A Method of High Density Recording on Flexible Magnetic Discs

Robert C. Franchini and Donald L. Wartner

Shugart Associates
Sunnyvale, California

To increase the amount of data stored on one flexible disc requires either more flux changes per inch of track or a better code; the latter is the obvious choice because the former is close to its physical upper limit

Ed Note: *Computer Design* readers will observe that this Tech Note is similar in many respects to another Note on double-density recording that appeared in the September issue of this magazine. Although the conclusions stated here are different from those of the previous author, both arguments are logical and well-presented. Therefore, the editors, who usually do not devote space to the same topic in consecutive months, thought both points of view were worthy of attention. Readers who design or use flexible disc drives or related equipment will want to consider both opinions carefully, and will be interested in a third view that will appear in a forthcoming issue.—WBR

During the last two years, the growth in use of flexible disc files (often called floppy discs) has been phenomenal. More and more applications are being found daily. Discs are useful in applications that require greater performance and reliability than tape cassettes or cartridges can provide, and that cost less per function than removable hard disc packs, such as keyed data entry, point-of-sale systems, remote batch terminals, microprogram loading, error logging, minicomputer program and data storage, and word processing.

In many of these applications, storage requirements are so extensive that they arouse interest in increasing the storage capacity of individual flexible discs. Increasing the linear density along the tracks on the disc surface is the easiest way to do this. To implement such a linear density change requires either an increase in the number of flux changes per inch (fcpi) or a more efficient code. The latter is the obvious choice in a disc file that is already operating at densities as high as 6400 fcpi.

Most flexible disc files today use frequency modulation (FM) encoding, which writes data bits (D) at the center of the bit cells, and clock bits at the leading edge of the bit cells. Each clock bit and each binary 1 is represented by a flux change on the disc's magnetic surface, that appears at the output of the detector circuit as a rectangular pulse. Binary 0s are represented by no change in flux and no pulse at the detector. Clock frequency alone is f, corresponding to a stream of pulses from the detector separated by time T1 when a series of 0s is read. When reading 1s, the frequency jumps to 2f, and the pulse separation becomes T/2.
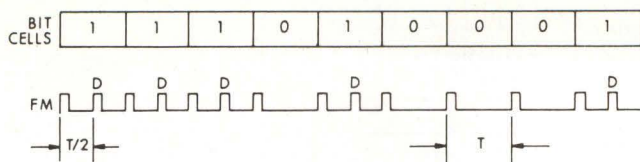
Fig. 1 FM encoding. In frequency modulation encoding of binary data, every bit cell has a clock pulse at its leading edge, and a 1 is stored in the form of a pulse at the center of the bit cell. Method imposes a penalty in the form of unneeded clock pulses
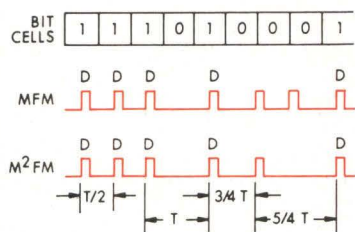

Fig. 2 More efficient codes. Modification of FM (top trace) eliminates clock pulse except between two cells containing 0s. Further modification (bottom trace) also eliminates clock pulse in second cell following a 1. Both codes eliminate unnecessary pulses, and permit recording on disc surface at higher density than ordinary FM

Modified Codes

In FM encoding a penalty is paid in clock bits (Fig. 1). Thus, if the code could be made more efficient by eliminating some of the clock bits, higher information densities could be achieved. One such technique, devised several years ago, is commonly called modified frequency modulation (MFM). This encoding scheme has been used successfully on high performance disc drives such as the IBM 3330 and 3340. Rules for MFM cause data bits (D) to be written at the center of the bit cells, as in FM, and clock bits to be written at the leading edge of the bit cells, but only when both the previous cell and the present cell store binary 0s. Since the sole function of clock bits is to synchronize the clock in the detector circuit in the ab-

sence of binary 1s in the data stream, they are not needed when 1s are present; leaving them out permits a reduction of the average spacing of data bits, thus increasing the apparent density of data stored on the disc without exceeding the physical limits for data and clock bits imposed by characteristics of the medium.

Additional modification of MFM enables more improvements in encoding efficiency through further optimization of the clock signals. This encoding scheme is commonly referred to as modified-modified frequency modulation (M^2FM); it writes data bits (D) at the center of the bit cell, as before, and clock bits at the leading edge of a bit cell, but only if the previous bit cell contains neither a 1 nor a clock bit, and if the present bit cell also does not contain a 1.

In both modified forms of FM a given amount of data can be recorded in half the space (Fig. 2); the basic period between pulses is halved, corresponding to the doubling of the write oscillator frequency, but the number of flux changes per inch is the same as for basic FM. Thus the bit cell in MFM or M^2FM is half the size of that for FM, and for a given track velocity the data transfer rate is doubled.

Disadvantages of Modified FM

Both MFM and M^2FM can double the amount of data stored on a given length of track or a given area of disc, and double the rate at which it is transferred to or from the disc. However, both codes also have disadvantages.

Because the MFM or M^2FM bit cell does not always contain a clock bit at its leading edge, encoding and decoding require more complex electronic circuitry. The most reliable method of separating MFM or M^2FM encoded data is through use of a phase-locked oscillator (PLO) (Fig. 3). Once synchronized, the PLO tracks recorded information and generates separate "windows" for data bits and clock bits. Staying in synchronization with small, slow variations in disc speed, it averages quick variations caused by bit shift.

Bit shift is characteristic of all magnetic discs and tapes, including single-density flexible discs, but can be more serious with MFM and
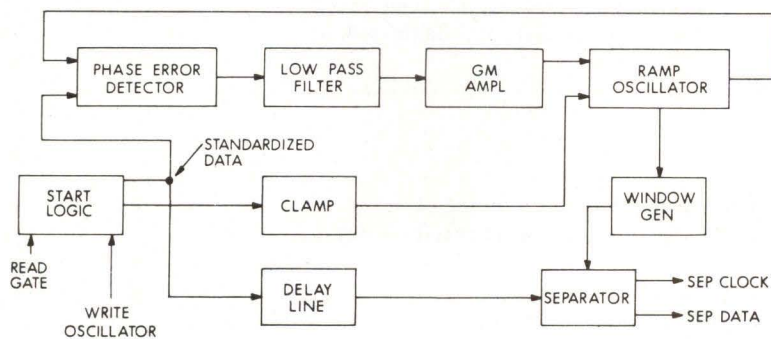

Fig. 3 Phase-locked oscillator. This circuit tracks recorded information and generates separate data and clock signals from it, even when, as with MFM and M^2FM, some recorded clocks have been omitted

M²FM. It is a real or apparent tendency for a flux reversal to be read slightly before or after its nominal time. Because the number of flux changes per inch is the same with these high density codes as with FM, the amount of bit shift in absolute terms is the same; but because the bit cell is half the size, the system can tolerate only half as much bit shift without error. Therefore, to reduce the amount of bit shift, a write precompensation circuit is added to the controller logic for a floppy disc drive to compensate for known bit shift patterns.

Since most bit shift is predictable, on the basis of pulse superposition theory, the data pattern can be analyzed in a shift register (Fig. 4) during the write process, and data and clock bits can be written either early, on time, or late. If the pattern would cause a late bit shift, the bits are written early; if the bit shift was early, the bits are written late. Then, during the read process, the bits appear more nearly on time within the PLO window.

Thus the principal disadvantages of both high density codes are minimized, leaving only the problem of deciding which one to use in a particular system. In our opinion, the right code is M²FM. In both codes, minimum data bit separation is $T/2$; but minimum clock bit separation in MFM is also $T/2$, whereas in M²FM it is $(3/4)T$. Although both data and clock bit streams are in the same channel, they can be considered separately. Both streams are subject to bit shift, which is inversely proportional to bit separation, whereas the effectiveness of write precompensation is directly proportional to bit separation. Since MFM encoding has both a $T/2$ clock bit and $T/2$ data bit separation, bit shifts in both fields have nearly the same magnitude. Consequently, the system designer must theoretically use a symmetric PLO window.

In M²FM encoding, however, the minimum separation of clock bits is greater than that of data bits, so that the clock field has less bit shift than the data field. The system designer can take advantage of this difference by designing his PLO data separator with an asymmetric window that has a 60/40 ratio. This asymmetry gives greater system margin and reliability.
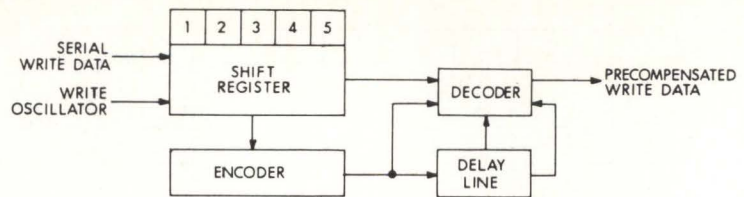
One factor that affects the ratio

Fig. 4 Write precompensation. Data pass through shift register before being written; if data pattern implies that excessive bit shift would occur when it is read, it is written slightly early or late to compensate for timing error
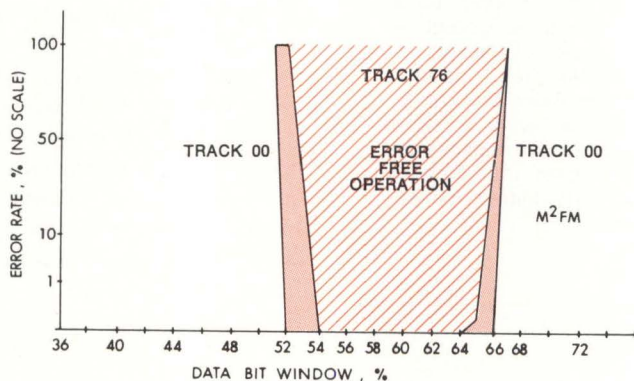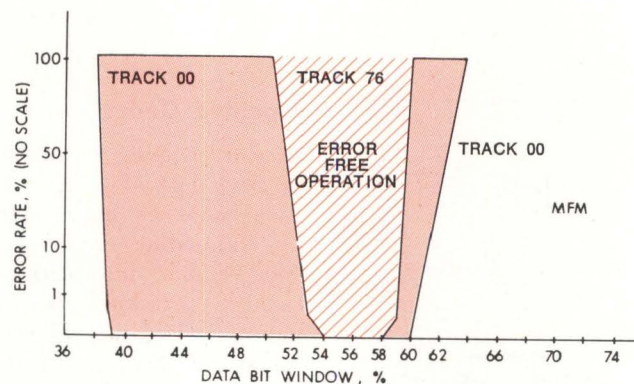

Fig. 5 Margin measurements. Data window for innermost track (crosshatched) is wider for M²FM than for MFM, and is more nearly equal to width of outermost track window, indicating substantially better reliability

of the two windows is the number of bit patterns that have to be checked to determine precompensation. MFM has four such patterns, while M²FM has only three; and one of the three requires much less precompensation (see Comparison of Write Precompensation). Measuring the effect of the window ratios in both encoding methods shows how much more tolerance is found in M²FM (Fig. 5).

These diagrams show the extent

to which the data window, as a percentage of the total bit period, could be varied without encountering errors. The extreme left and right edges of the tinted portion represent measurements of error rates at the disc's outermost track (Track 00). The innermost edges of the tinted portion show error rate measurements at the disc's innermost track (Track 76). The cross-hatched area represents error-free operation on

108

all tracks. In any given system the same read/write head serves both tracks as well as all intermediate ones, so the innermost track, which has the highest density, rules the system's performance.

For MFM the error-free windows are from 54 to 58% of the bit period; centered on 56%, the margin is only ±2%. As pointed out previously, theory predicts the center of this window at 50%. The difference is due to a small maladjustment in the test setup, which doesn't affect the narrowness of the window or the spread between inner and outer tracks. On the other hand, for M²FM the window extends from 54 to 64%, so that the margin is ±5%. The fact that the shaded portion of the M²FM diagram is only slightly wider than the crosshatched portion also shows that the error rate is more nearly the same for all tracks than it is in MFM.

Similar tests with the read/write head purposely misaligned relative to the track it was reading showed a much smaller error rate for a given misalignment, or a much larger tolerable misalignment for a given error rate, for M²FM than for MFM.

Summary

There are many approaches to increasing capacity of flexible disc drives. By far the simplest and most straightforward approach is to use more efficient encoding. Although a number of encoding schemes are available to the system designer, one that offers the greatest system margin and hence data handling reliability is most advantageous to him. We believe the M²FM encoding system satisfies this requirement, and our experience with flexible disc drive systems in the field supports our judgment. □