Radial tracking system of the Philips Compact Disc Audio System.

# COMPACT DISC DIGITAL AUDIO SYSTEMS

## *How the new digital audio 4.7-inch playback record system works*

### *By David Ranada*

COMPACT Disc players—the hottest new products to appear on the audio scene in quite some time—are probably the most complex and sophisticated electronic devices for the home ever offered to the consumer. They are designed to play back, with extraordinarily high audio fidelity, up to 75 minutes of music recorded on mirror-finish 12-centimeter (4.7″) acrylic discs. The system does not wear out or damage the discs while it is playing them since only a light beam contacts the disc to read it out; that light comes from a semiconductor laser.

CD players combine aspects of a home computer's disc drives (in the player's disc-drive and laser-tracking servos), a specially programmed microprocessor (to handle a player's control functions), a digital signal processor (to decode and correct the digital-audio bit stream), and a high-quality, high-precision digital-to-analog conversion system (to change the digital audio data into normal audio signals of high sonic quality). How do all these systems interact to make a CD player work?

**Digital Audio Basics.** A digital-audio recording/playback system converts a smoothly changing, analog waveform into a series of binary numbers describing the waveform, records those numbers as pulses on a suitable medium (magnetic tape or the master disc for Compact Disc pressings), plays back those pulses, converts them into numbers, and finally reconverts the numbers into a smoothly varying analog waveform (Fig. 1). How can such a system provide the high audio quality claimed by advocates of digital audio? After all, typical Compact Disc player specifications claim a dynamic range of more than 90 dB, flat (±0.2 dB) frequency response to 20,000 Hz, distortion of less than 0.01%, and *no* wow or flutter.

The answer lies in the way the two basic operations of digital audio are performed. The first operation is called sampling. If the amplitude of a waveform—any waveform—is somehow captured or frozen at regular, closely spaced intervals, the waveform is said to be sampled. The waveform in Fig. 1A gets sampled in Fig. 1B, the samples being the points at the ends of the vertical lines. Mathematical models of this activity say that as long as samples are made at least twice as frequently as the highest frequency contained in the signal being sampled, *no information is lost.* From the information contained in the samples the waveform can be exactly reconstructed, as long as the waveform did not contain any frequencies greater than the *sampling rate.* The Compact Disc system's designers wanted it to have a frequency response extending to 20,000 Hz, so the audio signal for Compact Disc master tapes is sampled 44,100 times per second. The sampling rate is 44.1 kHz, more than twice the highest frequency desired by a slight margin.

Quantization—the second basic operation of digital audio—is shown in Fig. 1B and 1C. It is simply the *measurement* of the amplitude of the samples, the results of which are shown here in arbitrary numerical units in both decimal and binary representation. Binary representation is used by analog-to-digital converter (ADC) circuits in digital-audio equipment.

Note that the measurements are not infinitely precise; there is some "rounding off" to the closest integer value. Making an infinitely precise measurement would generate a number with an infinite number of digits (or bits); there has to be a practical limit imposed to the resolution of the measurement. This limitation creates noise and distortion in the reproduced signal, but the noise and distortion can be reduced to any level by increasing resolution of the ADCs.

In the Compact Disc system, audio is encoded with 16 bits of resolution, meaning that the ADCs used in Compact Disc mastering can distinguish changes in voltage levels of 1 in $2^{16}$. This translates to 1 in 65,536 or the equivalent of 1″ in about 1 mile. With this degree of resolution, the signal-to-noise ratio or dynamic range of the CD system can be greater than 90 dB, the theoretical limit being a little better than 97 dB. Since quantization error is also responsible for nearly all the distortion in the digital audio process, it too is reduced to very low levels with 16-bit encoding.

The digital numbers from the ADC are converted into a pulse train (Fig. 1D) for recording on tape or on a Compact Disc master. Since it has only two possible values (high and low, corresponding to 1 and 0), this waveform is far more immune to distortion and noise in the recording and playback process than the original analog waveform. Distortions that can and do occur in the data-storage media (tape or optical disc) only rarely change the recorded pulse train enough to confuse the decoding circuitry as to whether a high or low value was intended.

In playback—this is where a CD player comes in—the original pulse waveform is recovered from the recorded waveform (Fig. 1E). After it is converted into the original numerical sequence (Fig. 1F) the series of digits is fed into a digital-to-analog converter (DAC) which changes them into discrete voltage levels (Fig. 1G). Because the numbers are clocked into the DAC at the same rate as they were generated at the ADC during the recording process, variations in the speed of the playback medium do not affect the reproduced waveform. Therefore, wow and flutter are eliminated!

An output-smoothing circuit removes the inaudible ultrasonic components that give the staircase waveform it's disjointed character. The filtered waveform makes up the output of the digital audio system, an output that is a very high-fidelity replica of the original input signal.

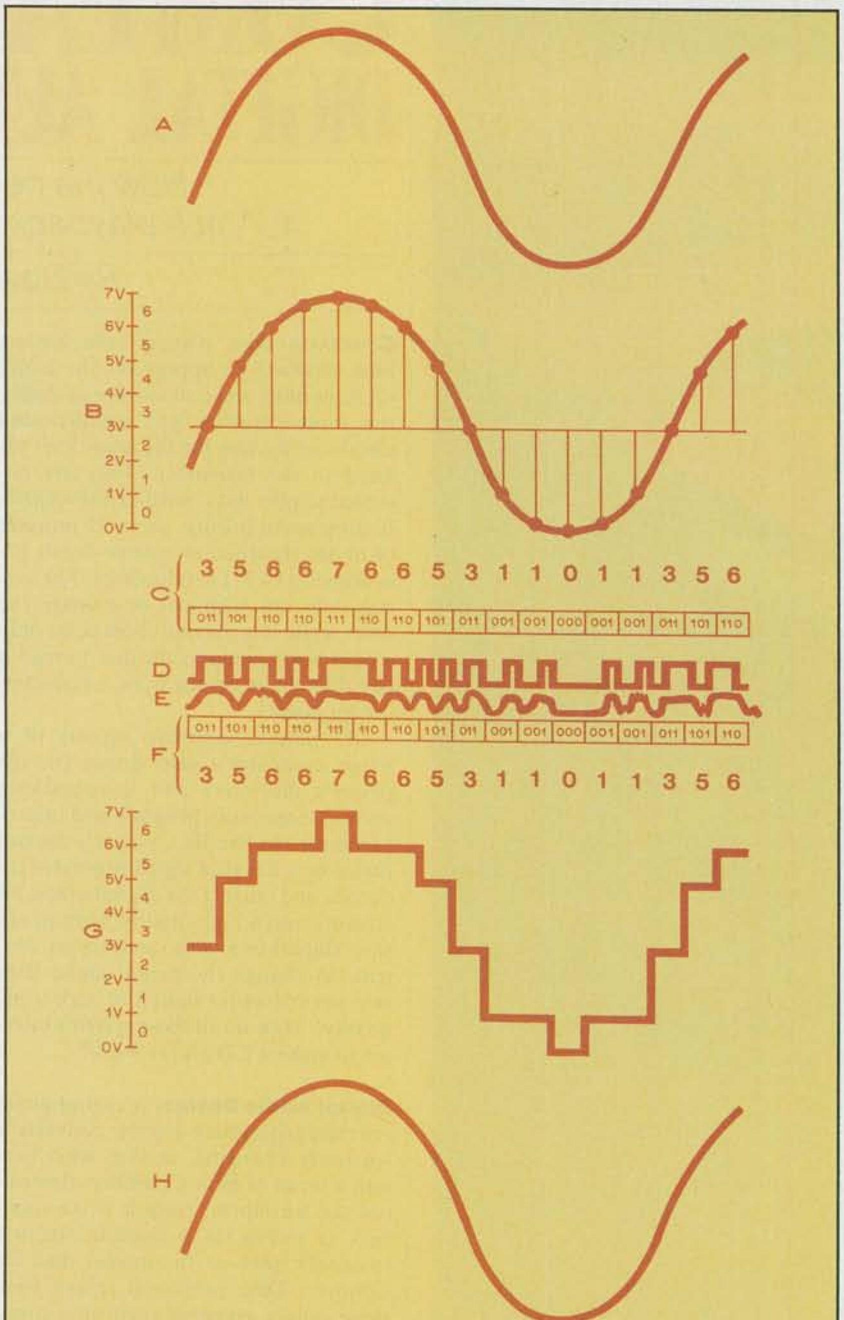


*Fig. 1. Basic operations of digital audio: sampling (B); quatization (C); recover of data (E,F); and reconversion of data into an audio signal (GH).*

**The Compact Disc and Player.** Before exploring some of the processing that goes on inside a Compact Disc player let's take a look at what a CD player has to play, the Compact Disc itself. Audio information is recorded on a CD pressing in a track of microscopic "pits" on the surface of a transparent acrylic disc 12 centimeters in diameter, as shown in an accompanying photo. These pits average slightly less than 1 micrometer in length; each revolution of the track is 1.6 micrometers away from the next. Data is contained in the *length* and *spacing* of the pits and the spaces ("lands") between them.

A cross-section of a CD pressing is shown in Fig. 2A. In it you can see that, though the pits (P) are made as indentations, they appear as "bumps" to the scanning laser beam. It comes up through the bottom of the disc, hits the aluminized signal surface (AL), and is reflected out again. Changes in the intensity of reflected light are detected (Fig. 2B) and converted to a digital bit stream (Fig. 2C). The pits are protected on one side by the bulk of the disc's plastic substrate (S), and on the other by a tough lacquer coating and a label (PR).

A typical mechanical arrangement of a Compact Disc player is shown in Fig. 3 and the illustration on page 40. It consists of an optical scanner, containing a semiconducter laser and a photodiode, mounted on a tracking device that allows the scanner to move across the face of the spinning disc. The disc spins counterclockwise in relation to the laser beam and the track of pits starts toward the center of the disc and moves toward the outside edge, both movements being opposite to those with analog (black-disc) records. The speed of rotation *slows down* between the start of a disc and its finish.

It starts out spinning at around 500 rpm and ends up spinning at about 200 rpm if the program extends close to the edge of the record. The speed is controlled by servo networks so that the track is always passing over the laser at between 1.2 and 1.4 meters per second. The steady decrease in rotation speed increases the amount of information that can be held on the disc. If the rotation rate were constant, the pits and alternating lands near the start of the disc would be unnecessarily large.

**Data on Disc.** It would be a stupendous and misleading oversimplification to say that digital-audio data gets recorded on a Compact Disc exactly as it emerged from the digital-audio master recorder. There is a good deal of digital signal manipulation on the audio data before it finds itself on the signal-surface of a Compact Disc. Most of this processing is undone by a CD player in its operations to recover the original digital-audio. Then why is all the processing necessary? The three main reasons are: to increase the playing time of a CD, to make it audibly immune to disc damage (scratches, fingerprints, small pressing defects), and to provide special digital control and data signals for advanced Compact Disc applications.

Figure 4, as complicated as it appears at first glance, summarizes all the main digital operations performed to a digital-audio signal on its way to the master Compact Disc. The digital data on a CD is gathered into units called frames and the diagram depicts the encoding process for one frame (which takes only about 136 microseconds to be scanned by a CD player's laser).

The digital-audio signal content of a CD frame consists of enough information to make up six sampling periods (Fig. 4A). Since there are 32 bits generated in one sampling period (16 bits per stereo audio channel every 22.68 microseconds) there are 32 × 6 = 192 audio-data bits per frame (Fig. 4B). Each 16-bit sample is broken down the middle into two 8-bit bytes, called *symbols* in CD terminology (Fig. 4C). One symbol contains the 8 most significant bits of the sample, the other contains the 8 least significant bits. This repartitioning of the data yields 24 symbols containing only digital-audio data per frame.

The first major signal-processing operation on the audio data during the encoding process is depicted in Fig. 4D and E. These two lines depict the functions that make the data stream immune to transmission errors due to disc defects or damage.

They start with the generation of redundant data that a CD player will use to detect and correct for any errors in the data stream as read by its laser (see box on page 45). Simply put, the redundant data is a mathematical summary of the digital-audio data contained in the frame and is a special form of parity-check code (for those of you familiar with digital data transmission). The first error-correction symbols to be generated are the four Q symbols which find their places in the middle of the frame.

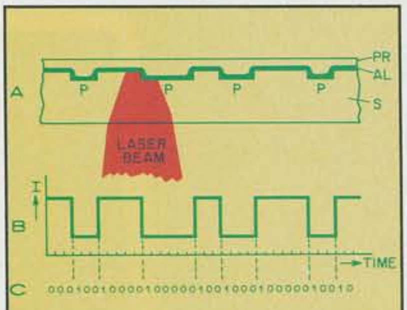


*Fig. 2. A 1.2-mm thick Compact Disc (A) consists of three layers: an acrylic substrate (S), an aluminized reflecting "signal surface" (AL), and a lacquer coating and label (PR). The pits (P) are pressed from above but the laser beam scans from below. The reflected intensity of the beam varies as the pits pass by (B) and these variations hold the digital-audio data (C).*

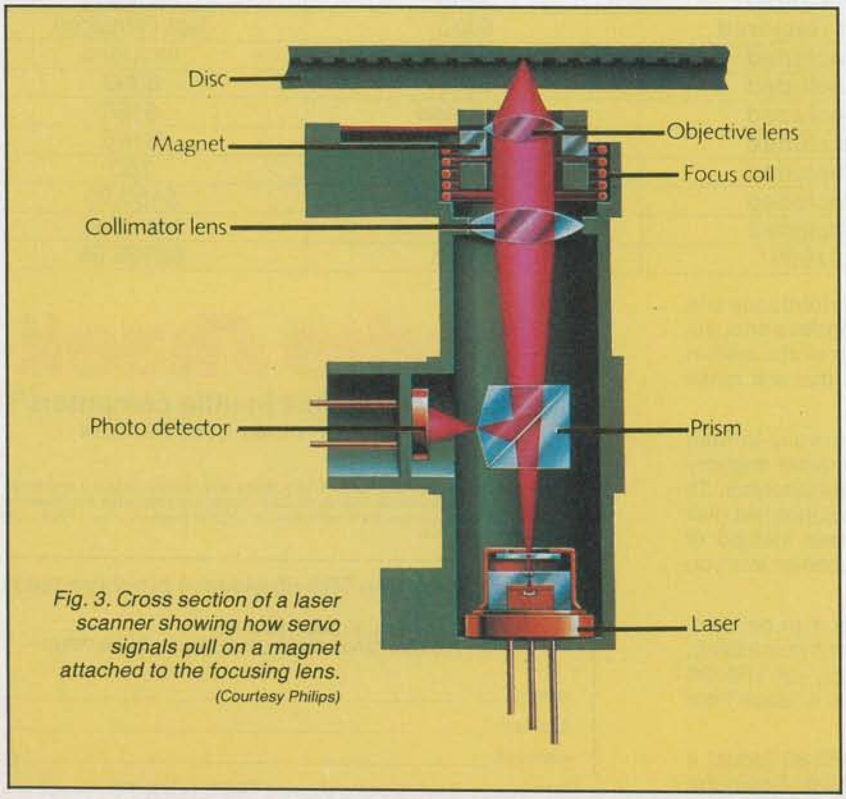


*Fig. 3. Cross section of a laser scanner showing how servo signals pull on a magnet attached to the focusing lens.*

*(Courtesy Philips)*

Next comes interleaving, a process for making the recorded signal immune to disc scratches and other relatively large-scale defects (see accompanying box). Interleaving scrambles the symbols from a large number of frames preceding and following the one being encoded in the diagram. This scrambling (depicted by the arrows between Fig. 4D and E) is performed in a very strictly defined way, which moves originally consecutive symbols to distantly separated frames, and which moves symbols from distant frames into the one under construction. The Q-parity symbols also undergo this shuffling process. No data is left out or repeated—the number of digital-audio symbols per frame is still 24, the number of Q-parity symbols is still 4—its just that 28 previously consecutive symbols find themselves strewn over about 40 different frames in a very regular, and undoable pattern.

After interleaving, four P-parity error-detection/correction symbols are calculated from the 24 audio and 4 Q-parity symbols now in the frame. This additional parity data is tacked onto the end of the frame. The whole parity/interleaving scheme is called CIRC (Cross-Interleave Reed-Solomon Code).

To the start of the frame is added one 8-bit symbol containing control and display (C&D) bits. At present this symbol is used only to hold data about the timings of each musical selection on a Compact Disc, and data signalling where each selection begins and ends. This information is used by a CD player to cue up individual selections.

Future applications of the C&D symbol include text data for lyrics; graphics data for computer-type displays on TV's and monitors; and ultimately video data in conjunction with a video frame-store system. The C&D symbol can also be made to carry a telephone-quality digitally encoded voice signal for sing- or play-along recordings. (Since the C&D symbols contain 8 bits and occur 7,350 times per second–once per frame—a voice signal with a signal-to-noise ratio of about 45 dB and a frequency range up to about 3,500 Hz can be carried.)

Players for all these advanced functions have not yet been issued. In fact, some or the most technically complex options, like computer graphics and video, are still in the planning stages.

**Modulations.** At this stage (Fig. 4E), the frame consists of 33 8-bit symbols of digital-audio (24 symbols), parity-check (8 symbols), and control/display data (1 symbol). This symbol sequence is then fed into a digital modulation circuit which converts, via a read-only-memory look-up table, each 8-bit symbol into a 14-bit word in a process called EFM (Eight-to-Fourteen Modulation). Essentially the 8-bit symbols modulate a bit pattern 14 bits wide (Fig. 4F and G). Why is this done?

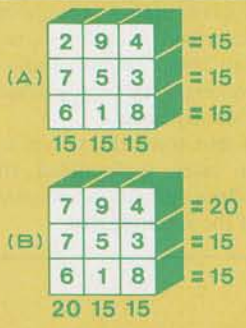
The full answer to that question is rather complex, but it boils down to the fact that about 25 per cent *more* information can be stored on a Compact Disc if the 8-bit symbols are converted to 14-bit words with certain special properties. To the CD user this increase of "information density" means longer

## ERROR CORRECTION

Analog stylus-in-a-groove recordings are very susceptible to all sorts of sonic degradations from defects in the manufacturing process or damage inflicted by the user. What record buyer hasn't had the unpleasant experience of a bad pressing or hasn't heard the noise added by fingerprints or scratches on an analog disc's surface. Digital audio offers salvation from these problems since specially encoded digital signals can be made immune to what are called "transmission errors." By the incorporation in the pit/land track on a CD of extra information which is mathematically redundant to the original digital-audio signal, a Compact Disc player can eliminate to a substantial degree the sonic side effects of even gross disc damage like scratches and fingerprints.

The theory of error-correction coding can get extremely involved, mathematically speaking, but the basic principles of how error correction is applied in the CD system can be illustrated by a simple analogy. Figure A illustrates what is called a magic square. As you might recall from grade-school arithmetic, a magic square is an array of numbers, the rows, columns, and diagonals of which all add up to the same number. In this case each row, column, or diagonal adds up to 15. If you were to transmit the numbers contained in the magic square along with the sum of any row or column, and if one of the numbers were to be incorrectly received, you would still be able to regenerate the magic square completely and exactly. Why? Because you know that each row or column or diagonal adds up to 15. If an incorrect number is received as in Fig. B, the rows, columns, and diagonals will not add up correctly. The offending number can be easily located and corrected. The location is at the intersection of the row and column which do not add up correctly; the correct value is that number which makes the sum come out correctly.

By transmitting the sum value along with the numbers in the magic square you were transmitting redundant information, a mathematical summary of the numbers contained in the square. The receiver is able, using this redundant information, to calculate and correct more than one transmission error even if one of the errors occurs with the sum value. Of course, correcting the sum value isn't helpful if all you are interested in is the magic square's numbers, but at least that this example shows that the transmission of redundant information makes the detection and correction of transmission errors much easier.



*Figs. A and B. Magic squares with and without the correct numbers.*

The Compact Disc system makes use of a similar, but much more complex, process to protect the digital-audio data from transmission errors. About 25 per cent of all the data encoded on a Compact Disc is mathematically redundant to the digital-audio information. These error-correction symbols are derived by binary computation in much the same way the sums for the magic square's rows and columns were obtained. Naturally the numerical data on a CD is not placed in magic squares—digitally encoded music doesn't create sequences of numbers that are easily fitted into magic squares. Instead, the CD encoding system creates unique patterns of bits which are hard to confuse with one another without being able to find and correct for errors.

No error correction system is perfect, however. Eventually enough errors can occur that exact reconstruction of the transmitted signal is impossible. In the magic square example, for instance, if four or more numbers were altered in transmission there would be no unique solution to the square. There may be several different arrays of numbers which will make up a valid magic square. Likewise with the Compact Disc system. Eventually there may be so many errors in the data stream, due to extensive disc damage, that exact correction of the digital-audio data cannot be performed. A CD player will then interpolate the correct data from valid data by taking an average of the number before the erroneous digits and the number after them.

In the CD system, if one out of every thousand data bits were incorrect (giving a bit error rate or BER of $10^{-3}$), interpolation will be performed about 1,000 times per minute. A BER of $10^{-3}$ corresponds to a very well scratched disc. If only 1 out of every 10,000 bits is in error (BER $= 10^{-4}$), a typical value for an undamaged CD pressing), the interpolation rate drops to about one 16-bit sample in every 10 hours of playing time.

There is a very small possibility that an undetected error will sneak through the CD error-correction system. With the magic square this would mean that several numbers were changed in such a way that the square is still "magic" but not the same one as transmitted. In digital audio, an undetected error can lead to a "click" in the reproduced sound. The CD error correction system will produce fewer than one click for every 750 hours of playing time with a disc having a BER of $10^{-3}$. In practical terms this means that a reasonably well cared for Compact Disc pressing will never generate clicks or pops in playback. The listener is spared from these by the actions of the Compact Disc error-correction system on the data stream. ◇

playing times. Accompanying this increase of playing time, EFM decreases the CD system's sensitivity to disc and optical-system defects and production tolerance variations.

The signal created in a CD player's photodiode detector circuitry by the EFM sequence of pits and lands has two more important properties: it has low dc (0-Hz) content and it is "self-clocking". Low dc content is important because the servo systems which keep a CD player's laser "on track" utilize low-frequency control and feedback signals. Any dc content in the signal from the photodiode detector can throw off the tracking of the laser. (Think of a very long pit or land as having substantially higher dc content than rapidly alternating, small pits and lands.) EFM has very little 0-Hz in its waveforms since the spacing and lengths of pits and lands is strictly controlled.

A CD player must synchronize its operation with the data coming from the disc and must also synchronize the rotation rate of the disc with the operation of the CD-decoding circuits. EFM has self-clocking properties allowing a CD player to determine which bit belongs to which symbol during decoding.

While the 8-bit symbols are converted into 14-bit EFM words, three additional bits are added to each word. These "merging bits" are calculated to further reduce the dc content of the signal. Finally, the end of the frame is marked by a sync signal of 24-bits length together with three merging bits. In all, a frame which started out as 192 audio bits now contains 588 channel bits. These are recorded onto a Compact Disc so that a 1 in the channel-bit sequence turns into the *edge* of a pit. A CD player's decoder circuits really don't care if they are scanning pits or lands; the circuits are sensitive to pit edges. It makes no difference whether a logic "1" is encoded as a pit or a land.

**Inside a CD Player.** A Compact Disc player has to undo all the encoding steps just described in order to recover the original 16-bit digital audio samples. The process, very simply described, follows this sequence:

1. Light from the player's semiconductor laser is focussed through a prism/lens optical system onto the surface of the spinning disc.

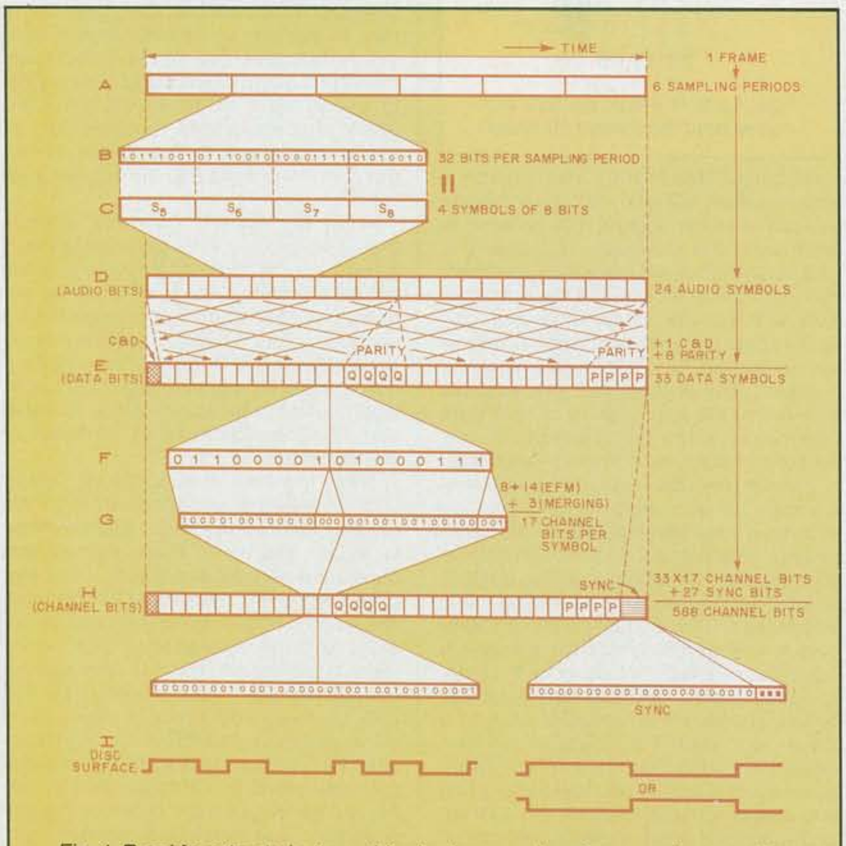


*Fig. 4. Read from top to bottom, this chart summarizes the operations applied to a digital audio signal before it gets on a Compact Disc. Read from bottom up, it summarizes the CD player recovery process.*

2. The reflected laser beam returns through the same optical system but is diverted to a light-sensitive detector (a photodiode array).

3. The photodiode output, after amplification and filtering, is converted into a squared-off pulse train by the use of comparator circuitry.

4. The pulse train of zeros and ones (corresponding to Fig. 4H) is demodulated according to the rules of EFM into 8-bit symbols carrying audio data, parity-check information, and C&D data.

5. Data control/error-detection-correction/memory circuits de-interleave and correct the data using the CIRC decoding rules.

6. Sixteen-bit digital-audio samples are assembled from the error-corrected symbols and sent to the 16-bit digital-to-analog converters where they are converted into analog voltages.

7. The staircase-shaped DAC outputs are smoothed by sharp-cutoff lowpass filters and the resulting *analog* signals make up a CD player's audio output.

This brief summary pays short shrift to the complexity and sophistication of the operations and circuitry involved. Semiconductor lasers are not universally used components, let alone in mass-market consumer products like CD players. Steps 4 and 5 make use of complex LSI chips, without which the building of a practical and affordable CD player would be impossible. For example, the three Sony-developed chips performing steps 4 and 5 and used by them and several other companies in their first CD players contain the equivalent of about 7,500 gates, not including a 16 kilobit RAM necessary for the processing. Sixteen-bit DACs weren't even available in low-cost monolithic form five years ago.

**Tracking.** That short list of CD-player functions also glosses over two of the most fascinating applications of servo-loop techniques: the turning of the disc and the tracking of the track of pits by the laser/optical system. These two operations are controlled by digital/analog/electromechanical servo loops of fascinating complexity. They are so intricate, in fact, that only a very broad description of them can be given here.

As mentioned before, the rotation rate of the disc is locked to the frame rate, which itself is determined by characteristics of the disc's data stream. Phase-lock techniques are used to keep the rate of disc rotation generally synchronized with the frame rate. Exact synchrony is not necessary because of the self-clocking nature of the data modulation scheme; EFM demodulation chips can determine where each bit belongs from the signal itself.

Servo-loop feedback techniques also keep the laser trained and focussed on the pit track. The precision of the servos systems that accomplish this is very high. Taking tracking as an example, the laser system must be guided over a spiral of pits with a pitch of only 1.6 micrometers. Disc eccentricity and player tolerance may cause the track to swing from side to side; the maximum permitted swing is 300 micrometers. Yet for accurate decoding, the tracking servo must allow the laser beam to follow the pits to with 0.01 micrometers and also must absorb the effects of vibrations and jolts applied to the player chassis.

When it comes to focussing the laser beam onto a CD's inner signal surface, the focus servo must be able to cope with disc warps of up to 1 millimeter. The focussing system, however, must maintain focus on the signal surface to tolerances 1,000 times smaller: ±1 micrometer.

A CD player performs such feats of electromechanical precision by monitoring the shape and/or intensity of the laser beam after it is reflected off the signal surface and back into the optical system. Deviations in the shape or intensity of the beam as detected by a photodiode array are used to move the optical system's objective lens nearer or farther from the disc surface to cope with focussing errors), or to move the optical system as a whole across the surface of the disc (to correct for tracking error).

**What's Next.** Even though the Philips/Sony Compact Disc is a tremendous engineering achievement, it is only part of the start of what will probably be a revolution in information-storage technology: the optical age. In its present form, a Compact Disc can hold about 16 *gigabits* of raw data. If the EFM modulation system is used to store computer data, a CD will hold about 9.14 gigabits or 1.14 gigabytes. If half of that capacity is used as redundant data for error detection and correction, one Compact Disc could hold about 570 megabytes of digital data, about 1,000 times the amount of a high-density floppy disc. (The error-correction system for the audio CD, as good as it is, still lets through too many interpolated and undetected errors to be tolerated in data-only applications, thus the very conservative assumption of 50 percent data redundancy.)

Since its small size increases production yields and lowers production costs over its larger (12") LaserVision videodisc counterpart, don't be surprised if the high information capacity of the CD system becomes a standard medium for mass distribution of software and data bases. The enormously larger data-carrying capacity of the 12" LaserVision system can even make it impractical for mass distribution purposes; who needs all that data anyway?

Software distribution won't be the only data application of the Compact

## INTERLEAVING

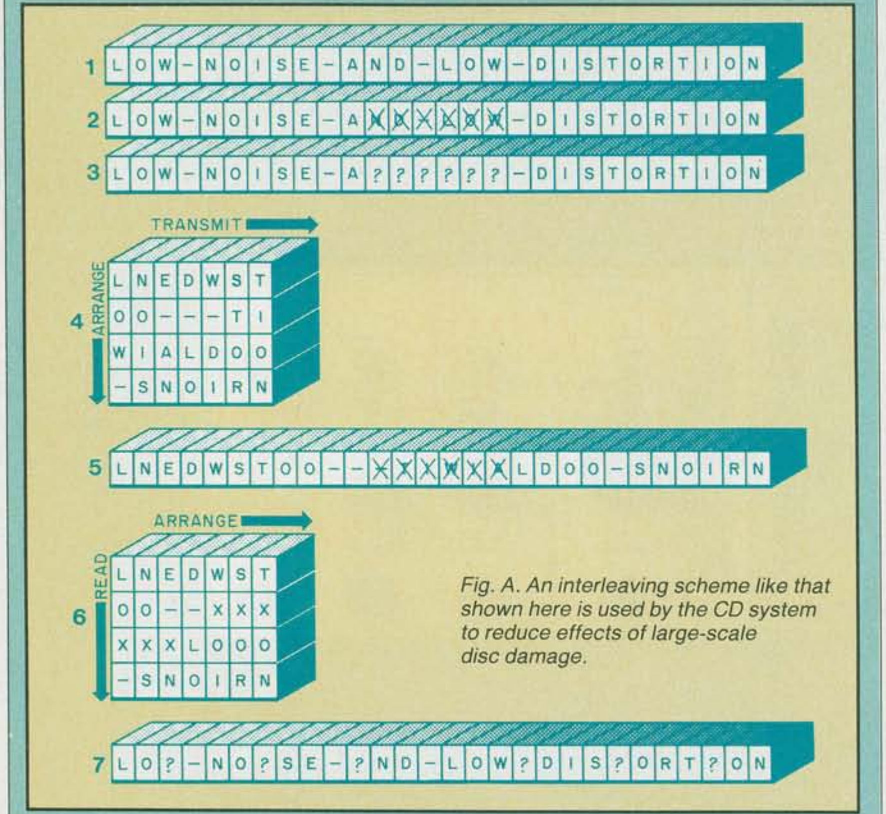


*Fig. A. An interleaving scheme like that shown here is used by the CD system to reduce effects of large-scale disc damage.*

Interleaving is a process used in the encoding of Compact Disc data that reduces the system's susceptibility to what are called "burst errors." These are errors in the data stream which occur to many bits in succession. They can result from interference to the scanning laser caused by relatively large disc damage (fingerprints, long scratches).

A simple example can easily demonstrate the effectiveness of interleaving (Fig. A). Suppose you transmitted the message LOW-DISTORTION-AND-LOW-NOISE, but a burst error of six characters occured during reception. The receiver would then have a hard time figuring out what was meant by LOW-NOISE-A?????-DISTORTION (Figure A1). You could have made the transmission more immune to the loss of several consecutive characters by first arranging them in columns, then transmitting them by rows (Figure A4). The transmitted signal is thus a regularly scrambled—interleaved—version of the original signal. If six successive characters are now obliterated, the receiver can deinterleave the message by rearranging it in rows and reading it out by columns (Fig. A5 and 6). The result is a much more intelligible signal; the burst error has been dispersed into separated "random" errors (Fig. A7).

Just as these random errors are much more easily handled by readers of English, they are more easily handled by a CD player's error-correction circuitry. In fact, readers of English supply a "context" to the received message to make it intelligible. Generally speaking, a CD player does the same thing with the redundant error-correction data placed on the disc. The redundant data provides a mathematically defined and standardized "context" in which to interpret random errors in the data stream.

In the CD encoding process interleaving is accomplished by writing the audio data into a digital memory and then reading it out in a different order—the equivalent to the row/column changeover in the example. A CD player performs the opposite operation in the de-interleaving process. ◇

Disc system. Efforts are being made at a number of companies to come up with a practical, inexpensive method of *recording* on Compact Discs and optical videodiscs. Not only will home digital-audio disc recorders become possible with the development of such technology but home computers could easily use a recordable CD for storage of large amounts of information. Even an unerasable CD recording system could be useful.

At present, research efforts are being directed towards a suitable laser/recording substrate combination. The semiconductor lasers used in audio CD players are too weak to directly "write" on Compact Discs.

All that relates to developments five to ten years away from reaching commercial reality. Much closer to appearing on your electronics dealer's shelves is a digital signal processor for audio signals. Already, relatively slow microcomputers in typical home computers are being put to good use in audio via specialized Fast-Fourier-Transform measurement system (like the IQS system Julian Hirsch uses in testing audio equipment). Now under development are ultra-fast, special-purpose arithmetic processors which, under control of a microprocessor, can be made to perform such audio tricks as filtering, reverberation simulation, compression, expansion, click and pop removal, and sound synthesis and mixing *while an audio signal is still in digital form.* Among the many advantages that all-digital signal processing promises are no audible degradation in signal to noise ratio or distortion levels, the capability of extraordinarily complex realtime operations on an audio signal, and, last but certainly not least, software based processing.

Imagine an audio component looking much like a popular video-game console. Just plug in a cartridge and the system becomes a digital ambience synthesizer, or a digital multi-band equalizer, or a tick/pop remover for playing these old-fashioned stylus/groove recordings. Technologically such a system is possible today; a commercial system may see the light of day as soon as two years from now. By then CD players will be supplied with direct digital-audio data outputs so a direct link from recording studio to a home stereo system's power amplifier will be possible. The CD players and recordings available today are the start of our digital audio future, a future which promises extraordinary operating flexibility, convenience, and—above all—sound quality. ◇