The Bell System Technical Journal

Vol. XVIII *July, 1939* *No. 3*

Frequency-Modulation: Theory of the Feedback Receiving Circuit

By JOHN R. CARSON

THIS paper may be regarded both as a continuation of a prior one by the writer and Thornton C. Fry [1] and as a companion of that by J. G. Chaffee [2] the inventor of the circuit under consideration. For an understanding of the present, an acquaintance with the prior paper [1] is absolutely necessary, since the fundamental analysis and the formulas there developed are too lengthy to be repeated here. References to that paper will be designated by (Ref.).

As the name implies, in the feedback circuit part of the incoming signal, after passing through a band-pass filter, a frequency detector [3] and a demodulator, is fed back through a variable frequency oscillator. The output of the variable frequency oscillator is connected to one branch of a modulator on the other branch of which the incoming high-frequency wave is impressed. While this method of feedback differs in some respects from that of the well known feedback amplifier, it is a fair inference that some if not all of the very important advantages of the feedback amplifier may also be present in the circuit under discussion. This inference is verified by the mathematical analysis of this paper.

After a brief development of the elementary theory and formulas of the feedback circuit as a receiver of frequency-modulated waves, the greater part of the paper is devoted to deriving formulas for the signal-to-noise power ratio—a criterion of fundamental importance in estimating the merits of the system. These are then compared with the corre-

[1] "Variable Frequency Electric Circuit Theory," *this Journal*, October 1937.

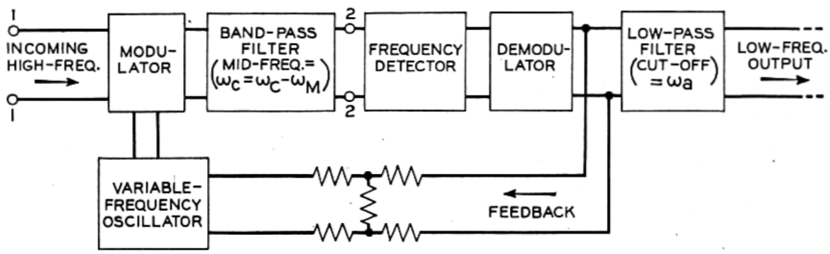
[2] "The Application of Negative Feedback to Frequency-Modulation Systems," *I. R. E. Proceedings*, May 1939; this issue of the *Bell Sys. Tech. Journal*.

[3] The function of the "frequency detector" is to detect or render explicit the variation of the "instantaneous frequency" of the frequency-modulated wave. A more precise term, therefore, would be "frequency variation detector," but for brevity the term used in the text is preferable.

sponding ratios formulated in (Ref.) for straight reception and also reception with amplitude limitation. In this way, as regards reduction of noise and "fading," the feedback circuit is found to have advantages comparable with those attainable by amplitude limitation.[4]

I

The receiving system operates as follows (see sketch):



Feedback receiving circuit.

The incoming frequency-modulated wave at terminals 1, 1 is taken as

$$E \exp\left(i\omega_C t + i\lambda \int^t s\,dt \right), \tag{1}$$

where E is the wave amplitude, ω_C the carrier frequency, λ the modulation index and $s = s(t)$ is the low-frequency signal which it is desired to recover.

This wave is impressed at terminals 1, 1 on one pair of terminals of a "product" modulator; on the other pair of terminals of the modulator there is impressed the output of a local variable-frequency oscillator:

$$M \exp\left(i\omega_M t + i\mu \int^t \sigma\,dt \right). \tag{2}$$

Here ω_M is the "carrier" frequency of the oscillator, μ (a positive real quantity) is the index of modulation of the oscillator and $\sigma = \sigma(t)$ is the low-frequency current fed back to the oscillator.

The output wave of the modulator is then equal to

$$c_1 EM \exp\left(i(\omega_C - \omega_M)t + i \int^t (\lambda s - \mu\sigma)dt \right) + c_1 EM \exp\left(i(\omega_C + \omega_M)t + i \int^t (\lambda s + \mu\sigma)dt \right). \tag{3}$$

[4] Armstrong, *Proc. I. R. E.*, May 1936, also see (Ref.).

The second term of (3) is suppressed by the band-pass filter.[5] Then writing $\omega_C - \omega_M = \omega_c$, it follows that the effective output wave is

$$c_1 EM \exp\left(i\omega_c t + i \int^t (\lambda s - \mu\sigma) dt \right). \tag{4}$$

ω_c is the intermediate carrier frequency and is always $< \omega_C$, the transmitted carrier frequency. The constant c_1 is a parameter depending on the characteristics of the modulator.

The wave (4) is transmitted through the band-pass filter, and the wave arriving at terminals 2, 2 is then

$$c_1 c_2 EM \exp\left(i\omega_c t + i \int^t (\lambda s - \mu\sigma) dt \right). \tag{5}$$

The parameter c_2 (taken as a constant) depends on the transmission characteristics from the modulator to terminals 2, 2.

Assuming an ideal frequency detector (see Ref.) the output to the terminals of the rectifier (or demodulator) is

$$c_1 c_2 c_3 EM \left(1 + \frac{\lambda s - \mu\sigma}{\omega_1} \right) \exp\left(i\omega_c t + i \int^t (\lambda s - \mu\sigma) dt \right). \tag{6}$$

Here the parameters c_3 and ω_1 depend on the characteristics of the frequency detector.

Finally assuming that

$$\frac{\lambda s - \mu\sigma}{\omega_1} < 1 \tag{7}$$

the low-frequency output of the rectifier [6] is

$$c_1 c_2 c_3 c_4 EM \left(1 + \frac{\lambda s - \mu\sigma}{\omega_1} \right). \tag{8}$$

If the constant term of (8) is suppressed and a fraction η of the rectified output is fed back to the oscillator we have, finally,

$$\sigma = \frac{m}{\mu} \frac{\lambda s}{1 + m}, \tag{9}$$

[5] Indeed the principal function of the band-pass filter is to suppress frequencies in the neighborhood of $\omega_C + \omega_M$.

[6] More generally a demodulator. In the present paper a straight-line rectifier is postulated for mathematical simplicity but the theory applies equally well to other forms of detection or demodulation.

where

$$m = c_1 c_2 c_3 c_4 \frac{\eta \mu E M}{\omega_1}, \\ = C \frac{\eta \mu E M}{\omega_1}. \tag{10}$$

The low-frequency current delivered to the receiver through the low-frequency output circuit proper differs from σ as given by (9) by a constant factor only.

From the foregoing we note that

$$\lambda s - \mu \sigma = \frac{\lambda s}{1 + m} \tag{11}$$

and that the "instantaneous frequency" of the intermediate high-frequency wave (4) is

$$\omega_c + \frac{\lambda s}{1 + m}. \tag{12}$$

Hereinafter, without any loss of generality, we suppose that $-1 \leqq s(t) \leqq 1$. Consequently the intermediate frequency-modulated wave has a frequency variation lying between $\pm \lambda/(1 + m)$, whereas in the incoming frequency-modulated wave, the frequency variation lies between $\pm \lambda$.

We note also from (9) that if the parameter m is large compared with unity, the low-frequency received wave is approximately given by

$$\sigma = \frac{\lambda}{\mu} s. \tag{13}$$

The recovered signal is thus (for large values of m) seen to be independent of the amplitude, E, of the incoming high-frequency wave; therefore, the system is insensitive to "fading."

II

We now take up the problem of calculating the relative low-frequency *noise* and *signal* powers, the ratio of which is of fundamental importance in appraising the merits of the receiving circuit. In this we shall closely follow the methods developed in Section IV (Ref.).

We suppose that at terminals 1, 1 there enters, in addition to the signal, a typical noise element

$$a_n \exp\left(i(\omega_C + \omega_n)t + i\alpha_n\right). \tag{14}$$

We write for convenience in the subsequent analysis

$$a_n = A_n E. \tag{15}$$

A_n is then the *relative* amplitude of the noise element, referred to the amplitude E of the high-frequency signal. We shall suppose throughout that A_n is small compared with unity; that is, the noise is small compared with the signal.

We further suppose that at terminals 2, 2 between the band-pass filter and the frequency detector there is introduced a second typical noise element

$$b_n \exp\,(i(\omega_c + \omega_n)t + i\beta_n), \tag{16}$$

which is entirely independent of the noise element (14). This may be regarded as caused by tube-noise in amplifiers (not shown in sketch). We write

$$b_n = B_n E \tag{17}$$

so that B_n is the relative amplitude of the noise element, referred to the amplitude E of the incoming signal wave. It also is assumed small compared with unity.

The total input to the frequency detector, neglecting the random phase angles, is then

$$c_1 c_2 EM \left[\exp\left(i \int^t \Omega\, dt \right) + A_n \exp\left(i \int^t (\Omega + \Omega_n{}^a) dt \right) + \frac{B_n}{c_1 c_2 M} \exp\left(i \int^t (\Omega + \Omega_n{}^b) dt \right) \right], \tag{18}$$

where

$$\begin{aligned} \Omega &= \omega_c + \lambda s - \mu\sigma \\ \Omega_n{}^a &= \omega_n - \lambda s \\ \Omega_n{}^b &= \omega_n - \lambda s + \mu\sigma. \end{aligned} \tag{19}$$

The output of the frequency detector is then (see Ref.)

$$c_1 c_2 c_3 EM \exp\left(i \int^t \Omega\, dt \right) \times \left[1 + \frac{1}{\omega_1} (\lambda s - \mu\sigma) + A_n \left(1 + \frac{1}{\omega_1} (\omega_n - \mu\sigma) \exp\left(i \int^t \Omega_n{}^a\, dt \right) \right) + \frac{B_n}{c_1 c_2 M} \left(1 + \frac{\omega_n}{\omega_1} \right) \exp\left(i \int^t \Omega_n{}^b\, dt \right) \right]. \tag{20}$$

After the output as given by (20) is rectified, the constant term suppressed, and only first powers in A_n and B_n retained, we get finally

$$\sigma = \left(\frac{m/\mu}{1+m}\right)\left[\left(\lambda s + A_n\left(\omega_1 + \omega_n - \frac{m}{1+m}\right)\lambda s\right)\cos\int^t \Omega_n{}^a\,dt + \frac{B_n}{c_1c_2M}(\omega_1+\omega_n)\cos\int^t \Omega_n{}^b\,dt\right]. \quad (21)$$

Now the right-hand side of (21) corresponds precisely with formula (64) (Ref.) on which the calculation of the relative low-frequency noise and signal powers is based. Consequently following the methods developed in Ref. and assuming A_n and B_n small we get

$$\overline{\sigma^2} = \left(\frac{m/\mu}{1+m}\right)^2\left[\lambda^2\overline{s^2} + \left(\frac{1}{3}\omega_a^2 + \omega_1^2 + \frac{\lambda^2\overline{s^2}}{(1+m)^2}\right)\omega_a N_a^2 + \left(\frac{1}{3}\omega_a^2 + \omega_1^2 + \overline{(\lambda s - \mu\sigma)^2}\right)\omega_a N_b^2/c_1^2c_2^2M^2\right]. \quad (22)$$

The *relative* low-frequency noise and signal powers are then (omitting the common factor $\left(\frac{m/\mu}{1+m}\right)^2$)

$$P_N = \frac{1}{3}\omega_a^3N_a^2\left[1 + 3\left(\frac{\omega_1}{\omega_a}\right)^2 + 3\frac{(\lambda/\omega_a)^2}{(1+m)^2}\overline{s^2}\right] + \frac{1}{3}\frac{\omega_a^3N_b^2}{c_1^2c_2^2M^2}\left[1 + 3\left(\frac{\omega_1}{\omega_a}\right)^2 + 3\frac{\overline{(\lambda s - \mu\sigma)^2}}{\omega_a^2}\right], \quad (23)$$

$$P_S = \lambda^2\overline{s^2}.$$

In these formulas N_a^2 is proportional to the noise power level in the neighborhood of the carrier frequency ω_C; it enters at the input terminals 1, 1 (see Ref. Appendix 2). N_b^2 is proportional to the noise power level in the neighborhood of the intermediate carrier frequency ω_c; it enters at terminals 2, 2. ω_a is the highest essential frequency in the low-frequency signal $s(t)$; it is the cut-off frequency of the low-pass filter.

Formula (22) is solvable (see Appendix) but a simple approximate solution, valid when the noise is small compared with the signal, is made possible by observing that under this restriction

$$\lambda s - \mu\sigma = \frac{\lambda s}{1+m} \quad (11)$$

to a good approximation. Introducing this approximation into P_N as given by (23) and writing

$$N^2 = N_a^2 + N_b^2/c_1^2c_2^2M^2, \tag{24}$$

we have

$$P_N = \frac{1}{3}\,\omega_a^3N^2\left(1 + 3\left(\frac{\omega_1}{\omega_a}\right)^2 + 3\,\frac{(\lambda/\omega_a)^2}{(1+m)^2}\,\overline{s^2}\right), \tag{25}$$
$$P_S = \lambda^2\overline{s^2}.$$

Now from the inequality, necessary for rectification,

$$\omega_1 > \frac{\lambda}{1+m}$$

it is seen that as the parameter m is increased, ω_1 may be reduced by the factor $1/(1+m)$. In accordance with this, we replace ω_1 by $\omega_1/(1+m)$ in (25) and get

$$P_N = \frac{1}{3}\,\omega_a^3N^2\left[1 + \frac{3}{(1+m)^2}\left(\left(\frac{\omega_1}{\omega_a}\right)^2 + \left(\frac{\lambda}{\omega_a}\right)^2\overline{s^2}\right)\right], \tag{26}$$
$$P_S = \lambda^2\overline{s^2}.$$

The noise power P_N can be still further reduced by eliminating ω_1 from (26) by a circuit arrangement explained in Ref. Section III; if this is done we get, instead of (26),

$$P_N = \frac{1}{3}\,\omega_a^3N^2\left(1 + 3\,\frac{(\lambda/\omega_a)^2}{(1+m)^2}\,\overline{s^2}\right),\quad P_S = \lambda^2\overline{s^2}. \tag{27}$$

We have now to compare the relative noise and signal powers of the feedback with (1) straight reception *without* feedback and (2) reception with *amplitude limitation.*

For *straight reception* (without feedback) we have (see equation (68), Ref.), corresponding to (26),

$$P_N = \frac{1}{3}\,\omega_a^3N^2\left(1 + 3\left(\frac{\omega_1}{\omega_a}\right)^2 + 3\left(\frac{\lambda}{\omega_a}\right)^2\overline{s^2}\right),\quad P_S = \lambda^2\overline{s^2}, \tag{28}$$

and corresponding to (27)

$$P_N = \frac{1}{3}\,\omega_a^3N^2\left(1 + 3\left(\frac{\lambda}{\omega_a}\right)^2\overline{s^2}\right),\quad P_S = \lambda^2\overline{s^2}. \tag{29}$$

When noise reduction is effected by *amplitude limitation* the corresponding relative noise and signal powers (see equation (78), Ref.) are

$$P_N = \frac{1}{3}\omega_a{}^3 N^2, \tag{30}$$
$$P_S = \lambda^2 \overline{s^2}.$$

If we assume as above that $-1 \leq s \leq 1$ and $\overline{s^2}$ is of the order of magnitude of 1/2, then in practical applications $\lambda/\omega_a \gg 1$ and $\omega_1 > \lambda$. On this basis comparison of (26) with (28) and (27) with (29) shows that, when $m \gg 1$, the noise power with *feedback* is very much smaller than *without feedback*, the ratio of the noise powers in the two cases being approximately $1/(1+m)^2$. (This assumes, of course, that N^2 is approximately equal in the two cases.)

Comparing, however, the noise power with *feedback* to that obtainable by *amplitude limitation*, it will be seen that in order to reduce the former to the order of magnitude of the latter it is necessary that

$$\frac{(\lambda/\omega_a)}{(1+m)} < 1. \tag{31}$$

From the preceding it is seen that the performance of the feedback circuit and the reduction in noise-power ratio obtainable depend in a fundamental manner on the parameter m, defined above by the formula

$$m = c_1 c_2 c_3 c_4 \frac{\eta\mu E M}{\omega_1}. \tag{32}$$

If the characteristics of the modulator rectifier and variable-frequency oscillator are stipulated, it is possible to calculate m in terms of these characteristics and the constants and connections of the network. It is experimentally determinable (among other ways) as follows:

Let the feedback circuit be opened between the low-pass filter and the variable-frequency oscillator, and let the filter be closed by an impedance equal to that of the oscillator as seen from the filter. Then $m = 0$ (since there is no low-frequency feedback to the oscillator) but m/μ is finite.

Denoting the value of σ under these circumstances by σ_1, it follows from (9) that

$$\sigma_1 = \frac{m}{\mu}\lambda s. \tag{33}$$

Consequently dividing σ_1 by σ, as given by (9), we have

$$1 + m = \sigma_1/\sigma,$$
$$m = \frac{\sigma_1 - \sigma}{\sigma}. \tag{34}$$

Stated in words, $1 + m$ is the reciprocal of the ratio of the values of σ *without* and *with* the low-frequency feedback into the oscillator. It should be noted that this requires that the band-pass filter transmit the frequency band 2λ centered on ω_c.

Appendix

In formula (22), the expression $\overline{(\lambda s - \mu\sigma)^2}$ has been replaced by $\lambda^2\overline{s^2}/(1 + m)^2$, its value when the noise is absent. When noise is present, but small compared with the signal, this should still give a good approximation for $\overline{\sigma^2}$. We now propose to derive an exact solution of (22); to this end we write

$$\lambda s - \mu\sigma = \frac{\lambda s}{1 + m} - n(t) \tag{1a}$$

which is always possible.

Now inspection of (1a) shows that $n(t)$ is the value of $\mu\sigma$ when $s = 0$; consequently $\overline{n^2} = \mu^2\overline{\sigma_0^2}$ where $\overline{\sigma_0^2}$ is given by (34). Furthermore, since s and n are entirely independent, $\overline{sn} = 0$, and

$$\overline{(\lambda s - \mu\sigma)^2} = \frac{\lambda^2\overline{s^2}}{(1 + m)^2} + \mu^2\overline{\sigma_0^2}\,. \tag{2a}$$

Substitution of (2a) in (22) gives for P_N, instead of (23),

$$P_N = \frac{1}{3}\,\omega_a{}^3 N^2 \left(1 + 3\left(\frac{\omega_1}{\omega_2}\right)^2 + 3\,\frac{(\lambda/\omega_a)^2}{(1 + m)^2}\,\overline{s^2} \right) + \omega_a \mu^2 \overline{\sigma_0{}^2} N_b{}^2/c_1{}^2 c_2{}^2 M^2. \tag{3a}$$

The second term is a second order quantity in the noise power and may therefore be neglected when the noise is small, as is assumed throughout this paper.