# ENCODING/DECODING TECHNIQUES DOUBLE FLOPPY DISC CAPACITY

Double-density storage increases standard disc capacity without requiring modifications to the drive unit. Evaluation of different encoding, decoding, and format schemes shows how each impacts critical design considerations

**John F. Hoeppner and Larry H. Wall** Shugart Associates, Incorporated, Sunnyvale, California

Ever increasing data storage requirements have caused designers to investigate new methods of expanding the 410k-byte single-density capacity of the standard 8" (20-cm) floppy disc. Double-density encoding, double-sided recording, and double-track density are the methods most often evaluated. Of these three, only double-density encoding has the advantage of doubling the system disc capacity with the same floppy disc drive used for single density. Additionally, improvements in head and media resolutions have made double-density encoding a highly reliable storage technique.

To double disc data density, one of several different encoding and decoding schemes can be implemented. Of these, modified frequency modulation encoding has become virtually the industry standard because of support by IBM as well as by several floppy disc controller chip manufacturers. With this method, twice as many data bits can be written on the disc medium than with single-density encoding. This increased storage impacts the design of the disc controller and the format of the medium. It also makes the decoding of data bits more susceptible to bit shift problems, in which bits are shifted away from their nominal positions. These effects must be carefully considered in the design of a double-density disc controller if soft error rates comparable to those of single-density drives (1 in $10^9$ bits) are to be maintained.

## Single-Density Encoding Schemes

The fact that double-density controllers are more complex in design is easily explained by examining the differences between single- and double-density encoding. Parameters of various single- and double-density schemes are compared in the Table.

Frequency modulation (FM), based on the IBM 3740 method of encoding (Fig 1), is the industry standard

**Single- and Double-Density Encoding Parameters**

| Parameter | Recording Technique FM | MFM | M²FM | GCR |
|---|---|---|---|---|
| Bit Cell | 4 μs | 2 μs | 2 μs | 1.6 μs |
| Flux Changes/Cell | 2 | 1 | 1 | 1 |
| Flux Changes/in | 6536 | 6536 | 6536 | 8170 |
| Data Rate (kilobits/s) | | | | |
| Storage Device | 250 | 500 | 500 | 625 |
| System | 250 | 500 | 500 | 500 |
| Frequency Ratios | 2/1 | 2/1 | 2.5/1 | 3/1 |
| Bit to Bit Spacings | 2 μs | 2 μs | 2 μs | 2 μs |
| | 4 μs | 3 μs | 3 μs | 3.2 μs |
| | | 4 μs | 4 μs | 4.8 μs |
| | | | 5 μs | |
| Diskette Capacity (kilobytes) | 410 | 820 | 820 | 820 |

method of single-density encoding. With this scheme, a clock bit is written at the beginning of each bit cell, and data bits are written between clock pulses. Each bit cell is 4-μs wide with the data bit written in the middle of the cell, or 2 μs after the clock bit. Effectively, this scheme amounts to two flux changes per bit cell.

To decode a data bit in single-density recording, the data separator generates a 2-μs data window 1 μs after the clock pulse. Therefore, the data window—centered around the expected position of the data bit—allows the presence or absence of a data bit to be detected. Due to the 2-μs size, bits, even if shifted, are still likely to remain inside the window.

A constant bit cell reference, provided by the clock bit, simplifies encoding and decoding with this scheme (Fig 2). Many large scale integrated (LSI) disc controller chips are available to handle single-density data encoding and disc drive interfacing. The data separator circuit required to decode single density usually consists of simple timing circuits that generate the 2-μs data window.

## Higher-Density Encoding Schemes

Three double-density encoding schemes are presently used to increase floppy disc capacity: modified frequency modulation (MFM), modified-modified frequency modulation (M²FM), and group coded recording (GCR). Each scheme increases disc capacity by replacing clock bits with data bits, using slightly different techniques. In each scheme, data rate as well as drive capacity is increased. The increased data rate might not affect controller design since most controllers incorporate a data buffer to transfer data asynchronously. Encoding schemes, however, do affect disc controller design.

### MFM Encoding

With available head and media technology, MFM coding is the most easily implemented and may bec the industry double-density standard. It is used in IBM System/34 and in all available double-density disc controller chips. MFM encoding doubles floppy data capacity to 820k bytes by replacing the clock positions (used in FM encoding) with data bits (Fig This scheme reduces the size of the bit cell by one-hal 2 μs, thereby doubling data bit capacity.

Clock bits are still used, but are written only wi data bits are not present in both the preceding the current bit cell. As a result, there is only one change per bit cell. Clock bits are written at the beg ning of the bit cell, while data bits are written in middle, or 1 μs after the bit cell's leading edge.

To decode data bits in MFM encoding, a data separa must generate a 1-μs data window and a 1-μs data w dow complement for a clock window. Because not ev bit cell has a clock pulse, the data clock windows c not be timed from the clock pulse. Instead, the d separator must continuously analyze the bit position side the windows so that the data clock windows rem synchronous with the data/clock bits. Ideally, the bit w appear at the center of the window. However, data b can shift outside the 1-μs data window due to bit-sh effects.

Consequently, data separator design, as well as overa disc controller design, for double-density encoding more complex than for single-density encoding. Prese LSI controller chips can handle the drive interface, doubl density encoding function, and bit-shift pattern detectio However, bit-shift compensation circuits and a high res lution analog data separator must be added (Fig 2

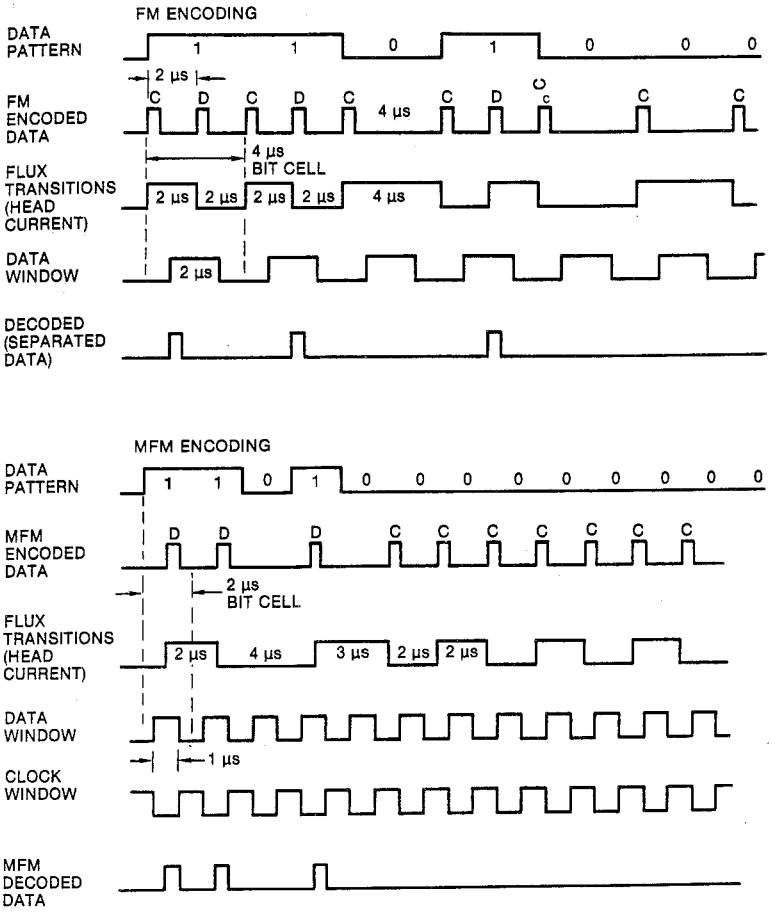



Fig 1 FM vs MFM encoding schemes. MFM encoding doubles number of data bits on disc media by replacing clock bits used in FM encoding with data bits. Scheme reduces bit cell width from 4 to 2 $\mu$s and reduces data window to 1 $\mu$s

Despite these constraints, disc controller design for MFM is simpler than that for either of the other two double-density encoding schemes.

## M²FM Recording Scheme

Until recently, M²FM has been the commonly used double-density encoding scheme, because the resolution of the medium and the read/write head was not adequate for the 1-$\mu$s data window used in MFM. In M²FM, a clock is written only if no data or clock bit is present in the preceding bit cell, and no data bit occurs in the current cell. Because clock pulses are relatively isolated on the medium, the effect of bit shift on clock pulses is minimal. Therefore, a narrower clock window can be used to decode the clock pulse. The width of the data window can thus be increased to 1.2 $\mu$s, which allows more margin for shifted data bits.

Today's ceramic based read/write heads have much better resolution than those used in the past. This head design reduces the effects of bit shift, and makes the window margin provided by M²FM unnecessary. Additionally, M²FM is subject to a droop problem, which occurs in the read amplifier circuit when a low frequency pattern is read.

## GCR Encoding

GCR encoding evolved from methods used in magnetic tape recorders. This method translates four data bits

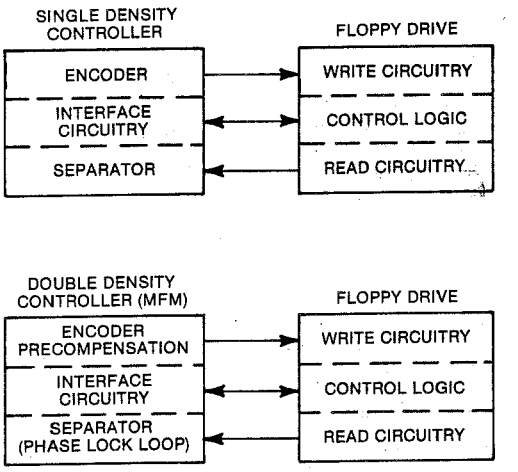



Fig 2 Single- vs double-density circuits. Effects of bit shift require more complex circuits for double-density floppy disc controllers than for single-density controllers. In addition to different encoding circuits, bit-shift compensation circuits are necessary, as well as high resolution analog data separator

into a 5-bit code during a write (Fig 3). During a read, the 5-bit code is retranslated to four data bits; no clock bits are generated. Using data rates specified by drive manufacturers, this scheme results in an 80% increase in density rather than a 100% increase. If the data rate is increased to provide double density, the number of flux changes per inch increases beyond reliable limits for floppy disc drives. This method requires more circuits to code, decode, and provide necessary lookup tables, and costs more than either of the other two double-density encoding schemes.

## Bit-Shift Calculations

Bit shift occurs on any NRZ (nonreturn to zero) recorded medium—single or double density. Using the NRZ method, data are recorded by reversing polarity of current flow for each flux transition. This shift is, however, more noticeable with double density due to smaller bit cells and resulting smaller data/clock windows. Some aspects of bit-shift phenomena are predictable; other aspects are not.

Predictable bit-shift effects result from normal read/write head operation. Data are written when the read/write head generates a flux change in the gap of the head, which causes a change in magnetization of the medium oxide. In reading, a current is induced into the read/write head when a flux transition on the medium is encountered. The current change is not instantaneous, since it takes a finite time to build up to peak and then to return to zero (Fig 4). If flux transitions are close together, the current buildup after one flux transition declines, but it does not reach zero before a second transition begins. When current pulses are summed by the read/write head, the peaks are shifted. A negative flux change, for example, appears late because it has been added to the remnant of a positive transition.

This type of bit shift is predictable. Spacing between bits results in greater bit shift on the inner tracks (43 to 76) of the floppy disc (Fig 5). On those tracks, bits can be expected to shift up to 350 ns. On the outermost tracks (00 to 42), bit shift is negligible.

However, other causes for bits shifting from the position where they are written are not predictable. Variations in disc drive rotational speed can cause bits to shift by a constant, but unpredictable, amount. Drives are specified as 360 r/min ±2% rotational speed. The 2% variation includes a 1% allowance for 120-Vac line frequency variations and a 1% allowance for belt pulley tolerances. These variations will shift bits by as much as ±40 ns from their originally written positions.

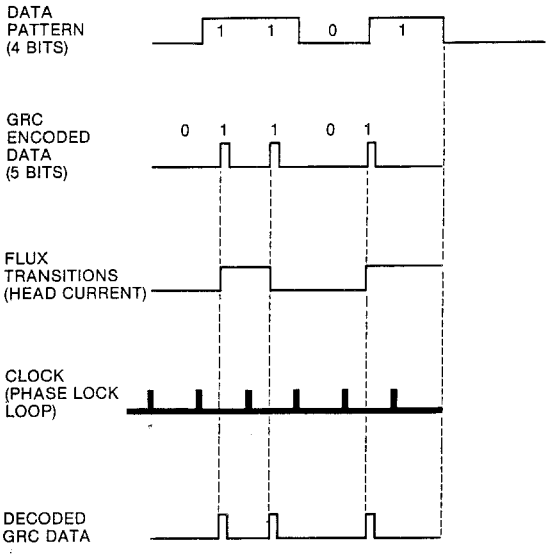



Fig 3 GCR encoding scheme. Four data bits are translated into 5-bit code during write; for example, data pattern 1101 is encoded into serial bit stream 01101, according to GCR encode rules. To decode, data window is generated around expected position of each bit. Result is serial read data of 01101, which must be decoded to 1101 by lookup table circuitry

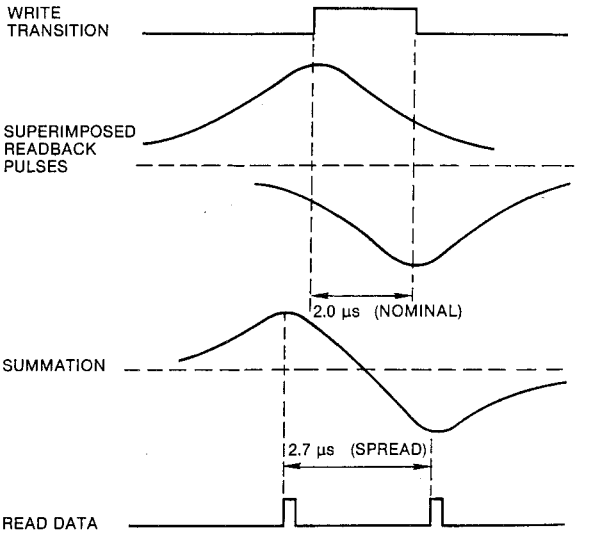



Fig 4 Bit-shift effects. During reading, bits are shifted predictably from their written positions as result of normal read/write head operation. When flux change occurs on medium, current is induced into read/write head. Current change is not instantaneous, since finite time is required to build up to peak and then to return to zero. If flux transitions are close together, current buildup after one flux transition declines, but does not reach zero before second transition begins. Peaks shift when current pulses are summed by read/write head

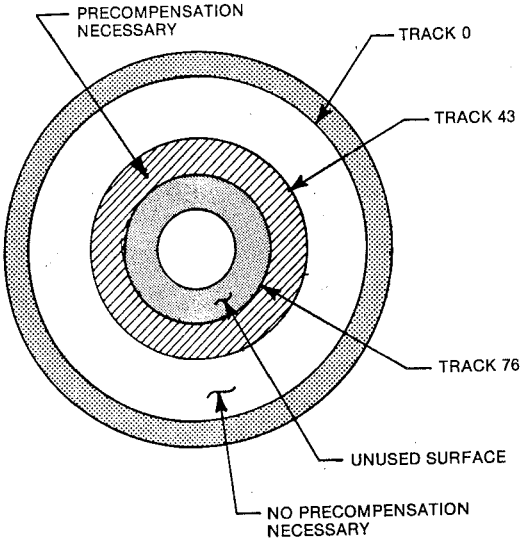



Fig 5 Bit shift vs track position. Significant bit shift occurs on innermost tracks of floppy disc (43 to 76) because bits on these tracks are packed closer together and therefore are more susceptible to bit shift. On outermost tracks (00 to 42) bit density is less, so that predictable bit shift is minimal

Instantaneous speed changes, although smaller than rotational speed changes, also cause bits to appear shifted; this shifting effect varies from drive to drive. Write over, or incomplete erasure of, previously recorded data can also shift bits. These shifts can range up to ±50 ns: again, they are variable and unpredictable. Other variable components of bit shift include electrical noise, radial alignment (offtrack), and nonsymmetry of the read/write head and associated electronics.

A bit shift of up to ±450 ns can be expected. This number is obtained using observed figures for predictable and unpredictable components of bit shift.

## Double-Density Controller Chip Design

Both predictable and unpredictable bit shifts must be accommodated in the design of a double-density disc controller if data are to be read reliably. Controller chips reliably perform MFM encoding and decoding. To compensate for predictable bit shift, however, compensation circuits must be added, and the data separator circuit must generate and position a high resolution data window so that a bit can be read reliably even if unpredictably shifted.

Despite the problems posed by bit shift, double-density controller design has been simplified with the recent introduction of LSI controller chips. For example, the NEC uPD765 and the Western Digital FD1791 perform MFM encoding and decoding as well as pattern detection to provide signals that identify which bits will be predictably shifted. These chips also perform serial/parallel conversion, generate cyclic redundancy check (CRC) characters for error detection, provide write current switches to reduce write current as density increases on the track, perform address mark detection for the IBM double-density format, and furnish intelligent status reporting. Both chips interface to commonly available 8-bit microprocessor based interface systems. Although they differ in software overhead and in hardware interface signals, the two chips are functionally the same.

Discrete logic circuit implementation of the same controller chip functions typically requires at least 60 to 70 integrated circuits (ICs), including a dedicated microprocessor, although greater format flexibility is possible than with the LSI chips, which require fixed gap lengths in the formats. Bipolar logic implementations of the same controller chip functions also result in a larger part count. Because some error detection functions can be performed in software, the bipolar parts count is somewhat less than that of discrete logic parts. Still, either method is substantially more complicated than the LSI chip implementation.

### Bit-Shift Compensation Circuits

Since some aspects of bit shift can be predicted, it is possible to compensate for this shift within the design of the disc controller. Two methods currently being used are precompensation and postcompensation.

With precompensation, bits are deliberately shifted in the direction opposite that of the expected shift. As data are being written, the LSI controller chip detects bit patterns. From these bit patterns, the controller calculates which bit will shift in which direction. Since bit shift is negligible for the first 43 tracks of data, the controller issues no precompensation signals until after track 43.

For example, a 4-bit pattern of 0110 on an inner track would cause the third bit to appear as much as 350 ns later than its nominal position. The controller chip, after detecting this late bit-shift pattern, would generate an early signal, indicating that the third bit should be written earlier to make it appear closer to its nominal position when read. Conversely, if the third bit were going to appear early, a late signal would be generated so that the bit could be written later.

How early or late the bit should be written is a function of its position in the data pattern and its track position, among other factors. The shift at middle and outer tracks can range from 50 to 350 ns. For optimum precompensation in all cases, 150 to 175 ns of precompensation is recommended for both early and late bits. Timing circuitry that achieves this precompensation includes counters, shift registers, delay lines, and 1-shot delay circuits.

For example, a 1-shot precompensation circuit (Fig 6) is connected to the controller chip early- and late-bit signals. Based on precompensation of 175 ns, a 300-ns delay is used for an on-time bit, a 125-ns 1-shot delay is used for a bit to be written early, and a 475-ns 1-shot is used for a bit to be written late. The trailing edge of

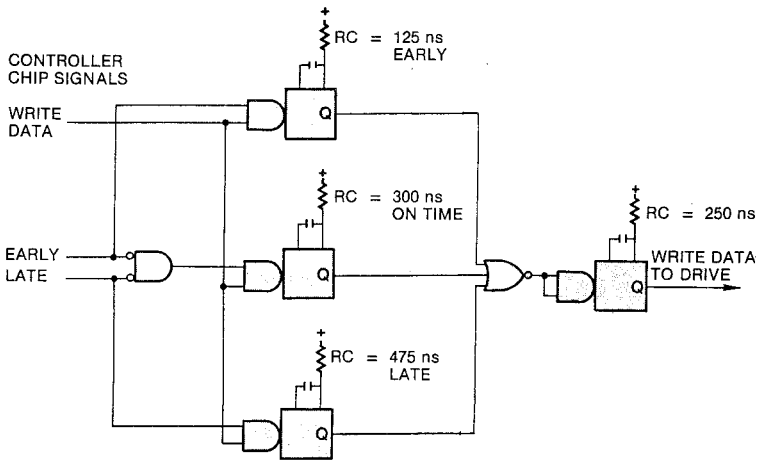



Fig 6 Precompensation circuit. To compensate for predictable bit shift, circuit writes early bits that would normally be shifted late during reading. Conversely, bit expected to appear early is written late

each of these 1-shot pulses fires a fourth 1-shot to provide a standard width write pulse to the disc drive.

Postcompensation alters the read signal rather than the write signal. Since bit shift is, in part, a function of read channel frequency response (phase characteristics), postcompensation circuitry changes the read signal after track 60 to compensate for bit shift. This IBM method is relatively new; floppy discs written on System/34 are incompatible with disc drives that do not use postcompensation.

## Data Separation

Substantial reductions in the amount of bit shift can be achieved using either a pre- or post-compensation method. However, unpredictable bit shift can still occur. Therefore, special design consideration must be given to the type and resolution of data separator used in reading data bits from the disc.

The bit stream transferred from the disc to the controller consists of composite clock and data bits. With single density, a data bit is decoded by a data window that is generated from the clock bit. In double density, the lack of consistent clock bits makes it impossible to generate a data window in this manner. Instead, the separator circuit must first determine the nominal position of clock and data bits and then generate a 1-$\mu$s clock and data window that is centered around the bit positions. The more accurately the bit position can be determined and the tighter the resolution of the data window, the lower the soft error rate of the disc.

The 1-$\mu$s data window must be centered on a bit that can potentially shift as much as $\pm$450 ns. This shift leaves less than a 50-ns margin to the edge of the window; a data bit appearing on the edge of the data window could be read as a clock bit. The total error from the data separator circuit must therefore be less than $\pm$50 ns.

To determine the nominal bit position around whic to center the window, the data separator must track dat bit frequency changes. In this manner, even if an u predictable bit shift occurs, the data separator can adju the window's position to compensate for the chang Otherwise, the shifted bit could be positioned outside th window. To remain within the error rate specified b drive manufacturers, not more than 1 in $10^9$ bits ca appear outside the window. With present technology only an analog data separator based on a phase-lock loo technique can provide the necessary reliability.

Digital data separators have lower resolution than a analog phase-lock loop type of separator, and canno accurately determine the nominal position of the dat bit around which to position the window. As a result error rates higher than those specified may result fron the use of digital data separators.

For clarity, the performance of a typical digital dat separator is examined. To generate a data window, crystal clock is divided down by a value preloaded i a counter to create a 50/50 data/clock window. Assuming that a 20-MHz clock is used, the smallest incremen to which the window can be adjusted is the least signifi cant bit (LSB) of the counter, which in this case is 5( ns. The separator determines the nominal bit position around which to center the data/clock window by sampling bit positions within the window. From sampling results, the data separator adjusts the window position using a feedback mechanism that changes the preload value in the counter. This effectively shifts the window position relative to the nominal bit position.

The 50-ns resolution, however, creates a problem. As shown in Fig 7, if a bit is shifted a maximum of 450 ns in one direction, the digital data separator will compensate by moving the window 1 LSB or 50 ns in that direction. If a subsequent bit is then shifted 450 ns in

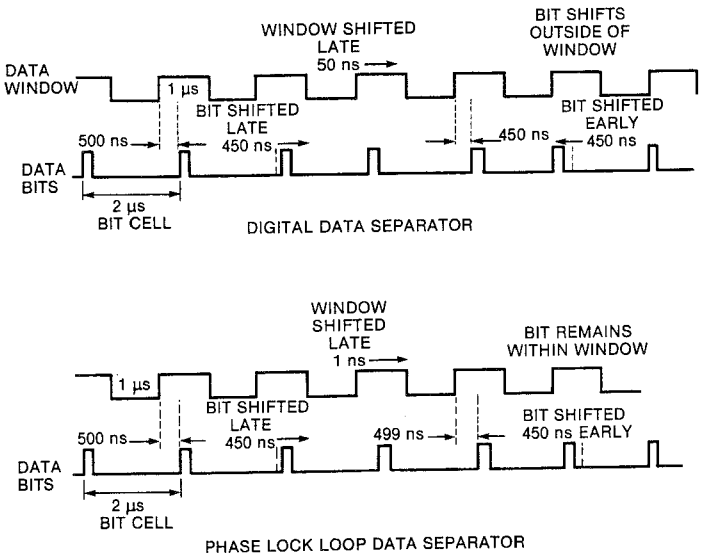



Fig 7 Data separator timing diagrams. Unpredictable bit shift is handled differently by digital and analog data separators. Typical digital data separator, using 20-MHz clock, provides 50-ns resolution. Assuming two worst case bit shifts in opposite directions, data bit would appear outside data window. Analog data separator, however, provides 1-ns resolution. As result, it can handle unpredictable bit shift more reliably than digital data separator

the opposite direction, that bit will appear 50 ns outside the window, resulting in a misread.

An analog phase-lock loop separator, on the other hand, has tighter resolution (±1 ns) and handles bit frequency changes more reliably (Fig 7). With this method, a phase-lock loop locks onto the basic frequency of data bits read off the disc, and determines nominal bit positions for data and clock bits by sampling every bit (clock and data). It uses the phase relationship between a bit and its window to vary the position of the window. By sampling each bit, the phase-lock loop determines the phase error between a bit and the frequency being generated.

Changes in the data window position depend both on an integration factor and on the amount of bit position error. For an integration factor of 100:1, the data window would move 1 ns per bit cell for each 100-ns change in bit position, until the entire error had been compensated.

Using the same example as described for the digital data separator, the phase-lock loop would detect the data bit shifted late by 100 ns. Then, the data window would be adjusted so that it appeared 1-ns late. If the next bit is shifted early, the analog separator would also detect that shift and position the next data window, not late, but early. In this manner, the phase-lock loop reliably tracks frequency changes. A block diagram of a typical phase-lock loop circuit is shown in Fig 8.

## Double-Density Formats

In double density, data are encoded differently from single density, but are formatted more or less identically on the disc. Each of the 77 data tracks on a standard 8" (20-cm) floppy disc is organized in data records, which are also referred to as sectors. Two methods of sectoring currently exist: hard and soft. In the more prevalent soft sectoring, the number of sectors and their length can vary. Optimum sector size for systems with small main memories and smaller data bases may be as low as 128 bytes. For those systems with larger main memories and large data bases, 256 or 512 bytes per sector may be optional.

Soft sectoring is supported by IBM in both double- and single-density controllers. The double-density format shown in Fig 9 is used in the IBM System/34 and will likely become the industry standard. It is also the format supported by available LSI disc controller chips. Similar in most respects to the standard single-density format, the double-density format has the same number of sectors (26), each with twice as many bytes (256).

Double-density format differs from single-density format in the way an address mark is detected. An address mark flags the beginning of every index, identification (ID), and data field. In both single- and double-density recording, the address mark is unique in terms of its clock pattern; the single-byte single-density address mark contains two or three missing clock pulses.

This method is modified for double density because dropping two or three clock bits in a row could cause the data separation circuit to lose synchronization. In double density, each of the first three bytes of a 4-byte address mark has one active data bit followed by four zero data bits. According to double-density encoding protocol, the three zero data bits should have an associated clock pulse. However, to make the address mark unique, the middle clock pulse is dropped. The respective address marks for index, identification, and data fields are made unique by their associated data pattern. For example, an index address mark has three C2 bytes (hexadecimal), followed by an FC byte; an ID address mark has three A1 bytes followed by an FE byte; and

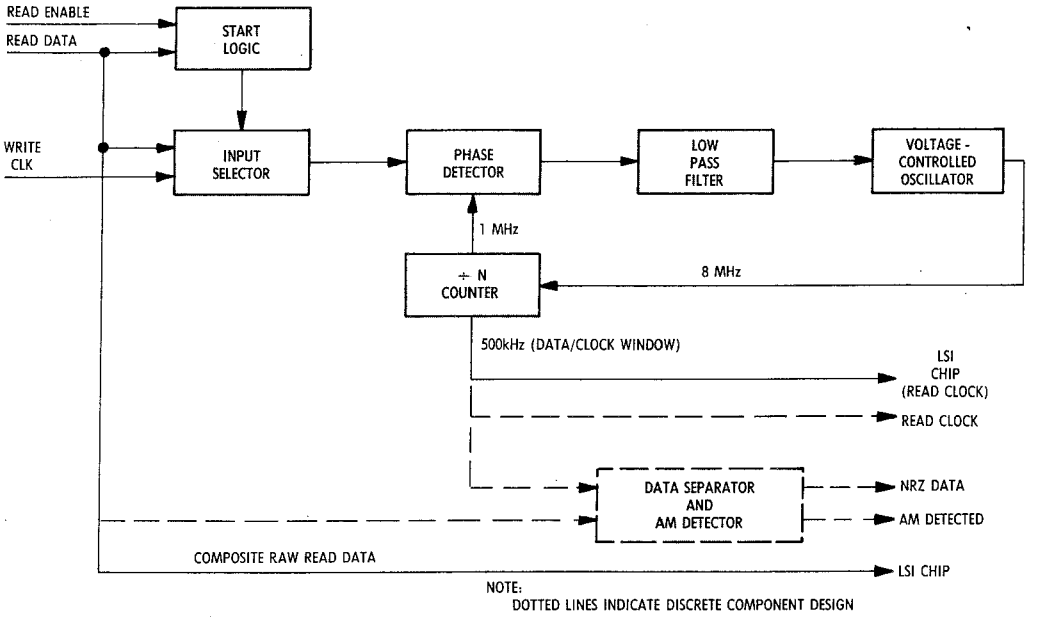



Fig 8 Analog phase-lock loop data separator. Input selector selects write clock signal as input to phase detector until start logic has detected synchronized area preceding address mark of formatted disc. At that point, input selector switches from write clock to read data as input to phase detector. Phase detector then begins sampling difference between read data and output of voltage controlled oscillator (VCO). A ÷ N counter continually divides VCO frequency down to same frequency as read data. Finally output of ÷ N counter becomes data/clock window

data has three A1 bytes followed by an FB byte. This bit pattern is automatically generated by the controller chip, as is the CRC character at the completion of the data and ID field.

Obviously, LSI controller chips simplify disc formatting, but offer some limitations on format flexibility. If a designer chooses to configure a custom format, certain design guidelines for gap length must be observed if data are to be recorded reliably.

Gap 1 of 22 bytes must be present if an index address mark is used to separate the index address mark from the first identification field. Gap 2 of 22 bytes, which separates the identification and data fields, is necessary to protect the ID field from erasure during a write. Gap 3 of 54 bytes is a speed tolerance gap, which again protects the ID field if the medium is interchanged between drives with different rotational speeds. Gap 4 is also a speed tolerance buffer. This gap length is determined by the difference between the format length and the actual track capacity, which may vary from drive to drive.

A double-density format (shown in Fig 9) is, with some exceptions, the same as a single-density format with gaps and data fields that are twice as long.

## Summary

Improvements in head and media resolution have made double-density encoding a reliable method of doubling the capacity of a floppy disc from 410k to 820k bytes. A minifloppy drive, using double density, can increase its capacity from 110k to 220k bytes. Research analysis into the problems of bit shift has enabled designers to reliably compensate for bit-shift effects so that a minimum soft error rate in excess of 1 in $10^9$ is possible with double-density encoding. LSI controller chips have made double-density system design less complex, less time consuming, and therefore less costly. In addition, double-density floppy discs with doubled capacity can be implemented on single-density drives. These capabilities

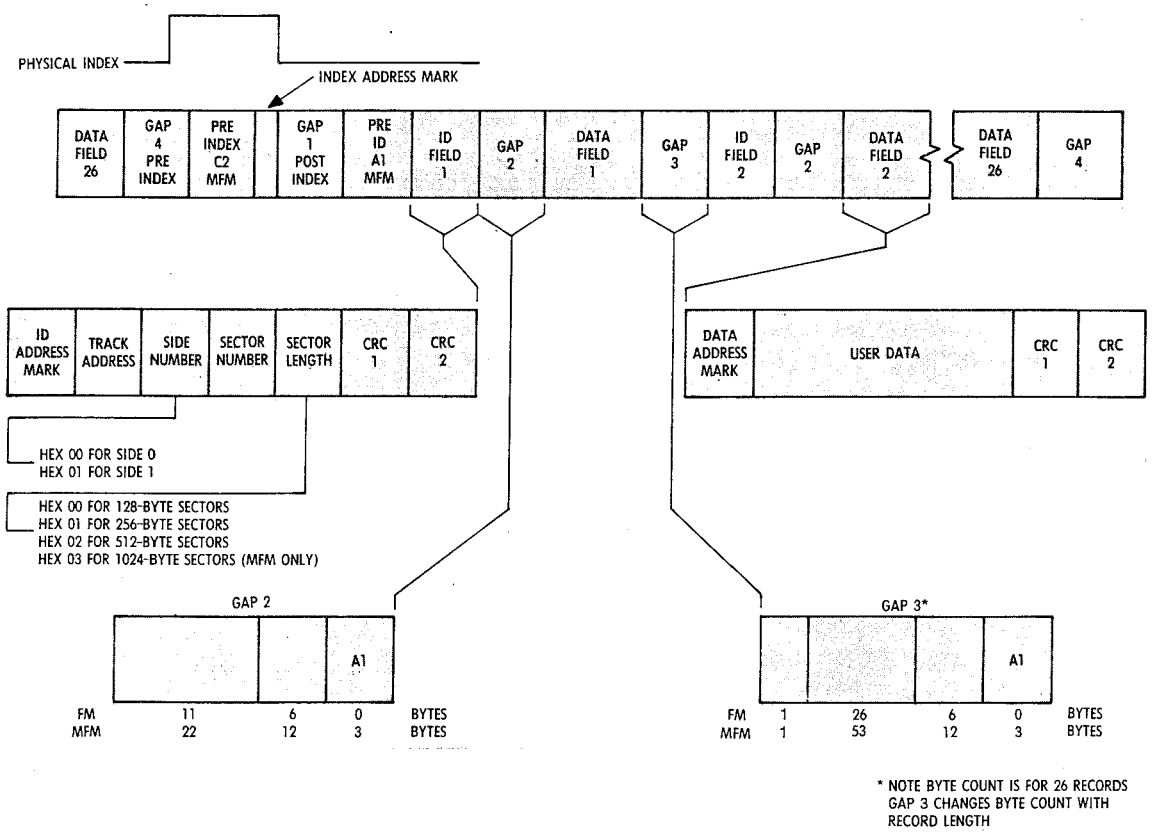



Fig 9 Double-density format. Defacto industry standard MFM double-density format. Within each sector, 4-byte address mark flags beginning of every index, identification, and data field. Address marks are distinguished from data by missing middle clock pulse in each of first three bytes. Index, ID, and data field address marks each have unique data clock pattern

should broaden the applications of floppy disc drives and make systems with multiple floppy disc drives more compact.

## Bibliography


R. C. Franchini and D. L. Wartner, "A Method of High Density Recording on Flexible Magnetic Discs," *Computer Design*, Oct 1976, pp 106-109

D. J. Kalstrom, "Simple Encoding Schemes Double Capacity of a Flexible Disc," *Computer Design*, Sept 1976, pp 98-102

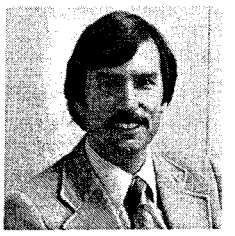
NEC Microcomputers, "uPD765 Single/Double Density Floppy Disk Controller Data Sheet," NEC Microcomputers, Inc, Wellesley, Mass, 1978

Shugart Assoc, *Double Density Design Guide*, Shugart Associates, Inc, Sunnyvale, Calif, Oct 1977

P. S. Sidhu, "Group-Coded Recording Reliably Doubles Diskette Capacity," *Computer Design*, Dec 1976, pp 84-88

Western Digital, "FD1791A/B Floppy Disk Formatter/Controller Data Sheet," Western Digital Corp, Newport Beach, Calif, July 1978

J. Worden, "Design Considerations for Dual-Density Diskette Controllers," *Computer Design*, June 1978, pp 103-110


*John F. Hoeppner as applications engineer for Shugart is responsible for the company's line of floppy and fixed disc products. His experience includes work in the areas of controller development, magnetic tape and disc recording, and mini and microcomputer software development.*



*Larry H. Wall is manager of product engineering at Shugart, and is responsible for the electrical components within floppy disc drives. His experience includes the design of an IBM compatible VFO. He holds a BSEE degree from California Polytechnic State University and is currently working on a masters degree in engineering management at Santa Clara University.*