Theory of Error Rates for Digital FM

By J. E. MAZO and J. SALZ

(Manuscript received June 29, 1966)

A general theory is presented for evaluating the error performance of a digital FM system in the presence of additive noise. The digital system considered is a conventional one employing a voltage-controlled oscillator as the modulator and a limiter-discriminator followed by a low-pass filter as the demodulator. Because of the nonlinear nature of the demodulation process, no adequate analytical techniques have been available to provide a satisfactory treatment. Adopting the notion of "clicks" used by S. O. Rice to study threshold effects in analog FM systems, we have succeeded in evolving a theory capable of predicting performance for a wide range of applications. While our theory reinforces some previously derived results for binary and for narrow-band systems, the results obtained here are not confined to these situations. In particular, the inefficiency of the FM discriminator as a detector for a large number of orthogonal signals is quantitatively evaluated, as well as the role of the post-detection filter. Some qualitative aspects of the error-causing mechanisms discussed in the paper are general, but quantitative results are confined to additive Gaussian noise and large signal-to-noise ratios.

I. INTRODUCTION

Theoretical investigations of FM receivers with analog input signals date back to J. R. Carson and T. C. Fry,[1] and to M. G. Crosby.[2] These investigators and others that followed them[3,4,5] were primarily concerned with the signal-to-noise (S/N) transfer attainable in FM receivers and the determination of threshold effects. Recently S. O. Rice,[6] and previously J. Cohn,[7] attacked the threshold problem in FM receivers from a fresh point of view by using the notion of "clicks." It has been observed that when the noise at the input of an FM receiver is increased beyond some value, the receiver "breaks," that is, for a given (S/N) at the input, a much poorer (S/N) at the output is measured than would be predicted from a linearized analysis of the receiver. Before the breaking point, clicks are heard in the output of an audio receiver. As the input noise is further increased, the clicks merge into a sputtering sound. Rice's approach is to relate this breaking point with the expected number of clicks per second at the output due to the added noise at the input.

While in analog application the criterion of (S/N) transfer is satisfactory, in digital data transmission it does not by itself furnish an adequate performance criterion. Usually performance is judged in terms of error rates which cannot be predicted from the (S/N) transfer for nonlinear receivers. The error rate clearly depends on the statistical distribution of the output noise. In good systems, the errors are very infrequent and are associated with rare peak noise conditions. The statistical structure of the occurrence of infrequent noise peaks and the manner in which they cause errors in FM receivers is the main subject of this paper. Some previous investigations of these effects have been carried out. For example, Bennett and Salz[8] have analyzed binary FM systems, including the effects of distortion. They derived formulas for the error rate without including the post-detection filter in their model. Since the error rates that they obtained for a well-designed *binary* system were close to the optimum obtainable for any receiver, they were able to conclude that the neglect of this filter was justified. Formulas are also available[9,10] for the probability distribution function of the instantaneous frequency of signal plus noise at the input to the post-detection filter for N-ary FM, but these equations are not very useful in predicting the performance of a practical FM system since the task of relating this distribution to the distribution at the output of the post-detection filter is apparently untractable. In a recent paper, Salz[11] considered a multilevel FM narrowband digital communications system where he included the post-detection filter in his analysis. However, the results assume that the post-detection filter did not perform significant selective processing of the detected signal.

In this paper, we shall develop a general theory from which the performance of FM receivers with arbitrary processing gain may be predicted. We shall view the conventional FM receiver, described in Section II, as a device for detecting digital signals and examine its properties in detail. In Section III, after approximating the post-detection filter by an ideal integrator, we show how clicks enter the problem.* Our assumptions and the ensuing mathematical model of the stochastic output are also stated there. The following section supplies the considerable amount of

* Cohn, Ref. (7), has also mentioned the application of the concept of clicks to explain errors in digital FM. Further, D. Schilling of Brooklyn Polytechnic Institute has called to the authors' attention that he is also investigating the relationship between clicks and error rates in FM.

mathematical detail needed to quantitatively substantiate the work of Sections V through VII. In particular, the notion of clicks will be used to explain the poor performance (compared to ideal) of this receiver to detect a large number of orthogonal signals. This phenomenon has also been mentioned by Wozencraft and Jacobs.[12] Another result of the present paper is to establish conditions under which the previous analyses reliably predict the performance of actual FM systems. The work of Refs. 8 and 11 will be supported and it will be shown that for multilevel wide-band systems the post-detection filter cannot be ignored. Finally, in Section VIII a discussion is given to suggest circumstances under which successive clicks will not be independent and an instructive example is given showing how this renders ineffective the additional selective filtering possible at the input when the frequencies are very widely spaced.

II. THE DIGITAL FM SYSTEM

A digital FM signal is readily produced by changing the frequency of an oscillator in response to a digital baseband signal. The voltage or current at the output of such an oscillator may be represented as

$$S(t) = A \cos \left[\omega_c t + \int_0^t s(t')dt' + \theta \right], \tag{1}$$

where A is a real amplitude, ω_c the angular center frequency of the oscillator, and θ is an initial phase angle. The digital information-bearing signal $s(t)$ is taken to be a piece-wise constant function of time representable as a random time series of the form

$$s(t) = \omega_d \sum_{n=0}^{n=\infty} a_n g(t - nT), \tag{2}$$

where $\{a_n, n = 0, 1, \cdots\}$ is a sequence of independent and identically distributed integer valued stochastic variables representing the data. For example, one might have $a_n = \pm 1$ with equal probability for binary systems. The function $g(t)$ is a rectangular pulse of unit amplitude and T seconds duration and ω_d is a proportionality constant relating frequency displacement to baseband signal voltage or current. The spectral properties of this FM wave have been extensively analyzed in Refs. 13 and 14.

Transmission and reception of the FM wave is accomplished as follows. The wave $S(t)$ is first processed by a transmitting filter, channel noise is added, and the result is processed again by a receiving filter assumed to be the inverse of the transmitting one. The signal is then detected *via* the limiter-discriminator and filtered at baseband before being synchronously sampled at $t = nT$ (using independent timing information) to de-

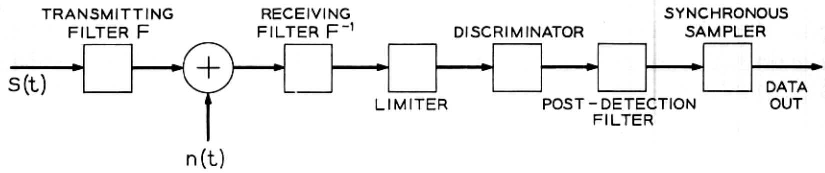

Fig. 1 — Block diagram of a digital FM receiver.

termine sequentially the values of $\{a_n\}$. We have illustrated these operations in block diagram form in Fig. 1. A detailed description of the blocks shown is given in Ref. 15. We shall state here in mathematical terms the assumed operation of the limiter-discriminator. Let the input to the limiter be written in terms of in-phase and quadrature components as

$$x'(t) \cos \omega_c t - y'(t) \sin \omega_c t \equiv R(t) \cos [\omega_c t + \varphi(t)], \tag{3}$$

where

$$R(t) = \sqrt{[x'(t)]^2 + [y'(t)]^2} \tag{4}$$

and

$$\varphi(t) = \tan^{-1} y'(t)/x'(t). \tag{5}$$

Then the output of the discriminator is taken to be

$$\frac{d\varphi}{dt} = \frac{x'(t)\dot{y}'(t) - y'(t)\dot{x}'(t)}{[x'(t)]^2 + [y'(t)]^2}, \tag{6}$$

where the dots denote differentiation with respect to time. The post-detection filter acts upon the quantity (6).

III. FORMULATION OF THE PROBLEM AND A MATHEMATICAL MODEL

We approximate the low-pass filter as an ideal integrator whose impulse response is unity for T' seconds and zero afterward. The duration T' is taken equal to the signaling time T and so no intersymbol interference occurs at the sampling times for a wave described by (1) and (2). The results obtained with this particular choice of filter should be representative of the results one would obtain with any low-pass filter of similar bandwidth.

The sampled output q' of the discriminator low-pass filter output is given by (7)

$$q' = \int_0^T \frac{x'(t)\dot{y}'(t) - y'(t)\dot{x}'(t)}{x'^2(t) + y'^2(t)}\, dt. \tag{7}$$

The in-phase and quadrature components occurring in (7) are now not those of the pure FM wave (1), but have the analogous components of zero mean noise added in as well. One may, by use of a rotating coordinate system, equally consider

$$q = q' - a_n \omega_d T = \int_0^T \frac{x(t)\dot{y}(t) - y(t)\dot{x}(t)}{x^2(t) + y^2(t)} \, dt \tag{8}$$

where $y(t)$ is now a zero mean quadrature noise process, while $x(t)$ is an in-phase noise process with mean A, the amplitude of the noise-free received FM wave. We now proceed formally with (8), defining a quantity

$$r(t) = y(t)/x(t). \tag{9}$$

Equation (8) is then rewritten as a path integral

$$q = \int_{r(0)}^{r(T)} \frac{dr(t)}{1 + r^2(t)} = \int d\varphi(t). \tag{10}$$

In (10) we have written $d\varphi = d(\tan^{-1} y/x)$, but of course we do not mean that φ is evaluated using some fixed branch of $\tan^{-1} y/x$ since this would give φ as a single valued function of y and x and would not allow for the fact that as we circle once about the origin in the xy-plane φ increases by 2π. The noise processes $y(t)$ and $x(t)$ wander about the xy-plane (see Fig. 2), usually staying close to their mean values but occasionally taking large excursions and encircling the origin. Each infinitesimal portion of the path contributes an amount $d\varphi$ volts to the output and all these small amounts from all the small portions of the path must be added together to form the total contribution q. It is easy to see that q depends on the path taken, not just on its endpoints. A simple mathematical reason for this is that the transformation (9) is undefined whenever $x(t) = 0$. Further, the paths taken in the xy-plane are random, and q is therefore, a random variable with some probability density related to the statistics

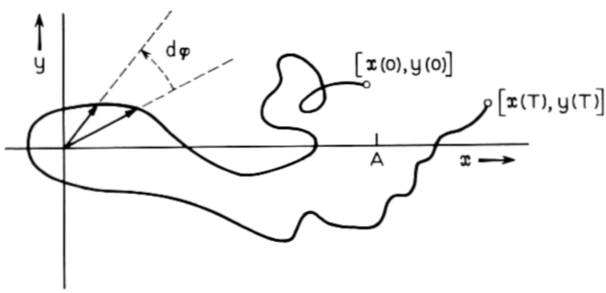

Fig. 2 — A possible path in the xy-plane traced by the noise from $t = 0$ to $t = T$.

of $r(t)$. Unfortunately, this probability density is not determined solely by the elementary statistics of $r(t)$. As will be seen, in addition to the elementary statistics of $r(t)$ the distribution of its singularities on the time axis enters the picture. The singularities of $r(t)$ are determined by the zero-crossings of $x(t)$. Thus, the behavior of FM receivers is intimately related to the structure of the zero crossings of the added noise.[16]

To see how to handle the situation, visualize the following hypothetical state of affairs. Suppose for $0 \leqq t \leqq T$ we have that $y(t) > 0$, and that $x(t)$ is positive for a while, decreases once through zero at $t = t_0$, and then remains negative. A possible plot of $r(t)$ versus t over the time interval is then shown in Fig. 3. For this particular path one has

$$q = \int_{r(0)}^{\infty} \frac{dr}{1 + r^2} + \int_{-\infty}^{r(T)} \frac{dr}{1 + r^2} = \int_{-\infty}^{\infty} \frac{dr}{1 + r^2} + \int_{r(0)}^{r(T)} \frac{dr}{1 + r^2} . \quad (11)$$

In (11) the straightforward interpretation of the integrals is meant. Evaluating the infinite integral one obtains for this path

$$q = \pi + \tan^{-1} r(T) - \tan^{-1} r(0),$$

where $\tan^{-1} x$ means the principal value inverse tangent function, $| \tan^{-1} x | \leqq \pi/2$. In general, one has the result that

$$q = \tan^{-1} r(T) - \tan^{-1} r(0) + n(T)\pi, \quad (12)$$

where $\tan^{-1} x$ again has the principal value interpretation and $n(T)$ is an integer (which may be positive, negative, or zero) which is related to the number of times $x(t)$ vanishes in the interal T and to the sign of $y(t)$ when $x(t)$ vanishes. For large signal-to-noise ratios it is clear that if $x(t)$ vanishes by going to zero from the positive side that it will almost immediately be followed by another vanishing of $x(t)$ in the other direction. If $y(t)$ has not changed, the contribution of the "return trip" to $n(T)$ will cancel the contribution from the previous crossing of the y-axis. On the other hand, if $y(t)$ does change sign so as to cause an encircling of the origin then the contribution to $n(t)$ will be the same as the previous crossing. The net contribution to $n(T)$ of a number of paths is shown in Fig. 4. The paths which have $\Delta n = \pm 2$ are immediately recognized as the "clicks" discussed by Rice.[6] The "clicks" are not the only contribution to $n(T)$ however. There is also a contribution because of the fact that at $t = 0$, when our process begins, we may be in the middle of a large noise fluctuation and be over in the left-half plane. Immediately afterwards, at $t = 0+$, we will experience a contribution of ± 1 to $n(T)$; a similar situation may prevail at time $t = T$, when a possibility exists of stopping the process immediately after we have crossed over to the left-half plane. We will show later that for large signal-to-noise ratios, these

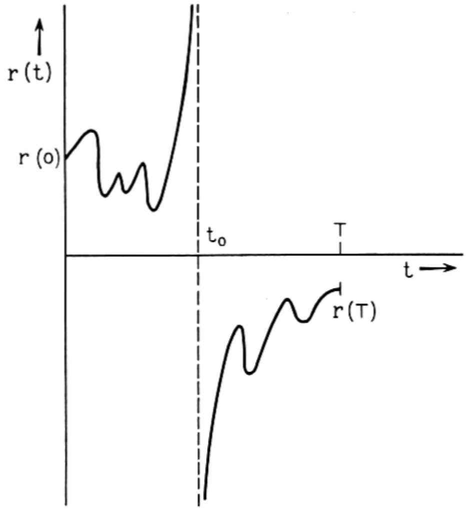

Fig. 3 — A possible sample function of $r(t)$.

end-effects may be neglected because they occur with a probability that is asymptotically negligible compared with the probability of a click.

An important fact to observe before proceeding with the analysis is that q can be decomposed into the sum of three random variables. The first two random variables appearing in (12) are continuous and bounded. Their probability densities are related to the elementary statistics of $x(t)$ and $y(t)$. The third random variable is a discrete one, whose probabilities are determined from the probabilities of zero-crossings of $x(t)$ and $y(t)$.

The remarks made thus far about the effect of noise on FM reception have been general; no assumptions have been made about the statistical nature of the additive disturbance. In order to obtain quantitative results some definite assumptions are necessary. For the remainder of the

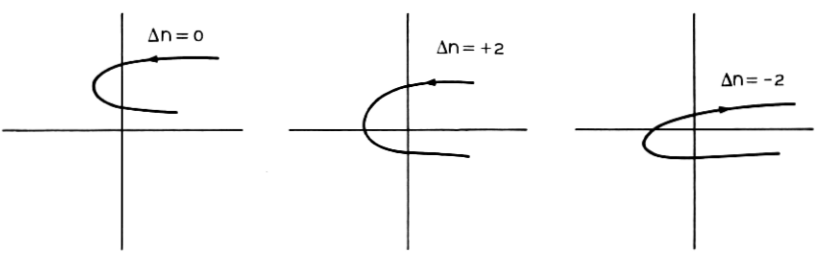

Fig. 4 — Net changes Δn in $n(T)$ caused by some typical paths in the xy-plane.

paper we shall set ourselves the task of studying the structure of the probability distribution of q when the input noise statistics are those of a Gaussian process having a symmetric spectral density about the carrier. From these distributions we determine the error rates as a function of the pertinent system parameters.

No attempt will be made in this paper to derive an exact probability density for the random variable q. This is not a mathematically tractable problem since it requires knowledge of the probability distribution of zero-crossings of random processes. This by itself has been an area of investigation for many years without too much success. The probability distribution of the zero-crossings of most elementary random processes is not currently known.

In order to permit an analysis of the model two assumptions are made, both of which we feel are quite reasonable. These two assumptions taken together state that the three random variables that determine q *via* (12) are all independent. We separate this statement into two assumptions because their individual justification stems from two different physical arguments, one having to do with bandwidth and the other with signal-to-noise ratio. The first assumption states that $\tan^{-1} r(T)$ and $\tan^{-1} r(0)$ are independent. For a flat Gaussian noise input this will be a good approximation if $T \geqq 1/W$, where W is the input noise bandwidth. Since T is also the signaling interval, and the correlation function of the input noise has its first zero at $t \sim 1/W$, the motivation for this assumption is clear. The second assumption, somewhat harder to justify, states that $n(T)$ is independent of the previous two random variables, and the clicks, which comprise $n(T)$, are independent from one another. This is clearly an assumption expressing an intuitive feeling that the clicks occur rarely and of sufficiently short duration. In general, they will be rare if the signal-to-noise ratio is large, and short if the bandwidth satisfies $W \geqq 1/T$ as required above.

These two assumptions plus the identification of crossings of the negative x-axis by the moving point in the xy-plane (as calculated by Rice) with the occurrence of a click shall constitute our working model of the output noise. An indication of how this model must be modified if the input noise spectrum is not relatively flat is given in Section VIII.

IV. THE BASIC DISTRIBUTIONS

Let y be a Guassian variable of zero mean, variance σ^2, and x be another independent Gaussian variable of mean A, variance σ^2.* Then the den-

* Recall that even though our $x(t)$ and $y(t)$ are not independent processes because the noise spectrum will not be symmetrical about $(\omega_c + a_n\omega_d)$, they are independent variables.

sity $\bar{p}(\bar{\varphi})$ where $\tan \bar{\varphi} = y/x$ and $\bar{\varphi}$ has the full range of 2π is well known and is given by Bennett,[17]

$$\bar{p}(\bar{\varphi}) = \frac{\exp(-\rho)}{2\pi} + \tfrac{1}{2}\sqrt{\frac{\rho}{\pi}} \cos \bar{\varphi} \exp(-\rho \sin^2 \bar{\varphi}) \cdot [1 + \operatorname{erf}(\sqrt{\rho} \cos \bar{\varphi})], \tag{13}$$

where $\rho = A^2/2\sigma^2$.

One fact which is implicitly contained in (13) is the probability P_L of finding the signal point in the left half of the xy-plane. However, an easier way to obtain P_L is as follows:

$$P_L = Pr(x < 0) = \tfrac{1}{2} \operatorname{erfc} \frac{A}{\sqrt{2}\,\sigma} \sim \frac{\exp(-\rho)}{2\sqrt{\pi}\sqrt{\rho}}. \tag{14}$$

Equation (14) will be of use in the arguments used to discard the "end effects" at $t = 0$ and $t = T$ spoken of earlier. Equation (13) also immediately yields the probability density $p(\varphi)$ for $\varphi = \tan^{-1}(y/x)$, $-\pi/2 < \varphi < \pi/2$. Indeed, we have

$$\begin{aligned} p(\varphi) &= \bar{p}(\varphi) + \bar{p}(\varphi + \pi) \\ &= \frac{\exp(-\rho)}{\pi} + \sqrt{\frac{\rho}{\pi}} \cos \varphi \exp(-\rho \sin^2 \varphi) \operatorname{erf}(\sqrt{\rho} \cos \varphi) \end{aligned} \tag{15}$$

for $|\varphi| \leqq \pi/2$.

Suppose φ_1 and φ_2 are two independent angles which have the density (15), and define an angle $\Phi = \varphi_1 - \varphi_2$, $|\Phi| \leqq \pi$. It will be of interest for us to determine the probability P_φ that Φ exceeds some angle $\varphi > 0$, i.e., we would like to determine

$$P_\varphi = \int_{-\pi/2}^{(\pi/2)-\varphi} d\varphi_2 \int_{\varphi_2+\varphi}^{\pi/2} d\varphi_1 p(\varphi_1) p(\varphi_2), \qquad \varphi > 0. \tag{16}$$

In general, one is unable to perform these integrations exactly, but since discussion has already been limited to the large S/N region, little will be lost if we make use of this in simplifying the evaluation of (16). The asymptotic evaluation is carried out in detail in the Appendix; we distinguish three cases:*

Case I; $0 < \varphi < \pi/2$:

$$P_\varphi \sim \frac{1}{\sqrt{8\pi}} \frac{\cot(\varphi/2)}{\sqrt{\cos \varphi}} \frac{\exp[-2\rho \sin^2(\varphi/2)]}{\sqrt{\rho}}. \tag{17a}$$

* In (17) the symbol "$\sim$" is used to denote asymptotic equality; this has also been used in (14). Also (17a) and (17c) do not hold if φ gets too close to the end points of the appropriate interval. As a rough rule, φ should not be closer than $1/\sqrt{\rho}$ radians to the end points.

Case II; $\varphi = \pi/2$:

$$P_\varphi \sim (\tfrac{1}{4}) \exp(-\rho). \tag{17b}$$

Case III; $\varphi > \pi/2$:

$$P_\varphi \sim \frac{\exp[-\rho(1 + \cos^2 \varphi)]}{2\pi\sqrt{\pi}\,\rho\sqrt{\rho}\,\sin\varphi\,\cos^2\varphi}. \tag{17c}$$

The most important characteristic of the result (17) is the dependence of the exponent on angle, since for large ρ the nonexponential factors are relatively slowly varying.

We should remark that for very small angles (15) is well approximated by the Gaussian curve

$$g(\varphi) = \sqrt{\frac{\rho}{\pi}} \exp(-\rho\varphi^2) \tag{18}$$

of zero mean and variance $1/2\rho$. The difference angle Φ would, for very small Φ, be well approximated by the difference of two independent Gaussian variables, each having the density (18). The quantity P_φ calculated on this basis agrees (asymptotically) with the small angle approximation of (17a).

The final item that we discuss in this section is the density of $n(T)$, or rather we discuss the density of that part of $n(T)$ that arises from the clicks ($\Delta n = \pm 2$), ignoring $\Delta n = \pm 1$ contributions. For this we need only take over some ideas and formulas from Rice.[6] We have that (ignoring $\Delta n = \pm 1$)

$$\pi n(T) = 2\pi N(T), \tag{19}$$

where $N(T)$ is the number of clicks that occur in time T. Following Rice, we assume that all clicks are independent and that those tending to increase (decrease) φ by 2π form a Poisson process with rate of occurrence $N_+(N_-)$. In general, with a modulated signal, N_+ and N_- are not equal. The probability density $p(z)$ of $z = N(T)$ is then given by

$$p(z) = \exp[-(N_+ + N_-)T] \sum_{k=-\infty}^{\infty} \delta(z - k) \left(\frac{N_+}{N_-}\right)^{k/2} \cdot I_k(2T\sqrt{N_+N_-}); \tag{20}$$

as may be shown by forming the discrete convolution of the densities of the positive and negative clicks. In (20) $\delta(\cdot)$ is the Dirac delta function and $I_k(\mu)$ is the modified Bessel function of integer order k, behaving for small μ as[18]

$$I_k(\mu) \underset{\mu \to 0}{\longrightarrow} \left(\frac{\mu}{2}\right)^{|k|} \frac{1}{|k|!}; \tag{21}$$

also

$$I_{-k}(z) = I_k(z).$$

The type of modulation that we are concerned with is when the instantaneous frequency deviates by ω_d from the carrier* for a time T, T being the signaling and processing interval. For this situation Rice gives for the average rates N_+ and N_- when the noise at the receiver input is Gaussian

$$\begin{aligned} N_+ = \tfrac{1}{2}\{\sqrt{r^2 + f_d^2}\,[1 - \operatorname{erf}\sqrt{\rho + \rho f_d^2/r^2}] \\ - f_d \exp(-\rho)[1 - \operatorname{erf}(f_d\sqrt{\rho}/r)]\} \end{aligned} \tag{22}$$

and

$$N_- = N_+ + f_d \exp(-\rho), \tag{23}$$

where†

$$\begin{aligned} r &= (1/2\pi)(\dot{\sigma}/\sigma) \\ \sigma^2 &= \operatorname{var} x = \operatorname{var} y \\ \dot{\sigma}^2 &= \operatorname{var} \dot{x} = \operatorname{var} \dot{y}. \end{aligned} \tag{24}$$

Under the assumption that f_d is positive we have asymptotically

$$\begin{aligned} N_+ &\sim \frac{1}{4\sqrt{\pi}} \frac{1}{\rho^{3/2}} \frac{r}{\left(1 + \frac{f_d^2}{r^2}\right)\left(\frac{f_d^2}{r^2}\right)} \exp[-\rho(1 + f_d^2/r^2)] \\ N_- &\sim N_+ + f_d \exp(-\rho). \end{aligned} \tag{25}$$

Thus, we see that for large ρ an ever greater majority of clicks occur in the negative direction ($f_d > 0$) and for our purposes of computing error rate the clicks in the positive direction may be neglected; i.e., we shall use

$$\left.\begin{aligned} N_+ &\sim 0 \\ N_- &\sim f_d \exp(-\rho) \end{aligned}\right\} \quad \text{for} \quad f_d > 0. \tag{26}$$

* We trust that no confusion will arise between r introduced in (24) and $r(t)$ introduced in (9).

† The case $\omega_d = 0$ corresponds to no modulation. Also, for ease of writing, we no longer explicitly consider the factor a_n.

For $f_d < 0$ the situation is reversed of course. We note that the effect of the clicks on a modulated carrier is to tend to make the measured frequencies appear closer to the carrier frequency than the transmitted frequencies. That is, confining oneself for the moment to *only errors caused by clicks*, frequencies transmitted higher (lower) than the carrier will be measured to be at that frequency or a lower (higher) one, when the noise is small.

Since we shall use approximation (26), the distribution (20) for $z = N(T)$ may be replaced by the simpler Poisson one, where the probability of getting exactly N (negative) clicks in time T is given by*

$$p[N(T)] = \frac{\exp(-N_-T)(N_-T)^{N(T)}}{[N(T)]!}. \tag{27}$$

Also the probability of getting M or more clicks is, for large signal-to-noise ratios, approximately the probability of getting exactly M clicks.

V. DISTRIBUTION OF OUTPUT AND PROBABILITY OF ERROR

Equations (14), (17), (26), and (27) provide the information required to calculate the distribution of q, (12). In principle we simply convolve the continuous density of $[\tan^{-1} r(T) - \tan^{-1} r(0)]$ with the discrete density of $n(T)\pi$. In Fig. 5, we have given a qualitative sketch of the result, neglecting end effects. This picture is intended to show that the density consists of a central lobe about the transmitted frequency extending to $\pm\pi$ on each side, which is the density of $[\tan^{-1} r(T) - \tan^{-1} r(0)]$, plus identically shaped lobes displaced by integral multiples of 2π toward lower frequencies (assuming $f_d > 0$). These displaced lobes are weighted by the probability of getting the appropriate number of clicks to effect the displacement. Thus, the lobe occupying the space $-2n\pi \pm \pi$ is weighted by the probability of getting exactly n clicks in time T. For $n = 0$ the weighting is essentially one, for large S/N. There are, strictly speaking, similar lobes and weightings on the opposite side as well, but these weights are, for large S/N, negligible compared to the *corresponding* lobe we have drawn. That is to say, the first lobe on the right (not shown in Fig. 5) has small probability compared to the first lobe on the left, but has a large probability compared to the second lobe on the left. Nevertheless, we have neglected to include it because we will generally be concerned with probabilities like $\Pr[\,|\,q - f_dT\,| > \varphi]$, and thus corresponding weights are important. We dwell on this point be-

* We confine ourselves to $f_d > 0$. Exactly analogous consideration apply to $f_d < 0$. The case $f_d = 0$ occurs if an odd number of frequencies are allowed.

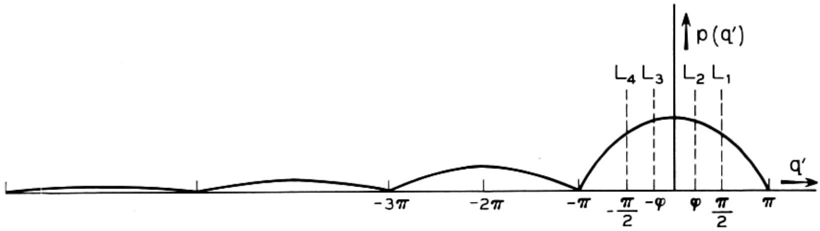

Fig. 5 — Qualitative sketch of density of q' (neglecting end effects) for $f_d > 0$. The dashed lines are for reference in the text.

cause it is conceivable that for some practical or conceptual application the neglect would not be justified.

The discussion given above is still not quite correct; it is modified when we include end effects. The principle correction that inclusion of end effects will cause is to add two more side lobes, one over the interval $[-2\pi,0]$ and the other over the interval $[0,2\pi]$. The weightings of these lobes certainly should not exceed the estimate given in (14), and this will be enough to exclude them for our purposes.

We now apply our results to some typical calculations. Consider the case of narrow band* FM (defined by $\Delta f_d T < \pi$), where one has J equally spaced frequencies of separation Δf_d crowded into a bandwidth W. The probability of error for any *one* of the frequencies† (not situated at the ends) is the area outside of the interval bounded by lines L_2 and L_3 in Fig. 5. If L_2 and L_3 are defined by $|q| = \varphi$ then the probability of error for such a frequency is, from (17a),

$$Pe = \frac{1}{\sqrt{2\pi}} \frac{\cot \varphi/2}{\sqrt{\cos \varphi}} \frac{\exp\left[-2\rho \sin^2 (\varphi/2)\right]}{\sqrt{\rho}}, \tag{28}$$

where, if one assumes that the bandwidth $W = J\Delta f_d$, one would take

$$\varphi = \frac{\pi W T}{J}. \tag{29}$$

Our requirement that $\Delta f_d T < \pi$ implies $J > 2$ for the narrow-band formula to be applicable (assuming $WT = 1$). Note $\sin^2 (\varphi/2)$ is less than $\frac{1}{2}$, and thus the exponent in (28) is $\exp[-k\rho]$, where $k < 1$. Now the contribution of the clicks to P_e is essentially the area A_L of the first side lobe which is by (26) and (27)

$$A_L = f_d T \exp(-\rho). \tag{30}$$

* Note the special sense in which the term is uesd here.
† The P_e for a frequency at the end is one-half the expression (28).

But expression (30) is, asymptotically, exponentially small compared to (28). Likewise, the area due to the side lobes caused by end effects is exponentially small, and the probability of error for narrow-band FM is given by (28). The result that the clicks do not asymptotically contribute to errors in narrow-band multilevel FM lends justification to a previous evaluation of this type system by Salz,[11] who considered the special narrow-band system with $WT = 1$. It is both interesting and gratifying that this result is in agreement with the result given in Ref. 11. In a later paper, Salz and Koll[19] report on experimental results which agree with the earlier theoretical results.

Next, consider the asymptotic evaluation of P_e for the case of orthogonal signals; this case corresponds to $(\Delta\omega_d)T = \pi$, and we assume that the thresholds are spaced midway between the frequencies. Thus, (for a frequency not on the edges) we have that the error probability is given by the area outside of that bounded between the lines L_1 and L_4. The contribution from the major lobe is, from (17b),

$$\tfrac{1}{2} \exp(-\rho).$$

In addition, the area of the first side lobe is asymptotically comparable to this and is

$$f_d T \exp(-\rho),$$

being weakly dependent on the frequency sent. In fact, for the nth signal $(J = 2n)$ we have for orthogonal signals that

$$f_d T = \frac{n}{2}, \qquad n = 1,2,\cdots,\frac{J}{2}.$$

The average error rate is then, for orthogonal signals (J of them, J even, and equally spaced signals and thresholds),

$$P_e = (\tfrac{1}{2}) \exp[-\rho] + (\tfrac{1}{4})(J/2 + 1) \exp(-\rho). \tag{31}$$

Equation (31) is indeed a surprising result. The first term of (31) is the probability of confusing the transmitted frequency with one of its nearest neighbors. The second term is the (average) probability of confusing it with its second nearest neighbor closest to the carrier. This is because the area from $(-\pi)$ to $(-3\pi/2)$ is, by application of (17b), negligible compared to the area from $(-3\pi/2)$ to $(-5\pi/2)$. Thus, it states that for the multilevel scheme considered here (a not unreasonable one) one is less likely to confuse a transmitted frequency with its nearest neighbors than one is to confuse it with a particular one of its second nearest neighbors. We see from (31) that the error rate from the continuous part of the output is comparable to the error rate caused by clicks.

As a final remark about the orthogonal system we see comparing (31) and (14) why end effects are neglected again.

For a final example, consider the wide-band situation where the signals are loosely packed in the band; i.e., $(\Delta\omega_d)T > \pi$. Now no errors will be caused by the continuous part of the output; only clicks will cause errors. If the frequencies are widely spaced a single click may not cause an error; several clicks during the time interval T may be required. Thus, suppose that the frequencies are spaced so that the phase differences of nearest neighbors is $(\Delta\omega_d)T = 2n\pi$, n being any positive integer. The probability of error will then be the probability of getting n (or more) clicks in time T, which from (26) and (27) behaves as

$$\begin{aligned} \frac{(f_d T)^n \exp(-n\rho)}{n!} &\geqq \frac{(n/2)^n \exp(-n\rho)}{n!} \\ &\geqq \tfrac{1}{2} \exp(-n\rho). \end{aligned} \tag{32}$$

The coefficient in (32) is at least as bad as for the orthogonal case, but the important item is the exponent. Superficially at least it appears that we have gained in performance by spacing the frequencies widely, since the exponential has changed from $e^{-\rho}$ from the minimum orthogonal case ($\Delta\omega_d T = \pi$) to $e^{-n\rho}$. One must realize, however, that one is talking about different ρ's here. The bandwidth for the case under consideration is essentially $2n$ times the minimum orthogonal one and therefore, for the same signal power, the nominal value of ρ has decreased $2n$, and one has in fact not gained a factor of n in the exponent. In addition to the bandwidth penalty, error performance has actually suffered too.

VI. COMPARISON WITH OPTIMUM

One can demonstrate how the FM discriminator compares with the optimum detector when used to detect orthogonal signals; i.e., when $\Delta\omega_d T = \pi$. It is known that when optimum detection is used for any orthogonal set of signals, the (exponential part of the) error rate behaves as $\exp[-E/N_0]$, where E is the signal energy (assumed common to all J levels) and $N_0/2$ is the (two-sided) spectral density of the noise. If we let S denote the average signal power, write $E = ST$, and estimate the total bandwidth W for large J by $W = J/(2T)$, we see that the ideal exponent becomes $\exp[-J\rho/2]$. However, we had seen that, regardless of the number of levels, the discriminator error rate for $\Delta\omega_d T = \pi$ behaves as $\exp[-\rho]$. Thus, we have lost a factor of J in the error exponent by substituting discriminator detection for matched filter detection.

An important conclusion may immediately be drawn concerning the performance of conventional FM receivers or detectors of orthogonal signals. Our results show that the receiver is indeed inferior in performance when compared with the optimum. This fact has been stated by Wozencraft and Jacobs[12] and the reasons are clear from our analysis. The FM receiver admits too much noise at its front-end which cannot be cleaned by the post-detection filter because of the nonlinear anomalies, namely the clicks. As a matter of fact, the amount of noise grows in direct proportion to the number of orthogonal signals, hence the inferior exponent.* The optimum detector is a bank of matched filters. The noise power at the output of each filter does not grow with the number of signals; it is a fixed constant determined by the bandwidth of the filter, which roughly needs be no broader than the symbol rate, $1/T$.

This poor performance of conventional FM receivers when used to detect data *might* be remedied by employing an FM with feedback system such as described in Refs. 20 and 21. The physical argument to support this contention is often stated as follows. In the absence of the feedback loop, the IF filter must be wide enough to pass the total swing of the incoming signal. However, since the feedback loop tracks the incoming frequency, this IF filter, whose width determines the noise variance, could be narrowed and less noise would be admitted.

This possibility of making use of FM with feedback to improve the error rate in digital systems has been suggested by Wozencraft and Jacobs.[12] Unfortunately a mathematical treatment of this difficult problem does not exist at present.

VII. EFFECT OF POST-DETECTION FILTER

In the previous sections we have discussed in detail the performance of an FM discriminator followed by a low-pass filter. The low-pass filter was approximated by an ideal integrator whose integration time was taken to be equal to the duration of an individual signaling interval. Formulas sufficient to evaluate the performance of multilevel FM without the post-detection filter have recently been developed by Mazo and Salz;[10] comparison of the results of the present paper with Ref. 10 will show the influence of filtering.

* Actually, these qualitative conclusions may be arrived at by the Gaussian approximation to the output noise. The reason why this works is apparent from (31) which gives P_e for orthogonal signals. The first term of (31) is not due to clicks but arises from the continuous part of the output noise. This is the part that the Gaussian approximation would tend to duplicate. The second term of (31) is due to clicks and has the same behavior with regard to ρ. Even if one could keep ρ constant as the number of levels J increased, the factor of J in the click contribution to (31) would still degrade performance.

Suppose that the angular frequency $\dot{\psi}$ is sent and we ask for the probability that the observed output is less than z, where $(\dot{\psi} - z) > 0$. It is shown in Ref. 10 that the probability P is essentially given by* (for large ρ)

$$P \cong \exp\left[-\rho \frac{(\dot{\psi} - z)^2}{z^2 + \dot{\sigma}^2/\sigma^2}\right]. \tag{33}$$

Consider the situation for orthogonal signals, or in fact for any signal set where the frequency spacing between the individual frequencies is fixed. One expects the ratio $\dot{\sigma}^2/\sigma^2$ to increase as the square of the total input bandwidth, hence as J^2, the square of the number of levels. Thus, for a large number of orthogonal levels the post-detection filter does very well in improving the error performance, changing the error rate from† (roughly) $\exp(-\rho/J^2)$ to $\exp(-\rho)$. One would certainly expect something like this to be true since, for a large number of levels, the noise bandwidth before the post-detection filter is much greater than the signal bandwidth at that point.

Another qualitative effect of the post-detection filter may be noted. From (33) we see that the distribution of output noise without the post-detection filter depends on the frequency sent, because of the factor $(z^2 + \dot{\sigma}^2/\sigma^2)$ in the exponent; the "spread" of the probability density will be roughly twice as great at the ends of the band than at the center, and thus without a post-detection filter one would not choose the frequencies to be equally spaced. We have seen that there is no such dependence of the error rate exponent on the transmitted frequency when the post-detection filter is present.

VIII. AN APPARENT PARADOX

At this point we have basically concluded our discussion of error rates in digital FM, based in part upon the theory of "clicks" in FM receivers. In particular, we have seen in Section VI that even when frequencies were widely spaced so that $\omega_d T$ is many multiples of 2π the error performance did not improve. The reason was noted to be that although the distance between frequencies increased, the noise admitted to the system increased by a corresponding factor. The latter is predicated on the assumption that the input bandpass filter is essentially a flat filter up to some cutoff frequency determined by the signal spectrum. It may be possible, however, to shape the front-end filter so that increasing the frequency separation does not cause a proportionate increase in the

* Equation (33) represents only the exponential part of P. Also (33) is true [see Ref. 10] only if $(\dot{\psi} - z)^2/[z^2 + \dot{\sigma}^2/\sigma^2] < 1$.

† Set $(\dot{\psi} - z)^2 \approx (\Delta f)^2$, $(\dot{\sigma}^2/\sigma^2) \approx J^2(\Delta f)^2$.

noise power admitted. We know that the power spectrum of the transmitted signal will have peaks at the transmitted frequencies of width of the order $(1/T)$. Suppose we have a notch filter then, with transmittance peaks at the possible frequencies of the appropriate width. The input noise power will be constant and therefore by choosing a large enough separation one can force the probability of error to be arbitrarily small, contradicting optimality considerations for reception of signals against a white Gaussian noise background.

Before giving what we feel is the correct answer to the stated paradox, we wish to explore some other considerations which, on the surface, might resolve the paradox without changing the basic assumptions of the model. One might first object that our argument was too heuristic; is the noise power really constant as the frequency separation increases? To answer this we have performed the following calculations. We have chosen transmitting and receiving filters so that the FM signal is strictly undistorted and then optimized the filters to minimize the variance of the noise admitted. This procedure is discussed in Ref. 11, and the results depend on the power spectrum of the noise. We then specialize to a binary system and, using (48) of Ref. 13 for the spectral density of a binary FSK wave train, calculate the noise admitted. The result shows that while the noise admitted does, in fact, increase as the frequency separation increases, it does so only logarithmically with the separation. Thus, the error probability still will decrease to an arbitrarily small value as the separation increases and from this point of view the question is still unresolved.

A second consideration is the following. The probability of error that we have calculated was based on asymptotic approximations to formulas given in Ref. 6. The results depended only on the amplitude of the received FM wave and the average noise power σ^2 at the input to the limiter-discriminator; if one allows transmitting and receiving filters the more relevant parameters are the average signal power on the line, P_{line}, and σ^2. However, the exact formulas of Rice also involve the quantity $\dot{\sigma}^2$ which is the average power in the derivative of the noise at the input to the limiter-discriminator (after the receiving filter).* Let $S(\omega)$ be the signal spectral density and $F(\omega)$ the transmittance of the receiving filter. Further, let us insist that the signal at the input to the limiter-discriminator be exactly the FSK wave described,† so the transmitting filter is the inverse of the receiving filter. We then have for a white noise background of N_0‡

* Rice, Ref. 6, uses the parameter $r = (1/2\pi)\ (\dot{\sigma}/\sigma)$.

† We emphasize that continuous phase at frequency transition is demanded, but nothing more.

‡ It is for such a noise background that the optimum results are known.

$$\sigma^2 = \frac{N_0}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2\, d\omega \tag{34a}$$

$$\dot{\sigma}^2 = \frac{N_0}{2\pi}\int_{-\infty}^{\infty} \omega^2\, |F(\omega)|^2\, d\omega \tag{34b}$$

$$P_{\text{line}} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{S(\omega)}{|F(\omega)|^2}\, d\omega. \tag{34c}$$

When one realizes that the spectrum of an FSK wave decreases at infinity like the fourth power of the frequency,[13] (34b) and (34c) imply that $\dot{\sigma}^2$ and the line power P_{line} cannot both be finite. Thus, suppose $\dot{\sigma}^2$ is finite. The convergence of the integral in (34b) implies that $|F(\omega)|^2$ must decrease at least like $1/\omega^{3+\varepsilon}$, $\varepsilon > 0$. The integral for P_{line} will, for large ω, look like

$$\int \frac{1}{\omega^4}\cdot \omega^{3+\epsilon} d\omega$$

which diverges. Likewise, the assumption of finite line power implies $\dot{\sigma}^2$ is infinite. An infinite $\dot{\sigma}^2$ certainly violates the conditions under which the asymptotic results of Rice's formulas hold. In particular, these formulae show that an infinite $\dot{\sigma}^2$ corresponds to an infinite average number of clicks per second (assuming such a language is still possible) and the FM discriminator will not work, in the strict sense. On the other hand, if we choose the evil of infinite line power then perfect performance is not surprising.

While the above theorem about $\sigma, \dot{\sigma}$, P_{line} is true from a mathematical point of view, it is almost irrelevant from an engineering point of view because it involves discussions of infinitely large frequencies, and does not really eliminate the paradox at all. We need merely precede the limiter-discriminator with a flat filter with a cutoff so high that the signal is *almost* undistorted. Since real discriminators work, this is not an unreasonable thing to assume. Now $\dot{\sigma}$ is finite, and although we may have to go to extremely large S/N ratios, the paradox is as entrenched as ever.

The resolution of the problem lies in a reinterpretation of Rice's calculation of the average number of crossings of the negative x-axis. We had assumed each crossing corresponds to an encirclement of the origin which is independent of all past and future encirclements. This is reasonable when the receiving filter is essentially flat across the whole received spectrum and the correlation time out of the receiving filter is small ($\sim 1/W$). However, if the input noise spectrum is chopped into a few slits or notches, correlations in the noise being processed in the detector will persist for a longer time and multiple encirclements of the origin can occur with essentially the same probability that one would normally associate with a single large excursion close to the origin.

To make our arguments more precise we consider a binary situation at almost zero rate, i.e., we have very narrow filters F_1 and F_2 about the frequencies $(\omega_c + \omega_d)$ and $(\omega_c - \omega_d)$, respectively. The bandwidth of these individual filters is of order $1/T$. The noise out of F_1 and F_2 can be written as

$$\begin{aligned} n_1(t) &= n_{1x}(t) \cos (\omega_c + \omega_d)t - n_{1y}(t) \sin (\omega_c + \omega_d)t \\ n_2(t) &= n_{2x}(t) \cos (\omega_c - \omega_d)t - n_{2y}(t) \sin (\omega_c - \omega_d)t, \end{aligned} \tag{35}$$

where $n_{1x}(t)$, etc., are independent baseband noise currents. If we assume that the frequency $(\omega_c + \omega_d)$ is being transmitted with amplitude A, then in a "coordinate system" following that frequency we have

$$\begin{aligned} x &= A + X \\ y &= Y, \end{aligned} \tag{36}$$

where

$$\begin{aligned} X &= n_{1x} + n_{2x} \cos 2\omega_d t + n_{2y} \sin 2\omega_d t \\ Y &= n_{1y} + n_{2y} \cos 2\omega_d t - n_{2x} \sin 2\omega_d t. \end{aligned} \tag{37}$$

A typical portion of the path that the noise traces out in the xy plane can be calculated from (36) and (37) and is shown in Fig. 6. Neglecting the time variations of $n_{1x}(t)$, etc., which vary on a time scale comparable

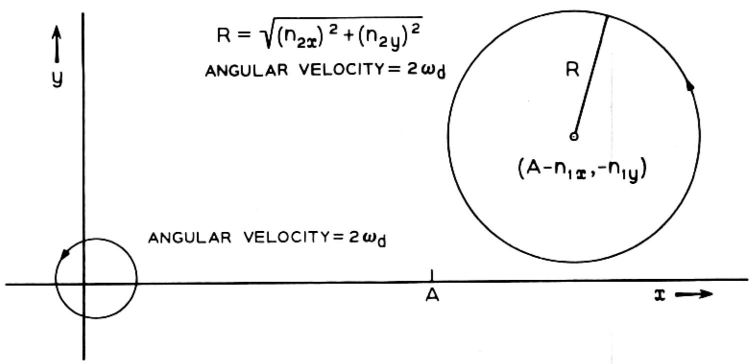

Fig. 6 — Small portions of some noise trajectories when receiving filter has two transmittance peaks.

to T, we see the path is a circle centered at $(A - n_{1x}, -n_{1y})$, of radius $\sqrt{n_{2x}^2 + n_{2y}^2}$, and counter-clockwise angular velocity of $(2\omega_d)$. If σ_1^2 and σ_2^2 denote the average noise powers out of F_1 and F_2, respectively, then the probability P that the circle is appropriately situated with a large enough radius to encircle the origin is given exactly by

$$P = 2f_d \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \exp(-\rho), \tag{38}$$

where $\rho = A^2/2(\sigma_1^2 + \sigma_2^2)$. For the case of a symmetrical spectrum about the carrier ($\sigma_1^2 = \sigma_2^2$), (38) is comparable to (26). However, our circle is rotating with frequency $2f_d$ and will have a constant radius for about T seconds; thus, it will complete $2f_dT$ revolutions in time T. As the frequencies are spread and notched filters are used the noise indeed does not increase proportionally, but the number of multiple encirclements of the origin that a click will make does increase as the separation. Thus, the filter shaping under discussion will affect the statistical structure of the clicks, preventing a violation of optimality.

Note added in Proof. A discussion of the click contribution to the error rate has been given very recently by J. Klapper in the RCA Review, June, 1966.

APPENDIX

Asymptotic Behavior of P_φ

We wish to record here an outline* of the details of the evaluation of (16) for large S/N so as to obtain the results given in (17). If we set

$$p(x) = \frac{1}{\pi} \exp(-\rho) + \sqrt{\frac{\rho}{\pi}} \cos x \exp(-\rho \sin^2 x) \cdot \operatorname{erf}(\sqrt{\rho} \cos x), \tag{39}$$

then according to (16) the required probability is written as

$$P_\varphi = \int_{-(\pi/2)}^{(\pi/2-\varphi)} dy \int_{(y+\varphi)}^{(\pi/2)} dx p(y) p(x). \tag{40}$$

If we define the distribution function

$$P(\xi) = \int_{-(\pi/2)}^{\xi} p(y) dy \tag{41}$$

* We do not explain the techniques used here for the asymptotic evaluation of integrals. The interested reader may wish to consult the subjects "saddle point method," "Laplace's method," "Watson's lemma" in Ref. 22.

and perform and integration by parts, (40) becomes

$$P_\varphi = \int_{-(\pi/2)}^{\pi/2-\varphi} P(y)p(y+\varphi)dy. \tag{42}$$

Our evaluation will be based upon approximating the functions $P(y)$ and $p(y)$ when ρ is large. In particular, from (39) we see

$$p(y) \sim \sqrt{\frac{\rho}{\pi}} \cos y \exp(-\rho \sin^2 y), \tag{43}$$

provided y is not close to $\pm\pi/2$. Integrating (43) yields

$$P(y) \sim \tfrac{1}{2}[\operatorname{erf} \sqrt{\rho} + \operatorname{erf}(\sqrt{\rho} \sin y)], \tag{44}$$

which will be a good approximation for large ρ except when y is near $-\pi/2$. These exceptional points will receive special consideration.

As a first example consider the case when $\varphi < \pi/2$. The integrand for (42) is shown symbolically in Fig. 7. Consider the contribution first from negative y. This is from (42), (43), and (44)

$$\frac{1}{2}\int_{-(\pi/2)}^{0} dy[\operatorname{erf}\sqrt{\rho} - \operatorname{erf}(\sqrt{\rho}\sin|y|)] \sqrt{\frac{\rho}{\pi}} \cos(y+\varphi) \cdot \exp[-\rho \sin^2(y+\varphi)]. \tag{45}$$

Next, approximate $\operatorname{erf}\sqrt{\rho}$ by unity to obtain

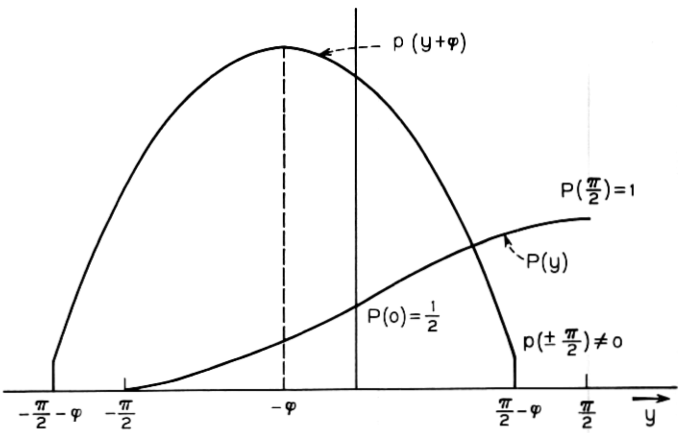

Fig. 7 — Symbolic representation of the factors in the integrand of (42) drawn for $\varphi < \pi/2$.

$$\frac{1}{2}\int_{-(\pi/2)}^{0} dy\, \text{erfc}\,(\sqrt{\rho}\sin|y|)\sqrt{\frac{\rho}{\pi}}\cos(y+\varphi) \cdot \exp[-\rho\sin^2(y+\varphi)] \tag{46}$$

and use the asymptotic expansion

$$\text{erfc}\, x \sim \frac{1}{\sqrt{\pi}x}\exp(-x^2). \tag{47}$$

The resultant integrand has a saddle point at $y = -\varphi/2$, and a routine saddle point evaluation will yield (17a) of the text. It is easy to verify that the error made by replacing erf $\sqrt{\rho}$ by unity in (45) creates an asymptotically small error. Likewise the neglect of positive y is asymptotically small for

$$\begin{aligned}\int_0^{(\pi/2)-\varphi} dyP(y)p(y+\varphi) &\leqq \int_0^{(\pi/2)-\varphi} dy\,p(y+\varphi) \leqq \left[\frac{\pi}{2}-\varphi\right]\left(\frac{1}{\pi}\right) \\ &\cdot\exp(-\rho) + \int_0^{(\pi/2)-\varphi} dy\sqrt{\frac{\rho}{\pi}} \\ &\cdot\cos(y+\varphi)\exp[-\rho\sin^2(y+\varphi)] \\ &\sim \frac{\exp[-\rho\sin^2\varphi]}{2\sqrt{\pi\rho}\sin\varphi},\end{aligned} \tag{48}$$

by Laplace's method. For $\varphi < \pi/2$ we have

$$2\sin^2(\varphi/2) < \sin^2\varphi$$

which proves our point. The addition of the term $(1/\pi)\exp(-\rho)$ in (48) provides a strict upper bound to $p(y+\varphi)$ and thus takes care of special considerations at the right end of $p(y+\varphi)$. At the left end point of the range of integration, $(-\pi/2)$, $p(y+\varphi)$ is still well approximated. The function $P(y)$ is, however, approximately

$$P(y) \approx \frac{\exp(-\rho)}{\pi}\left[\frac{\pi}{2}+y\right], \qquad y \approx -\frac{\pi}{2}. \tag{49}$$

Using (49) it is easy to obtain an estimate of the contribution of the left end point behaving as $\exp(-\rho)$ and this is asymptotically small. This ends our discussion for $\varphi < \pi/2$.

We give a somewhat more condensed outline for $\varphi = \pi/2$. The contribution of the middle of the range of integration is again approximated by (46) with $\varphi = \pi/2$. Using (47) and (46) immediately evaluates to

$$[\exp(-\rho)]/4.$$

Next, consider the error made at the right end point. Equation (43) holds to within a strip of order $1/\sqrt{\rho}$ from 0, after which $p(y + \varphi)$ behaves like $[\exp(-\rho)]/\pi$. Therefore, the error behaves like

$$\frac{1}{2} \cdot \frac{\exp(-\rho)}{\pi} \cdot \frac{1}{\sqrt{\rho}},$$

which is asymptotically small.

The left end point error is bounded by

$$-\int_{-(\pi/2)}^{0} \frac{\exp(-\rho)}{\pi} \left(\frac{\pi}{2} + y\right) \sqrt{\frac{\rho}{\pi}} \sin y \exp(-\rho \cos^2 y) dy \sim \frac{\exp(-\rho)}{\pi \sqrt{\pi \rho}},$$

which is again asymptotically small.

Our final case is $\varphi > \pi/2$, and this time end point contributions will not be small. The reason is that if one examines the integral representing the contribution from the middle of the range of integration, i.e.,

$$\frac{1}{2} \int_{-(\pi/2)}^{(\pi/2)-\varphi} dy [\operatorname{erfc} \sqrt{\rho} \sin |y| - \operatorname{erfc} \sqrt{\rho}] \sqrt{\frac{\rho}{\pi}} \cos(y + \varphi) \times \exp[-\rho \sin^2(y + \varphi)], \tag{50}$$

it is exponentially dominated by contributions near the end points. But in (50) our approximation to $P(y)$ vanishes faster than the correct $P(y)$ at $y = -\pi/2$, and our approximation to $p(y + \varphi)$ vanishes at $y = (\pi/2 - \varphi)$ while the true $p(y + \varphi)$ does not. This implies that the asymptotic evaluation of (50) will be asymptotically smaller than the correct contributions from the ends of the interval. The contribution from the right end is

$$\frac{1}{2} \int_{-(\pi/2)}^{(\pi/2)-\varphi} dy [\operatorname{erf} \sqrt{\rho} - \operatorname{erf} \sqrt{\rho} \sin |y|] \frac{\exp(-\rho)}{\pi} \sim \frac{\exp[-\rho(1 + \cos^2 \varphi)]}{4\pi \sqrt{\pi} \rho \sqrt{\rho} \sin \varphi \cos^2 \varphi}. \tag{51}$$

The lower limit of integration (51) is immaterial, as will be the upper limit in (52). For the contribution from the left end we have

$$\frac{e^{-\rho}}{\pi} \int_{-(\pi/2)}^{(\pi/2)-\varphi} dy \left[\frac{\pi}{2} + y\right] \sqrt{\frac{\rho}{\pi}} \cos(y + \varphi) \exp[-\rho \sin^2(y + \varphi)] \sim \frac{\exp -\rho(1 + \cos^2 \varphi)]}{4\pi \sqrt{\pi} \rho \sqrt{\rho} \sin \varphi \cos^2 \varphi}. \tag{52}$$

The sum of (52) and (51) yields (17c) of the text.

REFERENCES

1. Carson, J. R. and Fry, T. C., Variable Frequency Electric Circuit Theory with Application to the Theory of Frequency-Modulation, B.S.T.J., *16*, October, 1937, pp. 513–540.
2. Crosby, M. G., Frequency-Modulation Noise Characteristics. Proc. IRE, *25*, 1937, pp. 472–514.
3. Blachman, N. M., The Demodulation of a Frequency-Modulated Carrier and Random Noise by a Limiter and Discriminator. J. Appl. Phys., *20*, 1949, p. 38, 976.
4. Lawson, J. L. and Uhlenbeck, G. E., *Threshold Signals*, MIT Radiation Laboratory Series, *24*, McGraw-Hill Book Co., New York, 1950, Chap. 13.
5. Middleton, David, *Statistical Communication Theory*, McGraw-Hill Book Co., New York, 1960, Chap. 15.
6. Rice, S. O., Noise in FM Receivers, Chapter 25 of *Time Series Analysis*, M. Rosenblatt, (ed.), John Wiley & Sons, Inc., New York, 1963.
7. Cohn, John, Proc. N. E. C., (Chicago) *12*, 1956, pp. 221–236. We would like to thank S. O. Rice for pointing out this reference to us.
8. Bennett, W. R. and Salz, J., Binary Data Transmission by FM Over a Real Channel, B.S.T.J., *42*, 1963, pp. 2387–2426.
9. Salz, J. and Stein, S., Distribution of Instantaneous Frequency for Signal Plus Noise, IEEE Trans., *IT-10*, 1964, pp. 272–274.
10. Mazo, J. E. and Salz, J., Probability of Error for Quadratic Detectors, B.S.T.J., *44*, November, 1965, pp. 2165–2186.
11. Salz, J., Performance of Multilevel Narrow-Band FM Digital Communication Systems, IEEE Trans., *COM-13*, 1965, pp. 420–424.
12. Wozencraft, J. M. and Jacobs, I. M., *Principles of Communication Engineering*, John Wiley & Sons, Inc., New York, 1965.
13. Bennett, W. R. and Rice, S. O., Spectral Density and Autocorrelation Functions Associated with Binary Frequency Shift Keying, B.S.T.J., *42*, September, 1963, pp. 2355–2385.
14. Anderson, R. R. and Salz, J., Spectra of Digital FM, B.S.T.J., *44*, July–August, 1965, pp. 1165–1189.
15. Bennett, W. R. and Davey, J. R., *Data Transmission*, McGraw-Hill Book Co., New York, 1965.
16. Blachman, Nelson M., FM Reception and the Zeros of Narrow-Band Gaussian Noise, IEEE Trans., *IT-10*, 1964, pp. 235–241.
17. Bennett, W. R., Methods of Solving Noise Problems, Proc. IRE, *44*, 1956, pp. 609–638. See (253).
18. National Bureau of Standards, Handbook of Mathematical Functions, U. S. Government Printing Office, Washington, D.C., 1964.
19. Salz, J. and Koll, V. G., An Experimental Digital Multilevel FM Modem, IEEE Trans., *COM-14*, 1966, pp. 259–265.
20. Chaffee, J. G., The Application of Negative Feedback to Frequency Modulation Systems, B.S.T.J., *18*, July, 1939, pp. 403–437.
21. Enloe, L. H., Decreasing the Threshold in FM by Frequency Feedback, Proc. IRE, *50*, 1962, pp. 18–30.
22. Copson, E. T., *Asymptotic Expansions*, Cambridge University Press, 1965.