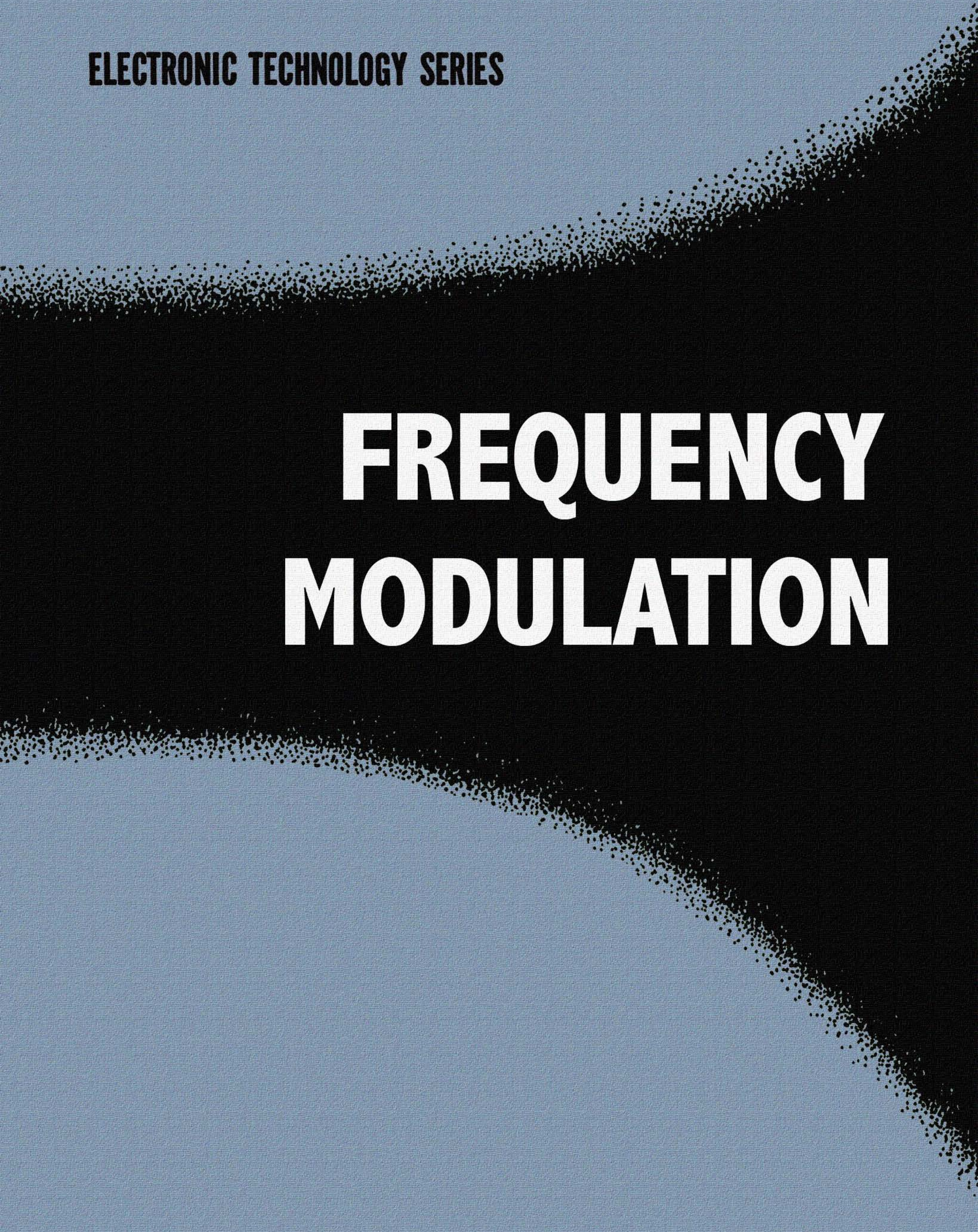
ELECTRONIC TECHNOLOGY SERIES



FREQUENCY MODULATION

a RIDER publication

$.90

Frequency Modulation

Edited by

Alexander Schure, Ph. D., Ed.D.

JOHN F. RIDER PUBLISHER, INC.
116 West 14th Street • New York 11, N. Y.

FIRST EDITION

Library of Congress Catalog Card No. 55-12396

Printed in the United States of America

PREFACE

The subject matters presented herein are designed to give a comprehensive review of the more important principles of frequency modulation. The concepts, designs, and theory differences inherent between f-m and a-m are evolved to permit logical extension of familiar a-m developments to the study of f-m. Thus the diagrams and schematics used to illustrate and simplify each topic permit application of familiar vacuum tube and conventional radio circuit theory. The methods of presentation are essentially those that have been successful in helping trainees understand f-m.

An entire chapter has been devoted to the fundamental ideas of f-m. Similarly, another entire chapter explains the production of f-m. The analyses presented therein are such as to render these subjects thoroughly comprehensible to the technician or junior engineer desiring a review of these essential concepts.

A detailed nonmathematical explanation has been given of phase modulation. Although phase modulation (pm) produces frequency modulation and thus gives the same end result, the techniques utilized in pm vary sufficiently to require thorough understanding by the electronic practitioneer.

The pertinent considerations of the propagation and reception of f-m complete the essential review elements. F-m detectors and limiters have been excepted, since these are the subject of a separate review booklet in this series.

Grateful acknowledgment is made to the staff of the New York Technical Institute for its preparation of the manuscript for this booklet.

New York, N. Y.
October, 1955

A. S.

CONTENTS

Chapter 1

FUNDAMENTAL CONCEPTS OF FM

1. FM and AM Compared

Sound may be transmitted over a distance by superimposing its instantaneous pressure variations upon a radio frequency (r-f) carrier wave. The sound is first converted to an electrical audio frequency (a-f) signal, which is in turn made to vary some characteristic of the carrier. The carrier's amplitude, frequency, or phase may be varied by a process called *modulation.* Frequency modulation (fm) and phase modulation (pm) are not distinctly different, since one cannot exist without the other and one can be used to generate the other.

Since the early days of radio, great strides were made in improving amplitude modulation (a-m) transmitters and receivers. Power output of radio broadcasting stations steadily increased while receivers were made more sensitive and more selective. These measures were aimed at overcoming several notable defects of radio communication: (a) loss of signal strength over long distances, (b) fading due to atmospheric changes, (c) poor reliability of communication during heavy sunspot activity, and (d) high signal-to-noise ratio required to override man-made and atmospheric noises in the receiver. Regarding the last, satisfactory reception with an a-m receiver is not assured until the signal-to-noise ratio at the antenna input connection can be made at least about 100:1 —a "brute force" approach to the solution of an annoying problem because the condition is brought about by the sheer force of the desired signal, rather than by techniques of reception. The problem is acute in large cities where the local noise is extremely difficult to obscure by the signal.

The advent of fm has virtually put an end to the noise problem. The very nature of frequency modulation makes it feasible to have receiver-input signal-to-noise ratios as low as 2:1 without objectionable interference in the output of the receiver. For the most part, both natural (atmospheric) and man-made noises are forms of a-m. Receivers designed to be insensitive to a-m—to respond to frequency or phase variation in the r-f carrier while ignoring amplitude changes—are remarkably quiet even in the vicinity of sparking motors, lightning discharges, and other forms of heavy "static." This one advantage of the f-m system counts heavily in its favor and, even if there were no others, would be sufficient justification for its use.

Although frequency modulation, as a process, does not possess better possibilities for high fidelity than a-m, improved reproduction of the transmitted signal is far more likely when wide-band fm is used due to certain indirect, less obvious factors to be discussed later in this chapter. Mention should be made at this point, however, that many a-m stations limit themselves to a total bandwidth of 15,000 cps due to a ruling of the FCC that requires transmitters to restrict the audio frequencies used for modulation to a maximum of ± 7500 cps if interference is being caused. With the present state of congestion in the r-f spectrum in the United States, it is doubtful whether more than a small percentage of the radio stations in this country exceed the 15-kc bandwidth resulting from the use of an a-f modulation up to 7500 cps, even though they may do so if it does not result in interference. This means that, in many cases, frequencies above 7500 kc are *not* transmitted; this is not a particularly severe restriction for speech transmission but does become noticeable, particularly to the trained ear, as a loss of fidelity in the transmission of orchestral and instrumental music.

However, the real practical limitation of a-m broadcast reception frequency response is in the *receivers*. These receivers are made selective so they will reject interference from undesired distant stations at night. This selectivity, plus deliberate frequency response limitations in the a-f amplifier and loudspeaker system (compromise with cost) make the frequency response of nearly all a-m broadcast receivers far less than that transmitted by any a-m broadcast station.

F-m broadcast station transmitters are required to be capable of modulation at frequencies up to 15,000 cycles per second. The chief reason for this extended range is that these stations operate at carrier frequencies in the very-high frequency portion of the spectrum where there is room for assignment of channels each 200 kc wide. Contrast this with the usual bandwidth of the a-m stations in this country and it is not difficult to see why higher fidelity is to be expected from f-m broadcast stations.

F-m stations other than the broadcast type also operate on frequencies above 30 mc. However, most f-m stations for *communications* employ a limited modulation-frequency range for speech only, and have a relatively small bandwidth. With these systems high-fidelity response is not a factor, but the signal-to-noise advantage of f-m is paramount.

Signals in the very high frequency range are not normally reflected from the ionosphere, and thus do not have the property of long-distance night-time transmission as do a-m broadcast signals. They are also attenuated by rough terrain and travel best over "line-of-sight" paths. Even in the daytime, a-m broadcast signals can usually be *detected* over a greater distance than standard f-m broadcast stations of the same power.

However, the noise-rejection features of the f-m system and its relative freedom from phase effects causing selective fading are such that the primary service area may be greater for standard f-m broadcasting than for standard a-m broadcasting. The "primary service area" of a broadcast station is defined by the FCC as "the area in which the ground wave [including space wave] is not subject to objectionable interference or objectionable fading." The absence of sky waves at very high frequencies prevents interarea interference in the f-m broadcast band.

Another advantage of fm is that interference between adjacent stations is greatly reduced. Tests conducted in various locations have shown that even when two f-m broadcast stations are using the same frequency, one of these will completely "take over" the receiver if its strength is somewhat more than twice that of the other. (For lower deviation than 75 kc, a greater ratio would probably be required.) Again, contrast this with a-m broadcasting, where the desired signal must have a strength better than 30 times that of the undesired signal for satisfactory reception.

The use of high radio frequencies for f-m transmitters and receivers adds another advantage which is not immediately apparent. At these frequencies antennas may be designed for very efficient response, particularly with respect to gain and directivity where they cannot be so constructed at the frequencies employed for a-m stations. The bulk of a half-wave dipole decreases as the frequency rises. This accounts for the large number of parasitic arrays used with very high frequency receiving equipment. At 100 mc, a half-wave dipole need be only 4.5 feet in length, whereas at 1000 kc a half-wave length is approximately 450 feet.

2. F-M Terminology

The work of studying or reviewing any technical subject is always facilitated if all the terms to be used in the explanations are clearly defined. It is necessary to remember just what f-m means in connection with the meanings of these terms. In frequency modulation, the audio signal causes the frequency of the r-f carrier to change so that this change in frequency may later be interpreted as (1) the original audio frequencies present in the modulating sound and (2) the various amplitudes of these audio frequencies.

With this in mind, some of the fundamental f-m terms are defined as follows:

Center frequency is the frequency of the emitted carrier when *no modulation* is present; it is also the carrier frequency at the instant when the modulating cycle is at its zero point. It is also called the *resting* or *idling* frequency of the system and is the one to which reference is made when the station frequency is given in the newspapers and radio magazines.

Frequency deviation may be formally defined as "the peak difference between the instantaneous frequency of the modulated wave and the carrier frequency." If a carrier center frequency is 100 mc and modulation causes the carrier frequency to shift to 100.01 mc, the deviation is said to be 10 kc (.01 mc).

The *frequency swing* of an f-m transmitter is the *range* over which the frequency of the carrier changes. For a given deviation as defined above, the frequency swing is twice this value. In the example, a deviation of 10 kc from the resting frequency would indicate that, during symmetrical modulation by a fixed audio signal, the frequency swing would be 20 kc. That is, the deviation of +10 kc and −10 kc around the center frequency constitutes a frequency swing of 20 kc. In the forthcoming chapters, when reference is made to a certain deviation, a + or − sign will precede the number to indicate whether the frequency excursion is higher or lower than the center frequency, respectively. On the other hand, no algebraic sign will be used when frequency swing is discussed; it is to be assumed by the reader that this term refers to total range over which the transmitter frequency is changing.

Definitions of other terms will follow the discussion of the f-m wave.

3. The F-M Wave

Figure 1 illustrates the manner in which an audio signal produces either a.m. or fm on a radio frequency carrier and presents a comparison of the "appearances" of the two waves in diagrammatic form. Assuming that the interval of observance of these waveforms is 1/1000 second, the audio frequency must be 2000 cps, since two cycles are completed in the interval. For the purposes of analysis, assume that the radio frequency is 50 kc (50,000 cps).

The audio modulating voltage (2) has the following values:

at points *A, C, E, G, I*	zero voltage
at points *B* and *F*	maximum positive voltage
at points *D* and *H*	maximum negative voltage

In amplitude modulation (3), it is seen that at the points of zero audio voltage there is no effect upon the "idling" amplitude of the carrier wave. Wherever the audio voltage rises to a positive maximum (*B* and *F*), the amplitude of the carrier correspondingly increases, and where the audio goes down to its negative peaks (*D* and *H*), the amplitude of the carrier is reduced. Thus, the changing audio voltage appears as a varying carrier amplitude; the carrier is then transmitted to the receiver, where the detector separates the audio voltage and restores it to its original form, preparatory to reproduction in the headphones or loudspeaker. The maximum amount of permissible carrier variations occurs when the carrier rises to twice its unmodulated voltage, as would occur at points *B* and *F* if sufficient audio power were available, and falls to zero on negative half-cycles. This is termed 100 per cent modulation. Any attempt to bring about more modulation than this leads to severe distorition and the production of undesirable spurious frequencies.

In the f-m wave (4), the amplitude of the carrier remains constant. The circuits are so designed as to cause the *frequency* of the carrier to vary with the audio voltage. As the positive peaks of audio are approached (*B* and *F*), the frequency of the carrier increases toward its peak frequency—the positive limit of deviation; swing of the instantaneous value of the a-f voltage toward negative audio maxima results in a negative carrier deviation toward its minimum

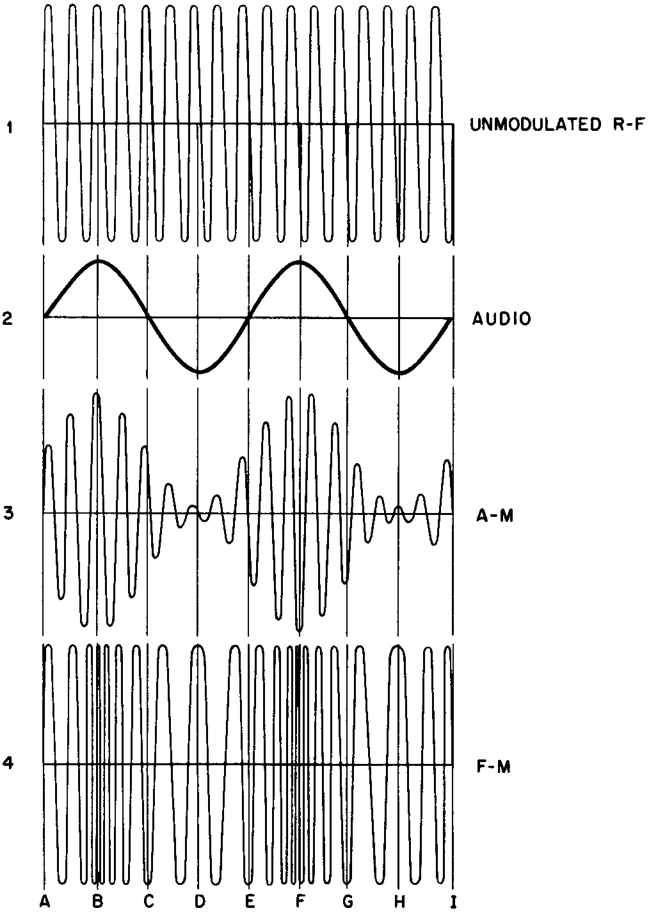


Fig. 1. Comparison between waveforms of a-m signal (3) and f-m signal (4), each derived from a single-frequency carrier (1) and modulating signal (2).

frequency. Evidently, at points *A, C. E. G,* and *I,* the carrier frequency is that of the unmodulated state, because the instantaneous audio voltage is zero at these times.

An assumed set of figures should help to clarify the phenomenon, and at the same time illustrate the use of the f-m terms previously defined:

Center Frequency	50 kc
Audio Frequency	2000 cps
Assumed + deviation	+75 cps
With symmetrical modulation,	
The deviation is	75 cps
The frequency swing is	150 cps

4. Secondary Definitions

Unlike a.m., fm does not, as a process, contain a set of conditions that rigorously define percentage modulation. From a purely theoretical point of view, any frequency swing is possible, provided that there is enough audio power to produce the swing, down to the point where the carrier frequency would be reduced to zero on negative half-cycles. For example, if the carrier frequency is 50 kc as in our illustration, the frequency deviation *could be* + and —50,000 cycles per second; on the downward sweep, of course, the frequency would reach zero, and on the positive half-cycle it would rise to 100,000 cps. Still, there would not have to be distortion if operating conditions were properly set. But, with this amount of frequency swing, the spectrum space required would be absolutely prohibitive since the deviation is equal to the carrier frequency. On the other hand the frequency swing permitted to f-m broadcast stations by the FCC, which is limited to 150 kc (±75 kc deviation), is practical because it represents such a small fraction of the carrier frequency. This swing is defined as "100 per cent modulation" and should be recognized as a limitation dictated by arbitrary choice rather than by technical requirements.

The 75-kc deviation figure applies only to f-m *broadcast* stations. As explained below, f-m stations in other services are assigned other maximum deviation values. F-m *communications* systems such as police and fire, radio, utilities, etc., are ordinarily limited to a deviation of 15 kc. In amateur service, "narrow-band fm" is defined as that which produces a bandwidth no greater than that necessary for an a-m signal with the same modulation frequency range. More will be said about this later in the book.

With these facts in mind, several other important f-m terms may now be defined:

Percentage modulation is the ratio of the actual frequency swing of the transmitter to the frequency swing arbitrarily defined as 100 per cent modulation by the FCC. For f-m stations operating between 88 and 108 mc (the so-called f-m broadcast band) 100 per cent modulation means a frequency swing of 150 kc; for the sound portion of standard television broadcast signals, 100 per cent modulation has been defined as a frequency swing of 50 kc.

As an example, consider a transmitter in which the modulating voltage produces a frequency of 20 kc. If this is a commercial f-m station, the percentage modulation is 20/150 = 13.3 per cent. If it is a television station, then the modulation percentage is 20/50 = 40 per cent.

Modulation index is used in determining how the sidebands of the transmitted signal are distributed around the carrier frequency, a matter to be discussed later in this book. It is found by dividing a given frequency deviation by the frequency of the audio modulating signal. For instance, if a modulating signal of 4000 cps produces a frequency deviation of +20 kc, the modulation index is:

$$20{,}000/4000 = 5$$

Deviation ratio is a quantity established by the FCC in this country for control of maximum frequency swings. It is obvious that, in any commercial system of communication, the bandwidth occupied by the station must be fixed by law to prevent overlapping of transmitted signals and concomitant interference. The deviation ratio is officially defined as the ratio of the maximum frequency deviation to the highest audio frequency used in modulating the transmitter. In this country, it is set up as follows:

Standard fm, 88 to 108 mc
Maximum deviation = ±75 kc
Highest audio frequency = 15,000 cps
Deviation Ratio = 75,000/15,000 = 5

Television sound

Maximum deviation = ± 25 kc
Highest audio frequency = 15,000 cps
Deviation Ratio = 25,000/15,000 = 1.67

5. Summary

In the light of what has gone before, it would be helpful to summarize the effect upon the carrier of an audio voltage used for frequency modulation.

Effect of the frequency *of the audio.* The number of periodic frequency changes which occur per second in the carrier is exactly equal to the audio frequency used for modulation. In the example of Fig. 1 the audio frequency is 2000 cps; thus the carrier swings through its peak deviations 4000 times in each second.

Effect of the loudness *of modulating sound.* Assuming that the magnitude of the audio voltage is proportional to the sound intensity, then the frequency swing is determined by the loudness of modulating sound. The louder the sound, the greater the frequency swing.

6. Review Questions

(1) Name the three wave characteristics that, when varied by a modulating source, may produce usable modulation.

(2) Explain the popular misconception that "fm is capable of higher fidelity than a.m."

(3) What is the most important advantage fm possesses over a.m?

(4) What is the highest audio frequency that may be used by a-m broadcast stations when a possibility of interference with other stations exists? by an f-m broadcast carrier under *any* conditions?

(5) Why are f-m broadcast stations permitted to use higher audio frequencies than a-m stations?

(6) Why is much more efficient antenna design possible for f-m reception than for a-m reception?

(7) Distinguish carefully between "frequency deviation" and "frequency swing."

(8) Explain the difference between "modulation index" and "deviation ratio."

(9) Why must the definition of 100 per cent frequency modulation be an arbitrary one?

(10) If the *frequency swing* of some mythical f-m station is 700 kc, its carrier frequency is 95 mc, and the highest audio frequency it will use for modulation is 20,000 cps, what is its deviation ratio?

Chapter 2

THE PRODUCTION OF FM

7. A-M Generators

The chain of events that leads to the emission of a radio signal from an antenna must start with an *oscillator;* this is a generator of radio-frequency energy. From this point onward through the chain are found amplifiers, control stages, and modulators.

To produce a.m., the modulator must somehow force the amplitude of the r-f voltage to vary in step with the modulating voltage. This behavior is easy to obtain because output amplitude depends upon so many things—plate voltages, screen voltages, control grid voltages, cathode voltages, and suppressor grid voltages—and variations in any one of these characteristics can cause amplitude changes.

For instance, consider a simple oscillator like that shown in Fig. 2. This drawing shows an Armstrong oscillator with a carbon-granule microphone in series with the tuning capacitor across the tank coil. When no sound is impressed on the diaphragm of the microphone, the radio-frequency output of the circuit will have a fixed amplitude and a frequency determined by the values of L_1 and C_1. When sound is allowed to act upon the microphone, the carbon granules become compressed and decompressed as the air waves strike the diaphragm. This changes the resistance of the microphone, which in turn affects the amplitude of the r-f current flowing in the tuned circuit. As this current changes, the output r-f amplitude changes in exact accordance with the sound waves. This is amplitude modulation. At the same time it should be noted that the frequency of the circuit does not change significantly, because only the resistance is changing; resistance variations have little effect upon frequency of oscillation. From this it may be concluded that the amount of amplitude variation obtained depends upon the intensity of the sound reaching the microphone and that there is little or no change in frequency.

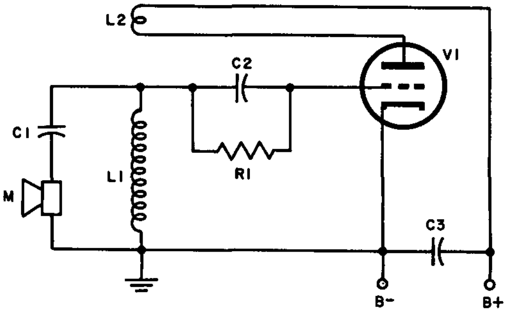


Fig. 2. Carbon microphone varies resistance across oscillator tuned circuit to provide amplitude modulaton.

Practical a-m radiotelehone transmitters use this principle in essence. That is, by having the sound produce a voltage or current which then is combined with the output of the r-f portion of the transmitter in such a manner as to cause amplitude variations of the emitted wave, amplitude modulation is realized.

8. Simple F-M Generator

By substituting a condenser-type microphone for the carbon granule microphone just discussed, the Armstrong oscillator becomes an f-m transmitter of a very primitive variety. The connection is shown in Fig. 3. The microphone is no longer in series with tuning capacitor *C1* but has been placed in parallel with the tuned circuit consisting of *C1* and *L1*. The frequency of the oscillator output is now governed by the constants *L1, C1* and the capacitance of the microphone *M*. When no sound is allowed to act upon *M*, the carrier remains at its center or resting frequency. As the compressional air waves due to the modulating sound reach the diaphragm of the microphone, it vibrates toward and away from the fixed back plate, changing its capacitance as it does so. As it moves inward, the total circuit capacitance rises because the capacitance of the microphone increases. Increased total capacitance results in reduced frequency since resonant frequency is inversely proportional to the square root of the capacitance in accordance wih the equation:

$$F_r = \frac{1}{2\pi\sqrt{LC}}$$

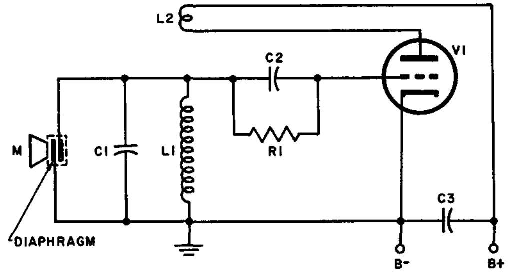


Fig. 3. Condenser microphone varies capacitance across oscillator tuned circuit to provide frequency modulation.

Likewise, when the diaphragm moves further from the position is occupies when at rest (away from the back plate) the total circuit capacitance decreases and the frequency rises. Figure 4 illustrates this pictorially. Compare the f-m wave thus created with that of Fig. 1 (4).

It is seen that the amount by which the frequency varies depends upon the loudness of the sound striking the microphone; the louder the sound, the more the frequency varies. With no sound reaching the microphone, the output of the oscillator is identical with that of the unmodulated wave with respect to frequency. As only tuned-circuit capacitance is being changed by the sound waves, the amplitude of the output wave is constant while only the frequency follows the pulses of the sound waves.

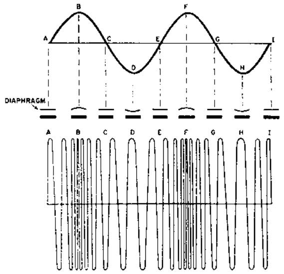


Fig. 4. A-f signal waveform, corresponding positions of diaphragm of Fig. 3, and effect on modulated signal waveform.

The simple circuits just discussed have been used for explanation purposes only and are not representative of actual methods of producing either a.m. or fm. In the case of the condenser microphone, however, the modulation method described, though crude and impractical as it stands, has been employed for generating fm by replacing the microphone with a *reactance tube* system. In the paragraphs that follow, this system is to be described; it will be shown that a vacuum tube circuit may be used to frequency-modulate a carrier by causing a sound wave to change the resonant frequency of an oscillator.

9. FM by Reactance-Tube Modulation

As a preliminary step in the analysis of reactance-tube modulation principles, the reader must be prepared to adopt a somewhat different viewpoint with respect to the constitution of capacitive and inductive circuits than the one to which he is probably accustomed.

It is relatively easy to show, both mathematically and experimentally, that the current flowing through an inductor connected to an a-c generator lags behind the applied voltage by 90 degrees (Fig. 5). This reaction is "felt" by the generator as it attempts to deliver power to the external circuit as a purely inductive load. Now the inductor is replaced by a "black box," the contents of which are unknown, and the generator is turned on once again. Suppose that the reaction from the "black box" is precisely the same as that provided by the inductor; that is, the current lags the voltage by 90 degrees. *Regardless of the contents of the "black box,"* the generator will still "feel" an inductive load and react in exactly the same manner as it did previously.

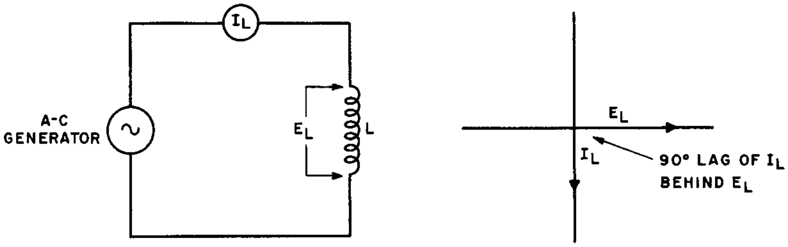


Fig. 5. Basic inductive circuit and vectorial relation of voltage and current.

From this point of view, then, any device or combination of devices that causes current to lag voltage by 90 degrees may be considered as an "inductor" with relation to its effect on other circuits; where such a combination produces a 90-degree current lead, it may be termed a "capacitor" even though true condensers may not be present at all; if our "black box" maintains an in-phase relationship between *E* and *I*, it constitutes a "resistor." With this in mind, reactance-tube modulation is much easier to follow.

Figure 6 illustrates the basic reactance tube arrangement. Two things should be noted:

a. *R1* and *C1* are connected across the oscillator tank circuit and thus *can* affect the frequency if it can be established that this series circuit acts either inductively or capacitively, or causes the tube to do so.

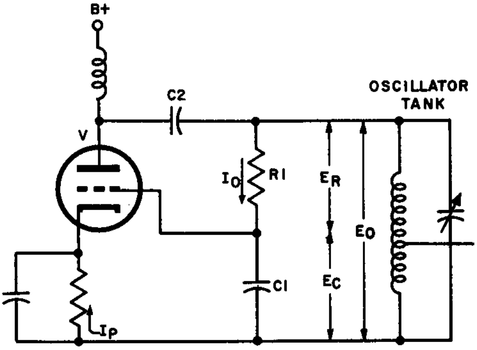


Fig. 6. Reactance tube circuit.

b. *R1* is a large resistor and *C1* is a relatively large capacitor for the frequency being used. This means that the resistance of *R1* is *much greater* than the capacitive reactance of *C1*. The circuit is, therefore, almost entirely resistive and the current (I_0) through the *R1C1* combination due to the voltage impressed across it by the oscillator tank circuit (E_0) may be considered to be in phase with E_0 (Fig. 7).

Since the current through a capacitor always leads the voltage applied across it, the potential E_c must lag behind the current through *C1* (I_0) by 90 degrees (Fig. 8). *C1* is terminated at one end by the grid of the vacuum tube and at the other end by the

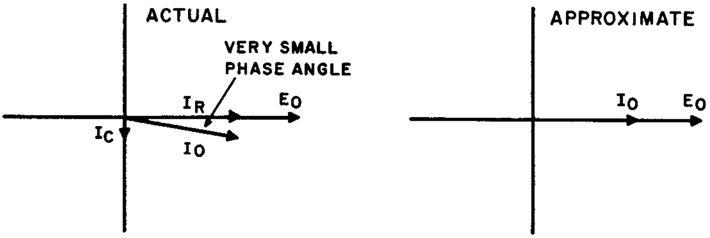
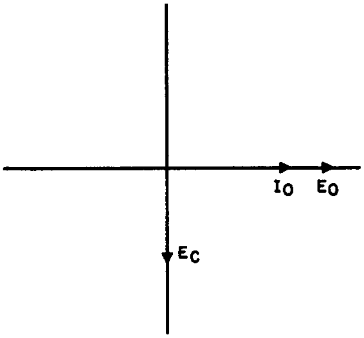


Fig. 7. *R1-C1 current voltage relations. Left exact; right equivalent when reactance of C is relatively small.*

cathode, so that the voltage developed across this capacitor really constitutes the applied grid voltage.

In a triode or any tube acting as a linear amplifier, the plate curent waveform is always identical with the grid voltage waveform and, in addition, is in phase with it. This is to be expected because a negative-going grid reacts on the plate current in a way such as to make the plate current negative-going as well; similarly, a positive-going grid causes a positive-going plate current. Thus, the plate current vector (I_p) may be added to the diagram in phase with the grid voltage E_c (Fig. 9).

Fig. 8. Relation of C1-voltage E_c to current I_0 and voltage E_0.



Examination of Fig. 9 now discloses that the plate current (I_p) of the tube lags behind the applied plate voltage (E_o) from the oscillator tank circuit by 90 degrees. From the oscillator's "point-of-view," then, it looks as if an inductor had been connected across the tank coil, because in an inductor this same relationship of E and I occurs. Therefore, the reactance tube (V), together with its associated resistive and capacitive components, *is an inductor* insofar as the oscillator circuit is concerned, and does contribute to the tuning or resonant frequency of the oscillator.

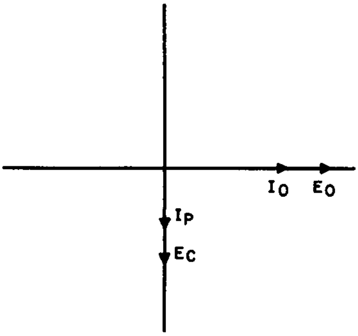


Fig. 9. Diagram of Fig. 8 with plate current I_p added.

If nothing is added to the circuit of Fig. 6, the reactance tube network simply adds inductance in parallel with the tank inductance. Inductors in parallel have a joint inductance *less* than that of either of the two components, just like paralleled resistors. Hence, the reactance tube circuit effectively raises the frequency of the oscillator above that of the oscillator tank circuit alone. So far, the reactance tube has simply added inductance to the oscillator, has lifted its frequency somewhat higher than it was before the tube was added, but has not accomplished frequency modulation of the oscillator output. To do this, an audio voltage introduced into the reactance tube circuit must cause the "inductance" to vary in step with the causative sound.

Let us add a second grid to the triode of Fig. 6 (see Fig. 10) and apply a fluctuating audio voltage to it. As the second grid is raised and lowered in potential, it causes the plate current (I_p) of the reactance tube to vary according to the same waveform. Thus, the current through the "inductor" formed by the reactance tube and its parts fluctuates in step with the audio signal developed at the input to the second grid.

The oscillatory voltage applied across this "inductor" has not changed, but the current flowing through it is being modified; the conclusion must be drawn that the "inductive reactance" of the tube circuit must be changing. That is,

$$I_o = E_0/X_{rt}$$

where X_{rt} is the apparent "inductive reactance" of the reactance tube circuit. Any variation in X_{rt} must be caused by fluctuations in L_{rt} (the apparent inductance of the reactance tube circuit). Now, the frequency of oscillation is approximately given by the equation:

$$f = \frac{1}{2\pi\sqrt{LC}}$$

L_{rt} forms a part of the total inductance of the circuit (L in the equation above) and, if this changes as has been indicated in the foregoing discussion, the frequency of oscillation must change in accordance. This is frequency modulation.

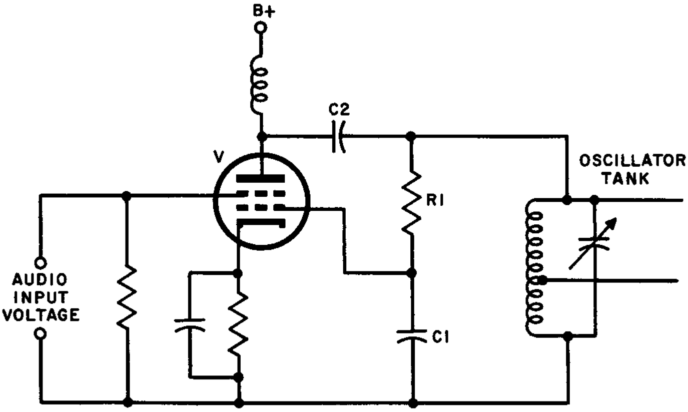


Fig. 10. Diagram of reactance tube controlling oscillator frequency by providing inductance across oscillator tank.

A reactance tube may be connected to produce an apparent "capacitance," rather than inductance with similar effect. (Figure 11 shows an explanatory sketch of such an arrangement.) In this connection, the capacitive reactance of *C1* is made much greater than the resistance of *R1* so that the *C1R1* combination is largely capacitive. This reverses the phase relationships of the previous case in that the grid voltage is now E_R. Because the current causing E_R is capacitive, it leads the initiating voltage from the oscillator tank circuit, thus producing a leading plate current I_p. The vectors associated with this condition should be drawn by the reader as an exercise.

10. Frequency Multiplication

Because fm is used exclusively at relatively high radio frequencies, it is both desirable and convenient to start the chain with an oscillator of low frequency deviated to a much smaller extent than the final, desired frequency swing of the transmitter is to be. This system has several important advantages:

a. Low frequency oscillators may be designed to be much more stable frequency-wise than high frequency oscillators.

b. It is comparatively easy to obtain small deviations. This requires less audio power and the deviations may be made much more linear if they are small.

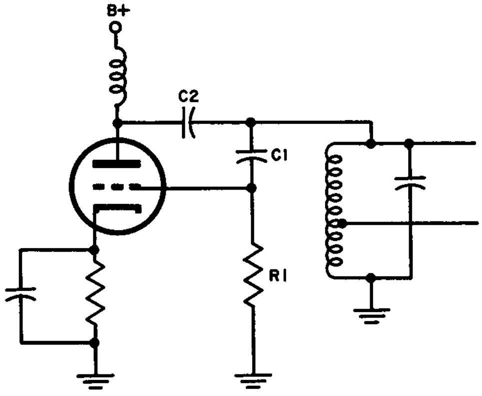


Fig. 11. Diagram of reactance tube that simulates capacitance.

To help clarify the operation in this process, consider BMT Radio Patrol Station KEA-911 in New York City. This station operates on a frequency of 44.34 mc. In designing the transmitter, an oscillator frequency of 2771.25 kc was chosen. To obtain the required output frequency, the oscillator signal is doubled four times in four separate doubler stages (Fig. 12). Thus:

$$2771.25 \times 2 \times 2 \times 2 \times 2 = 44{,}340 \text{ kc} = 44.34 \text{ mc}$$

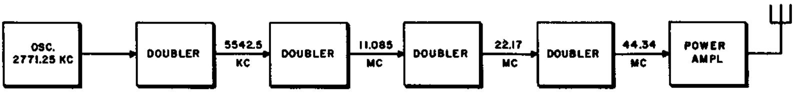


Fig. 12. Block diagram of crystal-controlled r-f section employing frequency doublers.

According to the license issued to this station, it has been assigned a maximum bandwidth of 40 kc with a maximum frequency swing of 30 kc (Deviation = ±15 kc). To arrive at this deviation, the frequency swing of the modulated oscillator need only be a total of 1.875 kc. The deviation in this case, of course, is half this value or ± 937.5 cycles per second. As the frequency of the oscillator is multiplied in the doubler stages, the absolute deviations are multiplied to the same degree so that it is possible to start with a low oscillator resting frequency and small deviations, and finish with a high output frequency having the assigned maximum deviation. The simple case given below should help explain this idea:

Given: Oscillator center frequency = 1000 kc
Maximum oscillator deviation = ± 1 kc
Frequency multiplication = 16 × by 4 doublers
Then:
Before multiplication, highest frequency to which oscillator is deviated = 1001 kc
Lowest frequency to which oscillator is deviated = 999 kc
Total swing = 2.0 kc

Now:
After multiplication 16 ×
Highest frequency to which signal is deviated (16 × 1001) = 16,016 kc
Lowest frequency to which signal is deviated (16 × 999) = 15,984 kc
Total swing = 32.0 kc

These figures illustrate that a 16 × frequency multiplication of the resting frequency also provides a 16 × multiplication of the total frequency swing from oscillator to output.

11. Frequency Multipliers

Frequency multipliers are radio frequency amplifiers adjusted to produce severe distortion of the input r-f waveform so as to generate a signal rich in harmonics. The circuit of a simplified doubler is given in Fig. 13. It is to be noted that grid-leak bias is used. Resistor R_g is chosen to provide about five to six times cut-off bias, placing the amplifier well into the Class C region. The tank circuit consisting of *L1* and *C1* is tuned to the frequency of the preceding stage; the second harmonic of this frequency appears

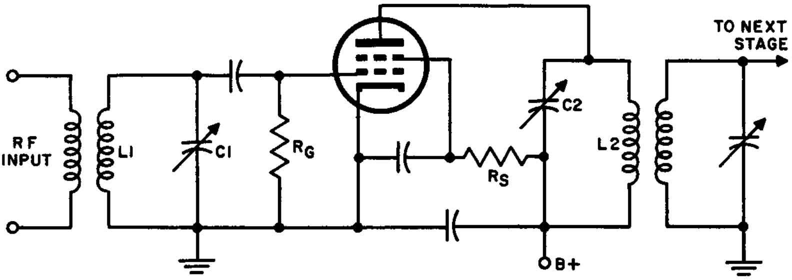


Fig. 13. Schematic diagram of frequency multiplier.

together with the fundamental frequency and other higher-order harmonics at the output of the amplifier tube but, since the *L2C2* tank circuit is tuned to twice the fundamental frequency, only this doubled frequency appears at the input to the next stage in significant strength.

The power output obtainable from a multiplier decreases as the attempt is made to use higher order harmonics. A properly adjusted doubler is capable of about 65 per cent efficiency, while a quadrupler cannot be expected to provide better than about 30 per cent efficiency. This accounts for the use of cascaded doublers in the f-m transmitter described in the preceding paragraph rather than triplers or quadruplers.

Figure 14 shows why Class C operation of multipliers results in the kind of distortion needed to generate the desired harmonics. The input waveform is sinusoidal. If no distortion were present—

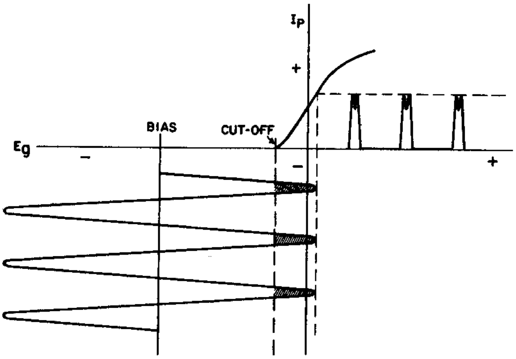


Fig. 14. How class C operation distorts waveform and thus generates harmonics.

of a condition found in Class A amplifiers—the output waveform would also be sinusoidal. In the case of the Class C amplifier, the plate current of the multiplier tube flows in short pulses to produce an overall current waveshape that does not resemble the input voltage; these sharp plate current peaks consist of a component of the fundamental frequency plus large percentages of the second, third, fourth, and higher order harmonic components, a desirable result in the light of the function of this kind of stage.

12. Frequency Stabilization

One of the essential specifications for frequency modulation by the reactance tube method is that the carrier frequency must return to the same mean value each time the modulation stops; this is the frequency assigned by the FCC as the center frequency of the transmitter. This method of frequency modulation inherently requires an oscillator whose frequency can be easily changed (deviated) by the modulator, hence a highly stable oscillator such as a crystal-controlled type cannot be used. At the same time, some control network *must* be associated with the transmitter which assures "on-frequency" operation without the need for constant attention.

By connecting a crystal-control circuit as illustrated in Fig. 15, a degree of stability that closely approximates that of the crystal-controlled oscillator may be achieved. The circuit functions as follows: (Frequencies indicated have been chosen in round numbers to facilitate the explanation.) The frequency of the transmitter oscillator, 10 mc, is multiplied by a series of doubler stages up to the output frequency of 100 mc. A part of the output is fed back

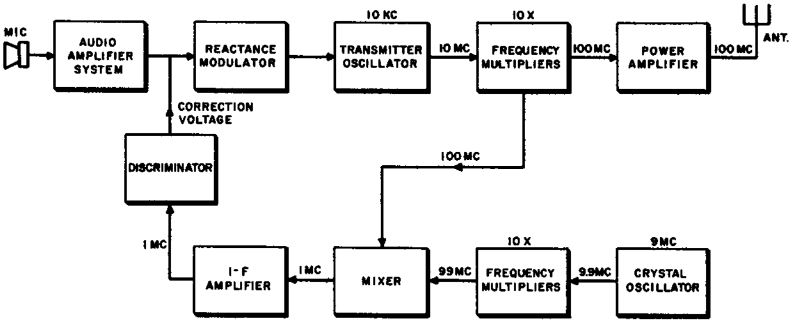


Fig. 15. Block diagram of f-m transmitter whose center frequency is stabilized by an afc system linked to a crystal-controlled r-f channel.

to a mixer stage in the lower row. In another section of the transmitter, a temperature controlled crystal oscillator generates a 9.9-mc signal, which is amplified and multiplied by a second set of doublers to 99 mc.

The two signals, at 100 mc and 99 mc respectively, are then combined in the mixer, the intermediate frequency output of 1 mc being transferred to an amplifier tuned to this signal for the purpose of building up its amplitude.

The discriminator is a special type of vacuum tube stage that produces a fixed d-c output voltage only as long as it receives a radio-frequency signal to which it is resonant. This fixed voltage is constantly fed to the grid of the reactance tube and serves as a reference potential which establishes the proper operating frequency of the transmitter oscillator.

The frequency of the crystal oscillator is presumed to be absolutely constant within tolerance limits prescribed by the FCC; but the transmitter oscillator may drift—an occurrence which the system must prevent. Suppose that its frequency rises, even slightly. Instantly, the intermediate frequency output of the mixer stage also rises, feeding an off-resonance signal to the discriminator. This stage now changes its d-c output voltage to the reactance tube to a new value. The new voltage causes the latter to change its effect on the transmitter oscillator in a direction which automatically restores the correct frequency. This system is precisely the same, in principle, as the automatic frequency control (afc) networks used in many television receivers.

13. Review Questions

(1) Explain why a simple modulation system for fm, like the use of a condenser microphone across the tuned circuit of the oscillator, would be unsatisfactory for any commercial f-m service.

(2) Does the use of the term "reactance" in connection with a reactance tube modulator imply that the tube may behave as either a capacitor or an inductor? Explain.

(3) A resistor of 100 ohms, a capacitor of 350 $\mu\mu$f, and an inductor of 40 microhenries are connected in series across a source of a-c signal. If the applied voltage is 1 volt, find the frequency of the signal that would cause 10 ma to flow through the series network.

(4) What phase relationship between current and voltage must exist in the load of an electron tube if the tube is to "see" the load as a pure inductor? As a pure capacitor? As a pure resistor?

(5) A "black box" is connected across a tuned circuit which generates an oscillatory voltage. The box contains a capacitor and a resistor in series. Will the frequency after this circuit is connected to the tank be higher or lower than it was before connection? Explain.

(6) Explain how the phase-shifting network in a reactance tube causes the tube to act as a capacitor. Use vectors.

(7) What are the advantages in starting an f-m transmitter chain with a low frequency rather than a high frequency oscillator?

(8) The center frequency of the oscillator of an f-m transmitter is 1.5 mc. The oscillator is followed by a tripler and two doublers. What is the output center frequency?

(9) In the transmitter described in question 8, what oscillator deviation is necessary to produce an output frequency swing of 30 kc?

(10) Explain why class C operation of frequency multipliers is essential. What would happen if Class A frequency multiplier operation were attempted?

Chapter 3

FROM PM TO FM

14. Advantage of Phase Modulation

Because intelligence may be conveyed from transmitter to receiver by causing the intelligence to vary one of the three wave characteristics—amplitude, frequency, or phase—it follows that communication should be possible by affecting the phase of a radio-frequency carrier in step with usual audio frequency variations produced by a modulator system.

This suggests a significant improvement over direct frequency modulation. As will be shown, phase modulation (pm) produces fm so that the end result is the same. But, where the use of a crystal controlled oscillator is not feasible when frequency modulating a carrier directly, the natural high frequency stability offered by crystal control *is* possible with pm.

Although an automatic frequency control system as described in paragraph 11 suffices to maintain stability standards for some f-m services where the requirements are not extremely rigorous, it does not take full advantage of the possibilities inherent in direct crystal-control. Herein lies the advantage of pm. for, with this system of ultimately achieving fm, direct crystal control is not only possible but is relatively easy to realize.

15. The Meaning of Carrier Phase-Shift

To help visualize the significance of a phase shift in a voltage or current, consider a radio frequency carrier voltage as shown in

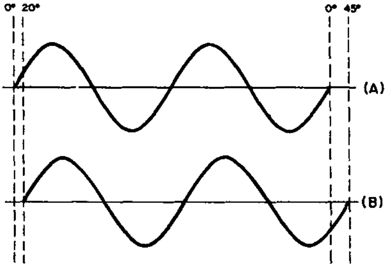


Fig. 16. (a) R-f waveform, (b) Same waveform undergoing phase shift.

Fig. 16. In (a) a certain carrier voltage is represented as starting out at an arbitrary zero-degree position, completing two cycles, and finishing at the zero-degree position. The same carrier is shown in (b) as starting out *20 degrees later than it originally did* and completing two cycles, the end of the last cycle occurring 45 degrees later than previously. It must be emphasized that these two diagrams portray the same carrier in two different conditions, *not* two different carriers.

It is to be assumed that some electrical arrangement added to the oscillator circuit has succeeded in shifting the phase of the carrier with respect to its resting phase by different amounts over two cycles (20 degrees and 40 degrees). Such a phase shift is tantamount to a change in frequency as is evident merely from inspection of the waveforms; yet it will be demonstrated that this action is possible even when the carrier frequency is held rigidly constant by a thermostatically controlled crystal. In short, although phase modulation results in frequency modulation, the frequency of the carrier is not swung or deviated in the same sense as it is when the

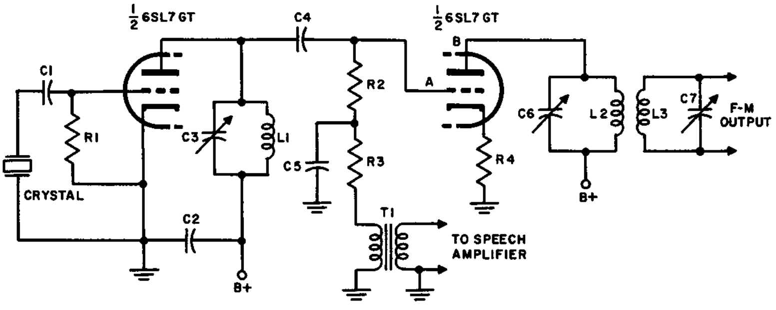


Fig. 17. Schematic diagram of a phase modulator.

fm is applied directly as in the reactance modulation system. In other words, the same frequency excursions are produced; but they are not caused by variations of oscillator tuned-circuit component values.

16. Fundamental Phase Modulator

The schematic diagram of a phase modulator system is given in Fig. 17. A crystal controlled oscillator provides a signal of fundamental frequency low enough to be subsequently multiplied by the frequency multiplier system as in the reactance tube arrangement. The output of the oscillator is applied to the control grid of the second section of the 6SL7GT at point *A* in the schematic. *R2, R3,* and the primary winding of *T1* provide the grid return path for this section. *T1* is an audio coupling transformer of the type designed for grid modulation. The combination formed by *L2* and *C6* is the plate tank circuit of the r-f amplifier section of the 6SL7GT and it is across this tuned circuit that the frequency-modulated output signal is to appear, to be fed later to the frequency multipliers.

The output voltage which appears across *L2-C6* is the vector resultant of two out-of-phase voltages impressed across the tank circuit via different paths. First, as is expected, the crystal oscillator output arriving at the grid (*A*) through coupling capacitor *C4* is amplified through normal tube action and is applied as a radio-frequency voltage to *L2C6,* which is tuned to the crystal frequency. The tank combination is to be considered, then, as a purely resistive load at this frequency, across which an output r-f voltage is generated by the amplification of the tube. Second, voltage of exactly the same frequency from the same source (the crystal oscillator) is developed across *L2-C6* by capacitive coupling through the interelectrode (grid-plate) capacitance *A* to *B*. This transfer is not affected by the usual tube action; that is, the path *A* to *B* is to be considered as an ordinary capacitor having no relation to the normal tube behavior.

If the applied grid voltage is represented by the reference vector E_g (Fig. 18a), then the amplified voltage resulting from tube action must be shown as a second vector 180 degrees out-of-phase with E_g (E_a), since the output voltage of a tube working into a resistive load is always of opposite phase from the input grid voltage (Fig. 18b).

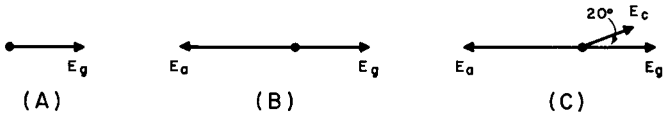


Fig. 18. Phase relations in Fig. 17.

The second output voltage developed across *L2-C6* leads the grid voltage by an angle determined by the capacitive reactance of the grid-to-plate capacitance from *A* to *B*. This conclusion is arrived at as follows: the r-f current flowing from *A* to *B* must lead the applied voltage E_g as a result of normal capacitor action; the angle of lead depends upon the ratio of the capacitive reactance of the grid-plate capacitor and the resistance of the resonant tank circuit *L2-C6*. Assume a lead angle of 20 degrees for this capacitive current. The voltage that appears across the tuned circuit must be in phase with this current because, here again, the *L2-C6* combination is resonant and acts as a purely resistive load. Thus, the capacitively-transferred voltage, E_c, must lead the applied grid voltage by 20 degrees as illustrated in Fig. 18c.

Because both of these voltages (E_a and E_c) are developed simultaneously across *L2-C6*, the total voltage across the tuned circuit is the vector sum of these individual voltages. The resultant is obtained by means of the usual vector parallelogram as shown in Fig. 19-a; the resultant is labeled E_0 and is depicted in the figure as lagging E_a by about 30 degrees.

As long as the amplified voltage E_a remains of constant amplitude, this relationship of phase angles is maintained. The picture changes, however, when the amplitude of E_a is caused to change by the application of a modulation voltage on the grid of the amplifier by the speech system. As the microphone of the modulator is

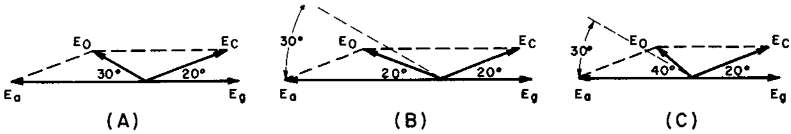


Fig. 19. Vector phase relations for Fig. 17, showing how phase modulation results.

actuated by sound input, the voltage applied to the control grid of the amplifier section of the 6SL7GT varies in synchronism with the sound. This causes *both* E_c and E_a to change in magnitude, but the variation in E_a is greater than that in E_c to the extent that the gain of the tube is greater than unity. Hence, E_a effectively increases and decreases with the swings of the modulating signal, producing a changing resultant as illustrated in Fig. 19 (b) and 19 (c). In Fig. 19 (b), the resultant output voltage, E_o, is shown as differing in phase from E_a by 20 degrees due to the elongation of the E_a vector, which in turn is due to a swing of the audio voltage in one direction. On the next half-cycle, when the audio voltage reaches its peak in the opposite direction, E_a shortens, causing a phase shift between E_o and E_a, so that the new phase angle is 40 degrees (Figure 19c). It is evident that the swinging audio voltage produces an output voltage which varies in phase over a total angle of about 20 degrees in this illustration. This is phase modulation.

The resting phase of this system might be defined as 150 degrees, the angle between the applied voltage E_g and the resultant voltage E_o. On positive swings of the audio voltage, E_a increases so that the new phase angle is 160 degrees; on negative swings E_a is reduced, changing E_o to a new phase of 140 degrees. These variations in the phase of the resultant output voltage are equivalent to fm as described in paragraph 14—fm that has been achieved without directly swinging the frequency of the r-f carrier but by varying the phase of the carrier output voltage with respect to its resting phase.

17. Effect upon PM of Audio Modulating Frequency

It has been shown that a modulating voltage impressed upon the control grid of the phase-shift amplifier gives rise to phase modulation. A direct proportionality between the amount of phase modulation thus produced and the magnitude of the modulating voltage is implicit in this relationship for this reason: The extent to which E_a changes depends directly upon the peak amplitude of the adio signal—the larger the signal, the greater the variation of E_a; consequently, since the phase of E_o, the output voltage, is proportional to the length of the E_a vector, the amount of phase-shift must be proportional to the audio voltage. This concept is common in fm as well as in pm.

The relationship between the audio *frequency* and the amount of phase modulation is not so obvious, however. It can best be understood by considering the sequence of events from the following point of view:

One complete phase swing (from + 10 degrees to — 10 degrees as in Fig. 19) occurs during one complete audio cycle. If the modulating frequency is, say, 1000 cycles per second, this swing must occur in a period of .001 second. Now assume that the audio frequency rises to 5000 cycles per second; the complete shift of phase now must take place in a much shorter period of time, .0002 second. A more rapid phase change means a higher rate of frequency change—hence more fm. This may be stated in the form of a simple but useful equation:

$\Delta F = f \Delta \theta$ where F = carrier frequency
f = modulating frequency
θ = phase angle in radians

Using this equation, it is easy to find out just how much fm can be expected from a given change of phase angle at a specific modulation frequency. Assume that, in the example used in the previous paragraph, the modulating frequency was 5000 cps and that the total phase shift ($\Delta \theta$) was ± 20 degrees.

The amount of fm would then be:

$\Delta F = 5000 \times .35$ radians (one radian = 57.3 degrees, so that 20 degrees = 20/57.3 radians = .35 radians, approximately) thus, ΔF = 1750 cps, or a total carrier frequency swing of 1750 cps.

Retaining the same phase angle change of .35 radians, should the audio modulating frequency now drop to 1000 cps, the frequency swing would fall to 350 cps. Thus, *even though the audio level did not change,* the amount of fm did change seriously as a result of the variation in audio *frequency*. This condition is obviously intolerable in a system that requires frequency excursions that are proportional to changes in audio voltage, not to changes in audio frequency. Thus, fm derived from pm has introduced into it a unique kind of distortion that must be canceled out before the system can be used.

18. Cancellation of the Frequency Factor

In the system illustrated in Fig. 17, *R2, R3,* and *C5* form a frequency compensation network designed to neutralize the effect that the frequency of the modulation voltage has upon the amount of frequency modulation as described in paragraph 16. The audio voltage output from the secondary of transformer *T1* may be considered as being applied across the series circuit consisting of *R3* and C5. At the higher audio frequencies, the voltage developed across *C5* is substantially less than it is at lower frequencies because the reactance of this capacitor decreases as the frequency rises. On the other hand, the audio voltage applied to the grid of the second section of the 6SL7GT as modulation is essentially the voltage which appears across *C5*. Consequently, as the modulation frequency increases, less audio voltage is fed to the control grid; as the frequency goes down, the grid voltage rises. With the proper choice of component values, this inverse frequency relationship completely cancels the frequency factor in the equation $\Delta F = f \Delta \theta$ so that the frequency swings of the r-f output from the system are exactly the same as they would be if the modulation were applied by the direct f-m method rather than the p-m method.

This action may be viewed in a slightly different manner: the value of *C5* is sufficiently large so that this capactitor has negligible reactance to the radio frequency generated by the crystal oscillator. Hence practically all of the r-f potential developed by the oscillator appears as a voltage drop across *R2*. *C5,* however, does carry a voltage that varies in magnitude in accordance with the audio frequency coming from the transformer *T1* as described above. *C5* is effectively in series with *R2* so that this audio voltage adds to and subtracts from the r-f voltage across the resistor, changing the bias on the 6SL7GT to accomplish modulation.

19. P-M Transmitters

The similarities and differences between a direct f-m transmitter such as that pictured in the block diagram of Fig. 15 and the p-m transmitter illustrated in Fig. 20 are evident upon inspection.

The p-m transmitter starts with a temperature-controlled crystal oscillator, which feeds the phase modulator (Fig. 17). In this transmitter, the final output resting frequency is to be within the f-m broadcast band—96 mc in this case. Assume that the conditions are arranged to produce a final frequency swing of 144 kc (deviation ± 72 kc). Thus the frequency of the cyrstal oscillator is 4 mc and the deviation produced by phase modulation starts at ± 3 kc. A tripler stage and three frequency doublers yield a total frequency multiplication of 24 times for both the oscillator frequency and the deviation as previously explained, yielding an output signal in this case of 96 mc with a deviation of ± 72 kc, as shown in the block diagram.

Significantly absent is the frequency controlling feedback network of Fig. 17. The need for this has been eliminated, of course, by the use of a crystal driver and phase modulation rather than direct fm.

20. Side Bands in FM

One of the factors that contribute to the spacing between a-m stations as specified by the FCC is the presence of side frequencies in the case of single tone modulation, or sidebands when the modulation is the result of speech or music. Side-frequency components occupy spectrum space; it will be remembered by the reader than a 1000-kc carrier amplitude modulated by a 1000-cps pure tone develops two side frequency components, one at 1001 kc and the other at 999 kc. Such a signal uses 2 kc of the spectrum. The situation becomes worse as the modulating frequency rises; for an audio tone of 5000 cps, for example, the side frequency components are spaced 5 kc above and below the carrier frequency, resulting in a bandwidth of 10 kc of occupied space. The bandwidth of an a-m

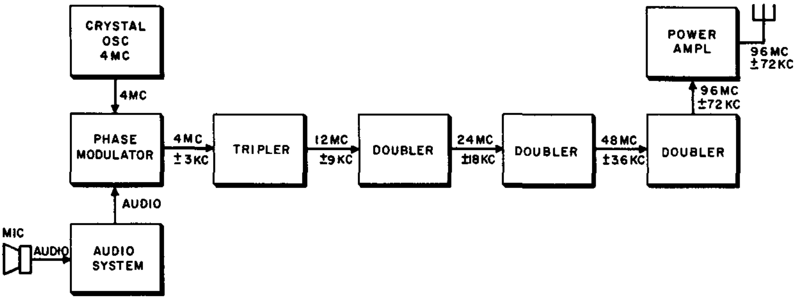


Fig. 20. Block diagram of phase-modulated transmitter.

station is thus given by: Bandwidth = 2 × highest modulating frequency.

It may be shown that, in general, a frequency modulated signal possesses not only the first order sidebands described above, but also second, third, fourth, etc., order sidebands that may cause interference in adjacent channels. In Fig. 21, the signal emitted by an f-m station using a low modulation index—the carrier plus the side frequency components—is illustrated; the pair of lines labeled *A* indicates the position and magnitude of the first order side-frequency components, those shown as *B* the second order side-frequency components, and so forth to those of the fourth order. It should be noted that each pair of side frequencies is separated from its neighbors by an amount equal to the audio modulating frequency that produces the side effects.

The question immediately arises of the bandwidth required for an f-m station if its sidebands extend out in higher and higher orders up and down the spectrum. With a modulation index of 0.5 or less, the power radiated in the form of side frequencies above the first order is negligible and need not be considered as a source of adjacent channel interference. On the other hand, the maximum modulation index (deviation ratio) for f-m broadcast stations in this country is five. When a modulation index of this magnitude is reached, the number of significant sidebands rises to the *eighth order*—eight sidebands on each side of the carrier for a total of *sixteen!* To help insure interference-free reception, the FCC has established the "guard band" arrangement, in which a stretch of 25 kc is inserted between f-m broadcast channels; this is a normally inactive region that serves simply to space the channels further apart.

When the modulation index is five, the bandwidth occupied by the transmitting station is 240 kc. How can this be reconciled

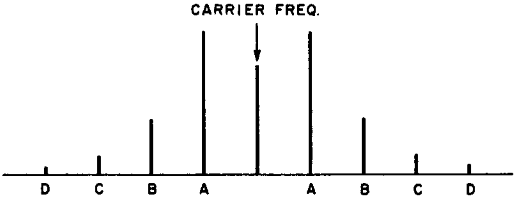


Fig. 21. Typical frequency components of an f-m wave.

with the allocation of f-m broadcast station frequencies, the channel width being 200 kc maximum? This question has three answers:

a. Very seldom does an audio frequency component of 15,000 cps have sufficient intensity to produce the full deviation of ± 75 kc and, unless it does so, the higher order sidebands shrink to practical insignificance.

b. Practical operating conditions have shown that the ± 25 kc guard band is sufficiently wide to absorb the higher order sidebands that do appear when the modulation index hits its normal peaks.

c. To avoid the risk of adjacent channel interference due to occasional bursts of higher order sidebands, the FCC follows the practice of assigning adjacent frequencies to f-m stations widely separated in geographical location.

21. Review Questions

(1) Differentiate between frequency modulation and phase modulation.

(2) Draw a schematic diagram of a typical phase modulator system.

(3) Explain the operation of the phase modulator schematic of question 2. Use vectors to illustrate the pertinent voltage relationships.

(4) Explain the relationship between the audio frequency modulating voltages and the amount of phase modulation obtained.

(5) State a mathematical equation used to determine the amount of fm that can be expected from a given change of phase angle at a specific modulation frequency.

(6) Explain how a frequency compensation network can be designed to neutralize the effect that the frequency of the modulation voltage has upon the amount of frequency modulation.

(7) List the similarities and differences between a direct f-m transmitter and a p-m transmitter. Illustrate by means of block diagrams.

(8) With a modulation index of five, the bandwidth occupied by the transmitting station is 240 kc. Reconcile this with the allocation of f-m broadcast channel widths of 200 kc maximum.

(9) Explain the purpose of "guard bands."

(10) What is meant by a "low modulation index"?

Chapter 4

THE PROPAGATION AND RECEPTION OF FM

22. General Information

Radio signals of all types, regardless of frequency or kind of modulation, are propagated by means of a radiated magnetic and electric field.

Radio frequency energy fed to a transmitting antenna produces electromagnetic and electric fields having the same frequency as the original energy. A substantial portion of the energy leaves the antenna in the form of energy usually termed—rather broadly—as *waves*. Once the waves have leaped into space, the paths they follow to the receiver are determined chiefly by their frequencies.

23. The Ground Wave

The ground wave is that portion of the total radiation which is directly affected by the electrical characteristics of the earth. Basically, there are three distinct classes of ground waves.

a. *The surface wave* is radiated energy that is actually guided by the earth and follows a curved path, sometimes over a considerable distance around the arc formed by the earth's surface. This wave is useful only at very low radio frequencies because the earth absorbs a significant amount of power from it when the frequency is high. Below about 500 kc, reliable communication may be maintained over distances up to 900 or 1000 miles exclusively through the medium of the surface wave. Regular a-m broadcasts cover a good portion of their service areas in this way.

b. *The direct wave* is, just as its name implies, the energy that travels in a more or less straight line from the transmitting antenna to the receiving system. Sometimes this direct wave is diffracted by the atmosphere so that it follows a somewhat curved path around the surface of the earth. We may call such a wave the "diffracted direct wave" or the "diffracted space wave."

c. *The earth reflected wave* is that part of the total energy that is directed downward and reflects from the surface of the earth to reach the receiving antenna along a path different from that of the direct wave. The direct wave and the reflected wave, when considered together, are often referred to as the "space wave."

At the frequencies used by f-m services—approximately 30 mc to 300 mc—the surface wave is of negligible importance; thus, f-m service areas are covered almost entirely by the space wave and diffracted space wave. As will be indicated in the next paragraph, the *sky wave,* which is so important for long distance communication, has no practical significance for fm.

24. The Sky Wave

Energy radiated upward by transmitting antennas may or may not return to the earth. The factors that determine the fate of a given wave are principally the frequency of the radiated energy, the condition of the upper layers of the atmosphere, sunspot activity, and a host of almost unpredictable daily and seasonal changes in humidity, atmospheric pressure, and atmospheric temperature. Up to about 30 mc, radio waves are returned to the ground by refraction that occurs in the uppermost layers of the atmosphere—in inosphere. This effect is seldom noticed above 30 mc.

Practically all f-m stations operate at frequencies higher than 30 mc. This means that reliance must be placed upon the space waves for f-m communication. The coverage of f-m stations is thus confined to relatively short distances, a condition that might appear to be disadvantageous at first glance—as indeed it is in some respects—but which is really beneficial when one considers frequency allocations throughout the country, as previously explained.

25. Line of Sight Transmissions

Although a small amount of bending is produced even in high frequency waves by the refraction of the tropospheric air (layer in contact with the earth), the constancy of this effect cannot be counted upon unless space-wave diffraction occurs. For broadcast, communication, and emergency f-m services, unreliability cannot be tolerated so, for practical purposes, the distance coverable by a high-frequency f-m emission can be considered to be "line-of-sight" in most cases.

By substituting the known heights of the receiving antenna (H_r) and the transmitting antenna (H_t) in the equation given below, the distance that can be covered by a line-of-sight transmission may be easily calculated.

$$D = 1.23\ (\sqrt{H_t} + \sqrt{H_r})$$

To illustrate the application of this equation,* consider the following problem: what distance will a signal cover by a line-of-sight route if the transmitting antenna is 1000 feet high and the receiving antenna is 10 feet high?

$$D = 1.23\ (\sqrt{1000} + \sqrt{10})$$
$$D = 1.23\ (31.6 + 3.16)$$
$$D = 42.8 \text{ miles}$$

By using the same equation, it may be shown that an increase in the height of the receiving antenna by 90 feet raises the range from 42.8 miles to 51.8 miles, a gain of 9 miles. If the transmitting antenna is lifted by the same distance, that is, 90 feet, the new coverage is only 44.5 miles, a net increase of 1.7 miles. This rather unexpected result is obtained because of the square-root relationship between distances and antenna height, a small change in the lesser of the two heights having a greater net effect upon the distance than the same modification of the higher antenna.

It is clear from the foregoing considerations that the receiving antenna height is a comparatively important factor of the receiving system and cannot be ignored if maximum satisfaction is to be gained from the f-m system.

26. The F-M Antenna

From the point of view of cost and ease of installation, the simplest satisfactory f-m antenna is a half-wave dipole (Fig. 22). It consists of two colinear quarter-wave rods separated from each

---

*Note: Normal refraction usually adds about 15 per cent to this distance. With high powers and directional antennas even greater distances (far in excess of line-of-sight) are possible.

other by $\frac{3}{4}$ to $1\frac{1}{2}$ inches at the center, an arrangement that provides a radiation resistance of approximately 72 ohms at the frequency for which it is cut.

The correct overall length of the dipole may be found from the following equation:

$$\text{Length in feet} = \frac{492 \times 0.94}{\text{frequency (mc)}} \text{ (final form of equation)}$$

This equation is derived from the general relation:

$$\text{wavelength} = \frac{\text{velocity of wave}}{\text{frequency}}$$

To illustrate the use of the equation, suppose an antenna is to be cut for the broadcast range from 88 mc to 108 mc. The center of this range is 98 mc. Substituting this number for frequency we have:

$$\text{Length} = \frac{492 \times .94}{98} = 4.72 \text{ feet}$$

$$\text{Length} = 4 \text{ feet, } 8.6 \text{ inches}$$

Thus, each quarter-wave rod should be cut close to 2 feet, 4-$\frac{1}{2}$ inches.

Transmitting antennas are usually horizontally polarized; this means that the receiving antenna should likewise be set up with its quarter-wave rods placed horizontally to obtain maximum induction from the passing wave.

A half-wave dipole has definite directional characteristics, because it responds more strongly to signals that arrive "broadside" to the antenna than those which come from directions off the ends

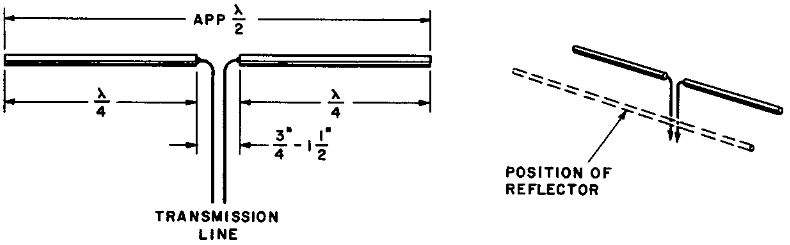


Fig. 22. Dipole antenna. Dash lines show reflector position.

of the rods. If the desired signal is from the north and strong interference of any kind is received from the south, there is a good likelihood that the latter will partially or entirely mask the desired signal.

To improve the "front-to-back" ratio (strengthening of forward response and weakening of response to signals from the opposite direction), a reflector may be added to the half-wave dipole as shown by the dash-lines in Fig. 22. This also sharpens forward directivity.

27. Transmission Lines

Many types of transmission lines are commercially available, among which two are especially favored: twin-lead parallel lines and flexible coaxial cable. (See fig. 23.) As in all power transfer systems, every effort should be extended to match impedances at both the antenna end and the receiver end. The average dipole without reflector has a center impedance of from 72 to 100 ohms; if 72-ohm coaxial cable or 72-ohm twin lead is used with such an antenna, care should be taken to make certain that the receiver input is designed to accept the 72-ohm line without much of an impedance mismatch, otherwise standing waves with their attending losses will be present on the line.

The same precautions must be observed in matching the 300-ohm line to the receiver. In this case, a special type of dipole known as the folded dipole is common. Such a dipole, consisting of two closely spaced half-wave lengths connected together at their ends, has a center impedance of 300 ohms. In areas where signal strength is high, none of these considerations is particularly critical; in fringe areas, however, even tiny losses may reduce the signal-to-noise ratio so that adherance to good installation practice will pay off in the form of greatly improved performance.

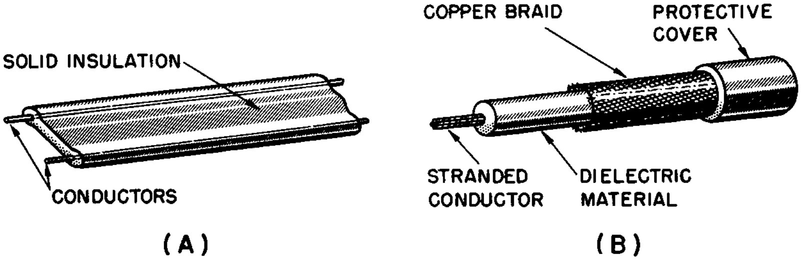


Fig. 23. (a) Twin-lead transmission line, (b) Coaxial cable.

28. Reduction of Interference

To one who has listened and enjoyed f-m broadcasts it is apparent that background noise and other types of interference are substantially less annoying with a properly operating f-m receiver than with an equivalent a-m receiver. The term "interference," as it is used here, includes atmospherics (statics), electrical noise, internal receiver noise, crosstalk from stations on adjacent channels, and beatnotes from transmitters operating near the frequency of the desired signal.

Internal receiver noise originates mainly from two effects, known as *thermal agitation* and the *shot effect.* All matter is in a state of random, haphazard molecular motion; free electrons in conductors, whether or not they are under the influence of an impressed emf, undergo the same kind of agitation and, since electrons in motion constitute an electric current, small, ever-varying emf's appear and disappear constantly. This is thermal agitation. The shot effect results from the constitution of the electron stream moving between the cathode and plate of the electron tubes in the receiver. If this stream were a homogeneous fluid such as is found in water pipes there would be no shot effect; but the fact is that the electron stream is composed of discrete particles, which do not arrive at the plate in uniform bunches but rather are subject to statistical variations that cause the number of electrons reaching the plate at one instant to be different from the number arriving at the next instant in time. This "unevenness" of flow causes tiny noise voltages, which are amplified in radio receivers to the point where they may become annoying.

When the two systems are compared under the same conditions, i.e., with a given carrier frequency and a modulation frequency of 15,000 cps, the noise level of the f-m receiver is more than 18 db lower than that of the a-m receiver. This figure is postulated on a deviation ratio of five. As the deviation ratio is decreased, the superiority of the f-m receiver is not as great but even under conditions where the deviation ratio is reduced to one, the advantage of the f-m system is still over 4 db.

When two a-m carriers having a frequency separation of about 15 kc or less are received simultaneously, the listener is often annoyed by a shrill beatnote caused by the heterodyning of the two radio frequencies in the audio range. Unless the desired signal is at least 75 to 100 times stronger than the interfering one, this beatnote will continue to be heard in the background. The actual effect of heterodyning is to create a new amplitude-modulated signal that may be almost as strong as (sometimes even stronger than) the desired signal.

Two interfering f-m stations produce the same amplitude-modulation beatnote. However, this interference becomes negligible when the desired signal is only *twice as strong* as the interfering one.

29. Narrow-Band FM

On the lower frequencies, where various communications services and radio amateurs operate, frequency modulation with limited deviation is permitted by the FCC. *Narrow-band fm,* as it is generally called, is defined as fm (or pm) that does not occupy a wider channel than an a-m emission having the same audio modulating frequencies. Essentially, this definition might be restated by saying that only the first order sidebands are to be emitted, since transmission on any of the higher-order sidebands would occupy a larger slice of the spectrum than an a-m station requires. We have seen that the bandwidth is limited to first-order sidebands by keeping the modulation index 0.5 or less.

A narrow-band f-m signal is not as effective as an a-m signal of the same r-f power with 100 per cent modulation. The amplitude of the first sideband with an index of 0.5 is not quite 25 per cent of the unmodulated carrier amplitude while in a 100 per cent modulated a-m transmitter the sideband amplitude is 50 per cent of the unmodulated carrier amplitude. This yields a power ratio of only 1:4, fm to a.m.—a substantially reduced radiating effectiveness.

30. Pre-emphasis and De-emphasis

To understand the desirability of pre-emphasis in f-m transmissions, reference should again be made to the equation given in paragraph 16.

$\Delta F = f \Delta \theta$ where: F = carrier frequency
f = audio modulating frequency
θ = phase angle

Since noise interference may produce indirect fm by shifting the phase of the carrier (fm from pm), this equation shows that high audio frequency noise produces more equivalent fm than does lower frequency noise. To this must be added the fact that signals in the higher audio frequency range are usually of relatively small amplitude and can produce little frequency swing. The result of these two effects is that the signal-to-noise ratio in the equivalent fm of an uncompensated transmitter is rather poor at the higher audio frequencies.

This condition may be corrected by overemphasizing the high audio frequencies during the modulation process—a form of predistortion called pre-emphasis. This improves the signal-to-noise ratio of the received signal in the upper audio ranges; however, the reverse process—de-emphasis of the higher audio frequencies—must take place in the receiver to restore the entire modulation to its original relative values over the audio range.

De-emphasis in the receiver must occur after demodulation is complete; thus, a de-emphasis network consisting of a resistor and capacitor (Fig. 24) is inserted between the discriminator or ratio detector and the first audio amplifier. The audio voltage coupled to the grid of the audio amplifier is that which appears across capacitor *C;* because this capacitor has a lower reactance for higher frequencies, less voltage is developed across it, and hence less voltage is applied to the next stage, when the audio frequencies are high. The resistor and capacitor form an *RC* network having a time constant of 75 microseconds. This conforms with transmission standards in which the impedance specified for this combination is that of an inductance-resistance (*RL*) filter having a time constant of 75 microseconds.

Pre-emphasis followed by de-emphasis accomplishes a significant improvement in the signal-to-noise ratio for the upper audio frequencies without changing the relative values of the different modulation components.

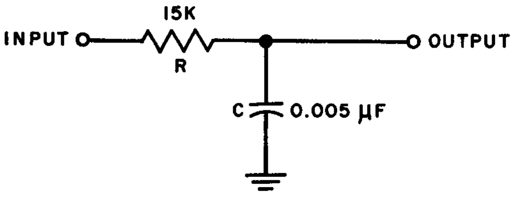


Fig. 24. De-emphasis circuit in receiver

31. The F-M Receiver

Modern f-m receivers are superheterodynes that appear very similar to equivalent a-m receivers (Fig. 25), except for the presence of the *limiter* stage or stages in the former. This close simiarity is superficial, however, because many differences exist which do not appear in the block diagrams. Some of these differences are: (a) The f-m receiver employs an entirely different system for demodulation. This stage may be a *discriminator,* a *ratio detector,* a *gated-beam detector,* or any one of a half-dozen special demodulator circuits that have been devised by manufacturers over the past years. (b) A-m sets do not contain limiters. This stage serves to remove or minimize amplitude variations in the converted f-m signal and is the most important single step toward the elimination of noise. The limiter may consist of one of two stages appearing in the i-f amplifier strip. (c) An r-f amplifier is more frequently used in f-m sets. (d) The intermediate frequency used in f-m receivers is considerably higher than the i-f in a-m sets. The standard a-m i-f is 455 kc, while most f-m manufacturers use 10.7 mc as the i-f for their receivers. (e) Since the band occupied by a wide-band f-m signal is much greater than that used in a.m., the frequency response of the i-f amplifier must be correspondingly wide.

32. R-F and Converter Considerations

The r-f amplifier and the receiving antenna have the same general function as in an a-m set: they contribute to the selectivity, improve the sensitivity and signal-to-noise ratio, and help to reject image interference.

Converter circuits, or separate oscillator-mixer arrangements, in f-m receivers resemble their counterparts in a-m sets closely. However, the stability of the oscillator is more important at the high frequencies used in fm because a small amount of frequency drift is much more noticeable in its detuning effect. Some of the precautions taken to avoid oscillator drift include the use of air trimmers rather than mica trimmers, temperature compensating capacitors in the tuning circuits, and low-loss insulation in the variable capacitors and coil forms.

33. The I-F and Limiter Requirements

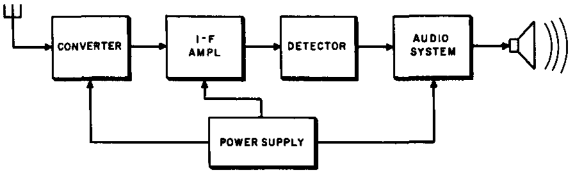
Intermediate frequency amplifier stages in f-m receivers must pass a wider band of frequencies than their equivalents in a-m equipment; the ideal frequency response for the f-m broadcast bands should be flat-topped over a range of about 150 kc and then drop off rapidly beyond this range. This kind of curve is difficult to obtain and, because the limiter stages that follow compensate in great measure for the inadequacies of the preceding i-f stages, the matter is not given as much attention as one might think necessary. In addition, the individual i-f stage gain is low as a result of the high operating frequency used—usually 10.7 mc—and attempts to improve the width of response usually lead to still further loss of gain.

It is important to note that the limiter must be fed a signal of adequate strength if it is to perform satisfactorily. If the input is weak, amplitude modulation reaches the demodulator. Thus, it is of utmost importance to design an f-m receiver with sufficient i-f gain to drive the limiter to plate saturation and cut-off at all times.

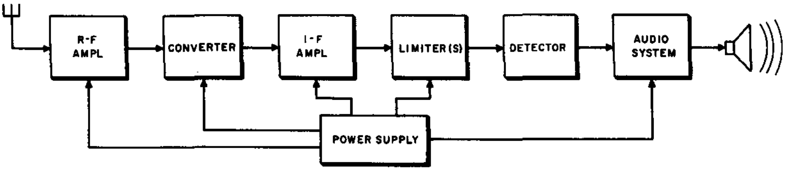
With enough i-f gain, the limiter can compensate for the lack of a flat-topped i-f response curve. It does this by clipping off the peaks of the over-amplified frequencies so that they are brought down to the same level as those which have not received as much amplification in the troughs of the i-f response curve.

34. The Demodulator

The demodulator in an f-m set is the equivalent of the second detector in an a-m superheterodyne. The general term "demodulator" is used here because there are many circuits and tubes de-



A-M RECEIVER



F-M RECEIVER

Fig. 25. Comparison between block diagrams of a-m receiver and f-m receiver.

signed for recovering the audio from a frequency modulated signal, each with its own specific name, such as discriminator, ratio-detector, gated-beam detector, locked-in oscillator detector, etc. Regardless of the circuit or tube used, an f-m demodulator provides optimum performance only if it yields output voltages that vary in strict accordance with the changes of carrier frequency introduced by the original modulation, and if it is unresponsive to noise. The popularity of the ratio detector indicates that it is still the preferred circuit in these respects; it is simple in construction, easy to align, linear in operation, and comparitively insensitive to amplitude variations. For circuit details and comprehensive descriptions of theory of operation of all f-m demodulators, the reader is referred to *F-M Limiters and Detectors,* edited by Alexander Schure.*

35. Audio Systems in F-M Receivers

The present trend toward high fidelity in broadcast and reproduction equipment calls for careful attention to the audio system of the f-m receiver. In high quality sets, every effort is made to secure an audio output that is faithful and rich. Thus, consciencious design includes push-pull output stages with relatively high percentages of inverse feedback, high quality coupling components, and multiple-speaker installations in well-designed enclosures.

36. Review Questions

(1) What are the differences between the surface wave, the direct wave, and the earth reflected wave?

(2) Above what frequency, approximately, does the sky wave cease to be of importance in radio transmission? Why?

(3) To increase line-of-sight transmission distance, which antenna should be raised—the transmitting or receiving antenna? Explain.

(4) To what length should the quarter-wave rods on an antenna used exclusively on 46 mc be cut?

(5) Primarily what kind of modulation, fm or a.m., is produced by atmospheric electricity and man-made electrical noise?

(6) Is the immunity to noise inherent in f-m reception greater or less at high deviation ratios than at low deviation ratios?

(7) What must be the ratio amplitudes of two interfering stations to permit a stronger f-m signal to dominate a weaker one completely? How does this compare with the same situation in a.m.?

*New York: John F. Rider Publisher, Inc. 1955.

(8) What are the advantages and disadvantages of narrow-band fm as compared to wideband fm?

(9) Why is pre-emphasis used in f-m transmissions?

(10) What is the recommended time constant for the de-emphasis network in f-m receivers?

(11) Explain why oscillator drift at the frequencies used for fm is more serious than the drift in a-m receivers.

INDEX