Communication by Phase Modulation[*]

MURRAY G. CROSBY[†], MEMBER, I.R.E.

Summary—Practical methods of generating and receiving phase modulation are described which open up the possibility of using phase modulation as a communication system. A new receiver is described which uses an off-neutralized crystal filter and provides a simple practical receiver which has not been heretofore available for phase modulation. Other methods of reception are described and discussed.

Propagation tests which were conducted between California and New York indicate that the propagation characteristics of phase modulation are substantially the same as those of amplitude modulation.

The noise characteristics of phase modulation are considered and it is shown that the signal-noise ratio at the output of the phase-modulation receiver is equal to the product of the phase deviation in radians and the carrier-noise ratio.

The chief advantage of phase modulation is realized at the transmitter where a power gain of about four-to-one is obtained and modulating equipment is reduced by the ability to modulate at a low level without the requirement of linearity in the stages following the modulator. The chief difficulty occurs at the receiver where the susceptibility to microphonics is increased and the circuits are slightly more complicated.

INTRODUCTION

DURING the course of the work on frequency modulation which has been described in previous publications,[1,2] development work on phase modulation was also carried on by the engineers of R.C.A. Communications, Inc. The results of the frequency-modulation propagation tests pointed the way to phase modulation as a means of eliminating the extreme distortion encountered when frequency modulation is transmitted over a multipath medium such as the ionosphere. This could be done without losing the advantages of frequency modulation with respect to the ease of modulation and the ability to use class C amplification in all the transmitter stages.

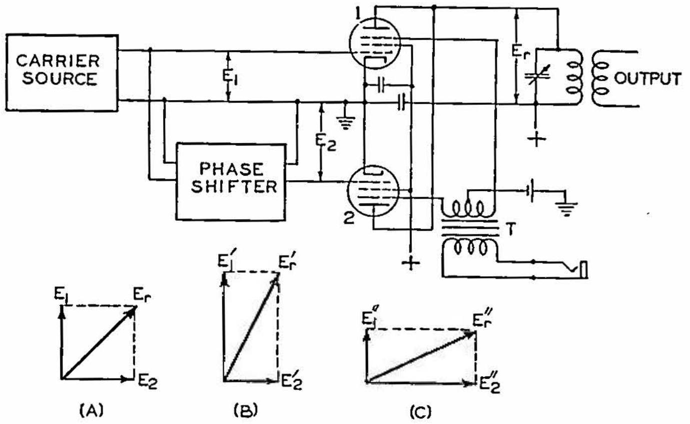

Fig. 1—Differential phase modulator.

Accordingly, in the propagation tests the transmitter was arranged to radiate phase modulation as well as frequency and amplitude modulation so that all three types could be compared. At the receiving end several different types of phase-modulation receivers were developed and given a working test which demonstrated their relative advantages. The results of this work indicated that phase modulation provides a new and important method of communication with many advantages. In the following, the experience received in that work is drawn upon to describe the more practicable methods of generating and receiving phase modulation with the object of placing the system upon a working basis as a means of communication.

THE PHASE MODULATOR

When a wave is phase modulated, its instantaneous phase is deviated from the position it would have taken if the modulation were not present. Such a phase shift may be introduced by passing the wave through a network which imparts a time delay to the wave so that the wave at the output of the network has a phase which is different from that at the input. The problem of generating phase modulation thus becomes a problem of causing the modulating wave to impart time delay to the wave in accordance with the modulating potentials. One method of doing this is shown schematically in Fig. 1. Voltage from the carrier source is fed directly to modulator tube 1. This voltage is represented by the vector E_1 of vector diagram (A) of Fig. 1. Phase-shifted, or time-delayed, voltage is fed to modulator 2 and is represented by vector E_2. The resultant of these two voltages is formed in the common plate circuit of the two modulator tubes and is represented by E_r. (The amplification effected in the modulator tubes is neglected in the vector diagrams.) The modulator tubes are differentially modulated by energy fed to transformer T. Two instantaneous positions of differential modulation are shown in diagrams (B) and (C) of Fig. 1. It can be seen that the resultant voltage is deviated in phase between limits which are determined by the phase separation of the voltages fed to the two modulator tubes.

In the phase modulator of Fig. 2, use is made of the fact that the phase of the output of a tuned amplifier varies as the tuning is varied. Steady carrier energy is fed amplifier 1 which has tuned circuit TC in its output. The tuning of TC is modulated by reactance

---

[*] Decimal classification: R148×R410. Original manuscript received by the Institute, June 16, 1938. Presented before Thirteenth Annual Convention, New York, N. Y., June 17, 1938.

[†] R.C.A. Communications, Inc., Riverhead, Long Island, New York.

[1] Murray G. Crosby, "Frequency modulation propagation characteristics," PROC. I.R.E., vol. 24, pp. 898–913; June, (1936).

[2] Murray G. Crosby, "Frequency modulation noise characteristics," PROC. I.R.E., vol. 25, pp. 472–514; April, (1937).

tube 2 which is given a 90-degree phase shift due to the resistance-capacitance phase shifter, R,C. This type of reactance tube is the same type as is used in automatic-frequency-control practice.[3,4] In the particular circuit of Fig. 2, the modulating potentials are applied to the suppressor grid of the reactance tube.

A simple type of phase modulator, developed by H. E. Goldstine,[5] takes advantage of the fact that the output of a crystal oscillator may be phase modulated by modulating one of the element voltages in the same manner that the ordinary tuned-circuit oscillator may be frequency modulated by modulating one of the element voltages. Apparently either type of oscillator circuit has some degree of reactance tube effect inherent in it, but in the case of the crystal oscillator the stability of the crystal prevents rapid frequency variations from taking place so that only phase deviations are effected.

The circuit of Fig. 3 shows a method of producing phase modulation in which a transmission line is employed.[6,7] By modulating the plate resistance of a tube which acts as the terminating impedance of the line, at some point on the line, or for a given length of line, the combination of the incident wave of voltage applied to the line and the variable amount of reflected voltage produces a resultant which varies in phase. This can be seen from the vector diagrams of Fig. 3(A) and (B) which show the vector relations between the incident and reflected waves of voltage for the point on the line at which these two voltages are 90 degrees out of phase. The incident wave E_i combines with the reflected wave E_r to form the resultant E when the terminating impedance of the line is such as to allow practically full reflection. When the terminating impedance is modulated to a value which reduces the reflected wave to the value given by E_r' in Fig. 3(B), the resultant voltage takes the new position E' which is shifted in phase by the amount Φ.

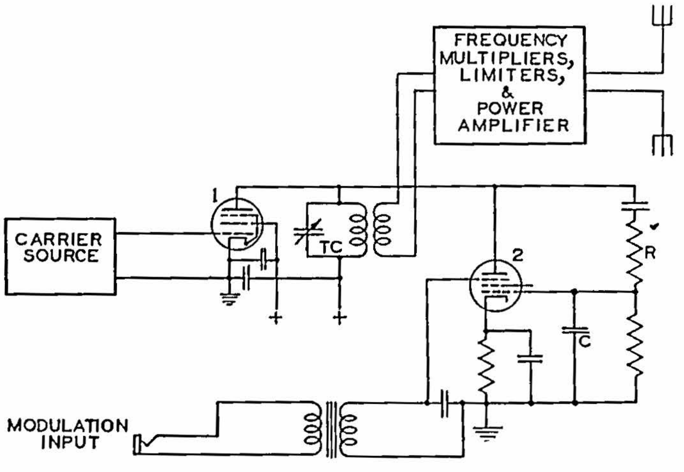

Fig. 2—Phase modulation by modulating a tuned circuit with a reactance tube.

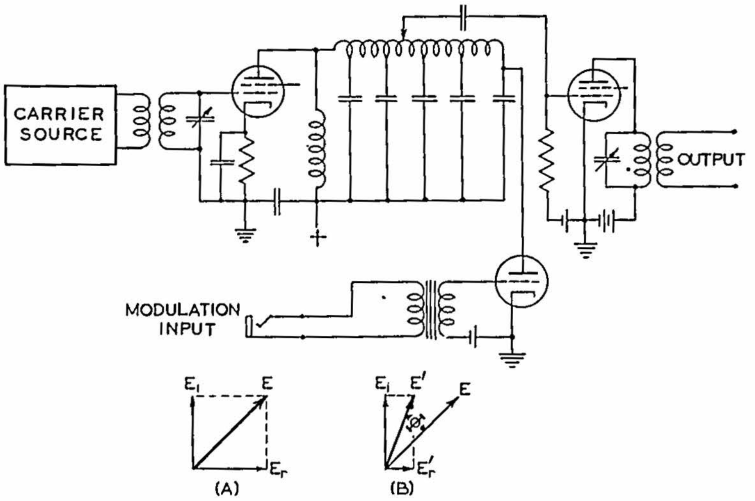

Fig. 3—Phase modulation by modulating the terminating impedance of a transmission line.

A frequency-modulated oscillator may be arranged to produce phase modulation by the application of a network in the modulating-potential circuit which passes these potentials with an amplitude proportional to their frequency.[8,9] This system has the advantage that a high degree of phase modulation is obtainable. It was successfully used in the phase-modulation propagation tests during the year 1931, but has the disadvantage that it lacks the stability of the master-oscillator systems which are used with the other methods.

With any of the above phase modulators, the most convenient arrangement is that in which a low degree of modulation is produced at the modulator and higher degrees are obtained by the use of frequency multiplication. Nonlinear distortion and concomitant amplitude modulation are reduced in this way. Concomitant amplitude modulation may be further reduced by the use of limiting in the stages following the modulators. The degree of modulation may also be increased by the use of cascade modulation[10] in which the radio-frequency output of one modulator is fed to the radio-frequency input of another which adds its phase deviation to that of the first.

---

[3] D. E. Foster and S. W. Seeley, "Automatic tuning, simplified circuits, and design practice," Proc. I.R.E., vol. 25, pp. 289–313; March, (1937).

[4] Other types of reactance tubes in this use are described in U.S. Patents No. 2,033,231, No. 2,087,428, and No. 2,012,710.

[5] U.S. Patent No. 2,111,587.

[6] U.S. Patent No. 2,085,418 also discloses how amplitude modulation may be produced by the same method.

[7] A somewhat similar phase-modulating system, employing a transmission-line section to modulate the tuned amplifier, is described by Austin Eastman, "Fundamentals of Vacuum Tubes," McGraw-Hill Book Company, (1937), page 362.

[8] See U.S. Patent No. 2,085,793.

[9] Hans Roder has also described this system in a discussion in Proc. I.R.E., vol. 20, p. 887; May, (1932).

[10] U.S. Patent No. 2,104,318.

THE PHASE-MODULATION RECEIVER

In order to receive a wave which is phase modulated, a converting circuit which converts the phase modulation into amplitude modulation is used. Then, in a manner similar to the process used in frequency-modulation reception, the amplitude modulation is detected by ordinary methods.

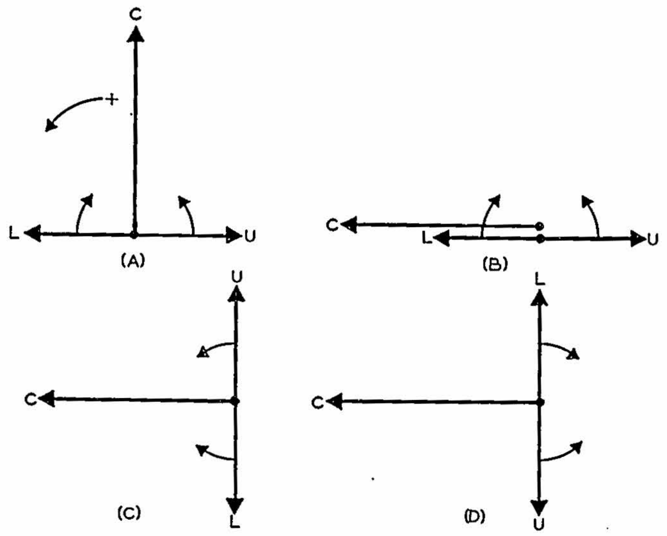

Fig. 4

An approximate explanation of the reason for the converting circuit is given by the following comparison between the carrier and side-frequency relations existing in phase and amplitude modulation: For a given instant of time, the vector relations between the carrier and side frequencies in amplitude modulation are as shown in Fig. 4(A). The phase relation of the three components is such that as the side frequencies rotate with respect to the carrier vector, they combine with the carrier in a manner to add and subtract from the carrier amplitude and thereby vary the resultant amplitude sinusoidally. The relation of the carrier and side frequencies in phase modulation for a given instant of time is as shown in Fig. 4(B). The carrier of the phase-modulated wave is shifted 90 degrees with respect to that of the amplitude-modulated wave. This shift causes the side frequencies to combine with the carrier in such a manner that the amplitude variation produced by one side frequency is canceled by an equal and opposite variation caused by the other side frequency so that only a phase variation of the resultant is produced. (In this approximate explanation, the small amount of amplitude modulation caused by neglecting the side frequencies having an order higher than the first will be neglected.) This phase variation may also be taken as an effective frequency variation since frequency modulation produces the same type of phase variation.

In view of this phase relation existing between the carrier and side frequencies of phase modulation, it can be seen that the following methods may be used to convert to amplitude modulation for subsequent detection: 1. Phase shifting the carrier with respect to the side bands. 2. Phase shifting the side bands with respect to the carrier. 3. Detecting each side band in combination with the carrier separately. 4. Detecting the phase variation as an effective frequency modulation by the use of a frequency-modulation receiver with an equalizing network which corrects for the frequency distortion encountered.

The receivers of the following sections utilize these methods of receiving phase modulation and are considered in order of the author's opinion of their practicality.

OFF-NEUTRALIZED CRYSTAL-FILTER RECEIVER

In this receiver the inherent properties of a simple crystal filter are utilized to convert the phase modulation into amplitude modulation for detection. It has been found that when a crystal filter of the type in which the holder capacitance is neutralized, is operated in the off-neutralized condition, it is capable of converting phase modulation into amplitude modulation. The conversion is effected either by shifting the phase of the side bands with respect to the carrier or by a single-side-band action which allows the separate detection of the side bands in conjunction with the carrier.

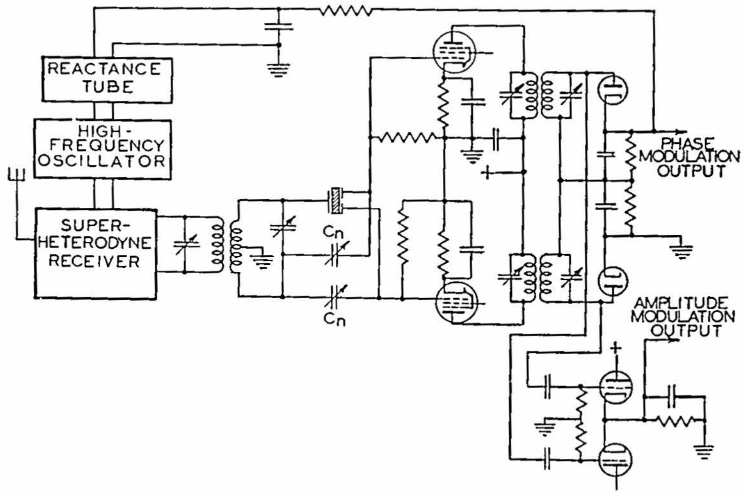

Fig. 5—Off-neutralized crystal-filter phase-amplitude-modulation receiver.

A schematic diagram of this type of receiver is shown in Fig. 5. The type of crystal filter used will be recognized as similar in some respects to that which has been used on amateur single-signal telegraph receivers for some time. A bridge circuit is arranged so that the capacitance of the crystal holder may be over- or underneutralized. One of the holder electrodes is split so as to make it possible to obtain both the over- and underneutralized outputs from the same crystal. It has been found that there is negligible reaction between the two neutralizing circuits with this arrangement and the two outputs are obtained substantially independent of each other.

The two crystal-filter outputs feed diode driver tubes which feed a differential detecting system for the detection of phase modulation and for obtaining automatic-frequency-control voltage. A pair of parallel-connected infinite-impedance diode detectors are also fed by the diode driver tubes for the purpose of detecting amplitude modulation.

The manner in which the off-neutralized crystal filter converts phase modulation into amplitude modulation may be explained as follows: Fig. 6 shows the simplified circuit of one of the filters. When the neutralizing condenser C_n is made equal to the capacitance of the crystal holder C_h the circuit acts as though C_h were removed. A simple resonance curve is then obtained for the input-output characteristic. When the neutralizing condenser is made less than the holder capacitance, the circuit is said to be underneutralized and acts as though C_h is only slightly reduced. The reactance characteristic of the crystal for this underneutralized condition (assuming zero crystal resistance) is as shown in Fig. 7(A). Since the input voltage is fed to the crystal and R_1 in series, the output across R_1 (assuming a constant input voltage) will be dependent on the impedance of the crystal and will have a characteristic as shown in Fig. 7(B). When a carrier and side bands are passed through this type of filter, with the carrier tuned to the peak frequency F_c, a major portion of the side bands appear in the flat portions of the characteristic on both sides of the carrier frequency and the rejection frequency F_1. Since the reactance which feeds R_1 is capacitive for these flat portions of the characteristic, the corresponding side bands will be shifted practically 90 degrees in phase while the carrier will be unshifted. Such a shift in phase relations converts the phase-modulated wave represented by Fig. 4(B) to the relation portrayed by Fig. 4(C) so that the relations are proper to produce amplitude modulation. The side frequencies in the vicinity of the rejection frequency are substantially eliminated so that for these lower-modulation frequencies the phase modulation is converted to amplitude modulation by the removal of one side band. The side frequencies in the immediate vicinity of the carrier frequency are exalted with the carrier so that an increased output might be expected from these modulation frequencies. However, this exaltation is compensated for by the phase shift being smaller in this region and the efficiency of conversion from phase to amplitude modulation being consequently less. Thus, in spite of this combination of single-side-band reception for some modulation frequencies and phase shifting of the side bands for the other modulation frequencies, a flat over-all output may be obtained without equalization. Although it is not indicated on the vector diagrams, a carrier exaltation is also effected by the filter. The degree of this carrier exaltation may be controlled by choice of R_1 (Fig. 6) and choice of Q of the crystal.

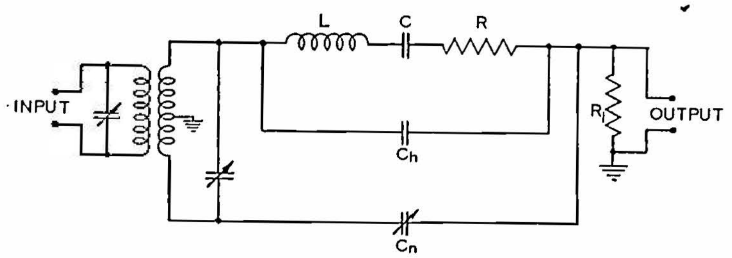

Fig. 6—Equivalent circuit of the neutralized crystal filter. L, C, and R = equivalent constants of the crystal; C_h = holder capacitance; C_n = neutralizing condenser.

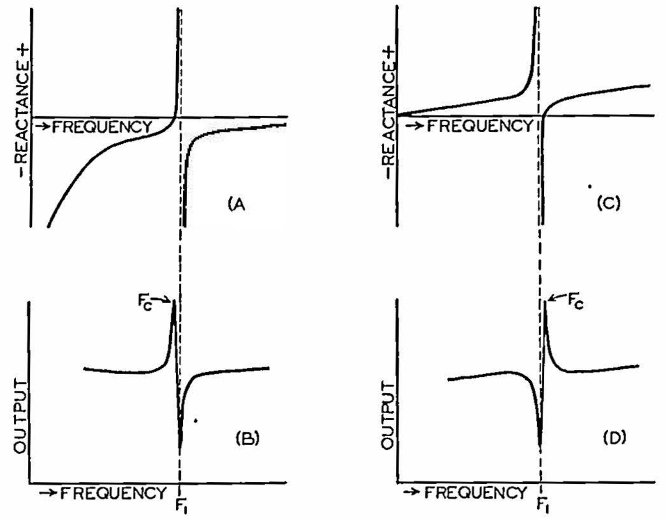

Fig. 7—Reactance and input-output characteristics of off-neutralized crystal filters. (A) and (B) = underneutralized; (C) and (D) = overneutralized.

When the neutralizing condenser is made larger than the holder capacitance, the filter is said to be overneutralized and the reactance characteristic of the crystal is changed as though the capacitance of the holder were replaced by an inductance. The reactance characteristic is then as shown in Fig. 7(C) and the corresponding filter input-output characteristic is shown in Fig. 7(D). It is seen that the reactance which feeds R_1 is inductive for the range of frequencies below the rejection frequency and above the carrier frequency. Thus the side bands are shifted 90 degrees with respect to the carrier, but in a direction opposite to that effected by the underneutralized filter. This converts the carrier and side frequencies of Fig. 4(B) to the relation shown in

Fig. 4(D). Comparing Figs. 4(C) and (D), it can be seen that the amplitude envelope is approaching a maximum point in the case of the underneutralized filter of Fig. 4(C) and a minimum point in the case of the overneutralized filter of Fig. 4(D). Hence the amplitude envelopes are 180 degrees out of phase and the detected outputs must be combined with a 180-degree phase reversal between them so as to make the outputs additive. This combination cancels amplitude modulation present on the incoming signal. In the circuit of Fig. 5, the phase reversal is effected by grounding the cathode of one of the diodes and making the other cathode the high potential point.

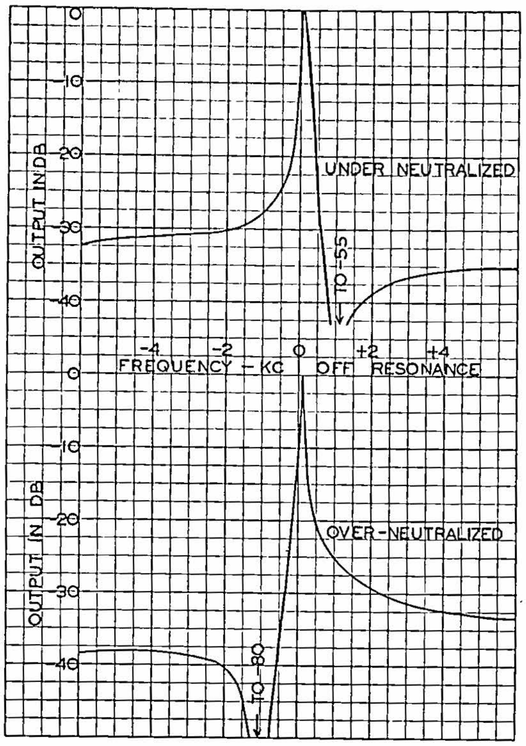

Fig. 8—Typical input-output characteristics of slightly off-neutralized crystal filters. Input held constant.

In the above explanation it was assumed that a high degree of off-neutralization was used. That is, the neutralizing condenser was adjusted well beyond the point of equality with the holder capacitance. When this is done, the rejection frequency occurs close to the carrier frequency and practically equal outputs are obtained from the upper and lower side bands which are disposed above the range between the carrier frequency and the rejection frequency. As the neutralizing condenser is adjusted closer to equality with the holder capacitance, the rejection frequency moves out away from the carrier frequency and the side band on the side of the rejection frequency is reduced with respect to the opposite side band. This is shown in Fig. 8 in which typical under- and overneutralized filter characteristics are shown for the case of a low degree of neutralization. The filters thus effect a single-side-band action as well as a carrier-exalting effect. As a consequence the reception of amplitude modulation is possible on the same receiver by combining the detector outputs in phase. This is done by means of the second pair of detectors of the infinite-impedance type in the circuit of Fig. 5. It has been found that this type of reception of amplitude modulation requires equalization to reduce the overaccentuated low-modulation frequencies. However, a simple equalizer such as a series condenser and a shunt resistance serves the purpose very well.

Adjusting for a low degree of off-neutralization is by far the preferred method of reception since it makes possible the reception of amplitude modulation as well as phase modulation and also allows the detection of a single side band in conjunction with the carrier of either type of modulation. The latter possibility sometimes aids in the reduction of interference.

Automatic-frequency-control energy may be taken from the combined detector output due to the frequency discrimination effected by the fact that the two filter characteristics have their rejection points on opposite sides of the carrier. Hence, as the frequency is varied away from the carrier frequency, the input to one detector decreases at a faster rate than the input to the other. This produces a frequency-discrimination characteristic as shown in Fig. 9. Because of the high selectivity of the crystal filters, automatic frequency control is practically a necessity on this type of receiver.

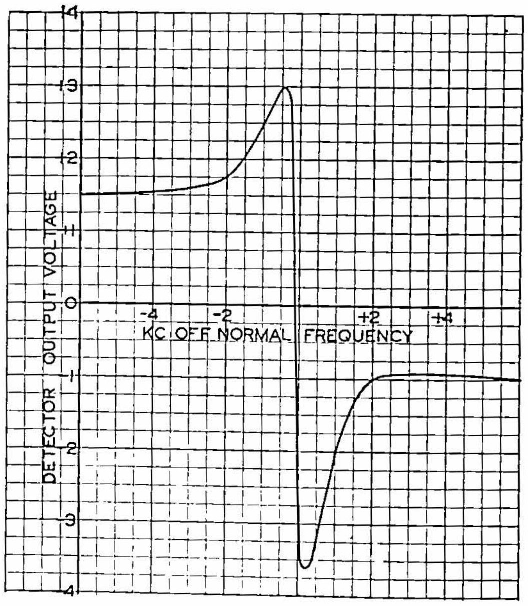

Fig. 9—Typical frequency-discrimination characteristic of off-neutralized crystal filters with diode driver tubes included.

AUXILIARY-CARRIER RECEIVERS

A receiver of this type is shown schematically in Fig. 10. Part of the incoming modulated wave is fed to a carrier filter which removes the side bands and makes available an unmodulated local carrier which may be combined with the phase-modulated signal so as to produce a resultant voltage which is amplitude modulated. Automatic volume control or limiting may be applied to the filtered carrier and it may be combined with the signal at a level such as to produce a carrier exaltation which eliminates the distortion caused by carrier fading. The vector diagram of Fig. 10(A) shows how the filtered and unfiltered voltages combine for the unmodulated condition. The phase adjuster in the filtered-carrier circuit is adjusted so that the two components of the filtered carrier E_1 and E_2, which appear on the push-pull detector input transformer, are 90 degrees out of phase with the unfiltered carrier E_s. This produces resultants E_3 and E_4, which are balanced in amplitude for the unmodulated condition. When the phase of the incoming wave E_s is shifted by the amount Φ as shown in Fig. 10(B), one of the resultants is modulated down in amplitude (E_3') and the other up (E_4'). A phase shift in the opposite direction produces differential amplitude modulation such that E_3 is modulated up and E_4 down. When this differential modulation is detected on the diode detectors, one of the diode resistors must be reversed with respect to the other so as to combine the detected outputs in push-pull as shown. Differential voltage for automatic frequency control is also available from the diode resistors and may be passed through a time-constant circuit to a reactance tube which controls the tuning of a frequency-converting oscillator of the receiver. Due to the high selectivity of the carrier filter, which is most conveniently a quartz-crystal filter, automatic frequency control is practically a necessity on this type of receiver also.

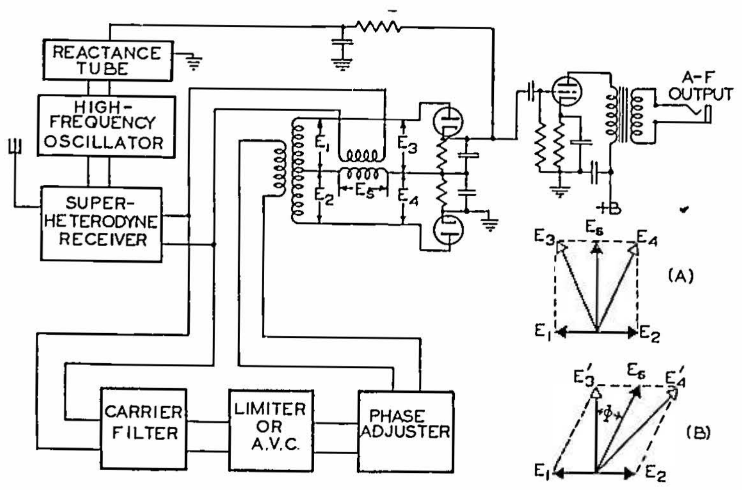

Fig. 10—Auxiliary-carrier phase-modulation receiver.

An alternative to the carrier filter and limiter of the receiver of Fig. 10 is a local oscillator. This local oscillator supplies carrier energy and it is the function of the automatic-frequency-control system to hold the local oscillator in phase synchronism with the incoming carrier. This places a rigid requirement on the automatic-frequency-control system, but it may be eased somewhat by employing a small amount of the incoming carrier to hold the local oscillator barely "in step." Just enough locking voltage is used to maintain phase synchronism, but not enough to allow the local carrier to follow the modulation on the incoming signal.

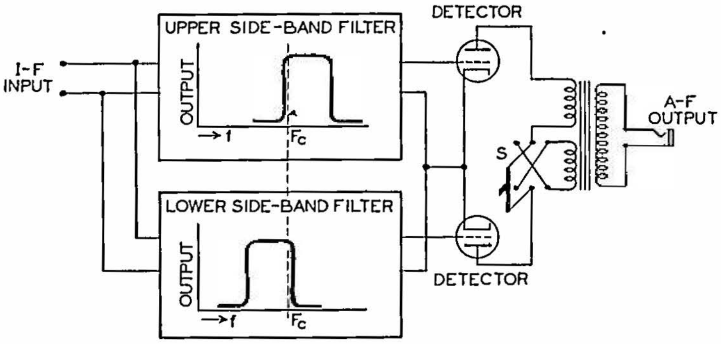

Fig. 11—Single-side-band phase-modulation receiver.

An ordinary autodyne detector may be tuned so as to receive phase modulation by using the "zero-beat" method of reception. The strength of the incoming signal is adjusted so as to hold the oscillating detector barely in step, so that the local oscillator does not appreciably follow the phase modulations on the signal. Thus the local oscillator provides a carrier which is phase shifted with respect to the incoming carrier and the resulting amplitude modulation is detected. This type of reception provides a simple receiver for monitoring, but is rather critical to tune unless provided with automatic frequency control, as described above, together with limiting or automatic gain control of the signal.

SINGLE-SIDE-BAND RECEPTION

In addition to the type of single-side-band reception described in connection with the receiver of Fig. 3, the receiver of Fig. 11 shows how conventional single-side-band filters may be employed to receive phase modulation. Two single-side-band filters are arranged to filter the carrier and each side band separately so that each detector detects the combination of the carrier and one side band at a time. This separate detection of the side bands prevents the output caused by one side band from canceling the output due to the other. By the use of the push-pull connection of the detector outputs, the two detector outputs combine in phase. The parallel combination of the detected outputs, obtained by throwing switch S, allows the reception of amplitude modulation.

EQUALIZED FREQUENCY-MODULATION RECEPTION

The arrangement of Fig. 12 shows how a frequency-modulation receiver may be used for the reception of phase modulation. When phase modulation is received on a frequency-modulation receiver, the inherent difference between frequency and phase modulation makes the audio-frequency output directly proportional to the audio frequency as shown by the solid line of Fig. 12(A). By passing such an output through an equalizing network which passes the audio frequencies inversely proportional to their frequency as shown by the dotted line of Fig. 12(A), the response is equalized so that the over-all output is flat for phase-modulation reception.

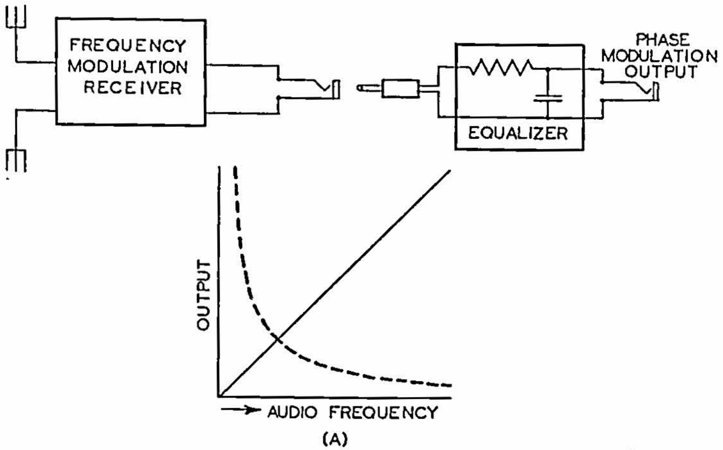

Fig. 12—Frequency-modulation receiver equalized for phase-modulation reception.

PROPAGATION CHARACTERISTICS OF PHASE MODULATION

The general conclusion of the California-to-New York propagation tests was that the propagation characteristics of phase modulation were substantially the same as those of amplitude modulation. If there was any difference between the two systems, it was too small to be detected by the program and tone observations which were applied to phase modulation in the same manner that they were applied to frequency modulation.[1] This would be expected since the side-band characteristics of phase modulation are not radically different from those of amplitude modulation.

Receivers of the off-neutralized crystal-filter and auxiliary-carrier type, as well as the corrected frequency modulation and single-side-band type, were used in the propagation tests. Aside from the power gain resulting in improved signal strengths, the most predominant improvement was that caused by the carrier exaltation effected by the phase-modulation receivers. When a carrier-exalting amplitude-modulation receiver was set up for comparison, the only difference that could be noticed between the two types of modulation was the difference in signal strengths.

It was found that the equalized-frequency-modulation type of phase modulation receiver is subject to rather extreme fading distortion when selective fading is encountered. The distortion took place during the fading minimums at which time a heavily overmodulated signal seemed to result. When this overmodulated signal was detected on the frequency modulation receiver and passed through the equalizing network which accentuated the lower modulation frequencies, the result was a rough, crunching noise bearing little relation to the applied modulation and having a level far above that of the applied modulation. Removal of the frequency-modulation limiter effected no improvement.

A receiver of the single-side-band type (Fig. 11) was set up using a single filter so that only one side band and the carrier could be received at a time. Filtered carrier was also provided so that carrier-exalted single-side-band reception was also possible. It was found that unless carrier-exalted reception was used, distortion due to the intermodulation between modulation frequencies was rather high. It was also found that this single-side-band reception was more susceptible to fading than the double-side-band systems. Since carrier exaltation was provided on both systems, the difference could only be attributed to a frequency diversity in which the probability of transmitting the signal by two side bands was greater than that in which only a single side band was used. This latter effect was noted on both the amplitude- and phase-modulation transmissions.

PHASE-MODULATION POWER GAIN

The carrier power gain effected by phase modulation over amplitude modulation for the unmodulated condition is the same as that for frequency modulation which is said to be four-to-one. This gain is due to the fact that the phase-modulation transmitter may be run continuously at its peak power output and upward modulation does not have to be provided for. When practicable systems are considered, it is found that the power gain ranges from approximately six-to-one for the most inefficient low-level amplitude modulation systems to about three-to-one for the high-level systems in which the modulating power is about equal to the radio-frequency input power. Thus the theoretical figure of four-to-one may be taken as an average value.

When the power gain is considered for the modulated condition, the magnitude of the signal-phase deviation and the signal-noise ratio must be considered. Since the receiver linearly converts phase deviations of the carrier into output voltages, the signal-noise ratio may be found by determining the ratio between the phase deviation of the signal and that of the noise. The effective phase deviation produced by the noise may be deduced by determining the manner in which the carrier and noise voltages combine to form a resultant, which will be both phase and amplitude modulated. This has been done in the previously published paper on frequency modulation[2] in which equation (5) of that paper gives the resultant of the frequency-modulated wave and the noise voltage. This equation may be changed to include the phase modulation case by merely substituting the phase deviation Φ for the modulation index F_d/F_m. The equation may also be changed to consider a single sinusoidal component of the noise resultant instead of the complete resultant of the spectrum,[11] so that the following equation results:

$$e = K \sin \left[\omega t + \Phi \cos pt + \tan^{-1} \frac{\sin (\omega_{na} t + \Phi \cos pt)}{C/n + \cos (\omega_{na} t + \Phi \cos pt)} \right] \tag{1}$$

in which K represents the amplitude envelope of the wave, $\omega = 2\pi \times$ carrier frequency, $\Phi =$ phase deviation of the signal, $\omega_{na} =$ angular velocity of the beat note between the carrier and the noise frequency, $p = 2\pi \times$ modulation frequency, $C =$ peak amplitude of the carrier, and $n =$ peak amplitude of the noise component.

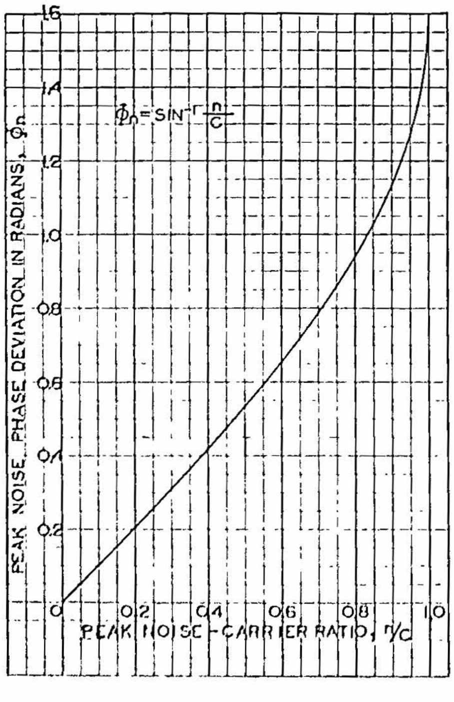

Fig. 13—Relation between the effective phase deviation produced by the noise and the noise-carrier ratio.

The final term in the phase angle of (1) describes the phase deviation produced by the noise. The condition under which maximum phase deviation occurs may be found by equating the first derivative of that term to zero. When this is done for the unmodulated case ($\Phi =$ zero), it is found that the peak phase deviation occurs when the phase angle between the noise and carrier voltages is equal to $\cos^{-1}(-n/C)$. Substituting this value of the phase angle in the original equation, the peak phase deviation due to the noise is found to be

$$\Phi_n(\text{peak}) = \tan^{-1} 1/\sqrt{(C/n)^2 - 1} = \sin^{-1} n/C. \tag{2}$$

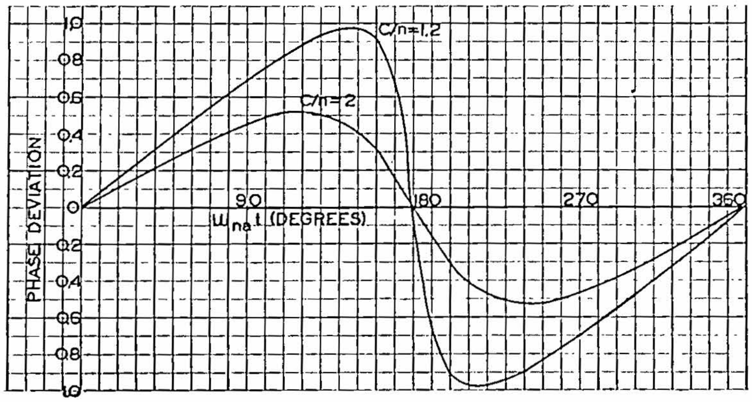

Fig. 14—Wave form of the phase deviation produced by the noise.

When (2) is plotted for various carrier-noise ratios, the curve of Fig. 13 is obtained. It is seen that for the low noise-carrier ratios, the peak phase deviation of the noise in radians is practically equal to the noise-carrier ratio. The curve gradually departs from this equality as the noise-carrier ratio approaches unity, where a peak phase deviation of 1.57 radians (90 degrees) is obtained. This departure from equality occurs mostly in the vicinity of a noise-carrier ratio of unity (at a carrier-noise ratio of one decibel, the departure is less than two decibels). Consequently, for most practical purposes, it may be assumed that the maximum value of the radians of phase deviation produced by the noise is about equal to the peak noise-carrier ratio.

The wave form of this phase deviation produced by the noise has characteristics somewhat similar to the wave form produced by the effective frequency deviation of the noise as plotted in Fig. 4 of the paper on frequency-modulation noise characteristics previously referred to.[2] This is shown in Fig. 14 in which one cycle of the noise phase deviation is plotted for carrier-noise ratios of 1.2 and 2. The wave

---

[11] See page 479 of footnote 2 for description.

form approaches a saw-tooth shape as unity carrier-noise ratio is approached and approaches a sinusoidal shape as the carrier-noise ratio is made large. This fact will undoubtedly cause the crest factor of the noise to increase in the vicinity of unity carrier-noise ratio. Such being the case, it can be seen that when root-mean-square values are considered, the phase deviation of the noise in radians will be more nearly equal to the noise-carrier ratio.

Knowing the phase deviation of the noise, the next step would be to compare the over-all transmissions of the noise for the cases of phase and amplitude modulation. However, this step is unnecessary since the noise component linearly produces a noise output voltage in the same manner for both systems. Consequently a comparison of the signal-noise ratios effected by the noise components in the two cases will be all that is required. In the case of amplitude modulation, the signal-noise ratio is known to be proportional to the product of the carrier-noise ratio and the modulation factor. In the case of phase modulation the signal-noise ratio will be equal to the signal phase deviation divided by the noise phase deviation. Hence (subscript "p" indicates the phase-modulation system),

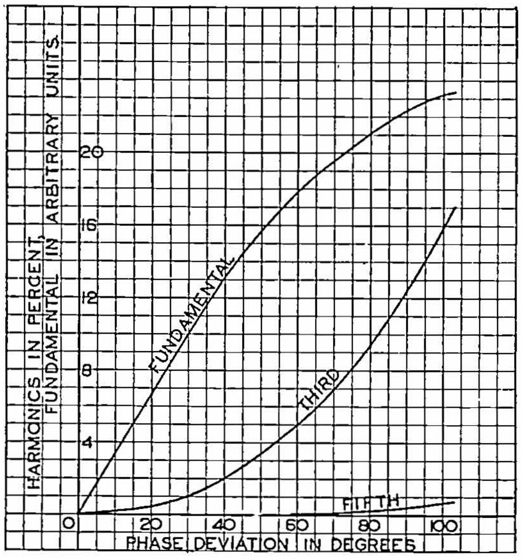

Fig. 15—Theoretically determined fundamental and harmonic outputs of a phase-modulation receiver.

$$S_p/N_p \text{ (peak values)} = \Phi/\sin^{-1} n/C \qquad (3)$$

or, for most practical purposes,

$$S_p/N_p \text{ (peak values)} = \Phi C/n. \qquad (4)$$

Thus the ratio between the signal-noise ratios of the phase- and amplitude-modulation receivers is, for most practical purposes, given by (subscript "a" indicates the amplitude-modulation system)

$$\frac{S_p/N_p}{S_a/N_a} \text{ (peak values)} = \frac{\Phi C_p/N_p}{MC_a/N_a} \qquad (5)$$

where M is the modulation factor in the amplitude-modulation system.

From (5) it can be seen that for equal carrier strengths, the ratio between the signal phase deviation in radians and the modulation factor of the amplitude-modulation system gives the ratio of the signal-noise ratios of the two systems. To evaluate this factor, it remains to determine the permissible signal phase deviation used in the phase-modulation system.

It happens that harmonic distortion appearing in the receiver output places a limitation on the phase deviation for the receivers of the type which depend upon phase shifting of the carrier or side bands, or upon a single-side-band action. An analysis of the output of these types of receivers, which the author intends to submit for publication at a future date, indicates that the output consists of the fundamental and only the odd harmonics since the even harmonics are balanced out by the push-pull detection system. The fundamental is proportional to the first-order Bessel function of the phase deviation $J_1(\Phi)$, the third harmonic is proportional to the third-order Bessel function $J_3(\Phi)$, and so on. Thus, by consulting the Bessel function tables, the curves of Fig. 15 are obtained showing the amplitude of the fundamental and harmonics as the phase deviation is varied. From these curves it can be seen that the phase deviation may be carried to about one radian or 57.3 degrees for a harmonic distortion of about 5 per cent. Although for high-fidelity program transmission, less than a radian of phase deviation might be used to keep the distortion down, and for low-quality systems a higher deviation and distortion would be allowable, for most purposes one radian could be considered an average value for power calculations.

Substituting a signal phase deviation of one radian and a modulation factor of unity in (5), it can be seen that for equal carrier-noise ratios the peak signal-noise ratios obtained from the two systems are about equal. Consequently all of the gain caused by phase modulation is effected by the increased carrier-noise ratio which is brought about by the increase in transmitter power of four-to-one. At a carrier-noise ratio of unity, about 4 decibels of this 6-decibel gain are lost, but at a carrier-noise ratio of 1 decibel, less than 2 decibels are lost. This loss would be less apparent if root-mean-square instead of peak values were considered.

ADVANTAGES AND DISADVANTAGES

The main advantage obtained by the use of phase modulation is realized at the transmitter. The reduction of the amount of modulating equipment required and the power gain obtainable present an advantage which greatly outweighs the small disadvantages which the system has. Since modulation may be accomplished at the master oscillator or its following stage where the level is low, and since the amplitude linearity of the power amplifier may be practically disregarded, many of the troubles encountered in the use of amplitude modulation are eliminated. In general it may be said that for a given complement of tubes in the transmitter, practically four times the carrier-power output can be realized by phase modulation as compared to amplitude modulation and this gain (6 decibels) is fully realized in terms of signal-noise ratio at the receiver.

At the receiver, the main advantage obtained is that caused by carrier exaltation. This, of course, is obtainable with amplitude modulation also. The receiver is somewhat more complicated due to the addition of the carrier-exalting circuits with their requirement of automatic frequency control. However, in the case of the off-neutralized crystal-filter receiver this complication is about the same as that encountered when automatic frequency control is applied to any type of receiver. Furthermore, it is the author's opinion that the improvement effected by carrier exaltation alone makes the added complication worth while.

The main difficulty encountered at the transmitter is the somewhat increased susceptibility to the introduction of alternating-current hum. However, almost this same susceptibility is present with the use of amplitude modulation because fading sometimes converts phase modulation into amplitude modulation.

The chief difficulty at the receiver is the extremely increased susceptibility to microphonics on the oscillators. Apparently most oscillators are being microphonically modulated with frequency modulation which does not show up in amplitude- or frequency-modulation reception, but when this small frequency deviation is received as phase modulation the effective phase deviation is very great. This is especially true in the case of the microphonics having low periods such as produced by bumps or jars of the oscillator tube or circuit. In the case of the ordinary high-frequency oscillator of a superheterodyne receiver receiving a signal in the vicinity of 10 megacycles, the microphonics produced by a person walking around in the room are strong enough to be only a few decibels less than the receiver output due to full modulation. This susceptibility requires special treatment of the high-frequency oscillator. Shock- and soundproofing may be employed together with careful design of the parts of the oscillator circuit to reduce the possibility of vibration. Another alternative is the use of a crystal oscillator for the high-frequency oscillator. Such a system requires either a tunable first intermediate frequency or the use of a low-frequency oscillator which heterodynes the crystal oscillator so that a stabilized beat output is obtained. The stabilized beat output acts as the high-frequency oscillator and is variable over the small range covered by the low-frequency oscillator. This alternative eliminates the necessity of soundproofing since the crystal oscillator is stable enough to be free from microphonics and the low-frequency oscillators are oscillating at a low enough frequency so that the frequency deviation due to the microphonics is small. However, such systems have the disadvantage of added complication and cost.

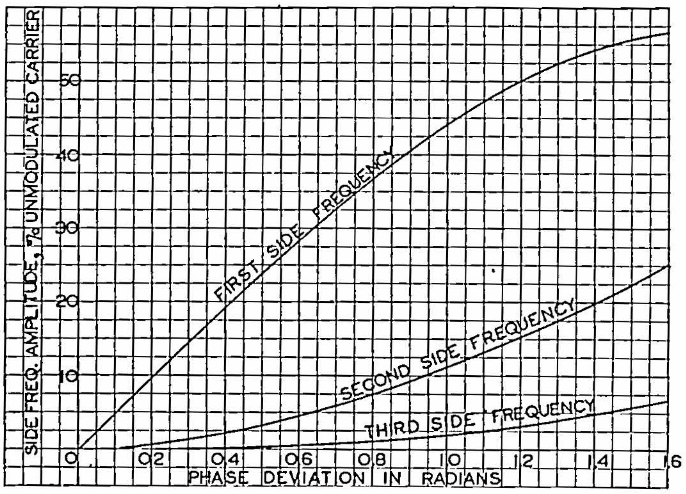

Fig. 16—Side-frequency amplitudes versus phase deviation for single-tone phase modulation.

It might be contended that the presence of the higher-order side frequencies of phase modulation place this type of modulation at a disadvantage as compared to amplitude modulation with respect to adjacent-channel interference. (See Fig. 16 showing the relative amplitudes of the side frequencies for single-tone phase modulation.) However, the following points dispute that contention:

1. The difference between the higher-order side frequencies produced by slightly overmodulated phase- and amplitude-modulation transmitters is likely to be quite small. Phase modulation does not have the well-defined limit of 100 per cent modulation beyond which all the higher-order side frequencies increase very rapidly.[12] Instead there is a gradual increase of the second side frequency with only a slight amount of third side frequency at the full modulation of one radian.

2. In an analysis of the side frequencies of phase and frequency modulation for the case of more than

---

[12] I. J. Kaar, "Some notes on adjacent channel interference," Proc. I.R.E., vol. 22, pp. 295–313; March, (1934).

one modulating frequency, which the author has submitted for publication,[13] it is shown that the presence of the low modulation frequencies in conjunction with the high modulation frequencies reduces the higher-order side-frequency amplitudes. In other words, the presence of a bass viol with a violin tends to reduce the higher-order side frequencies which would be produced if the violin were present alone. This phenomenon is accumulative so that the greater the number of modulating frequencies in the modulating wave, the more will the wave be confined to its channel.

3. The amplitudes of the higher modulation frequencies of program and voice modulation (excluding frequency inverting or other secrecy systems), which produce the out-of-channel interference, are known to be less than those of the lower modulation frequencies. Hence only a small amount of adjacent-channel interference should be caused by the higher-order side frequencies of these higher modulation frequencies. This situation also accentuates the effect mentioned in point number 2.

In view of the above points there is some doubt as to which system will produce the most adjacent-channel interference. At any rate, it can be seen that the difference will not be as great as the sinusoidal side-frequency resolution of phase modulation might indicate.

ACKNOWLEDGMENT

The author gratefully acknowledges the helpful suggestions and encouragement received from Messrs H. H. Beverage and H. O. Peterson during the course of this work.

He also wishes to accredit the promotive enthusiasm of Mr. C. W. Hansell together with the inventive contributions of Mr. Hansell and his staff at the Transmitting Design Laboratories of R.C.A. Communications, Inc. at Rocky Point, New York.

Mr. R. E. Schock assisted with the tests and development work at the receiving end.

Mr. J. W. Conklin was active in developing the transmitter which was used in the propagation tests, and handled the transmissions during the early part of the tests. Mr. T. J. Boerner handled the transmissions during the latter part of the tests.

---

[13] Murray G. Crosby, "Carrier and side-frequency relations with multi-tone frequency for phase modulation," *RCA Rev.*, vol. 3, pp. 103–106; July, (1938).