Spectrum Analysis of Digital Magnetic Recording Waveforms

ARNOLD L. KNOLL, MEMBER, IEEE

Abstract—Closed form expressions have been obtained for the power density spectra of signal waveshapes in use for digital magnetic recording. The "signal" is the magnetization versus distance profile imposed on the magnetic medium to encode binary ones and zeros. Two general classes of signals are distinguished in terms of neighboring bit-to-bit correlation. Standard recording methods such as saturation NRZ and phase modulation are considered, as well as techniques requiring bias current, such as sinusoidal frequency-shift modulation. The spectra are compared with frequency response curves for the reading head at several recorded bit densities. The more ideal bandwidth characteristics of the newer phase- and frequency-modulation techniques help to explain their improved performance at high bit densities.

Index Terms—Digital magnetic recording, frequency modulation, frequency-shift modulation, magnetic recording, NRZ recording, phase modulation, phase-shift modulation, power density spectrum, spectrum analysis.

I. Introduction

MAGNETIC RECORDING systems for computer or instrumentation applications employ many different waveshapes for encoding digital information in a form suitable for writing on magnetic tape and subsequent reading and recovery. A plot of tape magnetization versus distance along the tape for a single channel may reveal a waveshape approximating a square wave, in which the magnetization switches from positive to negative saturation whenever the binary information changes from a one to a zero (NRZ recording),[1] or a sinusoidal waveshape with several cycles of one wavelength representing a binary one and the same duration at another wavelength representing a binary zero (frequency-shift modulation).[2] The choice of a particular coding scheme in a given application depends upon many factors, among them the complexity of coding and decoding equipment, the allowable error rate, and the information rate or density of information required per unit tape length. In general, increased density of information may be achieved at the expense of an increased error rate, but the relationship is not linear for any given scheme, and each scheme has its optimum range of operation.

As in communications systems, a basis for comparison of various recording schemes can be found by examining how efficiently each makes use of the available bandwidth. Many of the newer techniques introduced to increase the information density in magnetic recording (e.g., phase modulation) are successful because they offer a better mating of the bandwidth of the recorded signal with the bandwidth limitations imposed by the writing and reading (mainly the reading) process than is achieved by the older techniques. An examination of the frequency spectrum of various recording schemes will therefore provide important information for purposes of comparison. This paper derives expressions for the frequency spectrum (more exactly, the power density spectrum) of standard recording schemes, as well as the newer high-density recording schemes. A summary of the recording schemes considered is presented in Table I, while Fig. 1 shows the waveshape for these schemes in the form of tape magnetization versus distance on the recorded tape. For the schemes in which the tape is saturated in either direction, the WRITE current versus time waveshape is the same as shown for the magnetization versus distance. In fact, the instantaneous transition from positive to negative values approximates the behavior of the current better than the magnetization (see Section III). In the sinusoidal schemes, linearity in the magnetization is achieved by superimposing a high-frequency bias current on the information waveshape, as in standard audio recording. Although the discussion is phrased in the context of tape recording, the results apply as well to other recording media, such as disks and drums.

Among the recording methods considered, schemes 1–4 are among the oldest and are discussed by Athey.[1] Scheme 5, also known as Manchester recording, is becoming fairly standard for high-density recording, especially on disk files. Scheme 6, introduced by Gabor,[3] may be regarded as a variation of scheme 5 (as in reference [1], page 54, where it is called diphase, and has ones and zeros interchanged with the representation in Fig. 1). It will be shown in Section III that the spectra for schemes 5 and 6 are the same over a long coding interval, although individual bits are coded differently. However, scheme 6 is also a special case of scheme 7; and indeed the spectrum for scheme 6 was obtained by simplifying the spectrum for scheme 7 in this special case. The general approach of scheme 7 was investigated by Chao.[4] The sinusoidal phase-modulation technique has been studied by Hopner,[5] while the frequency-modulation technique using sinusoidal signals was considered by Gillis[6] as well as Young[2] for application to magnetic recording.

A derivation of the spectral density for the sinusoidal frequency-shift modulation technique has been carried

Manuscript received March 7, 1967; revised July 17, 1967.
The author is with Honeywell Inc., Electronic Data Processing Division, Waltham, Mass. 02154.

TABLE I

SUMMARY OF SIGNAL WAVESHAPES FOR DIGITAL MAGNETIC RECORDING (SEE ILLUSTRATIONS IN FIG. 1)

1) *NRZ (Nonreturn to Zero) Saturation Recording:*
The tape is fully saturated in one direction or the other for one or more bit periods. The positive direction (for example) represents a binary ZERO while the negative direction represents a binary ONE.

2) *NRZ1 Saturation Recording:*
This is a variation of the NRZ method in which the direction of magnetization is switched at the beginning of a bit period when a binary ONE is being recorded but the magnetization retains its direction of saturation when a ZERO is recorded. This is probably the most common technique now in use in the computer field.

3) *RB (Return to Bias) Saturation Recording:*
The tape is normally saturated in the negative direction by a constant bias current. When a binary ONE appears in the input data, the magnetization is switched to the positive direction for a small portion of the bit period, and then returns to the negative direction.

4) *RZ (Return to Zero) Recording:*
This is a variation of the RB method in which there is no bias current, i.e., the tape is unmagnetized except for a small portion of the bit period associated with a binary ONE when the tape is saturated in the positive direction.

5) *Phase Modulation Saturation Recording:*
The basic coding signal consists of one cycle of a square wave over the bit period, i.e., the tape magnetization is saturated in one direction for the first half of the bit period and in the other direction for the second half. For a ZERO bit, the first half is positive and the second half is negative, while the phase is reversed for a ONE bit, i.e., the first half is negative and the second half positive. Note that for alternating bits, no switching occurs at the beginning of a bit period, but switching always occurs in the middle of the period.

6) *Frequency Doubling Saturation Recording:*
The same square wave as in method 5 is used to represent a ZERO bit. A ONE bit is represented by one-half cycle of a square wave at half the frequency, i.e., by positive or negative saturation throughout the bit interval. The sign of either signal is chosen so that switching always occurs at the beginning of a bit period. For a string of all ONES, no switching occurs in the middle of the interval.

7) *High-Frequency Modulation Saturation Recording:*
This is a variation of method 6 in which the frequency of the square wave representing a ZERO bit is k times the frequency of the square wave representing a ONE, where k is an integer greater than 2. Each bit period corresponding to a ZERO will contain $k/2$ cycles of the square wave.

8) *Sinusoidal Phase Modulation Linear Recording:*
Magnetization corresponds to one (or more) cycles of a sine wave per bit period for a ZERO bit and the same *sine* wave of opposite *sign* for a ZERO bit. The phase at the beginning of the bit interval is usually chosen to be zero.

9) *Sinusoidal Frequency Modulation Linear Recording:*
A ZERO bit is coded as a sine wave at one frequency, having r_a cycles per bit period, and a ONE bit as a sine wave at a second frequency, having r_b cycles per bit period. If r_a and r_b are integers, the frequencies are commensurate and phase relationships are preserved from period to period.

out previously by Bennett and Rice[7] for application to communications systems. They have considered general sinusoidal signals in which the frequencies are not necessarily commensurate with the bit period. For magnetic recording purposes, only commensurate frequencies have thus far been considered[2],[6] and these are the cases we have included (case 9, Tables I and III). The

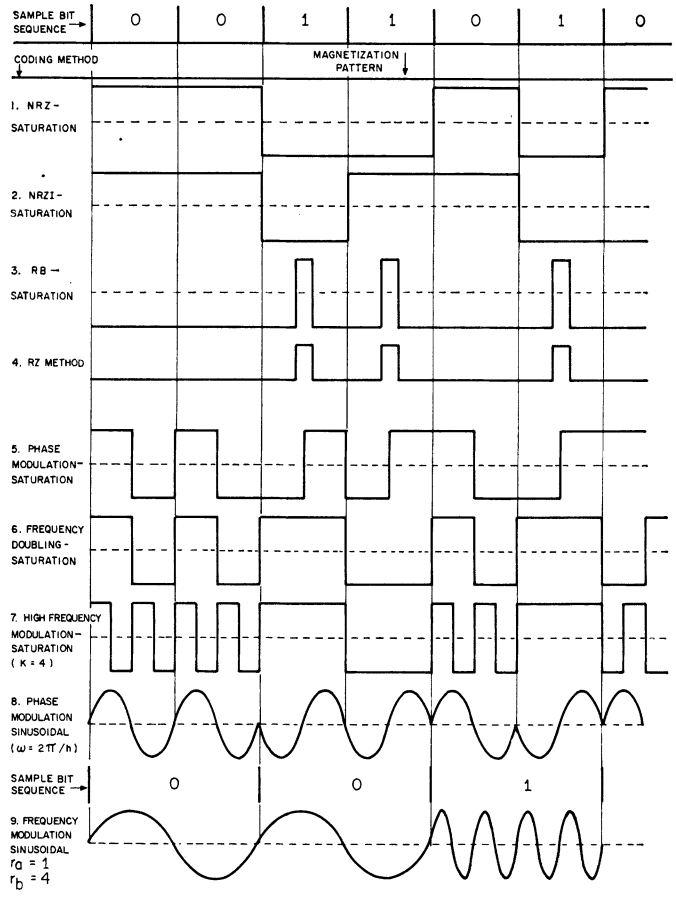

Fig. 1. Magnetization versus distance for various recording schemes (see Table I).

results of Bennett and Rice were discovered after the derivation given in this paper was completed, and were found to agree with the results of case 9 for commensurate frequencies. The method of derivation presented by Bennett and Rice[7] is not easily extended to the class 2 type of signals considered in Section II, in which the waveshapes representing successive bits are not statistically independent.

II. General Assumptions and Basic Equations

It is assumed that the recording and reproducing equipment are operated at a constant speed so that the discussion takes place in terms of frequency and time, rather than wavelength and distance. Since the information being recorded is random by nature, any measure of frequency distribution must be a statistical one. As explained by Bennett and Rice,[7] "the spectral density function, or power spectrum, of a random sequence of signals defines the distribution of average signal power vs. frequency. This information is useful in system design because it indicates the frequency band of most importance. . . . It does not tell us about important spectral components which may be associated with unlikely but possible specific signal sequences." With this limitation in mind, it is most reasonable to assume that the information to be recorded consists of a string of N bits, arriving at a rate of $1/h$ bits per second, with N going

to infinity. The nth bit ($n=0, 1, 2, \cdots, N-1$) is encoded as a signal lasting h seconds and denoted as $f_n(t)$ ($nh<t\leq(n+1)h$). h is called the bit period. $f_n(t)$ has the form $f_a(t)$ if the bit is a zero, and $f_b(t)$ if the bit is a one. When the ensemble of all possible information strings is considered, the nth bit is a random variable with equal probability of being a zero or a one, independently of the values of the other bits. The composite recorded signal $f(t)$, $0<t\leq Nh$, represents a sample function from a stochastic process whose spectral density $P(\omega)$ is given by[8]

$$P(\omega) = \lim_{N\to\infty} \frac{2}{Nh} E(|S(\omega)|^2) \tag{1}$$

$$S(\omega) = \int_0^{Nh} e^{-j\omega t} f(t)\,dt = \sum_{n=0}^{N-1} \int_{nh}^{(n+1)h} f_n(t) e^{-j\omega t}\,dt. \tag{2}$$

The symbol E represents the expectation or ensemble average of the function $S(\omega)\cdot S(\omega)^* = |S(\omega)|^2$, where the asterisk denotes complex conjugate, and ω is the frequency in radians.

For computational purposes, it is convenient to separate the specific cases under consideration into two classes, depending upon the behavior of the signals $f_a(t)$ and $f_b(t)$ as a function of the bit number n.

Class 1

The time behavior of the function $f_a(t)$ will be the same for each bit period in which it occurs, independent of n, and similarly for $f_b(t)$. Mathematically, this may be expressed by writing

$$\begin{aligned} f_a(t) &= f_a(t-nh) \\ f_b(t) &= f_b(t-nh) \end{aligned} \quad nh<t\leq(n+1)h, \quad \text{all } n. \tag{3}$$

Recording schemes 1, 3, 4, 5, 8, and 9 of Fig. 1 fall into this class.

Letting $t'=t-nh$, (2) may then be written

$$S_1(\omega) = \sum_{n=0}^{N-1} e^{-j\omega nh} \int_0^h f_n(t') e^{-j\omega t'}\,dt'. \tag{4}$$

Class 2

The time behavior of $f_a(t)$ will take one of two related forms, called $f_{a+}(t)$ and $f_{a-}(t)$ respectively, where $f_{a+}(t)=-f_{a+}(t)$. Similarly, $f_b(t)$ has two forms for which $f_{b-}(t)=-f_{b+}(t)$. Each of the forms is independent of the bit number, so that a relation such as (3) holds for each form; e.g.,

$$f_{a+}(t) = f_{a+}(t-nh) \quad nh<t\leq(n+1)h, \text{ etc.}$$

In other words, the signal representing a binary zero occurring during any bit period may have one of two waveshapes which differ by a sign reversal; and similarly for a binary one. The choice between plus and minus form during any bit period is determined by the waveshape form assumed during the previous bit period, with either choice being equally likely at the first bit ($n=0$). Thus, while successive bits are statistically independent, their coded representations are not. For signals in class 2, (2) becomes

$$S_2(\omega) = \sum_{n=0}^{N-1} e^{-j\omega nh} \int_0^h d_{n-1} f_{n+}(t') e^{-j\omega t'}\,dt' \tag{5}$$

where d_{n-1} takes on the value $+1$ or -1, depending upon the form of f_{n-1}. The four cases that may occur are summarized in Table II. They will be seen to apply to recording schemes 2, 6, and 7 in Fig. 1, with suitable definitions of f_{a+} and f_{b+}. One straightforward result that follows from Table II is that two "zeros" in succession always have the same waveshapes, while two "ones" in succession always have a sign change. Schemes 2, 6, and 7 in Fig. 1 do, indeed, exhibit this behavior.

General expressions for the spectral density of the two classes are obtained directly by evaluating $E(|S(\omega)|^2)$ for (4) and (5) and applying the summation calculus[9] to the resulting double summations. The details are summarized in the Appendix. One finally obtains the following:

for class 1,

$$P_1(\omega) = \lim_{N\to\infty} \frac{1}{2Nh} \left[\left| \frac{\sin^2\left(\frac{\omega hN}{2}\right)}{\sin^2\left(\frac{\omega h}{2}\right)} \right| \cdot (|I_a|^2 + |I_b|^2 + I_aI_b^* + I_bI_a^*) + N(|I_a|^2 + |I_b|^2 - I_aI_b^* - I_bI_a^*) \right], \tag{6}$$

for class 2,

$$P_2(\omega) = \lim_{N\to\infty} \frac{1}{Nh} \left[N(|I_{a+}|^2 + |I_{b+}|^2) + (N-1)\left\{ \{|I_{a+}|^2 - |I_{b+}|^2\} \cos(\omega h) + j\sin(\omega h)\{I_{a+}I_{b+}^* - I_{b+}I_{a+}^*\} \right\} \right], \tag{7}$$

TABLE II
NEIGHBORING BIT BEHAVIOR FOR CLASS 2 SIGNALS

If $f_{n-1}=$	then $d_{n-1}=$	and $f_n=$
1. $+f_{a+}$ (ZERO bit, $d_{n-2}=+1$)	$+1$	$+f_{n+}$ (either f_{a+} or f_{b+})
2. $f_{a-}=-f_{a+}$ (ZERO bit, $d_{n-2}=-1$)	-1	$-f_{n+}$ (either $-f_{a+}$ or $-f_{b+}$)
3. $+f_{b+}$ (ONE bit, $d_{n-2}=+1$)	-1	$-f_{n+}$ (either $-f_{a+}$ or $-f_{b+}$)
4. $f_{b-}=-f_{b+}$ (ONE bit, $d_{n-2}=-1$)	$+1$	$+f_{n+}$ (either f_{a+} or f_{b+})

where

$$I_n = \int_0^h f_n(t)e^{-j\omega t}dt \tag{8}$$

and the subscript n stands for a, b, $a+$, etc.

III. Specific Cases—Results and Discussion

With the general framework and equations presented in Section II, the derivation of the power density for specific cases is straightforward and involves mainly the evaluation of the integrals I_a and I_b for each case, assuming that the waveshapes f_a and f_b have been specified. The integrations are standard but tedious, and will not be carried out here. Table III defines the waveshapes assumed for the class 1 schemes in Table I in terms of c, the amplitude of magnetization at saturation, and presents the resulting expressions for the power density as a function of frequency $P(\omega)$. Table IV provides the same information for the schemes in class 2.

It will be noted that the waveshapes specified for the saturation recording schemes are idealized in that they assume instantaneous switching from positive to negative saturation, and vice versa. In reality, there is a finite transition region between positive and negative saturation which is a function of the recording medium as well as the WRITE current and head characteristics. This finite transition region will tend to reduce the high-frequency spectral content of all the signals considered. As seen in Section IV, more severe frequency limitations are imposed by the READ head at the bit densities considered here, which are well below the theoretical upper limits imposed by the finite transition region, especially for thin metallic media.[10] In addition, the newer recording schemes will be shown to be less dependent on the high-frequency content. However, the method developed here will still be applicable to less ideal waveshapes if the form of the transition curve can be specified analytically. This may become more important when heads with narrower gaps (down to 20 microinches effective gap width) become more prevalent.

In comparing Tables III and IV, it is seen that the spectra for NRZ and NRZ1 recording are the same (schemes 1 and 2) and the spectra for saturation phase modulation and frequency doubling (schemes 5 and 6) are also identical. This results from the fact that, for a random bit sequence, the waveshape over a long coding interval has the same form within each pair, although the coding of individual bits is different over any short interval (see Fig. 1).

Some of the schemes in Table III contain discrete spectral lines at particular frequencies. At these frequencies, the spectral density $P_1(\omega)$ becomes indeterminate when the limiting procedure indicated in (6) is carried out. If $P_1(\omega)$ is then evaluated at these frequencies by applying L'Hospital's rule,[11] it goes to infinity, implying that finite power is concentrated at these discrete frequencies. The presence of these discrete components could have been found by decomposing each signal into the sum of a random and nonrandom signal, as done by Bennett and Rice.[7] The nonrandom signal in each bit period is given by

$$f_s(t) = 1/2(f_a(t) + f_b(t)),$$

while the random signal is a choice between

$$f_r(t) = 1/2(f_a(t) - f_b(t))$$

and $-f_r(t)$. If the nonrandom signal is periodic from bit period to bit period, as it is in schemes 3, 4, and 9, the discrete components will correspond to the spectrum of the periodic waveform. No discrete components will be present when the nonrandom signal vanishes (when $f_b=-f_a$). Also, the class 2 signals contain no discrete components because of the equally likely occurrence of $f_a(t)$ in the form $f_{a+}(t)$ or $f_{a-}(t)$, and similarly for $f_b(t)$, so that no nonrandom periodic signal exists.

For schemes 3, 4, and 9, the distribution of power between the discrete and continuous spectrum can be easily determined by comparing the mean square power in a bit period given by

$$1/2\int_0^h [f_s(t)]^2dt \quad \text{and} \quad 1/2\int_0^h [f_r(t)]^2dt,$$

respectively. One finds that in schemes 4 and 9, the power is distributed evenly between the discrete and continuous spectrum; while in scheme 3, most of the power is at zero frequency due to the biasing level, and the discrete-to-continuous power ratio is given by $(h-g)/g$. In scheme 9, the discrete power is evenly divided between the two coding frequencies. In the other two schemes, the frequency distribution of the discrete components will correspond to that of a periodic nonsymmetrical rectangular waveshape.

Curves of $P(\omega)$ versus frequency are shown in Figs. 2 through 6. Related schemes are grouped together for easy comparison. The frequency $f=\omega/2\pi$ is normalized

TABLE III
WAVESHAPE AND POWER SPECTRA FOR CLASS 1 SIGNALS

Recording Scheme (See Fig. 1 and Table I)	$f_a(t)$	$f_b(t)$	$P_1(\omega)$ (See listed Fig.)
1) Saturation NRZ	$+c,\ 0 < t \leq h$	$-c,\ 0 < t \leq h$	$\frac{8c^2h}{(\omega h)^2}\sin^2\left(\frac{\omega h}{2}\right)$ (Fig. 2)
3) Return-to-Bias Saturation	$-c,\ \ 0 < t \leq h$	$-c,\ \ 0 < t \leq \frac{h-g}{2}$ $+c,\ \ \frac{h-g}{2} < t \leq \frac{h+g}{2}$ $-c,\ \ \frac{h+g}{2} < t \leq h$	$\frac{8c^2h}{(\omega h)^2}\sin^2\left(\frac{\omega g}{2}\right)$ (except for discrete spectral lines at the frequencies $\omega_k = (2\pi k/h)$, k an integer, i.e., $P_1(\omega) \to \infty$) (Fig. 2)
4) Return-to-Zero Saturation	$0,\ \ 0 < t \leq h$	$0,\ \ 0 < t \leq \frac{h-g}{2}$ $+c,\ \ \frac{h-g}{2} < t \leq \frac{h+g}{2}$ $0,\ \ \frac{h+g}{2} < t \leq h$	$\frac{2c^2h}{(\omega h)^2}\sin^2\left(\frac{\omega g}{2}\right)$ (except at frequencies $\omega_k = (2\pi k/h)$, where $P_1(\omega)$ has discrete spectral line)
5) Saturation Phase (Shift) Modulation	$+c,\ \ 0 < t \leq \frac{h}{2}$ $-c,\ \ \frac{h}{2} < t \leq h$	$-c,\ \ 0 < t \leq \frac{h}{2}$ $+c,\ \ \frac{h}{2} < t \leq h$	$\frac{8c^2h}{(\omega h)^2}\left[1-\cos\left(\frac{\omega h}{2}\right)\right]^2$ (Fig. 3)
8) Sinusoidal Phase (Shift) Modulation	$c\sin\frac{2\pi}{h}t,\ \ 0 < t \leq h$	$-c\sin\frac{2\pi}{h}t,\ \ 0 < t \leq h$	$\frac{8c^2h}{(2\pi)^2\left[1-\left(\frac{\omega h}{2\pi}\right)^2\right]^2}\sin^2\left(\frac{\omega h}{2}\right)$ (Fig. 3)
9) Sinusoidal Frequency (Shift) Modulation (See Section III) a) Signal Continuous Between Bits	$c\sin\frac{2\pi r_a t}{h},\ \ 0 < t \leq h$ (r_a is an integer)	$c\sin\frac{2\pi r_b t}{h},\ \ 0 < t \leq h$ (r_b is an integer, $r_b \neq r_a$)	$\frac{2c^2h}{(2\pi)^2}\left[\frac{r_a}{r_a^2-\left(\frac{\omega h}{2\pi}\right)^2}-\frac{r_b}{r_b^2-\left(\frac{\omega h}{2\pi}\right)^2}\right]^2\sin^2\left(\frac{\omega h}{2}\right)$ (except for discrete spectral lines at the frequencies $\omega_a \equiv \frac{2\pi r_a}{h}$ and $\omega_b \equiv \frac{2\pi r_b}{h}$ where $P_1(\omega) \to \infty$) (Figs. 5 and 6)
b) Signal Slope Continuous Between Bits	$c\cos\frac{2\pi r_a t}{h},\ \ 0 < t \leq h$	$c\cos\frac{2\pi r_b t}{h},\ \ 0 < t \leq h$	$\frac{2c^2h}{(2\pi)^2}\left(\frac{\omega h}{2\pi}\right)^2\left[\frac{1}{r_a^2-\left(\frac{\omega h}{2\pi}\right)^2}-\frac{1}{r_b^2-\left(\frac{\omega h}{2\pi}\right)^2}\right]^2\sin^2\left(\frac{\omega h}{2}\right)$ (except for discrete spectral lines at the frequencies ω_a and ω_b) (Fig. 6)

TABLE IV
WAVESHAPE AND POWER SPECTRA FOR CLASS 2 SIGNALS

Recording Scheme (See Fig. 1 and Table I)	$f_{a+}(t)$	$f_{b+}(t)$	$(P_{2\omega})$ (See listed Fig.)
2) Saturation NRZ1	$+c, \quad 0 < t \leq h$	$-c, \quad 0 < t \leq h$	$\frac{8c^2h}{(\omega h)^2} \sin^2\left(\frac{\omega h}{2}\right)$ (Fig. 2)
6) Frequency Doubling Saturation	$+c, \quad 0 < t \leq \frac{h}{2}$ $-c, \quad \frac{h}{2} < t \leq h$	$+c, \quad 0 < t \leq h$	$\frac{8c^2h}{(\omega h)^2}\left[1 - \cos\left(\frac{\omega h}{2}\right)\right]^2$ (Fig. 3)
7) High Frequency Modulation Saturation	$+c\begin{cases} 0 < t \leq \frac{h}{k} \\ \frac{2h}{k} < t \leq \frac{3h}{k} \\ \cdots\cdots\cdots\cdots \\ \frac{(k-2)h}{k} < t \leq \frac{(k-1)h}{k} \end{cases}$ $-c\begin{cases} \frac{h}{k} < t \leq \frac{2h}{k} \\ \frac{3h}{k} < t \leq \frac{4h}{k} \\ \cdots\cdots\cdots\cdots \\ \frac{(k-1)h}{k} < t \leq h \end{cases}$ k is an even integer	$+c, \quad 0 < t \leq h$	$\frac{8c^2h}{(\omega h)^2}\left[\tan\left(\frac{\omega h}{2k}\right)\cos\left(\frac{\omega h}{2}\right) - \sin\left(\frac{\omega h}{2}\right)\right]^2 \sin^2\left(\frac{\omega h}{2}\right)$ (Fig. 4)

in terms of the bit frequency $b = 1/h (f/b = \omega h/2\pi)$. The power density is also normalized by dividing out the factor $2c^2h/\pi^2$, and is plotted on a logarithmic scale. For the sinusoidal schemes (8 and 9), $P(\omega)$ is multiplied by a factor of 2 before plotting, since a sine wave contains half the power of a square wave having the same maximum amplitude.

The following characteristics are noted from the spectrum curves:

1) The curves for NRZ (NRZ1) and RB recording in Fig. 2 have the largest power concentration about zero frequency. All the other techniques have a null at zero frequency. Also, the major lobe for the RB case is much wider than the major lobe for NRZ (five times as wide in the case shown, where the ON period in RB is one-fifth of the bit period for a binary ONE). For the RB case, the discrete spectral components present at the frequencies $\omega_k = 2\pi k/h$, k an integer, are not shown.

2) In the phase-modulation techniques, the power is concentrated near the coding frequency $\omega = 2\pi/h$ (Fig. 3). Of course, the saturation technique produces much more significant sidelobes than the sinusoidal technique. If the magnetization transition region were taken into account in the saturation technique, the sidelobe levels would be reduced somewhat.

3) In the high-frequency modulation technique of Fig. 4 (with $k = 4$), two main lobes are seen centered at the two fundamental coding frequencies of $\omega_a = 4\pi/h$ and $\omega_b = \pi/h$, respectively. Although the sidelobe behavior is quite complex, the spectrum shows good separation between the main lobe contributed by the ONE bits and that due to the ZERO bits. In contrast, when $k = 2$ and the method reduces to frequency doubling (Fig. 3), the main lobe encompasses both normalized frequencies of interest (1/2 and 1) so that frequency discrimination may not provide a satisfactory detection method.

4) A result similar to point 3 is noted by comparing Figs. 5 and 6 for sinusoidal frequency modulation. Only in the latter case, where the coding frequencies differ by a factor of 4, can two distinct main lobes be observed for the continuous spectrum. It should be noted, however, that all sinusoidal frequency modulation schemes in which the coding frequencies are commensurate with the bit frequency will contain discrete spectral lines at the coding frequencies, each containing one-quarter of the total signal power.

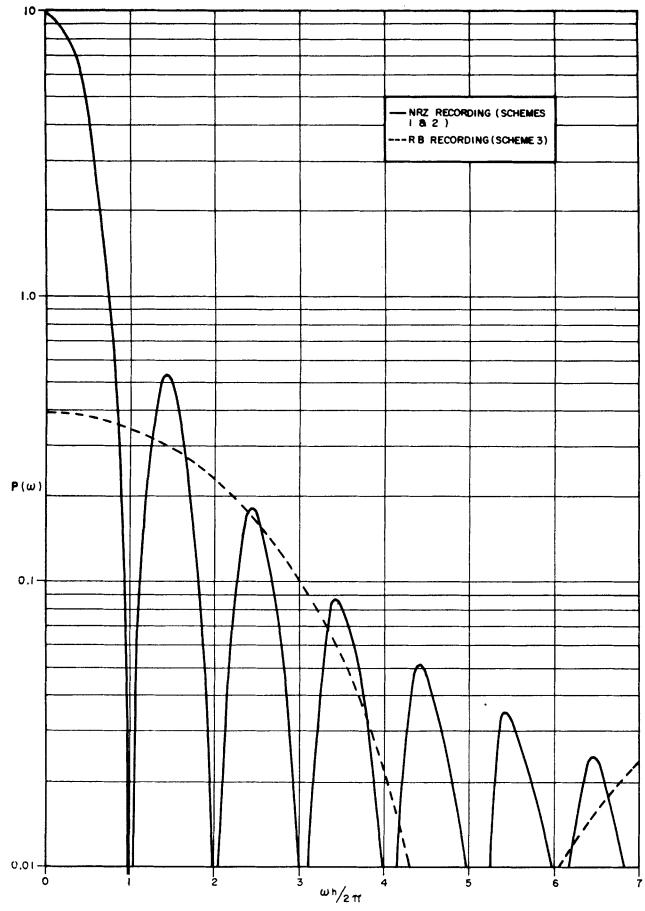

Fig. 2. Power density spectrum for saturation techniques.

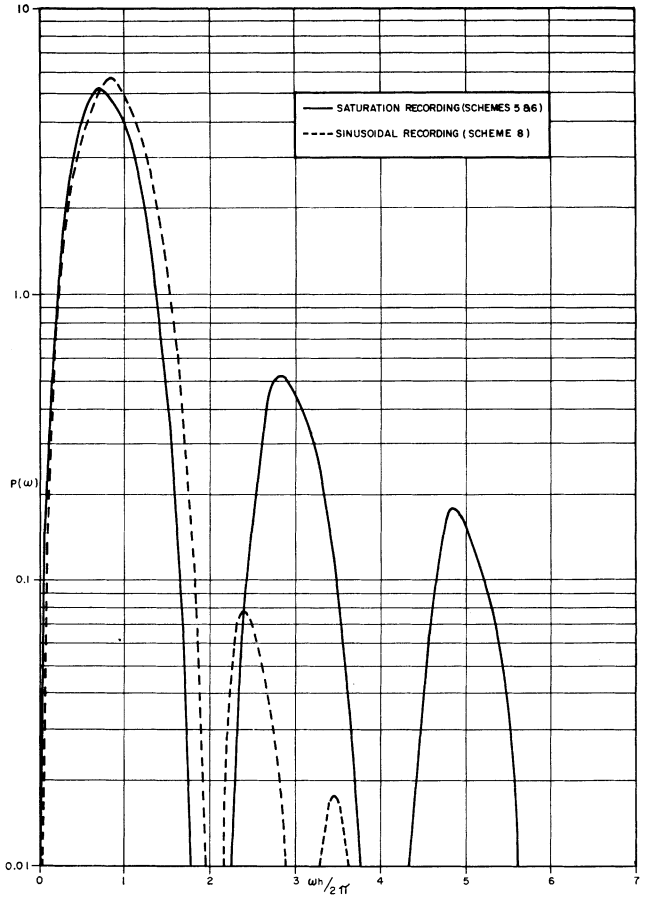

Fig. 3. Power density spectrum for phase-(shift)-modulation techniques.

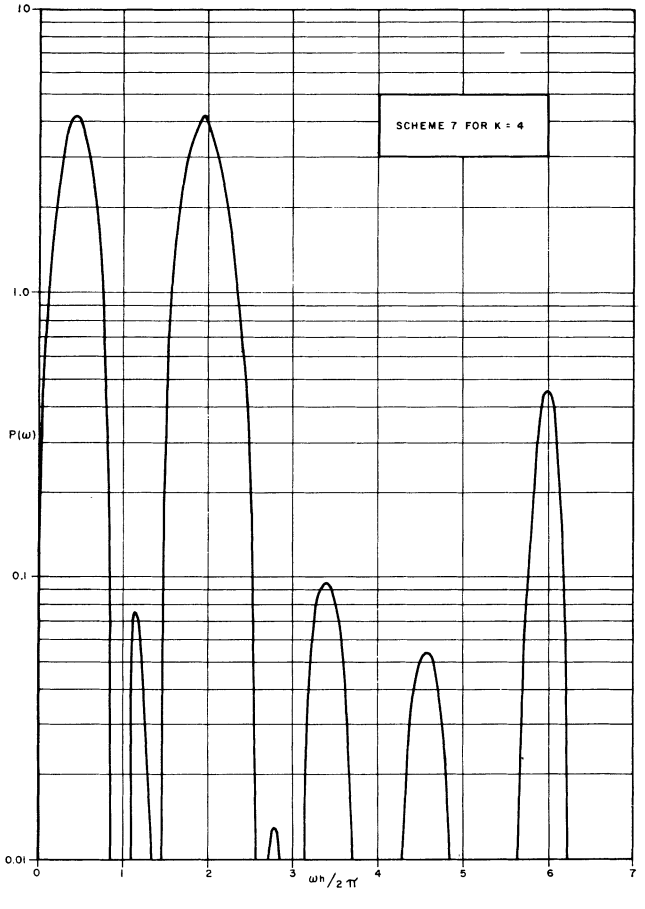

Fig. 4. Power density spectrum for high-frequency modulation saturation.

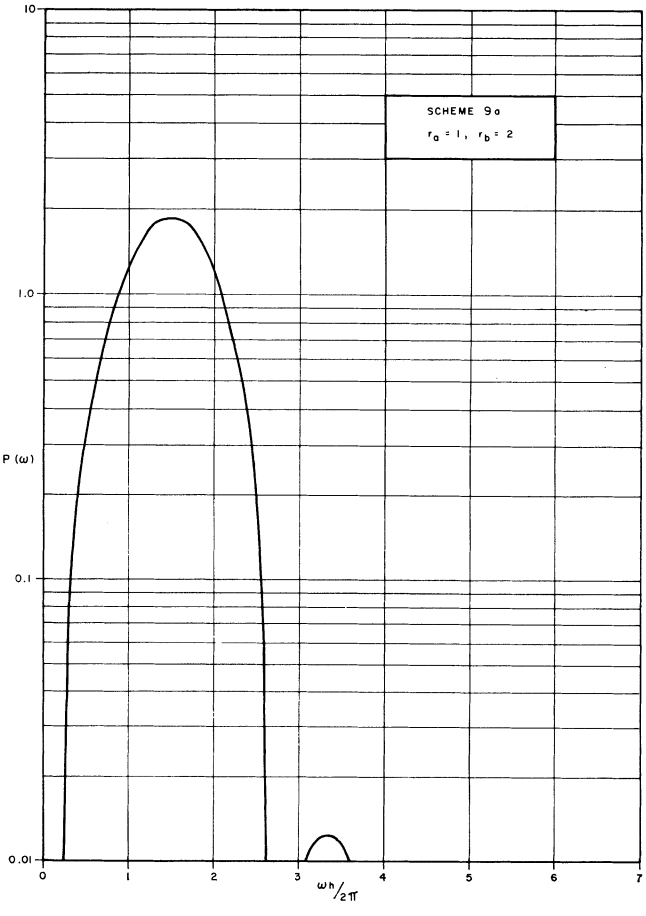

Fig. 5. Power density spectrum for sinusoidal frequency (shift) modulation: $r_b/r_a = 2$.

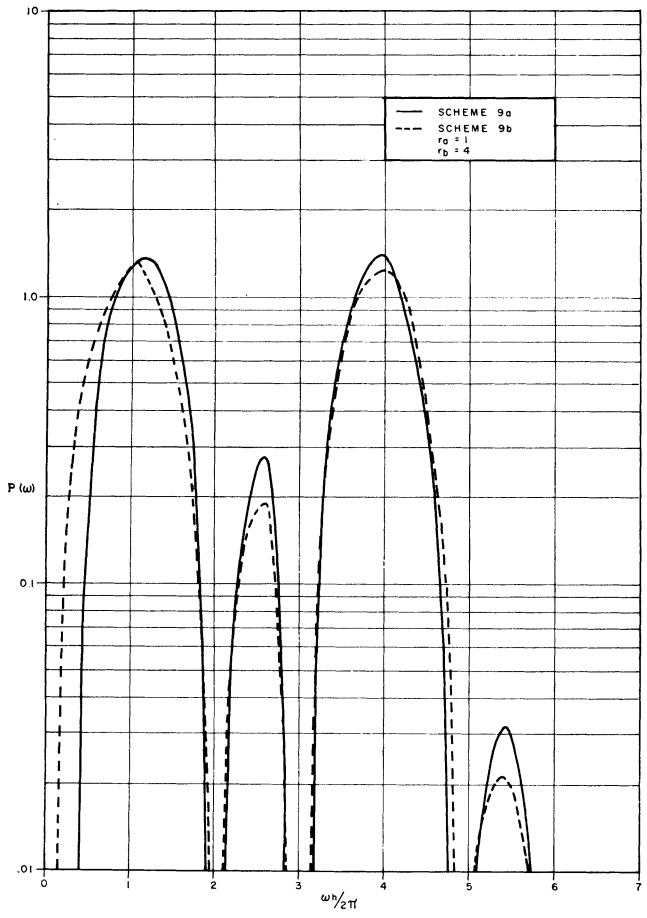

Fig. 6. Power density spectrum for sinusoidal frequency (shift) modulation: $r_b/r_a = 4$.

5) A comparison of the sine wave and cosine wave coding in Fig. 6 shows that the sine wave contains slightly more power in the sidelobes at higher frequencies. An explanation of this lies in the fact that in the case of cosine wave coding, the coded signal always has continuous slope in passing from one bit interval to the next, while the sine wave has only continuous magnitude. The difference is not very significant.

IV. Comparison of Spectra with Head Response

To appreciate fully the significance of the characteristics given in Section III, it is necessary to compare the spectra of the recorded signals with the frequency response of the filter representing the interaction between the recorded signal and the reading head. Of course, the voltage developed across the coils of the reading head is further amplified and modified by the electronics in the read-out system; however, the raw output signal has particular significance since the signal-to-noise ratio in any frequency band at the head output cannot be improved upon by further filtering. The squared frequency response of the reading head for a constant flux input at all frequencies is shown in Fig. 7 in terms of the normalized frequency used in the previous curves, and a tape velocity v of 120 in/s. The frequency response has the form[12]

$$V(\omega) = (\text{const.})\omega\left[\frac{1 - e^{-\omega d/v}}{\frac{\omega}{v}d}\right]\left[e^{-\omega a/v}\right] \cdot \left[L\left(\frac{g\omega}{2v}\right)\right] \tag{9}$$

where

d is tape depth
a is separation between tape and head

and

g is head gap width.

The gap loss function $L(x)$ in (9) is usually approximated by the function $(\sin x)/x$, but has been tabulated to greater accuracy by Wang.[12] In constructing Fig. 7, the effect of head to tape separation has been neglected, and the tabulated results of Wang for the gap loss function have been employed. Wang has carried out his calculations to $x = 1.25$, which includes part of the first sidelobe of the gap loss function, as well as the main lobe. The sidelobe only appears in the third curve in Fig. 7, at 3200 bits per inch (bpi). At frequencies near and above the first null, the primary signal loss is due to the gap loss, while at low frequencies, the loss is due to the differentiating action of the head, represented by the ω term in (9). All three curves in Fig. 7 actually represent the same frequency response, but have been normalized in terms of bit period h for comparison with the spectrum curves by substituting for the velocity $v = 1/hn$, where n is the number of bits per inch.

Since the coded bit sequence is a stochastic function, the power spectrum at the READ head output is given by

$$P_{\text{out}}(\omega) = P(\omega) \cdot |\, V(\omega)\,|^2 \tag{10}$$

and can be obtained by multiplying points on Figs. 2 through 6 with points on Fig. 7. This process has been carried out for NRZ and saturation phase-modulation recording, and the results are shown on Figs. 8 and 9, respectively. From these curves, the following conclusions can be drawn:

1) At low bit densities, many of the sidelobes present in the NRZ spectrum are preserved in the output spectrum; however, the near dc portion of the main lobe is highly attenuated, which accounts for much of the irregularities present in the output signal when long strings of zeros are present in the input. As the bit density is increased, more of the high-frequency content is also lost.

The loss of the high-frequency sidelobes in NRZ recording causes broad output pulses in place of the sharply defined, highly peaked pulses that are desirable for the peak detection methods usually employed with

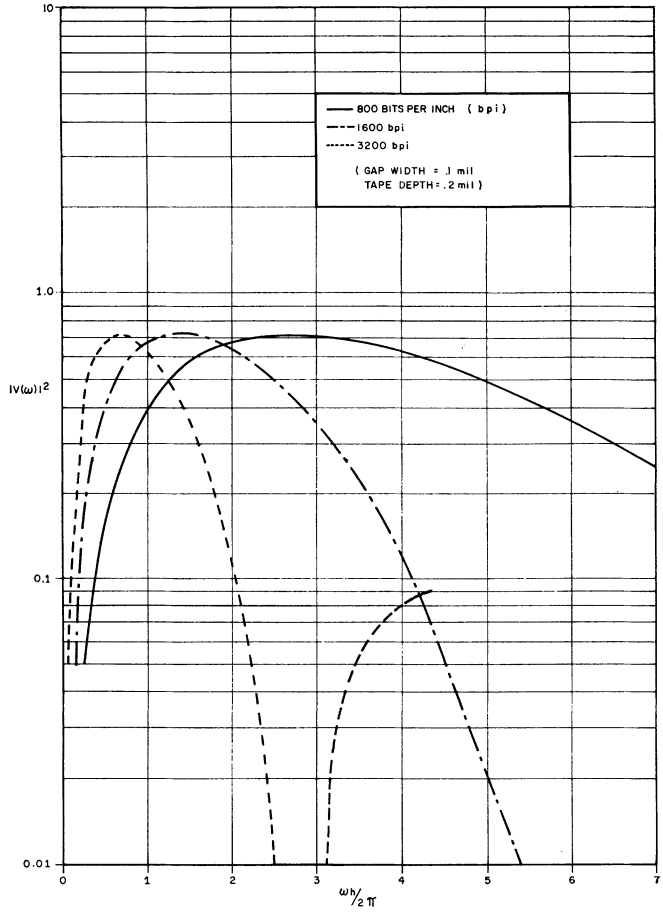

Fig. 7. Frequency response of reading head, normalized in terms of bit density.

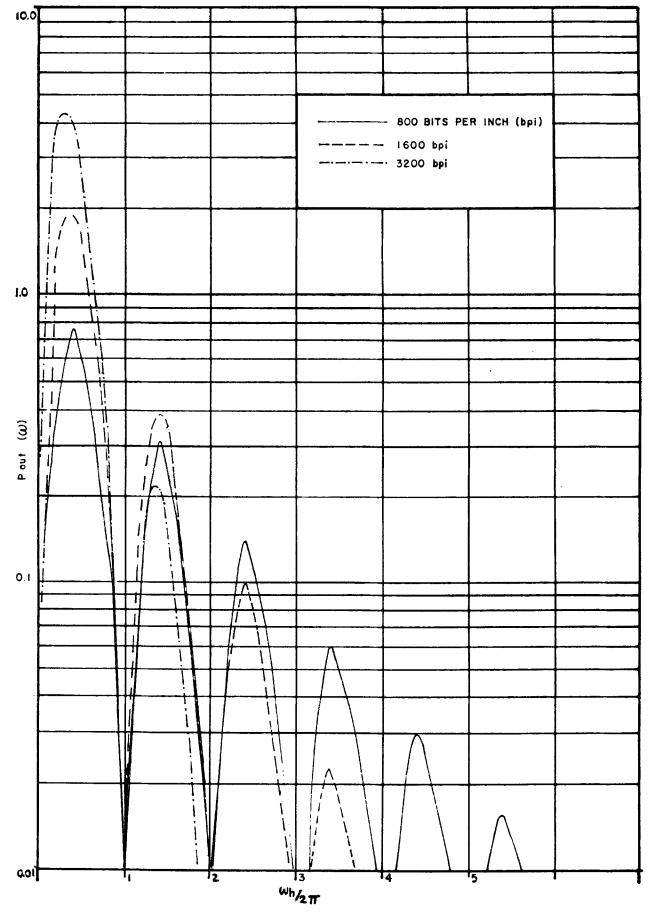

Fig. 8. Output power spectrum for NRZ1 recording (scheme 2).

this type of recording. (If the high-frequency content were preserved, the pulses would be the derivative of the tape magnetization, with its sharp transitions from positive to negative saturation.) As the density is increased, the increased overlap between pulses in neighboring bit intervals leads to large shifts in peak position within each bit interval for nonsymmetrical bit sequences and causes resolution problems. This phenomenon is usually referred to as intersymbol interference. With RB recording, the situation is even worse since the main lobe occupies such a broad spectrum and the sidelobes appear at correspondingly higher frequencies.

2) With the modulation techniques, near dc attenuation is not significant since little signal energy is present near dc. At high bit densities, the sidelobes of the signal produced by saturation phase modulation (Fig. 9) will be attenuated; however, for decoding purposes, the phase information contained in the main lobe component of the signal with power concentration between the normalized frequencies of 1/2 and 1 usually suffices for signal reconstruction, since the head response does not alter the phase characteristics of the signal, except for a uniform phase shift of 90° over all frequencies. At a bit density such that the main lobe of the head response drops off above a value of 2 (3200 bpi in Fig. 7), the main lobe of the head spectrum almost matches the main lobe of the signal spectrum. The output signal for sinusoidal phase modulation and saturation phase modulation will be almost equivalent at high bit densities, since the main lobes of their power spectra are so similar. Similar detection schemes should then be applicable to both modulation approaches.

If a long string of ones in phase modulation is compared to a long string in NRZ1 recording, it may be concluded that the *highest coding* frequency for NRZ1 is half that of phase modulation, in terms of the fundamental of the periodic square wave used for coding in each of the two cases. This consideration has led to the following statement in reference [1], page 55: "Basically, the highest frequency present in a phase shift modulation coding process is twice that in NRZ. The price of the extra bandwidth is gladly paid in many applications for the improved accuracy with which the data is derived." This statement can be misleading if applied to the bandwidth required in the output circuitry, in the light of the preceding discussion. We have

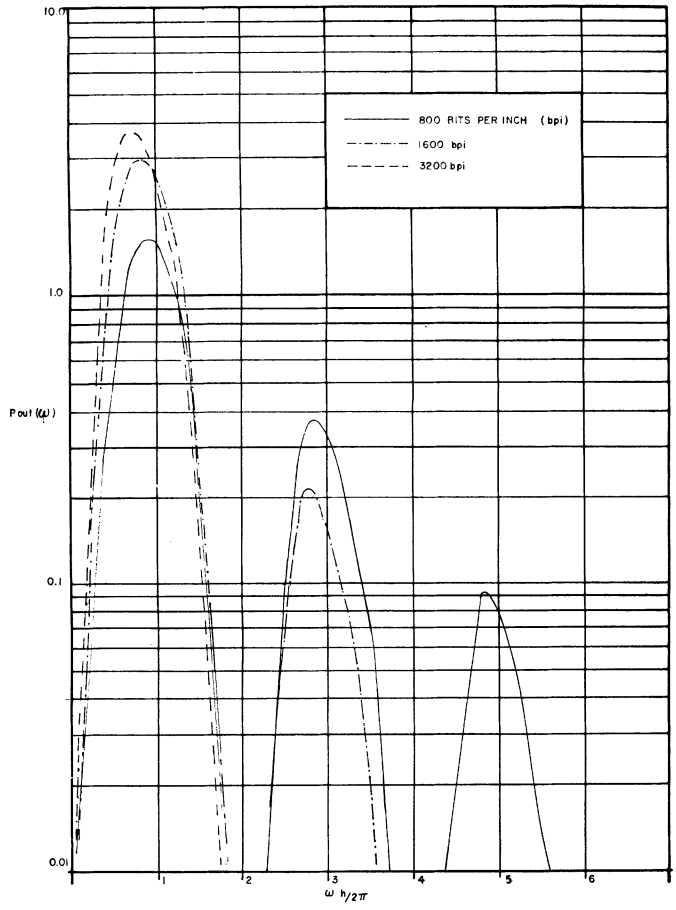

Fig. 9. Output power spectrum for saturation phase-modulation recording (scheme 5).

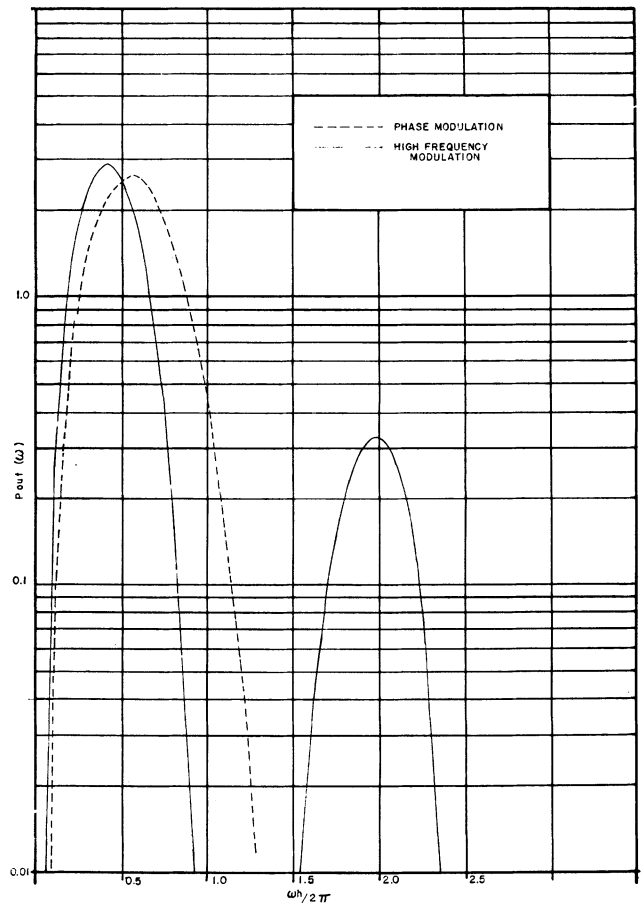

Fig. 10. Output power spectrum for schemes 5 and 7 at 6400 bpi.

indicated that the NRZ approach requires much greater bandwidth because of the necessity of preserving near dc information and high-frequency sidelobes. Athey's statement is correct when referring to the input circuitry required for generating the input current waveform to the recording head.

3) In considering the frequency-modulation techniques represented by Figs. 4 and 6, one might conclude that these FM techniques are useful only at lower bit densities than are the phase-modulation techniques, because of the presence of a main lobe at higher frequencies in the FM than in the PM techniques. The fact that the spectrum shows good separation between the basic two frequencies can be used to advantage, however, to achieve an even greater bit density. Since only odd harmonics of the lower frequency produce significant sidelobes, the upper frequency is chosen to be an even multiple of the lower frequency. Now, consider scheme 7 (Fig. 4). If the bit density and READ head parameters are chosen so that the first null for the reading head occurs slightly above the normalized frequency of 1, the output spectrum will preserve the low-frequency main lobe in the signal spectrum (centered near 1/2) while severely attenuating the main lobe centered near 2. This is shown in Fig. 10, in which one curve contains the result of multiplying Fig. 4 by the head response at 6400 bpi (same as curve C, Fig. 7, with the abscissa values divided by 2). The output signal will then contain a significant pulse for a ONE bit and low amplitude ripples for a ZERO bit, which can be further reduced by high-frequency filtering. This approach then resembles standard NRZ1 recording in that a significant pulse is only present for a ONE bit; however, it contains none of the low-frequency interference inherent in the NRZ1 approach. The broadening of the "one" pulses due to the absence of high-frequency sidelobes does not lead to the type of pulse-peak shifting present in NRZ1 recording, since a magnetization transition always occurs at the beginning of a bit period. Thus, the intersymbol interference in an asymmetrical bit sequence is much less serious.

Fig. 10 also shows that the main lobe for phase modulation has strong attenuation on its high-frequency side at the bit density of 6400 bpi. The power at the normalized frequency of 1 is down more than 7.5 dB from its value at 1/2, although the frequencies are almost equiv-

alent in the input spectrum (Fig. 3). While the phase-modulation approach requires that the first null of the head response be above the normalized frequency of 2 (3200 bpi in Fig. 7), the high-frequency approach of scheme 7 will operate for a null above 1 (6400 bpi), thus implying a factor of 2 improvement in bit packing density. The same reasoning applied to scheme 9 (Fig. 6) will lead to a packing density comparable to phase modulation.

Summary

Signal waveshapes in use for digital magnetic recording have been classified in one of two classes according to the behavior of the signal in neighboring bit periods. A general expression for the power density spectrum has been derived for each of the classes, assuming a random bit sequence, and applied to members of both classes. The waveshapes assumed were idealized square waves or sine waves. It may be valuable in future investigations to apply the equations to less idealized waveshapes and, in particular, to examine the effects of preemphasis on the high frequency portion of the input signal when linear recording techniques (sinusoidal signals) are used.[5]

Comparison of the signal spectra with the frequency response of the reading head showed the advantages inherent in modulation techniques which eliminate near dc signal power. Although phase-modulation techniques have received greater attention for application to magnetic recording, frequency-modulation techniques appear to show great promise. The latter techniques have a greater degree of freedom, allowing a wide choice of modulating frequencies and detection methods (e.g., FM discriminators, matched filters, peak detection as used by Chao,[4] etc.). In this paper, only commensurate modulating frequencies were considered; however, there is a growing literature in the communications technology field which deals with more general configurations. (See, for example, Anderson and Salz,[13] Von Baeyer and Tjhung,[14] as well as Bennett and Rice.[7]) Further study of the applicability of these results to magnetic recording is desirable.

A final word on nomenclature. No standard names have yet been adopted in the literature for the digital phase- and frequency-modulation techniques, although the names "phase-shift modulation" and "frequency-shift modulation" seem to be gaining favor. The latter names identify the techniques with the transmission techniques of phase-shift and frequency-shift keying (PSK and FSK), and help to distinguish them from the older analog techniques of frequency and phase modulation. The distinction begins to fade, however, when one considers low modulation frequencies with respect to a carrier, as in Von Baeyer and Tjhung.[14] In this paper, the various names were used interchangeably when it appeared that no confusion would result.

Appendix

In this Appendix, the power density spectrum $P(\omega)$ is derived for the two general classes of signals under consideration, starting with the Fourier integrals $S(\omega)$ given in (4) and (5).

Define $T(\omega) \equiv |S(\omega)|^2$.

From (4)

$$T_1(\omega) = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} e^{-j\omega h(n-m)} I_n I_m^*$$

where

$$I_n = \int_0^1 f_n(t) e^{-j\omega t} dt.$$

Then, from (1)

$$P_1(\omega) = \lim_{N\to\infty} \frac{2}{Nh} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} e^{-j\omega h(n-m)} E[I_n I_m^*]. \tag{11}$$

Now, if $n=m$, f_n is equally likely to be f_a or f_b, and

$$E[I_n I_n^*] = 1/2(|I_a|^2 + |I_b|^2).$$

If $n \neq m$, four combinations are equally likely, and

$$E[I_n I_m^*] = 1/4(|I_a|^2 + |I_b|^2 + I_a I_b^* + I_b I_a^*).$$

Equation (11) is most easily evaluated in terms of the variable $p=n-m$. Since there are $N-|p|$ terms in (11) for each value of p between $-(N-1)$ and $+(N-1)$, the summation calculus[9] and some tedious algebra lead to the result given in (6). Similarly, from (5),

$$T_2(\omega) = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} e^{-j\omega h(n-m)} d_{n-1} d_{m-1} I_{n+} I_{m+}^*$$

and

$$P_2(\omega) = \lim_{N\to\infty} \frac{2}{Nh} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} e^{-j\omega h(n-m)} \cdot E[d_{n-1} d_{m-1} I_{n+} I_{m+}^*]. \tag{12}$$

For this class of signals, the correlation between neighboring bit perlods must be accounted for in evaluating the expectation term in (12). With reference to Table II, one obtains the following:

For $n=m$,

$$E[(d_{n-1})^2 |I_{n+}|^2] = 1/2(|I_{a+}|^2 + |I_{b+}|^2).$$

For $n=m+1$,

$$E(d_{n-1} d_{n-2} I_{n+} I_{(n-1)+}^*] = 1/4(|I_{a+}|^2 - |I_{b+}|^2 + I_{b+} I_{a+}^* - I_{a+} I_{b+}^*).$$

For $n=m-1$

$$E[d_{n-1} d_n I_{n+} I_{(n+1)+}^*] = 1/4(|I_{a+}|^2 - |I_{b+}|^2 + I_{a+} I_{b+}^* - I_{b+} I_{a+}^*).$$

For $n-m \geq 2$, no correlation exists, and the product $d_{n-1}d_{m-1}$ is equally likely to take on the value $+1$ or -1 for any of the possible combinations of $I_{n+}I_{m+}{}^*$. Therefore,

$$E[d_{n-1}d_{m-1}I_{n+}I_{m+}{}^*] = 0, \quad |n-m| \geq 2.$$

Putting the above expectation values into (12), one obtains (7).

References

[1] S. W. Athey, *Magnetic Tape Recording*, NASA Technology Survey, Washington, D. C.: U. S. Government Printing Office, 1966, pp. 52–55.

[2] R. E. Young, "Frequency shift modulation," *Instruments and Control Systems*, vol. 38, pp. 109–112, April 1965.

[3] A. Gabor, "Digital signals at high density on magnetic tape," *Electronics*, vol. 32, pp. 72–75, October 16, 1959.

[4] S. C. Chao, "High density digital magnetic recording," *Proc. Internat'l Conf. on Magnetic Recording*, London, England, pp. 47–49, July 1964.

[5] E. Hopner, "High density recording using nonsaturation techniques," *IEEE Trans. Electronic Computers*, vol. EC-13, pp. 255–261, June 1964.

[6] J. E. Gillis, "A technique for achieving high bit packing density on magnetic tape," *IEEE Trans. Electronic Computers*, vol. EC-13, pp. 112–117, April 1964.

[7] W. R. Bennett and S. O. Rice, "Spectral density and autocorrelation functions associated with binary frequency-shift keying," *Bell Sys. Tech. J.*, vol. 42, pp. 2355–2385, September, 1963

[8] S. O. Rice, "Mathematical analysis of random noise," in *Noise and Stochastic Processes*, N. Wax, Ed. New York: Dover, 1954.

[9] R. W. Hamming, *Numerical Methods for Scientists and Engineers*. New York: McGraw-Hill, 1962, Ch. 3.

[10] D. W. Chapman, "Theoretical limit on digital magnetic recording density," *Proc. IEEE*, vol. 51, pp. 394–395, February 1963.

[11] I. S. Sokolnikoff, *Advanced Calculus*. New York: McGraw-Hill, 1939, p. 54.

[12] H. S. C. Wang, "Gap loss function and determination of certain critical parameters in magnetic data recording systems," *Rev. of Sci. Instr.*, vol. 37, pp. 1124–1130, September 1966.

[13] R. R. Anderson and J. Salz, "Spectra of digital F.M.," *Bell Sys. Tech. J.*, vol. 44, pp. 1165–1189, July 1965.

[14] H. J. Von Baeyer and T. T. Tjhung, "Effects of pulse shaping on digital F.M. spectra," *1966 Proc. Nat'l Electronics Conf.*, vol. 22. pp. 363–368.

Crosstalk (Noise) in Digital Systems

IVOR CATT

***Abstract*—As digital system speeds increase and their sizes diminish, it becomes increasingly important to understand the mechanism of signal crosstalk (noise) in interconnections between logic elements. The worst case is when two wires run parallel for a long distance. Past literature has been unsuccessful in explaining crosstalk between parallel wires above a ground plane, because it was assumed that only one signal propagation velocity was involved.**

This paper proves that a signal introduced at one end of a printed wire above a ground plane in the presence of a second parallel (passive) wire must break up into two signals traveling at different velocities. The serious crosstalk implications are examined.

The new terms *slow crosstalk* (SX), *fast crosstalk* (FX) and *differential crosstalk* (DX) are defined.

***Index Terms*—Crosstalk (noise) in digital systems, Directional Couplers, graphs of characteristic impedance and crosstalk in Stripline and Microstrip, interconnection of 1-ns logic gates, multilayer printed circuit boards, resistive paper analog for measuring *L*, *C* and *Z*.**

I. Introduction

THE PROPAGATION delay and signal rise time of logic circuits now available are in the region of 1 ns. To gain the speed advantage that these circuits offer us, it is necessary to reduce the physical size of a system to less than one cubic foot, so that the effective speed will not be greatly reduced by propagation delays in the interconnecting lines (reference [1], page 104). The problem of how to fit all the interconnections between logic elements into such a confined space is a difficult one, and multilayer printed circuit boards seem to be necessary.

Fig. 1 shows a cross section of such a board. In order to keep signal cross coupling (noise) down to a reasonable level, (say 10 percent of signal amplitude), at least one voltage plane must separate successive layers of signal lines. At a frequency of 1 GHz, skin depth in copper (reference [2], page 238) is 2×10^{-4} cm ($=0.00008$ in.). At 10 GHz, skin depth drops to 7×10^{-5} cm ($=0.00003$ in.). At present, it is not practicable to laminate copper less thick than 0.001 in. (which is more than ten times the skin depth). This means that in practice, a signal traveling down a signal line and returning by the immediately adjacent voltage plane(s) will not penetrate beyond the plane(s), and each voltage plane will screen signals above it from signals below it with negligible crosstalk through the voltage plane. So long as a signal is transmitted down between a signal line and the voltage plane(s) immediately above or below it, the only

Manuscript received August 29, 1966; revised July 5, 1967.

The author is with Sperry Semiconductor Division, Norwalk, Conn.