DECLASSIFIED 2020

RESTRICTED

NRL REPORT R-3231

FR-3231

COMPARISON OF AM AND NARROW-BAND FM UHF COMMUNICATION SYSTEMS

APPENDIX B

THEORETICAL CONSIDERATIONS

DECLASSIFIED by NRL Continuous Declassification Team
Date: 22 DEC 2014
Reviewer's name(s): [redacted]
Declassification authority: NAVY DECLASS GUIDE/NAVY DECLASS MANUAL 11 DEC 2012 D6 SERIES

DECLASSIFIED: By authority of:
5000A January 1958
Entered by: E. Bliss Code 2027

DISTRIBUTION STATEMENT A APPLIES
Further distribution authorized by UNLIMITED only.

NAVAL RESEARCH LABORATORY

WASHINGTON, D.C.

UNCLASSIFIED

RESTRICTED

DECLASSIFIED



DECLASSIFIED

RESTRICTED

NRL REPORT R-3231

COMPARISON OF AM AND NARROW-BAND FM UHF COMMUNICATION SYSTEMS

APPENDIX B
THEORETICAL CONSIDERATIONS

Robert M. Maiden

March 17, 1950

Approved by:

Mr. T. McL. Davis, Head, Radio Techniques Section
Mr. L. A. Gebhard, Superintendent, Radio Division II

NAVAL RESEARCH LABORATORY

CAPTAIN F. R. FURTH, USN, DIRECTOR

WASHINGTON, D.C.

DECLASSIFIED

RESTRICTED

DECLASSIFIED

[redacted]

DISTRIBUTION

BuShips	
Attn: Code 910B	10
Attn: Code 925	10
CNO	
Attn: Code Op-413-B2	5
BuAer	
Attn: Code EL-4	1
Attn: Code EL-92	1
Attn: Code TD-4	1
ONR	
Attn: Code 470	2
CO, ONR, Boston	1
SNLO, USNELO, Fort Monmouth	1
Dir., USNEL	2
CDR, USNOTS	
Attn: Reports Unit	2
ComOpDevFor	1
Supt., USNPGS	1
BAGR, CD, Wright-Patterson AFB	
Attn: CADO-D1	1
OCSigO	
Attn: Ch. Eng. & Tech. Div., SIGTM-S	1
CO, SCEL	
Attn: Dir. of Eng.	2
CG, USAF	1
CG, AMC, Wright-Patterson AFB	1
Attn: Eng. Div., Electronics Subdiv., MCREEO-2	1
CO, AMC, Watson Labs., Red Bank	
Attn: Ch. Eng. Div., WLENG	1
Attn: ENR	1
CO, Air Force Cambridge Res. Labs.	
Attn: ERRS	1
Attn: ERCTL	1
Dir., NBS	
Attn: CRPL	1
RDB	
Attn: Library	2
Attn: Navy Secretary	1
Naval Res. Sec., Science Div., Library of Congress	
Attn: Mr. J. H. Heald	2

DECLASSIFIED

[redacted]

RESTRICTED

DECLASSIFIED

CONTENTS

Abstract	vi
Problem Status	vi
Authorization	vi
INTRODUCTION	1
TRANSMISSION OF INFORMATION	1
MODULATION OF RADIO WAVES	2
IDEAL DOUBLE-SIDEBAND A-M WAVES	2
Phase Relations	4
Modulation Waveforms	4
IDEAL F-M WAVES	6
Side-Frequency Components	7
Sideband Phase Relations	9
Multitone FM	10
Bandwidth Considerations	11
PROPAGATION	12
A-M AND F-M RECEIVERS	18
A-M Receiver Fundamentals	18
F-M Receiver Fundamentals	21
MODULATION DISTORTION	22
General Considerations	22
Comparison Between F-M and A-M System Distortion	24
Tuned Circuit Distortion	25
Distortion In Modulators and Demodulators	28
DISTORTION ARISING FROM MULTIPATH TRANSMISSION	29
CHANNEL SPACING CONSIDERATIONS IN COMMUNICATION SYSTEMS	31

RESTRICTED

DECLASSIFIED

DECLASSIFIED

RESTRICTED

INTERFERENCE IN COMMUNICATION SYSTEMS 33

FLUCTUATION NOISE IN A-M RECEIVERS 35

FLUCTUATION NOISE IN F-M RECEIVERS 38

Ideal F-M System (No Amplitude Limiting) 40
Ideal F-M System (Perfect Amplitude Limiting) 42
Theoretical Output Signal-to-Noise Ratios (FM and AM) 43
Detuning Effects 45

PRE-EMPHASIS AND DE-EMPHASIS 47

IMPULSE NOISE INTERFERENCE 51

IMPULSE NOISE IN A-M RECEIVERS 53

Noise-Peak Limiter 53

IMPULSE NOISE IN F-M RECEIVERS 54

COMMON-CHANNEL INTERFERENCE IN M AND F-M SYSTEMS 56

Linear and Square-Law Detectors w . Two-Signal Input 57
Beat-Note Interference (AM) 58
F-M System (Two-Signal Input) 59
Beat-Note Interference (FM) and Comparison of FM and AM 61
Capture Effects 62

OFF-CHANNEL INTERFERENCE 64

SUMMARY COMPARISON OF A-M AND F-M CHARACTERISTICS 67

CONCLUSIONS 72

ACKNOWLEDGMENTS 73

RESTRICTED

DECLASSIFIED

DECLASSIFIED

ABSTRACT

The characteristics of a-m and f-m communication systems are discussed and compared from a theoretical viewpoint. Topics treated include capacity for transmission of information, spectrum of a-m and f-m waves, propagation, modulation distortion, multipath distortion, channel-spacing considerations, signal-to-noise ratios, effect of pre-emphasis and de-emphasis, performance in the presence of impulse noise, common-channel interference, off-channel interference, and cross modulation. A summary theoretical comparison of the two types of systems is presented in tabular form. The discussions apply directly to naval voice communications systems operating in the VHF/UHF bands, but are presented in sufficiently general form for application to other types of systems.

PROBLEM STATUS

This report completes this phase of the problem; work on another phase continues.

AUTHORIZATION

NRL Problem R01-12
NR 501-120

DECLASSIFIED

[redacted]

DECLASSIFIED

COMPARISON OF AM AND NARROW-BAND FM UHF COMMUNICATION SYSTEMS
APPENDIX B - THEORETICAL CONSIDERATIONS

INTRODUCTION

Modern radio communication systems generally achieve the transmission of information by modulation of a carrier wave, transmission of this modulated wave through the medium of propagation to the receiving point, and demodulation of the wave at the receiver to reproduce the information. A carrier wave contains but two basic quantities which may be modulated; its amplitude and its phase angle. All the various types of modulation are particular forms of amplitude modulation or phase-angle modulation. Of the many types of modulation in current use, amplitude-modulation (AM) and frequency-modulation (FM) are the two types which have been used most widely. Frequency modulation can be considered a special case of phase-angle modulation. The primary purpose of this report is to compare, on a theoretical basis, the advantages and disadvantages of these two forms of modulation for multichannel voice communication systems intended for naval use.

There are many factors which must be considered in a comparison of this sort. Several of them, such as size and weight of equipment required, complexity of circuitry, and cost, do not readily lend themselves to theoretical treatment. Other factors, such as bandwidth requirements, performance in the presence of interference, limitations imposed by propagation characteristics, and inherent distortion, can be more readily treated in a theoretical manner. While particular emphasis has been placed on comparison of AM and FM for use in multichannel naval voice communication systems operating in the 200-400 Mc frequency range, the material is presented in a sufficiently general form to be useful for determining the characteristics of other systems.

A brief discussion of the basic characteristics of a-m and f-m systems will be given, followed by consideration of the several factors which are important in comparison of modulation types particularly for naval UHF/VHF voice communication systems. A summary comparison presented in tabular form will be found at the end of the report. Frequent references to technical papers are included throughout, to indicate where additional information may be found. Particular emphasis has been placed on the subject of fluctuation-noise phenomena as they affect the systems, since it is this factor which usually sets a limit to the useful range of communication systems at the highest radio frequencies.

TRANSMISSION OF INFORMATION

The quantity of information which may be transmitted over a radio communication link is proportional to the product of the bandwidth of the system and the time during which

[redacted]

DECLASSIFIED

DECLASSIFIED

the system is available for transmitting information.[1,2] Systems employing different types of modulation are therefore equally effective in accomplishing the transmission of information, provided this product is the same for all systems considered. The foregoing statement is true only when no interfering signals are present. When interference is present, its power in relation to the signal power becomes a factor in determining the information-handling capacity of the system.[3,4,5] Thus f-m and a-m systems may be made equally effective in their capacity to transmit information in the absence of interference. When interference of any sort is present, however, the information being transmitted becomes altered in both systems. Comparison of f-m and a-m systems from the point of view of interference is presented in a later section of the report.

MODULATION OF RADIO WAVES

An arbitrary radio wave may be represented by the equation:

$$e = A \cos \theta, \tag{1}$$

where A is the instantaneous amplitude and θ is the instantaneous phase angle, each of which may be single-valued functions of time. If θ is a linear function of time and A is a constant, the equation represents an unmodulated carrier wave. It is to be noted that there are but two quantities which may be changed so as to modulate the carrier wave in accordance with an intelligence waveform. One method of modulation is to relate the instantaneous amplitude A to the modulating or intelligence waveform (am); another is to relate the instantaneous phase angle θ to the intelligence waveform (phase-angle modulation). It will be shown later that what is commonly known as fm can be considered a special case of phase-angle modulation, in which the instantaneous phase angle is related linearly to the time integral of the intelligence waveform.

IDEAL DOUBLE-SIDEBAND A-M WAVES

In amplitude modulation, the envelope amplitude of the radio-frequency carrier wave is made to change in accordance with the time variations of the intelligence signal to be transmitted. Figure 1 shows this relationship graphically for an arbitrary intelligence waveform. The unmodulated carrier may be represented mathematically by:

$$e_c = A \cos (\omega_c t + \phi), \tag{2}$$

where A is its instantaneous envelope amplitude, ω_c is equal to 2π times its frequency in cycles per second, ϕ is an arbitrary phase angle, and t represents time.

---

[1] Hartley, R. V. L., "Transmission of Information," B.S.T.J. Vol. 7, pp. 535-563, July 1928

[2] Gabor, D., "Theory of Communication," Jour. I.E.E. (London) Pt. III, Vol. 93, pp. 429-457, Nov. 1946

[3] "Bandwidth vs. Noise in Communication Systems," Electronics, Vol. 21, pp. 72-76, Jan. 1948

[4] "Theoretical Requirements for Communication Systems," Electronics, Vol. 21, p. 138, March 1948

[5] Tuller, W. G., "Theoretical Limitations on the Rate of Transmission of Information," Proc. I.R.E., Vol. 37, pp. 468-478, May 1949

DECLASSIFIED

DECLASSIFIED

A general intelligence waveform may be represented by

$$e_g = g(t), \qquad (3)$$

where g(t) is a single-valued function of time. The mathematical expression for the modulated waveform is[6,7,8]

$$e = A_o\,[1 + m\,g_n(t)]\,\cos\,(\omega_c t + \phi), \qquad (4)$$

where m is a constant which is called the modulation index, or modulation factor, and $g_n(t)$ is the normalized intelligence function, which varies proportionally with the intelligence waveform, but whose absolute value never exceeds unity. This restriction is placed on the modulation function to exclude the case of overmodulation. A value of m = 1 corresponds to full or 100 percent modulation. During intervals when the envelope of the modulated waveform, shown as a dashed line in Figure 1, exceeds the amplitude of the unmodulated carrier (A_o), the modulation is termed "upward"; and when the envelope amplitude is less than the unmodulated carrier amplitude, the modulation is termed "downward."

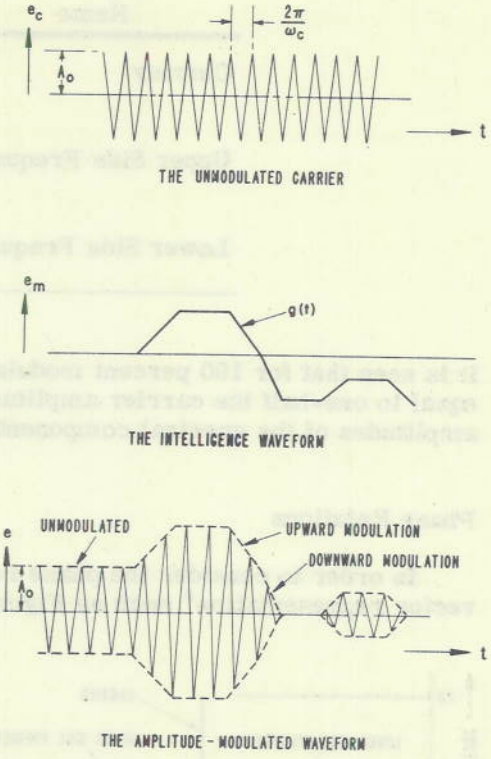


Figure 1

For the special case when $g_n(t)$ is a simple sinusoid, equation (4) may be expanded to show the presence of side frequencies. Expanding equation (4) with $g_n(t) = \cos pt$, where $p = 2\pi$ times the frequency of the modulating tone, one obtains:

$$e = A_o\,[1 + m\cos pt]\cos\,(\omega_c t + \phi) = A_o\cos\,(\omega_c t + \phi) + m\,A_o\,[\cos pt][\cos\,(\omega_c t + \phi)] \qquad (5)$$

$$= \underbrace{A_o\cos\,(\omega_c t + \phi)}_{\text{carrier}} + \underbrace{\frac{m\,A_o}{2}\cos\,[(\omega_c + p)t + \phi]}_{\text{upper side frequency}} + \underbrace{\frac{m\,A_o}{2}\cos\,[(\omega_c - p)t + \phi]}_{\text{lower side frequency}}\,. \qquad (5a)$$

The modulated wave is seen to be the sum of three waves with characteristics as follows:

---

[6] Heising, R. A., "Modulation in Radio Telephony," Proc. I.R.E., Vol. 9, pp. 305-352, August 1921

[7] Carson, J. R., "Notes on the Theory of Modulation," Proc. I.R.E., Vol. 10, pp. 57-64, Feb. 1922

[8] Hartley, R. V. L., "Relations of Carrier and Side-Bands in Radio Transmission," Proc. I.R.E., Vol. 11, pp. 34-56, Feb. 1923

DECLASSIFIED

DECLASSIFIED

RESTRICTED

Name	Amplitude	Frequency
Carrier	A_o	$\frac{\omega_c}{2\pi}$
Upper Side Frequency	$\frac{mA_o}{2}$	$\frac{\omega_c + p}{2\pi}$
Lower Side Frequency	$\frac{mA_o}{2}$	$\frac{\omega_c - p}{2\pi}$

It is seen that for 100 percent modulation (m = 1), the side frequencies have an amplitude equal to one-half the carrier amplitude. Figure 2 shows graphically the frequencies and amplitudes of the spectral components for 100 percent modulation.

Phase Relations

In order to consider the phase relations between the carrier and side frequencies, a vector representation[9] such as Figure 3 can be employed. Although these components have different frequencies, the instantaneous phase angles between their representative vectors may be defined. In Figure 3, the carrier vector is represented by OC, the upper side frequency by CE and the lower side frequency by CD. It is assumed that the observer is "riding" on the carrier vector as it rotates counter-clockwise at an angular velocity ω_c about O. Hence the vector OC appears stationary to the observer. The two vectors corresponding to the upper and lower side frequencies appear to rotate in opposite directions at the angular velocity of the modulating wave. These vectors have an amplitude equal to one-half the carrier amplitude for the case of 100 percent modulation, as previously mentioned. The resultant of the two side-frequency vectors is always in phase with the carrier vector, but varies in amplitude throughout the cycle of the modulation. Thus the resultant of the three vectors, carrier plus upper and lower side frequencies, is in phase with the carrier vector and lengthens and shortens at the modulating frequency.

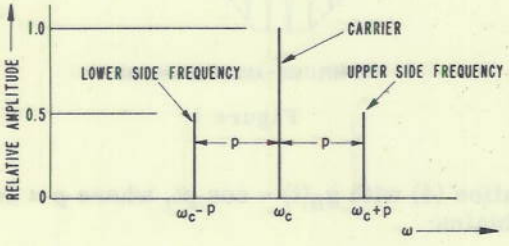


Figure 2 - Carrier and side frequencies in single-tone amplitude modulation (m = 1)

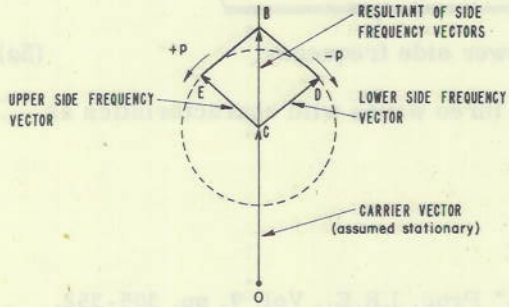


Figure 3 - Vector diagram of amplitude modulation (m = 1)

Modulation Waveforms

Figures 4a and 4b show the waveforms for single-tone am with 100 percent and 50 percent

[9] La Port, E. A., "Characteristics of Amplitude Modulated Waves," RCA Review, Vol. 1, p. 36, April 1937

RESTRICTED

DECLASSIFIED

DECLASSIFIED

modulation respectively. The modulation factor may be derived from the waveform in the case of single-tone modulation by using the following formulas:[10] (See Figure 4.)

$$m = \frac{A_{max} - A_o}{A_o} \quad \text{(Upward Modulation)}$$

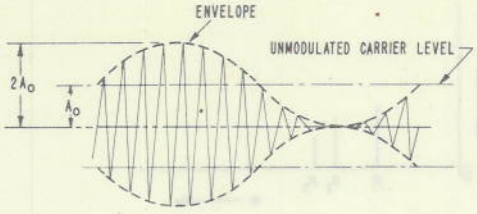
$$m = \frac{A_o - A_{min}}{A_o} \quad \text{(Downward Modulation)} \qquad (6)$$

Inspection of equations (5) and (5a) reveals that:

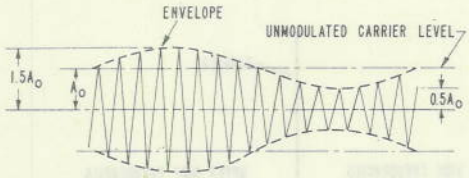
(a) The envelope of the modulated wave corresponds to the form of the modulating signal.

(b) The amplitudes of the side frequencies are equal and are proportional to the modulation factor m, and their frequencies differ from that of the carrier by the modulating frequency.

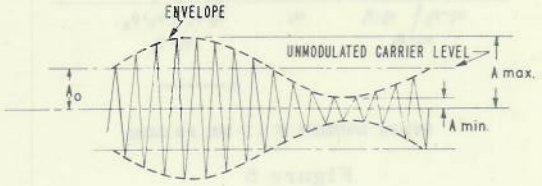
(c) The carrier-frequency energy is not changed by the modulation process, showing that the carrier does not contain intelligence. The process of modulation results in the generation of side-frequency energy which contains the intelligence.



(a) 100% SINGLE-TONE AMPLITUDE MODULATION



(b) 50% SINGLE-TONE AMPLITUDE MODULATION



(c) SINGLE-TONE AMPLITUDE MODULATION

Figure 4

(d) Of the total energy transmitted in a 100 percent tone-modulated wave, two-thirds is in the carrier and one-sixth is in each of the two side frequencies. The average power in a 100 percent modulated wave is one and one-half times the average power of the unmodulated carrier. The instantaneous peak power transmitted during 100 percent modulation is four times that of the unmodulated carrier, and the instantaneous peak voltage of the modulated wave for 100 percent modulation is twice that of the unmodulated carrier.

When the modulating signal consists of more than one tone, upper and lower side frequencies are generated for each tone in the modulating signal. Figure 5 shows the spectral components for a carrier modulated by a wave consisting of three tones. It is seen that the relative amplitudes of the tones are preserved in the side-frequency spectrum.

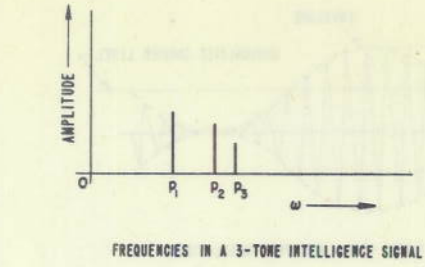
The intelligence conveyed by practical systems does not consist of simple sine waves of infinite duration, but rather of complex waveforms such as those which might be generated by the human voice speaking into a microphone. Any complex waveform, however, can be regarded as the resultant of the superposition of many sinusoids of proper amplitudes and phases, which cancel during intervals when the complex wave has zero amplitude

[10] Terman, F. E., "Radio Engineers' Handbook," p. 531, McGraw-Hill Book Company, Inc., New York and London, 1943

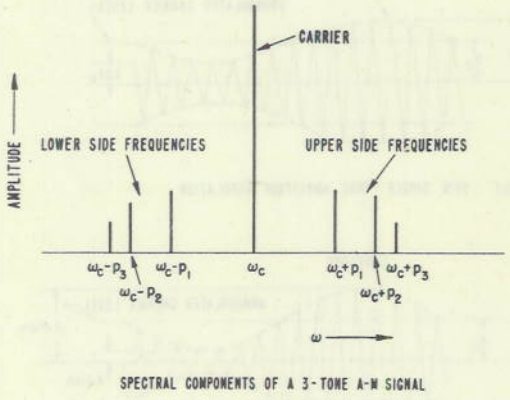
DECLASSIFIED

DECLASSIFIED

RESTRICTED

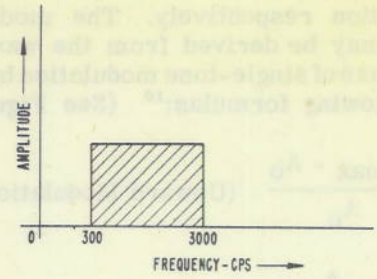


FREQUENCIES IN A 3-TONE INTELLIGENCE SIGNAL

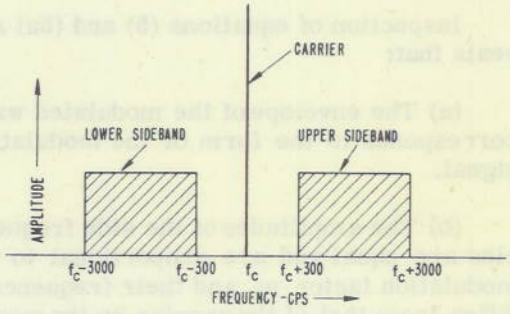


SPECTRAL COMPONENTS OF A 3-TONE A-M SIGNAL

Figure 5



BANDWIDTH FOR INTELLIGIBLE VOICE TRANSMISSION



SPECTRUM OF VOICE-MODULATED WAVE

Figure 6

and which add appropriately during intervals when the complex waveform has finite amplitude. These component sinusoids are harmonically related if the complex wave form is periodic, and form a continuous spectrum if the waveform is nonperiodic.

The human voice is transmitted satisfactorily even if only the components having frequencies between about 300 and 3000 cycles per second are permitted to modulate the carrier. When such a signal modulates a carrier, upper and lower side bands are generated as shown in Figure 6.

It may be seen by referring to Figures 2, 5, and 6 that the radio-frequency spectrum of an ideal double-sideband amplitude-modulated wave is always symmetrical about the carrier frequency even when the modulating waveform is not symmetrical about the time axis. It may be inferred from the foregoing that the radio-frequency bandwidth occupied by such a wave is equal to twice the highest modulating frequency. This is an important factor in designing communication systems, as it sets a minimum to the bandwidth requirement. Allowances for oscillator and circuit drifts and inaccuracies add to this primary requirement and will be discussed later.

IDEAL F-M WAVES

In frequency modulation, the instantaneous frequency of the carrier is made to vary in accordance with the instantaneous amplitude of the intelligence wave or signal to be transmitted, while the amplitude of the carrier wave is held constant. Figure 7 shows this relationship graphically. The f-m wave may be represented mathematically by

RESTRICTED

DECLASSIFIED

DECLASSIFIED

$$e_c = A_o \cos \theta(t), \qquad (7)$$

in which A_o is the constant amplitude of the wave and $\theta(t)$ is a function of time. The instantaneous frequency[11,12] is defined by:

$$f_i = \frac{\omega_i}{2\pi} = \frac{1}{2\pi}\frac{d\theta}{dt}. \qquad (8)$$

As the modulating signal varies in amplitude, the instantaneous frequency of the wave is made to vary about the center frequency, $f_c = \omega_c/2\pi$. The difference between the instantaneous frequency and the unmodulated carrier frequency when the modulating signal reaches its maximum amplitude is termed the maximum deviation and will be represented by $\Delta f_{max} = \Delta\omega_{max}/2\pi$. Thus the instantaneous frequency of a wave which is frequency-modulated by an intelligence signal represented by $g_n(t)$ becomes

$$f_i = \frac{\omega_i}{2\pi} = \frac{1}{2\pi}\frac{d\theta}{dt} = f_c + g_n(t)\,\Delta f_{max}. \qquad (9)$$

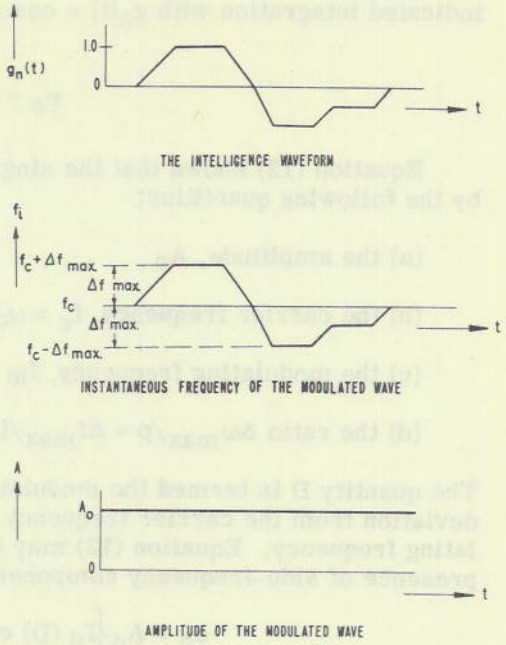


Figure 7

Multiplying equation (9) by 2π and integrating gives the value of θ in equation (7):

$$\theta = \omega_c t + \Delta\omega_{max}\int g_n(t)\,dt. \qquad (10)$$

Substitution of the value of θ from equation (10) into equation (7) yields the general equation of the frequency-modulated wave:

$$e_c = A_o \cos\left[\omega_c t + \Delta\omega_{max}\int g_n(t)\,dt\right]. \qquad (11)$$

Inspection of equations (10) and (11) reveals that the modulation or intelligence is contained only in the phase-angle term of the cosine function and not at all in the amplitude term. Thus frequency modulation can be considered to be a type of phase-angle modulation.

If the phase angle θ is made linearly proportional to the amplitude of the modulating signal, the wave is termed phase-modulated; but when, as in equation (11), the phase angle is linearly proportional to the time integral of the modulating signal, the wave is termed frequency-modulated. In the latter case, the instantaneous frequency of the modulated wave is linearly proportional to the instantaneous amplitude of the intelligence signal.

Side-Frequency Components

For the case when $g_n(t)$ is a simple sinusoid (single-tone frequency modulation), equation (11) may be expanded to show the presence of side frequencies. Performing the

---

[11] Carson, J. R., op. cit.

[12] Roder, H., "Amplitude, Phase and Frequency Modulation," Proc. I.R.E., Vol. 19, pp. 2145-2175, Dec. 1931 (Also see discussion in Proc. I.R.E., p. 884, May 1932.)

DECLASSIFIED

DECLASSIFIED

indicated integration with $g_n(t) = \cos pt$, one obtains:

$$e_c = A_o \cos\left(\omega_c t + \frac{\Delta\omega_{max}}{p} \sin pt\right) . \tag{12}$$

Equation (12) shows that the single-tone frequency-modulated wave is characterized by the following quantities:

(a) the amplitude, A_o

(b) the carrier frequency, $f_c = \omega_c/2\pi$

(c) the modulating frequency, $f_m = p/2\pi$

(d) the ratio $\Delta\omega_{max}/p = \Delta f_{max}/f_m = D$.

The quantity D is termed the modulation index and is equal to the ratio of the maximum deviation from the carrier frequency caused by the modulating tone, divided by the modulating frequency. Equation (12) may be expanded in terms of Bessel functions to show the presence of side-frequency components as follows:[13]

$$\begin{aligned} e_c = A_o \Big\{ & J_o(D) \cos \omega_c t \\ & - J_1(D)\,[\cos(\omega_c - p)t - \cos(\omega_c + p)t] \\ & + J_2(D)\,[\cos(\omega_c - 2p)t + \cos(\omega_c + 2p)t] \\ & - J_3(D)\,[\cos(\omega_c - 3p)t - \cos(\omega_c + 3p)t] \\ & \ldots (-1)^n J_n(D)\,[\cos(\omega_c - np)t + (-1)^n \cos(\omega_c + np)t] \Big\} . \end{aligned} \tag{13}$$

Inspection of equation (13) reveals the presence of the following frequency components:

Name	Amplitude	Frequency
Carrier	$J_o(D)\,A_o$	$\frac{\omega_c}{2\pi}$
First-Order Side Frequencies	$J_1(D)\,A_o$	$\frac{\omega_c + p}{2\pi}$ and $\frac{\omega_c - p}{2\pi}$
Second-Order Side Frequencies	$J_2(D)\,A_o$	$\frac{\omega_c + 2p}{2\pi}$ and $\frac{\omega_c - 2p}{2\pi}$
Nth-Order Side Frequencies	$J_n(D)\,A_o$	$\frac{\omega_c + np}{2\pi}$ and $\frac{\omega_c - np}{2\pi}$

[13] Hund, A., "Frequency Modulation," p. 350, McGraw-Hill Book Co., Inc., New York and London, 1942

DECLASSIFIED

DECLASSIFIED

It is evident that the single-tone modulated f-m wave consists of an infinite number of side frequencies symmetrically disposed about the carrier and spaced from it at integral multiples of the modulating frequency. Information is contained in each component of the f-m wave, as each term in the expression of equation (13) contains the modulation index D.

The power in the modulated wave is the same as that contained in the unmodulated carrier. The amplitudes of the carrier and side frequencies depend on the peak deviation and modulating frequency, as determined by the Bessel function of the first kind of argument D and order n. The spectrum of an f-m wave having a modulation index of 7 and a modulating frequency of 1000 cps is shown in Figure 8. Spectra for other conditions may be computed using tables of Bessel functions.[14,15]

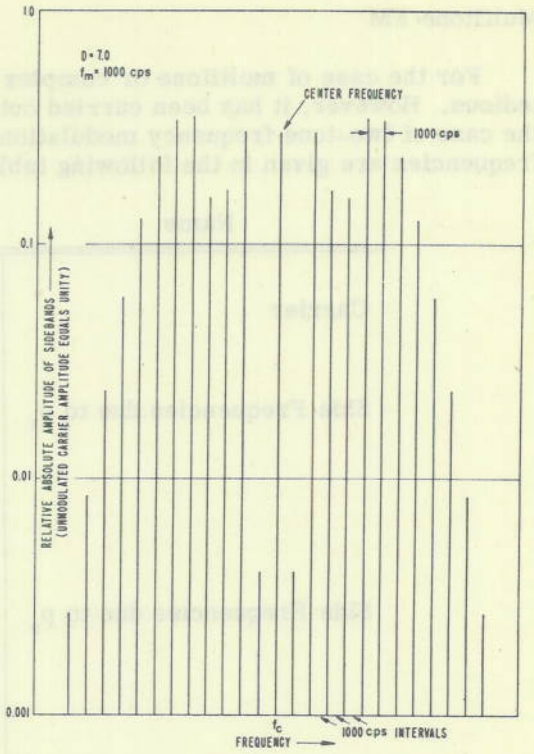


Figure 8 - Spectrum of a frequency-modulated wave

Sideband Phase Relations

The phase relations among the sidebands in an f-m wave may be discussed by expressing equation (12) in another form as follows:[15]

$$e_c = A_o \Big\{ J_o(D) \cos \omega_c t - [2J_1(D) \sin pt] \sin \omega_c t + [2J_2(D) \cos 2pt] \cos \omega_c t - [2J_3(D) \sin 3pt] \sin \omega_c t \ldots\ldots + [2J_{2m}(D) \cos 2mpt] \cos \omega_c t - [2J_{2m+1}(D) \sin (2m+1)pt] \sin \omega_c t \Big\}$$

(when m = 2, 3, 4.....), (14)

It may be seen that the resultants of the odd-order side-frequency pairs are displaced 90° with respect to the carrier, and the resultants of the even-order side-frequency pairs are in phase with the carrier. This is shown graphically in Figure 9. For the case when D is small (< 0.2) only the first sideband pair has significant amplitude and for this case it may then be said that the resultant of the significant side-frequency vectors is in quadrature with the carrier vector. It is on this premise that the Armstrong method of generating frequency-modulated waves is based.[16]

---

[14] Hund, A., op. cit., p. 352

[15] Hund, A., op. cit., p. 50 (equation 13 of this report is analogous to Hund's equation 23).

[16] Armstrong, E. H., "A Method of Reducing Disturbances in Radio Signaling by a System of Frequency Modulation," Proc. I.R.E., Vol. 24, pp. 689-740, May 1936

DECLASSIFIED

DECLASSIFIED

Multitone FM

For the case of multitone or complex frequency modulation, the spectral analysis is tedious. However, it has been carried out and conclusions are as stated below.[17,18,19] For the case of two-tone frequency modulation, the amplitudes of the components and their frequencies are given in the following table:

Name	Amplitude	Frequency
Carrier	$J_0(D_1)\, J_0(D_2)\, A_0$	$\frac{\omega_c}{2\pi}$
Side Frequencies due to p_1	$J_1(D_1)\, J_0(D_2)\, A_0$	$\frac{\omega_c \pm p_1}{2\pi}$
	$J_2(D_1)\, J_0(D_2)\, A_0$	$\frac{\omega_c \pm 2p_1}{2\pi}$
Side Frequencies due to p_2	$J_0(D_1)\, J_1(D_2)\, A_0$	$\frac{\omega_c \pm p_2}{2\pi}$
	$J_0(D_1)\, J_2(D_2)\, A_0$	$\frac{\omega_c \pm 2p_2}{2\pi}$
Combination Frequencies	$J_1(D_1)\, J_1(D_2)\, A_0$	$\frac{\omega_c + p_1 \pm p_2}{2\pi}$
	$J_1(D_1)\, J_1(D_2)\, A_0$	$\frac{\omega_c - p_1 \pm p_2}{2\pi}$
	$\vdots$	$\vdots$
	$J_1(D_1)\, J_2(D_2)\, A_0$	$\frac{\omega_2 + p_1 \pm 2p_2}{2\pi}$
	$J_1(D_1)\, J_2(D_2)\, A_0$	$\frac{\omega_c - p_1 \pm 2p_2}{2\pi}$
	$\vdots$	$\vdots$
	$J_2(D_1)\, J_1(D_2)\, A_0$	$\frac{\omega_c + 2p_1 \pm p_2}{2\pi}$
	$J_2(D_1)\, J_1(D_2)\, A_0$	$\frac{\omega_c - 2p_1 \pm p_2}{2\pi}$
	$\vdots$	$\vdots$

[17] Terman, F. E., op. cit., p. 582

[18] Crosby, M. G., "Carrier and Side Frequency Relations with Multitone Frequency or Phase Modulation," RCA Review, Vol. 3, pp. 103-106, July 1938

[19] Giacoletto, L. J., "Generalized Theory of Multitone Amplitude and Frequency Modulation," Proc. I.R.E., Vol. 35, pp. 680-693, July 1947

DECLASSIFIED

DECLASSIFIED

When the modulating signal contains more than two tones, the spectral analysis becomes exceedingly complex. It has been shown that for the case of multitone FM, the shape of the side-frequency spectrum is intimately associated with the wave shape of the modulating signal. A symmetrical side-frequency power spectrum arises from a modulating waveform which is symmetrical about the time axis, whereas an unsymmetrical side-frequency spectrum is present when the modulating signal is unsymmetrical about its time axis.[19]

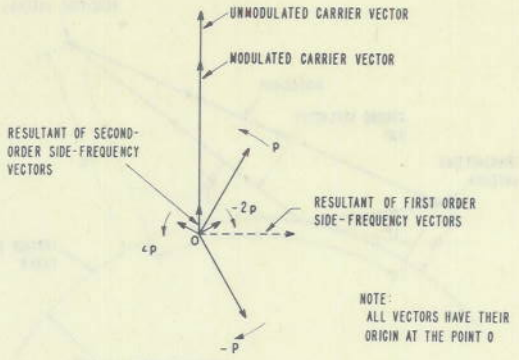


Figure 9 - Vector relations in an f-m wave (D = 1)

Bandwidth Considerations

The bandwidth occupied by an f-m wave having symmetrical single-tone modulation is theoretically infinite, since the spectrum of the wave contains an infinite number of side frequencies spaced symmetrically about the unmodulated carrier frequency at all integral multiples of the modulating frequency. The amplitudes of the side-frequency components decrease regularly beyond the region where the order of the side frequency is greater than the modulation index. The bandwidth occupied by the f-m wave may be considered as that region in which the amplitudes of its components exceed some arbitrary value. Considering only those components having amplitudes greater than 10 percent of the unmodulated carrier amplitude, the bandwidth of the wave may be expressed by the formula:

$$\Delta f_f = 2\Delta f_{max} + 2f_m, \qquad (15)$$

where Δf_{max} is the maximum deviation and f_m is the modulating frequency. If side-frequency components having amplitudes less than 10 percent of the unmodulated carrier amplitude may be neglected, and $\Delta f_{max}/f_m$ (= D) is less than unity, equation (15) may be written

$$\Delta f_f = 2f_m \qquad (15a)$$

as only the first sideband pair has an amplitude greater than 10 percent of the unmodulated carrier amplitude.

A more conservative criterion for the bandwidth occupied by an f-m wave is to consider the frequency boundaries for the group of components whose amplitudes are greater than 1 percent of the unmodulated carrier amplitude. In this case, there is no simple expression for the bandwidth occupied by the wave.[20,21]

Tables of Bessel functions may be used for accurate spectrum occupancy calculations. For instance, an f-m system having a peak frequency deviation of 7000 cycles per second and a maximum modulating frequency of 1000 cycles per second (modulation index equals 7.0) would require a bandwidth of 20,000 cycles per second, assuming that all components

---

[20] Hund, A., op. cit., Chapter I

[21] Corrington, M. S., "Variation of Bandwidth with Modulation Index in Frequency Modulation," Proc. I.R.E., Vol. 35, pp. 1013-1020, Oct. 1947

[redacted]

DECLASSIFIED

DECLASSIFIED RESTRICTED

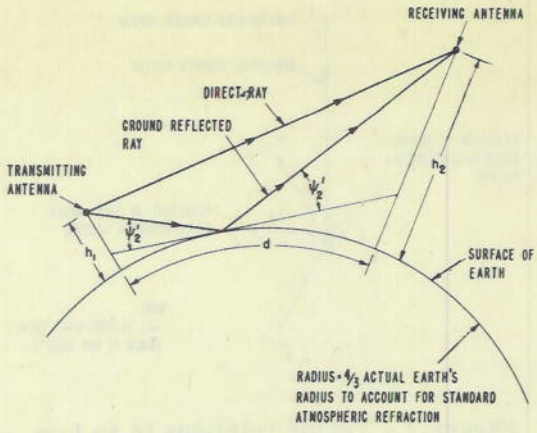


Figure 10 - Geometry of propagation paths

having an amplitude of less than 1 percent of the unmodulated carrier amplitude are neglected. If the maximum modulating frequency were 3000 cycles per second and the peak deviation 7000 cycles per second, a bandwidth of 30,000 cycles would be required.

PROPAGATION

The following consideration of radio wave propagation will be confined to the frequency range from 100 to 400 Mc. Figure 10 shows the geometrical arrangement of the transmitter and receiver in a typical radio link. In some naval installations, antenna heights above the surface of the earth may be less than 30 ft. (e. g., on submarines and PT boats), or in other shipboard installations as high as 180 ft. (CVB), and in fighter aircraft, greater than 40,000 ft. The intervening surface of the earth in Naval systems will generally be sea water. Transmission in this frequency range is usually by direct ray and ground-reflected ray, as shown in Figure 10, since reflections from the ionosphere rarely occur for frequencies above 60 Mc.[22]

Figures 11 through 15 show the attenuation-versus-distance curves for propagation of vertically polarized radio waves at a frequency of 330 Mc over sea water for several

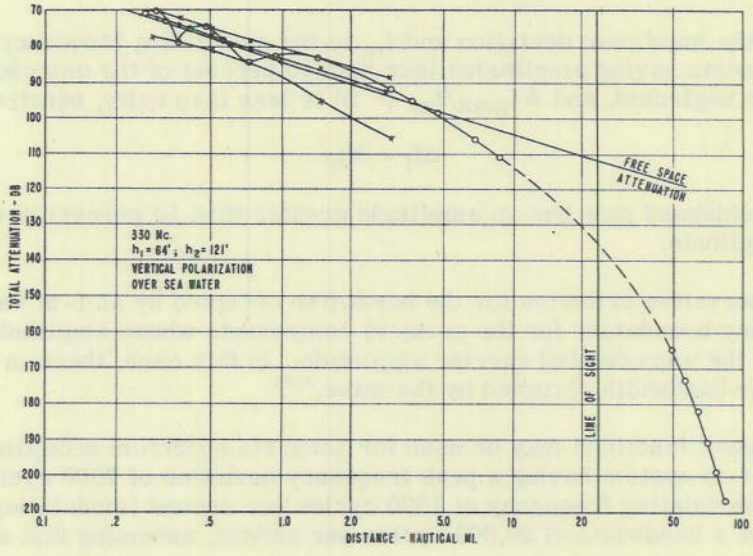


Figure 11

[22] Terman, F. E., op. cit., p. 754

RESTRICTED
DECLASSIFIED

DECLASSIFIED

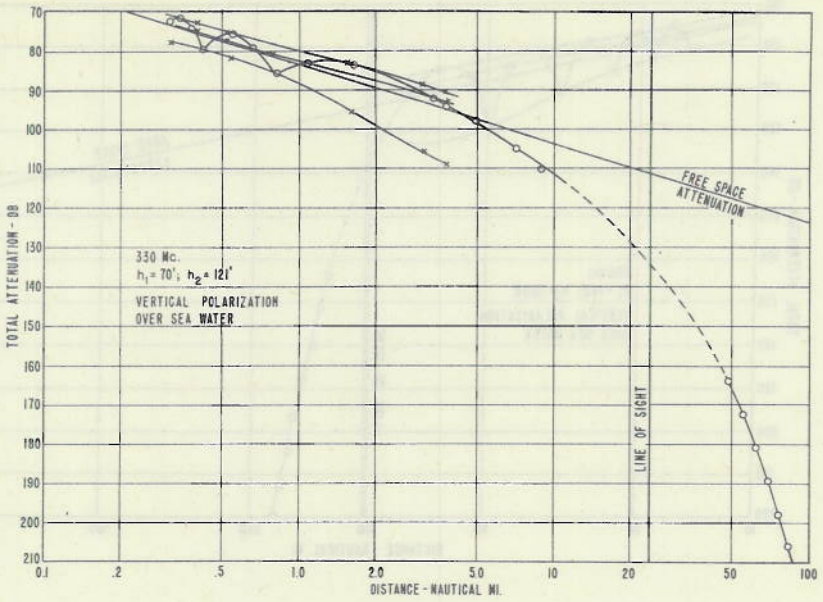


Figure 12

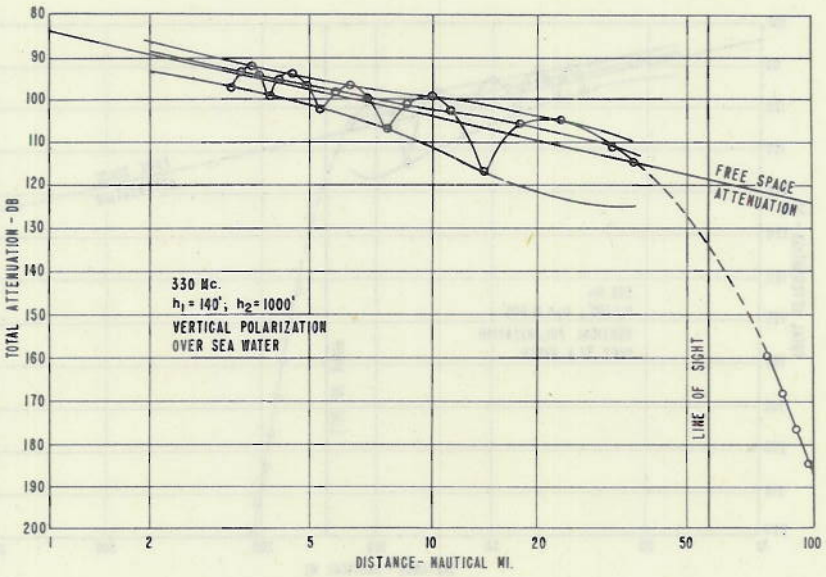


Figure 13

DECLASSIFIED

DECLASSIFIED

RESTRICTED

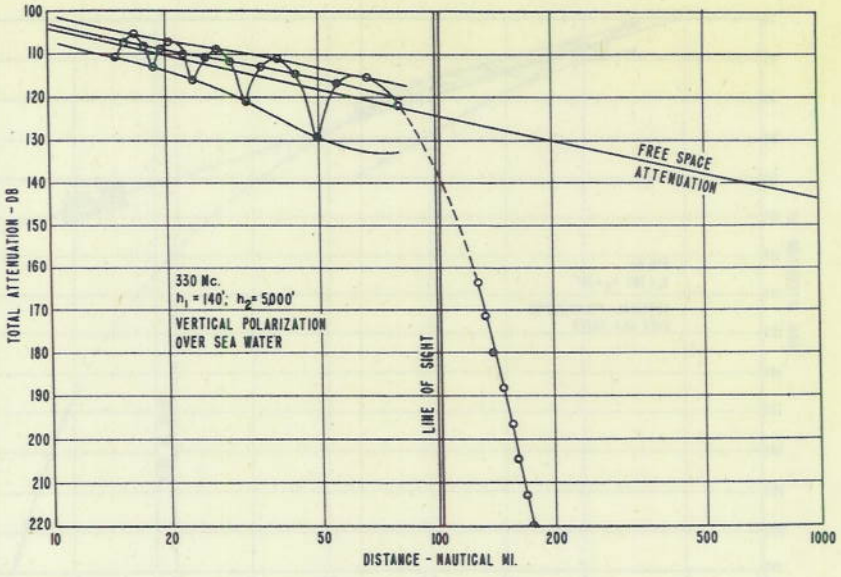


Figure 14

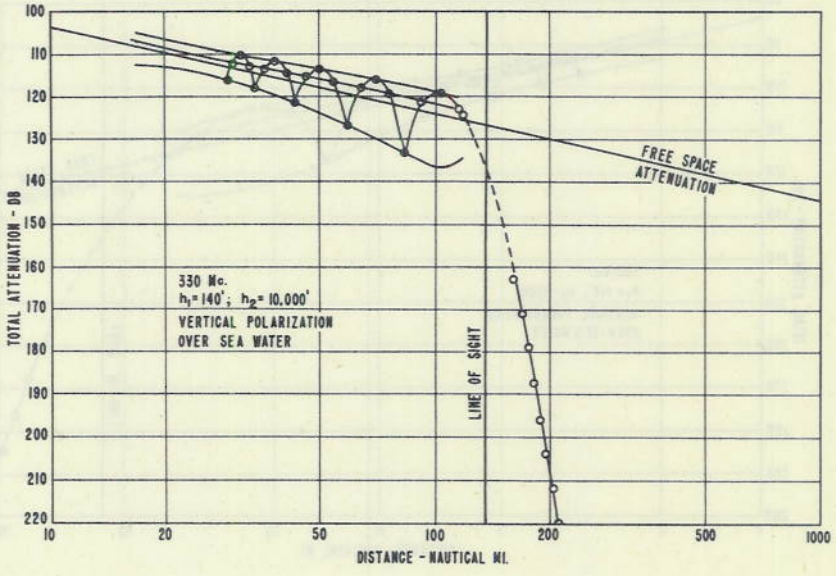


Figure 15

RESTRICTED

DECLASSIFIED

DECLASSIFIED

combinations of antenna heights. The ordinates give the attenuation between two vertical half-wave dipoles, and the abscissas give the distance between the dipoles measured along the surface of the earth. These curves were computed using curved-earth theory, with the earth's radius taken as 4/3 its actual value to allow for the effect of standard atmospheric refraction.[23] These curves apply to the general conditions existing in the f-m vs. a-m comparison tests conducted by the Naval Research Laboratory, under Problem R01-12.[24,25] It is emphasized that the curves described above give the attenuation to be expected only under ideal conditions. They apply to the transmission of an unmodulated carrier between the transmitting and receiving antennas; distortion or changes in effective modulation index which may occur in the case when the signal is modulated were not considered. Ducting and other anomalous effects often result in ranges much in excess of those predicted theoretically.[26,27] Propagation curves for ideal half-wave dipoles with horizontal polarization indicate much greater variations in field strength due to reflections for distances less than line-of-sight than indicated for vertical polarization.

Figures 16, 17, and 18 indicate the maximum range in nautical miles at which a ten-microvolt signal may be obtained at the terminals of a receiving half-wave dipole antenna for a given radiated power. These figures were derived from propagation curves, similar to those of Figures 11 through 15, which are presented in NRL Report R-3202.[28] The range plotted is based on the first intersection (least range) of the attenuation abscissa with the propagation curve for each attenuation shown. The ten-microvolt received signal level is considered a practical average limit for good voice communication in the 100-400 Mc frequency range based on present-day receiver characteristics. The dotted lines in Figures 16, 17, and 18 indicate that the data for the 100-200 Mc region were obtained by extrapolation from the data for 200 Mc and above. These curves show that approximately 20 to 70 watts of radiated power is required in the 100-400 Mc frequency range to provide a 10-microvolt signal at distances near the radio line-of-sight.

It may be concluded from the foregoing that to obtain a given range with a given antenna system, the transmitter power must be increased as the frequency is increased. It should be noted, however, that if the radiated power is <u>not</u> increased as the frequency is increased, the loss in range due to increase in frequency is relatively small in terms of distance, provided that the power used at the initial lower frequency is sufficient to obtain radio line-of-sight range. A radiated power of 70 watts should, on the average, be sufficient to obtain approximately radio line-of-sight coverage in the 100-400 Mc band under ideal conditions.

---

[23] Peake, H. J., "Calculation of the Attenuation of Electric Field Strength," NRL Report R-3202 (Unclassified), Dec. 1947

[24] McClenon, D., "Comparison of AM and Narrow-Band FM UHF Communication Systems - Appendix E - Shipboard Studies and Investigations," NRL Report R-3234 (Restricted), March 1948

[25] Hollweck, F. J., "Comparison of AM and Narrow-Band FM UHF Communication Systems - Appendix F - Airborne Studies and Investigations," NRL Report R-3235 (Restricted), June 1948

[26] Terman, F. E., op. cit., p. 762

[27] Schwartz, L. S., "Factors Affecting the Choice of Operating Frequencies for Line-of-Sight Systems," NRL Report R-3005 (Unclassified), Oct. 1946, pp. 36-44

[28] Peake, H. J., op. cit.

DECLASSIFIED

DECLASSIFIED

 RESTRICTED

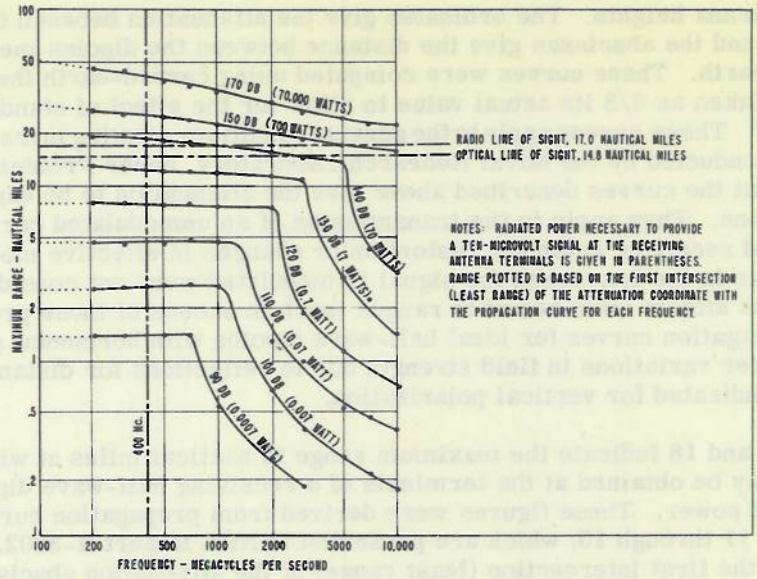


Figure 16 - Maximum range vs. frequency. Attenuation between transmitting and receiving half-wave dipoles as a parameter. Antenna heights, 30-70 ft.; vertical polarization; transmission over sea water.

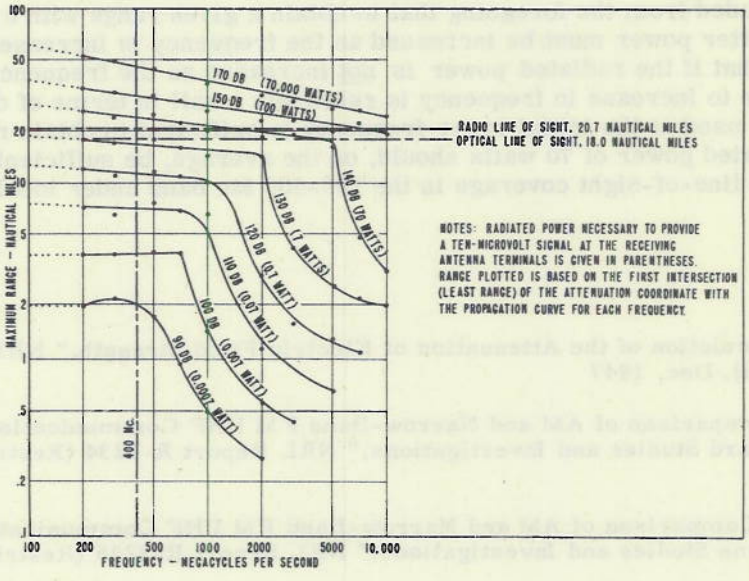


Figure 17 - Maximum range vs. frequency. Attenuation between transmitting and receiving half-wave dipoles as a parameter. Antenna heights, 70-70 ft.; vertical polarization; transmission over sea water.

RESTRICTED

DECLASSIFIED

DECLASSIFIED

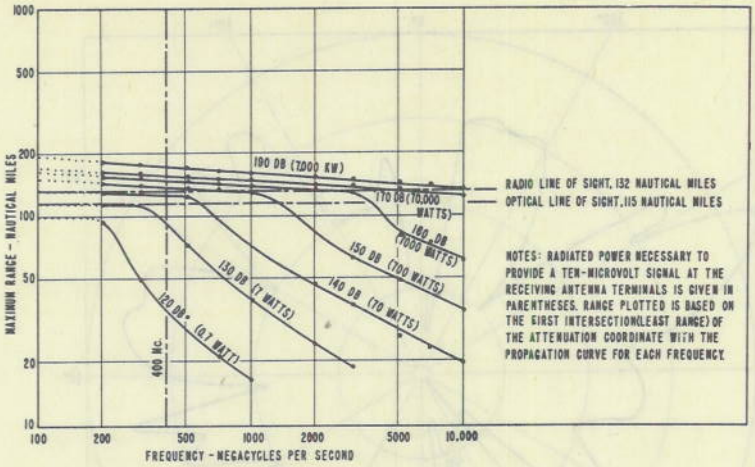


Figure 18 - Maximum range vs. frequency. Attenuation between transmitting and receiving half-wave dipoles as a parameter. Antenna heights, 70-10,000 ft.; vertical polarization; transmission over sea water.

The foregoing discussion is based on radiation and reception with ideal half-wave dipole antennas exhibiting a perfect pattern for vertical orientation of the poles. Measurements of the patterns of typical antennas installed on naval vessels, however, reveal considerable deviations from the ideal. Figure 19 shows the horizontal pattern of a typical vertical half-wave dipole installed on shipboard and operating at a frequency of 328.2 Mc. It is estimated that the variations from the average circular pattern are as great as ±10 db for this case. In order to insure a radio line-of-sight coverage in all directions with this type of antenna pattern at both transmitter and receiver, the power available must be of the order of 20 db greater than that calculated for radio line-of-sight with a system having ideal antennas. In other words, the power must be increased by a factor in the neighborhood of 100 times over the ideal case to allow for antenna pattern deficiencies. Losses in transmission lines and variations in equipment performance over the selected range of operating frequencies will also have some effect on transmitter power requirements.

Fading of VHF and UHF signals received at considerable distances from the transmitter can be caused by variations in the refractive index of the earth's atmosphere and changes in tropospheric reflections.[29] These changes are related to meteorological conditions and the degree of fading is different from day to night and from summer to winter. The fading may be sufficiently selective to cause serious distortion of wide-band signals, such as are used in television. The effects of fading would be more pronounced in wide-deviation f-m systems than in small-deviation systems. In general, the effects of fading would be much the same for f-m or a-m transmissions which occupy the same frequency bandwidth.

---

[29] Federal Communications Commission, TID Report 2.4.5., "Summary of Tropospheric Propagation Measurements and the Development of Empirical VHF Propagation Charts," Oct. 20, 1948

DECLASSIFIED

DECLASSIFIED

RESTRICTED

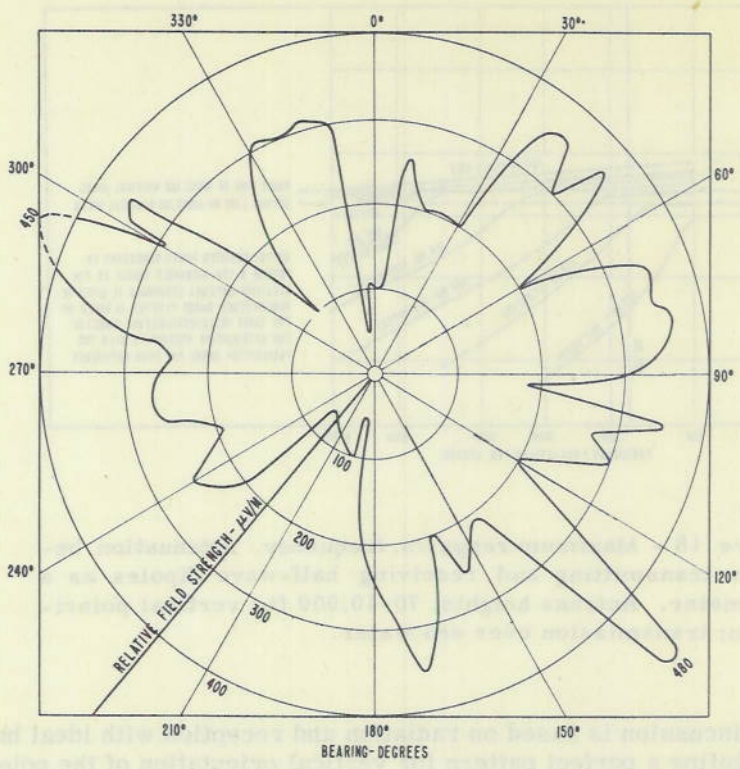


Figure 19 - Typical shipboard antenna pattern measured on a half-wave vertical dipole at 328.2 Mc (measured in horizontal plane)

A-M AND F-M RECEIVERS

A-M Receiver Fundamentals

A block diagram of a typical a-m receiver for speech signals is shown in Figure 20. The voltages induced in the antenna may have components at many frequencies with widely different amplitudes. The "desired" signal will be the amplitude-modulated wave appearing at the operating frequency to which the receiver is tuned. All other signals will then be "undesired" or interfering signals. Insofar as possible, the receiver must amplify only the desired signal from perhaps a few microvolts to a magnitude of volts, and then demodulate it. The demodulating process consists of deriving the modulating waveform from the modulated carrier. A-m demodulators can be considered as operating by causing a "beating" of the carrier with the sidebands to produce the difference frequencies corresponding to the component elements of which the modulating signal is composed.[30] This "beating" process may be considered as a multiplication or heterodyning of the sideband voltage by the carrier voltage and selection of the audio-frequency components of the

[30] Heising, R. A., op. cit.

RESTRICTED

DECLASSIFIED

DECLASSIFIED

product by a suitable filter.[31] Detection is usually effected by a half-wave diode rectifier or an equivalent.

The demodulated intelligence waveform is then amplified by a suitable audio amplifier and fed to a reproducing device such as a loudspeaker or headphones. The ability of a receiver to perform these functions is determined largely by its selectivity and amplification, although many of its other characteristics have definite importance.

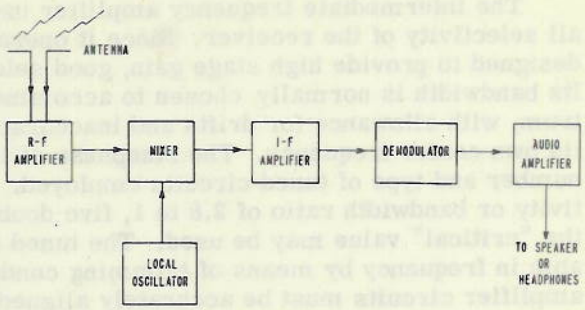


Figure 20 - Block diagram of a typical superheterodyne-type a-m receiver

The superheterodyne type of receiver is now the most widely used and this discussion will therefore be based on its characteristics. The principles underlying its operation are briefly as follows: The incoming signal is selected and amplified by the r-f amplifier and its circuits and fed to the mixer. A locally generated oscillation whose amplitude is usually large compared to the signal amplitude is also fed to the mixer. A beating or heterodyning process takes place in the mixer, whose output contains components at frequencies which include the sum and difference of the local oscillator and signal frequencies. These components are fed to the intermediate-frequency (i-f) amplifier, which amplifies selectivity either a sum or difference frequency component, depending on the design of the receiver. The amplified intermediate frequency is fed to the detector or demodulator which, with the audio amplifier and output device, completes the reception process.

The r-f amplifier serves to select the desired signal from the many signals of various other frequencies which may be present in the antenna circuit. It is designed to offer as much attenuation to the image and intermediate frequencies as possible. The tubes used are chosen to provide stable amplification and low noise, as well as minimum cross modulation. For the latter reason, remote cutoff tubes are often employed. The pass band of the circuits is made sufficiently wide to accommodate the significant portion of the signal spectrum, with proper allowance for the frequency drift and inaccuracies of the transmitter and receiver, including tracking errors and circuit drifts. In general the r-f circuits are tuned with variable air condensers, which are usually ganged to the local oscillator condenser. The number and type of tuned circuits preceding the mixer determines in part the degree of image and i-f attenuation obtained.

The mixer is chosen for optimum conversion efficiency consistent with low noise generation; its noise characteristics are particularly important in high-frequency receivers, where the r-f amplifier has relatively low gain.

The local heterodyne oscillator must furnish an accurate, constant-amplitude voltage to the mixer. Great pains are taken to make the local oscillator frequency as stable as possible and to insure more or less constant output amplitude over its tuning range. Crystal-controlled oscillators are often employed, particularly where unattended or remote-control operation on fixed channels is required. Receiver frequency stability is determined in large part by heterodyne oscillator stability.

---

[31] Harris, D. B., "Selective Demodulation," Proc. I.R.E., Vol. 35, pp. 565-572, June 1947

DECLASSIFIED

DECLASSIFIED

The intermediate frequency amplifier usually contributes most to the gain and overall selectivity of the receiver. Since it operates at a fixed frequency, it can usually be designed to provide high stage gain, good selectivity and good center frequency stability. Its bandwidth is normally chosen to accommodate the desired portion of the signal spectrum, with allowance for drifts and inaccuracies of the transmitter, local oscillator, and its own center frequency. The steepness of its selectivity curve is determined by the number and type of tuned circuits employed. For instance, to obtain a 60 to 6 db selectivity or bandwidth ratio of 2.8 to 1, five double-tuned transformers coupled slightly under the "critical" value may be used. The tuned circuits are usually made somewhat adjustable in frequency by means of trimming condensers or movable slugs in the coils. The amplifier circuits must be accurately aligned so as to provide the desired amplitude and phase characteristics for transmission of the modulated wave with low distortion.

The half-wave diode detector will give good performance when it is designed for and operated at high rectification efficiency, and when the ratio of the a-c load impedance to the d-c load resistance is greater than the average maximum modulation index of the wave being received. Input peak amplitudes greater than 1 to 2 volts are usually required for high rectification efficiency. Residual distortion as a result of variation of the efficiency of rectification with input amplitude is commonly less than 3 percent. The diode characteristic is, in general, essentially square-law for input amplitudes less than one volt and is essentially linear for input amplitudes greater than two volts, when thermionic diodes are used.

The audio amplifier is built with sufficient gain to provide the necessary power to drive the reproducing device. Its bandwidth is generally chosen to accomodate the frequency range of the intelligence. For adequate reproduction of male speech for naval communications, this bandwidth extends from about 300 to 3000 cps. Provision is sometimes made for reducing the audio bandwidth for the purpose of excluding interfering signals. This will cause some of the component frequencies of the intelligence to be lost, but may render the communication more intelligible than if the interference were heard along with the signal.

Receivers which are operated from a-c power sources feed some energy at the power frequency and its harmonics to the output device. This is generally referred to as "hum." It may enter the signal channel through plate, screen, grid or heater circuits or by inductive or capacitative coupling between circuit components. Hum voltages may amplitude- or frequency-modulate the local oscillator, causing undesirable output. The output hum level at full audio and i-f gain is usually kept more than 40 db below the standard signal output level (6 milliwatts into 600 ohms for Navy receivers). The output voltage of a receiver is not zero when all the gain controls are reduced to minimum and no input signal is present. There will normally be a residual output due to audio-amplifier hum. This factor influences the shape of receiver output curves and limits the measurements of output signal-to-noise ratios.

In addition to hum, another cause of residual output with no input signal is audio noise. In a receiver which employs a high-gain audio amplifier preceded by a potentiometer used as an audio gain control, the noise generated by the first audio amplifier tube may be sufficient to cause noise output, even if the audio gain control is set so that the grid circuit resistance is zero. Thus the audio noise is another factor which influences the output characteristics of a receiver.

DECLASSIFIED

DECLASSIFIED

F-M Receiver Fundamentals

A block diagram of a typical f-m receiver is presented in Figure 21. The operation of the components of f-m receivers is much the same as that of a-m receivers, with the exception of the limiter and demodulator. The f-m wave bears the desired intelligence not in its amplitude variations but in its instantaneous frequency variations, as shown previously. The ideal limiter serves the purpose of removing envelope amplitude variations which the incoming wave may contain or which it may acquire in its course through receiver pre-detection circuits. The demodulator circuit performs two functions; it converts the frequency variations of the incoming limited signal into equivalent amplitude variations through the medium of some form of frequency discriminator, and it detects or rectifies the resulting amplitude-modulated wave to produce the intelligence waveform of audio frequency.

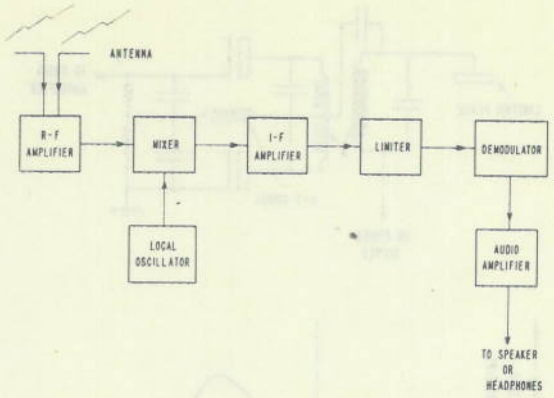


Figure 21 - Block diagram of a typical f-m receiver

A "side-tuned" type of f-m demodulator is shown in Figure 22. Conversion from frequency to amplitude variations occurs if the signal is tuned so that it swings over the sloping portion of the amplitude characteristic of the network. This type of demodulator is generally found in inexpensive receivers and is preceded by some simple form of amplitude limiter which serves the purpose of removing undesired amplitude variations.

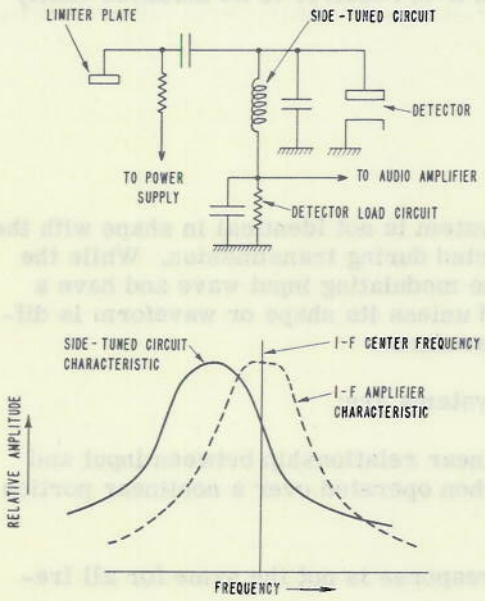


Figure 22 - Side-tuned discriminator relations

The Foster-Seely type of f-m demodulator, shown in Figure 23, provides amplitude cancellation at or near the center frequency. Its operation depends on the relative phase angles of the voltages in tuned transformer circuits. It must be preceded by a suitable limiter to remove amplitude variations, as it in itself has no amplitude-limiting properties.

The ratio detector[32] shown in Figure 24 is a modification of the Foster-Seely type which provides in itself some degree of amplitude limiting. It is a balanced type of f-m detector which has, in effect, a half-wave shunt diode limiter connected across its secondary. The ratio detector

[32] Seely, S. W. and Avins, J., "The Ratio Detector," RCA Review, Vol. VIII, pp. 201-236, June 1947

DECLASSIFIED

DECLASSIFIED RESTRICTED

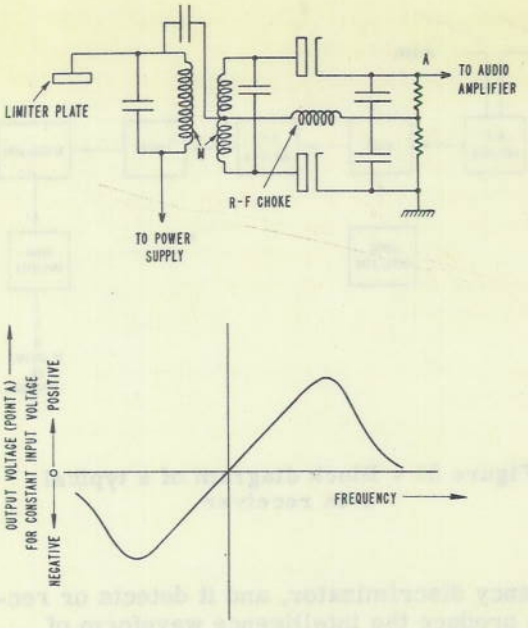


Figure 23 - Foster-Seely discriminator characteristics

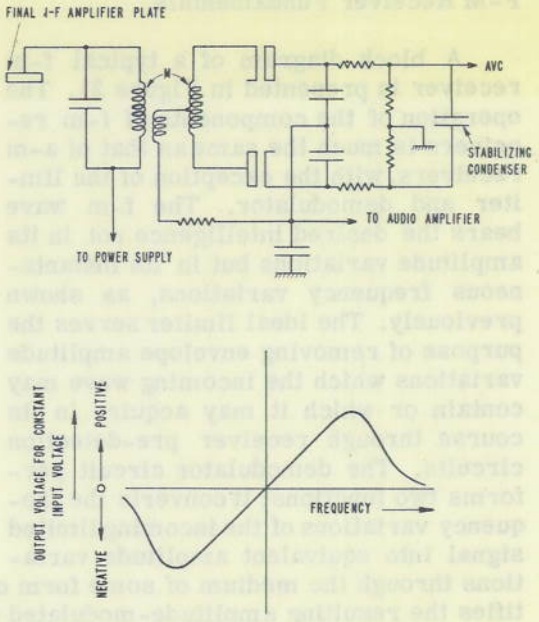


Figure 24 - Ratio detector characteristics

provides a convenient source of AVC voltage, gives fairly adequate performance without a separate limiter preceding it, and hence allows an a-m receiver to be modified easily for f-m operation.

MODULATION DISTORTION

General Considerations

When the output wave from a communication system is not identical in shape with the input modulating wave, it is said to have been distorted during transmission. While the output wave may show a delay in time relative to the modulating input wave and have a different amplitude, the output wave is not distorted unless its shape or waveform is different from that of the original input intelligence waveform.

Several types of distortion in communication systems are:

(a) distortion due to devices which have a nonlinear relationship between input and output signal amplitudes, e.g., vacuum tubes when operated over a nonlinear portion of their dynamic transfer characteristics,

(b) distortion due to circuits whose amplitude response is not the same for all frequencies of the input wave,

(c) distortion due to circuits which delay signals of different frequencies unequally.

RESTRICTED DECLASSIFIED

DECLASSIFIED

These three types of distortion are sometimes referred to as amplitude (or harmonic) distortion, frequency distortion, and time delay (or phase) distortion, respectively.[33] Distortion has also been classified into two categories termed nonlinear distortion or linear distortion, according to whether or not new components having frequencies other than those contained in the original waveform are generated.[34] The latter definition is somewhat confusing, particularly in the case of the transmission of modulated waves through selective circuits in which the phase or amplitude relationships of the components are selectively modified. While no new component frequencies are generated, the original ones have their relative phases and amplitudes modified; the envelope of the modulated wave (in the case of AM), and the instantaneous frequency of the modulated wave (in the case of FM) may contain new components which were not present in the input wave.

Amplitude distortion manifests itself by the generation of harmonics of the input waveform in the case of single sine wave input. In the case of more than one sine wave applied to the nonlinear device, both harmonics and sum and difference frequencies are present in the output waveform. Frequency and phase distortion in audio amplifiers does not result in the generation of components which were not present in the input wave; but since the relative amplitudes and phases of components originally present have been modified, the input and output waveforms have different shapes.

In the case of transmission of modulated waves through r-f and i-f amplifiers, frequency and phase distortion again appears as an alteration of the amplitudes of the carrier and sidebands of the modulated waves. The envelope (in the case of AM) or the instantaneous frequency (in the case of FM) may contain new harmonics and sum and difference frequencies.

The amount of distortion when harmonics of a single tone are generated is usually measured or calculated in terms of a distortion factor, which is equal to the ratio of the square root of the sum of the squares of the amplitudes of the harmonics generated to the rms amplitude of the fundamental output tone. This factor when expressed in percent is termed the "percent harmonic distortion." Distortion may also be rated by the "intermodulation" method,[35,36] in which the effect of a low-frequency sine wave on a high-frequency sine wave is measured. Intermodulation distortion figures are related to harmonic distortion figures, because the same fundamental factors are involved in producing the distortion.

In comparing f-m and a-m systems from the point of view of distortion, differences exist only in those circuits through which the modulated waves are transmitted or in which they are generated or detected. Distortion of audio waveforms, either in the transmitter or receiver, would be the same for both f-m and a-m systems, since, in general, the same waveforms and the same sorts of circuits and tubes are involved. It should also be noted that distortion can be held to very small values by proper design of system components.

---

[33] Terman, F. E., Radio Engineering, p. 226, McGraw-Hill, New York, 1947

[34] Goldman, S., "Frequency Analysis, Modulation, and Noise," p. 49, McGraw-Hill, New York, 1948

[35] Hilliard, J. K., "Distortion Tests by the Intermodulation Method," Proc. I.R.E., Vol. 29, pp. 614-620, Dec. 1941

[36] Warren, W. J. and Hewlett, W. R., "An Analysis of the Intermodulation Method of Distortion Measurement," Proc. I.R.E., Vol. 36, pp. 457-466, April 1948

DECLASSIFIED

DECLASSIFIED

Hence, distortion rarely becomes an overwhelming consideration. Distortion is, however, a cumulative item; it increases as more distortion producing elements are cascaded, and becomes an important factor in long relay systems having many repeaters.

The design of single-link communication systems to provide minimum distortion does not always result in optimum performance in other respects, such as signal-to-noise ratio. For example, an f-m system having a unity ratio of peak-deviation to half i-f bandwidth would be highly undesirable from the point of view of the distortion which would result from even slight frequency drifts and mistuning, but would be desirable from the point of view of signal-to-noise ratio, particularly in the case of weak input signals.

Comparison Between F-M and A-M System Distortion

The following remarks on the comparison of f-m and a-m systems from the distortion point of view will be made with regard to that part of the system in which a modulated wave is being transmitted. This will normally include the output stages of the transmitter and the r-f and i-f amplifiers in the receiver up to the demodulator.

F-m systems have a special advantage over a-m systems in regard to amplitude distortion (that caused by nonlinear tube characteristics). The nonlinear action of the tube distorts the modulation of an a-m wave, but does not affect the instantaneous frequency of the f-m wave, insofar as that part of the spectrum which is usually selected by succeeding stages is concerned.

Consider a nonlinear tube characteristic given by[37]

$$i_p = a_o + a_1(e_i) + a_2(e_i)^2 + \ldots\ldots a_n(e_i)^n, \tag{16}$$

where i_p is the plate current,

e_i is the input voltage, and

$a_o \ldots\ldots a_n$ are constants which determine the tube's transfer characteristic.

If a wave expressed by

$$e_i = A(t) \sin [\Omega_o t + \psi(t)]$$

is impressed on the tube, the plate current may be expressed by

$$i_p = \sum_{\nu=0}^{\infty} B_\nu (t) \sin \nu [\Omega_o t + \psi(t)]$$

In the above equation ν is an integer and $\psi(t)$ represents a phase variation which may be derived from a frequency variation. A(t) is the amplitude of the f-m wave which may be considered constant for the case of pure FM.

It can be shown that B_ν (t) is constant when A(t) is constant. Thus for the case of pure FM at the input, the output consists of a series of groups of components corresponding to the fundamental ($\nu = 1$) and harmonics ($\nu > 1$) of the input spectrum. The peak deviation of the harmonic waves is equal to ν times the input peak deviation. Since usually

[37] Güttinger, P., "Frequenzmodulation," p. 40, AG. Gebr. Leeman and Co., Zürich, 1947

DECLASSIFIED

only the fundamental of the spectrum is transmitted by succeeding circuits, the wave is not distorted.

The amplitude limiter is a highly nonlinear device, which produces little distortion of f-m waves but produces severe clipping of a-m waves. Thus amplitude distortion in the transmission of f-m waves is not normally important. Amplitude distortion in a-m systems must be avoided by careful choice of vacuum tubes and operating bias potentials. The proper use of AVC in a-m receivers is instrumental in avoiding much of this type of distortion.

The tuned circuits in communication systems cause distortion of the modulated waves because the relative amplitudes and phases of the input waves are not preserved during transmission. The distortion can appear as harmonics of the modulating tones, sum and difference tones, or as unequal output amplitudes for different modulating frequencies (similar to frequency distortion in audio amplifiers). Theoretical treatments of tuned circuit distortion have been carried farther for the case of f-m transmission than for a-m transmission. Most of these treatments, however, apply to the case of single-tone modulation and neglect detuning.

Tuned Circuit Distortion

Those characteristics of tuned circuits which are involved in distortion calculations are the frequency response and the phase response. For distortionless transmission of modulated waves, the frequency-response curve of the circuit must be constant for all frequencies contained in the wave, and the phase shift introduced by the circuit must either be zero for all frequencies involved (no time delay), or be a linear function of frequency (equal time delay for all components). There are no physically realizable tuned circuits which possess these ideal characteristics and hence all practical circuits introduce distortion to some extent. In general, freedom from tuned-circuit distortion is achieved by providing sufficient bandwidth in the tuned circuits for the more important components of the wave. For the same amount of distortion, it is generally necessary to provide a wider circuit bandwidth with respect to signal spectrum bandwidth in the case of f-m systems.

Two types of tuned circuits are commonly employed in communication systems, single-tuned circuits and double-tuned (two-tuned-coupled) circuits. Calculations of the distortion generated by these circuits have been made for the case of f-m transmission.[38,39,40] The parameters involved in calculating tuned-circuit f-m distortion for single-tone modulation are (1) the ratio of the peak frequency deviation of the modulated wave to the 3-db bandwidth of the tuned circuit, and (2) the ratio of the peak frequency deviation to the modulating frequency.

---

[38] Roder, H., "Effects of Tuned Circuits Upon a Frequency-Modulated Signal," Proc. I.R.E., Vol. 25, pp. 1617-1647, Dec. 1937

[39] Jaffee, D. L., "A Theoretical and Experimental Investigation of Tuned-Circuit Distortion in Frequency-Modulation Systems," Proc. I.R.E., Vol. 33, pp. 318-333, May 1945

[40] Philco Research Division, Philco CD-838, Fourth Quarterly Report on Contract W-36-039-sc-32114, "Radio Frequency Conversion as a Method of Radio Relaying," (Restricted) May 15, 1947

DECLASSIFIED

The distortion may be classified as "stationary" or "dynamic." The stationary distortion results from the nonlinearity of the phase and amplitude characteristics of the circuit; these are effective even when the frequency variation occurs at a slow rate. The dynamic part becomes important only when the ratio of the maximum frequency deviation to the modulating frequency is less than about 3 (as with high modulating frequencies). This distortion results from the inability of tuned circuits to follow rapid frequency variations. For the case of an accurately tuned signal, it has been shown that the stationary and dynamic parts of the distortion are in phase quadrature.[41]

In f-m systems, the amount of such distortion increases rapidly with modulating frequency. Distortion also increases in proportion to the logarithm of the number of tuned circuits cascaded. The double-tuned circuit generally causes greater distortion of the high modulating frequencies than the single-tuned circuit because of greater phase shift and sharper frequency cutoff, but causes less distortion of the low modulating frequencies, because of more constant response with change of frequency. Single- and double-tuned circuits cause equal distortion when the ratio of modulating frequency to the half 3-db bandwidth of the tuned circuit is approximately 1/8.

Distortion calculations made on the basis of single-tone sine waves are apt to be misleading, since normal intelligence functions usually contain many sine waves which are distributed throughout the audio band. The deviation corresponding to each sine wave in a group is necessarily smaller than the peak deviation caused by all of the waves in the group acting together. In addition, since the distortion is less for the lower modulating frequencies, distortion calculations based only on high-frequency single-tone modulation are apt to be somewhat pessimistic when applied to circuits normally used for the transmission of speech or music.

The nature of the waveform also affects the calculated distortion. Consider the case of a modulating signal made up of a large number of tones. The distortion of each tone may be computed separately and appropriately summed on an rms basis to obtain a value for the total distortion. If a similar calculation is made for the case of a smaller number of tones which produce a peak deviation equivalent to the peak deviation of the tones considered above, the calculated distortion in the former case (involving the greater number of tones) would be smaller. Presence of multiple tones would, however, lead to intermodulation distortion products. For the case of stationary distortion, the improvement of computed multiple-tone distortion over single-tone distortion varies as the 4th power of the ratio of the number of tones considered in the two cases.

An example of the distortion which may be expected from 5 double-tuned circuits cascaded is contained in the following two sets of data:

Distortion due to 5 Double-Tuned Circuits

	Case 1	Case 2
Modulating frequency	3 kc	3 kc
I-f 3-db bandwidth	60 kc	125 kc
Peak frequency deviation	20 kc	7 kc
Percent harmonic distortion	1 %	.001%

[41] Güttinger, P., op. cit., p. 50

[redacted]

DECLASSIFIED

DECLASSIFIED

It is seen that case 2 above, in which the i-f bandwidth is considerably wider and the peak frequency deviation much lower, has the lower distortion. In general, the greater the percent modulation of the receiver pre-detector pass band, the greater the distortion.

A recent development in the distortion field is the design of all-pass lattice networks to compensate for tuned-circuit distortion.[42] The following data show what may be achieved by this method:

Conditions

Center frequency of transformer	24.0 Mc
3-db bandwidth	0.86 Mc
Peak frequency deviation	10.3 Mc
Modulating frequency	60,000 cps
Transformer coupling (No detuning)	critical

Data

Order of Harmonic	% Distortion for a Single Transformer	% Distortion for a Single Transformer with Correcting Network
3rd	1.3	0.17
5th	0.34	0.34
7th	0.01	0.03
9th	0.03	0.01
11th	0.07	0.00

When the incoming signal is centered on the response characteristic of the tuned circuit(s), for the case of single-tone modulation, only odd harmonics of the modulating tone appear as distortion. This arises because of the normal symmetry of the phase-shift and amplitude-response characteristic of the network with respect to the unmodulated center frequency of the impressed wave. Detuning of the input modulated wave causes even harmonics of the modulating tones to appear as distortion products.[43,44]

Reduction of a-m distortion produced by tuned circuits may be achieved by the use of wide-band circuits or all-pass phase-compensating filters. In the former case, four times the peak frequency deviation for a modulating frequency of 15 kc will keep the distortion below a level which can be detected visually on a cathode-ray oscilloscope.[45]

Another type of distortion occurs in f-m tuned circuits which is similar to frequency distortion in audio amplifiers. The amplitude of the fundamental modulating tone after transmission through tuned circuits is not the same for all modulating frequencies. The

---

[42] Radio Corporation of America Quarterly Status Report No. 9, "Methods of Modulating and Multiplexing," Signal Corps Contract W-36-039-sc-32075, Aug. 6, 1948

[43] Jaffee, D. L., op. cit., p. 328

[44] Güttinger, P., op. cit., p. 55

[45] Jaffee, D. L., op. cit., p. 328

DECLASSIFIED

DECLASSIFIED

distortion which results is equivalent to pre-emphasis of the high modulation frequencies.[46] This distortion is nearly independent of the ratio of peak frequency deviation to the 3-db half i-f bandwidth when the modulating tone has a frequency of less than 1/4 the 3-db i-f bandwidth. The amplitude of the fundamental output is increased by a factor of approximately 1.26 (2 decibels) when the modulating tone has a frequency equal to 1/4 the 3-db bandwidth of a single-tuned circuit.

Quantitative comparison of tuned circuit distortion in a-m and f-m systems is not possible at this time because of the lack of information in the case of AM. In AM, the amplitudes and phases of the carrier and sidebands may be modified in traversing the tuned circuits. The envelope of the wave, which is determined by the amplitudes and relative phases of the components of the waves, is then modified, resulting in distortion. If the frequency-response curve of the circuits is nearly constant and the phase-shift versus frequency characteristic of the circuit is linear, distortion will be minimized. If the phase-shift curve is nonlinear but essentially symmetrical about the carrier frequency, as in the case of no detuning, a frequency distortion effect may be expected. If the phase-shift curve and the frequency response characteristic are unsymmetrical about the carrier frequency, such as might be the case when the carrier is detuned, harmonic distortion may be expected. When the carrier is attenuated with respect to one or both sidebands, the equivalent of overmodulation may occur, producing serious distortion. When one sideband is attenuated with respect to the carrier and the other sideband increases in amplitude, which may be caused by severe detuning, the output wave may become essentially a single-sideband wave and some distortion will be produced with a linear detector.

When more than one modulating tone is contained in the signal, intermodulation distortion components may be generated during detection. Despite the absence of detailed information on a-m distortion in tuned circuits, it may be surmised that the ratio of required circuit bandwidth to signal spectrum would be less in the case of a-m systems. This may be partially justified by the following argument. If the frequency-response and phase-shift characteristics of the tuned circuits in AM are nonlinear but symmetrical with respect to the carrier frequency, there will be no distortion other than frequency and phase distortion due to relative attenuation and phase-shift of the sidebands. For the same conditions in the case of FM, odd-order harmonic distortion is produced. Harmonic distortion is not produced in the case of AM except when corresponding sidebands have different amplitudes or unequal phase displacement relative to the carrier. Of these various distortions, harmonic distortion is generally the most serious.

Distortion in Modulators and Demodulators

Distortion due to nonlinear amplitude response in modulators and demodulators gives rise to harmonics and intermodulation frequencies. These distortions are generally less in FM than in AM. Reactance tube modulators and phase modulators can be made as linear as desired by providing narrow phase or frequency deviations at the modulator, followed by wide-band frequency multiplier stages. Distortion in a-m modulators depends directly on the dynamic transfer characteristic of the modulator stage. For undistorted modulation in AM, the modulator should have a square-law dynamic amplitude transmission characteristic. Such a characteristic cannot be perfectly obtained over the entire dynamic range of operation and hence imposes an inherent limit to the freedom from distortion which may be attained in a-m systems.

---

[46] Philco Research Division Report, op. cit., Appendix D

DECLASSIFIED RESTRICTED

DECLASSIFIED

While half-wave detection is usually employed in a-m receivers, a balanced detector is customary in FM. The balanced detector is effective in reducing even-harmonic distortion which may be generated due to nonlinearity of the detector characteristic. If desired, however, a-m receivers could be built with balanced detectors.

In summary, f-m waves are generally not affected by nonlinearities of vacuum tube characteristics, whereas a-m waves are affected. Tuned-circuit distortion imposes more limitations on the design of f-m systems than it does in the case of a-m systems. Frequency modulators can be made more distortion-free than amplitude modulators, and f-m and a-m demodulators are about on a par insofar as distortion is concerned.

DISTORTION ARISING FROM MULTIPATH TRANSMISSION

A type of modulation distortion which may occur in communication systems is that caused by the existence of more than one radio path between the transmitting and receiving antennas. The individual waves which reach the receiving antenna over the various paths, while all originating at the same transmitter and all bearing the same modulating intelligence, may be delayed in time by various amounts. They may also have the phases and amplitudes of their carrier and side frequencies modified due to absorption, scattering, or reflection phenomena, which may occur during transmission. The effect may be considered as a special form of interference, where the interfering waves are generated by the same transmitter which generates the desired wave. Early observations of these effects indicated that the distortion is more pronounced in f-m than in a-m transmission.[47] A-m distortion as well as f-m distortion has been observed in the case of long-distance transmission paths when reflections from the ionosphere caused the existence of multiple paths.[48,49,50] F-m distortion has also been observed in the case of short-distance radio links when hills, mountains, and large buildings having metal framework served as reflectors in causing multiple paths.[51,52]

Theoretical analysis of this type of distortion has been confined to the case of two paths having different lengths (and hence different time delays), when the modulating wave consists of a single sinusoid and when frequency-selective absorption and phase shift do not occur.

---

[47] Eckersley, T. L., "Frequency Modulation and Distortion," Experimental Wireless and Wireless Engineer, Vol. 7, pp. 482-484, Sept. 1930

[48] Potter, R. K., "Transmission Characteristics of a Short-Wave Telephone Circuit," Proc. I.R.E., Vol. 18, pp. 581-648, April 1930

[49] Crosby, M. G., "Frequency Modulation Propagation Characteristics," Proc. I.R.E., Vol. 24, pp. 898-913, June 1936

[50] Crosby, M. G., "Observations of Frequency-Modulation Propagation on 26 Megacycles," Proc. I.R.E., Vol. 29, pp. 398-403, July 1941

[51] Mayo, A. D. and Sumner, C. W., "FM Distortion in Mountainous Terrain," QST, Vol. 28, pp. 34-36, March 1944

[52] Corrington, M. S., "Frequency Modulation Distortion Caused by Multipath Transmission,'" Proc. I.R.E., Vol. 33, pp. 878-891, Dec. 1945

DECLASSIFIED

DECLASSIFIED

The parameters involved in determining the degree of distortion in both a-m and f-m systems are the following:

(a) the relative carrier frequency phase delay, $\omega(\Delta t)$,
where ω = angular velocity of the carrier wave
and Δt = time-delay difference between the two paths;

(b) the relative modulating-frequency phase delay, $p(\Delta t)$,
where p = angular velocity of the modulating wave;

(c) the ratio of the peak amplitudes of the waves corresponding to the two paths, $b = \frac{E_2}{E_1}$
where E_1 and E_2 are the peak unmodulated carrier amplitudes of the two waves.

The distortion is greatest when the relative carrier-frequency phase delay $\omega(\Delta t)$ is an odd multiple of π radians (corresponding to amplitude opposition of the carrier waves), and when the ratio of the unmodulated peak carrier amplitudes, $b = E_2/E_1$, is near unity (equality of wave amplitudes). The distortion increases with the relative time delay Δt, becoming severe when the relative modulating-frequency phase delay is several degrees.

The spectra of the two sinusoidally modulated waves which reach the receiver (in either FM or AM) will contain components of the same frequencies. The path-length difference introduces phase and amplitude differences in the components of the spectra, but the frequencies of the components are unaltered. With this fact in mind, the distortion may be likened to that caused by tuned circuits. It should be noted that in the case of FM, the instantaneous frequencies of the two received waves are different most of the time, due to the modulation. The instantaneous difference frequency is thus continually changing with time with FM, whereas with AM it remains fixed.

In a-m systems, two-path distortion in the case of single-tone modulation generally manifests itself as a change of the received carrier amplitude, accompanied by a change of the modulation index.[53] The effective modulation index can exceed unity for certain modulation frequencies and distortion in the form of harmonics of these frequencies results.

In f-m systems, the two signals combine to form a wave whose instantaneous frequency varies rapidly between very wide limits.[54,55,56] Sharp spike-shaped waveforms appear at the detector output when the input signal amplitude ratio approaches unity. The audio amplifier generally has the effect of smoothing out these sharply peaked waveforms, since it does not normally respond to the very high harmonic components which contribute to the sharpness of the spikes. For a given modulating frequency, f-m multipath distortion is roughly independent of the modulation-frequency phase delay when the peak-frequency deviation is large. When the peak-frequency deviation is small, the distortion increases as the modulation phase delay is increased, the distortion becoming appreciable for

---

[53] Gerks, I. H., (Collins Radio Company, Cedar Rapids, Iowa) "An Analysis of Distortion Resulting from Two-Path Propagation," Paper presented at I.R.E. Convention, New York City, 1949. (Proc. I.R.E. Vol. 37, pp. 1272-1277, Nov. 1949)

[54] Gerks, I. H., op. cit.

[55] Corrington, M. S., op. cit., (footnote 52 of this report)

[56] Meyers, S. T., "Nonlinearity in Frequency-Modulation Radio Systems due to Multipath Propagation," Proc. I.R.E., Vol. 34, pp. 256-265, May 1946

RESTRICTED
DECLASSIFIED

DECLASSIFIED

modulation phase-delay angles as low as two degrees. Under the worst condition of delay (modulation-frequency phase delay equal to 180 degrees) the distortion is greatest and is therefore most evident for the high modulating frequencies.

In an a-m system, serious harmonic distortion will not occur until over-modulation occurs, whereas in f-m systems the distortion initially increases with increasing peak frequency deviation, and then decreases as more and more of the distortion is distributed among higher harmonics which are attenuated by the audio amplifier.[57] A-m systems may thus be expected to give better performance in regard to multipath distortion than narrow-deviation f-m systems. A quantitative comparison of f-m and a-m systems in this respect must await further analysis of the a-m case and extended treatment of the f-m case.

In naval service, it is conceivable that multipath conditions could exist in a large task force, in plane-to-plane communications (particularly when the planes are at high altitudes), and when ships are operating near a shore line. Under these conditions, more than one path with significant time delay differences can exist and distortion might result, particularly for systems having high modulation-frequency components (e. g., television). For the case of UHF ship-to-ship voice communications, the transmission range is limited to line-of-sight distances, so that differences in modulation-frequency time delays over possible multiple paths would not normally be expected to be large enough to give rise to serious distortion in either a-m or f-m systems.

Studies of multipath distortion are now in progress which may provide further information on which to base a quantitative comparison of f-m and a-m systems in this respect.[58]

A-m multipath distortion may be reduced by the use of such devices as exalted carrier reception[59] and, in some cases, by judicious use of directive antennas. In f-m systems, multipath distortion may be minimized by the use of extremely wide bandwidth in the discriminator circuit.[60]

CHANNEL SPACING CONSIDERATIONS IN COMMUNICATION SYSTEMS

For the purpose of maintaining reliable communication service in a multichannel system, it is generally required that the amplitude of the response due to the various components of the desired signal remain within a 6-db range, so that the receiver working bandwidth is normally that between the 6-db attenuation points. Furthermore, it is usually required that the response of the receiver to a signal of the same amplitude on an adjacent channel shall be attenuated 40 to 60 db, or more, with respect to the desired signal. Design of a system to meet these requirements necessitates consideration of the channel spacing, selectivity ratio of the receiver, drifts and inaccuracies of the oscillators and tuned circuits employed, and the width of the frequency spectrum of the transmitted wave.

---

[57] Gerks, I. H., op, cit.

[58] Alford, A., "Multipath Phenomena and Their Effect on Communication Systems," Report No. 2 on Contract W19-122-ac-27 (Unclassified), Feb. 15, 1949

[59] "High Frequency Communication Methods," Final Report on Contract W-36-30-sc-32244, Paul Godley Company (Great Notch, N. J.), April 1, 1948

[60] Gerks, I. H., op. cit.; Arguimbau, L. B. and Granlund, J., "Sky-Wave FM Receiver," Electronics, Vol. 22, p. 101, Dec. 1949. (See also references (3), (4), and (5) on p. 103.)

DECLASSIFIED

DECLASSIFIED

RESTRICTED

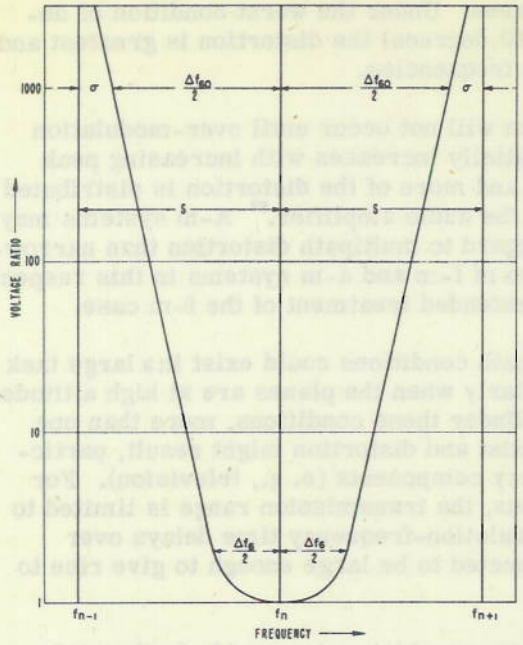


Figure 25 - Characteristics of a communication system

The 6-db bandwidth of the receiver must accommodate the combined maximum frequency inaccuracies and drifts of all the oscillators involved in the system, the inaccuracies and drifts of all the tuned amplifiers, and an allowance for the sidebands of the desired modulated wave. When the expected inaccuracies and drifts of the system are known and the 6-db bandwidth of the receiver has been chosen, the remaining two factors, selectivity ratio of the receiver and channel-spacing, bear a definite relation to each other.

Figure 25 shows the over-all selectivity characteristic of one receiving channel and the location of the adjacent channel transmitters in the frequency spectrum. The nominal maximum 60-db bandwidth of the receiver is labeled Δf_{60}, the nominal 6-db bandwidth of the receiver is labeled Δf_6, and the nominal channel spacing is labeled S. If the quantity σ in equation (17) is made equal to the maximum departure of the transmitter from its nominal value plus the nominal maximum drift of the receiver, plus the difference between the carrier frequency and the frequency of the highest significant sideband, the adjacent channel signal will be attenuated at least 60 db with respect to the desired signal under most conditions.

It is seen from Figure 25 that for satisfactory performance,

$$S \geq \frac{\Delta f_{60}}{2} + \sigma \tag{17}$$

or

$$S \geq \frac{\Delta f_6}{2}\left[\frac{\Delta f_{60/2}}{\Delta f_{6/2}} + \frac{\sigma}{\Delta f_{6/2}}\right] \tag{18}$$

or

$$S \geq \frac{\Delta f_6}{2}\left[R_{60/6} + \frac{2\sigma}{\Delta f_6}\right], \tag{19}$$

where $R_{60/6}$ is the selectivity ratio of the receiver. Equation (19) may be solved for $R_{60/6}$ or σ, giving

$$R_{60/6} \leq \frac{2}{\Delta f_6}(S - \sigma) \tag{20}$$

and

$$\sigma \leq S - \frac{R}{2}\Delta f_6 . \tag{21}$$

RESTRICTED

DECLASSIFIED

DECLASSIFIED

Equations (19), (20), and (21) are helpful in the specification of system parameters for 60-db adjacent-channel attenuation. The quantities contained within them are defined as follows:

σ = sum of the maximum expected transmitter and receiver frequency drifts and inaccuracies, plus the difference between the carrier frequency and the frequency of the highest significant sideband;

Δf_6 = 6-db over-all bandwidth of the receiver up to the final detector;

Δf_{60} = 60-db over-all bandwidth of the receiver up to the final detector;

S = channel frequency separation;

$R_{60/6}$ = ratio of 60-db to 6-db over-all bandwidths of the receiver.

Modulation suppression or capture effects result in an apparent increase in the selectivity of the receiver. Hence the adjacent channel attenuation in a system designed according to the preceding formulas may under some conditions be greater than 60 db, insofar as the output modulation is concerned. This, however, would not be true of heterodyne interference effects, if such occurred at a frequency within the useful output band of the receiver.

INTERFERENCE IN COMMUNICATION SYSTEMS

The types of interference which are encountered in communication systems may be classified as follows:

(a) Fluctuation noise (random noise).

(b) Impulse noise.

(c) Common-channel interference.

(d) Off-channel interference (including cross-modulation and adjacent-channel interference).

(e) Interference due to spurious receiver responses.

The sources and characteristics of these several types of interference will be discussed in the following paragraphs.

Fluctuation noise is generated in the resistive components of circuits and in the vacuum tubes of communication systems. It is caused by the thermal agitation of electrons in the networks and by random effects such as emission variations in vacuum tubes. The mean-square noise voltage at the terminals of an impedance is given by:[61]

$$\overline{e_n^2} = 4\,KT \int_{f_1}^{f_2} R(f)\, df, \qquad (22)$$

---

[61] Terman, F. E., op. cit., p. 476 (footnote 10 of this report).

DECLASSIFIED

DECLASSIFIED

where K is Boltzmann's constant,

T is the temperature of the circuit in degrees Kelvin,

R(f) is the resistive component (in ohms) of the impedance across which the noise is developed,

f is frequency in cycles per second, and f_1 and f_2 are the frequency limits of the band in which the noise is observed.

The energy of fluctuation noise at its source is uniformly distributed over the entire range of commonly employed audio and radio frequencies. The amplitudes and phases of the component noise voltages are completely random. The input circuits and vacuum tubes of receivers are the primary sources of the thermal agitation noise which is important in communication systems. Random noise generated in vacuum tubes is caused by differences in emission velocities of electrons from the cathode, random division of space current between plate and screen circuits, and random induction of voltages on the control grids. Tube noise often contributes more noise energy than does circuit thermal agitation noise, and these two effects taken together determine the minimum detectable signal for a communication system operating under the ideal condition of no external electrical noise.

Impulse noise is, in general, generated external to the receiver and is produced by electrical machinery, ignition systems, radar equipment, atmospheric static, etc., which generate short-duration high-amplitude pulses. Part of the spectrum of such pulses will lie in the pass band of the receiver, resulting in the familiar forms of impulse noise interference. Generally, the part of the spectrum received consists of approximately equal-amplitude components related in phase and equally spaced in frequency. The peak value of such impulse noise after being selectively amplified by a receiver has been shown to be directly proportional to the bandwidth of the receiver, and the rms value of the impulse noise is proportional to the square root of the bandwidth.[62]

Common-channel interference is usually generated by transmitters operating within the assigned frequency channel of the desired signal. Off-channel interference includes adjacent-channel interference and is caused by transmitters operating at frequencies outside the desired signal channel but having sufficient energy to overcome the attenuation provided against them by the selectivity of the receiver.

Interference to reception may be caused by transmitters operating on the image or intermediate frequencies of the receiver if they create a sufficiently strong signal to overcome the image and i-f attenuation provided by the receiver preselector circuits.

A-m and f-m receivers behave differently in the presence of these types of interference. A general idea of the reason for this may be obtained by reference to Figure 26. This figure shows a signal vector S, an interference vector I and the resultant of these two vectors R. The signal and interference vectors may be considered as variable quantities to allow consideration of either f-m or a-m signals with either f-m or a-m interference. The differences between the resultant vector R and the signal vector S represent the changes brought about by the presence of the interference I. In a-m systems the amplitude of the resultant R is of importance, while in f-m systems the time derivative of the angle ϕ, divided by 2π (the instantaneous frequency of the resultant vector R), is the

---

[62] Landon, V. D., "A Study of the Characteristics of Noise," Proc. I.R.E., Vol. 24, pp. 1514-1521, Nov. 1936

DECLASSIFIED

DECLASSIFIED

quantity which must be considered. An ideal amplitude-modulated signal would be represented by amplitude variations of the vector S as it rotates about 0 at a constant velocity, and an ideal frequency-modulated signal would be represented by variation in the angular velocity of rotation of the vector S while its amplitude remains constant.

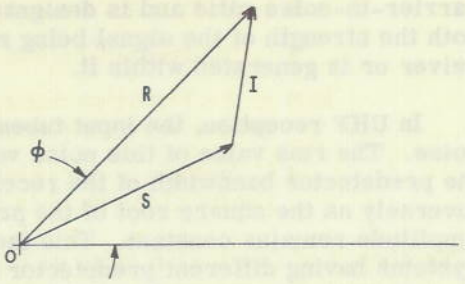


Figure 26 - Vector diagram illustrating the effect of interference in a-m and f-m systems

Addition of the interference vector I to the signal vector S results in variations of the resultant signal in both instantaneous angular velocity and amplitude, as evidenced by the variations in the amplitude of R and its angular rotational velocity. The effect of the vector I in determining the amplitude and phase of the resultant R depends on the amplitude of I. In a-m systems, the simplest means of reducing the effect of interference is to increase the amplitude of S with respect to I. In f-m systems, increasing the amplitude of S without changing I will reduce the magnitude of the phase variations of the resultant. In addition, when the signal is greater in amplitude than the interference, increasing the magnitude of the variations in angular velocity of the signal vector S (i. e., increasing the modulation index) will serve to reduce the effect of the interference vector on the phase variations of the resultant vector R. Thus it may be expected that wide-deviation f-m systems should be superior to narrow-deviation f-m systems with regard to their behavior in the presence of a given magnitude of interference at the input to the receiver, when the signal-to-interference ratio exceeds unity. It can be expected that f-m systems may be superior to a-m systems for the case of large peak-frequency deviations and signal-to-interference ratios in excess of unity.

If the interfering signal exceeds the desired signal in amplitude, the same general reasoning may be used with the signal and interference interchanged. The frequency variations of the signal would have little effect in determining the frequency variations of the resultant. There is a sharp change in the performance of f-m systems in the presence of interference in the region where the signal-to-interference ratio is near unity. This region of change from good performance to poor performance is termed the threshold. In a-m systems the threshold is not nearly so sharp, there being a relatively smooth deterioration in performance as the input signal-to-interference ratio passes from values above unity to values below unity.

FLUCTUATION NOISE IN A-M RECEIVERS

With efficient receiver designs, the fluctuation noise voltages which have their origin in the tubes and circuits of a receiver determine the minimum signal voltage which can be made intelligible by the receiver under conditions of no external interference. Stated another way, the fluctuation noise characteristics of the receiver generally determine the maximum reliable range of a communication system in the absence of external interference for given conditions of transmitter power, antenna gains, and propagation. The study of the fluctuation noise characteristics is thus of primary importance.

An a-m receiver can amplify the sum of the modulated carrier and noise voltages substantially linearly up to the point of demodulation, provided that no amplitude limiting or nonlinearity is present in the amplifiers preceding the demodulator. The ratio of the rms carrier voltage to the rms noise voltage just prior to the final detector is termed the

DECLASSIFIED

DECLASSIFIED

carrier-to-noise ratio and is designated x. It is to be noted that this ratio, x, depends on both the strength of the signal being received and the noise power which enters the receiver or is generated within it.

In UHF reception, the input tubes and circuits are the primary sources of fluctuation noise. The rms value of this noise voltage is normally proportional to the square root of the predetector bandwidth of the receiver. Thus the carrier-to-noise ratio, x, will vary inversely as the square root of the predetector bandwidth of the receiver if the input signal amplitude remains constant. This fact must be remembered when making comparisons of systems having different predetector bandwidths.

The combination of modulated carrier and noise voltage is detected, this process producing voltages having frequencies, in the general case, which fall in the range from 0 to 2β, where β is the half-bandwidth of that part of the receiver preceding the detector. The voltages which contribute to the noise output of the detector arise from the "beating" in the detector circuit of the noise components lying within the bandwidth of the receiver, with each other, with the carrier, and with the sidebands of the signal. The voltages producing the intelligence output result from the beating of the sidebands of the modulated wave with the carrier in the detection process. A distortion component of small magnitude is produced by the beating of the sidebands with each other.

The ratio of the rms voltage of the modulation to the rms voltage of the noise immediately after the detection process is termed the detected output signal-to-total-noise ratio and is designated $(S/N)_t$. The audio amplifier serves to select only those frequencies which fall within its transmission range and transmits these components to the output device. The ratio of the rms voltage of the modulation to the rms voltage of the noise at the output device is termed the system output signal-to-noise ratio and is designated $(S/N)_o$. The quantities defined above enter into theoretical treatments of noise in receivers.

The above signal-to-noise ratios are difficult to measure directly. The ratio of signal-plus-noise to noise is the quantity usually measured, utilizing generally available meters which, however, do not indicate true values of rms voltages for complex waveforms. Thus the comparison of experimental data with theory requires special measurements and calculations. Hum and audio noise voltages which are not usually considered in theoretical treatments contribute to output meter indications. Their effect on measurements must be carefully considered. In the following paragraphs, the noise characteristics of receivers will be discussed from a theoretical point of view in order to point out the approximate limits of receiver performance in the presence of fluctuation noise.

The relations between $(S/N)_o$ and x have been derived for a-m receivers having idealized response characteristics when the input signal is a single-tone amplitude-modulated wave, accurately tuned to the center of the response characteristic of the receiver.[63,64,65]

---

[63] Ragazzini, J. R., "The Effect of Fluctuation Voltages on the Linear Detector," Proc. I.R.E., Vol. 30, pp. 277-288, June 1942

[64] Burgess, R. E., "The Rectification of Signal and Noise by Linear and Square-Law Detectors," Paper No. R.R.B./C.93; Radio Department, National Physical Laboratory, dated March 24, 1944

[65] Wing, A. H., "Terminal Report on Frequency-Modulation vs. Amplitude Modulation for use in Airborne VHF Equipment," Problem No. 2, Service Project No. AC-230.03, Cruft Laboratory, Harvard University, OSRD Report No. 6280, Dec. 1, 1945

DECLASSIFIED

DECLASSIFIED

The input noise is considered as consisting of a great number of components having equal amplitudes and random phases uniformly distributed throughout the bandwidth of the receiver. The detected noise arises from beats between these components and the various components of the modulated wave. The signal-to-total noise ratio $(S/N)_t$ is given in terms of the carrier-to-noise ratio x as follows:

$$\left(\frac{S}{N}\right)_t = mx^2\left[\frac{2}{1 + x^2(2 + m^2)}\right]^{1/2}, \qquad (23)$$

where m is the index of modulation. This relation is shown graphically in Figure 27. Equation (23) was derived for the case of a square-law detector and is a good approximation for the case of a linear detector. The approximation is most accurate for small values of m and large values of x.

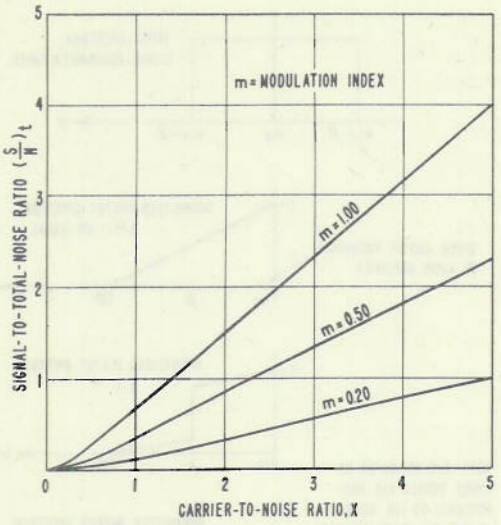


Figure 27 - Relations between carrier-to-noise ratio and signal-to-total-noise ratio for square-law detection of an a-m wave in the presence of fluctuation noise (these curves apply approximately to the linear detector).

The output signal-to-noise ratio $(S/N)_o$ is greater than the signal-to-total-noise ratio because of the effect of audio selectivity on the spectrum of the noise output from the detector. The detector output noise spectrum extends over the frequency range from 0 to 2β when the carrier amplitude is zero. As the carrier amplitude increases (x increases), the detector output noise spectrum shrinks so that it extends to a frequency equal to half the i-f bandwidth for large carrier-to-noise ratios. Figure 28 shows several detector output spectra for different values of x when an unmodulated carrier and a noise spectrum are present at the input to a square-law detector. The change in the shape of the noise spectrum is explained by the fact that the detected noise voltages which result from the input noise components beating among themselves, become insignificant in comparison with those which result from the input noise voltages beating with the carrier as the carrier increases in amplitude. This also explains the increase of detected noise when a carrier is tuned in.

When a modulated carrier is considered, similar detector output spectra are obtained. Since the audio amplifier accepts only a part of the detected noise energy, but in general almost all the detected modulation energy, the relation between the output signal-to-noise ratio and x can be approximated by:

$$\left(\frac{S}{N}\right)_o \cong mx^2\left\{\frac{2\beta}{\omega_c\left[1 + x^2(2 + m^2)\right]}\right\}^{1/2}, \qquad (24)$$

where 2β and ω_c are the "noise bandwidths" in radians per second of the i-f and audio amplifiers, respectively. (The noise bandwidth is defined as the width of the equivalent rectangle corresponding to the energy transmission characteristic of the amplifier.) Equation (24) involves two assumptions: That the audio-amplifier bandwidth is small compared to the half-bandwidth of the i-f amplifier, and that the audio-amplifier responds to the low-frequency part of the detector output noise spectrum where its amplitude is essentially uniform. These conditions usually exist in UHF receivers, where the i-f bandwidth is necessarily much larger than the audio-amplifier bandwidth to allow for drifts of

DECLASSIFIED

DECLASSIFIED

 RESTRICTED

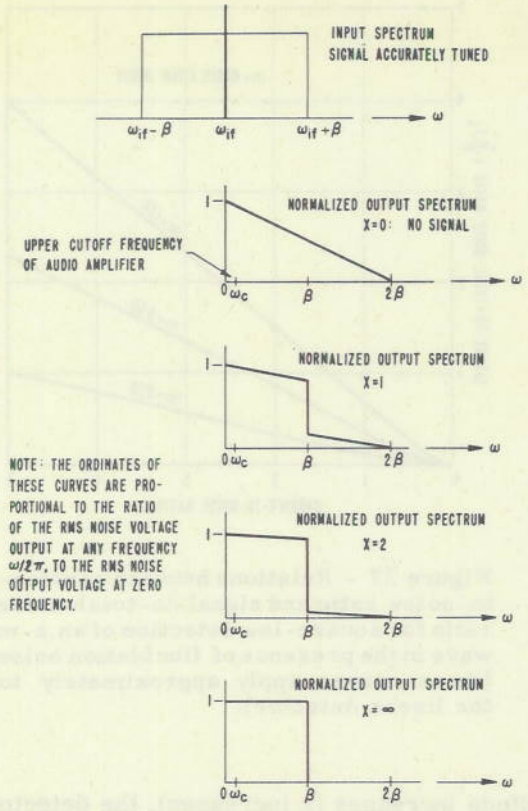


Figure 28 - Output mean square noise voltage spectrum from a square-law rectifier when the input spectrum is rectangular (from footnote reference 64)

oscillators and tuned circuits. Equation (24) may be simplified for large values of x and small values of m, giving:

$$\left(\frac{S}{N}\right)_o \cong mx\sqrt{\frac{\beta}{\omega_c}}. \qquad (25)$$

Thus for values of x which are large compared to unity, the output signal-to-noise ratio of an a-m receiver is proportional to the carrier-to-noise ratio, the index of modulation and the square root of the ratio of the half-bandwidth of the i-f amplifier to the bandwidth of the audio amplifier.

The effect of moderate detuning on the output signal-to-noise ratio of an a-m receiver can also be estimated. As the receiver is detuned, the signal output will be reduced in a manner corresponding to the shape of the nose of the i-f selectivity curve. The spectrum of the noise at the output of the detector is modified by this detuning process, resulting in the generation of high frequency components and the diminution of some of the noise components present before the detuning occurred. Figure 29 illustrates the change in the noise spectrum at the output of the detector caused by detuning. However, that part of the detector output spectrum to which the audio amplifier responds is not changed by the detuning process when β is large compared to ω_c, except that its amplitude may change in a manner similar to the change in signal output corresponding to the nose of the i-f selectivity curve. The output signal-to-noise ratio would therefore be changed only slightly by moderate detuning. Severe detuning will cause selective attenuation of the sidebands of the signal and extreme phase shifts, resulting in modulation distortion. When the carrier is detuned by an amount greater than the half-bandwidth of the i-f amplifier, the detector noise output will be reduced, ultimately producing the triangular detector output noise spectrum corresponding to zero signal input.

FLUCTUATION NOISE IN F-M RECEIVERS

Theoretical treatments of the transfer of signal and fluctuation noise through f-m receivers have been published by several authors. Crosby has treated the case of an ideal receiver with a perfect limiter preceding the demodulator circuits.[66] (A perfect limiter is one which removes all envelope amplitude variations from the incoming signal and noise voltages.) This analysis is valid only for carrier-to-noise ratios which are large compared

[66] Crosby, M. G., "Frequency Modulation Noise Characteristics," Proc. I.R.E., Vol. 25, pp. 472-514, April 1937

RESTRICTED

DECLASSIFIED

RESTRICTED
DECLASSIFIED

to unity. Carson and Fry have analyzed the case of an ideal receiver with and without a perfect limiter,[67] this analysis also being valid only for signal-to-noise ratios which are large compared to unity. Both of the above treatments were predicated on the use of linear rectifiers in the demodulator circuit. Algebraic difficulties arise when the carrier-to-noise ratio is near unity in the case of the linear detector. These difficulties do not appear in the case of square-law detectors.

Burgess,[68] in an analysis of signal-to-noise ratio in a-m receivers, found that both linear and square-law rectifiers provide essentially the same signal-to-noise ratio characteristics. The discriminator in an f-m receiver converts the received f-m wave into an equivalent a-m wave prior to detection. Therefore an analysis based on square-law detection in the f-m case can be expected to provide results which are very similar to those that would be obtained with a rigorous treatment based on linear detection over a wide range of carrier-to-noise ratios. Such an analysis will now be described.

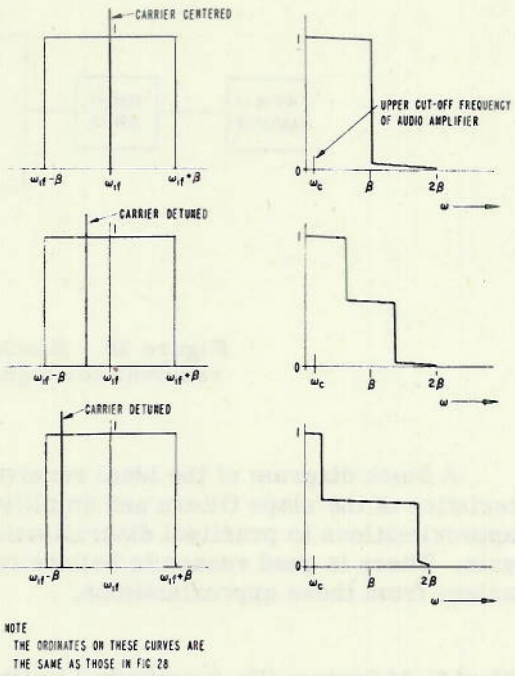


Figure 29 - Modification of total audio noise spectrum due to detuning

The analysis is predicated on an ideal receiver having rectangular i-f and audio-selectivity characteristics. The demodulator is assumed to consist of a dual-channel network fed from a common source, each channel having a linear slope filter serving as one element of a frequency discriminator. Each slope filter is terminated in a rectifier, and the output networks of these two rectifiers are connected in opposition to each other to form the detector circuit. The combined output voltage of the rectifiers is fed to the audio amplifier, which terminates in an output load. The case in which no limiter precedes the slope filters will be treated first.

It was shown above that the analysis is closely applicable to the case in which linear detectors are employed. The analysis has also been extended to show the effect of a perfect amplitude limiter on signal-to-noise ratio in f-m reception. The perfect limiter is assumed to be a device which removes all envelope amplitude variations from the composite wave resulting from the combination of a frequency-modulated signal and a random noise spectrum. A brief description of the method used for these analyses will be presented, as a complete formulation is too involved to include in this report. The method is similar to that used by other workers.[68,69]

---

[67] Carson, J. R. and Fry, T. C., "Variable Frequency Electric Circuit Theory with Application to the Theory of Frequency Modulation," Bell System Technical Journal, Vol. XVI, pp. 513-540, Oct. 1937

[68] Burgess, R. E., op. cit.

[69] Ragazzini, J. R., op. cit.

RESTRICTED
DECLASSIFIED

DECLASSIFIED

RESTRICTED

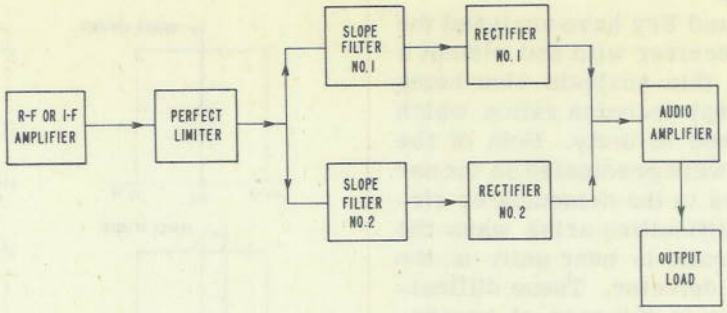


Figure 30 - Block diagram of idealized f-m receiver for signal-to-noise ratio analysis

A block diagram of the ideal receiver considered is given in Figure 30, and the characteristics of the slope filters and amplifiers are given in Figure 31. These are quite close approximations to practical discriminator characteristics, idealized to simplify the analysis. There is good reason to believe that little error in the signal-to-noise ratio values arises from these approximations.

Ideal F-M System (No Amplitude Limiting)

In the first case, where the limiter is absent (no limiting), the voltage at the output of the predetector amplifier consists of a pure frequency-modulated wave and a noise spectrum linearly combined. Its mathematical representation is as follows:

$$E_{if} = S \sin\left(\sigma_c t + \frac{\Delta\omega}{\Omega} \sin \Omega t\right) + \sum_n \frac{E_N}{\sqrt{\beta}} \sin(\sigma_n t + \phi_n), \tag{26}$$

where S is the peak amplitude of the fequency-modulated carrier, σ_c is the mean angular velocity of the carrier, $\Delta\omega$ is the peak deviation of the carrier in radians per second, and Ω is the angular velocity of the modulating wave. E_n is the rms value of the noise and hence $E_N/\sqrt{\beta}$ is the peak amplitude of a single component of noise of angular velocity σ_n. ϕ_n is the random phase angle of the nth noise component and n is a number which designates the noise component under consideration. Thus the noise is assumed to consist of a summation of a great many equal-amplitude sinusoidal components having random phase angles. The input rms signal-to-noise ratio is thus $x = S/\sqrt{2}E_N$.

The signal and noise voltages are then mathematically "passed" through the two slope filters. After the magnitudes and phases of the voltages incident on the rectifiers are found, the output voltages due to pairs of input components (at least one of which is a noise component) are computed, giving the detected alternating output noise voltages from each rectifier for each pair of input components. The rectifiers, if square-law, respond to the square of the envelope of the input wave. If the wave is of the form $\sum_\mu e_\mu \sin \omega_\mu t$, and the rectifier has a characteristic given by $e_o = \alpha$ (envelope)^2, the alternating component of the output will be

$$2\alpha \sum_{\substack{r \\ s<r}} \sum_s e_r e_s \cos(\omega_r - \omega_s)t,$$

RESTRICTED

DECLASSIFIED

DECLASSIFIED

which represents the summation of all the cross-products of the input components. The components from the two detectors having the same frequency and phase angle are subtracted. The mean square values of the net output voltages of the same frequency are then added, giving the spectral component of noise at each audio frequency. The mean square values are then summed (integrated) over the audio-amplifier bandwidth to obtain the mean squared output noise voltages. The mean squared signal output is then computed. The square root of the ratio of these quantities gives the output rms signal-to-noise ratio. The signal-to-noise ratios obtained with linear rectifiers will be essentially identical with the values obtained using square-law rectifiers, provided the peak frequency deviation is small compared to the half-bandwidth of the i-f amplifier.

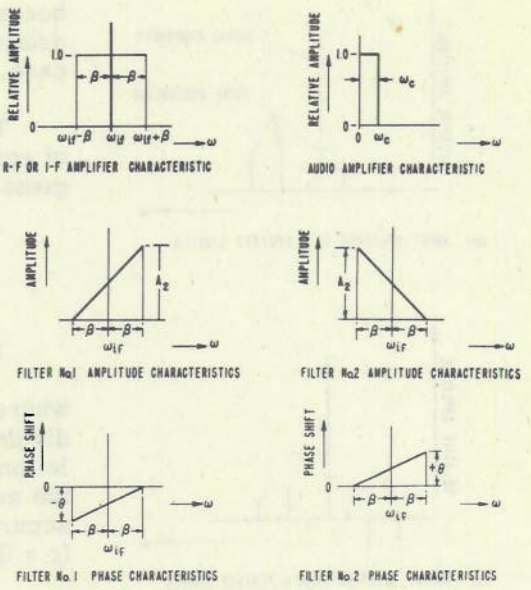


Figure 31

The following equation expresses the relation between input and output signal-to-noise ratios as derived by the above method for an f-m receiver without an amplitude limiter:

$$\left(\frac{S}{N}\right)_o^2 = \frac{2x^4d^2}{\left[\frac{2}{3} - \frac{(2-\gamma_c)^4}{24}\right] + 2x^2\gamma_c\left[4\rho^2 + 2d^2 + \frac{\gamma_c^2}{3}\right]} \quad (27)$$

(FM-No Limiter)

In the above equation

$\left(\frac{S}{N}\right)_o$ = rms output signal-to-noise ratio,

x = rms signal-to-noise (carrier-to-noise) ratio at the i-f amplifier output,

d = ratio of maximum signal-frequency deviation to half-bandwidth of i-f amplifier,

γ_c = ratio of audio bandwidth to half-bandwidth of i-f amplifier,

ρ = fractional detuning, or quotient of mean signal frequency minus center frequency of the demodulator (which is assumed to be the center frequency of the i-f amplifier) divided by the half-bandwidth of the i-f amplifier.

In this formula and those to follow, the following condition must be satisfied, $\rho + d + \gamma_c \leq 1$. This restriction involves little loss in generality when the audio bandwidth is small compared to the i-f bandwidth. The effects of detuning when it exceeds the above limitation cannot be calculated, but under such conditions the output signal is usually seriously distorted due to selective sideband attenuation in the i-f amplifier. Such distortion then

DECLASSIFIED

DECLASSIFIED

RESTRICTED

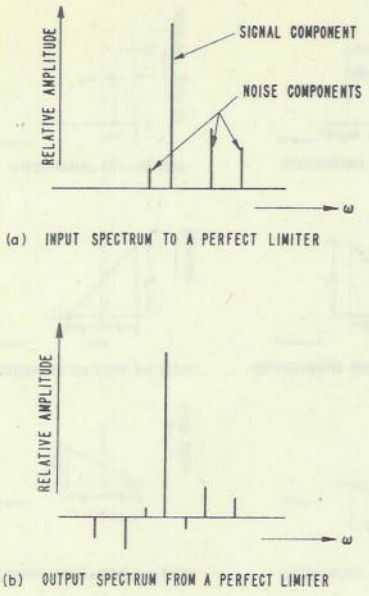


Figure 32 - Action of a perfect limiter on a combination of several sine waves (one large compared with the others)

becomes of major concern, and signal-to-noise ratio deterioration effects are of secondary interest in that case.

The quantity in the first bracket in the denominator of equation (27) may be eliminated when x is large. This gives

$$\left(\frac{S}{N}\right)_o^2 = \frac{x^2 R^2}{\gamma_c \left[4\left(\frac{\rho}{\gamma_c}\right)^2 + 2R^2 + \frac{1}{3}\right]}, \qquad (28)$$

(FM-No Limiting)

where R is the deviation ratio, or the maximum deviation divided by the audio bandwidth. (R equals mf, the modulation index, only when the modulating frequency equals the audio cutoff frequency.) When the input signal is accurately tuned to the center of the i-f characteristic ($\rho = 0$), equation (28) becomes

$$\left(\frac{S}{N}\right)_o = \frac{\sqrt{3}\, x R}{\sqrt{1 + 6R^2}} \sqrt{\frac{1}{\gamma_c}}. \qquad (29)$$

Ideal F-M System (Perfect Amplitude Limiting)

The analysis for the case when a perfect limiter precedes the dual-channel slope filters is as follows. The resultant of the combination of the frequency-modulated signal and the noise is expressed in terms of the composite envelope and phase angle. The envelope is then reduced to a constant-amplitude value (due to perfect amplitude limiting) and the spectrum of the resulting wave is found. It may be shown for the case in which the signal is large compared to the noise that each noise component is split into two spectral components, skew-symmetrical about the signal-frequency component. The skew-symmetrical arrangement indicates the 90-degree phase relationship of the f-m sideband components to the carrier. The amplitude of these noise components is halved relative to the signal by the action of the limiter. Figure 32 illustrates this process. The spectrum at the output of the limiter is passed through the slope filters, and the detection process is analyzed in the same manner as for the case with no limiter. The result of this analysis gives the following formula for the output rms signal-to-noise ratio for an f-m receiver with a perfect limiter:

$$L\left(\frac{S}{N}\right)_o = \frac{\sqrt{3}\, d x (x^2 + 1)\sqrt{\frac{1}{\gamma_c}}}{\left[12\rho^2 + 6d^2 + (x^2 + 1)^2 \gamma_c^2\right]^{1/2}}. \qquad (30)$$

This formula is more accurate for values of x large compared to unity and becomes inaccurate rapidly as x becomes smaller than 2.

RESTRICTED

DECLASSIFIED

DECLASSIFIED

Equation (30) reduces to

$$L\left(\frac{S}{N}\right)_o = \sqrt{3}\,x\,R\sqrt{\frac{1}{\gamma_c}} \qquad (31)$$

when $\rho = 0$ and x^2 is large compared to unity. Equation (31) agrees with Crosby's result for the same case. Comparing equations (29) and (31), it is seen that for large input signal-to-noise (carrier-to-noise) ratios, the improvement in output signal-to-noise ratio due to amplitude limiting is:

$$I_L = \frac{L\left(\frac{S}{N}\right)_o}{\left(\frac{S}{N}\right)_o} = \left[1 + 6R^2\right]^{1/2}. \qquad (32)$$

When $R > 2$, $I_L \cong \sqrt{6}\,R$. Carson and Fry[70] have derived a similar equation, using another method, which differs somewhat from equation (31). It is as follows:

$$I_L = \left[1 + \frac{3}{2}R^2\right]^{1/2}. \qquad (33)$$

The difference between these equations has not yet been resolved. Experimental measurements have indicated that equation (32) is more nearly correct.

Theoretical Output Signal-to-Noise Ratios (FM and AM)

The theoretical output signal-to-noise ratio obtainable with f-m and a-m receivers is presented graphically, for a practicable set of conditions, in Figure 33. Output rms signal-to-noise ratio is plotted against input rms carrier-to-noise ratio for the following cases:

Curve A - F-m receiver with a perfect limiter, using equation (30), with $d = 7.2/62.5$, $\rho = 0$, $\gamma_c = 3.0/62.5$.

Curve B - A-m receiver, using equation (24), with $m = 1$, $\beta = 62.5$ kc, $\omega_c = 3.0$ kc.

Curve C - F-m receiver without a limiter, using equation (27), with $d = 7.2/62.5$, $\rho = 0$, $\gamma_c = 3.0/62.5$.

The upper part of curve A, which is shown as a solid line, was plotted with values

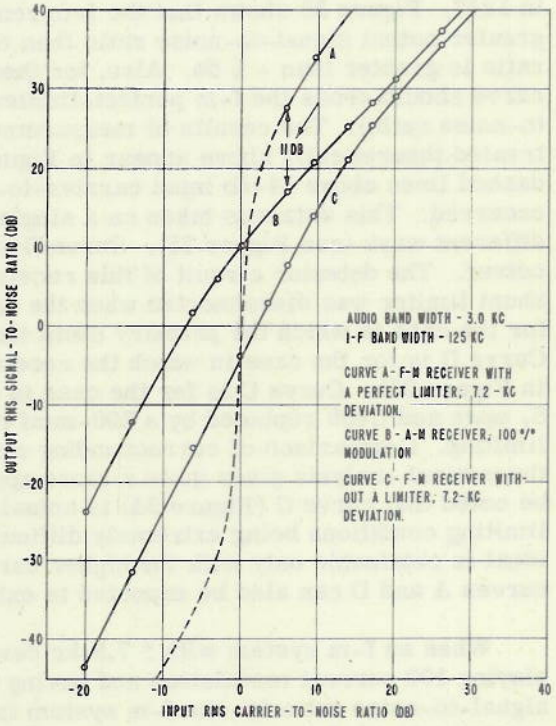


Figure 33 - Theoretical signal-to-noise ratio characteristics of ideal a-m and f-m receivers

[70] Carson, J. R. and Fry, T. C., op. cit.

DECLASSIFIED

DECLASSIFIED

derived from equation (30). The dashed portion of the curve was added by the following procedure. The dashed line was drawn as a continuation of the solid line so as to cross the no-limiting curve at zero db input signal-to-noise ratio. It was then extended through points which were as much below the no-limiting curve C as the solid portion of A was above C. The reason for this construction is as follows: If the input voltage to a perfect limiter consists of two sinusoids, the output of the limiter contains the same two sinusoids except that the weaker one is suppressed relative to the stronger one. If the carrier signal is considered equivalent to a sine wave and the noise as equivalent to another sine wave, the carrier can be expected to suppress the noise (due to limiting action) when the carrier-to-noise ratio is large. This is seen to be the case, as indicated by the upper portion of curve A. That is, for high input signal-to-noise (carrier-to-noise) ratios, the output signal-to-noise ratio is greater when amplitude limiting takes place than when no limiting is present. When the stronger input wave is the noise, it might be expected that the signal would be suppressed due to the action of the limiter. The dashed part of curve A was drawn accordingly. It is emphasized that the dashed portion of the curve constitutes an estimate. However, it is usually necessary to find the crossover point between the a-m and f-m curves by some procedure, even though approximate, when a-m to f-m comparisons are desired.

Curve A may be made to approach curve C asymptotically if the threshold of limiter action is properly chosen. Thus the limiter could be so designed that noise alone will not cause limiting while the addition of signal will cause the limiting action to commence.

The bandwidths and deviations used for calculating the graphs of Figure 33 are those employed with the RDZ and X-RDZ-2 receivers used in the a-m/f-m tests made at NRL in 1947. Figure 33 shows that the f-m receiver with a perfect limiter should produce a greater output signal-to-noise ratio than the a-m receiver when the input carrier-to-noise ratio is greater than + 1 db. Also, for theoretically ideal receivers of this type, the a-m curve should cross the f-m perfect-limiter curve at approximately + 11 db (output signal-to-noise ratio). The results of measurements made under conditions similar to the case treated theoretically above appear in Figure 34. The curves of this figure are shown as dashed lines above 14-db input carrier-to-noise ratio, as above this point audio overload occurred. This data was taken on a single receiver whose detector was modified four different ways (see Figure 35). Curve D pertains to an unmodified Model X-RDZ-2 receiver. The detector circuit of this receiver is shown in Figure 35a. The added primary shunt limiter was disconnected when the data of curve D was taken (S_1 open). Curve A is for the case in which the primary shunt limiter was connected in the circuit (S_1 closed). Curve B is for the case in which the receiver was modified for a-m operation, as shown in Figure 35b. Curve C is for the case in which the circuit of Figure 35a was used with S_1 open and C269 replaced by a 300-mmf capacitor, to approximate a condition of no limiting. Comparison of corresponding curves in Figures 33 and 34 indicates that the theoretical analysis gives quite a close approximation to the actual conditions. It should be noted that curve C (Figure 34) is actually for the condition of partial limiting, true no-limiting conditions being extremely difficult to obtain. Likewise, limiting approaching the ideal is obtainable only with the higher carrier levels in most practical equipment, and curves A and D can also be expected to exhibit the effects of partial limiting.

When an f-m system with ± 7.5-kc deviation is compared with an a-m system employing 100 percent modulation and having the same audio and i-f bandwidths, the output signal-to-noise ratio for the f-m system is greater than that of the a-m system by approximately 14 db, provided the input carrier-to-noise ratio is the same for each case and is large compared to unity. With 1000-cps modulation, the bandwidth of the spectrum of the f-m input signal is approximately 17 kc, while in the case of the a-m system it is

[redacted] RESTRICTED

DECLASSIFIED

DECLASSIFIED

RESTRICTED

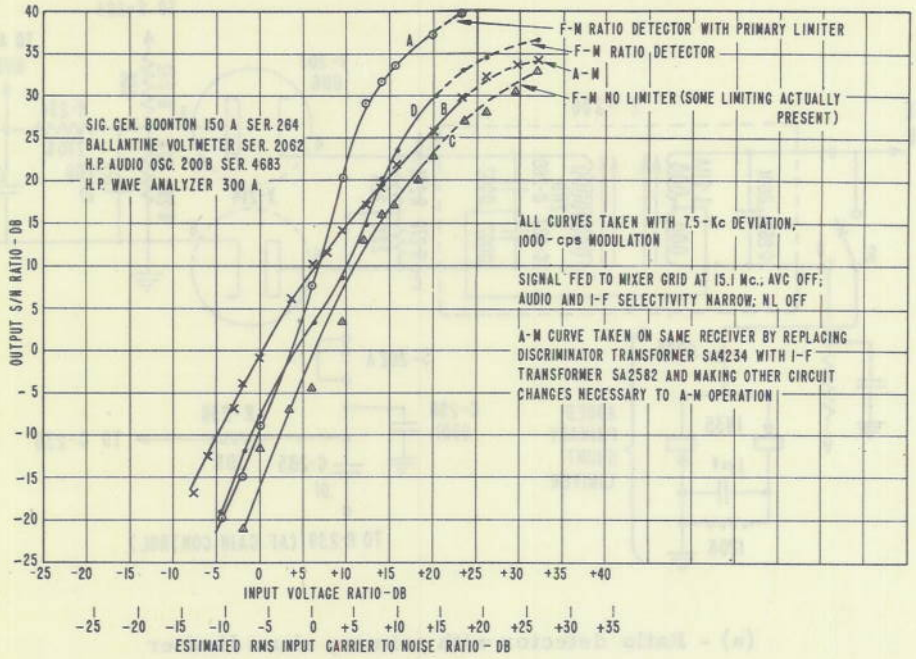


Figure 34 - Input voltage ratio vs. output S/N ratio - db
X-RDZ-2 Ser. 3239

only about 2 kc. If the deviation of the f-m system were reduced so that the f-m signal spectrum occupied the same maximum bandwidth as the a-m signal (6 kc), the difference in output signal-to-noise ratio would be greatly reduced. For the case of ± 3-kc deviation, the difference in output signal-to-noise ratio for high input levels would then become approximately 6.5 db in favor of FM. This degree of improvement is achievable only under ideal conditions and for large input carrier-to-noise ratios. It should be noted that for input carrier-to-noise ratios which are small compared to unity, the a-m system in all cases will provide the greater output signal-to-noise ratio.

Detuning Effects

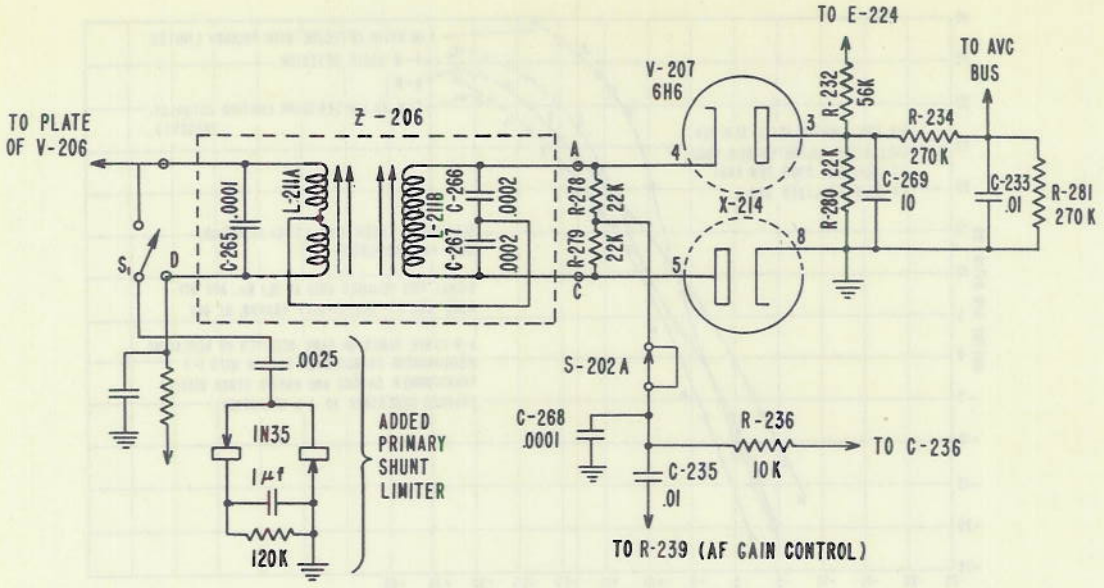
The effect of detuning the f-m signal so that its unmodulated carrier frequency does not correspond to the center frequency of the detector (discriminator) may be computed from equations (30) and (27), for the cases of perfect limiting and no limiting, respectively. The output rms signal-to-noise ratio is reduced when such detuning occurs. The extent of the deterioriation of output signal-to-noise ratio due to detuning is plotted in Figures 36 and 37. Inspection of these curves indicates that the amount of such deterioriation with an f-m receiver having a perfect limiter is greater for the lower values of carrier-to-noise ratio (weak signals), and becomes less serious when the carrier-to-noise ratio is large (strong signals). In the case of an f-m receiver with no limiter, the deterioriation with detuning is less pronounced than in the case of the receiver with perfect limiting, but is greater for the larger values of carrier-to-noise ratio.

The effect of detuning on the advantage of FM over AM with regard to output signal-to-noise ratio may be ascertained by reference to the data designated by arrows on Figures 33 and 36.

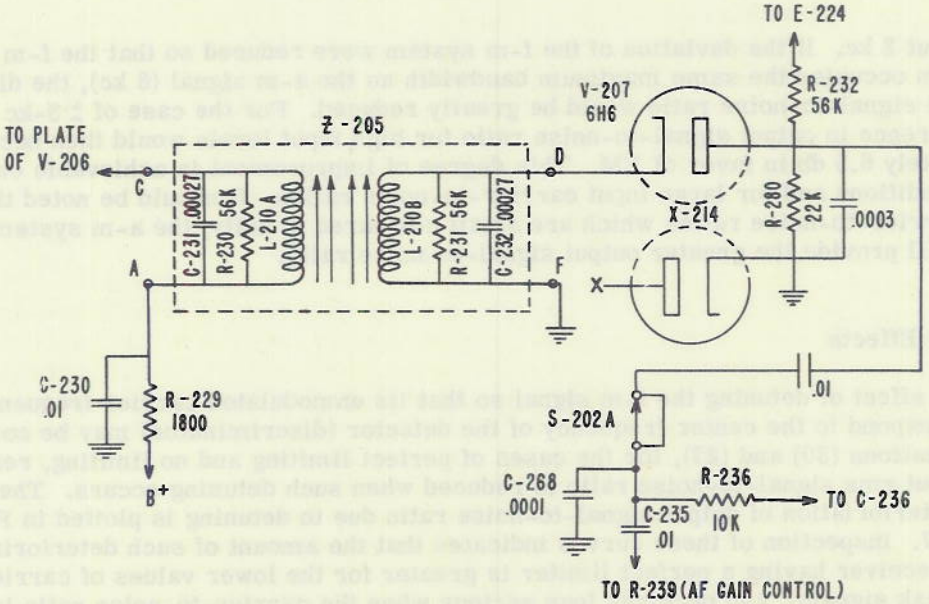
RESTRICTED

DECLASSIFIED

DECLASSIFIED
RESTRICTED



(a) - Ratio detector with primary shunt limiter



(b) - X-RDZ-2 modified for a-m measurements

Figure 35

RESTRICTED
DECLASSIFIED

DECLASSIFIED

It can be seen in Figure 33 that for an input carrier-to-noise ratio of + 6 db, the f-m system with a perfect limiter has an advantage in output signal-to-noise ratio of approximately + 11 db over an a-m system under conditions of perfect centering (no detuning). Figure 36 indicates that for the same carrier-to-noise ratio (+ 6 db), detuning to the extent of approximately 23 kc results in a deterioriation of 11 db in output signal-to-noise ratio for the f-m system. Such an amount of detuning can be quite common in a VHF or UHF operating system. Hence it may be concluded that consideration of detuning as a factor in such systems results in a reduction of their potential advantage over a-m systems, since in the latter negligible reduction in output signal-to-noise ratio results from detuning.

PRE-EMPHASIS AND DE-EMPHASIS

An improvement in signal-to-noise ratio in both a-m and f-m systems can be achieved by the use of pre-emphasis and de-emphasis techniques.[71] A simple circuit, called the pre-emphasis network, can be included in the audio system of a transmitter to provide greater transmission gain for high audio frequencies than for low audio frequencies. At the receiver, usually in the detector output circuit, another simple circuit called the de-emphasis network is then included. The de-emphasis network attenuates the high audio components of the modulating wave to a greater degree than the low audio components, in such a way as to exactly equalize the effect of the pre-emphasis network at the transmitter. From a theoretical point of view, the improvement in signal-to-noise ratio that can thus be realized is greater in f-m systems than in a-m systems.

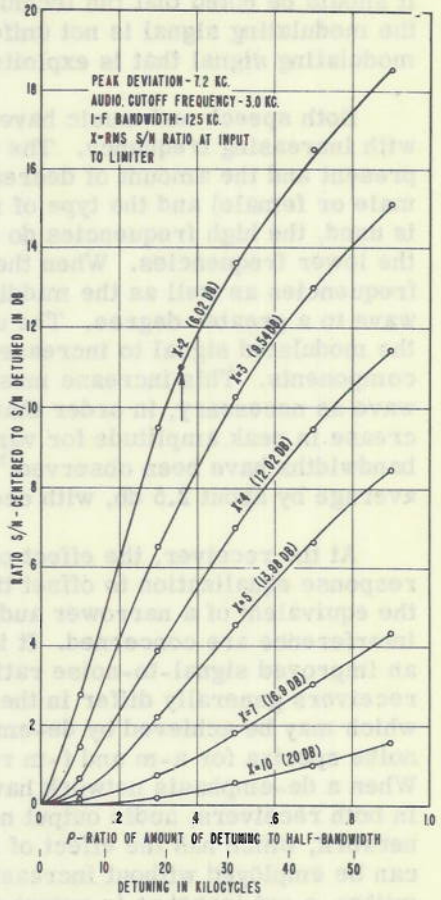


Figure 36 - Effect of detuning in an f-m receiver with a perfect limiter

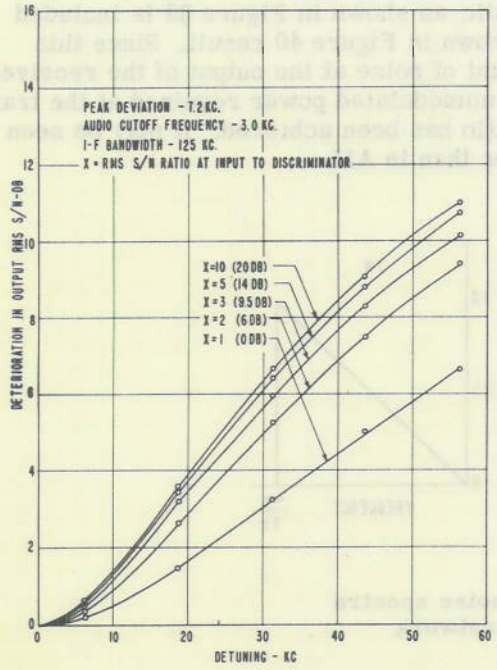


Figure 37 - Theoretical effect of detuning in an ideal f-m receiver without a limiter

[71] Hathaway, J. L., "High Frequency Pre-emphasis," Electronics, Vol. 12, pp. 29-32, Nov. 1939

DECLASSIFIED

DECLASSIFIED

It should be noted that the technique described above is useful only when the spectrum of the modulating signal is not uniform, and it is the nonuniformity of the spectrum of the modulating signal that is exploited to obtain the signal-to-noise ratio improvement.

Both speech and music have power spectra that, in general, decrease in magnitude with increasing frequency. The frequency ranges in which the maximum audio power is present and the amount of decrease at high frequencies depend on the type of speech (e.g., male or female) and the type of music (organ, orchestra, etc.). When no pre-emphasis is used, the high frequencies do not modulate the carrier wave to as great a degree as the lower frequencies. When the proper degree of pre-emphasis is employed, the high frequencies as well as the middle-range frequencies can be made to modulate the carrier wave to a greater degree. The use of a pre-emphasis network causes the peak value of the modulated signal to increase, because of the increased amplitudes of the high-frequency components. This increase must be offset by reducing the level of the entire modulating wave as necessary, in order that overmodulation may not take place. The amount of increase in peak amplitude for various types of program material having 5-kc and 12-kc bandwidths have been observed.[72] The effective modulation level was greater on the average by about 2.5 db, with occasional short-duration bursts of 4.5 db.

At the receiver, the effect of the de-emphasis network is to provide frequency-response equalization to offset the pre-emphasis at the transmitter, and also to provide the equivalent of a narrower audio bandwidth as far as random noise and other types of interference are concerned. It is this effective reduction in audio bandwidth that causes an improved signal-to-noise ratio with single-tone measurements. Since f-m and a-m receivers generally differ in the shape of their audio noise spectra, the improvement which may be achieved by de-emphasis also differs. Idealized strong-signal audio-output noise spectra for a-m and f-m receivers without de-emphasis are shown in Figure 38. When a de-emphasis network having a characteristic as shown in Figure 39 is included in both receivers, audio output noise spectra as shown in Figure 40 result. Since this network, which has the effect of reducing the amount of noise at the output of the receivers, can be employed without increasing the amount of unmodulated power required at the transmitter, a net increase in output signal-to-noise ratio has been achieved. It may be seen that the amount of noise reduction in FM is greater than in AM.

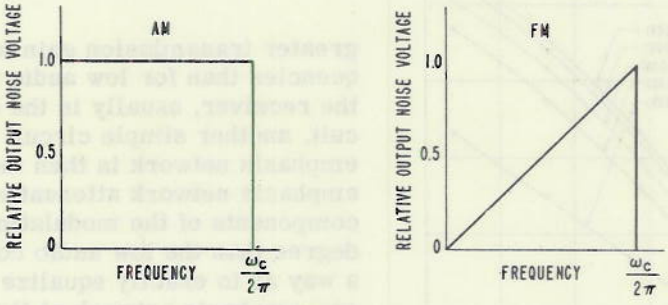


Figure 38 - Audio output noise spectra without de-emphasis network

[72] Crosby, M. G., "The Service Range of Frequency Modulation," RCA Review, pp. 349-371, Jan. 1940

DECLASSIFIED [redacted] RESTRICTED

DECLASSIFIED

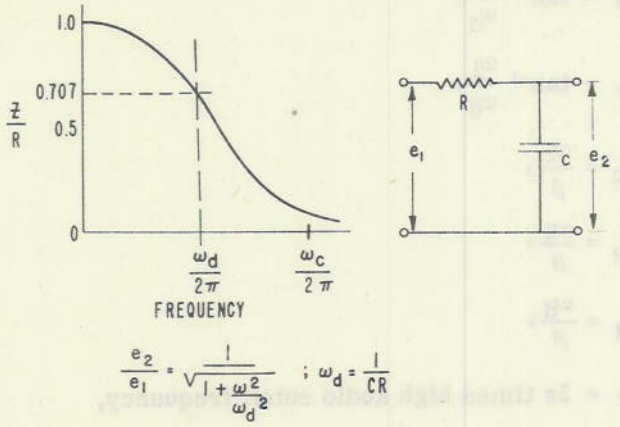


Figure 39 - De-emphasis network characteristics

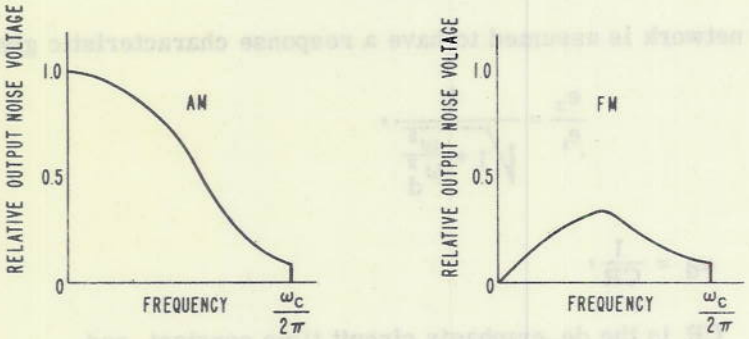


Figure 40 - Audio output noise spectra with de-emphasis

The following formulas have been derived by the author for computing the improvement in audio output signal-to-random-noise ratio that can theoretically be achieved in idealized a-m and f-m systems by using de-emphasis:

$$I_{AM}(db) = 10 \log_{10} \frac{\left\{2 + 2x^2(2+m^2)\right\}(\gamma_c - \gamma_l) - \frac{1}{2}(\gamma_c^2 - \gamma_l^2)}{\gamma_d^2\left[\frac{\left\{2 + 2x^2(2+m^2)\right\}(\theta_2 - \theta_1)}{\gamma_d} - \frac{1}{2}\log_e \frac{\gamma_d^2 + \gamma_c^2}{\gamma_d^2 + \gamma_l^2}\right]}; \quad (34)$$

$$I_{FM}(db) = 10 \log_{10} \frac{\frac{2}{3}(\gamma_c^3 - \gamma_l^3)(x^2+1)^2 + (\gamma_c - \gamma_l)(8\rho^2 + 4d^2)}{\gamma_d\left[2(x^2+1)^2\gamma_d\left\{(\gamma_c - \gamma_l) - \gamma_d(\theta_1 - \theta_2) + (8\rho^2 + 4d^2)(\theta_1 - \theta_2)\right\}\right]}, \quad (35)$$

RESTRICTED DECLASSIFIED

DECLASSIFIED

where $\theta_1 = \tan^{-1}\frac{\omega_c}{\omega_d}$,

$\theta_2 = \tan^{-1}\frac{\omega_l}{\omega_d}$,

$\gamma_c = \frac{\omega_c}{\beta}$,

$\gamma_e = \frac{\omega_l}{\beta}$,

$\gamma_d = \frac{\omega_d}{\beta}$,

ω_c = 2π times high audio cutoff frequency,

ω_l = 2π times low audio cutoff frequency,

ω_d = 2π times frequency at which response of de-emphasis network is 0.707 of its maximum value, and

β = pre-detector half bandwidth.

The de-emphasis network is assumed to have a response characteristic given by (see Figure 39):

$$\frac{e_2}{e_1} = \frac{1}{\sqrt{1+\frac{\omega^2}{\omega_d^2}}},$$

where $\omega_d = \frac{1}{CR}$,

CR is the de-emphasis circuit time constant, and

ω is 2π times any frequency in the audio pass band.

The above equations were derived for the same idealized conditions as assumed for equations (24), (27), and (30). The equation for I_{FM} should become inaccurate for values of x less than 2. This equation applies to the case of an ideal f-m receiver having a perfect limiter. It is assumed that the audio amplifier without de-emphasis has a rectangular response characteristic with a low-frequency cutoff angular velocity of ω_l and a high-frequency cutoff angular velocity of ω_c, as indicated above.

The following table gives values for the improvement in decibels due to de-emphasis, computed from the above formulas. No allowance has been made for the decrease in average modulation level which would be required at the transmitter because of pre-emphasis.

It is seen from this tabulation that the improvement in signal-to-noise ratio which can be achieved in the case of FM is superior, in general, to that for AM as the carrier-to-noise ratio x increases above unity. The improvement with AM is independent of the input signal-to-noise ratio for the values shown in the table. The improvement with FM becomes smaller for lower values of input signal-to-noise ratio. With AM, the improvement due to

DECLASSIFIED

DECLASSIFIED

Input Rms Carrier-to-Noise Ratio		Improvement due to De-Emphasis (Decibels)				
		$\omega_l = 0, \frac{\omega_c}{2\pi} = 3000$ cps; $\frac{\omega_d}{2\pi} = 1300$ cps		$\frac{\omega_l}{2\pi} = 300$ cps, $\frac{\omega_c}{2\pi} = 3000$ cps; $\frac{\omega_d}{2\pi} = 1135$ cps		
x	x (db)	AM	FM	AM	FM	FM (10-kc Detuning)
1.4	+ 3.0	3.0	3.4	4.0	4.4	—
2.0	+ 6.0	3.0	—	4.0	4.9	4.25
3.0	+ 9.5	3.0	4.8	4.0	5.7	—
5.0	+14.0	3.0	5.4	4.0	6.2	—

de-emphasis is not affected by small amounts of detuning, but with FM the improvement is reduced when the receiver is detuned. It may be concluded that the improvement in signal-to-noise ratio that can be achieved by de-emphasis is small for the case of narrow audio bandwidths. The use of pre- and de-emphasis in systems having wide post-detector bandwidths would produce correspondingly greater improvements.

Because of the rapid change in output signal-to-noise ratio in an f-m system for carrier-to-noise ratio values of 0 db to about + 10 db, an improvement of 4 decibels in output signal-to-noise ratio will be equivalent to perhaps 1-db increase in transmitter power without pre-emphasis. In the case of an a-m system, however, 4 decibels signal-to-noise ratio increase would be equivalent to about 4 db transmitter power increase above 0-db carrier-to-noise ratio. Emphasis networks would thus benefit an a-m system much more than they would an equivalent f-m system in the practical sense that they cause an improvement equivalent to an increase in transmitter power. Both systems would also benefit because of improved receiver performance with regard to transient noise disturbances.

IMPULSE NOISE INTERFERENCE

Impulse noise is a form of interference caused by pulses of energy which generally enter a communication receiving system from external sources. The interfering impulses usually have their origin in atmospheric disturbances (e. g., lightning), rotating machinery, radar equipment, electric razors, ignition systems, etc. The impulses may occur at a random or steady rate. The amplitude and shape of the impulses may be the same or different from impulse to impulse. Fluctuation noise may be considered as a special type of impulse noise in which the impulses occur so frequently that they overlap each other. The spectral composition of impulses is such that the components exhibit a general decrease in amplitude with increasing frequency. This results in a decrease in the average level of impulse noise with increasing radio frequency. Radar interference, however, may occur with high amplitudes at any frequency, since it has the spectrum of a pulse-modulated carrier. Impulse interference in the 200-400 Mc band is caused mainly by impulses whose duration is very short (of the order of one microsecond) and whose source is near the receiver. Because of its nature, impulse noise is usually treated in terms of peak amplitudes rather than rms values.

When an impulse appears at the input of either an f-m or an a-m receiver of the same bandwidth in the absence of a carrier, a transient wave train or oscillation appears at the

DECLASSIFIED

DECLASSIFIED

output of the i-f amplifier. The characteristics of this output oscillation are primarily a function of the properties of the receiver, provided that the duration of the input impulse is short compared to the time constants of the selective circuits. The envelope of the output wave train has a peak value which is directly proportional to the bandwidth of the receiver, and a duration which is inversely proportional to the bandwidth. The frequency of the oscillation is the center frequency of the predetection amplifier, if that amplifier has symmetrical selectivity characteristics. The shape of the envelope of the output wave train depends on the shape of the selectivity curve of the amplifier. The shapes of output impulses for various types of i-f selectivity characteristics have been observed and discussed by Landon.[73]

The response of selective amplifiers to short pulses may be analyzed as follows. When an impulse enters a receiver, the input tuned circuit "rings" at its resonant frequency, i. e., an oscillation builds up and decays at a rate determined by the time constant of the circuit. This wave train is amplified selectively by the following stages of the receiver. The short impulse is thus converted to a longer wave train, because of the time required to store energy in the circuit reactances and the time subsequently necessary to dissipate the stored energy in the circuit resistances. During amplification of the wave train, limiting may take place if the input impulse is of sufficient amplitude to cause any of the amplifier tubes to be driven beyond the region of linear operation. The effect of limiting is to extend the duration of the output wave train.

Grid current may flow in some of the tubes during part of the impulse wave train, causing an increase in grid bias if the grid circuit resistance is large. The grid circuit time constants of the amplifier stages must be kept as short as possible to reduce the time of desensitization caused by the flow of grid current. Changes in the grid bias of an amplifier may result in a change in the input capacity of the tube. This would cause the circuit attached to the tube to be detuned during the impulse, thus changing the frequency of the output pulse wave. It will be shown that this will cause increased interference in f-m receivers. Such detuning may be minimized by the use of high capacity (low L/C ratio) tuned circuits which will make the change in tube input capacity small compared with the total tuning capacity. The impulses at the output may have sufficient energy to cause operation of AVC circuits. This can also cause desensitization and detuning of the receiver.

The response of the receiver to impulse energy may also be analyzed by consideration of the spectrum of the impulse. That part of the spectrum of an impulse which falls within the selectivity curve of the receiver contains a number of components of the same phase, equally spaced in frequency and essentially equal in amplitude. The average frequency for these is the center frequency of the receiver. These components add in phase during the impulse and hence the peak value is proportional to the number of components accepted by the receiver. Thus the peak amplitude of the output impulse is directly proportional to the bandwidth of the receiver.[74]

When a carrier and impulse are present simultaneously, the noise wave train and the carrier are additive. If the noise amplitude at the output is much greater than the carrier amplitude, the output during the pulse is essentially the same as though the carrier were not present. When the noise amplitude is of the same order of magnitude as the carrier amplitude, the envelope amplitude of the output wave during the pulse depends on the phase relation between the noise and signal. The envelope amplitude of the carrier plus impulse

---

[73] Landon, V. D., "Impulse Noise in FM Reception," Electronics, Vol. 14, p. 26, Feb. 1941

[74] Crosby, M. G., op. cit., (footnote 66 of this report)

DECLASSIFIED

DECLASSIFIED

may vary from impulse to impulse, since the occurrence rate of the noise wave train will seldom have a fixed integral relationship to the signal. The instantaneous frequency of the output wave during the impulse will have some value between that of the signal carrier and the noise wave train, depending on their relative amplitudes. Such waveforms have been observed and reported by Landon.[75]

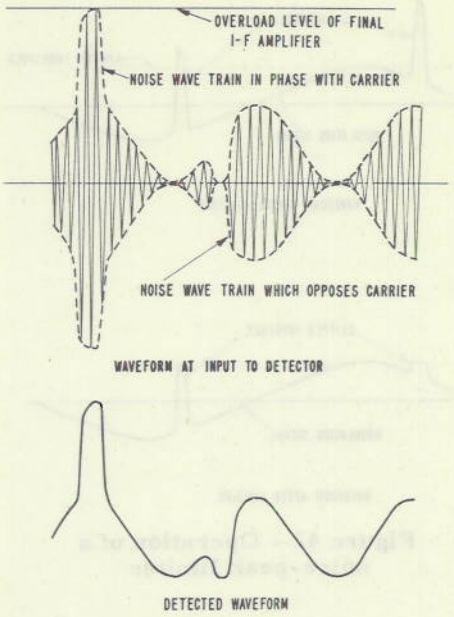


Figure 41 - Detection of impulse noise in a-m receiver

IMPULSE NOISE IN A-M RECEIVERS

The detector in the a-m receiver produces an output proportional to the envelope of the waveform which is applied to it. It is evident that slight detuning of the preceding tuned circuits or slight detuning of the signal from the center of the selectivity curve of the receiver would normally have little effect on the amplitude of the envelope of the output noise impulse. Accordingly one can expect little change in interference due to such detuning. (This is not the case in f-m reception, which will be discussed later.) The a-m detector output voltage which is fed to the audio amplifier consists of the desired modulation, plus deviations produced by those impulses which add to the modulated carrier and produce a change in the envelope. The amplitude of the detected impulse is limited in one direction by the saturation characteristics of the final i-f amplifier stages and in the other direction by the detector cutoff voltage. This is illustrated in Figure 41, which shows the envelope of the waveform of the input to the detector and the resulting detected output.

If the detector load circuit has a bandwidth equal to the half-bandwidth of the i-f amplifier, the output will be a very close reproduction of the envelope of the input waveform. Those components of this waveform at frequencies higher than the cutoff frequency of the audio amplifier will not be present in the output of the receiver. The audio amplifier is usually designed to provide a bandwidth sufficient to reproduce the intelligence waveform to a desired degree. Thus if a wide-band i-f amplifier is followed by a narrow-band audio amplifier, the noise impulses which are detected will be rounded off, decreased in amplitude, and lengthened, due to audio selectivity. The selectivity of the audio amplifier thereby assists in modifying the interference. The deterioration of intelligibility at the output of the receiver due to impulse interference is proportional to the energy of the output pulses in relation to the signal intelligence and thus depends on the amplitude, duration and repetition rate of the output pulses.

Noise-Peak Limiter

The impulse energy which reaches the output of the receiver may be reduced by the use of a "noise-peak limiter," a device which operates to remove the peaks of impulses which exceed some predetermined amplitude of the modulation. The device is usually

---

[75] Landon, V. D., op. cit., (footnote 73, of this report)

DECLASSIFIED

DECLASSIFIED
RESTRICTED

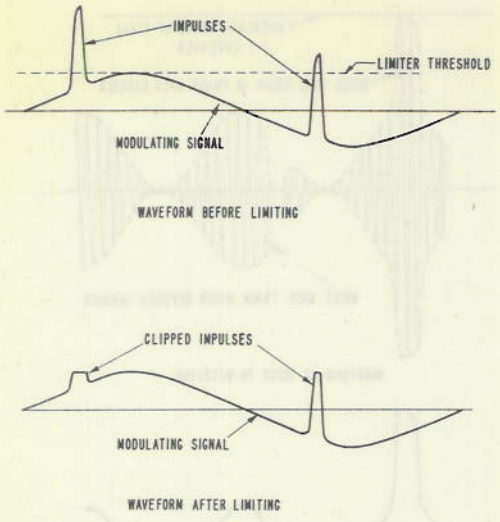


Figure 42 - Operation of a noise-peak limiter

located in the final detector load circuit, as it generally requires a relatively large signal for proper operation. Limiters which operate in circuits preceding the final detector are sometimes used in a-m receivers for telegraph signals. The operation of several types of noise peak limiters used in a-m receivers has been treated.[76]

Figure 42 gives an idealized picture of the operation of a simple noise-peak limiter. The peaks of the noise impulses which are greater than a predetermined threshold level are removed by the limiter. It should be noted that most peak limiters would have little effect on output impulses which are less than the average signal output. Limiters of this type are most effective in reducing impulse interference if they are located so that they may operate on the peaks of the impulses before they are broadened by the action of selective circuits, that is, peak limiting, to be most effective, should precede frequency selectivity.[76,77] For this reason, audio-type peak limiters should be located between the final detector and the audio amplifier. It is also important to keep the time constant of the detector load circuit as short as possible, to prevent lengthening of the impulse prior to the limiter.

In summary, detuning will have little effect on impulse noise interference in a-m reception. By the use of simple amplitude limiters, the peak impulse noise can be limited to the average value of the detected modulation. These limiters offer little improvement when the detected noise impulse is small compared to the detected modulation.

IMPULSE NOISE IN F-M RECEIVERS

The performance of f-m receivers in the presence of impulse noise differs somewhat from that of a-m receivers. The waveform of the impulse at the input to the demodulator is the same as that in an a-m receiver of the same bandwidth when no signal is present. When a signal is present, the effect of the impulse is to change the instantaneous frequency and amplitude of the signal wave during the time of occurrence of the impulse prior to the detector. Upward modulation of the signal wave by the noise may be removed by the amplitude limiter generally employed in f-m receivers, but downward amplitude modulation may be only partially eliminated. Also, the phase shifts produced by the impulse are still present after limiting, and the output of the f-m receiver will thus contain impulse energy.

Landon[78] has shown that the form of the output impulse is determined largely by the characteristics of the frequency discriminator and the phase relation between the pulse

---

[76] Toth, E., "Noise and Output Limiters," Electronics, Vol. 19, Nov. and Dec. 1946, also "Amplitude Limiters," NRL Report R-3375 (Restricted), Oct. 1948

[77] Wald, M., "Noise Suppression by Means of Amplitude Limiters," Wireless Engineer, Vol. 17, pp. 432-438, Oct. 1940

[78] Landon, V. D., op. cit., (footnote 73 of this report)

RESTRICTED
DECLASSIFIED

DECLASSIFIED

and the signal. Optimum performance is obtained when the over-all selectivity curve of the receiver (including the discriminator) is symmetrical about the desired signal. He emphasizes that this characteristic must be accurately symmetrical for a wide range of input-signal amplitudes to obtain optimum performance in the presence of impulse noise. The output circuits of the two halves of the usual discriminator must also have identical response characteristics.

Crosby[79] has shown that when the peak signal-to-noise ratio at the output of the i-f amplifier is large compared to unity, an ideal f-m receiver having a perfect limiter reduces the effect of impulse noise to a greater degree than an a-m receiver having the same selectivity. The improvement of an f-m receiver over an a-m receiver is given by the following formula:

$$\frac{\frac{S_f}{N_f}}{\frac{S_a}{N_a}} = 2\frac{F_d}{F_a}, \tag{36}$$

where S_f/N_f is the output peak-signal to peak-noise ratio in an f-m receiver, S_a/N_a is the output peak-signal to peak-noise ratio in an a-m receiver having the same selectivity as the f-m receiver, F_d is the peak deviation of the f-m signal (assumed to be equal to half the i-f bandwidth of the f-m receiver), in cycles per second, and F_a is the highest frequency of response of the audio amplifier.

For the case when the peak signal-to-noise ratio at the output of the i-f amplifier is small, it has been shown that the maximum value of the noise output from the discriminator is limited.[80] This arises from the fact that the instantaneous frequency of the wave applied to the discriminator cannot lie outside the pass band of the i-f amplifier. This effect is referred to as "frequency limiting," which may be compared with amplitude limiting as discussed previously. If the deviation of the f-m signal is equal to the half-bandwidth of the i-f amplifier, the peak noise output cannot exceed the peak signal output if the f-m receiver contains an effective amplitude limiter. Thus for this particular case, the limiter plus balanced discriminator produces an effect on the impulses similar to that produced by an amplitude limiter following the detector in an a-m receiver.

Smith and Bradley[81] have shown that the peak amplitude of the output impulse is on the average less than the peak signal output for this condition (peak deviation = half predetector bandwidth). This arises from the fact that the phase of the impulse wave train is random with respect to that of the signal and because the peak output impulse amplitude depends on this phase relation, as was demonstrated by Landon. The peak amplitude is also dependent on the instantaneous deviation of the signal from the center of the selectivity characteristic of the receiver at the time the impulse occurs. For a system having 75-kc deviation for maximum modulation and a de-emphasis time constant of 100 microseconds, the peak amplitude of the generated noise signal is normally about 13 percent of

---

[79] Crosby, M. G., op. cit., (footnote 66 of this report)

[80] Crosby, M. G., ibid.

[81] Smith, D. B. and Bradley, W. E., "The Theory of Impulse Noise in Ideal Frequency Modulation Receivers," Proc. I.R.E., Vol. 34, pp. 743-751, Oct. 1946

DECLASSIFIED

DECLASSIFIED

the program signal. When the peak signal deviation is less than half the i-f bandwidth, the output noise peaks can rise above the peaks of the signal output if the signal is detuned. In the latter case, a noise-peak limiter such as used in a-m receivers would serve to reduce the interference.

The comparative impulse noise performance of a-m and f-m receivers like the Models RDZ and X-RDZ-2 may be discussed according to the general treatment of impulse noise given above. When the peak impulse noise is small compared with the peak signal at the output of the i-f amplifiers, the f-m system with a perfect limiter and ideal discriminator should produce a greater output signal-to-impulse-noise ratio than the a-m receiver with an amplitude limiter in the detector load circuit. The relative impulse noise performance of a-m and f-m receivers is given by equation (36), provided that the response characteristics of the two receivers are identical, and the peak frequency deviation equals the half i-f bandwidth in the f-m case. The noise-peak limiter in the a-m receiver would have little effect on the noise in this case, since the noise impulses at the detector are small compared with the peak signal at the output of the detector.

When the peak noise at the output of the i-f amplifier is large compared with the peak signal at this point, the performance of the a-m receiver should be found superior for the case when the f-m peak deviation is substantially less than the half-bandwidth of the i-f amplifier. If the peak deviation of the f-m signal is approximately equal to the half-bandwidth of the i-f amplifier, the performance of the two receivers would be quite similar. It should be noted that if the impulses are sufficiently strong and occur often enough, they may cause the generation of a large AVC voltage which will tend to reduce the gain of a receiver and thus reduce the signal output. Hence, when a receiver is operating with strong impulse interference having a high rate of occurrence, the AVC networks should be switched out and manual gain adjustments used.

Generally speaking, the optimum performance of f-m receivers in the presence of impulse interference is obtained only when the receiver has a symmetrical selectivity characteristic with the signal centered thereon, an effective predetector amplitude limiter, and when the peak frequency deviation of the signal is approximately equal to the half-bandwidth of the i-f amplifier. When these conditions are not fulfilled sufficiently, the a-m receiver having an effective postdetector noise peak limiter will give superior performance. F-m receivers are generally superior to a-m receivers when the peak impulse noise is small compared with the peak signal at the output of the i-f amplifier.

COMMON-CHANNEL INTERFERENCE IN A-M AND F-M SYSTEMS

When an interfering a-m wave and a desired signal are received simultaneously in the same channel (assuming a receiver with a linear-law detector), the desired output is both masked and reduced owing to the presence of the interfering signal. The masking is caused by: (1) the generation of a "heterodyne" or "beat-note" having a fundamental frequency equal to the difference between the two carrier frequencies and with a considerable harmonic content at or near carrier level equality; (2) the presence of the modulation of the undesired signal in the output; (3) the generation of harmonics of both the undesired and desired modulation components; and (4) the presence of "cross-talk" consisting of combination tones of the components of both modulations and their harmonics. The reduction in desired signal output due to the presence of the undesired signal is sometimes called "capture effect" or "modulation suppression" and results from the interaction of the two signals in forming their resultant envelope. The magnitude of the masking components and the degree of reduction of the desired signal output are both dependent on the ratio of the undesired to desired input signal amplitudes.

DECLASSIFIED [redacted]

DECLASSIFIED

In the case of f-m reception, masking and capture effects caused by common-channel interference also occur, but the nature and magnitude of these effects are quite different from the a-m case. Many theoretical treatments of two-signal common-channel a-m and f-m interference have appeared in the literature.[82,83,84,85,86,87,88] In the following paragraphs, a qualitative discussion of common-channel interference effects in f-m and a-m systems will be presented. The behavior of linear and square-law detectors with two input signals will be described briefly to form a basis for the discussion.

Linear and Square-Law Detectors with Two-Signal Input

Two a-m waves having different carrier frequencies may be represented by their spectra, as shown in Figure 43. The carriers are designated by c and c′, the upper sidebands by u and u′, and the lower sidebands by l and l' respectively. All these components are assumed to fall within the predetector bandwidth of the receiver, as shown. They add to form the input voltage to the detector of the receiver. The detector produces an output voltage which is proportional to the envelope of the input voltage in the case of a half-wave linear detector, and is proportional to the square of the envelope of the input voltage in the case of the half-wave square-law detector. Thus the waveform and the spectrum of the detector output voltage are determined by the envelope of the detector input voltage.

It has been shown[89] that the square of the envelope of a group of waves contains components having all possible difference frequencies as well as a steady or d-c component. The difference frequency components which lie within the audio amplifier or postdetector bandwidth are of prime interest, since they include the desired modulating signal and may contain the heterodyne beat-note, as well as harmonics and cross-talk components which appear at the reproducing device as interference. In the case of square-law detection, the amplitudes of these difference frequency components are proportional to the product of the amplitudes of the two input frequency components which give rise to the particular

---

[82] Aiken, C. B., "The Detection of Two Modulated Waves which Differ Slightly in Carrier Frequency," Proc. I.R.E., Vol. 19, pp. 120-137, Jan. 1931, and Vol. 20, pp. 569-578, March 1932

[83] Aiken, C. B., "Theory of the Detection of Two Modulated Waves by a Linear Rectifier," Proc. I.R.E., Vol. 21, pp. 601-629, April 1933

[84] Moullin, E. B., "The Detection by a Straight Line Rectifier of Modulated and Heterodyne Signals," Experimental Wireless and Wireless Engineer, Vol. 9, pp. 378-383, July 1932

[85] Terman, F. E., "Linear Detection of Heterodyne Signals," Electronics, Vol. 1, pp. 386-387, Nov. 1930

[86] Corrington, M. S., "Frequency Modulation Distortion Caused by Common- and Adjacent-Channel Interference," RCA Review, Vol. 7, pp. 522-560, Dec. 1946

[87] Wheeler, H. A., "Common-Channel Interference Between Two Frequency-Modulated Signals," Proc. I.R.E., Vol. 30, pp. 34-50, Jan. 1942

[88] Keall, O. E., "Interference in Relation to Amplitude, Phase and Frequency Modulated Systems," Wireless Engineer, Vol. 18, pp. 56-63, Feb. 1941

[89] Burgess, R. E., op. cit., (footnote 64 of this report)

DECLASSIFIED

DECLASSIFIED

RESTRICTED

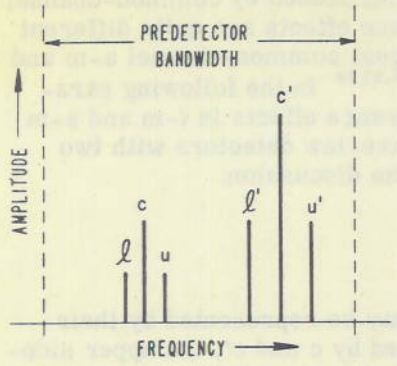


Figure 43 - Spectrum of the input voltage to an a-m detector in the case of common-channel interference

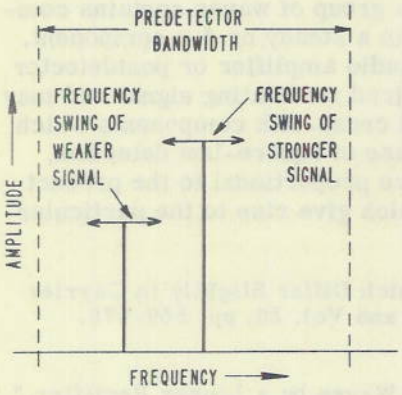


Figure 44 - Spectrum of the input voltage to the limiter of an f-m receiver in the case of common-channel interference

output frequency under consideration. The square-law detector may then be considered as a device which generates difference frequency components and in effect multiplies the amplitudes of the input components to produce the amplitude of each difference frequency component which appears in the output.

A linear detector behaves in a somewhat similar manner; however, it performs, in effect, the operation of taking the square root of the square of the envelope. Many other harmonics and cross-products are thereby generated in the case of the linear detector under multiple signal conditions, and their amplitudes for certain ideal cases can be expressed in terms of the higher mathematical functions. Useful approximations can be made when the modulation indices are small and when the amplitudes of the two carriers do not approach equality. In this case, the linear and square-law detectors behave similarly, except in one important respect. The amplitude of each difference frequency component in the case of the linear detector is proportional to the product of the amplitude of the two input components involved but is in addition inversely proportional to the square root of the sum of the squares of the amplitudes of all the input components. This has been shown[90] to give rise to the appearance of a capture effect when linear detection is employed, an effect which is not present when square-law detectors are employed. In the case of no interference, the linear detector traces the envelope of the input wave, yielding the desired signal modulation. In the case of the square-law detector, the fundamental modulation output component is generated by the carrier beating with the upper and lower sidebands, and second-harmonic distortion is generated by the upper sidebands beating with their counterparts among the lower sidebands. If more than one tone is present, many intermodulation frequencies are also generated by the square-law detector with double sideband a-m input.

Beat-Note Interference (AM)

When an interfering a-m wave is present, its components beat with all the components in the desired signal. Since the carrier is usually the strongest component of each of the two signals, the beat or heterodyne between carriers is normally the largest component of the interference. The carrier beat-note amplitude is proportional to the product of the two carrier amplitudes for the case of square-law detection and is proportional to the weaker of the two carrier amplitudes for the case of linear detection.

---

[90] Zeek, R. W., "Investigation and Analysis of 'Capture Effect' in Frequency Modulated and Amplitude Modulated Communication Systems," (Restricted), NRL Report R-3422, Feb. 1949 (see Appendices)

RESTRICTED

DECLASSIFIED

DECLASSIFIED

The carriers are normally fixed in frequency and have a constant amplitude in the case of a-m systems. The beat-note will then be a steady tone which will be apparent in the receiver output if the difference frequency lies within the audio-amplifier bandwidth. The cross-talk and harmonics which arise from beats between the weaker components of the input spectra result in output components in the detector output circuit which are smaller in amplitude than the heterodyne. The nature of the cross-talk depends on the difference between the two carrier frequencies. If the carriers are separated far enough in frequency, no audible intermodulation interference will be present, but if the carriers are separated by an amount only slightly greater than the audio bandwidth, the carrier beat-note will be missing and the cross-talk will take the form of inverted speech, which is sometimes referred to as "monkey chatter." If the two signals are separated far enough so that no heterodyne or audible interference is heard but both signals are still within the i-f bandwidth, the capture effect phenomenon still remains unless perfect square-law detection is attained.

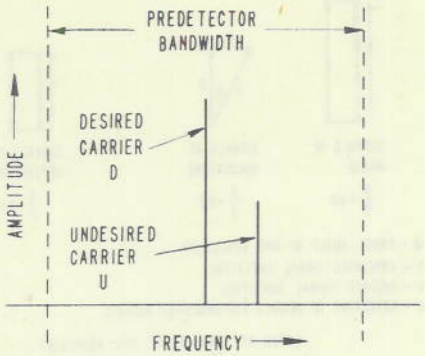


Figure 45 - Spectrum at input to limiter during common-channel interference (both signals unmodulated)

F-M System (Two-Signal Input)

In the case of FM, one may consider the desired and undesired input signals as represented by two carriers which vary in frequency but which remain constant in amplitude, as shown in Figure 44. As previously indicated, these two waves combine to form a resultant which undergoes both phase and amplitude variations. The envelope amplitude variations of this resultant are removed by a perfect limiter in the f-m receiver. Theoretically, a constant-amplitude variable-frequency wave is then passed on to the discriminator. When the signal is accurately tuned, a balanced discriminator produces an output voltage which is proportional to the deviation of the instantaneous frequency of the limiter output wave from the balance or center frequency of the discriminator. The components of the detector output wave which have frequencies lying within the audio bandwidth are passed on to the output of the receiver. Thus only the variation in instantaneous frequency of the resultant of the two input waves determines the nature of the output waveform.

The spectrum at the input to the limiter is shown in Figure 45, for the case when neither signal is modulated. The resultant of the two waves and its phase angle are shown in the vector diagrams of Figure 46. A rapid shift in the phase angle of the resultant occurs when the two waves are opposed to each other. This phase shift becomes greater as the amplitudes of the two waves approach equality. For a given ratio of input amplitudes, both the phase shift and the rate of phase change of the resultant are greatest when the two signals are in phase opposition and hence the most severe phase shifts correspond to the holes or dips in the envelope of the waves.

The discriminator output voltage is proportional to the rate of change of this phase angle, and takes the form of a pulse which recurs at a rate equal to the difference frequency of the two carriers. The spectrum of the output wave thus consists of components at the difference frequency and its harmonics. When one input wave is much smaller in amplitude than the other, the output pulse resulting from the phase shift of the resultant of the two input waves is broad, and hence most of the detector output energy is contained

DECLASSIFIED

DECLASSIFIED

RESTRICTED

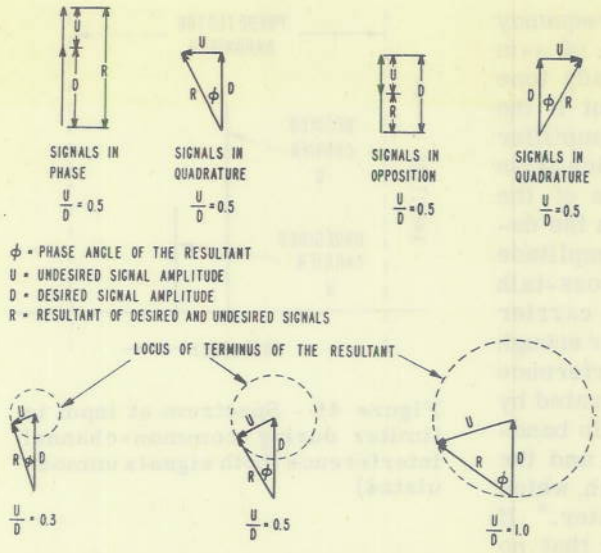


Figure 46 - Vector diagrams illustrating phase variations of the resultant of desired and undesired signals during common-channel interference

in the difference-frequency fundamental. As the amplitudes of the two input carriers approach equality, the detector output pulse becomes more sharply peaked and the energy in the harmonics of the difference frequency in the detector output increases.

When the input carrier waves are of equal amplitude, all the harmonic components of the detector output wave have equal amplitudes, as in the spectrum of a unit impulse. The audio amplifier passes only those components which fall within its response band and also modifies their relative phase angles. Thus the pulses which reach the output of the receiver will usually be rounded and broadened forms of the detector output pulses.

When the desired signal is modulated and the undesired signal is unmodulated (the desired signal having the larger amplitude), the detector output wave contains the desired modulation with pulses superimposed. The pulse peaks occur during intervals when the input waves are in opposition, i.e., when a dip or hole occurs in the envelope of the composite wave at the input to the detector. If the limiter does not function when these dips in the envelope occur (which may be the case for practical limiters), some of the envelope amplitude variations are also passed on to the discriminator. The detector output caused by these amplitude variations is of such polarity as to reduce the pulse output caused by detection of the rapid phase shift of the resultant. Thus the effect of imperfect limiting is to reduce the undesired signal carrier interference output of an f-m receiver.

As the desired signal is modulated, the rate of occurrence of the impulses changes in accordance with the change in the difference frequency between the two carriers. The maximum amplitudes of the impulses depend on the amplitude ratio of the two interfering signals and on their instantaneous frequency difference at the point of phase opposition. Thus, when the desired signal (in this case, the modulated signal) is the stronger, the detector output contains the modulation of the desired signal with a varying density of impulses of changing amplitude superimposed.

In the limit, when the two input signals have the same amplitude, the desired signal output component is reduced to one-half its interference-free amplitude and the interfering impulses have a theoretically infinite amplitude and an infinitesimal duration. The interfering impulses will produce a noise component in the receiver output which rises to a maximum at carrier-level equality, and falls off on both sides of this unity carrier-ratio condition. The effect of excellent and mediocre limiting on this phenomenon is shown in Figure 47.

As the undesired signal becomes stronger than the desired signal, the impulses become smaller in amplitude and the modulation of the desired signal is further suppressed or "captured." When both the desired and undesired signals are modulated, similar effects

RESTRICTED

DECLASSIFIED

RESTRICTED
DECLASSIFIED

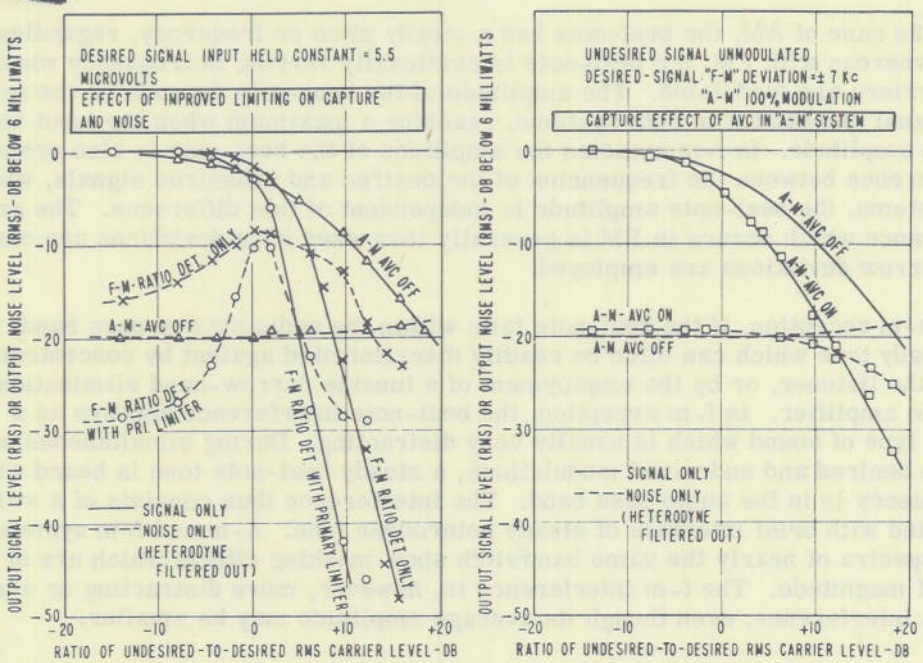


Figure 47 - Comparison of measured a-m and f-m co-channel (two-signal) effects (desired-signal depression and noise variation)

occur, so that as the undesired signal becomes the stronger, its modulation will predominate in the output and the desired signal will ultimately disappear into the noise level.

Beat-Note Interference (FM) and Comparison of FM and AM

With both carriers unmodulated, a heterodyne tone will appear in the receiver output, with a pitch corresponding to the pulse repetition frequency. This tone will have a harmonic content determined by carrier amplitude ratio and the number of harmonics which are within the audio amplifier response range. Modulating one or both carriers causes the heterodyne tone to change pitch in accordance with the modulation. Increasing the deviation of the f-m signals tends to reduce the interference in the receiver output, because a greater percentage of the interference appears in higher beat-note frequencies which are not passed by the audio amplifier. It should be noted that the use of a de-emphasis network in the system will reduce the interference, since it has the effect of reducing the audio bandwidth insofar as the interference is concerned.

The masking effects in a-m and f-m systems during conditions of common-channel interference may be compared as follows. The heterodyne or beat-note interference comprises the most significant part of the masking output in both f-m and a-m systems. Harmonics of the beat-note, as well as signal and heterodyne intermodulation components, also contribute to the masking effects in both f-m and a-m systems. There is a rise in noise output in the f-m case as the carrier equality point is approached, which is not encountered with AM.

RESTRICTED
DECLASSIFIED

DECLASSIFIED

RESTRICTED

In the case of AM, the beat-note has a steady pitch or frequency, regardless of modulation, whereas with FM, the beat-note is continually varying in frequency when one or both carriers are modulated. The amplitude of the beat-note depends on the ratio of the input signal amplitudes in both systems, reaching a maximum when the input signals are equal in amplitude. In f-m systems the amplitude of the beat-note is also proportional to the difference between the frequencies of the desired and undesired signals, whereas in a-m systems, the beat-note amplitude is independent of this difference. The amount of interference which occurs in FM is generally less when wide deviations are employed than when narrow deviations are employed.

In a-m reception, if the beat-note falls within the audio system pass band, it appears as a steady tone which can often be readily discriminated against by concentration on the part of the listener, or by the employment of a tunable narrow-band elimination filter in the audio amplifier. In f-m reception, the beat-note interference appears as a "swishy," spitting type of sound which is usually very distracting. During simultaneous pauses in both the desired and undesired modulations, a steady beat-note tone is heard with FM if its frequency is in the audio pass band. The interference thus consists of a swishy sound punctuated with brief intervals of steady heterodyne tone. A-m and f-m systems having signal spectra of nearly the same bandwidth show masking effects which are of the same order of magnitude. The f-m interference is, however, more distracting or annoying than the a-m interference, even though its average amplitude may be smaller.

A narrow-deviation f-m system might be expected to exhibit a lesser degree of interference than an a-m system when the difference between the desired and undesired signal frequencies is very small (less than 100 cps). Wide-deviation f-m systems should exhibit less common-channel interference than a-m systems, in proportion to the ratio of the peak frequency deviation to the audio cutoff frequency of the system.

De-emphasis may be expected to provide a reduction of the common-channel interference in both a-m and f-m systems. The amount of interference reduction should be nearly proportional to the reduction in effective audio bandwidth. When the system employs a rather narrow audio bandwidth, the effectiveness of a de-emphasis network may be rather small. A more complete and detailed discussion of common-channel interference effects in f-m and a-m systems may be found in NRL Report R-3460.[91]

Capture Effects

In addition to the masking effects discussed above, which occur during conditions of common-channel interference, the modulation suppression or "capture" effects will also occur. While the masking effects are the more evident when the difference frequencies between the undesired and desired signal components lie within the audio range, the capture phenomena become the more evident when the difference tone between the signal frequencies is inaudible, since it is then not overshadowed by the masking effects of the heterodyne and signal cross-talk.

The degree of capture or modulation suppression increases rapidly, particularly in f-m systems, when the ratio of the undesired-to-desired signal amplitudes exceeds unity. When the undesired-to-desired signal amplitude ratio is less than or equal to unity, the amount of modulation suppression lies between 0 db and 6 db. An exhaustive treatment

---

[91] Howe, W. E. W. and Zeek, R. W., "Investigation of Nonlinear Receiving Circuits Under Two-Signal Conditions" (Restricted), NRL Report R-3460, Nov. 15, 1949

RESTRICTED

DECLASSIFIED

DECLASSIFIED

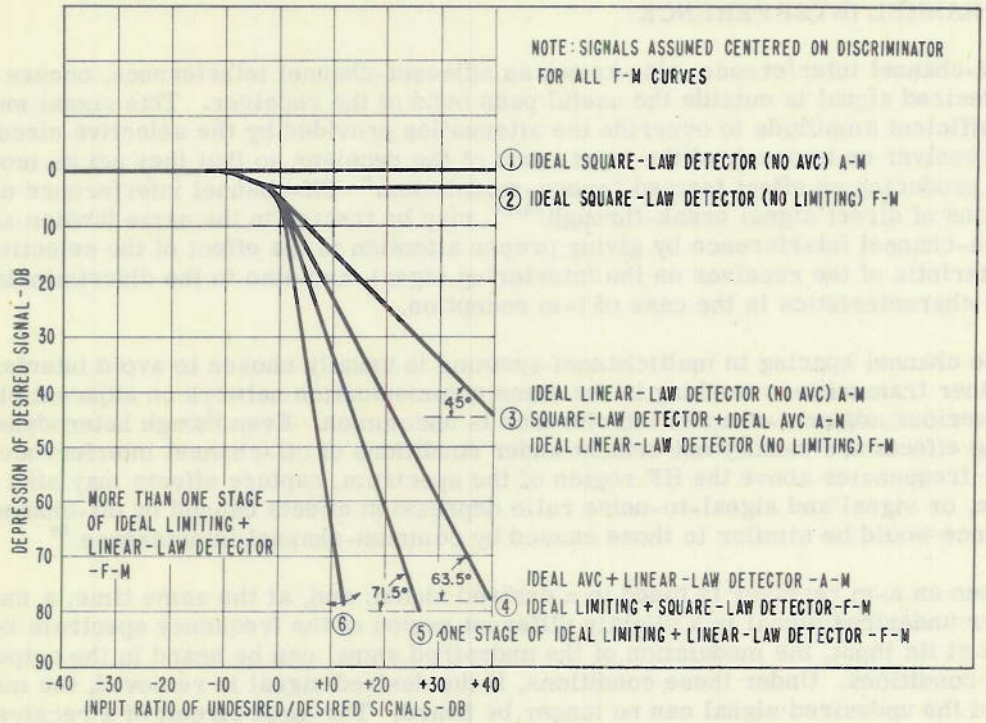


Figure 48 - Theoretical capture curves for f-m and a-m receivers

of capture phenomena has been presented in NRL Report R-3422,[92] hence only a brief comparison of f-m and a-m capture phenomena will be given in the following paragraphs.

Figure 48 illustrates the modulation suppression or capture effects for both f-m and a-m reception. The amount of reduction of the desired modulation output from its interference-free value in decibels is plotted as a function of the undesired-to-desired signal amplitude ratio for several different conditions. The characteristic most free of the capture effect is that exhibited by an a-m receiver without AVC and employing a perfect square-law detector. This curve is the horizontal line corresponding to 0 db reduction of the desired modulation for all values of the input signal ratio. Square-law detectors produce undesirable modulation distortion, however, and hence are not commonly employed in double sideband a-m receivers. The characteristic exhibited by the a-m receiver with a linear detector and with no AVC has a capture slope (see curve) of 45 degrees. When AVC is employed in an a-m receiver having a linear detector, the capture slope becomes approximately 63.5 degrees. The greatest degree of capture is that exhibited by an f-m receiver having a perfect limiter, linear detector, and accurately balanced discriminator. Practical a-m receivers may be expected to exhibit capture slopes of the order of 40 to 60 degrees, while practical f-m receivers may be expected to exhibit capture slopes of the order of 50 to 80 degrees. In its ability to obtain intelligence from a desired signal when a stronger undesired signal is simultaneously present, the a-m system has a decided advantage.

---

[92] Zeek, R. W., op. cit.

DECLASSIFIED

DECLASSIFIED

OFF-CHANNEL INTERFERENCE

Off-channel interference, also known as adjacent-channel interference, occurs when the undesired signal is outside the useful pass band of the receiver. This signal must have sufficient amplitude to override the attenuation provided by the selective circuits of the receiver or to overload the input tubes of the receiver so that they act as modulators, producing an effect termed "cross-modulation." Off-channel interference under conditions of direct signal break-through[93,94] may be treated in the same fashion as common-channel interference by giving proper attention to the effect of the selectivity characteristic of the receiver on the interfering signal, and also to the discriminator and limiter characteristics in the case of f-m reception.

The channel spacing in multichannel systems is usually chosen to avoid interference from other transmitters working in the same communication network on adjacent channels. Hence serious adjacent-channel interference is uncommon. Even though heterodyne-tone masking effects are usually not evident under conditions of off-channel interference at carrier frequencies above the HF region of the spectrum, capture effects may still occur. Capture, or signal and signal-to-noise ratio depression effects caused by off-channel interference would be similar to those caused by common-channel interference.[95]

When an a-m receiver is tuned to a desired signal, and, at the same time, a much stronger undesired signal in a slightly different region of the frequency spectrum is present at its input, the modulation of the undesired signal can be heard in the output under certain conditions. Under these conditions, if the desired signal is removed, the modulation of the undesired signal can no longer be heard. The input circuit of a receiver is usually designed to attenuate signals which lie outside the region of the spectrum occupied by the desired signal. However, when the undesired signal is so strong as to override the selectivity of the input circuit and to cause the input tube to operate in a nonlinear region of its transfer characteristic, this interference effect, which is termed cross-modulation, will occur.

Cross-modulation is caused primarily by third-order nonlinearity of the transfer characteristics of the input tube of the receiver. This form of interference normally is not evident in the reception of f-m signals with ideal receivers. The results of the following mathematical analysis of cross-modulation will be employed to discuss this effect in both f-m and a-m systems.[96]

Let a desired signal, represented by D sin ϕ_d, and an undesired signal, represented by U sin ϕ_u, be incident on a vacuum tube whose input-output (transfer) characteristic is given by the following expression:

$$e_o = a_o + a_1 e_i + a_2 e_i^2 + a_3 e_i^3 \ldots\ldots a_n e_i^n .$$

---

[93] Plusc, Igor, "Investigation of Frequency-Modulation Signal Interference," Proc. I.R.E., Vol. 35, pp. 1054-1059, Oct. 1947

[94] Goldman, S., op. cit.

[95] Corrington, M. S., op. cit., (footnote 86 of this report)

[96] Güttinger, P., op. cit.

[redacted]

DECLASSIFIED

DECLASSIFIED

In the above expressions, D, U, ϕ_d, and ϕ_u may be functions of time, thus denoting amplitude and frequency modulation of both signals; e_o is the output voltage from the vacuum tube, and e_i is the input or grid voltage. The a's are constants determined by the characteristics of the tube and the biasing voltages applied. Substituting D sin ϕ_d + U sin ϕ_u for e_i in the above expression gives:

$$\begin{aligned} e_o = a_o &+ a_1(D \sin \phi_d + U \sin \phi_u) \\ &+ a_2(D \sin \phi_d + U \sin \phi_u)^2 \\ &+ a_3(D \sin \phi_d + U \sin \phi_u)^3 \ldots\ldots \\ &\ldots + a_n(D \sin \phi_d + U \sin \phi_u)^n . \end{aligned}$$

Expansion of this expression using trigonometric identities and regrouping of terms leads to the following expression for the output voltage from the tube:

$$\begin{aligned} e_o = &\left\{a_o + \frac{a_2}{2}(D^2 + U^2)\right\} + a_2 D U \cos(\phi_d - \phi_u) \\ &+ \left\{a_1 D + \frac{3}{4} a_3 (D^3 + 2 D U^2)\right\} \sin \phi_d \\ &+ \left\{a_1 U + \frac{3}{4} a_3 (U^3 + 2 D^2 U)\right\} \sin \phi_u \\ &+ \frac{3}{4} a_3 D U \left\{D \sin(2\phi_d - \phi_u) + U \sin(2\phi_u - \phi_d)\right\} \\ &- \frac{a_2}{2}\left\{D^2 \cos 2\phi_d + U^2 \cos 2\phi_u\right\} - a_2 D U \cos(\phi_d + \phi_u) \\ &- \frac{3}{4} a_3 D U \left\{D \sin(2\phi_d + \phi_u) + U \sin(\phi_d + 2\phi_u)\right\} \\ &- \frac{a_3}{4}\left\{D^3 \sin 3\phi_d + U^3 \sin 3\phi_u\right\} + \ldots\ldots\ldots\ldots . \end{aligned} \tag{37}$$

Figure 49 shows the input and output spectral components for a tube in which cross-modulation occurs, for the case when the input consists of two carriers having frequencies f_d and f_u. When these carriers are modulated they would, of course, be accompanied by sidebands. These have purposely been omitted for the sake of clarity of the drawings. The frequencies labeled f in this figure are analogous to the phase angles represented by ϕ in the above analysis.

The tuned circuits of the receiver respond to the component represented by the third term of the above expression, which contains the factor sin ϕ_d, where ϕ_d is the phase angle of the desired signal. It is seen that the amplitude of this component depends on a_3, the coefficient of the third-order term in the series expansion representing the tube characteristic; D, the amplitude of the desired signal which may be a function of time; and U, the amplitude of the undesired signal which may also be a function of time. Thus the signal which is passed on to the demodulator contains amplitude variations dependent on the magnitude of both the desired and undesired signals, but with phase-angle variations (represented by ϕ_d) dependent only on the desired signal.

DECLASSIFIED

DECLASSIFIED

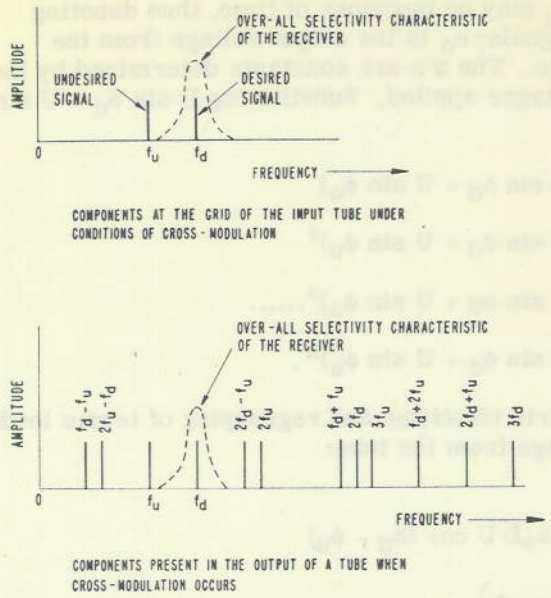


Figure 49

An a-m demodulator will respond only to the amplitude variations, and its output will contain the modulation of both the desired and undesired signal. It should be noted that since the amplitude term contains U^2 multiplied by D, the undesired modulation will not be present at the output of the receiver when the desired signal is not present (D = 0). Above a threshold level determined by the value of the subterms in D, the undesired modulation will increase rapidly in the output, approaching a rate directly proportional to U^2.

In the case of an ideal f-m receiver, the limiter will remove all variations in the amplitude portion of the third term above, and the wave which reaches the discriminator contains the phase variations of the desired signal only. In this case, the modulation of the undesired signal would not be observed in the output. If the f-m receiver does not contain a perfect limiter, the signal which reaches the discriminator contains amplitude variations which would give rise to the appearance of some undesired modulation in the output.

Analysis of the preceding expression indicates that there are some conditions under which cross-modulation can occur which were not treated above. For example, if the difference frequency $f_d - f_u$, represented by the second term of the expression, falls within the selectivity limits of the receiver, cross-modulation effects will occur in an f-m system, since the difference frequency component will contain the amplitude and frequency modulations of both signals.

When the undesired f-m signal is greater in amplitude than the desired signal at the input to the final detector system (including the limiter stages), the desired signal will be depressed or "captured" by the undesired signal. In this case, the first tube in the receiver will act as an imperfect limiter. A similar effect occurs in the a-m case but with a lesser degree of capture. These effects are identical with those encountered in the case of common-channel operation, but are modified by the adjacent-channel predetector selectivity of the receiver and are usually without heterodyne effects.

Cross-modulation effects are also caused by orders of nonlinearity higher than the third[97] and may originate in the second tube or the mixer under certain conditions. Variable-mu tetrodes or pentodes which have small third-order nonlinearity are usually employed in the input circuits of receivers to reduce cross-modulation.[98]

---

[97] Wheeler, H. A., "Two-Signal Cross Modulation in an FM Receiver," Proc. I.R.E., Vol. 28, pp. 537-548, Dec. 1940

[98] Ballantine, S. and Snow, H. A., "Reduction of Distortion and Cross-Talk in Radio Receivers by Means of Variable-mu Tetrodes," Proc. I.R.E., Vol. 18, pp. 2102-2127, Dec. 1930

[redacted]

DECLASSIFIED

DECLASSIFIED

Reduction in output signal-to-noise ratio appears under conditions of off-channel interference and is much more pronounced in f-m than in a-m systems. It may be explained as resulting in part from the capture effects discussed above. In UHF receivers, a large part of the noise originates in the plate circuit of the first r-f amplifier tube, in the second r-f amplifier tube (if this is provided), and particularly in the mixer tube. Nonlinear operation will generally be most evident in the first tube of the receiver, since the tuned circuits following the first tube offer further attenuation to the undesired signal, thereby protecting the next tube. Thus the undesired signal may cause a reduction in desired signal amplitude due to the signal-depression phenomenon. This, however, would not normally take place to a significant degree after the first tube, so that there would be no proportionate reduction of the noise which is generated in the tubes and circuits which follow the first tube. Under these conditions, the output signal-to-noise ratio may be reduced to a greater extent due to the presence of strong off-channel interference in f-m receivers than in a-m systems. The deterioration in signal-to-noise ratio of the desired signal may also result from the generation of noise due to the rapid phase shifts of the resultant signal which occurs in the f-m demodulator in the case of common-channel interference, when the undesired-to-desired signal amplitude ratio is near unity.

In summary, cross-modulation occurs in a-m reception and is caused primarily by third-order nonlinearity of the transfer characteristics of the input tube. Cross-modulation effects are virtually eliminated in f-m reception if an effective limiter is employed, except under certain special conditions. Noise-level increases and signal-to-noise ratio depression caused by strong off-channel interference are greater in f-m systems than in a-m systems.

SUMMARY COMPARISON OF A-M AND F-M CHARACTERISTICS

The tabular comparison of a-m and f-m system characteristics which follows contains, in brief form, summaries of the information which is more fully discussed in the body of this report.

PROPERTY UNDER CONSIDERATION	A-M CHARACTERISTICS	F-M CHARACTERISTICS
Capacity for transmission of information in the absence of interference.	Proportional to bandwidth of the transmission and to the time during which the system is available for transmitting information.	Proportional to bandwidth of the transmission and to the time during which the system is available for transmitting information.
Bandwidth of the spectrum of the radiated wave.	Twice the highest modulating frequency in the case of double-sideband AM; equal to the frequency of the highest modulating frequency in the case of single-sideband AM.	Twice the highest modulating frequency plus twice the peak frequency deviation for components not less than about 10% of unmodulated carrier amplitude.
Average power in the modulated wave.	One and one-half times the average power in the unmodulated carrier for 100% modulation.	Same as the average power in the unmodulated wave for any deviation.

DECLASSIFIED

DECLASSIFIED

PROPERTY UNDER CONSIDERATION	A-M CHARACTERISTICS	F-M CHARACTERISTICS
Symmetry of radiated spectrum.	Radiated spectrum is always symmetrical about the carrier frequency for double-sideband AM.	Radiated spectrum is symmetrical about the unmodulated carrier frequency only when the modulating waveform is symmetrical about the time axis.
Instantaneous peak power in the radiated wave.	Four times the instantaneous peak power in the unmodulated carrier (for 100% modulation).	Equal to the peak power in the unmodulated carrier.
Peak voltage of the modulated wave.	Twice the peak voltage of the unmodulated wave (for 100% modulation).	Same as the peak voltage of the unmodulated wave.
Fading of waves during transmission due to changes in tropospheric reflections.	Fading may be expected to be most serious in wide-band systems. Carrier fading most serious in effect.	Generally worse than AM, because f-m radiated spectrum usually much wider than for AM.
Modulation harmonic distortion arising from nonlinear tube characteristics in receiver r-f and i-f circuits.	Can occur if proper design precautions are not taken.	Does not normally occur.
Harmonic distortion due to nonlinear frequency-response and phase-response characteristics of tuned circuits.	Ratio of circuit bandwidth required to spectrum bandwidth of the modulated wave for a given amount of distortion is generally less than for the f-m case.	Harmonic distortion occurs when frequency- and phase-response characteristics of the tuned circuits are nonlinear (even if symmetrical) with respect to unmodulated carrier frequency. Circuit bandwidths required for the same amount of distortion are generally greater than for the corresponding case in a-m systems.
Modulation harmonic distortion due to modulator characteristics.	A fundamental limitation is the degree to which the dynamic characteristic of the modulator tube can be made square-law.	Distortion-free modulation may be achieved by employing small deviation during initial modulation, followed by wide-band frequency multiplication.
Modulation harmonic distortion due to demodulator characteristics.	Distortion is due to nonlinearities of the demodulator tube characteristic, and to unequal a-c and d-c loading of the demodulator tube (in the case of modulation percentages near 100).	Nonlinear discriminator characteristics and insufficient predetector bandwidth will introduce distortion. Otherwise essentially the same as AM.

[redacted]

DECLASSIFIED

DECLASSIFIED

PROPERTY UNDER CONSIDERATION	A-M CHARACTERISTICS	F-M CHARACTERISTICS
Distortion arising from multipath transmission.	Distortion does not occur until a condition of over-modulation sets in. Distortion should be less in any a-m system than in a narrow-deviation f-m system. A-m multipath distortion may be reduced by employing exalted carrier reception.	Distortion increases rapidly with increasing deviation and then decreases when the deviation becomes much greater than the audio bandwidth. Wide-deviation f-m systems may suffer less from multipath distortion than a-m systems having the same audio bandwidth. In the latter case, however, the f-m system would require a much wider bandwidth than the a-m system.
Frequency distortion due to tuned circuits.	Frequency distortion in a-m systems will generally produce the effect of low frequency pre-emphasis; i.e., greater transmission for the lower modulation frequencies.	Frequency distortion in f-m systems appears as a pre-emphasis of the high modulating frequencies.
Signal-to-random-noise ratio (under conditions of accurate tuning).	A-m receivers generally provide greater output signal-to-noise ratios than f-m receivers having the same i-f and audio bandwidths when the input carrier-to-noise ratio has low values (less than unity), generally corresponding to the case of weak-signal reception. A-m receivers may provide greater output signal-to-noise ratio than f-m receivers,which do not contain amplitude limitation. A-m systems exhibit a low rate of variation of output signal-to-noise ratio as the carrier-to-noise ratio varies (1 db per db above C/N ratio = 0 db; 2 db per db below this threshold).	F-m receivers having perfect amplitude limiters provide greater output signal-to-noise ratios than a-m receivers having the same audio and i-f bandwidths when the input carrier-to-noise ratio has values greater than unity, corresponding to the case of strong signal reception, and when the ratio of the peak frequency deviation to the audio bandwidth exceeds $1/\sqrt{3}$. The amount by which the f-m output signal-to-noise ratio exceeds the a-m signal-to-noise ratio under these conditions is proportional to the deviation ratio of the f-m system. F-m systems exhibit a very abrupt change in output signal-to-noise ratio as the carrier-to-noise ratio varies near unity.

DECLASSIFIED

DECLASSIFIED

[redacted]

PROPERTY UNDER CONSIDERATION	A-M CHARACTERISTICS	F-M CHARACTERISTICS
Signal-to-random noise ratio (under conditions of moderate detuning).	Moderate detuning of an a-m signal results in a very minute deterioration of the output signal-to-noise ratio for all values of carrier-to-noise ratio.	Moderate detuning of an f-m receiver with a perfect amplitude limiter results in serious deterioration of the output signal-to-noise ratio when the input carrier-to-noise ratio is less than about 5 to 1 (weak- and moderate-strength signals). The deterioration is less pronounced when the carrier-to-noise ratio is greater than 5 (strong signals). An f-m receiver having no amplitude limiting is not so seriously affected by detuning as a receiver with a limiter, as far as deterioration of signal-to-noise ratio is concerned. In the latter case, the deterioration is greater for stronger signals than for weaker one.
Effect of pre-emphasis and de-emphasis on output signal-to-noise ratio.	The output signal-to-noise ratio is improved in a-m systems by the use of pre-emphasis and the corresponding de-emphasis. The amount of improvement is greater for wider postdetector bandwidths and is roughly independent of the input carrier-to-noise ratio. The improvement is not seriously affected by detuning of the signal.	The improvement in output signal-to-noise ratio due to the use of pre-emphasis and de-emphasis in f-m systems is greater than that which can be achieved in a-m systems having the same bandwidth for the case of high input carrier-to-noise ratio and no detuning. The improvement is less for the low carrier-to-noise ratios than for the high carrier-to-noise ratios and decreases with detuning.

[redacted]

DECLASSIFIED

DECLASSIFIED

PROPERTY UNDER CONSIDERATION	A-M CHARACTERISTICS	F-M CHARACTERISTICS
Impulse noise interference (noise weaker than signal at i-f amplifier output).	The a-m system is generally inferior to the f-m system in the suppression of weak impulse noise, unless noise-limiter circuits are employed. Simple noise-peak limiters have little effect on weak impulse noise.	The f-m system with perfect limiter is generally superior to the a-m system in suppression of weak impulse noise, particularly when the tuned circuits and the discriminator are accurately tuned and have symmetrical characteristics. Departures from precise circuit tuning and symmetry, discriminator nonlinearity or detuning, and signal detuning cause a reduction of the effectiveness of the f-m system in suppression of weak impulse noise.
Impulse noise interference (noise stronger than signal at the i-f amplifier output).	A simple postdetector noise-peak limiter is very effective in reducing strong impulse noise interference. Its effectiveness is independent of detuning or tuned circuit misadjustments. A-m systems may be found superior to f-m systems in suppression of strong impulse noise.	The noise-reducing ability of f-m receivers having perfect limiters and which are otherwise ideal is severely reduced when the signal is detuned. A simple noise-peak limiter, if added in the detector output circuit of an f-m receiver would make it equivalent to an a-m receiver in performance.
Common-channel interference (masking effects).	The heterodyne between the two carriers is the strongest masking component in the audio output. This masking effect is present only when the carrier separation is less than the audio amplifier bandwidth. The beat-note in a-m is a steady tone whose amplitude is independent of carrier separation and can be discriminated against by concentration on the part of the listener or by employing a narrow-band audio-elimination filter.	In f-m systems, the beat-note is continually varying in frequency, producing a swishing, spitting sound which is difficult to discriminate against. The amplitude of the beat-note is proportional to the instantaneous frequency difference of the two f-m signals. Wide-deviation f-m systems should exhibit less common-channel masking effects than narrow-deviation systems.

DECLASSIFIED

DECLASSIFIED
RESTRICTED

PROPERTY UNDER CONSIDERATION	A-M CHARACTERISTICS	F-M CHARACTERISTICS
Common-channel interference (noise rise).	No appreciable change of noise level with increase of interfering-to-desired carrier ratio.	Noise level increases as much as 20 db at carrier amplitude-ratio equality.
Common-channel interference (capture or modulation suppression effects).	Modulation suppression due to the presence of a strong interfering signal (capture effect) is normally less pronounced in a-m systems than in f-m systems. A-m systems normally exhibit a capture slope between 40 and 60 degrees, depending upon the law of the detector and effectiveness of the AVC system.	Capture effect in f-m is more pronounced than in a-m systems. The capture slope usually lies between 50 and 80 degrees, depending on the law of the detectors, the effectiveness of limiting, and the linearity of the discriminator characteristic. Suppression of the modulation increases as the receiver is made more nearly ideal.
Off-channel interference (signal-to-noise ratio depression).	Under conditions of strong off-channel interference, the output signal-to-noise ratio of a-m receivers is not appreciably reduced.	The presence of strong off-channel interference causes a marked deterioration in the output signal-to-noise ratio of f-m receivers.
Cross-modulation.	Cross-modulation can occur in a-m systems if an interfering signal is strong enough to cause nonlinear operation of the input tubes of the receiver. Long-cutoff tubes must be employed in the input stages to minimize this effect.	Cross-modulation does not normally occur in f-m systems having a perfect amplitude limiter, under usual adjacent-channel operation conditions.

CONCLUSIONS

One cannot state that either an a-m or f-m system is more advantageous for general application. When the detailed conditions under which a system is to operate are known, including the type of intelligence to be transmitted, the importance of such considerations as power consumption, service area, distortion, performance in the presence of strong and weak interferences, etc., it may be feasible to specify the preferred type of modulation. Disadvantages and advantages of both systems must be carefully weighed in the light of their importance in the particular situation under consideration.

Wide-deviation f-m systems are capable of providing greater output signal-to-noise ratios than a-m systems when the signal is strong compared to the noise, providing the f-m receiver is accurately tuned to the signal. As the signal and noise amplitudes approach equality, the f-m system rapidly loses its advantage over the a-m system, and when the signal is weaker than the noise the a-m system will give the better output signal-to-noise ratio. A narrow-deviation f-m system exhibits only a small advantage in output signal-to-noise ratio over an equivalent a-m system for large input signal-to-noise ratios.

RESTRICTED
DECLASSIFIED

DECLASSIFIED

An a-m system may be capable of providing a greater voice communication range than an f-m system of the same bandwidth under practical conditions. Detuning affects the output signal-to-noise ratio of f-m systems to a much greater extent than it affects a-m systems. Distortion due to circuits with nonlinear amplitude response is greater in a-m systems than in f-m systems, but distortion arising during transmission of modulated waves through tuned circuits is greater in the case of FM. Properly designed f-m systems are more subject to capture effect than a-m systems. Cross-modulation generally does not occur in f-m systems, but the reduction of signal-to-noise ratio under conditions of strong off-channel interference is greater in f-m systems. These are only some of the differences between the characteristics of these two types of systems. These factors take on different degrees of importance when different applications are considered.

For naval multichannel voice communications systems operating in the UHF range the majority of the advantages favor AM over narrow-band FM. The a-m system will provide greater communication range, less deterioration due to detuning (an important consideration in fixed channel systems employing crystal control of frequency), less reduction of desired signal modulation due to common-channel interference (capture effect), better suppression of strong impulse noise, less tuned-circuit distortion (although this is not an important consideration), less deterioration of output signal-to-noise ratio under conditions of strong off-channel interference, and several other advantages. F-m system advantages include better signal-to-noise ratio for strong, accurately centered input signals, less distortion due to nonlinear tube characteristics (not an important consideration), better performance in the presence of weak impulse noise and freedom from cross-modulation effects. The a-m advantages outweigh the f-m advantages with regard to the naval application, where good weak-signal performance and immunity to strong interference are required.

An instance where FM would be more desirable is the case of commercial broadcasting requiring low modulation distortion, economical operation of high-power transmitters, high signal-to-noise ratio in a given area, and low signal-to-noise ratio outside the given area.

Another case where f-m systems would be preferable is in multiplex relay links involving numerous relay stations. Here the cross-modulation and distortion due to vacuum tube nonlinearities would limit the usefulness of a-m systems.

ACKNOWLEDGMENTS

The assistance provided by Mr. Robert W. Zeek, Radio Techniques Branch, Radio Division II; Dr. Duncan Harkin, formerly with Radio Division III; and Mr. Samuel F. George, Radio Division III, is gratefully acknowledged. Particular thanks are due to Mr. Emerick Toth, Radio Techniques Branch, Radio Division II, for his stimulating discussions and guidance.

* * *

DECLASSIFIED

[redacted]

[redacted]