NRL REPORT NO. R-31

AMPLITUDE AND FREQUENCY MODULATION FOR FACSIMILE TRANSMISSION AND OTHER APPLICATIONS

by

A. D. Watt

September 1947

Problem No. 39R01-11

Approved by:

L. A. Gebhard
Superintendent
Radio Division II

Commodore H. A. Schade, USN
Director
Naval Research Laboratory

FR-3154

NAVAL RESEARCH LABORATORY

Washington, D.C.

DISTRIBUTION

BuShips	
Attn: Code 925A	10
Attn: Code 925B	2
CNO	
Attn: Code 20B	1
Attn: Code 20D	1
Attn: Code 413C	1
ONR	
Attn: Code N482A	5
Dir., USNEL	2
RDB	
Attn: Library	2
Attn: Navy Secretary	1
Science and Technology Project	
Attn: Mr. J. H. Heald	2

ACKNOWLEDGMENT

Appreciation is expressed to Mr. C. C. Rae for his suggestions and assistance in the performance of experiments. Acknowledgment is also made to Mr. C. B. Davis for his suggestions and review of this report and to Mr. Morton Fischman who performed many of the computations and aided in the preparation of graphs.

CONTENTS

ABSTRACT

The bandwidth required for the transmission of intelligence by f-m waves is found to be as great or greater than that required by a-m transmission. The spacing of radio-frequency channels, however, is normally not a function of intelligence bandwidth, but is dependent on the rate at which the side band amplitudes decrease outside this band. When modulating functions with discontinuities are employed, the channel spacing required by frequency modulation may be less than that required by amplitude modulation.

The effect of noise and interfering signals on a-m waves and f-m waves of various deviation ratios is considered. General conclusions are formulated regarding the most desirable type of modulation for various modulating functions. Radio-circuit anomalies such as multipath, fading, and doppler effect are considered, and where possible the controllable factors are investigated for optimum operation.

PROBLEM STATUS

This report concludes the work on this phase of the problem. Work on the general problem (S1398) is continuing.

AMPLITUDE AND FREQUENCY MODULATION
FOR FACSIMILE TRANSMISSION AND OTHER APPLICATIONS

INTRODUCTION

In general the transmission of intelligence over a radio or wire circuit requires that some type of non-sinusoidal modulation function be employed. Sinusoidal and non-sinusoidal amplitude modulation have been treated rather thoroughly in the literature.* However, most analyses of carrier and side-band relations in frequency modulation are concerned only with the case of single sine-wave modulation.† Several exceptions are found‡ where square-wave modulation is treated, and also where the case of several applied sinusoids is considered. §

The following paragraphs contain mathematical derivations of the side-band distribution for several modulating wave shapes, as well as a graphical method which permits the solution of most modulating wave shapes. The output spectra for amplitude modulation and frequency modulation with the same modulating wave are presented graphically, and the interesting possibility that f-m transmission may produce less interference bandwidth than a-m transmission for certain modulating wave shapes is considered. Some other aspects such as signal-to-noise relations and the effect of a radio path on f-m transmission are also presented.

BANDWIDTH REQUIREMENTS

General Considerations

Numerous circuits and methods are available to produce frequency-modulated waves. Since we are primarily interested in the characteristics of the modulated carrier, we will concern ourselves with the wave after it is produced and neglect the method used, with one exception. It is important to note that the analysis used here does not apply to

---

* Terman, F. E., "Radio Enginner Handbook," McGraw-Hill Book Co., New York, 1943, p. 532; B. Salberg, "Frequency Spectra and Bandwidth of Amplitude-Modulated Radio Telegraph Transmission," NRL Report R-2426, 27 December 1944, and "Pulse Transmission Spectra," Naval Research Laboratory, 23 June 1941

† Carson, J. R., "Notes on the Theory of Modulation," Proc. IRE, Vol. 10, pp. 57-64, February 1922; Hans Roder, "Amplitude, Phase, and Frequency Modulation," Proc. IRE., Vol. 19, pp. 2145-2176, December 31

‡ Van Der Pol, Balth, "Frequency Modulation," Proc. IRE, Vol. 18, pp. 1194-1205, July 1930

§ Crosby, M.G., "Carrier and Side-Frequency Relations with Multi-Tone Frequency or Phase Modulation," RCA Review, Vol. 3, p. 103-106, July 1938 the type of frequency-shift keying using two separate oscillators alternately fed into an output circuit since in this type of keying rapid phase shifts occur when the output is shifted from one oscillator to the other.

Normally any modulated carrier can be represented by the equation

$$e = A \sin (F(t) + \theta) \tag{1}$$

where e is the instantaneous voltage at the output of the transmitter or at the point of reception. The quantity may be a function of time and determines the amplitude of the carrier; however, in the case of pure frequency modulation, A is constant and can be set equal to 1. The expression F(t) is a function of the instantaneous frequency, and θ determines the starting point of the carrier. For purposes of analysis we can rewrite equation (1) as

$$e = \sin \phi \tag{2}$$

which can be represented by the real part of the rotating vector in Figure 1.

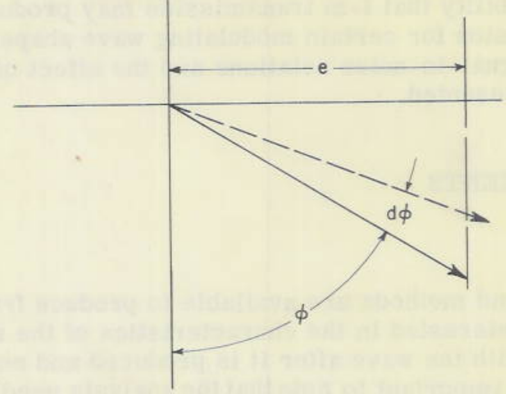


Figure 1

From Figure 1 we obtain the following relations

$$\omega_c = \frac{d\phi}{dt} \tag{3}$$

$$d\phi = \omega_c dt \tag{4}$$

$$\phi = \int \omega_c dt \tag{5}$$

where ω_c is the instantaneous frequency in radians per second. If ω_c is a constant we get

$$\phi = \omega_c t \tag{6}$$

which when substituted in (2) yields

$$e = \sin \omega_c t \tag{7}$$

If ω_c is not a constant we must use

$$e = \sin\int \omega_c \, dt. \tag{8}$$

ω_c can usually be represented as

$$\omega_c = \omega_o + \Delta\omega f(t) \tag{9}$$

where ω_o is 2π times the carrier frequency and $\Delta\omega$ is 2π times the maximum departure from the hypothetical mid frequency, i.e., the average frequency. The modulating function is f(t) and is normally considered as varying between limits of ± 1 if the function is symmetrical. If the function is not symmetrical a total variation of two is all that is required. Equation (8) now becomes

$$e = \sin(\omega_o t + \Delta\omega \int f(t)dt). \tag{10}$$

By trigonometric identity

$$e = \sin\omega_o t \cdot \cos(\Delta\omega\int f(t)dt) + \cos\omega_o t \cdot \sin(\Delta\omega\int f(t)dt). \tag{11}$$

Choosing f(t) such that its d-c component is zero, makes the $\int f(t)dt$ recurrent if f(t) is recurrent. Now $\cos(\Delta\omega\int f(t)dt)$ is also recurrent and can be represented as a series of cosine terms plus a d-c term. Equation (11) can now be written as

$$\begin{aligned} e = {} & A_o \sin\omega_o t + \sin\omega_o t \cdot (B_1 \cos\omega_a t + B_2 \cos 2\omega_a t + \ldots) \\ & + \cos\omega_o t \cdot (A_1 \sin\omega_a t + A_2 \sin 2\omega_a t + \ldots) \end{aligned} \tag{12}$$

Where B_n would be the magnitude of the nth harmonic of the $\cos(\Delta\omega\int f(t)dt)$ function, and A_n would be the magnitude of the nth harmonic of the $\sin(\Delta\omega\int f(t)dt)$ function. It immediately follows that

$$\begin{aligned} e = {} & A_o \sin\omega_o t + \frac{B_1}{2}\sin(\omega_o + \omega_a)t + \frac{B_2}{2}\sin(\omega_o + 2\omega_a)t + \ldots \\ & \frac{B_1}{2}\sin(\omega_o - \omega_a)t + \frac{B_2}{2}\sin(\omega_o - 2\omega_a)t + \ldots \\ & \frac{A_1}{2}\sin(\omega_o + \omega_a)t + \frac{A_2}{2}\sin(\omega_o + 2\omega_a)t + \ldots \\ & - \frac{A_1}{2}\sin(\omega_o - \omega_a)t - \frac{A_2}{2}\sin(\omega_o - 2\omega_a)t + \ldots \end{aligned} \tag{13}$$

Regrouping we obtain

$$\begin{aligned} e = {} & A_o \sin\omega_o t + \left[\frac{B_1}{2} + \frac{A_1}{2}\right]\sin(\omega_o + \omega_a)t + \left[\frac{B_1}{2} - \frac{A_1}{2}\right]\sin(\omega_o - \omega_a)t \\ & + \left[\frac{B_2}{2} + \frac{A_2}{2}\right]\sin(\omega_o + 2\omega_a)t + \left[\frac{B_2}{2} - \frac{A_2}{2}\right]\sin(\omega_o - 2\omega_a)t \\ & + - - - - \\ & + - - - - \end{aligned} \tag{14}$$

where A_o is the carrier amplitude,

$\frac{B_n}{2} + \frac{A_n}{2}$ is the amplitude of the nth side band above the carrier,

$\frac{B_n}{2} - \frac{A_n}{2}$ is the amplitude of the nth side band below the carrier.

As an aid in visualizing what has been done thus far, consider the case of a narrow rectangular pulse in Figure 2. The last two curves of Figure 2 are the ones to be represented by a series and it is rather obvious that the shape of these curves is dependent on $\Delta\omega \int f(t)\, dt$ (max), the maximum phase deviation, which will be represented as M. Since $\Delta\omega \int f(t)\, dt$ (max) is dependent on the pulse width, it follows that the output spectrum is a function of pulse width and frequency shift as might be expected.

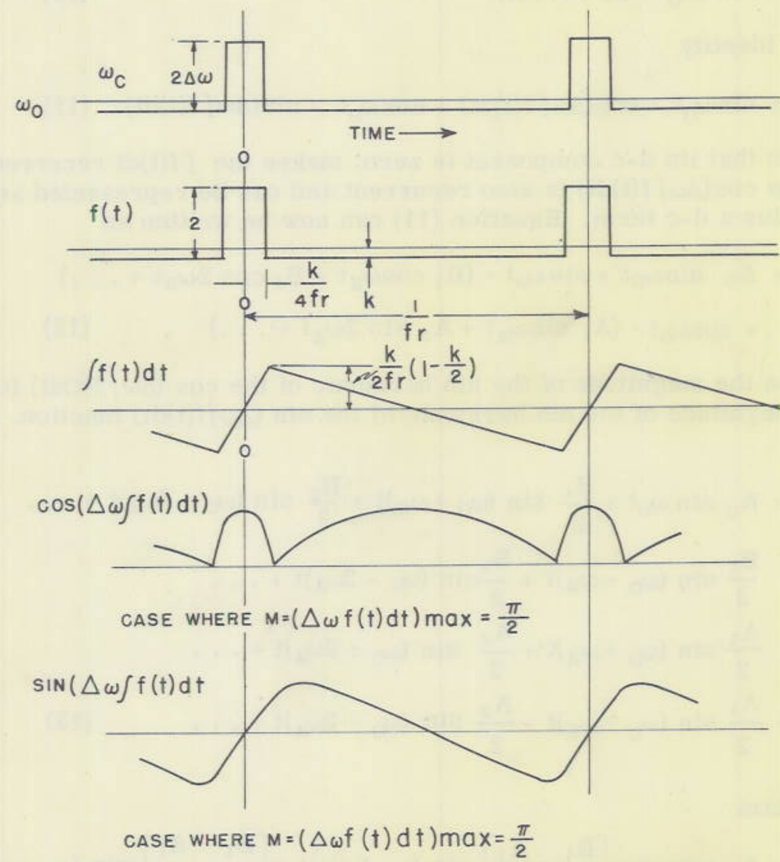


Figure 2 - Rectangular Pulse Frequency Modulation

The case of a rectangular pulse can be solved mathematically and should provide a good example to be carried through the operations indicated in the previous paragraphs. The derivation may seem rather lengthy; however, it should be noted that only rather simple mathematical functions are involved and most of the necessary steps are included.

Particular Solution for Rectangular Pulses and Square Waves

From Figure 2 we obtain the following equations:

$$f(t) = -k \qquad -\frac{1}{2f_r} < t < -\frac{k}{4f_r} \tag{15}$$

$$f(t) = 2-k \qquad -\frac{k}{4f_r} < t < \frac{k}{4f_r} \tag{16}$$

$$f(t) = -k \qquad \frac{k}{4f_r} < t < \frac{1}{2f_r} \tag{17}$$

For simplicity we shall change variables from t to x when integrating

$$\text{where } x = 2\pi f_r t \tag{18}$$

$$\text{or} \quad t = \frac{x}{2\pi f_r} \tag{19}$$

$$f(x)_1 = \int^{x/2\pi f_r} f(t)dt = -\frac{kx}{2\pi f_r} + C_1 \qquad -\pi < x < -\frac{\pi k}{2} \tag{20}$$

$$f(x)_2 = \int^{x/2\pi f_r} f(t)dt = \frac{x}{\pi f_r}\left(1 - \frac{k}{2}\right) \qquad -\frac{\pi k}{2} < x < \frac{\pi k}{2} \tag{21}$$

$$f(x)_3 = \int^{x/2\pi f_r} f(t)dt = -\frac{kx}{2\pi f_r} + C_2 \qquad \frac{\pi k}{2} < x < \pi \tag{22}$$

When $x = -\frac{\pi k}{2}$, (23)

$$f(x)_1 = +\frac{\pi k^2}{4\pi f_r} + C_1 \tag{24}$$

$$f(x)_2 = -\frac{\pi k}{2\pi f_r} + \frac{\pi k^2}{4\pi f_r}$$

Since $f(x)_1 = f(x)_2$

$$C_1 = -\frac{k}{2f_r} \tag{25}$$

$$f(x)_1 = -\frac{k}{2f_r}\left(1 + \frac{x}{\pi}\right) \tag{26}$$

and similarly $C_2 = \frac{k}{4f_r}$

$$f(x)_3 = \frac{k}{2f_r}\left(1 - \frac{x}{\pi}\right). \tag{27}$$

Since $f(x)_n = \int f(t)dt$, we can let (28)

$$F(x)_n = \cos \Delta\omega \int f(x)_n \tag{29}$$

By Fourier Series we get

$$F(x) = A_0 + B_1 \cos x + B_2 \cos 2x + \ldots \tag{30}$$

Note that the sine terms are absent because $F(x) = F(-x)$.

$$A_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(x)\, dx \tag{31}$$

and

$$B_n = \frac{1}{\pi} \int_{-\pi}^{\pi} F(x) \cos nx\, dx \tag{32}$$

$$F(x)_1 = \cos \Delta\omega \left[-\frac{k}{2f_r}\left(1 + \frac{x}{\pi}\right)\right] \tag{33}$$

$$F(x)_2 = \cos \Delta\omega \left[\frac{x}{\pi f_r}\left(1 - \frac{k}{2}\right)\right] \tag{34}$$

$$F(x)_3 = \cos \Delta\omega \left[\frac{k}{2f_r}\left(1 - \frac{x}{\pi}\right)\right] \tag{35}$$

In solving for A_0 and B_n's we shall do so for $F(x)_1$, $F(x)_2$, $F(x)_3$, separately.

$$A_{01} = \frac{1}{2\pi} \int_{-\pi}^{-\pi k/2} \cos \left[\frac{\Delta\omega k}{2f_r}\left(1 + \frac{x}{\pi}\right)\right] dx \tag{36}$$

$$= \frac{1}{2\pi}\left\{\cos \frac{\Delta\omega k}{2f_r} \int_{-\pi}^{-\pi k/2} \cos \frac{\Delta\omega k x dx}{2f_r \pi} - \sin \frac{\Delta\omega k}{2f_r \pi} \int_{-\pi}^{-\pi k/2} \sin \frac{\Delta\omega k x dx}{2f_r \pi}\right\} \tag{37}$$

$$= \frac{1}{2\pi}\left\{\cos \frac{\Delta\omega k}{2f_r}\left[\frac{2f_r\pi}{\Delta\omega k} \sin \frac{\Delta\omega k x}{2f_r\pi}\right]_{-\pi}^{-\pi k/2} + \sin \frac{\Delta\omega k}{2f_r}\left[\frac{2f_r\pi}{\Delta\omega k} \cos \frac{\Delta\omega k x}{2f_r\pi}\right]_{-\pi}^{-\pi k/2}\right\} \tag{38}$$

$$= \frac{1}{2\pi} \cdot \frac{2f_r\pi}{\Delta\omega k}\left\{\cos \frac{\Delta\omega k}{2f_r}\left[\sin\left(-\frac{\Delta\omega k^2}{4f_r}\right) + \sin\left(+\frac{\Delta\omega k}{2f_r}\right)\right]\right.$$

$$\left. + \sin \frac{\Delta\omega k}{2f_r}\left[\cos \frac{\Delta\omega k^2}{4f_r} - \cos \frac{\Delta\omega k}{2f_r}\right]\right\} \tag{39}$$

Letting $a = \frac{\Delta\omega k}{2f_r} = M \frac{1}{\left(1 - \frac{k}{2}\right)}$, where $M = \frac{\Delta\omega k}{2f_r}\left(1 - \frac{k}{2}\right)$ = maximum phase deviation,

$$A_{O1} = \frac{1}{2a}\left\{\cos a\left[\sin\frac{-ak}{2} + \sin a\right] + \sin a\left[\cos\frac{ak}{2} - \cos a\right]\right\} \quad (40)$$

$$= \frac{1}{2a}\left\{\cos a\,\frac{k}{2}\ \sin\ a - \sin a\,\frac{k}{2}\cos a\right\} \quad (41)$$

$$= \frac{1}{2a}\sin a\left(1 - \frac{k}{2}\right) . \quad (42)$$

$$A_{O1} = A_{O3}, \text{ therefore } A_{O1} + A_{O3} = \frac{\sin a\left(1 - \frac{k}{2}\right)}{a} \quad (43)$$

$$= \frac{\left(1 - \frac{k}{2}\right)}{M}\sin M . \quad (44)$$

$$A_{O2} = \frac{1}{2\pi}\int_{-\pi k/2}^{\pi k/2}\cos\left(\frac{\Delta\omega x}{\pi f_r} - \frac{\Delta\omega k x}{2\pi f_r}\right)dx \quad (45)$$

$$= \frac{1}{2\pi}\left\{\int_{-\pi k/2}^{\pi k/2}\cos\frac{\Delta\omega x}{\pi f_r}\cdot\cos\frac{\Delta\omega k x}{2\pi f_r} + \int_{-\pi k/2}^{\pi k/2}\sin\frac{\Delta\omega x}{\pi f_r}\cdot\sin\frac{\Delta\omega k x}{2\pi f_r}\right\} \quad (46)$$

$$= \frac{1}{2\pi}\left\{\left[\frac{\sin\frac{\Delta\omega}{\pi f_r}\left(1 - \frac{k}{2}\right)x}{\frac{2\Delta\omega}{\pi f_r}\left(1 - \frac{k}{2}\right)} + \frac{\sin\frac{\Delta\omega}{\pi f_r}\left(1 + \frac{k}{2}\right)x}{\frac{2\Delta\omega}{\pi f_r}\left(1 + \frac{k}{2}\right)}\right]_{-\pi k/2}^{\pi k/2}\right.$$

$$\left. + \left[\frac{\sin\frac{\Delta\omega}{\pi f_r}\left(1 - \frac{k}{2}\right)x}{\frac{2\Delta\omega}{\pi f_r}\left(1 - \frac{k}{2}\right)} - \frac{\sin\frac{\Delta\omega}{\pi f_r}\left(1 + \frac{k}{2}\right)x}{\frac{2\Delta\omega}{\pi f_r}\left(1 + \frac{k}{2}\right)}\right]_{-\pi k/2}^{\pi k/2}\right\} \quad (47)$$

$$= \frac{1}{\frac{2\Delta\omega}{f_r}\left(1 - \frac{k}{2}\right)}\left[\sin\frac{\Delta\omega}{\pi f_r}\left(1 - \frac{k}{2}\right)x\right]_{-\pi k/2}^{\pi k/2} \quad (48)$$

$$= \frac{1}{\frac{2\Delta\omega}{f_r}\left(1 - \frac{k}{2}\right)}\left\{\sin\frac{\Delta\omega k}{2f_r}\left(1 - \frac{k}{2}\right) + \sin\left[\frac{\Delta\omega k}{2f_r}\left(1 - \frac{k}{2}\right)\right]\right\} \quad (49)$$

$$= \frac{k}{2M}\ \sin M \quad (50)$$

$$A_0 = A_{01} + A_{02} + A_{03} = \frac{\sin M}{M} \cdot \tag{51}$$

Solving for B_n,

$$B_n = \frac{1}{\pi} \int_{-\pi}^{\pi} F(x) \cos nx \, dx \tag{52}$$

$$B_{n1} = \frac{1}{\pi} \int_{-\pi}^{-\pi k/2} \cos\left(\frac{\Delta\omega k}{2f_r} + \frac{\Delta\omega kx}{2\pi f_r}\right) \cos nx \, dx \tag{53}$$

$$= \frac{1}{\pi} \left\{ \cos \frac{\Delta\omega k}{2f_r} \int_{-\pi}^{-\pi k/2} \cos \frac{\Delta\omega kx}{2\pi f_r} \cos nx \, dx - \sin \frac{\Delta\omega k}{2f_r} \int_{-\pi}^{-\pi k/2} \sin \frac{\Delta\omega kx}{2\pi f_r} \cos nx \, dx \right\} \tag{54}$$

$$= \frac{1}{\pi} \left\{ \cos a \int_{-\pi}^{-\pi k/2} \cos a \frac{x}{\pi} \cos nx \, dx - \sin a \int_{-\pi}^{-\pi k/2} \sin a \frac{x}{\pi} \cos nx \, dx \right\} \tag{55}$$

$$B_{n1} = \frac{1}{\pi} \left\{ \cos a \left[\frac{\sin\left(\frac{a}{\pi} - n\right) x}{2\left(\frac{a}{\pi} - n\right)} + \frac{\sin\left(\frac{a}{\pi} + n\right) x}{2\left(\frac{a}{\pi} + n\right)} \right]_{-\pi}^{-\pi k/2} + \frac{1}{2} \sin a \left[\frac{\cos\left(\frac{a}{\pi} - n\right) x}{\left(\frac{a}{\pi} - n\right)} + \frac{\cos\left(\frac{a}{\pi} + n\right) x}{\left(\frac{a}{\pi} + n\right)} \right]_{-\pi}^{-\pi k/2} \right\} \tag{56}$$

$$= \frac{1}{2\pi} \left\{ \cos a \left[\frac{-\sin (a - n\pi) \frac{k}{2} + \sin (a - n\pi)}{\left(\frac{a}{\pi} - n\right)} + \frac{-\sin (a + n\pi) \frac{k}{2} + \sin (a + n\pi)}{\left(\frac{a}{\pi} + n\right)} \right] \right.$$

$$+ \sin a \left[\frac{\cos (a - n\pi) \frac{k}{2} - \cos (a - n\pi)}{\left(\frac{a}{\pi} - n\right)} + \frac{\cos (a + n\pi) \frac{k}{2} - \cos (a + n\pi)}{\left(\frac{a}{\pi} + n\right)} \right] \Bigg\} \tag{57}$$

$$B_{n1} = \frac{1}{2\pi} \Bigg\{ \frac{1}{\left(\frac{a}{\pi} + n\right)} \Big[\cos a \sin a \cos n\pi + \cos a \cos a \sin n\pi$$
$$- \cos a \sin \frac{ak}{2} \cos \frac{n\pi k}{2} - \cos a \cos \frac{ak}{2} \sin \frac{n\pi k}{2}$$
$$+ \sin a \cos \frac{ak}{2} \cos \frac{n\pi k}{2} - \sin a \sin \frac{ak}{2} \sin \frac{n\pi k}{2}$$
$$- \sin a \cos a \cos n\pi + \sin a \sin a \sin n\pi \Big]$$
$$+ \frac{1}{\left(\frac{a}{\pi} - n\right)} \Big[\cos a \sin a \cos n\pi - \cos a \cos a \sin n\pi$$
$$- \cos a \sin \frac{ak}{2} \cos \frac{n\pi k}{2} + \cos a \cos \frac{ak}{2} \sin \frac{n\pi k}{2}$$
$$+ \sin a \cos \frac{ak}{2} \cos \frac{n\pi k}{2} + \sin a \sin \frac{ak}{2} \sin \frac{n\pi k}{2}$$
$$+ \sin a \cos \frac{ak}{2} \cos \frac{n\pi k}{2} + \sin a \sin \frac{ak}{2} \sin \frac{n\pi k}{2}$$
$$- \sin a \cos a \cos n\pi - \sin a \sin a \sin n\pi \Big] \Bigg\} \tag{58}$$

$$B_{n1} = \frac{1}{2} \left\{ \frac{1}{a + n\pi} \left[\cos \frac{n\pi k}{2} \cdot \sin \left[a \left(1 - \frac{k}{2}\right)\right] - \sin \frac{n\pi k}{2} \cdot \cos \left[a \left(1 - \frac{k}{2}\right)\right]\right] + \frac{1}{a - n\pi} \left[\cos \frac{n\pi k}{2} \sin \left[a \left(1 - \frac{k}{2}\right)\right] + \sin \frac{n\pi k}{2} \cos \left[a \left(1 - \frac{k}{2}\right)\right]\right] \right\} \tag{59}$$

$$= \frac{1}{2} \left\{ \frac{1}{\frac{M}{\left(1 - \frac{k}{2}\right)} + n\pi} \left[\sin \left(M - \frac{n\pi k}{2} \right) \right] + \frac{1}{\frac{M}{\left(1 - \frac{k}{2}\right)} - n\pi} \left[\sin \left(M + \frac{n\pi k}{2} \right) \right] \right\} . \tag{60}$$

$$B_{n1} = B_{n3}$$

$$B_{n1} + B_{n3} = \left\{ \frac{\sin\left[M - \frac{n\pi k}{2}\right]}{\frac{M}{\left(1 - \frac{k}{2}\right)} + n\pi} + \frac{\sin\left[M + \frac{n\pi k}{2}\right]}{\frac{M}{\left(1 - \frac{k}{2}\right)} - n\pi} \right\} . \tag{61}$$

$$B_{n2} = \frac{1}{\pi} \int_{-\pi k/2}^{\pi k/2} \cos \frac{\Delta\omega\left(1 - \frac{k}{2}\right)x}{\pi f_r} \cdot \cos nx \, dx \tag{62}$$

$$= \frac{1}{2\pi} \left[\frac{\sin \frac{\left[\Delta\omega\left(1 - \frac{k}{2}\right) - n\right]x}{\pi f_r}}{\frac{\Delta\omega\left(1 - \frac{k}{2}\right)}{\pi f_r} - n} + \frac{\sin \frac{\left[\Delta\omega\left(1 - \frac{k}{2}\right) + n\right]x}{\pi f_r}}{\frac{\Delta\omega\left(1 - \frac{k}{2}\right)}{\pi f_r} + n} \right]_{-\pi k/2}^{\pi k/2} \tag{63}$$

$$= \frac{k}{2} \left\{ \frac{\sin\left[a\left(1 - \frac{k}{2}\right) - \frac{n\pi k}{2}\right]}{\left[a\left(1 - \frac{k}{2}\right) - \frac{n\pi k}{2}\right]} + \frac{\sin\left[a\left(1 - \frac{k}{2}\right) + \frac{n\pi k}{2}\right]}{\left[a\left(1 - \frac{k}{2}\right) + \frac{n\pi k}{2}\right]} \right\} \tag{64}$$

$$= \frac{k}{2} \left\{ \frac{\sin\left[M - \frac{n\pi k}{2}\right]}{\left[M - \frac{n\pi k}{2}\right]} + \frac{\sin\left[M + \frac{n\pi k}{2}\right]}{\left[M + \frac{n\pi k}{2}\right]} \right\} \tag{65}$$

$$B_n = \left[\frac{1}{\frac{M}{\left(1 - \frac{k}{2}\right)} + n\pi} + \frac{\frac{k}{2}}{\left[M - \frac{n\pi k}{2}\right]} \right] \sin\left[M - \frac{n\pi k}{2}\right]$$

$$+ \left[\frac{1}{\frac{M}{\left(1 - \frac{k}{2}\right)} - n\pi} + \frac{\frac{k}{2}}{\left[M + \frac{n\pi k}{2}\right]} \right] \sin\left[M + \frac{n\pi k}{2}\right] . \tag{66}$$

Investigation of sin $(\Delta\omega \int f(t)\,dt)$

$$G(x)_2 = \sin \frac{\Delta\omega \left(1 - \frac{k}{2}\right)x}{\pi f_r} \quad . \tag{67}$$

Since $G(x) = -G(-x)$ there are no cosine terms and no d-c term.

$$A_{n2} = \frac{1}{\pi} \int_{-\pi k/2}^{\pi k/2} \sin \frac{\Delta\omega\left(1 - \frac{k}{2}\right) x}{\pi f_r} \sin nx\, dx. \tag{68}$$

From equations (62) - (65) it follows that

$$A_{n2} = \frac{k}{2}\left\{ \frac{\sin\left[M - \frac{n\pi k}{2}\right]}{\left[M - \frac{n\pi k}{2}\right]} - \frac{\sin\left[M + \frac{n\pi k}{2}\right]}{\left[M + \frac{n\pi k}{2}\right]} \right\} \tag{69}$$

$$A_{n1} = \frac{1}{\pi} \int_{-\pi}^{-\pi k/2} - \sin\left(\frac{\Delta\omega k}{2f_r} + \frac{\omega k x}{2\pi f_r}\right) \sin nx\, dx \tag{70}$$

$$= -\frac{1}{\pi}\left\{ \sin \frac{\Delta\omega k}{2f_r} \int_{-\pi}^{-\pi k/2} \cos \frac{\Delta\omega k x}{2\pi f_r} \sin nx\, dx \right.$$

$$\left. + \cos \frac{\Delta\omega k}{2f_r} \int_{-\pi}^{-\pi k/2} \sin \frac{\Delta\omega k x}{2\pi f_r} \sin nx\, dx \right\} \tag{71}$$

$$= \frac{-1}{2\pi} \left\{ \sin \frac{\Delta\omega k}{2f_r} \left[\cos \frac{\left(\frac{\Delta\omega k}{2\pi f_r} - n\right) x}{\frac{\Delta\omega k}{2\pi f_r} - n} - \cos \frac{\left(\frac{\Delta\omega k}{2\pi f_r} + n\right) x}{\frac{\Delta\omega k}{2\pi f_r} + n} \right]_{-\pi}^{-\pi k/2} \right.$$

$$\left. + \cos \frac{\Delta\omega k}{2f_r} \left[\frac{\sin\left(\frac{\Delta\omega k}{2\pi f_r} - n\right) x}{\frac{\Delta\omega k}{2\pi f_r} - n} - \frac{\sin\left(\frac{\Delta\omega k}{2\pi f_r} + n\right) x}{\frac{\Delta\omega k}{2\pi f_r} + n} \right]_{-\pi}^{-\pi k/2} \right\} \tag{72}$$

$$= -\frac{1}{2}\left\{\sin a\left[\frac{\cos\left(\frac{ak}{2}-\frac{n\pi k}{2}\right)-\cos(a-n\pi)}{a-n\pi}-\frac{\cos\left(\frac{ak}{2}+\frac{n\pi k}{2}\right)+\cos(a+n\pi)}{a+n\pi}\right]\right.$$

$$\left.+\cos a\left[\frac{-\sin\left(\frac{ak}{2}-\frac{n\pi k}{2}\right)+\sin(a-n\pi)}{a-n\pi}+\frac{\sin\left(\frac{ak}{2}+\frac{n\pi k}{2}\right)-\sin(a+n\pi)}{a+n\pi}\right]\right\} \quad (73)$$

$$= -\frac{1}{2}\left\{\frac{1}{a+n\pi}\left[-\sin a\cos\frac{ak}{2}\cos\frac{n\pi k}{2}+\sin a\sin\frac{ak}{2}\sin\frac{n\pi k}{2}\right.\right.$$

$$\left.+\cos a\sin\frac{ak}{2}\cos\frac{n\pi k}{2}+\cos a\cos\frac{ak}{2}\sin\frac{n\pi k}{2}\right]$$

$$+\frac{1}{a-n\pi}\left[\sin a\cos\frac{ak}{2}\cos\frac{n\pi k}{2}+\sin a\sin\frac{ak}{2}\sin\frac{n\pi k}{2}\right.$$

$$\left.\left.-\cos a\sin\frac{ak}{2}\cos\frac{n\pi k}{2}+\cos a\cos\frac{ak}{2}\sin\frac{n\pi k}{2}\right]\right\} \quad (74)$$

$$= -\frac{1}{2}\left\{\frac{1}{a+n\pi}\left[-\cos\frac{n\pi k}{2}\sin a\left(1-\frac{k}{2}\right)+\sin\frac{n\pi k}{2}\cos a\left(1-\frac{k}{2}\right)\right]\right.$$

$$\left.+\frac{1}{a-n\pi}\left[\cos\frac{n\pi k}{2}\sin a\left(1-\frac{k}{2}\right)+\sin\frac{n\pi k}{2}\cos a\left(1-\frac{k}{2}\right)\right]\right\} \quad (75)$$

$$= +\frac{1}{2}\left\{\frac{1}{\frac{M}{\left(1-\frac{k}{2}\right)}+n\pi}\left[\sin\left(M-\frac{n\pi k}{2}\right)\right]-\frac{1}{\frac{M}{\left(1-\frac{k}{2}\right)}-n\pi}\left[\sin\left(M+\frac{n\pi k}{2}\right)\right]\right\}. \quad (76)$$

$$A_{n1} = A_{n3}$$

$$A_{n1} + A_{n3} = \left\{\frac{\sin\left(M-\frac{n\pi k}{2}\right)}{\frac{M}{\left(1-\frac{k}{2}\right)}+n\pi}-\frac{\sin\left(M+\frac{n\pi k}{2}\right)}{\frac{M}{\left(1-\frac{k}{2}\right)}-n\pi}\right\} \quad (77)$$

$$A_n = \left[\frac{1}{\frac{M}{\left(1-\frac{k}{2}\right)}-n\pi}+\frac{\frac{k}{2}}{M-\frac{n\pi k}{2}}\right]\sin\left[M-\frac{n\pi k}{2}\right]$$

$$- \left[\frac{1}{\dfrac{M}{\left(1 - \dfrac{k}{2}\right)} - n\pi} + \frac{\dfrac{k}{2}}{M + \dfrac{n\pi k}{2}} \right] \sin \left[M + \frac{n\pi k}{2} \right] \tag{78}$$

$$F(x) = \frac{\sin M}{M} + \sum_{n=1}^{n \to \infty} \left\{ \frac{M \sin \left[M - \dfrac{n\pi k}{2} \right]}{\left[M + n\pi \left(1 - \dfrac{k}{2}\right)\right]\left[M - \dfrac{n\pi k}{2} \right]} + \frac{M \sin \left[M + \dfrac{n\pi k}{2} \right]}{\left[M - n\pi \left(1 - \dfrac{k}{2}\right)\right]\left[M + \dfrac{n\pi k}{2} \right]} \right\} \cos nx \tag{79}$$

$$G(x) = \sum_{n=1}^{n \to \infty} \left\{ \frac{M \sin \left[M - \dfrac{n\pi k}{2} \right]}{\left[M + n\pi \left(1 - \dfrac{k}{2}\right)\right]\left[M - \dfrac{n\pi k}{2} \right]} - \frac{M \sin \left[M + \dfrac{2\pi k}{2} \right]}{\left[M - n\pi \left(1 - \dfrac{k}{2}\right)\right]\left[M + \dfrac{n\pi k}{2} \right]} \right\} \sin nx \tag{80}$$

Let

$$F(x) = \alpha + (\beta + \gamma) \cos nx, \text{ and} \tag{81}$$

$$G(x) = (\beta - \gamma) \sin nx, \text{ where } x = 2\pi f_r t. \tag{82}$$

Substituting back in Equation (7)

$$e = \sin \omega_0 t \, [\alpha + (\beta + \gamma)] \cos nx + \cos \omega_0 t \, (\beta - \gamma) \sin nx = \alpha \sin \omega_0 t \tag{83}$$

$$+ \frac{(\beta + \gamma)}{2} \sin (\omega_0 t + nx) + \frac{(\beta + \gamma)}{2} \sin (\omega_0 t - nx)$$

$$+ \frac{(\beta - \gamma)}{2} \sin (\omega_0 t + nx) - \frac{(\beta - \gamma)}{2} \sin (\omega_0 t - nx) \tag{84}$$

$= \alpha \sin \omega_0 t$ -- carrier

$+ \beta \sin (\omega_0 t + nx)$ -- upper side band

$+ \gamma \sin (\omega_0 t - nx)$ -- lower side band (85)

We now obtain the final equation for the rectangular pulse output spectrum

$$e = \frac{\sin M}{M} \sin \omega_o t + \sum_{n=1}^{n=\infty} \frac{M}{N + n\pi} \cdot \frac{\sin N}{N} \cdot \sin (\omega_o + n2\pi f_r)t$$

$$+ \sum_{n=1}^{n=\infty} \frac{N}{P - n\pi} \cdot \frac{\sin P}{P} \cdot \sin (\omega_o - n2\pi f_r)t \qquad (86)$$

Where

$$M = \frac{k\Delta\omega}{2f_r} (1 - k)$$

$$N = M - \frac{n\pi k}{2}$$

$$P = M + \frac{n\pi k}{2}$$

$$\omega_o = 2\pi f_o$$

f_o = carrier frequency

f_r = pulse repetition frequency

Evaluation of lower side band when $M = n\pi \left(1 - \frac{k}{2}\right)$

$$e = \frac{M \sin\left[M + \frac{n\pi k}{2}\right]}{\left[M - n\pi\left(1 - \frac{k}{2}\right)\right]\left[M + \frac{n\pi k}{2}\right]} \qquad (87)$$

differentiating numerator and denominator with respect to M

$$= \frac{M \cos\left[M + \frac{n\pi k}{2}\right] + \sin\left[M + \frac{n\pi k}{2}\right]}{\left[M - n\pi\left(1 - \frac{k}{2}\right)\right] + \left[M + \frac{n\pi k}{2}\right]} \qquad (88)$$

$$= \frac{n\pi\left(1 - \frac{k}{2}\right) \cos n\pi}{n\pi} \qquad (89)$$

$$= \left(1 - \frac{k}{2}\right) \qquad (90)$$

Square Wave Keying

Special case where k = 1

$$e = \frac{\sin M}{M} \sin \omega_0 t$$

$$+ \frac{M}{M + \frac{n\pi}{2}} \left[\frac{\sin \left[M - \frac{n\pi}{2} \right]}{\left[M - \frac{n\pi}{2} \right]} \right] \sin (\omega_0 + n2\pi f_r)t$$

$$+ \frac{M}{M - \frac{n\pi}{2}} \left[\frac{\sin \left[M + \frac{n\pi}{2} \right]}{\left[M + \frac{n\pi}{2} \right]} \right] \sin (\omega_0 - n2\pi f_r)t \qquad (91)$$

$$= \frac{\sin M}{M} \sin \omega_0 t$$

$$+ \frac{M}{M^2 - \left(\frac{n\pi}{2}\right)^2} \cos M \left[\sin (\omega_0 - n2\pi f_r)t - \sin (\omega_0 t + n2\pi f_r)t\right]$$

$$+ \frac{2M}{M^2 - \left(\frac{n\pi}{2}\right)^2} \sin M \left[\sin (\omega_0 - n2\pi f_r)t + \sin (\omega_0 + n2\pi f_r)t\right] \qquad (92)$$

where $M = \frac{\Delta\omega}{4f_r} = \frac{\pi}{2} m$ and n = order of side band.

Letting $M = \frac{\pi}{2} m$ where $m = \frac{\Delta f}{f_r}$ we obtain the final equation for the square wave spectrum.

$$e = \frac{2}{\pi} \left\{ \frac{\sin \left(\frac{\pi}{2} m\right)}{m} \sin \omega_0 t \right.$$

$$+ \frac{m}{m^2 - n^2} \cos \left(\frac{\pi}{2} m\right) \left[\sin (\omega_0 - n2\pi f_r)t - \sin (\omega_0 + n2\pi f_r)t\right] \text{ (where n is odd)}$$

$$\left. + \frac{m}{m^2 - n^2} \sin \left(\frac{\pi}{2} m\right) \left[\sin (\omega_0 - n2\pi f_r)t + \sin (\omega_0 + n2\pi f_r)t\right] \right\} \text{ (where n is even)} \qquad (93)$$

Equation (86) gives the amplitude and position of the various side frequencies for a rectangular pulse wave. To illustrate the manner in which this equation can be used, a particular example is worked out for a one millisecond pulse with a repetition rate of 25 cycles per second. The results are shown in Figures 3 to 7 where they can be compared with the frequency spectrum of an amplitude modulated carrier using the same one millisecond pulses as modulation. The method of obtaining this a-m spectrum will not be included since it is covered elsewhere.* It should be interesting to note that at the wider shifts i.e., larger M's, the spectrum begins to approach the case of two amplitude-modulated carriers; however, at the smaller shifts this relation no longer holds and there tends to be a concentration of the energy between the shift limits with the side frequency amplitudes dropping off rapidly outside the shift limits. The particular case where k = 1 or square wave keying is covered by equations (92) and (93). The output spectra for several different modulation indices are presented in Figure 8 where they are compared with the corresponding case of a-m square-wave keying.

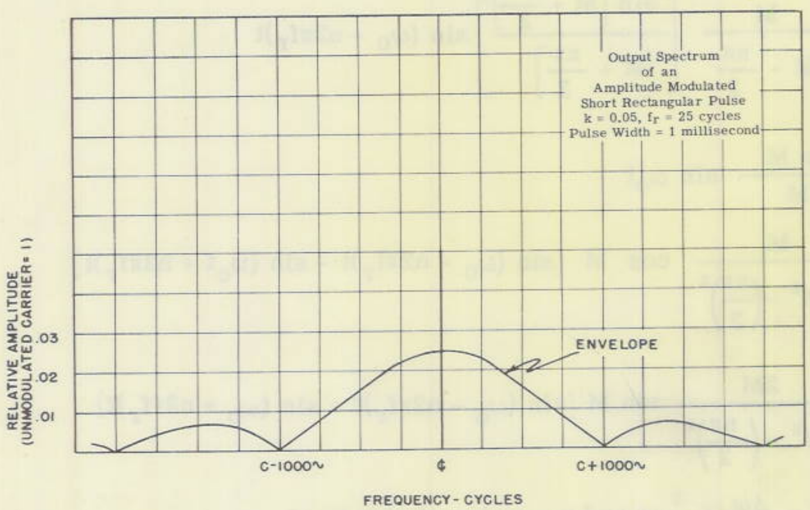


Figure 3

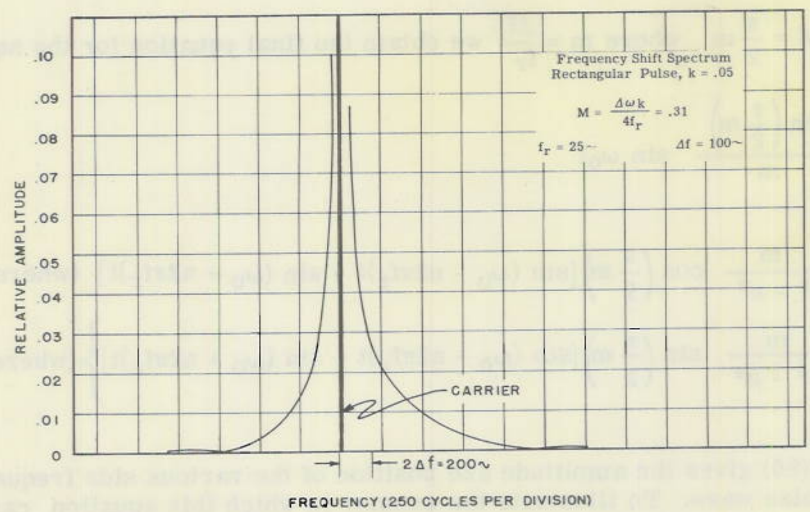
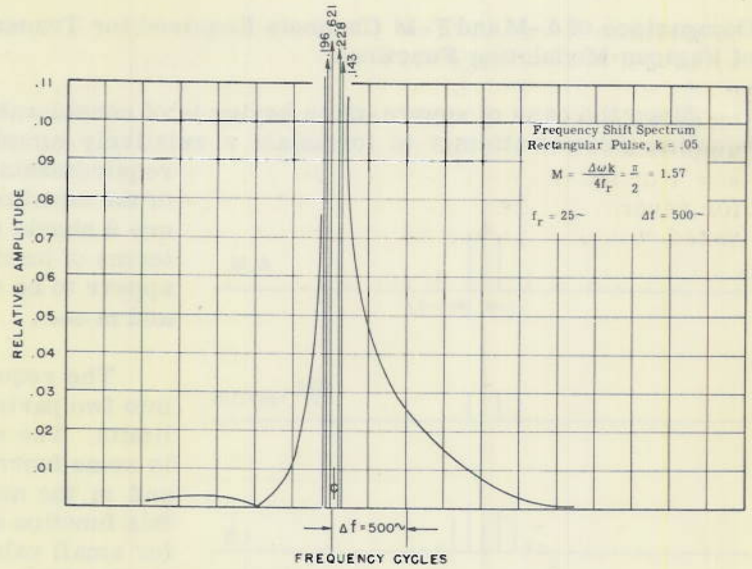


Figure 4

* Terman, F. E., "Radio Engineer Handbook," McGraw-Hill Book Co., New York, 1943, p. 532; B. Salzberg, "Pulse Transmission Spectra," Naval Research Laboratory, June 1941

Figure 5



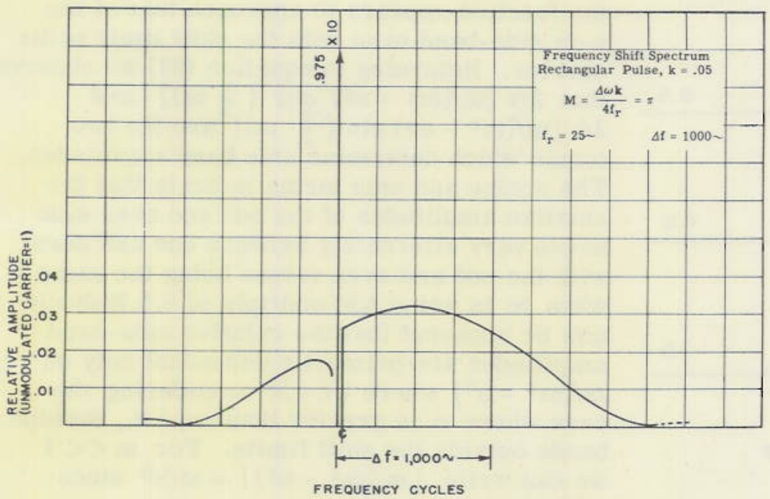
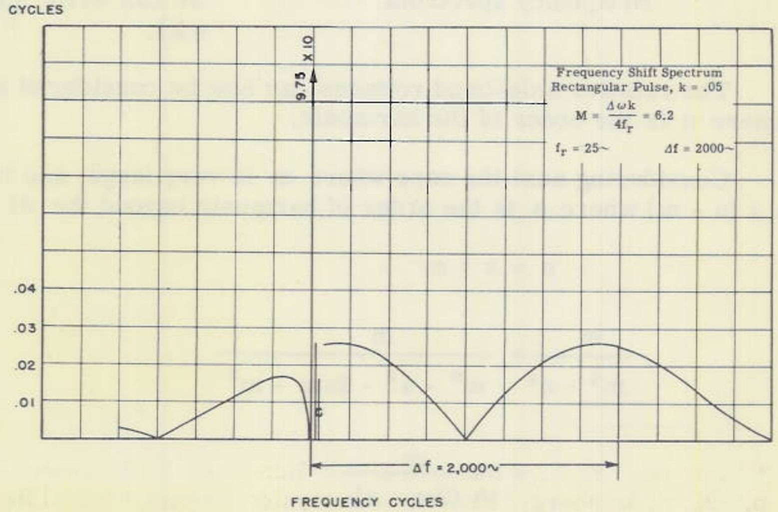


Figure 6

Figure 7



Comparison of A-M and F-M Channels Required for Transmission of Various Modulating Functions

Since the case of square-wave keying is of considerable interest we shall investigate equation (93) and attempt to formulate a relatively simple expression for bandwidth requirements. Before attempting any solution of the equation a careful examination of Figure 8 should reveal what we might expect in terms of bandwidth as well as what conditions appear to be approached at the limits $m \to 0$ and $m \to \infty$.

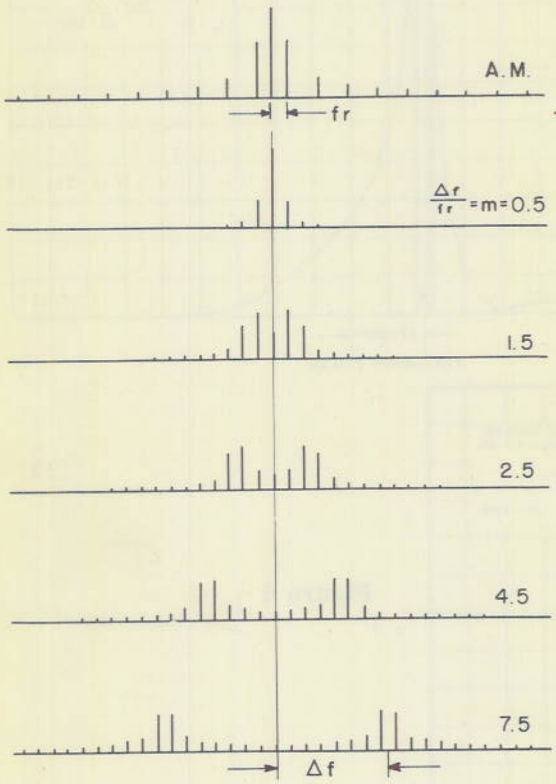


Figure 8 - Square-Wave Frequency Spectrum

The required bandwidth can be broken into two parts. The first is $2\Delta f$ or the shift limits. The second is $2f(n, m)$, where $f(n, m)$ is some function of n, the harmonic number, and m the modulation index. The value of this function appears to drop off rather rapidly for small values of m but, at the larger m's the function appears to approach that of the a-m side-band case with the shift limit as the carrier. Returning to equation (93) we observe that $2/\pi\,[m/(m^2 - n^2)\cos(\frac{\pi}{2}m)]$ and $2/\pi\,[m/(m^2 - n^2)\sin(\frac{\pi}{2}m)]$ are the two terms which determine side band amplitudes. The cosine and sine terms indicate that the relative amplitudes of the odd and even side bands vary alternately between one and zero with the odd and even values being the same when m is some odd multiple of 0.5. It should now be apparent that the relative side-band amplitudes are primarily dependent only on $m/(m^2 - n^2)$ where we are considering the case where n is greater than m, i.e., the side bands outside the shift limits. For $m << 1$ we can write $|m/(m^2 - n^2)| = m/n^2$ since $n \geq 1$.

The relative side-band voltages can now be considered as being distributed as $1/n^2$ where n is the order of the harmonic.

Considering next the case where m is very large, and introducing the relationship $a = (n - m)$ where a is the order of harmonic beyond the Δf limit, we obtain:

$$n = a + m \tag{94}$$

$$\frac{m}{m^2 - n^2} = \frac{m}{m^2 - a^2 - 2am - m^2} \tag{95}$$

$$= \frac{m}{-a\,(2m + a)} \tag{96}$$

Since $m \to \infty$ and a is considered as relatively small we approach

$$\left| \frac{m}{m^2 - n^2} \right| = \frac{1}{2a} \tag{97}$$

and the relative values equal 1/a where a is the order of the harmonic beyond the Δf limit point.

These considerations give some idea of the limits approached by the output spectra for various shifts. It is often desirable to be able to obtain rapidly the bandwidth beyond which the side-band amplitudes will be less than a given ratio to the unmodulated carrier. From equation (93) we obtain the following ratio of side-band amplitude to the unmodulated carrier amplitude:

$$R = \frac{\text{S. B. Ampl.}}{\text{Carrier Amp.}} = \frac{m}{m^2 - n^2} \tag{98}$$

Since we are interested in points where $n > m$ and R is taken as a positive ratio, we can rewrite (98) as

$$R = -\frac{m}{m^2 - n^2} \tag{99}$$

Solving for n, we obtain

$$n = \sqrt{m^2 + \frac{m}{R}} \tag{100}$$

If we now wish to get total bandwidth we multiply n by $2f_r$.

$$\text{B.W.} = 2f_r \sqrt{m^2 + \frac{m}{R}}\ , \tag{101}$$

where $m = (\Delta f / f_r)$ and R = ratio of side-band amplitude to unmodulated carrier.

As an example, consider a square-wave multiplex circuit where essentially square frequency-shift keying is employed.

$$f_r = 60 \text{ cycles},\ \Delta f = \pm\ 425 \text{ cycles}$$

$$m = \frac{\Delta f}{f_r} = \frac{425}{60} \doteq 7.$$

Assume we are interested in a 60-db voltage ratio, R = .001

$$\text{B.W.} = 120\sqrt{49 + 7000} = 10.1 \text{ kc} \tag{102}$$

As an interesting comparison we might take the same conditions with amplitude modulation and again assume essentially square wave keying. For square-wave a-m keying the first side-band amplitude should be 32 percent of the unmodulated full-on carrier and the relative amplitudes of the sideband decrease directly with n where n is again the order of the harmonic. From these observations we obtain

$$R = \frac{.32}{n} \tag{103}$$

$$n = \frac{.32}{R}$$

$$B.W. = \frac{2f_r \times .32}{R} = \frac{.64 f_r}{R} \;. \tag{104}$$

Substituting in the previous values of f_r and R we obtain

$$B.W. = .64 \times 60 \times 1000 = 38.4 \text{ kc} \tag{105}$$

It should be remembered that these values consider only the components introduced by the theoretical wave shape and that in practice the selectivity of the output circuits may reduce the bandwidth. On the other hand transients and other factors may increase the bandwidth considerably.

Minimum Bandwidth Required for Transmission of Intelligence

Equation (101) can be used to determine rapidly the bandwidth required for channel spacing or interference. Certain factors make it impossible to use this simple formula for relative side-band amplitudes when obtaining bandwidths required for the transmission of intelligence. If an exact solution is required, it is necessary to find the spectrum which will be produced for the specific case. The modulation function is then reformed by recombination of the side bands passed by the circuit or filter employed with due regard to phase and amplitude. This process is practical for many a-m waves; however, when f-m waves are concerned the process becomes very difficult and in many cases impractical. Recombination can be accomplished by employing the graphical method outlined in the following section; however, this too may be somewhat cumbersome.

When specifying the degree of fidelity with which an a-m wave will be passed by a given circuit, the number of side bands and the ratio of the largest to the smallest side band is often used as a figure of merit. Admittedly this is not an exact solution and may be misleading if care is not taken in interpreting the results obtained.

The bandwidth required to transmit f-m waves with small deviation ratios can readily be obtained by returning to equations (11), (12), and (13). It can also be seen from these equations that the odd side bands in an f-m wave originate as follows:

$$\sin (\Delta\omega \int f(t)dt) = A_1 \sin \omega_a t + A_3 \sin 3\omega_a t + \ldots \tag{106}$$

where

$$f(t) = k_1 \cos \omega_a t + k_3 \cos 3\omega_a t + k_5 \cos 5\,\omega_a t + \ldots \tag{107}$$

hence

$$\sin\left(\frac{\Delta\omega k_1}{\omega_a} \sin \omega_a t + \frac{\Delta\omega k_3}{3\omega_a} \sin 3\omega_a t + \ldots \right) =$$

$$A_1 \sin \omega_a t + A_3 \sin 3\omega_a t + \ldots \tag{108}$$

Assume a very small shift where M, the maximum deviation in radians, is less than 0.4 radian. Since $\sin x \doteq x$ when x is less than 0.4, we can rewrite equation (108) as

$$\frac{\Delta\omega k_1}{\omega_a} \sin \omega_a t + \frac{\Delta\omega k_3}{3\omega_a} \sin 3\omega_a t + \ldots = A_1 \sin \omega_a t + A_3 \sin 3\omega_a t + \ldots \tag{109}$$

From this we conclude that

$A_n = \frac{\Delta\omega k_n}{\omega_a n}$ or that the relative amplitude of the nth side band of an f-s signal of small shift is equal to (1/n) that of the corresponding a-m side band. It can also be seen from equation (109) that the bandwidth required for an f-s signal is the same as that required to pass an equivalent a-m signal provided the maximum deviation M is less than 0.4 radian.

Considerable difficulty is encountered when attempting to compare the bandwidths required for the transmission of intelligence by a-m and f-m carriers. This is probably due in part to the fact that different types of distortion are obtained when band limiting is applied to each of the signals. It has been shown that for small maximum deviations the f-m and a-m bandwidths required for intelligence are the same; however, for larger indices this is not true. When square-wave keying is considered, f-s waves with large indices appear to have frequency spectra very similar to the case of two a-m carriers alternately keyed. The bandwidth required for the a-m wave is equal to the frequency separation plus twice the highest modulating frequency. From these and other considerations an approximate formula for bandwidth required to transmit intelligence is given as

$$\text{B.W.} = 2(\Delta f + Mf_h) \tag{110}$$

where $2\Delta f$ is the total frequency shift and Mf_h is the highest modulating frequency required. It must be emphasized that this is an approximation; however, attempts at an exact solution indicate that it would be very complex and of little practical value. Experimental results included in a later section indicate that the bandwidths obtained by equation (110) are well within the limits required for design purposes.

While on the subject of band limiting it is interesting to consider briefly the process of reception and generalize on the effect of removing side-band components. The first process of reception is that of frequency discrimination or differentiation which produces an amplitude variation of the carrier. This signal is in turn detected or rectified and filtered so that the desirable detection products appear at the output. The process of differentiating can usually be performed mathematically on the side currents; however, the process of detection usually becomes very cumbersome and impractical. The formation of the modulation function at the detector can be considered as the beating together of the differentiated side currents. The summation of adjacent side-currents products produces the fundamental, and the summation of alternate side-current products produces the 2nd harmonic and so on. From these considerations it can be seen that the effect of band limiting of an f-m signal does not simply eliminate the components above a given frequency as is the case with a-m band limiting. There is instead an effect which reduces the amplitude of all of the components with the greatest reduction taking place at the higher frequencies and very little in general at the lower frequency components of the modulation function.

General Graphical Solution

The previous mathematical solution, although not extremely complex, is rather long and tedious. The question probably arises as to how complex the solution will become if a modulation function other than the rather simple rectangular pulse is considered. Returning briefly to equations (1) to (14), it can be seen that the operation up to equation (11) can be performed for most waves. When analyzing cos ($\Delta\omega \int f(t)dt$) and sin ($\Delta\omega \int f(t)dt$), the integration can normally be performed; however, the difficulty arises when attempting to break the sine and cosine functions into sine and cosine series by means of Fourier analysis. When the modulation function is a sine or cosine, these series expansions can be obtained readily by means of Bessel functions; however, this method becomes extremely laborious when most non-sinusoidal modulating functions are considered.

After observing the functions plotted in Figure 2, it becomes rather obvious that the operations could be performed graphically. The resulting sine and cosine functions can be broken up into simple sine and cosine terms by means of graphical analysis such as are described by Bryant, Correll and Johnson.* If a mechanical harmonic analyzer is available, it will provide a rapid means of analyzing recurrent wave forms.

Analysis of a Triangular F-M Wave

The triangular modulating function f(t) is plotted in Figure 9. If accurate results are desired, the modulating function should be drawn to a fairly large scale. The $\int f(t)dt$ is next drawn from f(t). The integral curve has a constant multiplier $\Delta\omega$ which determines the modulation index and maximum deviation M. The vertical scale of the integral curve will vary with M. It should be noted here that M is the maximum value of $\Delta\omega \int f(t)dt$ and also that it is in radians. The $\int f(t)dt$ maximum will have a particular value for each

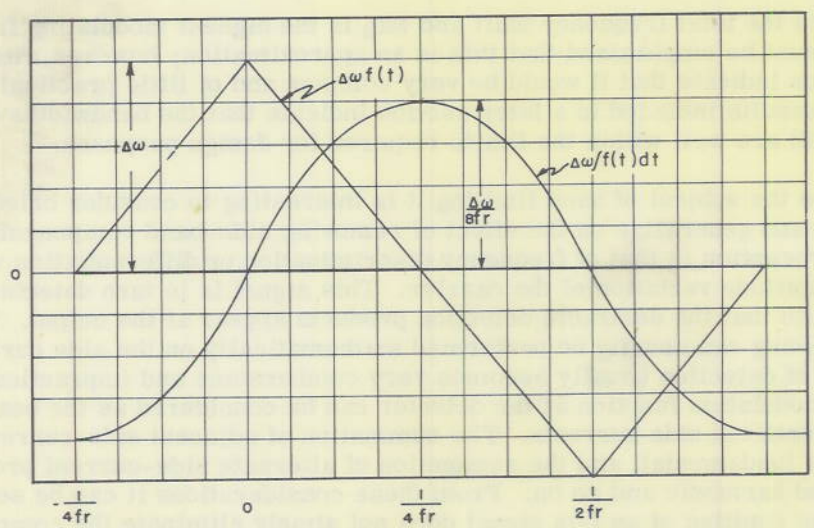


Figure 9 - Modulating Function and Integral

* Bryant, Correll and Johnson, "Alternating Current Circuits," McGraw-Hill Book Co., N. Y., 1939, pp. 461-491

wave shape. In this case it is ($1/8f_r$). The cosine and sine of the function can now be obtained by first preparing a linear scale in radians or degrees with sufficient divisions to permit accurate plotting. The scale must go from zero to M and must be long enough to go from the zero axis to the maximum of the integral curve. If the scale is longer, as it often will be, the scale is tipped slightly so that the zero is on the zero axis and the M employed intersects a horizontal line passing through the maximum point of the integral curve. Generally, it is desirable to do this all in tabular form up to the sine and cosine curves. The desired curves can now be plotted with the aid of a table of sine and cosine functions. The sine and cosine curves are plotted together for each value of M to facilitate harmonic analysis. (See Figures 10 to 13.) The curves are analyzed with a scale of one being a maximum and the "A" and "B" coefficients of the sine and cosine terms are those given in equation (12). Before plotting the output spectrum, the operations indicated by equation (14) should be performed. The output spectra for several different values of M are plotted in Figure 14.

The graphical method provides a means of visualizing why sine-wave frequency modulation requires more than one side band on either side of the carrier and also how the amplitude of the carrier is reduced and goes through zero at various deviations.

Another obvious use of the graphical analysis is to determine the amount of distortion produced by removing some of the side frequency components, reducing their relative amplitudes or shifting their phase.

The triangular wave of Figure 9 will be used as an example of the method used in reforming a wave after band limiting. The frequency spectra for the various ratios of ($\Delta f/f_r$) are shown in Figure 14 and the first example will be for m = 4.5. A theoretical filter which eliminates all but the first five side bands on each side of the carrier is considered. Recombination steps and the resulting triangular wave are illustrated by Figure 15. Figures 16 and 17 are similar to Figure 15 but for different conditions. An interesting comparison can be made between the final curves of these three figures. The initial curve (Figure 9) has an amplitude of 2.5 and it can be seen that the band limiting reduces the amplitude besides changing the shape.

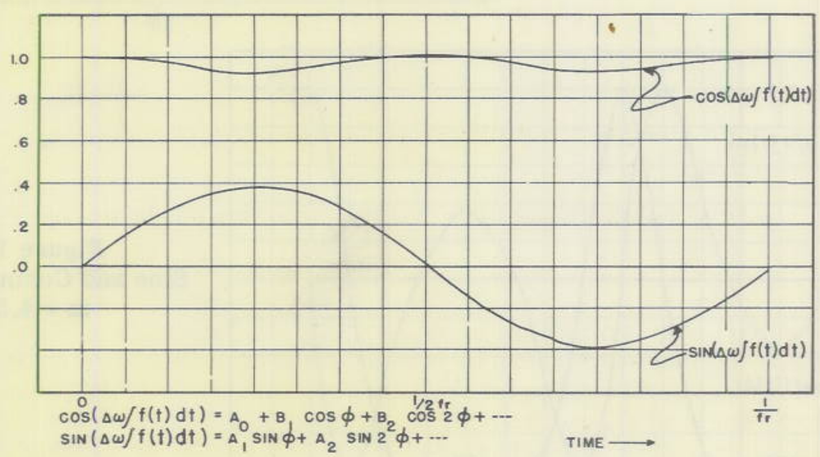


Figure 10 - Sine and Cosine Curves, m = 0.5

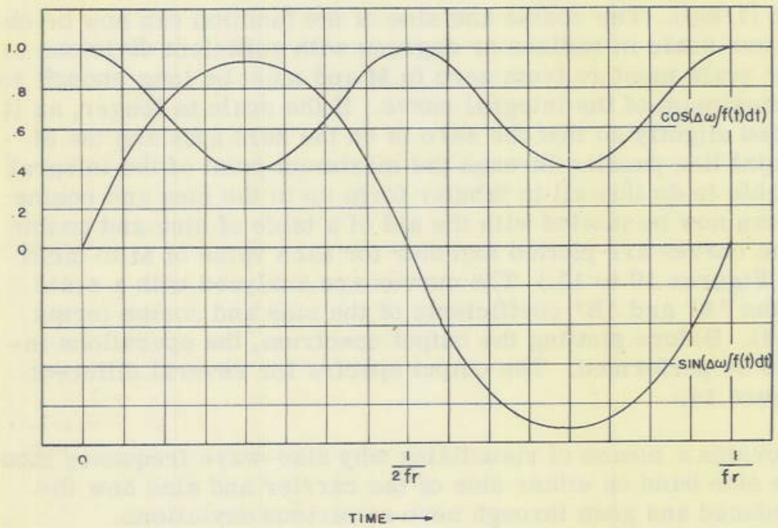


Figure 11
Sine and Cosine Curves
m = 1.5

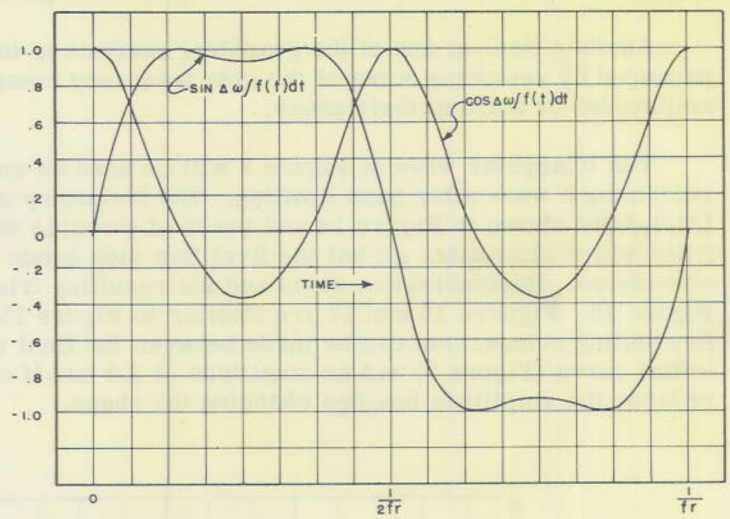


Figure 12
Sine and Cosine Curves
m = 2.5

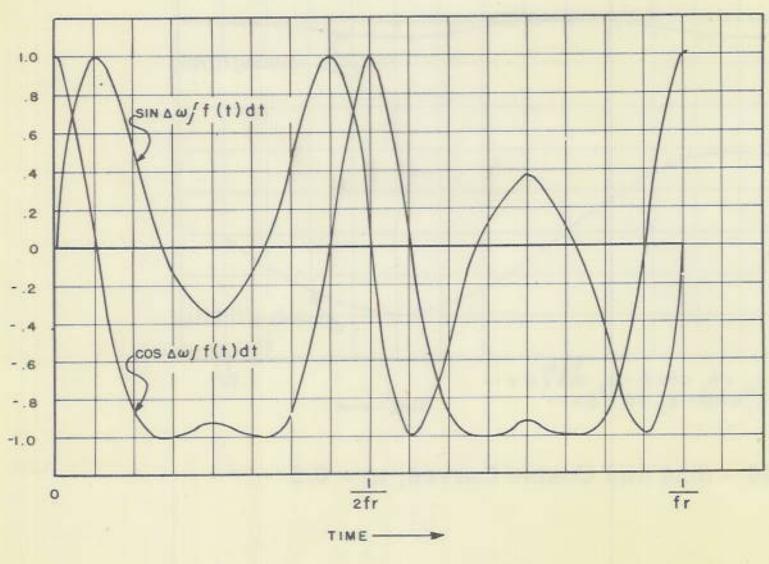


Figure 13
Sine and Cosine Curves
m = 4.5

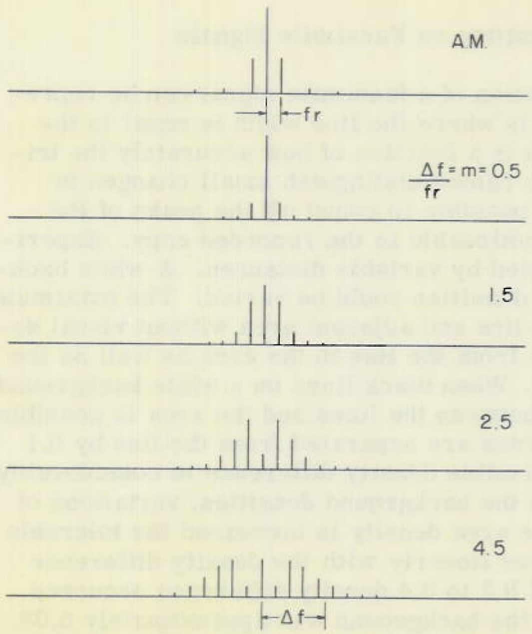


Figure 14
Triangular-Wave
Frequency Spectrum

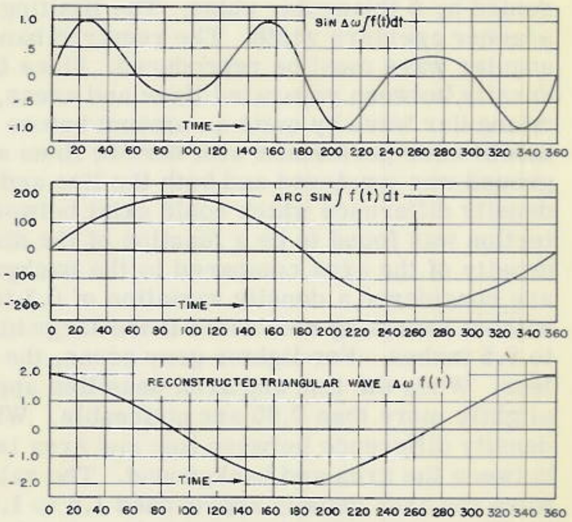


Figure 15
Reconstruction of Triangular F-M Wave
m = 4.5, Band Limited to First 5 Side Bands

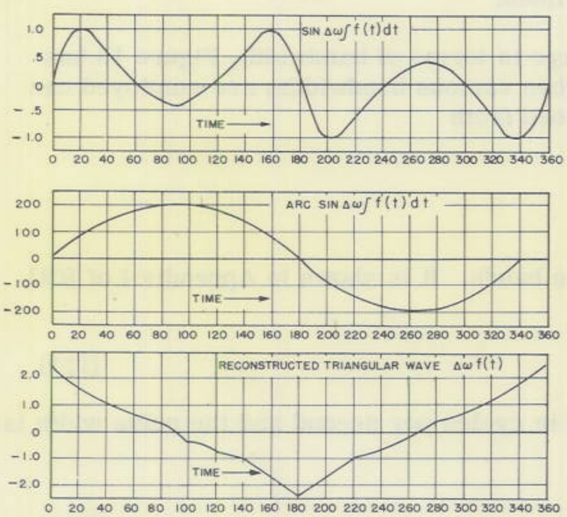


Figure 16
Reconstruction of Triangular F-M Wave
m = 4.5, Band Limited to First 7 Side Bands

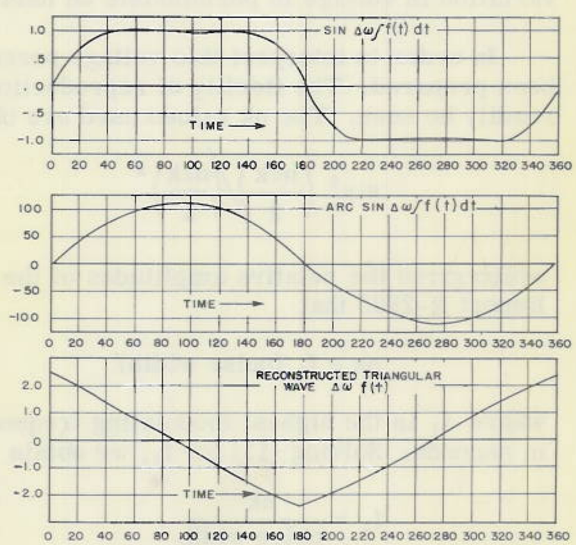


Figure 17
Reconstruction of Triangular F-M Wave
m = 2.5, Band Limited to First 7 Side Bands

Theoretical and Experimental Effect of Band Limiting on Facsimile Signals

It has been shown* that the modulation function of a facsimile signal can be represented by a triangular pulse. The limiting case is where the line width is equal to the scanner aperture width. The required bandwidth is a function of how accurately the triangular wave must be reproduced. Since the eye cannot distinguish small changes in density between separated lines and areas, it is possible to round off the peaks of the triangular wave by quite an amount before it is noticeable in the recorded copy. Experiments were performed with narrow lines separated by variable distances. A white background was employed and both the line and area densities could be varied. The maximum density difference which could exist between the line and adjacent area without visual detection was found to be a function of the distance from the line to the area as well as the density of the area compared to the background. When black lines on a white background are considered, a density variation of 0.3 to 0.4 between the lines and the area is possible without becoming noticeable if the large black areas are separated from the line by 0.1 to 0.5 inches. For lighter gray areas, the permissible density difference is considerably less. When the line and area densities approach the background densities, variations of slightly more than 0.05 are noticeable. When the area density is increased the tolerable density difference between line and area increases linearly with the density difference between the area and background. The values of 0.3 to 0.4 density difference occurred when the area density approached 1.6 to 1.8 and the background was approximately 0.03. The accuracy with which the pulse must be reproduced is also a function of the half-tone characteristics of the system involved. The equipment analyzed in NRL Report R-2885 will be used as an example since it is possible to check the theoretical conclusions with the results actually obtained.

Plate 9 of NRL Report R-2885 shows that approximately a straight-line relationship exists between the subject brightness and the output voltage. When the density differences given previously are converted into voltage, it is seen that approximately 10 to 15 percent variation in voltage is permissible on narrow lines.

In order to interpret this voltage percentage in terms of bandwidth, Figure 18 has been prepared. The fidelity of reproduction when various bandwidths are employed can readily be seen. The nk values used are obtained from

$$\sin^2\left(\frac{n\pi k}{2}\right)\Big/\left(\frac{n\pi k}{2}\right)^2$$

which gives the relative amplitudes of the side bands. It is shown in Appendix I of NRL Report R-2885 that

$$nk = f_1 \text{ (pulse width)} \qquad (111)$$

where f_1 is the highest modulating frequency in cycles per second and the pulse width is in seconds. Solving (11) for f_1, we obtain

$$f_1 = \frac{nk}{\text{pulse width}} . \qquad (112)$$

---

* NRL Report R-2885 "Determination of RC-120-B Facsimile Equipment Characteristics," 26 June 1946

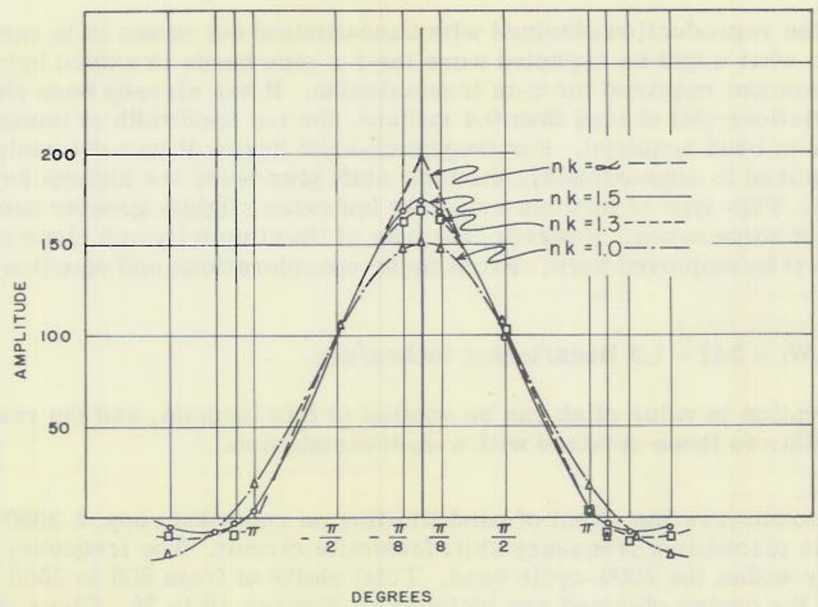


Figure 18 - Reconstructed Triangular A-M Pulse

From the manner in which the pulses are generated it is possible to state that

$$\text{pulse width} = \frac{2}{\text{lines/inch x inches/sec}} \qquad (113)$$

When this is substituted in (112) we obtain

$$f_1 = \frac{nk}{2} \text{ x lines/inch x inches/sec.} \qquad (114)$$

Returning to Figure 18 we see that the curves for nk = 1.5 and 1.3 are within the limits of possible error without visual degradation of the recorded copy.

Since a value of 1.3 will require the least bandwidth, this figure is chosen.

When high-quality half-tone pictures are considered, we obtain

$$f_1 = \frac{1.3}{2} \text{ x lines/inch x inches/sec.} \qquad (115)$$

The a-m bandwidth will be twice the highest modulating frequency. With black-and-white facsimile transmission it is possible to employ wave shaping at the receiver. This may permit the use of received waves which have suffered a greater amount of deformation from band limiting than is possible for high-quality half-tone pictures. If a lower definition picture is acceptable, it may be possible to employ values of nk from 1 to 1.3 with the higher values yielding a correspondingly higher quality picture. Experiments with the RC-120B facsimile equipment (NRL Report R-2885) have indicated that these conclusions are valid.

Frequency shift facsimile signals are more complex than those employing amplitude modulation; however, several cases have been investigated in an effort to obtain a relatively simple expression for bandwidth requirements. From Figures 15, 16 and 17 it appears that the reproduction obtained with band-limited f-s waves is in each case equal or superior to what might be expected were the f-s side bands to extend beyond the shift limits by the amount required for a-m transmission. It has already been shown that for maximum deviations (M) of less than 0.4 radians, the f-s bandwidth of intelligence is the same as the a-m band required. For frequency-shift keying it was determined that the bandwidth required is approximately the total shift plus twice the highest keying frequency, equation (110). This type of formula evidently indicates slightly greater bandwidth than is necessary in some cases. However, because of its simplicity and close approximation in most cases it is employed here. From these considerations and equation (115), we obtain

$$\text{B.W.} = 2\Delta f + 1.3 \text{ lines/inch} \times \text{inches/sec.} \tag{116}$$

The same variation in value of nk can be applied to this formula, and the results obtained should be similar to those obtained with a-m transmission.

In order to observe the effect of band limiting on recorded copy, a 2000-cycle band-pass filter was placed in a frequency shift facsimile circuit. The frequency was shifted symmetrically within the 2000-cycle band. Total shifts of from 500 to 2000 cycles were employed and the copies obtained are included as Figures 19 to 25. Close observation of the ladders on the turrets reveals that the rungs, not the larger shadows, are visible in Figures 19 and 20 and to a lesser extent in Figure 21. Beyond this shift the rungs can not be distinguished. Further increase in shift causes a considerable increase in distortion which is very objectionable in Figures 24 and 25.

Facsimile f-s channel widths may be more difficult to determine than those where frequency-shift keying is employed. When maximum deviations less than 0.4 are employed, it can readily be shown that the side band amplitudes will drop off by a factor of $(1/n^3)$ in contrast to the factor of $(1/n^2)$ for a-m transmission. When larger deviation ratios are employed, the channel will naturally be greater; however, outside the shift limits the side band amplitudes will drop off in amplitude faster than $(1/n^2)$ except for extremely large deviation ratios.

Bandwidth Requirement Conclusions

The bandwidth required for transmission of intelligence by frequency modulation, or frequency shift, is the same as amplitude modulation provided the maximum phase deviation does not exceed 0.4 radian. When the phase deviation exceeds 0.4 radian, the frequency modulation bandwidth will be greater.

The spacing of communication channels is normally not determined by the minimum bandwidth required, but rather by the rate at which the amplitude of the sidebands outside this band decreases. For modulating waves which basically consist of a fundamental and a large number of harmonics, such as pulses or square waves, the side bands may, and usually do, decrease in amplitude more rapidly for frequency modulation than they do for the normal amplitude-modulation case. This reduction in bandwidth required for channel spacing can be relatively great. An approximate reduction of 4 to 1 was indicated for the particular case of multiplex teletype investigated. The greater rate of reduction in sideband amplitude with frequency-shift transmission is due to the effective integration of the modulation function before it is applied to the carrier as a phase shift.

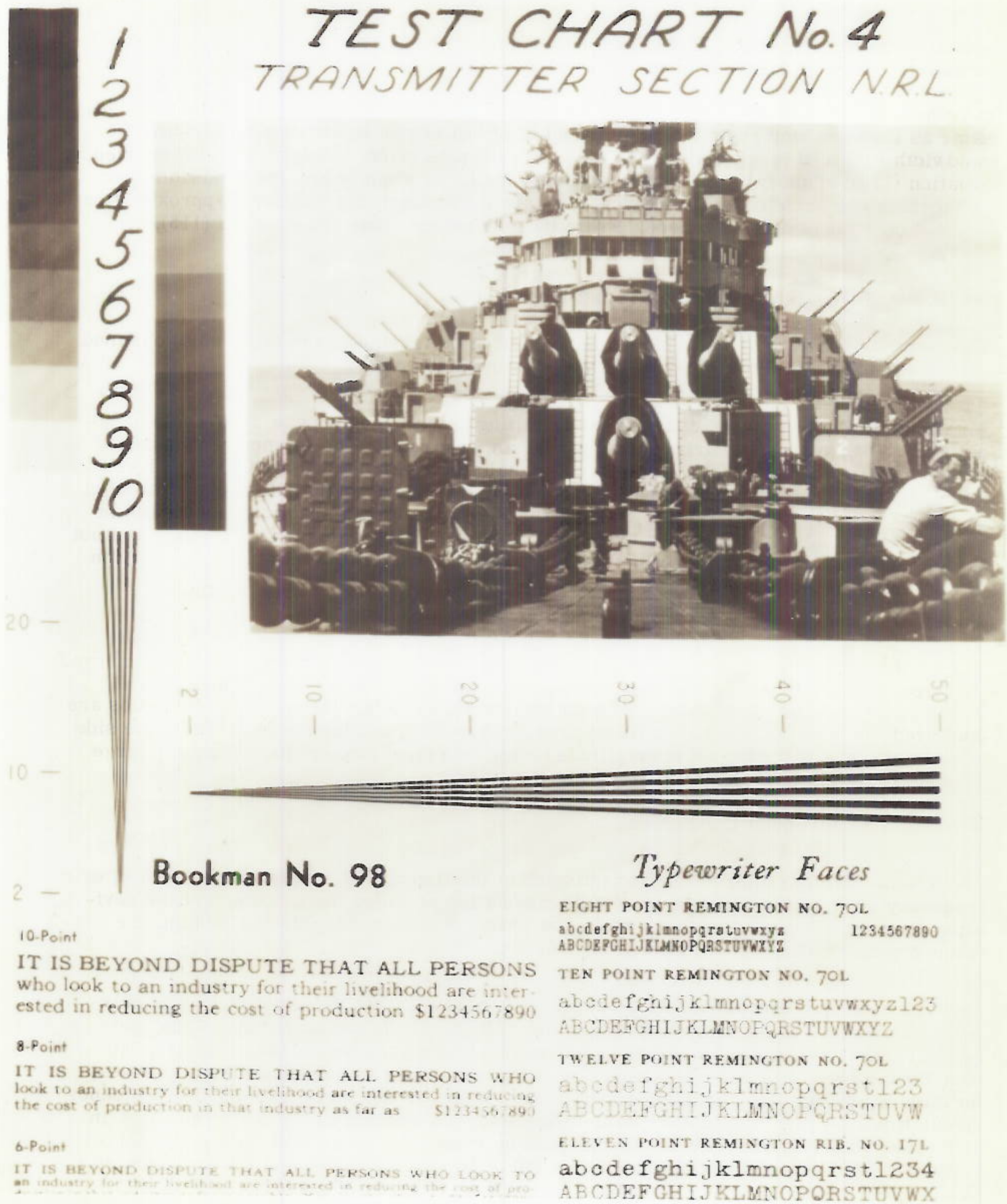


Fig. 19 - 2000~ Filter, 2ΔF = 500~

Fig. 20 - 2000~ Filter, 2ΔF = 750~

Fig. 21 - 2000~ Filter, 2ΔF = 1000~

Fig. 22 - 2000~ Filter, 2ΔF = 1250~

Fig. 23 - 2000~ Filter, 2ΔF = 1500~

Fig. 24 - 2000~ Filter, $2\Delta F = 1750$~

Fig. 25 - 2000~ Filter, $2\Delta F$ = 2000~

EFFECT OF NOISE AND INTERFERING SIGNALS

Theoretical Considerations

In all circuits there is a certain noise level which determines the weakest signal which can be received satisfactorily. Most interference or noise can be placed in one of two catagories (a) random noise and (b) impulse noise. As the name implies random noise has a purely random amplitude and can be considered as being composed of a continuous band of frequencies with random phases. Impulse noise, on the other hand, varies in amplitude in a more definite fashion and is characterized by high peak amplitudes of short duration. This type of noise can again be considered as a continuous band of frequencies; however, the phase relationship is definite in that the components may all add up together at definite times to produce large peak voltages.

With these concepts in mind it would appear that the noise energy present on a radio circuit could best be expressed in energy per bandwidth, or for convenience, as volts/cycle. Once this is determined the voltage produced at the output terminals of a selective circuit is volts/cycle x cycle bandwidth for impulse noise since the components are defined as adding directly. For random noise the output voltage is volts/cycle multiplied by $\sqrt{\text{cycle bandwidth}}$. The square root sign is necessary since the random phase distribution requires a root mean square or energy addition of the components.*

To determine the effect of noise on a f-s facsimile signal, consider the case where a steady-state signal is being received. Figure 26 contains two vectors which represent the signal and interfering frequencies.

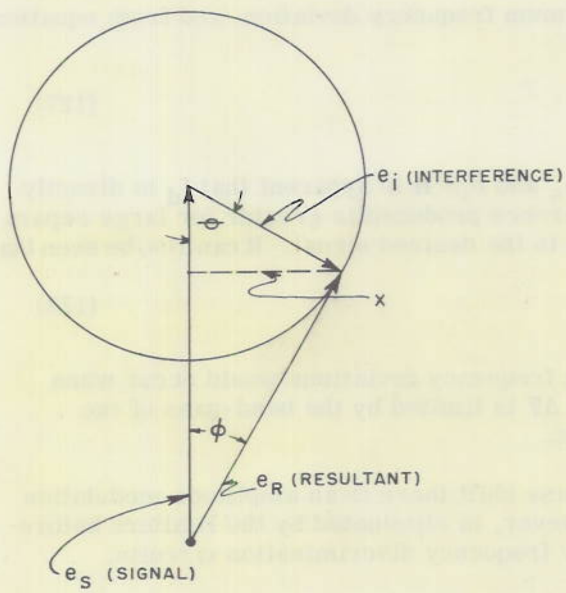


Figure 26
Addition of Signal and Interference

Assuming that the interfering signal is separated from the carrier by f_i cycles in general we can say

$$\frac{d\theta}{dt} = 2\pi f_i = \omega_i \,. \tag{117}$$

However, from Figure 26,

$$\sin\theta = \frac{x}{e_i} \tag{118}$$

$$\sin\phi = \frac{x}{e_R} \,. \tag{119}$$

For small values of θ and ϕ

$$\theta = \frac{x}{e_i} \tag{120}$$

and

$$\phi = \frac{x}{e_R} \,. \tag{121}$$

---

* Landon, V. D., "A Study of Noise Characteristics", Proc. IRE, Vol. 24, pp. 1514-1521, November 1936

The frequency deviation introduced in the signal is $(d\phi/dt)$ where

$$\frac{d\phi}{dt} = \frac{1}{e_R}\frac{dx}{dt} \quad . \tag{122}$$

Also

$$\frac{d\theta}{dt} = \frac{1}{e_i}\frac{dx}{dt} \tag{123}$$

and

$$\frac{dx}{dt} = e_i\frac{d\theta}{dt} \quad . \tag{124}$$

From equations (122) and (124)

$$\frac{d\phi}{dt} = \frac{e_i}{e_R}\frac{d\theta}{dt} \tag{125}$$

The expression $(d\phi/dt)$ is the frequency deviation introduced by the interference, and from Figure 26 it is readily apparent that the maximum frequency deviation introduced occurs when θ and ϕ are equal to zero. At this point

$$e_R = e_S - e_i \tag{126}$$

It now follows that $(d\phi/dt) = f_d$, the maximum frequency deviation, and from equations (117), (125), and (126) we obtain

$$f_d = \frac{e_i}{e_S - e_i}\, 2\pi f_i \tag{127}$$

where f_i is the difference frequency between e_S and e_i. It is apparent that f_d is directly proportional to f_i which means that the interference produced is greater for large separations than when the interfering signal is close to the desired signal. It can also be seen that

$$f_d = \frac{e_i}{e_S}\, 2\pi f_i \tag{128}$$

when $e_i << e_S$. It would also seem that infinite frequency deviations would occur when $e_i = e_S$. In practice this does not occur since ΔF is limited by the band-pass of the circuits following the clipper or limiter stages.

Figure 26 shows that in addition to the phase shift there is an amplitude modulation produced on the carrier. This modulation, however, is eliminated by the limiters before the signal is impressed on the slope circuit or frequency discrimination circuits.

If the carrier is considered as shifting between the limits of the band-pass filter, $2\Delta F$, the modulation introduced by the interfering signal is

$$M_f = \frac{fd}{\Delta F} = \frac{\frac{e_i}{e_S}\, 2\pi f_i}{\Delta F} \quad . \tag{129}$$

When the discriminator output passes through a low-pass filter, f_i is limited to the highest modulating frequency, f_a. From this it is apparent that for low signal-to-noise ratios the effect of an interfering signal is reduced as the shift ΔF is increased. It must be emphasized that this is all based on the assumption that the interfering signal is of much less magnitude than the noise. When the interfering signal is comparable to the desired signal this reduction of interference is no longer present.

The previous equation can be integrated over the proper limits to determine the effect of random and impulse noise on the signal; however, since this has been derived by Crosby* in a very complete manner, we will state some of the conclusions obtained there and see how they affect the particular problem in which we are interested. Figure 27 obtained from Crosby's article indicates that for large signal-to-noise ratios the deviation ratio (ratio of total deviation or input band pass to twice the highest modulating frequency) can be rather large and the resulting gain in signal to noise will be appreciable. For small signal-to-noise ratios it is evident that a deviation ratio of one gives more satisfactory results. This becomes more apparent when it is recalled that increasing the shift will increase the necessary bandwidth. This in turn causes an increase in noise voltage while the carrier signal strength remains the same. The following two equations from Crosby's article should give some idea of the signal-to-noise gain attainable provided the signal-to-noise ratio is great enough at the input of the receiver.

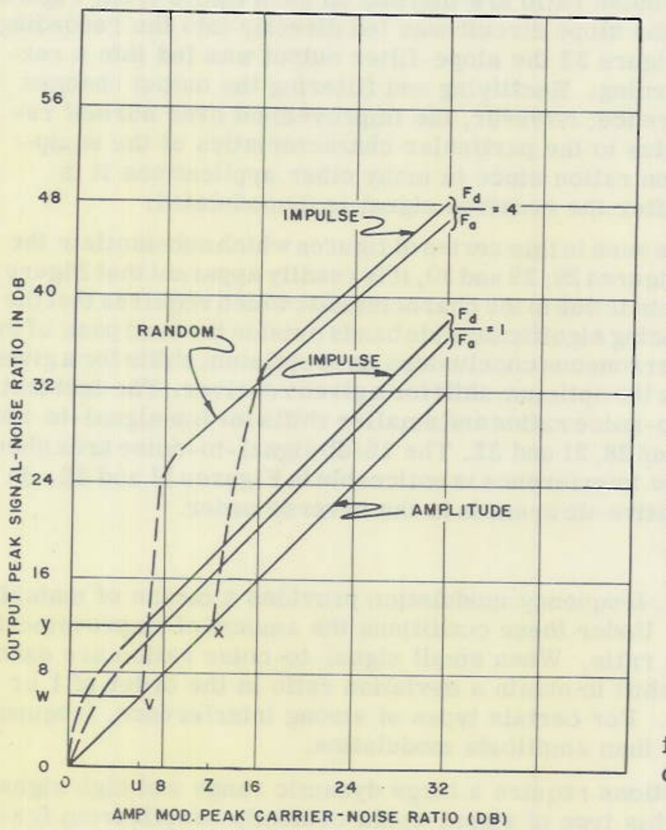


Figure 27 - Improvement in Signal-to-Noise Ratio Due to Frequency Modulation

$$\frac{E_{sig}}{E_{noise}} \text{ gain} = \sqrt{3}\,\frac{F_d}{F_a} \qquad (130)$$

for random noise where (F_d/F_a) is the deviation ratio,

$$\frac{E_{sig}}{E_{noise}} \text{ gain} = 2\,\frac{F_d}{F_a} \qquad (131)$$

for impulse noise. It should be noted that these equations are for receiver gain only since the gain attainable at the transmitter may vary with the type of signal employed and the design of the transmitter. It has been shown† that usable a-m facsimile recordings can be received with peak signal-to-noise

* Crosby, M. G., "Frequency Modulation Noise Characteristics," Proc. IRE, Vol. 25, pp. 472-514, April 1937

† NRL Report R-2885, "Determination of RC-120-B Facsimile Equipment Characteristics," 26 June 1946

ratios down to 8 db or less. This would indicate that a deviation ratio in the neighborhood of one or two would be best for facsimile transmission from the standpoint of noise, when reliability of operation is considered.

Experimental Results

A circuit was set up to determine experimentally the effect of noise on radiophoto equipment employing various shifts. The scanning speed was maintained constant at 13.1 inches per second and the noise was supplied by the output of an RBK receiver. The signal and noise values are both root mean square and are measured at the input to the band-pass filter. The signal-to-noise ratio was varied during transmission between the limits of -9 and +25 db. The width of the band-pass filter employed before the clipper. the shift employed, and the signal-to-noise ratio are marked on each figure from Figure 28 through Figure 32. The output of the slope circuit was fed directly into the recording circuit in Figures 28 to 31 while in Figure 32 the slope-filter output was fed into a rectifier and low-pass filter before recording. Rectifying and filtering the output changes the general appearance of the interference; however, the improvement over normal recording is very small. This may be due to the particular characteristics of the equipment and the relatively small deviation ratios since in many other applications it is necessary to have a low-pass filter after the received signal is demodulated.

Several interesting features can be seen in this series of figures which substantiate the theoretical conclusions. Comparing Figures 28, 29 and 30, it is readily apparent that Figure 28 has the least effective interference. This is due to the characteristic which requires that the shift be as large as possible without placing significant side bands outside the band pass of the filter. This characteristic may cause erroneous conclusions as to optimum shifts for a given service since the results may only give the optimum shift for a given receiver. The fact that large shifts are best for large signal-to-noise ratios and smaller shifts for low signal-to-noise ratios can be seen by comparing Figures 28, 31 and 32. The 15-db signal-to-noise area shows no interference in Figure 28 while some interference is noticeable in Figures 31 and 32. In contrast to this the legibility in the negative-db areas is in the reverse order.

Signal-to-Noise Conclusions

For large signal-to-noise ratios, frequency modulation provides a means of materially improving the signal-to-noise ratio. Under these conditions the amount of improvement is directly proportional to the deviation ratio. When small signal-to-noise ratios are encountered, it is necessary to reduce the shift to obtain a deviation ratio in the order of 1 or 2, if any improvement is to be obtained. For certain types of strong interference, frequency modulation may prove less desirable than amplitude modulation.

Certain types of modulating functions require a large dynamic range and high signal-to-noise ratios to be of any value. This type of signal would evidently benefit from frequency modulation with a high deviation ratio. Other types of modulating functions require only a small dynamic range and can operate satisfactorily with lower signal-to-noise ratios. Teletype which operates on an essentially on-off signal is an example of this type of modulating function. Circuits transmitting these functions will give more reliable operation if small shifts are used, and a deviation ratio of one will probably be the most desirable from the standpoint of noise and interference.

A comparison of Figures 28 and 31 reveals that although a slight amount of interference is noticeable in the 15 db area of Figure 31, it is not considered harmful. In contrast, if the -3 db areas are compared, a decided advantage is evident for Figure 31. This indicates further that a deviation ratio of approximately one is the most satisfactory for facsimile transmission. At present no reason can be seen for employing a modulation index of less than 0.4 radian since the bandwidth required does not decrease for smaller shifts.

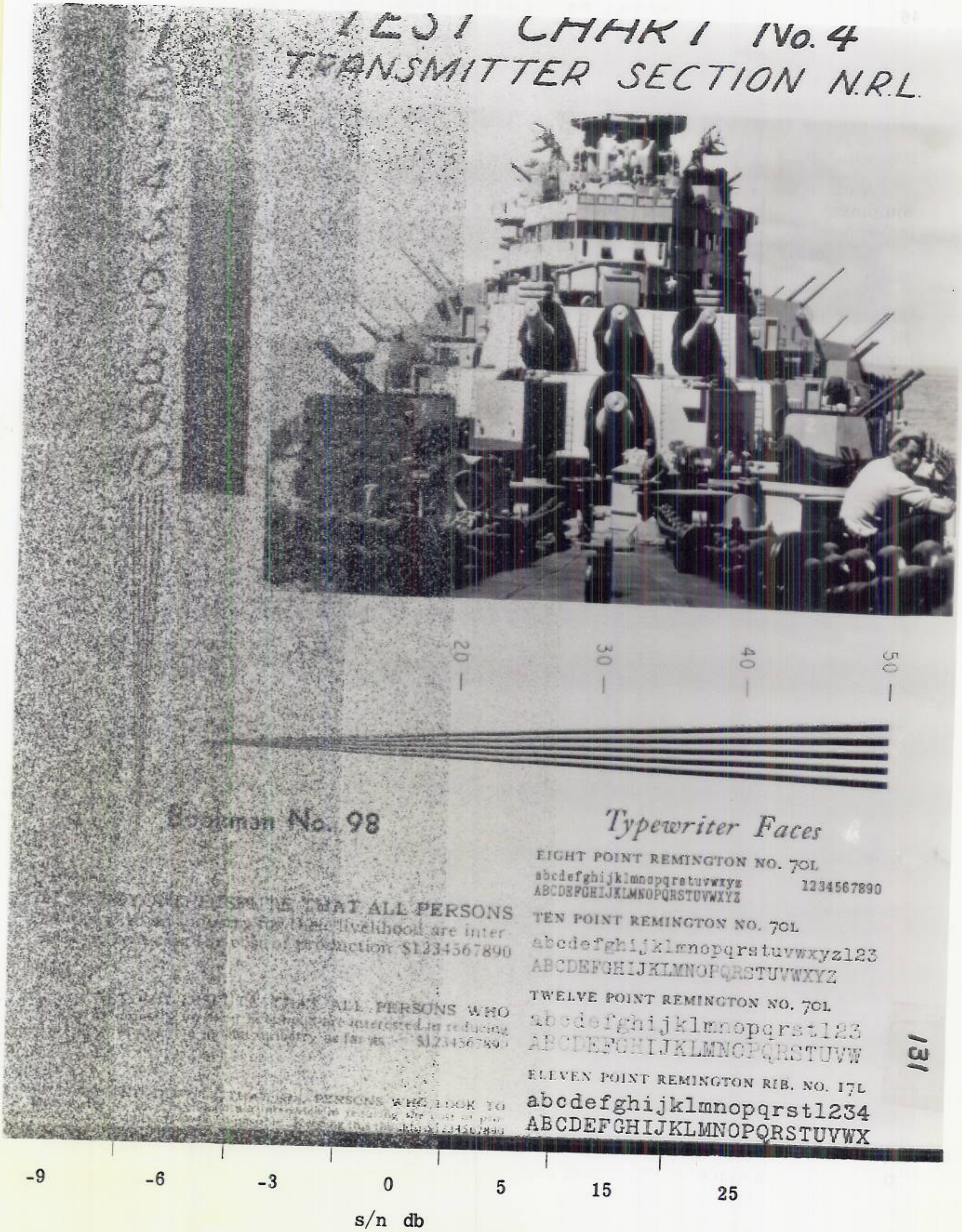
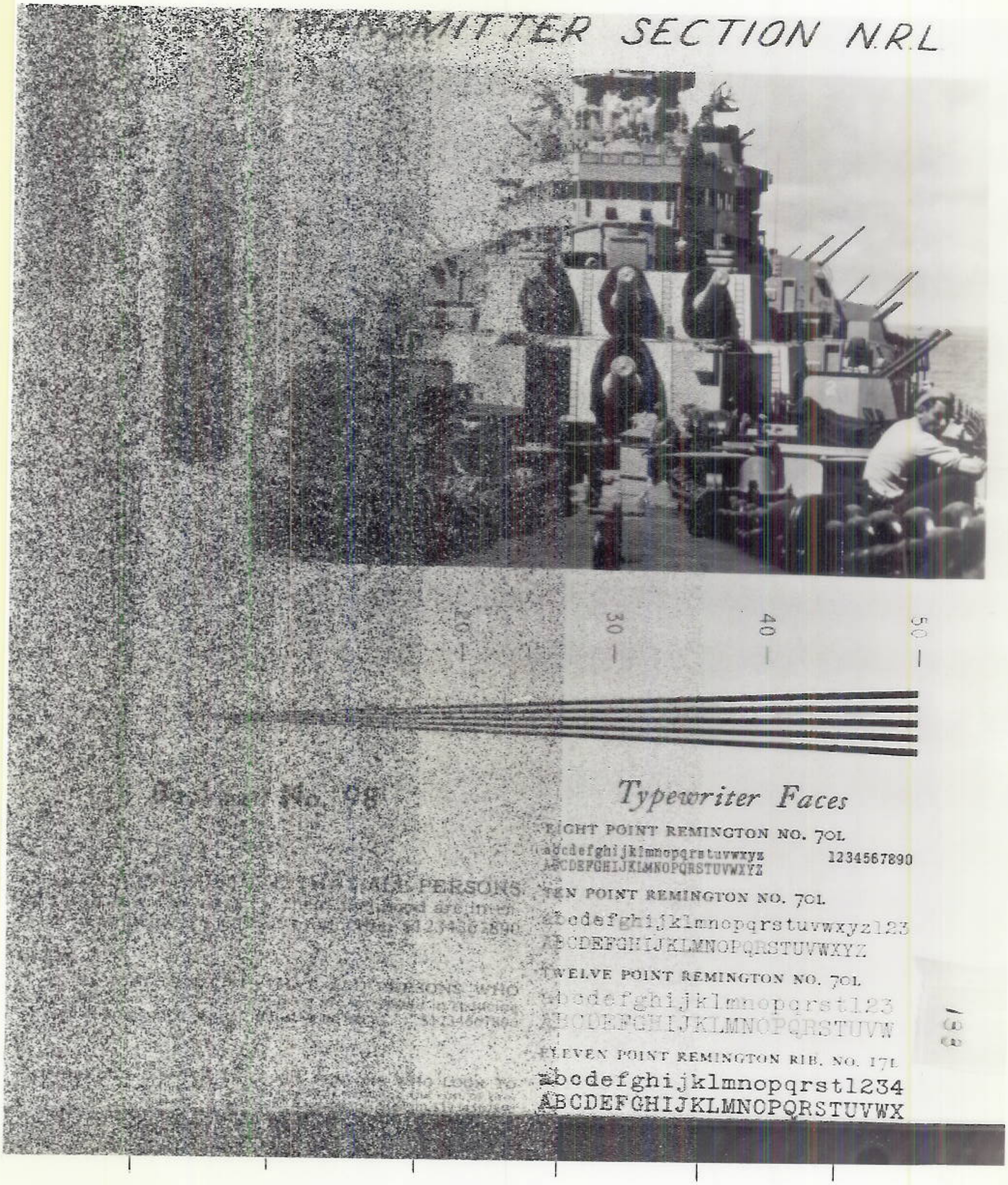


Fig. 28 - 4000~ Filter, 2ΔF = 1800~

Fig. 29 - 4000~ Filter, $2\Delta F$ = 1400~



-9 -6 -3 0 5 15 25

s/n db

Fig. 30 - 4000~ Filter, 2ΔF = 1000~

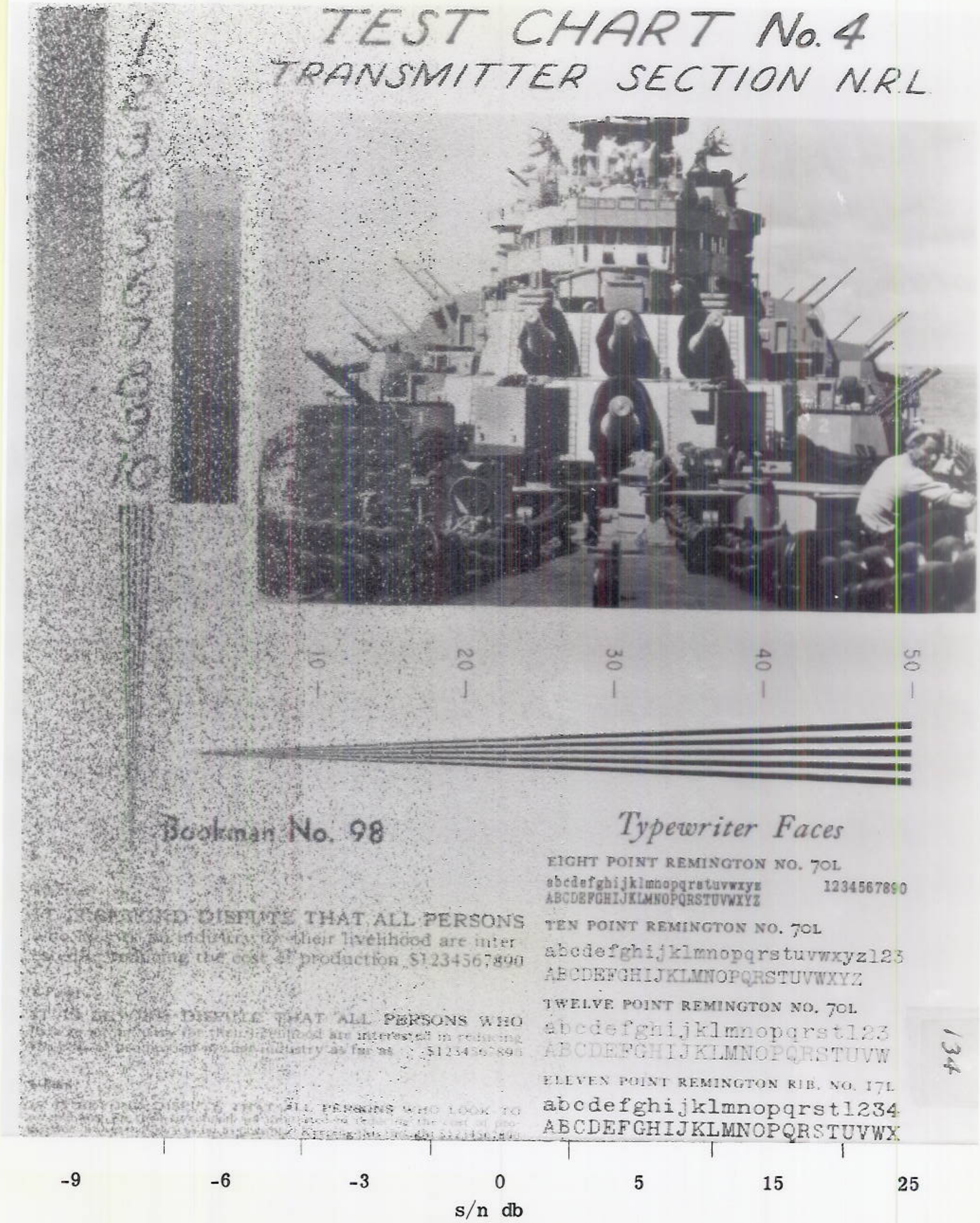


Fig. 31 - 2000~ Filter, 2ΔF = 1000~

-9 -6 -3 0 5 15 25

s/n db

Fig. 32 - 1000~ Filter, 2ΔF = 400~
Rect. & Filtered

EFFECT OF RADIO CIRCUIT ON SIGNALS

General Considerations

A brief review of the factors involved in long-distance high-frequency (3 to 30 Mc) transmission is desirable before attempting to analyze the effect of a radio circuit on facsimile or other similar signals. Transmission over long distances is normally only practicable by means of the sky wave which is reflected back to earth from the ionized layers in the upper atmosphere. The height of these ionized layers varies from 60 to 200 miles while the number of layers and their degree of ionization is dependent on a large number of factors including time of day, season and sun-spot cycle.

The exact mechanism of reflection can be found in several handbooks* as well as in many reports on ionospheric transmission. The signal at the receiver can be considered as consisting of the summation of a large number of signals which have arrived at the receiving antenna via many different paths. The complexity of the path becomes apparent when we realize the heterogeneous composition of the ionosphere in addition to the fact that the ionization density and effective height of the layers are constantly changing. Figure 33 which is taken from the "IRPL Radio Propagation Handbook," should aid in visualizing the method of sky wave transmission. The reflections and summations at the receiver give rise to deformations of the signal which may be classified in four main groups:

(1) Fading
(2) Multipath
(3) Doppler effect
(4) Echo effect.

Fading. - This is probably the most widely known phenomenon of ionospheric transmission. Fading can be devided into four main classes:

(A) Interference fading
(B) Polarization fading
(C) Absorption fading
(D) Skip fading.

The first two are primarily responsible for rapid fading which normally is the most annoying type of signal deformation. Since the limiters in an f-m system effectively eliminate the amplitude variations of all but the deepest fades, this phenomenon is of very little primary interest in an f-m system. Selective fading which can be caused by classes (A), (B) or (D) can cause serious degradation of the facsimile signal since it may eliminate significant side-band components. This type of distortion will probably be minimized by employing relatively small deviation ratios of one or less, since this would reduce the bandwidth and make it possible to utilize more nearly linear portions of the spectrum.

Multipath. - This is probably the most important type of distortion for facsimile and similar signals where the relative transmission time is an important factor. This type

---

* Terman, F. E., "Radio Engineer Handbook," McGraw-Hill Book Co., New York, p. 532, 1943; "IRPL Radio Propagation Handbook", Interservice Radio Propagation Laboratory, National Bureau of Standards, Washington, D.C., November 1943.

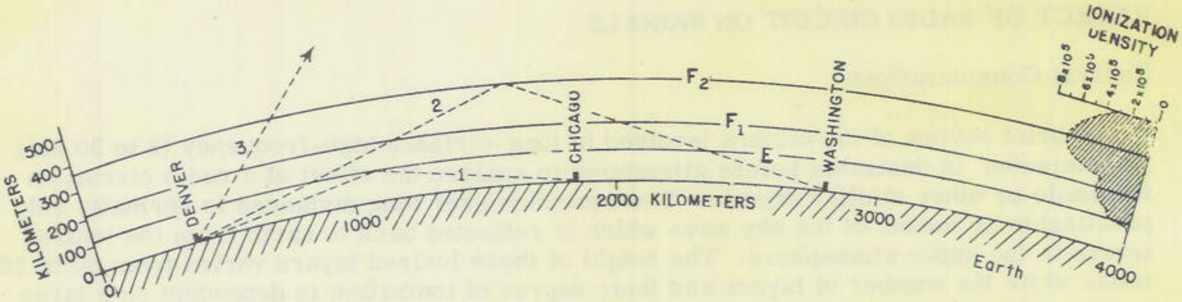


Figure 33

of distortion is produced when the paths over which the signals travel are considerably different in length. The time-delay differences encountered in practice normally vary between limits of 0.05 to 2 milliseconds. A 2-millisecond delay amounts to 0.026 inches on a facsimile receiver scanning at the rate of 13 inches per second. This distance is approximately one quarter the height of 8-point type, and will not permit intelligible copy of anything less than rather large type.

The amount of multipath distortion over a given circuit will vary with the frequency employed. In general the higher frequencies (18 to 20 megacycles) will suffer less than the lower frequencies. This is due to the fact that fewer paths exist as the maximum usable frequency is approached. In addition to choice of frequency it should be possible to reduce multipath effects by employing antennas with variable vertical directivity. This, however, is limited in application.

To determine the effect of multipath on f-m signals, consideration will be given to two cases:

(A) effect on a constant signal and
(B) effect on variable-frequency signal.

The phase between signals over various paths may change and introduce a variation in amplitude which will be removed by the limiters. If the shift is rapid a slight variation in frequency will occur and the output or recorded density will vary. This effect is very slight and normally can be neglected. For case (B) consider the transmitter as scanning a black area and suddenly changing to white. At the receiver both frequencies will be received at the same time for a period equal to the delay-time difference. The effect on the receiver can be treated in the same manner as an interfering signal which has already been covered in a preceding section.

It has been shown* that when two frequencies are being received at the same time, the instantaneous frequency f_i can be represented by the following equation:

$$f_i = \frac{\omega}{2\pi} + \frac{\mu}{\dfrac{\cos 2\pi\mu t + 1/x}{\cos 2\pi\mu t + x} + 1} \qquad (132)$$

* Corrington, Murlan S., "Frequency Modulation Distortion Caused by Common and Adjacent-Channel Interference," RCA Review, Vol. VII, No. 4, pp. 522-560, December 1946

where

ω = $2\pi f$ (carrier)

$$x = \frac{e\ (\text{interfering})}{e\ (\text{carrier})}$$

μ = frequency separation between signals

Since the output of a discriminator is proportional to the instantaneous frequency, the receiver output can be represented by equation (132) times a constant. A graph of this function is shown in Figure 34. From equation (132) it is evident that the intensity of the interference increases directly with the frequency separation μ. The frequency with which the interfering pulses occur is also a direct function of μ; however, their shape is entirely dependent on the relative intensities x. When multipath conditions exist, two or more paths are present over which the radio waves may travel. The travel times are usually different resulting in several frequencies being received at the same time. The least amount of interference occurs when the signal over one of the paths is much stronger than the rest. Under actual conditions the effective transmission of each path varies with time due to various phenomena and as a result first one and then another path may predominate. Since there will be a space displacement at each change-over it can be seen that the effect of changing paths will always be evident. Ideally, when path changes must be accepted, it is desirable to minimize the interference which will occur when the signals become of comparable amplitude and also if possible to reduce the number of changes occurring. Controlling the change-over would probably require complex directive antenna arrays and will not be considered further here. To minimize the amounts of interference during change-over in a given circuit with definite transmission characteristics, two variables are within the control of the operators.

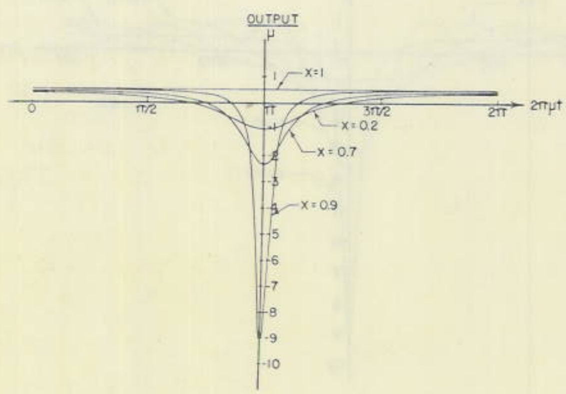


Figure 34 - Instantaneous Frequency Resulting From Interfering Signal

(A) The speed of intelligence transmission and

(B) the modulation index.

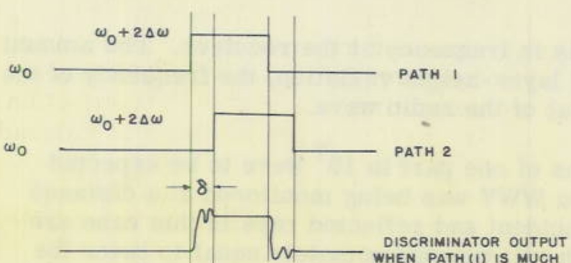


Figure 35 - Effect of Multipath on Frequency-Shift Transmission

Normally it is desirable to maintain (A) as large as possible. The question then arises as to the effect of varying the modulation index when multipath conditions exist on the circuit which is to be employed. Figure 35 should aid in visualizing what is taking place when a very simple type of repeating pulse is being transmitted over a two-path circuit. Referring to Figure 28 and equation (132) it is evident that the beats shown on the last curve of Figure 35 are directly proportional to μ or $2\Delta\omega$ as well as x, the ratio of signal strengths. Since the pulse amplitude is also directly proportional to the frequency shift, it is evident that when no filtering is present at the discriminator output the effect of multipath will be somewhat the same for different modulation indices. It should be emphasized that this conclusion is based on simplified assumptions which are only approximated in actual circuits. In the section on noise and interference it was pointed out that a low-pass filter which will just pass the highest component required by the modulating function is very desirable. A low-pass filter will reduce the amplitude of the multipath interference provided the interference has components which are higher in frequency than those required for satisfactory reproduction of the received signal. This indicates that deviation ratios of one or greater are desirable when multipath is presented. The amount of interference reduction will depend on the type of signal being received. In Corrington's article* the results of band limiting are calculated and plotted up in form similar to Figure 34 and are reproduced here as Figure 36. One of the interesting results of the analysis of the interfering wave spectrums is that the dc or average value will give an output equal to that of the larger signal by itself provided all the sine-wave terms due to multipath interference are eliminated by filtering.

From the above considerations it appears that deviation ratios greater than one should be employed to minimize multipath distortion. For frequency shift keying where essentially two instantaneous frequencies are employed it is considered desirable that the separation $2\Delta F$ should be greater than the highest frequency component passed by the filter following the discriminator. Before deciding on any fixed shift for a given service, it is necessary that all factors be considered and properly weighted according to circuit conditions.

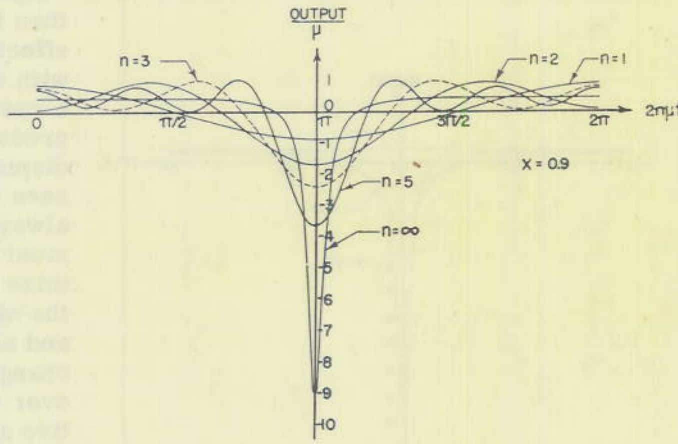


Figure 36
Interference Curves with Low-Pass Filter

Doppler Effect. - During certain periods the effective height of the ionized layers varies rather rapidly with time. Because of this, a steady frequency at the transmitter will vary slightly in frequency at the receiver. The amount of frequency deviation depends on the rate of layer-height variation, the frequency of the signal, and the angles of departure and arrival of the radio wave.

It was shown† that frequency deviations of one part in 10^7 were to be expected during transitional periods when radio station WWV was being monitored at a distance of 13 miles. It should be apparent that the incident and reflected rays in this case are almost vertical, and the change in path length will be approximately equal to twice the change in layer height. In a simplified case where virtual heights are considered, the following relations may be obtained:

$$\frac{L}{2} = \frac{h}{\sin\theta} \tag{133}$$

* Corrington, Op. cit.

† Lopham, Evan G., "Monitoring the Standard Radio-Frequency Emissions," RP766 Journal of Research of the National Bureau of Standards, Vol. 14, pp. 227-238, March 1935

where L is the path length, h the virtual layer height and θ is the angle the ray makes with the horizontal and may be as small as 5 or 6 degrees for long-distance transmission. Differentiating equation (133) with respect to time we obtain

$$\frac{dL}{dt} = \frac{2}{\sin\theta} \frac{dh}{dt}. \tag{134}$$

From equation (134) it is evident that in the simple case considered the frequency deviation introduced is inversely proportional to the sine of θ. With frequency deviations of one part in 10^7 for vertical incidence it follows that for an angle of 5 or 6 degrees the frequency deviation will be one part in 10^6.

In practice the conditions can be more complex than those considered, and the frequency deviations encountered may be one or two parts per million. This will amount to 20 or 40 cycles if a 20-megacycle signal is employed. To minimize this effect it is necessary that the frequency shift, Δf, employed in a circuit be large compared to the doppler shift.

Echo Effect. - As the name implies this disturbance arises when the transmitted ray strikes a reflecting area and is scattered or reflected in directions other than would be expected from straight refraction. This effect is likely to be most annoying when the receiver is relatively close to the transmitter. For certain conditions it is possible that rather strong signals may be received from directions other than that of the transmitter. Two sources* give rather complete accounts of what may be expected and describe the mechanism involved in producing echoes; also included are recordings of echoes with time delays up to 20 or 30 milliseconds.

The relative importance of echoes is not very great in normal point-to-point communication since their effect can be eliminated by proper use of directive antenna arrays. The magnitude of the echoes can also be minimized by proper choice of transmitting frequency.

Radio Circuit Conclusions

Most long-distance radio circuits are complex in nature and act in many different ways to deform signals transmitted over them. The f-m waves are affected somewhat differently than a-m waves; however, it is possible to reduce some of the harmful effects by proper choice of frequency and frequency shift. No simple rule can be given for treating the various types of deformation since some factors may produce improvements from one standpoint and at the same time cause another type of distortion to become more harmful. In view of this a rather thorough balancing of the factors mentioned in this section should be performed before attempting to improve any given circuit.

GENERAL CONCLUSIONS

When selecting the types of modulation to be employed for a given service, many factors must be carefully balanced. It has been shown that the type of function being transmitted is an important factor to consider. Functions requiring large dynamic ranges

---

* Taylor, A. Hoyt, and L. C. Young, "Studies of Echo Signals," Proc. IRE, Vol. 17, pp. 1491-1507, September 1929; C. F. Edwards and Karl G. Jansky, "Echoes from Near-by Short-Wave Transmitters," Proc. IRE, Vol. 29, pp. 322-329, June 1941

and large signal-to-noise ratios may often benefit when frequency modulation with relatively large shifts is employed. Other functions which can tolerate rather large amounts of interference may require narrow shifts, i.e. deviation ratios of approximately one. It is also possible that a-m transmission may be preferable is some cases.

When the modulating function contains discontinuities and transients with high harmonic content it is often desirable to employ frequency modulation since the channel spacing required may be less than that required by amplitude modulation.

The effect of a radio circuit on the choice of the type of modulation and index is rather complex. Fading requires that services such as facsimile employ frequency shift or some other equivalent method to prevent undesired changes in density. The doppler effect requires that the shift be great enough to minimize the 20 to 40-cycles doppler shift which may be encountered. Transmitter and receiver stabilities must also be considered; however, they can be controlled rather well if proper precautions are taken.

Multipath may require deviation ratios greater than two, especially in cases such as frequency-shift keying where it is desirable to eliminate the beat between mark and space frequencies.

In general when selecting a frequency shift for a given set of conditions, it is desirable to first determine the minimum signal-to-noise ratio which can be tolerated. Next it is necessary to determine the maximum deviation ratio which can be employed and still maintain a signal voltage which is about twice the noise voltage at the input to the limiter stage. It may then be necessary to modify the shift obtained by this analysis to compensate for the effects of the other types of distortion such as multipath and doppler shift or to meet the stability requirements of the transmitter and receiver employed.

Several other important factors which may affect the choice of a modulation method are size, weight, and complexity of equipment; however, since these factors vary considerably with each design they will not be discussed here.

* * *