Rate Optimization for Digital Frequency Modulation

By J. E. MAZO, HARRISON E. ROWE, and J. SALZ

(Manuscript received June 12, 1969)

The data rate of a multilevel digital FM system is optimized subject to fixed RF bandwidth, signal-to-noise ratio, and output error rate. The possibility of optimizing such a system was first considered by J. R. Pierce at Bell Telephone Laboratories. He made the observation that it is possible to send many levels slowly or fewer levels rapidly for an FM wave of fixed RF bandwidth and error rate, and that there must be a choice of signaling rate and number of levels that optimize the data rate. The rigorous treatment of this problem is the subject of this paper. The mathematical model we analyze uses frequency-shift keying at the transmitter and ideal discrimination detection with an integrate-and-dump circuit as the post-detection filter. Our results are exhibited graphically showing the various dependencies among the pertinent system parameters.

I. INTRODUCTION

In this paper we optimize the information rate (subject to certain constraints) of a multilevel digital FM system. This problem of delivering the maximum information through an FM system has recently been formulated by J. R. Pierce.[1] Specifically, he considered how one should choose the baseband signaling rate and the number of levels to get the most information through the channel, subject to fixed bandwidth, fixed RF signal-to-noise ratio, and fixed output error rate. This optimization has recently been carried out under the assumption that the conventional FM receiver can be linearized.[2] Small-noise linear FM theory is satisfactory when analyzing analog systems, but has its well known pitfalls in digital applications.

The purpose of this paper is to reexamine this problem more rigorously, paying particular attention to the anomalies (clicks) which can result from the nonlinear character of the receiver. In order to do this we must choose a particular mathematical model for digital

FM which is amenable to analysis. Such a model uses frequency-shift keying (FSK) at the transmitter and ideal discrimination detection with an integrate-and-dump circuit as the postdetection filter. The noise at RF is assumed to possess gaussian statistics. Although realizable FM systems do not exactly conform to this ideal mathematical model, we feel that the results predicted with the use of this model are applicable to real FM systems. In any case, the numerical results agree well with those derived from the linear theory. According to our present calculations, this is due to the circumstance that the optimum number of levels leads to small enough deviations so that the contribution of the clicks to the error rate can be neglected.

II. ANALYSIS

Consider an n-level FSK communication system with a sample rate $N = 1/T$, square-wave modulation, and a level separation (in frequency) Δf. Such a system would yield a data rate R given by

$$R = N \log_2 n = 1.443 \, N \ln n \text{ bits/s}, \tag{1}$$

and, according to Carson's rule, occupy a bandwidth*

$$B = N + (n - 1)\Delta f. \tag{2}$$

The FM signal plus gaussian noise enters a receiver consisting of an ideal RF filter (bandwidth B), limiter, discriminator, integrator (integration time T), and sampler (sampling rate N). The sampler outputs are simply the successive values of the instantaneous phase of the modulated wave following each (rectangular) modulation pulse, and would be separated by multiples of

$$\Delta\phi = 2\pi \frac{\Delta f}{N} \text{ radians} \tag{3}$$

in the absence of noise.

The simplicity of the present system (that is, the finite-time integrator post-detection filter) has permitted a fairly rigorous determination of the probability of error for high RF signal-to-noise ratio.[4] It is shown in Ref. 4 that the parameter $\Delta\phi$ given in equation (3) plays a very important role in the theory of error rates for digital FM. In particular, it is known that if $\Delta\phi < \pi$ (or equivalently, $\Delta f/N < \frac{1}{2}$), then it is the smooth noise at the baseband output which determines the error

* Comparison with the exact FSK spectra for $n = 2, 4, 8$ suggests that this approximation is valid for present purposes.[3]

rate; while if $\Delta\phi > \pi$ ($\Delta f/N > \frac{1}{2}$), then the clicks dominate, which is the basic reason for the probability of error taking on different forms in these two cases.

The optimum systems considered here are shown to correspond to the $\Delta f/N < \frac{1}{2}$ case, for which clicks are unimportant. Therefore we take the probability of error* P as given by twice equation (17a) of Ref. 4, with $\phi \rightarrow \Delta\phi/2 = \pi\, \Delta f/N$;

$$P \sim \frac{1}{(2\pi\rho)^{\frac{1}{2}}} \frac{\cot\left(\frac{\pi}{2}\frac{\Delta f}{N}\right)}{\left(\cos\left(\pi\frac{\Delta f}{N}\right)\right)^{\frac{1}{2}}} \exp\left[-2\rho\sin^2\left(\frac{\pi}{2}\frac{\Delta f}{N}\right)\right], \qquad \rho \gg 1, \quad \frac{\Delta f}{N} < \frac{1}{2}, \tag{4}$$

and subsequently verify that $\Delta f/N$ is indeed less than $\frac{1}{2}$ for the resulting optimum systems. Here ρ is the RF signal-to-noise ratio in the frequency band B. We treat the asymptotic approximation (for large ρ) of equation (4) as an equality in the following.

For fixed error rate P and RF signal-to-noise ratio ρ, equation (4) determines $\Delta f/N$. Rewriting equation (2),

$$\frac{B}{N} = 1 + (n - 1)\frac{\Delta f}{N}; \tag{5}$$

substituting equation (5) into equation (1),

$$\frac{R}{B} = \frac{1.443 \ln n}{1 + (n-1)\frac{\Delta f}{N}} \text{ bits/cycle.} \tag{6}$$

We set the derivative of equation (6) equal to zero, determining the optimum number of levels n_0 and maximum rate R_0 .

$$n_0(\ln n_0 - 1) = \frac{1}{\Delta f/N} - 1. \tag{7}$$

$$\frac{R_0}{B} = \frac{1.443}{n_0(\Delta f/N)}. \tag{8}$$

Alternatively, once the optimum number of levels n_0 has been de-

---

* For multilevel output samples, most errors will be to adjacent levels. Assuming that something like the Gray code is used, the symbol probability of error P of equation (4) will be approximately the bit probability of error for the final reconstructed binary signal.

termined via equations (4) and (7), we may express the other parameters of the (optimum) system in terms of n_0 only:

$$\frac{\Delta f}{N} = \frac{1}{n_0(\ln n_0 - 1) + 1}, \tag{9}$$

$$\frac{R_0}{B} = 1.443 \left[\ln n_0 + \frac{1}{n_0} - 1 \right] \text{bits/cycle}, \tag{10}$$

$$\frac{B}{N} = \frac{n_0 \ln n_0}{n_0(\ln n_0 - 1) + 1}. \tag{11}$$

Note that the restriction $\Delta f/N < \frac{1}{2}$ implies via equation (7) that

$$n_0 \geqq 4. \tag{12}$$

Finally, the Shannon capacity for the RF channel is

$$\frac{C}{B} = 1.443 \ln (1 + \rho) \text{ bits/cycle}. \tag{13}$$

III. RESULTS

Figures 1 to 7 illustrate the parameters of optimum multilevel FM systems using a finite-time integrator as a post-detection filter for two representative error rates ($P = 10^{-6}, 10^{-8}$).

The solid curves of Fig. 1 show the optimum number of levels n_0 versus the RF signal-to-noise ratio in dB, $10 \log_{10} \rho$, for the two values of P. The curves terminate at $n_0 = 4$, according to equation (12).

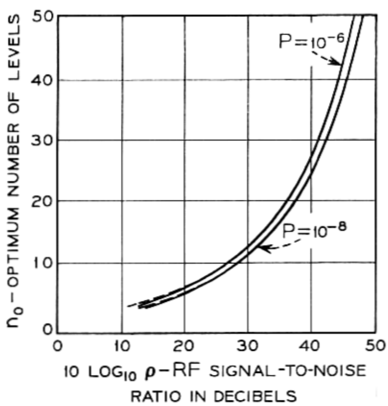

Fig. 1 — Number of levels for maximum data rate versus RF signal-to-noise ratio. Dashed lines indicate small-angle approximations.

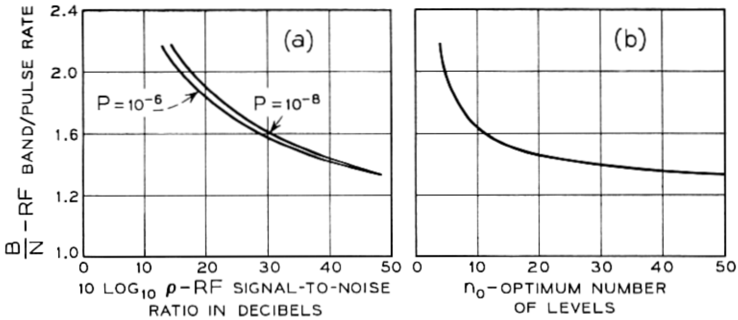

Fig. 2 — Bandwidth expansion factor for maximum data rate.

n_0 increases rapidly as ρ increases, for fixed P. The small-angle approximation for the trigonometric functions in equation (4) is shown by the dashed curves of Fig. 1; in this approximation changing P simply translates the curves of Fig. 1 horizontally. This is a reasonable approximation for the smallest n_0 permitted [by equation (12)], for the values of P of interest here.

Figures 2, 3, 4, and 5 show optimum system parameters plotted against two horizontal scales:

(*i*) 10 $\log_{10}\rho$—the RF signal-to-noise ratio in dB. Two plots are shown, for $P = 10^{-6}, 10^{-8}$. Using the small-angle approximation in equation (4), changing P translates these curves horizontally. This horizontal axis is the parameter of most direct physical interest.

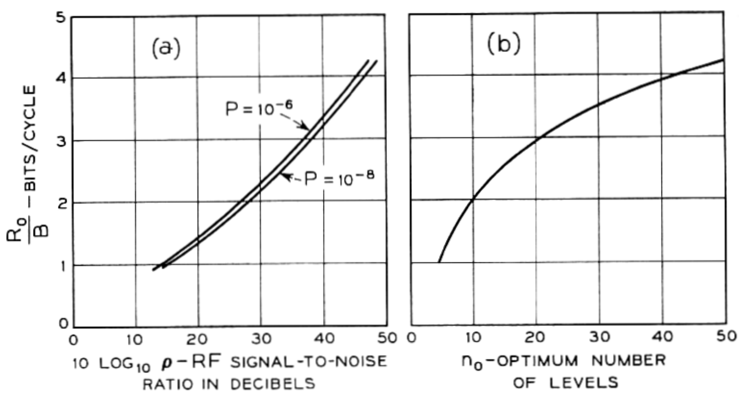

Fig. 3 — Maximum data rate per unit RF bandwidth.

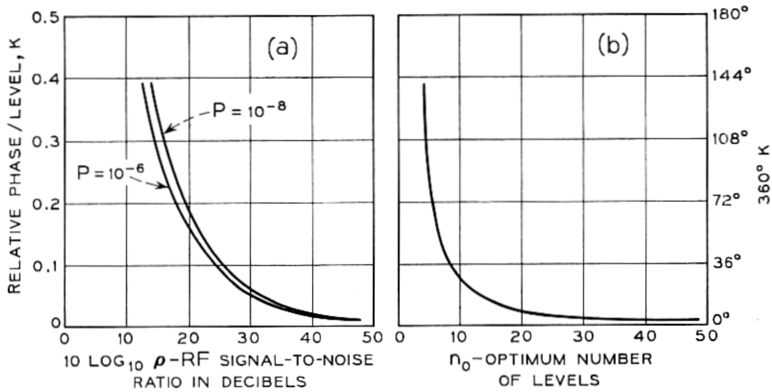

Fig. 4 — Relative phase shift per level in one sample interval for optimum systems.

(*ii*) n_0—the optimum number of levels, determined from Fig. 1. Here a single universal plot suffices rigorously for all P [That is, without small-angle approximations in equation (4)].
The vertical axes show:

Figure 2—B/N, the bandwidth expansion factor, roughly* one-half the ratio of RF to base-bandwidth. This factor varies from about 2 at small ρ or n_0 , to an asymptotic limit of 1 as ρ, $n_0 \rightarrow \infty$. For large ρ, n_0 we have small-index phase modulation, with only the first sideband significant. Even for the smallest ρ, n_0 considered here the bandwidth expansion is moderate.

Figure 3—R_0/B, the normalized maximum rate in bits per cycle. This quantity increases monotonically with ρ, n_0 .

Figure 4—$360 \cdot \Delta f/N$ represents the relative phase change in degrees corresponding to a change in modulation of one level.

Figure 5—$360\ (n - 1)\ \Delta f/N$ represents the maximum relative phase change in degrees in one sampling interval, corresponding to a change in modulation from the lowest to the highest level. The maximum value for this quantity, occurring for the smallest ρ, n_0 (that is, $n_0 = 4$) is not far from 360°. As ρ, n_0 increase, the maximum phase change becomes small for optimum systems.

Within the small-angle approximation, discussed in connection with Fig. 1, changing P merely shifts the horizontal (dB) axes of Fig. 1 and Figs. 2(a) to 5(a). Let us adopt the $P = 10^{-6}$ curves as standard,

---

* This is because the square-wave modulation assumed here is not strictly band-limited; in fact, its spectrum falls off so slowly that its rms bandwidth is infinite.

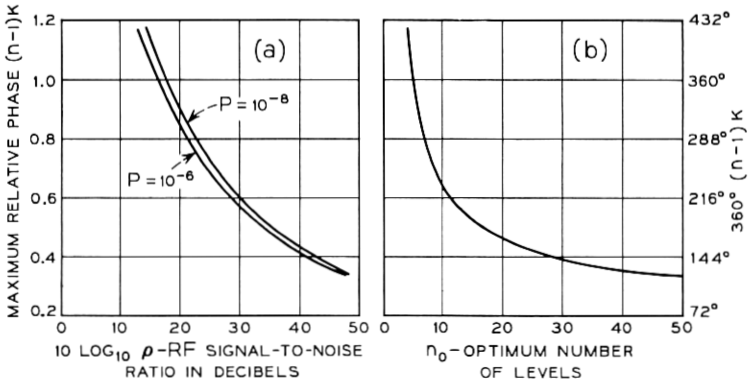

Fig. 5 — Maximum relative phase shift in one sample interval for optimum systems.

and plot the number of dB to be added to the 10 $\log_{10} \rho$ axes a sa function of P. This is shown in Fig. 6. We remark that this is only an approximation, and will begin to fail sooner as P decreases.

Finally, Fig. 7 compares the maximum data rate for the multilevel FM system with the Shannon capacity of the RF channel. The optimum data rate ranges from about 19 to 27 percent of the ideal RF channel capacity, for error probabilities P between 10^{-6} and 10^{-8}.

We have so far dealt with optimum systems. However, the number of levels may be fixed by other constraints, so that suboptimum systems are of interest. For example, it may not be practical to have the large number of levels required for optimum systems at large RF signal-to-

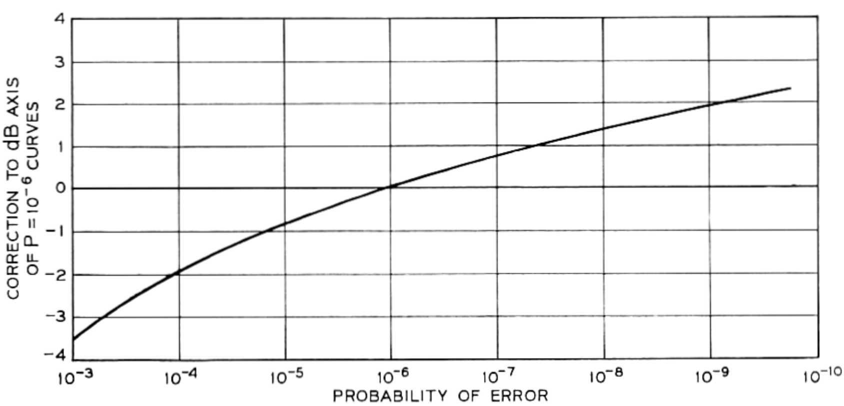

Fig. 6 — Correction for modifying $P = 10^{-6}$ curves to other error probabilities.

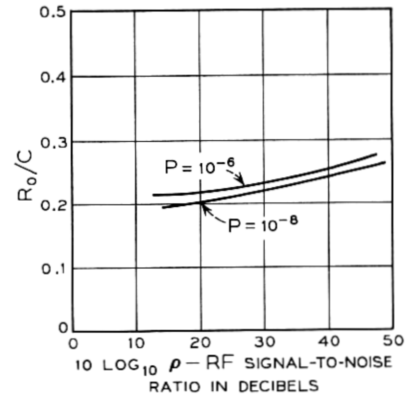

Fig. 7 — Ratio of maximum data rate to Shannon capacity.

noise ratios ρ; we may be restricted to 8 (or 16) levels, and it is necessary to determine how much the data rate will be reduced. Now rather than maximizing R by varying N and n in equation (1) subject to the constraints of equations (2) and (4), we fix n in equations (5) and (6). Figures 8 and 9 show the optimum rate R_0/B versus 10 $\log_{10}$ ρ [given

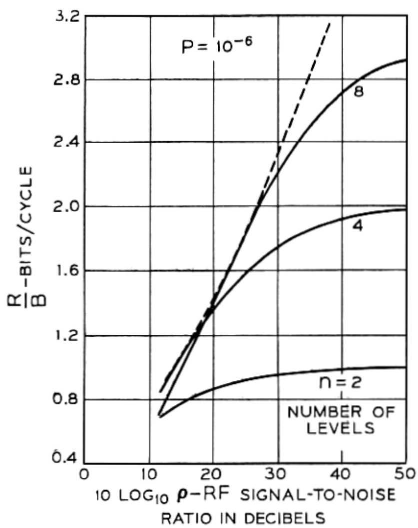

Fig. 8 — Best data rate for suboptimum systems with two, four, and eight levels compared to maximum data rate for optimum system. Dashed line—maximum data rate for optimum system, R_0/B (see Fig. 3).

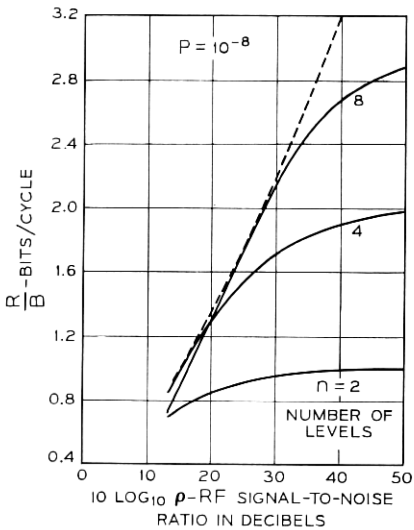

Fig. 9 — Best data rate for suboptimum systems with two, four, and eight levels compared to maximum data rate for optimum system. Dashed lined—maximum data rate for optimum system, R_0/B (see Fig. 3).

also in Fig. 3(a)], together with the rates for two, four, and eight levels, determined from equation (6) with $n = 2$, 4, and 8 for $P = 10^{-6}$, 10^{-8} in equation (4). While eight levels is strictly optimum only at the point of tangency between the R_8 and the R_0 curves, we see that the optimum is fairly broad. The corresponding bandwidth expansion factors are found from equation (5).

IV. DISCUSSION

We have presented the results of Figs. 1 through 9 as continuous curves. Actually, only isolated points of these curves are significant, since the number of levels must be integral. These continuous curves should consequently be replaced by appropriate "staircase" functions, but the difference will be significant only for small numbers of levels (that is, at low RF signal-to-noise ratios).

The present theory excludes two- and three-level systems. Naively, one might try to extend the present results to these cases by equation (17c) and Fig. 5 of Ref. 4. This may not be accurate for the error rates considered here ($P = 10^{-6}$, 10^{-8}), because the RF signal-to-noise ratio ρ becomes small, and the basic results of Ref. 4, that is, equations (17), (26), and (27), are asymptotic as ρ becomes large. However, for very much smaller error rates, for example, $P \approx 10^{-30}$, it is possible that this approach would be productive.

It would be desirable to extend the present results to binary and ternary systems; this will require a different or improved approach from the asymptotic evaluation of Ref. 4 for the error probability. It seems likely that clicks will dominate the error behavior for optimum two- and three-level systems.

The principal limitation in the present treatment (aside from the assumptions of the model, such as a finite-time integrator post-detection filter) lies in our lack of knowledge of the precise way in which the basic result for the probability of error P (equation (4) above) fails. We have merely assumed that this result holds for signal-to-noise ratios down to about 10 dB, independently of P or $\Delta f/N$. This provides additional motivation for further study of the asymptotic theory of Ref. 4.

REFERENCES

1. Pierce, J. R., unpublished work.
2. Rowe, H. E., unpublished work.
3. Lucky, R. W., Salz, J., and Weldon, E. J., *Principles of Data Communication*, New York: McGraw-Hill, 1968.
4. Mazo, J. E. and Salz, J., "Theory of Error Rates for Digital FM," B.S.T.J., *45*, No. 9 (November 1966), pp. 1511–1535.