSB-AM, CSSB-PM, and CSSB-FM incoherently because the baseband postprocessor can calculate the $r_e(t)$ function using $r_{h_i}(t)$ and $r_{h_q}(t)$ with Equation (5-66).

The incoherent AM/FM demodulator has the disadvantage of not being able to demodulate those cases where the phase function has nonzero mean value. Since the output of the frequency discriminator is equivalent to the derivative of the phase function, $r_f(t)$ was expected to have a wider bandwidth than the original phase function.

5.6 Summary

This chapter considered three different structures to demodulate various types of modulated carriers. Two of these, the AM/PM demodulator and the quadrature demodulator, are coherent in the sense that they require prior knowledge of the received carrier phase angle and frequency. The third system, the AM/FM demodulator, is an incoherent demodulator because the carrier phase is not necessary. This demodulator cannot demodulate any arbitrary modulated carrier because it requires certain conditions for the phase angle.

The quadrature demodulator was found to be superior in performance when compared with the AM/PM and AM/FM demodulators because it can recover almost any type of modulated carrier and does not present a troublesome phase detector.

This chapter did not consider the details of optimizing the reception in the presence of noise. This is an optimum reception problem and is beyond the scope of this dissertation.

CHAPTER VI
CONCLUSIONS

This dissertation studied the problem of obtaining transmitter and receiver structures capable of modulating and demodulating carriers following a variety of modulation laws. The general structure consisted of two modulators and a baseband preprocessor at the transmitter, and two demodulators and a baseband postprocessor at the receiver. The baseband preprocessor changed the message signal into two real-valued signals that were the complex envelope components. These signals were used to modulate the carrier. At the receiver, a pair of demodulators recovered the complex envelope components and the baseband postprocessor estimated the message.

The types of modulated carriers under consideration were amplitude modulation (AM), double-sideband linear modulation (DSB-LM), single-sideband linear modulation (SSB-LM), phase modulation (PM), frequency modulation (FM), compatible single sideband amplitude modulation (CSSB-AM), compatible single sideband phase modulation (CSSB-PM), and compatible single sideband frequency modulation (CSSB-FM).

Chapter III studied some properties of the complex envelope components; namely, the magnitude and phase functions, and the conjugate functions. New results were found for the autocorrelation function of the complex envelope components for most types of modulated carriers under consideration with

the case of a zero-mean gaussian random process signal. Also found were the second moment bandwidths for those signal components whose autocorrelation functions were available.

Chapter IV analyzed the transmitter problem. Various modulator structures were considered and were shown to be equivalent to two basic structures. These structures were named AM/PM modulator and the quadrature modulator, respectively. The PM/PM modulator, a combination of two phase modulators, was shown to be similar to the AM/PM modulator. Important results were the bandwidths that must be allowed for the different complex envelope component signals and the parameter dynamic range limitations imposed by physical and mathematical constraints. The quadrature modulator was found to be the most versatile of all the systems under consideration because it offered the widest parameter dynamic range and the quadrature component signals did not required a lot of bandwidth. This modulator had the disadvantage of requiring more signal baseband processing steps than either the AM/PM or the PM/PM modulators for some modulation laws. The AM/PM and PM/PM modulators weak points were the DSB-LM and SSB-LM phase functions that required a lot of bandwidth compared to the other cases.

The discussion in Chapter IV also pointed out that defining bandwidth in terms of distortion is better than defining it in terms of the second moment bandwidth. Examples were presented where the second moment bandwidth was too small for satisfactory results. In most cases that could be corrected

by adding a constant or using a scale factor. This constant was dependent on the type of modulation, on the complex envelope component, and in some cases, on the message itself.

The receiver problem was discussed in Chapter V from a qualitative point of view. Only two systems seemed to be practical. These were the AM/PM demodulator and the quadrature demodulator. The quadrature demodulator was superior to the AM/PM demodulator because it did not require wide bandwidths as the AM/PM did for the DSB-LM and SSB-LM phase functions. The phase detector output of the AM/PM demodulator was limited to the principal value range. Consequently, if the received phase function exceeds $\pm\pi$, the output of the detector will exhibit phase jumps that make the received phase function discontinuous. This makes the phase function occupy more bandwidth than necessary. These phase discontinuities get worse in the presence of noise or if the receiver and transmitter are not synchronized.

In conclusion, the concept of universal modulation and universal demodulation has been shown to be a practical possibility.

APPENDIX A
AUTOCORRELATION FUNCTION OF $\hat{x}(t)/x(t)$

Define $r(t)$ as

$$r(t) = \frac{\hat{x}(t)}{x(t)} \tag{A-1}$$

which has an autocorrelation function

$$R_r(\tau) = E\left[r(t)r(t+\tau)\right]$$

$$= E\left\{\frac{\hat{x}(t)\hat{x}(t+\tau)}{x(t)x(t+\tau)}\right\}. \tag{A-2}$$

Assume that $x(t)$ is a zero mean gaussian noise process with variance $R_x(0)$ and normalized autocorrelation function $\rho(\tau)$. By definition

$$R(\tau) = \int_{-\infty}^{\infty}\int\int\int\int \frac{x_1 x_2}{x_3 x_4}\, p(x_1, x_2, x_3, x_4)dx_1 dx_2 dx_3 dx_4, \tag{A-3}$$

where

$$x_1 = \hat{x}(t), \tag{A-4}$$

$$x_2 = \hat{x}(t+\tau), \tag{A-5}$$

$$x_3 = x(t), \tag{A-6}$$

and

$$x_4 = x(t+\tau). \tag{A-7}$$

The covariance matrix of $p(x_1, x_2, x_3, x_4)$ is

$$\Lambda = R_x(0) \begin{bmatrix} 1 & \rho(\tau) & 0 & -\hat{\rho}(\tau) \\ \rho(\tau) & 1 & \hat{\rho}(\tau) & 0 \\ 0 & \hat{\rho}(\tau) & 1 & \rho(\tau) \\ -\hat{\rho}(\tau) & 0 & \rho(\tau) & 1 \end{bmatrix} \tag{A-8}$$

which has a determinant:

$$|\Lambda| = R_x^4(0)\left[1 - \rho^2(\tau) - \hat{\rho}^2(\tau)\right] \tag{A-9}$$

and an inverse

$$\Lambda^{-1} = \frac{1}{\left[1 - \rho^2(\tau) - \hat{\rho}^2(\tau)\right] R_x(0)} \cdot$$

$$\cdot \begin{bmatrix} 1 & -\rho(\tau) & 0 & \hat{\rho}(\tau) \\ -\rho(\tau) & 1 & -\hat{\rho}(\tau) & 0 \\ 0 & -\hat{\rho}(\tau) & 1 & -\rho(\tau) \\ \hat{\rho}(\tau) & 0 & -\rho(\tau) & 1 \end{bmatrix} \tag{A-10}$$

The four dimensional gaussian probability density function is

$$p(x) = \frac{|\Lambda^{-1}|^{\frac{1}{2}}}{(2\pi)^2} \exp\left[-\vec{x}'\Lambda^{-1}\vec{x}/2\right], \tag{A-11}$$

where

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \tag{A-12}$$

and $\vec{x}'$ denotes the transpose of $\vec{x}$. Substituting Equations (A-8), (A-9), and (A-5) in Equation (A-10) yields

$$p(x) = \frac{1}{(2\pi)^2 R_x^2(0)\left[1 - \rho^2(\tau) - \hat{\rho}^2(\tau)\right]} \cdot$$

$$\cdot \exp\left[- \frac{\left[x_1^2 + x_2^2 + x_3^2 + x_4^2 - 2\left[\rho(x_1x_2 + x_3x_4) - \hat{\rho}(x_1x_4 - x_2x_3)\right]\right.}{2R_x(0)\left[1 - \rho^2(\tau) - \hat{\rho}^2(\tau)\right]} \right] \tag{A-13}$$

Substitution of Equation (A-13) into Equation (A-3) yields after simplification

$$R_r(\tau) = \frac{1}{2\pi\left[1 - \rho^2(\tau)\right]^{3/2} R_x(0)}$$

$$\cdot \int_{-\infty}^{\infty}\int \left[\frac{\rho(\tau)R_x(0)\left[1 - \rho^2(\tau) - \hat{\rho}^2(\tau)\right] + \hat{\rho}^2(\tau)\left[\hat{\rho}^2(\tau)x_4^2 - x_3x_4\right]}{x_3x_4} \right]$$

$$\cdot \exp \frac{-(x_3^2 - 2\rho(\tau)x_3x_4 + x_4^2)}{2R_x(0)\left[1 - \rho^2(\tau)\right]} \, dx_3 dx_4 \cdot \tag{A-14}$$

This equation can be separated as three integrals

$$I_1 = \frac{\rho(\tau)\, R_x(0)\left[1 - \rho^2(\tau) - \hat{\rho}^2(\tau)\right]}{1 - \rho^2(\tau)} \cdot \frac{1}{2\pi\sqrt{1 - \rho^2(\tau)}\, R_x(0)}$$

$$\int_{-\infty}^{\infty}\int \frac{1}{x_3 x_4} \exp\left[\frac{-\left[x_3^2 - 2\rho(\tau)x_3x_4 + x_4^2\right]}{2R_x(0)\left[1 - \rho^2(\tau)\right]}\right] dx_3 dx_4, \qquad \text{(A-15)}$$

$$I_2 = \frac{\hat{\rho}^2(\tau)\rho(\tau)}{(1 - \rho^2(\tau))} \cdot \frac{1}{2\pi\sqrt{1 - \rho^2(\tau)}\, R_x(0)}$$

$$\int_{-\infty}^{\infty}\int \frac{x_4}{x_3} \exp\left[\frac{-\left[x_3^2 - 2\rho(\tau)x_3x_4 + x_4^2\right]}{2\, R_x(0)\left[1-\rho^2(\tau)\right]}\right] dx_3 dx_4, \qquad \text{(A-16)}$$

and

$$I_3 = \frac{\hat{\rho}^2(\tau)}{1 - \rho^2(\tau)} \quad \frac{1}{2\pi\sqrt{1 - \rho^2(\tau)}\, R_x(0)}$$

$$\int_{-\infty}^{\infty}\int \exp \frac{-(x_3^2 + x_4^2 - 2\rho(\tau)x_3x_4)}{2R_x(0)\left[1 - \rho^2(\tau)\right]} dx_3 dx_4. \qquad \text{(A-17)}$$

These three integrals can be recognized as:

$$I_1 = \frac{\rho(\tau)R_x(0)\left[1 - \rho^2(\tau) - \hat{\rho}^2(\tau)\right]}{1 - \rho^2(\tau)} \cdot E\left[\frac{1}{x_3 x_4}\right] \qquad \text{(A-18)}$$

$$I_2 = \frac{\hat{\rho}^2(\tau)\,\rho(\tau)}{1 - \rho^2(\tau)}\, E\left[\frac{x_4}{x_3}\right] \qquad \text{(A-19)}$$

$$I_3 = \frac{\hat{\rho}^2(\tau)}{1 - \rho^2(\tau)} \cdot \qquad \text{(A-20)}$$

The expectations required for evaluating I_1 and I_2 were

calculated by Gaarder[A].

$$E\left[\frac{1}{x_3 x_4}\right] = -\frac{1}{R_x(\tau)} \ln\left[1 - \rho(\tau)\right] \qquad \text{(A-21)}$$

$$E\left[\frac{x_4}{x_3}\right] = \rho(\tau) \qquad \text{(A-22)}$$

Substitution of Equations (A-14) through (A-18) in Equation (A-10) yields after simplification:

$$R_r(\tau) = \frac{-\left[1 - \rho^2(\tau) - \hat{\rho}^2(\tau)\right] \ln\left[1 - \rho^2(\tau)\right] - \hat{\rho}^2(\tau)\left[1 - \rho^2(\tau)\right]}{1 - \rho^2(\tau)}. \qquad \text{(A-23)}$$

[A] N.T. Gaarder "The covariance of the frequency of a narrowband gaussian random process," IEEE Trans. on Comm. Tech. Vol. COM-14, pp. 659-666, October 1966.

APPENDIX B
COMPUTER PROGRAMS

The following is a sample FORTRAN program used to calculate the intermodulation distortion for the CSSB-PM quadrature modulator.

```
C GENERAL PROGRAM TO CALCULATE THE INTERMODULATION DISTORTION
C  AT THE OUTPUT OF AN IDEAL RECEIVER
C   DEFINE    XM = MESSAGE SIGNAL
C             XH = HILBERT TRANSFORM OF THE MESSAGE
C             XA = REAL COMPLEX ENVELOPE COMPONENT, (I(T))
C             XB = IMAGINARY COMPLEX ENVELOPE COMPONENT, (Q(T))
C             XC = FILTERED I(T)
C             XD = RILTERED Q(T)
       DIMENSION XM(1024),XH(1024),XA(2048),XB(2048),XC(2048),
      &  XD(2048),INV(256),S(256),M(3),XAA(1024),XBB(1024),
      &  OABM(50),OCBM(50),
      &  OCDM(50)
       DATA PI,DELT,M(1),M(2),M(3),IFSET/3.14159,9.765625E-4,
      &  10,0,0,1/
C        DELT IS THE TIME BETWEEN SAMPLES
       PDT=2.0*PI*DELT
C GENERATE THE TEST MESSAGE, XM, AND ITS HILBERT TRANSFORM
       DO  1  I=1,1024
       RID=PDT*FLOAT(I-1)
       XM(I)=0.5*(COS(2.0*RID)+COS(5.0*RID))
     1 XH(I)=0.5*(SIN(2.0*RID)+SIN(5.0*RID))
C GENERATE THE COMPLEX ENVELOPE QUADRATURE COMPONENTS
C  FOR DIFFERENT VALUES OF THE PM MODULATION INDEX (DP)
       DO  2  K=1,9,2
       DP=FLOAT(K)*PI/10.0
       DO  2  I=1,1024
       ARG=-DP*XH(I)
C PREPARE THE ARRAY FOR FAST-FOURIER TRANSFORM
C      STORE REAL PART IN ODD CELLS
C      STORE IMAGINARY PART IN EVEN CELLS
       XA(2*I-1)=EXP(ARG)*COS(DP*XM(I))
       XA(2*I)=0.0
       XB(2*I-1)=EXP(ARG)*SIN(DP*XM(I))
     3 XB(2*I)=0.0
C      STORE THE ORIGINAL TIME ARRAYS
```

Figure B.1. Sample computer program.

```
      DO  31  I=1,1024
      XAA(I)=XA(2*I-1)
   31 XBB(I)=XB(2*I-1)
C     CALCULATE THE FFT OF XA AND XB
      CALL HARM (XA,M,INV,S,IFSET,IFER)
      IFSET=2
      CALL HARM (XB,M,INV,S,IFSET,IFER)
C     THE ARRAY XA AND XB ARE IN THE FREQUENCY DOMAIN
C      SIMULATE THE FILTERING ACTION IN INCREMENTS OF 5 HZ
      DO  4  KI=1,15
      N=KI*5+1
      N1=1024-N+1
      N11=N1+1
      N2=N+1
C     THE PASSBAND OF THE FILTER IS N
      DO  5  I=1,N
      XC(2*I-1)=XA(2*I-1)
      XC(2*I)=XA(2*I)
      XD(2*I-1)=XB(2*I-1)
    5 XD(2*I)=XB(2*I)
      DO  6  I=N2,N1
      XC(2*I-1)=0.0
      XC(2*I)=0.0
      XD(2*I-1)=0.0
    6 XD(2*I)=0.0
      DO  7  I=N11,1024
      XC(2*I-1)=XA(2*I-1)
      XC(2*I)=XA(2*I)
      XD(2*I-1)=XB(2*I-1)
    7 XD(2*I)=XB(2*I)
C    RETURN XC AND XD TO THE TIME DOMAIN
      CALL HARM (XC,M,INV,S,-2,IFER)
      CALL HARM (XD,M,INV,S,-2,IFER)
      DO  8  I=1,1024
      XC(I)=XC(2*I-1)
    8 XD(I)=XD(2*I-1)
C    SIMULATE THE RECEIVER
C    CONSIDER THE CASE WHERE Q(T) IS FILTERED
      CALL DEMOD (XAA,XD,OADM,S,INV,M)
C    CONSIDER THE CASE WHERE I(T) IS FILTERED
      CALL DEMOD (XC,XBB,OCBM,S,INV,M)
C    CONSIDER THE CASE WHERE BOTH ARE FILTERED
      CALL DEMOD (XC,XD,OCBM,S,INV,M)
      WRITE (6,200) K,KI
  200 FORMAT(1H0,10X,'MOD CONSTANT',I3,'BANDWIDTH',I3)
      STOP
      END
```

Figure B.1. continued.

```
      SUBROUTINE DEMOD (XAA,XBB,OUTM,S,INV,M)
C   THIS SUBROUTINE SIMULATES AN IDEAL PM RECEIVER
C      XAA(I) IS THE I(T) COMPLEX ENVELOPE COMPONENT
C      XBB(I) IS THE Q(T) COMPLEX ENVELOPE COMPONENT
C      OUTM(I) IS THE MAGNITUDE SPECTRUM OF THE OUTPUT OF
C         THE IDEAL RECEIVER
      DIMENSION XAA(1),XBB(1),S(1),INV(1),M(1),OPR(2048)
     &,OUTM(1)
      PI=3.14159
      PI2=PI/2.0
      PI22=2.0*PI
      DO  1  I=1,1024
      IF (XAA(I),EQ.0.0) XAA(I)=1.0E-10
      OPR(2*I-1)=ATAN2(XBB(I),XAA(I))
    1 OPR(2*I)=0.0
C   MAKE A PHASE CORRECTION IF NECESSARY
      DO  21  I=2,1024
      DF=OPR(2*I-1)-OPR(2*I-3)
      IF (DF.GT.PI2) OPR(2*I-1)=OPR(2*I-1)-PI22
      IF (DF.LT.PI2) OPR(2*I-1)=OPR(2*I-1)+PI22
   21 CONTINUE
      CALL HARM (OPR,M,INV,S,2,IFER)
      DO  2  I=1,50
      OUTM(I)=SQRT(OPR(2*I-1)**2+OPR(2*I)**2)
    2 CONTINUE
C   CALCULATE THE DISTORTION
C        POWER IN THE MESSAGE FREQUENCIES
      POWER=OUTM(3)**2+OUTM(6)**2
C   POWER IN THE DISTORTION
      TOTAL=OUTM(2)**2+OUTM(4)**2+OUTM(5)**2
      DO  3  I=7,22
    3 TOTAL=OUTM(I)**2+TOTAL
      DIST=SQRT(TOTAL/POWER)*100.0
      WRITE (6,100) DIST
  100 FORMAT (1H0,3X,'IM DIST',2X,f11.6,1X,'X100')
      RETURN
      END
```

Figure B.1. continued.

The subroutine HARM is part of the IBM scientific subroutine library. This subroutine calculates the Fast Fourier Transform of an array.

The listing of a sample program to calculate the power distribution for the CSSB-PM quadrature components is shown in Figure B.2.

```
C   GENERAL PROGRAM TO CALCULATE THE POWER DISTRIBUTION
C       BETWEEN SIDEBANDS AS A FUNCTION OF THE EQUIVALENT
C       FILTER BANDWIDTH.
C   DEFINE   XM=MESSAGE SIGNAL
C            XH=HILBERT TRANSFORM OF THE MESSAGE
C            XA=REAL COMPLEX ENVELOPE COMPONENT, I(T)
C            XB=IMAGINARY COMPLEX ENVELOPE COMPONENT, Q(T)
C            XC=FILTERED I(T)
C            XD=FILTERED Q(T)
      DIMENSION XM(1024),XH(1024),XA(2048),XB(2048),XC(2048)
     &,XD(2048),XE(2048),INV(256),S(256),M(3),XAA(1024),
     &XBB(1024)
      DATA PI,DELT,M(1),M(2),M(3),IFSET/3.14159,9.7656E-4,
     & 10,0,0,1/
C   DELT IS THE TIME BETWEEN SAMPLES
      PDT=2.0*PI*DELT
C   GENERATE THE TEST MESSAGE, XM, AND ITS HILBERT TRANSFORM
      DO  1  I=1,1024
      RID=PDT*FLOAT(I-1)
      XM(I)=0.5*(COS(2.0*RID)+COS(5.0*RID))
    1 XH(I)=0.5*(SIN(2.0*RID)+SIN(5.0*RID))
C   GENERATE THE COMPLEX ENVELOPE QUADRATURE COMPONENTS
C       FOR THE DIFFERENT VALUES OF THE PM MODULATION INDEX
      DO  2  K=1,9,2
      DP=FLOAT(K)*DP/10.0
      DO  3  I=1,1024
      ARG=-DP*XH(I)
C   PREPARE THE ARRAY FOR THE FAST-FOURIER TRANSFORM
C   STORE REAL PART IN ODD CELLS
C   STORE IMAGINARY PART IN EVEN CELLS
      XA(2*I-1)=EXP(ARG)*COS(DP*XM(I))
      XA(2*I)=0.0
      XB(2*I-1)=EXP(ARG)*SIN(DP*XM(I))
    3 XB(2*I)=0.0
C   STORE THE ORIGINAL TIME ARRAYS
      DO  31  I=1,1024
      XAA(I)=XA(2*I-1)
   31 XBB(I)=XB(2*I-1)
C   CALCULATE THE FFT OF XA AND XB
      CALL HARM (XA,M,INV,S,IFSET,IFER)
      IFSET=2
      CALL HARM (XB,M,INV,S,IFSET,IFER)
C   THE ARRAY XA AND XB ARE IN THE FREQUENCY DOMAIN
C   GENERATE THE CONTROL SIGNAL SPECTRUM
      DO  32  I=1,1024
      XE(2*I-1)=XAA(I)
   32 XE(2*I)=XBB(I)
      CALL HARM (XE,M,INV,S,2,IFER)
      CALL AREAS (XE)
C   STUDY THE EFFECT OF BANDLIMITING THE COMPONENT SIGNALS
```

Figure B.2. Sample computer program used to calculate the sideband suppression factor.

```
      DO  4  KI=1,IST
      WRITE (6,200) K,KI
  200 FORMAT(1H0,10X,'MOD CONSTANT',I3,'BANDWIDTH',I3)
      N=KI*5+1
      N1=1024-N+1
      N2=N+1
      N11=N1+1
C   THE PASSBAND OF THE IDEAL FILTER IS N
      DO  5  I=1,N
      XC(2*I-1)=XA(2*I-1)
      XC(2*I)=XA(2*I)
      XD(2*I-1)=XB(2*I-1)
    5 XD(2*I)=XB(2*I)
      DO  6  I=N2,N1
      XC(2*I-1)=0.0
      XC(2*I)=0.0
      XD(2*I-1)=0.0
      XD(2*I)=0.0
      DO  7  I=N11,1024
      XC(2*I-1)=XA(2*I-1)
      XC(2*I)=XA(2*I)
      XD(2*I-1)=XB(2*I-1)
    7 XD(2*I)=XB(2*I)
C   RETURN XC AND XD TO THE TIME DOMAIN
      CALL HARM (XC,M,INV,S,-2,IFER)
      CALL HARM (XD,M,INV,S,-2,IFER)
C   THE ARRAYS ARE IN THE TIME DOMAIN
C    FORM THE COMPLEX ENVELOPE USING THE COMPONENT SIGNALS
C       AND STORE THE REAL PART IN THE ODD CELLS AND THE
C       IMAGINARY PART IN THE EVEN CELLS
      DO  8  I=1,1024
      X=XC(2*I-1)
      Y=XD(2*I-1)
C   CASE WHERE ONLY Q(T) IS FILTERED
      XC(2*I-1)=XAA(I)
      XC(2*I)=Y
C   CASE WHERE ONLY I(T) IS FILTERED
      XD(2*I-1)=X
      XD(2*I)=XBB(I)
C   CASE WHERE BOTH I(T) AND Q(T) ARE FILTERED
      XE(2*I-1)=X
    8 XE(2*I)=Y
C   CHANGE EVERY ARRAY TO THE FREQUENCY DOMAIN
      CALL HARM (XC,M,INV,S,2,IFER)
      CALL HARM (XD,M,INV,S,2,IFER)
      CALL HARM (XE,M,INV,S,2,IFER)
      CALL AREAS (XC)
      CALL AREAS (XD)
      CALL AREAS (XE)
    4 CONTINUE
    2 CONTINUE
      STOP
      END
```

Figure B.2. continued.

```
      SUBROUTINE AREAS (XA)
      DIMENSION XA(2048),Y(1024)
C   THIS SUBROUTINE DETERMINES WHAT PERCENTAGE OF THE TOTAL
C       POWER IS IN THE LOWER SIDEBAND OF THE MODULATED
C       CARRIER BY CALCULATING THE POWER IN THE NEGATIVE
C       FREQUENCY REGION OF THE COMPLEX ENVELOPE SPECTRA.
C       THE POWER IS PROPORTIONAL TO THE AREA UNDER THE COM-
C       PLEX ENVELOPE POWER SPECTRA.
      ARIP=0.0
      ARIN=0.0
      DO  1  I=1,1024
    1 Y(I)=XA(2*I-1)**2+XA(2*I)**2
      Y(1)=0.0
C   FIND THE AREA UNDER THE POSITIVE FREQUENCY REGION
      DO  2  I=2,512
    2 ARIP=ARIP+Y(I)
C   FIND THE AREA UNDER THE NEGATIVE FREQUENCY REGION
      DO  3  I=513,1024
    3 ARIN=ARIN+Y(I)
      RATIO=ARIN/(ARIN+ARIP)
      WRITE (6,100) ARIP,ARIN,RATIO
  100 FORMAT (1H0,10X,'UPPER AREA',1PE16.5,5X,'LOWER AREA',
     & 5X,1PE16.5,5X,'RATIO',0PF10.3)
      RETURN
      END
```

Figure B.2. continued.

REFERENCES

1. A. B. Carlson, Communication Systems, New York: McGraw-Hill, 1968.

2. H. S. Black, Modulation Theory, Princeton, N. J.: Van Nostrand, 1953.

3. J. Dugundji, "Envelopes and pre-envelopes of real waveforms," IRE Trans. Inf. Theory, Vol. IT-4, pp. 53-57, March 1958.

4. E. Bedrosian, "The analytic signal representation of modulated waveforms," Proc. IRE, Vol. 50, pp. 2071-2076, October 1962.

5. H. B. Voelcker, "Toward a unified theory of modulation Part I: Phase-envelope relationships," Proc. IEEE, Vol. 54, pp. 340-353, March 1966.

6. H. B. Voelcker, "Toward a unified theory of modulation Part II: Zero manipulation," Proc. IEEE, Vol. 54, pp. 735-755, May 1966.

7. G. B. Lockhart, "A spectral theory for hybrid modulation," IEEE Trans. Commun, Vol. COM-21, pp. 790-800, July 1973.

8. L. J. Giacoletto, "Generalized theory of multitone amplitude and frequency modulation," Proc. IRE, Vol. 35, pp. 680-693, July 1947.

9. O. G. Villard, "Composite amplitude and phase modulation," Electronics, Vol. 21, pp. 86-89, November 1948.

10. K. H. Powers, "The compatibility problem in single-sideband transmission," Proc. IRE, Vol. 48, pp. 1431-1435, August 1960.

11. L. R. Kahn, "Compatible single sideband," Proc. IRE, Vol. 49, pp. 1503-1527, October 1961.

12. P. F. Panter, Modulation, Noise, and Spectral Analysis, New York: McGraw-Hill, 1965.

13. H. E. Rowe, Signal and Noise in Communication Systems, New York: Van Nostrand, 1965.

14. J. B. Thomas, An Introduction to Statistical Communication Theory, New York: J. Wiley, 1969.

15. M. Schwartz, Information Transmission, Modulation, and Noise, New York: McGraw-Hill, 1970.

16. R. M. Glorioso and E. H. Brazael, "Experiments in 'SSB-FM' communication systems," IEEE Trans. Commun., Vol. COM-13, pp. 109-116, March 1965.

17. J. E. Mazo and J. Salz, "Spectral properties of single sideband angle modulation," IEEE Trans. Commun. Tech., Vol. COM-16, pp. 52-62, February 1968.

18. R. E. Kahn and J. B. Thomas, "Some bandwidth properties of simultaneous amplitude and angle modulation," IEEE Trans. Inf. Theory, Vol. IT-11, pp. 516-520, October 1965.

19. L. W. Couch, "Synthesis and analysis of real single-sideband signals for communication systems," Ph.D. dissertation, Univ. of Florida, Gainesville, Fl., 1968.

20. L. W. Couch, "Signal-to-noise ratio out of an ideal FM detector for SSB-FM plus Gaussian noise at the input," IEEE Trans. Commun., Vol. COM-17, pp. 591-592, October 1969.

21. W. D. Meewezen, "Interrelation and combination of various types of modulation," Proc. IRE, Vol. 48, pp. 1824-1832, November 1960.

22. G. D. Cain, "Compatibility concepts for modulation systems," Ph.D. dissertation, Univ. of Florida, Gainesville, Fl., 1970.

23. P. E. Leuthold and C. Thoeny, "Amplitude modulation by means of two-dimensional semirecursive digital filter," IEEE Trans. Commun., Vol. COM-22, pp. 1172-1179, September 1974.

24. J. B. Thomas and E. Wong, "On the statistical theory of optimum demodulation," IEEE Trans. Inf. Theory, Vol. IT-6, pp. 420-425, September 1960.

25. W. C. Lindsey, Optimum coherent linear demodulation," IEEE Trans. Commun. Tech., Vol. COM-13, pp. 183-190, June 1965.

26. C. J. Boardman and H. L. Van Trees, "Optimum angle modulation," IEEE Trans. Commun. Tech., Vol. COM-13, pp. 452-469, October 1965.

27. H. L. Van Trees, "Analog communication over randomly time varying channels," IEEE Trans. Inf. Theory, Vol. IT-12, pp. 51-63, January 1966.

28. R. E. Kahn and J. B. Thomas, "Bandwidth properties and optimum demodulation of single-sideband FM," IEEE Trans. Commun. Tech., Vol. COM-14, pp. 113-117, April 1966.

29. J. K. Holmes, "Optimum frequency demodulation in the presence of Gaussian noise," Proc. IEEE, pp. 241-242, February 1968.

30. C. N. Kelly and S. C. Gupta, "Discrete-time demodulation of continuous-time signals," IEEE Trans. Inf. Theory, Vol. IT-18, pp. 488-500, July 1972.

31. S. Prasad and A. K. Mahalanabis, "Finite lag receivers for analog communication," IEEE Trans. Commun., Vol. COM-23, pp. 204-212, February 1975.

32. J. P. Costas, "A mathematical analysis of the Kahn compatible single-sideband system," Proc. IRE, Vol. 46, pp. 1396-1401, July 1958.

33. B. F. Logan and M. R. Schroeder, "A solution to the problem of compatible single-sideband transmission," IRE Trans. Inf. Theory, Vol. IT-8, pp. s252-s259, September 1962.

34. N. Abramson, "Nonlinear transformations of random processes," IEEE Trans. Inf. Theory, Vol. IT-13, pp. 502-505, July 1967.

35. S. C. Plotkin, "FM bandwidth as a function of distortion and modulation index," IEEE Trans. Commun. Tech., Vol. COM-15, pp. 467-470, June 1967.

36. O. Shimbo and C. Loo, "Digital computation of FM distortion due to bandpass filters," IEEE Trans. Commun. Tech., Vol. COM-17, pp. 571-574, October 1969.

37. E. Bedrosian and S. O. Rice, "Distortion and crosstalk of linearly filtered angle-modulated signals," Proc. IEEE, Vol. 56, pp. 2-13, January 1968.

38. B. D. Weblock and J. K. Roberge, Electronic Components and Measurements, Englewood Cliffs, N. J.: Prentice-Hall, 1969.

39. L. R. Rabiner and B. Gold, Theory and Application of Digital Signal Processing, Englewood Cliffs, N. J.: Prentice-Hall, 1975.

40. J. L. Dubois and J. S. Aagaard, "An experimental SSB-FM System," IEEE Trans. Commun. Syst., Vol. CS-12, pp. 222-229, June 1964.

41. R. D. Barnard, "On the spectral properties of single-sideband angle modulated signals," Bell Syst. Tech. J., Vol. 43, pp. 2811-2838, November 1974.

42. D. Middleton, An Introduction to Statistical Communication Theory, New York: McGraw-Hill, 1960.

43. B. A. Bowen, "The transform method for nonlinear devices with non-Gaussian noise," IEEE Trans. Inf. Theory, Vol. IT-13, pp. 326-328, April 1967.

44. K. S. Miller, R. I. Bernstein, and L. E. Blumenson, "Generalized Raleigh processes," Q. Applied Math., Vol. XVI, pp. 137-138, 1958.

45. G. A. Korn and T. M. Korn, Mathematical Handbook for Scientists and Engineers, New York: McGraw-Hill, 1968.

46. K. S. Miller, "Moments of complex Gaussian processes," Proc. IEEE, Vol. 56, pp. 83-84, January 1968.

47. D. E. Norgaard, "The phase-shift method of single-sideband generation," Proc. IRE, Vol. 44, pp. 1718-1735, December 1956.

48. H. Chireix, "High power outphasing modulation," Proc. IRE, Vol. 23, pp. 1370-1392, November 1935.

49. D. Slepian, "On bandwidth," Proc. IEEE, Vol. 64, pp. 292-300, March 1976.

50. W. K. Squires and E. Bedrosian, "The computation of single sideband peak power," Proc. IRE, Vol. 48, pp. 123-124, January 1960.

51. G. Clark, "A comparison of current broadcast amplitude modulation techniques," IEEE Trans. Broadcasting, Vol. BC-21, pp. 25-31, June 1975.

52. C. M. Puckette, W. J. Butler, and D. A. Smith, "Bucket brigade transversal filters," IEEE Trans. Circuits Syst., Vol. CAS-21, pp. 502-510, July 1974.

53. D. Childers and A. Durling, Digital Filtering and Signal Processing, St. Paul, Mn.: West, 1975.

54. R. W. Schafer and L. R. Rabiner, "A digital signal processing approach to interpolation," Proc. IEEE, Vol. 61, pp. 692-702, June 1973.

55. A. M. Zayezdnyy and V. L. Lyanders, "Promising circuits for the detection of AM and FM or PM signals," Telecomm. and Radio Eng., Vol. 25, No. 10, pp. 34-37, October 1971.

56. E. Bedrosian, "A product theorem for the Hilbert transform," Proc. IEEE, Vol. 51, pp. 368-369, May 1963.

57. H. Taub and D. L. Schilling, Principles of Communication Systems, New York: McGraw-Hill, 1966.

58. A. J. Viterbi, Principles of Coherent Communications, New York: McGraw-Hill, 1966.

59. W. C. Lindsey and M. K. Simon, Telecommunication Systems Engineering, Englewood Cliffs, N. J.: Prentice-Hall, 1973.

60. H. B. Voelcker, "Demodulation of SSB signals via envelope detection," IEEE Trans. Commun., Vol. COM-14, pp. 22-30, February 1966.

61. J. V. Hanson, "Complex detection: a waveform preserving technique for SSB demodulation," IEEE Trans. Commun., Vol. COM-20, pp. 1097-1100, December 1972.

BIOGRAPHICAL SKETCH

Jorge A. Cruz-Emeric was born on March 6, 1951, in San Juan, Puerto Rico. In July, 1968, he was graduated from Republic of Colombia High School, Rio Piedras, Puerto Rico. The author received a degree of Bachelor of Science in Electrical Engineering from the Mayaguez Campus of the University of Puerto Rico, Mayaguez, Puerto Rico, in June, 1972. After his graduation he worked as electrical engineer for the Puerto Rico Water Resources Authority, Puerto Rico's electric utility company, until August, 1972, when he accepted the position of Instructor in the Department of Electrical Engineering at the University of Puerto Rico, Mayaguez Campus. In June, 1973 he was granted a leave of absence with financial aid to pursue graduate studies at the University of Florida. In January, 1974, he accepted a scholarship from the Economic Development Administration of Puerto Rico. In June, 1974, he obtained the Master of Engineering from the University of Florida.

Mr. Cruz-Emeric is married to the former Sara Hilda Quiñones-Santos and has a daughter, Sara Lydia. He is a member of Tau Beta Pi, the College of Engineers, Architects and Surveyors of Puerto Rico, and student member of the Institute of Electrical Electronic Engineers. In addition he is a licenced graduate engineer in the island of Puerto Rico.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Leon W. Couch, Chairman
Associate Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Donald G. Childers
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William A. Yost
Associate Professor of Speech and Psychology

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1976

Dean, College of Engineering

Dean, Graduate School