MICROWAVE RECEIVERS WITH ELECTRONIC WARFARE APPLICATIONS

JAMES BAO-YEN TSUI

SciTech Publishing, Inc
Raleigh, NC

SciTech Publishing, Inc
Raleigh, NC

Original Edition 1986
Reprint Edition 2005 with corrections

SciTech Publishing
911 Paverstone Drive, Suite B
Raleigh, NC 27615
Phone: 919-847-2434
Fax: 919-847-2568
www.scitechpub.com

10 9 8 7 6 5 4 3 2

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Tsui, James Bao-yen.
Microwave receivers with electronic warfare applications / James Bao-yen Tsui.
p. cm.
Originally published: New York : Wiley, 1986.
Includes bibliographical references and index.
ISBN: 1-891121-40-5
ISBN 13: 978-1-891121-40-1
1. Microwave receivers. 2. Microwaves—Military applications.
I. Title.
[UG485.T75 1992] 92-1257
623.7'3—dc20 CIP

Preface

The purpose of this book is to provide an introduction to microwave receivers. It is written at a level appropriate for advanced undergraduates and first-year graduate students. The rapid development of radar, communication, and weapons guidance systems generates an urgent need for microwave receivers to detect possible threats at the earliest stage of a military mission. The electronic counter countermeasures adapted in modern radar systems make the detection of these radars difficult. The microwave receivers used to intercept these signals must be able to meet these challenges. Thus, microwave receivers have become an important research area because of their applications to electronic warfare. Except for microwave power devices, almost all microwave devices can be used in microwave receiver designs. The advance in microwave receiver development can be considered to be a result of device and logic circuit research. On the other hand, many of the new microwave devices and new ideas in microwave systems are developed as by-products of specific receiver designs. Therefore, this book is written primarily for both microwave receiver engineers and microwave device engineers. A systems engineer needs to know the characteristics of microwave devices, and a device engineer should have enough knowledge about the systems that use his devices. The dependence of receiver performance on device characteristics is discussed in this book.

This book can be divided into two parts: the general properties and the structures of microwave receivers. The properties of microwave receivers are discussed in Chapters 1–3 and in Chapter 12. The structures of different receivers are discussed in Chapters 4–11. The basic properties of microwave devices used in receivers are discussed where appropriate.

The material in this book has been collected from many technical articles, discussions with colleagues and engineers from private industry, as well as the author's experience gained in working on microwave receivers. The recent

receiver, device, and logic developments are especially important for this book. For example, the successful research on radiofrequency (RF) delay lines has made some special receivers possible. Research on microwave receivers and devices is performed in many government and private laboratories. The Avionics Laboratory of the Air Force Wright Aeronautical Laboratories (AFWAL) has contributed a fair share on microwave receiver research.

This book can be considered to augment another book by this author, *Microwave Receivers and Related Components*, published in 1983 by the U.S. Government Printing Office. Approximately one-third of that book is devoted to receivers and the rest to components. In this book, the discussion is concentrated on receivers, and additional receiver designs are introduced; only the basic operating principles of components are discussed. However, cross references are provided for readers interested in components and devices.

The management at the Avionics Laboratory provided excellent support and guidance in the microwave receiver research. Without its support, this book could not have been written. Special thanks are extended to Dr. Charles Krueger, Col. David Niebauer, Mr. William Bahret, and Mr. Paul Hadorn from the Avionics Laboratory and Mr. Joseph Hoffmann, formerly with the Avionics Laboratory, for their valuable advice. I am also indebted to my collegues Mr. Will Brumfield, Mr. Rudy Shaw, Mr. Robert Davis, Mr. Joseph Caschera, Mr. Gerd Schrick, Dr. Richard Sanderson, Mr. Nicholas Pequignot, Mr. Eugene Salzman, and Miss Patti Ballentine for their assistance and consultation. I would like to thank Dr. Gordon Little of the University of Dayton Research Institute for his valuable discussion of Bragg cell receivers. The editing and proofreading of my manuscripts done by Mrs. Christine Carter is greatly appreciated. Special thanks go to Miss Carolyn Ray of AFWAL Technical Library for her assistance in the literature search.

Last, but certainly not least, I wish to thank my wife, Susan, and my children, David and Lisa. Without their encouragement and understanding, it would have been impossible to write this book.

JAMES BAO-YEN TSUI

Centerville, Ohio

Contents

Chapter 1

Introduction

1.1. HISTORICAL REVIEW (1–8)

Since radar was invented during World War II for military uses, its applications have been extended rapidly to other areas. Today radar is used at airports to guide airplanes to safe landings in fog and storms. Airplanes and ships also use radar as a guidance system through poor visibility conditions. Meteorologists show radar maps daily on television for weather forecasting. In military applications, radar has become a vital piece of equipment. It is used not only to search for and detect hostile aircraft, ships, and vehicles but also to guide weaponry.

In military operations, once a new weapon is developed, another will be invented to counteract the first one. Radar development was no exception. Once radar was introduced in World War II, ways to interfere with and defeat its effectiveness were under investigation. These approaches are often referred to as electronic countermeasures (ECMs). Chaff and jammers were popular methods to inhibit a hostile radar. In order to effectively interfere with a radar, one must at least know whether one is a radar target. Therefore, an electronic warfare (EW) receiver (also referred to as an intercept receiver) is needed to detect the existence of hostile radar signals.

Usually, detailed information about a hostile radar is not available. Thus it is impossible to design an intercept receiver as effective as the receiver in the radar system. However, the signal strength at the input of the intercept receiver is much stronger than at the radar receiver because the distance traveled by the radar signal from the source to the intercept receiver is half the distance from the source to the radar receiver, as shown in Figure 1.1. Therefore, if an intercept receiver is properly designed, it can detect the radar signal effectively.

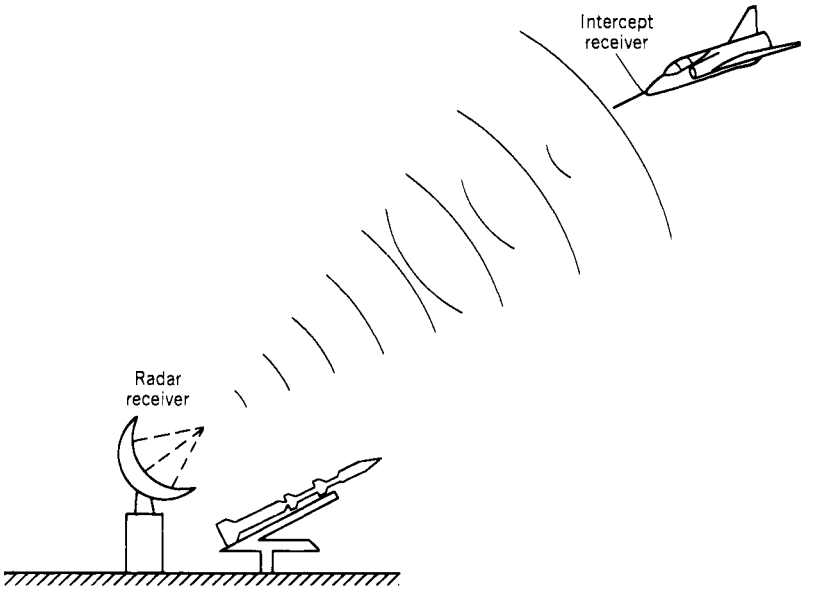


Figure 1.1. An intercept receiver and the victim radar.

The electronic signal environment has become more complicated as radar technology has improved. There are more radars in operation, and, in addition, the radar signals themselves are more sophisticated. As a result, the intercept receivers used to detect the radar signals must be sophisticated enough to cope with them.

1.2. BASIC UNITS OF AN EW RECEIVING SYSTEM

An electrical signal receiving system usually contains an antenna, a receiver, a signal processor, and some display units. A television set or a radio is a good example of a receiving system. The antenna couples the electromagnetic energy from free space and feeds it to the input of the receiver.

The basic functional units of an EW receiving system are shown in Figure 1.2. The receiver receives radio-frequency (RF) signals and generates video outputs. The digitizer converts the video outputs to digital information. The processor analyzes the digital information generated by the receiver to abstract the necessary information. The display unit presents this information in visual or audio forms to the operator. Sometimes the outputs from the signal processor can control some special-function units that will automatically take the necessary action, such as turning on jammers.

The digitizing unit following the receiver shown in Figure 1.2 can be considered as part of the receiver or as part of the signal processor. The receivers

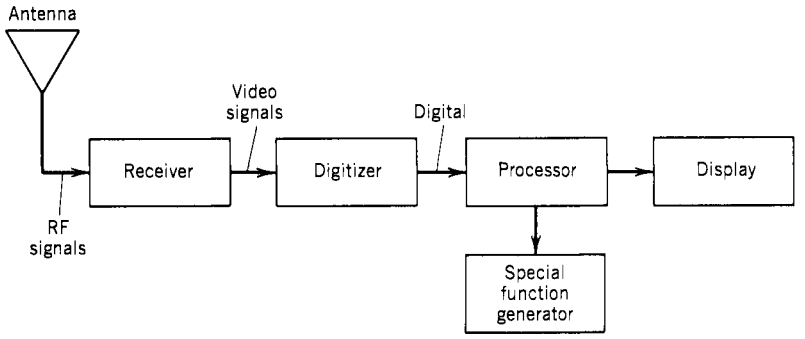


Figure 1.2. Basic EW receiver system.

discussed in this book contain the subsystem receiver and the digitizer. In other words, the receivers discussed take RF signals as input and generate digital words as outputs. Since the detailed circuitry of a digitizer can be very complicated, it is impractical to discuss it in detail. However, the operating principles of the digitizer will be discussed. A wide-band intercept system is shown in Figure 1.3. According to this figure, the subsystem "receiver" covers only the RF converter through the intermediate-frequency (IF) processor.

1.3. CLASSIFICATION OF EW RECEIVERS THROUGH FREQUENCY RANGE (9)

There are basically three ways to classify EW receivers: by their operating frequency, by their applications, and by their structures. In this book, the receivers will be discussed according to their structures. According to their operating frequency, EW receivers can be classified as very high frequency (VHF), ultrahigh frequency (UHF), microwave, and extremely high frequency (EHF) receivers. Sometimes they are simply divided into two groups: communication intercept receivers and radar intercept receivers. There is no clear separation in frequency range between the communication and radar operations. In general, the communication frequency range is in the lower spectrum range and the radar operating frequency is in the higher frequency range. Of course, there is an overlap of the frequency range in which both communication receivers and radar operate.

The microwave frequency range has been divided into bands, with each band designated by a letter or letters. The original purpose was to ensure military secrecy; subsequently the letter designations were adapted for peacetime use. Since these band designations are difficult to remember, new designations have been assigned to frequencies below 40 GHz. Both the new and the old designations are presently used. Table 1.1 lists both the designations below 40 GHz.

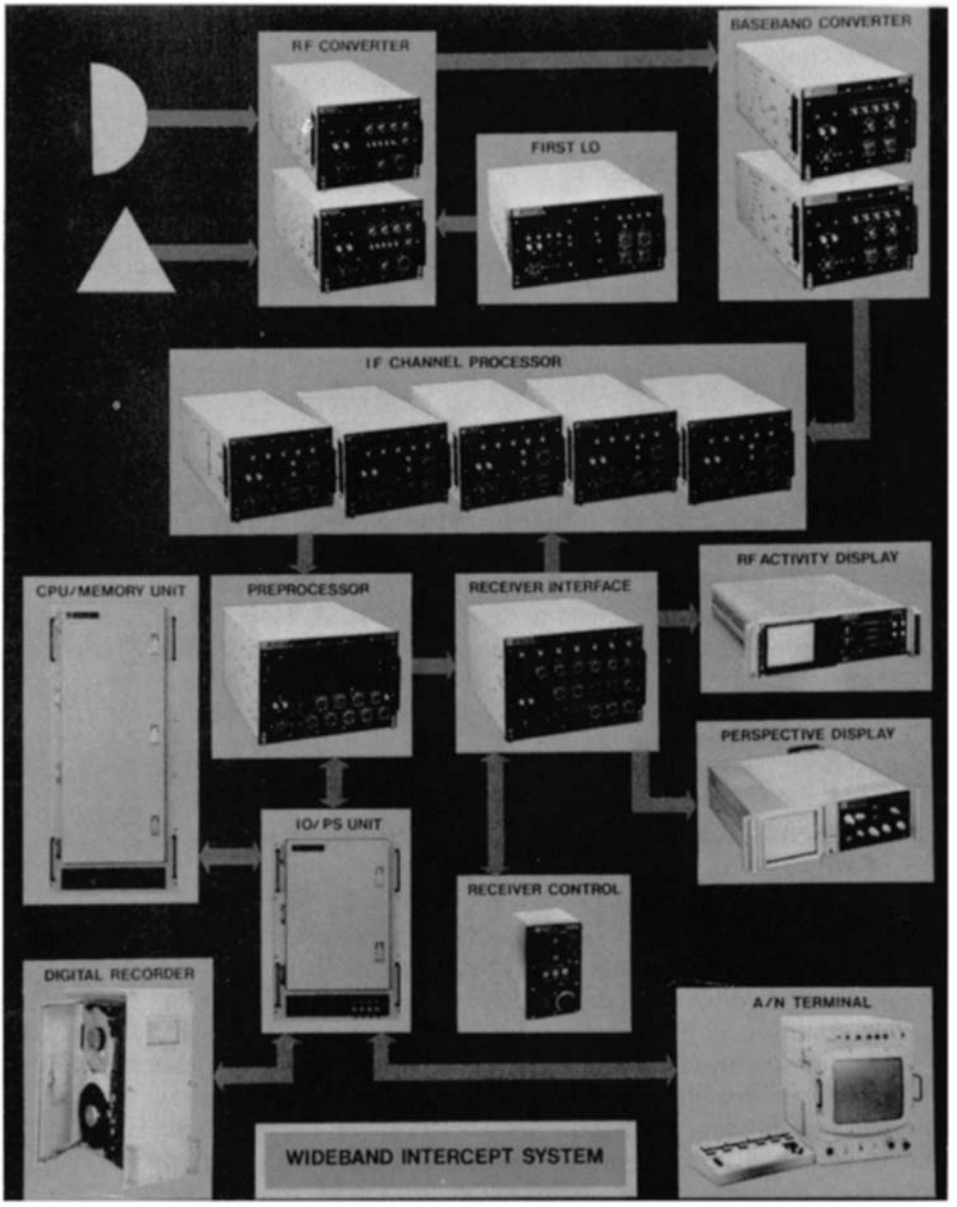


Figure 1.3. Wide-band intercept system. (Courtesy of Watkins-Johnson Co.)

TABLE 1.1. Frequency Band Designations Below 40 GHz

Previous frequency designations	UHF	L	S		C		X	Ku	K		Ka
Current frequency designations	C	D	E	F	G	H	I	J		K	
Frequency (GHz)	0.5	1.0	2.0	3.0	4.0	6.0	8.0	10.0	12.4	18.0	20.0	26.5	40.0

TABLE 1.2. Frequency Band Designations Above 40 GHz

Designation	Q	U	V	E	W	F	D	G
Frequency (GHz)	33–50	40–60	50–75	60–90	75–110	90–140	110–170	140–220

The frequency bands above 40 GHz are listed in Table 1.2. It should be noted that under the current band designations some of the same letters are used for bands below as well as above 40 GHz. For example, the letter D represents the 1–2-GHz band in the low-frequency designation and the 110–170-GHz band in the high-frequency designation. However, actual frequencies can always be used to resolve the ambiguity in frequency band designations.

1.4. CLASSIFICATION OF EW RECEIVERS THROUGH APPLICATIONS (10)

According to their applications, EW receivers can be classified as radar warning receivers (RWRs), radar homing and warning receivers (RHWRs), electronic countermeasure (ECM) receivers, electronic support measure (ESM) receivers, and electronic intelligence (ELINT) receivers. Each receiver type will be briefly discussed in the following paragraphs.

RWR. An RWR is used to detect weapon radar and provide warning to the pilot. A weapon control radar becomes a threat to an aircraft only when the main beam of the radar is directed toward the aircraft. Since there is substantial energy in the main beam of a radar, a receiver with moderate sensitivity is sufficient. Knowing the frequency of the hostile radar will provide useful information for sorting and identification. However, it is not essential to have fine-frequency information for a warning application. In order to cover all the threats, the receiver must have wide frequency and spatial coverage.

RHWR. The function of an RHWR is almost identical to that of an RWR. However, in addition to the warning function, the receiver includes a homing function that will provide passive guidance toward the hostile radar. Passive

ranging or location information is often required for standoff operation and weapon release. In order to provide the homing function, the receiver should generate accurate angle of arrival (AOA) information. The instantaneous angle and frequency coverage of the homing function need not be very wide; however, for the warning function, both the frequency and angle coverage need to be wide.

ECM Receivers. An ECM receiver provides information to a jammer to make it function in the most efficient way. To concentrate the energy of a jammer toward the victim radar, accurate frequency as well as AOA information on the hostile radar is needed. Another important factor of an ECM receiver is that the receiver must measure the input signal as soon as possible so that the jammer can set the desired parameters and jam the same pulse it receives. Thus the receiver often obtains all the information from the leading edge of the pulse. The receiver needs only to cover the frequency and the AOA range of the victim radar.

ESM Receivers. An ESM receiver is used to obtain all the information from an electronic order battle (EOB). In other words, the receiver should collect information on all radars in the battlefield. Therefore, the receiver needs to cover wide frequency ranges and wide spatial angles. In order to detect radars that are not pointed toward the receiver, the side lobes of the radar should also be detected. Therefore, the sensitivity of the receiver should be very high. In order to provide enough information for a signal processor to sort the environment, the receiver must measure all the parameters of the input signals with fine-grain resolution. In particular, the receiver should generate fine-frequency and AOA resolutions.

ELINT Receivers. An ELINT receiver is used to receive special signals, isolate them, and perform a fine-grain analysis on them. The input frequency range of the receiver should be wide enough to cover the signals of interest. However, the instantaneous bandwidth of the receiver need only be wide enough to cover one or a few signals of interest. The receiver should provide very fine measurements on all the parameters of the hostile radar. If the data cannot be analyzed at the collection station, it can be stored and analyzed at a later time. Thus the receiver does not have to process signals in near real time, as do the other receivers discussed.

1.5. CLASSIFICATION OF EW RECEIVERS THROUGH STRUCTURES (10–13)

According to the receiver structures (see Table 1.3), they can be classified as

crystal video,
superheterodyne (superhet),

instantaneous frequency measurement (IFM),
channelized,
compressive (microscan), and
Bragg cell receivers.

The discussion in this book will concentrate on the operating principles of receivers and their designs. Therefore, the receivers are discussed according to their structures. Each type of receiver is discussed in a separate chapter.

Crystal video and superhet receivers can be considered in a ready production stage. IFM receivers can be considered well developed. The other three types, channelized, compressive, and Bragg cell, are still under development at various stages. Each receiver has its own figure of merit. However, only the anticipated performances of the latter three types of receivers will be discussed here. It should also be emphasized that these anticipated performances are subject to change if some new devices or technologies are discovered. The performances of these receivers are listed in Table 1.3. They are only qualitative estimations rather than actual data.

In the following paragraphs, the potential applications of each type of receiver will be discussed briefly.

Crystal video receivers can be used as RWRs and RHWRs because of their small size and wide frequency coverage.

Superhet receivers can be used in almost any application because of their high selectivity, sensitivity, and dynamic range. However, because of their narrow-input bandwidth, this type of receiver is usually used in conjunction with other types of receivers in EW applications.

IFM receivers can be used as RWRs, RHWRs, and ECM receivers because

TABLE 1.3. Electronic Warfare Receivers

	Channel	Compressive	Optical	IFM	Crystal	Superhet
Instantaneous frequency bandwidth	Good	Good	Good	Excellent	Excellent	Poor
Capability to handle simultaneous signals	Good	Good	Good	Poor	Poor	Poor
Frequency resolution	Good	Good	Good	Good	Poor	Excellent
Sensitivity	Good	Good	Fair to good	Fair to good	Poor to Fair	Excellent
Dynamic range	Good	Fair to good	Fair	Fair to good	Fair	Excellent

of their wide input frequency range and small size. They may be used in conjunction with other receivers and as ESM and ELINT receivers.

Channelized receivers can be used as ESM and ELINT receivers because of their superior performance. Receivers with coarse channelization may be used as RWRs and RHWRs because of their potentially small size.

Compressive receivers can be used as ESM and ELINT receivers because of their high performance.

Bragg cell receivers can be used as ECM, ESM, and ELINT receivers because of their potentially high performance and small size.

1.6. FUTURE TREND OF EW RECEIVERS (14–19)

The research and development of EW receivers may concentrate on the channelized, compressive, and Bragg cell receivers because they are still in the developmental stages and have the potential for high performance. Development of new microwave devices and breakthroughs in present microwave devices will be the key factor to impact future EW receiver designs.

The complexity of the electromagnetic environment will push the utilization of hybrid EW receivers. A hybrid receiver consists of several types of generic receivers, since one kind of receiver may not be able to solve a specific problem. These receivers will function together as one receiving unit.

The advancing technology in microwave integrated circuits (MICs) will provide the potential to make the size of EW receivers very small, at the same time improving their reliability. In the past, microwave systems consisted of a large number of devices, each performing a very specific function. Thus the overall system is very complicated. The concept of supercomponents, using the MIC approach, was developed to integrate many components into one device. The supercomponents will be designed and manufactured by device engineers rather than by systems engineers. This approach will optimize interfacing for best performance, maximum reliability, and reduction in size and weight. In the future, MICs and supercomponents will be widely adapted in EW receiver systems.

The development of very high speed integrated circuits (VHSIC) will not only revolutionize the designs of digitizing circuits but may also provide the possibility of digitizing the input RF signals directly. If the input signals can be digitized directly, special processors should be built to analyze the signals. This approach may change the overall receiver design concepts, but it may not happen in the near future.

Developing an EW receiver is a very expensive and time-consuming task. It may take many years from the time an idea is conceived to the time an experimental model is made. Usually, after the experimental receiver is built, one will find many "obvious mistakes." Some of these mistakes can be vital to the original idea. If one can adequately model the performance of a new design before

building it, the obvious mistakes may be revealed or even avoided. Unfortunately, in order to generate an effective modeling program, one needs to have some experience in actually building receivers. From past experience in various receiver developments and the advance in computer technology, it is now possible to generate some useful receiver modeling programs.

1.7. ORGANIZATION OF THE BOOK

This book contains 12 chapters. In Chapter 2, the general characteristics of a microwave receiver and the common terms used in this field will be discussed. In Chapter 3, the general parameters measured by a wide-band RF receiver will be discussed. The major difference among various receiver structures is in the frequency measurement scheme. The other parameter measurement schemes are somewhat similar. Therefore, these schemes of measuring parameters other than frequency will be included in Chapter 3.

Six types of generic receivers will be discussed in Chapters 4–9. The operating principle of each receiver is emphasized. Some microwave components used in the receiver are also discussed whenever appropriate. The performances and applications of the receivers will only be discussed briefly.

In Chapter 10, the possibility of using different generic receivers to form some kinds of hybrid receivers that can achieve some special performance requirements will be discussed. In Chapter 11, the EHF receivers will be discussed. Although the structure of EHF receivers are included in the six generic receivers, the receivers are built differently. Whereas at lower microwave frequencies the receivers are mainly built in coaxial cable, stripline, or microstrip forms, EHF receivers are built in waveguide, fin line, or suspended stripline forms. Therefore a separate chapter is devoted on this subject. In the last chapter, the procedures of evaluating an EW receiver will be discussed. The receiver characteristics discussed in Chapter 2 will be measured experimentally in Chapter 12.

1.8. SPECIFIC REMARKS

Many abbreviations and acronyms are used in EW and related areas. To familiarize the readers with these terms, they are also used in this book. The first time an abbreviation is used in a chapter it will be spelled out, although some of them are commonly used. These abbreviations are also listed in Table 1.4 for quick reference.

Some of the figures presented in this book are generated from computer modeling programs. They are used to illustrate certain ideas and to help the reader to understand. In general, these simulated results are very close to experimental data.

TABLE 1.4. Table of Abbreviations

A/D	Analog to digital	NEP	Noise-equivalent power
APD	Avalanche photodetector	PA	Pulse amplitude
AOA	Angle of arrival	PC	Printed circuit
CCD	Charged coupled device	PIN	*p*-material–intrinsic–*n*-material
CMC	Convolve–multiply–convolve	PLL	Phase-locked loop
DF	Direction finding	POI	Probability of intercept
DOA	Direction of arrival	PRF	Pulse repetition frequency
ECM	Electronic countermeasures	PRI	Pulse repetition interval
EHF	Extremely high frequency	PW	Pulse width
ELINT	Electronic intelligence	RF	Radio frequency
EM	Electromagnetic	RHWR	Radar homing and warning receiver
EOB	Electronic order battle		
ESM	Electronic support measures	ROM	Read-only memory
EW	Electronic warfare	RWR	Radar warning receiver
FET	Field-effect transistor	SAW	Surface acoustic wave
FFT	Fast Fourier transform	S/H	Sample and hold
GGG	Gadolinium gallium garnet	S/N	Signal-to-noise ratio
IF	Intermediate frequency	TOA	Time of arrival
IFM	Instantaneous frequency measurement	TOD	Time of departure
		TRF	Tuned radio frequency
IOC	Integrated optic circuit	TSS	Tangential sensitivity
I–V	Current versus voltage	TWT	Traveling wave tube
LO	Local oscillator	VCO	Voltage-controlled oscillator
LSB	Least significant bit	VHSIC	Very high speed integrated circuit
MCM	Multiply–convolve–multiply		
MDS	Minimum discernible signal	VLSIC	Very large scale integrated circuit
MIC	Microwave integrated circuit		
MMIC	Microwave monolithic IC	VSWR	Voltage standing wave ratio
MOS	Metal–oxide semiconductor	YAG	Yttrium aluminum garnet
MSB	Most significant bit	YIG	Yttrium iron garnet

REFERENCES

1. L. N. Ridenour, *Radar System Engineering* (Radiation Laboratory Series), McGraw-Hill, New York, 1947.
2. S. N. Van Voorhis, *Microwave Receivers* (Radiation Laboratory Series), McGraw-Hill, New York, 1948.
3. J. F. Reintjes and G. T. Coate, *Principles of Radar*, McGraw-Hill, New York, 1952.
4. D. K. Barton, *Radar System Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1964.
5. R. S. Berkowitz, Ed., *Modern Radar*, Wiley, New York, 1965.
6. F. E. Nathanson, *Radar Design Principles*, McGraw-Hill, New York, 1967.
7. M. I. Skolnik, *Radar Handbook*, McGraw-Hill, New York, 1970.
8. L. B. Van Brunt, *Applied ECM*, EW Engineering Inc., Dunn Loring, VA, 1982.
9. T. K. Ishii, *Microwave Engineering*, Ronald Press, New York, 1966.
10. J. F. Hoffmann and J. B. Y. Tsui, Survey of EW receiver techniques, SPIE meeting, Washington, D.C., May 1–2, 1984.

11. F. Rappolt and N. Stone, Receivers for signal acquisition, *Microwave J.*, **20**, 29 (January 1977).
12. T. Harper, Hybridization of competitive receivers, Watkins Johnson Co. Tech Notes, **7**(1), Jan/Feb 1980.
13. C. B. Hofmann and A. R. Baron, Wideband ESM receiving systems, *Microwave J.*, **23**, 24 (September 1980).
14. J. B. Cochrane and F. A. Markl, Broadband building blocks shape tomorrow's warning receivers, *Microwaves*, (November, 1980).
15. J. H. Collins and P. M. Grant, A review of current and future components for electronic warfare receivers, *IEEE Trans. Microwave Theory Techniques*, **MTT-29**, 395–403 (1981).
16. Microwave receiver versus spectrum analyzer, *Microwave J.*, Application Note, 50 (August 1973).
17. Multiband receiver, *Microwave J.*, **26** 228 (May 1983).
18. K. D. Gilbert and J. B. Sorci, Jr., Microwave supercomponents fulfill expectations, *Microwave J.*, **26**, 67 (November 1983).
19. W. Tsai, R. Gray, and A. Graziano, The design and performance of supercomponents: High density MIC modules, *Microwave J.*, **26**, 81 (November 1983).

Chapter 2

Characteristics of Microwave Receivers

2.1. INTRODUCTION

In this chapter, some common terms representing the characteristics and performance of EW receivers will be discussed. These terms include sensitivity, dynamic range, probability of intercept (POI), and throughput rate. Although used on a daily basis, some of these terms are not clearly defined and others have too many definitions. This chapter will discuss the different definitions and try to clarify them. Among all the terms, emphasis will be placed on receiver sensitivity and dynamic range. Important concepts used to design receivers will also be included.

In the discussion of receiver sensitivity, the effect of video bandwidth will be included. The sensitivity is a function of the ratio of radio-frequency (RF) bandwidth to video bandwidth. Curves will be provided to predict the receiver sensitivity.

Most of the receiver measurements will be discussed in Chapter 12. However, the measurement of false alarm rates will be discussed in this chapter, because it is closely related to the sensitivity of the receiver. The equations used to calculate the sensitivity of a receiver are also used to calculate the false alarm rate. Examples will be presented to demonstrate utilization of the equations and curves.

2.2. THERMAL NOISE AND NOISE FIGURE (1–6)

In most microwave receivers developed today, there is a video detector (or detectors) that converts RF energy to video signals. The sensitivity of a receiver

is limited either by the characteristics of the video detector or by the internally generated noise of the receiver. If the RF gain in front of the video detector is high enough, the sensitivity of the receiver is limited by the receiver noise level; otherwise, the sensitivity is determined by the video detector. Thermal noise generated by the thermal motion of electrons in all components of a receiver is always present. Noise generated by a resistor R can be represented by a noise generator in series with the resistor. Maximum power transfer from a generator to a load occurs when the load impedance is matched to the generator impedance. Available power refers to the power that would be delivered to a matched load. The available thermal noise power at the input of a receiver can be expressed as

$$N_i = kTB \tag{2.1}$$

where k is Boltzmann's constant ($=1.38 \times 10^{-23}$ J/°K,) T is the temperature of resistor R (R is not included in the above equation) in kelvin, (refs. 4, 6) B is the bandwidth of the receiver in hertz, and N_i is noise power in watts. The power level in a receiver system is often very low and is expressed in milliwatts. Another common expression of power is in dBm, which is defined as

$$P\,(\text{dBm}) = 10 \log(P) \tag{2.2}$$

where the P on the right side is power in milliwatts. For $P = 1$ mW, P(dBm) = 0. When $P > 1$ mW, P(dBm) is positive. When $P < 1$ mW, P(dBm) is negative. The thermal noise at room temperature (where $T = 290$°K can be represented in dBm as

$$P\,(\text{dBm}) = -174\ \text{dBm/Hz} \qquad \text{or} \qquad P\,(\text{dBm}) = -114\ \text{dBm/MHz} \tag{2.3}$$

Equation (2.3) is derived by substituting the values of k and T in Eqs. (2.1) and (2.2). These two values are often used in determining receiver sensitivity. It is common practice to calculate receiver sensitivity at room temperature because the sensitivity and false alarm rate of a receiver is often measured by connecting a signal source or a matching load to the input of the receiver. However, if the input of a receiver is an antenna facing the cold sky, the temperature T could be much lower. The noise power of a practical receiver is always higher than the thermal noise of an ideal receiver because noise is introduced from every component in the receiver. The noise figure is defined as

$$F = \frac{N_o}{GN_i} = \frac{\text{noise output of practical receiver}}{\text{noise output of an ideal receiver at temperature } T} \tag{2.4}$$

where N_o is the noise output from the receiver, G is the RF gain of the receiver, and N_i is the input thermal noise ($=kTB$). To standardize the definition, the Institute of Radio Engineers specifies that room temperature ($T = 290$°K) be used in Eq. (2.4) (refs. 5, 6). Since an ideal receiver adds no noise of its own

to the signal being processed, the noise output is the input noise amplified by the gain G of the receiver.

The gain of a receiver is defined as

$$G = S_o/S_i \tag{2.5}$$

where S_o and S_i are the available output and the input signal powers, respectively. Combining Eqs. (2.4) and (2.5), one obtains

$$F = \frac{S_i/N_i}{S_o/N_o} = \frac{\text{signal-to-noise ratio at input of receiver}}{\text{signal-to-noise ratio at output of receiver}} \tag{2.6}$$

Since the input signal-to-noise ratio (S_i/N_i) is always greater than the output signal-to-noise ratio (S_o/N_o), F is always greater than unity. The noise figure F can be expressed in decibels as

$$F\,(\text{dB}) = 10\log(F) \tag{2.7}$$

If there are N amplifiers connected in cascade, the noise figure can be expressed as

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \cdots + \frac{F_N - 1}{G_1 G_2 \cdots G_{N-1}} \tag{2.8}$$

where $G_1, G_2, \ldots$ and $F_1, F_2, \ldots$ are the gain and noise figure of the first, second, ... amplifiers, respectively, and they are expressed in power ratios rather than in decibels.

The derivation of Eq. (2.8) can be illustrated with two amplifiers connected in cascade as shown in Figure 2.1. The noise at the output of amplifier 2 can be written as

$$\begin{aligned} N_o = {} & \text{noise from amplifier 1 at output of amplifier 2} \\ & + \text{noise introduced by amplifier 2} \end{aligned} \tag{2.9}$$

The noise contributed from amplifier 1 is $kTBF_1$ amplified by G_1 and G_2. The noise generated by amplifier 2 is G_2kTBF_2, which contains the noise G_2kTB generated from amplifier 1. Thus the noise generated from amplifier 2 alone is

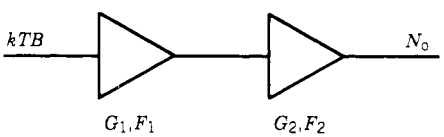


Figure 2.1. Two amplifiers in cascade.

$G_2kTBF_2 - G_2kTB$. Therefore Eq. (2.9) can be written as

$$N_o = kTBF_1G_1G_2 + G_2kTBF_2 - G_2kTB \tag{2.10}$$

From the definition of noise figure given in Eq. (2.4), the overall noise figure of the two cascade amplifiers can be expressed as

$$F = \frac{kTBF_1G_1G_2 + G_2kTBF_2 - G_2kTB}{kTBG_1G_2} = F_1 + \frac{F_2 - 1}{G_1} \tag{2.11}$$

The contribution of the second amplifier to the overall noise figure may be negligible if the gain of the first amplifier is very high. This derivation can be extended to an amplifier chain of more than two stages as shown in Eq. (2.8).

If the first component in an amplifier chain is a passive device, that is, a filter followed by an amplifier, there are two ways to calculate the system noise figure. One way is to use Eq. (2.8) directly. The gain of the first component is less than 1 (or a minus value in decibels) and the noise figure of the component is equal to the inverse of the gain. The other way is to use the second component as the first one in Eq. (2.8) to calculate the noise figure and then multiply the noise of the first component (or add the loss of the first component in decibels) to the calculated value to obtain the overall noise figure.

For example, Figure 2.2 shows an amplifier chain with three components. The first is a filter with a 3-dB loss; the second is an amplifier with a 20-dB gain and a 4-dB noise figure; and the third is a 2-dB attenuator. Using the first approach, $G_1 = \frac{1}{2}(-3\text{ dB})$ and $F_1 = 2(3\text{ dB})$, $G_2 = 100(20\text{ dB})$ and $F_2 = 2.5(4\text{ dB})$, $G_3 = (1/1.58)(-2\text{ dB})$ and $F_3 = 1.58(2\text{ dB})$, the calculated result is $F = 2 + 3 + 0.012 = 5.012(7\text{ dB})$.

Using the second approach, $G_1 = 100(20\text{ dB})$ and $F_1 = 2.5(4\text{ dB})$ and $G_2 = (1/1.58)(-2\text{ dB})$ and $F_2 = 1.58(2\text{ dB})$, $F = 2.5 + 0.006 = 2.506(4\text{ dB})$. The overall noise figure $F = 2 \times 2.506 = 5.012(3 + 4 = 7\text{ dB})$.

In the first approach, the percentage of noise contributed by each component can be found easily. The first component contributes 39.9% (2/5.012%), the second component contributes 59.9% (3/5.012%), and the third component contributes 0.2% (0.012/5.012%). This is a very useful relation in receiver design. It shows which component contributes most of the noise to the system.

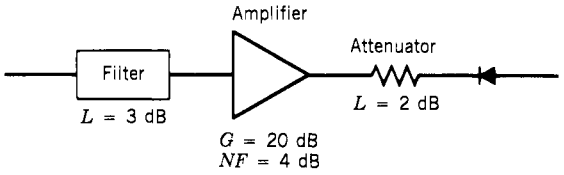


Figure 2.2. A simple RF chain.

2.3. TANGENTIAL SENSITIVITY (7–23)

The tangential sensitivity (TSS) of a receiver is measured through visual display on an oscilloscope monitoring the output of the diode detector or the output of the video amplifier following the detector. The input must be a pulsed signal. On the scope display, when the minimum of the noise trace in the pulse region is roughly tangential to the top of the noise trace between pulses, as shown in Figure 2.3, the receiver is at its TSS. Frohmaier (ref. 7) experimentally determined that at the TSS the signal is 8 dB above the noise level at the output of the detector with a standard deviation of 0.4 dB. Experiments carried out by Williams (ref. 8) indicate that the spread among a group of observers in setting to a tangential level is likely to be 1 dB about the mean. The TSS is a very widely used criterion because of its extreme simplicity and its ability to give a convenient comparison of the sensitivities of widely different receivers. Its disadvantage is that in some receivers the output of the detector or the video amplifier is not easily monitored.

The TSS depends on the RF bandwidth (B_R), the video bandwidth (B_V), the noise figure of the receiver, and the characteristics of the detector. In a receiver, B_R is almost always greater than B_V. The discussion on TSS in this section is based on Lucas's article (ref. 9). The signal strength TSS required at the input of a detector to produce the tangential sensitivity at the output is derived by

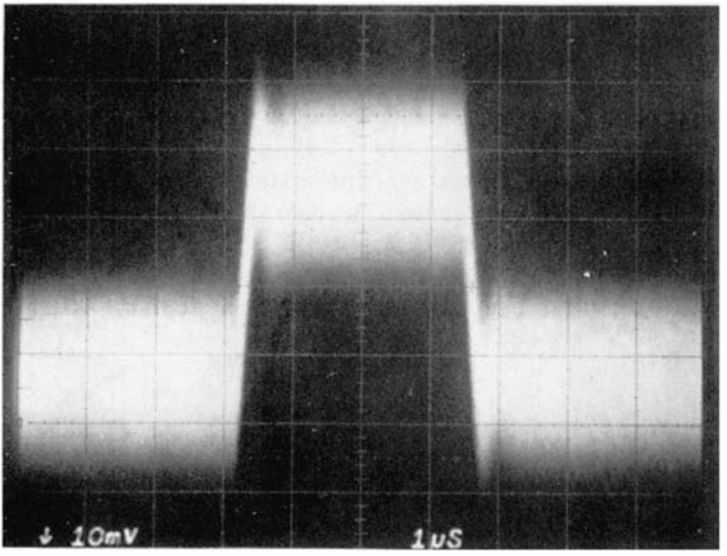


Figure 2.3. Detector output at tangential sensitivity.

Lucas. The results are rewritten in a different format as

$$\text{TSS} = -114 + 10\log F_T + 10\log\left(3.15B_R + 2.5\sqrt{2B_RB_V - B_V^2 + \frac{AB_V}{(G_TF_T)^2}}\right)$$

$$\text{for } B_V \leqslant B_R < 2B_V \tag{2.12}$$

$$\text{TSS} = -114 + 10\log F_T + 10\log\left(6.31B_V + 2.5\sqrt{2B_RB_V - B_V^2 + \frac{AB_V}{(G_TF_T)^2}}\right)$$

$$\text{for } B_R \geqslant 2B_V \tag{2.13}$$

where G_T and F_T are the overall gain and noise figure from the input of the receiver to the detector and B_R and B_V are the RF and video bandwidth measured in megahertz. The value -114 is from Eq. (2.3) and is the thermal noise floor of a 1-MHz bandwidth system at room temperature. The constant A is related to the diode characteristic and the noise figure of the video amplifier following the diode. It can be expressed as

$$A = \frac{4F_V}{kTM^2} \times 10^{-6} \tag{2.14}$$

where F_V is the noise figure of the video amplifier expressed in power ratio and M is the figure of merit of the diode detector and can be expressed as (ref. 10)

$$M = \beta/\sqrt{R} \tag{2.15}$$

where β is the detector sensitivity in volts per watts and R is the dynamic impedance of the diode in ohms.

The value of A can be determined in one of the following ways:

1. If the figure of merit M (or β and R) and the noise figure F_V are given, A can be calculated from Eqs. (2.14) and (2.15).

2. It can be measured experimentally. If the video detector and the video amplifier under test are considered as the only components in a receiver and there is no gain or loss in front of the detector, then $G_T = F_T = 1$ in Eqs. (2.12) and (2.13). Under this condition, A is the only dominant term in these equations and these two equations can be approximated by

$$\text{TSS} = -110 + 10\log\sqrt{AB_V} \tag{2.16}$$

By measuring the TSS experimentally, A can be determined by Eq. (2.16). The error introduced in this approximation is usually less than the measurement error.

The sensitivity of a receiver can be considered under two different situations:

1. *RF Gain-Limited Case.* Under this condition, there is not enough RF gain in front of the detector, and the sensitivity of the receiver is determined primarily by the detector.

2. *Noise-Limited Case.* Under this condition, there is adequate RF gain before the detector. The sensitivity of the receiver is limited by the noise figure F_T rather than by the detector.

If it is arbitrarily chosen that

$$\frac{AB_V}{(G_T F_T)^2} < 0.2(2B_R B_V - B_V^2)$$

or

$$(G_T F_T)^2 > \frac{AB_V}{0.2(2B_R B_V - B_V^2)} \tag{2.17}$$

the term $AB_V/(G_T F_T)^2$ in Eqs. (2.12) and (2.13) can be neglected and the error introduced will be less than 0.4 dB. The two equations can be written as

$$\text{TSS} = -114 + 10 \log F_T + 10 \log(3.15B_R + 2.5\sqrt{2B_R B_V - B_V^2})$$

$$\text{for } B_V \leqslant B_R < 2B_V \tag{2.18}$$

$$\text{TSS} = -114 + 10 \log F_T + 10 \log(6.31B_V + 2.5\sqrt{2B_R B_V - B_V^2})$$

$$\text{for } B_R \geqslant 2B_V \tag{2.19}$$

Equations (2.18) and (2.19) are commonly used since in most receiver designs there is usually enough RF gain provided in front of the detector. In Eq. (2.19), the value of $\gamma = B_R/B_V$ can spread over a wide range, especially in wide-band receivers. Often, it is desirable to read the TSS graphically. Equation (2.19) can be written as

$$\text{TSS} = [-114 + 10 \log(6.31 + 2.5\sqrt{2\gamma - 1})] + 10 \log F_T + 10 \log B_V$$

$$= T(\gamma) + 10 \log F_T + 10 \log B_V \quad \text{dBm} \tag{2.20}$$

The value in the square brackets, expressed as $T(\gamma)$, versus the value of γ is plotted in Figure 2.4. To find the TSS value, just read the value $T(\gamma)$ from the curve and add $10 \log F_T + 10 \log B_V$.

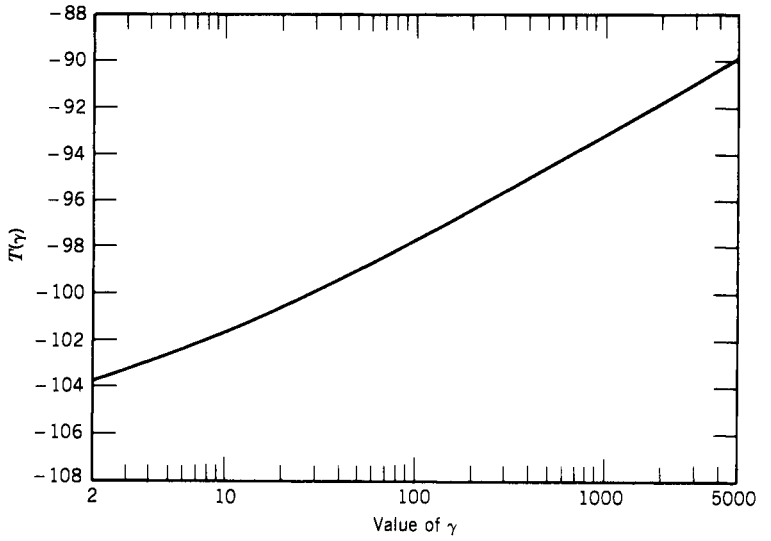


Figure 2.4. $T(\gamma)$ versus γ plot.

Now let us find out how much gain is required in a receiver to make the sensitivity noise limited. This problem can be solved from Eq. (2.17). By substituting Eq. (2.16) into (2.17), the gain required can be written as

$$G_T(\text{dB}) \gtrsim T_d - (-114 + 10 \log F_T + 10 \log\sqrt{2B_R B_V - B_V^2}) \quad \text{dB} \quad (2.21)$$

where T_d is replacing TSS in Eq. (2.16) as the TSS of the video detector. The value in parentheses in Eq. (2.21) represents the noise floor of the receiver. It should be noted that this noise floor includes the effect of video filter bandwidth B_V. Later, the noise floor is often referred to the input of a receiver, which does not include the effect of B_V. Thus Eq. (2.21) indicates that the total gain required to make the receiver noise limited is approximately equal to the gain needed to bring the noise floor of the receiver to the TSS level of the detector. This is an expected result that is very important in receiver designs.

For example, if the noise figure in a receiver is 10 dB with an RF bandwidth B_R of 10 MHz and a video bandwidth B_V of 1 MHz, the noise floor is about -98 dBm ($-114 + 10 + 6$). If the total loss in the RF chain before the detector is 16 dB and the TSS of the detector is -35 dBm, the total minimum gain required is approximately 79 dB ($98 - 35 + 16$).

In a receiver, it is highly desirable to use one bandwidth—the *effective bandwidth*—to calculate the sensitivity. However, from Eqs. (2.12) and (2.13) or (2.18) and (2.19), it is obvious that an effective bandwidth cannot be defined easily. But when the receiver is noise limited and the RF bandwidth B_R is much

greater than the video bandwidth B_V, Eq. (2.19) can be written as

$$\begin{aligned} \text{TSS} &= -114 + 10 \log F_T + 10 \log(2.5\sqrt{2B_R B_V}) \\ &= -114 + 10 \log F_T + 4 + 10 \log\sqrt{2B_R B_V} \quad \text{dBm} \end{aligned} \tag{2.22}$$

The 4 dB in Eq. (2.22) can be regarded as the input signal-to-noise ratio required to produce the TSS of the receiver (ref. 11). The term $\sqrt{2B_R B_V}$ can be regarded as the effective bandwidth of the receiver, which agrees with Klipper's results (ref. 12). The error introduced by neglecting the term $6.31B_V$ and B_V in the square root of Eq. (2.19) is less than 1 dB when γ is greater than 46.

2.4. FALSE ALARM RATE AND PROBABILITY OF DETECTION (2, 25–30)

In order to operate a receiver in a satisfactory manner, a certain threshold must be set up to keep the false alarm rate below a desired level. Although the TSS can be easily measured, it cannot easily provide the information on false alarm and probability of detection. In this section, the sensitivity of the receiver from the viewpoint of false alarm and probability of detection will be discussed.

The following discussion is based on Skolnik's book (ref. 2). The probability of detection as a function of false alarm rate and input signal-to-noise ratio has already been derived; however, the effect of video bandwidth B_V is not included. The effect of B_V is included in reference 25, which will be discussed in the next section, and curves are generated for convenient applications.

The noise voltage V entering an RF filter is assumed Gaussian with a variance ψ_0; the mean value of V is equal to zero. Let the threshold be set at V_T as shown in Figure 2.5. The average time interval between crossings of the threshold by

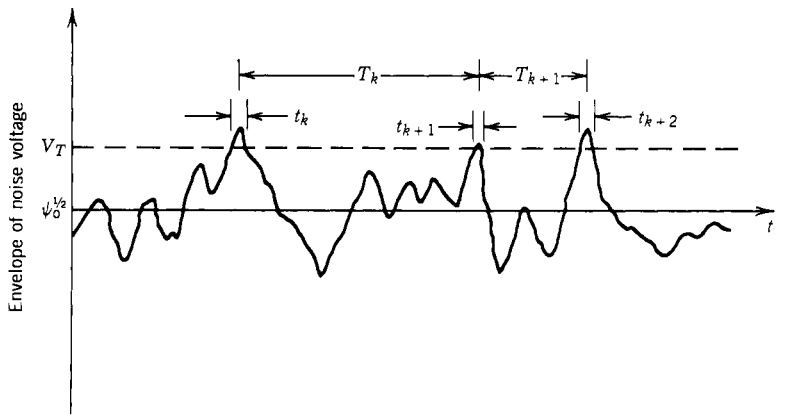


Figure 2.5. Envelope of receiver output illustrating false alarms due to noise.

noise alone is defined as the false alarm time T_{fa} in seconds:

$$T_{fa} = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} T_k \tag{2.23}$$

where T_k is the time between crossings of the threshold voltage V_T by the noise envelope as shown in Figure 2.5. The probability of false alarm P_{fa} may be defined as the duration of time the envelope is actually above the threshold to the total time of observation. It can be expressed as

$$P_{fa} = 1/T_{fa}B_R \tag{2.24}$$

where B_R is the RF bandwidth in hertz.

If Gaussian noise were passed through a narrow-band IF filter, the probability density of the envelope of the noise voltage output is

$$p(E) = \frac{E}{\psi_0} \exp\left(-\frac{E^2}{2\psi_0}\right) \tag{2.25}$$

where E is the amplitude of the envelope of the filter output. The probability that the noise voltage envelope will exceed the threshold voltage V_T is defined as the probability of false alarm

$$\begin{aligned} P_{fa} &= \int_{V_T}^{\infty} p(E)\, dE \\ &= \int_{V_T}^{\infty} \frac{E}{\psi_0} \exp\left(-\frac{E^2}{2\psi_0}\right) dE = \exp\left(-\frac{V_T^2}{2\psi_0}\right) \end{aligned} \tag{2.26}$$

If a sine wave signal of amplitude A is present along the noise in the receiver, the probability density becomes

$$p(E) = \frac{E}{\psi_0} \exp\left(-\frac{E^2 + A^2}{2\psi_0}\right) I_0\left(\frac{EA}{\psi_0}\right) \tag{2.27}$$

where $I_0(z)$ is the modified Bessel function of zero order and augmented z and is defined as

$$I_0(z) = \sum_{n=0}^{\infty} \frac{z^{2n}}{2^{2n} n! n!} \tag{2.28}$$

The probability of detection is

$$P_d = \int_{V_T}^{\infty} p(E)\, dE \tag{2.29}$$

The threshold V_T in Eq. (2.29) is the same value of V_T in Eq. (2.26), but the probability density function is defined in Eq. (2.27) rather than in (2.25).

To illustrate the process of threshold detection, the probability density functions of Eqs. (2.25) and (2.27) are plotted in Figure 2.6. In this figure, an arbitrary threshold V_T is chosen. The shaded area represents the probability of false alarm and the area under the curve $S/N = 1$ and to the right of V_T represents the probability of detection. If the threshold V_T is increased to reduce the probability of the false alarm, the probability of detection will be reduced also.

The probability of detection P_d as a function of the signal-to-noise ratio (S/N) and probability of false alarm are shown in Figure 2.7.

The following steps are taken to obtain Figure 2.7. The value of ψ_0 is setting to 1. Choose a false alarm rate P_{fa}; from Eq. (2.26), a value of V_T is obtained in terms of P_{fa}. This value of V_T is used in Eq. (2.29). The signal amplitude A in Eq. (2.27) and the noise variance ψ_0 can be related to the signal-to-noise ratio as

$$A/\sqrt{\psi_0} = \sqrt{2S/N} \tag{2.30}$$

Assign a value of signal-to-noise ratio. The value of A can be determined from Eq. (2.30). The final step is to carry out the integration in Eq. (2.29) to obtain the P_d value. The actual integration of the equation is carried out through numerical method by changing the variable E to $1/x$ in order to bound the limit of the integration from $1/V_T$ to 0. The value of $I_0(EA/\psi_0)$ is evaluated from Eq. (2.28).

The scale on the P_d axis is generated through the normal distribution

$$\phi(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \tag{2.31}$$

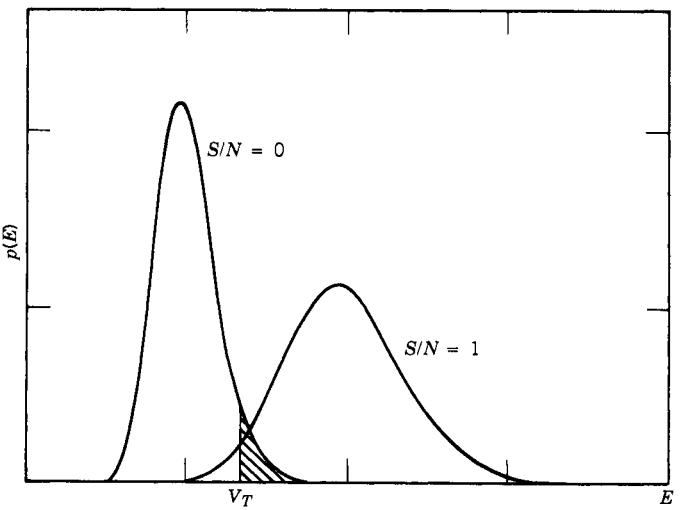


Figure 2.6. Probability density functions for noise alone and for signal plus noise.

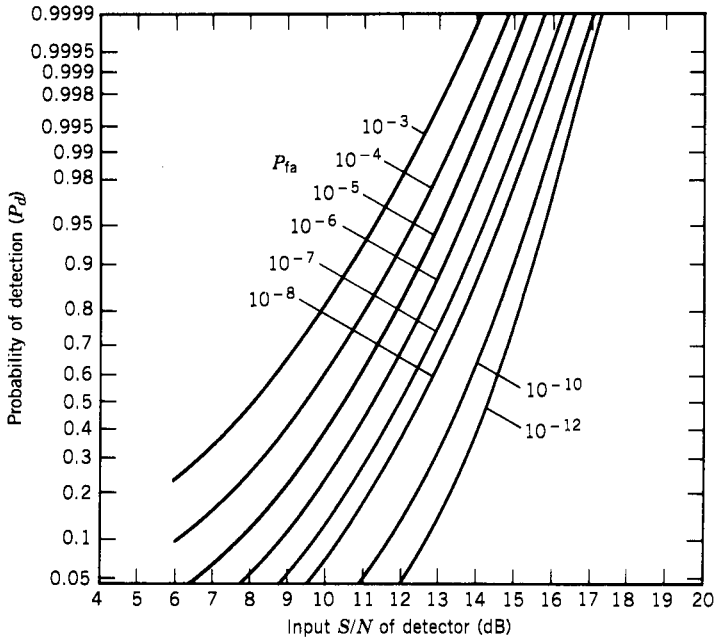


Figure 2.7. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm B_V not included.

The value of $\phi(y)$ can be found in mathematical tables (see ref. 24). It should be noted that the signal-to-noise ratio used in Figure 2.7 is at the input of the detector. It does not represent the noise at the input of a receiver. In order to find the S/N required at the input of a receiver, the noise figure of the receiver must be used to transfer the detector S/N to the receiver input S/N. Although Figure 2.7 provides very valuable information in a receiver design, the effect of the video bandwidth B_V is not included. In the above discussion, the value of B_V is usually greater than $\frac{1}{2}B_R$ but less than B_R.

This result is suitable for most superhet receiver applications because in a superhet receiver the RF and the video circuits usually have approximately the same bandwidth. If many pulses are received and integrated by a receiver's video circuit, the sensitivity of the receiver can be improved. This is a common practice in radar receivers. Curves and equations (refs. 26–30) are available to find the sensitivity of a receiver if the false alarm rate, probability of detection, and number of pulses are given. However, this result is not suitable for EW receiver application because the RF bandwidth is usually much wider than the video bandwidth. In addition, the input pulses are received and encoded one by one, so the advantage of integration by the video circuit does not exist in an EW receiver. For this reason, curves for EW receiver applications will be generated in the next section.

The results obtained in Figure 2.7 are within 1 dB compared to the results presented in reference 2 but closer to the results from reference 28. The difference may be caused by the different approximations used.

2.5. EFFECT OF VIDEO BANDWIDTH ON FALSE ALARM RATE AND PROBABILITY OF DETECTION (24–36)

The effect of video bandwidth B_V is included in the probability density function $p(E)$ derived by Emerson (ref. 25). For $\gamma = B_R/B_V < 4$,

$$p(E) = \frac{K_1}{K_2} \sum_{j=0}^{\infty} \beta_j \psi^j \left(\frac{K_1 E}{K_2} \right) \tag{2.32}$$

where

$$\psi^j(x) = \frac{d^j}{dx^j} [x^{k_4 - 1} \exp(-x)] \tag{2.33}$$

$$\beta_0 = \frac{1}{\Gamma(K_4)} \tag{2.34}$$

$$\beta_1 = \beta_2 = 0 \tag{2.35}$$

$$\beta_3 = \frac{K_4[2 - (K_1 K_3 / K_2^2)]}{6\Gamma(K_4 + 3)} \tag{2.36}$$

Equation (2.33) can be written separately as

$$\psi(x) = x^{K_4 - 1} \exp(-x) \tag{2.37}$$

$$\psi'(x) = [(K_4 - 1)x^{K_4 - 2} - x^{K_4 - 1}] \exp(-x) \tag{2.38}$$

$$\psi''(x) = [(K_4 - 1)(K_4 - 2)x^{K_4 - 3} - 2(K_4 - 1)x^{K_4 - 2} + x^{K_4 - 1}] \exp(-x) \tag{2.39}$$

where

$$x = K_1 E / K_2 \tag{2.40}$$

$$P_{fa} = \int_{V_T}^{\infty} p(E)\, dE \quad \text{with } S/N = 0 \tag{2.41}$$

and

$$P_d = \int_{V_T}^{\infty} p(E)\, dE \quad \text{with } S/N > 0 \tag{2.42}$$

where S/N is included in K_1, K_2, K_3, and K_4. It should be noted that

$$\int_0^\infty p(E)\,dE = 1 \tag{2.43}$$

as expected from the definition of probability density, and

$$\begin{aligned}\int_{V_T}^\infty p(E)\,dE &= \frac{K_1}{K_2}\int_{K_1V_T/K_2}^\infty p(x)\,dx = \int_{K_1V_T/K_2}^\infty [\beta_0\psi(x) + \beta_3\psi'''(x)]\,dx \\ &= \frac{1}{\Gamma(K_4)}\int_{K_1V_T/K_2}^\infty x^{K_4-1}\exp(-x)\,dx \\ &\quad + \frac{K_4[2 - (K_1K_3/K_2^2)]}{6\Gamma(K_4+3)}\psi''(x)\Big|_{K_1V_T/K_2}^\infty\end{aligned} \tag{2.44}$$

where $\Gamma(z)$ is the gamma function defined as

$$\Gamma(z) = \int_0^\infty x^{z-1}\exp(-x)\,dx \tag{2.45}$$

and

$$K_1 = 1 + \frac{S}{N} \tag{2.46}$$

$$K_2 = \frac{1}{(1+\gamma^2/2)^{1/2}}\left[1 + \frac{S}{N}\left(\frac{1+\gamma^2/2}{1+\gamma^2/4}\right)\right] \tag{2.47}$$

$$K_3 = \frac{4}{2+3\gamma^2/4}\left[1 + 3\frac{S}{N}\left(\frac{2+3\gamma^2/4}{2+\gamma^2/4}\right)^{1/2}\right] \tag{2.48}$$

$$K_4 = \frac{K_1^2}{K_2} \tag{2.49}$$

By substituting K_1, K_2, K_3, and k_4 into Eqs. (2.42) and (2.43), the values of P_{fa} and P_d can be obtained. To improve the accuracy of the integration, when $P_{fa} > 0.5$, the following relation is adapted:

$$\int_{V_T}^\infty p(E)\,dE = 1 - \int_0^{V_T} p(E)\,dE \tag{2.50}$$

The integrals are calculated through a numerical method. Results of $\gamma = 1$ are shown in Figure 2.8.

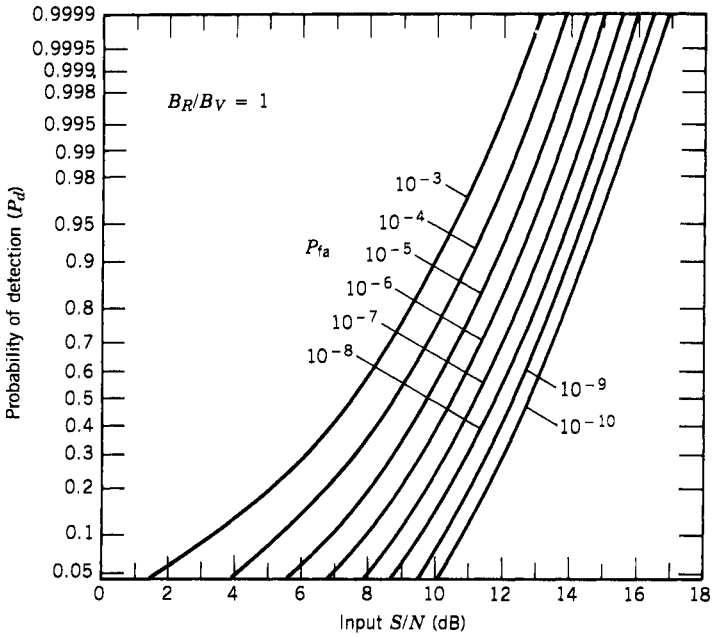


Figure 2.8. Probability of detection for a sine wave in noise as a function of the signal-to-noise power ratio and the probability false alarm for $\gamma = 1$.

The case with $\gamma > 4$ is discussed as follows and these are the most popular cases in EW receivers: For $\gamma = B_R/B_V > 4$,

$$p(E) = \frac{1}{\sqrt{K_2}} \sum_{j=0}^{\infty} \alpha_j \phi^j \left(\frac{E - K_1}{\sqrt{K_2}} \right) \tag{2.51}$$

$$\phi^j(x) = \frac{d^j}{dx^j} \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2} \right) \right] \tag{2.52}$$

where

$$\alpha_0 = 1 \tag{2.53}$$

$$\alpha_1 = \alpha_2 = 0 \tag{2.54}$$

$$\alpha_3 = -K_3/6K_2^{3/2} \tag{2.55}$$

$$x = (E - K_1)/\sqrt{K_2} \tag{2.56}$$

and K_1, K_2, and K_3 are given in Eqs. (2.46)–(2.48). Therefore

$$p(x) = \frac{1}{\sqrt{2\pi K_2}} \exp\left(-\frac{x^2}{2}\right) - \frac{K_3}{6K_2^{3/2}} \frac{d}{dx} \phi''(x) \tag{2.57}$$

where

$$\phi''(x) = \frac{1}{\sqrt{2\pi}} (x^2 - 1)\exp\left(-\frac{x^2}{2}\right) \tag{2.58}$$

Before calculating the P_d and P_{fa} versus the input S/N curves, the effects of γ on the probability density function $p(x)$ will be illustrated. The effects of γ can be shown in Figures 2.9 and 2.10. In Figure 2.9, the probability density function with $S/N = 0$ is plotted against normalized output voltage Y for different γ values. Figure 2.10 shows a similar plot with $S/N = 2$. From these curves, the effect of γ can be readily seen. The higher the γ value, the narrower the probability density function. The video filter reduces the noise resulting from the wide-RF bandwidth.

The mathematical approach to solve this problem is similar to the procedures discussed above:

$$P_{fa} = \frac{\int_{V_T}^{\infty} p(E)\, dE}{\int_0^{\infty} p(E)\, dE} \quad \text{with } \frac{S}{N} = 0 \tag{2.59}$$

$$P_d = \frac{\int_{V_T}^{\infty} p(E)\, dE}{\int_0^{\infty} p(E)\, dE} \quad \text{with } \frac{S}{N} > 0 \tag{2.60}$$

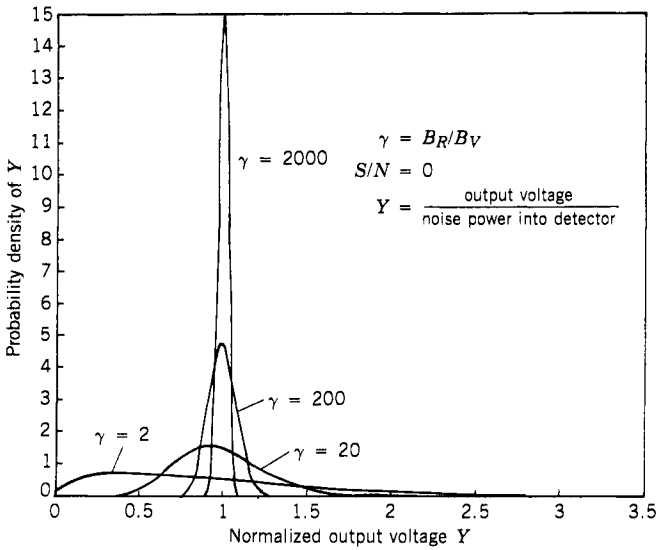


Figure 2.9. Output probability densities for the Gaussian system (signal absent).

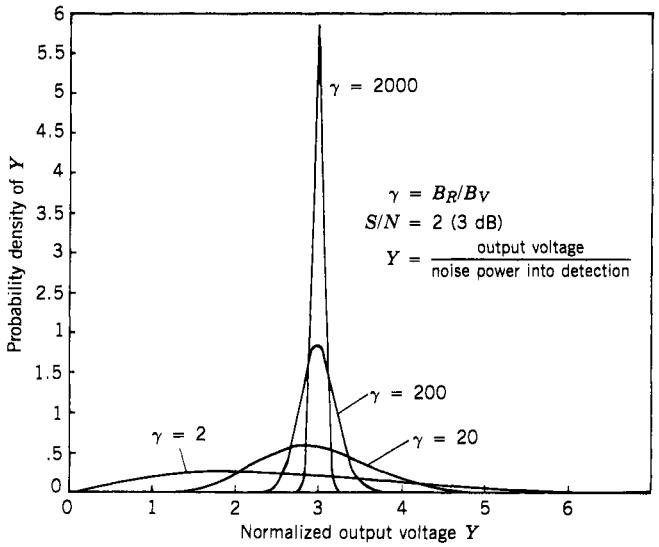


Figure 2.10. Output probability densities for the Gaussian system $S/N = 2$).

The integrals are calculated as

$$A \equiv \int_0^\infty p(E)\,dE = \sqrt{K_2}\int_{-\sqrt{K_4}}^\infty p(x)\,dx$$

$$= \frac{1}{\sqrt{2\pi}}\frac{K_3}{6K_2^{3/2}}(K_4 - 1)\exp\left(-\frac{K_4}{2}\right) + \frac{1}{\sqrt{2\pi}}\int_{-\sqrt{K_4}}^\infty \exp\left(-\frac{x^2}{2}\right)dx \tag{2.61}$$

$$B = \int_{V_T}^\infty p(E)\,dE = \sqrt{K_2}\int_{K_5}^\infty p(x)\,dx$$

$$= \frac{1}{\sqrt{2\pi}}\frac{K_3}{6K_2^{3/2}}(K_5^2 - 1)\exp\left(-\frac{K_5^2}{2}\right) + \frac{1}{\sqrt{2\pi}}\int_{K_5}^\infty \exp\left(-\frac{x^2}{2}\right)dx \tag{2.62}$$

where

$$K_5 = (V_T - K_1)/\sqrt{K_2} \tag{2.63}$$

The result of A should be unity by the definition of the probability density function. However, only limited terms are used in Eq. (2.51), and the result of A is not unity but very close to it.

The integral in Eqs. (2.61) and (2.62) are closely related to the normal probability of integral

$$\operatorname{erf}\left(\frac{y}{\sqrt{2}}\right) = \frac{1}{\sqrt{2\pi}} \int_{-y}^{y} \exp\left(-\frac{x^2}{2}\right) dx \tag{2.64}$$

where erf is the error function. The two integrals in Eqs. (2.61) and (2.62) are evaluated from the relations

$$\begin{aligned}\frac{1}{\sqrt{2\pi}} \int_{y}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx &= \frac{1}{\sqrt{2\pi}} \left[\int_{0}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx - \int_{0}^{y} \exp\left(-\frac{x^2}{2}\right) dx\right] \\ &= \frac{1}{2} - \frac{1}{2\sqrt{2\pi}} \int_{-y}^{y} \exp\left(-\frac{x^2}{2}\right) dx = \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{y}{\sqrt{2}}\right)\end{aligned} \tag{2.65}$$

The $\operatorname{erf}(y/\sqrt{2})$ is calculated through approximation (ref. 24).

To have better accuracy for large values of P_d, let us define

$$\begin{aligned}C \equiv \int_{0}^{V_T} p(E)\, dE &= \sqrt{K_2} \int_{-\sqrt{K_4}}^{K_5} p(x)\, dx \\ &= -\frac{1}{\sqrt{2\pi}} \frac{K_3}{6K_2^{3/2}} (K_5^2 - 1)\exp\left(-\frac{K_5^2}{2}\right) + \frac{1}{\sqrt{2\pi}} \frac{K_3}{6K_2^{3/2}} (K_4 - 1)\exp\left(-\frac{K_4}{2}\right) \\ &\quad + \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{K_4}}^{K_5} \exp\left(-\frac{x^2}{2}\right) dx\end{aligned} \tag{2.66}$$

The integral in Eq. (2.66) is calculated using the numerical method.

$$\begin{aligned}&\text{For } P_d < 0.5, \quad P_d = B/A \\ &\text{For } P_d > 0.5, \quad P_d = (A - C)/A\end{aligned} \tag{2.67}$$

The results of $\gamma = 1$–3000 are shown in Figures 2.12–2.35. Figure 2.11 is the same as Figure 2.7, where the effect of B_V is not included. Figure 2.12 is the same as Figure 2.8.

Different integration methods were used to calculate these curves, and their results are well within 1 dB. From these curves, one realizes that the higher the γ value, the less the S/N value required to achieve a certain probability of detection. This effect implies that the video bandwidth is capable of smoothing the noise from the wide-RF bandwidth. The negative values of S/N (in decibels) in the curves mean that the signal is below the noise level. However, the signal plus noise is still above the noise level.

The equations used to generate the curves in Figures 2.11 to 2.35 do not include the characteristics of the detector diode. Therefore the receiver should be

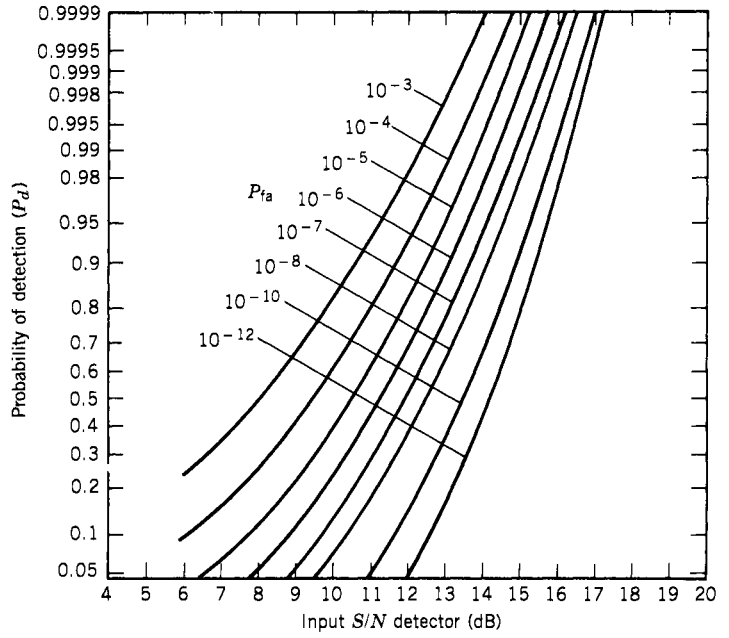


Figure 2.11. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm B_V not included.

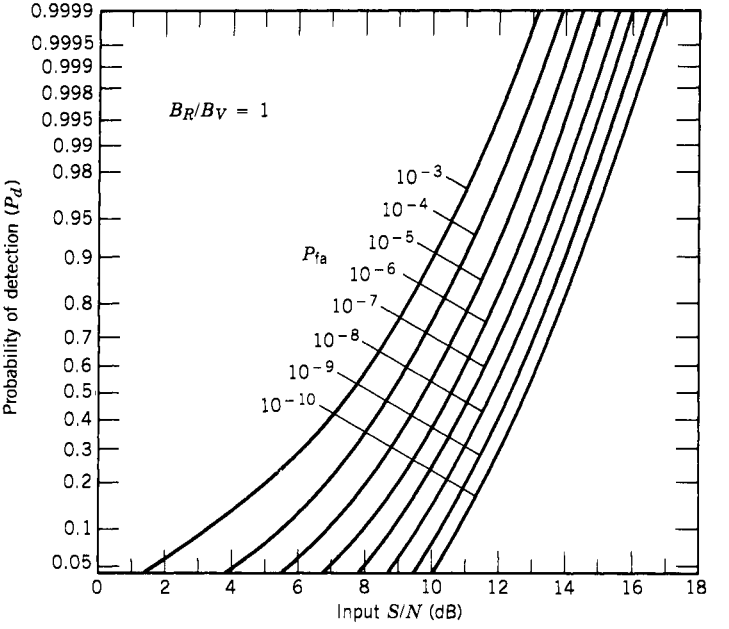


Figure 2.12. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 1$.

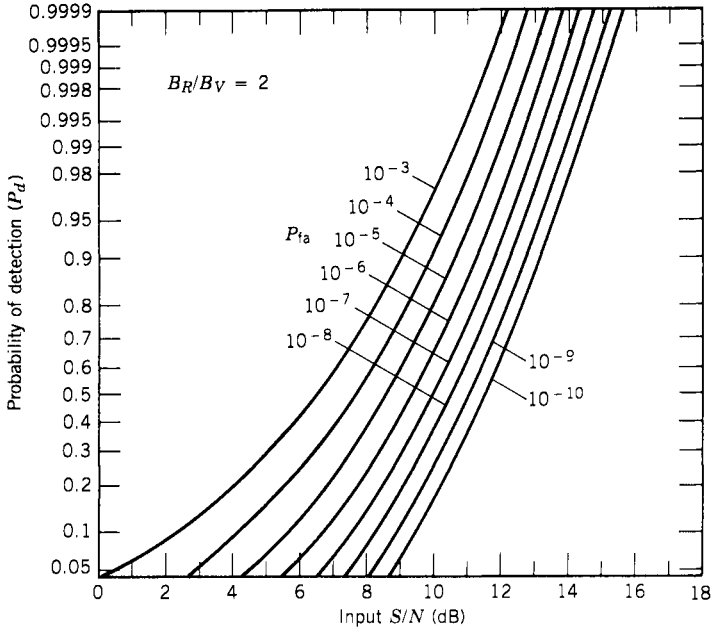


Figure 2.13. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 2$.

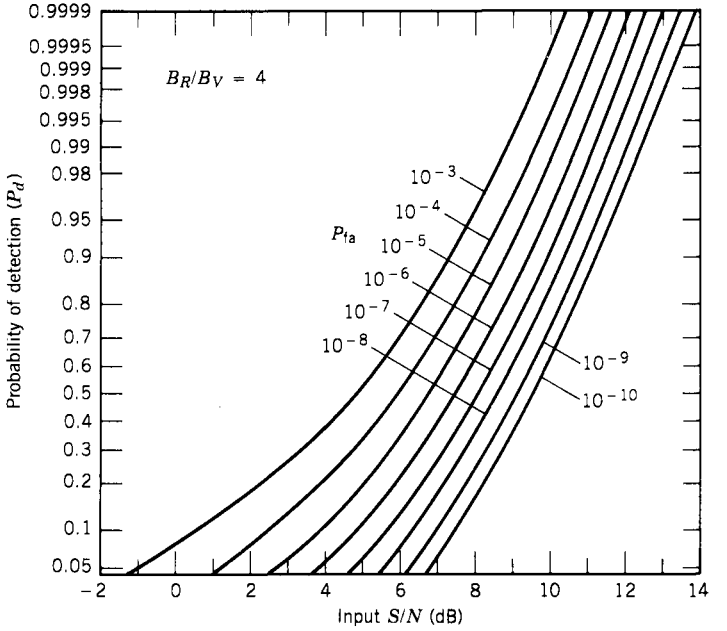


Figure 2.14. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 4$.

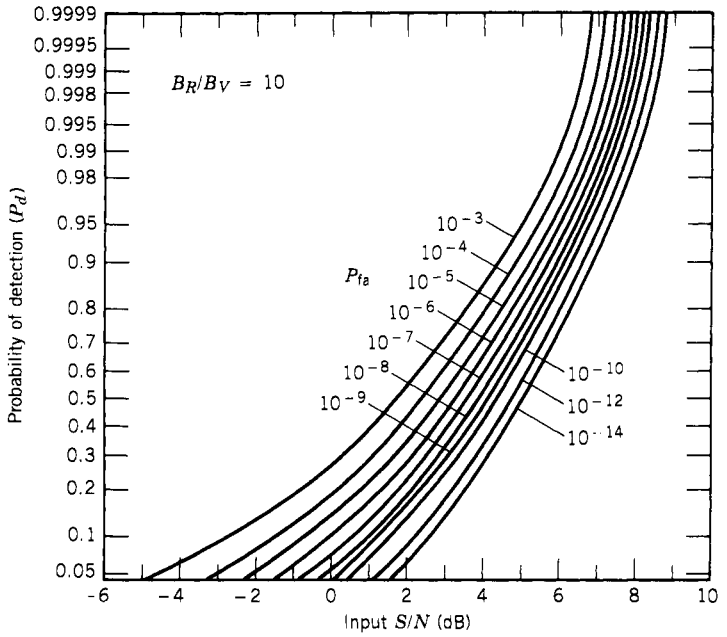


Figure 2.15. Probability of detection for a sin wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 10$.

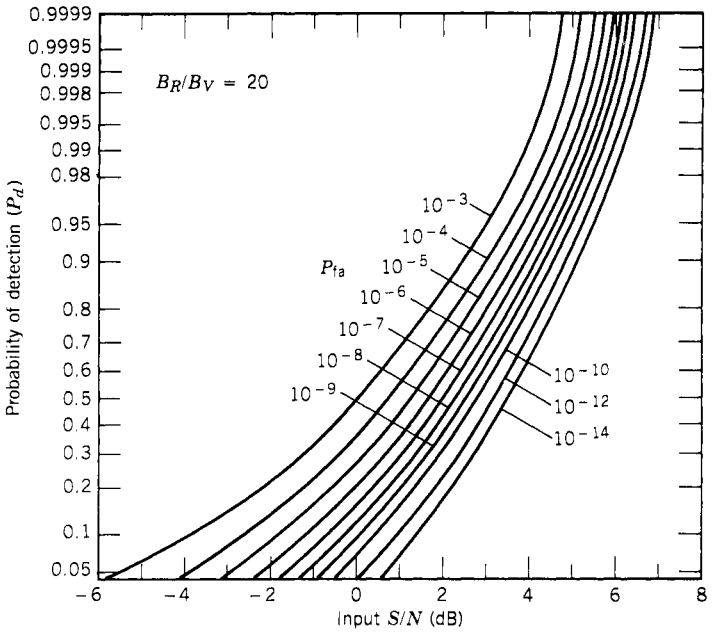


Figure 2.16. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 20$.

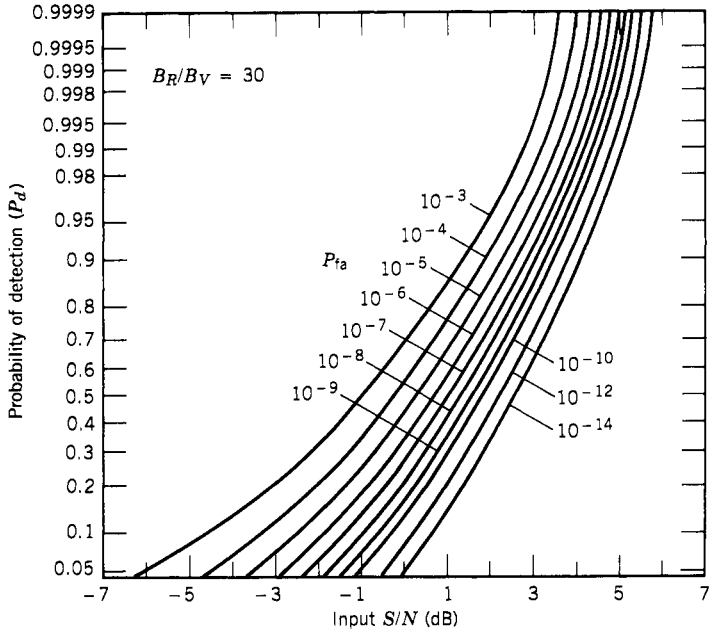


Figure 2.17. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 30$.

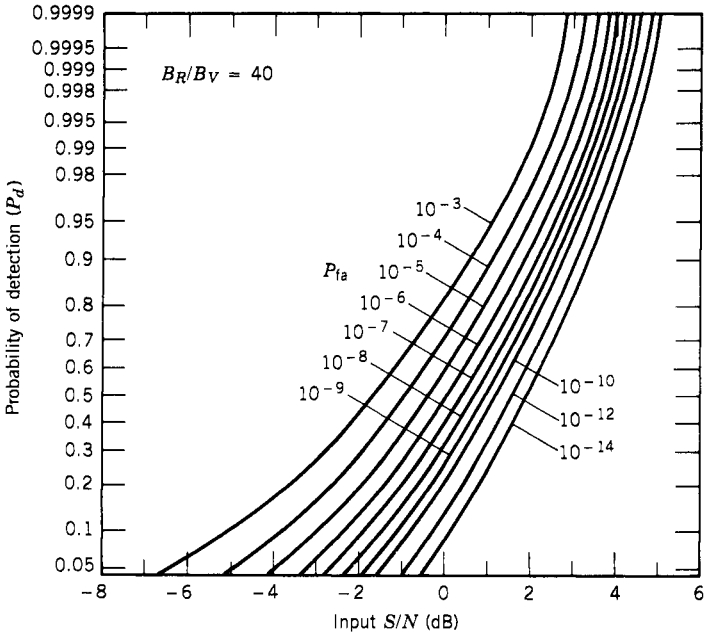


Figure 2.18. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 40$.

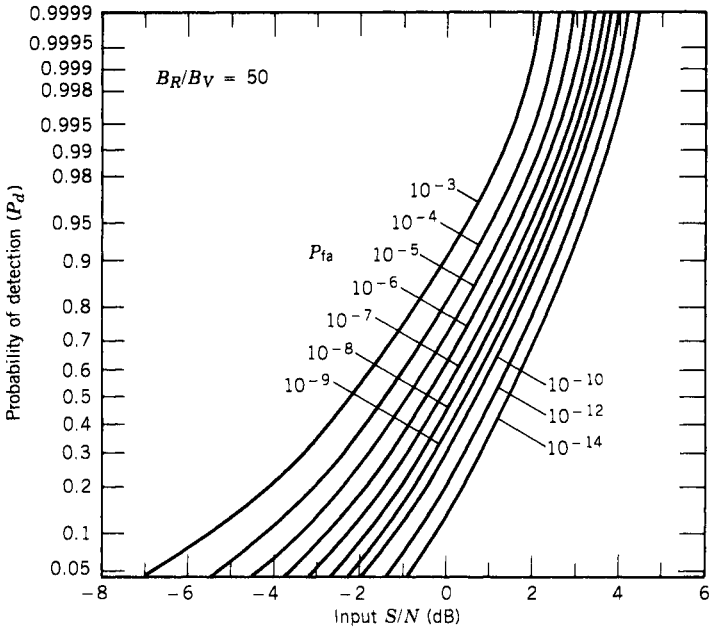


Figure 2.19. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power)ratio and the probability of false alarm with $\gamma = 50$.

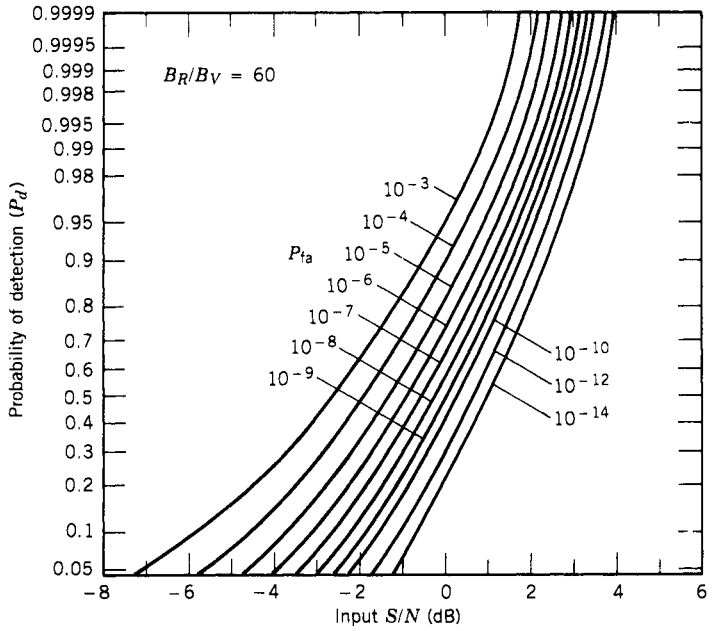


Figure 2.20. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 60$.

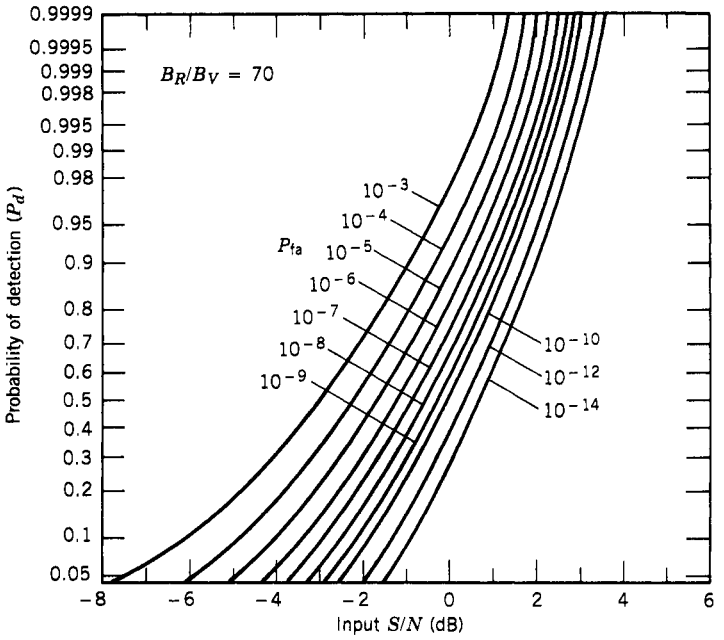


Figure 2.21. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 70$.

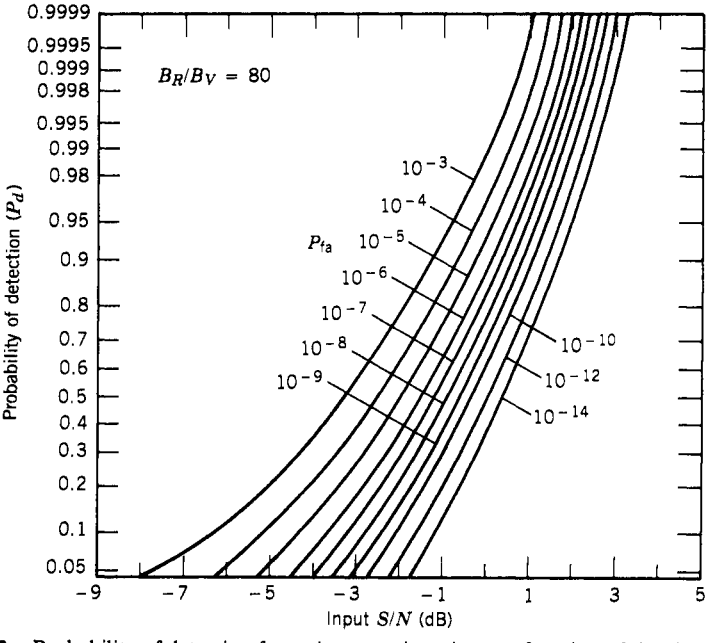


Figure 2.22. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 80$.

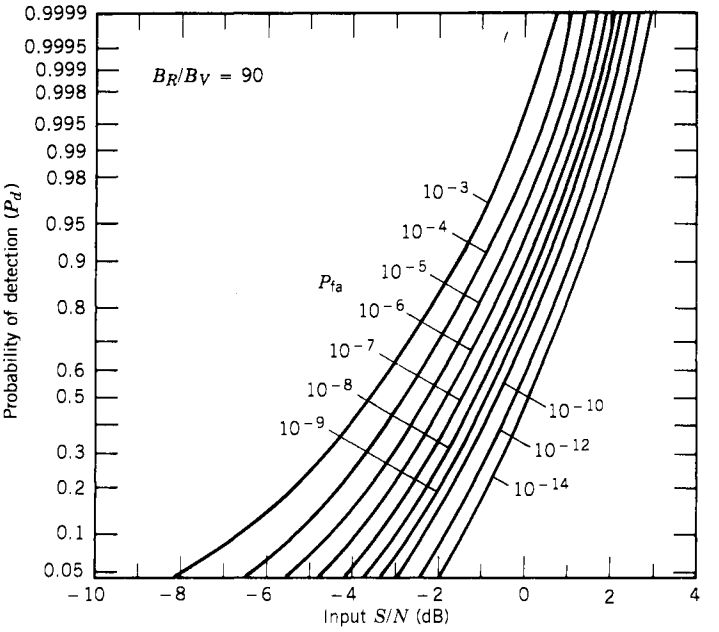


Figure 2.23. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 90$.

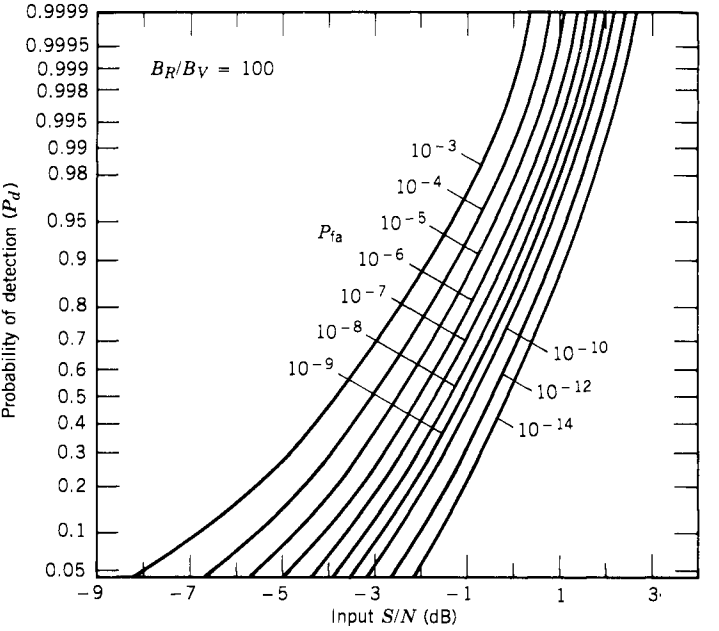


Figure 2.24. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 100$.

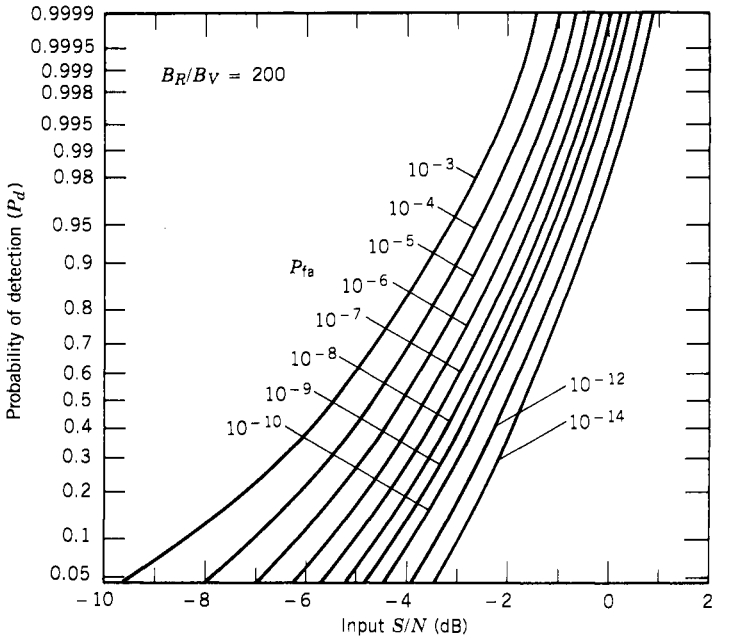


Figure 2.25. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 200$.

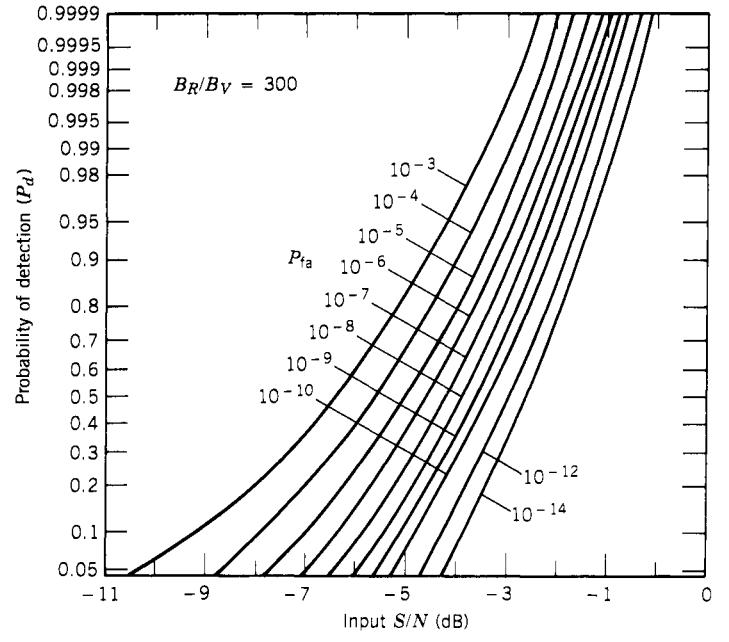


Figure 2.26. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 300$.

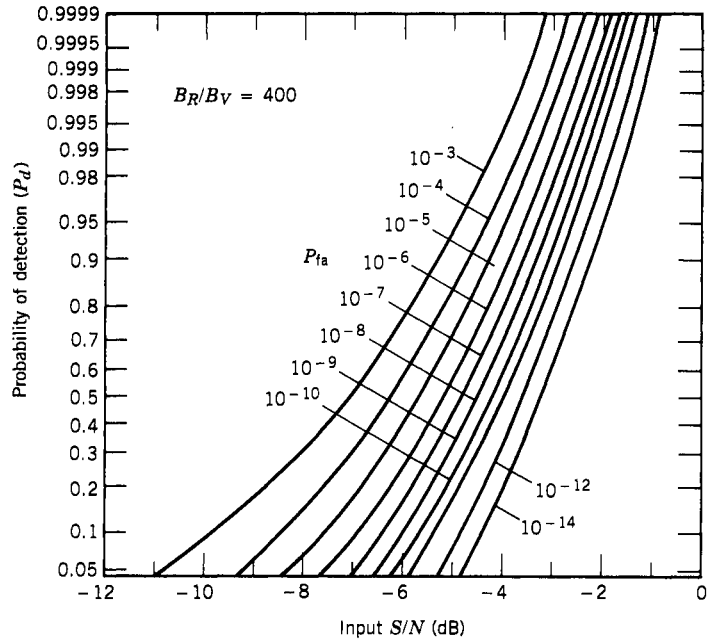


Figure 2.27. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 400$.

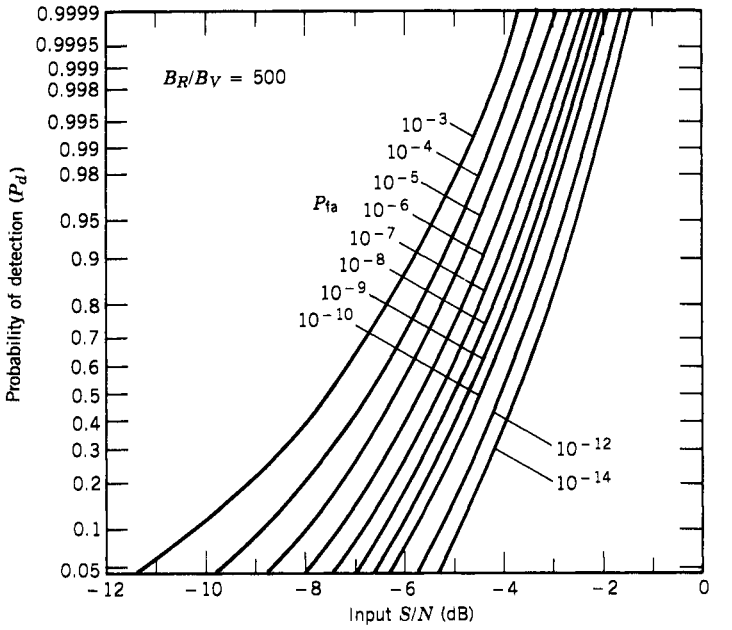


Figure 2.28. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 500$.

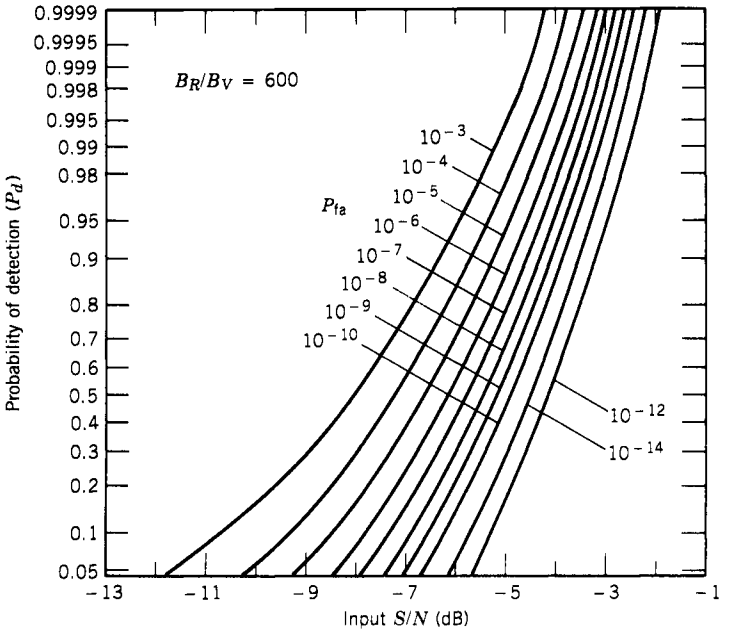


Figure 2.29. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 600$.

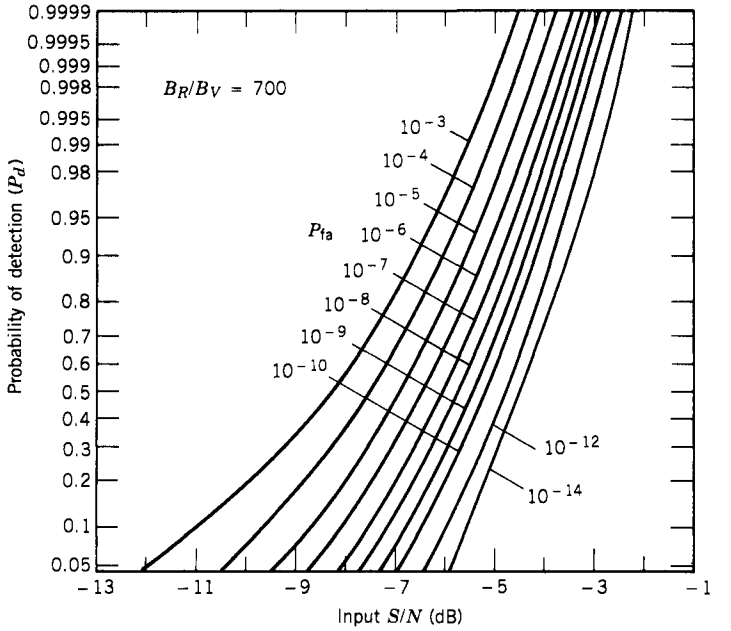


Figure 2.30. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 700$.

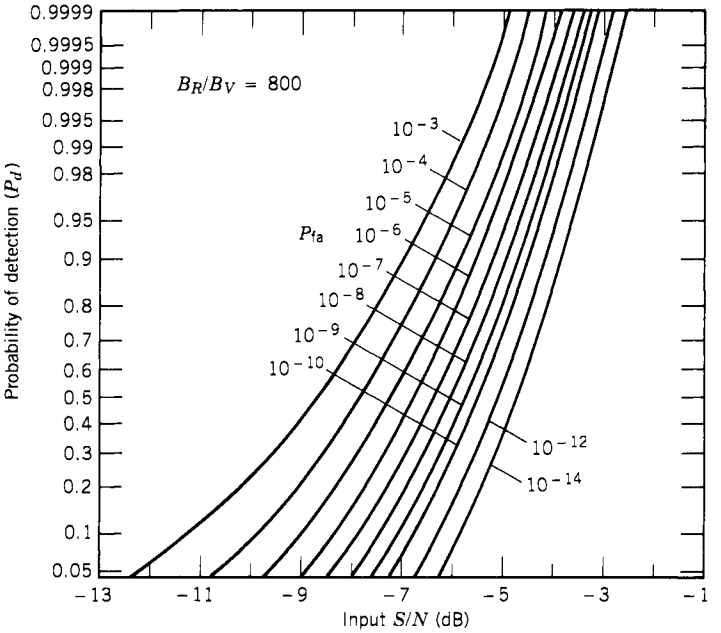


Figure 2.31. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 800$.

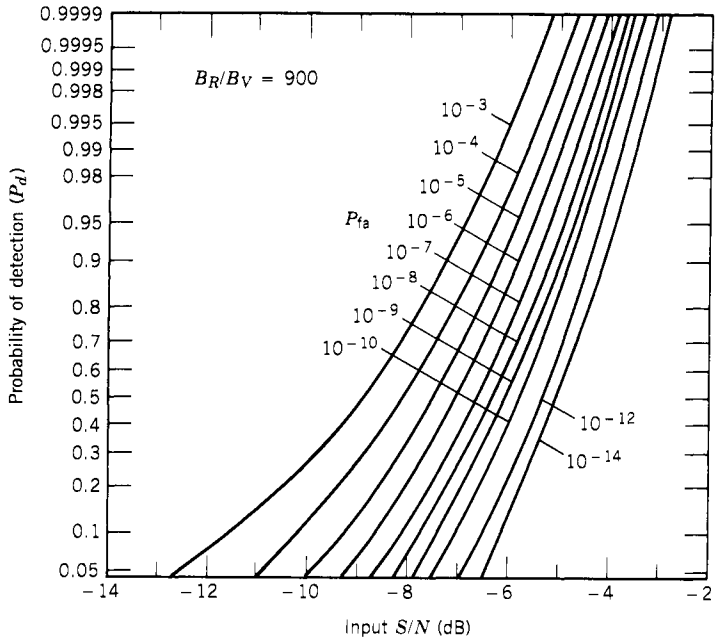


Figure 2.32. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 900$.

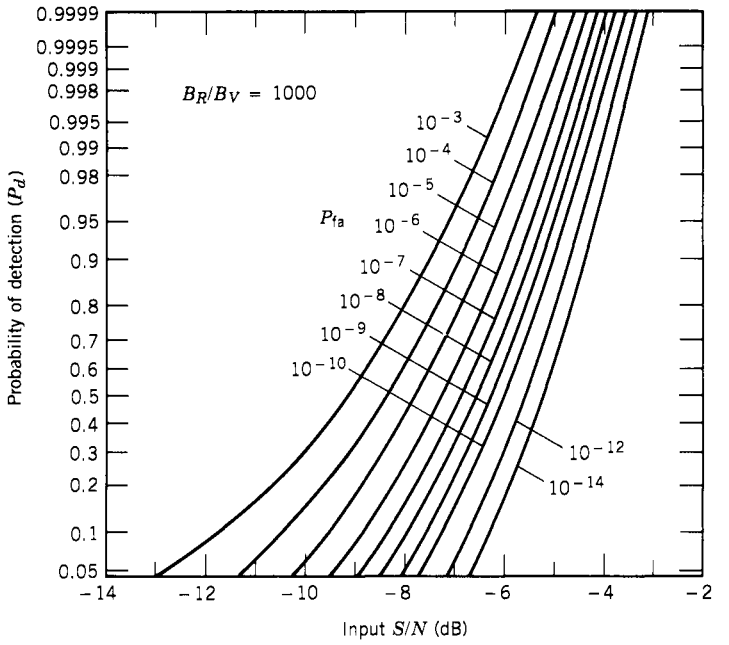


Figure 2.33. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 1000$.

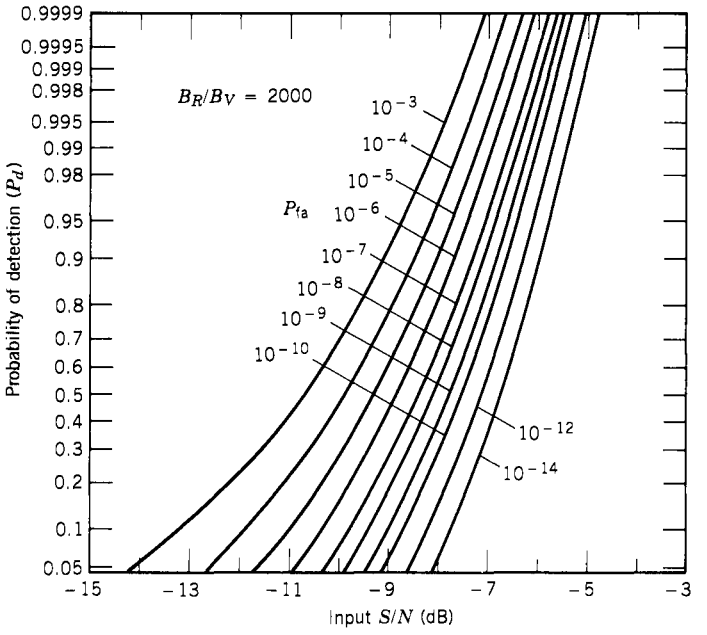


Figure 2.34. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 2000$.

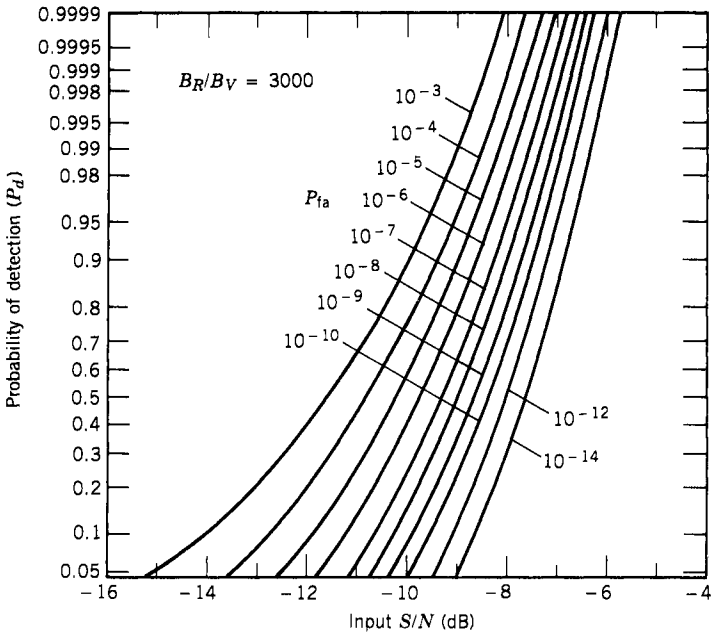


Figure 2.35. Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm with $\gamma = 3000$.

in the noise-limited case. In other words, the RF gain in front of the receiver is high enough that the effect of the detector can be neglected. The use of these curves will be demonstrated in the next section.

2.6. EXAMPLE OF RECEIVER SENSITIVITY CALCULATION

An example will be presented here to demonstrate the use of the equations and figures presented in this chapter. In this example, an instantaneous frequency measurement (IFM) receiver with the following specifications is used:

$$B_R = 2000 \quad \text{MHz} \quad (2000\text{–}4000\ \text{MHz})$$
$$B_V = 1 \quad \text{MHz}$$
$$G_T = 51 \quad \text{dB}$$
$$F_T = 10 \quad \text{dB}$$

The value of $\gamma = B_R/B_V$ is 2000. The noise floor is at -71 dBm ($-114 + 10 \log B_R + F_T = -114 + 33 + 10 = -71$). The gain G_T will amplify the noise level to -20 dBm, which is high enough to neglect the effect of the detector.

The TSS can be obtained from Figure 2.4 and Eq. (2.20):

$$T(\gamma) = -91.8$$

$$\text{and } \log B_V = 0. \text{ Since } B_V = 1,$$

$$\text{TSS} = 10 - 91.8 = -81.8 \quad \text{dBm}$$

Next let us determine the sensitivity of the receiver from a given probability of detection and false alarm rate. This sensitivity is sometimes referred to as the operational sensitivity. If one false alarm in 100 sec is the maximum acceptable rate and a probability of detection P_d of 99% is required, from Eq. (2.24),

$$P_{fa} = \frac{1}{100 \times 2000 \times 10^6} = 5 \times 10^{-12}$$

The S/N required to generate a P_d of 0.99 is about -5.6 dB (from Figure 2.34). Therefore the corresponding signal strength at the input of the receiver is

$$-114 + 10 \log B_R + 10 + (-5.6) = -76.6 \text{ dBm}$$

which is about 5.2 dB above the TSS.

2.7. PROBLEMS OF FALSE ALARM RATE MEASUREMENTS

Measurements of receivers will be discussed in Chapter 12. However, since the false alarm measurement is closely related to the results discussed in Sections 2.4 and 2.5, it will be included in this chapter.

To determine the operational sensitivity of a receiver, the probability of detection and the false alarm rate must be measured. It is relatively easy to measure the probability of detection, but under normal operating conditions, the false alarm of a receiver is often too low to be measured accurately.

The probability of detection of a microwave receiver is normally measured by applying pulsed signals of low amplitude and known pulse repetition frequency (PRF) to the input of the receiver. The data-ready pulse (threshold broken) at the receiver output is monitored. By increasing the input signal power level, the frequency f of the data-ready pulse will approach that of the input signal PRF. The ratio of f/PRF is the probability of detection at that specific input power level.

Theoretically, the false alarm can be measured in a similar manner. With the input of the receiver properly terminated, the data-ready events at the output of the receiver are counted. The total number of false alarms counted divided by the associated measurement time interval is the false alarm during that

period. However, under normal operating conditions, the false alarm rate is relatively low, from one false alarm per many minutes to one false alarm per several hours. The low false alarm rate creates measurement difficulty. During a long measurement interval, noise generated in the laboratory and disturbance in the power supply can cause very high error in the measurement. In addition, it is very undesirable to measure a receiver for a long period of time because it is time consuming.

For example, if one measures two false alarms in 2 hr, it would be inappropriate to claim that the false alarm rate is 1/hr because the observation time is too short to provide a dependable statistical result. On the other hand, if zero false alarm is measured in 10 min, the false alarm rate is still unknown. It is desirable to measure the false alarm at a higher rate such as a few to tens of false alarms per seconds or minutes.

One way of increasing the false alarm rate is to lower the threshold at the output of the video detector. This adjustment will change the operating condition of the receiver, and it is very undesirable because the false alarm rate cannot be obtained at the operating condition of the receiver. Another approach is to connect a noise generator with a known noise level at the input of the receiver to increase the false alarm rate. This method will improve the false alarm rate measurement accuracy, but the false alarm rate at the normal operating conditions must be calculated when the noise source is replaced by a matched load. This latter approach will be discussed in Sections 2.8 and 2.9.

2.8. EXPERIMENTAL SETUP FOR FALSE ALARM MEASUREMENTS (37)

The experimental setup is shown in Figure 2.36. A calibrated noise source is applied to the input of the receiver through a variable attenuator A, which is used to control the amount of noise input to the receiver. A frequency counter or an event counter is used at the output of the receiver to count the data-ready pulses.

If a matched load is connected to the input of the receiver under test, since $kTBG$ is the input noise amplified by the gain of the amplifier, the noise added by the receiver N_r is

$$N_r = N_o - kTBG = (F - 1)kTBG \tag{2.68}$$

where N_o is the total noise at the output of the receiver and F is the noise figure of the receiver.

A noise generator can be specified in decibels or in noise temperature, and these two quantities are related by

$$10 \log \frac{T_1 - T}{T} = M \quad \text{dB} \tag{2.69}$$

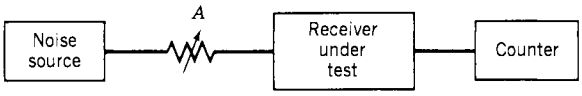


Figure 2.36. Basic experimental setup for false alarm measurement.

where T_1 is the equivalent noise temperature of the noise generator, T is room temperature, and M is the excess noise ratio of the noise generator expressed in decibels.

If N_2 represents the total noise at the output of the receiver with the noise generator at the input of the receiver and N_1 represents the noise output with a matching load at the input, the ratio N_2/N_1 can be expressed as

$$R = \frac{N_2}{N_1} = \frac{\begin{matrix}\text{(noise from input} \\ \text{termination)} \times G +\end{matrix} \quad \begin{matrix}\text{(receiver} \\ \text{noise)} +\end{matrix} \quad \begin{matrix}\text{(noise from noise} \\ \text{generator)} \times G\end{matrix}}{\text{(noise from input termination)} \times G + \text{(receiver noise)}}$$

$$= \frac{kTB_R G + (F-1)kTB_R G + k(T_1 - T)B_R G}{kTB_R G + (F-1)kTB_R G}$$

$$= 1 + \frac{T_1 - T}{FT} = 1 + \frac{M}{F} \tag{2.70}$$

where F expressed in power ratio is the noise figure of the experimental setup in Figure 2.36, which includes the attenuator A and the noise figure of the receiver, and M is the noise source expressed in power ratio. If the noise figure of the receiver is F_T, then

$$F\,(\text{dB}) = F_T\,(\text{dB}) + A\,(\text{dB}) \tag{2.71}$$

From Eq. (2.70), it is obvious that the noise figure of the receiver and the attenuator A determine the amount of noise introduced into the receiver.

2.9. THEORETICAL CALCULATIONS OF FALSE ALARM RATE (25)

In this discussion, only the cases of $\gamma(B_R/B_V)$ greater than 4 will be discussed because the results do not change rapidly with γ values; the false alarm rate is a statistical measurement, and a very accurate measurement result will not be expected. Therefore, results obtained from curves with γ close to the desired value should provide sufficient accuracy.

To calculate the false alarm rate, the probability density function in Eq. (2.51) can be used. The K_2 and K_3 values in Eqs. (2.47) and (2.48) can be modified

by setting $S/N = 0$. Thus,

$$K_2 = \frac{N^2}{(1 + \gamma^2/2)^{1/2}} \tag{2.72}$$

$$K_3 = \frac{4N^3}{2 + 3\gamma^2/4} \tag{2.73}$$

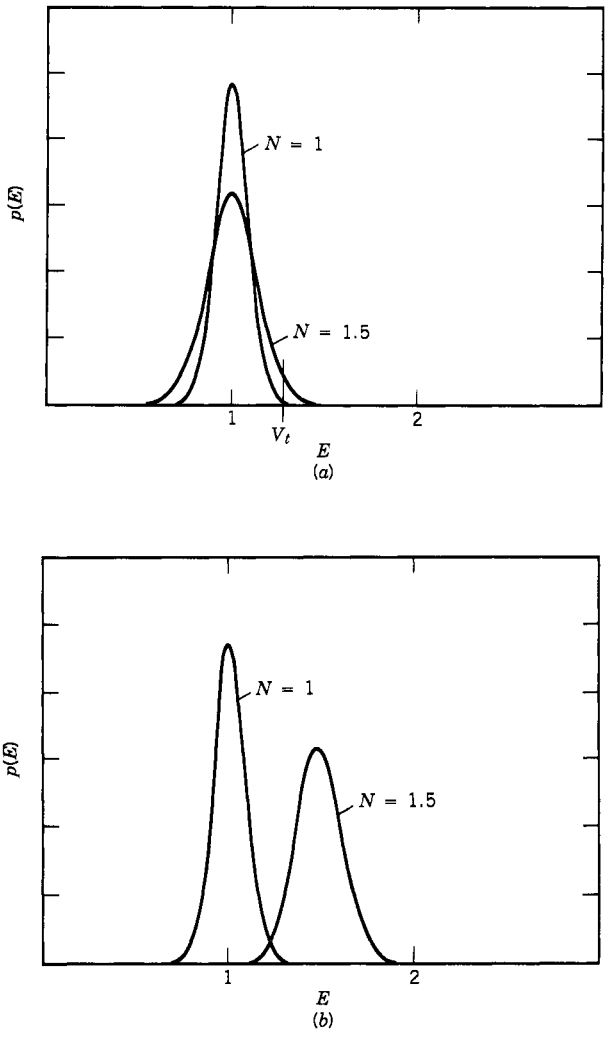


Figure 2.37. $p(E)$ versus E for $\gamma = 200$ and different N: (*a*) ac-coupled case: (*b*) dc-coupled case.

The value of K_1 in Eq. (2.46) is receiver dependent.

$$K_1 = 1 \quad \text{for ac-coupled video output}$$
$$K_1 = N \quad \text{for dc-coupled video output} \tag{2.74}$$

The plots of $p(E)$ versus E for $\gamma = 200$ and different N values are shown in Figure 2.37. Figure 2.37*a* shows the ac-coupled case where V_T is the threshold

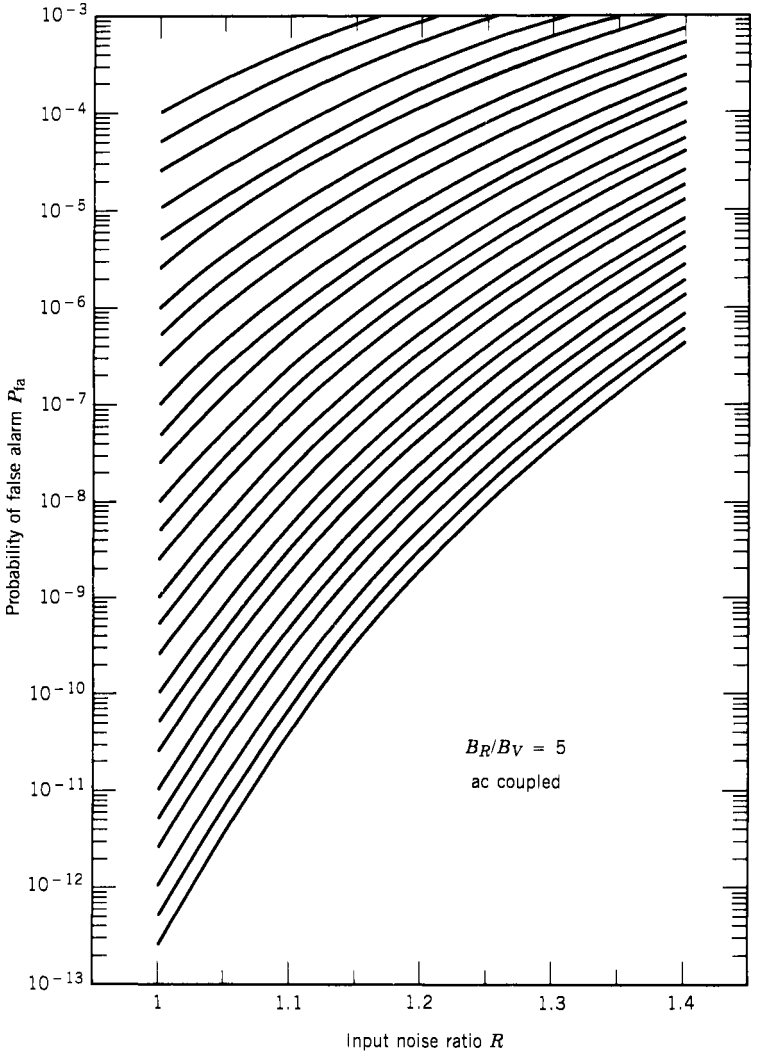


Figure 2.38. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 5$, ac coupled.

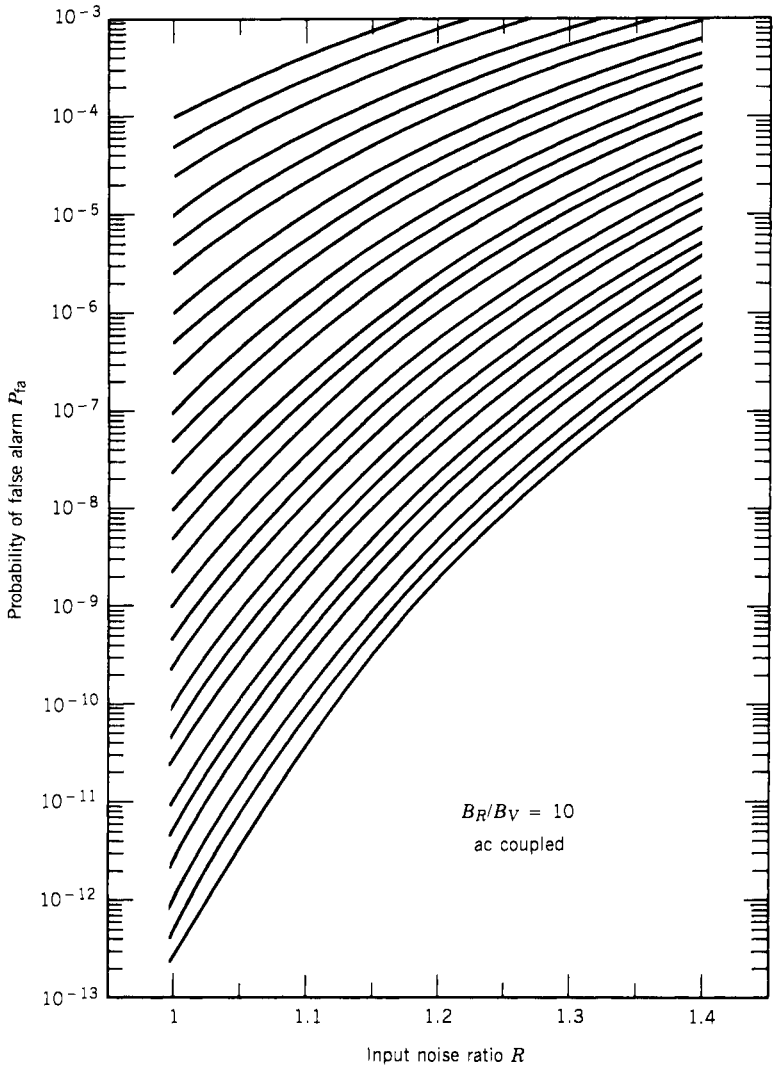


Figure 2.39. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 10$, ac coupled.

used to determine the false alarm rate. Figure 2.37*b* shows the dc-coupled case. As expected, the higher the N value, the wider the probability function spreads. For the dc-coupled case, in addition to the spread of $p(E)$, the dc level of the detector output is also shifted.

The next step is to find the false alarm rate defined by Eq. (2.59). The mathematics used to derive the results are quite similar to the calculations described in

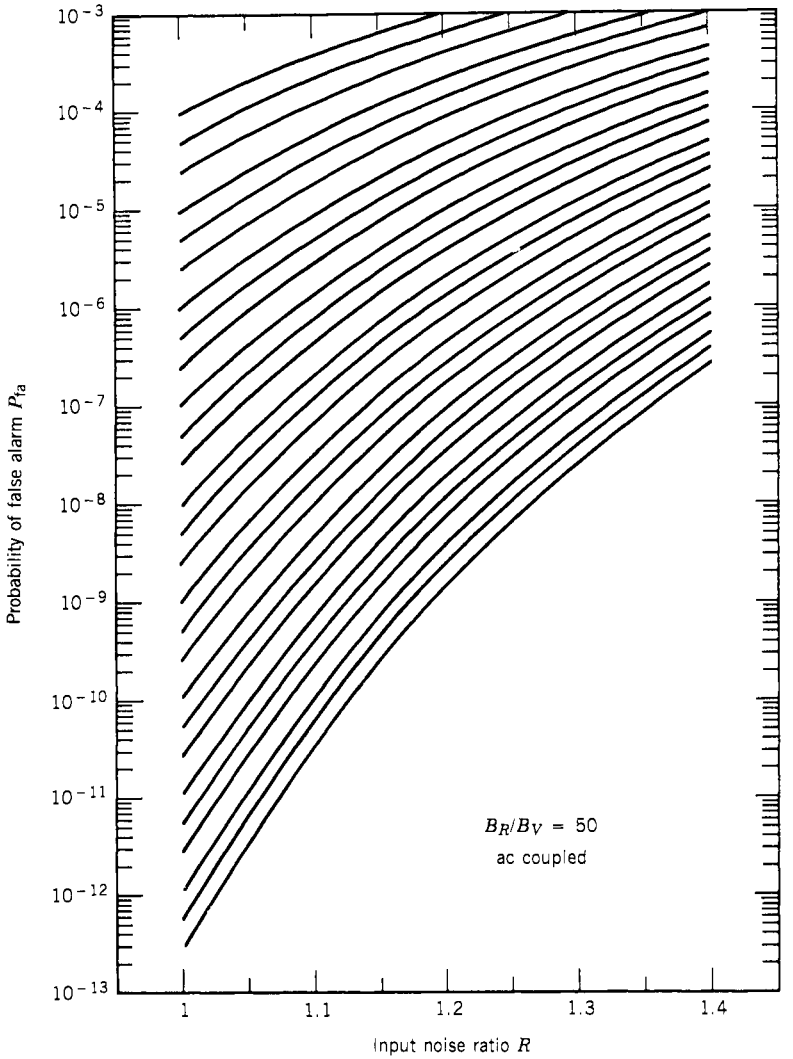


Figure 2.40. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 50$, ac coupled.

Section 2.5. The results are shown in Figures 2.38–2.50 for different γ values. Figures 2.38–2.44 show the results for ac-coupled cases, whereas Figures 2.45–2.50 show the results for dc-coupled cases. Since the results do not depend on the characteristics of the video detector, the receiver sensitivity is in the noise-limited condition.

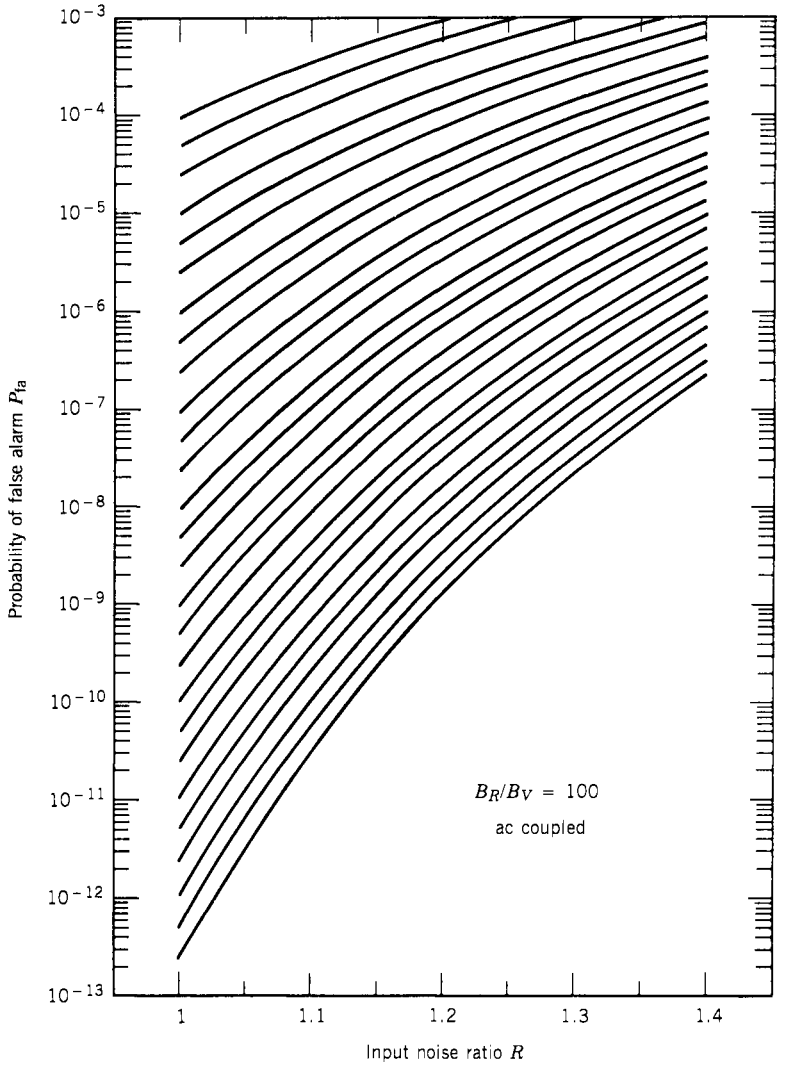


Figure 2.41. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 100$, ac coupled.

2.10. EXAMPLE OF FALSE ALARM RATE MEASUREMENT

An example is presented to demonstrate the use of the curves obtained from the last section. If a noise source of 15 dB is used, as shown in Figure 2.51 on page 59, the overall RF gain before the detector is approximately 62.6 dB (30.6 − 3 + 36 − 1), which will provide enough gain to make the receiver sensitivity noise limited. The noise figure of the receiver is approximately

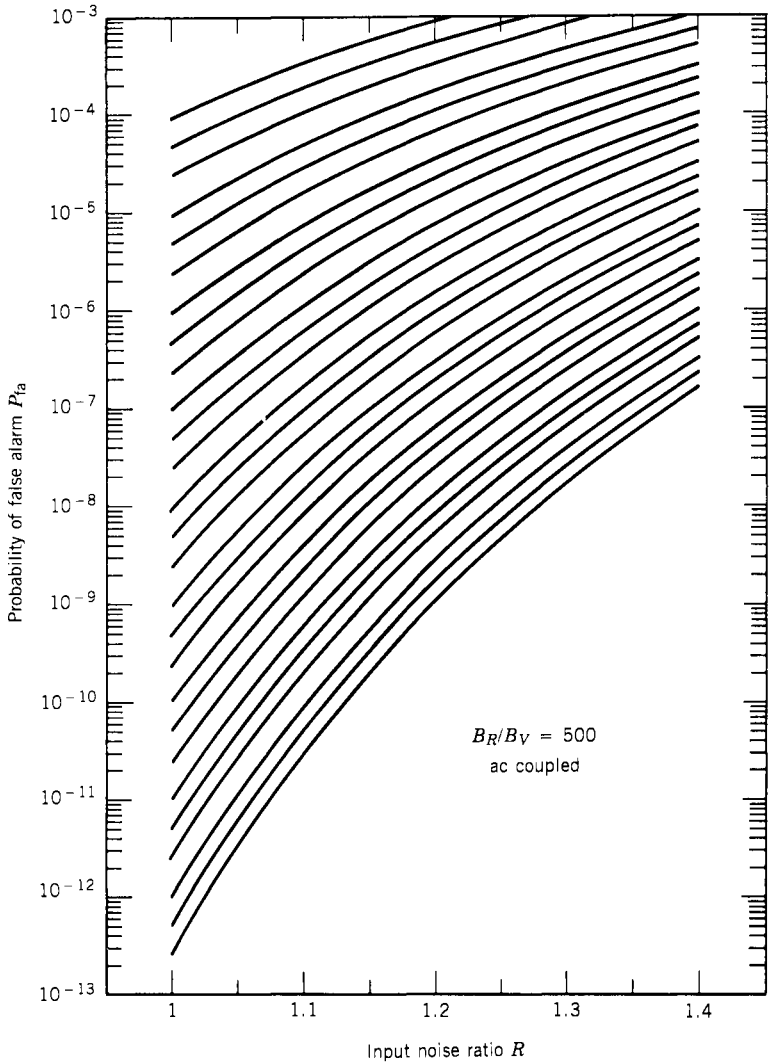


Figure 2.42. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 500$, ac coupled.

6.5 dB. The γ value is 2000 and the video amplifier is ac coupled. If the attenuator A is set at 21 dB, the total noise figure is 27.5 dB (6.5 + 21). Under this condition, the false alarm rate measured is 1.7 Hz, or 1.7 false alarms in 1 sec. The corresponding time between false alarms is

$$T_{fa} = 1/1.7 = 0.59 \quad \text{sec}$$

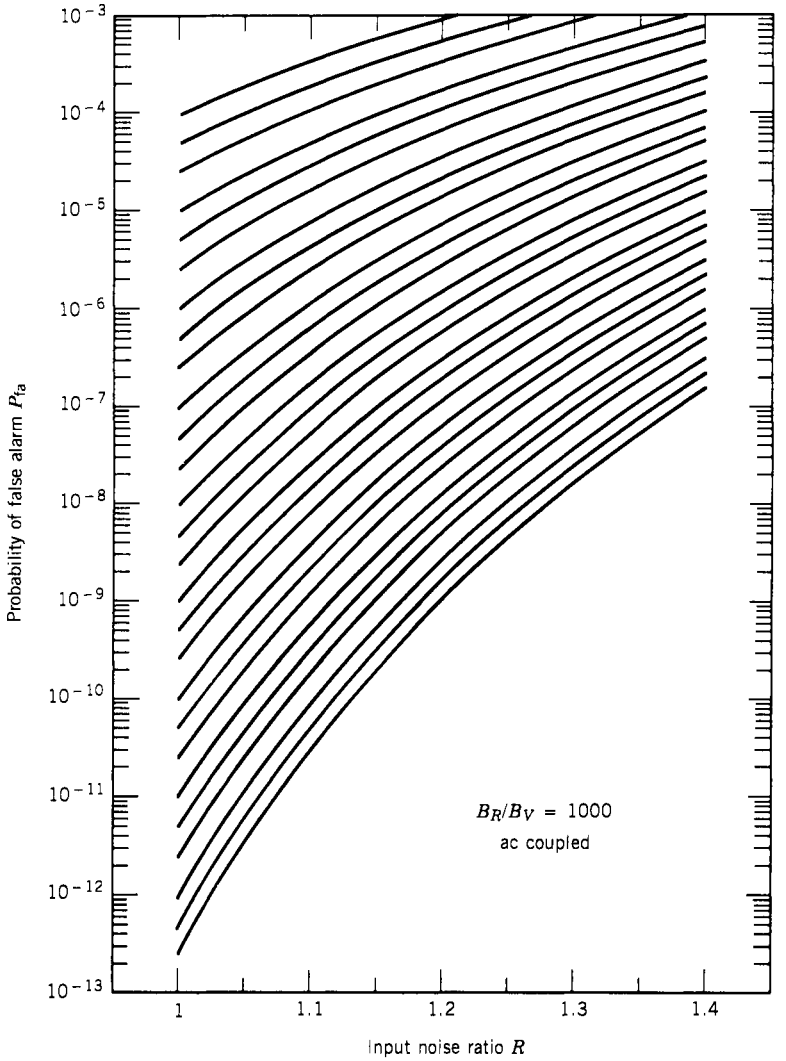


Figure 2.43. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 1000$, ac coupled.

Substituting $F = 27.5$ dB (562.3) and $M = 15$ dB (31.6) into Eq. (2.70), one obtains

$$R = \frac{N_2}{N_1} = 1 + \frac{31.6}{562.3} = 1.056$$

The probability of false alarms is obtained from Eq. (2.24):

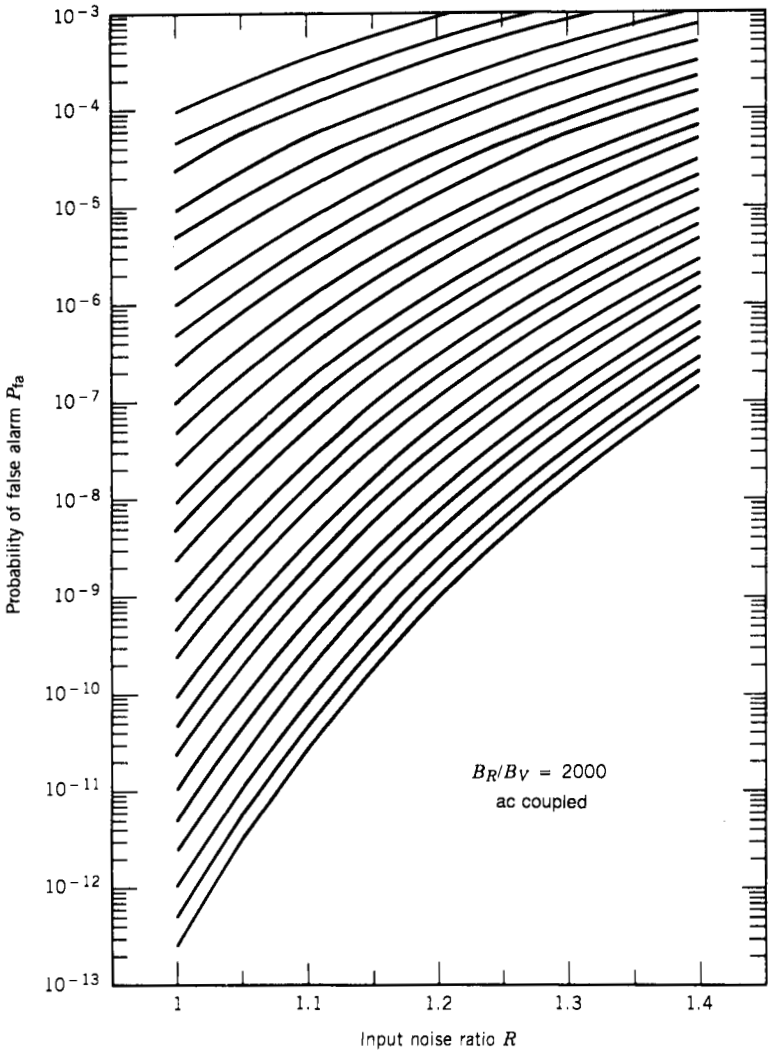


Figure 2.44. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 2000$, ac coupled.

$$P_{fa} = \frac{1}{T_{fa}B_R} = \frac{1.7}{2000 \times 10^6} = 8.5 \times 10^{-10}$$

From Figure 2.44, a point corresponding to the probability of false alarm $P_{fa} = 8.5 \times 10^{-10}$ and $R = 1.056$ is located on the figure. The P_{fa} of $R = 1$ is extrapolated as 1.1×10^{-10} by following the direction of the curves. The

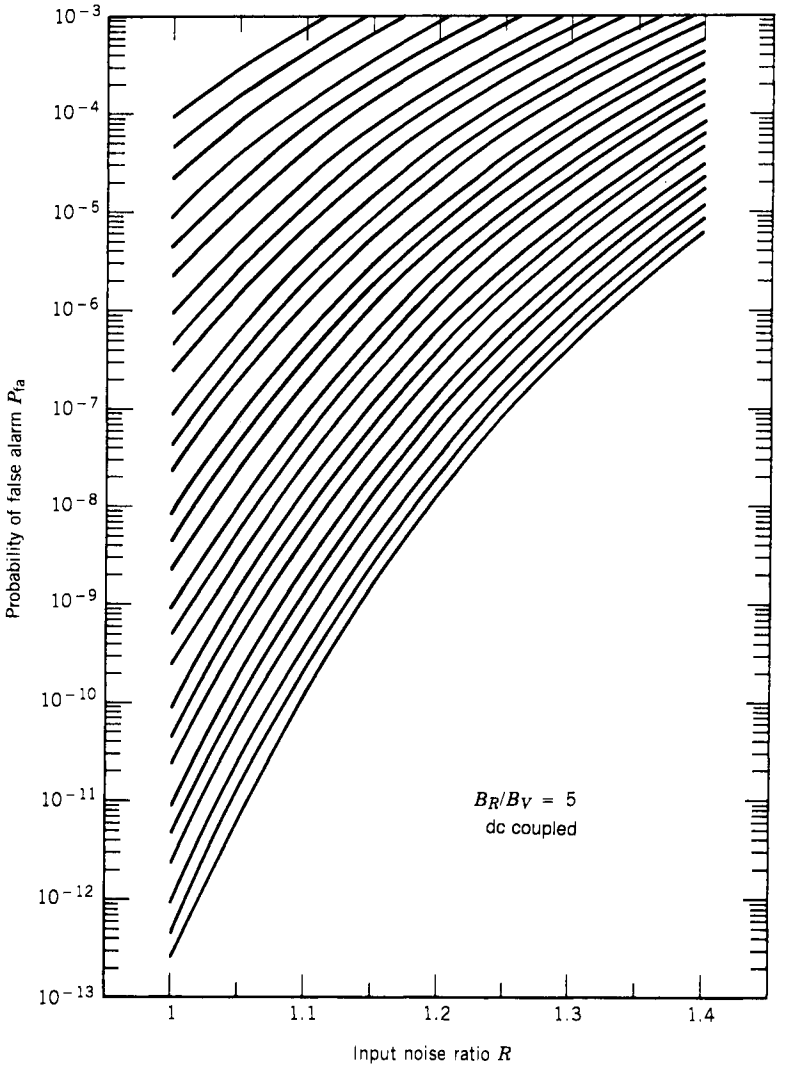


Figure 2.45. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 5$, dc coupled.

time between false alarms T_{fa} is

$$T_{fa} = \frac{1}{P_{fa}B_R} = \frac{1}{1.1 \times 10^{-10} \times 2 \times 10^9} = 4.5 \quad \text{sec}$$

Thus the receiver should generate one false alarm every 4.5 sec under normal operating condition, in agreement with the measured results. In this example,

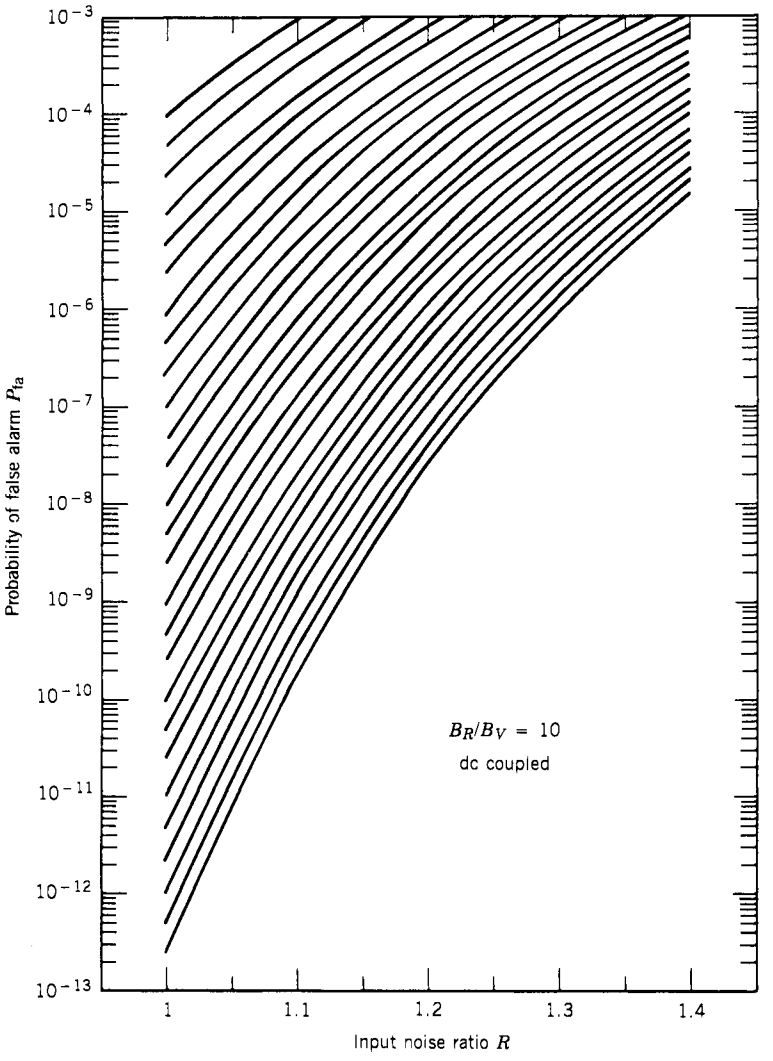


Figure 2.46. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 10$, dc coupled.

the false alarm rate under normal operating conditions can be measured directly and the above procedure is not needed. This special example is presented here because the calculated false alarm rate at operating conditions can be experimentally verified.

In general, from Figures 2.38–2.50, one can observe that the smaller the R value, the more accurate the extrapolated results. However, the R value must be large enough to generate easily measurable false alarm rates. If the value of R

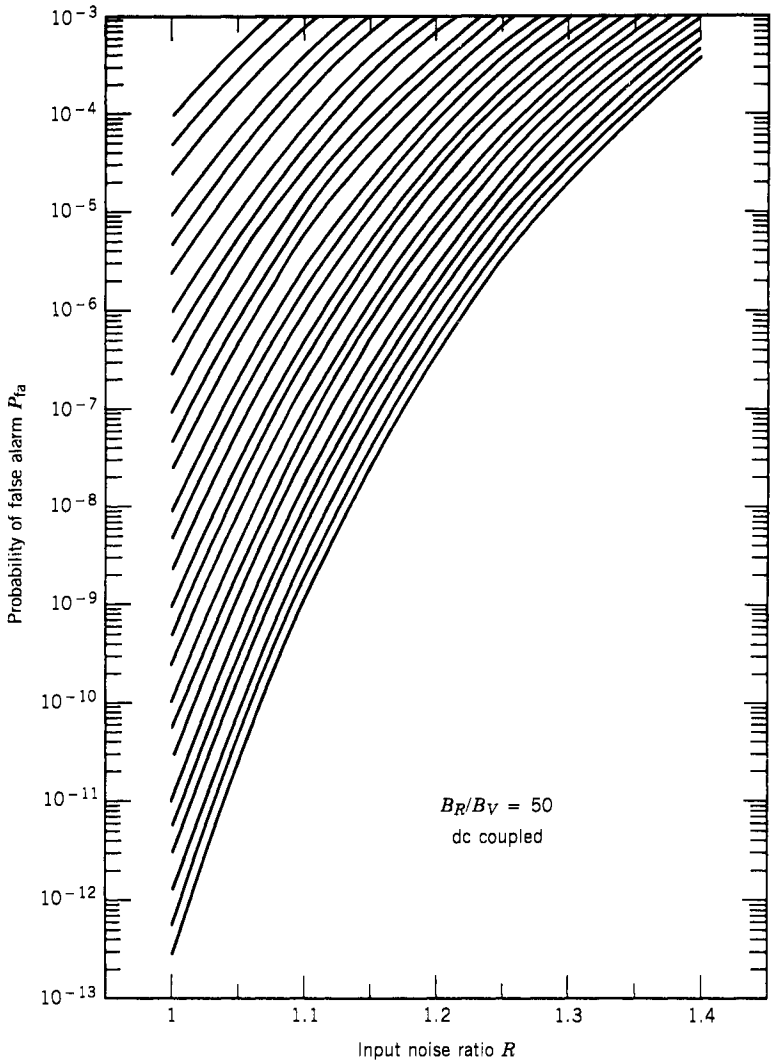


Figure 2.47. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 50$, dc coupled.

is too large, substantial errors can be introduced. For the dc-coupled case, when the γ value is large, it is difficult to obtain accurate results because the slopes of the curves are too steep. In an actual dc-coupled receiver, a slight change in input noise level will induce a large change in the false alarm rate. One can also consider this situation from a different point of view. If the threshold of the receiver is moved up slightly, a large reduction in the false alarm rate will result.

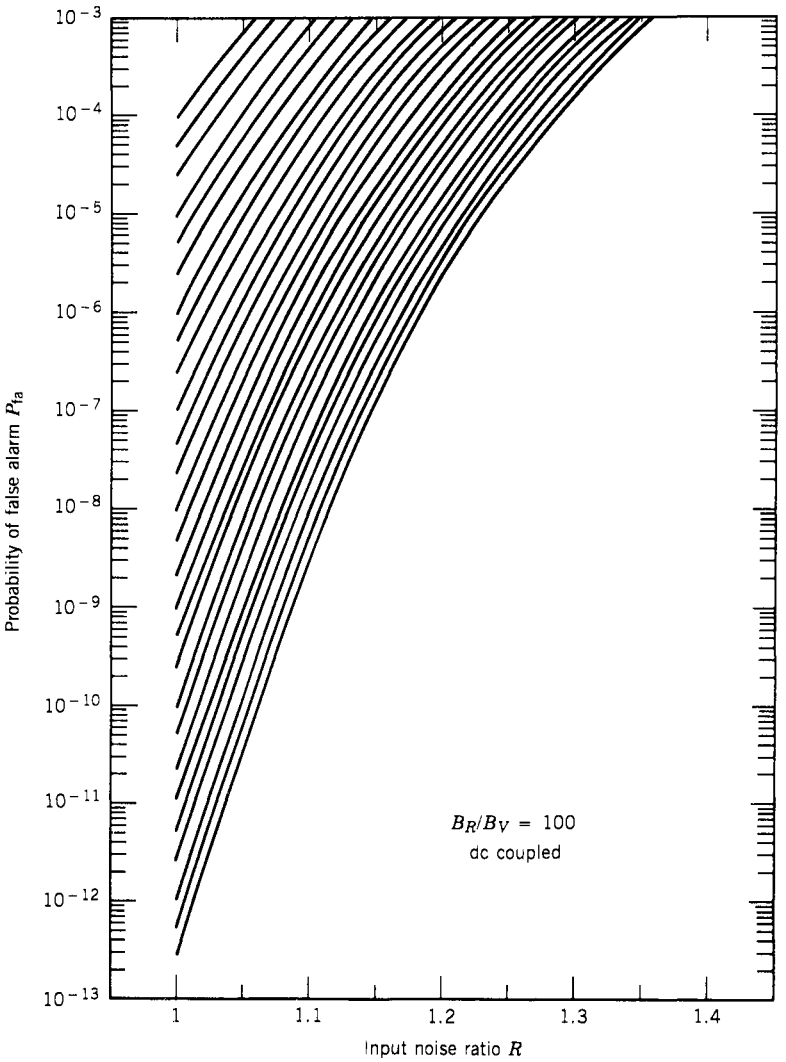


Figure 2.48. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 100$, dc coupled.

Therefore, a receiver with a large γ value can be designed to have a very low false alarm rate with little degradation in sensitivity. This phenomenon is especially true for a dc-coupled receiver. In a dc-coupled receiver with a large γ value, that is, an IFM receiver, it is often not necessary to measure false alarm rates under normal operating conditions because it is too low to be of concern.

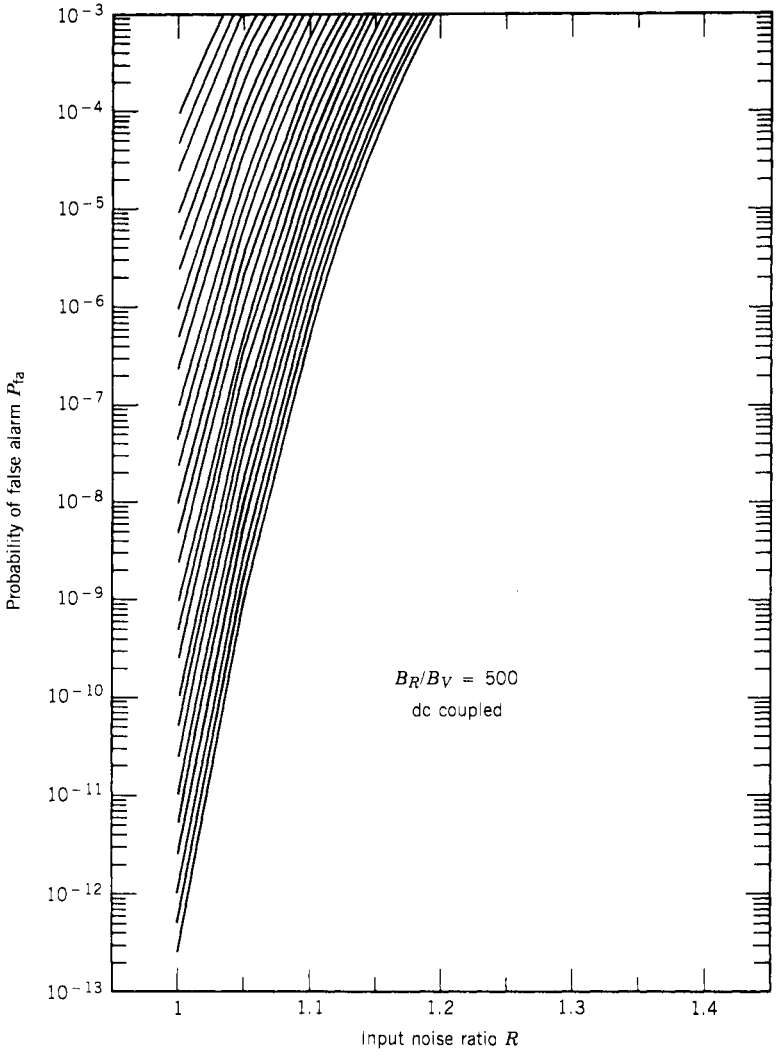


Figure 2.49. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 500$, dc coupled.

2.11. INTRODUCTION TO DYNAMIC RANGE

Dynamic range is a term commonly used to indicate the input signal amplitude range that the receiver can process properly. Unfortunately, there is no universal definition given to this term. The lower limit of the dynamic range is the sensitivity of the receiver, and there is no agreement in this sensitivity level as discussed in the previous sections. Among the different standards used are minimum discernible signal (MDS), which implies $S/N = 1$ (0 dB), tangential sensitivity, and operational sensitivity. Worse yet, the upper limit of the input signal

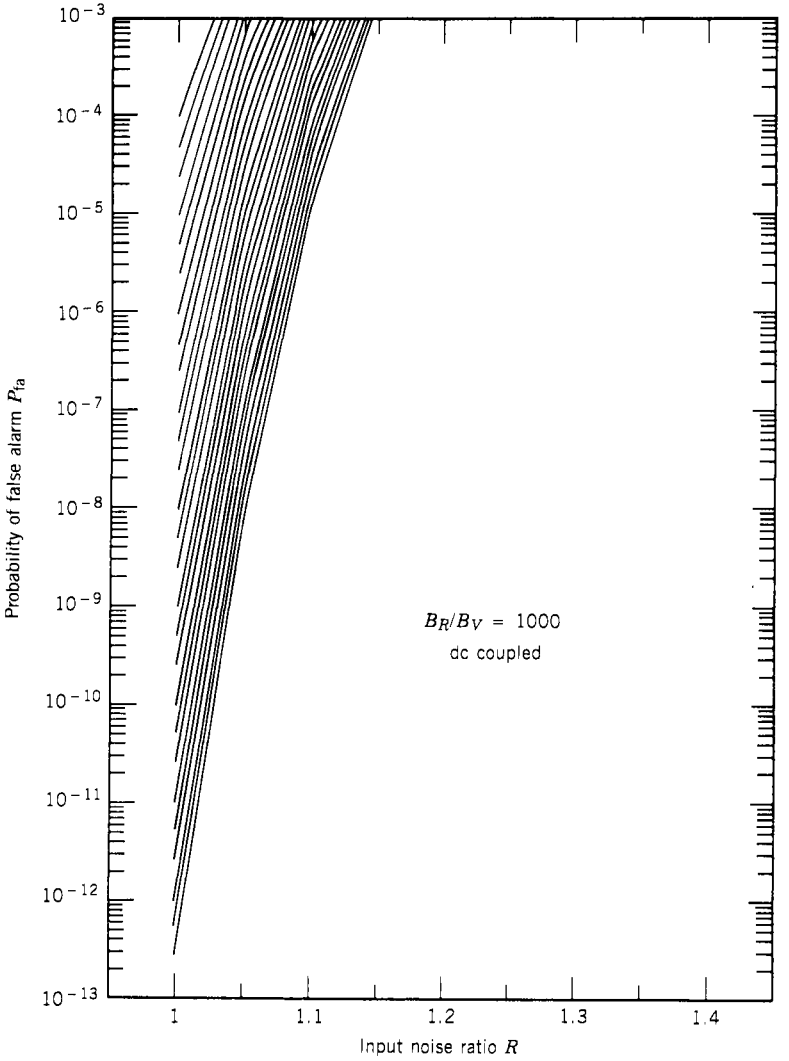


Figure 2.50. Probability of false alarm versus input signal-to-noise ratio with $\gamma = 1000$, dc coupled.

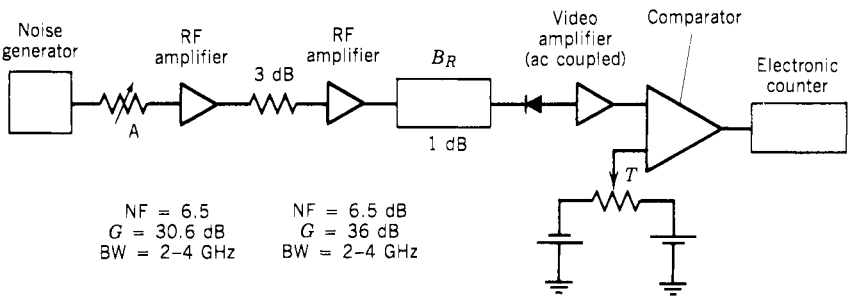


Figure 2.51. Experimental setup for false alarm measurement.

that a receiver can process is not well defined either. Therefore, it is possible for one engineer to claim that a receiver has a 50-dB dynamic range whereas another engineer can claim that the same receiver has only a 30-dB dynamic range.

It is difficult to define a universal dynamic range for all receivers. Instead, in the following sections some of the commonly used dynamic ranges will be discussed. The discussion will place emphasis on the upper limit of the dynamic range, since the lower limit and the sensitivity have been discussed earlier.

One of the upper limits of the dynamic range is the 1-dB compression point. There are other common definitions, such as single-signal dynamic range, two-signal spur-free dynamic range, and two-signal instantaneous dynamic range.

2.12. ONE-dB COMPRESSION POINT (38–44)

To prevent a microwave receiver from generating spurious data, the components in the receiver must operate in linear region. In other words, the output amplitude must be linearly proportional to the input signal amplitude. The amplitude measurement circuit in a receiver is often designed to take this effect into account. If the amplitude is measured in decibels, when the input signal increases 1 dB, the output signal will also increase by 1 dB. This result is shown in Figure 2.52. In Figure 2.52, the gain of a receiver as shown in Eq. (2.5) can be expressed in decibels as

$$G\,(\text{dB}) = 10 \log(S_o/S_i) \tag{2.75}$$

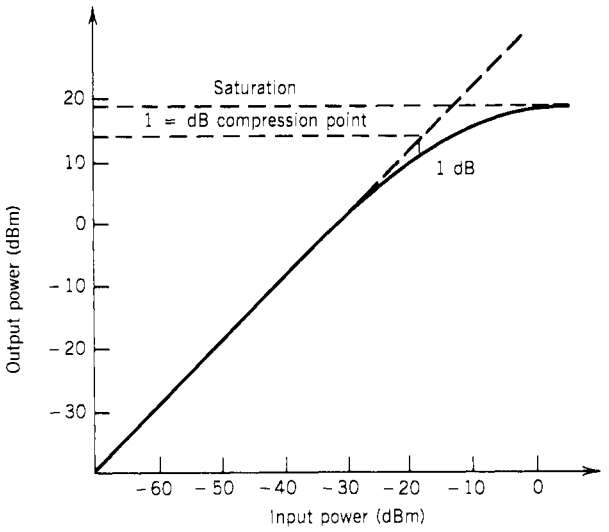


Figure 2.52. Input and output signal relation of a receiver (or amplifier) gain.

or

$$S_o\,(\text{dBm}) = G\,(\text{dB}) + S_i\,(\text{dBm}) \tag{2.76}$$

From Eq. (2.76), one can see that the output is linearly related to the input. If both the x and y axes have the same scale, the slope of the straight line is 1. The gain can be read from the difference on the axes. The amplification gain in Figure 2.52 is 30 dB.

However, in a receiver, if the input signal strength continues to increase, some of the linear components (i.e., amplifiers and mixers) in the receiver will start to saturate. As a result, the output of the receiver will no longer increase linearly with the input signal. For example, when the input increases 1 dB, the output will increase less than 1 dB as shown in Figure 2.52. When the output of the receiver deviates 1 dB from its linear region, the input level (or the output level) is called the 1-dB compression point. The 1-dB compression point in Figure 2.52 is approximately 14 dBm in reference to the output. This definition is commonly used in linear microwave components as well as in receivers that measure signal amplitude.

2.13. SINGLE-SIGNAL DYNAMIC RANGE

When a receiver provides only the frequency information and does not measure the amplitude information on the input signal, the dynamic range can be determined by its frequency measurement capability. Since the frequency measurement is considered to be the primary performance of an EW receiver, the dynamic range is often determined by this capability, even the receiver measures signal amplitude. Under this situation, one may quote two dynamic ranges for the same receiver. One is related to its amplitude measurement capability and the other is related to its frequency measurement capability.

The lower limit of the dynamic range is usually defined as the weakest signal level where the measured frequency error is within a certain predetermined range. The upper limit of the dynamic range is the strongest signal level where the measured frequency error is within the same predetermined range. Let us use an IFM receiver that measures both the frequency and amplitude of the input signal to demonstrate the dynamic range. If the receiver can measure input signals from -65 to -10 dBm with a ± 3-MHz accuracy, the dynamic range of this receiver is often referred to as 55 dB, although the amplitude measurement circuit may have only a 30-dB linear range. In order to avoid confusion in the above example, both dynamic ranges of the receiver should be given.

It will be discussed later that some receivers (i.e., crystal video and IFM receivers) can measure only one signal at a time even though there are simultaneous signals at the input. In some other receivers, if there are simultaneous input signals, the receiver will measure all of them. The receivers with the simultaneous signal capability should report one signal if there is only one input

signal. However, when some linear components in the receiver are driven into the nonlinear region, additional signals (spurs) may appear at the output of the receiver. For example, a mixer may generate strong intermodulation products and an amplifier may produce second harmonics when they are saturated. Often, the dynamic range of a receiver is referred to as the single-signal spur-free dynamic range. Within this dynamic range, if one signal is present at the input of the receiver, the receiver will not generate spurious signals.

2.14. TWO-TONE SPUR-FREE DYNAMIC RANGE AND THIRD-ORDER INTERMODULATION (3, 39–44)

From the terms *two-tone spur-free dynamic range*, one can easily interpret that when two simultaneous signals are within this range, the receiver will not generate any spurious signals. Since two signals can have many different frequency and amplitude combinations, it is desirable to standardize the input conditions to evaluate the performance of a receiver. The general way to define the input condition is that the two signals of the same amplitude and their frequencies are in the input bandwidth of the receiver. Under this condition, the most likely generated spurs are the third-order intermodulation harmonics. Therefore, let us first discuss the third-order intermodulation harmonics. Since the generation of a second harmonic is similar to that of the third harmonic, it will also be discussed here.

If a receiver works in its linear region, two signals at the input will produce two signals at the output and no harmonically related spurious signals will be generated. When the input signals are strong and drive some of the components in the receiver into the nonlinear region, the relationship between the input and output of the receiver can be written as

$$V_o = a_1 V_i + a_2 V_i^2 + a_3 V_i^3 + \cdots \tag{2.77}$$

where V_o and V_i are the output and input voltage, respectively, and a_1, a_2, $a_3, \ldots$ are constants.

The second term in Eq. (2.77) will generate a second harmonic that doubles the frequency of the input signal. Usually, the second harmonic is not a major concern in receivers with less than octave bandwidth since the second harmonic will fall outside the band of the receiver. If the receiver bandwidth is over an octave, the second-order harmonic is of great concern.

Third-order intermodulation products are produced by two simultaneous input signals of different frequencies. To simplify the discussion, the two signals can be expressed as

$$V_i = \cos \omega_1 t + \cos \omega_2 t \tag{2.78}$$

It is assumed that the two signals have the same amplitudes and phases. The third term $(a_3 V_i^3)$ of Eq. (2.77) generates terms containing

$$\cos(2\omega_1 t - \omega_2 t) \qquad \text{and} \qquad \cos(2\omega_2 t - \omega_1 t)$$

where ω_1 and ω_2 are the angular frequencies of the two input signals. The third-order intermodulation harmonics are shown in frequency domain as in Figure 2.53.

The third-order intermodulation products are of the greatest concern since they are the lowest-order intermodulation products that can fall within the passband of a receiver. That is why the third-order intermodulation products are often used as the upper limit of the two-tone spur-free dynamic range. The third-order intermodulation is also a very important performance characteristic in RF amplifiers; therefore, it is often specified in an RF amplifier.

The second- and third-order intermodulations can be expressed in an input versus output plot as shown in Figure 2.54. The input versus output of the fundamental frequency is represented by the straight line of unit slope. The second-order product has a slope of 2:1, whereas the third-order products have a slope of 3:1. Their asymptotes intercept the asymptote of the fundamental, and the intercept points are referred to as the second- and third-order intercept points. The third-order intercept point is often given in an amplifier specification. Knowing the second- and third-order intercept points and the input level, the amplitude of the second- and third-order intermodulation products can be calculated from Figure 2.54. The third-order intermodulation products can be obtained through the intersection of two straight lines: the third-order line and a line parallel to the y axis. The third-order line passes the point $x = Q_3 - G$ and $y = Q_3$ with a slope of 3:1, which can be expressed as

$$\frac{y - Q_3}{x - (Q_3 - G)} = 3 \tag{2.79}$$

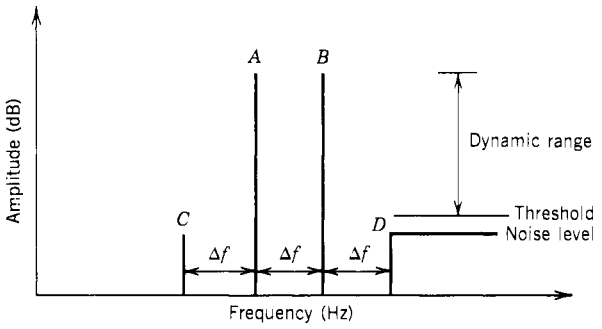


Figure 2.53. Third-order intermodulation harmonics.

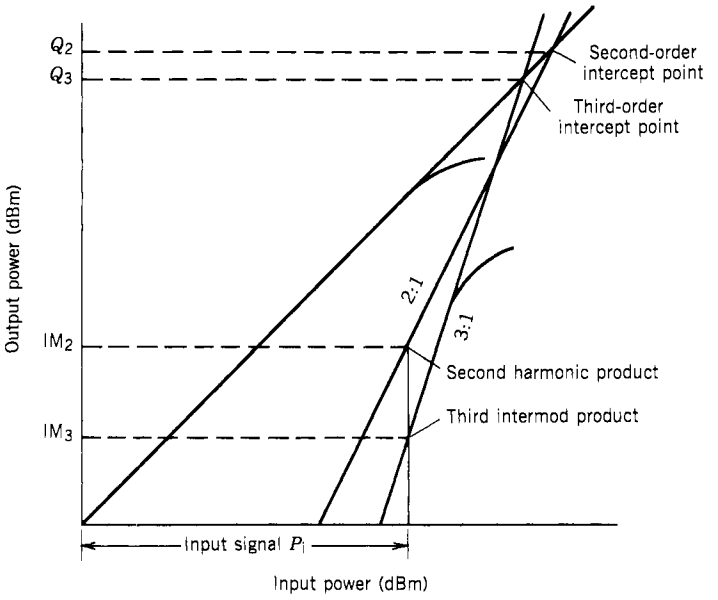


Figure 2.54. Second- and third-order intercept points.

The line parallel to the y axis is

$$x = P_i \tag{2.80}$$

All the quantities in Eqs. (2.79) and (2.80) are in decibels or in dBm because in Figure 2.52 both the input and the output are in dBm. Combining these two equations, one obtains the third-order intermodulation referenced to the output as

$$y = \mathrm{IM}_3 = 3(P_i + G) - 2Q_3 \quad \mathrm{dBm} \tag{2.81}$$

where P_i is the input power level of either signal since both signals have the same amplitude, G is the gain of the receiver in the linear region, and Q_3 is the third-order intercept points referenced to the output.

Similarly, the second harmonic product (referenced to the output) is

$$\mathrm{IM}_2 = 2(P_i + G) - Q_2 \quad \mathrm{dBm} \tag{2.82}$$

where Q_2 is the second-order intercept point referenced to the output.

From Eqs. (2.81) and (2.82), the amplitude of the intermodulation can be calculated if the input signals and the intercept point are known. Next, let us use the second- and third-order intermodulation products to determine the dynamic range. The most reasonable assumption is that if the intermodulation

amplitude equals the noise level of the receiver, the input signal is at the upper limit of the dynamic range. The noise level is

$$N_i = F_T - 114 + 10 \log B_R \quad \text{dBm} \tag{2.83}$$

where F_T is the receiver noise figure and B_R is the RF bandwidth in megahertz. Let the third-order intermodulation equal N_i; then

$$\text{IM}_3 = N_i + G \quad \text{dBm} \tag{2.84}$$

In the above equations, the IM_3 is referenced to the output and the receiver noise N_i is referenced to the input level. Substituting Eq. (2.81) into (2.84), one obtains

$$N_i + G = 3P_i + 3G - 2Q_3 \quad \text{dB} \tag{2.85}$$

The corresponding input level is

$$P_I \equiv P_i = \tfrac{1}{3}(N_i - 2G + 2Q_3) \quad \text{dBm} \tag{2.86}$$

If the two-tone spur-free dynamic range (DR) is defined from this input power level to the noise floor, then

$$\text{DR} = P_I - N_i = \tfrac{2}{3}(Q_3 - G - N_i) \quad \text{dB} \tag{2.87}$$

Using a similar approach, the input signal level that generates a second harmonic at the noise floor is, from Eq. (2.82),

$$P_I = \tfrac{1}{2}(N_i + Q_2 - 2G) \quad \text{dBm} \tag{2.88}$$

The corresponding dynamic range is

$$\text{DR} = \tfrac{1}{2}(Q_2 - N_i - 2G) \quad \text{dB} \tag{2.89}$$

In Eqs. (2.87) and (2.89), the lower limit of the dynamic range is the noise floor. The threshold of a receiver should be above the noise floor. In other words, the useful dynamic range of a receiver is always lower than that given in these equations.

Both results in Eqs. (2.87) and (2.89) show that the higher the intercept points, the higher the dynamic range. Yet, on the other hand, the higher the amplification gain, the lower the dynamic range. Therefore, in receiver designs, the sensitivity and the dynamic range must be carefully compromised.

For the more general cases where the two input signals under consideration are of unequal magnitude, the third-order levels may be computed through

approximation from the following formulas (ref. 39):

$$\mathrm{IM}_3 = 2P_1 + P_2 - 2Q_3 + 3G \quad \text{dBm} \quad \text{for } 2f_1 - f_2 \tag{2.90}$$

$$\mathrm{IM}_3 = 2P_2 + P_1 - 2Q_3 + 3G \quad \text{dBm} \quad \text{for } 2f_2 - f_1 \tag{2.91}$$

where f_1 and f_2 are the signal frequencies with amplitudes of P_1 and P_2, respectively.

To conclude this section, let us consider two examples.

1. A receiver has the following specifications:

$$G = 40 \text{ dB}, \quad Q_3 = 30 \text{ dBm}, \quad F_T = 4 \text{ dB}, \quad \text{and} \quad B_R = 100 \text{ MHz}$$

Then

$$N_i = 4 + (-114) + 10 \log B_R = -90 \quad \text{dBm}$$

The input power that will generate a third-order intermodulation product equal to the noise floor is, from Eq. (2.86),

$$P_I = \tfrac{1}{3}(-90 - 80 + 60) = -36.7 \quad \text{dBm}$$

The corresponding dynamic range is from Eq. (2.87):

$$\mathrm{DR} = -36.7 - (-90) = 53.3 \quad \text{dB}$$

2. If $P_1 = -40$ dBm and $P_2 = -43$ dBm, then the two intermodulation products are from Eqs. (2.90) and (2.91):

$$\mathrm{IM}_3 = 2 \times (-40) + (-43) - 2 \times 20 + 3 \times 30 = -73 \text{ dBm} \quad \text{for } 2f_1 - f_2$$

$$\mathrm{IM}_3 = 2 \times (-43) + (-40) - 2 \times 20 + 3 \times 30 = -76 \text{ dBm} \quad \text{for } 2f_2 - f_1$$

2.15. INTERCEPT POINTS OF CASCADE AMPLIFIERS (45–48)

Usually in a microwave receiver there are many RF amplifiers connected in cascade in front of the detector to achieve high sensitivity. The overall gain of the receiver is simply the product of the individual amplifier, which can be written as

$$G_T = G_1 G_2 G_3 \cdots G_N \tag{2.92}$$

or expressed in decibels as

$$G_T\,(\text{dB}) = G_1\,(\text{dB}) + G_2\,(\text{dB}) + \cdots + G_N\,(\text{dB}) \tag{2.93}$$

where G_1, G_2, . . . are the gain of the first, second, · · · amplifiers. The noise figure of cascade amplifiers can be found through Eq. (2.8).

In this section, the overall intercept points will be calculated if the intercept points of each individual amplifier are given. The intercept points can be derived from the nonlinear input versus output relation as given in Eq. (2.77). Starting with one amplifier,

$$V_o = a_1 V_i + a_2 V_i^2 + a_3 V_i^3 + \cdots \tag{2.77}$$

The input contains two equal amplitude signals of angular frequencies ω_1 and ω_2,

$$V_i = V \sin \omega_1 t + V \sin \omega_2 t \tag{2.94}$$

Substituting Eq. (2.94) into (2.77), one obtains

$$\begin{aligned} V_o = {} & a_1 V(\sin \omega_1 t + \sin \omega_2 t) \\ & + a_2 V^2[-\tfrac{1}{2}\cos 2\omega_1 t - \tfrac{1}{2}\cos 2\omega_2 t + \cos(\omega_1 - \omega_2)t - \cos(\omega_1 + \omega_2)t + 1] \\ & + a_3 V^3[\sin^3 \omega_1 t + \sin^3 \omega_2 t + \tfrac{3}{2}\sin \omega_1 t + \tfrac{3}{2}\sin \omega_2 t - \tfrac{3}{4}\sin(2\omega_1 + \omega_2)t \\ & - \tfrac{3}{4}\sin(2\omega_2 + \omega_1)t + \tfrac{3}{4}\sin(2\omega_1 - \omega_2)t + \tfrac{3}{4}\sin(2\omega_2 - \omega_1)t] \end{aligned} \tag{2.95}$$

If the input impedance of the amplifier is R, which matches the impedance of the signal generator, the input power is

$$P_i = \left(\frac{V}{2\sqrt{2}}\right)^2 \Big/ R_s = \frac{V^2}{8R_s} \tag{2.96}$$

In the above equation, the input voltage is divided equally between the input and the source impedances. The output power calculated from the fundamental frequency term to the load impedance of R_o is

$$P_1 = GP_i = \frac{a_1^2 V^2}{8R_o} \tag{2.97}$$

Let us assume that $R_o = R_s$, then from Eqs. (2.96) and (2.97) the gain can be

written as

$$G = a_1^2 \tag{2.98}$$

The power output corresponding to frequency $2\omega_1$ or $2\omega_2$ (the second harmonics) is obtained from Eq. (2.95) as

$$P_2 = \frac{a_2^2 V^4}{8 \times 4R_s} = \frac{a_1^2 V^4}{64R_s^2} \frac{2a_2^2 R_s}{a_1^2} = \frac{2a_2^2 R_s}{a_1^2} GP_i^2 \tag{2.99}$$

The power output of the third-order intermodulation that corresponds to the frequency $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ is

$$P_3 = \frac{9a_3^2 V^6}{8 \times 16R_s} = \frac{a_1^2 V^6}{512R_s^3} \frac{9 \times 4R_s^2 a_3^2}{a_1^2} = \frac{36a_3^2 R_s^2}{a_1^2} GP_i^3 \tag{2.100}$$

Let Q_{2i} be a certain value of P_i such that $P_1 = P_2$. Thus, Q_{2i} corresponds to the second-order input intermodulation intercept point. Here the subscript i is used to indicate that the intercept point is referenced to the input where the Q's in Section 2.14 are referenced to the output of a receiver. By equating Eqs. (2.97) and (2.99),

$$GQ_{2i} = \frac{2a_3^2 R_s}{a_1^2} GQ_{2i}^2 \tag{2.101}$$

or

$$\frac{1}{Q_{2i}} = \frac{2a_2^2 R_s}{a_1^2} \tag{2.102}$$

Let Q_{3i} be a certain value of P_i such that $P_1 = P_3$. Thus Q_{3i} is the third-order input intermodulation intercept point. By equating Eqs. (2.97) and (2.100)

$$GQ_{3i} = \frac{36a_3^2 R_s^2}{a_1^2} GQ_{3i}^3 \tag{2.103}$$

or

$$\frac{1}{Q_{3i}} = \frac{6a_3 R_s}{a_1} \tag{2.104}$$

In the above equations, the intercept points are expressed in terms of constants a_1, a_2, and a_3 and the input and output impedances. Now, let us discuss the cascade amplifiers. Like the derivation of the noise figure, two amplifiers in

cascade as shown in Figure 2.55 are used to derive the composite intercept points. The general results of multiple staged amplifiers can be extended from the results of the two amplifiers. The first amplifier can be expressed as

$$V_1 = b_1 V_i + b_2 V_i^2 + b_3 V_i^3 + \cdots \tag{2.105}$$

and the second amplifier as

$$V_2 = c_1 V_1 + c_2 V_1^2 + c_3 V_1^3 + \cdots \tag{2.106}$$

where b_i and c_i are constants, and V_i and V_1 are the input and output of amplifier 1 and V_2 is the output of amplifier 2. The output of amplifier 1 is also the input of amplifier 2. Substituting Eq. (2.105) into (2.106),

$$\begin{aligned} V_2 = {} & c_1(b_1 V_i + b_2 V_i^2 + b_3 V_i^3) + c_2(b_1 V_i + b_2 V_i^2 + b_3 V_i^3)^2 \\ & + c_3(b_1 V_i + b_2 V_i^2 + b_3 V_i^3)^3 \end{aligned} \tag{2.107}$$

which can be rearranged as

$$V_2 \cong b_1 c_1 V_i + (b_2 c_1 + b_1^2 c_2) V_i^2 + (b_3 c_1 + b_1^3 c_1) V_i^3 + \cdots \tag{2.108}$$

where the term $2b_1 b_2 c_2 V_i^3$ belonging to the third term is neglected because its amplitude is small, since it results from two nonlinear terms b_2 and c_2. Comparing Eq. (2.108) with (2.77), the composite constants of the cascade amplifiers are

$$a_1 = b_1 c_1 \tag{2.109}$$

$$a_2 = b_2 c_1 + b_1^2 c_2 \tag{2.110}$$

$$a_3 = b_3 c_1 + b_1^3 c_1 \tag{2.111}$$

and

$$a_2^2 = (b_2 c_1)^2 + (b_1^2 c_2)^2 + 2 b_2 c_1 b_1^2 c_2 \tag{2.112}$$

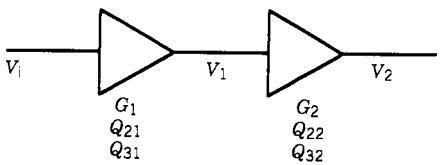


Figure 2.55. Two amplifiers in cascade.

The overall second-order intermodulation is given by Eqs. (2.102), (2.109), and (2.112) as

$$\frac{1}{Q_{2iT}} = \frac{2a_2^2 R_s}{a_1^2} = \frac{2b_2^2 c_1^2 R_s}{b_1^2 c_1^2} + \frac{2b_1^4 c_2^2 R_s}{b_1^2 c_1^2} + \frac{4b_2 c_1 b_1^2 c_2 R_s}{b_1^2 c_1^2}$$

$$= \frac{2b_2^2 R_s}{b_1^2} + \frac{2b_1^2 c_2^2 R_s}{c_1^2} + 2\sqrt{\frac{2b_2^2 R_s}{b_1^2}\frac{2c_2^2 R_s b_1^2}{c_1^2}} \tag{2.113}$$

Use the relation in Eq. (2.102) again. Equation (2.113) can be written as

$$\frac{1}{Q_{2iT}} = \frac{1}{Q_{2i1}} + \frac{G_1}{Q_{2i2}} + 2\sqrt{\frac{G_1}{Q_{2i1}Q_{2i2}}} \tag{2.114}$$

which can be again written as

$$\sqrt{\frac{1}{Q_{2iT}}} = \sqrt{\frac{1}{Q_{2i1}}} + \sqrt{\frac{G_1}{Q_{2i2}}} \tag{2.115}$$

where Q_{2i1} and Q_{2i2} are the second-order intercept points of amplifiers 1 and 2, respectively, and G_1 is the gain of amplifier 1.

The third-order intercept point is from Eqs. (2.104), (2.109), and (2.111).

$$\frac{1}{Q_{3iT}} = \frac{6a_3 R_s}{a_1} = \frac{6b_3 c_1 R_s}{b_1 c_1} + \frac{6b_1^3 c_1 R_s}{b_1 c_1} = \frac{1}{Q_{3i1}} + \frac{G_1}{Q_{3i2}} \tag{2.116}$$

where Q_{3i1} and Q_{3i2} are the third-order intercept points of amplifiers 1 and 2, respectively. Equations (2.115) and (2.116) can be generalized as

$$\sqrt{\frac{1}{Q_{2iT}}} = \sqrt{\frac{1}{Q_{2i1}}} + \sqrt{\frac{G_1}{Q_{2i2}}} + \sqrt{\frac{G_1 G_2}{Q_{2i3}}} + \cdots \tag{2.117}$$

and

$$\frac{1}{Q_{3iT}} = \frac{1}{Q_{3i1}} + \frac{G_1}{Q_{3i2}} + \frac{G_1 G_2}{Q_{3i3}} + \cdots \tag{2.118}$$

Equations (2.117) and (2.118) are used to calculate the intercept points for a cascade chain of amplifiers. The unit used in the above equation is power in watts or milliwatts. In the above derivation, the intercept points are referenced to the input of each component. However, the intercept points of commercial components are usually specified with respect to the output level. The input intercept point can be obtained by substracting the gain (in decibels) of the component from the output intercept point.

If any component in the RF chain in front of the detector is passive, a high

intercept point (i.e., 100 dBm) can be assigned to the component, which will not affect the overall intercept point and then use Eqs. (2.117) and (2.118) to calculate the overall intercept points. An example in the next section will demonstrate the use of these equations.

2.16. GRAPHIC METHOD OF DETERMINING NOISE FLOOR AND INTERCEPT POINT (49)

A graphic method to determine the noise floor and the third-order intercept point of a receiver will be discussed in this section. The results obtained from this approach are not very accurate under some specific conditions. However, the results will provide a clear picture of the limitation of the noise floor and the third-order intercept point. The detailed discussion will be illustrated through an example. The steps are listed briefly below:

1. Find the absolute noise floor of the receiver.
2. Plot the intrinsic noise floor of each component with respect to the absolute noise floor.
3. Plot the input third-order intercept point of each component.
4. From the lowest intrinsic noise floor, plot the relative noise floor.
5. From the relative noise floor and the input third-order intercept point, determine the upper limit of the signal level limited by each component.
6. From the minimum range between relative noise to the upper signal level, plot the relative upper signal limit.
7. From the input intrinsic noise floor and the upper signal limit, determine the overall input third-order intercept point.

To demonstrate the graphic method, let us use the following example. Figure 2.56 shows an RF chain of a receiver in front of the video detector. Components 1 and 5 are filter banks with different bandwidths. The third-order intercept points of components 2, 3, and 4 are listed. Amplifiers 6–9 are placed after multiplexing filter 5, where simultaneous signal conditions will not be considered. Thus the third-order intercept points of amplifiers 6–9 are not of concern and are not listed.

The first row in Figure 2.56 represents the component number. The second row lists the noise figure F of each component. The third row lists the gain G of each component. The fourth row lists the third-order intercept point Q_3 of each component referenced to the output level in. The input third-order intercept points Q_{3i} are obtained by subtracting the gain from the output third-order intercept points of the component and are listed in the fifth row. The units of these components are decibels or dBm. The numbers in parenthesis are the equivalent power or amplitude. The bandwidth of filter 5 is 81.25 MHz.

Component	1	2	3	4	5	6	7	8	9
F	7	6	3	8	7	6.5	6.5	6.5	6.5
G	−7 (0.2)	−6 (0.25)	15 (31.6)	5	−7	10	10	10	10
Q_3		6	−3	30					
Q_{3i}	100	12 (15.8)	−18 (0.0158)	25 (31.6)					

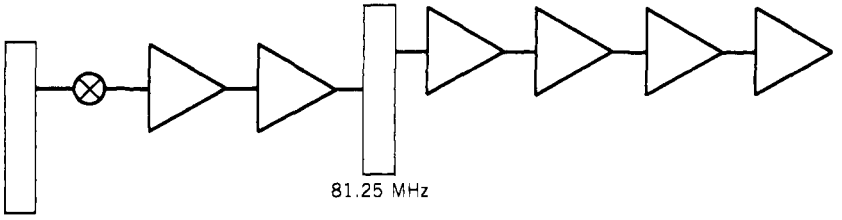


Figure 2.56. RF chain before detector.

The noise figure, from Eq. (2.8), is 16.5 dB. Thus the noise at the input of the receiver is

$$-114 + 10\log(81.25) + 16.5 = -78.4 \quad \text{dBm}$$

The overall input third-order intercept point can be found from Eq. (2.118) as

$$\frac{1}{Q_{3iT}} = \frac{0.2}{15.8} + \frac{0.2 \times 0.25}{0.0158} + \frac{0.2 \times 0.25 \times 31.6}{316}$$

$$= 0.0126 + 3.15 + 0.005 = 3.168$$

$$Q_{3iT} = 0.316 \text{ mW} = -5 \quad \text{dBm}$$

In this calculation, the term $1/Q_{3i1}$ is neglected because of the large value of Q_{3i1}, which can be assigned as 100 dB.

From Eq. (2.86), the input signal level generating third-order intermodulations equal to the noise level can be calculated. The gain G in Eq. (2.86) is 0, since in the above calculation Q is referenced to the input. Thus

$$P_I = \tfrac{1}{3}[-78.4 + 2 \times (-5)] = -29.5 \quad \text{dBm}$$

The graphic method is illustrated step by step in Figure 2.57.

1. To find the absolute noise level,

$$A = -114 + 10\log(81.25) = -94.9 \quad \text{dBm}$$

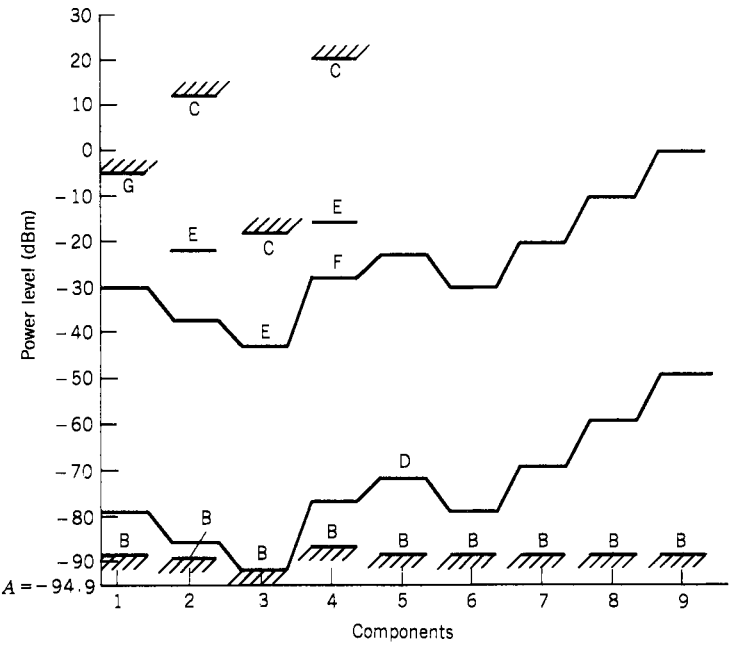


Figure 2.57. Graphic method determining the noise level and third-order intercept point.

2. Mark the intrinsic noise floor B of each component by drawing a line above the absolute noise floor by the amount of the noise figure of the component.
3. Mark the input third-order intercept point C of each component.
4. From the lowest intrinsic noise floor (component 3), add (or substract) the gain of each component to find the relative noise floor D. The relative input noise floor is about -78 dBm, which is close to the calculated value -78.4 dBm.
5. From the relative noise floor D and the third-order intercept point C, find the upper limit of signal level E. At signal level E, the third-order intermodulation is equal to the relative noise floor. Signal level E is two-thirds the distance DC above the intrinsic noise floor D.
6. Obtain the lowest dynamic range by measuring the minimum distance from D to E. In this example, the minimum occurs at component 3. From the lowest E value, draw the relative upper signal level F by adding (or subtracting) the gain of each component. The input signal upper limit level is about -30 dBm, whereas the calculated value is -29.5 dBm.
7. Find the overall input third-order intercept point G by drawing a line one-half the distance DF above level F at the input. The Q_{3iT} value obtained here is about -5 dBm.

The results obtained from the graphic method are very close to the calculated results. From this figure, one can see that the noise floor D is limited by the third component, and the upper limit of the dynamic range is also limited by the third component. The signal level at each stage is clearly shown. The upper limit and the noise floor is determined by one component only. However, in Eqs. (2.8) and (2.118), the noise figure and the input third-order intercept are affected by all the components in the circuit.

The problem with this graphic approach is that when two or more components limit the noise floor or the upper signal level at the same time, the graphic method will not show the effect. However, through Eqs. (2.8) and (2.118), it is expected that under these conditions the noise floor will move upward and the upper signal level will move downward because of the contribution of all components.

2.17. TWO-SIGNAL INSTANTANEOUS DYNAMIC RANGE

If a receiver can receive simultaneous signals, that is, channelized, microscan, and Bragg cell receivers, the simultaneous input signals often interfere with each other. Let us assume that the two signals are not of the same aplitude: a strong one and a weak one. If the strong signal drives the receiver into the nonlinear region, it will suppress the weak one. Under this condition, both the amplitude measurements will be inaccurate. Both signals will be measured less than the actual values because of the saturation effect. If the suppression effect is very strong, the receiver will miss the weak signal.

The two-signal instantaneous dynamic range is defined as the maximum amplitude separation between two simultaneous signals such that the receiver can measure both signals correctly. Often the frequency measurements of the signals are of prime consideration. Therefore, the instantaneous dynamic range is considered as the maximum amplitude separation between two signals such that the receiver can measure both their frequencies correctly. In general, the instantaneous dynamic range of a receiver is a function of the frequency separation between the two signals. If the inputs are pulsed signals, their spectrum will be spread in the frequency domain. For example, a pulse of amplitude A and width T can be written as

$$\begin{aligned} S(t) &= A \quad \text{for } -\tfrac{1}{2}T < t < \tfrac{1}{2}T \\ &= 0 \quad \text{elsewhere} \end{aligned} \tag{2.119}$$

Its Fourier transform is

$$S(f) = \int_{-T/2}^{T/2} A \exp(-j\omega t)\, dt = AT \frac{\sin \pi f T}{\pi f T} \tag{2.120}$$

where f is frequency and $\omega = 2\pi f$. The power spectrum $S(f)^2$ plotted in decibels for $A = 1$ is shown in Figure 2.58. When the inputs are two simultaneous pulses, the weak signal must not be buried in the side lobes of the strong one because the receiver cannot detect the weak one. Therefore, in evaluating instantaneous dynamic range, the pulse width of the input signals must be long enough such that this condition will not occur. Continuous-wave or long pulsed signals are often used to evaluate instantaneous dynamic range instead of pulsed signals. Furthermore, the minimum frequency separation between the two signals must be greater than the frequency resolution of the receiver; otherwise, the receiver cannot differentiate them.

A typical result of instantaneous dynamic range is shown in Figure 2.59. The farther the two signals are separated in frequency, the wider the dynamic range because filters are often used in receivers to separate signals close in frequency. The dynamic range versus frequency plot resembles the filter shape used in the receiver when the two signals are close in frequency. When the two signals are separated far apart in frequency, the filter effect will no longer show, and the dynamic range is more or less a constant value determined by the linear active components used in the receiver.

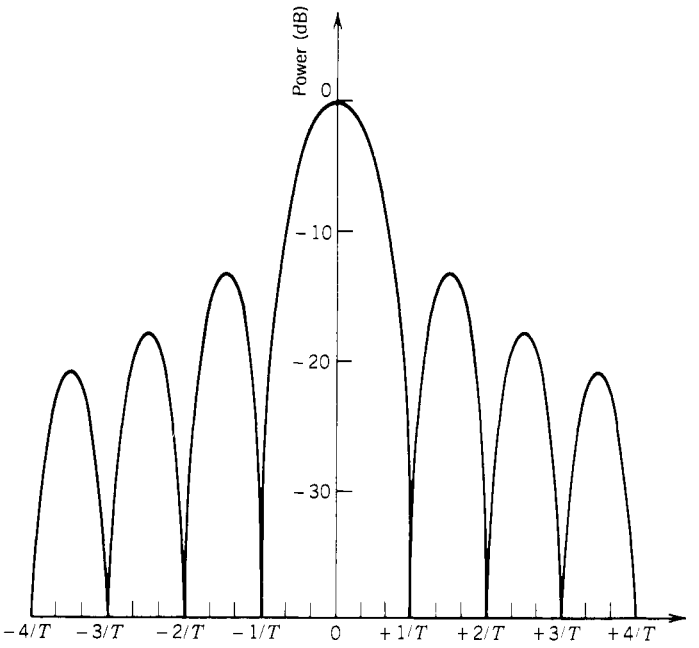


Figure 2.58. Power spectrum of a pulsed signal.

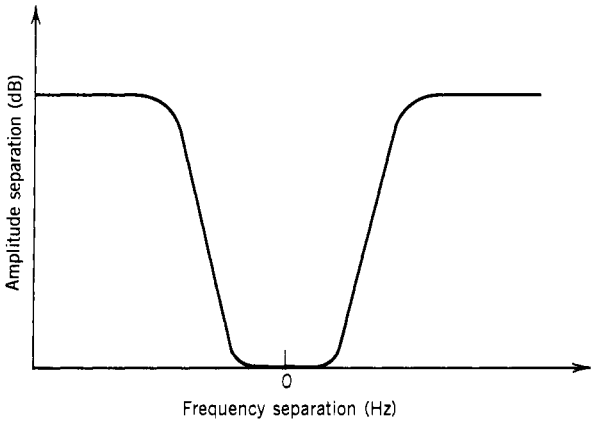


Figure 2.59. Typical two-signal dynamic range.

2.18. PROBABILITY OF INTERCEPT (50, 51)..

The term POI is used in connection with EW receivers. It tells the percentage of pulses the receiver will collect in a certain signal environment. Of course, this is a complicated problem. It depends on the density and distribution of the pulses in the environment. Therefore, POI is a very loosely defined term.

The POI can be determined from two aspects: spatial and frequency range. If the receiving antenna covers only a section of the entire area of interest, the antenna can be used in a search mode to cover the entire area. The receiver will miss some pulses in this time-sharing mode. If the antenna covers an angle of θ_A and the angle of interest is θ_T, the POI is

$$\text{POI}_{\text{AOA}} = \theta_A/\theta_T \tag{2.121}$$

If the instantaneous bandwidth of the receiver does not cover the entire frequency range of interest, the receiver must be time shared among the different frequency ranges. If B_R is the input bandwidth of the receiver and B is the overall spectrum of interest, the POI of the above situation can be easily determined as

$$\text{POI}_{\text{BW}} = B_R/B \tag{2.122}$$

The overall POI is

$$\text{POI} = \text{POI}_{\text{AOA}} \times \text{POI}_{\text{BW}} \tag{2.123}$$

When the antenna coverage is equal to or greater than the area of interest and the instantaneous bandwidth equals the input bandwidth, the overall POI is 100% from Eq. (2.123). However, it must be emphasized that even under this

condition the receiver can still miss some pulses. One situation is that some of the signals are outside the instantaneous dynamic range as discussed in the above section. Another situation is that if one signal is followed very closely by a second signal, the second may be missed by the receiver. This will be discussed in the next section. It is almost always possible to put up certain signal conditions so that a receiver will miss some pulses.

2.19. THROUGHPUT RATE, SHADOW TIME, AND SIGNAL DELAY TIME

The throughput rate is applicable to receivers dealing primarily with pulsed signals. It tells the maximum pulse rate the receiver can process. If the receiver measures the pulse width (PW) information, the throughput rate often depends on the PW of the input signal, since the receiver must wait until the end of the pulse to generate the PW information. To define a throughput rate independent of PW, the receiver should be evaluated with signals of short PW. The PW should be short enough that the receiver process time depends on the intrinsic property of the receiver rather than on the PW.

If a receiver can process several simultaneous signals, then the throughput rate is difficult to measure. When two signals are very close in time or are time coincident, both of them will be properly processed by the receiver. Under this condition, one way to create a meaningful throughput rate is to use simultaneous signals equal to the number of signals the receiver can process simultaneously; then the PRF is increased until the receiver starts to drop pulses, and the highest PRF the receiver can process is the throughput rate.

The throughput time is defined as the inverse of the throughput rate. Its unit is often expressed in microseconds, and this is just a different way to express the maximum pulse rate the receiver can process.

Shadow time is another way to present the signal processing rate of a receiver. It is usually specified as the minimum time from the trailing edge of a first pulse to the leading edge of a second pulse such that the receiver can process both of them. As discussed above, the shadow time also depends on the PW when the PW is over a certain limit. Thus, signals of short PW should be used to evaluate shadow time.

Another parameter often of interest to receiver engineers is the delay time of the receiver. It takes a certain time for a signal to pass through a receiver and be encoded in digital form. In most receivers, a data-ready pulse is often generated at the output of the receiver to tell a following processor that data is ready to be taken out. The signal delay time is often measured from the lead of the RF pulse to the data-ready pulse. This delay time is not necessarily related to the throughput time because the receiver can work in a pipeline form. For example, if a 2-μsec delay line is used in a receiver, the signal delay time will be longer than 2 μsec. However, the receiver can have a throughput time of only 1 μsec, which is shorter than the signal delay time, because the receiver may accept and process

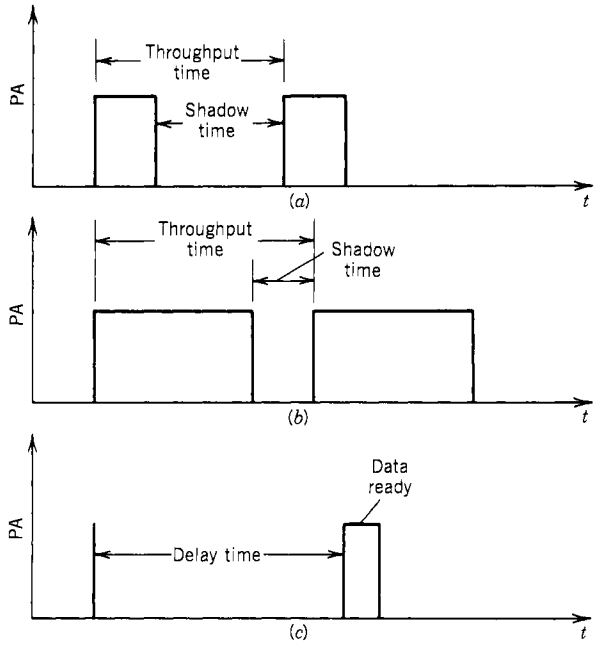


Figure 2.60. Throughput, shadow, and delay times in a receiver: (*a*) short pulses; (*b*) long pulses; (*c*) delay time.

another signal before the first signal is completely encoded. Figure 2.60 shows the different times discussed in this section.

2.20. SUMMARY

The characteristics and performance of receivers are discussed in this chapter. Emphasis is placed on sensitivity and dynamic range because they are the most important factors in a receiver. Tangential sensitivity and operational sensitivity are the commonly used terms because they can be measured experimentally. The noise figure with the relative equations are presented, and they are very useful in receiver designs. A method of measuring false alarm rate is presented in this chapter because the equations used are closely related to the sensitivity calculations.

There are many different definitions of dynamic ranges. For a systems engineer, the commonly used ones are the single-signal and two-tone spur-free dynamic ranges. They can be measured experimentally and will be discussed in Chapter 12. Some terms that are commonly used in EW receivers, such as POI and throughput rate, are also discussed.

In addition, a graphic method of determining receiver sensitivity and

dynamic range is introduced. Although this method may be somewhat inaccurate under certain conditions, it will provide a clear view of the noise floor and dynamic range limitations of every component in a receiver.

REFERENCES

1. T. K. Ishii, *Microwave Engineering*, Ronald Press, New York 1966.
2. M. I. Skolnik, *Introduction to Radar Systems*, Chapter 8, McGraw-Hill, New York, 1962.
3. D. K. Barton, *Radar System Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1964.
4. R. Pettai, *Noise in Receiving Systems*, Wiley, New York, 1984.
5. *IRE Dictionary of Electronics Terms and Symbols*, Institute of Radio Engineers, New York, 1961.
6. IRE standards on electron tubes: definitions of terms 1957, *Proc. IRE*, **45**, 983–1010 (1952).
7. J. H. Frohmaier, Noise performance of a three-stage microwave receiver, *Electronic Tech.*, **37**, 245–246 (1960).
8. D. William, Visual measurement of receiver noise, *Wireless Engineer*, **24**, 100–104 (April 1947).
9. W.J. Lucas, Tangential sensitivity of a detector video system with R.F. pre-amplification, *Proc. IEE*, **113**, 1321–1330 (1966).
10. S. Sareen, Threshold detectors for nanosecond fault or level detection in RF system, Application Note, p. 40, Aertech Industries, Sunnyvale, CA, 1975.
11. The criterion for the tangential sensitivity measurement, Hewlett Packard Application Note 956-1, Oct. 1973.
12. L. Klipper, Sensitivity of crystal video receivers with RF preamplification, *Microwave J.*, **8**, 85–92 (1965).
13. James B. Y. Tsui, Tangential sensitivity of EW receivers, *Microwave J.*, **24**, 99 (Oct. 1981).
14. Spectrum analysis—Noise figure measurement, Hewlett Parkard Application Note 150-9, April 1976.
15. D. L. Adamy, Calculate receiver sensitivity, *Electronic Design*, **25**, 118 (December 6, 1973).
16. S. E. Lipsky, Calculate the effects of noise on ECM receivers, *Microwaves*, **13**, 65 (October 1974).
17. G. Franklin and T. Hatley, Dont eyeball noise, *Electronic Design*, **24**, 184 (November 22, 1973).
18. D. L. Johnston, IFM and channelized receiver detection sensitivity curves, The International Countermeasures Handbook, 463, EW Communications Inc. Palo Alto CA June 1977.
19. A. R. Martin and R. F. Cobb, A guide to acquisition receiver selection and performance, *Microwave J.*, **11**, 63 (June 1968).
20. F. Rappolt and N. Stone, Receivers for signal acquisition, *Microwave J.*, **20**, 35 (January 1977).
21. J. C. Harp, Receiver performance: What does sensitivity really mean, *Defense Electronics Technology*, 96 (May 1978).
22. R. I. Buick, Tangential sensitivity of a microwave receiver, *Microwave J.*, **11**, 33 (September 1968).
23. C. P. Heinzman and D. D. Tang, Broadband crystal video detector, *Microwave J.*, **8**, 93 (1965).
24. H. B. Dwight, *Tables of Integrals and Other Mathematical Data*, 4th ed., MacMillan Co., New York, 1961.
25. R. C. Emerson, First probability densities for receiver with square law detectors, *J. Applied Physics*, **24**(9), 1169–1176 (September 1953).
26. D. K. Barton, *Radars, Vol. 2, The Radar Equation*, Artech House, Inc., Dedham, VA, 1974.
27. D. P. Meyer and H. A. Mayer, *Radar Target Detection Handbook of Theory and Practice*, Academic Press, New York, 1973.
28. G. H. Robertson, Operating characteristics for a linear detector of CW signals in narrow-band Gaussian noise, *Bell System Technical J.*, **46**, 755–774 (1967).
29. W. J. Alberhseim, A closed-form approximation to Robertson's detection characteristics, *Proc. IEEE*, **69**, 839 (1981).

30. D. W. Tufts and A. J. Cann, On Albersheim's detection equation, *IEEE Trans. Aerospace Electronic System*, **AES-19**, 643–645 (1983).
31. J. V. DeFranco and W. L. Rubin, *Radar Detection*, Prentice Hall, Englewood Cliffs, NJ, 1968.
32. S. O. Rice, Mathematical analysis of random noise, *Bell System Technical J.*, **23**, 282–332 (1944); **24**, 46–156 (1945).
33. J. E. Marcum, A statistical theory of target detection by pulsed radar, mathematical appendix," *IRE Trans. Information Theory*, **IT-6**, 145–267 (1960).
34. J. Tsui and R. Shaw, Operational sensitivity of EW receivers, NAECON, p. 1234, May 18–20, Dayton OH, 1982.
35. J. Tsui and R. Shaw, Sensitivity of EW receivers, *Microwave J.*, **25**, 115 (November 1982).
36. J. C. Harp, What does receiver sensitivity mean? *Microwave System News*, **8**, 54 (July 1978).
37. Noise figure primer, Hewlett-Packard Co., Application Note No. 57, January 1965.
38. Solid state amplifiers, Watkins-Johnson Company, June 1979.
39. Solid state microwave amplifiers, Aertech Industries Catalog No. 5978, Sunnyvale, CA.
40. Microwave component amplications—designing with modular amplifiers, Avantek, Santa Clara, CA, 1976.
41. F. March, C. Wheeler, and R. Disman, Microwave transistor amplifiers, Watkins-Johnson Co. Application Note 100378A, June 1973.
42. Designing with GPD amplifiers, ATP-1002, Avantek, Santa Clara, CA, June 1972.
43. Thin-film cascadable amplifiers, Watkins Johnson Co., Palo Alto, CA, 1977.
44. D. L. Cheadle, Cascadable amplifiers, Watkins-Johnson Co. Tech Notes, **6**(1) (January/February 1979).
45. RF and microwave components, Anzac, Div. of Adams-Russell Co., Inc., Burlington, MA, 1984.
46. S. Fast, Advanced receiver modeling methods, AFAL-TR-76-265, Air Force Wright Aeronautical Laboratories, WPAFB, OH, 1976.
47. D. E. Norton, The cascading of high dynamic range amplifiers, *Microwave J.*, **16**, 57 (June 1973).
48. R. K. McDowell, High dynamic range receiver parameters, Watkins-Johnson Tech Note, 7(2) (March/April 1980).
49. S. Ploss, System Research Laboratories, private communication.
50. B. R. Hatcher, Intercept probability and intercept time, *Electronic Warfare*, **8**, 95 (March/April 1976).
51. B. R. Hatcher, EW acquisition systems, probability of intercept and intercept time, Watkins-Johnson Tech Note, **3**(3) (May/June 1976).

Chapter 3

Parameters Measured by EW Receivers

3.1. INTRODUCTION

This chapter discusses the quantities that an EW microwave receiver measures. The digital words a receiver generates and passes to a digital processor usually include five quantities. They are

frequency,
pulse amplitude (PA),
pulse width (PW),
time of arrival (TOA), and
angle of arrival (AOA),

as shown in Figure 3.1.

Except for the frequency measurements, all these quantities are measured with similar schemes, even in different kinds of receivers. Therefore, they will be discussed in this chapter.

The frequency information is very important for both sorting and jamming. By comparing the frequency of the pulses received, pulse trains of various radars can be sorted out. Knowing the frequency of the victim radar, the jammer can concentrate its energy in the desired frequency range. The theoretical aspect of frequency measurements will be discussed here; the actual measurement schemes will be discussed separately, in the chapters devoted to the particular receiver.

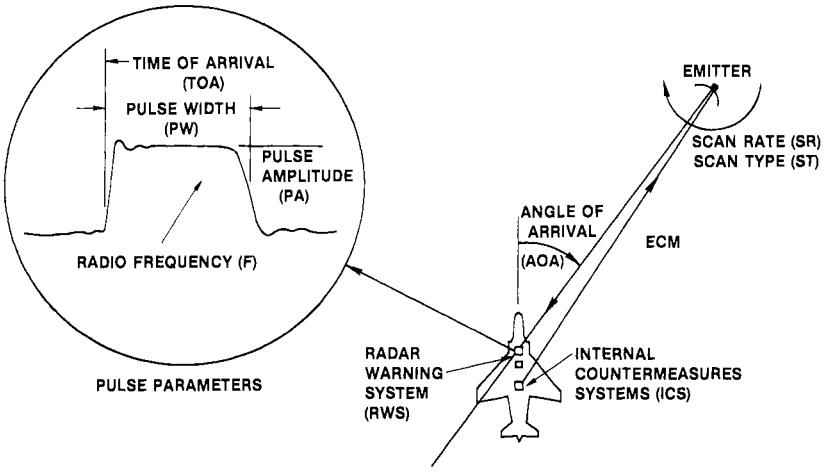


Figure 3.1. Information in an input radar pulse.

PA can be used to generate the scan pattern of some radars. Theoretically, the PA information may be used to estimate the distance of a radar, but the accuracy of the distance measured through PA is quite poor. Accurate PA information may be helpful in generating some effective jamming schemes, that is, inverse gain jamming, where the jammer power is inversely proportional to the PA measured.

PW can be used to provide coarse information on the type of radars. For example, generally speaking, weapon radars have short pulses.

TOA can be used to generate the pulse repetition frequency (PRF) of a radar. It can also be used to predict the TOA of the next pulse for a jammer.

Strictly speaking, the AOA information can be divided into two categories: the elevation and azimuth. Since the AOA information is difficult to measure, an EW receiver will usually measure the AOA only in the azimuthal direction.

Among these parameters, AOA is the most difficult and expensive to measure. It often requires several receivers matched either in amplitude or in phase for measurement. However, AOA is the most valuable parameter because a hostile radar can change the first four parameters, but it cannot easily change the AOA information received by an EW receiver.

Sometimes the polarizations of the input signals are also measured. This information can be used to select jamming schemes, that is, cross-polarization where the polarization of the jamming signal is 90° with respect to the input signal. This parameter is commonly measured by using antennas of different polarizations. By comparing the output amplitudes of these antennas, the polarization of the input signals can be obtained. Since the polarization measurement is quite similar to the amplitude AOA measurement, it will not be discussed in this chapter.

Another parameter of interest in EW receiver measurements is the range to

the emitter. This information can be used in many applications, for example, for threat avoidance and weapon delivery. The range measurement by an EW receiver is often referred to as passive ranging because there is no radiation involved as in radar. The most effective passive ranging approach is through accurate AOA measurement and a relative long base line to calculate the distance to the emitter. Since the AOA measurement will be discussed, passive ranging will not be discussed.

3.2. RESOLUTION, ACCURACY, AND ROOT MEAN SQUARE VALUES

Since the outputs of the receivers discussed in this book are in digital format, the least significant (LSB) of the output also represents the resolution of the parameter measured. The resolution does not represent the accuracy of the measurement, but it is related to it. The accuracy seldom matches the resolution. Often the accuracy is many times the resolution. For example, if a receiver has a frequency resolution of 1 MHz, the frequency accuracy of the receiver will be greater than or at best equal to 1 MHz. The accuracy could even be 10 MHz.

Only when the quantity to be measured is known, can the accuracy of the receiver measurement be determined. The accuracy is determined by the difference between the measured value and the input value. Therefore, in some measurements, a calibration signal is required to determine the accuracy of the measurement. In order to improve the accuracy of a measurement, a number of measurements are often required. Let us use the frequency measurement as an example. (This approach also applies to other parameter measurements.) If N measurements are made on the same input signal, each measured value is represented by f_i, the average frequency is given by

$$f_a = \frac{\sum_{i=1}^{N} f_i}{N} \tag{3.1}$$

where i is an integer $= 1, 2, \ldots, N$. The average value should equal the input frequency. If the average value does not equal the input frequency, the measurement is often referred to as biased. Sometimes a correction factor can be added to adjust the bias.

The difference between each individual measurement and the average can be used to determine the quality of the measurement. Assume the difference frequency to be

$$\Delta f_i = f_i - f_a \tag{3.2}$$

A root mean square (rms) value can be derived as

$$\Delta f_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\Delta f_i)^2} \tag{3.3}$$

If f_a in Eq. (3.2) is replaced by the input frequency, the rms is called the standard deviation. A small rms value means that the measurements made by the receiver are consistent. A large rms value represents the opposite. The causes of the large rms value can be divided into two categories. First, the quantity f_i has a wide variation. Second, most of the f_i values are very close together, but once in awhile an f_i is far from the rest of the values. Although both situations are undesirable, the latter cause is worse, because a few frequencies generated outside a group of measurements with frequencies close together may confuse the signal processor following the receiver.

3.3. ACCURACY OF FREQUENCY MEASUREMENTS (1)

The theoretical limit of the frequency measurement is discussed in this section. The rms frequency error (Δf_{rms}) of a perfectly rectangular pulse modulated by an RF wave is given (ref. 1) as

$$\Delta f_{rms} = \frac{\sqrt{3}}{\pi\tau(2E/N_0)^{1/2}} \tag{3.4}$$

where τ is the pulse width, E is the signal energy related to the signal power S through

$$S = E/\tau \tag{3.5}$$

and N is the noise power per unit bandwidth, which is related to the total noise N and bandwidth B through

$$N = N_0 B \tag{3.6}$$

Equation (3.4) shows that to obtain accurate frequency measurement, the PW must be long and the signal-to-noise ratio, must be high. The accuracy of frequency measurement relating to the PW is a very important factor, especially in receivers with narrow filters (i.e., channelized, compressive, and Bragg cell receivers).

If a receiver is designed to report frequency on a pulse-by-pulse basis, the rule of thumb is that the frequency accuracy, measured through a narrow-band system, is limited by the inverse of the PW. It is very difficult, if not impossible, to design a receiver with frequency accuracy better than the inverse of the PW. However, an IFM receiver can provide higher frequency accuracy than that measured by a narrow-band system because an IFM receiver does not contain narrow-band filters.

3.4. PULSE AMPLITUDE MEASUREMENTS (1)

If the input signal is pulse-modulated RF, it is often desirable to measure the PA. Theoretically, PA may be used to obtain a gross range of information on an emitter. However, the distance measured from this method is very inaccurate because a 6-dB difference will cause 100% error in distance. A change of the emitter antenna direction will provide a large change in the PA measured. PA can also be used as a sorting parameter to predict the scan pattern of a radar. A fine-grain amplitude may reveal some modulation pattern of the victim radar antennas and provide useful jamming information. If the amplitude of an RF-modulated system is A, the signal power will be

$$S = \tfrac{1}{2}A^2 \tag{3.7}$$

The rms value of the amplitude error is related to the noise power as

$$N = (\Delta A_{\text{rms}})^2 \tag{3.8}$$

From Equations (3.7) and (3.8), the relative error in PA measurement is (ref. 1)

$$\frac{\Delta A_{\text{rms}}}{A} = \frac{1}{(2S/N)^{1/2}} \tag{3.9}$$

Equation (3.9) clearly states that the stronger the signal-to-noise ratio, the more accurate the PA measurement.

PA is usually measured in decibels for two reasons. First, it can cover a wide dynamic range. If the amplitude is measured in a linear scale, it is very difficult to cover the entire dynamic range of the receiver. For example, a 60-dB dynamic range means a voltage change of three orders, which is hard to measure on a linear scale. The second reason is that the difference in decibels is equivalent to the ratio of the two signals. This difference can provide AOA information on the input signals in an amplitude comparison system, which will be discussed in Section 3.10.

The PA should be measured at the peak of the pulse, and special techniques must be used to locate the peak. The leading edge of the pulse may have different rise time, which depends on the characteristics of the radars. It can range from a few nanoseconds to microseconds. The trailing edge of the pulse also changes from a few nanoseconds to microseconds. Therefore, the peak of the pulse is around the center portion of the pulse. Sometimes PA is also affected by multipath. Multipath means that the signal reaches the receiver through different paths, i.e., is reflected from some objects. The reflected signals will add in or out of phase to the direct signal depending on the path difference. This phenomenon will affect the PA. Multipath also affects the PW of the input signals because the path length of the reflected signals may be quite different from that of the direct beam, which will contribute to a very slow decay time.

Peak detector circuits can be built in different forms. They fall into two classes: open loop and closed loop. These classes are illustrated in Figure 3.2.

The open-loop type depends solely on the diode characteristics. This type of peak detector is not very linear. Its performance is PW dependent. The rise time suffers at the expense of the long storage time of the diode.

The closed-loop type offers greatly improved linearity and independence from PW over the results obtainable with the open-loop peak detector. The problem with the closed-loop type is that it is very difficult to simultaneously obtain fast response time (i.e., 0.1 μsec settling time) and loop stability. Another problem with the peak detector approach is that any overshoot on the video signal may provide an erroneous peak value.

In the peak detector approaches, an electronic switch must be added to the circuits. The switch is controlled by another circuit that senses the trailing edge of the pulse. The charge on the holding capacitors in Figure 3.2 will be discharged at the trailing edge of the pulse so that the peak detector is ready to process a second signal. This additional circuit will add complexity to the peak detector approaches.

Another approach to measuring PA is using an analog-to-digital (A/D) converter to digitize the video output. The A/D converter digitizes (or samples) the output video pulse at a constant rate. The sampled value is compared with the previous one. If the previous value is stronger than the present one, the PA is represented by the previous value. Figure 3.3 shows a typical A/D converter approach.

The A/D converter must be fast enough to ensure that the peak of the pulse

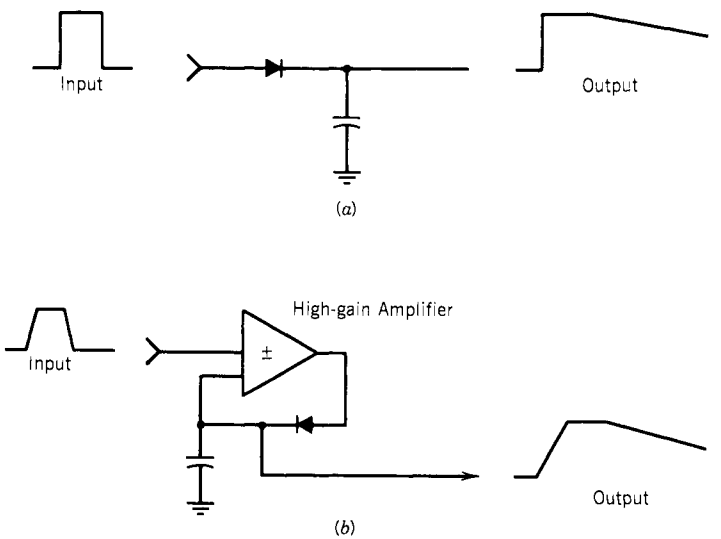


Figure 3.2. Classes of peak detector circuits: (*a*) open loop; (*b*) closed loop.

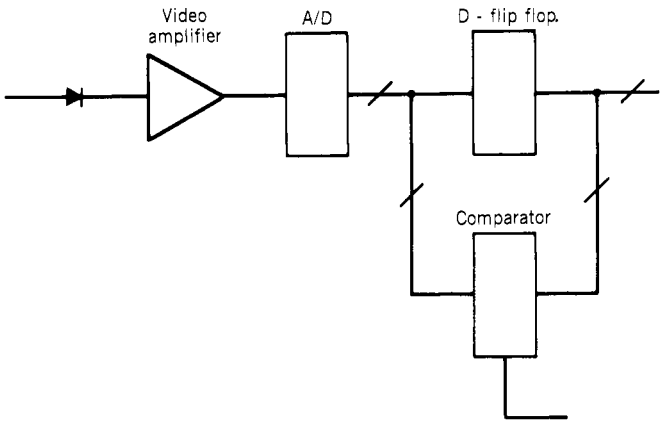


Figure 3.3. Peak detection circuit with an A/D converter.

will not be missed between adjacent samples. If the A/D converter samples at an overshoot of the video output, an erroneous peak value will be generated. Another problem with this approach is that the high-speed clock required to drive the A/D converter may add noise to the output pulse.

Another approach that works in conjunction with the PW measurement will be discussed in the next section.

3.5. MULTIPLE TRIGGERING PROBLEM (2–6)

In many EW receivers, there is only one simple threshold at the output of the video amplifier. When an input signal is above the threshold, the receiver will start to measure all the parameters on the pulse. When the input signal is below the threshold, the receiver will not respond. However, when an input signal is close to the threshold of the receiver, since there is noise associated with the input signal, the signal may trigger the receiver many times, as shown in Figure 3.4. As a result, the receiver will report many input signals with very close TOAs and short PWs. When this situation occurs, the other parameters measured are very close or may even be the same, since they are generated from the same signal. This is a very undesirable phenomenon, because the digital processor following the receiver must sort all of pulses measured.

This problem can be solved by introducing a double-threshold scheme. The circuit is shown in Figure 3.5*a*, and the two thresholds are shown in Figure 3.5*b*. The input signal must cross the upper threshold to be declared a legible input and trigger the receiver to start the measurements of all parameters. The signal must drop below the lower threshold to be declared the end of the signal.

The separation between the two thresholds can be arbitrarily chosen. If the separation between the two thresholds is too small, the multiple triggering

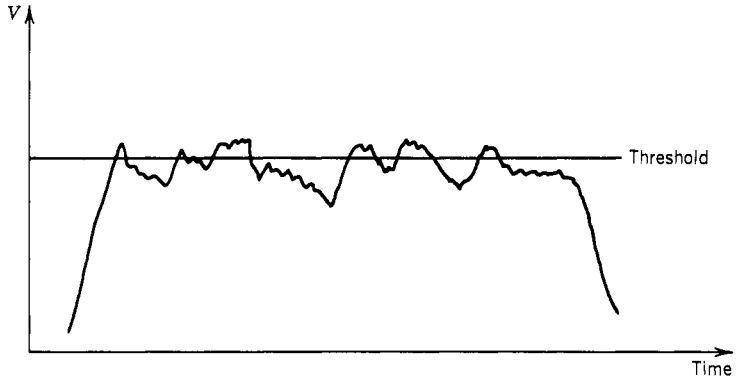
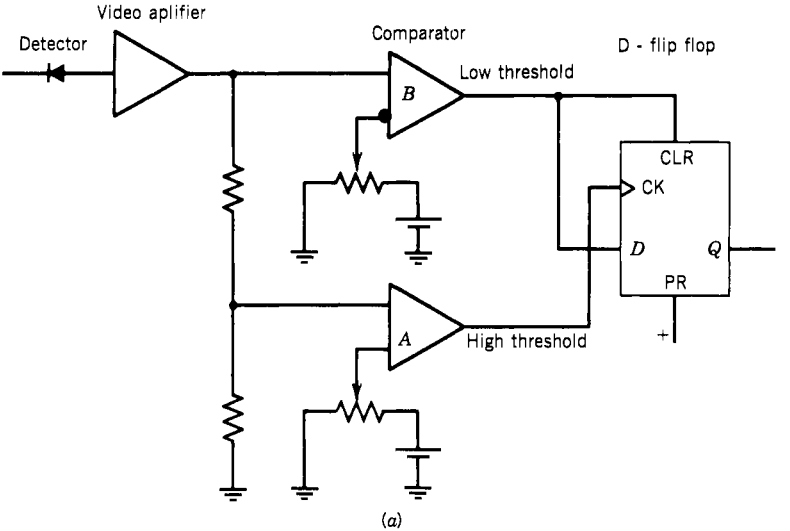
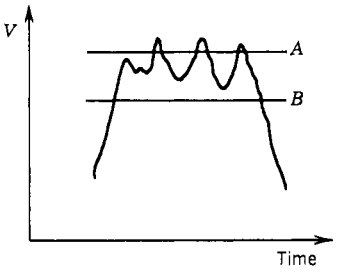


Figure 3.4. Signal close to threshold.



(*a*)



(*b*)

Figure 3.5. Dual-threshold detection circuit: (*a*) circuit; (*b*) thresholds.

problem will occur. If the separation is too large, the sensitivity of the receiver will suffer because the upper threshold determines the sensitivity of the receiver. In general, the separation is about 3 dB.

3.6. PULSE WIDTH MEASUREMENTS (1, 7, 8)

Pulse width is another important sorting parameter a receiver can measure. The PW can provide some gross information on the type of radar emitter. However, as discussed in Section 3.4, PW is affected by the multipath problem. To measure the PW, there are two major concerns: The first is how to define the PW and second, the bandwidth of the receiver must be wide enough to preserve the pulse shape. Since the pulse amplitude is measured in dBm, the PW is often defined at points 3 dB below the peak of the pulse. The minimum video bandwidth required to properly process the signal can be approximated by (ref. 7)

$$B_V = 0.35/t_r \tag{3.10}$$

where t_r is the rise time of the pulse, often defined as the time from 10 to 90% of the PA on the leading edge. In Equation (3.10), only the video bandwidth is considered as the limiting bandwidth, since the RF bandwidth is usually wider than the video bandwidth. The bandwidth problem will be further discussed in Section 4.5.

The time delay rms error measured on the leading edge of the pulse can be expressed as (ref. 1)

$$\Delta T_{rms} = t_r/(2S/N)^{1/2} \tag{3.11}$$

Equation (3.11) shows that the steeper the leading edge and the higher the signal-to-noise ratio, the better the accuracy of the measurement on the leading edge. The same argument holds true for the trailing edge of the pulse. Intuitively, one can draw the same conclusion, that is, the sharper the leading and trailing edges, the better the pulse shape is defined and the easier it is to make the measurement.

The simplest way to measure the PW is to compare the PA with a fixed threshold as shown in Figure 3.6. The problem with this circuit is that the PW measured depends on the signal strength. Therefore, the desirable PW may not be obtained.

Another approach to measure the PW is to use a peak detector to hold the peak value of the input pulse. Then this peak value is negatively offset by an amount equal to 3 dB, and the peak value compared with a delayed version of the video signal as shown in Figure 3.7. The output is a logic-level signal representing the desired PW. The difficulty with this approach is in building the peak detector, as discussed in Section 3.4.

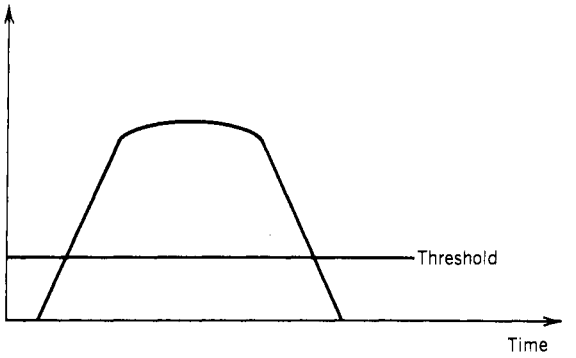


Figure 3.6. Simple scheme to measure pulse width.

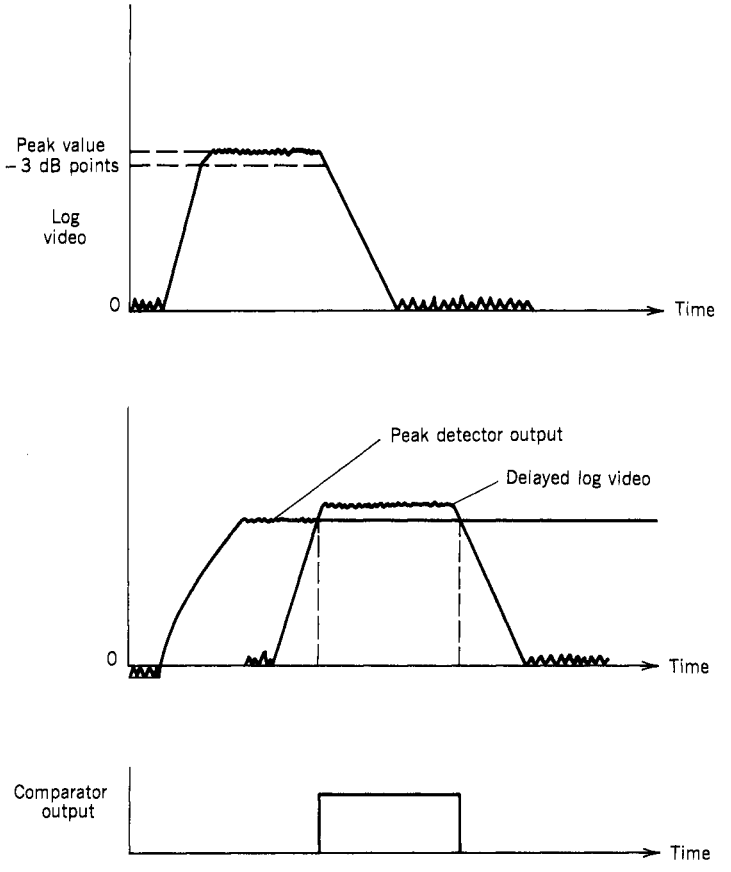


Figure 3.7. Peak and delayed signal comparison scheme to measure pulse width.

To avoid the problems of peak detectors, delay lines and comparators are used to determine the -3 dB of the video output. Figure 3.8 shows the schematic of this approach. The basic idea is to generate a set of successively delayed video outputs. The delay time between adjacent videos is less than the minimum pulse width expected. In the special example shown in Figure 3.8, five comparators and four delay lines are used. Video output 1 (the top one) is not delayed,

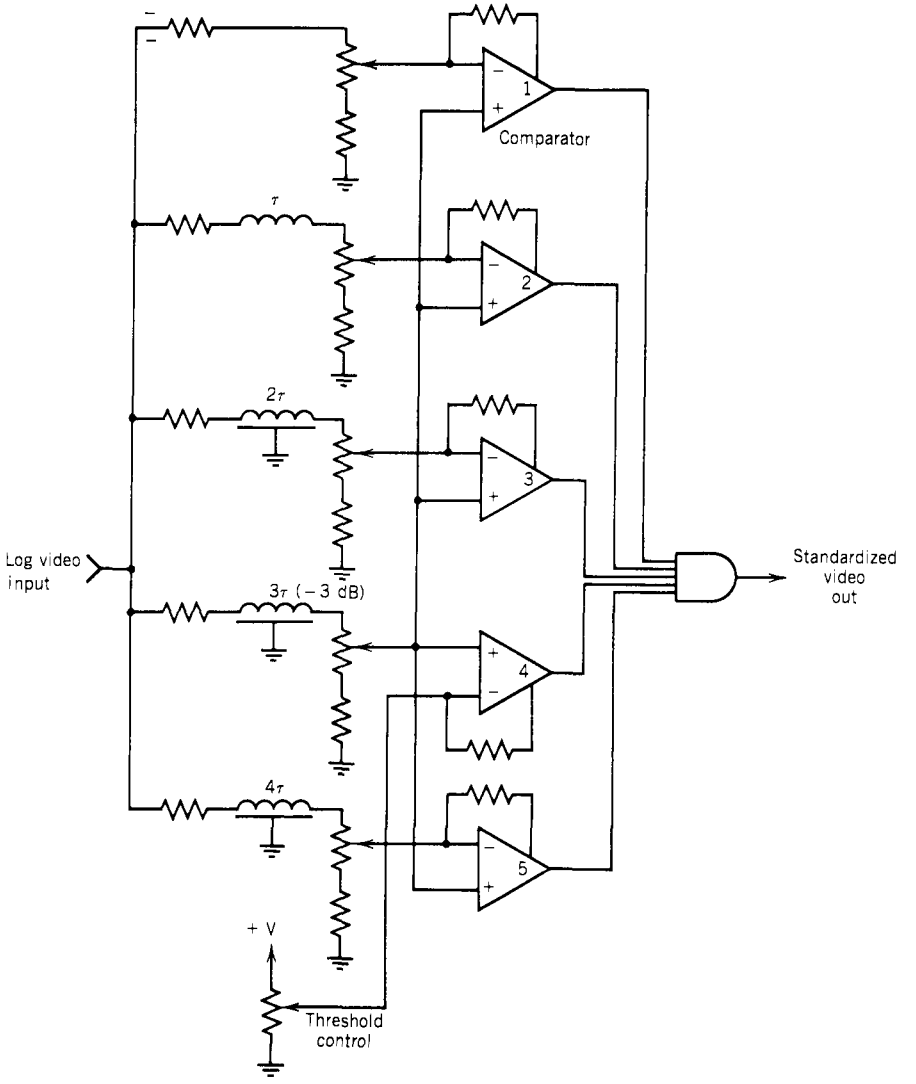


Figure 3.8. Delay lines and comparator circuits to measure PW.

and the other four outputs are delayed successively by τ. One of these videos (input of comparator 4) is offset by 3 dB and compare against all other videos. Comparators 1, 2, 3, and 5 compare the individual input with the input of 4, and 4 compares the input video with a threshold. The output of all the outputs are ANDed together. The result is a logic signal that enters the "true" state when the reference video pulse reaches the -3 dB on the leading edge and leaves the true state when the video falls below the -3 dB on the trailing edge. Figure 3.9*a* illustrates the conditions of a pulse with fast rise time (fast relative to the delay increment). Figure 3.9*b* illustrates the conditions of a pulse with a low rise time. In Figures 3.9*a* and 3.9*b*, the output is a logic signal (referred to as standard video pulse), and its width represents the PW.

In the above PW measurement scheme, the standard video pulse generated is under the peak of the pulse. The leading and trailing edges are at -3 dB. Therefore, if the center of the standard video pulse is used as a reference time to sample the video output, the PA can be obtained. Since one cannot predict the

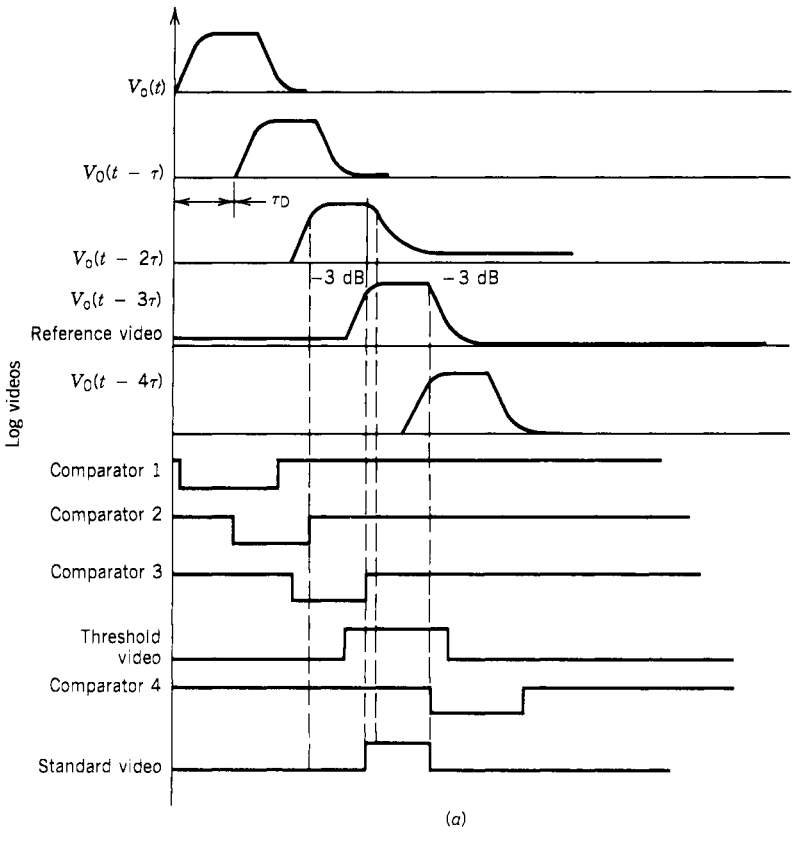


Figure 3.9. PW waveforms and measured results: (*a*) PW with fast rise time; (*b*) PW with slow rise time.

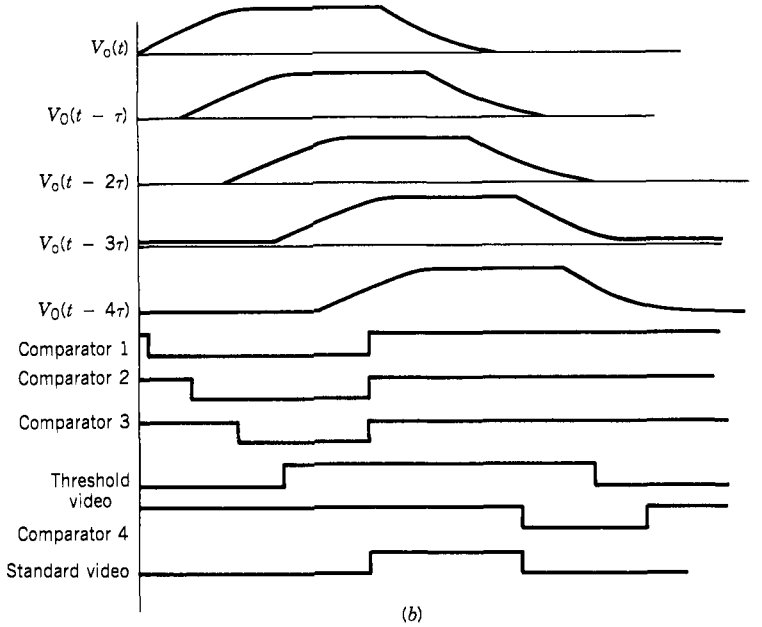


Figure 3.9. (Continued).

center of the standard video pulse, a sampling window is often chosen after a fixed time delay from the leading edge of the standard video pulse. The PA measured by this approach is quite accurate.

Finally, the bit pattern used to represent the PW will be discussed. The PW can range from a few hundred nanoseconds to continuous wave (CW). It is impractical to wait until the end of a CW signal to process it. In general, one can predetermine a maximum PW, and any signal with a PW longer than this designated value is declared a CW signal. In most receivers, the maximum PW is programable. Since the PW covers such a wide range, it is desirable to report the PW in a logarithmic scale. In other words, a short pulse is measured with fine time resolution and a long pulse is measured with coarse time resolution. For example, a pulse of a few hundred nanoseconds can be measured with 50-nsec resolution, whereas a pulse of tens of microseconds can be measured with 1 μsec resolution.

3.7. TIME OF ARRIVAL MEASUREMENTS

The time of arrival (TOA) measurement provides a time reference to all the received pulses. A free-running clock can be used to time the leading edge of the pulse. The TOA information can be used for two purposes. First, TOA can be used as a sorting parameter to generate pulse repetition frequency (PRF) information. Sometimes the term *pulse repetition interval* (PRI), the inverse of

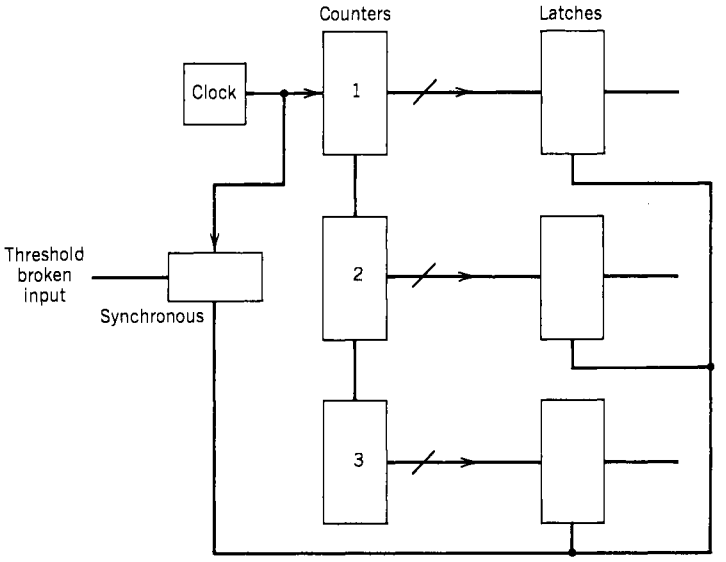


Figure 3.10. TOA-generating circuit.

PRF, is used instead of PRF. The TOA is a very useful sorting parameter, especially against radar with stable PRFs. The second application of TOA is to provide a jammer with accurate time information.

The basic measurement of TOA is the same as the PW measurement; therefore, the accuracy of TOA measurement is also governed by Eq. (3.11). Since the TOA is measured by a free-running clock, its absolute value is not very important, but the difference between the TOA (also referred to as ΔTOA) is very useful. To measure the scan pattern of a radar, the overall clock period is in general longer than the scan time of the radar. The radar scan time is usually on the order of seconds. Therefore, the overall clock period usually lasts a few seconds before it laps around. For example, if the TOA resolution is 1 μsec, a total of 24 TOA bits will be 16.8 sec (2^{24} μsec). Usually in a receiver, the TOA output occupies more bits than the other measured parameters.

Figure 3.10 shows a typical approach of generating the TOA information. On the left is a clock that triggers three counters. The synchronizer takes the threshold broken signal as an input and generates a pulse synchronous with the clock. This synchronizing pulse in turn triggers the three latches and registers the TOA of the input signal.

3.8. ANGLE OF ARRIVAL MEASUREMENTS (9–16)

In an EW system, the angle of arrival (AOA, sometimes referred to as direction of arrival, DOA) information is extremely important. Since the AOA is obtained from the position of the emitter, this is the only parameter a hostile

emitter cannot change easily. Thus, the AOA becomes the most reliable sorting parameter in an EW receiving system, especially when the hostile radar intentionally varies its RF and PRF.

Seven ways to measure AOA will be discussed:

1. The simplest approach is to use a narrow-beam antenna to search the direction of interest. However, this is not a very popular approach for EW applications.
2. AOA is measured through amplitude comparison.
3. AOA is measured through phase comparison.

Approaches 2 and 3 are the most common in EW applications. In both approaches, a multiple number of antennas and receivers are needed. Besides, the antennas and the receivers must be matched in amplitude or phase. Thus, the AOA information becomes the most costly parameter to be measured.

4. AOA is measured through Doppler frequency shift.
5. AOA is measured through TOA difference.
6. AOA is measured through a microwave lens
7. AOA is measured through multiple-beam arrays and beam-forming networks.

There are many sophisticated approaches to obtain the AOA information, especially in radar applications. However, many of the approaches are too complicated to be adapted in an EW receiving system.

3.9. AOA MEASURED BY NARROW-BEAM ANTENNA AND SIDE LOBE CANCELLATION (1, 11, 17, 18)

The easiest approach to measure AOA is to use a narrow-beam antenna and search the direction of interest. When the antenna is pointing in the direction of an emitter, the signal at the receiver is the strongest. Theoretically, by monitoring the signal strength while the antenna is searching, the AOA can be obtained. In practice, there are many side lobes in the antenna pattern in addition to the main lobe. It is difficult to differentiate a signal in the direction of a side lobe from that of the main lobe. One way to solve this problem is to use two antennas and two receivers. One receiver is connected to the narrow-beam antenna and the other is connected to an omnidirectional antenna as shown in Figure 3.11. When the output from the narrow-beam antenna is stronger than that of the omnidirectional antenna, the signal is in the direction of the main lobe of the narrow-beam antenna. Otherwise, the signal is from one of the side lobes, and the AOA of the signal is not known. This scheme is referred to as side lobe cancellation.

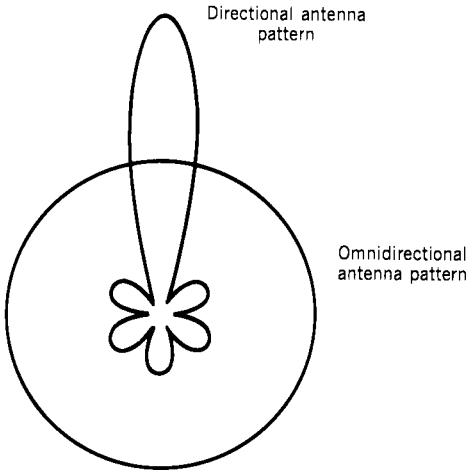


Figure 3.11. Side lobe cancellation AOA scheme.

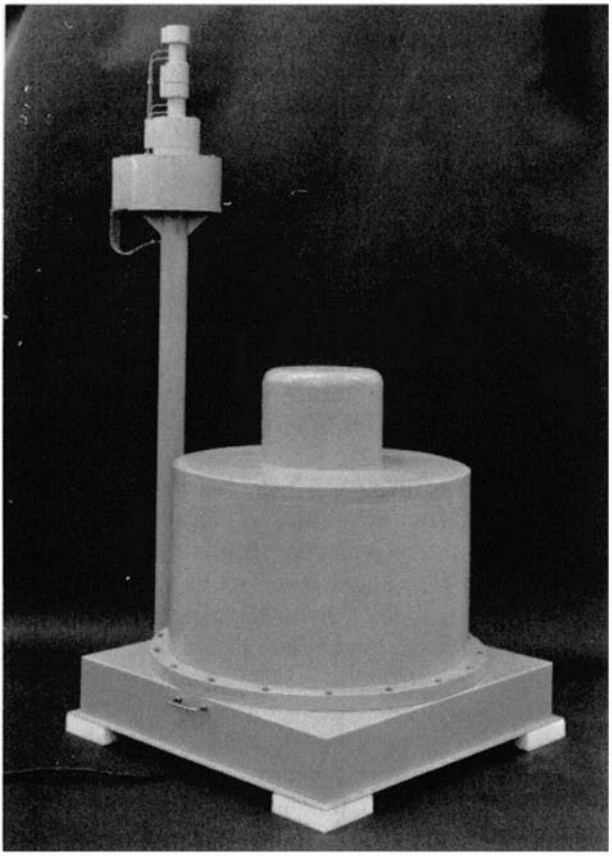


Figure 3.12. Side lobe cancellation system. (Courtesy of Watkins-Johnson Co.)

The accuracy of the AOA is related to the antenna beamwidth and the signal strength. It can be expressed as (ref. 1)

$$\delta\theta = \frac{\sqrt{3}\lambda}{\pi D(2E/N_0)^{1/2}} \tag{3.12}$$

where $\delta\theta$ is the angular error, λ is the wavelength, D is the linear dimension of the antenna measured in the plane of angle θ, E is the signal energy, and N_0 is the noise power per unit bandwidth as shown is Eqs. (3.5) and (3.6). Equation (3.12) shows that in order to have high AOA accuracy, the antenna must be very long.

Figure 3.12 shows a side lobe cancellation system. The taller antenna is omnidirectional. The radome of the scanning antenna is shown at the center of the figure.

This AOA measurement approach is not suitable for EW applications in general. In an EW system, it is desirable to cover a wide angle of interest instantaneously. But in this approach, the probability of intercept (POI) is low since the antenna has to search the direction of interest sequentially. This approach cannot process simultaneous signals under certain conditions. For example, if a weak signal is the main beam of the directional antenna and a strong signal is outside of its main beam, the output of the omnidirectional antenna may be stronger than that of the directional antenna and may mask the signal in the main beam. The other deficiency of this approach is that the antenna is usually too big to be used in an EW system. This is especially true for airborne applications.

3.10. AOA MEASURED BY AMPLITUDE COMPARISON (10, 19)

Amplitude comparison is very simple. It uses a multiple directional antenna–receiver system, and each antenna is pointing toward a different direction. The antenna that points the nearest to the direction of the emitter will generate the strongest output from the receiver connected to this antenna. By comparing the different amplitudes of these outputs, the AOA information can be obtained.

A four-quadrant amplitude comparison direction finding (DF) system is shown in Figure 3.13. Four antennas point in four directions. This system can be frequency independent, which means it may cover a wide frequency range. The actual frequency bandwidth of this system depends on the antennas, receivers, and the amplitude tracking among them. Wider bandwidths have high probability of intercept. Unless the receivers have the capability to separate signals by frequency, a wide-bandwidth AOA system suffers a high probability of being contaminated by signals arriving at the antennas simultaneously.

Assume that the amplitudes A_t from the four antennas can be written as

$$A_t = A \exp[-K^2(\theta - \alpha)^2] \tag{3.13}$$

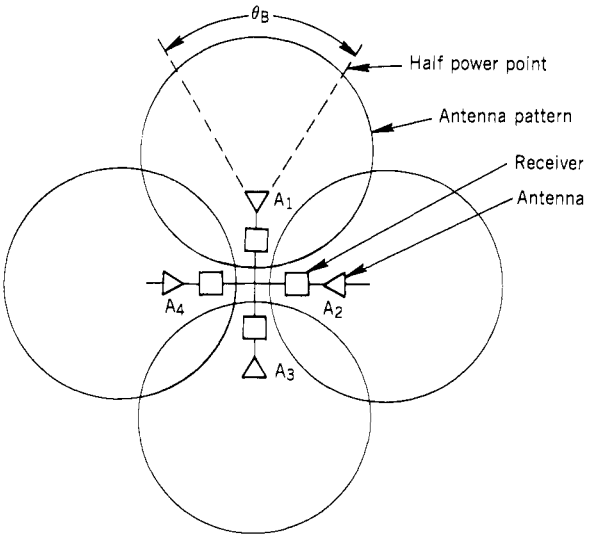


Figure 3.13. A four-quadrant amplitude AOA system.

where A is the relative amplitude of antenna peak power, θ is the bearing angle measured from the antenna axis, α is the beam offset angle, and

$$k^2 = 2.776/\theta_B^2 \tag{3.14}$$

where θ_B is the antenna beamwidth measured between half-power points. The constant 2.776 comes from the relation that when $\alpha = 0$ and $\theta = \frac{1}{2}\theta_B$, the amplitude from the antenna gain is $0.5A$. For the system in Figure 3.13,

$$\alpha_1 = 0, \quad \alpha_2 = 90°, \quad \alpha_3 = 180°, \quad \text{and} \quad \alpha_4 = 270°$$

the AOA information is obtained from the ratio of the amplitudes of two adjacent antennas (i.e., through antennas 1 and 2):

$$R = \frac{A_1 \exp(-k^2\theta^2)}{A_2 \exp[-k^2(\theta - \alpha_2)^2]}$$

$$= \frac{A_1}{A_2} \exp(-2k^2\alpha_2\theta + k^2\alpha_2^2) \tag{3.15}$$

where $\alpha_1 = 0$.

The R ratio can be expressed in decibels as

$$R\,(\text{dB}) = 10\log(A_1/A_2) - 8.68k^2\alpha_2\theta + 4.34k^2\alpha_2^2 \tag{3.16}$$

The error slope is, by definition, the rate of R with respect to θ; therefore, differentiation of Eq. (3.16) gives the error slope as

$$\frac{d}{d\theta}[R\,(\text{dB})] = -8.68k^2\alpha_2 \tag{3.17}$$

The error slope is a function of antenna beamwidth and the squint angle between the two antennas. It is independent of the incident angle θ. In other words, the error slope is constant for all θ provided the antenna pattern can be approximated by the exponential form. If the antenna amplitude imbalance is measured in decibels, then the angular error is

$$\Delta\theta = \frac{A_2/A_1}{8.68k^2\alpha_2} \tag{3.18}$$

Therefore, system accuracy is directly proportional to the amplitude imbalance between the two antennas and inversely proportional to the error slope.

In a practical system, neither the antennas nor receivers are perfectly balanced. Thus in Eq. (3.18), the amplitude imbalance should include the receiver

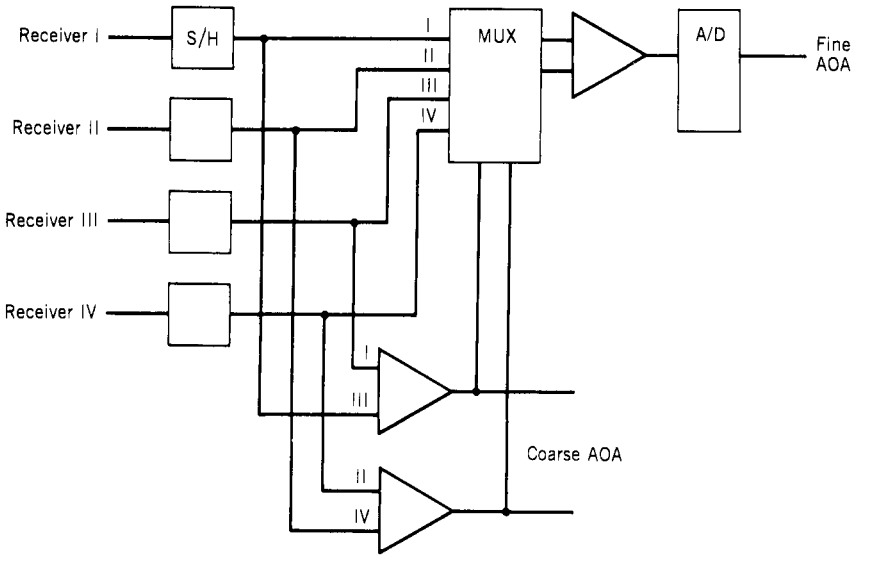


Figure 3.14. Comparison logic for a four-quadrant AOA system.

amplitude mismatch. Often this kind of system is calibrated against an input signal with known AOA information.

Figure 3.14 shows a typical four-quadrant amplitude comparison system with logic circuits. The amplitudes from four receiver outputs are stored in the four sample-and-hold (S/H) circuits. The amplitudes of two opposite antennas are compared first to determine which quadrant the signal comes from. This information provides the two most significant bits (MSB) of the AOA and also controls a multiplexer. The inputs of the multiplexer are from the outputs of the four receivers. The coarse AOA bits will select two adjacent antennas with the strongest outputs and send them to the following differential amplifier where the fine AOA information is determined. Since the receiver outputs are generally expressed in decibels, the differential amplifier, which amplifies the difference of two input signals in Figure 3.14, will provide the ratio of the two antenna outputs, which determines the AOA information. The A/D converter will change the AOA into digital information.

3.11. PHASE COMPARISON (INTERFEROMETRY) AOA SYSTEM (10, 20–22)

A simple two-element phase comparison system is shown in Figure 3.15. The two antennas are separated by distance d. If a wave is coming in at angle θ from the bore side, the phase ϕ can be expressed as

$$\phi = \frac{2\pi d \sin \theta}{\lambda} \tag{3.19}$$

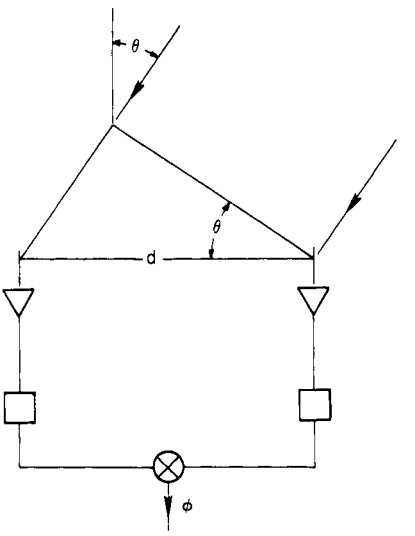


Figure 3.15. Two-element phase comparison AOA system.

where d is the distance between the two antennas, θ is the incident angle as shown in Figure 3.15, and λ is the wavelength of the incident wave. If the phase ϕ can be measured, the incident angle θ can be determined.

From Eq. (3.19), it is obvious that to calculate the AOA, the frequency (or wavelength) must be known. For a fixed ϕ value, the longer the distance d and the shorter the λ, the more accurate the angle θ can be measured. However, there is a maximum distance between the two antennas. Beyond this value, it will cause ambiguity in θ. The maximum phase shift ϕ must be within 2π to avoid the ambiguity. The maximum distance can be obtained by changing θ from $+90°$ to $-90°$ and keeping the change of ϕ within 2π.

$$\Delta\phi = 2\pi = \frac{2\pi d_{\max}[\sin 90° - \sin(-90°)]}{\lambda} \tag{3.20}$$

or

$$d_{\max} = \tfrac{1}{2}\lambda \tag{3.21}$$

The maximum distance between two antennas without causing ambiguity is half the input signal wavelength. If more accuracy is desirable without ambiguity, a multiple number of antennas with nonuniform spacings can be used as shown in Figure 3.16. Whereas the long distances between antennas provide the AOA accuracy, the short distances can resolve the ambiguity. The AOA error of a phase comparison system is contributed from three possible sources: the frequency, phase measurement error, and noise. The AOA error of a two-channel phase comparison system is discussed in references 10 and 20, and the results are presented here.

The contribution of phase errors to rms AOA error is

$$\varepsilon_\phi = \frac{\lambda\sigma_\phi}{\sqrt{2}\pi d} \tag{3.22}$$

where σ_ϕ is the rms channel phase tracking error.

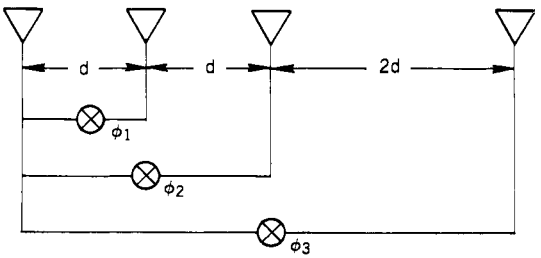


Figure 3.16. Multiple-element phase comparison AOA system.

The contribution of the direction finding (DF) error by frequency inaccuracy is

$$\varepsilon_f = \sin\theta \frac{\Delta f}{f} \tag{3.23}$$

where Δf is the rms frequency error.

The contribution of the thermal noise to rms AOA error is

$$\varepsilon_n = \frac{\lambda}{\pi d(2S/N)^{1/2}} \tag{3.24}$$

where S/N is the signal-to-noise ratio.

The overall rms AOA error can be written as

$$\varepsilon = \sqrt{\varepsilon_\phi^2 + \varepsilon_f^2 + \varepsilon_n^2} \tag{3.25}$$

In general, the AOA result obtained from a phase comparison system is more accurate than an amplitude comparison system.

Another ambiguity problem in a phase comparison system is shown in Figure 3.17. For a one-dimensional antenna array, if the two incident waves in Figure 3.17 have the same frequency and both have the same angle θ with respect to the bore side of the array, they will produce the same phase angle ϕ. The phase comparison AOA system cannot differentiate one incident wave from the other. To solve this problem, a two-dimensional array will be required.

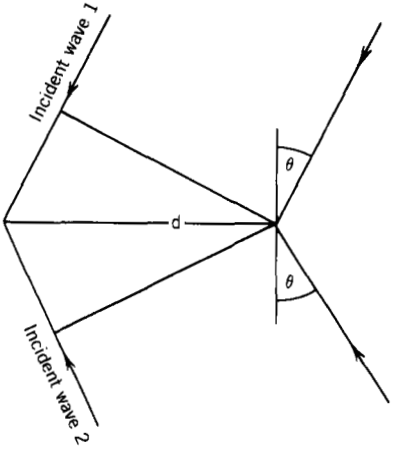


Figure 3.17. Ambiguity caused by waves from opposite sides of antenna.

3.12. AOA MEASUREMENT THROUGH DOPPLER FREQUENCY SHIFT (23–28)

If the antenna of a receiver is moving in a certain direction (say x) with velocity v, the frequency of the input signal will be changed by the Doppler frequency shift. The Doppler frequency shift depends on the direction of the emitter with respect to the line of motion as shown in Figure 3.18. If the distance between the emitter and receiver is D, this distance can be expressed in terms of phase by

$$\phi = 2\pi D/\lambda \tag{3.26}$$

The Doppler frequency f_d is generated by the time change of phase angle ϕ

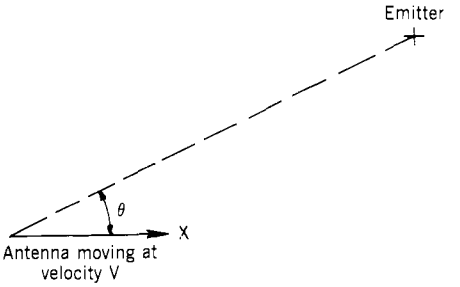


Figure 3.18. Doppler frequency shift induced by motion of a receiving antenna.

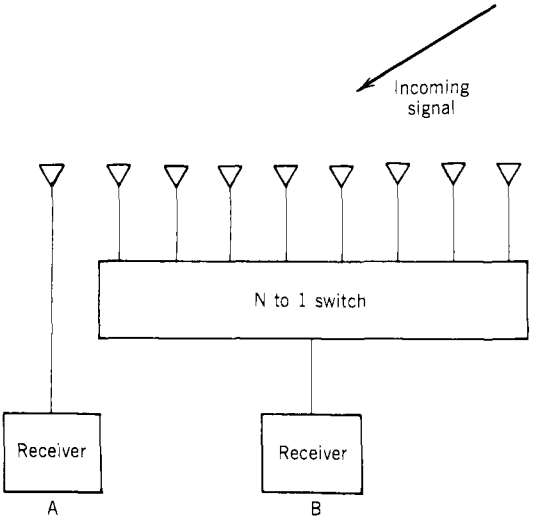


Figure 3.19. High-speed switching scheme to simulate an antenna motion.

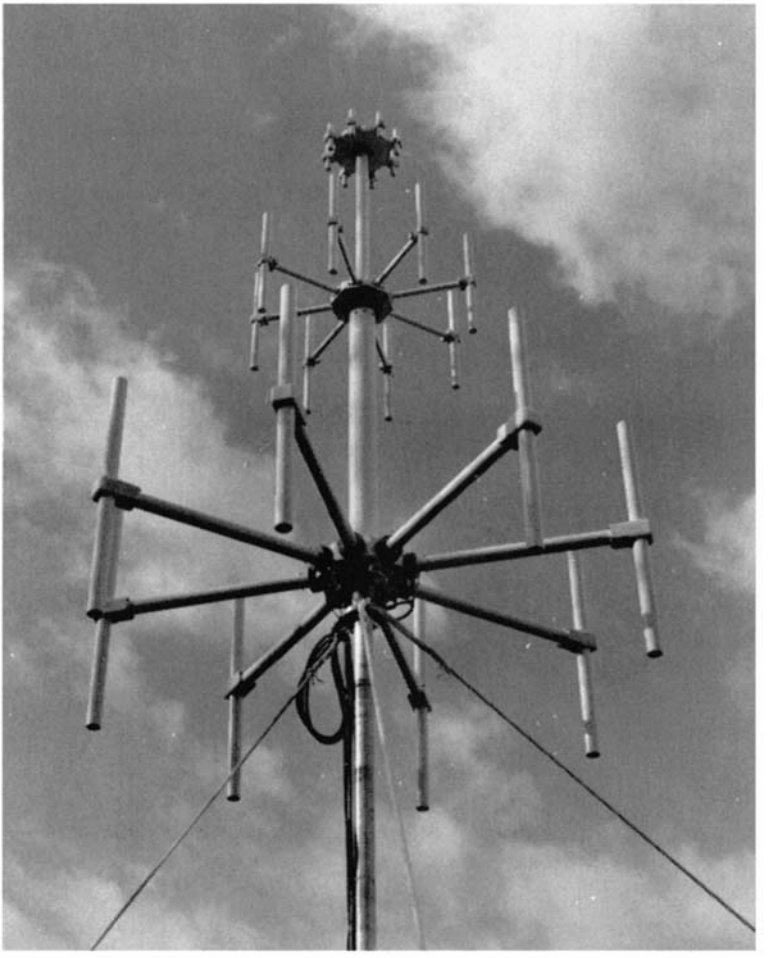


Figure 3.20. Circularly arranged switching antennas to obtain AOA information. (Courtesy of Northern Scientific Laboratories, Division of General Instrument Corp.)

along the direction D:

$$f_d = \frac{1}{2\pi}\frac{d\phi}{dt} = \frac{1}{\lambda} v \cos\theta = \frac{f_0 v}{C}\cos\theta \tag{3.27}$$

where v is the velocity of the moving antenna, f_0 is the frequency of the emitter, C is the velocity of light, and θ is the angle between v and line of sight D.

The signal frequency f_0 can be measured by another stationary receiver. If the velocity v is known, the AOA information can be obtained from Eq. (3.27).

It is impractical to physically move the antenna to generate any noticeable

Doppler frequency shift. However, the antenna motion can be stimulated by switching the receiver rapidly among an array of antennas as shown in Figure 3.19. Receiver A is connected to a fixed antenna. Receiver B is switching from antenna 1 to N sequentially through an N-to-1 switch to simulate a linear motion. The frequency difference between the two receivers can be used to determine the AOA information. The ambiguity shown in Figure 3.17 can also occur in this scheme. This approach has been applied to the VHF-to-UHF frequency range, and the antennas are arranged in a circle, as shown in Figure 3.20. In this arrangement, there is no stationary antenna, and the AOA information can be obtained with proper processing. With the advancements in microwave devices, this approach may be used at microwave frequencies.

3.13. AOA MEASUREMENT THROUGH DIFFERENTIAL TOA

If the TOA of a pulsed signal can be measured very accurately, it can be used to obtain AOA information. As shown in Figure 3.21, the time difference Δt for the pulsed signal to reach the two antennas is related to the incident angle θ by

$$\Delta t = (d/C)\cos\theta \tag{3.28}$$

where d is the distance between the two antennas and C is the speed of light.

As discussed in Figure 3.16, if the angle coverage is 360°, there is an ambiguity problem. Equation (3.28) shows that the longer the distance, the more accurate the measurement. Since the leading edge of a radar pulse has finite rise time, it is difficult to measure the TOA accurately. Therefore, this approach is not widely adapted for relatively short baseline systems.

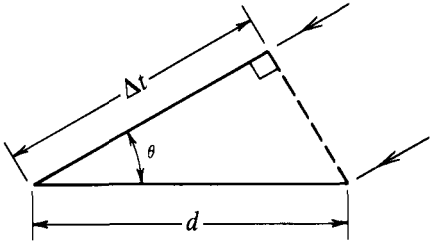


Figure 3.21. AOA measurement through TOA difference.

3.14. AOA MEASUREMENT BY MICROWAVE LENS (29–31)

The microwave lens discussed here is an analogy to an optic lens. It can be built in stripline or microstrip forms. The basic operating principle is illustrated in Figure 3.22. The heart of a microwave lens is a beam-forming network, a special shaped metal sheet sandwiched between two dielectric plates. Ground planes

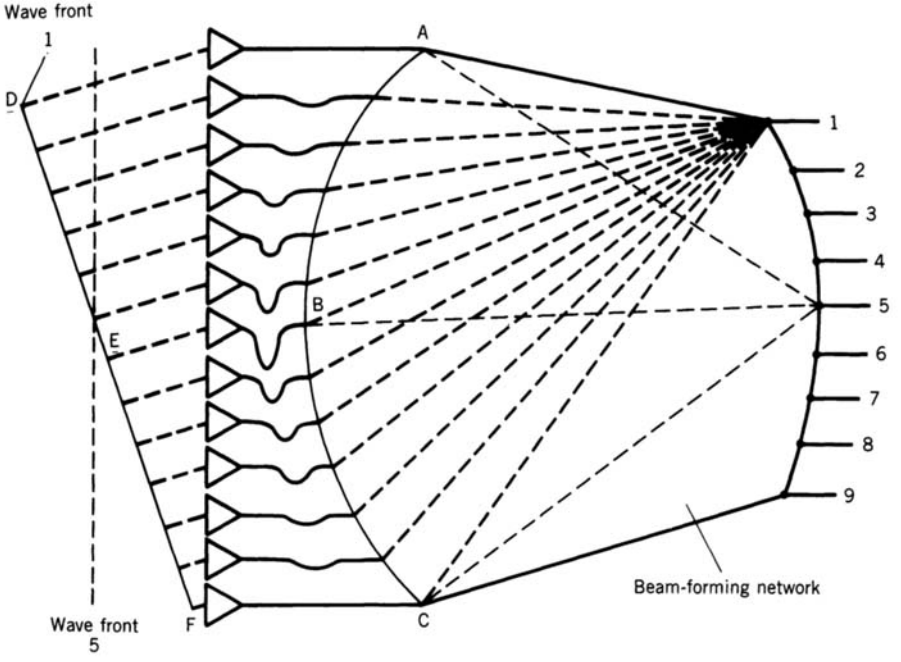


Figure 3.22. Basic configuration of a microwave lens.

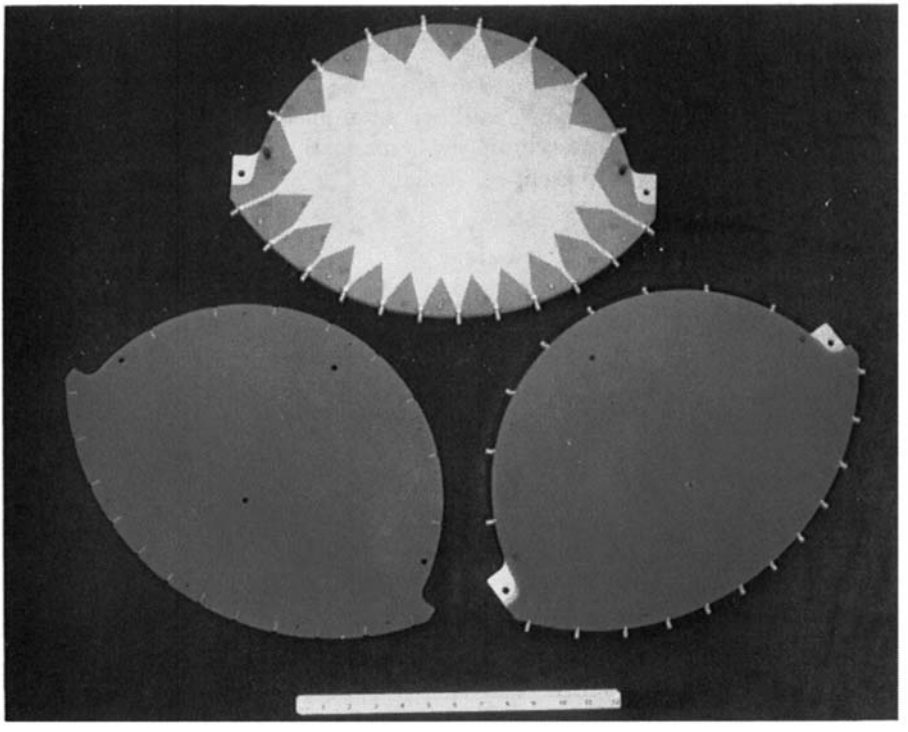


Figure 3.23. Microwave lens for AOA measurement. (Courtesy of Raytheon Co.)

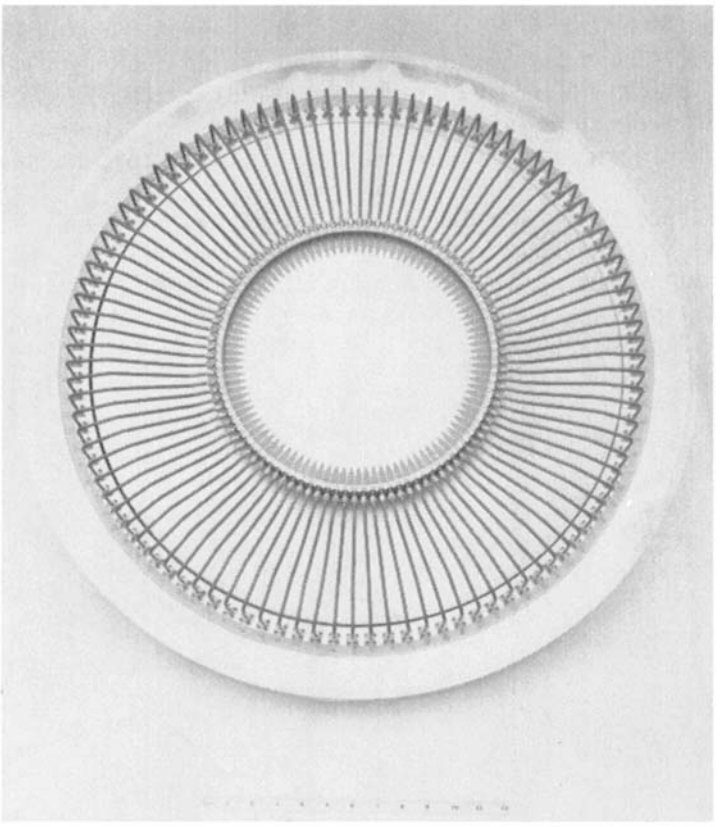


Figure 3.24. Circular microwave lens. (Courtesy of Raytheon Co.)

are added on the outside of the dielectric plates. Input antennas are connected on the left side of the beam-forming network. The outputs are connected to the right side of the beam-forming network. To simplify the discussion, assume that all the input antennas are aligned in a straight line and all the feeding coaxial cables are of the same length.

The beam-forming network will provide different delays from a certain input port to a certain output port, so that a wave coming from a certain direction will exert at a certain output. For example, an incident wave (labeled 5, corresponding to output port 5) with a wave front parallel to the plane of the antenna array will arrive at the lens arc ABC in equal phase. If the lens is designed with path length $A5 = B5 = C5$, the incident wave 5 will emerge from output port 5. If wave 1 is incident on the antenna array and the path length is designed as such that $DA1 = EB1 = FC1$, this wave will emerge from port 1. It should be noted that only three ports (A, B, and C) are considered here because in the lens design only three points on the input arc can be used to generate the design equations.

If more than three points are used, it will generate more equations than unknowns, which cannot be solved. A microwave lens is shown in Figure 3.23. Figure 3.24 shows a circular microwave lens. The incoming signal will arrive at the opposite side of the lens. If an isolator is used at each port to separate the input from the output, the lens can cover 360° azimuth. However, in this arrangement, a large number of receivers are required to cover the desired AOA range.

This approach is basically frequency independent, and thus the microwave lens is a wide-band system. Since input signals from different directions will reach different ports, this approach can measure AOA information on simultaneous signals and separate them according to their AOA. However, to process simultaneous signals with relatively fine AOA resolutions, many receivers are required.

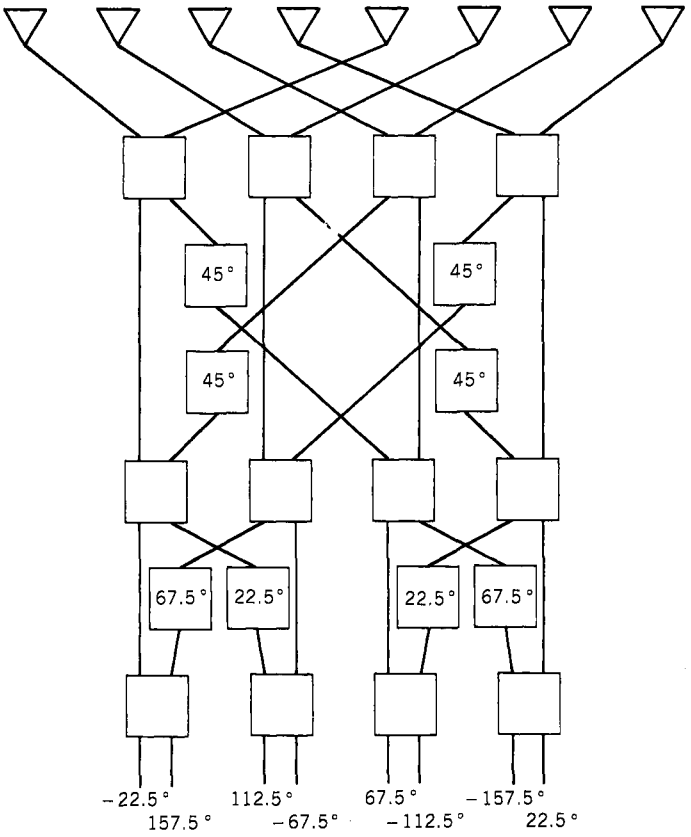


Figure 3.25. Eight-element multiple-beam AOA system. (Based on ref. 31.)

3.15. MULTIPLE-BEAM AOA SYSTEM (11, 32–34)

A multiple-beam AOA system is quite similar to a microwave lens. The system basically consists of three parts: the input antenna array, the beam-forming network, and the receivers. This system was first discussed by Bulter in 1960 (ref. 32), and the network was described by matrices; therefore, the multiple-beam AOA system is often referred to as a Bulter matrix.

Like a microwave lens, the multiple-beam system can measure the multiple-input signals. The difference between a multiple-beam AOA system and a microwave lens is in the beam-forming network. Whereas a microwave lens uses a continuous structure, a multiple-beam system uses discrete components. Usually, in a multiple-beam system, the number of outputs equal the number of input antennas. The maximum number of output ports is often limited by the number of receivers that can be made available.

The beam-forming network consists of hybrids and phase shifters. An eight-element feed system is shown in Figure 3.25. The system consists of 12 hybrid couplers and 8 phase shifters. In this example, the eight-element array forms eight output beams that lie to the right and left of the broadside direction. The output terminals are labeled to identify the beams.

3.16. ANTENNAS (35–38)

The antennas discussed in this section are limited to the ones commonly used in connection with the AOA measurement system discussed in Sections 3.8–3.15. The narrow-beam antennas discussed in Section 3.8 are generally large in size.

The most commonly used antenna in an amplitude or phase comparison system is a spiral antenna, as shown in Figure 3.26. The spiral antenna is fabricated with printed circuit (PC) technology on a dielectric board. The feeding point is at the center of the antenna. The frequency coverage is often over an octave. It covers approximately 90° in spatial angle. The diameter of the antenna is approximately one-half wavelength of the lowest operating frequency. The antenna in Figure 3.26 covers 8–40 GHz. The gain of the spiral antenna is slightly higher at the higher frequency range. Specially designed antennas can work in an extremely high frequency (EHF) range. Since the spiral antenna is fabricated on a printed circuit board, it is basically flat in structure. This structure is suitable for flush mounting on an airborne platform. This kind of antenna is very popular for EW applications.

Horn antennas are often used with the microwave lens AOA system. The antennas are arranged in one- and two-dimensional array. Figure 3.27 shows a two-dimensional array. In general, the horn antenna has higher gain than a spiral antenna but less spatial coverage. This antenna array feeds two microwave lenses to measure AOA in azimuth. The antennas lined in the vertical direction are used to narrow the beamwidth in the elevation direction to increase the gain.

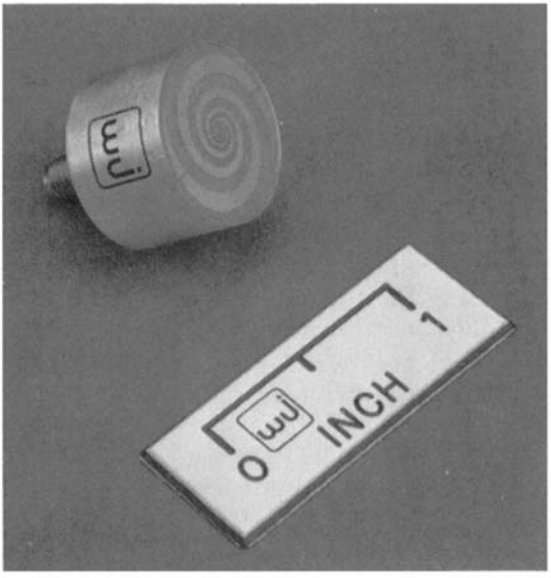


Figure 3.26. A spiral antenna. (Courtesy of Watkins-Johnson Co.)

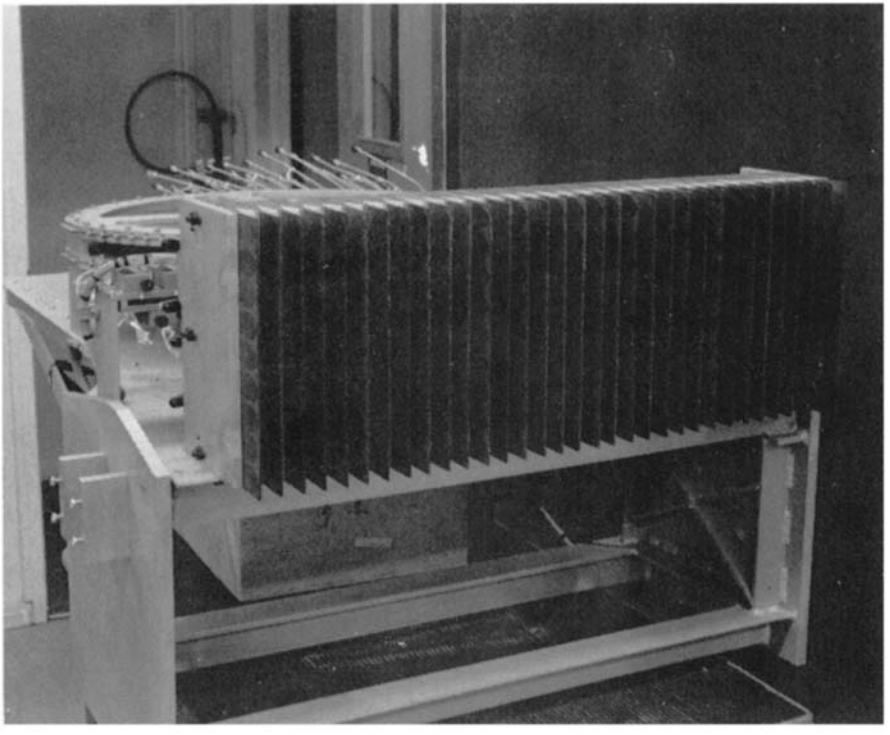


Figure 3.27. Horn antenna array feeding a microwave lens. (Fabricated by Raytheon Co., Courtesy of Avionics Laboratory, AFWAL.)

3.17. SUMMARY

The theoretical limit of the frequency measurement is discussed. The longer the PW and the higher the signal-to-noise ratio, the more accurate the frequency measurement. The PA, PW, and TOA measurement schemes are commonly used in different types of receivers.

To measure the AOA information over 360° azimuth instantaneously, multiple receivers with either amplitude or phase matching are required. Thus the AOA information is the most difficult parameter to measure, but it is also the most valuable information because this is the only parameter a hostile radar cannot vary easily. The frequency and AOA information are the most reliable parameters used in an EW digital processor to sort and identify radar signals.

REFERENCES

1. M. I. Skolnik, *Introduction to Radar Systems*, Chapter 10, McGraw-Hill, New York, 1962.
2. J. J. Risko, Two-level threshold circuit, U.S. Patent 4,139,851, February 1979.
3. S. Mathieu, Apparatus for analyzing a physical quantity, U.S. Patent 4,291,356, September 1981.
4. W. Acker, Signal monitor system, U.S. Patent 4,317,080, February 1982.
5. R. K. Galpin, Zero-crossing comparators with threshold validation, U.S. Patent 4,352,999, October 1982.
6. L. R. Murray and T. M. Wu, Two-level threshold circuitry for large-scale integrated circuit memories, U.S. Patent 4,435,658, March 1984.
7. Hot carrier diode video detectors, Hewlett-Packard, Application Note 923, Palo Alto, CA.
8. R. Hilquist, formerly with Watkins-Johnson Co., private communication.
9. H. G. Hopkins and P. Horner, Direction-finding site errors at very high frequencies, *Proc. IRE*, **96**, 321–345 (1949).
10. L. G. Bullock, G. R. Oeh, and J. J. Sparagna, An analysis of wideband microwave monopulse direct-finding techniques, *IEEE Trans. Aerospace Electronic System*, **AES-7**, 188–202 (1971).
11. M. I. Skolnik, Ed., *Radar Handbook*, McGraw-Hill, New York, 1970.
12. D. F. Yaw, C/D band direction finding antennas, Technical Memorandom EVTM-75-044, Westinghouse Electric Corp., Defense and Electronic Systems Center, Baltimore, April 1975.
13. A. R. Baron, K. P., Davis, and C. P. Hoffmann, Passive direction finding and signal location, *Microwave J.*, **25**, 59 (1982).
14. J. E. Evans, Aperture sampling techniques for precision direction finding, *IEEE Trans. Aerospace Electronic Systems*, **AES-15**, 891–895 (1979).
15. J. B. Harrington, Improving system and environmental DF accuracy, Watkins-Johnson Co. Tech Note, **9**(1) (January/February 1982).
16. J. B. Harrington, DF system calibration and correction techniques, Watkins-Johnson Co., Tech Note, **9**(2) (March/April 1982).
17. T. Harper, Airborne rotary DF antenna systems, Watkins Johnson Co., Tech Notes, **2**(2) (March/April 1975).
18. J. E. Hill, Antenna polarization, Watkins-Johnson Co., Tech Note, **6**(4) (July/August 1979).
19. W. Cohen and C. M. Steinmetz, Amplitude- and phase-sensing monopulse system parameters, *Microwave J.*, Part I, **2**, 27 (October 1959); part II, **2**, 33 (November 1959).
20. R. F. Morrison, Jr., and N. M. Sarachan, Digital direction finder utilizing binary array, U.S. Patent No. 3,213,453, October 19, 1965.
21. V. P. Ipatov and A. V. Titov, Uniqueness and accuracy of phase measurements in the case of two frequency radiation, *Radio Engineering Electronic Physics*, **18**, 140–144 (January 1974).

22. A. V. Titov, Characteristics of two methods of obtaining single valued phase readings in the case of multifrequency radiation, *Radio Engineering Electronic Physics*, **19**, 132–135 (April 1974).
23. C. W. Earp and R. M. Godfrey, Radio direction-finding by the cyclical differential measurement of phase, *IEE Proc. Radio Communication Convention*, **94**, 705–721 (1947).
24. J. A. Fantoni and R. C. Benoit, Jr., Applying the Doppler effect to direction finder design, *Electronic Industries Tele-Tech*, **16**, 75 (January 1957); **16**, 66 (February 1957).
25. D. E. N. Davies, A fast electronically scanned radar receiving system, *J. Brit. IRE*, **21**, 305–318 (April 1961).
26. H. W. Reddien and R. J. Masak, Doppler scanning landing guidance system development, Proc. Inst. Navigation National Conf. pp. 191–205, April 14–16, 1971.
27. R. A. Rosein and L. L. Sanders, The performance of the Doppler microwave landing system in multipath environment, AGARD Symposium on Air Traffic Control Systems, 26–29, June 1972.
28. R. S. Barratt, The UK approach to MLS, *Microwave J.*, **19**, 19 (October 1976).
29. W. Rotman and R. F. Turner, "Wide-angle microwave lens for line source application, *IEEE Trans. Antennas Propagation*, **AAP-11**, 623632 (November, 1963).
30. D. H. Archer, J. Prickett, Signal spectrum analyzer, U.S. Patent 3,735,256, May 22, 1973.
31. R. C. Hansen, Ed., *Microwave Scanning Antennas*, Vol. 1, *Apertures*, Academic Press, New York, 1964.
32. J. L. Butler, Multiple beam antennas, Internal Memo RF 3849, Sanders Associates, Nashua NH, Jan. 8, 1960.
33. J. P. Shelton and K. S. Keller, Multiple beams from linear array, *IEEE Trans. Antenna Propagation*, **AP-9**, 154–161 (March 1961).
34. A. Mosko, An introduction to wideband two-channel direction finding systems, Part I, *Microwave J.*, **27**, 2nd ed., 91 February 1984).
35. H. Jasik, *Antenna Engineering Handbook*, 2nd ed., McGraw-Hill, New York, 1984.
36. W. L. Weeks, *Antenna Engineering*, McGraw-Hill, New York, 1968.
37. E. C. Jorden, *Electromagnetic Waves and Radiating Systems*, Prentice Hall, Englewood Cliffs, NJ, 1964.
38. J. D. Krause, *Antennas*, McGraw-Hill, New York, 1950.

Chapter 4

Crystal Video Receivers

4.1. INTRODUCTION

A crystal video receiver is the simplest in structure among all electronic warfare (EW) receivers. It was also the first microwave receiver developed. This kind of receiver can cover a very wide frequency bandwidth, but it does not have the capability to determine the frequency of the input signals. It can only tell whether there is pulsed microwave energy in the input band of the receiver. Generally speaking, a crystal video receiver cannot detect continuous-wave (CW) signals because the video amplifier following the detector is often ac coupled.

Since the input radio-frequency (RF) bandwidth of a crystal is wide, the sensitivity of the receiver is relatively low, as discussed in Chapter 2. Therefore, the dynamic range of the receiver is moderate. Since the receiver cannot separate input signals by frequency, it cannot detect the existence of simultaneous signals.

This kind of receiver is used in EW applications because of its extremely small size and simplicity. An airborne crystal video receiver system is shown in Figure 4.1. In this figure, there are four spiral antennas with four crystal receivers to cover the four quadrants. A stub antenna, a signal processor, two display units, and a power supply are also included. The four receivers are the four rectangular boxes at the center. Their small size can be readily appreciated if one compares the size of other receivers (see Chapters 5–9).

4.2. BASIC PRINCIPLE OF OPERATION (1–3)

The key component in a crystal video receiver is a crystal detector. The schematic of a simple crystal video receiver is shown in Figure 4.2. The output of an antenna feeds directly to a detector. Following the detector is a video amplifier that

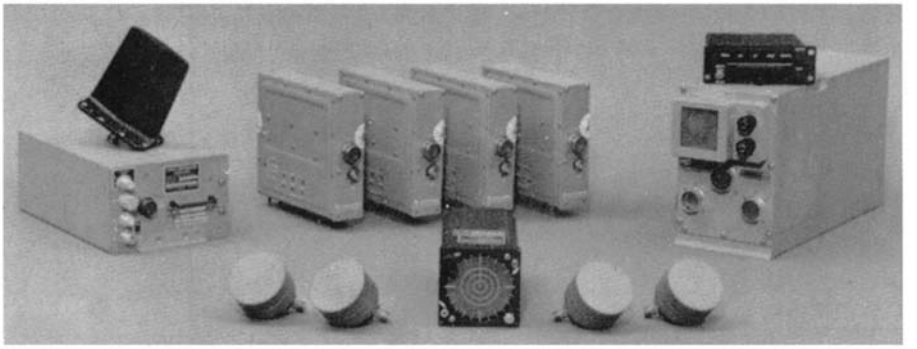


Figure 4.1. Crystal video receiver in an airborne system. (Courtesy of Applied Technology, Div. of Litton.)

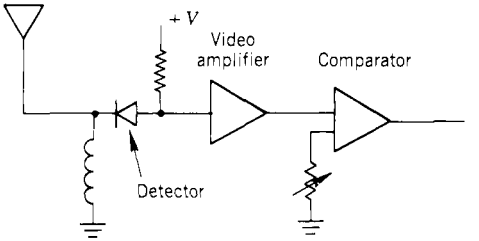


Figure 4.2. A basic crystal video receiver.

amplifies the detector output and passes the video signal to the input of a comparator. An input signal with enough energy will cross the threshold of the comparator and be detected. Without the video amplifier, the video signal from the detector is often too weak to be detected directly through a comparator.

In general, the sensitivity of a receiver depends on the property of the crystal detector and the noise figure of the video amplifier as discussed in Chapter 2. In other words, the sensitivity is limited by the RF gain instead of by the noise in the receiver. The most common approach to improve the sensitivity of a crystal video is to improve the property of the diode detector and reduce the video amplifier noise. Of course, the most effective way to improve the sensitivity is to add RF amplifiers in front of the detector. Since a crystal video receiver often covers a very wide frequency bandwidth, it is usually difficult to find RF amplifiers to cover the desired bandwidth. However, advances in amplifier technology may make such an approach possible.

4.3. PRACTICAL CONSIDERATIONS IN CRYSTAL VIDEO RECEIVERS (4, 5)

A crystal receiver can measure the following parameters:

pulse amplitude (PA),
pulse width (PW),
time of arrival (TOA), and
angle of arrival (AOA).

In general, the parameter measurement circuits in a crystal video receiver are rather simple in structure in order to keep the receiver size small and the cost low. Therefore, one must keep in mind that only limited improvements are worth adding to a practical crystal video receiver. If more capability is desired from a receiver, some other type of receiver should be chosen instead of improving the performance of a crystal video receiver.

One of the major efforts in designing a crystal video receiver is to improve its sensitivity. The easiest approach is to reduce the loss between the antenna and the receiver. Since a crystal video receiver is small, it can be installed very close to the antenna, a common practice in utilizing crystal video receivers.

Equation (2.13) shows that the tangential sensitivity (TSS) of a receiver without RF gain is related to the video bandwidth B_V; the narrower the B_V, the higher the sensitivity. Thus it is desirable to keep the video bandwidth as narrow as possible, yet wide enough to pass the shortest pulse anticipated.

Improving the impedance matching network between the antenna and the detector will ideally transfer maximum energy to the detector. Thus a properly designed matching network in front of the detector will improve the receiver sensitivity (see Section 4.7).

Bias current can be applied to the detector diode to improve the receiver sensitivity. In general, the optimum bias current is frequency dependent, and proper bias current must be determined according to the RF chosen. This subject will be further discussed in the next section.

Sometimes a frequency multiplexer or a tunable filter is added in front of a crystal video receiver. This kind of filter will limit the input RF bandwidth and provide coarse frequency information on the signal received. The filter can also reduce the noise bandwidth. However, the additional inserion loss introduced by the filter will partially offset the noise bandwidth reduction effect. A crystal video receiver with a tunable RF filter in front of the detector is also called a tuned RF (TRF) receiver.

In addition to improving the sensitivity of a crystal video receiver, it is often required to extend the receiver capability to measure CW signals. Most of the video amplifier following the detector is ac coupled, and a CW signal will appear at the output of the detector as a dc voltage, which cannot pass

through an ac-coupled video amplifier. A dc-coupled amplifier will eliminate this problem. However, the output of a dc-coupled amplifier may drift, which will change the signal level at the input of the comparator. If the output level is drifted lower, the sensitivity of the receiver will be reduced. If the output level is drifted higher, the false alarm rate of the receiver will increase. Another problem of a dc-coupled amplifier is that there is dc noise at the output of the detector. A dc amplifier will amplify this noise and reduce the receiver sensitivity. Advancements in dc-coupled video amplifier technology should reduce the drift problem and make it possible for crystal video receivers to receive CW signal.

Another common approach to add CW detection capability is to put an amplitude modulator in front of the detector. The modulator will introduce amplitude modulation on CW signals. The modulated CW signal will produce pulsed video signals after detection and will pass an ac-coupled amplifier. Usually, one CW signal will block the crystal receiver from intercepting other pulsed signals.

Multiple crystal video receivers can be used to measure the AOA of an input signal. Crystal video receivers can measure the AOA only through amplitude comparison because in a phase interferometric AOA system, the frequency of the input signal must be measured, and a crystal video receiver does not have frequency measurement capability. Using crystal video receivers to measure the AOA is the simplest approach. In an amplitude AOA measurement system, logarithmic (log) amplifiers must be added after the video amplifiers. This log amplifier is often referred to as a video log amplifier, which is different from an RF log amplifier. A video log amplifier takes a video signal as input and generates a video output. An RF log amplifier takes RF as input and generates a video output. In both amplifiers, the output represents the log of the input signal. The RF log amplifier can provide RF gain and improves the sensitivity and dynamic range of a receiver. However, they usually cannot cover the RF bandwidth of a crystal video receiver. A detailed discussion of AOA measurements can be found in Section 3.10. The schematic of a crystal video receiver with CW and AOA capabilities is shown in Figure 4.3.

In the following sections, the crystal detectors, video amplifiers, and video log amplifiers will be discussed.

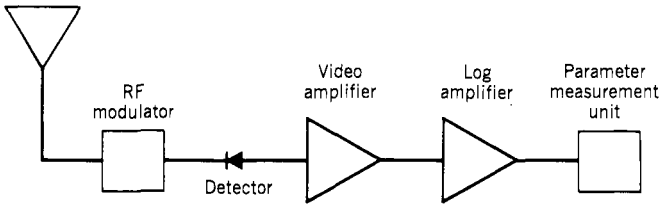


Figure 4.3. Crystal video receiver with amplitude modulator and log amplifier.

4.4. CHARACTERISTICS OF DETECTORS (4–9)

The performance characteristics used to describe detector diodes are current–voltage (I–V) curve, TSS, video resistance, voltage sensitivity, and figure of merit.

A. I–V Curve

The current–voltage relationship of ideal diodes obeys the well-known diode equation:

$$I = I_s\left(\exp\frac{qV}{nkT} - 1\right) \tag{4.1}$$

where I is the current through the diode, V is the voltage applied across the diode, I_s is the saturation current, n is a constant greater than but close to unity, k is Boltzmann's constant ($=1.38 \times 10^{-16}$ erg/°K), T is the temperature (in kelvin), and q is the electronic charge ($=1.601 \times 10^{-19}$ C). For many calculations, $n = 1$ is used.

A typical I–V curve is shown in Figure 4.4. When a diode is used as a detector, it usually operates in the forward biase region. A real diode also exhibits parasitic resistance and capacitance, which limits its operation at high frequencies.

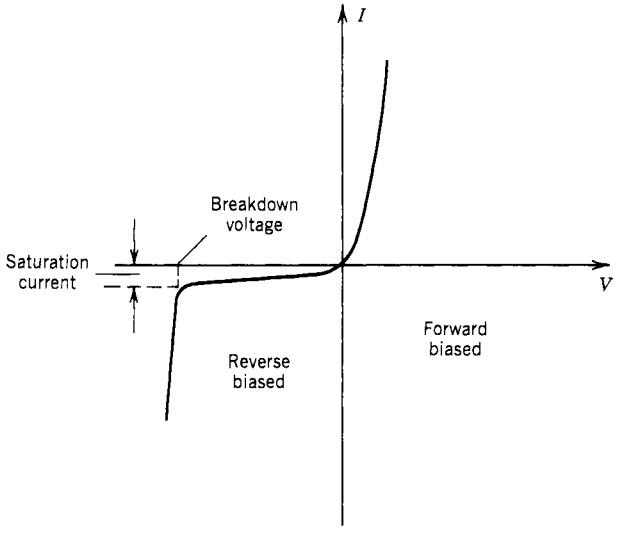


Figure 4.4. Typical I–V curve of a diode.

B. Tangential Sensitivity

Tangential sensitivity (TSS) refers to an input signal level that will provide an 8-dB signal-to-noise ratio (S/N) at the output of a video amplifier following the detector. Detailed discussion can be found in Section 2.3. The TSS value does not only depend on intrinsic diode parameters. There are many other factors that affect the measured TSS value of a given diode. The most important factors are

1. RF and bandwidth,
2. video bandwidth,
3. diode dc bias current,
4. test mount or circuit, and
5. video amplifier noise figure following the diode.

To obtain maximum sensitivity at any given frequency, the diode is usually biased in the forward direction. However, bias introduces noise in the diode and reduces the diode video resistance. These effects exert a competitive influence on TSS; therefore, the bias value must be stated for a given TSS. A typical TSS versus dc bias current of a detector diode is shown in Figure 4.5. The TSS increases with the increase of dc bias current. By further increasing the bias current, the TSS starts to decrease.

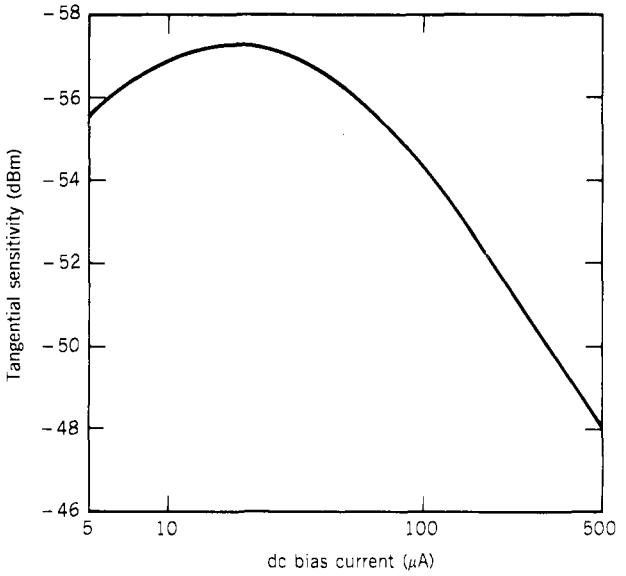


Figure 4.5. Effect of dc bias on TSS. Test conditions: Video load impedance R_L, 38 kΩ; video bandwidth B^V, 2 MHz; equivalent noise resistance of video amplifier (R^A), 500 Ω; RF, 10 GHz. (Based on ref. 5.)

C. *Equivalent Circuit*

The equivalent circuit of a detector diode is shown in Figure 4.6, where L_p and R_s are series inductance and resistance, respectively; C_p is the package capacitance; and C_j and R_j are the junction capacitance and resistance, respectively. The parameters C_p and L_p are circuit elements that depend on the diode geometry and the circuit configurations in which the diode is used. The resistors R_s and R_j and capacitor C_j are dependent on diode material and the junction property.

Video resistance is defined as the sum of the diode series resistance R_s and the junction resistance R_j:

$$R_V = R_s + R_j \tag{4.2}$$

The junction resistance R_j is simply the small-signal low-frequency dynamic resistance of the diode, and it is a function of the dc bias current. Resistance R_j can be obtained by differentiating the diode I–V relationship and is given by

$$R_j = \frac{nkT}{q(I_d + I_s)} \tag{4.3}$$

where I_d is the bias current.

D. *Voltage Sensitivity*

Voltage sensitivity specifies the slope of the output video versus the input power; that is,

$$V = \gamma P_i \tag{4.4}$$

The value of γ depends on the load resistance, signal level, and RF. It is particularly sensitive to the signal level, which must be kept in the square-law range of the diode. A higher value of γ means better sensitivity.

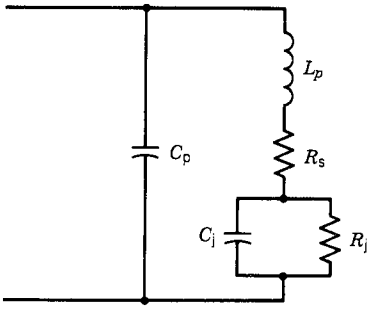


Figure 4.6. Equivalent circuit of a diode detector.

E. Figure of Merit

The figure of merit is an old "measure" of the "supposed efficiency" of a video detector diode and is given by

$$M = \frac{\gamma}{\sqrt{R_V + R_A}} = \frac{\beta R_V}{\sqrt{R_V + R_A}} \tag{4.5}$$

where R_A is the video amplifier series noise resistance and β is the current sensitivity of the diode ($\beta = \gamma/R_V$).

However, sometimes the figure of merit is defined as

$$M = \frac{\gamma}{\sqrt{R_V}} = \beta\sqrt{R_V} \tag{4.6}$$

which means $R_A = 0$ in Eq. (4.5). This condition assumes that the noise contributed from the video amplifier is negligible compared to that of the diode, which is usually true.

4.5. VIDEO BANDWIDTH (5, 9)

The video bandwidth required for pulse recovery depends on the nature of the information to be gained from the pulse. For example, sometimes peak pulse detectability is more significant than pulse shape; other times leading edge information and pulse shape are more desirable. The formal approach in finding the required bandwidth is through the definition of effective bandwidth, defined as (ref. 9)

$$B_e^2 = \frac{(2\pi)^2 \int_{-\infty}^{\infty} f^2 |S(f)|^2 \, df}{\int_{-\infty}^{\infty} |S(f)|^2 \, df} = \frac{(2\pi)^2}{E} \int_{-\infty}^{\infty} f^2 |S(f)|^2 \, df \tag{4.7}$$

where $S(f)$ is the Fourier transform of the input signal and E is the signal energy. Therefore, the effective bandwidth is input signal dependent.

The deviation of the pulse position (ΔT) is related to the effective bandwidth (B_e) by

$$\Delta T = \frac{1}{B_e\sqrt{2E/N_o}} \tag{4.8}$$

where N_o is noise power per unit bandwidth. Equation (4.8) indicates that the wider the effective bandwidth, the less the deviation in pulse position measurement.

There are several empirical equations that are often used to obtain the required video bandwidth (ref. 5). If maximum pulse detectability, which means maximum S/N is desired, the video bandwidth is inversely proportional to the pulse width,

$$B_V = 1/\text{PW} \tag{4.9}$$

for a rectangular pulse and rectangular low-pass filter. If a simple RC filter is used to provide the video bandwidth, the maximum S/N occurs when the filter's upper 3-dB video frequency is approximately

$$f_{u(3dB)} = 0.25/\text{PW} \tag{4.10}$$

In other applications, it may be desirable to resolve the rise time (t_r) of the pulse. The required upper 3-dB frequency is

$$f_{u(3dB)} = 0.35/t_r \tag{4.11}$$

which is usually higher in frequency than that required for maximum S/N.

If the video circuit is ac coupled, the low-frequency cutoff of the video bandwidth will affect the amplitude of the pulse. The amplitude of the pulse will droop. The amplitude droop and the low-frequency 3-dB point is related by

$$f_{L(3dB)} = \frac{\text{droop}\%}{600\ \text{PW}} \tag{4.12}$$

If the video amplifier is dc-coupled $f_{L(3dB)} = 0$, there is no amplitude droop problem.

4.6. TYPES OF DIODE DETECTORS (10–21)

Generally speaking, detector diodes can be classified into two types according to their operating conditions: the square-law detectors (also referred to as the small-signal type) and the linear detectors (also referred to as the large-signal type, or peak detector). Although one type of detector is called linear, a detector is indeed a nonlinear device that rectifies the RF signal. In a square-law detector, the output voltage is proportional to the square of the input voltage, or directly proportional to the input power. The relation can be expressed as in Eq. (4.4):

$$V = \gamma P_i$$

where γ is a constant and P_i is the input power. In a linear detector, the output

voltage is

$$V = kV_i \tag{4.13}$$

where k is a constant and V_i is the input voltage.

A detector will contain both the square-law region and the linear region. Figure 4.7 shows a typical detector output voltage versus input power. At low input power levels, the detector operates at the square-law region, whereas at high input levels, the detector operates at the linear region. In a crystal video receiver, the detector often operates in both regions depending on the input power. If the detector is used in a receiver with enough RF gain, and saturation occurs at the last amplifier stage, the input at the detector may be kept at a constant level. Under this condition, the detector can operate at either the linear region or the square-law region. The detector is usually designed to operate at the linear region because of the high saturation power level of an RF amplifier.

The most popular detectors for microwave applications are Schottky and tunnel diodes. A Schottky diode uses the Schottky barrier, a metal-to-semiconductor junction, to form the diode. The current in a Schottky barrier junction depends on majority carrier conduction, in contrast to the minority carrier operation of ordinary *pn* junction diodes. Since there are no minority carrier storage effects involved, the devices are potentially capable of operating at frequencies approaching the reciprocal of the dielectric relaxation time of the semiconductor crystal. For a practical device, this frequency will be on the order of 1000 GHz.

Schottky diodes can be made using two approaches: the point contact diode and the planar Schottky diode (or hot carrier diode). The point contact

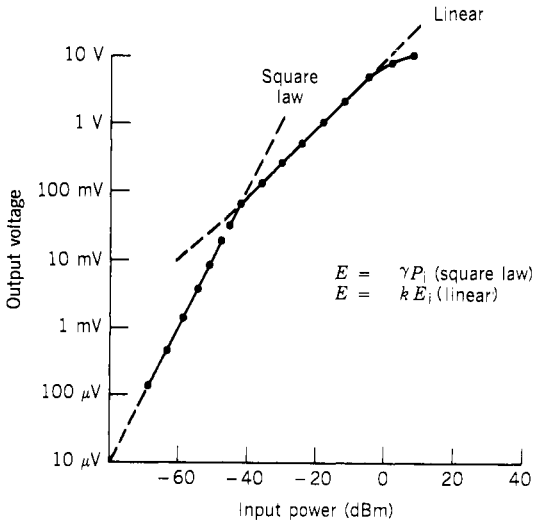


Figure 4.7. Typical detector output voltage versus input power. (Based on ref. 10.)

diode is fabricated by pressing a fine metal point into the surface of the semiconductor, sometimes followed by an electrical or mechanical forming operation. A hot carrier diode is fabricated by depositing a metal film on a prepared surface of the semiconductor. A photometallurgically built Schottky diode is shown in Figure 4.8. In this figure, A is a heavily doped single-crystal waver, B is a lightly doped epitaxial silicon layer, C is a layer of protective glass (SiO_2), D is the metal semiconductor junction, and E is a layer of suitable metal deposited on the surface of the semiconductor. The choice of the metal is the most important factor in controlling the Schottky barrier height and consequently the saturation. Contact metal (F) is plated at the top and bottom of the device. The obvious advantages of the hot carrier diode over the point contact diode are closer control of geometry and better resistance to mechanical shock. It is difficult, however, to make hot carrier diodes with capacitances as small as that of a point contact diode. This capacitance and the series resistance of the diode determine the upper operation frequency of the diode.

A tunnel diode is a *pn* junction device whose doping density is made purposely high in order to produce a very narrow junction across which electrons can tunnel easily. Esaki discussed the tunneling effect in 1948 (ref. 16); thus the diode is also called an Esaki diode. The tunneling phenomenon is a majority carrier effect. The tunneling time of carriers through a potential energy barrier is governed by the quantum transition probability per unit time. The tunneling time is very short, permitting the use of tunnel diodes well into the millimeter wave range. The $I-V$ curve of a typical tunnel diode is shown in Figure 4.9. When a tunnel diode is operated as a detector, it is often biased in the forward direction at a voltage somewhat lower than the voltage corresponding to the peak current I_p. However, a bias current in a tunnel diode is not critically required.

The performance of Schottky diodes versus tunnel diodes as detectors is listed below (see also ref. 20):

1. If properly biased, Schottky diodes have better sensitivity.
2. Tunnel diodes have faster rise time, especially at higher power levels (0 dBm). The rise time of Schottky diodes can be improved by loading the diode with resistance, thus sacrificing sensitivity.

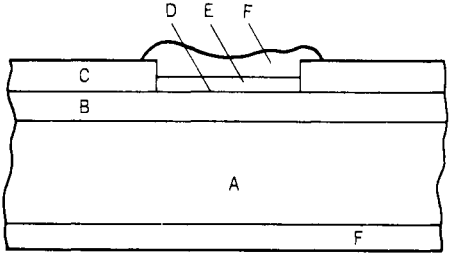


Figure 4.8. Cross section of a Schottky diode. (Based on ref. 18.)

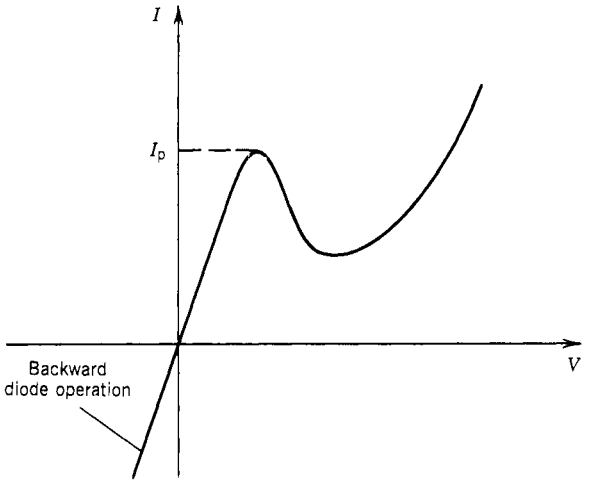


Figure 4.9. Typical $I-V$ curve of tunnel diode.

3. It is easier to build dc-coupled detector log amplifiers by using tunnel diodes because no dc bias is required. In a dc-coupled Schottky detector, the dc bias is often eliminated in order to reduce the drift problem. A common method for applying dc bias is shown in Figure 4.10. The dc bias effect is subtracted by the differential amplifier. The second diode in the circuit is used only to compensate the drift in the detector. The $I-V$ curve of the two diodes should be closely matched over the operating temperature range. Even with this arrangement, the error introduced is sometimes quite large.

4. Tunnel diodes have better tracking accuracy versus temperature than that of Schottky diodes.

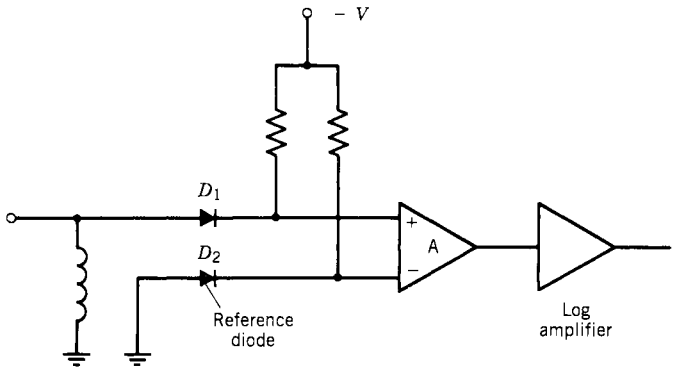


Figure 4.10. A dc-coupled amplifier.

5. The input VSWR of a tunnel diode can be kept below 2.5 through proper circuit design. The input VSWR of Schottky diodes can be improved by increasing the dc bias current, which in turn decreases the detection sensitivity.

Sareen (ref. 20) concluded that in system applications where the overall parameters need to be optimized, a tunnel diode is superior to the Schottky detector. However, if only one or two of the discussed parameters are important, a Schottky may be the logical choice.

4.7. DIODE DETECTOR CIRCUITS AND VIDEO AMPLIFIERS (4, 5, 7)

A typical diode detector circuit is shown in Figure 4.11. The circuit on the left side of the diode is the RF circuit. The circuit on the right side of the diode is the video circuit. If the diode is represented by its equivalent circuit as shown in Figure 4.6, the circuit in Figure 4.11 can be represented in Figure 4.12*a* with its RF and video circuits shown in Figures 4.12*b* and 4.12*c*, respectively.

A. Video Circuit

In the video circuit, R_L represents a load or input resistance of a video amplifier and C_A represents the amplifier input capacitance as well as all the stray and particularly the cable capacitances present in the video circuit. These R and C elements will impose a limit on the upper cutoff frequency of this circuit, which is given by

$$f_{u(3dB)} = \frac{1}{2\pi R_T C_T} \tag{4.14}$$

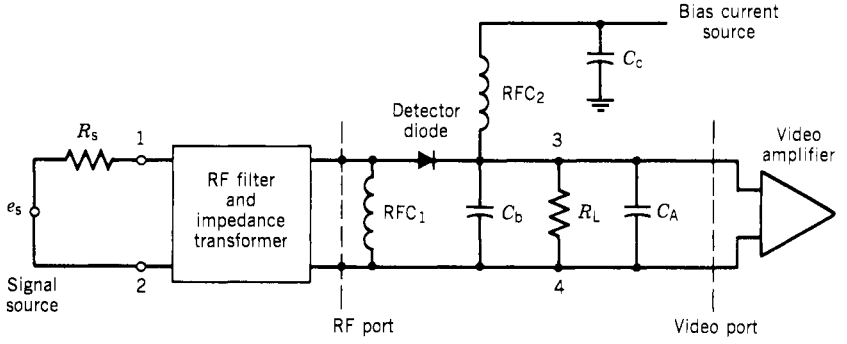


Figure 4.11. Typical detector circuit. (based on ref. 5)

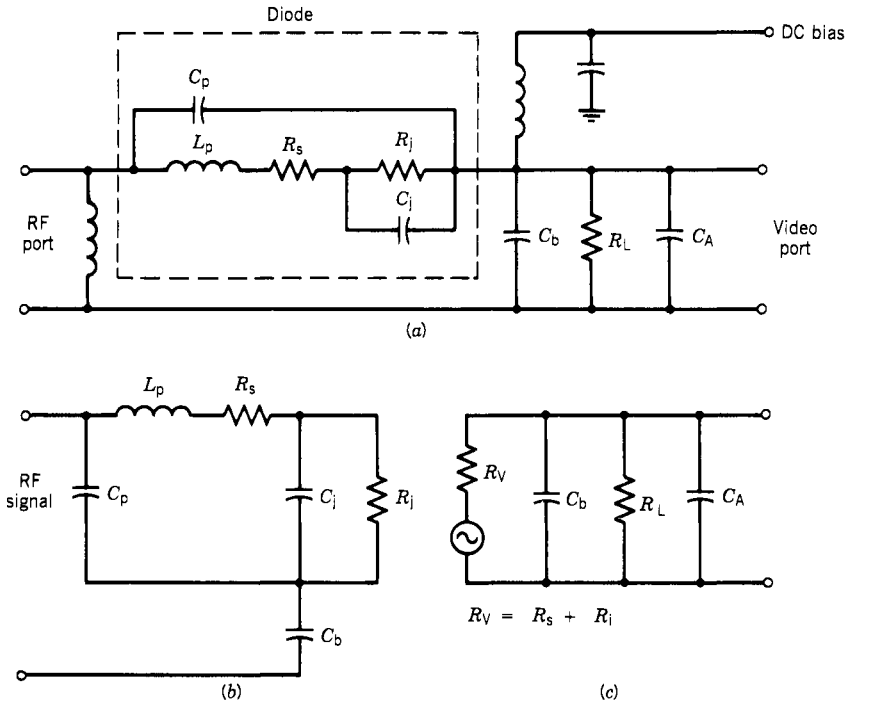


Figure 4.12. Video detector equivalent circuits: (*a*) complete video detector circuit; (*b*) equivalent circuit at RF port; (c) equivalent circuit at video port. (based on ref. 5)

where

$$R_T = \frac{R_V R_L}{R_V + R_L} \tag{4.15}$$

and

$$C_T = C_b + C_A \tag{4.16}$$

In most video detector designs, this circuit has a greater influence on the effective video bandwidth than the bandwidth of the video amplifier. The $R_T C_T$ time constant can be reduced by reducing all the element values within certain limits. A severe reduction in the value of the RF bypass capacitance C_b will lead to poor RF–video isolation and a decrease in the signal level delivered to the diode. The reactance of this capacitor at the operating frequency should be kept to less than 10% of the RF impedance of the diode. At low RF and wide video bandwidths, this capacitor can be replaced by a low-pass filter structure as shown in Figure 4.13 with a cutoff frequency $f_c < \frac{1}{2} f_{rf}$. Since this

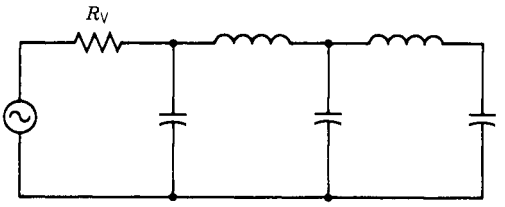


Figure 4.13. Low-pass video coupling structure. (Based on ref. 5.)

filter is required to pass short pulses, the design should be for flat time delay rather than for equal ripple.

A reduction in stray capacitance and in the input capacitance of the video amplifier is always advisable, and in wide-band designs this reduction is often necessary. One suitable technique that is particularly effective when cable interconnections must be made between the detector circuit and the video amplifier is shown in Figure 4.14. In this circuit, the original cable capacitance C_0 is reduced by feedback to

$$C_e = C_0(1 - A_V) \tag{4.17}$$

where A_V is the total gain from the input of the amplifier to the first shield and must be less than 1.

Alternately, either R_L or R_V can be reduced to increase the video bandwidth in Eq. (4.14). The amount that R_L can be reduced is often limited if voltage amplification is desired since the output voltage of the detector is maximized by making R_L large. The resistance R_V of the diode can be lowered by increasing the bias current. Although this results in reduced sensitivity, as shown in Figure 4.5, it may nevertheless be needed to achieve the required video bandwidth. A reduction of video resistance has other beneficial effects. For instance, it can be adjusted to be the optimum source resistance value for minimum noise figure of the video amplifier.

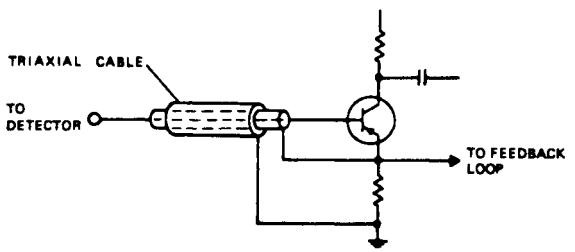


Figure 4.14. Circuit for reducing cable capacitance. (Based on ref. 5. Courtesy of Hewlett Parkard Co.)

B. *RF Circuit*

The RF circuit consists essentially of a filter structure that is designed to match the signal source resistance, usually 50 Ω, to the diode junction resistance R_j. The equivalent circuit of the diode at the RF port is shown in Figure 4.12*b*. The presence of R_s and C_j introduces a loss of signal that is dependent on frequency and the magnitudes of R_s, C_j, and R_j.

This dependence is given by

$$L_{dB} = 10 \log \left(1 + \frac{R_s}{R_j} + \omega^2 C_j^2 R_s R_j\right) \tag{4.18}$$

Since R_j is a function of bias current, that is,

$$R_j \simeq 28/I_d \tag{4.19}$$

the above equation can be restated as

$$L_{dB} = 10 \log \left(1 + \frac{R_s I_d}{28} + \frac{28\omega^2 C_j^2 R_s}{I_d}\right) \tag{4.20}$$

from which it is obvious that the RF loss can be minimized at any given frequency by biasing the diode. Bias also affects the RF impedance of the diode, and the impedance obtained at a bias corresponding to maximum TSS is not necessarily optimum for achieving broadband RF matching to the diode.

Assuming that the RF bypass capacitor C_b is properly chosen so that its reactance can be neglected, the RF admittance of the remaining circuit can be expressed as

$$Y = j\omega C_p + \left(R_s + j\omega L_p + \frac{1}{j\omega C_j + 1/R_j}\right)^{-1} \tag{4.21}$$

This admittance has two potential resonant frequencies. Neglecting R_s and R_j, these resonant frequencies can be expressed approximately as

$$\omega_s \simeq 1/\sqrt{L_p C_j} \tag{4.22}$$

$$\omega_p \simeq \sqrt{1/L_p C_j + 1/L_p C_p} \tag{4.23}$$

The presence of these resonances makes it more difficult to achieve broadband RF matching structures for the diode.

To simplify broadband matching, the impedance of the diode can be reduced by using more dc bias than would normally be required for maximum TSS. Although there will be a reduction of sensitivity, the trade-off may be worthwhile in broadband systems. Sometimes a shunt resistor across the diode is

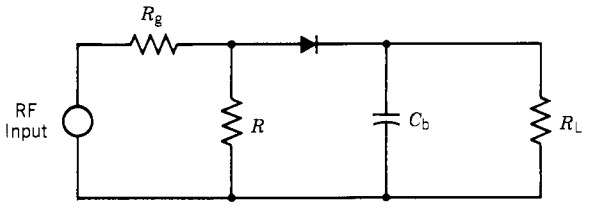


Figure 4.15. Resistive matching of detector diode (R = matching resistor, C_b = video capacitor.)

added to reduce the VSWR as shown in Figure 4.15 (refs. 4, 18). Sensitivity may be traded for VSWR by adjusting the value of the shunt resistor. For example, improving the match with a shunt resistor of 1000–300 Ω causes more loss due to power absorbed by the resistors. However, sensitivity loss due to reflections from an unmatched diode is 1 or 2 dB worse than that due to the matching network using a 300-Ω resistor.

4.8. VIDEO LOGARITHMIC AMPLIFIERS (22–24)

A log amplifier is often used after a video amplifier in a crystal receiver to provide a logarithmic response because PA is measured in decibels. If a differential amplifier that amplifies the difference of two input signals is used after two log amplifiers each with input A and B, respectively, the ratio of the input signals is obtained because

$$\log A - \log B = \log(A/B) \tag{4.24}$$

This is the basic relation used in an amplitude AOA measurement system.

A basic log circuit can be built as shown in Figure 4.16 where A is an operational amplifier. An operational amplifier has a very high input impedance and high gain. The input voltage of the amplifier is near zero. Therefore, the input current I_i must equal the feedback current I_f:

$$I_i = -I_f \tag{4.25}$$

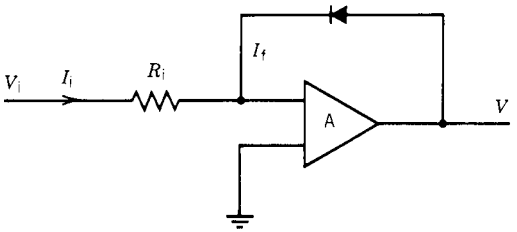


Figure 4.16. Basic log amplifier.

But

$$I_i = \frac{V_i}{R_i} \tag{4.26}$$

$$I_f = I_s\left(\exp\frac{qV}{kT} - 1\right) \simeq I_s \exp\frac{qV}{kT} \tag{4.27}$$

where V_i is the input voltage, R_i is the input resistance, V is the voltage at the output of the operational amplifier and the voltage across the feedback diode, q is the electronic charge, k is Boltzmann's constant, and T is at room temperature.

Substituting Eqs. (4.26) and (4.27) into (4.25) one can obtain

$$V = \frac{2.3kT}{q}\log\left(\frac{V_i}{R_i I_s}\right) = \frac{2.3kT}{q}\left(\log\frac{V_i}{R_i} - \log I_s\right) \tag{4.28}$$

Therefore, the output of the amplifier is proportional to the log of the input voltage. If the input voltage is the output of a square-law detector, the input voltage is proportional to the input RF power. Under this condition, the output of the log amplifier is proportional to the log of the input power. The log response of a log amplifier is quite accurate. Figure 4.17 shows the output voltage versus the input power of an ideal log amplifier. The direction of the diode determines the operation polarity of the circuit.

Temperature drift is a major problem in log amplifier. From Eq. (4.28) one can see that there are two separate temperature effects to be compensated: a

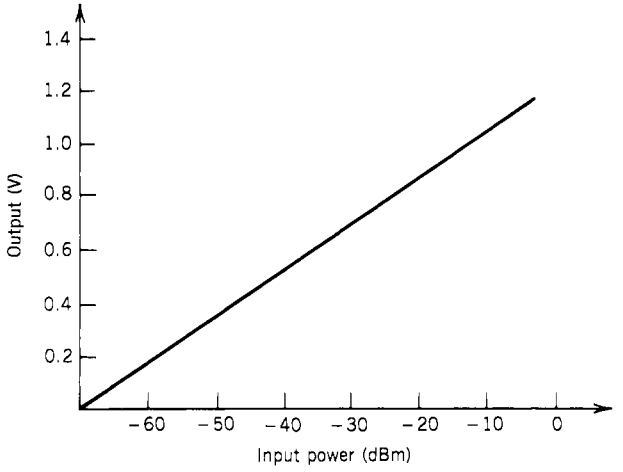


Figure 4.17. Input versus output of an ideal log amplifier.

temperature-sensitive scale factor ($2.3kT/q$) and a temperature-sensitive offset term ($2.3kT \log I_s/q$) (ref. 24). The temperature-sensitive scale factor can be compensated by a temperature-sensitive resistor. The temperature-sensitive offset term can be removed or reduced by a constant current source and a second, matched diode, as shown in Figure 4.18.

The response time of the log amplifier is usually of great concern. The rise time of the log amplifier depends on the video bandwidth. The rise time of a few nanoseconds can be accomplished in some log amplifiers. Whereas the rise time can be made very fast, the falling time of a log amplifier is relatively long. The falling time depends on the input pulse amplitude and width. The long falling time can prevent the log amplifier from detecting a weak signal following a strong one as shown in Figure 4.19.

A video log amplifier cannot improve the dynamic range of a receiver. The dynamic range of a receiver is determined by the diode detector. However,

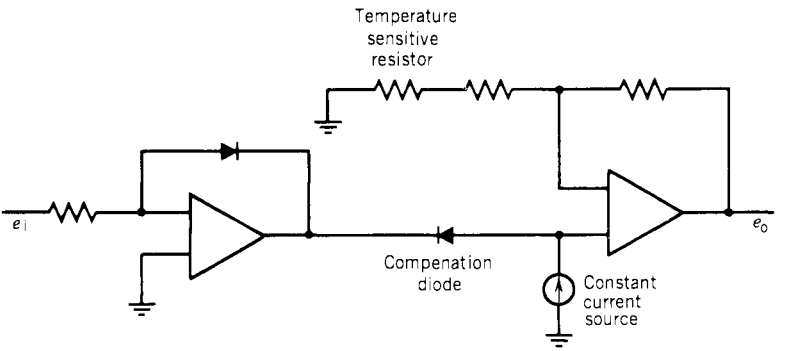


Figure 4.18. Compensation circuit of a log amplifier. (Based on ref. 24.)

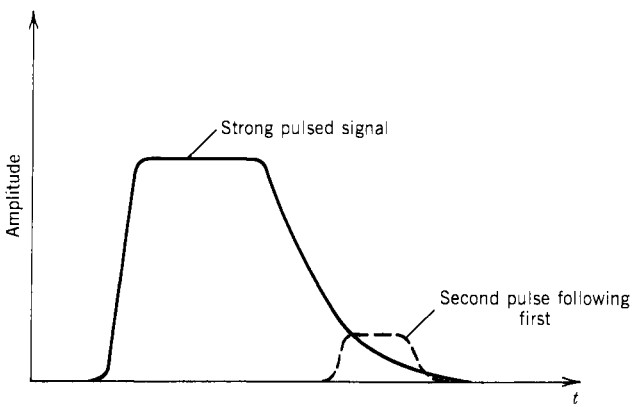


Figure 4.19. Typical pulse output from a log amplifier.

an RF log video amplifier can improve the dynamic range of a receiver (see Section 5.7).

4.9. SUMMARY

A crystal video receiver has relatively low sensitivity and dynamic range. It does not have fine frequency measurement capacity. However, due to its simplicity in structure, it is often used as a warning receiver such as that used in an automobile to detect traffic radar. Due to its small volume and relatively low cost, it is sometimes used as an AOA measurement system in conjunction with other EW receivers.

REFERENCES

1. S. N. Van Voorhis, *Microwave Receivers* (Radiation Laboratory Series), McGraw-Hill, New York, 1947.
2. L. B. Van Brunt, Applied ECM, EW Engineering Inc., Dunn Loring, VA, 1982.
3. W. E. Ayer, Characteristics of crystal video receivers employing RF preamplifion, Technical Report No. 153-3, Stanford Electronics Laboratories, Stanford University, September 20, 1956.
4. Impedance matching techniques for mixers and detectors, Application Note 963, Hewlett-Parkard Co., Palo Alto, CA, July 1974.
5. Hot carrier diode video detectors, Application Note 923, Hewlett-Parkard Co., Palo Alto, CA.
6. W. J. Lucas, Tangential sensitivity of a detector video system with r.f. preamplification, *Proc. IEEE*, **113**, 1321–1330 (1966).
7. The 33800 series mixer/detector module, Application Note 921, Hewlett-Parkard Co., Palo Alto, CA.
8. H. D. Mills, On the equation $i = i_0[\exp \alpha(V - R_i) - 1]$, *IBM J.*, **11**, 553–554 (1967).
9. M. I. Skolnik, *Introduction to Radar Systems*, Chapter 10, McGraw-Hill, New York, 1962.
10. Dynamic range extension of Schottky detectors, Application 956-5, Hewlett-Parkard Co., Palo Alto, CA, October 1975.
11. A. M. Cowley and H. O. Sorenson, Quantitative comparison of solid-state microwave detectors, *IEEE Trans. Microwave Theory Techniques*, **MTT-14**, 588–602 (1966).
12. H. A. Watson, *Microwave Semiconductor Devices and Their Circuit Applications*, Chapter 11, McGraw-Hill, New York, 1969.
13. A. van der Ziel, *Solid-State Physical Electronics*, McGraw-Hill, New York, 1968.
14. S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed., Chapter 8, Wiley, New York, 1981.
15. W. Schottky, *Nataiwiss.*, **26**, 843 (1938).
16. L. Esaki, New phenomenon in narrow germanium *p-n* tunnel junction, *Phys. Rev.*, **109**, 603–604 (1958).
17. B. A. Miller, T. P. Miles, D. C. Cox, A design technique for [illegible]ing a microwave tunnel diode amplifier in strip line, *IEEE Trans. Microwave Theory Techniques*, **MTT-15**, 554–561 (1967).
18. P. A. Szente, S. Adam, and R. B. Riley, Low-barrier Schottky diode detectors, *Microwave J.*, **19**, 42 (February 1976).
19. C. A. Burrus, Backward diodes for low level millimeter wave detection, *IEEE Trans. Microwave Theory Techniques*, **MTT-11**, 357–362 (1963).
20. S. Sareen, Schottky versus tunnel diode detectors in crystal video applications, Aertech Application Note, Sunnyvale, 1975.

21. S. Sareen, Threshold detectors for nanosecond fault or level detection in RF systems, Aertech Application Note, Sunnyvale, 1975.
22. D. Sheingold and F. Pouliot, The hows and whys of log amps, *Electronic Design*, **3**, 52 (February 1, 1974).
23. A. Helfrick, Build high-gain, wide-range log amps, *Electronic Design*, **6**, 116 (March 15, 1974).
24. J. G. Graeme, G. E. Tobey, and L. P. Huelsman, Eds., *Operational Amplifiers: Design and Applications*, McGraw-Hill, New York, 1971.

Chapter 5

Superheterodyne and Homodyne Receivers

5.1. INTRODUCTION

A superheterodyne receiver (often referred to as a superhet receiver) uses a very clever idea to improve the sensitivity and selectivity of a receiver. As discussed in Chapter 4, the most effective way to improve the sensitivity of a receiver is to add amplifiers in front of the detector. However, radio-frequency (RF) amplifiers are sometimes not available at the operating frequency. The superheterodyne circuit invented by Edwin H. Armstrong in 1918 improved the sensitivity of a receiver by adding amplification at a lower frequency. This approach linearly transforms the input RF signal to a lower frequency where amplifiers and narrow-bandwidth filters are readily available. The information in the input signal will be retained, since the transform is linear. The amplifiers may improve the receiver sensitivity from gain-limited case to noise-limited case, and the filters will limit the noise bandwidth. Increasing the amplification gain and reducing the noise floor will greatly improve the sensitivity of a receiver.

A superhet receiver is the most commonly used receiver in communication because of its high sensitivity and selectivity. Almost all the commercial radio and radar receivers are of the superhet type. In electronic warfare (EW) applications, the superhet receivers are used to isolate an input signal and measure its fine-grain information. Since it is relatively easy to phase match between two identical superhet receivers because of their narrow bandwidth, they are commonly used in a phase interferometry system to measure the angle of arrival (AOA) of the input signals. A superhet receiver is shown in Figure 5.1. In this receiver, there are several switchable IF filters of different bandwidths. The proper IF bandwidth should be selected according to the application. The

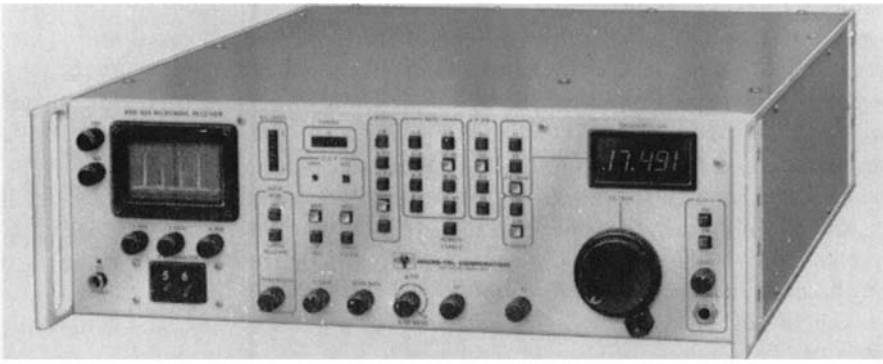


Figure 5.1. A superhet receiver. (Courtesy of Microtel Corp.)

frequency tuning is displayed in digital format, and the input signals are displayed on a scope.

To linearly transform the input frequency in a superhet receiver, two basic microwave components or subsystems are required. They are a mixer and a local oscillator. There are many different types of mixers and oscillators, and they will be discussed later in this chapter.

5.2. PRINCIPLE OF OPERATION (1–5)

A basic superhet receiver is shown in Figure 5.2. This receiver contains five components: a mixer, a local oscillator (LO), an intermediate-frequency (IF) filter, an IF amplifier, and a video detector. The LO generates a continuous-wave (CW) signal of frequency f_L. If the input signal frequency is f_s, the mixer will shift f_s to f_I, which is the difference frequency of f_L and f_s. This procedure is often called down conversion, because the IF is lower than the input frequency f_s. For some special applications, the IF may be higher than the input frequency f_s, and this procedure is referred to as up conversion.

The IF filter following the mixer is a bandpass filter that is used to pass the desired IF signal and to stop all other frequencies generated in the mixer.

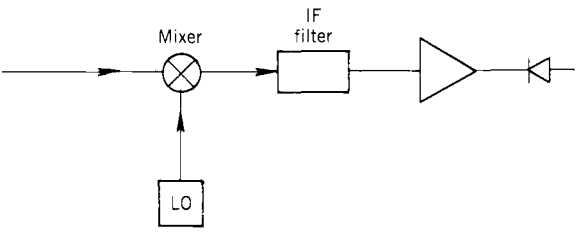


Figure 5.2. A basic superhet receiver.

Although the mixer can linearly transform one frequency to another frequency, strictly speaking, it is not a linear device. When two frequencies f_s and f_L are present at the input of the mixer, many intermodulation (intermod) frequencies will be generated at the output. Even if there is only one input signal, the mixer will produce many intermod frequencies. These frequencies will confuse the receiver measurement and should be filtered out.

The IF filter is also part of the frequency measurement circuit, because if f_L and f_I are known, the input signal frequency can be determined. The center frequency of the filter is usually fixed, and the frequency of the LO is tuned to let the difference frequency $f_L - f_s$ (or $f_s - f_L$) pass through the filter.

The IF amplifier following the IF filter will provide most of the gain of the receiver. This gain will increase the sensitivity of the receiver. Of course, the mixer in front of the receiver will degrade the noise figure of the receiver, but the IF gain added may change the sensitivity from a gain-limited case to a noise-limited case, which will greatly improve the receiver sensitivity. The following example will demonstrate the improvement of the sensitivity. If in a certain superhet receiver the mixer has an insertion loss of 6 dB followed by an IF filter of 1 MHz bandwidth with a 1-dB loss and an IF amplifier with a noise figure of 5 dB, the overall noise figure of the receiver is approximately 12 dB (6 + 1 + 5). If the video bandwidth is also 1 MHz and the IF gain is high enough that the receiver is noise limited, the tangential sensitivity can be calculated from Eq. (2.18) as

$$\text{TSS} = -114 + 12 + 10\log(5.65) = -94.5 \quad \text{dBm}$$

Without the IF amplifier, the receiver sensitivity depends on the crystal detector, which may be in the region of -40 to -50 dBm. The advantage of the IF gain can be readily seen from this simple example.

Following the IF amplifier is a crystal video detector. The detector performs the same function as any other detector in a microwave receiver that converts microwave energy to a video signal. A video amplifier following the detector is often used to amplify the video signal for further processing. The video signals are often visually displayed or converted into sound as in a commercial radio receiver. In an EW receiver, a comparator is often used after the video amplifier to detect the existence of input signals. As discussed in Chapter 3, when the input signal is near the threshold, the noise riding on the signal may trigger the comparator many times even with one input signal. This multiple triggering problem will confuse the following processor. Special circuits such as those discussed in Chapter 3 may be required to eliminate this problem.

5.3. INTERMODULATION GENERATED IN A MIXER (6–16)

In this section, the intermod frequencies generated by a mixer and their effects to the receiver will be discussed. The principle of operation and the structures of mixers will be discussed in Section 5.8.

Since a mixer is a nonlinear device, it will generate many other frequencies besides the desired frequency $\pm(f_L - f_s)$ or $f_L + f_s$. All these undesired frequencies are called spurious responses, or simply spurs. The frequencies of these spurs are very predictable. They can be written as

$$f_I = Mf_1 + Nf_2 \tag{5.1}$$

where f_1 and f_2 are used to represent the input frequencies instead of f_L and f_s. Using f_1 and f_2 is more generalized because one frequency can be considered to be the signal and the other the LO frequency; M and N are integers (either positive or negative) and the output frequency f_I must be a positive value.

Although Eq. (5.1) is extremely simple, it is tedious to calculate all the output frequencies. A spur chart is often used to predicate the spurious frequencies. Assuming $f_2 > f_1$ and dividing Eq. (5.1) by f_2, one obtains

$$\frac{f_I}{f_2} = M\frac{f_1}{f_2} + N \tag{5.2}$$

Plotting f_I/f_2 versus f_1/f_2 with different M, N values, the spur chart can be obtained. It should be noted that the desired output frequencies are

$$f_I = -f_1 + f_2 \quad \text{for } M = -1, N = 1 \tag{5.3}$$

and

$$f_I = f_1 + f_2 \quad \text{for } M = 1, N = 1 \tag{5.4}$$

Equation (5.3) represents a down-conversion, whereas Eq. (5.4) represents up-conversion.

To keep f_I positive, the spur chart for the down-conversion case is limited in the region

$$0 < f_I/f_2 < 1 \qquad \text{and} \qquad 0 < f_1/f_2 < 1 \tag{5.5}$$

whereas for the up-conversion case

$$1 < f_I/f_2 < 2 \qquad \text{and} \qquad 0 < f_1/f_2 < 1 \tag{5.6}$$

A spur chart for $f_I = f_2 - f_1$ containing all the spurious frequencies generated up to the 6 by 6 order is shown in Figure 5.3. To simplify the notation, let $f_1 = L$ and $f_2 = H$. The regions on the output line $f_2 - f_1$ where there are no lines across are the preferable areas for down conversion because there is no spur past these regions. However, if a wide-bandwidth IF is desired, there are almost always some spurs in the IF bandwidth. As a general rule, the higher the order of the spurs, the smaller their amplitude. An area containing some higher-order spurs usually is acceptable for wide-band down conversion.

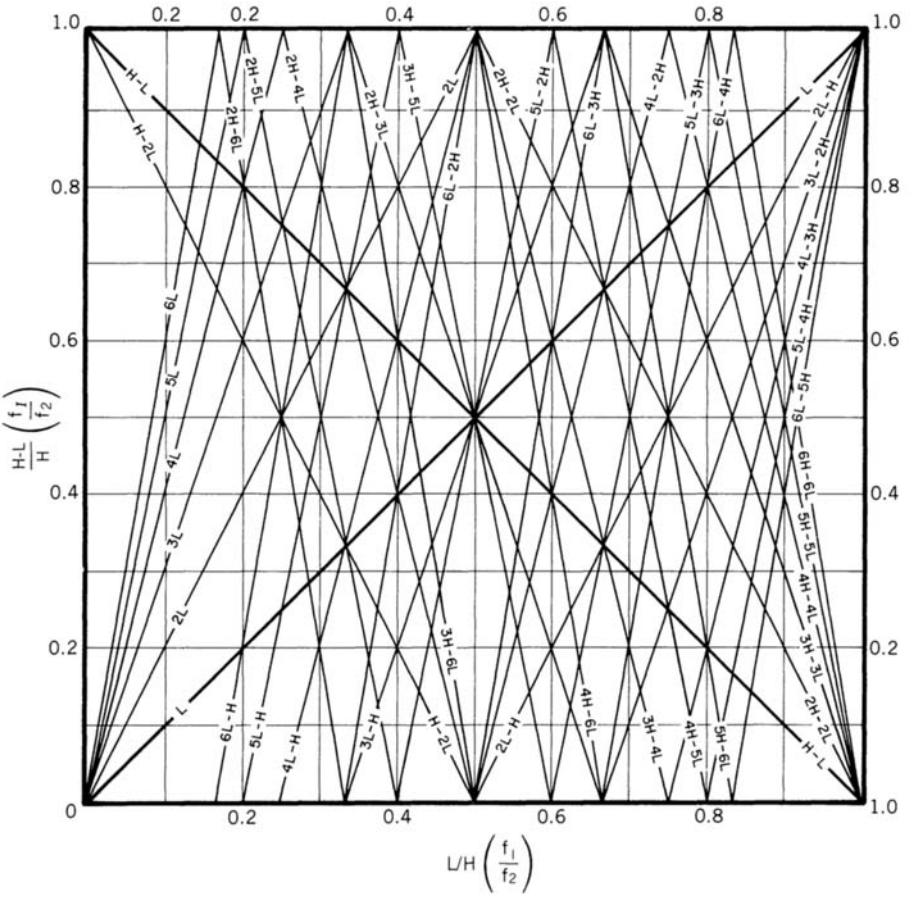


Figure 5.3. Down-conversion mixer spur chart.

Let us use an example to demonstrate how the spur chart is used and some of the spur problems of a superhet receiver. If the input signal is in the 450–550-MHz range and the IF filter is centered at 100 MHz with a bandwidth of 10 MHz, the LO frequency must sweep a range of 100 MHz that can be either 100 MHz above or below the input signal. Let us assume that the LO is 100 MHz below the input signal. Therefore, the LO will sweep from 350 to 450 MHz. The spur should be checked across the entire frequency band by the spur chart. Let us find the spurs at the center of the input bandwidth where

$$H = 500 \text{ MHz} \qquad L = 400 \text{ MHz} \qquad H - L = 100 \text{ MHz}$$

Therefore,

$$L/H = 0.8 \qquad (H - L)/H = 0.2$$

which locates the center of the input and output bands. The IF filter is from 95 to 105 MHz; thus the range of L/H is 0.79 (395/500) to 0.81 (405/500). Similarly, the range of $(H - L)/H$ is from 0.19 to 0.21. One can draw a square with the left lower corner at 0.79, 0.19 and the right upper corner at 0.81, 0.21. In the square, there are two spurs: one caused by $5H - 6L$, the other by $4L - 3H$. The problem caused by these spurs is not obvious at first glance. Since the output frequencies of the spurs are at the center of the IF filter with amplitudes much smaller than the desired output, it will not cause any confusion to the measurement procedure.

Now let us change the input condition slightly. If the input signal H is 540 MHz instead of 500 MHz, and the LO frequency is sweeping to determine the input frequency. Of course, if the LO is at 440 MHz, the IF will pass through the filter, indicating that the input frequency is at 540 MHz. However, if the LO is at 430 MHz, the $4L - 3H$ will be 100 MHz, which will also pass the IF filter. Thus the receiver will report an input signal at 530 MHz with amplitude equal to the amplitude of the $4L - 3H$ spur, which is wrong. In a similar manner, at a certain input frequency the $5H - 6L$ product will pass the IF filter and be reported as an input signal. Generally speaking, for one input signal the mixer will generate many spurs and the receiver will report them as individual signals. The dynamic range of the receiver is limited by these spurs.

Another spur generated by a mixer to confuse the receiver output is when the input signal is 100 MHz below the LO frequency. In the above example, if there is a signal at 300 MHz, when the LO is at 400 MHz, the IF will pass the output filter and the receiver will declare that an input signal is at 500 MHz. This is referred to as the image problem. The solution to this problem is to install a bandpass filter with a bandwidth from 450 to 550 MHz in front of the mixer. This filter will block input signals outside the band of interest. Although this filter will solve the image problem, it will not eliminate the spurious problem.

5.4. PRESELECTOR (TRACKING RF FILTERS) (16)

The best way to reduce the spur problem in a superhet receiver is to use mixers with very low spurious output. Many mixers are designed to accomplish this purpose. In some wide-band (i.e., 100–500-MHz) frequency converters, this is still the only approach to reduce the spur levels. However, many of the lower-order spurs [small value of M and N in Eq. (5.1)] are quite large in amplitude, and it is impossible to neglect them in a receiver with high dynamic range.

Another very effective approach to reduce the spurious problem is to install a tunable filter in front of the mixer. The filter is often referred to as a preselector or tracking RF filter. The preselector will be tuned at the same rate as the LO or it will track the LO. The center frequency of the preselector is related to the LO frequency as

$$f_s = f_L \pm f_I \tag{5.7}$$

which is the same as the input frequency related to the LO and IF frequencies.

Using the example in the above section, if the tunable filter has a bandwidth of 10 MHz (the same bandwidth as the IF filter) with a tunable range of 450–550 MHz, it will eliminate most of the spurs. For example, the input signal at $H = 540$ MHz will generate a 100-MHz (the $4L - 3H$) spur with $L = 430$ MHz. If the input filter tracks the LO, when the LO is at 430 MHz, the center frequency of the tunable filter is at 530 MHz (100 MHz higher than the LO) with a bandwidth of 10 MHz. Under this condition, the input signal of 540 MHz will be filtered out, and it cannot mix with the LO to generate a spur to reach the mixer.

Most superhet receivers have a tunable filter in front of the mixer to reduce the spur problem as shown in Figure 5.4. The preselector can be in one of the following forms:

1. For manually tuned superhet receivers, variable capacitors are commonly used. Two variable capacitors are controlled by the same tuning mechanism. One capacitor is used to change the LO frequency and the other one is used to tune the center frequency of the filter. This scheme is often used in AM and FM radios.

2. For receivers with a fixed number of channels, fixed filters and fixed LO frequencies are used. Each time the receiver is tuned to a new channel, a matched filter and an LO frequency pair are chosen. This scheme is often used in channelized receivers because the channel bandwidth may be too wide for a tunable filter to have the same bandwidth. This scheme is also used in television sets.

3. The most commonly used preselector in a superhet receiver is a yttrium iron garnet (YIG) filter combined with a YIG oscillator. The operating principle of the YIG filter and oscillator will be discussed later. The preselector is typically a multistage YIG filter with a bandwidth of between 20 and 70 MHz depending on the operating frequency. A wider bandwidth is available by staggering the

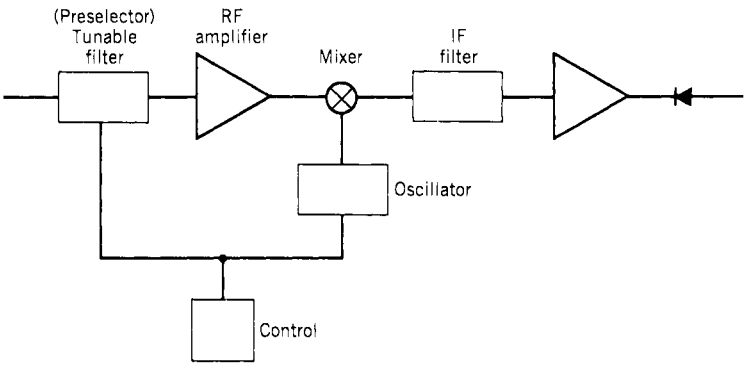


Figure 5.4. Superhet receiver with preselector.

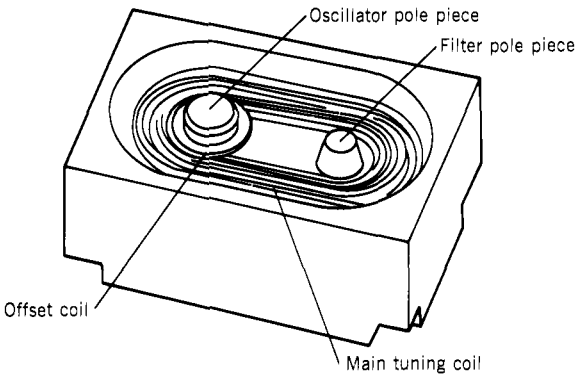


Figure 5.5. Preselector and YIG oscillator in one magnetic housing. (Based on ref. 16.)

resonant frequency of each YIG sphere. The tracking between the filter and the oscillator must be carefully maintained so that the misalignment between the filter and the oscillator will be small compared to the RF filter bandwidth. One of the reasons for mistracking is the different ambient temperatures of the two YIG devices. Temperature compensation schemes can be applied to keep the YIG spheres in the oscillator and the filter in a close temperature range. In some cases, the filter can be adjusted externally through a dc bias current. Periodic adjustment of the filter is required to ensure correct filter center frequency relative to the LO frequency. In some special designs, the YIG spheres of the filter and the YIG spheres of the oscillator are controlled by one single tuning coil in a single housing unit as shown in Figure 5.5. The magnetic circuit consists of a magnetic shell with two pole pieces but only a single main tuning coil. The preselector YIG sphere is mounted in one magnetic gap and the oscillator YIG sphere is located in the other. The main tuning coil tunes the two devices together across the frequency range. A small offset coil can be wound around one of the poles. Current supplied to the offset coil can be used to adjust the tracking between the two devices. Since the YIG spheres are in the same housing, the temperature difference between them is kept at a minimum. This arrangement can cover an operating frequency of 1–18 GHz.

5.5. YIG FILTERS (17–26)

Yttrium iron garnet is a ferrite material. To build a filter, a YIG sphere of typically 20–50 mil in diameter is put in the center of two coils as shown in Figure 5.6*a*. The input coil is in the y–z plane and the output coil is in the x–y plane. These two coils are coupled through the YIG sphere. Without the sphere, the two coils are not coupled because their axes are perpendicular to each other. If a dc magnetic field with proper field strength is applied along the

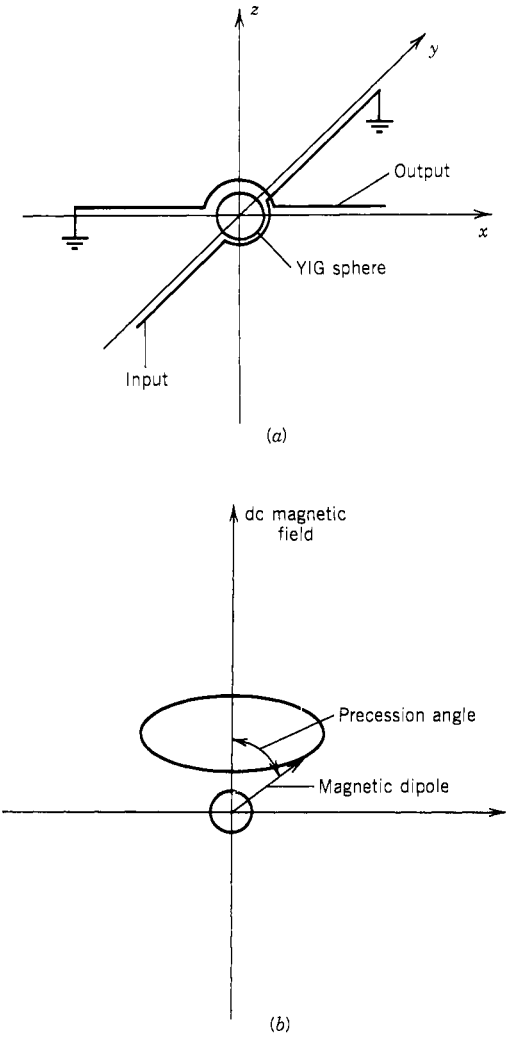


Figure 5.6. A YIG filter: (*a*) single-stage YIG filter; (*b*) ferromagnetic resonance.

z axis and an RF magnetic field is applied in the x–y plane, resonance will occur in the sphere coinciding with the applied microwave frequency as shown in Figure 5.6*b*. As this occurs, the applied RF power is coupled from the input coil to the output coil through the YIG sphere. The characteristic of the filter depends on the coupling between the coil and the sphere and the sphere itself. In a multistage YIG filter, the dc magnetic field containing the YIG spheres must be uniform and the frequency response of each of the spheres must be tracked.

In a YIG filter, when the input power increases to a certain level, the output power stops to follow the increase of input power linearly. Thus, a YIG filter has a limiting level. Under the limiting condition, the precession angle (the angle between the magnetic dipole and the dc magnetic field) of the electrons in the YIG ferrite reaches its maximum and no further increase is possible. Therefore, there is no further increase in coupling between the input and output coils. The output power no longer increases linearly beyond a certain input power. The YIG limiter is also discussed in Sections 6.4 and 7.19.

A typical passband shape of a YIG filter is shown in Figure 5.7. There is more than one resonant mode occurring in a YIG crystal, and these modes lead to spurious resonances. These spurious responses are shown in Figure 5.7.

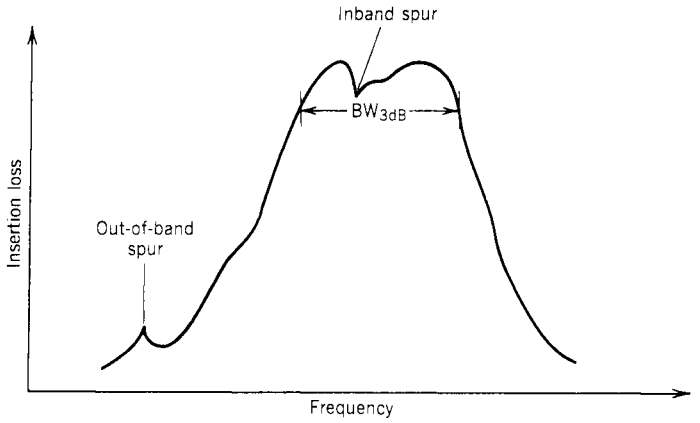


Figure 5.7. Typical response of a YIG filter.

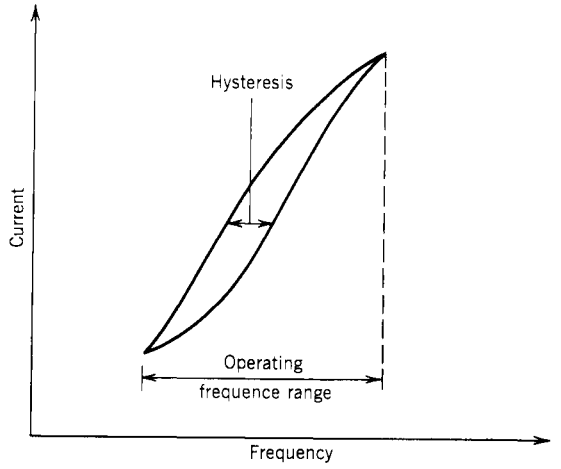


Figure 5.8. Typical hysteresis effect of a YIG filter.

Since the center frequency of a YIG filter is tuned by changing the current in the tuning coil, which in turn changes the magnetic field, the tuning speed of a YIG filter is relatively slow. The tuning time is typically from tens of microseconds to a few milliseconds. Thus YIG filters and YIG oscillators are not suitable for superhet receivers with very fast tuning rates. There is hysteresis effect in tuning a YIG filter (shown in Fig. 5.8). The same current does not produce the same center frequency to the filter. The center frequency of the YIG filter depends on the amplitude of the current as well as on the direction of the tuning.

5.6. RF AMPLIFIERS IN FRONT OF MIXER

If the first element in a superhet receiver is a mixer, the noise figure of the receiver is primarily limited by the insertion loss of the mixer and the noise figure of the first IF amplifier. Further increasing the IF amplification gain beyond a certain limit will not improve the sensitivity of the receiver. The only way to further increase the sensitivity of a superhet receiver is to reduce the insertion loss of the mixer. However, the improvement in performance of a mixer is somewhat limited.

The most effective way to improve the sensitivity of a receiver is to add an RF amplifier in front of the mixer. Although wide-band RF amplifiers are difficult to fabricate, as mentioned in Chapter 4, in superhet receivers the RF bandwidth is usually narrower than that of a crystal video receiver, and using an RF amplifier in front of the mixer is a common practice. Since the input frequency band of a superhet receiver is tunable over the entire frequency of interest, many relatively narrow band RF amplifiers can be used to cover the entire frequency band. Each amplifier is switched in to cover a special input frequency band. Although the RF amplifier in front of the mixer improves the receiver sensitivity, it often degrades the dynamic range of the receiver because the additional gain provided by the amplifier may drive some components to saturation. Therefore, a compromise between sensitivity and dynamic range must be carefully evaluated.

In general, if sensitivity of a receiver is of primary interest, an RF amplifier is often used in front of the mixer. On the other hand, if dynamic range is of primary interest, the RF amplifier is often omitted, and the first component of the receiver is a mixer. Radio-frequency amplifiers are available even in millimeter frequency ranges.

5.7. LOGARITHMIC AMPLIFIERS (26–28)

If a superhet receiver is designed to measure pulse amplitude (PA), in addition to the linear amplifier, an RF log amplifier will be used. Thus, in most superhet receivers, there are usually a linear output channel and a log output channel. In

the log channel, an IF log amplifier is often used. An RF log amplifier is often called an IF log amplifier when it is operated in the IF circuit of a superhet receiver. The term RF is used here to distinguish this amplifier from the video log amplifier discussed in Section 4.8. An RF log amplifier is different from a video log amplifier. The former will take an RF signal as input and generate a video signal, whereas the latter will take a video as input and generate a video signal. In both amplifiers, the outputs are proportional to the log of the input signal.

An RF log amplifier is made of many stages of RF amplifiers. These amplifiers are connected in cascade with a detector at the output of each amplifier as shown in Figure 5.9. The outputs from the detectors are amplified and summed in a log circuit. In this figure, there are limiters at the output of each diode detector. The video limiters eliminate the constraint of designing an RF amplifier stage with controlled limiting characteristic while maintaining the desired linear response. The video limiting will begin while the RF amplifier is still operating in the linear region. Thus this approach is less dependent on the limiting characteristics of the RF amplifiers.

The output from the summing circuit is proportional to the log of the power of the RF input signal. Whereas a weak input signal is amplified through the chain of amplifiers and detected by the last few detectors, a strong signal will be detected by all the detectors. The dynamic range of the receiver can be improved with this arrangement because each detector only covers a portion of the RF range. Experimentally, one detector can approximate an ideal logarithmic with ± 1-dB deviations over a 12-dB range. Therefore, the amplification and dynamic range of each stage must be held below this value to ensure log accuracy. If a 3-dB coupler is used at the output of each stage to couple the energy to the detector, the amplifier in each stage must accommodate this loss.

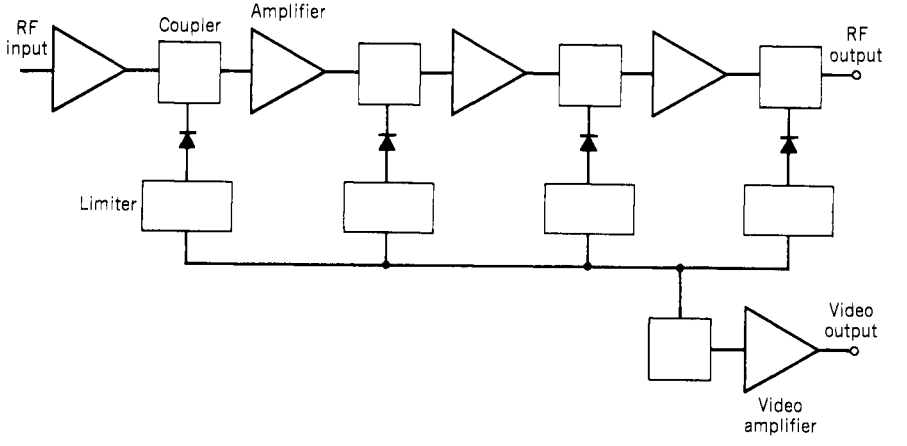


Figure 5.9. An n-stage successive detection RF log amplifier.

The RF log amplifier must contain a sufficient number of stages to provide the desired overall dynamic range.

Since there is a finite delay between the RF amplifier outputs, these delays should sometimes be compensated after each detector. Therefore, a different delay time is introduced at the output of each detector such that the video signals arriving at the log circuit are time coincident.

Another approach to building RF log amplifiers is through RF compression as shown in Figure 5.10. Between each amplifier stage there is a diode limiter to compress the input RF signal. At the final stage, there is a diode detector to convert the RF signal to a video signal related to the log of the input signal. Each stage will convert approximately 10–12 dB to ensure log accuracy and a sufficient number of stages to provide the desired dynamic range. The difference between this RF log amplifier and the former one is that the RF output from the compressive log amplifier is the log of the input signal, whereas the RF output of the successive detection log amplifier is usually limited to a constant level.

As discussed in Section 4.8, the response time of an RF log amplifier is of great concern. Whereas the rise time may be short, the trailing time is often long. The same concern applies to RF log amplifiers. Using detectors with smaller parasitic capacitance may shorten the trailing time. The RF log amplifier can be either ac or dc coupled. The ac-coupled ones can only measure the amplitude of pulsed signals, whereas the dc-coupled ones can measure both pulsed signals and CW signals. The disadvantage of a dc-coupled amplifier is that sometimes the baseline will drift. Occasional readjustment of the output level may be required. The advance in RF log amplifier technology will reduce the drift problem.

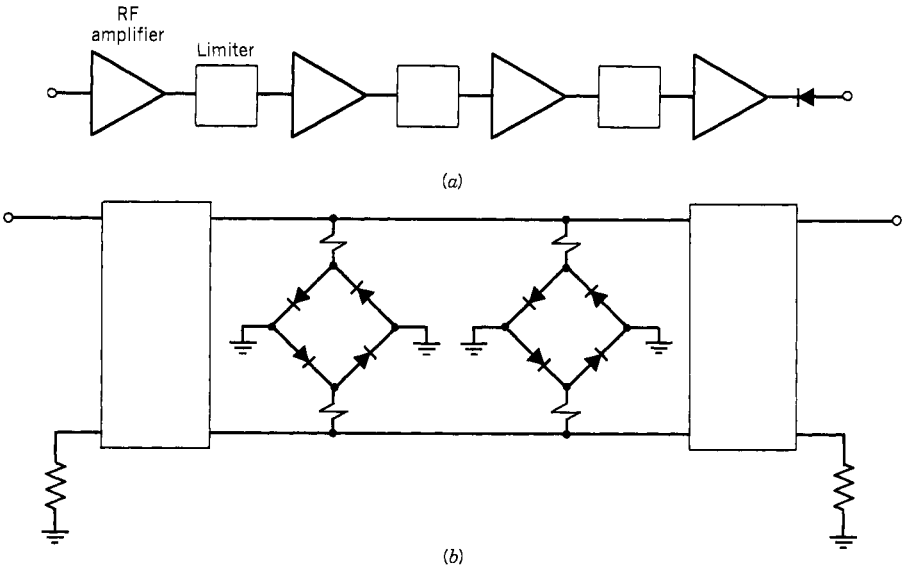


Figure 5.10. An *n*-stage compressive log amplifier. (*a*) log amplifier (*b*) limiter.

5.8. OPERATING PRINCIPLE AND CHARACTERISTICS OF MIXERS (29–45)

The operating principle of a mixer can be described from a nonlinear relation as

$$I = a_0 + a_1 V + a_2 V^2 + \cdots + a_n V^n \tag{5.8}$$

where I is the current flowing through the diode, $a_0, a_1, \ldots, a_n$ are constants, and V is the applied voltage across the diode. This equation can be considered as a power series approximation of the diode equation (4.1).

The applied voltage includes the LO input ($V_L \sin \omega_L t$) and the input signal ($V_s \sin \omega_s t$). Thus

$$V = V_L \sin \omega_L t + V_s \sin \omega_s t \tag{5.9}$$

For simplicity, zero-phase angles are assumed for both the LO and the input signal.

Substituting Eq. (5.9) into (5.8) results in

$$I = a_0 + a_1(V_L \sin \omega_L t + V_s \sin \omega_s t) + a_2(V_L \sin \omega_L t + V_s \sin \omega_s t)^2 + \cdots \tag{5.10}$$

The first two terms in Eq. (5.10) represent a dc component and the original signal, respectively. The third term can be expressed as

$$\begin{aligned} a_2(V_L \sin \omega_L t + V_s \sin \omega_s t)^2 = {} & a_2 V_L^2 \sin^2 \omega_L t + a_2 V_s^2 \sin^2 \omega_s t \\ & + a_2 V_L V_s[\cos(\omega_L - \omega_s)t - \cos(\omega_L + \omega_s)t] \end{aligned} \tag{5.11}$$

The last term in Eq. (5.11) generates the desired frequency shifts $\omega_L - \omega_s$ (which can be represented as $\omega_s - \omega_L$ if $\omega_s > \omega_L$) and $\omega_L + \omega_s$.

Another way of looking at the mixer action is to consider the diodes in the mixer as sampling switches. In general, the LO power in a mixer is higher than the input signal power. As a rule of thumb, the LO power should be 10 dB higher than the highest input signal power anticipated. The LO will control the diode and make it an on–off switch and sample the input signal waveform accordingly. This sampling effect is illustrated in Figure 5.11. For simplicity, the LO input of the diode can be considered as a square wave, and it will control the conductivity of the diode. The output of the diode is shown in Figure 5.11*d*. The Fourier analysis of the diode output will reveal the two important frequency components $\omega_L - \omega_s$ and $\omega_L + \omega_s$. Of course, other frequency components are present in the mixer output.

The conversion loss of a mixer is defined as the ratio of IF output power to signal input power. It measures the efficiency with which a mixer converts RF energy to IF energy. It depends on the losses in the RF circuitry surrounding the mixer, the diodes, and the bias level to the diodes.

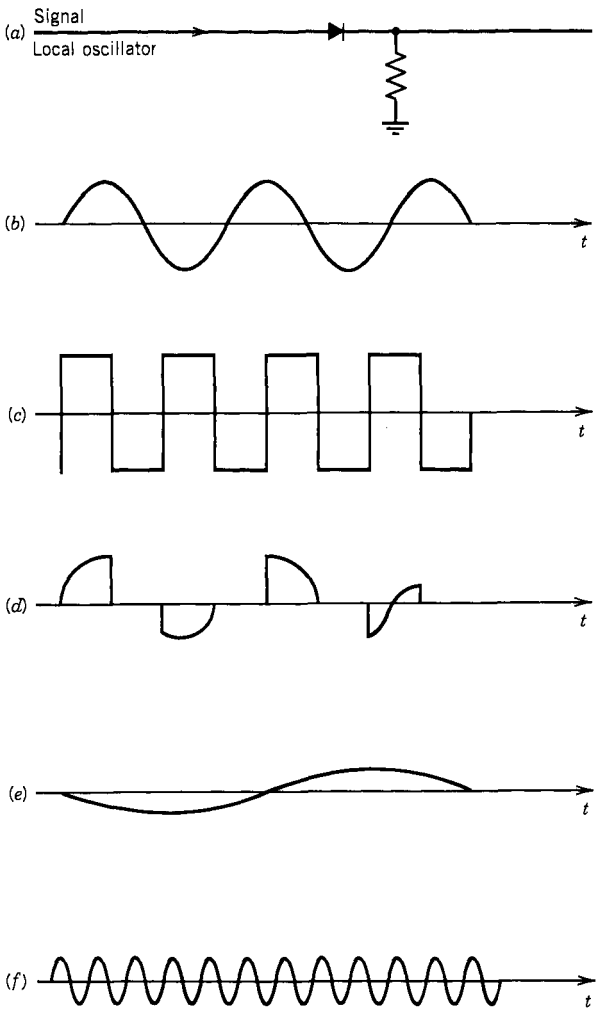


Figure 5.11. Mixer as a sampling switch: (*a*) simple diode mixer; (*b*) input signal (*c*) LO input: (*d*) diode output: (*e*) $\omega_L - \omega_s$: (*f*) $\omega_L + \omega_s$.

The noise figure of a mixer is equal to or slightly greater (0.5 dB) than its conversion loss. Often an IF amplifier follows mixer as an integrated part of the mixer. Under this condition, the noise figure of the mixer should include the noise of the amplifier as

$$\mathrm{NF} = L_c + \mathrm{NF}_{\mathrm{If}} \tag{5.12}$$

where L_c is the conversion loss of the mixer in decibels and $\mathrm{NF}_{\mathrm{If}}$ is the noise of the IF amplifier in decibels. Since the insertion loss is directly related to the

noise figure, it should be kept as low as possible. This is especially important for receivers without an RF amplifier in front of the mixer because the insertion loss limits the noise figure of the receivers.

Another important parameter of a mixer is isolation. There are three ports in a mixer. The isolation measures the power leaking from one port to another one. In many mixers, the isolations between the RF–IF and LO–IF are given by manufacturers on their data sheets because the IF is the desired signal. However, the isolation between the other two ports is also important. Poor RF–LO isolation may lead to the LO radiating through the RF port.

Since the mixer in a receiver is considered to be a linear component, its dynamic range is defined similar to that in an amplifier. The second- and third-order intercept points of a mixer are usually specified. The dynamic range can be calculated by the equations in Sections 2.11–2.15. However, a mixer is not a true linear device. The spurious outputs generated by the mixer must be counted in addition to the spurs generated by the saturation effect.

In Sections 5.9–5.14 different kinds of mixers will be discussed: single-diode mixers, balanced mixers, double balanced mixers, image rejection mixers, image-enhanced mixers, and harmonically pumped mixers.

5.9. SINGLE-DIODE MIXERS (46)

The single-diode mixer is the simplest diode mixer. It contains a diode, a directional coupler, an RF choke, and a low-pass filter as shown in Figure 5.12. The LO power is applied through the directional coupler. The impedance match between the diode and the input transmission line determines the voltage standing wave ratio (VSWR) for the signal input and the amount of LO power appearing at the input of the signal port. The low-pass filter is used at the IF output to block the RF and LO frequencies. The RF choke is needed at the input side of the diode to prevent loss of IF energy in the input port. In this

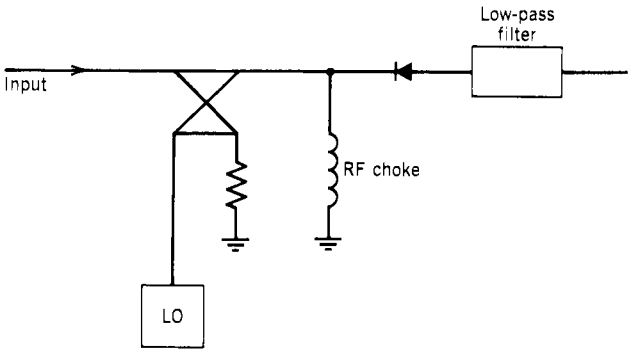


Figure 5.12. Single-diode mixer.

arrangement, the output frequency must be lower than both the RF and LO frequencies.

The performance of a single-diode mixer is relatively poor. The isolation between the RF and LO ports is generally low and the VSWR is relatively high. There is no special circuitry to reduce the amplitudes of spurious responses. More LO power is necessary as a result of the relatively poor coupling factor of the directional couplers. This kind of mixer is seldom used in any receiver.

5.10. BALANCED (SINGLE BALANCED) MIXERS (46–48)

A balanced mixer contains two single-diode mixers and a 3-dB hybrid used to supply the RF and LO power to the two diodes as shown in Figure 5.13. The two diodes are connected in opposite directions. The balanced mixer offers better performance over a single-diode mixer, including reduced spurious responses, cancellation of the dc component at the IF output, and convenient separation of RF and LO inputs.

The outputs of the 3-dB hybrid can be either 90° or 180° out of phase. If the two outputs of the hybrid are 90° out of phase, the RF inputs reflected from the diodes will emerge from the LO port, while the reflected LO power will come from the RF input port. Thus, the input VSWR of the mixer is low (typically less than 1.5) at either of the RF and LO ports. The isolation between the RF and LO ports is rather poor (typically around 7 dB). In general, the harmonics and intermodulations of the RF and LO ports are suppressed. Filters are usually required to minimize the spurious responses.

If a 180° hybrid is used in the balanced mixer, the spurious responses derived from the even harmonics of one of the input signals will be suppressed. It is

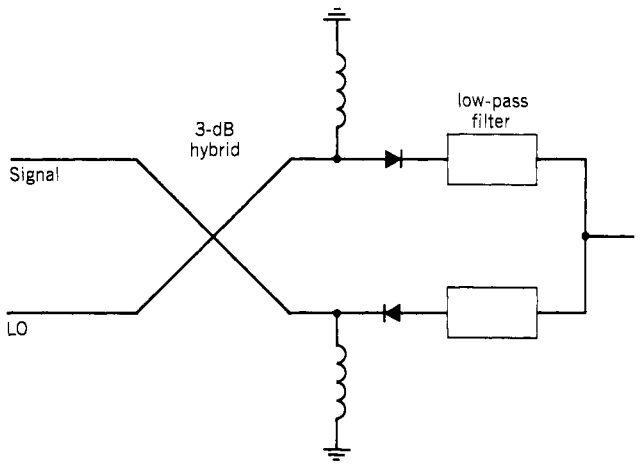


Figure 5.13. Balanced mixer.

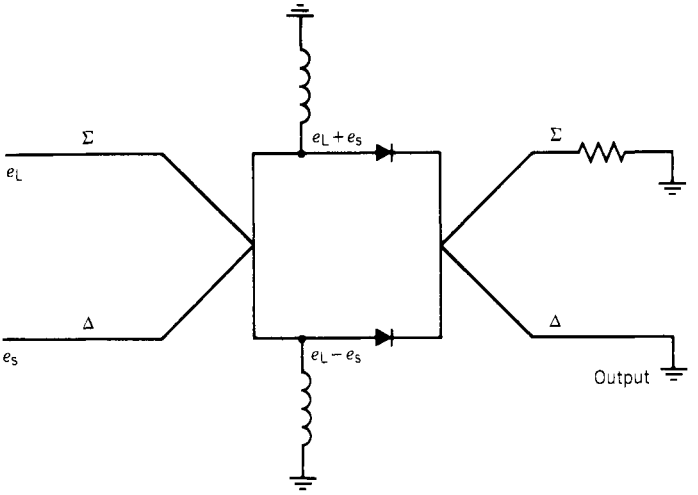


Figure 5.14. Balanced mixer with 180 hybrids.

usually designed to suppress the spurs derived by the even harmonics of the LO signal as shown in Figure 5.14. In this arrangement, the LO power will be in phase at the diodes while the RF signal will be 180° out of phase. At the output of the mixer another 180° hybrid is used. The LO effect will be reduced because the two e_L inputs at the difference terminal (Δ) are 180° out of phase, but the signal e_s is in phase. Reversing the LO and RF input ports, the mixer will suppress the even harmonics of the RF input signal. The degree of supression depends on the matching of the diodes and the balance of the hybrids. The RF input reflected by the diode will emerge from the RF port, while the reflections of the LO input will come from the LO port. As a result, the isolation between the RF and LO of this type of mixer is very good (typically 20 dB or more). This same property, however, causes the VSWR for both the RF and LO ports to depend on the diode match to the transmission line. The VSWR is relatively poor (typically 2). Filters are usually required to reduce the spurious responses.

5.11. DOUBLE BALANCED MIXERS (49–55)

Double balanced mixers using four mixer diodes to suppress spurious responses are the most commonly used in receivers because they provide superior performance over balanced mixers. As early as 1939, Caruthers (ref. 49) discussed various four-diode modulator circuits. The schematic of a double balanced mixer is shown in Figure 5.15. The symmetry provided by the four-diode array formed by uniting two mixers ensures complete isolation between the RF and LO ports when all diodes are perfectly matched. As a result, all three

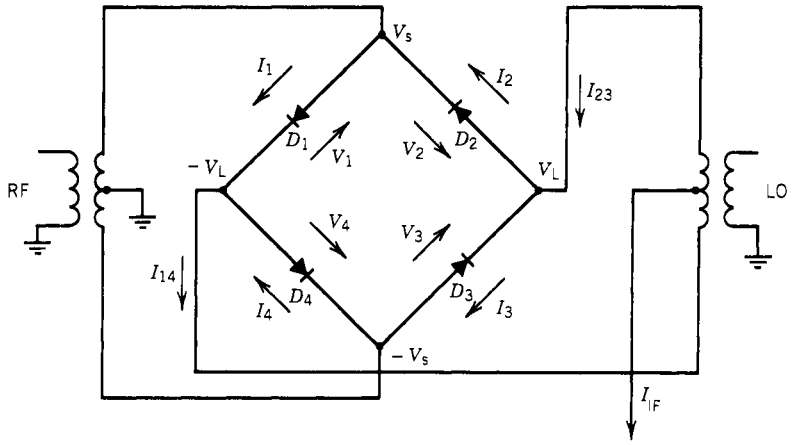


Figure 5.15. Double balanced mixer.

ports (RF, LO, and IF) are well isolated over a wide frequency range. For mixers under 500 MHz, typical isolations are greater than 30 dB. The isolation will deteriorate at higher frequencies because it is difficult to maintain the circuit symmetry required to provide the isolations. The isolations are typically greater than 12 dB for mixers of up to 12 GHz.

Since the circuit in Figure 5.15 is symmetrical, the RF and LO ports are interchangeable. The double balanced mixer suppresses all the spurious responses derived from the even harmonics of both the RF and LO frequencies. The two-tone third-order intermodulation product in a double balanced mixer is lower than that of a single balanced mixer, since less RF voltage appears across each diode for a given input power level. In other words, double balanced mixers have higher dynamic range.

To derive the IF output of the double balanced mixer, assume voltages V_1, V_2, V_3, and V_4 are applied across the four diodes D_1, D_2, D_3, and D_4. The diode currents I_1, I_2, I_3, and I_4 are related to the voltage through Eq. (4.1):

$$I_1 = I_s[\exp(\alpha V_1) - 1] \tag{5.13}$$

$$I_2 = I_s[\exp(\alpha V_2) - 1] \tag{5.14}$$

$$I_3 = I_s[\exp(\alpha V_3) - 1] \tag{5.15}$$

$$I_4 = I_s[\exp(\alpha V_4) - 1] \tag{5.16}$$

where α is the diode slope parameter and I_s is the saturation current. The differential conductance for each diode may be written as

$$g_1 = dI_1/dV_1 = \alpha I_s \exp(\alpha V_1) \tag{5.17}$$

$$g_2 = dI_2/dV_2 = \alpha I_s \exp(2V_2) \tag{5.18}$$

$$g_3 = dI_3/dV_3 = \alpha I_s \exp(\alpha V_3) \tag{5.19}$$

$$g_4 = dI_4/dV_4 = \alpha I_s \exp(\alpha V_4) \tag{5.20}$$

Since, in general, the LO power is higher than the RF power, it is assumed that the conductance of the diodes is modulated by the LO power alone. Thus, for the conductance calculations in Eqs. (5.17)–(5.20), the voltages can be written as

$$V_1 = V_2 = V_3 = V_4 = V_L \cos \omega_L t \tag{5.21}$$

Then

$$g_1 = \alpha I_s \exp(\alpha V_L \cos \omega_L t) \tag{5.22}$$

$$g_2 = \alpha I_s \exp(\alpha V_L \cos \omega_L t) \tag{5.23}$$

$$g_3 = \alpha I_s \exp(-\alpha V_L \cos \omega_L t) \tag{5.24}$$

$$g_4 = \alpha I_s \exp(-\alpha V_L \cos \omega_L t) \tag{5.25}$$

In Eqs. (5.24) and (5.25), the V_L is negative because in diodes 3 and 4 the direction of the diodes and the assumed current are in opposite directions. The current in the diodes can be approximated by

$$I_1 = g_1 V_1 \tag{5.26}$$

$$I_2 = g_2 V_2 \tag{5.27}$$

$$I_3 = g_3 V_3 \tag{5.28}$$

$$I_4 = g_4 V_4 \tag{5.29}$$

To calculate the currents in the diode, the signal will be included in the expressions of voltages. Thus

$$V_1 = V_L \cos \omega_L t + V_s \cos \omega_s t \tag{5.30}$$

$$V_2 = V_L \cos \omega_L t - V_s \cos \omega_s t \tag{5.31}$$

$$V_3 = V_L \cos \omega_L t + V_s \cos \omega_s t \tag{5.32}$$

$$V_4 = V_L \cos \omega_L t - V_s \cos \omega_s t \tag{5.33}$$

Substituting Eqs. (5.17)–(5.20) and (5.30)–(5.33) into Eqs. (5.26)–(5.29), one obtains

$$I_1 = \alpha I_s \exp(\alpha V_L \cos \omega_L t)(V_L \cos \omega_L t + V_s \cos \omega_s t) \tag{5.34}$$

$$I_2 = \alpha I_s \exp(\alpha V_L \cos \omega_L t)(V_L \cos \omega_L t - V_s \cos \omega_s t) \tag{5.35}$$

$$I_3 = \alpha I_s \exp(-\alpha V_L \cos \omega_L t)(V_L \cos \omega_L t + V_s \cos \omega_s t) \tag{5.36}$$

$$I_4 = \alpha I_s \exp(-\alpha V_L \cos \omega_L t)(V_L \cos \omega_L t - V_s \cos \omega_s t) \tag{5.37}$$

The currents I_{23} and I_{14} in Figure 5.15 are

$$\begin{aligned} I_{23} &= I_2 + I_3 \\ &= \alpha V_L I_s[\exp(\alpha V_L \cos \omega_L t) + \exp(-\alpha V_L \cos \omega_L t)] \cos \omega_L t \\ &\quad - \alpha V_s I_s[\exp(\alpha V_L \cos \omega_L t) - \exp(-\alpha V_L \cos \omega_L t)] \cos \omega_s t \end{aligned} \tag{5.38}$$

$$\begin{aligned} I_{14} &= I_1 + I_4 \\ &= \alpha V_L I_s[\exp(\alpha V_L \cos \omega_L t) + \exp(-\alpha V_L \cos \omega_L t)] \cos \omega_L t \\ &\quad + \alpha V_s I_s[\exp(\alpha V_L \cos \omega_L t) - \exp(-\alpha V_L \cos \omega_L t)] \cos \omega_s t \end{aligned} \tag{5.39}$$

The current from the IF port is

$$\begin{aligned} I_{If} &= I_{14} - I_{23} \\ &= 2\alpha V_s I_s[\exp(\alpha V_L \cos \omega_L t) - \exp(-\alpha V_L \cos \omega_L t)] \cos \omega_s t \\ &= 4\alpha V_s I_s[\sinh(\alpha V_L \cos \omega_L t)] \cos \omega_s t \end{aligned} \tag{5.40}$$

Using the formula

$$\sinh(A \cos \theta) = 2 \sum_{k=0}^{\infty} N_{2k+1}(A) \cos(2k + 1)\theta \tag{5.41}$$

where the $N(A)$ are the modified Bessel functions of the first kind of order k. [Here $N(A)$ is used to represent the Bessel function rather than the conventional $I_k(A)$ to avoid confusion with the current I.] Equation (5.40) can be written as

$$\begin{aligned} I_{If} &= 4\alpha V_s I_s[2N_1(\alpha V_L) \cos \omega_L t \cos \omega_s t + 2N_3(\alpha V_L) \cos 3\omega_L t \cos \omega_s t \\ &\quad + 2N_5(\alpha V_L) \cos 5\omega_L t \cos \omega_s t + \cdots] \\ &= 4\alpha V_s I_s\{N_1(\alpha V_L)[\cos(\omega_L - \omega_s)t + \cos(\omega_L + \omega_s)t] \\ &\quad + N_3(\alpha V_L)[\cos(3\omega_L - \omega_s)t + \cos(3\omega_L + \omega_s)t] \\ &\quad + N_5(\alpha V_L)[\cos(5\omega_L - \omega_s)t + \cos(5\omega_L + \omega_s)t] + \cdots\} \end{aligned} \tag{5.42}$$

From Eq. (5.42), it can be seen that the total current at the IF only contains frequency terms $mf_L \pm f_s$, where m is an odd integer. Thus the harmonics derived from the even harmonics are suppressed.

To improve the dynamic range of double balanced mixers, sometimes two diodes connected in series, instead of one diode, are used in each of the four branches of Figure 5.15. Although the intermodulation level of the mixer can be reduced, more LO power is required to drive the mixer, since it must provide twice the voltage to control two diodes in series. This is one of the major disadvantages in high-dynamic-range mixers, especially in a receiver where LO power is scarce.

5.12. IMAGE REJECTION MIXERS (56)

For a given IF frequency ω_I and LO frequency ω_L, there are two RF signals ($\omega_s = \omega_L \pm \omega_I$) that can produce the same IF output, as shown in Figure 5.16. If one of these RF signals is the desired input, the other one is referred to as an image because they are mirror images of each other with respect to the LO frequency. The image frequency only exists in down converters. It does not exist in up converters. The problem caused by the image frequency is that one cannot determine the frequency of the input signal because it can be either higher or lower than the LO frequency. The common solution to this problem is to put a bandpass filter in front of the mixer to pass only one frequency band.

An image rejection mixer will separate the desired RF inputs from their image frequency. The mixer has two IF outputs: one contains the desired signal, the other contains the image. Two simple mixers are combined to form an image rejection mixer, as shown in Figure 5.17. The RF signal is fed to the mixer through a 90° hybrid, and a 90° hybrid is also used to couple the IF out. The LO signal is applied to the mixers through an in-phase power divider. Signals with frequency above the LO frequency will come out from one IF, and signals with frequency below the LO frequency will come out from the other IF. Although the concept of this approach is rather attractive, the performance of the mixer depends on the match between the two mixers.

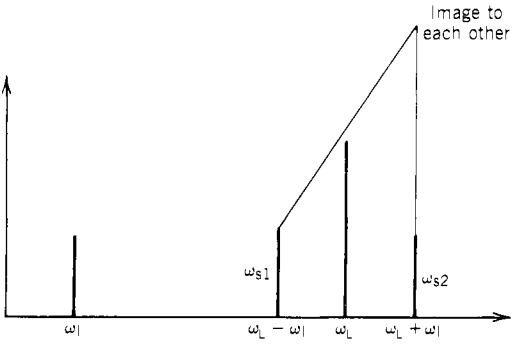


Figure 5.16. Image frequency in a mixer.

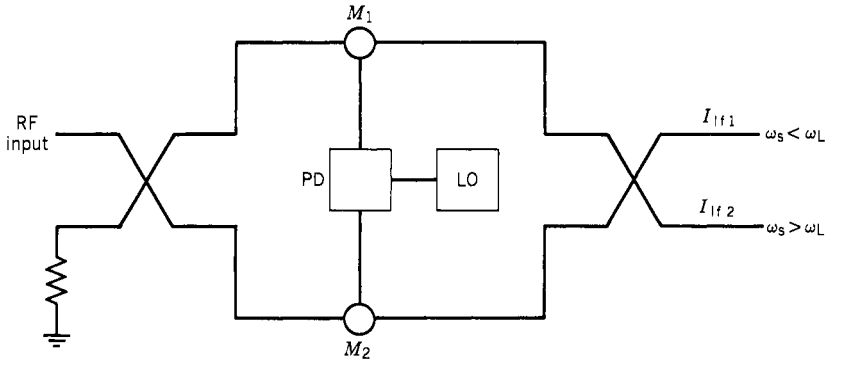


Figure 5.17. Image rejection mixer.

To simplify the analysis, two single-diode mixers are considered, although they are usually balanced mixers. The conductances of the diodes are [see Eq. (5.22)]

$$g_1 = g_2 = \alpha I_s \exp(\alpha V_L \cos \omega_L t) \tag{5.43}$$

where α is the diode slope parameter, I_s is the diode saturation current, and V_L and ω_L are the amplitude and angular frequency of the LO signal, respectively. The voltage across the diodes are

$$V_1 = V_L \cos \omega_L t + V_s \cos \omega_s t \tag{5.44}$$

$$V_2 = V_L \cos \omega_L t + V_s \sin \omega_s t \tag{5.45}$$

where V_s and ω_s are the amplitude and angular frequency of the input RF signal, respectively. The currents at the output of the diode mixer are

$$I_1 = g_1 V_1 \tag{5.46}$$

$$I_2 = g_2 V_2 \tag{5.47}$$

Since only the terms containing $\omega_L - \omega_s$ are of interest, the current I_1 and I_2 can be simplified as

$$I_1 = A \cos(\omega_L - \omega_s)t \tag{5.48}$$

$$I_2 = -A \sin(\omega_L - \omega_s)t \tag{5.49}$$

where A is a constant. The current at the outputs of the second 90° hybrid are:

For $\omega_s < \omega_L$

$$I_{If1} = I_1 + I_2 \exp(-\tfrac{1}{2}j\pi) = A\cos(\omega_L - \omega_s)t - A\sin[(\omega_L - \omega_s)t - \tfrac{1}{2}\pi]$$
$$= 2A\cos(\omega_L - \omega_s)t \tag{5.50}$$

$$I_{If2} = I_1 \exp(-\tfrac{1}{2}j\pi) + I_2$$
$$= A\cos[(\omega_L - \omega_s)t - \tfrac{1}{2}\pi] - A\sin(\omega_L - \omega_s)t = 0 \tag{5.51}$$

For $\omega_s > \omega_L$

$$I_{If1} = A\cos(\omega_s - \omega_L)t + A\sin[(\omega_s - \omega_L)t - \tfrac{1}{2}\pi] = 0 \tag{5.52}$$

$$I_{If2} = A\cos[(\omega_s - \omega_L)t - \tfrac{1}{2}\pi] + A\sin(\omega_s - \omega_L)t$$
$$= 2A\sin(\omega_s - \omega_L)t \tag{5.53}$$

From Eqs. (5.50)–(5.53), it can be concluded that if $\omega_s = \omega_L - \omega_I$, the output appears at I_{If1}, and if $\omega_s = \omega_L + \omega_I$, the output appears at I_{If2}, which means that the signal and its image are separated into two IF outputs. This mixer can be used to improve the instantaneous RF bandwidth of a receiver if two IF processors are used.

5.13. IMAGE-ENHANCED MIXERS (57–62)

An input signal to a mixer will produce a spurious response at its image frequency. Let $\omega_s = \omega_L \pm \omega_I$. Then its image is at $\omega_L \mp \omega_I$. If ω_s mixes with the second harmonics of ω_L, the following results will be obtained:

$$2\omega_L - \omega_s = 2\omega_L - (\omega_L \pm \omega_I) = \omega_L \mp \omega_I \tag{5.54}$$

which indicates that any signal that mixes with the second harmonic of the LO frequency will generate its own image frequency. There is useful energy in the image frequency. This energy at the image frequency can be recovered to produce additional power, thereby decreasing conversion loss. This is the basic operating principle behind the image-enhanced mixer. The insertion loss of the mixer can be reduced by 1–2 dB through the image-enhanced approach.

For a single balanced mixer with a 180° hybrid, the generated image appears at the RF input port. A filter with a bandwidth to pass through only the input signal but that reflects the image energy back to the mixer is placed at the RF input port. By adjusting the electric length between the filter and the mixer, the image can be reflected back to produce IF energy in the proper phase to minimize the mixer conversion loss. This phasing is very critical, since degradation in the conversion loss can result from improper tuning.

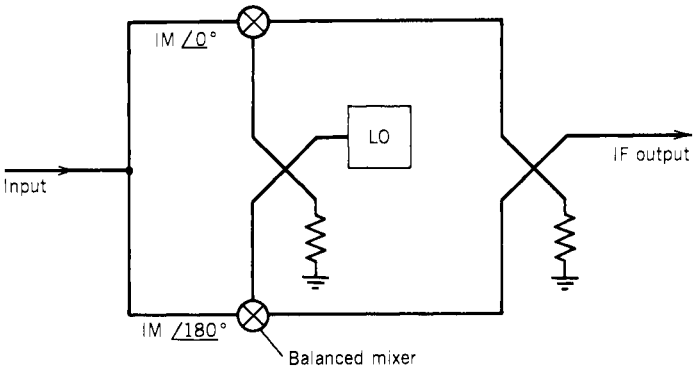


Figure 5.18. Image-enhanced mixer.

The above approach is suitable only for narrow-band mixers. For wide-bandwidth mixers, especially for mixers with overlapping signal and image bandwidths, filtering at the RF input port is not practical. A different approach with two single balanced mixers arranged as shown in Figure 5.18 can reflect the image frequency. The RF inputs are fed in phase to the two mixers, whereas the LO inputs and IF outputs are coupled through 90° hybrids. This configuration caused the image signals generated by the two balanced mixers at the RF input ports to be 180° out of phase with each other. Since the LO frequencies at the mixers are 90° out of phase, the second harmonic of the LO will be 180° out of phase. The image frequencies are generated by mixing the input signal with the second harmonic of the LO frequencies; thus they are 180° out of phase. This results in a maximum image current at the input port and is equivalent to a short-circuit condition for the image current. This short circuit will reflect the image back to the mixer to enhance the desired output. The short-circuit condition is not bandwidth limited because the electric length between the diodes of the two balanced mixers is very short and there is essentially no phase difference between them. Thus this arrangement provides wide bandwidth.

5.14. HARMONICALLY PUMPED MIXERS (48, 63, 64)

Harmonically pumped mixing means that the IF output is generated by mixing the signal and the harmonics of the LO frequency. These outputs are always present in a conventional mixer. They are considered to be spurious responses and are usually suppressed. In the harmonically pumped mixers, they are considered the desired outputs and therefore are enhanced. Harmonic mixing has been used primarily at millimeter wave frequencies where LO sources are either unavailable or prohibitively expensive. Sometimes harmonic mixing is used in superheterodyne receivers with wide input bandwidth and limited LO

tuning range. In the latter application, fundamental mixing is used when the input frequency is low, and harmonic mixing is used when the input frequency is high. The sensitivity of this kind of receiver degrades at higher frequencies as harmonic mixing is used.

If only even-order harmonics are of interest, the antiparallel diode pair arrangement shown in Figure 5.19 can provide the following advantages:

1. reduced conversion loss by suppressing the fundamental mixing products,
2. lower noise figure through suppression of LO noise sidebands,
3. suppression of direct video detection, and
4. inherent self-protection against large peak inverse voltage burnout.

The analysis of the antiparallel diode pair will be discussed in a similar manner as the double balanced mixer. The differential conductance for each diode is

$$g_1 = \alpha I_s \exp(-\alpha V) \tag{5.55}$$

$$g_2 = \alpha I_s \exp(\alpha V) \tag{5.56}$$

The composite differential conductance is

$$g = g_1 + g_2 = 2\alpha I_s \cosh(\alpha V) \tag{5.57}$$

Assuming the conductance is controlled by the LO input only,

$$\begin{aligned} g &= 2\alpha I_s \cosh(\alpha V_L \cos \omega_L t) \\ &= 2\alpha I_s[N_0(\alpha V_L) + 2N_2(\alpha V_L) \cos 2\omega_L t + 2N_4(\alpha V_L) \cos 4\omega_L t + \cdots] \end{aligned} \tag{5.58}$$

where α is the diode slope parameter, V_L and ω_L are the LO voltage and angular frequency, and $N_k(\alpha V_L)$ is the modified Bessel function of the first kind. The

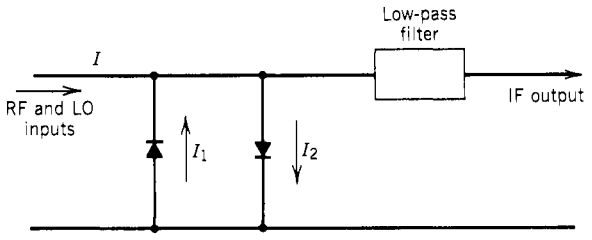


Figure 5.19. Antiparallel diode pair mixer.

current I in the mixer can be written as

$$\begin{aligned} I = gV &= g(V_L \cos \omega_L t + V_s \cos \omega_s t) \\ &= A \cos \omega_L t + B \cos \omega_s t + C \cos 3\omega_L t + D \cos 5\omega_L t \\ &\quad + E \cos(2\omega_L + \omega_s)t + F \cos(2\omega_L - \omega_s)t + G \cos(4\omega_L + \omega_s)t \\ &\quad + H \cos(4\omega_L - \omega_s)t + \cdots \end{aligned} \tag{5.59}$$

It can be seen that the total current only contains frequency terms $mf_L \pm nf_s$ where $m + n$ is an odd integer (i.e., $m + n = 1, 3, 5, \ldots$). The desired signals are the terms containing $2\omega_L + \omega_s$ and $2\omega_L - \omega_s$. In this approach, the fundamental product and all the mixing products containing the odd harmonics of the LO frequency are suppressed. Thus the antiparallel diode pair arrangement can only be used for even harmonic mixing.

5.15. STABILITY OF OSCILLATORS (65–76)

Another important component in a superhet receiver is the LO used to pump the mixers. The LO must have enough power to switch the diodes in the mixer. This section will discuss the stability of LOs. In the following sections, the YIG, voltage-controlled oscillator (VCO), phase-locked loop (PLL), and direct frequency synthesizer will be discussed. In recent superhet receiver developments, emphasis is placed on the LOs. Superhet receivers with various searching and tuning modes have been accomplished through proper LO designs. The general requirements of LOs are fast tuning speed, accurate frequency setting, and good spectrum purity. If special tuning modes are needed, the LO is the major component to be designed to accommodate the requirement.

The stability of an oscillator is very important. Any instability in an LO will be introduced in the IF output and contaminate the signal to be measured. The stability of an oscillator can be divided into two parts: long and short term. The long-term stability is usually expressed as parts per million of frequency change per hour, day, or even year. It represents a reasonably predictable phenomenon that depends on temperature change and the aging of devices used in the frequency control element. The long-term stability is usually measured in the time domain as frequency versus time.

The short-term frequency fluctuations refer to all elements causing frequency changes about the nominal frequency of less than a few seconds duration. Fluctuations of this nature are more conveniently viewed in the frequency domain than in the time domain. In an LO, the short-term stability is more important than the long-term stability.

An ideal sine wave source can be written as

$$V(t) = V_0 \sin 2\pi v_0 t \tag{5.60}$$

where V_0 is the nominal amplitude and v_0 is the nominal frequency. Here v is used to represent the frequency of the oscillator and f will be used to represent the fluctuations of the frequency v or it can be referred to as the Fourier frequency.

To account for the noise components at the output of a real sine wave generator, Eq. (5.60) can be modified as

$$V(t) = [V_0 + E(t)] \sin[2\pi v_0 t + \phi(t)] \tag{5.61}$$

where $E(t)$ is the amplitude fluctuation and $\phi(t)$ is the phase fluctuations. Figure 5.20*a* shows a sine wave that is perturbed for short instance by noise. These types of noise are loosely referred to as "glitches." This same signal can be represented in the frequency domain, as in Figure 5.20*b*. This kind of plot is called the power spectrum. The power spectrum has a high value at v_0 and a low value at v_s, where $v_0 = 1/T_0$ and $v_s = 1/T_s$. Here T_0 and T_s are the time intervals shown in Figure 5.20*a*.

Some noise will cause the instantaneous frequency to "jitter" around v_0 with the probability of being higher or lower than v_0. Usually a "pedestal" is associated with the v_0, as shown in Figure 5.21. A typical RF signal with the sideband spectrum is shown in Figure 5.22. The sideband includes the pedestal as well as some noise of constant repetition.

The power spectrum, often referred to as the RF spectrum, is very useful in many applications. The RF spectrum can be separated into two independent

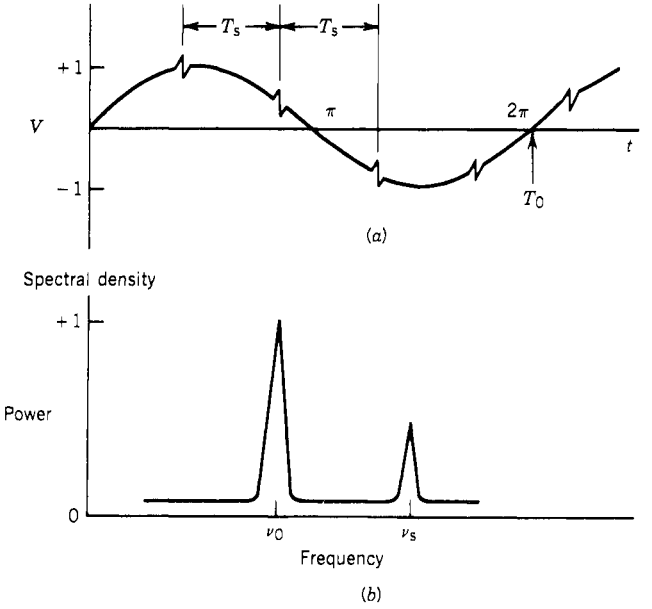


Figure 5.20. Sine wave perturbed for short instances by noise. (*a*) time domain (*b*) frequency domain.

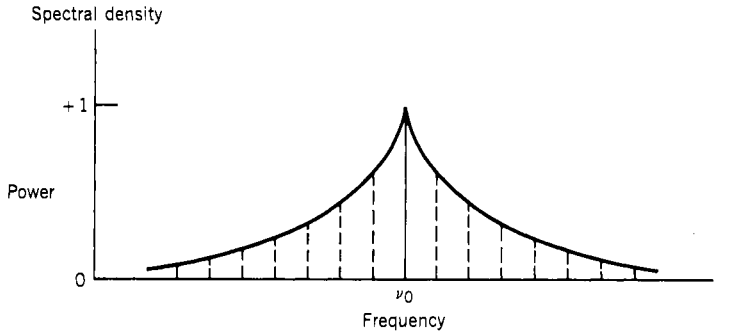


Figure 5.21. Frequency ν_0 with jitter around it.

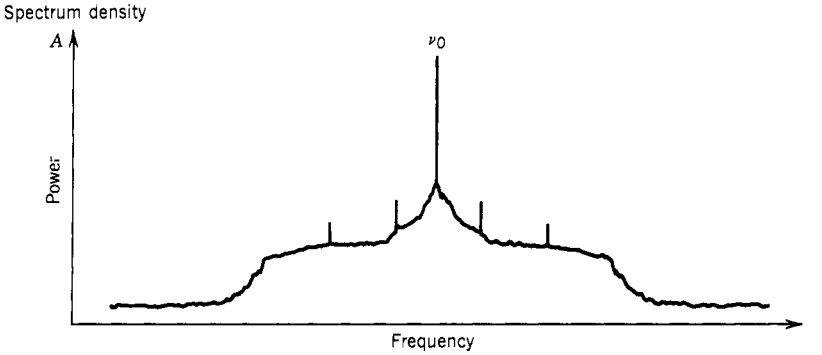


Figure 5.22. Typical RF source with sideband spectrum. (Ref. 69, Courtesy of Hewlett Packard Co.)

spectra: the spectral density of $E(t)$ and that of $\phi(t)$. Unfortunately, in a given RF spectrum, it is impossible to determine whether the power at different Fourier frequencies is a result of amplitude fluctuations $E(t)$ or phase fluctuations $\phi(t)$. In general, the spectral density of $E(t)$ is negligibly small, and the total modulation of the phase fluctuation is also small (mean square value is much less than 1 rad); thus the RF spectrum has approximately the same shape as the phase spectral density. The phase spectral density is the Laplace transform of the phase $\phi(t)$. The main difference in the representation of RF spectrum and phase spectrum is that the RF spectrum includes the fundamental signal (Fig. 5.22) and the phase spectral density does not (Fig. 5.23). Another major difference is that the RF spectrum is a power spectral density and is measured in watts per hertz. The phase spectral density involves no power measurement to the signal and had the unit rad/Hz.

The long-term stability of a typical crystal oscillator is shown in Figure 5.24. The gradual change in frequency over days or months is known as aging. Time domain stability is typically specified for a 1-sec average. Shorter or longer

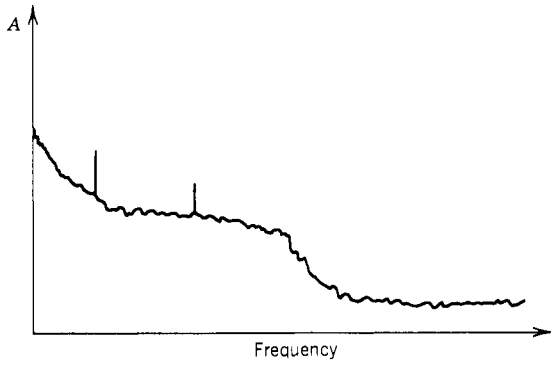


Figure 5.23. Phase spectral density. (Ref. 69, Courtesy of Hewlett-Packard Co.)

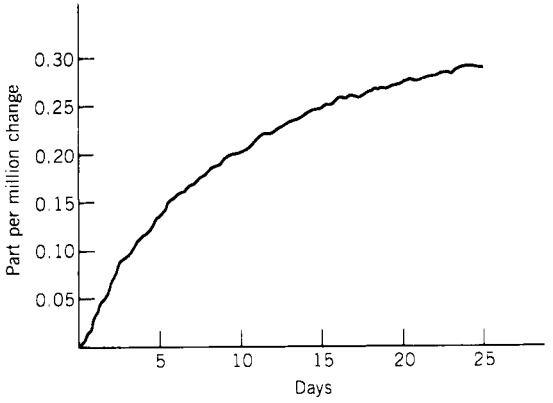


Figure 5.24. Time domain stability of the fractional frequency change over time (days) starting from a point of calibration. (Ref. 73, Courtesy of Hewlett-Packard Co.)

averaging times may be required in an accuracy computation for some applications. A short-term frequency stability in time domain is shown in Figure 5.25. In this figure, the time is from 10^{-3} to 10^{3} sec. The spectral density is shown in Figure 5.26. At 10 Hz away from the carrier, the sideband energy is below -120 dB in comparison with the carrier frequency.

Another effect that causes frequency shift in an oscillator is often referred to as frequency pulling. Frequency pulling is caused by a change in load impedance of the oscillator. When the load of the oscillator changes, the frequency and the output power level may change. A pulling figure is defined by a frequency deviation from the load-matched condition to the VSWR of 1.5. If the load of an oscillator varies over a wide range in a certain system, an isolator is often used at the output of the oscillator to reduce the frequency shift caused by the pulling effect. When the supply voltage to the oscillator varies, the output frequency may also change. This phenomenon is referred to as a pushing effect

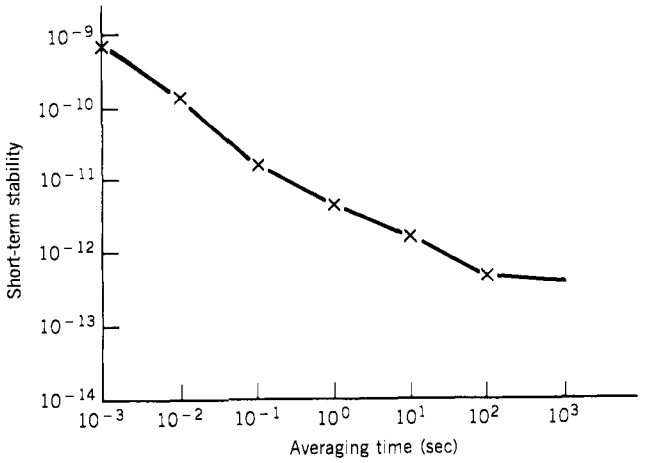


Figure 5.25. Time domain stability (short term) for specific average time. (Ref. 73, Courtesy of Hewlett-Packard Co.)

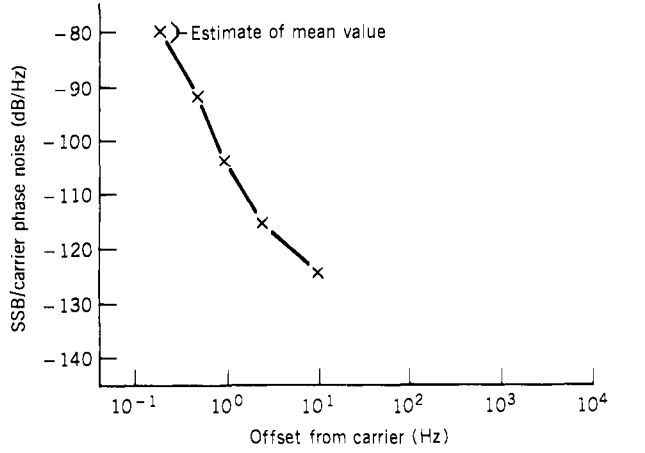


Figure 5.26. Phase spectral density at specific offsets from the carrier. (ref. 73, courtesy of Hewlett-Packard Co.)

and can be expressed in hertz per volts. In some oscillators, internal voltage regulators reduce the pushing effect.

5.16. YIG-TUNED OSCILLATORS (77–84)

The LO of a superhet receiver must be able to be tuned over the same frequency range as the receiver input bandwidth. The most commonly used ones are the YIG oscillators and the VCOs.

When installed in the proper magnetic field with suitable coupling, a YIG

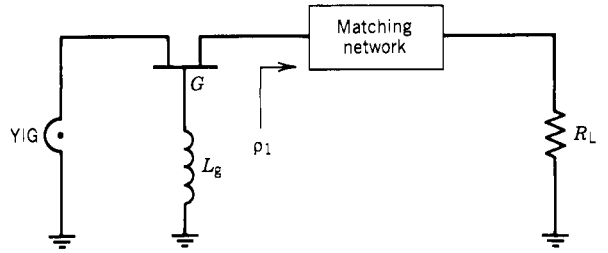


Figure 5.27. Basic YIG-tuned oscillator.

sphere will behave like a tunable microwave cavity with a circuit Q on the order of 1000–8000. In an oscillator, the spectral purity is directly related to the circuit Q. Therefore, when a YIG sphere is used as the frequency-determining element in a negative-resistance oscillator, it provides excellent AM and FM noise characteristics. Since the advance of field-effect transistor (FET) technology, most of the YIG-tuned oscillators of up to 20 GHz use FET devices.

A basic circuit of a YIG-tuned oscillator is shown in Figure 5.27. The YIG sphere used as the tuning element is connected to the source of the FET. It is desirable that an inductive admittance be presented to the oscillator to avoid spurious oscillations with the inductive coupling loop. This requirement limits the FET in a common gate configuration with a common lead inductor (L_g) as the feedback element.

The resonant frequency is linearly related to the magnetic biasing field according to the expression

$$f_0 = \gamma_e H_0 \tag{5.62}$$

where γ_e is the gyromagnetic ratio ($=2.8$ MHz/G where G represents magnetic field in Gauss) and H_0 is the magnetic biasing field. In general, the tuning curves (tuning current versus output frequency) of YIG-tuned FET oscillators are extremely linear and will deviate from the ideal straight line by $\pm 0.05\%$. The power output remains flat within ± 1.5 to ± 3.0 dB over the tuning range.

In many YIG-tuned oscillators, a low-inductance FM coil is added to the main coil. The coil is in close proximity to the YIG sphere and is used to fine tune the oscillator for phase-locked applications or provides frequency modulation. Since tuning the YIG oscillator is accomplished by changing the biasing current, which in turn changes the magnetic field around the YIG sphere, the tuning speed is relatively slow. In general, it takes milliseconds to tune a YIG oscillator to cover its entire frequency range.

5.17. VOLTAGE-CONTROLLED OSCILLATORS (VCOs) (85–99)

In a VCO, a varactor diode usually serves as a voltage-controlled capacitor in a tuning circuit to control the frequency of the oscillator. The active device

in a VCO can be a Gunn, Impatt diode, or transistor with appropriate biasing and feedback circuitry. The major advantage of a VCO is its extremely fast tuning speed, which can be measured in nanoseconds. The limiting factor of the tuning speed is the ability of the external voltage driver circuit to change the voltage across the varactor diode. This is primarily controlled by the driver impedance and the bypass capacitors in the tuning circuit.

Tuning curves for VCOs are basically nonlinear due to both the capacitance versus voltage characteristic of the varactor and the varying impedance of the negative-resistance device. The curve, however, is quite smooth and monotonic (the output frequency is single valued at any voltage and continuously increases for a continuously increasing tuning voltage). To tune the frequency of a VCO very rapidly over a wide range, the Q of the tuning circuit cannot be very high. In general, the Q of a VCO is below 50. The FM noise, varying approximately inversely proportional to the loaded Q of the tuning circuit, will be relatively high. In comparison with a YIG-tuned oscillator, the FM noise of a VCO is about 20 dB higher.

The varactor diode, the tuning element of a VCO, is a semiconductor *pn* junction device under reverse-bias condition. The capacitance is a function of the applied voltage, which can be expressed as (refs. 90, 91)

$$C(V) = \frac{C(0)}{(1 + V/\phi)^{\gamma}} \tag{5.63}$$

where $C(0)$ is the junction capacitance at 0 V, V is the applied voltage, ϕ is the contact potential, and γ is a constant and approximately equal to 0.5. The value of γ depends mainly on the doping of the device. In an LC tuning circuit, the oscillation frequency is

$$f = \frac{1}{2\pi\sqrt{LC}} \tag{5.64}$$

Substituting Eq. (5.63) into (5.64) and using the relations that $V \gg \phi$ and $(V + \phi)^{\gamma} \gg C(0)\phi^{\gamma}/C$, the oscillation frequency can be approximated by

$$f = \frac{V^{0.5\gamma}}{2\pi\sqrt{LC(0)\phi^{\gamma}}} \tag{5.65}$$

If $\gamma = 1$, it implies that the resonant frequency changes with the square root of V. If $\gamma = 2$, as in an hyperabrupt varactor (ref. 90), the frequency will change linearly with the applied voltage.

A typical VCO tuning curve is shown in Figure 5.28. The nonlinear tuning curve can be corrected through a linearizer. The linearizer has an E_0 versus E_i transfer function matching the VCO tuning curve through piecewise linear

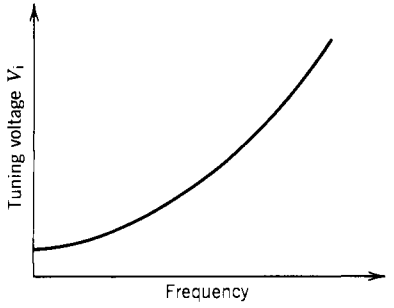


Figure 5.28. Typical VCO tuning curve.

sections as shown in Figure 5.29. Here E_0 is the tuning voltage applied to the VCO. In general, the more the piecewise sections, the more the tuning curve approaches a straight line.

Since the VCO can be tuned extremely fast, the frequency response to the tuning must be evaluated. If the VCO frequency is switched from one value to another, the frequency is difficult to measure during the transient period and it is even hard to define frequency during this time. Therefore, only the initial and final frequencies are of interest. Figure 5.30*a* shows the VCO tuning voltage as a function of time, whereas Figure 5.30*b* shows a typical VCO frequency response versus time. The time Δt_0 may be considered as a settling time, the time required to tune the frequency within a specified tolerance of the final value. The postdrift is shown as Δf and is defined as the frequency change between arbitrarily defined times t_1 and t_2. Time t_1 has been defined from 10 μsec to 1 sec as the short-term post-tuning drift period. Long-term posttuning drift is often defined for the time period from 1 sec to 1 hr.

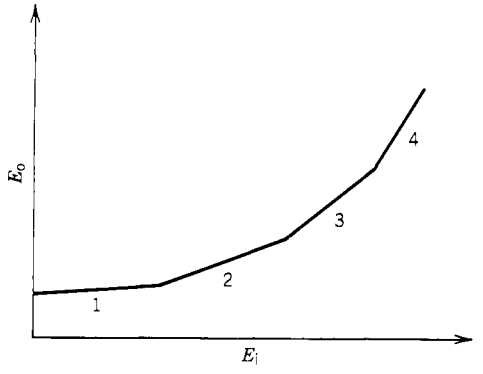


Figure 5.29. Linearizer transfer function.

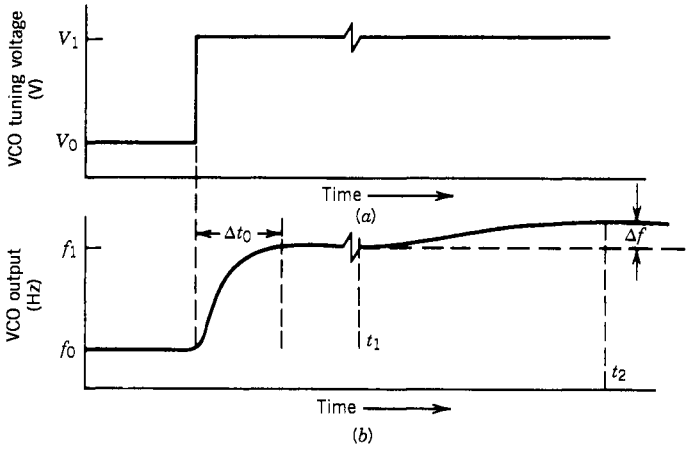


Figure 5.30 Typical VCO frequency versus time response. (Ref. 99, courtesy of Hewlett-Packard Co.)

5.18. OSCILLATORS WITH PHASE-LOCKED LOOPS (100–107)

The frequency of the YIG-tuned oscillators and VCOs discussed in the above sections can be tuned to any desired value within the frequency range of the device. However, in this open-loop tuning scheme, the accuracy and stability of the frequency is limited by the loaded Q and the control circuits. If high-frequency accuracy and stability are desired, the tunable oscillator must be phase locked to a stable oscillator (i.e., a crystal-controlled oscillator). This scheme is called frequency synthesizing. A frequency synthesizer can translate the stable frequency of a precision frequency standard to one of thousands or millions of frequencies over a spectrum range. Two methods have been used in frequency synthesis: direct and indirect. Direct synthesis is accomplished by multiplexing, dividing, and frequency mixing, whereas the indirect method derives its frequency from one or more phase-locked loops (PLLs). The stability and accuracy in either case is derived from the reference sources.

A PLL is a feedback system whose function is to force the output frequency of a tunable oscillator to be coherent with a certain reference frequency. "Coherent" means highly correlated in both frequency and phase. The basic idea is to compare the frequency and phase of the reference signal to the output of the tunable oscillator. If the two signals differ in frequency and/or phase, an error voltage is generated and applied to the oscillator, causing it to correct in the direction required to decrease the difference. The correction procedure continues until phase lock is achieved; after that the oscillator will continue to track the reference signal. The general requirements of a PLL is that it must very quickly lock up to the reference signal and yet be capable of ignoring short-

term jitter inadvertantly introduced to the reference. When an oscillator with a PLL is first turned on, the oscillator should be tuned through some other means to a frequency close to the desired value so that the PLL can fine tune the oscillator to the desired value.

A simple PLL is shown in Figure 5.31. The output frequency f_o from the oscillator is divided by N through a programmable divider, which means the value of N can be changed with external controls. A phase detector compares the instantaneous phase difference between the reference frequency f_r and f_o/N. To compare the phase, f_r must equal f_o/N or $f_o = Nf_r$. Therefore, when N changes, the output frequency f_o changes. For example, if $f_r = 100$ kHz and $N = 99$–198, the output frequency f_o can be changed from 9.9 to 19.8 MHz in 100 kHz steps. To generate higher frequencies, a frequency multiplier can be placed at the output of the oscillator output as shown in Figure 5.31.

In microwave applications, there is no readily available divider to divide the microwave frequency. In this case, the most common PLL is the one shown in Figure 5.32. There are two input frequencies f_r and f_m. If the output frequency f_o is greater than f_m, then the mixer output frequency $f_d = f_o - f_m$. The difference frequency f_d is low enough that a frequency divider can divide

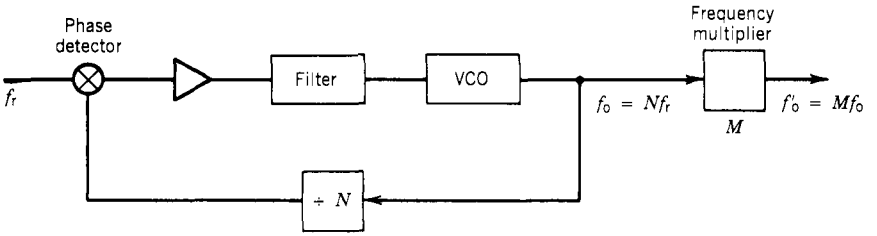


Figure 5.31. Basic phase-locked loop.

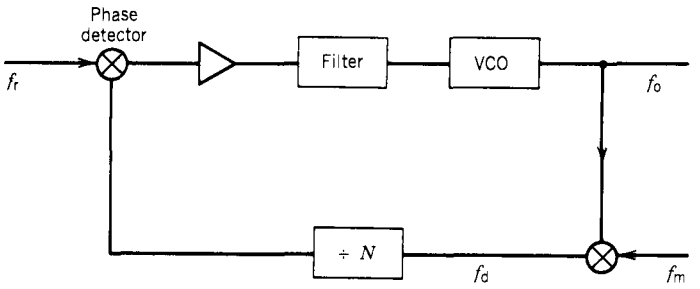


Figure 5.32. Phase-locked loop by frequency mixing.

it. For the phase detector to work, the following condition must be held:

$$f_r = (f_o - f_m)/N \tag{5.66}$$

or

$$f_o = f_m + Nf_r \tag{5.67}$$

To keep f_o stable, both f_r and f_m should be stable. Frequency f_m can be generated by multiplying the output frequency of an oscillator with a PLL as shown in Figure 5.31. In such an arrangement, two PLLS will be used, but the frequency f_m can be one of many values generated from the reference f_r, and the output frequency f_o is more versatile.

5.19. DIRECT FREQUENCY SYNTHESIZERS (108)

The frequency switching speed of oscillators stabilized with the PLL discussed in Section 5.18 is limited by the loop bandwidth, which seldom exceeds a few megahertz. In some applications, faster switching speed is needed and a direct synthesizer may satisfy these requirements.

The basic idea of a direct synthesizer is that there are N oscillators operating simultaneously. The outputs of these oscillators are connected to an N to 1 switch. The desired frequency is selected through the switch. The selection speed is primarily determined by the operating speed of the switch, which is fast. However, this kind of arrangement is very expensive and impractical for use to generate many frequencies. A practical approach to build a direct frequency synthesizer is to generate frequency by successive frequency mixing, filtering, and dividing–multiplexing operations. This modified approach is quite similar to the basic direct synthesizer design. Many fixed-frequency sources are available all the time. The output frequency is obtained by selecting the proper frequencies as the inputs of certain mixers to generate the desired value. As discussed earlier in this chapter, there are many spurious frequencies generated by a mixer. It is important to keep the undesired frequencies below a certain specific level. The frequency selection speed depends on the RF switching speed and the RF propagating through the filters. In general, the operating speed of a direct synthesizer ranges from tens to hundreds of nanoseconds.

A basic direct frequency synthesizer is shown in Figure 5.33. A stable reference oscillator is fed through a harmonic generator to produce many harmonic-related frequencies. These frequencies are then separated by a bank of filters. This approach is just one way to produce many simultaneous frequencies. Of course, these frequencies can be generated by other means. These frequencies are fed into a switching matrix. The switching matrix output frequencies f_x and f_y can be separately selected by external logic control. The f_x output is fed directly to a mixer, and the f_y frequency is divided by N and then fed to the

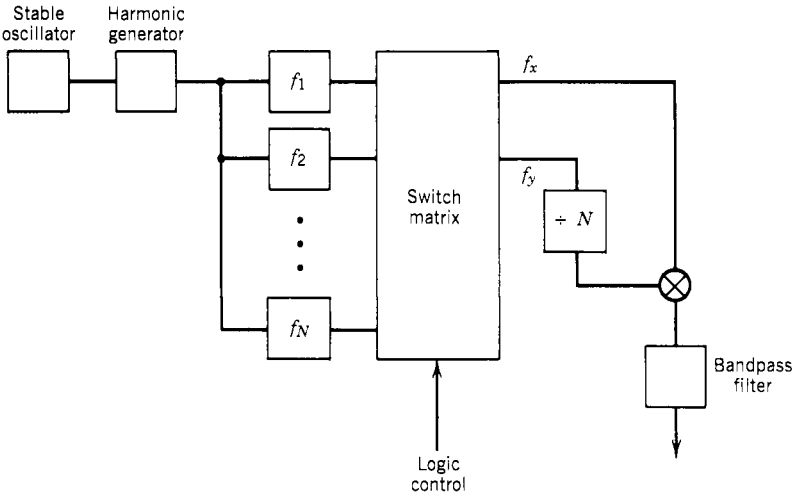


Figure 5.33. Basic direct frequency synthesizer. (Based on ref. 108.)

other input of the mixer. The output of the mixer can be selected as $f_x + f_y/N$, which can be any one of the N frequencies. The spurious frequencies generated by the mixer can be filtered out by the output bandpass filter if they are outside the frequency of interest. However, if the spurs are in the frequency band of interest, they cannot be filtered out. Thus the spurious frequencies generated in this type of synthesizer can be high.

The above discussion clearly reveals that the direct frequency synthesizer requires many frequency sources. These sources must be on all the time, although they will not be used all the time. As a matter of fact, they will be used only a small fraction of the time. Therefore, the direct frequency synthesizer design is very complex, including a large number of amplifiers, mixers, multipliers, dividers, and filters. The efficiency of a direct synthesizer is rather poor.

From the above discussion, it is obvious that in both the indirect and direct frequency synthesizers the output frequency is discretely rather than continuously tuned.

5.20. CRYSTAL BANDPASS FILTERS (109–112)

To measure the input signal frequency accurately, the frequency of the LO must be accurate. In addition, the center frequency of the IF filter must be accurate and the filter bandwidth must be narrow, especially for measuring CW signals. The most commonly used IF filters are crystal filters. A crystal filter is an electric bandpass filter in which at least one of the components is piezoelectrically resonant. Since a resonator is a mechanical device, the operational frequency of the crystal filter is generally below 100 MHz. That is why many superhet receivers use 60 MHz as the IF processing channel.

The common material used in a crystal filter is quartz, which has a very high Q. The equivalent circuit of a crystal resonator is shown in Figure 5.34. The operational frequency of the quartz crystal resonator depends on the crystal cut. The operational frequency of a quartz crystal ranges from 1.2 kHz to 100 MHz.

The crystals can be incorporated in the filter design as resonant elements. One way of using the crystal in a filter is shown in Figure 5.35. This kind of circuit is also referred to as a lattice structure. The circuit can be analyzed as a generalized lattice network with the equivalent circuit in Figure 5.34 representing the crystals. Figure 5.35 shows the schematic of a single section of a filter, and similar sections can be connected in cascade to provide the desired characteristics of a crystal filter.

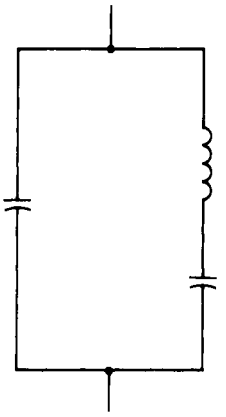


Figure 5.34. Equivalent circuit of a crystal resonator.

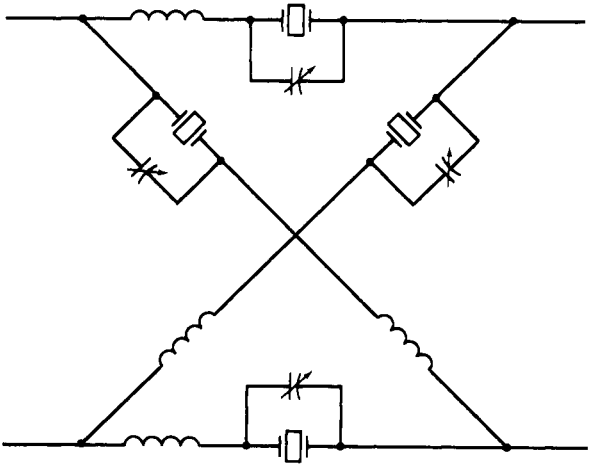


Figure 5.35. Lattice circuit for filter with crystals and series inductance in each arm.

5.21. EXAMPLE OF A SUPERHETERODYNE RECEIVER (113, 114)

In this section, an example of a superhet receiver will be presented. The example is based on a Watkins and Johnson Co. TN1000, a RF to IF converter. The input of the converter is from 0.5 to 18 GHz. The output is centered at 1 GHz with a bandwidth of 500 MHz. The RF to IF gain is 15 dB, and the noise figure is under 20 dB.

The block diagram is shown in Figure 5.36. The input is divided into six channels through a 1 to 6 microwave switch. One of the channels covers 0.5–8 GHz, which is referred to as the low band. The 8–18-GHz range, which is referred to as the high-frequency channel, is divided into five 2-GHz channels.

The input from the low-frequency channel is filtered through a low-pass filter with a cutoff frequency of 8.25 GHz and mixed by mixer M1. The LO input to M1 is tunable from 10.5 to 18 GHz. The output of M1, which is centered at 10 GHz, is fed through a bandpass filter to reduce the spurious frequencies generated by M1. This 10-GHz output is down converted to the desired 1 GHz by M2 and is fed to the input of a single-pole double-throw (SP2T) output switch. In this portion of the design, the input is first up converted to 10 GHz and then down converted to 1 GHz to minimize spurious frequencies generated by the mixers.

The high-frequency channel is divided into five 2-GHz subchannels. The outputs of these subchannels are selected by a SP5T switch and fed into mixer M3. The LO frequency to M3 is divided into two tuning bands. For an input in the 8–14-GHz range, the LO is tuned from 11.5 to 17.5 GHz, and for an input in the 14–18-GHz range, the LO is tuned from 10.5 to 14.5 GHz. The output from M3 is at a fixed frequency of 3.5 GHz, with a bandpass filter to reduce the spurious frequencies generated by mixer M3. This 3.5-GHz frequency is down

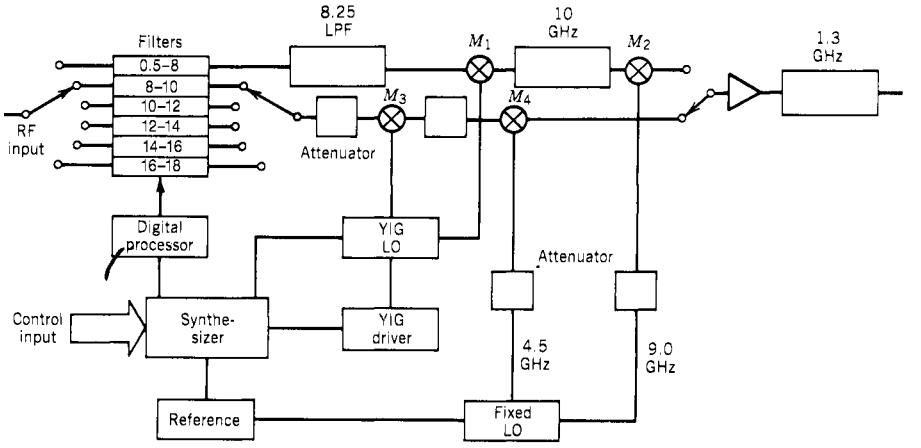


Figure 5.36. Block diagram of RF to IF converter.

converted to 1 GHz through M4. The output of M4 is fed to the input of the SP2T output switch.

The output from the final 2–1 switch is further filtered through a 1.3-GHz low-pass filter, and amplifiers are added to provide the required gain. To improve the VSWR, isolators at the inputs and outputs of the mixers are sometimes added.

The digital processor in Figure 2.36 performs the following functions: controls bus interface, generates YIG tuning voltage, and controls band switching. The RF to IF converter is controlled by external digital words. The digital processor will receive external commands and act accordingly.

The reference source is a crystal oscillator at 100 MHz used as the reference frequency to generate all the LO frequencies in the system. The fixed LO source unit takes the 100-MHz reference signal and is phase locked to it to generate a 4.5-GHz frequency as the LO for mixer M4. In this unit, there is also a frequency doubler that doubles the 4.5 GHz to provide a 9-GHz frequency as the LO for mixer M2.

The synthesizer, the YIG driver, and the YIG oscillator work together to provide the tuning LO frequency covering the range 10.5–18 GHz. The synthesizer takes the digital command words from the digital processor as input and D/A converts them to a voltage to tune the YIG LO to a desired frequency. It also takes the reference and the sampled YIG LO frequencies as inputs to phase lock the LO frequency to the reference. After the synthesizer tunes the LO close to the desired frequency, the sampled LO frequency and the reference frequency are compared through several PLLS to fine tune the LO to the desired frequency. The final LO frequency is phase locked to the reference frequency, a crystal-stabilized source. Therefore, the tuning frequency of the LO is stabilized. The minimum tuning step of the frequency synthesizer is 500 kHz.

In this design, there are no RF amplifiers in front of the mixers. Although adding amplifiers in front of the mixers can improve the noise figure of the RF to IF converter, it will reduce its dynamic range.

If this converter is used as a superhet receiver, the minimum frequency resolution is about 500 MHz. If fine-frequency resolution is desired, the common approach is to down convert the output frequency to a lower frequency, that is, 60 MHz, and add a narrow-band filter at the output to measure the input frequency. The frequency resolution is determined by the bandwidth of the last filter. The minimum bandwidth of the last filter should be wider than the minimum tuning step, which is 500 kHz. Finer frequency resolution can be achieved, if fine tuning is provided in the receiver following this converter.

The tuning of a superhet receiver to automatically lock on an input signal can be rather tricky. The output of the IF filter can be monitored, and this output is then used to tune the receiver to maximize the IF output. If the input is a CW signal, this approach should perform satisfactorily. If the input is a pulsed signal, the receiver usually cannot respond fast enough within the pulse. If the input is a pulse train of the same frequency, the receiver can search and

lock on some pulses at a later time. If the input is a pulsed signal, additional difficulties in amplitude measurements will occur because of the leading and trailing edge transient effects. Transient effect will be discussed in Chapter 7.

5.22. SCANNING SUPERHETERODYNE RECEIVERS

Because of its narrow input bandwidth, a superhet receiver has the highest sensitivity and dynamic range of all EW receivers. However, this narrow bandwidth will severely limit the POI. To cover a wider input bandwidth, the receiver can be made to scan a certain bandwidth at a fast rate repeatedly. This kind of receiver is often referred to as a scanning superhet receiver. Strictly speaking, a scanning superhet receiver does not improve the POI, but it covers a wider bandwidth in a regular rate.

The fastest scan rate achievable without sacrificing the sensitivity of the receiver is approximately B^2, where B is the bandwidth of the final IF filter. This relation can be reasoned as follow. If a filter bandwidth is B, the inverse response time is $1/B$. The fastest rate it scans the bandwidth B is B^2 [or $B/(1/B)$]. If the IF bandwidth of a receiver is 10 MHz, the fastest scan rate is 100 MHz/μsec. If the scan rate is faster than this value, the sensitivity of the receiver will be degraded. If the input bandwidth is 10 GHz, it will take the receiver 100 μsec to scan across the whole band without degrading the sensitivity.

A scanning superhet receiver can be used as a special receiver to intercept long pulsed signals or CW signals. If an integration circuitry with proper time constant is used at the output of the video detector, the receiver can be made very sensitive to CW or high-duty cycled signals. It should be emphasized that the sensitivity improvement in this manner is caused only by the integration circuitry after the video detector. If a proper threshold is chosen, the receiver will not respond to short pulses but will only detect the existence of CW or high-duty cycled signals.

5.23. HOMODYNE RECEIVERS

A homodyne receiver can be considered as a special superhet receiver. This kind of receiver also has a mixer and LO to down convert the input signal to a lower IF. The main difference between a superhet and a homodyne receiver is in the LO. In a homodyne receiver, the LO frequency is derived from the input signal as shown in Figure 5.37. In this figure, the input signal is divided into two paths. One portion of the signal is used to mix with an oscillator frequency f_I to generate the LO frequency $f_R + f_I$. This LO frequency is then mixed with the input signal f_R, again to produce an IF signal f_I. Since the output from mixer 2 usually covers a wide frequency range, it is impossible to place a filter at the output of the mixer to reduce the spurs generated by mixer 2. Therefore,

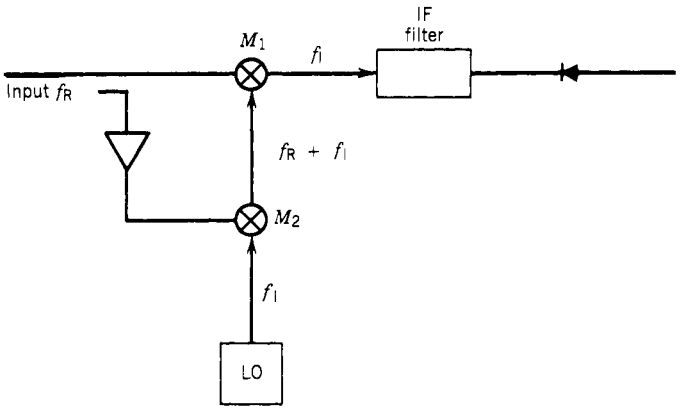


Figure 5.37. Basic homodyne receiver.

mixer 2 must be a special mixer that will suppress the output frequency $f_R - f_I$. Sometimes this kind of mixer is also called an image rejection mixer.

If there is no input signal, the two mixers will isolate the fixed frequency f_I from reaching the IF filter. When there is an input signal of any frequency f_R, a signal will be detected at the output of the IF filter. This receiver can cover a wide input frequency range. However, it does not provide frequency information on the input signal. Therefore, a homodyne receiver will function like a crystal video receiver that can only detect the existence of an input signal.

5.24. SOME POSSIBLE APPLICATIONS OF HOMODYNE RECEIVERS

Since a homodyne receiver cannot measure the frequency of an input signal, it has limited use. Thus a homodyne receiver is not widely used in EW applications. One of the potential applications is to detect the existence of spread spectrum signals. For example, signals of wide bandwidth, that is, chirp signals and phase code signals, are difficult to detect by conventional receivers because only a portion of the signal may enter the input bandwidth of an intercept receiver. Even in a receiver that covers a wide instantaneous bandwidth, the input signal may be spread into many output channels, that is, in a channelized receiver, and the output from each channel is still low. In a homodyne receiver, no matter how the input signal varies, the LO frequency always has the same variation. Therefore, the output of mixer 1 will always be a simple CW signal at frequency f_I. A signal of wide bandwidth will be detected as an ordinary pulsed signal with a simple carrier frequency.

Another possible application of a homodyne receiver is to use a narrowband circuit to measure an input signal over a wide frequency range. For example, if a wide-band, high-dynamic-range RF log amplifier is difficult to

build, one can design a narrow-band RF log amplifier with center frequency at f_1 to measure signal amplitudes at the output of a homodyne receiver. However, in such an application, one must carefully evaluate the overall design, including the homodyne receiver.

5.25. SUMMARY

The performance of a superhet receiver can be considered excellent in every aspect except the low POI. Its best feature is that a superhet receiver can be used to separate and isolate a signal of interest and to perform a fine-grain analysis on the signal. It is the most commonly used receiver in many microwave areas. Superhet receivers are most commonly used in communication and radar applications.

In EW applications, a superhet receiver is often used in conjunction with some other types of wide-band receivers. The wide-band receiver can be used as a cueing receiver to find the frequency of the signals of interest. The superhet receiver can then be tuned to the desired frequency to do a fine-grain analysis on the signals. One such application is to measure the AOA information of an input signal through an interferometric system. Because of the relatively narrow bandwidth, phase matching between different superhet receivers is easier to achieve than in a wide-band receiver. In an amplitude comparison AOA system, the superhet receivers are also commonly used.

Recent developments in superhet receivers are in the areas of frequency synthesizers and control circuitry. The synthesizer can improve the tuning accuracy and the tuning speed of a superhet receiver. The control circuitry can improve the quality of the signals displayed, signal processing, and receiver versatility. The advances in microwave components and logic circuits will improve the receiver performance and reduce size.

REFERENCES

1. S. N. Van Voorhis, *Microwave Receivers*, Radiation Laboratory Series, Vol. 23, McGraw-Hill, New York, 1948.
2. M. I. Skolnik, *Radar Handbook*, Chapter 5, McGraw-Hill, New York, 1970.
3. P. J. Meier, H. E. O'Kean, and E. W. Sard, Integrated X-band sweeping superheterodyne receiver, *IEEE Trans. Microwave Theory Techniques*, **MTT-19**, 600–609 (1971).
4. E. J. Crescenzi, Jr., R. W. Oglesbee, and R. A. Chappell, Integrating components for new front-end design, *Microwaves*, **13**, 35 (August 1974).
5. C. E. Dexter and R. D. Glaz, HF receiver design, Watkins Johnson Co. Tech Notes, **5**(2) (September/October 1977).
6. T. T. Brown, Mixer harmonic chart, *Electronic Buyer's Guide*, pp. R-58, 59, June 1953.
7. M. Y. Huang, R. L. Buskirk, and D. E. Carlile, Select mixer frequencies painlessly, *Electronic Design*, **25**, 103–109 (April 12, 1976).
8. H. W. Pollack and M. Engelson, An analysis of spurious response levels in microwave receivers, *Microwave J.*, **5**, 72 (December 1962).

9. R. Nitzberg, Spurious frequency rejection, *IEEE Trans. Electromagnetic Compatibility*, **EMC-6**, 33–36 (1964).
10. L. M. Orloff, Intermodulation analysis of crystal mixer, *IEEE, Proc.*, **52**, 173–179 (1964).
11. J. H. Lepoff and A. M. Cowley, Improved intermodulation rejection in mixers, *IEEE Trans. Microwave Theory Techniques*, **MTT-14**, 618–632 (1966).
12. J. T. Herishen, Diode mixer coefficients for spurious response pridiction, *IEEE Trans. Electromagnetic Compatibility*, **MEC-10**, 355–363 (1968).
13. A. M. Yousif and J. G. Gardiner, Multifrequency analysis of switching diode modulators under high level signal conditions, *Radio Electronic Engineer*, **41**, V.D.C. 621.372.233 (January 1971).
14. E. F. Beane, Prediction of intermodulation levels as function of local oscillator power, *IEEE Trans. Electromagnetic Compatibility*, **MEC-13**, 56–63 (1971).
15. Z. I. Bogan and J. B. Y. Tsui, Computer modeling of diode circuits and its applications, Air Force Avionics Laboratory Technical Report, AFAL-TR-77-266, March 1978.
16. J. C. Papp and R. V. Jackson, YIG-tuned integrated devices, Watkins-Johnson Co. Tech Notes, **4**(5) (September/October 1977).
17. R. E. Tokhein and J. C. Hover, The four port YIG filter, *Microwave J.*, **13**(9) (September 1970).
18. P. S. Carter, Magnetically-tunable microwave filter using single crystal yttrium-iron-garnet resonators, *IEEE Trans. Microwave Theory Techniques*, **MTT-9**, 252–260 (1961).
19. H. Skeie, Nonreciprocal coupling with single-crystal ferrites, *IEEE Trans. Microwave Theory Techniques*, **MTT-12**, 587–594 (1964).
20. R. L. Comstock, Synthesis of filter-limiters using ferrimagnetic resonators, *IEEE Trans. Microwave Theory Techniques*, **MTT-12**, 599–607 (1974).
21. P. S. Carter, Equivalent circuit of orthogonal-loop-coupled magnetic resonance filters and bandwidth narrowing due to coupling inductance, *IEEE Trans. Microwave Theory Techniques*, **MTT-18**, 100–105 (1970).
22. M. Igarashi and Y. Naito, Properties of a four-port nonreciprocal circuit utilizing YIG on strip line filters and circilator, *IEEE Trans. Microwave Theory Techniques*, **MTT-20**, 828–833 (1972).
23. M. Igarashi and Y. Naito, Theoretical analysis of magnetic resonance nonreciprocal circuits-limitations of 3 dB bandwidth and available range, *IEEE Trans. Microwave Theory Techniques*, **MTT-22**, 821–829 (1974).
24. J. Helszain, Scattering parameters of looped coupled YIG resonators, *Microwave J.*, **21**, 53 (December 1978).
25. G. J. Hurst, YIG and what it means to microwave instruments, *Marconi Instrumentation*, **15**, 76–79 (Spring 1977).
26. YIG devices, Watkins-Johnson Company, Palo Alto, CA, 1980.
27. RF and microwave components, Anzac Div. Adams Russel, Burlington, MA, 1982.
28. Private communication, R. Knittle, Hughes Aircraft Co., Fullerton, CA.
29. J. R. Reynolds and M. R. Rosenxweig, Learn the language of mixer specification, *Microwaves*, **17**, 72 (May 1978).
30. S. M. Fukuchi and R. B. Mouw, Broad-band double balanced mixer/modulators, Parts I, II, *Microwave J.*, **12** (March/May 1969).
31. H. D. Mills, On the equation $i = i_0[\exp \alpha(V - R_i) - 1]$, *IBM J.*, **11**, 553–554 (September 1967).
32. E. M. Rutz-Philipp, Power conversion in nonlinear resistive elements related to interference phenomena, *IBM J.*, **11**, 544–552 (September 1967).
33. A. R. Derr, A technique for determining the local oscillator waveforms in a microwave mixer, *IEEE Trans..Microwave Theory Techniques*, **MTT-23**, 828–831 (1975).
34. 17 most asked questions about mixers, Q & A, No. 1 in a series, Mini-Circuit, A Division of Scientific Components Corp., Brooklyn, NY.
35. D. Cheadle, Selecting mixers for best intermod performance, *Microwaves*, **12**, 48, 58 (November/December 1973).
36. D. Neuf and D. Brown, What to look for in mixer specs, *Microwaves*, **13**, 48 (November 1974).

37. D. L. Cheadle, Measure mixer noise with your power meter, *Microwaves*, **14**, 48–49 (March 1975).
38. D. P. Howson, Minimum conversion loss and input match conditions in the broad-band mixer, *Radio Electronic Engineer*, **42**, 237–242 (May 1972).
39. J. Cohen, Mixer conversion loss versus local oscillator waveshape, *Microwave J.*, **21**, 71 (October 1978).
40. C. W. Gerst, Jr., New mixer designs boost D/F performance, *Microwaves*, **12**, 60 (October 1973).
41. E. J. Crescenzi, Jr., R. W. Oglesbee, and R. A. Chappell, Integrating components for new front-end design, *Microwaves*, **13**, 35 (August 1974).
42. R. B. Hirsch, Mixers, mixer-preamps, and limiters, *Microwave J.*, **18**, 55 (October 1975).
43. E. J. Crescenzi, Jr., and F. A. Marki, Miniature X-band amplifiers and down converters, Watkins Johnson Co. Tech Notes, **5**(6) (November/December 1978).
44. B. C. Henderson, Mixers: part 1, Watkins Johnson Co. Tech Notes, **8**(2) (March/April 1981).
45. H. H. Cross, Automatic bias circuit for mixers, *Microwave J.*, **26**, 226 (May 1983).
46. H. Howe, Jr., Stripline circuit design, Artech House, Dedham, MA, 1974.
47. G. Begemann, An X-band balanced fin-line mixer, *IEEE Trans. Microwave Theory Techniques*, **MTT-26**, 1007–1011 (1978).
48. A. R. Kerr, Noise and loss in balanced and subharmonically pumped mixers: part I—theory and part II—Application, *IEEE Trans. Microwave Theory Techniques*, **MTT-27**, 938–950, December 1979.
49. R. S. Caruthers, Copper oxide modulators in carrier telephone systems, *Bell System Technical J.*, **18**, 315–337 (1939).
50. H. A. Watson, *Microwave Semiconductor Devices and Their Circuit Applications*, McGraw-Hill, New York, 1969.
51. R. Phlieger, A new MIC double balanced mixer with RF and IF band overlap, *IEEE MTTS International Microwave Symposium*, pp. 301–303, June 1973.
52. T. Smith and J. Wright, Spurious performance of microwave double-balanced mixers, Watkins Johnson Application Note, August 1975.
53. H. Ogawa, M. Aikawa, and K. Morita, K-band integrated double balanced mixer, *IEEE Trans. Microwave Theory Techniques*, **MTT-28**, 180–184 (1980).
54. B. C. Henderson, Predicating intermokulation supression in double-balanced mixers, Watkins Johnson Co. Tech Notes, **10**(4) (July/August 1983).
55. B. C. Henderson, Reliably predict mixer IM suppression, *Microwave RF*, **22**, 63 (November 1983).
56. J. B. Cochrane, Thin film mixers team up to block out image noise, *Microwaves*, **16**, 34 (March 1977).
57. K. M. Johnson, X-band integrated circuit mixer with reactively terminated image, *IEEE Trans. Microwave Theory Techniques*, **MTT-16**, 388–397 (1968).
58. A. A. M. Salah, *Theory of Resistive Mixer*, M.I.T. Press, Boston, 1971.
59. L. E. Dickens and D. W. Maki, A new phased type image enhanced mixer, *IEEE G-MTT International Symposium Digest Tech.*, pp. 149–151, 1975.
60. L. E. Dickens and D. W. Maki, An integrated circuit balanced mixer, image, and sum enhanced, *IEEE Trans. Microwave Theory Techniques*, **MTT-23**, 276–281 (1975).
61. M. E. Hines, Image conversion effects in diode mixers, *IEEE MTT-S International Microwave Symposium*, pp. 487–490, 1977.
62. M. R. Barber, Noise figure and conversion loss of the Schottky barrier mixer diode, *IEEE Trans. Microwave Theory Techniques*, **MTT-15**, 629–635 (1967).
63. M. V. Schneider, W. W. Snell, Jr., Harmonically pumped stripline down converter, *IEEE Trans. Microwave Theory Techniques*, **MTT-23**, 271–275 (1975).
64. M. Cohn, J. E. Degenford, and B. A. Newman, *Harmonic mixing with an antiparallel diode pair*, *IEEE Trans. Microwave Theory Techniques*, **MTT-23**, 667–673 (1975).
65. L. S. Cutler and C. L. Searle, Some aspects of the theory and measurement of frequency fluctuations in frequency standards, *Proc. IEEE*, **54**, 136–154 (1966).

66. J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healy, D. B. Lesson, T. E. McGuinigal, J. A. Mullen, W. L. Smith, R. Sydnor, R. F. C. Vessot, and G. M. R. Winkler, Characterization of frequency stability, NBS Technical Note 394, U.S. Dept. of Commerce/National Bureau of Standards, October 1970.
67. J. H. Shoaf, D. Halford, and A. S. Risley, Frequency stability specifications and measurement: high frequency and microwave signals, NBS Technical Note 362, U.S. Dept. of Commerce/National Bureau of Standards, January 1973.
68. D. A. Howe, Frequency domain stability measurement: a typical introduction, NBS Technical Note 679, U.S. Dept. of Commerce/National Bureau of Standards, March 1976.
69. Understanding and measuring phase noise in the frequency domain, Application Note 207, Hewlett-Parkard, October 1976.
70. A. L. Lance, W. D. Seal, F. G. Mendoza, and N. W. Hudson, Automated phase noise measurements, *Microwave J.*, **20**, 87 (June 1977).
71. S. Abrams, Sideband noise problems in receiver testing, *Microwave J.*, **23**, 51 (January 1980).
72. C. J. Grebenkemper, Local oscillator phase noise, Watkins Johnson Tech Note, **8**(6) (November/December 1981).
73. Fundamentals of quartz oscillators, Application Note 200-2, Hewlett-Packard, October 1977.
74. E. J. Baghdady, R. N. Lincoln, and B. D. Nelin, Short-term frequency stability characterization, theory and measurement, *IEEE Proc.*, **53**, 704–722 (1965).
75. E. J. Baghdady, J. A. Mullen, and B. D. Nelin, Short-term frequency stability, *IEEE Proc.*, **53**, 2110–2111 (1965).
76. B. G. Malcolm, Frequency pulling: what happens to oscillator stability with load variations, *Microwave System News*, **5**, 75 (June/July 1975).
77. J. M. Golio and C. M. Krowne, New approach for FET oscillator design, *Microwave J.*, **21**, 59 (October 1978).
78. R. J. Trew, Design theory for broad-band YIG-tuned FET oscillators, *IEEE Trans. Microwave Theory Techniques*, **MTT-27**, 8–14 (1979).
79. J. C. Papp, YIG-tuned FET oscillator design 8-18 GHz, Watkins Johnson Co. Tech Notes, **6**(5) (September/October 1979).
80. T. L. Heyboer and F. E. Emergy, YIG tuned GaAs FET oscillators, *IEEE MTT-S, International Microwave Symposium*, pp. 48–50, 1976.
81. P. M. Ollivier, Microwave YIG-tuned transistor oscillator amplifier design: application to C-band, *IEEE J. Solid State Circuits*, **SC-7**, 54–60 (1972).
82. T. Ruttan, X-band GaAs FET YIG-tuned oscillator, *IEEE MTT-S, International Microwave Symposium*, 264–265, 1977.
83. YIG-tuned bulk GaAs oscillators, Watkins Johnson Co. Tech Notes, **1**(5) (September/October 1974).
84. W. Green, W. Wilser, and K. Zublin, Ka-band YIG-tuned GaAs oscillators, Watkins Johnson Co. Tech Notes, **2**(5) (September/October 1975).
85. K. M. Johnson, Microwave varactor tuned transistor oscillator design, *IEEE Trans. Microwave Theory Techniques*, **MTT-14**, 564–572 (1966).
86. Voltage controlled oscillators, Frequency Sources, Inc., Catalog, 1974.
87. R. N. Buswell, Voltage controlled oscillators in modern ECM systems, Watkins Johnson Co. Tech Notes, **1**(6) (November/December 1974).
88. J. B. Fuller, Frequency-agile local oscillator, *Microwave J.*, **17**, 31 (May 1974).
89. R. N. Buswell, Linear VCOs, Watkins Johnson Co. Tech Notes, **3**(2) (March/April 1976).
90. R. M. Beach, Hyperabrupt varactor-tuned oscillators, Watkins Johnson Co. Tech Notes, **5**(4) (July/August 1978).
91. R. Beach, High speed linear oscillators, *Microwave J.*, **21**, 59 (December 1978).
92. E. G. Chalmers and R. E. Carman, Wide-band VCO uses Gunn varactor tuning, *Microwave System News*, **9**, 75 (January 1979).
93. M. Dydyk, A step-by-step approach to high-power VCO design, *Microwaves*, **18**, 54 (February 1979).
94. W. Wagner, Oscillator design by device line measurement, *Microwave J.*, **22**, 43 (February 1979).

95. F. Pergal, Detail a Coltitts VCO as a tuned one-port, *Microwaves*, **8**, 110 (April 1979).
96. R. J. Traynor and W. Cshuerch, New VCO test equipment aids manufacturing, *Microwaves*, **20**, 57 (January 1981).
97. C. D. Corbey, R. Davies, and R. A. Gough, Wide-band varactor tuned coaxial oscillators, *IEEE Trans. Microwave Theory Techniques*, **MTT-24**, 31–39 (1976).
98. E. C. Niehenke and R. D. Hess, A microstrip low-noise x-band voltage-controlled oscillator, *IEEE Trans. Microwave Theory Techniques*, **MTT-27**, 1075–1079 (December 1979).
99. Measuring the tuning step transient response of VCOs to 18 GHz, Hewlett-Packard Application Note 174-13, November 1974.
100. V. Manasewitsch, *Frequency Synthesizer Theory and Design*, Wiley, New York, 1980.
101. Phase locked loop systems, 2nd ed., Motorola Semiconductor Products, Inc., Phoenix AZ, August 1973.
102. F. M. Gardner, *Phase Lock Techniques*, Wiley, New York, 1979.
103. G. B. Shelton, Phase locking increases range of wide-range phase controller, *Electronic Design*, **16**, 96 (August 2, 1973).
104. P. G. Tipon, New microwave synthesizers that exhibit broader bandwidths and increased spectral purity, *IEEE Trans. Microwave Techniques*, **MTT-22**, 1246–1254 (1974).
105. S. R. Kurtz, Mixers as phase detectors, Watkins Johnson Tech Notes, **5**(1) (January/February 1978).
106. S. F. Wetenkamp and K. J. Wong, Transportation lag in phase-locked loops, Watkins Johnson Tech Notes, **5**(3) (May/June 1978).
107. G. D. Alley and H. C. Wang, An ultra-low noise microwave synthesizer, *IEEE Trans. Microwave Theory Techniques*, **MTT-27**, 969–974 (1979).
108. C. E. Foster and T. G. Simon, Microwave frequency synthesizers, *Watkins Johnson Co. Tech Notes*, **2**(3) (May/June 1975).
109. W. P. Mason, *Physical Acoustics: Principles and Methods*, Vol. 1, Chapter 5, Academic Press, New York, 1964.
110. A. I. Zverev, *Handbook of Filter Synthesis*, Chapter 9, Wiley, New York, 1967.
111. T. J. Lukaszek, Mode control and related studies of VHF quartz filter crystal, *IEEE Trans. Sonics and Ultrasonics*, **SU-18**, 238–246 (1971).
112. M. E. Van Valkenburg, *Introduction to Modern Network Synthesis*, Chapter 12, Wiley, New York, 1980.
113. Operation and maintenance manual for TN-1000/WJ 1840 0.5-18 GHa tuner, Watkins Johnson Co., June 15, 1982.
114. C. D. Dexter, Digitally controlled VHF/UHF receiver design, Watkins Johnson Co. Tech Notes, **7**(3) (May/June 1980).

Chapter 6

Instantaneous Frequency Measurement (IFM) Receivers

6.1. INTRODUCTION (1)

An IFM receiver was discussed as early as 1948 by Earp (ref. 1). The output of the first IFM receiver was displayed in polar coordinates on a scope. The angle of the polar display represents the input signal frequency and the amplitude represents the signal strength. Although modern IFM receivers still use the same fundamental operating principle, the frequency of the input signal is often represented in digital format. Figure 6.1 shows an IFM receiver with digital frequency, pulse amplitude (PA), and pulse width (PW) displays.

The characteristics of IFM receivers make them potentially suitable for electronic warfare (EW) applications. This kind of receiver has a wide instantaneous radio-frequency (RF) bandwidth. Sometimes the bandwidth is more than an octave. The receiver can measure short pulses with high frequency accuracy (i.e., 1 MHz resolution on 100 nsec pulse). It has moderately high sensitivity and dynamic range. Modern IFM receivers can be compact in size and lightweight. However, there is one major deficiency in IFM receivers; the receiver can only read out one input signal at a time. If multiple signals arrive at the input of the receiver simultaneously, the best result one can expect is that the receiver reports only one of the input frequencies correctly. Sometimes the receiver will report an erroneous frequency that is not related to either one of the input signals. The erroneous frequency data generated by the receiver may confuse the digital processor that follows the receiver. Therefore, it is desirable to identify all erroneous data generated by the receiver to alert the processor. However, no such capability currently exists. The other approach

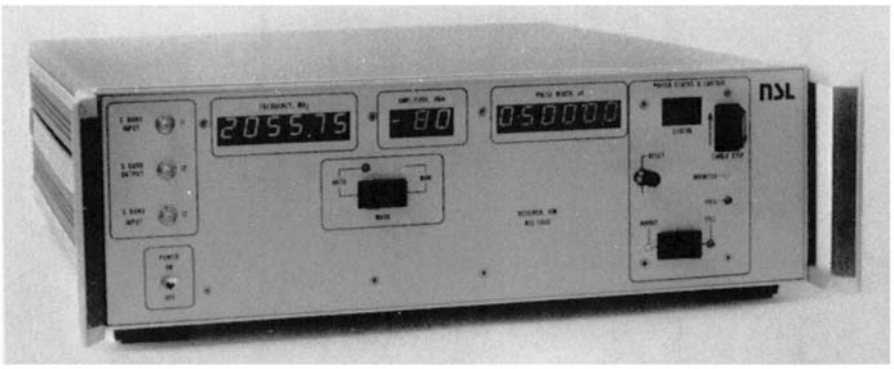


Figure 6.1. A digital IFM receiver. (Fabricated by Northern Scientific Laboratories, Division of General Instrument Corp. Courtesy of Avionics Laboratory, AFWAL.)

is to have a detection circuit sensing the existence of simultaneous signals. If the simultaneous signal flag is up, the processor will be aware that the frequency report might be wrong. This information may be used in the processor to make necessary precautions.

In this chapter, the operating principles of IFM receivers, some trade-offs in the designs, and several schemes for digitizing the outputs of the video amplitude will be discussed. Since an IFM receiver can measure only one input signal at a time, one continuous-wave (CW) signal will block the receiver from measuring other input signals. This CW problem will be discussed. The overlapping signal problem and some schemes to detect the existence of overlapping signals will also be discussed.

6.2. PRINCIPLE OF OPERATION (1, 2)

An IFM receiver uses delay lines to compare the phase of the input signal to measure its frequency. A sinusoidal wave is considered split into two paths, one path delayed a constant time with respect to the other one as shown in Figure 6.2a. There is a phase difference between the two outputs caused by the constant time delay. The phase difference is frequency dependent as shown in Figures 6.2*b*, *c* which show two sinusoidal waves of different frequencies. The relative phase angle between the delayed and the undelayed waves is

$$\theta = \omega\tau \tag{6.1}$$

In an IFM receiver, the frequency of the input signal is determined by measuring this phase difference. If the phase angle is measured and the delay time τ is known, the frequency can be determined through Eq. (6.1).

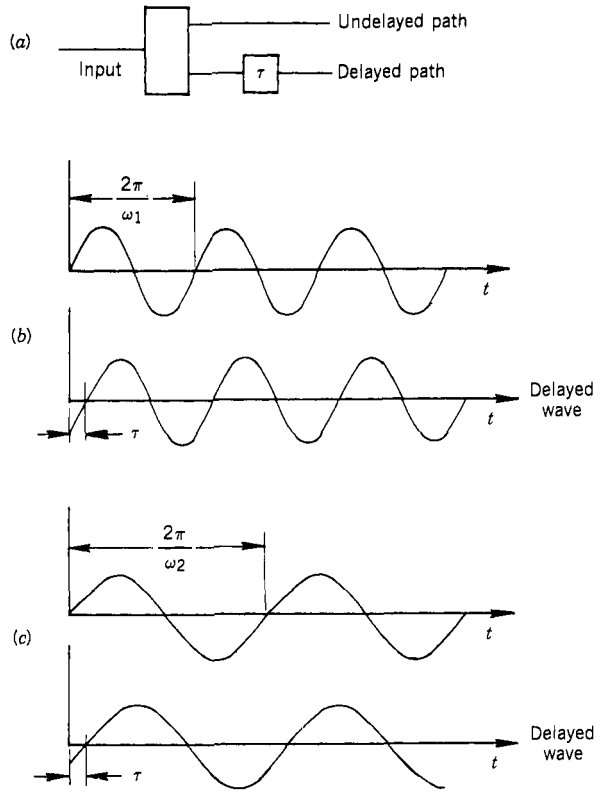


Figure 6.2. Phase relation of sinusoidal waves with constant time delay. (*a*) signal paths; (*b*) high frequency; (*c*) low frequency.

In an IFM receiver, the phase angle θ is measured using the relations

$$E = A \sin \theta \tag{6.2}$$

$$F = A \cos \theta \tag{6.3}$$

where E and F are two voltages and A represents the amplitude of the input signal. The phase angle can be measured from the amplitudes of E and F. In a practical IFM receiver, the major portion of the circuitry is to produce the above relations. In some of the earlier IFM receivers, the voltages E and F are measured on the X and Y axes of an oscilloscope. The trace of the point (X, Y) is a circle. The radius of the circle represents the amplitude of the signal, and the angle in the polar display represents the phase angle because

$$\theta = \tan^{-1}(E/F) = \omega\tau \tag{6.4}$$

In a modern IFM receiver, the amplitude information in Eqs. (6.2) and (6.3)

is often eliminated through a limiting amplifier, and only the frequency information is retained. Thus the results in Eqs. (6.2) and (6.3) are only used to generate digital frequency information.

6.3. BASIC COMPONENTS IN AN IFM RECEIVER (3–10)

The RF section of a basic IFM receiver consists of five components: a power divider, a delay line, a phase correlator (phase discriminator), four diode detectors, and two differential amplifiers. The power divider splits the input signal into two paths. The delay line produces the necessary time delay required in Eq. (6.1). The phase correlator provides the proper phase shift to both the delayed and undelayed input signals and combines them in special ways. The detectors will change the input RF signals to video signals and at the same time perform mathematical multiplications. The differential amplifiers will perform mathematical subtractions. The five components are combined to generate the relations in Eqs. (6.2) and (6.3). The detail operating principles of the phase correlator, detectors, and differential amplifiers will be discussed through an example.

A phase correlator is a passive component that has two inputs and four outputs. There are many different ways to build a phase correlator. In general, a phase correlator includes hybrid circuits (90° or 180° units). Sometimes an in-phase power divider is also used. Figure 6.3 shows a simple RF section of an IFM receiver. Figure 6.4*a* shows one of the phase correlator designs, and Figure 6.4*b* shows a different phase correlator with the input power divider and delay line. The phase correlator in Figure 6.4*a* contains three 90° phase shifters and one 180° phase shifter. To simplify the discussion, let us neglect the amplitude of the signal and use unity for all the signals in the analysis. Let us also assume that in the hybrids the signal from the coupled output (going straight across from the input) does not produce any change in phase with respect to the input signal. The signal from the transmitted output (through the diagonal path) will shift the specified phase with respect to the input signal. Strictly speaking, the

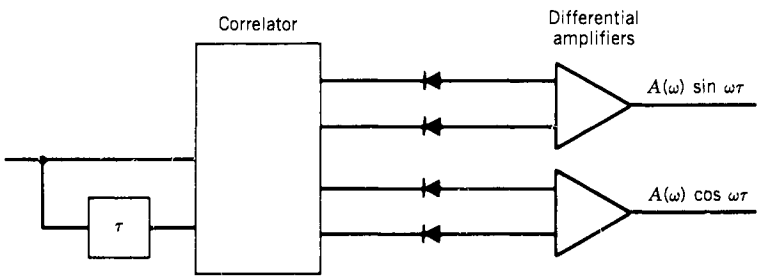


Figure 6.3. Basic IFM receiver.

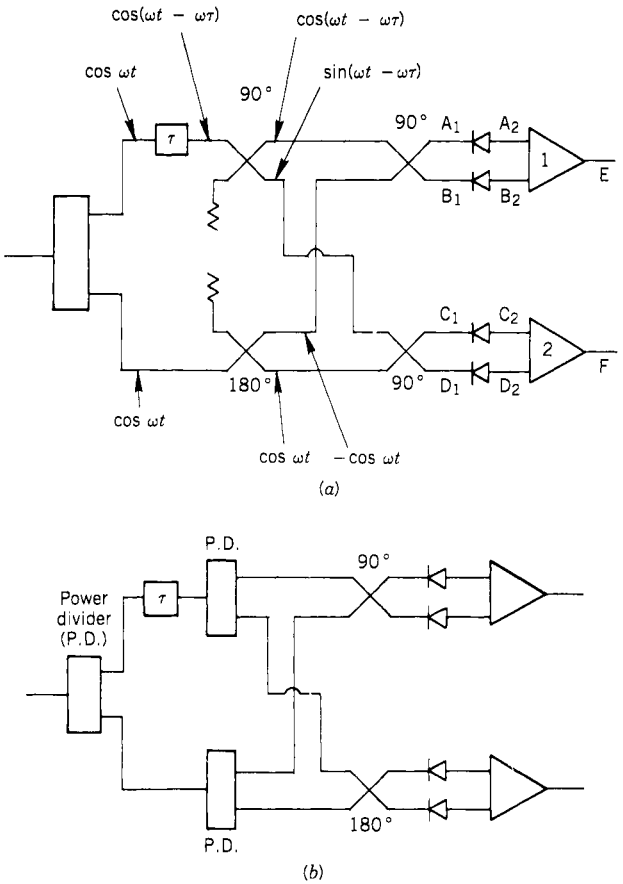


Figure 6.4. Phase correlators: (*a*) with four hybrids (1–180°, 3–90°); (*b*) with two power dividers and two hybrids.

outputs of the coupled path and the transmitted path of a hybrid do not have the same phase or the specified phase with respect to the input signal. However, this assumption will simplify the mathematics and is sufficient for this discussion.

For a $\cos \omega t$ input signal, the signals after the power divider and the first two hybrids are labeled in Figure 6.4*a*. The signal after the second two hybrids and before the diode detectors are

$$A_1 = \cos(\omega t - \omega\tau) - \sin \omega t \tag{6.5}$$

$$B_1 = \sin(\omega t - \omega\tau) - \cos \omega t \tag{6.6}$$

$$C_1 = \sin(\omega t - \omega\tau) + \sin \omega t \tag{6.7}$$

$$D_1 = -\cos(\omega t - \omega\tau) + \cos \omega t \tag{6.8}$$

The outputs of the detectors are derived from the squaring of the above equations because the detectors are square-law devices. The outputs after the detectors are

$$A_2 = \cos^2(\omega t - \omega\tau) - 2\cos(\omega t - \omega\tau)\sin\omega t + \sin^2\omega t \tag{6.9}$$

$$B_2 = \sin^2(\omega t - \omega\tau) - 2\sin(\omega t - \omega\tau)\cos\omega t + \cos^2\omega t \tag{6.10}$$

$$C_2 = \sin^2(\omega t - \omega\tau) + 2\sin(\omega t - \omega\tau)\sin\omega t + \sin^2\omega t \tag{6.11}$$

$$D_2 = \cos^2(\omega t - \omega\tau) - 2\cos(\omega t - \omega\tau)\cos\omega t + \cos^2\omega t \tag{6.12}$$

If at the outputs of the detectors there are low-pass filters that will stop the high-frequency terms in the above equations, the results can be written as

$$A_2 = 1 - \sin\omega\tau \tag{6.13}$$

$$B_2 = 1 + \sin\omega\tau \tag{6.14}$$

$$C_2 = 1 + \cos\omega\tau \tag{6.15}$$

$$D_2 = 1 - \cos\omega\tau \tag{6.16}$$

The outputs of the differential amplifiers 1 and 2 are

$$E = B_2 - A_2 = 2\sin\omega\tau \tag{6.17}$$

$$F = C_2 - D_2 = 2\cos\omega\tau \tag{6.18}$$

which are the desired results. It should be emphasized that the results in Eqs. (6.17) and (6.18) can be obtained through many different combinations of hybrids and power dividers.

Figure 6.4*b* shows another example of the phase correlator. In this example, the phase correlator consists of two in-phase power dividers, one 90° hybrid, and one 180° hybrid. One should be able to trace an input signal similar to that of Figure 6.4*a* and obtain the same outputs as Eqs. (6.17) and (6.18).

Figure 6.5 shows another IFM receiver arrangement. In this arrangement, two mixers are used after the two hybrids. The outputs from the two hybrids are combined through the two inputs (the signal and local oscillator ports) of the mixers. The outputs are from the intermediate-frequency (IF) port of the mixers. Each mixer will perform the function of a 180° hybrid, detectors, and differential amplifier. The diodes in the mixers perform the mathematical squaring. This kind of arrangement is especially suitable for lower-frequency operation because the mixers have better performance, that is, high isolation between the ports. Since each mixer will replace three components, this design may be more economic.

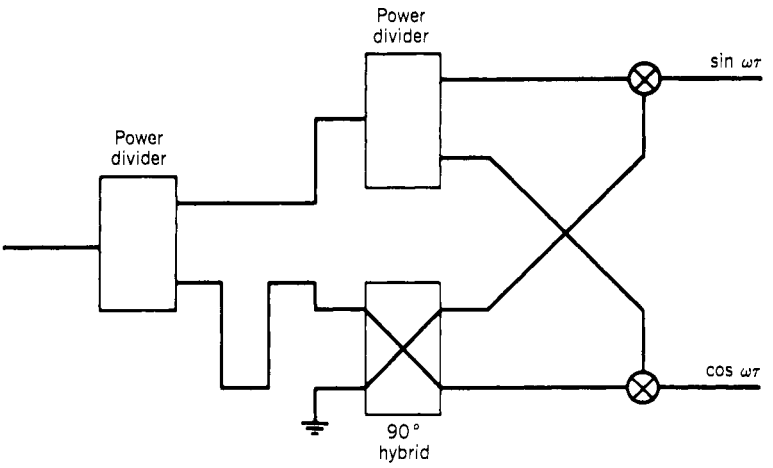


Figure 6.5. IFM receiver with two mixers as detectors.

In many IFM receiver designs, in order to improve the sensitivity of a receiver, enough amplification must be provided before the detectors. The amplification will make the detectors operate in the linear region. Mathematically, a detector operating in a linear region cannot perform the squaring operation. However, although the detectors are operating in a linear region, the IFM receiver still functions properly. The explanation is that detectors are basically nonlinear devices and convert RF signals to dc signals. They will perform the required mathematical squaring operation whether they operate in the nonlinear or linear region.

6.4. LIMITING AMPLIFIERS (11–25)

In an operational IFM receiver, there is usually a limiting amplifier in front of the power divider, delay line, and phase correlator. The limiting amplifier can improve the sensitivity of the receiver by providing RF gain to process weak input signals. In general, a signal must reach the limiting level in order to be processed by the receiver. When the input signal to a limiting amplifier is below a certain threshold, the amplifier will amplify the input signal as an conventional RF amplifier. However, when the input signal reaches a certain threshold level, the output will not increase but will stay at a constant level. The characteristic of an ideal limiting amplifier is shown in Figure 6.6*a* and a typical output is shown in Figure 6.6*b*. In generating Figure 6.6*b*, the limiter function was approximated by

$$V_o = \frac{\alpha}{\pi} \tan^{-1} \frac{\pi V_i}{\alpha} \tag{6.19}$$

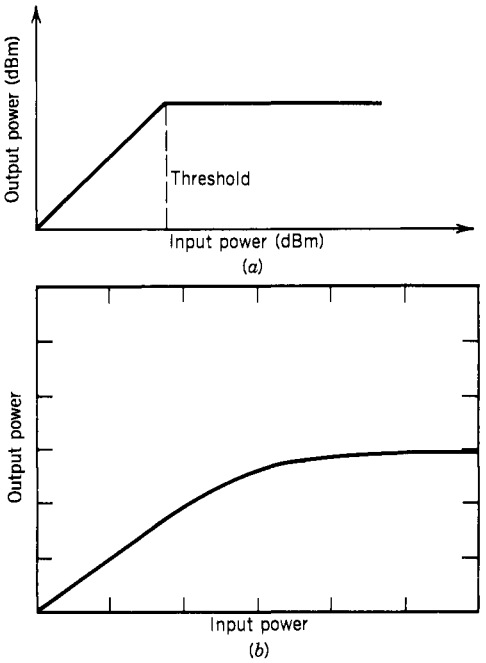


Figure 6.6. Input and output relation of a limiting amplifier: (*a*) Ideal response. (*b*) Typical response.

where α is a constant related to the limiting power level of the device and V_i and V_o are the input and output voltage, respectively. The logarithmic response of Eq. (6.19) is plotted in Figure 6.6*b*.

In a properly designed IFM receiver, the minimum signal to be detected should reach the knee of the curve. Stronger input signals will be limited at the constant-output level. In Figure 6.6, it is obvious that the amplitude information of the input signal will be lost after the limiter. A limiting amplifier will also affect the PW. For example, if the input signal has long leading and trailing edges, the limiting amplifier will shorten both the leading and trailing edges and distort the PW information.

From Figure 6.6, one can readily see that the limiting amplifier is a nonlinear device. The output of a limiting amplifier will no longer be a simple sinusoidal wave. However, the basic operating principle derived for an IFM receiver is based on a simple sinusoidal wave. In an IFM receiver, there is usually a low-pass filter after the limiting amplifier to reduce the harmonics and restore the signal to a sinusoidal wave. Sometimes the bandwidth of the power divider and the phase correlator used in the correlator will function as a low-pass filter.

The major difference between a limiting amplifier and a conventional linear amplifier is that in a limiting amplifier the output is very close to a constant value after saturation. The output of a conventional amplifier has more amplitude variation after saturation. Sometimes the amplitude may even decrease

with increasing input power after saturation. If properly designed, a linear amplifier can be used in saturation to replace a limiting amplifier.

It is important to notice that a limiting amplifier can be used only at the input of an IFM receiver. It should not be used in other kinds of receivers because the limiting amplifier will generate intermodulation products when overlapping signals are present. Since an IFM receiver only reads one signal at a time, the limiting amplifier will not disturb the performance of the receiver but will improve it.

A limiting amplifier expresses a capture effect on simultaneous signals. If two simultaneous signals are present at the input of the amplifier, a strong one and a weak one, the limiter will suppress the weak one more than the strong one. For example, if the strong one is 3 dB higher than the weak one at the input of the limiting amplifier, at the output of the limiting amplifier, the strong signal may become approximately 6 dB higher than the weak signal. In other words, the strong signal captures the weak signal. An intuitive explanation is that the strong signal has a higher probability of being clipped by the limiter. When the strong signal is clipped by the limiter, its fundamental frequency still exists, but the weak signal is totally masked during the period of clipping off. Figure 6.7

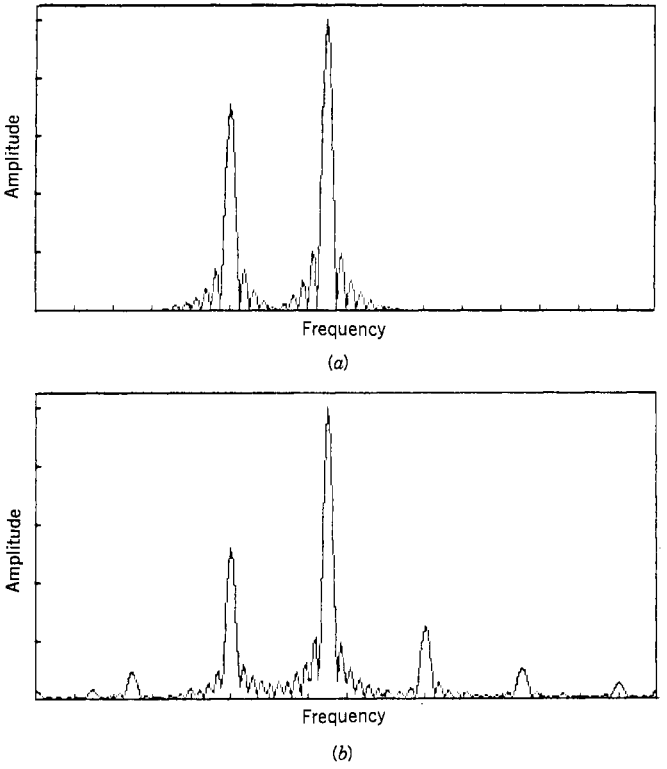


Figure 6.7. Simulated capture effect of a limiter. (*a*) Input signals to the limiter. (*b*) Output of limiter.

shows the results of a computer simulation of a limiter. The limiter effect is described by Eq. (6.19). The inputs are 3 dB apart in power level whereas the outputs are approximately 6 dB apart. Intermodulations are also generated at the outputs due to the nonlinear effect. The intermodulation can be used to detect simultaneous signals and will be discussed in Section 6.16.

The capture effect will reduce the probability of generating erroneous frequency data by the IFM receiver under simultaneous signal conditions because the stronger signals have a better chance of being reported after the limiting amplifier.

A limiting amplifier consists of two components: RF amplifiers and limiters. The amplifiers provide the desired gain, and the limiters provide the limiting effect on the output signal. The amplifiers and the limiters are arranged properly in a chain.

The limiters can be divided into two groups according to the devices used in them. One kind uses diodes as the limiting device and is called a diode limiter; the other kind uses a YIG sphere as a limiting device and is called a YIG limiter. The diode limiter uses the diode characteristic to accomplish the limiting effect. Figure 6.8 shows a two-diode limiter. The diodes are connected in parallel in reverse directions across a transmission line. When the input signal is above a certain threshold, one diode will clip the positive side of the signal and the other diode will clip the negative side of the signal. Therefore, the input level is limited.

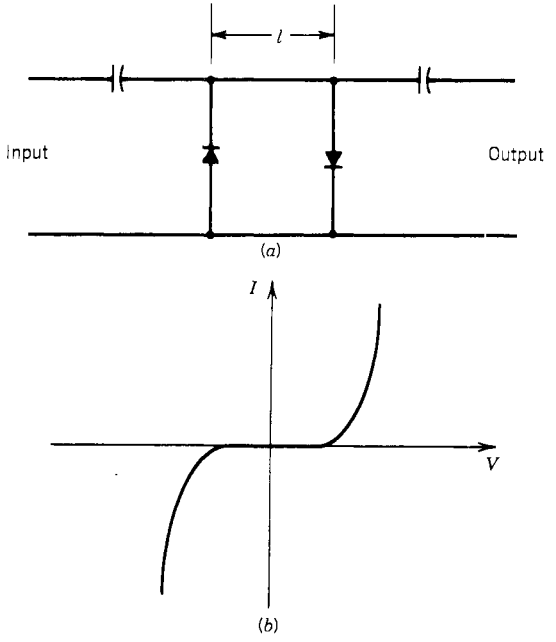


Figure 6.8. Diode limiter. (*a*) Schematic diagram. (*b*) I–*V* characteristic curve.

The excess input power is reflected back by the diodes. The diode limiters are usually used in limiting amplifiers.

A YIG limiter uses YIG spheres to achieve the limiting effect. The input and output of the limiter are coupled through YIG spheres, which are placed in a dc magnetic field of proper magnitude. The coupling is accomplished through magnetic spin waves in the YIG spheres excited by the input signal. (This effect is also discussed in Section 5.5.) When the input signal is under a certain threshold, the coupling will increase with increasing input signal strength. When the input signal is above the threshold, the coupling will start to saturate. Therefore, the output of the limiter will start to level off. When a pulsed signal passes a YIG limiter, the limiter cannot respond immediately. The leading edge of the pulse is not properly limited, and it will "leak" through the limiter. (Further discussion of YIG limiters can be found in Section 7.19.)

In general, the bandwidth of a YIG limiter is narrower than that of a diode limiter because the resonant frequency of the spin waves is related to the dc-applied magnetic field. The structure of a YIG limiter is more complicated than that of a diode limiter because a dc magnetic field must be properly applied. Therefore, YIG limiters are not often used in a limiting amplifier.

Another common use of limiters is to protect a receiver from damage by high-input RF signals. A limiter is often placed at the input of a microwave receiver as the first component. High input power will be limited by the limiter. The output from the limiter is low enough not to cause any damage to the receiver.

6.5. POWER DIVIDERS (26–37)

There are many different ways to divide input power into differential paths, for example, through a T-section or T-section with resistors. The commonly used ones are referred to as Wilkinson power dividers and were first discussed in 1960 (ref. 26). The Wilkinson power dividers can be broadband and have low voltage standing wave ratios (VSWR) at both the input and output ports. A schematic of this kind of power divider is shown in Figure 6.9. The input and output impedances of the power divider are Z_0. The input signal is applied to n-parallel quarter-wavelength transmission lines of characteristic impedance Z_{01} each. At the output ports, resistors of value R are connected to a common point b. In an actual power divider, the output ports are physically close together. The resistor R is determined as

$$R = Z_0 \tag{6.20}$$

The characteristic impedance of the transmission line is

$$Z_{01} = \sqrt{n} Z_0 \tag{6.21}$$

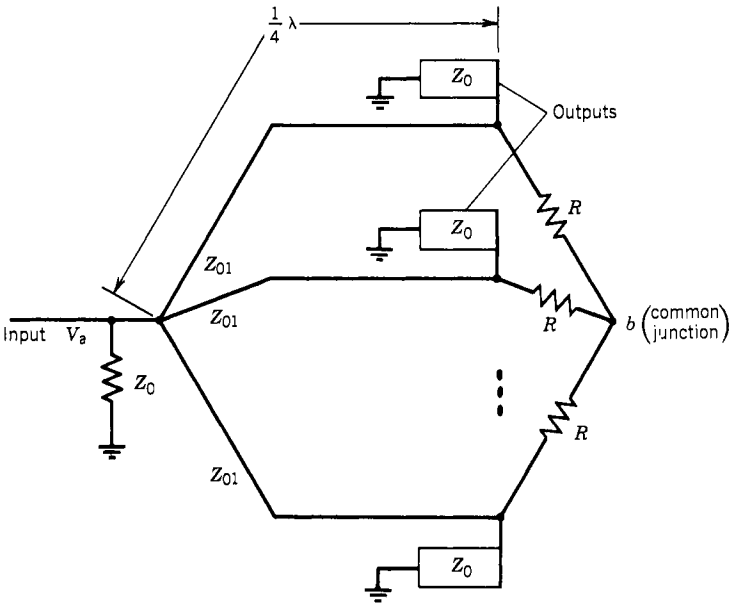


Figure 6.9. Wilkinson's power divider.

The operating principle of the Wilkinson power divider is that when a signal is applied at the input, it divides by virtue of symmetry into n equiphase equiamplitude ports. No power is dissipated by the resistors when matched loads are connected to the outputs, since all the transmission lines will be at the same potential. However, if a reflection occurs at one of the output terminals, the reflected signal will split; part of it will travel directly to the remaining output terminals via the resistors, and the remainder will travel back to the input, splitting again at the junction of the transmission lines and then returning to the remaining output terminals. Thus the reflected wave arrives at each remaining output terminal in two paths, and the length difference between the two paths will be 180° when the transmission lines are $\frac{1}{4}\lambda$ in length, where λ is the wavelength at the center frequency of the power divider. If resistor R, characteristic impedance Z_{01}, and input and output impedance Z_0 fulfill the relations given by Eqs. (6.20) and (6.21), the reflected signals from the two different paths will completely cancel each other at all the loads. Therefore, there is good isolation between the output ports.

6.6. DELAY LINES (38)

In an IFM receiver, the length of the delay line must be very accurate. If the delay line length is not accurate, the frequency calculated from Eq. (6.1) could

be erroneous. To operate an IFM receiver in a wide temperature range, the length of the delay lines must be kept constant over the operating temperature. One common approach to keep the delay line at constant length is to build an oven of constant temperatures around the delay line. Thus the delay line temperature is independent of the surrounding temperatures.

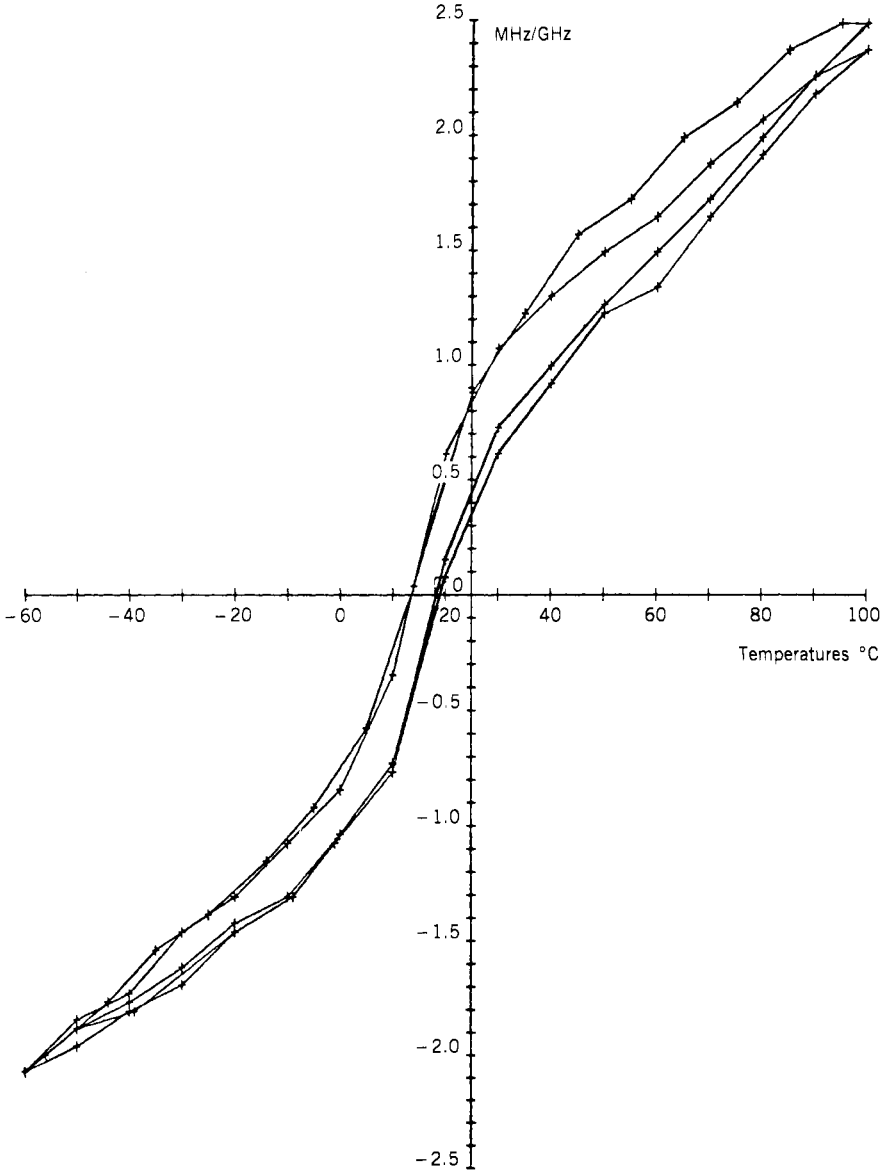


Figure 6.10. Frequency shift versus temperature for Teflon semirigid cable. (Ref. 38; courtesy of Northern Scientific Laboratories, Division of General Instrument Corp.)

Advances in dielectric material makes the length of the delay line less temperature dependent. Many recent IFM receivers use delay lines without a surrounding oven. Figures 6.10–6.12 show the results of three popular delay lines used in IFM receivers. Figure 6.10 shows the results of a Teflon semirigid cable. Figure 6.11 shows the results of a SiO_2 line, and Figure 6.12 shows the results

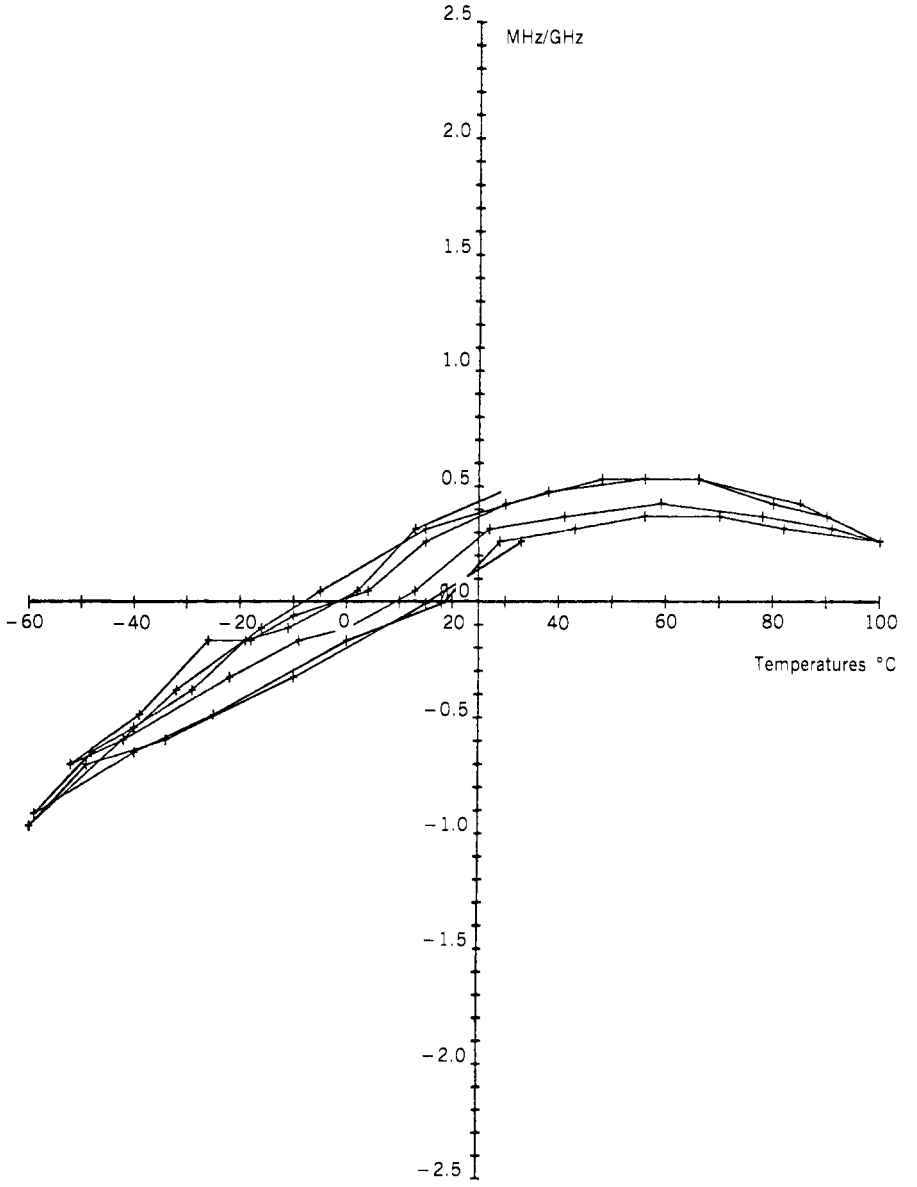


Figure 6.11. Frequency shift versus temperature for SiO_2 delay line. (Ref. 38 : courtesy of Northern Scientific Laboratories, Division of General Instrument Corp.)

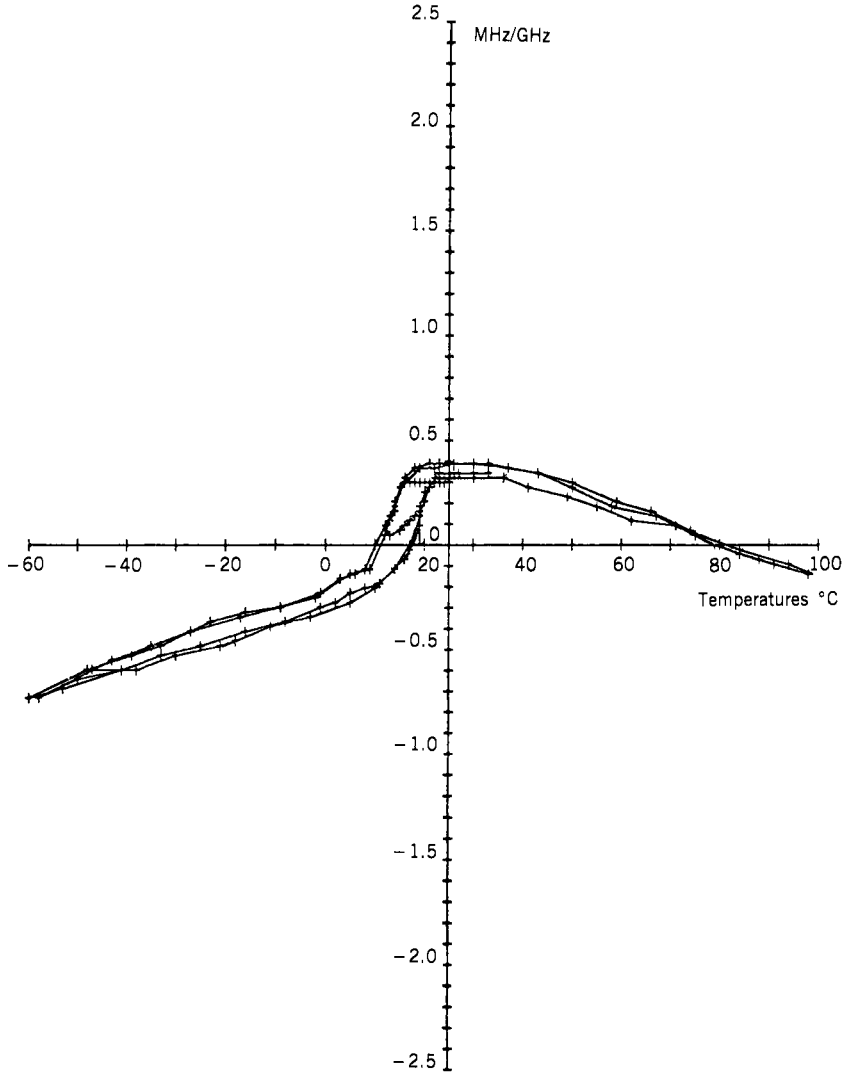


Figure 6.12. Frequency shift versus temperature for Gore expanded Teflon cable. (Ref. 38; courtesy of Northern Scientific Laboratories, Division of General Instrument Corp.)

of a Gore expanded Teflon cable. In these figures, the abscissa is the temperature and the ordinate is the change in megahertz per gigahertz instead of the change in line length. In all three cases, the results show hysteresis effect. The less the deviation in the ordinate, the less the frequency shift with temperature. The results in Figures 6.11 and 6.12 are better than that of Figure 6.10.

6.7. IFM RECEIVER WITH MULTIPLE CORRELATORS (39–44)

An operational IFM receiver usually contains several phase correlators with different length of delay lines. The longest line will provide the fine-frequency resolution, whereas the shorter lines are used to resolve the frequency ambiguities. From Eqs. (6.1)–(6.3), it is obvious that when $\omega\tau$ is greater than 2π, the value of ω cannot be uniquely identified from E and F. The shortest delay line in an IFM receiver is limited by the bandwidth of the receiver. If the bandwidth is ω, then the short delay line is limited by

$$\omega\tau < 2\pi \qquad \text{or} \qquad \tau < 2\pi/\omega = 1/f \tag{6.22}$$

The longest delay line is used to generate the fine-frequency information. Its length must be shorter than the minimum PW the receiver is designed to measure. Otherwise, the delayed and undelayed portions of the input signal with minimum PW will never overlap, and the receiver will not be able to measure its frequency.

An IFM receiver with four phase correlators is shown in Figure 6.13. The

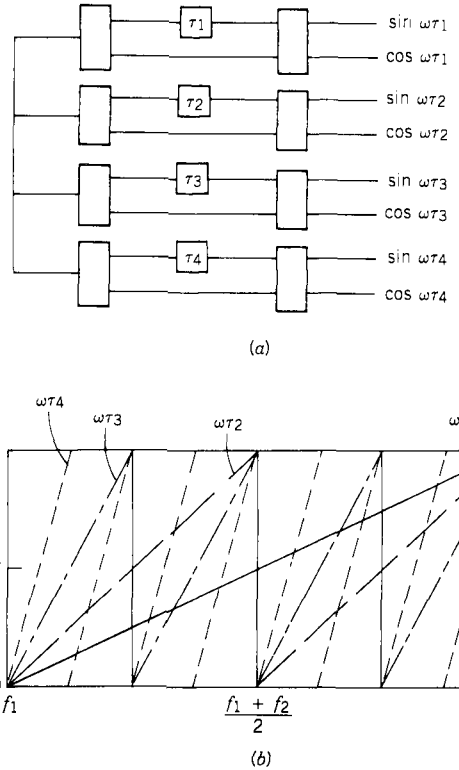


Figure 6.13. IFM receiver with four correlators.

input signal is separated into four parallel paths. Following each of the four outputs there is a basic IFM receiver with different lengths of delay line. Each basic unit will generate a pair of sin $\omega\tau$ and cos $\omega\tau$ outputs. The frequency of input signal is measured from the values of the four output pairs. In this special example, the delay line length is related as $\tau_4 = 2\tau_3 = 4\tau_2 = 8\tau_1$. The phase $\omega\tau$ versus frequency for the four lines are shown in Figure 6.13*b*. If one delay line generates only 2 bits of frequency information, and the total frequency coverage is 2 GHz, from Figure 6.13*b*, one can see that τ_1 will provide a 500-MHz resolution, τ_2 a 250-MHz resolution, τ_3 a 125-MHz, and τ_4 a 62.5-MHz resolution.

The higher the number of bits a correlator generates, the higher the input signal-to-noise ratio required, because to generate large number of bits, the sin $\omega\tau$/cos $\omega\tau$ outputs must be divided into many fine sections and their amplitudes compared. The noise on the outputs can cause error in the compared results. In practical receiver design, one delay line will generate a maximum of 6 bits of frequency information.

The total number of phase correlators in a practical IFM receiver depends on the design of the receiver. Let us use an example to demonstrate the different approaches. Consider an IFM receiver with an input frequency of 2–4 GHz. The frequency resolution is 1.25 MHz, and the minimum PW to be measured is 100 nsec. The total resolution cells is 1600 (2000/1.25), and 11 bits are required to display the frequency data.

In the above example, a minimum of two correlators are theoretically needed. If only two correlators are used, the first one will generate 5 bits and the second one will generate 6 bits of frequency data. The noise in the second correlator will cause a small frequency error, since a 1-bit error corresponds to 1.25 MHz. The noise in the first correlator will cause a large frequency error, since a 1-bit error corresponds to 80 MHz, and such a high frequency error is normally unacceptable. To have decent sensitivity, usually more than two correlators should be used. The correlators with short delay lines will only generate a few bits. It should be emphasized that the correlator with the longest delay line must generate the maximum number of bits. Otherwise, the capability of the receiver to measure minimum PW will suffer.

One practical approach to building the above IFM receiver is to use four correlators. The longest delay line is 25 nsec, which corresponds to an unambiguous frequency of 40 MHz. Using 6 bits to represent 40 MHz will produce the required 1.25 MHz. The other three delay lines can be 6.25, 1.563, and 0.391 nsec, which correspond to $\frac{25}{4}$, $\frac{25}{16}$, and $\frac{25}{64}$. The unambiguous frequency corresponding to the shortest delay line is 2560 MHz. The first three correlators combined will generate the five most significant bits (MSB) of frequency. The last correlator will generate the six least significant bits (LSB) of frequency data.

Another approach to building the IFM receiver is to use seven correlators. The longest delay line is also 25 nsec. The other six delay lines are 12.5, 6.25, 3.125, 1.563, 0.781, and 0.391 nsec, which correspond to $\frac{25}{2}$, $\frac{25}{4}$, $\frac{25}{8}$, $\frac{25}{16}$, $\frac{25}{32}$, and $\frac{25}{64}$. As in the above case, the last correlator generates the six LSB frequency

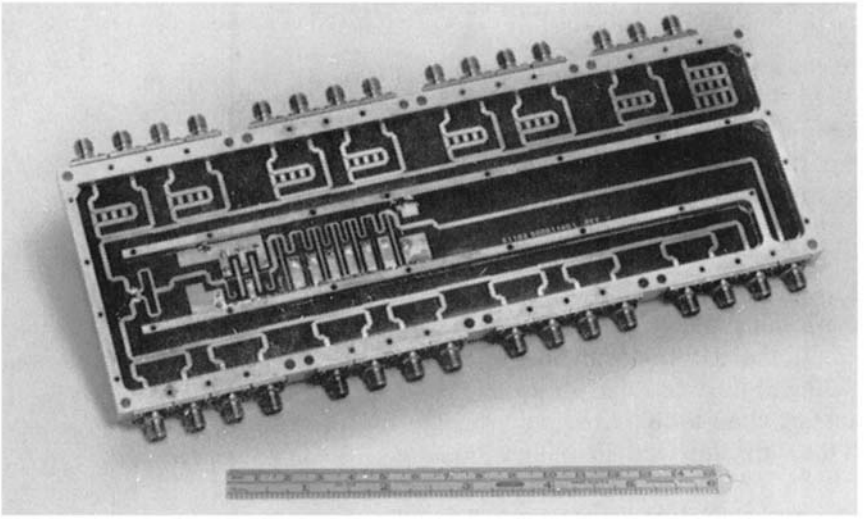


Figure 6.14. Stripline circuit for a seven-correlator IFM receiver. (Courtesy of Northern Scientific Laboratories, Division of General Instrument Corp.)

data and the first six correlators will generate the five MSB frequency data. This approach may have slightly higher sensitivity compared to the four-delay-line design because the phase correlators with short delay lines are only required to generate 1 bit of frequency information. The frequency error caused by noise is slightly lower in a correlator generating 1 bit of frequency information than in a correlator generating 2 bits of frequency information.

The layout of the stripline circuit of a seven-correlator IFM receiver is shown in Figure 6.14. The input is located at the lower right corner. The input signal will pass a low-pass filter at the center portion of the circuit. Four output connectors, two from the top of the circuit and the opposite two from the bottom of the circuit, will make the input power divider and the delay line combination. Among the four output connectors in one group, two of them are grounded and the other two are the inputs to the phase correlator. In this figure, there are a total of seven pairs of delayed and undelayed paths.

In the four-phase correlator approach, the input signal after passing through the limiting amplifier will be divided into four paths to the four correlators. In the seven-delay-line approach, the input signal after the limiting amplifier will be divided into seven paths. In comparing these two approaches, it is obvious that the first approach will use less hardware and have a slightly lower noise figure. However, the digital encoder in the second approach is simpler than that in the first, and less sensitive to noise except for the correlators with the longest delay lines. Therefore, the trade-off between the two approaches depends on the overall cost of RF and digital components and the complexity of the designs. Of course, there are many other ways to build the IFM receiver.

6.8. FREQUENCY DIGITIZING WINDOW

To convert the sin $\omega\tau$/cos $\omega\tau$ outputs to digital words, a time window must be assigned to measure the outputs. The sampling window is very narrow, usually in the nanosecond range. The time to sample the sin $\omega\tau$/cos $\omega\tau$ output is very important. In most IFM receivers, there is a video pulse that is used to generate the PA and PW information. This video pulse or the PA pulse is often used to control the frequency sampling time.

There are two basic ways to sample the sin $\omega\tau$/cos $\omega\tau$ outputs. One approach is to use the video pulse to generate a triggering pulse as shown in Fig. 6.15. The leading edge of this triggering pulse is used as the sampling window to sample the outputs from the differential amplifier at a fixed time after the leading edge of the video pulse has crossed the threshold. The sampling window must be close to the threshold crossing time in order to process short pulses. At the same time, the sampling window must be far enough from the threshold crossing so that the video pulse will reach steady state. This is especially true for pulses with slow rising times. The sampling window must be a compromise between these two requirements.

Another way is by sampling the pulse amplitude (PA) output and the sin $\omega\tau$/cos $\omega\tau$ outputs continuously and synchronously at a fixed rate. When the PA signal crosses the threshold, the corresponding sampled sin $\omega\tau$/cos $\omega\tau$ outputs can be used to generate the frequency information. However, to avoid measuring the frequency on the leading edge of the pulse, two sampled PA values, the one just crossing the threshold and the next one, are used to determine the desired sampling window. The sin $\omega\tau$/cos $\omega\tau$ outputs corresponding to the higher of the two PA values will be used to calculate the frequency information. If the sampling rate is high, short pulses can be processed. The advance in logic circuit speed makes this sampling scheme more popular.

There are various ways to encode the outputs from the differential amplifiers. The three commonly used approaches are amplitude comparison, resistor ring, and direct digitizing. Each will be discussed in the following sections.

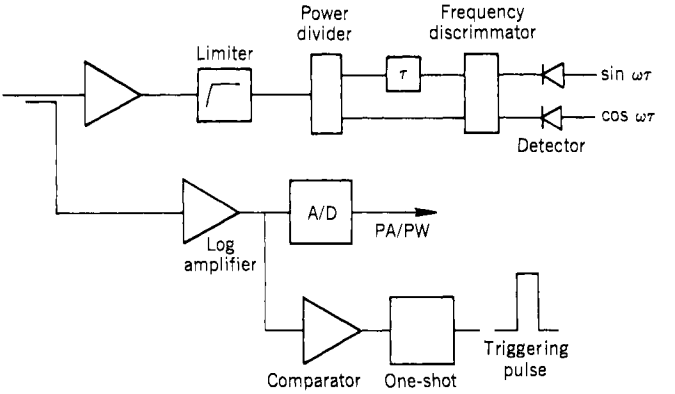
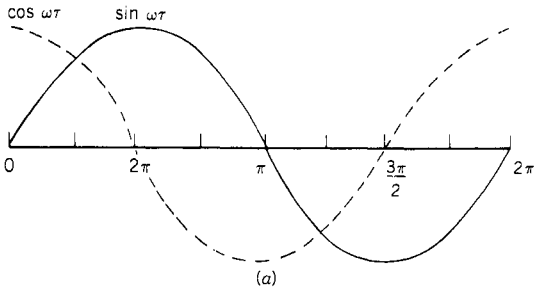


Figure 6.15. IFM receiver with PA measuring circuit.

6.9. FREQUENCY MEASUREMENT THROUGH AMPLITUDE COMPARISON

Amplitude comparison is one of the conventional ways to convert sin $\omega\tau$/cos $\omega\tau$ outputs to digital frequency information. The amplitude comparison scheme compares the sin $\omega\tau$/cos $\omega\tau$ outputs with some fixed reference voltages and also between themselves. Figure 6.16*a* shows the outputs sin $\omega\tau$ and cos $\omega\tau$ versus $\omega\tau$. These outputs can be rectified through diodes and compared to produce the results in Figure 6.16*b*. Figure 6.16*b* lists the desired bit pattern as a function of $\omega\tau$. This special pattern is a 4-bit Gray code. A Gray code changes only 1 bit from one frequency bin to an adjacent one. A Gray code can prevent encoding confusion at bin edges. To generate a logic 1, the conditions in Figure 6.16*c* must be fulfilled. To implement this arrangement, hardware comparators and voltage references are used. To generate a logic 1 at bits 4 and 3 in Figure 6.6*c*,



(*a*)

bit																
4	1								0							
3	1				0				0				1			
2	1		0		0		1		1		0		0		1	
1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1
bin no.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

(*b*)

bit	conditions to generate to logic "1"
4	sin $\omega\tau < 0$
3	cos $\omega\tau > 0$
2	$\lvert\cos \omega\tau\rvert - \sin \omega\tau \rvert > 0$
1	$\lvert\cos \omega\tau\rvert - \lvert\sin \omega\tau\rvert > (\cos 22.5° - \sin 22.5°) = 0.54$

(*c*)

Figure 6.16. Frequency digitizing through amplitude comparators: (*a*) sin $\omega\tau$/cos $\omega\tau$ outputs: (*b*) Typical four-bit Gray code; (*c*) condition to generate bit pattern in (*b*).

the outputs of sin $\omega\tau$ and cos $\omega\tau$ can be compared with 0 V. To generate a logic 1 at bit 1, the equation

$$|\cos \omega\tau| - |\sin \omega\tau| > 0.54 \tag{6.23}$$

or

$$|\sin \omega\tau| - |\cos \omega\tau| > 0.54 \tag{6.24}$$

must be fulfilled. Equations (6.23) and (6.24) can be implemented through voltage dividers and comparators. The values of 0.54, which equals cos 22.5° − sin 22.5°, can be implemented by a resistive voltage divider of ratio 0.54 at the outputs of the sin $\omega\tau$/cos $\omega\tau$. If more than four frequency bits are desired, they can be generated through the same approach, but the circuits will get progressively more complicated.

6.10. FREQUENCY MEASUREMENT THROUGH RESISTANCE RING (45)

Another approach to generate the digital frequency information is through the use of the resistance ring. The schematic diagram is shown in Figure 6.17.

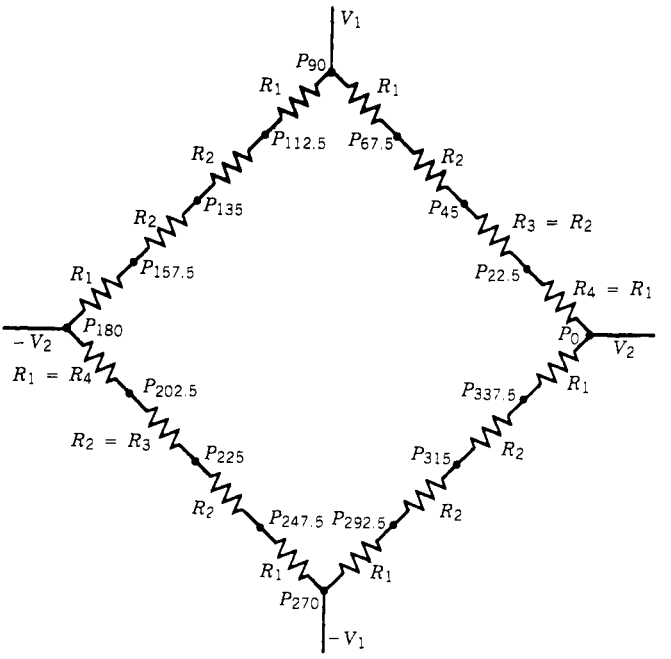


Figure 6.17. Frequency digitizing through resistance ring.

The outputs of the sin $\omega\tau$ and the cos $\omega\tau$ are applied to the two diagonal points V_1 and $-V_1$ and V_2 and $-V_2$ of the resistance network. Between each two input terminals there is a set of equivalued resistors to form a resistor ring. At the junction of each pair of resistors is an output terminal. Each facing pair of output terminals are connected to the inputs of a comparator, which compares the magnitude of the two inputs and generates a digital output. Although the resistor layout in Figure 6.17 is square, it is often referred to as a resistor ring because it divides 360 electric degrees into equal regions.

The resistor values can be determined from Figure 6.17. To simplify the discussion, let us use the special example in this figure to find the resistor values. The general case will be extended from this example. Since in this special example there are four resistors between points V_1 and $-V_1$ and V_2 and $-V_2$, the angle ϕ is defined as

$$\phi = 90°/4 = 22.5° \tag{6.25}$$

The resistor values must satisfy the following conditions. When $\phi = 0$, the voltage across V_1 and $-V_1$ should be zero. When $\phi = 22.5°$, the voltage across points $p_{292.5}$ and $p_{112.5}$ should be zero. When $\phi = 45°$, the voltage across the points p_{315} and p_{135} should be zero. In general, when the angle ϕ increases 22.5°, the voltage between certain opposite points should be zero. These pairs of points are rotated counterclockwise starting from V_1 and $-V_1$. The resistor R_1 can be calculated from the voltages at $p_{112.5}$ and $p_{292.5}$. When these two voltages are equal, the following equation is obtained.

$$V_1 - \frac{V_1 + V_2}{R} R_1 = -V_1 + \frac{V_1 + V_2}{R} R_1 \tag{6.26}$$

where R is the total resistance between points V_1 and $-V_1$ or V_2 and $-V_2$. Therefore, from Eq. (6.26),

$$R_1 = \frac{V_1}{V_1 + V_2} R = \frac{1}{1 + V_2/V_1} R \tag{6.27}$$

But

$$V_1 = 0.5 \sin(22.5°) \tag{6.28}$$

and

$$V_2 = 0.5 \cos(22.5°) \tag{6.29}$$

Substituting Eqs. (6.28) and (6.29) into (6.27), one obtains

$$R_1 = 0.293R \tag{6.30}$$

The values of R_2, R_3, and R_4 can be obtained through a similar procedure:

$$R_2 = \frac{1}{1 + \cot 45^\circ} R - R_1 = 0.207R \tag{6.31}$$

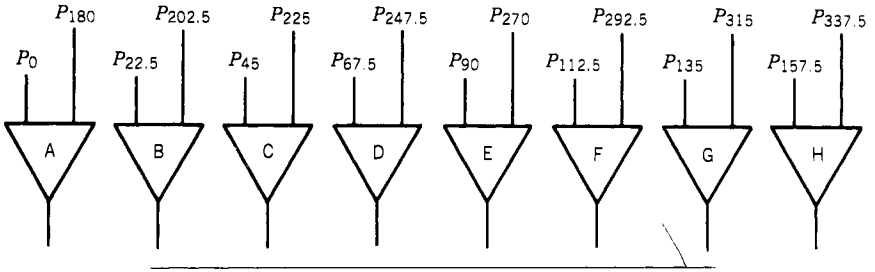
$$R_3 = \frac{1}{1 + \cot 67.5^\circ} R - R_1 - R_2 = R_2 = 0.207R \tag{6.32}$$

$$R_4 = \frac{1}{1 + \cot 90^\circ} R - R_1 - R_2 - R_3 = R_1 = 0.293R \tag{6.33}$$

With these R values, the outputs from the comparators are shown in Figure 6.18. In this figure, the comparator output is 1 when the left input is stronger. This is also a Gray code. A read-only memory (ROM) following the comparators will take these inputs and convert to a binary code.

If there are N resistors between points V_1 and V_2, the angle is defined as

$$\phi = 90^\circ/N \tag{6.34}$$



Angle (°)	A	B	C	D	E	F	G	H
0	1	1	1	1	0	0	0	0
22.5	1	1	1	1	1	0	0	0
45.0	1	1	1	1	1	1	0	0
67.5	1	1	1	1	1	1	1	0
90.0	1	1	1	1	1	1	1	1
112.5	0	1	1	1	1	1	1	1
135.0	0	0	1	1	1	1	1	1
157.5	0	0	0	1	1	1	1	1
180.0	0	0	0	0	1	1	1	1
202.5	0	0	0	0	0	1	1	1
225.0	0	0	0	0	0	0	1	1
247.5	0	0	0	0	0	0	0	1
270.0	0	0	0	0	0	0	0	0
292.5	1	0	0	0	0	0	0	0
315.0	1	1	0	0	0	0	0	0
337.5	1	1	1	0	0	0	0	0

Figure 6.18. Output from the eight comparators versus angle.

The resistors can be calculated as follows:

$$R_N = \frac{1}{1 + \cot N\phi} R - \frac{1}{1 + \cot(N-1)\phi} R \tag{6.35}$$

Since $R_1 = R_N$, only half of the resistors need be calculated.

The N resistors in the resistor will divide the 360 electric degrees into $4N$ equal regions. A total of $2N$ comparators are needed. In actual receiver design, the maximum number of N is 16, which will divide the circle into 64 portions. A maximum of 6 bits will be generated per correlator, which agrees with the earlier discussion in Section 6.9.

6.11. FREQUENCY DIGITIZING THROUGH A/D CONVERTERS

Another approach to digitize the outputs of the correlator is rather straightforward. The outputs of the $\sin \omega\tau$ and $\cos \omega\tau$ are digitized through analog-to-digital (A/D) converters. The digitized data will be fed into some ROM. ROM will take the inputs and perform the following operation:

$$\omega = \frac{1}{\tau} \tan^{-1} \frac{\sin \omega\tau}{\cos \omega\tau} \tag{6.36}$$

Therefore, the frequency of the input signal is calculated and displayed directly.

6.12. PULSE ON CW SIGNAL CONDITION (46)

From the above discussion, it is obvious that an IFM receiver can report only one signal at a time. Under simultaneous signal conditions, the best the receiver can do is to report one of the input frequencies correctly. If there is a CW signal at the input of an IFM receiver, the receiver always has a simultaneous signal condition when pulsed signals arrive. Thus, theoretically one CW signal can prevent an IFM receiver from encoding any other signals. The conventional solution to this problem is to put a tunable notch filter in front of the receiver to filter out the CW signal. Once a CW signal is detected and its frequency is measured, the filter is tuned to the frequency of the CW signal to block out the CW signal.

Another way to solve the CW signal problem in an IFM receiver is to separate the CW from pulsed signals and process them separately. Figure 6.19 shows the schematic for this approach. The RF front end of the receiver remains the same. At the output of each diode detector, the signal path is separated into two paths: one ac coupled and one dc coupled. The capacitors (C) are used to block the CW signals, which will appear as dc voltage at the output of the detector. The dc-coupled path will contain the CW and pulsed signal, whereas the ac-coupled path contains only the pulsed signals. The clamping diode D in the ac path will change the reference level of the pulsed signal to 0 V because after the coupling

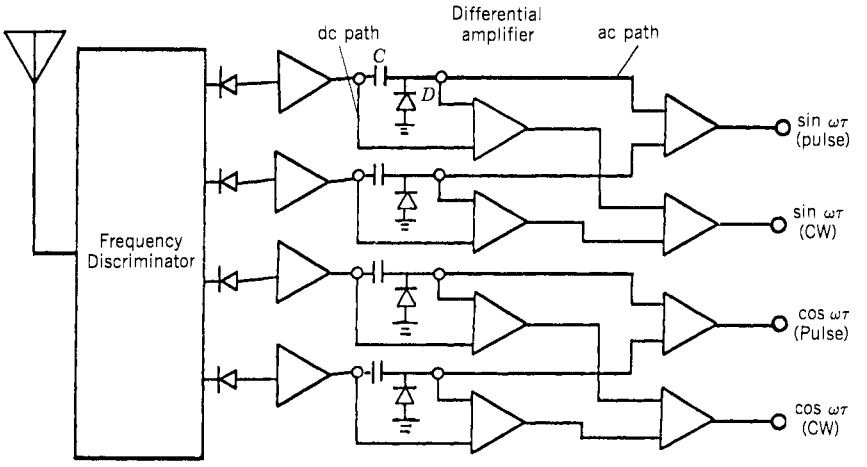


Figure 6.19. IFM receiver with CW–pulse separation circuit.

capacitor C the reference level of the pulse signal is below 0 V, as shown in Figure 6.20*a*. The reference level depends on the pulse amplitude and the pulse repetition interval (PRI). In Figure 6.20*a*, the positive areas should equal the negative areas. The clamping diodes raise the reference level to 0 V, as shown in Figure 6.20*b*. The output from the differential amplifier equals the difference of the dc- and ac-coupled outputs. Therefore, the output leaves only the dc components, which correspond to CW signals, because the pulsed signals from the ac and dc paths will cancel each other.

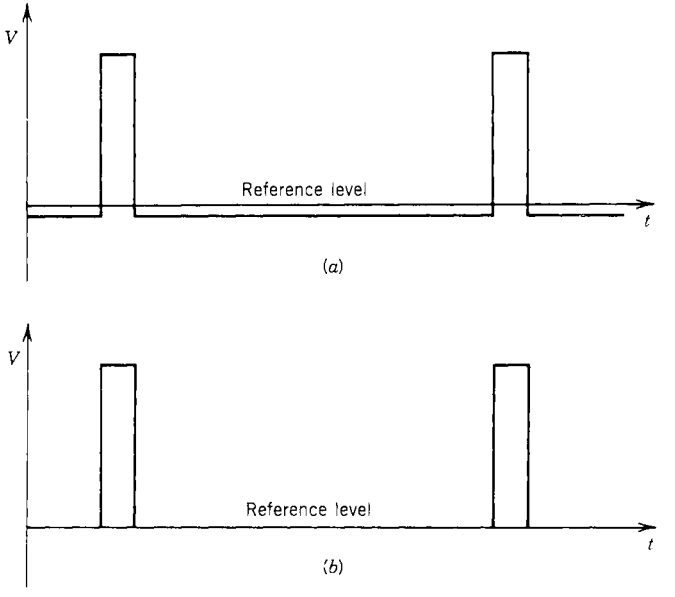


Figure 6.20. Outputs of ac-coupled path: (*a*) without clamping diode; (*b*) with clamping diode.

Two digitizing circuits will be used to encode the CW and pulsed signals separately. This circuit can separate only CW and pulsed signals. It cannot separate one CW signal from another CW signal nor can it separate two pulsed signals. For example, if two CW signals and one pulse signal are present at the input of the receiver, this scheme can separate out the pulsed signal and process it, but it cannot separate the two CW signals. On the other hand, if one CW signal and two simultaneous pulsed signals are present, this circuit can separate the CW signal and process it, but it cannot separate the two pulsed signals. The accuracy of the frequency readings on both the CW and pulsed signals may be slightly degraded because at the outputs of the detectors there is some interaction between the two signals, and they will slightly affect each other's frequency reading.

6.13. SIMULTANEOUS SIGNAL PROBLEMS (47, 48)

In this section, the problem of two time-coincident pulses arriving at an IFM receiver will be discussed. To eliminate confusion, let us define simultaneous signals and overlapping signals. When the leading edges of two signals are time coincident they are referred to as simultaneous signals. When the leading edges of two signals are not time coincident, they are referred to as overlapping pulses. Of course, simultaneous signals are also overlapping pulses. However, here overlapping pulses are only referred to as pulses with non-time-coincident leading edges.

The question of interest to systems engineers is that when two signals arrive at an IFM receiver simultaneously, what is the probability that the receiver will generate correct frequency information. Many laboratory tests have been carried out with two simultaneous signals on IFM receivers. Experimental data with more than two simultaneous signals will not be considered here because there will be too many different combinations of input conditions, and it is difficult to draw conclusions. The general rule in determining the performance of the IFM receiver under test with two simultaneous signals is that if the receiver can encode one frequency correctly, the performance is considered satisfactory.

Before discussing the test results, it is important to point out that the receiver under test uses the amplitude comparison scheme to determine frequency as discussed in Section 6.9. The outputs of the differential amplifiers are sampled at some constant delay time after the video pulse breaks the receiver threshold. It is possible that a receiver with other frequency encoding schemes may have slightly different results. Since the analog portion of the IFM receivers are similar, these results will provide a general idea of how an IFM receiver functions under various simultaneous signal conditions.

The test results show that when two simultaneous signals of amplitude within approximately 10 dB of each other arrive at an IFM receiver, the receiver will sometimes produce erroneous frequency data. The experimental results measured from the IFM receiver is shown in Figure 6.21. In this experiment,

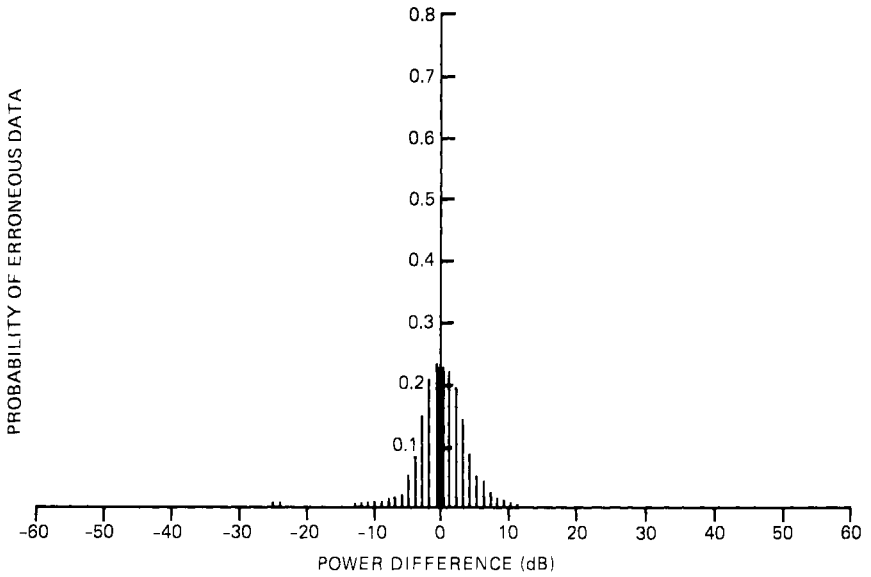


Figure 6.21. Erroneous frequency produced by typical IFM receiver with simultaneous signals.

an erroneous data is assumed when the output signal is more than 10 MHz away from either input signal. It should be noted that even when two signals are of the same amplitude, the probability of reporting erroneous data by the receiver is approximately 23%. It seems a little difficult to explain why the receiver can report one frequency correctly when two signals are of the same amplitude. To explain this phenomenon, the signal strengths from the inputs to the frequency measurement circuits should be considered. In general, there is gain variation across the frequency range, and the signal strengths at the input of the phase correlators may not equal in amplitude any more. The capture effect of the limiting amplifier will manifest the amplitude difference of the two signals. Thus, even if two signals are of the same amplitude at the input of the receiver, they may not be the same amplitude at the input of the frequency measurement circuits, and the receiver will encode the stronger signal correctly.

It was also found that when the two input signals are at certain frequencies, the receiver never generated erroneous data at any amplitude combination. It should be noted that in these tests the signal amplitudes are usually changed in 1-dB steps. For example, if one signal is set at frequency f_1 and the other at frequency f_2, when the signal at f_1 is stronger than f_2, the receiver will report f_1. If the signal at f_2 is increased in 1-dB steps, the receiver will still report f_1, and then suddenly start to report f_2. The receiver never reported any erroneous data in this test.

To explain this phenomenon, an IFM receiver with a single-phase correlator is used. If two simultaneous signals with frequencies ω_1 and ω_2 are present at

the input of the receiver, the outputs from the differential amplifiers are

$$E = A_1 \sin \omega_1 \tau + A_2 \sin \omega_2 \tau \tag{6.37}$$

$$F = A_1 \cos \omega_1 \tau + A_2 \cos \omega_2 \tau \tag{6.38}$$

where A_1 and A_2 are the amplitudes of signals 1 and 2, respectively. The above equations can be obtained through the same procedure demonstrated in Eqs. (6.5)–(6.18). In general, the receiver will report the vector sum of θ_1 and θ_2, as shown in Figure 6.22*a* where

$$\theta_1 = \omega_1 \tau \tag{6.39}$$

$$\theta_2 = \omega_2 \tau \tag{6.40}$$

If the input signal frequencies are related by

$$\theta_2 = \theta_1 \pm 180 \tag{6.41}$$

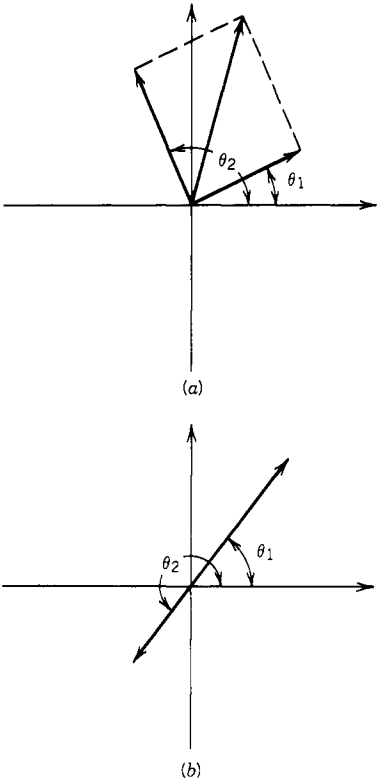


Figure 6.22. Vector representation of frequencies: (*a*) a general case: (*b*) collinear case.

as shown in Figure 6.22*b* (the collinear case), the receiver will always read the frequency of the stronger signal correctly.

However, it is suspected that if the amplitudes of the two signals are adjusted in a continuous manner rather than in 1-dB steps, the receiver may generate erroneous data on some measurements when the two vectors are not collinear.

6.14. OVERLAPPING SIGNALS PROBLEMS (47, 48)

Overlapping pulses are defined as pulses with leading edges that are not time coincident. In most IFM receivers, if the sampling window is controlled by the threshold break of the PA pulse, the frequency is measured about 60–100 nsec after the threshold is exceeded. If a second pulse arrives after the completion of the frequency measurement of the first pulse, the second pulse does not affect the frequency reading of the first signal. However, the receiver may miss the second pulse. The discussion here will concentrate on the case where the first signal triggers the frequency measurement circuit and the second pulse arrives before the frequency measurement of the first signal is complete. If the second signal is much weaker than the first, then the second pulse will not disturb the measurement. If the second signal is stronger than the first, then it will affect the measurement. Figures 6.23 and 6.24 show two sets of data with different delays. In Figure 6.23, the delay time is approximately 20 nsec; the results are not far from that of the time-coincident leading edges (Fig. 6.21). However, the

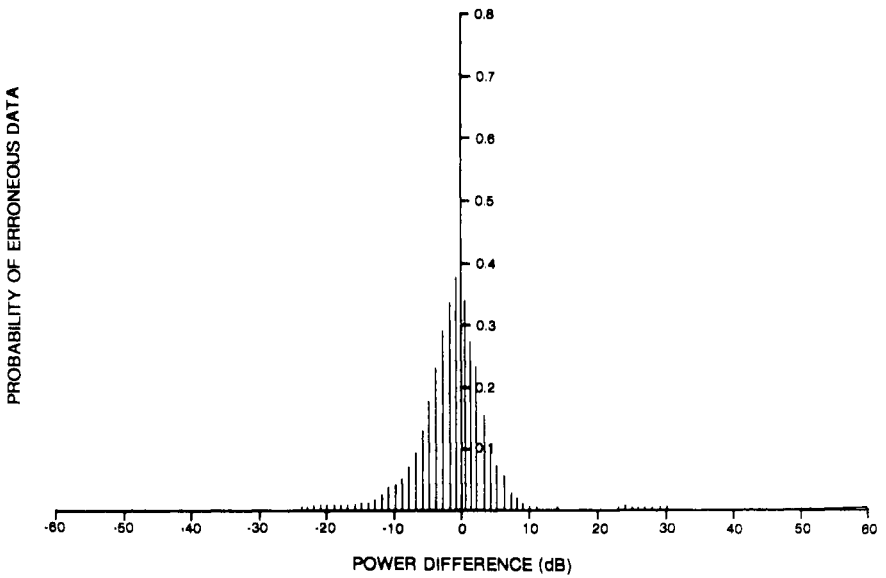


Figure 6.23. Probability of erroneous data with signal leading edge separation of 20 nsec.

probability of generating erroneous data is slightly higher. In Figure 6.24, the delay time is about 80 nsec, which is close to the sampling time of the receiver under test. The probability of generating erroneous data is very high when the second signal is stronger than the first. The reason for such poor performance is that when the second signal arrives, the output from the phase discriminator changes from one voltage level to a second voltage level. The first voltage corresponds to the frequency of the first signal, and the second voltage corresponds to the frequency of the second signal. It was observed that the transition from the first voltage to the second is not a smooth one, as expected. There is usually a "dip" in the voltage at the transition period, which will be discussed in more detail later (Section 6.17). The first signal will trigger the frequency measurement circuit of the receiver. When the receiver actually measures the frequency, the transition occurs and the measurement circuit sample some highly unpredictable voltage values. This effect will generate a very high probability of erroneous data.

The time delay between the two signals is very critical in relation to the generation of erroneous data from an IFM receiver. When the leading edge of the second signal is very close to the first one, the situation is quite similar to the simultaneous pulses. When the delay time between the leading edges increases, the probability of producing erroneous data also increases. Further increasing the delay time between the leading edges of the two signals will cause the probability of erroneous data to decrease slightly. When the second signal arrives after the sampling time, there is no longer erroneous data generated. Figure 6.25 shows a typical result of the "worst" probability as a function of the

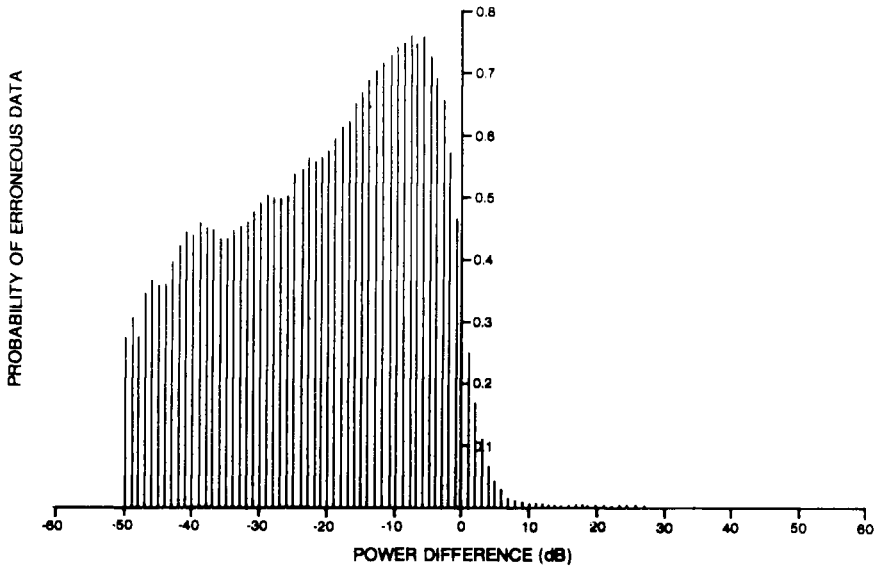


Figure 6.24. Probability of erroneous data with signal leading edge separation of 80 nsec.

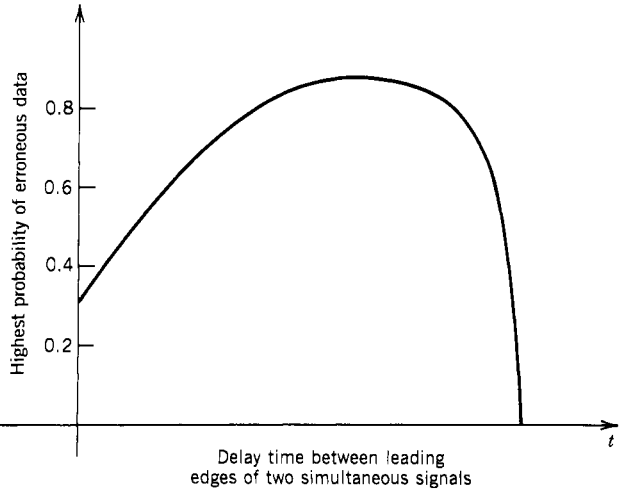


Figure 6.25. Highest probability of erroneous data generated by an IFM receiver versus the leading edge delay time.

leading edge delay time. It should be noted that in this figure only a few data points are available and the curve is extrapolated.

When the frequency is measured on a continuous sampling basis, the frequency error produced on overlapping pulses depends on the criterion of frequency encoding. The first approach to encode the frequency is that when the PA crosses the threshold, the PA data is compared with the second sampled data, and the digitized data corresponding to the higher PA value of the two samplings is used to calculate the frequency. The second approach is to monitor the PA video pulse and compare all the PA outputs at consecutive samplings. The data from $\sin \omega\tau/\cos \omega\tau$ corresponding to the maximum PA are used to calculate the frequency.

Although there is no experimental data presented here, the results can be reasoned under each individual case. The results on overlapping pulses from the first approach should be quite similar to the results shown in Figures 6.23–6.25 because the frequency is measured close to the leading edge of the $\sin \omega\tau/\cos \omega\tau$ pulses. In the second frequency-encoding scheme, the receiver will probably read the frequency of the stronger signal of the two input signals because the stronger signal will capture the limiting amplifier. However, if the two signals are of the same amplitude, the second frequency-encoding scheme may generate slightly higher percentages of erroneous frequency readings.

6.15. POSSIBLE SOLUTIONS TO SIMULTANEOUS/OVERLAPPING SIGNAL PROBLEMS

An ideal IFM receiver should not report any erroneous frequency information. However, with today's technology, this requirement cannot be achieved. When

overlapping signals arrive at the receiver, there is always a certain probability that the receiver will produce erroneous frequency readings. These erroneous frequency readings can create a tremendous burden on a signal processor that follows the receiver. The processor may process the erroneous data as correct data. Since these erroneous readings do not represent true signals, the processor can waste time manipulating the data. Worse yet, the processor may treat them as true signals and make wrong decisions.

Since an ideal IFM receiver cannot be built, it would be desirable to generate a flag that indicates erroneous data. If a processor receives this flag, it can neglect the reported frequency. However, such a capability cannot be implemented at the present time. The only approach that can be implemented is to detect the existence of simultaneous/overlapping signals. It should be emphasized again that even under simultaneous/overlapping signal conditions, an IFM receiver can produce a very high percentage of correct frequency information. Therefore, the processor should process all frequency data generated by the receiver. But when the data does not make sense and the overlapping signal flag is raised at the same time, the data can be considered erroneous and discarded.

In the following sections, several simultaneous/overlapping signal detection circuits will be discussed. Test results on some of the circuits will also be presented.

6.16. SIMULTANEOUS SIGNAL DETECTION THROUGH INTERMODULATION EFFECT (49)

The most straightforward approach for detecting simultaneous signals involves using a mixer to detect the existence of intermodulation (intermod) frequencies. The arrangement is shown in Figure 6.26. If there is only one input signal, the output from the mixer will be a dc voltage. If more than one signal is present at the input of the mixer, intermod frequencies will be generated at the output of the mixer. The low-pass filter in Figure 6.26 will block the input signals but pass the intermod frequencies. The detector will convert these frequencies to corresponding dc voltage and compare it against a fixed threshold. Although the filter after the mixer is called a low-pass filter, it should block dc voltages.

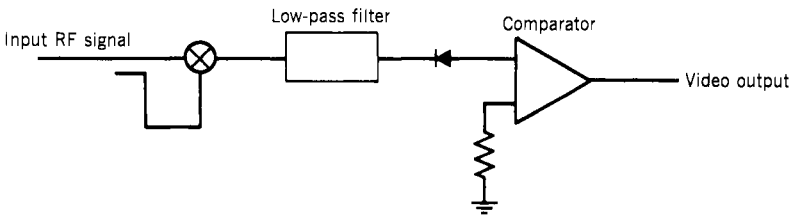


Figure 6.26. Simultaneous signal detection with mixer detector.

Even with one input signal, the mixer also produces a dc voltage, and this voltage should be blocked. This approach has some limitations—for example, if the input frequency range of an IFM receiver is 2–4 GHz, to block the fundamental frequency to the detector, the low-pass filter cutoff frequency must be less than 2 GHz. This filter will pass most simultaneous signal responses. If the input frequency range of the receiver is 2–6 GHz, to stop the input signals, the low-pass filter must also have a cutoff frequency of 2 GHz. In this arrangement, the simultaneous signal detector cannot detect signals separated by more than 2 GHz.

In the above arrangement, it is desirable to have strong intermod frequencies when simultaneous signals are present because it is easier to detect them. In order to have two signals generating strong intermod frequencies, the two signals must be close in amplitude. Therefore, this detection scheme will effectively flag erroneous data provided the two signals are close in amplitudes. Thus, this circuit should be very helpful for detecting simultaneous signals but not very sensitive to overlapping pulses. Since the overlapping time between the two overlapping signals is short, the intermod frequencies generated will be relatively short in time. This detection circuit is not very sensitive to such conditions, especially when the two signal amplitudes are far apart.

Since most IFM receivers have a limiting amplifier, which is a nonlinear device, before the phase correlator, it is possible to detect the existence of simultaneous signals by measuring the output of the limiting amplifier. In this approach, the limiting amplifier can replace the mixer and save some hardware. Figure 6.27 shows the basic arrangement. The low-pass filter and detector are required to perform the same function as that shown in Figure 6.26.

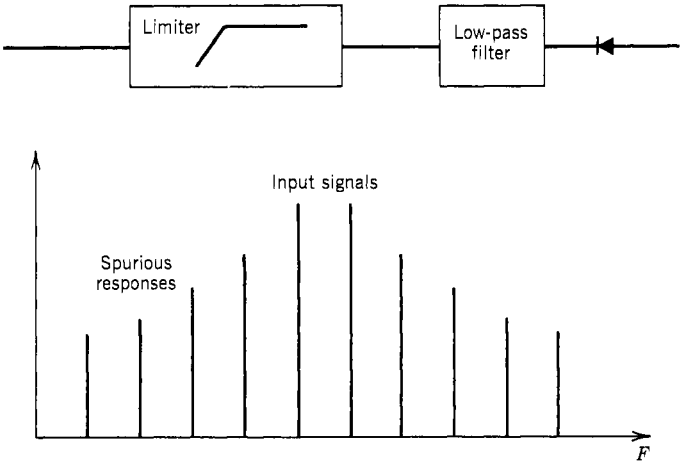


Figure 6.27. Using limiting amplifier for simultaneous signal detection.

6.17. DETECTION OF OVERLAPPING SIGNALS

It was mentioned above that an IFM receiver will generate a high percentage of erroneous data if a weak signal is followed by a strong one. Therefore, it is desirable to detect this condition. In order to discuss the overlapping pulse condition, the correlator output voltage V_1, V_1', V_2, and V_2' will be discussed first. (See Fig. 6.28.) Let us assume that

$$V_1 = \sin \omega_1 \tau \tag{6.42}$$

$$V_1' = \cos \omega_1 \tau \tag{6.43}$$

$$V_2 = \sin \omega_2 \tau \tag{6.44}$$

$$V_2' = \cos \omega_2 \tau \tag{6.45}$$

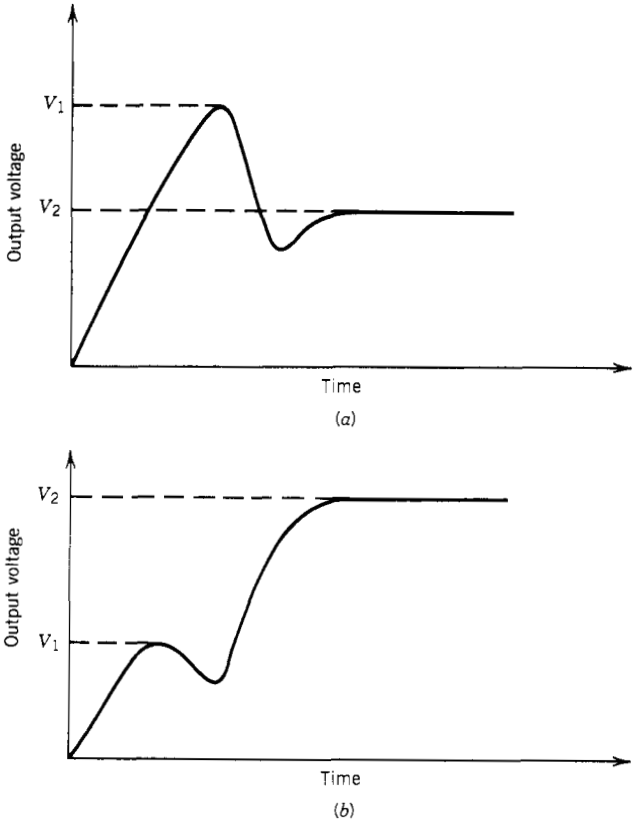


Figure 6.28. Voltage transient caused by overlapping pulses: (*a*) transient from voltage V_1 to V_2 ($V_1 > V_2$); (*b*) transient from voltage V_1 to V_2 ($V_1 < V_2$).

where V_1 and V_1' are the voltages corresponding to the first signal with angular frequency ω_1 and V_2 and V_2' are the voltages corresponding to the second signal with angular frequency ω_2 and τ is the delay time of the delay line. The amplitudes of the two signals are assumed to be unity. If $V_1 > V_2$ (or, equivalently, $V_1' < V_2'$), the output is shown in Figure 6.28*a*. If $V_1 < V_2$ (or, equivalently, $V_1' > V_2'$), the output is shown in Figure 6.28*b*. Between the transitions, there is usually a dip in the output voltage. This dip is very prominent in some cases and shallow in other cases, depending on the input conditions. The reason for this dip phenomenon may be explained as follows: When the second signal reaches approximately the same amplitude as the first and generates intermod frequencies, it is anticipated that the intermods will take energy away from the primary input signals and cause the dip in the output voltage.

Taking the square of V_1 and V_1' and summing them, the result should be a constant voltage because

$$V_1^2 + V_1'^2 = \sin^2\omega_1\tau + \cos^2\omega_1\tau = 1 \tag{6.46}$$

$$V_2^2 + V_2'^2 = \sin^2\omega_2\tau + \cos^2\omega_2\tau = 1 \tag{6.47}$$

However, since there is always a dip in the output voltage, the result will also have a dip as shown in Figure 6.29. Although Figure 6.28 can be observed on the scope, Figure 6.29 cannot be observed unless the mathematical operations in Eqs. (6.46) and (6.47) are performed. From Figure 6.28, it is obvious that when there are overlapping signals, there is almost always a dip in the $\sin \omega\tau$ (or $\cos \omega\tau$) that reflects on the output voltage. Alternatively, if there is a dip, then there are overlapping signals. Several methods will be used to detect the existence of this dip. They will be discussed in the following sections.

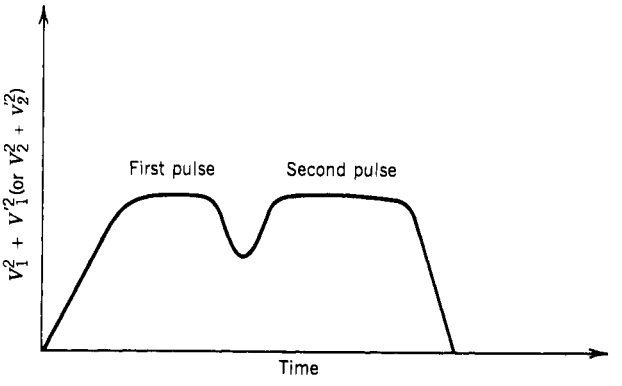


Figure 6.29. Output of $\sin^2 \omega\tau + \cos^2 \omega\tau$ with transient in the transition.

6.18. DIFFERENTIATOR DETECTION SCHEME FOR DETECTING OVERLAPPING SIGNALS (50, 51)

The most straightforward approach to detect dip is to build a differentiator circuit. A simple differentiator circuit is shown in Figure 6.30. It consists of a capacitor C and a resistor R_0 connected in series:

$$e_i = e_c + Ri = e_c + RC\frac{de_c}{dt} \tag{6.48}$$

where

$$R = R_1 + R_0 \tag{6.49}$$

$$e_o = \frac{R_0}{R}(e_i - e_c) = R_0 C\frac{de_c}{dt} \tag{6.50}$$

where e_i is the input voltage, e_c is the voltage across the capacitor C, R_1 is the input resistance and e_o is the output voltage. When the rise and fall time of the pulse are each $\gg RC$, and then

$$e_o \simeq R_0 C\frac{de_i}{dt} \tag{6.51}$$

Although Eq. (6.51) represents a differentiation function, the circuit in Figure 6.30*a* does not have any gain, and it is difficult to detect the dip.

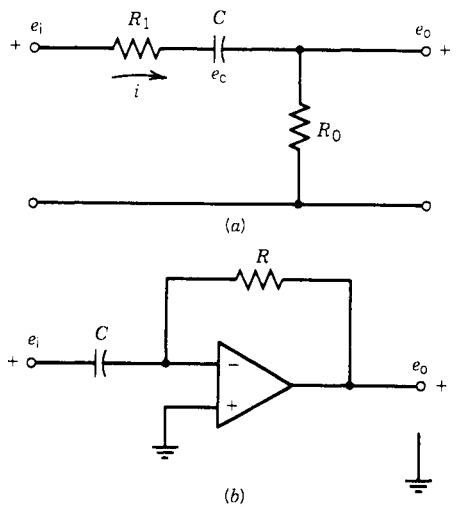


Figure 6.30. Differentiating circuits: (*a*) capactive differentiator; (*b*) using operational amplifier.

A differentiator can be built by using an operational amplifier as shown in Figure 6.30*b*. This circuit can provide some high gain to improve the sensitivity of the circuit. This circuit usually is too slow to effectively detect the dip in the video output.

6.19. SAMPLE/HOLD AND COMPARE SCHEME FOR DETECTING OVERLAPPING SIGNALS (52)

In this approach, the video output is sampled at the leading edge. The sampled value is held and compared with the rest of the pulses as shown in Figure 6.31. To eliminate false reports when only one signal is present, the signal at point *B* (Figure 6.31*c*) must be properly attenuated. If the signal *S* is larger than *B*, as shown in Figure 6.31*a*, there is only one signal at the input of the receiver. If signal *B* is larger than *S* (Figure 6.31*b*), then one signal is followed by a second one, and the second signal arrives before the first signal is properly encoded.

To make this scheme function properly, both $A \sin \omega\tau$ and $A \cos \omega\tau$ should be used because when one output is high, the other is low. Usually the dip in the high output caused by overlapping is more prominent.

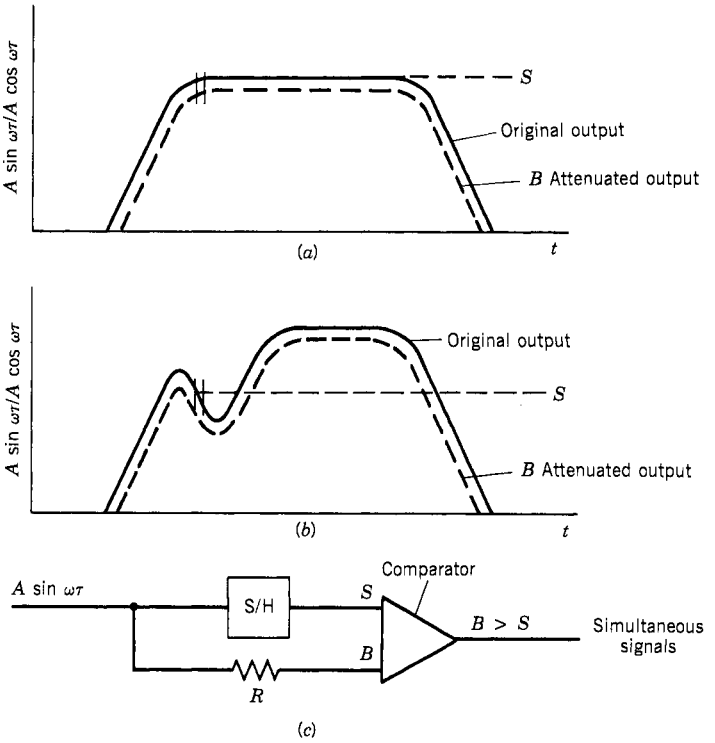


Figure 6.31. Sample/hold and compare scheme.

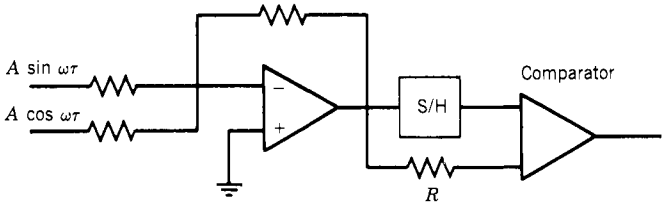


Figure 6.32. Modified sample/hold and compare scheme.

This circuit is difficult to align. When $A \sin \omega\tau$ (or $A \cos \omega\tau$) is close to zero, the condition in Figure 6.31*a* must be held in order to avoid the generation of false reports. If the circuit is so adjusted, it is usually not sensitive enough to detect overlapping signals because the attenuation is usually very high in order to eliminate false reports under single-signal conditions.

This approach can be modified slightly as shown in Figure 6.32. Instead of using comparator circuits at the $\sin \omega\tau$ and $\cos \omega\tau$ outputs, these outputs are combined through an operational amplifier and one sample/hold and compare circuit is used. Although this modification improves the performance of the detection circuit slightly, it also needs occasional adjustments.

6.20. DELAY AND COMPARE SCHEME FOR DETECTING OVERLAPPING PULSES

Another approach to detect the existence of dip in video output caused by overlapping signals is to use a delay and compare scheme. In this approach, the video outputs are first combined into one output through an operational amplifier. Their output from the operational amplifier is divided into two paths, undelayed and delayed, as shown in Figure 6.33. If there is only one input signal, the undelayed output is always larger than the delayed one except at the end of the pulse. If there are overlapping signals, the delayed signal may be stronger than the undelayed one at the dip. If the output from the comparator is gated at the proper time to detect this crossover and neglects the crossover at the end of the pulse, the presence of overlapping pulses can be detected, as shown in Figure 6.33*b*.

The first shortcoming of this approach is that the position of the dip depends on the time of arrival of the second pulse. To optimize this circuit, the delay time of the delay line should match the dip position, which is not predictable. The second problem with this approach is that the dip in some cases is not very prominent. To suppress the false reports on a single signal, the delayed video output must be properly attenuated. If the dip is not deep enough, this circuit may not be able to sense the existence of the dip and therefore fails to detect the existence of simultaneous signals.

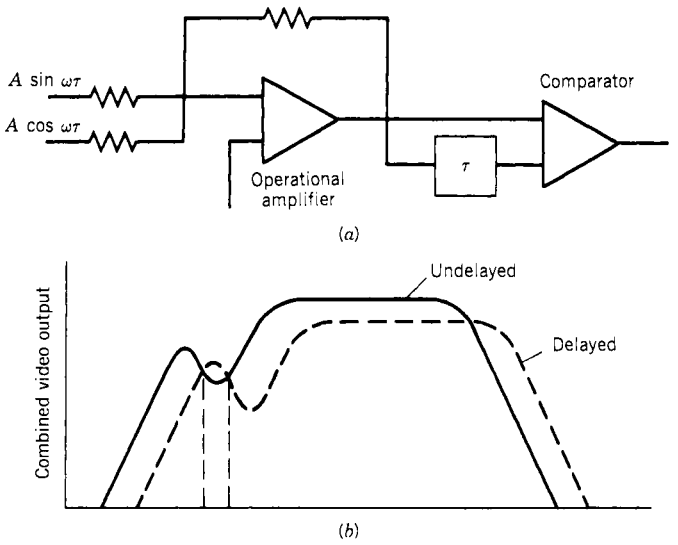


Figure 6.33. Delay and compare scheme: (*a*) circuit; (*b*) output.

The circuits discussed in this section and Section 6.19 have some common problems:

1. The circuits required many adjustments, and occasional readjustments, that is, attenuation on the delayed pulse and the change of delay time.

2. If the circuits were adjusted for minimum false reports, their sensitivity for detecting simultaneous signals also degraded. This difficulty arose from the fact that under various pulse overlapping conditions, the dip on the leading edge of the video pulse changed, and it was very difficult to accommodate all conditions.

6.21. DIGITAL DETECTION SCHEME FOR OVERLAPPING PULSES (53)

Although a digital detection scheme approach has demonstrated better performance than the circuits mentioned in the above sections, it is by no means ideal. This approach uses $\sin^2 \omega\tau + \cos^2 \omega\tau = 1$. If one sums the square of the $\sin \omega\tau$ and $\cos \omega\tau$ outputs, a constant-output level should be maintained. If the output level at the sampling window is lower than this constant value, it will indicate an overlapping signal condition. The summing and squaring of the two video outputs can be accomplished through A/D converters and digital manipulations. However, the problem is not this simple. Although there is usually a limiting amplifier in front of an IFM receiver to amplify the input signals to saturation, the sum of the squared video outputs is frequency dependent. In other words, the outputs from the differential amplifiers are

$A(\omega, S) \sin \omega\tau$ and $A(\omega, S) \cos \omega\tau$ rather than $\sin \omega\tau$ and $\cos \omega\tau$. The symbol $A(\omega, S)$ means the video amplitude depends on the input frequency and its signal strength S. The amplitude change of the video signal is due to the amplitude variations of the frequency response of the RF components used in the receiver. A typical $A(\omega, S)$ versus ω curve is shown in Figure 6.34. It should be noted that Figure 6.34 is not a measured curve but a hypothetical one because this curve cannot be measured directly. It can only be derived through digitizing the $\sin \omega\tau$ and $\cos \omega\tau$ outputs followed by the proper mathematical operation. A curve derived using this scheme is shown in Figure 6.35. The variation in the amplitude is caused by the digitizing effect. The curve $B(\omega)$ is a threshold curve.

Figure 6.34. Typical $A(\omega, S)$ versus ω curve.

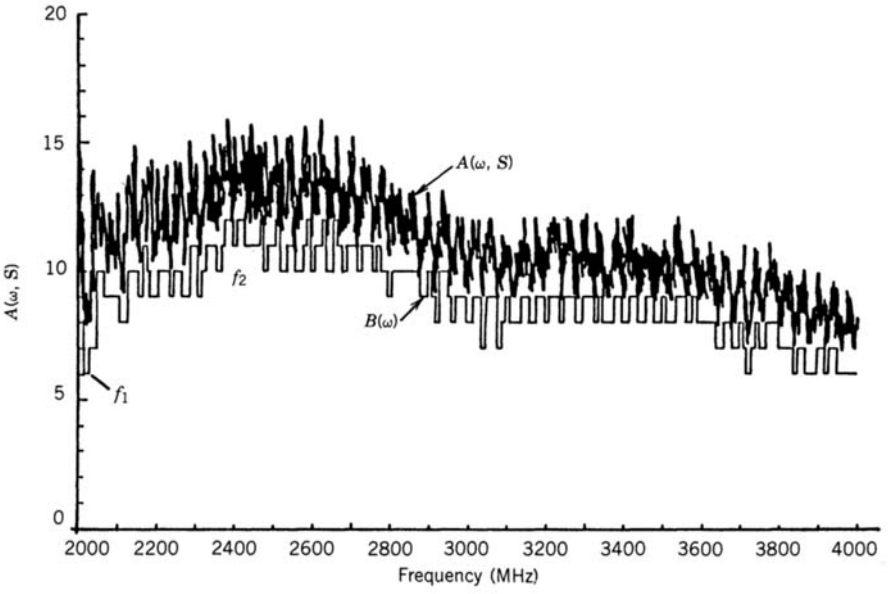


Figure 6.35. Measured $A(\omega,S)$ versus ω curve.

It is generated by subtracting a constant value from $A(\omega, S)$. If the value of $A(\omega, S)$ is above the threshold $B(\omega)$, the frequency reading should be correct. If $A(\omega, S)$ is below $B(\omega)$, there are overlapping signals at the input of the receiver.

To implement this idea in hardware, a ROM is used to store the threshold values $B(\omega)$. These threshold values are read from the ROM through the reported frequency of the receiver. Since the threshold does not vary rapidly with ω, not all the frequency bits are needed to address the ROM. For example, an IFM receiver with 11 bits of frequency information may only need 8 bits as the address bits of the threshold ROM. A/D converters are used to digitize the video outputs. Another ROM is used to implement the function

$$A^2(\omega, S)\sin^2\omega\tau + A^2(\omega, S)\cos^2\omega\tau = A^2(\omega, S) \tag{6.52}$$

The A/D sampling window and the receiver frequency sampling window should be matched in time. The measured $A(\omega, S)$ value is compared with the threshold value $B(\omega)$. The schematic diagram of this implementation is shown in Figure 6.36. This circuit can detect the overlapping signals from 70 to 95% of the time. As expected, this detection circuit will miss overlapping pulse conditions where there is no significant dip in the video output. This circuit will also miss pulses at a frequency where the threshold is low. For example, if the true input is f_2 (in Fig. 6.35) where the threshold is high and the erroneous frequency the receiver reported is at f_1, and since the true frequency is unknown, the overlapping signal detection circuit will compare the threshold at f_1 where the threshold is low. The $A(\omega, S)$ value may be lower than the threshold $B(\omega)$ at f_2, but it may be higher than the threshold at f_1; therefore, the overlapping signal condition will be missed.

The three methods discussed in the above sections are used to make decisions about the overlapping signal condition at the time of measurement. Another possible way to determine the overlapping pulses is to compare one measured frequency with an adjacent one in time. If the two readings are the same, the frequency is correct. If the two frequencies are different, there is the possibility of overlapping pulses. This kind of detection scheme may cause confusion when one frequency is correct and the other is wrong or when the input signal is frequency modulated.

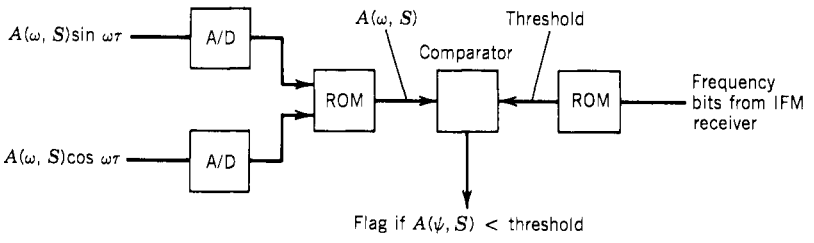


Figure 6.36. Digital simultaneous signal detection scheme.

6.22. RESAMPLING THE FREQUENCY INFORMATION

When a weak signal is followed by a strong one, the transient caused by the strong signal will affect the frequency measurements. This was discussed in former sections. However, if one can delay the frequency measurement time, the frequency of the strong signal (the second one) will be obtained because the strong signal will eventually suppress the weak signal and capture the measurement circuit. This kind of circuit can be implemented in an IFM receiver with an overlapping signal detector. If the overlapping detection circuit detects the existence of overlapping signals at the time a frequency sampling is made, the sampling gate can be delayed a short time (e.g., 50 nsec) and another measurement made. If the overlapping condition is as shown in Figure 6.37, the chance of obtaining correct frequency information on the resample is very high. There are many different ways to implement the resample scheme. A repetitive sampling scheme can be used, each sampling accompanying the output of the overlapping signal.

The resampling time cannot be delayed too long from the first frequency sampling window. If the delay is too long, when the input pulse is short, the second sampling may miss the pulse.

The following paragraph gives some idea on the results of resampling. A digital overlapping signal detector as discussed in Section 6.21 is used to control a resampling circuit. If overlapping signals are detected, the resampling circuit will take another measurement. The second measurement is also checked by the overlapping signal detector. The results are shown in Figures 6.38 and 6.39. The dotted lines represent the results with the resampling circuit on, and

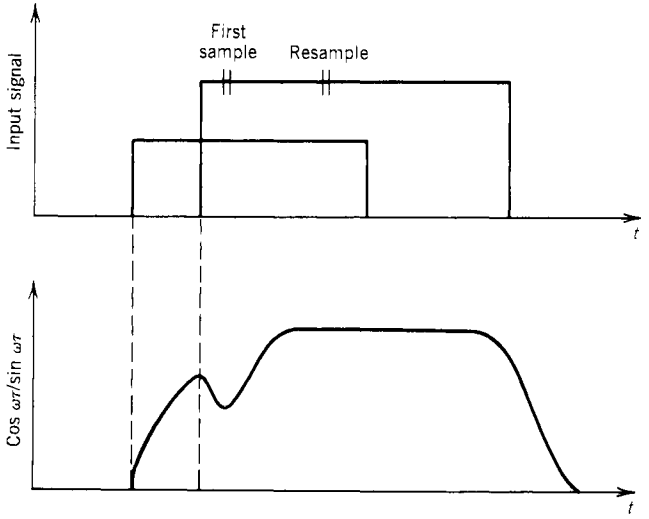


Figure 6.37. Time domain of resampling scheme.

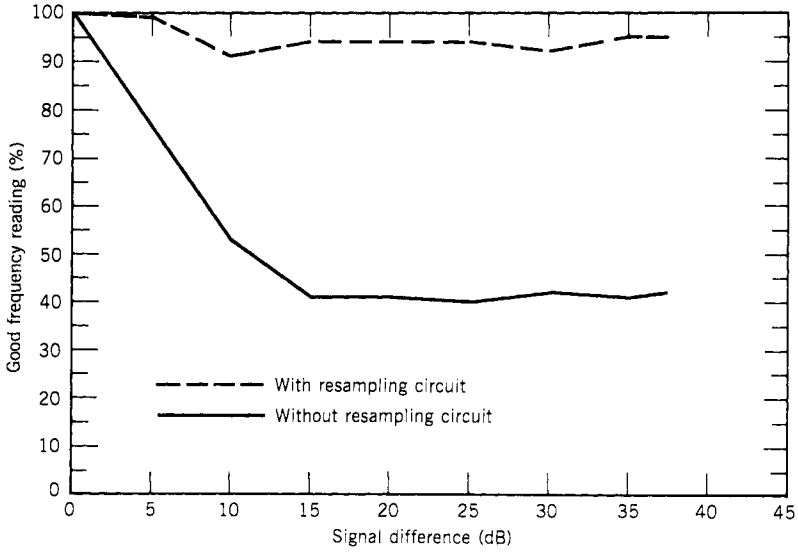


Figure 6.38. Test results of IFM receiver with frequency resampling circuit (second signal delay, 80 nsec).

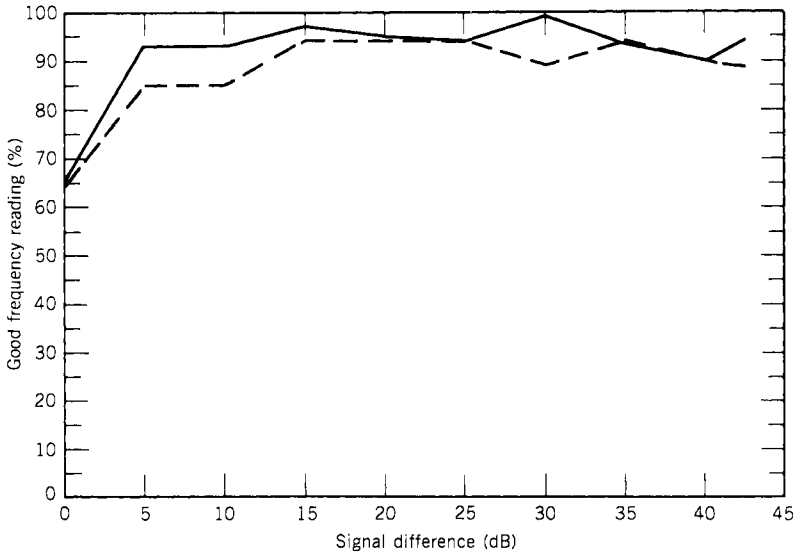


Figure 6.39. Test results of IFM receiver with frequency resampling circuit (second signal delay, 50 nsec).

the solid lines represent the results without the resampling circuit. The x axis represents the power difference in decibels, and the y axis gives the percentage of correct frequency readings.

In Figure 6.38, the results show that without resampling, when the second signal is stronger than the first by more than 15 dB the receiver reports only slightly higher than 40% correct frequency data. With the resampling circuit, the results improve to over 90% correct. Figure 6.39 shows a similar result. The major difference between Figures 6.38 and 6.39 is the input conditions. In Figure 6.38, the second signal is delayed 80 nsec from the first signal, whereas in Figure 6.39, the second signal is delayed by 50 nsec. In Figure 6.39, the resampling does not improve the results at all. However, the receiver reported approximately 90% correct data without the resampling circuit, which is rather high except when the two signals are less than 5 dB of each other. This implies that dip in the video output may not be significant enough to disturb the frequency measurement. Thus it is difficult to detect it.

From this experiment, one can see that when the IFM receiver reports a high percentage of correct frequency readings, the resampling circuit does not help much. But if the receiver reports a high percentage of erroneous data, the resampling circuit can help to improve the performance of the receiver, which is a desirable situation.

6.23. SUMMARY

An IFM receiver has the potential of being used in EW systems because it can cover very wide bandwidths, can provide fine-frequency resolution on short pulsed signals, and has moderately high sensitivity. One extremely attractive factor of IFM receivers is their very simple structure; therefore, the receivers can be made very compact, an important factor for EW applications. However, it has the problem of reporting information on only one signal at a time. It may report erroneous frequency information when simultaneous/overlapping signals are present. Several possible methods of improving the detection schemes of overlapping signal condition of the receiver have been discussed. Results show that the overlapping signal detection circuit combined with the resampling scheme can improve the performance of an IFM receiver under overlapping signal conditions. Further research to improve the receiver's capability to process simultaneous/overlapping signals is needed. The ideal solution is to design an IFM receiver with the capability to report multiple signals.

REFERENCES

1. C. W. Earp, Frequency indicating cathode ray oscilloscope, U.S. Patent 2434 914, January 27, 1948.
2. R. C. Cumming and G. A. Myers, Performance of receivers and signal analyzers using broadband frequency-sensitive devices, Technical Report No. 1905-1, Stanford Electronic Laboratories, SU-SEL-66-125, Stanford University, March 1976.

3. M. W. Wilkens and W. R. Kincheloe, Jr., Microwave realization of broadband phase and frequency discriminators, Technical Report No. 1962/1966-2, Stanford Electronics Laboratories, SU-SEL-68-057, November 1968.
4. G. A. Myers and R. C. Cumming, Theoretical response of a polar-display instantaneous-frequency meter, *IEEE Trans. Instrumentation Measurement*, **IM-20**, 38–48 (1971).
5. D. Heaton, The systems engineer's primer on IFM receivers, *Microwave J.*, **14**, 71 (February 1971).
6. S. J. Gourse and E. A. Worrell, Wide-band IFM flyable brassboard, Final Engineering Report, Litton Amecom, College Park, Maryland, November 1974.
7. N. M. Blachman, The effect of noise polar-display instantaneous frequency measurement, *IEEE Trans. Information Measurement*, **IM-25**, 214–221 (1976).
8. R. Grossbach, Degradation of polar-discriminator performance by non-ideal components, *Microwave J.*, **17**, 53 (December 1974).
9. D. L. Saul, Design a Ka-band polar frequency discriminator, *Microwave*, 74 (April 1976).
10. U. H. Gysel and J. P. Watjen, Wide-band frequency discriminator with high linearity. *IEEE MTT-S. Int. Microwave Symp.*, 1977.
11. R. V. Garver and J. A. Rosado, Broad-band TEM diode limiting, *IEEE Trans. Microwave Theory Techniques*, **MTT-10**, 302–310 (1962).
12. M. J. Rodriguez and D. Weissman, A microwave power limiter, *IEEE Trans. Microwave Theory Techniques*, **MTT-10**, 219–220 (1962).
13. F. E. Emergy, Solid-state limiting amplifiers, Watkins Johnson Co., Tech Notes, **5**(5) (September/October 1978).
14. H. Suhl, The nonlinear behavior of ferrites at high microwave signal levels, *Proc. IRE*, **44**, 1270–1284 (1956).
15. R. V. Garver and D. Y. Tseng, X-band limiting, *IRE Trans. Microwave Theory Techniques*, **MTT-9**, 202 (1961).
16. S. N. Stitzer, H. Goldie, and P. R. Emtage, Frequency-selective high power YIG limiters, *IEEE MTT-S International Microwave Symposium Digest*, pp. 272–274, 1976.
17. J. Helszajn, R. W. Murray, E. G. S. Davidson, and R. A. Suttrie, Waveguide subsidiary resonance ferrite limiters, *IEEE Trans. Microwave Theory Techniques*, **MTT-25**, 190–196 (1977).
18. P. R. Emtage and S. N. Stitzer, Interaction of signals in ferromagnetic limiters, *IEEE Trans. Microwave Theory and Technique*, **MTT-25**, 210–213 (1977).
19. G. S. Uebele, Characteristics of ferrite microwave limiters, *IRE Trans. Microwave Theory Techniques*, **MTT-7**, 18–23 (1959).
20. J. Brown, Ferromagnetic limiters, *Microwave J.*, **4**, 74 (November 1961).
21. S. N. Stitzer, H. Goldie, and P. S. Carter, Jr., X-band YIG limiters for FM/CW radar, *Microwave J.*, **20**, 35 (December 1977).
22. K. L. Kotzebue, Frequency-selective limiting, *IRE Trans. Microwave Theory Techniques*, **MTT-10**, 516–520 (November 1962).
23. R. C. S. Chien, Frequency-selective limiter and its application in a filter bank receiver, *Technical Note*, 1973-1, Lincoln Laboratory, MIT, April 13, 1973.
24. A. J. Giarola, D. R. Jackson, R. W. Orth, and W. P. Robbins, A frequency selective limiter using magnetoelastic instability, *Proc. IEEE Letter*, **5**, 593–594 (1967).
25. D. R. Jackson and R. W. Orth, A frequency selective limiter using nuclear magnetic resonance, *Proc. IEEE*, **55**, 36–45 (1967).
26. E. J. Wilkinson, An N-way Hybrid power divider, *IEEE Trans, Microwave Theory Techniques*, **MTT-52**, 116–118 (1960).
27. L. Parad and R. Moynihan, Split-tee power divider, *IEEE Trans. Microwave Theory Techniques*, **MTT-13**, 91–93 (1965).
28. G. F. Fike, Designing hybrid couplers with uneven power splits, *Microwaves*, 64 (November 1973).
29. C. Y. Ho, Hybrids in planar microwave integrated circuit, *Hybrid Microelectronics Seminar Notes*, Motorola, Phoenix, AZ, December 1976.
30. N. Nagai, E. Mackawa, and K. Ono, New *n*-way hybrid power dividers, *IEEE Trans. Microwave Theory Techniques*, **MTT-25**, 1008–1012 (1977).

31. H. Howe, Jr., Simplified design of high power, *n*-way in-phase power divider/combiners, *Microwave J.*, **22**, 51 (December 1979).
32. S. B. Cohn, A class of broadband three-port TEM-mode hybrids, *IEEE Trans. Microwave Theory Techniques*, **MTT-16**, 110–116 (1968).
33. R. B. Ekinge, A new method of synthesizing matched broadband TEM-mode three-ports, *IEEE Trans. Microwave Theory Techniques*, **MTT-19**, 81–88 (1971).
34. P. C. Goodman, A wide-band strip line matched power divider, *IEEE MTT-S International Microwave-Symposium*, pp. 16–20, 1968.
35. H. Y. Yee, F. C. Chang, and N. F. Audeh, *N*-way TEM-mode broad-band power divider, *IEEE Trans. Microwave Theory Techniques*, **MTT-18**, 682–288 (1970).
36. J. J. Taub and G. P. Kurpis, A more general *n*-way hybrid power divider, *IEEE Trans. Microwave Theory Techniques*, **MTT-17**, 406–408 (1969).
37. M. J. Head, Synthesize lumped element in phase power dividers, *Microwave J.*, **23**, 111 (May 1980).
38. W. B. Sullivan, Northern Scientific Laboratory, Fairfield, NJ. Private communication.
39. W. B. Sullivan, Frequency accuracy of digital frequency discriminator, part 1, Northern Scientific Laboratories, Fairfield, NJ, October 12, 1983.
40. W. B. Sullivan, Data compression receiver, Association of Old Crows, 12 Annual Electronic Warfare Symposium, pp. 23–25, September 1975.
41. Simultaneous signal performance of digital frequency discriminators, Anaren Microwave Inc., Syracuse, NY, April 1980.
42. *Digital Frequency Discriminators*, Technical Handbook, Pub. M1824-08, Anaren Microwave Inc., Syracuse, NY, 1977.
43. Performance of DFD's in the presence of noise, Pub. M1820-08, Anaren Microwave Inc., Syracuse, NY.
44. B. L. Bowler, Trade-offs in digital IFM receiver design, Pub. M1836-38, Anaren Microwave, Inc., presented at Joint RADC/Empire AOC Technical Seminar, Griffiss AFB, November 4, 1982.
45. C. W. Gerst and L. J. Paciorek, Angle digitizer, U.S. Patent No. 4277 748, July 7, 1981.
46. J. B. Y. Tsui and G. H. Schrick, Instantaneous frequency measurement (IFM) receiver with capability to separate cw and pulsed signals, U.S. Patent No. 4, 194,206, March 18, 1980.
47. R. L. Shaw and J. B. Y. Tsui, IFM receiver test and evaluation, Technical Report AFAL-TR-79-1049, Air Force Avionics Laboratory, Wright-Patterson Air Force Base, April 1979.
48. P. W. East, Design techniques and performance of digital IFM, *IEE Proc.*, **129**, 154–163 (1982).
49. J. B. Y. Tsui and R. Shaw, Simultaneous signal detection for Instantaneous frequency measurement (IFM) receivers by detecting the intermodulations product from non-linear devices. U.S. Patent 4,426,648.
50. Reference data for radio engineers, 4th ed., 460, International Telephone and Telegraph Corporation. American Book-Stratford Press, New York, 1956.
51. J. G. Graeme and G. E. Tobey, Ed., *Operational Amplifiers Design and Applications*. Chapter 6. McGraw-Hill, New York, 1971.
52. J. B. Y. Tsui, R. Shaw, and J. Caschera, Simultaneous signal detection capability for instantaneous frequency measurement (IFM) receiver. U.S. Patent No. 4,336,541, June 1979.
53. J. B. Y. Tsui, R. Shaw, J. Cisar, and T. Ratliff, Instantaneous simultaneous signal detecting. *Microwave J.*, **25**(12), 118 (December 1982).

Chapter 7

Channelized Receivers

7.1. INTRODUCTION (1–6)

To widen the radio-frequency (RF) coverage of a receiver, the most straightforward approach is to use brute force to build a large number of parallel narrow-band receivers with adjacent frequencies. A channelized receiver is built on this basic idea, which uses a large number of contiguous filters to sort the input signal. An input signal will pass a certain filter according to its frequency. By measuring the outputs of the filters, the frequency of the input signal can be determined. Although the idea is simple, this kind of receiver is bulky and expensive to fabricate because of the large number of filters required. With the advances in microwave integrated circuits (MIC) and surface acoustic wave (SAW) filters, the realization of a channelized receiver becomes practical.

The idea of channelization has been applied to many receiver designs. If a receiver is built to cover 2–18 GHz, one approach is to divide the input frequency range into sixteen 1-GHz parallel bands. Although this is a channelized approach, the receiver not considered to be channelized. It is only considered as a channelized front end. The channelized receiver discussed here not only uses the filters to separate input signals but also encodes the frequency to digital data. A channelized receiver is shown in Figure 7.1. On the left side, there are eight parallel channels; each channel has 16 fine-frequency slots. The upper right unit is the RF front-end down converter and the lower right unit is the power supply.

Research and development in channelized receivers has concentrated on filter designs. Although digitizing circuits following the filters to encode the input signals have been built, it seems that a systematic study to investigate the digitizing circuits is needed. Special effort is required to shrink the size of this kind of receiver.

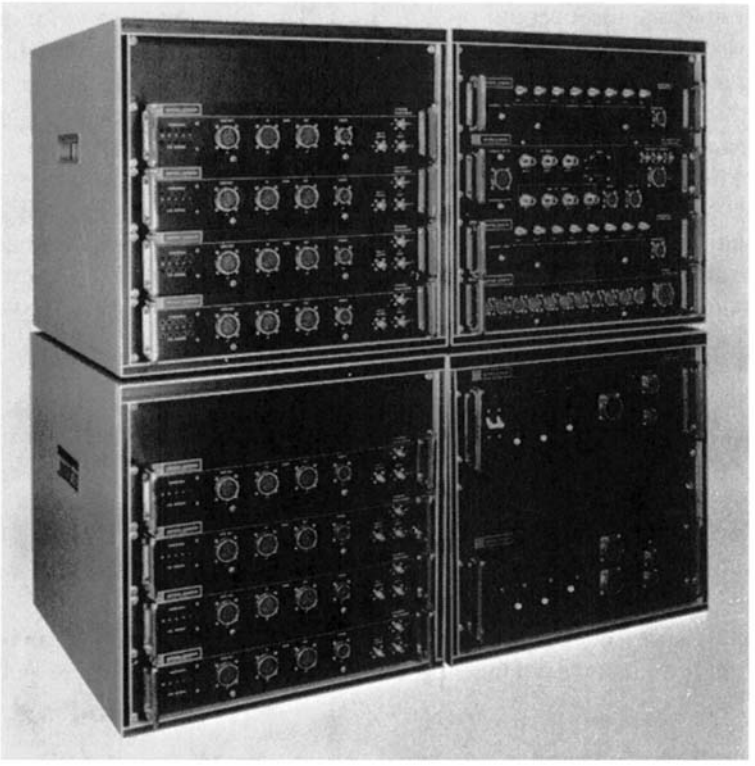


Figure 7.1. Channelized receiver. (Fabricated by Watkins-Johnson Co. Courtesy of Avionics Laboratory, AFWAL.)

In this chapter, some filter banks that can be used in a channelized receiver will be discussed. Some trade-offs and practical considerations will also be discussed. This chapter also includes a discussion of different approaches to encode the filter outputs.

7.2. DESIGN CONSIDERATIONS OF FILTER BANKS

In a channelized receiver, there are many parallel output slots. There are many components and circuits associated with each slot. These components include amplifiers, log amplifiers, limiters, video detectors, and video amplifiers. The size and cost of these components are the most important factors to be considered in choosing a channelized receiver. At present, components below 1000 MHz are relatively small in size and of low cost. Therefore, it is desirable to use filters under 1000 MHz as the fine-frequency measurement slot filters. However, the change in component technology may change these filter requirements.

Another important factor in selecting filters for a channelized receiver is determined by the outputs of the filters. The filters should accomplish two purposes. The first is to separate the input signals into different filters. If a strong and a weak signal are close in frequency, it is desirable to separate them in such a way that the strong one will not disturb the weak one. Under this requirement, the filters must have a sharp cut-off frequency (or a sharp "skirt"). Therefore, the filters will have many sections.

The second requirement is that the outputs from the filters will have less transient effect in time domain. The transient effect can be explained by the nature of the frequency domain of the pulsed signal as discussed in Section 2.17. For example, a square pulse of amplitude A and width T can be written as

$$S(t) = \begin{cases} A & \text{for } -\frac{1}{2}T < t < \frac{1}{2}T \\ 0 & \text{elsewhere} \end{cases} \tag{7.1}$$

where t is time. Its corresponding Fourier transform is

$$S(f) = \int_{-T/2}^{T/2} A \exp(-j\omega t)\, dt = AT \frac{\sin \pi fT}{\pi fT} \tag{7.2}$$

where f respresents frequency and $\omega = 2\pi f$. If this approach is carried out at different time instances along a pulse, a three-dimensional figure of a pulse

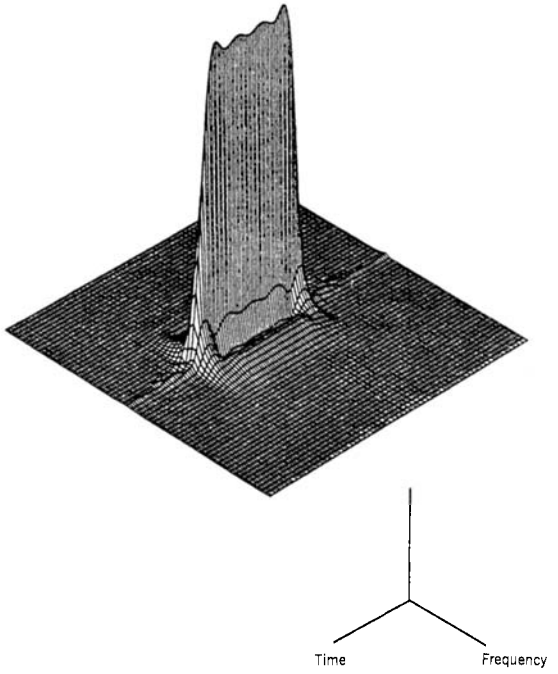


Figure 7.2. Frequency, amplitude (linear) versus time plot of a pulse.

signal is shown in Figure 7.2. Figure 7.3 shows the same result with amplitude in logarithmic scale.

Therefore, if a pulsed signal with a power spectrum as shown in Figure 7.2 and 7.3 passes the filter band, its energy will spread to many channels. The effect is demonstrated in Figure 7.4. A pulsed signal at the center of the filter bank will produce outputs not only from the center filter but also from the rest of the filters. The outputs from all the filters do not resemble the input, which is flat in amplitude. The output signals from the filters have larger amplitudes at the leading and trailing edges of the pulse, which is referred to as the transient effect (or rabbit ears). This transient effect is critical in channelized receiver design. In general, the more the energy spreads to other channels, the more complicated is the encoder design.

As a general rule, the narrower the filter bandwidth and the more sections the filter has, the more pronounced the transient effect. To have less transient effect, the filter used in a channelized receiver should have relatively wide bandwidth and less section. These requirements contradict the first requirement

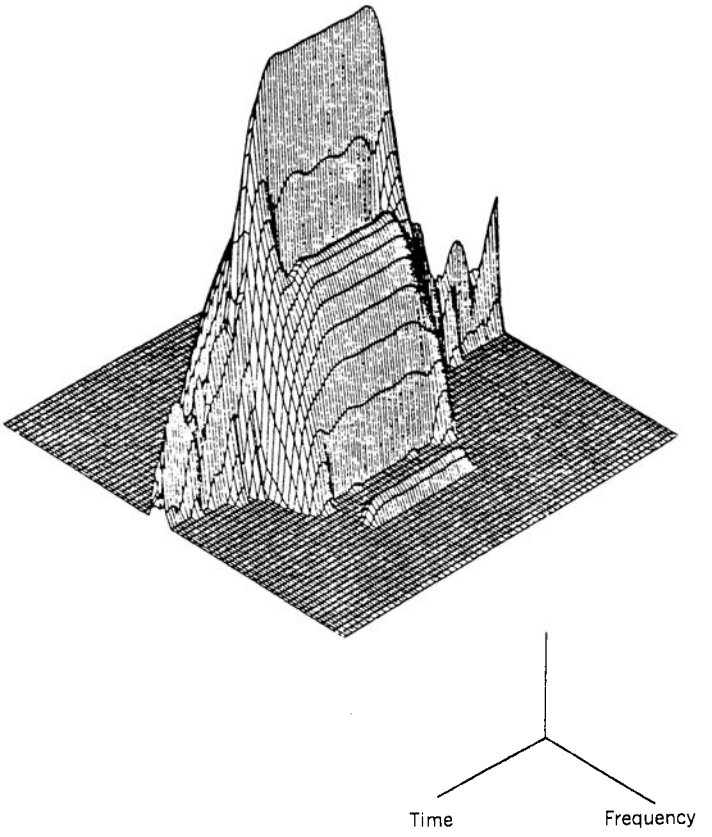


Figure 7.3. Frequency, amplitude (logarithmic) versus time of a pulse.

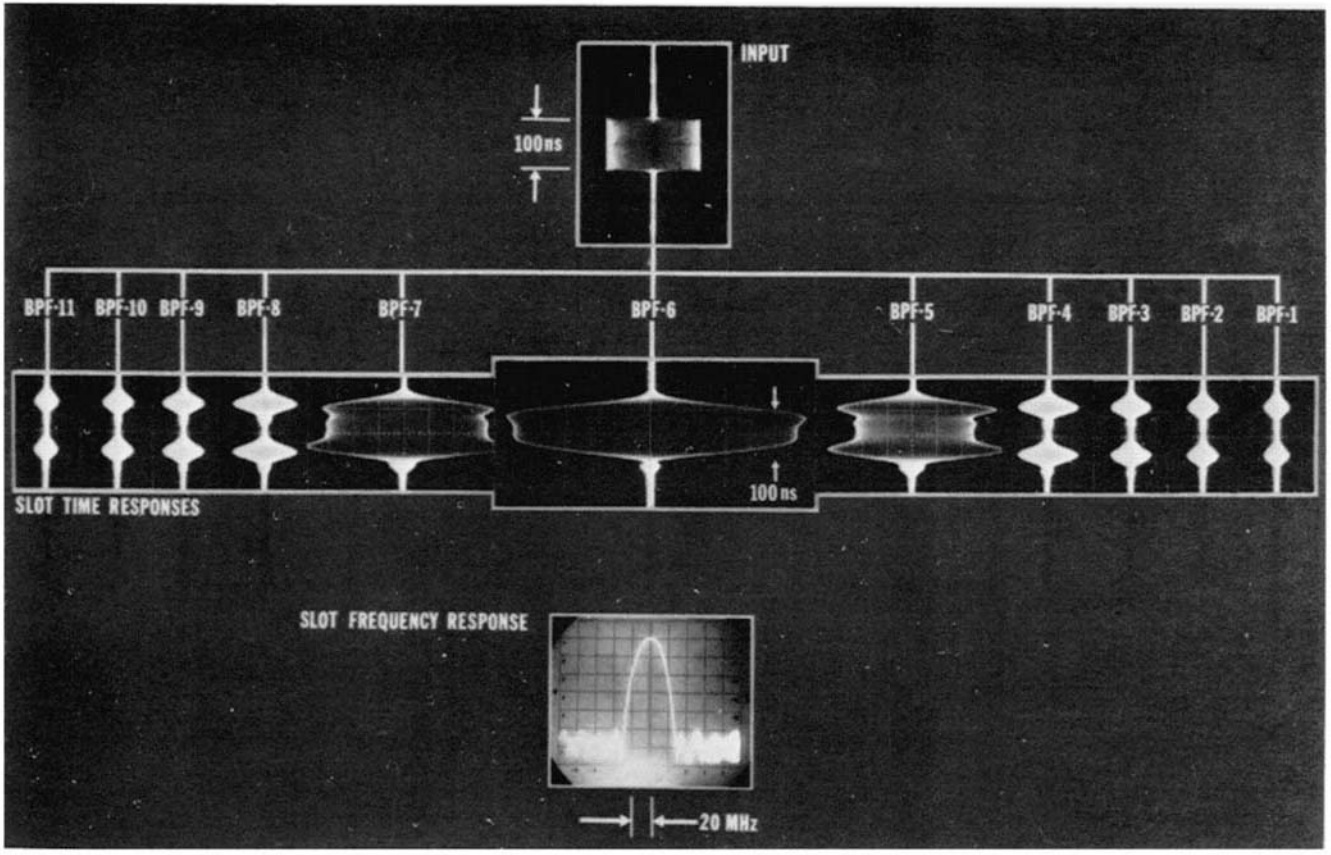


Figure 7.4. Output from a filter bank with one input signal. (Courtesy of Texas Instruments Inc.).

that the filters should separate the input signals into different slots. The choice of filters depends on the frequency resolution and the minimum pulse capability of the receiver, but at present there is no optimum selection procedure for filters used in channelized receivers.

The filter chosen from past experience was usually a composite, that is, two filters connected in cascade. The first filter has a relatively narrower bandwidth and fewer sections. This filter determines the transient effect of the signal. A second filter has wider bandwidth and more sections. This filter is used to provide the sharp skirt to separate input signals. As an example, if the minimum pulse to be processed is 100 nsec and the frequency resolution is 10 MHz, as an estimation, the first may have one to two sections with a bandwidth of 10–20 MHz. The second filter may have six to eight sections with a bandwidth from 40–60 MHz.

Other nonconventional filters (i.e., SAW filters) have this criterion as a general design guideline. Sometimes a Gaussian filter is considered desirable, because there is no sharp transient effect in this kind of filter. Hopefully, further research will provide more information concerning the selection of filters. The other problem in a channelized receiver is how to combine the individual filters into a filter bank. Filter banks will be discussed in the following sections.

7.3. FREQUENCY MULTIPLEXING THROUGH POWER DIVIDER (7–14)

It appears that the design of a frequency multiplexer could be easily accomplished by simply designing the required number of filters individually and then connecting them in series or in parallel. However, in such an approach, the input voltage standing wave ratio (VSWR) to the filter bank is very poor. The filters will interfere with each other. As a result, their characteristics will be different from what they are individually designed for.

The basic idea behind filter multiplexing is that if there are n filters connected in parallel, as shown in Figure 7.5*a*, or in series, as shown in Figure 7.5*b*, the input admittance and impedance should be unity (for a normalized system):

$$Y_i = Y_1 + Y_2 + \cdots + Y_n = 1 \tag{7.3}$$

or

$$Z_i = Z_1 + Z_2 + \cdots + Z_n = 1 \tag{7.4}$$

where Y_i and Z_i are the input admittance and impedance of the filter bank and $Y_1, Y_2, \ldots Z_1, Z_2, \ldots$ are the input admittance and impedance of each filter. Consequently, the sum of the real parts of the input admittances (or input impedances) must be constant and the sum of the reactive parts must be zero. In actual design, it is difficult to fulfill these conditions.

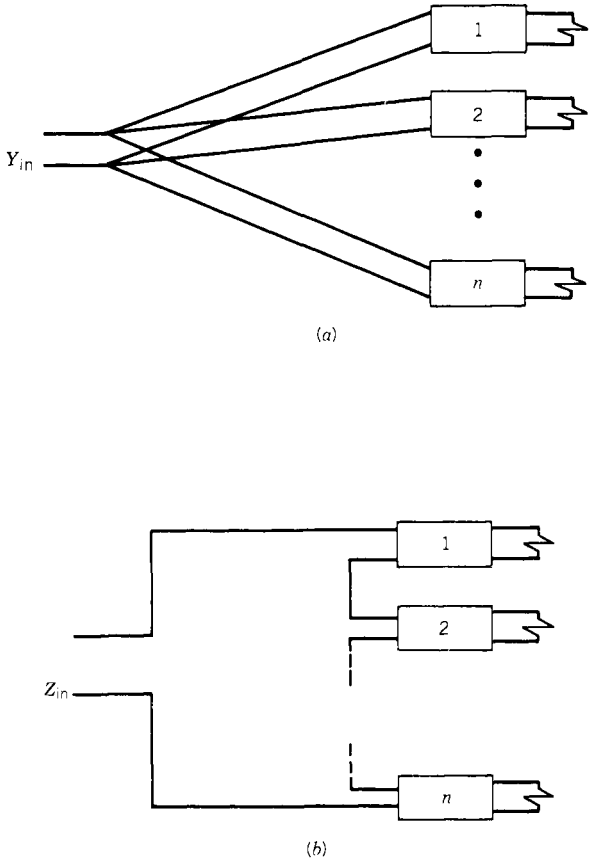


Figure 7.5. *N* filters connected in a filter bank: (*a*) in parallel: (*b*) in series.

One of the most common ways to achieve frequency multiplexing is through a power divider. At the outputs of the power divider, filters of different center frequencies are connected as shown in Figure 7.6. The input to the power divider is separated equally into its outputs. The signal will pass the filter with the appropriate center frequency. At all the other outputs of the power divider, the signal will be reflected back. This design can be considered a brute force approach. The advantage of this design is its simplicity and the low input VSWR if a resistive power divider (i.e., a Wilkinson power divider) is used. The disadvantage is that the insertion loss is high. The input signal is divided equally among the power divider outputs, which contributes a high loss. Every time the input is divided into 2, there is a minimum loss of 3 dB. Filter banks with many outputs can be fabricated with this design. The insertion loss must be recovered by additional amplification.

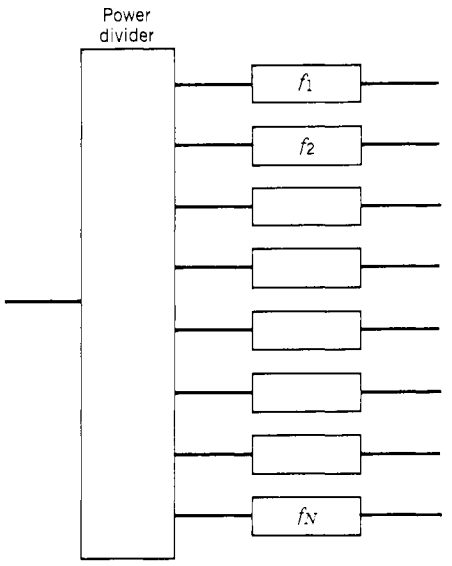


Figure 7.6. Frequency multiplexing through power divider.

7.4. FREQUENCY MULTIPLEXING THROUGH DIRECTIONAL FILTERS (7)

A directional filter is a four-port device that is specially designed for frequency multiplexing. Figures 7.7*a*,*b* show the four ports of a directional filter with the coupling between ports 1–4 and 1–2. The power incident at port 1 emerges from port 4 with the frequency response of a bandpass filter, while the remaining power emerges from port 2 with the response of a band-reject filter. Port 3 is isolated, and the input from port 1 does not reach port 3.

A form of stripline directional filter is known as the traveling wave loop directional filter (shown in Fig. 7.8). The directional filter in Figure 7.8 has two resonators. The number of resonators can be increased to improve the out of band frequency rejection. The mean circumference of the loops is a multiple of 360° at the midband frequency. Coupling between the loops is achieved through the two parallel sides. At the center of the operating frequency, the input signal from port 1 will excite a clockwise traveling wave in each of the loops. The traveling wave in the last loop excites a signal at port 4. At frequencies far from the resonator frequency of the loops, an input signal from port 1 will reach port 2 through the direct path.

The directional filters can be connected in cascade to form the frequency multiplexer shown in Figure 7.9. If the input signal contains frequencies f_a, f_b, f_c, f_d, and f_e, the f_a will emerge from filter *a*, and the remaining frequencies will pass filter *b*. The frequency f_b will emerge from filter *b*, and so on. The major

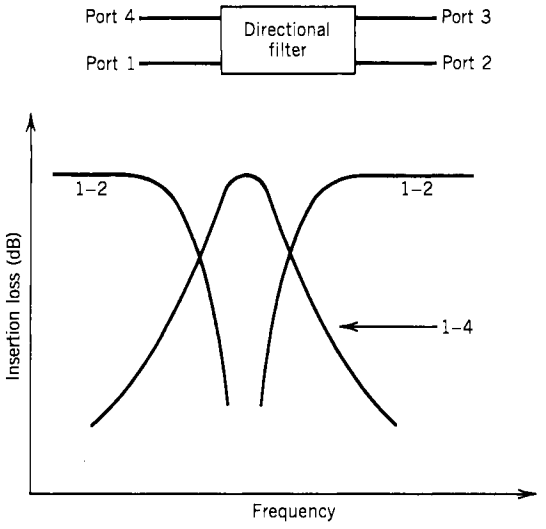


Figure 7.7. Frequency response of a directional filter.

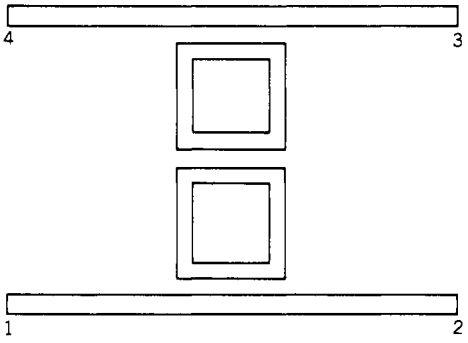


Figure 7.8. Traveling wave loop directional filter.

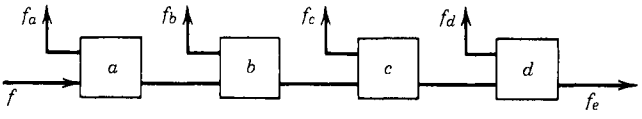


Figure 7.9. Multiplexer using directional filters.

drawback of the directional filter is that each resonator of the filter has two different orthogonal modes; and if more than two resonators are used in one filter, the tuning of the filters may be difficult. Another deficiency is that the longer the distance a certain frequency travels, the higher the insertion loss.

7.5. FREQUENCY MULTIPLEXING THROUGH MICROWAVE LENS (15–18)

The microwave lens discussed here is an analogy to an optic lens. It has a two-dimensional configuration built in a stripline form. The lens was original designed to measure angle of arrival (AOA) as discussed in Section 3.14. In the AOA application, the inputs of the lens are connected to individual antennas.

The lens can be changed as a frequency multiplexing device by modifying its inputs. The input network is a delay line, as shown in Figure 7.10. In Figure 7.10a, the input signal is coupled through a power divider and parallel delay

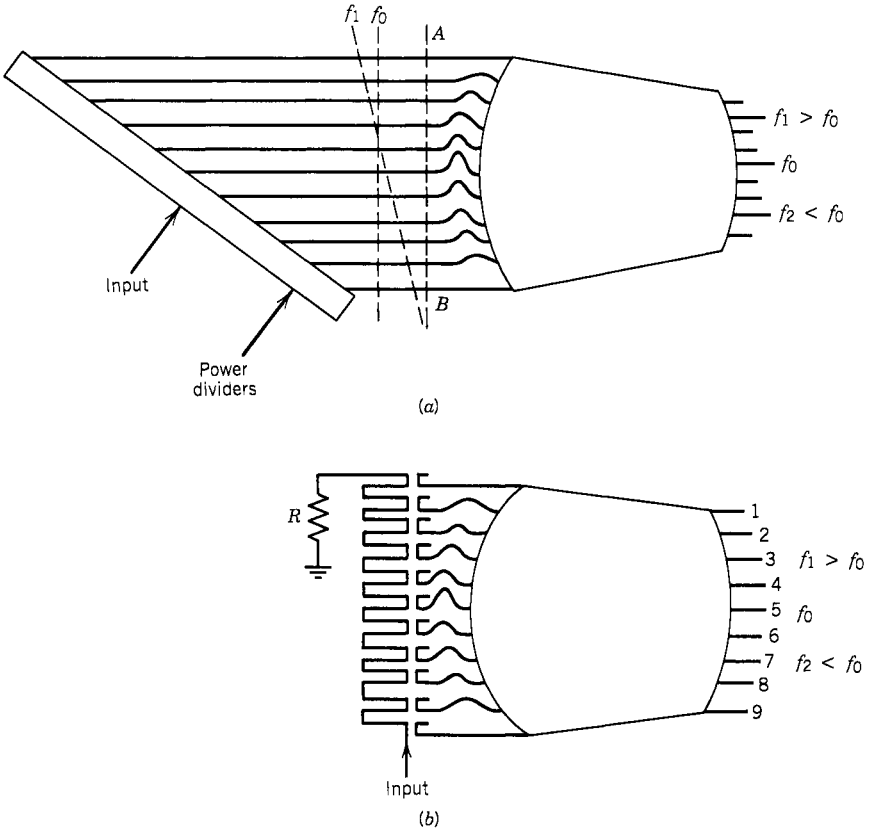


Figure 7.10. Microwave lens with delay line input network functioning as frequency multiplexer. (*a*) parallel feed; (*b*) series feed.

lines of different length. In Figure 7.10b, the input signal is coupled out from the delay line through hybrid couplers and fed to the inputs of the microwave lens. Each coupled output from the delay line is delayed consecutively by a length L. If

$$L = n\lambda_0 \tag{7.5}$$

where n is an integer and λ_0 is the midband wavelength corresponding to frequency f_0, the input signal with frequency f_0 will arrive at the input arc in phase and emerges from the center port. If the input frequency $f > f_0$, the wave front is tilted and emerges from output ports in the upper half of the lens. If the input frequency $f < f_0$, it will come out from ports in the lower half of the lens.

A breadboarded microwave lens with a delay line input network is shown in Figure 7.11. In the device, there are 11 input ports and 8 output ports. The outputs versus frequency are shown in Figure 7.12. The side lobes are relatively high, but the performance can be improved by increasing the number of input ports and properly weighting each input. To keep the size of the lens small, the operation frequency is usually in the gigahertz range. The number of input and output ports is also limited by the actual size of the device. The bandwidth of the

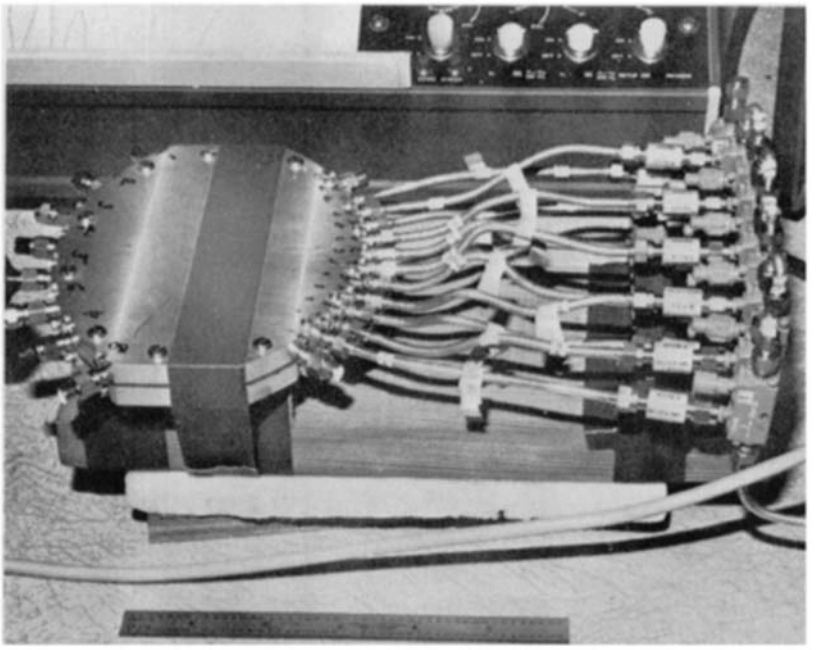


Figure 7.11. Microwave lens with input network. (Fabricated by Raytheon Co. Courtesy of Avionics Laboratory, AFWAL.)

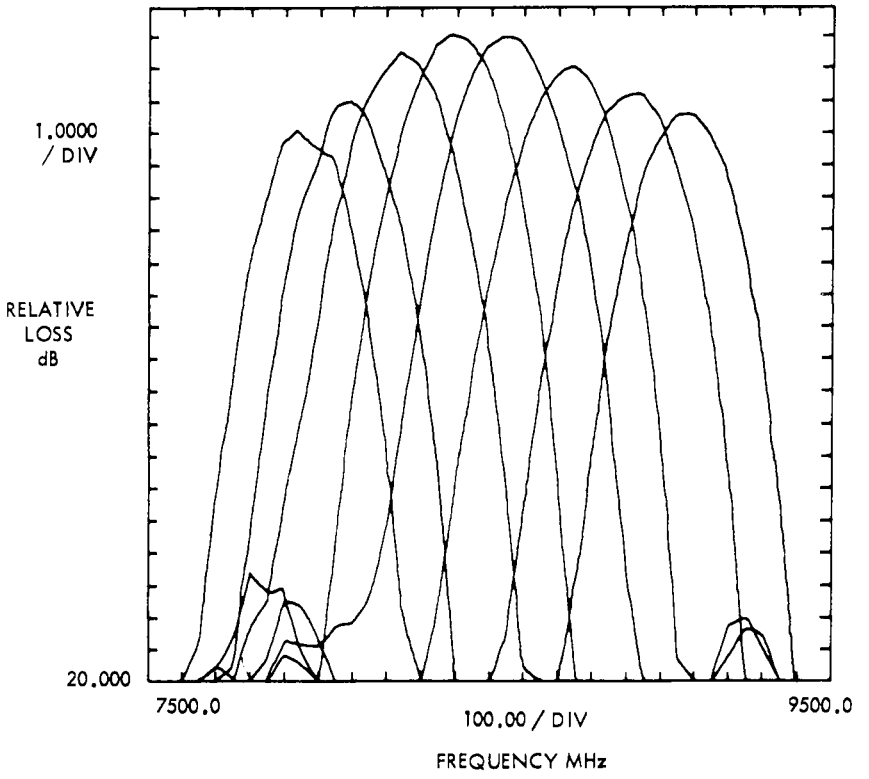


Figure 7.12. Microwave lens outputs versus frequency: minimum loss, 8.16 dB.

output can be changed by changing the n value in Eq. (7.5). It is impractical to make microwave lenses of reasonable size for frequency multiplexing under 1 GHz because the size will be very large. However, this design idea can be applied to SAW technology, which can be kept under 1 GHz with relatively small size.

7.6. HYBRID COUPLERS AND THEIR APPLICATIONS TO FREQUENCY MULTIPLEXING (7, 19–33)

A hybrid coupler is a four-port passive microwave device. Figure 7.13 shows a parallel line hybrid coupler. The input from port 1 will emerge from ports 2 and 3, while 4 is an isolation port. The two conductors 1–2 and 3–4, which are about $\frac{1}{4}\lambda_0$ in length, where λ_0 is the wavelength of the midband, are placed close to each other. The operating principle of the hybrid coupler can be explained with the help of Figure 7.14. The input energy can be coupled from conductors 1–2 to 3–4 in two ways: inductive and capacitive couplings. The

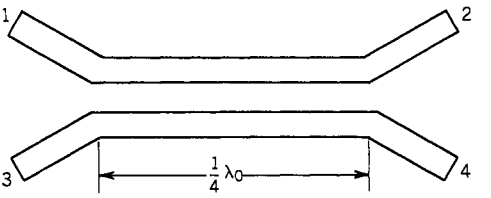


Figure 7.13. Parallel line hybrid coupler.

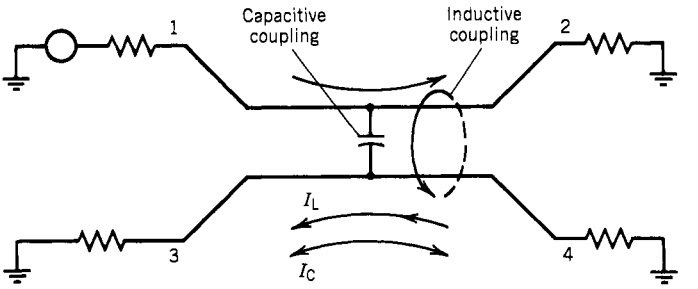


Figure 7.14. Capacitive and inductive couplings between two parallel conductors.

inductive coupling induces current I_L flowing from 4 to 3, whereas the capacitive coupling induces current I_c flowing from the center of the conductor to both 3 and 4. If the device is properly designed so that the amplitude of I_L and I_c are equal, then the current at port 4 will cancel and the current at port 3 will add up. If port 1 is the input, 2 will be the transmitted port, 3 will be the coupled port, and 4 will be the isolated port.

The output at port 3 leads the output from 2 by 90° in phase; thus the device is called a 90° hybrid coupler. A general expression for a 90° hybrid coupler is

$$V_2 = V_{\text{trans}} = \cos\theta \exp[-j(\beta l + \varepsilon)] \tag{7.6}$$

$$V_3 = V_{\text{coupled}} = j \sin\theta \exp[-j(\beta l + \varepsilon)] \tag{7.7}$$

where the input at port 1 has unit amplitude, β is the propagation constant $(2\pi/\lambda)$, l is the coupled length, ε is a small phase dispersion error term, and θ is the coupling angle. For an ideal case ($\varepsilon = 0$ and $l = \frac{1}{4}\lambda$),

$$V_2 = \frac{-j}{\sqrt{2}} \cos\theta \tag{7.8}$$

$$V_3 = \frac{1}{\sqrt{2}} \sin\theta \tag{7.9}$$

If the amplitudes of V_2 and V_3 are not equal, the device is called a directional coupler.

Hybrid couplers can be used for frequency multiplexing as shown in Figure 7.15. The input signal is divided equally in amplitude between ports 2 and 3 where two filters with same center frequency f_a are connected. The outputs from these two filters are combined by another hybrid coupler, and its output is at port a. The frequency rejected by these two filters emerges from port 4 but not from port 1 because at port 1 the two reflected waves are 180° out of phase. Port 4 can be considered to be the input signal to the next circuit, as shown in Figure 7.15. Since the reflected power does not go back to the input, the VSWR of the frequency multiplexer is relatively low.

The disadvantages of this approach are (1) the number of filters is doubled and (2) the frequency coming out of the last filters will have higher insertion loss because it travels the longest distance.

Hybrid couplers are also used in mixers (discussed in Chapter 5) and in phase correlators (discussed in Chapter 6).

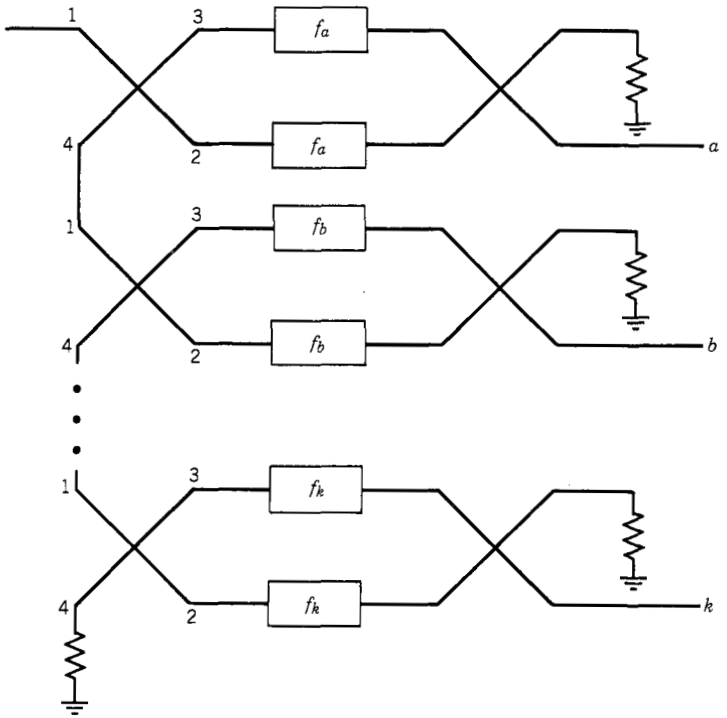


Figure 7.15. Frequency multiplexing using hybrid coupler.

7.7. DIELECTRIC FILTER BANK (34–53)

Before discussing dielectric filter banks, let us first take a look at an individual dielectric filter. It is known that the electromagnetic (EM) field stays inside dielectric or ferrimagnetic materials when they have free space boundaries. Materials of high dielectric constant or high permeability can be used to fabricate resonators. The EM field of a given resonant mode will be confined in or near the resonator. The field outside the resonator attenuates to negligible values at a distance less than the free-space wavelength. Therefore, the main loss in the resonator is caused by the loss of the dielectric material rather than by radiation. The unloaded Q_u of such a resonator is given by

$$Q_u = 1/\tan \delta \tag{7.10}$$

where $\tan \delta$ is the loss tangent of the material. Materials of high dielectric constant are of primary interest. Polycrystalline TiO_2 ceramic, also referred to as rutile, has a dielectric constant of about 100 and an unloaded Q_u of 5000–10,000. Thus this material is very attractive for fabricating dielectric filters. However, the main problem with TiO_2 is that the dielectric constant veries rapidly with temperature. It changes about 1000 ppm/°C. The resonant frequency is related to the dielectric constant (ε_r) by

$$f \propto \sqrt{\varepsilon_r} = \sqrt{\varepsilon_{r0} + \Delta\varepsilon_r} \simeq \varepsilon_{r0} + \tfrac{1}{2}\Delta\varepsilon_r \tag{7.11}$$

where ε_{r0} is the dielectric constant at temperature T_0 and $\Delta\varepsilon_r$ is the change in ε_r due to temperature. From the second term in Eq. (7.11), it can be derived that a 1000-ppm/°C change in dielectric constant will cause approximately a 500-ppm/°C change in frequency shift. For example, if the center frequency of a dielectric filter is at 3000 MHz, a temperature change of 20 °C will cause the center frequency to shift by 30 MHz ($500 \times 10^{-6} \times 3000 \times 10^6 \times 20$), which is unacceptable in most receiver applications. A ceramic material ($Ba_2Ti_9O_{20}$) (ref. 43) with $\varepsilon_r = 40$ and a temperature stability of 0–2 ppm/°C can be used to build dielectric filters with sufficient temperature stability. Another material with a commercial name of K38 has a typical ε_r of 38, $\tan \delta = 4 \times 10^{-4}$, and a temperature coefficient of -50 ppm/°C (ref. 49). Because of the availability of the material with a high dielectric constant and temperature stability, dielectric filters have been developed for possible receiver applications.

A common dielectric resonator is a cylindrical disk, of diameter D and height L as shown in Figure 7.16. When $L < D$, $TE_{01\delta}$ is the dominant mode in the resonator. The electric field E and magnetic field H are shown by the solid and dotted lines, respectively. In $TE_{01\delta}$ the first two subscript integers denote the waveguide mode and δ is the noninteger ratio ($=2L/\lambda_g < 1$, where λ_g is the waveguide wavelength).

The resonant frequency of the dielectric resonator has been calculated by many authors (refs. 34, 40–43). However, there is no closed-form solution to

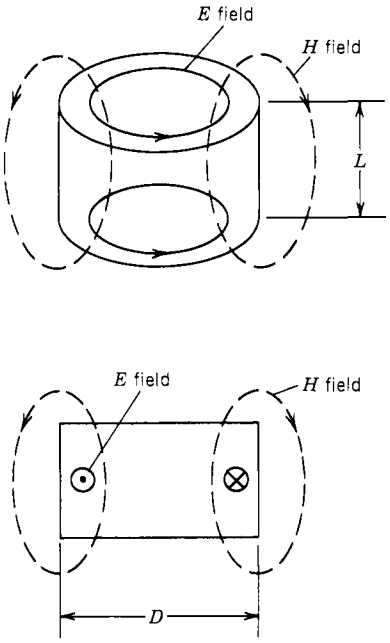


Figure 7.16. $TE_{01\delta}$ mode of dielectric resonator.

this problem. The resonant frequency of the dominant mode in a dielectric resonator placed in free space has been computed as (ref. 41)

$$\frac{J_0'(hD/2)}{hJ_0(hD/2)} + \frac{K_0'(pD/2)}{pK_0(pD/2)} = 0 \tag{7.12}$$

$$\beta \tan \frac{\beta L}{2} = \gamma \tag{7.13}$$

where

$$\beta^2 = \varepsilon_r k_0^2 - h^2 = k_0^2 + p^2 \tag{7.14}$$

$$\gamma^2 = h^2 - k_0^2, \qquad k_0 = \omega_0 \sqrt{\varepsilon_0 \mu_0} \tag{7.15}$$

where J_0 and K_0 are the Bessel and modified Hankel functions of order zero, ω_0 is the angular resonant frequency, D and L are the diameter and height of the dielectric disc, ε_r is the dielectric constant, and μ_0 and ε_0 are the permeability and permittivity of free space. The numerical results are plotted in Figure 7.17.

From the results of Figure 7.17, it is obvious that the higher the resonant frequency, the smaller the resonator. Masse and Pucel (ref. 49) used the K38 to build a four-resonator bandpass filter with a center frequency of 9.65 GHz

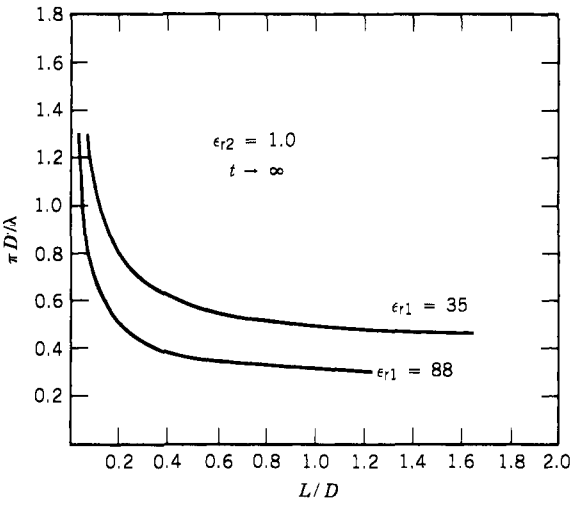


Figure 7.17. Numerical results of dielectric resonator frequency. (Based on Itoh and Rudokas, ref. 41.)

and a bandwidth of 50 MHz. Plourde and Linn (ref. 43) used $Ba_2Ti_9O_{20}$ to build a three-section bandpass filter with a center frequency at 4.11 GHz and a bandwidth of 26 MHz.

Lammers and Stiglitz (ref. 50) used dielectric (strontium titanate) resonators to build a 96-channel multiplexer that covers approximately 1 GHz bandwidth centered at 5 GHz. Each channel contains one resonator. The input to the multiplexer is fed through a transmission line, and the resonators are placed close to the transmission line to couple the energy out. Since the dielectric thermal coefficient of strontium titanate is about 2000 ppm/°C, the center frequency of the filters drifts approximately 10 MHz°C. Although the frequency shift is too high to have any practical applications, the feasibility of using dielectric resonators to make a multiplexer is demonstrated.

Harrison and Cohn (ref. 53) adapted a similar approach to couple the input signal from the feeding transmission line and built a two-resonator frequency multiplexer. They also built a five-resonator frequency multiplexer. The basic configuration of the five-resonator multiplexer is shown in Figure 7.18. The multiplexer contains four channels. The five dielectric resonators of each channel are placed in one housing. There is a coupling window at one end of the housing to couple the energy from the transmission line, which is at the center of the device, to the dielectric resonators. At the other end of the housing there is an output and a coupling adjustment mechanism. The insertion loss of each channel is about 10 dB with a bandwidth of 20 MHz and center frequency of about 8.1 GHz. At the band crossing, the insertion loss is about 3 dB below the center of the band. The frequency outputs are shown in Figure 7.19.

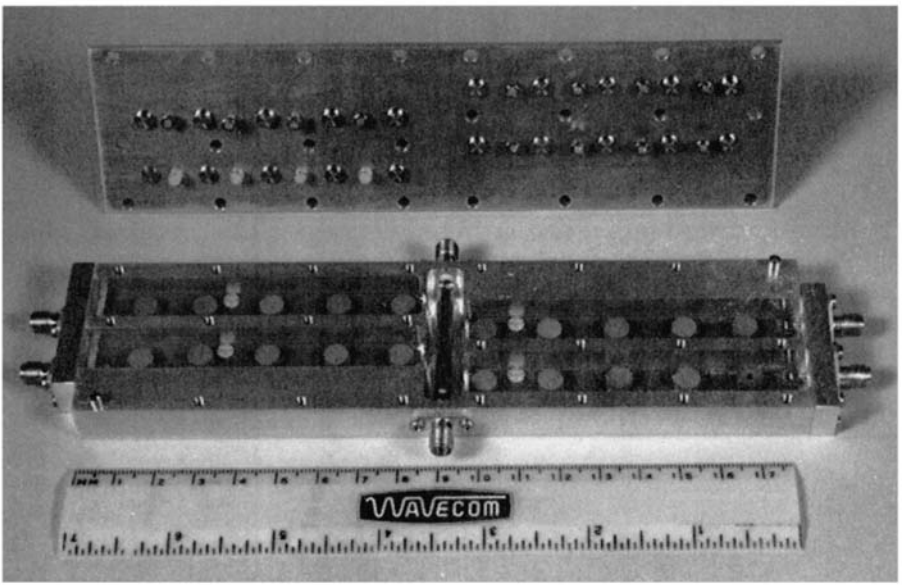


Figure 7.18. Basic configuration and output frequency of a four-channel dielectric resonator multiplexer. (Fabricated by Microwave Sources Inc. Courtesy of Avionics Laboratory, AFWAL.)

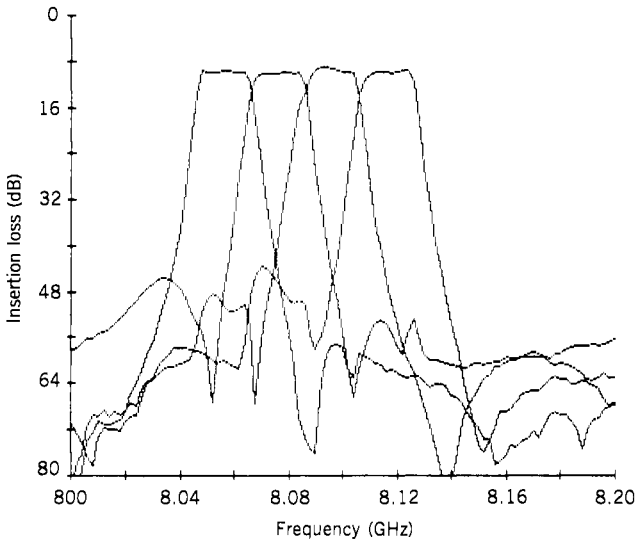


Figure 7.19. Output frequency responses of a five-dielectric-resonator, four-channel multiplexer.

To keep the dielectric filter at a reasonable size, the operating frequency of the filter is relatively high. Most of the dielectric filters operate above 2 GHz, and ordinarily they operate at around 10 GHz.

From the discussions in Section 7.2, it is obvious that the operating frequency of microwave lenses and dielectric filters are too high to be adopted in a channelized receiver unless low-cost amplifiers to a range of 10 GHz are available. A mixer can be used at the output of each filter to down convert the frequency for further processing, but the mixers with the proper local oscillators (LOs) will increase the complexity of the receiver design. However, the dielectric filter multiplexers may be used in receiver front-end channelization because of their high operating frequency.

7.8. SURFACE ACOUSTIC WAVE (SAW) FILTERS (54–63)

A SAW filter is a bandpass filter fabricated on a piezoelectric substrate. The input electric signal to the SAW filter is changed to an acoustic wave, filtered in the acoustic domain, and changed back to an electric signal at the output of the filter. The advantage of using acoustic waves to build a filter is the extremely small size and the possibility of mass production through integrated circuit techniques. The frequency response of the filter can be designed with great flexibility. A SAW filter can operate from a few megahertz up to 1 GHz.

The input signal to a SAW filter is converted to acoustic waves through a transducer. A typical interdigital transducer is fabricated on the surface of a piezoelectric surface as shown in Figure 7.20. When an electric voltage is applied to the transducer, the electric field between the individual fingers causes the surface of the piezoelectric material to distort. Therefore, an acoustic wave

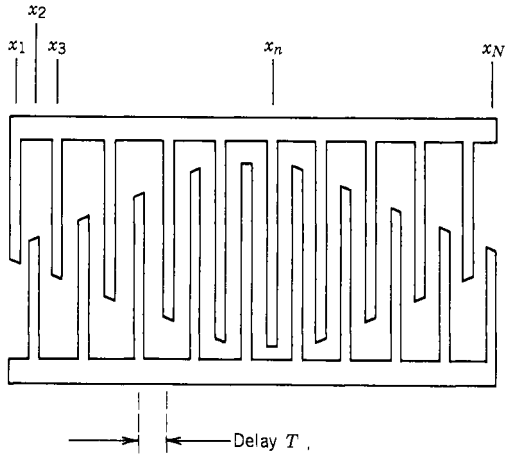


Figure 7.20. Typical interdigital transducer with apodization.

is induced on the surface of the substrate. The acoustic wave will travel in both directions perpendicular to the fingers. The overall acoustic wave is the summation of all the individual waves generated by the electrodes. Since the overlap of the electrodes varies in the wave propagation direction (referred to as apodization), as shown in Figure 7.20, the contribution to the overall acoustic wave is proportional to the electrode overlap.

The frequency response of the filter can be expressed as (ref. 56)

$$H(f) = \sum_{i=1}^{n} a_i \exp(j2\pi X_i f/v) \tag{7.16}$$

where a_i is the overlap weighting at each electrode, X_i is the location of the electrode, $j = \sqrt{-1}$, v is the surface acoustic wave velocity, and f is the frequency. If the distance between two adjacent electrodes is a half-wavelength, the corresponding delay time is T, and the transducer is symmetric with $2n + 1$ uniformly spaced electrodes, the transfer function is of the form

$$H(f) = 1 + 2 \sum_{i=1}^{n} a_i \cos(2\pi f T_i) \tag{7.17}$$

From this equation, it can be seen that $H(f)$ is a real function and is dispersionless for any transducer that is symmetric. Its frequency response has a repetition pattern as shown in Figure 7.21. The repetition is due to the discrete nature of the electrodes because the wave is sampled only at distinct points rather than sampling the entire acoustic wave.

The most common way to build a SAW filter bank is to use a power divider to separate the input into many parallel channels. At each output, a SAW filter of different frequencies is connected as discussed in Section 7.3. However, some

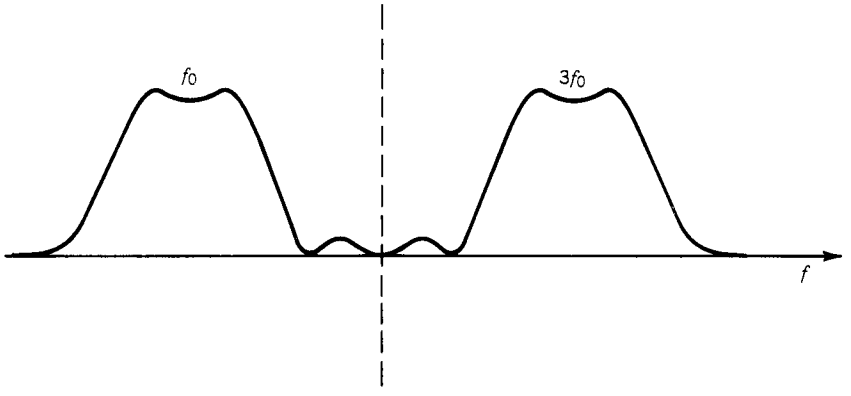


Figure 7.21. Frequency response of an ideal SAW filter.

special ways of multiplexing can be designed to suit SAW filters. These multiplexing schemes may be part of the filter designs. Several of the multiplexing schemes are discussed below.

1. The input transducers of SAW filters with different frequencies can be connected in parallel. A wide-band amplifier is used to drive these filters. One amplifier is used to drive four contiguous SAW filters (in ref. 60). The output impedance of the amplifier is designed to match the input impedance of the combined SAW filters for maximum power transfer. The amplifier is specially designed, and the operating bandwidth of this approach is limited.

2. Another approach uses the series and parallel combination connections of transducers as shown in Figure 7.22. In this figure, two of the input transducers are connected in series and then they are connected in parallel. When the transducers are connected in series, the input capacitance will be reduced. The exact arrangement of connection depends on the individual impedance of each input transducer. In this approach, the impedance may not be properly matched because it is limited by the individual transducer impedance.

3. The SAW filters can be multiplexed through a multistrip coupler (refs. 62, 63). Multistrip lines in front of a transducer form a 3-dB coupler as shown in Figure 7.23. A surface wave incidental on a multistrip coupler in track B will emerge from tracks A and B with equal amplitudes. The output from A is advanced 90° in phase relative to that of track B. This idea can be applied to frequency multiplexing as shown in Figure 7.24*a*. With a delay time τ built in

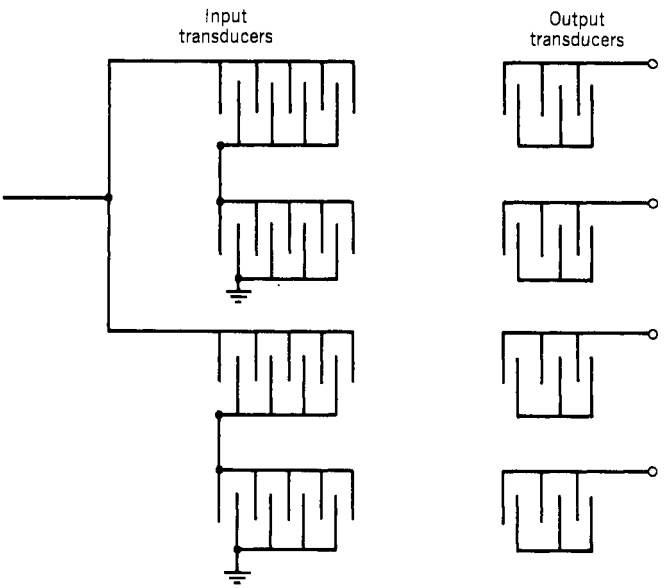


Figure 7.22. Series–parallel connections of input transducers of a SAW filter.

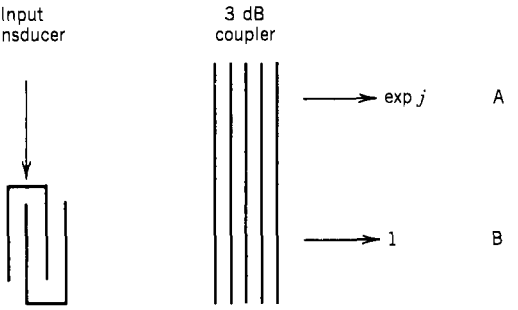


Figure 7.23. A 3-dB multistrip coupler.

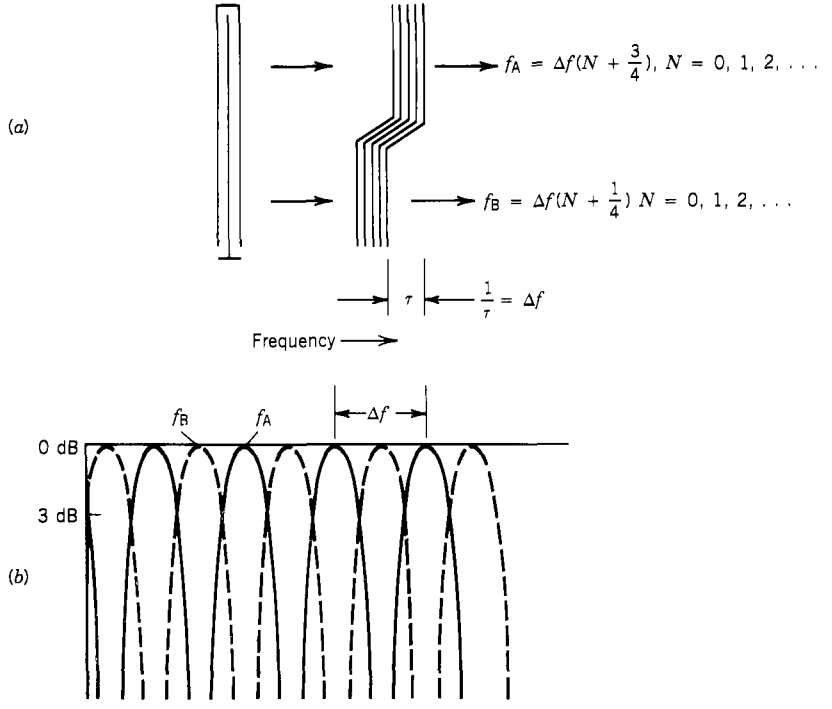


Figure 7.24. Frequency response of offset multistrip coupler, solid lines, track A; broken lines, track B; (*a*) filter configuration; (*b*) frequency response. (Based on Van de Vaart and Solie, ref. 63.)

the multistrip coupler, the frequency outputs from A and B are shown in Figure 7.24*b*. This idea can be used repeatedly to build a multiple-channel SAW multiplexer. A four-channel multiplexer is shown in Figure 7.25. The input is a long transducer, and the outputs are four separated transducers. In this special arrangement, the out-of-band rejection is rather poor.

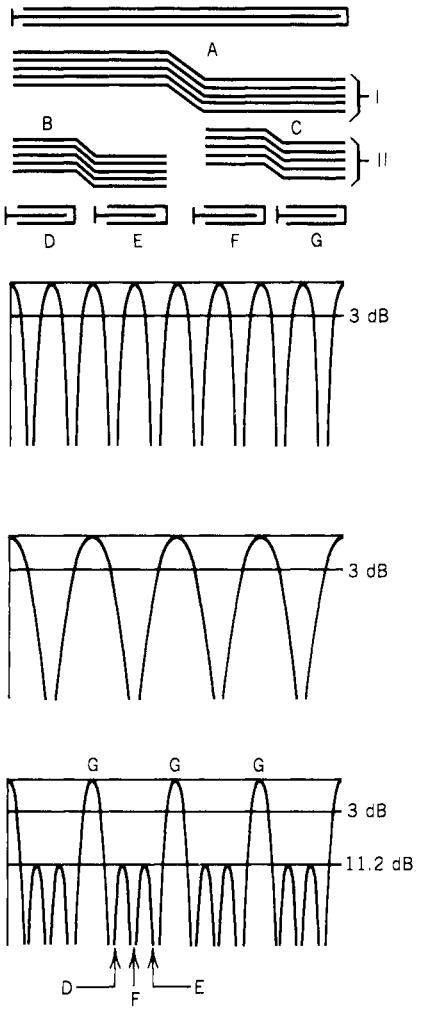


Figure 7.25. Scheme of a four-channel multistrip coupler multiplexer together with predicted transfer function. (Based on Van de Vaart and Solie, ref. 63.)

4. Multiplexing can be accomplished through reflectors as shown in Figure 7.26. Frequency-dependent reflectors are used to reflect different frequencies to different output ports. The reflectors can be either metal strips or grooves etched in the substrate. The spaces between the reflectors are not uniform. The high frequency is deflected to output 1, and the low frequency is deflected to output 3. Since the input transducer is bidirectional, deflectors on the opposite side can be added to obtain more outputs. In this arrangement, the outputs have different delay times. Some receiver design can take this advantage. If equal delay time is required, the output transducers can be properly placed.

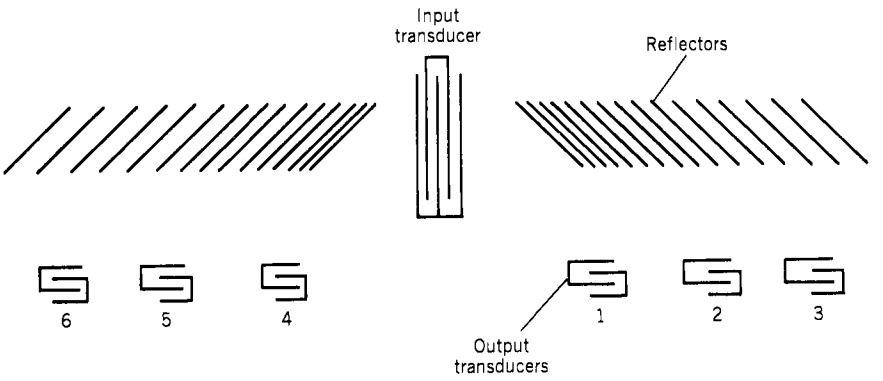


Figure 7.26. SAW multiplexing through frequency-dependent reflectors.

Although the above multiplexing schemes are very cleverly designed, it is difficult to accomplish the frequency multiplexing and at the same time provide the desired filter responses. Therefore, if the SAW filter shapes of a multiplexer are given, the general approach is still to design each filter individually and multiplex them through a power divider.

7.9. CHARACTERISTICS OF SAW FILTERS (64–69)

The SAW filter is one of the most promising filters for building channelized receivers. It is small in size, has the proper frequency range, has tailored frequency response, and is potentially suitable for mass production. It also has some special characteristics that are in many ways different from a conventional filter. These characteristics have serious impacts on receiver designs. They can be discussed in both the frequency and time domains.

The insertion loss of a SAW filter is very high, and this is one of the worst deficiencies of the filter. Although one can design a special SAW filter with low insertion loss, a SAW filter bank with some given filter response usually has very high insertion loss (approximately 20 dB). The excess loss must be recovered by additional amplifiers, which will increase the power consumption and the cost of the receiver.

The frequency response of a single SAW filter is shown in Figure 7.27. To demonstrate the difference in frequency responses between a SAW filter and a conventional filter, note that the SAW filter performance displayed in Figure 7.27 is inferior to a typical SAW filter. With today's technology, a typical SAW filter will have lower out-of-band side lobes. The out-of-band attenuation of the filter does not increase monotonically as in a conventional filter. Instead, there are many side lobes. These side lobes will limit the dynamic range of the receiver. If the filters used to build a receiver have 50-dB side lobes below the

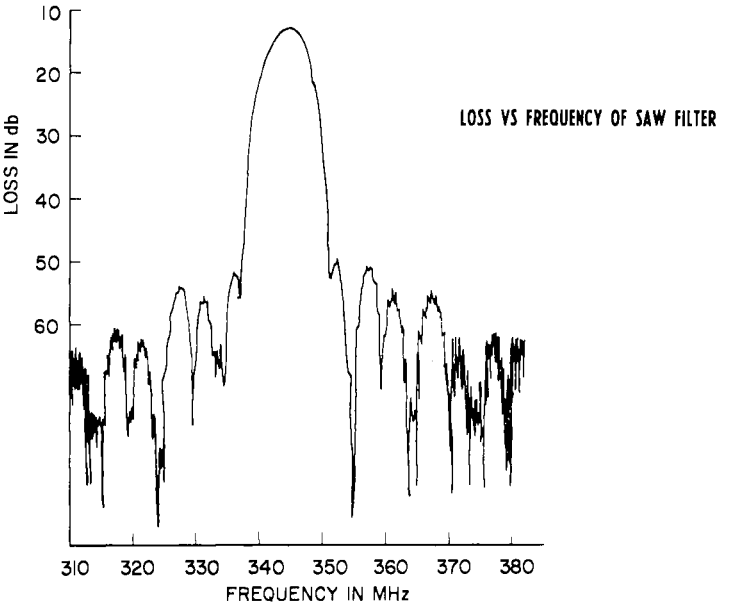


Figure 7.27. Frequency response of a SAW filter.

main output, the receiver's dynamic range is less than 50 dB, since a certain signal-to-noise ratio (S/N) is required to produce a decent false alarm rate.

The delay time in a SAW filter is rather long, in the microsecond range. A time domain response of a SAW filter is shown in Figure 7.28. The outputs in Figure 7.28 are in the following order: the direct feedthrough, the main signal, triple transit, and fifth transit. Sometimes other spurious responses follow the

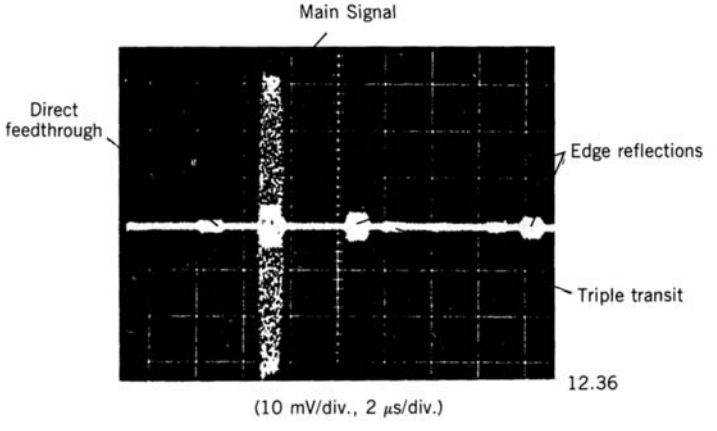


Figure 7.28. Time domain response of a SAW filter.

main signal. The direct feedthrough is caused by the capacitive coupling between the input and output transducers. The triple and fifth transits are caused by mismatch at the transducers. The signal is reflected back and forth from the transducers and comes out at a later time. The spurious responses are caused by edge reflections from the substrate. The outputs following the main signal are undesirable because they slow down the receiver processing time. The receiver must wait for all the outputs following the main signal to fade below a certain level before it can measure another signal. It should be noted that the results shown in Figure 7.28 are obtained from a filter with extremely poor time response to illustrate the various outputs. In general, the time response is much better. Filters with spurious responses of less than 50 dB in time domain are achievable.

If the frequency of the input signal is measured through the SAW filters and other parameters [i.e., pulse width (PW) and pulse amplitude (PA) are measured through a different path], time delay should be added to the PA and the PW measuring paths in order to match the long delay time in the SAW filter and synchronize the output information on the input pulses.

7.10. COARSE-FREQUENCY CHANNELIZATION

Before the discussion of fine-frequency channelization, let us use an example to discuss the down-converting and coarse-frequency channelization. If the desired receiver covers the frequency range 2.0–4.0 GHz with a desired frequency resolution of 10 MHz, the obvious approach is to build 200 parallel contiguous filters, each with a bandwidth of 10 MHz. However, it is impractical to build filters to cover the entire frequency bandwidth. A more reasonable approach is to divide the input band into many parallel channels (say 10 parallel channels of 200 MHz each) and down convert them to some intermediate-frequency (IF). To avoid second harmonics, the frequency range of the channels must be less than an octave. In this example, the IF range should be above 200–400 MHz. All the IF should be the same frequency. In this arrangement, one can design, fabricate, and evaluate 1 channel with twenty 10-MHz filters. If the result is satisfactory, one can duplicate the other 9 channels. It is not necessary to design all 10 channels.

In designing coarse channels and down converting, one should consider the possibility of sharing local oscillators (LOs). In the above example, if the down-converted frequency range is 400–600 MHz, the 2000–2200-MHz and 3000–3200-MHz bands can share one LO of 2600 MHz. The 2000–2200-MHz frequency range is converted inversely, that is, 2000 MHz is converted to 600 MHz and 2200 MHz is converted to 400 MHz. The 3000–3200-MHz band is down converted without inversion. Figure 7.29 shows such an arrangement. Following each channel, there are twenty 10-MHz filters to further process the input signal.

The bandwidth of the filter not only determines the frequency resolution of a channelized receiver but also determines the shortest PW the receiver can

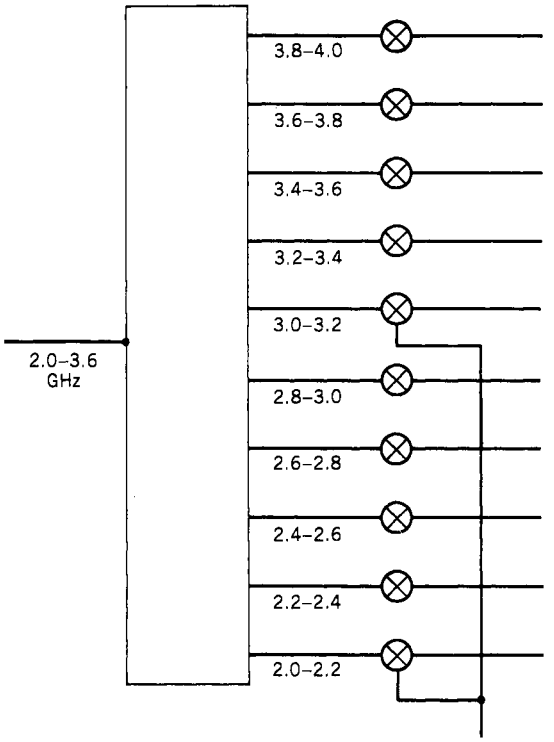


Figure 7.29. Example of coarse-frequency channelization.

measure. The rising time of a filter impulse response is inversely proportional to the bandwidth of the filter. The transient effect introduced on a pulse passing a filter causes some difficulty in measuring the signal frequency. The general rule to determine the filter bandwidth is that the bandwidth is approximately equal to the inverse of the minimum PW. For example, if the minimum PW the receiver is required to measure is 1 μsec, the filter bandwidth should be greater than 1 MHz.

If the above receiver is used to monitor a frequency band of 0.5–18 GHz, the input from 0.5–18 GHz will be divided into 2-GHz bands from 0.5–2, 2–4, 4–6, . . . , 16–18 GHz, as shown in Figure 7.30. With the exception of the 2–4-GHz band, all other bands will be converted to 2–4-GHz frequency range to match the input frequency of the receiver. The bands adjacent to the 2–4-GHz band are usually not directly converted to the 2–4-GHz range but are converted up to some frequency and then down converted to 2–4 GHz in order to avoid the severe spurious products produced by the mixers, as shown in Figure 7.30*a*.

The other way to convert an adjacent band is to divide 0.5–2 GHz into 0.5–1 and 1–2 GHz. The 0.5–1 GHz can be converted to 2.5–3 GHz, and the

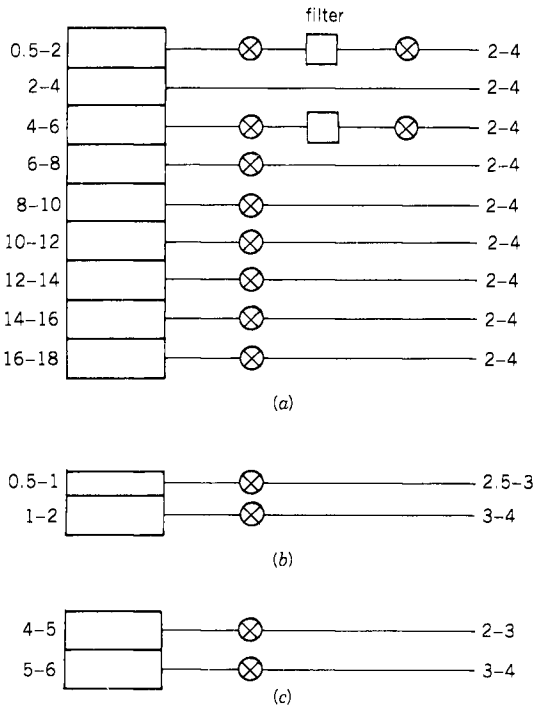


Figure 7.30. RF coarse-frequency channelization: (*a*) general channelization and down conversion; (*b*, *c*) down conversion through subchannelization.

1–2 GHz can be converted to 3–4 GHz, as shown in Figure 7.30*b*. The 4–6-GHz band can be processed in a similar manner, as shown in Figure 7.30*c*. In all the approaches, two mixers and local oscillators are required to convert one adjacent channel.

7.11. FREQUENCY MEASUREMENT CONSIDERATIONS IN CHANNELIZED RECEIVERS

It is often considered that once the channelization of a channelized receiver is built, the frequency of the input signal can be determined easily. By measuring the outputs of the fine-frequency filters, the signal frequency can be determined. However, this is not true. The research and development in channelized receivers are concentrated on how to determine the signal frequency after the channelization. From Figure 7.4, one can see that one signal will produce outputs in many filter outputs. One often argues that the filter has the strongest output representing the signal frequency. However, one must consider that the main advantage of a channelized receiver is to process simultaneous signals. If the filter with the strongest output is identified as the signal frequency, one might miss a

weak signal in a different slot filter. The frequency measurement scheme in a channelized receiver should take all these aspects into consideration.

To save hardware, amplifiers should be added in front of the filter bank instead of after the filters, as shown in Figure 7.31. One amplifier in Figure 7.31*a* can replace 20 amplifiers in Figure 7.31*b*. However, the amplifier in front of the filter bank must operate in the linear region; otherwise the receiver will produce spurious outputs. The amplifiers after the filters do not have this limitation because spurious frequencies generated in the slot filter usually will not affect the frequency measurement. Therefore, the amplifiers used in front of the filter bank usually have high power output, and the amplifiers used in the slot have low power output. As a general design rule, amplifiers should be added in front of the filter bank whenever possible. In the slot, amplifiers should be added to amplify the weakest anticipated signal to a certain output level—the threshold level.

At the output of each slot filter, there is a video detector followed by a video amplifier. The frequency of the input signal is determined by the video outputs of the slot filters. The frequency measurement scheme in a channelized receiver

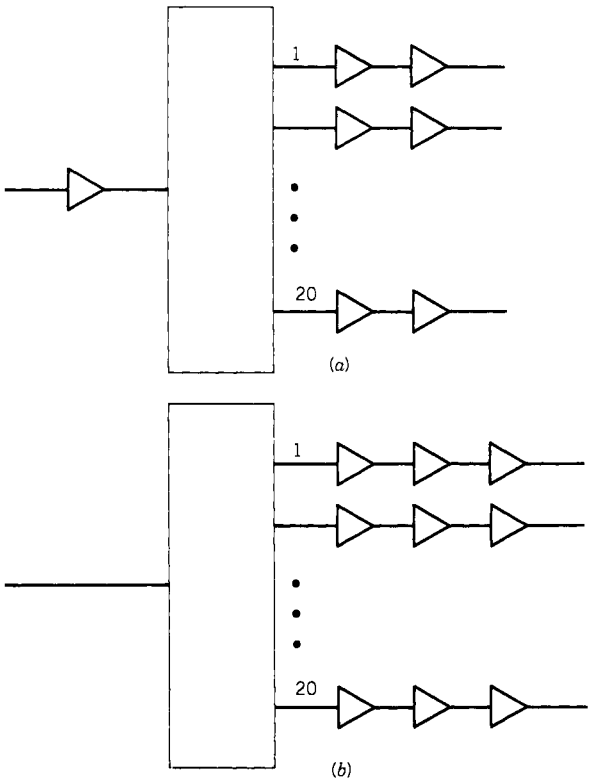


Figure 7.31. Comparison of placing amplifiers (a) in front of filter bank and (b) after filters.

can be ascribed to two approaches. The first approach can be considered a frequency domain measurement. In this approach, the outputs from adjacent filters are compared. The second approach can be considered a time domain measurement. In this approach, the output of an individual filter is used to determine whether the signal is inside or outside the center of the filter. There are several schemes used in the time domain approach. Each frequency-determining scheme will be discussed in the following sections.

Of course, the frequency of the input signal can be measured at the output of the slot filters by a frequency-sensitive circuit. The discussion of such measurements will be included in Chapter 10.

7.12. FREQUENCY DETERMINATION THROUGH FREQUENCY DOMAIN COMPARISON

An input signal will spread in the frequency domain and generate outputs from many adjacent filters as shown in Figure 7.4. The frequency-encoding scheme discussed in this section is through amplitude comparison among different filter outputs, and that is why it is referred to as a frequency domain comparison scheme. From Figure 7.4, it is obvious that the input signal is in the filter with the highest output. If the outputs of adjacent filters are compared with one another, the frequency of the input signal can be determined. In this approach, the signal amplitude must be maintained at the output of the slot frequency. The signal cannot be amplified to a constant level because the amplitude must be retained in order to perform amplitude comparison among different slots. A simple frequency domain comparison scheme is shown in Figure 7.32. There are eight filter slots and seven comparators. Let us assume that if the upper input to the comparator is stronger than the lower one, the output of the comparator is 1; otherwise it is 0. The output from a comparator passes an inverter; then it is ANDed with the output of the next lower comparator, and there are only six AND gate outputs (b–g in Figure 7.32). If the input signal is in slot C, the outputs from comparators 1 and 2 will be 0, and the outputs from comparators 3, 4, 5, 6, and 7 are 1. One can easily trace the outputs from the AND gates and find that only AND gate c will produce a 1 and the rest of the AND gates will produce 0's. Therefore, the frequency of the input is determined.

If another signal is in slot F, the outputs of the comparators depend on the signal amplitudes in slots C and F. Let us assume that the outputs of the comparators are listed in order from the top as 0, 0, 1, 1, 0, 1, and 1. The outputs from the AND gates c and f are 1, and the outputs from the rest of the AND gates are 0. Therefore, this arrangement is also applicable to simultaneous-signal conditions.

In general, the input signal must be above a certain threshold to be detected, and a threshold circuit should be added in each frequency filter output. The output of the threshold comparator can be ANDed with the output of AND gates b, c, . . . , g in Figure 7.32. The output of AND gates will have a 1 output

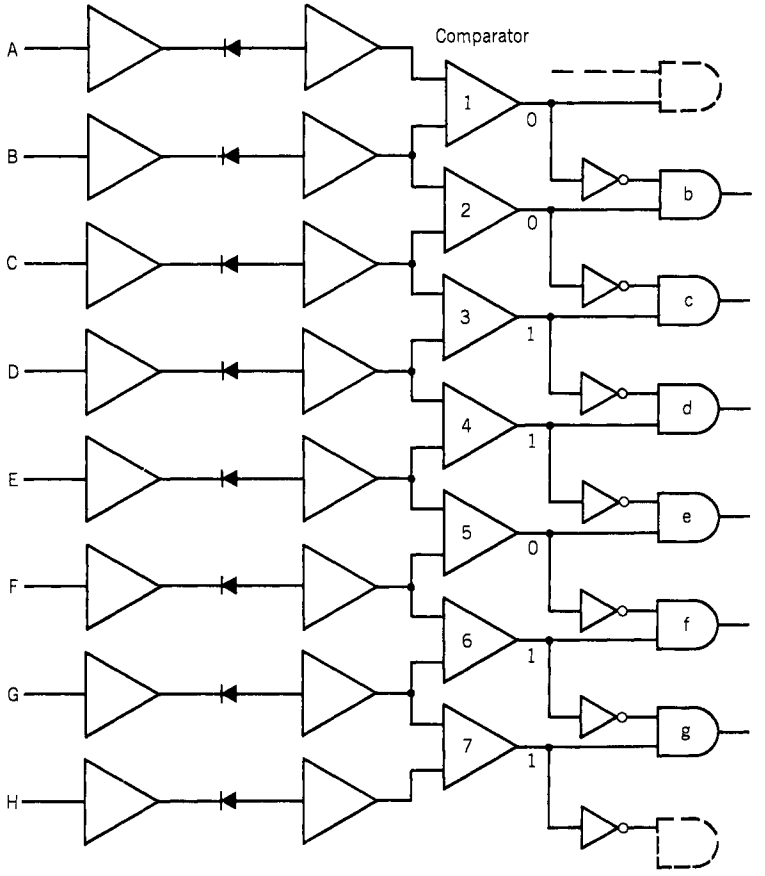


Figure 7.32. Simple amplitude comparison scheme.

only when the input signal is in the slot and above threshold. This threshold circuit can avoid an output caused by noise triggering the comparators.

Although the above design seems reasonable, a critical problem exists—when to sample the outputs from the comparators. In general, there are two approaches. The first approach is to sample the outputs of the comparators at a fixed delay time after threshold crossing. However, in general, the outputs will not cross the thresholds at the same time. The usual approach is that whenever one threshold is crossed, the output from that slot will compare with the outputs of its neighboring slots.

The other approach is to sample the outputs at a constant rate. If the input signal crosses threshold, the sampled outputs will be used to make comparisons. This approach also has the possibility of making wrong decisions. This problem can be explained as follows. Let us again assume that the input signal is in slot C. The output from slot C may not be higher than that of B and D during the entire

period of the pulse. It is possible that at some instant of time the output from B or D is higher than that of C. One of the causes of this phenomenon may be slightly different delay time in these slots. This problem will cause the frequency encoder to give wrong frequency information. One possible solution to this problem is to compare the filter outputs integrated over a period of time instead of comparing the output at one instant of time. To implement this idea, few consecutive output readings in time can be averaged, and then the amplitude is compared against its neighboring slots. This modification will complicate the design.

In the frequency comparison scheme, the characteristics of the detectors and video amplifiers should be closely matched with one another. Any discrepancy among the detectors may cause erroneous frequency readings. Since each detector is operating at a different frequency range, the amplitude matching among the slots is often hard to achieve.

7.13. FREQUENCY DETERMINATION THROUGH TIME DOMAIN COMPARISON

In this approach, the frequency of the input signal is determined solely from one filter slot and does not involve the outputs from adjacent slots. The term time domain comparison is used because the frequency is often determined by the filter output as a function of time. A unique problem will be created by this scheme, that is, the width of each encoded frequency slot. The ideal case is that each slot has the same width and shares a common boundary with adjacent slots, as shown in Figure 7.33. If an input signal is close to a boundary, it will either be in one slot or in its neighboring one. However, in a practical design, the slot width is signal dependent. The slot width will change slightly with input amplitude and PW. Therefore, it is impossible to build slots sharing common boundaries.

In an actual channelized receiver, if one tries to design filter slots sharing common boundaries, it will result in one of the two conditions shown in Figure 7.34*a* or *b*. In 7.34*a*, there is a gap between two slots. An input signal that falls in the gap will not be detected. As a result, the receiver will miss a signal, which

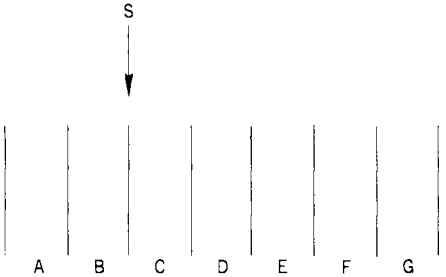


Figure 7.33. Contiguous filter slot with common boundaries.

is a very undesirable situation. In Figure 7.34*b*, there is a slight overlap between the two slots. If an input signal falls in this region, it will be detected as two signals, which is undesirable but may be tolerable.

One way to solve the above problem is through overlapping slots. Figure 7.35 shows such an arrangement. There are a total of nine slots, A–I. Each filter overlaps its adjacent ones by one-third of its width. One input signal can be detected in one or two slots. The bottom portion of Figure 7.35 shows the corresponding frequency bins, which are numbered 1–17. If slot B is the only one that has an output, the signal frequency is encoded as bin 3. If slots B and C both have outputs, the signal frequency will be encoded as bin 4, which is under B and C. There are nine slots and 17 frequency bins. Therefore, this approach improves the frequency resolution without adding hardware. The

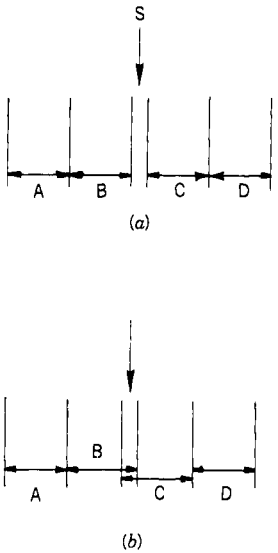


Figure 7.34. Filter slot boundary conditions. (*a*) with gap (*b*) with overlap

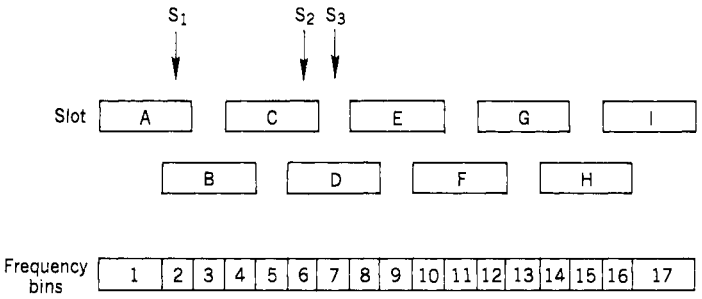


Figure 7.35. Slot with one-third overlapping on adjacent boundaries.

two end slots can be properly narrowed to provide the correct frequency resolution. A common way to do this is to install an additional filter at each end of the filter bank and use these filters to provide the correct frequency resolution for the two end slot filters.

The limitation to read simultaneous signals by this encoding scheme is discussed as follows. Suppose one signal is at S_1 and another signal is at S_2, as shown in Figure 7.35. Signal S_1 will trigger slots A and B and S_2 will trigger C and D. Four signals will be read as in bins 2, 3, 4, and 5 that are obviously wrong. If the second signal is at S_3, then the correct frequencies of bins 2 and 7 will be reported. Therefore, this encoding approach can correctly encode simultaneous signals separated by five frequency bins or more.

In the time domain comparison scheme, usually two filters are used in each slot. The first filter is used to separate signals as discussed in Section 7.2. The second filter is used to determine the signal frequency. Following the first filter, there are amplifiers. If the PA is measured at the slot output, IF log amplifiers will be used in the slot. If the PA is not measured at the slot output, the input signal is often amplified to a constant level either by linear amplifiers or by limiting amplifiers because it is easier to process signals with constant amplitudes.

In the following sections, some of the approaches used to encode the input signal frequencies will be discussed.

7.14. DUAL-DETECTION SCHEME TO DETERMINE FREQUENCY (70)

In this approach, the signal after the first filter and the amplifier chain is divided into two paths. There is one filter in each path. The two filters have the same center frequency but a different shape factor as shown in Figure 7.36. Let us assume that filter A_1 has a narrower bandwidth compared to filter A_2 and the

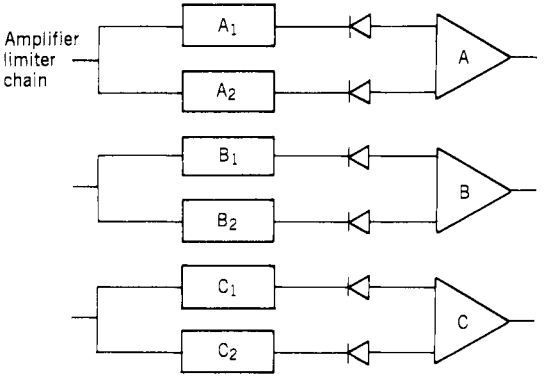


Figure 7.36. Dual-detection scheme.

two filters have the same number of poles. An attenuator of proper value is put in series with filter A_2 to reduce the output amplitude. At the outputs of the filters, there are two detectors. The characteristics of these two detectors should be closely matched. Their outputs are compared through a comparator.

The operating principle of the dual-detection scheme is illustrated in Figure 7.37. The frequency responses of the two filters are shown in this figure. The output from filter A_2 is attenuated by the attenuator in its path. The two cross points of the two responses determine the bandwidth of the slot. When the input is in the frequency range f_1-f_2, the output from filter A_1 is larger than that of filter A_2, and the output from the comparator is positive. When the input signal is outside the frequency range f_1-f_2, the output from A_1 is smaller than that from A_2, and the output from the comparator is negative. Therefore, by monitoring the output of the comparator, it can be determined whether there is a signal in the slot.

The problem is not as simple as described in the frequency domain. The transient phenomenon in the two filters must be considered. Since the two filters have different bandwidths, the transient responses are also different. The wider the filter bandwidth, the shorter the transient time. In other words, the rising time of a signal passing filter A_2 is shorter than that of filter A_1. The outputs from the two filters are shown in Figure 7.38. In Figure 7.38*a*, the input signal is insider slot A_1, whereas in Figure 7.38*b*, the input signal is outside slot A_1. These outputs are used only to demonstrate the effect of the transient. In the bottom of the two figures, the corresponding comparator outputs are shown.

From the comparator outputs, it is relatively easy to detect a signal in the center portion of the slot. But when the signal is outside of the slot, sometimes the comparator will generate positive outputs at the leading and trailing edges of the pulse, which are usually referred to as rabbit ears. The detection circuit following the comparator must be able to recognize an output generated by a signal in the slot and ignore the ears. If the circuit cannot ignore the rabbit ears, it will report false frequency from that slot. The simplest approach is to make

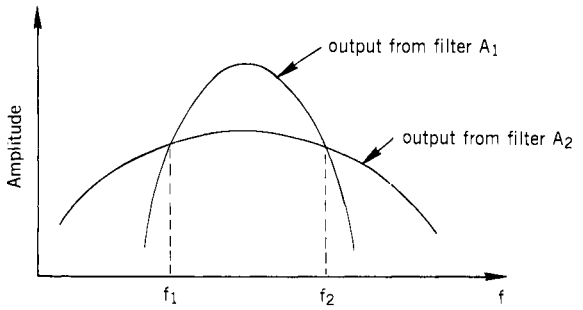


Figure 7.37. Output from dual-detection filters in frequency domain.

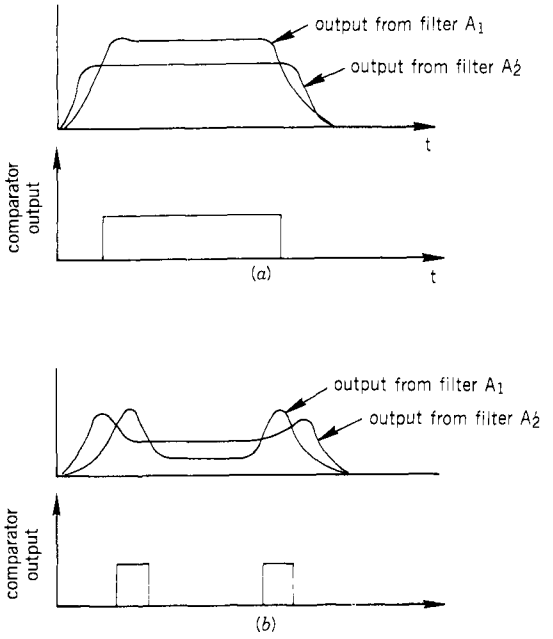


Figure 7.38. Output from slot filters and comparator. (*a*) signal in slot (*b*) signal out of slot.

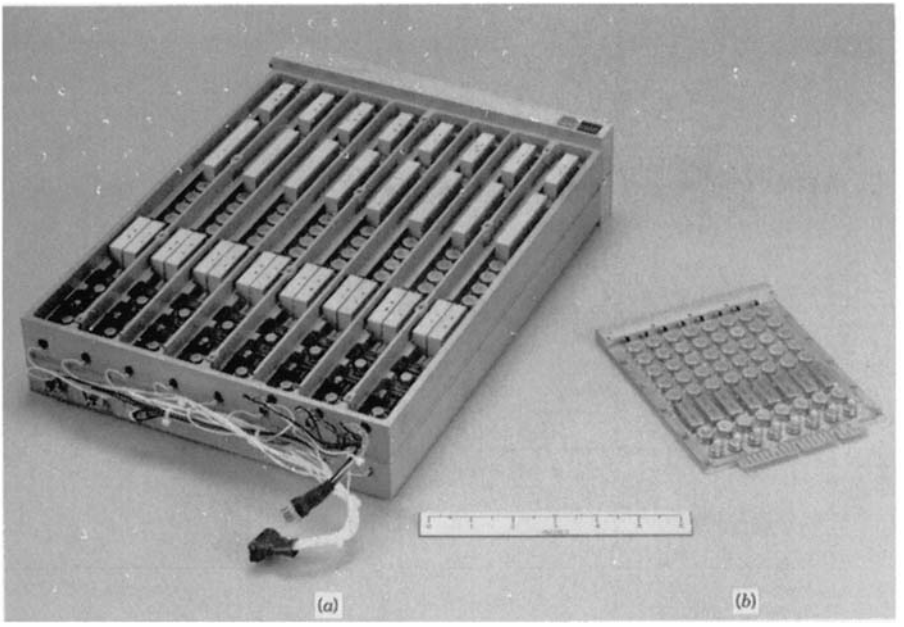


Figure 7.39. Dual detection assemblies. (*a*) with conventional filters (*b*) with SAW filters. (Fabricated by Watkins Johnson Co. Courtesy of Avionics Laboratory, AFWAL.)

the decision between the two ears. If the input signal PW decreases, the time between the two ears also decreases. When the two ears are very close together, there is not enough time between them to properly detect the signal. This effect determines the minimum PW the receiver can process.

A channelized receiver using the dual-detection scheme is shown in Figure 7.39. In Figure 7.39*a*, there are a total of 16 slots: 8 in the top container, 8 in the bottom container. The figure contains two 1:8 power dividers. At the output of the power divider, there are two filters connected in cascade to form the composite filter discussed in Section 7.2. The power divider and the two filters form the filter bank. Following the filters are the amplifier limiter chain and the two

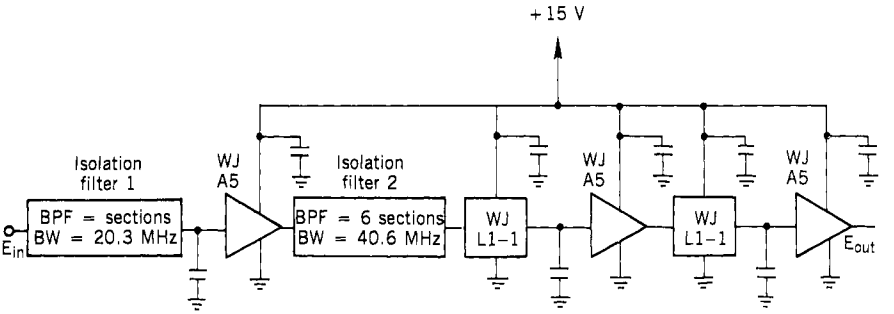


Figure 7.40. Conventional filter slot and amplifier limiter chain.

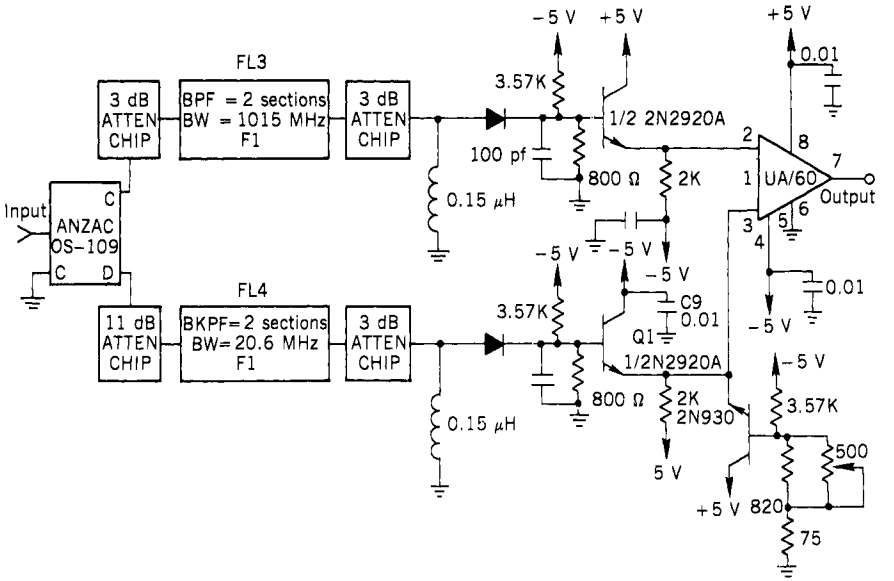


Figure 7.41. Dual detection with conventional filters.

parallel filters, which are used for the dual-detection scheme. At each output of the two parallel filters there is a detector followed by a comparator. Figure 7.39*b* shows a total of eight slots with SAW filters as the dual-detection filters. In this figure, the filter bank is not included. It only contains the amplifier limiter chain, the dual-detection filters, detectors, and the comparator. The

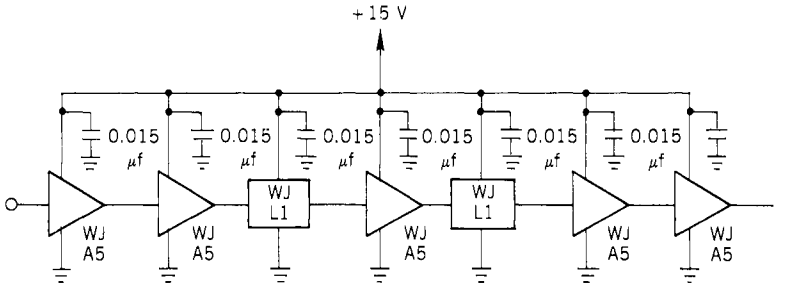


Figure 7.42. SAW filter slot and amplifier limiter chain.

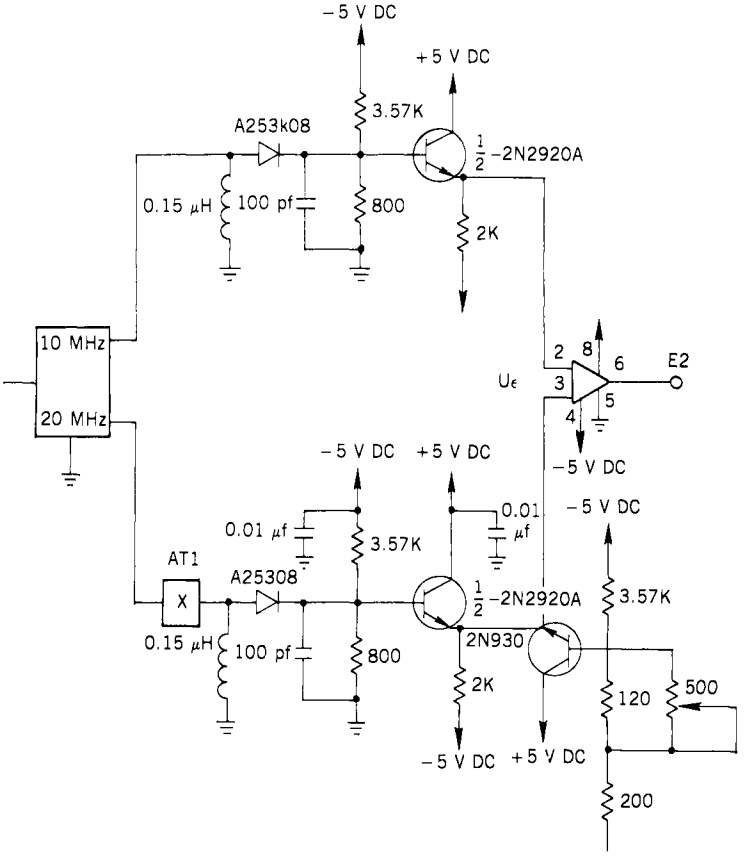


Figure 7.43. Dual detection with SAW filters.

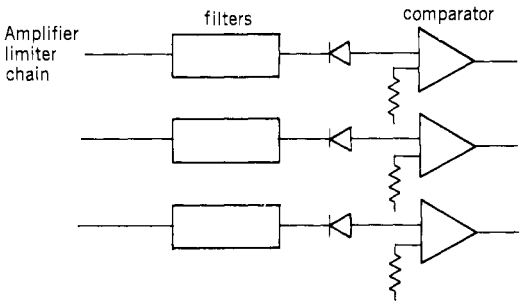


Figure 7.44. Simplified dual-detection scheme

size reduction due to the SAW filters is clearly shown. The detailed circuits are shown in Figures 7.40–7.43. Figure 7.40 shows one slot of the filter bank, and Figure 7.41 shows the dual-detection circuits and the comparator. Figures 7.42 and 7.43 show the equivalent portion of the SAW filter version. It should be noted that the slot assembly with SAW filters is smaller in size, but there are two more amplifiers in each slot to compensate for the high insertion loss in the SAW filters.

The frequency slots discussed in the above sections are designed to have one-third bandwidth overlap on adjacent slots, as discussed in Section 7.13 and illustrated in Figure 7.35. With this design, one input signal can only trigger two adjacent slots as expected.

The dual detection can be simplified by eliminating the wide-band filter as shown in Figure 7.44. The output from the filter is detected and compared with a constant voltage. The output from the comparator is similar to that of Figure 7.38. The deficiency of this design is that if the output level from the amplifier limiter chain changes, the output from the comparator will change and cause frequency inaccuracy. In the dual-detection scheme, when the output level of the amplifier chain varies, both detector outputs will change, and the amplitude variation will be somewhat compensated.

7.15. ENERGY DETECTION SCHEME TO DETERMINE FREQUENCY (71)

The basic operating principle of an energy detection scheme is similar to the simplified dual-detection scheme discussed in Section 7.14. The major difference between these two schemes is in the circuits that follow the diode detector. This approach is intended to eliminate the rabbit ears generated in the dual-detection scheme. The circuit is shown in Figure 7.45. The output from the diode detector is compared with a constant voltage. The output of the comparator is amplified and stored in a capacitor C. The voltage across the capacitor is again compared with another fixed threshold. The output of the second comparator is used to determine whether the input signal is in the slot.

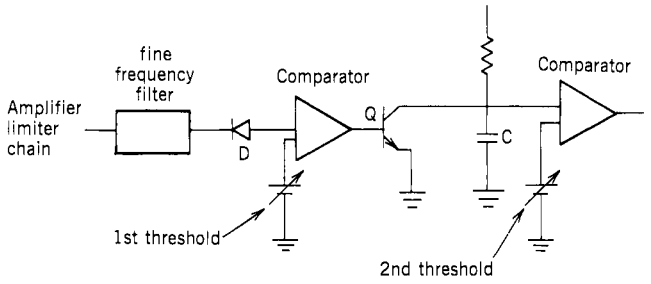


Figure 7.45. Energy detection scheme for determining frequency.

If the diode detector output is above the first threshold, transistor Q will be biased off, and capacitor C will be charged. If the output from the detector is below the threshold, transistor Q will be switched on and the charge on capacitor C will be drained through the transistor. If the input signal is in the slot, the output from the diode detector and the voltage across capacitor C are shown in Figure 7.46*a*, and the output from the second comparator will be positive. If the signal is outside the slot but close to it, the output from the diode detector and the voltage across capacitor C are shown in Figure 7.46*b*. In this figure, although the first threshold is crossed twice, the second threshold is never

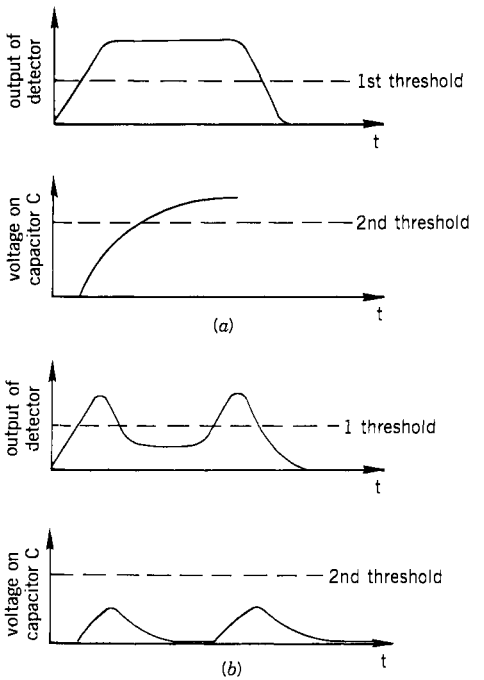


Figure 7.46. Outputs from energy detection scheme. (*a*) signal in slot (*b*) signal out of slot.

crossed, and the output of the second comparator is always zero. The advantage of this design is to eliminate the timing circuit following the comparator used in the dual-detection scheme. However, if the signal amplitude at the output of the amplifier limiter chain varies, the frequency accuracy will suffer.

As discussed in the above section, when the input signal PW decreases, the leading and trailing edges will close and the capacitor will be charged above the second threshold. Under this condition, an input signal outside the slot will be measured as a signal inside the slot. In other words, the detection circuit makes wrong decisions on short pulsed signals.

7.16. PEAK VALLEY COMPARISON SCHEME TO DETERMINE FREQUENCY (72)

In this approach, the output of the diode detector is divided into two parallel paths. A delay line is placed in one of the paths, as shown in Figure 7.47. The amplitudes of the two signals are properly adjusted and fed to the inputs of a comparator. The output of the comparator is used to determine whether an input signal is in the slot filter. The purpose of this design is to detect the existence of the ears caused by the transient of the filter. If rabbit ears are detected, the input signal will be declared outside the slot. If rabbit ears are not detected, the input signal will be considered inside the slot.

The inputs and outputs of the comparator are shown in Figure 7.48. If the input signal is inside the slot, as shown in Figure 7.48*a*, the transient effect caused by the filter is not very prominent. Under this condition, the valley of the undelayed signal is higher than the peak of the delayed signal and the comparator output is positive. If the input signal is outside the slot, the transient effect caused by the filter is very prominent. Under this condition, the peak of the delayed signal is higher than the valley of the undelayed signal, and the corresponding output from the comparator is shown in Figure 7.48*b*. To determine whether the input signal is inside the slot, the output of the comparator must be sampled at the proper time. Therefore, a timing circuit following the comparator is needed to sample its output. When the PW decreases, this encoding will also generate false information because the interval available for proper sampling will also decrease.

Since the two inputs to the comparator are both from the detector output, the amplitude variation at the output of the amplifier limiter chain should have

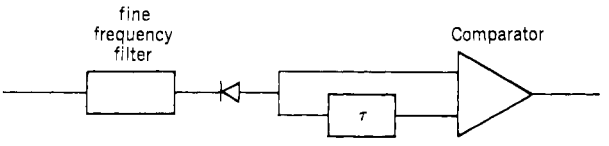


Figure 7.47. Peak valley comparison scheme to determine frequency.

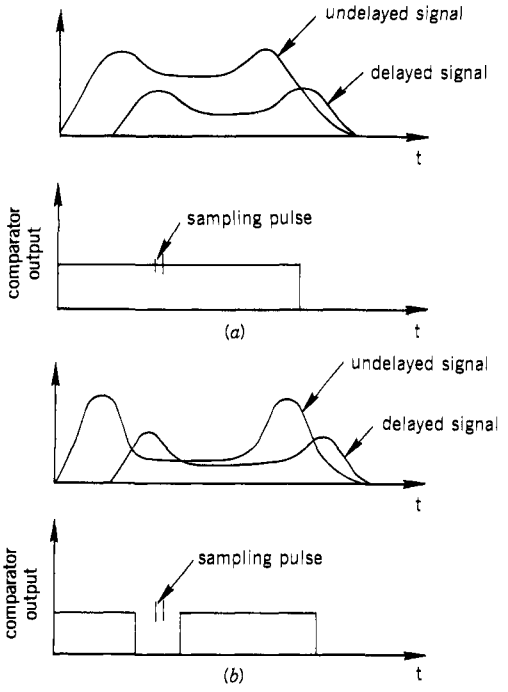


Figure 7.48. Inputs and outputs of comparison in a peak valley comparison scheme. (*a*) signal in slot (*b*) signal out of slot.

limited effect. However, the variation of the leading edge of the input signal will cause different transient effects. Thus the shape of the leading edge may cause frequency measurement inaccuracy.

7.17. FREQUENCY DETERMINATION WITH SLOTS CONTAINING AMPLITUDE INFORMATION

Most of the frequency determination schemes discussed in the above sections do not measure the PA information at the slot level. Therefore, the input signal is usually amplified to a constant level before it reaches the diode detector. This constant power level at the inputs of the detectors decreases the possibility of generating erroneous frequency information. As a result, the designs of the frequency determination circuits is simplified. If IF log amplifiers are used in every slot, the output from the log amplifier is signal amplitude dependent. Theoretically, the basic approaches discussed in the above sections can be applied to determine signal frequency with amplitude variation. However, the possibility of generating erroneous frequency could be higher and the circuits more complicated than the circuits measuring only the signal frequency.

The frequency detection schemes discussed will occasionally produce some spurious outputs even with constant amplitude at the inputs of the video detectors. Therefore, it is reasonable to assume that when the input signals applied to the video detector circuits have different amplitude, the problem of generating spurious outputs and erroneous frequency information will be more prominent.

One possible approach to measure signal frequency containing amplitude information is to use both frequency domain and time domain frequency-determining schemes. The input signal should be qualified in both domains before its frequency is declared. Although this combination of frequency domain and time domain approaches may be helpful in determining the input signal frequency, the circuit design would be more complicated.

7.18. MODELING OF CHANNELIZED RECEIVERS (73–77)

To reduce risk and save time in designing a channelized receiver, one of the approaches is to use a computer to model the design before fabrication. Of course, computer modeling should also be applied to other types of receivers. However, at present, the modeling of channelized receivers is more developed than the modelings of other types of receivers.

The major function in modeling a microwave receiver design is to determine the signal shape at the output of each component and predict the performance of the receiver. The components of concern used in the modeling program of a channelized receiver are filters, limiters, and log amplifiers. The effect of some linear devices (i.e., amplifiers and attenuators) can be modeled by simply multiplying a constant by the signal.

There are many kinds of filters. The ones most commonly used in microwave receivers are the Butterworth, Chebyshev, elliptical, and SAW filters. The poles and zeros of the first three types of filters can be located theoretically. There are two ways to calculate the output of a filter. One way is to operate in time domain and is referred to as convolution integration:

$$Y(t) = \int_0^t R(\tau)H(t - \tau)\, d\tau \tag{7.16}$$

where $Y(t)$ is the output of the filter, $R(\tau)$ is the input signal, and $H(t)$ is the transfer function of the filter.

The other way to calculate the filter output is through multiplication of the input signal and the filter transfer function in the frequency domain. The result is then converted to time domain through an inverse fast Fourier transform (FFT). This operation is more versatile than the convolution integration because it can use actual S parameters measured on a filter to predict the output signal shape, where S parameters are a popular means for characterizing n-port

networks at high frequency (ref. 77). Figure 7.49 shows the experimental and calculated results of a SAW filter, and they match very well.

The limiters and log amplifiers can be modeled through mathematical approximation in the time domain. It simply multiplies the input signal to generate the output. If the frequency domain response is of interest, an FFT operation will generate the frequency data.

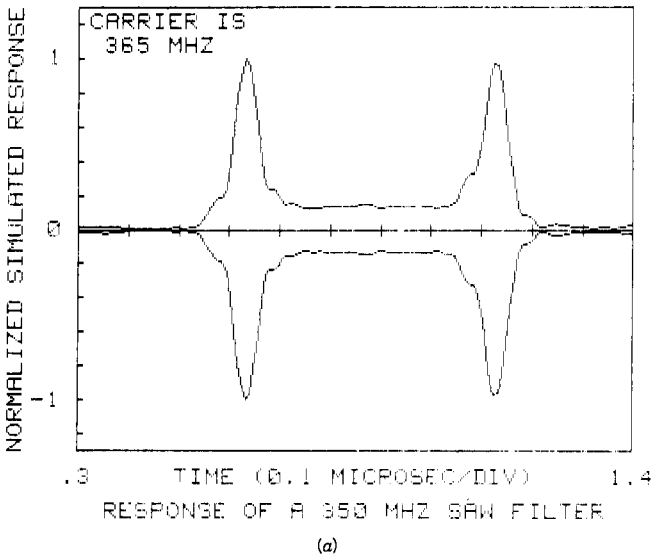


Figure 7.49. SAW filter output (*a*) modeled (*b*) experimental.

7.19. FREQUENCY-SELECTIVE LIMITERS (78–90)

In building wide-band receivers, it is very desirable to have a frequency-selective limiter. When simultaneous signals of different frequencies enter the limiter, they are limited independently. The performance of an ideal frequency-selective limiter is shown in Figure 7.50. Each signal is limited individually as though there are no other signals. Such a device is ideal for wide-band receiver applications, since there is no intermodulation generated and all the output signals are at a constant level. However, theoretically, this is almost impossible because a limiter is a nonlinear device and intermodulations will be generated from a nonlinear device. Strictly speaking, the entire channelized receiver can be considered a frequency-selective limiter. An yttrium iron garnet (YIG) limiter exhibits some of the characteristics of a frequency selective limiter.

A YIG limiter uses a YIG sphere (or spheres) to couple between the input and output ports. Its structure is like that of a YIG filter (see Section 5.5). In a YIG limiter, the subsidiary resonance mode of operation is chosen because it provides a reasonably low limiting threshold over broad instantaneous bandwidth. The theory of subsidiary resonance limiting was discussed in reference 90.

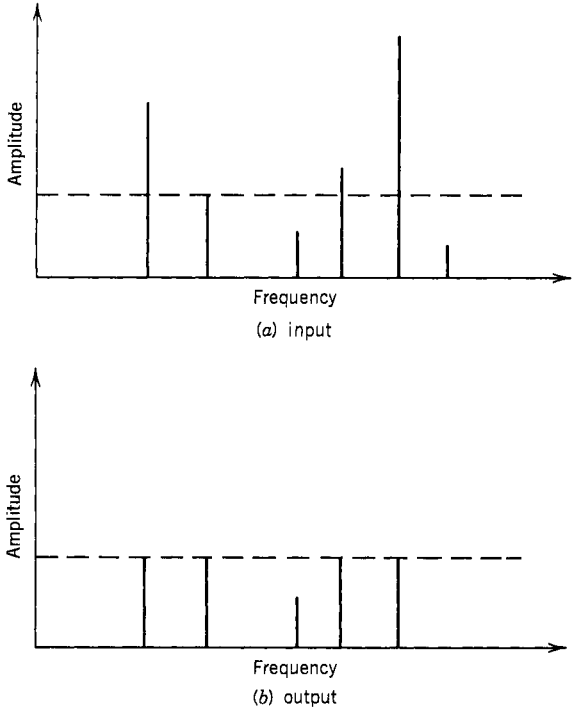


Figure 7.50. Inputs and outputs of an ideal frequency-selective limiter.

The external static magnetic field is adjusted so that spin waves exist at one-half the signal frequency f. A symbolic representative of a YIG limiter is shown in Figure 7.51. There are two resonators of frequency f and $\frac{1}{2}f$ that are coupled through a nonlinear element. When the input signal at frequency f is below a certain threshold, it will be coupled to the output as if the subharmonic resonator is nonexistent. When the input power is above the threshold, energy will be coupled to the subharmonic resonator $f/2$. Under this condition, part of the RF energy is coupled through the subharmonic resonator to the crystal lattice vibration and dissipated as heat. Part is reflected to the input source because of the impedance mismatch caused by the added equivalent conductance in the f resonant circuit by the $f/2$ resonant circuit. These combined effects result in the limiting characteristic.

In a wide-band YIG limiter, many subharmonic resonators with frequencies $f_1/2, f_2/2, \ldots, f_n/2$ can be imaged to couple in the main resonator. The limiting action is frequency selective because input signals separated in frequency more than a few spin wave linewidths apart do not couple to the same subharmonic resonator. Therefore, each input signal is limited independently. However, when the input signals with frequencies close together (typically 20 MHz) are applied to the YIG limiter simultaneously, then capture effect where the strong signal suppresses the weak one will still occur. When a pulse passes through a YIG limiter, the leading edge of the pulse will not be properly limited. This is referred to as the leakage effect. A strong pulse will have a short but prominent leakage effect. This effect degrades the performance of the limiter on short pulses.

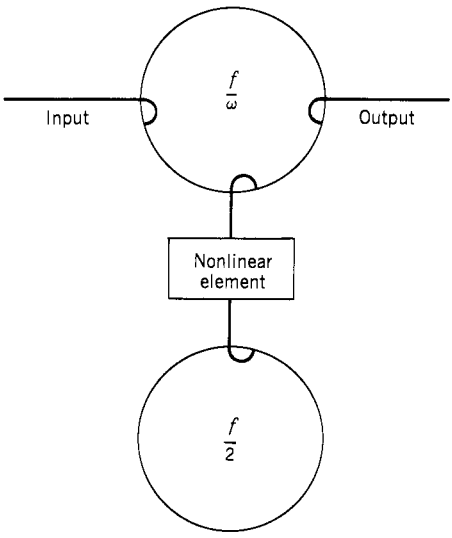


Figure 7.51. Symbolic representation of a YIG frequency-selective filter.

7.20. SUMMARY

A channelized receiver can be considered to be many fix tuned superheterodyne receivers operating in parallel. It has all the good characteristics of a superhet receiver: high sensitivity, wide dynamic range, and fine-frequency resolution. The bandwidth of the receiver is proportional to the amount of hardware used in the receiver. The more channels built, the more bandwidth the receiver can cover. But the large number of parallel channels also make the receiver bulky and expensive.

The most critical portion of a channelized receiver is to determine which slot contains the input signal. Although several approaches have been presented in this chapter, a thoroughly theoretical study may improve the frequency-determining scheme. Using SAW filters in a channelized receiver may reduce the size and possibly the cost, but the SAW filters also seem to degrade the performance of the receiver. Further research in filter technology is required to improve receiver performance. Finally, reducing the receiver size is another research topic to be investigated.

REFERENCES

1. T. Harper, High probability of intercept receivers, Watkins Johnson Co. Tech Notes, **2**(4) (July/August 1975).
2. T. Harper, New trends in EW receivers, *Countermeasures*, 34–38 (December/January 1976).
3. The channelized receiving systems, Staff Report, *Microwave System News*, **6**, 63 (December 1975–January 1976).
4. P. Hennessy and J. D. Quick, The channelized receiver comes of age, *Microwave System News*, **9**, 36 (July 1979).
5. T. Harper, Hybridization of competitive receivers, Watkins Johnson Co. Tech Notes, **7**(1) (January/February 1980).
6. C. B. Hoffmann and A. R. Baron, Wideband ESM receiving systems, Part I, *Microwave J.*, **24**, 24 (September 1980); Part II, 57 (February 1981).
7. G. L. Matthaei, L. Young, and E. T. M. Jones, *Microwave Filters, Impedance-Matching Networks and Coupled Structures*, McGraw-Hill, New York, 1964.
8. E. G. Cristal and G. L. Matthaei, A technique for the design of multiplexers having contiguous channels, *IEEE Trans. Microwave Theory Techniques*, **MTT-12**, 88–93 (1964).
9. R. J. Wenzel, Application of exact synthesis method to multichannel filter design, *IEEE Trans. Microwave Theory Techniques*, **MTT-13**, 5–15 (1965).
10. C. Buntschuh, Octave bandwidth, high directivity microstrip coupler, RACE-TR-73-396, Contract F30602-72-C-0282, AD 777320, January 1974.
11. J. I. Haine and J. D. Rhodes, Direct design formulas for asymmetric band-pass channel diplexers, *IEEE Trans. Microwave Theory Techniques*, **MTT-25**, 807–812 (1977).
12. J. D. Rhodes and R. Levy, A generalized multiplexer theory, *IEEE Trans. Microwave Theory Techniques*, **MTT-27**, 99–111 (1979).
13. J. D. Rhodes and R. Levy, Design of general manifold multiplexers, *IEEE Trans. Microwave Theory Techniques*, **MTT-27**, 111–123 (1979).
14. J. D. Rhodes and S. A. Alseyab, A design procedure for band-pass channel multiplexer connected at a common junction, *IEEE Trans. Microwave Theory Techniques*, **MTT-28**, 246–253 (1980).

15. W. Rotman and R. F. Turner, Wide-angle microwave lens for line source application, *IEEE Trans. Antenna Propagation*, **AP-11**, 623–632 (1963).
16. D. H. Archer and J. Prickett, Signal spectrum analyzer, U.S. Patent 3,735,256, May 22, 1973.
17. R. C. Hansen, Ed., *Microwave Scanning Antennas*, Vol. 1, *Aperture*, Academic Press, New York, 1964.
18. R. Brooks, Acoustic wave diffractor for array processing, *Applied Optics*, **22**, 2810–2816 (1982).
19. E. M. T. Jones and J. T. Bolljahn, Coupled-strip-transmission line filters and directional couplers, *IRE Trans. Microwave Theory Techniques*, **MTT-4**, 75–81 (1956).
20. J. K. Shimizu and E. M. T. Jones, Coupled-transmission-line directional couplers, *IRE Trans. Microwave Theory Techniques*, **MTT-6**, 403–410 (1968).
21. Anaren, publication M9001-67 by Anaren Microwave Inc., New York, 1967.
22. J. P. Shelton, J. Wolfe, and R. C. Van Wagoner, Tanden couplers and phase shifters for multi-octave bandwidth, *Microwaves*, **4** (April 1965).
23. I. J. Bahl and P. Bhartia, Characteristics of inhomogeneous broadbide-coupled strip lines, *IEEE Trans. Microwave Theory Techniques*, **MTT-28**, 529–535 (1980).
24. S. Rehnmark, High directivity CTL-couplers and a new technique for the measurement of CTL-coupler parameters, *IEEE Trans. Microwave Theory Technique*, **MTT-25**, 1116–1121 (1977).
25. B. Sheleg and B. E. Spielman, Broadband directional couplers using microstrip with dielectric overlays, *IEEE Trans. Microwave Theory Techniques*, **MTT-22**, 1216–1219 (1974).
26. P. P. Toulios and A. C. Todd, Synthesis of symmetrical TEM-mode directional couplers, *IEEE Trans. Microwave Theory Techniques*, **MTT-13**, 536–544 (1965).
27. A. Podell, A high directivity coupler technique, *IEEE* GMTT-1970, Int. Microwave Symposium, pp. 33–36, May 1970.
28. R. Levy, Zolotarev branch-guide couplers, *IEEE Trans. Microwave Theory Techniques*, **MTT-21**, 95–99 (1973).
29. J. W. Carr, Balanced-line microwave hybrids, *Microwave J.*, **16**(5), 49–52 (May 1973).
30. S. Rehnmark, Wide-band advanced line microwave hybrids, *IEEE Trans. Microwave Theory Techniques*, **MTT-25**, 825–830 (1977).
31. F. C. DeRonde, Recent developments in broadband directional couplers on microstrip, *IEEE MTT-S Int. Microwave Symposium*, pp. 215–217 (1975).
32. J. E. Dalley, A strip line directional coupler utilizing a nonhomogeneous dielectric medium, *IEEE Trans. Microwave Theory Techniques*, **MTT-17**, 706–712 (1969).
33. S. D. Shamasundara and K. C. Gupta, Sensitivity analysis of coupled microstrip directional coupler, *IEEE Trans. Microwave Theory Techniques*, **MTT-26**, 788–794 (October 1978).
34. S. B. Cohn, Microwave band-pass filters containing high-Q dielectric resonators, *IEEE Trans. Microwave Theory Techniques*, **MTT-16**, 218–227 (1968).
35. W. B. Harrison, A miniature high-Q band-pass filter employing dielectric resonators, *IEEE Trans. Microwave Theory Techniques*, **MTT-16**, 210–218 (1968).
36. T. D. Iveland, Dielectric resonator filters for application in microwave integrated circuits, *IEEE Trans. Microwave Theory Techniques*, **MTT-19**, 643–652 (1971).
37. S. J. Naumann and J. C. Sethares, Microwave dielectric resonators, Air Force Oambridge Research Laboratories, AFCRL-65-867, November 1965.
38. M. A. Gerdine, A frequency stabilized microwave band-rejection filter using high dielectric constant resonators, *IEEE Trans. Microwave Theory Techniques*, **MTT-17**, 354–359 (1969).
39. H. Y. Yee, Natural resonance frequency of microwave dielectric resonator, *IEEE Trans. Microwave Theory Techniques*, **MTT-13**, 256 (1965).
40. Y. Konishi, N. Hoshino, and Y. Utsumi, Resonant frequency of a $TE_{01\delta}$ dielectric resonator, *IEEE Trans. Microwave Theory Techniques*, **MTT-24**, 112–114 (1976).
41. T. Itoh and R. S. Rudokas, New method for computing the resonant frequencies of dielectric resonators, *IEEE Trans. Microwave Theory Techniques*, **MTT-25**, 52–55 (1977).
42. M. W. Pospiezalski, On the theory and application of the dielectric post resonator, *IEEE Trans. Microwave Theory Techniques*, **MTT-25**, 228–231 (1977).

43. J. K. Plourde and D. F. Linn, Microwave dielectric resonator filters utilizing $Ba_2Ti_9O_{20}$ ceramics, *IEEE MTT-S, Int. Microwave Symposium*, pp. 290–293, 1977.
44. M. W. Pospieszalski, Cylindrical dielectric resonators and their applications in TEM line microwave circuits, *IEEE Trans. Microwave Theory Techniques*, **MTT-27**, 233–238 (1979).
45. M. R. Stiglitz and J. C. Sethares, A hybrid ferrimagnetic dielectric microwave filter, *Proc. IEEE*, **55**, 1734–1735 (1967).
46. J. C. Sethares and M. R. Stiglitz, Visual observation of high dielectric resonator modes, *Applied Optics*, **8**, 2560–2562 (1969).
47. A. Karp, H. J. Shaw, and D. K. Winslow, Circuit properties of microwave dielectric resonators, *IEEE Trans. Microwave Theory Techniques*, **MTT-16**, 818–828 (1968).
48. M. R. Stiglitz, Frequency tuning of futile resonators, *Proc. IEEE*, **54**, 413–414 (1966).
49. D. J. Masse and R. A. Pucel, A temperature stable band-pass filter using dielectric resonators, *Proc. IEEE*, **60**, 730–731 (1972).
50. U. H. Lammers and M. R. Stiglitz, A 96-channel multiplexer with dielectric resonators, *Microwave J.*, **22**, 59 (October 1979).
51. C. L. Ren, Waveguide band-stop filter utilizing $Ba_2Ti_9O_{20}$ resonators, *IEEE, International Microwave Symposium*, **MTT-5**, 227–229 (1978).
52. K. Wakino, T. Nishikawa, H. Matsumoto, and Y. Ishikawa, Miniaturized band-pass filters using half-wave dielectric resonators with improved spurious response, *IEEE, International Microwave Symposium*, **MTT-5**, 230–232 (1978).
53. W. H. Harrison and S. B. Cohn, Development of dielectric resonator filter bank, Wavecom, No. 30810, May 27, 1983.
54. R. H. Tancrell and M. G. Holland, Acoustic surface wave filters, *Proc. IEEE*, **59**, 393–409 (1971).
55. C. S. Hartmann, D. T. Bell, and R. C. Rosenfeld, Impulse model design of acoustic surface-wave filter, *IEEE Trans. Microwave Theory Techniques*, **MTT-21**, 162–175 (1973).
56. R. H. Tancrell, Analytic design of surface wave band-pass filters, *IEEE Trans. Sonics Ultrasonics*, **SU-21**, 12–22 (1974).
57. H. Engan, Surface acoustic wave multielectrode transducers, *IEEE Trans. Sonics Ultrasonics*, **SU-22**, 395–401 (1975).
58. M. H. El-Diwany and C. K. Campbell, Modification of optimum impulse response techniques for application to SAW filter design, *IEEE Trans. Sonics Ultrasonics*, **SU-24**, 277–279 (1977).
59. J. P. Reilly, C. K. Campbell, and M. S. Suthers, The design of SAW band-pass filters exhibiting arbitrary phase and amplitude response characteristics, *IEEE Trans. Sonics Ultrasonics*, **SU-24**, 301–305 (1977).
60. R. R. Jones, J. Schellenberg, W. K. Tanski, and R. A. Moore, Transplexing SAW filters for ECM, *Microwaves*, Part 1, **14**, 43 (December 1974); Part 2, **15**, 68 (January 1975).
61. C. Huber, J. Lane, B. A. Newman, J. T. Godfrey, C. H. Grauling, and R. A. Moore, A low side lobe SAW continuous filter bank using MDC LiTaO, *Ultrasonics Symposium Proc. IEEE*, Cat. 77CH1264-ISU, pp. 568–572, 1977.
62. L. P. Solie, A surface acoustic wave multiplexer using offset multistrip couplers, *Ultrasonics Symposium Proc.*, pp. 153–156, 1974.
63. H. Van de Varrt and L. P. Solie, Surface acoustic wave multiplexing techniques, *Ultrasonics Symposium Proc. IEEE*, Cat. 75CH0994-4SU, 1975.
64. C. H. Krueger and J. B. Y. Tsui, Application of surface acoustic wave filter to channelized receivers, NAECON, 1975.
65. G. L. Matthei, D. Y. Wong, and B. P. O'Shaughnessy, Simplifications for the analysis of interdigital surface-wave devices, *IEEE Trans. Sonics Ultrasonics*, **SU-22**, 103–114 (1975).
66. G. L. Matthaei, F. Barman, and E. B. Savage, Acoustic surface wave resonator for bandpass filter applications, *IEEE MTT-S International Microwave Symposium*, 1976.
67. H. L. Ragen, Surface acoustic wave resonator filters, *IEEE MTT-S, International Microwave Symposium*, 1976.
68. G. L. Matthaei, E. B. Savage, and F. Barman, Synthesis of acoustic surface wave resonator filters using any of various coupling mechanisms, *IEEE MTT-S, International Microwave Symposium*, 1977.

69. E. McCune, SAW filters improve channelized receivers, *Microwaves RF*, **23**, 103 (January 1984).
70. R. Hollis, Double detection filter techniques test report, Watkins Johnson Co., Palo Alto, CA, 1973.
71. D. Lundy and K. Grover, Motorola Co., private communication.
72. W. Daniels and T. Higgins, Texas Instruments Inc., private communication.
73. T. Murakami and R. W. Sonnenfeldt, Transient response of detectors in symmetric and asymmetric side band systems, *RCA Review*, **16**, 581–611 (December 1955).
74. K. W. Henderson and W. H. Kautz, Transient response of conventional filter, *IRE Trans. Circuit Theory*, **CT-5**, 333–347 (1958).
75. T. H. Lecklider, Program gives filter time response electronic design, *Electronic Design*, **9**, 192 (April 1974).
76. J. B. Y. Tsui, J. E. Adair, J. E. Hawkins, and S. J. LaFleur, Transient response of filters, Air Force Avionics Laboratory Technical Report AFAL-TR-77-249, December 1977.
77. "S-parameter design," Hewlett Packard Application Note 154 Palo Alto, CA, April 1982.
78. A. I. Zverev, *Handbook of Filter Synthesis*, Chapter 7, New York, 1967.
79. S. N. Stitzer and H. Goldie, Frequency selective high power YIG limiters, *IEEE MTT-S, International Microwave Symposium*, 1976.
80. J. Helszajn, R. W. Murray, E. G. S. Davidson, and R. A. Suttrie, Waveguide subsidiary resonance ferrite limiters, *IEEE Trans. Microwave Theory Techniques*, **MTT-25**, 190–196 (1977).
81. P. R. Emtage and S. N. Stitzer, Interaction of signals in ferromagnetic limiters, *IEEE Trans. Microwave Theory Techniques*, **MTT-25**, 210–213 (1977).
82. S. N. Stitzer, H. Goldie, and P. S. Carter, X-band YIG limiters for FM/CW radar, *Microwave J.*, **20**, 35 (December 1977).
83. R. C. S. Chien, Frequency selective limiter and its application in a filter bank receiver, Technical Note 1973-1, Lincoln Laboratory, MIT, April 13, 1973.
84. A. J. Giarola, D. R. Jackson, R. W. Orth, and W. P. Robbins, A frequency selective limiter using magnetoelastic instability, *Proc. IEEE*, **5**, 593–594 (1967).
85. D. R. Jackson and R. W. Orth, A frequency selective limiter using nuclear magnetic resonance, *Proc. IEEE*, **55**, 36–45 (1967).
86. K. L. Kotzebue, Frequency selective limiting, *IRE Trans. Microwave Theory Techniques*, **MTT-10**, 516–520 (1962).
87. J. Brown, Ferromagnetic limiters, *Microwave J.*, **4**, 74 (November, 1961).
88. R. V. Garver and D. Y. Tseng, X-band limiting, *IRE Trans. Microwave Theory Techniques*, **MTT-9**, 202 (1961).
89. G. S. Uebele, Characteristics of ferrite microwave limiters, *IRE Trans. Microwave Theory Techniques*, **MTT-7**, 18–23 (1959).
90. H. Suhl, The nonlinear behavior of ferrites at high microwave signal levels, *Proc. IRE*, **44**, 1270–1284 (1956).

Chapter 8

Compressive (Microscan) Receivers

8.1. INTRODUCTION (1)

The name *compressive receiver* is given to this kind of receiver because a dispersive delay line (DDL) is used to compress the input radio-frequency (RF) signal to a narrow pulse. It is also referred to as a *microscan receiver* because a fast sweeping local oscillator (LO) is used to convert the input signals to frequency-modulated (FM) signals. The idea of using a DDL to measure frequency was patented in 1960 (ref. 1). The advances in surface acoustic wave (SAW) technology and high-speed logic circuits have revitalized the interest in developing compressive receivers.

A compressive receiver is a wide-band receiver with fine-frequency resolution. It has the capability to process simultaneous signals. Whereas the detected outputs from a channelized receiver are in parallel, the detected outputs from a compressive receiver are narrow pulses arriving in series in time domain. By measuring the positions of these compressed pulses, the frequency of the input signals can be determined. Since the detected pulses are very narrow and very close in time, high-speed logic circuits are required to process them.

The research and development efforts in compressive receivers have been concentrated on the DDL and sweeping LO. Although there are still improvements to be made in these areas, the major deficiency in the receiver is in the digitizing circuits that take the detected video outputs as inputs and generate digital words to represent the parameters of the input RF signals. Although experimental compressive receivers demonstrated that the digitizing circuits can be built, there is still room for improvement. An experimental compressive receiver is shown in Figure 8.1. The RF section is at the bottom of the picture,

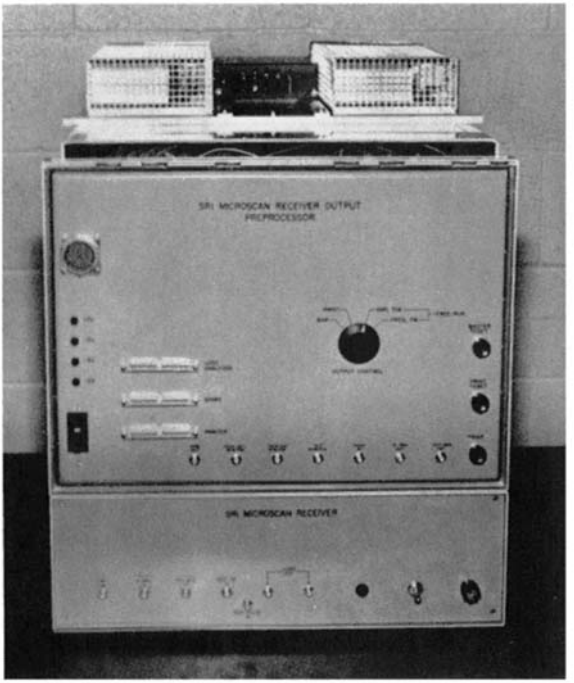


Figure 8.1. Experimental compressive receiver. (Fabricated by SRI International, courtesy of Avionics Laboratory, AFWAL.)

and its outputs are video signals. The top unit is the digitizing section, which converts the video signals to digital words. The relative sizes of these two units reflect their complexities.

In this chapter, the operating principle of compressive receivers, some key components, and some key circuits for processing the short compressed pulses will be discussed.

8.2. PRINCIPLE OF OPERATION (2–11)

Before discussing compressive receivers, the mechanism of compressing an FM signal into a pulse will be explained. The detailed mathematical expressions will be derived in later sections. Assume that an FM signal changes frequency linearly from f_0 to f_1 in the time interval t_0 to t_1 with a frequency versus time slope of m, as shown in Figure 8.2*a*. If this signal enters a DDL with a frequency versus time slope of $-m$, as shown in Figure 8.2*b*, the output from the DDL will be a burst of energy at time $t_0 + t_1$. This phenomenon can be explained as follows. The leading edge of the signal with frequency f_0 entering the DDL at t_0 is delayed by t_1, while the trailing edge of the signal with frequency f_1 entering

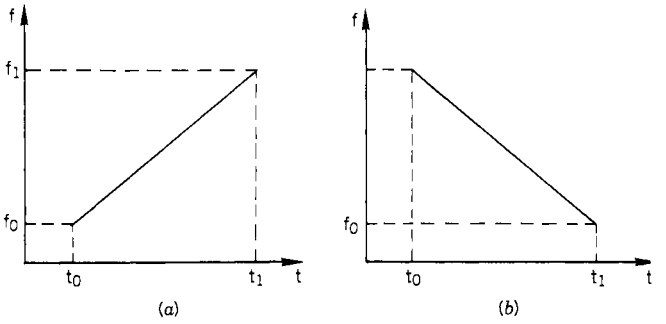


Figure 8.2. Input FM signal and matched DDL. Frequency versus time for (*a*) the input FM signal and (*b*) the matched DDL.

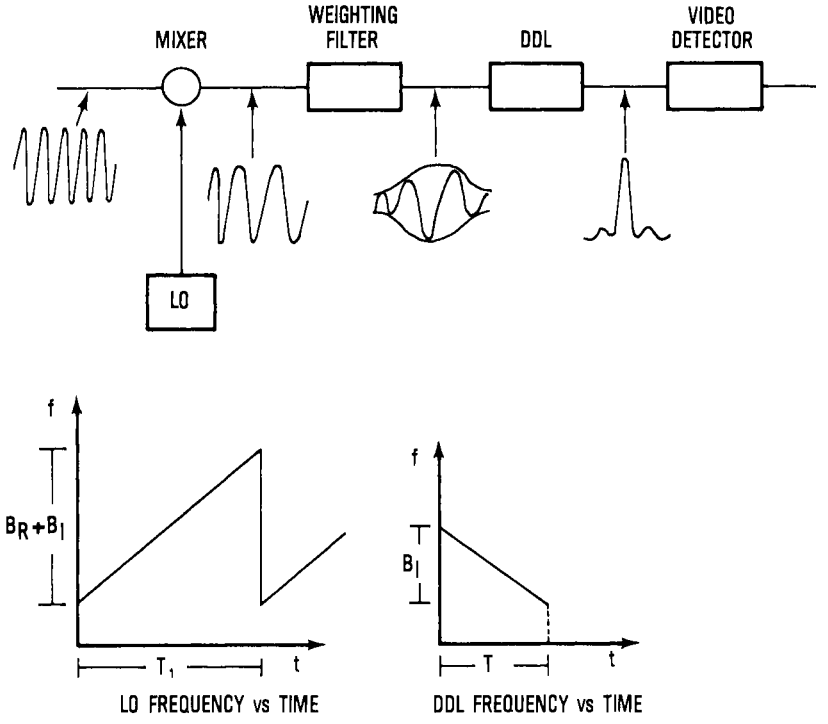


Figure 8.3. Basic compressive receiver.

the DDL at t_1 is delayed by t_0. Therefore, both the leading and trailing edges emerge from the DDL at $t_0 + t_1$. Thus the entire FM pulse from t_0 to t_1 is compressed and comes out at time $t_0 + t_1$. The frequency range $f_1 - f_0$ is the bandwidth of the DDL, and the time difference $t_1 - t_0$ is the dispersive delay time of the DDL. Compressive receivers use this basic idea to compress the input signal into short pulses.

A simple compressive receiver is shown in Figure 8.3. The first component in a compressive receiver is a mixer. The LO to the mixer is an FM signal. A weighting filter is used after the mixer to modify the output signal from the mixer. The output from the weighting filter passes through the DDL and is compressed into narrow RF pulses. The video detector will convert the RF pulses to video pulses. By measuring the positions of the output video pulses in a scan, the frequency of the input signals can be measured.

Assume that the input bandwidth of the receiver is B_R and the DDL bandwidth is B_I, which is also the intermediate-frequency (IF) bandwidth of the receiver. The frequency versus time slope of the sweeping LO matches the slope of the DDL. The LO should sweep a total frequency range equal to the sum of B_R and B_I. If the input signal is a continuous wave (CW), the output from the mixer will be an FM signal with a slope matching that of the DDL. The purpose of the weighting filter is to control the output waveform from the mixer, which in turn controls the shape of the compressed pulse from the DDL for side lobe suppression. It should be noted that both the weighting filter and DDL are linear devices. Thus their positions in the receiver can be exchanged without affecting the output pulse. In some designs, the weighting function is built into the DDL, and the two components become one. The DDL will compress the FM-modulated pulse into a very narrow pulse that is detected by the video detector.

To demonstrate how a compressive receiver reads frequency, three CW signals with frequencies f_a, f_b, and f_c are shown in Figure 8.4*a*. The heavy portion of the line should be ignored for the present discussion. Signals with frequencies f_a and f_c are at the upper and lower band edges of the receiver and signal with frequency f_b at the center of the band. In this special case, it is assumed that B_R equals B_I. In general, B_R is not necessarily equal to B_I. Also assume that the LO has zero flyback time as shown in Figure 8.4*b*. If the LO frequency is f_0, the input signal f_a is converted to frequency $f_0 + f_a$ and f_c is converted to frequency $f_0 + f_c$, as shown in Figure 8.4*c*. In Figure 8.4*c*, the portions of the signals in the band B_I are intercepted by the receiver and will be compressed by the DDL. The portions outside B_I will be missed by the receiver. In order for B_I to intercept both $f_0 + f_a$ and $f_0 + f_c$ and fill the entire bandwidth B_I, the LO must sweep the bandwidth $B_R + B_I$.

Let us examine the signals in the DDL a little more closely. To avoid confusion, the sweep period of the LO is represented by T_1, and the dispersive delay time of the DDL is T. From Figure 8.4*c*, it is seen that frequency $f_0 + f_a$ enters the delay line first followed by $f_0 + f_b$ and $f_0 + f_c$. This time difference is frequency dependent. The frequency range of all three signals intercepted by B_I are the same. Therefore, if the three signals have the same amplitude, the compressed outputs from the DDL should be identical. However, since frequency $f_0 + f_a$ first enters the DDL, the corresponding compressed pulse will leave the DDL first followed by pulses corresponding to frequencies $f_0 + f_b$ and $f_0 + f_c$, as shown in Figure 8.4*d*. It should be emphasized that the order of the output pulses depends only on the input frequencies f_a, f_b, and f_c. Thus, by measuring the positions of the pulses, the frequency of the input signals can be determined.

Some important factors in Figure 8.4 will be discussed here. First, the outputs

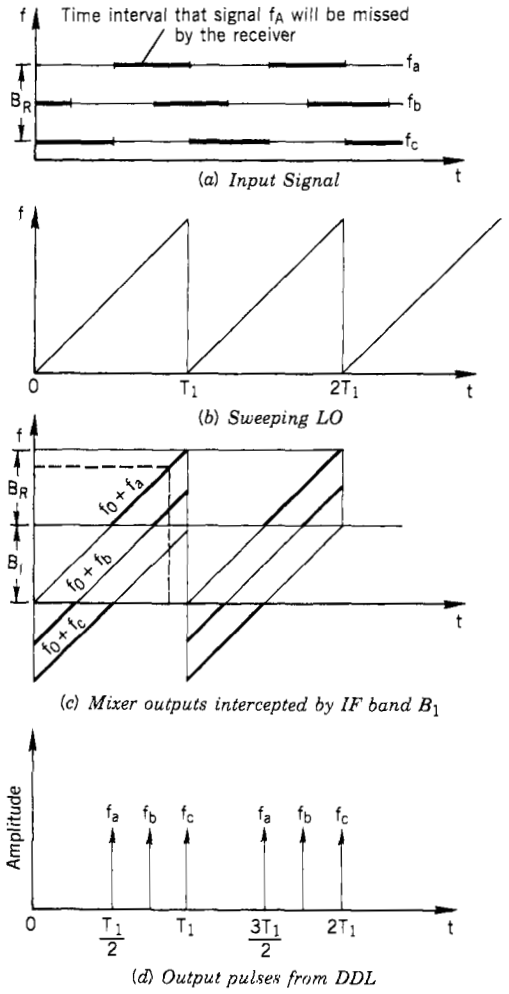


Figure 8.4. Time and frequency relation of a compressive receiver with $B_R = B_I$.

from the DDL occur only in the time from $\frac{1}{2}T_1$ to T_1 and from $\frac{3}{2}T_1$ to $2T_1$, and so on. During the time period from 0 to $\frac{1}{2}T_1$ and T_1 to $\frac{3}{2}T_1$, there are no outputs. This situation will be slightly different when B_R is not equal to B_I. Second, the compressive receiver does not have 100% probability of intercept (POI). For example, if a pulsed signal of frequency f_a arrives between time $\frac{1}{2}T_1$ to T_1 (shown by the heavy line in Fig. 8.4*a*), the output from the mixer is outside the IF bandwidth B_I, and the receiver will miss the signal. Similar arguments hold true for f_b between $\frac{3}{4}T_1$ and $\frac{5}{4}T_1$ and for f_c between T_1 and $\frac{3}{2}T_1$. In general, a pulsed signal of PW $< \frac{1}{2}T_1$ with any frequency in the input band of the receiver can be missed by the receiver if it arrives at a certain time period. If the input

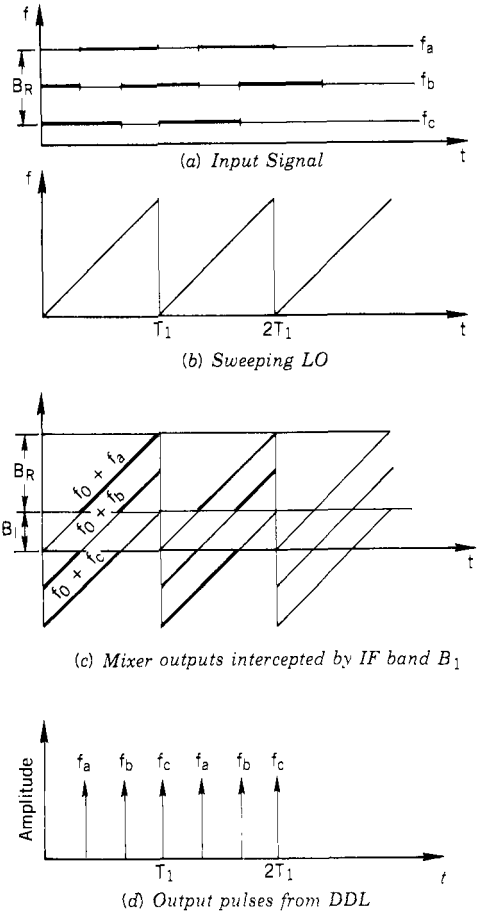


Figure 8.5. Time and frequency relation of a compressive receiver with $B_R = 2B_I$.

signal PW $> \frac{1}{2}T_1$, the receiver will always intercept portions of the pulse. However, if the input signal does not fill up the DDL, the signal will be processed by the receiver with less sensitivity. The probability of intercept of the compressive receiver will be discussed later.

If the LO sweeps less than the required bandwidth, as shown by the dotted lines in Figure 8.4*c*, the receiver will miss portions of the input signal. If one follows the dotted lines, it can be seen that the signal $f_0 + f_c$ does not fill the entire bandwidth B_I. Under this condition, the receiver will intercept f_c with a degraded sensitivity. Thus it is interesting to note that when the LO does not sweep enough bandwidth, the receiver intercepts signals near the input band edges with degraded sensitivity. Stated another way, a compressive receiver can receive signals outside of its designed bandwidth with less sensitivity and less POI.

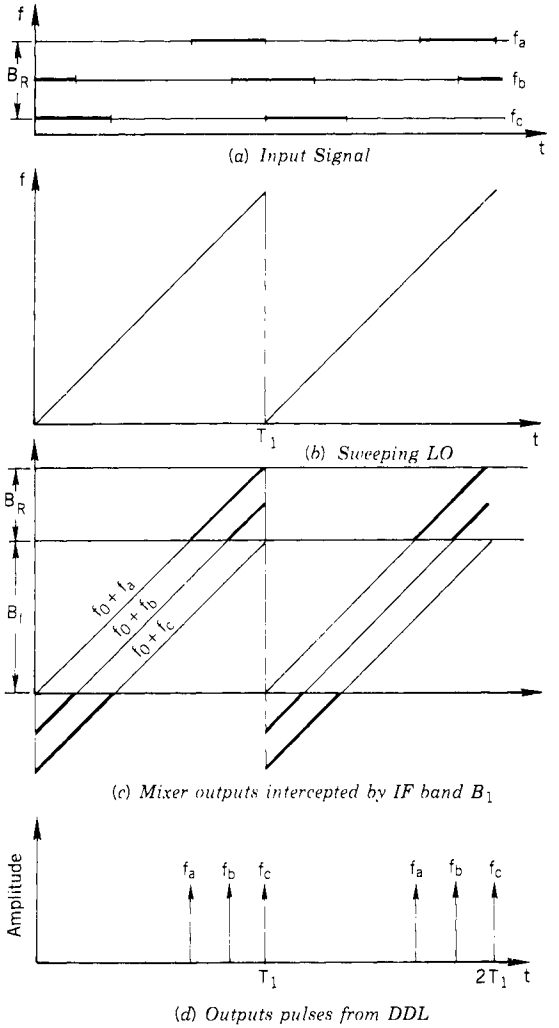


Figure 8.6. Time and frequency relation of a compressive receiver with $B_R = \frac{1}{2}B_I$.

In a compressive receiver, the IF bandwidth and the RF bandwidth are not necessarily equal. They can have almost any relation. Figure 8.5 shows the time and frequency relation of a compressive receiver with $B_R = 2B_I$ and LO period $T_1 = 3T$. In this arrangement, the input bandwidth of the receiver is wider than the arrangement in Figure 8.4 if the same DDL is used but the POI is degraded. This configuration is often used to improve the input bandwidth of a compressive receiver, especially in communication compressive receivers, since most of the input signals are CW and the POI is not of primary concern. During the period from 0 to $\frac{1}{3}T_1$ and T_1 to $\frac{4}{3}T_1$, there are no outputs from the DDL.

Figure 8.6 shows the time frequency relation of a compressive receiver with $B_R = \frac{1}{2}B_I$ and LO period $T_1 = \frac{3}{2}$. This arrangement improves the POI. However, in this arrangement, the IF bandwidth B_I is very large but the receiver input bandwidth is relatively narrow and the POI is still not 100%. Therefore, this approach is seldom used to improve the POI of a receiver. However, with a slight modification, this approach can achieve 100% POI (the detail will be discussed in Section 8.21).

8.3. MATHEMATICAL ANALYSIS (9, 10)

From the above discussion, it appears that the output from the DDL is a pulse with zero PW. However, this is not true. The compressed pulse has a certain PW and shape. In this section, mathematical analysis will be used to further elaborate the signal passing through the DDL and to predict the output pulse shape. The effect of the weighting function on the output pulse shape is also presented.

As mentioned above, assume that the bandwidth of the DDL is B_I and the dispersive delay time is T. Let us start the analysis from the output of the mixer. A pulsed FM signal with constant amplitude can be expressed as

$$S(t) = \begin{cases} \cos(\omega_0 t + \frac{1}{2}\mu t^2) & \text{for } -\frac{1}{2}T < t < \frac{1}{2}T \\ 0 & \text{elsewhere} \end{cases} \tag{8.1}$$

where ω_0 is the center angular frequency of the DDL, t is time, and μ is the scan rate, which can be expressed as

$$\mu = 2\pi B_I / T \tag{8.2}$$

Equation (8.1) can be written in exponential form as

$$S(t) = \begin{cases} \exp[j(\omega_0 t + \frac{1}{2}\mu t^2)] & \text{for } -\frac{1}{2}T < t < \frac{1}{2}T \\ 0 & \text{elsewhere} \end{cases} \tag{8.3}$$

since in this form it is easier to perform the mathematical operation.

The transfer function of the DDL can be expressed as

$$H(\omega) = \exp\left[j \frac{(\omega - \omega_0)^2}{2\mu} \right] \tag{8.4}$$

The signal at the output of the DDL is given by

$$G(\omega) = H(\omega)S(\omega) \tag{8.5}$$

where $S(\omega)$ is the Fourier transform of $S(t)$ of Eq. (8.3), and

$$S(\omega) = \int_{-T/2}^{T/2} S(t)\exp(-j\omega t)\, dt \tag{8.6}$$

where ω is the angular frequency.

To find the time domain output of the DDL, one needs to do an inverse Fourier transform of $G(\omega)$:

$$g(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(\omega)\exp(j\omega t)\, d\omega \tag{8.7}$$

By substituting Eqs. (8.3)–(8.6) into (8.7), one obtains

$$g(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)\left\{\int_{-T/2}^{T/2} \omega(\tau)\exp[j(\omega_0\tau + \tfrac{1}{2}\mu\tau^2)]\exp(-j\omega\tau)\, d\tau\right\} \times \exp(j\omega t)\, d\omega \tag{8.8}$$

In Eq. (8.8), $w(\tau)$ represents the effect of the weighting filter. Strictly speaking, the weighting filter should be a function of frequency. However, the weighting filter can be written as a function of time because after the mixer a CW input signal becomes a linear FM signal where the frequency is linearly proportional to time. The order of integration of Eq. (8.6) can be rearranged as

$$g(t) = \frac{1}{2\pi}\int_{-T/2}^{T/2} w(\tau)\exp[j(\omega_0\tau + \tfrac{1}{2}\mu\tau^2)]\left\{\int_{-\infty}^{\infty} H(\omega)\exp[j\omega(t-\tau)]\, d\omega\right\} d\tau \tag{8.9}$$

The integration in the braces of Eq. (8.9) with the substitution of $H(\omega)$ by Eq. (8.4) can be written as

$$\int_{-\infty}^{\infty} H(\omega)\exp[j\omega(t-\tau)]\, d\omega = \int_{-\infty}^{\infty} \exp\left[\frac{(\omega-\omega_0)^2}{2\mu}\right]\exp[j\omega(t-\tau)]\, d\omega$$

$$= \int_{-\infty}^{\infty} \exp\left\{j\left[\frac{\omega^2}{2\mu} + \left(-\frac{\omega_0}{\mu} + t - \tau\right)\omega + \frac{\omega_0^2}{2\mu}\right]\right\} d\omega \tag{8.10}$$

To integrate Eq. (8.10), assume

$$a^2 \equiv \frac{1}{2\mu} \tag{8.11}$$

$$2ab \equiv \frac{-\omega_0 + \mu(t-\tau)}{\mu}$$

$$b = \frac{-\omega_0 + \mu(t-\tau)}{2a\mu} = \frac{-\omega_0 + \mu(t-\tau)}{\sqrt{2\mu}} \tag{8.12}$$

and

$$c \equiv \frac{\omega_0^2}{2\mu} \tag{8.13}$$

After substituting Eqs. (8.11)–(8.13) into (8.10) and completing the square, Eq. (8.10) can be rewritten as

$$\int_{-\infty}^{\infty} \exp\{j[(a\omega + b)^2 + c - b^2]\}\, d\omega$$

$$= \exp[-j(b^2 - c)] \int_{-\infty}^{\infty} \exp[j(a\omega + b)^2]\, d\omega \tag{8.14}$$

Using the relation

$$\exp(jx) = \cos x + j \sin x \tag{8.15}$$

and

$$\int_0^{\infty} \sin(a^2x^2)\, dx = \int_0^{\infty} \cos(a^2x^2)\, dx = \frac{\sqrt{\pi}}{2a\sqrt{2}} \tag{8.16}$$

the integration in Eq. (8.14) can be evaluated, and the result is

$$\int_{-\infty}^{\infty} H(\omega)\exp[j\omega(t-\tau)\, d\omega = \sqrt{2\pi\mu}\, \exp\{j[\omega_0(t-\tau) - \tfrac{1}{2}\mu(t-\tau)^2 + \tfrac{1}{4}\pi]\} \tag{8.17}$$

After substituting Eq. (8.17) into (8.9), one obtains

$$g(t) = \sqrt{\frac{\mu}{2\pi}} \exp\left[j\left(\omega_0 t - \frac{\mu t^2}{2} + \frac{\pi}{4}\right)\right] \int_{-T/2}^{T/2} w(\tau)\exp(j\mu\tau t)\, d\tau \tag{8.18}$$

where τ is a dummy variable.

The term $\exp[j(\omega_0 t - \frac{1}{2}\mu t^2 + \frac{1}{4}\pi)]$ represents the frequency of the output pulse. The center frequency of the output pulse from the DDL is at ω_0, which is also the center frequency of the DDL, but the phase is shifted by $\frac{1}{4}\pi$. The amplitude of the output signal can be obtained by integrating the function

$$R(t) = \int_{-T/2}^{T/2} w(\tau)\exp(j\mu\tau t)\, d\tau \tag{8.19}$$

The constant $\sqrt{\mu/2\pi}$ is neglected during this discussion. Equation (8.19) states that the amplitude of the output pulse from the DDL is the Fourier transform of the weighting function. For the unweighted signal, $w(\tau) = 1$, and the integration of $R(t)$ yields

$$R(t) = \frac{\sin(\frac{1}{2}\mu Tt)T}{\frac{1}{2}\mu Tt} \tag{8.20}$$

This is the product of a sinc function ($\sin x/x$) with the delay time T. If this output is detected by a log video detector, the output has many side lobes. The first side lobe is about 13 dB below the main lobe. This first side lobe can be found as

$$\frac{\sin x}{x} = 1 \qquad \text{for } x = 0 \quad \text{(the main lobe)}$$

$$\frac{\sin x}{x} = \frac{-2}{3\pi} \quad \text{for } x = \frac{3\pi}{2} \quad \text{(the first side lobe)}$$

and

$$10 \log \left(\left[\frac{\sin x}{x} \right]^2_{x=3\pi/2} \Big/ \left[\frac{\sin x}{x} \right]^2_{x=0} \right) = -13.5 \quad \text{dB} \tag{8.21}$$

The rest of the side lobes decrease monotonically (see Fig. 7.2). Therefore, for one input signal, the compressive receiver will generate a string of pulses, a pulse with many side lobes in the time domain, rather than a single pulse. To measure the frequency of the input signal, the position of the main lobe must be measured. The side lobes of the output pulse will limit the two-signal dynamic range because they tend to mask signals whose amplitudes are less than the side lobe of another signal. For an unweighted DDL, the largest side lobe is about 13 dB below the main lobe; therefore, the dynamic range is limited to about 13 dB.

The usual way of improving the dynamic range of a compressive receiver is to suppress the side lobes by adding a weighting filter before (or after) the DDL. The weighting function can also be built in the DDL, that is, in a SAW DDL. One of the popular weighting functions is a cosine square on the pedestal, which can be expressed as

$$w(t) = K + (1 - K)\cos^2(\pi t/T) \tag{8.22}$$

where K is a constant between 0 and 1. Substituting $w(t)$ into Eq. (8.19) and carrying out the integration yields

$$R(t) = \frac{T \sin(\frac{1}{2}\mu Tt)}{2} \left(\frac{1 + K}{\frac{1}{2}\mu Tt} + \frac{(1 - K)\frac{1}{2}\mu Tt}{\pi^2 - (\frac{1}{2}\mu Tt)^2} \right) \tag{8.23}$$

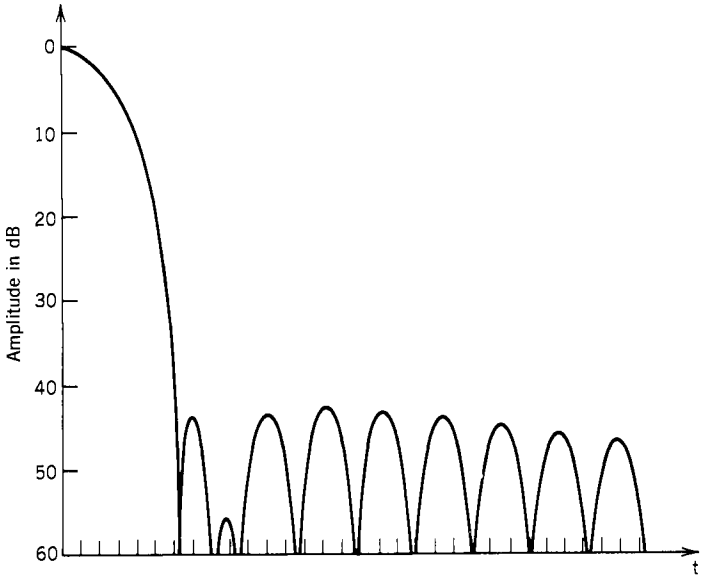


Figure 8.7. Output of a Hamming weighting filter.

When $K = 0.08$, this weighting function is referred to as the Hamming weighting function, which will reduce the side lobe to 42.8 dB below the main lobe. The power spectrum output from a Hamming weighting function is shown in Figure 8.7. Although the Hamming weighting function suppresses the time side lobes, it broadens the main lobe. Broadening the main lobe decreases its measurement accuracy in the time domain, which in turn decreases the frequency measurement accuracy of a compressive receiver. This is a trade-off in compressive receiver design. Many other different weighting functions reduce the side lobes of the output pulse and improve the dynamic range of the receiver.

8.4. BASIC EQUATIONS

In this section, the relationships between different parameters of a compressive receiver will be discussed. The characteristics of the compressive pulse that are very important to a compressive receiver are closely related to the characteristics of the DDL. These results will be listed in Table 8.1.

If the bandwidth of the DDL in a compressive receiver is B_I and the dispersive delay is T, the sweeping LO must scan at a rate to match the frequency versus time slope of the DDL. The scan rate is

$$S = B_I/T = \mu/2\pi \tag{8.24}$$

Table 8.1. Key parameters and their relations in a compressive receiver

RF bandwidth	B_R
DDL bandwidth	B_I
Dispersive delay time	T
Scan Rate	$S = B_I/T$
Oscillator scan width	$B_T = B_R + B_I$
Time to scan the total bandwidth (without retracing time)	$T_1 = \dfrac{B_R + B_I}{B_I} T$
Time to scan the total bandwidth (with retracing time T)	$T_1 = \dfrac{B_R + B_I}{B_I} T + T_R$
Frequency resolution ($k > 1$)	$\Delta f = k/T$
Maximum number of pulses from DDL per scan	$N = B_R/\Delta f = B_R T/k$
Period DDL has output per scan	$T_0 = B_R T/B_I$
Output pulse width from DDL	$P = k/B_I$
Output pulse rate	$R = B_I/k$

If the input bandwidth of the receiver is B_I, as mentioned in Section 8.2, the LO must sweep

$$B_T = B_R + B_I \tag{8.25}$$

Therefore, the total scan time is

$$T_1 = \frac{B_T}{S} = \frac{(B_R + B_I)T}{B_I} \tag{8.26}$$

If the sweeping LO has a retrace time of T_R, the total time is

$$T_1 = \frac{(B_R + B_I)T}{B_I} + T_R \tag{8.27}$$

The frequency resolution of the receiver is determined by the width of the main lobe of the compressed pulse. From the $\sin(\frac{1}{2}\mu Tt)$ term in Eqs. (8.20) and (8.23), one can find that the PW of the main lobe is proportional to $1/T$. In other words, the longer the delay time T and the narrower the main lobe, the finer is the frequency resolution the receiver can achieve. One can simply write the frequency resolution Δf as

$$\Delta f = k/T \tag{8.28}$$

where k is a proportionality constant. The value of k is usually between 1 and 2, and depending on the weighting function and the frequency detection scheme.

The maximum number of output pulses from the DDL equals the total RF

bandwidth B_R divided by the frequency resolution. Thus

$$N = B_R/\Delta f = B_R T/k \tag{8.29}$$

which is proportional to the time bandwidth product of the DDL.

From Figures 8.3–8.5 and the discussion in Section 8.2, it is already known that the compressed pulses come out of the DDL for only part of the sweeping time T_1. If T_0 is the time period for which the delay line has output, then

$$T_0 = \frac{B_R}{B_I} T \tag{8.30}$$

This relation is obvious from Figures 8.3–8.5. The output pulse width from the DDL equals the time period T_0 divided by the total number of pulses N, which can be expressed as

$$P = T_0/N \tag{8.31}$$

Substituting Eqs. (8.29) and (8.30) into (8.31), one obtains

$$P = k/B_I \tag{8.32}$$

The above relation can also be derived from Eq. (8.20). When $\mu Tt/2 = \pi$ or $t = 2\pi/\mu T$, $\sin(\mu Tt/2) = 0$, which labels the first zero of the output pulse in time domain, the width of the pulse can then be approximated by

$$P = \frac{2\pi}{\mu T} k \tag{8.33}$$

Substituting Eq. (8.24) into (8.33), one obtains

$$P = k/B_I \tag{8.32}$$

The speed of the pulses leaving the DDL is

$$R = N/T_0 = B_I/k \tag{8.34}$$

The above relationships are listed in Table 8.1.

8.5. DESIGN CONSIDERATIONS

The basic equations derived in Section 8.4 contain most of the important relationships for compressive receiver design. In this section, the applications of these relationships to receiver design will be demonstrated. To simplify the discussion, assume that the constant $k = 1$, since some numerical examples can be readily discussed.

Equation (8.28) indicates that the longer the dispersive time of the DDL, the finer the frequency resolution. The frequency resolution is determined by the dispersive delay time. If a frequency resolution of 1 MHz is desired, a dispersive time of 1 μsec should be used. However, the dispersive time of the DDL is also determined by the minimum PW the receiver processes. If the dispersive time is 1 μsec, it takes at least a 1-μsec pulse to fill up the DDL. The best a 100-nsec pulse can achieve is to fill 10% of the DDL. Thus the sensitivity of the receiver to process a 100-nsec pulse is at least 10 dB below its normal sensitivity. Therefore, the frequency resolution of a compressive receiver is closely tied to its minimum PW capability, as expected. The frequency resolution of the receiver should be related to the minimum PW by the approximation.

$$\Delta f \cong 1/\mathrm{PW}_{\min} \tag{8.35}$$

From Section 8.2, it is known that the wider the bandwidth of the DDL, the wider the input bandwidth of the compressive receiver, if the POI is kept constant. Therefore, one should anticipate that to widen the receiver bandwidth, an obvious solution is to increase the bandwidth of the DDL. However, widening the bandwidth of the DDL will create two effects on the compressed pulses that will affect the receiver design. First, Eq. (8.32) shows that the wider the bandwidth of the DDL, the narrower the output pulse. The video circuit following the detector must have enough bandwidth to pass the narrow pulse. If the bandwidth B_I of the DDL is 1 GHz, the output PW will be 1 nsec. To properly process this PW, the video bandwidth must be approximately 1 GHz which is not a trivial task. If the peak of the pulse shifts 1 nsec, the frequency error will be one frequency resolution unit.

Second, Eq. (8.34) shows that the output rate from the DDL is proportional to the bandwidth B_I. The logic circuit following the detector must have the speed to match the bandwidth. If $B_I = 1$ GHz, the logic following the detector must also operate at 1 GHz. If the available logic speed is half of B_I, say 500 MHz, a two-phase clock each 500 MHz will be used to achieve the 1-GHz logic speed, and this approach will complicate the logic design. With present logic speed, the bandwidth of the DDL in a compressive receiver seems limited by the logic speed rather than by the DDL technology.

In the above discussion, if the value of k is greater than unity, the requirements on the video bandwidth and the logic speed in the compressive receiver will not be as stringent, but the frequency resolution will suffer. However, this

relation can be used in designing compressive receivers with wide input bandwidth. It will be further discussed in Section 8.20.

Another important fact in designing a compressive receiver is that the output from the mixer can either chirp up or chirp down in frequency. As long as the frequency versus time slope of the DDL matches the chirp rate as shown in Figure 8.2, theoretically there is no difference in the up or down frequency chirp. However, in an actual DDL, the longer the delay time and the higher the frequency, the higher the insertion loss. To equalize the insertion loss over the desired frequency range, the higher input frequency to the DDL should have a short delay time. In other words, the output from the mixer should be chirped up in frequency.

As mentioned in Chapter 5, in addition to the desired frequency output, there are other frequency outputs from the mixer. However, only the desired frequency will match the frequency versus time slope of the DDL. Other frequencies from the mixer do not match the slope of the DDL, and their energy is not compressed into a pulse in the time domain. Therefore, the spur products from a mixer cause less problems in a compressive receiver than do other kinds of receivers.

8.6. DISPERSIVE DELAY (COMPRESSIVE) LINES (12–16)

The first component to choose in a compressive receiver is the DDL. The line must have the desired bandwidth and dispersive delay time. The bandwidth times the delay time is referred to as the time bandwidth product of the delay line. It is desirable that the insertion loss of the DDL be uniform across the operating frequency band. Since in a DDL higher insertion loss occurs at higher frequency, and the longer the delay time, the higher frequency is usually designed with short delay times and the lower frequency with longer delay times. Sometimes an equalizer is required to introduce additional insertion loss at the desired frequency to obtain a uniform insertion loss across the band.

There are many approaches to produce a DDL. Some of these approaches are still in the research stage. Although magnetostatic waves (MSWs) and crimped coaxial DDLs have been demonstrated in the laboratory, at the present time, further research is needed to adapt them to a compressive receiver. With today's technology, these two delay lines are not easy to build to meet specifications, that is, the desired frequency range, dispersive delay time, and frequency versus time slope. The commonly used DDLs are made of electromagnetic (EM) waves and SAW devices.

8.7. FOLDED TAPE MEANDER LINES (17–19)

The folded tape meander line is an EM DDL. A time bandwidth of 1000 can be achieved through this approach. The operation frequency can approach 4 GHz. This kind of DDL can only be fabricated one at a time, and there are no mass

production schemes available at the present time. The physical configuration of a folded tape meander line is shown in Figure 8.8. The meander line is totally immersed in dielectrics. In practice, this is accomplished by two sheets of dielectric material of thickness D separating the meander line from the ground planes above and below and individual pieces of dielectric material of proper size separating the meander line turns from each other. Dielectric strips are also placed at the four sides of the DDL.

The phase shift for each turn of the meander line (a turn is defined as a single piece of conducting tape of width L with periodic structures) is given implicitly by (ref. 17)

$$\cot^2\left(\frac{\omega L}{2C}\right) = \frac{\gamma - \cos\theta - \gamma' \cos 2\theta + \cdots}{\gamma + \cos\theta + \gamma' \cos 2\theta + \cdots} \cot^2\left(\frac{\theta}{2}\right) \tag{8.36}$$

where ω is the angular frequency; L is the width of the meander line ($L = \frac{1}{4}\lambda_0$); C is the speed of light in the dielectric material; γ, γ' are coupling parameters determined by dimensions S, B, W, and D (as shown in Figure 8.8); and θ is the phase shift at frequency ω. The time delay of a single turn can be obtained by differentiating Eq. (8.36) with respect to ω:

$$\tau = d\theta/d\omega \tag{8.37}$$

The delay time versus frequency is shown in Figure 8.9 with different values of γ. Smaller value of γ will provide tighter coupling.

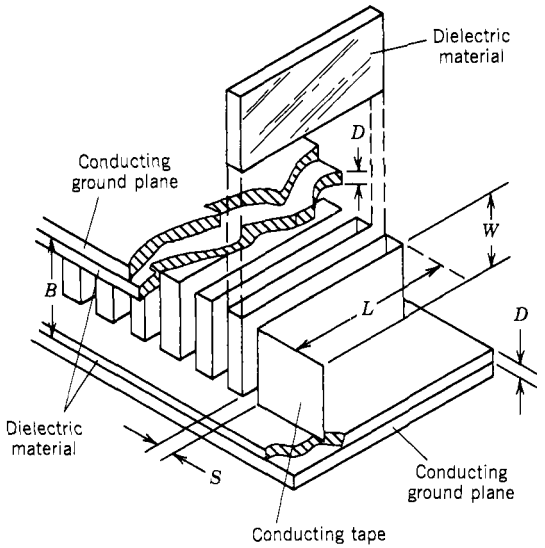


Figure 8.8. Folded tape meander line. (Based on Hewitt, ref. 17.)

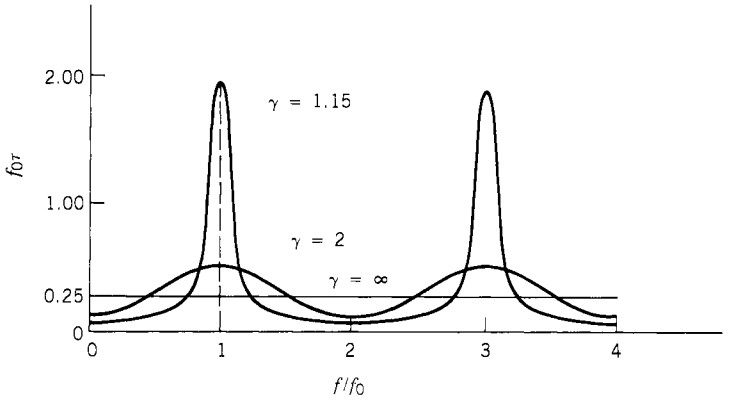


Figure 8.9. Delay versus frequency for a single meander line. (Based on Hewitt, ref. 17.)

The relationship between γ, γ' and the ratio W/S are plotted in Figures 8.10 and 8.11 for four values of D/S. It is possible to construct a meander line with a delay accuracy of 1% by using the values of γ and γ' in Figures 8.10 and 8.11. The total delay of a meander line at a given frequency is the sum of all the component delays of each turn at that frequency. Figure 8.12 shows the delay time varying with the three sections of a meander line at frequencies f_1, f_2, and f_3. The three

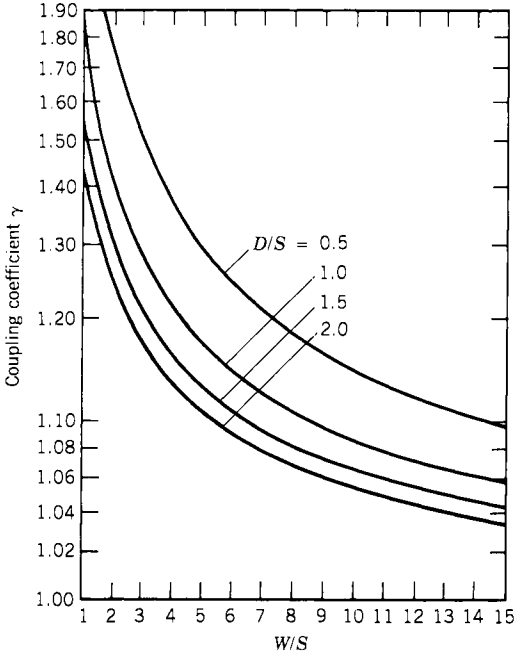


Figure 8.10. Parameter γ versus W/S and D/S. (Based on Hewitt, ref. 17.)

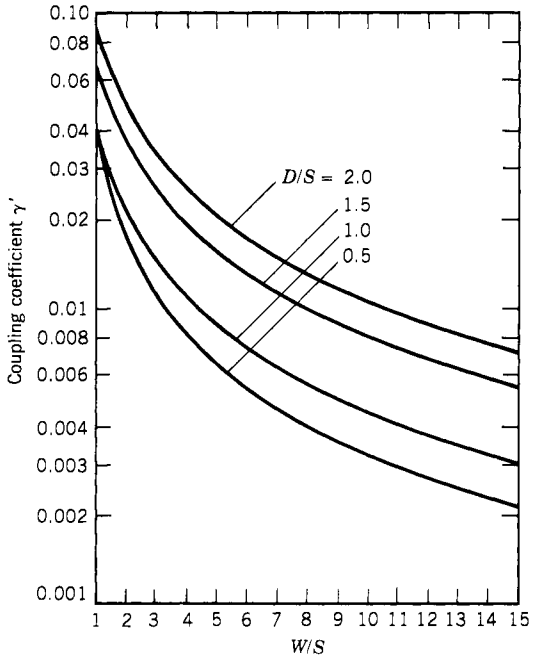


Figure 8.11. Parameter γ' versus W/S and D/S. (Based on Hewitt, ref. 17.)

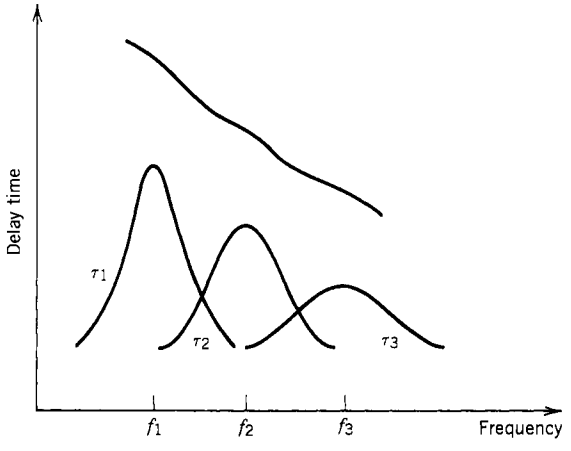


Figure 8.12. Linear delay synthesis using three meander line sections.

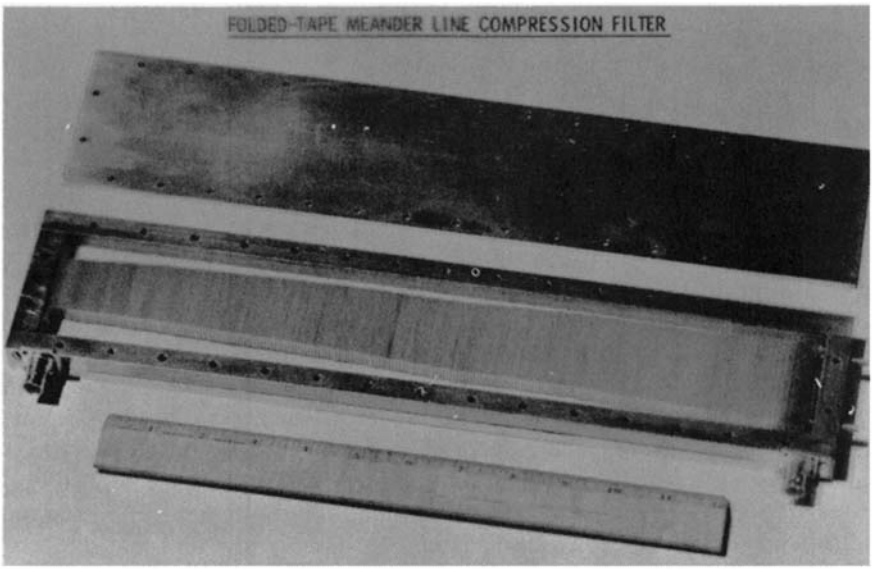


Figure 8.13. EW meander line. (Fabricated by SRI International, courtesy of Avionics Laboratory, AFWAL.)

sections added will give the overall delay characteristics. The total delay time of the line is the sum of $\tau_1 + \tau_2 + \tau_3$. In general, to design a meander line, the delay time τ_k of any single section made of k turns with $f = f_k$ will be determined first. The synthesis of a desired delay function is accomplished by manipulating the number of sections; the center frequency of each section, which is related to the dimension L; the number of turns in each section; and the γ, γ' for each section.

Dispersive delay lines in gigahertz can be fabricated with this method. A metallic ribbon is used to fabricate this kind of DDL. The dimensions of W and S in Figure 8.8 are determined first. The length L is changed from turn to turn. After the metal ribbon is made into the proper shape, dielectric sheets of proper dimensions are inserted between the turns. Figure 8.13 shows a meander line. Some fine adjustments on the meander line can be made by moving the dielectric materials around it on a trial-and-error basis to improve the frequency versus time linearity of the line. This procedure increases the complexity of manufacturing folded tape meander lines.

8.8. SURFACE ACOUSTIC WAVE (SAW) DISPERSIVE DELAY LINES (20–31)

In a SAW DDL, the input electric signal is changed into an acoustic signal by an input transducer. The acoustic signal is delayed dispersively in the substrate; and at the output of the delay line, the signal is transformed back into an electric

signal. Since the velocity of SAWs is approximately five orders of magnitude less than that of the EW wave, the size of the SAW DDL is very small in comparison with an EM DDL. The common materials used for the SAW devices are $LiNbO_3$, $LiTiO_3$, and quartz. A time bandwidth product of 1000 has been achieved by using SAW technology. The operating frequency ranges from tens of megahertz to a few gigahertz. There are several approaches to designing the SAW DDLs. Two common designs will be discussed here. In the first design, the acoustic wave travels from the input transducer directly to the output transducer. In the second design, the acoustic wave is reflected from some frequency-dependent grating to reach the output transducer. Probably, SAW DDLs will be the most commonly used delay lines in a compressive receiver in the near future because of the relative ease in manufacturing them and their small size.

A basic configuration of a DDL is shown in Figure 8.14. In this DDL, the high frequency has short delay time and the low frequency has long delay time. The two transducers are facing each other on a substrate. The electrodes on the transducers are nonuniformly distributed, as shown in Figure 8.15. The electrodes are metal strips deposited on the surface of the substrate. The locations of the electrode can be determined as (ref. 20)

$$x_n = Vt_n \tag{8.38}$$

where V is the average surface wave velocity. The surface wave has different velocities with the metal finger deposition on the surface. If the metal electrode width to the center-to-center spacing is equal everywhere to a constant β, and

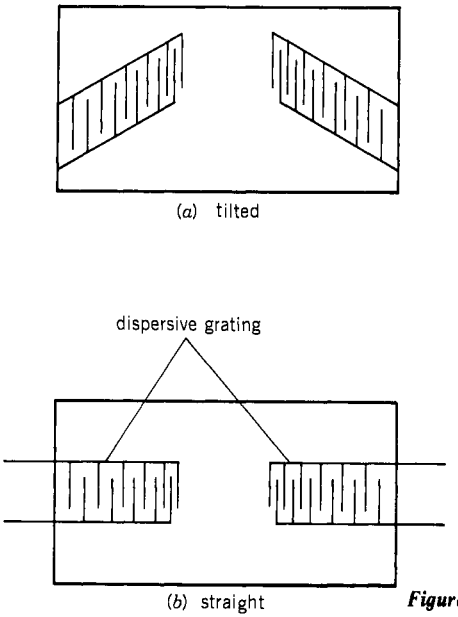


Figure 8.14. Down-chirp DDL transducer geometry.

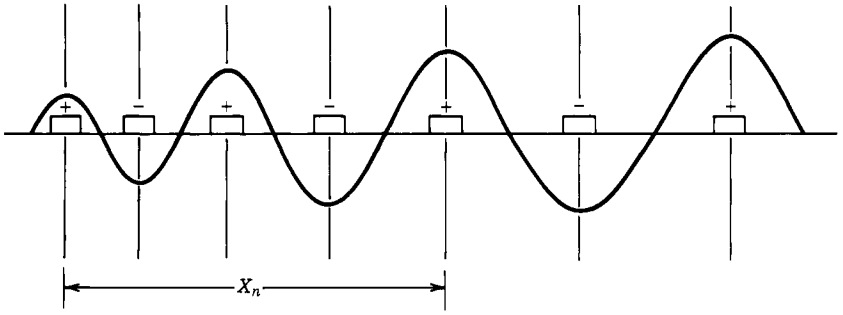


Figure 8.15. Electrodes on a transducer producing an FM chirp.

V_m and V_0 are the surface wave velocities under the electrode and without the electrode, respectively, then the average velocity can be written as

$$V = \frac{V_0 V_m}{V_m + \beta(V_0 - V_m)} \tag{8.39}$$

The t_n in Eq. (8.38) can be expressed as

$$t_n = \frac{T(f_0 + \frac{1}{2}B)}{2B}\left[1 - \sqrt{1 - \frac{2Bn}{T(f_0 + \frac{1}{2}B)^2}}\right] \tag{8.40}$$

where T is the total differential delay time, f_0 is the center frequency, B is the bandwidth of the transducer, and n is an integer. In Figure 8.15, the waveform generated by the electrodes are related to the distance between them. As shown in Figure 8.14, since the high-frequency ends of the electrodes are close together, the high frequency in the DDL will have a short delay time.

The second approach to make a SAW DDL is with reflectors fabricated on the substrate. As shown in Figure 8.16, the input and output transducers are slanted dispersive structures. High frequency is launched at the outside portions of the transducer, and the low frequency is launched at the center portion of the transducer. Facing the transducers there are dispersive gratings. The gratings can either be metal strips or grooves etched in the surface of the substrate. The high-frequency path is shorter than the low-frequency path, and the delay time is linearly related to frequency. In this kind of design, one more dimension is introduced as compared to Fig. 8.14. It is easier to design a DDL with a high time bandwidth product by using the reflective structure because the additional degree of freedom is available. Dispersive delay lines with a bandwidth around 500 MHz and a delay time of a few hundred nanoseconds can be achieved by SAW technology.

With present technology, the major deficiencies of a SAW DDL are the high insertion loss and the limited dynamic range in the time domain. The excess

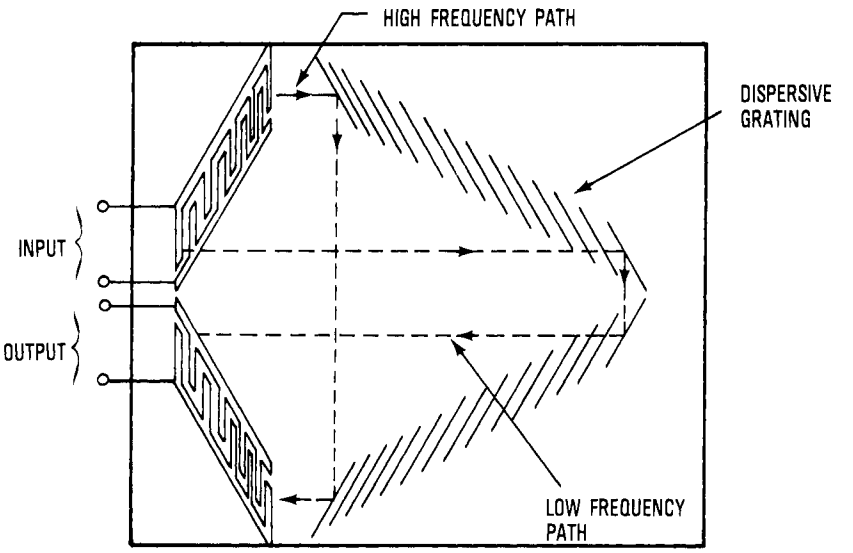


Figure 8.16. SAW DDL with reflector structure.

insertion loss must be recovered by additional amplifications, which are costly and power consuming. The triple transit in the DDL will limit the dynamic range of the receiver. As previously mentioned, only the signal with the correct frequency versus time slope will be compressed by the DDL. Signals with a frequency versus time slope different from the DDL will not be properly compressed. Therefore, their outputs will be relatively low. In a SAW DDL, only the fundamental frequency output has the correct frequency versus time slope to match the input signal. The triple-transit output has a dispersive delay time three times that of the fundamental output. Therefore, the input signal will not be compressed properly by the triple-transit effect. Although the triple transit is very low at the output of the DDL, for designing a receiver with high dynamic range, this is one of the limiting factors. Figure 8.17*a* shows the frequency response of a SAW DDL, and Figure 8.17*b* shows the simulated results of the output of a SAW DDL. The first output is the compressed output from the fundamental frequency. The output that follows, which is spread in the time domain and is about 52 dB below the main signal, is caused by the triple transit. This triple transit and the side lobes usually limit the dynamic range of the receiver.

Dispersive delay lines with wider bandwidth and longer dispersive delay times are desirable because the present bandwidth of the SAW DDL matches state-of-the-art logic speed. However, it is anticipated that the logic speed will increase with time in the near future. To make compressive receivers with wide bandwidths and high POIs, the DDL might be the limiting factor in the future.

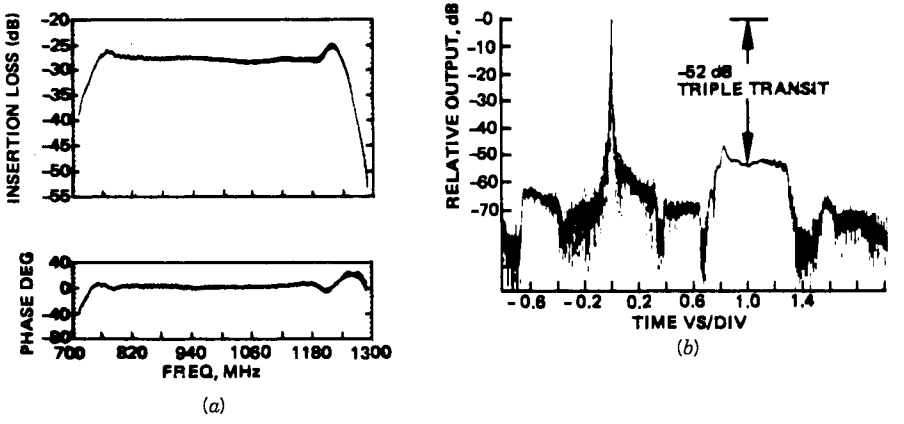


Figure 8.17. Output of a SAW DDL. (*a*) Frequency response of the SAW compression line. (*b*) Simulated recompressed pulse characteristics of the SAW chirp filters. (Courtesy of Hughes Aircraft Co.)

A longer dispersive delay time is also desirable because the dispersive time can be used to trade the processing logic speed (see Section 8.20).

8.9. CRIMPED COAXIAL DISPERSIVE DELAY LINES (16)

Crimped coaxial DDLs are fabricated by introducing discontinuities (crimps) on semirigid microwave coaxial cables. The crimps are introduced by squeezing the outher conductor of the semirigid cable either by a tube cutter or by a crimper, as shown in Figure 8.18. The degree of crimp is very important. It should be deep enough to reflect the input signal without cricking the outer conductor and interrupting the current flow, which would cause very high insertion losses.

These crimps change the capacitance of the conductor and thus the impedance at those points. If a signal of single frequency travels along the coaxial cable, energy will be reflected back at each discontinuity of the cable. These reflected signals will interfere constructively at the input port after some time delay. If the positions of the discontinuities are properly located, the delay time

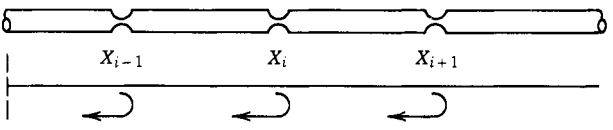


Figure 8.18. Crimped coaxial cable DDL.

introduced will vary linearly with the input signal frequency. The recurrence relationship used in calculating crimp positions is

$$X_{i+1} = X_i + \frac{\sqrt{\varepsilon_r}C}{2f_L + 2(f_H - f_L)/L} \tag{8.41}$$

where X_i is the length along the delay line where the crimp is introduced; L is the total length of the line; f_H and f_L are the high and low frequencies of the DDL, respectively; C is the speed of light; and ε_r is the relative dielectric constant of the cable.

This kind of DDL is relatively easy to fabricate and has 50 Ω impedance automatically. This technique is suited to generate a delay line of wide bandwidth with relatively short dispersive delay time because long delay times will have very high insertion losses. Dispersive delay times of 100 nsec and bandwidths of 1–2 GHz have been accomplished. Since this delay line is a one-port device, a circulator must be used to separate the input and the output. This is one of the major disadvantages.

8.10. SWEEPING LOCAL OSCILLATORS (32–35)

The sweeping LO must sweep the range that equals the RF and IF bandwidths with the frequency versus time slope matching the DDL. Usually the LO is made of a voltage-controlled oscillator (VCO) (discussed in Section 5.17). A YIG oscillator is often too slow to be used as the LO of a compressive receiver. One of the major difficulties in designing the LO is that the frequency versus time of a VCO cannot be easily measured. One of the schemes to evaluate an LO is to build a unit consisting of the LO and a matched DDL. By measuring the output pulse from the DDL, the performance of the LO can be indirectly determined. In this measurement scheme, the DDL is used as a reference unit. Therefore, the frequency versus time slope of the DDL should be made as linear as possible because the DDL is used as a reference in measuring the LO performance. The characteristics of a DDL can be measured by using a network analyzer. The frequency versus time slope of the LO should also be designed as linear as possible to match the DDL. After the LO and DDL are assembled, the compressed pulse from the DDL is displayed on an oscilloscope. If the compressed pulse is deviated too far from the desired output waveform, it means that the linearity of the LO needs to be improved. However, in a practical setup, the performance of the DDL will not be perfect; sometimes the characteristics of both the LO and the DDL will be adjusted to improve the output waveform. Therefore, the fabrication procedure is rather complicated.

A common method of generating an FM chirp from a VCO is applying a ramp voltage as the tuning voltage. The output frequency from the VCO will

then be a function of time. In general, the frequency versus voltage curve of a VCO is not linear. Therefore, applying a linear voltage versus time on the VCO will not generate a linear FM chirp. Linearizing circuits are often required to improve the linearity of a VCO. One approach to design a linearizing circuit is to measure the steady-state frequency versus voltage of the VCO and then divide the frequency axis into many equal sections, as shown Figure 8.19. If the voltage applied to the VCO changes from V_1 to V_2 to V_3, and so on, in equal time, the frequency output from the VCO will be nonlinear with respect to time t. A typical voltage versus time curve is shown in Figure 8.20. In the figure, the voltage applied does not change linearly with respect to time. The voltage versus time is shown on the right side of the figure. Theoretically, the voltage versus

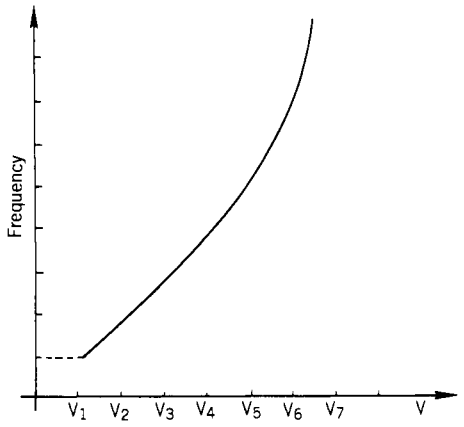


Figure 8.19. Frequency versus voltage of a VCO.

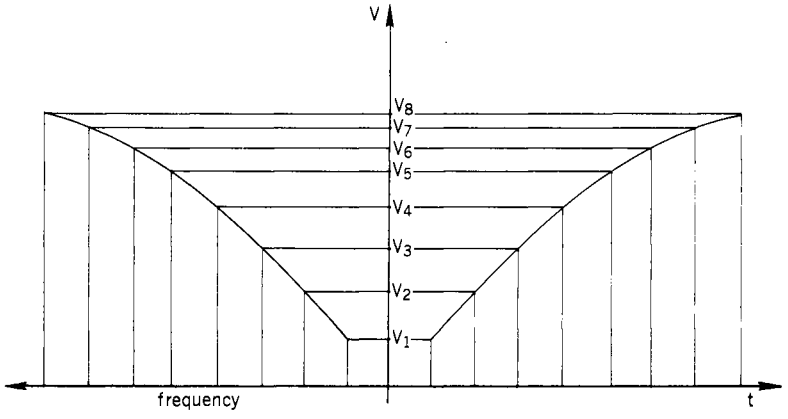


Figure 8.20. VCO tuning voltage versus time and frequency.

time curve can be stored in a memory circuit to generate the desired voltage shape. However, due to the short response time required to generate the voltage, it is often impractical to use a memory circuit.

In Figure 8.20, the voltage versus time curve is somewhat exponential, similar to the voltage on a charging capacitor. Thus, one of the linearizing circuits (ref. 35) can be achieved through a charging capacitor, as shown in Figure 8.21. Prior to the start of the sweep, the switch S is closed, clamping the capacitor to voltage $-V$. At the start of the sweep, the switch is opened and the capacitor starts to charge exponentially toward a voltage V_0. If the VCO frequency versus voltage characteristics were exactly exponential, applying this capacitor to the VCO would cause the frequency to vary linearly with time.

The characteristics of a VCO, however, are not exactly exponential. To correct the deviations from exponential characteristics, the voltage V_0 is varied while C is charging from one fixed value to another during the sweeping period. One way to change voltage V_0 during the sweeping period is to use many transistor switches Q's, as shown in Figure 8.22. In this figure, each switch is connected to a constant voltage supply through a resistor. The switches are controlled by a clock synchronized to the sweeping of the LO. When the switch is turned on, a constant current will flow through for the duration of the switching pulse. The constant current can be adjusted by changing the resistor in series with the switch. The constant current applied to capacitor C results in a net current that will change the voltage across the capacitor. By adjusting the resistors in the circuit, the voltage across the capacitor will be made to match the desired tuning voltage. In practice, the resistors are adjusted such that the desired output compressed pulse is obtained. Sometimes increasing the number of tuning stages can improve the linearity of the VCO. Therefore, the fabrication of the LO of the microscan receiver is rather complicated if a VCO is used to generate the desired frequency.

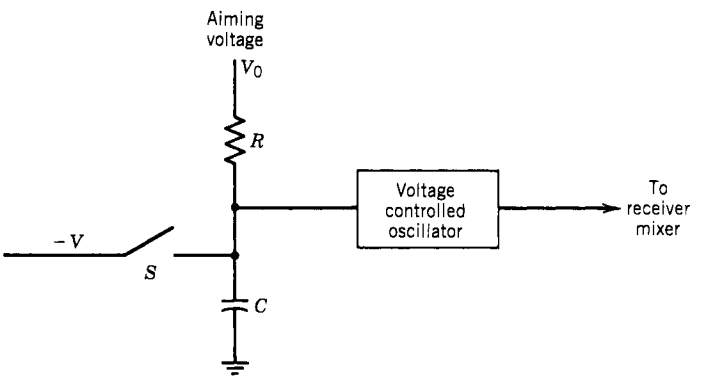


Figure 8.21. VCO driving technique.

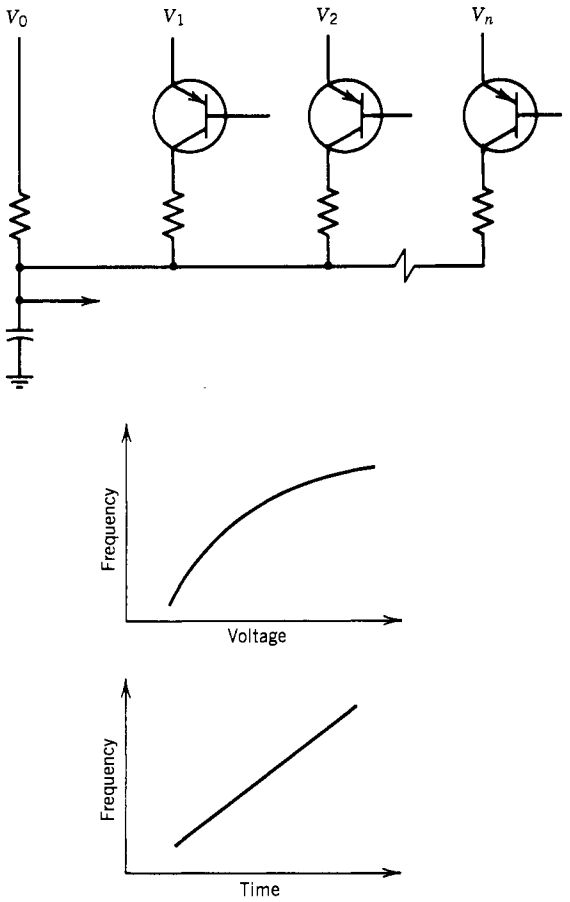


Figure 8.22. VCO driving with multiple tuning voltage.

8.11. SWEEPING LOCAL OSCILLATORS MADE OF DISPERSIVE DELAY LINES (36)

Another approach to generate the sweeping frequency is to input a pulse through a DDL. If an RF pulse with a frequency matching the center of the DDL is applied at the input of a DDL, the output will be a sweeping frequency. Figure 8.23 shows such an LO arrangement. The mixer is used to shift the DDL output frequency to the desired range. Since the frequency domain representation of an RF short pulse is a wide frequency with the center at the carrier frequency of the pulse, the DDL will delay one end of the frequency with respect to the other end, and a shirp frequency will be generated. Theoretically, the input to the DDL can be a very short video pulse. However, the spectrum component at the center

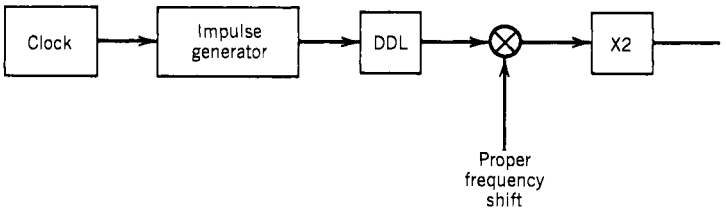


Figure 8.23. Sweeping LO made of DDL.

frequency of the DDL will be low. As a result, the output from the DDL is also low. Therefore, a short RF pulse rather than a video pulse is used as the input of the DDL.

One of the advantages of this approach is that the generated frequency versus time can be predicted. Since the frequency versus time of the DDL can be measured and it matches its generated frequency, the LO performance can be determined. Thus the frequency versus time of the LO can be properly designed to match the compressive line. Another advantage with this approach is that there is no retrace time as in the VCO approach. As soon as the last frequency leaves the DDL, the beginning frequency can start to come out of the DDL.

The disadvantages of this approach are (1) the output from the DDL is low and (2) the DDL with a wide time bandwidth product is required. To avoid confusion, refer to the DDL used for the LO as an LO DDL and the DDL used for compressing FM signals as a compressive DDL. Since the energy in a short RF is rather low and the insertion loss of a DDL is usually high, the output from the LO DDL will be very low. High-gain RF amplifiers are needed to amplify the output from the LO DDL to a reasonable level so that it can drive a mixer. The frequency versus time slope of the LO DDL must match that of the compressive DDL. However, the bandwidth of the LO DDL is the sum of the RF and IF bandwidths. Thus its bandwidth is wider than that of the compressive DDL. The dispersive delay time of the LO DDL is also longer. In a compressive receiver, if the RF bandwidth is equal to the IF bandwidth, the time bandwidth of this DDL is four times that of the compressive DDL. Therefore, the LO DDL will be the one to limit the bandwidth of the receiver if the processing logic is not the limiting factor.

When the RF bandwidth equals the IF bandwidth in a compressive receiver, the compressive DDL can also be used as the LO DDL. Figure 8.24 shows such an arrangement. In this figure, two compressive DDLs are connected in cascade to make an LO DDL. The first DDL will stretch the input RF pulse into a chirp signal. The second DDL will increase the dispersive delay time between the two end frequencies. The mixer shifts the output to the desired frequency range. The frequency doubler following the mixer will double the frequency range. Therefore, the output from this arrangement will satisfy the LO requirements. The difficulties with this approach is the relatively high insertion loss in the DDLs

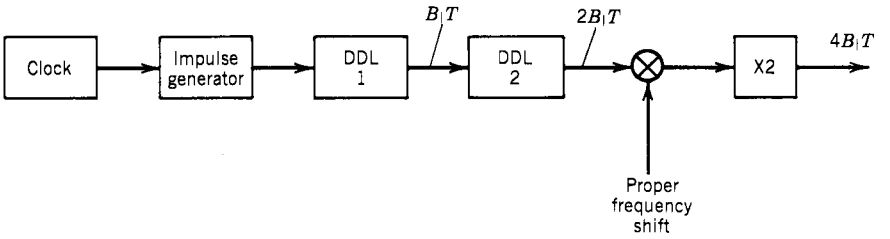


Figure 8.24. LO using cascade DDLs and frequency doubler.

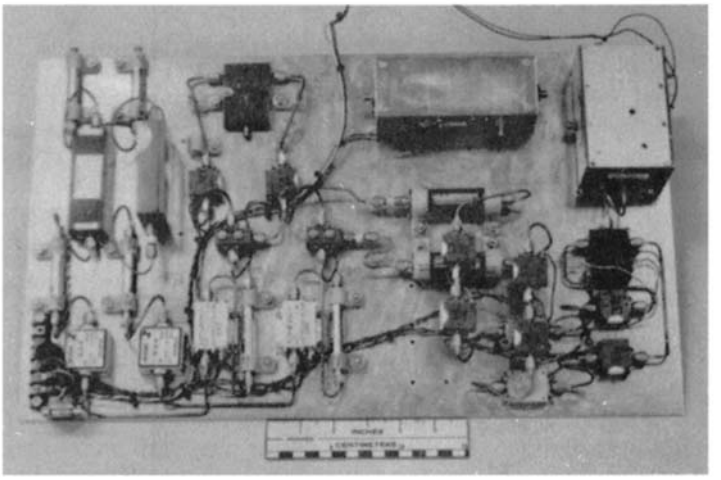


Figure 8.25. Experimental setup of a chirp LO with DDL. (Courtesy of Hughes Aircraft Co.)

and the low efficiency of the frequency doubler. This idea can be extended to make one DDL with a dispersive delay time doubling that of the compressive DDL but with the same frequency coverage. A frequency doubler can be used to extend the frequency to the desired range as shown in Figure 8.23. With this approach, the time bandwidth product of the LO DDL is twice that of the compressive DDL rather than four times.

Figure 8.25 shows an experimental setup of such a chirp LO. Two LO frequencies are generated, and they are used in a compressive receiver with interlace scan, which will be discussed in the next section. Figure 8.26 shows the input and output of the DDL. The inputs are shown as pulses in the time domain (Fig. 8.26*a*), and the outputs are chirped signals in time domain (Fig. 8.26*b*). Figures 8.26*c* and *d* show the results of Figure 8.26*b* in the frequency domain.

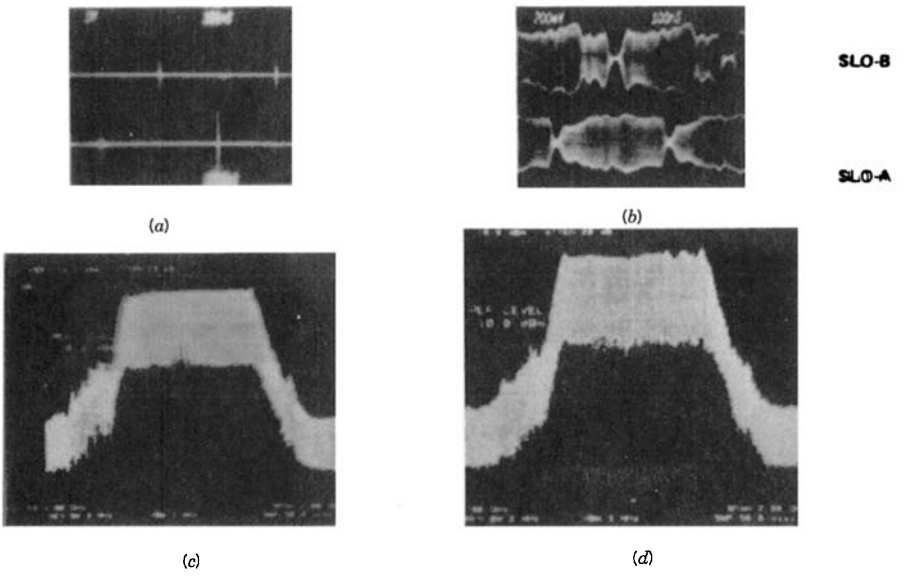


Figure 8.26. Inputs and outputs of DDLs (*a*) impulse input (*b*) output in time domain (*c*)(*d*) outputs in frequency domain. (Courtesy of Hughes Aircraft Co.)

8.12. COMPRESSIVE RECEIVER WITH INTERFACE SCAN

First, compare the POI of a compressive receiver against a scanning superhet receiver. If a superhet receiver with a frequency resolution Δf is used to scan over a certain RF bandwidth, the receiver must scan with a speed less than the impulse response time of the IF filter. As discussed in Section 5.22, the maximum scan rate without degrading a receiver sensitivity is $(\Delta f)^2$. Compared to a compressive receiver of the same frequency resolution, the scan rate of the compressive receiver is B_I/T, where $T \simeq 1/\Delta f$. Therefore, a compressive receiver can scan $B_I T(B_I/\Delta f)$ times faster than a superhet receiver of the same frequency resolution. In other words, a compressive receiver has a better POI than a superhet receiver by a factor equal to the time bandwidth product of the DDL.

Although the compressive DDL has improved POI, the POI is still less than 100%. As shown in Figures 8.3–8.5, if a pulsed signal arrives at the receiver with a certain frequency and at a certain time, the receiver will miss the input signal. To improve the POI, an interlace scan can be added to the receiver. Theoretically, if the RF bandwidth is greater than the IF bandwidth, an interlace scan can be used to improve the POI. However, in practical receivers, the interlace scan approach often applies to compressive receivers with RF bandwidth B_R equal to the IF bandwidth B_I. The basic configuration of an interlace scan is shown in Figure 8.27. An additional set of LOs and mixers are required. If the sweeping LO does not have any retrace time, the additional scan will start in

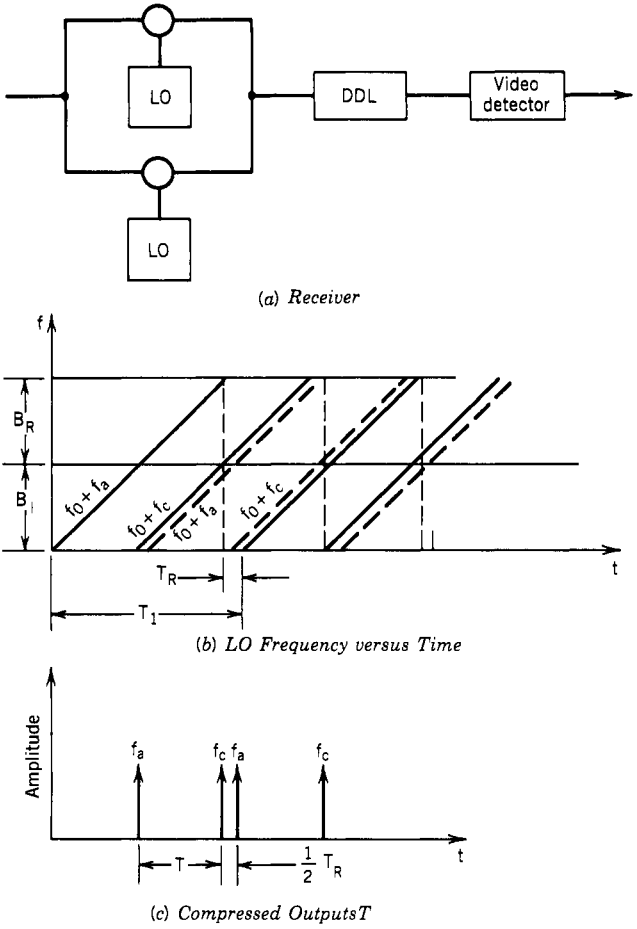


Figure 8.27. Compressive receiver with interlace scan.

the center of the two scans. Instead of only half of the period T_1 having output, the pulses will come out the DDL over the entire period T_1.

As shown in Figure 8.27, any input pulse will be intercepted by the receiver. In other words, the receiver will not miss any pulse. However, a short pulse may be intercepted by two consecutive scans, and the DDL is not filled by the pulse. Under this condition, the pulse is received with a degraded sensitivity. Although the interlace scan improves the POI of a compressive receiver, more hardware is required. The logic circuits following the detector must process twice the number of outputs. At the input, the signal is divided into two parallel paths so additional amplification is required. These are some of the disadvantages of a compressive receiver with an interlace scan.

8.13. REQUIREMENTS OF LOGARITHMIC AMPLIFIER

Like all the RF receivers, the RF signal in a compressive receiver must be converted to a video signal for further processing. To measure the amplitude of the input signal, a logarithmic (log) amplifier should be used at the output of the DDL to convert the output RF pulse to a video pulse. A simple diode detector usually does not provide enough dynamic range to process signals over a wide amplitude range. Discussion of the log amplifier is included in Section 5.5. However, because there are special concerns of the log amplifiers in a compressive receiver, the requirements of a log amplifier in a compressive receiver will be discussed here.

In addition to the linearity and dynamic range requirements, the bandwidth of a log amplifier in a compressive receiver must be wide enough to process the narrow compressed pulses. The recovery time of the log amplifier must be very short. The video bandwidth of the log amplifier should be approximately equal to the bandwidth of the DDL. If the bandwidth of the DDL is 500 MHz, the output pulse is approximately 2 nsec. The bandwidth of the log amplifier must be about 500 MHz. If the bandwidth is too narrow, several adverse effects will occur. First, the amplitude of the input pulse will be measured incorrectly because the compressed pulse will be wide and lose its amplitude fidelity.

Second, the accuracy of the input signal frequencies will be degraded because of the widening of the compressed pulse. The frequency of the input signal is determined by measuring the center position of the compressed pulse. Widening the video pulse by the log amplifier will degrade the position measurement of the pulse.

Third, the receiver may miss signals when they are close in frequency, a major problem in a receiver. Signals close in frequency come out of the DDL as pulses close in time. If the recovery time of the log amplifier is long, the log amplifier will detect the first pulse but may miss the next pulse in time, especially when the second pulse is lower in amplitude than the first. As a result, the receiver will miss the signal corresponding to the second pulse.

From the above discussion, it is obvious that the video bandwidth of the log amplifier is extremely important (in addition to the linearity requirement). The time delay caused by the individual RF amplifiers used in the log amplifier must be carefully compensated. If the delay time is not properly compensated, the staggering effect will also widen the output video pulse. Therefore, although there is only one log amplifier required in a compressive receiver as compared to many log amplifiers required in a channelized receiver, this one log amplifier must be carefully designed.

8.14. SIMPLE THRESHOLD DETECTION SCHEME

The outputs from a log amplifier following the DDL in a compressive receiver are many short pulses. There are also side lobes associated with these pulses.

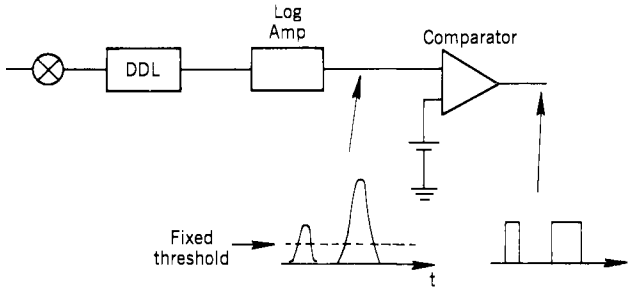


Figure 8.28. Simple detection of a compressive receiver.

To measure the frequency of the input signal, one must measure the center of the main pulse but neglect the side lobes. This is quite similar to the channelized receiver where one input signal will generate outputs from many adjacent filters. The frequency is determined by measuring the filter with the strongest outputs and neglecting the outputs from adjacent filters. In a compressive receiver, there are several approaches to determine the positions of the main lobes while neglecting the side lobes. First, let us discuss some simple detection schemes.

The simplest approach is to compare the output from the DDL with a fixed threshold, as shown in Figure 8.28. When a compressed pulse breaks the threshold, it will be declared a legible output. There are some obvious shortcomings with this approach. First, the amplitude of the output pulse changes with the input signal level. When the signal is strong, the base of the pulse becomes wider, and it is difficult to determine the center of the pulse. Second, this detection scheme may not differentiate the main lobe from side lobes, especially when the side lobes are close to the threshold.

Another approach is to compare the output pulse with many comparators with different thresholds, as shown in Figure 8.29. In this figure, the weak

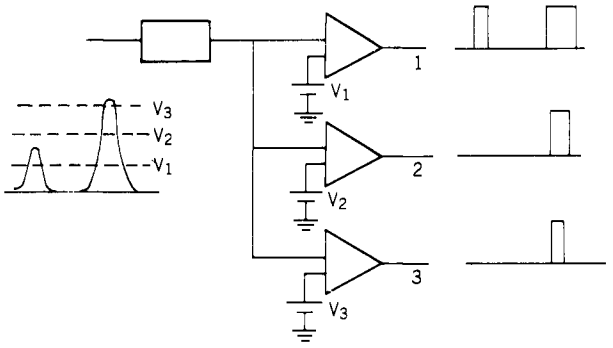


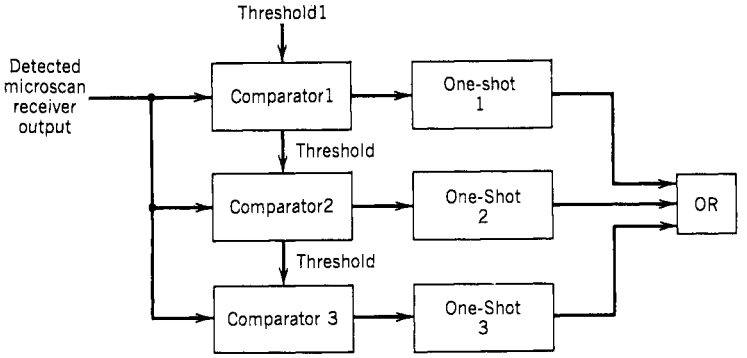
Figure 8.29. Multiple-threshold detection of a compressive receiver.

signal only crosses the lower threshold V_1 and generates a narrow pulse, whereas the strong signal crosses all three thresholds and generates a narrow pulse from the output of comparator 3. The center of the pulse can be determined by the narrowest pulse generated by the comparator with the highest threshold. However, the side lobes may still be detected as a signal through the lower threshold.

From the above discussion, it is obvious that the simple threshold detection scheme cannot effectively avoid the side lobes. Therefore, a filter with proper weighting should be added to suppress the amplitude of the side lobes. The following schemes are designed to suppress the side lobes when the weighting filter cannot fully reduce them.

8.15. CENTROID DETECTION THROUGH CONSECUTIVE COMPARATOR (37)

The concept of the peak detector is quite similar to the multiple-threshold detection circuit discussed in the above section. In this circuit, one-shot logic rates are used to suppress side lobes. The basic circuit is shown in Figure 8.30. The video output pulses from the log amplifier are applied to a number of high-speed comparators operating in parallel. Following each comparator, there is a one-shot logic gate. The outputs of these one shots are ORed together. The thresholds of the comparators are set at different levels. The differences between any two adjacent comparators are approximately equal. When a threshold is exceeded, the comparator output goes high and triggers the one-shot circuit that follows. The output pulse generated by the one shot is of fixed duration. The duration is set in such a way that it is longer than the period of a side lobe but shorter than the period of a main lobe. If a main lobe crosses the threshold, the center of the pulse will be determined by the trailing edge of the OR gate, as shown in Figure 8.30*b*. If a side lobe exceeds certain thresholds, the output from the one shot will overlap the one-shot output of the following comparator



(*a*)

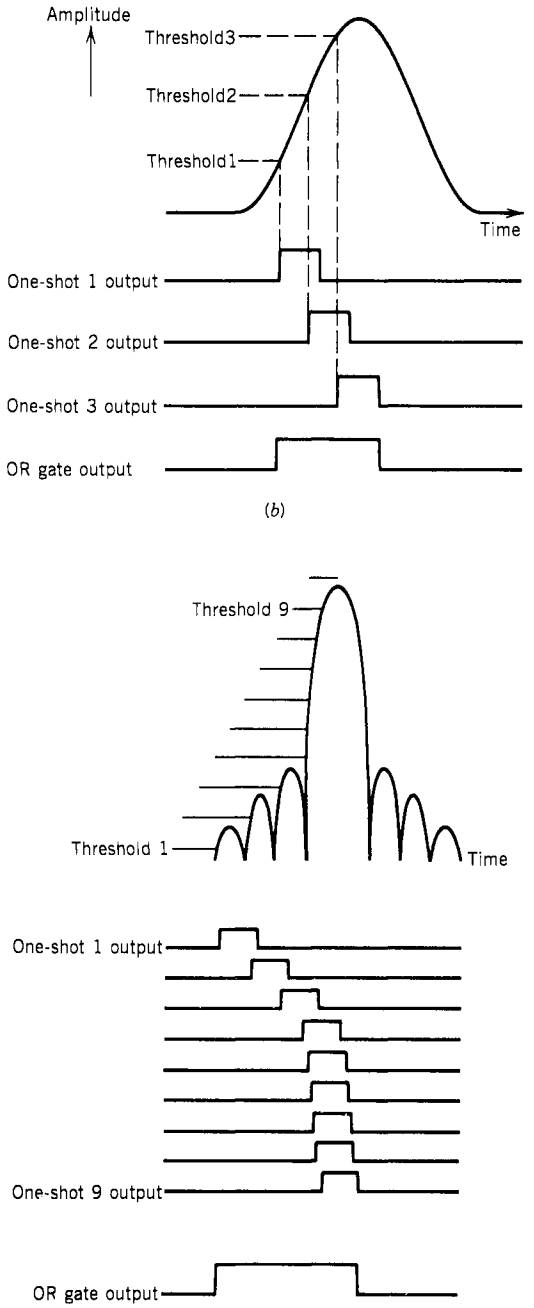


Figure 8.30. Practical peak detection scheme: (*a*) circuit diagram; (*b*) timing diagram; (*c*) compressed pulse output.

and will not generate a trailing edge at the output of the OR gate. Therefore, the side lobes of a compressed video pulse will be masked by this arrangement. Figure 8.30*c* shows the output of a compressed pulse with side lobes. Although there are only 3 thresholds shown in Figure 8.30*b* and 10 thresholds in Figure 8.30*c*, in an actual design, many thresholds are required. By dividing the expected amplitude range into more levels, the accuracy of the position of a pulse peak can be improved.

When the input signal is weak, the amplitude of the compressed pulse is also small. If the first threshold is barely crossed, the trailing edge of the comparator (also the trailing edge of the OR gate) output may be slightly behind the peak of the pulse. This problem can be solved by dividing the comparators into two groups. The outputs of the odd-numbered comparators are ORed together, and the outputs of the even-numbered comparators are ORed together. The two outputs from the OR gates are ANDed together. The output of the AND gate corresponds to the output of the next highest threshold triggered, which has less dependence on the amplitude level. However, this design with two groups decreases the sensitivity of the receiver by one threshold level because to have an output, the compressed pulse must cross the first threshold in both groups. This is similar to the noise triggering problem discussed in Section 3.5.

Since the amplitude of the compressed pulse is digitized by the comparators, the signal strength of the input signal can be obtained from the outputs of the comparators. No additional amplitude measurement circuits are required.

8.16. PEAK DETECTION THROUGH DELAY, SUMMING, AND COMPARE SCHEME (38)

The concept of the delay and summing scheme (ref. 38) is illustrated in Figure 8.31. In this figure, there are two delay lines, 1 and 2. Each is tapped at many (in this discussion, assume three) places. The individual delay time in the two delay lines can be considered mirror images of each other. In other words, the first delay in line 1 equals the last delay in line 2, the second delay in line 1 equals the second to the last delay in line 2, and so on. Each individual delay time is approximately equal to the output pulse width. In this discussion, the width of the compressed pulse is assumed to be 2 nsec and the delay times of line 1 are 2, 2, 1, and 1 nsec. The delay times of delay line 2 are 1, 1, 2, and 2 nsec. The last three delayed outputs from line 1 are summed together and compared to the sum of the first three delayed outputs of line 2. Since operational amplifier A works as an inverter and B as a summing device, the output of B is the difference of inputs of delay line 1 and delay line 2. At the center of the output pulse, the sum of inputs from line 1 equals the sum of inputs from line 2. Just before the center of the output pulse, one side is greater than the other, and just after the center of the pulse, the reverse is true. Therefore, at the center of the pulse, the output from B is changed from one polarity to another. In other words, there is a zero crossing at the center of the main pulse. However, when the side lobes

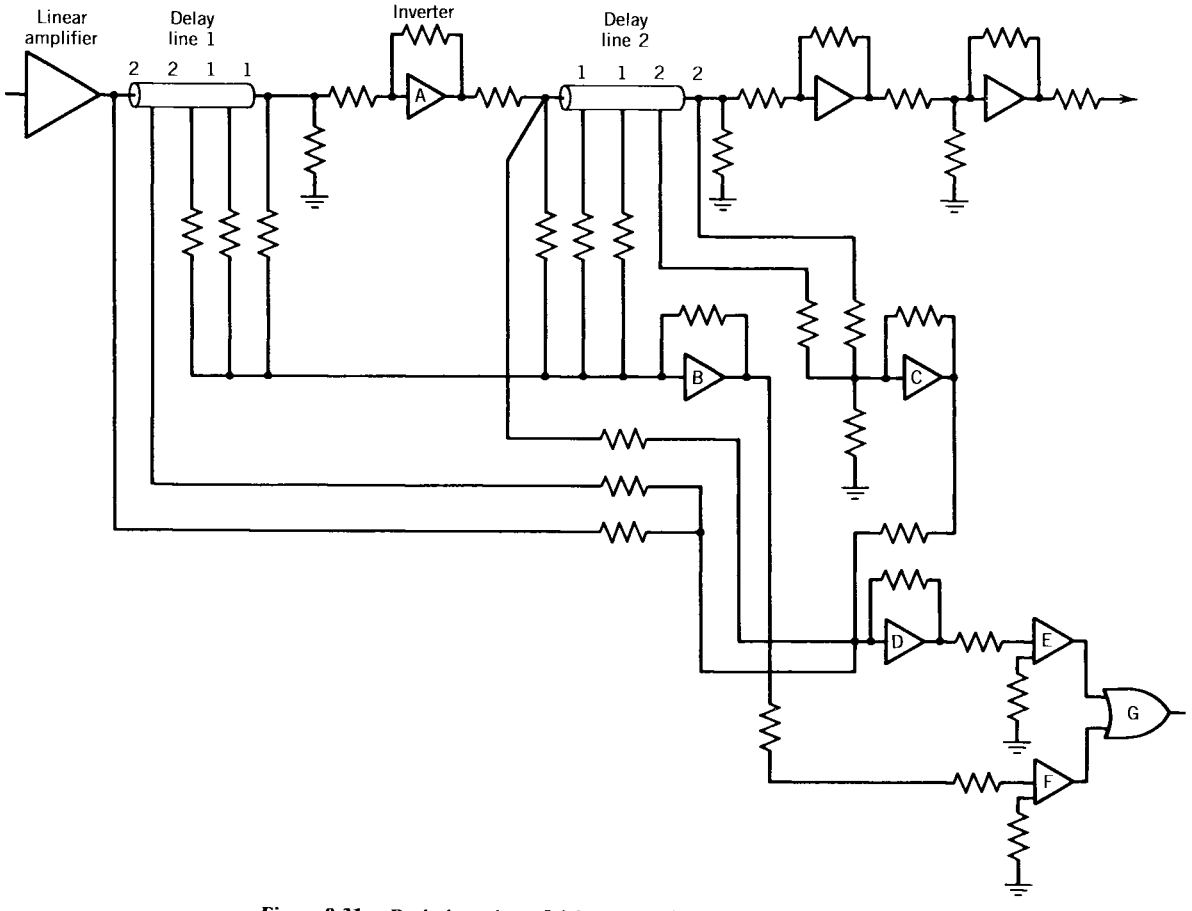


Figure 8.31. Peak detection of delay, summing and comparison scheme.

pass through the two delay lines, there are also zero crossings. The output of B is shown in Figure 8.32 as "centroid" detector. The outputs in Figure 8.32 are simulated results. From this output, it is difficult to determine which zero crossing is at the center of the main lobe of the output pulse. To determine the correct zero crossing, additional circuits are required.

The additional circuits to separate the center zero crossing of the main lobe from the other zero crossings are also included in Figure 8.31. The operational amplifier D is used to perform this function. The inputs to D can be divided into three parts. The first part of the input is from the first two taps of delay line 1. The second part is from the last two taps of delay line 2, which passes inverter C. Since the second part of the input passes two inverters, the output from inverter C has the same polarity as the first part of the input. The third part of the input to D is from the output of inverter A, which represents the compressed output pulse. The third part of the input has opposite polarity from the first two parts. The sum of the first two inputs is shown in Figure 8.32, the "mean level." The third input is marked as the output pulse. The output from operational amplifier D is the difference of these two curves, labeled "enable." This output is used with the centroid detector output to determine the position of the main lobe and neglects the zero crossings of the side lobes.

The outputs from operational amplifiers B and D are analog information.

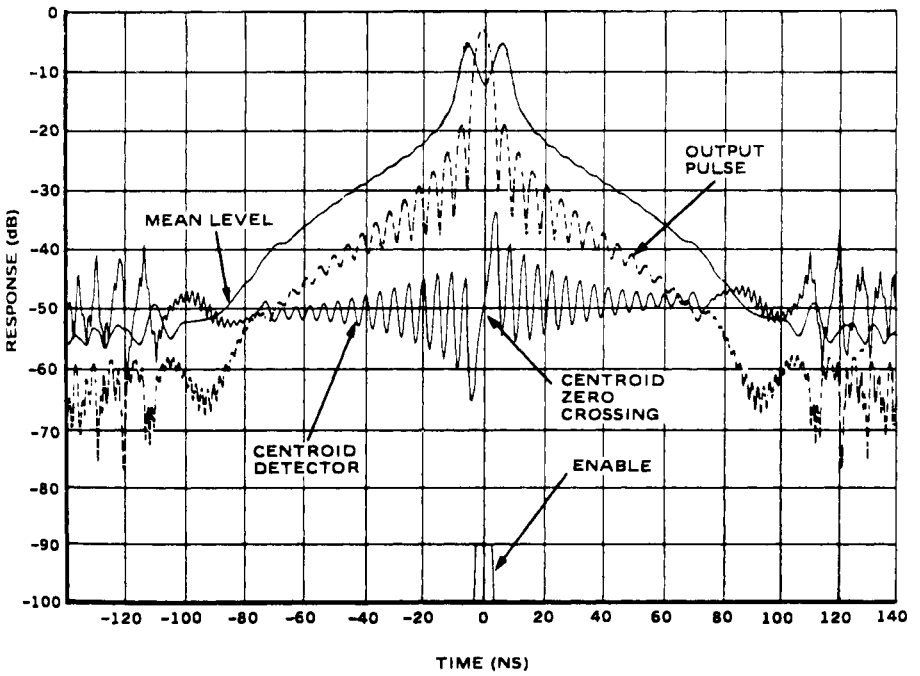


Figure 8.32. Output from peak detection circuit. (Courtesy of Hughes Aircraft Co.)

To convert them to digital information, two comparators are used following B and D. The output from comparator E is marked as an enable pulse at the bottom of Figure 8.32. The zero crossing within the enable pulse represents the center of the pulse. The leading edges of the output pulse from comparator F are the zero crossings. The outputs from B and D are ANDed together through gate G. The leading edge of the output pulse from gate G represents the center of the compressed pulse.

Although this approach provides the center of the compressed pulse, it does not measure its amplitude. The amplitude of the input signal must be measured through some other circuits. Since the center of the compressed pulse is determined by this approach, the digitizing window for the amplitude measurement circuit is readily determined.

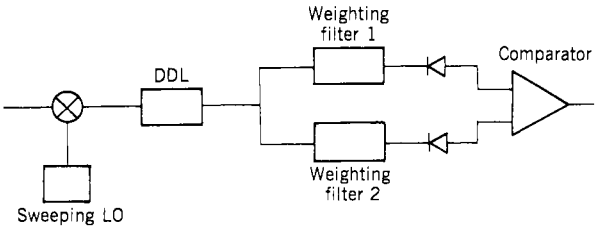
8.17. SIDE LOBE SUPPRESSION THROUGH DUAL-DETECTION SCHEME

Another way to differentiate the main lobe from the side lobes of the outputs of a DDL is through a dual-detection scheme. This approach can be considered the analogy of the dual detection in a channelized receiver as discussed in Section 7.14. The difference is that in Section 7.14 the comparison is made in frequency domain, and here the comparison is made in time domain. In this approach, the weighting filter is placed after the DDL. Since both the DDL and the weighting filter are linear components, their order in a compressive receiver does not affect the results of the output.

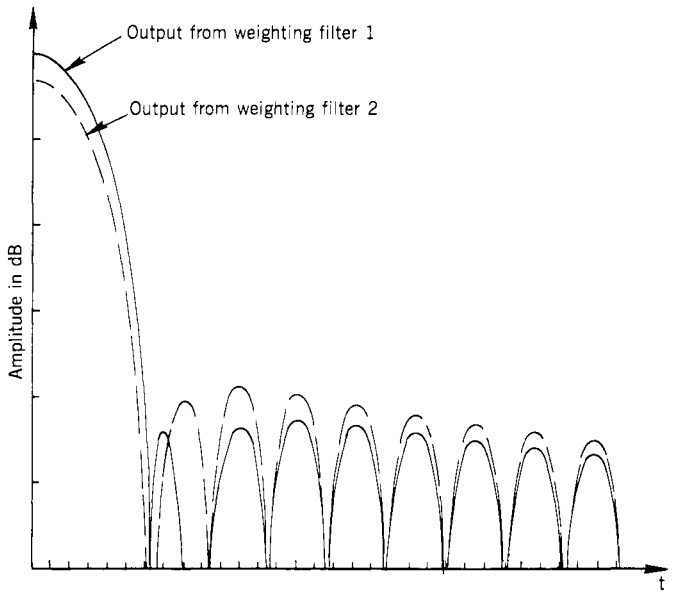
Instead of using one weighting filter, the output from the DDL is divided into two parallel paths, and two parallel weighting filters with different weighting factors are used, as shown in Figure 8.33. The outputs from the two weighting filters are detected and fed into a comparator. If the weightings are properly chosen, the output from the comparator should contain only the main lobe, and the side lobes will be suppressed.

In this discussion, assume that both weighting filters are cosine square on pedestal [Eq. (8.22)] with different k values. For instance, $k = 0.06$ for filter 1 and $k = 0.14$ for filter 2. The outputs from the two weighting filters are shown in Figure 8.33*b*. The output from the comparator is shown in Figure 8.33*c*. Theoretically, this approach does not generate any outputs from the response of the side lobes, which should simplify the processing circuits that follow. The center of the output pulse should be the center of the compressed pulse.

The disadvantages of this approach are that two weighting filters and detectors (maybe log amplifiers) are required and the width of the output pulse is wider than the main lobe. Therefore, the frequency accuracy of the receiver may be degraded. The center of the pulse must still be determined. If the tracking between the two channels is not properly maintained, the side lobes may cause the comparator to generate outputs that will cause the receiver to produce spurious outputs.



(*a*)



(*b*)

(*c*)

Figure 8.33. Dual-detection scheme for compressive receiver: (*a*) basic arrangement; (*b*) filter outputs; (*c*) comparator output.

8.18. PARTIALLY INTERCEPTED PULSES AND OTHER SIGNAL PARAMETER MEASUREMENTS BY A COMPRESSIVE RECEIVER

Besides the frequency information, a compressive receiver can be used to measure other parameters. If these parameters are measured after the frequency separation, the quality of these parameters will be improved because the possibility of contamination from simultaneous signals is greatly reduced. Before discussing these measurements, however, the problem of partially intercepted pulses in a microscan receiver will be discussed.

Under normal operating conditions, an input pulse will be partially intercepted by the receiver at both the leading and trailing edges of the pulse. These partially intercepted pulses may cause inaccuracy in the PA measurement. If the receiver is not properly designed, even the frequency measurement may be disturbed by the partially intercepted pulses and produce spurious signals. Theoretically, the center of the compressed pulse generated by a partially filled compressive delay line will be the same as that of a fully filled line. Thus, the frequency measurement of the pulse will not be disturbed. However, in a practical DDL, the center of the compressed pulse of a partially filled line and a fully filled line may be different. This will cause frequency measurement error.

The video output from a compressive receiver is displayed scan by scan in Figure 8.34. In Figure 8.34*a*, the frequency shifts caused by the partially filled DDL at the pulse leading and trailing edges are not far from the fully filled line. Under this condition, the partially filled output can be detected by amplitude comparison of two pulses of the same frequency from two consecutive scans. The outputs from all the scans will be treated as one input signal. In Figure 8.34*b*, the compressed pulse positions from a partially filled DDL is far from that of a filled DDL. Then the leading and trailing edges of the input pulse may be

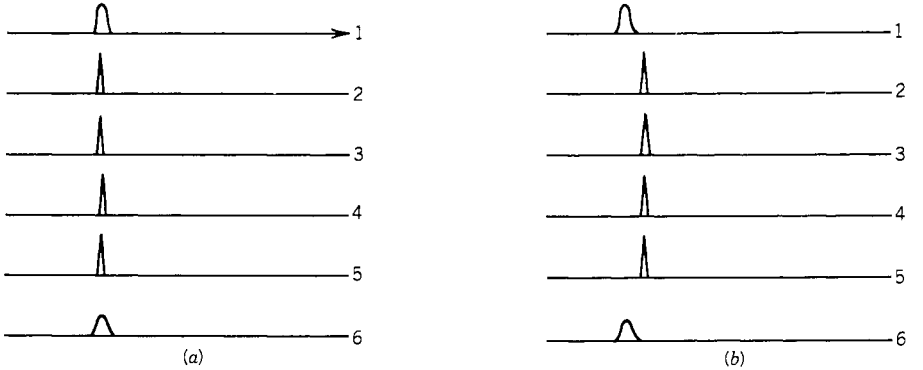


Figure 8.34. Video outputs from compressive receiver: Partially filled line has (*a*) no frequency deviation and (*b*) large frequency deviation.

treated as short input pulses with different frequencies. Therefore, this effect will cause the receiver to generate spurious signals. There are two approaches to avoid this problem. First, the characteristics of the DDL can be improved so that the pulse positions from a partially and fully filled DDL are very close. Second, the frequency limits can be widened on comparing outputs from consecutive scan outputs. The first approach is preferred because the second approach may miss signals close in frequency.

Once the video outputs from partially filled DDLs are properly processed, the measurements of parameters other than frequency are straightforward. However, it is desirable to obtain these parameters from the output of a fully filled DDL. Especially for the PA measurement, the output obtained from a partially filled DDL will be lower than the real value. If an input pulse is very short and cannot fill the compressive DDL, the measurement accuracy on this signal will be degraded.

AOA and PA Measurements. The PA of the input signal can be measured from the video output of the compressed pulse. The measurement circuits must have enough video bandwidth to process the short pulses from the DDL. Since at the output of the DDL the amplitude and the phase of the input signals are both maintained, the AOA information can be obtained by either amplitude or phase comparison schemes. The PA information from different quadrants can be compared to obtain AOA. If the AOA is measured by phase comparison, the phase information can only be obtained before video detection. To improve the phase measurement accuracy, the compressed pulse is often amplified to saturation. However, the PA information on the compressed pulse will be lost.

PWand TOA Measurements. In a microscan receiver, it is common practice to generate TOA from the first compressed pulse and time of departure (TOD) from the last compressed pulse even if they are outputs of a partially filled DDL. The PW is determined by subtracting the TOA from the TOD. At the output of the compressed DDL, the position of the compressed pulses within one scan depends only on the input signal frequency and not on the TOA of the signals. Therefore, in one scan, one cannot determine which pulse arrives first. In other words, the TOA and PW resolution is the same as the scan time. A compressive receiver with fine-frequency resolution will require a DDL with long delay time, which in turn will produce poor TOA and PW resolution.

8.19. DIGITIZING CIRCUIT CONSIDERATIONS

The compressed RF pulse from the DDL is converted to video pulses through a log amplifier, and this video information must be converted to digital information to feed a digital processor. The detailed circuit can be rather complicated and is beyond the scope of this book. Generally speaking, with the state-of-the-

art technology, the size of the digitizing circuit is larger than the microwave section of the compressive receiver, as shown in Figure 8.1. Some general discussions of how the video pulses should be processed will be briefly presented in this section.

Since a long pulse of fixed frequency can be intercepted by a compressive receiver in many consecutive scans, it is by no means desirable to report the signal from each scan because there is too much redundant information. In general, only three scans on one intercepted input pulse are needed (see Fig. 8.34). The beginning scan (1) and the last scans (6) are used to provide the TOA and PW information of the input pulse. One additional scan (usually scan 2) may be required to provide the frequency and PA information if the beginning scan does not fill the compressive DDL. One way to implement such an approach is to use a shift register whose clock rate is synchronized with the receiver scan. Intercepted signals are loaded into the shift register. The location of the information in the shift register represents the signal frequency. On the next scan, the receiver output is compared with the content of the shift register within a certain tolerance. One of three conditions can occur. First, if the shift register does not contain the information, the output from the receiver is considered a new signal. Second, if the shift register contains the same frequency information (within the predetermined tolerance), the receiver output is considered a continuation of the same signal. In Figure 8.34, the signals in scans 2–5 are the continuation signals. The output of scan 2 is the first continuation signal. Third, if the receiver does not have an output, but the shift register contains the information, the information in the shift register is considered an old signal. The output of a scan can contain new signals for some frequencies but continuation or old signals for other frequencies.

The amplitude of a new signal must be compared with the first continuation signal. If both have the same amplitudes, it means that the new signal fills the compressive DDL or nearly fills the DDL. Under this condition, only the new and old signals are needed to be stored in a parameter register because they contain all the necessary information to describe a pulse. The frequency, PA, and TOA are obtained from the new signals, and the TOD is obtained from the old signals. The PW of the signal can be calculated from the difference TOD − TOA. If the amplitude of the new signal is less than the first continuation signal, the first continuation signal is used to generate the PA and frequency information. If the frequency in the first continuation signal is different from the new signal but within the comparison tolerance, the frequency information is often generated from the first continuation signal because the first continuation fills the DDL and provides more accurate frequency information. If the input signal is a short pulse and is intercepted by only two consecutive scans, the PA measured may be less than the true value. The minimum PW the receiver can process with fidelity should at least fill the compressed DDL once.

The information generated from the above approach is stored in a parameter register. The size of the parameter register should match the number of simultaneous signals to be processed. If more signals than the number of the param-

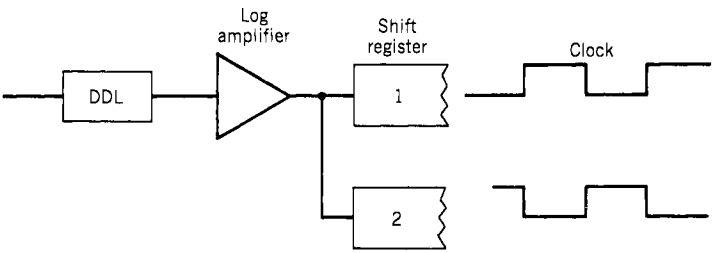


Figure 8.35. Two-phase clock to increase logic speed.

eter registers are intercepted during a scan, some of the signals will not be stored in the register. As a result, these signals will be lost.

As mentioned earlier in this section, the clock of the shift register must match the output pulse rate from the DDL. In some compressive receivers, the frequency bandwidth of the DDL is higher than that of the shift register. In other words, the clock of the shift register cannot match the output of the DDL. One way to increase the logic speed of the shift register is to use a multiple-phase clock, as shown in Figure 8.35. In this figure, there are two parallel shift registers. The clock used to drive the two shift registers are 180° out of phase. The outputs from the DDL are fed into the two shift registers in an alternate way. If each shift register is operated at 500 MHz, the equivalent operating speed of this arrangement is 1 GHz. The disadvantage of this approach is that more shift registers are required.

8.20. TRADING OF TIME–BANDWIDTH PRODUCT FOR WIDER COMPRESSED PULSE

The compressed pulse from the DDL of a compressive receiver is usually very narrow. The video circuit following the DDL must have enough bandwidth to faithfully process the pulses. Sometimes, it is desirable to widen the compressed PW for easier signal processing. However, the compressed PW is inversely proportional to the IF bandwidth, which means that either the receiver has narrower input bandwidth or the receiver has lower POI. Both results are undesirable. There is an approach to widen the compressed pulse and keep the same IF bandwidth. This approach uses a DDL with a dispersive delay time longer than the desired value. Instead of generating the high-frequency resolution limited by the delay time of the DDL, coarse frequency is produced. As a result, the design does not take the full advantages of the bandwidth of the DDL. The equivalent effect of this design is to increase the k value in Eq. (8.28). This approach can be further explained by the following example.

If the desired B_I is 1000 MHz, and the frequency resolution is 10 MHz, a DDL with $B_I = 1000$ MHz and $T = 100$ nsec is needed to satisfy the requirements. The compressed PW is 1 nsec. However, if the dispersive delay time of the

DDL is increased to 1000 nsec and the compressed pulse is still kept at 1 nsec, the frequency resolution will be increased to 1 MHz. In the latter case, if the video bandwidth is narrow, which will extend the compressed PW to 10 nsec, the resolution of the receiver will decrease to 10 MHz. A more effective way to widen the compressed PW is to use a weighting filter in series with the DDL.

The main disadvantages of this approach are (1) a DDL with a high time–bandwidth product is difficult to fabricate and (2) a short input pulse will not be able to fill the DDL and, as a result, the PA information produced on a short input pulse will be less than the true value. In the above example, theoretically, a 100-nsec pulse can fill the 100-nsec DDL, but it can only fill 10% of the 1000-nsec DDL. The TOA resolution of a compressive receiver with this design is degraded. In the above discussion, the TOA resolution is degraded from 100 to 1000 nsec.

8.21. GENERALIZED COMPRESSIVE SCHEMES (25, 40–42)

The arrangement discussed in Section 8.2 can be considered a special case. There are two general arrangements that can be used to perform Fourier transforms. These two approaches are referred to as the multiply–convolve–multiply (MCM) and convolve–multiply–convolve (CMC). A DDL will perform the convolution to the input signal in the time domain, and a mixer will perform a multiplication to the input signal in the time domain. Figure 8.36*a* shows an MCM arrangement, and Figure 8.36*b* shows a CMC arrangement. The mathematical expression of the MCM approach can be derived as follows.

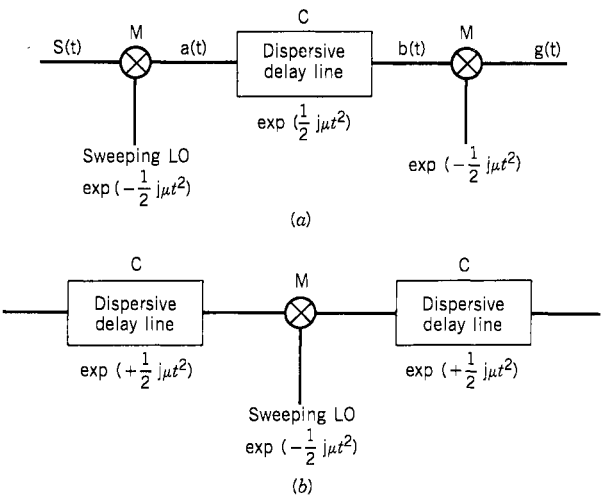


Figure 8.36. Common compressive schemes: (*a*) MCM configuration; (*b*) CMC configuration.

In Figure 8.36*a*, if the input signal is $S(t)$, the output of the first mixer is

$$a(t) = S(t)\exp(-\tfrac{1}{2}j\mu t^2) \tag{8.42}$$

where μ is the scan rate defined in Eq. (8.2). The output from the delay line is

$$\begin{aligned} b(t) &= \int_{-\infty}^{\infty} a(\tau)h(t-\tau)\,d\tau \\ &= \int_{-\infty}^{\infty} S(\tau)\exp(-\tfrac{1}{2}j\mu\tau^2)\exp[\tfrac{1}{2}j\mu(t-\tau)^2]\,d\tau \end{aligned} \tag{8.43}$$

where $h(t)$ is the transfer function of the DDL.

The output from the second mixer can be written as

$$g(t) = b(t)\exp(-\tfrac{1}{2}j\mu t^2) \tag{8.44}$$

Substituting Eq. (8.43) into (8.44) yields

$$\begin{aligned} g(t) &= \int_{-\infty}^{\infty} S(\tau)\exp(-\tfrac{1}{2}j\mu\tau^2)\exp[\tfrac{1}{2}j\mu(t-\tau)^2]\exp(-\tfrac{1}{2}j\mu t^2)\,d\tau \\ &= \int_{-\infty}^{\infty} S(\tau)\exp(-\tfrac{1}{2}j\mu t\tau 2)\,d\tau \end{aligned} \tag{8.45}$$

If $\omega = \mu t$, the result in Eq. (8.45) indicates that the output from the second mixer is the Fourier transform of the input signal $S(t)$. The CMC approach can produce a similar result.

Although the above analysis shows that in both approaches the output is the Fourier transform of the input signal, components with different bandwidth can achieve different results. Let us use CW as the input signals to discuss their applications as compressive receivers. For the CMC configuration, the first DDL will provide different delay times to the CW signals according to their frequencies. Since the input signals are CW, this delay will not provide any significant effects on the input signals. The mixer and the second DDL will perform as a conventional compressive receiver. Therefore, this configuration is not widely adapted in a compressive design.

For the MCM configuration, the first two components perform the same function as an ordinary compressive receiver. The second mixer will shift the center frequency of output pulses. Let us use an example to discuss this approach. In this example, 100% POI can be achieved without the interlace scan scheme discussed in Section 8.12.

Figure 8.37 shows a compressive receiver with the IF bandwidth B_I equal to twice the RF input bandwidth B_R. The LO to the first mixer has a bandwidth

equal to $B_{I/2}$. Figure 8.37 shows the time and frequency relation of this compressive receiver. Since the first mixer output is always in the passband of the DDL, the receiver will have a 100% POI.

From Figure 8.37*c*, one should notice that all three signals enter the DDL at the same time, which is different from the cases discussed in Section 8.2. Since the three frequencies in Figure 8.37*c* cover different frequency ranges, they will leave the DDL at different times, as shown in Figure 8.37*d*. Therefore, by measuring the positions of the compressed pulses, the frequency of the input signals can be determined. One should notice that the center frequency of each compressed pulse is different because each input signal uses different portions of the DDL. In this special case, the center frequency of f_a is higher than the center frequencies of f_b and f_c. If an RF log amplifier is used after the DDL to convert

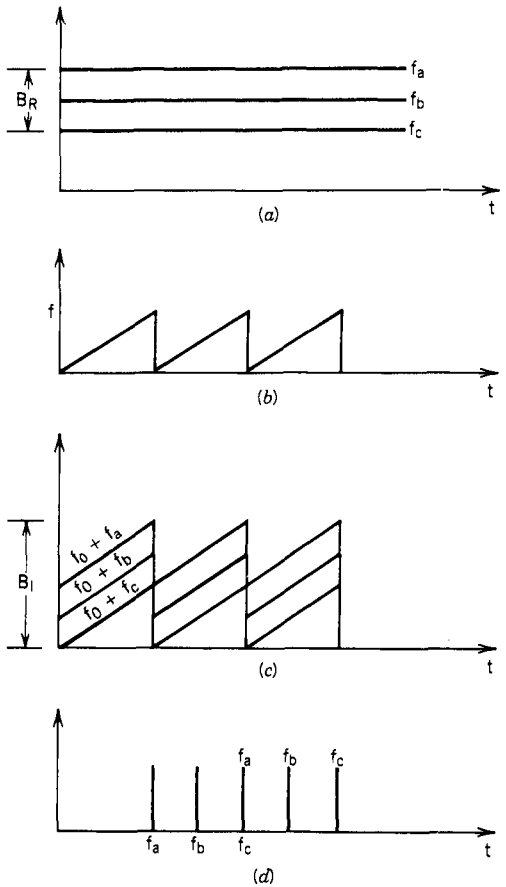


Figure 8.37. Time and frequency relation of compressive receiver with MCM configuration and $B_I = 2B_R$: (*a*) input signals; (*b*) sweeping LO; (*c*) first mixer output intercepted by DDL; (*d*) output pulses from DDL.

the compressed pulses to video signals, the bandwidth of the log amplifier must be wide enough to cover the center frequency of all the compressed pulses. This requirement will complicate the log amplifier design.

If the LO of the second mixer is also a chirp signal of the same sweeping rate as the first mixer, the center frequencies of the compressed pulses can be converted to the same frequency. An RF log amplifier with a bandwidth to process the compressed pulse can then be used to convert the compressed pulses to video signals. The frequency resolution of this kind of compressive receiver is determined by part of the dispersive delay time. In this example, the frequency resolution is equal to half of the dispersive delay time of the DDL. The compressed output pulse rate is equal to the RF bandwidth, which is less than the bandwidth of the DDL.

8.22. SUMMARY

The outputs of a compressive receiver are in series, in contrast to a channelized receiver where all the outputs are in parallel. Therefore, potentially, a compressive receiver will need less hardware, but high-speed logic is required to process the outputs of the receiver. A compressive receiver has the following general properties.

The bandwidth is moderately high. Presently, it is limited by the speed of the processing logic devices. The frequency resolution is limited by the delay time of the DDL and the video bandwidth to process the compressed pulse.

The sensitivity of the receiver is determined by its noise figure and the frequency resolution. Increasing the time–bandwidth product of the DDL only widens the bandwidth of the receiver, which in turn improves the POI. The sensitivity of a compressive receiver can be improved through longer dispersive delay times, which is related to the frequency resolution of the receiver. However, the time–bandwidth product does not improve the sensitivity of the receiver beyond that of a superhet receiver.

The dynamic range of the receiver is limited by the side lobes generated at the outputs of the DDL. Properly designed weighting filters can reduce the side lobes.

A compressive receiver processes a CW signal as a pulsed signal, and no special circuits are required. The receiver will intercept the CW signal and divide the signal into many scans and process them as pulsed signals. Therefore, this kind of receiver should be very useful to intercept communication signals.

The compressive receiver is basically still in the developing stage. The major difficulties are in the digitizing circuits that convert the video pulses to digital information.

REFERENCES

1. W. D. White, Signal translation apparatus utilizing despersive network and the like, for panoramic reception, amplitude-controlling frequency response, signal frequency gating, frequency-time domain conversion, etc. U.S. Patent 2,954,465, September 27, 1960.

2. W. R. Kincheloe, The measurement of frequency with scanning spectrum analyzers, Technical Report 557-2, Systems Techniques Laboratory, Stanford Electronic Labs, Stanford University, Air Force Contract AF30(602)-2398, October 1962.
3. W. R. White and I. M. Saffitz, Compressive receivers, Airborne Instruments Lab. Div. of Cutler-Hammer, Inc., *Topics in Electronics*, Vol. 3, 1962.
4. J. R. Klauder, A. C. Price, S. Darlington, and W. J. Albersheim, The theory and design of chirp radars, *Bell System Technical J.*, **39** (July 1960).
5. R. G. Sweet and W. R. Kincheloe, Jr., A real-time scanning spectrum analyzer using a tapped sonic-delay line filter, Technical Report 1967-1, SU-SEL-64-038, Systems Techniques Laboratory, Stanford Electronic Lab, Stanford University, June 1964.
6. R. G. Sweet, H. Hewith, and W. R. Kincheloe, Jr., A high resolution rapid scan receiver and signal recorder for communications frequencies, Technical Report 1967-2, SU-SEL-66-111, Systems Techniques Lab. Stanford Electronics Lab, Stanford University, December 1966.
7. T. Harper, New trends in EW receivers, *Countermeasures*, 33 (December/January, 1975/1976).
8. C. B. Hoffman and A. R. Baron, Wide-band ESM receiving systems, *Microwave J.*, part 1, **24**, 24 (September 1980); part 2, **25**, 57 (February 1981).
9. M. Bernfeld, C. E. Cook, J. Paolillo, and C. A. Palmieri, Matched filtering, pulse compression and wave form design, *Microwave J.*, part 1, **7**, 57–64 (October 1964); part 2, **7**, 81–90 (November 1964); part 3, **7**, 70–76 (December 1964); part 4, **8**, 73–81 (January 1965).
10. C. E. Cook and M. Bernfeld, *Radar Signals and Introduction to Theory and Applications*, Chapter 7, Academic Press, New York, 1967.
11. D. Nichols, Transform receivers for ECM applications, *Microwave J.*, **25**, 113 (October 1982).
12. J. C. Adam, J. H. Collins, and J. M. Owens, Magnetostatic wave group-delay equilizer, *Electronics Letters*, **9**, 557–558, (November 1973).
13. T. A. Martin, R. B. Ducharme, and R. L. Sawin, A microwave surface acoustic wave dispersive IMCOM, AFCRL-TR-74-0622, Air Force Cambridge Research Laboratories, December 1974.
14. J. D. Adam, M. R. Daniel, and C. E. Nothnick, MSW dispersive delay lines in a compressive receiver, *Proc. IEEE Ultrasonic Symposium*, pp. 533–536, (82CH1823-4), 1982.
15. F. R. Morgenthaler, Apparatus for processing electromagnetic wave energy, U.S. Patent 3,895,324 July 15, 1975.
16. H. W. Fuller, Broadband MW pulse compression using crimped coax delay lines, *Microwave J.*, **23**, 52 (April 1980).
17. H. S. Hewitt, A computer designed 720 to 1 microwave compression filter, *IEEE Trans. Microwave Theory Techniques*, **MTT-15**, 687–694 (1967).
18. A. E. Dunn, A pulse compression filter employing a microwave helix, Stanford Electronics Lab, Tech Report 557-1, Stanford University, October, 1960.
19. A. E. Dunn, Realization of microwave pulse compression filters by means of folded-taped meander lines, Stanford Electronics Lab, SEL-62-113 (TR-557-3), Stanford University, October 1962.
20. W. R. Smith, H. Gerard, and P. B. Snow, Highly dispersive acoustic filter study, Research and Development Technical Report ECOM-0046-F, U.S. Army Electronics Command, Fort Monmouth, N.J. September 1973.
21. B. R. Potter, Surface acoustic wave slanted correlators for linear FM pulse compressor, *IEEE MTT International Microwave Symposium Digest.* (77CH1219-5), pp. 318–320, June 1977.
22. W. R. Smith, H. M. Gerard, and W. R. Jones, Analysis and design of dispersive interdigital surface wave transducers, *IEEE Trans. Microwave Theory Techniques*, **MTT-20**, 458–471 (1972).
23. H. Matthews, *Surface Wave Filters*, Wiley, New York, 1977.
24. B. J. Hunsinger and S. Datta, Termination of surface acoustic wave velocity and impedance differences between metal strips and free surface regions of metallic gratings, Interim Report RADC-TR-81-4, Rome Air Development Center, Griffiss Air Force Base, February 1981.
25. M. A. Jack, P. M. Grant, and J. H. Collins, The theory, design and applications of surface acoustic wave Fourier-transform processors, *Proc. IEEE*, **68**, 450–468 (1980).
26. E. G. S. Paige, Dispersive filter: their design and application to pulse compression and temporal transformations, *IEEE Conference Publication*, 109, 167–180, September 1973.

27. R. C. Wissiamson, Properties and applications of reflective array devices, *Proc. IEEE*, **64**, 702–710 (1976).
28. J. M. Alsup, Prime transform SAW device, *Proc. IEEE Ultrasonics Symposium* (75CH0994-4SU), pp. 377–380, 1975.
29. H. W. Gerard, O. W. Otto, and R. D. Weglein, Development of a broadband reflective array 10000:1 pulse compression filter, *Proc. IEEE Ultrasonics Symposium* (74CH0896-1SU), pp. 192–201, 1974.
30. C. O. Newton and E. G. S. Paige, Surface acoustic wave dispersive filter with variable linear, frequency time slope, *Proc. IEEE Ultrasonics Symposium* (76 CH1120-5SU), pp. 424–427, 1976.
31. J. B. Harrington and R. B. Nelson, Compressive intercept receiver uses SAW devices, *Microwave J.*, **17**, 57 (September 1974).
32. R. N. Buswell, Voltage controlled oscillators in modern ECM systems, Watkins Johnson Tech Notes, **1**(6) (November/December 1974).
33. J. B. Fuller, Frequency-agile local oscillator, *Microwave J.*, 31 (May 1974).
34. R. Beach, High speed linear oscillators, *Microwave J.*, **21**, 59 (December 1978).
35. R. E. Winkelman and H. F. Schmid, SRI International, private communication.
36. V. H. Estrick and D. Fawcett, Hughes Aircraft Co., private communication.
37. B. Graig and H. F. Schmid, SRI International, private communication.
38. R. Gaudagnolo and D. Fawcett, Hughes Aircraft Co., private communication.
39. J. B. Y. Tsui, Dual detection scheme for compressive receivers, U.S. Patent 4,200,840, 29 April 1980.
40. C. Atzeni, G. Manes, and L. Masotti, Programmable signal processing by analogue chirp transformation using SAW devices, *IEEE Ultrasonics Symposium Proc.*, (75 CH0994-5SU), pp. 371–376, 1975.
41. C. Atzeni, G. F. Manes, and I. Masotti, Real time network analysis using dual analogue chirp transform, *Electronics Letters*, **12**, 248–249 (1976).
42. M. A. Jack, G. F. Manes, P. M. Grant, C. Atzeni, L. Masotti, and J. H. Collins, Real time network analysers based on SAW dirp transform processors, *IEEE Ultrasonics Symposium Proc.*, (76CH1120-5SU), pp. 376–381, 1976.

Chapter 9

Bragg Cell Receivers (Optical Processors)

9.1. INTRODUCTION (1–9)

Optical processing, in general, involves the use of optical techniques to analyze signals and perform special signal-processing functions. A Bragg cell receiver is one kind of optical processor that, in essence, performs a Fourier transform on input signals to obtain the frequency information. In this chapter, the discussion will be limited to the operation and performance of Bragg cell receivers.

Although the principles of optical processing have been known for many years, it was only recently that the requisite technologies developed sufficiently to make Bragg cell receivers feasible. Of particular importance is the development of the laser, since high brightness in the optical wave is essential to achieve satisfactory performance. In addition, the spatial coherence of lasers can be used to generate special detection schemes to improve the dynamic range of the Bragg cell receiver. Other significant technological advances over the last decade include development of Bragg cells with large time–bandwidth products and availability of one- and two-dimensional photodetector arrays with a large number of detectors. While much work remains to be done, especially in the area of improving the dynamic range of the receiver and the digitizing circuits that convert the video signals to digital information, these recent developments in optical technology have encouraged the exploration of using Bragg cells in electronic warfare (EW) receivers.

The most attractive aspect of using a Bragg cell to build a microwave receiver is its potential for extremely small size and low cost. A Bragg cell is equivalent to the filter bank in a channelized receiver. A Bragg cell can achieve a time–bandwidth product of 1000, which is equivalent to a filter bank with 1000

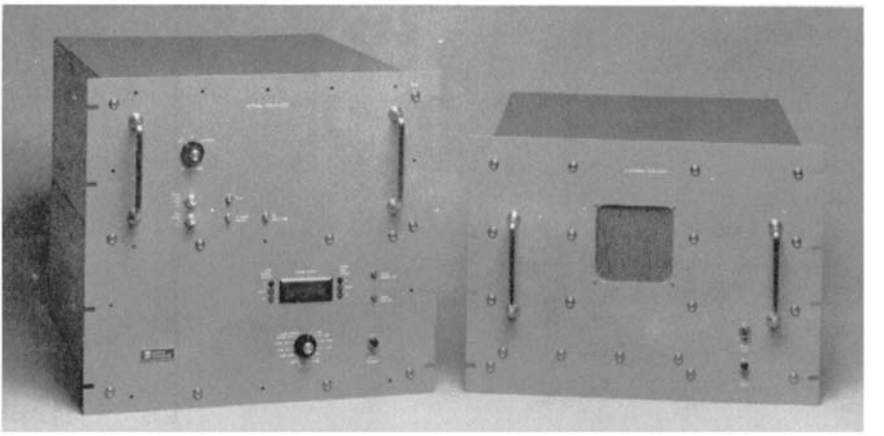


Figure 9.1. Bragg cell receiver. (Fabricated by Applied Technology, Division of Litton, Courtesy of Avionics Laboratory, AFWAL.)

parallel channels. Furthermore, the development of integrated optical circuits (IOCs) in which the laser source, Bragg cell, and optical and detector arrays are fabricated on a single chip would make the optical portion of the receiver extremely small. Figure 9.1 shows a Bragg cell receiver. The optical bench and the photodetectors are in the left unit. The inputs to the left unit are RF signals and the outputs are video signals. The right unit contains the digitizing circuits that convert the video signals to digital information.

The operating principles of Bragg cell receivers will be described along with the various detection schemes employed. The basic designs of an IOC will also be discussed. Finally, an interferometric approach to improve the dynamic range of the Bragg cell receiver will be discussed.

9.2. OPTICAL FOURIER TRANSFORM (10, 11)

Many types of optical signal processors, including Bragg cell receivers, perform their function by spatially modulating the phase or amplitude of an optical beam with the input signal, as shown at the input plane of Figure 9.2. An optical

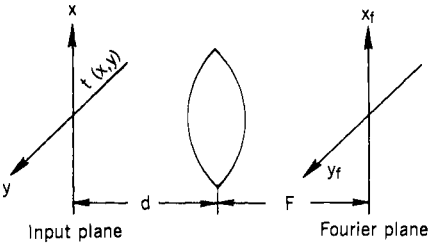


Figure 9.2. Basic optical Fourier transform.

Fourier transform is performed on the spatially modulated optical beam to display the frequency domain characteristics of the signal being processed as a spatial distribution of the light. To illustrate this phenomenon, consider the situation depicted in Figure 9.2. A collimated light beam passes through an input plane with amplitude modulation mask $t(x, y)$ and through a lens located a distance d behind the input plane. The light amplitude distribution in the lens focal plane, or Fourier plane, represents the Fourier transform of $t(x, y)$. This relationship can be expressed as

$$U(x_f, y_f) = \frac{A \exp[j(k/2F)(1 - d/F)(x_f^2 + y_f^2)]}{j\lambda F} \times \iint_{-\infty}^{\infty} t(x, y) \exp\left[-j\frac{2\pi}{\lambda F}(xx_f + yy_f)\right] dx\, dy \tag{9.1}$$

where x, y are coordinates in the input plane and x_f, y_f are coordinates in the Fourier plane. In the above equation, k is the wave number ($= 2\pi/\lambda$), λ is the wavelength of the incident light, d is the distance between the object plane and the lens, F is the distance between the lens and its focal plane, A is input light amplitude, $t(x, y)$ is the modulation of the input plane, and $U(x_f, y_f)$ is the complex optical wave amplitude in the Fourier plane.

In Eq. (9.1), the transformation differs from a true Fourier transform by the phase factor preceding the integral. But when $F = d$, the phase factor disappears and Eq. (9.1) represents the exact Fourier transform. In any case, the optical intensity distribution $I(x_f, y_f) = UU^*$ represents the power spectrum of the modulation $t(x, y)$; that is,

$$I(x_f, y_f) = UU^* = \frac{A^2}{\lambda^2 F^2}\left|\iint_{-\infty}^{\infty} t(x, y) \exp\left[-j\frac{2\pi}{\lambda F}(xx_f + yy_f)\right] dx\, dy\right|^2 \tag{9.2}$$

where U^* is the complex conjugate of U. If the usual square-law photodetectors are used at the Fourier plane, the power spectrum, instead of the Fourier transform, will be detected.

In a basic Bragg cell receiver, the Bragg cell is usually a one-dimensional device. Assume that the amplitude of the power spectrum is unity. The output at the Fourier plane of a one-dimensional system can be obtained from Eq. (9.1) as

$$U(x_f) = \int_{-\infty}^{\infty} t(x) \exp(-j2\pi f_x x)\, dx \tag{9.3}$$

where $f_x = x_f/\lambda F$.

For a sinusoidal amplitude grating of length l, $t(x)$ can be expressed as

$$t(x) = \left(\frac{1}{2} + \frac{m}{2}\cos 2\pi fx\right)\text{rect}\,\frac{x}{l} \tag{9.4}$$

where m is the modulation depth, f is the spatial frequency of the grating (equal to the inverse of the grating spacing periodicity), and rect is the rectangle function defined as

$$\text{rect}\,\frac{x}{l} = \begin{cases} 1 & \text{when } |x| < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} \tag{9.5}$$

The amplitude mask $t(x)$ represents the input signal. Substituting $t(x)$ in Eq. (9.3) yields

$$U(f_x) = \int_{-\infty}^{\infty} \left(\frac{1}{2} + \frac{m}{2}\cos 2\pi fx\right)\text{rect}\,\frac{x}{l}\exp(-j2\pi f_x x)\,dx \tag{9.6}$$

Equation (9.6) can be evaluated expediently using the Fourier transform convolution theorem, which states that the Fourier transform of a product ab is equal to the convolution of the Fourier transforms of a and b. Defining $\mathcal{F}(z)$ as the Fourier transform of z, we have

$$\mathcal{F}(ab) = \mathcal{F}(a) * \mathcal{F}(b) \tag{9.7}$$

where $*$ represents convolution.

It is straightforward to show that

$$\mathcal{F}(a) = \mathcal{F}(\tfrac{1}{2} + \tfrac{1}{2}m\cos 2\pi fx) = \tfrac{1}{2}\delta(f_x) + \tfrac{1}{4}m\delta(f_x - f) + \tfrac{1}{4}m\delta(f_x + f) \tag{9.8}$$

where $\delta(f_x)$ is the Dirac delta function defined as

$$\begin{aligned} &\delta(x - x_0) = 0 \quad \text{for } x \neq x_0 \\ &\int_{-\infty}^{\infty} g(x)\delta(x - x_0)\,dx = g(x_0) \end{aligned} \tag{9.9}$$

Let

$$\mathcal{F}(b) = \mathcal{F}\left(\text{rect}\,\frac{x}{l}\right) = l\,\text{sinc}(lf_x) \tag{9.10}$$

where

$$\text{sinc}(lf_x) = \frac{\sin \pi lf_x}{lf_x} \tag{9.11}$$

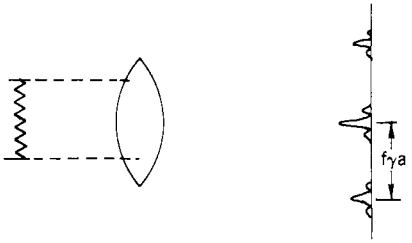


Figure 9.3. Fourier transform of a rectangular window with sinusoidal grating.

The convolution of $\mathcal{F}(a)$ and $\mathcal{F}(b)$ is then

$$\mathcal{F}(a) * \mathcal{F}(b) = \tfrac{1}{2}[\text{sinc}(lf_x) + \tfrac{1}{2}m \text{ sinc } l(f_x - f) + \tfrac{1}{2}m \text{ sinc } l(f_x + f)] \tag{9.12}$$

Therefore, the output at the Fourier plane can be written as

$$U(x_f) = K\left[\text{sinc}\frac{lx_f}{\lambda F} + \frac{m}{2}\text{sinc } l\left(\frac{x_f}{\lambda F} - f\right) + \frac{m}{2}\sin l\left(\frac{x_f}{\lambda F} + f\right)\right] \tag{9.13}$$

where K is a constant.

The output optical wave amplitude consists of three sinc functions as depicted in Figure 9.3. The light intensity on the detector is proportional to the square of $U(x_f)$ or

$$I(x_f) = |U(x_f)|^2 \tag{9.14}$$

From Eq. (9.13), it is obvious that the zero-order output is at the center of the Fourier plane, whereas the first-order outputs are located at positions

$$x_f = \pm f\lambda F \tag{9.15}$$

The displacements of the first-order outputs from the zero-order output are proportional to the grating frequency f when λ and F are fixed. By measuring the position of the light spot, the grating frequency f can be obtained. The intensity of the light spot is the square of a sinc function, which is the Fourier transform of the input window.

9.3. BRAGG DIFFRACTION (12–19)

Although the discussion in the above section illustrates the basic operating principles of Bragg cell receivers, the results are strictly valid only for thin sinusoidal amplitude and phase gratings. However, in most Bragg cell applications, the acoustic wave (and hence the grating) extends over a significant distance along the light propagation path. In this case, the modulation mask is

regarded as a thick grating. An important consequence of this fact is that the incident angle of the optical wave onto the acoustic waveform must be appropriately adjusted to achieve maximum energy transfer to the first-order output in the Fourier plane.

The important parameters for describing Bragg cell operation are illustrated in Figure 9.4. In this figure, a sinusoidal acoustic wave of frequency f with corresponding wavelength Λ propagates with velocity V_s along the x direction in an optically isotropic Bragg cell of refractive index n. A collimated optical wave of free-space wavelength λ impinges on the sound field at angle θ_i from the z axis, as measured in the Bragg cell. The width of the acoustic wave along the z axis (and hence the thickness of the resultant elasto-optically formed phase grating) is L. Analysis of Maxwell's wave equations shows that two optical waves emerge from the cell, the transmitted incident beam and the first-order diffracted beam deflected at angle θ_d from the z axis as measured in the Bragg cell. For maximum energy transfer to the first-order beam, it is necessary that

$$\theta_i = \sin^{-1}(\lambda f/2nV_s) \tag{9.16a}$$

and that

$$\theta_d = \theta_i \tag{9.16b}$$

By analogy to X-ray diffraction in crystals, Eqs. (9.16a) and (9.16b) are called the Bragg conditions, and the incidence angle θ_i satisfying Eq. (9.16a) is called the Bragg angle. In the case of typical Bragg cell optical processors, $\theta_i \ll 0.1$ rad and Eq. (9.16b) can be approximated as

$$\theta \simeq \lambda f/2nV_s \tag{9.17}$$

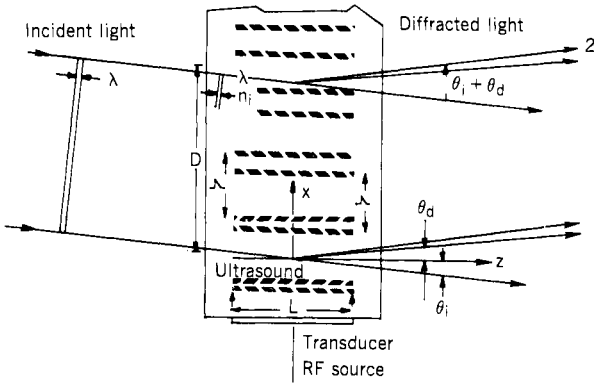


Figure 9.4. Traveling wave acousto-optic modulator. (Courtesy of Avionics Laboratory, AFWAL.)

In terms of the external incidence angle θ_i', one can write

$$\theta_i' \simeq \lambda f / 2V_s \tag{9.18}$$

and

$$\theta_d' \simeq \lambda f / 2V_s \tag{9.19}$$

The optical wave exiting the Bragg cell is composed of two parts, one whose phase is identical with that of the incident wave and one whose phase is linearly modulated over the x-axis extent of the Bragg cell. The Fourier transform lens then yields an optical amplitude distribution in the focal plane consisting of two sinc functions. The displacement of the first-order spot from the undeflected spot is given by

$$x_f = F(\theta_i' + \theta_d') = F\lambda f / V_s \tag{9.20}$$

and is seen to be proportional to the input acoustic frequency f and focal length F.

The Bragg conditions can also be obtained through the following argument. The acoustic wave is represented as a set of equally spaced, partially reflecting mirrors (see Fig. 9.5*a*). In this arrangement, $\theta_i = \theta_d$, and the deflected wave will be a superposition of waves reflected from the series of mirrors. The deflected wave will have maximum strength when the waves reflected from each mirror

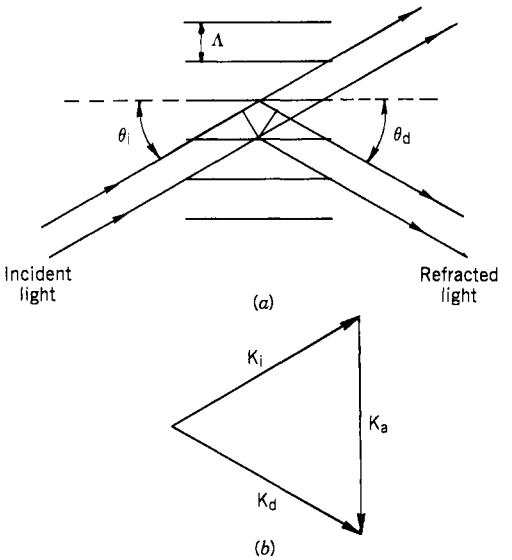


Figure 9.5. Reconstruction geometry. (*a*) Reflection pattern (*b*) conservation of momentum.

are all in phase. This will occur when the optical path difference between waves reflected from successive mirrors is one optical wavelength:

$$2\Lambda \sin \theta_i = \lambda \tag{9.21}$$

or

$$\theta_i = \theta_d = \sin^{-1} \frac{\lambda f}{2V_s} \simeq \frac{\lambda f}{2V_s} \tag{9.22}$$

which is similar to the results in Eqs. (9.18) and (9.19). Figure 9.5*b* shows the conservation of momentum argument of the Bragg condition. The incident optical wave vector $\mathbf{k}_i$, the deflected optical vector $\mathbf{k}_d$, and the acoustic wave vector $\mathbf{k}_a$ are related through

$$\mathbf{k}_i + \mathbf{k}_a = \mathbf{k}_d \tag{9.23}$$

In practice, the incident optical angle is fixed at the angle that fulfills Eq. (9.15) at the center of the operational band. In this case, the deflected beam will change angle according to the acoustic frequency, but the incident angle stays the same. This is true because natural diffraction spreading of the acoustic waves automatically provides a range of incidence angles for a fixed optical beam.

If multiple signals are input to a Bragg cell, there will be more deflected beams with each deflection as given in Eq. (9.20). However, if too many signals are present in a Bragg cell simultaneously and the total amount of deflected light is beyond a certain level, the outputs will be affected by one another. In this case, the input signals have driven the Bragg cell to a nonlinear region, and interference among the signals will occur.

9.4. BRAGG CELLS (13, 15, 20–29)

A Bragg cell is made of a light-transmitting crystal with an input transducer at one end of the cell, as shown in Figure 9.6. The input transducer changes the input RF signal to acoustic form and launches it into the crystal as an acoustic wave. In the crystal, the acoustic wave produces variations in the refractive index of the crystal. This refractive index modulation causes the laser beam deflection. The transducer is a piece of piezoelectric material sandwiched between two conductors and bonded onto one end of the crystal.

The input signal is applied to the two conductors through an impedance matching network. The thickness of the transducer depends on the operating frequency of the Bragg cell. For an input frequency of 1 GHz, the thickness is around 1 μm. The size of the transducer is determined by the outer conductor. At higher operating frequencies, smaller transducers are required, implying that less acoustic energy will be transmitted to the Bragg cell. Therefore, the higher

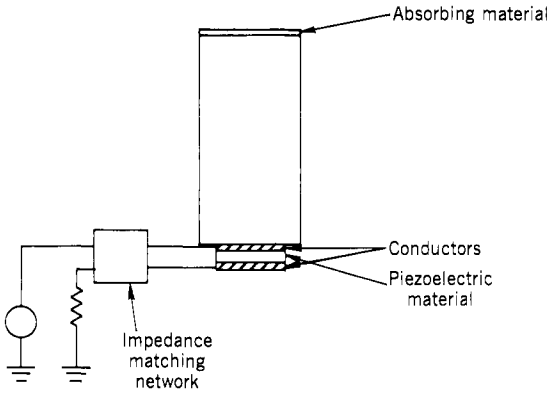


Figure 9.6. Basic Bragg cell configuration.

the operating frequency of a Bragg cell, the lower the efficiency. The electric field applied across the piezoelectric material will create a mechanical vibration that will couple into the crystal. At the opposite end of the crystal, there is an absorbing material that will prevent the acoustic wave from being reflected back. Spoiling the reflection at the end of the Bragg cell can also be accomplished by roughening the crystal itself or by cutting the end of the Bragg cell at an angle so that the reflected acoustic wave will not reflect in the same direction. Due to the suppression of the reflected acoustic waves, it is seen that Bragg cells use traveling acoustic waves rather than standing acoustic waves to carry out their function.

The piezoelectric material used in the transducer can be one of several materials. The most commonly used are lithium niobate ($LiNbO_3$) and zinc oxide (ZnO). The conducting material deposited on the piezoelectric material can be either gold or aluminum. The input impedance of the transducer is rather low, and a matching network must be used to match the impedance to 50 Ω for maximum power transfer. A Bragg cell with matching network is shown in Figure 9.7.

One basic physical requirement necessary for the Bragg cell to work is that the thickness L of the Bragg cell (along the light direction), the acoustic wavelength Λ, the light wavelength λ, and the index of refraction n be related approximately by (ref. 15)

$$L \geqslant 2n\Lambda^2/\lambda \tag{9.24}$$

The acoustic velocity in solids is typically about 6×10^3 m/sec. Then for a 1-GHz input signal, the acoustic wavelength is approximately 6 μm with a light wavelength of 0.6328 μm (He–Ne laser) and a value of $n = 3.31$ (GaP material), the thickness of the Bragg cell should be greater than 0.38 mm. In general, the thickness of a Bragg cell is from a few millimeter to about 10 mm. To

Figure 9.7. Bragg cell. (Courtesy of Applied Technology, Division of Litton.)

keep the acoustic loss in the Bragg cell low, the optical window is kept very close to the transducer of the Bragg cell.

The time–bandwidth product of a Bragg cell is the product of the transient time and the operating bandwidth. The transient time is the window width of the Bragg cell divided by the acoustic velocity. The time–bandwidth product depends on both the manufacturing technology and the material used. The maximum number of parallel channels in a Bragg cell receiver equals the time–bandwidth product of the Bragg cell used. The bandwidth of the receiver equals the bandwidth of the Bragg cell. The frequency resolution of the receiver is inversely proportional to the transient time (τ) as

$$\Delta f = k/\tau \tag{9.25}$$

where k is a constant and has a value in the range 1.2–2. If $k = 1$, the total number of parallel channels equals the time–bandwidth product of the Bragg cell.

The time–bandwidth product is not the only parameter used to choose the Bragg cell material. The efficiency of the Bragg cell and the laser used in the Bragg cell receiver can also affect the material choice. For example, GaP has a very wide time–bandwidth product (~ 750) and can operate at high RF frequencies with Ne–He and solid-state lasers. However, it does not work with an argon laser because GaP does not transmit blue light.

The efficiency of a Bragg cell is defined as the ratio of the diffracted light power to the input light power. The diffracted light power is dependent not only on the Bragg cell itself but also on the RF input power. To make the efficiency of a Bragg cell RF power independent, the efficiency is defined in percent per watt. For example, an efficiency of 30% per watt means that the Bragg cell will diffract 30% of the input light beam with 1 W RF input power. An efficiency of 200% per watt theoretically means that the Bragg cell will diffract all the input light with 0.5 W RF input power. In general, however, the efficiency is defined under linear operating condition.

The light output from a Bragg cell is linearly related to the input RF power only up to approximately 10% efficiency. Above this limit, the diffraction will become highly nonlinear. This limits the linear dynamic range of the cell. A Bragg cell with high efficiency requires low RF input driving power; thus a high-efficiency Bragg cell is desirable.

As discussed in Section 9.3, the Bragg angle condition cannot be exactly matched over a wide frequency range. The efficiency is determined for each frequency according to the ultrasonic transducer angular radiation pattern intensity (W) in the direction required for the Bragg condition. This limits the interaction length and corresponding efficiency for a given bandwidth. For a simple uniform transducer, the angular spectrum is the sinc^2 function. Then, for normal diffraction,

$$W(F, F_m) = \text{sinc}^2\left[\frac{1}{2}\frac{L}{L_0}(FF_m - F^2)\right] \tag{9.26}$$

where F is the frequency normalized to the midband frequency and F_m is the frequency at which the Bragg angle is matched, L_0 is the acousto-optic characteristic length at the midband frequency, and

$$L_0 = \Lambda^2 n/\lambda \cos\theta_i \tag{9.27}$$

Normal diffraction band shapes are shown in Figure 9.8.

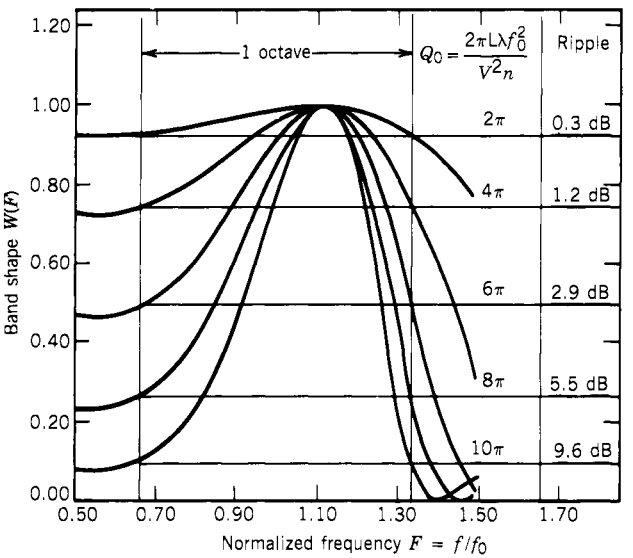


Figure 9.8. Acousto-optic band shapes for normal matching with octave bandwidth. (Based on Hecht, ref. 15.)

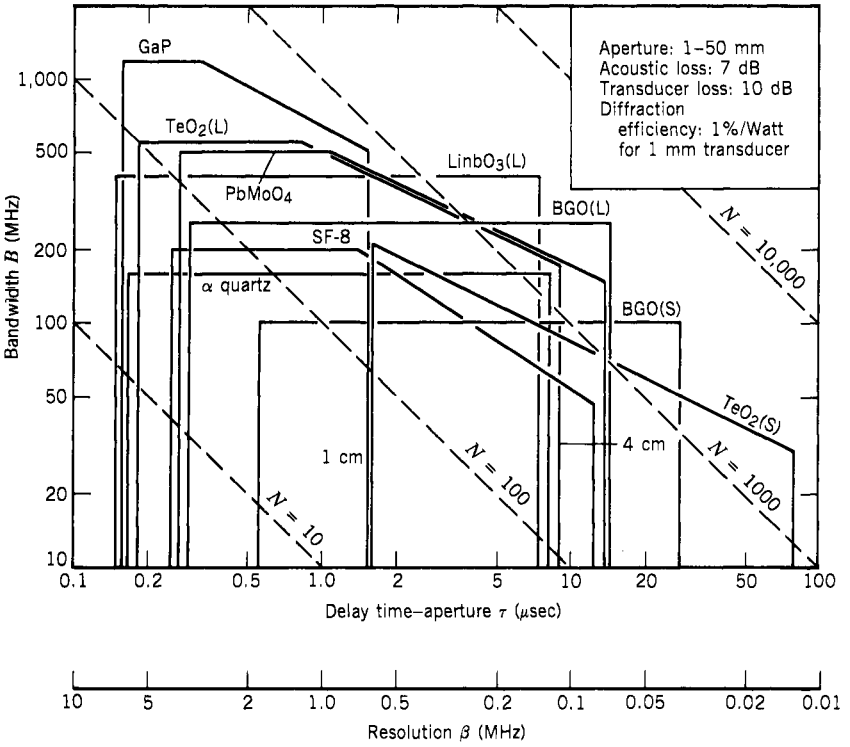


Figure 9.9. Bandwidth resolution contours. (Courtesy of Applied Technology, Division of Litton.)

The time–bandwidth products of some commonly used Bragg materials are shown in Figure 9.9. The frequency bandwidth is shown on the vertical scale and the time delay τ is shown on the horizontal axis. The inverse of τ is plotted as the resolution (Δf) in megahertz. The dotted lines show the values of the time–bandwidth products. Although Figure 9.9 provides the general limitations of the Bragg cell materials, recent developments have made possible Bragg cells that provide better predicted values. For example, using shear mode acoustic wave and proper transducer design, a $LiNaO_3$ Bragg cell with 2000 MHz bandwidth and 200–300 nsec window time has been fabricated.

9.5. WEIGHTING EFFECT IN A BRAGG CELL (15)

It is shown in Eq. (9.2) that the output of a Bragg cell is the Fourier transform of the window. If the window is a one-dimensional rectangular function, the output at the Fourier plane is a sinc function. This means that the light spot will have a main lobe and many side lobes. If only one signal is measured by a Bragg cell receiver, the main lobe can be easily detected as the center of all the detected

frequencies. If multiple signals are present in a Bragg cell, it is necessary to distinguish the main lobes from side lobes. The complexity of the problem is increased when wide dynamic range is required from the receiver. This same problem happens in channelized and compressive receivers.

In a Bragg cell receiver, the simplest solution seems to be adding a proper weighting function to the Fourier transform to reduce the side lobes. The weighting filter in a Bragg cell receiver is a spatial one and is generally termed as apodization. It can be produced by an absorbing material with a desirable spatial attenuation distribution installed in front or behind a Bragg cell. This arrangement changes the spatial light intensity of the laser beam that can be used to convolve with the Bragg cell window to reduce the side lobes at the Fourier plane. In addition, there are a number of naturally occurring weighting functions. To see how they affect the spit size and its side lobe, the following discussion is presented.

Equation (9.3) is rewritten here as

$$U(x_f) = \int_{-\infty}^{\infty} t(x)\exp(-j2\pi f_x x)\,dx$$

where $t(x)$ represents the composite natural weighting function composed of the window function, the acoustic amplitude attenuation function, and the optical beam profile. The "window" function $t_1(x)$ can be written as

$$t_1(x) = \text{rect}\left(\frac{x}{D} - \frac{1}{2}\right) \tag{9.28}$$

where D is the height of the Bragg cell. The window is extended from $x = 0$ to $x = D$. The acoustic amplitude attenuation function $t_2(x)$ is

$$t_2(x) = \exp(-\alpha x) \tag{9.29}$$

where α is a frequency-dependent loss factor in nepers per second. The optical beam amplitude profile $t_3(x, T)$ is calculated as

$$t_3(x, T) = \exp\left[-4T^2\left(\frac{x}{D} - \frac{1}{2}\right)^2\right] \tag{9.30}$$

where $T = D/(2W_0)$ and W_0 is the half-width of the laser beam at $1/e^2$ of the center intensity, where e is the base of the natural logarithm. When $T = 0$, the optical beam is uniform. The combined weighting function is the product $t_1t_2t_3$:

$$t(x) = \exp\left[-\alpha x - 4T^2\left(\frac{x}{D} - \frac{1}{2}\right)^2\right]\text{rect}\left(\frac{x}{D} - \frac{1}{2}\right) \tag{9.31}$$

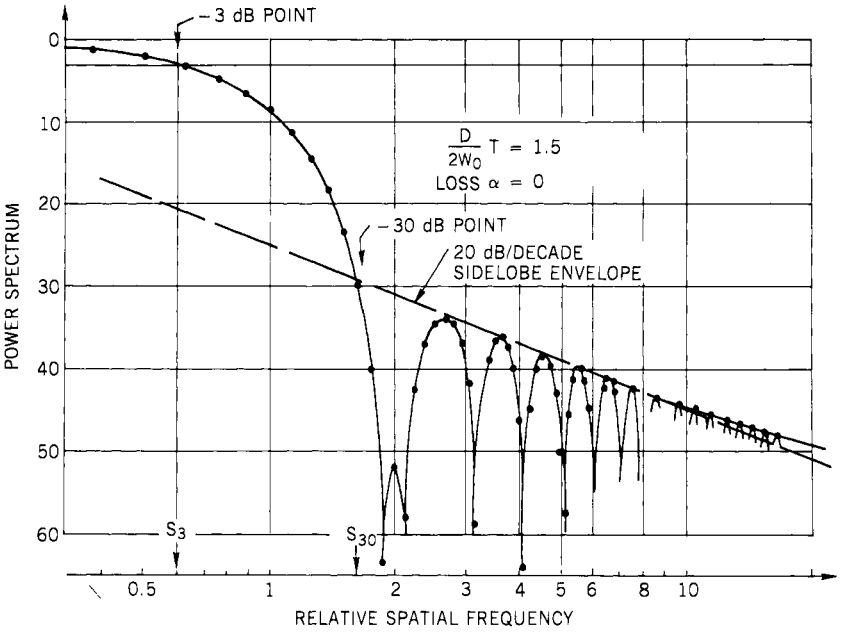


Figure 9.10. Spectrum of truncated Gaussian. (Based on Hecht, ref. 15.)

Substituting Eq. (9.31) into (9.3), the Fourier transform can be obtained from numerical integration with a computer. A typical spectrum is shown in Figure 9.10, where $T = 1.5$ and $\alpha = 0$. As expected, the side lobes are reduced in amplitude, and the main lobe is broadened. For this special arrangement, all side lobes are lower than 34 dB relative to the central peak, and the side lobe peaks drop off at the expected rate of 20 dB/decade (or 6 dB/octave). The slight rise of the side lobes at high relative spatial frequency (>8) shown in Figure 9.10 is due to the computation and is not a physical effect.

9.6. GENERAL CHARACTERISTICS OF PHOTODETECTORS (30 (30–40)

The last component to be discussed in the optical domain of the Bragg cell receiver is the photodetector. Photodetectors are placed in the Fourier plane to detect the position of the light spot, which is related to input signal frequency. At the outputs of the photodetectors, the light signals are converted back to electrical signals where further processing is needed to convert the signals to digital information. In this section, the discussion will concentrate on the general characteristics of photodetectors. Different types of photodetectors will be discussed in the next section.

Photodetectors can be divided into two groups according to their responses to light. One kind is a power detector, and its output is a voltage or current proportional to the optical power input of the detector. The other kind is an energy detector whose output is proportional to the total energy accumulated during a specified time period. The noise floor of these detectors are defined somewhat differently.

A power photodetector is similar to a square-law detector in microwave receivers. The output current is proportional to the input optical power. The dynamic range of a photodetector is often defined as extending from the noise level to saturation. The noise floor is specified in terms of the noise-equivalent power (NEP), defined as the root-mean-square (rms) incident optical power required to give an output rms signal current (or voltage) equal to the rms noise current (or voltage) (refs. 30, 31). If the photodetector responsivity R is defined as

$$R = \frac{\text{output current (or voltage)}}{\text{input optical power}} \tag{9.32}$$

The NEP can be written as

$$\text{NEP} = N/R \tag{9.33}$$

where N is the rms noise. For silicon photodetectors, R is approximately 0.4 A/W.

With this definition, the NEP is given in watts. The NEP is used to describe a specific detector under specific operating conditions. Generally, it is a function of detector bias current, illumination (optical) wavelength, and electrical (video) frequency bandwidth because the noise and responsitivity depend on these parameters. The NEP is often specified at some low frequency (e.g., 1 kHz) with 1 Hz bandwidth and at the peak response wavelength. Usually the NEP will be independent of frequency from approximately 1 kHz to some upper frequency limit—typically 10–100 MHz.

Another frequently encountered definition of the noise-equivalent power is

$$\text{NEP}' = N/\sqrt{B_v}R \tag{9.34}$$

where B_v is the video bandwidth. In this case, the NEP has the units of watts per hertz$^{1/2}$. The NEP′ is numerically equivalent to NEP if the video bandwidth in NEP specification is 1 Hz.

To determine the electronic noise floor of a detector when the NEP is specified for the specific detector operating conditions, one simply uses Eq. (9.33). However, if the NEP is specified at 1 Hz bandwidth or if NEP′ is given, it is necessary to make the assumption that the noise spectral density (mean square noise per unit bandwidth) is frequency independent. Under this situation, the electric noise floor is computed by Eq. (9.34).

For example, the RCS C38016 silicon PIN diode preamplifier module is specified to have an NEP' of 3×10^{-12} W/$\sqrt{\text{Hz}}$ and a responsitivity of 10^4 V/W at 900 nm wavelength. If this detector is operated at 1 MHz bandwidth, the rms electric noise floor is 3×10^{-5} V ($3 \times 10^{-12} \times 10^4 \times \sqrt{10^6}$). To obtain the optical noise floor at 633 nm (He–Ne laser line), the responsitivity at 633 nm is needed. Taking R (633 nm) to be 60% of R (900 nm), Eq. (9.33) yields an NEP of 5×10^{-9} W, ($3 \times 10^{-5}/(0.6 \times 10^4)$) or -53 dBm. This number is often used to estimate the sensitivity of a Bragg receiver.

In integrating photodetectors where charge transfer techniques are used for readout, the NEP is not generally used as a figure of merit because the dominant noise sources are different and the output is proportional to energy rather to power. In such a device, the incident radiant flux generates charge carriers that are stored in capacitors during the integration period. After the integration period, the stored carriers are clocked out. One way to clock them out is with charge-coupled device (CCD) shift registers. Figure 9.11 shows an equivalent circuit of such a detector array. The noise is generally described in terms of noise electrons produced in the photodetection readout process. One way to characterize the noise floor is to compute the number of noise electrons N associated with a given detector element. Equivalent to the NEP, noise floor can be defined as the average optical power incident to the element over integration time that would result in a stored charge of N electrons. Assuming the

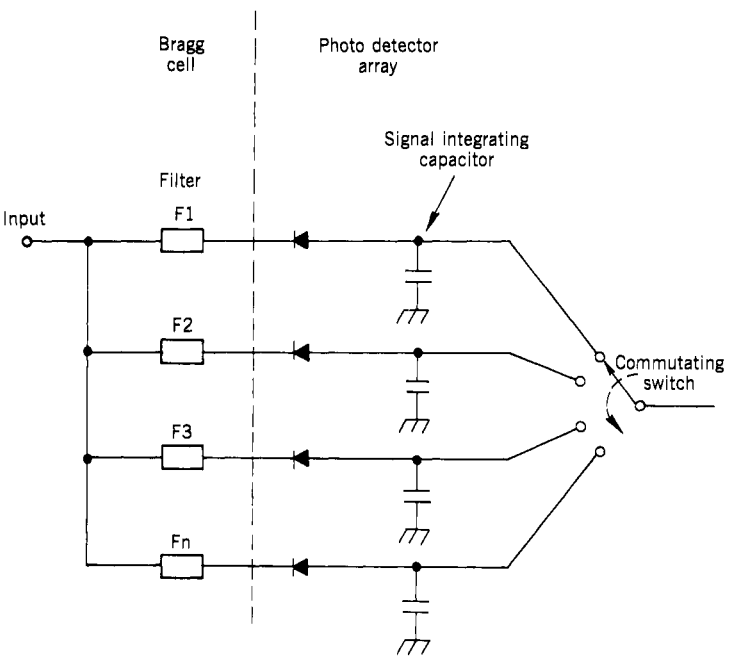


Figure 9.11. Electrical equivalent of photodetectors.

photon-to-charge-carrier conversion factor (quantum efficiency) is η, then P_0, equivalent to the NEP, can be written as

$$P_0 = \frac{N_e hf}{\eta\tau} \text{ W} \tag{9.35}$$

where h is the Planck's constant ($=6.625 \times 10^{-34}$ J-sec) f is the frequency of the incident light, and τ is the integration time in seconds.

If the number of noise electrons N_e is given, the value of P_0 can be computed for use in system performance calculations. For example, if $N_e = 1000$ and $f = 4.74 \times 10^{14}$ Hz ($\lambda = 633$ nm) for 1 msec integration time with $\eta = 0.5$,

$$P_0 = \frac{1000 \times 6.625 \times 10^{-34} \times 4.74 \times 10^{14}}{0.5 \times 10^{-3}}$$

$$= 6.28 \times 10^{-13} \text{ W} = 6.28 \times 10^{-10} \text{ mW} = -93 \text{ dBm}$$

From this result, one may draw the conclusion that the energy photodetector has a lower noise floor than a power photodetector. This is not really true because the noise floor depends on the integration time. A better way to explain this phenomenon is the longer the integration time, the lower the noise floor, if there is an input signal at the photodetector. In other words, the energy photodetector can provide better sensitivity in detecting long pulses or CW signals. Therefore, they are suitable for communication receivers. If the input signal is a short pulse and the integration time is 1 μsec, the NEP will be -63 dBm, which is 30 dB higher than the above value (-93 dBm). Although the noise floor may be slightly lower than in a power photodetector, if other characteristics are taken into consideration, that is, response time, a power photodetector may be preferred over an energy type in some applications.

Whereas a power photodetector will respond as an ordinary microwave video detector, the sensitivity of an energy photodetector will depend on the input pulse width (PW). Assume the integration time is T_s, the Bragg cell window is τ, and T_s is greater than τ. The three kinds of input signals are

$$\text{PW} > T_s, \qquad \tau < \text{PW} < T_s, \qquad \text{and} \qquad \text{PW} < \tau$$

The power on the detector as a function of time is shown in Figure 9.12 for signals of the same PA. The sensitivity of the receiver decreases as the total received energy (areas under the curves in Fig. 9.12) decreases. The sensitivity expressed in dB can be approximately represented as

$$\text{PW} > T_s \qquad \text{full sensitivity } S \tag{9.36}$$

$$\tau < \text{PW} < T_s \qquad S + 10\log(T_s/\text{PW}) \tag{9.37}$$

$$\text{PW} < \tau \qquad S + 10\log(T_s/\tau) + 20\log(\tau/\text{PW}) \tag{9.38}$$

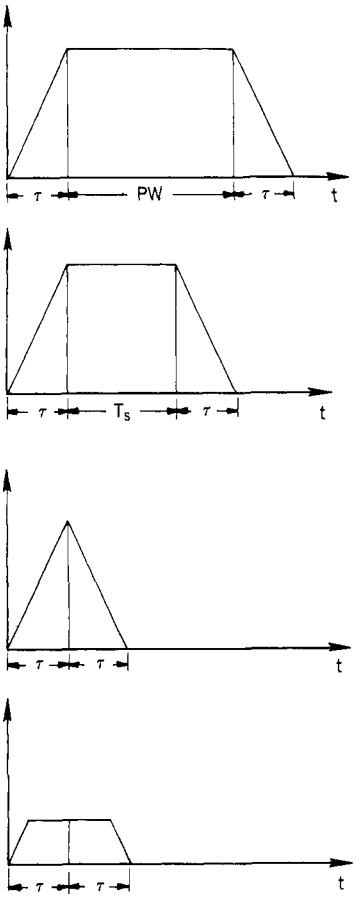


Figure 9.12. Energy on a detector for different signal pulse width.

The factor 20 in Eq. (9.38) is due to energy reduction as well as to spectrum spreading. The results are shown in Figure 9.13.

Let us use an example to demonstrate this discussion. Assume that the integration time is 10 μsec, the Bragg cell window is 1 μsec, and the full sensitivity of the receiver is -60 dBm for a CW signal.

For PW $>$ 10 μsec, the sensitivity is -60 dBm.
For PW = 5 μsec, $S = -57$ dBm $(-60 + 10 \log 10/5)$.
For PW = 0.1 μsec, $S = -30$ dBm$(-60 + 10 \log 10/1 + 20 \log 1/0.1)$.

Of course, for PW $< \tau$, the sensitivity of a Bragg cell receiver with a power photodetector will also degrade.

Like a channelized receiver, it is desirable to add amplification in front of the photodetectors to improve the receiver sensitivity. The amplifiers used in a

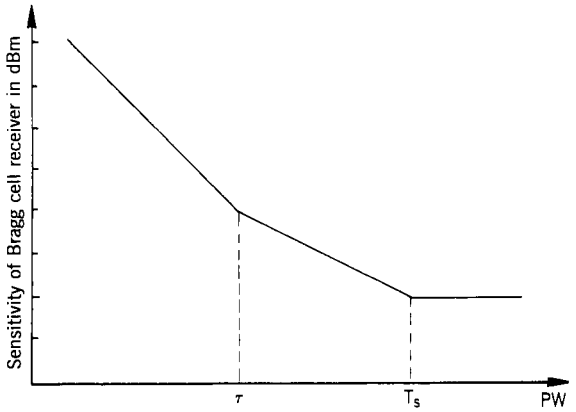


Figure 9.13. Bragg cell receiver sensitivity versus input pulse width.

Bragg cell receiver must amplify light signals instead of electric signals. A light amplifier usually consists of two functions: photoemission and photoamplification. Photoemission in vacuum through a photodiode has been one of the preferred means of detecting low optical power because of the low dark current noise. The favorite method of amplification has been secondary emission, as in a photomultiplier tube. The desired photoamplifier must have many parallel channels and be small in size so that it can be placed in the Fourier plane. Successful development of the light amplifier may improve the sensitivity of a Bragg cell receiver.

9.7. TYPES OF PHOTODETECTORS (30–39)

There are many different types of photodetectors according to their operating principles, such as photoconductive and photovoltaic types. In Bragg cell receivers, photodetectors of the photovoltaic type are the most commonly used. The three detector subtypes generally employed included PIN (*p* material–intrinsic–*n* material) photodiodes, avalanche photodiodes (APD), and charged coupled devices (CCD). The first two types are power detectors, and the last is an energy detector.

A photodiode is a *pn* junction device and usually is reverse biased. The bias voltage will create a depletion region with a high electric field. However, this field is not large enough to cause avalanche. When photons are incident on a photodiode, electron hole pairs are generated. These electron hole pairs are separated by the electric field in the depletion region. For high-speed operation, the depletion region must be kept thin to reduce the transit time. On the other hand, the depletion layer must be thick enough so that most of the incident photons are absorbed, thereby providing high quantum efficiency. Thus the response time and quantum efficiency in a photodetector must be properly

compromised. The PIN diode with an intrinsic layer between the p and n materials is one of the most commonly used photodetectors because the depletion region thickness, which is also the intrinsic layer thickness, can be controlled to provide the desired frequency response and quantum efficiency. This type of photodetector has the potential to be used in a Bragg cell receiver.

APDs are operated at high reverse bias voltage so that avalanche multiplication takes place. This multiplication provides current gain, which improves the sensitivity of the diode. Consequently, the NEP of an APD is lower than that of a PIN photodiode. Presently, there are two areas of concern in employing APDs in Bragg cell receivers. First, the applied bias voltage required by the APD is in the several-hundred-volt range. Although this does not have an adverse effect on the performance of the receiver, it is an inconvenience. Second, the APD output current is sensitive to ambient temperature. This effect will influence the performance of the receiver. The APD temperature problem must be solved before they can be routinely used in Bragg cell receivers.

A CCD array consists of a large number of closely spaced metal–oxide semiconductor (MOS) diodes. The CCD array can be in one- or two-dimensional forms. The diodes are operated in a depletion mode. When photons are incident on a CCD, they will create electron hole pairs, and the charges will be stored in charge packets at the semiconductor surface. In readout, these charge packets are transferred through the MOS diode chain by appropriate clock voltage applied to the input gate, as shown in Figure 9.14. Figure 9.14*a* shows the basic

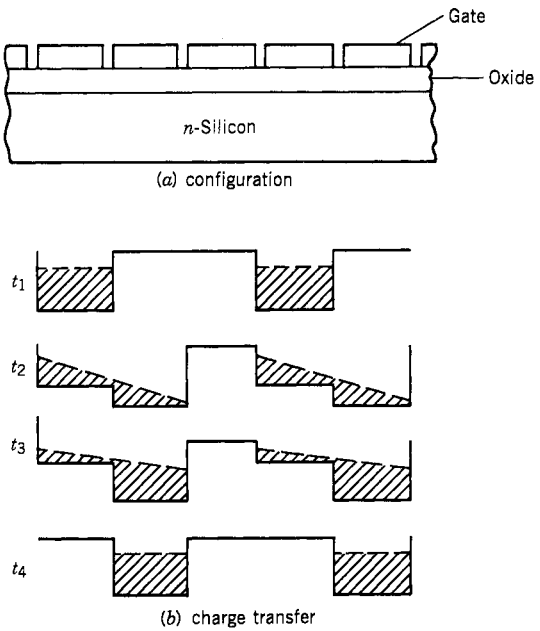


Figure 9.14. Charge-coupled device.

diode structure, and Figure 9.14*b* shows the charge transfer from one diode to a neighboring one. Since photodetectors can be densely packed in a CCD array, such a detection system can be put directly at the Fourier plane of a Bragg cell receiver.

9.8. DISCRETE PHOTODETECTORS AND FIBER OPTICS (41, 42)

Photodetectors can also be divided into two groups according to whether they come in discrete packages or in arrays. Although both kinds can be used in Bragg cell receivers, their performance levels can be quite different. Discrete photodetectors and their application in Bragg cell receivers are discussed in this section.

Popular discrete photodetectors include silicon planar PIN photodiodes and avalanche diodes. Discrete photodiodes usually have high dynamic range and fast response time. However, due to packaging problems, it is difficult to

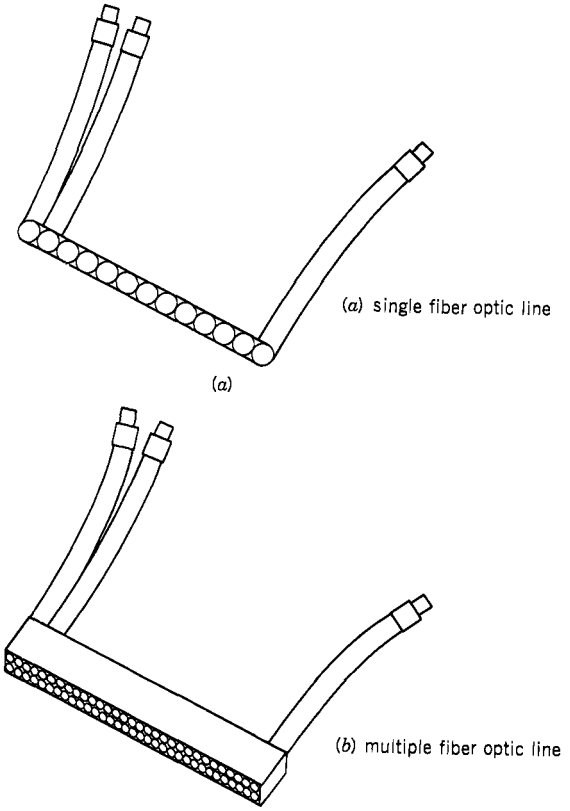


Figure 9.15. Arrangement of fiber optics.

put the diodes directly on the Fourier plane. One way to use discrete diodes in a Bragg cell receiver is to convey the light from the Fourier plane to the diodes through fiber optics. An optical fiber is a long polished glass cylinder that will transmit light without loss by leakage through the cylinder walls provided that all surface reflections occur at an angle exceeding the critical angle. There are two approaches to using fiber optics. One way is to use fiber sizes comparable to that of the light with the fibers arranged in a single array, as shown in Figure 9.15*a*. Another way is to use fibers that are much smaller than that of the light spot. The fibers are arranged in compact groups, as shown in Figure 9.15*b*. In this arrangement, one diode is fed by many fibers. In the first approach, if one fiber is broken, the corresponding channel will be totally blanked out. Another disadvantage of this approach is then when the light falls between two fibers, a portion of the light will be lost. If rectangular fibers are available, this loss of sensitivity between channels will be reduced.

In the second arrangement, if one fiber is broken, the diode will only lose part of the light. The resultant intensity error may be misinterpreted by subsequent processing circuitry resulting in erroneous information. Since a large number of fine fibers have higher probability of having broken fibers, the single-fiber arrangement shown in Figure 9.15*a* is often used at present.

The major defficiency of using fiber optics is the large volume required by the fiber optics and detectors; the detection system may be many times larger than the optical bench. The large receiver size in Figure 9.1 is due to the fiber optics used.

9.9. PHOTODETECTOR ARRAYS

A photodetector array contains many photodetectors in either a linear or a two-dimensional array. The photodetecting elements in an array can be packed very close together. For example, a Reticon RL 1024C linear array contains 1024 elements with approximately 1 mil center-to-center spacing. Therefore, the detector array can be placed directly in the Fourier plane. Thus it is reasonable to assume that photoarrays will be used in Bragg cell receivers. There are many different kinds of linear arrays, and they will be classified here according to how the information is coupled out.

A. Parallel Arrays

The outputs from the photodetectors are connected individually to outside circuits. This kind of array is similar to discrete photodetectors, and they are power detectors. The isolation between the adjacent elements should be high in order to produce the high dynamic range of the receiver. If the technology in parallel detector arrays advances, they should be able to replace discrete photodetectors in Bragg cell receivers due to their strong package advantage.

B. Series Array

A series photodetector array usually consists of many photodetectors and has only one output. The detectors in the array are usually energy photodetectors rather than power detectors. The output from each detector is switched on sequentially, as shown in Figure 9.16. If there are 1024 detectors in an array and the switching time is 1 μsec, reading the entire array will take 1.024 msec. While a detector is not connected to the output circuit, it integrates the reaching light power. A series array has two disadvantages if it is used in a Bragg cell receiver. First, it provides poor time resolution. Time of arrival (TOA) and pulse width (PW) cannot be measured properly for pulsed signals. In the above example, the TOA and PW resolution is about 1 msec, which is grossly inadequate. The second disadvantage is that a series array is not suitable to detect short pulses. When a short pulse reaches a detector, the detector will respond for the period of the pulse or the Bragg cell window time, whichever is longer. When the signal disappears, the detector will integrate noise. Therefore, the sensitivity for detecting short pulses will be low. Series array photodetectors are more appropriately used in communication receivers where signals have long duration.

C. Series–Parallel Array

A series–parallel array can be considered as many individual series arrays connected in parallel. For example, a series-parallel array can have 160 detectors and 16 parallel outputs. Each output will switch to 10 detectors sequentially. Therefore, the 10 detectors actually form a series array, as shown in Figure 9.17. If the switching time is 1 μsec, a time resolution of 10 μsec can be achieved. Because there are a total of 16 parallel outputs, the output circuit will only process 16 outputs. Thus the design of the output circuit will be simpler than a total parallel approach.

D. Random-Access Array

In a random-access array, the number of outputs is low, one or two. The output sequence of the array can be programmed. For example, an array containing 512 elements may have two outputs, one connected to all the even-numbered detectors and the other connected to all the odd-numbered detectors, as shown

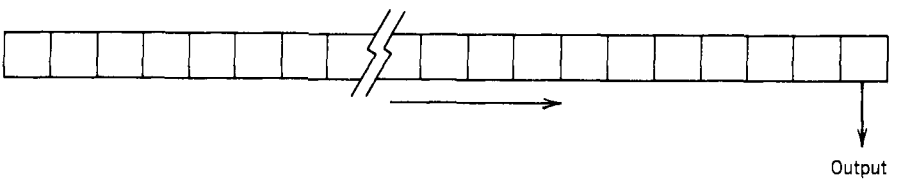


Figure 9.16. Series photodetector array.

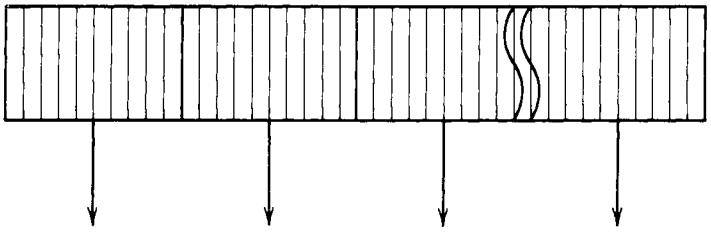


Figure 9.17. Series–parallel photodetector array.

in Figure 9.18. The two outputs can be programmed separately. For example, one output is programmed to sequentially scan the even detectors from 2 to 10 and the other output to look at a single detector repetitively. The smaller the total number of detectors examined, the shorter the revisit time. If the switching time is 1 μsec, the time periods in the above arrangement are approximately 10 μsec for the first output and 2 μsec for the second. This kind of detector array may be useful for ELINT Bragg cell receivers. If a priori information is available, the array can be programmed to output the desired signals.

E. Two-Dimensional Array

A two-dimensional photodetector array contains rows (or columns) of detectors. Some television cameras use a two-dimensional photo array. One way of using a two-dimensional array in a Bragg cell receiver is to use one dimension (say the row) to represent frequency and the other to represent time. During a certain time period, such as 1 μsec, the entire row of information will shift downward and the new data will start to collect on the top row. A two-dimensional photodetector array with frequency information is shown in Figure 9.19.

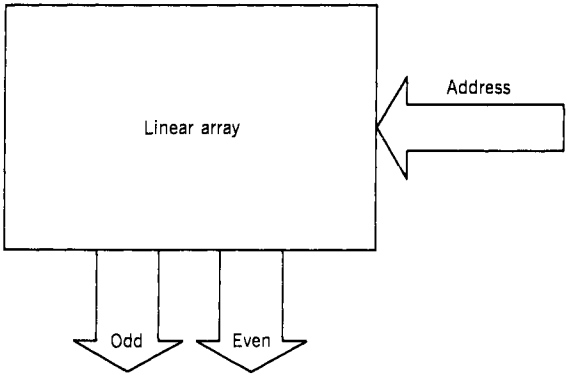


Figure 9.18. Random-access photodetector array.

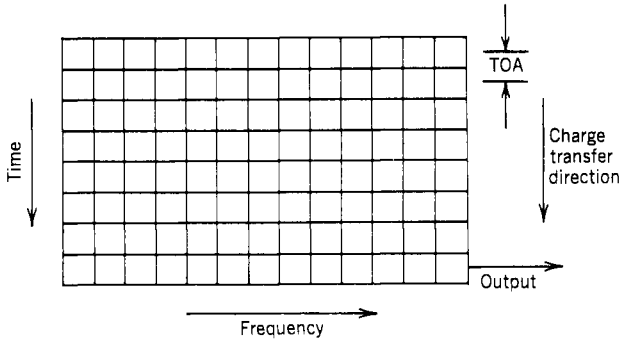


Figure 9.19. Two-dimensional array.

This arrangement will provide good frequency resolution and high TOA resolution. However, when the entire array is filled with information, the individual photodetector must be clocked out in series, which takes time. Therefore, the receiver will spend a small portion of time collecting information and a much longer time processing it. As a result, the probability of intercept (POI) of the receiver will be rather poor. However, if proper processing can be developed, the two-dimensional array may be used in a two-dimensional Bragg cell receiver to process, for instance, frequency and AOA.

With today's technology, it seems that, among all the photodetector arrays discussed above, the only arrays that are suitable for EW Bragg cell receiver applications are the parallel and the series–parallel arrays. The outputs from a parallel array can be processed in a similar way as the outputs of a channelized receiver.

9.10. LASER SOURCES (32, 43–46)

There are two common laser sources used in Bragg cell receivers: the He–Ne gas laser and the GaAlAs semiconductor laser. The gas laser generates coherent light from electronic transitions occuring between discrete energy levels in the lasing gas. The He–Ne laser has a wavelength at 0.6328 μm, and the output is very stable. The divergence of the laser beam is very small. A reasonable sized He–Ne laser source will generate 2–5 mW light power. One advantage of the He–Ne laser is that the light is in the visible region of the spectrum; thus the alignment of the optical bench is relatively easy. Since a He–Ne laser is rather large, it is most often used in experimental Bragg cell receivers.

A semiconductor laser is made from a *pn* junction diode. The quantum transitions are determined by the band structure of the laser material. The laser action is produced by simply passing a forward current through the diode itself. The laser is very small in size, typically on the order of 0.1 mm in length. Thus the divergence of the laser beam is generally larger than that of the He–Ne laser.

As a result, the lens used to form the beam from the laser must be very close to the laser itself. The output wavelength is about 0.8300 μm, which is in the infrared spectral region. The output power is in the 10–20-mW region. High-output lasers can improve the sensitivity of the Bragg cell receiver. The large power combined with the small size makes semiconductor lasers very attractive for Bragg cell receiver applications. One minor disadvantage of the semiconductor laser is that the output wavelength is sensitive to temperature change. The shift in wavelength will reflect as a frequency error, as shown in Eq. (9.19). Therefore, temperature compensation circuitry is often necessary. Since the technology in semiconductor lasers is advancing very rapidly, the temperature problem might be solved soon.

9.11. POWER BRAGG CELL RECEIVER (13, 47–49)

The term power Bragg cell receiver is used because the detector output current is proportional to the input RF power. The key components in a power Bragg cell receiver are the laser, the Bragg cell, an optical bench, photodetectors, and the digitizing circuits that convert the outputs from the photodetectors to digital information. The simplest arrangement of a Bragg cell receiver is shown in Figure 9.20. In addition to the Bragg cell, the optical bench contains a beam expander, a collimator, and a Fourier transform lens. Figure 9.21 shows a compact photobench with all the components except the digitizing circuits.

The laser beam is converted to a collimated sheet beam by the beam expander and collimator. The beam is incident on the Bragg cell at the Bragg angle. The output is focused on the photodetectors through the Fourier transform lens.

For the present discussion, assume that the photodetectors are either discrete photodetectors fed through optical fibers or a parallel photodetector array. The outputs from the photodetectors are video signals. Therefore, only video amplification is required. This is different from a channelized receiver where RF

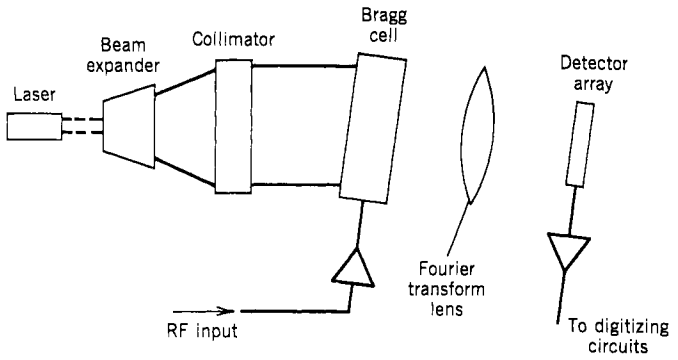


Figure 9.20. Basic Bragg cell receiver.

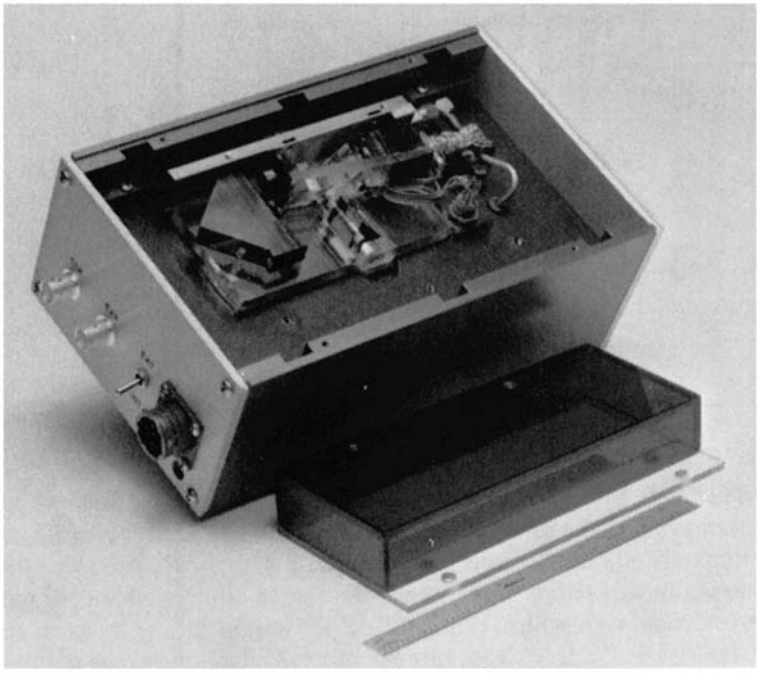


Figure 9.21. Optical bench of a Bragg cell receiver.

amplifiers and fine-frequency determining filters can be used in front of the video detectors to improve frequency-encoding accuracy. One of the most common approaches to obtain fine-frequency information from the photodetector is to compare the amplitude of adjacent channels. In this kind of digitizing circuit, channel gain should be properly matched; otherwise, erroneous frequency data may result.

The sensitivity of a power Bragg cell receiver is calculated differently from that of a conventional microwave receiver. Let us use an example to demonstrate the sensitivity. If the output power of a laser diode is 10 dBm, the transmission loss in the optical path is 3 dB (the transmission loss does not include the efficiency of the Bragg cell, which will be calculated separately), and the noise floor of the photodetector is at -53 dBm. To produce -53 dBm power at the detector, the Bragg cell can have an equivalent loss of 60 dB, which corresponds to a deflection efficiency of 1×10^{-6} of the input light. If the efficiency of the Bragg cell is 20% per watt, the required RF input at the Bragg transducer is 0.005 mW (-23 dBm). If a signal-to-noise ratio of 15 dB is required, the minimum required power at the Bragg cell is -8 dBm. If the RF amplifier has a gain of 55 dB, the sensitivity of the receiver is -63 dBm.

The upper limit of the dynamic range of is determined by the nonlinearity of the Bragg cell. The two-signal third-order intermodulation products are

Table 9.1. Power level in a typical optical bench

	Minimum Signal	Maximum Signal
Input laser power	10 dBm	10 dBm
Optical loss	3 dB	3 dB
Bragg cell diffraction	45 dB	17 dB
RF input power (20% per watt)	−8 dBm	20 dBm[a]
Light power on detector	−38 dBm[b]	−10 dBm

[a]Limited by linearity required.
[b]Set by photodetector noise floor.

10 $\log(\frac{1}{36}\eta^2)$ dB below the desired signal, where η is the diffraction efficiency (ref. 13) in general, the diffraction efficiency should be kept under 2%, which is approximately 50 dB below the desired signal. With a 20% per watt Bragg cell, the maximum allowable input power at the transducer of the Bragg cell is 0.1 W, or 100 mW (20 dBm). With a 55-dB RF amplification, the corresponding input signal is −35 dBm. From this example, the dynamic range of the receiver is 28 dB, from −63 to −35 dBm. Table 9.1 shows the losses in the optical bench up to the photo detectors.

There are several ways to improve the sensitivity of a power Bragg cell receiver. The most common ways are to decrease the noise level of the photodetectors and to increase the laser power. The improvement in sensitivity of these two approaches is increased on a decibel-by-decibel basis. For example, if the laser power is increased by 5 dB, the sensitivity of the receiver will also improve by 5 dB. Therefore, the dynamic range of the receiver is also improved by 5 dB. Of course, to decrease the transmission loss of the optical path will improve the sensitivity. However, the improvement is limited because the loss in a well-designed system is quite good, less than 3 dB. Improving the efficiency of the Bragg cell will decrease the required RF input power to the Bragg cell, which should simplify the RF design, but it would not improve the sensitivity and dynamic range significantly because the upper limit is around 2% diffraction efficiency. The next section will discuss an interferometric approach to improve the dynamic range of a Bragg cell receiver.

9.12. INTERFEROMETRIC BRAGG CELL RECEIVER (50, 51)

The linear dynamic range of a power Bragg cell receiver is limited on the low end by the photodetector. As discussed in the example in Section 9.11, the dynamic range is about 28 dB. A very effective way to improve the dynamic range of a Bragg cell receiver is by the interferometric approach (also referred

to as coherent detection). The fundamentals of the interferometric approach are discussed in this section, and the expected improvement in dynamic range over that of a power Bragg cell system is discussed in the next section.

A photodetector is a square-law device, which means that the output current of a photodetector is proportional to the input optical power. In a power Bragg cell receiver, the optical power is proportional to the input RF power. Therefore, the output current of the photodetector is proportional to the input RF power. If the input RF power changes 10 dB, the output of the photodetector changes 20 dB. Thus the dynamic range of the detector output should be twice that of the input signal range. In the interferometric Bragg cell approach, the optical power from the Bragg cell is proportional to the input RF voltage. Thus the photodetector output is proportional to the input voltage. When the input RF power changes 10 dB, the output of the photodetector also changes 10 dB. The required dynamic range of the photodetector equals that of the input signal range.

An interferometric scheme built around a Mach–Zehnder interferometer (shown in Fig. 9.22) is used to demonstrate the basic idea. A laser beam is split into two paths with a beam splitter, and each beam is reflected by a mirror toward a beam combiner. The beam splitter–mirror–combiner arrangement is in either a rectangular or a square form. This arrangement produces equal path lengths for the two arms of the interferometer. In each path there is a Bragg cell. One of the cells is a signal cell where the input signal is applied. The other cell is a reference cell where a known signal is applied.

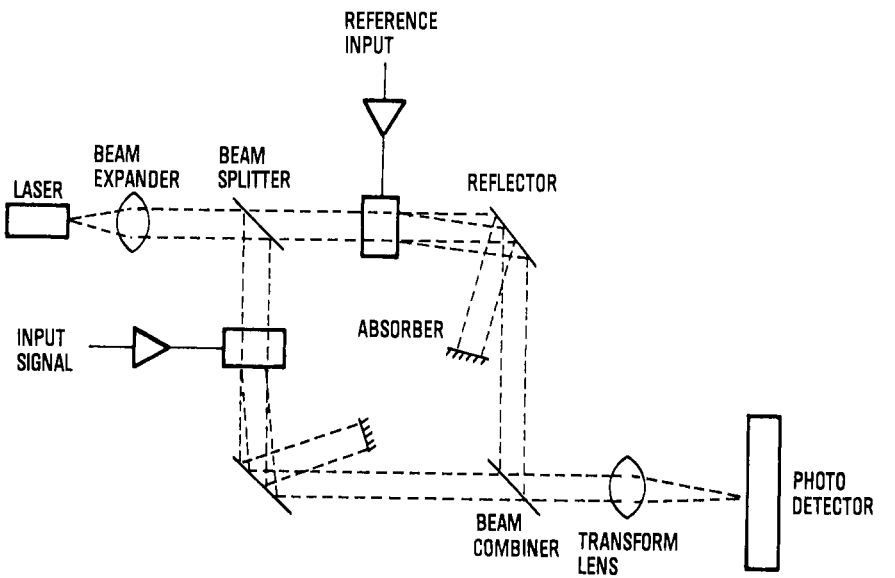


Figure 9.22. Mach–Zehnder interferometric Bragg cell scheme.

In Figure 9.22, the beam expander forms the laser beam into slits at the signal and reference Bragg cells to match the acoustic beams. The diffracted light from each Bragg cell is combined by the beam combiner, and appropriate lenses are placed after the interferometer to form a spot for each spectral component on a photodetector array or fiber array–detector combination.

The reference input signal is always present. If the reference were not present all the time, the probability of intercept (POI) of the receiver would degrade. After the light passes through a Bragg cell, its optical frequency (laser beam wavelength) will shift by the RF signal in the Bragg cell. Thus the output light from the reference Bragg cell can be written as

$$A_{r0} = A_r \cos(\omega + \omega_r)t \tag{9.39}$$

where A_r and ω are the amplitude and angular frequency of the light incident to the reference Bragg cell and ω_r is the angular frequency of the reference signal. Similarly, the output from the signal cell can be written as

$$A_{s0} = A_s \cos(\omega + \omega_s)t \tag{9.40}$$

where ω_s is the angular frequency of the input signal.

The light falling on each photodetector will be a combination of signals from the reference and signal Bragg cells, as described in Eqs. (9.39) and (9.40). Since the detector is a square-law device, its output is proportional to the square of the inputs. Thus

$$\begin{aligned} I &= [A_r \cos(\omega + \omega_r)t + A_s \cos(\omega + \omega_s)t]^2 \\ &= \text{dc terms} + KA_rA_s \cos(\omega_r - \omega_s)t \end{aligned} \tag{9.41}$$

where I is the photodetector current, K is a constant, and the high-frequency terms are presumed to be filtered out following the photodetector. Thus the output current is proportional to the amplitude of the input signal, as anticipated. The output frequency is the difference of the signal and the reference frequencies. It is interesting to note that the outputs from the photodetectors are not video signals any more but RF signals. Therefore, the output signals must again be detected by diode detectors. However, RF amplifiers and/or RF log amplifiers can be added before the diode detector. One potential RF chain extending from the photodetector to the diode detector is shown in Figure 9.23. One can almost consider this approach equivalent to a channelized receiver. Here, however, the frequency in every slot is the same, but in a channelized receiver the frequency in each slot is different.

To discuss the frequency at the output detector, let us use the following example to demonstrate the results. If the Bragg cell receiver has a bandpass of 750–1250 MHz, the input to the signal Bragg cell must also be 750–1250 MHz. Assume that the input bandpass to the reference Bragg cell extends from 800 to 1300 MHz. A simple approach is to assume that the input to the reference

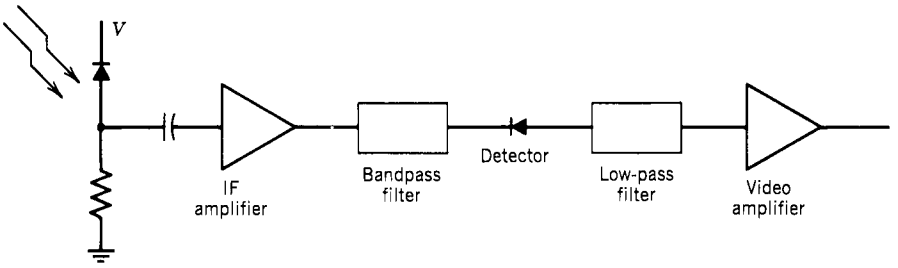


Figure 9.23. Slot assembly in an interferometric Bragg cell receiver.

Bragg cell is a chirp signal. The chirp bandwidth is 500 MHz, and the chirp time is less than the Bragg cell window time. Under this condition, the reference beam will be on the detector array all the time. Let the photodetector array have 500 detector elements. The center of the modulating frequency is 800.5 MHz at one end of the detector array and 1299.5 MHz at the other end. The modulating frequencies of the remaining detectors are each shifted by 1 MHz between 800.5 and 1299.5 MHz. It is presumed that the optical parts are aligned so that both the outputs from the reference and the signal cells fall on the photodetector array.

The light from the signal Bragg cell is present at a certain detector only when a signal is present. If the input signal frequency is 750.5 MHz, the output from the signal Bragg cell will fall on the first detector where the reference frequency is 800.5 MHz. Thus the difference frequency is 50 MHz. If the input signal is 1000.5 MHz, the corresponding output from the signal Bragg cell will fall on the detector where the reference is 1050.5 MHz. Thus the difference frequency is also 50 MHz. Figure 9.24 is used to demonstrate that the output of every photodetector has the same RF frequency.

The output frequency can be chosen by the designer. If the output frequency is too high, however, the cost of the RF components used in the channel will be high. Furthermore, it is relatively difficult to fabricate high-frequency circuits. If the output frequency is too low, the available bandwidth will be narrow and may distort the pulse shape of the input signals.

If the input to the reference Bragg cell is a repetitive chirp signal, the frequency spectrum falling on the photodetector array is not continuous. A Fourier transform of the repetitive chirp signal will show that many discrete frequency components will be presented. These frequency components will mix with the side lobe of the signal input at the photodetectors to produce a signal at IF. This phenomenon will limit the dynamic range of the Bragg cell receiver. Some other approaches may be employed to correct this deficiency.

The reference Bragg cell and the associated electronics to generate the reference signal can be eliminated from the system. The original laser beam without the frequency shift could be used as the reference beam on the photodetectors. Under this arrangement, the outputs from the detectors would have the same

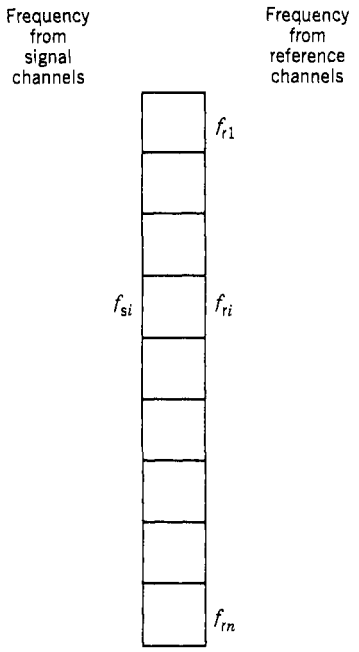


Figure 9.24. Frequency shift on array photodetectors.

frequency as the input signals. Although this arrangement uses only one Bragg cell, it is not a preferred approach because the output frequency from the photodetector would be very high, (possibly, in the gigahertz range). It is difficult to process outputs from the photodetectors with such high frequencies. In addition, each output would have a different frequency, and the results resemble a channelized receiver.

Although the interferometric scheme can be used to improve the dynamic range of the Bragg cell receiver, the receiver is very complex. One additional Bragg cell is needed. In addition to the photodetectors, RF circuit and video detectors will be needed.

9.13. DYNAMIC RANGE IMPROVEMENT BY INTERFEROMETRIC BRAGG CELL SCHEME (13, 50, 51)

From the discussion in Section 9.12, one might conclude that by using the interferometric scheme the dynamic range of a Bragg cell receiver can be doubled in terms of decibels, that is from 28 to 56 dB. However, the actual dynamic range improvement realized will be less than this estimate. The following discussion on the improvement in dynamic range is based on Shah (ref. 50, 51).

The maximum light level at the photodetector is limited by the available laser power and the acceptable intermodulation and spurious signal levels.

Let the total transmission of the optical system be Q and the laser power be P_0. Also assume that the maximum diffraction efficiency for the signal Bragg cell is limited by the third-order intermodulation requirements. The intermodulation is (ref. 13)

$$\text{IM} = -10 \log(\tfrac{1}{36}\eta^2) \quad \text{dB} \tag{9.42}$$

below the desired signal, where η is the diffraction efficiency. If the permissible third-order intermodulation products at maximum signals are no larger than the detection noise floor, the IM from Eq. (9.42) is equal to the system dynamic range (DR). Thus the maximum permissible diffraction efficiency can be written as [from Eq. (9.42) with DR = IM]

$$\eta = 6 \times 10^{-\text{DR}/20} \tag{9.43}$$

The optical power P_s reaching a photodetector can be calculated for the power spectrum case as

$$P_{sp} = P_0 \eta Q = 6QP_0(10^{-\text{DR}/20}) \tag{9.44}$$

where Q is optical transmission.

In the interferometric case, P_s is the same as in Eq. (9.44) except the laser beam has to go through two beam splitters. Therefore, only 25% of P_0 in Eq. (9.44) would be available. Also, the useful signal is produced as interference between the signal and reference beam. The reference beam is generated using a reference Bragg cell whose diffraction efficiency η_r could be from 10 to 70% depending on the fracture limit of the cell. In general, the nonlinearities in the reference cell will not affect the performance of the receiver. The diffracted reference beam spreads over all photodetector channels. If the system frequency resolution is Δf and the total bandwidth is B_R, the channel separation is approximate $2.5B_R/\Delta f$. The factor 2.5 comes from assuming that it takes 2.5 frequency resolutions to separate two signals. Thus, with an identical optical transmission Q, the light intensity from the reference path incident on the photodetector is

$$P_R = QP_0\eta_r \Delta f/(4 \times 2.5B_R) \tag{9.45}$$

A reduction in light intensity by a factor of 4 occurs due to beam splitter–combiner action. The light intensity from the signal path is the same as in Eq. (9.44) except an additional factor of 4 reduction occurs due to the beam splitter–combiner.

The useful signal is the result of interference between light from the signal and reference paths. The peak-to-peak signal is the square root of 4 times the product of reference and signal light intensities at the detector. The rms value of the signal, P_s, can be obtained by dividing the peak-to-peak value by $2\sqrt{2}$ and

is

$$P_{si} = \frac{QP_0[6(10^{-DR/20})\eta_r \Delta f/(2.5B_R)]^{1/2}}{2\sqrt{2}} \tag{9.46}$$

If we assume that the lower limit of the dynamic range is the NEP of the photodetector, the dynamic range in the power spectrum scheme, DR_p, is obtained from Eq. (9.44) by setting $DR = 10 \log(P_{sp}/NEP)$

$$DR_p = \frac{20}{3} \log \frac{6QP_0}{NEP} \tag{9.47}$$

The dynamic range of the interferometric scheme, DR_I, is [from Eq. (9.46)] by setting $DR = 20 \log(P_{si}/NEP)$

$$DR_I = \frac{40}{3} \log \frac{6QP_0\sqrt{\eta_r \Delta f/120B_R}}{NEP} \tag{9.48}$$

The numerical factor in front of Eq. (9.48) is twice that of Eq. (9.47) because of the interferometric effect. Now let us use an example to demonstrate the dynamic range of both cases under the following assumptions.

$$P_0 = 10 \text{ mW} \qquad Q = 0.9 \qquad \eta_r = 0.7 \qquad B_R = 500 \text{ MHz}$$

$$\Delta f = 5 \text{ MHz} \qquad \text{NEP} = -53 \text{ dBm } (5 \times 10^{-6} \text{ mW})$$

The calculated dynamic range of the noninterferometric scheme is 46.9 dB, and the dynamic range of the interferometric scheme is 65.6 dB. The improvement is obvious. The dynamic ranges calculated are referred to the noise floor.

In the above discussion, the two beam splitters are assumed to split the light in half. The question raised here is what is the optimum beam splitting ratio in the Mach–Zehnder interferometric scheme. To answer this question, it is assumed that when the product of the amplitudes from the signal path and the reference path is maximum, the photodetector will generate maximum output. Assume that the fraction of light deflected by the first beam splitter to the signal path is α, so that the fraction transmitted to the reference path is $1 - \alpha$. Similarly, the second beam splitter passes fraction β of the light from the signal path and deflects fraction $1 - \beta$ of the light in the reference path. The light from the signal path to the detector is

$$P_s = P_0 \eta Q \alpha \beta \tag{9.49}$$

The light from the reference path to the detector is

$$P_r = P_0 \eta_r Q(1 - \alpha)(1 - \beta)\, \Delta f/2.5B_R \tag{9.50}$$

The product of Eqs. (9.49) and (9.50) is

$$P_s P_r = P_0^2 \eta \eta_r Q^2 \alpha\beta(1 - \alpha)(1 - \beta)\,\Delta f/2.5B_R \tag{9.51}$$

To find the maximum of Eq. (9.51), one needs to take the partial differentials with respect to α, β and set the results to zero. This yields

$$\alpha = 0.5 \qquad \text{and} \qquad \beta = 0.5 \tag{9.52}$$

which is the same result used earlier in this section, that is, both beam splitters have the same ratio 0.5.

The Bragg cell receiver using the interferometric scheme is still in the development stage. The RF circuits between the photodetectors and the video detectors still need to be studied. The digitizing circuits behind the video detectors also need to be developed; however, the digitizing circuits used in channelized receivers can be used as references.

9.14. TWO-DIMENSIONAL OPTICAL PROCESSOR (52, 53)

Another possible application of Bragg cell receivers in the EW field is to simultaneously measure the frequency and angle of arrival (AOA) of intercepted signals. Since the Bragg cell is basically a two-dimensional device, one dimension can be used for frequency measurement and the other dimension can be used to measure another parameter, for example, AOA. The theory of operation is identical to the AOA interferometry scheme discussed in Chapter 3. Actually, one can consider the Bragg cell approach as a special output processing scheme for the interferometric AOA system. The basic scheme is shown in Figure 9.25. In this example, the Bragg cell has four input transducers, and each one is connected to an antenna. Each transducer generates its individual acoustic wave in the Bragg cell. The phase differences among the four antennas depends on the AOA of the input signal, and the four acoustic waves in the Bragg cell will retain these phase relationships.

If a laser beam is incident on the Bragg cell at the proper angle, the output will read both the frequency and the AOA of the input signal. In Figure 9.25, the ordinate represents the frequency information, and the abscissa represents the angular information. The angle information measured from this scheme is obtained from the phase relation of the four inputs. The angle measured in this arrangement depends on the signal frequency. The higher the input frequency, the more the light deflection along the abscissa, even though the incident angle is the same. The frequency and AOA outputs shown in Figure 9.25*b* exhibit the keystone effect. The AOA information should be calibrated accordingly.

The input transducers on the Bragg cell are not necessarily equally spaced. As discussed in Chapter 3, the smallest antenna spacing is less than one-half of the shortest wavelength to be measured in order to eliminate ambiguity prob-

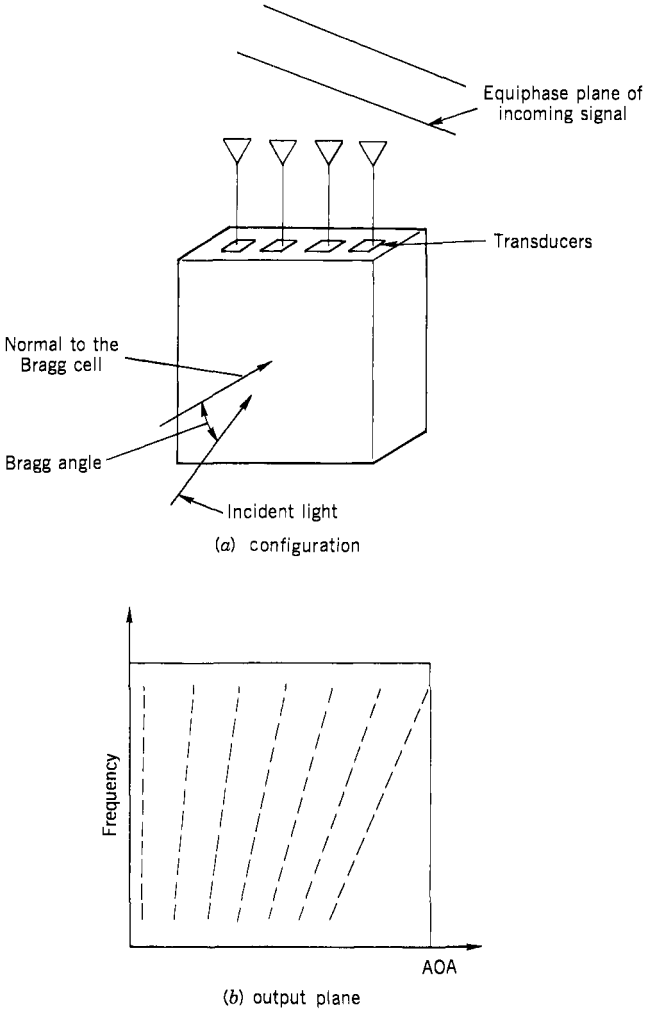


Figure 9.25. Bragg cell used to measure frequency and AOA.

lems. Larger antenna spacings are required to generate fine angle information. If the antennas used on the Bragg cell receiver are nonuniformly spaced, the input transducers on the Bragg cell must also be spaced accordingly; the spacings of the input transducers and the spacings of the antennas must have the same ratio.

Usually the number of antennas used in a phase interferometric system is minimized to make the system simple. The limited number of antennas will generate large side lobes in the AOA measurement. Thus, even if only one signal is present at the input of the two-dimensional Bragg cell receiver, there are many side lobe peaks in the output plane. The peaks are present both in the

frequency and AOA directions. If the Bragg cell receiver processes only one signal at a time, the highest peak will represent the frequency and the AOA of the input signal. If the receiver is used to process simultaneous signals, substantial processing is required to locate the proper peaks to represent the correct frequency and AOA information.

If the output plane features a two-dimensional photodetector array, in addition to the difficulty in locating the correct peaks and avoiding the side lobes, the output rate is also slow. Any input signal in the Bragg cell will be projected on the Fourier plane, but if the photodetector array is clocked out in series, usually it takes several milliseconds to output the information. Since the input signals may stay in the Bragg cell for a very short time, on the order of 1 μsec, and the output time is in milliseconds, the receiver will miss many input signals. In other words, the POI of the two-dimensional Bragg cell receiver is very poor. The high processing speed of a Bragg cell is slowed down by the electronics that follows. Only when the optical outputs at the Fourier plane can be properly processed can the two-dimensional Bragg cell receiver have any practical value for EW applications.

9.15. OTHER PARAMETERS MEASURED BY A BRAGG CELL RECEIVER

The Bragg cell receiver can also be used to measure the AOA of the input signals through an amplitude comparison scheme. The amplitudes of the input signals are retained at the Fourier output plane. Therefore, the relative amplitudes of the input signals can be deduced. If multiple receivers are used with directional antennas facing different directions, the AOA of the input signals can be obtained by amplitude comparison.

The pulse width (PW) and time of arrival (TOA) of the input signal can be measured at the output of a Bragg cell. As mentioned in Chapter 7, when the filter is narrow, it will disturb the PW and TOA measurement. The same is true in Bragg cell receivers. When the window delay time is long, it will disturb the leading and trailing edges of the input pulse. This phenomenon can be explained as follows. Assume that the Bragg cell window is τ and the frequency resolution is approximately $1/\tau$. The PW can be either greater or less than τ. The signal in the Bragg cell is shown in Figure 9.26. In Figure 9.26*a*, the PW is greater than the window time τ. The total time that the signal is in the window equals the PW. In Figure 9.26*b*, the PW is less than the window time τ, and the total time that the pulse is in the Bragg cell equals the window time. Therefore, the minimum PW the receiver can measure is equal to the window time of the Bragg cell.

It should be noted that when the input signal does not fill the entire window, a spectrum spreading effect will occur in which energy is spread to neighboring detectors. This same effect also occurs at the leading and trailing edges of an input pulse.

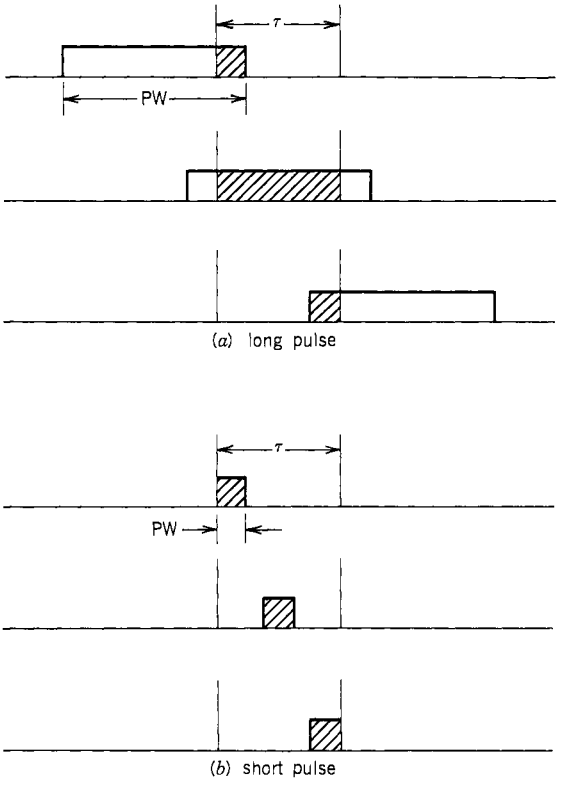


Figure 9.26. Pulse in a Bragg cell.

From Figure 9.26, it is also noted that the amplitude of the input signal should be measured when the Bragg cell window is fully filled. If the signal is too short to fill the window, the PA measurement will be less than the true value. Therefore, to measure all input signals properly, the Bragg cell window should be equal to or less than the minimum PW anticipated. To measure the PA, one needs to sample the outputs from the Bragg cell in short time intervals and compare them to obtain the leveled value, which respresents the true amplitude of the input signal.

9.16. INTEGRATED OPTICAL BRAGG CELL RECEIVERS (54–62)

The power Bragg cell receiver discussed in this chapter can be implemented in an integrated optical circuit (IOC) form. The advantages of using integrated optics for the Bragg cell receiver are the extremely small size and the potentially low production cost. In addition, the IOC Bragg cell receiver will be less susceptible to mechanical vibration. The entire Bragg cell receiver, which includes

laser source, Bragg cell, acoustic transducer, detector array, and optical lens system, can conceivably be integrated on a single chip. At present, all components except the laser and the detector array can be fabricated on the chip directly. The laser and the detectors are fabricated separately and butt-end coupled onto the chip. Therefore, the IOC Bragg cell receiver is potentially more rigid than a conventional optical bench form.

The operating frequency of an IOC Bragg cell receiver is usually lower than that of a bulk wave Bragg cell receive because

1. surface acoustic waves propagate more slowly than bulk waves, and it is easier to sustain a bulk wave at the higher frequency;
2. propagation loss is higher in surface waves than in bulk waves, and the higher the frequency, the higher the loss;
3. the interaction strength between the surface acoustic wave and the light in the waveguide decreases at higher frequencies; and
4. it is difficult to fabricate surface wave transducers capable of operating at high frequencies.

Since the operating frequency is lower for integrated optics, the input bandwidth of an IOC Bragg cell receiver is often narrower because the input bandwidth of a Bragg cell is usually kept less than an octave. However, if the receiver is small and inexpensive, many such receivers can be used in parallel to cover a wider input bandwidth.

The operating principle of the IOC Bragg cell receiver is exactly the same as the conventional bulk wave Bragg cell receiver discussed in Section 9.11. A solid-state diode laser is used as the light source because of its small size and

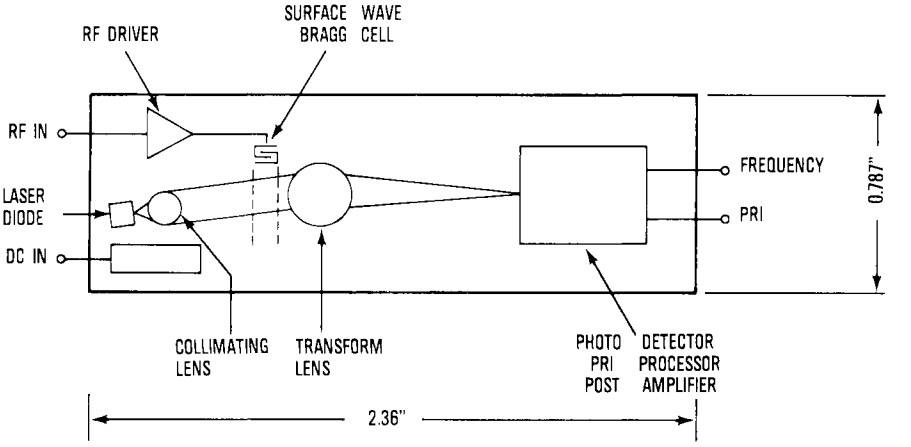


Figure 9.27. Schematic of integrated optic Bragg cell receiver.

high output power. The light is transmitted in a light waveguide rather than in free space. The light waveguide is a thin layer of material with a higher index of refraction than the substrate in which the light is trapped by total internal reflection at the interface of the waveguide. The Bragg cell utilizes surface acoustic waves rather than bulk acoustic waves as the diffracting grating. The surface wave interacts with the light in the waveguide on the surface of the substrate. There are two lenses in an IOC Bragg cell receiver. One lens is used between the laser and the Bragg cell to expand and collimate the laser beam. The other lens is used between the Bragg cell and the detector array as the Fourier transform lens. A photodetector array is often used in an IOC Bragg cell receiver because of its small size. The schematic of a power IOC Bragg cell receiver is shown in Figure 9.27. The entire chip can be a few square centimeters in size.

9.17. COMPONENTS AND PREDICTED PERFORMANCE OF IOC BRAGG CELL RECEIVERS (54–62)

A. Materials

The most commonly used material for fabricating IOC Bragg cell receivers is lithium niobate ($LiNbO_3$) because the material has superior piezoelectric and electro-optic coupling characteristics and therefore is one of the prime materials used for SAW devices and bulk wave Bragg cells. Optical waveguides on the substrate can be produced by diffusion of metals, that is, titanium (Ti), in place of some of the niobium (Nb). The undiffused waveguide is several micrometers thick and exhibits insertion losses of about 0.3–1.0 dB/cm. An electro-acoustic transducer can be deposited on the substrate directly because of the strong piezoelectric coupling, and a frequency bandwidth of 1 GHz may be possible. However, the laser and the photodetector array cannot be fabricated on the same substrate. They must be hybrid coupled to the waveguide in a compact fashion. Butt-end coupling techniques are being developed with approximately 50% efficiency.

Two other materials that have the potential to be used to fabricate IOC Bragg cell receiver are gallium arsenide (GaAs) and silicon (Si). Gallium arsenide material has the potential to be used to fabricate the laser source, the lenses, the Bragg cell (because GaAlAs itself is a piezoelectric material), and the photodetector array. In other words, there is the potential that the entire IOC Bragg cell receiver can be made in monolithic form on a single chip. However, a different material would have to be used for the light waveguide. In addition, the piezoelectric effect is very weak with this material; a piezoelectric material overlay on the substrate may be required to provide the means to generate a surface acoustic wave. Extensive research on this material is required in order to develop a monolithic IOC Bragg cell receiver.

Silicon is an attractive material because silicon wafer processing is a well-established technology in integrated circuit design. Another main advantage of using silicon as the substrate is that the photodetectors can be fabricated directly on the substrate. However, the laser would have to be coupled to the waveguide because a laser cannot be fabricated on the silicon chip. Since silicon absorbs visible light, the optical waveguide cannot be fabricated on the chip directly. A silicon dioxide (SiO_2) layer on top of the substrate to isolate the optical waveguide is required. Corning 7059 glass is a typical material, having insertion losses of 0.2–0.6 dB/cm with an index of reflection of 1.57. Since Si and SiO_2 are not good piezoelectric materials, zinc oxide (ZnO) is generally used as the transducer material.

B. Bragg Cell Transducers

Since the light is concentrated in the optical waveguide in an IOC Bragg cell receiver, the transducer of the Bragg cell is used to generate a surface acoustic wave. Therefore, the basic structure of the transducer is the same as in SAW filters; metal fingers on piezoelectric material are used to generate acoustic waves. To increase the RF bandwidth, multiple transducers, each tuned to a specific RF band, are often used. Each transducer can be fabricated at a slightly different angle with respect to the optical beam to have the proper Bragg angle for matching the operating frequency. This arrangement is referred to as a stagger tilted transducer array and is shown in Figure 9.28.

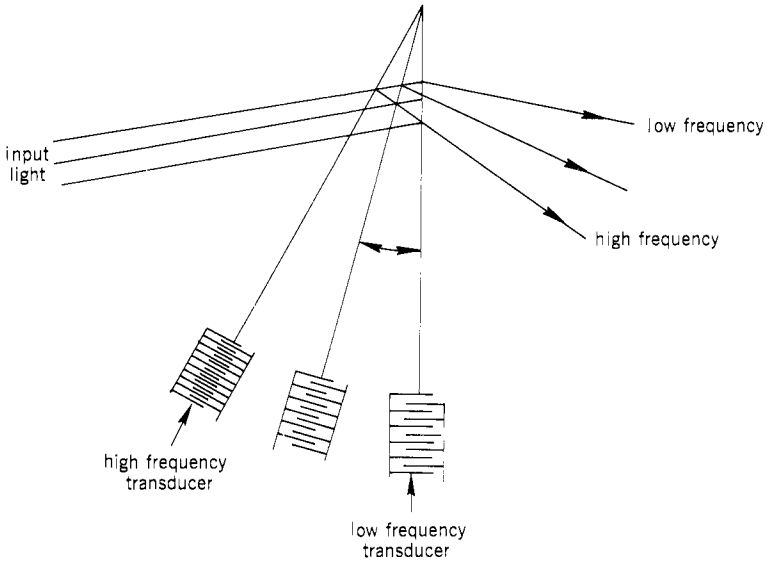


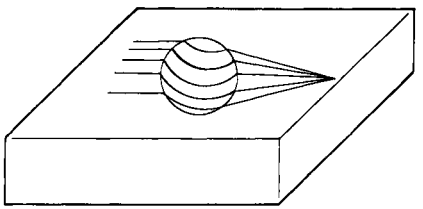
Figure 9.28. Angular stagger tilted transducer array.

C. *Lenses*

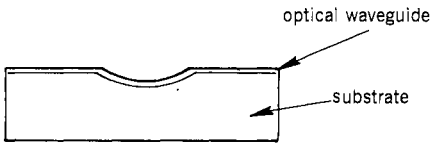
Three basic types of lenses are used in IOC Bragg cell receivers: the Luneburg lens, the geodesic lens, and the holographic lens. The Luneburg lens is formed by depositing a material with an index of refraction higher than that of the waveguide material on top of the waveguide. Light inside the waveguide is refracted by this material and changes the propagation direction. For the silicon substrate with a glass waveguide, Luneburg lenses are made possible by using Ta_2O_5 ($n = 2.1$) and Nb_2O_5 ($n = 2.2$) thin films. However, for GaAlAs and $LiNbO_3$ substrates, since the index of refraction of the waveguide is rather high, it is not presently practical to use Luneburg lenses.

At present, the commonly adapted lens in IOC Bragg cell receivers is a geodesic lens. A geodesic lens physically changes the light path length to achieve a focusing effect. Since this approach is independent of the index of refraction, it can be used on any substrate. A geodesic lens is formed by grinding a quasi-spherical depression in the substrate as shown in Figure 9.29. The light entering the side of the lens travels a shorter distance than the light entering the center of the lens. The differential length of light path focuses the incoming parallel light beam to a point. The grinding and polishing of this kind of lens is tedious and time consuming with existing technology. Sometimes the lens must be fabricated first. Then the positions of the other components, that is, the laser, Bragg cell, and photodetector array, have to be adjusted because the performance of the lens is not easily controlled. The performance of the lenses must be determined experimentally after they are fabricated. Some of these problems are already partially solved.

Another approach in collimating and focusing the light in an IOC Bragg cell receiver is to use holographic lenses. In this approach, a grating is made on the



(*a*) perspective



(*b*) side view

Figure 9.29. Geodesic lens.

surface of the substrate, which diffracts the laser beam to the desired direction with the desired beam form. The grating can be made in two ways: depositing parallel ridges or etching parallel grooves on the substrate surface through a mask. Figure 9.30 shows a linear chirped grating lens focusing a collimated incident guide wave to a point. Theoretically, the holographic lens approach is suitable for mass production because the technology is similar to that used in fabricating integrated circuits. This approach needs further development.

D. Photodetectors

In the existing IOC Bragg cell receiver, the photodetector array is butt coupled at the end of the substrate. At the outputs of the detectors, the light is converted to video signals. To digitize the outputs and further process them, the approaches discussed in Chapter 7 can be used.

E. A Developmental Model

Optical bench assemblies (the portion of the IOC Bragg cell receiver from the up laser to the photodetector array) have been fabricated experimentally (ref. 60). An input bandwidth of 400 MHz with a 4-MHz frequency resolution have been achieved. A series photodetector array was used to demonstrate the feasibility. The dynamic range achieved is comparable to that of a bulk wave Bragg cell receiver. Figure 9.31 shows the actual IOC Bragg cell receiver. Unless the digital processor portion of the Bragg cell receiver can be made compact, the integrated optical bench really does not have much advantage over a bulk wave device in size reduction because the optical portion will be a small percentage of the total receiver. However, if the digital circuits can be made in very large scale integrated (VLSI) circuit form, the size of the optical subsystem may be of some

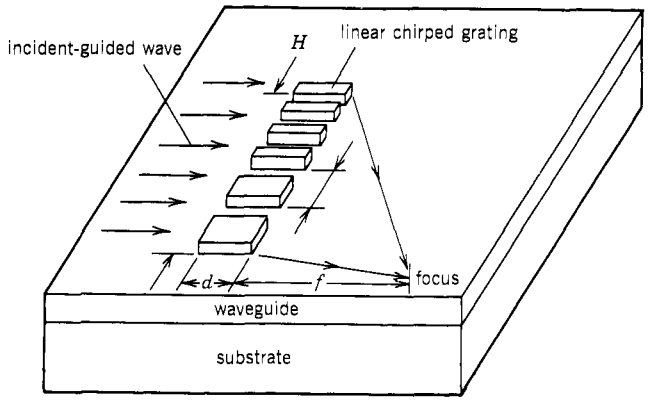


Figure 9.30. Chirped grating lens structure. (Based on Warren et al., ref. 62).

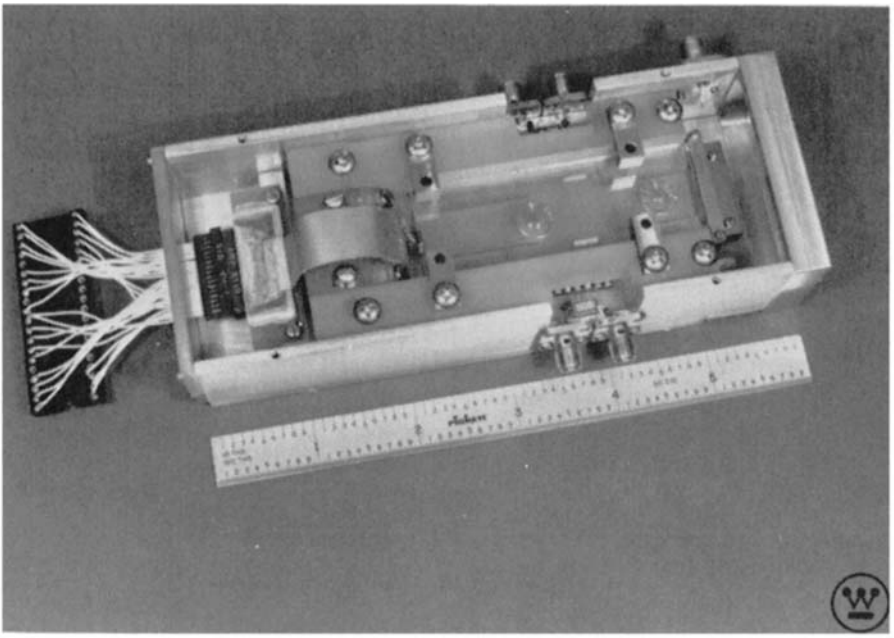


Figure 9.31. IOC Bragg cell receiver. (Fabricated by Westinghouse, courtesy of Avionics Laboratory, AFWAL.)

concern. The major advantage of the IOC Bragg cell receiver over a bulk wave Bragg cell receiver may be the potentially easy fabrication.

9.18. SUMMARY

A Bragg cell receiver has the potential of being used as an EW receiver. It can cover a wide instantaneous bandwidth of possibly over 1 GHz, and the frequency resolution can be very fine. However, to process short pulses, the frequency resolution is governed by the minimum PW of the input signals rather than by the performance of the Bragg cell. At present, the sensitivity and dynamic range of the Bragg cell receiver is limited by the photodetectors. With present detector technology, the dynamic range is rather low for EW receiver applications. Theoretically, the dynamic range of the Bragg cell receiver can be improved with the interferometric approach. However, further development is required to demonstrate the feasibility.

A Bragg cell receiver can be considered a channelized receiver where the channelization is accomplished through optical means. One Bragg cell is equivalent to several hundred filters connected in parallel. In a power Bragg cell receiver, RF amplifiers cannot be added after the frequency channelization to

improve the sensitivity, but light amplification may be added in front of the photodetectors to accomplish this. In an interferometric Bragg cell receiver, RF amplifiers can be added after channelization to improve the sensitivity of the receiver. Therefore, the interferometric Bragg cell receiver is very much like the channelized receiver discussed in Chapter 7.

A Bragg cell receiver has potential performance characteristics that are very desirable for EW applications. Although the feasibility of the receiver has been demonstrated, extensive research and development, especially in the digitizing circuits, is critically needed to realize its full capability. To take the advantage of the small size of the optical bench, the digitizing circuits must be made small in size using VLSI circuit technology.

REFERENCES

1. H. H. Hopkins, On the diffraction theory of optical images, *Proc. Roy. Soc. A.*, **217**, 408 (1953).
2. E. L. O'Neill, Spatial filtering in optics, *IRE Trans Information Theory*, **IT-2**, 56–65 (1956).
3. A. V. Lugt, Coherent optical processing, *Proc. IEEE*, **62**, 1300–1319 (1974).
4. W. T. Maloney, Acoustooptical approaches to radar signal processing, *IEEE Spectrum*, **6**, 40–48 (1969).
5. Analyzer to rely on light waves to detect rf signals, *Electronics*, 29 (December 22, 1977).
6. S. Petiotis, Acousto-optics light the path to broadband ESM receiver design, *Microwaves*, **16** 54 (September 1977).
7. A. E. Spezio, Acousto-optics for electronic warfare applications, Proc. 12th Annual ASILOMAR Conference on Circuits Systems and Computers, November 6–8, 1978.
8. R. A. Coppock, R. F. Croce, and W. L. Regier, Bragg cell RF signal processing, *Microwave J.*, **21**, 62 (September 1978).
9. I. J. Abramovitz, N. J. Berg, and M. W. Casseday, Signal processing with acousto-optics, *Optical Spectra*, **15**, 52 (September 1981).
10. J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, New York, 1968.
11. H. Kogelnik, Coupled wave theory for thick hologram gratings, *Bell System Technical J.*, **18**(9), 2909–2947 (November 1969).
12. I. C. Chang, Acousto-optic devices and applications, *IEEE Trans. Sonics Ultrasonics*, **SU-23**, 2–22 (1976).
13. D. L. Hecht, Multifrequency acousto-optic diffraction, *IEEE Trans. Sonics Ultrasonics*, **SU-24**, 7–18 (1977).
14. D. L. Hecht, Acousto-optic device techniques 400-2300 MHz, *Ultrasonics Symposium Proc. IEEE* (77CH1264-ISU), 1977.
15. D. L. Hecht, Spectrum analysis using acousto-optic devices, *Optical Engineering*, **16**, 461–466 (1977).
16. L. B. Lambert, M. Arm, and A. Aimette, Electro-optical signal processors for phased array antennas, *Optical Electrooptical Information Processing*, J. T. Tippet, Ed., Chapter 38, MIT Press, Cambridge MA, 1965.
17. C. E. Thomas, Optical spectrum analysis of large space bandwidth signals, *Applied Optics*, **5**, 1782–1790 (1966).
18. D. Psaltis and B. V. K. Vijaya Kumar, Acousto-optic spectral estimation: a statistical analysis, *Applied Optics*, **20**, 601–605 (1981).
19. C. F. Quate, C. D. W. Wilkinson, and D. K. Winslow, Interaction of light and microwave sound, *IEEE Proc.*, **53**, 1604–1623 (1965).

20. A. H. Meitzler, Piezoelectric transducer materials and techniques for ultrasonic devices operating above 100 MHz, *Ultrasonic Transducer Materials*, O. E. Mathial, Ed., Plenum, New York, 1971.
21. J. D. Larson III and D. K. Wilson, Ultrasonically welded piezoelectric transducers, *IEEE Trans. Sonics Ultrasonics*, **SU-18**, 142–152 (1971).
22. D. J. Torrieri, Signal processing theory of Bragg cells, U.S. Army Material Development and Readiness Command, Technical Report CM/CCM-81-1, May 1981.
23. E. K. Sittig, Design and technology of piezoelectric transducers for frequencies above 100 MHz, *Physical Acoustics*, Vol. IX, W. P. Mason and R. N. Thurston, Ed., Academic Press, New York, 1972.
24. D. A. Pinnow, Guided lines for the selection of acousto-optic materials, *IEEE Trans. Quantum Electron*, **QE-6**, 223–238 (1970).
25. N. Uchida and N. Niizeki, Acousto-optic deflection materials and techniques, *Proc. IEEE*, **61**, 1073–1092 (1973).
26. I. C. Chang and D. L. Hecht, Doubling acousto-optic deflector resolution utilizing second-order birefringement diffraction, *Applied Physics Letters*, **27**, 517–518 (1975).
27. M. J. Delaney and S. K. Yao, Wideband acoustic-optic Bragg cell, IEEE Ultrasonics Symposium Proc., pp. 408–412, San Diego, CA October 1982.
28. W. R. Klein and B. D. Cook, Unified approach to ultrasonic light deffraction, *IEEE Trans. Sonics Ultrasonics*, **SU-14**, 123–134 (1967).
29. E. D. Young and S. K. Yao, Design considerations for acousto-optic devices, *Proc. IEEE*, **69**, 54–64 (1981).
30. RCA electro-optics handbook, Chapter 10, Technical Series EOH-11 RCA/Commercial Engineering, Harrison, NJ, 1974.
31. W. L. Wolfe and G. J. Zissis, *The Infrared Handbook*, Chapters 11, 12, Office of Naval Research, Department of the Navy, Arlington, VA, 1978.
32. S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed., Chapter 13, Wiley, New York, 1981.
33. J. Muller, Thin silicon film *p-i-n* photodetector using internal reflection method, *IEEE Trans. Electron Devices*, **ED-17**, 342–347 (1970).
34. K. M. Johnson, High speed photodiode signal enhancement at avalanche breakdown voltage, *IEEE Trans. Electron Devices*, **ED-12**, 55–63 (1965).
35. L. K. Anderson and B. J. McMurtry, High speed photodetectors, *IEEE Proc.*, **54**, 1335–1349 (1966).
36. N. K. Sheridon and M. A. Berkovitz, "Optical processing with the Ruticon," *Optical Information Processing*, Vol. 83, pp. 68–75, Society of Photo-Optical Eng., Bellingham, WA, 1976.
37. J. P. Y. Lee, Acousto-optic spectrum analysis of radar signals using an integrating photodetector array, *Applied Optics*, **20**, 595–600, (1981).
38. G. M. Borsuk, Photo detectors for acousto-optic signal processors, *Proc. IEEE*, **69**, 100–118 (1981).
39. D. F. Barbe, Imaging devices using the charge-coupled concept, *Proc. IEEE*, **63**(1), 38–67 (1975).
40. R. H. Dyck and W. Steffe, Effects of optical crosstalk in CCD image sensors, *Proc. International Conf. Applications of CCD's*, San Diego, CA, pp. 1–55, October 1978.
41. H. Hodara, Ed., *Fiber and Integrated Optics*, Vol. 1, Crane, Russak, New York, 1977.
42. J. J. Pan, Fiber optics marches into microwave systems, *Microwave J.*, 93 (August 1982).
43. A. Yariv, *Quantum Electronics*, 2nd ed., Wiley, New York, 1975.
44. C. A. Burrus, H. C. Casey, Jr., and T. Y. Li, "Optical sources," in S. E. Miller and A. G. Chynoweth, Eds., *Optical Fiber Communication*, Academic Press, New York, 1979.
45. S. Jacobs, M. Sargent, J. F. Scott, and M. O. Scully, Eds., *Laser Applications to Optics and Spectroscopy*, Addison-Wesley, Reading, MA, 1975.
46. A. Yariv, *Introduction to Optical Electronics*, 2nd ed., Holt, Rinehart and Winston, New York, 1976.
47. T. M. Turpin, Spectrum analysis using optical processing, *IEEE Proc.*, **69**, 79–82 (1981).

48. M. W. Casseday, Wideband acousto-optic Bragg cells, *IEEE Ultrasonics Symposium Proc.* (BICH1689-9), pp. 735–739, 1981.
49. L. C. Lennert, I. C. Chang, D. L. Steinmetz, W. Brooks, and F. Langdon, A high-performance GH-bandwidth acousto-optic spectrum analyzer, *SPIE Proc. Int. Society Optical Eng.*, **352**, 10–16 (1983).
50. M. L. Shah, E. H. Young, A. Vander Lugt, and M. Hamilton, Interferomatric Bragg cell spectrum analyzer, *IEEE Ultrasonics Symposium Proc.*, pp. 743–745, 1981.
51. M. L. Shah, J. R. Teague, R. V. Belfatto, D. W. Thamson, and E. H. Young, Wideband interferometric acousto optic Bragg cell spectrum analyzer, *IEEE Ultrasonics Symposium Proc.*, pp. 740–742, 1981.
52. R. F. Croce, Acousto-optic processor for simultaneous spatial spatial and spectral analysis, M1641, GTE Sylvania Electro-optics Organization, April 1977.
53. Angle of arrival sensors using acousto-optical processing, Applied Technology Division of Itek Corp., N00173-76-C-0333 (A002), March 1977.
54. M. L. Dakss, L. Kuhn, P. F. Heidrich, and B. A. Scott, Grating coupler for efficient excitation of optical guided waves in thin films, *Applied Physics Letters*, **16**, 523–525 (1970).
55. D. B. Anderson, R. L. Davis, J. T. Boyd, and R. R. August, Comparison of optical waveguide lens technologies, *IEEE Trans. Quantum Electronics*, **QE-13**, 275–282 (1977).
56. J. T. Boyd and C. L. Chen, An integrated optical waveguide and charge-coupled device image array, *IEEE Trans. Quantum Electronics*, **QE-13**, 282–287 (1977).
57. M. C. Hamilton, D. A. Wille, and W. J. Miceli, An integrated optical RF spectrum analyzer, IEEE Ultrasonics Symposium, October 1976.
58. D. B. Anderson, J. T. Boyd, M. C. Hamilton, and R. R. August, An integrated-optical approach to the Fourier transform, *IEEE Trans. Quantum Electronics*, **QE-13**, 268–275 (1977).
59. D. B. Anderson, Integrated spectrum analyzer: an imminent chip, *IEEE Spectrum*, **15**, 22–29 (1978).
60. M. C. Hamilton, Wide-band acousto-optic receiver technology of electronic warfare systems, *J. Electronic Defense*, 50–55 (January/February 1981).
61. G. C. Righini, V. Russo, S. Sottini, and G. Toraldo di Francia, Geodesic lenses for guided optical waves, *Applied Optics*, **12**, 1477–1481 (1973).
62. C. Warren, S. Forouhar, W. S. C. Chang, and S. K. Yao, Double ion exchanged chirp grating lens in lithium niobate waveguides, *Applied Physics Letters*, **43**, 424–426 (1983).

Chapter 10

Hybrid and Cueing Receivers

10.1. INTRODUCTION

The utilization of several kinds of receivers combined together to accomplish a certain desired performance will be discussed in this chapter. In the previous chapters, different kinds of receivers have been discussed. Each kind has its unique characteristics. The characteristics of the six generic receivers are briefly summarized below for convenient reference.

Crystal Video Receiver. The structure of this receiver is very simple; the sensitivity and dynamic range are low; it does not provide frequency information and cannot separate simultaneous signals; it has wide-input bandwidth and is often used to measure angle of arrival (AOA) information through an amplitude comparison scheme.

Superheterodyne (Superhet) Receiver. The structure of this receiver is moderately complex; the sensitivity and dynamic range are extremely high; the input bandwidth is narrow; and it is suitable for measuring continuous-wave (CW) signals and isolating a certain signal in a narrow frequency range.

Instantaneous Frequency Measurement (IFM) Receiver. The structure of this receiver is relatively simple; the sensitivity and dynamic range are moderately high; the input bandwidth is very wide; it can measure short pulses with high-frequency resolution, but it may generate erroneous frequency information when simultaneous signals are present.

Channelized Receiver. The structure of the channelized receiver is very complicated and the size is large; the sensitivity and dynamic range are high; the input bandwidth depends on the number of parallel channels; the larger the number of parallel channels, the wider the input bandwidth; it can process simultaneous signals. The frequency resolution of this receiver depends on the bandwidth of the filters.

Compressive Receiver. The structure of this receiver is complicated, and high-speed logic circuits are required; the sensitivity and dynamic range are moderately high; the input bandwidth is moderately wide; it can process simultaneous signals; this receiver is still in the experimental stage.

Bragg Cell Receiver. The structure of the Bragg cell receiver is complicated, but the size is potentially small; the sensitivity is high; the dynamic range is low with the conventional approach; the input bandwidth is wide; it can process simultaneous signals. Further developments are required to realize the potential of this kind of receiver.

To accomplish some specific missions, often one kind of receiver cannot fulfill the requirements. But the combination of a few of the above receivers may produce the desired performance. Since some of the receivers are still in the development stage, their overall performance has not been fully explored. Therefore, there is no certain way to predict how the receivers should be combined. Some examples will be used to demonstrate the utilization of combined receivers to achieve a certain performance.

10.2. CHANNELIZED–IFM RECEIVER COMBINATION

In an EW receiver, if the desired instantaneous input bandwidth is very wide and the frequency resolution is very fine, a channelized, compressive, or Bragg cell receiver can be used to fulfill such a requirement. If a channelized or Bragg cell receiver is used, the number of parallel channels is very large. For example, if the input bandwidth is 1 GHz and a frequency resolution of 1 MHz is desired, a receiver with approximately 1000 channels is required, which will be very complicated and large in size. If a compressive receiver is used, the time–bandwidth product of the dispersive delay line (DDL) is very large, which is difficult to accomplish with present technology. In addition, the logic speed used in the digitizing circuits is also very high. If a Bragg cell is used, its time–bandwidth product are very large, and the number of parallel channels which is beyond state-of-the-art technology.

The pulse width (PW) to be processed by an EW receiver usually ranges from less than 1 μsec to CW. A narrow-band filter will introduce transient effects on both the leading and trailing edges of the signal. If the input signal has a long PW and the transient period occurs only over a small portion of the entire

pulse, the above three kinds of receivers will measure the fine-frequency resolution without much problem. However, when a short pulse arrives at the receiver and the transient period introduced is over the entire pulse, it is very difficult, if not impossible, for the receiver to determine the signal frequency accurately. Therefore, channelized, compressive, or Bragg cell receivers are not suitable for generating fine-frequency resolution on short pulses.

It has been discussed in Chapter 6 that an IFM receiver can measure short pulsed signals with fine-frequency resolution. However, it may generate erroneous frequency information when simultaneous signals are present at the input of the receiver. Using the example discussed above, one can design a receiver using ten 100-MHz IFM receivers to cover the total 1-GHz bandwidth. In front of each IFM receiver, there is a passband filter matching the bandwidth of the IFM receiver to prevent out-of-band signals from reaching the IFM receiver, as shown in Figure 10.1.

In front of the receiver there are bandpass filters and radio-frequency (RF) amplifiers. The outputs of the amplifiers are also fed to a video processor to generate PA, PW, and TOA information. The data generated by each individual IFM receiver will be encoded and passed to a digital signal processor. In this arrangement, the probability of simultaneous signals at the input of the receiver is greatly reduced because of the narrow-bandwidth filters (about 100 MHz) in front of the IFM receivers. Since the IFM receiver only covers 100 MHz bandwidth, the structure of the IFM receiver can be relatively simple. An IFM receiver with one single delay line may be enough to produce the required frequency resolution.

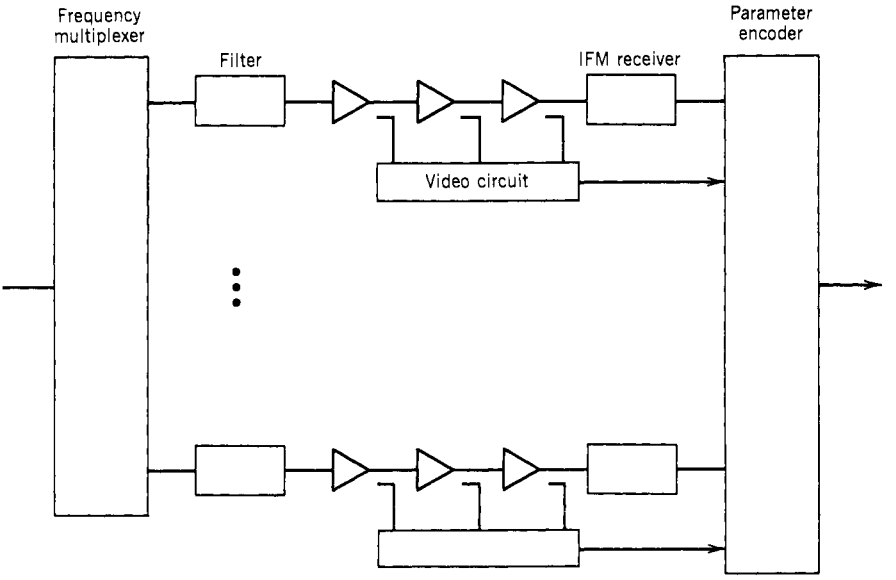


Figure 10.1. A basic channelized–IFM receiver.

If 20 filters and IFM receivers are used to build the channelized–IFM receiver combination to cover the same 1-GHz bandwidth, then the performance of the receiver will be slightly better than the above approach because the probability of simultaneous signals present at the input of the IFM receivers will be reduced because of the narrower filter bandwidth. To determine the optimum number of filters and IFM receivers, examine in detail the combination of a filter bank followed by IFM receivers.

In a conventional IFM receiver, the filter in front of the phase correlator is very wide, on the order of gigahertz. This wide-bandpass filter will not produce significant distortion to the input signals to cause frequency measurement difficulty. The relatively narrow bandwidth filters used in a channelized–IFM approach may cause frequency measurement difficulty.

To determine the bandwidth of the filters in front of the IFM receiver, one must choose between two parameters: instantaneous dynamic range and minimum PW. If one parameter is chosen, the other one will be tightly restricted. If a filter is chosen according to the desired instantaneous dynamic range, the frequency versus insertion loss of the filter is determined. Thus the bandwidth and the number of poles of the filters are determined. The transient period at the leading and trailing edges of the pulse are also determined. The IFM receiver should measure the frequency of the signal after the leading edge transient and before the trailing edge transient. If the pulse is too short, the leading and trailing edge transients will be close together. When the steady state between the two transients is too short to be measured, the receiver cannot properly process the signal anymore. Therefore, when the dynamic range of a channelized–IFM receiver is determined, its capability to process minimum PW is also determined.

If the filters in front of the IFM receivers are chosen by the minimum PW requirements, the transient periods of the leading and trailing edges on the input signals caused by the filters must be limited. This choice in turn determines the bandwidth and the number of poles of the filter. In other words, the shape of the filters is also determined. Thus, if the minimum PW is decided, the instantaneous dynamic range of the receiver is also determined.

Another problem of a channelized–IFM receiver is to determine how the receiver processes an input signal when the signal falls between two adjacent channels. An ideal approach would be to have a sharp boundary between the two filters. If the input signal falls on one side of the boundary, one IFM receiver will process it, otherwise another IFM receiver will process it. At any time, one signal is never processed by two IFM receivers. This approach can be accomplished by comparing the output of two adjacent filters to determine which filter has the stronger output. Then an IFM receiver can be assigned to measure the input signal. However, if two signals are present at two adjacent channels, the weak one may be missed by the receiver.

In another approach, in order not to miss an input signal in an adjacent filter, two adjacent IFM receivers must respond to the input signal. In other words, one input signal between two adjacent filters will be processed by

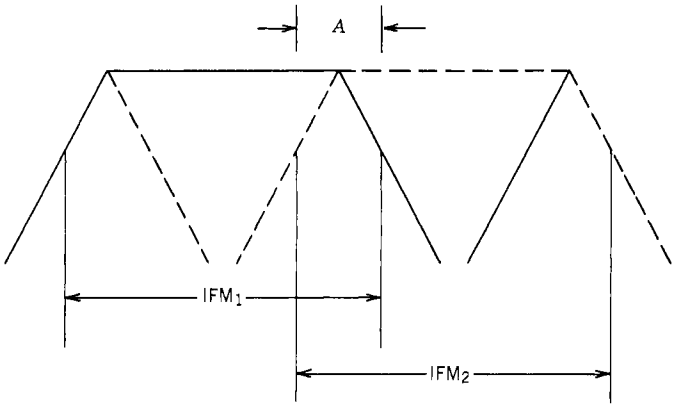


Figure 10.2. IFM receivers to measure signals in the overlap region of a channelized receiver.

two IFM receivers. One obvious disadvantage of this approach is that one IFM receiver will be generating redundant information. Worse yet, if the two IFM receivers generate slightly different frequency information, the receiver may report two signals when only one signal is present. The additional signal reported will confuse the digital processor following the receiver.

One way to solve this problem is to make IFM receivers slightly wider than the channel filter width, as shown in Figure 10.2. The overlapped regions is *A*. If a signal falls in region *A*, both IFM receivers will process the signal. After the frequency is measured by two adjacent IFM receivers, a decision will be made whether to report the frequency reading of one of the IFM receivers and suppress the other. When the signal is inside region *A*, only one frequency will be reported. When two frequency readings are not in region *A*, then both frequencies should be reported.

10.3. TRANSIENT RESPONSES AFTER FILTERS (1–3)

As discussed in Section 7.2, when a pulsed signal passes through a filter, the filter will cause the signal to ring at the leading and trailing edges of the pulse, and this is usually referred to as the transient effect. If an IFM receiver makes the measurement during the transient, a wrong frequency may be recorded. Figure 10.3 shows the frequency responses of several Butterworth filters. The filters are all centered at 800 MHz with a bandwidth of 80 MHz. In Figure 10.3*a*, *b*, *c*, and *d*, the filter has 6, 8, 10, and 12 poles, respectively.

The transient effect can be calculated through a fast Fourier transform (FFT) program if the input and the transfer function of the filter is known. Figure 10.4 shows the amplitude responses of a pulsed signal of 0.5 μsec, and the carrier frequency of the signal is at 740 MHz, which is 60 MHz from the center of the filter. Figure 10.4*a* shows the input signal. Figures 10.4*b*, *c*, *d*, and *e* show the outputs of the signal from the 6-, 8-, 10-, and 12-pole filters.

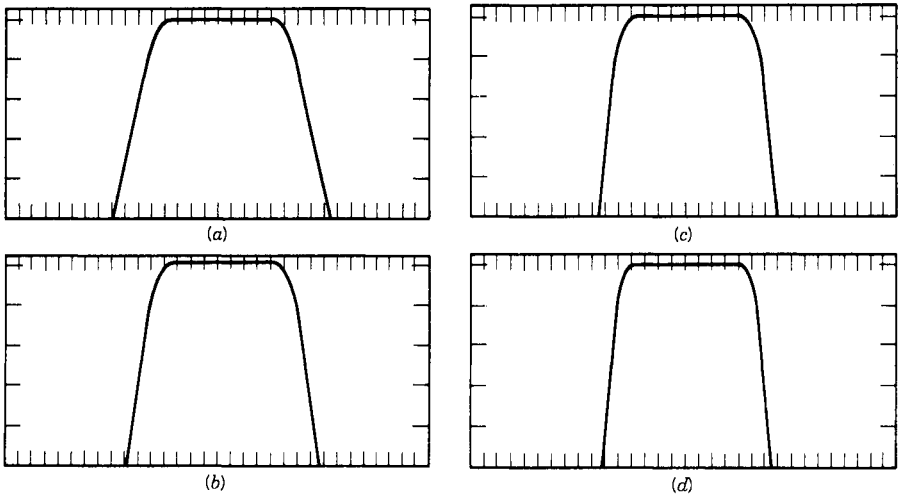


Figure 10.3. Frequency response of filters: (*a*) 6-pole filter; (*b*) 8-pole filter; (*c*) 10-pole filter; (*d*) 120-pole filter.

From these results, it is obvious that a filter with more poles has more out-of-band rejection, but the transient period is also longer. The out-of-band rejection of the filter is directly related to the instantaneous dynamic range of the receiver. The more out-of-band rejection a filter can provide, the higher the instantaneous dynamic range of the receiver, especially for two input signals close in frequency. Filters with narrower bandwidth can separate signals close in frequency; however, the transient effect is also prominent. These results are consistent with the discussions presented in previous sections.

Now let us discuss the frequency change after a pulsed signal passes a band-pass filter. Theoretically, the frequency can be obtained by the zero crossing of the output FFT results. The instantaneous frequency equals the inverse of the time between two adjacent crossings. However, accurate zero crossings are difficult to obtain from the FFT outputs. The following method (based on ref. 1) can be used to find the instantaneous frequency.

Taking the signal center frequency ω_0 as a reference, the asymmetrical low-pass filter transfer function can be written for an offset of $\Delta\omega$ as

$$H(\omega - \Delta\omega) = H_p(\omega - \Delta\omega) + jH_q(\omega - \Delta\omega) \tag{10.1}$$

where $\Delta\omega$ is the difference frequency between the frequency of the filter and the signal frequency, $H(\omega)$ is the filter low-pass equivalent transfer function, and $H_p(\omega - \Delta\omega)$ and $H_q(\omega - \Delta\omega)$ are the low-pass transfer functions of the in-phase and quadratic components of the impulse response (ref. 2).

The h_p, h_q components of the asymmetrical low-pass impulse response can be determined directly by recognizing the Fourier transform pair given as

$$H(\omega - \Delta\omega) \leftrightarrow \exp(-j\,\Delta\omega t)h(t) = h_p(t) + jh_q(t) \tag{10.2}$$

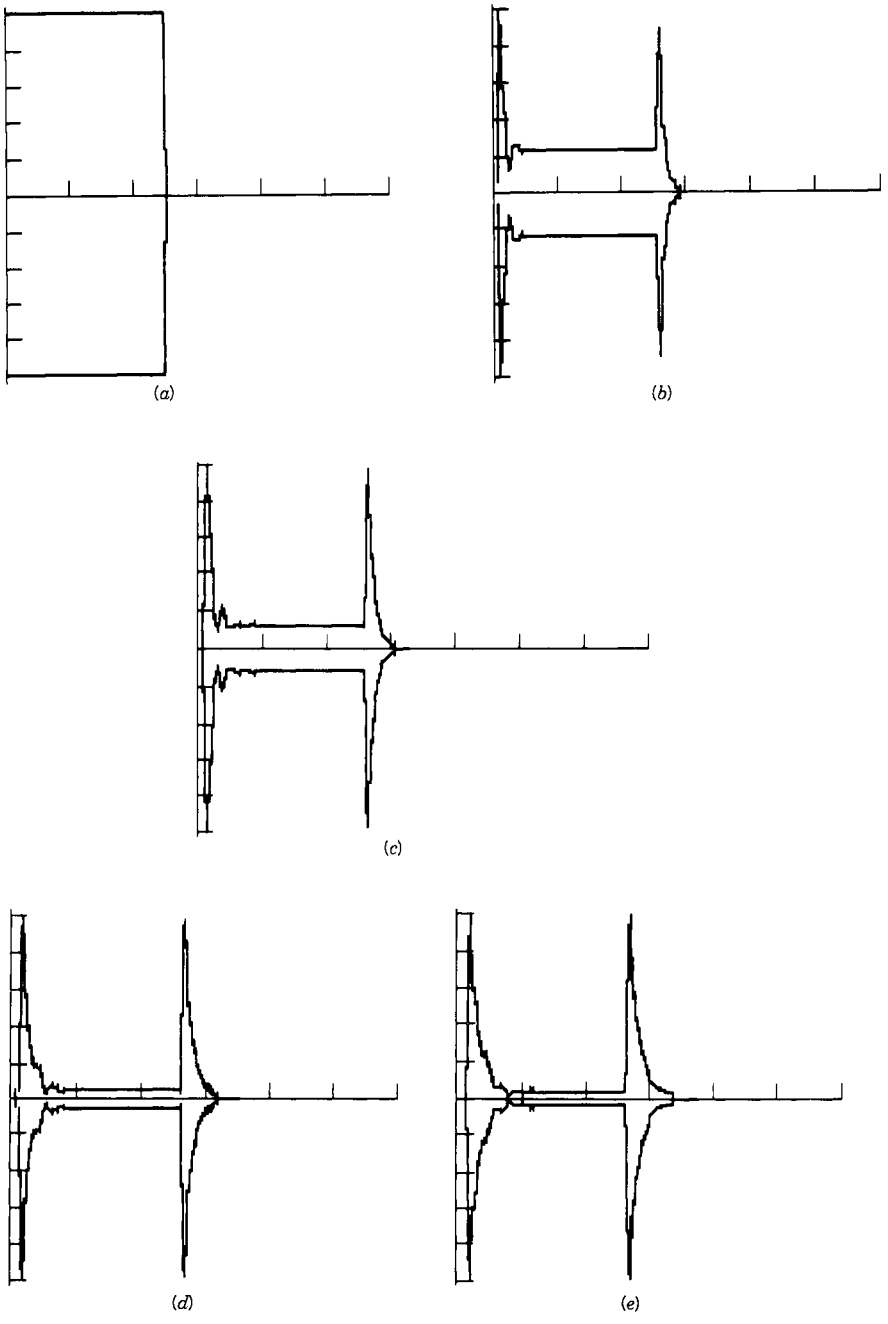


Figure 10.4. Outputs of filters with a center frequency at 800 MHz for a 0.5μs pulse with a carrier frequency of 740 MHz: (*a*) input signal; (*b*) 6-pole filter; (*c*) 8-pole filter; (*d*) 10-pole filter: (*e*) 12-pole filter.

where $h(t)$ is the low-pass impulse response of the filter. Using Euler's identity, the asymmetrical $h(t)$ reduces to

$$h_p(t) = h(t)\cos(\Delta\omega\, t) \qquad h_q(t) = -h(t)\sin(\Delta\omega\, t) \tag{10.3}$$

The operation defined in Eq. (10.3) forms the basis for the asymmetrical transient analysis.

In general, a signal input with in-phase and quadrature components $[p_s(t)$ and $q_s(t)]$ applied to a filter with complex impulse responses $[h_p(t)$ and $h_q(t)]$ will give in-phase and quadrature output components $[y_p(t)$ and $y_q(t)]$ with $*$ representing convolution,

$$\begin{aligned} y_p(t) + jy_q(t) &= (p_s + jq_s) * (h_p + jh_q) \\ &= (p_s * h_p - q_s * h_q) + j(p_s * h_q + q_s * h_p) \end{aligned} \tag{10.4}$$

or

$$y_p(t) = p_s * h_p - q_s * h_q, \qquad y_q(t) = p_s * h_q + q_s * h_p$$

The filter output $y(t)$ referenced to the signal center frequency ω_0 can be written as

$$y(t) = 2y_p(t)\cos(\omega_0 t) - 2y_q(t)\sin(\omega_0 t) \tag{10.5}$$

which can be represented by a phasor as

$$y(t) = 2\sqrt{y_p^2 + y_q^2}\cos[\omega_0 t + \tan^{-1}(y_q/y_p)] \tag{10.6}$$

where the envelope $e(t)$ is given by

$$e(t) = 2\sqrt{y_p^2 + y_q^2} \tag{10.7}$$

and the phase angle $\theta(t)$ as

$$\theta(t) = \tan^{-1}(y_q/y_p) \tag{10.8}$$

The instantaneous frequency $\omega_i(t)$ can be obtained by differentiating Eq. (10.8) with respect to t:

$$\omega_i(t) = \frac{d\theta}{dt} = \left(y_p \frac{dy_q}{dt} - \frac{dy_p}{dt} y_q\right) \Big/ (y_p^2 + y_q^2) \tag{10.9}$$

The frequency responses of the same signal passing the filters in Figure 10.3 are shown in Figure 10.5. In this figure, the input signal frequency is used as reference.

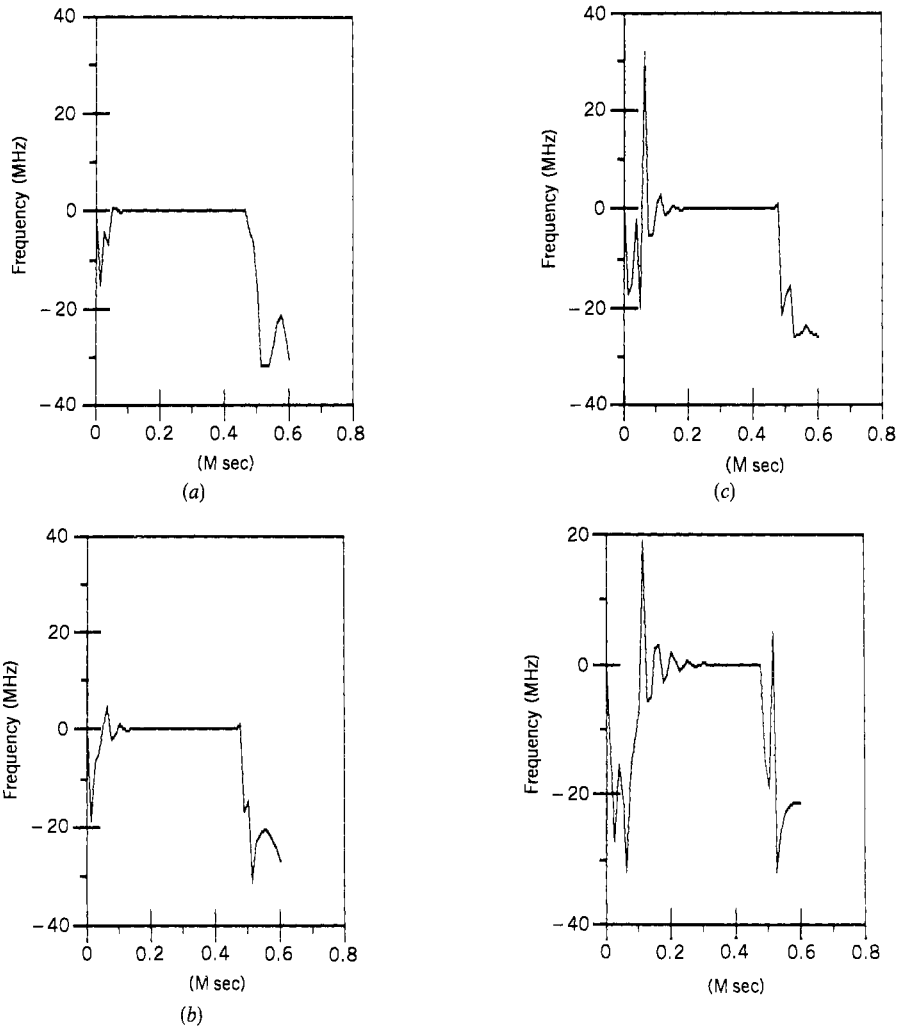


Figure 10.5. Instantaneous frequency of a signal pass through a filter. (*a*) 6-pole filter, (*b*) 8-pole filter, (*c*) 10-pole filter, (*d*) 12-pole filter.

These results, obtained from numerical analysis, will provide the input to the IFM receiver following the filter. As expected, the higher the number of poles, the more prominent the frequency shift and the longer the transient time. On the trailing edge, the amplitude of the signal is rather weak after 500 nsec, especially for the filter with a large number of poles. The frequency shifts after 550 nsec can be neglected because the corresponding amplitudes will be in the noise level. The amplitude and frequency responses from the filter outputs are the key information for designing a channelized IFM receiver.

10.4. CUEING RECEIVER CONCEPT

A cueing receiver can be considered to be a receiving system because it usually contains several kinds of receivers. The concept of a cueing receiver is to use one kind of receiver to obtain some coarse information on the input signals, and then use a different receiver in the receiving system to measure the detailed information on the signals. Thus, a cueing receiver is another example of the hybrid receiver concept applied to an EW receiver system. Like the hybrid receiver discussed above, a cueing receiver can have many different approaches. The system design depends on the specific requirements.

Let us use another example to discuss the concept of cueing receivers. In this special case, the system is designed to measure the AOA of the input signals. It was stated in Chapter 3 that it takes a very complicated receiver system to measure the AOA of an input signal. The instantaneous input bandwidth of the receiver must be wide enough to intercept the signals, and the AOA information should not be contaminated by simultaneous signals. Multiple channelized, compressive, or Bragg cell receivers with many antennas can be used to accomplish such a requirement. However, the size and cost of such a receiver system will be very high because a number of these receivers must be used either through amplitude or phase comparison to obtain the AOA information. These receivers are complicated in design, and it is difficult to obtain amplitude or phase matching among them.

A four-quadrant amplitude comparison system with four crystal video receivers can measure AOA over a wide instantaneous bandwidth, but it is susceptible to simultaneous signal contamination. On the other hand, a four-quadrant superhet receiver system with a relatively narrow bandwidth can generate AOA information with less simultaneous signal contamination, but it has low probability of intercept (POI). To combine the wide input frequency bandwidth with low probability of simultaneous signals, a cueing idea can be adapted.

The concept of a cueing receiver system is shown in Figure 10.6. This system consists of one omnidirectional antenna and four quadrant antennas. The receiver portion can be divided into three parts: the coarse-frequency measurement unit, the logic and local oscillator (LO) unit, and the AOA measurement unit.

The input signals from the omnidirectional antenna are first sorted by the coarse-frequency unit, which is a wide-band receiver with simultaneous signal capability. The coarse-frequency unit can be a channelized, compressive, or Bragg cell receiver. After the coarse frequency is determined, the frequency information can be sent into the steering logic unit, which in turn selects a proper frequency from the LO unit. The frequency from the LO will convert the input frequency to the passband of the AOA unit, which could be narrow superhet receivers. These superhet receivers can be used to measure the AOA through either amplitude or phase comparison. This arrangement will fulfill the requirements with one wide-bandwidth receiver having simultaneous signal capability and several narrow-band receivers.

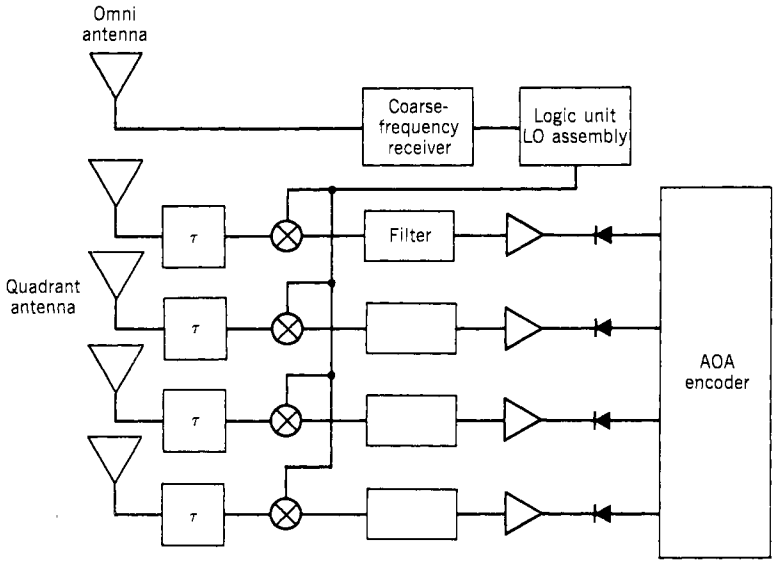


Figure 10.6. Cueing receiver system concept to measure AOA information.

The coarse-frequency measurement, the steering logic circuits, and selecting the LO frequency take finite time. To not miss the input signals, the input signals must be properly delayed. When the input signal reaches the AOA measurement unit, all LO frequencies are properly set. The example in Figure 10.6 can only measure the AOA of one input signal. If multiple-signal capability is needed, multiple superhet receivers are needed, and the receiver system will be rather complicated.

The concept of the cueing receiver is not new. However, its development is only starting because there were no available RF delay lines to store the RF signals temporarily. The advances in RF delay lines make the cueing receiver possible. Different types of delay lines will be discussed in the following sections.

10.5. GENERAL CHARACTERISTICS OF DELAY LINES (4)

In the cueing receiver system discussed above, one of the key components is the delay lines. It is very difficult to fabricate wide-band RF delay lines with microsecond delay times and very low insertion losses. It is desirable to have the delay line bandwidth cover the entire bandwidth of the receiver; otherwise, a number of delay lines will be used in parallel to cover the entire bandwidth, and the system will be complicated.

In a cueing AOA system, if the AOA is measured through a quadrant amplitude comparison system, the amplitude response of all the quadrant delay lines must be matched over the operating frequency range and the temperature

range. The amplitude mismatch among different quadrants can be stored in memory for calibration. If the AOA information is measured by phase comparison, the phases of all the quadrants, including all the delay lines, must be matched over the operating frequency and temperature ranges. For example, if a delay line operates at 1 GHz with a delay time of 2 μsec, the total delay measured in wavelength is 2000 ($10^9 \times 2 \times 10^{-6}$). Since each wavelength has 360°, the total phase change in the delay line is 7.2×10^5 (360×2000). If the phase match in an interferometric system is 10°, the accuracy of the phase match must be 1.39×10^{-5}. In other words, the delay line length must be controlled within this limit. The phase error can be measured, stored, and used for calibration. However, the phase error among all the different quadrants must be within 180°. If the phase error is over 180°, even calibration cannot correct it anymore. Especially if the receiver system is required to operate in a wide temperature range, the delay lines must be carefully designed and the delay lines should be packed closely together. It is probably advisable to use calibration signal sources to occasionally check the system.

10.6. ELECTROMAGNETIC DELAY LINES (5, 6)

There are many different kinds of delay lines. They are the electromagnetic (EM), SAW, bulk wave, magnetostatic surface wave, and fiber optic delay lines.

An EM delay line can be a coaxial cable. The delay time is proportional to the length of the line. The velocity of the wave propagation for a transverse electromagnetic (TEM) mode is

$$V = C/\sqrt{\varepsilon_r} \tag{10.10}$$

where C is the velocity of light and ε_r is the relative dielectric constant of the insulator between the two conductors.

If the length of the line is l, the delay time

$$\tau = l/V = l\sqrt{\varepsilon_r}/C \tag{10.11}$$

The delay time can be increased by increasing l or ε_r.

The characteristic impedance of the line is

$$Z = \frac{60}{\sqrt{\varepsilon_r}} \ln \frac{b}{a} \quad \Omega \tag{10.12}$$

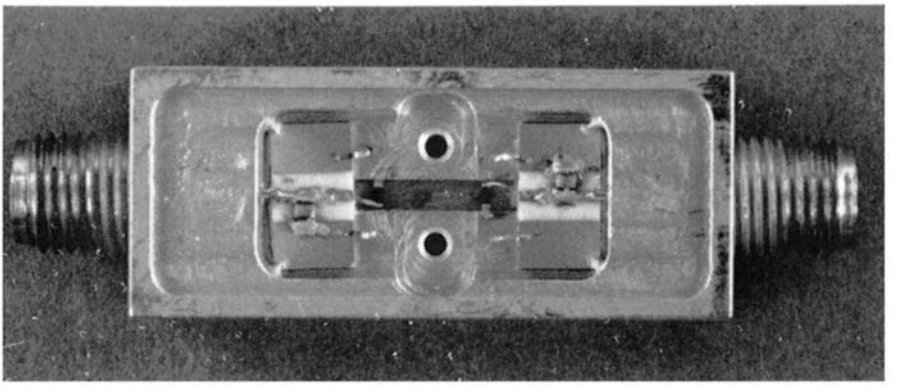
where b is the radius of the inner wall of the outer conductor and a is radius of the center conductor.

The insertion loss of the delay line is proportional to the delay time. Since in a cueing receiver system the delay time is rather long, in the microsecond

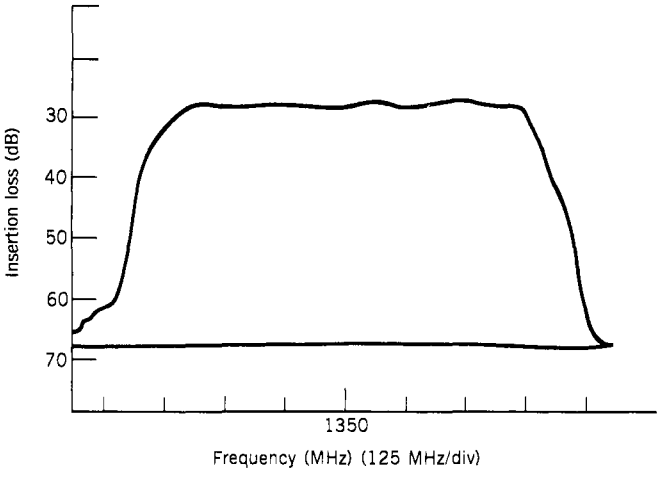
range, an EM delay line is usually not suitable to this application because the length of line is long and the insertion loss is high.

10.7. SURFACE ACOUSTIC WAVE (SAW) DELAY LINES (7–11)

A SAW delay line can be considered to be very wide band SAW filter (as discussed in Section 7.7). The input electric signal is changed into a SAW through a transducer, and the acoustic wave travels on the substrate and is changed back into an electric signal at the output by another transducer. The operating



(*a*)



(*b*)

Figure 10.7. SAW delay line. (*a*) delay line (*b*) insertion loss versus frequency (Courtesy of Texas Instruments Inc.)

frequency of a SAW delay line is higher than a SAW filter because it does not need long transducers to obtain the desired filter bandwidth. In general, the operating frequency of the SAW delay line is below 2 GHz with a bandwidth of less than 1 GHz. The insertion loss of a SAW delay line is high. Figure 10.7 shows the actual device (*a*) and the frequency response of a SAW delay line (*b*). The bandwidth is about 700 MHz centered at 1350 MHz, and the insertion loss is about 30 dB.

The triple transit and other time domain spurious responses of the delay line (discussed in Section 7.7) also limit the performance of a receiver. A receiver using the SAW delay line is usually designed with a dead period after processing a pulsed signal. This period is used to wait until the triple transit response is over. If the receiver is properly designed, it will only miss pulses with a frequency close to another signal and arriving behind it within the triple transit of the delay line. Thus the triple transit will not severely degrade the POI of the receiver. Since the SAW delay lines are very small in size, they can be packed close together to have small temperature differences for better amplitude and phase tracking. Figure 10.8 shows a package with five SAW delay lines. In this figure, amplifiers are provided at both the input and output of the delay lines to compensate for the high insertion loss. With present technology, it seems that SAW delay lines are the most promising delay lines for cueing receivers.

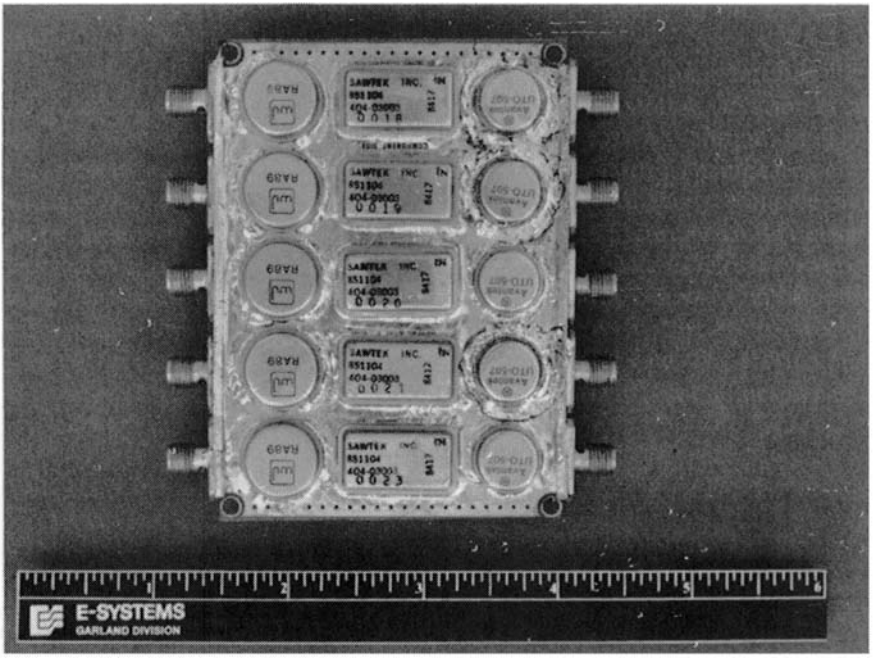


Figure 10.8. Closely packed SAW delay lines. (Courtesy of E-systems.)

10.8. SOLID-STATE BULK WAVE DELAY LINES (12, 13)

A solid-state bulk delay line operates on almost the same principle as a SAW delay line. An input transducer converts the input electric signal to an acoustic signal, and an output transducer converts the acoustic signal back to the electric signal. The only difference between a solid-state bulk wave delay line and a SAW delay line is that the acoustic wave is a bulk acoustic wave rather than a surface wave. Since the bulk wave has higher velocity than the surface wave, a bulk wave device has a potentially higher operating frequency and wider bandwidth than a SAW device.

The transducers are commonly made of piezoelectric (i.e., ZnO) film sandwiched between two metal films, as shown in Figure 10.9. The transducers are bonded on well-polished parallel surfaces of the delay line. Microwave matching networks are usually provided to transform the impedance of the transducers to approximate 50 Ω to match the characteristic impedance of the transmission lines used in the system. The acoustic wave velocity varies from several hundred to thousands of meters per second. For example, the sound velocity in quartz is 5970 m/sec. In general, the attenuation constant (dB/μsec) of bulk wave is smaller than that of its surface wave at the same frequency. Using $LiNbO_3$ as an example, the attenuation of the bulk wave is 0.1 dB/μsec versus the surface wave of 1.07 dB/μsec at 1 GHz. Thus, theoretically, the bulk wave delay line can operate at a higher frequency range in comparison to SAW delay lines. Moore et al. (ref. 12) claimed that with low-loss acoustic media such as sapphire and yttrium aluminum garnet (YAG) and high-efficiency transducers from ZnO films, bulk wave acoustic delay lines are available through the X band (8–12 GHz). They predicted that a delay line with a 1-μsec delay time at 10 GHz with 45 dB insertion loss and 20–40% bandwidth can be achieved.

The performance of a bulk device is similar to that of a SAW device. The dynamic range is limited, and the device also suffers with triple-transit response. This kind of delay line has the potential to be used in a cueing receiver at higher operating frequencies in the future.

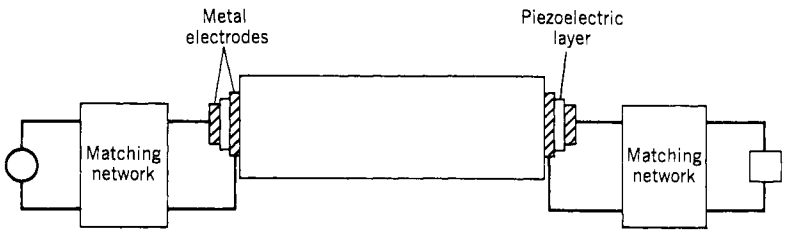


Figure 10.9. Bulk wave delay line.

10.9. MAGNETOSTATIC SURFACE WAVE DELAY LINES (14–20)

In the magnetostatic surface wave delay line, the input electric signal is transformed to a magnetostatic wave in a magnetic material, traveling a certain distance and coupled out in electric form again at the output. Yttrium iron garnet (YIG) is the most popular delay line material. The delay line is usually fabricated on a single crystal substrate located in a magnetic field, and the wave travels on the surface of the material. Therefore, this kind of delay line is sometimes referred to as the magnetostatic surface wave delay line. The wave travels along the surface in the $\mathbf{H} \times \hat{n}$ (cross product) direction, where $\mathbf{H}$ is the direction of the applied magnetic field, and $\hat{n}$ is the unit vector outwardly directed and normal to the slab, as shown in Figure 10.10. A signal traveling along the top surface of the slab will be reflected and travels back on the lower surface. The attentuation versus frequency response for a magnetostatic surface wave is shown in Figure 10.11. The attenuation is almost linearly increased with an increase of operating frequency.

Another form of the magnetostatic wave delay line is made of YIG film epitaxially grown on gadolinium gallium garnet (GGG) substrate. The YIG film thickness can range from 1 to 100 μm. A YIG thin-film delay line is shown in Figure 10.12. The input and output transducer can be a microstrip line that couples the electromagnetic field into and out of the YIG layer. A typical transducer is shown in Figure 10.13.

The performance of a magnetostatic delay line is quite similar to that of a SAW delay line or a bulk wave solid-state delay line. The direct coupling and the multiple time transit responses are also present in the magnetostatic delay line. The operating frequency of the magnetostatic delay line ranges from 1 to approximately 20 GHz. The delay time can be varied by changing the applied magnetic field intensity, as shown in Figure 10.14. An important factor should

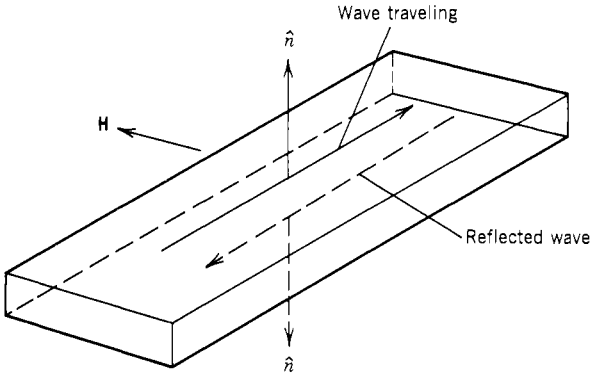


Figure 10.10. Magnetostatic wave delay line.

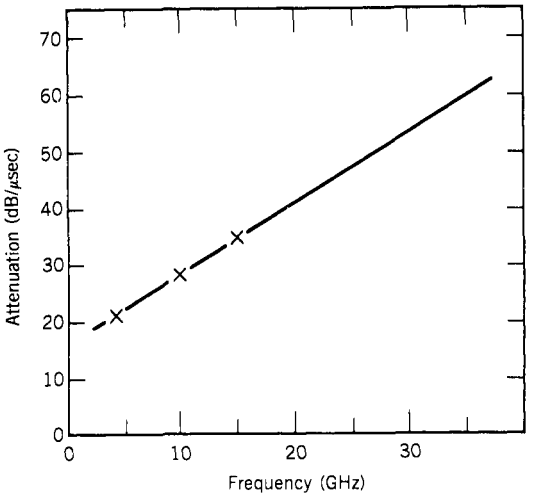


Figure 10.11. Attenuation versus frequency for magnetostatic surface wave. (Based on Merry and Sethares, ref. 14.)

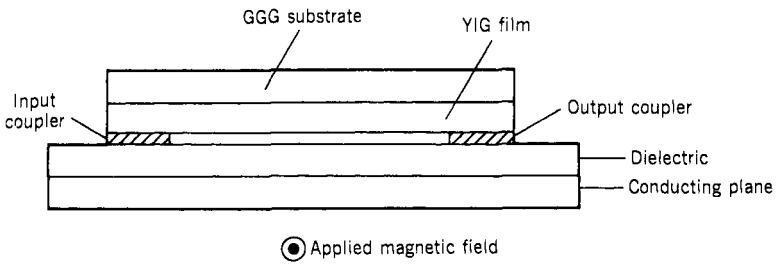


Figure 10.12. A thin-film magnetostatic wave delay line.

be pointed out in this figure. For a fixed delay line length, the delay time increases for higher operating frequencies. Therefore, strictly speaking, the magnetostatic delay line is dispersive. However, the delay time versus frequency is not a linear relation, which is desirable in a dispersive delay line (DDL) used in a compressive receiver. Thus, the bandwidth of the magnetostatic delay line is limited by the delay time deviation as well as by the insertion loss. The delay time of a magnetostatic delay line is in the 10–1000 nsec range.

The principal attractive feature of a magnetostatic delay line is its potentially high operating frequency, up to 20 GHz. Further development is required to make the magnetostatic delay line adaptable in a cueing receiver.

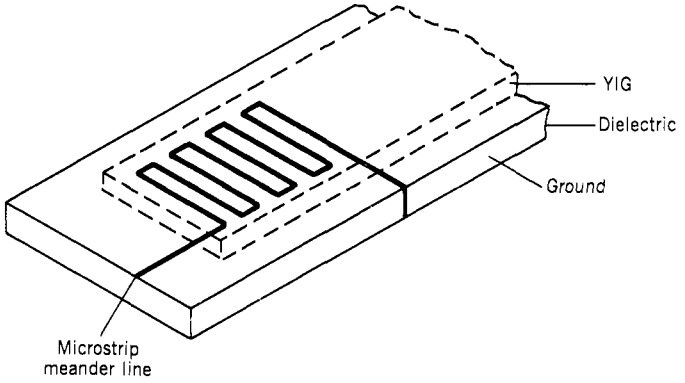


Figure 10.13. A typical magnetostatic delay line transducer.

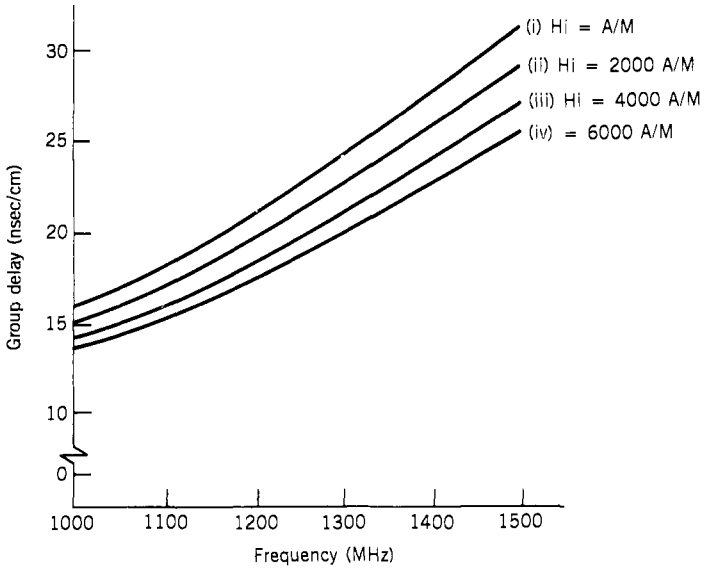


Figure 10.14. Theoretical plot of group delay against frequency in a magnetostatic delay line. (Based on Adam et al., ref. 15.)

10.10. FIBER OPTIC AND SUPERCONDUCTIVE DELAY LINES (21–30)

A fiber optic transmission line can transmit optic signals with very low insertion losses; a single-mode fiber optic is capable of accurately transmitting modulated signals with losses of less than 0.5 dB/km. The velocity of light traveling in the fiber optics is also given by Eq. (10.10). This velocity is about 2×10^8 m/sec. Therefore, to produce microsecond delay, the line must be very long. However, the diameter of the fiber optic line is very small, about 0.6–1.2 mm, and several

hundred meters of fiber optic line will not occupy a very large volume. The important factors concerning fiber optics are insertion loss and dispersion.

Attenuation in fiber optics results from a number of wavelength-dependent mechanisms, including scattering, absorption, and externally induced losses, such as those due to bending. Each of these losses can be small over a particular range of optical wavelengths. The combined effect of these losses restricts the operation of fiber systems to a relatively narrow spectral band from the visible to the near infrared (0.6–1.6 μm), which includes the wavelengths of semiconductor lasers. While for single-mode fiber optics the attenuation is under 0.5 dB/km, for multimode fiber optics the attenuation is about 2.3 dB/km (ref. 24). The modulation bandwidth of multimode fiber is most severely limited by modal dispersion. Multimode fiber can guide a number of transverse modes that do not, in general, travel at the same group velocity. As a result, the fidelity with which high-frequency components of an analog signal may be transmitted is limited. The length–bandwidth product of multimode fiber optics is about 1500 MHz-km. For single-mod fiber optics, a length–bandwidth product of 100 GHz-km can be achieved. Since the delay time required in a cueing receiver is a few microseconds, either multimode or single-mode fiber optics can be considered for cueing receiver applications.

If the fiber optics is used to delay electric signals, the input signals must first be converted to optical signals. There are two ways to modulate the light. One way is to use a light valve, that is, a Bragg cell controlled by the RF input at the output of the laser, as discussed in Chapter 9. The other way is to directly modulate a laser diode. The light output from a laser diode is related to the bias

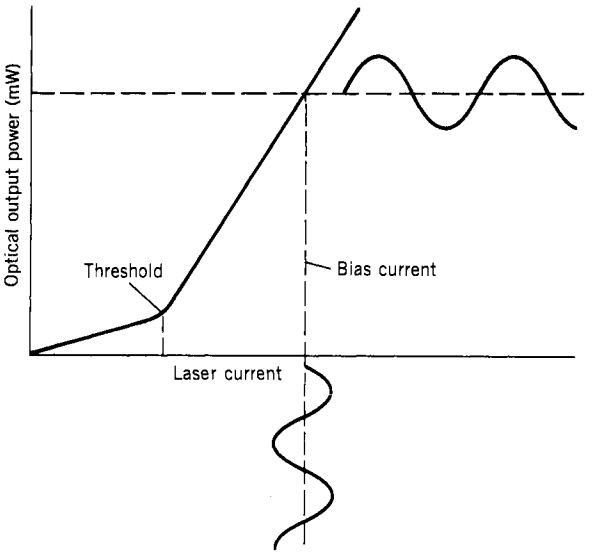


Figure 10.15. Laser output versus input current.

current, as shown in Figure 10.15. The laser is biased to a quiescent point, and the RF is used to modulate the bias current, which in turn will modulate the light output from the laser. This is the preferred approach because of its small size compared to light valve modulation. To obtain modulation with linear intensity, the laser should be biased above threshold, and the modulation should be superimposed on the biased current. At the output of the fiber optic delay line, a photodetector must be used to convert the light signals back to electric signals. The block diagram of a fiber optic delay line is shown in Figure 10.16.

The operating frequency of the fiber optic delay line depends on the input and output devices and the length–bandwidth product. Although the length–bandwidth product of the fiber optics is rather high, the operating frequency of a fiber optic delay line, including the input and output devices, can potentially achieve 10 GHz. The dynamic range of the fiber optic delay line is also limited by the input and output devices. Experimentally, a dynamic range of 30 dB

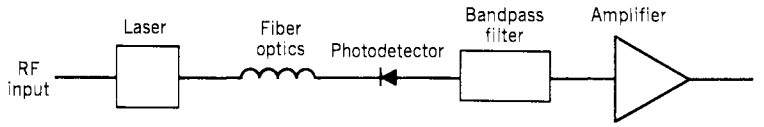


Figure 10.16. Block diagram of fiber optic delay line.

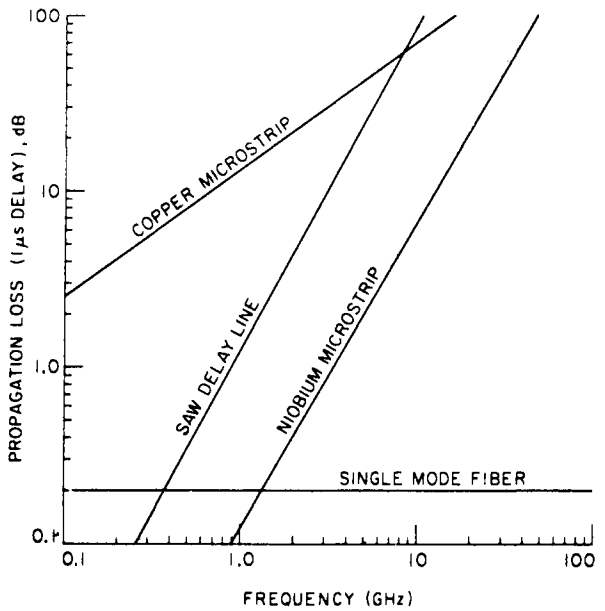


Figure 10.17. Propagation loss per 1 μsec of delay as a function of signal frequency. (Based on Jackson et al., ref. 23.)

can be achieved. The fiber optic delay line can be considered in the development stage. At present, the noise introduced in the fiber optic is rather high. Since a fiber optic delay line uses very complicated input and output devices, and the size of the delay line itself is large in comparison with other delay lines discussed above, the fiber optics may have limited potential for use in a cueing receiver.

Another type of delay line is made of superconducting materials. For example, superconducting niobium strip transmission lines on sapphire and silicon dielectric substrates have been used to make passive delay lines, capacitively coupled resonators, and linear frequency-modulated (FM) chirp filters. Like the fiber optic delay line, to provide a microsecond delay time, the delay line will be quite long. However, the insertion loss of a superconducting delay line is relatively low. For example, the loss of niobium thin film at 4.2 K and 3 GHz is 0.01% that of copper at room temperature and 0.1% of the loss of copper film at 4.2 K. Although this low loss may sound attractive for delay line applications in a cueing receiver, a superconducting delay line must operate at very low temperature, that is, at near liquid helium temperature. At the present time, the low temperature requirement limits the adaptation of superconducting delay lines in cueing receiver applications. Figure 10.17 shows the propagation loss various delay lines.

10.11. CUEING RECEIVER WITH AMPLITUDE AOA MEASUREMENT

In this section, an actual example of a cueing receiver will be briefly discussed. This system uses an amplitude comparison scheme to measure the AOA of the input signals. To simplify the design and to save hardware, the design is slightly modified from the discussion in Section 10.3. Instead of using four narrow-band receivers to measure the AOA of the input signals, one receiver is used. The receiver is switched sequentially among the four quadrant antennas. To compensate the switching time of the narrow-band receiver, four delay lines with different delay times are used, as shown in Figure 10.18. The narrow-band receiver is initially connected to antenna 1. When an input signal is detected by the omnidirectional antenna, its frequency is measured, and the logic unit will provide the proper LO frequency to the four mixers. The AOA receiver will start to measure the amplitude of the input signals sequentially from antennas 1, 2, 3, and 4. The differential delay time between two quadrant antennas matches the receiver switching time. In this arrangement, since only one receiver is used to measure the amplitude of all four quadrants, only the delay lines need to be amplitude matched.

Figure 10.19 shows a cueing receiver built with above scheme. The coarse-frequency measurement unit is a channelized receiver with SAW filters. The delay lines are also SAW delay lines. The shortest delay time is 1.3 μsec, and the differential delay time is 250 nsec. In addition to the narrow AOA receiver, there is also a fine-frequency measurement receiver with pulse parameter

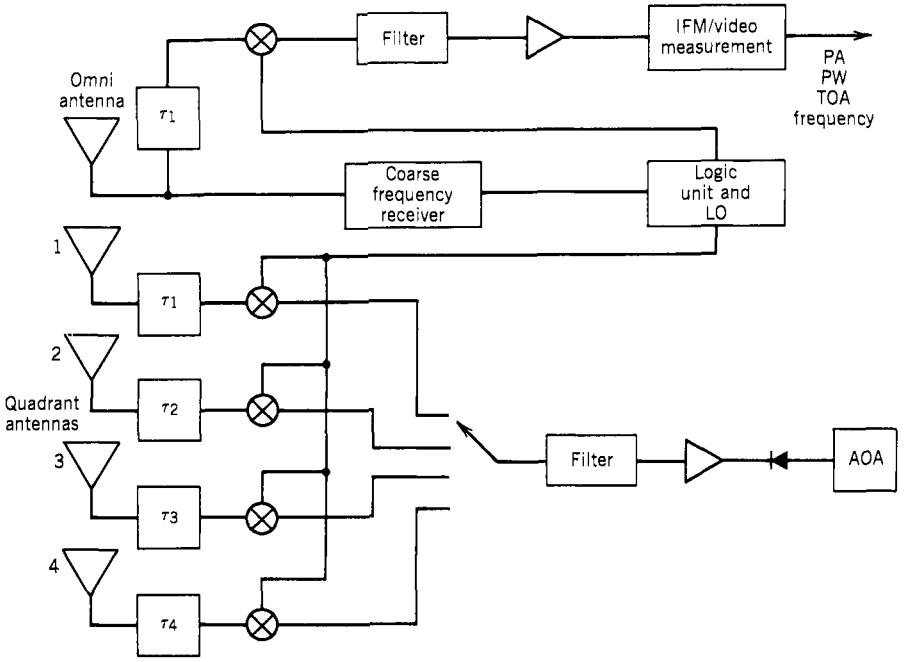


Figure 10.18. Modified cueing AOA system.

measurement capability. The fine-frequency receiver is basically an IFM receiver with one correlator. The fine-frequency receiver is connected to the omnidirectional antenna, as shown in Figure 10.18.

The logic circuit takes the digital outputs from the coarse-frequency unit as inputs and assigns a proper LO frequency to the mixers. The intermediate frequency (IF) from the mixer output is in the passband of the AOA and fine-frequency measurement receivers. The LO frequencies generate from a direct frequency synthesizer. Therefore, the frequency selection time is kept at a minimum.

The block diagram in Figure 10.18 shows only one AOA channel and one IFM fine-frequency–video measurement unit for easier demonstration of the cueing receiver. If simultaneous signals will be processed by the receiver, multiple AOA and fine-frequency receivers are required. The receiver shown in Figure 10.19 is rather large—there are multiple fine-frequency measurement receivers and multiple AOA receivers. Thus the receiver can receive simultaneous signals. The logic circuits of such a design are complicated because they must be able to assign different receivers to different input signals. Since there are multiple frequency synthesizers, one frequency can be assigned to one input signal. The number of simultaneous signals the receiver can process should equal the number of fine-frequency measurement receivers and AOA receivers. The complexity of the frequency selection logic circuits depends on

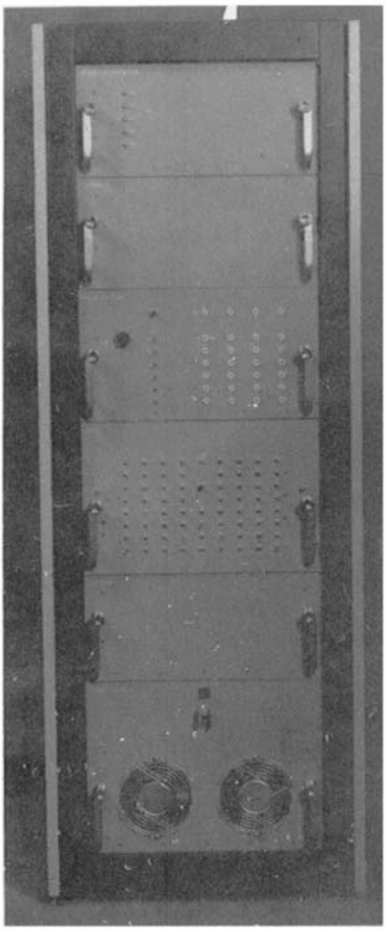


Figure 10.19. Cueing AOA receiver. (Fabricated by Texas Instruments Inc., courtesy of Avionics Laboratory, AFWAL.)

the number of simultaneous signals the receiver can process. It seems that complexity of the logic circuits can be increased tremendously if the number of simultaneous signals handled increases only modestly. On the other hand, if the receiver is designed to process small numbers of simultaneous signals, the logic circuit design can be relatively simple.

10.12. DIGITAL RF RECEIVERS (31–37)

The operating speed of digital circuits is getting faster because of the intensive research and development in this area. The advance in GaAs technology makes the gigahertz logic circuits possible. In the future, if the logic circuits are fast

enough, it will be possible to digitize the RF input signals directly and further process them in digital form. For example, an FFT can provide the frequency information on the input signals.

With today's technology, it is possible to digitize the RF input signals below 500 MHz directly through the A/D converter. A microwave receiver can be built with such an approach. Figure 10.20 shows a block diagram of such a microwave receiver. The input RF signal is down converted to a baseband, which is from dc to approximately half the A/D converter frequency, so that the A/D converters can be used to digitize the signals. The input bandwidth of this kind of receiver can be doubled through the I and Q channels, which are 90° out of phase as shown in Figure 10.20. For example, if the A/D converter operates at 500 MHz, the input bandwidth of the RF receiver will be less than 250 MHz, which is limited by the Nyquist criterion, that is, a minimum of two samples per cycle. However, using the I–Q channel approach, a bandwidth close to 500 MHz can be achieved.

Limited by today's A/D converter technology and the postprocessing to analyze the digitized information, the number of bits used in an RF receiver is usually quite limited (say 1 bit). Under this condition, the dynamic range of the receiver is quite limited. If the total number of bits is N, the dynamic range of the receiver can be approximated by

$$\mathrm{DR} = 20 \log 2^N = 6N \tag{10.10}$$

because there are a total of 2^N different voltage levels. The frequency resolution of the receiver will be determined by the total sampling length. The limited sampling length also creates side lobes, which limit the receiver dynamic range. Although a proper weighting function can reduce side lobes, the weighting function will also widen the main lobe, as discussed in Chapter 8. The following

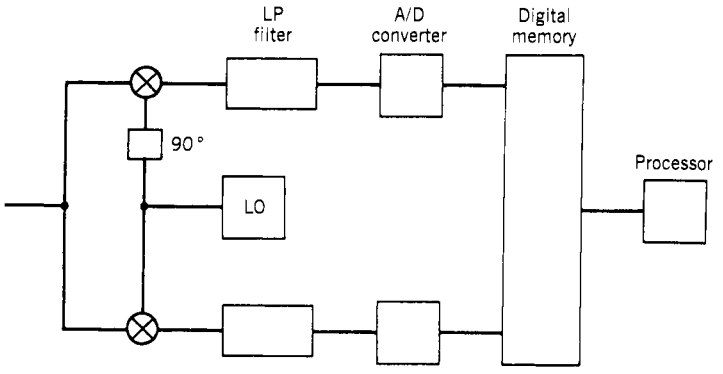


Figure 10.20. Block diagram of RF memory.

example will provide some general information on a digitizing receiver.

Sampling speed, 200 MHz	Input bandwidth <100 MHz <200 MHz with I and Q channels
Number of bits, 8	Dynamic range <48 dB
Sampling length, 1 μsec	Frequency resolution >1 MHz, side lobes are at 1.5, 2.5, 3.5 MHz, and so on, away from the main lobe

Processing the output digital data is obviously not a trivial task. Both hardware and software need to be developed. One of the most important factors is the availability of very high speed logic circuits. If proper digital processing is available, this kind of receiver can provide special information on the input signals. For example, the intrapulse information, that is, chirp rate and phase modulation, can be obtained from such a receiver. Although it is not likely that a general-purpose digital microwave receiver can be developed in the near future, some special-purpose RF digital receivers may be possible with present technology. At present time, microwave memory units are built with this basic idea. An experimental unit is shown in Figure 10.21.

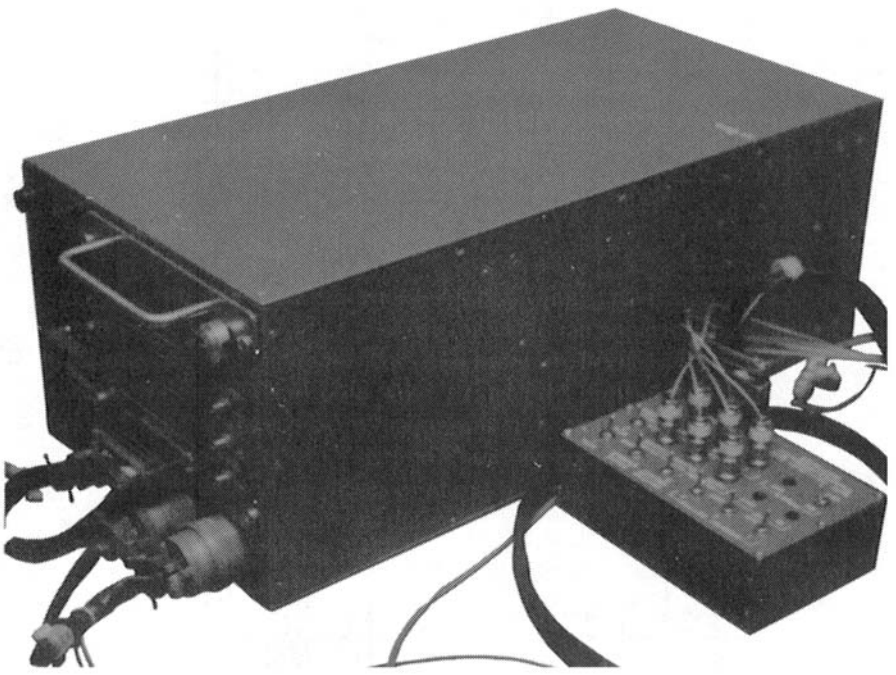


Figure 10.21. Digital microwave memory unit. (Fabricated by Whittaker Corp., courtesy of Avionics Laboratory, AFWAL.)

The A/D converters used in this unit are 1 bit at 200 MHz. The digitized data is stored in a memory unit, but the processing of the information is not at near real time. Therefore, the microwave memory unit can only record input information over a period of time but cannot process the input signals continuously like a conventional microwave receiver.

10.13. SUMMARY

The concept of the hybrid receiver can be used to design a receiver system to suit certain performance requirements. A number of possible combinations can be originated. As the electronic signal environment is getting more complicated, it might be difficult to process the signals in the environment with one kind of receiver. Some kind of hybrid combination might be a reasonable approach to solve the receiver problem.

It is anticipated that delay lines will play an important role in the hybrid receiver. It is impractical to build receivers to obtain detailed information on every input signal over a wide frequency range because such a receiver will be very complicated and large in size. It will be reasonable to build a receiver with a coarse-frequency unit over a wide bandwidth and then assign some fine parameter measurement units to further process the signals. In such an arrangement, delay lines must be used to delay the input signal to provide time for the fine parameter measurement receivers to be properly assigned.

REFERENCES

1. W. S. McCormick, The use of the instantaneous frequency transient in the design and optimization of the channelized receiver and instantaneous frequency measurement (IFM) versions of the passive EW receiver, 1984 USAF-SCEE Summer Faculty Research Program, August 17, 1984.
2. D. S. Humphreys, *The Analysis and Synthesis of Electrical Filters*, pp. 81–87, Prentice-Hall, Englewood Cliffs, NJ, 1970.
3. L. Franks, *Signal Theory*, pp. 84–85, Prentice-Hall, Englewood Cliffs, NJ, 1969.
4. L. Jacobson, Specifying delay lines, Allen Avionics Inc., Mineola, NY, 1979.
5. W. H. Hayt, *Engineering Electromagnetics*, 2nd ed., McGraw-Hill, New York, 1967, p. 283.
6. *Reference Data for Radio Engineers*, 5th ed., Chapter 22, Howard W. Sams, Indianapolis, IN, 1972.
7. A. J. Slobonik, Jr., and J. H. Silva, Ultra-flat UHF delay line modules, RADC-TR-77-257, Rome Air Development Center, Air Force Systems Command, July 1977.
8. L. A. Coldren and H. J. Shaw, Surface wave long delay lines, *IEEE Proc.*, **64**, 598–609 (1976).
9. A. J. Slobodnik, Acoustic surface design data, Air Force Cambridge Research Laboratories, Cambridge, MA, 1900.
10. W. Daniels, Texas Instruments, private communication.
11. I. C. Chang, Acoustoptic devices and applications, *IEEE Trans. Sonics Ultrasonics*, **SU-23**, 2–22 (1976).
12. R. A. Moore, R. N. Sundelin, G. Borsuk, J. Lane, C. Huber, and S. Lieberman, Broadband matched and low triple transit microwave delay line, *Proc. Ultrasonics Symposium IEEE* (78CH-1344-ISU), 1978.

13. R. N. Sundelin, R. A. Moore, B. R. McAvoy, and S. Lieberman, Gigahertz bandwidth submicrosecond low spurious microwave bulk acoustic delay line, *Proc. Ultrasonics Symposium*, pp. 161–164, September 1979.
14. J. B. Merry and J. C. Sethares, Low-loss magnetostatic surface waves at frequencies up to 15 GHz, *IEEE Trans. Magnetics*, **MAG-9**, 527–529 (1973).
15. J. D. Adam, J. H. Collins, and J. M. Owens, Magnetostatic wave group delay equipizer, *Electronics Letters*, **9**, 557–558 (1973).
16. L. K. Brundle and N. J. Freedman, Magnetostatic surface waves on a Y.I.G. slab, *Electronics Letters*, **4**, 132–134 (1968).
17. D. Colliver, C. Hilsum, B. D. Joyce, J. R. Morgan, H. D. Ress, and J. R. Knight, Experimental observation of magnetostatic modes in a Y.I.G. slab, *Electronics Letters*, **6**, 434–436 (1970).
18. J. D. Adam, Delay of magnetostatic surface waves in Y.I.G., *Electronics Letters*, **6**, 718–720 (1970).
19. J. D. Adam, M. R. Daniel, and D. K. Schroder, Magnetostatic wave devices move microwave design into gigahertz realm, *Electronics*, **53**, 123–128 (May 8, 1980).
20. J. C. Sethares, Magnetostatic wave time delays for phased array technology, presented at Naval Research Laboratory Phased Array Technology Workshop, September 9–10, 1980.
21. H. F. Wolf, *Handbook of Fiber Optics: Theory and Applications*, Garland STPM Press, New York, 1979.
22. W. T. Tsang, R. A. Long, and J. P. Van der Ziel, Low current threshold stripe buried-heterostructure laser with self-aligned current injection stripes, *Applied Physics Lett.*, **34**, 644, 1979.
23. K. P. Jackson, S. A. Newton, B. Moslehi, M. C. Chapin Cutter, J. W. Goodman, and H. J. Shaw, Optical fiber delay-line signal processing, *IEEE Trans. Microwave Theory Techniques*, **MTT-33**, 193–210 (1985).
24. E. O. Rausch, Georgia Tech Research Institute, private communication.
25. P. D. Lazay and A. D. Pearson, Developments in single-mode fiber design, material, and performance at Bell Laboratories, *IEEE J. Quantum Electron*, **QE-18**, 504–510 (1982).
26. K. Wilner and A. P. van der Heuvel, Fiber-optic delay lines for microwave signal processing, *Proc. IEEE*, **64**, 805–807 (1976).
27. C. T. Chang, J. A. Cassaboom, and H. F. Taylor, Fibre-optic delay-line devices for RF signal processing, *Electronics Letters*, **13**, 678–680 (1977).
28. S. E. Miller and A. G. Chynoweth, Eds., *Optical Fiber Telecommunications*, Academic Press, New York, 1979.
29. S. A. Reible, Wideband analog signal processing with superconductive circuits, *IEEE Proc. Ultrasonics Symposium*, **1**, 190–201 (1982).
30. J. T. Lynch, R. S. Withers, A. C. Anderson, P. V. Wright, and S. A. Reible, Multigigahertz-bandwidth linear-frequency-modulated filters using a superconductive stripline, *Applied Physics Letters*, **43**, 319–321, August 1983.
31. P. H. Saul, A. Fairgrieve, and A. J. Fryers, Monolithic components for 100 MHz data conversion, *IEEE J. Solid-State Circuits*, **SC-15**, 286–290 (1980).
32. C. Lin, Microwave-frequency intensity-modulation and gain-switching in semiconductor injection lasers, Optical Technology for Microwave Applications, Proc. SPIE, pp. 2–11, May 1–2, 1984.
33. T. R. Joseph, W. E. Stephens, and B. U. Chen, Fiber optic rf links, Optical Technology for Microwave Applications, Proc. SPIE, pp. 52–53, May 1–2, 1984.
34. D. Meignant and M. Binet, New ultra-high-speed GaAs strobed comparators, *Electron Lett.*, **19**, 67–68 (1983).
35. L. C. Upadhyayula, W. R. Curtice, and R. Smith, Design, fabrication, and evaluation of 2- and 3-bit GaAs MESSFET analog-to-digital converter IC's, *IEEE Trans. Electron Devices*, **ED-30**, 2–10 (1983).
36. R. E. J. van de Grift and R. J. van de Plassche, A monotithic 8-bit video A/D converter, *IEEE Trans. Solid State Circuits*, **SC-19**, 374–378 (1984).
37. R. A. Becker, C. E. Woodward, F. J. Leonberger, and R. C. Williamson, Wide-band electro-optic guided-wave analog-to-digital converters, *IEEE Proc.*, **72**, 802–819 (1984).

Chapter 11

Extremely High Frequency (EHF) Receivers

11.1. INTRODUCTION (1)

EHF receivers are discussed in a separate chapter because there are differences in the structure of the receivers, that is, the transmission lines, components, and antennas used in EHF receivers. Generally speaking, most of the receivers discussed in previous chapters are below 18–20 GHz. At these frequencies, the most popular transmission lines are strip and microstrip lines. Most of the components used to build EW receivers in this frequency range are also available. Above this frequency range, there are many transmission lines available, and many of the receiver components have limited performance. Advancement in the development of the components is so fast that it is difficult to assess the state-of-the-art performance of some of the devices.

The EHF range is from 30 to 300 GHz, and the corresponding wavelength is from 10 to 1 mm, which is also referred to as the millimeter wave. However, in 1976, The Institute of Electrical and Electronics Engineers declared 40–300 GHz as the nominal millimeter wave frequency. Thus, any frequency above 300 GHz with a wavelength of less than 1 mm is called submillimeter wave.

The basic concept of building EHF receivers is rather simple. The two common approaches are with crystal video receivers and wide-band down converters. In the down-converter approach, base receivers can be used to process the intermediate-frequency (IF) outputs. Since the major limitations in fabricating EHF receivers are the EHF devices and components, the performance of some of the EHF components will be discussed first. For example, if the local oscillator (LO) frequency source used in a down converter is limited in output power level or has spectrum impurity, the performance of the LO will directly affect the performance of EHF receivers.

11.2. ATMOSPHERIC EFFECT ON EHF TRANSMISSION (2, 3)

Atmospheric effect in the frequency range below 20 GHz is of less concern. However, in the EHF region, atmospheric attenuation has very important impacts on radar and communication applications. Figure 11.1 shows atmospheric attenuation as a function of frequency. The upper attenuation curve is measured at sea level, and the lower curve is at a 4-k altitude. The attenuation peaks are due to oxygen and water vapor absorption. The minima are referred to as windows. The common windows below 100 GHz are between 30 and 40 GHz and around 90 GHz. The windows are potentially used for radar and communications. The attenuation peaks are sometimes referred to as anti-windows, which can be used for short-distance covert communications or satellite communications.

The effect of atmospheric attenuation on radar operation can be demonstrated by a very simple example. In this example, the radar detection range as a function of peak transmitting power at some different discrete frequencies is calculated (ref. 2):

$$R^4 = \frac{P_t \tau G^2 \lambda^2 \rho \sigma n \exp(2\alpha R)}{(4\pi)^3 k T_0 F_n (B\tau)(S/N) L_s} \tag{11.1}$$

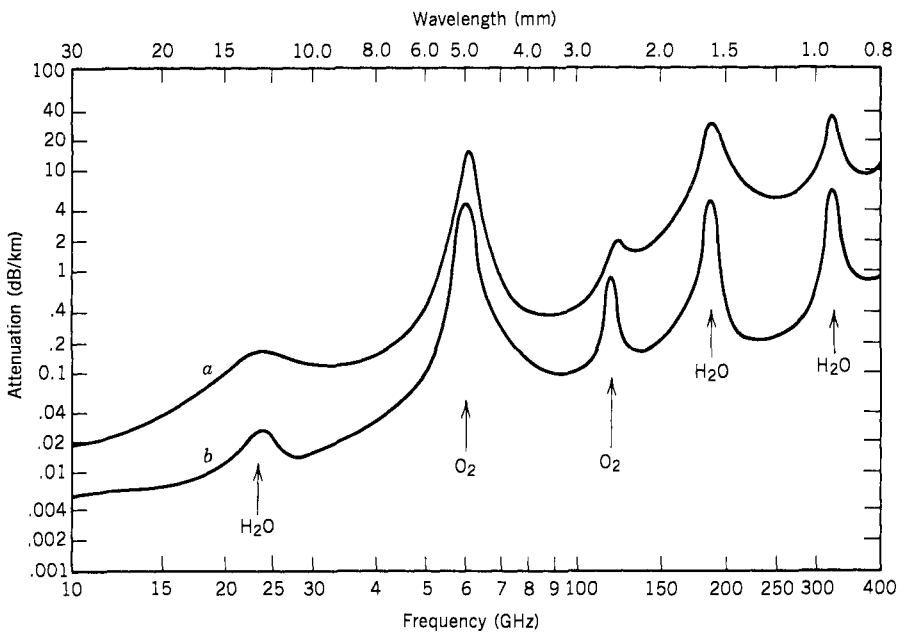


Figure 11.1. Average atmospheric absorption versus frequency. (*a*) at sea level, $P = 760$ mm Hg. $T = 20°C$, $\rho\ H_2O = 7.5\ g/m^3$; (*b*) at 4 km above sea level, $T = 0°C$, $\rho\ H_2O = 1\ g/m^3$.

where R is the range of radar detection in meters and P_t is the peak transmitted power change (let the power change from 1 W to 1 MW) in logarithmic scale.

Some arbitrary values are chosen for this calculation.

τ	Pulse width ($=200$ nsec)
G	Antenna gain ($=44.5$ dB)
ρ	Antenna efficiency (≈ 1)
σ	Radar cross section of target ($=10\ m^2$)
k	Boltzmann's constant ($=1.38 \times 10^{-23}$ J/K)
n	Number of hits integrated ($=1$)
T	Temperature ($=290$ K)
F_n	Noise figure ($=2$, $F = 3$ dB)
B_R	Receiver noise bandwidth ($=6$ MHz)
S/N	Signal-to-noise ratio required at receiver output ($=25$, $S/N = 14$ dB)
L_s	System loss ($=3$, $L = 5$ dB)
λ	Wavelength (in meters)
α	Atmospheric attenuation (in dB/km read from Fig. 11.1)

The detection range is calculated every 10 GHz for each input power. The accuracy used in the calculation is within 10 m. The results are shown in Figure 11.2. This figure by no means provides accurate radar ranging at the EHF

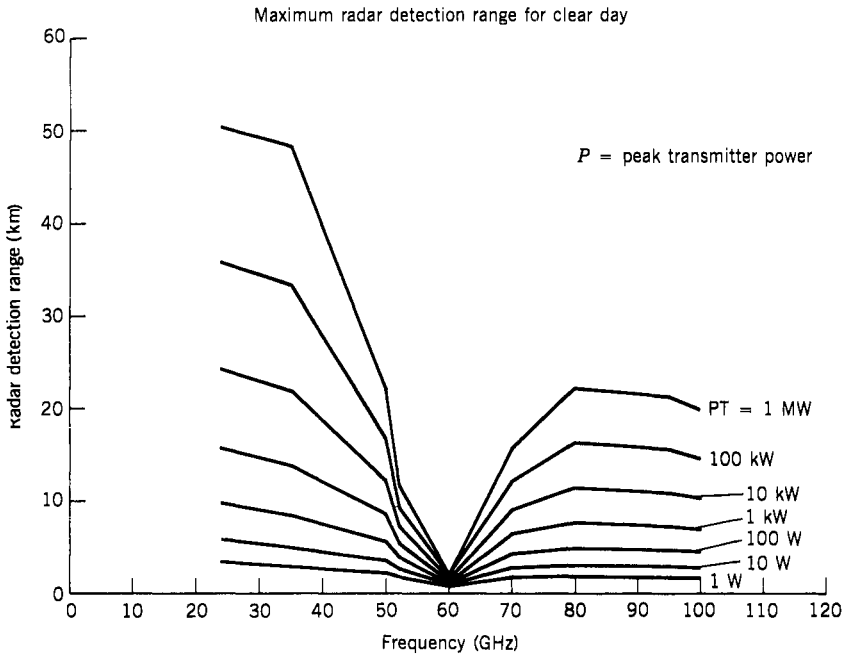


Figure 11.2. Radar range affected by atmospheric attenuation.

region; instead it gives some first-order effects on radar ranging by atmospheric attenuation. The important factor is that when atmospheric attenuation is high, the radar range is very short even though the transmitted power is very high. Therefore, these curves can provide some brief information on the frequency range an EHF EW receiver should cover. Some millimeter wave radar and missile guidance research and development can be found in reference 3.

11.3. TRANSMISSION LINES USED IN EHF RECEIVERS

There are many different kinds of transmission lines used in the EHF region. Traditionally, rectangular waveguides are the most popularly used transmission lines in the microwave frequency range because of their low insertion loss. However, with the advances in millimeter wave integrated circuits (MMIC), other transmission lines are also used because of their small size. The basic concepts of some transmission lines are not new, but the transmission lines are made available because of the development in high-frequency materials. These transmission lines will be discussed below.

A. Waveguides (1, 4–7)

Traditionally, waveguides are used by microwave engineers in both transmitters and receivers. Waveguides have rather low insertion losses and high power-handling capacity, but the size of a waveguide is large, and it is difficult to use because it is rigid. Figure 11.3 shows a section of a standard rectangular waveguide. For example, the outer dimensions of a low-frequency waveguide (WR-2300) is 23.250 × 11.750 in. Therefore, waveguides have been almost completely replaced by cables in receiver research under 20 GHz, even in laboratory setups.

In the EHF receiver area, standard waveguides continue to be the most commonly used transmission lines up to 220 GHz because of their effective shielding and relative ease of fabrication for low-volume applications where

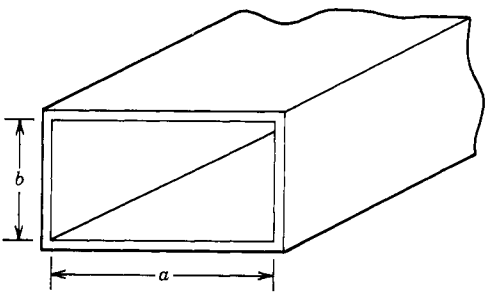


Figure 11.3. Standard rectangular waveguide.

cost is not the major concern. However, for frequencies above 100 GHz, RF loss due to ohmic skin effect becomes very significant. For example, losses at the 60-GHz waveguide (WR-15) are 0.038 dB/in., while at the 140-GHz waveguide (WR-7) the losses are four times as great. Furthermore, the small waveguide dimensions make fabrication difficult. For example, tolerances for WR-7 waveguide are listed at ± 0.00025 in. because of the small inside guide dimension: 0.065×0.0325 in. Although lower frequency designs of rectangular waveguide components can be mechanically scaled up 140–220 GHz, RF loss and bandwidth performance deteriorates considerably.

An oversized waveguide (Fig. 11.4) has lower transmission loss and high fabrication toterance because the cross section of the waveguide is increased. However, care must be taken to prevent the propagation of the higher-order mode. Components designed for the oversized waveguide have been less successful, and although mechanical tolerance is relaxed somewhat, RF design is considerably more difficult because the guiding structure admits a higher-order propagation mode.

The general trend of EHF transmission lines seems that whenever another form of transmission line is applicable to receiver developments other than waveguides, ridged rectangular waveguides will be replaced.

B. Strip and Microstrip Lines (8–17)

The transmission lines commonly used in the lower microwave frequencies are the strip and microstrip lines. One reason for the wide acceptance of these transmission lines is their simplicity. Another reason is their suitability for incorporating different circuits, particularly active devices. Although there are no clear definitions for strip and microstrip lines, it is commonly agreed by

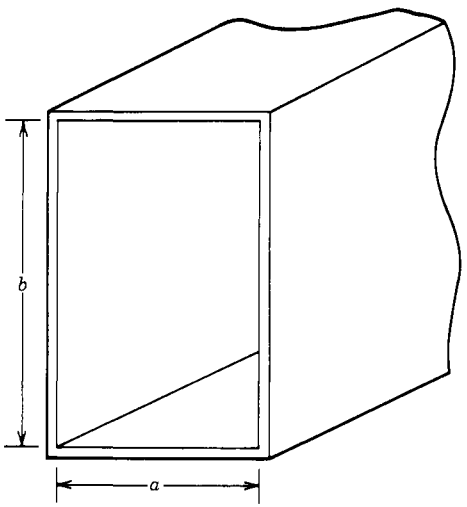


Figure 11.4. Oversized waveguide.

scientists and engineers in the microwave field that a conductor printed on a dielectric substance and sandwiched by another homogeneous dielectric parallel plane is called a strip line (or balanced strip line), as shown in Figure 11.5. Conducting plates are on both sides of the dielectric plates. If the printed line is on one side of a dielectric sheet with a shield conductor on the other side, as shown in Figure 11.6, the circuit is referred to as a microstrip line. It is also called an open strip line. The microstrip lines are enclosed in metal containers. When the operating frequency gets higher, the fundamental problems, that is, radiation loss, dispersion, and higher mode, of ordinary microstrip lines will become more severe. Although these problems can be controlled by choosing thinner substrates at higher operating frequencies, this approach will degrade the Q factor, compound tolerance problems, and restrict the range over which the characteristic impedance can be varied. Therefore, although strip and microstrip lines are very popular for use in lower microwave frequencies, they are not used much in the EHF range. Microstrip is seldom used for applications above 60 GHz. Attenuation in microstrip has been measured with an indicated value greater than 1.5 dB/in. at 140 GHz. However, in MMIC applications, the distance between elements is usually short, and microstrip lines are sometimes used.

C. *Slotlines (18–23)*

The slotline is formed by etching a narrow slot in the metallization on one side of the surface of the substrate, as shown in Figure 11.7. There is no conducting

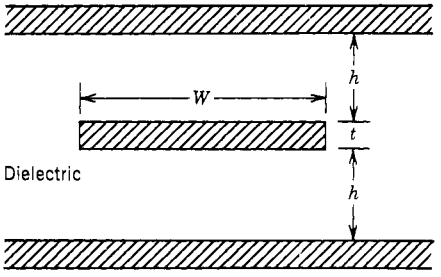


Figure 11.5. Strip line.

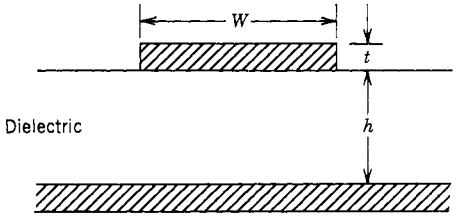


Figure 11.6. Microstrip line.

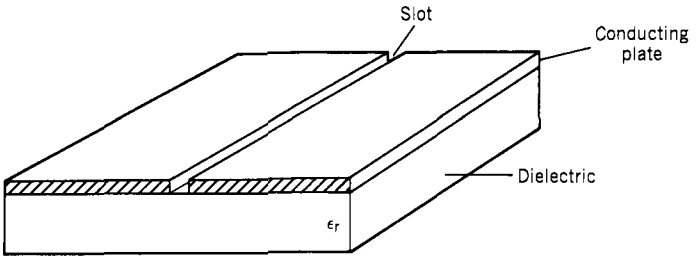


Figure 11.7. Slotline.

plate on the other side of the substrate. The energy is carried in both the dielectric substrate and the free space above it. A voltage difference exists between the slot edges. The electric field extends across the slot, and the magnetic field is perpendicular to the slot, as shown in Figure 11.8*a*. The longitudinal view in Figure 11.8*b* shows that in the air regions the magnetic field lines curve and return to the slot at half-wavelength intervals. The current paths on the conducting surface are shown in Figure 11.8*c*. The surface current density is the greatest at the edges of the slot and decreases rapidly farther from the slot. Since the dielectric constant ε is greater than 1, more energy is carried through the dielectric substrate than in the air. The availability of slotlines can improve the flexibility of microwave designs.

Increasing the operating frequency of slotlines can be achieved by reducing the gap width of the slot. Above 100 GHz, for example, the slot width would be less than 3 mil. The proximity of the conductors increases the already high loss.

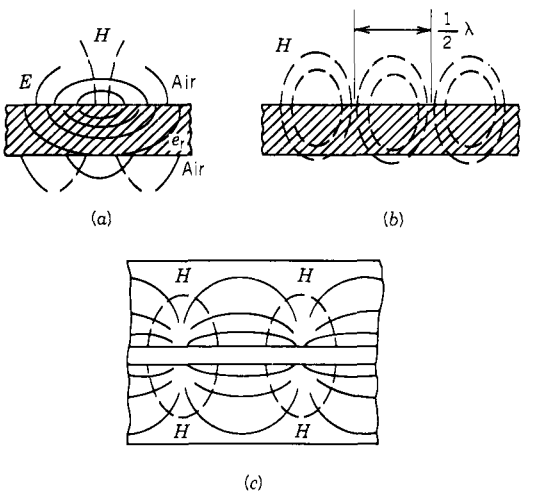


Figure 11.8. Field and current distribution of slotline: (*a*) field distribution in cross section perpendicular to wave propagation; (*b*) H field in longitudinal section; (*c*) current distribution on metal surface. (Based on Cohn, ref. 18.)

Implementation of passive components at high microwave frequencies is considered extremely difficult in slotlines because of the confined space and high loss. The slotlines are not considered to be main transmission lines in EHF, but sometimes they are used to change from one type of transmission line to another.

D. Fin Lines (24–31)

Fin lines are popular transmission lines in EHF receiver designs and can be considered to be slotlines enclosed in rectangular metallic waveguides. Generally, there are two kinds of fin lines: insulated and grounded. Figure 11.9*a* shows an insulated fin line, and Figure 11.9*b* shows a grounded fin line. For an insulated fin line, the fins are printed on a dielectric substrate insulated from the waveguide. Using this kind of fin line, bias may be introduced for active components mounted on the substrate. The RF continuity between the fins and the waveguide wall can be achieved by choosing the thickness of the broad

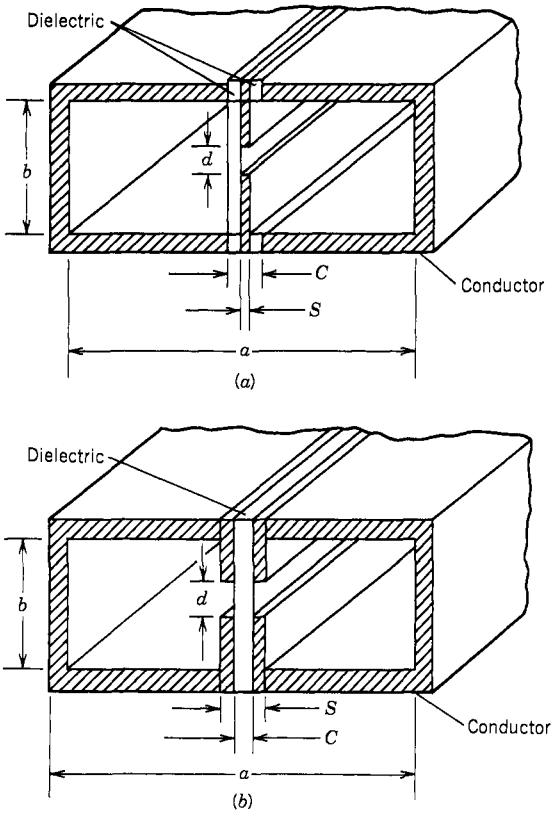


Figure 11.9. Fin lines: (*a*) insulated; (*b*) grounded.

walls to be a quarter-wavelength or by selecting $c \ll a$, where c is the dielectric thickness and a is the major inner dimension of the waveguide, as shown in Figure 11.9.

For a grounded fin line, the fins are printed on both sides of a single dielectric substrate. Since the fins are directly grounded to the metal waveguide walls, it is impractical to apply dc bias to active components on the substrate. The grounded fin line is applicable only to passive devices, but it has lower loss than the insulated fin line.

The dimensions of the metal waveguide enclosing the fin lines are not as tightly controlled as a regular waveguide. Besides, there can be many components enclosed in the waveguide. Therefore, although a metal waveguide is used in a fin line configuration, it is still more popular than a regular waveguide in EHF applications.

The major disadvantage of the fin line is transmission loss. Due to the concentration of the RF field and the conduction currents, losses are higher than in waveguides even if dielectric substrate losses were neglected. Fin line technology should be suitable for the 20–40-GHz range, with good results for some applications up to 60 GHz and potentially acceptable performance up to 100 GHz.

E. Suspended strip line (32, 33)

A suspended strip line can be considered to be a microstrip line enclosed in a rectangular metallic waveguide. The circuit is printed on a dielectric substrate sandwiched in a conducting waveguide. Figure 11.10*a* shows a single-registry type and Figure 11.10*b* shows a double-registry type. In Figure 11.10*b*, both

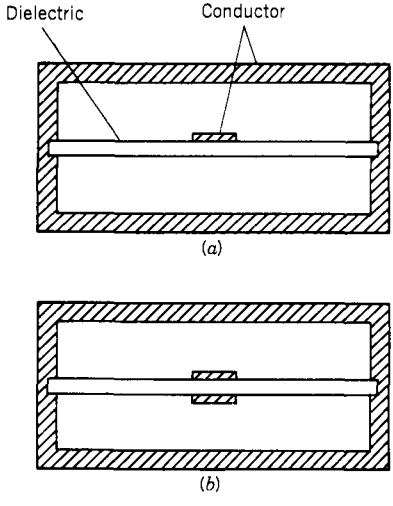


Figure 11.10. Suspended strip lines: (*a*) single-registry type (*b*) double-registry type.

printed conductors are at the same potential and can be treated as one conductor. Because these lines are doubly shielded, they have potentially lower loss than the microstrip line. The effective dielectric constant of these lines is only slightly greater than unity because most of the cross section is air. The transmission mode is very near TEM because the configuration is similar to a coaxial transmission line. As in the fin line structure, the dimensions of the waveguide do not have to be as accurate as conventional waveguides at EHF. The loose tolerances on the strip line itself allow simple fabrication. Implementation of passive components has largely been limited to filters. One material commonly used to fabricate fin lines and suspended strip lines is commercially referred to as Duroid and has a relative dielectric constant of approximately 2.2.

F. Dielectric Transmission Lines (34–39)

The dielectric transmission lines can be divided into two groups: dielectric waveguide and image lines. A simple dielectric waveguide is a piece of dielectric material with a high dielectric constant and low loss. Its cross section can be either rectangular or cylindrical, as shown in Figure 11.11. The rectangular guide will allow easy transition to the standard waveguide with minimal attenuation. The high dielectric constant will retain the electromagnetic energy inside the dielectric material. The simple dielectric waveguide is suitable for fabricating only passive EHF components because it is difficult to provide a ground plane for active devices. A dielectric image guide can be used to remedy this deficiency.

G. Dielectric Image Guides (40, 41)

A dielectric image guide is a rectangular dielectric waveguide formed on a conductor, as shown in Figure 11.12. The addition of a metal ground plane in an

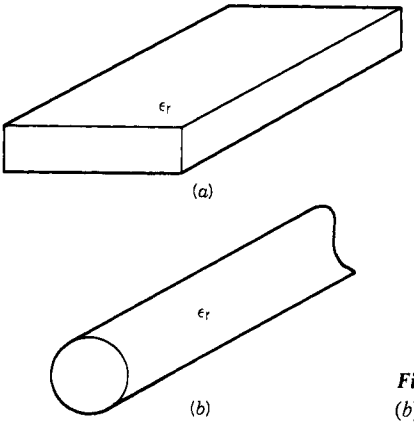


Figure 11.11. Dielectric waveguides: (*a*) rectangular; (*b*) cylindrical.

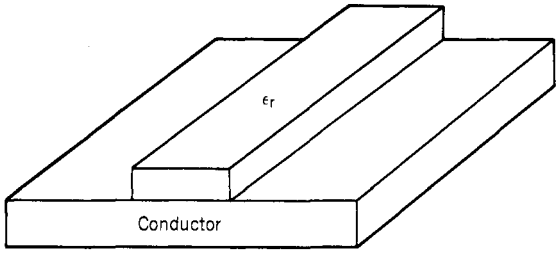


Figure 11.12. Dielectric image guide.

image guide improves the mechanical integrity of the dielectric as well as reducing the propagation loss. Another image guide, called an insulated dielectric waveguide, is shown in Figure 11.13. It is a dielectric waveguide that is separated from the conductor by another dielectric sheet. This layer is comprised of lower dielectric constant materials than the guide dielectric to ensure that most of the energy propagates within the main guide. Research indicates that losses are acceptably low and not acutely sensitive to size variations. Losses for image guides at 140 GHz are typically two-thirds of that measured in dielectric waveguides under similar conditions. This is due to the fact that the factor-of-2 reduction in dielectric cross section dominates over the increase in loss introduced by adding the imaging ground plane. At higher frequencies, this slight advantage disappears as ohmic loss increases. The conductor on the image line will provide a grounding plane for the possibility of mounting active devices in the circuit.

H. Summary (42)

The most readily observable and critical transmission property is insertion loss. A graph depicting the range of attenuation for each transmission line in the frequency range 120–150 GHz is shown in Figure 11.14. Most transmission lines exhibit roughly similar loss, with the exception of microstrip and slot lines.

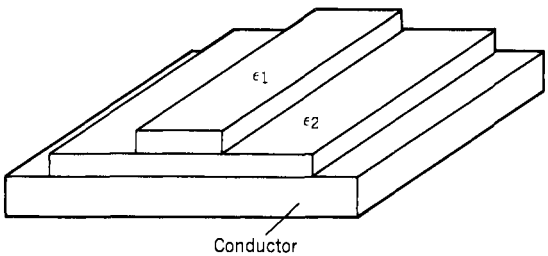


Figure 11.13. Insulated dielectric image guide.

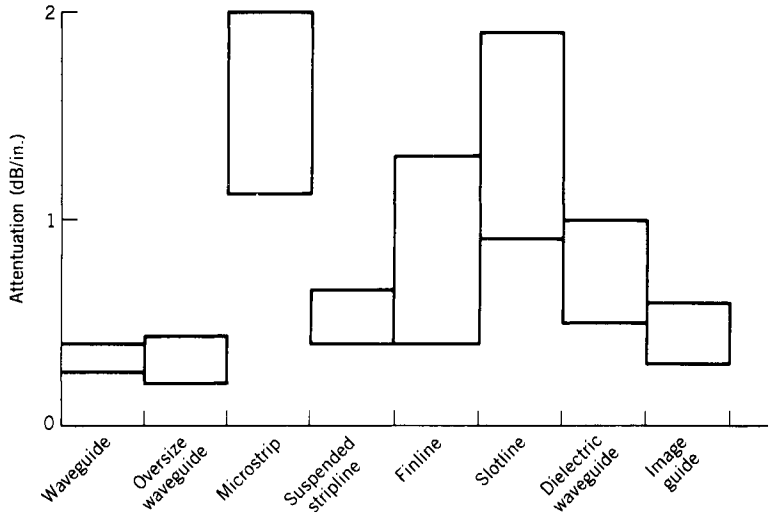


Figure 11.14. Attenuation over 120–150 GHz. (Based on Raue et al., ref. 42.)

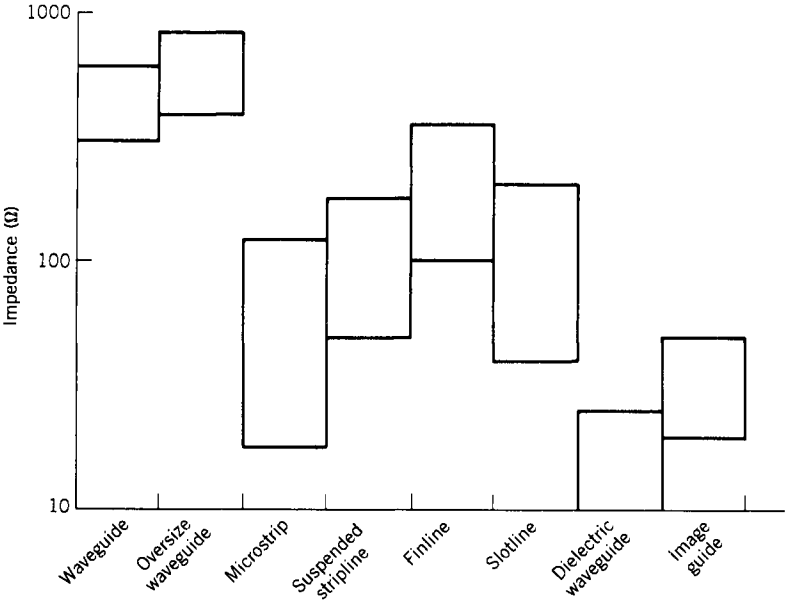


Figure 11.15. Impedance over 120–150 GHz. (Based on Raue et al., ref. 42.)

Another important property of transmission lines is their characteristic impedance. Low-impedance implementations tend to be more useful since interfacing to active devices such as diodes is less difficult because of the low impedance of the diode. Impedance ranges over 120–150 GHz for the transmission lines discussed above are shown in Figure 11.15.

11.4. EHF PASSIVE COMPONENTS

The most commonly used passive components in EHF receivers are power dividers, hybrid couplers, and filters. The operating principles of these circuit elements are the same as in the lower microwave frequency range. Since the transmission lines are different, they will be briefly discussed here.

A. Power dividers and hybrid couplers (43-45)

In the EHF range, power dividers and hybrid couplers are often used to build power amplifiers. Since the transistors and diodes have limited power capacity, to generate high-power components, many devices are used in parallel. For example, to build a power amplifier, power dividers or couplers are used to split the input power before the devices and combine the power after the devices, as shown in Figure 11.16. In an EHF receiver, these components are often used with filters for frequency channelization. Different forms of 3-dB couplers are shown in Figure 11.17. Since the wavelength is short, the size of these components can be very small.

B. Filters (45–48)

Filters are used in EW EHF receivers for wide-band frequency channelization. The bandwidth is usually over 1 GHz and up to 10 GHz. The common design

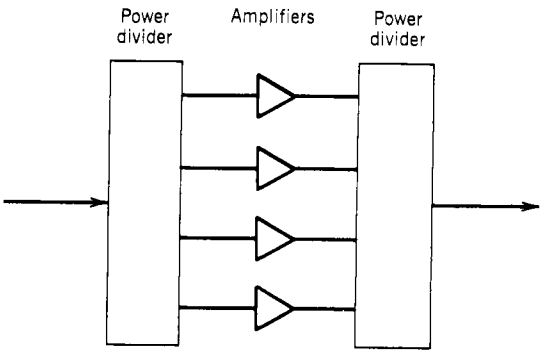


Figure 11.16. Power amplifier arrangement.

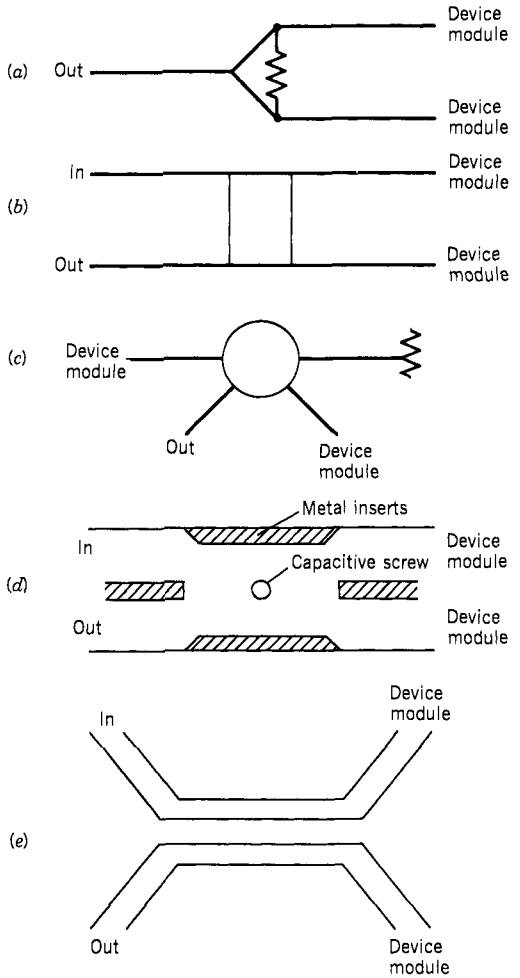


Figure 11.17. Forms of 3-dB couplers: (*a*) two-way Wilkinson divider; (*b*) 3-dB branch line coupler; (*c*) rat race coupler; (*d*) waveguide short slot coupler; (*e*) dielectric waveguide coupler. (based on Chang and Sun, ref. 45.)

is to put a metal insert in the E plane of a rectangular waveguide to form a filter. The metal strip in the E plane changes the inductance of the waveguide. A four-resonator E-plane metal insert filter is shown in Figure 11.18. Another kind of filter is a parallel coupled resonator filter, as shown in Figure 11.19. Figure 11.19*a* shows a simple bandpass filter excharge, whereas 11.19*b* represents a frequency diplexer. This kind of filter is often used as an IF filter in an EHF receiver. The same problem exists in the building of an EHF frequency multiplexer; that is, connection of the individual filters as in the lower microwave range. When individual filters are connected in parallel or in series to form a

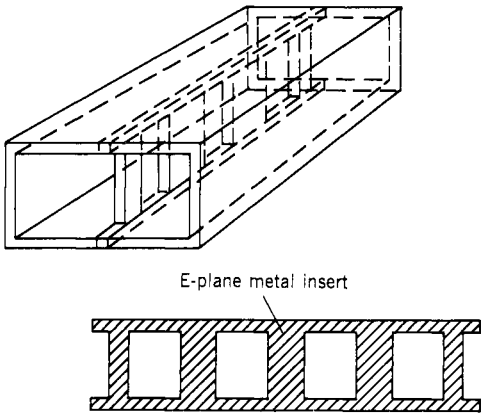
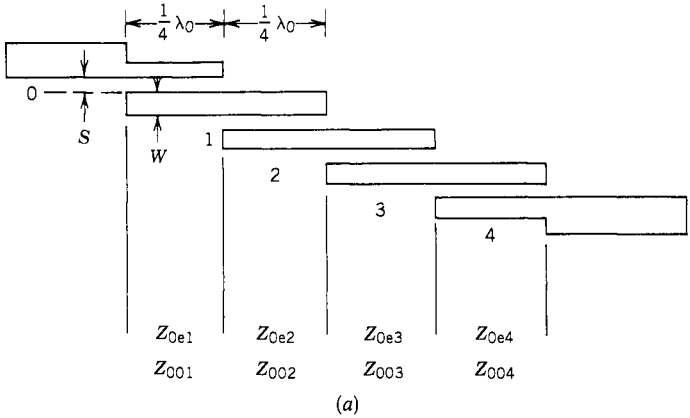


Figure 11.18. E-plane metal insert filter. (Based on Vahldieck et al., ref. 47.)



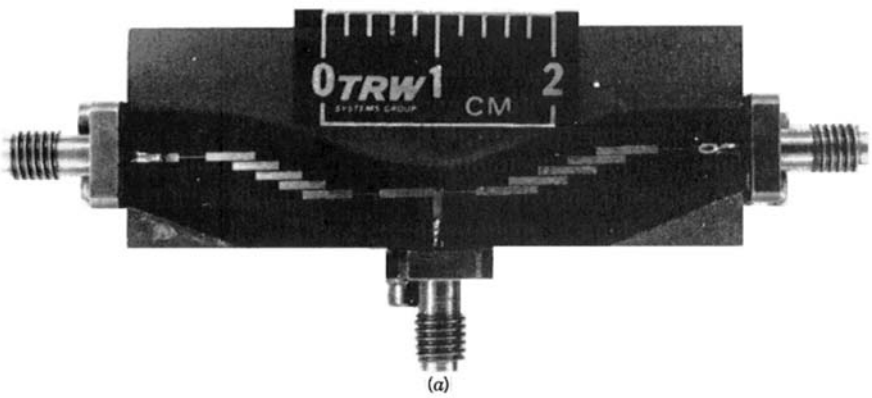


Figure 11.19. Parallel coupled resonator filters: (*a*) simple filter; (*b*) frequency diplexer.

frequency multiplexer, the interference among the filters will degrade the performance of the filters.

11.5. NONRECIPROCAL DEVICES (49–60)

Strictly speaking, a nonreciprocal device is also a passive device. In a nonreciprocal device, the energy flow from a certain port k to another port l is different from that from port l to k. Ferrite material is often used to make such a device. The two common reciprocal devices are isolators and circulators.

A. Isolators

An isolator is a two-terminal device. The energy flowing from port 1 to port 2 has very low insertion loss, but the energy flowing from port 2 to port 1 has very high insertion loss, referred to as the isolation of the device. Isolators are mainly used in microwave frequency range to improve impedance matching. Since the device and component technology in lower microwave frequencies is advancing and better impedance matching can be accomplished, isolators are not popularly used. Sometimes, they are used in conjunction with mixers at lower microwave frequencies to provide better impedance matching in order to reduce spurious outputs from the mixer. Sometimes, attenuators can be used to replace isolators because an attenuator has wider bandwidth. Although an attenuator has higher insertion loss, often power is not a major concern at the lower microwave frequency range. In the EHF range, isolators instead of attenuators are used often to improve the impedance matching of the circuits.

There are two kinds of isolators, Faraday rotation and field displacement types. Both isolators are in waveguide form. The Faraday rotation isolator consists of a 45° twist rectangular section and a circular section of waveguide. In the circular section of waveguide, there is a ferrite rod. A dc magnetic field is applied along the axis of the ferrite. A piece of resistance film is inserted at the end of the rectangular waveguide as shown in Figure 11.20. The E field of the input from port 1 is rotated 45° by the twist section and then rotated back in the ferrite rod. It will pass the isolator to port 2 without loss because the electric field is perpendicular to the resistance film. When the input is from port 2 to port 1, the electric field is rotated by the ferrite and the twist section in the same direction. Thus, the electric field from port 2 is rotated 90° in the isolator and is absorbed by the resistance film because the E field is parallel to it.

A field displacement isolator is shown in Figure 11.21. A piece of ferrite is placed along the E field in a rectangular waveguide. The ferrite is close to one side wall of the waveguide, and a dc magnetic field is applied along the E field of the electromagnetic wave. When the wave travels from port 1 to port 2, the RF magnetic field is rotating counterclockwise in the ferrite slab. When the wave travels from port 2 to port 1, the RF magnetic field is rotating clockwise in the ferrite slab. The electrons in the ferrite will resonate clockwise with the RF

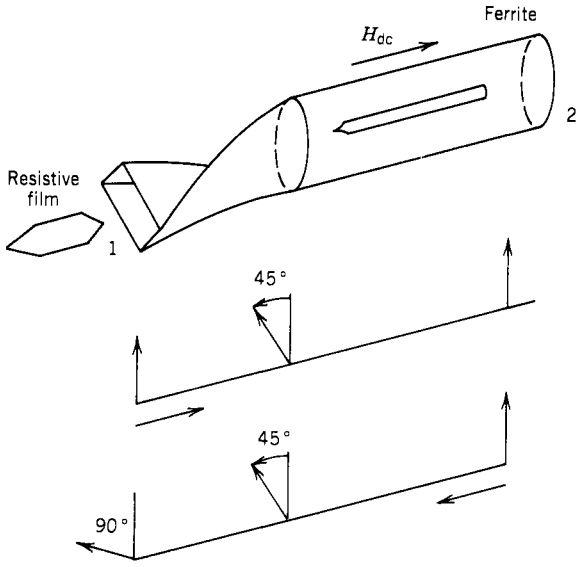


Figure 11.20. Faraday rotation ferrite isolator.

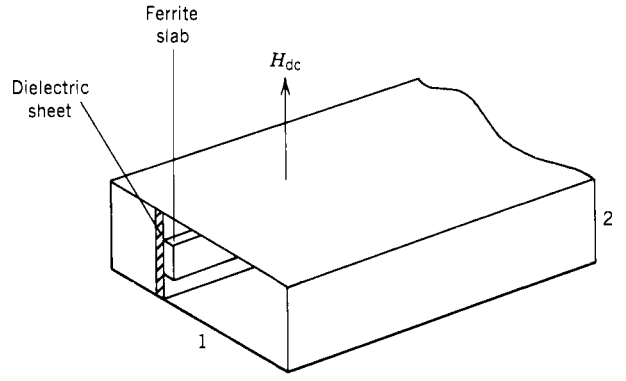


Figure 11.21. Field displacement isolator.

magnetic field and absorb the wave traveling from port 2 to port 1 but provide no absorption to the wave traveling from port 1 to port 2.

B. Circulators

A microwave circulator is a nonreciprocal ferrite device that contains three or four ports. The input from port n will emerge from port $n + 1$ but not from port $n - 1$. A three-port ferrite junction circulator, usually called the Y-junction circulator, is most commonly used. They are available in either rectangular

waveguide or strip line forms. The signal flow in the three-port circulator is assumed as $1 \rightarrow 2$, $2 \rightarrow 3$, and $3 \rightarrow 1$, as shown in Figure 11.22. If port 3 is terminated with a matched load, the device becomes an isolator, the two-terminal device shown in Figure 11.23. The input from port 1 will reach port 2, but the input from port 2 will be absorbed at port 3 by the matched load.

If the bandwidth of a circulator is wide, it can be used in frequency channelization. It is difficult to design frequency contiguous filters with a low-input voltage standing wave ratio (VSWR), as shown in Figure 11.24*a*. One way to improve the VSWR is to separate the frequency contiguous filters into two groups and connect them through a circulator, as shown in Figure 11.24*b*. In the two banks of filters, bank 1 contains the even-numbered filters and bank 2 contains the odd-numbered filters. Since frequency separation between filters in each bank is increased, the coupling between adjacent filters will decrease, and the VSWR at the input of the filter bank will improve. The input from port 1 will reach port 2. If the signal frequency matches the frequency of any filter in bank 1, it will pass through the proper filter. Signals whose frequencies do not match

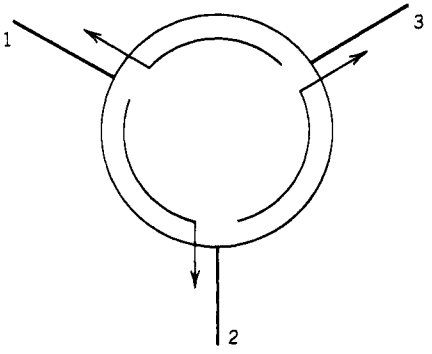


Figure 11.22. Signal flow in a circulator.

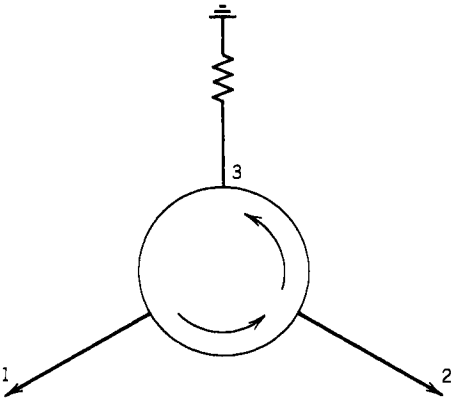


Figure 11.23. Circulator working as isolator.

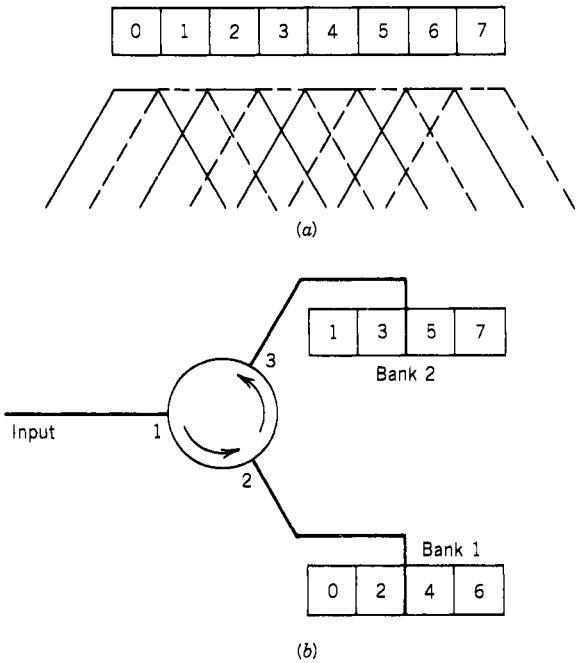


Figure 11.24. Parallel filter connection of contiguous frequency: (*a*) contiguous filter bank: (*b*) use of a circulator to connect two filter banks.

the bank 1 filters will be reflected and reach port 3. This reflected signal will pass the proper filter in bank 2.

A strip line Y-junction ferrite circulator is shown in Figure 11.25. It consists of two ferrite cylinders filling the space between a metallic conducting center disk and two conducting ground plates. The connections to the center disk are in the form of three strip line center conductors attached to the disk at points 120° apart from its circumference. A dc magnetizing field is applied parallel to the axis of the ferrite cylinders. In this simple arrangement, the device will perform as a circulator.

Experimentally, it was found that the Y circulator has some, but not all, of the properties of a low-loss transmission cavity. The maximum isolation occurred at almost the frequency at which the insertion loss was minimum. These observations suggest a resonance of the center disk structure as being an essential feature of the operation of the circulator. The lowest-frequency resonance of the circular disk structure is the dipolar mode, as shown in Figure 11.26. The electric field vector E is perpendicular to the plane of the disk. The E field, in the upper ferrite disk, is shown in circles with a dot or cross in the center to represent the outward and inward direction of the paper, respectively. The E-field vectors, in the lower ferrite disk, are mirror images of the upper one. The RF magnetic field vector H lies parallel to the plane of the disk, shown as solid lines in Figure

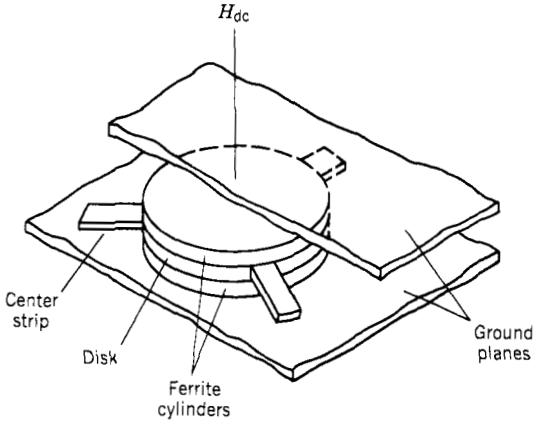


Figure 11.25. Strip line Y-junction circulator. (Based on Fay and Comstock, ref. 56.)

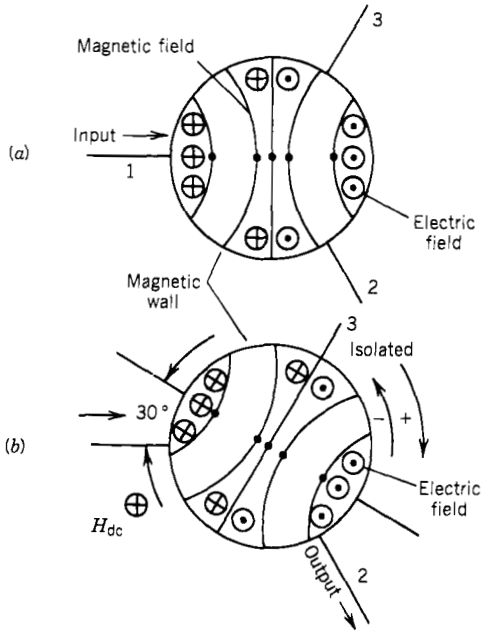


Figure 11.26. E and H fields in Y-junction circulator: (*a*) dipolar mode of dielectric disk, $H = 0$; (*b*) analogous pattern of magnetized disk. (Based on Fay and Comstock, ref. 56.)

11.26. The H field curves over the edge of the metal disk and continues back on the underside in the opposite direction, as the solid lines indicate, and back again on the upper side to make a closed loop. Therefore, the electric and magnetic fields in the upper and lower ferrite disks are the same, and analysis of the device need only be concerned with one ferrite disk. Figure 11.26*a* shows the E and H fields of the resonance mode without the dc magnetic bias field. Port 1

is the input. Ports 2 and 3, if open circuited, will see voltages that are 180° out of phase with the input voltage. The input power from port 1 will be divided equally among ports 2 and 3. Under this condition, there is no circulation effect in the device.

The standing wave pattern of Figure 11.26*a*, in which the field varies as $\exp(j\omega t)$, can be generated by two counterrotating field patterns of the same configuration. Each of these patterns would involve an RF magnetic field pattern that is circularly polarized at the center of the disk, becomes more elliptical as the radius increases, and is linearly polarized at the edge of the disk. If a dc magnetic biasing field is applied in the direction of the axis of the disk, the relative permeabilities μ_r^+ in the clockwise direction and μ_r^- in the counterclockwise direction are different, and μ_r^+ is larger than μ_r^-. As a result, the clockwise wave will rotate more than the counterclockwise wave, and a net rotation will result, as shown in Figure 11.26*b*. Under this condition, port 3 is situated at the voltage null of the disk and the voltages at ports 1 and 2 are equal. The device is equivalent to a transmission cavity between ports 1 and 2, and port 3 is isolated. This simple explanation illustrates how a circulator works.

11.6. EHF AMPLIFIERS (61–66)

Although the EHF signals can be amplified by some two-port negative-resistance devices (Gunn and IMPATT), the bandwidth of these amplifier is narrow. Since these devices have two ports, a circulator is needed to separate the input and the output. They can operate in a stable mode or as injection-locked oscillators. These amplifiers may be suitable for some special applications, but in general, they are not suitable for EHF EW receivers.

The most popular FET in the microwave region is the GaAs metal Schottky-barrier FET (MESFET). Its basic structure is shown in Figure 11.27. An *n*-type epitaxial GaAs layer is grown on a semi-insulating GaAs substrate. The source, gate, and drain are evaporated on the *n*-type GaAs. The source and gate are reverse biased, and the source and drain are forward biased. The majority of carriers (electrons) flow from the source to the drain. A depletion region is created under the gate by the voltage on the gate. When the reverse bias on the source and gate increases, the height of the depletion layer increases, which will increase the resistance between the source and the drain. Therefore, the gate source voltage controls the current flow from the drain to the source, which goes through the load resistor R (Fig. 11.27*b*). The amplification is thus accomplished.

FETs can be used to build EHF amplifiers. Experimental balanced cascade FET amplifiers can work up to 60 GHz with a noise figure of 7 dB. FET oscillators of up to 60 GHz have also been demonstrated in the laboratory. The research in this area is very active. New results and improvements will probably be reported by researchers on a continuous basis.

At present, it is anticipated that EHF receivers below 60 GHz will use FET amplifiers as the first stage. Receivers above 90 GHz will use mixers as the

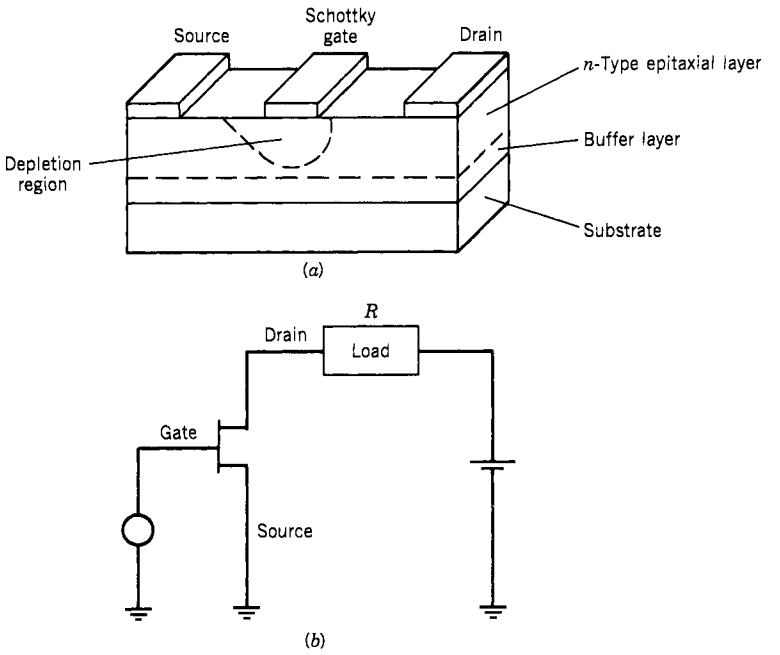


Figure 11.27. Metal Schottky barrier FET. (*a*) diode (*b*) circuit.

first stage. Between the frequency range 60–90 GHz, the first component used in a receiver probably will depend on the future developments of the devices.

11.7. EHF MIXERS (67–80)

The availability of EHF mixers made the development of channelized EHF receivers with high sensitivity possible. Wide-band mixers operating in the 100-GHz frequency range can be achieved. If amplifiers at certain EHF ranges are not available, mixers will be used as the first component of the receiver.

In general, there are two common approaches to fabricating EHF mixers. The first is a crossbar strip line mixer, as shown in Figure 11.28. The mixer consists of two diodes (located at the center of the figure). The RF signal is fed from a waveguide perpendicular to the plane containing the figure. The LO is fed from the right through a suspended strip line on the back of the substrate. A sliding short is furnished as an impedance matching element for the LO port. The IF output is on the left side of the figure. The IF filter is used to stop the RF and LO frequencies from reaching the IF output. An impedance matching circuit is provided before the IF filter. A typical insertion loss versus frequency is shown in Figure 11.29. The input frequency is from 86 to 106 GHz, and the IF output is from 1 to 21 GHz, which covers 20 GHz. The insertion loss is between 5 and 8 dB.

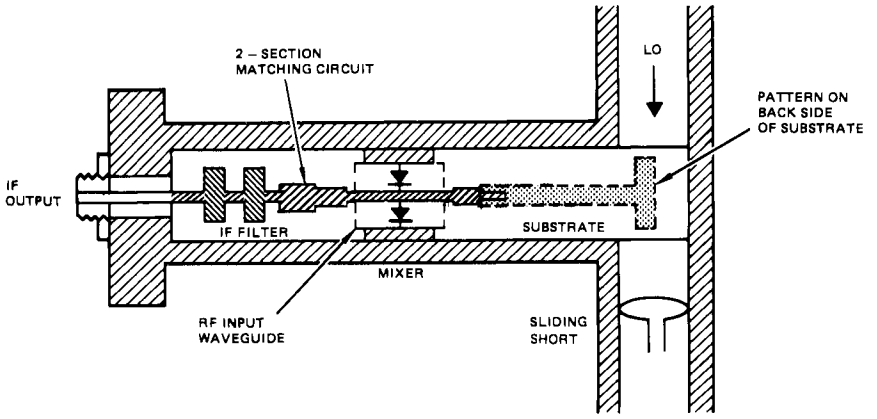


Figure 11.28. Cross-bar strip line mixer. (Courtesy of TRW Inc.)

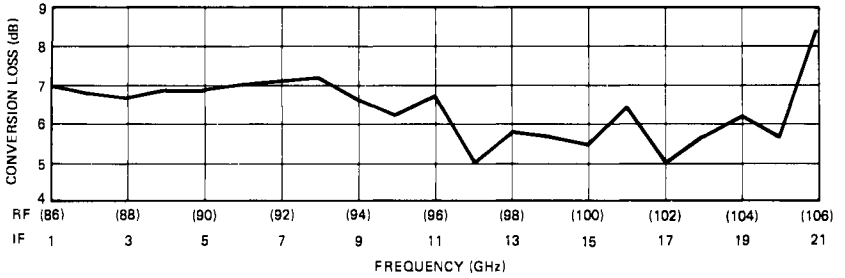


Figure 11.29. Typical performance of cross-bar strip line mixer. (Courtesy of TRW Inc.)

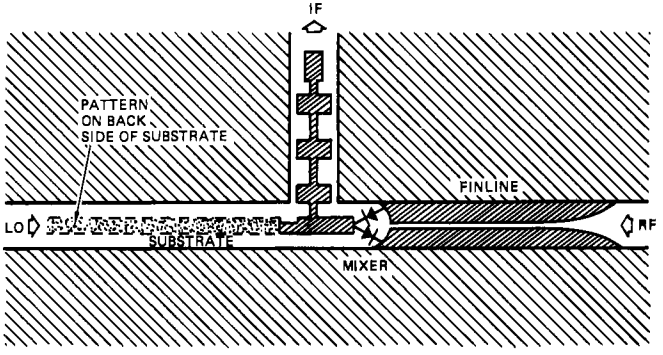


Figure 11.30. Fin line mixer. (Courtesy of TRW Inc.)

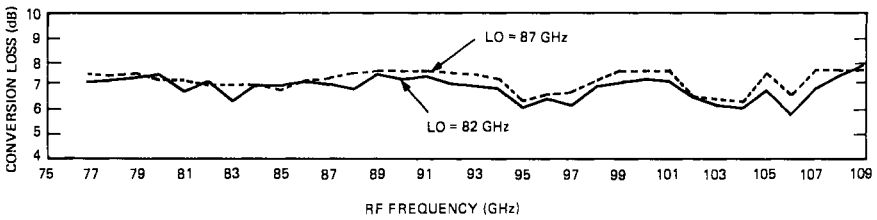


Figure 11.31. Typical performance of fin line mixer. (Courtesy of TRW Inc.)

Another kind of mixer, the fin line mixer, is shown in Figure 11.30. Two diodes are located at the center of the figure. The RF input is fed from the right side through the fin line to the two diodes. The LO is fed from the left side through the suspended strip line on the back of the substrate. The IF output is coupled from the top through an IF filter. Typical performance of a fin line mixer is shown in Figure 11.31. The insertion loss is between 7 and 8 dB. The input bandwidth is from 75 to 109 GHz, which covers the entire W band.

11.8. SOLID-STATE EHF SOURCES (81–96)

The revival of EHF research and development in recent years is attributed to the advent of new technology, especially the EHF sources. The development of field-effect transistors (FET), Gunn devices, IMPATT diodes, traveling wave tubes (TWT), magnetrons, and gyrotrons provide new signal sources for the frequency range. Although not all these sources will be used in receiver developments, they can be used as sources in radar transmitters. EW receivers, are needed only when EHF radars are available. Therefore, these sources will directly or indirectly affect the development of EW receivers. Since only solid-state EHF sources are suitable for LO sources in receivers, because of their small size, the discussion will be limited to solid-state sources.

A. FET Sources

If proper feedback circuitry is provided to an FET, it will perform as an oscillator. An FET signal source generates less power than a bulk device, that is, a Gunn oscillator, at EHF. Therefore, it is not popularly used as a frequency source, but it is often used as a frequency source at frequencies below 20 GHz.

B. Gunn Oscillators

A transferred electron device, first discovered by J. B. Gunn in 1963, is often referred to as the Gunn device. When a dc voltage is applied to a bar of *n*-type GaAs or InP, the current first increases linearly with voltage. When the average electric field increases beyond a threshold field of several kilovolts per centimeter, oscillation will start. The time of this oscillation is approximately equal

to the transient time of the carriers from the cathode to the anode. Subsequent research showed that these oscillations are a manifestation of the transferred electron effect.

Gunn oscillators are available at frequencies of from about 30 GHz to approximately 100 GHz. The power generated by a Gunn device is relatively low in comparison with an IMPATT diode of the same frequency range, but the spectrum is clean. Therefore, Gunn oscillators are the most popularly used LO sources in EHF receivers. To improve the stability of the LO frequency, Gunn oscillators can be phase locked to a frequency source with a crystal reference.

C. IMPATT Diode

An impact avalanche transit-time (IMPATT) silicon *pn* junction diode normally operates in a reverse biased mode. As the reverse bias voltage is increased, some electrons and holes in the region of high electric field will gain enough energy to impact the electrons and knock them off the atoms in the crystal lattice. This impact ionization will create electron–hole pairs. The electrons and holes thus created will be accelerated by the electric field and will cause additional impact ionization, which will result in avalanche breakdown. If the doping levels and *p* and *n* layer thicknesses are properly selected, the transit time delay associated with the electrons and holes moving across their respective drift region will generate RF energy in a chosen frequency.

IMPATT diodes can generate high power at EHF. The power capability is due to several factors. First, the device operates at the maximum electric field strength and carries enormous current densities. Second, good electronic efficiency can be achieved in the device. Third, silicon has good thermal conductivity, and thus heat can be readily removed.

Although the IMPATT diode can generate relatively high power, the spectrum purity is rather poor because the oscillation is caused by avalanche ionization. The IMPATT diode is often used in laboratories as a wide-band signal source. It has limited applications in EHF receivers because of its spectrum impurity. However, when the frequency measurement accuracy of an EHF receiver is not very stringent and the LO frequency is one at which the Gunn oscillator cannot produce enough power, the IMPATT diode is sometimes used as an LO.

11.9. BASIC EHF RECEIVER DESIGN CONCEPTS (97–103)

The EHF region covers a very wide frequency range (hundreds of gigahertz). Therefore, an EHF EW receiver should also cover a very wide frequency range to have a high probability of intercept (POI). Fortunately, the problem may not be as difficult as in the lower microwave frequency range because at least in the near future, the signal density in the EHF region is expected to be rather low. In addition, as shown in Figure 11.2, the possibility of building radar in

the antiwindow frequency range is questionable because of the high atmospheric attenuation. Therefore, whether EW receivers should cover the antiwindow region is also questionable.

Based on the above discussion, there are two approaches, at the present time, to building EHF receivers: crystal video and wide-band channelization. A crystal video receiver can be built to cover a very wide bandwidth. It can be built either in waveguide form or in other forms. If the crystal video receiver is built in waveguide form, its bandwidth is usually limited by the bandwidth of the waveguide used, because the detector diodes can cover a wider bandwidth than the waveguide. The problem of the crystal video receiver is its low sensitivity and the lack of capability to provide frequency information on the input signals. However, the receiver can be very simple and small in size.

Wide-band channelization can selectively down convert a certain frequency range of interest to a lower frequency range where amplification can be provided to improve the receiver sensitivity. In addition, other kinds of receivers can be added at the IF output to further measure the input signals. To make the channelization simple, the number of channels must be kept at a minimum. This implies that the IF bandwidth will be very wide. If the input frequency is below 60 GHz, low-noise RF amplifiers may be added before the mixer to further improve the sensitivity. If the input frequency is higher than 90 GHz, with today's technology the first component in the receiver should be a mixer because the amplifier at this frequency is not yet available. The disadvantage of wide-band channelization is the relatively complicated structure in comparison with a crystal video receiver.

The antenna beamwidth at EHF is usually very narrow, in contrast to the wide antenna beams at the lower microwave frequency range. To widen the spatial coverage, often the antenna beamwidth is deliberately widened. Since the antenna is very small in size at the EHF region, as a result, the antenna may become an integral part of the receiver. This may create some sensitivity measurement problems since one cannot feed a signal to the input of the receiver anymore. Instead, the input signal must be radiated to the input antenna. The input power must be carefully calibrated. Usually, an antenna with known gain and a power meter can be used to calibrate the input radiation power.

11.10. CONVENTIONAL CRYSTAL VIDEO RECEIVERS

This kind of receiver can be used only as a radar warning receiver (RWR) because it does not provide frequency information on the input signals and its sensitivity is low. The receiver is built in waveguide form, and the fundamental idea is very simple. One end of the waveguide is made into an antenna, and at the other end a video detector is installed. If the receiver is designed to receive signals of different polarizations, two receivers with orthogonal antennas can be used. The outputs of the two receivers can be combined at the video outputs. Figure 11.32 shows the schematic diagram of such a receiver. An actual receiver

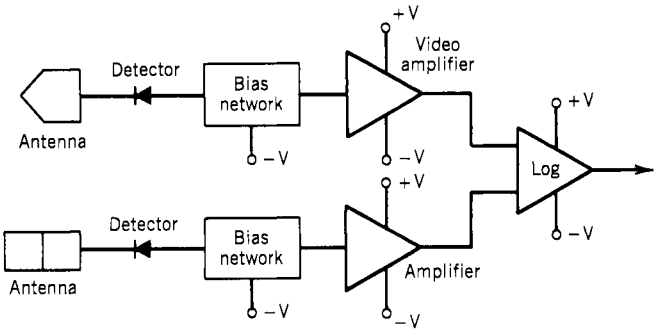


Figure 11.32. Schematic of an EHF crystal video receiver.

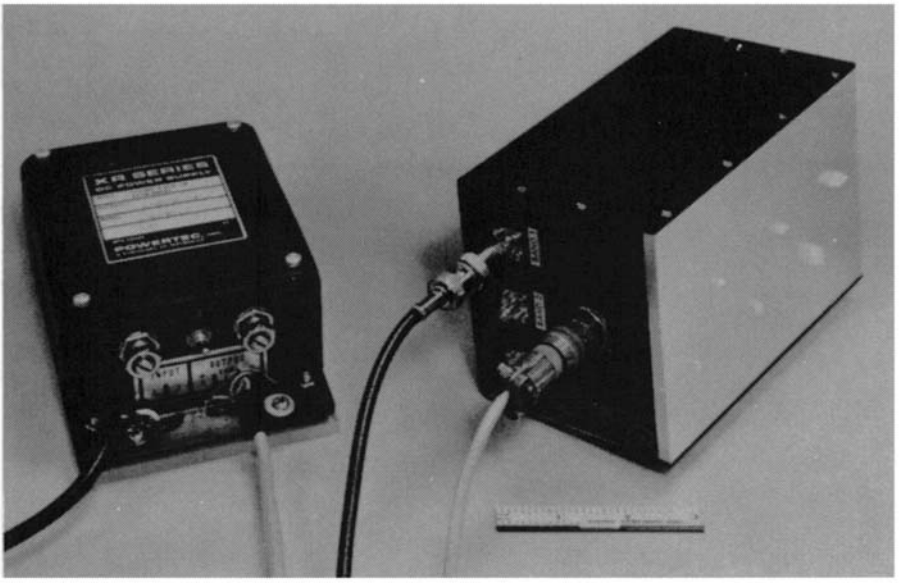


Figure 11.33. Three-band EHF receiver. (Fabricated by Hughes Aircraft Co., courtesy of Avionics Laboratory, AFWAL.)

is shown in Figure 11.33. The receiver covers three waveguide bandwidths with six inputs. Two inputs cover the same frequency range: One input is vertically polarized, and the other is horizontally polarized. The antenna is made by inserting a wedge-shaped dielectric piece in the opening of a rectangular waveguide that is flush with the conducting plate. The purpose of the dielectric insert is to widen the input beamwidth. The outputs from the two antennas of the same frequency bands are combined into one video output. To improve the sensitivity of the receiver, bias current is applied to the diode detectors.

11.11. INTEGRATED ANTENNA DETECTORS (104–110)

This kind of EHF receiver was first patented in 1967. It can be considered a crystal video receiver. The receiver consists of a spiral antenna and a detector. The spiral antenna is a conventional, covering the frequency band of interest. The following changes are made in the antenna: First, a diode detector is mounted at the conventional feeding port, the center of the antenna. Second, one of the outer end of the spiral arm is grounded, and the other outer end of the spiral arm is used as an output of the receiver. Figure 11.34 shows such a receiver.

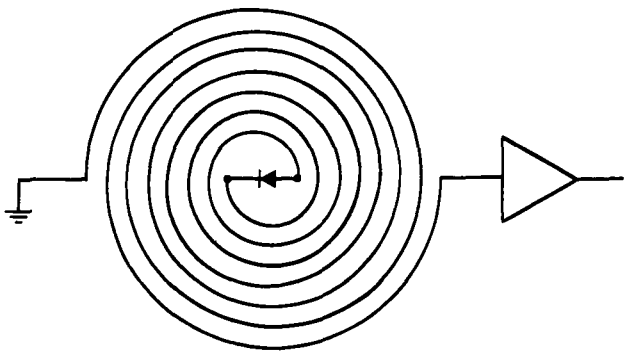


Figure 11.34. Integrated antenna detector.

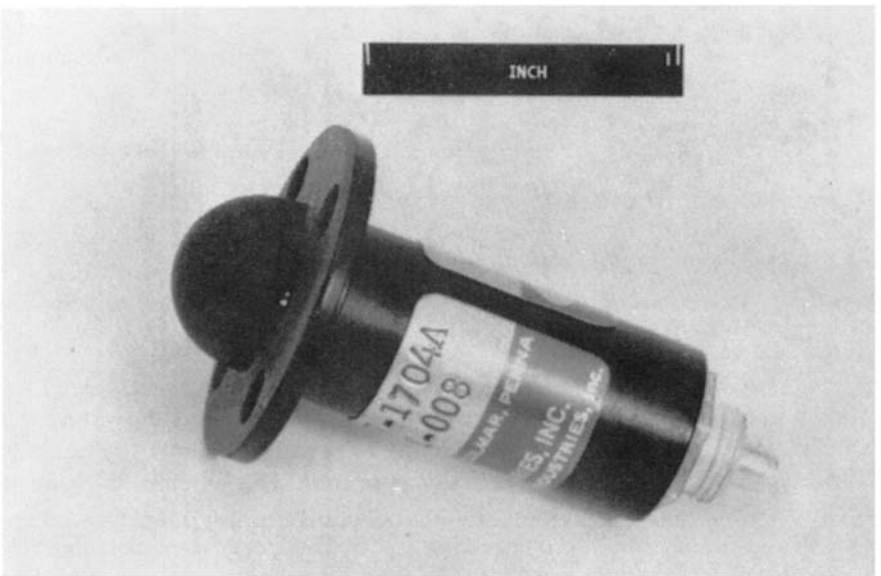


Figure 11.35. Actual integrated antenna detector (combined cavity backed spiral antenna and detector; 18–100 GHz). (Courtesy of American Electronic Laboratories, Inc.)

The name *antector*, standing for "antenna detector," is also used to describe this kind of receiver. The bandwidth of the receiver is the same as the bandwidth of the spiral antenna. Thus the bandwidth can be extremely wide and over several octaves. An actual integrated antenna detector receiver with a radome is shown in Figure 11.35.

The advantages of this kind of receiver are its extremely wide bandwidth, very small size, and simple and rigid structure. The only other component used in this receiver is a video amplifier. The major disadvantage of this receiver is the relatively low sensitivity limited by the detector.

If an LO is injected in the diode, it may convert the detector into a mixer. IF amplifiers can be added in front of the detector to improve the sensitivity of the receiver. However, this approach will increase the complexity of the receiver because an LO and an RF amplifier must be added.

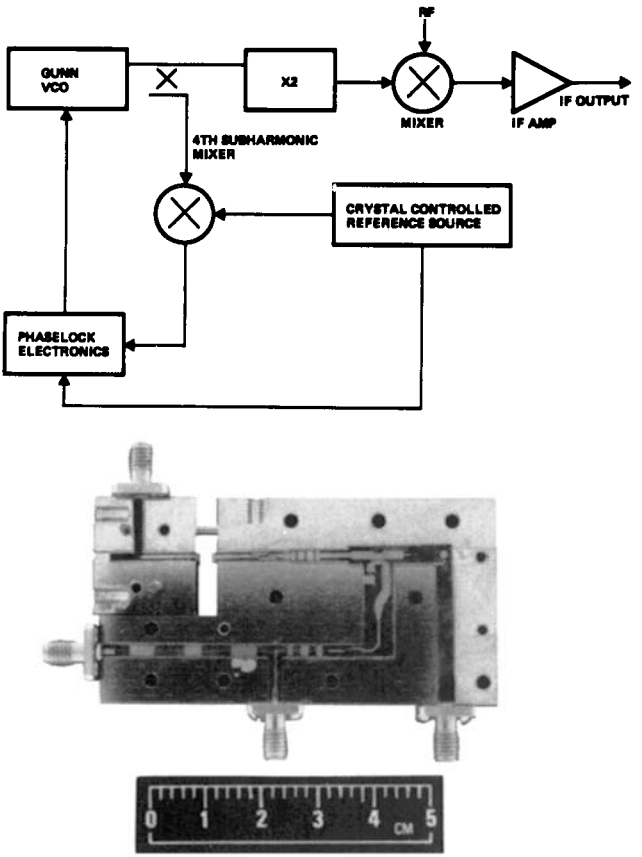


Figure 11.36. EHF down converter. (Courtesy of TRW Inc.)

11.12. WIDE-BANDWIDTH CHANNELIZATION

In this approach, the input frequency is first channelized into many parallel outputs by a frequency multiplexer. At the outputs of each output, a down converter is used to convert the outputs into a common IF frequency range like a conventional channelized receiver. A conventional microwave receiver can be used to measure the frequency of the input signals. Sometimes a video detector can also be placed at each output of the down converter to sense the activity of the band.

Channelization at the EHF is the same as at the microwave frequency. The difference is the technology used. In the EHF region, MMIC technology can be used. Figure 11.36 shows a EHF down converter. In this figure, the input is fed from a waveguide (not shown). The rest of the circuit is in MMIC form. The mixer is a cross bar type in suspended stripline form with two diodes. The LO is a Gunn oscillator, and it can be phase locked to a reference signal. However, the phase-locked loop is not shown in this figure. A frequency doubler is used to obtain the desired frequency since the fundamental frequency of the Gunn oscillator in MMIC form is not high enough. A high-pass filter is fabricated in the circuit, which will limit the fundamental frequency.

At the outputs of the wide-band channelizer, instantaneous frequency measurement (IFM) receivers can be used to further process the input signals An IFM receiver can cover a very wide bandwidth and provides fine-frequency information on the input signals. If the signal density at EHF is low, overlapping pulses will not cause a big problem to the IFM receiver.

11.13. SUMMARY

Advances in millimeter wave components have revived research and development in EHF receivers. There are presently two general approaches to fabricate EHF receivers: the crystal video receiver and wide-band channelization. In both approaches, the small size of the receiver is very much emphasized. The miniaturization includes not only the receivers but also the antennas. The advances in MMIC will have a big impact on EHF receiver development. In the future, EHF receivers may be made in integrated circuit forms.

REFERENCES

1. H. P. Westman, Ed., *Reference Data for Radio Engineers*, 5th ed., Howard W. Sams, Indianapolis, IN, 1972.
2. M. I. Skolnik, *Introduction to Radar Systems*, McGraw-Hill, New York, 1962.
3. K. J. Button and J. C. Wiltse, Eds., *Infrared and Millimeter Waves*, Academic Press, New York, 1981.
4. T. K. Ishii, *Microwave Engineering*, Ronald Press, New York, 1966.
5. R. E. Collin, *Foundations for Microwave Engineering*, McGraw-Hill, New York, 1966.

6. F. J. Tischer, Experimental attenuation of rectangular waveguides at millimeter wavelengths, *IEEE Trans. Microwave Theory Techniques*, **MTT-27**, 31–37 (1979).
7. H. J. Hindin and J. J. Taub, Oversize waveguide directional coupler, *IRE Trans. Microwave Theory Techniques*, **MTT-10**, 394–395 (1962).
8. H. A. Wheeler, Transmission line properties of parallel strips separated by dielectric sheet, *IEEE Trans. Microwave Theory and Techniques*, **MTT-13**, 72–185 (1965).
9. M. Caulton and H. Sobol, Microwave integrated circuit technology—A survey, *IEEE J. Solid State Circuits*, **SC-5**, 292–303 (1970).
10. H. A. Wheeler, Transmission line properties of a strip line between parallel plates, *IEEE Trans. Microwave Theory Techniques*, **MTT-26**, 866–876 (1978).
11. I. J. Bahl and R. Garg, A designer's guide to strip line circuits, *Microwaves*, **19**, 90 (January 1980).
12. H. A. Wheeler, Transmission line properties of a strip on a dielectric sheet on a plane, *IEEE Trans. Microwave Theory Techniques*, **MTT-25**, 631–647 (1977).
13. H. E. Stinehelfer, An accurate calculation of uniform microstrip transmission lines, *IEEE Trans. Microwave Theory Techniques*, **MTT-16**, 439–444 (1968).
14. E. Yamashita and R. Mittra, Variational method for the analysis of microstrip lines, *IEEE Trans. Microwave Theory Techniques*, **MTT-16**, 251–255 (1968).
15. E. Yamashits, Variational method for the analysis of microstrip-like transmission lines, *IEEE Trans. Microwave Theory Techniques*, **MTT-16**, 529–535 (1968).
16. H. Garg, Stripline like microstrip configuration, *Microwave J.*, **22**, 103 (April 1979).
17. R. A. Pucel, D. J. Masse, and C. P. Hartwig, Losses in microstrip, *IEEE Trans. Microwave Theory Techniques*, **MTT-16**, 342–350 (1968).
18. S. B. Cohn, Slot line on a dielectric substrate, *IEEE Trans. Microwave Theory Techniques*, **MTT-17**, 768–788 (1969).
19. E. A. Mariani, C. P. Heinzman, J. P. Agries, and S. B. Cohn, Slot line characteristics, *IEEE Trans. Microwave Theory Techniques*, **MTT-17**, 1091–1096 (1969).
20. N. Samardrija and T. Itoh, Double layered slot line for millimeter wave integrated circuits, *IEEE Trans. Microwave Theory Techniques*, **MTT-24**, 827–831 (1976).
21. K. C. Gupta, R. Garg, and I. J. Bahl, *Microstrip Lines and Slotlines*, Artech House, Dedham, MA, 1979.
22. R. Vogel, Microstrip slotline components for microwave ICs, *Microwave J.*, **23**, 83 (May 1980).
23. M. Aikawa and H. Ogawa, A new MIC magic T using coupled slotlines, *IEEE Trans. Microwave Theory Techniques*, **MTT-28**, 523–528 (1980).
24. L. P. Schmidt and T. Itoh, Characteristics of generalized fin-line for millimeter wave integrated circuits, *Int. J. Infrared Millimeter Waves*, **2**, 427–436 (1981).
25. P. J. Meier, Integrated fine line millimeter components, *IEEE Trans. Microwave Theory Techniques*, **MTT-22**, 1209–1216 (1974).
26. L. D. Cohen and P. J. Meier, E-plane mm wave circuits, *Microwave J.*, **21**, 63 (August 1978).
27. A. M. K. Saad and K. Schunemann, A simple method for analyzing fin line structures, *IEEE Trans. Microwave Theory Techniques*, **MTT-26**, 1002–1007 (1978).
28. R. N. Bates, M. D. Coleman, S. J. Nightingale, and R. Davies, E-planes drop millimeter costs, *Microwave System News*, **10**, 74 (December, 1980).
29. A. Beyer and K. Solbach, Fin line ferrite isolator for integrated millimeter wave circuits, IEEE MTT-Symposium, pp. 296–298, University of Duisburg, Germany, June, 1981.
30. H. Meinel and H. Callsen, Fin line pin diode attenuators and switches for 94 GHz range, *Electronics Letters*, **18**, 541–542 (June, 1982).
31. L. W. Koh, I. M. H. Williamson, M. L. Nyss, O. Gadoury, and D. E. Wheeler, Ka-band fin line SPDT switch, *Microwave J.*, **26**, 105 (June, 1983).
32. P. J. Meier, Two new integrated circuit media with special advantages at millimeter wavelengths, *IEEE G-MTT Symposium Digest*, pp. 221–223, May 22–24, 1972.
33. D. Rubin and A. R. Hislop, Millimeter-wave coupled line filter—design techniques for sustended substrate and microstrip, *Microwave J.*, **23**, 67 (October, 1980).
34. E. A. J. Marcatili, Dielectric rectangular waveguide and directional coupler for integrated optics, *Bell System Technical J.*, **48**, 2071–2102 (September, 1969).

35. M. V. McLevige, T. Itoh, and R. Mittra, New waveguide structures for millimeter wave circuits, *IEEE Trans. Microwave Theory Techniques*, **MTT-23**, 788–794 (1975).
36. H. Jacobs, G. Novick, C. M. Locascio, and M. M. Chrepta, Measurement of guide wavelength in rectangular dielectric waveguide, *IEEE Trans. Microwave Theory Techniques*, **MTT-24**, 815–820 (1976).
37. R. Knox, Dielectric waveguide microwave integrated circuits—An overview, *IEEE Trans. Microwave Theory Techniques*, **MTT-24**, 806–814 (1976).
38. T. Itanami and S. Shindo, Channel dropping filter for millimeter wave integrated circuits, *IEEE Trans. Microwave Theory Techniques*, **MTT-26**, 759–764 (1978).
39. G. M. Lindgren, Coupler design in open dielectric waveguide with web registration, *IEEE MTT-Symposium*, pp. 11–13, 1981.
40. D. D. King, Dielectric image line, *J. Applied Physics*, **23**, 699 (1952).
41. T. Itoh, Inverted strip dielectric waveguide for millimeter wave integrated circuits, *IEEE Trans. Microwave Theory Techniques*, **MTT-24**, 821–827 (1976).
42. J. E. Raue et al., Defense and Space Systems Group, TRW Inc., private communication.
43. K. L. Russell, Microwave power combining techniques, *IEEE Trans. Microwave Theory Techniques*, **MTT-27**, 472–478 (1979).
44. A. A. Saleh, Planar electrically symmetric N-way hybrid power dividers/combiners, *IEEE Trans. Microwave Theory Techniques*, **MTT-28**, 555–563 (1980).
45. K. Chang and C. Sun, Millimeter-wave power-combining techniques, *IEEE Trans. Microwave Theory Techniques*, **MTT-31**, 91–107 (1983).
46. Y. Konishi and K. Uenakada, The design of a bandpass filter with inductive strip–planar circuit mounted in waveguide, *IEEE Trans. Microwave Theory Techniques*, **MTT-22**, 869–873 (1974).
47. R. Vahldieck, J. Bornemann, F. Arndt, and D. Grauerholz, W-band low insertion loss E-plane filter, *IEEE Trans. Microwave Theory Techniques*, **MTT-32**, 133–135 (1984).
48. K. Solbach, The status of printed millimeter-wave E-plane circuits, *IEEE Trans. Microwave Theory Techniques*, **MTT-31**, 107–121 (1983).
49. R. F. Soohoo, *Theory and Application of Ferrites*, Prentice-Hall, Englewood Cliffs, NJ, 1960.
50. B. A. Auld, The synthesis of symmetrical waveguide circulators, *IRE Trans. Microwave Theory Techniques*, **MTT-7**, 238–247 (1959).
51. U. Milano, J. Saunders, and L. Davis, A Y-junction strip line circulator, *IRE Trans. Microwave Theory Techniques*, **MTT-8**, 346–351 (1960).
52. T. K. Ishii and J. B. Y. Tsui, Millimeter-wave field displacement type isolator with short ferrite strips, *Proc. IRE*, **49**, 975 (1961).
53. T. K. Ishii and J. B. Y. Tsui, Effects for ferrite strip mounting positions on millimeter wave isolator characteristics, *IRE Trans. Microwave Theory Techniques*, **MTT-9**, 362 (1961).
54. F. Wang, T. K. Ishii, and J. B. Y. Tsui, Ferrimagnetic resonance of single crystal barium ferrite in the millimeter wave region. *J. Applied Physics*, **32**, 1621 (1961).
55. H. Bosma, On strip line Y-circulator at UHF, *IEEE Trans. Microwave Theory Techniques*, **MTT-12**, 61–72 (1964).
56. C. E. Fay and R. L. Comstock, Operation of the ferrite junction circulator, *IEEE Trans. Microwave Theory Techniques*, **MTT-13**, 15–27 (1965).
57. J. W. Simon, Broadband strip line transmission Y-junction circulators, *IEEE Trans. Microwave Theory Techniques*, **MTT-13**, 335–345 (1965).
58. Y. S. Wu and F. J. Rosenbaum, Wide-band operation of microstrip circulators, *IEEE Trans. Microwave Theory Techniques*, **MTT-22**, 849–856 (1974).
59. Y. Ayasli, Analysis of wide band strip line circulators by integral equation technique, *IEEE Trans. Microwave Theory Techniques*, **MTT-28**, 200–209 (1980).
60. T. Miyoshi and S. Miyanchi, The design of planar circulators for wide band operation, *IEEE Trans. Microwave Theory Techniques*, **MTT-28**, 210–214 (1980).
61. G. C. Dacey and I. M. Ross, Unipolar field-effect transistor, *Proc. IRE*, **41**, 970–979 (1953).
62. G. C. Dacey and I. M. Ross, The field effect transistor, *Bell System Technical J.*, **34**, 1149–1189 (1955).

63. C. A. Liechti, Microwave field effect transistors—1976, *IEEE Trans. Microwave Theory Techniques*, **MTT-24**, 279–300 (1976).
64. C. A. Liechti, GaAs FET technology: a look into the future, *Microwaves*, **17**, 44–49 (October, 1978).
65. S. Y. Liao, *Microwave Devices and Circuits*, Prentice-Hall, Englewood Cliffs, NJ, 1980.
66. S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed., Wiley, New York, 1981.
67. L. E. Dickens, Low conversion loss millimeter wave mixers, *Digest MTT Symposium*, pp. 66–68, 1973.
68. W. Snell and M. V. Schneider, Millimeter wave thin film downconverter, *IEEE Trans. Microwave Theory Techniques*, **MTT-24**, 804–806 (1976).
69. A. Hislop and R. T. Kihm, A broad band 40–60 GHz balanced mixer, *IEEE Trans. Microwave Theory Techniques*, **MTT-24**, 63–64 (1976).
70. L. Yuan, Design and performance analysis of an octave bandwidth waveguide mixer, *IEEE Trans. Microwave Theory Techniques*, **MTT-25**, 1048–1054 (1977).
71. D. Held and A. R. Kerr, Conversion loss and noise of microwave and millimeter wave mixer, Part 1 and 2, *IEEE Trans. Microwave Theory and Techniques*, **MTT-26**, 49–61 (1978).
72. W. Kelly and G. T. Wrixon, Conversion losses in Schottky barry diode mixers in the submillimeter region, *IEEE Trans. Microwave Theory Techniques*, **MTT-27**, 665–672 (July 1979).
73. J. Calviello, Advanced devices and components for millimeter and submillimeter systems, *IEEE Trans. Electron Devices*, **ED-26**, 1273–1281 (1979).
74. R. Kawasaki and M. Akaike, A broad band second harmonic mixer covering 76–106 GHz, *IEEE Trans. Microwave Theory Techniques*, **MTT-26**, 425–427 (1978).
75. C. Gupta and J. D. Conte, Trends in mm-wave mixer designs, *Microwave J.*, **27**, 83 (April 1984).
76. C. Chao, A. Contolatis, S. A. Jamison, and P. E. Bauhahn, Ka-band monolithic GaAs balanced mixers, *IEEE Trans. Microwave Theory Techniques*, **MTT-31**, 11–15 (1983).
77. R. Tahim, G. M. Hayashibara, and K. Chang, Design and performance of W-band broad band integrated circuit mixers, *IEEE Trans. Microwave Theory Techniques*, **MTT-31**, 277–283 (1983).
78. P. Meier, E-plane components for a 94 GHz printed circuit balanced mixer, *Digest MTT Symposium*, pp. 267–269, 1980.
79. L. Bui, Broadband planar balanced mixers for millimeter wave applications, *Digest MTT symposium*, pp. 204–205, 1982.
80. P. Parrish, A. G. Gardiasmenos, and I. Galin, 94 GHz beam-lead balanced mixer, *IEEE Trans. Microwave Theory Techniques*, **MTT-29**, 1150–1157 (1981).
81. J. B. Gunn, Microwave oscillations of current in III–V semiconductors, *Solid State Communications*, **1**, 89–91 (1963).
82. J. A. Copeland, CW operation of LSA oscillator diodes—44 to 88 GHz, *Bell System Technical J.*, **46**, 284–287 (1967).
83. S. Y. Narayan and F. Sterzer, Transferred electron amplifiers and oscillators, *IEEE Trans. Microwave Theory Techniques*, **MTT-18**, 773–783 (1970).
84. T. P. Lee and C. A. Burrus, A millimeter wave quadrupler and up-converter using planar diffused gallium arsenide varactor diodes, *IEEE Trans. Microwave Theory Techniques*, **MTT-16**, 287–296 May (1968).
85. M. V. Schneider and W. W. Snell, Jr., A scaled hybrid integrated multiplier from 10 to 30 GHz, *Bell System Technical J.*, **50**, 1933–1942 (1971).
86. R. Hamilton, R. D. Fairman, S. I. Long, M. Ornori, and F. B. Fank, InP Gunn effect for millimeter wave amplifiers and oscillators, *IEEE Trans. Microwave Theory Techniques*, **MTT-24**, 775–780 (1976).
87. T. Takada, Hybrid integrated frequency multipliers at 300 and 450 GHz, *IEEE Trans. Microwave Theory Techniques*, **MTT-26**, 733–737 (1978).
88. J. Archer, Millimeter wavelength frequency multipliers, *IEEE Trans. Microwave Theory Techniques*, **MTT-29**, 552–557 (1981).

89. D. Smith, Fully integrated W-band microstrip oscillator, *Electronics Letters*, **19**, 222–223 (1983).
90. C. Rauscher, High frequency doubler operation of GaAs field effect transistors, *IEEE Trans. Microwave Theory Techniques*, **MTT-31**, 462–473 (1983).
91. M. Smith, Miniature mm-wave source advances phase-lock technology, *Microwave System News*, **14**, 84 (January 1984).
92. W. Haydl, Fundamental and harmonic operation of millimeter wave Gunn diodes, *IEEE Trans. Microwave Theory Techniques*, **MTT-31**, 879–889 (1983).
93. W. T. Read, A proposed high frequency negative resistance diode, *Bell System Technical J.*, **37**, 401–446 (1958).
94. C. A. Lee, R. L. Batdorf, W. Weigmann, and G. Kaminsky, The Read diode—an avalanche, transit-time, negative resistance oscillator, *Applied Physics Letters*, **6**, 89–91 (1965).
95. P. A. Blakey, B. Culshaw, and R. A. Giblin, Comprehensive models for the analysis of high efficiency GaAs IMPATTS, *IEEE Trans. Electron Devices*, **ED-25**, 674 (1978).
96. Special issue on solid-state microwave millimeter-wave power generation, amplification, and control, *IEEE Trans. Microwave Theory Techniques*, **MTT-27** (1979).
97. J. Whelehan, Low noise millimeter-wave receivers, *IEEE Trans. Microwave Theory Techniques*, **MTT-25**, 268–280 (1977).
98. K. Chang, K. Louie, A. J. Grote, R. S. Tahim, M. J. Mlinar, G. M. Hayashibara, and C. Sun, V-band low noise integrated circuit receiver, *IEEE Trans. Microwave Theory Techniques*, **MTT-31**, 149–154 (1983).
99. T. Oxley, Image guide and microstrip integrated W-band receivers, *Microwave J.*, **26**, 117 (1983).
100. A. Raisanen, An ultra low-noise Schottky mixer receiver at 80–120 GHz, *Digest*, 6th International IR & MM waves, pp. w3–8, Miami, Florida, 1981.
101. C. Hu and A. Denning, A broad-band low-noise receiver at W-band, *IEEE MTT-S Digest*, MM Symposium, pp. 111–113, 1981.
102. N. Erickson, A 200–350 GHz heterodyne receiver, *IEEE Trans. Microwave Theory Techniques*, **MTT-29**, 557–561 (1981).
103. W. Wilson, Submillimeter-wave receivers—A status report, *IEEE Trans. Microwave Theory Techniques*, **MTT-31**, 873–878 (1983).
104. E. M. Turner, Combined antenna and converter circuit, U.S. Patent 3,246,245, April 12, 1966.
105. J. R. Copeland, W. J. Robertson, and J. C. Golfert, Combined antenna and tunnel diode converter circuit, U.S. Patent 3,296,536, January 3, 1967.
106. S. N. Andre and E. M. Davis, Printed circuit spiral antenna having amplifier and bias feed circuits integrated therein, U.S. Patent 3,509,465, April 28, 1970.
107. K. M. Jagdmann and H. R. Phelan, Direct fed spiral antenna, U.S. Patent 3,949,407, April 6, 1976.
108. K. R. Moser, Radiated input mixer, U.S. Patent 4,287,603, September 1, 1981.
109. R. P. Flam, Integrated spiral antenna–detector device, U.S. Patent 4,319,248, March 9, 1982.
110. O. P. Gandhi, Microwave dosimeter, U.S. Patent 4,368,472, January 11, 1983.

Chapter 12

Measurements of EW Receivers

12.1. INTRODUCTION

The evaluations of EW receivers will be discussed in this chapter. Many of the characteristics of receivers are discussed in Chapters 2 and 3, and their actual measurement will be included in this chapter. Examples will be used to illustrate how a receiver is evaluated, and its results will be displayed. The examples are obtained from many different receivers.

Generally speaking, an EW receiver should be tested under different conditions in order to obtain its performance. Although a general approach to evaluate all types of receivers is desirable, it is likely that some special tests are required for some special performance. It is almost impossible to cover all the measurements performed on a receiver. Thus only some typical tests will be discussed.

An EW receiver usually covers a wide frequency and dynamic range. It will take many input conditions to fully evaluate the receiver. Therefore, a computer-controlled system must be used to test a receiver. Another concern is how to present the data. Of course, the simplest way is to print the data on a pulse-by-pulse basis. But this approach does not provide a perspective view of the receiver. The best approach is to present the data in graphics. The discussion in this chapter is based on the experimental setup in the Air Force Avionics Laboratory.

12.2. DIFFERENT TYPES OF TESTS

The types of tests discussed in this chapter can be divided into three categories, defined as follows.

The first test, referred to as a laboratory test, is limited to two input signals. This test provides the most information on a receiver. It evaluates the frequency, pulse amplitude (PA), pulse width (PW), and time-of-arrival (TOA) measurement capabilities of the receiver. It also determines the sensitivity, one-signal dynamic range, and spur-free and instantaneous dynamic ranges. However, the AOA measurement capability of the receiver cannot be evaluated with this test.

The second test is a simulator test. A simulated electronic environment is generated through many signal generators. The environment can be either static or dynamic. The outputs of the receiver can be processed in two ways. If a digital processor is available, the outputs from the receiver can be processed directly by the processor, and the data will be displayed at near real time. The advantage of this approach is that the receiver can be tested under different input conditions. The disadvantage is that the performance of the digital processor must be taken into consideration. The outputs of the receiver can also be stored in memory for further analysis. This will eliminate the effect of the digital processor, but the test will take longer time.

The third test, a field test, entails testing the receiver against some real radar. The differences between the simulator test and the field test are as follows:

1. In the simulator test, the signals are preprogrammed; in the field test, the receiver operator has no control over the input signals.
2. In the field test, an antenna is used; in the simulator test, the input signals are fed to the receiver through cables.
3. The AOA can be evaluated more realistically in a field test; in the simulator test, the AOA is simulated. The simulation of a wave front to evaluate a phase comparison AOA system is not an easy task.
4. A simulator test costs less and is easier to handle than a field test.

The discussion presented here will concentrate on laboratory tests because they provide the most information. In the laboratory test, measurements of frequency, sensitivity, and dynamic range are emphasized. For the simulator and field tests, since their results are similar, only the results of the field test will be presented.

12.3. LABORATORY TEST SETUP

The heart of the laboratory test is a computer. The computer will perform four basic functions:

1. Control the input signal conditions that is setting the input frequency, PA, and PW;

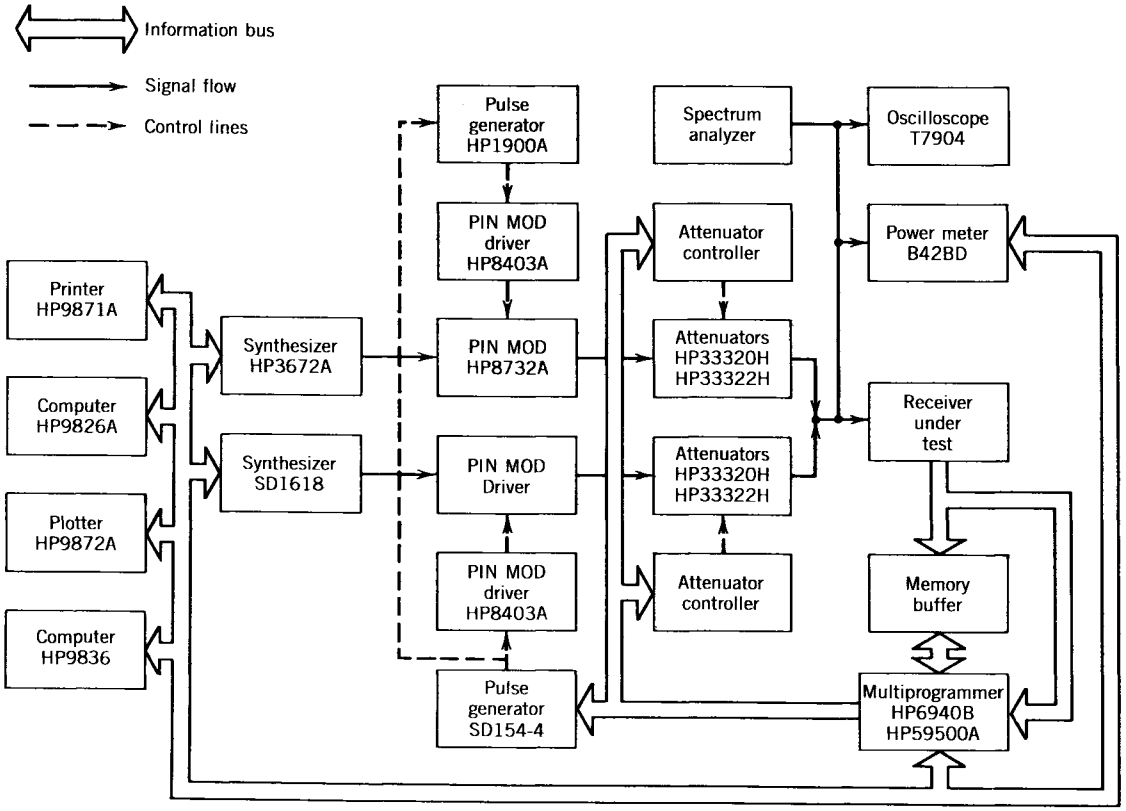


Figure 12.1. Experimental setup (automated receiver testing flowchart).

2. collect the digital outputs from the receiver and convert them to meaningful parameters;
3. compare the output data against the input information; and
4. present the data in a form that easily represents the performance of the receiver.

The actual experimental setup is shown in Figure 12.1. The thick lines represent the information bus, the dashed lines are control lines, and the solid lines are signal paths. The computer is the main control unit. The printer, plotter, spectrum analyzer, oscilloscope, and power meter can be considered as support equipment. The two synthesizers are the signal generators. The pulse generator and pin modulators with their drivers are used to generate pulses from the signal sources. Attenuators with controllers are used to control the PA of the input signals. However, the PA can also be changed by the attenuators associated with the signal generators. The multiprogrammer is an interface unit between the receiver under test and the computer. All the data collected are presented in graphic form.

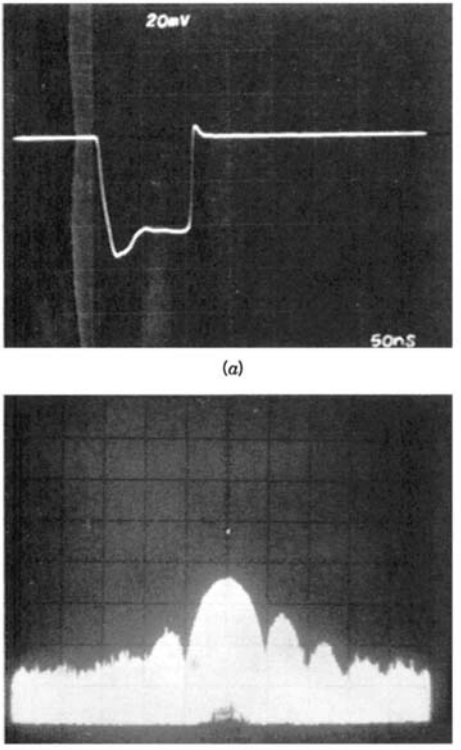
12.4. CALIBRATIONS

Before the actual tests are started, the laboratory setup should be calibrated. Since the signals are generated from frequency synthesizers, they require no calibration. Usually, the PW generated by the pin modulators is rather accurate, above 200 nsec. In other words, the RF PW equals the control PW from the pulse generator. When the PW is below 200 nsec, the PW of the pulse generator may not equal the PW of the RF signal. In addition, it is difficult to measure the actual PW of a narrow pulse because of the irregularity of the pulse shape. Figure 12.2 demonstrates this phenomenon. Therefore, the RF is often measured by a spectrum analyzer to calibrate the short pulse. Figure 12.2*b* shows the power spectrum output of the pulse in Figure 12.2*a*. The pulse is determined by the spectrum width of the main lobe. The PW is related to the width of the main lobe by

$$\text{PW} = 2/W \tag{12.1}$$

where W is the width of the main lobe.

The output power from the frequency synthesizer is leveled over a wide frequency range. However, from the output of the signal generators to the input of the receiver, there are several RF components, and their performance is usually frequency dependent. Therefore, the power at the input of the receiver needs to be calibrated. The power meter is used to calibrate the input power. The input power is measured across the frequency of interest at a fixed frequency interval. The power measurements are used to generate a set of calibration constants stored in the computer. Figure 12.3 shows uncalibrated input power.



(a)

(b)

Figure 12.2. Display of a 100-nsec pulse. (*a*) time domain (*b*) frequency domain.

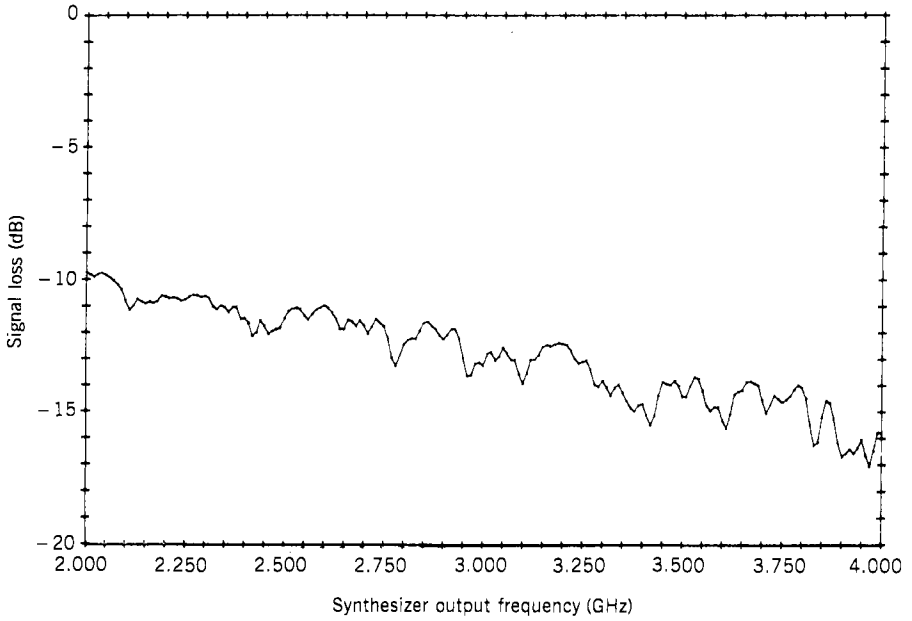


Figure 12.3. Uncalibrated input power.

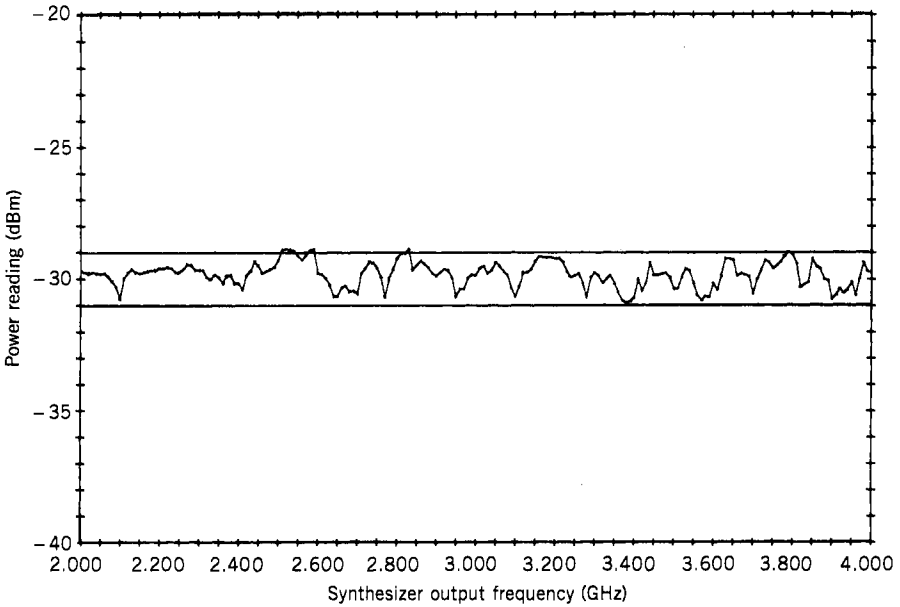


Figure 12.4. Calibrated input power.

The power is lower at a higher frequency, as expected. Figure 12.4 shows the results of the calibrated input. The power level is within ± 1 dB.

12.5. FREQUENCY ACCURACY MEASUREMENT

Frequency measurement is one of the most important factors in an EW receiver. It is also the first one to be measured. The results of this test are used in the sensitivity and dynamic range measurements. Usually, this measurement is performed at different power levels and various PWs. From these results, a general idea of the sensitivity and minimum PW capability of the receiver can be obtained.

During the test, one pulsed signal of known frequency is sent to the receiver, and the measured frequency output of the receiver is collected immediately. The input frequency versus the difference (error) frequency is plotted at each measurement. Although the results can be plotted as input versus output frequency, the presentation usually does not have enough resolution to display the frequency variations when testing a receiver of wide bandwidth. Therefore, the input versus frequency error approach is a better presentation.

If the error frequency is consistent, as shown in Figure 12.5, usually one reading is required at each frequency. However, in some receivers, the frequency error is not consistent, as shown in Figure 12.6, and occasionally high-frequency

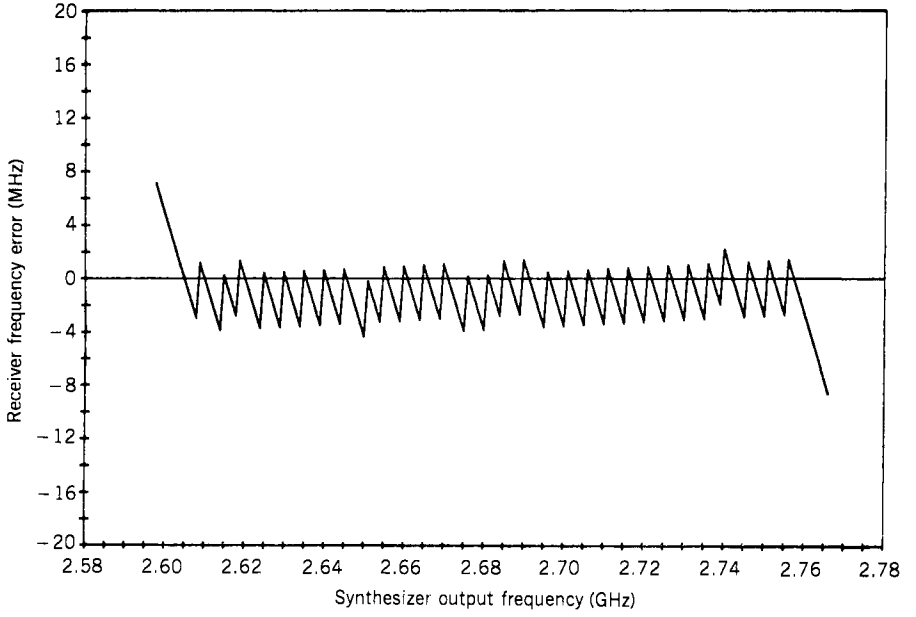


Figure 12.5. Frequency error measurement.

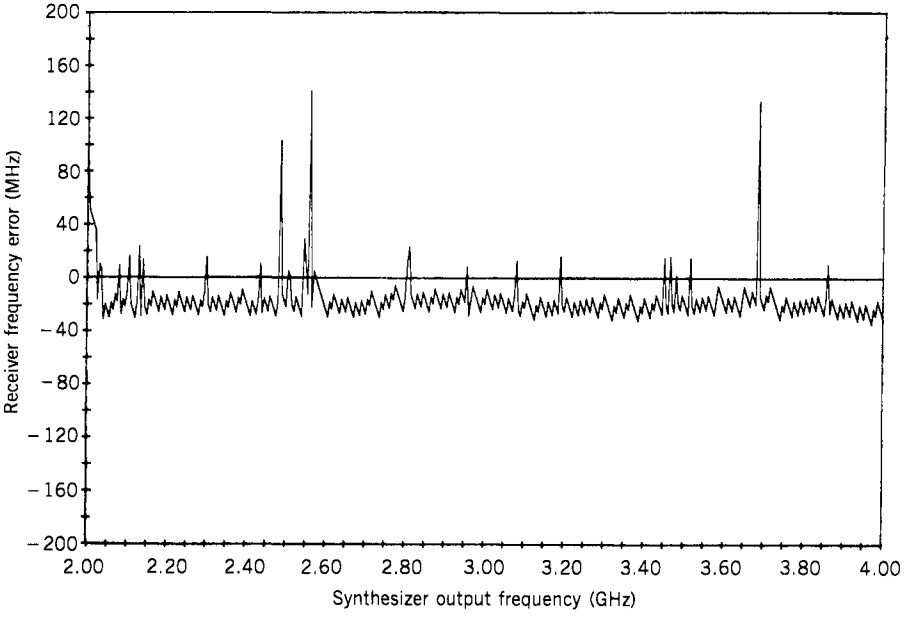


Figure 12.6. Frequency error measurement with some large errors.

errors are measured. This is very undesirable in a receiver because the erroneous frequency data may create confusion in the digital processor following the receiver. Receivers with such inconsistent results are often measured many times under the same input conditions. The average error frequency versus input frequency is plotted in addition to the single-pulse case. Figure 12.7 shows the results of a frequency error averaging measurement. In this plot, 100 signals were sent to the receiver at each frequency step. The average error frequency is obtained from 100 measurements. The results are relatively smooth compared with the results in Figure 12.6.

The same receiver is used to generate the results in Figures 12.6 and 12.7. If a receiver is properly designed, the magnitude of the positive error frequency should be approximately equal to the magnitude of the negative error frequency. In other words, the overall average of the error frequency should be close to zero. If the average is not close to zero, the frequency measured is considered biased. The bias can be corrected in the software of the receiver testing program. The bias in Figure 12.6 is corrected through the software, as shown in Figure 12.7.

Figure 12.8 shows the results of the frequency measurement at an input power level close to the sensitivity of the receiver. At the center of the receiver band, the sensitivity is lower and a hole (missing pulses) is formed. This means that the input power is too low to be detected. Thus, this curve can give a rough idea of the sensitivity of the receiver.

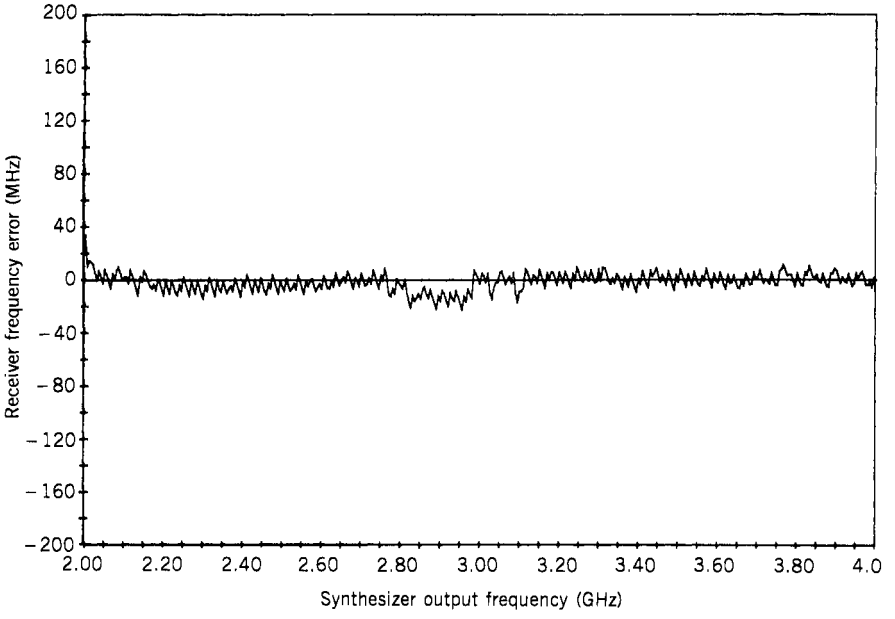


Figure 12.7. Average error frequency versus input frequency: number of samples per increment = 100.

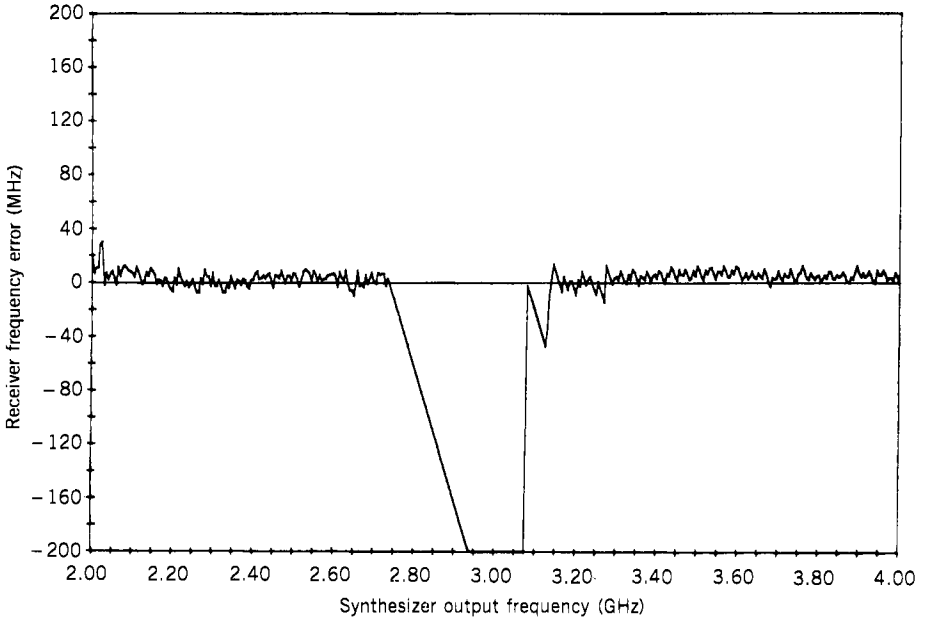


Figure 12.8. Frequency error plot with input signal close to receiver sensitivity level: number of samples per increment = 100.

12.6. STANDARD DEVIATION FREQUENCY MEASUREMENT

The standard deviation frequency measurement applies only to receivers where the standard frequency measurement shows inconsistent results, as illustrated in Figure 12.6. Standard deviation is similar to the average frequency error measurement with the exception of the data being presented differently.
The average frequency is obtained as

$$f_{\text{avg}} = \sum_{i=1}^{N} f_i \Big/ N \tag{12.2}$$

and the standard deviation is defined as

$$f_{\text{sd}} = \sqrt{\sum_{i=1}^{N} (f_i - f_{\text{avg}})^2 \Big/ N} \tag{12.3}$$

where N is the total number of measurements, i is an integer ($=1, 2, \ldots, N$); f_i is the output frequency of the ith measurement, f_{avg} is the average frequency from the N measurements, and f_{sd} is the frequency standard deviation at any particular frequency. The smaller the standard deviation, the more consistent the frequency measurement. Figure 12.9 shows the standard deviation frequency

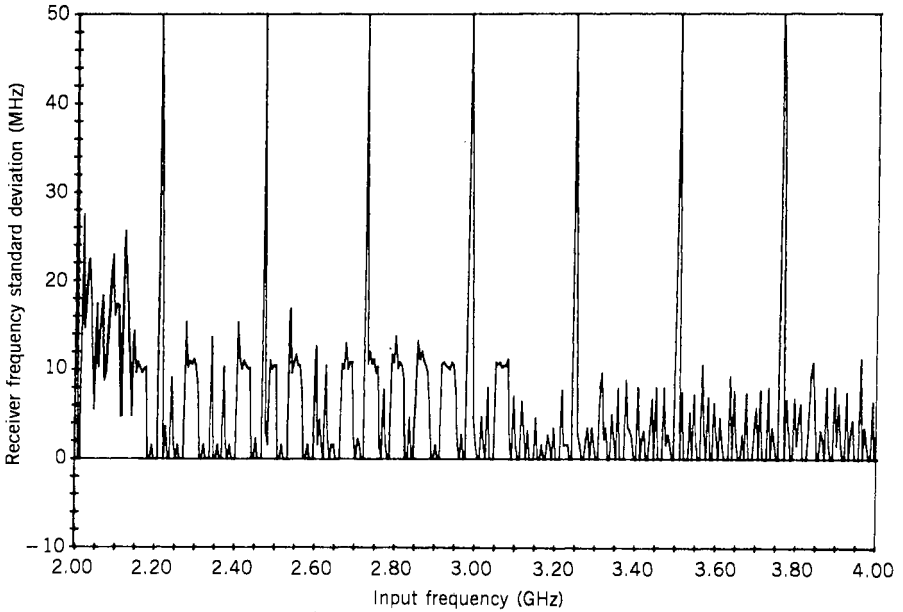


Figure 12.9. Standard deviation frequency.

versus input frequency. In this experiment, a total of 100 measurements are taken at each frequency step. From this figure, one can see that at some frequencies the frequency measurement is quite consistent, but at others, the frequency error measured is not consistent. The inconsistent measurements occur approximately every 250 MHz. It is anticipated that this receiver will cause many problems if it is ever used in an operating system. This inconsistent performance of a receiver cannot be easily revealed from Figures 12.6 and 12.7.

The standard deviation approach can be applied to other parameter measurements. However, other measured results have shown that receivers have more problems measuring frequency consistently than any other parameter. Therefore, the standard deviation measurements are seldom used on the rest of the parameter measurements.

12.7. SENSITIVITY MEASUREMENTS

There are many different definitions of receiver sensitivity. However, in an operating receiver, there are two ways to measure the sensitivity. The simple approach is to detect the data-ready flag at the output of the receiver. The data-ready flag is also referred to as threshold broken. If the input signal level crosses the threshold, the data-ready flag will be up. The other approach is to check the frequency reading as well as the data-ready flag.

The sensitivity measured here is defined as the minimum input power required for the receiver to have a "good" frequency reading. The good frequency reading is determined from the results of the frequency measurements. If the frequency measurement of the receiver is consistent, the good frequency can be easily defined. If the receiver frequency measurement is not consistent, the good frequency reading must be defined from a percentage point of view. That is, at the same input condition, many measurements will be taken; if a certain percentage of frequency is within a certain range, the frequency can be considered good. Although this definition of sensitivity is not mentioned in Chapter 2, it is one of the most useful for system engineers because the receiver will be used to collect data from input signals above this sensitivity level since the data obtained will be dependable.

The sensitivity of the receiver is shown in Figure 12.10. In this measurement, the frequency is qualified as good if it is within a 10-MHz frequency tolerance range. If the measured frequency is 90% good, the input power level is marked as the sensitivity of the receiver. However, the actual measurement to obtain the results is oversimplified. At a certain power level, 10 input signals are fed to the receiver under test. If 9 of the 10 pulses have good frequency readings, the input power will be considered the sensitivity level.

In some receivers, the sensitivity of the receiver depends on the PW of the input signal. Under this condition, the PW must be specified. Usually, when the PW is above a certain value, the sensitivity is independent of PW.

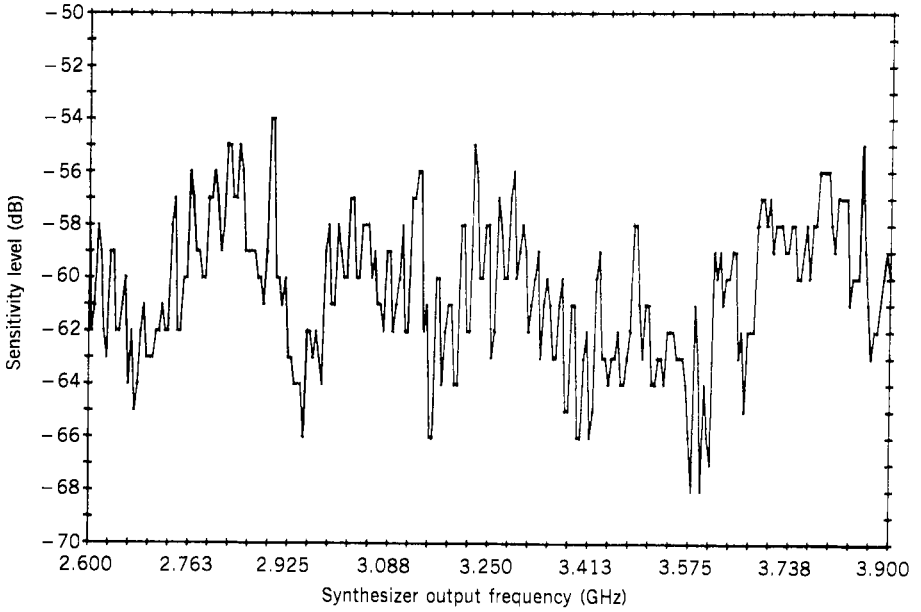


Figure 12.10. Sensitivity versus frequency.

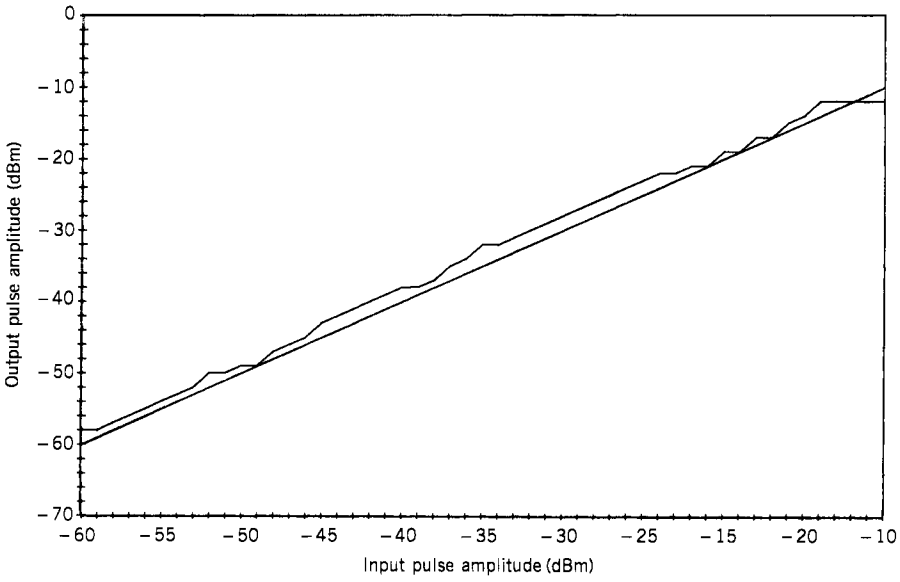


Figure 12.11. Output PA versus input PA.

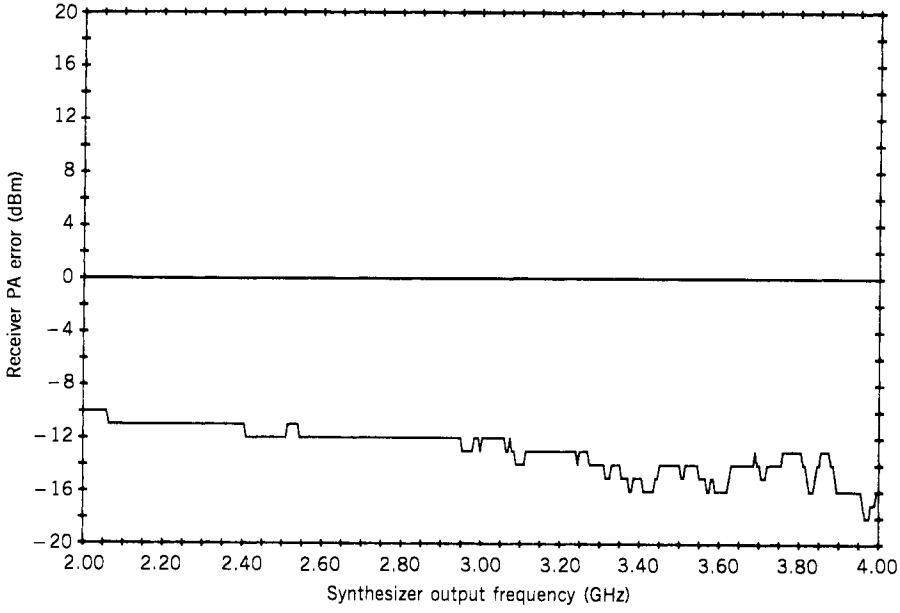


Figure 12.12. PA versus Frequency.

12.8. EVALUATION OF PA MEASUREMENTS

Basically, the amplitude measurement circuit in a receiver depends on two factors. The measurement is a function of input amplitude and frequency. Therefore, the usual way to evaluate the PA measurement circuit is to choose an input frequency and increase the input signal power while the PA output is monitored. Input versus output is shown in Figure 12.11. The slope of this curve is very important. If the slope is correct and the absolute value of the PA is incorrect, the error can be easily corrected by changing a constant in the PA conversion procedure. At the high input level, the PA measurement circuit usually saturates. This curve shows the linear dynamic range of the PA measurement circuit. The straight line represents the ideal input PA versus output PA.

Another approach is to keep the input power constant and monitor the PA output while changing the frequency across the band. At each frequency, the PA is plotted, as shown in Figure 12.12. This measurement shows the PA measured at different frequencies but at constant input power level. The straight line crossing 0 dBm is the reference input level.

12.9. EVALUATION OF PW MEASUREMENTS

In general, the PW measurement circuit in a receiver is not sensitive to frequency. Therefore, the output PW versus input PW will be plotted. The PW is often increased in a nonlinear piecewise fashion. For example, from 100 nsec to 1 μsec, the stepping increment is 100 nsec, and from 1 to 10 μsec, the stepping increment is 1 μsec, and so on. Figure 12.13 shows the results of a PW test.

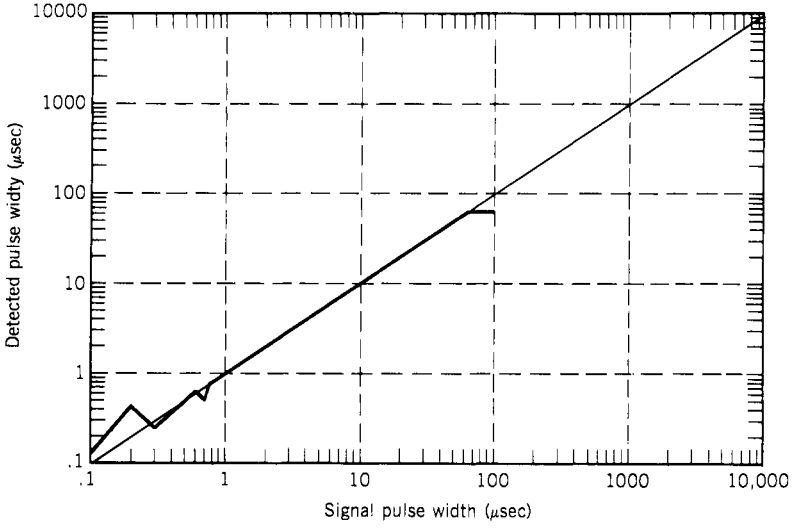


Figure 12.13. PW accuracy measurement.

The straight line represents the ideal PW measurement capability. At a short PW, the actual measurement deviates from the ideal case because of quantization error. At the top of the curve, the maximum PW is reached, and the output PW will not increase anymore.

12.10. SINGLE-SIGNAL DYNAMIC RANGE

Measurement of dynamic range determines the receivers ability to detect a single pulse consistently. In other words, it reveals whether or not a receiver drops pulses or generates spurs. At a certain frequency, the input power is incremented at 1-dB steps, and the number of pulses detected by the receiver is monitored. If the receiver detects more than one signal, a number representing the number of pulses detected is plotted on the graph at the power level and frequency output. If the receiver misses an input signal, a zero will be plotted. If one signal is detected, the output frequency is then checked against the input frequency. If the frequency is outside the tolerance, an × is plotted on the graph. If the receiver performs satisfactorily, no mark will be plotted. Figure 12.14 shows the results of such a measurement. In this particular plot, since there are no X marks, the frequency measurement is always correct. The receiver generated many extra signals around the sensitivity level of the receiver. This phenomenon is caused by multiple triggering and is discussed in Section 3.5.

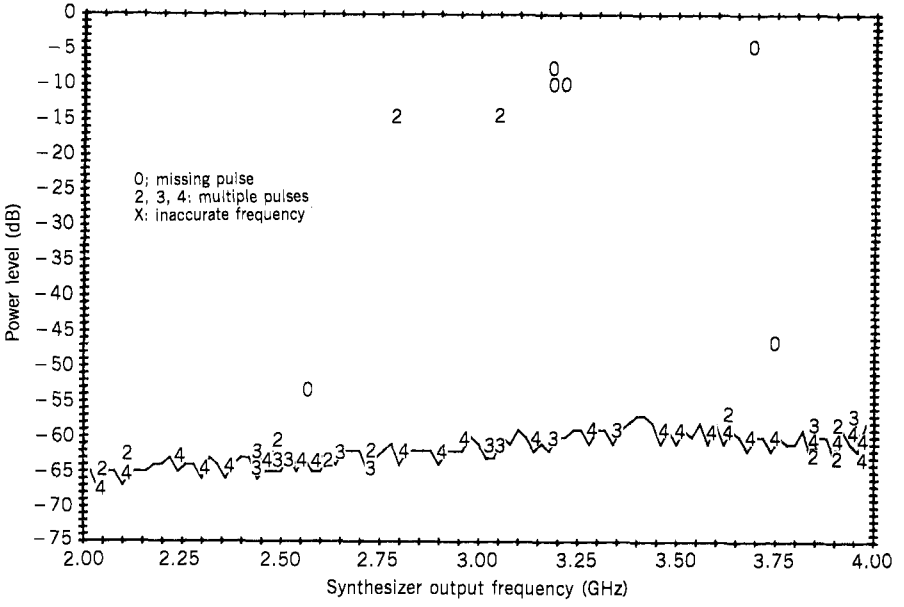


Figure 12.14. Single-pulse dynamic range.

12.11. TWO-SIGNAL SPUR-FREE DYNAMIC RANGE

In this test, two signal generators are used. The two input frequencies are separated by a constant value. When one frequency increases a certain amount, the other also increases by the same amount. Thus the difference frequency is kept constant. This difference frequency is selected according to the ability of the receiver to separate two signals. For example, if the receiver can read two signals with a minimum separation of 25 MHz, the difference frequency used in this test must be greater than 25 MHz.

During this test, both signal powers are increased or decreased by the same amount at the same time. In other words, the two signals must be kept at the same power level. The signals are also kept at the same PW and are time coincident. Both signals are set slightly below the receiver sensitivity level. Their power levels are incremented in 1-dB steps while the frequency output is monitored. If both signals are detected by the receiver, the lower limit of the dynamic range is marked. It should be pointed out that the two signals are checked for frequency accuracy to determine the lower limit of the dynamic range. When the signals are above this sensitivity limit, the input power is still increased in 1-dB steps. At each power level, the following checks are made and labeled accordingly:

0 The receiver misses both input signals

1 The receiver misses one input signal

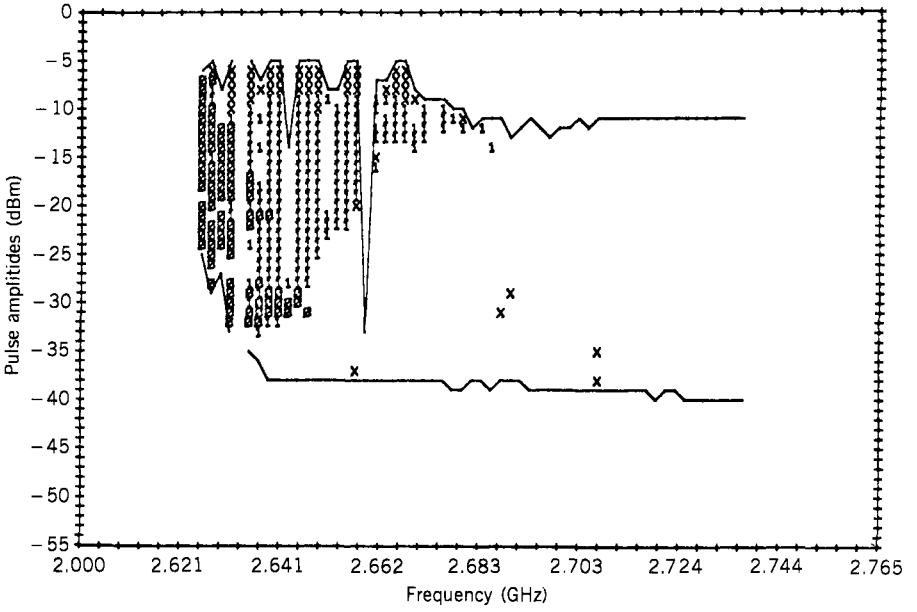


Figure 12.15. Two-signal spur-free dynamic range.

× The receiver receives both signals, but one or both of them are outside the frequency tolerance

The upper limit of the dynamic range is reached when more than two signals are reported by the receiver. Figure 12.15 shows an example of such a test. The two horizontal lines represent the upper and lower limits of the spur-free dynamic range of the receiver. It is important to note that the frequency plotted in this figure is the average values of the two input frequencies.

If the receiver performs properly, it will encode both input signals correctly until a spurious response or the third-order intermodulation is encountered. But for an improperly designed receiver, other conditions may occur, as shown in Figure 12.15.

12.12. INSTANTANEOUS DYNAMIC RANGE

The instantaneous dynamic range test evaluates the capability of a receiver to detect two simultaneous signals of different amplitudes and frequencies. Two pulsed signals are used in this test. The two pulsed signals are time coincident and of the same PW.

The first input signal is set at the center of the frequency band of the receiver (say f_1) at a power level close to the top limit of the receiver's dynamic range.

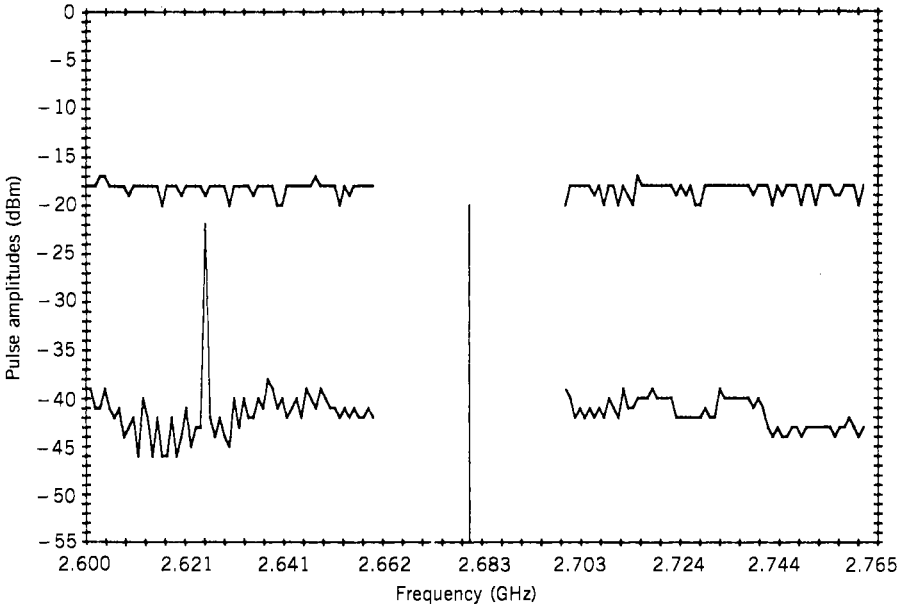


Figure 12.16. Two-signal instantaneous dynamic range.

The second signal is set at frequency f_2, and the power level is set slightly below the sensitivity of the receiver. While the first signal is held constant, the power level of f_2 is increased in 1-dB steps. When the receiver detects both signals correctly, the lower limit of the dynamic range is marked. Now keeping f_2 at a constant level and f_1 increasing 1 dB, if the receiver still encodes both signals correctly, f_1 is increased another dB until the receiver misses the weak signal f_2. The upper limit of the dynamic range is then marked at the highest f_1 power level at which both signals are read correctly.

Figure 12.16 shows measured results of such a test. It should be noted that the frequency plotted on the lower limit of the dynamic range represents the power level and the frequency of f_2. However, the frequency plotted on the upper limit represents the amplitude of the first signal f_1. The frequency of f_1 remains constant at the center of the receiver frequency band.

12.13. FIELD TESTS

In field tests, the receiver is set up in the field to test against some radar. In this test, in addition to the other parameters, the AOA information can be obtained. In a field test, the input signal cannot be controlled easily, and thus the receiver under test must be above a certain standard. Otherwise, with unknown input signals and a receiver of poor performance, the results will be difficult to evaluate.

In a field test, unless a very sophisticated signal processor is available, the test results must be stored temporarily and analyzed at a later time. Some typical

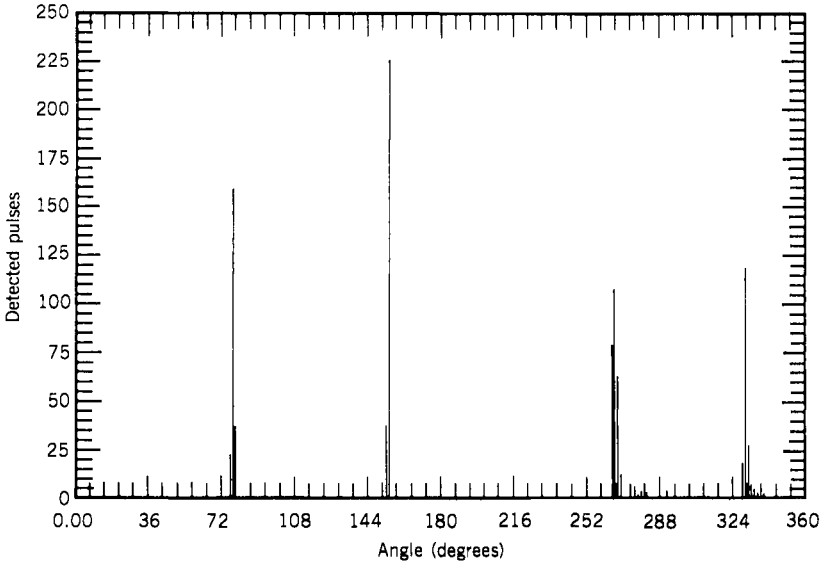


Figure 12.17. AOA measurement.

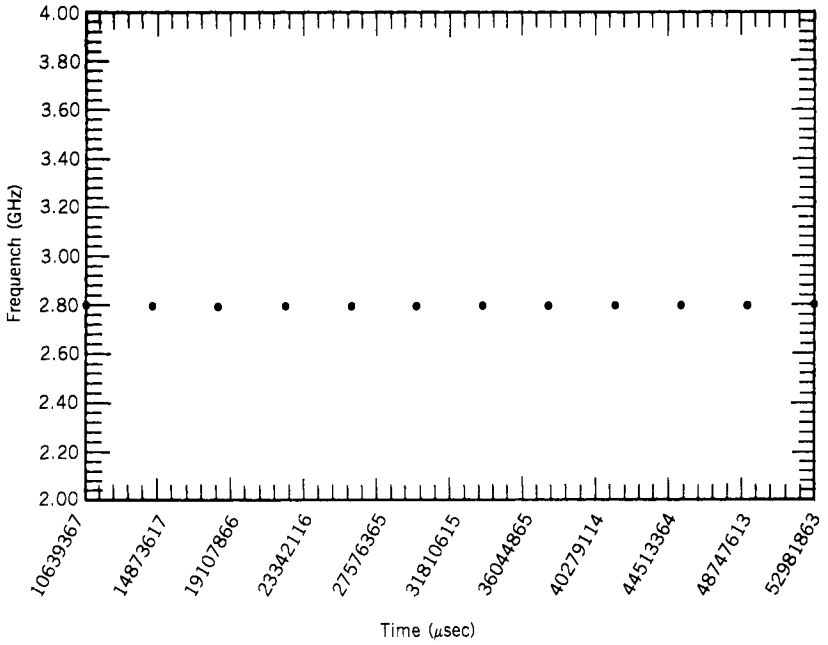


Figure 12.18. Frequency versus coarse TOA.

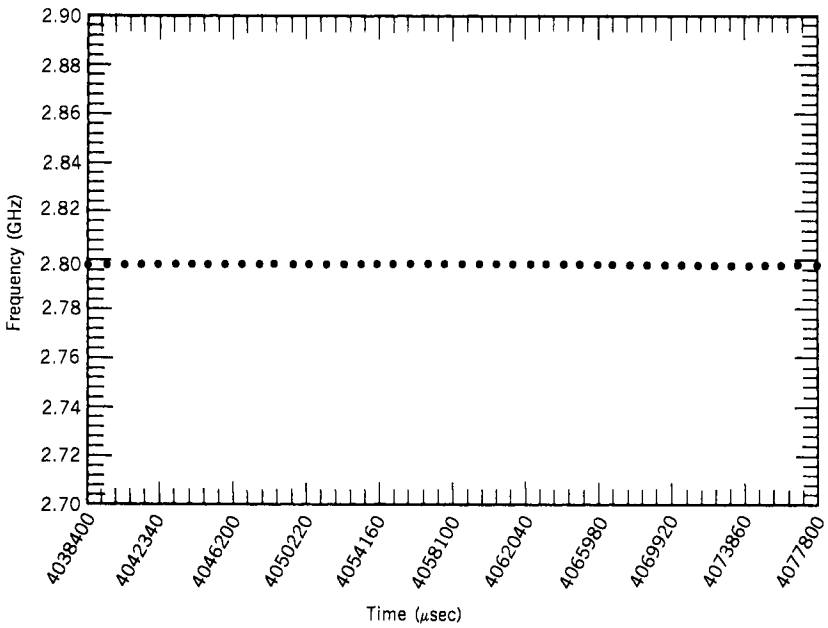


Figure 12.19. Frequency versus fine TOA.

test results will be presented here. Figure 12.17 shows the results of the AOA measurement. In this measurement, a total of 1024 pulses are collected and stored. The total number of pulses versus AOA is plotted. Figure 12.18 shows the results of frequency versus TOA. The time resolution in this figure is too coarse to show the fine detail. The same data are plotted again in Figure 12.19. In this figure, a much finer time resolution is used. One data point between time 40279114 and 44513364 μs shown in Figure 12.18 is plotted in Figure 12.19. The results in Figure 12.18 represent the scan rate of the radar and the results in Figure 12.19 represent the PRI of the radar.

12.14. SUMMARY

The receiver test procedures are used to determine the performance of an EW receiver. Although the detailed performance of an EW receiver is difficult to measure, because of the wide frequency bandwidth and high dynamic range the receiver covers, the measurement procedures discussed in this chapter will provide an overall performance evaluation of an EW receiver.

In general, if a deficiency is discovered in a certain receiver, a detailed evaluation should be performed so that one can understand the causes of the deficiency and generate solutions to correct it.

After all, a receiver is like a human being. There is no perfect person, and there is also no perfect receiver. One can always choose certain input signal conditions to make a receiver produce erroneous data or miss signals, even if the receiver is only a paper design.

REFERENCES

1. C. E. Foster III and K. J. Wong, Automated microwave receiving testing, *Military Electronics/Countermeasures*, 28, June 1978.
2. R. L. Shaw and J. B. Y. Tsui, IFM receiver test and evaluation, Avionics Lab, Technical Report AFWAL-TR-79-1049, Air Force Wright Aeronautical Laboratories, April 1979.
3. D. S. Cooley and R. M. Hume, An automated receiver test concept, *IEEE MTT-S International Microwave Symposium*, pp. 140–142, June 21–23, 1977.
4. K. J. Allen and J. B. Y. Tsui, Automated testing speeds EW receiver evaluation, Microwaves and RF, 24 p. 113 March 1985.

Appendix

Physical Constants

Constant	Symbol	Value	Unit
Permittivity of free space	ε_0	8.854×10^{-12}	Farad/meter
Permeability of free space	μ_0	1.257×10^{-6}	Henry/meter
Speed of light in vacuum	c	3×10^{8}	Meters/second
Free-space wave impedance	$\sqrt{\mu_0/\varepsilon_0}$	376.7	Ohms
Boltzmann constant	k	1.38054×10^{23}	Joule/°K
Electron charge	e	-1.6021×10^{-19}	Coulomb
Planck's constant	h	6.625×10^{-34}	Joule-sec

$$\log a = \frac{1}{2.3} \ln a$$

INDEX

C

D

E

S

T