

## Chapter 7

# Optical Storage Media

Current magnetic data storage carriers take the form of floppy disks or hard disks and are used as secondary storage media. Here, low average access time and adequate capacity can be offered for a reasonable price. However, since audio and video, either in compressed or uncompressed form, require higher storage capacity than other media, the storage cost for such continuous media data using traditional storage carriers is essentially higher.

Optical storage media offer a higher storage density at a lower cost. The *Audio Compact Disk*, the successor to *Long Play Disks* (LPs), is a commercially successful product in the entertainment industry. The computer industry has profited from this development, especially when audio and video should be stored digitally in the computer. This technology has been *the* main catalyst for the whole development of multimedia in computing because it is used in multimedia external devices. For example, external devices such as video recorders and DAT recorders (Digital Audio Tape) can be used for multimedia systems. The actual integration into the system is difficult, but not impossible [HS91, RSS90, SHRS90]. For this reason we will discuss the optical storage medium in more detail in this chapter. Other data storage media will not be considered here because they do not have the special properties which should be taken into account with respect to integrated multimedia systems.

This chapter provides an overview of the fundamentals for optical storage media.

Moreover, some analog and WORM (Write Once Read Many) systems will be briefly discussed. The CD-ROM and CD-ROM/XA are explained as they are derived from the CD-DA. Further developments with respect to multimedia such as CD-I, Photo-CD and DVI are presented. In addition to the *read-only* CD developments, there already exist techniques for writing to CD-WO and CD-MO. At the end of this chapter, the relationships between current CD technologies are shown and further developments are mentioned.

## 7.1 History

The video disk in the form of the *Video Long Play* (VLP) has been available for quite some time and was described in detail by 1973. The read-only video disk has not become a commercial success, although there are several *Write-Once* (WO) optical disks of different sizes and formats available on the market. These initial developments were based on analog techniques with the highest quality requirements at a moderate cost.

Ten years later, towards the end of 1982, the *Compact Disk-Digital Audio* (CD-DA) was introduced. This optical disk allows the digital storage of stereo-audio information at a high level of quality. N.V. Philips, in cooperation with the Sony Corporation, specified the CD-DA and developed the basic technology [MGC82, DG82, HTV82, HS82]. The CD-DA specification was published as the *Red Book* [Phi82] on which all other CD formats are based. In the five years since its introduction, approximately 30 million CD-DA recording devices and over 450 million CD-DA disks were sold.

In 1983, the extension of compact disk technology to storage of computer data was announced by N.V. Phillips and the Sony Corporation, and it was presented for the first time in November 1985. This *Compact Disk-Read Only Memory* (CD-ROM) specification was described in the *Yellow Book* [Phi85] and later became the standard ECMA-119 [ECM88]. The standard ECMA-119 specifies the CD-ROM physical format. The CD-ROM logical format is specified in the ISO Standard 9660 and came from the *High Sierra Proposal* (a proposal of industrial companies). It allows data access over file names and directories.

The *Compact Disk-Interactive* (CD-I) was announced by N.V. Phillips and the Sony Corporation in 1986. The CD-I specification is described in the *Green Book* and includes, besides the standard CD technology, a complete system description [Phi88]. In 1987, the *Digital Video Interactive* (DVI) was publicly presented. Here, the algorithms for compression and decompression of audio and video data stored on a CD-ROM are of importance.

In 1988, the *Compact Disk-Read Only Memory / Extended Architecture* (CD-ROM/XA) was announced. N.V. Phillips, the Sony Corporation and Microsoft specified the digital optical data carrier for several media and published the specification at the CD-ROM conference in Anaheim, California (USA) in 1989 [Phi89].

Since the beginning of 1990, developments in *Compact Disk-Write Once* (CD-WO) technology, as well as *Compact Disk-Magneto Optical* (CD-MO), have been well-known. The specifications of CD-WO and CD-MO are covered in the *Orange Book* [Phi91].

## 7.2 Basic Technology

In principle, optical storage media use the intensity of reflected laser light as an information source. A laser beam of approximately 780 nm wave length can be focused at approximately 1  $\mu\text{m}$ . In a polycarbonate *substrate layer* we encounter holes, corresponding to the coded data, which are called *pits*. The areas between these pits are called *lands*. Figure 7.1 shows a cut through an optical disk along a track. In the middle of the figure, the lands and pits are schematically presented.

The *substrate layer* is covered with a thin *reflective layer*. The laser beam is focused on the reflective layer from the substrate layer. Therefore, the reflected beam has a strong intensity at the lands. The pits have a depth of 0.12  $\mu\text{m}$  (from the substrate surface). The laser beam is lightly scattered at the pits, meaning it is reflected with a weak intensity. The signal, shown in Figure 7.1, denotes schematically the intensity of the reflected beam – a horizontal line is drawn as the threshold value. Hence, according to Figure 7.1, a compact disk consists of:

- The label.

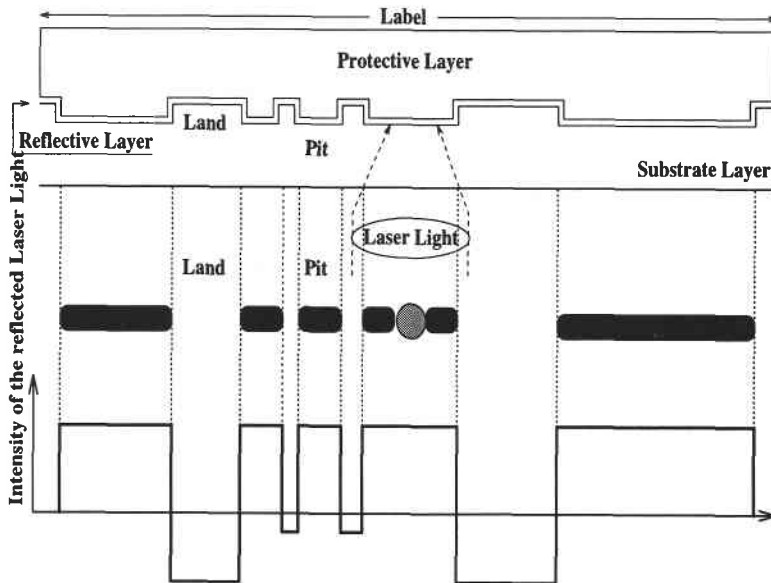


Figure 7.1: *Cut through an optical disk along the data trace. A schematic presentation with the layers (above), the “lands” and the “pits” (in the middle), and the signal (below).*

- The protective layer.
- The reflective layer.
- The substrate layer.

An optical disk consists of a sequential order of these pits and lands allocated in one track. Figure 7.2 shows an enlarged cut of such a structure.

In contrast to floppy disks and other conventional secondary storage media, the entire optical disk information is stored in one track. Thus, the stored information can be easily played back at a continuous data rate. This has advantages for audio and video data, as they are continuous data streams.

The track is a spiral. In the case of a CD, the distance between the tracks is  $1.6 \mu\text{m}$ . The track width of each pit is  $0.6 \mu\text{m}$ . The pits themselves have different lengths. Using these measurements, the main advantage of the optical disk in comparison

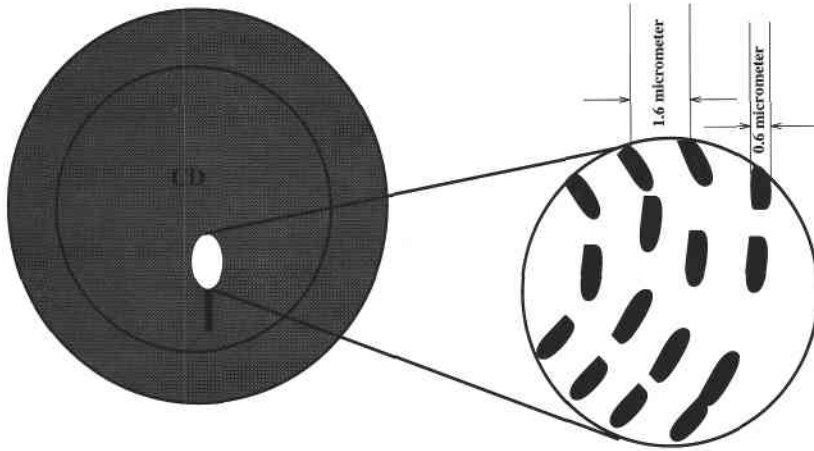


Figure 7.2: Data on a CD as an example of an optical disk (track with “lands” and “pits”).

to magnetic disks is that on the former 1.66 data bits per  $\mu\text{m}$  can be stored. This results in a data density of 1,000,000 bits per  $\text{mm}^2$ , which implies 16,000 tracks per inch. In comparison, a floppy disk has 96 tracks per inch.

While magnetization can decrease over time and in the case of tapes, for example, cross talk can occur, these effects are unknown in optical disks. Hence, this medium is very good for long-term storage. Only a decomposition or change of the material can cause irreparable damage. According to current knowledge, such effects will not occur.

The light source of the laser can be positioned at a distance of approximately one millimeter from the disk surface, and hence, it does not have to be positioned directly on the disk, respectively near the surface, as is the case with magnetic hard disks. This approach reduces friction and increases the life span of the involved components.

### 7.3 Video Disks and Other WORMs

The video disk, in the form of *Laser Vision*, serves as the output of motion pictures and audio. The data are stored in an analog-coded format on the disk; the reproduced data meet the highest quality requirements. The Laser Vision disk has a diameter of approximately 30 cm and stores approximately 2.6 Gigabytes.

Due to the similarities to LP records for audio information, the video disk was originally called the *Video Long Play* disk. It was described for the first time in Phillips' 1973 Technical Review [Phi73].

The motion picture on the video disk is encoded as frequency modulation, and the audio signal is mixed with the video signal. Figure 7.1 shows the principle of the recorded data. The main information of the mixed audio-video signal is the time at which the signal has the value zero. Hence, each zero cross-point corresponds to a change between a pit and a land on the disk. Such a change can occur at any time, and is written to the disk in a non-quantized form, i.e., the pit length is not quantized. Therefore, this method is time-continuous and *analog*.

Since the video disk was designed as *Read Only Memory* (ROM), many different write-once optical storage systems have come out, known as the *Write Once Read Many* (WORM) disk. An example is the *Interactive Video Disk*. This disk is played at a *Constant Angular Velocity* (CAV). On each side, 36 minutes of audio and video at a rate of 30 frames per second can be stored and retrieved. One can also store around 54,000 studio quality images per side.

Write-once storage media have a capacity between 600 MBytes and approximately 8 Gigabytes. The diameter of the disks is between 3.5 and 14 inches. The main advantage of a WORM disk, compared to other mass storage media, is the ability to store large amounts of data which may not be changed later, i.e., an archive which is secure. To increase capacity, *juke-boxes* are available, which allow the stocking of several disks and lead to capacities of over 20 Gigabytes.

The following interesting peculiarities occur when using a WORM disk:

- *Media Overflow* can occur when a WORM disk is nearly full. Prior to writing

data to a WORM disk, it must be determined if the recorded data can fit onto the disk and/or if the data should be stored on different disks. Further, it is required that if the data need to be stored on more than one physical disk, the time point at which data should be written to another physical disk must be determined. This approach is especially important for continuous media because these media streams can only be interrupted at certain points.

- *Packaging* refers to the problem of fixed-block structures in WORMs. Only data sentences of a given size can be written. For example, if the block size is 2,048 bytes, and only one byte is written, 2,047 bytes will be recorded without any information content.
- *Revision* refers to the problem of subsequently marking invalid areas. For example, during changes of a document invalid areas are created on the disk(s). These areas have to be subsequently rewritten, which implies a document may be distributed over several WORMs after a while. Here, the distribution of the document over several disks should not disturb the output of the data stream of a continuous medium.

There are also other problems with WORM disks such as the following: besides the number of incompatible WORM disk formats, most multimedia systems lack adequate software support for the WORM disks, which results in a poor integration of WORM technology into the computer environment.

## 7.4 Compact Disk Digital Audio

### 7.4.1 Preliminary Technical Background

The CD has a diameter of 12 cm; the disk is played at a *Constant Linear Velocity* (CLV). Therefore, the number of rotations per time unit depends on the particular radius of the accessed data. The spiral-shaped CD track consists of approximately 20,000 windings. In comparison, an LP disk has only approximately 850 windings.

Information is stored according to the principle shown in Figures 7.1 and 7.3. The length of the pits is always a multiple of  $0.3 \mu\text{m}$ . The transition from pit to land

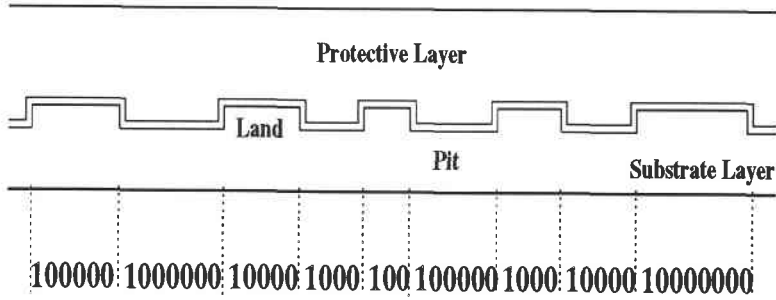


Figure 7.3: *Lands and pits with their related digital data stream.*

and from land to pit corresponds to the coding of a 1 in the data stream. A 0 is coded as no transition. Figure 7.3 shows a data stream as a sequence of lands and pits, and below it is the corresponding digital data stream.

### Audio Data Rate

The audio data rate can be easily derived from the given sample frequency of 44.1 kHz and the 16-bit linear quantization. The stereo-audio signal obeys the pulse-code modulation rules and the following audio data rate is derived:

$$\begin{aligned}
 \text{Audio data rate}_{CD-DA} &= 16 \frac{\text{bits}}{\text{sample}} \times 2 \text{ channels} \times 44100 \frac{\text{samples}}{s \times \text{channel}} \\
 &= 1,411,200 \frac{\text{bits}}{s} = 1,411,200 \frac{\text{bits/s}}{8 \text{ bits/byte}} \\
 &= 176.4 \frac{\text{kbytes}}{s} \cong 172.3 \frac{\text{Kbytes}}{s}
 \end{aligned}$$

Analog LPs and cassette tapes have a signal-to-noise ratio between 50 dB and 60 dB. The quality of the CD-DA is substantially higher. As a first approximation, we can assume 6 dB per bit during the sampling process. Hence, with 16-bit linear sampling, we obtain the following:

$$S/N_{CD-DA} \cong 6 \frac{\text{dB}}{\text{bit}} \times 16 \text{ bits} = 96 \text{ dB}$$

The signal-to-noise ratio is exactly 98 dB.



## Capacity

A CD-DA play time is at least 74 minutes. With this value, the capacity of a CD-DA can be easily determined. The following example shows the computation of a capacity for pure audio data without taking into consideration additional information, such as error correction:

$$\begin{aligned} \text{Capacity}_{CD-DA} &= 74 \text{ min} \times 1,411,200 \frac{\text{bits}}{\text{s}} = 6,265,728,000 \text{ bits} \\ &= 6,265,728,000 \text{ bits} \times \frac{1}{8 \frac{\text{bits}}{\text{byte}}} \times \frac{1}{1024 \frac{\text{bytes}}{\text{Kbyte}}} \times \frac{1}{\frac{\text{Kbytes}}{\text{Mbyte}}} \cong 747 \text{ Mbytes} \end{aligned}$$

### 7.4.2 Eight-to-Fourteen Modulation

Each change from pit to land and from land to pit corresponds to the coding of a  $1$  which is sent across a communication channel as the *channel bit*.

Pits and lands may not follow too closely one after another on a CD-DA since the resolution of the laser is not sufficient to read such direct pit-land-pit-land-pit sequences (i.e., 11111 sequences) correctly. Therefore, an agreement was negotiated that at least two lands and two pits always occur consecutively in a sequence. Hence, between two  $1$ s there always exist at least two  $0$ s.

On the other hand, pit or land sequences are not allowed to be too long; they must keep a maximal distance. Otherwise, no phase-correct synchronization signal (clock) can be derived. Hence, the maximal length of the pits and lands is limited. At most, ten  $0$ s (as the channel bits) can follow one after another.

For these reasons, the bits written on a CD-DA, in the form of pits and lands, do not directly correspond to the actual information; before recording, *Eight-to-Fourteen Modulation* is applied [HS82]. Using this transformation, the regularity of the minimal and maximal distances is met.

Eight-bit words are coded as 14-bit values. Given the minimal and maximal allowable distances, there are 267 valid possibilities. 256 possibilities are used. For example, the code includes the following two entries shown in Table 7.1.

Audio Bits	Modulation Bits
00000000	01001000100000
00000001	10000100000000

Table 7.1: *Two values out of the code table.*

<b>Audio Bits</b>	00000000					00000001			
<b>Modulation Bits</b>	0	100	1000	100000		10000	100000000		
<b>Filling Bits</b>	0	10				100			
<b>Channel Bits</b>	0	10	0	100	1000	100000	100	10000	100000000
<b>On the CD-DA</b>	l	pp	p	lll	pppp	llllll	ppp	lllll	ppppppppp

Table 7.2: *Integration of the filling bits.*

Through the direct consecutive insert of the modulation bits (14-bit-values), the minimal allowed distance of two bits and the maximal distance of ten bits may still not be followed. Therefore, three additional bits are inserted between two consecutive modulation symbols so that the required regularity can be met. The filling bits are chosen depending on the neighboring bits. Table 7.2 shows an example (the audio bits are taken from Table 7.1) to clarify the integration of the filling bits and the whole transformation process from an 8-bit audio word to its CD-DA representation of lands and pits.

### 7.4.3 Error Handling

The goal of error handling on a CD-DA is the detection and correction of typical error patterns [HTV82]. A typical error, a consequence of a scratch and/or pollution, can be characterized as a *burst error*.

On the first level, a two-stage error correction is implemented according to the Reed-Solomon algorithm: for every 24 audio bytes, at the first stage, individual byte errors are recognized and corrected; at the second stage, double byte errors are recognized and corrected; also, other fault bytes in the sequence can be recognized, but they cannot be corrected using this approach.

On the second level, real consecutive data bytes are distributed over several frames. A *frame* consists of 588 channel bits, which correspond to 24 audio bytes. Therefore, the audio data is stored on the CD-DA in an interleaved form. This means that due to a burst error, only parts of data are modified.

An error rate of  $10^{-8}$  is specified. More precisely, the burst errors, which are distributed over a maximal 7 frames, can be recognized and corrected. This corresponds to a track length of 7.7 mm. For example, a 2-mm-diameter hole in a CD-DA means that the audio data can still be played correctly. Nevertheless, experiments have shown that not all devices correct each error according to the given specification. The above described method for error correction is known as *Cross Interleaved Reed-Solomon Code* (CIRSC).

#### 7.4.4 Frames, Tracks, Areas and Blocks of a CD-DA

Audio data, error correction, additional control and display-bytes, and a synchronization pattern all constitute *frames*.

- *Audio data* are divided into two groups, each consisting of 12 audio bytes. They contain the high and low bytes of the left and right channels.
- *Error-detection* and *error-correction* bytes (4 bytes per frame) are inserted, according to the above description, at both stages.
- Each frame has a *control* and a *display* byte. It consists of eight bits, which are marked with *P*, *Q*, *R*, *S*, *T*, *U*, *V* and *W* (*subchannel bits*). These subchannel bits are drawn together over 98 frames for each subchannel and used together. Hence, there are eight subchannels, each having 98 bits, of which 72 are used for the actual information. The 98 frames together build a *block*. Unfortunately, sometimes the blocks are also called frames.

So the *P*-subchannel is used to differentiate between a CD-DA (with audio data) and a CD with other computer data. The *Q*-subchannel is used, for example, in the following:

- The *lead-in area* for storage of the directory content.

- The rest of the CD-DA for a specification of the relative time inside a track, and for the absolute time specification of the CD-DA.
- The *synchronization pattern* determines the beginning of a frame. The pattern consists of twelve *1s* and twelve *0s* as channel bits, and three filling bits.

Table 7.3 shows an overview of frame components with corresponding bytes and bits.

	Audio Bits	Modul.Bits	Fil.Bits	Ch.Bits	Together
Synchronization			3	+ 24	= 27 bits
Control & Display		i.e.	(14 + 3)		= 17 bits
12×Data	12 × 8	i.e. 12×	(14 + 3)		= 204 bits
4×Error Handling		i.e. 4×	(14 + 3)		= 68 bits
12×Data	12 × 8	i.e. 12×	(14 + 3)		= 204 bits
4×Error Connection		i.e. 4×	(14 + 3)		= 68 bits
Frames Together					= 588 bits

Table 7.3: *Components of a frame.*

Using these data, different data streams with corresponding rates can be recognized [MGC82]:

- The *audio bit stream* (also called the *audio data stream*) carries  $1.4112 \times 10^6$  bits/s. Here, only the 16-bit quantized samples are taken.
- The *data bit stream* consists of the audio bit stream, including the control and display bytes, and necessary bytes for error handling. The number of bits can reach the value of  $1.94 \times 10^6$  bits/s.
- The channel bit stream includes the data bit streams with the *Eight-to-Fourteen Modulation*, the *filling bits* and the *synchronization bits*. The data rate is approximately  $4.32 \times 10^6$  bits/s.

A CD-DA consists of the following three *areas*:

- The *lead-in area* includes the directory of the CD-DA. Here, the beginning of the individual tracks are registered.
- The *program area* includes all tracks of the CD-DA. Here, the actual data are stored.
- At the end of each CD-DA there is a *lead-out area*. This area is used mainly to help the play-recorder when the reader head accidentally goes beyond the program area.

The program area of each CD-DA can consist of 99 *tracks* of different lengths. It includes at least one track and each track incorporates at most one song, or a sentence of a symphony. If the program area is randomly accessed (e.g., access 5th song), the reader head is positioned to the beginning of the particular track (e.g., position to the beginning of the 5th track).

Each track can have (according to the *Red Book*) several *index points*. Therefore, at certain places, direct positioning can occur. Mostly, only two a priori defined *Index Points (IPs)* are used, the  $IP_0$  and  $IP_1$ .  $IP_0$  marks the beginning of each track.  $IP_1$  defines the beginning of the audio data inside each track. The area between  $IP_0$  and  $IP_1$  is called *track pregap*. CD-DA disks possess a track pregap of two to three seconds for each piece.

Another structure, called *block* (Figure 7.4), was introduced in addition to frames and tracks, but it does not have a special meaning in CD-DA technology. It is introduced at this point only because of comparability reasons to other CD technologies discussed in the next sections. A CD-DA block includes 32 frames and stores 2,352 bytes.



Figure 7.4: CD-DA block layout according to the “Red Book.”

### 7.4.5 Advantages of Digital CD-DA Technology

Errors on a CD-DA can be caused by damage or pollution. The CD-DA is not sensitive, with respect to uncompressed audio, to the usual appearance of reading errors. The CD-DA specification, in the form of the *Red Book*, serves as the basis for all optical CD storage media. For example, *Eight-to-Fourteen Modulation* and the *Cross Interleaved Reed-Solomon Code* are always used. Hence, a fundamental specification has been developed which is used in many systems and means compatibility for many systems.

The disadvantage of the CD-DA technology is that its achieved reliability is too low for the storage of computer data. This led to a further development of the CD technology.

## 7.5 Compact Disk Read Only Memory

The *Compact Disk Read Only Memory* (CD-ROM) was designed as the storage format for general computer data – in addition to uncompressed audio data [Che86, FE88, Hol88, LR86, OC89]. Further, CD-ROM technology has been planned to be the basis for storage of other media (e.g., video) [KSN<sup>+</sup>87, Wil89]. This was specified in the *Yellow Book* by N.V. Phillips and the Sony Corporation [Phi85]. The *Yellow Book* was later accepted as the ECMA standard [ECM88].

CD-ROM tracks are divided into *audio* (corresponding to the CD-DA) and *data* types. One track itself may either contain audio only or data only. A CD-ROM can contain both types of tracks, tracks with audio and other tracks with data. In such a mixed form, the data tracks are usually located at the beginning of the CD-ROM and then followed by the audio tracks. Such a CD is called a *Mixed Mode Disk* (see Figure 7.11).

### 7.5.1 Blocks

A CD-DA has an error rate of less than  $10^{-8}$  and allows random access to individual tracks and index points. The use of a CD-ROM with its general-purpose computer

data requires much better error correction and random access to a data unit with a higher resolution than the track.

This data unit is called a *block*, meaning the *physical block*. In the ISO 9660 standard, there also exists the notion of a *logical block*. The logical block has similar properties to the sectors of other media and file systems. A CD-ROM block consists of 2,352 bytes of a CD-DA block. Hence, the de facto CD-DA standard serves as the basis for the de facto CD-ROM standard.

Out of the 2,352 bytes of a block, 2,048 or 2,336 bytes (depending on whether computer data or audio data are stored on CD-ROM) can be used for user data. The remaining bytes are used for the identification of random access, as well as for another error correction layer, thereby lowering the error rate further.

75 blocks per second are played back, each consisting of 32 frames. Each frame is 73.5 bytes (588 bits).

$$\text{Block} = 1411200 \frac{\text{bits}}{\text{s}} \times \frac{1}{75} \text{s} \times \frac{1}{8 \text{bits/byte}} = 2352 \text{bytes}$$

Figure 7.5 shows the data hierarchy of a CD-ROM with audio data.

### 7.5.2 Modes

The CD-ROM was specified with the following goal: it should serve to hold uncompressed CD-DA data and computer data. This goal is achieved by introducing two modes: *mode 1* and *mode 2*.

#### CD-ROM Mode 1

CD-ROM mode 1 serves as the actual storage of computer data (Figure 7.6). The block contains 2,048 bytes for information storage out of the available 2,352 bytes. The 2,352 bytes are split into the following groups:

- 12 bytes for *synchronization*, i.e., for the detection of the block beginning.