Rewritable optical disk drive technology

by P. Asthana
B. I. Finkelstein
A. A. Fennema

Optical disk drives provide an effective solution to the growing need for removable high-capacity storage. In this paper, we review the technology behind the optical disk drives used in IBM's optical storage systems. The basic physics of data recording and readout and the engineering of the primary building blocks of an optical drive (the optical head, the servo system, and the data channel) are discussed. We also outline the technological directions of future optical drives as they must continue to improve in performance and capacity.

Introduction

The ability to store and transfer large quantities of information, be it data, text, or image, is becoming a crucial part of the business process in today's information-intensive world. The amount of data processed by businesses has been increasing steadily, driven by the increased computerization of business processes, the trend toward richer context data (such as embedded images or video), and even the rise of the World Wide Web (which promotes the storage of information in digital form). As the amount of data that is processed increases, so too does the amount of data that must be stored, which increases the need for cost-effective data storage. The need to transfer context-rich information such as video or multimedia files also places severe and in some cases impractical demands on floppy disks or commercial communication lines.

Removable storage systems provide a good match to the need for cost-effective data storage and transfer. The marginal cost of increasing storage is very small—i.e., extra storage space is easily acquired by using additional inexpensive media cartridges. Removable storage systems include floppy disks, removable magnetic disk drives, magnetic tape, and optical disks.

When storage needs exceed approximately a gigabyte, the most cost-effective storage is provided by either tape or optical storage. In situations where random access to the data is subsequently required, optical storage is the preferred solution. The random-access capabilities are approaching those of low-end magnetic disk drives. Furthermore, optical disks are extremely robust and relatively impervious to wear or to standard magnetic fields.

For a workstation or personal computer user, a single optical drive connected to the computer may be sufficient to meet storage needs. For users with larger storage needs, optical libraries (or "jukeboxes") are available in several sizes and can provide hundreds of gigabytes of storage capacity in a relatively small footprint. An optical library will have one or more drives and a large number of cartridges. For example, IBM's largest optical library currently contains four drives and more than a hundred

0018-8646/96/$5.00 © 1996 IBM

cartridges. Random access to any cartridge is possible through the use of an automated picker mechanism.

Optical drives are available in a number of technologies. Rewritable optical drives are based on either magneto-optical (MO) or phase-change technologies. Write-once, read-many (WORM) drives are suitable for archiving data; they can either be true ablative WORM (in which marks are ablated in the media) or firmware WORM (or MO-WORM), in which an MO disk is used but specific instructions on the media and specialized firmware in the drive prevent the possibility of rewriting. Many security-conscious customers (such as those in the financial, legal, or government fields) prefer true ablative WORM.

Standardization of the disk format has been an important issue in ensuring customer acceptance of optical drive technology. IBM has been very active in defining the international standards. The first-generation rewritable capacity and format adopted by the International Organization of Standards (ISO) was the 650MB capacity for the 5.25-in. and 128MB capacity for the 3.5-in. optical cartridges. These capacity points are referred to in the industry as 1× capacity points for each of the two form factors. In this paper we refer to 1×, 2×, or 4× capacity points when discussing disk capacity, because this terminology can conveniently be applied to either the 3.5-in. or the 5.25-in. form factors. The second-generation standard (or 2× capacity), approved by ECMA in 1991 and submitted to the ISO, was the 1.3GB capacity for the 5.25-in. and the 230MB capacity for the 3.5-in. optical disk drives. The third-generation standard is 2.6 GB (or 4× capacity) for the 5.25-in. form factor, and it is likely to be >500 MB for the 3.5-in. form factor.

Typically, optical drives in the 3.5-in. form-factor size are used for personal storage applications. The optical disk cartridge of a 3.5-in. drive is the same size as a standard floppy disk cartridge. For jukebox applications, the 5.25-in. form-factor optical drives are typically used.

IBM has a long history of key contributions in the field of optical storage. The basic formulation used in all current magneto-optic disks was invented at the IBM research laboratories in Yorktown in 1971 [1] (for which the inventors were recently awarded the National Medal of Technology by President Clinton). In addition to basic work in optical storage, IBM also took an active role in product development, and was the first company to deliver a 3.5-in. form-factor rewritable optical drive to the market in 1991. Within two years, IBM had expanded its product line to include 5.25-in. drive products and optical library products. The 5.25-in. products were used in IBM's optical libraries and provided a competitive advantage, since they uniquely offered both magneto-optic and ablative WORM capability. From a business perspective, however, the stand-alone drive products could not provide sufficient financial returns to warrant further investment, and IBM decided to halt the manufacture and sale of optical drives in 1995. IBM continued its manufacture and sale of optical library systems (which command good margins) using vendor drives.

In this paper we consider the technology behind magneto-optical drives. Optical recording contains sufficient subtleties that thorough coverage would require a book of some substance. Therefore, given the scope of this paper, the discussions herein are necessarily of limited depth. Still, a sufficient technical depth is maintained to give the reader an appreciation for the technology of optical recording and the design of the drive. The drive technology can be broken down into four main sections: the optical disk media on which information is recorded; the optical head, which involves the laser and optics; the servo portion, which involves the focusing, seeking, and track-following functions; and the channel, which involves the recorded signal-detection and processing functions. These sections are discussed separately in this paper. Finally, the paper discusses future technology issues and specific areas that must be addressed in order to achieve higher storage capacities and performance.

Magneto-optical media and basic recording physics

In an optical drive, the diverging beam from the laser diode is collimated by the optics in the head and directed toward the optical disk. An objective lens is used to focus the laser beam to a diffraction-limited spot on the media. Data recording is achieved through a thermomagnetic process [2], which is also known as Curie point writing, since it relies on the threshold properties of the Curie temperature of magnetic materials. In this process, the energy within the focused optical spot heats the recording material past its Curie point (about 200°C), a threshold above which the magnetic domains of the material are susceptible to moderate (about 300 gauss) external magnetic fields. Note that magneto-optic recording films possess vertical magnetic anisotropy; i.e., the magnetic domains are always oriented vertically. Application of an external magnetic field is used to set the state of the magnetization vector (which represents the polarization of the magnetic domains) in the heated region to either "up" (a "one" bit) or "down" (a "zero" bit). When the material is cooled below the Curie point, this orientation of the magnetic domains is fixed. This recording cycle has been shown to be highly repeatable (more than one million cycles) in any given region without degradation of the material. This is an important aspect if the material is to be claimed as fully rewritable.

In any practical recording process, it is necessary to have a sharp threshold for recording. This ensures the stability of the recorded information both to environmental conditions and during readout.

Thermomagnetic recording is an extremely stable process. At room temperature (and below), the magnetic domains in magneto-optic recording films are affected only by fields greater than several kilogauss (in comparison, the information stored on a magnetic floppy disk is affected by magnetic fields as low as a hundred gauss). The coercivity of a magneto-optic material remains high until it has almost reached the Curie temperature. Near the Curie temperature (about 200°C), the coercivity drops rapidly by two or three orders of magnitude as the magnetic domain structure becomes disordered. Readout of the recorded information can safely be achieved with a laser beam of about 2 mW power at the disk, a power level which is high enough to provide good signal strength at the detectors, but low enough not to affect the recorded information because any resulting media heating is far below the Curie threshold.

During readout, the magnetic state of the recorded bits of information is sensed through the polar Kerr effect by a low-power linearly polarized readout beam. In this effect, the plane of polarization of the light beam is rotated slightly (0.5°) by the magnetic vector. The direction of rotation, which defines whether the bit is a "one" or a "zero," is detected by the readout detectors and channel. Although the tiny amount of Kerr rotation results in a very small signal modulation superimposed on a large dc bias, the technique of differential detection permits an acceptable signal-to-noise ratio (SNR) to be achieved. This is discussed in the section on channel issues.

Early MO recording media, such as manganese bismuth (MnBi) thin films [3], were generally crystalline in nature. The magnetic domains, which followed the crystalline boundaries, were irregular in shape [4]. The crystalline nature of the films caused optical scattering of the readout signal, and the irregular domains led to noise in the recorded signal. The combination degraded the SNR sufficiently to make polycrystalline magneto-optic media impractical. The discovery at IBM of magneto-optic materials based on the rare earth/transition metal (RE/TM) alloys [1] in 1976 provided a practical material system for rewritable magneto-optic recording. These materials were amorphous, and thus allowed acceptable signal-to-noise ratio to be obtained. Most commercial magneto-optic films today are based on terbium iron cobalt (TbFeCo).

Optical head

The purposes of the optical head are to transmit the laser beam to the optical disk, to focus the laser beam to a diffraction-limited spot, and to transmit readout signal information from the optical disk to the data and servo detectors. The design of the optical head must accommodate a number of performance trade-offs and take into account the imperfections and variability of cost-effective optical components and lasers. In this section, the design and functionality of an optical head are examined. The section begins with a discussion of the laser diode, since several of the optical components are chosen to match the properties of the laser diode.

The laser diode is the key component in optical storage, whether the recording technology is magneto-optic, ablative WORM, or phase-change. The current optical drives use gallium arsenide semiconductor lasers that emit in the 690-nm wavelength region. The lasers are typically rated to have a maximum continuous output power in the 40-mW range (the reasons for the choice in power rating are discussed later). In order to ensure good wavefront quality, the lasers must be index-guided.

In a laser diode, light is emitted from the facets, which are the cleaved ends of the waveguide region of an index-guided laser. The facet dimensions are small enough (of the order of a few micrometers) for diffraction to take place as light is emitted from the facet [5]. As a result, the output beam has a significant divergence angle. In nearly all commercial laser diodes, the width of the facet (i.e., the dimensions parallel to the p-n junction plane) is typically about 10 μm, and the height (or the direction perpendicular to the junction plane) is about 1 μm. As a result of this aspect ratio, the divergence angles are unequal in the directions parallel (about 10°) and perpendicular (about 25°) to the laser junction. The spatial profile of the laser beam some distance from the laser is thus elliptical. A further characteristic of the output beam from a laser diode is that it has astigmatism. That is, the focal point of the component of the beam parallel to the junction is different from that of the component of the beam perpendicular to the junction. The design of the optical components to compensate for this ellipticity and astigmatism is discussed below.

The basic layout of the optical head is shown in **Figure 1**. The laser is mounted in a heat sink, which, along with the head materials and the interface, is carefully chosen to achieve an athermal response. This is to prevent any alignment shifts resulting from thermal expansion. The output from the laser diode is collimated by lens 1. A prism-like optical element called a "circularizer" is then used to reduce the ellipticity and astigmatism of the laser beam.

The beam then passes through a polarizing beam splitter, which reflects part (30%) of the beam toward a detector and transmits the rest toward the disk. The output of the laser is linearly polarized in the direction parallel to the junction (which is referred to as P-polarization). The ratio of the intensity of the P-polarization component of the emitted light to the ratio of the intensity in the S-polarization component (linearly polarized in the direction perpendicular to the junction) is >25:1. This polarizing beam splitter is designed to transmit 70% of the P-polarized light and 100% of the

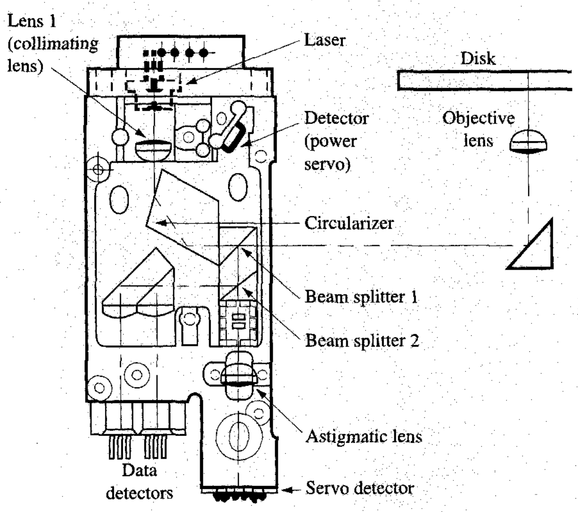

Figure 1

Drawing of the optical head in a typical magneto-optic disk drive. Shown is a split optics design that consists of a fixed set of components (laser, detectors, and polarizing optics) and a movable system consisting of a beam bender and an objective lens. Reprinted from *Laser Focus World*, Vol. 30, No. 1 (January 1994), p. 123. Copyright PennWell Publishing Company. Used by permission.

S-polarized light. The light that is reflected is incident on a light detector which is part of a power servo loop designed to keep the laser at a constant power. Without a power servo loop, the laser power will fluctuate with time as the laser junction heats up, which can adversely affect read performance.

The beam transmitted by the beam splitter travels to a turning (90°) mirror, called a beam bender, which is mounted on a movable actuator. During track-seeking operations, this actuator can move radially across the disk. The beam reflected by the turning mirror is incident on an objective lens (also mounted on the actuator) which focuses the light on the disk. This type of optical head design, in which the laser, the detectors, and most of the optical components are stationary while the objective lens and beam bender are movable, is called a split optics design. In early optical drive designs, the entire optical head was mounted on an actuator and moved during seeking operations. This led to slow seek times (200–300 ms) because of the mass on the actuator. A split optics design, which is possible because coherent light can be made highly collimated, lowers the mass on the actuator and thus allows much faster seek times (20 to 50 ms).

The size of the focal spot formed by the objective lens on the disk depends on the numerical aperture (NA) of the lens and the amount of overfill of the lens (i.e., the amount by which the diameter of the incident collimated beam exceeds the aperture of the objective lens). The numerical aperture is given by $NA = n \sin \theta_{max}$, in which n is the refractive index of the lens and θ_{max} is the incidence angle of a light ray focused through the margin (or rim) of the aperture stop of the lens.

The beam of light incident on the objective lens usually has a Gaussian electric field (or irradiance) profile. The profile of the focused spot is a convolution of the incident intensity profile and the aperture of the objective lens. A relationship exists among the Gaussian profile of the incident beam, the relative size of the objective lens aperture, and the focused spot size and intensity profile [4, 6, 7]. The net result is that overfilling the lens aperture will reduce the size of the focused spot at the cost of losing optical energy outside the aperture and increasing the size of the side lobes. For large amounts of overfill, the side lobes can be large enough to cause crosstalk in the readout signal. Optimization of the amount of overfill [4] yields an approximate spot diameter of

$$D = 1.18\lambda/NA, \tag{1}$$

in which λ is the wavelength of light and NA is the numerical aperture of the objective lens. The depth of focus, z, of the focal spot is given by

$$z = 0.8\lambda/(NA)^2. \tag{2}$$

The depth of focus defines the accuracy with which the objective lens position must be maintained with respect to the disk surface. The smaller the depth of focus, the less tolerance the system has for media tilt, and the more difficult the job of the focus servo system. Thus, trying to reduce the spot size (always a goal, since a smaller spot allows a higher storage density) by increasing the NA of the lens becomes impractical beyond an NA of about 0.6. At present, NAs of objective lenses in almost all optical drives are about 0.55.

The objective lens also acts as a collector lens for the light that is reflected from the disk. This reflected light is used for the servo systems and, during reading, contains the readout information. The reflected light follows the incident path up to the fixed optical element. Beam splitter 1 reflects all of the S-polarized light and 30% of the P-polarized light in the direction of the servo and data detectors. The portion of the light that is transmitted by the beam splitter is, unfortunately, focused by the collimating lens back into the facet of the laser. This feedback light causes a number of problems in the laser, even though the net amount of feedback does not exceed about 7% of the output light for most magneto-optic media.

For most lasers, the effects of feedback are to decrease the threshold and increase the slope of the power–current

(PI) curve. These effects are not really problematic; however, in some quantum well lasers, we have seen that the effects of optical feedback can cause severe problems. As shown in **Figure 2**, the presence of optical feedback causes the PI curve to roll over (the power at which roll-over occurs is inversely proportional to the amount of feedback). This phenomenon could be a serious problem if the roll-over point occurs below the maximum required erase or write power. Lasers have to be screened by the vendor to ensure that this does not happen. The optical feedback can also cause the laser to jump randomly between two or more cavity modes, which causes a random amplitude fluctuation in the output. This amplitude noise can be a serious problem, so techniques such as injection of high-frequency current [8] must be used to control the laser noise.

The light reflected by beam splitter 1 is further split by beam splitter 2 into servo and data components. The light reflected by the beam splitter is incident on the data detectors. Two detectors are used to implement what is known as differential detection. This is discussed further in the section describing the read process and data path. The light transmitted through beam splitter 2 is incident onto a special multi-element servo detector and is used to generate the servo signals. The mechanism by which these signals are generated is the subject of the section on the servo system. It is not by any means an overstatement to point out that the development of high-quality servo systems has played a crucial role in making high-capacity optical disk drives a reality.

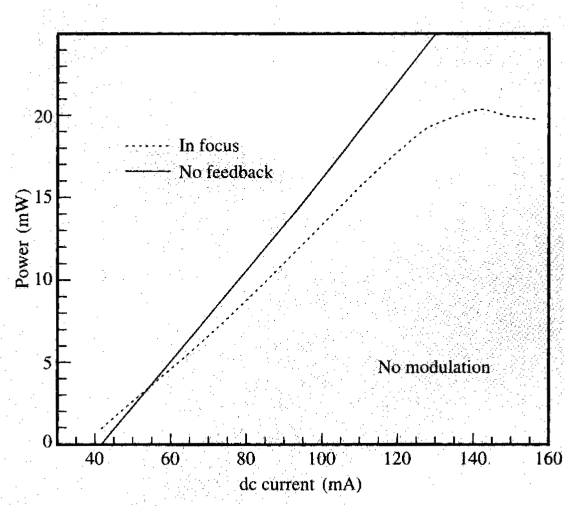

Figure 2

Effects of feedback on the PI (power-current) curve of a quantum well laser. When there is no feedback, the PI curve is linear with current, but when the objective lens is in focus, and there is feedback into the laser cavity, the PI curve rolls over. This phenomenon, which has not been observed in nonquantum well lasers, is troubling in that the laser may never reach its required operating power because of feedback.

Servo

The great advantage of disk storage over tape is the ability to quickly access all of the data on the surface of the disk. Data are typically stored in concentric tracks, which are further divided into sectors within a track. To access a given sector of data, the read/write element must be moved to the desired track and positioned there for the desired information to pass over the head as the disk rotates.

Low-density (e.g., 100 tracks per inch) disk drives that use removable media can control the position of the read/write element by reference to some datum in the drive itself. The position of the disk on the motor spindle changes slightly each time the disk is inserted, and there are additional tolerance variations from drive to drive. These variations limit the accuracy with which a track can be followed, and therefore the number of tracks per disk.

In higher-density disk drives, it is necessary to control the position of the read/write element by reference to the disk itself. For example, non-optical disk drives prewrite each disk with position control signals and track and sector information; optical disks have a prestamped continuous groove (as in a phonograph record) to provide information on the relative track location, as well as embossed data pits for track and sector identification.

Figure 3 shows a block diagram of an optical drive servo control system. The light beam from the laser passes through a series of transducers that modify its position and allow it to focus on the disk. The return beam contains information on the focus and position of the spot. The feedback signals derived from the servo detectors allow the system to maintain control of the beam.

- *Position control of the laser beam*

The extremely high track densities (18000 tpi) on optical disks require that the laser spot position be controlled to within a fraction of a micrometer. Additionally, the spot must be capable of being moved to any track on the disk. An actuator large enough to move across the entire disk surface will be too massive to quickly adjust to the rapid changes in the track position as the disk spins. Consequently, a compound actuator is used to control the radial position of the laser beam on the disk. A fine actuator with a very low mass is used to adjust the spot position over a limited range, but with great speed. A coarse actuator responds more slowly, but it has a much wider range of motion in order to position the spot at any radius on the disk surface.

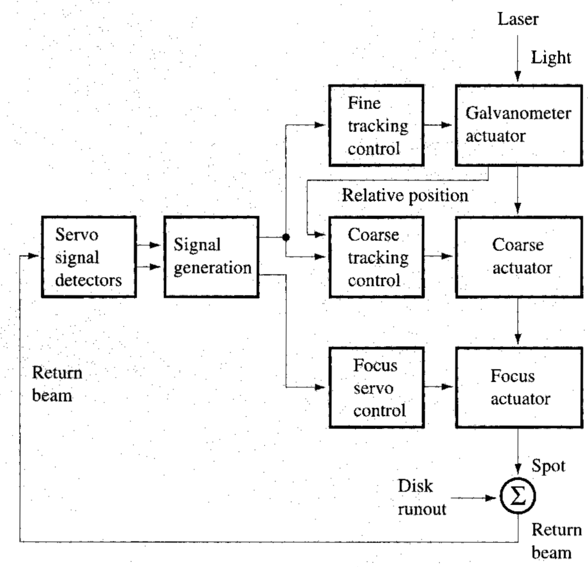

Figure 3

Block diagram of the servo control system for an optical disk drive. Information on the position and focus of the beam is derived from the returning light and used to control the beam through a series of transducers.

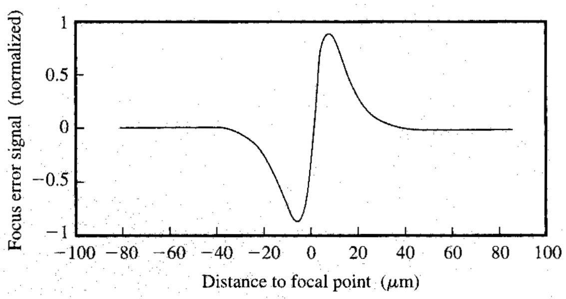

Figure 4

The focus error signal is generated by an astigmatic detection method, which indicates the direction and amount of defocus in the laser spot.

The spot size in optical recording has a diameter of less than a micrometer. To maintain this small spot size, the position of the objective lens must be constantly adjusted to correct for the axial motion of the disk surface as the disk spins. The lens position is controlled by a servomechanism in order to minimize the misfocus. This system is described below, followed by an explanation of the laser beam position control system used for seeking and track following.

Focus control

The focus control system requires a feedback signal that accurately indicates the degree and direction of focus error. To generate a focus error signal, an astigmatic lens is used to focus onto a quadrant detector a portion of the light reflected from the disk. In perfect focus, the focal spot is equally distributed on the four elements of the quad detector. However, if the lens is not in focus, the focal spot on the detector is elliptical because of the optical properties of the astigmatic lens. The unequal distribution of light on the detector quadrants generates a focus error signal (FES). This signal is normalized with respect to the light level to make it independent of laser power and disk reflectivity. An example of the FES is shown in **Figure 4**. This curve is often called the "S-curve" because of its shape.

Ideally, the FES would indicate the error in lens position throughout its entire range of motion. This would provide the control system with continuous information on the direction and amount of lens movement required to bring the beam into focus. However, the astigmatically generated FES is less than ideal. Although the focus actuator can move the lens over a range of several millimeters to match the disk movements during rotation, the FES is valid only for about one thousandth of this range. Figure 4 shows the FES as it relates to the relative position of the disk and lens. Because of the signal's narrow active range, a special focusing operation is needed. The system must move the lens to hunt for the "S-curve"; once it is detected, the servo uses this signal to provide feedback to the actuator, thereby locking the system in focus.

The focus actuator typically consists of an objective lens positioned by a small linear voice coil motor. The coils are preferably mounted with the lens to reduce moving mass, while the permanent magnets are stationary. The lens can be supported either by a bobbin on a sliding pin or by elastic flexures. The critical factors in the design are range of motion, acceleration, freedom from resonances, and thermal considerations.

Once the spot is focused on the active surface, it must find and maintain position along the desired track. This is the role of the tracking servo.

Tracking control

The same quadrant detector that is used to generate the focus signal can be used to generate the tracking error signal (TES). The beam returning from the disk contains first-order diffraction components; their intensity depends on the position of the spot on the tracks, and varies with the light falling along one axis of the quadrant detector.

The TES is the normalized difference of the current from the two halves of the detector, and it peaks when the spot passes over the cliff between a land and a groove (**Figure 5**).

The TES gives a valid indication of the lateral distance of the spot from the center of the track, but only within a limited range. As the beam moves onto the adjacent grooves, the slope of the signal reverses (Figure 5). This cyclical nature of the TES means that the direction of the spot over tracks cannot be derived from this signal alone. Consequently, some systems also make use of the total light intensity to provide a signal in quadrature to the TES. If provided, the quadrature signal in combination with the TES indicates the direction of the spot motion relative to the disk at any time. This additional information provides superior control when terminating a seek.

As mentioned above, the tracking mechanism consists of a coarse and a fine actuator. In high-performance drives, the coarse actuator consists of a linear voice coil motor driving a rail-mounted carriage. The fine actuator acts to produce small radial displacements of the laser spot on the disk. This is usually done by generating an independent radial motion of the objective lens. Because the beam overfills the lens, moving the lens results in a corresponding motion of the spot on the disk. A small voice coil motor with the coils attached to the lens assembly provides the driving force. Typically, a relative position error (RPE) sensor indicates the displacement of the spot with respect to the coarse carriage. This signal is useful for coordinating the motion of the coarse and fine actuators. The actuator is shown in **Figure 6**.

The compound tracking actuator configuration has advantages over a single actuator not only in track-following performance, but also when seeking between tracks. Small track displacements can be handled primarily with the fine actuator, so reading/writing operations can occur before the coarse actuator settles into position. Concluding a long seek operation is similar: Seek time is limited by the coarse actuator accelerations, but the seek can be concluded as soon as the fine actuator locks the spot position onto the track. Several different control architectures are appropriate, such as state space multiple input multiple output (MIMO), frequency-selective crossover, or master–slave configurations.

We have described the basic servo system above. There are, however, certain challenges to the implementation that arise in practice. Some of these are discussed below.

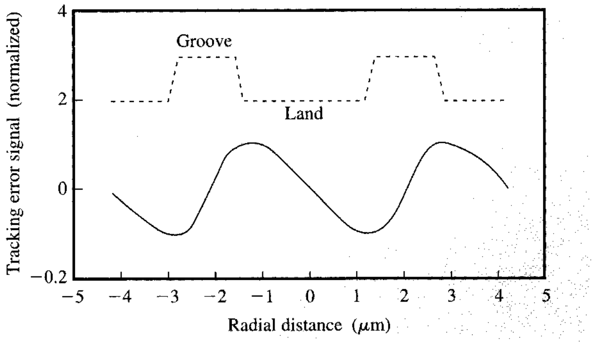

Figure 5

The tracking error signal is derived from the interaction of the laser spot with the grooves of the disk. The information is used to keep the spot centered on the track as the disk spins, and to seek between tracks.

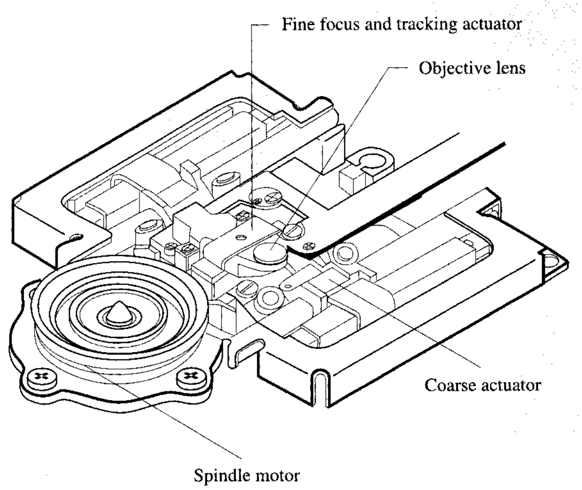

Figure 6

Schematic of a typical actuator system in an optical drive. The focus actuator and the compound fine and coarse tracking actuators adjust the position of the objective lens. The coarse tracking actuator is rail-mounted, and the focus actuator is a small voice coil motor.

- *Servo implementation challenges*

Servo signal quality

Any imperfections in the wavefront impinging on the disk surface can result in unwanted variations in the servo signals. An example is feedthrough from the tracking signal into the focus signal. The presence of tracks on the disk disturbs the focus error signal, producing a false focus shift. This systematic error in focus must be removed by measuring the error in the FES when in best focus, and then electronically subtracting the offset from the FES. Best focus is determined by either looking at the amplitude of the data signal from the disk, or finding the focus offset where the TES is largest. The maximum read signal occurs when the laser spot is smallest along a track; the maximum TES occurs when the spot is smallest across the track. Because of astigmatism in the beam, these points often do not correspond.

The unwanted feedthrough of tracking information into the focus signal also affects seek operations. As the speed of the actuator changes during a seek, the frequency of the feedthrough into the focus signal changes. This false signal can easily saturate the electronics as it passes through certain frequency bands where the gain of the focus compensation is high. Consequently, the electrical characteristics of the focus servo must be changed during seeking in order to prevent loss of focus and degradation of servo performance.

Both the FES and the TES are affected by surface debris on the disk, by the presence of prewritten or ablatively written data, and by optical variations of the media material. The servo system must incorporate special means for eliminating the spurious signal components or their effects.

Variations due to removable media

The use of a removable medium certainly gives a great advantage to magneto-optical disk drives, allowing the transportation of large amounts of data in a very dense package. But the use of removable, inexpensive media produces several challenges for the servo systems: First, the chucking of the disk on the spindle produces variations in concentricity, resulting in a radial track motion with each rotation of the disk. Second, any debris or imperfections of the chucking mechanism or disk will tilt the media; the disk rotation converts this tilt to cyclical surface motions spanning thousands of micrometers. These two effects are present in all disk drives, but the combination of high densities and removable plastic media demands high levels of performance from the focus and tracking systems in order to follow these motions and maintain the spot position within a fraction of a micrometer.

The ability to accurately follow the radial and axial motions of the spinning disk results directly from the quality of the focus and tracking actuators. The servo system must have high bandwidths in order to reject the errors due to shock, vibration, and media imperfections. The limitation to achieving high bandwidth is usually the resonance modes of the actuator. As the actuators are reduced in size, the frequencies of the resonances become higher and the achievable bandwidth of the system rises. The smaller mass of these actuators permits them to accelerate faster with less power, improving seek time as well.

Seek time reduction

Hard disk magnetic drives can improve their data access times by distributing the data among several disk platters in one assembly. A multiple arm head can access any of the disks; the same quantity of data can be accessed by a shorter mechanical stroke in a shorter time from multiple disks than from a single disk. The removable nature of optical media prevents a similar strategy from being easily implemented: In optical drives, seek time improvements are usually the result of improving the system components and reducing limitations.

There are several constraints to the seek performance of a system. Most important is the maximum acceleration and deceleration rate of the (coarse) actuator. The maximum acceleration from the coarse actuator is a function of the largest current that can be applied, the force constant, the mass, and the ability of the actuator to dissipate heat. The performance is also constrained by the highest track-crossing velocity that the optics, detectors, and preamplifier can resolve.

A primary technique for improving the acceleration of an actuator is to reduce its moving mass. Early optical drives carried the entire optical head assembly with the coarse actuator. A significant mass improvement resulted from leaving the bulk of the optical elements fixed, and moving only a prism and the objective lens. As described earlier, the objective lens was moved axially with the disk for focusing, and moved radially for fine tracking.

The moving mass can be further reduced by removing the mechanism for fine tracking from the objective lens. Fine tracking is then provided by a tilting mirror galvanometer mounted with the other fixed optical elements. There are several advantages besides the lower moving mass. First, there is better mechanical isolation between the two systems. Second, the galvanometer is small and capable of excellent servo bandwidth. Finally, it simplifies the focus actuator construction. There are however, several disadvantages: First, the active range of the tilting mirror is only about 10 μm, as opposed to 100 μm for the radially moving lens method. This requires that the coarse servo system be capable of following the track to within this range. Finally, the beam returning to the TES detectors is offset with this tilt, and the amount of offset varies with coarse position. This offset in the TES with galvanometer angle and coarse position complicates the task of accurately following the track center.

The optical head, servo, and actuators that we have discussed so far cover the essential optomechanical part of the drive. Next we consider the data path electronics and the mechanism of formatting, recording, and reading out data.

Optical recording and channel

In many ways, the recording channel/data path for an optical storage device is identical to that for tape, floppy, and hard disk drives (**Figure 7**). The data are sent over a standard interface to the drive, which encodes and writes the data to the disk. On readback, the signal is read, equalized, and the data detected. However, many differences do exist in the actual implementation, for instance in laser power control and embossed sector headers. The following is a brief description of the blocks that constitute an optical storage device. Actual implementations differ from drive to drive, and items such as error-recovery procedures and data-recovery procedures are beyond the scope of this work.

The process of storing information begins with data being sent over a standard interface to the optical drive. The majority of rewritable optical drives use the small computer systems interface (SCSI) as the standard protocol for sending command information and data to the drives. As the data are received by the drive, they are stored in buffer RAM. This storing of data continues until sufficient data exist in the buffer to start writing the data to the disk.

As of the writing of this paper, all magneto-optical drives use fixed-block architecture and embossed headers. The term *fixed-block architecture* means that data are divided into discrete blocks (sectors) of either 512 bytes or 1024 bytes prior to encoding and storing on the disk. For example, in order to store a 2000-byte record on a 1024-byte-per-sector disk, two sectors would be used.

Another distinguishing feature of optical storage is that the disk is hard-formatted at the time of manufacture. The position of the sector and its address is embossed into the substrate during molding. The fact that the sector information is hard-coded into the disk greatly increases the reliability of the system, but at a cost in flexibility. A disk meeting a given standard with a fixed sector size cannot be used as any other type of disk.

- *Write process and data path*

After the data have been segmented into discrete sector sizes, the error correction and control (ECC) bytes are added. The high-power Reed–Solomon ECC of optical drives reduces the error rate from 1×10^{-5} bytes in error per byte read to better than 1×10^{-13} bytes in error per byte read. Data are modulation-encoded following the addition of the ECC information. Optical disks use run-length-limited (RLL) modulation codes such as 0,3 and 2,7 to improve detection. Disks using 1× and 2× standards are encoded using the 2,7 code, whereas the new 4× disks will use a 1,7 code to compensate for the higher density.

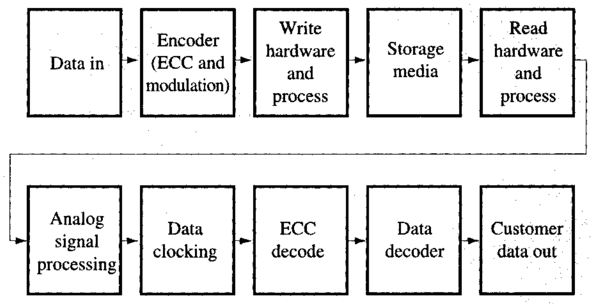

Figure 7

Model of the general digital recording channel.

Finally, special characters are inserted into the bit stream. These special characters include a synchronization character to tell where the data start, specific bytes to phase-lock the clock, and resynchronization characters to help align the data for ECC purposes. In total, 1024 bytes of user data are stored as 1360 bytes on the disk [4].

The output of the encoder is a data bit stream to be written on the disk. However, this bit stream cannot be fed directly to the laser for writing because of the thermal nature of the write process. In pulse position modulation (PPM), the information is stored in the position of a mark center (a mark is made for each "one" bit). At low rpm, a mark is made using a short, high-power pulse. However, at high rpm, such a strategy suffers from thermal interference. The heat from one pulse influences the size and position of the next mark written because not enough time has elapsed for the energy deposited on the disk to dissipate. In this case, an isolated mark is written at one power, and if the next mark is within a certain distance (time or space), a lower power is used, compensating for the elevated media temperature.

The problem with thermal interference is greater in pulse width modulation (PWM). Not only are marks written at various distances from one another, but the marks vary in length because information is encoded in the edge position. To compensate for this interference, a complex laser waveform must be generated in which the widths and powers are varied within the mark.

Laser power to the disk must be controlled very tightly. Writing is accomplished by delivering high laser power to the media, changing the physical characteristics of the disk to store the encoded bit stream. Powers at the disk range between 8 and 20 mW, depending on media type and the linear velocity of the disk. This power at the disk translates to between 16 and 40 mW at the laser because

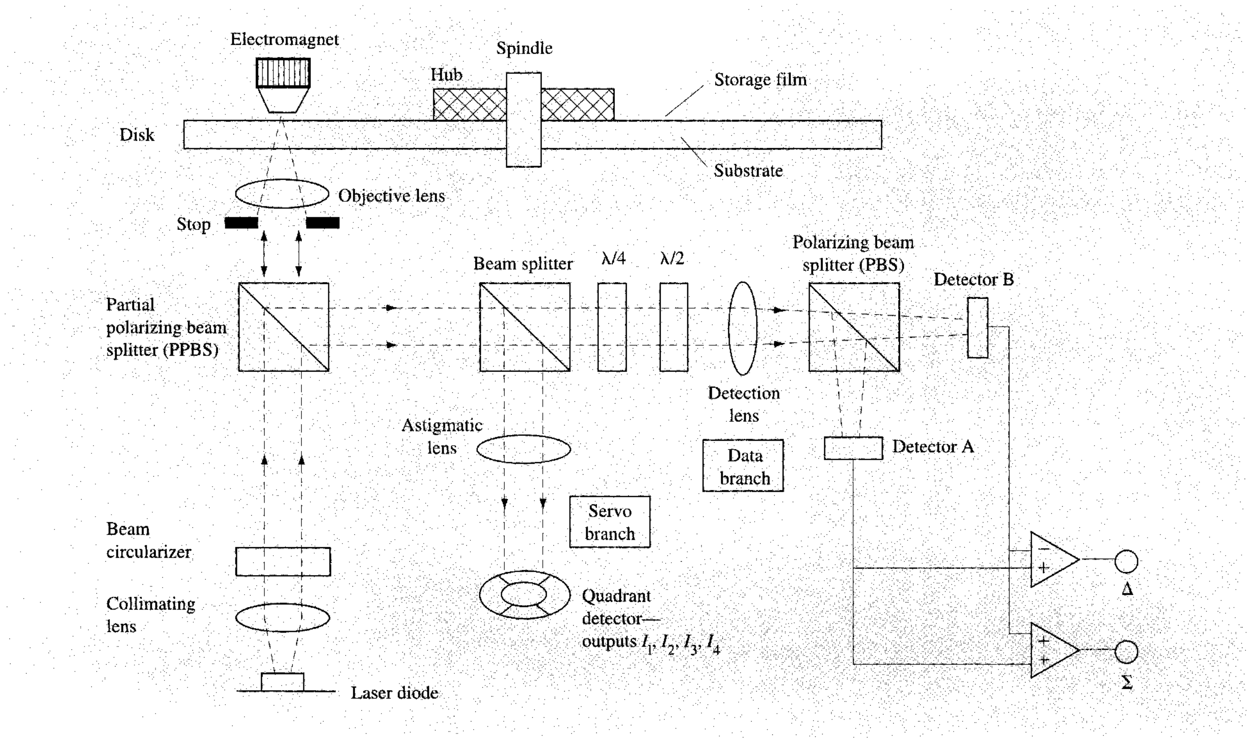

Figure 8

Block diagram of a typical optical head showing the differential detection function.

of head efficiency. However, the laser should be rated at high power to allow margins for degradation of the laser and contamination of the optics and media. The laser pulses also have very fast rise and fall times (<3 ns) to create very sharp thermal gradients, reducing write noise [9].

Laser power also must be tightly controlled during readback. The laser is an active probe: If the read power is too low, the SNR will be too low for adequate data reliability; if the read power is too high, the recorded information is in danger of erasure. Compact disk (CD), optical read-only memory (OROM), and write once, read many (WORM) disks usually require 0.5 mW of laser read power. Magneto-optical disks usually require between 1.5 and 2.0 mW of laser read power. Another responsibility of the laser driver is to supply a high-frequency modulation (HFM) current to the laser. This current suppresses mode-hopping and feedback noise by forcing the laser to be multi-mode. In general, increasing the high-frequency injection current reduces the laser noise. As a practical matter, however, the injection current cannot be made arbitrarily large, because it may then violate FCC limits on allowable radiation from computer accessories.

- *Read process and data path*

On readback, the light reflects from the disk and travels to a set of photodetectors. For WORM and OROM disks, the signal is intensity-modulated, and thus the sum of the currents from the two detectors is amplified prior to processing. The contrast between marks and no-mark is quite large for these two media types, leading to high signal-to-noise ratios.

For magneto-optic media, the readout process depends upon the polar Kerr effect. Linearly polarized light is incident on the medium, and its polarization is rotated by $\pm\theta_k/2$, the sign depending upon whether the medium is magnetized up or down. The magnitude of θ_k is typically $0.5°–1°$, which produces a small signal amplitude compared to the other forms of optical recording. Because laser noise and reflectivity noise are large compared to the signal, satisfactory SNRs are difficult to achieve. What makes MO storage systems possible is differential detection.

The output signal in an MO recording system is the signal from the light falling on one detector minus the signal from the light falling on the other detector (Figure 5). This "difference amplifier" combination represents a classic differential detection system. By placing a polarizing beam splitter at 45° to the incident polarization, the two data detectors receive the signals* [10]:

$$d_1 = I_0/2[\cos(\theta_k/2) - \sin(\theta_k/2)]^2 \simeq I_0/2(1 - \theta_k), \quad (3)$$

$$d_2 = I_0/2[\cos(\theta_k/2) + \sin(\theta_k/2)]^2 \simeq I_0/2(1 + \theta_k), \quad (4)$$

where $d_{1,2}$ refers to the detector signals, I_0 is the incident intensity, and $\theta_k/2$, the rotation angle, is assumed to be small. The readout signal is taken as

$$s = (d_1 - d_2)/(d_1 + d_2). \quad (5)$$

This signal is relatively insensitive to intensity noise either in the laser or the reflectivity of the disk. However, it is very sensitive to polarization noise, which may be introduced by polarization-sensitive diffraction, by substrate birefringence effects, by inhomogeneities in the magneto-optic films, or by other polarization-sensitive components [10].

In **Figure 8**, the currents from the diodes (Detectors A and B) are amplified and then subtracted in the differential (Δ) amplifier. The diode currents are also added in a summing amplifier (Σ) and used for intensity-modulated reading. This configuration has at least two advantages over a totem-pole configuration: More flexibility in diode biasing/placement is achieved, and the gain for each leg can be independently adjusted. However, a lower-noise configuration exists. In a totem-pole configuration, the currents from the photodetectors are first subtracted directly and then amplified. In this case, the electronic noise is halved (one amplifier, not two), but the capacitance of the circuit may be slightly higher. The choice of configuration is another trade-off in drive design.

In writing, a string of "one" and "zero" bits is sent to the encoder and then to the laser driver, and is then written on the disk. On readback, the signal has been corrupted by noise and distortion, and by intersymbol interference. In short, the read signal does not accurately resemble the stored information. Optical drives use equalization in an attempt to restore waveform quality. Equalization compensates for signal degradation by increasing the high-frequency components to improve resolution and restricting bandwidth to reduce noise. One advantage of optical recording over magnetic recording is that the phase of the transfer function is linear. Equalization improves the signal-to-noise ratio at the data detector, allowing significant improvements in capacity.

Another issue (besides the relative signal and noise spectra) in equalization is the detection technique required for the modulation used. In pulse position

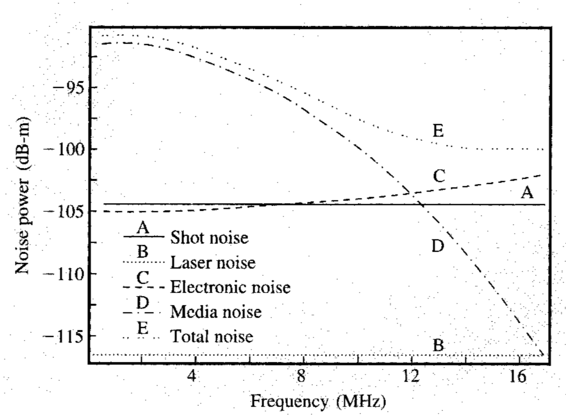

Figure 9

Noise power as a function of frequency measured at the pre-amp output.

modulation (PPM), the information is stored in the distance between mark centers. All marks are the same size, but vary in position. The current MO standards (90 mm 1× and 2×, 130 mm 1× and 2×) all use this type of modulation. These systems use PPM because variations in write power, media thermal sensitivity, and defocus do not change the location of the center of the mark, only the SNR of the system. To detect a "one" or "zero," the centers of the marks are detected by differentiating the signal, turning peaks into zero crossings. Differentiating in the time domains is equivalent to multiplying by the frequency in the frequency domain. This multiplication greatly reduces low-frequency noise, but increases high-frequency noise. **Figure 9** shows typical noise traces at the output of the preamp. **Figure 10** shows the effect of differentiating the noise floors; the net improvement in system SNR exceeds 3 dB. Selectively increasing boost while decreasing bandwidth further improves SNR.

The 130-mm and 90-mm drives based on 4× standards use pulse width modulation (PWM). In PWM, information is stored in the position of the edges of the marks, as opposed to the centers as in PPM recording. Because a mark has two edges but only one center, the coding efficiency is significantly higher for PWM recording, if the marks can be written satisfactorily. The shape of equalization for PWM depends on which detection method is chosen. If a threshold channel (dc channel) is used, edges are detected as threshold crossings. Since these channels need dc response, the noise prior to equalization is the same as that from the pre-amp. (Another difference between an optical head and a magnetic head is that an optical system does pass dc.)

* T. Strand, Almaden Research Laboratory, private communication, 1993.

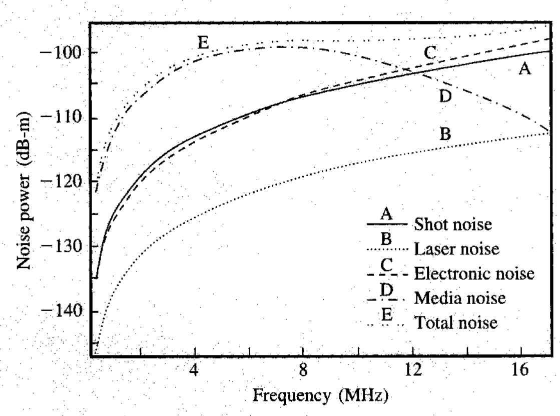

Figure 10

Noise power as a function of frequency measured at the differentiator output.

An alternative detection scheme is to double-differentiate the signal and then detect zero crossings. This detection method (ac channel) has a tremendous amount of high-frequency boost, which can degrade system SNR, especially in devices with appreciable electronic noise. The advantage of an ac channel is that the system is less susceptible to baseline wander, birefringence variation, and certain types of defects.

Spectral shaping helps achieve optimal performance. Equalization involves trade-offs among ideal readback waveforms, noise, and defect performance. In the final analysis, equalization improves data reliability, allowing the high linear density of current products.

Subsequent to equalization, the analog readback signal must be converted into digital information. The function of the data detector is to convert the signals from the equalizer into asynchronous binary data which can later be used to synchronize a phase-locked loop (PLL) to generate the resultant synchronous data. This is generally accomplished by first generating a qualification window by threshold-detecting the peaks of the normal equalizer output. This qualification window is particularly important with long wavelengths to enable discrimination between the differentiated signal and unwanted noise-induced zero crossings. Since the peaks of the normal signal correspond to the zero crossings of the differentiated signal, the qualification window should be centered with respect to the zero crossings. The width of the qualification window is a function of the threshold setting (usually 50%) and the dynamics of the tracking threshold.

Because signal amplitude decreases with decreases in recorded wavelength (resolution), and the baseline of the signal also varies, the threshold must be designed to track the signal dynamically while maintaining the desired threshold. This requires a balance between keeping the threshold dynamic enough to track the data and providing adequate smoothing to maintain a good threshold SNR. This generally requires parameter switching to maintain the design point across banded media.

Finally, the digital qualification window is coupled with the zero-crossing detector output of the differentiated signal using the AND logic operation to produce the asynchronous data pulses where each data pulse represents a detected mark (or encoded logical "one"). These asynchronous data pulses are applied to the PLL for generation of synchronous clock and data.

Synchronization of the PLL is performed by affixing a short-wavelength preamble to the beginning of each sector. With the PLL in acquisition mode (wide bandwidth), the PLL acquires frequency and phase-lock during this time. Once phase-lock is acquired, the PLL bandwidth is switched to tracking mode (low bandwidth) so that the clock timing error is insensitive to variations in the data pattern. This synchronous clock is then used to produce the corresponding synchronous data by comparing the input data pulse transition times to a synchronized data detection window. As long as the jitter on the asynchronous data pulse input is less than $\pm T/2$ (T being the clock period), the synchronous data output correctly represents the asynchronous data input. If, however, the jitter exceeds the $\pm T/2$ limit, an error occurs which leads to the function of the ECC decoder.

The process of converting channel bits back into customer data bytes is basically the reverse of the encoding/ECC steps. Processing starts with detection of the sector mark. After reading the sector IDs to ensure operation on the proper sector, the data are clocked into the decoder, which removes the modulation code. The remaining special characters are stripped from the data prior to converting from a serial data stream to a one-byte-wide parallel stream. Next, the data are fed into the ECC alignment buffer, and any errors are corrected. Defects up to 40 bytes long can easily be corrected.

Once data have been read from the disk, they are stored in buffer RAM. This buffer RAM is used as a holding place until sufficient data exist to be transferred over the interface. Since data are read off the disk at a rate significantly slower than most of the standard interface protocols, it is important to buffer large packets of data and burst these data packets to the host. By accurately controlling these packets, the SCSI bus will be made available for other devices sharing the bus.

This concludes a discussion on the technology currently present in the optical drives used in IBM optical library products. Optical drives are still a relatively new commercial technology (the first rewritable optical drives were commercially developed and marketed less than a decade ago). A considerable amount of technical growth is possible for optical drives. Some of the potential improvements in optical drive technology are discussed in the next section.

Future technology

There are three primary directions for future work on optical drives: increasing capacity, improving performance specifications (such as data rate and seek time), and reducing the entry cost of the drive. Some of the possible methods for achieving these goals are discussed briefly in the following sections.

• *Increasing capacity*

Arguably, the principal attribute of optical recording has been the high capacities that can be achieved on a single removable and robust platter. Continually increasing the disk capacity is a paramount development goal for optical drive products. A number of ways in which increased capacity may be obtained are outlined below.

Reducing the laser wavelength From Equation (1) we note that the spot size of the beam is directly proportional to the laser wavelength. Reducing the spot size by reducing the wavelength will result in a higher areal density. Current optical products use lasers that emit in the 670-to-690-nm (red) wavelength range. Optical drive manufacturers are awaiting lasers that emit in the blue region of the visible wavelength (430 nm), since optical drives equipped with such lasers will have twice as much storage capacity (about 5 GB on a single 5.25-in. disk platter) as drives with red lasers.

The potential for using blue lasers for optical storage took a major step forward with IBM's invention in 1990 of a high-efficiency frequency-doubled solid-state laser that could be directly modulated at high data rates. In this laser, the output of an infrared pump laser is fed into a mode-matched potassium niobate cavity where frequency doubling and gain can occur [11]. The fact that the laser could be directly modulated (i.e., modulation of the blue output light could be achieved simply by modulating the pump laser) at the kind of data rates found in optical storage was a breakthrough in the use of second-harmonic-generation material for lasers. Using this laser, IBM recently demonstrated magneto-optic recording at a record linear density of 2.4 Gb/in.^{2} [12].

The ideal blue laser source, from the viewpoint of compactness and direct modulation capability, would be a semiconductor laser diode. A research group at 3M Company has made tremendous progress and recently demonstrated a semiconductor laser made from the zinc selenide material system that could emit light in the 430-nm range [13]. However, this laser can operate continuously (CW mode) for any usable length of time only if it is cooled to cryogenic temperatures. Semiconductor blue lasers feasible for use in commercial storage systems are perhaps five years away.

Reduction of the laser wavelength alone will not provide the kind of gains in areal density needed to ensure that optical storage remains competitive with alternate technologies. Thus, other technologies will have to be implemented in conjunction with, or independent of, a reduction in laser wavelength.

Increased NA As noted earlier, the spot size is inversely proportional to the NA of the objective lens, but the depth of focus of the objective lens is inversely proportional to the square of the NA. The depth of focus is an important parameter to maximize, since it sets the acceptable degree of tilt of the media and determines the quality of the focus actuator and servo system (the smaller the depth of focus, the better the actuator and servo must be). As servo systems and actuators become more sophisticated, the use of objective lenses with numerical apertures greater than the present 0.55 becomes more feasible. It is expected that the next jump will be to an NA of 0.62, which will provide a spot 12% smaller in diameter than would an objective lens of 0.55 NA.

Optical superresolution Optical superresolution takes advantage of the diffraction properties of focused beams. For example, if the collimated beam of light incident on the objective lens is in the shape of an annulus, or ring (which can be accomplished simply by blocking a circular region at the center of the beam), the spot size of the focused beam will be smaller than if no light were blocked. What has happened is that energy has shifted from the central or zero order to higher-order side lobes. Provided these side lobes are not too big, this technique is an effective way of obtaining smaller spot sizes [14]. However, blocking a portion of the beam may cause considerable loss of optical power.

Partial response maximum likelihood (PRML) read channel The read channel of a data storage device is the set of electronics that process the detected readback signals from the disk and extract the digital data from the (often noisy) analog waveform. Currently, optical drives use an analog channel to perform this processing, which works well at low linear densities but starts to become ineffective as the linear densities of the recorded marks become much higher than about 40 kilobits per inch. At high linear densities, a phenomenon known as intersymbol interference (ISI) smears the analog waveform and makes it more difficult to distinguish the separate marks,

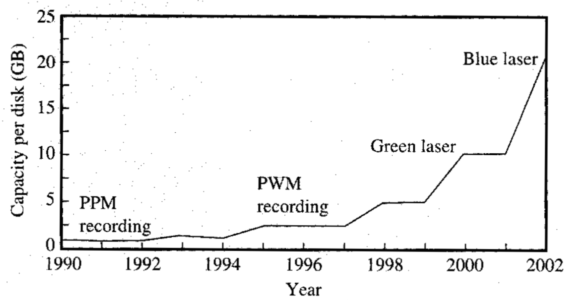

Figure 11

Projected gains in optical storage capacity per 5.25-in. disk cartridge.

especially in the presence of any noise. A PRML channel is a digital channel that uses a sophisticated algorithm to sample the waveform, filter out the noise, and extract the digital information. As such, the PRML channel can allow substantially higher linear densities to be achieved than an analog channel.

Recording on both grooves and tracks Most current optical disks consist of tracks separated by grooves. Data can be recorded in either the grooves or the lands (tracks), though at present most disks use land recording. Clearly, the storage capacity could be doubled if both the grooves and the land regions could be used for data recording. The problem with this otherwise elegant idea is the crosstalk between data on adjacent lands and grooves. Recently, several companies have suggested crosstalk cancellation algorithms that are effective enough to allow adjacent groove and land recording to be accomplished without noise and crosstalk.

An extrapolation of the potential capacity per 5.25-in. optical disk cartridge is shown in **Figure 11**. It is very important when choosing any storage enhancement technology to ensure that backward compatibility is maintained. From IBM's perspective, the drive should be able to read and write one previous generation of disks and read two previous generations. Backward compatibility requirements may limit the adoption of certain combinations of technologies to enhance storage density.

Increasing performance

Two key performance attributes of optical drives that will see improvement in the future are the data rate and seek times. While tremendous progress has been made in improving the seek times and data rates over the years (as shown in Figure 1), this is still an active area of improvement, given that the data rate and seek time are two key performance specifications often used in the competitive comparison of optical drives.

Increasing data rate There are three methods for increasing the readout data rate. One is to increase the linear density of recorded bits, the second is to spin the media faster, and the third is to use sophisticated read cache algorithms. The increase in linear density is a by-product of the increase in capacity, as discussed above. Increasing the rpm of the media (currently of the order of 3600 rpm industrywide) is a more complicated matter than simply spinning the motor faster.

Spinning the disk faster will require improvements in the disk, the drive servo mechanisms, and the electronics. The disk must be made more robust because of the additional mechanical stress on it due to the increased disk acceleration. The disk accelerations in both the radial and axial directions increase in proportion to the square of the rpm. At the higher rpms, keeping the spot in focus and on track becomes more difficult, requiring improvements in servo bandwidth to ensure sufficiently rapid response as time margins shrink. Achieving the clock rates required for the increased data rates will constitute another challenge. Today's 650MB/side disks (2× capacity) spinning at 3600 rpm require a clock frequency of 50 MHz. Future 1.3GB/side (4× capacity) systems spinning at 4800 rpm will have clock rates in excess of 110 MHz. Great care will have to be taken in controlling lead lengths, module placement, and parasitic capacitances. It is well known that readout rates can be improved using a RAM cache. However, it is not just the size of the cache that is important, but also the specific cache algorithm used. Development of advanced cache algorithms is therefore an important area for research.

In the preceding paragraph, we discussed techniques for improving the read data rate. It is also important to consider increasing the write data rate. Some of these techniques such as higher rpm or the use of a cache can also be applied to the increase of the write data rate, but there are areas unique to the write process that can be improved. One of these is the number of passes required to perform a write operation. In magneto-optic recording, the write process takes three passes: On the first revolution or pass, the track to be recorded is erased; on the second pass, the information is recorded on the track; and the third pass is used to verify that the data were actually written on the track. The erase and write passes may be combined by using a technique called direct overwrite (DOW), in which erasure and recording occur on the same pass. This may be accomplished by using two laser beams laterally displaced along the track direction, by using magnetic field modulation [15], or by using light intensity modulation [16]. The latter two techniques are the most likely to be applied, and their implementation is eagerly awaited. The separate verify pass may be eliminated by using a two-beam optical head to implement direct read after write (DRAW).

Improving seek time Continued improvements in seek times will be made possible by continuing the trend toward lower mass actuators. Advances in control algorithms will allow the system to use the full capabilities of this hardware. Both subjects are briefly described below.

- *Lowering actuator mass* One likely method of reducing the moving mass of the coarse tracking actuator is to move the focus actuator from the coarse carriage to the fixed optical elements. An objective lens on the carriage would still be required, but the voice coil motor and flexures could be removed to the stationary assembly for use with a focus adjustment lens. Although this method adds to the complexity of the optical design, it will allow higher accelerations and corresponding reductions in seek times. Further on the horizon are nonmoving actuators, such as those using acousto-optic or electro-optic beam deflection. These types of motionless actuators have the promise of very high bandwidths, but the disadvantage of cost and a limited number of resolvable spots.
- *Advanced algorithms* We discuss below an algorithm that can be used to optimize the seek performance within the hardware constraints.

  The motion of an actuator during a seek can be described by a graph showing the actuator position and velocity at every moment in time (**Figure 12**). The fastest seek possible between two tracks, given acceleration and speed constraints, is described by a unique path on the velocity/position graph. Nonlinearities and disturbances due to shock and vibration will disrupt the seek motion and cause the velocity and position of the actuator to deviate from the optimum path. The function of the seek control system is to return the actuator to the desired path in order to minimize the seek time.

  Many excellent techniques and algorithms exist for exercising this control of the actuator, i.e., determining an appropriate acceleration to apply to the actuator in order to re-intercept the optimum path. Most of these techniques are based on linear control methods. However, the exact solution is a nonlinear equation because of the square root term. Equation (6) shows the exact solution for the case of negligible friction in a sampled system. The calculation is easily implemented by a digital controller; it can usually be done faster than the analogous calculations for the linear control methods.

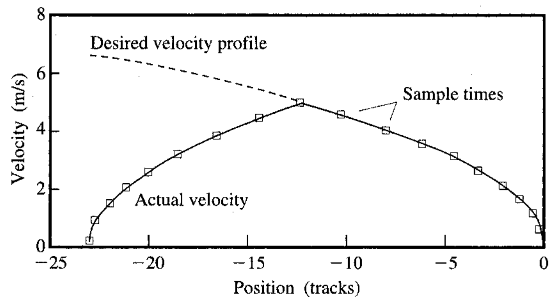

Figure 12

Velocity vs. position graph for a seek. The optimum seek, given the hardware limitations, traces a unique path on this graph. The closer the path can be followed, the better the seek performance.

$$a_{\text{intercept}}(x, v) = \frac{v}{T} + \frac{a_{\text{desired}}}{2} + \sqrt{\left(\frac{a_{\text{desired}}}{2}\right)^2 + 2\frac{a_{\text{desired}}}{T^2}x - \frac{a_{\text{desired}}}{T^2}v}, \tag{6}$$

where $a_{\text{intercept}}$ is the deceleration required to intercept the desired deceleration profile at the next sample point, x is the current distance from the target track, v is the current velocity, a_{desired} is the ideal deceleration profile rate, and T is the sample period.

The seek control system using this direct intercept method is therefore able to measure the position of the actuator at every sample, estimate its velocity, and determine the exact control acceleration necessary to follow or return to the optimum control path. Within the constraints of the drive hardware, this maximizes the seek performance.

- *Reducing cost*

Optical drives provide high-capacity removable storage at a low cost per megabyte, particularly if the storage needs are large enough to require multiple disks. However, a limitation to high-volume sales of optical drives is the relatively high entry price of the drives. For example, at the time of this writing, a 1.3GB magneto-optical drive had a street price of about $1400. Further price reductions are needed before the drive can appeal to a mass market. The prices on 3.5-in. optical drives have fallen rapidly (in the $450 range in mid-1996, down from $800 in 1994). However, these prices are still above the level required for mass penetration. To achieve large volumes, it is estimated that drive prices must drop to a level below

$200. Although optical drives offer the lowest cost per megabyte of any random-access storage device, it appears that many personal computer users place less weight on the cost per megabyte and more weight on the entry price.

Summary

In this paper, we have discussed an overview of the technology found in the magneto-optical drives developed by IBM. The current technology has made possible extremely reliable high-density optical recording. The future of optical recording is technically very exciting, with most of the significant capacity and performance increases still to come.

Acknowledgments

The authors are grateful to Tim Strand of the IBM Almaden Research Laboratory, and to John Bish and Terry Kise of the IBM Development Laboratory in Tucson for their contributions to this paper. The authors are also grateful to Terry McDaniel of the San Jose Laboratory for his help in reviewing this paper, and to John Blair and Ron Stahl of the Tucson Laboratory for contributions of some figures.

References

1. P. Choudhari, J. Cuomo, R. Gambino, and T. McGuire, U.S. Patent 3,949,387, 1976.
2. L. Mayer, *J. Appl. Phys.* **29,** 1003 (1958).
3. D. Chen, J. Ready, and G. Bernal, *J. Appl. Phys.* **39,** 3916 (1968).
4. A. Marchant, *Optical Recording*, Addison-Wesley Publishing Co., Reading, MA, 1990.
5. Zh. I. Alferov, V. Andreev, E. Portnoi, and M. Trukhan, *Sov. Phys.-Semiconduct.* **3,** 1107 (1969).
6. G. Bouwhuis, J. Braat, A. Huijser, J. Pasman, G. van Rosmalem, and K. Schouhamer Immink, *Principles of Optical Disk Systems*, Adam Hilger Ltd., Bristol, England, 1985.
7. J. Goodman, *Introduction to Fourier Optics*, McGraw-Hill Book Co., Inc., San Francisco, 1968.
8. A. Arimoto, M. Ojima, and N. Chinone, *Appl. Opt.* **25,** 1 (1986).
9. P. Asthana and B. Finkelstein, *Proceedings of the IEEE/LEOS (Laser and Electro-Optics Society) Annual Meeting*, Boston, November 16, 1992, p. 331.
10. M. Mansuripur, G. Connell, and J. W. Goodman, *J. Appl. Phys.* **53,** 4485 (1982).
11. W. Kozlovsky, W. Lenth, E. E. Latta, A. Moser, and G. Bona, *Appl. Phys. Lett.* **56,** 2291 (1990).
12. J. Hurst and W. Kozlovsky, *Jpn. J. Appl. Phys.* **32,** 5301 (1993).
13. M. Haase, J. Qiu, J. DePuydt, and H. Cheng, *Appl. Phys. Lett.* **59,** 1272 (1991).
14. T. Milster and C. Curtis, *Appl. Opt.* **31,** 6272 (1991).
15. Y. Murakami, J. Nakajama, M. Mieda, A. Takahashi, K. Ohta, and T. Ishikawa, *Proceedings of MORIS 92*, Tucson, December 1992, p. 341.
16. T. Tokunaga, Y. Nakaki, T. Fukami, and K. Tsutsumi, *Proceedings of MORIS 92*, Tucson, December 1992, p. 357.

Received June 7, 1995; accepted for publication June 24, 1996

Praveen Asthana *IBM Storage Systems Division, 5600 Cottle Road, San Jose, California 95193 (pasthana@vnet.ibm.com).* Dr. Asthana is a senior engineer and investment analyst at the IBM Storage Systems facility in San Jose, California. He received a B.S.E.E. degree from the California Institute of Technology in 1985, an M.B.A. from the University of Southern California in 1987, and a Ph.D. from the University of Southern California in electrical engineering in 1990. He has worked in optical drive development at the Tucson laboratory and was team leader for the writable compact disk (CD-R) project. Dr. Asthana currently works in IBM Finance as an investment analyst for future technology and product investments at the Storage Systems Division.

Blair I. Finkelstein *IBM Tape and Optics Business Unit, 9000 South Rita Road, Tucson, Arizona 85744 (MANGIC at TUCVM2).* Dr. Finkelstein has more than ten years of research and development experience at IBM Tucson in magnetic tape, floppy disks, holographic storage, and optical disk storage. He has received 11 patents, has published 16 times in the storage field, and is on the program committee for the 1996 Optical Data Storage Conference. Dr. Finkelstein received a Ph.D. degree in optical science from the University of Arizona.

Alan A. Fennema *IBM Tape and Optics Business Unit, 9000 South Rita Road, Tucson, Arizona 85744 (kd7fz@vnet.ibm.com).* Mr. Fennema is a senior engineer in the tape and optical product development facility in Tucson. He received a B.S. degree in electrical engineering in 1978, and since then has been employed at the IBM Tucson engineering and development facility. In 1991 he received an IBM Outstanding Achievement Award for his work on optical disk drives. Mr. Fennema has received 21 patents for inventions in optical and tape systems, and has authored numerous publications.