ISSN 0016-2523

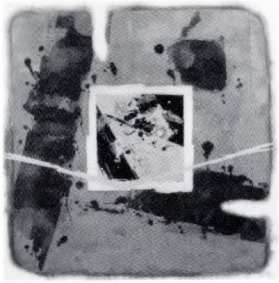
FUJITSU

SCIENTIFIC & TECHNICAL JOURNAL

Winter 1990 VOL.26, NO.4

Special Issue on Fujitsu File Devices

FUJITSU Sci. Tech. J., **26**, 4, pp. 259-438, Kawasaki, February, 1991



The Issue's Cover:
The Universe by Yukinobu KUDOH

The earth is composed of limited, practical elements. But above our heads, the sky extends into infinite space. Toward the greater progress of human beings, we are continuously striving to launch out into the universe. To achieve this dream, many countries are aggressively wrestling with space development.

Fujitsu is also contributing its state-of-the-art technologies to space development. Just a few of our useful R&D applications are research into robot simulators and robots for working in space and development of varied devices for biotechnological experiments carried out in space. Fujitsu is working toward the future of mankind, the earth, and the heavens.

FUJITSU Scientific & Technical Journal is published quarterly by FUJITSU LIMITED of Japan to report the results of research conducted by FUJITSU LIMITED, FUJITSU LABORATORIES LTD., and their associated companies in communications, electronics, and related fields. It is the publisher's intent that FSTJ will promote the international exchange of such information, and we encourage the distribution of FSTJ on an exchange basis. All correspondence concerning the exchange of periodicals should be addressed to the editor.

FSTJ can be purchased through KINOKUNIYA COMPANY LTD., 3-17-7 Shinjuku, Shinjuku-ku, Tokyo 160-91, (Telex No. 2424344, answerback KINOKUNI J).

The price is US $7.00 per copy, excluding postage.

FUJITSU LIMITED 1015 Kamikodanaka, Nakahara-ku,
Kawasaki 211, Japan
Cable Address: FUJITSULIMITED KAWASAKI
Telephone: +81-44-777-1111
Facsimile: +81-44-754-3562

Printed by MIZUNO PRITECH Co., Ltd. in Japan

FUJITSU

SCIENTIFIC & TECHNICAL JOURNAL

Winter 1990 VOL.26, NO.4

Special Issue on Fujitsu File Devices

CONTENTS

Featuring Papers

UDC 681.327.634:681.515
FUJITSU Sci. Tech. J., **26**, 4, pp. 261-270(1991)

File Technology Overview

• Minoru Fujino • Junshiro Sugihara • Masao Suzuki

File devices are used to store and retrieve the huge amounts of information handled in computer systems. A memory hierarchy consists of magnetic disks at the center, surrounded by various devices that focus on performance, long-term storage or economy.

This paper explaines the position of Fujitsu file devices in the memory hierarchy and discusses their performance. This paper also covers some key technologies of magnetic disks and the efforts made to achieve higher reliability.

UDC 681.327.634:681.515
FUJITSU Sci. Tech. J., **26**, 4, pp. 271-279(1991)

F1700 File Controller Unit

• Hiroyuki Takizawa • Yasuo Kurihara • Joji Kikuchi

This paper describes the Fujitsu's latest product F1700B/S File Control Unit.

The conventional control units were designed to fully utilize the performance of disk devices. The advanced F1700B/S File Control, however, has been designed to enhance the file subsystem from the point of data management in addition to high performance.

This paper also describes an overview, functions/performance and hardware structure of the advanced file controller. It also discusses reliability, maintenance, firmware structure and future expanded functions.

UDC 681.327.634
FUJITSU Sci. Tech. J., **26**, 4, pp. 280-290(1991)

F6427H Magnetic Disk Subsystem: HAYABUSA

• Takashi Koike • Toshio Negoro • Teiji Yoshida

This paper outlines the history of Fujitsu's development of large disk subsystems, discusses recent market trends regarding large disk subsystems, and explains why the F6427H subsystem was developed. It then discusses the features of the F6427H subsystem, the system performance and reliability, and the technologies used in the F6427H magnetic disk drive to achieve high efficiency and performance.

UDC 681.327.634
FUJITSU Sci. Tech. J., **26**, 4, pp. 291-295(1991)

F6490 Magnetic Disk Subsystem: DIA

• Tomohisa Oyama • Yuji Ogawa • Kazeo Sugiyama

An array disk is a data storage subsystem that runs several of magnetic disk drives in a parallel configuration to achieve a very high transfer rate. High reliability is also achieved by utilizing a parity disk drive.

DIA or Disk in array is the first array disk subsystem for Fujitsu, achieving a transfer rate of 36 Mbyte/s.

This paper describes the configurations and performance of the subsystem in addition to a consideration on reliability.

UDC 681.327.67
FUJITSU Sci. Tech. J., **26**, 4, pp. 296-305(1991)

F6631 Solid State Disk: High-Speed Virtual Disk Unit

• Hajime Sugiura • Etsuo Morita • Soichiro Nagasawa

Fujitsu has developed the F6631 Solid State Disk (SSD) using 1-Mbit DRAMs as a storage element. This unit has a mass-storage capacity of 1 Gbyte and data transfer rate up to 72 Mbyte/s. In addition, the parallel processing disks with a redundant disk is used for the non-volatile data facility to ensure high reliability.

This paper describes the technology used in the F6631 SSD to provide the numerous functions, high performance, and high reliability.

UDC 681.327.634
FUJITSU Sci. Tech. J., **26**, 4, pp. 306-315(1991)

Fujitsu Small Magnetic Disk Drives

• Ikuo Kitamura • Norihito Aramaki • Takeo Masuda

The progress in the area of small magnetic disk drive is very rapid. This paper outlines the latest Fujitsu small disk drives; M2671P (8-inch, 2 600 Mbytes), M2652P/S (5-inch, 2 000 Mbytes), M262*X*S (3.5-inch, 520 Mbytes). It describes the general specification, interfaces and features of Fujitsu's disk drives, beginning with the history of the drives.

This paper covers magnetic heads, recording media, read/write and servo technologies to achieve higher capacity and speed, with the first priority on reliability.

Experimental data are presented on these technologies.

UDC 681.327.636
FUJITSU Sci. Tech. J., **26**, 4, pp. 316-320(1991)

F1751/F6470 Magnetic Tape Subsystem

• Yoshikazu Nakamura

This paper describes the functions and advantages of the F1751/F6470 Magnetic Tape Subsystem on which tape cartridge conforms to the industry standard.

It also outlines an 8-inch form factor M2481A/B Magnetic Tape Drive that uses the same cartridge.

UDC 538.975:681.327.634
FUJITSU Sci. Tech. J., **26**, 4, pp. 337-352(1991)

Thin Film Disk Technology

• Shoji Ishida • Kazuyuki Seki

High-density recording requires magnetic recording media with high coercivity, high magnetic saturation, and low noise. This paper reports metal-sputtered thin film disk media developed by Fujitsu to meet these requirements. Beginning with the features of cobalt-based alloys, substrates, sputtering process, and electromagnetic characteristics are described with analyses of the fine surface. Texturing, the protective layer, and lubrication are discussed in conjunction with durability and reliability with experimental results. These metal-sputtered media are installed on the latest Fujitsu large-capacity magnetic disk drives such as the 8-inch, 2.6-Gbyte F6427H subsystem.

UDC 681.327.636
FUJITSU Sci. Tech. J., **26**, 4, pp. 321-329(1991)

F6455 Magnetic Tape Library System

• Satoru Ohtsuka • Hajime Sugiura • Tetsuo Komura

Fujitsu has developed a library system using single-reel tape cartridges as the storage media. This library system can contain up to 5152 data cartridges, and provides storage capacity of up to 1 TB. The average cartridge transport time is 10 s.

This paper describes the latest Fujitsu library system concept, configuration and features.

UDC 538.975:681.327.634
FUJITSU Sci. Tech. J., **26**, 4, pp. 353-364(1991)

Thin Film Head Technology

• Mitsumasa Oshiki • Shigemitsu Hamasaki

This paper overviews the development and manufacturing of thin film heads in Fujitsu for high-performance magnetic disk drives. The design concepts, including slider and suspension design, are first discussed based on experiments and computer simulations.

A description of the wafer process follows, giving details for photoresist patterning and permalloy plating. Slider machining and assembly are also briefly described.

UDC 681.327.634†681.327.68
FUJITSU Sci. Tech. J., **26**, 4, pp. 330-336(1991)

F6443D Magneto-Optical Disk Drive Subsystem

• Ryosuke Kudou • Hiroshi Ichii • Akio Futamata

The large storage capacity, low cost per bit, and high reliability of optical disk drives make them an attractive alternative to magnetic disks and tapes for future external nonvolatile data storage applications.

A large storage capacity, high-speed access, F6443D magneto-optical disk drive subsystem has been developed. This is the first product of its type to be marketed. The new subsystem can store 36 Gbytes in a space of only 1 m^2 and can access any part of the media within five seconds.

This subsystem is expected to boost system efficiency in general-purpose computer applications.

UDC 621.318:681.327.634
FUJITSU Sci. Tech. J., **26**, 4, pp. 365-377(1991)

Structual Design for High-Performance Magnetic Disk Drives

• Keiji Aruga • Yoshifumi Mizoshita • Mitsuhisa Sekino

This paper introduces technology related to the mechanical assemblies of high-performance magnetic disk drives.

Guidelines are given for the optimum design of rotary head actuators for high-speed access. The minimization of off-track error and positioning erros caused by disturbances are also discussed, in addition to contamination control ensuring the reliability of a low-flying head. This paper also gives examples of how computer-aided engineering (CAE) is applied to mechanical design. Examples include magnetic field analysis, vibration analysis, and thermal deformation analysis using the finite element method.

UDC 681.327.634:681.58
FUJITSU Sci. Tech. J., **26**, 4, pp. 378-390(1991)

Digital Servo Control for Head-Positioning of Disk Drives

• Susumu Hasegawa • Kazuhiko Takaishi • Yoshifumi Mizoshita

A digital servo controller has been developed for head positioning in magnetic disk drives. In order to achieve precise head positioning and stable fast access operation, the actuator velocity must be estimated correctly. A state space control that uses the state estimator is introduced in this paper.

An access control method called SMART (Structural Vibration Minimized Acceleration Trajectory) has also been developed. By using this SMART control, the high harmonics of actuator drive are damped and thus both the residual vibration immediately after access and flying height fluctuation of the head are decreased. This paper describes the design and the experimental results of the state space controller and the SMART access control.

UDC 538.975:681.327.634
FUJITSU Sci. Tech. J., **26**, 4, pp. 391-403(1991)

Signal Processing for High Density Magnetic Recording

• Takashi Aikawa • Hiroshi Mutoh • Takao Sugawara

This paper introduces some of Fujitsu's developments in signal processing for high density magnetic recording. First, signal processing systems and a new method of signal simulation are explained. Then, developments made in coding methods and waveform equalization are discussed. Finally, noise reduction and future developments (e.g. the use of Viterbi detection) are discussed.

UDC 621.318:681.327.634
FUJITSU Sci. Tech. J., **26**, 4, pp. 404-414(1991)

Flying Head Assemblies

• Seizi Yoneoka • Takeshi Ohwe • Yoshifumi Mizoshita

The key to developing high-performance magnetic disk drives with a large capacity and a fast head access is to improve the slider s flight characteristics by modifying the head slider and head support.

Use of a negative pressure head slider can reduce wear at CSS, and use of a microslider can improve durability at the rated disk velocity while stabilizing flight.

A newly developed inline suspension spring greatly reduces head vibration during accessing, and enables fast access. It has a load beam that is stiffened in the access direction by use of a frame structure and has a new pivot that reduces the generated moment.

UDC 621.318:681.327.634
FUJITSU Sci. Tech. J., **26**, 4, pp. 415-427(1991)

Head-Disk Interface

• Takayuki Yamamoto • Minoru Takahashi
• Masayoshi Shinohara

As the flying height of magnetic heads decreases, the possibility of head-disk contact increases. New acoustic emission (AE) sensors for detecting head-disk contact have been developed. These devices are smaller and more sensitive than conventional ones. Using these new sensors, a method of evaluating the head-disk interface has been developed, and the influence of organic gas and dust on the head-disk interface was studied. It was found that component outgassing markedly increases the number of head-disk contacts, and can cause a failure in the head-disk interface. Wear on the slider and carbon overcoat on the disk was also studied, and the chemical change of the slider surface due to frictional heat was observed.

UDC 621.318:681.327.634
FUJITSU Sci. Tech. J., **26**, 4, pp. 428-438(1991)

Perpendicular Magnetic Recording

• Junzo Toda • Katsumi Kiuchi • Hiroaki Wakamatsu

Perpendicular magnetic recording was investigated using a high-efficiency probe-type head and a high signal-to-noise ratio double-layer hard disk. It was found that the read output and the linear bit density sharply increased as the head-to-medium spacing decreased. Particularly, the output value reaches a level more than twice that of longitudinal recording.

The sharp magnetic transition in perpendicular recording does not cause an increase in noise at high densities. Low-noise media superior to longitudinal recording media can be realized by controlling the magnetic domain structure of the underlayer.

Consequently, this paper discusses the feasibility for an ultrahigh areal density of over 1.55 bits per square micron (1-Gbit/inch^2).

Preface

Special Issue on Fujitsu File Devices

● Tokio Tatsuta
Executive Director

It is a great pleasure for me to see published a special FSTJ issue on Fujitsu File Devices, and to be able to deliver our new file product "HAYABUSA". These two events are of great significance to me as I have been involved in the file device industry for many years. Instead of discussing technical aspects and products here, I would like to outline Fujitsu's file device business.

Fujitsu file devices have the following significant features:
A product lineup wider than that of any other manufacturer.
Unique concepts and technology.

Fujitsu supplies many types of computers ranging from personal computers to supercomputers. Fujitsu's file device group also supplies many types of products, for example, magnetic disk units, magnetic tape units, magnetic tape libraries, optical disk units, and semiconductor file units. Many of these products are also supplied on an OEM basis.

Fujitsu's first file device was the unique crossbar relay memory. This device is still in use in the FACOM 128 Relay Computer System in the Ikeda Memorial Hall at our Numazu Factory; this product is the only one of its kind.

We next developed magnetic drums and tapes for file devices. In the late 1950s, these devices were used in FACOM systems that were mainly delivered to universities and laboratories. These devices were also used at Fujitsu in design automation systems.

Let me mention some of the products and events which contributed to the growth of Fujitsu's file device business.

F603 Magnetic Tape Unit (Early 1960s)

Before this unit was developed, magnetic tape was driven by pinch rollers, and the performance and reliability of magnetic tape units were poor. Then came a completely new concept—a magnetic tape drive with a single capstan connected directly to a motor which started and stopped itself. Since then, the single capstan drive system has become standard throughout the industry.

F631 Magnetic Disk Unit (Late 1960s)

This was Fujitsu's first magnetic disk drive. It was quite a large unit that consisted of a stack of fifty-one 660 mm-diameter disks. The unit was connected to an F230-50/60 system and helped form the first online banking system.

The F631 was driven by powerful electro-hydraulic stepping motors made by FANUC LTD. These motors were so powerful that they shock the floor when they operated.

Founding of NIPPON PERIPHERALS LIMITED (1973) (Development of the NP-20 Magnetic Disk Drive)

NIPPON PERIPHERALS LIMITED (NPL) was established as a joint venture between FUJITSU LIMITED and Hitachi, Ltd. to develop file device products. Our first joint product was the NP-20 Winchester-type Magnetic Disk Drive. This device was similar to the IBM 3340, and became a best-selling product because of its high performance and reliability.

The founding of NPL was a remarkable event in that two competitors were working together. The venture not only resulted in the NP-20, but also promoted good relations between the two companies.

Initial OEM sales of the NP-20 to MEMOREX and BASF, and sales of the well-known F613 Magnetic Tape Drive to MEMOREX succeeded in building a foundation for future overseas OEM sales.

Eagle-1 Magnetic Disk Unit (Early 1980s)

This world-famous device uses a 10.5-inch disk and was jointly developed with FUJITSU LABORATORIES LTD. The device employs a completely new concept and has changed the direction of magnetic disk drive development. The uniqueness of the Eagle-1 is not only in its 10.5-inch disk, but also in the technologies used in the new concept disk enclosure.

I heard that when IBM engineers evaluated the unit, they said, "This is what we wanted to make." The outstanding performance and reliability of the Eagle-1 and its companion 8-inch model, the Swallow 2, have prompted the slogan, "Fujitsu disks never break."

HAYABUSA Magnetic Disk Unit (1990)

Fujitsu is now delivering the Hayabusa (F6427). This product is the culmination of all our knowledge and effort. Unlike conventional large disk systems, it has excellent performance, high reliability, and superb cost-effectiveness. We are sure this product will satisfy our customers.

The products and events that I have mentioned above are all based on Fujitsu's original concepts and technologies. We are fully aware of how important it is to satisfy our customers' requests when we plan and develop new products. No enterprises can produce successful products in this industry unless their engineers consider the customers' needs and opinions.

What will the future of file devices be? Will magnetic recording technology and magnetic disks remain as the nucleus of file devices? These questions are shared by all of us in the file device industry.

In the hierachy of file devices, magnetic disks are positioned somewhere between semiconductor memory devices and optical disks. I think that, in the future, semiconductors will have a greater impact on the position of magnetic disks.

However, as long as the performance of magnetic disks continues to improve at its current pace, they will retain their current position as the main type of file device in computer systems. Magnetic disk technology will be combined, not replaced, with semiconductor technology to achieve higher performance for file devices. I think the ultimate file system is one that is transparent to the user.

I am sure that we can reach this target if we at Fujitsu continue to work together to develop the technology and improve file devices in terms of function, speed, and reliability.

UDC 681.327.634:681.515

File Technology Overview

• Minoru Fujino • Junshiro Sugihara • Masao Suzuki

(Manuscript received June 14, 1990)

File devices are used to store and retrieve the huge amounts of information handled in computer systems. A memory hierarchy consists of magnetic disks at the center, surrounded by various devices that focus on performance, long-term storage or economy. This paper explains the position of Fujitsu file devices in the memory hierarchy and discusses their performance. This paper also covers some key technologies of magnetic disks and the efforts made to achieve higher reliability.

1. Introduction

Data accumulated by both large-scale systems and small-scale systems with workstations generally triples every five years.

This increase in data volume necessitates more efficient processing and is further accelerated by an expanding system scale, networking, and increases in the number of users. Fujitsu continues to develop technology to meet the needs of its customers in these areas.

2. Overview of the file hierarchy

Figure 1 shows the file hierarchy towards which Fujitsu is working. The upper portion has a higher access speed but a limited capacity. The lower portion has a large capacity but a slower access. Fig. 2 shows the functional relationships within the hierarchy.

Because of their high capacity, low cost, and nonvolatility, magnetic disks are the most widely used type of external memory device.

Remarkable improvements have been made in magnetic disk technology, but in comparison with the speeds of CPUs and main memory, their access time remains slow. To help close this gap, semiconductor disk units and disk cache devices have been developed. These units use semiconductor memory to enable high-speed access of data on magnetic disk. The position of the disk cache in the file hierarchy seems to be established, and its use is becoming common even in small systems.

The data base assist function of a magnetic disk controller enables efficient execution of large programs, for example, those used with a relational database, by sharing functions in controller units.

The disk array unit transfers data at high speed by using several magnetic disks in parallel and by transferring data simultaneously. This method will become more widespread.

Magnetic tapes are used to archive infrequently used data, for example, data for backing up magnetic disks, data transportations, and weekly or monthly batch processing. Libraries could be used to help automate the backing up of large databases and so save labor.

Rewritable optical disks are a product of high-capacity memory technology, but their performance does not compare with magnetic disks. One of the advantages of optical disks is their portability. Also, in terms of cost per bit, optical disk library units are more favorable than magnetic disks. However, because fewer application programs are available, optical disks are limited to use as a secondary memory of magnetic disk systems, and to the storage of image and experimental data in large systems. In the personal computer field, where the introduction of new standards is relatively easy, optical disks will be

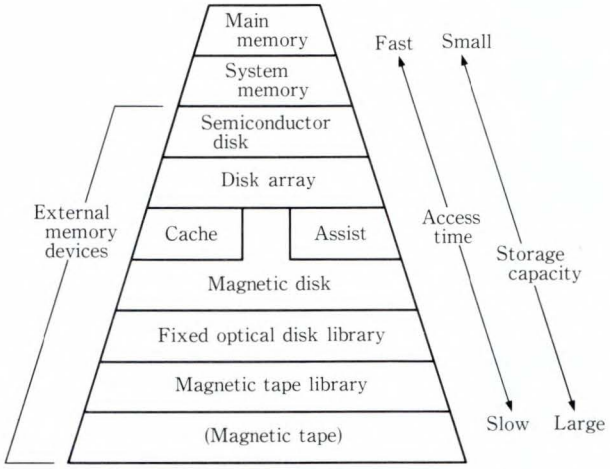


Fig. 1—Memory hierarchy in large computer systems.

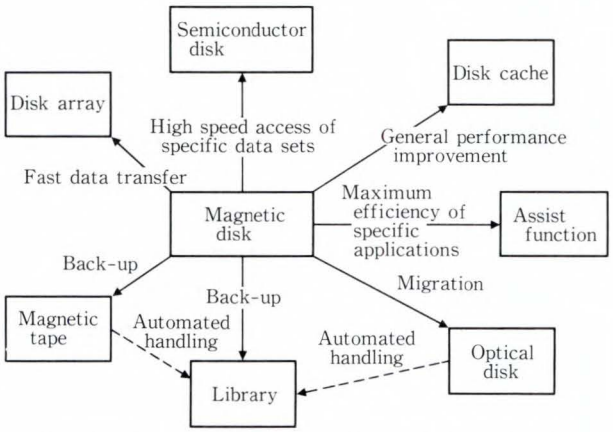


Fig. 2—Functional inter-relationships in memory hierarchies.

more widely used as a replacement for magnetic disks.

The price-to-performance relationship of some memory devices for large systems is shown in Fig. 3. Economical and efficient systems can be configured by using this price-performance data when combining memory devices.

Memory hierarchy is becoming more complex, making system configuration and file management more difficult. Automatic control by software should, therefore, be used to reduce the customer's workload. Instead of restricting control to hierarchy management, application programs must be developed that are not affected by new hardware and logical access interfaces.

Fujitsu makes full use of the know-how obtained from customer operations, and is developing software which can be expanded as needed.

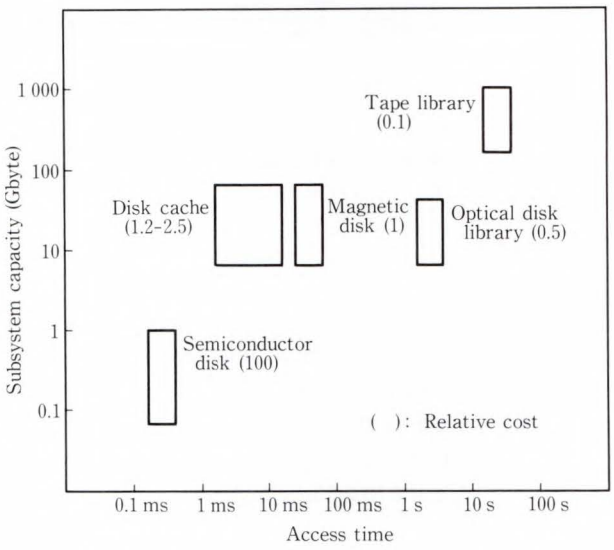


Fig. 3—Performance and cost comparison.

3. Improvements in subsystem performance

Magnetic disk units are nonvolatile and economical memory devices. However, because they have moving parts, their access time is much slower than that of main memory.

To allow for the relatively slow access time and to achieve optimum efficiency, jobs are executed in parallel so that the CPU need not wait for input-output operations.

A channel unit is used to control input-output and to transfer data independently from the CPU. When the channel and CPU capabilities are balanced, the system operates at maximum efficiency. Today's high-speed processing units require a large amount of data. This means that a large number of channels are needed – more than can be physically provided. This situation is referred to as the "channel bottleneck".

To solve this problem and enhance performance, the approaches described below have been taken.

3.1 Solutions to the channel bottleneck

To make better use of channels, dynamic channel path selection and asynchronous data transfer were introduced.

The data transfer rate can be greatly increased by using optical channels. Optical channels also enable longer cable lengths and eliminate the inconvenience of bulky cables. The maximum cable length has been extended to

1 km for magnetic disk devices, and to 2 km for magnetic tape devices. This limit used to be 120 m for both types of device. The maximum length can be further extended to 20 km by using a special adapter.

The off-loading of processing to a disk subsystem enhances the efficiency of data flow in the channels. Only processed data is transferred through the channels, and the data is transferred asynchronously with magnetic disk drive operation.

It is important to determine what types of processing should be off-loaded to the lower level subsystem, and to closely match application programs to the subsystem.

Operations suitable for off-loading include database retrieval, data dumping, and the copying of certain types of data. The final results are usually only a small part of the whole, which greatly increases the efficiency of channel use. This is especially true of retrieval from a relational database.

Data transfer from magnetic disk to magnetic tape is essentially performed without CPU assistance, and should be processed by a lower level subsystem. For this type of process, the channel, rather than the data flow path, becomes the reception counter for job requests and result reception.

In contrast to off-loading, there is the concept of centralized data processing. This involves making a large amount of memory accessible to the CPU and putting all data required for processing close to the CPU. The system memory is a semiconductor memory shared by several host CPUs.

Data input-output between these memory devices and magnetic disks is executed sequentially at high speed. A disk array is suitable for this type of application, a typical example of which is an array connected to a supercomputer.

As shown in Fig. 4, a disk array[1] provides high performance by configuring several magnetic disks in parallel and then tanasferring data simultaneously. By using a parity disk, data on a faulty disk can be recovered instantaneously. This enhances subsystem reliability. The disk array is extensively used in scientific and engineering computations, and is expected to be used in a broader range of applications in the future.

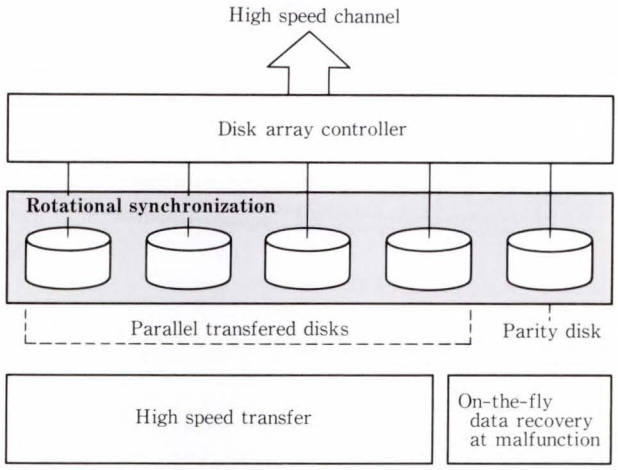


Fig. 4—Disk array[1]

3.2 Solutions to the access gap

The access gap can be resolved by using semiconductor memory. The packing density of semiconductor memory is becoming quite high and the cost is dropping significantly.

Processors are also becoming compact and inexpensive. These factors have promoted the general use of semiconductor disks and disk caches.

A semiconductor disk unit[2] emulates the magnetic disk drive in semiconductor memory. Unlike magnetic disk drives, the response time of semiconductor disk units is expected to enable constant access regardless of the load. Guaranteed access performance can be achieved for a small number of very frequently accessed data sets. The semiconductor disk unit could, therefore, solve the system bottleneck. Data in semiconductor memory is backed up by memory with a back-up battery and a small disk drive.

The disk cache function is based on the theory that disk access is not widespread in terms of time and location. During the first access, the neighboring data is also written into semiconductor memory. During the subsequent access, data is read from this memory if it has already been stored there. Data is written directly to the magnetic disk, bypassing the cache memory. This method is referred

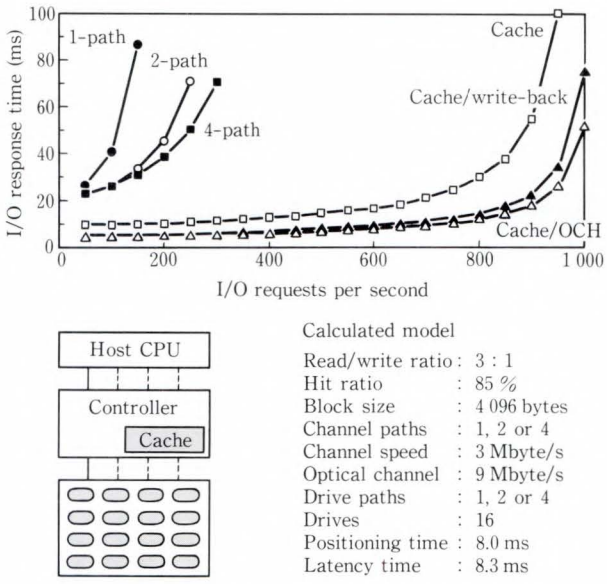


Fig. 5—Performance comparisons.

to as write-through writing. For further improvement, write-back writing using a cache memory is performed. In this method, data is written to a nonvolatile memory that is backed up by a small battery. This data is then written to the magnetic disk when the control unit is free.

Many small magnetic disk units now have a cache memory and are used as read ahead buffers.

Figure 5 shows the performance of systems using multiplexed channels and drives, and a system using a disk cache. As the figure shows, the effect of the disk cache is significant.

4. Downsizing of magnetic disks

4.1 Hayabusa magnetic disk subsystem

Figure 6 shows the development of Fujitsu's magnetic disk units in terms of capacity per spindle for the period 1983 to 1992.

The Hayabusa F6427H disk subsystem[3),4)], Fujitsu's top-of-the-line magnetic disk drive, is now being supplied for use with larger computers. This compact subsystem incorporates the latest technology for high-density magnetic disk drives. Compared to conventional systems, its floor space requirement is 40 percent less and its power consumption is 27 percent less.

The F6427H combines an 8-inch disk

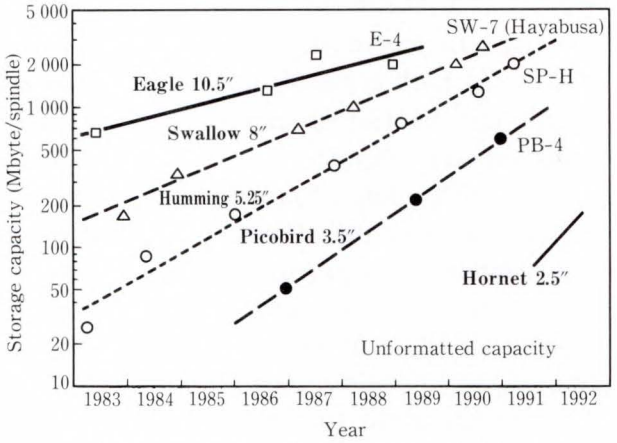


Fig. 6—Development of Fujitsu disk drives.

enclosure, a power supply, and electronic circuits in a single drive module. Each cabinet contains sixteen modules enabling a total capacity of 30 Gbytes. A faulty module can be replaced without stopping the other modules. All parts except the disk modules are duplicated and have multiple paths. This approach greatly enhances reliability and availability.

4.2 Small disk drives[5)]

As shown in Fig. 6, 8-inch, 5-inch, and 3.5-inch disks are already in use, and the use of 2.5-inch and 1.8-inch disks is not far off. As personal computers gradually began to incorporate small disk units, a large market of new products was created. High production by dedicated manufacturers led to keen competition in pricing and performance.

Because of the competition and the need to overcome size limitations, new technologies have been energetically applied to the development of smaller disks. Five-inch disk drives with capacities of more than 1-Gbyte have been developed. The use of a smaller disk makes it easier to achieve high-speed rotation and positioning. This further contributes to the downsizing and modularity of the disk drive. In laptop computers, disk drives must be even smaller, consume less power, and be more robust.

Another trend is towards embedded intelligent controllers. An intelligent controller performs complex operations with simple commands to reduce the load on the host.

Some examples of these operations are command queuing, automatic recovery from errors, and read-ahead buffer capabilities. In the latest 3.5-inch and 2.5-inch disk units, these features are built-in and are equivalent to features in large system controllers having a disk cache.

The most widely used interface standard for intelligent controllers is the Small Computer System Interface (SCSI). The use of a fast SCSI and a wide SCSI is being proposed to remove limitations on transfer speed and the number of units that can be linked. The IPI-2 and IPI-3 interfaces are used in middle and high-end systems.

5. Magnetic tape

The use of magnetic tape for data compatibility and data transfer is no longer significant. There is a greater need for data migration, long-term data retention (archiving), evacuation of data against disaster, and data back-up to magnetic disk.

Because of its greater capacity, higher transfer speed, and better ergonomics, the IBM 3480-type cartridge magnetic tape unit is replacing conventional open-reel magnetic tape units.

Data compression has also become essential. This is a code conversion that compresses data according to the frequency of use of each character. Characters used more frequently are converted to shorter codes. The compression ratio depends on the data, but a typical ratio of more than 50 percent can be achieved. This greatly increases the storage capacity of magnetic tape reels.

Fujitsu supplies high-performance magnetic tape subsystems[6] for large computers, and supplies compact low-priced magnetic tape units for medium to low-end systems. The auto loader function is becoming indispensable for improved ergonomics.

The greatest factor in the choice of magnetic tape used for personal computers and workstations is the price. Also, Digital Audio Tape (DAT) and 8-mm video devices are being developed.

6. Optical disks

Optical disk systems have random access capability, large capacity, and media interchangeability. These systems are positioned between magnetic disks and magnetic tape.

The commercial application of rewritable optical disks has begun, but their use is still limited to a small number of fields. Optical disks are mainly used by dedicated systems where the low cost per bit of optical disks are an important consideration, for example, electronic filing systems and secondary volumes for magnetic disks.

Library devices play an important role in operations in large computer systems.

7. Magnetic tape and optical disk libraries

Magnetic tape still plays an important role in the back-up of magnetic disks. Some computer centers have tens of thousands of tapes, and the handling of these tapes incurs considerable expense. Fujitsu's library units can be used to automate backup while eliminating errors and reducing costs.

The requirements in this field are large capacity, low cost per byte, small physical size, and high-speed data transfer. Also, the media must be able to tolerate long storage periods, and the read/write unit must be accessible for

Table 1. Cartridge type magnetic tape drives

Model	Features	Load/unload	Data transfer rate	Auto-loader/library
F6470	IBM 3480 compatible Stand-alone type	7 s/7 s	3.0 Mbyte/s	6 or 24 cartridge auto-loader F6455 magnetic tape library (5 200 cartridges max)
M2481	IBM 3480 cartridge compatible 8-inch form factor	7 s/7 s	1.5 Mbyte/s	10 cartridge auto-loader Compact library planned (280 cartridges max)

Table 2. Optical disk drives and libraries

Model	Features	Capacity	Access time	Data transfer rate	Remarks
F6441	12-inch write-once Single medium	2.8 Gbytes	250 ms	Read: 0.8 Mbyte/s Write: 0.4 Mbyte/s	–
F6442	12-inch write-once Library	88 Gbytes	250 ms	Read: 0.8 Mbyte/s Write: 0.4 Mbyte/s	32 cartridges max Unload and load time: 13 s avg.
M2502B	12-inch write-once Single medium	3.6 Gbytes	200 ms	Read: 0.8 Mbyte/s Write: 0.4 Mbyte/s	–
M2505B	5-inch write-once 5-inch form factor	600 Mbytes	100 ms	Read: 0.7 Mbyte/s Write: 0.35 Mbyte/s	–
F6443	8-inch rewritable Fixed media library	35.6 Gbytes	200 ms	Read: 0.7 Mbyte/s Write: 0.2 Mbyte/s	8 fixed media Head traverse time: 5 s avg.

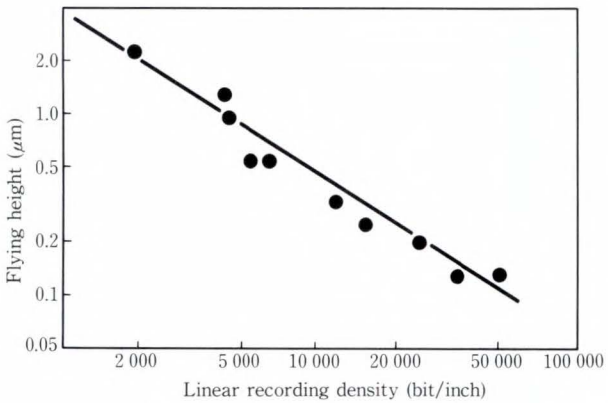


Fig. 7—Trend in head flying height.

the long periods required to transfer data.

Fujitsu supplies magnetic tape libraries, optical disk libraries, and stand-alone units that meet the above requirements[7),8)]. These products are listed in Tables 1 and 2.

8. Magneitc disk reliability

The reliability and performance of file systems, especially magnetic disk units, have markedly improved. At their initial stage of development, file systems had a Mean Time Between Failure (MTBF) rating of several thousand hours. Now, Fujitsu's latest models have an MTBF rating of 200 thousand hours.

8.1 Head-disk interface

The greatest problem with magnetic disk units is the possibility of a head crash. In a head crash, the head and media touch, the media is damaged, and stored data is lost. Continuous efforts to improve reliability have reduced such occurrences.

The Head-Disk Interface (HDI)[9)] is being studied to identify faults, prevent crashes, and understand the physical of disk drives. To reduce the size of drives and increase the recording density, the head flying height must be lowered (see Fig. 7). This increases the importance of HDI.

Fujitsu's activities related to HDI and some of the tools developed to improve Fujitsu products are explained below.

1) There is a probability that the head will touch the media instead of floating above it. To detect such a contact, an ultra-compact, highly sensitive piezo-electric element is mounted on the back of the head. At the initial stage of development, this element was mounted on every head in order to monitor the design, manufacture, and assembly of the head and media.
2) A drive is placed in a special chamber with reduced air pressure in order to lower the head flying height. This enables the accelerated test on head crashes without changing other operational conditions.
3) Sliders and head supports are designed by computer simulation to be tolerant of high-speed seeking and flying fluctuations[10)]. A tool which dynamically measures the head flying height during seeking has been developed to check designs.
4) A method of capturing the small amounts of gas produced by the various materials used in disk enclosures (DEs) has been developed. Also, a material which does not produce harmful gasses has been selected for use in DEs.

8.2 Product quality assurance

When product development planning begins, Fujitsu's inspection department establishes an evaluation for the new product. For a new technology, two or three evaluation plans are draw up.

The design engineer checks that the designed product meets the original requirements. After this check (called the Design Verification Test), the product is passed to the manufacturing section where the testing department subjects the product to a series of running and environmental tests based on the approved evaluation plans.

Before the product is shipped, several hundred units are given continuous running tests for several months to confirm the MTBF. After testing, the units are disassembled and thoroughly inspected. If a failure occurs or the symptoms of a failure are discovered, production is halted until the problem is corrected. Nothing is shipped unless it passes all of the tests.

After regular shipment begins, sample units are picked from daily production and subjected to running tests. These units are then disassembled and inspected. This procedure ensures consistent quality in the manufacturing process.

To maintain high quality during assembly, clean rooms are monitored and controlled by dust counters, and all personnel are given extensive training regarding cleanliness.

9. Future outlook

This section gives an overview of anticipated developments in magnetic disk technology.

9.1 High density

As shown in Fig. 8, the recording density of magnetic disks has increased 100 fold over the past 20 years, and currently stands at 100 Mbytes per square inch. The increase in density for small disks has been more rapid than that for larger disks. A density of 1-Gbyte square inch, already announced by IBM, will be reached if this increase continues at the current rate.

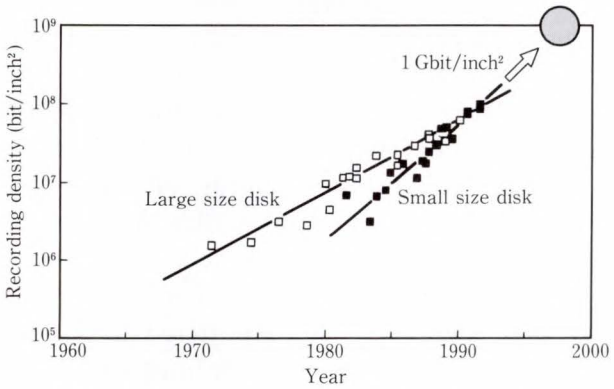


Fig. 8—Improvements in recording densities.

9.2 PPI

In recent development of small-size disks, there has been tremendous competition to increase capacity without increasing the disk size (form factor). One solution is to stack as many disks as possible in the available space (platters per inch or ppi). This requires reductions in the disk thickness and head size. Reductions in the size of the head slider are also welcome from the HDI standpoint. (If the head slider is smaller, the impact energy of head crashes is less.)

9.3 Disks

Metal-sputtered disks[11] have almost replaced particulated disks, which were the norm for many years. It has been shown that recording density can be improved by increasing the coercivity to 1 800 Oe (22.6 A/m) or more. The main tasks now are in developing a head that can write on a high-coercivity disk, and in improving the signal-to-noise ratio by reducing disk-noise.

From the HDI standpoint, a major breakthrough will be the development of an extremely thin and durable protective layer, and the development of a thin, rigid, and smooth substrate material having an appropriate texture.

9.4 Heads

Most heads in use today are of the ferrite monolithic type. However, the transfer speed of these heads is limited because of their large inductance. When combined with a metal-

sputtered disk, this type of head has insufficient magnetization reversal capability (i.e. its writing capability is insufficient). For small disks, which have a relatively slow transfer speed, the Metal In Gap (MIG) head has become popular. For transfer speeds of more than 3 Mbytes per second, thin-film heads[12] are beginning to be used.

To further increase the recording densities of high-coercivity metal-sputtered disks (see Fig. 9), the development of a head with a high writing capability is essential. Development of materials having a high saturation flux density, such as amorphous CoZrCr or plated CoFe, is under study.

Magneto-Resistive (MR) heads have been under investigation for a long time, and densities of 1-Gbit per square inch using these heads have recently been reported.

These developments will not only benefit small disks with slow rotational speeds, but will also enable an optimal design by splitting read and write operations. There are many problems to be resolved regarding high-density MR heads, for example, the control of Barkhausen noise.

9.5 Perpendicular recording

High density perpendicular recording involves a new technique[13]. Progress has been made in longitudinal recording, and its performance compares very favorably with that of perpendicular recording.

Figure 10 shows the results of experiments and computer simulations on these two recording methods. The figure shows that perpendicular recording is superior when the head flying height is less than 0.1 μm. Technology is approaching the point where such an ultralow flying height is practicable. However, many obstacles must be overcome before implementation is possible, and research is continuing.

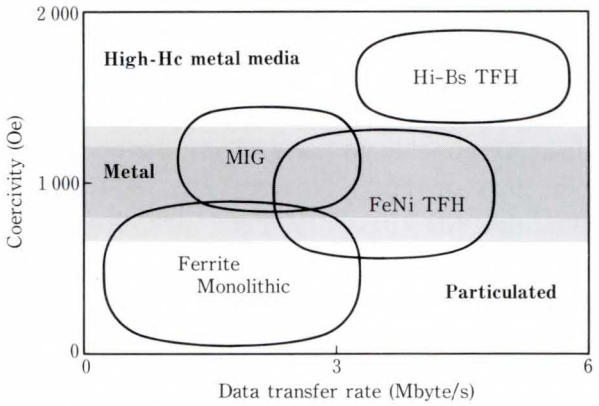


Fig. 9—Magnetic heads and recording media.

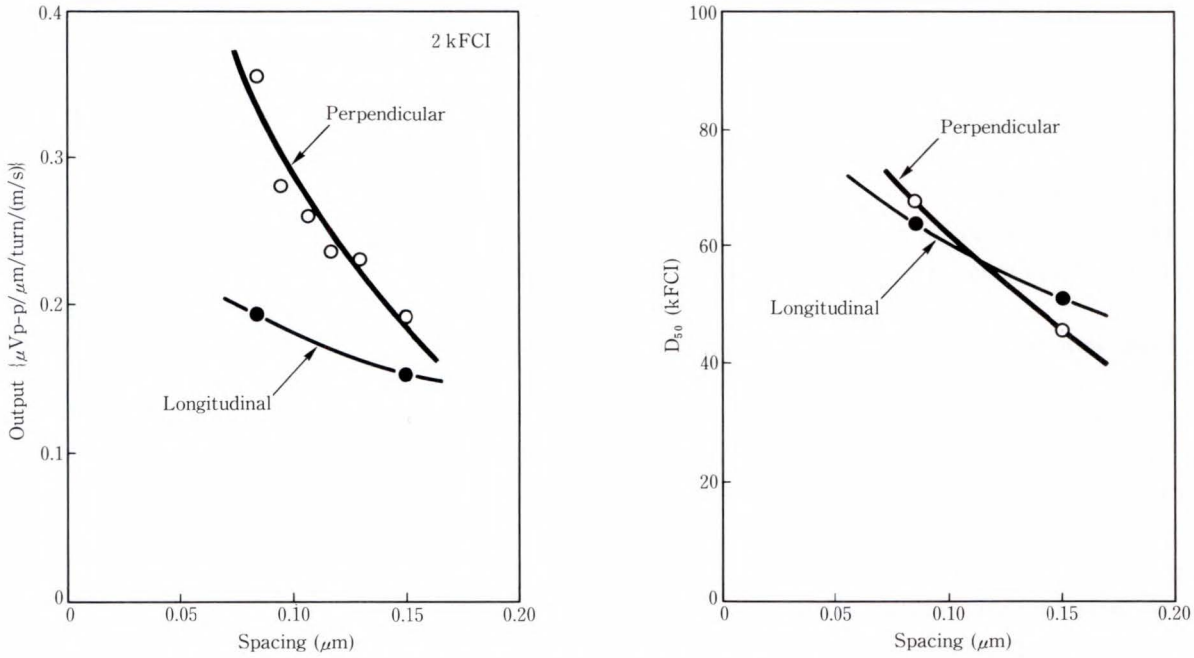


Fig. 10—Longitudinal and perpendicular recording.

9.6 Signal processing

As the recording density rises, signal quality deteriorates because of interference from adjacent data and because of decreases in the signal-to-noise ratio.

Signal deterioration can be reduced by processing the signal before writing and after reading.

1) Equalization and precompensation[14]

Equalization of the read signal makes read signal processing easier and more precise. In precompensation, the predictable distortion of the read signal is compensated by making appropriate changes to the write signal.

2) Run length limited (RLL) coding

By adding redundancy to the code and instead excluding the disadvantageous code combination of interference helps higher density recording.

3) Error correction code (ECC)

By adding redundancies based on a mathematical algorithm, errors can be detected by the invalid codes they produce, and the original data can then be recalculated.

These are typical processing operations. More complex operations aided by developments made in LSI and MPU technolgoy will be incorporated.

4) Most likelihood decoding

Most likelihood decoding involves determining the most likely code of a read signal from the preceding and following waveforms. Also, ways to reproduce a low signal-to-noise ratio signal are currently being developed.

9.7 Drive design and servo control

Magnetic disk units are becoming more common in personal and laptop computers, and are being used in move diversified physical environments. Computers are now exposed to greater temperature changes, vibration, and shock.

Power consumption and acoustic noise are also important concerns. Tracking precision must be enhanced to the maximum to enable higher densities. A higher access speed is also required.

Efforts are underway to satisfy these conflicting requirements[15]. Most noteworthy is the application of the digital signal processor (DSP)[16].

The DSP can be used to do the following:

1) Measure environmentally induced changes in characteristics, and then adjust the control parameters.
2) Measure or estimate vibration, and then make compensations to reduce the vibration.
3) Reduce vibration and noise by smoothing acceleration and deceleration, and reduce power consumption and temperature increases.

10. Conclusion

Magnetic disk units will have higher recording densities, become more compact, and be used in more diverse applications.

Semiconductor memory will continue its sharp growth. The difference in cost per bit between semiconductor memory and magnetic disk drives is more than 100 times what it used to be, but it is now on the order of several tens of times.

The optical disk unit is being established as a large capacity memory unit having media exchangeability, transportability, and mass-production capability. Optical disks will become widespread as prices drop due to technological breakthroughs.

Optical disks, conventional magnetic tape units, DAT, and 8-mm video tapes can be used as backup media. Each of these types of media has their advantages. The competition to determine which products survive in the marketplace is expected to continue for some time.

Will the memory heirarchy become more complex or will it be simplified through the use of best-fit technology?

In either case, Fujitsu is committed to providing the best cost performance and the best solutions as an across-the-board manufacturer.

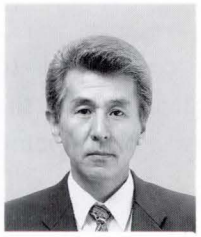
References

1) Oyama, T., Ogawa, Y., and Sugiyama, K.: F6490 Magnetic Disk Subsystem: DIA. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File

Devices), pp. 291-295 (1991).
2) Sugiura, H., Morita, E., and Nagasawa, S.: F6631 Solid State Disk: High-Speed Virtual Disk Unit. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 296-305 (1991).
3) Takizawa, H., Kurihara, Y., and Kikuchi, J.: F1700 File Controller Unit. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 271-279 (1991).
4) Koike, T., Negoro, T., and Yoshida, T.: F6427H Magnetic Disk Subsystem: HAYABUSA. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 280-290 (1991).
5) Kitamura, I., Aramaki, N., and Masuda, T.: Fujitsu Small Magnetic Disk Drives. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 306-315 (1991).
6) Nakamura, Y.: F1751/F6470 Magnetic Tape Subsystem. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 316-320 (1991).
7) Ohtsuka, S., Sugiura, H., and Komura, T.: F6455 Magnetic Tape Library System. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 321-329 (1991).
8) Kudou, R., Ichii, H., and Futamata, A.: F6443D Magneto-Optical Disk Drive Subsystem. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 330-336 (1991).
9) Yamamoto, T., Takahashi, M., and Shinohara, M.: Head-Disk Interface. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 415-427 (1991).
10) Yoneoka, S., Ohwe, T., and Mizoshita, Y.: Flying Head Assemblies. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 404-414 (1991).
11) Ishida, S., and Seki, K.: Thin Film Disk Technology. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 337-352 (1991).
12) Oshiki, M., and Hamasaki, S.: Thin Film Head Technology. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 353-364 (1991).
13) Toda, J., Kiuchi, K., and Wakamatsu, H.: Perpendicular Magnetic Recording. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 428-438 (1991).
14) Aikawa, T., Mutoh, H., and Sugawara, T.: Signal Processing for High Density Magnetic Recording. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 391-403 (1991).
15) Aruga, K., Mizoshita, Y., and Sekino, M.: Structural Design for High-Performance Magnetic Disk Drives. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 365-377 (1991).
16) Hasegawa, S., Takaishi, K., and Mizoshita, Y.: Digital Servo Control for Head-Positioning of Disk Drives. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 378-390 (1991).

Minoru Fujino
Information Processing Devices Group
FUJITSU LIMITED
Bachelor of Mechanical Eng.
Osaka University 1963
Specializing in File Devices Design

Masao Suzuki
Information Processing Devices Group
File Devices Div.
FUJITSU LIMITED
Bachelor of Control Eng.
Tokyo Institute of Technology 1966
Specializing in File Devices Design



Junshiro Sugihara
Information Processing Devices Group
File System Div.
FUJITSU LIMITED
Bachelor of Mechanical Eng.
Waseda University 1965
Specializing in File Devices Design

UDC 681.327.634:681.515

F1700 File Controller Unit

• Hiroyuki Takizawa • Yasuo Kurihara • Joji Kikuchi

(Manuscript received June 29, 1990)

This paper describes the Fujitsu's latest product F1700B/S File Control Unit.
The conventional control units were designed to fully utilize the performance of disk devices. The advanced F1700B/S File Control, however, has been designed to enhance the file subsystem from the point of data management in addition to high performance.
This paper also describes an overview, functions/performance and hardware structure of the advanced file controller. It also discusses reliability, maintenance, firmware structure and future expanded functions.

1. Introduction

Conventional file controllers have been positioned on the data path from the channel to the disks. They interpret commands from channels, translate them into disk drive-oriented commands and perform various other tasks. In addition to these functions, file controllers are now equipped with disk cache functions, which control disk drives and can manage data on the disks.

The F1700 File Control Unit (FCU) was developed and distributed as a controller for the F6425G/H Disk Drive Unit[1]. This FCU will satisfy the market's needs for higher file functions and performance. The controller functions have been improved so that now, rather than being simple file controllers, their functions can be expanded as file processors.

Conventional controller operations were based on a single microprogram (control programs). Communication with host channels, subordinate device controls, and cache memory controls were all performed by serial operations. Therefore, the same number of control directors (controller processors) as the maximum number of simultaneous read and write operations was required. However, there has been a rise in users' needs for expanded functions and parallel operations. Expansion of file processing functions has been become absolutely necessary. Therefore, Fujitsu has investigated the overall functions of controllers, and has decided on an entirely new basic structure for controllers as file subsystems that can greatly expand and improve the functions of conventional file controllers.

The basic controller functions together with the improvement of maintenance functions has been divided into five parts. These functions have been built into modules, each with a microprocessor structure. One or two modules are implemented on a single PC-board.

For the basic technology, the F1700 FCU introduces CMOS 20-Kgate LSI. Each module has a common bus structure that enables communication among modules. The common bus structure is duplexed from the viewpoint of fault-tolerance.

The functions and structure of the modules are explained in detail in section 3.3, but these can be summed up as follows:

1) Channel adapter: CA
   Module for communication between channels
2) Device adapter: DA
   Module for communication between devices
3) Resource manager: RM
   Module for managing resources
4) Service adapter: SA

Module for maintenance

5) Cache function engine: CFE
Module for supporting extended functions.

Each of these modules is managed by independent control programs. They work together as the subsystem while communicating among modules and sending inquiries to the RM for device use status.

As explained in chapter 3, it is not necessary to alter the overall structure to conform to environmental changes and function expansion. Rather, a broad range of applications can be achieved simply by partial module modifications and addition of functions.

2. Functions/performance

2.1 Disk drives

The following magnetic disk drives can be connected to the F1700B/S.
F6425C4/B4/K4/L4/M4/N4/GA4/GB4/HA4/HB4, and F6427HA/HB.

For the disk drives of the F6425GA4/GB4/HA4/HB4 and the F6427HA/HB, four-access paths (these are called four-path cross calls) are available for each drive. In other cases, two-access paths (cross calls) are available.

Disk drives that can be connected to the F1700B are, at maximum, 128 spindles. It is possible to construct file systems that have large storage capacities of up to 240 Gbytes.

Disk drives consisting of a maximum of 64 spindles can be connected to the F1700S.

There are no restrictions on mixing different drives for the F1700B/S subsystem configuration as long as they are not mixed in the same string.

Mixing different drives in the same string can be defined as follows:

1) F6425B4/L4/N4
They can be mixed within a string of F6425C4, K4, or M4.
2) F6425GA4/HA4 with F6425GB4/HB4
They can be mixed if the same number of paths are provided (2 or 4).
3) F6427HA/HB cannot be mixed with other drives.

2.2 Channels for F1700B/S

The maximum number of channels is 32 for the F1700B and eight for the F1700S. The types of channels which can be connected are as follows:

1) 3 Mbyte/s and 4.5 Mbyte/s Block Multiplexer Channel devices (hereafter called the BMC).
2) 6 Mbyte/s and 9 Mbyte/s Optical Channel devices (hereafter called the OCH).
3) 9 Mbyte/s OCH is supplied by Fujitsu M-780 systems and 6 Mbyte/s OCH is supplied by Fujitsu M-760 systems.

The F6427HA/HB requires a channel that has a data transfer speed of 4.5 Mbyte/s or more. When the F1700B/S is equipped with a disk cache feature, data transfer between the mainframe memory and disk cache is executed at the data transfer speed of the connected channel. The data transfer speed should be faster than that of the disk drives, regardless of the type of disk drive.

The aggregate data transfer speed of the F1700B is 72 Mbyte/s at maximum. That is, 16 BMCs at a transfer rate of 4.5 Mbyte/s are capable of simultaneous data transfer. At maximum, eight OCHs at a transfer rate of 9 Mbyte/s are capable of simultaneous data transfer.

The total data transfer speed of the F1700S is 36 Mbyte/s. The F1700S can transfer data simultaneously with a maximum of four channels. In addition, the BMC and OCH can be mixed within the F1700B/S.

2.3 Disk cache feature

The disk cache is an optional feature of the F1700B/S. The disk cache is designed to improve the average response time of I/O operations for large amounts of data stored in the disk drives. This option makes use of the characteristic that there will be a high probability of access to the same track within a specific time period. Once accessed, the data on the same track is read to memory in the controller (disk cache memory). This is called staging. Thereafter, access requests for this track are processed using the disk cache memory. This is called hit operations. This shortens the positioning time of

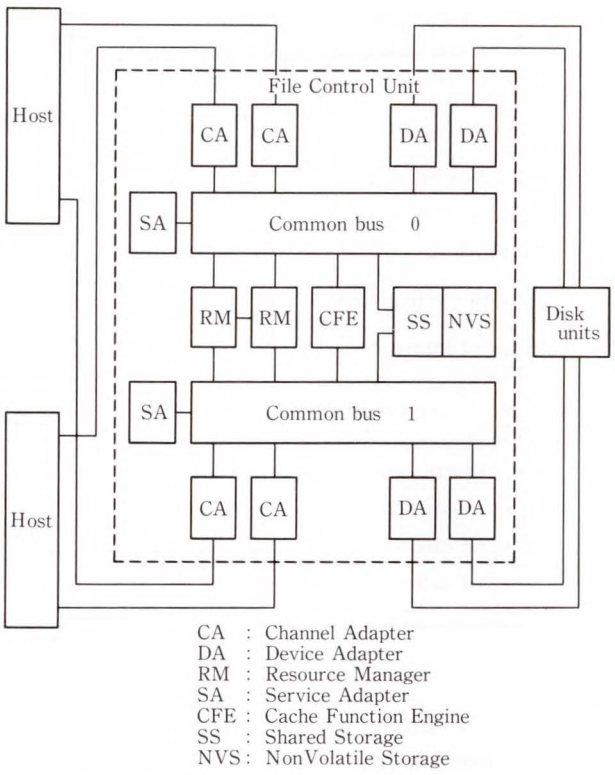


Fig. 1—Internal logical configuration of FCU.

disk drives and greatly reduces the response time[2)-5)].

Furthermore, the F1700B/S is equipped with a nonvolatile storage (NVS) to enable high-speed writing. Once data to be written is placed in the NVS, the I/O operations from the host are completed. Data in the NVS is written back to a specified drive in between I/O operations from the host. This is called write-back processing (see Fig. 1).

The capacity of disk cache memory is 32 Mbytes to 512 Mbytes for the F1700B, and 32 Mbytes to 256 Mbytes for the F1700S. For both models, the nonvolatile storage capacity is 4 Mbytes. When the disk cache feature is used, Disk Cache Support Programs (DCSP) are required corresponding to the Fujitsu M-series operating systems OS IV/F4 MSP, OS IV/X8 FSP, and UTS/M.

These programs are necessary for specifying the target disk drives, data set and write-through/write-back operation mode and for managing statistical data.

2.4 I/O queuing functions

To speed up I/O requests from the host, information such as the location of a specific I/O data recorded on the disk drive is memorized as a task control block (TCB) in the FCU. Up to seven TCBs can be created and queued for each device. If any of the resources such as channel and, device paths and devices that are needed for processing with the F1700B/S are temporarily unavailable, a TCB can be created. This TCB can be dequeued and executed as the resources become available. This function is called queuing. This function makes it possible to increase the efficiency of the channel and device paths.

2.5 Centralized resource management

The F1700B/S has a mechanism for centralized management of the resources necessary for processing. Information such as the device status, channel path status, TCB, reserve information, statistics information, and error status is stored in the shared memory area.

This mechanism together with the previously described I/O queuing function reduces overhead and improves efficiency. The RM is the module for this role in the F1700B/S.

Furthermore, because the occurrence of a failure on the RM can cause a fatal error, duplex RMs are implemented as hot standby.

2.6 Maintenance information management

The F1700B/S contains and maintains on a built-in hard disk the configuration information of the model, addresses, number of drives, string configuration and cache memory capacity of the drives and type of channel connected to it. This information is read from the built-in hard disk when the F1700B/S is powered on and used as control information. This configuration information can be referenced and updated by hardware maintenance programs and a maintenance tool called the Fujitsu Maintenance Device (FMD). For example, with the F6425GA4/GB4/HA4/HB4 and the F6427HA/HB, drives can be added without stopping user's operation. This is called "hot plug-in". This "hot plug-in" makes it easy to alter the system configuration, thus

24-hour operating systems.

3. Configuration

The FCU introduces a bus structure that can connect additional modules to facilitate function expansion and performance upgrade. Figure 1 shows the logical configuration of the FCU. To achieve high performance and reliability, two common buses are provided and configured so that modules can communicate using common buses.

3.1 Common bus

The common bus is a communication bus among the various modules. The common bus is a dualized synchronized bidirectional bus. Each common bus has a data transfer rate of 36 Mbyte/s for one-way and is capable of connecting up to 128 modules. Common buses have special functions. For example, two-destination addressing makes it possible to communicate simultaneously with two different modules. There is also a broadcast function which enables communication with 16 modules with consecutive addresses without specifying each address.

3.2 Microprocessor

The FCU has been designed so that communication between modules via the common bus is done by standard procedures so that some overhead is generated from time to time. Each module is equipped with high-speed microprocessors to reduce this overhead. These high-speed microprocessors were developed exclusively for the FCU using CMOS 20-Kgate LSI. As the specifications in Table 1 indicate, the development of powerful microprocessors makes it possible to maintain high performance in a modularized structure. The microprocessors are commonly used in each module in the FCU.

Table 1. Microprocessor specifications

Item	Specification	Remarks
Instruction bit width	32 bits	ECC bits added
Internal bus bit width	16 bits	–
ALU bit width	16 bits	–
Cycle time	110 ns	–
Execution cycle count	1, 2, or 3	Mostly 1 cycle
Number of internal registers	256	16 bit-wide
Number of external registers	Up to 128	16 bit-wide
Control storage addressing	Up to 1 Mwords	–
Table storage addressing	Up to 1 Mwords	–
Number of interrupt levels	8	–
Number of subroutine stacks	64	–

3.3 Modules

3.3.1 Channel adapter (CA)

This module is used for communications with the host system channels. This module receives requests from the host system, distributes them to the appropriate modules in the FCU, and executes data transfer with the host system. A single channel adapter module has two channel ports, but only one of these operates at a time. Initially, the FCU only supported CA for controlling the block multiplexer channel (BMC) interface, but when the Fujitsu M-series mainframe supported optical channels (OCHs), the FCU project team developed an optical channel interface adapter (OCA). The FCU currently supports two types of channel adapter modules such as the BMC and OCH.

3.3.2 Device adapter (DA)

This module controls the disk drives through the bidirectional controller interface (BCI). This module transfers data between devices. One device adapter module can control up to 64 logical drives.

3.3.3 Resource manager (RM)

The RM constantly monitors the resource operating status. The RM assesses the resource structure as to how many CA and DA are connected, whether a cache is installed, and how much capacity (in units of megabytes) the cache memory has. This module assigns the appropriate modules needed for each processing requested from the host. If the module assigned is busy with another task, the request is queued. When resources become available, the task which has been queued is dequeued and the signal to process the task is sent to the module. The RM

releases the module when it completes its task, then prepares for the next task. The RM manages the physical and logical configuration of the subsystem as well as restricting data transfer that exceeds the common-bus capacities. In this way, it manages the resources within the FCU and performs efficient task operation.

3.3.4 Service adapter (SA)

This module initializes the various modules, performs initial microprogram loading, monitors operating status, detects errors, and disconnects modules that have malfunctions.

In addition to maintenance and diagnosis of the various modules, this module also provides a human-machine interface for maintenance and diagnosis. Furthermore, the service adapter records physical configuration information and error data on a 5 1/4-inch built-in hard disk as maintenance information.

3.3.5 Cache function engine (CFE)

The CFE is part of the cache function, which can be provided as an option. The CFE consists of a dedicated microprocessor and dedicated memory for recording tables (Hush table and least-recently-used table) needed for cache control.

Based on the signals from other modules, the CFE judges hit-and-miss, allocates cache memory, and deallocates cache memory for data infrequently used.

3.3.6 Shared storage (SS)

The SS is a high speed memory with a large capacity that can be accessed from any module. The SS is currently used as data memory for the cache. The SS is divided into 16-Kbyte blocks. This makes it easy to manage allocation of memory blocks for separate tasks and disconnection of bad blocks.

The SS consists of CMOS 1 Mbit DRAM. The SS can have a capacity of 32-512 Mbytes, which can be increased in units of 32 Mbytes.

3.3.7 Nonvolatile storage (NVS)

The NVS is used for high-speed write operations in the disk cache (write-back cache function). The NVS uses a CMOS 256-Kbit SRAM and has a capacity of 4 Mbytes. The data to be updated is first stored in the NVS, then written back to the specified disk drive late. The NVS is backed up by batteries to retain the data until it is written back to the disk drive even if power fails. Figure 2 shows the outline of the write-back cache operation.

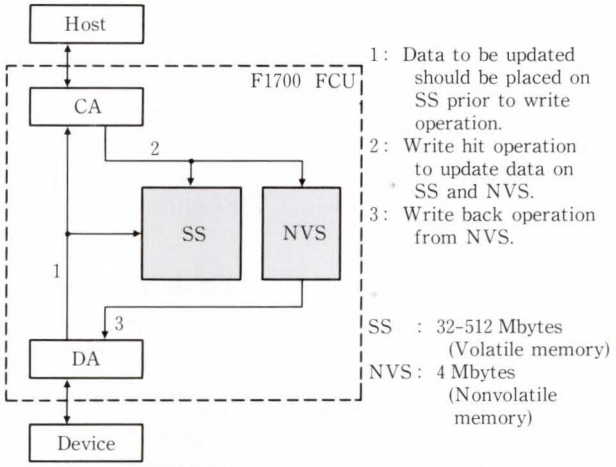


Fig. 2—Write-back operation using NVS.

4. Reliability

The FCU offers a high degree of reliability demanded for disk subsystems. The FCU also offers greatly upgraded overall function and performance. The FCU has a dualized common-bus system and alternate module allocation so that a single malfunction of the hardware will not cause a fatal system failure.

Since the RM requires centralized management to manage resources, primary and standby RMs are provided. If the primary RM fails, the operation is immediately switched to the standby RM. The same function is provided for memorized resource configuration tables. Each common bus is equipped with a clock generator to prevent FCU failure caused by a clock system failure.

The FCU has a dual power supply system. Two different power supplies are provided independently for each common bus, thus preventing common bus failures due to a power failure. The DC power supply for the modules such as the RM and CFE connected to both common buses is supplied by diode ORed of the power supply. Therefore, a single power failure will never cause all modules to fail.

In this way, extensive considerations have been taken to insure that one single malfunction of the FCU logical circuits or power supply will

not cause total FCU failure. This means that an extremely high degree of reliability has been achieved.

5. Maintenance

The FCU has a 4-line, 80-column LCD on its maintenance panel.

This LCD displays the module configuration and status so that the customer engineer can perform the appropriate maintenance operations according to the displayed information. The maintenance panel is also equipped with a General Purpose Interface Bus (GP-IB) connector. An FMD (Fujitsu Maintenance Devices) can be connected through this connector. The FMD is a laptop PC with special maintenance software. The FMD is used for internal diagnosis, remote maintenance, module configuration definition, subsystem configuration definition, and microprogram download.

The SA module in the FCU records error data on a 5 1/4-inch hard disk. The FMD checks and references this information so that the customer engineer can locate the failed area and pinpoint those areas that may fail in the future. These FCU features facilitate fast repairs and preventative maintenance. The FCU is also equipped with a time of day (TOD) clock.

Error data is recorded on a hard disk with the time the error occurred. This makes it easy to check errors and determine their relation to the error data stored in the host.

6. Logical structure of microprogram

6.1 Classification of modules

The F1700B/S microprogram modules can be classified according to function. There are task-operating modules and task-managing modules. The task-operating modules include the CA which controls the channel path operations and the DA which controls the device path operations. The management modules include the RM which performs concentrated management of resources, CFE which manages cache memory allocation, and SA which manages the disk subsystem configuration. Based on internal information and data flow, the modules can be classified into two types: single bus modules and dual bus modules. The CA, DA, and SA are single bus modules. The RM and CFE are dual bus modules. Single bus modules handle operation depending on a path. Dual bus modules manage shared information. Each of these modules transfers control data through the common bus in the FCU. The module corresponding to the cache memory segment is called the SS. This module is a task operating module and a dual bus module. However, the SS is the only module that consists of only hardware. The SS is a passive module with no microprocessor. For processing tasks using the cache memory, control data is set by the CA or DA. Data transfer is also performed by the CA or DA.

6.2 Task control block (TCB)

Input/output requests from the host are queued and processed as task-control-blocks (TCB) in the FCU. The track position information and task procedure of specified drives are recorded in the TCB. The TCB is created not only for taking I/O requests from the host but also for processing tasks in the FCU that re-

Table 2. Flow of device access operations

Communication modules	Operation
CH → CA	Receives information of a track where data to be processed is located in the specified drive.
CA → RM	Creates TCB based on information received. Temporarily disconnects CH.
RM → DA	Instructs starting of locating data on a specified drive.
DA → DV	Selects a specified drive, instructs starting of positioning, initiates single DV operation.
DV → DA	Reports completion of positioning.
DA → RM	Reports completion of positioning.
RM → CA	Instructs connection with CH.
CA → CH	CH connecting
CA → RM	Reports completion of CH connection.
RM → CA/DA	Coupling instruction for read/write data.
CH+CA/DA+DV	Data processing.
CH → CA	Task completed.
CA → DA	Task completed.
DA → RM	Task completed.
RM → CA/DA	Instructs release of resource.

Table 3. Flow of read-hit operations

Communication module	Operation
CH → CA	Receives information of a track where data to be processed is located in the specified drive.
CA → RM	Creates TCB.
CA → CFE	Requests checking to see if the specified track is in the cache memory.
CFE → CA	Track is in cache memory (hit processing).
CA → SS	Reads data from cache memory.
CH → CA	Task completed.
CA → CFE	Task completed.
CA → RM	Task completed.
RM → CA	Instructs release of resource.

quires resources. Up to 1 024 TCBs can be created for the subsystem. Up to eight TCBs can be created for each device (normally, one of these would be operating while the remaining seven would be queued). As these TCBs are transferred, multiple task-operating modules work together to execute I/O requests from the host.

6.3 Device access operation

Table 2 shows one of the general processing flows for an input/output operation of reading or writing data directly from or to a disk drive. Here, CH means a channel and DV means a device. The arrows show the flow of control information and data.

6.4 Cache access operation (read hit)

Table 3 shows the general processing flow for a read-hit operation in which the specified data is in the disk cache memory.

7. Expanded functions

As explained in the previous sections, the F1700 FCU series has a modular structure designed for flexibility to adjust to future function modifications and additions without changing the basic architecture. Starting with the disk cache function and nonvolatile storage (NVS) for high speed writing, already in the market, there are plans to further expand the functions in the F1700 series.

7.1 Dual volume control function (DVCF)

Several approaches to protect disk data have been considered. Among them, duplicating disk data based on disk cache and high speed write operation is scheduled.

This function would hold the same data in two different drives. Thus, if one of the drives fails, the remaining data on the drives would be used to continue the task in progress. Currently, Fujitsu's system combines hardware and software to create a DVCF. However, Fujitsu plans to enable processing tasks in the subsystem to minimize the software load.

7.2 Asynchronous data transfer function

Currently, the data transfer process in the drive is synchronized with the rotation of the disk media (except in cache operations). The rotational speed of the disk directly affects the performance of the subsystem in terms of access time so that it should be faster. However, it is still difficult to reduce the amount of time required for data processing between data records on a track because communication between the FCU and channel should be completed within this time period. Furthermore, system expansion in recent years demands that an external storage unit be located away from the mainframe. This means that the turn-around time for communication between the FCU and channel will become longer.

To satisfy these demands, the input/output operations should be separated into two parts and processed asynchronously. One process for communication with the channel and the other process for communication with the disk drive. With asynchronous data transfer, the input/output operations can be performed independently with the turn-around time for the channel and the rotational speed of the disk drives.

From a technical viewpoint, the disk cache function is a very specialized application of asynchronous data transfer, since data can be transferred to or from the disk cache memory in hit cases without having to access the disk drive at this time. This disk cache function is therefore applicable to all other input/output operations as the asynchronous data transfer function.

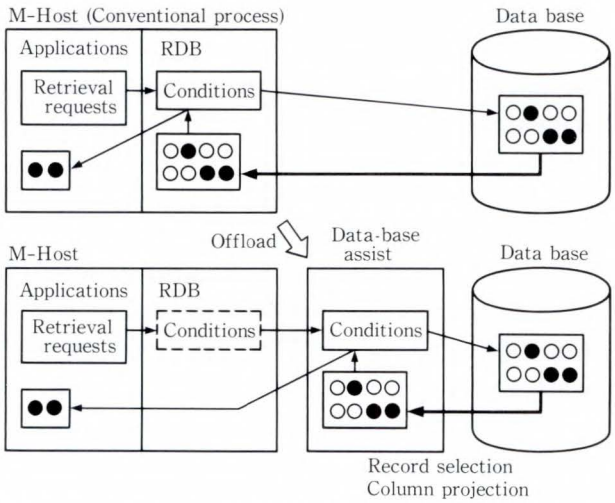


Fig. 3—Data base assist function (Offloading data-base retrieval functions to File Control Unit).

Furthermore, this asynchronous data transfer function would also facilitate effective use of disk capacity and extended cable length which would make the system construction more flexible.

7.3 Data base functions

In recent years, as analysis of internal operations in the host computer has focused on data-base processing, which provides a rather large load, ways to reduce this processing load are being explored[6)-8)].

Since most data-base operations involve referencing and sorting of data, data-base processing could be speeded up and the host load reduced if the processing could be performed in the file subsystem (see Fig. 3).

7.4 Control of various file devices

As previously stated, the F1700 FCU has a flexible hardware structure that will be able to accommodate newly developed file devices by altering the DA section. It will be possible to back-up files in the file system. That is, the F1700 FCU is connected to back-up devices such as magnetic tape and optical disks together with magnetic disks[9)].

All of the functions outlined above are designed to accommodate future expansion. There are still a few technical issues to be solved, but Fujitsu is devoting all its efforts in improving the file subsystem technology. Fujitsu will soon achieve further expansion of performance that will meet the needs and expectations of all Fujitsu users.

8. Conclusion

Fujitsu has developed the F1700B/S FCU as the control unit for the F6427H disk drive unit. The F1700B and F6427H subsystem can provide approximately 240 Gbytes of storage capacity. As the users' system configuration grows yearly, the disk subsystem should be shared by multiple systems. To meet this demand, this subsystem can provide better performance such as 32 channel paths, 16 device paths, aggregate throughput of 72 Mbyte/s, and 128 disk drive configuration.

Because of the excellent architecture of the F1700 FCU series, function expansions in the near future such as the data-base assist function and other various type of device control are very easy to implement.

Fujitsu will continue to develop external storage subsystems as state-of-the-art products based on the flexible architecture of the advanced file controller.

References

1) Yanagida, T.: Double-pole Development for Magnetic Disk Drives and Manifold Applications for Storage Devices. (in Japanese), *NIKKEI COMPUTER*, 1989-06-06, pp. 61-70.
2) Kaneko, S., Kikuchi, J., and Tanaka, R.: Disk Cache System. (in Japanese), *FUJITSU*, **34**, 2, pp. 253-261 (1983).
3) Kaneko, S.: Disk Cache System for High-speed Disk Access in Large General-Purpose Computers and Performance Evaluation. (in Japanese), *NIKKEI ELECTRONICS*, **364**, 1985-03-11, pp. 205-232.
4) Kuwabara, K.: Proven Effectiveness of Disk Cache and Solid State Disk to Improve the System Performance. (in Japanese), *NIKKEI ELECTRONICS*, **364**, 1985-03-11, pp. 159-187.
5) Hirano M.: Looking on Disk Cache Function to Fill the Access Gap. (in Japanese), *NIKKEI COMPUTER*, 1982-03-22, pp. 71-85.
6) Kuwabara, K.: Development of the Data Base Machine by the Hardware Implementation of Search/Sort Processing. (in Japanese), *NIKKEI*

ELECTRONICS, **458**, 1988-10-17, pp. 114-115.
7) Yanagida, T.: Development of DB/DC for the Large Amount of Transaction Processing. (in Japanese), *NIKKEI COMPUTER*, 1990-10-15, pp. 172-175.
8) Makinouchi, A., Ishikawa, H., and Suzuki, F.: Architecture of Multi-media Data Base System. Res. Rep. Inf. Proc. Soc., Jpn., (DATA BASE SYSTEM), 85-DB-50-4 (1985).
9) Inoue, Y., Ogawa, Y., and Kaneko, S.: External Storage Devices. (in Japanese), *FUJITSU*, **38**, 5 (Special Issue: Information Processing Devices), pp. 347-362 (1987).

Hiroyuki Takizawa
Engineering Dept.
File System Div.
FUJITSU LIMITED
Bachelor of Electronics Eng.
Keio University 1969
Specializing in Development of Magnetic Storage Systems

Joji Kikuchi
Engineering Dept.
File System Div.
FUJITSU LIMITED
Bachelor of Electronics Eng.
Hokkaido University 1972
Specializing in Development of Magnetic Storage Systems

Yasuo Kurihara
Engineering Dept.
File System Div.
FUJITSU LIMITED
Bachelor of Applied Electronics Eng.
University of Electrocommunications 1973
Specializing in Development of Magnetic Storage Systems

UDC 681.327.634

F6427H Magnetic Disk Subsystem: HAYABUSA

• Takashi Koike • Toshio Negoro • Teiji Yoshida

(Manuscript received June 4, 1990)

This paper outlines the history of Fujitsu's development of large disk subsystems, discusses recent market trends regarding large disk subsystems, and explains why the F6427H subsystem was developed. It then discusses the features of the F6427H subsystem, the system performance and reliability, and the technologies used in the F6427H magnetic disk drive to achieve high efficiency and performance.

1. Introduction

In 1981, Fujitsu developed the first disk enclosure (DE) for the F6421 magnetic disk drive. To improve reliability, this DE[1] contains all mechanical recording components, for example, the magnetic media and head, in a clean environment that is completely isolated from the external environment. Fujitsu has continued to develop large magnetic disk drives using this DE technology[2),3)]. Some examples of these products are: the F6425A4/B4/C4 magnetic disk drive and the F1774C disk controller (developed in 1983); and the F6425M4/N4 magnetic disk drive with double storage capacity and the F1774D disk controller (developed in 1984).

In 1988, Fujitsu developed the following: the F1700A file controller with improved disk controller functions, the F6425GA4/GB4 with a storage capacity equal to the F6425C4/B4 but with a higher performance, and the F6425HA4/HB4 with a tripled storage capacity.

Since the F6421 series, Fujitsu has been developing large capacity, high performance, and high reliability products using 10.5-inch DE technology.

Meanwhile, the performance of general-purpose computers has been growing; thus large-capacity, high-performance, and high reliability magnetic disk drives are required to form the nucleus of external storage equipment. Also, because of social factors such as the rapidly rising land prices in metropolitan areas, there is a growing demand for small energy-saving units. To meet these demands, Fujitsu developed the F6427H magnetic disk drive using an 8.25-inch DE instead of the conventional 10.5-inch DE. The DE technology used for this drive enables a recording density of 69.3 Mbit/sq inch, which is double the recording density of the F6425/HA4/HB4 disk drive.

The F6427H magnetic disk drive and the F1700 file controller form a file subsystem. This subsystem can be used in the Fujitsu M-series general-purpose computers and VP2000-series supercomputers as a successor to the F6425 magnetic disk subsystem.

This paper discusses the features, specifications, and technologies of the F6427H subsystem.

2. F6427H subsystem

2.1 Features of the subsystem

Up to sixteen DEs can be mounted on the F6427H magnetic disk drive in units of four DEs. Each DE has a storage capacity of 1.89 Gbytes. The storage capacity per unit can be from 7.56 Gbytes to 30.24 Gbytes in units of 7.56 Gbytes. Two types of F6427H magnetic disk drives are available: type HA and type HB.

Type HA has internal string controllers, whereas type HB does not. One string consists of one HA unit and one HB unit. Each string has a storage capacity of up to 60.48 Gbytes.

Table 1. F6427H subsystem specifications

Magnetic disk subsystem		F6427							
		HA 1	HA 2	HA 3	HA 4	HB 1	HB 2	HB 3	HB 4
String controller		Four units				None			
DE/unit		4	8	12	16	4	8	12	16
Storage capacity	/Unit (Gbytes)	7.56	15.12	22.68	30.24	7.56	15.12	22.68	30.24
	/DE (Gbytes)	1.89 (Format)							
Medium configuration	Cylinders	2 655 + 1 (Alternate) + 2 (Customer engineer) + 3 (Surface analysis)							
	Heads	15 (data) + 1 (servo)							
Positioning time	Maximum	25 ms							
	Average	12 ms							
	Minimum	2 ms							
Latency time		6.91 ms (4 340 rpm)							
Data transfer rate		4.5 Mbyte/s							
Data format		CKD variable							
Subsystem	Controller	F1700A/B/S							
	Control DE (max)	F1700A/S : 64 F1700B : 128							

Fig. 1 – F6427H subsystem.

One subsystem consists of up to four strings and has a storage capacity of up to 241.92 Gbytes.

Table 1 lists the approximate specifications for the F6427H subsystem. Figure 1 shows the subsystem.

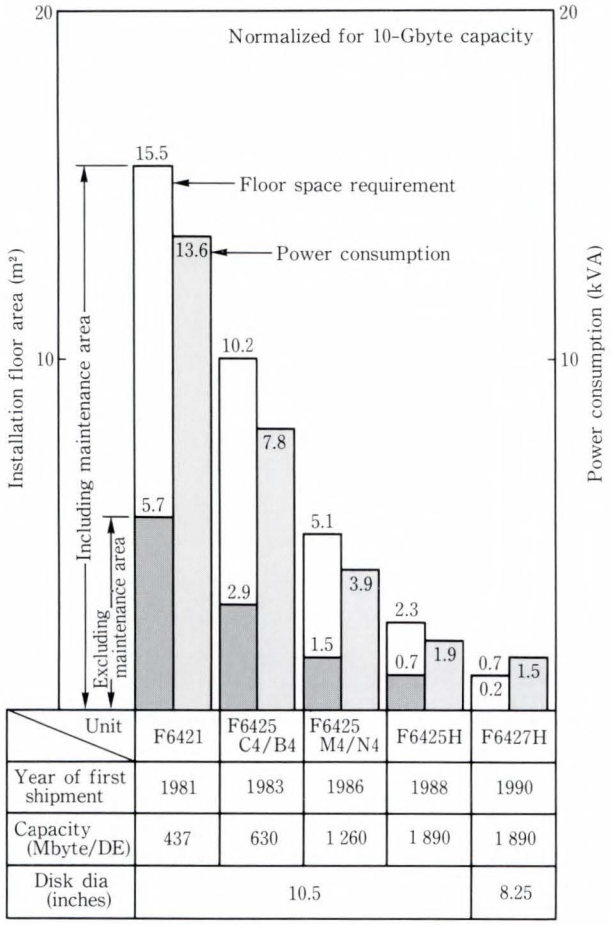
The DE for the F6427H magnetic disk drive consists of an advanced metal sputtered device and a thin-film head. Fujitsu has achieved high density recording in a low-power small-frame drive for the first time using 8.25-inch DE technology. A high-speed rotary actuator and a high-speed direct-drive DC spindle motor were developed for the drive. A digital signal processor (DSP) is used in the control circuit to enable precise control, and to achieve an average positioning time of 12 ms and a transfer rate of 4.5 Mbyte/s.

For high reliability, the DE is completely sealed to maintain a dust-free internal environment.

The drive is a drive module (DM) combining a DE, drive control PC-board, drive power supply, and drive cooling fan. The drive can be easily installed or removed from the unit in order to change the system configuration or perform maintenance.

The F6427H can accomodate up to 32 DMs in four columns of eight (see section 3.1), and each DM takes up the minimum of space. Figure 2 shows how the installation space and power consumption per 10 Gbytes have fallen in the last ten years.

To improve throughput and reliability, the F6427H subsystem has a 4-path crosscall function to enable any four drives within the same string to be simultaneously read and written.



Unit	F6421	F6425 C4/B4	F6425 M4/N4	F6425H	F6427H
Year of first shipment	1981	1983	1986	1988	1990
Capacity (Mbyte/DE)	437	630	1 260	1 890	1 890
Disk dia (inches)	10.5				8.25

Fig. 2—Floor space requirement and power consumption.

The error rate was reduced dramatically by using a multi-burst error correcting code (ECC) based on the Reed-Solomon coding method.

2.2 Performance of the subsystem

The F6427H subsystem has a storage capacity of 1.89 Gbytes/drive and 242 Gbytes/subsystem. Such a large capacity causes an increase in I/O operations and can cause problems with path and device capacities.

The path capacity problem can be solved by increasing the number of drive paths. To ease the path capacity problem and improve throughput, the F6427H subsystem has a 4-path crosscall function and a dynamic crosscall path (DXP) function.

The data transfer speed of the F6427H magnetic disk drive is 1.5 times that of the F6425H, and its average access speed is about 22 percent higher.

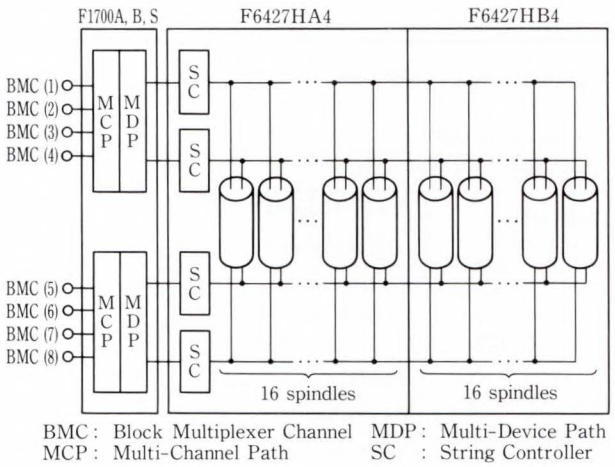


Fig. 3—Four-path crosscall configuration.

Because of mechanical restrictions, the head positioning time and latency time could not be reduced below the milli-second level. This is a major hurdle to increased device capacity. To ease device capacity problems in the F6427H subsystem, an optional disk cache featuring a high-speed write function and a large-capacity buffer memory can be installed in the file control units.

2.2.1 Four-path crosscall

Fujitsu used a 2-path crosscall function with the DXP function for the first time in the F6425C4/B4 (shipped in 1983). The DXP function consists of an extended reserve function and dynamic reconnection function.

To solve the path capacity problem caused by increasing the storage capacity of each drive and subsystem, the 4-path crosscall function can be added as an option to F6425G/H models shipped after the end of 1987. To meet the needs of subsystems which continue to grow in scale, the 4-path crosscall function was installed in the F6427H magnetic disk drive. Figure 3 shows an example of a 4-path crosscall configuration. Like the 2-path crosscall function, the 4-path crosscall function further improves system performance when combined with the DXP function.

In conventional systems, if there is more than one path between the host system and drive, a reserved drive can be activated or a mechanically operated drive can be reactivated

using the currently used path or a reserved path. The extended reserve function, however, enables a reserved drive to be accessed from another free path in the same path group. The dynamic reconnection function switches from the disconnected access path to another free path in the same group at reconnection after a seek or sector search.

If a path group is defined for the magnetic disk drive to be accessed, this function works under a host system having the enhanced channel facility (ECF).

2.2.2 Disk cache

Because of mechanical restrictions, the minimum access time of a magnetic disk drive is at the millisecond level; this is more than one million times slower than semiconductor memory devices.

These restrictions cannot easily be overcome; therefore, to dramatically improve the subsystem performance, other approaches had to be taken. So, in 1981, Fujitsu developed a disk cache for the F6421 magnetic disk subsystem[1].

The operating principle of the disk cache is to copy into the cache memory of the file controller the data in the magnetic disk drive that is frequently accessed by the host system. The copied data is reaccessed by direct transfer from cache memory instead of from the drive[4),5)]. Data in the cache memory is replaced according to a least recently used (LRU) algorithm to give priority to frequently accessed data. If the host system requests this kind of data, the data is transferred directly from the cache memory. This eliminates the mechanical operation involved in a magnetic disk access, and greatly reduces the I/O response time[6)].

Disk cache technology prevents increases in the amount of access in large-capacity subsystems (e.g. the F6427H subsystem) from increasing the I/O response time. This technology, therefore, improves the subsystem performance.

The cache memory of the F6427H subsystem can be extended up to 512 Mbytes (F1700B) in units of 32 Mbytes according to the system size and cost-performance requirements.

2.2.3 High-speed write function

Previously, Fujitsu used the write-through method for disk cache write processing. In this method, data is simultaneously written into cache memory and magnetic disk by a mechanical operation. Using this method, it was not possible to reduce the I/O response time for write processing. To solve this problem, the high-speed write function was developed.

In this method, data is simultaneously written into cache memory and nonvolatile memory.

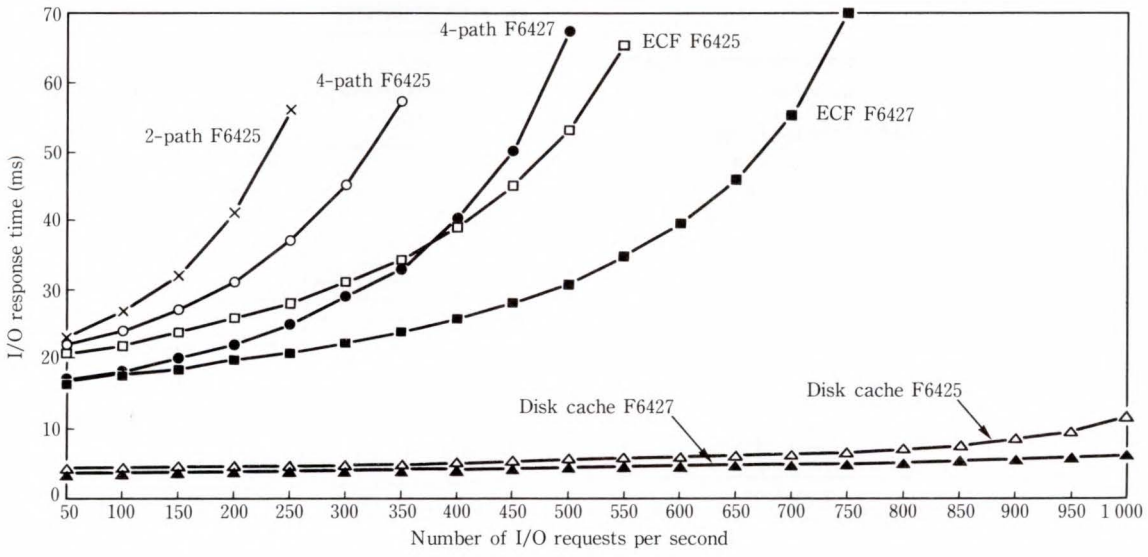


Fig. 4—Comparison of F6427 and F6425 functions.

When the write operation is finished, intervention by the host system is also terminated to reduce the I/O response time for write processing. Note that magnetic disk write processing does not require host system intervention because write back processing is automatically processed in the file controller.

The F6427H subsystem has a 4-Mbyte nonvolatile memory.

The F6427H magnetic disk drive is superior to the F6425 magnetic disk drive in both transfer speed and average access time. Figure 4 shows the results of simulations of each function of the F6425 and F6527H magnetic disk subsystems. These simulations assume that there are 32 drives per string and 4 Kbytes per transfer block. The cache function, including the high-speed write function, was simulated. As shown in the figure, when the 4-path function is used (with or without ECF support) the F6427 throughput is about 50 percent higher than that of the F6425.

2.3 Reliability of the subsystem

This section describes how the F6427H subsystem reliability has been improved.

The number of components in the file controller has been greatly reduced by using the latest technology. The power supply of each director and cache mechanism is independent or duplexed. The nonvolatile memory is backed up by a battery that enables retention of written data for up to 20 days without an external supply.

The number of components in the string controller has been reduced by using large-scale integrated circuits (LSI). For high speed and reliability, the four string controllers operate independently of each other and have no common parts.

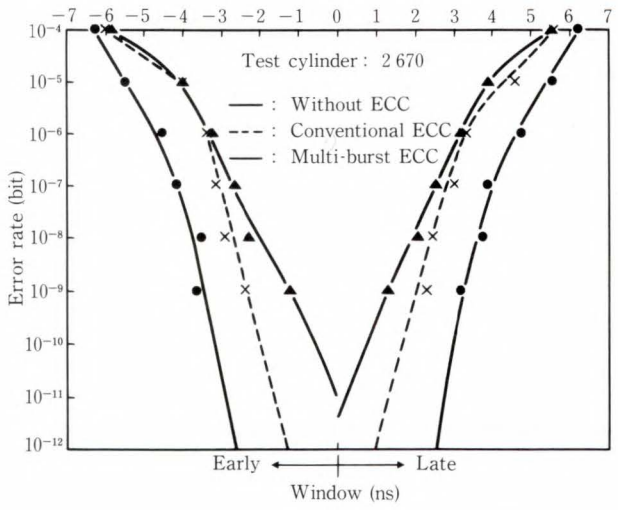


Fig. 5—Relationship between phase margin and error rate.

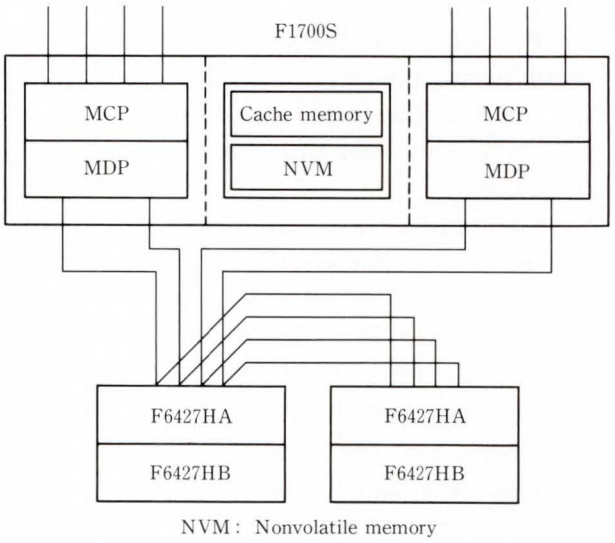


Fig. 6—F1700S 4-device path configuration.

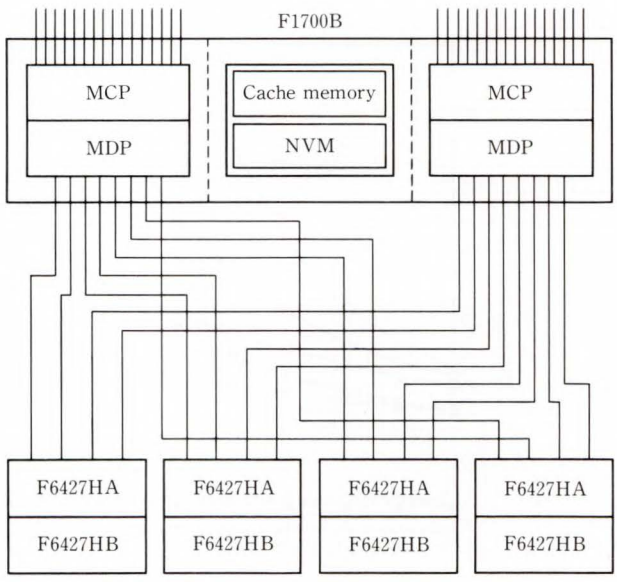


Fig. 7—F1700B 16-device path configuration.

Multi-burst error correction coding (ECC) based on Reed-Solomon coding is used to increase the read error margin.

To record code data, the error rate after correction must be 10^{-12} or better. To achieve this level using the current ECC of the F6425 magnetic disk drive, the device error rate must be 10^{-8} or better. For the F6427H magnetic disk drive, the device error rate should be 10^{-5} or better.

Since the actual device error rate is the same as in the F6425, the margin will increase by 10^3.

Figure 5 shows the relationship between the phase margin and error rate. The figure shows that the multi-burst ECC reduces the error rate to 1/1 000 that of the conventional ECC used in the F6425.

2.4 Subsystem configuration

The F6427H subsystem basically has a 4-path crosscall configuration, but various other systems can be created according to requirements.

The F1700S file controller is a low-cost model of the F1700 series. It can control up to 64 drives (120.96 Gbytes). Figure 6 shows an example of a 4-device path configuration using the F1700S. The F1700B file controller has improved functions for large-scale systems and can have a 4- to 16-device path configuration. It can control up to 64 drives when in the 4-device path configuration, and up to 128 drives (241.92 Gbytes) when in the 8-device path and higher configurations. Figure 7 shows an example of a 16-device path configuration using the F1700B.

3. F6427H magnetic disk drive

3.1 Unit structure

The installation space required by this unit has been minimized.

The cabinet contains a stack of two columns and eight rows that can accommodate up to sixteen drive modules. The mother board at the back has connectors for the drive modules.

There are two types of units: type HA, which has four string controllers, and type HB,

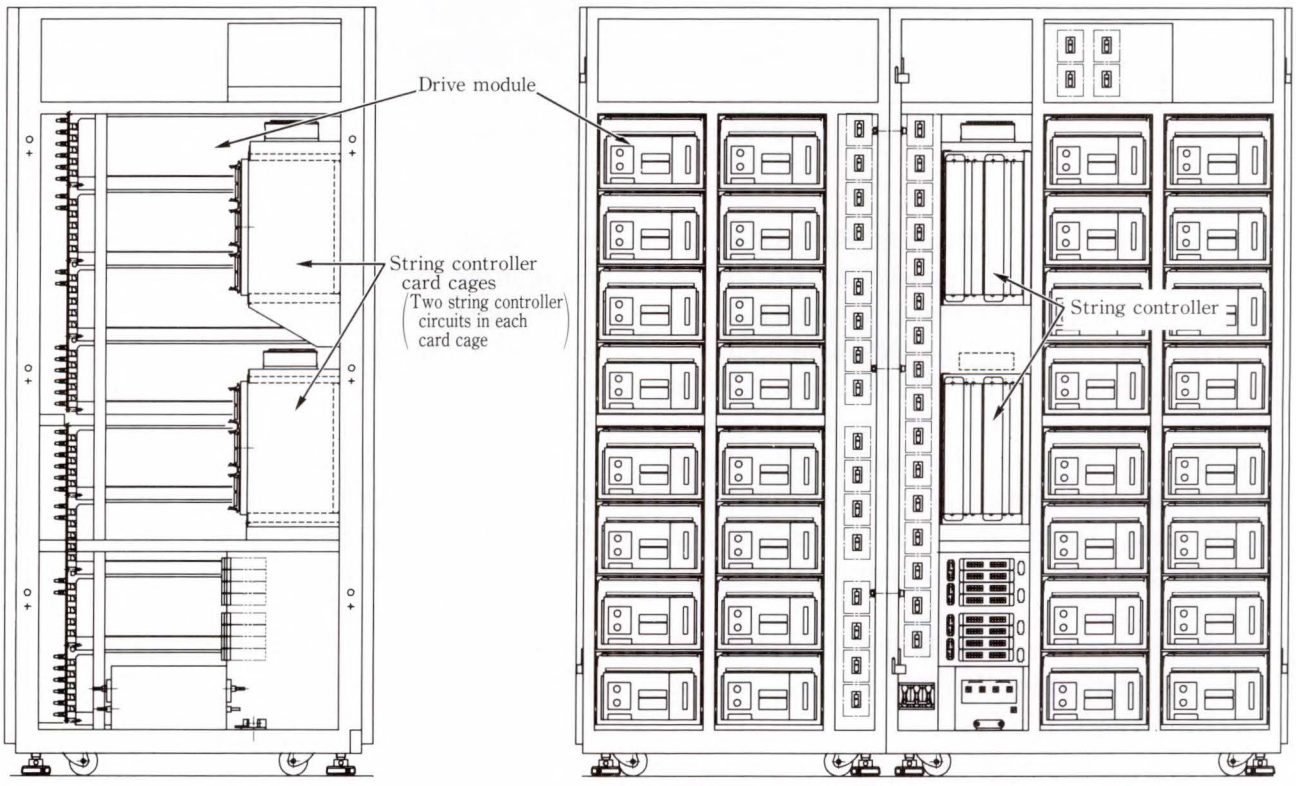


Fig. 8—Internal view of F6427H.

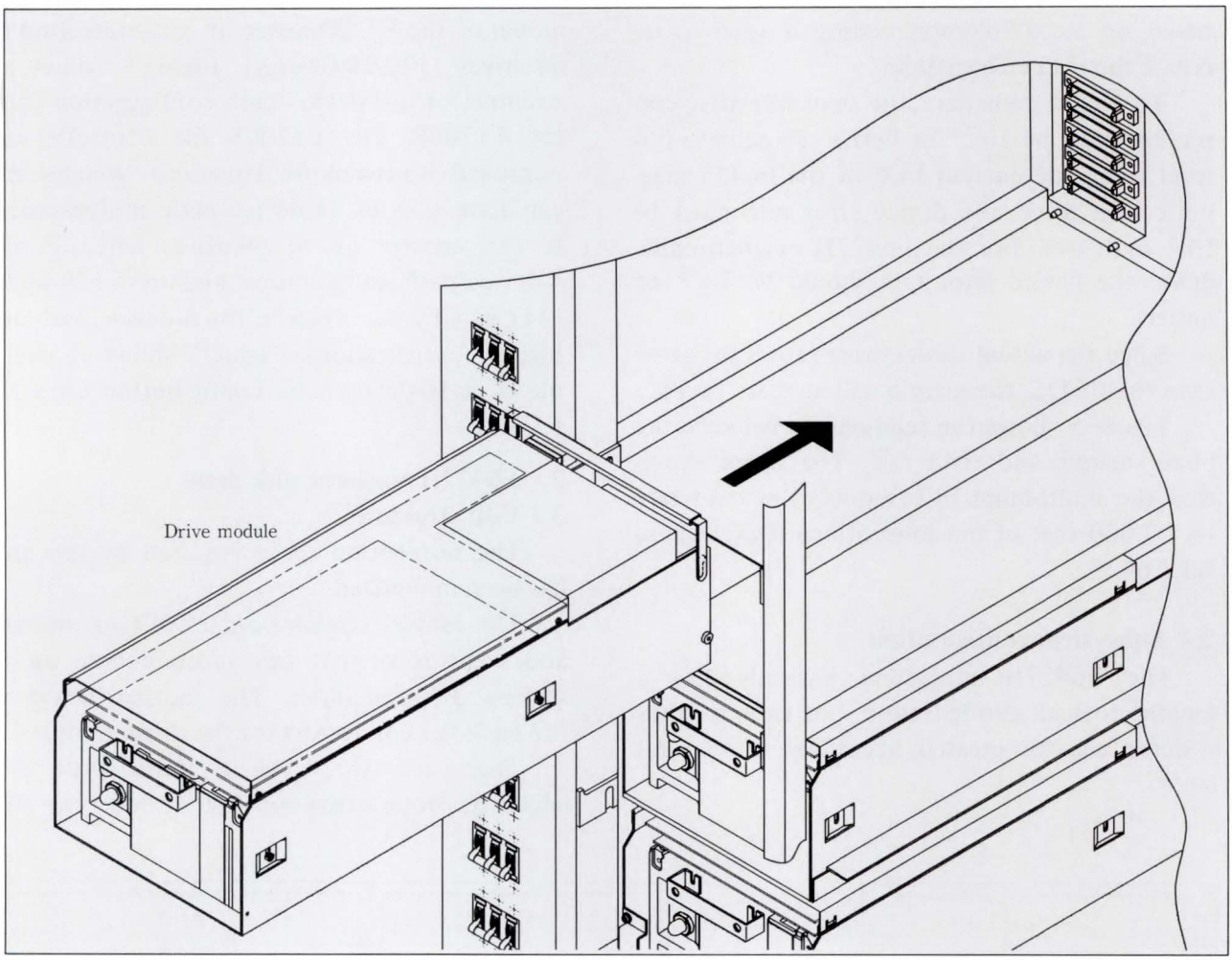


Fig. 9—Mounting the drive module.

which has no string controllers. Figure 8 shows the inside of the front panel. The HA unit is on the right, and the HB unit is on the left.

The HA unit has the PC-board and the power supply unit for string controllers in the gate at the left of the drive stack. The power supply unit has the same height and depth as a standard PC-board and is easily installed in its gate. There are gates at the top and bottom, and each gate contains two pairs of string controllers.

The HB unit has exactly the same structure as the HA unit except that it is narrower because it has no string controller.

The drive is a drive module (DM) assembly consisting of a DE, drive control PC-board, and drive power supply unit. This assembly can easily be installed or removed for maintenance or system expansion.

Figure 9 shows how the drive module is installed in the subsystem.

3.2 Drive module

To increase the storage capacity of the unit while minimizing the space used, drives must be closely mounted together on a unit. Therefore, a high-density combined drive module design was used. Figure 10 shows the drive module. The DE is loaded with drive control PC-boards. The drive power supply unit is behind the DE and the cooling fan is at the end of the unit.

Also, at the rear, there are connectors to transfer signals between the drive module and cabinet and to supply AC power to the drive

Fig. 10—Drive module.

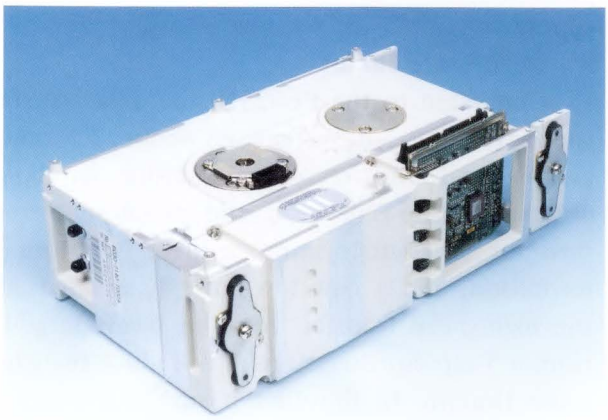


Fig. 11—F6427H DE.

module. The drive module has a duct surrounded by a metal plate to efficiently radiate heat from the drive. To minimize thermal interference between drives, air is taken in from the front of each drive and exhausted by a ventilation fan at the back.

3.2.1 DE structure

For the F6427H magnetic disk unit, a special DE was developed using 8-inch technology. Figure 11 shows the DE for the F6427H. This DE uses an 8.25-inch medium instead of the conventional 10.5-inch medium used for Fujitsu's large magnetic disk units. This change makes it more compact and reduces size and power consumption.

Table 2 lists the specifications of this DE.

Table 2. F6427H DE specifications

Storage capacity	Spindle	1 890 Mbytes (format)
	Track	47 476 bytes (format)
Data transfer rate		4.78 Mbyte/s
Disk	Diameter	8.25 inches
	Number	9 platters
Recording density	Track density	2 080 tpi
	Linear density	33 310 bpi
Average positioning time		12 ms
Disk rotational speed		4 340 rpm

A low-noise metal sputtered medium with a strong coercive force and a thin-film head is used to achieve a high recording density. In addition, the DE incorporates a quick-access in-line gimbal, the medium is protected with a carbon film, and the lubricant used was carefully selected to enable the stable low-level floatation required for high-density recording.

The voice coil motor (VCM) has a vertical moving coil. To reduce vibration and enable easy access, its pivot of rotation is aligned with the carriage's center of gravity.

The spindle motor consumes moderate amounts of power even during high-speed operation. The DE has a rigid structure and its resonant frequency is set high so that high-speed positioning does not cause vibration.

3.2.2 Drive controller

The drive controller rotates the disk according to commands from the string controller, and executes positioning and read/write operations. Figure 12 shows the block diagram of the drive controller.

To improve the S/N ratio of low-level signals (e.g. read/write signals), the wires between the drive controller PC-board and the DE were reduced by mounting the board directly on the DE. The controller consists of a logic circuit, servo circuit, and read/write circuit.

1) Logic circuit

The logic circuit has four parts: a controller for the string controller interface, a controller for read/write operations, a servo controller for the voice coil and spindle motors, and a disk rotation detector. The interface controller and

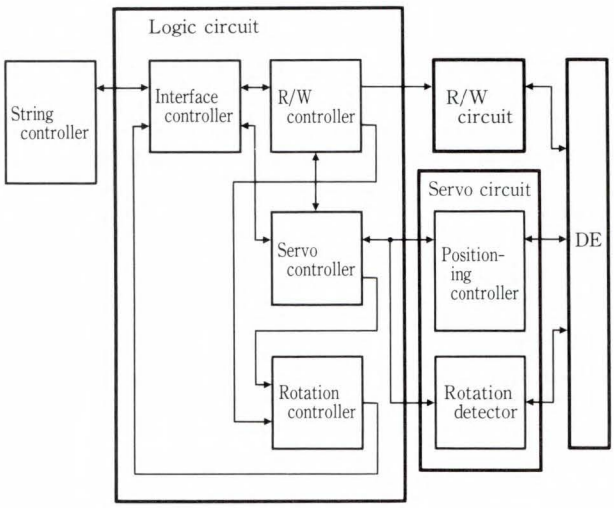


Fig. 12—Drive controller.

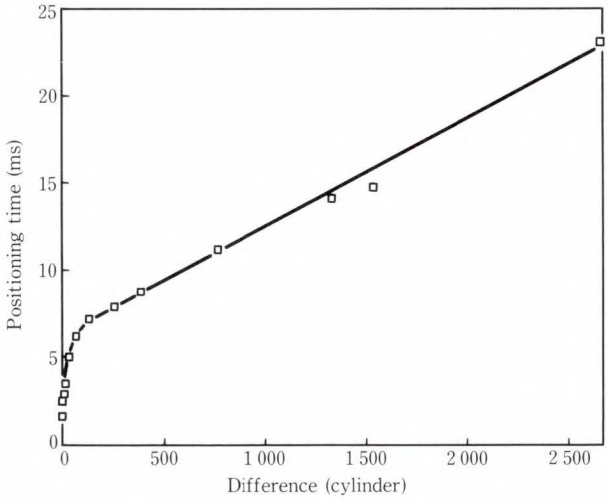


Fig. 13—Relationship between positioning time and difference distance between head and data positions.

servo controller have separate microprocessors (MPUs) to enable flexible and precise control.

2) Servo circuit

The servo circuit consists of a positioning controller and a rotation controller. The positioning controller (voice coil motor controller) uses a digital signal processor (DSP) to provide the precise control necessary for high-speed access. Figure 13 shows the relationship between the positioning time and the distance between the head and data positions.

The rotation controller has a dedicated MPU, and the spindle motor is driven at high-efficiency by a pulse width modulated (PWM) supply.

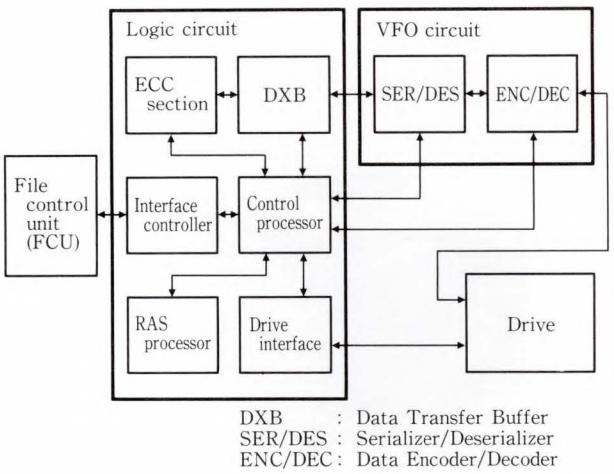


Fig. 14—String controller.

3) Read/write circuit

Compared to the F6425H, the data transfer speed of this subsystem is about 1.5 times greater and the recording density is about double. A thin-film head and metal sputtered medium are used to improve the S/N ratio and resolution, and a new low-noise IC is used for the high-speed transfer thin-film head. In addition, a 5-tap equalizer is used to reduce the edge noise (unique to thin-film heads) and produce stable read/write characteristics.

3.3 String controller

The string controller can control a string of up to 32 drives under the F1700 file controller. The F6427HA magnetic disk drive has four string controllers and enables each drive in the string to use up to four paths. This greatly improves the string throughput and reliability. Figure 14 shows the block diagram of the string controller.

The string controller contains a logic circuit and a variable frequency oscillator (VFO) circuit.

1) Logic circuit

The logic circuit has six parts: a processor to control the overall operation of the string controller (this processor is called control processor), a controller for the file control unit (FCU) interface, a drive interface, a data transfer

buffer for format control by the control processor, an ECC section to generate the ECC and to detect and correct data errors, and a RAS processor which provides RAS functions such as self-diagnosis of the string controller and microprogram loading into the control processor at power-on.

2) VFO circuit

The VFO circuit has two parts: a serial/parallel converter which converts read/write data between the FCU and drive, and a data encoder/decoder which supplies a control clock signal to each section in the string controller. This signal is synchronized with a servo signal or with the reading of faw data from the drive. In a write operation, the encoder/decoder encodes the serial data from the serial/parallel converter to 1/7 run length limited (RLL) code, and then sends it to the drive. In a read operation, the encoder/decoder decodes the data from the drive back into serial data, and then sends it to the serial/parallel converter.

Large-scale ICs are packaged at high density in the string controller to reduce the average wire length and improve the data transfer speed. The multi-burst ECC assures high data quality despite the high recording density. The system uses the same 1/7 RLL method for data recording as the current F6425G/H. However, the coding method has been modified and the number of error propagation bits has been reduced to improve quality.

4. Conclusion

Fujitsu has developed a disk enclosure for the F6427H, and has successfully achieved the high recording density of 69.3 Mbit/sq inch using a thin-film head and an 8.25-inch metal-sputtered medium.

The disk enclosure has a storage capacity of 1.89 Gbytes, and the F6427H subsystem 241.92 Gbytes. For high performance operation, the subsystem also has a 4-path cross call function, a high-speed write function, and a disk cache. In the field of magnetic disk subsystems, there is a growing market demand not only for high performance, large capacity, and low bit-costs, but also for small physical size, small floor-space requirements, and low power consumption.

To meet these demands, Fujitsu is developing new advanced technologies and using them to create new products.

References

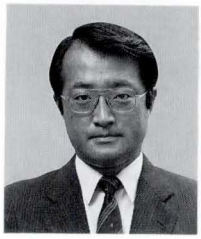
1) Miura, Y., Suzuki, M., and Sugihara, J.: FACOM 6425 Disk Subsystem and Future Disk Storage Technology. *FUJITSU Sci. Tech. J.*, **23**, 4 (Special Issue on Computer Systems), pp. 279-295 (1987).
2) Mizoshita, Y., and Matsuo, N.: MECHANICAL AND SERVO DESIGN OF A 10 INCH DISK DRIVE. *IEEE Trans. Magnetics*, **MAG-17**, 4, pp. 1387-1391 (1981).
3) Fujino, M., Sugihara, J., and Ogawa, S.: Magnetic Disk Storage. *FUJITSU Sci. Tech. J.*, **21**, 4, pp. 395-407 (1985).
4) Kaneko, S., Kikuchi, J., and Tanaka, R.: Disk Cache System. (in Japanese), *FUJITSU*, **34**, 2, pp. 253-261 (1983).
5) Kaneko, S.: Disk Cache System for High-speed Desk Access in Large General-Purpose Computers and Performance Evaluation. (in Japanese), *NIKKEI ELECTRONICS*, **364**, 1985-03-11, pp. 205-232.
6) Inoue, Y., Ogawa, Y., and Kaneko, S.: External Storage Devices. (in Japanese), *FUJITSU*, **38**, 5 (Special Issue: Information Processing Devices), pp. 347-362 (1987).

Takashi Koike

Engineering Dept.
File System Div.
FUJITSU LIMITED
Bachelor of Electrical Eng.
Seikei University 1971
Master of Electrical Eng.
Seikei University 1973
Specializing in Magnetic Storage Systems

Teiji Yoshida

Engineering Dept.
File System Div.
FUJITSU LIMITED
Bachelor of Communication Eng.
Tokai University 1974
Specializing in Magnetic Storage Systems



Toshio Negoro

Engineering Dept.
File System Div.
FUJITSU LIMITED
Bachelor of Electrical Eng.
Nagoya University 1971
Master of Electrical Eng.
Nagoya University 1973
Specializing in Magnetic Storage Systems

UDC 681.327.634

F6490 Magnetic Disk Subsystem: DIA

• Tomohisa Oyama • Yuji Ogawa • Kazeo Sugiyama

(Manuscript received June 3, 1990)

An array disk is a data storage subsystem that runs several of magnetic disk drives in a parallel configuration to achieve a very high transfer rate. High reliability is also achieved by utilizing a parity disk drive.

DIA or Disk in array is the first array disk subsystem for Fujitsu, achieving a transfer rate of 36 Mbyte/s.

This paper describes the configurations and performance of the subsystem in addition to a consideration on reliability.

1. Introduction

Applications that require processing of large amounts of data at high speed are increasing. Example applications are image processing using vector processors and engineering calculations that require fast input and output of large amounts of data. The DIA subsystem was developed to meet these needs.

As the name indicates, DIA or Disk in array is a magnetic disk subsystem based on an array disk concept.

2. Array disk concept[1)]

An array disk subsystem is a number of disk drives configured into an array. Simultaneous read/write operations occur within the array disk. Since data is handled in parallel, a high transfer rate can be achieved.

Assume, for example, that eight disk drives of 1.89-Mbyte storage capacity and 4.5 Mbyte/s transfer rate are configured into an array. The array will have a 15-Gbyte capacity and 36 Mbyte/s transfer rate.

In addition, the following configuration achieve higher data integrity;

1) An additional "parity disk drive" is provided.

Parity data is generated from the data distributed to the eight drives and saved on a parity disk drive. Even if one of the eight drives fails, the data can be reconstructed by using the parity data saved on the parity disk drive.

2) An additional "spare disk drive" is provided.

The data of a failed disk drive is reconstructed to the spare disk drive. When the data of a failed drive is reconstructed on the spare disk drive, the spare disk drive is used instead of the failed disk drive.

As configured above, an array disk offers a very high transfer rate and fault tolerant disk subsystem.

3. Subsystem outline

The DIA subsystem consists of an F6490A array disk controller and an F6490B array disk unit. Figure 1 shows the logical block diagram. Figure 2 shows the subsystem.

3.1 F6490B array disk unit

This unit consists of ten high-performance 8-inch magnetic disks (M2671 or Swallow 7) developed by Fujitsu. This unit operates as a single disk drive. Eight disks are used for data, one disk is used for the parity disk drive, and one disk is used for the spare disk drive.

A spindle synchronization function keeps

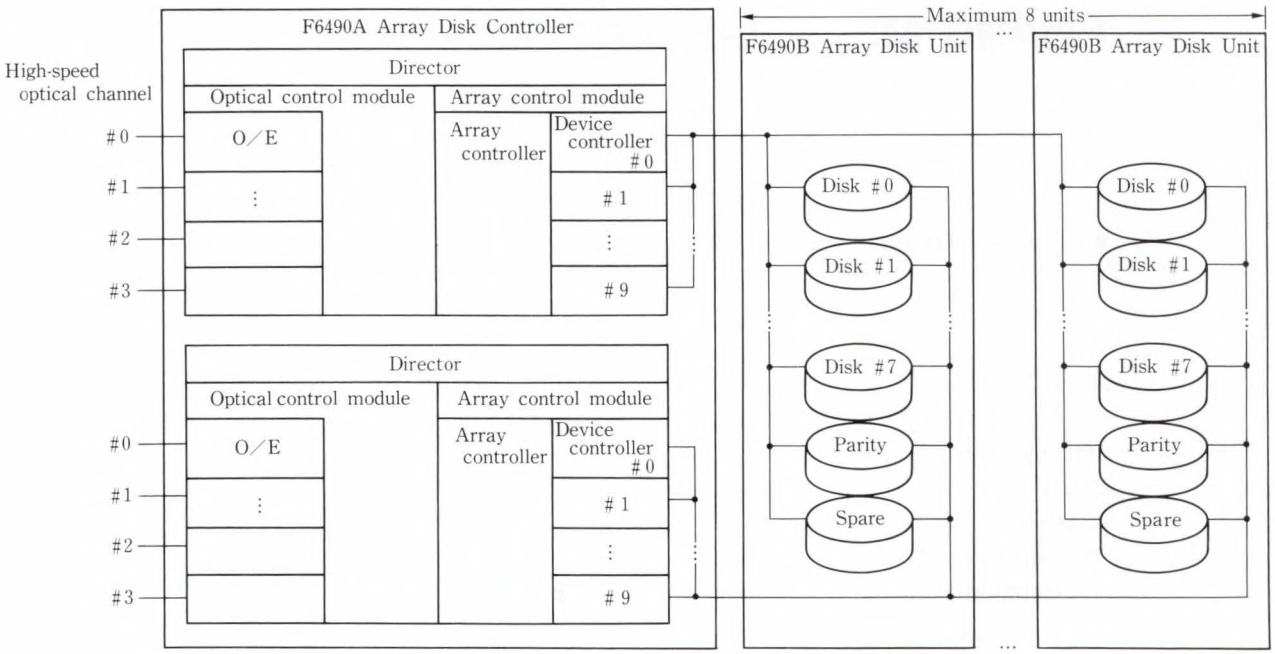


Fig. 1—Diagram of the subsystem.

the rotation of each disk synchronized. Data can be read or written simultaneously at the target positions on all disk drives, which minimizes the access time.

Each disk drive is integrated in a drive module consisting of a disk enclosure, control board, power supply, and fan (see Fig. 3). The drive module can be removed easily from the array disk unit for maintenance.

3.2 F6490A array disk controller

This controller consists of two directors (or controllers in a narrow sense). Each director consists of two modules: An optical control module to control the optical channel interface and an array control module to control the array disk units. The latter is divided into ten device controllers to control each disk drive and an array controller to assemble and disassemble the data to or from each device controller.

The two directors in an array disk controller are independent of each other. The directors can control two array disk units at the same time. This configuration enables array disk units to be accessed from one director even if the other director becomes inoperable.

The dynamic path reconnection function

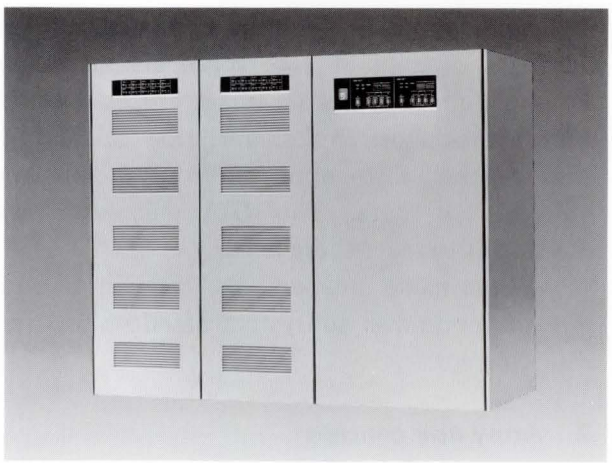


Fig. 2—F6490 Magnetic Disk Subsystem.

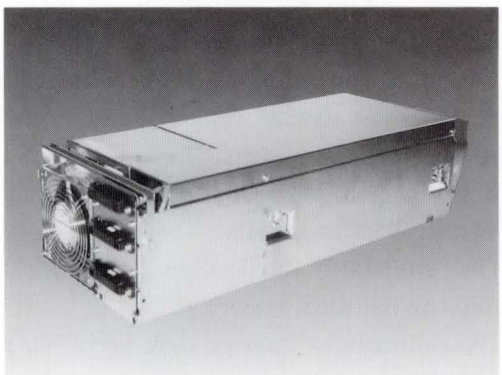


Fig. 3—Drive module.

Table 1. General specification for M2671 and F6490

M2671 magnetic disk drive

Disk Size	8 inches
Storage capacity	2.6 Gbytes (Unformatted)
Average positioning time	12 ms
Latency	6.9 ms
Data transfer rate	4.5 Mbyte/s
Interface	IPI-2
Dual port	Standard equipped
Spindle synchronization	Standard equipped

F6490 magnetic disk subsystem

Storage capacity for subsystem		15-120 Gbytes (15 Gbytes increments)
Data transfer rate		Maximum 36 Mbyte/s
Dimensions	F6490A array disk controller	$815^{l} \times 815^{b} \times 1\,690^{h}$ (mm)
	F6490B array disk unit	$560^{l} \times 815^{b} \times 1\,690^{h}$ (mm)
Weight	F6490A array disk controller	350 kg
	F6490B array disk unit	400 kg

enables data to be transferred through one of the directors, whichever is available. This function assures high subsystem throughput.

Since up to eight array disk units can be connected to the array disk controller, the subsystem capacity can range from 15 Gbytes to 120 Gbytes.

In addition, up to eight channels can be connected to the array disk controller and several host systems may share the subsystem.

Table 1 lists the performance specifications of the DIA subsystem (F6490) and the disk drive M2671 (or Swallow 7).

3.3 High-speed data transfer

A high-speed optical I/O interface and high-speed optical channel were developed to enable high-speed data transfer between the channel and DIA. The transfer rate is 36 Mbyte/s. The transmission distance is up to 1 km. The connection cables are light-weight for easy installation. To support the high-speed optical I/O interface, Fujitsu developed two types of LSI chips, one with 20 000 gates and one with 75 000 gates.

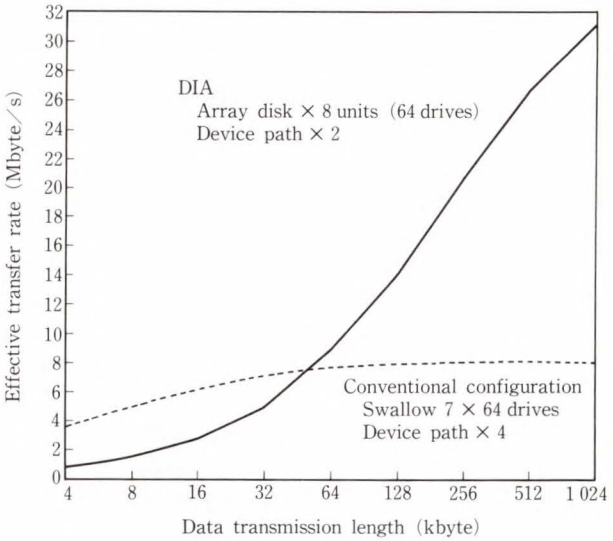


Fig. 4—Comparison of effective transfer rate.

4. Performance evaluation

A simulation was made to compare the effective transfer rate between the DIA subsystem and a conventional disk subsystem. Sixty-four disk drives were used for the simulation to make each subsystem configuration almost the same. Figure 4 shows the result.

The conventional disk subsystem has a better transfer rate if the data transmission length is short because four disk drives can be accessed at the same time. As the data transmission length becomes larger, the DIA subsystem has a better transfer rate, as expected.

5. Highly reliable subsystem

To reduce the maintenance time and keep the system fault-tolerant, parity and spare disk drives are provided.

5.1 Parity disk drive for higher reliability

The DIA receives data from a channel and divides it into bytes, then stores it on eight disk drives. The array disk controller generates a parity byte from every eight bytes and stores it on the parity disk drive.

The parity byte is used to reconstruct data if a read error or permanent malfunction occurs on one of the disk drives. Since data is reconstructed immediately, data transfer continues without any delay.

The principle of generating a parity byte

and reconstructing data can be simplified as follows;

1) The data written on the eight disk drives is added and the sum is written on the parity disk drive.
2) If any one of the eight disk drives fails, the data can be reconstructed by substracting the data stored on the seven other drives from the parity byte.

Actually, a bit-wise exclusive-OR operation is executed.

The parity disk drive ensures error-free data unless more than one disk drive fails. Data cannot be reconstructed if an uncorrectable read error or permanent malfunction occurs on more than one disk drive at the same time.

5.2 Spare disk drive for higher serviceability

The two directors in the array disk controller always monitor each disk drive in the array disk units. A faulty disk drive is automatically detached from the DIA subsystem upon detection of any permanent malfunction or inadequate statistical information of read and seek errors.

After detaching the faulty drive, the subsystem automatically reconstructs the data of the faulty disk drives on a spare disk drive. The data is reconstructed without any host activity. The data can be reconstructed within 30 min if the controller is dedicated to the data reconstruction. Since the subsystem assures exactly the same data on both the disk drive that failed and the spare disk drive, the software need not be aware that data reconstruction occurred.

After completion of data reconstruction on a spare disk drive, the data is accessible even if one more disk drive fails again.

The faulty drive replaced by the spare disk drive can be maintained without disturbing the system operations.

5.3 MTDL (Mean Time to Data Loss)[2)]

Even if one of the disk drives fails, the DIA subsystem is still operable. A new criterion of MTDL replaces the conventional MTBF (Mean Time Between Failures).

An array disk unit consists of ten disk drives and some shared components such as the operator panel and line filter. The data will not be accessible if one of the shared components fails.

The MTDL of an array disk unit is obtained from the MTDL of the ten disk drives and the MTDL of the shared components.

The MTDL of the disk drives is:

$$\mathrm{MTDL\,(D)} = \frac{\text{MTBF of 8 disk drives}}{\text{Possibility of additional drive failure during data reconstruction}}$$

$$= \frac{\mathrm{MTBF} \times \mathrm{MTBF}}{8 \times 9 \times \text{time used for reconstruction}} .$$

The MTBF represents the MTBF of a single disk drive.

The MTDL of the shared components is the same as the MTBF of the shared components. The MTDL of the shared components is:

$$\mathrm{MTDL\,(C)} = \mathrm{MTBF\,(C)}$$

The MTDL of an array disk unit must be considered in both failures. The MTDL of an array disk unit is:

$$\mathrm{MTDL\,(A)} = \frac{1}{\frac{1}{\mathrm{MTDL\,(D)}} + \frac{1}{\mathrm{MTDL\,(C)}}} .$$

An example MTDL is calculated for the configuration described up to here.

Assume and calculate:

MTBF of a single disk drive: 100 000 h
Time used for data reconstruction: 12 h
Failure rate of the shared components: 200 fit

The MTDL are approximately:

$$\mathrm{MTDL\,(D)} = 1\,320 \text{ year}$$

$$\mathrm{MTDL\,(C)} = \frac{10^9}{200} = 570 \text{ year}$$

$$\mathrm{MTDL\,(A)} = 400 \text{ year}.$$

As calculated here, it is very rare that data is not accessible.

6. Conclusion

The DIA is the first Fujitsu disk subsystem that uses the array disk concept. Software for the DIA subsystem, high-speed optical channels, and magnetic disk drives were developed. The array disk controller was jointly developed with Intellistor Inc. of the United States.

The DIA is designed for large computers system, especially vector processors. The demand for an array disk subsystem, featuring high-speed data transfer, is also increasing for engineering workstations. The technology developed for the DIA subsystem will be used for these demands. In addition, 5.25- or 3.5-inch magnetic disk drives will be used to meet space and cost requirements.

References

1) Kim M.Y.: Synchronized Disk Interleaving. *IEEE Trans. Comput.*, **C-35**, 11, pp. 987-988 (1988).
2) Patterson, D.A., Gibson, G., and Katz, R.H.: A Case for Redundant Arrays of Inexpensive Disks (RAID). Univ. Calif., Berkeley CA, UCB/USD 87/391, 1987, 24p.

Tomohisa Oyama
Engineering Dept.
File System Div.
FUJITSU LIMITED
Bachelor of Electronic Eng.
Ibaraki University 1973
Specializing in Magnetic Storage Systems

Yuji Ogawa
Engineering Dept.
File System Div.
FUJITSU LIMITED
Bachelor of Electronic Eng.
Aoyama Gakuin University 1971
Master of Electronic Eng.
Aoyama Gakuin University 1973
Specializing in Magnetic Storage Systems

Kazeo Sugiyama
Development Dept.
FUJITSU PROGRAM LABORATORIES LIMITED
Bachelor of Electronic Eng.
Shizuoka University 1972
Specializing in Magnetic Storage Systems

UDC 681.327.67

F6631 Solid State Disk: High-Speed Virtual Disk Unit

• Hajime Sugiura • Etsuo Morita • Soichiro Nagasawa

(Manuscript received June 4, 1990)

Fujitsu has developed the F6631 Solid State Disk (SSD) using 1-Mbit DRAMs as a storage element. This unit has a mass-storage capacity of 1 Gbyte and data transfer rate up to 72 Mbyte/s. In addition, the parallel processing disks with a redundant disk is used for the non-volatile data facility to ensure high reliability.
This paper describes the technology used in the F6631 SSD to provide the numerous functions, high performance, and high reliability.

1. Introduction

In recent computer systems, the data bases have become large and great amounts of data with various characteristics are stored without coherent organization. In this environment, data access is complicated and two contradictory requirements are made on external storage devices, especially on magnetic disk drive units. These requirements are large storage capacity and high-speed access. In general, the larger the storage capacity, the lower the storage cost. At the same time, the access speed is also reduced. Therefore, it is impossible for one storage device to meet these two requirements at the same time.

To resolve this problem, it is necessary to use a small performance-oriented storage device for frequently accessed files requiring fast access and a capacity-oriented storage device accessed at average speed.

The Solid State Disk (SSD) is a performance-oriented external storage device that executes the I/O processing of the magnetic disk drive unit on the semiconductor memory. It is located at the top of the memory hierarchy of the external storage device[1].

Because all positioning operations are processed electrically, the access gap between main storage and the magnetic disk drive, which is five digits or more in speed, can be supplemented.

The important technological element of an SSD is the high-speed processing of the control unit and the highly reliable non-volatile data facility.

As described above, a SSD processes I/O operation electrically. Consequently the response time depends most on the processing speed of the controller. Therefore, it is very important to use a high-speed microprocessor for the controller and high-speed internal bus.

Also, since the original data is saved only on the non-volatile data facility when the unit power is turned off, full countermeasures must be taken against data loss due to hardware failure.

Fujitsu developed the F6630 SSD using 256-Kbit DRAMs in 1985 and the F6631 in 1989. For the F6631, a 1-Mbit DRAM provides a large-capacity memory. CMOS 20-Kgate LSIs are used in the controller to minimize the size of the unit and to improve the functions and performance.

2. General description

The SSD is controlled as a small-capacity magnetic disk drive unit from the host system. The data in the disk unit is directly transferred

to the main memory in the CPU without mechanical operation such as in a real magnetic disk unit. In addition, because there is no rotational delay, this disk unit can guarantee a constant high-speed access time, without being affected by the I/O load of the channel unit, controller, or other devices. The user only has to store the data that requires high-speed response in the SSD to enable reliable access to the data at high speed.

At present, the transaction files in online systems mostly use the SSD.

2.1 Characteristics

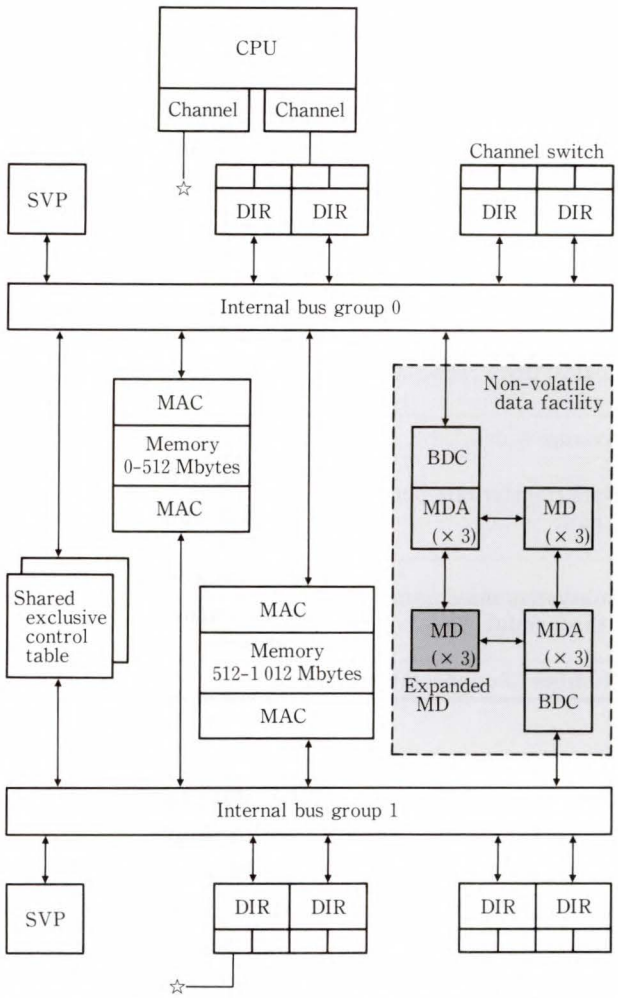
The F6631 SSD is connected to the CPU data channel via the controller: director (DIR). This unit has a maximum of eight DIRs and a maximum storage capacity of 1-Gbyte using 1-Mbit DRAMs.

Figure 1 shows the internal configuration of the F6631 SSD. The following explains the characteristics of this unit in reference to the figure.

2.1.1 High performance

A maximum of eight directors can perform complete simultaneous data transfer. When a Fujitsu M-780 optical channel unit[2] is connected, the data transfer speed is 9 Mbyte/s per connection path and 72 Mbyte/s per total subsystem. The directors and memory device are installed in one cabinet to reduce the processing time of the interface between directors and memory device. The average access speed is 0.2 ms, which is equivalent to the execution time of the positioning command string (SEEK, SET SECTOR, SET FILE MASK) in a real disk drive unit.

The memory in the unit is partitioned into a maximum of 32 logical drives. The storage capacity of the logical drive is set according to the number of cylinders allocated. Each logical drive, which is recognized by the software as a magnetic disk drive unit (F6425 disk drive unit) having fewer cylinders, can be shared by several host systems in the same way as a real magnetic disk drive unit. Each logical drive can be accessed from up to eight host systems simultaneously. The throughput when the drive is shared by host systems can be improved by a factor of up to eight over that of a real magnetic disk drive unit.



SVP : SerVice Processor
DIR : DIRector
MAC : Memory Access Controller
BDC : Backup Disk Controller
MD : Micro Disk
MDA : Micro Disk Adapter

Fig. 1—F6631 internal module configuration.

2.1.2 Nonvolatile data facility

The nonvolatile data facility which retains data while the unit power is off, is one of the most important components of the SSD.

The main component of the non-volatile data facility is a micro disk drive. The F6631 uses a parallel processing disk system with a redundant disk (parity disk). The redundant disk restores data from the rest of the disks when a failure occurs in one micro disk drive. This function facilitates the highly reliable non-volatile data facility.

In addition, in case of a power failure, an optional battery unit can be used to ensure the

Table 1. Performance specification comparison

Item		F6631A1/A2	F6630B/C/E
Technology	Memory	1-Mbit DRAM	256-Kbit DRAM
	Logic circuit	CMOS 20-Kgate LSI	TTL logic
Number of directors (DIR)		4/8 units	Up to 4 units
Emulation devices		F6425 magnetic disk drive	F6425 magnetic disk drive
Memory capacity		64-1024 Mbytes 64 Mbytes increment	32-256 Mbytes 16 Mbytes increment
Number of logical drives		Up to 32	Up to 8
Logical drive capacity		Allocation by number of cylinders (1-1 468 cylinder)	Selection from 1-, 2-, 4-, 8-, 16-, 32-, 64-, 128-Mbytes
Average access time		0.2 ms	0.3 ms
Data transfer rate per path	At BMC connection	4.5 Mbyte/s	4.5 Mbyte/s
	At optical channel connection	9.0 Mbyte/s	Not applicable
Subsystem maximum throughput	At BMC connection	36 Mbyte/s	18 Mbyte/s
	At optical channel connection	72 Mbyte/s	Not applicable
Number of connectable channels		16	16

operation of the non-volatile data facility.

2.1.3 Highly reliable technology

Almost all of the hardware components, including the internal data processing bus, power supply system, and battery unit, are duplicated. During subsystem operation, the subsystem itself constantly monitors the status of the non-volatile data facility or duplicated battery unit and power supply system. When an abnormality is detected, the subsystem sends a warning message to the host system to prompt maintenance to prevent a serious problem.

2.1.4 Easy maintenance

Maintainability is improved using the maintenance panel, which has a large liquid crystal display (LCD). All maintenance can be performed using menus.

The following operations are performed from the maintenance panel: Setting the number of logical drive partitions and their capacity, setting device addresses on the channel interface, and obtaining maintenance information.

2.2 Performance specifications

The performance specifications of the F6631 and F6630 are compared in Table 1. The performance specifications of the F6631 logical drive and F6425 magnetic disk drive unit are compared in Table 2.

Table 2. Comparison of logical drive and real disk drive drive

Item	SSD (F6425 J4/K4 emulation)	F6425 K4/L4 real disk drive
Storage capacity	Up to 630 Mbytes	630 Mbytes fixed
Track capacity	47 476 bytes	47 476 bytes
Track/cylinder	15 tracks	15 tracks
Cylinder/spindle	Up to 885	885 + 1 (alternate) + 1 (surface analysis)
Average access time	0.2 ms	23.3 ms
Seek time	(0)	(15.0 ms)
Average rotational delay	(0)	(8.3 ms)
Data transfer rate	4.5/9 Mbytes	3.0 Mbyte/s

3. Emulation method

An SSD that uses semiconductor memory as a storage media must emulate an existing disk drive unit.

The F6631 is designed to emulate the F6425 series magnetic disk drive unit. This

unit is Fujitsu's high-end magnetic disk drive unit and has the COUNT-KEY-DATA (CKD) format.

As shown in Fig. 1, the memory cards are connected to the internal buses via the MACs. The memory area is recognized as a linear memory space continuing from address 0 to the maximum address by DIR and other components.

Data is directly read and written specifying the memory address on the internal bus.

For the F6631, only DIR performs emulation concerned with the F6425 track format described below. Other components control it as linear memory space. This system localizes the components affected when emulating disk units having different track formats.

3.1 Logical drive concept

For general magnetic disk drive units, the subsystem is configured with 16 to 32 disk drives connected under one or more controllers. Regarding the F6631, one unit emulates the whole subsystem including several disk drives, using the logical drive concept.

The memory space can be divided into from 1 to 32 areas. The partitioned areas are named logical drives #0 to #n. The operating software or application software and channel unit assume the logical drive to be a small-capacity F6425 magnetic disk drive unit when performing I/O operations. The storage capacity is allocated to the logical drive in units of the cylinders of the magnetic disk.

The capacity is freely allocated to each logical drive in the range from one cylinder to the maximum number of cylinders of the emulated disk drive unit. In addition, logical drives with different capacities can coexist in one unit to meet the various needs of users. An example of the memory partitioning is shown in Fig. 2.

3.2 Track emulation

Each logical drive space allocated on the memory is partitioned and managed in units of tracks. The actual allocation is performed in an ascending order of cylinder and track address

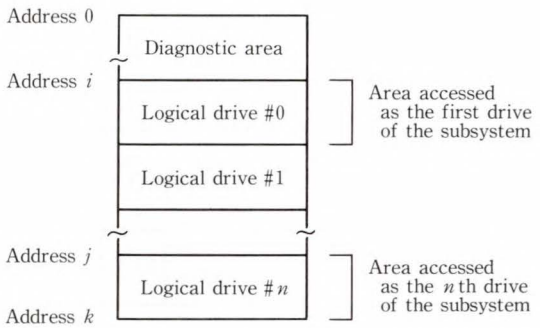


Fig. 2—Memory partitioning by logical drive.

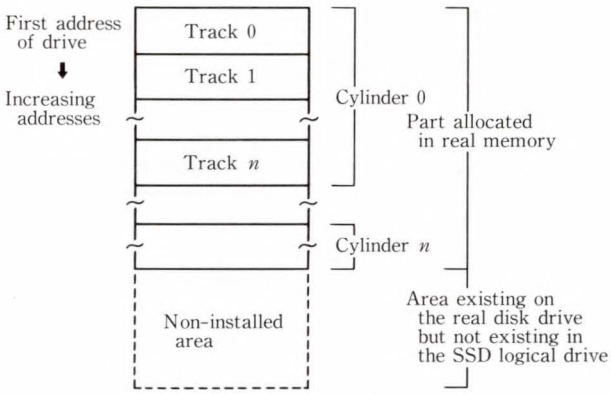


Fig. 3—Track layout.

in the same way as the real drive.

The logical drive partition and track emulation method is shown in Fig. 3.

Details of the track format is described in Fig. 4.

Each field of the track is in the form of a replacement for one track of the magnetic disk by memory. The home address (HA), record 0 (R0: COUNT, DATA field) comes first and record n (Rn: COUNT, KEY, DATA field) follows. The gap for recognizing a field limit is also emulated. This gap, which is not required for recognizing records in the SSD, is required for keeping track capacity compatibility with the F6425 magnetic disk unit. This is because a CKD format which is recorded by variable length recording, the effective total capacity of the track changes according to the number of records existing on one track.

3.3 Track directory

The track directory is a special field of the SSD for directly positioning the records

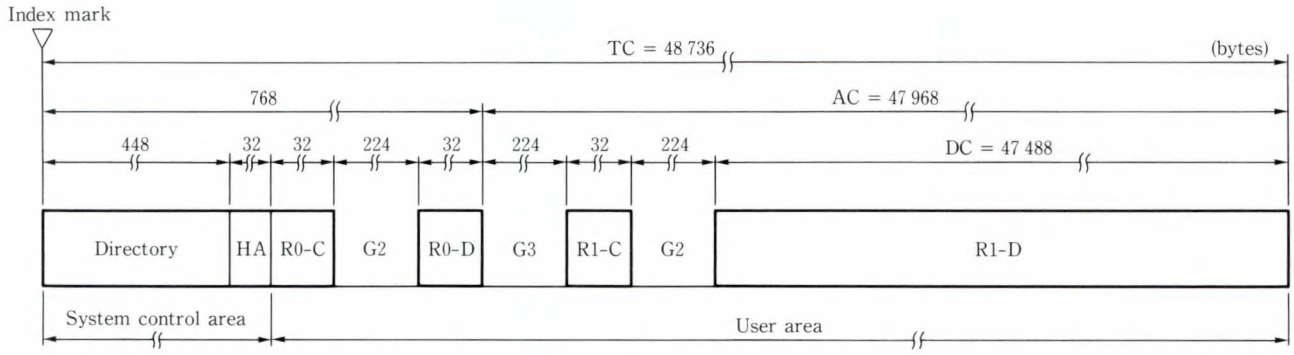


Fig. 4—Overview of track format.

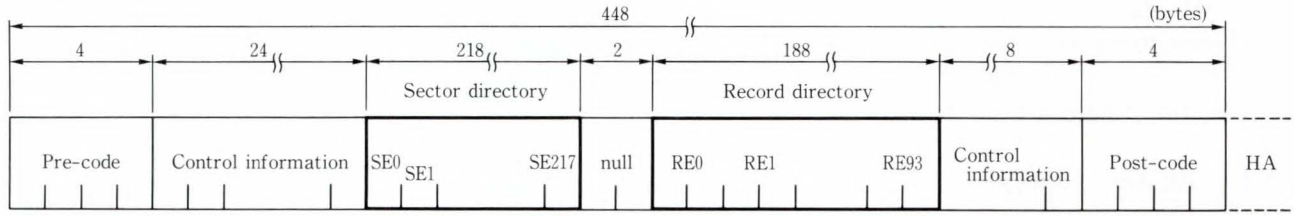


Fig. 5—Track directory format.

efficiently and at high speed with the SET SECTOR commands. The track directory is recorded at the position corresponding to the G1 gap (gap between the index mark and home address) in the track format of the real disk.

It consists of the sector directory and record directory. The relationship between the track directory configuration and records recorded on the track is shown in Fig. 5.

The sector directory has entries from sector 0 to sector 217. The contents of each entry is the pointer for the entry of the record directory which corresponds to the record number recorded at the pertinent sector position.

The record directory has entries from record 0 to 93. Record 93 is equivalent to the maximum number of records that can be written in the F6425 disk drive unit. Each entry of the record directory consist of two bytes and it describes the displacement of the header position of the COUNT field of the pertinent record from the index mark.

The track directory is updated every time a record is created or deleted on the pertinent track.

When accessing by the host system, the DIR calculates the memory address of the access track from the parameters of the SEEK commands. The DIR then reads the track directory of the pertinent track into the DIR in advance. The DIR obtains the nearest record number by referring to the sector directory from the parameters of the SET SECTOR commands. Then it calculates the displacement from the index mark of the record to be accessed based on this number.

With two directories, DIR completely emulates the processing by the SET SECTOR command and enables direct high-speed positioning to an arbitrary record on the track.

3.4 Non-installed areas and dummy data set

The storage capacity of the F6631 logical drive can be freely defined from one cylinder to n cylinders. However, the capacity is usually less than that of the real F6425 and memory is not allocated in area bounded by the dotted line in Fig. 3. This area is called the non-installed area.

The tracks in this area maintain compatibility with the actual F6425 by emulating as if only the home address and record 0 exist.

If the disk storage area is managed by the volume table of contents (VTOC), a dummy data set should be allocated to the non-installed area before use to prevent the user program accessing the non-installed area. (MSP, which is an operating system for larger Fujitsu computers, is managed by the VTOC.)

4. Performance improvement

The F6631 logical drive is shared by a maximum of eight DIRs. Each DIR has two channel interface switching mechanisms, enabling a logical drive to be connected to a maximum of 16 channels.

The F6631 provides the following performance improvements by using the characteristics of the semiconductor memory as a storage element.

4.1 Cross-call function

To maintain compatibility with existing magnetic disk drives, the shared exclusive management of the host access of the SSD must also have the equivalent functions as that of the real magnetic disk drive. However, for an SSD capable of high-speed data transfer without mechanical operation, the time for shared exclusive management occupies a large part of the total I/O processing time.

Therefore, a special bus is assigned for each internal bus group. Using this special bus, the table access for shared exclusive management is completely separated from the data transfer by normal read and write operations.

These special buses minimize the reduction in performance due to increased overhead for exclusive management processing.

The F6631 also supports the extended reserve functions.

In a configuration where the SSD is shared by systems, the extended reserve function manages the channel routes coming from one host CPU as a path group. If the RESERVE command for occupying the logical drive is issued from one of the channel routes in a path group, this function manages the pertinent logical drive to be reserved from that path group. The host system having the pertinent path group can access the logical drive from any path in the group.

This management not only increases the efficiency of using paths, but also enables access from other paths if one path becomes abnormal during reservation.

4.2 Multi-access function

The multi-access function is used to increase throughput using the characteristics of a random access semiconductor memory as the memory media.

For the real magnetic disk drive, only one input-output request from the host system can be processed at one time on one physical drive. For the F6631, if the cylinder address to be accessed is different, up to eight input-output requests can be processed simultaneously for one logical drive. This function allows a maximum of eight times the throughput over that of the real F6425 disk drive unit.

If a new I/O request is issued to the cylinder processing another I/O request, the second request waits for processing with a command retry procedure.

When one host system has reserved the pertinent logical drive, the multi-access function replies with "device busy" to an I/O request from another host system to make the request wait at that host system.

If several DIRs simultaneously access different logical drives, up to eight DIRs can simultaneously transfer data without "device busy" being issued.

5. Nonvolatile data facility

As the SSD uses a semiconductor memory as the storage media, it is necessary to have a mechanism for retaining data when the subsystem power is off. This mechanism is called a nonvolatile data facility. It is also an essential function of the SSD to maintain compatibility with the F6425 real disk drive in operation.

The parallel processing disk system with a redundant disk (parity disk), which uses a micro

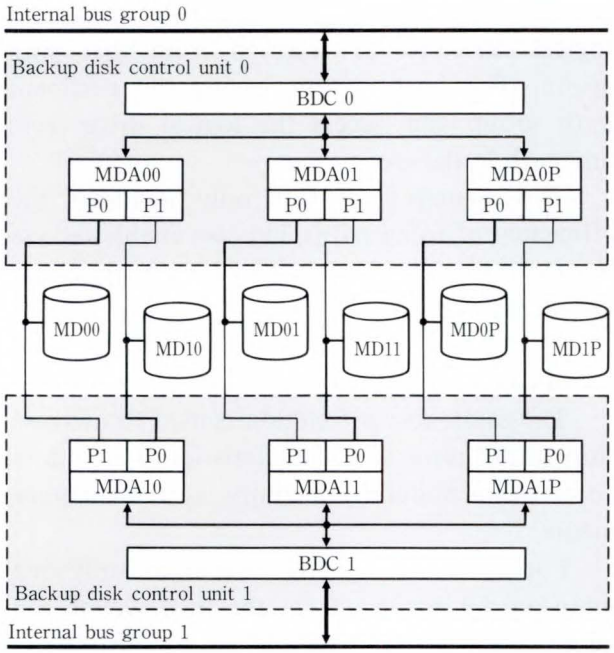


Fig. 6—Non-volatile data facility.

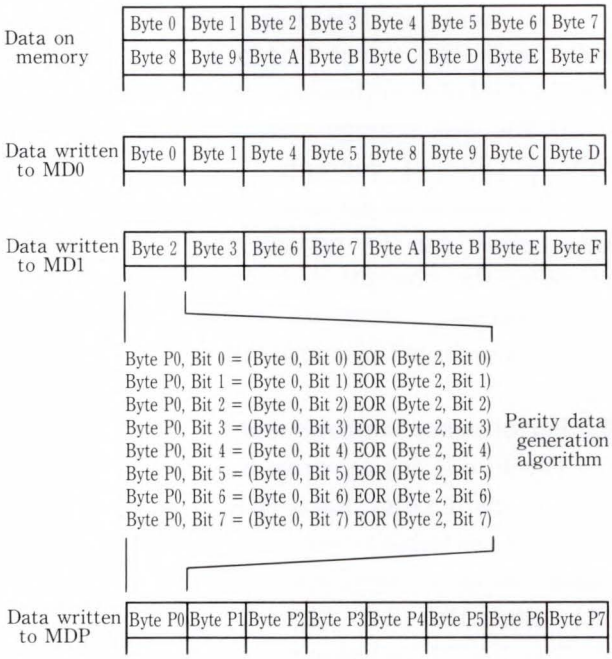


Fig. 7—Generation of parity data.

disk drive with an unformatted storage capacity of 380 Mbytes, is used in the nonvolatile data facility of the F6631.

5.1 Operation of the non-volatile data facility

When the unit power is off, the user data is saved in the nonvolatile data facility. When the power supply of the subsystem is turned on, the micro code of each component of the subsystem is loaded, then data is restored from the non-volatile data facility into the memory.

The user data is stored in several micro disk drives, including the redundant disk, described later. Even if one micro disk drive fails during restoration, the original data can be regenerated from the remaining micro disk drives.

When the operation terminates and turning off the subsystem power is indicated, the data in the memory is saved in the nonvolatile data facility first. Then the subsystem is actually turned off.

5.2 Parallel processing disk system with redundant disk

5.2.1 Hardware configuration

A detailed diagram of the non-volatile data facility in Fig. 1 is shown in Fig. 6. A simplified layout of the data in the micro disk drives is given in Fig. 7.

Micro disk drives MD00, MD01, and MD0P are installed in the basic configuration and micro disk drives MD10, MD11, MD1P are added when the storage capacity exceeds 512 Mbytes.

The configuration consisting of one backup disk controller (BDC) and three micro disk adapters (MDA) is called the backup disk control unit. Two backup disk control units are included in one SSD. Each backup disk control unit is connected to the internal control bus to constitute a device cross-call configuration for micro disk drive groups.

MDAs are disk controllers corresponding to each micro disk drives. BDC has a parallel data transfer function and a function to control the redundant disks for three MDAs/micro disk drives. The contents of the 512-Mbyte memory are stored alternately in the two micro disk drives, MD00 and MD01, from the head of the memory in two-byte units. Then a parity information of two-byte pair is automatically generated and stored in MD0P.

Usually the MD0X micro disk drive group saves or restores data through the first backup

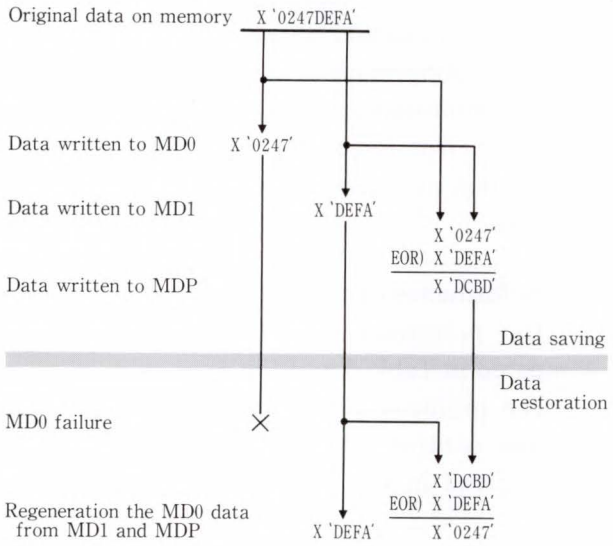


Fig. 8—Example of data recovery.

disk control unit. The MD1X micro disk group saves or restores data through the second backup disk control unit. However, if a failure occurs in one of the backup disk control units, the micro disk drive in the group saves or restores the data via another backup control unit.

5.2.2 Data regeneration at micro disk drive failure

Figure 8 shows the data regeneration procedure at micro disk drive failure.

Supposing that X'0247DEFA' is written in the memory, X'0247' is stored in MD0, and X'DEFA' is stored in MD1. In addition, X'DCBD', the exclusive OR of the two, is stored in MDP.

If data is lost due to MD0 failure at data restoration, the BDC regenerates the data in MD0 by exclusive ORing data in MDP and MD1.

5.3 Battery backup unit

The battery unit is an option for securing the non-volatile data facility in the event of a power failure. A shield type compact lead-acid battery is used for the F6631.

If a power failure is detected during subsystem operation, the battery unit is automatically activated and waits one minute for recovery from the power failure. During this period, the memory device and all the subsystem components are kept in the active power mode.

If the power failure is recovered within one minute, the system switches from the battery backup mode to the normal power mode and continues operation.

If the power failure exceeds one minute, the non-volatile data facility starts storing memory data in the micro disk drives.

After saving the data, the non-volatile data facility turns off the subsystem power and disconnects the battery unit.

If the power failure is recovered while data is being saved, the nonvolatile data facility continues subsystem operation by making remaining data in the memory valid, after saving the data in the micro disk drives.

6. RAS function

Because of its characteristics, essential files for system operation are mostly stored in the SSD. Therefore, high reliability and flexible maintenance is required.

6.1 High reliability design in the hardware configuration

To ensure high reliability, the following is considered essential to prevent the subsystem becoming unavailable due to a single failure.

6.1.1 Duplex access paths and power supplies

The subsystem internal bus is separated into two groups. The data transfer paths to the memory device and the micro disk drive are fully duplicated. Furthermore, functional components are also duplicated and arranged for each bus group.

6.1.2 Duplicated power supply

The power supply to the memory device and micro disk drive is fully duplicated to cope with failures in the power supply unit.

Each power supply provides power to each bus group. The bus groups can turn on or off independently. This system enables semi-active maintenance of the subsystem.

6.1.3 Data integrity

The memory device has an error correction code (ECC), which automatically corrects the errors of single bit memory error and detects errors exceeding two bits of memory.

Memory soft errors are automatically rewritten by the patrol function described below.

6.2 Self-diagnostic function (patrol function)

The F6631 continuously diagnoses the resources in the subsystem, even while the device is operating in the online mode. This function is called patrol.

If an error is detected during patrol, this function terminates the I/O request from the host system with a temporary equipment error, edits the detailed information of the error in its sense byte, and posts it to the host system to prompt a response from the field engineer.

The interval of patrol error report is variable. However, it is usually reported at ten-minute intervals.

6.2.1 Patrol and error report method in the memory device

Memory is patrolled at 455 μs intervals in 32-byte units in parallel with the access from the host system.

Because four memory access controllers (MACs) patrol different areas, the patrol interval per 1 Gbyte of memory is about one hour.

If a 1-bit error is detected in memory, MAC restores the correct data obtained by ECC and reads it again to determine whether it was a hard error or a soft error. If the memory error is not recovered by correcting the data, MAC determine it to be a hard error and reports the patrol error to the host system.

6.2.2 Patrol of micro disk drives

For the micro disk drive, the read check for the full area and the write function check using the diagnostic area are performed.

If a sector that cannot be read is detected by the read check, an alternative sector is automatically assigned.

The micro disk drives are patrolled at cycles of about 30 min.

6.2.3 Patrol of power supply unit and battery unit

The status of the duplicated power supplies and fans is always monitored.

For the battery unit, the output voltage is measured by connecting a dummy load for several seconds to the battery unit once every two hours. It residual capacity is measured.

If an abnormality is detected, the patrol error is immediately reported to the host system. The error is continuously reported to the host system at a certain cycle until the problem is fixed.

7. Performance considerations

The performance of F6630 and F6631 is compared in Table 3 and Fig. 9.

The primitive performance of the SSD and the disk control unit (F1774C with disk cache) is compared in Fig. 9. This chart compares the

Table 3. Comparison of overall performance

Item	F6631	F6630
Single path configuration	636	549
2-path configuration	1 260	1 088
4-path configuration	2 479	2 137
8-path configuration	4 549	Not applicable

This table shows the number of random read operations of 4 096-byte data for one second.
For both devices, the number of logical drives is eight. The CPU continuously responds to each logical drive I/O processing request.

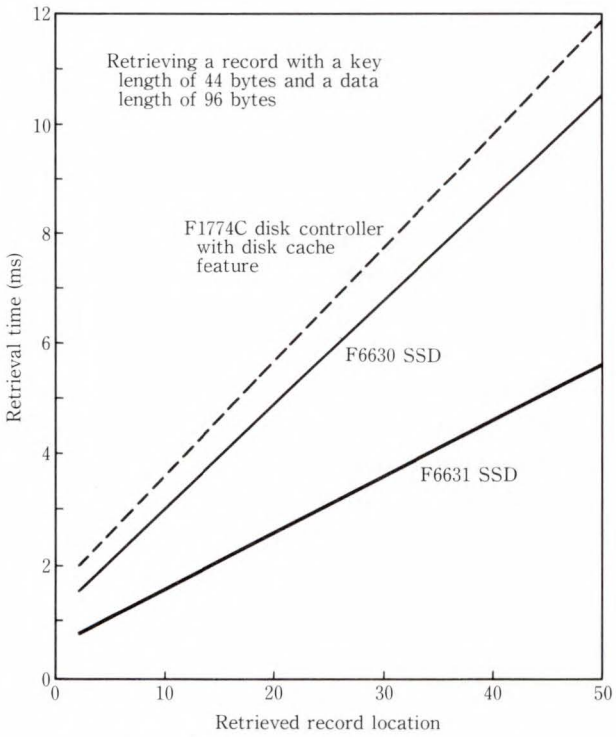


Fig. 9—Comparison of primitive performance.

time for accessing the object record when the record is searched sequentially under a single logical drive configuration.

Table 3 shows an example of comparing the total performance of the SSD.

The number of connecting paths and the total number of processes done by the unit is compared by performing random read operations of 4 096-byte data.

The eight-path configuration enables about seven times as many operations as a single-path configuration.

8. Conclusion

The characteristics of the current SSD are very fast, but expensive, and small in capacity.

Therefore, the SSD is used mainly for high performance that requires high-speed data processing rather than low cost. The market has been limited to large systems such as Fujitsu M-780.

On the other hand, even in medium-large systems whose performance have rapidly improved, the amount of data that must be processed at high speed is rapidly increasing. The need for a SSD having a good balance between cost and performance is rapidly increasing.

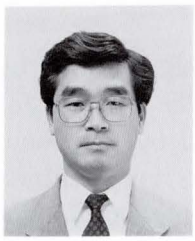
The need for a SSD having a good balance between cost and performance is rapidly increasing.

To meet the market requirements, the difference in storage cost between an SSD and a real magnetic disk drive unit must be reduced.

The task is to make systems at reduced cost by achieving a breakthrough in memory technology.

References

1) Inoue, Y., Ogawa, Y., and Kaneko, S.: External Storage Devices. (in Japanese), *FUJITSU*, **38**, 5 (Special Issue: Information Processing Devices), pp. 347-362 (1987).
2) Okamoto, K.: A Long Distance Interface between Mainframe and its peripherals. (in Japanese), Proc. 38th Natl. Conv., Inf. Proc. Soc., Jpn. 1989, p. 1882.



Hajime Sugiura
Planning Dept.
Research and Planning Div.
FUJITSU LIMITED
FUJITSU Technical College 1973
Specializing in Magnetic Storage Systems

Soichiro Nagasawa
Engineering Dept.
File System Div.
FUJITSU LIMITED
Bachelor of Science
Waseda University 1979
Master of Science
Waseda University 1981
Specializing in Microcode Design of Magnetic Disk Controller

Etsuo Morita
Engineering Dept.
File System Div.
FUJITSU LIMITED
Bachelor of Electronic Eng.
Nagoya University 1965
Master of Electronic Eng.
Nagoya University 1967
Specializing in Magnetic Storage Systems

UDC 681.327.634

Fujitsu Small Magnetic Disk Drives

• Ikuo Kitamura • Norihito Aramaki • Takeo Masuda

(Manuscript received June 19, 1990)

The progress in the area of small magnetic disk drive is very rapid. This paper outlines the latest Fujitsu small disk drives; M2671P (8-inch, 2 600 Mbytes), M2652P/S (5-inch, 2 000 Mbytes), M262*X*S (3.5-inch, 520 Mbytes). It describes the general specification, interfaces and features of Fujitsu's disk drives, beginning with the history of the drives.
This paper covers magnetic heads, recording media, read/write and servo technologies to achieve higher capacity and speed, with the first priority on reliability. Experimental data are presented on these technologies.

1. Introduction

Magnetic disk drives first came into wide use in the second half of the 1960s. The first disk drives were for disks 14 inches in diameter and approached the size of an office desk, yet had a storage capacity of no more than 10 Mbytes. As magnetic disk technology steadily advanced, however, magnetic disks soon became the main external memory for large-scale computers.

Soon, smaller media came into use. First were 10.5-inch disks, followed by 8-inch disks. The wave of office automation in the 1980s made the 5.25-inch disk very popular. Storage capacity was greatly expanded and soon 3.5-inch media appeared. Today, we are witnessing the development of disk drives the size of a cigarette carton (2.5 inches) with a capacity of as much as 100 Mbytes.

Behind this progress has been the steady development of head and media technology, increases in recording density through ultraprecise head positioning, and the advancement of VLSI devices and other semiconductor technology.

2. Improvements in Fujitsu disk drives

Improvements in Fujitsu disk drives are shown in Fig 1. In 1981, when the industry standard disk was 14 inches, Fujitsu introduced drives for 10.5-inch high performance disks[1]. That series came to be used as main external files for the company's large-scale computers.

Shortly afterwards, drives for 8-inch disks appeared. As first, these drives looked like toys, but they incorporated the technology of the larger disk drives and quickly came into general use. While maintaining the original form factor, 8-inch drives were to attain a storage capacity of up to 2.6 Gbytes.

The 8-inch drives were quickly followed by 5.25-inch and 3.5-inch drives. The demand for such drives was greatly affected by the popularity of personal computers. As Fig. 1 shows, small disk drives underwent technological innovations much more rapidly than did large disk drives. The need for compact, economical disks promoted extensive disk drive downsizing.

This article introduces three of Fujitsu's latest drives, and outlines their technical features. These drives are called the Swallow 7, the Super Humming, and the Picobird 4[note].

3. Description of Fujitsu disk drives

Figure 2 shows the three drives: the Swallow 7, Super Humming, and Picobird 4. Table 1 lists their overall specifications. The sections which follow describe each drive.

Note: Swallow, Super Humming, and Picobird are the code names given to Fujitsu's 8-, 5.25-, and 3.5-inch disk drives.

3.1 Swallow 7 (M2671P)

The Swallow 7 is an 8-inch high-performance disk drive developed for use with mainframes and high-performance workstation file servers. Fujitsu is at the point of adapting it to serve as a disk array for supercomputers[2].

1) Standard 8-inch drive with a storage capacity of 2.6 Gbytes (unformatted)
2) Minimum 2-ms, average 12-ms high-speed seek
3) Short latency time of 6.9 ms (4 340 rpm) with low power consumption
4) High-speed data transfer of 4.78 Mbyte/s
5) IPI-2 interface conforming to ANSI standards

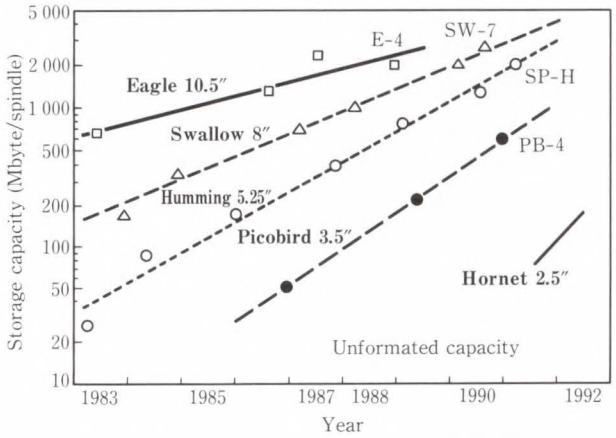


Fig. 1—Development of Fujitsu disk drives.

3.2 Super Humming (M2652P/S)

The Super Humming was developed for use with mainframes, supercomputers, workstations, and high-performance personal computers.

While maintaining the standard 5.25-inch disk form factors and power requirements, the Super Humming features a large capacity and

Table 1. General specification

Specification	Swallow 7 M2671P	Super Humming M2652P/S	Picobird 4 M262*X*S
Disk size (inch)	8	5.25	3.5
Storage capacity (Mbytes)			
Unformatted	2 648	2 000	–
Formatted	–	1 600	520
Number of platters	9	12	6
Number of heads	15 + 1	20 + 1	11 + 1
Positioning time (ms)			
Minimum	2	2	3
Average	12	12	12
Maximum	23	22	25
Rotational speed (rpm)	4 340	5 400	4 400
Latency (ms)	6.91	5.56	6.82
Data transfer rate (Mbyte/s)			
Disk drive	4.78	4.758	2.4-3.0
SCSI async mode	–	–	3 max
sync mode	–	10 max	5 max
Interface	IPI-2	IPI-2/SCSI	SCSI

Note: Formatted capacity is for 512 byte/block.

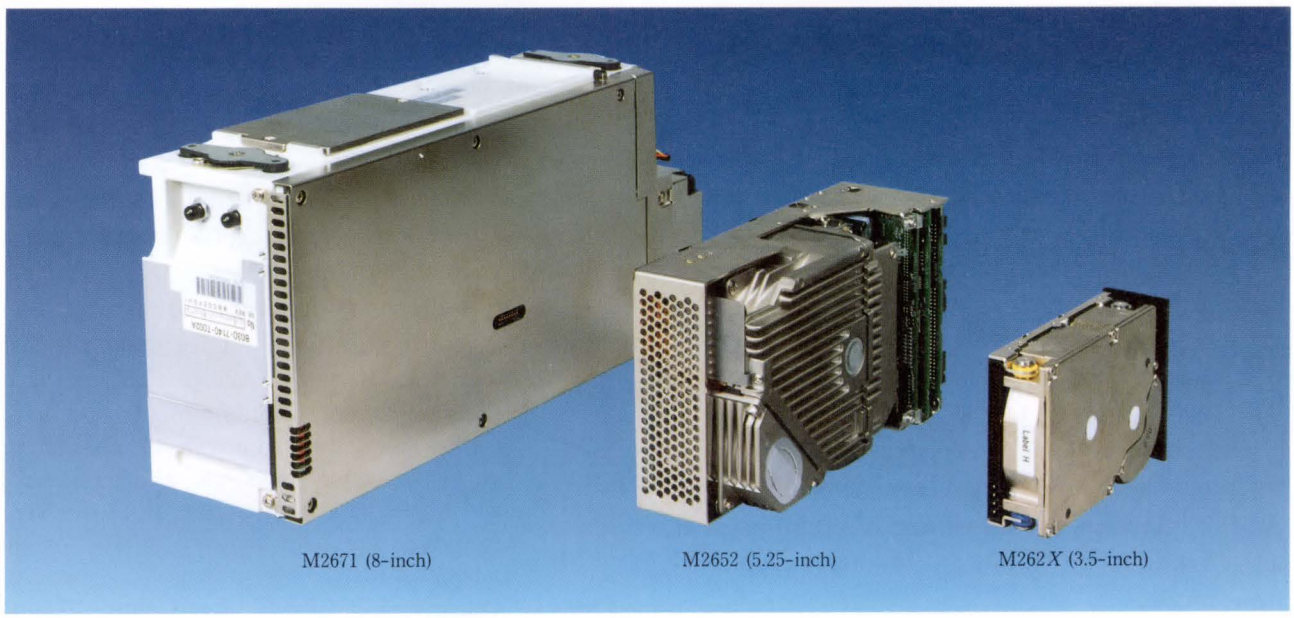


Fig. 2—Small disk drives.

fast data access. It supports two types of interfaces, IPI-2 (Intelligent Peripheral Interface-2) and SCSI (Small Computer System Interface).

1) 5.25-inch form factor with a capacity of 2.0 Gbytes (unformatted)
2) Minimum 2-ms, average 12-ms high-speed seek
3) Latency time reduced to 5.65 ms (5 400 rpm)
4) Fast data transfer rate
   i) Maximum 10 Mbyte/s on an SCSI data bus (synchronous mode)
   ii) 4.75 Mbyte/s (head to media)
5) IPI-2 interface conforming to ANSI standards
   A dual port is installed as standard.
6) SCSI interface conforming to ANSI standards
   Advance installation of SCSI-2 (ANSI X3T9/89-042) standard specifications[7].
7) Standard installation of spindle synchronization functions.
   Models vary with the number and type of port used:
   i) Single port, single-ended
   ii) Single port, differential
   iii) Dual port, differential.

3.3 Picobird 4 (M262*X*S)

Though the disks are only 3.5 inches, the Picobird 4 disk drive has a 520-Mbyte capacity. This high-performance drive is more than adequate for use with office computer systems and workstations.

1) Format 520-Mbyte capacity
   The drive is equipped with six disks and uses constant-density recording[note 1].
2) High-speed access
   Minimum 3-ms, average 12-ms high-speed seek, featuring a new acutuator motor and digital signal processor (DSP).
3) Latency time is only 6.82 ms (4 400 rpm).
4) High-speed data transfer
   When reproducing recorded data, the transfer rate is between 2.4 Mbyte/s and 3.0 Mbyte/s. At the interface, the maximum is 3.0 Mbyte/s in asynchronous mode and 5.0 Mbytes in synchronous mode.

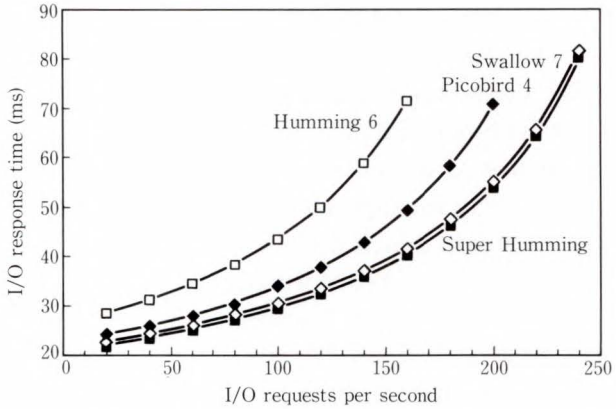


Fig. 3—Performance simulation.

3.4 Performance comparison

Figure 3 is a performance simulation of the Swallow 7, Super Humming and Picobird 4, together with that of the Humming 6 as a reference. The simulation shows the extent to which response time is increased as the frequency of I/0 requests increases. The simulation clearly shows that the Swallow 7 and Super Humming have a greatly superior I/O request processing capability. Compared with the 5-inch Humming 6[note 2], the Picobird 4 also has superior capability.

4. Interface

The small computer system interface (SCSI) has, as the name suggests, become popular for use with small computers. Its greatest assets are low cost and easy handling. Up to now, an SCSI has not been adequate for large, high-performance systems. However, improvements have been made since the SCSI-2 was proposed. In anticipation of the SCSI-2, Fujitsu is to provide a high-performance system, rather than just a high-performance drive.

The intelligent peripheral interface 2 (IPI-2), a device-specific command set for magnetic disk

Note 1) Constant-density recording means data is recorded at a density approximately constant on all tracks. Outer tracks have a larger capacity and higher data transfer rate. In conventional recording, the track capacity is based on the innermost track.

Note 2) Capacity 1.26 Gbytes, access time 14.5 ms, 3 600 rpm, data transfer speed 3 Mbyte/s.

drives, has numerous advantages. A small command overhead and few limitations on transfer speed help the drive achieve full performance. The IPI-2 is also compatible with most large host interfaces.

Fujitsu has given IPI-2 the position of a major interface for large, high-performance computer systems.

4.1 SCSI Interface

Table 2 lists commands supported by the Super Humming and Picobird 4. The following features are incorporated in the interface to upgrade system performance and reliability and to minimize the system's load.

4.1.1 Upgraded performance

1) 256-Kbyte data buffer

A 256-Kbyte buffer makes possible a system configuration that uses high-speed data transfer capabilities efficiently.

2) Read-ahead cache

When data reading is completed, the read-ahead cache automatically reads the next data block and writes it to a data buffer. If the next command requests that data, the data can be transferred directly from the buffer.

3) Defective block skipping

During formatting, a logical data block can be reallocated in a physical sequence by skipping a defective block. This makes possible high-speed data processing without generating additional latency because of the defective block.

4) SCSI-2 logical specification supported

Several SCSI-2 commands are supported such as READ/WRITE LONG and WRITE SAME as well as extended mode parameters, inquiry data and sense information. This provides enhanced system tunability and diagnostic capability.

5) Reduced overhead

SCSI overhead has been improved to less than 500 μs by a fine-tuned micro code and enhanced protocol chip.

4.1.2 Upgraded system reliability

1) 64-bit error correction code (ECC)

2) Parity protection

Parity protection is used on the SCSI interface data bus, the internal data bus, and the control bus.

4.1.3 Reduced system load

1) Continuous block processing

Table 2. List of SCSI commands

○: Supported

SCSI command	Super Humming	Picobird 4
a) Control/sense commands		
TEST UNIT READY	○	○
INQUIRY	○	○
READ CAPACITY	○	○
MODE SELECT	○	○
MODE SELECT EXTENDED	○	○
MODE SENSE	○	○
MODE SENSE EXTENDED	○	—
NO OPERATION	○	○
REZERO UNIT	○	○
START/STOP UNIT	○	○
RESERVE UNIT	○	○
RELEASE UNIT	○	○
PRIORITY RESERVE	○	○
REQUEST SENSE	○	○
WRITE RAM	○	—
READ RAM	○	—
LOG SELECT	○	—
LOG SENSE	○	—
READ DEVICE CHARACTERISTICS	○	—
b) Data access commands		
READ	○	○
READ EXTENDED	○	○
WRITE	○	○
WRITE EXTENDED	○	○
WRITE AND VERIFY	○	○
VERIFY	○	○
SEEK	○	○
SEEK EXTENDED	○	○
SET LIMITS	○	○
c) Format commands		
FORMAT	○	○
REASSIGN BLOCKS	○	○
READ DEFECT DATA	○	○
d) Test/diagnostic commands		
SEND DIAGNOSTIC	○	○
RECEIVE DIAGNOSTIC RESULTS	○	○
WRITE BUFFER	○	○
READ BUFFER	○	○
READ LONG	○	○
WRITE LONG	○	○
WRITE SAME	○	○
SPACE ID AND READ DATA	○	—
RECOVER ID	○	○
RECOVER DATA	○	○

Blocks are assigned logical block addresses so that the host can access data regardless of track and cylinder boundaries.

2) Automatic alternate block reassignment

When a defective data block is detected during reading or writing, an alternate block is assigned automatically.

3) Error recovery

The system attempts recovery from all types of errors. Recoverable read errors are corrected in the data buffer so that only error-free data is transferred. This frees the initiator (host) from all error recovery processing.

4.1.4 Miscellaneous features

1) Programmable block length

The most appropriate block length can be specified in 2-byte increments.

2) Multihost support

The SCSI has multi-initiator support in addition to reserve/release and command stack functions.

4.2 IPI-2 interface

The Swallow 7 and Super Humming IPI-2 have the following features:

4.2.1 Upgraded performance

1) Processing time has been shortened by reducing the overhead for command proccessing to 10-20 μs.
2) High-speed data transmission of 10 Mbyte/s is possible using a 16-bit data bus.
3) A dual port is supplied as standard, improving subsystem throughput.

4.2.2 Upgraded system reliability

1) An internal bus parity check improves system reliability.

4.2.3 Reduced system load

1) Field commands for data bus control are supported, enabling flexible formatting.
2) Optional parameters and short RPS for load RPS/position commands are supported to optimize system processing.

4.2.4 Miscellaneous features

1) The interface supports up to 130 Kbytes per track, and 1 024 sectors per track.
2) A D-connector to prevent EMC is available as an option.
3) Spindle synchronization is supported.

5. High reliability and head-disk interface (HDI)

Fujitsu ensures that a high degree of reliability is built into its products. The Swallow 7, Super Humming, and Picobird 4 all feature an MTBF[note)] rating of 200 000 h, a marked increase over previous models. The excellent MTBF rating is a result of Fujitsu's long experience and know-how.

5.1 Design considerations

The following considerations guide its design.

1) The drive was designed logically to enable prior detection and correction of errors by adding a parity check and a powerful ECC.
2) VLSI devices and microprocessors were incorporated to reduce the number of components and to achieve lower power requirements.
3) A new equalizer in the read channel was incorporated to improve the read error rate.

5.2 HDI reliability

Head crashes have always been the biggest problem with magnetic disk drives. Normally, the head flies above the disk, without touching it. A head crash occurs if the head touches the disk and damages it. The result is a loss of data on the disk.

To achieve higher recording densities, the head must be as close to the media as possible. Designers must continue to consider reliability as they attempt to bring the head closer to the media.

Fujitsu's approach to this problem was to introduce the concept of an HDI. The HDI is described in detail in a separate article[3]. Fujitsu has taken the following measures to ensure the reliability of its drives.

5.2.1 Head and media production policies

To achieve higher magnetic densities and expand disk drive capacity, Fujitsu has strategies for developing both magnetic heads and

Note: Mean Time Between Failures. The MTBF is calculated to be the total operating time per total number of disk drive failures. The MTBF must not be misconstrued as the life of the drive.

recording media[4),5)] internally. Table 3 lists selected combinations of heads and media for each drive best suited for performance and cost effectiveness.

1) Metal sputtered media

For metal sputtered media, substrate texturing, a protective layer, and lubrication are the keys to HDI. In production, Fujitsu maintains strict control and supervision over these key items.

2) Magnetic head flexure and slider design

Fujitsu design incorporates a magnetic head support and a slider that withstands fluctuations in the azimuth-angled flying height. These features are best suited for achieving high speed seek and low flying height. Figure 4 shows the head flexure[6)]. It also shows fluctuations in the flying height due to acceleration and deceleration during seek operations and how fluctuation is reduced.

The shape of the heads used in the Picobird 4 and Super Humming was specifically designed to reduce the distance between disks and to in-

Table 3. Heads and media outlined

Specification	Swallow 7	Super Humming	Picobird 4
Write/read head	Thin film	Thin film	MIG (Metal In Gap)
Recording media	Metal sputter	Metal sputter	Metal sputter
Flying height (μm)	0.16	0.13	0.13
Bit density (bpi)	33 310	50 257	46 383
Track density (tpi)	2 098	1 780	1 751

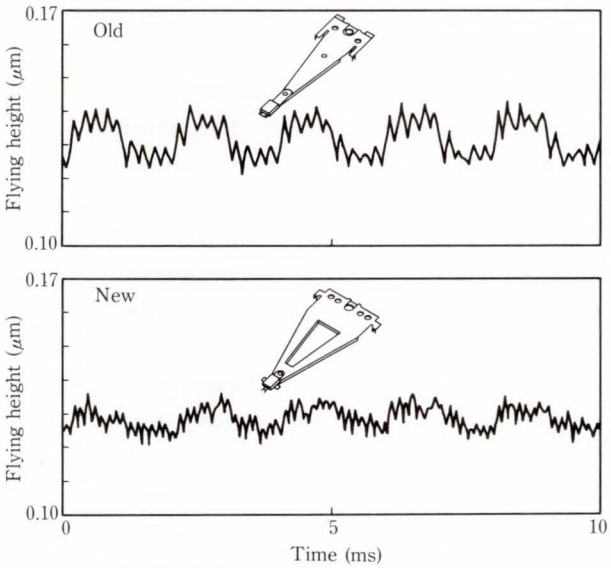


Fig. 4—Old and new flexures and their flying height fluctuation.

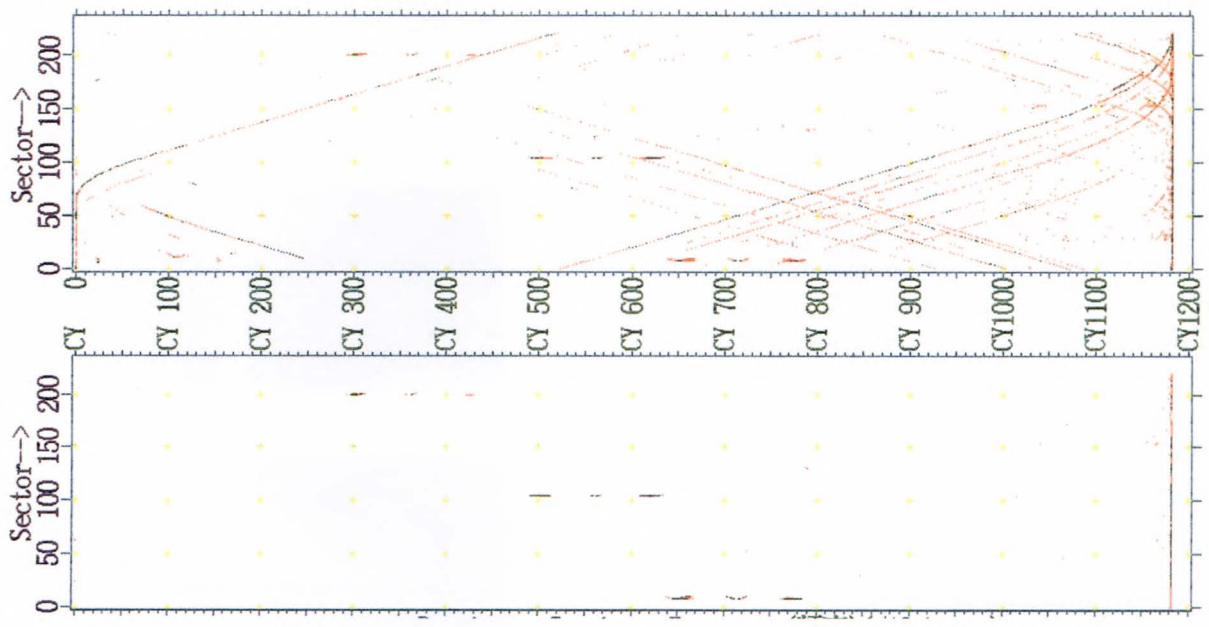


Fig. 5—Example of head touch mapping at early developmental stage.
(The torus of the disk is displayed as a rectangular with cylinder and sector numbers. The upper chart plots the degree of impact by colors, the lower chart the frequency both for 10 h running. Parabolic loci are head touches due to vibration during acceleration and deceleration. Isolated islands are fixed protuberances.)

crease the number of disks. These contributed much to increasing drive capacity.

5.2.2 Complete monitoring procedure

1) Test equipment

Fujitsu plants employ high-quality test equipment on a par with first-rate research laboratories.

This equipment is used to tackle the problems of prototype development and field failure analysis. It is also used for continuous control of the thickness of the lubricant coating, which should be just over several nanometers.

2) High-sensitivity piezoelectric sensor[3)]

Fujitsu has developed an ultra-small piezo electric sensor installed on the head. The sensor monitors head touches with a high degree of sensitivity and is used in the production line and development. Figure 5 maps head touches. The torus of the disk is opened and depicted as a rectangle. The parabolic trajectories give evidence of head touches caused by vibrations while the disk is undergoing high acceleration or deceleration. Isolated islands represent fixed protusions. Of course, corrective measures for these were taken early in the development and no such defective equipment was ever shipped.

3) Outgassing analysis[3)]

Fujitsu has set up ways to collect and analyze the tiny amounts of organic gas occurring inside the disk enclosure. Rubber packing materials that do not produce outgassing are carefully selected and the production processing is painstakingly monitored. The generation of even the tiniest amount of organic gas within the disk drive can cause dust to accumulate on the head and result in a head crash. Therefore, these precautions were absolutely indispensable.

5.3 Inspection before shipment and production-line control

5.3.1 MTBF verification

Before the first shipment, several hundred units were tested for several months to verify the MTBF rating and to correct any problems.

5.3.2 On-going reliability test (ORT)

Each day, a number of drives were selected from the production line and subjected to continuous running tests. This monitoring is followed by disassembling the units for quality analysis.

5.3.3 Improvement of clean rooms

Fujitsu clean rooms are equipped with high resolution dust counters to constantly monitor cleanliness. In addition, personnel keep work areas immaculate. They also have meetings to propose and implement new measures to ensure cleanliness.

6. Other key technologies

6.1 Read/write channel

To improve the read/write signal-to-noise (S/N) ratio caused by higher tpi/bpi, and to correct the waveform, which is unique to thin-film heads, Fujitsu has incorporated a 5-tapped equalizer channel[7)]. With thin-film heads, signals known as negative-edge noises are generated before and after the main signals in the isolated waveform of the head. These negative-edge noises distort the main signal.

Fujitsu's equalizer generates signals that have the reverse polarity of the negative-edge noises

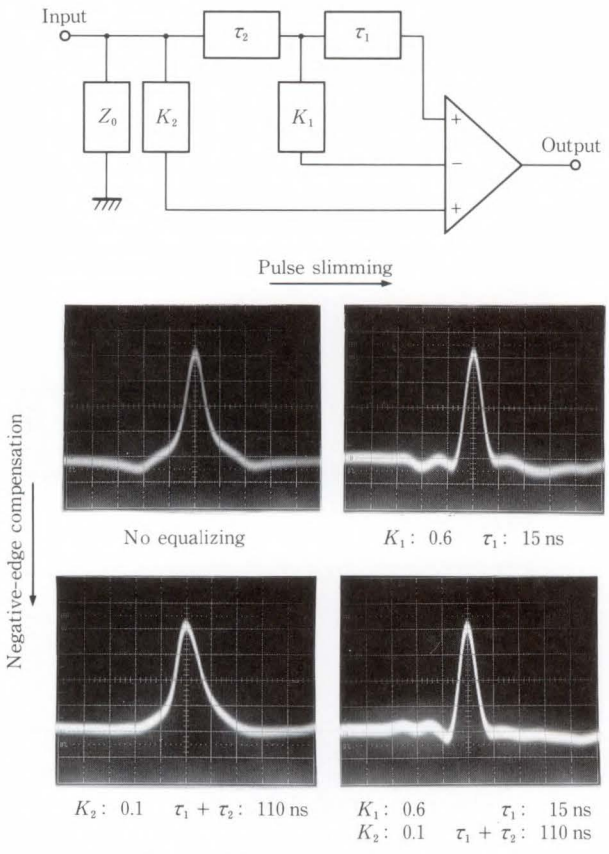
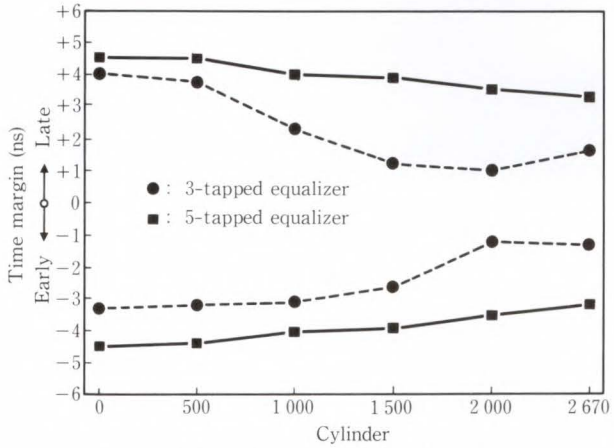


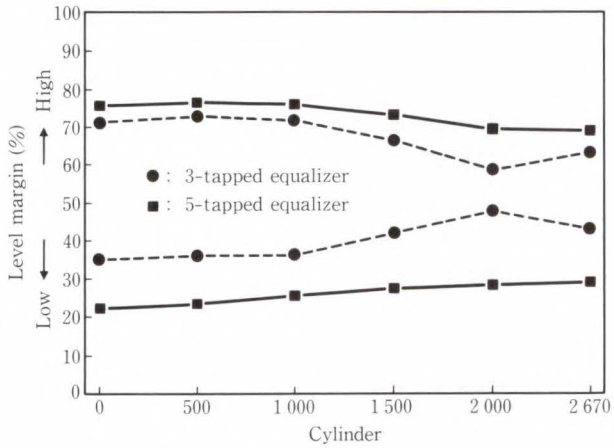
Fig. 6—Five tapped equalizers.

to cancel them out. This reduces the peak shift and eliminates distortion.

Distortion and peak shift phenomena vary with the head position, so the level of the equalizer's corrective signal is also varied with the head position. This ensures a highly dependable read circuit. Figure 6 shows the equalizer circuits and waveforms. Figure 7 shows improvements in read-time margins and read-level margins.



a) Time margin vs cylinder



b) Level margin vs cylinder

Fig. 7—Margin improvement by 5-tapped equalizer.

6.2 Servo control circuit

The drives introduced in this article use digital servo control by means of a DSP to upgrade performance and reliability. The control circuit configuration is shown in Fig. 8. The special features of this circuit are described below.

1) DSP control[8)]

Compared with analog control, digital control using the processor enables more delicate parameter control. For example, the processor can set a current wave pattern during drive acceleration, which includes only small high-frequency harmonic components, and can suppress vibration and reduce acoustic noise. Figure 9 shows reduced settling after a seek operation, compared with the conventional

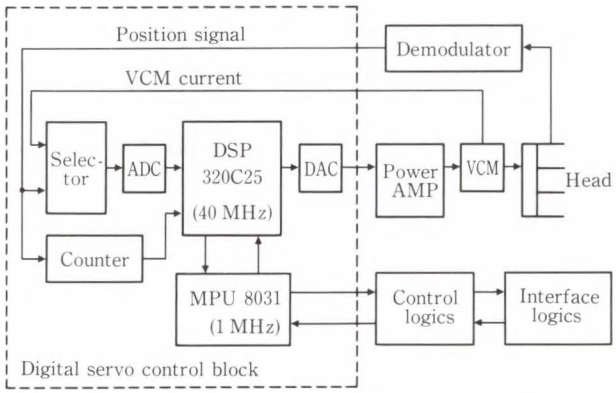
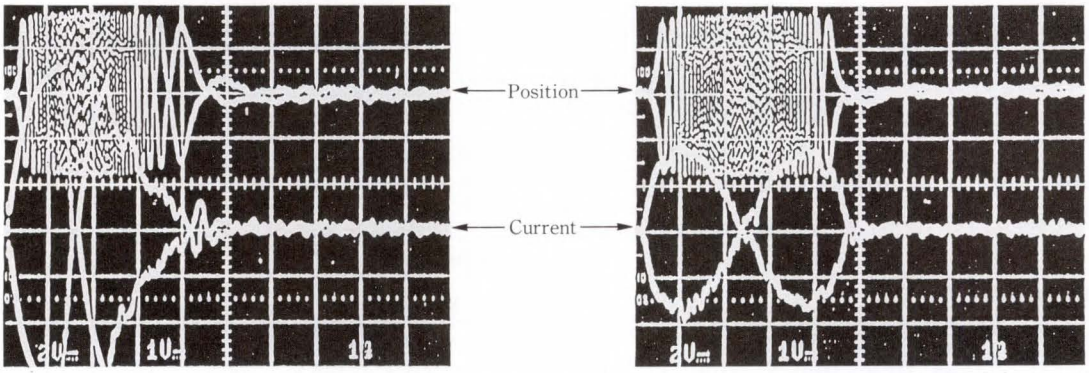


Fig. 8—Digital servo schema for head positioning.



a) Conventional control (slow settling)

b) Optimum current control by DSP (fast settling)

Fig. 9—DSP controlled seek.

method, which is no longer appropriate for faster access times.

The processor also monitors changes in power voltage, mechanical force, inconsistent servo data, and setup conditions. Monitoring enables control values to be corrected in response to changes in voltage and ambient temperature. This upgrades the reliability of seek operations.

2) Reduced number of components

Conventional servo circuits consisted of operational amplifiers and analog switches. Digitization has enabled the number of components to be greatly reduced and has reduced power consumption. This increases reliability.

6.3 Actuator

The actuator holds the heads and moves them to target tracks at high speed. Actuator operation has been improved in several ways.

1) The actuator motor has been downsized and power efficiency improved.
2) The degree of overall intertia and rigidity have been improved.
3) Thermal distortion, the greatest cause of off-tracking, has been reduced.

In new disk enclosures, the actuator consists of a one-piece structure of aluminum or magnesium.

This new structure has produced the following results.

1) Inertia has been reduced by 80 percent compared to previous models (70 percent in weight), giving the Swallow 7 an access time of 12 ms.
2) Irreversible distortion does not occur.
3) Reversible distortion has been analyzed by computer simulation to determine the best actuator shape. This has helped reduce the thermal off-track to only one third that of previous models. The actuator is now capable of high-density recording at 2 000 tpi.

6.4 Actuator motor

The actuator motor has been designed for the shortest access time under conditions of limited space and lower power consumption.

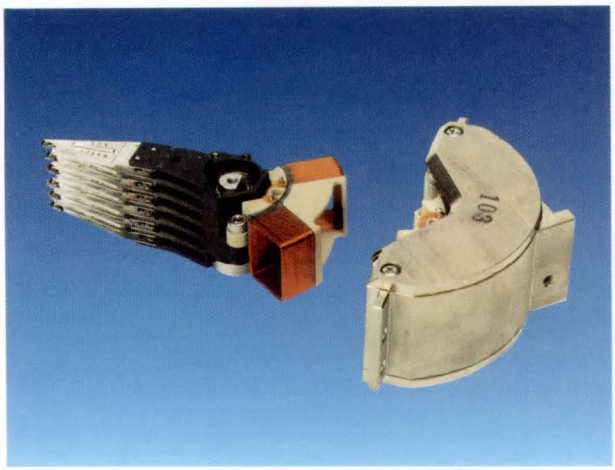


Fig. 10—Actuator in Picobird 4.

Figure 10 shows the actuator motor used in the Picobird 4. Neodymium iron magnets are placed around two coils to create a high magnetic density and a coil efficiency of 65 percent. (Coil efficiency refers to the effective length of power generation along the entire length of the coil. Normal coil efficiency is 40-45 percent.)

6.5 Spindle motor

A DC-Hall motor is used for the spindle. With a microprocessor to control the rectification timing, an adequate torque can be generated at both low and normal temperatures. Power consumption is reduced by pulse width modulation (PWM) control using MOSFETs. Current fluctuations and noise are minimized by reversing the regenerated current through diodes.

7. Conclusion

Magnetic disk drives will continue to evolve toward greater recording densities and compactness, maintaining and extending their popularity as convenient, efficient storage devices.

Fujitsu will continue to work to develop leading-edge technology to create innovative and attractive products.

References

1) Mizoshita, Y., and Matsuo, N.: MECHANICAL AND SERVO DESIGN OF A 10 INCH DISK DRIVE. *IEEE Trans. Magnetics*, **MAG-17**, 4, pp. 1387-1391 (1981).

2) Oyama, T., Ogawa, Y., and Sugiyama, K.: F6490 Magnetic Disk Subsystem: DIA. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 291-295 (1991).
3) Yamamoto, T., Takahashi, M., and Shinohara, M.: Head-Disk Interface. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 415-427 (1991).
4) Oshiki, M., and Hamasaki, S.: Thin Film Head Technology. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 353-364 (1991).
5) Ishida, S., and Seki, K.: Thin Film Disk Technology. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 337-352 (1991).
6) Yoneoka, S., Ohwe, T., and Mizoshita, Y.: Flying Head Assemblies. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 404-414 (1991).
7) Aikawa, T., Mutoh, H., and Sugawara, T.: Signal Processing for High Density Magnetic Recording. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 391-403 (1991).
8) Hasegawa, S., Takaishi, K., and Mizoshita, Y.: Digital Servo Control for Head-Positioning of Disk Drives. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 378-390 (1991).

Ikuo Kitamura
Engineering Dept.
File Device Div.
FUJITSU LIMITED
Bachelor of Electric Communication Eng.
The University of Electro-Communications 1970
Specializing in Mechanical Eng.

Norihito Aramaki
Engineering Dept.
File Device Div.
FUJITSU LIMITED
Bachelor of Information Eng.
Kyushu University 1972
Specializing in Read/Write Circuit and Interface

Takeo Masuda
Engineering Dept.
File Device Div.
FUJITSU LIMITED
Bachelor of Electronic Eng.
Yokohama National University 1973
Master of Electronic Eng.
Yokohama National University 1975
Specializing in Electronic Circuit Design

UDC 681.327.636

F1751/F6470 Magnetic Tape Subsystem

• Yoshikazu Nakamura *(Manuscript received August 7, 1990)*

This paper describes the functions and advantages of the F1751/F6470 Magnetic Tape Subsystem of which tape cartridge conforms to the industry standard.
It also outlines an 8-inch form factor M2481A/B Magnetic Tape Drive that uses the same cartridge.

1. Introduction

For more than 20 years, Fujitsu has developed and manufactured many types of tape devices for Fujitsu computer systems, OEM customers, and the PCM marketplace.

Magnetic tape units are not only for magnetic disk drive data back-up archival uses, but can also be used for data set interchange between different computer systems. Initial software installation is just one example of this interchange.

For a long time, open reel magnetic tape units were used in computer systems. Since the introduction of the IBM 3480-type Cartridge Tape, it has come to replace open reel tape and is becoming the standard tape for data interchange due to easier handling.

Although the cartridge tape is smaller than open reel, it has increased recording density and higher data transfer rates. Additionally, the tape subsystem has several new functions.

Using the same cartridge, Fujitsu developed a cartridge tape subsystem, the F6470 tape unit and the F1751 Tape Controller, and started shipment in 1987.

In October of 1989, Fujitsu introduced the low cost and compact M2481A/B Cartridge Tape Unit and maintains data interchangeability with the F6470.

The following is a brief explanation of Fujitsu's cartridge tape subsystems, which have many new functions not obtained by previous open reel tape units.

2. F1751/F6470 Magnetic Tape Subsystem

Figure 1 shows the F1751/F6470.

The general specification is also listed in Table 1. Installation space and performance are greatly improved over open reel tape subsystems.

2.1 System configuration

The maximum F1751/F6470 Magnetic Tape Subsystem configuration is shown in Fig. 2. One F1751 Tape Controller can control a maximum of four F6470 Tape Units each of which has two tape drives.

A configuration of one controller and its connected tape units is known as a "string". Since two "strings" can be connected together, the maximum configuration of the subsystem is two controllers with sixteen tape drives. In a subsystem configuration with two controllers connected to each other, even if one of the controllers has a malfunction or is disconnected for maintenance, the other controller can continue to control all sixteen tape drives for higher subsystem availability. In addition, if the load of one of the controllers increases, the other controller can balance the load in order to achieve better throughput.

2.2 Characteristics of F1751 Tape Controller

The F1751 Tape Controller consists of three functional areas; the host interface control (HIC), buffer memory (BUFF), and formatter (FMT) (see Fig. 3).

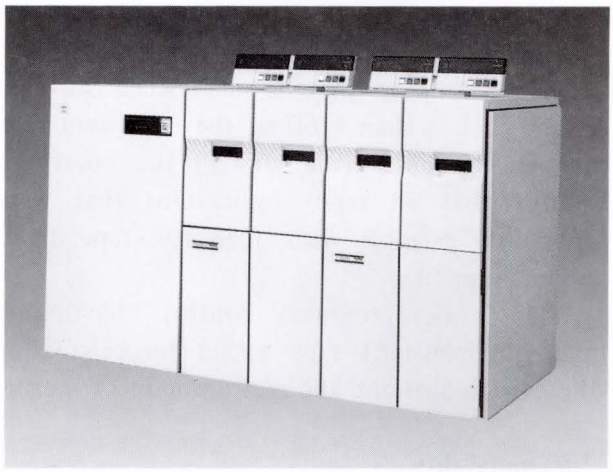


Fig. 1—F1751/F6470 Magnetic Tape Subsystem.

Table 1—General specification for F6479 Magnetic Tape Drive

Number of tracks	18
Recording scheme	GCR
Recording density	37 871 bpi
IBG length	2 mm
Data transfer rate	3 Mbyte/s
Tape speed	2 m/s
Load/unload time	7 s
Rewind time	48 s

The HIC and FMT are controlled by microprogramming techniques. Control information is exchanged through the control information table (CIT) located in the BUFF. This architecture provides separation during host channel command execution or data transfer from tape drive operations. For the read or write error recovery, the FMT and BUFF combinations can greatly reduce the host load.

The BUFF has a capacity of 4 Mbytes and is partitioned into 512 Kbytes for each tape transport. This simple buffer management algorithm minimizes the overhead in the controller while at the same time maintaining high performance with any number of tape drives connected.

2.3 Data compression

Data compression is a technique that encodes data bytes into fewer data bits, thereby allowing large amounts of data to be stored on less magnetic tape. Historically, these data

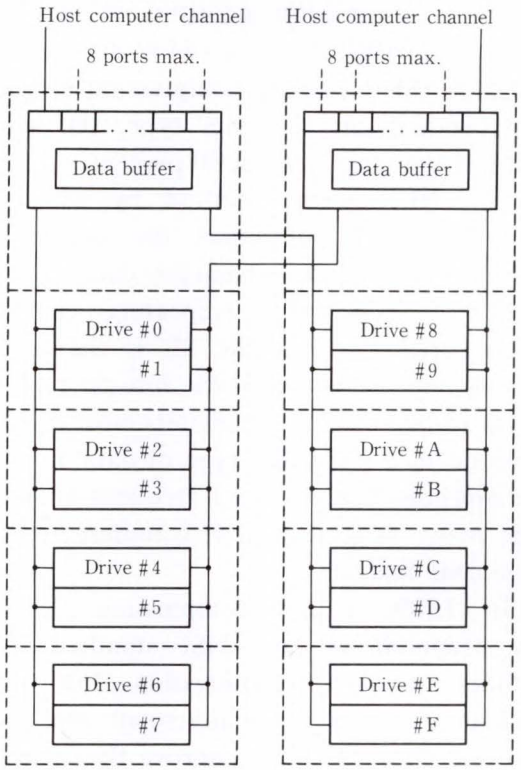


Fig. 2—Maximum system configuration.

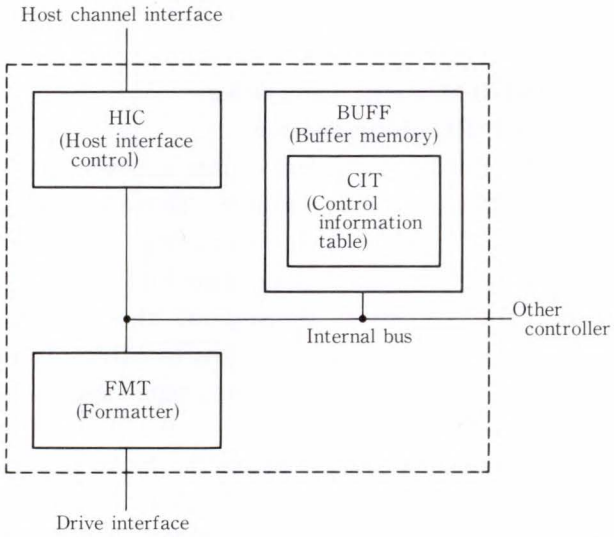


Fig. 3—F1751 Magnetic Tape Controller block diagram.

compression algorithms were handled by the host software, which resulted in excessive overhead.

To solve this problem, Fujitsu developed a data compression option for the F1751 that implements the algorithm in the hardware. This hardware can process data without any additional time loss in the host overhead. The

hardware compression algorithm is a combination of run length encoding (RLE) and statistical encoding (SE). Though compression efficiency is dependent on the data type, the average compression ratio is about 50 percent.

The RLE is the method by which the controller hardware writes the compression indication code, the character data, and the count of consecutive repeating characters in the original data. The SE is the method of assigning as few as three bits to represent a byte (8-bit) of those characters most frequently found in a particular type of data; however, the overhead for the least frequent characters is as many as sixteen bits to represent a byte of original data.

In 1989, a new compression algorithm was proposed as an ANSI standard which includes not only compression, but also a method of combining many small data blocks into one large block to improve the efficiency of the tape. The F1751/F6470 will include this standard compression feature in the near future.

2.4 F6470 Magnetic Tape Drive

2.4.1 Recording method

The F6470 Tape Drive uses a new recording method called Double Density NRZI (DDNRZI)[1] with high coercivity tape and 18 tracks of recorded data. The F6470 has six times the recording density of the previous open reel tape and data transfer rates up to 3 Mbyte/s, even though the tape speed of 2 m/s is relatively slow.

The read or write magnetic head is very important in achieving higher recording densities[2]. Thin film technology heads are used in the F6470. The read head consists of a magneto-resistive (MR) element. The write head is an inductive head to which a very narrow pulse current is applied to reduce heat dissipation because thin film technology heads have relatively high resistance windings.

2.4.2 F6470 tape path and control method

The F6470 is basically a streaming tape drive in which the tape is driven reel to reel directly[3]. There are no capstans or vacuum columns in the tape path such as are found in open reel tape drives.

The start-stop time is small since the tape inertia is less than 1/60 of the open reel tape drives. The buffer memory in the controller assures that all tape applications that were previously possible with open reel tape drives are maintained.

The microprogram control techniques maintain constant tape speed by calculating the tape radius and the reel moment of inertia.

2.5 Operability

To provide better operability, the F6470 has a tape loading mechanism that can load the cartridge automatically if the tape is inserted from the front loading window. If an unload command is issued from the host CPU, the tape is rewound and the cartridge is ejected such that it can easily be taken out by hand.

Additionally, the F6470 Tape Drive has an

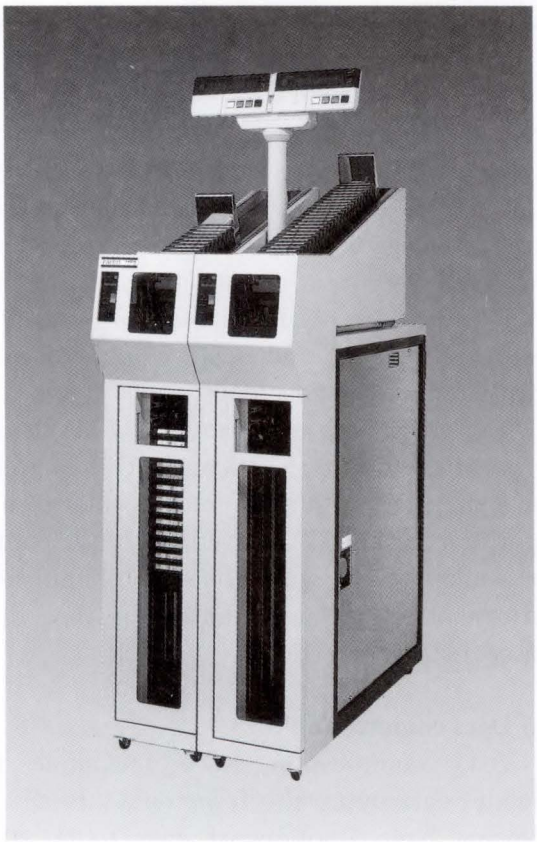


Fig. 4—Autoloader for 24 cartridges.

8-character alphanumeric display which is used to indicate the cartridge volume serial number when tape is requested to be mounted, to indicate cleaning messages, or to indicate unit check codes. All of these help the operator. Tape path cleaning is as easy as inserting a cleaning cartridge.

2.6 Auto cartridge loader

The F6470 has two types of auto cartridge loader options which can stack 6 or 24 cartridges at once. These options are most often used in magnetic disk drive back-up applications to reduce the number of cartridge mounts required to an operator. Figure 4 shows the 24 cartridge autoloader option installed on an F6470.

2.7 Reliability

To achieve a compact tape unit, Fujitsu has packaged two tape drives in each F6470 cabinet. Every component, including the power supply is duplicated, and they all are therefore isolated from each other, except for the pneumatic pump and the main power input circuit which are shared by the two drives.

The two controllers can be connected to each other, which results in a full duplex subsystem with high reliability. In addition, if an FMT (see Fig. 3) should malfunction after data has been stored in the buffer memory, the other FMT can execute the write operation for the tape drive.

2.8 Maintainability

The alphanumeric display panel indicates that all errors can be detected by the tape drive as error codes. Most repair can be done by merely replacing the component according to the error code, which reduces trouble shooting and repair time.

In order to identify problems which the tape drive finds difficult to detect, a lot of sense information is provided so that the log record has enough information. Also, the online test programs are provided in order to diagnose overall functionality of the subsystems.

Table 2. Physical parameters of IBM3480-type Cartridge

Type	Single reel
Dimensions	101.6 x 127 x 25.4 mm
Tape width	12.7 mm (0.5 inches)
Tape length	168 m
Storage capacity	200 Mbytes

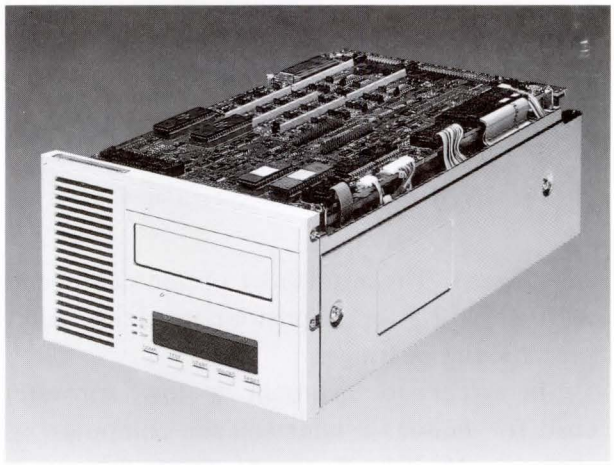


Fig. 5—M2481A/B Cartridge Tape Drive.

3. IBM3480 cartridge

The physical parameters of the IBM3480-type Cartridge are listed in Table 2.

The magnetic material[4] of the tape media is dioxide chromium (CrO_2) which is different from previous open reel tape (Fe_2O_3). This new media achieves 1.5 times the coercivity (approx. 550 Oe) of the older media. Therefore, even though the tape length is only 168 m, the smaller cartridge can store 200 Mbytes compared to the 150 Mbytes capacity of the largest open reel tape.

4. M2481A/B Cartridge Tape Drive

The M2481A/B is a compact 8-inch form factor tape drive as shown in Fig. 5. The tapes written by M2481A/B have interchangeability with the F6470.

The compact M2481A/B Tape Drives have many features of the bigger F6470 such as an air-supported tape guide and read or write head for longer tape durability, an alphanumeric message display, data transfer rates up to 3 Mbyte/s, and an autoloader option

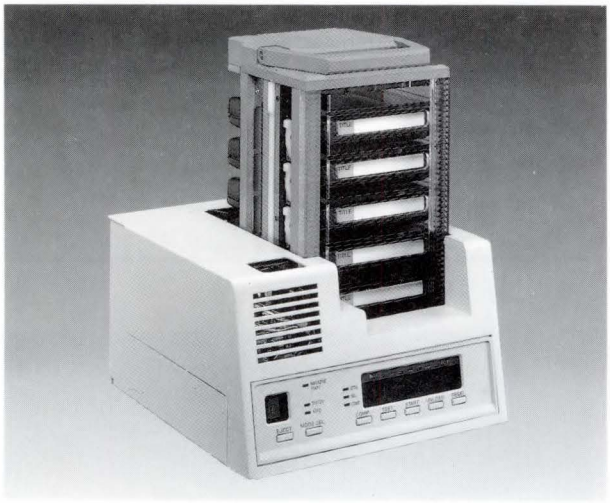


Fig. 6—Autoloader for 5 cartridges.

(see Fig. 6). The autoloader option can operate with either a 5- or 10-position magazine used to prestack the cartridges.

In order to maintain a low subsystem cost for mid-range and smaller computer systems, the M2481A/B was designed with a single controller and interface path which does not support the duplex configuration of the F6470.

5. Conclusion

Although the recently proposed data compression algorithm can increase storage capacity up to 400 Mbytes per cartridge, it is difficult to say whether this cartridge tape drive will meet customers' needs given by the current rapid growth of magnetic disk drive capacity.

Market pressure to increase the capacity and performance of tape subsystems will likely result in double density drives within the next few years.

Another impact on tape drives is the downsizing of computer systems used for distribution network processing systems. This trend will soon require tape drives that are more compact and cost-effective.

Fujitsu will continue to make every effort to provide compact and low-cost tape drives to meet these needs.

References

1) Schneider, R.C.: Write equalization in high-linear-magnetic recording. *IBM J. Res. Develop.*, **29**, 6, pp. 563-568 (1985).
2) Cannon, D.M., Dempwolf, W.R., Schmalhorst, J.M., Schelledy, F.B., and Silkensen, R.D.: Design and performance of a magnetic head for a high-density tape drive. *IBM J. Res. Develop.*, **30**, 3, pp. 270-277 (1986).
3) Winarski, D.J., Chow, W.W., Bullock, J.G., Froehlich, F.B., and Osterday, T.G.: Mechanical design of the cartridge and transport for the IBM 3480 Magnetic Tape Subsystem. *IBM J. Res. Develop.*, **30**, 6, pp. 635-644 (1986).
4) Brandshaw, R., Bhushan, B., Kalthoff, C., and Warne, M.: Chemical and mechanical performance of flexible magnetic tape containing chromium dioxide. *IBM J. Res. Develop.*, **30**, 2, pp. 203-216 (1986).

Yoshikazu Nakamura
Tape Subsystem Dep.
File System Div.
FUJITSU LIMITED
Bachelor of Electronics Eng.
Tokyo Electrical Communication University 1967
Specializing in Tape Drive Engineering

UDC 681.327.636

F 6455 Magnetic Tape Library System

• Satoru Ohtsuka • Hajime Sugiura • Tetsuo Komura

(Manuscript received May 8, 1990)

Fujitsu has developed a library system using single-reel tape cartridges as the storage media. This library system can contain up to 5 152 data cartridges, and provides storage capacity of up to 1 TB. The average cartridge transport time is 10 s.
This paper describes the latest Fujitsu library system concept, configuration and features.

1. Introduction

The library system consists of a read/write device (basic component), a media warehouse, and a media selection and carrying mechanism. The library system uses automated mounting and warehouse facilities for data storage media[1), 2)]. The library system promotes efficient use of magnetic disk storage and provides automatic data backup.

Fujitsu developed the F6460 Cartridge Library System[3)] in 1982 and the F6453 Cartridge Tape Library[4)] in 1986.

These library systems[4),5)] solve the following problems that arise as the amount of data handled by computers increases:

1) Increased space required for additional magnetic disk units.
2) Manual operation and management difficulties due to increased magnetic disk data backup processing and processing time.
3) Increased operation errors due to increased media handling.

The Fujitsu library systems use non-standard, unique storage media. Their application is limited to disk backup operations. User requirements for library systems have diversified with increased performance and capacity of magnetic disk drives.

A recent market study shows that the following functions are necessary:

1) Improved performance of library systems to match the improvements in disk performance
2) Direct use by application programs (eliminate the limitation of the applications)
3) Data exchange with other systems.

These requirements indicate a need for new library systems to be developed.

Fujitsu developed the F6470 Magnetic Tape Unit[6)] in 1987. This Magnetic Tape Unit uses high-performance, compact, convenient, and international standard magnetic tape cartridges as its storage media. The recording format of the F6470's single-reel type data cartridge also conforms to international standards, enabling data exchange with other systems.

The data transfer rate has been improved to 3 Mbyte/s which matches that of the disk drive (F6425). Capacity of the data cartridge has been increased to 205 Mbytes. Read/write operations of the F6470 are performed via the controller data buffer. This architecture reduces data access overhead time even when the data is directly read/written intermittently by application programs (start/stop mode processing). For this reason, this cartridge will replace most open-reel tapes and become a standard media for magnetic tape applications.

In line with the current market trend and consumer demand for library systems, Fujitsu

has developed the F6455 Magnetic Tape Library (MTL) system. The basic component is the F6470.

Chapter 2 describes the features of the MTL. Chapter 3 discusses the control software.

2. F6455 Magnetic Tape Library (MTL) features

The F6455 MTL main components are Directors, Magnetic Tape Units and Accessors (see Fig. 1).

1) Director (DIR)

The DIR works as a subsystem controller. It communicates with LIBSP (see chapter 3) and executes channel commands and controls the lower devices (MTUs and ACCs). Each DIR has a data input-output function and an ACC control function. The basic subsystem configuration includes two DIRs. In the maximum configuration, up to 4 DIRs can be installed (see Table 1). For I/O processing, a transfer of 1.5 Mbyte/s to 4.5 Mbyte/s can be selected to meet the channel data transmission capability.

2) Magnetic tape unit (MTU)

The MTU is a read/write device which has a load and unload mechanism for a tape cartridge. Tape load or unload time is 11 s and tape rewind time is 48 s (see Table 2). The basic subsystem configuration includes 8 MTUs. In the maximum configuration, up to 16 MTUs can be installed (see Table 1).

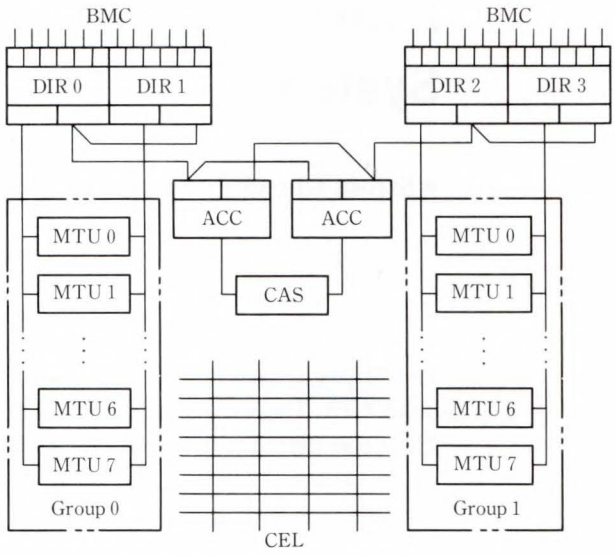
3) Accessor (ACC)

The ACC is a cartridge carrying mechanism. Two ACCs (right and left ACCs) are mounted. The right and left ACCs are placed so that they move horizontally on the same rail and they operate alternately to reduce the movement time. The average movement time is 10 s (see Table 2 and Fig. 2).

The F6455 MTL has the following five main features.

2.1 High-speed and large-capacity backup processing

The F6470 Magnetic Tape Unit used as the basic component provides optimum data transfer capability for high-speed (3 Mbyte/s) backup processing. The F6455 MTL can contain up to 5 152 data cartridges (205 Mbytes per volume),



Group 0: Accessed only by DIR 0 and DIR 1
Group 1: Accessed only by DIR 2 and DIR 3
ACC : Accessor — CEL : Cell
BMC: Block multiplexer channel — DIR : Director
CAS : Cartridge access station — MTU: Magnetic tape unit

Note: This configuration is available to Models 40B, 51B, 62B, 73B and 84B

Fig. 1—Subsystem configuration.

providing a subsystem storage capacity of up to 1 056 Gbytes (see Table 1).

In addition, there is a data compression feature as a standard function and this increases the data cartridge capacity (twice as large on the average). This feature enables the subsystem storage capacity to be increased to 2 000 Gbytes. A 50 percent compression efficiency (on the average) can be obtained from the following two compression methods:

1) Compression of continuous character strings (run length encoding) which converts repetitive codes to a data format consisting of the repetition count and the code.
2) Frequent character compression (statistical encoding) which converts characters with a high frequency of occurrence to special codes.

2.2 Efficient system operation

Cartridge handling is automated. This minimizes operator work such as mounting/demounting, carrying, and tape storage, the bottleneck of conventional systems.

Table 1. Models

Model	Maximum storage capacity (Gbyte)	Number of cartridges (unit)	Number of devices installed		
			DIR	MTU	ACC
10B	134.8	658	2	4	2
20B	250.5	1 222	2, 4	8	2
21B	266.5	1 300	2	4	2
30B	366.1	1 786	4	12	2
31B	382.1	1 864	2, 4	8	2
32B	398.1	1 942	2	4	2
40B	481.7	2 350	4	16	2
41B	497.7	2 428	4	12	2
42B	513.7	2 506	2, 4	8	2
43B	529.7	2 584	2	4	2
51B	613.3	2 992	4	16	2
52B	629.3	3 070	4	12	2
53B	645.3	3 148	2, 4	8	2
54B	661.3	3 226	2	4	2
62B	744.9	3 634	4	16	2
63B	760.9	3 712	4	12	2
64B	776.9	3 790	2, 4	8	2
65B	792.9	3 868	2	4	2
73B	876.5	4 276	4	16	2
74B	892.5	4 354	4	12	2
75B	908.5	4 432	2, 4	8	2
76B	924.5	4 510	2	4	2
84B	1 008.1	4 918	4	16	2
85B	1 024.1	4 996	4	12	2
86B	1 040.1	5 074	2, 4	8	2
87B	1 056.1	5 152	2	4	2

2.3 Flexible subsystem configuration

There are twenty-six possible configurations (135 Gbytes to 1 056 Gbytes) through combining cabinet frames in which cartridge storage cells (CELs) or components (DIR, MTU, ACC) are mounted (see Tables 1 and 3, and Figs. 1 and 3). The optimum storage capacity and subsystem configuration can be selected according to the system requirement. The subsystem magnetic tape unit as well as data cartridges can be shared by two or more host systems. Flexible conversion paths facilitate possible system upgrading, such as storage capacity expansion, after installation.

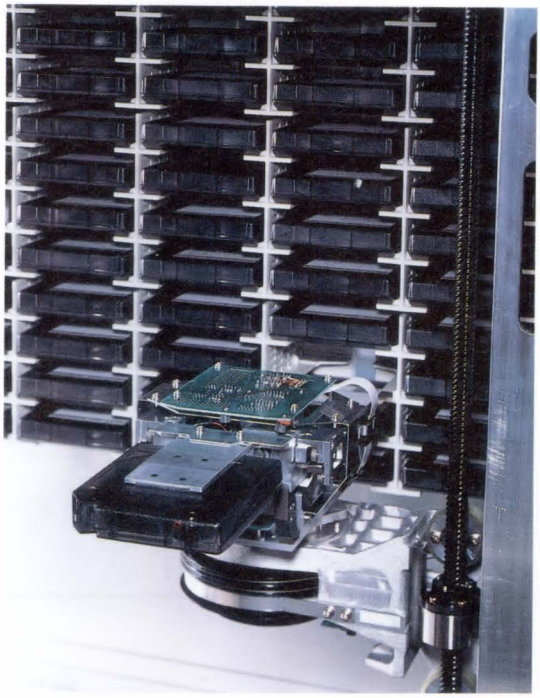


Fig. 2—Handling mechanism.

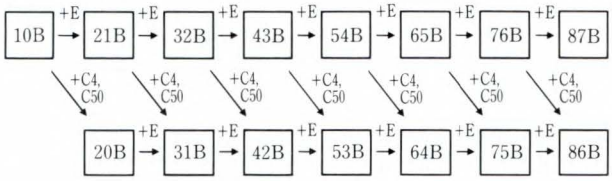
Table 2. System specifications

Item		Specification
No. of cartridges stored		658-5 152
Storage capacity		134.8-1 056.1 Gbytes
No. of tracks/cartridge		18
Data transfer speed		3 Mbyte/s (director to MTU)
Data recording method		GCR
Data compression feature		Standard equipped
Recording density		37 871 bpi
Interblock gap (IBG)		2.0 mm
Tape running speed		2.0 m/s
Tape rewind time		48 s
Tape load/unload time		11 s
Cartridge transport time		10 s average
Components	Directors	2, 4
	MTUs	4, 8, 12, 16
	Accessors	2
Cartridge access station		1 each: up to 10 cartridges can inserted and ejected
Connected channels		BMC with transfer rate of 1.5, 2, 3, and 4.5 Mbyte/s
No. of channel connections		Up to 24 per subsystem
Subsystem width		Approximately 5-16 m

Table 3. Frame configuration

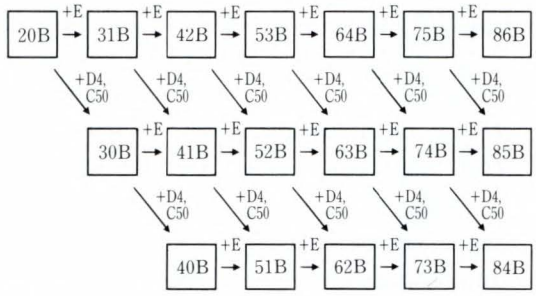
Frame type	Component				Remarks
	DIR	MTU	ACC	CEL	
F6455 A2	2	–	1	–	Includes DIR 0/1 and R-ACC
F6455 B0	–	–	1	–	Includes L-ACC only
F6455 B2	2	–	1	–	Includes DIR 2/3 and L-ACC
F6455 C4	–	4	–	–	Includes the PDU (power distribution unit)
F6455 D4	–	4	–	–	
F6455 A30	–	–	–	5	Includes CAS and R-garage
F6455 B30	–	–	–	89	Includes L-garage
F6455 C50	–	–	–	564	R: right L: left
F6455 E	–	–	–	642	

1) Available to two DIRs models



Note : Model 10B consists of A2 + A30, C4 + C50, and B0 + B30
Two DIRs models cannot be upgraded to four DIRs models

2) Available to four DIRs models



Note : Model 20B consists of A2 + A30, two C4 + C50, and B2 + B30

Fig. 3—Upgrading path.

2.4 High reliability

Redundant read/write mechanisms, accessors, and controllers improve subsystem reliability and enable maintenance without stopping the system (active maintenance).

1) MTU multiplexing

The minimum configuration consists of 4 MTUs and the maximum configuration of 16 MTUs. If an MTU fails, the defective MTU is blockaded and I/O operation is switched to an alternate MTU by the operating system's dynamic device reconfiguration function.

MTUs are maintained in units of two.

2) ACC duplexing

The ACC is duplexed. Should one accessor fail, subsystem operation is automatically switched to single-accessor mode operation. The defective ACC is stored in a garage and the other ACC resumes operation. The defective ACC can be checked and repaired in the garage and put back into service without stopping the system.

The control unit and power unit of the ACC are also duplexed (right and left). This enables active maintenance for recovery work after processing by the alternate ACC. In this subsystem, a total system failure caused by the ACC will only occur if both ACCs are defective. The probability of this failure is assumed to be one occurrence in 70 000 h of operation.

3) DIR duplexing and internal paths

The basic configuration consists of two DIRs. Two DIRs control up to eight MTUs and two ACCs. Each DIR contains an MTU data I/O path. If one DIR fails, the other DIR continues processing.

Each DIR has an independent power unit. Each DIR is maintained independently.

2.5 Cartridge operation mechanism

The standard cartridge entry and exit mechanism (CAS: cartridge access station) handles up to ten cartridges. An optional mass entry and exit mechanism is available for handling up to 90 cartridges.

The cartridge entry mechanism in CAS has a bar code reader mechanism for bar code labeled cartridges. This mechanism enables the software to manage correspondence between logical medium identifiers (VSN: volume serial numbers) which are magnetically recorded on the media, and medium identifiers which can be recognized by the operator.

3. Software

The hardware components of F6455 MTL are controlled by the software (LIBSP: Library

system support program) which works by the host CPU in the Block Multiplexer Channel (BMC) interface.

The following characteristics have been realized by the development, positioning of LIBSP as the subsystem which works under the operating system:

1) Flexible software recovery can be executed for errors which cannot be recovered through hardware recovery and thus high reliability is realized.
2) Expansion of the hardware composition and change in the future can easily be achieved.
3) Control becomes possible for library units other than MTL making common interface with main hardware components as library such as ACC and CAS.

LIBSP has the following five functions.

3.1 Warehouse management

LIBSP manages the media in the MTL using the X, Y, and Z cell addresses. The coordinates are stored in a control data set (CDS). The cartridge and cell positions are kept unique by coordinate management (see Fig. 4). The main functions of the warehouse management are volume management, cell management, and write protection control.

3.1.1 Volume management

The warehouse management function manages all volumes stored in the MTL cells. The volumes are divided into an active volumes group which is accessable by systems and an inactive volumes group which cannot be used. Users can access inactive volumes only when they are changed to active volumes by LIBSP.

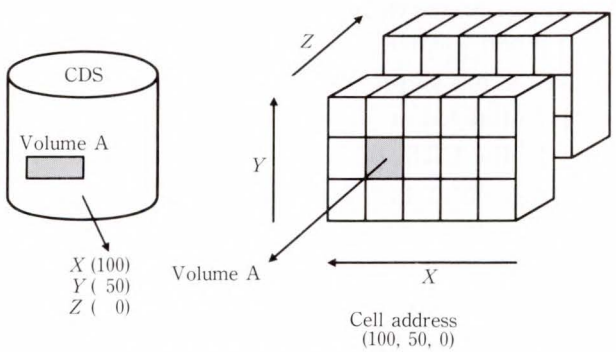


Fig. 4—Warehouse management with MTL.

Active volumes are further divided into private volumes and scratch volumes. Private volumes can be accessed only by specific users and scratch volumes can be accessed by any user. The private or scratch attribute is set during cartridge initialization.

A scratch volume is changed to a private volume when requested by a user. A private volume can be changed to a scratch volume when the cartridge is reinitialized or when a cartridge attribute change utility is used.

3.1.2 Cell management

The warehouse management function manages the cell use status (in use, not in use, and the volume attributes described above). LIBSP supports volume labels in the standard format (SL format) only.

3.1.3 Write protection control

The warehouse management controls volume write protection.

The following methods are available for write protection:

1) Physical protection using the write protection feature for cartridges
2) Logical protection by LIBPS and MTL, which inhibits the write operation when cartridges are mounted. The attribute of the cartridge is defined by use of the cartridge attribute change utility.

Logical protection is validated when LIBSP indicates that the MTL inhibits the write operation. This indication is made when the mount request for the relevant cartridge is issued after the user indicates write inhibition in LIBSP. Physical protection has priority over logical protection.

3.2 Device management

The device management facility manages the MTL hardware components. The device management facility has the following three functions.

3.2.1 Carrying cartridges

The device management facility accepts a mount request from a system and directs the ACC to move the cartridge.

When cartridge movement is successful, the device management facility monitors the MTU ready condition (loaded ready). If cartridge

movement fails, the device management facility checks the current position of the cartridge. Depending on the current position of the cartridge, the device management may move the cartridge again by using the ACC (recovery of cartridge movement).

When the cartridge is to be moved from or to an MTU, the MTU status (whether the MTU contains a cartridge) is checked and recovery processing is then performed on the basis of the ACC information.

If one ACC fails during cartridge movement recovery processing, a message is issued to inform the operator that the ACC cannot be used. Processing resumes using the other ACC. The unusable ACC is recovered after the customer engineer has repaired it. If both ACCs fail, the entire subsystem is disabled. If a mount request is issued when the subsystem is faulty, processing terminates with an error. LIBSP also controls the ACC paths. If one of the ACC paths is closed due to an error, processing resumes using the other path.

3.2.2 MTU control

LIBSP manages the MTU use status (whether the MTU is usable). It also manages the jobs and host for the MTU being used. The LIBSP operator command (VARY) is used to change an MTU from an unusable state to a usable state.

1) MTU blockade

When LIBSP recognizes that an MTU is unusable, it automatically blockades the MTU. Information about the blockaded MTU is reported to the operator by a system message. If a cartridge is left in a blockaded MTU, a system message is issued to inform the operator that the cartridge is to be removed from the MTU by the customer engineer.

2) MTU contention

If an MTU that is already being used by a job is requested by another job, completion of the current job has priority. If the MTU is still busy after five minutes, the operator is prompted for subsequent processing. If the operator chooses to wait, five more minutes are given for the job to be completed. If the operator chooses to cancel, the requested job is canceled.

3) DDR

MTU control automatically performs dynamic device reconfiguration (DDR) processing. The MTU subject to DDR is automatically selected so that processing can continue.

4) MTU cleaning

Normally, MTU heads are automatically cleaned after a head cleaning request is issued by the MTU. However, when cooperating jobs are processed and extension of processing time due to head cleaning is not desired, the operator can issue a command to clean the heads. The operator must specify the device address of the MTU.

3.2.3 Cartridge insertion and ejection control

Users cannot manually load cartridges into the library system. The CAS is used to insert or eject cartridges.

1) Cartridge insertion

There are two cartridge insertion modes. One mode inserts a small number of cartridges in a normal operation. The ordinary CAS is used in this mode.

The other mode inserts a large number of cartridges. The mass entry and exit mechanism is used in this mode.

2) Cartridge insertion operation

The insertion operation is divided into two types:

a) Insertion with label check
The cartridge is used immediately after it is inserted.

b) Insertion without label check
A large number of cartridges can be processed quickly.

3) Active cartridge

Insertion with label check is performed via the MTU. The MTU checks the label information of the inserted cartridge (active cartridge).

When a cartridge which has not been initialized, or has shipped from the factory is inserted, it is initialized and then stored in a cell.

In the CDS, the cartridge is registered by its volume serial number after initialization. When an initialized cartridge is inserted, the cartridge is registered by its volume serial number in the CDS.

4) Inactive cartridge

For insertion without label check, cartridges

are stored into the empty cells directly, not via the MTU, and cartridges are maintained as inactive cartridges.

General users cannot access inactive cartridges. LIBSP assigns virtual volume identifiers to such cartridges that can be initialized easily using this virtual volume ID.

A mount request cannot be made using a virtual volume identifier. For a cartridge inserted without label check, its volume serial number must be reported to CDS by using the operator command CHKCELL, or the cartridge must be initialized by using a utility.

5) Cartridge ejection operation

Ejection can be achieved by either an operator command or a utility. The operator command is used for ejection in an emergency or to eject a single cartridge. The utility is used to eject a large number of cartridges or those having consecutive volume serial numbers. Mass insertion and ejection is done only by the utility.

Normally, the entry and exit ports are used properly according to the user requirements. However, when one port cannot be used, the other port is used for insertion and ejection. In cooperation with the MTL, the unusable entry and exit port is closed regardless of the subsystem. The closed port can be made usable by operator command (VARY) after a customer engineer has serviced the port.

3.3 Mount/Demount Control

A cartridge requested by a system is automatically mounted/demounted. This section explains the mount/demount control.

3.3.1 Cartridge mount control

Cartridge mount processing is started with the mount message issued by the operating system (OS) (see Fig. 5).

After issuing a cartridge mount request, the system waits for the ready interrupt from the MTU. When the MTU subject to the mount request is not included in the library system, LIBSP entrusts the cartridge mount operation to the operator, as with conventional magnetic tape units. If the MTU subject to the mount request cannot be used, a usable MTU is automatically

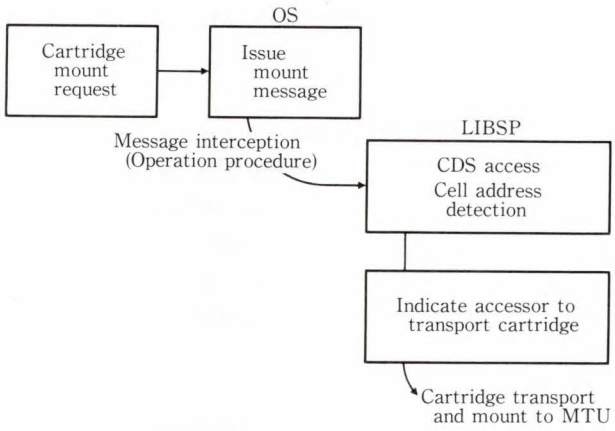


Fig. 5—Cartridge mount control.

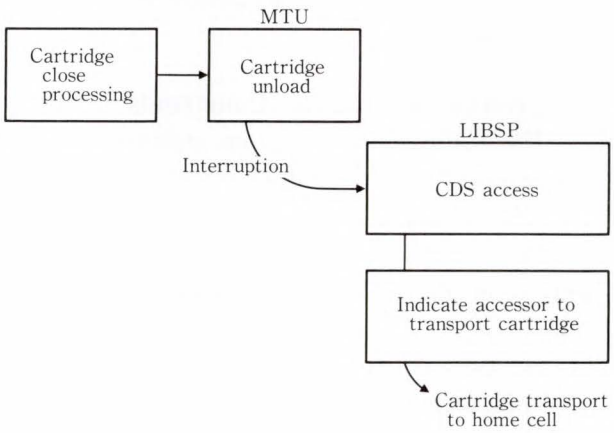


Fig. 6—Cartridge demount control.

selected and processing resumes. If an error occurs during cartridge loading, a reload is automatically tried for the same MTU.

If an MTU becomes unusable due to a hardware failure, LIBSP issues an error message to the operator. The MTU then enters the disabled state.

3.3.2 Cartridge demount control

Cartridge demount processing begins at the cartridge close processing by the system. Actual demount processing begins when completion of cartridge unloading is posted from the MTU (see Fig. 6).

3.4 Control Data Set Recovery Procedure

The MTL is managed by CDS created on a magnetic disk unit. The MTL cannot be controlled if CDS is destroyed.

The CDS recovery function is provided for this case. CDS can be recovered by the following two methods:

1) Without journal data

CDS is recovered using the MTL hardware information and the cartridge information in the MTL. Each cartridge in the MTL is checked one by one and CDS is created according to their contents. Much time is required for recovery by the above method.

2) With journal data

This method recovers CDS by using the CDS backup data collected in a cycle and the journal data collected every time the cartridge information in CDS is changed. Recovery by this method is quick.

3.5 Utilities and Operator Commands

The following subsections outline the utilities and operator commands.

3.5.1 Utilities

The utilities are used to manage the cartridges in the MTL. Each utility is explained below.

1) Ejection utility

This utility is used to eject cartridges. It enables ejection for consecutive volume serial numbers, ejection for each cartridge owner, and ejection depending on the number of mount operations.

2) Cartridge attribute change utility

This utility specifies the cartridge attribute (private or scratch) and which cartridges are write-protected.

3) Initialization utility

This utility initializes cartridges. Cartridges in the MTL must be initialized with this utility.

4) List utility

This utility prints the cartridge and cell information.

3.5.2 Operator commands

LIBSP provides the following operator commands for MTL operation and maintenance.

1) Ejection command

This command is used for ejection in an emergency or to eject a single cartridge.

2) Device information display command

This command displays the status of the ACCs, MTUs, and CAS of the MTL. For the CAS, it displays the insertion status. For an MTU, it displays the use status, the host using the MTU, and the volume serial number used. For an ACC, it displays whether the ACC is operable.

3) Device maintenance command

This command changes the status of the ACCs, MTUs, and CAS. It changes an unusable ACC, MTU, or CAS to the usable state. If a host fails, the MTU which was used by this host can be made usable by another host using this command.

4) Entry command

This command specifies the insertion mode and the start and stop of insertion. If a host fails during insertion, the host can be switched so that insertion can resume.

5) CDS backup synchronization command

This command synchronizes journal data with CDS backup. The execution of utilities and operator commands (except for mount processing) is suppressed while this command is being executed.

6) Error log collection command

This command collects the error log information of the MTL in the error log data set.

7) CDS matching command

This command creates CDS according to the volume serial numbers of the cartridges in the MTL.

4. Conclusion

Fujitsu has developed comprehensive and advanced library system products for the external storage hierarchy. The F6455 Magnetic Tape Library System using cartridge storage is cost-effective and can handle many applications. For this reason, we expect this system to become the nucleus of external storage devices.

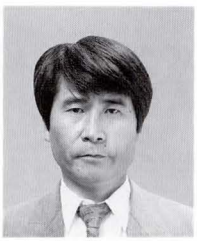
System security has also become important. For example, the automated system management for data integrity has been required in addition to automatic operation that merely eliminates the need for operators.

In order to meet such market demands, future library systems must be able to smoothly cooperate with existing third-party system

management programs, while maintaining operation compatibility and consistent user interface.

References

1) Johnson, C.T.: The IBM 3850: A Mass Storage System with Disk Characteristics. *Proc. IEEE,* **63**, 8, pp. 1166-1170 (1975).
2) Harris, J.P., Rohde, R.S., and Arter, N.K.: The IBM 3850 Mass Storage System: Design Aspects. *Proc. IEEE,* **63**, 8, pp. 1171-1176 (1975).
3) Sugiura, H., Moroto, S., and Suzuki, K.: FACOM 6460 Cartridge Library System (CLS). (in Japanese), *FUJITSU,* **34**, 6, pp. 1015-1024 (1983).
4) Inoue, Y., Ogawa, Y., and Kaneko, S.: External Storage Devices. (in Japanese), *FUJITSU,* **38**, 5 (Special Issue: Information Processing Devices), pp. 347-362 (1987).
5) Sugiura, H., and Hayakawa, A.: External Storage Devices. (in Japanese), *FUJITSU,* **37**, 2 (Special Issue: FACOM M-780 Computer Systems), pp. 148-154 (1986).
6) Nakamura, Y.: F1751/F6470 Magnetic Tape Subsystem. *FUJITSU Sci. Tech. J.*, **26**, 4 (Fujitsu File Devices), pp. 316-320 (1991).



Satoru Ohtsuka

Engineering Dept.
File System Div.
FUJITSU LIMITED
Bachelor of Electronic Eng.
University of Electro-Communication 1972
Master of Communications Eng.
University of Electro-Communications 1974
Specializing in Magnetic and Optical Storage Systems

Hajime Sugiura

Planning Dept.
Research and Planning Div.
FUJITSU LIMITED
FUJITSU Technical College 1973
Specializing in Magnetic Storage Systems

Tetsuo Komura

Development Dept.
Software Div.
FUJITSU LIMITED
FUJITSU Technical College 1976
Specializing in Development of Basic Software Design

UDC 681.327.634†681.327.68

F6443D Magneto-Optical Disk Drive Subsystem

• Ryosuke Kudou • Hiroshi Ichii • Akio Futamata

(Manuscript received June 6, 1990)

The large storage capacity, low cost per bit, and high reliability of optical disk drives make them an attractive alternative to magnetic disks and tapes for future external nonvolatile data storage applications.

A large storage capacity, high-speed access, F6443D magneto-optical disk drive subsystem has been developed. This is the first product of its type to be marketed. The new subsystem can store 36 Gbytes in a space of only 1 m^2, and can access any part of the media within five seconds.

This subsystem is expected to boost system efficiency in general-purpose computer applications.

1. Introduction

Optical disk drives are evolving from read-only, write-once types to drives that have rewrite capability. Much attention has been focused on the third generation of optical disk drives and their potential to change the hierarchical structure of external storage devices.

This paper describes the F6443D magneto-optical disk drive subsystem, which is designed for use with general-purpose computers. Unlike previous library systems, this subsystem consists of eight disks mounted on each spindle. This arrangement provides both high-speed access and a large-storage capacity.

2. Development background

2.1 System requirements

Computers continue to handle growing volumes of data, and systems are supporting more magnetic disk drives than ever before. In larger systems, the cost of magnetic disk drives now accounts for more than half of the total system cost. This trend has led to a continuing need for a random-access file having a lower cost per bit and a smaller footprint than that of magnetic disk drives.

Optical disk drives have attracted much attention because their recording densities can be more than several times higher than that of magnetic disk drives. Users of write-once, read-only optical disk drives have great expectations for the new rewritable optical drives. Fujitsu has now developed magneto-optical technologies that bring the rewritable optical disk drive to commercial use with high reliability.

The introduction of the magneto-optical disk drive subsystem has two purposes. The first of these purposes is to address the needs of research institutions for the handling of large volumes of information. The second is to demonstrate the practicality of using optical disk drives in large computer systems for which reliability is critical.

Figure 1 shows the position of the subsystem in the external storage hierarchy[1]. Conventionally, magnetic disk drives, magnetic tape libraries, and write-once optical disk libraries are used to store large amounts of data[2]. In terms of storage capacity and access time, the subsystem can be viewed as external storage that

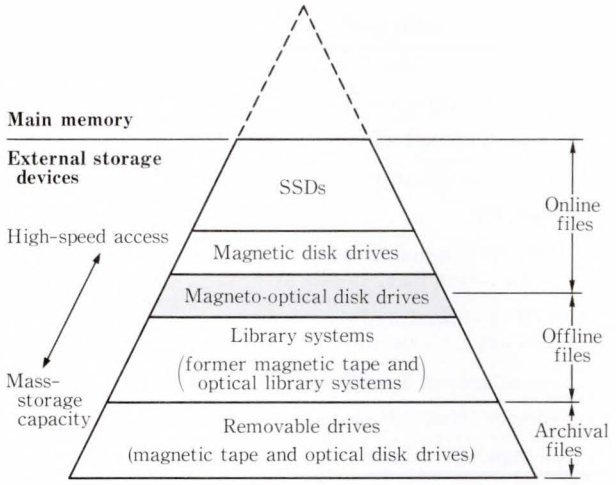


Fig. 1—External storage hierachy.

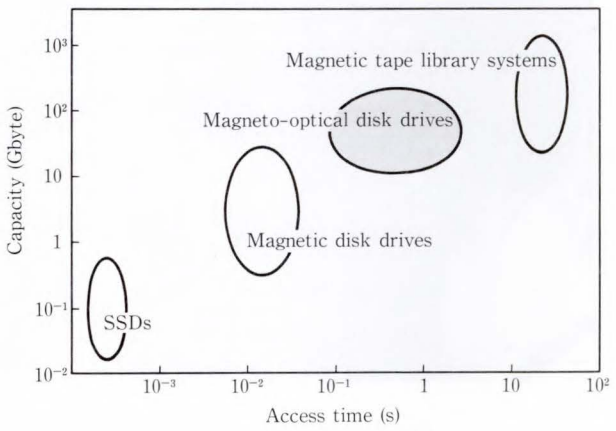
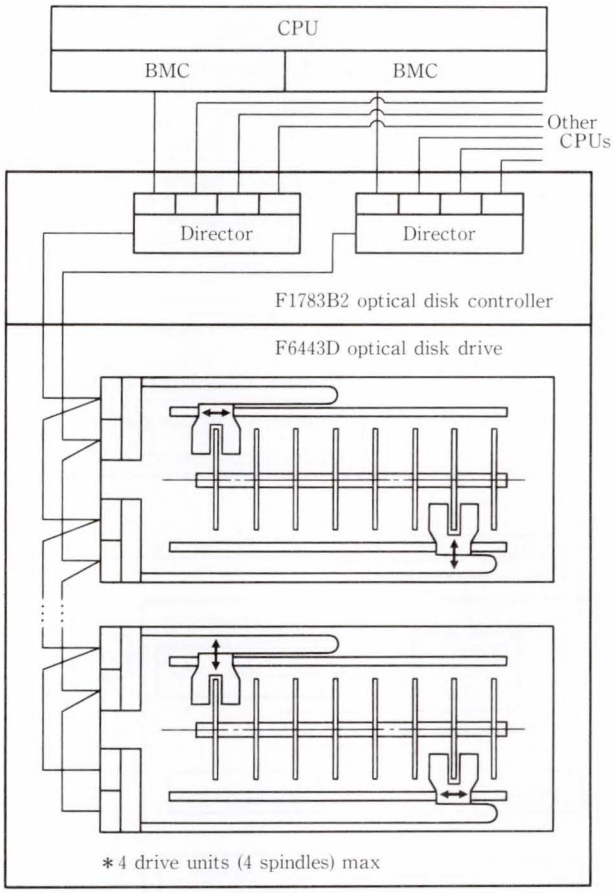


Fig. 2—Storage capacity and access time.



BMC: Block Multiplexer Channel

Fig. 3—Example of subsystem configuration.

fills the gap between magnetic disk drives and magnetic tape libraries as shown in Fig. 2. This subsystem provides a storage capacity several times higher than that of magnetic disk drives and an access speed several times faster than that of conventional libraries.

These features primarily recommend the subsystem for use as a replacement for magnetic disk drives containing data that is accessed relatively infrequently[3].

3. Subsystem

3.1 Subsystem configuration

The subsystem consists of an F1783B2 optical disk controller and an F6443D optical disk drive (see Fig. 3). The drive can contain up to four drive units. In the maximum configuration, the subsystem can store about 36 Gbytes of data. Despite this huge storage capacity, this compact subsystem saves space (the dimensions are given in Table 1).

3.2 F1783B2 optical disk controller

The F1783B2 optical disk controller contains a pair of directors in a single cabinet for control of up to four drive units. The controller is connected to a Fujitsu M-series system via block multiplexer channels (BMCs). Its cross-call capability enables the access path to be duplex. The controller can be optionally connected to up to four BMCs.

Because the subsystem records coded data, the controller uses high-speed real-time Error Checking and Correction (ECC) for high-performance error correction. LSI circuits and microprocessors are used to provide compact and advanced control circuits. The advanced

Table 1. Specifications of magneto-disk drive subsystem

Recording capacity/subsystem (formatted)		8.9 Gbytes to 35.6 Gbytes (1 to 4 spindles)
Outside dimensions (mm)		1 157 (*l*) x 930 (*b*) x 1 400 (*h*)
Drive unit	Number of disks	8 (8-inch diameter)
	Number of heads	4 (2 heads/actuator)
	Rotational speed	1 800 rpm
	Access time	200 ms (on-media av.) 5 s (media-to-media av.)
	Data transfer speed	979 kbytes/s (read) 326 kbytes/s (write)
Controller	Number of channel-paths	2 (standard) to 8 (option)
	Number of device-paths	2 (device cross-call available)
	Uncorrectable error rate	less than 10^{-12} (with ECC)

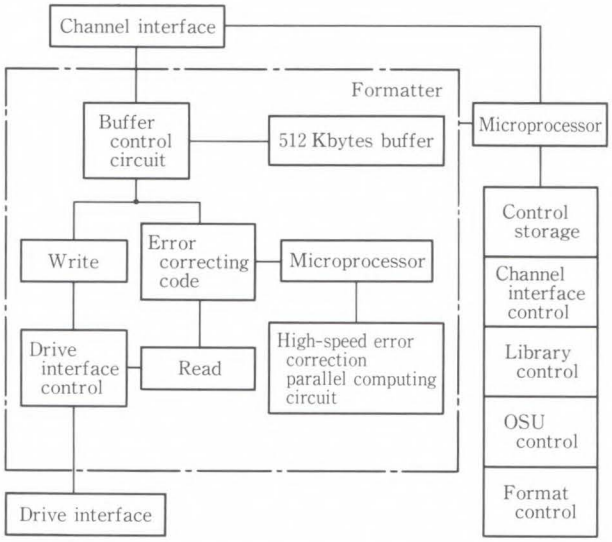


Fig. 4—F1783B2 block diagram.

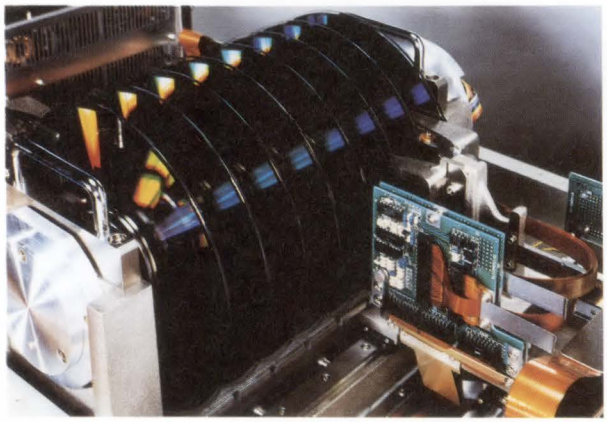


Fig. 5—Fixed disk drive.

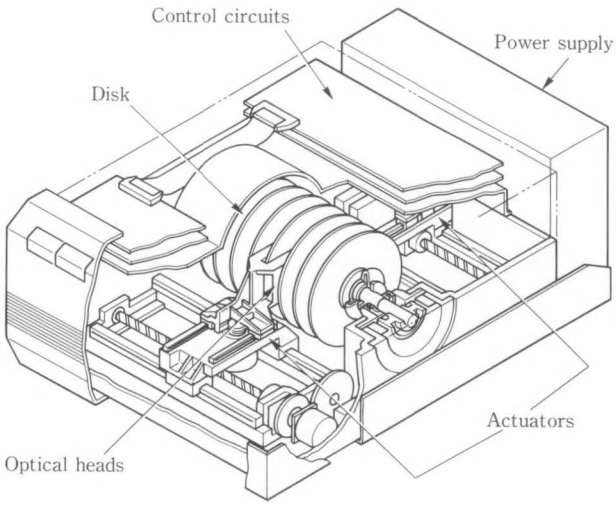


Fig. 6—Drive unit.

intelligence concepts implemented in the controller include comprehensive error recovery management in the subsystem, random data reading/writing, and automatic alternate-block management. Figure 4 is the block diagram of the controller.

The controller supports a library command scheme to enable efficient software-based access scheduling.

3.3 Drive unit

The features of the drive unit are summarized below.

1) Fixed-disk drive

The drive unit contains eight eight-inch optical disks on a single shaft (as in a magnetic disk drive) for high-speed access (see Fig. 5). The disks have glass media substrates. To improve reliability, the entire mechanical assembly,

including the disks, optical heads, and actuators, is fully enclosed to protect against dust.

Figure 6 shows the structure of the drive unit.

2) Dual actuators

The drive unit has two actuators which position a pair of optical heads over the specified tracks on the specified disk. This obviates the need for head rotation when accessing the opposite side of the disk. Because each actuator has an independent power supply, control circuit, and signal processing circuit, the subsystem can control the two actuators independently (separate logical device addresses) and, therefore, efficiently. If either actuator fails, the other can be used instead. This redundancy feature greatly improves subsystem reliability.

Two stages of access control are implemented for the optical heads: media-to-media access for disk selection, and on-media access for accessing tracks on the selected disk.

On-media access is performed using coarse and fine control. Fine control is performed by the Fujitsu-ILT (Inter Linked Tracking) lens control tracking scheme. ILT covers a wider range than conventional fine tracking schemes, which makes both the circuitry and mechanism simpler and less expensive.

Media-to-media access is supported by high-reliability features such as high-precision ball-bearing screws and a function that keeps the heads from moving between media at non-normal positions.

3) Low power consumption and small size

The F6443D optical disk drive saves space and power. It requires only 60 percent of the power per unit capacity required by the F6427H large storage capacity magnetic disk drive currently shipped by Fujitsu. Excluding the maintenance area, the F6443D requires only 68 percent of the space required by the F6427H.

4. Drive technology

4.1 Magneto-optical recording principle

In magneto-optical recording, data is magnetically recorded on the disk perpendicular to the disk magnetic recording layer (see Fig. 7). When data is recorded or erased, a small beam of laser light is directed onto the disk to locally increase the disk temperature and thus lower the coercive force of the recording layer. At the same time, a weak magnetic field is applied to the disk to turn the direction of magnetization in that part of the recording layer. At room temperature, the recording layer has a coercive force large enough to allow retention of the recorded data.

The phenomena in which the polarization plane of the reflected beam rotates slightly clockwise or counter-clockwise with the direction of magnetization of the recording layer is known as the Kerr effect. This angle of rotation can be detected as a signal state change and is used to read the recorded data.

4.2 Magneto-optical heads

Figures 8 and 9 show a magneto-optical head and its internal structure.

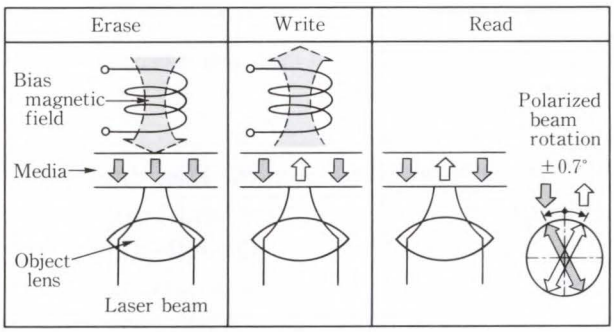


Fig. 7—Principle of magneto-optical recording.

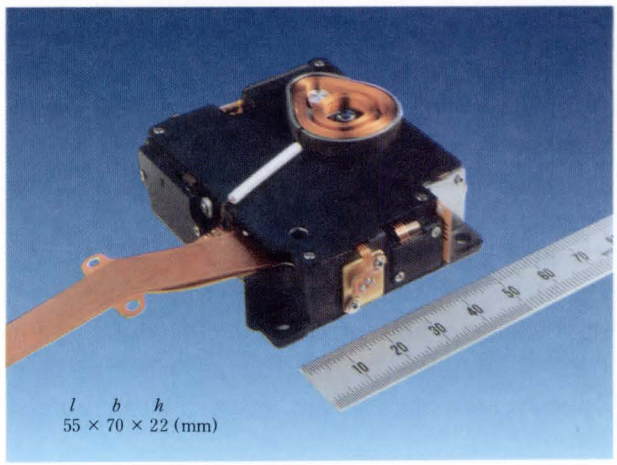


Fig. 8—Magneto-optical head.

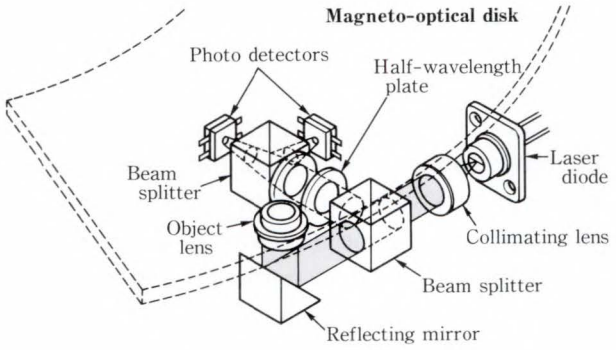


Fig. 9—Internal structure of optical head.

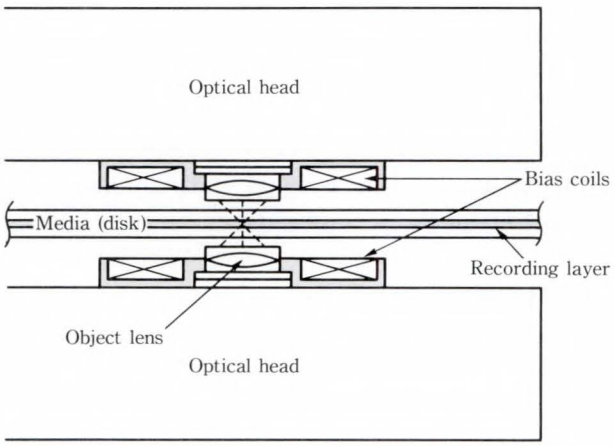


Fig. 10—Configuration of face-to-face type magneto-optical heads.

The light emitted from the laser diode passes through the collimating lens and beam splitter. The beam is deflected by the reflecting mirror, and then focused onto the recording layer of the disk by the object lens. When recorded data is read, the polarized beam reflected from the recording layer is directed back through the beam splitter along the same route. The beam is split into two beams by the splitter after passing through a half-wavelength plate. Finally, the magneto-optical signal is detected differentially by two photo-detectors.

The magneto-optical heads have the following feautres:

1) Bias magnetic field application

Magneto-optical recording writes and erases data using a bias magnetic field. Usually, this field is applied by placing an electromagnet opposite the head on the other side of the disk (see Fig. 7).

Optical heads are positioned on both sides of

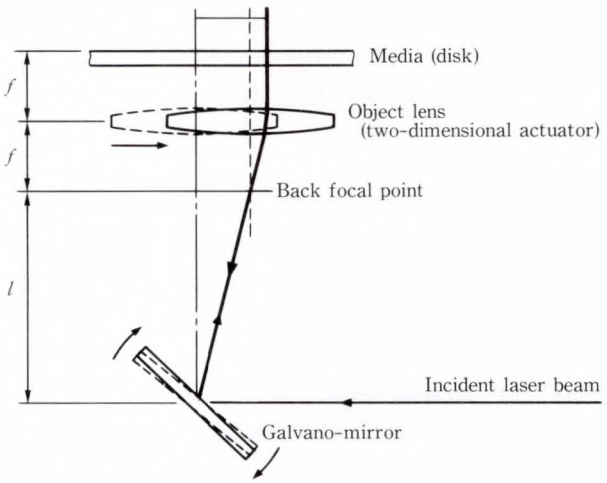


Fig. 11—Principle of interlinked tracking (ILT).

the disk (see Fig. 10). The bias magnetic field is produced by thin bias coils placed around the object lens. This arrangement enables push-pull drive of two bias coils, which reduces head volume, head mass, and power consumption. Push-pull drive achieves a better than 75 percent power reduction over conventional drive arrangements.

2) Lens tracking control

A stepping motor is used for coarse access. Because the positioning precision of a stepping motor is very rough, coase access using a stepping motor can be done only with optical heads having a wide tracking range. A new technique called ILT is used to overcome this problem (see Fig. 11). A galvano-mirror is driven in step with the object lens so that the laser beam always passes through the rear focal point. This tracking method keeps the laser beam from going off-axis even if the object lens moves over a wide range[4].

3) Magneto-optical signal reproduction

Magneto-optical signal reproduction requires precise control of the optical system, signal processing, and the associated circuits that detect the minute Kerr rotation angles. A low-noise amplifier that can be embedded in the head, and techniques for ensuring precise control of wavefront aberrations of optical components and the phase of polarized light have been developed. These techniques, coupled with the new medium design concepts outlined in the next section, give reproduced magneto-

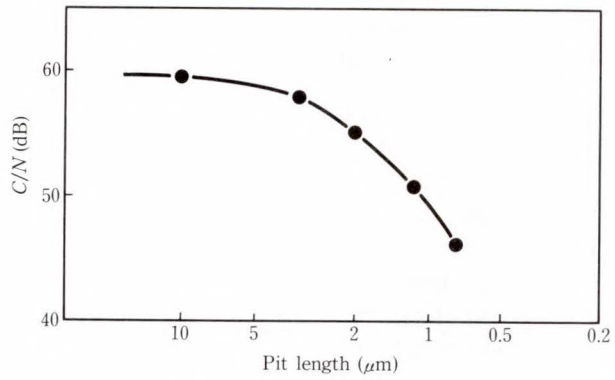


Fig. 12—Carrier-to-noise ratio vs. pit-length.

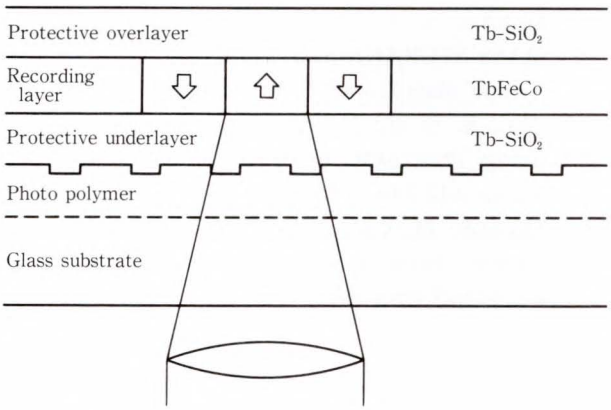


Fig. 13—Cross section of magneto-optical disk.

optical signals a carrier-to-noise ratio (C/N) of about 50 dB even for the shortest recorded pit length (see Fig. 12). This is more than adequate for recording.

4.3 Magneto-optical disks

1) Disk composition

The eight-inch media substrates are made of special reinforced glass. The beam guide grooves are made from an ultraviolet-cured photo-polymer (2p) and are formed by replication from a master disk. The glass substrate resists mechanical deformation and extreme ambient conditions better than plastic substrates.

As shown in Fig. 13, the recording layer is a thin alloy film of terbium (Tb), iron (Fe), and cobalt (Co). This film enables easy formation of perpendicularly magnetized films by sputtering. Being amorphous, this film eliminates the noise caused by grain boundaries. However, because it is a rare-earth metal, terbium is easily oxidized. To overcome this problem, the protective layer

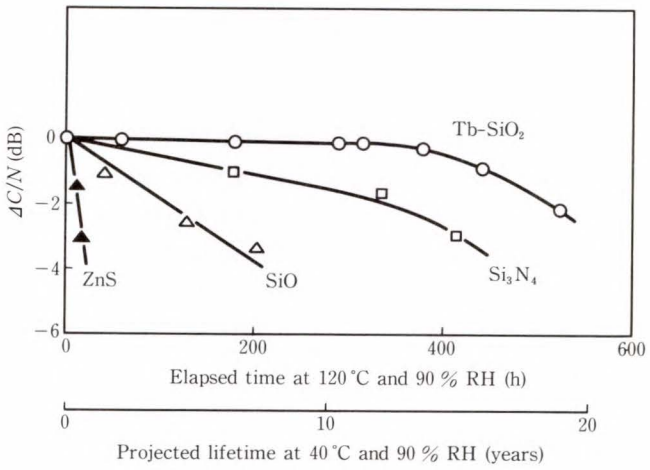


Fig. 14—Results of accelerated lifetime test.

described below was developed.

2) Protective layer

A Tb-SiO_2 protective layer was developed to prevent recording layer oxidation. This stable protective layer uses the characteristics of Tb to oxidize and capture externally intruding oxygen. As Fig. 14 shows, this protective Tb-SiO_2 layer performs much better than other protective layers, and ensures a disk life of ten years or more under normal conditions of use[5].

Because it has a refractive index larger than the substrate material, the protective Tb-SiO_2 layer also improves the signal quality by enabling the rotation angle of the polarization plane due to multiple beam interference to be increased up to 0.7 degrees.

5. Support software

The magneto-optical disk drive subsystem is controlled by software run on a Fujitsu M-series or VP-series CPU.

The Optical Storage Manager/Extensions (OSM/E) is a basic software program that accurately controls optical head movement. It also provides functions for managing disk space and data transfer between the drives and main memory.

The two sides of each disk constitute a "volume", which is identified by a volume serial number. This interface enables disk access by specification of just the volume serial number or dataset name.

Some of the other support software prod-

ucts are: the Automatic Data Migration Facility (ADF), which supervises data migration and recall between the magnetic and optical disk drives, the Distributed System Manager (DSM), which manages optical disk space; and the Resource Access Control Facility (RACF), which provides dataset security. These and other support software products enable construction of a flexible system that meets the needs of the computer operation center.

6. Conclusion

This paper has introduced an economical, high-volume magneto-optical disk drive subsystem that can complement magnetic disk drives.

Because it is a newcomer to the mass-storage device market, the magneto-optical disk drive has several performance constraints. However, progress in elementary techniques for high-performance optical heads and disks is steady, and there is no indication that these problems are insurmountable. It is expected that the magneto-optical disk will be applied to more and more areas of computing and data processing.

References

1) Inoue, Y., Ogawa, Y., and Kaneko, S.: External Storage Devices. (in Japanese), *Fujitsu*, **38**, 5 (Special Issue: Infomation Processing Devices), pp. 347-362 (1987).
2) Tajiri, H., and Abe, S.: THE DEVELOPMENT OF MASS STORAGE SYSTEM IN FUJITSU. 7th IEEE Sympo. Mass Storage Systems, 1985, pp. 94-99.
3) Kaneko, S.: CONSIDERATIONS FOR FUJITSU MASS STORAGE SYSTEMS. Dig. Papers 9th IEEE Sympo. Mass Storage Systems, 1988, pp. 83-87.
4) Arai, S. et al.: Interlinked Tracking Servo Technology *Proc. SPIE*, **695** (Optical Mass Data Storage II), pp. 141-146, (1986).
5) Miyazaki, M. et al.: A new Protective Film Magneto-Optical TbFeCo Media. *J. Appl. Phys.*, **61**, 8, pp. 3326-3328 (1987).

Ryosuke Kudou
Engineering Dept.
File Device Div.
FUJITSU LIMITED
Bachelor of Mechanical Eng.
Chiba University 1972
Specializing in Optical Disk Drive Development

Hiroshi Ichii
Engineering Dept.
File System Div.
FUJITSU LIMITED
Bachelor of Electronics Eng.
Kobe University 1973
Specializing in Optical Disk Controller Development

Akio Futamata
Optical Disks Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Electrical Eng.
Tohoku University 1973
Specializing in Magneto-Optical Disk Drive

UDC 538.975:681.327.634

Thin Film Disk Technology

• Shoji Ishida • Kazuyuki Seki *(Manuscript received June 6, 1990)*

High-density recording requires magnetic recording media with high coercivity, high magnetic saturation, and low noise. This paper reports metal-sputtered thin film disk media developed by Fujitsu to meet these requirements. Beginning with the features of cobalt-based alloys, substrates, sputtering process, and electromagnetic characteristics are described with analyses of the fine surface. Texturing, the protective layer, and lubrication are discussed in conjunction with durability and reliability with experimental results. These metal-sputtered media are installed on the latest Fujitsu large-capacity magnetic disk drives such as the 8-inch, 2.6-Gbyte F6427H subsystem.

1. Introduction

High-density magnetic recording requires magnetic recording media with high coercivity and high-saturation magnetic flux density. To meet these requirements, thin film magnetic media with higher coercivity and higher magnetization than the conventional particulated oxide media have been developed. Above all, thin film media consisting of cobalt-based alloys using the sputtering method (metal-sputtered media) have been put to practical use.

Fujitsu has developed high-coercivity and low-noise metal-sputtered media, which are mounted on large-capacity magnetic disk drives. The drives include the 8-inch Hayabusa (F6427H)[1], the 5.25-inch Humming, the 3.5-inch Picobird. Hayabusa is a high-density, large-capacity magnetic disk drive for large computer systems. The metal-sputtered media and thin film head have been combined to achieve superior drive performance.

This paper reports metal-sputtered media developed by Fujitsu and the technological trends in sputtering process, electromagnetic characteristics, reliability and durability.

2. Features of metal-sputtered media

2.1 General features

This section discusses the developmental background and general features of these metal-sputtered media.

1) There are markets that use many small-diameter media (5 inches or less) such as hard disk drives for personal computers.
2) Manufacturing tools such as the sputtering system have been developed at the same time to mass-produce hundreds of thousands of media a month. The media are approaching conventional particulated oxide media in cost.
3) For small-diameter media, the speed of the head relative to the disk is low and the reliability is easier to improve than for large disks. This made it easier for the media to come onto the market.
4) The sputtering method can select alloy components more easily than the plating method. Corrosion resistance has been improved significantly by adding Cr, Pt or Ta.
5) Because these media have excellent magnetic properties, they are suitable for high-density recording. The coercivity (H_c) is about 1 000 Oe (12.6 A/m). Higher coercivity is being investigated by CoCrTa alloys[2]. Because the residual magnetic flux density is high (8 000 G; 0.8 T), even small-diameter disks with low rotational speed have high reproduced output.
6) The media noise is relatively high. The zig-zag domain wall occurs in the magnetic

transition layer, and the noise increases as the recording density becomes higher[3)-5)]. However, since the reproduced output is high, the media usually has a better overall signal-to-noise ratio.

7) Although the manufacturing process is relatively simple, it is devised in various ways. The manufacturing process features a texturing technique to roughen the substrate surface. The texturing technique prevents head stiction and provides shape anisotropy for the magnetic properties in circumferential orientation[6)-9)].

2.2 Features of Fujitsu metal-sputtered media

1) The magnetic material uses CoCrTa alloys. In optimal process conditions, the coercivity is 1 200-1 400 Oe (15-17.6 A/m), and the product of film thickness and residual magnetic flux density ($t \cdot Mr$) is 400-500 $\mu m \cdot G$.
2) The medium noise has been reduced to half of that of the currently common material CoNiCr.
3) The optimal texture ensures a glide height of 0.08 μm.
4) The protective layer consists of carbon with high-affinity liquid fluorocarbon lubricant applied. Surface texturing roughness and lubricant thickness are optimized to form a low-wear layer free of head stiction.
5) Higher coercivity can be obtained. When coercivity is 1 800 Oe (22.6 A/m) and the spacing is 0.10 μm, high-density recording of 54.7 kFCI (D_{50}) has been confirmed.

3. Process technology

3.1 Outline of the manufacturing process

As shown in Fig. 1, the media have a multi-layered structure. The structure consists of an under layer (chromium), a magnetic layer (cobalt alloy), and a protective layer (carbon). Figure 2 shows the flow of the general manufacturing process. Electroless Ni-P plated aluminum is used for the hard, non-magnetic substrate.

After the plated surface is polished, the surface is intentionally roughened (textured)

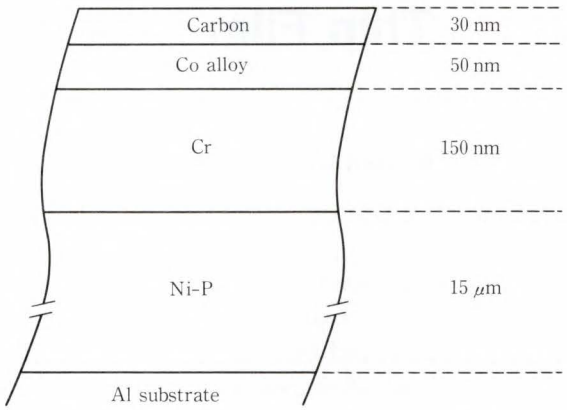


Fig. 1—Structure of metal-sputtered media.

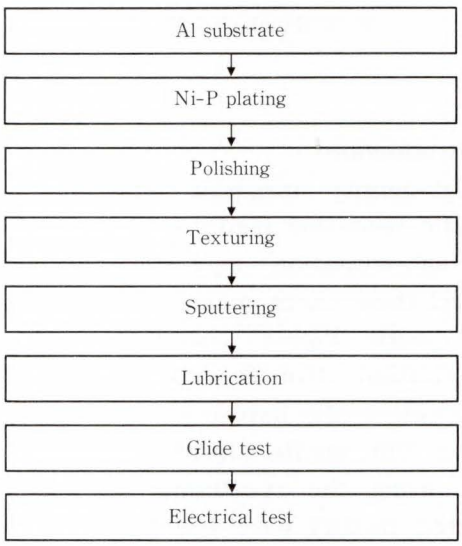


Fig. 2—Flow of the processes for manufacturing metal-sputtered media.

by a lapping tape or free-abrasive procedure. As a result, shape anisotropy is given to the magnetic properties in circumferential orientation. Also head stiction is prevented. However, too much surface roughness causes the head to become unstable so that it touches the media. Therefore the proper roughness must be chosen.

Next, Cr, CoCr alloy, and carbon film is sequentially formed at the same time on both sides by sputtering. The lower layer Cr facilitates the epitaxial growth of Co and helps to provide magnetic anisotropy in the plane of the magnetic layer.

The protective carbon layer has lubricating ability. Cases have increased in which liquid fluorocarbon lubricant is applied to the surface

to improve durability.

The media are finally checked for glide height, electrical characteristics, and media defects.

The manufacturing process is automated. Figure 3 shows part of the line that produces Fujitsu 8-inch metal-sputtered disks.

3.2 Substrate and texturing

The substrate material consists of A5086 alloys or Al of 99.9% or higher purity with 4% Mg added. A very small quantity of Cu and Zn are also added to improve the throwing power of Ni-P electroless plating. The Ni-P plating is normally about 15 μm thick and the Vickers hardness is about 500. The surface is finely polished, and its average surface roughness (R_a) is 2 nm to 3 nm.

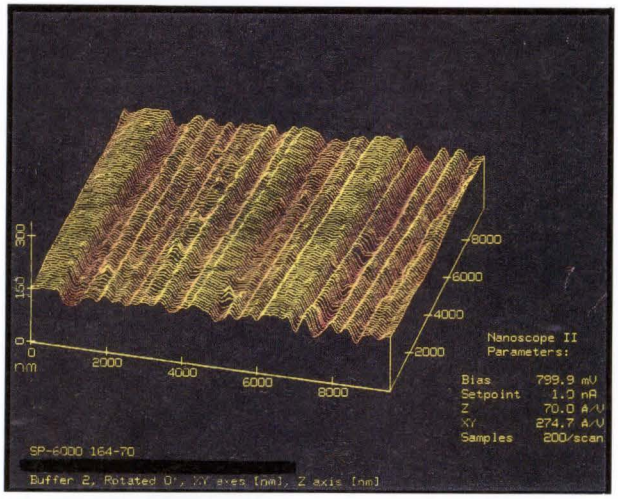
The dynamic mechanical characteristics have also been improved to achieve lower head flying height. The plated Ni-P is finished with a run-out of about 10 μm and an acceleration of about 10 m/s^2.

The plated Ni-P is not magnetic. However, preheating the Ni-P gradually crystallizes the amorphous material and magnetizes it. When the plated Ni-P is heated at 280 °C for one hour, it is magnetized to about 10 G.

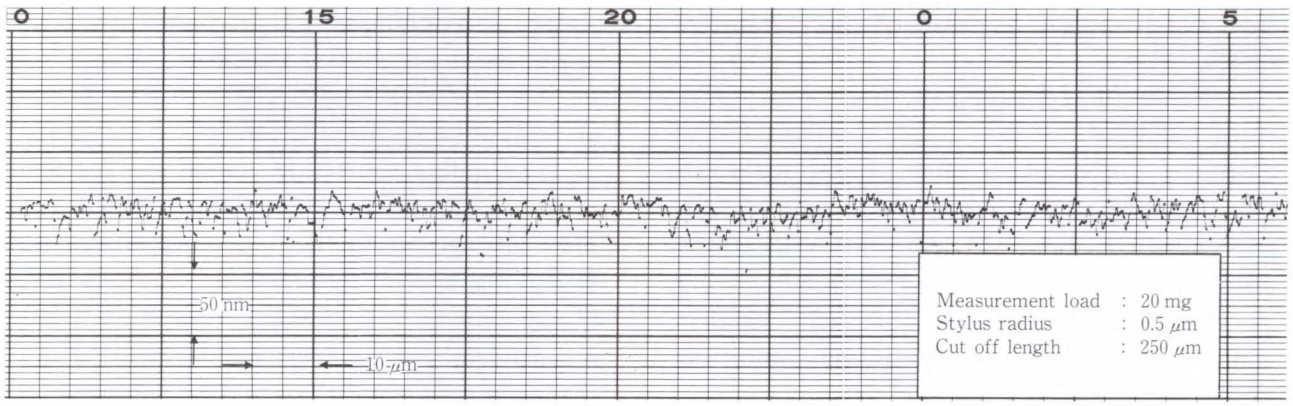
To improve the magnetic properties, such as coercivity and squareness, the substrate preheating temperature for sputtering tends to be raised as much as possible. The magnetization of Ni-P substrate is now the subject of investigation.

Surface treatment called texturing is adopted to reduce the contact area between

Fig. 3—Manufacturing line for metal-sputtered disk.



a) Scanning Tunneling Microscopy (STM)



b) Mechanical stylus

Fig. 4—Texture profiles.

the head slider and the disk surface to decrease the stiction and friction force. Lapping tape or free-abrasive gives circumferential fine-grained roughness on the substrate surface. Figure 4 shows the texturing profile with the mechnical stylus and STM. Although texturing reduces the friction, too much roughness degrades the head flying characteristics. To ensure a head flying height of 0.10 μm, an average surface roughness of 10 nm or less is required with a peak height of 50 nm or less.

Of the metal-sputtered media processing technologies, this texturing technique is the key to determining the head flying characteristics and medium reliability. This texturing technique is complicated, and contains a lot of know-how.

3.3 Sputtering

Cr, Co alloy, and carbon are sequentially sputtered in an Ar atmosphere on both sides of the textured substrate at the same time. Mass-production sputtering normally uses more than one DC magnetron cathode for high-rate deposition. The important parameters for the sputtering conditions include the substrate temperature, Ar pressure, and residual impurity gas pressure. This system is fully automated and has functions such as those to preheat substrates in the loading chamber, and to remove residual gas using large cryo-pumps.

There are two types of sputtering systems. One is the stationary type, used for deposition when the substrate is stationary facing the circular target. The other is the in-line type, used to put many substrates on the rectangular tray and to pass it in front of the rectangular target.

The stationary type is free from cross-contamination because it has an independent chamber configuration. Because no substrate moves during sputtering, better magnetic properties can be obtained with little locational dispersion. The type and pressure of gas can be changed independently for each chamber, and a wide range of sputtering conditions can be selected.

In the in-line type, the relative position of the substrate and target changes momentarily.

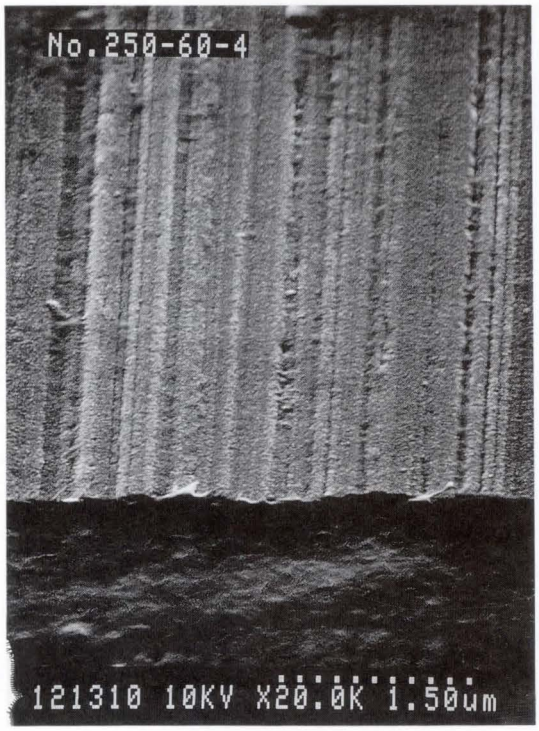


Fig. 5—SEM photograph of medium surface.

The incident angle of the sputtering particles also changes. This system is subjected to oblique-incidence sputter deposition in the direction of substrate conveyance. Uniaxial anisotropy occurs in the conveyance direction, where coercivity and squareness become high. When this uniaxial anisotropy occurs, two-cycle modulation appears in the envelope of the reproduced waveform. However, texturing the substrate creates circumferential anisotropy, and modulation also disappears.

This uniaxial anisotropy created by texturing is considered as circumferential shape anisotropy caused by the shadowing mechanism of the sputtered particles[9),10)]. Figure 5 shows an SEM photograph of the medium surface.

The in-line sputtering system enables simultaneous processing of multiple media and provides high productivity. The mass-production sputtering system becomes increasingly larger in capacity by enhancing the heat and exhaust mechanisms. A sputtering system with a throughput of more than 500 3.5-inch disks per hour has also been developed, allowing high-volume production with reduced cost.

3.4 Protection and lubrication

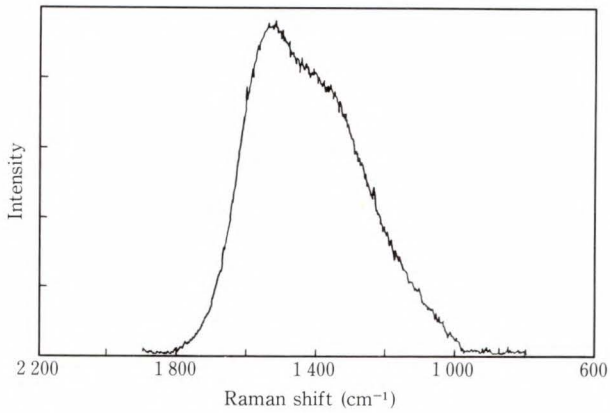
The metal-sputtered media must be protected and lubricated to improve corrosion resistance, to reduce friction, and to improve durability against head contact during start-stop (CSS). Corrosion resistance has been improved by changing the alloys from CoNi based alloys to ternary alloys with superior corrosion resistance such as CoNiCr, CoCrPt and CoCrTa. Generally, carbon has been used as an good protective layer[11)-13)]. The sputtered carbon layer is found not to be crystalline but amorphous through X-ray diffraction analysis and electron-beam diffraction analysis.

However, in laser Raman spectroscopy, the layer is often identified as a structure called diamond-like carbon film. The diamond-like carbon film has hardness, thermal conductivity, and insulation close to diamond, and is crystallographically amorphous. In terms of the electrical structure of the carbon film, however, it is considered as SP^2 distributed in an SP^3 structure. The film with weak SP^3 property and strong SP^2 property is called glassy carbon film, which is more fragile than diamond-like carbon film.

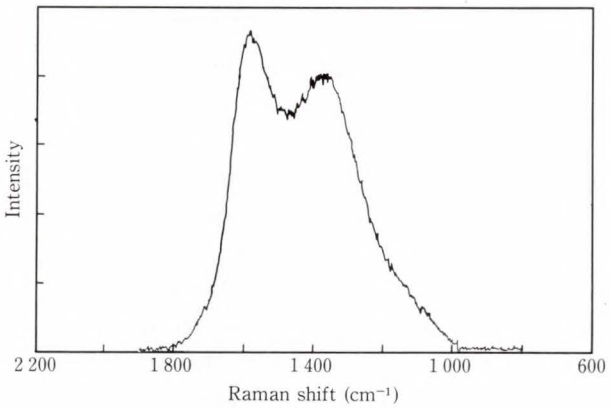
Figure 6 shows the laser Raman spectra of diamond-like carbon film and glassy carbon film. The peak of the spectrum of the diamond-like carbon film is at about 1 530-1 550 cm^{-1}. The diamond-like carbon film has an asymmetrical Raman band with a shoulder band at about 1 400 cm^{-1}. The glassy carbon film has broad Raman lines at about 1 590 cm^{-1} and 1 360 cm^{-1} [14),15)]. A mixture of diamond-like carbon films and glassy carbon films is often produced. When the sputtering temperature is higher, the glassy carbon films are more likely to be produced.

A hydrocarbon layer is considered to form on the surface of carbon film, absorbing moisture and contaminants in the air, and providing lubricating ability. However, if the hydrocarbon layer on the surface is destroyed during continuous head sliding (or CSS), the friction increases and wear may result.

Cases have increased in which liquid fluorocarbon lubricant is applied on the surface



a) Diamond-like carbon film



b) Glassy carbon film

Fig. 6—Laser spectrum of carbon films.

to suppress friction and wear caused by continuous sliding. The liquid fluorocarbon lubricant has a functional end group with a double carbon bond at the end-terminal of the molecule. The thickness of the applied lubricant is normally about 2 nm because the thickness must match the surface roughness created by texturing. Applying too much lubricant increases the stiction force with the head and makes it impossible to start the disk rotating.

Protective ceramic layers such as SiO_2 [16)] and ZrO_2 [17)] are being developed. The protective layers are harder than carbon and their friction coefficients are not changed by continuous CSS.

4. Magnetic properties and electromagnetic characteristics

4.1 Magnetic properties of CoCrTa disk

Cobalt has a high saturation magnetic flux density of 18 kG (1.8 T) and a high

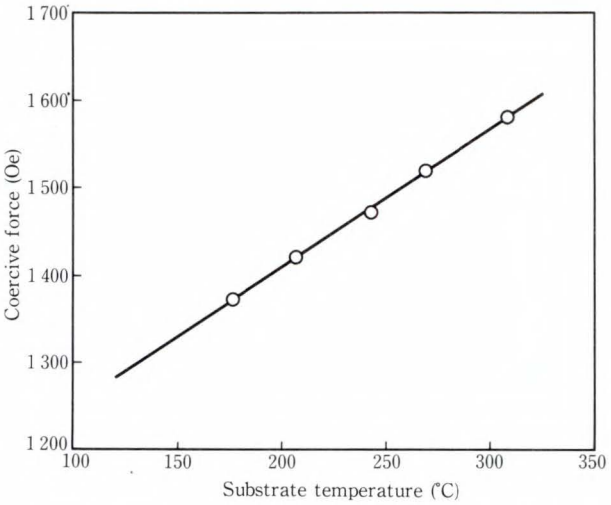


Fig. 7—Substrate temperature and coercive force.

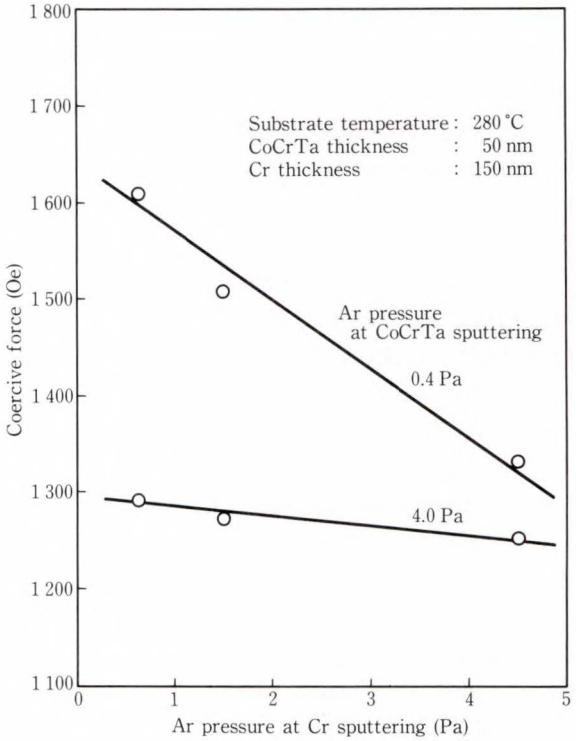


Fig. 8—Ar pressure and coercive force.

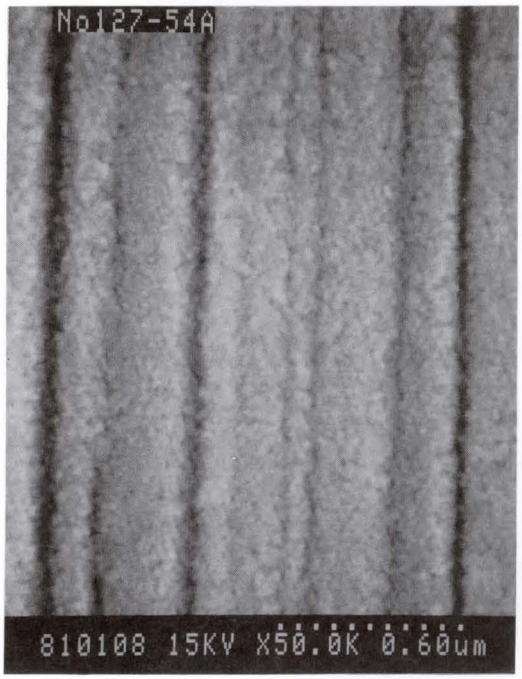
magnetocrystsalline anisotropy energy of 4.2×10^5 J/m^3. The Co-based alloyed thin film produced by sputtering is often a set of polycrystals several tens of nanometers in size. The thin film is close to the diameter of the particles that form the single domain structure. If the axis of easy magnetization (C-axis of the hcp structure) is oriented in the plane of the film, high coercivity and hysteresis characteristics with good squareness can be expected.

From these viewpoints, the alloy components and the sputtering process have been studied. Currently, CoNiCr is being mainly used for metal-sputtered disks with coercivity of up to 1 000 Oe (12.6 A/m). CoCrTa has attracted attention as a higher-coercivity alloy.

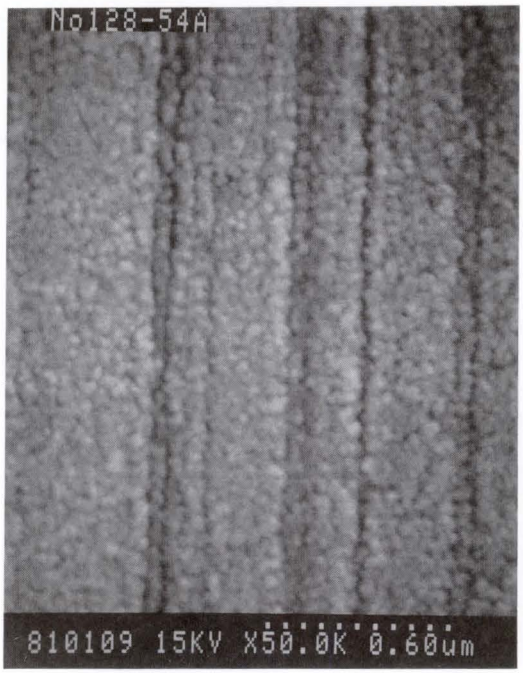
This section discusses the relationship between the sputtering conditions and magnetic properties of CoCrTa. The typical components of Co alloy are $Co_{86}Cr_{12}Ta_2$. Figure 7 shows the relationship between the substrate temperature and coercivity. The substrate temperature is measured immediately before Cr is sputtered using an infrared radiation thermometer. H_c increases in proportion to the substrate temperature ranging from 150 °C to 300 °C.

Figure 8 shows the relationship between the Ar pressure and coercivity H_c. When the Ar pressure for Cr sputtering is lower, H_c becomes higher. When the Ar pressure of CoCrTa is also lower, H_c rises significantly.

Figure 9 shows an FE-SEM photograph of the surface. The Ar pressures for Cr sputtering and CoCrTa are changed for observation. The surface morphology is determined by the gas pressure for Cr sputtering. If the gas pressure for Cr sputtering is low, the crystallite on the surface is small and smooth. If the gas pressure is high, a columnar structure of several tens of nanometers in diameter appears. This is because collisions between the sputtered particles and Ar atoms increases and the mobility of the sputtered particles on the substrate decreases. Because it does not depend on CoCrTa gas pressure, the strong epitaxial property of the lower layer (Cr) may dominate the epitaxial growth of CoCrTa. Figure 10 shows the noise characteristics of media with gas pressure for Cr sputtering. The magnetic properties are the same. When the gas pressure for Cr sputtering is high, the medium noise level is low. This is because when the gas pressure is higher, the grain boundary isolation increases with space generated between the columnar structures.

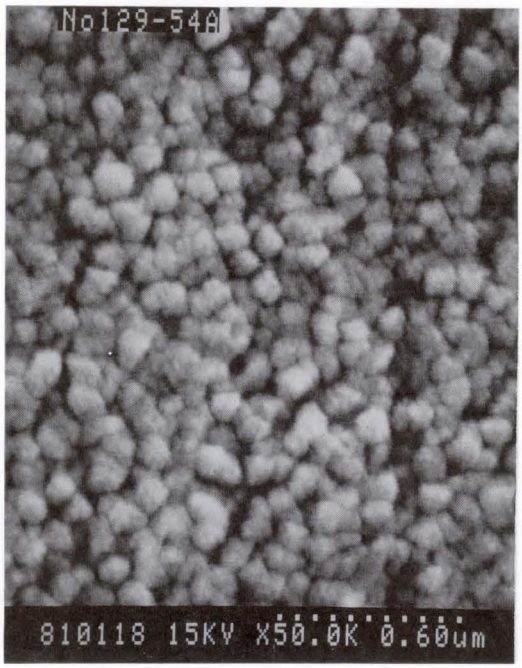


a) Ar at Cr sputtering: 0.4 Pa
Ar at CoCrTa sputtering: 0.4 Pa



c) Ar at Cr sputtering: 0.4 Pa
Ar at CoCrTa sputtering: 4.0 Pa

b) Ar at Cr sputtering: 4.0 Pa
Ar at CoCrTa sputtering: 0.4 Pa



d) Ar at Cr sputtering: 4.0 Pa
Ar at CoCrTa sputtering: 4.0 Pa

Fig. 9—Ar pressure and medium surface taken by SEM.

Figure 11 shows the relationship between the magnetic film thickness and coercivity H_c. As the film thickness increases, H_c decreases monotonically. This may be because the interac-

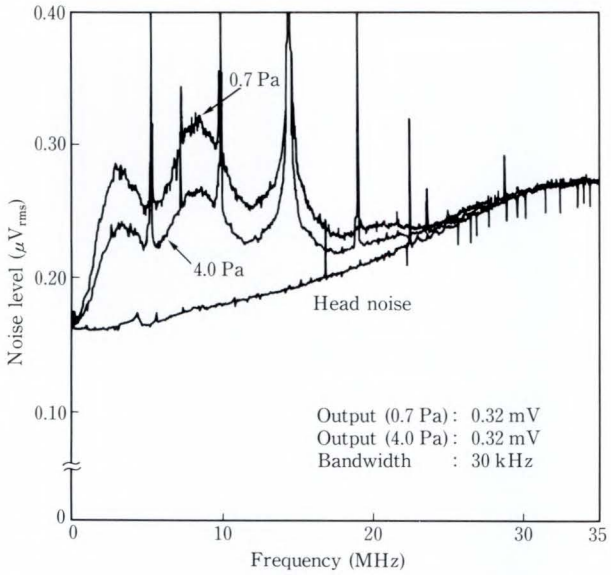


Fig. 10—Ar pressure and medium noise for sputtering Cr.

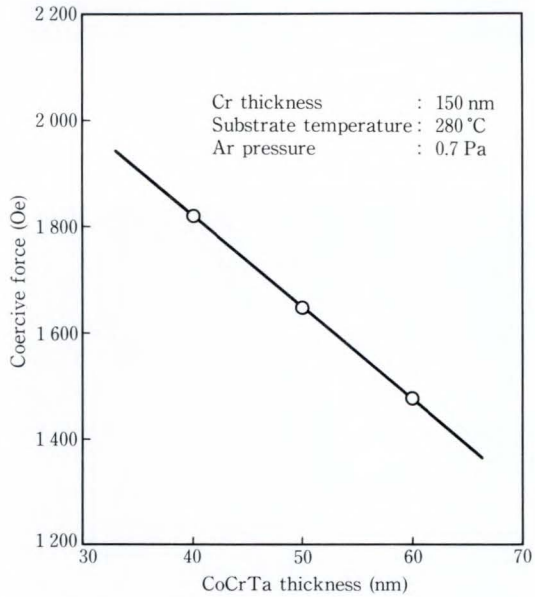


Fig. 11—Magnetic film thickness and coercive force.

tion between the Co particles is strengthened as the film thickness increases.

When a negative DC bias voltage is applied to the substrate during sputtering, H_c is known to become high[18),19)].

Figure 12 shows the effect of the applied DC bias. Applying a proper bias voltage to sputter Cr and CoCrTa increases H_c by hundreds of Oe (several A/m). This may be because the collision between Ar gas and the substrate surface facilitates segregation of Cr. The sputtering conditions have been researched to ensure that coercivity is 1 400 Oe (17.6 A/m) and $t \cdot Mr$ is 500 $\mu m \cdot G$.

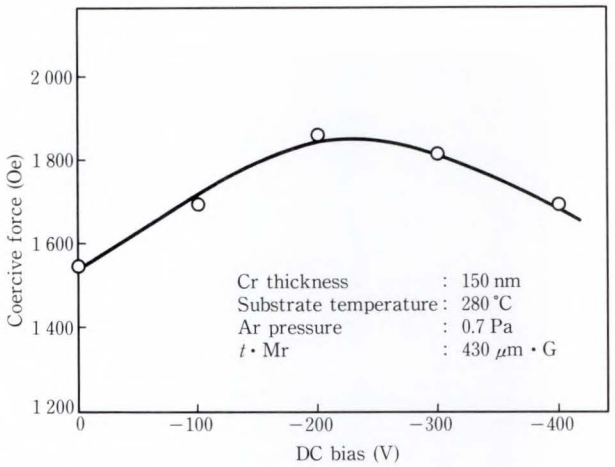
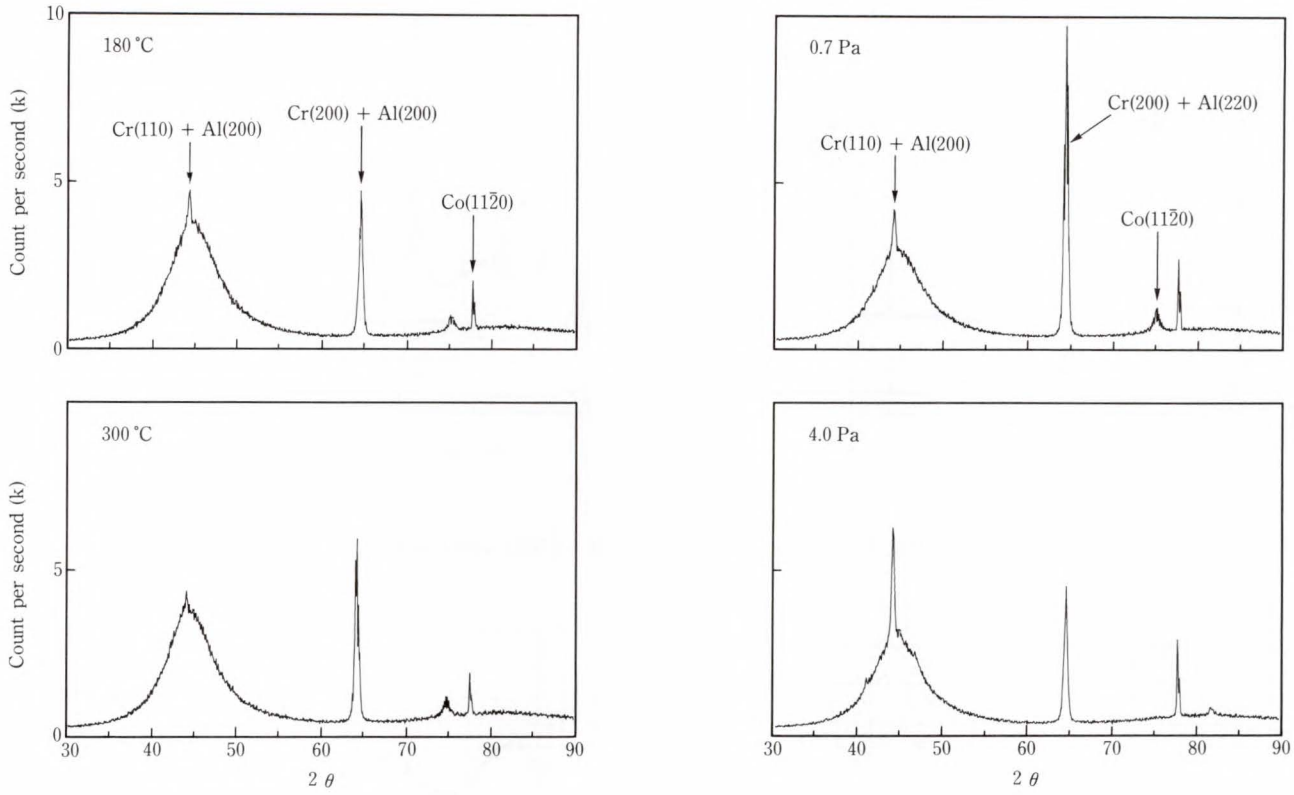


Fig. 12—Substrate DC bias voltage and coercive force.

4.2 Analysis of the crystal structure

X-ray diffraction analysis and electron-beam diffraction analysis are used for the crystallographic study of epitaxial growth of hcp-Co on the bcc-Cr surface[20)-23)]. However a clear understanding of the crystal orientation cannot be obtained due to the difference in the sample produced. The Cr and CoCrTa crystal structures discussed in section 4.1 were analyzed by X-ray diffraction.

Figure 13 shows the diffraction analysis pattern when the substrate temperature and Ar gas pressure are changed. Cr (110), Cr (200), and Co ($11\bar{2}0$) are oriented in parallel on the medium surface. When the temperature is higher and the gas pressure is lower, the orientation of Cr (200) and Co ($11\bar{2}0$) is stronger. At higher temperature and lower gas pressure, hcp-Co is oriented to the C axis resulting in high coercivity in-plane. Cr (200) and Co ($11\bar{2}0$) may be under epitaxial growth because atomic arrays become equal when they overlap at an angle. Strictly speaking, however, the atomic arrays do not match. Thus, the Co layer becomes thick and the epitaxial property may become weak. The size of Cr crystallite obtained from the half-value width of the diffraction peak of Cr (200) is about 35 nm assuming there is no



a) Parameter: Substrate temperature

b) Parameter: Ar gas pressure

Fig. 13—X-ray diffraction pattern of Cr (200 nm) and CoCrTa (50 nm).

stress in the film. When the Ar pressure is high, a columnar structure appears on the Cr layer. The size of the crystallite is rather small (about 25 nm).

On the other hand the CoCrTa crystallite is much smaller (about 13 nm) than the Cr crystallite. To develop media with higher coercivity and lower noise, the relationship between the microstructure and magnetic properties of thin film must be studied.

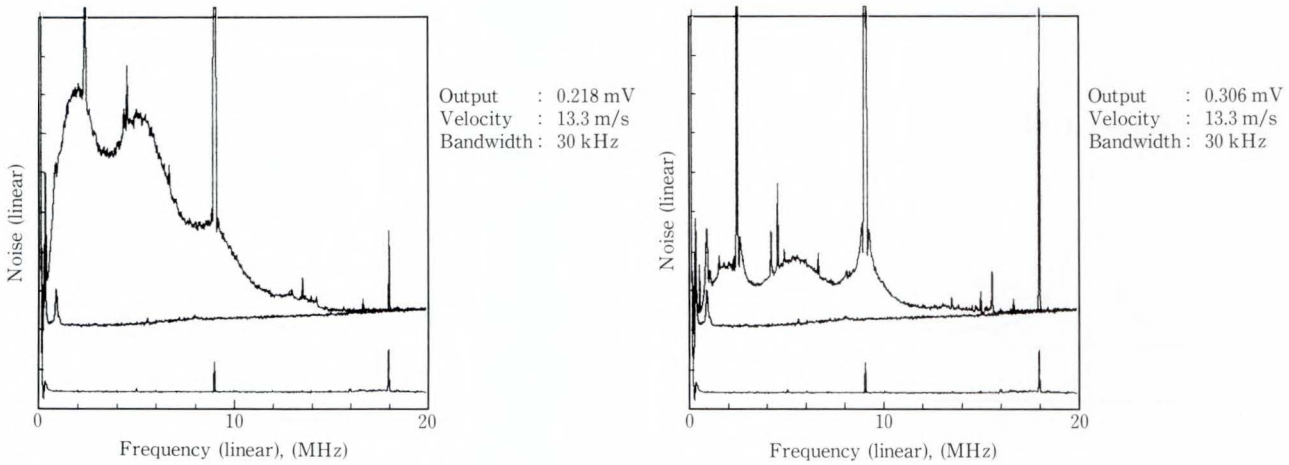
4.3 Electromagnetic characteristics of CoCrTa media

As discussed in section 4.1, the CoCrTa disk can have higher coercivity with a low-medium noise. Figure 14 compares the medium noise with conventional magnetic material, CoNiCr. At a write frequency of 9 MHz (34.4 kFCI) with a noise band-width of 20 MHz, the media noise of CoNiCr is 8.98 μV_{rms}. The total signal-to-noise ratio of CoNiCr is 17.5 dB. The media noise of CoCrTa is 4.92 μV_{rms} and the total signal-to-noise ratio is 23.7 dB. The media noise is reduced by half and the signal-to-noise ratio is improved by about 6 dB.

The reduction of medium noise by adding Ta cannot be clearly analyzed. This may, however, be because segregation of Cr enhances grain boundary isolation and weakens the exchange interaction on the grain boundaries.

Figure 15 lists the typical electromagnetic characteristics of a Hayabusa magnetic disk drive[1)] using a CoCrTa disk and thin film head with a flying height of 0.16 μm. Figure 16 shows the relationship between the phase margin and error rate. Excellent electromagnetic characteristics of the CoCrTa disk contribute to the improved data reliability for the Hayabusa magnetic disk drive.

As discussed in section 4.1, optimized sputtering conditions can ensure a high coercivity of 1 800 Oe (22.6 A/m). Figure 17 shows the recording density characteristics when a CoCrTa disk with a coercivity of 1 800 Oe (22.6 A/m) is combined with the thin film head. The thin film head has a flying height of 0.07 μm (spacing



a) CoNiCr b) CoCrTa

Fig. 14—Noise characteristics of CoNiCr and CoCrTa.

Inner (r 65 mm)					
V_{2F}	Res	O/W	W_{50}	D_{50}	S/N
0.539	73.5	−30.2	35.2	34.3	29.1
mV_{PP}	%	dB	ns	kFCI	dB

Fig. 15—Electromagnetic characteristics of Hayabusa magnetic disk drive.

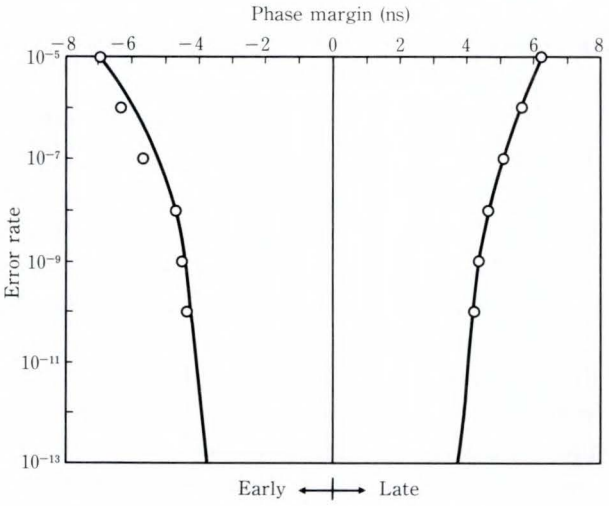


Fig. 16—Error rate for Hayabusa magnetic disk drive.

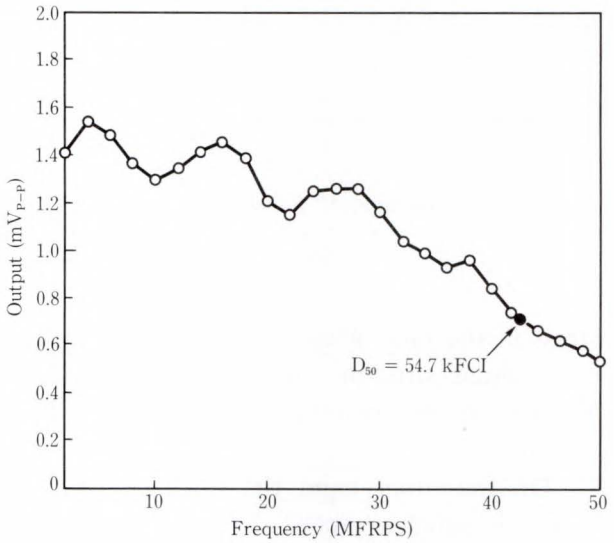


Fig. 17—Recording density characteristics of metal-sputtered disk with 1 800 Oe (22.6 A/m).

of 0.10 μm with protective layer thickness of 30 nm added). The gap is 0.35 μm. High-density recording (D_{50} = 54.7 kFCI) is achieved.

Research to improve the coercivity and reduce the media noise of the metal-sputtered disk has been pursued in earnest. Recently, various reports of noise reduction have been made. Non-magnetic Cr layers are being used to divide the Co magnetic layer into multiple layers to weaken the exchange interaction in the direction of the film thickness[24)-26)]. IBM has combined the high-coercivity metal-sputtered media (H_c: 1 600-1 800 Oe; 20-22.6 A/m) and inductive write head and MR read head. IBM has confirmed that this combination ensures a super-high-density recording of 1 Gbit/sq·inch[27)]. Improvement in the performance of metal-sputtered disk is one of the key technologies for achieving high-density magnetic recording in the future.

5. Reliability and durability

The disk to be mounted in the hard disk drive must be excellent in durability and reliability. The durability and reliability of metal-sputtered disk include: corrosion resistance, friction characteristics, head flying characteristics, and wear durability.

5.1 Corrosion resistance

In the initial stage of metal-sputtered disk development, binary alloy CoNi was used. The disks did not always have good corrosion resistance. Ternary alloys with Cr added such as CoNiCr, CoCrPt, and CoCrTa came to be used, and the corrosion resistance was improved significantly. Figure 18 shows an evaluation of the corrosion resistance of the CoCrTa disk with a carbon protective layer. A change in the magnetic properties and number of defects in the environment of 80 °C, 80 %RH was measured. The magnetic properties did not change and the number of defects did not increase over 2 000 h, and no practical problems occurred.

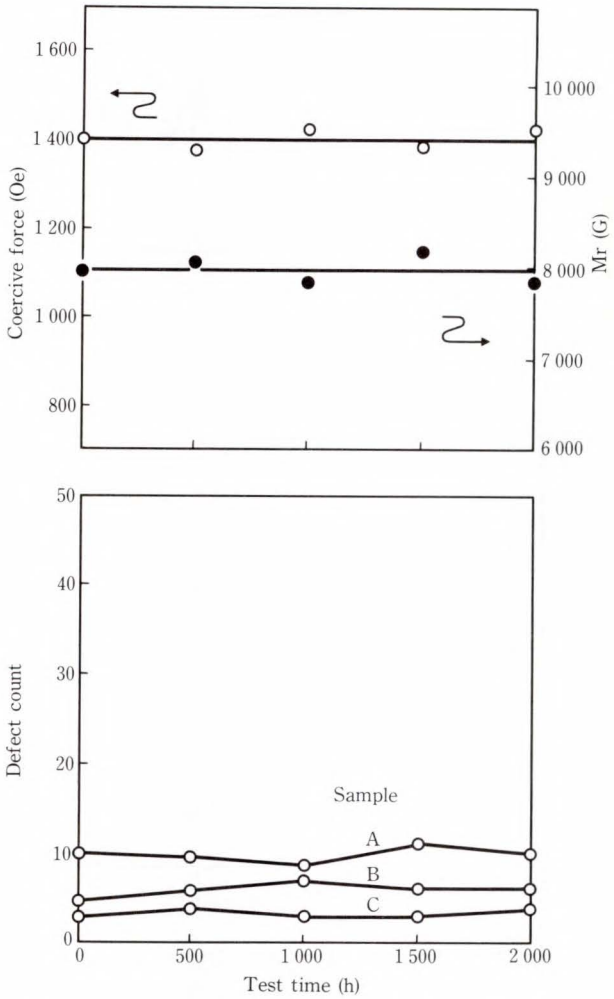


Fig. 18—Corrosion resistance evaluation (80 °C, 80 %RH).

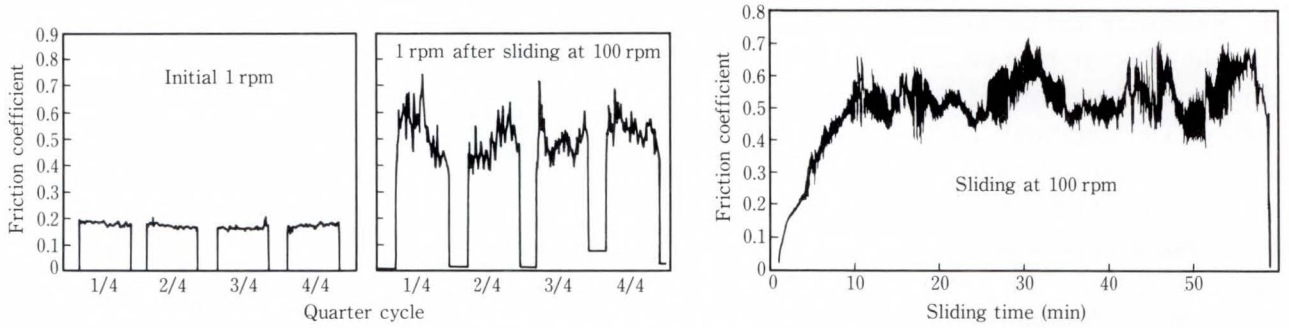
5.2 Friction characteristics

As discussed in section 3.4, the hydrocarbon layer on the surface of the carbon film has lubricating ability. However, head sliding easily destroys the hydrocarbon layer and increases friction. A practical problem is that if CSS is repeated, the friction coefficient will exceed 1. The spindle motor may not start up or the carbon layer may wear and cause the head to crash.

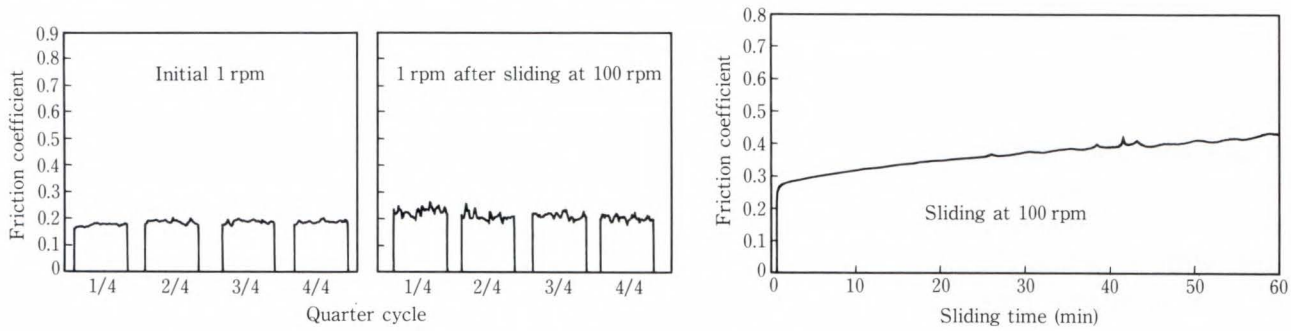
Since lubrication depends on moisture absorption, controling it is difficult and unstable. Therefore, attempts have been made to actively apply lubricant to the carbon surface. Perfluoropolyethers with functional end groups having double carbon bonds at end-terminal of molecule are known as being suitable lubricants for carbon.

Figure 19 shows the difference of change in friction coefficient for continuous constant-speed sliding, with and without lubricant. In Fig. 19, the left chart indicates the static and dynamic friction coefficient for every quater turn at 1 rpm in initial state (before sliding). The right chart indicates the change in the friction coefficient following continuous sliding at 100 rpm for one hour. The middle chart indicates the friction coefficient at 1 rpm, again after continuous sliding. This figure indicates that the increase in friction coefficient due to continuous sliding can be suppressed by applying lubricant. The change in friction characteristics caused by continuous sliding is called the slide-μ characteristic, from which the change in the friction characteristics caused by CSS can be predicted.

Applying too much lubricant causes head stiction when rotation stops. The lubricant thickness has an area for both optimized stiction and friction in relation to surface roughness



a) Without lubricant



b) With lubricant

Fig. 19—Friction changes by continuous sliding.

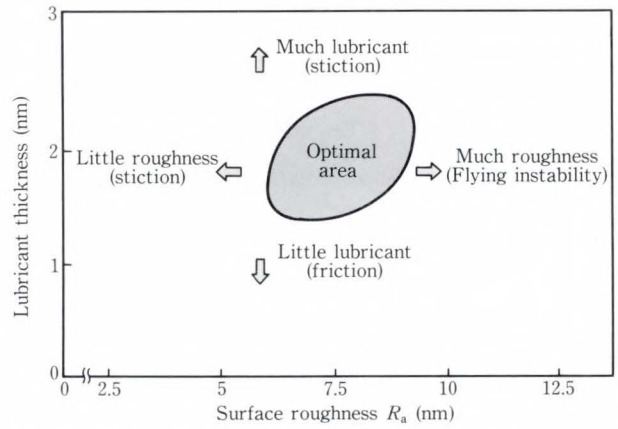


Fig. 20—Optimal area of surface roughness and lubricant.

by texture as shown in Fig. 20. Too much lubricant causes stiction problems. Poor lubricant causes friction problems for sliding. If the surface roughness is too low, even a little lubricant causes stiction. As the surface roughness increases, the lubricant thickness can become larger until stiction is caused. Too much surface roughness causes glide height problems (head-touches).

To satisfy this relationship R_a and the lubricant thickness must strictly be controlled. If the lubricant is applied to the carbon layer, the process also requires the removal of contamination from the carbon surface and lubricant bonding.

5.3 Head flying characteristics

The head flying characteristics are determined mainly by surface roughness created by texturing. Figure 21 shows the take-off flying height measured using a highly sensitive acoustic emission sensor (AE). When the average surface roughness R_a is 8 nm or less, the take-off height is 0.07 μm or less.

Figure 22 shows the result of measuring the take-off over an 8-inch disk. A take-off height of 0.08 μm can be guaranteed for the whole area by making texture uniformly. The take-off height characteristics do not simply depend on the average surface roughness R_a but on the projection height of the surface. To ensure a glide height of 0.08 μm, the maximum projection height must be 50 nm or less. Burnishing is very useful for removing abnormal projections caused by texturing or sputter

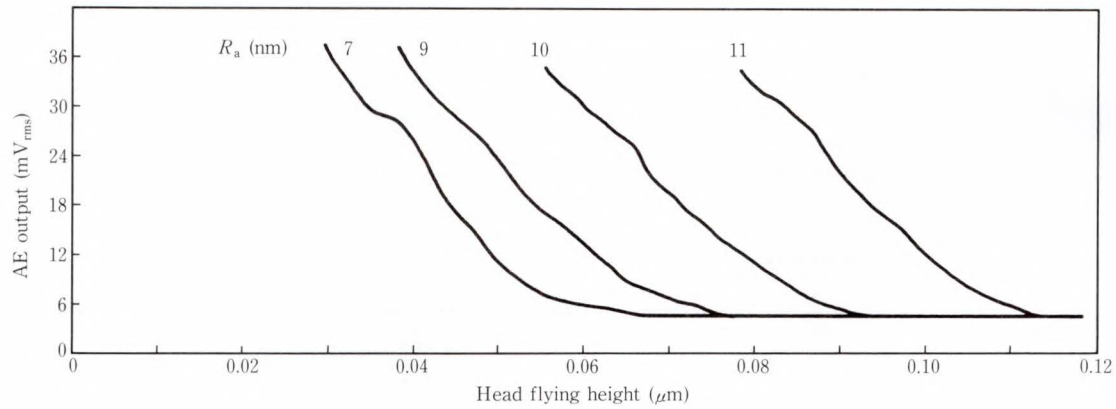


Fig. 21—Surface roughness and take-off flying height.

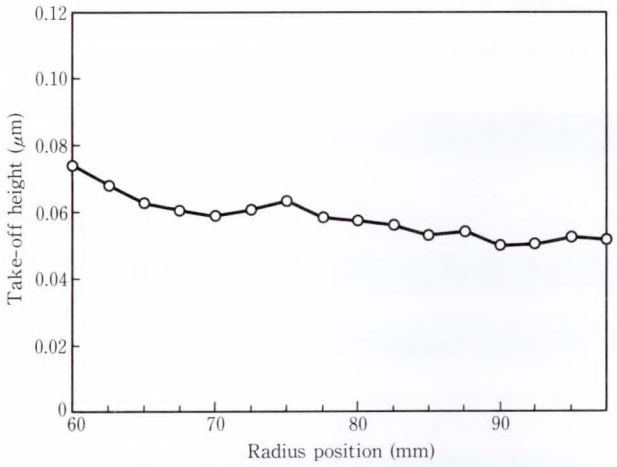


Fig. 22—Take-off height over an 8-inch disk.

deposition and particles adhering to the surface.

The head flying characteristics depend literally on the shape of projections on the surface. To lower the flying height, a compromise must be reached between the friction characteristics and head stiction. Some hard disk drives that incorporate a head load/unload mechanism without CSS are available, and indicate a future possibility.

Instead of aluminum substrate, a chemically strengthened glass substrate and ceramic substrate have been developed as smooth substrates to lower the flying height. This is also attracting considerable interest.

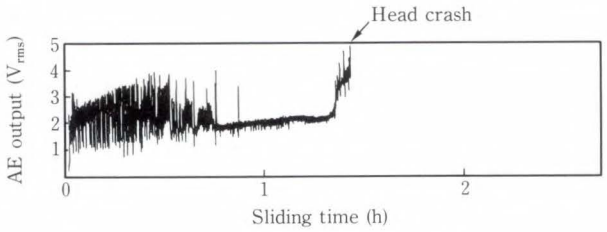
5.4 Wear durability

As an accelerated method for evaluating CSS and the durability against head crashes, a fast sliding test using a high-load head and backward sliding has been used.

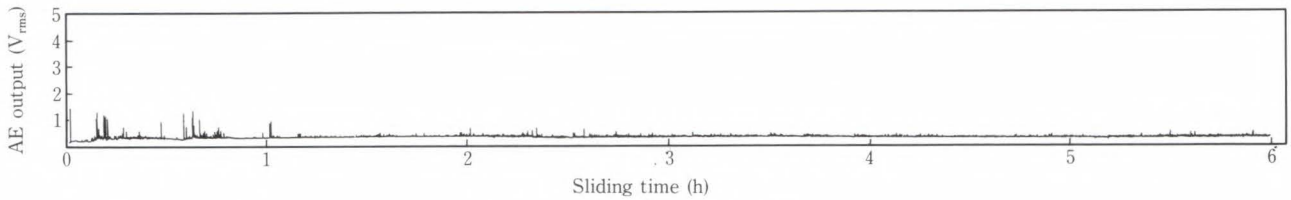
Recently a method for evaluating durability has been proposed[28]. This method uses an actual head slider in low pressure air, reducing the head flying height in near contact to the disk at high-speed.

Figure 23 shows the result of durability evaluation using this depressured sliding method. Contact between the head and disk (head-touch) is detected with a highly sensitive AE. Particulated disks that have been used for a long time are also included for comparison. When sliding under low pressure states, the AE signal is disturbed. However, then AE signal becomes steady. This is because removal of projections from the disk surface, and accumulation on and removal from the head slider are repeated, resulting in stable sliding. When the lubricant is removed from the surface and wear starts, the worn particles accelerate wear by themselves. The AE signal is increased and a clear sliding scratch occurs, leading to head crash. Under the conditions where a head crash occurs on particulated media after about one hour (under low pressure of 2.7 Pa), metal-sputtered media do not crash even after six hours of sliding. Only glossiness appears on the slider, showing excellent durability.

Figure 24 shows the wear depth of the slider and the accumulation of worn particles on the slider side. The metal-sputtered disk used in this figure is a result of optimized texture, sputtering carbon, and lubrication, giving

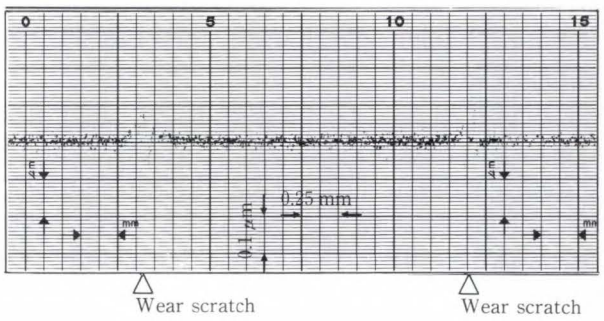


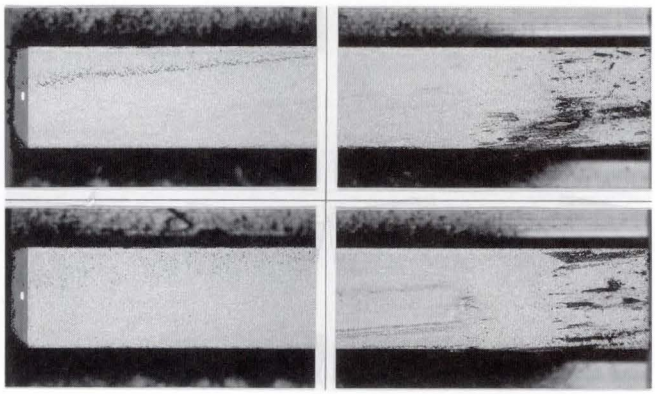
a) Particulated disk



b) Metal-sputtered disk

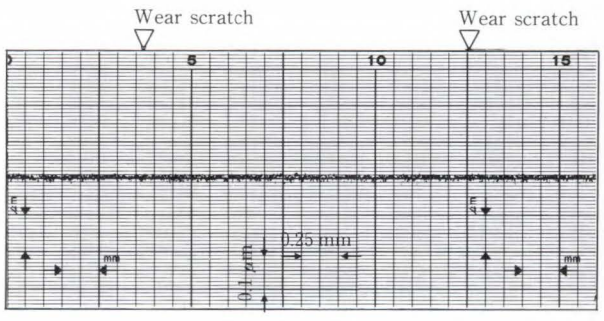
Fig. 23—Evaluation of durability in depressured chamber.

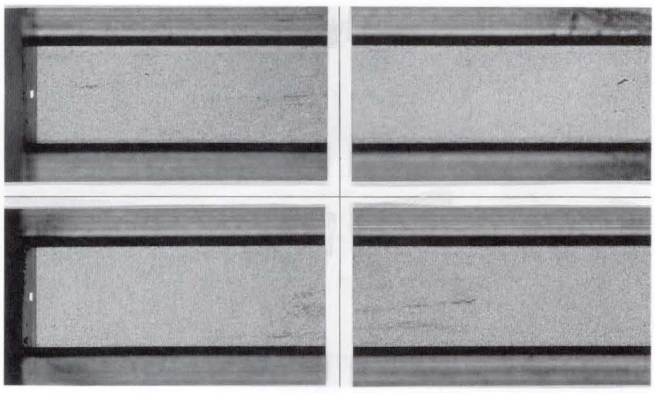




(× 100)

a) Particulated disk





(× 100)

b) Metal-sputtered disk

Fig. 24—Media wear depth and head slider surface.

it high durability. If the carbon layer is brittle or the lubrication is poor, head crash occurs easily, even on metal-sputtered disks.

Figure 25 shows the result of CSS durability of the Hayabusa magnetic disk drive using 8-inch metal-sputtered disk. Even to 20 000 CSS opera-

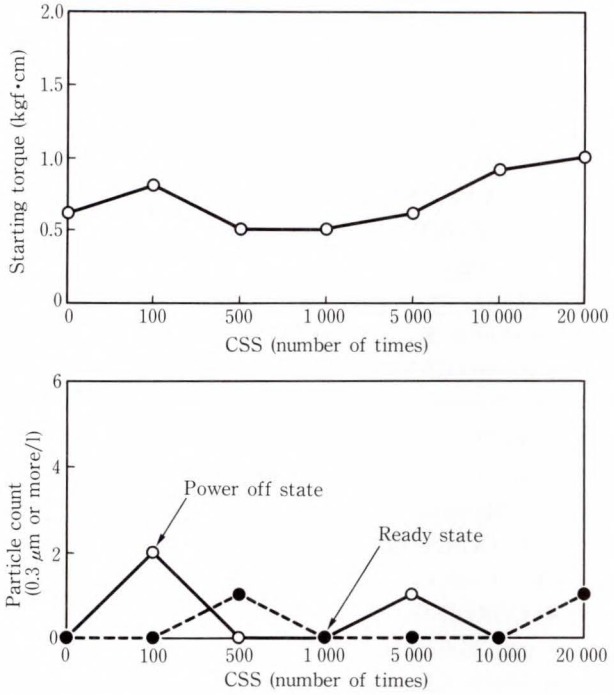


Fig. 25—CSS durability of Hayabusa magnetic disk drive.

tions, the starting torque does not increase and only a few particles of 0.3 μm or more are counted. The durability is very stable. This metal-sputtered disk having excellent durability contributes to the high mean time between failures (MTBF) of the Hayabusa magnetic disk drive.

6. Conclusion

This paper has discussed the technology specific to thin film magnetic disks, one of the key technologies to support high-density, higher capacity magnetic disk drives. Especially, metal-sputtered disks of cobalt based alloys developed by Fujitsu using sputtering have been reported along with recent trends in technology. Disks featuring CoCrTa of higher coercivity and lower noise are assembled onto the 8-inch large-capacity disk drive, Hayabusa.

On further study of the technologies of higher-coercivity, lower media noise, head-disk interface and tribology, the performance of disk drives is expected to increase.

References

1) Koike, T., Negoro, T., and Yoshida, T.: F6427H Magnetic Disk Subsystem: HAYABUSA. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 280-290 (1990).

2) Allan, J.C., and Fisher, R.D.: A COMPARISON OF THE MAGNETIC AND RECORDING PROPERTIES OF SPUTTERED TERNARY ALLOYS FOR HIGH DENSITY APPLICATIONS. *IEEE Trans. Magnetics*, **MAG-23**, 1, pp. 122-124 (1987).

3) Baugh, R.A., Murdock, E.S., and Natarajan, B.R.: MEASUREMENT OF NOISE IN MAGNETIC MEDIA. *IEEE Trans. Magnetics*, **MAG-19**, 5, pp. 1722-1724 (1983).

4) Tong, H.C., Ferrier, R., Chang, P., Tzeng, J., and Parker, K.L.: THE MICROMAGNETICS OF THIN-FILM DISK RECORDING TRACKS. *IEEE Trans. Magnetics*, **MAG-20**, 5, pp. 1831-1833 (1984).

5) Arnoldussen, T.C., and Tong, H.C.: ZIGZAG TRANSITION PROFILES, NOISE, AND CORRELATION STATISTICS IN HIGHLY ORIENTED LONGITUDINAL FILM MEDIA. *IEEE Trans. Magnetics*, **MAG-22**, 5, pp. 889-891 (1986).

6) Rossi, E.M., McDonough, G., Tietze, A., Arnoldussen, T., Brunsch, A., Doss, S., Henneberg, M., Lin, F., Lyn, R., Ting, A., and Trippel, G.: Vaccum-deposited thin-metal-film disk. *J. Appl. Phys.*, **55**, 6, pp. 2254-2256 (1984).

7) Arnoldussen, T.C., Rossi, E.M., Ting, A., Brunsch, A., Schneider, J., and Trippel, G.: OBLIQUELY EVAPORATED IRON-COBALT AND IRON-COBALT-CHROMIUM THIN FILM RECORDING MEDIA. *IEEE Trans. Magnetics*, **MAG-20**, 5, pp. 821-823 (1984).

8) Judge, J.S., and Speliotis, D.E.: THE EFFECT OF TEXTURIZING ON THE MAGNETIC AND RECORDING PROPERTIES OF PLATED RIGID DISKS. *IEEE Trans. Magnetics*, **MAG-23**, 5, pp. 3402-3404 (1987).

9) Simpson, E.M., Narayan, P.B., Swami, G.T.K., and Chao, J.L.: EFFECT OF CIRCUMFERENTIAL TEXTURE ON THE PROPERTIES OF THIN FILM RIGID RECORDING DISKS. *IEEE Trans. Magnetics*, **MAG-23**, 5, pp. 3405-3407 (1987).

10) Teng, E., and Ballard, N.: ANISOTROPY INDUCED SIGNAL WAVEFORM MODULATION OF DC MAGNETRON SPUTTERED THIN FILM DISKS. *IEEE Trans. Magnetics*, **MAG-22**, 5, pp. 579-581 (1986).

11) King F.K.: DATAPOINT THIN FILM MEDIA. *IEEE Trans. Magnetics*, **MAG-17**, 4, pp. 1376-1379 (1981).

12) Kurosawa, H., Mitani, T., and Yonezawa, T.: AP-

PLICATION OF DIAMOND LIKE CARBON FILMS TO METALLIC THIN FILM MAGNETIC RECORDING MEDIA. *IEEE Trans. Magnetics*, **MAG-23**, 5, pp. 2410-2412 (1987).

13) Khan. M.R., Heiman, N., Fisher, R.D., Smith, S., Smallen, M., Hughes, G.F., Veirs, K., Marchon, B., Ogletree, D.F., Salmeron, M., and Siekhaus, W.: CARBON OVERCOAT AND PROCESS DEPENDENCE ON ITS MICROSTRUCTURE AND WEAR CHARACTERISTICS. *IEEE Trans. Magnetics*, **MAG-24**, 6, pp. 2647-2649 (1988).
14) Sato, T., Furuno, S., Iguchi, S., and Hanabusa, M.: Deposition of Diamond-like Carbon Films by Pulsed-Laser Evaporation. *Jpn. J. Appl. Phys.*, **26**, 9, pp.L1487-L1488 (1987).
15) Ramsteiner, M., and Wagner, J.: Resonant Raman scattering of hydrogenated amorphous carbon: Evidence for π-bonded carbon clusters. *Appl. Phys. Lett.*, **51**, 17, pp.1355-1357 (1987).
16) Suganuma, Y., Tanaka, H., Yanagisawa, M., Goto, F., and Hatano, S.: PRODUCTION PROCESS AND HIGH DENSITY RECORDING CHARACTERISTICS OF PLATED DISKS (Invited). *IEEE Trans. Magnetics*, **MAG-18**, 6, pp. 1215-1220 (1982).
17) Yamashita, T., Chen, G.L., Shir, J., and Chen, T.: SPUTTERED Zro_2 OVERCOAT WITH SUPERIOR CORROSION PROTECTION AND MECHANICAL PERFORMANCE IN THIN FILM RIGIT DISK APPLICATION. *IEEE Trans. Magnetics*, **MAG-24**, 6, pp. 2629-2634 (1988).
18) Tani, N., Hashimoto, M., Ishikawa, M., Ohta, Y., Nakamura, K. and Itoh, A.: High Coercivity Force CoNiCr/Cr Sputter Hard Disk. (in Japanese), Inst. Electron. Commun. Engrs., Jpn., **CPM-88**, 92, pp.23-28 (1988).
19) Duan, S.L., Artman, J, O., Lee, J.W., Wong, B., and Laughlin, D.E.: EFFECT OF SPUTTERING CONDITIONS, ANNEALING AND THE MICROSTRUCTURE OF Cr UNDERLAYER ON THE MAGNETIC PROPERTIES OF CoNiCr/Cr THIN FILMS. *IEEE Trans. Magnetics*, **25**, 5, pp. 3884-3886 (1989).
20) Lee, H.J., and Baral, D.: EFFECT OF CRYSTAL TEXTURE ON THE MAGNETIC PROPERTIES OF THIN HCP Co-Ni FILMS. *IEEE Trans. Magnetics.* **MAG-21**, 5, pp. 1477-1479 (1985).
21) Chen, G.L.: NEW LONGITUDINAL RECORDING MEDIA Co_x Ni_y Cr_z FROM HIGH RATE STATIC MAGNETRON SPUTTERING SYSTEM. *IEEE Trans. Magnetics*, **MAG-22**, 5, pp. 334-336 (1986).
22) Ishikawa, M., Tani, N., Yamada, T., Ota, Y., Nakamura, K., and Itoh, A.: FILM STRUCTURE AND MAGNETIC PROPERTIES OF CoNiCr/Cr SPUTTERED THIN FILM. *IEEE Trans. Magnetics*, **MAG-22**, 5, pp. 573-575 (1986).
23) Ohno, T., Shiroishi, Y., Hishiyama, S., Suzuki, H., and Matsuda, Y.: MODULATION AND CRYSTALLOGRAPHIC ORIENTATION OF SPUTTERED CoNi/Cr DISKS FOR LONGITUDINAL RECORDING. *IEEE Trans. Magnetics.* **MAG-23**, 5, pp. 2806-2811 (1987).
24) Murdock, E.S., Natarajan, B.R., and Walmsley, R.G.: NOISE PROPERTIES OF MULTILAYERED Co-ALLOY MAGNETIC RECORDING MEDIA. *IEEE Trans. Magnetics*, **26**, 5, pp. 2700-2705 (1990).
25) Lambert, S.E., Howard, J.K., and Sanders, I.L.: REDUCTION OF MEDIA NOISE IN THIN FILM METAL MEDIA BY LAMINATION. *IEEE Trans. Magnetics*, **26**, 5, pp. 2706-2708 (1990).
26) Hata, H., Hyohno, T., Fukuichi. T., Yabushita, K., Umesaki, M., and Shibata, H.: MAGNETIC AND RECORDING CHARACTERISTICS OF MULTILAYER CoNiCr THIN FILM MEDIA. *IEEE Trans. Magnetics*, **26**, 5, pp. 2709-2711 (1990).
27) Yogi, T., Tsang, C., Nguyen, T.A., Ju, K., Gorman, G.L., Castillo, G.: LONGITUDINAL MEDIA FOR 1 Gb/in^2 AREAL DENSITY. *IEEE Trans. Magnetics*, **26**, 5, pp. 2271-2276 (1990).
28) Terada, A., Ohtani, Y., Kimachi, Y., and Yoshimura, F.: Wear Properties of Lubricated Medium Surface under High Velocity Head Sliding. Tribology and Mechanics of Magnetic Storage Systems, STLE, SP25, pp. 69-71 (1988).

Shoji Ishida
Media Manufacturing Dept.
File Device Div.
FUJITSU LIMITED
Bachelor of Electrochemical Eng.
Yokohama National University 1969
Master of Electrochemical Eng.
Yokohama National University 1971
Specializing in Magnetic Media Technology

Kazuyuki Seki
Media Manufacturing Dept.
File Device Div.
FUJITSU LIMITED
Bachelor of Electrical Eng.
Science University of Tokyo 1975
Specializing in Magnetic Media Technology

UDC 538.975:681.327.634

Thin Film Head Technology

• Mitsumasa Oshiki • Shigemitsu Hamasaki *(Manuscript received June 13, 1990)*

This paper overviews the development and manufacturing of thin film heads in Fujitsu for high-performance magnetic disk drives. The design concepts, including slider and suspension design, are first discussed based on experiments and computer simulations.
A description of the wafer process follows, giving details for photoresist patterning and permalloy plating. Slider machining and assembly are also briefly described.

1. Introduction

Magnetic disk drives have constantly been improved by increasing the recording density, and the data transfer rate, and reducing the device size.

These improvements have been made possible by improvements in the areal recording density, packaging, and signal processing techniques. In particular, areal recording density cannot be increased significantly without improvements in the thin film magnetic recording media and thin film head technologies.

This paper explains the technologies used in fabricating thin film heads for high-performance magnetic disk drives. The emphasis is on technologies for designing and manufacturing thin film heads.

Improvements in the recording density and transfer rate cannot be achieved for conventional bulk ferrite magnetic heads because they have reached the limits of their performance with regard to the machining processes and magnetic characteristics in the high-frequency domain.

Thin film heads have more precise electromagnetic circuitry than conventional heads. The inductance is 20 percent or less that of conventional heads. The heads have excellent high-frequency characteristics and begun to be used for high-performance magnetic disk drives. Thin film heads are manufactured using vacuum deposition, photolithography, and other similar techniques used in semiconductor wafer processing[1].

2. Designing thin film heads for magnetic disk drives

Figure 1 is a conceptual view of the head. Figure 2 is a conceptual cross-section of the magnetic circuitry.

Like conventional monolithic heads, thin film heads have, an electromagnetic transducing element in the slider end that consists of thin magnetic films, thin film conductors, and insulating films.

Because a few hundred to more than one thousand elements are formed on a three- or four-inch wafer at the same time, thin film heads can be mass-produced.

2.1 Characteristics of thin film heads

Compared to the conventional monolithic heads, thin film heads have the following characteristics:

1) Low inductance

Semiconductor manufacturing techniques have reduced the size of the magnetic circuits and coils in the head, reducing the inductance to 20 percent or less that of a conventional monolithic head. The inclusion of a high-speed IC enables high-speed switching of the recording current, producing a sharp magnetic transition layer on the recording medium. These characteristics make the head quite suitable for high-density recording and high-speed data transfer.

In general, increasing the data transfer rate widens the signal frequency band, increasing the impedance noise. Low inductance of the thin

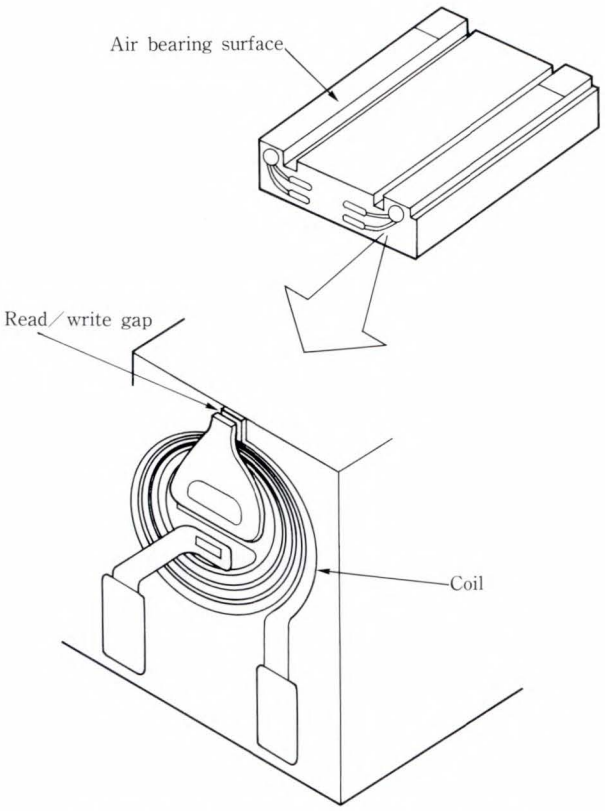


Fig. 1—Conceptual view of thin film head.

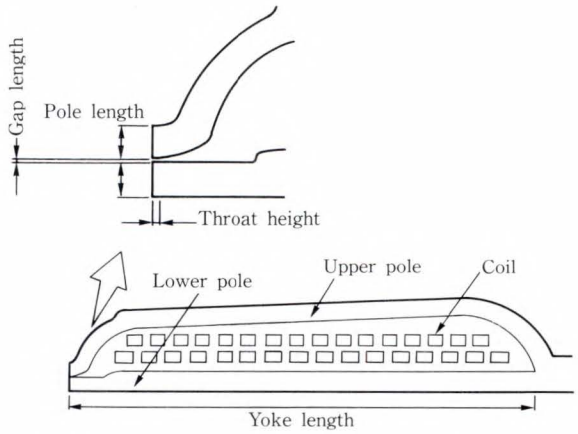
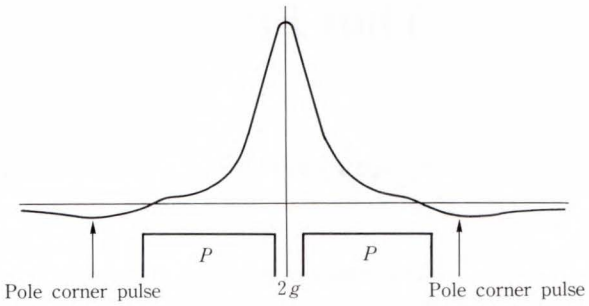


Fig. 2—Cross-section of thin film head magnetic circuitry.

film head, however, reduces head noise, producing an excellent S/N ratio. Fujitsu's thin film head has a low inductance of 600 nH even for coils consisting of 32 turns.

2) Pole corner pulses

The finite length magnetic pole in the thin film head generates reversed polarity pulses around the reproduced main waveform at locations corresponding to the magnetic pole corners, as shown in Fig. 3.



P = Pole length
$2g$ = Gap length

Fig. 3—Isolated waveform and pole corner pulses.

These pole corner pulses greatly influence the pulse width, output voltage, resolution, and peak shifts. They are vital factors in determining the signal quality[2].

Pole corner pulses cause undulations in regenerated output frequency characteristics. Fujitsu sets the magnetic pole length so that the regenerated output reaches a local maximum at the maximum operating frequency.

3) Narrow the track width

Photolithography based on semiconductor manufacturing techniques is used to fabricate thin film heads. Since the core width of the thin film head is not machine-cut as in the conventional monolithic head, a very fine core is easy to produce to 8 μm wide or less.

2.2 Designing thin film heads

2.2.1 Determining electromagnetic characteristics

1) Magnetic circuitry and coil

The cross-section of the magnetic circuitry shown in Fig. 2 was assumed. The authors optimized the design using computer simulation to estimate the magnetic saturation point and magnetic field from the magnetic pole tip. Figure 4 shows an example of simulation.

Increasing the number of coil turns gives a higher reproduced output. This, however, increases both the electrical resistance and inductance, causing the total S/N ratio to saturate, as shown in Fig. 5.

Increasing the number of coil turns causes a

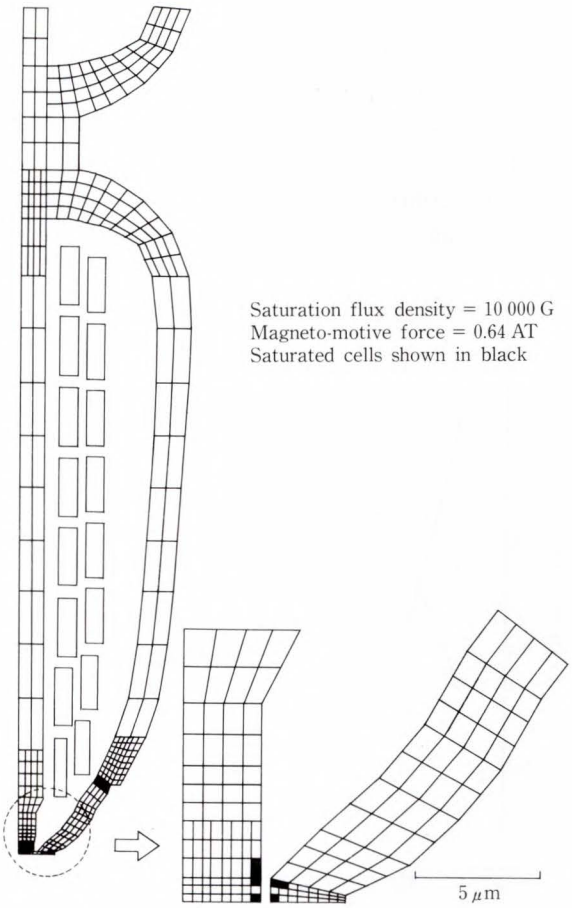


Fig. 4—Magnetic saturation in pole tip (computer simulation).

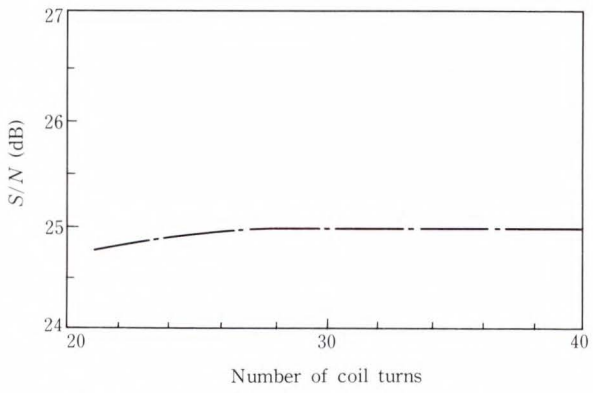


Fig. 5—Number of coil turns and S/N ratio.

longer magnetic yoke, because a finer coil is difficult to produce due to the limitations of wafer processing. Read/write efficiency is reduced for a longer yoke length, as shown in Fig. 6. Therefore, the design must be optimized taking the total S/N ratio and wafer processing limitations into account.

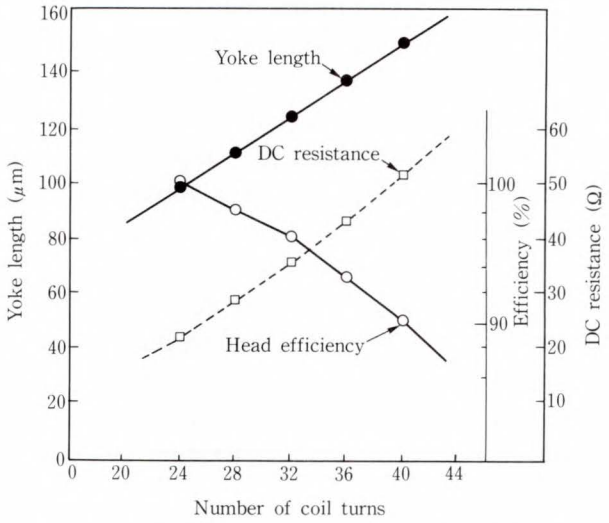


Fig. 6—Number of coil turns and head parameters.

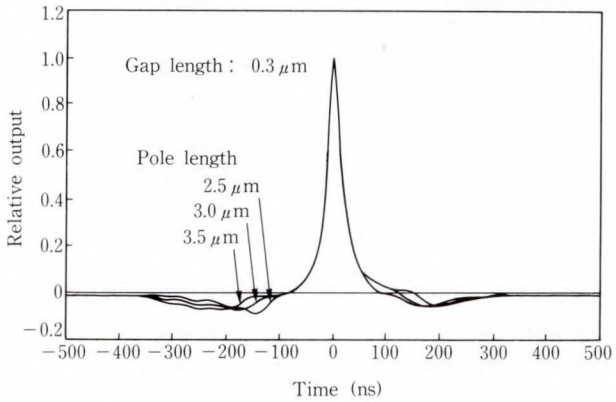


Fig. 7—Pole length and isolated waveform.

2) Magnetic pole length

As mentioned in section 2.1, pole corner pulses occur at magnetic pole corners. Figure 7 shows the results of simulation indicating that the isolated waveform varies with the magnetic pole length. Pole corner pulse locations are also found to vary with the length of the magnetic pole. Since the pole corner pulse causes a peak shift to vary for the recorded data pattern, the magnetic pole length must be correctly determined in relation to the recorded data pattern[3].

The write performance is better if the magnetic poles are thicker and the magnetic field from the tip of the head is higher. To optimize the pole length, the authors simulated a waveform using various recording media parameters and head parameters.

3) Gap length and throat height

The gap length and throat height are other important parameters determining the electromagnetic transducing characteristics. Reducing the gap length makes the regenerated signal sharper, but degrades the recording performance.

If the throad height is too large, a satisfactory magnetic field is not generated and the overwrite characteristics and regeneration efficiency deteriorate. Conversely, if it is too small, excessively high magnetic fields are generated, causing recording demagnetization and reduced output and resolution.

Using the waveform simulation, the authors evaluated a prototype to obtain optimum values for the gap length and throat height.

Control of the gap length and the throat height is extremely important to stabilize the characteristics of thin film heads. Fujitsu sets the nominal tolerance for the gap length to within ±0.05 μm and the throad height to within ±1 μm.

4) Flying height

The flying height greatly influences signal quality and reliability. Signal quality improves as the flying height decreases with reduced spacing loss. However, the possibility for head-to-disk interference increases.

The flying height for heads with high-performance magnetic disk drives should be determined so as to achieve simultaneously high signal quality and high reliability. The disk drives will be fully evaluated to ensure sufficient reliability.

5) Results of evaluation

The major electromagnetic characteristics to be evaluated are the dependence of regenerated output on the current (I/V), the overwrite characteristics (O/W), the frequency dependence, and the regenerated isolated waveform.

Figure 8 shows the I/V and O/W characteristics of the thin film head developed for the HAYABUSA (F6427H) disk drive[4]. The I/V characteristics showed excellent results, free of recording demagnetization.

The O/W characteristics are also sufficient with −30 dB at the specified write current of 30 mA.

Figure 9 shows the dependence of regenerated output on frequency. The pole corner pulse causes undulations as predicted. D_{50} is 37 kFCI.

The isolated waveform shown in Fig. 10 closely matches the simulated waveform in

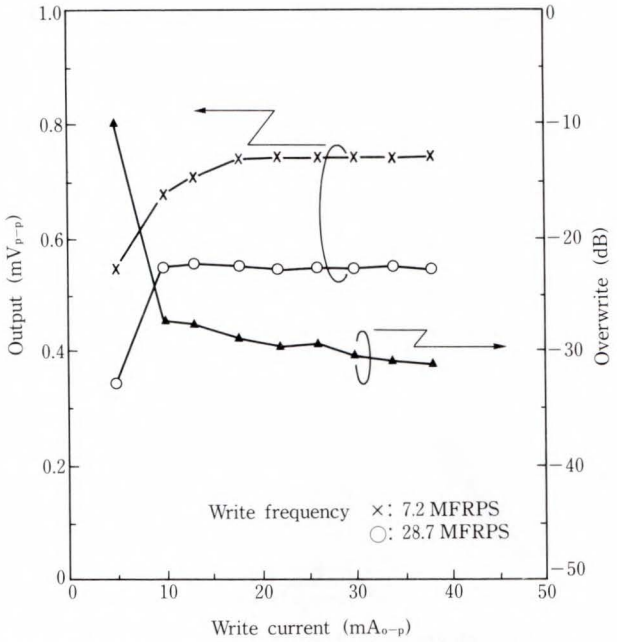


Fig. 8—Write current dependence of regenerated output and overwrite.

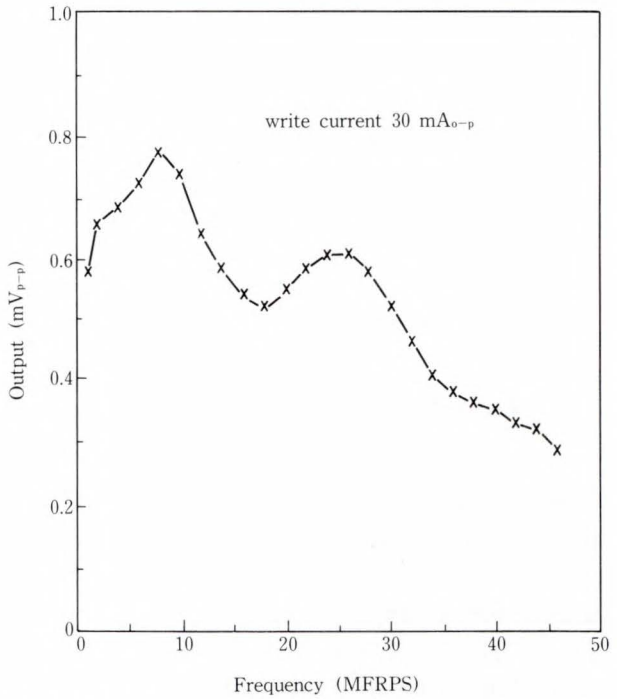
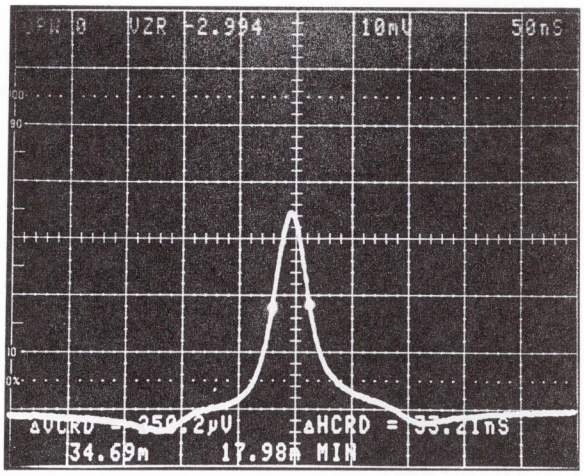


Fig. 9—Frequency dependence of regenerated output.

Fig. 7 (pole length: 3.0 μm, and gap length: 0.3 μm).

Of the thin film head evaluation items, popcorn noise is very important in addition to the electromagnetic characteristics. Popcorn noise occurs when data is read immediately after write operation. It is attributable to the unstable magnetic domain structure of the magnetic yoke[5]. Figure 11 shows an observed example of popcorn waveform.

The authors clarified the relationship between the error rate and the popcorn noise counts and level, and succeeded in developing wafer processing that suppresses popcorn noise.



1.0 MFRPS W_{50} = 33.2 ns (0.99 μm)

Fig. 10—Regenerated isolated waveform.

2.2.2 Flying stability

1) Designing the slider

The slider must be designed to fly stably over the recording medium to assure high reliability. As shown in Fig. 12, as the disk drives are becoming smaller, the sliders are also becoming smaller. The damage that the recording medium suffers from the head touching is proportional to the kinetic energy of the head. If the slider is light, the collision energy is low and the medium is less seriously damaged.

Manufacturing small, high-precision sliders requires high-precision manufacturing techniques. Fujitsu has worked to manufacture sliders by balancing wafer processing and machining techniques. If the early Winchester head slider is assumed to have a mass of 100 percent, the mass is 30 percent for the current industry standard slider and 12 percent for Fujitsu's latest production model.

2) Designing the suspension

Fujitsu has developed inline head suspension (see Fig. 13) suitable for high-speed seek mo-

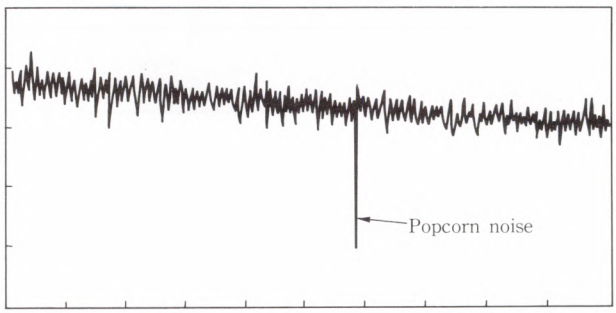


Fig. 11—Popcorn noise.

a) Winchester slider b) Standard thin film head slider c) Fujitsu thin film head slider

Fig. 12—Changes in slider shape.

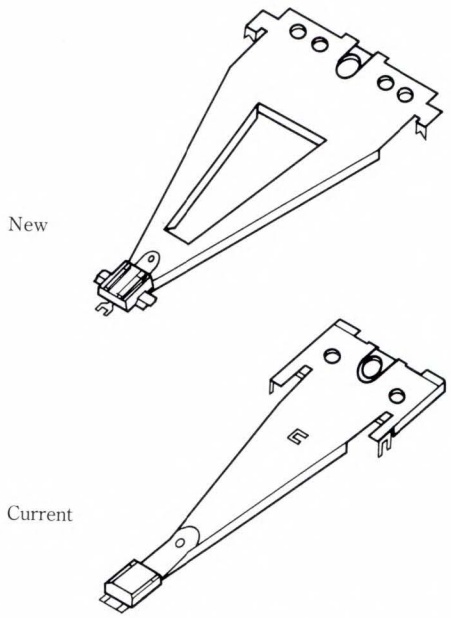


Fig. 13—Newly developed inline head assembly.

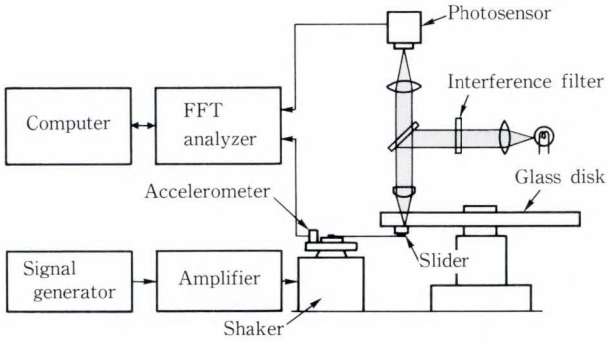


Fig. 14—System for measuring the head flying height.

tion, which is explained in a separate article in this issue[6].

3) Evaluating the flying characteristics

Evaluating the flying height and posture of the head is important for signal quality and device reliability.

The flying charcteristics are evaluated by flying the head on a rotating glass disk at normal speed. Figure 14 is a block diagram of a measurement system developed by Fujitsu. The flying height is measured by ultraviolet light interference. This provides measurement with greater precision than conventional methods using visible light[7].

The head is vibrated vertically or radially (in the seek direction) over the glass disk so that fluctuations in the flying height can also be measured.

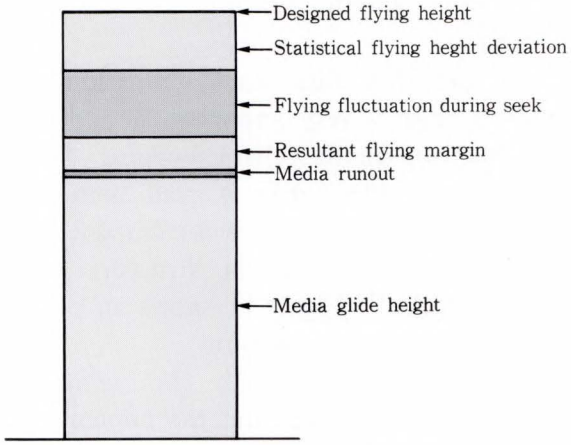


Fig. 15—Flying margin budget.

To test device reliability, the head flying margin is also examined. The flying margin is calculated by subtracting the following factors from the flying height: the static flying tolerance, media glide height, media runout, and flying height fluctuation due to head seek. After studying the margin, the head and recording media allowing a sufficiently positive flying margin are finally designed (see Fig. 15).

3. Thin film head manufacturing

The process for manufacturing thin film heads can be divided into three stages (see Fig. 16).

1) Wafer processing, in which thin magnetic films, thin film conductors and insulation layers are formed in prescribed shapes on the substrate.
2) Machining, in which the wafer is shaped into a slider that floats at a certain distance above the recording medium.
3) Assembly, in which the slider is mounted on a support spring to keep it on the medium.

3.1 Wafer process

3.1.1 Wafer processing

Figure 17 outlines the wafer process. Al_2O_3 film is formed by sputtering on an alumina-based ceramic substrate. The film is completed by mirror finishing. A seed layer is then formed on the board in a vacuum so that permalloy can be electroplated for the lower magnetic pole.

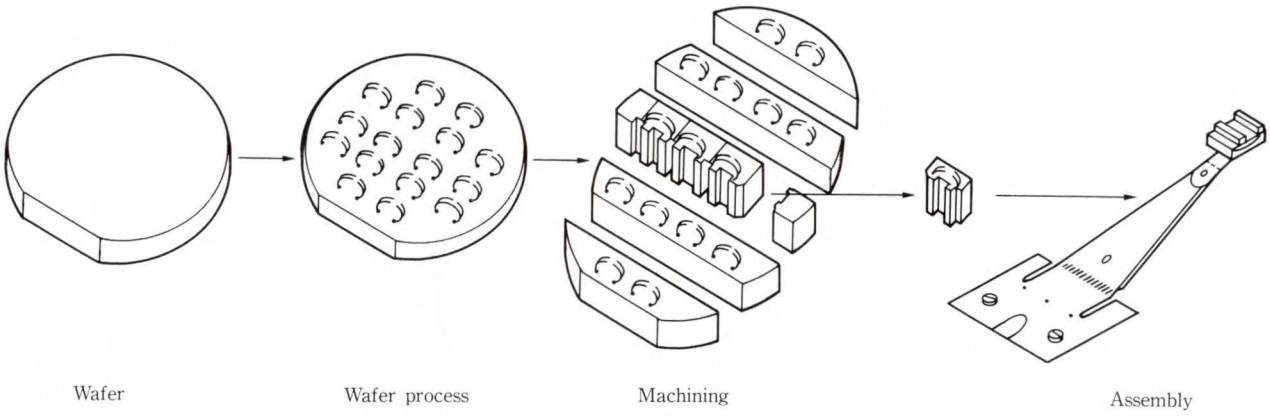


Fig. 16—General process for manufacturing thin film head.

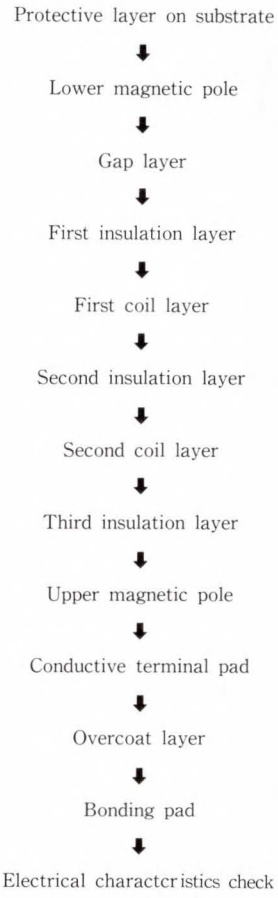
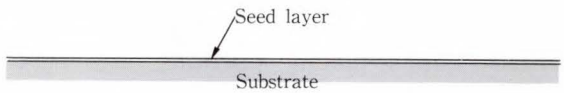
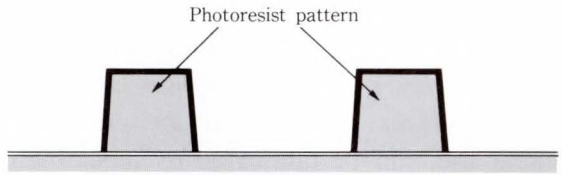


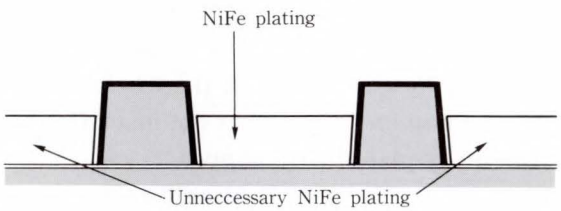
Fig. 17—Wafer processing flow.



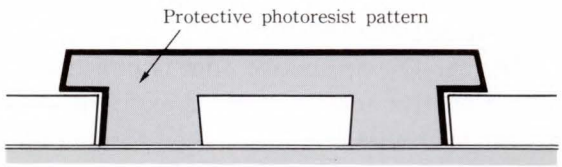
a) Seed layer deposition



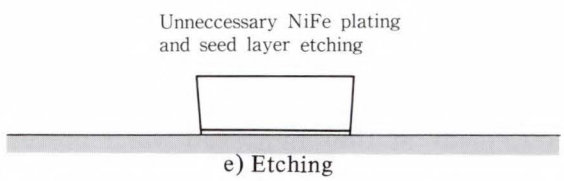
b) Photoresist patterning



c) Electroplating



d) Photoresist repatterning



e) Etching

Fig. 18—Magnetic pole formation.

A photoresist pattern is fabricated using photolithography and a permalloy film is selectively plated to the specified thickness to form a magnetic pole. Unnecessary portions of the film are removed by chemical etching to finish the magnetic pole[8]. Figure 18 shows this process.

Flux enhancement and upper magnetic layers are formed in the same way. Magnetic

pole formation is discussed later in detail.

After a lower magnetic pole is prepared, an Al_2O_3 film is sputtered to provide a magnetic gap. An etched hole is prepared at the back end of the lower magnetic pole to touch the upper magnetic pole.

An insulation layer is formed on the Al_2O_3 gap layer to level the height differences in the lower magnetic pole. The insulation layer is formed by positive photoresist, exposing, developing and shaping the photoresist into the specified form, and is cured at about 200 °C.

After the first insulation layer has been made, a conductive seed layer is deposited in a vacuum to fabricate a copper-plated coil. Unwanted copper-plated portions are covered with a photoresist pattern for selective copper plating. After plating, the seed layer is removed by ion milling to complete coil formation. Figure 19 shows this process.

The insulation layer on the coil is formed in the same way as other insulation layers. Two- and three-layer coils, which are currently in general use, can be fabricated by repeating coil and insulation layer formation the required number of times.

After the coil and insulation layers have been fabricated, the upper magnetic pole is formed in the same way as the lower magnetic pole. The conductive pad for the insulation layer is selectively plated with copper.

An Al_2O_3 layer is sputtered to protect the thin film head elements during machining. Because elements rise 20 μm to 30 μm above the substrate surface, bias sputtering is required to offset the differences[9].

The Al_2O_3 protective film is polished until the conductive terminal pad is exposed. A gold pad for lead wire bonding is then formed on top of the exposed conductive pad.

Together with intermediate inspection, inductance and electric resistance are finally measured to complete thin film processes.

3.1.2 Photoresist patterning

The most important technique for forming an element in thin film head wafer process is photoresist patterning. The core width of the magnetic poles and the coil laminating must be controlled with high precision.

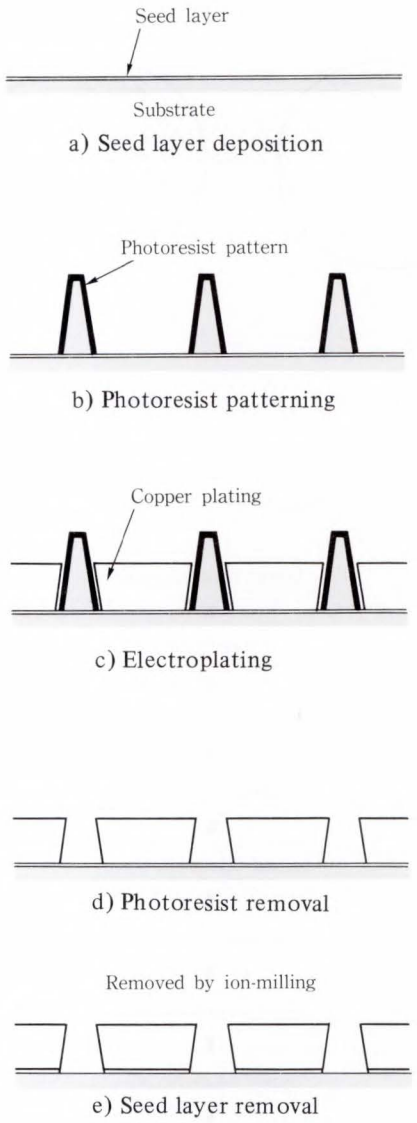
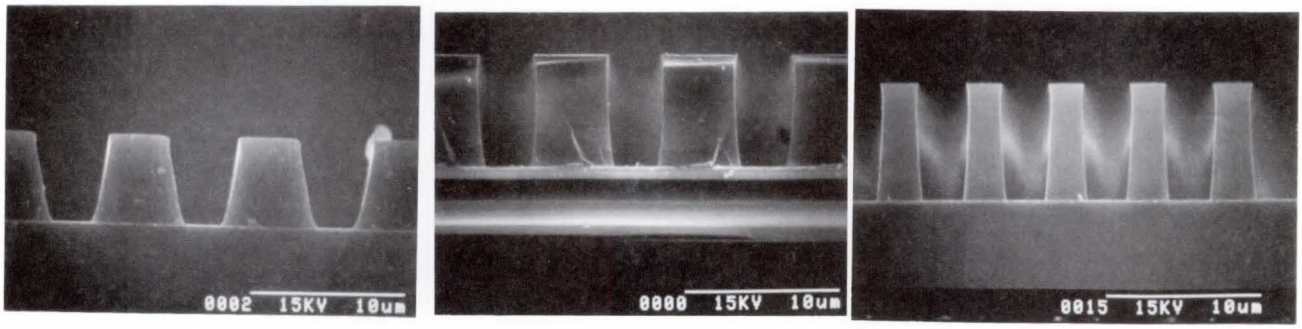


Fig. 19—Coil layer formation.

Proximity exposure, generally used to expose photoresist in thin film head manufacturing, has gradually been replaced by projection exposure because coils are now made of multiple layers with very fine patterns. In projection exposure, an image of the photomask is focused on the surface of the photoresist, with no danger of mask collison. By selecting the focusing point, a sharp cross sectional pattern can be created.

Figure 20 shows a cross-section of the three photoresist patterns for forming a coil by projection exposure.



a) Ordinary photoresist pattern b) CEL used c) New high-aspect resist pattern

Fig. 20—Cross-section of photoresist by projection exposure.

1) Ordinary photoresist pattern

This is the profile of a 5.5 μm thick pattern using an ordinary positive photoresist. The cross section of the pattern is tapered toward the substrate surface and has an angle of about 80°. When the photoresist pattern is used to plate a coil, the cross-sectional area decreases by about 20 percent compared to ideal rectangular cross section, and the coil resistance increases accordingly.

2) Contrast enhanced layer (CEL)

The pattern profile when the CEL is applied to the top surface of the photoresist is then exposed for development, forming a photoresist pattern with an angle of 90°. The CEL thus helps in forming a good pattern profile.

3) New high-aspect resist pattern

Based on the foregoing results, the authors designed a new photomask for coils. Figure 20 c) shows the result of patterning 8 μm-thick photoresist, verifying that 3 μm lines and spaces can be formed easily with a wall angle of 90°.

This technique makes it easy to form a coil with a cross section of 3 x 5 μm^2. Coils with 32 turns and a resistance of 20 Ω or less can be manufactured reliably, providing high performance and a good S/N ratio.

3.1.3 Magnetic pole formation

Plated permalloy is currently used for magnetic poles because it has good magnetic characteristics with high magnetic permeability and saturation flux density, and it has good productivity when used to form films.

The process involves a plating bath with Fe ions added to Watt plating solution, which is widely used in nickel plating[10]. The alloy composition tends to vary when plating an alloy such as permalloy. The plating conditions are therefore determined taking into account the fluctuations in chemical blends and control of the environment. Such conditions were set to minimize the fluctuations in the alloy composition at any portion of the upper magnetic pole with 20-μm undulation.

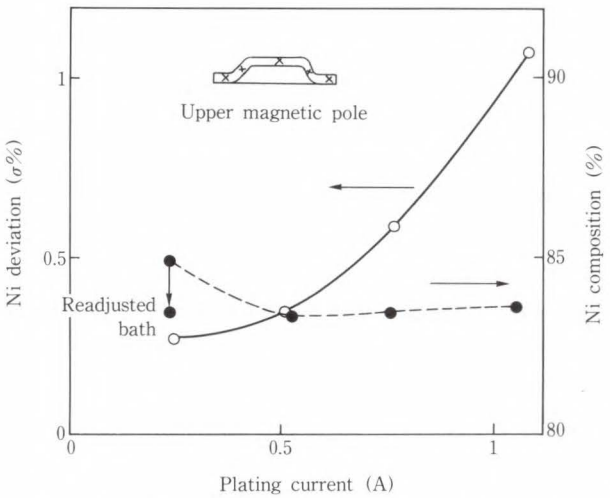


Fig. 21—Plating current and permalloy composition of upper magnetic pole.

Figure 21 shows the dependence of the composition (Ni%) of the upper magnetic pole and its deviation (σ%) versus the plating current. As the figure shows, the composition depends little on the plating current. However, deviation quickly increases when a high current is used to increase the plating rate to improve productivity.

The composition distribution of permalloy is known as allowing the irreversible displacement of the magnetic domain walls and causing

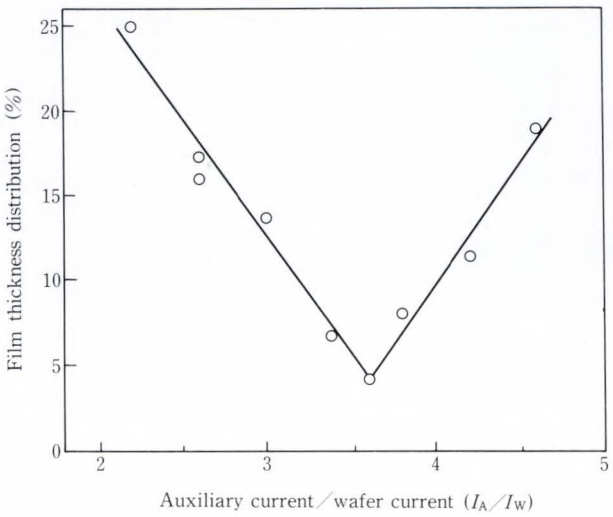


Fig. 22—Current dependence of film thickness distribution.

output distortion and phase shift as well as popcorn noise. Low-current plating is desirable to ensure uniform alloy composition.

The figure also shows that low-current plating makes a Ni-rich film with the existing bath. To avoid Ni-rich composition, the authors added Fe ions and readjusted the composition to obtain the desired Ni percentage, resulting in a permalloy pole with negative magnetostriction as close to 0 as possible.

Since several hundred to a thousand thin film head elements are simultaneously processed on one substrate, the film thickness and composition must be strictly controlled over the substrate. The authors installed an auxiliary electrode encircling the substrate and having 1.4 times the size of the substrate.

The wafer current I_W was fixed and an auxiliary electrode current I_A was varied to ensure uniformity of the film thickness over the wafer. As Fig. 22 shows, the optimum value is $I_A/I_W = 3.6$.

By optimizing the wafer and auxiliary electrode current, the auxiliary electrode shape, and stirring conditions, the authors were able to achieve a film thickness uniformity of 5 percent or less and a composition distribution of 0.3 percent or less.

3.2 Machining process

Figure 23 outlines the machining process.

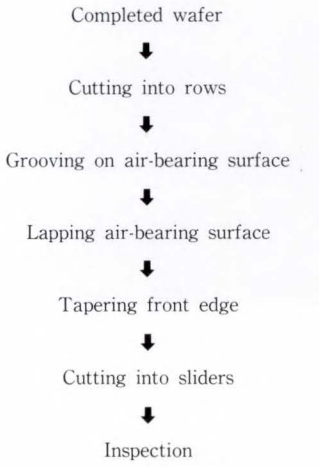


Fig. 23—Slider machining flow.

The wafer-processed substrate is cut into rows, each of which consists of several sliders lined up side by side.

The air-bearing surface, which faces the disk, is grooved and lapped. At the same time, the throat height is controlled to the specified value. A resistance monitoring technique was introduced to allow the throat depth to be accurate to ±1 μm where optical measurement is inadequate.

The tapered edges are lapped to complete row processing.

Each row is then separated into individual sliders.

The slider dimensions are finally inspected and the flatness of the air-bearing surface is checked by laser interference. Figure 24 shows the finished slider.

3.3 Assembly process

Figure 25 outlines the assembly and evaluation processes. Leads are bonded to the completed slider, and protective resin is coated and cured. The electrical resistance and inductance are then measured. The suspension is bonded to the back of the sliders. Leads are formed into the specified shape and secured on the suspension, and the spring loading of the suspension is inspected.

The flying height and posture are inspected on a glass disk. The electromagnetic characteristics are tested by recording and reproducing data on an actual disk media.

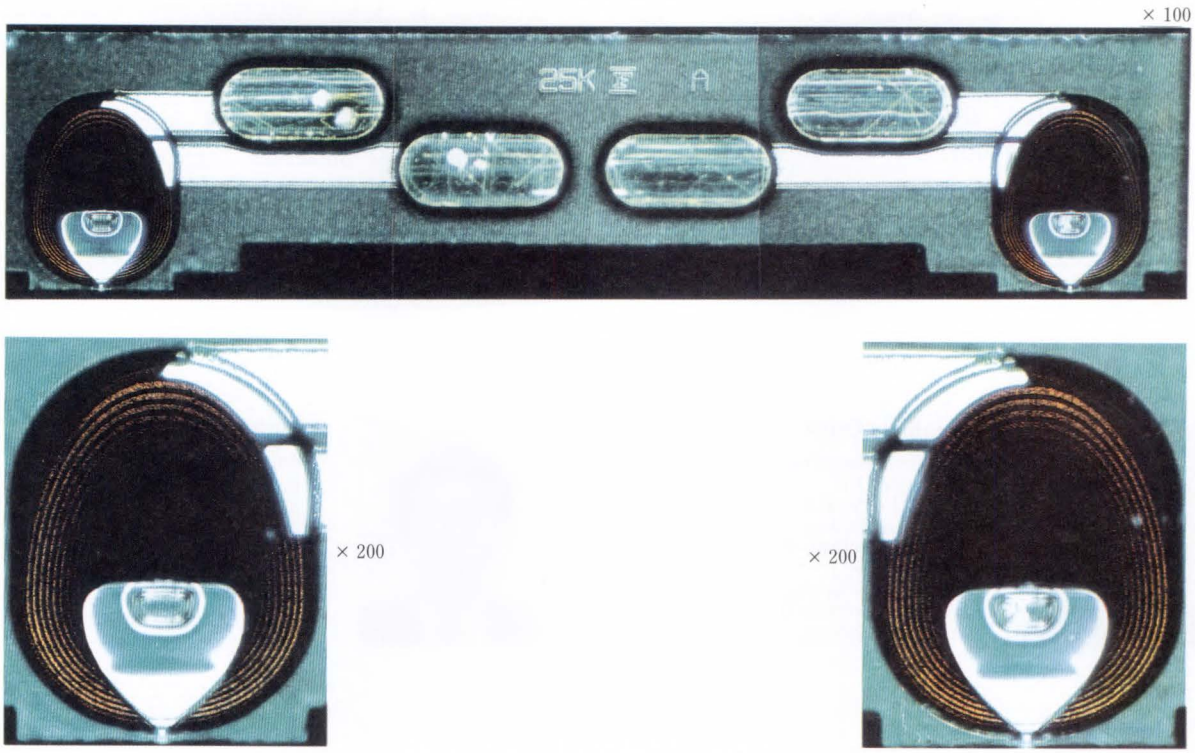


Fig. 24—Completed thin film head elements.

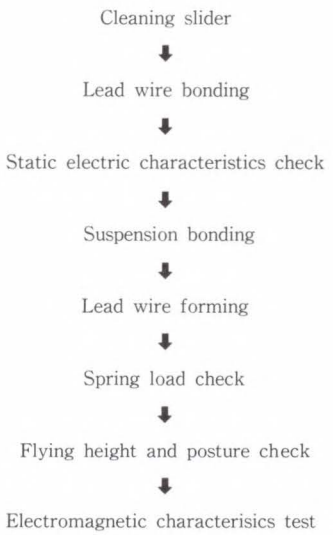


Fig. 25—Assembling flow.

4. Conclusion

This paper discussed technologies for designing and manufacturing thin film heads. Using thin film heads and thin film disks manufactured in the above ways, the authors attained high-density recording at 33 100 bit/inch and 2 078 track/inch for the Hayabusa (F6427H) disk drive[4].

Fujitsu will continue working to achieve even higher recording densities for higher performance disk drives. This includes the development of magnetic materials with high-saturation flux density, and magnetoresistive thin film heads as well as the manufacturing processes for high-volume production.

References

1) Kuwahara, K.: Large Capacity Magnetic Disk Drive. (in Japanese), *NIKKEI ELECTRONICS*, **378**, 1985-09-23, pp. 177-202.
2) Muto, H., Aikawa, T., and Sasaki, M.: Reference Material. Meet., Inst. Electr. Commun. Engrs., Jpn., MR84-3, (1984).
3) Kakehi, A., Oshiki, M., Aikawa, T., Sasaki, M., and Kozai, T.: A THIN FILM HEAD FOR HIGH-DENSITY RECORDING. *IEEE Trans. Magnetics*, **MG-18**, 6, pp. 1131-1133 (1982).
4) Koike, T., Negoro, T., and Yoshida, T.: F6427H Magnetic Disk Subsystem: HAYABUSA. *Fujitsu Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 280-290 (1991).
5) Klaassen, K.B., and van Peppen, J.C.L.: DELAYED RELAXATION IN THIN FILM HEADS, *IEEE Trans. Magnetics*, **25**, 5, pp. 3212-3214 (1989).

6) Ohwe, T., Yoneoka, S., Aruga, K., Yamada, T., and Mizoshita, Y.: A DESIGN OF HIGH-PERFORMANCE INLINE HEAD ASSEMBLY FOR HIGH-SPEED ACCESS. *IEEE Trans. Magnetics*, **26**, 5, pp. 2445-2447 (1990).
7) Mizoshita, Y., Aruga, K., and Yamada, T.: DYNAMIC CHARACTERISTICS HEADSLIDER. *IEEE Trans. Magnetics*, **MG-21**, 5, pp. 1509-1511 (1985).
8) Jones, Jr., R.E.: "IBM3370 Film Head Design and Fabeication". *IBM Disk Storage Technol.*, 1980, pp. 6-7.
9) Kennedy, T.N.: Sputtered insulator Film Contouring over Substrate Topography. *J. Vac. Sci. Technol.*, **13**, 6, pp. 1135-1137 (1976).
10) Wolf, I.W., and McConnel, V.P.: Nickel-Iron Alloy Electrode Deposits for Magnetic Shielding. Proc. Amer. Electroplater's Soc., pp. 215-218 (1956).

Mitsumasa Oshiki
Head Manufacturerering Dept.
File Device Div.
FUJITSU LIMITED
Bachelor of Physics
Gakushuin University 1969
Dr. of Physics
Gakushuin University 1974
Specializing in Thin Film Head Fabrication Process

Shigemitsu Hamasaki
Engineering Dept.
File Device Div.
FUJITSU LIMITED
Bachelor of Instrument and Control Eng.
Kobe University 1972
Specializing in Magnetic Recording

UDC 621.318.681.327.634

Structural Design for High-Performance Magnetic Disk Drives

• Keiji Aruga • Yoshifumi Mizoshita • Mitsuhisa Sekino

(Manuscript received June 4, 1990)

This paper introduces technology related to the mechanical assemblies of high-performance magnetic disk drives.
Guidelines are given for the optimum design of rotary head actuators for high-speed access. The minimization of off-track error and positioning errors caused by disturbances are also discussed, in addition to contamination control ensuring the reliability of a low-flying head. This paper also gives examples of how computer-aided engineering (CAE) is applied to mechanical design. Examples include magnetic field analysis, vibration analysis, and thermal deformation analysis using the finite element method.

1. Introduction

Magnetic disk drives are widely used as the primary file storage device in computer systems ranging from high-end computers to personal computers. This wide use is due to the rapid progress made in magnetic disk drive performance. Recording density has been increasing at least tenfold decade. Compact 5.25-inch disk drives with storage capacities in excess of one gigabyte are already available. This high recording density technology has stimulated a trend towards physically smaller drives (downsizing) in keeping with the downsizing of computer systems.

This progress has been made possible by integrating magnetic, electronic, and mechanical technologies. The performance and reliability of disk drives largely depend on the disk enclosure (DE) mechanisms. During the last ten years, access time, track width, and head flying height, which are closely related to the mechanical design, have been decreased by a factor of one-half to two-thirds[1]. The top-of-the-line models in 1990 feature a track width of 10 μm (2 500 tracks per inch), a seek time of 10 ms, and a head flying height of 0.1 μm.

The achievement of these levels requires highly precise engineering of the disk spindle, head actuator, and air-circulation system. This paper focuses on the latest progress made in the design of these three components.

Downsizing has promoted technical advances which have brought about improvements in mechanical characteristics. At the same time, downsizing has created new technical problems. For example, portability requires stricter head positioning control due to external shock and vibration.

This paper also gives some examples of how computer-aided engineering (CAE) is applied to the design process. (The use of CAE is indispensable to the improvement of design quality.)

2. Design of high-speed head actuators

The linear head actuator, once the only type available, is being replaced by the much more compact rotary actuator[2]. Inline head support[3] significantly reduces the actuator moment of inertia and is now a practicable proposition. Conventionally, the rotary actuator is less rigid than the linear actuator, making it unsuitable for high-speed seeking. But now,

downsizing has increased the resonance frequency of drive components. In addition, progress made in vibration analysis techniques has improved dynamic stiffness design.

The sections below discuss the design of inline rotary actuators.

The design objectives are as follows:

1) A motor suitable for high-speed operation
2) Good controllability and rigidity at high speeds.

More specifically, the target is an average seek time of 10 ms.

2.1 Optimum design of basic dimensions

A higher speed rotary actuator requires that the moment of inertia of moving parts be minimized and the torque constant maximized. The torque and the amount of head displacement can be increased by increasing the radius of the driving coil and the radius of head rotation. However, these changes will significantly increase the total moment of inertia.

The radius of head rotation and the size of driving coil required to achieve the maximum acceleration constant can be found using the following equation:

$$K_\alpha = \frac{Bl\,(r_1)\,r_0 r_1}{J\,(r_0, r_1)}\ , \qquad \cdots\cdots\cdots (1)$$

where, K_α : Acceleration constant (head acceleration/drive current)

J : Total moment of inertia of all moving parts

r_0 : Radius of head rotation

r_1 : Radius of coil rotation

Bl : Force constant (force/drive current).

The denominator of K_α is the drive current. The power consumption is obtained by multiplying the drive current by the supply voltage. Therefore, K_α can also be used as an index of head acceleration per unit of power consumption.

The skew angle (angle of head tilt to the track) must not be too large in relation to the disk size and actuator stroke. The maximum allowable skew angle determines the minimum value of r_0.

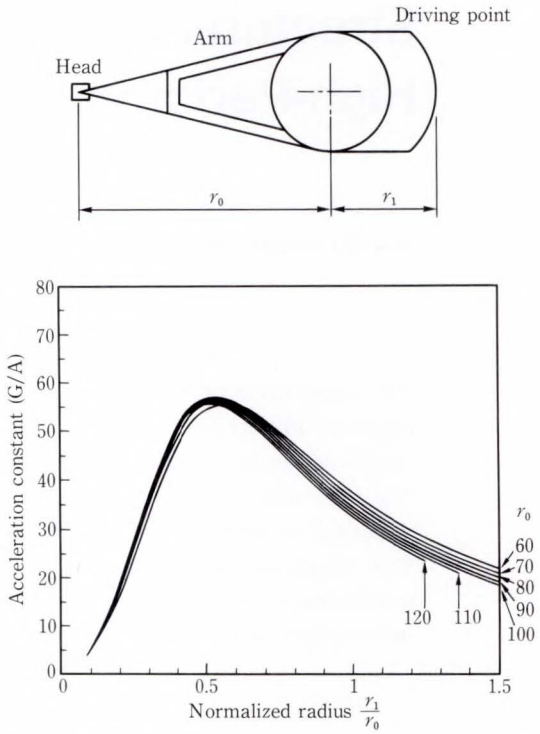


Fig. 1—Acceleration constant versus r_1/r_0 ratio for rotary actuators.

The proposed mechanism uses an inline actuator with a vertical flat coil. Figure 1 shows the results of calculations made to find the optimum coil radius for rotary head actuators. Head assemblies are mounted on the tips of the arms, and the coil is supported by upper and lower support boards. In this structural model, the arm has a simple truss structure with a constant truss width, and the head assemblies are assumed to be identical. The moment of inertia is calculated as a function of r_0 and r_1 to obtain K_α (see Equation 1). The results of K_α calculation in various r_0, r_1 show that there is an optimum coil radius r_1 for any value of r_0. The results shown in Fig. 1 indicate that K_α is maximum when r_1/r_0 is about 0.5. On this basis, the rough dimensions for the actuator can be decided.

If for example, r_0 = 70 mm (corresponding to 5.25-inch drive), the optimum design would have an equivalent mass of about 10 g at the head tip. This value is about one-tenth that of conventional linear actuators.

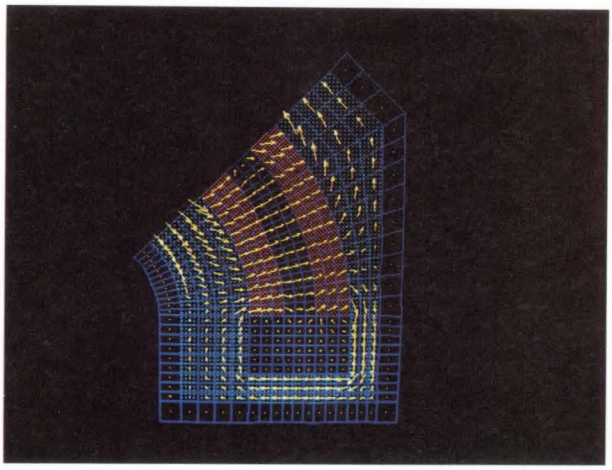


Fig. 2—Finite element magnetic field analysis of the magnetic circuit.

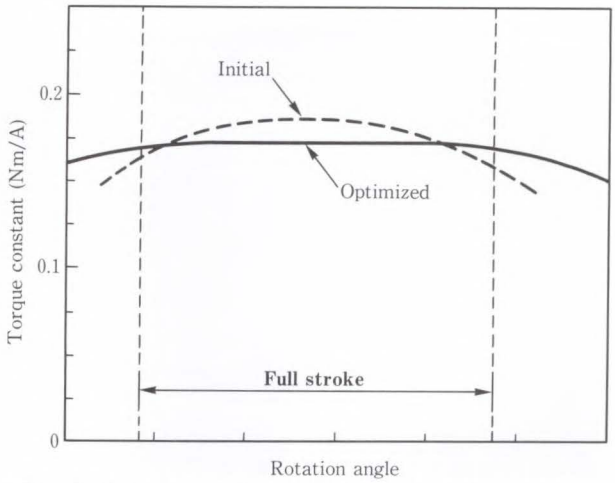


Fig. 3—Rotation angle versus torque constant for two types of magnet assemblies (FEM model).

2.2 Design of the magnetic circuit and coil

Once the basic dimensions have been decided, the motor can be designed.

The optimization of power consumption has been studied[4]. The linearity of the torque constant over the rotation range affects controllability, and research has also been done in this area[5].

Magnet assemblies and coils were designed using computer simulation based on the following methods:

1) Magnetic field analysis using the finite element method (FEM)
2) Torque constant calculation using the results of the FEM.

Half of the magnet assembly was modelled using two-dimensional analysis with the symmetrical boundary condition (see Fig. 2). A high power rare-earth magnet of neodymium, iron, and boron having an energy product of 40 MGOe (320 J/m^3) was used.

Once the magnetic flux density of the air gap was obtained, the torque constant is calculated by numerically integrating the magnetic flux density at the coil.

The general formula for calculating the torque constant is:

$$K_\tau(\theta) = \frac{n}{\Delta\phi \Delta r_0} \int_0^{z_0} \int_{\phi_0+\theta}^{\phi_1+\theta} \int_{r_i}^{r_0} B_g(\phi, z, r) \times r dr d\phi dz , \quad \cdots\cdots\cdots (2)$$

where, $\Delta\phi = \phi_1 - \phi_0$

$\Delta r = r_0 - r_i$

ϕ_0 to ϕ_1 : Angle of coil area (generally, multiple areas)

r_i to r_0 : Radius of coil area

θ : Rotation angle of actuator

ϕ : Angle of the magnetic circuit

B_g : Magnetic flux density of air gap (radial component)

n : Number of effective turns of coil

r : Radius of coil rotation

z_0 : Effective coil height.

When Equation (2) is approximated two-dimensionally and then numerically integrated, the torque constant at actuator angles can be calculated.

Figure 3 shows the relationship between the rotation angle and the torque constant for a conventional assembly (dashed line) and an optimized assembly (solid line) as determined by FEM analysis. Because the magnetic flux density of a conventional assembly is greatest at the magnet center, the torque constant is also greatest at the center of the stroke, resulting in a nonlinearity of about seven percent. The nonlinearity of the optimized assembly is less than three percent and was achieved by changing the shape of the magnet and coil. Linearity improvement, however, tends to lower the peak value of the torque constant. The linearity

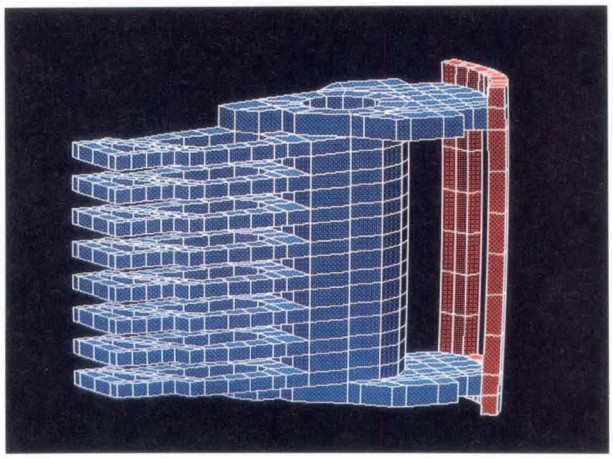
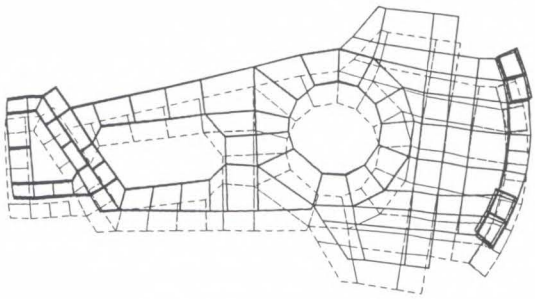
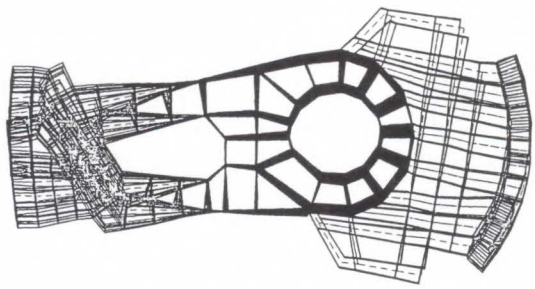


Fig. 4—Solid model of actuator for finite element structural analysis.



a) Frequency: 2.5 kHz



b) Frequency: 6.0 kHz

Fig. 5—Typical actuator vibration modes.

and peak torque constant must, therefore, be optimized.

2.3 Simulation of vibration characteristics

The next step was to design the detailed actuator structure and improve the vibration characteristics[6),7)].

The basic dimensions were determined as

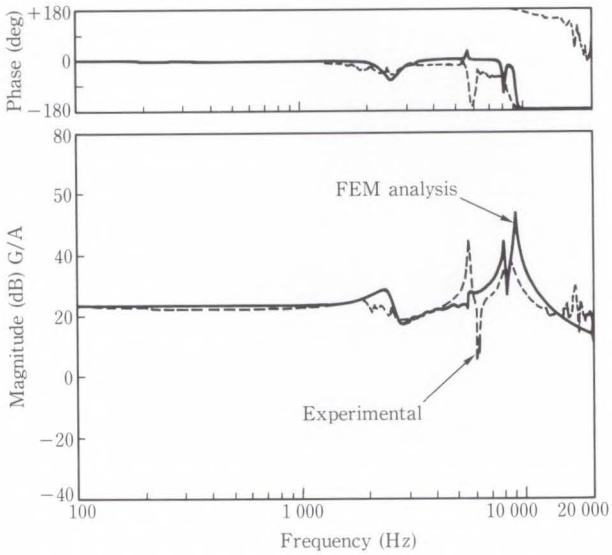


Fig. 6—Mechanical transfer function of arm acceleration/drive current.

explained in section 2.1. Figure 4 shows the solid model of the actuator structure.

This model is used for the following calculations:

1) Mass properties such as weight, balance, and moment of inertia
2) Finite element vibration analysis.

In finite element structural analysis, the model takes the stiffness of the fixed shaft and the rotor bearings into consideration.

Figure 5 shows typical vibration modes obtained by eigenvalue analysis. The vibration due to the mass of the rigid rotor and the spring constant of the bearing occurs at 2.5 kHz. The fundamental vibration mode of rotor deformation in the seek direction is at 6 kHz. These frequencies can be handled by ordinary servo-mechanisms, which having bandwidths of several hundred Hz.

Figure 6 shows the mechanical transfer function of arm acceleration as obtained from eigenvalue analysis. The figure shows the calculated and observed responses, which agree fairly well.

Using these results, the open loop transfer function of the servosystem was then calculated (see Fig. 7). This transfer function is the product of three terms: the actuator transfer function in Fig. 6, the transfer function of the new inline

head assembly[8], and the transfer function of the lead-lag compensator. The bandwidth (open-loop zero crossing frequencies) of the servomechanism is 700 Hz, wider than that of conventional systems. The servomechanism, however, is completely stable and has a phase margin of 40 degree. Although size differences may also have to be considered, the optimized design has rotary actuator characteristics equal to or better than those of a linear actuator, e.g. reference 9.

Figure 8 shows the waveform obtained during an actual seek using the high-speed actuator for 5.25-inch drive and the SMART controller[10]. The figure shows that the target average seek time of 10 ms can be achieved, and that the positioning accuracy after the seek is satisfactory.

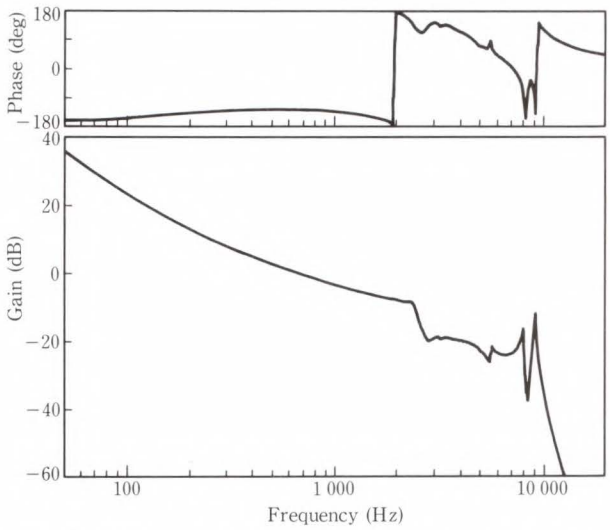


Fig. 7—Open loop transfer function calculated using mechanical model (FEM analysis) and lead-lag compensator.

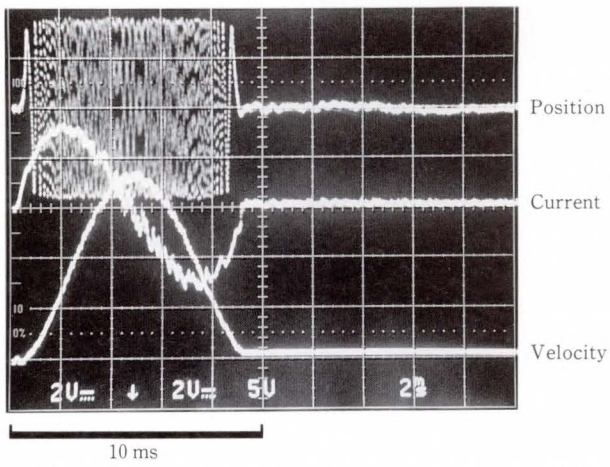


Fig. 8—Typical seek operation using the optimized actuator.

3. Precision head positioning mechanism

To avoid degarding the signal-to-noise ratio, the total positioning error for a 10 μm track width must be about 1 μm.

The following section discuss the problems that affect positioning accuracy.

3.1 Off-track minimization

Off-track error refers to the misalignment between the servo-head and data heads in a dedicated servo system. Off-track error is a major problem that must be solved in order to achive a high track density. Off-track error is caused by thermal deformation in the mechanical structure. It can be compensated to some degree by using a sector servo system, which reads position information from the data area. However, there are many restrictions on using this type of control for fast seeking. Therefore, off-track error must be trackled by reducing the thermal distortion.

Figure 9 shows an example measurement of

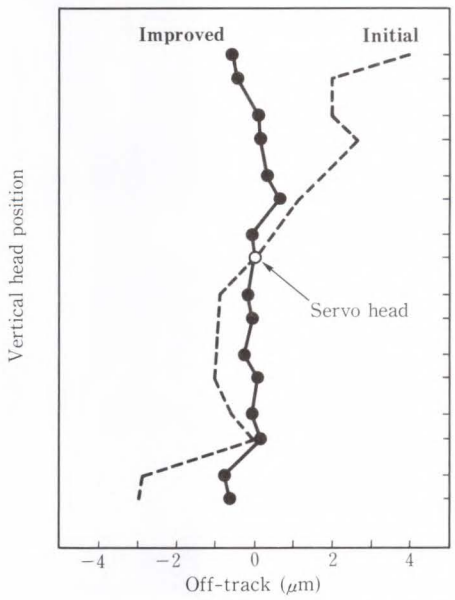


Fig. 9—Example measurement of off-track error for all heads over the temperature range of 0-50 °C.

the off-track error for all heads over the temperature range from 0 °C to 50 °C.

The major causes of off-track error are:

1) Deformation caused by partial temperature differences in the disk enclosure
2) Deformation caused by differences in the thermal expansion ratio of the components of the drive.

Disk drives operate over a wide range of temperatures. Deformation caused by differences in thermal expansion ratios is often the dominant cause of off-track error. If all components are made using the same material, no deformation occurs. The disk platter is made

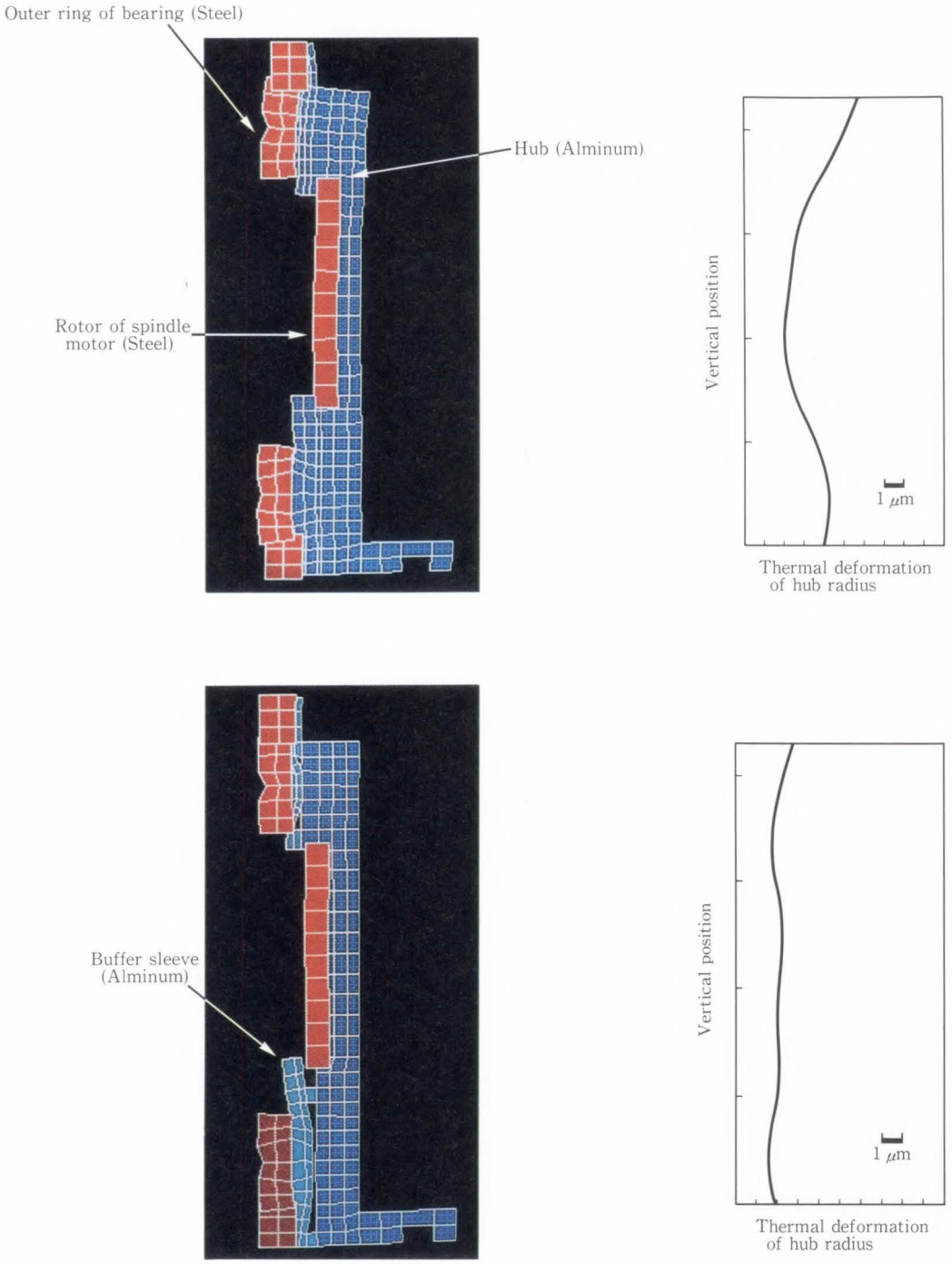


Fig. 10—Finite element thermal deformation analysis for evaluation of the buffer sleeve.

of an aluminum alloy; but aluminum cannot be used for the bearings. Making all parts using the same thermal expansion ratio is, therefore, not practicable. The major problem is the large difference in the thermal expansion ratios of the aluminum spindle (or actuator) and the steel bearings. The dotted line in Fig. 9 indicates the off-track error caused by thermal deformation of the bearing fitting.

If thermal deformation is absorbed by the head arm or disk mounting section, off-track error is suppressed. A low-thermal distortion mechanism was designed to eliminate thermal deformation (see Fig. 10). The effect of the buffer sleeve on the thermal deformation absorption mechanism was simulated using FEM.

Figure 10 shows the result of thermal deformation analysis of the spindle hub using an axisymmetric element. Figure 10 clearly shows the thermal deformation that occurs when the hub is cooled from 50 °C to 0 °C. In the conventional structure, the bearing fitting is deformed by about 3.5 μm.

However, on a structure using the buffer sleeve, thermal deformation is absorbed and the disk mounting area is deformed by 0.5 μm at the most. Because this sleeve reduces stiffness, the resonance frequency of the spindle may be adversely affected. Finite element analysis enables the thermal deformation and rigidity to be balanced. Although this example is for a spindle, the same concept can be applied to the actuator bearing, where the of deformation absorption is the same.

A prototype disk enclosure was manufactured as discussed above and then evaluated. The off-track error of the prototype was considerably less at about 0.7 μm (see the solid line in Fig. 9). Although off-track error is generally dispersive, the dispersion is comparatively small in this mechanism.

3.2 Improved disturbance resistance

Since small disk drives are used in portable personal computers, high tolerance to external shocks and vibration is required.

Positioning error is caused by external shock and vibration, and by the reaction of the actuator drive force during high-speed seeking. If the loop gain of the servo controller is infinite, there is no residual error. The loop gain, however, is limited by its stability.

Positioning errors caused by disturbance and seek reaction force are basically the same in terms of disk movement. Figure 11 shows the control model including the disk and drive movement.

3.2.1 Linear actuator: acceleration feedforward compensation[11)]

Because the seek reaction force in the linear actuator is in the seek direction, a positioning error is easily observed. The linear actuator is susceptible to external vibration in the seek direction. The severest positioning error is found in the dual actuator on a single spindle. When one actuator is on-track, another actuator may seek at high speed. As a result, an on-track actuator positioning error is caused by the movement of the disk in the seek direction. These movement are called mechanical interactions[12)].

Feedforward compensation (see dashed line of Fig. 11) was developed using an accelerometer to minimize mechanical interaction and other disturbances. The measured base acceleration in the seek direction is used for compensation.

The compensation current is,

$$i_{comp} = \frac{m\alpha}{Bl}, \qquad \ldots\ldots\ldots\ldots\ldots (3)$$

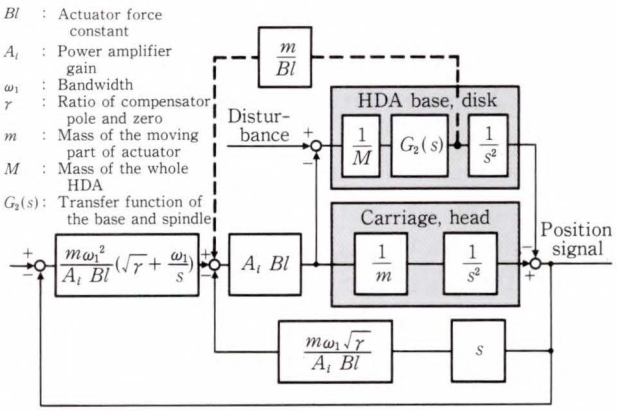
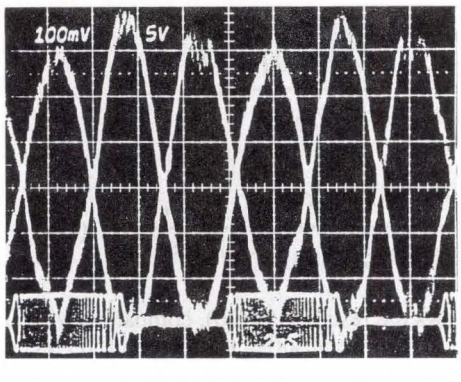
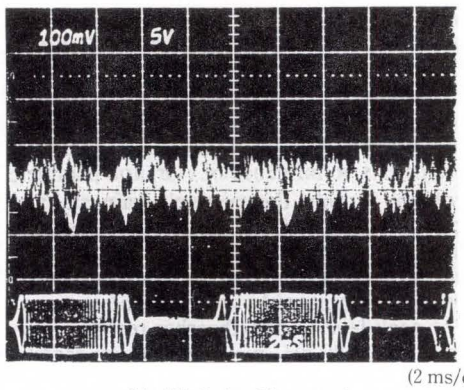


Fig. 11—Block diagram of head positioning model including movement of the entire disk enclosure.



a) Without feedforward



b) With feedforward

Fig. 12—Effect of acceleration feedforward on interactions between dual actuators.

where, α : Acceleration measured at the base.

This type of control is equivalent to the infinite mass of the disk enclosure. Figure 12 shows the effect of this type of control on the worst-case interaction. The positioning error caused by interaction without feedforward control is about 2 μm_{0p}. Compensation reduces the positioning error to about 0.4 μm.

3.2.2 Rotary actuator uncoupled supporting system[13)]

The major feature of the fully balanced rotary actuator is that a translational disturbance does not cause a positioning error. (For the rotary actuator, only the drive rotation causes a positioning error.) However, if the actuator is supported on a vibration isolator, the coupling of the isolator reduces this advantage. The worst positioning error occurs at the resonance frequency of the isolator,

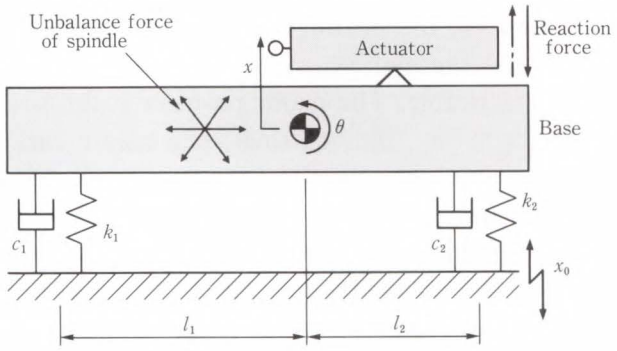


Fig. 13—Model of head positioning for rotary actuators under different disturbances.

and is caused by inadequate damping in the rubber. Some types of high damping rubber are available, but their temperature characteristics do not meet the requirements. However, by careful design, these problems can be compensated.

Figure 13 shows a simplified model of the disturbance that causes positioning errors in the rotary actuator. The control model is the same as the linear model shown in Fig. 11.

Improvements made in the tolerance of the balanced rotary actuator to disturbance is based on uncoupled support. This involves aligning the movement caused by the spring constant of the isolation rubber with the drive's center of gravity. Uncoupled support, therefore, supports the drive while preventing translational disturbances from rotating the drive. Complete uncoupled support, although theoretically possible, is unrealistic.

The degree of coupling is given by:

$$\nu = 1 - \frac{k_2 \, l_2}{k_1 \, l_1},$$

where, $\nu = 0$: fully uncoupled, $\nu = 1$: fully coupled.

Figure 14 shows how the positioning error changes with the frequency of translational-disturbance for a coupling ratio of 0.7 and 0.1. The figure shows that the coupling error falls by about 20 dB when the coupling ratio is reduced to 0.1. This clearly demonstrates the effect of uncoupled support.

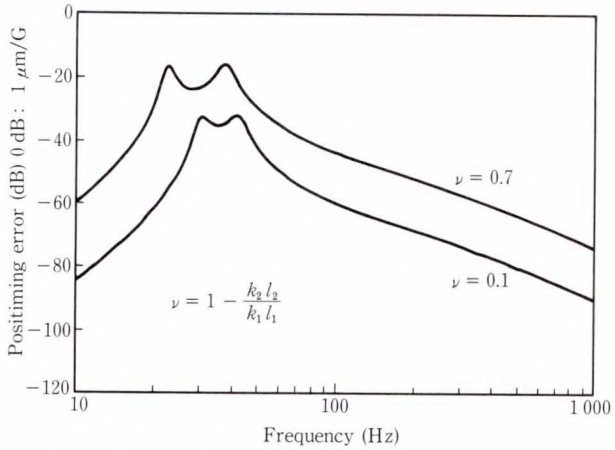


Fig. 14—Positioning error due to external translational vibration for two values of coupling ratio.

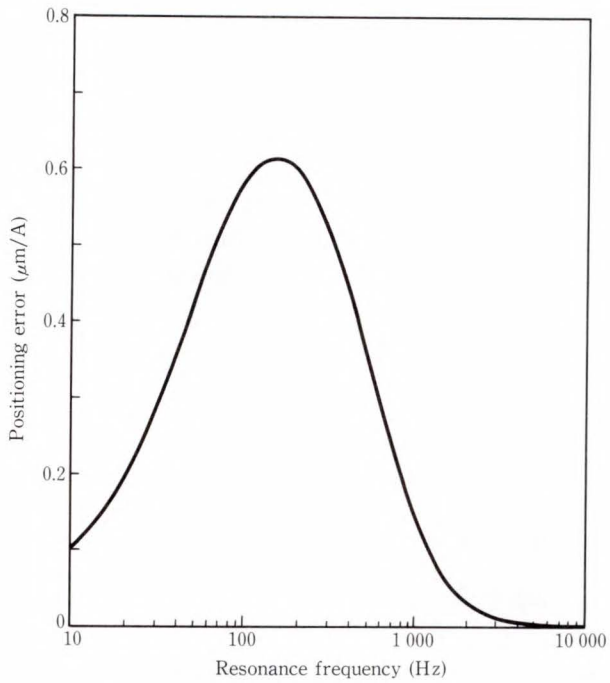


Fig. 15—Positioning error caused by seek reaction force versus resonance frequency of isolator.

Positioning error caused by the seek reaction force is the same as in the linear actuator. Uncoupled support has no effect on this positioning error. Although the use of a rotational accelerometer for compensation is effective, it is slightly exaggerated in small drives.

The effect of decreasing the error by changing the resonance frequency of the isolator is discussed below.

The final tracking error can be found by multiplying the amplitude of disk movement in the seek direction by the tracking error transfer function of the servosystem. Figure 15 shows the relationship between the positioning error and the resonance frequency of the isolator. The figure shows that the positioning error is worst when the resonance frequency of the isolator is about a quarter of the servo bandwidth (700 Hz in this case). If the required positioning accuracy is 0.2 μm, the resonance frequency should be below 30 Hz or above 700 Hz. In this calculation, the damping ratio is 0.1 in terms of the characteristics of the higher temperature rubber. If the damping ratio can be increased, the allowable range of resonance frequency increases.

4. Contamination control

To ensure reliability at a head flying height of 0.1 μm, the probability of contacts between flying head and disk must be decreased. The major causes of contacts are variations in flying height, and contamination in the disk enclosure. The following section dicuss contamination control, which involves the reduction of particle production in the disk enclosure, and the quick removal of particles.

4.1 Recirculation filtration

To purge particles from the DE, recent drive drives commonly use self-recirculating filtration using the air flow generated by disk rotation[14),15)]. Figure 16 shows a simplified model of this system. The breather filter is used to adjust the pressure in the disk enclosure. This model also takes the small leakages of air from and into the disk enclosure into consideration.

The equation for particle purging is:

$$V_0 \frac{dN}{dt} = -N(Q_\eta + q) + N_{\text{out}} q_i + N_{\text{gen}} . \quad \ldots\ldots\ldots (4)$$

The first term indicates the sum of the number of particles trapped by the filter and the number of particles purged to the outside

environment. The second term indicates the number of particles intruding from the outside environment due to small leaks. The third term indicates particles generated in the disk enclosure.

The leakage flow rates, static pressure of the low-pressure region, filter flow rate, and differential pressure are as follows:

$$q = pC_H, \quad \cdots\cdots\cdots (5)$$

$$q_i = p_0 \{C_L + C_b(1 - \eta_b)\}, \quad \cdots\cdots (6)$$

$$p_0 = \frac{\Delta p C_H}{C_H + C_L + C_b}, \quad \cdots\cdots\cdots (7)$$

$$Q = C\Delta p, \quad \cdots\cdots\cdots (8)$$

$$\Delta p = \lambda \frac{1}{2} \rho U^2, \quad \cdots\cdots\cdots (9)$$

where, N : Number of particles per unit volume

V_0 : Interior volume

Q : Filter flow rate

η : Filter trapping efficiency

Δp : Differential pressure generated by disk rotation

p_0 : Static pressure of low-pressure region

C : Flow conductance of circulation filter $= \frac{A}{\sigma}$

A : Effective filter area

σ : Filter pressure loss per unit of flow velocity

C_H : Leakage flow rate at high pressure region

C_L : Leakage flow rate at low pressure region

U : Velocity of disk circumference

ρ : Air density

λ : Dynamic pressure to static pressure conversion efficiency

N_{out} : Particle concentration outside the enclosure

N_{gen} : Number of generated particles in DE per unit of time.

Subscript b indicates value at the breather filter.

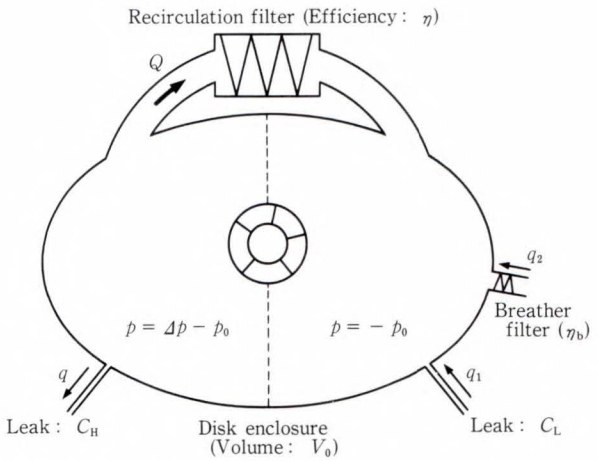


Fig. 16—Equivalent model of disk enclosure filtration.

The differential pressure (Δp) in Equation (9) depends on the flow rate (Q). The relationship between Δp and Q is known as the fan loading curve. This relationship depends on the flow route configuration and many other parameters. To avoid complication, simplified Equation (9) is used for relatively low flow rates. The conversion efficiency λ is difficult to ascertain, but value from experiments is typically is within about 0.1 to 0.5.

By integrating Equation (4), the change in the number of particules over time is found to be:

$$N = \left(N_0 - \frac{N_{out} q_i + N_{gen}}{Q\eta + q}\right) e^{-\frac{Q\eta + q}{V_0} t} + \frac{N_{out} q_i + N_{gen}}{Q\eta + q}, \quad \cdots\cdots\cdots (10)$$

where N_0 : Initial particle concentration.

Equation (10) shows that a larger filter flow rate, higher trapping efficiency, and smaller DE volume will improve purging. Generally, as the trapping efficiency improves, the pressure loss is increased and the flow rate is reduced. In the actual media, the product $Q\eta$ is higher for a low pressure loss low-efficiency media than for a high pressure loss high-efficiency media (e.g. efficiency of 99.97 percent).

The attainable cleanliness depends on the leakage rate and amount of particles

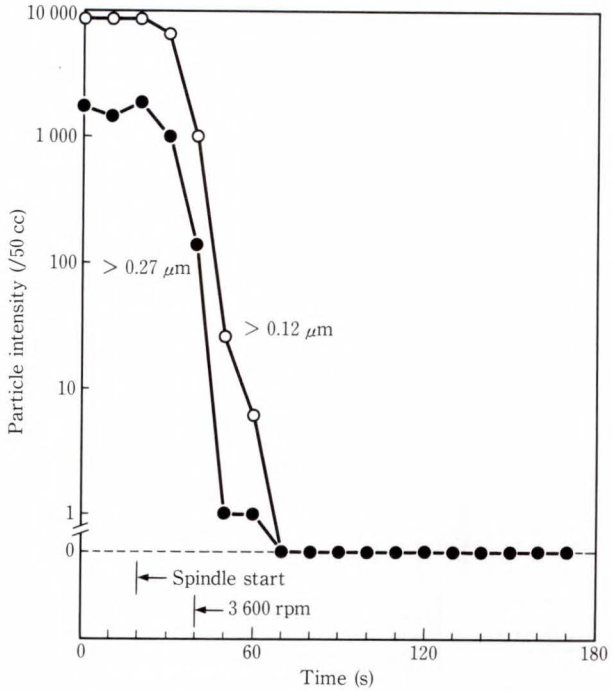


Fig. 17—Particle purging of a 10-inch drive.

Table 1. Calculated and measured values for purge characteristics (one-tenth life)

Rotational speed	Calculated	Measured
3 600 rpm	4.0 s	6 s
2 000	13.1	17

Note: Results of measurement for a 10.5-inch disk enclosure.

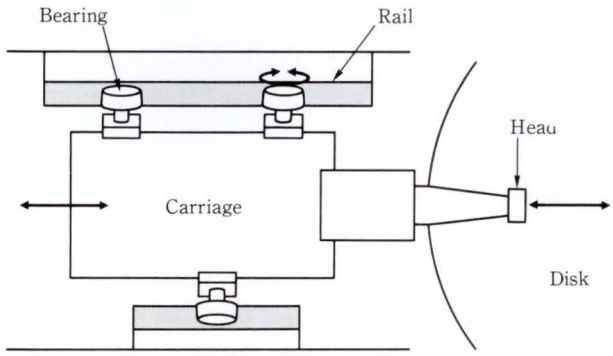


Fig. 18—Typical linear guide mechanism for head actuator.

generated in the disk enclosure. Cleanliness can also be improved by increasing the filter's flow rate.

The one-tenth life (time required for particle concentration to be reduced to one-tenth) has been used to evaluate the particle purge rate.

Neglecting small leaks, the one-tenth life can be represented as follows:

$$T_{10} = \frac{V_0}{2.303 Q \eta} \quad \cdots\cdots\cdots (11)$$

If V_0 and η are considered as constants, the one-tenth life is also an indicator of the filter flow rate.

Figure 17 shows the results of a purge test for an actual disk enclosure.

The number of particles becomes zero within about 30 s after the spindle reaches its rated rotational speed. In theory, particles are purged as depicted by a straight line on a semi-logarithmic chart. Table 1 lists the calculated and measured values for one-tenth life of particle purging. Because of the difficulty in estimating λ (dynamic pressure to static pressure conversion efficiency), the calculated and measured values differ slightly; however, the correlation is strong enough to validate the above design theory.

4.2 Reduced particle generation in the linear guide mechanism

As discussed above, increasing the filter flow rate is one way to improve cleanliness. It is also important to reduce N_{gen} (number of generated particles) in Equation (10). The main sources of particles are the points of friction; in this case, the rolling bearings. The spindle and rotary actuator bearings can be easily sealed off using a magnetic fluid seal, but this is not possible for the linear actuator bearings.

Figure 18 shows a typical linear guide mechanism.

The main particle sources are:

1) Bearing grease
2) Contact between the outer bearing ring and the rail.

To eliminate these particle sources, a new guide mechanism[16] was developed.

A shaft bearing with a magnetic seal on

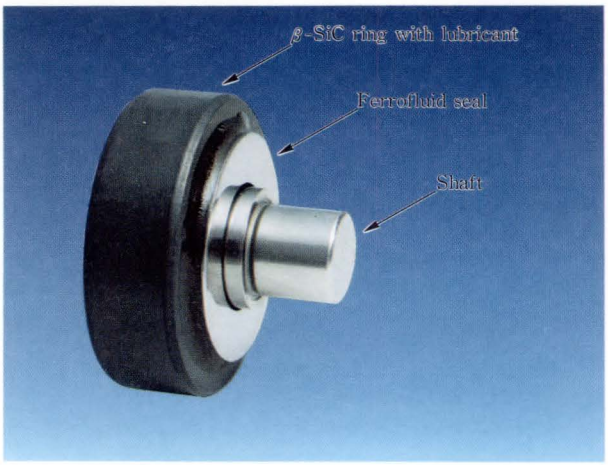


Fig. 19—Ceramic bearing for linear guide mechanism.

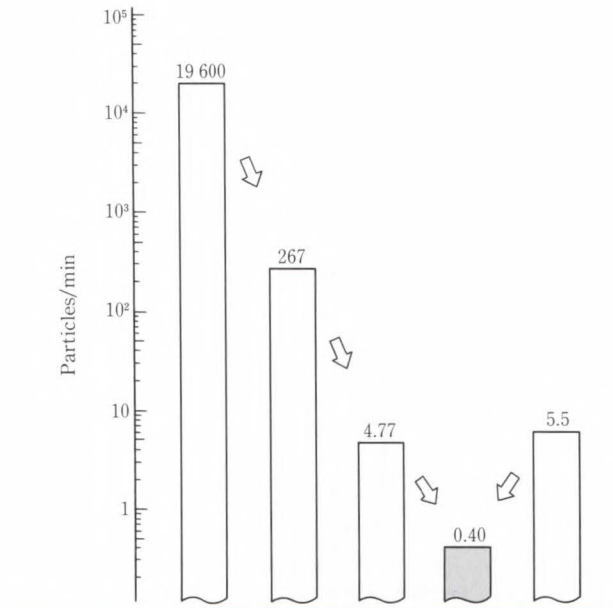
one end is used to eliminate particle source 1) (see Fig. 19). Decreasing particle source 2) is not so easy. The rolling contact area is worn minutely due to setting angle error or sliding caused by the inertial force during quick acceleration. Conventionally, both bearings and rails are made of steel. The combination of bearings and rails generates many particles. Experiments have been made on combinations of fine ceramic materials that may be harder and have a superior wear resistance.

Many accelerated tests have been done using pin-on-disk and bearing-on-disk. Figure 20 shows the results of experiments on bearing-on-disk. The combination of silicon nitride (Si_3N_4) and silicon carbide (SiC) with a fluorocarbon lubricant reduces particle generation to less than 1/10 000 of the figure for steel. The guide mechanism for this combination consists of a precision lapped guide rail (Si_3N_4) and a bearing with a porous SiC outer ring impregnated by fluorocarbon lubricant.

this guide mechanism has been used in tests to measure the number of particles generated under conditions equivalent to continuous seek operations. The results of these tests show that the number of generated particles has been lowered to 1/20 or less, and thus prove the effectiveness of this guide mechanisum.

5. Conclusion

The following techniques for high-performance magnetic disk drives have been



Guide rail	Material	Stainless steel (440 c)	Silicon carbide (SiC)	Silicon nitride (Si_3N_4)
	Surface roughness	Less than 0.04 μm Rz (precision lapped)		0.7 μm Rz
Bearing	Material	Stainless steel (440 c)		SiC impregnated with lubricant

Fig. 20—Particle generation for various combinations of bearing and rail materials.

developed:

1) Optimum design of high-speed head actuators
2) Minimization of thermal off-track errors
3) Minimization of disturbance-caused positioning errors
4) Contamination control of disk enclosure.

These techniques enable the realization of fast, high-density, and high-reliability magnetic disk drives. There are still many areas to be worked on.

The main areas of importance regarding mechanical assemblies are:

1) More compact and precise mechanisms
2) Full off-track error compensation for the actuator
3) Higher rotational speeds.

Within the next few years, a track width of less than 10 μm and a flying height of less than 0.1 μm will be realized. Research and development will be continued to achieve these goals.

References

1) Fujino, M., Sugihara, J., and Ogawa, S.: Magnetic Disk Storage. *FUJITSU Sci. Tech. J.*, **21**, 4, pp. 395-407 (1985).
2) Heath, J.S.: Design of a Swinging Arm Actuator for a Disk File. *IBM J. Res. Develop.*, **20**, 4, pp. 389-397 (1976).
3) Gatzen, H.H., and Hughes, G.F.: Flight Attitude and Takeoff/Landing Behavior of a Miniature Winchester Head for Rotary Actuators. *Soc. Tribol. Lubric. Engrs.*, **SP-22**, pp. 133-137 (1987).
4) Cooper, E.S.: MINIMIZING POWER DISSIPATION IN A DISK FILE ACTUATOR. *IEEE Trans. Magnetics*, **MAG-24**, 3, pp. 2081-2091 (1988).
5) Winfrey, R.C., Riggle, C.M., Bernett, F., Read, J., and Svendsen, P.: DESIGN OF A HIGH PERFORMANCE ROTARY POSITIONER FOR A MAGNETIC DISK MEMORY. *IEEE Trans. Magnetics*, **MAG-17**, 4, pp. 1392-1395 (1981).
6) Kakizaki, T., Kogure, K., and Mitsuya, Y.: Deisgn and Dynamic Analysis of a High Performance Rotary Actuator for a Disk Storage. Proc. 1st Int. Symp. Design Synthe. Conf., 1984, pp. 813-818.
7) Ohta, H., Naruse, J., and Hirata, T.: Vibration Reduction of Magnetic Disk Drives with Swinging Arm Actuator. (in Japanese), Trnas. *Jpn. Soc. Mech. Engrs.*, Ser.-C, **50**, 459, pp. 2113-2123 (1984).
8) Ohwe, T., Yoneoka, S., Aruga, K., and Mizoshita, T.: A DESIGN OF HIGH PERFORMANCE INLINE HEAD ASSEMBLY FOR HIGH SPEED ACCESS. *IEEE Trans. Magnetics*, **26**, 5, pp. 2445-2447 (1990).
9) Franklin, G.F., Powell, J.D., and Workman, M.L.: Digital Control of Dynamic Systems. 2nd ed., Addison Wesley, 1990, 718p.
10) Hasegawa, S., Takahashi, K., and Mizoshita, Y.: Digital Servo Control for Head-Positioning of Disk Drives. *FUJITSU Sci. Tech. J.*, **26**, 4 (Special Issue on Fujitsu File Devices), pp. 378-390 (1991).
11) Aruga, K., Mizoshita, Y., Iwatsubo, M., and Hatagami, T.: Acceleration Feedforward Control for Head Positioning in Magnetic Disk Drives. *Jpn. Soc. Mech. Engrs.*, Int. J., Ser. III, **33**, 1, pp. 35-41 (1990).
12) Mizukami, M.: Mechanical Interaction Supression in a Dual-Actuator by Trapezoidal Driving. 16th IMCSS Symp. 1987, pp. 321-326.
13) Aruga, K., Yamada, T., Imamura, T., and Mizoshita, Y.: A Study on Head Positioning Accuracy in Magnetic Disk Drives with Rotary Head Actuator. (in Japanese), *Jpn. Soc. Mech. Engrs.*, Proc. C, **900**-14, pp. 463-465 (1990).
14) Mizoshita, Y., and Matsuo, N.: MECHANICAL AND SERVO DESIGN OF A 10 INCH DISK DRIVE. *IEEE Trans. Magnetics*, **MAG-17**, 4, pp. 1387-1391 (1981).
15) Hattori, S., Takekado, S., and Kurihara, E.: Design of Filteration Mechanism in Magnetic Disk Drives. (in Japanese), *Jpn. Soc. Mech. Engrs., Proc.*, **860**-6, pp. 22-24 (1986).
16) Maeda, H., Mizoshita, Y., Aruga, K., and Yamada, T.: Ceramic Guide Mechanism of Linear Head Actuators for Magnetic Disk Drives. *Soc. Tribol. Lubric. Engrs.*, **SP-29**, pp. 85-90 (1990).

Keiji Aruga
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Mechanical Eng.
University of Tokyo 1975
Master of Mechanical Eng.
University of Tokyo 1977
Specializing in Mechanism of Magnetic Disk Drives

Yoshifumi Mizoshita
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Precision Eng.
Hiroshima University 1972
Master of Precision Eng.
Hiroshima University 1974
Specializing in Mechanical Control Engineering

Mitsuhisa Sekino
Engineering Dept.
File Device Div.
FUJITSU LIMITED
Bachelor of Mechanical Eng.
Keio University 1975
Master of Mechanical Eng.
Keio University 1977
Specializing in Mechanism Design of Magnetic Disk Drives

UDC 681.327.634:681.58

Digital Servo Control for Head-Positioning of Disk Drives

• Susumu Hasegawa • Kazuhiko Takaishi • Yoshifumi Mizoshita

(Manuscript received June 25, 1990)

A digital servo controller has been developed for head positioning in magnetic disk drives. In order to achieve precise head positioning and stable fast access operation, the actuator velocity must be estimated correctly. A state space control that uses the state estimator is introduced in this paper.

An access control method called SMART (Structural Vibration Minimized Acceleration Trajectory) has also been developed. By using this SMART control, the high harmonics of actuator drive are damped and thus both the residual vibration immediately after access and flying height fluctuation of the head are decreased. This paper describes the design and the experimental results of the state space controller and the SMART access control.

1. Introduction

Digital servo control, i.e. servomechanism control using digital computers, was used only for large systems in early days, such as aircraft altitude control and the control of robot arms. However, as high-integration, high-performance and low-priced microprocessors have become available in recent years, the use of digital servo control is expanding to car and domestic electrical appliances, and office-automation equipment.

The general advantage of digital servo control over analog servo control is that it has intelligent functions. The digital servo control can make control more flexible, and also make control circuits compact.

For head positioning control of magnetic disk drives, the write/read head must be positioned accurately on the data track, which is 10 μm wide or less. The head, which floats at 0.1 μm to 0.3 μm on the magnetic recording disk surface, moves at a high acceleration of 50 G to 100 G.

The first advantage by the digital servo control is the introduction of vibration control that suppresses the resonance of the mechanical parts (sliders, gimbals and arms) caused by high-speed motion, and that in turn prevents noise and degradation of performance.

The second advantage facilitates automatic adjustment of control parameters that absorb the variation of the mechanical characteristics for each disk enclosure and the fluctuation depending on the time and external environment[1]. This function cannot be achieved by the conventional analog control where parameters are fixed. Digital servo control also improves productivity independently of device types because common control programs can be used.

The use of integrated digital circuits further fosters the downsizing trend in the late magnetic disk drives (from 5.25 to 3.5 and 2.5 inches).

As a result, use of the digital servo control for head positioning will greatly improve reliability and reduce cost of mass-produced disk drives.

This paper describes the introduction of state space control in the tracking operation and vibration suppression control in the access operation on the head positioning system.

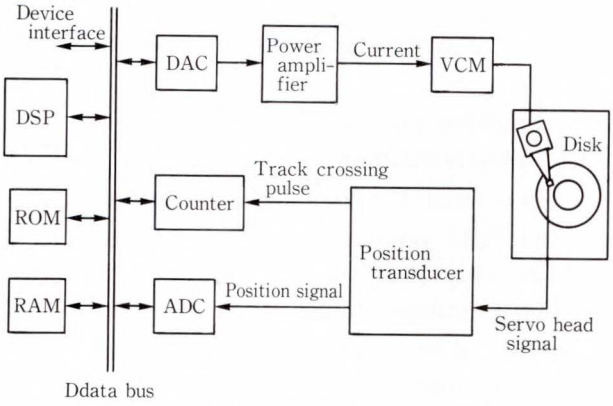


Fig. 1—Digital servo system for head positioning.

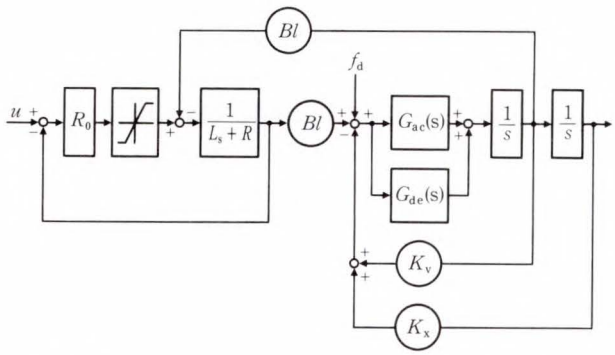


Fig. 2—Model of power amplifier, motor and plant.

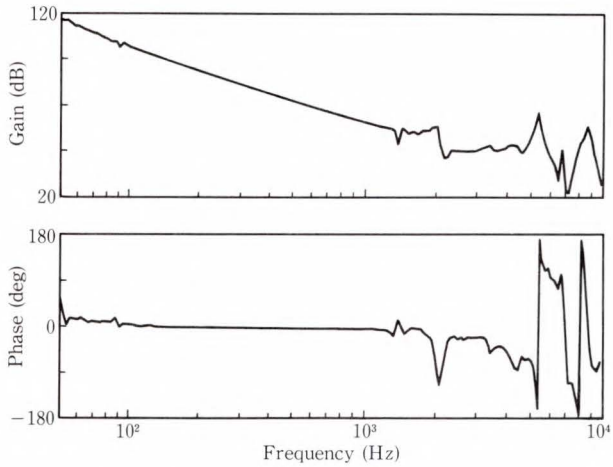


Fig. 3—Mechanical transfer function of actuator.

2. Control systems

2.1 Configuration of the digital servo systems

Figure 1 shows the configuration of the digital servo system for head positioning of the magnetic disk drives. The servo pattern is pre-recorded on the servo disk surface for dedicated servo systems, and also on a part of the sectors of the data disk surface for sector servo systems. This servo pattern is transformed to analog position signals and track crossing pulses by the position transducer.

The controller consists of the digital signal processor (DSP), analog-to-digital converter (ADC), digital-to-analog converter (DAC) and counter. The drive signal output from the controller is converted to current by the power amplifier and the current drives the voice coil motor (VCM).

The servo system for head positioning is connected to the host controller through the bus line.

2.2 Servo model

Figure 2 is a block diagram of the detailed model of the disk drive servo system. In the diagram, f_d indicates the distrubance to the system and K_v or K_x is the proportional factor of the force depending on velocity or position. The mechanical transfer functions $G_{ac}(s)$ and $G_{de}(s)$ indicate the vibration characteristics of the actuator and disk enclosure, respectively.

The vibration characteristics of the actuator are an important factor that determines the servo band in general. In order to improve positioning accuracy, the servo band must be widened, but this is restricted by the mechanical resonance of the actuator.

Figure 3 shows an example of the mechanical transfer function of the actuator for the magnetic disk drive. Mechanical resonance appears at 1.3, 2.0 and 5.2 kHz. These incidences of resonance can be reduced using notch or low-pass filters. Even then, the system oscillates as the servo band increases because phase rotation occurs near the servo band frequency. Therefore, the upper limit of the servo band in which stable servo characteristics are provided without generating oscillation is assumed to be about 500 Hz to 700 Hz.

2.3 Problems

This section describes several problems to be considered in designing a control system

for head positioning.

2.3.1 Vibration of the mechanical parts

The mechanical vibration of magnetic disk drives can be roughly classified into two areas: low frequency vibration that is lower than the servo bands (500 Hz to 700 Hz) and high frequency vibration that is higher than the servo bands. The former is caused by the anti-vibration support mechanism of the drive and the latter comes mainly from the actuator mechanism.

The low frequency vibration can be suppressed by control systems if the vibration can be measured (this paper does not describe the specific countermeasures). Basically, the high frequency vibration cannot be suppressed by a control loop. Consequently, to reduce the positioning error caused by high frequency resonance, either the rigidity of mechanical parts must be increased, or an access operation which does not generate this mechanical resonance should be introduced.

2.3.2 Elimination of stationary offset force

The offset force applied to the servo system consists of the reaction force of the flexible printed circuit board (FPC) for signal connection to the actuator, and the wind pressure due to the air flow caused by the high-speed rotation of the disk. This offset force can be eliminated on the control loop by only applying an integrator to the control loop if the offset force is weak. However, if the offset force is strong and the integrator is saturated, feedback for correcting the offset force is required.

2.3.3 Velocity measurement

To perform the state feedback control, the velocity of the actuator is required, but in the magnetic disk drive this velocity is not directly observed. Therefore, there must be a method for estimating the velocity from the measured position and VCM drive current.

2.3.4 VCM inductance

In the conventional access operation, because the maximum acceleration had to be applied, a time lag due to VCM coil inductance occurred. Consequently high-speed access was prevented.

The authors have solved these problems by using the new tracking and access systems (described in chapter 3 and subsequent chapters).

3. Tracking control mechanism

3.1 Velocity estimation

In general, the basic head positioning control is the state feedback, that is, the PID control which uses position, integrated position and velocity independent of analog and digital control. The state value observed in the magnetic disk drive is the position (position error) only. In this case, the velocity must be estimated.

For this purpose, the following two methods are used:

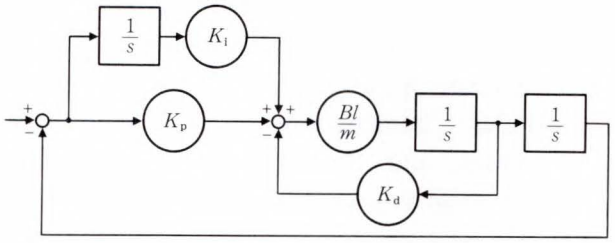
1) Differential values of the position
2) Combination of position differential values and current integral values.

Method 2) was used in the conventional analog control system.

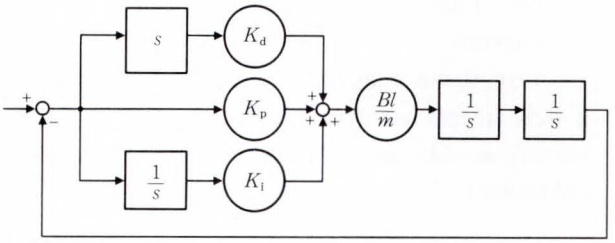
Figures 4 a) to c) show block diagrams of the idealized states feedback control and the PID control using the above methods. Figure 5 shows the simulation results of the open loop characteristics about these three control methods. Here the feedback coefficients use the values of the optimal regulator[2]. As a result, when using Method 1), it is found that the peak of phase margin (phase difference angle in the gain intersection and stability index of digital servo systems) cannot be restricted and the zero cross frequency also shifts to the high frequency band. Even when using Method 2), the peak of phase margin is smaller than the idealized states feedback control. This is because the velocity obtained from the differentiation of position error signals is the relative velocity to the disk rather than the absolute velocity.

3.2 Installation of state estimator

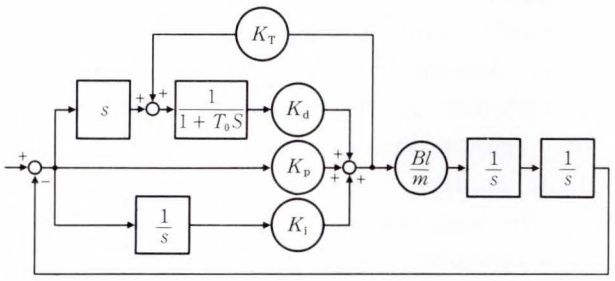
The state feedback control can be configured easily by estimating the state value of the plant on a realtime basis if digital control with high-speed arithmetic unit is assumed[3),4)]. Therefore, the designers have installed a state space control that uses this state estimator in the head positioning system.



a) Idealized states feedback system



b) PID control system (differential position)



c) PID control system
(differential position + current integral)

Fig. 4—Tracking control systems.

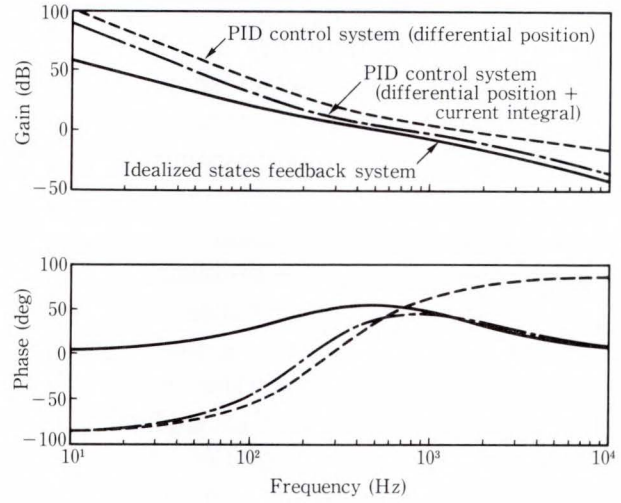


Fig. 5—Open loop transfer function (continuous system).

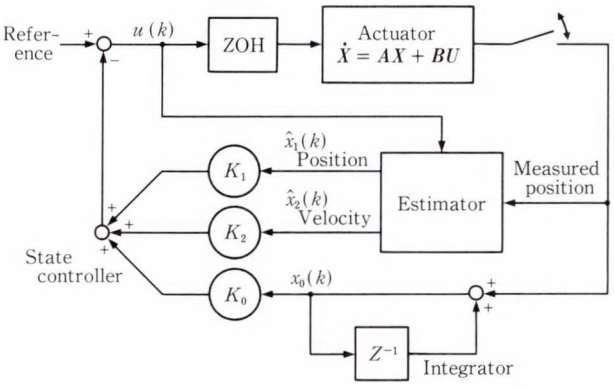


Fig. 6—Block diagram of state space control.

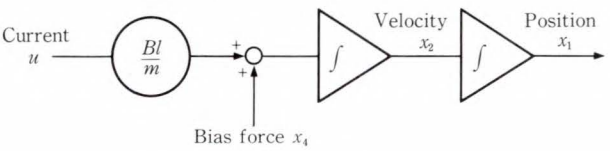


Fig. 7—Model of the actuator.

Figure 6 is a block diagram of this state space control. To design the state estimator, the model of the object to be controlled must first be set. In this case, a double integrator model with fixed bias force is assumed as shown in Fig. 7 (see Appendix 1 for a specific control model and the representation of an equation of the state estimator).

The control rule for the controller of Fig. 6 is expressed by the following equation:

$$u(k) = -K_1\hat{x}_1(k) - K_2\hat{x}_2(k) - K_0x_0(k),$$

where, $\hat{x}_1(k)$ is estimated position, $\hat{x}_2(k)$ is estimated velocity and $x_0(k)$ is integrated position. Feedback gains K_0 to K_2 can be determined from the simulation of frequency response based on the pole placement or optimal regulator theory (see Appendix 2).

3.3 Simulation results

Figure 8 shows the simulation results of the open loop characteristics of the state space control (sampling frequency 30 kHz) using this state estimator. The feedback gain is obtained by the optimal regulator ($\gamma = 9$, design

zero cross frequency of 200 Hz), and the estimator cutoff frequency is 3 kHz ($\zeta = 0.5$). Consequently, the diagram shows that a sufficient phase margin is obtained at a servo band of 500 Hz by using the estimator control.

4. Access control mechanism

4.1 Mechanical resonance problems

In a conventional analog servo control system, optimal control is performed based on the deceleration curve. This curve is designed so that drive current can be applied up to the limit of the power amplifier at acceleration and so that coil power consumption can be minimized at deceleration[2].

However, the resonance (main resonance of 2 kHz to 5 kHz) is easily excited in the control based on the conventional deceleration curve when there are high harmonics in the acceleration. As a result, the residual vibration continues immediately after the access terminates, and the settling time is delayed or the positioning accuracy is degraded (see Fig. 9).

Further, for a head slider floating at a spacing of 0.1 μm to 0.3 μm, the resonance of the high frequency band causes floating fluctuation.

For these reasons, authors have developed a new access control system[5] called Structural vibration Minimized Acceleration Trajectory (SMART).

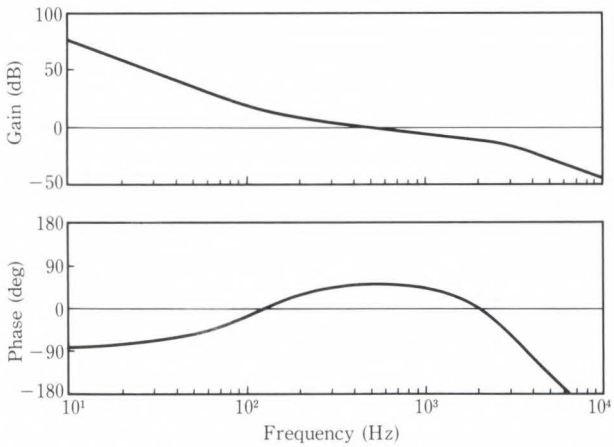
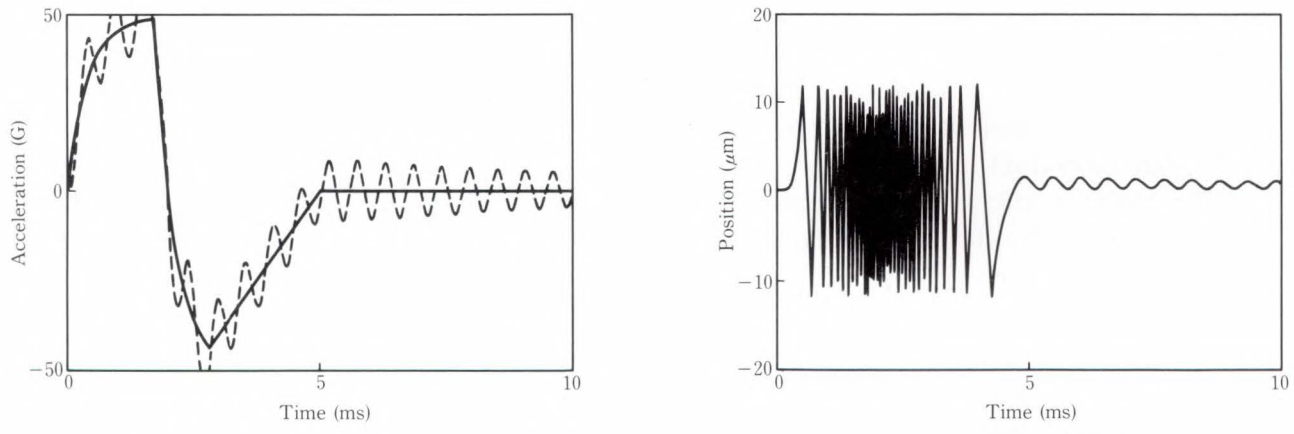


Fig. 8—Open loop transfer function using state estimator (simulation).

4.2 Design of the SMART control system

The actuator driving systems that have been proposed to date use the trapezoidal driving current[6], combine the minimum trajectories of power dissipation and jerk[7], and use the residual vibration as the estimation standard[8]. But they cannot always be used easily by the actual digital control.

The basic design concept of SMART control is to formulate "smooth motion" that is unlikely to generate vibration and is in a form in which the motion can be handled easily by digital signal processing. Accordingly, the authors set the square integral of the differentiated acceleration $d\alpha/dt$ (rate of acceleration change) as cost function P and obtained the solution



a) Acceleration and resonance of actuator

b) Residual vibration

Fig. 9—Resonance of actuator by use of conventional trajectory.

that minimizes this cost function.

$$P = \int_0^{T_0} (\frac{d\alpha}{dt})^2 \, dt \, ,$$

assuming that the access time is T_0 and the access distance is a. The initial and terminal conditions are listed below.

$t = 0$: $x_1(t) = a$ $x_2(t) = 0$ $x_3(t) = 0$,

$t = T_0$: $x_1(t) = 0$ $x_2(t) = 0$ $x_3(t) = 0$,

where, position: $x_1(t)$ velocity: $x_2(t)$ acceleration: $x_3(t)$. From this, when using the optimal control theory to normalize the solution, the SMART state values (position, velocity and acceleration) are expressed in the following polynomials of time (see Appendix 3).

$$x_1(t) = -60\, a \left\{ \frac{1}{10}(\frac{t}{T_0})^5 - \frac{1}{4}(\frac{t}{T_0})^4 + \frac{1}{6}(\frac{t}{T_0})^3 \right\}.$$

$$x_2(t) = -60 \frac{a}{T_0} \left\{ \frac{1}{2}(\frac{t}{T_0})^4 - (\frac{t}{T_0})^3 + \frac{1}{2}(\frac{t}{T_0})^2 \right\}.$$

$$x_3(t) = -60 \frac{a}{T_0^2} \left\{ 2(\frac{t}{T_0})^3 - 3(\frac{t}{T_0})^2 + (\frac{t}{T_0}) \right\}.$$

Figure 10 shows a diagram plotting each state value of the SMART control against time. The acceleration varies smoothly and the acceleration and deceleration peaks are located at $t/T_0 = 0.21$ and 0.79. Such an equation normalized for time is very easy to use for digital control programming.

4.2.1 Determination of the access time

Immediately before the access operation is performed, the SMART control must predetermine how much time it takes for the head to move to the desired track.

As an example, Fig. 11 a) shows the relationship (design value) between the number of access tracks and its access time when the maximum acceleration change rate (maximum gradient of the current) is fixed independently

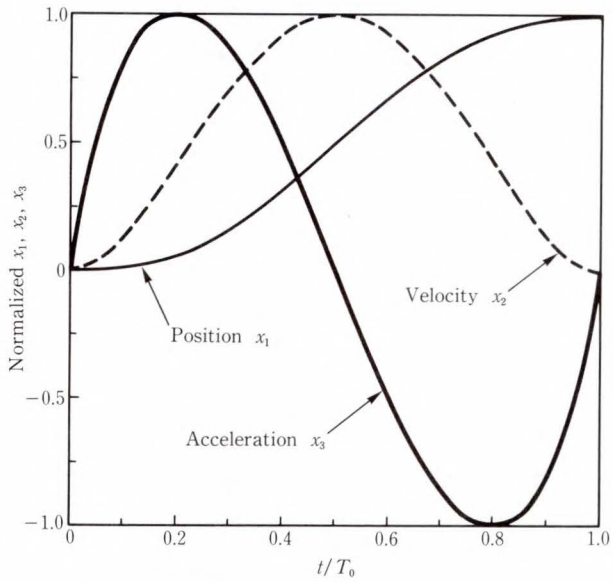


Fig. 10—Optimal trajectories (SMART).

of the number of access track. In this case, the access time is proportional to the cubic root of the access distance. Figure 11 b) shows the maximum velocity and current (proportional to the acceleration) in each track access. The diagram shows that the maximum current increases linearly for the access time required for each track access.

4.2.2 Control algorithm

Figure 12 is a block diagram of the SMART control.

The control law is as follows:

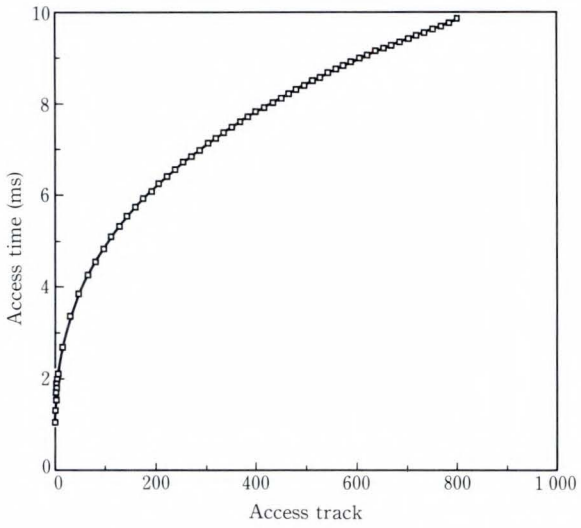
$u(k) = K_v \{ V_t - \hat{x}_2(k) \} + K_f x_3(k)$,

where, V_t is the desired velocity (SMART velocity) trajectory calculated for each sampling period. $K_f x_3(k)$ is the feedforward current (SMART acceleration) for its velocity trajectory.

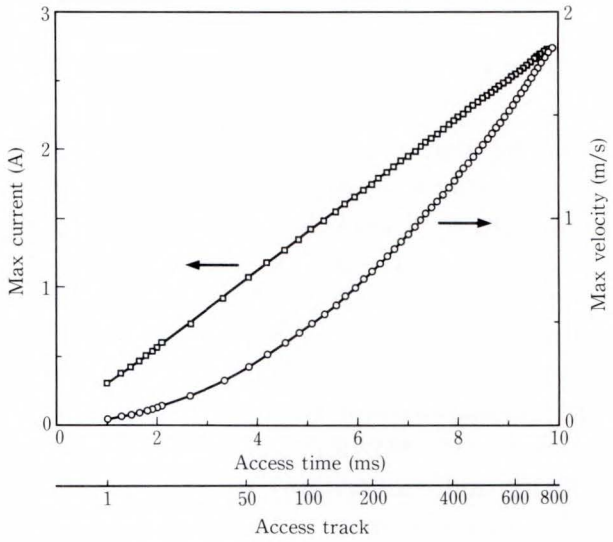
4.3 Comparing SMART and conventional trajectories

4.3.1 Acceleration spectrum

The SMART control feature reduces the high harmonics of the actuator due to smooth motion. Figure 13 compares the acceleration waveform spectrum of the SMART and conventional trajectories. For the SMART control, the diagram shows that the vibrational energy decreases by about 20 dB in the vicinity of the resonance frequency of the actuator (2 kHz),



a) Access time vs access track



b) Max current and max velocity

Fig. 11—Design of access operation (SMART, max di/dt = const.).

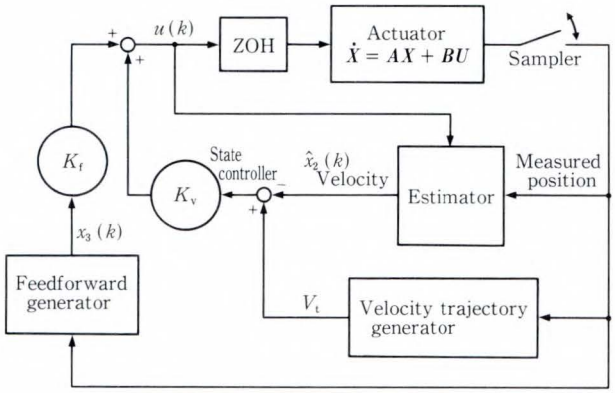
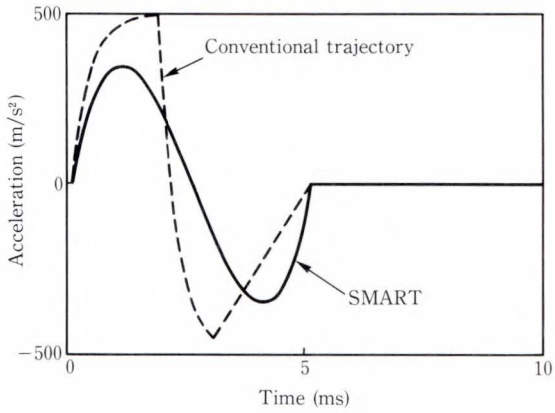
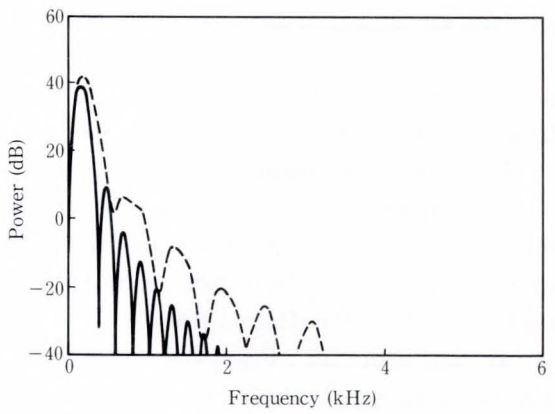


Fig. 12—Block diagram of access control (SMART).



a) Acceleration



b) Power spectrum

Fig. 13—Comparison with conventional trajectory.

and that actuator resonance is not easily excited.

4.3.2 Coil power consumption

Table 1 compares the maximum acceleration (peak current) and coil power consumption of the SMART and conventional control systems when the access distance and time are fixed.

In this table, the value for the conventional system is the actual measured value for the specific case when the seek distance is 1.5 mm. The shortest time control and the minimum power consumption control cannot actually be realized due to coil inductance.

The table shows that the coil power

Table 1. Power consumption comparison

Trajectory	Maximum acceleration	Coil power consumption
SMART	$\frac{10}{\sqrt{3}}\frac{a}{T_0^2}$ (1)	$\frac{120}{7}R(\frac{m}{Bl})^2(\frac{a^2}{T_0^3})$ (1)
Minimum time	$4\frac{a}{T_0^2}$ (0.69)	$16R(\frac{m}{Bl})^2(\frac{a^2}{T_0^3})$ (0.93)
Minimum power	$6\frac{a}{T_0^2}$ (1.04)	$12R(\frac{m}{Bl})^2(\frac{a^2}{T_0^3})$ (0.7)
Cycloid	$2\pi\frac{a}{T_0^2}$ (1.09)	$2\pi^2R(\frac{m}{Bl})^2(\frac{a^2}{T_0^3})$ (1.15)
Conventional*	$8\frac{a}{T_0^2}$ (1.39)	$30R(\frac{m}{Bl})^2(\frac{a^2}{T_0^3})$ (1.75)

Bl : Motor force constant m : Moving mass
R : Coil resistance
* : Seek distance is 1.5 mm
Number in ( ) is a ratio to SMART value.

Table 2. DSP specifications

Instruction execution time	100 ns
Data word	16-bit fixed decimal point
Multiplier	32-bit (16 x 16 bits)
Shifter	16-bit parallel

Table 3. Specification of ADC and DAC

	Number of bits	Conversion time
ADC	8	3 μs
DAC	12	4 μs

consumption for the SMART control is suppressed to 60 percent of that in the conventional system. If the acceleration varies in the cycloidal curve, both the peak current and coil power consumption become higher than those of the SMART control system.

5. DSP implementation and results

5.1 DSP and peripheral circuit

5.1.1 DSP

To perform the state estimation and control operation in a limited sampling period, economical DPSs with fast execution speed must be used. Table 2 lists an example of the DSP specifications.

5.1.2 ADC and DAC

The resolution of the ADC must be less than 1/10 that of the desired positioning accuracy and a fast conversion time is required.

To obtain a drive current resolution that corresponds to the positional resolution, the DAC must have a larger number of bits than the ADC. The output range of DAC must cover the dynamic range needed for access operation. Table 3 lists examples of the ADC and DAC.

5.1.3 Position signal demodulator

Two-phase position signals are demodulated from four types of magnetic patterns that were shifted in the radial direction and in the rotational direction. This demodulator circuit consists of the AGC amplifier, SYNC pulse detector, peak holder and PLL circuit. The circuit has been integrated on one chip.

5.2 Control program structure

Figure 14 is a basic flowchart of the designed DSP control program. The design sampling period is determined by the timer interrupt function of the DSP chip. Tasks are divided into two: tracking and seek routines. The state is estimated by both routines. In the seek routine, however, the desired velocity and acceleration trajectory (SMART) must also be calculated.

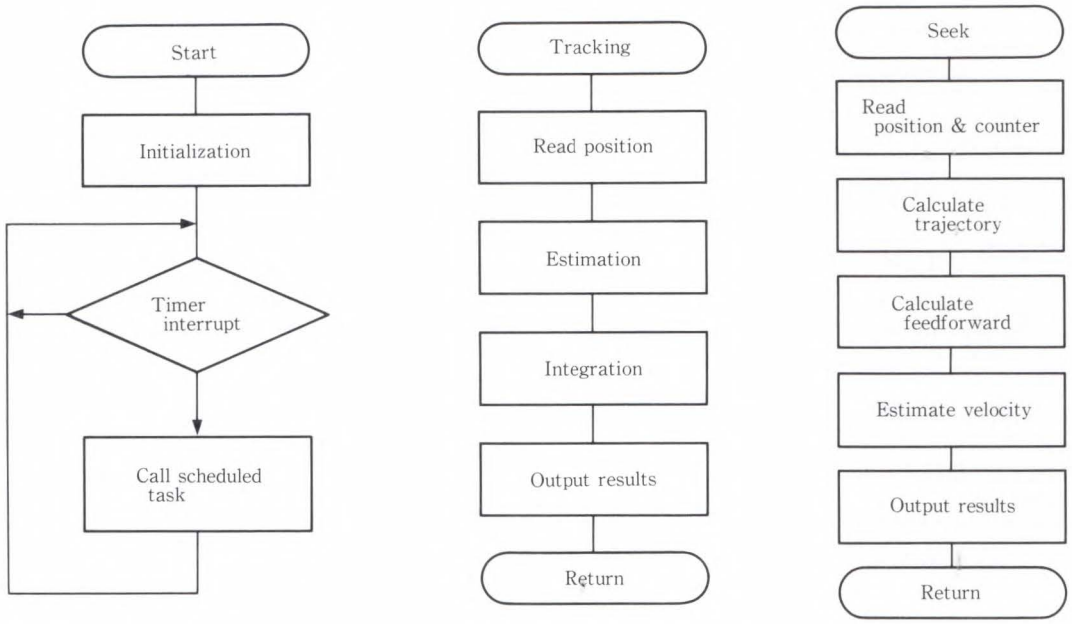
5.3 Tracking control results

Figure 15 shows the open loop characteristics for the state space control achieved by a digital control using the DSP. This figure shows that the results match the simulation well (see Fig. 8) and that a stable system (phase margin of about 48 °) can be obtained.

For the purpose of comparison with the simulation, the notch or low-pass filter for suppressing the gain of the high frequency band resonance was not applied to the system.

5.4 Access control results

Figure 16 shows the trajectory of each state value for the access operation (SMART control) achieved using DSP. In this figure, normalized time t/T_0 is determined from normalized distance x/a using the memory table. The diagram shows that t/T_0 increases linearly during access operation and that the head moves on the SMART trajectory as specified



a) Back ground task b) Scheduled task

Fig. 14—Basic flowchart of digital servo system.

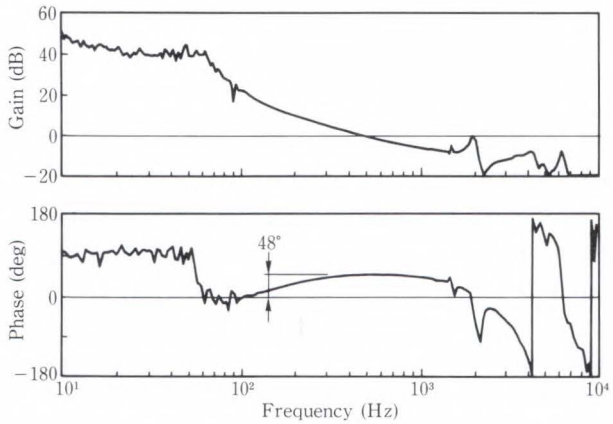


Fig. 15—Open loop transfer function using state estimator (experiment).

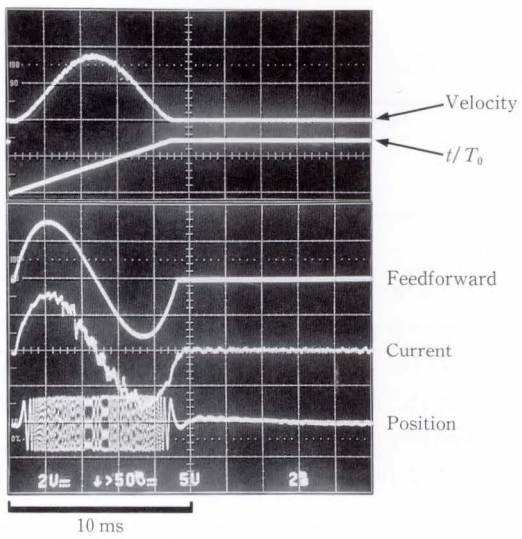


Fig. 16—Experiment with SMART trajectory (average track seek).

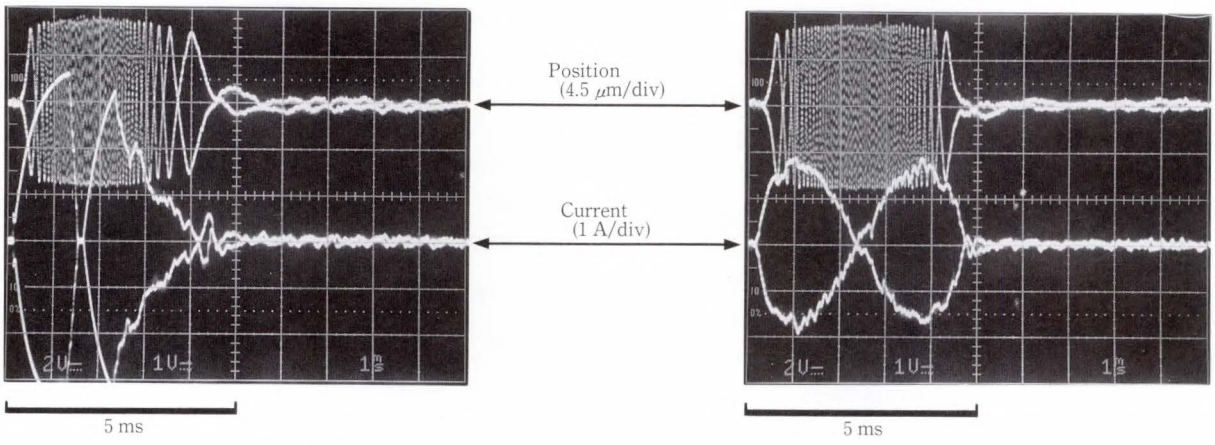
in the theory. The figure also shows that the power amplifier is not saturated during seek operation.

5.4.1 Residual vibration

Figure 17 compares the positioning error of the SMART for access operation. This example compares the access time under the same conditions in both control systems (100 tracks access). The figure shows that the residual vibration is suppressed to less than ±0.5 μm by the SMART control.

5.4.2 Flying height fluctuation

In the SMART control system, because the driving current varies smoothly, the fluctuation of head flying height is suppressed during access operation. Figure 18 is an example for the 8-track seek, to show that the flying



a) Conventional trajectory b) SMART

Fig. 17—Comparison of residual vibration.

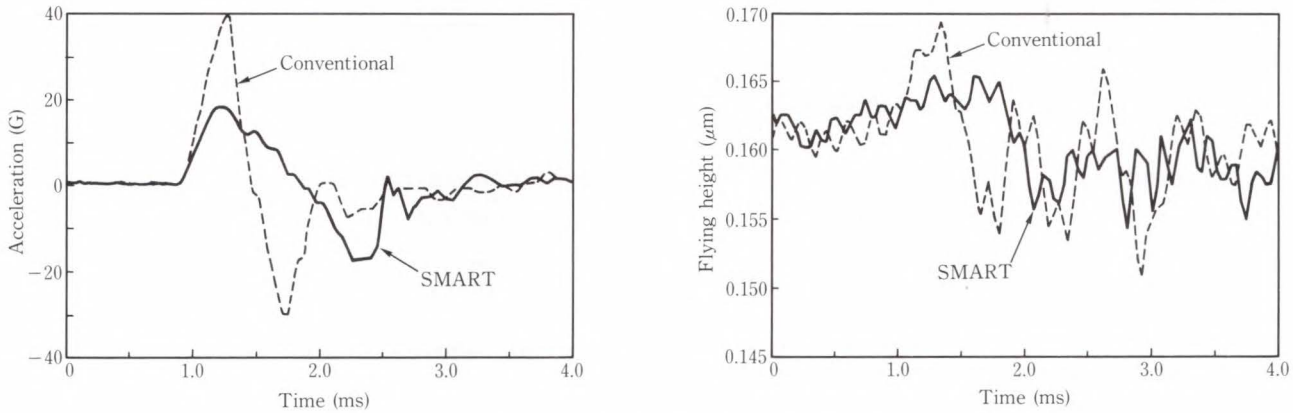


Fig. 18—Comparison of flying height fluctuation.

height fluctuation of the head is reduced to about half as much as previously achieved.

6. Conclusion

The authors have developed a digital servo control system that enables high-speed access and highly accurate positioning for the head positioning of magnetic disk drives.

This paper has described the state space control system that uses a state estimator as the design method of the track following control and access control.

This paper has also reported the design of a new access control system called SMART. The SMART control prevents resonance of the mechanical parts and enables high-speed and highly accurate positioning.

The trend toward high-recording-density magnetic disk drives will continue. The track density will increase from the current 2 000 tpi (track per inch) to 4 000-8 000 tpi within the next decade.

To increase the positioning accuracy in order to handle the above densities, the head positioning mechanism will change from a current mechanism, where plural data heads are moved using the single actuator, to piggy-back actuators that independently control the data heads using piezo-actuators. Therefore, the servo control system will change from the conventional single servo loop control to the double servo loop control that extends the servo bands up to kilohertz.

As for the control circuit, the execution

speed of the one-chip DSP with a large amount of memory will increase up to about three times the current speed. As the cost of the DSP is reduced, it is expected that the application of the intelligent digital control will extend from high-end disk drives to low-end disk drives.

For the control features, intelligent functions that make use of the advantages of digital control will advance. The automatic adjustment of servo parameters or learning control will be provided to compensate the fluctuation of Bl (motor force constant) and the automatic detection of the resonance frequency of the mechanical parts. The servo circuit monitors changes in the external environment, such as temperature, and power voltages. Adaptive control that includes the model of the plant and the parameter adjustment mechanism will be introduced in many cases.

7. Appendices

Appendix 1: Tracking control model

The discrete state equation for the control model shown in Fig. 7 are expressed in the matrices form:

$$x(k+1) = \boldsymbol{\Phi} x(k) + \boldsymbol{\Gamma} u(k),$$
$$y(k+1) = \boldsymbol{H} x(k+1),$$

where $u(k)$ is the control input. The state vector becomes:

$$x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_4(k) \end{bmatrix},$$

where, position: $x_1(k)$, velocity: $x_2(k)$, and bias force: $x_4(k)$.

Output vector $\boldsymbol{H}$ becomes:

$$\boldsymbol{H} = [\,1\ 0\ 0\,].$$

Discrete system matrix $\boldsymbol{\Phi}$ and discrete control input matrix $\boldsymbol{\Gamma}$ appear as:

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & 1 & 1/2 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{\Gamma} = \begin{bmatrix} T^2 (\dfrac{Bl}{2m}) K_c \\ T^2 (\dfrac{Bl}{m}) K_c \\ 0 \end{bmatrix},$$

where, T is the sampling period and K_c is the position gain. Also Bl is motor force constant and m is moving mass.

The following conversions are used:

$$x_2(k) \leftarrow T x_2(k), \quad x_4(k) \leftarrow T^2 x_4(k).$$

The characteristics of the device used in the experiement are listed in Table A 1.

So, the equation of the sate estimator becomes:

$$\hat{x}(k+1) = [\boldsymbol{\Phi} - \boldsymbol{L}\boldsymbol{H}]\,\hat{x}(k) + \boldsymbol{\Gamma} u(k) + \boldsymbol{L} x(k),$$

where, $\hat{}$ means an estimated value. $\boldsymbol{L} = [L_1 L_2 L_4]^T$ is the correction coefficient matrix, and is determined according to the design response frequency of the estimator.

Appendix 2: Discrete type optimal regulator

The state equation becomes:

$$x(k+1) = \boldsymbol{\Phi} x(k) + \boldsymbol{\Gamma} u(k),$$

where, integrated position: $x_0(k)$, position: $x_1(k)$, and velocity: $x_2(k)$.

The control input that minimizes cost function J is calculated from the above equation by obtaining the solution of Riccati's steady-state equation.

$$J = \sum_{k=0}^{\infty} \{\, x(k)^T \boldsymbol{Q} x(k) + u(k)^T \boldsymbol{R} u(k) \,\}.$$

In the above equation, matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ become:

$$\boldsymbol{Q} = \begin{bmatrix} {\omega_0}^3 & 0 & 0 \\ 0 & {\omega_0}^2\sqrt{\gamma - 2\sqrt{\gamma}} & 0 \\ 0 & 0 & \omega_0\sqrt{\gamma - 2\sqrt{\gamma}} \end{bmatrix}^2,$$

$$\boldsymbol{R} = (\frac{Bl}{m})^2 .$$

Riccati's steady-state equation appears as:

$$\boldsymbol{P} = \boldsymbol{\Phi}^T \boldsymbol{P} \boldsymbol{\Phi} + \boldsymbol{Q} - \boldsymbol{\Phi}^T \boldsymbol{P} \boldsymbol{\Gamma} (\boldsymbol{R} + \boldsymbol{\Gamma}^T \boldsymbol{P} \boldsymbol{\Gamma})^{-1} \boldsymbol{\Gamma}^T \boldsymbol{P} \boldsymbol{\Phi}.$$

As a result, the control input of the optimal

Table A1. Characteristics of the control actuator

Bl	Motor force constant	2.1	N/A
m	Moving mass	0.010 1	kg
T	Sampling period	40×10^{-6}	s
K_c	Position gain	$= K_d K_p / T_{pm}$	A/m
K_d	Demodulation gain	10	V/track
T_{pm}	Track pitch	12×10^{-6}	m/track
K_p	Power amplifier gain	1	A/V

regulator becomes:

$$u = -(\boldsymbol{R} + \Gamma^T \boldsymbol{P} \Gamma)^{-1} \Gamma^T \boldsymbol{P} \boldsymbol{\Phi} x(k).$$

Appendix 3: Normalization of the SMART trajectory

First, a third order actuator model in which the differential value $(d\alpha/dt)$ of acceleration (α) is applied to a double integrator model (see Fig. A 1).

The state equation appears as:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0\,1\,0 \\ 0\,0\,1 \\ 0\,0\,0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \dfrac{Bl}{m} \end{bmatrix} u,$$

where, position: x_1, velocity: x_2, and acceleration: $x_3 = \alpha$.

The cost function becomes:

$$P = \int_0^{T_0} u^2 dt.$$

Assuming that the access time is T_0 and the access distance is a, the boundary conditions appear as:

$$x(0) = \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \quad x(T_0) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Upon using adjoining state vector $\boldsymbol{P}$, the Hamiltonian H becomes:

$$H = \boldsymbol{P}^T(\boldsymbol{A}x + \boldsymbol{B}u) + u^2,$$

here,

$$\boldsymbol{A} = \begin{bmatrix} 0\,1\,0 \\ 0\,0\,1 \\ 0\,0\,0 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 0 \\ 0 \\ \dfrac{Bl}{m} \end{bmatrix}.$$

Therefore, optimal control input $u(t)$ appears as:

$$u = -\frac{1}{2}\boldsymbol{B}^T \boldsymbol{P}.$$

And, the canonical equations of motion become:

$$\begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix} = \boldsymbol{D} \begin{bmatrix} x \\ p \end{bmatrix}, \quad \boldsymbol{D} = \begin{bmatrix} \boldsymbol{A} & -\dfrac{1}{2}\boldsymbol{B}^T \\ 0 & -\boldsymbol{A}^T \end{bmatrix}.$$

Assume that the characteristic equations of the $\boldsymbol{A}$ and $\boldsymbol{D}$ are $g\boldsymbol{A}$ (s) and $g\boldsymbol{D}$ (s). Their relationship becomes:

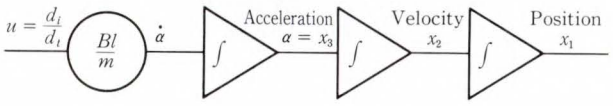


Fig. A1—Third order model of the actuator.

$$g\boldsymbol{D}(\mathrm{s}) = g\boldsymbol{A}(\mathrm{s})\, g\boldsymbol{A}(-\mathrm{s}),$$

$$g\boldsymbol{D}(\mathrm{s}) = -\mathrm{s}^6.$$

Therefore, the eigenvalue of Hamiltonian matrix $\boldsymbol{D}$ becomes:

$$\mathrm{s} = 0 \text{ (sextuple root)}.$$

Finally, assume that C_1 to C_6 are unknown numbers. The optimal solution is written below:

$$x_1(t) = C_1 t^5 + C_2 t^4 + C_3 t^3 + C_4 t^2 + C_5 t + C_6,$$

$$x_2(t) = 5C_1 t^4 + 4C_2 t^3 + 3C_3 t^2 + 2C_4 t + C_5,$$

$$x_3(t) = 20C_1 t^3 + 12C_2 t^2 + 6C_3 t + 2C_4.$$

C_1 to C_6 are determined according to the boundary conditions. In the end, the acceleration, velocity and position are presented as functions of normalized time (t/T_0) as listed below:

$$x_1(t) = -60a \left\{ \frac{1}{10}\left(\frac{t}{T_0}\right)^5 - \frac{1}{4}\left(\frac{t}{T_0}\right)^4 + \frac{1}{6}\left(\frac{t}{T_0}\right)^3 \right\},$$

$$x_2(t) = -60\frac{a}{T_0} \left\{ \frac{1}{2}\left(\frac{t}{T_0}\right)^4 - \left(\frac{t}{T_0}\right)^3 + \frac{1}{2}\left(\frac{t}{T_0}\right)^2 \right\},$$

$$x_3(t) = -60\frac{a}{T_0^2} \left\{ 2\left(\frac{t}{T_0}\right)^3 - 3\left(\frac{t}{T_0}\right)^2 + \left(\frac{t}{T_0}\right) \right\}.$$

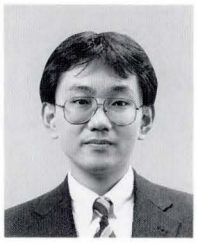
References

1) Hasegawa, S., Mizoshita, Y., and Ueno, T.: Design of Head-Positioning Controller Using DSP. (in Japanese), Trans. 944th Symp., JSME, 1987, pp. 16-18.

2) Mizoshita, Y., and Futamata, A.: Head-Positioning Servo Design for Disk Drives. *FUJITSU Sci. Tech. J.*, **18**, 1, pp. 101-105 (1982).
3) Franklin, G., Powell, J., and Workman, M.L.: Digital Control of Dynamic Systems. 2nd ed., Addison-Wesley, 1990, pp. 703-747.
4) Stich, M.C.: Digital Servo Algorithm for Disk Actuator Control. Conf. Appl. Motion Control, 1987, pp. 35-41.
5) Hasegawa, S., Mizoshita, Y., Ueno, T., and Takaishi, K.: Fast Access Control of Head Positioning Using a Digital Signal Processor. Proc. SPIE/SPSE Symp. Stor. Retriev. Syst. Appl., **1248**, 1990, pp. 104-113.
6) Mizukami, M.: Mechanical Interaction Suppression in a Dual-Actuator by Trapezoidal Driving. Proc. Incremental Motion Control Syst. Devices, 1987, pp. 321-326.
7) Cooper, E.S.: MINIMIZING POWER DISSIPATION IN A DISK FILE ACTUATOR. *IEEE Trans. Magnetics*, **24**, 3, pp. 2081-2091 (1988).
8) Meckel, P., and Seering, W.: Active Damping in a Three-Axis Robotic Manipulator. *Trans. ASME, J. VASRD*, **107**, 1, pp. 38-46 (1985).

Susumu Hasegawa
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Mechanical Eng.
Keio University 1978
Master of Mechanical Eng.
Keio University 1980
Specializing in Control Engineering

Yoshifumi Mizoshita
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Precision Eng.
Hiroshima University 1972
Master of Precision Eng.
Hiroshima University 1974
Specializing in Mechanical Control Engineering



Kazuhiko Takaishi
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Electronic Eng.
Doshisha University 1989
Specializing in Control Engineering

UDC 538.975:681.327.634

Signal Processing for High Density Magnetic Recording

• Takashi Aikawa • Hiroshi Mutoh • Takao Sugawara

(Manuscript received June 4, 1990)

This paper introduces some of Fujitsu's developments in signal processing for high density magnetic recording. First, signal processing systems and a new method of signal simulation are explained. Then, developments made in coding methods and waveform equalization are discussed. Finally, noise reduction and future developments (e.g. the use of Viterbi detection) are discussed.

1. Introduction

Recently, there have been remarkable increases in the capacity and operation speed of magnetic disk drives. This rapid progress is largely due to improvements in high-density recording and high-speed data transfer. In the last ten years, areal density has been increased ten times of 10-100 Mbit/inch^2. As the recording area for each bit becomes smaller, the signal becomes buried in noise.

The progress made in signal processing technology has contributed to improvements in the reading of information. It can be said that this contribution has mainly been due to new coding and waveform equalization technologies. Coding technology has progressed from the MFM coding to 2-7 coding, and then to 1-7 coding. In the field of waveform equalization technology, pulse slimming by a cosine equalizer was introduced. In the last five years, thin-film heads have been put into use, and the double-cosine equalizer for these new heads was developed.

In line with the development of these coding techniques and equalizing circuits, Fujitsu has developed a simulation technique. This simulation technique is used to develop various signal processing methods and design optimum circuit constants.

The main topic of this paper is the technology incorporated in a new 100 Mbit/inch^2 disk drive. This drive has an areal density of 100 Mbit/inch^2, a linear density of 48 000 BPI (bit/inch), a track density of 2 100 TPI (track/inch), and a data transfer rate of 4.76 Mbyte/s.

2. Signal processing circuit systems and simulation technology

This chapter outlines the signal processing systems used in magnetic disk drives, and describes the margin simulation technology developed by Fujitsu.

Figure 1 outlines the widely used phase discrimination method of recording and retrieving information on magnetic disks. Each peak of the readout waveform corresponds to a written 1. If the pulse corresponding to a written peak (peak pulse) is within the period of the window pulse, the data is determined to be 1. If it is outside the window, the data is determined to be 0.

To obtain peak pulses, the readout waveform is differentiated and the zero-crossing positions of the signal are detected. The peak pulses are shifted from their original positions because of various factors. One of these factors is jitter noise, which is indicated by ΔT_j in Fig. 1. Another factor is the peak shift due to interference between adjacent output

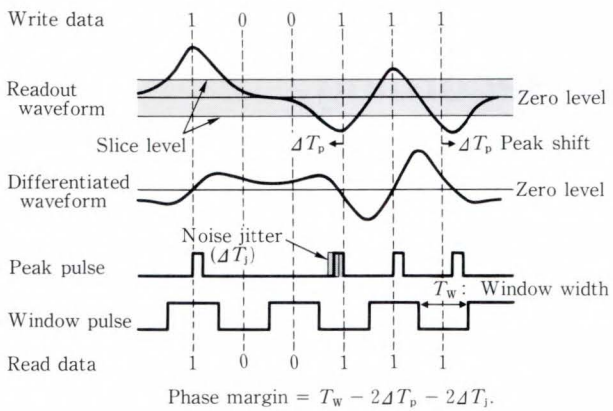


Fig. 1—Peak detection.

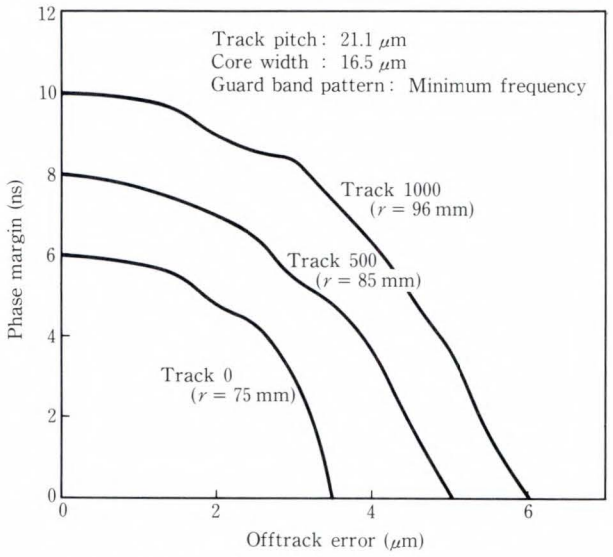


Fig. 2—Offtrack dependence of phase margin.

signals. The peak shift is indicated by ΔT_p in Fig. 1. If a peak pulse is shifted by the above factors until the pulse falls within the period of the adjacent window pulse, an error occurs. Without error correction, the error probability on a current disk drive is 10^{-9} or less.

When the recording density is increased, the resultant reductions in the signal-to-noise ratio and resolution may cause jitter noise or peak shift. These effects may also be caused by signal deteriorations due to positioning error (offtrack error).

To compensate for signal deterioration due to offtrack errors, signal processing systems are designed with safety margins. One of these margins, the phase margin, is shown in Fig. 1. Figure 2 shows the result of an experiment on the relationship between the phase margin and offtrack error. If the phase margin is increased, the offtrack tolerance also increases.

Consequently, enlargement of the phase margin is a key point for signal processing design.

As described above, a magnetic disk drive regards the zero-crossing point of the differentiated waveform to indicate a 1 . Therefore, if consecutive 0s are written, a peak pulse is generated not at a peak in the readout waveform but at a noise peak. To distinguish such a noise-generated peak from a real peak, the original waveform is sliced at a specific level and the amplitudes are then observed (see Fig. 1). If the range of slice level setting used to read information without error is wide, the signals will be of good quality. This range is called the level margin. Both the phase margin and the level margin are important factors in the reliability of a disk drive.

Figure 3 shows the block diagram of a readout channel[1]. The automatic gain control (AGC) circuit maintains the signal amplitude at a specified level. The waveform equalizer (described in chapter 4) reduces the peak shift. The filter circuit removes high-frequency noise from the signal and improves the signal-to-noise ratio. Peak positions are determined by both peak detection and amplitude detection.

The phase and level margins are closely related to the signal-to-noise ratio, resolution, and differentiator constant. These margins are also closely related to the coding, and waveform equalization. Simulation technology is an important tool used to obtain the optimum constants of these circuits and the optimum signal processing method[1)-3)].

Figure 4 shows the block diagram of a newly developed margin simulator. The phase and level margins are calculated from an isolated waveform, pattern data, write compensation (described later), noise spectrum, circuit type (equalizer or filter), and circuit constants. The waveform pattern created from the isolated waveform and data pattern is converted to frequency domain information using a fast

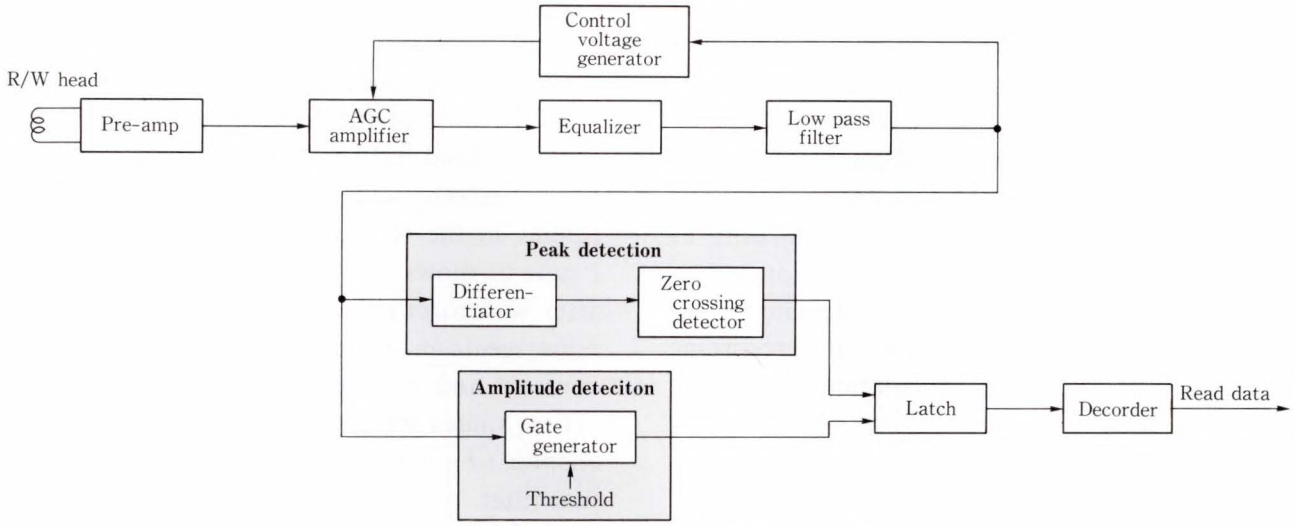


Fig. 3—Block diagram of readout channel[1].

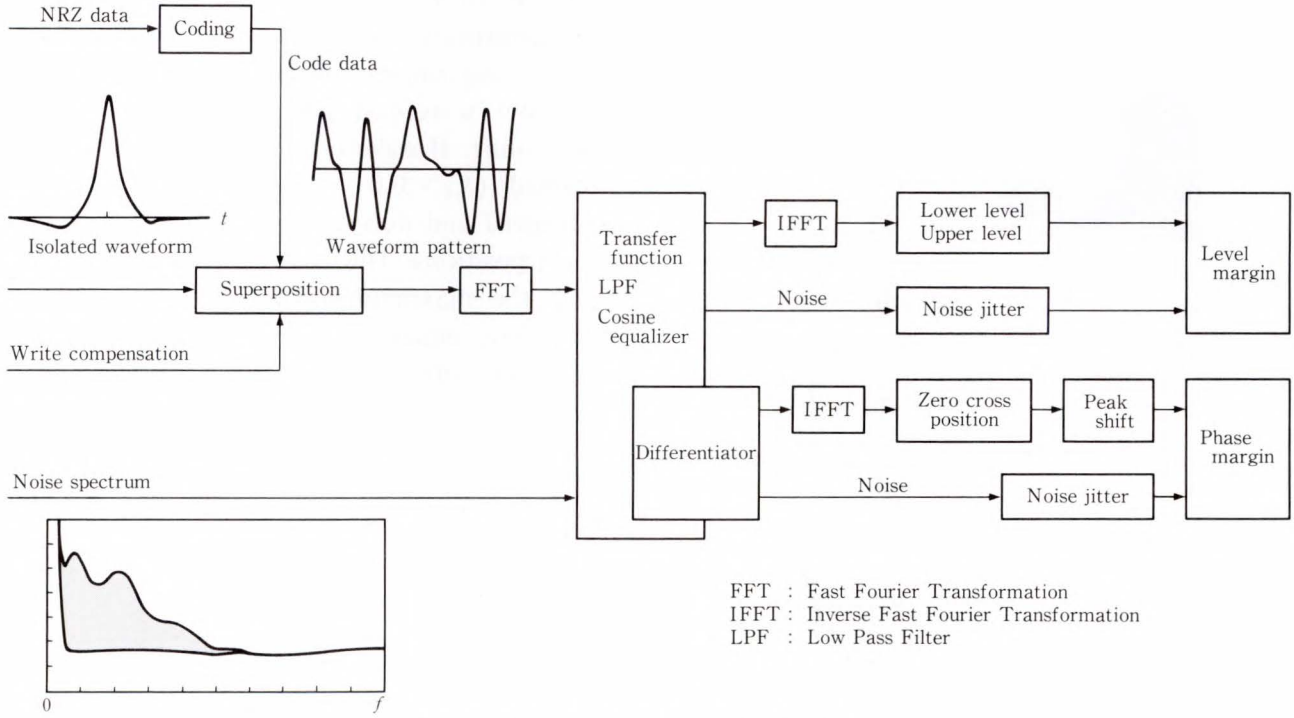


Fig. 4—Block diagram of margin simulator.

Fourier transformatron (FET). To obtain the margins, the converted information and noise spectrum are used to calculate the differentiated waveform, peak shift, and noise jitter. In these calculations, the frequency domain information is converted to time domain information using an inverse fast Fourier transform (IFFT). In the simulation, noise jitter is assumed to have a Gaussian distribution.

The features of this simulator are as follows:

1) An arbitrary isolated waveform can be input into the simulation from an outside source or from a calculation using special functions of the program. The simulator can also accept isolated waveforms from a digitizer or sampling oscilloscope, thus allowing simulation using waveforms from an actual head.
2) The gain and phase response of the real circuits can be entered into the simula-

tion. This enables simulation of the effects of characteristics that cannot be represented by numeric expressions. For example, the phase characteristics of the delay line used in a waveform equalizer can be input to calculate its effect on margins.

3) The output waveform and peak shift at an arbitrary position (i.e. a position before or after equalization) can be calculated.

Figure 5 shows an example of a recording density characteristics calculation based on isolated waveform data from a digitizer. This example simulates the characteristics of the thin-film magnetic head of a 100 Mbit/inch^2 disk drive. Because of the finite pole-face length, the thin-film head has negative side-lobes on its isolated readout signal. These side-lobes induce ripples in the recording density characteristics. Figure 6 shows the measured noise spectrum after waveform equalization, and the calculated noise spectrum. In this figure, K_1 is an equalizer constant and τ_1 is the delay of the delay line. (These values are described in detail in the next chapter.) F_c is the cut-off frequency of the low pass filter.

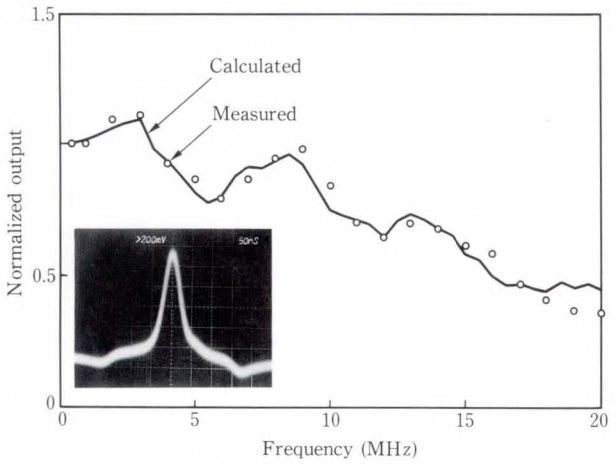


Fig. 5—Frequency characteristics simulated for an isolated waveform.

The noise spectrum is measured at $K_1 = 0.5$. In the calculation, K_1 is changed from 0.0 to 0.6. As shown in Figs. 5 and 6, the simulated characteristics closely match the actual results of the experiment. Figure 7 shows an example of how waveform equalization improves the peak shift. If waveform equalization is not performed ($K_1 = 0.0$), the peaks are alternately advanced and delayed from their original (correct) positions. The peak shift is about 3.8 ns, and the maximum noise jitter is ±3.8 ns. If waveform equalization is performed ($K_1 = 0.5$), the peak shift is reduced to 1.6 ns. This remain-

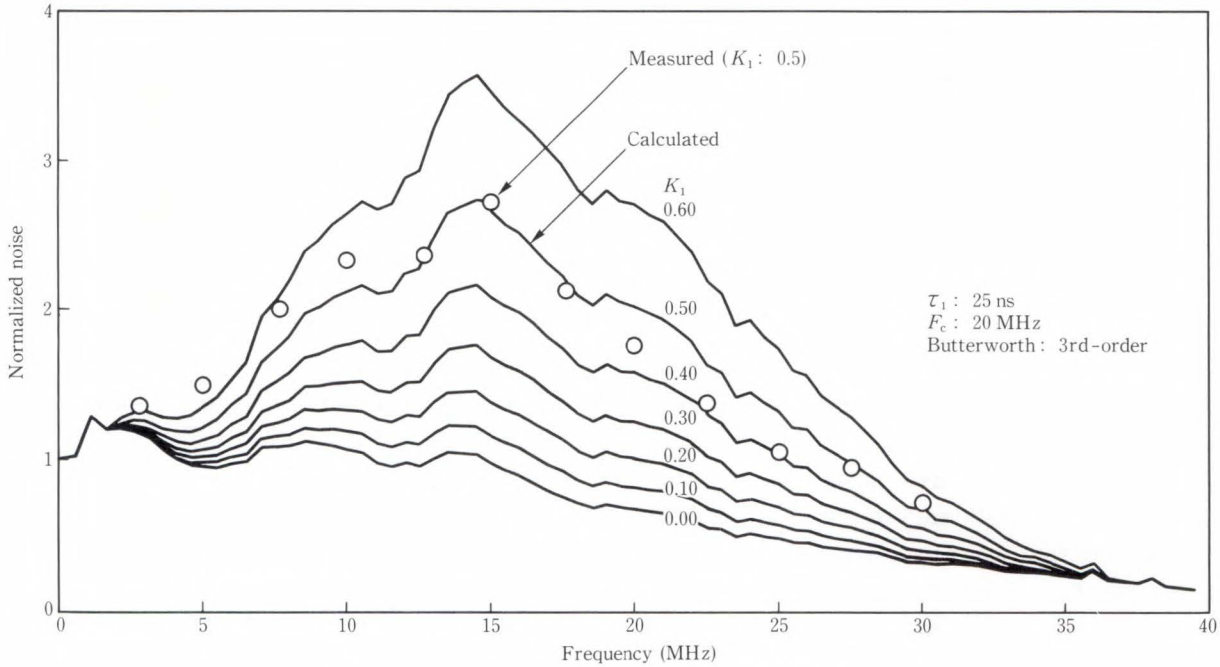


Fig. 6—Example of noise spectrum calculation.

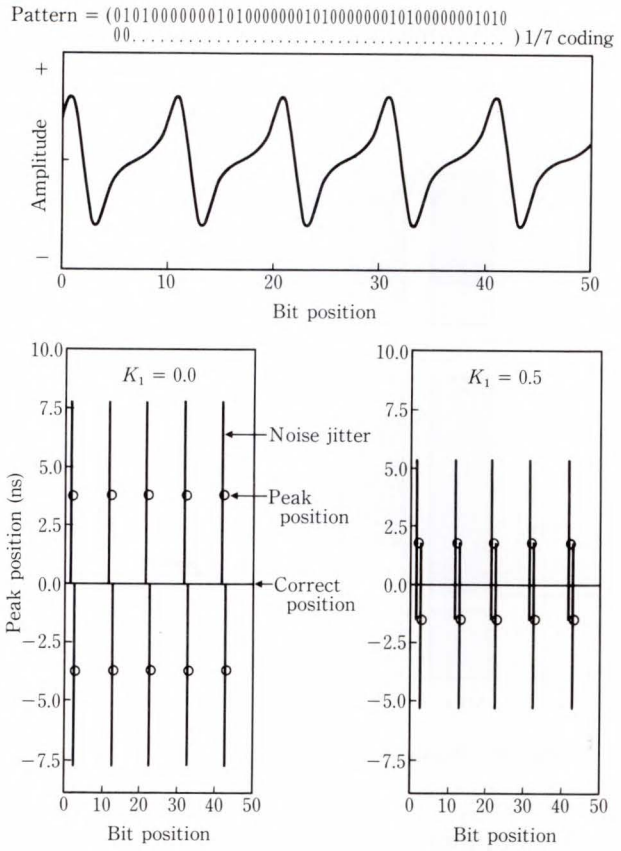


Fig. 7—Example of peak shift calculation.

ing peak shift is mainly due to the negative side-lobes of the thin film head. The calculation results described in the following chapters are based on this simulator.

The following is a brief explanation of the write compensation shown in Fig. 4.

The write circuit estimates the timing of peak shifts, and reduces or enlarges the intervals between the bits to be written. Figure 8 shows the write current and readout waveforms for write pattern 0 ... 0110 ... 0. Because of the intersignal interference, the interval between the two ls is extended. Therefore, the duration of the write current pulse is decreased during information recording. However, writing at reduced intervals increases the maximum recording density on the disk. This results in increased interference from output signals and reduced signal output. Therefore, current disk drives combine write compensation and waveform equalization to suppress the peak shift due to interference.

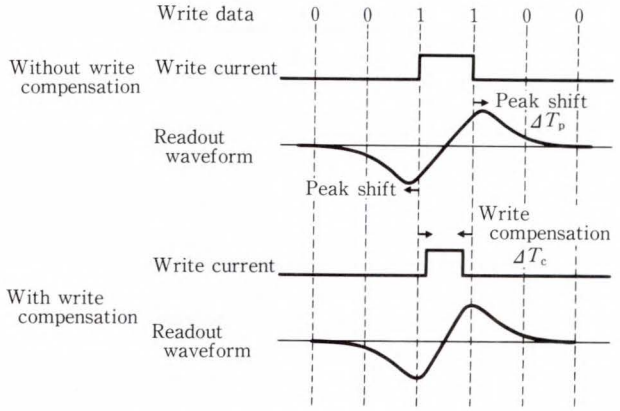


Fig. 8—Write compensation.

3. Coding methods

In general, a magnetic disk drive does not record information in the format in which it was sent from the upper-level device. When recording information is received, this information is converted according to certain rules.

The main requirements for coding are as follows:

1) The minimum time interval between two transitions (T_{min}) to be recorded on disk should be as long as possible.
2) The maximum time interval between two transitions (T_{max}) should be short enough to enable self-clocking.
3) The window width (T_w) should be large.

Table 1 shows the main coding methods. Coding parameters are usually represented in the format of "d, k, m, n". In this format, "d" indicates the minimum number of consecutive zeros in the converted information string, "k" the maximum number of consecutive zeros. "m" the number of bits to be converted, and "n" the number of converted bits. The 1-7 coding method converts 2 bits to 3 bits. The minimum and maximum number of consecutive zeros in the converted bits are 1 and 7 respectively. If the data bit time interval before conversion is T, the data bit time interval after conversion is $2T/3$. This value is equal to the window width (T_w). The minimum time interval between two transitions (T_{min}) is $4T/3$. The T_{min} value is 1.33 times larger than the data bit time interval (T) before conversion. This means that the maximum recording density on the

Table 1. Examples of modulation codes

Coding	T_{min}/T	T_{max}/T	T_w/T	$(d\ k\ m\ n)$
NRZI	1	∞	0.5	0 ∞ 1 1
MFM	1	2	0.5	1 3 1 2
8/9	0.89	3.56	0.89	0 3 8 9
2-7	1.5	4	0.5	2 7 2 4
1-7	1.33	5.33	0.67	1 7 2 3

T : Data bit time interval
T_{min} : Minimum time interval between two transitions
T_{max} : Maximum time interval between two transitions
T_w : Data detection window width
d : Minimum number of consecutive zeros
k : Maximum number of consecutive zeros

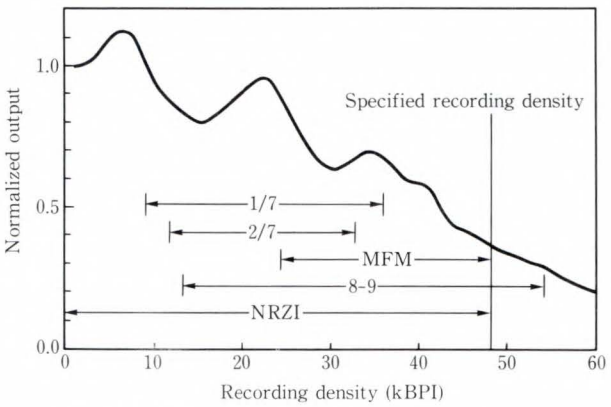


Fig. 9—Recording densities of coding methods.

disk can be decreased to 1/1.33 of the original recording density.

Figure 9 shows the recording density regions of four codings. In this figure, the recording density before conversion is 48 000 BPI and is indicated as the specified recording density. The recording density region depends on the type of coding. The maximum recording density for modified frequency modulation (MFM) coding is the specified recording density. The minimum recording density is half the maximum recording density. For the 1-7 coding described above, the maximum recording density is 36 000 (48 000/1.333) FCI (fluxchanges/inch), and the minimum recording density is 9 000 (36 000/4) FCI. Thus, 1-7 coding and 2-7 coding use lower recording densities than MFM coding when recording on disks.

Figure 9 suggestes that 2-7 coding may be

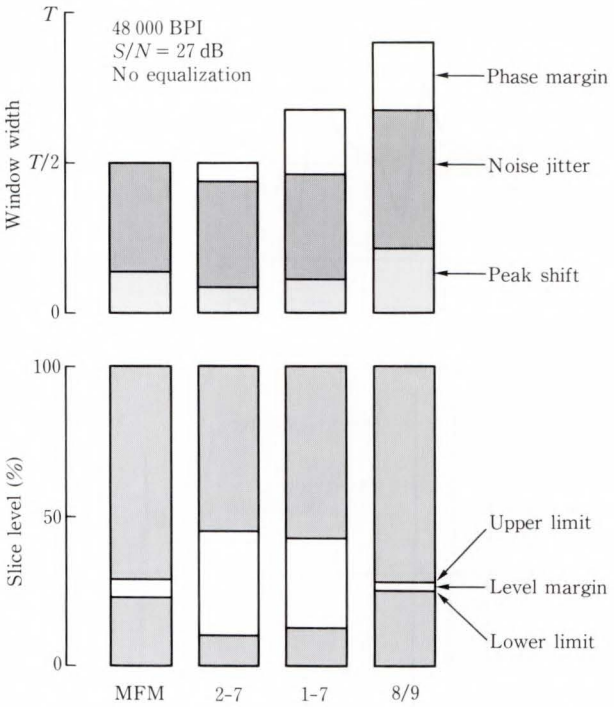


Fig. 10—Phase and level margins of coding methods.

the optimum method because the region between the minimum and maximum recording densities is narrow and the maximum recording density is low. If the window width is considered, the 1-7 coding may be advantageous. However, a single coding method cannot satisfy all the requirements, and optimum coding depends on the head and disk characteristics. To select a coding method, simulation and experiments are required.

Figure 10 compares the performance of MFM, 2-7, 1-7, and 8/9 codings. This figure indicates the phase and level margins for an error rate of 10^{-9}. The frequency band of the signal reproduction system is assumed to be twice the maximum read-write frequency.

The upper graph shows the phase margins. In the graph, T is the data bit time interval before conversion, and the height of each bar indicates the window width. In the reproduction system used for the simulation, it was found that noise jitter has a large influence regardless of the type of coding. When MFM coding is used, the phase margin is zero. When 8/9 coding is used, the window is wide but noise jitter has a large influence. In addition, because the mini-

mum time interval between two transitions is small, the peak shift is large and the phase margin is small. There is little difference between 2-7 and 1-7 codings in terms of peak shift and noise jitter. Because 1-7 coding provides a wider window than 2-7 coding, it can allows a larger phase margin.

The lower graph shows the level margins. Because MFM and 8/9 coding have the smallest time interval between two transitions, they also have the smallest level margins. This is because the output signal is low due to intersignal interference. On the other hand, 1-7 and 2-7 codings use a comparatively large minimum time interval between two transitions. Thus these codings can have higher level margins because there is less interference between adjacent signals.

Of these four coding methods, 2-7 and 1-7 codings have optimum balance between phase and level margins. (These two methods are the most common methods in current use.) The 1-7 coding is being used in recent Fujitsu models because it allows a wide window width for high-speed data transfer. For example, the 100 Mbit/inch^2 disk drive has a data transfer rate of 4.76 Mbyte/s and a window width of 17.5 ns.

4. Waveform equalization

Waveform equalization slims the readout waveform (see Fig. 11) in order to reduce the peak shift and interference from adjacent waveforms. Figure 11 shows the circuit of the reflecting-type cosine equalizer used for waveform equalization. The delay line is terminated with matching impedance Z_0 at the input.

The transfer function of this equalizer is as follows:

$$H(\omega) = 1 - K_1 \times \cos(\omega\tau_1),$$

K_1 : Equalization constant

τ_1 : Delay of delay line.

This circuit is generally called a cosine equalizer[4]. It compensates deteriorations in the high frequency characteristics of the head and disk and produces smooth frequency characteristics in the frequency band used.

Recently, an increasing number of thin-film magnetic heads are being used to enable high-speed data transfer. The readout waveform of a thin-film head has pulses, (called negative pulses), at the positions corresponding to its pole edges. These pulses appear as ripples in the recording density characteristics (see Fig. 5). If this waveform is equalized by a conventional cosine equalizer, the ripple is enhanced and an undesired peak shift is generated. The ripple positions are closely related to the pole-face length. The thiner the pole-face length, the higher the density of the first peak in the characteristics[5]. However, the pole-face length is also closely related to the write capability of the head. If the pole-face length is too short, information cannot be written on the disk. For a thin film head, it is important to balance the read performance with the write capability, especially for high density media having a high coercive force.

The 100 Mbit/inch^2 disk drive uses 1-7 coding. The maximum and minimum recording densities for disks in this drive are 36 000 FCI and 9 000 FCI respectively. To remove peaks from this density region, the pole-face length must be 0.7 μm or less.

If a head of this length is made using currently

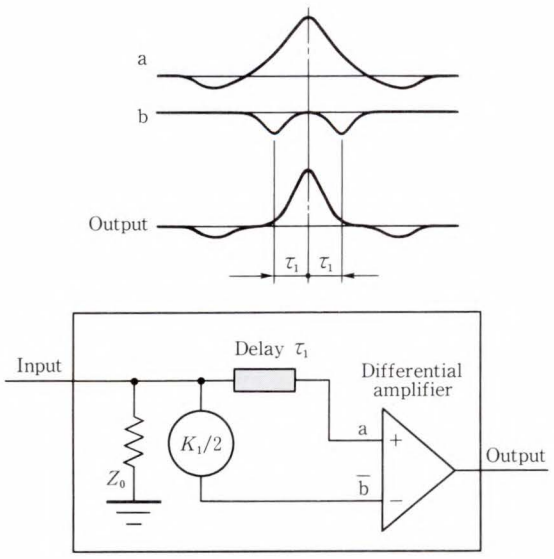


Fig. 11—Cosine equalizer.

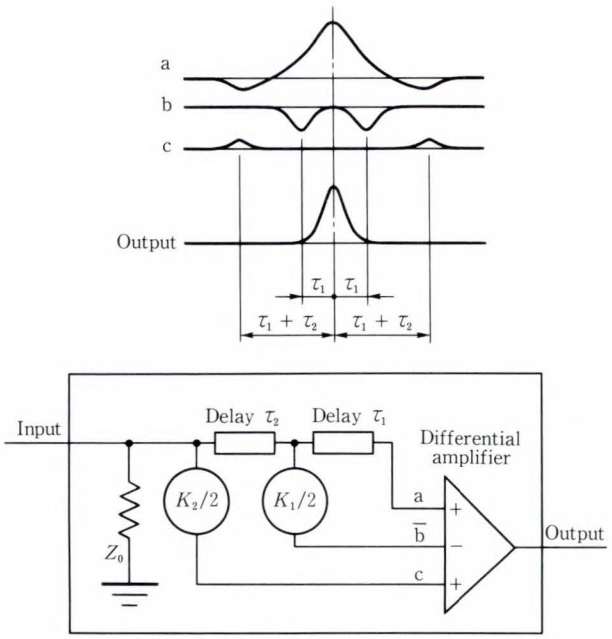


Fig. 12—Double cosine equalizer.

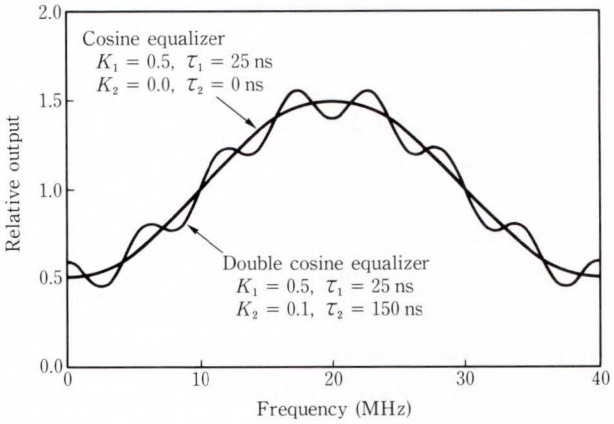


Fig. 13—Frequency characteristics of double cosine ezualizer.

used materials, the write capability is not sufficient for a high coercive force media (1 400 Oe, or 11 kA/m). To ensure sufficient write capability, the pole-face length was set at 3.5 μm; thus, three peaks appear in the region (see Fig. 5). To cope with this problem, a new equalizer was developed. The new equalizer eliminates negative pulses and reduces the width of the readout signal (see Fig. 12).

The new equalizer consists of two delay lines, two amplitude dividers, and a differential amplifier. When $K_2 = 0$ and $\tau_2 = 0$, the equalizer becomes a conventional cosine equalizer. To reduce the side-lobes, the equalizer generates reverse polarity pulses that are synchronous with the negative pulses, and then adds the two. Two delay lines $(\tau_1 + \tau_2)$ are used, the delay being equal to the time between a positive and a negative pulse of an isolated waveform.

The transfer function is as follows:

$$H(\omega) = 1 - K_1 \times \cos(\omega\tau_1) + K_2 \times \times \cos\{\omega(\tau_1 + \tau_2)\},$$

K_1, K_2 : Equalization constant

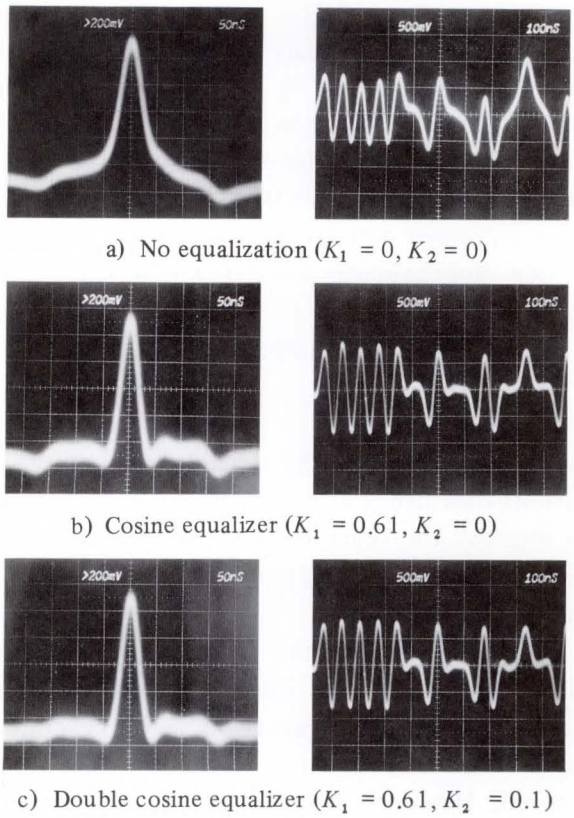
τ_1, τ_2 : Delays of delay lines.

Because of its transfer function, this new equalizer is called a double cosine equalizer.

Figure 13 shows the frequency characteristics of this equalizer. It has a peak response with symmetrical lower-amplitude variations at the sides of the peak. The peaks of these variations are designed to overlap the low points in the recording density characteristics of the head and disk.

Figure 14 shows examples of waveforms processed by the double cosine equalizer, and the effect of this design on the recording density characteristics of the 100 Mbit/inch^2 disk drive. Graph a) in Fig. 14 shows the state without equalization. Graph b) shows the state after conventional pulse slimming. Graph c) shows the state after double cosine equalization. This figure shows that the peak levels are equalized and the amplitude of variations, which determines the lower threshold level, is decreased. The double cosine equalizer also eliminates ripples in the recording density characteristics to smoothen the characteristics of the recording density region.

Figures 15 and 16 show simulations of how the double cosine equalizer improves the phase and level margins in the 100 Mbit/inch^2 disk drive. The configuration is shown in Fig. 12 with $\tau_1 = 25$ ns, and $\tau_2 = 150$ ns. The optimum value of constant K_1, which determines the pulse width, was calculated to be 0.5. The horizontal axis shows the value of K_2, which indicates the level of the waveform to be added at the negative pulse position. When K_2 is close to 0.1, both the phase and level margins are maximum.



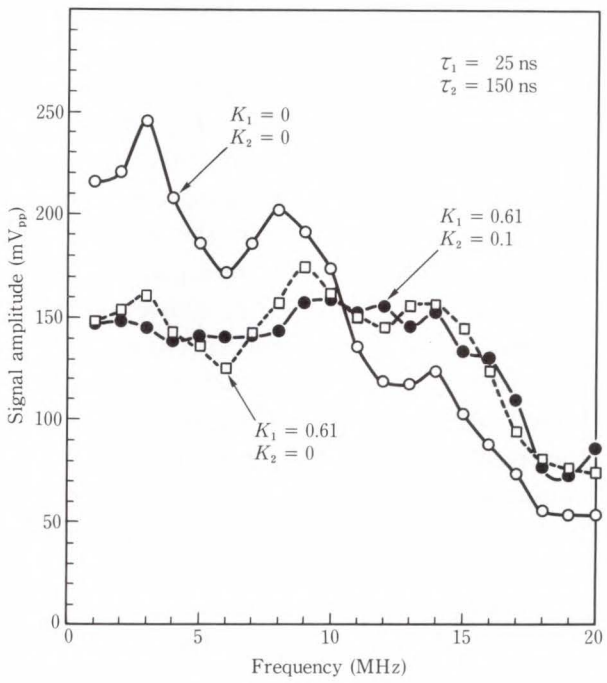


Fig. 14—Equalized waveforms and frequency characteristics.

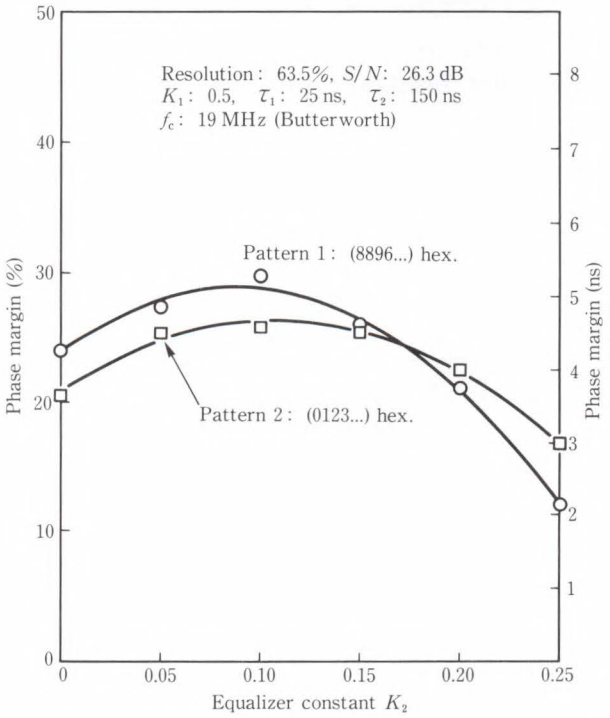


Fig. 15—Phase margin dependence on equalizer constant K_2.

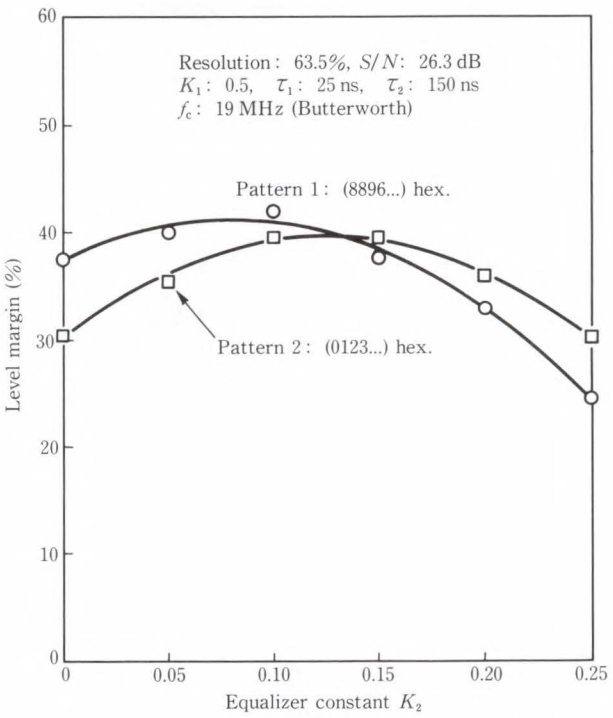


Fig. 16—Level margin dependence on equalizer constant K_2.

The double cosine equalizer increase the phase margin by about 1 ns (ten percent of the data window). It also increases the level margin by between five and ten percent.

Figure 17 shows the measured phase margins of pattern 2 (see Fig. 15). The graph shows the error rates measured at different window positions. The phase margins correspond to the time intervals between the two curves in the figure. The double cosine equalizer increases the phase margin by 1.3 ns above the margin obtained using a conventional cosine equalizer. This increase is almost equal to the value obtained in the simulation. Figure 17 also shows the result of the previously mentioned write compensation. The use of write compensation increases the phase margin by an additional 1 ns.

5. Noise reduction

This chapter discusses the noise source of disk drives.

As described above, jitter noise has a large influence on margins. Noise can be classified into the following three general categories:

Disk noise, head impedance noise, and preamplifier noise.

Figure 18 shows how preamplifier noise has been reduced. The noise voltage of the MB4114, Fujitsu's latest integrated circuit for thin film heads, is 0.54 $nV_{rms}/\sqrt{Hz}$. In comparison with other noise sources, this figure is relatively small. Figure 19 shows the noise spectrum of the 100 Mbit/inch^2 disk drive when equipped with the MB4114.

An approximate breakdown of the total noise is as follows:

Head impedance noise: 50 percent
Disk noise: 35 percent
Preamplifier noise: 15 percent

This shows that the preamplifier noise in current disk drives is relatively small. Therefore, to improve the signal-to-noise ratio, it is important to precisely analyze the head-to-preamplifier interface.

Figure 20 shows the equivalent circuit of the head/preamplifier interface. Because it is relatively small in comparison with the circuit input resistance, R_c, head resistance *rh* has been ignored. However, the number of coil turns is being increased to obtain higher head outputs. The thin-film head is regarded as having a low inductance and as being suitable for high-speed data transfer; but, thin-film heads have at least 32 turns and their inductance, at 0.5 μH to 0.7 μH, is close to that of a bulk head. Also, the resistance of these heads is about

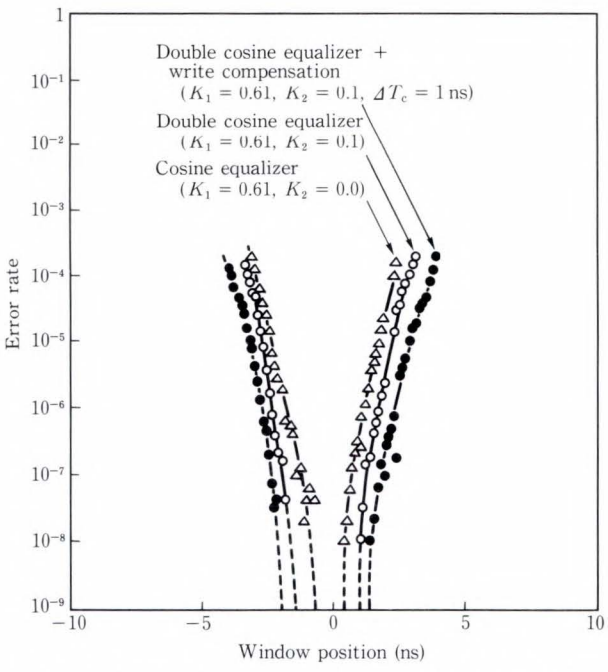


Fig. 17—Error rate characteristics of 100 Mbit/inch^2 disk drive.

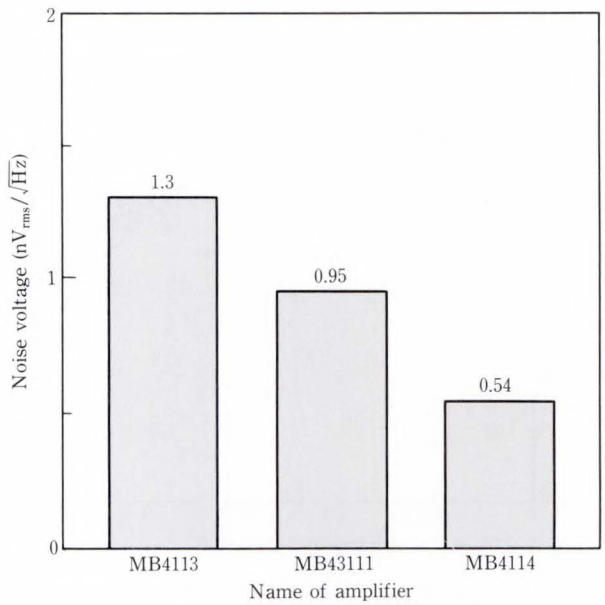


Fig. 18—Preamplifier noise.

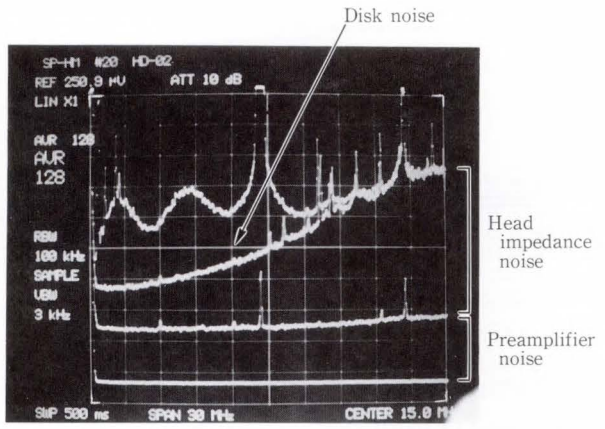
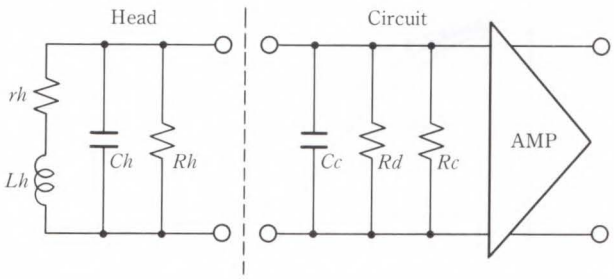


Fig. 19—Noise spectrum (100 Mbit/inch² disk drive).



$$|H(j\omega)| = \frac{1}{\sqrt{(1-(\omega/\omega_0)^2 + rh/R)^2 + 4\zeta^2(\omega/\omega_0)^2(1 + R\,rh\,C/Lh)^2}}$$

$$\theta = \tan^{-1}\frac{-2\zeta(\omega/\omega_0)(1 + rh/4\zeta^2 R)}{1-(\omega/\omega_0)^2 + rh/R},$$

$\omega_0 = \frac{1}{\sqrt{LhC}}$: Resonant angular frequency

$\zeta = \frac{1}{2R}\sqrt{\frac{Lh}{C}}$: Damping factor.

$C = Ch + Cc \qquad R = Rh /\!/ Rd /\!/ Rc$

rh : Head resistance — Rh: Head loss resistance
Lh: Head inductance — Rd: Damping resistance
Ch: Head capacitance — Rc: Circuit input resistance
Cc: Stray capacitance + circuit capacitance

Fig. 20—Equivalent circuit of head-to-preamplifier interface.

32 Ω. Therefore, it is important to take the head resistance into consideration.

The gain versus frequency characteristics of the equivalent circuit shown in Fig. 20 for two values of head resistance (0 Ω and 32 Ω) are shown in Fig. 21.

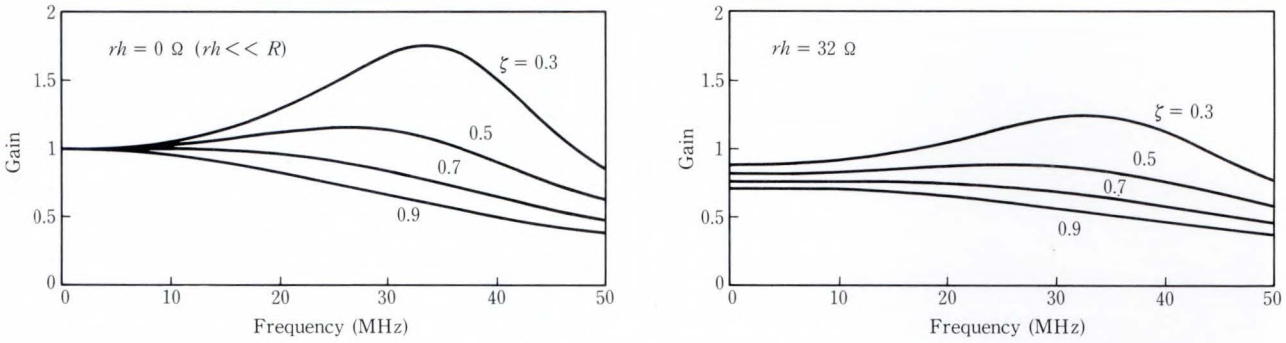
When the resistance of the thin-film head is included in simulations (i.e. $rh = 32$ Ω), it reduces the gain in the low-frequency region. This means that the head resistance reduces the transmission efficiency of head output.

When the head resistance is halved to 16 Ω, the output signal and disk noise increase by about six percent. But this change in head resistance and decreases the impedance noise by about 25 percent. This is because the thermal noise of the head resistance is reduced. The overall effect on the signal-to-noise ratio at the preamplifier output is simulated to be an improvement of about 1.5 dB. A reduction of head resistance is, therefore, a significant improvement in terms of signal processing.

6. Future development

If recording densities continue to increase, the head output will naturally fall regardless of improvements made in the head and disk system. This section briefly discusses future development regarding data detection.

The conventional peak detection system (differentiation and zero-level slice) is not very tolerant of poor signal-to-noise ratios. This fact is stimulating the application of communication technology to magnetic recording. Use of Viterbi detection, the Maximum Likelihood Sequence Detection (MLSD) for convoluted codes, is being investigated. In this method, the interference of readout signals in magnetic recording is regarded as a type of inter-code interference. However, Viterbi detection becomes impractically complex when applied to the interference between output signals having arbitrary values and durations such as are found in magnetic recording. To ease the complexity, the combination of partial response and Viterbi detection[6] and a path feedback method[7] have been suggested. The former method assigns a specific correlation between signals by waveform equalization, and the latter method reduces the restricted length by path feedback. In short, the solution is to improve the waveform equalizer or to improve the Viterbi decoder. Figure 22 shows the simulated relationship between the signal-to-noise ratio and error rate for a 68 percent resolution and 10 percent modulation. It was confirmed that the Viterbi detection method can extract information from a signal having a signal-to-noise ration of about 6 dB less than



(Lh = 0.6 μH, Rh = 350 Ω, C = 30 pF)

Fig. 21—Frequency characteristics (amplitude) of signal source.

conventional bit-by-bit detection.

Recently, areal recording densities of 1 Gbit/inch^2 have been achieved in experiments by combining partial response and Viterbi detection[8].

7. Conclusion

The progress made in signal processing over the last ten years has mainly been due to margin simulator, new coding methods (e.g. 2-7 and 1-7 coding), waveform equalization (e.g. cosine and double cosine equalization) and the reduction of pre-amplifier noise.

If recording densities continue to increase, head output will fall and disk defects will become significant. Therefore, it will become increasingly important to develop signal processing methods (e.g. Viterbi detection and new error correction systems) that can overcome poor signal-to-noise ratios and multiple disk defects.

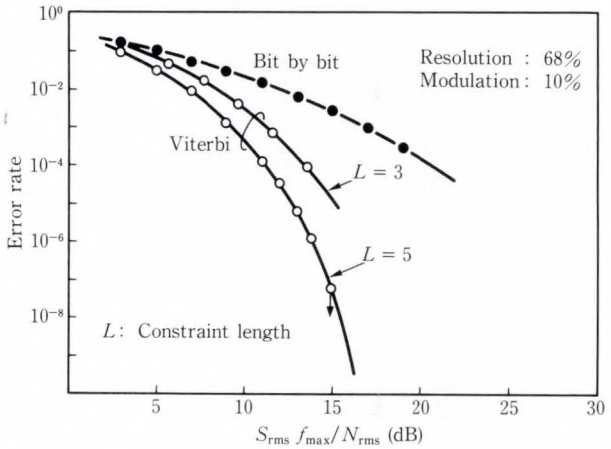


Fig. 22—Signal-to-noise ratio and error rate using Viterbi detection.

References

1) Aikawa, T., Mutoh, H., and Sugawara, T.: AN EXPERIMENTAL STUDY OF SIGNAL EQUALIZATION FOR THIN FILM HEADS. *IEEE Trans. Magnetics*, **MAG-22**, 5, pp. 1209-1211 (1986).
2) Mutoh, H., Aikawa, T., and Sasaki, M.: A STUDY OF PHASE MARGIN AND LEVEL MARGIN OF A THIN FILM HEAD. (in Japanese), Nat. Conv. Rec., 1984, IECE, Jpn., p. 1-218.
3) Mutoh, H., Aikawa, T., and Sasaki M.: A Consideration for Signal Processing of Thin Film Head. (in Japanese), IECE Tech. Rep., **MR84-3**, pp. 19-26 (1984).
4) Kameyama, T., Takanami, S., and Arai, R.: IMPROVEMENT OF RECORDING DENSITY BY MEANS OF COSINE EQUALIZER. *IEEE Trans. Magnetics*, **MAG-12**, 6, pp. 746-748 (1976).
5) Kakehi, A., Oshiki, M., Aikawa, T., Sasaki, M., and Kozai, T.: A THIN FILM HEAD FOR HIGH DENSITY RECORDING. *IEEE Trans. Magnetics*, **MAG-18**, 6, pp. 1131-1133 (1982).
6) Wood, R.W., and Petersen, D.A.: Viterbi Detection of Class IV Partial Response on a Magnetic Recording Channel. *IEEE Trans. Commun.*, **COM-34**, 5, pp. 454-461 (1986).
7) Matsushita, K., Iketani, A., Ide, A., and Yamamitsu, G.: Path Feedback Viterbi Detection without Opening Eye. (in Japanese), IECE Tech. Rep., **MR87-38**, pp. 21-27 (1987).
8) Howell, T.D., MeCown, D.P., Diola, T.A., Tang. Y-S., Hense, K.R., and Gee, R.L.: ERROR RATE PERFORMANCE OF EXPERIMENTAL GIGABIT PER SQUARE INCH RECORDING COMPONENTS. *IEEE Trans. Magnetics*, **26-5**, pp. 2298-2302 (1990).

Takashi Aikawa
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Electronics Eng.
Keio University 1972
Master of Electronics Eng.
Keio University 1974
Specializing in Magnetic Recording Systems

Hiroshi Mutoh
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Electronics Eng.
Nihon University 1980
Master of Electronics Eng.
Nihon University 1982
Specializing in Magnetic Recording

Takao Sugawara
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Electronics Eng.
Iwate University 1984
Specializing in Signal Processing in Magnetic Recording

UDC 621.318:681.327.634

Flying Head Assemblies

• Seizi Yoneoka • Takeshi Ohwe • Yoshifumi Mizoshita

(Manuscript received June 8, 1990)

The key to developing high-performance magnetic disk drives with a large capacity and a fast head access is to improve the slider's flight characteristics by modifying the head slider and head support.

Use of a negative pressure head slider can reduce wear at CSS, and use of a microslider can improve durability at the rated disk velocity while stabilizing flight.

A newly developed inline suspension spring greatly reduces head vibration during accessing, and enables fast access. It has a load beam that is stiffened in the access direction by use of a frame structure and has a new pivot that reduces the generated moment.

1. Introduction

There is a demand for magnetic disk drives that have larger storage capacities, compact structures, and a faster head access. A high-performance model with a recording density of 100 Mbit/inch^2 and an average access time of less than 10 ms has been developed.

To increase recording density, the media and its protecting layer are becoming thinner, which makes the media more fragile. Also, the probability of head/disk contact is increasing because the flying height of the head slider is being reduced. For fast access, high-performance magnetic disk drives now exert an acceleration approaching 100 G on the head slider. This acceleration causes the head suspension spring to vibrate, which in turn causes considerable fluctuations in head spacing, and increases the change of contact between the head and disk. Vibration of the suspension spring also prevents fast access and precise positioning.

This paper describes ways to improve flight stability and HDI reliability by using a negative-pressure head slider[1)] and a micro-slider. This paper also describes how improvements in head support enable faster access speeds.

2. Head sliders

The standard head slider for magnetic disk drives is currently 4 mm long with a double-rail and a flat tapered structure. The flight stability and durability of this type of positive-pressure head slider are balanced by regulating the loading force. To reduce spacing fluctuations, the loading force must be increased to stiffen the air film. To improve durability, however, the loading force must be decreased to prevent head and disk wear. Another disadvantage of the positive pressure head slider is the difficulty in improving the relationship between the flying height and disk velocity (e.g. takeoff characteristics and the ratio between the flying heights at the inner and outer tracks).

One way to overcome these problems is to use a negative-pressure head slider. This type of slide can stabilize flight, reduce wear at CSS, and optimize the flying ratio mentioned above. Another way is to miniaturize the slider, which would reduce wear at CSS and reduce impacts during head flight, and thereby improve HDI reliability. This approach would also reduce space fluctuations.

2.1 Negative pressure head slider

Figure 1 compares a negative-pressure head

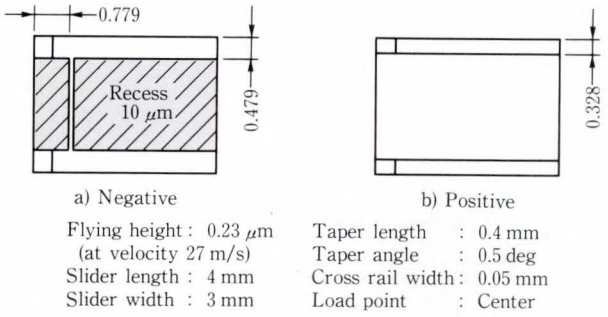


Fig. 1—Slider specifications.

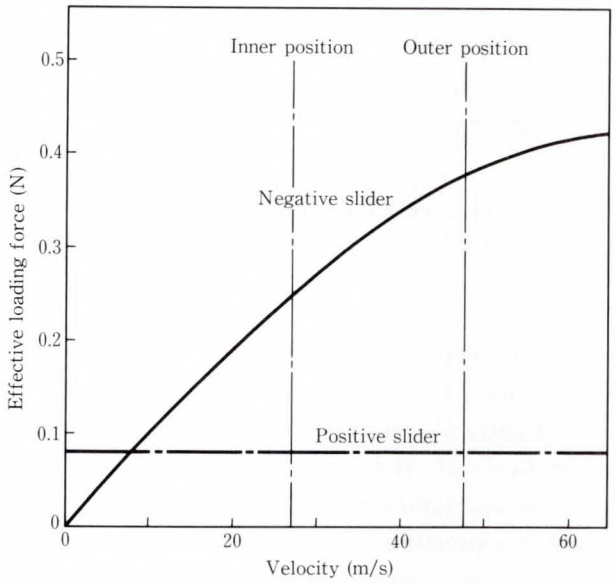


Fig. 2—Effective loading force dependence on velocity.

slider with a positive-pressure head slider. Both sliders are 4 mm long and are supported by a suspension spring. The loading force was set to 0 N for the negative-pressure head slider and to 0.08 N for the positive-pressure head slider.

The floating force of the positive-pressure head slider is caused by a wedge effect at the side rails. This force balances the loading force of the slider suspension spring. The suction of the negative-pressure head slider is caused by air expanding at the recess behing the cross rail between the side rails near the leading edge. (This suction also contributes to balance.)

The recess in the negative-pressure head slider is much farther away from the medium than the side rails; therefore, the suction generated at the recess is not greatly affected

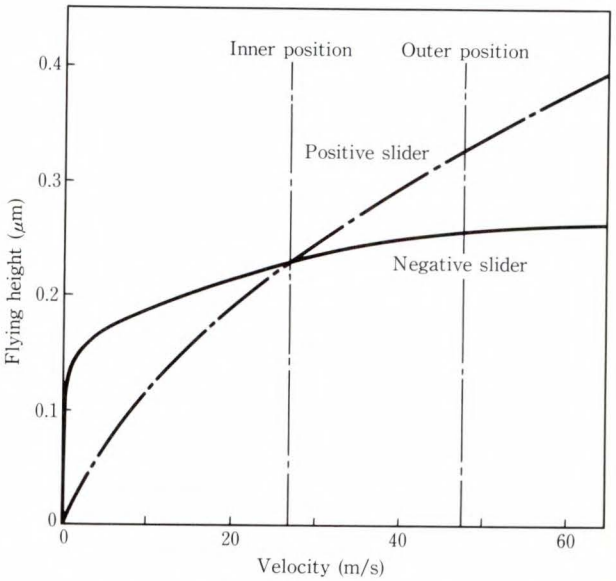


Fig. 3—Flying height dependence on velocity.

by spacing fluctuation. At a constant disk velocity, the suction acts as a loading force. The effective loading force is the sum of the loading force and the suction.

2.1.1 Reducing CSS wear

Figure 2 shows the relationship between the disk velocity and the effective loading force. This relationship was obtained using FEM analysis of a modified Reynolds equation. This equation takes into consideration the effect of the molecular mean free path of air[2]. The flight characteristics of a negative-pressure head slider is greatly affected by the depth of the recess[3]. In general, the smaller the recess, the lower the disk velocity at which the suction peaks. In this example, the recess depth is 10 μm and the suction gradually increases as the velocity increases from zero to the velocity at the outermost track.

This recess depth enables the slider to take off from the disk at a very low disk velocity (see Fig. 3), which reduces the duration of head/disk sliding at CSS.

During CSS, the impact between the head and disk was measured using an AE (acoustic emission) sensor attached to the head arm (see Fig. 4). The positive pressure head slider generated considerable impact at velocities below 7.5 m/s. The disk asperity was estimated to be 0.09 μm. For the negative-pressure head

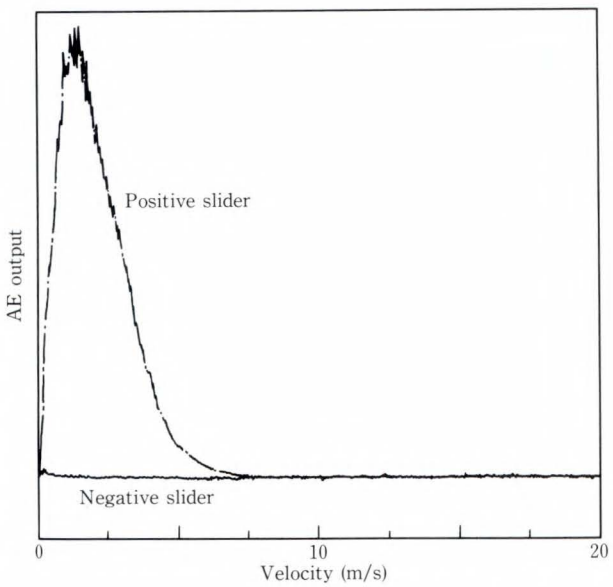


Fig. 4—AE output dependence on velocity.

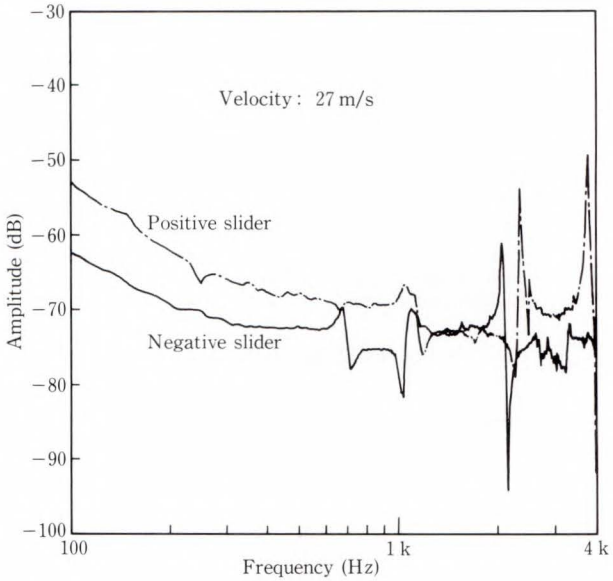


Fig. 5—Frequency response of spacing fluctuation to acceleration of head arm in access direction.

slider, the impact at CSS was so gentle as to be scarcely detectable. CSS durability is greatly improved by suppressing the effective loading force until the slider reaches or exceeds the disk asperity.

2.1.2 Improving the relationship between disk velocity and flying height

Figure 3 also shows that, compared to conventional sliders, the flying height of the negative-pressure head slider slightly increases as it moves from the inner to outer tracks. This occurs because the suction increases as the velocity increases, which counterbalances the increase in the floating force.

The negative-pressure head slider obviates the extra spacing in the outer tracks and so avoids reductions in recording density. This type of slider achieves an ideal relationship between the disk velocity and flying height.

2.1.3 Reducing spacing fluctuations

A slider achieves stable flight by increasing the effective loading force. A negative-pressure head slider can produce a large effective loading force at the rated disk velocities for which stable flight is required, while keeping it to a low level during head/disk sliding at CSS.

Over the rated velocity range from inner to outer tracks, the simulated effective loading force of the negative-pressure head slider was more than three times that of a positive-pressure head slider (see Fig. 2).

Figure 5 shows the measured transfer function of the spacing fluctuation due to access acceleration at the arm for positive- and negative-pressure head sliders. The head suspension springs of the two types have the same shape and only differ in the spring pressure. The spacing fluctuation was obtained by measuring the intensity of monochromatic light interfered between a transparent glass disk and the slider[4]. The gain was set to 0 dB at 1 μm/G. Even at the velocity of the inner most track, the spacing fluctuation due to access acceleration of the negative-pressure head slider could be reduced by about 10 dB below that of the positive-pressure head slider. The spacing fluctuation can be reduced even further on outer tracks where the effective loading force is higher.

2.2 Micro-slider

A small slider has a low weight and inertia; therefore, the loading force can be reduced while maintaining stable flight. Reducing the slider size improves flight stability and reduces head/disk deterioration during CSS and head flight.

2.2.1 Flight stability and CSS durability

If the disk velocity is low at a start or stop and the slider maintains contact with the disk, the slider and disk wear will depend on the loading force and normalized force imposed on the unit rail area.

If the slider size is reduced while keeping the same shape and the flight stability (spacing fluctuation ratio due to disk turbulence) constant, the loading force can be reduced in proportion to the cube of the slider size.

The normalized force decreases almost in proportion to the slider size (see Fig. 6). This indicates that by reducing the slider size, CSS durability can be improved. This also improves flight stability.

Figure 6 also shows that the normalized force increases as the flying height decreases. To enable a higher recording density, the slider will have to fly lower, so the slider size should be reduced to prevent the normalized force from increasing.

2.2.2 Reducing impact on flying slider

The air film must be kept stiff to reduce spacing fluctuation and prevent contact between the head and disk. However, making the air film too stiff will increase the force of impacts between the head and disk.

Collisions for each size of slider against a disk projection were simulated. Then, the resultant contact force, generated heat, and temperature rises were calculated. Figure 7 shows the model used for the simulation.

The stiffness of the air bearing supporting the slider was calculated using the perturbation method[5], and was treated as linear. Figure 8 shows the contact force between the projection and an $Al_2O_3 \cdot TiC$ ceramic slider having dimensions similar to the standard slider (4 x 3.2 x 0.85 mm), and having the same flight-stability response to disk turbulence. The horizontal axis indicates the distance between the projection and the slider's leading edge. Regardless of size, the slider first contacts the projection at about 2/3 the slider's length from the slider's leading edge. The 5.5 mm slider bounces on the projection three times. The 1 mm slider, however, is lighter and is easily pushed away by the projection. It then passes the projection in a short time relative to its size without rebound.

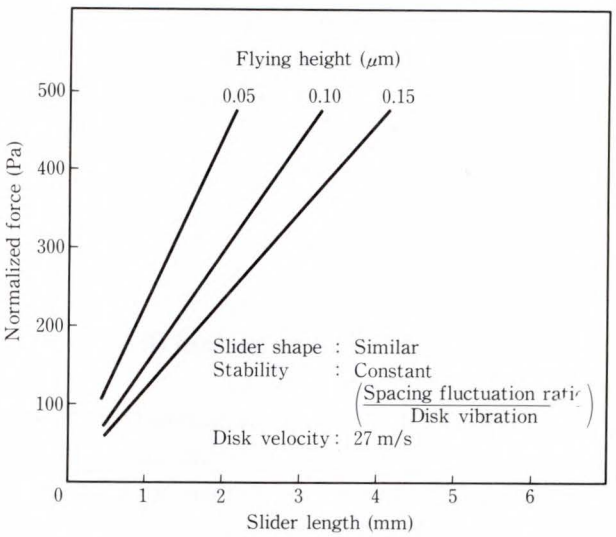
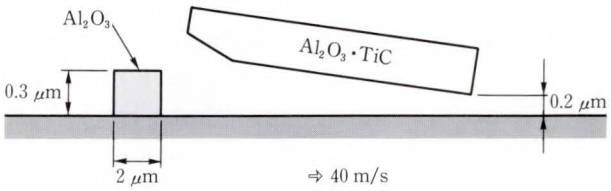


Fig. 6—Slider size and normalized force.



Projection
Height : 0.3 μm (flying height + 0.1 μm)
Diameter : 2 μm
Material : Al_2O_3
Penetration hardness : 11.4 GPa

Fig. 7—Collision calculation model.

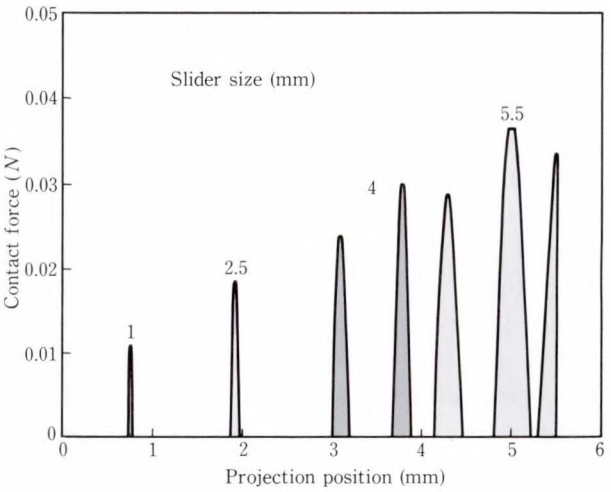


Fig. 8—Contact force.

By using a Blok Equation (1), the flash temperature at the instant of contact between the slider and projection can be estimated from the interface pressure[6].

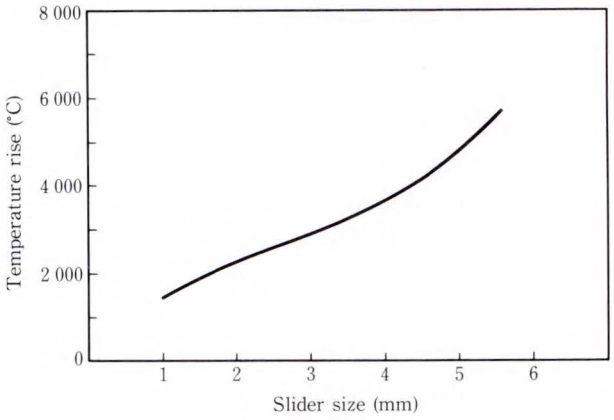
$$T = 2\mu PK\left(\frac{BV}{\pi k\rho C}\right)^{1/2}, \quad \ldots\ldots (1)$$

where, T : Flash temperature
μ : Coefficient of friction (0.3)
P : Interface pressure
K : Proportion of generated heat transferred to the moving body (–unity)
B : Length of projection contact area in the sliding direction (2×10^{-6} m)
V : Disk velocity (40 m/s)
k : Thermal conductivity (7.69 W/m·K at 500 °C)
ρ : Density (4 250 kg/m^3)
C : Specific heat (970 J/kg·K).

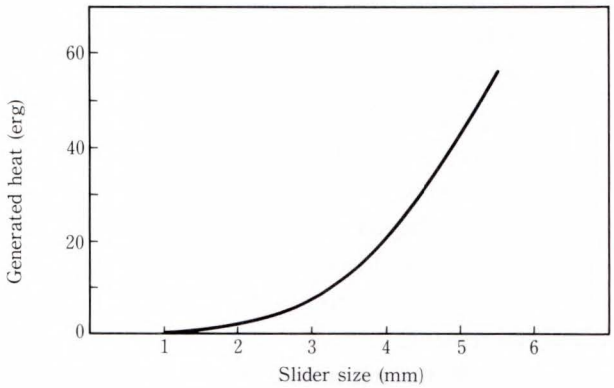
Figure 9a) shows the flash temperature at collision. The thermal conductivity of $Al_2O_3 \cdot TiC$ at 500 °C was used in the above calculation. As the temperature rises, thermal conductivity falls. So the real temperature rise is higher than in the calculation. The flash temperature at the contact point may well rise above 2 300 °C, which is the melting point for alumina. Such a high temperature may deteriorate the disk lubricant, allowing damage to the head and disk.

The amount of disk damage depends on the flash temperature and on the total heat generated by the collision. The frictional force is calculated by multiplying the contact force (see Fig. 8) by the coefficient of friction. The total heat generated by a collision is obtained by multiplying the shaped area in Fig. 8 by the coefficient of friction 0.3 in the case shown in Fig. 9 b) . Reducing the slider size dramatically reduces the amount of generated heat. In this example, the heat generated by a collision could be reduced by more than 90 percent by halving the size of the 4 mm standard slider; this is an extremely simple but effective way of extending the head/disk reliability.

Figure 10 shows the results of a durability



a) Temperature rise



b) Generated heat

Fig. 9—Temperature rise due to a collision.

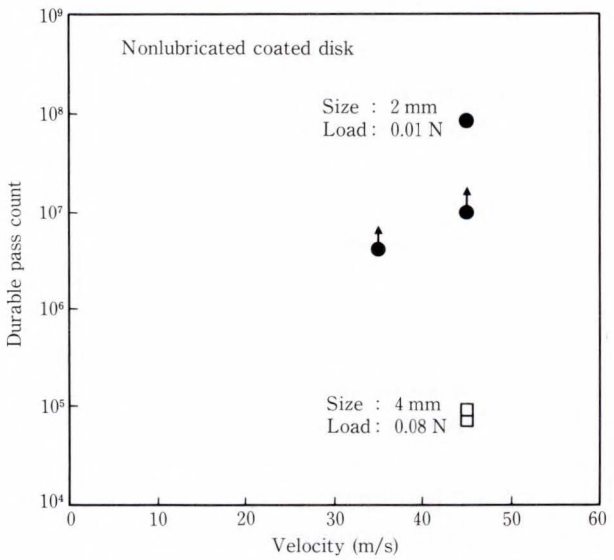


Fig. 10—Reverse-direction drag endurance.

evaluation of head/disk sliding. The number of passes was counted until the level of AE signals generated by sliding suddenly increased. The evaluation was performed for 2 mm and

4 mm heads. To speed up the evaluation, the slider was run in the reverse direction to maintain sliding, and a coated disk without lubricant was used. The arrows indicate that the test was suspended midway because of time restrictions. By reducing the slider size by 50 percent, the sliding durability of the head and disk was improved more than 1 000 times. This improvement occurred because the weight was reduced to 1/8, the loading force to 1/8, and the normalized force on the unit rail area to 1/2 their original values.

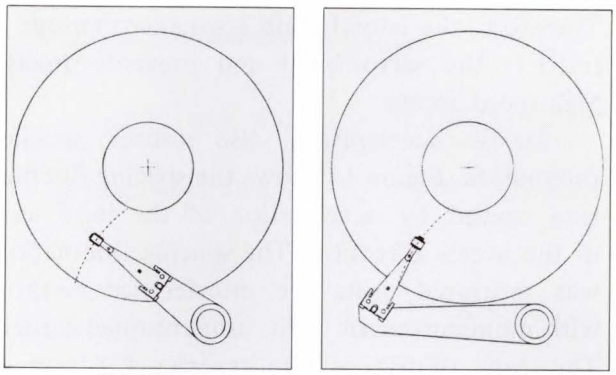
3. Head support system

For fast access, the head support system and suspension spring must be improved in addition to the slider flight characteristics. Figure 11 shows two head support systems. In the conventional system, the suspension spring is attached at an acute angle to the actuator arm. This makes the actuator arm unnecessarily long, and the resultant high inertia prevents fast access. In the recently developed inline support system, however, the suspension spring is inline with the actuator's rotation center. This simplifies the actuator structure and reduces the inertia of the actuator arm, thereby enabling fast driving of the head arm.

The suspension spring of the current inline head is the same as the one used in the conventional system except that it has been rotated 90 degrees to the slider (see Fig. 12). Thus, the access direction of the current inline system is perpendicular to the long axis of the suspension spring, a direction in which the spring is weak. Therefore, although the actuator of the current inline support system is suited to high-speed access, its associated suspension spring is not[7].

3.1 Dynamics of inline head suspension springs

Slider vibration in the access direction caused by access acceleration prevents fast access and degrades the head positioning precision. Figure 13 shows the frequency response of slider vibration in the access direction due to acceleration of the head arm in



a) Inline b) Conventional

Fig. 11—Head support systems.

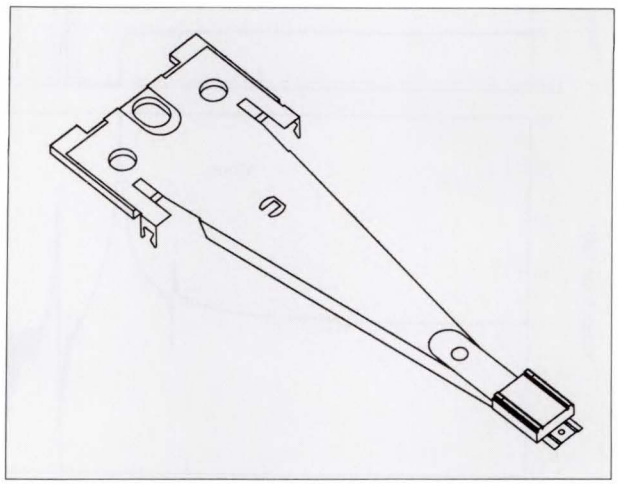


Fig. 12—Conventional inline head.

the access direction for the current inline head. An LDA (Laser Doppler Anemometer) was used to measure the slider vibration, and an accelerometer with a piezoelectric element was used to measure the arm acceleration.

Also, the vibration in the access and vertical directions at certain points on the head was measured using an LDA and an LDV (Laser Dopper Vibrometer) with vibration measurement directions offset by 90 degrees. The results of the mode analysis are shown in Fig. 14 (the mode numbers correspond to the ones shown in Fig. 13). Slider vibration in the access direction is caused by lateral main mode 3 and torsional modes 2 and 5 (which include lateral movement). The lateral main resonance (3.75 kHz) in mode 3 makes the slider vibrate more than 100 times as much as the arm. For stable control, the servo band is generally set to 1/10 or less of the resonance frequency.

Therefore, the lateral main resonance in mode 3 restricts the servo band and prevents precise high-speed access.

Access acceleration also causes spacing fluctuation. Figure 15 shows the spacing fluctuation caused by acceleration of the head arm in the access direction. The spacing fluctuation was measured using the interference method with monochromatic light, as mentioned earlier. The gain (0 dB) of the transfer function is the spacing fluctuation (1 μm) divided by the acceleration (1 G) of the head arm in the access direction. Figures 13 and 15 show that the slider vibration in the access direction and the spacing fluctuation are closely related. In modes 2, 3, and 5, the peak gain from the base line is within 3 dB. This correlation proves that the spacing fluctuation due to access is attributable to the factor shown in Fig. 16[8]. Spacing fluctuation at access acceleration

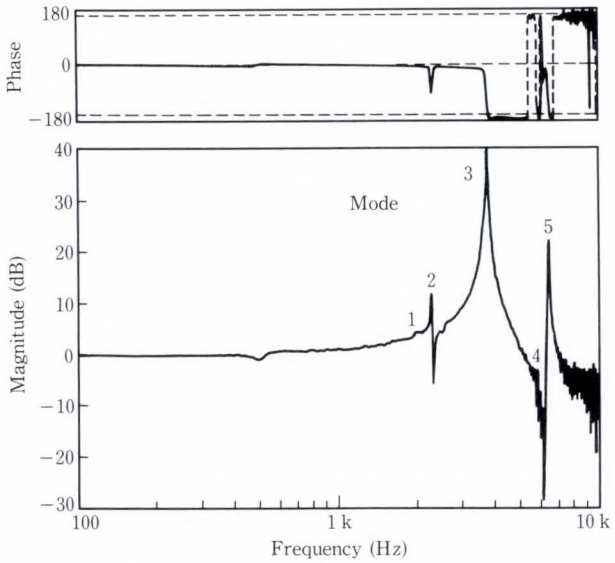


Fig. 13—Vibration of inline head slider due to head arm acceleration in access direction (lateral vibration: 0 dB = 1 G/G).

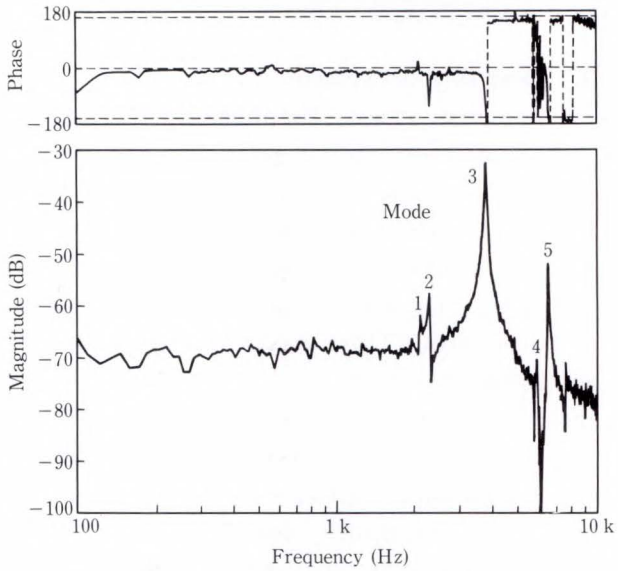
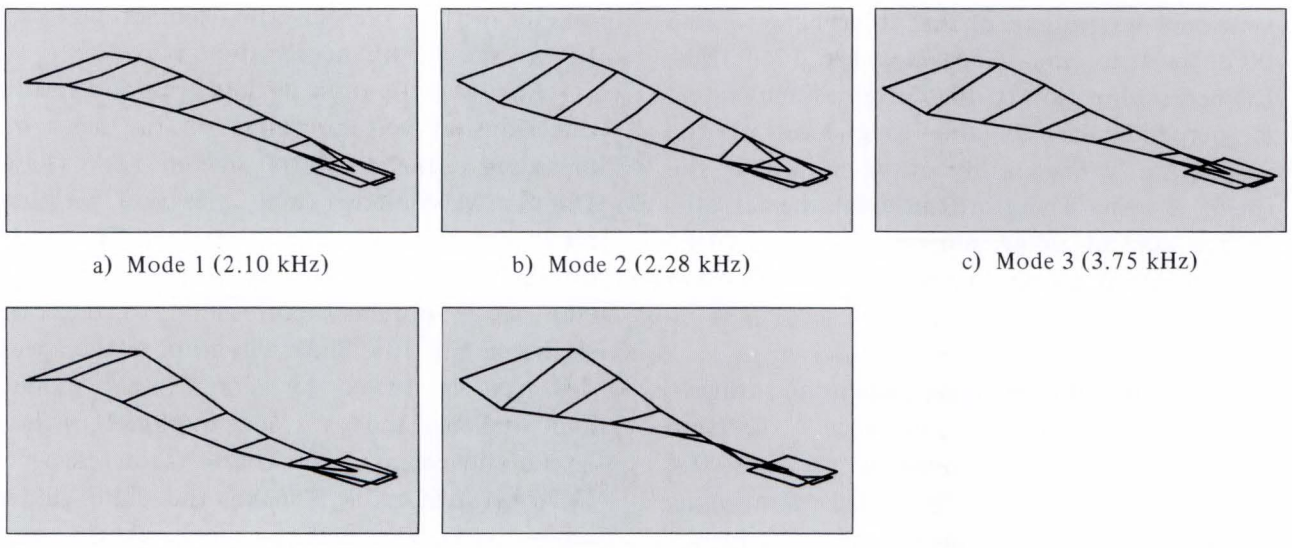


Fig. 15—Spacing fluctuation of inline head due to head arm acceleration $\{0\ \text{dB} = 1\ \mu\text{m/G} = 0.102\ \mu\text{m}\cdot(\text{m/s}^2)\}$.



a) Mode 1 (2.10 kHz)

b) Mode 2 (2.28 kHz)

c) Mode 3 (3.75 kHz)

d) Mode 4 (5.88 kHz)

e) Mode 5 (6.49 kHz)

Fig. 14—Vibration modes of inline head.

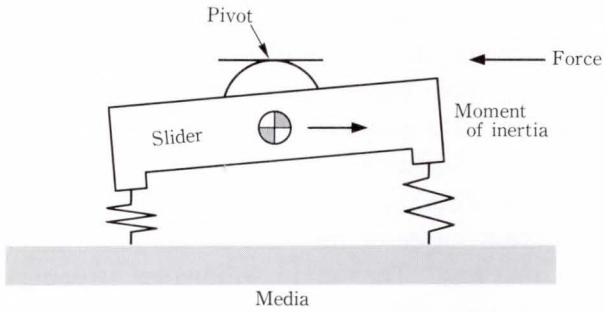


Fig. 16—Spacing fluctuation due to access[8].

is caused by slider rotation. This occurs because the slider's pivot point and its center of gravity are not aligned. Suppressing the lateral vibration of the suspension spring will reduce the spacing fluctuation. Reducing the offset between the slider's pivot point and its center of gravity also reduces the spacing fluctuation.

3.2 High-stiffness inline head suspension mechanism

A new inline head[9] that can be used for high-speed accessing has been developed. The new head features a suspension spring that is stiffened in the access direction to reduce vibrations caused by high-speed access. The new suspension spring, shown in Fig. 17, has a wider base. To maintain the equivalent translating mass of the suspension spring that is translated into the slider position, the suspension spring has a hollow center, and is made of thinner material. Use of the thinner material also helps prevent increases in the spring constant.

Figure 18 shows the relationship between the base width and resonance frequency of the suspension spring as obtained by FEM analysis.

By widening the base and making the hollow center, the resonance frequency in the lateral mode, which greatly affects fast tracking, can be raised without increasing the spring constant and equivalent translating mass, and without significantly decreasing the resonance frequency in the bending mode.

The high-stiffness inline head also reduces the rotational moment caused by access. The conventional suspension spring has its pivot

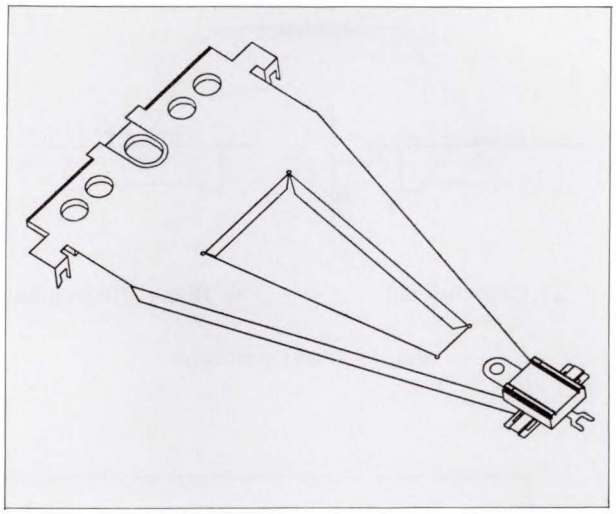


Fig. 17—High-stiffness inline head.

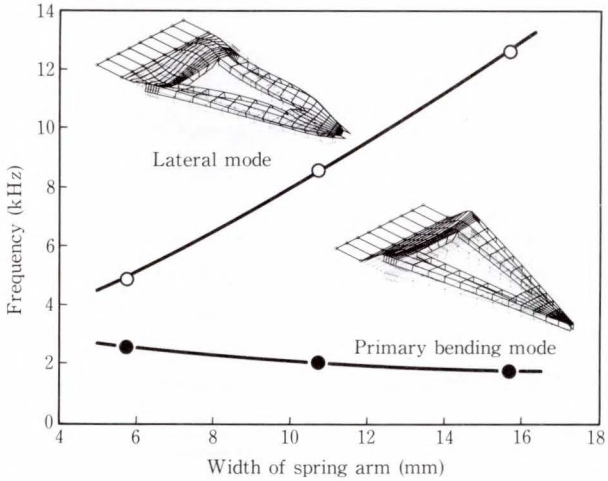
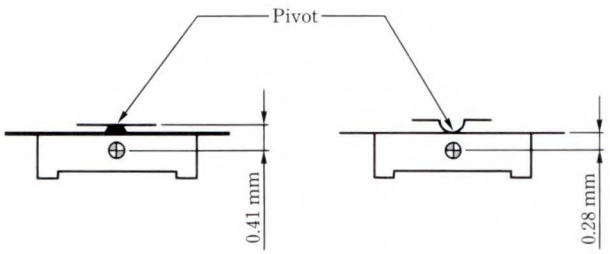


Fig. 18—Dependency of spring arm resonance frequency on base width.

projection on the gimbal side. Therefore, the access force is applied to a point removed from the slider back by the pivot projection height, resulting in a large rotational moment. The pivot projection of the high-stiffness inline head, however, is on the support spring side. This reduces the rotational moment and the resultant spacing fluctuation due to access. Figure 19 outlines the pivot structure of the high-stiffness inline head. Changing the pivot position in this way reduces the spacing fluctuation due to access by about 3.3 dB.

Figure 20 shows the spacing fluctuation caused by access acceleration on the current inline head and the high-stiffness inline head.



a) Conventional b) High-stiffness inline

Fig. 19—Pivot structures.

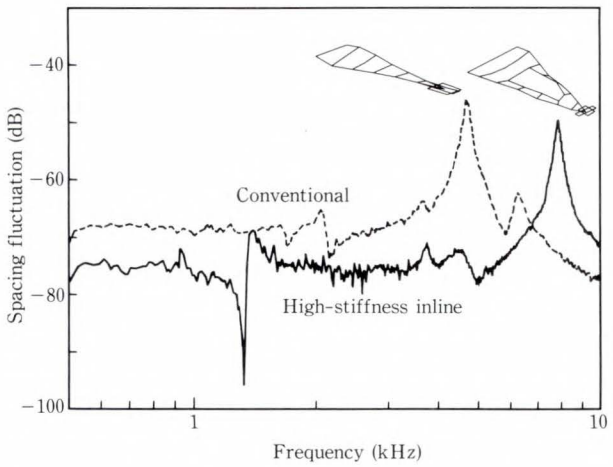
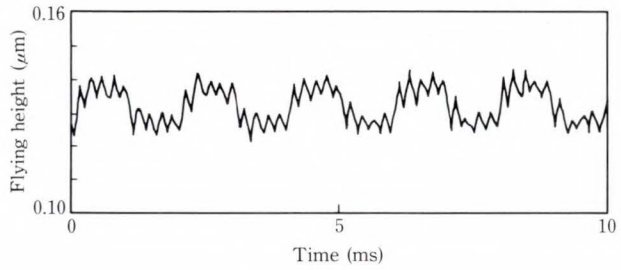
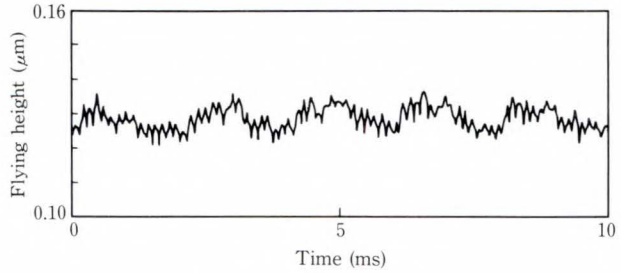


Fig. 20—Spacing fluctuation due to seek vibration.



a) Current inline



b) High-stiffness inline

Fig. 21—Spacing fluctuation due to square wave vibration.

The sliders used in the measurements were 80 percent the size of the 4 mm standard model. The high-stiffness inline head has a 14 mm wide support spring, a hollow center, and the modified pivot.

The diagrams in Fig. 20 show the measured vibration mode of the head at the peak spacing fluctuations. The main resonance frequency of the high-stiffness inline head was calculated using FEM analysis to be 11 kHz, but was 8 kHz in actual measurements. This represents a difference of about 30 percent. However, by comparing mode shapes, the vibration mode of the head was determined to be the lateral main resonance mode of the suspension spring.

The lateral main mode resonance frequency of the high-stiffness inline head is about twice that of the current inline head. This was achieved by widening the spring base and making the center hollow. This increase in resonance frequency enables a doubling of the servo band, and therefore enables faster access.

Below the main lateral resonance frequency, the spacing fluctuation decreased because the new pivot reduced the rotational moment due to slider access by 3.3 dB.

Figure 21 shows the spacing fluctuation when a 500 Hz square-wave acceleration in the access direction was applied to the inline heads. The spacing fluctuation of the high-stiffness in line head is about half that of the current inline head.

4. Conclusion

The recording density of magnetic disk drives has reached 100 Mbit/inch^2, and in the future will surpass that of current optical disks (400 Mbit/inch^2). To realize such a high recording density, the slider will have to fly at below 0.05 μm. Therefore, in addition to improvements in HDI reliability, a reduction in the slider's spacing fluctuation is also required. These objectives can be achieved by using a negative-pressure head slider and by reducing the slider size. For faster access, slider flight must be further stabilized and

the suspension spring must be strenthened against acceleration in the access direction.

A negative-pressure head slider was found to have the following advantages:

1) The relationship between the disk velocity and flying height can be adjusted as needed.
2) Impact between the head and disk at CSS can be completely eliminated by optimizal design.
3) A very stiff air bearing can be formed at the rated disk velocity to reduce the spacing fluctuation.

Despite these advantages, during flight, the conventional negative-pressure head slider risks a strong impact with dust (because its air film is very stiff), and can easily cause a head crash. These problems may prevent the negative-pressure head slider from being put to immediate practical use.

If the slider size is reduced, slider weight and inertia will decrease and the effective loading force required for stable flight will also be reduced. This will reduce the force of collisions between the slider and dust, thus solving the problem of head crash.

Current sliders are of the millimeter order. However, because the transducer mounted on the slider occupies less than one-thousandth of the slider's volume, there is considerable room for miniaturization.

To improve HDI reliability, a fine crown shape[10] that causes minimal stick-slip at CSS will be introduced. To realized a constant flying height over all tracks, other air bearing surface shapes which minimize spacing deviation due to the skew angle will be introduced. (Some examples of these shapes are the side-step[11] and the cross cut[12].)

Microsized multifunctional negative-pressure head sliders with complex air bearing shapes will be used in the future.

A high-stiffness inline head for fast access was developed by using FEM structural analysis and fine vibration measurement using an LDA and LDV. The new head has a widened suspension spring with a hollow center. It has been stiffened against acceleration in the access direction and has a new pivot structure. Because of these changes, the lateral resonance frequency has been doubled to 8 kHz and the spacing fluctuation due to access acceleration has been reduced to half that of the conventional inline head. This high-stiffness inline head makes it possible to take advantage of an inline support system suitable for a high-speed actuator drive.

References

1) White, J.W.: Flying Characteristics of the "Zero-Load" Slider Bearing. *Trans. ASME*, **105**, pp. 484-490 (1983).
2) Burgdorfer, A.: The Influence of the Molecular Mean Free Path on the Performance of Hydrodynamic Gas Lubricated Bearings. *Trans. ASME*, Ser., **D**, **81**, pp. 94-100 (1959); *ASME J. Basic Eng.*, pp. 94-100 (1959).
3) Yoneoka, S., Yamada, T., Aruga, K., Ooe, T., and Takahashi, M.: FAST TAKE-OFF NEGATIVE PRESSURE SLIDER. *IEEE Trans. Magnetics*, **MAG-23**, 5, pp. 3464-3466 (1987).
4) Yamada, T., Mizoshita, Y., and Aruga, K.: Spacing Fluctuation of Flying Head Sliders in Track Accessing (Measurement Using Optical Interferometer). *ASLE*, **SP-21**, pp. 82-84 (1986).
5) Ono, K.: Dynamic Characteristics of Air Lubricated Slider Bearing for Noncontact Magnetic Recording. *Trans. ASME J. Lubri. Technol.*, **97**, 2, pp. 250-260 (1975).
6) Ettles, C.M. McC.: Possible Flash Temperatures in Slider and Recording Disk Transient Contact. *ASLE, Trans.*, **29**, 3, pp. 321-328 (1986).
7) Yoneoka, S., Ooe, T., Aruga, K., Yamada, T., and Takahashi, M.: DYNAMICS OF INLINE FLYING-HEAD ASSEMBLIES. *IEEE Trans. Magnetics*, **25**, 5, pp. 3716-3718 (1989).
8) Aruga, K. et al.: Spacing Fluctuation of Flying Head Sliders in Track Accessing (Forced Vibration Analysis Using Finite Element Method). *ASLE*, **SP-21**, pp. 79-86 (1986).
9) Ohwe, T., Yoneoka, S., Aruga, K., Yamada, T., and Mizoshita, Y.: A DESIGN OF HIGH PERFORMANCE INLINE HEAD ASSEMBLY FOR HIGH SPEED ACCESS. *IEEE Trans. Magnetics*, **27**, 1990, (in print).
10) Lee H.J., Hempstead, R.D., and Weiss, J.: STUDY OF HEAD AND DISK INTERFACE IN CONTACT START STOP TEST. *IEEE Trans. Magnetics*, **25**, 5, pp. 3722-3724 (1989).
11) White, J.W.: Magnetic Head Air Bearing Slider

Assemblies Utilizing Transverse Pressurization Contours. U.S. Patent 4 673 996, 1985-04-29 Filed, and 1987-06-16 Date of Patent.

12) Daito, H.: Magnetic Head Slider. Utility Model in Japan, 1985-186 562, Released on 1985-12-10.

Seizi Yoneoka
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Electrical Eng.
Yokohama National University 1974
Master of Electrical Eng.
Yokohama National University 1976
Specializing in Magnetic Recording

Takeshi Ohwe
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Mechanical Eng.
Niigata University 1981
Specializing in Magnetic Recording

Yoshifumi Mizoshita
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Precision Eng.
Hiroshima University 1972
Master of Precision Eng.
Hiroshima University 1974
Specializing in Mechanical and Control Engineering

UDC 621.318:681.327.634

Head-Disk Interface

• Takayuki Yamamoto • Minoru Takahashi • Masayoshi Shinohara

(Manuscript received June 4, 1990)

As the flying height of magnetic heads decreases, the possibility of head-disk contact increases. New acoustic emission (AE) sensors for detecting head-disk contact have been developed. These devices are smaller and more sensitive than conventional ones. Using these new sensors, a method of evaluating the head-disk interface has been developed, and the influence of organic gas and dust on the head-disk interface was studied. It was found that component outgassing markedly increases the number of head-disk contacts, and can cause a failure in the head-disk interface. Wear on the slider and carbon overcoat on the disk was also studied, and the chemical change of the slider surface due to frictional heat was observed.

1. Introduction

In magnetic disk drives, the spacing between the head and disk is reduced as much as possible to increase the recording density. As the spacing decreases, contact between the head and disk becomes more likely. Therefore, the head-disk interface (HDI) is being studied to prevent head-disk interface failure.

In HDI technology, it is important to know how the lubricant layer will rupture, how the protective layer will wear, and when head crashes occur. HDI technology covers the lubrication and protective layer on the magnetic disk, dust or contamination in the disk enclosure, and flying stability of the magnetic head.

This paper describes a method of detecting head-disk contact, tools for testing flying stability of the magnetic head in various atmospheres, and sliding tests made in a rarefied atmosphere.

This paper then presents the results of HDI evaluation using these methods.

These results include the influence of gas and dust on the flying characteristics of the magnetic head. Special attention is given to how contamination of the head and disk surfaces affect head-disk contacts. The cleaning effect of the slider surface during head-disk contacts occurring during seeks is decribed, and the results of sliding tests under various environments are presented.

Lastly, this paper describes the chemical change of the head slider surface caused by friction and wear, and how head sliding affects the roughness of the disk surface.

2. Evaluation tools for HDI

The lower the flying height of the magnetic head, the more important flying stability evaluation becomes for the achievement of stable flying characteristics.

This section describes a new head-touch sensor developed to detect head-disk contacts. This section also describes a tester developed to evaluate the effect of environmental factors on the flying characteristic of the magnetic head.

2.1 Head touch sensor

In a magnetic disk system, the magnetic head flies at a submicron spacing from the disk surface on a high-velocity air flow created by the rotating disk. If the flying height is decreased without compensatory measures, the probability of contact increases, the head and disk surfaces become contaminated, and intermittent contact gives way to continuous contact or head crash. A head crash seriously damages the disk surface,

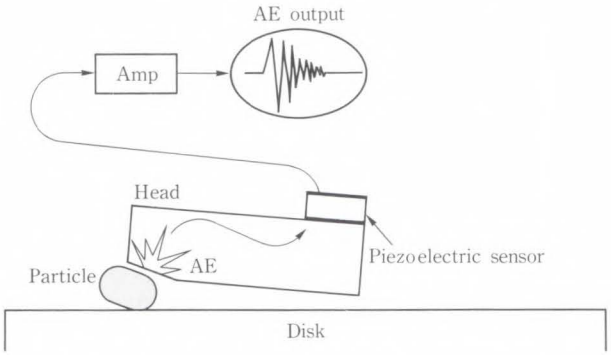
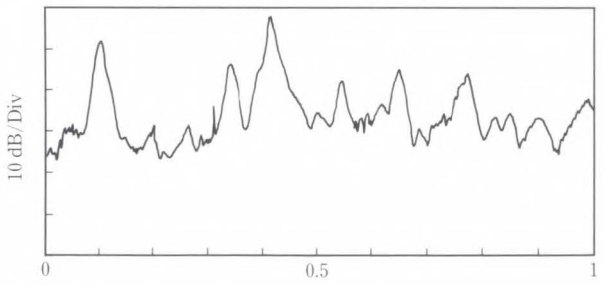
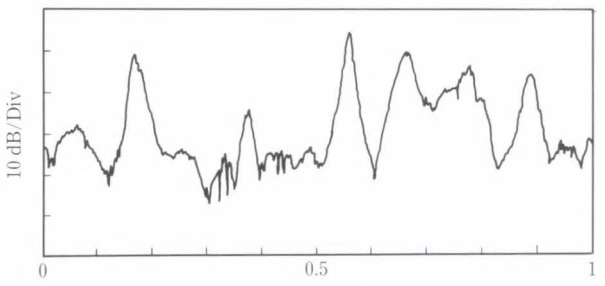


Fig. 1—Head/disk contact detection system.



a) Ferrite slider



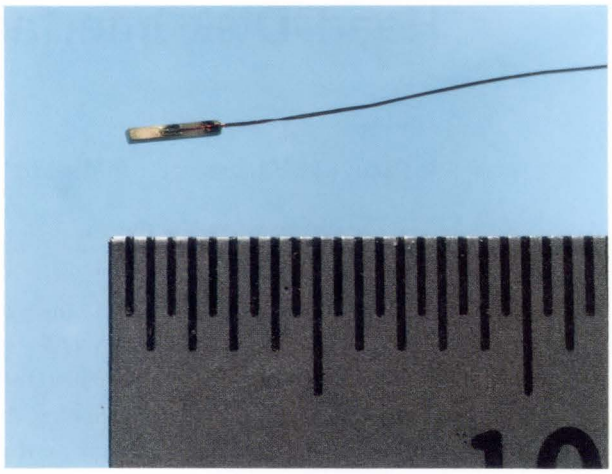
b) Alumina slider

Fig. 2—Spectrum when passing a projection.

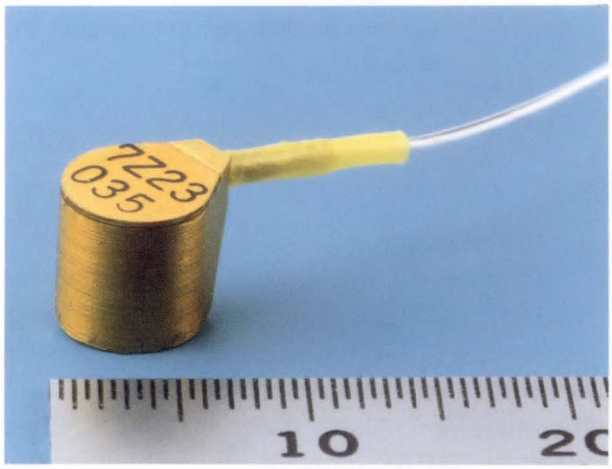
and data may be lost or some other serious error may occur. Therefore, head-disk contact, (head touch) must be detected with high sensitivity in the early stage of contact.

Commercially available acoustic emission (AE) sensors have been used for head touch detection. However, these sensors are not sensitive enough to detect a head touch. Therefore, new AE sensors and a new amplifier were designed. The new sensors have higher sensitivity and a lower price than conventional sensors.

Figure 1 shows the head touch detection system. When the head collides with flying dust



a) Piezoelectric sensor



b) New type AE sensor

Fig. 3—Head touch sensors.

or a disk projection, an acoustic emission is emitted. To detect this acoustic emission, the AE sensor[1] is mounted directly on the slider. The spectral pattern of the acoustic emission depends on the material and shape of the slider. Figure 2 compares the acoustic spectrums of two sliders having the same shape but made of different materials. The peak frequencies of the spectrum depend on the elasticity on specific gravity of the material. An AE sensor mounted on the slider has high sensitivity because the sensor and signal source are close together. However, because it is difficult to shield the sensor from electromagnetic fields, external electrical noise can easily cause interference. To solve this problem and to increase sensitivity, differential amplification was used and a specific peak frequency was selected.

Figure 3a) shows the new AE sensor. This head-mounted high-sensitivity AE sensor, called a piezoelectric sensor, is thinner than the lead in a mechanical pencil and does not significantly increase the head weight.

Figure 3b) shows another type of high-sensitivity AE sensor mounted on an arm. This sensor is housed in a case and is easily handled. This type of sensor has a disadvantage in that the signal quality is degraded by attenuation and mechanical vibration because the AE sensor is mounted on the arm and is distant from the signal source. This problem was solved by matching the mechanical resonance frequency determined by the slider material and shape with that determined by the size of the AE sensor. In addition, it was possible to detect head touches of all the heads in the disk enclosure with only one AE sensor by analyzing the propagation of acoustic emission in the disk enclosure.

Figure 4 compares the output from various sensors when the disk rotation velocity is gradually reduced. (The same heads and disks are used.) The output of the new AE sensor rapidly increases as the speed of disk rotation is reduced. This indicates that the new sensor has a high sensitivity to head touches.

The two types of high-sensitivity AE sensors are used for various purposes. For example, they are used to manage surface roughness and contamination of disks in the manufacturing process, and to monitor head touches when the disk drive is in operation. This improves the reliability of the evaluation of head flying and head touches.

2.2 HDI testing in various environments

The spacing between the magnetic head and disk is reduced to increase recording density. Flying reliability is greatly affected, however, by environmental factors such as gas, temperature, and humidity. The conventional method of environmental testing is to check the failure rate of many magnetic disk drives running in a thermostatic cabinet. However, no significant results can be obtained even over an extended period of testing. Therefore, an HDI tester which can control the ambient gas was developed to monitor contacts and the disk surface. Figure 5 shows the configuration of the HDI tester. A spindle is installed in an enclosure in which clean air is circulated at a controlled temperature and humidity. During the test, the AE sensor mounted on the head detects head-disk contacts.

The head and disk surface are monitored by a CCD camera attached to a microscope. The AE, camera image, temperature, and humidity signals are continuously recorded on a video

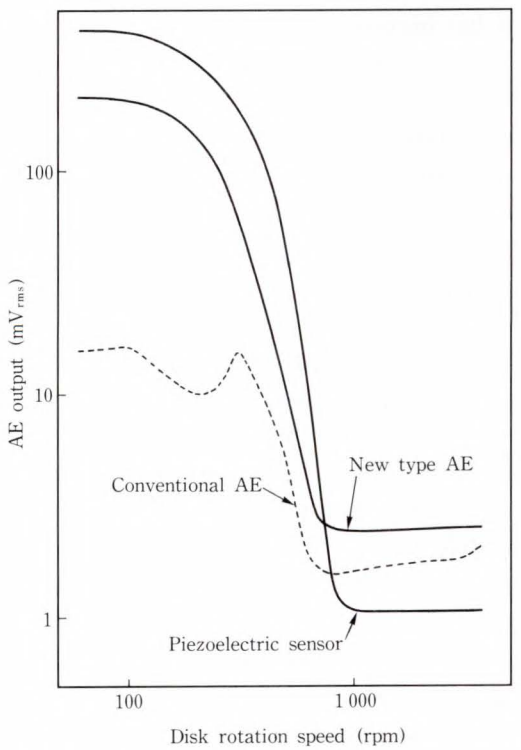


Fig. 4—Comparison of sensitivity of AE sensors.

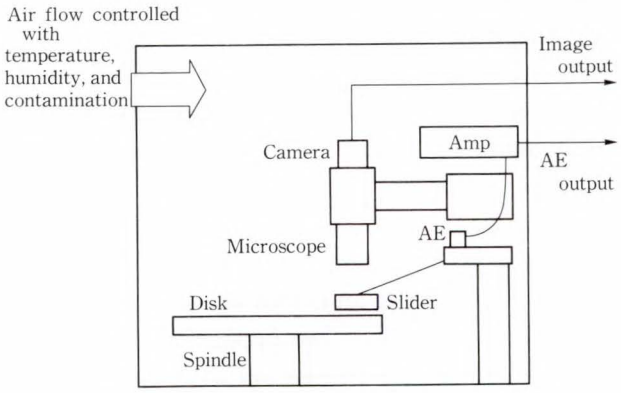


Fig. 5—Configuration of HDI evaluator.

tape recorder so that necessary data can be reproduced at any time.

This HDI tester enables a series of head crash processes to be observed. The ambient gas, temperature, humidity, disk, and head can be changed to investigate how quickly head-disk contacts increase, damage to the disk surface, and what happens if contact decreases gradually.

The HDI tester is used to check the relationship between the head-disk spacing and the environmental factors, and to analyze the head crash mechanism.

2.3 Sliding tests in rarefied air

A magnetic head contacts the disk when the disk is stopped, or is being started or stopped. Therefore, the wear characteristic of the disk have to be evaluated. The wear characteristics can be evaluated by conducting contact start and stop (CSS) and sliding tests.

In the CSS test, the disk is accelerated to full speed from rest, continued at full speed for a short time, then decelerated to a stop. This cycle causes repeated intermittent sliding contact during acceleration and deceleration. This test can be conducted using an actual head, disk, and magnetic disk drive but it takes a long time.

The high-velocity sliding test achieves high-velocity sliding by narrowing the slider rail width of the magnetic head or by reversing the rotational direction to reduce flotation. This test can be executed in a short time but it is hard to correlate the test results with normal operation because a special head is used for the test.

The sliding test in rarefied air was developed to solve these problems[2]. The magnetic head flies above the disk due to the pressure generated by the flow caused by the rotating disk. High-velocity sliding can be achieved by reducing this pressure. The object of the sliding test in rarefied air is to reduce the ambient pressure and achieve high-velocity sliding with almost no flight pressure on the slider surface. Using the tester, a stable contact run can be attained at high velocity using an actual head.

Figure 6 outlines the pressure-reducing slide tester. The disk is fixed on a spindle and the head is fixed on an X-Z stage. A rotary pump reduces the internal pressure of the chamber to 1 Pa or less. The disk rotates at 3 600 rpm or more, which causes high-velocity sliding. The head-disk sliding is sensed by the new AE sensor. The disk surface can be monitored on a real time basis, using a microscope equipped with a CCD camera. This tester also has a gas inlet to enable running tests to be conducted in various gas environments.

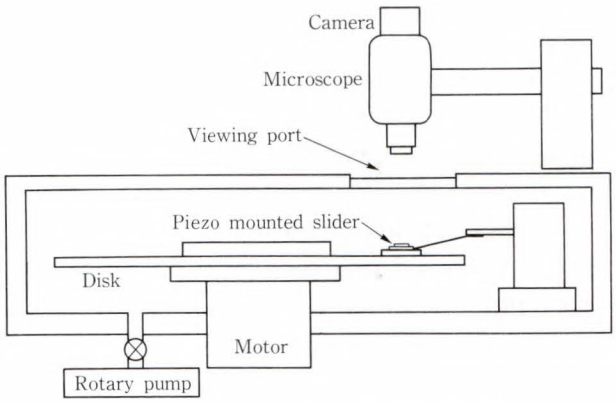


Fig. 6—Pressure-reducing slide tester.

3. Environmental factors and flying reliability of magnetic heads

3.1 Influence of temperature, humidity, and other factors

Disk enclosures inevitably contain materials that generate small particles or gas. In particular, organic materials, such as those found in gaskets, grease, and adhesives, release organic vapors throughout the disk enclosure. Plasticizers and organic siloxane from silicone gaskets are especially likely to affect the head-disk interface because they are released in relatively large amounts and spread very quickly[3)-5)]. Gasket materials that produce little or no outgassing are available but they are too expensive for practical use.

Therefore, thermally treated silicone gaskets are widely used. This section presents the results of studies on the influence of silicone gasket outgassing on the head-disk interface.

First, the head was positioned at the inner portion of the disk and the disk rotated. Head-seek was repeated for 8 h, followed by non-seek flying at an inner portion for 16 h. These two

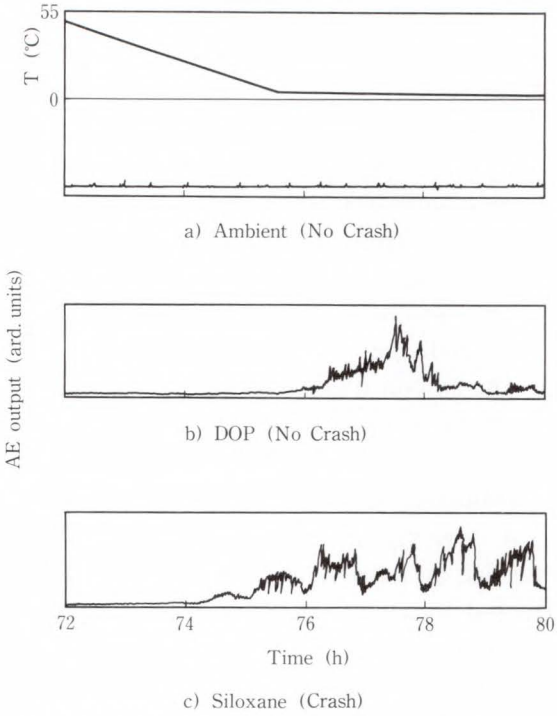


Fig. 7—Effects of temperature on head touch.

tests were equivalent to the head movement of one day's use. The temperature in the test chamber was maintained at 55 °C for the first three days, and at 5 °C for the next three days. The relative humidity in the test chamber was less than ten percent.

In this study, a particulate disk with perfluoropolyether lubricant and Al_2O_3·TiC head were used. Dioctylphthalate (DOP) and octamethylcyclo tetrasiloxane (siloxane) were used to simulate the plasticizer and organic siloxane that outgas from the silicone gasket.

In the tests made in ambient atmosphere, no head-disk contacts occurred over the range of 55 °C to 5 °C, {see Fig. 7a)}. When DOP was introduced in the test chamber, no head-disk contacts occurred at 55 °C, but AE output increased as the temperature was reduced {see Fig. 7b)}. Head-disk contacts started to occur at about 12 °C. No head-disk contacts were observed during non-seek flying, and AE output decreased to zero on the last day. After the test, the inner position of the disk surface was slightly abraded, but no head crash had occurred.

When siloxane was introduced in the test chamber no head-disk contacts occurred at

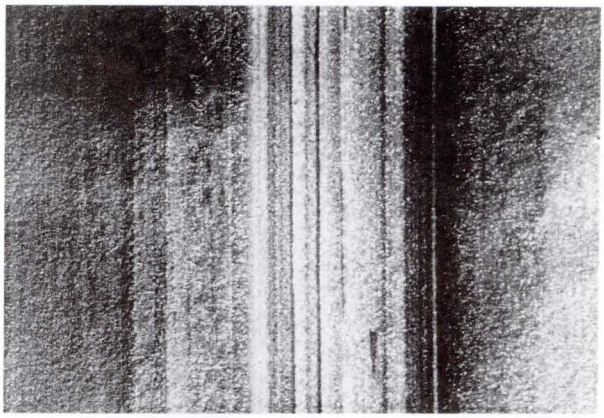


Fig. 8—Head-crashed disk surface.

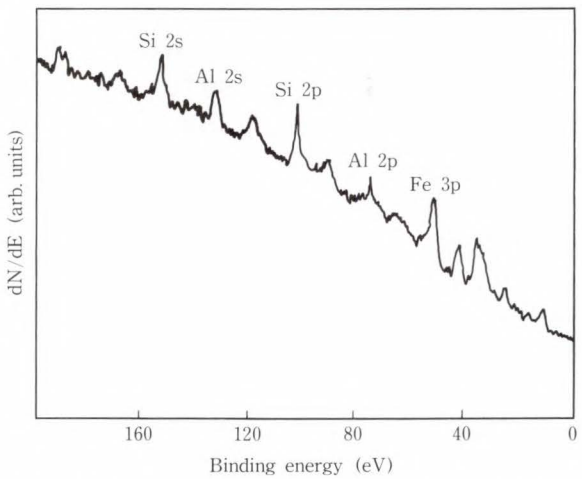


Fig. 9—XPS spectrum of head-crashed disk surface.

55 °C {see Fig. 7c)}. Head-disk contacts occurred at about 30 °C and were observed during non-seek flying. Repeating the head-seek operation increased the frequency of head-disk contacts. Eventually, the disk surface became damaged and a head crash was observed. Figure 8 shows a SEM image of the scratch on the disk caused by the head crash. The disk surface after the tests was analyzed using X-ray photoelectron spectroscopy (XPS). A typical XPS spectrum is shown in Fig. 9. The only place where elemental silicon was detected was at the mark left by the head crash. The binding energy of Si-2p is 102.5 eV, that of Si is 99.4 eV, and that of oxidized silicon is 102.9 eV[6]. This suggests that the siloxane on the disk was oxidized by the heat generated by head touch.

The head touch generation mechanism will

now be discussed using the results of the above tests, and a model of the mechanism will be described. The number of head-disk contacts increased when the temperature was lowered from 55 °C to 5 °C. This occurred because the organic materials evaporated at the high temperatures and then condensed on the disk at the low temperatures to contaminate the disk and slider surfaces and thereby cause head touch. When the temperature was lowered from 22 °C to 5 °C, there was no increase of head touch even in a siloxane atmosphere. This fact proves the aforementioned mechanism, that is, the evaporation of contaminant needs to accelerate the head touch.

Head-disk contacts occurred at 12 °C in a DOP atmosphere, and at 30 °C in a siloxane atmosphere. Liquid with a low surface tension spreads easily over a solid surface. Siloxane, with a surface tension of 0.02 N/m, spreads more easily over a disk than DOP, which has a surface tension of 0.036 N/m. At low temperatures, siloxane is a thin (solid) film on the disk surface, while DOP is in the form of droplets.

In a DOP atmosphere, the disk and slider surfaces were contaminated at low temperatures. This contamination induced head-disk contact. During a head seek, however, this tendency decreased because the droplets of DOP on the disk surface were removed by the head, and significant liquid contaminants were not deposited. This is supported by the fact that slider surface contaminants in a DOP atmosphere are liquid-like. In this experiment, a quartz slider was used to observe the rail surfaces.

In a siloxane atmosphere, head-disk contacts increased when the head seek operation was repeated, and occurred even during non-seek flying. This is because of solid contaminants on the disk and head surfaces are not easy to remove. The localized flash temperature caused by a head-disk contact probably exceeds has been reported to exceed 600 °C[7)-9)]. Preliminary experiments suggested that polysiloxane changes to silicon oxide at 350 °C. These silicon oxide particles were hard enough to abrade the disk surface.

The head crash mechanism may be summarized as follows:

1) Organic gas is vaporized at high temperatures.
2) Vaporized gas condenses on the disk surface at low temperatures and contaminates the slider and disk surfaces.
3) Contaminants on the disk and slider cause head-disk contact which generates heat. This heat products silicon oxide particles from the siloxane.
4) The silicon oxide particles increase the probability of head-crash.

3.2 Influence of dust and contaminants[10)]

A low flying height makes the head-disk interface an extremely critical area that is easily affected by small particles in the air. However, quantitative data on the effect of varying types and amounts of dust are not currently available.

To investigate the effect of particles on the head-disk interface, a contamination system which disperses controlled concentrations of micro-sized glass heads onto a disk was developed. This system was then used to evaluate the head flying characteristics and durability of the contaminated disks. Finally, the effect of dust removal on disk durability was evaluated in a high-velocity sliding test.

3.2.1 Contamination system

The contamination system disperses a controlled amount of particles into the air flowing over the disk.

This system has the following features:

1) The particle concentration in the chamber is measured using a particle counter. The mixing parameters are regulated to control the particle concentration.
2) The air flow is regulated using a flow meter.
3) The mixer mixes and disperses the particles in the air.
4) Particles are trapped by an air filter in the exhaust.

Particle concentration was varied from 10^4 to 10^7 at cleanliness class over several hours using glass beads of an average diameter of 1.6 μm.

SEM observation after contamination indicated that if a disk is stored in air contaminated

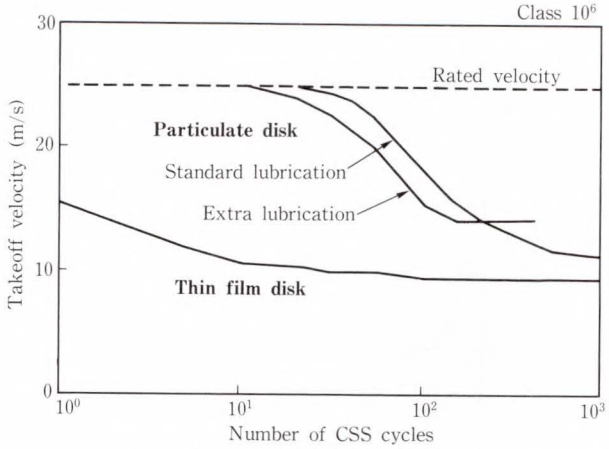


Fig. 10—CSS count versus takeoff velocity.

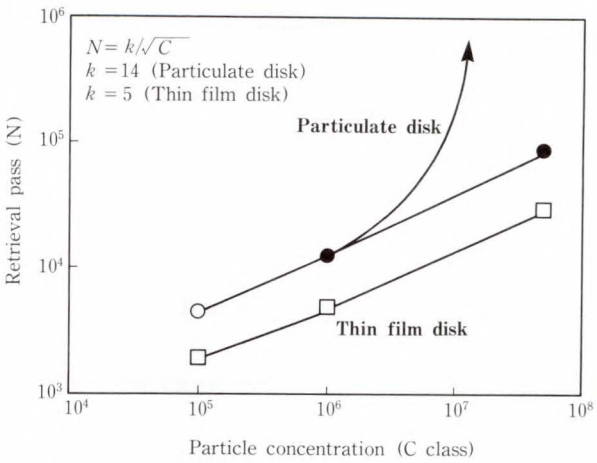


Fig. 11—Dust contamination concentration versus retieval passes.

with dust, the dust sticks evenly to the disk surface. The dust concentration on the disk increases in proportion to the storage time.

3.2.2 Flying characteristics at contact start and stop

In general, a magnetic head can not easily fly above a dust-contaminated disk. The output of the AE sensor was measured to evaluate the head flying characteristics from start to stop of disk rotation. AE output rapidly increases soon after the disk starts to rotate, and then decreases with increasing rotational velocity. AE finally reaches a constant value that is nearly equal to the inherent noise of the sensor system. This level is reached when the head achieves stable flying. The velocity at this point is called the take-off velocity.

Figure 10 shows the relationship between the take-off velocity and CSS count in a CSS test on contaminated disks. At the beginning of a test using a thin film disk, the head took off at 16 m/s. However, when CSS was repeated, the head took off more easily. Eventually, the take-off velocity became the same as for an uncontaminated disk.

In the same test using a particulate disk, the head did not take off at the rated velocity of 25 m/s, but took off gradually as the number of CSS cycles increased. These results indicate that dust can be removed from a thin film disk more easily than from a particulate disk.

3.2.3 High-velocity sliding test

To check the influence of dust, a high-velocity sliding test with AE monitoring was conducted using a head with narrow slider rails on a contaminated disk. When the sliding test was conducted on a contaminated disk, AE output first increased and then decreased as the number of slides was increased. Finally, the AE level became the same as that of an uncontaminated disk. Thus, the influence of dust may have been completely eliminated at this level. The retrieval pass is defined by the number of slides.

Figure 11 shows the relationship between retrieval passes and the particle concentration to which the disk was exposed. The arrow indicates that head-crash occurred on the least lubricated disk before AE output recovered.

The relationship between retrieval passes and particle concentration becomes linear regardless of the disk type; this can be represented by the following equation:

$$N = k / \sqrt{C},$$

where

N: Number of retrieval passes

C: Particle concentration

k: Coefficient of resistance to particle removal.

To determine how difficult it is to remove dust from a disk surface, k in the above equation was calculated. From this calculation, it was found that dust can be removed from a thin film disk about three times more easily than from a particulate disk.

The above test results show the influence of particle contamination, and provide valuable information for questions regarding the manufacture of disk drives, for example, which combinations of head and disk have the highest tolerance to dust?, and what are the optimum control methods?

4. Wear of slider materials and protective layer on disk

4.1 Chemical changes in the slider surface due to sliding[11),12)]

The sliders of flying heads are mostly made of MnZn ferrite for monolithic heads, and $Al_2O_3 \cdot TiC$ ceramic for thin-film heads. After a sliding test, the slider surface is not only mechanically damaged but also chemically damaged. However, very few detailed studies on friction and wear of such materials have been conducted. A surface analytical technique was used to study the structural and compositional change in the slide surface.

4.1.1 MnZn ferrite

Polycrystalline MnZn ferrite with three different averabe grain sizes (10, 30, and 60 μm) were processed into 7 mm diameter pins with semi-spherical ends. Sliding tests using iron disks in normal atmosphere at a load of 0.23 N and a sliding velocity of 2 m/s were conducted. It was found that wear resistance increased with increasing grain size. Using an SEM, it was also found that the grain boundaries had worn preferentially. This indicates that grain boundary destruction is the main factor in wear of the MnZn ferrite. Since a small grain has a large grain boundary, it may wear faster.

Figure 12 shows the results of analyzing the pin surface before and after a sliding test using Auger electron spectroscopy (AES). The pin surface definitely lost a slight amount of Zn during the sliding test but the Mn concentration did not change. Then, the Mn and Zn concentration in the wear debris was analyzed using an X-ray microanalyzer (XMA). The Zn concentration was somewhat higher in the wear debris than in the surface layer. This is because Zn was lost preferentially. These results suggest that the changes in surface

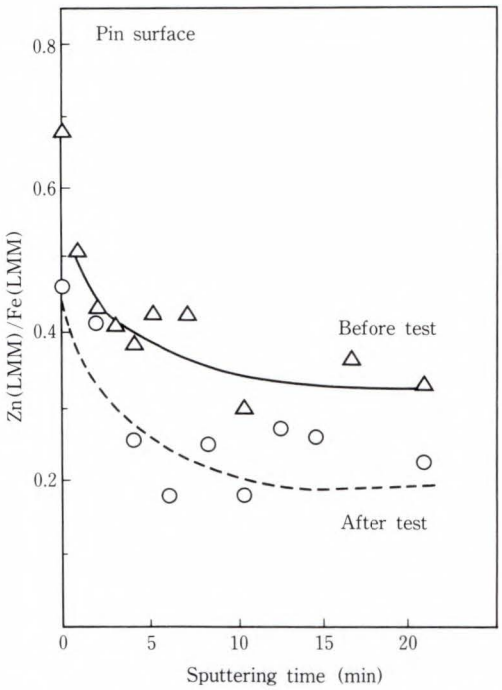


Fig. 12—Auger intensity ratio of Zn to Fe obtained from a worn MnZn ferrite pin.

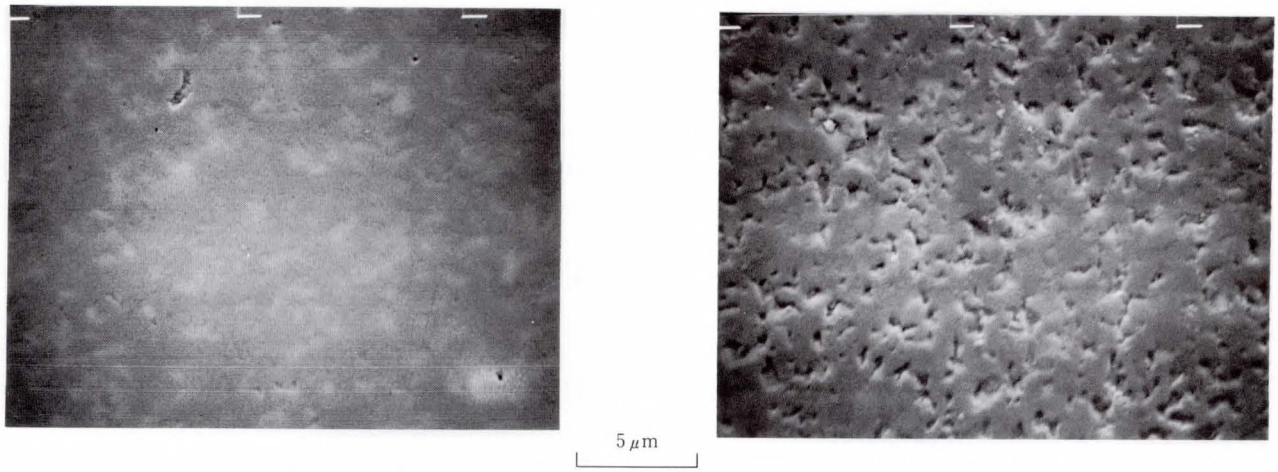
composition took place before the formation of the wear debris. Selective removal of Zn rich portions occurred due to the wear.

4.1.2 $Al_2O_3 \cdot TiC$

Titanium carbide in $Al_2O_3 \cdot TiC$ is known to be preferentially worn by sliding. The mechanism of preferential wear was investigated by analyzing the slider surface after the sliding test. $Al_2O_3 \cdot TiC$ sliders and carbon overcoated thin-film disks were mounted in the aforementioned pressure-reducting slide tester. The load was 0.15 N and the sliding speed was 12 m/s. The tests were conducted in air and argon atmospheres.

After the sliding test in air, the TiC portions of the $Al_2O_3 \cdot TiC$ were partially discolored into interference colors of green and blue. The longer the sliding distance, the larger the discolored area. The slider did not discolor in an argon atmosphere.

Figure 13 shows SEM photographs of the $Al_2O_3 \cdot TiC$ slider surface before and after the sliding test. The slider surface was highly polished before sliding. Surface analysis by XMA showed that the bright portions were TiC grains,



a) Before sliding test

b) After sliding test

Fig. 13—Al_2O_3·TiC slider surface before and after sliding test.

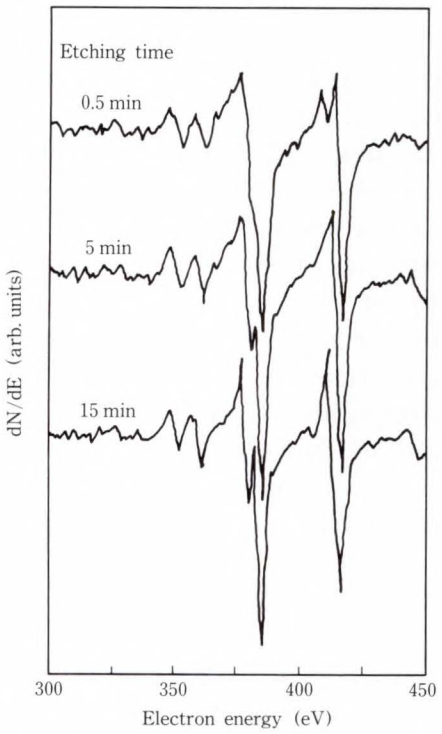


Fig. 14—Changes in Ti Auger profile with depth.

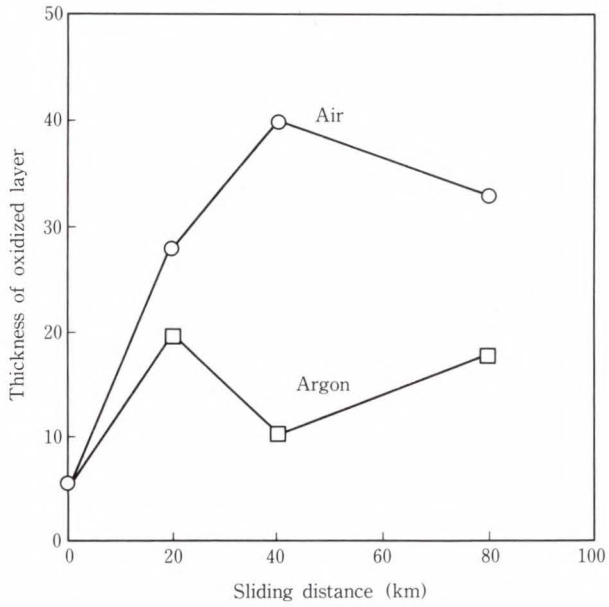


Fig. 15—Thickness of oxidized layer.

and during sliding test, these TiC grains were selectively lost. TiC is harder than Al_2O_3 at room temperatures but is more brittle at high temperatures. Frictional heat at the slider surface may have made the TiC portions brittle and caused them to wear. Preferential wear of TiC also occurred in a friction test in an argon atmosphere but the effect was not so marked as in air.

After the sliding test, the slider surface was analyzed using AES. Qualitative analysis indicated that the Al_2O_3·TiC surface had also lost some carbon during the sliding test. Therefore, the Ti-LMM Auger spectral profile was measured. Figure 14 shows how the Ti-LMM spectral profile changes with depth when the slider surface is sputter-etched after the sliding test in air. One minute of sputter etching is equivalent to 5 nm of SiO_2 etching. At the sliding surface layer (sputter etched for 0.5 min), the high-energy peak is separated and titanium exists in

its oxidized form. When sputter etching was continued for 5 min, both titanium oxide and titanium carbide were left. When sputter etching was continued for 15 min, however, the titanium carbide profile could be seen. This indicates that frictional heat oxidized the TiC in the Al_2O_3 · TiC surface.

Figure 15 shows the depth of the oxidized layer in the slider surface after the test in air and argon atmospheres as determined using AES. Here, the thickness of the oxidized layer is defined by the sputter etching time when the peak intensity ratios of Ti-LMM and C-KLL become 0.9. This is indicated by the SiO_2 conversion rate. The value of 0.9 is the peak intensity ratio of the fresh Al_2O_3 · TiC slider surface before the sliding test.

The oxidized layer after the sliding test in air was thicker than the sample tested in argon. This proves that the TiC reacted with oxygen in the air and changed into oxide.

A block of Al_2O_3 · TiC was heated in air from 400 °C to 1 000 °C over 30 min, and the temperature at which the oxide layer segregated was determined using X-ray diffraction. At 700 °C, the amount of TiO_2 increased while the amount of TiC decreased. The formation of an oxide layer on the slider surface indicates that the local flash temperature on the slider surface was at least 700 °C.

The results can be summarized as follows:

1) TiC grains were preferentially worn on the Al_2O_3 · TiC surface during the sliding test.

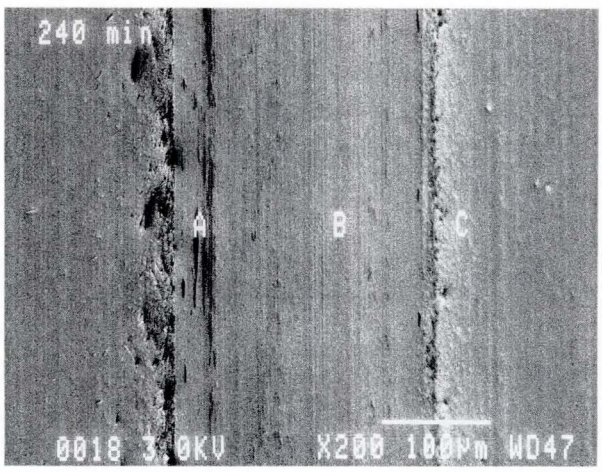


Fig. 16—SEM photograph of worn surface.

2) Preferential wear is attributable to TiC oxidation by frictional heat.

Frictional heat chemically changes the surface layer of the slider material. This must be taken into account when developing future contact recording systems.

4.2 Wear characteristics of the carbon overcoated thin film disk

A 5.25-inch Co alloy thin film disk was used. The disk surface was coated with a 20 nm carbon, film, the surface of which was lubricated with perfluoropolyether. The slider was made of Al_2O_3 · TiC, the applied load was 0.06 N, and the disk rotation speed was 3 600 rpm.

After performing the sliding test for 5 min, comparatively wide black stripes (about 100 μm wider than the rail width) could be seen. AES analysis confirmed that the stripes were made of carbon powder. The SEM photograph in Fig. 16 shows the disk surface after sliding for 240 min. Part B is smoother than the non-sliding parts where is outside the black lines. The surface roughness was measured using a contact profilometer at four sliding traces in the radial direction under the following conditions.

Stylus diameter: 0.5 μm

Measuring length: 100 μm

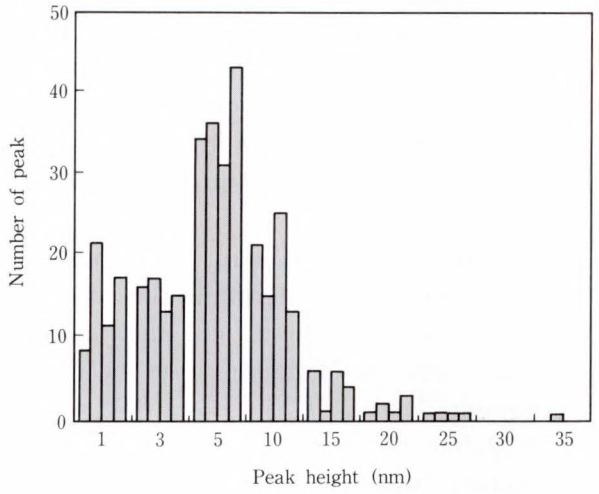
Figure 17 shows histograms on the distribution of peaks before the sliding test, and after five and 120 min of sliding. The highest peak in the non-sliding part was 35 nm. However, 5 min of sliding had removed all peaks higher than 25 nm, and 10 min of sliding had removed all peaks higher than 20 nm. After 120 min of sliding, the distribution did not change. Peaks of 20 nm or higher were removed at the beginning of sliding. All remaining peaks were gradually worn.

Figure 18 shows the relationship between various roughness parameters and the sliding time.

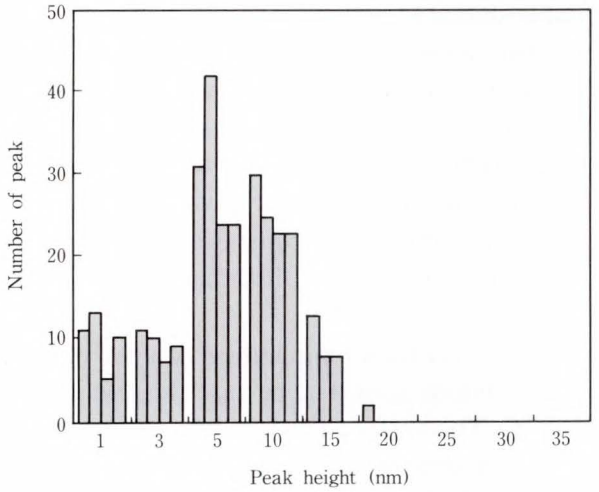
The parameters are as follows:

R_{max} : Distance between the highest peak and lowest valley

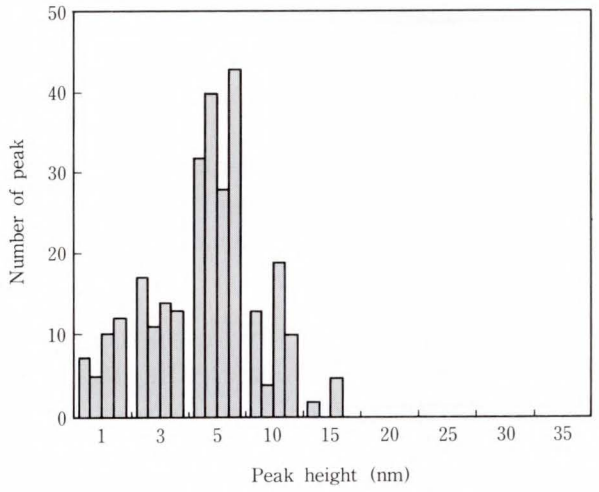
R_z : Distance between the averages of five highest peaks and five lowest valleys



a) Before sliding test



b) After 5 min of sliding



c) After 120 min of sliding

Fig. 17—Distribution of peak height before and after sliding test.

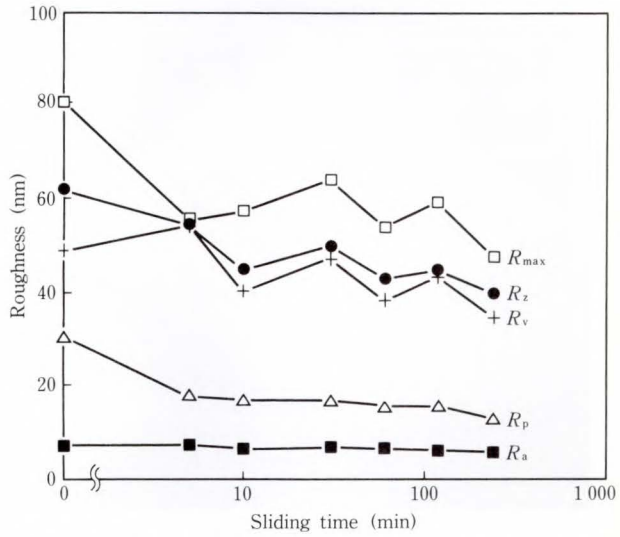


Fig. 18—Changes in roughness with increasing sliding time.

R_v : Distance between the lowest valley and the center line

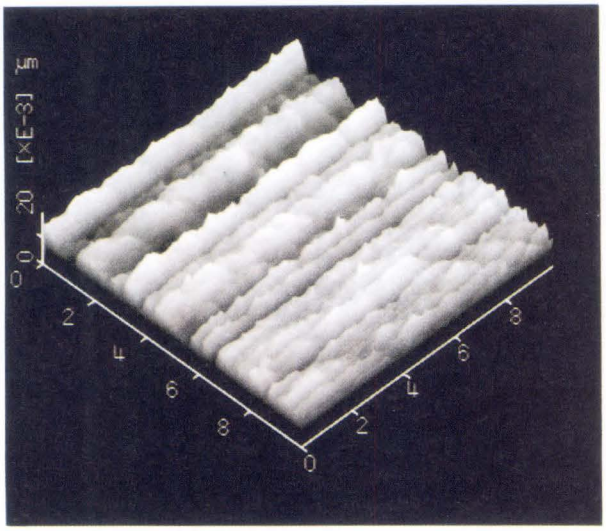
R_p : Distance between the highest peak and the center line

R_a : Average roughness at the center line.

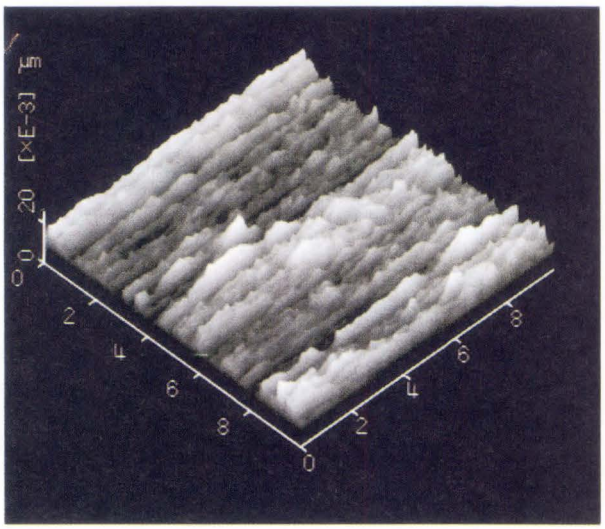
R_p decreased from 31 nm to 18 nm in 5 min, then continued to decline. R_z and R_{max} showed a large initial decrease and then continued to fluctuate (probably because the R_v value was always changing). Overall, these two values generally decreased.

R_v fluctuation may be attributable to powder filling the valleys. R_a hardly changes. Figure 19 shows result of observing the sliding surface using a scanning tunneling microscope (STM). The STM image shows micro-sized ridges and valleys. These features are caused by sliding but cannot be measured with a contact profilometer or SEM.

Figure 20 shows the thickness of the lubricant film during sliding as measured using an XPS. The thickness of the film decreased from 2.4 nm to 1.9 nm during the first 10 min of sliding, and continued to decrease at a rate that was inversely proportional to the logarithm of time. The rapid decrease in the early stage may be due to mechanical loss of lubricant at the projections caused by sliding.



a) Before sliding



b) After sliding

Fig. 19—STM images of disk surface before and after sliding test.

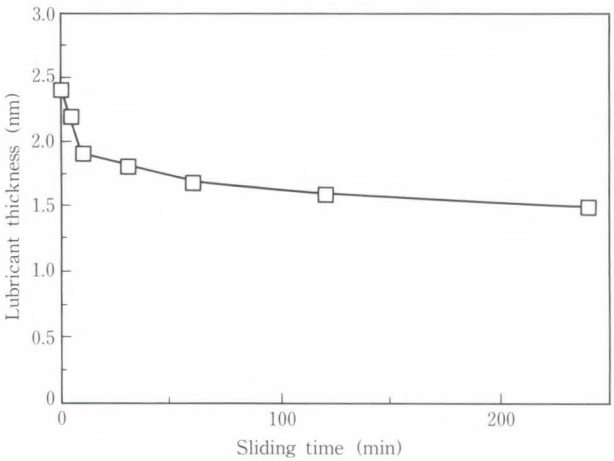


Fig. 20—Lubrication loss versus sliding time.

5. Conclusion

After analyzing the frequency of acoustic emission generated by the head-disk interface, new acoustic emission sensors were developed. These sensors are cheaper than conventional ones yet are highly sensitive to head-disk contacts.

An HDI tester was built to improve the reliability of tests, determine the effects of environmental conditions on the head-disk interface, and to determine the head crash mechanism.

A sliding tester which can change the test atmosphere was also developed. This tester reduces the flying height of the magnetic head by reducing the air pressure. This tester enables stable sliding at high speed.

The influence of gases released in the disk enclosure was investigated. It was found that the selection of drive components, and the control of gas types and quantity significantly affect the reliability of magnetic disk drives.

It was found that dust was removed more easily from a thin film disk than from a particulate disk.

The local flash temperature of an $Al_2O_3 \cdot TiC$ slider sliding against a particulate disk exceeds 700 °C. The surface layer of the slider is chemically changed by the frictional heat. The lubricant film became thin and the carbon film lost its projections.

6. Acknowledgement

The wear characteristics of MnZn ferrite were studied by Mr. T. Yamamoto, one of the authors, during his work at Toyota Technological Institute (TTI). We thank Prof. Nakajima, Associate Prof. Honda, and Assistant Prof. Imada of TTI for their guidance.

References

1) Yeak-Scranton, C.E.: NOVEL PIEZOELECTRIC TRANSDUCERS TO MONITOR HEAD-DISK INTERACTIONS. *IEEE, Trans. Magnetics*, **MAG-22**, 5, pp. 1011-1016 (1986).
2) Terada, A., Ohtani, T., Kimachi, Y., and Yoshikawa, F.: Wear Properties of Lubricated Medium Surface under High Velocity Sliding. Tribol. Mechanics of Magnetic Storage System, **SP25**, pp. 69-73 (1988).

3) Prenosil, M.: Contamination Control in Disk Drive Development. *Res. Develop.*, May, pp. 100-103 (1988).
4) Cormia, R.D., Martin, B., and Zoz, D.A.: Drive-level Organic Contamination Test. *Proc. Head-Disc Technol. Rev.*, November 13 (1988).
5) Yamamoto, T. Takahashi, M., and Shinohara, M.: Influence of Component Outgassing on the Head-Disk Interface. Tribol. Mechanics of Magnetic Storage System, **SP29**, pp. 92-94 (1990).
6) Grundner, M., and Jacob, H.: Investigations on Hydrophilic and Hydrophobic Silicon (100) Wafer Surface by X-Ray Photoelectron and High-Resolution Electron Energy. *Loss-Spectroscopy. Appl. Phys.*, **A.39**, 2, pp. 73-82 (1986).
7) Bhushan, B.: Magnetic Head-Media Interface Temperatures Part 1-Analysis. *Trans ASME, J. Tribol.*, **109**, pp. 243-251 (1987).
8) Bhushan, B.: Magnetic Head-Media Interface Temperatures Part 2-Application to Magnetic Tapes. *Trans ASME, J. Tribol.*, **109**, pp. 252-256 (1987).
9) Ettles, C.M.: Possible Flash Temperature in Slider and Recording Disk Transient Contact. *ALSE Trans.* **29**, 3, pp. 321-328 (1986).
10) Takahashi, M., Shinohara, M., Yamamoto, T., and Tsuchiya, K.: An Investigation of Head-Disk Contact Using Disks Contaminated with Glassbeads. *Proc. SPIE*, **1248**, Storage and Retrieval Systems and Application, pp. 125-132 (1990).
11) Yamamoto, T., Imada, Y., and Nakajima, K.: The Effect of Grain Size on the Wear Beheviour of Mn-Zn Ferrite. *Proc. Inst. Mech. Eng.* , 1987-5, pp. 671-677 (1987).
12) Yamamoto, T., Tsuchiya, K., and Toda, J.: Wear of Al_2O_3·TiC. (in Japanese), Proc. JSLE Okayama, 1988, pp. 29-32.

Takayuki Yamamoto
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Material Physics
Osaka University 1981
Master of Material Science
Toyota Technological Institute 1987
Specializing in Tribology for Head-Disk Interface

Minoru Takahashi
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
FUJITSU Technical College 1972
Specializing in Tribology for Head-Disk Interface

Masayoshi Shinohara
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Applied Chemical Eng.
Tokyo Institute of Technology 1969
Specializing in Magnetic Recording

UDC 621.318:681.327.634

Perpendicular Magnetic Recording

• Junzo Toda • Katsumi Kiuchi • Hiroaki Wakamatsu

(Manuscript received June 4, 1990)

Perpendicular magnetic recording was investigated using a high-efficiency probe-type head and a high signal-to-noise ratio double-layer hard disk. It was found that the read output and the linear bit density sharply increased as the head-to-medium spacing decreased. Particularly, the output value reaches a level more than twice that of longitudinal recording. The sharp magnetic transition in perpendicular recording does not cause an increase in noise at high densities. Low-noise media superior to longitudinal recording media can be realized by controlling the magnetic domain structure of the underlayer.
Consequently, this paper discusses the feasibility for an ultrahigh areal density of over 1.55 bits per square micron (1-Gbit/inch^2).

1. Introduction

Perpendicular recording, which in principle enables high density recording[1], is being developed. A ultrahigh linear bit density of 25 kbit/mm (600 kFCI) has been obtained by combining a probe-type head and a double-layer medium[2]. It has been reported, in flexible disk systems, that a signal-to-noise ratio adequate for practical use can be obtained at a track width of 3.8 μm that is a track density of 250 track/mm (6 350 TPI)[3].

The application of perpendicular recording to hard disks has been under examination[4]. Because the high-density characteristics of perpendicular recording greatly depend on the spacing between the head and medium, it has been difficult to apply perpendicular recording to conventional hard disk drives. However, recent miniaturization of magnetic disk drives allows the use of an ultralow-flying head.

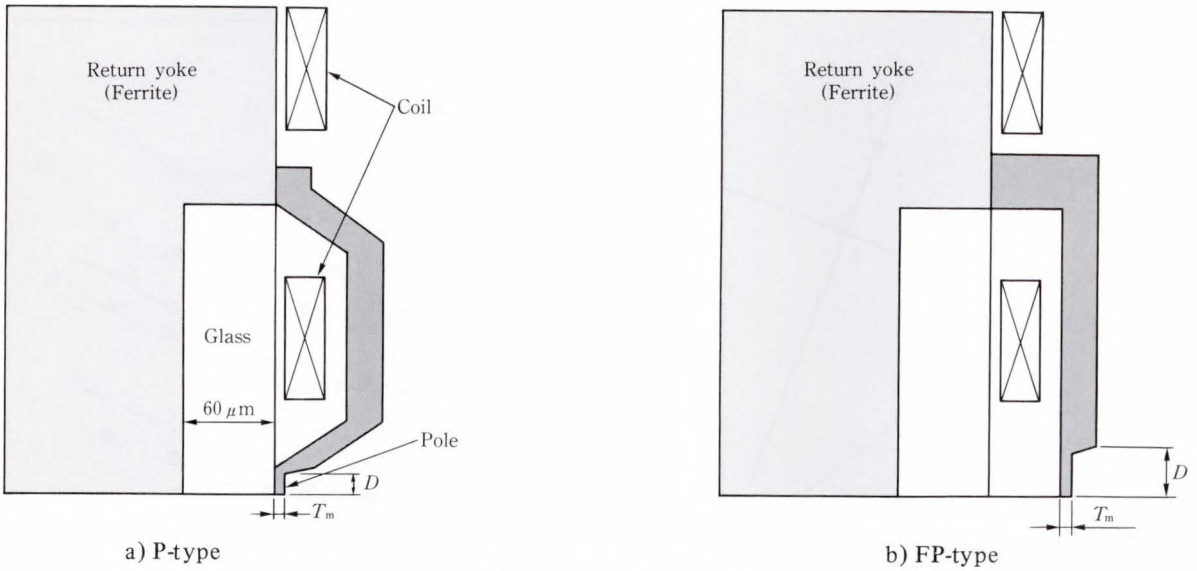
A high-efficiency probe-type thin film head and a high signal-to-noise ratio double-layer hard disk were developed to examine the possibility of high density recording.

This paper describes the structures of the new head and disk, an investigation into the high signal-to-noise ratio and high resolution recording required for areal density improvement, and discusses problems to be solved and their countermeasures.

2. Structures of probe-type thin film head and double-layer hard disk

To enable high recording densities, the write and read efficiency can be improved by constructing a closed magnetic circuit in combination with a probe-type head and a double-layer medium. Two types of probe head with a ferrite return yoke were developed (see Fig. 1 for FP- and P-type head and also Table 1). The magnetic pole is made of amorphous CoZrCr[5] that generates a strong magnetic recording field. The magnetic pole of the FP-type head is flat; therefore, it has an advantage over the P-type head in that it enables easy fabrication of a narrow track width. The magnetic pole and the return yoke edge are 60 μm apart. This wide gap reduces the waveform distortion due to the readback signal at the return yoke edge[6].

A double-layer medium with various magnetic characteristics was prepared. Figure 2 shows the medium cross-section. Table 2 lists the medium specifications. A NiFe underlayer, Ti interlayer, CoCr recording layer, and carbon protective layer are laminated on a disk sub-



a) P-type

b) FP-type

Fig. 1—Probe type thin film head.

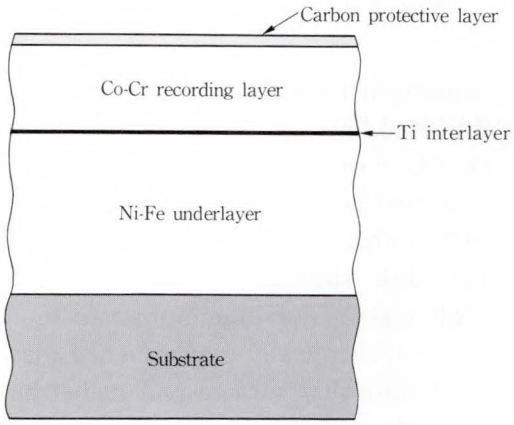


Fig. 2—Double-layer medium.

Table 1. Probe-type thin film head specifications

Return yoke	NiZn ferrite
Coil	33 turns (30 Ω)
Pole	Amorphous CoZrCr (B_s = 1.5 T) Thickness, T_m : 0.4 μm Pole depth, D : 6 μm Track width: 2-11 μm
Flying height	0.1 μm (v = 13 m/s)
Inductance	0.94 μH

Table 2. Double-layer media specifications

	Thickness δ (μm)	$H_c\perp$ (kA/m)	B_s (T)
CoCr	0.1-0.15	16-175	0.12-0.73
Ti	0, 0.02	–	–
NiFe	1-10	–	0.9
Substrate	Glass or NiP plated aluminum (130 mm dia)		

strate with a diameter of 130 mm. The NiFe underlayer is deposited by electroplating or RF sputtering. The other layers are deposited by RF sputtering. The saturation flux density (B_s) and perpendicular coercivity ($H_c\perp$) of the CoCr layer are controlled by the target composition and sputtering conditions.

3. Examination of recording density improvement

3.1 Read output

Read output on the perpendicular recording is explained by the magnetic circuit model, consisting of a head and medium[7)-8)]. The read output is affected by the operating point (A or B), which is defined as the intersection of the magnetic circuit load line and the B-H curve of the CoCr layer (see Fig. 3). To increase the read output, the flux density on the operating point must be increased. There are two approaches to increase it. One is to increase the perpendicular coercivity ($H_c\perp$) and another is to increase the permeance factor (P_r) which is the gradient of the load line. The permeance factor can be increased by reducing the head-to-medium spacing. Figure 4 shows the dependence of the permeance factor on the head-to-medium spacing.

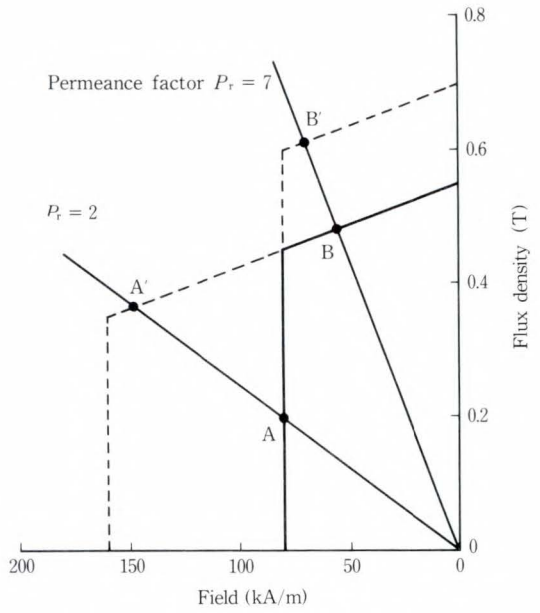


Fig. 3—Operating point of Co-Cr film.

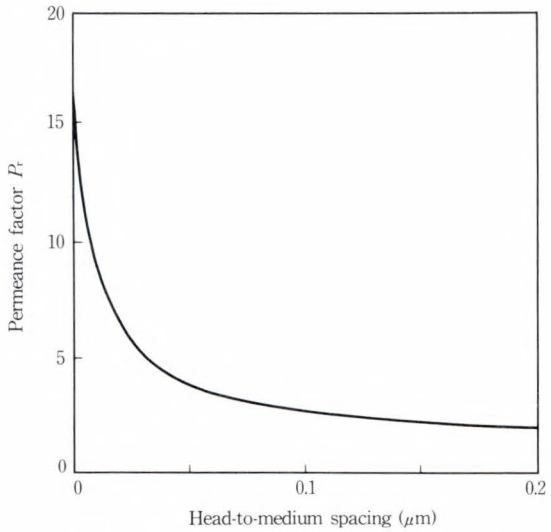


Fig. 4—Dependence of P_r on spacing.

With the current hard disk system, the head-to-medium spacing (flying height + medium protection film thickness + magnetic pole recess depth) is 0.1-0.2 μm and the permeance factor is about 2. The operating point for a CoCr layer having an ideal squareness increases along the A-A′ load line in proportion to the perpendicular coercivity ($H_c\perp$). To increase the read output, therefore, high coercivity is required up to about 160 kA/m (2 000 Oe). In the spacing less than 0.05 μm, however, the read output does not depend on the coercivity but on

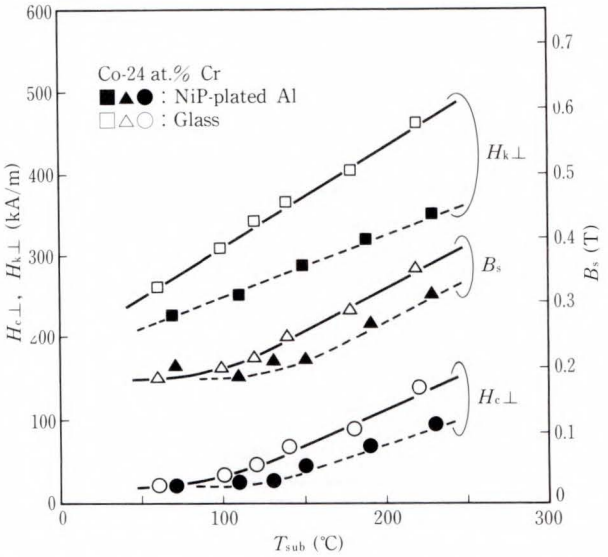


Fig. 5—Dependence of magnetic properties of NiP plated Al and glass substrate on substrate temperature (δ_{CoCr} = 0.15 μm).

the saturation flux density (B_s) as shown with the B-B′ load line.

On the basis of the magnetic circuit model analysis described above, studies were conducted to realize a high coercivity medium of 160 kA/m.

The disk substrates were examined first. The NiP plated aluminum substrate in current use, and a chemically strengthened glass substrate of smoother surface and higher hardness were investigated. Figure 5 shows the magnetic properties of the Co-24 at.%Cr films on NiP plated aluminum substrate and the glass substrate. A CoCr film having excellent magnetic properties, i.e. high coercivity ($H_c\perp$), high perpendicular anisotropy field ($H_k\perp$), and high saturation flux density (B_s) can be obtained on a glass substrate. This can be inferred from the fact that Cr segregation within the column[9] is promoted because the surface temperature during sputtering of the glass substrate is higher than that of the NiP plated aluminum substrate. This difference is due to the lower thermal conductivity of the glass substrate. A chemically strengthened glass substrate shows promise as a way to improve the magnetic properties of a CoCr film.

Figure 6 shows the dependence of the coercivity and saturation flux density on the sub-

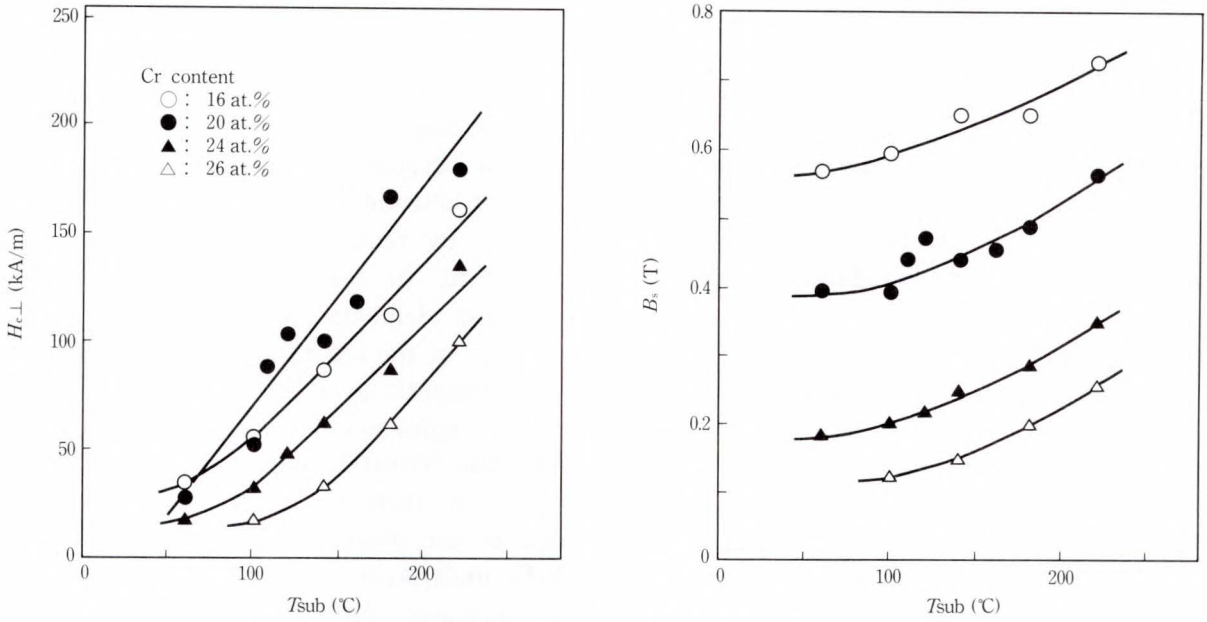


Fig. 6—Dependences of H_c and Bs on substrate temperature ($\delta_{CoCr} = 0.15\ \mu m$).

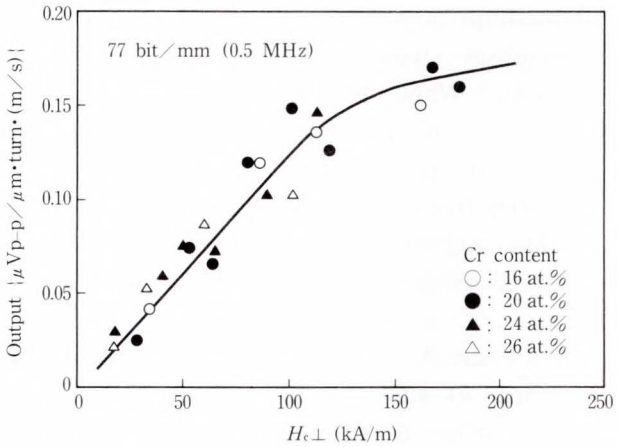


Fig. 7—Dependence of output on H_c ($\delta_{CoCr} = 0.15\ \mu m$, Head-to-medium spacing = 0.15 μm.

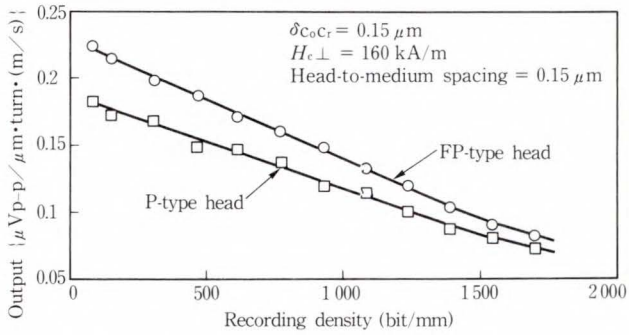


Fig. 8—Recording characteristics of FP- and P-type head.

strate temperature (T_{sub}) for the CoCr film with a 16-26 at.%Cr content on the glass substrate. The coercivity and saturation flux density increase with substrate temperature. For various substrate temperatures, the coercivity reaches the maximum value at a 20 at.%Cr content. A preferable high-coercivity film can be obtained by setting the target composition to 20 at.%Cr and by raising the substrate temperature above 180 °C.

The recording characteristics were measured using a probe-type head. Figure 7 shows the dependence of the read output on the coercivity at a spacing of 0.15 μm. The read output increases in proportion to coercivity ($H_c\perp$) independently of the Cr content, and is saturated when the coercivity exceeds 120 kA/m (1 500 Oe). This result agrees with the result of the magnetic circuit model analysis shown in Fig. 3. The saturation of the output implies that the operating point is at the shoulder of the B-H curve.

Figure 8 compares the recording characteristics of a P- and FP-type head combined with a high-coercivity medium. The read efficiency of the FP-type head is 20 percent higher than that of the P-type head. Figure 9 shows the normalized output, which remains constant even at a 2 μm track width.

Figure 10 shows the dependence of the read output on the spacing for an FP-type head with

a high-coercivity medium. The read output sharply increases as the spacing decreases. This is because of the increase in the magnetic interaction between the head and medium. An output of 0.4 $\mu V_{p\text{-}p}/\mu m \cdot turn \cdot (m/s)$ is obtained at a spacing of 0.08 μm. This value is more than twice that achieved using a longitudinal recording head and medium.

3.2 Medium noise

To improve the signal-to-noise ratio, the read output must be increased and the noise be reduced. It has been reported that the noise in metal thin-film media for longitudinal recording increases with the recording density increased because the noise is generated by the zigzag transition[10]. The zigzag transition is suppressed by reducing the magnetic grain size. Therefore, to reduce the media noise, nonmagnetic elements are added to the recording material or the recording layer is laminated with a nonmagnetic layer[11),12)].

In perpendicular recording media, the magnetic transition is very sharp and does not cause an increase in noise. However, noise due to the magnetic domain structure of the NiFe underlayer is noticeable[13)]. Read sensitivity increases with the underlayer thickness. A thick underlayer, however, causes an increase in noise level. Figure 11 shows examples of medium noise spectra for various underlayer thicknesses with recording layer thickness constant. When the underlayer thickness is increased, an increase in noise occurs at a specific frequency that is independent of the recording frequency.

The magnetic domains of underlayers were observed using the Bitter method, and the noise spectra of a medium without CoCr layers was measured. It was found that the increase in noise at the specific frequency is caused by the stripe domains of the underlayer that have perpendicular anisotropy, especially the

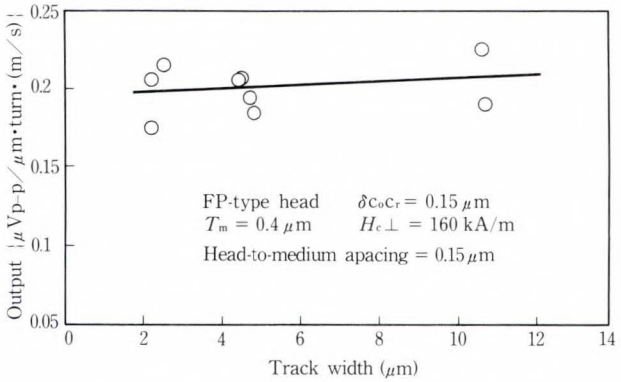


Fig. 9—Track width and read output.

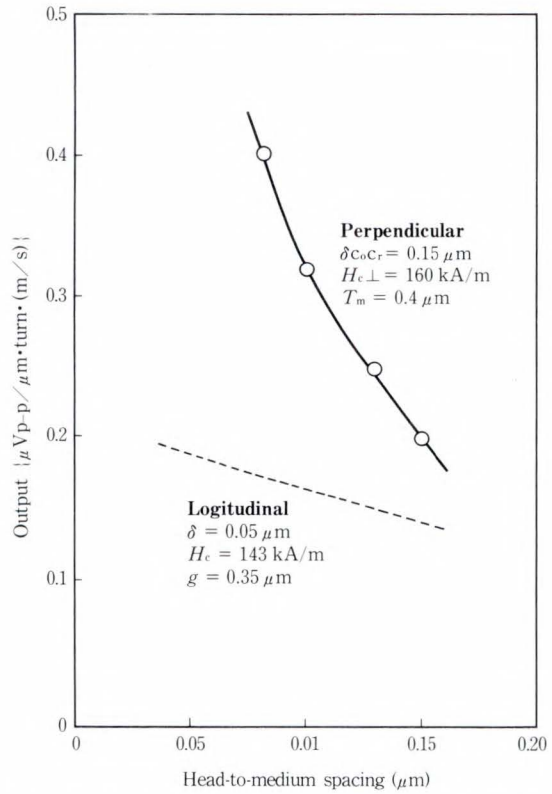


Fig. 10—Dependence of output on spacing in perpendicular and longitudinal recording.

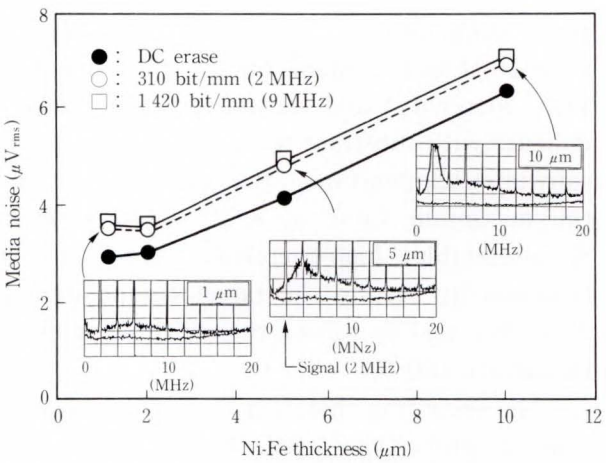


Fig. 11—Dependence of media noise on Ni-Fe thickness (δ_{CoCr} = 0.15 μm).

stripe domains that extend in the disk radius direction (see Fig. 12)[14].

The medium noise can be reduced by controlling the magnetic domain structure, for example, by adjusting the direction of stripe domains that extend in the circumferential direction or by reducing the number of domain walls. Figure 13 shows the controlled underlayer domain structures and noise spectra. The noise at the specific frequency decreases and a medium noise of 4 μV_{rms} was obtained (bandwidth = 20 MHz). This value is the same as that achieved using a CoCrTa longitudinal recording medium, which is noted for its low noise characteristics. The noise level can be further reduced by reducing the grain size of the recording layer and underlayer, and by reducing the exchange coupling between CoCr magnetic grains.

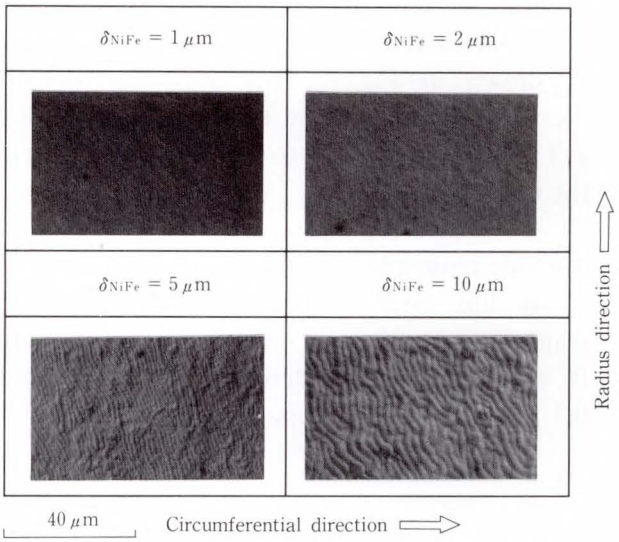


Fig. 12—Domain structures of underlayers having different thicknesses.

3.3 Resolution

A report on single-layer media states that a CoCr layer on a Ti layer has satisfactory crystal orientation and magnetic properties[4]. In conventional double-layer media, a 0.02 μm thick Ti interlayer is deposited between the NiFe underlayer and CoCr recording layer. However, this interlayer weakens the head-to-medium interaction, and is detrimental to high density recording at low spacing.

A medium without a Ti interlayer was prepared and the recording density character-

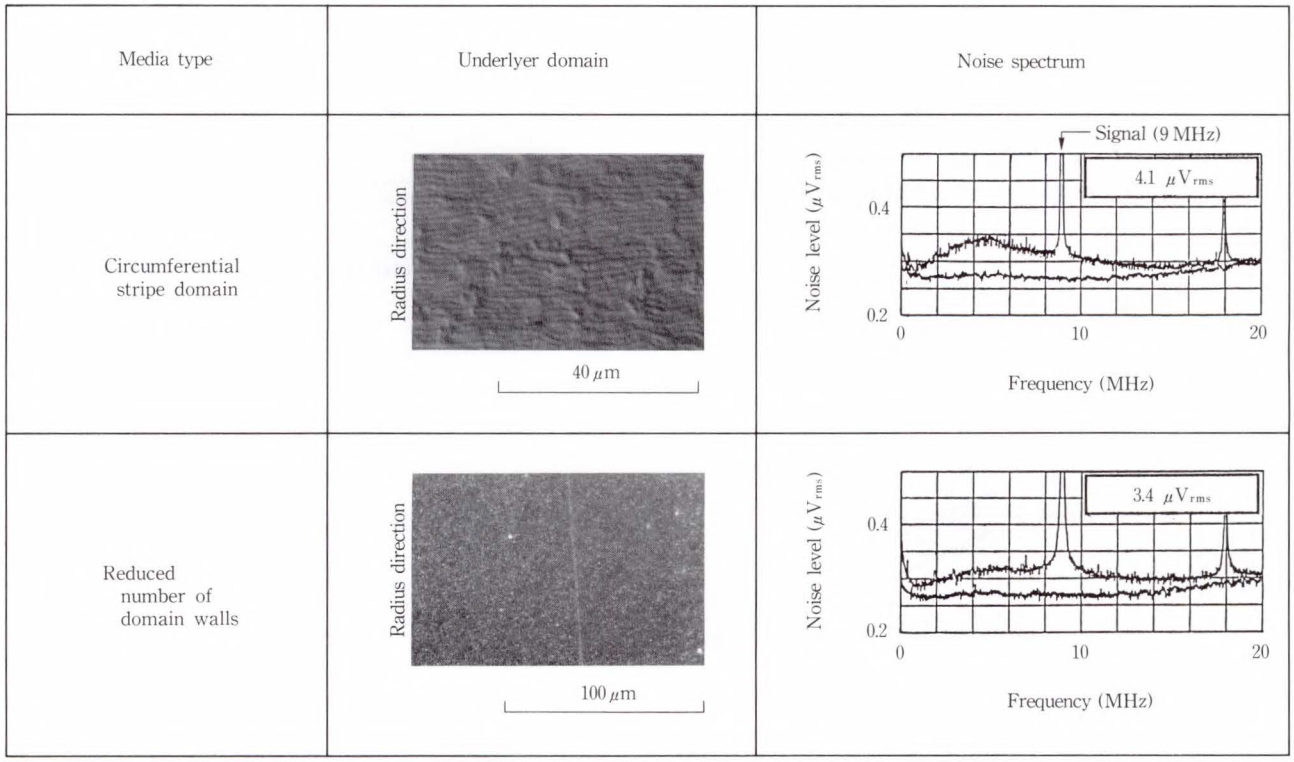


Fig. 13—Noise spectra of media with controlled underlayer domain.

istics were examined using a head with a pole thickness of 0.4 μm. Figure 14 compares D_{50} of media without a Ti interlayer with that of media with a Ti interlayer. When the CoCr thickness is 0.15 μm, D_{50} of a medium with a Ti interlayer at a spacing of 0.08 μm is only 1 890 bit/mm (48 kFCI), while that of a medium without a Ti interlayer is 2 360 bit/mm (60 kFCI). When the CoCr thickness is 0.1 μm and the Ti interlayer is eliminated, D_{50} increases to 2 800 bit/mm (71 kFCI). Then D_{50} of 4 000 bit/mm (100 kFCI) was obtained by reducing the pole thickness and the head-to-medium spacing. This value agrees with the data reported in a study on a nonflying head[15].

When the CoCr film thickness was 0.15 μm, there was little difference in the read output between a medium with a Ti layer and a medium without a Ti layer. However, when the CoCr thickness was 0.1 μm, the output of the medium without a Ti layer was only 70 percent the output of the medium with a Ti layer.

The CoCr perpendicular coercivities of these media were measured using the Kerr effect. The results showed that the coercivity of the surface layer was independent of the existence of the Ti layer. Then, the coercivity throughout the layer depth was measured by ion-milling the CoCr layer. Figure 15 shows the results. A medium with a Ti layer has a high coercivity of 160 kA/m (2 000 Oe) throughout the CoCr initial layer. A medium without a Ti layer has a high coercivity of 160 kA/m on the surface, but the coercivity sharply decreases near the CoCr/NiFe boundary. The coercivity of CoCr layers thicker than 0.1 μm is constant.

Figure 16 shows the relationship between the effective coercivity and read output. The effective coercivity is defined as the mean of the coercivity measured at the surface and at a depth of 0.05 μm. The read output is proportional to the coercivity in the same manner as shown in Fig. 7. This indicates that the decrease in the output of a medium without a Ti interlayer is due to the low coercivity of the CoCr initial layer.

3.4 Side cross-talk characteristics

It has been reported that, because the magnetic transition of perpendicular recording is very sharp, the side-written width is only 0.1 μm and the side cross-talk is relatively low[16].

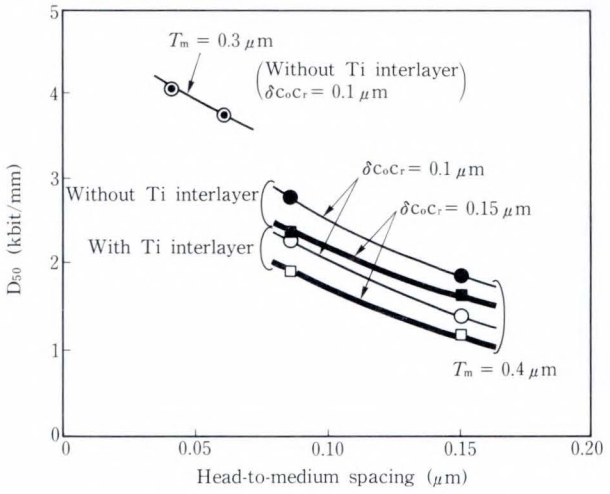


Fig. 14—Dependence of D_{50} on spacing.

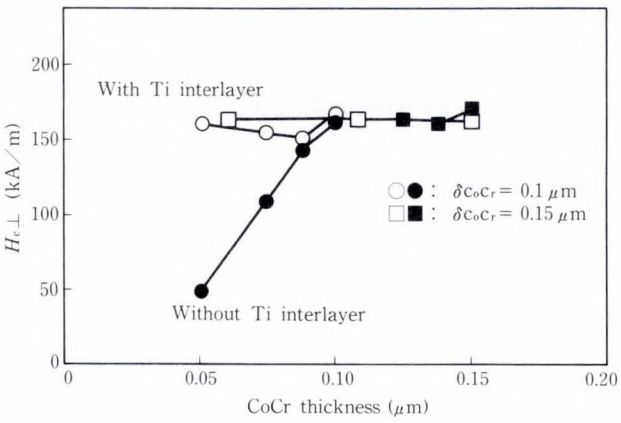


Fig. 15—Change in coercivity with CoCr depth.

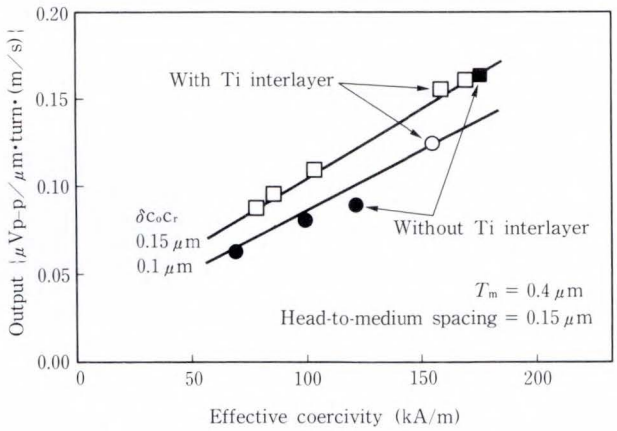
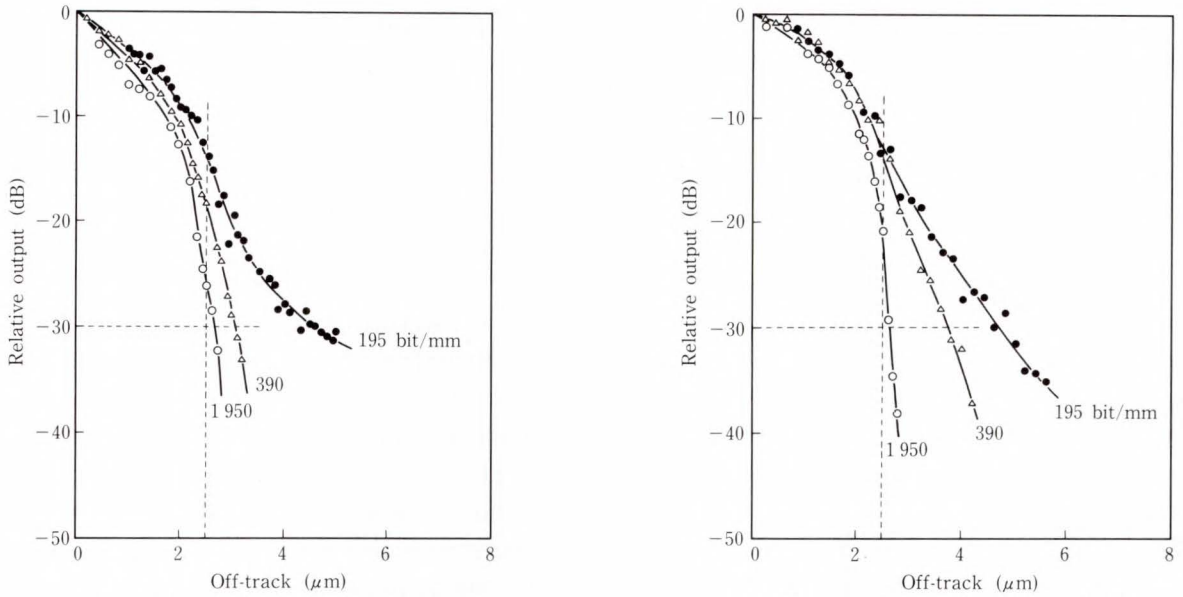


Fig. 16—Output and effective coercivity.



a) Head-to-medium spacing = 0.08 μm

b) Head-to-medium spacing = 0.13 μm

Fig. 17—Side cross-talk characteristics (track width = 2.5 μm).

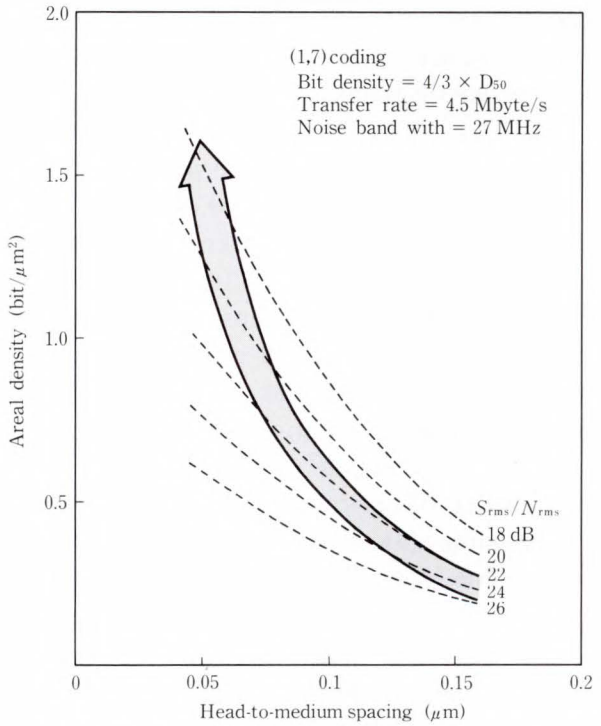


Fig. 18—Potential for ultrahigh areal density.

Figure 17 shows the side cross-talk characteristics of the FP-type head with a 2.5 μm track width. For recording densities over 390 bit/mm (10 kFCI), a cross-talk level of less than −30 dB is obtained with a narrow guard-band between adjacent tracks of 0.5 μm with a head-to-medium spacing of 0.08 μm.

3.5 Head-to-medium spacing and areal recording density

The areal recording density was estimated from the experimental data on the FP-head with the high coercivity disk. Figure 18 shows the relationship between areal density and head-to-medium spacing. The signal-to-noise ratio was calculated using the signal output at D_{50} and the noise in the 27 MHz band by assuming the code was (1, 7) and the transfer rate was 4.5 Mbyte/s. An areal density of 0.78 bit/μm² (500 Mbit/inch²) can be realized at a 0.08 μm spacing and a 0.4 μm pole thickness with a signal-to-noise ratio of 21 dB. The areal density can be increased above 1.55 bit/μm² (1 Gbit/inch²) by reducing the spacing and pole thickness and by using low signal-to-noise ratio signal processing.

4. Ultrahigh density to come

To put the ultrahigh-density recording described above into practical use, several problems must be solved.

The first is the reliability problem of the head and disk due to the ultralow spacing. A flying height of 0.1 μm is already used in small disk drives because the contact energy is reduced due to the low relative velocity.

To reduce the contact energy, the load force and slider size must be reduced. Lubrication on the extremely thin disk protection film is also a key to good reliability.

The second problem is the output reduction or fluctuation observed with a narrow track width[17]. This is caused by the closure domain of the magnetic pole. The output reduction or fluctuation can be suppressed to some extent by changing the pole shape and by controlling the anisotropy field of the magnetic pole. Laminating the magnetic pole with nonmagnetic layers is also effective in controlling the magnetic domain structure.

The third problem is head impedance noise. Medium noise decreases as the track width is reduced, but head impedance noise remains constant. As a result, the ratio of impedance noise to the total noise increases as the track width narrows. The inductance of a probe-type head having a ferrite return yoke is greater than that of a ring-type thin film head for longitudinal recording. Therefore, the probe-type head generates more noise than the ring-type thin film head. To reduce the resistance and inductance, the coil must be miniaturized.

The fourth problem is poor off-track allowance in the presence of prewritten information[18]. Because perpendicular recording has a very small side-write area, a slight off-track causes noise due to incomplete erasure of prewritten information. However, a small side-write area enables the track width to be accurately defined, which is essential for high track density recording. Perpendicular recording is superior when applied to disk drives with high head positioning accuracy.

The fifth problem is the erasure of information by external magnetic fields. An external magnetic field of only 240 A/m (3 Oe) can be sufficiently concentrated to erase information by the magnetic pole of a probe-type head.

Magnetic shielding is effective against this problem. (In an experiment, the output signal did not decrease in an external field of over 1.6 kA/m (20 Oe) when the head was shielded.)

5. Conclusion

A high-efficiency probe-type head and a high signal-to-noise ratio double-layer hard disk for perpendicular magnetic recording were developed, and the possibility of high density recording was examined.

The read output is proportional to the coercivity of the medium and sharply increases as the head-to-medium spacing is reduced. It was confirmed that a high output (more than twice that of longitudinal recording) can be obtained from a high-coercivity medium.

The noise in perpendicular recording is caused dominantly by the magnetic domain structure of the NiFe underlayer. A noise level equivalent to that in a CoCrTa longitudinal low noise recording medium was obtained by controlling the domain structure. The noise can be further reduced by reducing the grain size of the recording layer and underlayer, and by reducing the exchange coupling between CoCr magnetic grains.

A D_{50} of 2 800 bit/mm (71 kFCI) was obtained at a spacing of 0.08 μm using a head with a pole thickness of 0.4 μm. Reducing the spacing and the pole thickness enables a D_{50} of more than 4 000 bit/mm (100 kFCI).

The areal recording density was estimated on the basis of the above examination. It was found that perpendicular magnetic recording has the potential for an ultrahigh areal density of over 1.55 bit/μm^2 (1 Gbit/inch^2), which exceeds that of optical disks.

Magnetic disk memory has played a leading part in file systems because of their excellent speed, capacity, and bit-cost. The availability of perpendicular magnetic recording offering ultrahigh areal density suggests that the magnetic disk memory will maintain its position for a long time to come.

References

1) Iwasaki, S., and Nakamura, Y.: AN ANALYSIS FOR THE MAGNETIZATION MODE FOR HIGH DENSITY MAGNETIC RECORDING. *IEEE Trans. Magnetics*, **MAG-13**, 5, pp. 1272-1277 (1977).
2) Yamamoto, S., Nakamura, Y., and Iwasaki, S.: EXTREMELY HIGH BIT DENSITY RECORDING

WITH SINGLE-POLE PERPENDICULAR HEAD. *IEEE Trans. Magnetics*, **MAG-23**, 5, pp. 2070-2072 (1987).

3) Numazawa, J., Yoneda, Y., Aruga, F., and Horiuchi, T.: READ/WRITE CHARACTERISTICS OF NARROW TRACK WIDTH SINGLE-POLE HEAD. (in Japanese), *IEICE Tech. Rep., Mag. Rec.*, **MR88**-14., pp. 1-8 (1988).
4) Kobayashi, K., Toda, J., and Yamamoto, T.: High Density Perpendicular Magnetic Recording on Rigid Disks. *FUJITSU Sci. Tech. J.*, **19**, 1, pp. 99-126 (1983).
5) Hata, K., and Takahashi, Y.: ENVIRONMENTAL RESISTIVITY OF AN AMORPHOUS Co-Zr THIN FILM HEAD. (in Japanese), Nat. Conf. Rec., 1985, Semiconductor Dev. Mater., IECE, Jpn., 1-42, 1985, p. 42.
6) Koshikawa, T., Kanda, E., Kanamine, M., Hosono, K., and Takagi, H.: Edge-Noise Evaluation of Single-Pole Thin Film Head. *J. Mag. Soc., Jpn.*, **13**, Suppl. Sl, pp. 523-527 (1989).
7) Kiuchi, K., Wakamatsu, H., Suzuki, F., and Takagi, H.: HIGH-ENERGY CoCr THIN FILMS SPUTTERED ON GLASS DISKS FOR PERPENDICULAR RECORDING. *IEEE Trans. Magnetics*, **24**, 6, pp. 2341-2343 (1988).
8) Kobayshi, K.: Analysis of Output Voltage in Perpendicular Recording. (in Japanese), *J. Mag. Soc., Jpn.*, **12**, 3, pp. 479-484 (1988).
9) Maeda, Y., and Asahi, M.: SEGREGATION IN SPUTTERED Co-Cr FILMS. *IEEE Trans. Magnetics*, **MAG-23**, 5, pp. 2061-2063 (1987).
10) Aoi, H., Tsuchiya, R., Shiroishi, Y., and Matsuyama, H.: A STUDY ON THIN-FILM-MEDIA NOISE BY MAGNETIC DOMAIN OBSERVATION. *IEEE Trans. Magnetics*, **24**, 6, pp. 2714-2717 (1988).
11) Shiroishi, Y., Matsuda, Y., Yoshida, K., Suzuki, H., Ohno, T., Tsumita, N., Ohura, M., and Hayashi, M.: READ AND WRITE CHARACTERISTICS OF Co-ALLOY/Cr THIN FILMS FOR LONGITUDINAL RECORDING. *IEEE Trans. Magnetics*, **24**, 6, pp. 2730-2732 (1988).
12) Murdock, E.S., Natarajan, B.R., and Walmsley, R.G.: NOISE PROPERTIES OF MULTILAYERED Co-ALLOY MAGNETIC RECORDING MEDIA. *IEEE Trans. Magnetics*, **26**, 5, pp. 2700-2705 (1990).
13) Ouchi, K., Akao, Y., and Iwasaki, S.: High Rate Fabrication of Perpendicular Magnetic Recording Medium. (in Japanese), IEICE Tech. Rep., Group Meet. Mat., MR84-6, pp. 9-26 (1984).
14) Kiuchi, K., Wakamatsu, H., Suzuki, F., and Kanamine, M.: Effects of Underlayer Thickness and External Fields on Noise in Perpendicular Double-Layered Media. *J. Mag. Soc., Jpn.*, **13**, Suppl. Sl, pp. 301-305 (1989).
15) Nakamura, Y., Ouchi, K., Yamamoto, S., and Watanabe, I.: RECORDING CHARACTERISTICS OF PERPENDICULAR MAGNETIC HARD DISK MEASURED BY NON-FLYING SINGLE-POLE HEAD. *IEEE Trans. Magnetics*, **26**, 5, pp. 2436-2438 (1990).
16) Numazawa, J., Yoneda, Y., Aruga, F., and Horiuchi, T.: Side Crosstalk Response in Perpendicular Magnetic Recording with Narrow Track Width Single-Pole Head. (in Japanese), *J. Mag. Soc., Jpn.*, **13**, 2, pp. 121-124 (1989).
17) Nakamura, Y., Yamakawa, K., and Iwasaki, S.: ANALYSIS OF DOMAIN STRUCTURE OF SINGLE POLE PERPENDICULAR HEAD. *IEEE Trans. Magnetics*, **MAG-21**, 5, pp. 1578-1580 (1985).
18) Beaulieu, T.J., Seagle, D.J., Meininger, M.A., and Spector, C.J.: TRACK DENSITY LIMITATION FOR DUAL-LAYER PERPENDICULAR RECORDING IN A RIGID DISK ENVIRONMENT. *IEEE Trans. Magnetics*, **25**, 5, pp. 3369-3371 (1989).

Junzo Toda
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Electrical Eng.
Osaka University 1974
Master of Electrical Eng.
Osaka University 1976
Specializing in Magnetic Recording

Hiroaki Wakamatsu
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Electronic Eng.
Doshisha University 1980
Specializing in Magnetic Recording

Katsumi Kiuchi
File Memory Laboratory
FUJITSU LABORATORIES, ATSUGI
Bachelor of Information Eng.
Yokohama National University 1982
Master of Information Eng.
Yokohama National University 1984
Specializing in Magnetic Recording

FUJITSU

SCIENTIFIC & TECHNICAL JOURNAL

VOL.26, NOs.1-4

INDEX

• SUBJECT INDEX • AUTHOR INDEX

Subject Index

(Ⓘ mark is Invited paper.)

P

S

T

U

Author Index

(Ⓘ mark is Invited paper.)

K

M

N

O

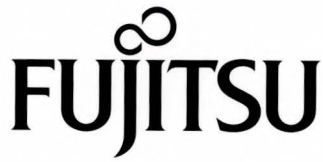
S

T

U

W

Y


FUJITSU

Overseas Offices

Abu Dhabi Office
Suite 802, A1 Masaood Tower,
Seikh Hamdan Street, Abu Dhabi,
U.A.E.
Telephone: (971-2)-333440
FAX: (971-2)-333436

Algiers Office
9, Rue Louis Rougie Chateau Neuf,
EL Biar, Alger 16030, Algeria
Telephone: (213)-2-78-5542
Telex: 408-67522

Amman Office
P.O. Box 5420, Ammán, Jordan
Telephone: (962)-6-662417
FAX: (962)-6-662417

Bangkok Office
Room 405, 4th Floor, Dusit
Thani Bldg., 1-3, Rama IV.
Bangkok, Thailand
Telephone: (66-2)-236-7930
FAX: (66-2)-238-3666

Beijing Office
Room 2101, Fortune Building
5 Dong San Huan Bei-lu,
Chao Yan District, Beijing,
People's Republic of China
Telephone: (86-1)-501-3261
FAX: (86-1)-501-3260

Bogotá Office
Cra. 13 No. 27-50,
Edificio Centro Internacional
Tequendama, Oficina 326/328
Bogotá, D.E. Colombia
Telephone: (57-1)-286-7061
FAX. (57-1)-286-7148

Brussels Office
Avenue Louise 176, Bte 2
1050 Brussels, Belgium
Telephone: (32-2)-648-7622
FAX: (32-2)-648-6876

Hawaii Branch
6660 Hawaii Kai Drive, Honolulu,
Hawaii 96825, U.S.A.
Telephone: (1-808)-395-2314
FAX: (1-808)-396-0059

Indonesia Project Office
16th Floor, Skyline Bldg.,
Jalan M.H. Thamrin No. 9,
Jakarta, Indonesia
Telephone: (62-21)-3105710
FAX: (62-21)-3105983

Jakarta Office
16th Floor, Skyline Bldg.,
Jalan M.H. Thamrin No. 9,
Jakarta, Indonesia
Telephone: (62-21)-333245
FAX: (62-21)-327904

Kuala Lumpur Office
Letter Box No. 47, 22nd Floor,
UBN Tower No. 10, Jalan P.
Ramlee, 50250, Kuala Lumpur,
Malaysia
Telephone: (60-3)-238-4870
FAX: (60-3)-238-4869

London Office
2, Longwalk Road, Stockley Park,
Uxbridge, Middlesex, UB11 1AB,
England
Telephone: (44-81)-573-4444
FAX: (44-81)-573-2643

Munich Office
c/o DV18 Siemens AG, Otto-
Hahn-Ring 6, D-8000, München
83, F.R. Germany
Telephone: (49-89)-636-3244
FAX: (49-89)-636-45345

New Delhi Office
Mercantile House, 1st Floor,
15, Katsurba Gandhi Marg
New Delhi-110 001, India
Telephone: (91-11)-331-1311
FAX: (91-11)-332-1321

New York Office
680 Fifth Avenue, New York,
N.Y. 10019, U.S.A.
Telephone: (1-212)-265-5360
FAX: (1-212)-541-9071

Paris Office
Bâtiment Aristote, 17
Rue Olof Palme 94006,
Creteil Cedex, France
Telephone: (33-1)-4-399-0897
FAX: (33-1)-4-399-0700

Shanghai Office
Room 1504 Ruijin Bldg.,
205 Maoming Road South,
Shanghai, People's Republic
of China
Telephone: (86-21)-336-462
FAX: (86-21)-336-480

Taipei Office
Sunglow Bldg., 66, Sung Chiang
Road, Taipei, Taiwan
Telephone: (886-2)-561-7715
FAX: (886-2)-536-7454

Washington, D.C. Office
1776 Eye Street, N.W.,
Suite 880, Washington, D.C.,
20006, U.S.A.
Telephone: (1-202)-331-8750
FAX: (1-202)-331-8797

Overseas Subsidiaries

		Telephone	Facsimile
FKL Dong-Hwa Ltd.	338-13, Daehong-Ri, Sunghwan-Eub, Chunwon-Gun, Chungnam, Republic of Korea	(82-417)-581-0701	(82-417)-581-0700
Fujian Fujitsu Communications Software Ltd.	Wu, Li Ting, Fu Ma Lu, Fuzhou, Fujian Province, People's Republic of China	(86-591)-560070	(86-591)-560022
Fujitsu America, Inc.	3055 Orchard Drive, San Jose, CA 95134-2017, USA	(1-408)-432-1300	(1-408)-432-1318
Fujitsu Australia Ltd.	475 Victoria Ave., Chatswood, NSW 2067, Australia	(61-2)-410-4555	(61-2)-411-8603
Fujitsu Australia Software Technology Pty. Ltd.	1st Floor, Techway House, 18 Rodborough Road, French Forest, N.S.W. 2086, Australia	(61-2)-936-1111	(61-2)-975-2899
Fujitsu Business Communication Systems, Inc.	3190 Mira Loma Ave., Anaheim, CA 92806, USA	(1-714)-630-7721	(1-714)-630-7660
(Sales Headquarter)	2411 West 14th Street, Tempe, AZ 85281, USA	(1-602)-921-5900	(1-602)-921-5999
Fujitsu Canada, Inc.	6280 Northwest Drive, Mississauga, Ontario, Canada L4V 1J7	(1-416)-673-8666	(1-416)-673-8677
Fujitsu Component (Malaysia) Sdn. Bhd.	No. 1, Lorong Satu, Kawasan Perindustrian Parit Raja, 86400 Batu Pahat, Johor, Malaysia	(60-7)-482-111	(60-7)-481-771
Fujitsu Component of America, Inc.	3330 Scott Boulevard, Santa Clara, CA 95054-3197, USA	(1-408)-562-1000	(1-408)-748-7655 or (1-408)-727-0355
Fujitsu Customer Service of America, Inc.	11085 N. Torrey Pines Rd. La Jolla, CA 92037, USA	(1-619)-457-9900	(1-619)-457-9968
Fujitsu Deutschland GmbH	Frankfurter Ring 211, 8000 München 40, F.R. Germany	(49-89)-323-780	(49-89)-323-78100
Fujitsu do Brasil Ltda.	Rua Manoel da Nóbrega, 1280-2 Andar, C.E.P. 04001, São Paulo, SP, Brazil	(55-11)-885-7099	(55-11)-885-9132
Fujitsu España, S.A.	Edificio Torre Europa 5ª, Paseo de la Castellana, 95, 28046 Madrid, Spain	(34-1)-581-8000	(34-1)-581-8300
Fujitsu Europe Ltd.	2, Longwalk Road, Stockley Park, Uxbridge, Middlesex, UB11 1AB, England	(44-81)-573-4444	(44-81)-573-2643
Fujitsu Europe Telecom R&D Centre Limited	2, Longwalk Rd., Stockley Park, Uxbridge, Middlesex, UB11 1AB, England	(44-81)-576-0286	(44-81)-573-2643
Fujitsu Finance (U.K.) PLC	2, Longwalk Road, Stockley Park, Uxbridge, Middlesex, UB11 1AB, England	(44-81)-569-1628	(44-81)-573-2643
Fujitsu France S.A.	Bâtiment Aristote 17, Rue Olof Palme 94006, Creteil Cedex, Paris, France	(33-1)-4-399-4000	(33-1)-4-399-0700
Fujitsu Hong Kong Ltd.	Room 2521, Sun Hung Kai Centre, 30 Harbour Road, Hong Kong	(852)-8915780	(852)-5721724
Fujitsu Imaging Systems of America, Inc.	3 Corporate Drive, Commerce Park, Danbury, CT 06810, USA	(1-203)-796-5400	(1-203)-796-5665 or (1-203)-796-5723
Fujitsu International Finance (Netherlands) B.V.	Officia 1, De Boelelaan 7, 1083 HJ Amsterdam, The Netherlands	(31-20)-465996	(31-20)-15123
Fujitsu Italia S.p.A.	Via Melchiorre, Gioia No. 8, 20124 Milano, Italy	(39-2)-657-2741	(39-2)-657-2257
Fujitsu Korea Ltd.	9th Floor, Korean Reinsurance Bldg., 80, Susong-Dong, Chongro-Gu, Seoul Special City, Republic of Korea	(82-2)-739-3281	(82-2)-739-3294
Fujitsu Microelectronics Asia Pte. Ltd.	No. 2, Second Chin Bee Road, Jurong Town, Singapore 2261, Singapore	(65)-265-6511	(65)-265-6275
Fujitsu Microelectronics, Inc.	3545 North First Street, San Jose, CA 95134-1804, USA	(1-408)-922-9000	(1-408)-432-9044
Fujitsu Microelectronics Ireland Limited	Greenhills Centre, Greenhills Road, Tallaght, Dublin 24, Ireland	(353-1)-520744	(353-1)-520539
Fujitsu Microelectronics Italia S.r.l.	Centro Direzionale, Milanofiori, Strada No. 4-Palazzo A2, 20090 Assago-Milano, Italy	(39-2)-824-6170	(39-2)-824-6189
Fujitsu Microelectronics Limited	Hargrave House, Belmont Road, Maidenhead, Berkshire SL6 6NE, U.K.	(44-628)-76100	(44)-628-781484
Fujitsu Microelectronics (Malaysia) Sdn. Bhd.	Pesiaran Kuala Selangor, Seksyen 26, 40000 Shah Alam, Selangor Darul Ehsan, Malaysia	(60-3)-511-1155	(60-3)-511-1227
Fujitsu Microelectronics Pacific Asia Limited	Rooms 616-617, Tower B, New Mandarin Plaza, 14 Science Museum Road, Tsimshatsui East, Kowloon, Hong Kong	(852-3)-723-0393	(852-3)-721-6555
Fujitsu Mikroelektronik GmbH	Lyoner Strasse 44-48, Arabella Centre 9. OG/A, 6000 Frankfurt/Niederrad 71, F.R. Germany	(49-69)-66320	(49-69)-6632122
Fujitsu Network Switching of America, Inc.	4403 Bland Road, Somerset Park, Raleigh, NC 27609, USA	(1-919)-790-2211	(1-919)-790-8376
Fujitsu New Zealand Ltd.	6th Floor, National Insurance House, 119-123 Featherston Street, Wellington, New Zealand	(64-4)-733-420	(64-4)-733-429
Fujitsu Nordic AB	Torggatan 8, S-171 54 Solna, Sweden	(46-8)-764-7690	(46-8)-28-0345
Fujitsu Philippines, Inc.	2nd Floor, United Life Bldg., Pasay Road, Legaspi Village, Makati, Metro Manila, Philippines	(63-2)-85-49-51	(63-2)-817-7576
Fujitsu (Singapore) Pte. Ltd.	200, Cantonment Road, #11-01 South Point, Singapore 0208, Singapore	(65)-224-0159	(65)-225-5075
Fujitsu Systems Business of America, Inc.	2986 Oakmead Village Court, Santa Clara, CA 95054, USA	(1-408)-988-8012	(1-408)-492-1982
Fujitsu Systems of America, Inc.	12670 High Bluff Drive, San Diego, CA 92130-2013, USA	(1-619)-481-4004	(1-619)-481-4093
Fujitsu (Thailand) Co., Ltd.	60/90 (Nava Nakorn Industrial Estate Zone 3) Moo 19, Phaholyothin Road, Thambon Klongnung, Amphur Klongluang, Pathumthani 12120, Thailand	(66-2)-529-2630	(66-2)-529-2581
Fujitsu Vitória Computadores e Serviços Ltda.	Avenida Nossa, Senhora da Penha, 570, 8°Andar, Paraia do Canto-Vitória-Espirito Santo, Brazil	(55-27)-225-0355	(55-27)-225-0954
Information Switching Technology Pty. Ltd.	Level 32, 200 Queen Street, Melbourne 3000, Australia	(61-3)-670-4755	
Intellistor, Inc.	2402 Clover Basin Drive, Longmont, Colorado 80503, USA	(1-303)-682-6400	(1-303)-682-6401
Tatung-Fujitsu Co., Ltd.	5 Floor Tatung Bldg., 225, Nanking East Road 3rd Section, Taipei, Taiwan	(886-2)-713-5396	(886-2)-717-4644

(The information here as of January 18, 1991)

FUJITSU LIMITED

6-1, Marunouchi 1-chome, Chiyoda-ku, Tokyo 100, Japan

Phone: National (03) 216-3211 International (Int'l Prefix) 81-3-216-3211 Telex: J22833 Cable: "FUJITSULIMITED TOKYO"