F-M Noise and Interference

This survey discusses the f-m principle in noise reduction, the extent to which frequency-modulation is effective in reducing various types of interference, compares a-m systems with f-m systems, and concludes that the distribution and phase relationships of sidebands largely explain success of "staticless" reception

By
STANFORD GOLDMAN
General Electric Co., Bridgeport, Conn.

PROBABLY the outstanding reason for the growing use of frequency modulation is its characteristically low noise and interference levels. It is the purpose of the present article to give elementary explanations of these effects, and to present simple formulas for practical use in the calculation of f-m noise and interference.

A few fundamental facts about modulation will not be out of place at this point. An audio signal consisting of a pure audio tone of frequency μ and intensity proportional to a^2, may be written as $a \cos 2\pi\mu t$. If this signal is used for amplitude modulation of the r-f carrier, $A \sin 2\pi ft$, the resulting amplitude modulated signal is

$$\begin{aligned} &A(1 + k\,a \cos 2\pi\mu t) \sin 2\pi ft \\ &= A[\sin 2\pi ft + \frac{1}{2} ka \sin 2\pi(f+\mu)t \\ &+ \frac{1}{2} ka \sin 2\pi(f-\mu)t] \qquad (1) \end{aligned}$$

where k is a constant, depending on the depth of modulation.

The result of amplitude modulation has thus been the generation of two sidebands in the frequency spectrum. These are displaced from the carrier on either side by the audio frequency μ and have a magnitude equal to $ka/2$ times the carrier magnitude. The quantity ka is called the modulation factor, and, in amplitude modulation, can never exceed 1.

If, on the other hand, the audio signal, $a \cos 2\pi\mu t$, is used for frequency modulation of the r-f carrier, $A \sin 2\pi ft$, the resulting frequency modulated signal* is

$$\begin{aligned} &A \sin[2\pi ft + (D/\mu) \sin 2\pi\mu t] \\ &= A\,J_0(D/\mu) \sin 2\pi ft \\ &+ J_1(D/\mu)[\sin 2\pi(f+\mu)t - \sin 2\pi(f-\mu)t] \\ &+ J_2(D/\mu)[\sin 2\pi(f+2\mu)t + \sin 2\pi(f-2\mu)t] \\ &+ J_3(D/\mu)[\sin 2\pi(f+3\mu)t \\ &- \sin 2\pi(f-3\mu)t] + \ldots\ldots \qquad (2) \end{aligned}$$

The result of frequency modulation has thus been the generation of sidebands displaced from the carrier frequency not only by the audio frequency, but by all harmonics† of the audio as well. The magnitudes of the sidebands are no longer simply proportional to the modulation factor divided by 2, as was the case in amplitude modulation, but are now proportional to the quantities $J_n(D/\mu)$. These quantities are called Bessel's functions, and vary somewhat like damped sine waves. A group of them is shown in Fig. 1.

The instantaneous frequency of the f-m signal, $A \sin[2\pi ft + (D/\mu) \sin 2\pi\mu t]$ is

$$\begin{aligned} &\frac{1}{2\pi}\frac{d}{dt}[2\pi ft + (D/\mu) \sin 2\pi\mu t] \\ &= f + D \cos 2\pi\mu t \qquad (3) \end{aligned}$$

The maximum frequency deviation occurs at that point of the audio cycle when $\cos 2\pi\mu t = \pm 1$. At that instant, according to Eq. (3), the frequency deviation from the carrier is equal to D. The extent of frequency modulation is measured by this maximum frequency deviation. In amplitude modulation, there is a theoretical limit to the extent of modulation, namely when the depth of modulation is equal to the carrier magnitude. There is no corresponding theoretical limit (except zero frequency) to the extent of frequency modulation. In practice, the extent of modulation, in frequency modulation, is limited by the government's ruling on maximum frequency swing or by the modulation capabilities of the transmitter. The modulation factor in f-m is defined as the ratio of the frequency deviation to the maximum permitted frequency deviation.

In Fig. 2 are shown amplitude and frequency modulated waves and their sideband components.‡ In Fig. 3 are shown the differences in sideband composition for various modulating and maximum deviation frequencies.

The General Problem of Interference

In amplitude modulation, as is well known, the ratio of audio interference to audio signal is in general the same as the ratio of r-f interference to r-f signal arriving at the second detector. This is not true in frequency modulation. In frequency modulation, the stronger of the two r-f signals arriving at the limiter grid, tends to remove the audio effect of the weaker.

The reason for this is illustrated in Fig. 4, where the two r-f signals are represented by rotating vectors, as is customary in studies of modulation. Let us suppose that A is the desired signal and B is the interference. The resultant of the two is R, and the rate of change of the angle ϕ_R is the total effective frequency modulation.

* In a uniform sine wave, such as $\sin \omega t$, the frequency is $1/(2\pi)$ times the rate of change of the argument of the sine function, or in other words

$$f = \frac{1}{2\pi}\frac{d}{dt}(\omega t) = \frac{\omega}{2\pi}$$

If the sine function is not uniform, it is customary to define its instantaneous frequency in an analogous manner as $1/(2\pi)$ times the rate of change of the argument of the sine function. In accordance with this definition, the f-m signal in Eq. (2).

$$A \sin\left[2\pi ft + \left(\frac{D}{\mu}\right) \sin\ 2\pi\mu t\right]$$

is shown by Eq. (3) to have a frequency $(f + D \cos 2\pi\mu t)$. This is the justification for saying that the expression in Eq. (2) is frequency modulated with the audio signal $a \cos 2\pi\mu t$. The constant, D, represents the maximum frequency deviation of the signal from that of the unmodulated carrier, f. This maximum deviation occurs when $\cos 2\pi\mu t = \pm 1$. The constant, D, is proportional to the audio signal strength, a, as well as to the frequency sensitivity of the transmitter.

† The "harmonic" sidebands, if proportioned as in Eq. (2) produce no distortion in an f-m signal. On the other hand, absence of harmonics in these proportions will cause distortion.

‡ Note that amplitude modulation does not change the carrier energy, but it adds sideband energy. On the other hand, frequency modulation decreases the carrier energy and puts energy into the sidebands in such a way that the total (carrier + sideband) energy is independent of the extent of modulation.

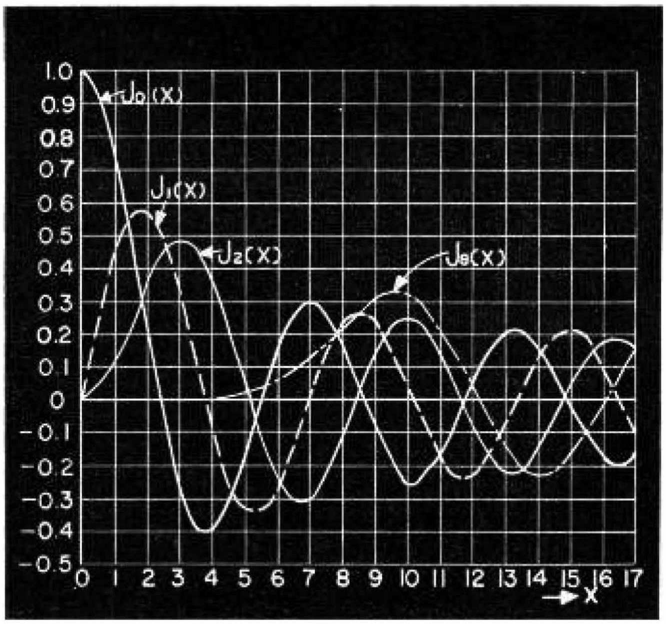

Fig. 1—A group of Bessel functions, plotted for order 0, 1, 2, and 8, and in which $x = D/\mu$. To use, we must determine x from known values of D and μ

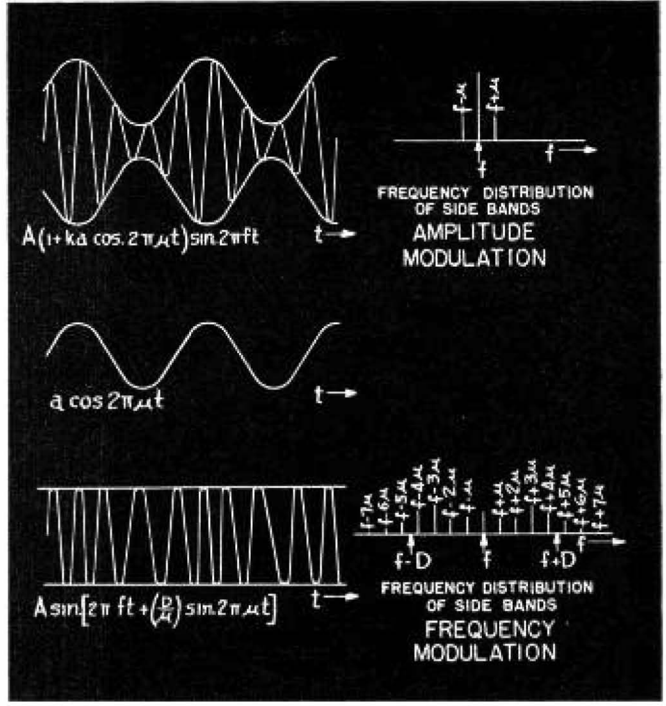

Fig. 2—Diagram of modulated waves and their sideband components for amplitude modulation (top) and frequency modulation. Modulating wave is sinusoidal

It is apparent from the figure that even if the angular variation of B is thousands of degrees, it will not cause much change in ϕ_R, for the maximum angle between R and A cannot exceed $\tan^{-1} B/A$. Therefore, if the modulation of A has a large deviation ratio (ratio of the maximum deviation frequency to the audio frequency which is the case in wide band frequency modulation so that A (and therefore R) has several complete revolutions in one audio cycle, the relative effect of B on the overall frequency modulation will be very small; considerably smaller than the value of B/A might lead one to expect from amplitude modulation experience. Furthermore the elimination of the effect of B becomes more complete as the maximum deviation frequency of A is increased.

If A is the desired f-m signal and B the interference, then the audio signal will be quite free from interference so long as A is greater than B during all portions of the audio cycle. If the relative value of B is then increased, there is a rapid rise in the amount of audio interference when B approaches the value of A, and by the time B exceeds A during all portions of an audio cycle, the interference has completely eliminated the signal. There is thus a sharp transition from good signal to poor signal as the relative value of A to B is decreased, such as would occur at a critical distance away from the A transmitter. Since the interference, B, is likely to have amplitude modulation as well as frequency modulation, the transition is not quite as sharp as it would otherwise be, but it is still very striking.

The foregoing discussion gives a general idea of the action of f-m in reducing interference. It is next desired to derive quantitative formulas for the action of f-m. This is quite involved mathematically in the general case. However, as will be seen in the following sections, in the most important practical case when the signal is considerably greater than the interference and a good limiter is used, the derivations become quite simple.

A different type of noise reduction exhibited by wide band f-m receivers, but which can also be achieved in a-m receivers, is described in footnote on page 42.

The Simplest Case of Interference

Probably the simplest type of interference in a radio receiver is that produced by a harmonic wave of fixed amplitude and frequency, such as an unmodulated carrier, $B \sin 2\pi gt$, whose intensity is small relative to that of the desired signal. Let us first find what interference this produces in an amplitude modulated signal receiver. Let the signal carrier be $A \sin 2\pi ft$. Then the resultant is

$$\begin{aligned} A \sin 2\pi ft + B \sin 2\pi gt &= A \sin 2\pi ft \\ &+ B \sin 2\pi ft \cos 2\pi(g-f)t \\ &+ B \cos 2\pi ft \sin 2\pi(g-f)t \\ &= [A + B \cos 2\pi(g-f)t] \sin 2\pi ft \\ &+ B \sin 2\pi(g-f) \cos 2\pi ft \\ &= \sqrt{A^2 + 2AB \cos 2\pi(g-f)t + B^2} \sin \left\{ 2\pi ft \right. \\ &\left. + \tan^{-1}\left[\frac{B \sin 2\pi(g-f)t}{A + B \cos 2\pi(g-f)t}\right] \right\} \end{aligned} \tag{4}$$

As the result of a Taylor series expansion the amplitude of the resultant is then, approximately,

$$\begin{aligned} &\sqrt{A^2 + 2AB \cos 2\pi(g-f)t + B^2} \\ &= A[1 + (B/A) \cos 2\pi(g-f)t] \end{aligned} \tag{5}$$

provided that A is much larger than B. Comparison of the right side of Eq. (5) with the left side of Eq. (1) indicates that the interference causes an effective interference modulation factor of B/A. Furthermore, the interference modulation frequency is $(g-f)$.

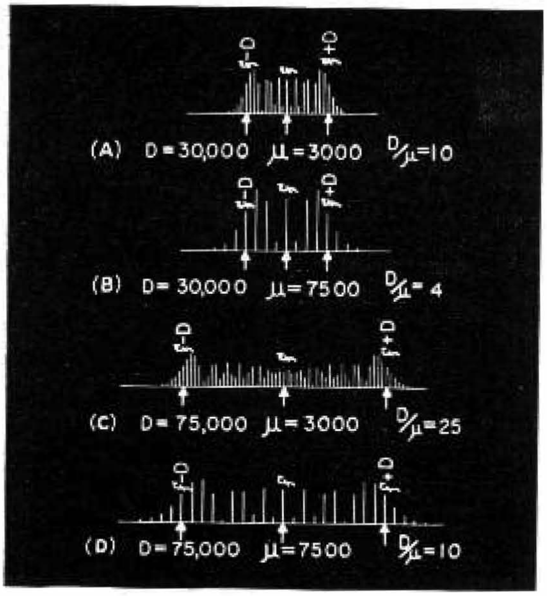

Fig. 3—Spectrum distribution in frequency-modulated system, for various deviation frequencies, D, and modulating frequencies, μ. Compare top and bottom curves for which $D/\mu = 10$ in both cases

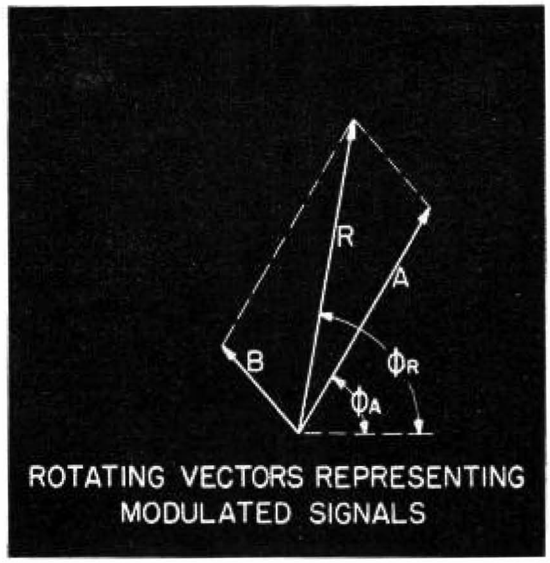

Fig. 4—Rotating vectors representing f-m signals. Vector B is the interfering signal, A is the desired signal, R is the resultant

Let us next see what effective interference modulation, the same signal $B \sin 2\pi gt$ produces in a frequency-modulated signal receiver with the same signal carrier. To find this, we refer again to Eq. (4). If it is again assumed that A is much larger than B, then, approximately

$$\sin\left[2\pi ft + \tan^{-1}\left[\frac{B \sin 2\pi(g - f)t}{A + B\cos 2\pi(g-f)t}\right]\right] = \sin\left[2\pi ft + (B/A)\sin 2\pi(g-f)t\right] \quad (6)$$

Therefore, by comparison with the left side of Eq. (2), we see that the modulation factor is

$$\frac{B}{A}\cdot\left(\frac{g-f}{D}\right) \quad (7)$$

where D is the maximum frequency deviation of the f-m system. The modulation factor of the interference in frequency modulation is consequently less than in the corresponding a-m case by the ratio $(g-f)/D$. It is therefore clear that the interference is reduced in the same proportion as the maximum deviation of the f-m system is increased, as already indicated in our general discussion of interference in the preceeding section. Equation (6) shows that audio interference signal has the frequency $(g-f)$ just as in the a-m case. It also shows the further important fact that the modulation factor of the interference is directly proportional to the audio frequency $(g-f)$.

Common Channel Interference

Common channel interference is the interference between the desired signal and an interfering signal of approximately the same carrier frequency. The modulation produced on the desired carrier, by the interfering carrier and its sidebands is the measure of common channel interference.

For the sake of simplicity, we shall assume that both the interfering and the desired carriers are unmodulated. Then, in accordance with the analysis of the preceding section, the modulation factor of the interference is B/A in the amplitude-modulation case, and is $B\ (g-f/AD$ in the frequency-modulation case. Here, as before, B and A are the relative signal strengths of the interfering and signal carriers, $g-f$ is the difference frequency between the carriers, and D is the maximum frequency deviation of the f-m system. In the case of wide band frequency modulation, D is greater than the highest audio frequency passed by the receiver, so that common channel interference in f-m is necessarily less than it is in amplitude modulation. This is a distinct advantage of frequency modulation.

Present f-m transmission standards call for an accentuation of high audio tones at the transmitter, with an equalization of the system by a corresponding decrease in high frequency response at the receiver. This decrease in high frequency response at the receiver is approximately linear above 1500 cps. The f-m carrier difference-frequency interference is therefore limited to 1500 B/AD regardless of the pitch of the carrier difference-frequency. At the frequencies at which frequency modulation is now broadcast (42 Mc to 50 Mc), the pitch of the difference-frequency will vary rapidly throughout the audio range at all times due to frequency drift so that 1500 B/AD is a good approximation to the average common channel interference. By a more elaborate study*, it can be shown that this approximation for the average inter-

* S. Goldman, IRE Convention, June, 1940.

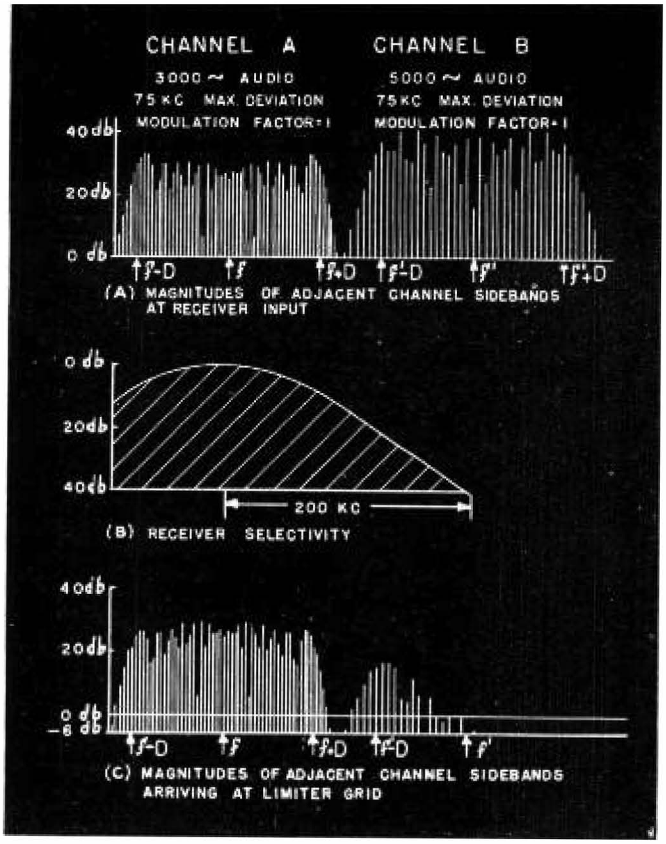

Fig. 6—Amplitude-frequency spectrum of frequency-modulated signals in two adjacent channels, for interfering signal in adjacent channel 6 db greater than desired signal

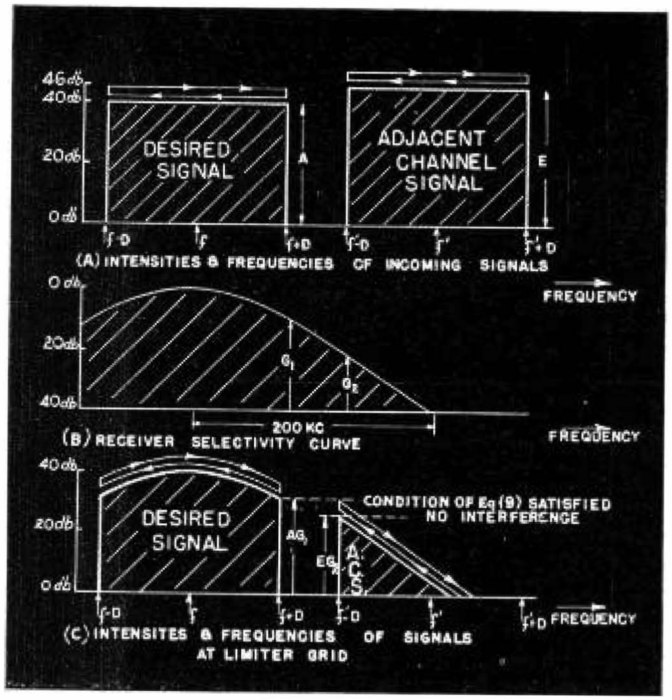

Fig. 5—Graphical study of adjacent channel sideband interference, for adjacent channel 6 db above desired channel signal. Shaded areas indicate desired and undesired signals

ference is still valid even when both carriers are modulated by a normal program.

With present f-m standards, D is normally 75,000 cps. Therefore f-m common channel interference is

$$\frac{1500}{75000} \cdot \frac{B}{A} = \frac{1}{50} \cdot \frac{B}{A} \qquad (8)$$

or about one fiftieth of what it would be for the corresponding case in amplitude modulation.

Adjacent Channel Interference

In the case of adjacent channel interference on present f-m channels, the carrier difference-frequency note is far above audibility, so that it no longer causes interference. However, the higher order sidebands in frequency modulation extend so far away from the carrier that interaction between sidebands may be of audible frequencies. The selectivity of the receiver also enters into the adjacent channel interference picture, so that all in all it is quite a different story from the common channel case.

There are two ways of treating f-m adjacent channel interference, one of which may be called static and the other dynamic. Both of these treatments must be used for a complete picture. In the static treatment we consider the carrier and sidebands as shown in Fig. 5A and the selectivity of the receiver. The magnitude of sidebands getting to the limiter grid as a result of the selectivity, is shown in Fig. 5C. If we know the magnitudes of these sidebands, we can calculate the interaction of audible frequencies. It is by no means obvious that the interaction of sidebands in frequency modulation will give rise to audio or how large this audio will be. However, it has been shown* that the interaction of two adjacent channel signals will give audio of the difference frequencies of the adjacent channels' sidebands and of amounts proportional to the products of the magnitudes of the sidebands multiplied by the factor audio difference frequency, divided by the maximum deviation frequency of larger signal carrier at limiter. This derivation is rather involved and will not be reproduced here.

In the dynamic treatment of f-m adjacent channel interference, we consider the carriers of constant intensity but varying in frequency as shown in Fig. 6. If now, during any appreciable portion of an audio cycle, the interfering carrier intensity arriving at the limiter grid exceeds that of the signal carrier, then, as shown previously in discussing the general problem of interference, the signal is ruined as a high quality signal. If, on the other hand, *the level of the desired signal exceeds that of the interfering signal at the limiter grid during all portions of the audio cycle* then it may be shown mathematically for all normal signals that the adjacent channel interference will be at least 60 db below the signal level. The condition indicated in italics thus is both necessary and sufficient for an f-m signal free from adjacent channel interference. This condition may be expressed by the following formula:

* S. Goldman, IRE Convention, June, 1940.

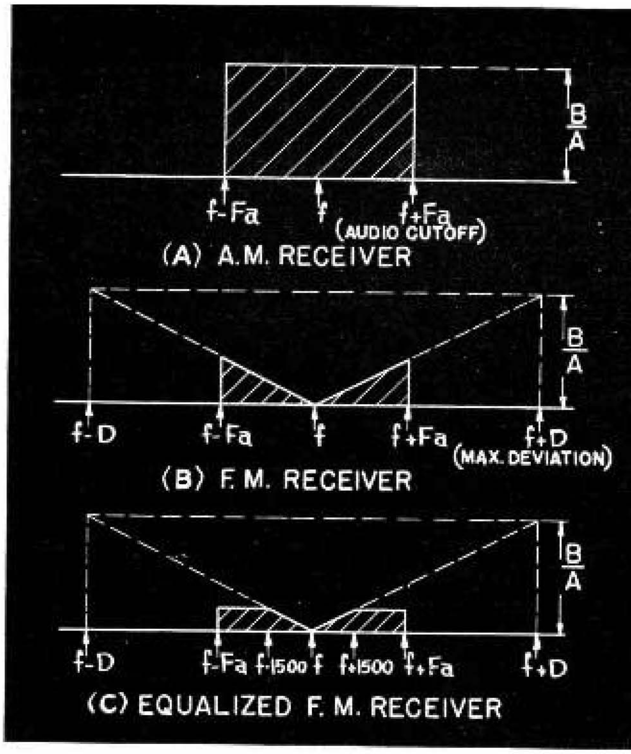

Fig. 7—Shaded areas are proportional to the effective values of sideband noise for three different types of receivers

$$AG_1 > EG_2 \qquad (9)$$

where A is the level of the desired signal at the input of the receiver,

G_1 is the gain of the receiver at the frequency of maximum deviation of the signal toward the adjacent channel,

E is the level of the adjacent channel interfering signal at the input of the receiver, and

G_2 is the gain of the receiver at the frequency of maximum deviation of the adjacent channel toward the signal channel.

According to Eq. (9), there must be enough receiver selectivity in the frequency range between the frequencies of maximum deviation of the signals toward each other, to take care of the difference in level of the desired and adjacent channel signals at the input of the receiver.

Random and Impulse Noise

While any undesired response of a radio receiver, such as hum, adjacent channel interference, etc., may be called noise, we wish to consider here particularly the two kinds of noise‡ usually called random noise and impulse noise. Random noise is a general type of noise such as the interchannel noise of a very sensitive radio receiver. This noise is due to a continuous distribution of nondescript radio frequency signals of unrelated phases. Interaction between these radio frequency noise signals and a strong signal carrier gives rise to audible noise. Interaction between the r-f noise signals themselves, also gives rise to audible noise, but this is of much smaller amount in amplitude modulation, and is negligible in frequency modulation, if a carrier is present.

The audio noise consists of terms of the type described by Eq. (5) and (B/A), in amplitude modulation and terms described by Eqs. (6) and (7) in frequency modulation. The shaded areas in Fig. 7 show the magnitudes of the audio effects of these sidebands in accordance with (B/A) and Eq. (7).

Since the phases of the r-f signals in random noise are unrelated, the total noise modulation factor is equal to the square root of the sum of the squares of all these terms. Thus the noise in amplitude modulation is proportional to the square root of the band-width. This may be expressed by the formula

$$\text{Noise} = K\sqrt{F_a} \qquad (10)$$

where K is a constant, and F_a is the highest audio frequency reproduced by the receiver. It then follows from Fig. 7 that in frequency modulation,

$$\text{Noise} = K\sqrt{\int_0^{F_a}\left(\frac{g-f}{D}\right)^2 d(g-f)}$$

$$= \frac{K\,F_a}{\sqrt{3}\,D}\sqrt{F_a} \qquad (11)$$

Consequently for random noise, frequency modulation is superior to amplitude modulation by a (voltage) factor of $\sqrt{3}\,D/F_a$ in case there is no receiver equalization. For the equalized receiver, the improvement ratio of frequency modulation over amplitude modulation is approximately $D/1500=50$ for random noise, if $D=75{,}000$ cps.

Let us next consider impulse noise, for those cases in which the signal exceeds the noise peaks. This is noise characterized by high peaks of short duration, such as are generated by automobile ignition systems, and by many important natural and man-made sources. The components of impulse noise are essentially spread uniformly over the transmission band of a receiver in a way similar to the frequency spread of random noise. However, the phases of impulse noise components are not spread at random.† It is clear physically and may be demonstrated by Fourier analysis that when an impulse is at its peak, its frequency components must be in phase. Therefore the total noise modulation factor of *impulse noise* components is the sum of that of the individual components, and is not an r-m-s value. Consequently, in amplitude modulation systems, the peak voltage of the resultant of the frequency components, the peak voltage of the transmitted impulse noise, is directly proportional to the bandwidth of the transmission system of the receiver. This is quite different from the case of random noise, in which the peak voltage is proportional to the square root of the bandwidth of the transmission system. Wide band reception makes impulse peaks higher and therefore makes it more likely that they will

‡ The principal formulas of this section were originally derived by Crosby, *Proc. I.R.E.*, April, 1937.

† V. D. Landon, *Proc. I.R.E.*, Nov. 1936, p. 1514.

exceed the signal level and thus cause serious interference. Consequently, in weak signal areas, ignition noise may be more disturbing in wide band frequency modulation than in narrow band frequency modulation or in amplitude modulation.

It may readily be shown mathematically that the peak frequency deviation caused by impulse noise, or, in other words, the peak f-m impulse noise, is likewise proportional to the sum of the effects of the components in the band width of the transmission system. This band width is the audio frequency transmission band of the receiver, since any higher frequency noise components are lost in the audio amplifier and speaker. In the case of any single impulse, both the frequency modulation and amplitude modulation effects of the impulse depend upon the phase of the r-f signal carrier at the instant that the impulse occurs. These phase effects will, however, average out in any overall picture.

The relative values of impulse noise are therefore.

$$\text{Impulse noise} = k\,F_a \tag{12}$$

for amplitude modulation and

$$\text{Impulse noise} = k \int_o^{F_a} \left(\frac{g-f}{D}\right) d\,(g-f) = k\,\frac{F_a^2}{2D} \tag{13}$$

for frequency modulation. Consequently frequency modulation is superior to amplitude modulation for impulse noise by a (voltage) factor, $2D/F_a$ in case there is no receiver equalization. For the equalized receiver, the improvement ratio of frequency modulation over amplitude modulation is approximately $D/1500 = 75{,}000/1500 = 50$ for impulse noise. This is the same ratio as for random noise.

Discussion

In the foregoing treatment, formulas were derived to show the amount of interference and noise reduction in frequency modulation systems. These formulas were derived on the assumption that the r-f signal reaching the limiter tube was considerably greater than the interference. This is quite a justifiable assumption in deriving the interference reduction obtained in frequency modulation systems, for if the condition is not satisfied and the interference at the limiter tube is greater than the signal, then as was pointed out, there is, except in special cases,* no reduction in interference at all, but rather an increase in it. The derivations also assumed that only frequency modulation gave rise to audio in the f-m receiver. In other words, a perfect limiter was assumed. The formulas derived thus show to what extent frequency modulation reduces interference under conditions when it does so best.

The real explanation of frequency modulation's effectiveness is to be found in the distribution and phase relations of the f-m sidebands. To understand this, suppose that an audio signal *amplitude modulates* an r-f carrier of say two megacycles and that this modulated carrier is amplified by a high gain amplifier which has uniform amplification in the entire frequency range from zero to four megacycles. If the output of this amplifier is observed on an oscilloscope, the noise will be tremendous and will mask the signal unless the latter is very large. If, however, the amplifier is tuned and only passes a signal in the region within ±10 kc of the carrier, the noise will be greatly reduced while the signal will be practically unaffected, so that a tremendous improvement in the signal to noise ratio will be observed. This shows how frequency selectivity reduces noise. On the other hand, mistuning of an i-f amplifier will demonstrate very quickly how frequency selectivity reduces adjacent channel interference.

The next question which naturally arises is why frequency modulation reduces interference. Prior to an investigation, it might be supposed that the limiter tube, by cutting off noise peaks, was the important factor in noise reduction. This, however, is not the answer; because noise peaks which exceed the signal will be just about as noisy in f-m as in a-m receivers and it is only when the signal reaching the limiter exceeds the interference that frequency modulation is effective. The real and very important increase of the limiter is that it strips the signal of amplitude modulation and allows f-m to do its work.

In the case of an f-m signal there is a combined system of amplitude, frequency, and phase selectivity. The selective circuit in this case is the discriminator or slope detector in combination with the audio amplifier of restricted band width. If a signal comes to this detector with amplitude, frequency, and phase relations of its components such as shown on the right side of Eq. (2), these components will combine in such a way as to give maximum frequency shift at an audio frequency rate and consequently maximum audio output. If on the other hand the amplitude, frequency, and phase relations of the components are not those shown on the right side of Eq. (2), they will not combine efficiently to give frequency shift at an audio frequency rate. The combination of carrier and sidebands on the right side of Eq. (2) has the property that it will produce extremely large frequency shifts at a low audio frequency rate as compared with what would be produced by a random distribution of carrier and sideband components of the same energy. Thus the slope detector in combination with the audio amplifier effectively selects those signals with amplitude, frequency, and phase relations of their components similar to those shown on the right side of Eq. (2).

The explanation of the effectiveness of frequency modulation in reducing noise and interference may therefore be considered as a generalized type of selectivity. It would be interesting to consider possible uses of other types of selective circuits. A different type of combined amplitude, phase and frequency selectivity already used in radio receivers is the demodulation of a weak carrier by a strong one at the second detector.

* In the case of impulse noise of very high peaks of very short duration, there is another effect which works to the advantage of wide band f-m systems, which in effect, however, is not peculiar to frequency modulation.

Short duration impulses are always spread out in the i-f amplifier of a receiver, the amount of spread being inversely proportioned to the band width of the i-f amplifier. This is a well-known property of tuned circuits. Now a sharp impulse will cause a tremendous frequency shift in the f-m signal, corresponding with, although not necessarily proportional to, the high amplitude peak. However, the band widths of most i-f amplifiers and the outputs of most slope detectors in use are limited to very little more than that produced by a 75 kc. frequency shift. A band width of ±75 kc from the carrier is nevertheless still sufficiently wide to reduce the duration of the impulse to a very short time. The maximum audio output produced by an impulse is thus limited, whereas its duration is kept short. The overall effect of highly peaked, very short duration impulses is therefore greatly reduced in wide band f-m receivers.

This phenomenon is of considerable practical importance, but as already mentioned, it is not peculiar to frequency modulation. Wide band a-m receivers with amplitude limiting, would show the same property.