Ideal Binary Pulse Transmission by AM and FM

By E. D. SUNDE

(Manuscript received May 18, 1959)

In binary pulse transmission by carrier amplitude or frequency modulation it is ordinarily desirable, both for efficient bandwidth utilization and for improved performance under adverse noise conditions, to use bandpass channels of the minimum practicable bandwidth, as determined by considerations of intersymbol interference and filter design. It is shown that intersymbol interference can be avoided in binary pulse transmission by FM without the need for a wider channel band than in double-sideband AM, for equal pulse transmission rates. Explicit general expressions are derived for the appropriate shaping of the bandpass channel and for the shapes of received pulses, for cases in which rectangular binary pulses are transmitted by FM, without premodulation or postdetection pulse shaping by low-pass filters. Illustrative comparisons are made of binary pulse transmission by AM and FM for two special cases of general interest in communication theory and pulse-system design. The more general case of partial pulse shaping by premodulation and postdetection low-pass filters is also considered.

The performance of FM and AM systems in the presence of noise depends on the division of channel shaping between transmitting and receiving filters. The optimum division with FM and AM is determined for random noise, and comparisons are made of signal-to-noise ratios for optimized FM and AM systems. It is shown that there is a single universal relation between error probability and signal-to-noise ratio, applying to an infinite universe of optimized baseband systems and optimized AM systems with ideal synchronous detection, and that this relation is the same as for baseband transmission over an idealized flat channel of minimum bandwidth. The analysis indicates that, with binary FM and appropriate postdetection low-pass filters, it is possible in principle to realize an improvement in signal-to-noise ratio over bipolar double-sideband AM with synchronous detection (phase reversal), for equal channel bandwidths, average signal power and pulse transmission rates, although this may not be feasible with practicable filters.

TABLE OF CONTENTS

I. INTRODUCTION

Transmission of digital or analog information by binary rather than by multilevel pulses offers significant advantages in systems design. For one thing, it simplifies the implementation of regenerative repeaters and various kinds of terminal equipment, such as carrier modulators and demodulators, and devices for timing-wave provision, coding and storing of messages and automatic error-checking or correction. For another thing, binary pulse transmission imposes less severe requirements on the transmission medium with respect to signal-to-noise ratio, amplitude and phase deviations over the channel band, and tolerable transmission-level variations. Because of these advantages, binary rather than multilevel pulse transmission is ordinarily the more practical and economical method, even in existing channel facilities designed primarily for voice or other analog transmission, where consideration of the rather high signal-to-noise ratio alone would permit a much greater number of pulse amplitudes and, thus, a substantially greater channel capacity than could be economically realized.

The three principal methods of binary pulse transmission by carrier modulation now in use are double- and vestigial-sideband AM, in the form of "on-off" keying with envelope detection, and FM in the form of "frequency-shift" keying. With synchronous or homodyne detection in AM, other methods are feasible that afford a bandwidth saving or improved signal-to-noise ratio or, like FM, have the advantage over "on-off" AM that they facilitate rapid automatic compensation of transmission-level variations. Among these binary methods are bipolar double-sideband AM, also referred to as phase reversal or two-phase modulation, and bipolar vestigial-sideband AM. Another method is bipolar double-sideband AM on each of two carriers at quadrature, also referred to as quadrature double-sideband AM or four-phase modulation.

For optimum performance in binary pulse transmission by AM or FM it is essential that the transmission-frequency characteristics of the channels be appropriately shaped with respect to amplitude and phase, so that intersymbol interference is avoided or at least reduced to a practicable minimum. A second requirement for optimum performance in the presence of noise is an appropriate division of channel shaping between transmitting and receiving filters. In addition, there are various other requirements not pertaining to the channel, such as exact timing in the transmission and reception of pulses and ideal AM and FM modulators and demodulators.

The purpose of this presentation is a determination of these optimum channel characteristics and the optimum signal-to-noise ratios for various error probabilities in binary pulse transmission by AM and FM, with particular emphasis on FM.

The analysis of both analog and digital pulse transmission for FM is more complex than it is for AM, since FM is a nonlinear modulation method. For this reason, the sideband spectrum of a given signal is wider than it is in double-sideband AM, and a wider bandpass channel is required for analog signal transmission without distortion.

In analog transmission it is possible to realize improved performance in the presence of noise in exchange for the increased bandwidth. In binary pulse transmission, however, it is ordinarily desirable, both for optimum performance under adverse noise conditions and for efficient bandwidth utilization, to use the minimum channel bandwidth practicable from the standpoint of intersymbol interference and filter design.

While a wider bandpass channel than in AM is required for distortionless analog transmission, this does not preclude the possibility that, under appropriate conditions, pulses can be transmitted by FM with no intersymbol interference, without the need for a greater channel band than is required in double-sideband AM. This depends on the possibility of controlling pulse distortion resulting from bandwidth limitation so that zero points in the received pulses occur at uniform intervals from the peak pulse amplitude. When pulses are transmitted by AM, this can be accomplished by appropriate shaping of the transmission-frequency characteristic of the channel, as shown elsewhere.[1,2] The analysis is extended herein to ideal binary pulse transmission by FM.

A basic criterion for performance in digital transmission through noise is the error probability as related to the signal-to-noise ratio, which has been dealt with elsewhere for baseband transmission,[3] "on-off" double-sideband AM with envelope detection,[4,5] bipolar AM or carrier phase modulation with synchronous detection[5,6] and frequency modulation.[5,7] In the above analyses random noise is assumed and the signal-to-noise ratio is stated in terms of signal power during a steady "mark" or "space," which, in bandlimited channels, is usually not equal to average signal power, even in binary phase or frequency modulation systems. Moreover, ideal flat baseband or bandpass channels of the minimum bandwidth required to avoid intersymbol interference are assumed or implied in AM, although they are not practicable in actual systems. In the case of FM, no consideration is given to bandwidth requirements and channel shaping for optimum performance.

As an aid in systems design, specific consideration is given in Sections II through XI of this presentation to appropriate bandwidths and channel shaping for AM and FM systems. The remainder of the analysis is concerned with signal-to-noise ratios as related to channel shaping and to appropriate filter shaping for optimum performance in the presence of random noise.

II. FREQUENCY AND AMPLITUDE MODULATION BY BINARY PULSES

The original signal ordinarily would consist of rectangular baseband pulses of duration T, with a negative polarity to indicate "space" and a positive polarity to indicate "mark," or conversely. These rectangular pulses may be applied directly to the frequency modulator, or they may be applied to a premodulation low-pass filter for preshaping.

The modulator output is applied to a bandpass channel with a certain transmission-frequency characteristic, which, in the ideal case, would be symmetrical about the midband frequency and have a linear phase characteristic. The envelope of the received pulses at the channel output and the frequency modulation of the carrier within the envelope depend on the shape of the modulating pulses and on the transmission-frequency characteristic of the bandpass channel. With an ideal detector at the receiving end, the demodulated signal is proportional to the time derivative of the phase of the carrier within the envelope, i.e., to the "instantaneous frequency deviation."

Detection of the phase derivative is facilitated by conventional frequency discriminators or zero-crossing detectors when the channel bandwidth is narrow in relation to the carrier frequency, but this is not a basic theoretical requirement if appropriate detectors are postulated. Nor, with appropriate ideal balanced FM detectors as assumed herein, is a limiter necessary for elimination of amplitude modulation effects, although it is highly desirable with unbalanced detectors and for prevention of undesirable effects of sudden level changes.

At the detector output, a postdetection low-pass filter may be used for final pulse shaping, but it is not essential for this purpose. Such a filter may be required to eliminate unwanted demodulation products (carrier ripple), but the bandwidth required for this can be much greater than that of the bandpass channel, particularly when the channel bandwidth is small in relation to the midband frequency. This condition can always be realized by frequency translation before demodulation, which may also be required for optimum performance with conventional frequency discriminators or zero-crossing detectors.

A more important function of the postdetection filter in conventional analog signal transmission is elimination of higher frequency noise components, in order to realize the inherent FM noise advantage. In binary pulse transmission with a bandpass channel of no greater bandwidth than is required to avoid intersymbol interference, as considered here, the noise advantage that can be derived from the use of a low-pass filter may be rather limited. To realize a significant noise advantage the filter must, in this case, have the appropriate shape, depending on its bandwidth, to avoid excessive intersymbol interference, as will be shown later.

In frequency modulation, the signal applied to the input of a bandpass communication channel is of the general form

$$E_i(t) = \sin[\omega_0 t + \varphi + \psi_i(t)], \tag{1}$$

where ω_0 is the radian frequency of the unmodulated carrier, φ is the carrier phase and $\psi_i(t)$ is related to the modulating voltage $V_i(t)$ by

$$\psi_i(t) = \bar{\omega}_1 \int_0^t V_i(t)\,dt, \tag{2}$$

with $\bar{\omega}_1$ the frequency deviation in radians per second per volt.

In the case of bipolar binary pulse transmission, the original signal, $V_0(t)$, is ordinarily in the form of rectangular pulses of amplitude V_0 and duration T, of either positive or negative polarity. In this case, the original signal is constant during a signal interval of duration T and is given by

$$V_0(t) = \pm V_0\,. \tag{3}$$

In general, with a premodulation low-pass filter, the carrier modulating voltage, $V_i(t)$, differs from $V_0(t)$. Without such a filter, as assumed here, $V_i(t) = V_0(t) = \pm V_0$ and

$$\begin{aligned} \psi_i(t) &= \pm\bar{\omega}_1 \int_0^t V_0 \, dt \\ &= \pm\bar{\omega}t, \end{aligned} \tag{4}$$

where the frequency deviation $\bar{\omega}$ is

$$\bar{\omega} = \bar{\omega}_1 V_0 . \tag{5}$$

The voltage applied to the input of the bandpass channel in this case is, in accordance with (1),

$$E_i(t) = \sin [(\omega_0 \pm \bar{\omega})t + \varphi]. \tag{6}$$

Equivalent performance could accordingly be obtained if the outputs of two oscillators of frequency $\omega_0 + \bar{\omega}$ and $\omega_0 - \bar{\omega}$ were gated by the voltage $V_0(t)$, so that carrier step pulses of duration T and one or the other of the above two frequencies would be applied directly to the bandpass channel. If the latter method is actually used, the two oscillators must be interlocked to avoid excessive phase discontinuities and resultant transmission impairments that would otherwise be likely to occur in switching from one oscillator to the other.

With phase modulation, rather than frequency modulation as considered above, $\psi_i(t)$ in (1) is related to the modulating voltage by

$$\psi_i(t) = \psi_1 V_i(t), \tag{7}$$

where ψ_1 is the phase modulation in radians per volt.

In the case of bipolar binary pulse transmission, (1) becomes

$$\begin{aligned} E_i(t) &= \sin [\omega_0 t + \varphi \pm \psi_1 V_i(t)] \\ &= \sin (\omega_0 t + \varphi) \cos [\psi_1 V_i(t)] \pm \cos (\omega_0 t + \varphi) \sin [\psi_1 V_i(t)], \end{aligned} \tag{8}$$

where the negative sign is used for a space and the positive sign for a mark, or conversely.

The first component in (8) is independent of the pulse polarity. In an optimized system this component must be minimized and the second component, which depends on the pulse polarity, be maximized. The optimum condition is obtained when ψ_1 is so chosen that $\psi_1 \hat{V}_i = \pi/2$, where $\hat{V}_i$ is the peak amplitude of $V_i(t)$. In the particular case of rectangular modulating pulses, $V_i(t) = \hat{V}_i = V_0$ and (8) becomes

$$E_i(t) = \pm\cos (\omega_0 t + \varphi), \tag{9}$$

which represents a sudden phase reversal from space to mark.

In amplitude modulation, the signal applied to the bandpass channel is

$$E_i(t) = [a_0 + a_1 V_i(t)/\hat{V}_i] \cos(\omega_0 t + \varphi), \tag{10}$$

where a_0 and a_1 are constants that determine the degree of modulation, as discussed below for two special cases.

In unipolar or "on-off" binary pulse transmission, $a_1 = a_0$ and, in the particular case of rectangular modulating pulses,

$$\begin{aligned} E_i(t) &= a_0(1 \pm 1) \cos(\omega_0 t + \varphi) \\ &= 0 \qquad \text{for space} \\ &= 2a_0(\cos \omega_0 t + \varphi) \qquad \text{for mark.} \end{aligned} \tag{11}$$

In bipolar AM, $a_0 = 0$ and, for rectangular modulating pulses, (10) becomes

$$E_i(t) = \pm a_1 \cos(\omega_0 t + \varphi), \tag{12}$$

which is identical to (9) with $a_1 = 1$.

With phase reversal or bipolar AM, the signal can be recovered with the aid of a product demodulator, i.e., by homodyne or synchronous detection. To this end, a synchronous demodulating carrier, $\cos(\omega_0 t + \varphi)$, must be derived from or controlled by the signal, which may entail more complicated instrumentation at the receiving end than is required with frequency modulation. Unipolar AM permits the use of simple envelope detection in exchange for a sacrifice in signal-to-noise ratio compared to the other methods. A further disadvantage of unipolar AM is that it is more susceptible to errors during sudden level changes than is bipolar AM or FM.

With any of the above modulation methods the shape of the received pulses depends on that of the modulating pulses and on the transmission-frequency characteristic or "shaping" of the bandpass channel. The appropriate shaping for avoiding intersymbol interference is well known[1,2] for baseband transmission and amplitude and phase modulation systems, and is determined in the following sections for binary FM.

The particular case of rectangular modulating pulses will be considered in detail, and explicit expressions will be derived for appropriate channel shaping to avoid intersymbol interference. The more complicated cases of premodulation and postdetection pulse shaping will be discussed later.

III. TRANSMITTED FREQUENCY-SHIFT WAVE

Let a continuing "space" be represented by a steady-state transmitted wave

$$E_s^{o}(t) = \sin[(\omega_0 - \bar{\omega})t + \varphi] \tag{13}$$

and a continuing "mark" by

$$E_m^o(t) = \sin [(\omega_0 + \omega)t + \varphi], \tag{14}$$

where t is the time from the beginning of a signal element of duration T and is related to the time with respect to the midpoint of a signal element by

$$t = t_0 + T/2. \tag{15}$$

With (15) in (13) and (14):

$$\begin{aligned} E_s^o(t_0) &= \sin [(\omega_0 - \bar{\omega})(t_0 + T/2) + \varphi] \\ &= \sin [(\omega_0 - \bar{\omega})t_0 + \varphi_0 - \bar{\omega}T/2], \end{aligned} \tag{16}$$

$$E_m^o(t_0) = \sin [(\omega_0 + \bar{\omega})t_0 + \varphi_0 + \bar{\omega}T/2], \tag{17}$$

where

$$\varphi_0 = \varphi + \omega_0 T/2. \tag{18}$$

It will be assumed that

$$\bar{\omega}T = \pi, \tag{19}$$

in which case the frequency difference between mark and space in cycles per second is $2\bar{\omega}/2\pi = 1/T$, or equal to the bit-rate. This assumption need not be made at this point, but it turns out later to be a condition for avoiding intersymbol interference and simplifies the analysis. With the above assumption,

$$E_s^o(t_0) = -\cos [(\omega_0 - \bar{\omega})t_0 + \varphi_0], \tag{20}$$

$$E_m^o(t_0) = +\cos [(\omega_0 + \bar{\omega})t_0 + \varphi_0]. \tag{21}$$

Assume that a single mark of duration T is preceded and followed by a continuing space. The resultant transmitted wave can be regarded as made up of two components. One is a steady-state component given by the following expression applying for $-\infty < t_0 < \infty$:

$$E_s^o(t_0) = -\cos [(\omega_0 - \bar{\omega})t_0 + \varphi_0]. \tag{22}$$

The other is a transient $E_{sm}^o = -E_s^o + E_m^o$ given by the following expression applying for $-T/2 < t_0 < T/2$:

$$\begin{aligned} E_{sm}^o(t_0) &= \cos [(\omega_0 - \bar{\omega})t_0 + \varphi_0] + \cos [(\omega_0 + \bar{\omega})t_0 + \varphi_0] \\ &= \cos (\omega_0 t_0 + \varphi_0) 2 \cos \bar{\omega}t_0 . \end{aligned} \tag{23}$$

The spectrum of $E_{sm}{}^o$ is given by

$$\begin{aligned} S_{sm}{}^o &= 2\int_{-T/2}^{T/2} \cos(\omega_0 t_0 + \varphi_0)\cos\bar{\omega} t_0\, e^{-i\omega t_0}\, dt_0 \\ &= e^{i\varphi_0}\int_{-T/2}^{T/2} \cos u t_0 \cos\bar{\omega} t_0\, dt_0 \\ &\qquad + e^{-i\varphi_0}\int_{-T/2}^{T/2} \cos(2\omega_0 + u)t_0 \cos\bar{\omega} t_0\, dt_0, \end{aligned} \tag{24}$$

where $u = \omega - \omega_0$ is the frequency from midband.

When the bandwidth of the channel is small in relation to the midband frequency, the spectrum need only be considered for $u \ll \omega_0$. The second integral in (24) can then be disregarded in comparison with the first, and the amplitude of the spectrum becomes independent of φ_0, i.e., independent of the phase of the carrier with respect to the modulating pulse. On this assumption, the amplitude of the spectrum of the carrier envelope at the frequency u from the midband frequency ω_0 is given by the first integral in (24), and becomes

$$S_{sm}{}^o(u) = S^o(u) = \frac{T}{2}\left[\frac{\sin(\bar{\omega} - u)T/2}{(\bar{\omega} + u)T/2} + \frac{\sin(\bar{\omega} + u)T/2}{(\bar{\omega} + u)T/2}\right]. \tag{25}$$

With $\bar{\omega}T = \pi$ in accordance with (19), (25) becomes

$$S^o(u) = S^o(-u) = \frac{T}{2}\,\frac{4}{\pi}\,\frac{\cos(\pi u/2\bar{\omega})}{1 - (u/\bar{\omega})^2}. \tag{26}$$

For $u = \pm\bar{\omega}$, the value of (26) is*

$$S^o(\pm\bar{\omega}) = \frac{T}{2}. \tag{27}$$

IV. FREQUENCY-SHIFT PULSE TRANSMISSION CHARACTERISTIC

Let the phase characteristic of the channel be assumed to be linear, and let $A(u)$ be the amplitude characteristic of the channel as a function of the frequency $u = \omega - \omega_0$. At the channel output, the spectrum of the received wave resulting from $E_{sm}{}^o(t_0)$ given by (23) is then

$$S(u) = A(u)S^o(u). \tag{28}$$

If the amplitude characteristic is symmetrical, i.e., if $A(-u) = A(u)$, the spectrum is also symmetrical; i.e.,

$$S(-u) = S(u). \tag{29}$$

* This result is obtained by determining the value of the limit 0/0 as $u \to \bar{\omega}$.

When the above conditions are satisfied, the shape of a received pulse in response to $E_{sm}^{o}(t_0)$ is given by[2]

$$E_{sm}(t_0) = \cos [\omega_0 t_0 + \varphi_0]\bar{E}_{sm}(t_0), \tag{30}$$

where $\bar{E}_{sm}$ is the envelope of the received pulse. The shape of the envelope is the same as that of a demodulated pulse in AM, when the pulse spectrum is $S(u)$, and is given by[2]

$$p(t_0) = \bar{E}_{sm}(t_0) = \frac{1}{\pi}\int_{-\omega_0}^{\infty} S(u) \cos ut_0 \, du, \tag{31}$$

where the lower limit can, for practical purposes, be replaced by $-\infty$, since $S(u) \approx 0$ for $u = -\omega_0$ when $\bar{\omega} \ll \omega_0$. The received wave in response to the steady-state component $E_s^{o}(t_0)$ given by (22) is

$$\begin{aligned} E_s(t_0) &= -A(-\bar{\omega}) \cos [(\omega_0 - \bar{\omega})t_0 + \varphi_0] \\ &= -A(-\bar{\omega})[\cos (\omega_0 t_0 + \varphi_0) \cos \bar{\omega} t_0 + \sin (\omega_0 t_0 + \varphi_0) \sin \bar{\omega} t_0], \end{aligned} \tag{32}$$

where $A(-\bar{\omega}) = A(\bar{\omega})$ is the amplitude of the transmission-frequency characteristic $A(u)$ at $u = \mp\bar{\omega}$, i.e., at the frequencies $\omega_0 \mp \bar{\omega}$.

The resultant received wave when a mark of duration T is preceded and followed by a continuing space is

$$E(t_0) = E_s(t_0) + E_{sm}(t_0) \tag{33}$$

$$\begin{aligned} &= -\cos (\omega_0 t_0 + \varphi_0)[A(-\bar{\omega}) \cos \bar{\omega} t_0 - p(t_0)] \\ &\quad -A(-\bar{\omega}) \sin (\omega_0 t_0 + \varphi_0) \sin \bar{\omega} t_0 . \end{aligned} \tag{34}$$

This can be written in the form

$$E(t_0) = \bar{E}(t_0) \cos (\omega_0 t_0 + \varphi_0 + \psi_0), \tag{35}$$

where the envelope is given by

$$\begin{aligned} \bar{E}(t_0) &= A(-\bar{\omega})\{[\cos \bar{\omega} t_0 - \mu p]^2 + \sin^2 \bar{\omega} t_0\}^{1/2} \\ &= A(-\bar{\omega})\{1 + \mu^2 p^2 - 2\mu p \cos \bar{\omega} t_0\}^{1/2} \end{aligned} \tag{36}$$

and the phase modulation ψ_0 is given by

$$\tan \psi_0(t_0) = -\frac{\sin \bar{\omega} t_0}{\cos \bar{\omega} t_0 - \mu p}, \tag{37}$$

where $p = p(t_0)$ and

$$\mu = 1/A(-\bar{\omega}) = 1/A(\bar{\omega}). \tag{38}$$

With an ideal detector, the received signal is proportional to $\psi_0' = d\psi_0/dt$, which becomes

$$\psi_0' = -\bar{\omega} \frac{1 - \mu p \cos \bar{\omega} t_0 + (\mu p'/\bar{\omega}) \sin \bar{\omega} t_0}{\sin^2 \bar{\omega} t_0 + [\cos \bar{\omega} t_0 - \mu p]^2}, \tag{39}$$

where $p' = dp(t_0)/dt_0$.

Relation (39) gives the frequency deviation from the midband frequency ω_0, in which case a continuing space is represented by a frequency deviation $-\bar{\omega}$ and a continuing mark by $\bar{\omega}$. If the frequency $\omega_0 - \bar{\omega}$ is instead used as reference, the frequency deviation at the receiving end is

$$\psi'(t_0) = \bar{\omega} + \psi_0'(t_0). \tag{40}$$

The ratio $\psi'(t_0)/2\bar{\omega}$ represents the pulse transmission characteristic of the channel in response to a sudden frequency shift $2\bar{\omega}$ of duration T, from $\omega_0 - \bar{\omega}$ to $\omega_0 + \bar{\omega}$, and is given by

$$P(t_0) = \frac{\mu}{2} \frac{\mu p^2 - \cos \bar{\omega} t_0 - (p'/\bar{\omega}) \sin \bar{\omega} t_0}{\sin^2 \bar{\omega} t_0 + (\cos \bar{\omega} t_0 - \mu p)^2}. \tag{41}$$

From (39) or (41) the conditions for binary pulse transmission without intersymbol interference can be established, as is discussed in the next section.

V. IDEAL FREQUENCY-SHIFT TRANSMISSION CHARACTERISTICS

In order to transmit binary pulse trains without intersymbol interference, it is necessary that the transmission characteristic $P(t_0)$ for a single mark or pulse as considered in the preceding section be zero at sampling instants $t_0 = \pm mT$, $m = 1, 2, 3$, etc., and that, at $t_0 = 0$, $P = 1$.

In view of (19), $\cos(\pm m\bar{\omega}T) = (-1)^m$ and $\sin(\pm m\bar{\omega}T) = 0$. Hence at sampling points (41) becomes, for $m \neq 0$,

$$P(mT) = \frac{1}{2} \frac{\mu p(mT)}{\mu p(mT) + (-1)^m} \tag{42}$$

and

$$P(0) = \frac{1}{2} \frac{\mu p(0)}{\mu p(0) - 1}. \tag{43}$$

Thus $P(mT) = 0$ provided

$$p(mT) = 0, \tag{44}$$

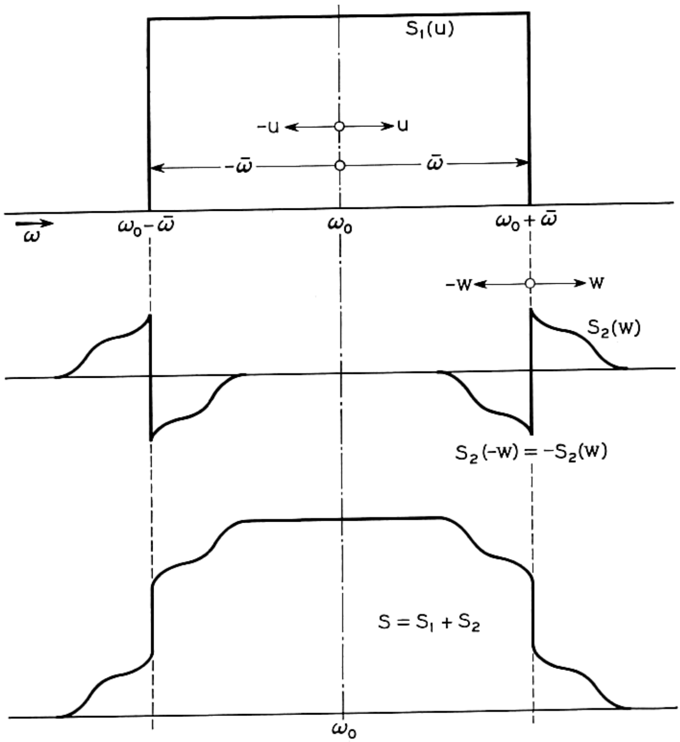

Fig. 1 — Properties of spectra at detector input such that, in double-sideband amplitude modulation, $p(0) = 1$, $p(mT) = 0$, $T = \pi/\bar{\omega}$, $m = 1, 2, 3, \cdots$.

and $P(0) = 1$ provided

$$\mu p(0) = 2. \tag{45}$$

The above conditions (44) and (45) can be satisfied provided that $S(u)$, in addition to being symmetrical as required by (31), has the further property illustrated in Fig. 1. This property is the same as that required for double-sideband transmission of pulses at intervals T without intersymbol interference,[1,2] and can be satisfied by an infinite variety of spectra. Among these it is convenient from the standpoint of theoretical evaluation of $p(t_0)$ to assume spectra of the form shown in Fig. 2 given by the following expressions.

In the range $0 < u < \bar{\omega} - \omega_x$:

$$S(u) = S(-u) = \frac{T}{2}. \tag{46}$$

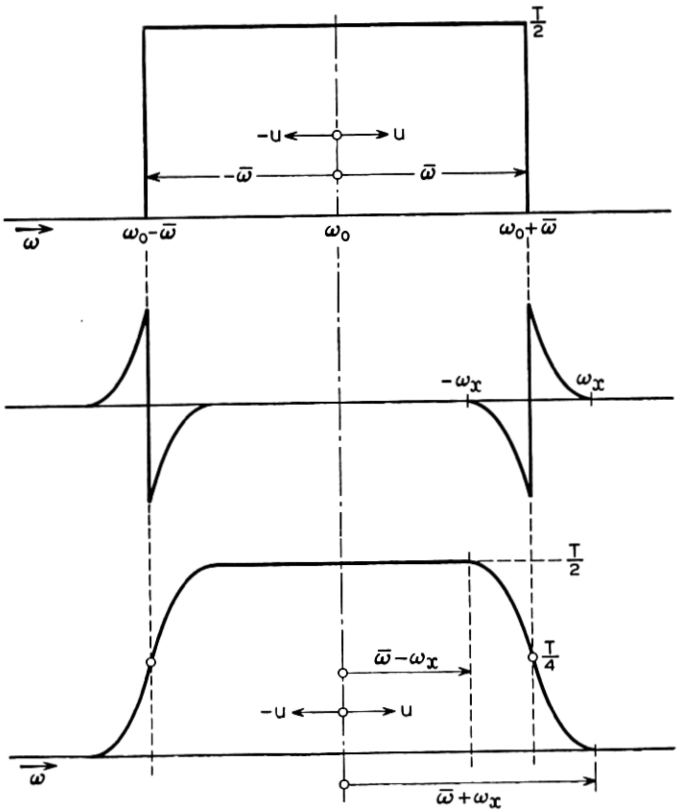

Fig. 2 — Special case of spectra with properties indicated in Fig. 1.

In the range $\bar{\omega} + \omega_x > u > \bar{\omega} - \omega_x$:

$$S(u) = S(-u) = \frac{T}{2} \cos^2 \left(\frac{\pi}{4} \frac{u - \bar{\omega} + \omega_x}{\omega_x} \right). \tag{47}$$

At $u = \bar{\omega}$, (47) becomes

$$S(\bar{\omega}) = S(-\bar{\omega}) = \frac{T}{4}. \tag{48}$$

The required amplitude characteristic of the channel obtained from (28) is

$$A(u) = \frac{S(u)}{S^o(u)}. \tag{49}$$

In view of (27) and (48), $A(-\bar{\omega}) = \frac{1}{2}$, so that (38) gives

$$\mu = 2. \tag{50}$$

When the spectrum is given by (46) and (47), evaluation of (31) gives[2]

$$p(t_0) = \bar{E}_{sm}(t_0) = \frac{T\bar{\omega}}{\pi} \frac{\sin \bar{\omega} t_0}{\bar{\omega} t_0} \frac{\cos \omega_x t_0}{1 - (2\omega_x t_0/\pi)^2}, \tag{51}$$

where $T\bar{\omega}/\pi = 1$, in view of (19), so that

$$\begin{aligned} p(t_0) &= \frac{\sin \bar{\omega} t_0}{\bar{\omega} t_0} \frac{\cos \omega_x t_0}{1 - (2\omega_x t_0/\pi)^2} \\ &= 1 \qquad \text{for} \qquad t_0 = 0 \\ &= 0 \qquad \text{for} \qquad t_0 = mT, \qquad m = 1, 2, 3, \cdots. \end{aligned} \tag{52}$$

Hence, $\mu p(0) = 2$ and $p(mT) = 0$ so that conditions (44) and (45) are satisfied, and single binary pulses can be transmitted without intersymbol interference. This is also the case for pulse trains, as is shown in Section XI.

To find the shape of the received pulses, it is necessary to employ (41) for other values of t_0 than $t_0 = 0$ and mT, as is illustrated in the next sections for two limiting cases of general interest.

VI. SPECIAL CASE OF FLAT SPECTRUM AT DETECTOR INPUT

The amplitude characteristic $A(u)$ of the channel and the frequency-shift pulse transmission characteristic $P(t_0)$ will be determined here for a channel of minimum bandwidth, in which case the spectrum $S(u)$ will be flat for $0 < u < \bar{\omega}$, and will be zero for $u > \bar{\omega}$.

With sharp cutoffs at $u = \pm\bar{\omega}$, ω_x will be zero in (46) and (47), so that

$$S(u) = S(-u) = \frac{T}{2} \qquad 0 \leqq u \leqq \bar{\omega} \tag{53}$$

$$= 0 \qquad u > \bar{\omega}. \tag{54}$$

The required amplitude characteristic of the channel, as obtained from (49), is

$$\begin{aligned} A(u) &= \frac{\pi}{4} \frac{1 - (u/\bar{\omega})^2}{\cos(\pi u/2\bar{\omega})} \\ &= \frac{\pi}{4} \qquad \text{for} \qquad u = 0 \\ &= 1 \qquad \text{for} \qquad u = \pm\bar{\omega} \\ &= 0 \qquad \text{for} \qquad |u| > \bar{\omega}_0. \end{aligned} \tag{55}$$

In the case of double-sideband AM, the spectrum of a pulse of duration T with respect to the midband frequency is

$$s^{\circ}(u) = \frac{T}{2} \frac{\sin (uT/2)}{uT/2}, \tag{56}$$

where, as before, it has been assumed that the channel bandwidth is small in relation to the midband frequency.

To obtain a flat spectrum of amplitude $T/2$ at the detector input, the required amplitude characteristic of the channel is

$$\begin{aligned} a(u) &= \frac{uT/2}{\sin (uT/2)} = \frac{\pi u/2\bar{\omega}}{\sin (\pi u/2\bar{\omega})} \\ &= 1 \qquad \text{for} \quad u = 0 \\ &= \frac{\pi}{2} \qquad \text{for} \quad u = \bar{\omega}. \end{aligned} \tag{57}$$

The amplitude characteristics $A(u)$ and $a(u)$ are shown in Fig. 3.

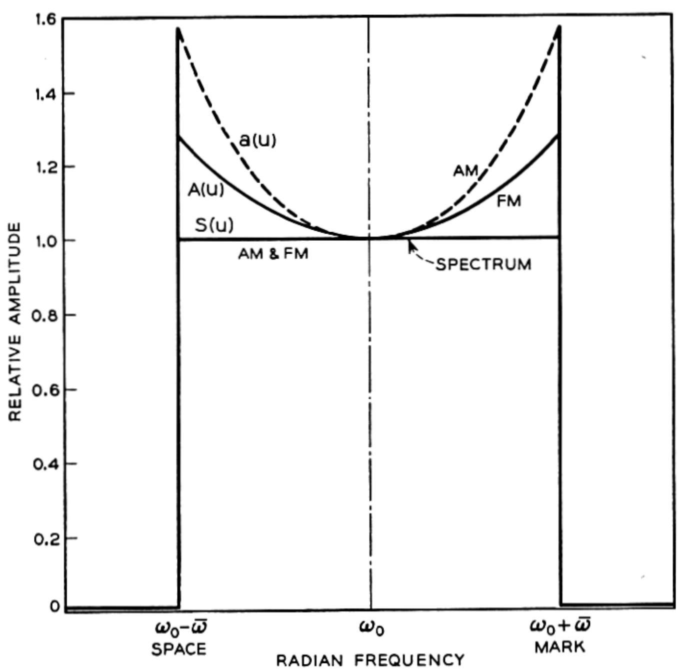

Fig. 3 — Ideal transmission-frequency characteristics of bandpass channels in AM and FM for rectangular modulating pulses and flat spectrum of minimum bandwidth at detector input.

For purposes of direct comparison, the value of $A(u)$ as obtained from (55) has been normalized to $A(u) = 1$ for $u = 0$, by multiplication by $4/\pi$.

With a channel characteristic as given by (55), $p(t_0)$, as obtained from (52) with $\omega_x = 0$, is

$$p(t_0) = \frac{\sin \bar{\omega}t_0}{\bar{\omega}t_0}, \tag{58}$$

$$p'/\bar{\omega} = \frac{1}{\bar{\omega}t_0}\left(\cos \bar{\omega}t_0 - \frac{\sin \bar{\omega}t_0}{\bar{\omega}t_0}\right). \tag{59}$$

With (58) and (59) in (41), the frequency-shift pulse transmission characteristic of the channel becomes

$$P(t_0) = \frac{\sin \bar{\omega}t_0}{\bar{\omega}t_0} \; \frac{3\,\dfrac{\sin \bar{\omega}t_0}{\bar{\omega}t_0} - 2\cos \bar{\omega}t_0}{1 + 4\,\dfrac{\sin \bar{\omega}t_0}{\bar{\omega}t_0}\left(\dfrac{\sin \bar{\omega}t_0}{\bar{\omega}t_0} - \cos \bar{\omega}t_0\right)}. \tag{60}$$

The function $P(t_0)$ is given in Table I and shown in Fig. 4.

The term $\sin \bar{\omega}t_0/\bar{\omega}t_0$ in (60) represents the pulse-transmission characteristic for double-sideband transmission over a channel with an amplitude characteristic $a(u)$. In Fig. 4 the double-sideband AM and the frequency-shift transmission characteristics are compared. It will be noted that they differ appreciably in shape, but have the common properties of unit amplitude at $t_0 = 0$ and zero amplitude at intervals such that $\bar{\omega}t_0 = m\pi$, $m = 1, 2, 3, \cdots$.

VII. SPECIAL CASE OF RAISED COSINE SPECTRUM AT DETECTOR INPUT

In actual communication systems, channels with sharp cutoffs as assumed in the previous example are impracticable for various reasons, such as excessive phase distortion near the band edges and relatively

TABLE I — FREQUENCY-SHIFT TRANSMISSION CHARACTERISTIC $P(t_0)$ FOR FLAT SPECTRUM OF MINIMUM BANDWIDTH

$\bar{\omega}t_0/\pi$	0	0.20	0.40	0.60	0.80	1.0
0	1.0	0.7860	0.5468	0.4182	0.2746	0
1	0	−0.3026	0.0072	0.1629	0.1454	0
2	0	−0.1540	−0.0330	0.0940	0.0985	0
3	0	−0.1023	−0.0339	0.0646	0.0743	0
4	0	−0.0766	−0.0303	0.0488	0.0597	0
5	0	−0.0611	−0.0267	0.0391	0.0499	0
10	0	−0.0304	−0.0160	0.0194	0.0274	0
19	0	−0.0160	−0.0091	0.0101	0.0151	0

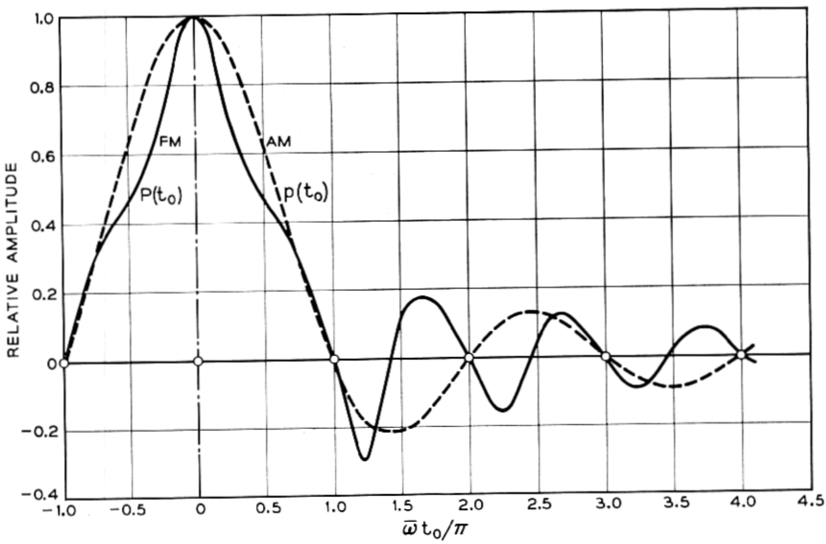

Fig. 4 — Shapes of demodulated pulses in FM and AM for rectangular modulating pulses and for flat spectrum of minimum bandwidth at detector input, as in Fig. 3.

large oscillations in the received pulses that entail precise synchronized sampling at fixed intervals to avoid intersymbol interference. A preferable type of channel characteristic is one that results in a "raised cosine" spectrum of the received wave in response to the transmission of a single rectangular pulse of duration T, as considered below. With this type of channel characteristic, the bandwidth is twice the minimum possible considered in the previous limiting case.

With $\omega_x = \bar{\omega}$ in (47),

$$\begin{aligned} S(u) = S(-u) &= \frac{T}{2} \cos^2 (\pi u / 4\bar{\omega}) \qquad u \leqq 2\bar{\omega} \\ &= 0 \qquad\qquad\qquad\qquad u > 2\bar{\omega}. \end{aligned} \tag{61}$$

The required amplitude characteristic of the channel as obtained from (49) is

$$\begin{aligned} A(u) &= \frac{\pi}{4} \frac{1 - \cos (\pi u/2\bar{\omega})}{2 \cos (\pi u/2\bar{\omega})} [1 - (u/\bar{\omega})^2] \\ &= \frac{\pi}{4} \qquad \text{for} \quad u = 0 \\ &= \tfrac{1}{2} \qquad \text{for} \quad u = \bar{\omega} \qquad (\mu = 2) \\ &= 0 \qquad \text{for} \quad u \geqq 2\bar{\omega}. \end{aligned} \tag{62}$$

The amplitude characteristic required for double-sideband transmission without intersymbol interference is, in this case, given by

$$\begin{aligned} a(u) &= \frac{\cos^2 (\pi u/4\bar{\omega})\pi u/2\bar{\omega}}{\sin (\pi u/2\bar{\omega})} \\ &= \frac{\pi u/4\bar{\omega}}{\tan (\pi u/4\bar{\omega})} \\ &= 1 \quad \text{for} \quad u = 0 \\ &= \frac{\pi}{4} \quad \text{for} \quad u = \bar{\omega} \\ &= 0 \quad \text{for} \quad u \geqq 2\bar{\omega}. \end{aligned} \tag{63}$$

The amplitude characteristics $A(u)$ and $a(u)$ are shown in Fig. 5. For convenient direct comparison, the value of $A(u)$ obtained from (62) has been normalized to $A(u) = 1$ at $u = 0$.

With $\omega_x = \bar{\omega}$ in (52),

$$p(t_0) = \frac{\sin 2\bar{\omega}t_0}{2\bar{\omega}t_0[1 - (2\bar{\omega}t_0/\pi)^2]}, \tag{64}$$

where the relation $\cos \bar{\omega}t_0 \sin \bar{\omega}t_0 = \frac{1}{2} \sin 2\bar{\omega}t_0$ has been used.

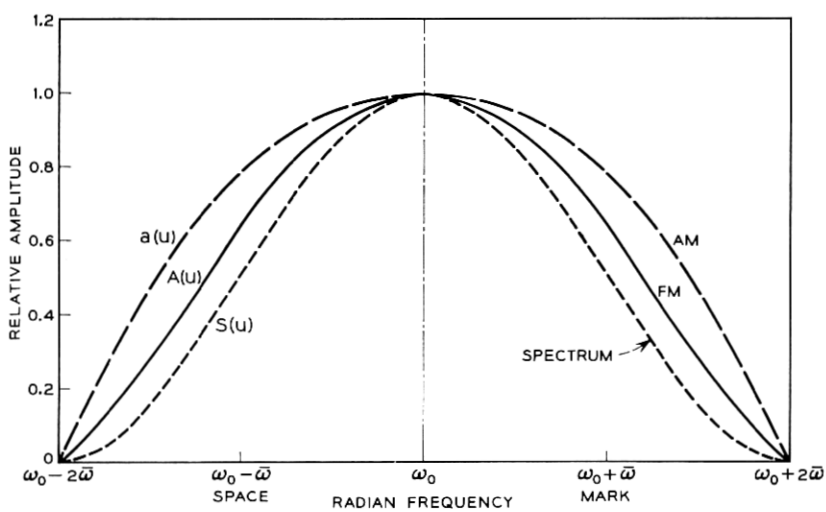

Fig. 5 — Ideal transmission-frequency characteristics of bandpass channels in FM and AM for rectangular modulating pulses and raised cosine spectrum at detector input.

TABLE II — FREQUENCY-SHIFT TRANSMISSION CHARACTERISTIC $P(t_0)$ FOR RAISED COSINE SPECTRUM AT DETECTOR INPUT

$\bar{\omega}t_0/\pi$		0.2	0.4	0.5	0.6	0.8
0	1.0	0.8053	0.5769	0.4887	0.4030	0.2035
1	0	−0.0175	0.0301	0.0265	0.0109	−0.0108
2	0	0.0035	−0.0048	−0.0053	−0.0027	0.0025
3	0	−0.0012	0.0016	0.0019	0.0011	−0.0009
4	0	0.0005	−0.0007	−0.0009	−0.0005	0.0004

Differentiation of (64) yields

$$\frac{p'}{\bar{\omega}} = 2\,\frac{\cos 2\bar{\omega}t_0}{2\bar{\omega}t_0[1 - (2\bar{\omega}t_0/\pi)^2]} - 2\,\frac{\sin 2\bar{\omega}t_0[1 - 3(2\bar{\omega}t_0/\pi)^2]}{(2\bar{\omega}t_0)^2[1 - (2\bar{\omega}t_0/\pi)^2]^2}. \qquad (65)$$

For $\bar{\omega}t_0 = \pi/2$, $p'/\bar{\omega} = -3/(2\pi)$.

With (64) and (65) in (41) the frequency-shift transmission characteristic $P(t_0)$ given in Table II and shown in Fig. 6 is obtained, for a channel with an amplitude characteristic $A(u)$ as shown in Fig. 5.

For comparison with $P(t_0)$, Fig. 6 also shows the pulse-transmission characteristic $p(t_0)$ obtained from (64) for double-sideband transmission over a channel with the amplitude characteristic $a(u)$ shown in Fig. 5. In both cases, the oscillations in the received pulses are quite small, and for this reason the transmission-frequency characteristics shown in Fig. 5 are preferable to those in Fig. 3 in practicable AM and FM systems.

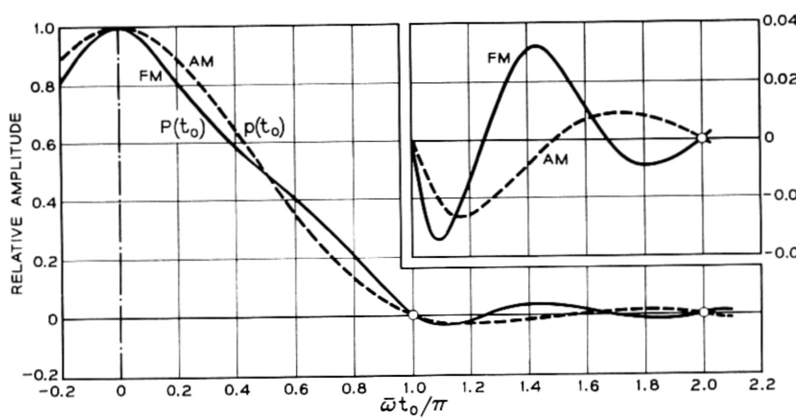

Fig. 6 — Shapes of demodulated pulses in FM and AM for rectangular modulating pulses and raised cosine spectrum at detector input, as in Fig. 5.

VIII. PREMODULATION PULSE SHAPING

In the preceding analysis the modulating pulses were assumed to be rectangular in shape and of duration T, in which case the modulator input in response to a change from space to mark was given by (23) and the corresponding spectrum by (24). In the more general case of premodulation pulse shaping, these equations are replaced by

$$E_{sm}^{o}(t_0) = 2 \cos (\omega_0 t_0 + \varphi_0) \cos \psi_i (t_0), \tag{66}$$

$$S^{o}(u) = 2 \int_0^{\infty} \cos \psi_i (t_0) \cos ut_0 \, dt_0 , \tag{67}$$

where, as before, the second component in (24) has been neglected, and the phase modulation ψ_i is related to the modulating voltage $V_i(t_0)$ by (2), or

$$\psi_i(t_0) = \bar{\omega}_1 \int_0^{t_0} V_i(t_0) \, dt_0 . \tag{68}$$

The above relations apply provided a continuing space is represented by a constant frequency deviation $-\bar{\omega}$ and a continuing mark by $\bar{\omega}$. To this end, it is necessary that the individual modulating pulses $V_i(t_0)$ overlap and be of such form that

$$\sum_{n=-\infty}^{\infty} V_i(t_0 + nT) = V_i(0). \tag{69}$$

For example, the latter relation is satisfied when impulses are applied to a flat low-pass filter of bandwidth $\bar{\omega}$ and linear phase, resulting in a modulating voltage $V_i(t) = V_i(0) (\sin \bar{\omega}t_0)/\bar{\omega}t_0$. The simpler case of overlaps between adjacent pulse intervals only is illustrated in Fig. 7.

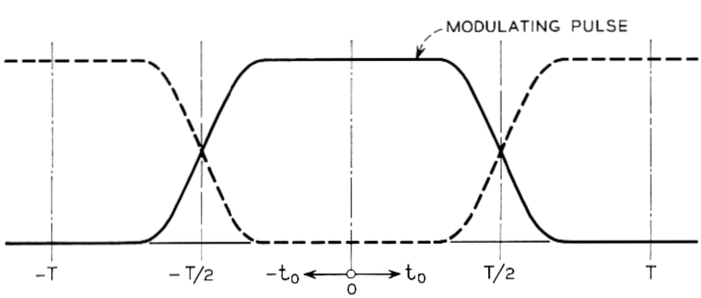

Fig. 7 — Modulating pulses with overlap between adjacent pulse intervals.

The amplitude characteristic of the channel required for a spectrum $S(u)$ at the detector input is, as before, given by

$$A(u) = \frac{S(u)}{S^o(u)}, \tag{70}$$

To avoid intersymbol interference, it is necessary to satisfy (45); i.e., $\mu p(0) = 2$. To this end, it is necessary, in accordance with the discussion in Section V, that $S(\pm\bar{\omega}) = T/4$ and that $A(\bar{\omega}) = \frac{1}{2}$. Hence, a requirement imposed on $S^o(u)$ as given by (67) is that

$$S^o(\bar{\omega}) = \frac{T}{2}. \tag{71}$$

There is an infinity of pulse shapes, $V_i(t_0)$, that satisfy (69), and the corresponding $\psi_i(t_0)$ can be determined formally from (68) and, in turn, $S^o(u)$ can be determined from (67). The principal problem is to determine pulse shapes, other than the rectangular shape considered previously, which also have a spectrum that satisfies (71). The solution of this problem will not be attempted here, but two pulse shapes of general interest in pulse transmission theory will be considered.

A familiar example of modulating pulses that overlap into an infinite number of pulse intervals is represented by the idealized pulse shape obtained by applying impulses to an ideal flat low-pass filter with linear phase characteristic. It will be assumed that the filter bandwidth is $\bar{\omega}$, in which case the modulating voltage is

$$V_i(t_0) = V_i(0)\,\frac{\sin \bar{\omega}t_0}{\bar{\omega}t_0}. \tag{72}$$

With (72) in (68), and with $\bar{\omega}_1 V_i(0) = \bar{\omega}$,

$$\psi_i(t_0) = \mathrm{Si}\,(\bar{\omega}t_0) \tag{73}$$

where Si is the sine integral function.

In Fig. 8 the phase modulation, ψ_i, obtained from (73) is compared with that for rectangular modulating pulses, for which $\psi_i = \bar{\omega}t_0$ for $\bar{\omega}t_0 \leqq \pi/2$ and $\psi_i = \pi/2$ for $\bar{\omega}t_0 > \pi/2$.

With (73) in (67) and $\bar{\omega}t_0 = \tau$, $\bar{\omega}T = \pi$, $a = u/\bar{\omega}$:

$$S^o(u) = \frac{T}{2}\,\frac{4}{\pi}\int_0^\infty \cos \mathrm{Si}\,(\tau) \cos a\tau \, d\tau. \tag{74}$$

For $\tau > 17$ the following approximation applies:

$$\mathrm{Si}\,(\tau) \cong \frac{\pi}{2} - \frac{\cos \tau}{\tau}. \tag{75}$$

The peak amplitudes, $\hat{p}(0)$, required to produce an error when an impulse occurs midways between sampling points is greater than when they occur at a sampling point by a factor of $2^{1/2}$ in FM and a factor of 2 in AM. With a Gaussian amplitude distribution of the pulses, the probability of an error from a pulse midway between two sampling points is in the order of 1 per cent of the probability of an error from a pulse at a sampling point in the case of FM, and is substantially smaller for AM. Hence, virtually all the errors will be caused by pulses that occur near sampling points. The AM advantage over FM for equal error probability is 3 db for impulses that occur at sampling points, and would be expected to be only slightly greater, about 4 db when impulses occurring at all instances with respect to a sampling point are considered.

The above comparisons apply without a postdetection low-pass filter in FM. With an optimum bandpass receiving filter characteristic in FM, the reduction in peak impulse noise afforded by low-pass filter would be expected to be about the same as the reduction in average random noise.

APPENDIX C

Optimum Receiving Filter Characteristic

The optimum receiving filter characteristic in AM and in FM without a postdetection low-pass filter can be determined from the solution of the more general case considered here, of FM with a postdetection filter.

In the latter case, the optimum $R(u)$ is obtained when the product of the two integrals in (125) is a minimum, or for the minimum value of the product:

$$J = J_1 J_2 , \qquad (256)$$

where J_1 and J_2 are functions of $R(u)$ given by

$$J_1 = \int_{-\infty}^{\infty} L^2(u) R^2(u) (1 + u/\bar{\omega})^2 \, du , \qquad (257)$$

$$J_2 = \int_{-\infty}^{\infty} \frac{S^2(u)}{R^2(u)} \, du = 2 \int_0^{\infty} \frac{S^2(u)}{R^2(u)} \, du . \qquad (258)$$

In (257), $L(-u) \neq L(u)$, so that it is convenient to resolve the integrand into one component with even symmetry with respect to u and one with odd symmetry. The integral of the latter component vanishes and that of the component with even symmetry becomes

$$J_1 = \int_0^{\infty} H^2(u) R^2(u) \, du , \qquad (259)$$

raised cosine pulse, which overlaps into adjacent pulse intervals, and is considered in the next section. With the latter type of pulse it turns out that $S^o(\bar{\omega}) = 0.994\ T/2$. It can thus be conjectured that there is some intermediate shape of overlapping modulating pulses for which (71) is satisfied. For practical purposes, this is the case with modulating pulses given by (64) or for the raised cosine pulses considered in the next section.

IX. RAISED COSINE MODULATING PULSES

Raised cosine modulating pulses have the shape indicated in Fig. 9, and can be derived conveniently by appropriate gating of a biased steady-state sine wave, with the total pulse duration, $2T$, equal to one cycle of the sine wave, which may have advantages from the standpoint of instrumentation.

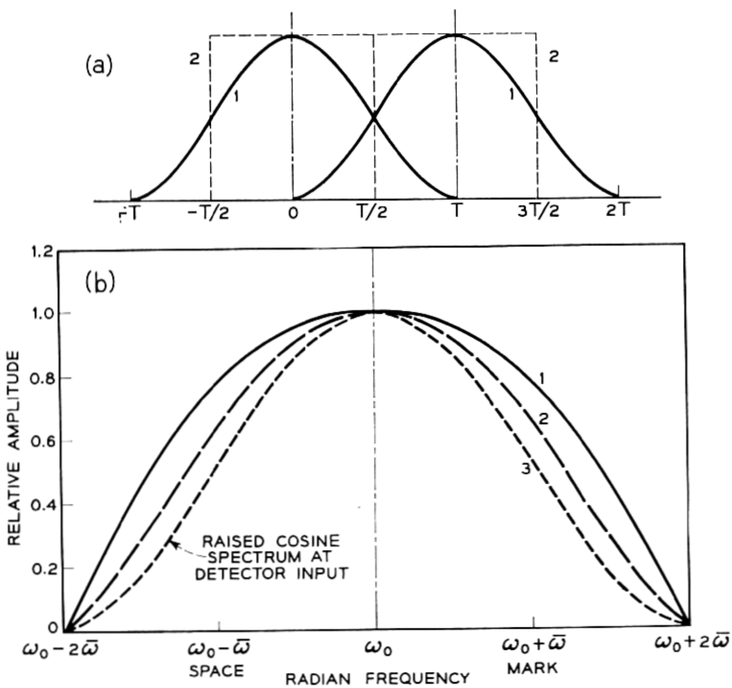

Fig. 9 — Comparison of transmission-frequency characteristics of bandpass channel in FM for raised cosine modulating pulses (solid curve) and rectangular modulating pulses (dashed curve), with raised cosine spectrum at detector input (dotted curve) in both cases. Pulse shapes are same as those for FM in Fig. 6.

With a raised cosine modulating voltage,

$$
\begin{aligned}
V_i(t_0) &= \frac{V_0}{2}[1 + \cos(\pi t_0/T)] \qquad -1 < t_0/T < 1 \\
&= 0 \qquad\qquad\qquad\qquad\qquad\quad -1 > t_0/T > 1.
\end{aligned}
\tag{78}
$$

In accordance with (2), the resultant phase modulation is, for $\bar{\omega} t_0 \leqq \pi$,

$$
\begin{aligned}
\psi_i(t_0) &= \frac{\bar{\omega}_1 V_0}{2} \int_0^{t_0} [1 + \cos(\pi t_0/T)]\, dt_0 \\
&= \tfrac{1}{2}(\bar{\omega} t_0 + \sin \bar{\omega} t_0).
\end{aligned}
\tag{79}
$$

For $\bar{\omega} t_0 \geqq \pi$, $\psi_i(t_0) = \psi_i(T) = \pi/2$.

The spectrum of $\cos \psi_i(t_0)$ is obtained from (67) with the upper limit equal to T, since $\cos \psi_i(t_0) = 0$ for $t_0 > T$. The following relation is thus obtained:

$$
S^o(u) = 2 \int_0^T \cos \tfrac{1}{2}(\bar{\omega} t_0 + \sin \bar{\omega} t_0) \cos u t_0 \, dt_0 . \tag{80}
$$

With $\bar{\omega} t_0 = x$, $\bar{\omega} T = \pi$ and $u/\bar{\omega} = a$, (80) becomes

$$
\begin{aligned}
S^o(u) &= \frac{2}{\bar{\omega}} \int_0^\pi \cos(x/2 + \tfrac{1}{2}\sin x) \cos ax \, dx \\
&= \frac{T}{\pi} \int_0^\pi \cos(\nu_1 x + \tfrac{1}{2}\sin x)\, dx \\
&\quad + \frac{T}{\pi} \int_0^\pi \cos(\nu_2 x + \tfrac{1}{2}\sin x)\, dx ,
\end{aligned}
\tag{81}
$$

where

$$
\nu_1 = \tfrac{1}{2} - a, \qquad \nu_2 = \tfrac{1}{2} + a.
$$

The above relation can be written

$$
S^o(u) = T\,[J_{\nu_1}(-\tfrac{1}{2}) + J_{\nu_2}(-\tfrac{1}{2})], \tag{82}
$$

where $J_\nu(z)$ is a so-called Anger function, which is associated with Bessel functions and is defined by[8]

$$
J_\nu(z) = \frac{1}{\pi} \int_0^\pi \cos(\nu x - z \sin x)\, dx, \tag{83}
$$

$$J_\nu(z) = \frac{\sin \nu\pi}{\nu\pi}\left[1 - \frac{z^2}{2^2 - \nu^2} + \frac{z^4}{(2^2 - \nu^2)(4^2 - \nu^2)} - \frac{z^6}{(2^2 - \nu^2)(4^2 - \nu^2)(6^2 - \nu^2)} + \cdots\right] + \frac{\sin \nu\pi}{\pi}\left[\frac{z}{1^2 - \nu^2} - \frac{z^3}{(1^2 - \nu^2)(3^2 - \nu^2)} + \frac{z^5}{(1^2 - \nu^2)(3^2 - \nu^2)(5^2 - \nu^2)} - \cdots\right]. \tag{84}$$

The spectrum as a function of a obtained from (82) and (84) is given in Table III, together with that for a rectangular modulating wave, as obtained from (26).

TABLE III — VALUES OF $(2/T)S^o(u)$ FOR RAISED COSINE AND RECTANGULAR MODULATING PULSES

$a = u/\bar{\omega}$	0	0.5	1	1.5	2
Raised cosine.	1.58	1.39	0.994	0.544	0.282
Rectangular.	1.28	1.20	1.000	0.720	0.425

To obtain a raised cosine spectrum at the detector input, as given by (61), the required amplitude characteristic of the bandpass channel obtained from (70) is

$$A(-u) = A(u) = \frac{\cos^2(\pi u/4\bar{\omega})}{2[J_{\nu_1}(-\frac{1}{2}) + J_{\nu_2}(-\frac{1}{2})]}, \tag{85}$$

which gives for various values of $a = u/\bar{\omega}$

$a = u/\bar{\omega}$:	0	0.5	1	1.5	2
$A(u)$:	0.63	0.61	0.503	0.275	0

Since $A(\bar{\omega}) = 0.503$, rather than 0.50, the factor $\mu = 2$ in (43) is replaced by $\mu = 2.012$. The peak pulse amplitude at $t_0 = 0$ as obtained from (43) is, in this case,

$$P(0) = \frac{1}{2}\frac{2.012}{2.012 - 1} = 0.994.$$

Thus, intersymbol interference in this case results in a slight reduction in the peak amplitude of a pulse.

In Fig. 9 the above channel characteristic is compared with that for rectangular modulating pulses. In both cases the shape of the demodulated pulses would be as shown in Fig. 6 for FM, aside from the slightly smaller peak amplitude, 0.994 rather than 1.00. It can be shown that, with raised cosine modulating pulses, the error involved in neglecting the second integral in (24) is not appreciable, even when the total channel band width is equal to the midband frequency. Hence, the channel shaping shown in Fig. 9 for raised cosine modulating pulses applies without restriction on channel bandwidth relative to midband frequency.

X. POSTDETECTION PULSE SPECTRA AND FILTERING

At the detector output a low-pass filter may be desirable for final pulse shaping, elimination of unwanted demodulation products or higher-frequency noise components, as noted in Section II. The appropriate transmission-frequency characteristic of such filters depends on the spectra of the demodulated pulses, as discussed here.

Let $P(t_0)$ be the shape of the pulses at the detector output in FM as given for rectangular modulating pulses by (41) and illustrated in Figs. 4 and 6 for two special cases. These pulses have a baseband spectrum given by

$$S_0(\omega) = \int_{-\infty}^{\infty} P(t_0) e^{-i\omega t_0}\, dt_0 \tag{86}$$

or, since $P(-t_0) = P(t_0)$, by

$$S_0(\omega) = 2 \int_0^{\infty} P(t_0) \cos \omega t_0 \, dt_0 . \tag{87}$$

In accordance with the definition in Section IV, $P(t_0) = \psi'(t_0)/2\bar{\omega}$. Hence, for $\omega = 0$, (87) becomes

$$\begin{aligned} S_0(0) &= \frac{1}{2\bar{\omega}} \int_{-\infty}^{\infty} \psi'(t_0)\, dt_0 \\ &= \frac{1}{2\bar{\omega}} [\psi(\infty) - \psi(-\infty)] . \end{aligned} \tag{88}$$

In view of (19), the phase change caused by transmission of a mark preceded and followed by a continuing space, is $\psi(\infty) - \psi(-\infty) = 2\bar{\omega}T = 2\pi$. Hence (88) becomes

$$S_0(0) = T. \tag{89}$$

In the case of AM, the spectrum of the baseband pulses is equal to the spectrum at the detector input above the midband frequency multiplied by a factor of 2, because of the direct addition of the two sideband spectra. From Fig. 2 it follows that, in this case, $S_0(0) = T$. Relation (89) thus shows that the dc component of the demodulated pulses is the same in FM as in AM. That is, the areas under the FM and AM pulses shown in Figs. 4 and 5 are equal.

For equal spectra at the detector input in AM and FM, the pulse shape $P(t_0)$ has a nonlinear relation (41) to the pulse shape $p(t_0)$ in AM. For this reason, the bandwidth of the demodulated pulses will be greater in FM than in AM, and, in view of the appearance of $p(t_0)$ in the denominator of (41), the bandwidth is theoretically infinite. For this reason, part of the spectrum will be eliminated by any postdetection low-pass filter, and intersymbol interference is thereby introduced unless the filter has an appropriate amplitude characteristic.

With the aid of a postdetection low-pass filter having an amplitude characteristic $A_0(\omega)$ and a linear phase characteristic, it is possible to modify the spectrum S_0 into a desired spectrum S_m with such properties that intersymbol interference is absent. To this end, the amplitude characteristic would be so chosen that

$$A_0(\omega)S_0(\omega) = S_m(\omega). \tag{90}$$

For example, by appropriate choice of $A_0(\omega)$ the pulse shape shown in Fig. 4 for FM could be modified into that shown for AM, or into other shapes. The principal difficulty resides in the determination of the spectrum $S_0(\omega)$ from (87), which entails numerical integration, in view of the fairly complicated expressions for $P(t_0)$.

The spectrum obtained by numerical integration of (87) is given in Table IV for the special case of a flat spectrum of minimum bandwidth at the detector input, as considered in Section VI. In the numerical integration, contributions to the integral were neglected for $\bar{\omega}t_0 = \tau > 20$.

The above spectrum is shown in Fig. 10, together with the baseband

TABLE IV — $S_0(\omega)/T$ FOR FLAT SPECTRUM AT DECTECTOR INPUT

$u/\bar{\omega}$	0	0.25	0.5	0.75	0.9
0	1*	0.8779	0.7561	0.6254	0.5385
1	0.4755	0.2932	0.0626	−0.2370	−0.4609
2	−0.1321	0.2871	0.2031	0.4186	−0.0696
3	0.0396	−0.0233	−0.0529	−0.0220	0.0418
4	0.1105	0.0650	0.0270	0.0003	—

* Based on (89); compares with computed value of 0.99999654.

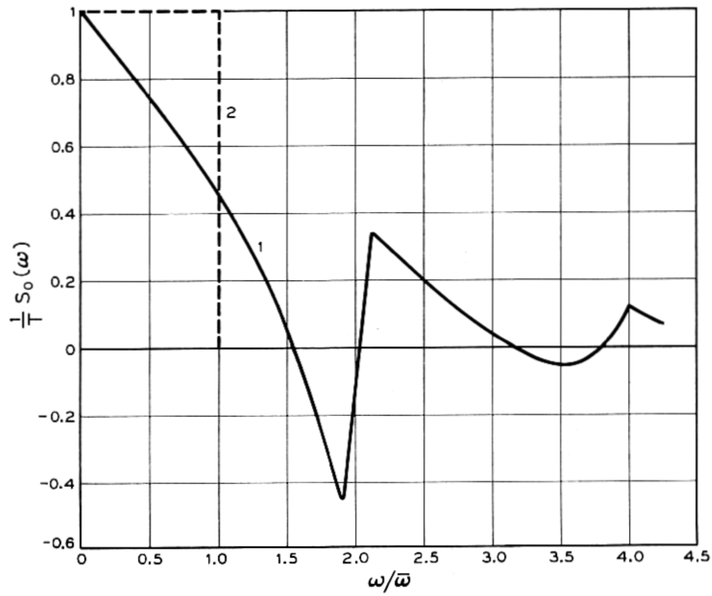

Fig. 10 — Spectra, $S_0(\omega)$, of demodulated FM (solid curve) and AM (dashed curve) pulses shown in Fig. 4 for flat spectrum at detector input.

spectrum of the AM pulses. It will be noted that the spectrum is negative in certain ranges for $\omega > 1.6\bar{\omega}$. This places a restriction on the choice of the modified spectrum S_m and on the filter bandwidth, if $A_0(\omega)$ is to be positive and finite for all values of ω in the filter band. To this end, it is necessary that the filter bandwidth be less than $1.6\bar{\omega}$.

In Fig. 11 is shown the amplitude characteristic $A_0(\omega)$ of the post-detection low-pass filter obtained from (90) when $S_m(\omega)$ is assumed to be equal to the AM baseband spectrum. With the amplitude characteristic $A_0(\omega)$ shown in Fig. 11, the FM pulses shown in Fig. 4 would be converted into pulses of the same shape as shown for AM.

With a raised cosine spectrum at the detector input, as considered in Section VII, the spectrum of the demodulated pulses obtained by numerical integration of (87) is given in Table V. In the numerical integration, contributions to the integral for $\bar{\omega}t_0 = \tau > 5$ were disregarded.

The above spectrum is shown in Fig. 12, together with that of the AM pulses.

The circumstance that $S_0(\omega)$ is negative in the approximate range $1.9\bar{\omega} < \omega < 2.1\bar{\omega}$ in this case limits the filter bandwidth to less than $1.9\bar{\omega}$, if $A_0(\omega)$ is to be positive and finite for all values of ω in the filter band.

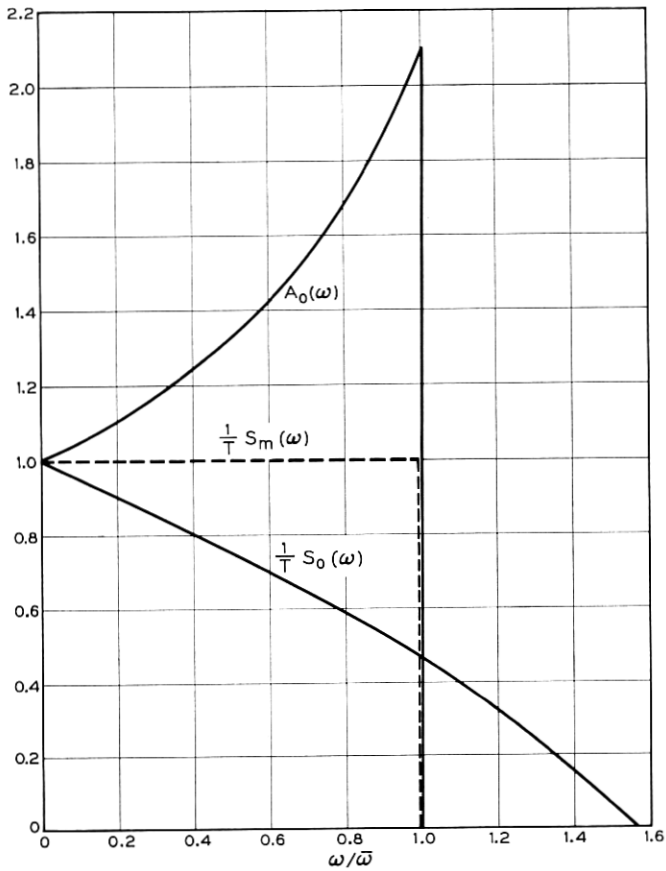

Fig. 11 — Amplitude characteristic, $A_0(\omega)$, of postdetection low-pass filter required to convert spectrum $S_0(\omega)$ into modified spectrum $S_m(\omega)$.

TABLE V — $S_0(\omega)/T$ FOR RAISED COSINE SPECTRUM AT DETECTOR INPUT

$u/\bar{\omega}$	0	0.25	0.5	0.75	1.0
0	1*	0.9492	0.8092	0.6147	0.4159
1	0.4159	0.2438	0.1063	0.0208	−0.0069
2	−0.0069	0.0131	0.0582	0.0926	0.0742
3	0.0742	0.0302	0.0093	0.0032	0.0052
4	0.0052	0.0096	0.0121	0.0114	0.0084

* From (89); actual computed value = 0.99999951.

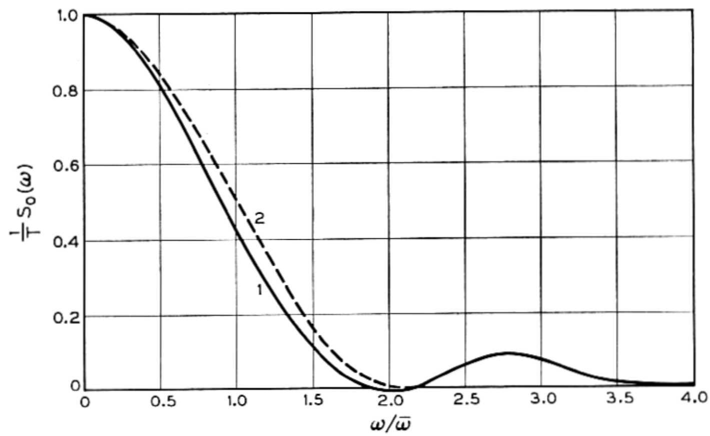

Fig. 12 — Spectra $S_0(\omega)$ of demodulated FM (solid curve) and AM (dashed curve) pulses shown in Fig. 6 for raised cosine spectrum at detector input.

By way of illustration, an appropriate type of modified spectrum of bandwidth $1.75\bar{\omega}$ is given by the following expressions.

For $0 < \omega/\bar{\omega} < \frac{1}{4}$:

$$S_m(\omega) = T. \tag{91}$$

For $\frac{1}{4} < \omega/\bar{\omega} < \frac{7}{4}$:

$$S_m(\omega) = T \cos^2 \left(\frac{\pi}{4} \frac{4\omega/\bar{\omega} - 1}{3} \right). \tag{92}$$

The above spectrum S_m is shown in Fig. 13, together with the spectrum $S_0(\omega)$ and the amplitude characteristic $A_0(\omega)$ of the low-pass filter obtained from (90). The shape of the pulses at the output of the filter for the above spectrum $S_m(\omega)$ can be obtained from (52) with $\omega_x = \frac{3}{4}\bar{\omega}$, but it does not differ significantly from that shown in Fig. 6 for AM.

With a low-pass filter having a linear phase and an amplitude characteristic $A_0(\omega)$ as shown in Fig. 13, intersymbol interference is avoided, and some improvement in signal-to-noise ratio is realized by elimination of higher-frequency noise components in the detector output, as will be shown later.

XI. PULSE TRAINS

In the preceding analysis, transmission of a single mark of duration T was assumed. When a pulse train consisting of a sequence of marks and spaces is transmitted, (33) is modified into the following expression

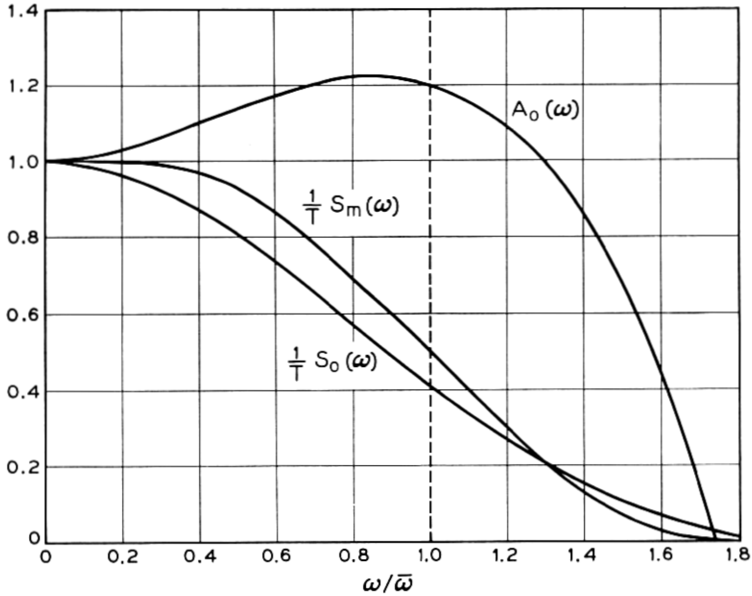

Fig. 13 — Amplitude characteristic, $A_0(\omega)$, of postdetection low-pass filter required to convert spectrum $S_0(\omega)$ of Fig. 12 into a modified spectrum $S_m(\omega)$.

for the received wave at the detector input:

$$W_i(t_0) = -\cos(\omega_0 t_0 + \varphi_0)\,[A(-\bar{\omega})\cos\bar{\omega}(t_0 - mT) - \sum_{m=-\infty}^{\infty} a_m p(t_0 - mT)] - A(-\bar{\omega})\sin(\omega_0 t_0 + \varphi_0)\sin\bar{\omega}(t_0 - mT), \tag{93}$$

where $a_m = 0$ for a space and 1 for a mark, and m is an integer. The above expression can be written in the form

$$W_i(t_0) = \bar{W}_i(t_0)(\cos\omega_0 t_0 + \varphi_0 + \Psi_0), \tag{94}$$

where the pulse train envelope is given by the following expression in place of (36):

$$\bar{W}_i(t_0) = A(-\bar{\omega})[1 + \mu^2 \sum a_m^2 p^2(t_0 - mT) - 2\mu \sum a_m p(t_0 - mT)\cos\bar{\omega}(t_0 - mT)]^{1/2}, \tag{95}$$

where Σ indicates summation between the limits $m = -\infty$ and $m = \infty$, as in (93). Expression (37) for the phase Ψ_0 is replaced by

$$\tan\Psi_0 = \frac{\sin\bar{\omega}(t_0 - mT)}{\cos\bar{\omega}(t_0 - mT) - \mu\sum a_m p(t_0 - mT)}. \tag{96}$$

Expression (39) for the time derivative of Ψ_0 is replaced by

$$\Psi_0' = -\frac{\bar{\omega}}{D}\Big[1 - \mu \sum a_m p(t_0 - mT) \cos \bar{\omega}(t_0 - mT) + \frac{\mu}{\bar{\omega}} \sum a_m p'(t_0 - mT) \sin \bar{\omega}(t_0 - mT)\Big], \tag{97}$$

where

$$D = \sin^2 \bar{\omega}(t_0 - mT) + [\cos \bar{\omega}(t_0 - mT) - \mu \sum a_m p(t_0 - mT)]^2. \tag{98}$$

Expression (41) is replaced by the following expression for the demodulated pulse train:

$$W(t_0) = \frac{\mu}{2D}\Big\{\mu\Big[\sum a_m p(t_0 - mT)\Big]^2 - \sum a_m p(t_0 - mT) \cdot \cos \bar{\omega}(t_0 - mT) - \frac{1}{\bar{\omega}} \sum a_m p'(t_0 - mT) \sin \bar{\omega}(t_0 - mT)\Big\}. \tag{99}$$

At the sampling points $t_0 = 0$, $\sin \bar{\omega} mT = 0$ and $\cos \bar{\omega} mT = (-1)^m$, so that (99) becomes

$$W(0) = \frac{\mu}{2} \frac{\sum a_m p(mT)}{\mu \sum a_m p(mT) - (-1)^m}. \tag{100}$$

Since $p(mT) = 0$ except for $m = 0$, (100) becomes

$$W(0) = \frac{\mu}{2} \frac{a_0 p(0)}{\mu a_0 p(0) - 1}, \tag{101}$$

where $\mu = 2$ and $p(0) = 1$, so that

$$\begin{aligned} W(0) &= \frac{a_0}{2a_0 - 1} \\ &= 1 \quad \text{for} \quad a_0 = 1 \text{ (mark)} \\ &= 0 \quad \text{for} \quad a_0 = 0 \text{ (space)}. \end{aligned} \tag{102}$$

There is thus no intersymbol interference when a pulse train is transmitted.

In (99) the denominator D depends on the composition of the pulse train. For this reason the shape of the demodulated pulse train between sampling points cannot be obtained by direct superposition of individual demodulated pulses, such as those shown in Fig. 4 and Fig. 6.

XII. AVERAGE SIGNAL POWER

The preceding analysis was concerned with the ideal shaping of the over-all transmission-frequency characteristic of the bandpass channel required to avoid intersymbol interference in FM and AM systems. This over-all shaping will ordinarily be divided between bandpass filters at the transmitting and of the receiving ends. While the division between the two ends is immaterial from the standpoint of intersymbol interference, it does affect signal power at the transmitting end and interference of various kinds at the receiving end. In order to determine the optimum proportioning and shaping between transmitting and receiving ends, it is thus necessary to consider both signal power and interference.

In analog systems for voice transmission and other purposes, the peak signal power is ordinarily substantially greater than the average signal power, by 10 db or more, and is usually a limitation in systems design. In binary pulse systems with representative transmission-frequency characteristics,* however, peak signal power is not much greater than average signal power, and the latter is ordinarily a limitation, either from the standpoint of repeater design or interference with other systems. For this reason, average signal power will be considered here in comparing FM and AM binary pulse systems.

Let the amplitude characteristic of the transmitting filter be $T(u)$ and that of the receiving filter be $R(u)$, in which case $T(u)R(u) = A(u)$, where $A(u)$ is the over-all amplitude characteristic and u the frequency from midband.

Let the peak amplitude of the carrier for a continuous mark or space at the output of the transmitting filter be E, in which case the signal power for a continuous mark is

$$P_m = \frac{E^2}{2}. \tag{103}$$

In Appendix A it is shown that the average signal power for random pulse trains with FM and with bipolar AM are given by

$$P_{\text{FM}} = P_m \frac{4}{\pi T} R^2(\bar{\omega}) \int_{-\infty}^{\infty} \left[\frac{S(u)}{R(u)}\right]^2 du, \tag{104}$$

$$P_{\text{AM}} = P_m \frac{2}{\pi T} R^2(0) \int_{-\infty}^{\infty} \left[\frac{S(u)}{R(u)}\right]^2 du, \tag{105}$$

where T is the pulse interval or duration, and $S(u)$ is the spectrum at

* This excludes idealized flat channels of minimum bandwidth.

the channel output, or detector input, in response to the transmission of a single pulse, i.e., in response to the transmission of a mark preceded and followed by a continuing space.

By way of illustration, in the case of a raised cosine spectrum as given by (61)

$$P_{\mathrm{FM}} = P_m \frac{T}{\pi} R^2(\bar{\omega}) \int_{-2\bar{\omega}}^{2\bar{\omega}} \frac{\cos^4 (\pi u/4\bar{\omega})}{R^2(u)} du, \tag{106}$$

$$P_{\mathrm{AM}} = P_m \frac{T}{2\pi} R^2(0) \int_{-2\bar{\omega}}^{2\bar{\omega}} \frac{\cos^4 (\pi u/4\bar{\omega})}{R^2(u)} du. \tag{107}$$

If the receiving filter has a half-cosine shape as given by

$$\begin{aligned} R(u) &= \cos \pi u/4\bar{\omega} && \text{for} \quad -2\bar{\omega} < u < 2\bar{\omega} \\ &= 0 && \text{for} \quad |u| > 2\bar{\omega} \end{aligned} \tag{108}$$

expressions (106) and (107) become, with $R(\bar{\omega}) = \frac{1}{2}^{1/2}$ and $R(0) = 1$,

$$\begin{aligned} P_{\mathrm{FM}} = P_{\mathrm{AM}} &= P_m \frac{T}{2\pi} \int_{-2\bar{\omega}}^{2\bar{\omega}} \cos^2 \pi u/4\bar{\omega}) \, du \\ &= P_m . \end{aligned} \tag{109}$$

In this particular case the average signal power of a random pulse train in both FM and bipolar AM is equal to the signal power for a continuous mark or space.

Consider next a flat receiving filter, in which case

$$\begin{aligned} R(u) &= 1 && \text{for} \quad -2\bar{\omega} < u < 2\bar{\omega} \\ &= 0 && \text{for} \quad |u| > 2\bar{\omega}. \end{aligned} \tag{110}$$

In this case (106) and (107) yield

$$\begin{aligned} P_{\mathrm{FM}} &= P_m \frac{T}{\pi} \int_{-2\bar{\omega}}^{2\bar{\omega}} \cos^4 \pi u/4\bar{\omega} \, du \\ &= \tfrac{3}{2} P_m , \end{aligned} \tag{111}$$

$$P_{\mathrm{AM}} = \tfrac{1}{2} P_{\mathrm{FM}} = \tfrac{3}{4} P_m . \tag{112}$$

Finally, let the receiving filter have a raised cosine shape as given by

$$\begin{aligned} R(u) &= \cos^2(\pi u/4\bar{\omega}) && \text{for} \quad -2\bar{\omega} < u < 2\bar{\omega} \\ &= 0 && \text{for} \quad |u| > 2\bar{\omega}. \end{aligned} \tag{113}$$

In this case $R(\bar{\omega}) = \frac{1}{2}$ and

$$P_{\text{FM}} = P_m \frac{T}{4\pi} \int_{-2\bar{\omega}}^{2\bar{\omega}} du \\ = P_m, \tag{114}$$

$$P_{\text{AM}} = 2P_{\text{FM}} = 2P_m. \tag{115}$$

The above relations show that, for equal signal power in FM and bipolar AM during transmission of a continuing mark (or space), there may be appreciable difference in average signal power for a random pulse train, depending on the shape of the transmitting filter. For this reason, signal power during a continuing mark, which is often used in specifying signal-to-noise ratio, may not be an appropriate reference signal power.

To determine the optimum division of channel shaping between transmitting and receiving ends, it is necessary to consider the effect of random noise or other interference, such as impulse noise. The effect of any particular type of interference depends on the shape of the receiving filter, as discussed in the following sections for random noise. Impulse noise is discussed briefly in Appendix B.

XIII. RANDOM NOISE IN FM AND AM SYSTEMS

Certain basic equations relating to noise and interference on FM and AM are given in Appendix B and applied to the particular case of single-frequency interference. In the case of a sinusoidal interfering voltage at a frequency u from midband and of amplitude $e(u)$ at the input of the receiving filter, rms interference in FM and bipolar AM taken in relation to the peak-to-peak signal amplitude at the detector output is given by

$$\bar{\eta}_{\text{FM}} \cong \frac{\bar{e}(u)R(u)}{2ER(\omega)} (1 + u/\bar{\omega}), \tag{116}$$

$$\bar{\eta}_{\text{AM}} = \frac{\bar{e}(u)R(u)}{2ER(0)} \tag{117}$$

where $\bar{e}(u) = e(u)/2^{1/2}$, $R(u)$ is the amplitude characteristic of the receiving filter and E is the peak amplitude of the carrier for a continuing mark or space. The above relation for $\bar{\eta}_{\text{FM}}$ is a first-order approximation applying if $\bar{e}(u)$ is small in relation to E, as is required for transmission without excessive error rates.

The equations give the rms amplitude of the interfering voltage at the detector output, taken in relation to the peak-to-peak difference in pulse amplitudes for mark and space, considering all possible phases of the interfering voltage equally probable. The slicing or threshold level is ordinarily half this difference, and rms interference voltage taken in relation to the slicing level would be twice as great.

Random noise can be regarded as the sum of a very large number of sinusoidal waves of different frequencies, with both the amplitude and phase of each component wave varying with time. For a single sinusoidal wave at a frequency u from midband and peak amplitude e, the rms amplitude is $\bar{e} = e/2^{1/2}$. When the sinusoidal wave varies in amplitude with time there will be a certain rms amplitude $\mathbf{e}$ over a long interval, and a corresponding average noise power $\mathbf{e}^2$ for each sinusoidal component. The corresponding average noise power per unit of bandwidth at the receiving filter input will be designated $n(u)$. In the case of white thermal noise as assumed in the following, $n(u) = n$ is independent of u.

The ratio of noise power to signal power at the output of the detector can be obtained from (116) and (117) for single-frequency waves, by integration of the noise power density over the channel band. Thus, in the case of FM, the ratio of average output noise power N_0 to the output peak-to-peak signal power $\hat{S}_0$ between mark and space, as obtained by integration of (116), becomes

$$(N_0/\hat{S}_0)_{\mathrm{FM}} = \frac{n}{8P_m R^2(-\bar{\omega})} \int_{-\infty}^{\infty} R^2(u)(1 + u/\bar{\omega})^2 \, du \tag{118}$$

$$= \frac{n}{8P_m R^2(-\bar{\omega})} \int_{-\infty}^{\infty} R^2(u)(1 + u^2/\bar{\omega}^2) \, du, \tag{119}$$

where the last expression follows since $R(-u) = R(u)$, so that the integral of $R^2 2u/\bar{\omega}$ vanishes.

In case of bipolar AM, the corresponding ratio obtained by integration of (117) is

$$(N_0/\hat{S}_0)_{\mathrm{AM}} = \frac{n}{8P_m R^2(0)} \int_{-\infty}^{\infty} R^2(u) \, du. \tag{120}$$

Relations (119) and (120) can be expressed in terms of average signal power in FM and AM with the aid of relations (104) and (105). The following expressions are thus obtained:

$$(N_0/\hat{S}_0)_{\mathrm{FM}} = \frac{2n}{4\pi T P_{\mathrm{FM}}} \left[\int_{-\infty}^{\infty} R^2(u)(1 + u^2/\bar{\omega}^2) \, du \right] \cdot \left[\int_{-\infty}^{\infty} \frac{S^2(u)}{R^2(u)} \, du \right], \tag{121}$$

$$(N_0/\hat{S}_0)_{\mathrm{AM}} = \frac{n}{4\pi T P_{\mathrm{AM}}} \left[\int_{-\infty}^{\infty} R^2(u)\, du \right] \left[\int_{-\infty}^{\infty} \frac{S^2(u)}{R^2(u)}\, du \right]. \tag{122}$$

The above expressions for FM apply without a postdetection low-pass filter. As discussed in Section X, it is possible to modify the pulse shape in FM without causing intersymbol interference, with the aid of a low-pass filter of appropriate amplitude characteristic $A_0(\omega)$ that depends on the bandwidth. Such a filter reduces the noise power in a narrow band at ω by the factor $A_0^2(\omega)$. On a carrier basis, this is equivalent to multiplying the noise power in a narrow band at a frequency u from

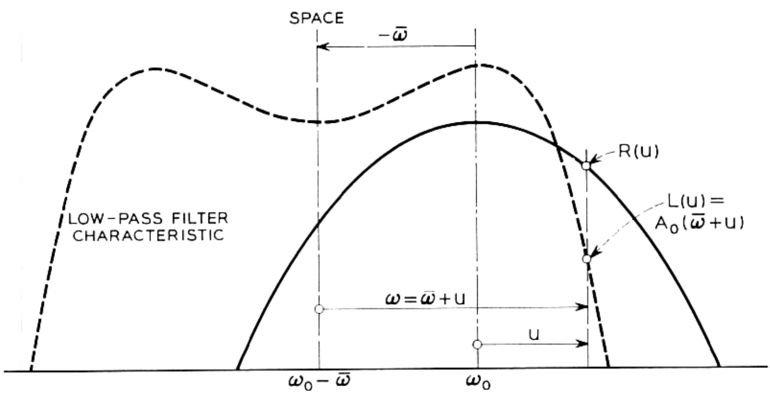

Fig. 14 — Frequency modulation with postdetection low-pass filter having transmission characteristic $A_0(\omega)$. Noise power in the narrow band at frequency u from midband is reduced by the factor $L^2(u) = A_0^2(\bar{\omega} + u)$ during transmission of space with carrier at $\omega_0 - \bar{\omega}$, as assumed above.

midband by $A_0^2(\bar{\omega} + u)$ during transmission of a space, as indicated in Fig. 14, or by $A_0^2(-\bar{\omega} + u)$ during transmission of a mark. This equivalent representation on a carrier basis is legitimate, provided the carrier power at the detector input is substantially greater than the noise power, in which case the ratio $N_0/\hat{S}_0$ obtained from (121) is much smaller than one.

The following notation indicated in Fig. 14 will be used:

$$L(u) = A_0(\bar{\omega} + u). \tag{123}$$

With such a filter, (118) is modified into

$$(N_0/\hat{S}_0)_{\mathrm{FM}} = \frac{n}{8P_m R^2(-\bar{\omega})} \int_{-\infty}^{\infty} L^2(u) R^2(u)(1 + u/\bar{\omega})^2\, du, \tag{124}$$

and (121) is modified into

$$(N_0/\hat{S}_0)_{\text{FM}} = \frac{2n}{4\pi T P_{\text{FM}}} \left[\int_{-\infty}^{\infty} L^2(u) R^2 (1 + u/\bar{\omega})^2 \, du \right] \cdot \left[\int_{-\infty}^{\infty} \frac{S^2}{R^2} \, du \right], \tag{125}$$

where L, R and S are functions of u.

Comparison of (125) with (121) shows that, for a given $R(u)$, a post-detection low-pass filter reduces the ratio $(N_0/\hat{S}_0)_{\text{FM}}$ by the factor

$$\rho = \frac{\int_{-\infty}^{\infty} L^2 R^2 (1 + u/\bar{\omega})^2 \, du}{\int_{-\infty}^{\infty} R^2 (1 + u^2/\bar{\omega}^2) \, du}, \tag{126}$$

where L and R are functions of u.

XIV. OPTIMUM AM SYSTEMS

The minimum value of $(N_0/\hat{S}_0)$ for a given average power $P = P_{\text{AM}}$ is obtained when the product of the two integrals in (122) is a minimum. As shown in Appendix C, this is the case when $R(u)$ is such that the two integrals are equal, in which case the optimum $R(u)$ is given by

$$R^o(u) = cS^{1/2}(u), \tag{127}$$

where c is an arbitrary constant independent of u that can be chosen to give $R(u)$ the appropriate dimension.

With (127) in (122), the optimum ratio $N_0/\hat{S}_0$ becomes

$$(N_0/\hat{S}_0)^o = \frac{n}{4\pi T P} \left[\int_{-\infty}^{\infty} S(u) \, du \right]^2. \tag{128}$$

When $S(u)$ has the properties previously discussed and illustrated in Fig. 1 and Fig. 2, the integral in (128) is always equal to the area under the rectangle in the upper part of these figures, or $2\bar{\omega}T/2 = \bar{\omega}T = \pi$. Furthermore, with (127) in (105), it follows that $P_m = P_{\text{AM}} = P$. Hence, for all spectra $S(u)$ of the form previously assumed, (128) becomes

$$(N_0/\hat{S}_0)^o = \frac{N(\bar{\omega}T)^2}{4\pi T P} = \frac{N}{4P} = \frac{N}{4P_m}, \tag{129}$$

where

$N = n\bar{\omega} =$ average noise power in a flat band $\bar{\omega}$ at input of receiving filter,
$P =$ average signal power at receiving filter input,
$P_m =$ signal power for continuous mark (or space),
$N_0 =$ average noise power at detector output,
$\hat{S}_0 =$ peak-to-peak signal power at detector output.

In (129) N is the average noise power in a flat band equal to that of the minimum possible bandwidth $\bar{\omega}$ over which pulses can be transmitted at intervals $T = \pi/\bar{\omega}$ without intersymbol interference, i.e., the noise power in a band $1/2T$ cps. With the above definition of N and for the above pulse transmission rate, (129) applies for all spectra at the detector input with the properties illustrated in Figs. 1 and 2. There is thus no noise penalty, but only a bandwidth penalty, in modifying the spectrum that was indicated in Fig. 2 and discussed previously to obtain pulses whose shape is more appropriate shape than that of systems with the minimum possible bandwidth. Moreover, (129) also applies for optimum bipolar baseband systems with base band spectra equal to those shown on the right-hand side of ω_0 in Figs. 1 and 2. Thus there is no noise penalty in bipolar double sideband AM, but the bandwidth of the carrier channel is twice that of the baseband channel.

This two-fold increase in bandwidth for a given transmission rate can be overcome by providing two independent channels on two carriers at quadrature, a method sometimes referred to as four-phase transmission. With 3-db reduction in noise power, because of the two-fold reduction in bandwidth, and with 3 db less signal power per channel, so that the average signal power P is the same as for a single channel in bipolar AM, (129) applies with N defined as above An alternate means of avoiding the two-fold increase in bandwidth is to use bipolar vestigial sideband AM[1,2] with homodyne detection, in which case (129) also applies. The last two methods are thus equivalent to bipolar baseband transmission, both as regards bandwidth and signal-to-noise ratio.

At the detector output the pulses may be bipolar or may be biased into unipolar (on-off) pulses, and (129), in terms of the peak-to-peak signal power $\hat{S}_0$ at sampling instants, applies regardless of any bias.

In the particular case of a raised cosine spectrum as given by (71), the over-all amplitude characteristic of the channel is given by (63) or

$$a(u) = \frac{\pi u/4\bar{\omega}}{\tan(\pi u/4\bar{\omega})}. \tag{130}$$

The optimum receiving filter characteristic as obtained from (127) is

$$R^o(u) = \cos(\pi u/4\bar{\omega}). \tag{131}$$

The corresponding optimum transmitting filter characteristic is $T^o = a(u)/R^o$, or

$$\begin{aligned} T^o(u) &= \frac{\pi u/4\bar{\omega}}{\sin(\pi u/4\bar{\omega})} \\ &= 1 && \text{for} \quad u = 0 \\ &= \frac{\pi}{4}\,2^{1/2} \cong 1.12 && \text{for} \quad u = \bar{\omega} \\ &= \frac{\pi}{2} && \text{for} \quad u = 2\bar{\omega}. \end{aligned} \tag{132}$$

These characteristics are shown in Fig. 15.

In the case of vestigial-sideband AM, the carrier would be at $\omega_0 + \bar{\omega}$ or $\omega_0 - \bar{\omega}$ rather than at ω_0. Pulses can then be transmitted at twice the double-sideband rate, provided homodyne detection is used so that the effect of the quadrature component is eliminated.[2] The optimum shape of the receiving filter is again given by (131), but the shape of the associated optimum transmitting filter is modified, since the spectrum at the

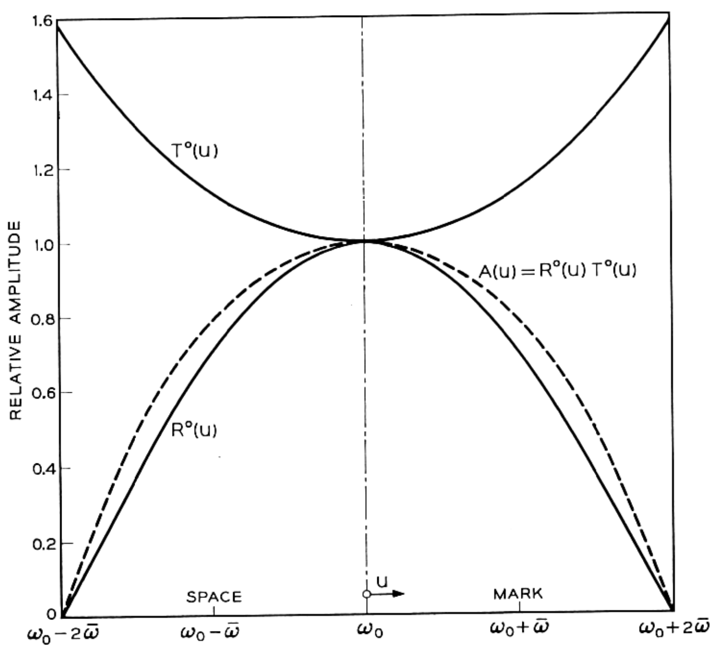

Fig. 15 — Double-sideband AM with raised cosine spectrum at detector input: R^o = optimum shape of receiving filter; T^o = optimum shape of transmitting filter; $A = R^oT^o$ = combined transmission characteristic.

channel input is now caused by a rectangular carrier pulse of frequency $\omega_0 \pm \bar{\omega}$ rather than ω_0. The optimum transmitting filter characteristic is given by

$$T^o(u) = \cos(\pi u/4\bar{\omega}) \frac{\pi(u \pm \bar{\omega})/4\bar{\omega}}{\sin[\pi(u \pm \bar{\omega})/4\bar{\omega}]}. \qquad (133)$$

The positive signs apply when the carrier is at $\omega_0 - \bar{\omega}$ and the negative signs apply when it is at $\omega_0 + \bar{\omega}$; the expression applies for $-2\bar{\omega} \leqq u \leqq 2\bar{\omega}$.

If the carrier is assumed at $\omega_0 + \bar{\omega}$,

$$\begin{aligned} T^o(u) &= \cos(\pi u/4\bar{\omega}) \frac{\pi(u - \bar{\omega})/4\bar{\omega}}{\sin[\pi(u - \bar{\omega})/4\bar{\omega}]} \\ &= 0 \qquad \text{for} \quad u \leqq -2\bar{\omega} \\ &= \frac{\pi}{4} 2^{1/2} \qquad \text{for} \quad u = -\bar{\omega} \\ &= \frac{\pi}{4} 2^{1/2} \qquad \text{for} \quad u = 0 \\ &= \frac{1}{2} 2^{1/2} \qquad \text{for} \quad u = \bar{\omega} \\ &= 0 \qquad \text{for} \quad u \geqq 2\bar{\omega}. \end{aligned} \qquad (134)$$

In actual systems, it may be expedient from the standpoint of design to employ transmitting and receiving filter characteristics that differ from the optimum characteristics shown in Fig. 15. This results in some penalty in signal-to-noise ratio, as shown below.

By way of illustration, it will be assumed that the transmitting filter has the shape shown in Fig. 5 for the channel transmission-frequency characteristic, in which case the receiving filter could be flat and given by

$$\begin{aligned} R(u) &= 1 \qquad \text{for} \quad -2\bar{\omega} \leqq u \leqq 2\bar{\omega} \\ &= 0 \qquad \text{for} \quad -2\bar{\omega} > u > 2\bar{\omega}. \end{aligned} \qquad (135)$$

In this case, (122) for AM becomes

$$N_0/\hat{S}_0 = \frac{n}{4\pi TP} \left[\int_{-2\bar{\omega}}^{2\bar{\omega}} du \right] \left[\int_{-2\bar{\omega}}^{2\bar{\omega}} \frac{T^2}{4} \cos^4(\pi u/4\bar{\omega})\, du \right]$$

$$= \frac{3}{2} \frac{N}{4P} \qquad (136)$$

$$= 2 \frac{N}{4P_m}, \qquad (137)$$

where N, P and P_m are defined as in (129) and the last relation follows from (112) for the above case; i.e., $P = \frac{3}{4}P_m$.

Comparison of (136) with (129) shows that, for equal average signal power P, the ratio $N_0/\hat{S}_0$ is greater here than for optimum division of filter shaping by a factor of $\frac{3}{2}$ (or 1.8 db), while comparison of (137) with (129) shows that, for equal continuous mark power P_m, it is greater by a factor of 2 (or 3 db). This illustrates that the penalty incurred in departing from the optimum division of channel shaping between transmitting and receiving filters can be significant.

XV. OPTIMUM FM SYSTEMS

The minimum ratio $N_0/\hat{S}_0$ is obtained when the product of the two integrals in (121) is a minimum. As shown in Appendix C, this is the case when $R(u)$ is such that the two integrals are equal, resulting in the following expression for the optimum $R(u)$ without a postdetection low-pass filter:

$$R^o(u) = cS^{1/2}(u)(1 + u^2/\bar{\omega}^2)^{-1/4}, \tag{138}$$

where c is an arbitrary constant. The corresponding optimum $N_0/\hat{S}_0$ obtained with (138) in (125) is

$$(N_0/\hat{S}_0)^o = \frac{2n}{4\pi TP}\left[\int_{-\infty}^{\infty} (1 + u^2/\bar{\omega}^2)^{1/2}S(u)\,du\right]^2 \tag{139}$$

$$= \frac{N}{4P}\lambda^o, \tag{140}$$

where $P = P_{\mathrm{FM}}$, N is defined as in (129) and

$$\lambda^o = 2\left\{\frac{\left[\int_{-\infty}^{\infty} (1 + u^2/\bar{\omega}^2)^{1/2}S(u)\,du\right]^2}{\left[\int_{-\infty}^{\infty} S(u)\,du\right]^2}\right\}. \tag{141}$$

Comparison of (140) with (129) shows that, for equal average signal power, the optimum ratio $N_0/\hat{S}_0$ is greater in FM than in bipolar AM by the factor λ^o. Inspection of (141) shows that $\lambda^o > 2$, without a postdetection low-pass filter.

With (138) in (126), the factor ρ, by which the noise power is reduced by a postdetection low-pass filter, becomes

$$\rho = \frac{\int_{-\infty}^{\infty} L^2(u)S(u)\,\dfrac{(1 + u/\bar{\omega})^2}{(1 + u^2/\bar{\omega}^2)^{1/2}}\,du}{\int_{-\infty}^{\infty} (1 + u^2/\bar{\omega}^2)^{1/2}S(u)\,du}, \tag{142}$$

where $L(u)$ is defined by (123).

With a postdetection low-pass filter, (140) is replaced by

$$(N_0/\hat{S}_0)^o = \frac{N}{4P}\gamma, \tag{143}$$

where

$$\gamma = \lambda^o \rho \tag{144}$$

is the factor by which the optimum ratio $N_0/\hat{S}_0$ differs in FM from bipolar AM, for equal average signal power at the receiving filter input.

The above factor γ is not the minimum (optimum) factor with a postdetection low-pass filter, but rather the factor applying when a low-pass filter is applied to a system in which $R(u)$ is optimized without such a filter. If a postdetection low-pass filter of specified amplitude characteristic $A_0(\omega)$ is assumed, the optimum $R(u)$ is related to $A_0(\omega)$, as discussed below.

With a postdetection low-pass filter, the minimum ratio $N_0/\hat{S}_0$ in terms of average signal power is again obtained when $R(u)$ is such that the two integrals in (125) are equal. As shown in Appendix C, $R(u)$ is in this case given by

$$R^o(u) = c2^{1/4}S^{1/2}(u)[H(u)]^{-1/2}, \tag{145}$$

where

$$H(u) = [L^2(u)(1 + u/\bar{\omega})^2 + L^2(-u)(1 - u/\bar{\omega})^2]^{1/2} \tag{146}$$

$$= [A_0^2(\bar{\omega} + u)(1 + u/\bar{\omega})^2 + A_0^2(\bar{\omega} - u)(1 - u/\bar{\omega})^2]^{1/2}. \tag{147}$$

In the last relation, $A_0(\bar{\omega} \pm u)$ designates the amplitude characteristic of the low-pass filter at $\omega = \bar{\omega} \pm u$.

The optimum ratio obtained with (145) in (125) can be written

$$(N_0/S_0)^o = \frac{N}{4P}\gamma^o, \tag{148}$$

where

$$\gamma^o = \frac{\left[\int_{-\infty}^{\infty} H(u)S(u)\,du\right]^2}{\left[\int_{-\infty}^{\infty} S(u)\,du\right]^2} \tag{149}$$

$$= \frac{\left[\int_{0}^{\infty} H(u)S(u)\,du\right]^2}{\left[\int_{0}^{\infty} S(u)\,du\right]^2}, \tag{150}$$

where (150) follows from (149) since $H(-u) = H(u)$ and $S(-u) = S(u)$.

For the reasons discussed in Section X, the bandwidth of the low-pass filter must be less than $2\bar{\omega}$. Hence, for $u \approx \bar{\omega}$, $A_0(\bar{\omega} + u) = 0$, so that the first term in (147) vanishes and

$$\begin{aligned} H(u) &\cong A_0(0)(1 - u/\bar{\omega}) \qquad \text{for} \quad u/\bar{\omega} \cong 1 \\ &= 0 \qquad \text{for} \quad u = \bar{\omega}. \end{aligned} \tag{151}$$

Thus, for $u \cong \bar{\omega}$, (145) becomes

$$\begin{aligned} R^o(u) &\cong 2^{1/4} c \left[\frac{S(u)}{A_0(0)(1 - u/\bar{\omega})} \right]^{1/2} \qquad \text{for} \quad u/\bar{\omega} \cong 1 \\ &= \infty \qquad \text{for} \quad u = \bar{\omega}. \end{aligned} \tag{152}$$

The corresponding transmitting filter characteristic is $T^o(u) = A(u)/R^o(u) = 0$ for $u = \bar{\omega}$.

With $R^o(\pm\bar{\omega}) = \infty$, the noise power at the detector input would become infinite and the signal-to-noise ratio at the detector input would be zero. Hence, the basic premise of adequately high signal-to-noise ratios underlying the representation in Fig. 14 and expression (125) would be violated. In this case, $N_0/\hat{S}_0$, without a postdetection low-pass filter as given by (121), would become infinite. To limit the noise power at the detector input, so that (125) is a legitimate approximation, it is necessary to modify $R^o(u)$ near $u = \pm\bar{\omega}$ in such a way that $R^o(\pm\bar{\omega}) \neq \infty$ and $N_0/\hat{S}_0$, as obtained from (121) without a postdetection filter, becomes appropriately small. The value of $N_0/\hat{S}_0$ obtained from (125) after such modification of $R(u)$ will be greater than that obtained from (150), but may be smaller than that given by (144). The factor γ^o is thus to be regarded as a lower bound that cannot be fully realized but may be closely approached, at least for small ratios N/P in (148), by appropriate modification of $R^o(u)$ near $u = \pm\bar{\omega}$. (An example of such modification is indicated by the dotted curves in Fig. 17, to be discussed later.)

XVI. OPTIMUM FM SYSTEMS OF MINIMUM BANDWIDTH

In the limiting case of a channel of the minimum possible bandwidth, as considered in Section VI, the channel characteristic is given by (55). When normalized to unit amplitude for $u = 0$, this characteristic is

$$A(u) = \frac{1 - (u/\bar{\omega})^2}{\cos(\pi u/2\bar{\omega})}. \tag{153}$$

With a spectrum as given by (53) and (54), the optimum receiving filter characteristic obtained from (138) becomes, with c chosen as $(2/T)^{1/2}$,

$$R^o(u) = (1 + u^2/\bar{\omega})^{-1/4}. \tag{154}$$

The corresponding optimum receiving filter characteristic,

$$T^o(u) = A(u)/R^o(u),$$

is

$$T^o(u) = \frac{[1 - (u/\bar{\omega})^2](1 + u^2/\bar{\omega}^2)^{1/4}}{\cos(\pi u/2\bar{\omega})}. \tag{155}$$

The above expressions apply for $-1 < u/\bar{\omega} < 1$.

The factor λ^o defined by (141) becomes

$$\lambda^o = 2\frac{\left[\int_{-\bar{\omega}}^{\bar{\omega}} (1 + u^2/\bar{\omega}^2)^{1/2}\, du\right]^2}{\left[\int_{-\bar{\omega}}^{\bar{\omega}} du\right]^2} \tag{156}$$
$$= \tfrac{1}{2}[2^{1/2} + \log_e(1 + 2^{1/2})]^2$$
$$\cong 2.65.$$

This corresponds to about 4.2 db disadvantage in signal-to-noise ratio for an optimum FM system without a postdetection low-pass filter, as compared to an optimum bipolar AM system, for equal average signal power.

In the above case the spectrum $S_0(\omega)$ of the demodulated pulses is as shown in Fig. 10. With a low-pass filter having the amplitude charactertic shown in Fig. 11, this spectrum is converted into a flat spectrum of the minimum permissible bandwidth. With the above type of filter, $L(u) = A_0(\bar{\omega} + u) = 0$ for $u > 0$. Since $S(u) = 0$ for $\bar{\omega} < u < -\bar{\omega}$, (142) becomes

$$\rho = \frac{\int_{-\bar{\omega}}^{0} A_0^2(\bar{\omega} - u)\, \frac{(1 + u/\bar{\omega})^2}{(1 + u^2/\bar{\omega}^2)^{1/2}}\, du}{\int_{-\bar{\omega}}^{\bar{\omega}} (1 + u^2/\bar{\omega}^2)^{1/2}\, du} \tag{157}$$

$\cong 0.38$ (by numerical integration).

With the above low-pass filter, (144) gives

$$\gamma = \lambda^o \rho \cong 1.0, \tag{158}$$

so that the signal-to-noise ratio is virtually the same as for an optimum bipolar binary AM system, for equal average signal power.

The factor γ^o given by (150) becomes

$$\gamma^o = \frac{\left[\int_0^{\bar{\omega}} A_0(\bar{\omega} - u)(1 - u/\bar{\omega})\, du\right]^2}{\left[\int_0^{\bar{\omega}} du\right]^2} \tag{159}$$

$$\cong 0.60 \text{ (by numerical integration).}$$

For reasons mentioned in Section XV, the minimum factor γ that can be realized will be less than that given by (158) but greater than that given by (159), which is to be regarded as a lower bound. Since this minimum factor γ is less than one, some advantage in signal-to-noise ratio can be realized with FM, as compared to an optimum bipolar baseband or AM system with synchronous detection, for equal average signal power at the input of the receiving filter. This advantage would be small, and is principally of theoretical interest as an indication that a noise advantage can be derived from the unavoidable two-fold increase in channel bandwidth with FM as compared to baseband transmission or equivalent AM methods.

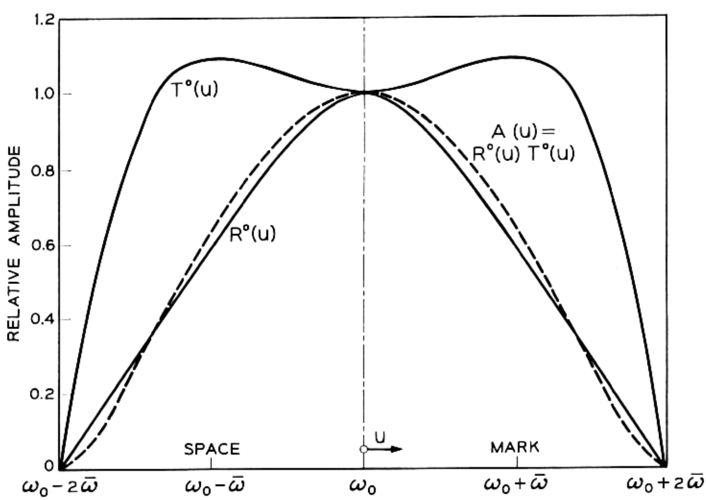

Fig. 16 — Frequency modulation with raised cosine spectrum at detector input and no postdetection low-pass filter: R^o = optimum shape of receiving filter; T^o = optimum shape of transmitting filter; $A = R^oT^o$ = combined transmission characteristic.

XVII. OPTIMUM FM SYSTEMS WITH RAISED COSINE SPECTRUM

With a raised cosine spectrum at the detector input, as considered in Section VII, the over-all amplitude characteristic of the bandpass channel is given by (62). When normalized to unit amplitude for $u = 0$, the characteristic is

$$A(u) = \frac{1 + \cos(\pi u/2\bar{\omega})}{2\cos(\pi u/2\bar{\omega})}(1 - u^2/\bar{\omega}^2). \tag{160}$$

The optimum characteristic of the receiving filter as obtained from (138) is, in this case,

$$R^o = \frac{\cos(\pi u/4\bar{\omega})}{(1 + u^2/\bar{\omega}^2)^{1/4}}. \tag{161}$$

The corresponding optimum characteristic of the transmitting filter is $T^o = A(u)/R^o$, or

$$\begin{aligned} T^o &= \frac{\cos(\pi u/4\bar{\omega})}{\cos(\pi u/2\bar{\omega})}(1 - u^2/\bar{\omega}^2)(1 + u^2/\bar{\omega}^2)^{1/4} \\ &= 1 \qquad\qquad\qquad \text{for} \quad u = 0 \\ &= \frac{4}{\pi}(\tfrac{1}{2})^{1/4} \cong 1.09 \qquad \text{for} \quad u = \bar{\omega} \\ &= 0 \qquad\qquad\qquad \text{for} \quad u = 2\bar{\omega}. \end{aligned} \tag{162}$$

The above filter characteristics are shown in Fig. 16.

With a raised cosine spectrum in (141),

$$\begin{aligned} \lambda^o &= 2\left\{\frac{\left[\int_{-2\bar{\omega}}^{2\bar{\omega}} (1 + u^2/\bar{\omega}^2)^{1/2}\cos^2(\pi u/4\bar{\omega})\,du\right]^2}{\left[\int_{-2\bar{\omega}}^{2\bar{\omega}} \cos^2(\pi u/4\bar{\omega})\,du\right]^2}\right\} \\ &= 2\left[\int_0^2 (1 + x^2)^{1/2}\cos^2(\pi x/4)\,dx\right]^2 \\ &\cong 2.8 \text{ (by numerical integration).} \end{aligned} \tag{163}$$

With the aid of a postdetection low-pass filter, the noise power in the output is reduced by the following factor ρ obtained from (142):

$$\rho = \frac{\int_{-2\bar{\omega}}^{2\bar{\omega}} G(u)\cos^2(\pi u/4\bar{\omega})\,du}{\int_{-2\bar{\omega}}^{2\bar{\omega}} (1 + u^2/\bar{\omega}^2)^{1/2}\cos^2(\pi u/4\bar{\omega})\,du}, \tag{164}$$

where

$$G(u) = L^2(u) \frac{(1 + u/\bar{\omega})^2}{(1 + u^2/\bar{\omega}^2)^{1/2}}. \qquad (165)$$

The factor γ^o given by (150) in this case becomes

$$\gamma^o = \frac{\left[\int_0^{2\bar{\omega}} H(u) \cos^2 (\pi u/4\bar{\omega})\, du\right]^2}{\left[\int_0^{2\bar{\omega}} \cos^2 (\pi u/4\bar{\omega})\, du\right]^2}, \qquad (166)$$

where, as in (147),

$$H(u) = [A_0^2(\bar{\omega} + u)(1 + u/\bar{\omega})^2 + A_0^2(\bar{\omega} - u)(1 - u/\bar{\omega})^2]^{1/2}.$$

The optimum receiving filter characteristic is, in accordance with (145),

$$R^o(u) = c2^{1/4} \cos (\pi u/4\bar{\omega})[H(u)]^{-1/2}. \qquad (167)$$

With a raised cosine spectrum at the detector input, the spectrum of the demodulated pulses is as shown in Fig. 12. For a low-pass filter with a transmission-frequency characteristic as shown in Fig. 13, approximate values of $L(u) = A_0(\bar{\omega} + u)$, $G(u)$ and $H(u)$ are given in Table VI by way of illustration.

TABLE VI — ILLUSTRATIVE VALUES OF L, G AND H

	$-u/\bar{\omega}$							
	−0.75	−0.5	0	0.5	0.75	1.0	1.5	2.0
$L(u)$	0	0.65	1.2	1.15	1.05	1.00	1.15	1.2
$G(u)$	0	0.8	1.44	0.30	0.05	0	0.185	0.51
$H(u)$	0.27	1.1	1.7	1.1	0.27	0	0.58	1.2

For the particular low-pass filter above, the optimum receiving filter characteristic obtained from (167) and the corresponding transmitting filter characteristic $T^o(u) = A(u)/R^o(u)$ are as shown in Fig. 17, where they are normalized to unity for $u = 0$. It will be noted that, for $u = \pm\bar{\omega}$, $R^o = \infty$ and $T^o = 0$. Hence these optimum characteristics cannot be realized physically, though they can be approached, as indicated by the dotted lines near $u = \pm\bar{\omega}$.

For the above low-pass filter, numerical integration of (164) and (166) gives $\rho = 0.5$ and $\gamma^o = 0.84$.

The significance of the above various numerical results are, in summary, as follows.

The factor $\lambda^o = 2.8$ indicates about a 4.5-db disadvantage in signal-to-noise ratio, for an optimum FM system without a postdetection filter as compared to an optimum bipolar AM or baseband system, for equal average signal power at the input of the receiving filter. The factor $\rho = 0.5$ indicates a 3-db improvement in signal-to-noise ratio obtained with the aid of the particular postdetection low-pass filter assumed above, so that the above FM disadvantage is reduced to 1.5 db; i.e., $\gamma = \lambda^o \rho = 1.4$.

In accordance with the discussion in Section XV, a somewhat lower factor γ could be realized when the division of channel shaping between transmitting and receiving bandpass filters is optimized for this particular low-pass filter. The lower bound is represented by $\gamma^o = 0.84$, and the minimum value that could be realized with an appropriate modification in filter characteristics near $u = \bar{\omega}$ would be greater than 0.84 but less than $\gamma = 1.4$. With the modification indicated in Fig. 17, it turns out

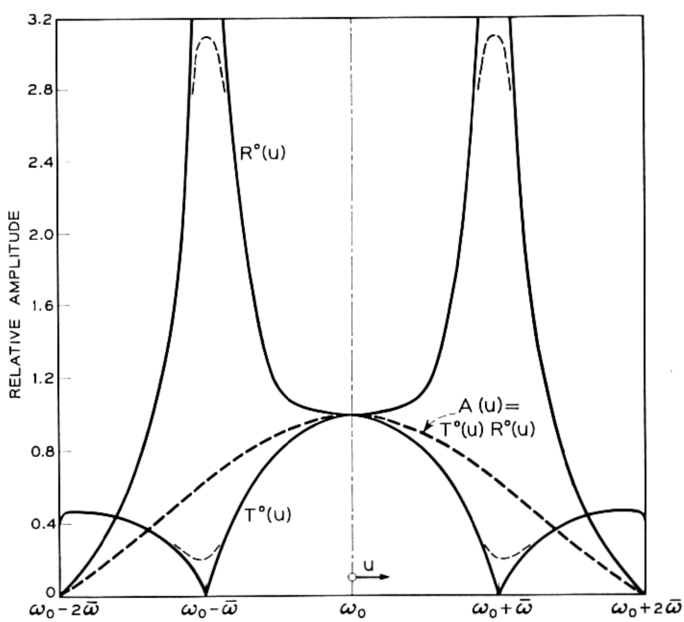

Fig. 17 — Frequency modulation with raised cosine spectrum and post-detection low-pass filter with transmission characteristic $A_0(\omega)$ shown in Fig. 13: R^o = optimum shape of receiving filter; T^o = optimum shape of transmitting filter; $A = R^o T^o$ = combined transmission characteristic.

that the ratio $N_0/\hat{S}_0$ obtained from (125) is smaller than that for bipolar AM by a factor $\gamma \cong 0.94$. Thus, with optimum design in FM, the signal-to-noise ratio would be very nearly the same as for an optimum bipolar AM or baseband system.

If the spectrum $S_0(\omega)$ shown in Fig. 12 is converted into a flat spectrum of bandwidth $2\bar{\omega}$ by an appropriate low-pass filter, it turns out that $\gamma^o \cong 0.64$, corresponding to about a 2-db advantage in signal-to-noise ratio over an optimum bipolar AM or baseband system.

As an illustration of the penalty incurred in departing from the optimum division of channel shaping between transmitting and receiving bandpass filters, it will be assumed that $R(u)$ is given by (135), as previously considered for AM. In this case, the receiving filter is flat between $-2\bar{\omega} < u < 2\bar{\omega}$, and (125) becomes

$$N_0/\hat{S}_0 = \frac{2n}{4\pi TP} \int_{-2\bar{\omega}}^{2\bar{\omega}} L^2(u)(1 + u/\bar{\omega})^2 \, du \cdot \int_{-2\bar{\omega}}^{2\bar{\omega}} \frac{T^2}{4} \cos^4 (\pi u/4\bar{\omega}) \, du, \tag{168}$$

which can be written as

$$N_0/\hat{S}_0 = \frac{N}{4P} \gamma, \tag{169}$$

with

$$\gamma = \left(\frac{1}{\bar{\omega}}\right)^2 \int_{-2\bar{\omega}}^{2\bar{\omega}} L^2(u)(1 + u/\bar{\omega})^2 \, du \cdot \int_0^{2\bar{\omega}} \cos^4 (\pi u/4\bar{\omega}) \, du \tag{170}$$

$$\cong 1.66 \text{ (by numerical integration)},$$

where the numerical result applies for a postdetection low-pass filter with the same amplitude characteristic as assumed previously.

The value $\gamma = 1.66$ corresponds to a 2.2-db disadvantage in signal-to-noise ratio as compared to an optimum bipolar AM system, and about a 0.7-db disadvantage compared to an optimum FM system ($\gamma = 1.4$). With the above type of flat receiving filter, about a 1.8-db penalty in signal-to-noise ratio was incurred in AM, as shown in Section XIV.

XVIII. SIGNAL-TO-NOISE RATIOS AND ERROR PROBABILITIES

In the case of baseband transmission or AM with homodyne detection, the probability of exceeding the rms noise amplitude by a specified factor follows the normal law.

If the polarity of the noise voltage is specified, (i.e., positive or negative) the probability of exceeding the rms noise amplitude by a factor k is

$$p = \tfrac{1}{2} \operatorname{erfc} (k/2^{1/2}), \tag{171}$$

where erfc is the error function complement.

If the noise amplitude at a sampling instant t is $[N_0(t)]^{1/2}$, an error will occur if the ratio $[N_0(t)/\hat{S}_0]^{1/2}$ exceeds $\frac{1}{2}$ in the presence of a space or negative pulse, or exceed $-\frac{1}{2}$ in the presence of a mark or positive pulse. The rms ratio, $(N_0/\hat{S}_0)^{1/2}$, must be held smaller by an appropriate factor $1/k$. The probability of an error in a digit is accordingly $p = p_e$, as given by (171), provided

$$(N_0/\hat{S}_0)^{1/2} = \frac{1}{2k}. \tag{172}$$

In accordance with (129),

$$(P/N)^o = \tfrac{1}{4}(\hat{S}_0/N_0). \tag{173}$$

From (172) and (173),

$$k = (P/N)^{1/2}. \tag{174}$$

Hence, for an optimum AM system as assumed above, the error probability in binary bipolar pulse transmission is, with (174) in (171),

$$p_e = \tfrac{1}{2} \operatorname{erfc} (P/2N)^{1/2}. \tag{175}$$

In Table VII the signal-to-noise ratios P/N obtained from (175) are shown for various probabilities of an error in the digit. In accordance with (129), under the optimum condition the average signal power P in bipolar AM is equal to the continuous mark power P_m.

Ideal synchronous detection, as assumed above, cannot be fully realized with symmetrical bipolar AM methods. Derivation of a demodulating carrier from the signal wave, or adequate phase control of a locally supplied carrier, entails some increase in either signal power, noise power or bandwidth, depending on the particular method used, and thus entails a somewhat greater ratio P/N than that given in Table VII.

In the case of unipolar or "on-off" baseband or double-sideband AM with homodyne detection, the maximum tolerable peak noise power is 6 db less than it is with bipolar transmission, but the average signal power is reduced by 3 db, so that the ratio P/N must be increased by 3 db, as indicated for unipolar AM in Table VII. The average signal power is, in this case, 3 db less than the signal power during a continuous mark.

TABLE VII — OPTIMUM SIGNAL-TO-NOISE RATIOS P/N IN DECIBELS FOR BINARY AM AND FM SYSTEMS[A]

Probability of an Error in a Digit	AM with Ideal Synchronous Detection[B]		FM with Ideal Frequency Discriminator Detection		
	Bipolar (Two-Phase)	Unipolar (On-Off)[C,D]	No Low-Pass Filter[E]	With Low-Pass Filter[E,F]	With Low-Pass Filter[F,G]
10^{-2}	7.3	10.3 (11.3)	11.8	8.8	6.5
10^{-4}	11.4	14.4	15.9	12.9	10.6
10^{-6}	13.6	16.6 (16.9)	18.1	15.1	12.8
10^{-8}	15.0	18.0	19.5	16.5	14.2
10^{-10}	16.0	19.0	20.5	17.5	15.2
10^{-12}	17.0	20.0	21.5	18.5	16.2

P = Average signal power at input of receiving filter
N = Average noise power in flat band $W = 1/2T$ cps at input of receiving filter
T = Interval between pulses, in seconds

Notes:

[A] Signal-to-noise ratios in terms of noise power $2N$ in double-sideband channel of bandwidth $2W$ are 3 db smaller than in table.
[B] Applies for baseband transmission, double-sideband, double-sideband on two carriers at quadrature and vestigial-sideband AM.
[C] Signal-to-noise ratios in terms of steady mark power are 3 db greater than in table.
[D] Values in brackets are for double-sideband with optimum envelope detection.
[E] Band-pass filter shaping as in Fig. 16.
[F] Low-pass filter shaping as in Fig. 13.
[G] Band-pass filter shaping as in Fig. 17.

For the reasons discussed in Section XIV, the optimum signal-to-noise ratios shown in Table VII for bipolar AM apply for bipolar baseband transmission, for bipolar double-sideband AM (phase reversal or two-phase transmission), for bipolar quadrature double-sideband AM (four-phase transmission) and for bipolar vestigial-sideband AM. Similarly, the ratios shown for unipolar AM with synchronous detection apply for unipolar baseband, double-sideband, quadrature double-sideband and vestigial-sideband transmission. Furthermore, with optimum division of channel shaping between transmitting and receiving filters, the above optimum signal-to-noise ratios apply for all spectra of the pulses at the detector input with the properties illustrated in Figs. 1 and 2. Moreover, the signal-to-noise ratios apply not only for pulses of duration T equal to the pulse interval, as assumed in the previous analysis, but also for pulses of shorter duration. When the duration of the pulses is less than T, the receiving filter characteristic remains unchanged, but the shape of the transmitting filter is modified because of the different spectrum of the modulating pulses. This also applies with other than rectangular shapes of the modulating pulses.

There are thus an infinite number of optimum AM systems with a performance from the standpoint of signal-to-noise for a given error probability equivalent to that of a baseband system of the minimum possible bandwidth, provided ideal homodyne (synchronous) detection is used, in which the pulse train is applied to a product demodulator together with a constant demodulating carrier of proper phase.

The ratios shown in Table VII for unipolar AM also apply in a first approximation to "on-off" double-sideband AM with envelope rather than homodyne detection. However, in the latter case both the average noise power and the probability distribution at the detector output differ between mark and space.[4] During a mark, the rms noise amplitude, and the probability that this noise amplitude is exceeded by a factor k, is virtually the same as it is with homodyne detection, for large signal-to-noise ratios. During a space, however, the rms noise amplitude is increased by a factor of $2^{1/2}$, and the probability that the rms amplitude is exceeded by a specified factor k follows the Rayleigh law

$$p = e^{-k^2} \tag{176}$$

rather than the Gaussian law (171).

For the above reason, the optimum slicing or threshold level is not one-half the peak pulse amplitude, but slightly greater, depending on signal-to-noise ratio and error probability. The optimum slicing levels with binary double-sideband AM and envelope detection and the corresponding optimum signal-to-noise ratios versus error probability have been determined elsewhere,[4] and are indicated in Table VII for two cases. For an error probability of 10^{-6} the optimum threshold level is about 52 per cent of the peak pulse amplitude and the signal-to-noise ratio is about 0.3 db greater than for unipolar AM with homodyne detection. Hence, for error probabilities in the range ordinarily considered acceptable, the difference in signal-to-noise ratio with envelope and homodyne detection is insignificant.

Comparison of binary FM and AM on the basis of signal-to-noise ratios is legitimate provided that, for a given ratio $N_0/\hat{S}_0$, the error probability is the same in FM and AM. For high signal-to-noise ratios, this is approximately the case, since the normal law (171) is then closely approximated in FM.[9] On this premise, comparison on the basis of signal-to-noise ratios is legitimate for small error probabilities.

In accordance with the discussion in Section XV, the optimum signal-to-noise ratio in binary FM without a postdetection filter is related to the optimum ratio in bipolar binary AM by

$$(P/N)^o_{\text{FM}} = \lambda^o (P/N)^o_{\text{AM}}, \tag{177}$$

where $\lambda^o \cong 2.65$ (4.2 db) for a flat spectrum as considered in Section XVI and $\lambda^o \cong 2.8$ (4.5 db) for a raised cosine spectrum at the detector input, as considered in Section XVII. In Table VII the latter case is assumed as being the more representative, and the signal-to-noise ratios are taken 4.5 db greater than they are in bipolar AM.

With the aid of an appropriate postdetection low-pass filter, the signal-to-noise ratio can be improved such that

$$(P/N)_{\text{FM}}^o = \rho\lambda^o(P/N)_{\text{AM}}^o. \tag{178}$$

With a postdetection filter having an amplitude characteristic as shown in Fig. 13, $\rho \cong 0.5$, corresponding to 3-db improvement in signal-to-noise ratio, and this case is assumed in Table VII.

As discussed in Sections XV, XVI and XVII, when a postdetection low-pass filter is used, it is possible with FM to realize some improvement in signal-to-noise ratio over bipolar AM. In this case, a lower bound is given by

$$(P/N)_{\text{FM}}^o = \gamma^o(P/N)_{\text{AM}}^o, \tag{179}$$

where $\gamma^o \cong 0.84$ with the type of filter shown in Fig. 13. This corresponds to about an 0.8-db improvement in signal-to-noise ratio over bipolar AM and entails transmitting and receiving bandpass filter characteristics as indicated in Fig. 17. This lower bound is given in the last column of Table VII, but it cannot be fully realized, for reasons discussed in Section XV.

As noted before, the optimum signal-to-noise ratios given in Table VII for synchronous AM are universal and apply to a variety of optimized systems, including the special case of ideal flat channels of minimum bandwidth assumed in other analyses of baseband,[3] synchronous AM or PM systems.[5,6] With appropriate allowance for different definitions of signal power (continuous mark versus average power) and of the bandwidth used in specifying noise power (flat single-sideband versus flat double-sideband), the results given in Table VII for synchronous AM conform with those in the above references.

In the case of FM, however, the optimum signal-to-noise ratio depends on several factors that need not be considered in AM, such as the shape of the spectrum at the detector input and the shape of the post-detection low-pass filter. There is thus no universal optimum signal-to-noise ratio for a given error probability in FM, and the ratios given in Table VII for FM apply for the particular conditions indicated. For this reason, significant comparisons cannot be made with signal-to-noise ratios for FM given elsewhere[5,7] that are based on simplified mathematical models that ignore the various factors above. One analysis[5] indicates about a 1-db advantage of bipolar AM (or phase reversal) over FM for an error probability of 10^{-4}.

XIX. SUMMARY

It has been shown that binary pulses can be transmitted without intersymbol interference by FM over a channel of the same bandwidth as is required for double-sideband AM. To this end, a first requirement is that the total frequency shift between space and mark be equal to the pulse transmission rate, for example 500 cps for 500 bits per seconds. A second basic requirement is that the pulses at the input of the frequency modulator have the appropriate shape, a condition that is met with rectangular modulating pulses of duration equal to the interval between pulses. A third requirement with rectangular modulating pulses is that the channel bandwidth be small in relation to the midband frequency. A fourth requirement is that the bandpass channel must have the appropriate amplitude-versus-frequency characteristic and a linear phase characteristic.

The appropriate amplitude characteristic of the bandpass channel is not the same as for AM, nor is the shape of the received pulses the same. By way of illustration, a comparison is made in Fig. 3 of the amplitude characteristic of the bandpass channels for FM and AM, for a channel of the minimum possible bandwidth, if intersymbol interference is to be avoided. With the channel characteristics shown in Fig. 3, the transmission of a single pulse, (i.e., transmission of a mark, preceded and followed by a continuing space) will give rise to a flat frequency spectrum at the detector input with both FM and AM, as indicated in Fig. 3. Such a flat spectrum at the detector input will give rise to a pulse at the detector output, but the pulse shape is not the same in FM and AM, as illustrated in Fig. 4. However, a common property of the pulse shapes shown in Fig. 4 is that they have zero points at intervals T equal to the duration of the rectangular modulating pulses. Thus, pulses can be transmitted at these intervals without intersymobl interference by FM or AM.

In actual pulse systems, channels of the minimum possible bandwidth are not practicable for various reasons. In Fig. 5 comparison is made of the appropriate amplitude characteristics with FM and AM for channels with twice the minimum bandwidth. In this case, transmission of a single pulse by FM or AM gives rise to a "raised cosine" spectrum at the detector input and to pulse shapes at the detector output, as shown in Fig. 6. Because of the small oscillations in the tails of the received pulses, the channel characteristics shown in Fig. 6 are desirable for FM and AM systems.

In the above illustrations, rectangular modulating pulses were assumed. With modulating pulses of other shapes that overlap between pulse intervals, certain restrictions are imposed on the shape of the pulses if intersymbol interference is to be avoided, which renders determination of the exact appropriate shapes difficult. For example, $(\sin x)/x$ modulating pulses, which are often considered in AM, are inappropriate in FM. However, with raised cosine modulating pulses, as shown in Fig. 9, it is possible by appropriate channel shaping to virtually avoid intersymbol interference. With such modulating pulses and a channel characteristic as shown in Fig. 9, the shape of the demodulated pulses will be virtually the same as those shown in Fig. 6 for FM. Although intersymbol interference cannot be avoided with raised cosine modulating pulses, it is small enough to be disregarded (less than 1 per cent).

The pulses at the output of the FM detector, such as those shown in Figs. 4 and 6, have baseband spectra of infinite bandwidth, as in Figs. 10 and 12. It is possible to modify the shape of the pulses in such a way that intersymbol interference is not introduced, with the aid of a postdetection low-pass filter having the appropriate amplitude characteristic. For example, the FM pulse of Fig. 4 has a spectrum as shown in Fig. 10. This spectrum can be converted into a flat spectrum of minimum bandwidth with the aid of a postdetection low-pass filter having the amplitude characteristic shown in Fig. 11 and a linear phase characteristic. The pulse shown in Fig. 4 for FM would thereby be converted into the same shape as shown for AM.

The ideal amplitude characteristics of the bandpass channels in FM, such as those exemplified in Figs. 3 and 5, are obtained with the aid of an appropriate combination of transmitting and receiving bandpass filters. From the standpoint of intersymbol interference, as considered above, the division of channel shaping between these filters is immaterial, but there is an optimum division from the standpoint of performance in the presence of noise.

By way of example, for the over-all amplitude characteristics of the bandpass channels shown in Fig. 5 for AM, the optimum division of channel shaping between transmitting and receiving filters is shown in Fig. 15 for the case of random noise. As discussed in Section XVIII, with optimum division of channel shaping and ideal synchronous (homodyne) detection, there is an infinity of optimum AM systems with performance equivalent to that of a baseband system of the minimum possible bandwidth, as regards signal-to-noise ratio for a given pulse transmission rate and error probability.

With an over-all channel characteristic as shown in Fig. 5 for FM, the optimum division of channel shaping is as shown in Fig. 16, assuming no postdetection low-pass filter. Such an optimum FM system has about a 4.5-db disadvantage in signal-to-noise ratio compared to an optimum bipolar AM or phase reversal system, for equal average signal power. By providing a postdetection low-pass filter with a transmission characteristic as shown in Fig. 13 and linear phase, the signal-to-noise ratio is improved about 3 db, and the FM disadvantage of 4.5 db is reduced to about 1.5 db. This assumes the same division of bandpass channel shaping as without a low-pass filter. However, the optimum division with the above low-pass filter is different, and is approximately as indicated in Fig. 17, except near $\omega_0 \pm \bar{\omega}$. With an optimum division, it appears possible, in principle, to realize an advantage in signal-to-noise ratio of at most 0.8 db over an optimum bipolar AM or baseband system, for equal average signal power.

With appropriate postdetection low-pass filters of the minimum permissible bandwidth, it is possible in principle to realize an advantage in signal-to-noise ratio of at most 2 db over an optimum bipolar AM or baseband system. However, the above FM advantages cannot be fully attained in practice. They are principally of theoretical interest in that they indicate that an advantage in signal-to-noise ratio can be derived from the unavoidable two-fold increase in channel bandwidth with FM as compared to baseband transmission, or equivalent AM methods.

XX. ACKNOWLEDGMENTS

The writer is indebted to H. O. Pollak for a critical review that resulted in a radical modification of Sections VIII and X, to Miss Marion C. Gray for supervision of computer programming carried out by Miss D. G. Becker to determine the pulse spectra given in Section X, to S. O. Rice for pointing out the Anger function relations used in Section IX and for helpful suggestions at other points in the analysis, and to G. Wascheck for many editorial comments.

APPENDIX A

Pulse Train Envelopes and Average Signal Power

In Section IV the transmission of a single mark or pulse was considered, and the resultant wave at the channel output (detector input) was, in this case, given by (35) and (36), or

$$E(t_0) = \cos(\omega_0 t_0 + \varphi_0 + \psi_0)\bar{E}(t_0), \tag{180}$$

where

$$\bar{E}(t_0) = A(-\bar{\omega})[1 + \mu^2 p^2(t_0) - 2\mu p(t_0) \cos \bar{\omega} t_0]^{1/2} \tag{181}$$

and

$$\mu = 1/A(-\bar{\omega}) = 1/A(\bar{\omega}). \tag{182}$$

When a train of pulses is transmitted, the resultant envelope is given by (95) or

$$\begin{aligned} \mathbf{W}(t_0) = A(-\bar{\omega}) \Big[1 &+ \mu^2 \sum_{m=-\infty}^{\infty} a_m^2 p^2(t_0 - mT) \\ &- 2\mu \sum_{m=-\infty}^{\infty} a_m p(t_0 - mT) \cos \bar{\omega}(t_0 - mT) \Big]^{1/2}, \end{aligned} \tag{183}$$

where T is the interval between pulses and

$$\begin{aligned} a_m &= 0 \qquad \text{for space} \\ &= 1 \qquad \text{for mark.} \end{aligned} \tag{184}$$

In view of (184), the rms value of the envelope at a particular time, t_0, with respect to a sampling point is for equal probability of marks and spaces:

$$\begin{aligned} \bar{W}(t_0) = A(-\bar{\omega}) \Big[1 &+ \frac{\mu^2}{2} \sum p^2(t_0 - mT) \\ &- \mu \sum p(t_0 - mT) \cos \bar{\omega}(t_0 - mT) \Big]^{1/2}, \end{aligned} \tag{185}$$

where the limits of the summations are as in (183).

The mean squared value of the envelope taken over a pulse interval T is

$$\begin{aligned} \bar{W}^2 &= \frac{1}{T} \int_{-T/2}^{T/2} \bar{W}^2(t_0)\, dt_0 \\ &= A^2(-\bar{\omega}) \Big[1 + \frac{\mu^2}{2} \int_{-T/2}^{T/2} \sum p^2(t_0 - mT)\, dt_0 \\ &\qquad - \mu \int_{-T/2}^{T/2} \sum p(t_0 - mT) \cos \bar{\omega}(t_0 - mT)\, dt_0 \Big], \end{aligned} \tag{186}$$

which can be transformed into

$$\bar{W}^2 = A^2(-\bar{\omega}) \left[1 + \frac{\mu^2}{2T} \int_{-\infty}^{\infty} p^2(t_0)\, dt - \frac{\mu}{T} \int_{-\infty}^{\infty} p(t_0) \cos \bar{\omega} t_0\, dt_0 \right]. \tag{187}$$

In (32) for $p(t_0)$, the lower limit of integration, $-\omega_0$, can be replaced by $-\infty$, since $S(u) = 0$ for $u \leqq -\omega_0$, in which case (32) can be written

$$p(t_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} 2S(u) \cos ut_0 \, du \tag{188}$$

or, by inversion,

$$2S(u) = \int_{-\infty}^{\infty} p(t_0) \cos ut_0 \, dt_0 . \tag{189}$$

In view of (189), the last-bracket term in (187) with $u = \bar{\omega}$ becomes $-2\mu S(\bar{\omega})/T$. Since $S(\bar{\omega}) = T/4$, in accordance with (48), and $\mu = 2$, in accordance with (50), the last-bracket term becomes -1, and (187) simplifies to

$$\bar{W}^2 = A^2(-\bar{\omega}) \frac{\mu^2}{2T} \int_{-\infty}^{\infty} p^2(t_0) \, dt_0 . \tag{190}$$

In view of (189),

$$\int_{-\infty}^{\infty} p^2(t_0) \, dt_0 = \frac{1}{2\pi} \int_{-\infty}^{\infty} [2S(u)]^2 \, du. \tag{191}$$

Hence (190) can be written

$$\bar{W}^2 = A^2(-\bar{\omega}) \frac{\mu^2}{\pi T} \int_{-\infty}^{\infty} S^2(u) \, du. \tag{192}$$

This is the mean squared value of the envelope at the output of a channel with an over-all amplitude characteristic $A(u)$. If the transmitting filter is assumed to have an amplitude characteristic $T(u)$ and the receiving filter to have a characteristic $R(u)$, then

$$T(u)R(u) = A(u). \tag{193}$$

The spectrum at the transmitter output, or at the input of the receiving filter in a channel without transmission loss, is in this case $T(u)S(u) = S(u)/R(u)$. The mean squared value of the envelope at the transmitter output is obtained by replacing $A(\bar{\omega})$ by $T(\bar{\omega})$ and $S(u)$ by $S(u)/R(u)$ in (192), so that in this case

$$\bar{W}_0^2 = T^2(\bar{\omega}) \frac{1}{\pi T} \left[\frac{R(\bar{\omega})}{A(\bar{\omega})}\right]^2 \int_{-\infty}^{\infty} \left[\frac{S(u)}{R(u)}\right]^2 du, \tag{194}$$

where $A(\bar{\omega}) = \frac{1}{2}$.

When a continuous mark with unit amplitude of the carrier at the input of the transmitting filter is transmitted, the envelope of the carrier at the output is $T(\bar{\omega})$. It will be assumed that the carrier amplitude at the input is so chosen that the peak amplitude at the output is E for a continuous mark or space. With $T(\bar{\omega})$ replaced by E in (194), and with $A(\bar{\omega}) = \frac{1}{2}$,

$$\bar{W}^2 = E^2 \frac{4R^2(\bar{\omega})}{\pi T} \int_{-\infty}^{\infty} \frac{S^2(u)}{R^2(u)}\, du. \tag{195}$$

The average signal power within the envelope is, in view of the factor $\cos(\omega_0 t_0 + \varphi_0 + \psi_0)$ in (180), smaller by a factor of $\frac{1}{2}$, or

$$P_{\mathrm{FM}} = \frac{E^2}{2} \frac{4R^2(\bar{\omega})}{\pi T} \int_{-\infty}^{\infty} \frac{S^2(u)}{R^2(u)}\, du. \tag{196}$$

In bipolar AM (183) is replaced by

$$\bar{W}(t_0) = A(0) \sum_{m=-\infty}^{\infty} a_m p(t_0 - mT), \tag{197}$$

where $a_m = -1$ or 1.

Equation (190) is replaced by

$$\begin{aligned} \bar{W}^2 &= A(0)^2 \frac{1}{T} \int_{-\infty}^{\infty} p^2(t_0)\, dt_0 \\ &= A^2(0) \frac{2}{\pi T} \int_{-\infty}^{\infty} S^2(u)\, du, \end{aligned} \tag{198}$$

and (194) is replaced by

$$\bar{W}_0^2 = T^2(0) \frac{2}{\pi T} \left[\frac{R(0)}{A(0)}\right]^2 \int_{-\infty}^{\infty} \frac{S^2(u)}{R^2(u)}\, du. \tag{199}$$

When the peak amplitude of the transmitted carrier for a continuous mark is E, the average signal power in this case is, with $A(0) = 1$,

$$P_{\mathrm{AM}} = \frac{E^2}{2} \frac{2}{\pi T} R^2(0) \int_{-\infty}^{\infty} \frac{S^2(u)}{R^2(u)}\, du. \tag{200}$$

APPENDIX B

Interference in AM and FM Systems

B.1. *General*

B.1.1 *Frequency Modulation*

Let a carrier of frequency $\omega_0 - \bar{\omega}$ representing a space be transmitted, and let its peak amplitude at the input of the receiving filter be E. If the receiving filter has an amplitude characteristic $R(u)$ as a function of the frequency u from midband, the signal at the detector input is

$$e_s(t_0) = -ER(-\bar{\omega})[\cos(\omega_0 - \bar{\omega})t_0 + \varphi_0]. \tag{201}$$

An interfering voltage at the detector input can be written in the general form

$$e_i(t_0) = r_i(t_0)\cos(\omega_0 t_0 + \varphi_0) + q_i(t_0)\sin(\omega_0 t_0 + \varphi_0). \tag{202}$$

The signal at the detector input in the presence of interference is then

$$\begin{aligned} e_s + e_i = {} & -\cos(\omega_0 t_0 + \varphi_0)[ER(-\bar{\omega})\cos\bar{\omega}t_0 - r_i(t_0)] \\ & -\sin(\omega_0 t_0 + \varphi_0)[ER(-\bar{\omega})\sin\bar{\omega}t_0 - q_i(t_0)]. \end{aligned} \tag{203}$$

The phase at the detector input in the presence of interference is given by

$$\tan\psi_{0,i} = -\frac{\sin\bar{\omega}t_0 - \mu_i q_i}{\cos\bar{\omega}t_0 - \mu_i r_i}, \tag{204}$$

where

$$\mu_i = \frac{1}{ER(-\bar{\omega})}. \tag{205}$$

The demodulated signal in presence of interference is proportional to $d\psi_{0,i}/dt_0$, which becomes

$$\begin{aligned} \psi_{0,i}' = -\frac{\bar{\omega}}{D}\Big[& 1 - \mu_i(r_i\cos\bar{\omega}t_0 + q_i\sin\bar{\omega}t_0) \\ & + \frac{\mu_i}{\bar{\omega}}(r_i'\sin\bar{\omega}t_0 - q_i'\cos\omega t_0) + \frac{\mu_i^2}{\bar{\omega}}(q_i'r_i - r_i'q_i)\Big], \end{aligned} \tag{206}$$

where $r_i' = dr_i/dt_0$, $q_i' = dq_i/dt_0$ and

$$\begin{aligned} D &= (\sin\bar{\omega}t_0 - \mu_i q_i)^2 + (\cos\bar{\omega}t_0 - \mu_i r_i)^2 \\ &= 1 + \mu_i^2(r_i^2 + q_i^2) - 2\mu_i(r_i\cos\bar{\omega}t_0 + q_i\sin\bar{\omega}t_0). \end{aligned} \tag{207}$$

The frequency deviation with respect to the frequency $\omega_0 - \bar{\omega}$ is $\psi_i' = \bar{\omega} + \psi_{0,i}'$, and the ratio $\eta_i = \psi_i'/2\bar{\omega}$ becomes

$$\begin{aligned} \eta_i(t_0) = \frac{\mu_i}{2D}\Big[& \mu_i(r_i^2 + q_i^2) - r_i\cos\bar{\omega}t_0 - q_i\sin\bar{\omega}t_0 \\ & - \frac{1}{\bar{\omega}}(r_i'\sin\bar{\omega}t_0 - q_i'\cos\bar{\omega}t_0) - \frac{\mu_i}{\bar{\omega}}(q_i'r_i - r_i'q_i)\Big]. \end{aligned} \tag{208}$$

This is the amplitude of the interfering voltage taken in relation to the peak amplitude of a demodulated pulse.

At sampling points $t_0 = mT$, $\sin \bar{\omega} t_0 = 0$ and $\cos \bar{\omega} t_0 = (-1)^m$. At these points

$$\eta_i(mT) = \frac{\mu_i}{2D}\left[\mu_i(r_i^2 + q_i^2) - (-1)^m\left(r_i - \frac{1}{\bar{\omega}} q_i'\right) - \frac{\mu_i}{\bar{\omega}}(q_i' r_i - r_i' q_i)\right], \tag{209}$$

$$D(mT) = -(-1)^m 2\mu_i r_i + \mu_i^2(r_i^2 + q_i^2). \tag{210}$$

If second-order terms are neglected, which is permissible for adequately low amplitudes of the interfering voltage, the above expression reduces to

$$\eta_i(mT) \cong -\frac{1}{2ER(-\bar{\omega})} \frac{r_i(mT) - q_i'(mT)/\bar{\omega}}{1 - (-1)^m r_i(mT)/ER(-\bar{\omega})} \tag{211}$$

$$\cong -\frac{1}{2ER(-\bar{\omega})}[r_i(mT) - q_i'(mT)/\bar{\omega}]. \tag{212}$$

B.1.2 *Bipolar AM or Two-Phase Modulation*

In bipolar AM with homodyne detection, let a carrier $-E \cos(\omega_0 t_0 + \varphi_0)$ represent a space and a carrier $E \cos(\omega_0 t_0 + \varphi_0)$ a mark. The signal plus interference at the detector input during a space in this case is

$$e_s + e_i = -\cos(\omega_0 t_0 + \varphi_0)[ER(0) - r_i] + \sin(\omega_0 t_0 + \varphi_0) q_i \tag{213}$$

The demodulated output after elimination of high-frequency demodulation products by low-pass filtering is

$$V_{s+i} = -\tfrac{1}{2}[ER(0) - r_i(t_0)]. \tag{214}$$

If a space is represented by 0 at the output, rather than by $-ER(0)/2$, the demodulated output is $V_{s+i} + ER(0)/2$, or $r_i/2$. The resultant interference taken in relation to the amplitude $ER(0)$ of a mark is

$$\eta_i = \frac{r_i(t_0)}{2ER(0)}. \tag{215}$$

B.1.3 *Unipolar AM with Envelope Detection*

In unipolar or "on-off" pulse transmission with envelope detection, the demodulated interference voltage in the presence of a space (zero carrier) is

$$V_i^{(s)} = (r_i^2 + q_i^2)^{1/2}. \tag{216}$$

The amplitude of the interfering voltage taken in relation to the amplitude of a demodulated mark is

$$\eta_i^{(s)} = \frac{(r_i^2 + q_i^2)^{1/2}}{ER(0)}. \tag{217}$$

When a mark is transmitted, the voltage at the detector input is

$$e_{s,i}^{(m)} = \cos(\omega_0 t_0 + \varphi_0)[ER(0) + r_i] + \sin(\omega_0 t_0 + \varphi_0) q_i \tag{218}$$

and at the detector output is

$$\begin{aligned} V_{s,i}^{(m)} &= \{[ER(0) + r_i]^2 + q_i^2\}^{1/2} \\ &\cong ER(0) + r_i . \end{aligned} \tag{219}$$

The amplitude of the interfering voltage taken in relation to the amplitude of a demodulated mark in this case is

$$\eta_i^{(m)} = \frac{r_i}{ER(0)}. \tag{220}$$

B.2 *Single Frequency Interference*

In the particular case of a sinusoidal interfering voltage of frequency $\omega_0 + u$ and amplitude $e(u)$ at the input of the receiving filter, the interfering voltage at the detector input is

$$e_i(t_0) = e(u)R(u) \cos[(\omega_0 + u)t_0 + \varphi_1] \tag{221}$$

$$\begin{aligned} &= e(u)R(u) \cos(\omega_0 t_0 + \varphi_0) \cos(ut_0 + \varphi) \\ &\quad - e(u)R(u) \sin(\omega_0 t_0 + \varphi_0) \sin(ut_0 + \varphi), \end{aligned} \tag{222}$$

where $\varphi = \varphi_1 - \varphi_0$.

In this case,

$$\begin{aligned} r_i(t_0) &= e(u)R(u) \cos(ut_0 + \varphi), \\ q_i(t_0) &= -e(u)R(u) \sin(ut_0 + \varphi). \end{aligned} \tag{223}$$

B.2.1 *Frequency Modulation*

In the case of FM, (212) becomes

$$\eta_i \cong -\frac{e(u)R(u)}{2ER(-\bar{\omega})}\left(1 + \frac{u}{\bar{\omega}}\right) \cos(ut_0 + \varphi). \tag{224}$$

The rms interference in FM with all phases, φ, equally probable is

$$\begin{aligned}\bar{\eta}_i &\cong \frac{e(u)R(u)}{2ER(-\bar{\omega})}\left(1+\frac{u}{\bar{\omega}}\right)\left[\frac{1}{\pi}\int_0^\pi \cos^2(ut_0+\varphi)\,d\varphi\right]^{1/2} \\ &= \frac{\bar{e}(u)R(u)}{2ER(-\bar{\omega})}\left(1+\frac{u}{\bar{\omega}}\right),\end{aligned} \tag{225}$$

where

$$\bar{e}(u) = \frac{e(u)}{2^{1/2}}. \tag{226}$$

B.2.2 *Bipolar AM*

For bipolar AM, (215) gives

$$\eta_i = \frac{e(u)R(u)\cos(ut_0+\varphi)}{2ER(0)}, \tag{227}$$

and the rms value with all phases φ equally probable is

$$\bar{\eta}_i = \frac{\bar{e}(u)R(u)}{2ER(0)}. \tag{228}$$

B.2.3 *Unipolar AM with Envelope Detection*

For unipolar AM with envelope detection, (223) in (217) yields, for the interference during a space,

$$\eta_i^{(s)} = \frac{e(u)R(u)}{ER(0)} = 2^{1/2}\frac{\bar{e}(u)}{ER(0)} \tag{229}$$

and (220) gives, for interference during a mark,

$$\eta_i^{(m)} = \frac{e(u)R(u)}{ER(0)}\cos(ut_0+\varphi), \tag{230}$$

with an rms value

$$\bar{\eta}_i^{(m)} = \frac{\bar{e}(u)R(u)}{ER(0)}. \tag{231}$$

B.3 *Impulse Noise*

In the case of idealized impulse noise, the interfering voltage is of the general form

$$\begin{aligned}e_i(t_0) &= p_i(t_0)\cos(\omega_0 t_0+\varphi_i) \\ &= p_i(t_0)\cos\varphi\cos(\omega_0 t_0+\varphi_0) - p_i(t_0)\sin\varphi\sin(\omega_0 t_0+\varphi_0),\end{aligned} \tag{232}$$

where $\varphi = \varphi_i - \varphi_0$ and $p_i(t_0)$ is the envelope of the noise pulse at the detector input. The shape of the latter will depend on the receiving filter characteristic, $R(u)$. The phase φ represents the difference between the phase of the carrier within the noise pulse envelope and the phase of the carrier within the signal pulse envelope. This phase difference depends on the instant at which the impulse occurs and can have any value.

B.3.1 *Frequency Modulation*

Comparison of (232) with (202) shows that, in this case,

$$\begin{aligned} r_i(t_0) &= p_i(t_0) \cos \varphi, \\ q_i(t_0) &= -p_i(t_0) \sin \varphi \end{aligned} \tag{233}$$

In (208) the various quantities become:

$$r_i^2 + q_i^2 = p_i^2(t_0), \tag{234}$$

$$r_i \cos \bar{\omega} t_0 + q_i \sin \bar{\omega} t_0 = p_i(t_0) \cos (\bar{\omega} t_0 + \varphi), \tag{235}$$

$$r_i' \sin \bar{\omega} t_0 - q_i' \cos \bar{\omega} t_0 = p_i'(t_0) \sin (\bar{\omega} t_0 + \varphi), \tag{236}$$

$$q_i' r_i - r_i' q_i = 0. \tag{237}$$

With these values in (208), the amplitude of a noise pulse after demodulation becomes

$$\eta_i(t_0) = \frac{\mu_i}{2D} \left[\mu_i p_i^2(t_0) - p_i(t_0) \cos (\bar{\omega} t_0 + \varphi) - \frac{p_i'(t_0)}{\bar{\omega}} \sin (\bar{\omega} t_0 + \varphi) \right], \tag{238}$$

$$D = 1 + \mu_i^2 p_i^2(t_0) - 2\mu_i p_i(t_0) \cos (\bar{\omega} t_0 + \varphi). \tag{239}$$

An error will occur if, at the sampling instant, $\eta_i(t_0) \geqq \frac{1}{2}$, which gives the following relation for determining the peak amplitude of a noise pulse at the detector input that will produce an error in the demodulated signal:

$$\begin{aligned} \tfrac{1}{2}\mu_i \left[\mu_i p_i^2(t_0) - p_i(t_0) \cos (\bar{\omega} t_0 + \varphi) - \frac{p_i'(t_0)}{\bar{\omega} \sin (\bar{\omega} t_0 + \varphi)} \right] \\ \geqq \tfrac{1}{2}[1 + \mu_i p_i(t_0) - 2\mu_i p_i(t_0) \cos (\bar{\omega} t_0 + \varphi)] \end{aligned} \tag{240}$$

With $\mu_i = 1/ER(-\bar{\omega})$, this can be written

$$\frac{p_i(t_0) \cos (\bar{\omega} t_0 + \varphi) - \frac{1}{\bar{\omega}} p_i'(t_0) \sin (\bar{\omega} t_0 + \varphi)}{2ER(-\bar{\omega})} = \tfrac{1}{2}. \tag{241}$$

B.3.2 *FM vs. Bipolar AM*

In the case of bipolar AM, the corresponding relation is

$$\frac{p_i(t_0)\cos(\omega t_0 + \varphi)}{2ER(0)} = \tfrac{1}{2}. \tag{242}$$

If it is now assumed that the impulses occur at sampling instants $t_0 = 0$, then $p_i'(t_0) = 0$ and the following relations apply.

For FM:

$$\frac{p_i(0)\cos(\bar{\omega} t_0 + \varphi)}{2ER(-\bar{\omega})} = \tfrac{1}{2}. \tag{243}$$

For bipolar AM:

$$\frac{p_i(0)\cos(\bar{\omega} t_0 + \varphi)}{2ER(0)} = \tfrac{1}{2}. \tag{244}$$

The peak amplitudes of the impulses that will cause errors are thus smaller in FM than in AM by the factor

$$\frac{\hat{p}(0)_{\text{FM}}}{\hat{p}(0)_{\text{AM}}} = \frac{R(-\bar{\omega})}{R(0)}. \tag{245}$$

In accordance with (104) and (105), the following relation applies between the average signal powers in FM and AM:

$$\frac{P_{\text{FM}}}{P_{\text{AM}}} = 2\left[\frac{R(-\bar{\omega})}{R(0)}\right]^2. \tag{246}$$

From (245) and (246) it follows that

$$\frac{\hat{p}(0)_{\text{FM}}}{\hat{p}(0)_{\text{AM}}} = \left(\frac{P_{\text{FM}}}{P_{\text{AM}}}\right)^{1/2} \tfrac{1}{2}^{1/2}. \tag{247}$$

Thus, for equal probability of errors when the impulse occurs at sampling instants and equal average signal power $P_{\text{FM}} = P_{\text{AM}}$, the impulses at the detector input can be greater in AM than in FM by a factor $2^{1/2}$, corresponding to 3 db.

As another limiting case, assume that the impulses occur midways between sampling points corresponding to $\bar{\omega} t_0 = \pi/2$ or $t_0 = \pi/2\bar{\omega}$. In this case (241) and (242) become

for FM and $t_0 = \pi/2\bar{\omega}$:

$$\frac{-p_i(t_0)\sin\varphi + (1/\bar{\omega})p_i'(t_0)\cos\varphi}{2ER(-\bar{\omega})} = \tfrac{1}{2}; \tag{248}$$

for AM and $t_0 = \pi/2\bar{\omega}$:

$$\frac{-p_i(t_0)\sin\varphi}{2ER(0)} = \tfrac{1}{2}. \tag{249}$$

Since φ may have any value, it is permissible in (249) to substitute $\varphi_1 = (\varphi - \pi/4) = \cos\pi/4 \sin\varphi + \sin\pi/4 \cos\varphi$, in which case, for AM

$$\frac{-p_i(t_0)(\sin\varphi + \cos\varphi)}{2ER(0)} = \tfrac{1}{2}^{1/2}. \tag{250}$$

The following approximation applies for a representative pulse shape* and $t_0 = \pi/2\bar{\omega}$:

$$-\frac{1}{\bar{\omega}}p_i'(t_0) \cong p_i(t_0) \cong \tfrac{1}{2}p_i(0), \tag{251}$$

where $p_i(0)$ is the peak pulse amplitude at the detector input, which will occur at a time $t_0 = \pi/2\bar{\omega}$ from a sampling point.

Hence (248) and (250) can be written

for FM:

$$-p_i(0)\,\frac{\sin\varphi + \cos\varphi}{ER(-\bar{\omega})} \cong 1; \tag{252}$$

for AM:

$$-p_i(0)\,\frac{\sin\varphi + \cos\varphi}{ER(0)} \cong 2^{1/2}. \tag{253}$$

The peak amplitudes that will cause errors are thus smaller in FM than in AM by the factor

$$\frac{\hat{p}(0)_{\mathrm{FM}}}{\hat{p}(0)_{\mathrm{AM}}} \cong \frac{R(-\bar{\omega})}{2^{1/2}R(0)}. \tag{254}$$

In view of (246),

$$\frac{\hat{p}(0)_{\mathrm{FM}}}{\hat{p}(0)_{\mathrm{AM}}} = \frac{1}{2}\left(\frac{P_{\mathrm{FM}}}{P_{\mathrm{AM}}}\right)^{1/2}. \tag{255}$$

Thus, for equal probability of errors when the impulses occur midways between sampling points and equal average signal power $P_{\mathrm{FM}} = P_{\mathrm{AM}}$, the peak amplitude of the impulses at the detector input can be greater in AM than FM by a factor of 2, corresponding to 6 db.

* In the case of a pulse shape obtained with a raised cosine spectrum, $p_i(t_0) = \frac{1}{2}\,p_i(0)$ and $p_i'(t_0)\bar{\omega} = 0.475\,p_i(0)$.

The peak amplitudes, $\hat{p}(0)$, required to produce an error when an impulse occurs midways between sampling points is greater than when they occur at a sampling point by a factor of $2^{1/2}$ in FM and a factor of 2 in AM. With a Gaussian amplitude distribution of the pulses, the probability of an error from a pulse midway between two sampling points is in the order of 1 per cent of the probability of an error from a pulse at a sampling point in the case of FM, and is substantially smaller for AM. Hence, virtually all the errors will be caused by pulses that occur near sampling points. The AM advantage over FM for equal error probability is 3 db for impulses that occur at sampling points, and would be expected to be only slightly greater, about 4 db when impulses occurring at all instances with respect to a sampling point are considered.

The above comparisons apply without a postdetection low-pass filter in FM. With an optimum bandpass receiving filter characteristic in FM, the reduction in peak impulse noise afforded by low-pass filter would be expected to be about the same as the reduction in average random noise.

APPENDIX C

Optimum Receiving Filter Characteristic

The optimum receiving filter characteristic in AM and in FM without a postdetection low-pass filter can be determined from the solution of the more general case considered here, of FM with a postdetection filter.

In the latter case, the optimum $R(u)$ is obtained when the product of the two integrals in (125) is a minimum, or for the minimum value of the product:

$$J = J_1 J_2 , \tag{256}$$

where J_1 and J_2 are functions of $R(u)$ given by

$$J_1 = \int_{-\infty}^{\infty} L^2(u) R^2(u) (1 + u/\bar{\omega})^2 \, du , \tag{257}$$

$$J_2 = \int_{-\infty}^{\infty} \frac{S^2(u)}{R^2(u)} \, du = 2 \int_0^{\infty} \frac{S^2(u)}{R^2(u)} \, du . \tag{258}$$

In (257), $L(-u) \neq L(u)$, so that it is convenient to resolve the integrand into one component with even symmetry with respect to u and one with odd symmetry. The integral of the latter component vanishes and that of the component with even symmetry becomes

$$J_1 = \int_0^{\infty} H^2(u) R^2(u) \, du , \tag{259}$$

where

$$H^2(u) = L^2(u)(1 + u/\bar{\omega})^2 + L^2(-u)(1 - u/\bar{\omega})^2. \tag{260}$$

When a small variation, $\delta R(u)$, is made in $R(u)$, the resultant variation in J is

$$\begin{aligned} \delta J &= J_2 \delta J_1 + J_1 \delta J_2 \\ &= J_2 \int_0^\infty 2R(u)H^2(u)\delta R(u)\, du \end{aligned} \tag{261}$$

$$-J_1\, 2 \int_0^\infty 2\left[\frac{R^2(u)}{R^3(u)}\right] \delta R(u)\, du. \tag{262}$$

The optimum $R(u)$ is obtained when $\delta J = 0$, or

$$\int_0^\infty \left[J_2 R(u) H^2(u) - \frac{2J_1 S^2(u)}{R^3(u)}\right] \delta R(u)\, du = 0, \tag{263}$$

which is the case when

$$J_2 R(u) H^2(u) - \frac{2J_1 S^2(u)}{R^3(u)} = 0 \tag{264}$$

or

$$R(u) = R^o(u) = \frac{c 2^{1/4} S^{1/2}(u)}{H^{1/2}(u)}, \tag{265}$$

where $c = (J_1/J_2)^{1/4}$ is a constant.

With (265) in (258) and (259),

$$J_1 = c^2 2^{1/2} \int_0^\infty S(u) H(u)\, du, \tag{266}$$

$$J_2 = \frac{1}{c^2}\, 2^{1/2} \int_0^\infty S(u) H(u)\, du, \tag{267}$$

$$\begin{aligned} J_1 J_2 &= 2\left[\int_0^\infty S(u) H(u)\, du\right]^2, \\ &= \frac{1}{2}\left[\int_{-\infty}^\infty S(u) H(u)\, du\right]^2. \end{aligned} \tag{268}$$

The optimum ratio, N_0/S_0, obtained with (268) in (125) becomes

$$(N_0/S_0)^o{}_{\text{FM}} = \frac{n}{4\pi T P_{\text{FM}}} \left[\int_{-\infty}^\infty S(u) H(u)\, du\right]^2. \tag{269}$$

In the case of FM without a postdetection filter, $L(u) = 1$ and, in this case,

$$H^2(u) = 2(1 + u^2/\bar{\omega}^2), \tag{270}$$

in which case (265) gives (138).

In the case of AM, the term $u^2/\bar{\omega}^2$ is absent in (270) and

$$H^2(u) = 2, \tag{271}$$

so that (265) gives (127).

REFERENCES

1. Nyquist, H., Certain Topics in Telegraph Transmission Theory, Trans. A.I.E.E., **47,** April 1928, p. 617.
2. Sunde, E. D., Theoretical Fundamentals of Pulse Transmission, B.S.T.J., **33,** May 1954, p. 721; July 1954, p. 987.
3. Oliver, B. M., Pierce, J. R. and Shannon, C. E., The Philosophy of PCM, Proc. I.R.E., **36,** November 1948, p. 1324.
4. Bennett, W. R., Methods of Solving Noise Problems, Proc. I.R.E., **44,** May 1956, p. 609.
5. Montgomery, G. F., A Comparison of Amplitude and Angle Modulation for Narrow-Band Communication of Binary-Coded Messages in Fluctuation Noise, Proc. I.R.E., **42,** February 1954, p. 447.
6. Cahn, C. R., Performance of Digital Phase Modulation Communication Systems, I.R.E. Trans., **CS-7,** May 1959, p. 3.
7. Turin, G. L., Error Probabilities for Binary Symmetric Ideal Reception Through Nonselective Slow Fading and Noise, Proc. I.R.E., **46,** September 1958, p. 1603.
8. Watson, G. N., *A Treatise on the Theory of Bessel Functions*, Cambridge Univ. Press, London, 1944.
9. Rice, S. O., Statistical Properties of a Sine Wave Plus Random Noise, B.S.T.J., **27,** January 1948, p. 109.