Optical Data Storage

Tom D. Milster
•Optical Sciences Center, The University of Arizona, Tucson, AZ 85721, USA
Q1
e-mail: milster@arizona.edu

Abstract

In its most general definition, optical data storage is simply using light as a tool to store and retrieve data. Optical data storage is found in popular consumer products. Compact Discs (CDs), Digital Versatile Discs (DVDs), and MiniDiscs (MDs), are all forms of optical data storage. More advanced forms of optical data storage include high-speed devices and library products. All optical data storage devices use optical principles to achieve high data density, rugged packaging, reliable information retrieval, and cost-effective production. There are many forms of optical storage media and many types of optical systems used to scan data. This chapter discusses the basic principles of optical data storage, types of commercial optical media available in 2003, several performance parameters, and some interesting prospects for future systems.

Keywords

optical data storage; optical recording; compact disks; digital versatile disks; optical disks; optical memory; optical servos.

1

Inspiration for the Invention and Early Development

The optical disk was envisioned in 1958 by an eclectic engineer named David Paul Gregg, who patented the idea in 1961 [1–3]. At that time, magnetic disks were in their infancy, and there was no low-cost "videodisk" system that could be marketed as a consumer product. As told by the inventor, inspiration for the videodisk system came from the picture shown in Fig. 1, which is a photograph from a trade magazine that appeared in the mid 1950s. The photograph was produced from an early scanning electron microscope. The narrow lines in the lower-left corner of the picture, which were written by an electron beam, are 0.030 μm wide and are separated by 0.070 μm. [For a sense of reference, if we were to magnify a 1.0-μm-wide line to the length of one yard, the edge thickness of a piece of paper under the same magnification would be about the length of a football field. The lines in Fig. 1 would be about 1 in wide. ($1\ \mu m = 10^{-6}\ m = 0.000039$ in)] Gregg saw the picture and, while on horseback in Mexico, imagined a rotating plastic disk with tracks of data marks read by an inexpensive optical system. As originally envisioned, a master disk is first written with an electron beam. The master disk is used to produce low-cost replicas, which are then sold to consumers.

In many ways, Gregg was prophetic. The compact discs (CDs), a mature development and improvement of Gregg's invention, is an immensely successful consumer product. Digital versatile discs (DVD) systems are advanced versions of CDs. Today's most advanced systems, which use blue laser diode optical systems, can utilize electron-beam machines to make master disks. Someday, it may be possible to record information with the 0.030-μm line width observed in Fig. 1, which would result in a disk with data density several thousand times the data density found on DVDs.

Early development of the optical disk probably started with Gregg's employment at Mincon, a 3 M company, in the early 1960s. Greg left Mincon and formed a small company called Gauss Electrophysics in 1964 that continued his optical disk development. Within a few

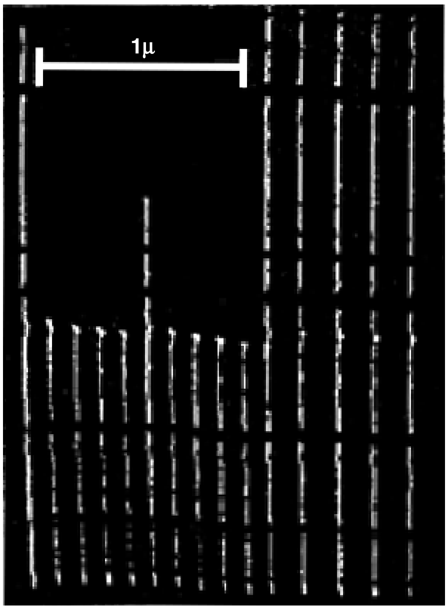


Fig. 1 This picture of lines written by an electron beam was the inspiration for the invention of the videodisk in the late 1950s. Lines in the lower left-hand corner of the picture are 0.030 µm wide and spaced by 0.070 µm. If it were possible to reliably record data at this density, a 130-mm diameter optical disk would have several thousand times the capacity of a digital versatile disc (DVD)

years, other companies became interested in the idea. For example, Gregg and his colleague, Keith Johnson, made a complete technical disclosure to Philips NV in 1967, in hopes of obtaining funding for development of a consumer video disk. Philips declined, but Tim Scott, another Gauss employee, courted entertainment giant MCA into investing in the idea. MCA's motivation was to distribute 11 000 movie titles from its massive film library as consumer video products. MCA, in 1978, with cooperation from Philips, eventually produced the DiscoVision system that used replicated disks called *LaserDiscs*. However, the LP-sized disks were problematic. Production problems with contamination and process control limited consumers to only 50 movie titles. In addition, customers would typically return LaserDiscs several times before one worked with acceptable quality. Pioneer, who arranged a joint effort with MCA in 1977, eventually improved the quality of the LaserDiscs with better manufacturing practices. In 1979, IBM joined with MCA to form a joint venture called DiscoVision Associates (DVA). This company continued to manufacture disks until competition with VHS and Pioneer's superior product forced closure of its Carson, California, plant in 1981. DVA continues to reap benefits from optical disk technology in the form of patent royalties and licenses, although Pioneer now wholly owns it.

For the mass consumer market in the United States, disk technology was finally accepted with the advent of the Philips/Sony Compact Disc (CD) product in 1983. According to the best information available to the author, these companies (and any others manufacturing optical disks) still pay royalties to DVA, based in large part to Gregg's patent. An interesting and detailed history of the optical disk is provided in Reference [3].

2

Data Marks and Spaces: The Information Carriers

Digital information is stored on optical disks in the form of arrangements of data marks in spiral tracks. Small sections of CD and DVD surfaces are illustrated in Fig. 2, which also displays representations of laser spots that are focused on the surfaces to write and read data. CDs typically use a 1.6-µm track pitch, which is the

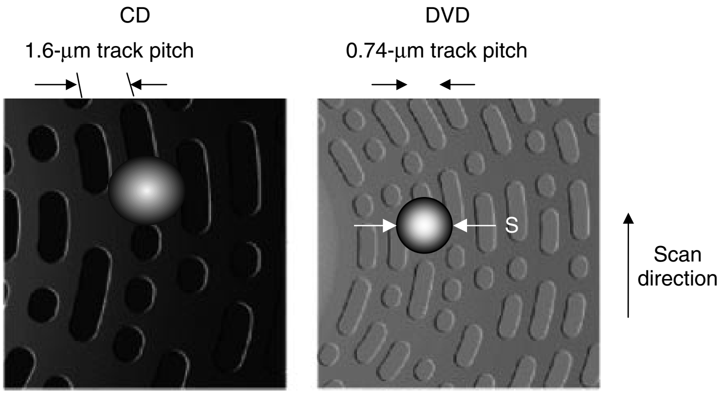


Fig. 2 Small sections of a compact disc (CD) and a digital versatile disc (DVD) are displayed relative to the laser spots that are focused onto them during recording and readout

radial distance between turns of the spiral tracks. Data marks are nearly one-half as wide as the track pitch. Lengths of data marks and spacings between marks are determined by the encoding scheme [4, 5] used to translate user data into mark patterns along each track, which is described in more detail in Sect. 5.8. The width of the CD laser spot is slightly smaller than the track pitch. DVD media are similar to CD media, except that the track pitch is smaller (0.74 µm), data marks are shorter and narrower, and the laser spot diameter *s* is smaller. Since there are more data marks per unit area on a DVD compared to a CD, the DVD holds more data.

3
Optical Data Storage Basic Principles

Storage and retrieval of data on optical disks is described in two simple steps. First, data marks are recorded on a surface. Data marks are prerecorded, like on music CDs, or they are recorded by users on blank disks, like with CD-recordable products. The second step is retrieval of information from the disk, where a light beam scans the surface. Modulation in the reflected light is used to detect the data-mark pattern under the scanning spot.

The process for exposing data marks on a recordable optical disk is shown in Fig. 3, where an input stream of digital information is converted with an encoder and modulator into a drive signal for a laser source. The laser source emits an intense light beam that is directed and focused onto the surface with illumination optics. As the surface moves under the scanning spot, energy from the intense scan spot is absorbed, and a small, localized region heats up. The surface, under the influence of heat beyond a critical writing threshold, changes its reflective properties. Modulation of the intense light beam is synchronous with the rotation, so a circular track of data marks is formed as the surface rotates. The scan spot is moved slightly as the surface rotates to allow another track to be written on new media during the next revolution.

Data marks on prerecorded disks are fabricated by first making a master disk with the appropriate data-mark pattern.

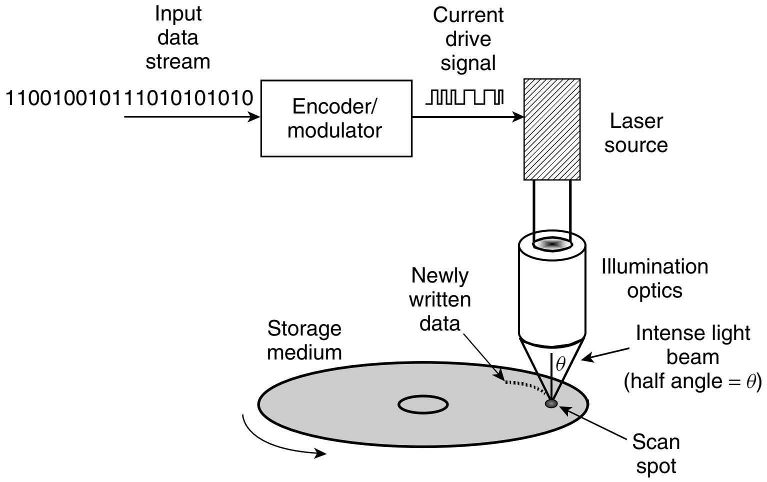


Fig. 3 The process of recording data onto an optical disk starts with the user input data stream converted to a current drive signal for the laser diode. Intense pulses from the laser cause physical changes in the surface of the recording medium as the disk spins, which result in spiral tracks of data marks

Masters for prerecorded CDs and DVDs are often exposed in a similar manner to exposing data marks on recordable optical disks, except that the light-sensitive layer is designed to produce pits in the master that serve as data marks in the replicas. Inexpensive replicas of the master are made with injection-molding equipment.

Readout of data marks on the disk is illustrated in Fig. 4, where the laser is used at a constant output power level that does not heat the data surface beyond its thermal writing threshold. The laser beam is directed through a beam splitter into the illumination optics, where the beam is focused onto the surface. As the data marks to be read pass under the scan spot, the reflected light is modulated. Modulated light is collected by illumination optics and directed by the beam splitter to servo and data optics, which converge the light onto detectors. The detectors change light modulation into current modulation that is amplified and decoded to produce the output data stream.

4

Commercial Media

There are several types of optical disks, which are differentiated by the type of data marks on the recording layer. The most popular disks are based on pit-type, magneto-optic, phase-change, and dye-polymer data-mark technologies. Several commercial trade names are associated with the four technologies, as shown in Table 1. In this section, basic data-mark technologies are reviewed, and commercial formats are listed.

Pit-type data-mark technology for CD read-only memory (ROM) and DVD-ROM products is based on a very simple scattering phenomenon. Most CDs, like music and data distribution CDs, are ROM disks. The small data-mark pits are arranged in spiral tracks around the center of the disk, as shown in Fig. 2. The pit lengths are about 1 to 3 μm long. The pits along a track are nearly uniform and measure about 0.5 to 0.8 μm

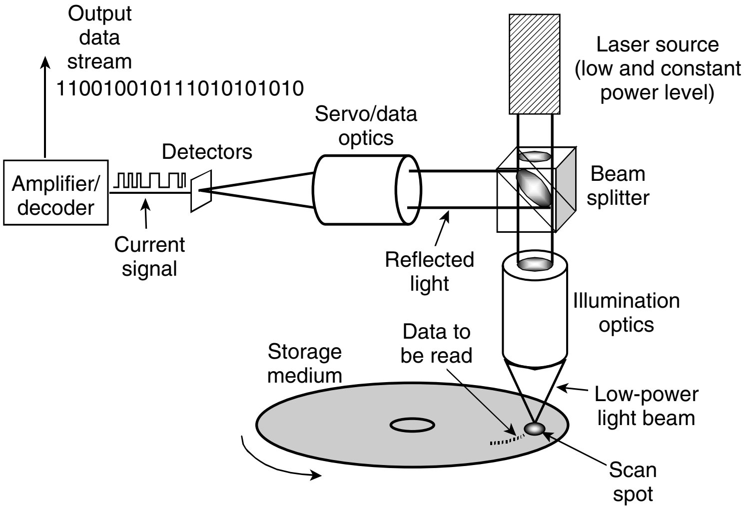


Fig. 4 A constant, low-power laser beam scans a data track to readout data from the disk. Reflected light is modulated by the data-mark pattern. The reflected light is directed to servo and data detectors with a beam splitter, which convert the light modulation in a current signal that is then decoded

Tab. 1 Four technologies of commercial optical disks

Disk technology	***Description***
CD-ROM or DVD-ROM	Compact disc (CD) and digital versatile disc (DVD) products use pit-type technology. CD and DVD products are read-only memories (ROMs), that is, they are used for software or entertainment distribution and cannot be used for recording information.
CD-R	Compact-disk-recordable (CD-R) products use dye-polymer technology. CD-R products can be used for recording information, but, once the information is recorded, it cannot be erased and reused.
CD-RW, DVD-RW, DVD +RW, DVD – RW, DVD-RAM, BD	Compact-disk-rewriteable (CD-RW), certain DVD products and the Blu-ray Disk (BD) use phase-change technology. Data can be erased and the disks reused.
MO	Erasable disks using magneto-optic (MO) technology are popular for workstation environments. Data can be erased and the disks reused.

wide. As the light spot passes over a pit, most of the reflected light scatters away from the illumination optics. The remaining light collected by the objective lens is small compared to the amount of light that gets collected when the spot is over a smooth portion of the track, where the disk surface acts as a mirror to the focused light. The data signal is derived from the detector that senses the amount of collected light, as shown in Fig. 4.

The amount of light scattered from each pit depends on the depth of the pit and the size of the laser beam illuminating it. A simplistic explanation is that the portion of the laser spot reflected from the pit exhibits a phase change because of the additional path that the light must traverse compared to the portion of the laser spot that is reflected from the surrounding flat area of the recording surface. The two portions of the spot interfere destructively upon propagation of the light back to the objective lens if the effective depth of the pit is one-fourth of the illuminating laser wavelength. In practice, the pit depth profile is designed to not only provide good data signal modulation but also good tracking performance, as explained in Sect. 5.6, which is optimized at a slightly different pit depth [6].

In order to increase the amount of light reflected to the detector, the entire recording surface of the CD, including both pits and areas between pits, is coated with aluminum. The aluminum is then coated with a lacquer or other protective layer, onto which the label is printed. The read out optical system focuses light through the clear surface of the disk, as shown Fig. 5. The data-mark patterns and tracks are actually located nearer the label than the clear side through which the user views the rainbow diffraction pattern that forms because of the close radial spacing of the tracks. The plastic CD substrate is 1.2 mm thick, which is designed so that contamination, like fingerprints or scratches, on the surface of the disk does not adversely affect disk performance during read out [7]. DVD substrates are only 0.6 mm thick, which implies that DVDs are intrinsically more sensitive to contamination than CDs. (Although DVDs are intrinsically more sensitive to contamination than CDs, a more powerful error correction code on DVDs minimizes problems associated with contamination.)

Dye-polymer or dye-monomer technology is used in CD-R products. Dye polymers/monomers are organic films that are

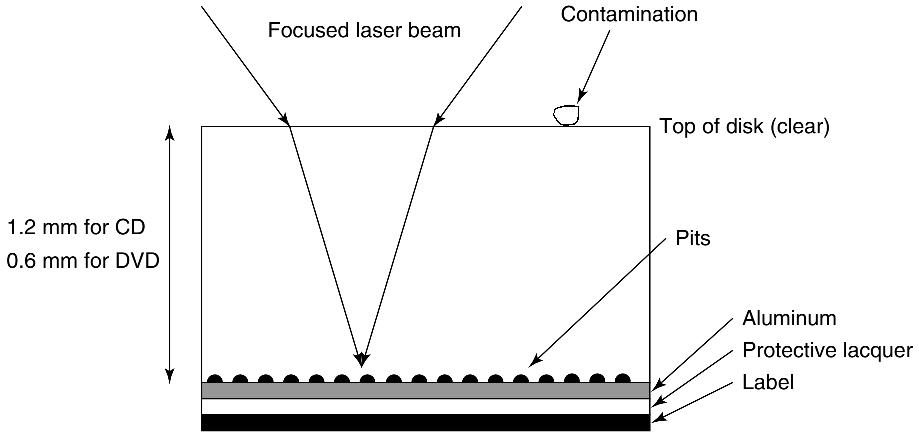


Fig. 5 A CD is used so that the focused laser light is passed through the clear side of the substrate and illuminates the data-mark pattern on the other side. The thickness of the cover layer is designed to reduce effects of contamination on the surface of the disk, like dust or fingerprints

ablated to form pits along tracks [8–10]. To form a pit, a high-power focused spot locally heats a micron-sized area. The dye polymer absorbs a large percentage of the laser energy. Owing to the low thermal conductivity of dye polymers, extremely high temperatures are reached. Although the exact mechanism of pit formation in CD-R is not known, a simple explanation is that, in the heated area, the dye material is vaporized or heated to the point that material flows to form a pit. To read data, a low-power laser beam scans the track, and the collected light is sensed with a simple detector. Light scattering from the data marks modulates the reflected light, which is similar to the phenomenon described above for pit-type media. Since the recording process for CD-R's is destructive, the user only writes data marks once. Data marks cannot be erased and rewritten.

There are three classifications of dyes used to make CD-Rs, which are cyanine, metallized azo, and phthalocyanine. At a laser wavelength of 0.78 µm, where CD players are designed to operate, there are only slight differences in performance between the dyes during the writing process. All the dyes absorb laser light and heat the recording surface. In addition to the dye layer, CD-Rs have a reflective layer, like the CD-ROM. However, the reflective layer is usually silver or gold instead of aluminum.

Different combinations of dyes and reflective layers influence the visual appearance of the CD when viewed from the clear side of the substrate. Since the visible spectrum is of a shorter wavelength than the recording laser wavelength for CDs, the dye usually appears with a characteristic semitransparent color. The absorption spectra of several dyes as measured from commercial disks are shown in Fig. 6. Table 2 lists some combinations of dyes, reflectors, and the resulting synthesized

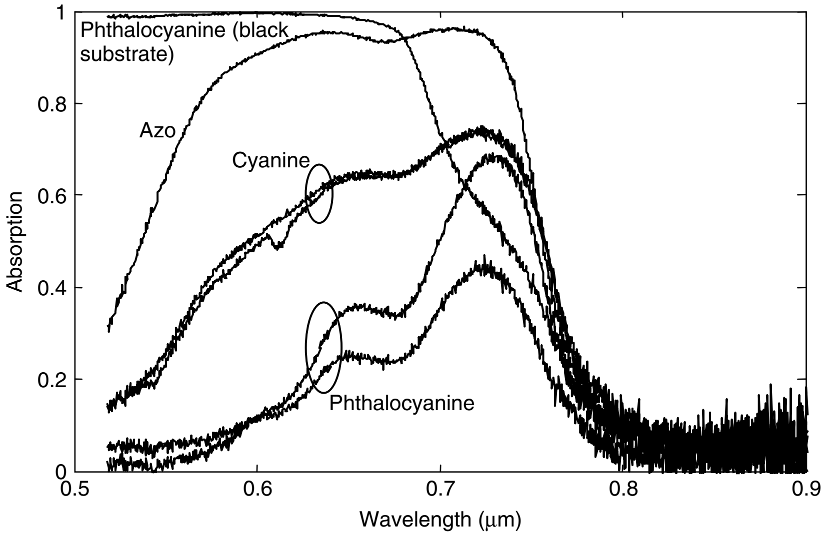


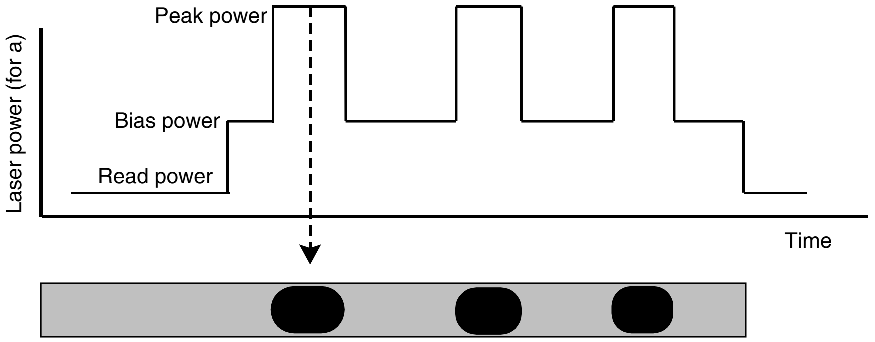
Fig. 6 Absorption spectra for several commercially available dyes used on CD-R media. Dyes absorb between 10 to 20% at the 0.78-µm wavelength of the laser diode, but they vary substantially in absorption characteristics in the visible range between 0.5 to 0.68 µm

Tab. 2 CD-R materials, reflective layer, and the resulting color as observed through the clear side of the substrate

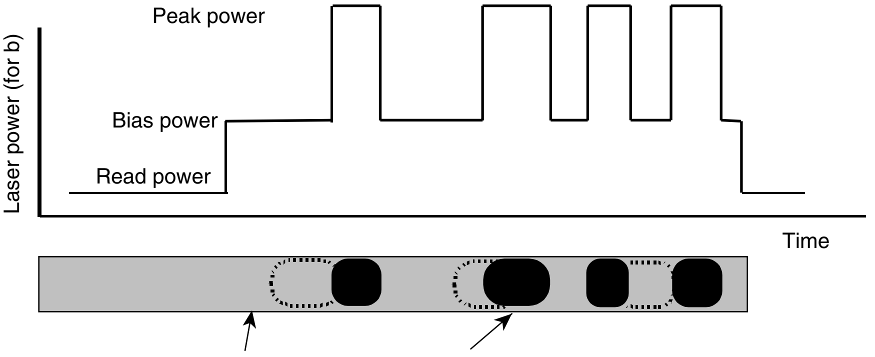
Dye material	***Disk color***	
	Gold reflector	***Silver reflector***
Cyanine	Green	Green/blue
Phthalocyanine	Gold	Light green
Metallized azo	–	(Dark) blue
Advanced phthalocyanine	Gold	–
Formazan (hybrid cyanine/phthalocyanine combination)	Green/gold	–

color to the eye. For example, the green CD-R, the cheapest combination, uses the cyanine dye [11]. By itself, the cyanine dye is blue in color, but, together with the gold reflector, the recording surface appears green. Cyanine's ability to maintain reflectivity is poor, which gives it an expected lifetime of only about 10 years. Improved-formula cyanine dyes in combination with silver reflectors (blue synthesized color) show better performance, which is better than twenty-year lifetime after recording. The gold-colored CD-R uses a phthalocyanine dye and a gold reflector. The dye is transparent by itself, so the gold color shines through. Modulation in the reflected light caused by writing on the gold medium is the best of all CD-R media, and lifetime of such CD-R's is said to be over 100 years [12]. Phthalocyanine dye with a silver reflector is the most commonly available combination, which appears light green in color. The disk substrate is normally transparent to visible wavelengths. Some "black CDs" have a substrate that is opaque to visible radiation shorter than 0.68 μm, as shown in Fig. 6, but exhibit similar absorption characteristics to other disks at longer wavelengths closer to the laser at 0.78 μm. The black-disk absorption spectrum shown in Fig. 6 uses a phthalocyanine dye. Blue media are made of azo dyes. Like cyanine, the azo dye is blue, but azo disks use a silver reflector, which results in a blue synthesized color.

CD-RW products use a different recording layer material than CD-Rs. In addition to allowing the user to write data marks, data marks can also be erased with multiple cycles before degradation. CD-RWs use phase-change technology, which is based on differences of the crystalline and amorphous states of semimetal alloys, like AgInSbTe or GeSbTe [13–15]. To record or erase phase-change data marks, a high-power focused spot locally melts the medium in micron-sized regions as the disk spins. The thermal cycle of the local regions determines if the region will stabilize in a crystalline or an amorphous state. By controlling the energy in the focused spot, the thermal cycle and the state of the material can be controlled. For example, a high-power laser pulse and rapid cooling quenches the material into the amorphous state, as shown in Fig. 7, when the laser is pulsed to the peak-power level. A lower-power laser beam and slow cooling anneals the material into a crystalline state. Usually, marks are in the amorphous state and the background is in the crystalline state. CD-RW media are "write dark" media, which means that the amorphous state of the data marks does not reflect as much light as the crystalline background. Some commercial media are "write bright," in that the recording layer is initially in the amorphous state, and bright crystalline marks are written on the dark amorphous background [16]. In Fig. 7(a), a virgin track is exposed to a certain mark pattern. Three clearly defined, dark amorphous marks are formed along the track,



(a) Before overwrite



(b) After overwrite

Fig. 7 The process of recording data marks on CD-RW media. (a) A laser power modulation signal with three peaks creates a pattern of three amorphous marks in a bright crystalline background. The peak laser pulses are use to quench the medium into the amorphous (low reflectivity) state. (b) Formation of a new data pattern, where a different laser pulse pattern is used. In the new pattern, a series of five pulses creates the five new data marks. The bias power between pulses anneals the medium into the crystalline (high-reflectivity) state, which erases the old data

where each data mark corresponds to a peak-power laser pulse. The bias power ensures that regions between marks anneal into the high-reflectivity crystalline state. When new data are written on the track, as shown in Fig. 7(b), a similar sequence of laser pulses are used, except that the laser pulses correspond to the new data-mark pattern. Old data marks are overwritten and replaced with either crystalline or amorphous material of the new pattern. The phase-change process inevitably involves a mechanical deformation of the material. Therefore, the number of direct overwrite cycles is limited to several thousand. Like a CD player, the detector in a CD-RW player simply senses the amount of collected light. The data signal is derived from the detector current.

The readout signal contrast is optimized in a CD-RW player by adding several thin-film layers around the recording material. The effect of the layers is to produce a thin-film reflective filter. A typical configuration of the storage layers in CD-RW products is shown in Fig. 8. The laser beam focuses

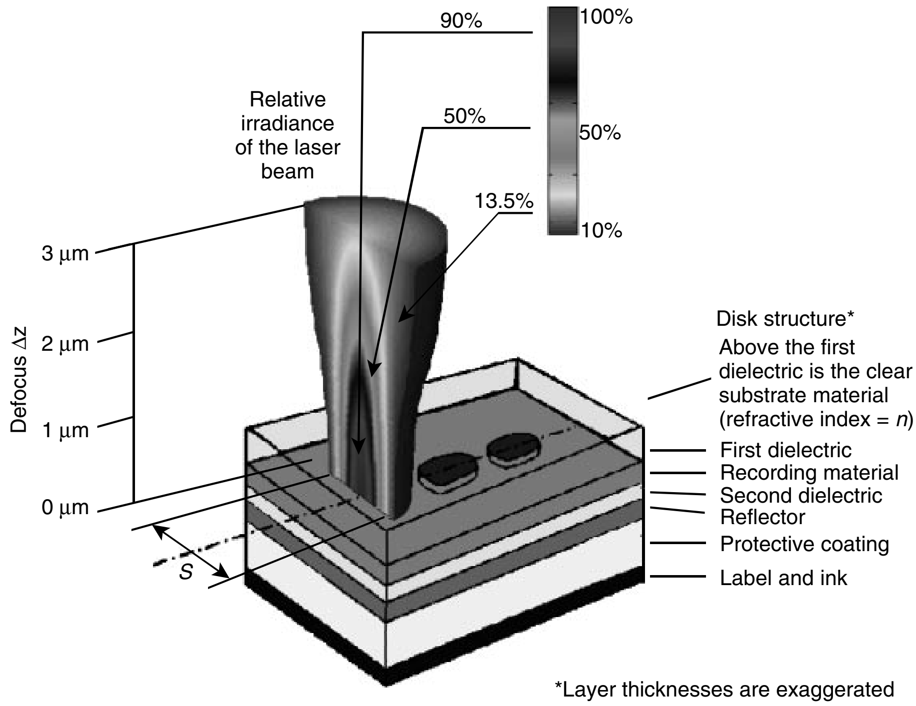
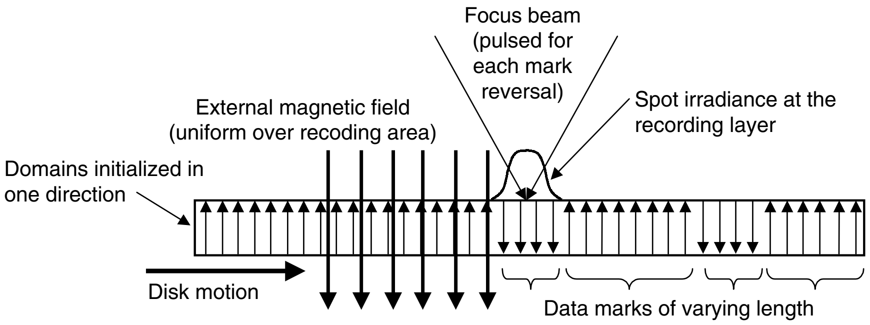


Fig. 8 Detail of the laser spot irradiance distribution near the focus of the data marks. Energy falls off rapidly with defocus Δz, and the spot has a finite width s. Multiple thin-film layers are used to enhance recording and readout characteristics of CD-RW media

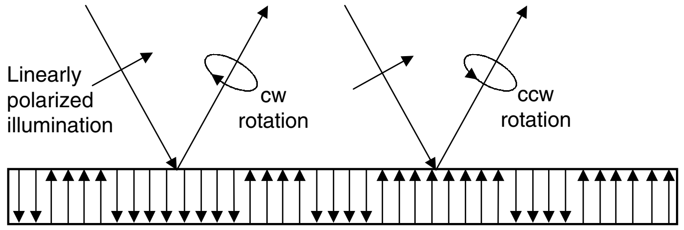
through the clear substrate material and into the thin-film recording layers. The first layer is a transparent dielectric. The second layer is the recording material, and the third layer is another layer of transparent dielectric. The fourth layer is a reflector. The thicknesses of these four layers are designed to tune the reflective properties for maximum signal contrast, and they are also adjusted to provide adequate absorption so that a reasonable amount of laser power can be used for writing [17]. A protective coating and a label and ink layer separate the thin films from the environment.

Magneto-optic (MO) products store information in small magnetic data marks, which are about the same size as pits on a CD. The recording layer is initially erased so that all magnetic domains are aligned in one direction perpendicular to the recording surface, as shown in Fig. 9(a). In this configuration, the magnetic domains are extremely stable. A large magnetic field of several thousand oersteds is required to overcome the magnetic moment of the domains. The magnetic field required to reorient domains is called the *coercivity* [18]. To record data marks, a high-power focused spot locally heats the recording surface. Heat reduces coercivity and magnetic domains in the region of the focus spot are reoriented with an external magnetic field. When the laser beam is switched to low power between data marks, the recording layer is not heated, and the

OE011



(a) MO recording



(b) MO readout

Fig. 9 (a) Magneto-optic recording involves heating the recording layer with the focused laser spot in order to reduce the layer coercivity. Domains in this small region can then be flipped with an externally applied magnetic field in order to form the data-mark pattern. Once the bits cool, they are frozen in place until heated again. (b) During readout, a low-power focused spot illuminates the data marks. Linearly polarized light from the laser is rotated either clockwise or counterclockwise, depending on domain orientation. The data detector senses the polarization change and converts this information into a current signal

external magnetic field has no effect on domain orientation. As described in Fig. 3, the laser beam is modulated between high power and low power as the disk spins in order to write a pattern of data marks along each track. Each mark contains magnetic domains oriented in the opposite direction compared to the magnetic domains of the background [19]. The marks have the property that, as a low-power focused light spot passes over it, polarization of the reflected light is rotated, as shown in Fig. 9(b). Polarization rotation on reflection is due to the polar Kerr effect [20]. When the laser beam illuminates a data mark with domains oriented away from the laser beam, linear incident polarization is rotated slightly in the counterclockwise direction. When the laser beam illuminates the region between data marks, linear incident polarization is rotated slightly in the clockwise direction. In order to detect the data signal, a detector is used to sense change in polarization of the reflected light. (Differential detection is the most common detector geometry for MO media. The reflected light is directed through a polarizing beam splitter, where the surface of the beam splitter is oriented to split light equally for unrotated light. Clockwise or counterclockwise rotation of the polarization imbalances the split. The difference between light levels after the split can easily indicate the change in polarization.) For example, an indication that the reflected light is rotated in the counterclockwise direction implies that the laser spot illuminates a data mark. In order to erase data, the external magnetic field in Fig. 9(a) is reversed, and the laser beam heats an entire section of the track. A major difference between CD and MO products is that the MO marks are produced in a track with an almost undetectable change in the topology of the track. That is, there is almost no mechanical deformation of the track as the marks are recorded or erased. This property enables MO products to exhibit over one million erase cycles with little if any degradation in performance [21, 22].

A collection of the available CD and CD-like formats are listed in Table 3, along with their associated data-mark technology [23]. DVDs used for movie or data distribution are pit-type ROM disks. The track pitch and pit size is smaller than in CDs, as shown in Fig. 2. DVD products can also be erasable, and there are a multitude of formats available, as listed in Table 4. Erasable DVD products use phase-change technology. DVD-R products, like CD-R, use a write-once dye-polymer recording layer. Unlike CDs, DVDs can use more than one storage layer per disk. DVDs can be double sided, use two layers on one side, or use two layers on each side. Adding layers increases the total capacity of the disk. Newer technologies, like the Blu-ray Disc (BD) described in Sect. 7, are introduced into the marketplace with erasable phase-change technology.

5
Technology

Several important aspects of optical data storage technology are associated with the optical-mechanical-electrical system that is used to write and read data to the disk. This section reviews basic concepts necessary to understand how these systems work.

5.1
Data Density and Spot Size

Capacity of an optical disk is determined by its *data density*, which is the number of bits of information stored per unit area on the surface, and the recording area. Data density is often specified in gigabits (10^9 bits) per square inch of recording surface area (Gb in^{-2}). For example, a 0.65-gigabyte (GB) CD has a recording area of about 14.5 in^2, so the data density is $(0.65)(8)/14.5 = 0.36$ Gb in^{-2}, where one byte = 8 bits.

A fundamental limitation to the data density is due to the size of the focused laser beam that illuminates the surface. Figure 8 shows a detailed picture of the laser irradiance approaching the surface, where irradiance is defined as the laser power per unit area. Ideally, maximum irradiance is located at the recording material, along with the smallest spot size s. As the distance increases away from the ideal focus, the spot size increases and the peak irradiance decreases. A defocus distance Δz of only 3 µm dramatically reduces peak irradiance and increases spot size. An approximate formula used to estimate the ideal spot size is $s = \lambda/(\sin\theta)$, where θ is the marginal ray angle of the illumination optics, as shown in Fig. 3. Spot size s is the full width of the irradiance distribution at the $1/e^2$ (13.5%) irradiance level relative to the peak. The value of $\sin\theta$ is often called the *numerical aperture* or NA of the optical system. CD systems exhibit $\lambda = 0.78$ µm and $NA = 0.47$, which produce a spot size of 1.7 µm.

OE011

Tab. 3 CD formats

Format	*Data-mark technology*	*Characteristics*
CD-ROM	Pit	Computer storage medium with the capacity of up to 700 MB
CD-R	Dye polymer	Write-once computer and audio storage >700 MB
CD-RW	Erasable phase change	Erasble >700 MB storage with >1000 erase cycles
CD+G (CD + Graphics)	Pit	Audio CD plus additional graphics and/or text information recorded in R-W subcodes
CD+Midi	Pit	Audio CD plus MIDI music information in the R-W subcodes to enable sounds from CD to be revoiced or remixed through the MIDI compatible device.
CD-3 (3″CD)	Pit	Audio CD with the smaller diameter – (3″) with playing time reduced to 20 min. Also called *CD-single*
CD-A (CD Audio)	Pit	Original, audio compact disc, containing up to 74 minutes (suggested by "Red Book") of stereo digital audio along with 8 subcode tracks labeled P-W
CD-E (Erasable)	Erasable phase change	Audio or data CD which can be recorded and erased many times
CD-EG (CD Extended Graphics)	Pit	Enhancement to the CD+G format adding 256 colors and instant mix of two pictures
CD-I (CD-Interactive)	Pit	Extension of the CD-ROM format aimed specifically to the consumer market. System offers high-resolution graphics, still and (recently) moving pictures, and stereo sound. The CD-I player also plays back CD-A, CD+G and Photo CD.
CD-I ready	Pit	A CD-A that contains additional data "Hidden" in a space before the first track. Loaded into a CD-I player the disc offers many features of full CD-I.
CD-ROM XA (Extended Architecture)	Pit	Development of the CD-ROM that has been designed to meet the multimedia and interactive needs. Audio, graphics, and (some) video information have been added to original CD-ROM format.
CD-V (CD Video)	Pit	Also known as the LaserDisc containing up to 2 h of analog video and digital audio information. There are three formats 12″, 8″, and 5″.
CD-V single	Pit	A version of the CD-V containing up to 5 min of video plus 20 min of CD audio only.
CD-WORM (Write Once, Read Many)	Dye polymer	Audio CD allowing direct recording (only once) of musical information. Used also for storing of computer data or copying of CD-ROMs.
CDTV	Pit	A version of the CD-ROM discs written in the AMIGA language with the interactive and graphics possibilities.

Tab. 3 *(Continued)*

Format	***Data-mark technology***	***Characteristics***
MINI DISC	Magneto-optical	A SONY 2″ rewritable magneto-optical disc that is using data reduction system to record up to 74 min of audio information (possible use for computer data storage).
Photo CD	Dye-polymer	A KODAK development that is using a CD-WORM system to record up to 100 still pictures. (The player is compatible with the CD-I system). To be able to read these disks, CD Player must be "multisession".
Video CD	Pit	CD using MPEG-1 encoding process to compress video (including feature films) on CD. The picture quality is higher than in standard CD-ROMs, but additional hardware (•MPEG encoder) is necessary.

Q2

Tab. 4 DVD formats

Format	***Data-mark technology***	***Characteristics***
DVD-5	Pit	Single-layer 4.7 GB read-only DVD.
DVD-9	Pit	Dual-layer 7.95-GB read-only DVD. Both layers are read from one side of the disk.
DVD-10	Pit	Double-sided 8.7-GB DVD. Must turn disk over to read second side.
DVD-18	Pit	Dual-layer, double-sided 17.1-GB read-only DVD. Can read two layers from each side.
DVD Video	Pit	One of DVD-5 through DVD-18 with MPEG-1 or MPEG-2 video files, audio, subpictures, and navigation data.
DVD Audio	Pit	DVD-5 with high-quality audio files.
DVD-ROM	Pit	One of DVD-5 through DVD-18 with computer-friendly file formats.
DVD-R	Dye-polymer	Write-once 4.7 GB/side
DVD+R	Dye-polymer	Similar to DVD-R, except designed to be compatible with DVD+RW
DVD-RAM	Erasable phase change	Erasable computer-friendly random access with 4.7 GB/side. Number of erase cycles >100 000. Not compatible with all players.
DVD-RW	Erasable phase change	Erasable with better compatibility than DVD-RAM and >1000 erase cycles.
DVD+RW	Erasable phase change	Similar to DVD-RW, except designed to be compatible with DVD-ROM and DVD Video players
DVD-VCD	Pit	CD-V authored on a DVDR/W. Audio has to be resampled to 48 kHz.
DVD-SVCD	Pit	SVCD authored on a DVDR/W. Higher quality video than DVD-VCD. Audio has to be resampled to 48 kHz like the DVD-VCD.
DVD-MP3	Erasable phase change	MP3s burned on a DVDR/W.
MiniDVD	Pit	DVD format on a CD-R(W) instead of a DVD disc. MiniDVD is also sometimes called *cDVD*. A miniDVD only fits about 15 min video on a 650-MB CD-R(W). This is not a supported format.

DVD systems exhibit $\lambda = 0.65\,\mu\text{m}$ and $NA = 0.60$, which produce a spot size of 1.1 µm. BD systems use $\lambda = 0.405\,\mu\text{m}$ and $NA = 0.85$, which produce a spot size of 0.48 µm.

5.2
Thermal Recording

In order to write data onto the spinning disk, the laser must be pulsed to a high-power level. The time duration of the high-power pulse determines the length of the data mark that is written onto the surface. Laser writing is possible because the medium is thermally sensitive, that is, the medium exhibits a thermal threshold [8]. Below the threshold, medium properties do not change significantly. Above the threshold, a physical change occurs in the medium.

Figure 10 shows lines of constant temperature, which are called *isotherms*, generated on an aluminum surface for a 200-ns (200×10^{-9} s) focused laser pulse. The surface is moving at 10 ms^{-1}, so the isotherms are spread out along the scan direction. This profile is representative of those found in DVD optical disks. Notice that the end of the pulse generates a wider isotherm than at the beginning of the pulse, due to the fact that heat builds up and spreads out in the direction perpendicular to the scan. This effect is called *thermal blooming*, and is a serious problem, especially in magneto-optic systems, if not corrected by varying the properties of the laser pulse [24]. For example, a corrected laser pulse might have a higher power level at the beginning of the pulse than at the end of the pulse. Figure 10 indicates that, if the threshold temperature of the medium is equal to the 200 °C isotherm, a data mark of approximately 0.6 µm wide and 2.5 µm long will be written at this location on the surface of the disk.

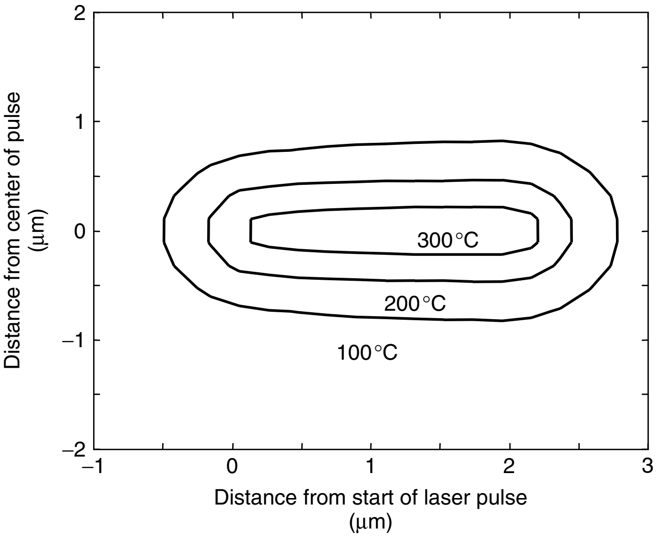


Fig. 10 Since optical disk media are sensitive to a thermal threshold, simulated thermal contours of a laser spot scanning the recording surface show isotherms that predict the mark size and shape

5.3
Frequency Response and Equalization

The ultimate limit to the size of the data marks on the recording surface is determined by the frequency response of the optical system. S*patial frequency* is $1/T$, where T is the period of the data-mark pattern. As the period decreases, spatial frequency increases. The frequency response is understood simply by recognizing the behavior of the reflected light and how the reflected light is collected by the objective lens. For example, Fig. 11 shows the reflected light distribution for a periodic pattern of data marks along a track. The reflected light consists of three cones. The direct reflection is the central cone. The two outer cones are called *diffracted orders*. They are very similar to the central cone in appearance, but they are spread apart by angle ψ. As T decreases, ψ increases, and the diffracted orders spread more widely apart. ψ is also directly proportional to the laser wavelength. Shorter-wavelength lasers exhibit smaller ψ.

When the spot scans over data marks, the optical phase of each diffracted order changes, but the phase of the central cone does not change. The phase difference between the diffracted orders and the central cone produces a modulation in the overlap area because of interference; that is, as

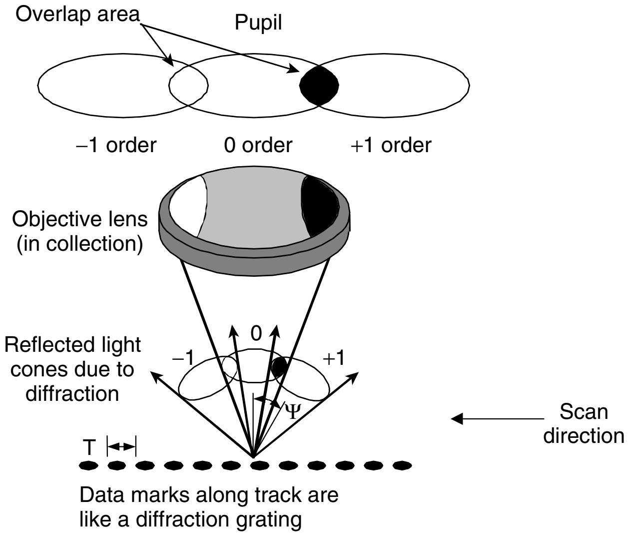


Fig. 11 The data-mark pattern reflects light in a diffraction pattern consisting of three primary cones, which are the zero and ± first diffraction orders. As the data pattern moves under the laser spot, relative phases of the ± first diffracted orders change with respect to the zero order. In the overlap area, the phase difference produces an interference effect that modulates the irradiance level. This light modulation is then converted into a current signal by the detectors. The amount of overlap area, and hence the amplitude of the data signal, depends on the spatial frequency $1/T$ of the data-mark pattern. Higher-frequency patterns produce smaller overlap area

the spot scans over data marks, the overlap areas get brighter or darker as a function of the relative position between the spot and each mark. Brightness of the central cone does not vary. Therefore, the contrast of the signal modulation received at the detectors is determined by the amount of overlap area. More overlap area produces a higher contrast data signal. As T decreases, so does the overlap area. At some critical mark period, there is no overlap and, consequently, no signal modulation at the detector. This critical mark period is called the *resolution limit*, T_R, of the optical system. A numerical value for T_R is found from

$$T_R = \frac{\lambda}{2NA} = \frac{s}{2},$$

where NA is the numerical aperture of the objective lens and s is the spot size.

Figure 12 shows the *modulation transfer function* for the optical system, which measures the contrast of the current signal modulation versus the spatial frequency of the data-mark pattern. Maximum modulation is observed for long marks. The contrast reduces gradually to the resolution limit. The maximum mark period in practical devices is well above the resolution limit T_R. For example, the shortest mark period in CDs is about 1.8 μm, which is about a factor of 2 longer than the resolution limit of $T_R = 0.78/(2 \times 0.47) = 0.83$ μm. Of course, real data patterns are more complicated than simple periodic patterns, but each real data pattern can be decomposed into a collection of weighted periodic patterns. Therefore, the modulation transfer function is also useful in describing system behavior for real data patterns.

When the real data pattern contains both high-frequency and low-frequency components, a significant contrast difference exists in the current signal. These differences in contrast make detecting signal transitions difficult. In order to minimize the contrast difference, electronic circuits are often employed during signal amplification, as shown in Fig. 4. The electronic

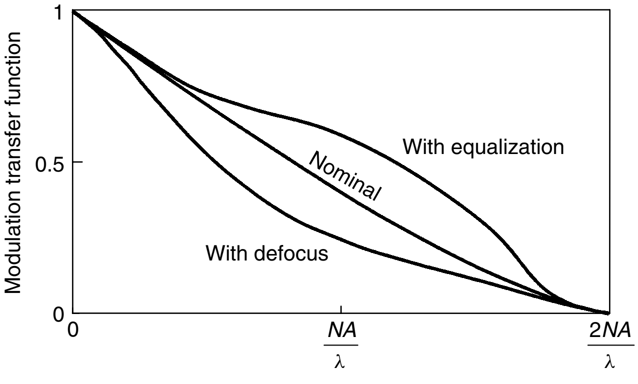


Fig. 12 The modulation transfer function is a plot of signal contrast versus spatial frequency of the data-mark pattern. Defocus blurs the laser spot and reduces contrast, especially in the midfrequency range. Equalization circuits can be used to boost the transfer function after the light modulation has been converted into a current signal

circuits partially *equalize* the modulation transfer function and provide more reliable signal decoding [25]. An example of an equalized modulation transfer function is shown in Fig. 12, where the high-frequency contrast is boosted with respect to the low-frequency contrast. Unfortunately, physical limitations of electronic circuitry and noise considerations do not permit ideal equalization, which exhibits uniform contrast for all spatial frequencies out to the resolution limit.

5.4
Effects of Defocus

An example of the effects of defocus is shown in Fig. 13, where irradiance of a DVD-like system with $NA = 0.6$ and $\lambda = 0.65$ µm is plotted for several values of Δz. In focus at $\Delta z = 0$, the spot is well confined and only a small fraction of the spot energy is contained in the diffraction rings surrounding the central lobe. At $\Delta z = 0.5$ µm, the peak irradiance is reduced slightly and a small amount of energy is shifted to the first diffraction ring closest to the central lobe. At $\Delta z = 0.5$ µm, changes in the spot shape dramatically affect device performance. However, as defocus increases beyond $\Delta z = 0.5$ µm, peak irradiance degrades rapidly and a significant amount of energy is spread into the diffraction rings. An estimate of the allowable *depth of focus* is $\Delta z = \pm\lambda n/(4\sin^2\theta) = \pm\lambda n/(4NA^2)$, where n is the refractive index of the disk substrate. For example, with $n = 1.5$, $\Delta z = \pm 0.67$ µm. The effect of defocus on an unequalized modulation transfer function is shown in Fig. 12, where the midfrequency response of the system suffers severe degradation.

5.5
Servo Optics

The size of the focused spot is very small in the direction along the tracks, which allows many data marks to be written for each revolution of the disk. Since the

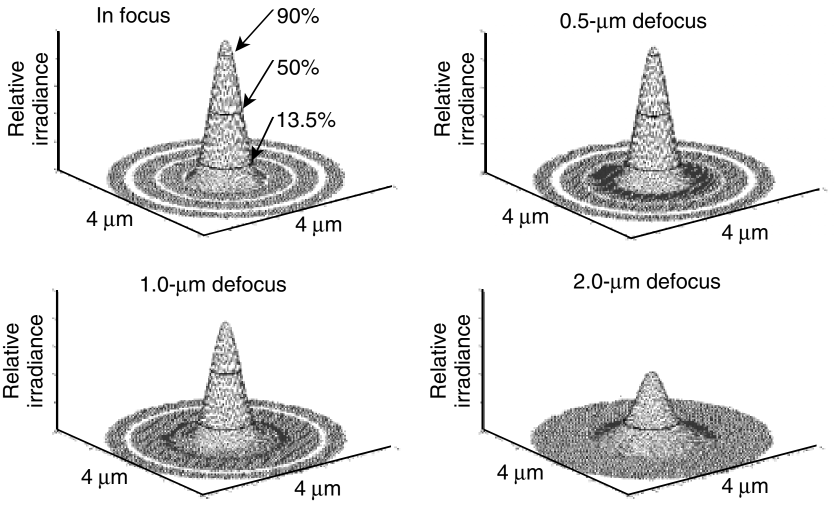


Fig. 13 The relative irradiance distribution of a focused DVD-like laser spot is shown with different amounts of defocus. More than 0.5 µm of defocus significantly degrades the peak irradiance and spot quality

light spot is also small in the direction perpendicular to the track, tracks are spaced closely together. In CD systems, the track pitch, where pitch is defined as the center-to-center track spacing in the radial direction, is typically 1.6 µm. In DVD systems, track pitch is 0.74 µm. Data-mark width is typically less than one-half the track pitch in order to reduce the effects of cross talk from marks on adjacent tracks. The optical spot must be centered over the marks as the disk spins in order to obtain maximum signal amplitude at the decoding electronics. A typical requirement is that the spot must be kept on track center to better than 1/10th the track pitch, or 0.16 µm for CD systems and 0.07 µm for DVD systems. In the focus direction, the spot must be controlled to better than 1/10th the depth of focus, which is about 0.25 µm for CDs and 0.13 µm for DVDs [26, 27].

This demanding control of the spot center and focus position is complicated by the fact that the optical disk and the electric motor that rotates the disk suffer from loose tolerances that induce large variations in the track position as the disk spins [28]. For example, thicknesses of CDs vary by more than 50 µm. Registration errors during the molding process typically offset the center of the track radii by more than 30 µm from the center of the disk rotation. Wobble of the motor shaft induces variations in the focus position by several 100 µm.

In commercial optical data storage systems, position control is accomplished with closed-loop feedback servos. A basic diagram that illustrates the servo technique is shown in Fig. 14. The difference between the desired spot position and an error signal that is derived from the actual spot position is amplified with some gain *G* and used as input to an actuator H [29]. The actuator is usually a mechanical device, like a voice coil, which moves an optical element that, in turn, repositions the spot in either the focus direction or across the tracks. The feedback sensor instantly determines the spot position, and the new information is fed back into the control loop.

Both focus and tracking actuators are usually combined into one mechanical unit that moves the objective lens. A photograph of an actuator assembly from a commercial CD player is shown in Fig. 15. The objective lens is mounted in a suspension that has a range of motion of a few millimeters in both the focus and tracking directions. The flexure of the suspension is very stiff with respect

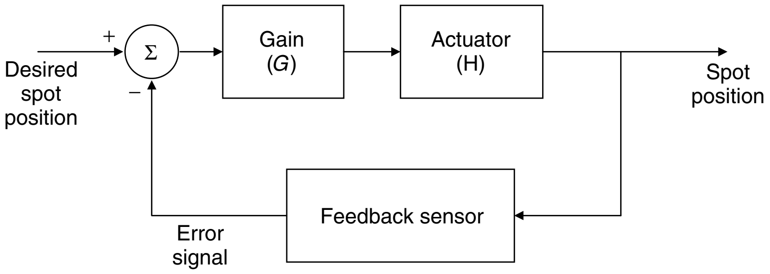


Fig. 14 A basic diagram of the servo loop used in optical storage devices shows the gain (*G*) of the drive electronics producing current for the actuator (H), which positions the lens over the data track. A feedback sensor provides an error signal for robust control of spot position

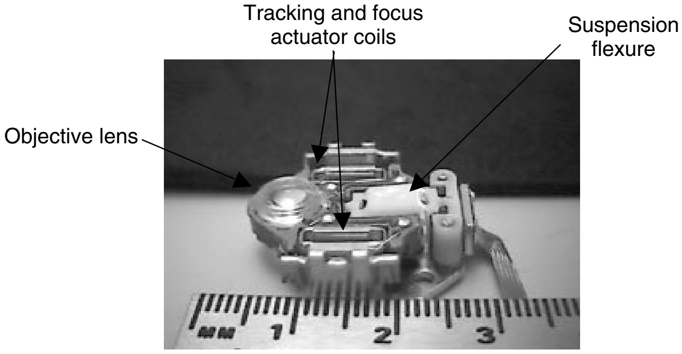


Fig. 15 A photo of a commercial CD actuator assembly that illustrates the objective lens, suspension flexure, and the coils used to move the suspension. The scale in the figure shows in millimeters the distance between the smallest adjacent lines. The entire lens and actuator assembly is less than 30 mm long, excluding the electrical connections

to motion in any other direction. The fixed part of the suspension also has permanent magnets mounted on it that are aligned with electric coils on the moving part of the actuator assembly. As electric current is passed through the coils, the induced magnetic field presents a force on the permanent magnets and moves the suspension. If the tracking coils are activated, the suspension moves in the cross-track direction. If the focus coils are activated, the suspension moves in the focus direction. By moving the suspended objective lens, position of the focus spot on the disk is changed. When the actuators are combined with a servo loop, accurate control of the spot position is possible, even in extreme environments like those found in portable disk players.

5.6
Feedback Sensors

An important part of the servo loop pictured in Fig. 14 is the feedback sensor. In fact, it is not possible to control the spot position better than the sensor can detect position errors. Usually, separate tracking and focus sensors are implemented in optical data storage devices. The error signal generated from the tracking sensor is called the *tracking error signal* (TES), and the error signal generated from the focus sensor is called the *focus error signal* (FES) [7]. The following TES and FES sensors are described:

Push–pull tracking is a method to provide a TES using a groove pattern on the disk, where the period of the grooves is equal to the track pitch.

Three-spot tracking is a method to provide a TES by detecting the signals from three spots focused onto the disk, where the central spot is centered over a track and two neighboring spots are slightly each side of the central spot in the direction across the track●. Q5

Astigmatic focusing is a method to provide a FES by fabricating a small amount of astigmatism into the servo optics.

Generation of a TES signal using grooves is shown in Fig. 16. In addition to the data marks, each track contains a land and groove area. Data marks

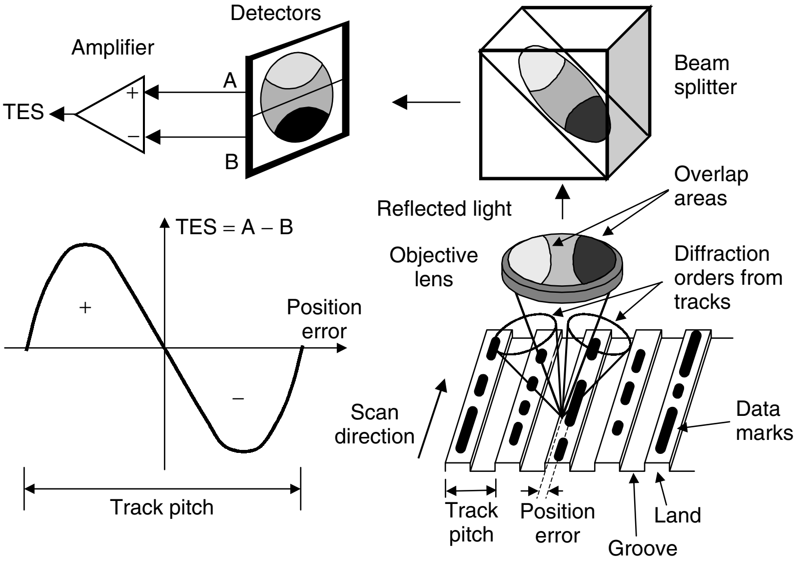


Fig. 16 The push–pull tracking error signal (TES) is generated by using a slit-cell servo detector and sensing the difference in light level between the cells. Since grooves of the disk diffract light like a grating, the diffracted orders overlap. In this case, spatial frequency of the grating is fixed, and the diffraction occurs in an orthogonal direction compared to diffraction from data marks. As the laser spot moves off track, the relative phase change in the diffracted orders produces bright and dark patterns on the detector. The TES difference signal indicates the relative off-track location of the laser spot

are usually written in the land areas that are closer to the objective lens than the grooves. The collection of lands and grooves forms a diffraction grating in the cross-track direction. As with single-frequency data-mark patterns, the land/groove diffraction grating produces separated cones in the reflected light because of diffraction. When the spot is centered on the groove, the phase of each diffracted order is equal, so the overlap areas are of equal brightness. When the spot is off center, the phase of the diffracted orders changes, and brightness of the overlap regions become unbalanced. This brightness asymmetry is detected with a split-cell detector. Current signals A and B are subtracted to form the TES. When the spot moves in one direction off center, the total power on detector A becomes brighter than the total power on detector B. The TES signal is positive. As the spot continues to move in the same direction, the detector signals become more unbalanced, which creates a more positive TES. If the spot moves in the opposite direction, the detectors become unbalanced in the opposite sense, which creates a negative TES. Near the center of the track, the TES is linear and provides a good quality feedback signal that is directly proportional to the position error. The TES is periodic with a period equal to the track pitch. This type of TES signal is called *push–pull tracking*, which is descriptive of the light-pattern behavior on the detectors as the spot moves off •track center [30]. Q6
Push–pull tracking is often used in CD-R

and CD-RW players, where some form of tracking reference is necessary before data can be written.

Both regular data-mark patterns and grooves produce diffracted orders in the reflected light. These orders overlap at the objective lens, as shown in Fig. 17. Diffracted orders from the data-mark patterns spread in the direction parallel to the scan direction, and the amount of spread is proportional to the data-mark frequency. Diffracted orders from the grooves spread in the direction perpendicular to the scan direction, and their separation is constant. Modulation is observed in the overlap areas between the data-mark orders and the objective lens as the spot scans along the track. In addition to the modulation because of the data pattern, brightness changes can be observed in the overlap areas between the groove orders and the objective lens as the spot moves off track.

A second method used to generate a TES is shown in Fig. 18, which is called *three-spot tracking* [31]. In addition to a central laser spot, two additional laser spots are generated by the illumination optics. These three spots are imaged onto separate detectors A, B, and C by the servo optics. The leading spot is imaged onto detector B, and the trailing spot is imaged onto detector A. The central spot, which is used to detect the data signal, is imaged onto detector C. The brightness of the spots at the detectors is determined by the amount of overlap between the spot and the data marks. When the spot is centered over a data mark, its corresponding light level at the detectors is reduced the most. When the data spot is centered over a track, the leading and trailing spots are slightly offset from the center in opposite directions and by an equal amount. Their brightness at the detectors is equal, and the difference between detector signals A and B is zero. When the spots are slightly off track, as shown in Fig. 18, the trailing spot is now overlapping less of the data

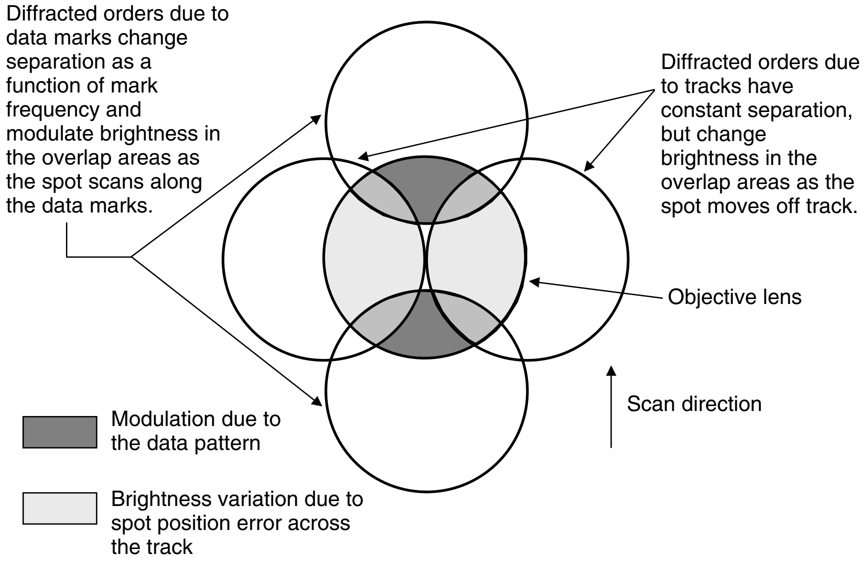


Fig. 17 Diffracted orders from the disk contain information about the data-mark pattern and the grooves

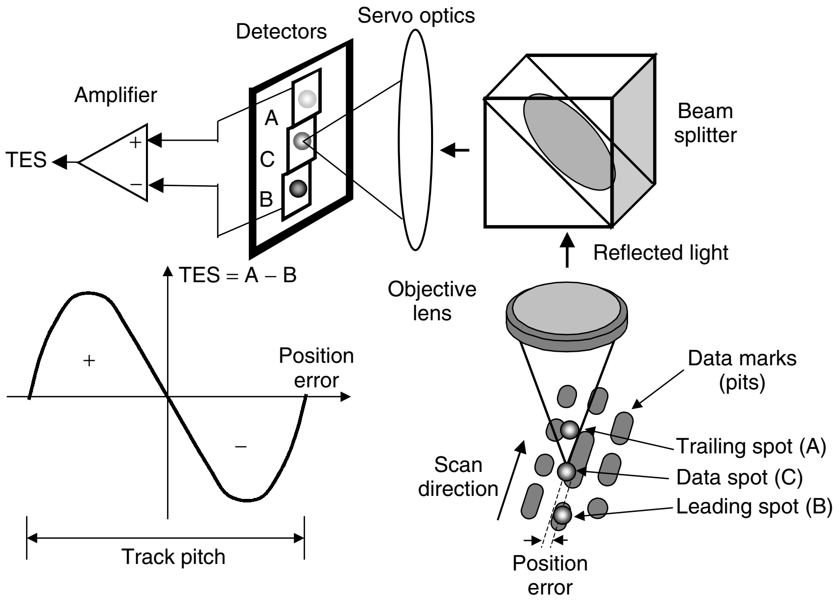


Fig. 18 Three-spot tracking uses two auxiliary laser spots that ride edges of the track as the disk spins. The spots are reimaged onto separate detector elements. If the data track is not centered, the amount of light reflected from each auxiliary spot changes. One detector spot becomes brighter and the other dims. A difference signal produces a reliable TES

mark and the leading spot is overlapping more of the data mark. Therefore, the brightness on detector A increases, and the brightness on detector B decreases. The difference between detector currents A and B is now positive. When the spot on the disk shifts in the opposite direction, the difference current becomes negative. The TES is generated from the difference between the currents from detectors A and B, and provides a good quality feedback signal that is directly proportional to the position error in the center of the track. As in the case of push–pull tracking, the TES for three-spot tracking is periodic with a period equal to the track pitch. Unlike the push–pull technique, three-spot tracking requires that the spots be reimaged onto the detectors. Three-spot tracking is often used in music CD players, where there is no land and groove pattern and the TES must be generated from only the data marks.

The popular *astigmatic focus* method to generate an FES is shown in Fig. 19 [32]. The reflected light is directed into the servo lens, which affects light on the detectors in a special way, that is, the light spot on the detector plane changes shape as a function of the disk defocuses. When the disk is too close to the objective lens, the light spot elongates along the right diagonal on detector quadrants A and C. When the disk is in focus, the light spot is circular. When the disk is too far from the objective lens, the light spot elongates along the left diagonal on detector quadrants B and D. Summing diagonal quadrants and then subtracting the results creates the FES signal. If the

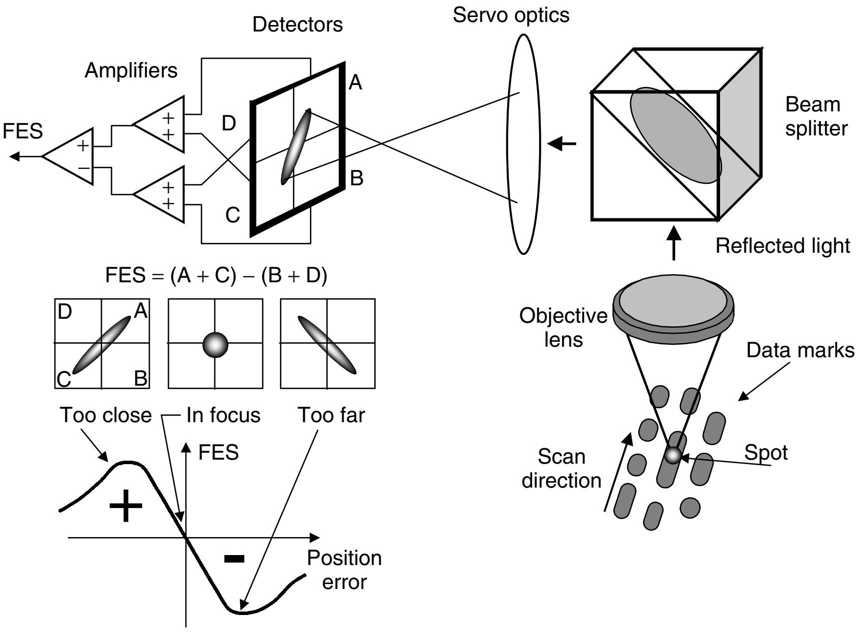


Fig. 19 The astigmatic focusing technique uses a special lens in the servo optics before a quadrant-cell detector. The lens introduces a small amount of astigmatism into the beam along a diagonal direction on the detector. As the spinning disk goes into and out of focus, the astigmatism forces the light spot to change shape. A difference signal from the detector quadrants produces an FES that indicates the amount of defocus

disk is too close to the objective lens, the FES is positive. When the disk is in focus, the FES is zero. If the disk is too far from the objective lens, the FES is negative. Near the focus condition, the FES is nearly linear and provides a good quality feedback signal for the servo loop. The elongated spot behavior is due to a small amount of astigmatism fabricated into the servo lens. Astigmatism is a difference in the focusing power in diagonal directions, and is similar to astigmatism that commonly occurs in the eye.

5.7
Noise and Jitter

As shown in Fig. 12, the zero of the modulation transfer function determines the resolution limit of the optical system. In practice, it is not possible to obtain this limit due to noise. Noise limits the ability of the detection electronics to determine the proper bit pattern from the detector signal. Sources of noise include reflectivity variations across the disk, photon noise in the laser beam, detector noise, and other sources [33, 34].

A detector signal with noise is displayed in Fig. 20 for a large spot and a small spot. The same amount of random noise is assumed for both signals. In order to detect the bit pattern, a threshold level is established on the basis of the signal amplitudes. When the signal level falls below the threshold, a *transition* occurs. The *data bit* value changes from 1 to 0 at the transition marked in Fig. 20, where

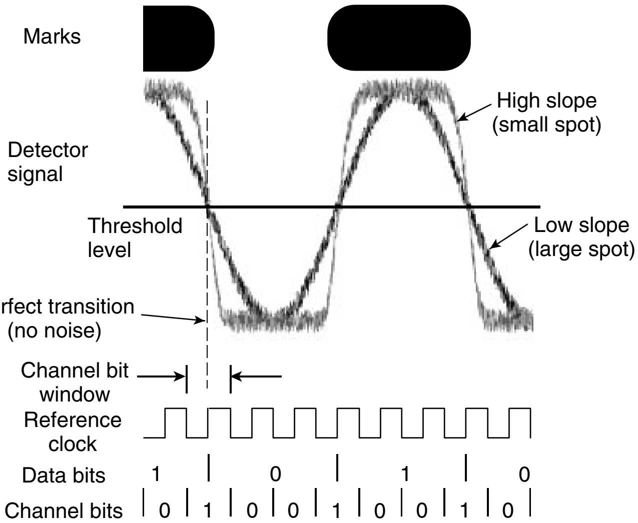


Fig. 20 Detector signals with noise are shown for scanning data marks with two spot sizes. The width of the channel-bit window is determined by the reliability of detecting a transition across a threshold level. Transitions determine the positions of channel-bit 1s in the data pattern, so data marks and spaces can each represent more than one channel bit

there is one data bit for each mark and one data bit for each space between marks. This minimum mark length is a function of the spot size *s*, and is generally found to be 0.6 *s*. Marks and spaces can be longer than the minimum mark length, but they are not shorter.

In very simple recording schemes, the data bit represents the desired output data stream. In practice, the transition signals the change of a *channel bit* value, where more than one channel bit is present for each data bit [35, 36]. It is the channel bits that determine the output data stream and the data density. Each channel bit is defined by a *channel bit window*. Size of the window is determined by the minimum window width before noise degrades the reliability of detecting the transition.

A magnified portion of the transition region is shown in Fig. 21. The amplitude noise σ_N on the signal creates uncertainty σ_W in the position of the transition. A detection error occurs if the noise shifts the transition from the ideal window into a neighboring window. The window size is usually specified so that uncertainty σ_W produces no more than one error per 10^4 transitions. Variation of the transition within the timing window is called *jitter*. As shown in Fig. 21, the amount of width variation (or jitter) is a function of signal slope. That is, $\sigma_W = \sigma_N/m$, where m is the signal slope [37]. Small spots yield high slope, small σ_W and short channel-bit windows.

5.8
Data Coding and Formatting

The different ways of organizing 1 s and 0 s on the disk are called *formats* [38]. There are several different formats in use today, as illustrated in Tables 3 and 4, with new

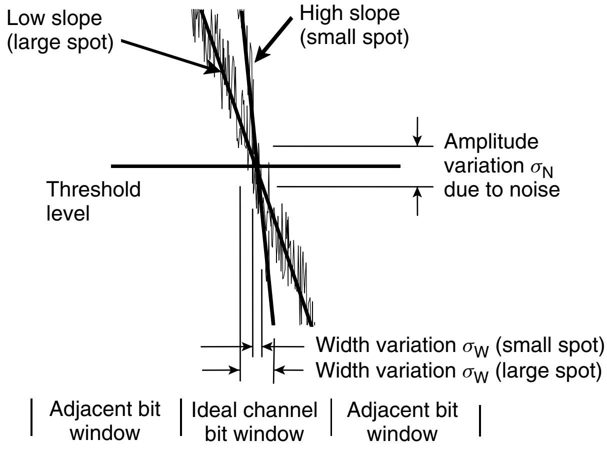


Fig. 21 Small spots produce high slopes in the transition region, and the signal is less affected by noise as compared to the system with the larger spot. Therefore, laser systems that generate smaller spots can pack more channel bits into the data-mark sequence

ones being invented all the time. Some are more popular than others; some require special drives to access them, while others are compatible with each other to some degree. This section describes the CD-DA (digital audio) format. Details of other formats for CDs are available through the standards set by the industry in a "rainbow" of reference books [39]. That is, a particular book "color" corresponds to a particular standard. For example, CD-ROM format follows the "Yellow Book," CD-DA follows the "Red Book," and CD-R and CD-RW follow the "Orange Book." These standards are important because they insure interchangeability between different players. The differences between CD formats are only in the system information that is written with the data, which amounts to a very small portion of the total pattern. All CD formats share the same basic structure.

The set of rules used to convert user data bits into their physical data-mark representation and back again are called *channel codes*. The channel code for CD-ROM is called *eight-to-fourteen modulation* (EFM) [35, 36]. EFM interprets user's data along with error correction data, address data, synchronization data, and other content into the stream of channel bits recorded in the data-mark pattern. The minimum number of 0 s between logical 1s is set by the jitter requirement, as explained in Sect. 5.7•, and the maximum number of 0s is set by the need to provide a synchronization signal for the reference clock shown in Fig. 20. The conversion of an eight-bit user byte under these restrictions leads to a fourteen-bit channel sequence, from which this code scheme derives its name. A fourteen-bit channel sequence is called a *symbol*. For example, the user byte 00000000 is encoded as 01001000100000, and the user byte 00011111 is encoded as 00100000010000. During readout, the EFM decoder of the CD-ROM works in the opposite direction, as shown in Fig. 4, converting the current signal into a binary data stream, which is

Q7

then cleared of any miscellaneous data by the drive's electronics. An example of an EFM sequence is shown in Fig. 22, where there are a minimum of 2 zeros following each transition and a maximum of 10 zeros following each transition.

5.9
Configurations for Optical Media

Optical media are produced in several different configurations. Figure 23 displays four configurations that are in commercial use or have been tested in laboratories. The most common configuration is the single-layer disk, like the CD, where data are recorded in a single-storage layer.

In order to increase data capacity of the disk, several layers are used. Each layer is partially transmitting, which allows a portion of the light to penetrate throughout the thickness of the layers. The scan spot is adjusted by refocusing the illumination optics so that only one layer is read out at a time. Some of the DVD formats in Table 4 use two layers on one side of the disk.

Data are also recorded in volumetric configurations [40, 41]. Like with the multiple-layer disk, the scan spot is refocused throughout the volume of material to access information. Volumetric configurations offer the highest efficiency for data capacity, but they are not easily paired with simple illumination optics.

The final configuration is to place the information on a flexible surface, like ribbon or tape [42, 43]. Like magnetic tape, the ribbon is pulled under the scan spot and data are recorded or retrieved. Flexible media have about the same capacity efficiency as volumetric storage. The advantage of a flexible medium over a volumetric medium is that no refocusing is necessary. The disadvantage is that a moderately complicated mechanical system must be used to move the ribbon.

5.10
Laser Sources

The semiconductor laser diode is a key technology element in establishing the optical data storage industry. Although both the communications industry and the optical data storage industry use laser diodes, the latter consumes orders of magnitude more diodes than the former. Each CD player on the market uses an AlGaAs laser diode that operates with a wavelength around 0.780 μm. These small light sources are important because they

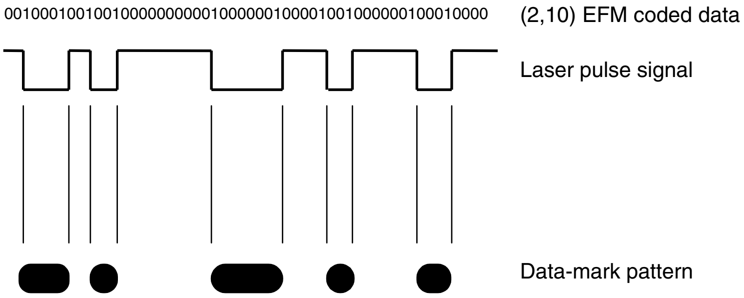


Fig. 22 An eight-to-fourteen (EFM) modulation code produces a laser pulse signal and data-mark pattern that exhibits a specific number of zeros between each transition

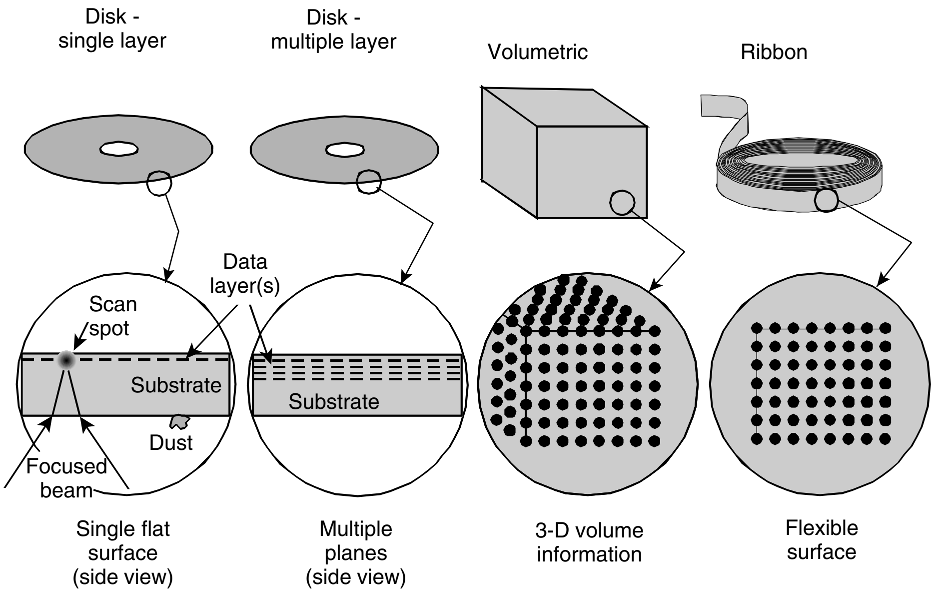


Fig. 23 Four possible configurations of optical storage media include a single-layer substrate, like a CD, a multiple-layer substrate, like a DVD, volumetric configurations, like holographic and bit-wise storage, and ribbon (tape) media

can emit a relatively bright beam, they are reliable, and they can be directly modulated with simple electronics.

A drawing illustrating the most basic structural attributes of a CD laser diode is shown in Fig. 24. It is a double heterostructure device, whose GaAs active layer, where injected electrons and holes combine to produce photons, is bordered by *p*- and *n*-doped layers of GaAlAs that confine the carriers to the active region. The layers are grown epitaxially on an *n*-doped GaAs substrate. A 2- to 3-μm wide metal contact on the top of the chip provides a path for the injected current. The circuit is closed by the metal contact on the bottom of the highly conducting substrate. The dimensions of a typical laser diode chip are of length 300 to 400 μm and height and width 100 μm. The refractive index of GaAs is higher than GaAlAs, so that the structure also serves as a waveguide, confining the electromagnetic field produced by the coherent stimulated emission of photons in the vertical dimension to about 1 μm. In the lateral direction, the confinement of the field is effected by the presence of the injected electrons, which raises the refractive index. (This is called *gain guiding*.) In more complex structures, the lateral confinement of the field is effected by a more complex epitaxial growth and etching process that yields a narrow stripe of GaAs bounded laterally by epitaxially grown material of a lower refractive index. (This is called *index guiding*.)

The cleaved facets of the laser diode act as mirrors, reflecting about 30% of the light generated in the active stripe back into it. Without this feedback, laser action is not possible. The transverse dimensions of the active stripe waveguide region are such that the waveguide supports only

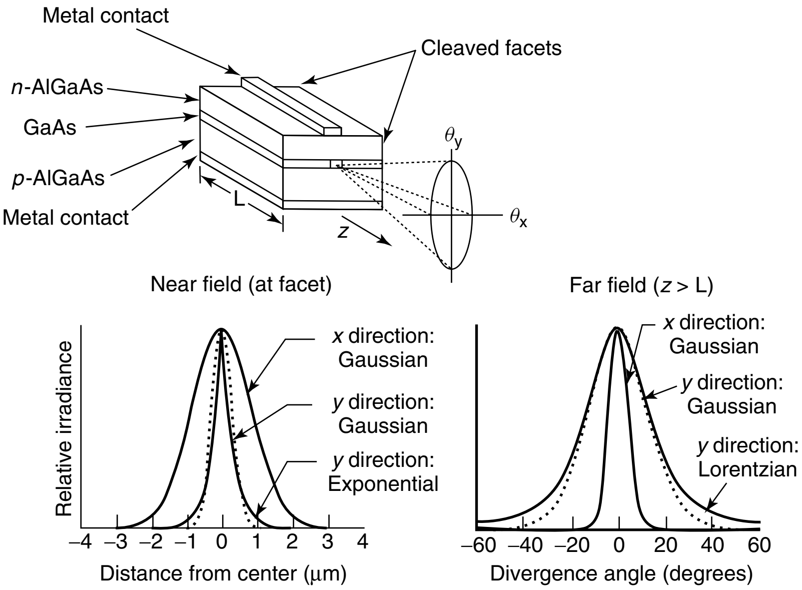


Fig. 24 A laser diode is used as the light source in optical data storage devices. Cleaved facets, a vertical waveguide structure, and either lateral gain or index confinement define the resonator cavity of the diode. The emission from the diode is not circular. Near the facet, the long dimension (x) of the emission is along the junction plane. As the beam propagates away from the facet, diffraction forces the beam shape to be about three times wider in the y-dimension than in the x-dimension. The Gaussian approximation made to the mode profiles in the y-direction slightly underestimates the power at wide divergence angles far from the facet

one spatial mode. The laser cavity, on the other hand, supports several longitudinal modes. These modes are spaced in frequency by $c/2nL$, where c is the velocity of light in free space, n is the refractive index of GaAs (3.5), and L is the length of the laser chip. Around 0.78 μm, the corresponding wavelength separation of the longitudinal modes is about 0.0003 μm. At a constant temperature, the laser oscillates in only one longitudinal mode, the one closest to the peak of the gain curve. Well above the threshold current for oscillation, the linewidth of the emission can be as small as 10 MHz. But neither the gain nor the oscillation frequency is constant with temperature. The laser can therefore jump from one longitudinal mode to another as the temperature changes, which is a phenomenon called *mode hopping*. The new mode may produce more or less power than its predecessor, depending on where it lies under the spectral gain curve. This mode change produces a source of laser noise in optical drives. Another form of laser noise is a consequence of feedback of light into the laser cavity from reflections taking place outside the laser diode cavity. In the optical paths illustrated in Figs. 3 and 4, some of the light reflected from the disk is transmitted back through the beam splitter and into the diode cavity. Thus, the data layer itself forms an external cavity for laser oscillation. The problem is so serious that MO drives modulate the laser current at a high radio frequency

(300–500 MHz) to force the laser into multiple longitudinal-mode oscillation, such that the average output power does not change over time scales characteristic of optical storage (tens of nanoseconds). The rapid modulation of the current never allows the laser to settle down into just one longitudinal mode. The higher the gain in the laser, the more sensitive it is to external feedback. Most CD, DVD, and BD systems utilize a polarizing beam splitter combined with a quarter-wave plate in order to limit feedback into the laser cavity and provide maximum power for writing data marks. The combination effectively produces an optical isolator, so that reflected light from the disk is completely directed by the beam splitter into the servo and data optics.

The mode structure of the laser diode is important for two reasons. First, the source must exhibit good spatial coherence in the transverse mode structure because most tracking servo techniques depend on an interference effect between diffracted orders, as explained in Sect. 5.6. Single transverse mode behavior is commonly achieved in commercial diodes over a wide range of operating conditions. Feedback effects from light returning to the laser influence laser output and increase laser noise, which is a source of jitter. High-frequency modulation that reduces laser noise also produces a low temporal coherence, due to the relatively large number of modes observed over the bandwidth of the data detection electronics.

The diverging wavefronts emanating from the output facet are not circular. Owing to the 1 × 3 μm emission area at the facet, the light expands anamorphically. Typically, the divergence angles of a beam emitted from a laser diode are three times greater in the vertical dimension than in the lateral dimension. Divergence correction with prisms corrects the beam to a nearly circular profile. Laser diodes also exhibit astigmatism that must be corrected by the optics of the optical head in a storage device. Gain-guided lasers exhibit 10 to 15 μm of astigmatism and index-guided lasers about 2 to 3 μm. These numbers are a measure of the axial separation of the two line images in the geometrical optics limit that would be formed if the diverging wavefront from the laser were turned into a converging one by a perfect lens at unit magnification. In most cases, the anamorphic beam correction accomplished with correcting prisms also compensates for the astigmatism.

The operating power depends on whether data are being written to the disk or data are being detected during readout. Since optical storage media are thermally sensitive, a relatively high power is required for writing data. For example, CD-R and CD-RW media typically require 5 mW to 10 mW at the disk surface during writing. Since there are losses in the optical system associated with shaping the laser diode beam and directing the beam with beam splitters, the typical efficiency of the optical path is around 50% [44]. Therefore, laser diodes that operate with powers greater than 20 mW are generally required. For systems that write data faster than standard playing time, more power is required because the disk spins faster. During readout, the required laser power is greatly diminished. Typical read-only systems require only 0.5 mW.

Since the resolution limit is improved by using short-wavelength lasers, modern DVD systems use strained multiple quantum-well diodes produced in AlGaInP/GaInP by MOCVD [45]. These diodes typically emit with a wavelength from 0.635 to 0.680 μm. Power levels from the diodes reach 35 to 50 mW. BDs use

InGaN violet laser diodes with operating wavelengths around 405 nm [46, 47].

6
Performance

Three important performance characteristics of optical data storage devices are the capacity, data rate, and access time.

Capacity is the maximum amount of data that can be stored on a single disk. Capacity is usually specified in terms of gigabytes (GB), or 10^9 bytes. (1 byte $=$ 8 bits).

Data rate is the number of digital bits per second that are recorded or retrieved from a device during transfer of a large data block. Data rate is usually specified in terms of megabits per second (Mbps), or 10^6 bits per second.

Access time is the latency experienced between when a request is made to access data and when the data starts flowing through the communication channel. Access time is usually specified in terms of milliseconds, or 10^{-3} s.

Together, data rate and access time determine the *throughput* of the device for large files. That is, throughput determines the time required to locate and transmit data to and from the storage device.

The data rate can be different for writing and reading data on a disk. During writing, the data rate is determined by the highest medium velocity that produces clearly defined marks. During reading, the data rate is determined by the highest medium velocity that produces sufficient signal-to-noise ratio. One straightforward way to increase data rate is to use more than one laser beam at a time. The increase in data rate is nearly proportional to the number of beams. A consumer CD product based on using multiple beams for readout that are generated with a diffraction grating has been shown to dramatically improve data rate without large increases in disk rotation rate [48].

The access time is determined by the mechanical latency due to the disk rotation. The highest latency is the time it takes the disk to make one revolution. Reduction of latency requires spinning the disk faster.

Important considerations for storage are the performance requirements of new and existing applications. An illustrative example is found in the CD/DVD marketplace. In 1991, the CD-ROM exhibited a capacity of 0.64 GB and a data rate of 1.2 Mbps. Although today's CD-ROM has the same or slightly more capacity, market forces have driven the data rate to over 50 Mbps. The increased data rate of CD-ROM drives may have, in part, been responsible for the delayed market acceptance of DVD-ROM. Introductory DVD products exhibit a data rate of 10 Mbps. Thus, computer applications eagerly accept higher speed devices.

A serious limitation exists with disk-based optical data storage. As the data rate increases, the playing time for a fixed capacity decreases. Applications that require long playing times (and correspondingly high capacities) must use multiple disks. For example, a CD-ROM drive operating at 50 Mbps takes only 102 s to read the entire disk. Correspondingly, a hypothetical DVD-ROM drive operating at 400 Mbps (a similar speed multiplier compared to the fast CD drive) takes less than 100 s to read a 4.7-GB disk.

A useful figure of merit is the *capacity-rate product* (CRP), which is the product of the capacity in gigabyte and the data rate in megabits per second. The CRP and other performance characteristics of disk-based products are given in Table 5. The data-rate speedup factor is shown as "1X" or "40X," where 1X refers to the data rate of products first introduced into the marketplace, like the CD-ROM in 1991. 40X refers to a data rate that is 40 times faster that the 1X rate. Also included in Table 5 are preliminary data concerning the Blu-ray product, which is now available [49].

7
Optical Data Storage for High-definition Television (HDTV)

Optical data storage is commercially successful in the form of CDs for audio and software distribution and DVDs for video distribution. CDs and DVDs look very similar because the fundamental optical technology for both devices is the same. This similarity is also true for the next generation of optical data storage, which may be used for digital home theater recording and HDTV distribution. However, CDs, DVDs, and next-generation products are different in terms of specific optical components in the drive, in how data are managed and in details of the disk structure used to store the information. These differences allow a larger volume of data to be recorded on each successive generation. Larger data volumes translate into higher quality video and longer playing time. In this section, a description of several technical aspects associated with optical disk technology is presented.

Table 6 displays a comparison of optical disk technologies in terms of several parameters. Optical parameters include laser wavelength, objective lens numerical aperture, protective layer thickness, and free working distance. Data management parameters include data rate, video format, HDTV play time, and bitrate scheme. Disk structure parameters are user data capacity, minimum channel-bit length, and track-to-track spacing. Three next-generation optical disk technologies are listed. They are the HD-DVD, the Advanced Optical Disk (AOD), and Blu-ray (BD). All three next-generation technologies have the potential to become commercially successful in terms of distributing HDTV. The AOD and Blu-ray also have the potential for recording HDTV.

In order to maximize disk capacity, the optical system uses high NA and short λ. For maximum contamination protection, the protective layer should be as thick as possible. However, the combination of thick protective layer and high NA is not easily accomplished. High NA systems are sensitive to changes in substrate thickness and disk tilt. Manufacturing variations create thickness nonuniformities, which are usually a small percentage of the total disk thickness. Motor instabilities induce tilt as the disk spins. A DVD optical spot degraded by a disk thickness variation is shown in Fig. 25(b), as compared to an ideal focused spot in Fig. 25(a). Energy from the central portion of the

Tab. 5 Capacity-rate product (CRP)

Parameter	*CD-1X*	*CD-40X*	*DVD-1X*	*DVD-40X*	*Blu-ray 1X*
Capacity (GB)	0.64	0.64	4.7	4.7	20
Data rate (Mbps)	1.2	48	10	400	25
CRP	0.77	30.7	47	4700	500
Retrieval time (min)	70	1.7	62.7	1.6	106.7

Tab. 6 Parameters for HD-video storage with optical disks and D-VHS

Parameter	*Video CD*	*DVD*	*Warner HD-DVD*	*Advanced optical disk*	*Blu-ray (BD)*
Laser wavelength (λ in μm)	Infrared (0.78)	Red (0.65)	Red (0.65)	Blue (0.065)	Blue (0.405)
Objective lens numerical aperture (NA)	0.45–0.5	0.60–0.65	0.65	0.65	0.85
User data capacity, in gigabytes (GB)	0.68	4.7 one side 9 both sides (same for recordable)	4.7 one side 9 both sides (read only)	Recordable: 20 one side 49 both sides read only: 15 one side 30 both sides	23–27 one side 50 both sides
Data rate, in megabits per second (Mbps)	1.44	10	10	(13)[a]	36
Video format	MPEG-1 or MPEG-2	MPEG-2	MPEG-4 (or enhanced MPEG-2)	(MPEG-2)[b]	MPEG-2
Bit-rate scheme	Fixed	Variable	Variable	Variable	Variable
Approximate HDTV[c] play time (hours)	Not practical	Not practical	2(+) one side 4(+) both sides	2(−) one side 4(−) both sides	2(+) one side 4(+) both sides
Protective layer thickness (mm)	1.2	0.6	0.6	0.6	0.1
Free working distance (mm)	1.2	1.0	1.0	1.0	0.05–0.10

Channel-bit length[d] in microns (μm)	0.277	0.13	0.13	(0.08)[e]	0.047–0.053
Track-to-track spacing in microns (μm)	1.6	0.74	0.74 Q3	(0.46)[f] ●	0.32
Cartridge	N	N	N	N	Y
Association	Sony, Philips	Sony, Philips	AOL Time Warner	Toshiba, NEC	Hitachi, MEI, Philips, Pioneer, Sony, Thompson, Sharp, LG Electronics, Samsung

[a] Higher bit rates are certainly possible for both recording and readout. The number in parentheses is an average. [A demonstration of high bit-rate recording is given in Ishii et al. (**2002**) "GeSbTe Phase-Change Material for Blue-Violet Laser at High Linear Speed," *Jpn. J. Appl. Phys.* **41** (Part 1 No. 3B), 1691–1692.]

[b] Not clearly defined.

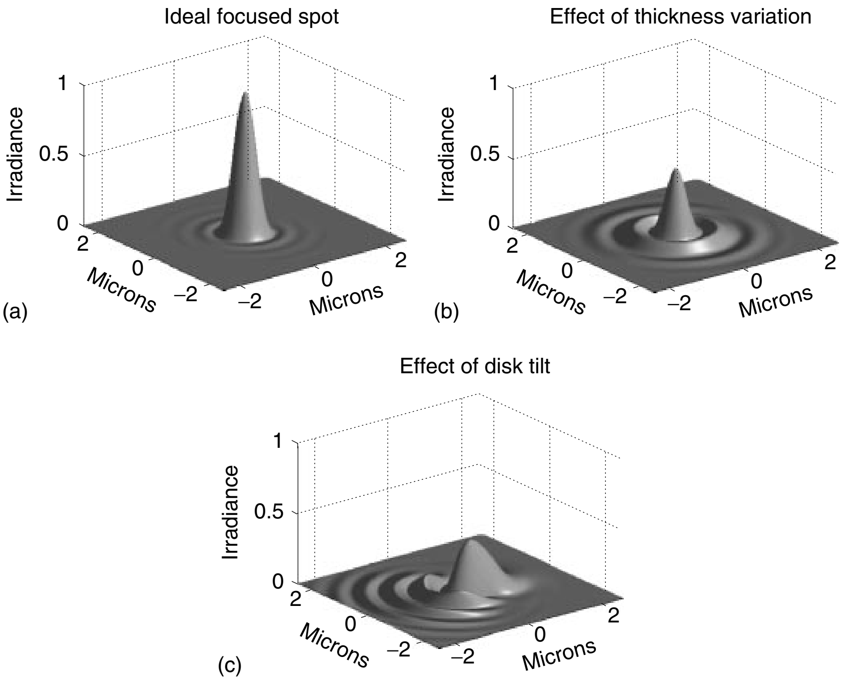
[c] Typical screen resolution is 1080i/720p.

[d] Channel-bit length for optical disks are approximately the minimum mark length divided by three.

[e] Estimated.

spot is redistributed to concentric rings, which degrade the quality of the read out signal. This effect is called *spherical aberration*. Figure 25(c) shows an optical spot degraded from disk tilt. Energy redistributes asymmetrically, which also degrades the read out signal. Tilt causes *coma*, which is another form of aberration. In order to limit these effects, the substrate is made as thin as possible without sacrificing contamination protection.

The most conservative technology listed in Table 6 is the Video CD. Its thick protective layer, relatively low NA, and long laser wavelength produce a stable system that is not very sensitive to environmental factors like dust and scratches. The ideal spot size is about 0.78/0.5 = 1.6 μm. Although the cover layer is thick at 1.2 mm, the sensitivities to thickness variations and disk tilt are low because of the low NA. DVD technology uses a shorter-wavelength laser, higher NA optics, and a thinner protective layer. The combination of short wavelength and higher NA produce a spot size of about 1.1 μm. The protective layer had to be made thinner because the sensitivity to thickness variations and disk tilt is too high otherwise. DVDs are slightly more sensitive to dust and scratches than CDs. The net effect is not great because higher NA reduces the focal depth and DVDs have a more robust error-management strategy.



Q4

Fig. 25 •The optical stylus must be nearly perfect in order to record and playback the microscopic data marks on optical disks. Thickness variations in the disk substrate induce spherical aberration, as shown in (b), which distributes energy from the central peak into the surrounding diffraction rings. Disk tilt induces coma, which redistributes the spot energy asymmetrically, as shown in (c). In both cases, the electrical data signal recovered from the disk is seriously degraded

The HD-DVD system is exactly the same as a DVD in terms of the optical parameters. The AOD and Blu-ray systems both use a new blue laser source that emits 0.405-μm light. The AOD system uses the same protective layer thickness as a DVD, and it uses the same NA objective lens. Owing to the short wavelength, the spot size for the AOD is about 0.62 μm. Sensitivity to dust and scratches is about the same as a DVD. Sensitivities to thickness variations and disk tilt are worse, however, due to the wavelength difference. In fact, in order to increase capacity of the AOD to 20 GB, margins like tilt sensitivity will need to be significantly improved in the AOD disk and player systems, as compared to DVD players. The Blu-ray system uses both higher NA and thinner cover layer. The spot size is $0.405/0.85 = 0.48$ μm, which is the smallest spot size of all the technologies listed in Table 6. However, because of the high NA, the protective layer had to be made thin to limit sensitivity to thickness variations and disk tilt. Therefore, Blu-ray disks are sensitive to dust and scratches. The free working distance is nearly the same for all technologies except Blu-ray. Blu-ray systems utilize more complicated lens systems due to the high NA, so working distance had to be reduced. The integrity of this reduced working distance is not clear at this time.

To the user, all generations of optical disks look very similar. They all are round disks that are approximately 120 mm in diameter, have a central mounting hole and are approximately 1.2 mm thick. Through many years of experience with CDs, this format has proven effective and mechanically reliable. However, the manner in which data layers are arranged on the disk depends on the technology used. For example, the CD uses a simple 1.2-mm thick substrate, as shown in Fig. 26(a). Data are recorded on only one side of the disk, through the clear 1.2-mm substrate, which also serves as the protective layer. DVDs, HD-DVDs, and AODs use the format shown in Fig. 25(b), where two 0.6-mm substrates are bonded together and the data are recorded on the bond side of each substrate. DVDs also allow two more layers per side (A, B in Fig. 26(b)), where the layers are separated by a thin adhesive spacer. The two layers are fabricated before bonding at the same time as the individual 0.6-mm substrates. Like the CD, data are recorded and read through the clear substrates. It is likely that the HD-DVD and AOD will also take advantage of this multiple-layer concept. A potential implementation of the Blu-ray disk is shown in Fig. 26(c), where the protective layer is very thin at 0.1 mm. In this case, data are recorded on the substrate, which does not serve as the protective layer. Instead, a protective layer resin is spun on and hardened or a thin protective sheet is bonded on the substrate. Initially, Blu-ray has only one data layer. The next BD product will probably be a two-layer single-sided disk. Eventually, the dual-layer dual-sided disk may be produced. Because of the thin protective layer, the Blu-ray disk must also be used with a cartridge. If Blu-ray follows the development of DVD-RAM, the product will first be released in a cartridge and then migrated to disks that do not need a cartridge. The difference in the Blu-ray disks that do not need a cartridge is the addition of a thin, hard protective coating.

The logical organization of data on the disk and how those data are used are considerations for data management. Data management considerations have important implications in the application of optical disk technology to storage for

OE011

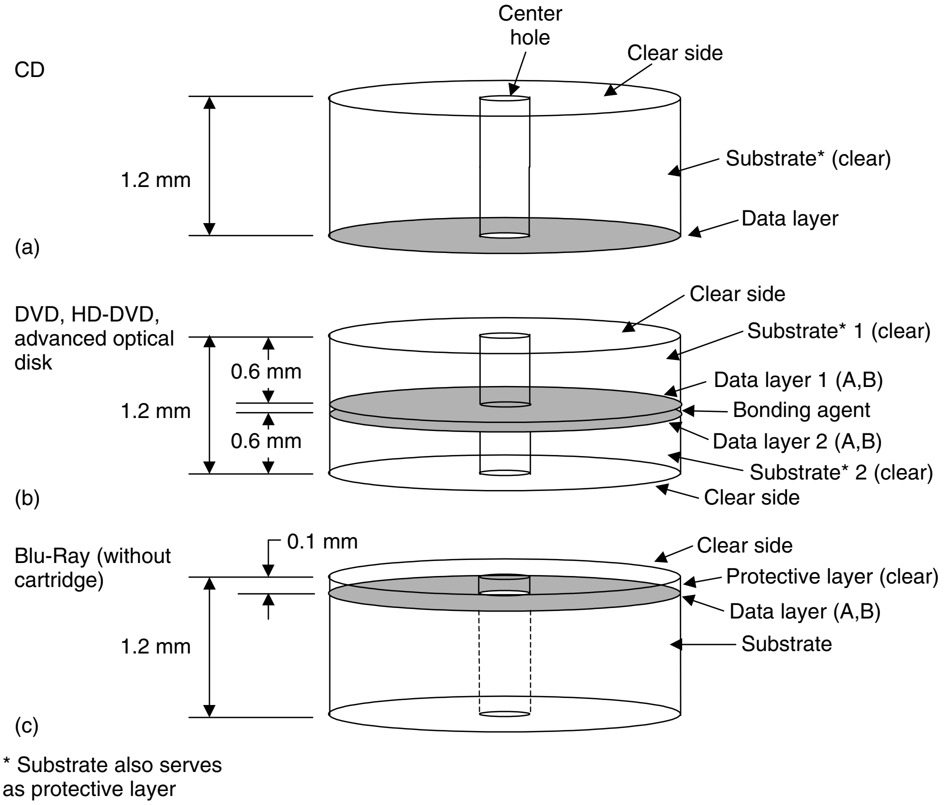


Fig. 26 Basic differences are shown between the CD, DVD-like, and Blu-ray (BD) systems. All disks are 120 mm in diameter and 1.2 mm thick. The CD records only on one side, while the DVD-like systems record data in the center of the disk. Blu-ray (BD) systems record data through a thin protective layer near the outside of the disk surface

HDTV. For example, simply using a more advanced error correction scheme on DVDs allows a 30% higher disk capacity compared to CDs. Data rate, video format, bit-rate scheme, and HDTV play time are all data management issues.

There is a basic difference in data management between CDs and DVDs. Since CDs were designed for audio, data are managed in a manner similar to data management for magnetic tape. Long, contiguous files are used that are not easily subdivided and written in a random-access pattern. Efficient data retrieval is accomplished when these long files are read out in a contiguous fashion. To be sure, CDs are much more efficient that magnetic tape for pseudorandom access, but the management philosophy is the same. On the other hand, DVDs are more like magnetic hard disks, where the file structure is designed to be used in a random-access architecture, that is, efficient recovery of variable length files is achieved. In addition, the original error correction strategy for CDs was designed for error concealment when listening to audio, where DVDs utilize true error correction. Later generations of optical disks also follow the DVD model. For example, the error correction scheme in Blu-ray is designed to allow random file manipulation, like DVD. It also includes a more robust scheme called *Picket ECC* that partially counters the increased sensitivity of Blu-ray to dust and scratches by emphasizing correction of burst errors.

The random-access nature of DVDs allows very efficient methods for data compression. For example, MPEG-2 with variable bit rate allows data to be read out from the disk *as they are required*, rather than supplying data at a constant rate. Slowly moving scenes, like love scenes or conversations, require much less information per frame than a fast-moving car chase or explosion. In these fast-moving scenes, the maximum amount of information per scene is limited only by the maximum data rate of the player. For HDTV, acceptable picture quality is obtained by using MPEG-2 with a maximum data rate of about 13 to 25 Mbps for most scenes. During a slow scene, not as many files are accessed, and much less storage area on the disk is used. This architecture leaves room on the disk for the data associated with faster-moving scenes.

Fixed-rate schemes, like magnetic tape, supply data at a constant rate, no matter what the requirements of the scene. During fast-moving scenes, the data stream from the tape supplies an adequate data rate. The tape speed and data rate for these devices are set by the upper limit of the scene requirements. Since the tape does not slow down during slower scenes, the data stream is ''padded'' at these times with useless information that takes up valuable storage area on the tape. Overall, the random-access architecture of optical disks is a much more efficient way of using the available storage area. That is, optical disks do not require as many gigabytes of user data capacity for an equivalent length and quality HDTV presentation.

It is not practical to store HDTV on CDs and DVDs with MPEG-2. For CDs, special multiple-beam readout or high-velocity disk drives could produce the data rate, which is an advantage of the fixed bit-rate scheme. However, the play time would be only a few minutes, at best. DVDs are not capable of the 13-Mbps random data rate to support MPEG-2. The AOD exhibits acceptable data rate and reasonable user data capacity for up to 2 h of HDTV per side compressed with variable bit-rate MPEG-2. Blu-ray has a slightly higher capacity and data rate. The 2-h play time for HDTV with Blu-ray in Table 6 is really a specification for real-time recording, which is not easily compressed into an efficient variable-rate scheme. Blu-ray should easily provide 2 h or longer of prerecorded HDTV per side compressed with MPEG-2.

MPEG-2 is a technique for compressing video data and replaying the data associated with certain *rules* that are defined in the MPEG-2 specifications. The action of the optical disk system is not to compress data or interpret the video information rules. Instead, the optical disk system only stores and retrieves data on command from the video operating system. Therefore, as video operating systems and associated compression technology become more advanced, no fundamental changes are required to the optical disk system. MPEG-4 technology is an advanced video compression scheme that utilizes advanced prefiltering and postfiltering, in addition to a rule-based algorithm. Estimated improvement in compression is a around a factor of 3 beyond MPEG-2.

The HD-DVD is simply a common DVD player that is being used within an advanced video operating system implementing MPEG-4, or similar compression. In this environment, the Warner HD-DVD system may provide up to 2 h of HDTV per data layer. Interestingly, if MPEG-4 or similar compression technology is successful, it can also be applied to the AOD and Blu-ray systems. However, at this time, MPEG-4 is meant as a format only for prerecorded video. The computation power required to process an HDTV movie into MPEG-4 is currently beyond the consumer market. Given the recent history of advances in the home computer market, it may not be long before an MPEG-4 disk recorder is also available.

8
Future Systems

Future illumination optical systems will use high NA and shorter λ in order to obtain smaller spot size and higher data density. The effects of using higher *NA* and shorter λ are shown in Fig. 27. For example, if the NA of a DVD system is increased to $NA = 0.85$, the spot size is reduced by 30%. However, the allowable defocus is reduced by 50%. Alternatively, a blue laser operating at $\lambda = 0.405$ µm and $NA = 0.60$ achieves nearly a 40% reduction in best-focus spot size at a penalty of reducing the allowable defocus by the same 40% factor. In general, it is desirable to decrease wavelength rather than increase NA due to the difficulty of decreased depth of focus.

Parameters of three generations of optical disk products are shown in Table 6. These systems are evolutionary products. Shortening laser wavelength and increasing NA reduce spot size and increase capacity. The Blu-ray system, which operates at $\lambda = 0.405$ µm and $NA = 0.85$, provides a capacity of 27 GB per layer.

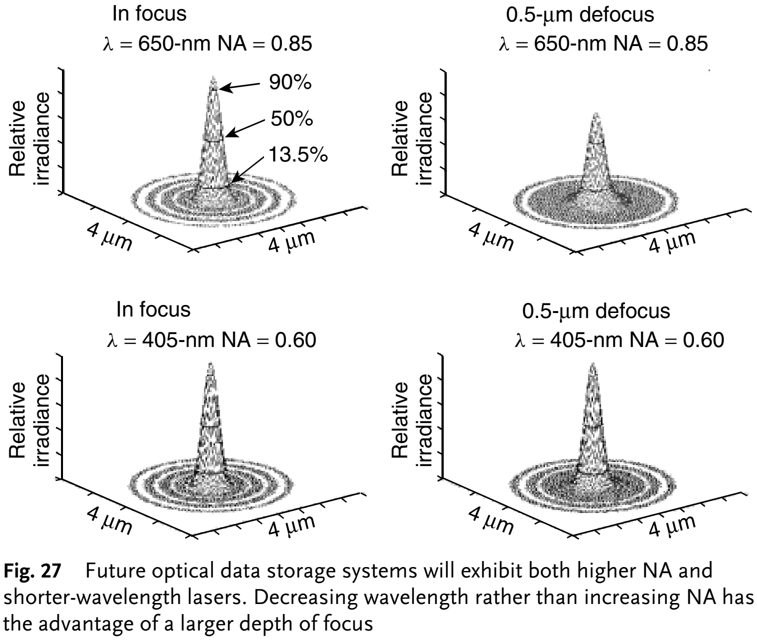


Fig. 27 Future optical data storage systems will exhibit both higher NA and shorter-wavelength lasers. Decreasing wavelength rather than increasing NA has the advantage of a larger depth of focus

However, the Blu-ray system is near the limit for conventional optical systems with standard optical materials. For example, increasing NA beyond 0.85 is possible with a conventional optical system, but the engineering challenges are substantial. In addition, most plastic substrates exhibit poor transmission below 0.400 μm. Even if a new laser diode becomes available with a wavelength shorter than 0.400 μm, it is not clear whether substrate, detector, and media technologies can support it. Instead, recent research points to two promising technologies that may provide the fourth-generation optical disk.

The first technology is called *near-field optics* [50, 51]. Near-field optics use a transducer, like a small hole in a metal film or a special lens element, to produce a light spot that is smaller than the ideal spot size given by $s = \lambda/NA$. However, the trade off for smaller spot size is that the recording layers now must be in proximity to the transducer. The evanescent energy in the spot that couples from the transducer to the recording layers falls off exponentially with distance.

Invented by Prof. Gordon Kino and colleagues at Stanford University, the solid immersion lens (SIL) is under investigation as a possible candidate for near-field transducer [52]. The basic SIL system is shown in Fig. 28, where the optical system is supplemented with a hemispherical lens element. When the focused light from the objective lens enters the SIL, the velocity of the light slows down according to n, the index of refraction of the lens. Marginal ray angle θ is not deviated by the hemisphere as it enters the lens material. Since the laser frequency does not change, the effective wavelength of the light reduces and the spot size is now given by $s = \lambda/(n \sin\theta) = \lambda/NA_{\text{eff}}$, where NA_{eff} is the *effective numerical aperture.* In laboratory systems, NA_{eff} approaching 2.0 have been demonstrated [53]. When coupled with a blue laser diode, the

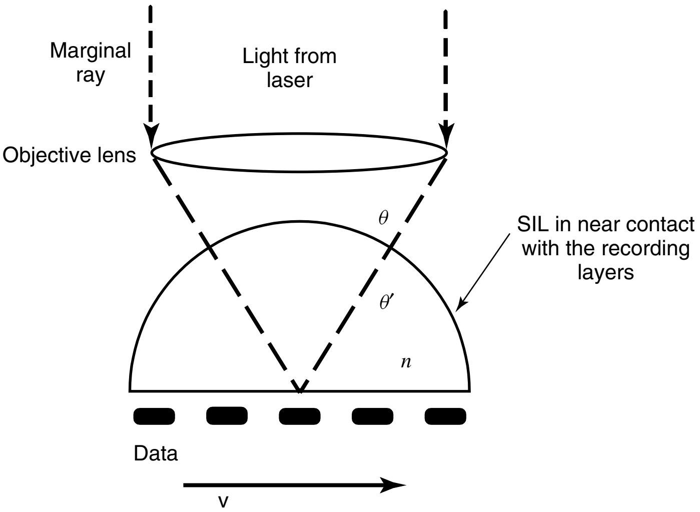


Fig. 28 A solid immersion lens (SIL) system uses an image-centric hemisphere in near contact to the recording layers. The SIL increases the effective numerical aperture of the system by a factor of the lens refractive index *n*, and it decreases the focused spot size by the same amount

potential of an SIL system is to increase capacity beyond a factor of 4 above the Blu-ray system. However, control of the gap that separates the SIL from the recording layers is a difficult engineering problem, especially if the optical disk is removable from the optical drive.

A second possible technology for fourth-generation optical disks is called *magnetically amplifying magneto-optical system* (MAMMOS), which is similar to MO systems described in Sect. 4. MAMMOS technology takes advantage of the fact that the primary limitation to resolution in optical data storage systems is reading data, as Q8 explained in •Sect. 5.3. With a pulsed laser and a modulating external field, magnetic domains can be written in the recording layer that are much smaller than the resolution limit [54]. Readout of these marks in a MAMMOS system is illustrated in Fig. 29, where a multiple-layer MO stack is used. Each MO layer reacts differently to the heat deposited by the laser beam. The bottom layer, which is called *the recording layer*, contains the written information in the form of small bits. This layer has a high coercivity, and it is not easily affected by the relatively low-temperature profiles generated by the readout beam. The top layer is the expansion layer, and it has a low coercivity, among other special properties. The middle layer is a thin nonmagnetic layer. When the readout beam heats the expansion layer, magnetic energy from the recording layer couples into the expansion layer and forms an expanded copy of the recording layer in it. Only a small region of the storage layer around the center of the laser spot is copied. Expansion of the bit pattern produces a magnified image in the expansion layer. To the readout optical system, it appears that the light spot travels over relative large marks, which produce good signal-to-noise ratio. Capacity of MAMMOS systems have been demonstrated to be about three times greater than Blu-ray disks [55]. Potential difficulty in MAMMOS systems lies in economically producing disks and player systems.

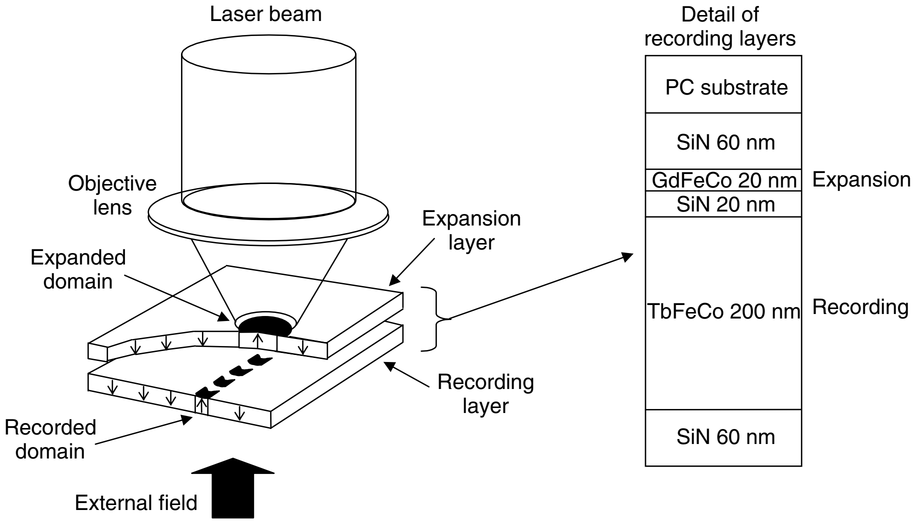


Fig. 29 The magnetically amplifying magneto-optical system (MAMMOS) uses a special configuration of magnetic layers to produce good signal readout from a very small data-mark pattern

Other systems are also worthy of mention because disruptive technologies are always possible. For example, near-field superresolution structures (SuperRENS) combine a nonlinear optical material layer with a conventional phase-change media structure [56]. This system has the advantage of using near-field optical effects with a conventional readout system. It is not necessary to maintain a small gap, as required in the SIL systems. Also, volumetric storage systems show promise. Instead of recording only on one or two layers, volumetric bit-wise systems store data on several hundred layers through the thickness of the disk [57]. Volumetric bit-wise system may need to use nonlinear properties of the recording layers in order to record data marks without interference from other layers. Volumetric holographic storage is also of interest, where data are stored throughout the volume as a collection of gratings [58]. Finally, optical data storage and magnetic disk storage may converge into *hybrid* recording, which uses the optical beam only as a heat source to lower the coercivity for magnetic writing [59]. Hybrid readout is accomplished with magnetic sensors, and hybrid recording may use near-field optics.

Glossary

Access Time: The time required to move an optical (or magnetic) head to an arbitrary data track on a disk. Access time is usually specified in terms of milliseconds, or 10^{-3} s.

Astigmatic Focusing: Method to provide a FES by fabrication of a small amount of astigmatism into the servo optics.

Blu-ray Disc (BD): Erasable phase-change technology using a blue laser diode. The capacity of BD disks is over 20 GB.

Capacity: The maximum amount of data that can be stored on a single disk. Capacity is usually specified in terms of gigabytes (GB), or 10^9 bytes (1 byte = 8 bits).

Capacity-rate Product (CRP): The product of the capacity in GB and the data rate in Mbps.

CD-DA (Digital Audio): A data format used for encoding digital audio.

Channel Bit: A bit in the data stream of information that is recorded onto the optical disk. A channel bit differs from the physical pits or data marks on the disk in that a single pit may represent more than one channel bit.

Channel Bit Window: Defines a channel bit as falling within a small time interval (called a *window*) during the recording or playback sequence. During playback, transitions in the electrical readout signal caused by the optical spot passing over the edge of a data mark signals the change of a channel bit value. That is, the transition signals a zero-to-one or one-to-zero change in the stream of channel bits. Owing to noise and instabilities in the system, there is some uncertainty in the transition. The channel bit window allows the transition to occur within an interval wide enough to allow for reliable detection. The size of the channel bit window is the determining factor for the minimum spacing of bits along the track.

Channel Codes: Set of rules used to convert user data bits into their physical data-mark representation and back again. For CDs, it is called *eight-to-fourteen modulation* (EFM).

Coercivity: The magnetic field required to reorient domains.

Compact Discs (CD): Single-layer optical storage medium with a capacity of around 0.7 GB. The diameter of a CD is about 120 mm, and the substrate is 1.2 mm thick. The CD recording head uses a near-infrared laser diode and an objective lens with a 0.47 numerical aperture.

Coma: An aberration induced by tilt of the disk with respect to the beam focused through it by the recording head.

Data Density: The number of bits of information stored per unit area on the surface. Often specified in gigabits (10^9 bits) per square inch of recording surface area (GB in^{-2}).

Data Rate: The number of digital bits per second that are recorded or retrieved from a device during transfer of a large data block. Data rate is usually specified in terms of megabits-per-second (Mbps), or 10^6 bits per second.

Depth of Focus: The amount of axial displacement from the focal point that does not change the system output: $\Delta z = +/-\lambda n/(4 \sin^2\theta) = +/-\lambda n/(4\text{NA}^2)$.

Diffracted Orders: The two outer cones of light reflected off of a collection of tracks (recordable media) or a collection of data marks (ROM media).

Digital Versatile Discs (DVD): Multiple-layer optical storage medium with a single-layer capacity of about 4.7 GB. The diameter and thickness of a DVD is the same as a CD, but the data layers on a DVD are in the center of the 1.2 mm thickness. The DVD uses a red laser and a 0.6 numerical aperture objective lens in the recording head.

Equalize: Compensate for the reduced amplitude of high-frequency data bit patterns using electrical filters in the readout electronics.

Focus Error Signal (FES): Error signal generated from the focus sensor.

Formats: Different ways of organizing ones, zeros, and system information on the disk.

Hybrid Recording: Uses the optical beam only as a heat source to lower the coercivity for magnetic writing.

Index of Refraction (n): Ratio of the speed of light in vacuum to the speed of light in a material.

Irradiance: Laser power (Watts) per unit area (square meters).

Isotherms: Lines that represent points of constant temperature on an object.

Jitter: A variation of the data signal transition from the center of the timing window. Usually, jitter is Gaussian distributed about the center and the Gaussian width is specified as a percentage of the width of the window.

Magneto-Optic (MO): The information is stored as data marks in small magnetic domains. The orientation of the magnetic domains within a data mark are mostly uniform and differ from the orientation of magnetic domains in the land area outside the marks.

MiniDisc (MD): A magneto-optical data storage medium made by Sony for recording and playing music on a 2.5-in disk.

Mode Hopping: A term used to describe the behavior of a laser diode that is usually induced by reflected light from the disk entering the laser cavity. Specifically, it is the jumping from one longitudinal mode to another. Mode hopping can also occur as temperature changes.

Modulation Transfer Function: Measures the contrast of the data signal modulation versus the spatial frequency of the data-mark pattern.

Numerical Aperture (NA): Often known as value $\sin\theta_{\max}$, where $\theta_{\max}$ is the apex angle of the cone of light converging, in air, to the focal spot.

NA$_{\text{eff}}$: The effective numerical aperture when the focus cone is measured inside a dielectric material, like the glass element of a solid immersion lens. The expression for the effective numerical aperture includes the index of refraction n of the dielectric. That is, $NA_{\text{eff}} = n \sin\theta_{\max}$.

Objective Lens: The lens element in the recording head that focuses light onto the data layer.

Oresteds: A unit of magnetic field intensity.

Optical Data Storage: Using light as a tool to store and retrieve data.

Pit-type Data-mark Technology: A data storage method where small data-mark pits are arranged in spiral tracks around the center of a disk. When a light spot passes over a pit, most of the reflected light scatters away from the illumination optics. The remaining light is collected by the objective lens is small compared to the amount of light that gets collected when the spot is over a smooth portion of the track, where the disk surface acts as a mirror to the focused light.

Pitch: The center-to-center track spacing in the radial direction.

Polar Kerr Effect: The effect that causes polarization rotation on reflection from a magnetically polarized medium.

Push-pull Tracking: Method to provide a TES using a groove pattern on the disk, where the period of the grooves is equal to the track pitch.

Resolution Limit (T_R): The mark period where there is no signal modulation at the detector: $T_R = \lambda/2\text{NA} = s/2$. Longer mark periods will produce signal modulation at the detector.

Spherical Aberration: An optical effect caused by focusing through the optical disk. If not corrected, energy from the central portion of the spot is redistributed to concentric rings, which degrade the quality of the read out signal.

Spatial Frequency: How often a signal repeats in space. It is the inverse of the period of the bit spacing.

Spot Size: The full width of the irradiance distribution at the $1/e^2$ (13.5%) irradiance level relative to the peak.

Symbol: A fourteen-bit channel sequence.

Thermal Blooming: When heat builds up and spreads out in the direction perpendicular to the scan, it causes the end of the pulse to generate a wider isotherm than at the beginning of the pulse.

Three-spot Tracking: Method to provide a tracking-error signal. The signals are processed from three spots focused onto the disk, where the central spot is centered over a track and two neighboring spots are slightly each side of the central spot in the direction across the track.

Throughput: Determines the time required to locate and transmit data to and from the storage device.

Track Pitch: The radial distance between turns of the spiral tracks.

Tracking Error Signal (TES): The error signal generated from the tracking sensor.

Write Bright: Term used when talking about phase-change disks. The recording layer is initially in the amorphous state, and bright crystalline marks are written on the dark amorphous background.

Write Dark: Term used when talking about phase-change disks. The amorphous state of the data marks does not reflect as much light as the crystalline background.

References

[1] Gregg, D. P. *Transparent Recording Disc*, U.S. Patent 3,430,966, issued March 4, 1969.

[2] Bellis, M. *David Paul Gregg and the Optical Disc*, from http://inventors.about.com/library/inventors/blopticaldisk.htm, Nov. 7, 2002.

[3] Cellitti, D. R. (2002), World on a silver platter: a brief history of optical storage systems, *Widescreen Review: Laser Magic 1988 Guide*, •ISSN 10804722, p. 14. (An updated copy of this article can be found at http://www.oz.net/~blam1/DiscoVision/WRLM98_main.htm, Nov. 11.) Q9

[4] Howe, D. G. (1997), Data reliability and errors, Chapter 2 in T. W. McDaniel, R. H. Victora (Eds.), *Handbook of Magneto-Optical Data Recording*. NJ: Noyes Publications.

[5] Immink, K. A. S. (1999), *Codes for Mass Data Storage Systems*. The Netherlands: Shannon Foundation Publishers.

[6] Kubo, M., Harada, S., Takeshima, H., Kobayashi, T., Ohmori, T., Kobayashi, Y. (**1993**), *Jpn. J. Appl. Phys.* **32** (Part 1, 11B), 5329–5334.

[7] Marchant, A. B. (1990), *Optical Recording: A Technical Overview*. Reading, MA: Addison-Wesley.

[8] Tieke, B., Langereis, G. R., Meinders, E. R., Kablau, J. G. F., Woudenberg, R., van Kollenburg, R. A. J. (**2002**), *Jpn. J. Appl. Phys.* **41** (Part 1, 3B), 1735–1738.

[9] Burgess, A. N., Evans, K. E., Mackay, M., Abbot, S. J. (**1987**), *J. Appl. Phys.* **61**(1), 74–80.

[10] Huh, Y. J., Kim, J. S., Nam, T. Y., Kim, S. C. (**1997**), *Jpn. J. Appl. Phys.* **36**(Part 1, 1B), 403–409.

[11] CD Dye, a web document published by CD Media World, Nov. 11, 2002, http://www.cdmediaworld.com/hardware/cdrom/cd_dye.shtml.

[12] Stinson, D., Ameli, F., Zaino, N. (1995), *Lifetime of Kodak Writable CD and Photo CD Media*, a white paper available from Digital and Applied Imaging Division of Eastman Kodak Company, 1669 Lake Avenue, Rochester, New York; See also http://www.cd-info.com/CDIC/Technology/CD-R/Media/Kodak.html.

[13] Handa, T., Tominaga, J. Haratani, S., Takayama, S. (**1993**), *Jpn. J. Appl. Phys.* **32**(Part 1, 11B), 5226–5229.

[14] Terada, M., Furuya, K., Okamura, T., Morimoto, I., Nakao, M. (**1993**), *Jpn. J. Appl. Phys.* **32**(Part 1, 11B), 5219–5222.

[15] Hamada, E., Takagishi, Y., Yoshizawa, T., Fujii, T., Negishi, R., Nakajima, T. (**2000**), *Jpn. J. Appl. Phys.* **39**(Part 1, 2B), 785–788.

[16] Tyan, Y. S., Preuss, D. R., Vazan, F., Marino, S. J. (**1986**), *Jpn. J. Appl. Phys.* **59**(3), 716–719.

[17] Suzuki, M., Furuya, K., Nishimura, K., Mori, K., Morimoto, I. (1990), Disk structure for high performance phase change erasable optical disk, in M. de Haan, Y. Tsunoda (Eds.), *Optical Data Storage 1990*, Proceedings of SPIE Vol. 1316, pp. •374–381. Q10

[18] Mee, C. D., Daniel, E. D. (1987), *Magnetic Recording Volume I: Technology*. New York: McGraw-Hill.

[19] Hurst Jr, J. E., McDaniel, T. W. (1997), Writing and erasing in magneto-optical recording, Chapter 7 in T. W. McDaniel, R. H. Victora (Eds.), *Handbook of Magneto-Optical Data Recording*. NJ: Noyes Publications.

[20] Wright, C. D. (1997), The magneto-optical readout process, Chapter 8 in T. W. McDaniel, R. H. Victora (Eds.), *Handbook of Magneto-Optical Data Recording*. NJ: Noyes Publications.

[21] Kryder, M. H. (**1985**), Magneto-optic recording technology, *J. Appl. Phys.* **57**(8, Pt. 2B), 3913–3918.

[22] Brucker, C. (1997), Magneto-optical thin film recording materials in practice, Chapter 5 in T. W. McDaniel, R. H. Victora (Eds.), *Handbook of Magneto-Optical Data Recording*. NJ: Noyes Publications.

[23] A portion of this information was obtained from web document http://www.cilect.org/NCDRFF.htm, Nov. 12, 2002.

[24] Ide, H., Toda, T., Kirino, F., Maeda, T., Kugiya, F., Mita, S., Shigenatsu, K. (**1993**), *Jpn. J. Appl. Phys.* **32**(Part 1, 11B), 5342–5348.

[25] Yokogawa, F., Miyanabe, S., Ogasawara, M., Kuribayashi, H., Tomita, Y., Yamamoto, K. (**2000**), *Jpn. J. Appl. Phys.* **39**(Part 1, 2B), 819–823.

[26] Brouwhuis, G., Braat, J. J. M. (**1978**), *Appl. Opt.* **17**(13), 1993–2000.

[27] Finkelstein, B. I., Childers, E. R. (1991), The effects of focus misregistration on optical disk performance, in J. J. Burke, T. A. Shull, Nobutake• (Eds.), *Optical Data Storage 1991*, Proceedings of SPIE Vol. 1499, pp. 438–449. *Q11*

[28] Getreuer, K. W., Grassens, L. J. (1997), Servos and actuators, Chapter 3 in T. W. McDaniel, R. H. Victora (Eds.), *Handbook of Magneto-Optical Data Recording*. NJ: Noyes Publications.

[29] Dorf, R. C. (1986), *Modern Control Systems*. Reading, MA: Addison-Wesley.

[30] Braat, J. J. M., Brouwhuis, G. (**1978**), *Appl. Opt.* **17**(13), 2013–2021.

[31] Stork, D. G. *CD ROM Apparatus for Improved Tracking and Signal Sensing*, U.S. Patent 5,642,341 issued June 24, 1997.

[32] Cohen, D. K., Gee, W. H., Ludeke, M., Lewkowicz, J. (**1984**), *Appl. Opt.* **23**(4), 565–570.

[33] Finkelstein, B. I. (1997), Sources of noise in magneto-optical readout, Chapter 9 in T. W. McDaniel, R. H. Victora (Eds.), *Handbook of Magneto-Optical Data Recording*. NJ: Noyes Publications.

[34] Peng, C., Mansuripur, M. (**1998**), *Appl. Opt.* **37**(5), 921–928.

[35] Immink, K. A. S. (**2001**), *IEEE J. Selected Areas Commun.* **19**(4), 756–764.

[36] Immink, K. A. S. (**1990**), *Proc. IEEE* **78**(11), 1745–1759.

[37] Howe, D. G. (1986), Signal-to-noise ratio (SNR) for reliable data recording, in R. P. Freese, A. A. Jamberdino, M. R. de Haan (Eds.), *Optical Mass Data Storage*, Proceedings of SPIE Vol. 695, pp. •255–261. *Q12*

[38] Good reference sites to download information about CD and DVD formats are http://www.disctronics.co.uk/technology/index.htm or http://www.discusa.com/cdref/cdbooks/books.htm.

[39] These books can be obtained from Philips International B.V., System Standards & Licensing, Building SFF 8, Glaslaan 2, 5616 LW Eindhoven, or P.O. Box 80002, 5600 JB Eindhoven, The Netherlands, Fax: +31-40-2732113, URL: www.licensing.philips.com

[40] Ashley, J., Bernal, M. P., Burr, G. W., Coufal, H., Guenther, H., Hoffinagle, J. A., Jefferson, C. M., Marcus, B., Macfarlane, R. M., Shelby, R. M., Sincerbox, G. T. (**2000**), *IBM J. Res. Dev.* **44**(3), 341–368.

[41] Hunter, S., Kiamilev, F. Esener, S., Parthenopoulos, D. A., Rentzepis, P. M. (**1990**), *Appl. Opt.* **29**(14), 2058–2066.

[42] Tokumaru, H., Okumura, H., Arai, K., Kawamura, N., Yoshimura, S. I. (**1998**), *Jpn. J. Appl. Phys.* **37**(Part 1, 4B), 2241–2244.

[43] Leibbrandt, G. W. R., Kahlman, J. A. H., van Rosmalen, G. E., Vrehen, J. J. (1997), Optical tape system: evaluation of recorder and media, in H. Birecki, J. Z. Kwiecien (Eds.), *Optical Data Storage 1997*, Proceedings of SPIE Vol. 3109, pp. •106–115. *Q13*

[44] Milster, T. D., Benedict, M. K., Stahl, R. P. (1990), Laser diode requirements for magneto-optical storage devices, in M. de Haan, Y. Tsunoda (Eds.), *Optical Data Storage 1990*, Proceedings of SPIE Vol. 1316, pp. •143–149. *Q14*

[45] Hiroyama, R., Inoue, D. Nomura, Y., Shono, M., Sawada, M. (**2002**), *Jpn. J. Appl. Phys.* **41**(Part 1, 2B), 1154–1157.

[46] Piprek, J., Nakamura, S. (**2002**), *IEE Proc.-Optoelectron.* **149**(4), 145–151.

[47] Hatakoshi, G., Onomura, M., Saito, S., Sasanumaand, K., Itaya, K. (**1999**), *Jpn. J. Appl. Phys.* **38**(Part 1, 3B), 1780–1785.

[48] http://www.kenwoodtech.com/white_paper6.html, Nov. 11, 2002.

[49] Furumiya, S., Minamino, J. I., Miyashita, H., Nakamura, A., Shouji, M., Ishida, T., Ishibashi, H. (2002), Optical disk recording system of 25 GB capacity, in T. Hurst, S. Kobayashi (Eds.), *Optical Data Storage 2001,* Proceedings of SPIE Vol. 4342, pp. Q15 •186–193.

[50] Milster, T. D. (**2000**), *Proc. IEEE* **88**(9), 1480–1490.

[51] Milster, T. D. (**2001**), *Opt. Eng.* **40**(10), 2255–2260.

[52] Mansfield, S. M., Studenmund, W. R., Kino, G. S., Osato, K. *Opt. Lett.* **18**(4), 305–307 (**1993**).

[53] Shinoda, M., Saito, K., Kondo, T., Ishimoto, T., Nakaoki, A. (2002), High density near field readout over 50 GB capacity using a solid immersion lens with high refractive index, Paper WC.2, *Technical Digest of the Joint International Symposium on Optical Memory and Optical Data Storage,* Waikoloa, Hawaii, 7–11 July, IEEE Catalog Number 02EX552, pp. 284–286.

[54] Awano, H., Ohnuki, S., Shirai, H., Ohta, N., Yamaguchi, A., Sumi, S., Torazawa, K. (**1996**), *Appl. Phys. Lett.* **69**(27), 4257–4259.

[55] Awano, H., Sekine, M., Tani, M., Kasajima, N., Ohta, N., Mitani, K., Takagi, N., Sumi, S. (**2000**), *Jpn. J. Appl. Phys.* **39**(Part 1, 2B), 725–728.

[56] Tominaga, J., Hiroshi, F., Sato, A., Nakano, T., Fukaya, T., Atoda, N. (**1998**), *Jpn. J. Appl. Phys.* **37**(Part 2, 11A), L1323–L1325.

[57] Kawata, Y., Nakano, M., Lee, S. C. (**2001**), *Opt. Eng.* **40**(10) 2247–2254.

[58] Ashley, J., Bernal, M.-P. Burr, G. W., Coufal, H., Guenther, H., Hoffnagle, J. A., Jefferson, C. M., Marcus, B., Macfarlane, R. M., Shelby, R. M., Sincerbox, G. T. (**2000**), *IBM J. Res. Dev.* **44**(3), 341–368.

[59] Ota, N., Sekine, M., Awano, H., Imai, S., Hohlfeld, J., Rasing, T. (2002), Hybrid recording capability toward tera bit/in^2 and 100 Gbps storage device, Paper TuC.1, *Technical Digest of the Joint International Symposium on Optical Memory and Optical Data Storage,* Waikoloa, Hawaii, 7–11 July, IEEE Catalog Number 02EX552, pp. 156–158.

QUERIES TO BE ANSWERED BY AUTHOR (SEE MARGINAL MARKS Q..)

IMPORTANT NOTE: Please mark your corrections and answers to these queries directly onto the proof at the relevant place. Do NOT mark your corrections on this query sheet.

Q1. Please provide the telephone number and fax number of the author.
Q2. Please spell out this abbreviation at the first instance.
Q3. Please provide the footnote for this citation.
Q4. Please clarify if part (a) also needs to be mentioned in the caption.
Q5. Please clarify if some word is missing in this sentence.
Q6. Please clarify if we can change the sentence ‘moves off track center’ to ‘moves off track from the center’.
Q7. We have changed the Section number to 5.7 here because ‘jitter requirement’ has been explained in that section. Please clarify if this is fine.
Q8. Data reading is explained in Sect. 6. Please clarify if we may change this section reference to ‘Sect. 6’.
Q9. Please provide the publisher’s name and the place of publication for this reference.
Q10. Please provide the publisher’s name and the place of publication for this reference.
Q11. Please provide the editor’s forename and provide the publisher’s name and the place of publication for this reference.
Q12. Please provide the publisher’s name and the place of publication for this reference.
Q13. Please provide the publisher’s name and the place of publication for this reference.
Q14. Please provide the publisher’s name and the place of publication for this reference.
Q15. Please provide the publisher’s name and the place of publication for this reference.