Disc File Education

December 1988

W J P Case

ID (Information Development)
MP 095 Ext 5524

IBM Internal Use Only

Contents

Preface

Further reading.

Further reading to assist in comprehension of IBM disc files can be found from page 663 of the 25 year anniversary publication of the Journal of Research and Development, available in the library and from many of the people in Storage Products.

There is a book, published by GPD in 1980, called Disk Storage Technology, (No. GA 26-1665-0) which includes many points of interest to people who do not have a great deal of knowledge about disk technology. Copies of this publication are available through the library.

A book published by SRA called Introduction to IBM Direct Access Storage Devices by Marilyn Bohl contains much information of assistance in understanding many of the terms and procedures used in IBM disc drives.

What is a disc drive

Disc files are information storage devices which utilise a rotatable disc with concentric data tracks containing the information, a head for reading or writing data on the various tracks and an actuator connected by a support arm assembly to the head for moving the head to the desired track and maintaining it over the track centreline during read or write operations. The movement of a head to a desired track is referred to as "track seeking" or accessing, while maintaining the selected head over the cwentreline of the desired track during a read or write operation is referred to as "track following". This is only applicable to closed loop servo systems.

The actuator is typically a "Voice Coil Motor" (VCM) in present day drives. This comprises a coil moveable through a magnetic field of a permanent magnet stator. The application of current to the VCM causes the coil, and thus the attached head, to move rapidly. The accelleration of the coil is proportional to the applied current so that there is no current in the coil if the head is perfectly stationary over a desired track.

The information is stored as bits of digital information recorded as transitions of the magnetisation on the surface of the disc.

Analogies

Capacity

The total capacity of a disc drive might be as much as the storage required for 50 telephone directories.

Flying

A head flying over the surface of a disc is comparable to a 747 jumbo jet flying over a football pitch with less than 1 foot between the jumbo and the pitch. With a stray football left on the pitch the jumbo would crash. This requires that the environment of the enclosure for the heads and discs has to be exceptionally clean. To achieve this cleanliness the Disc Enclosure (DE) has to be sealed from the normal atmosphere, with only an absolute filter between inside and outside of the DE.

Disc File History

The following 8 charts show how the technology of disc drives has changed in the 2 decades before 1980.

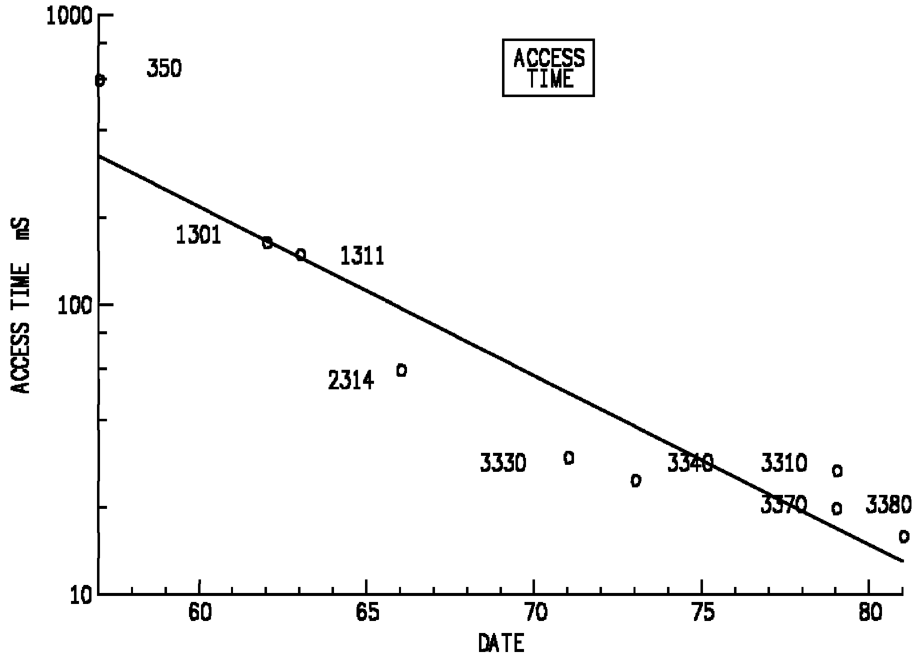


Figure 1. Access time against time

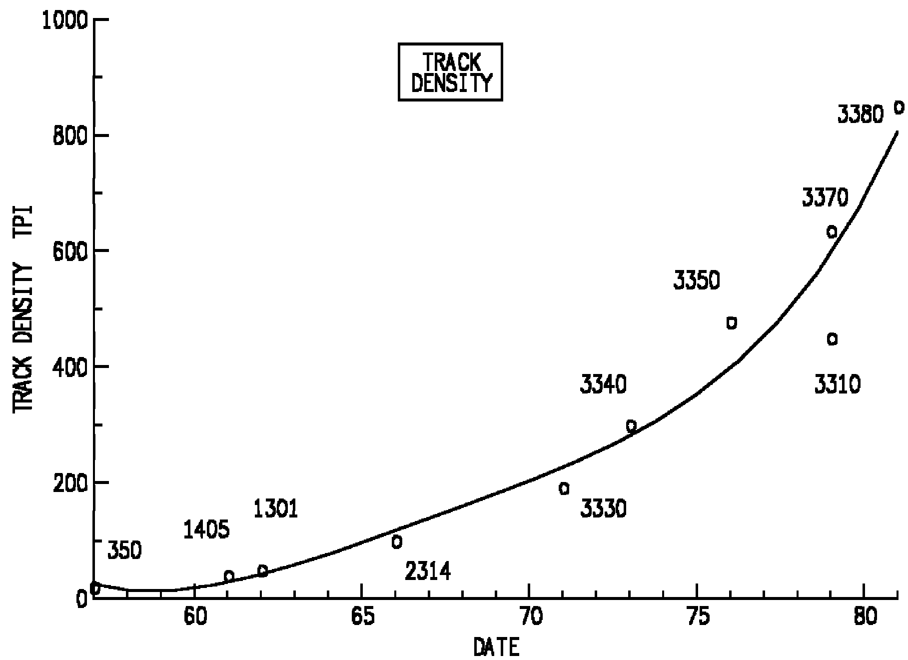


Figure 2. Tracks per inch against time

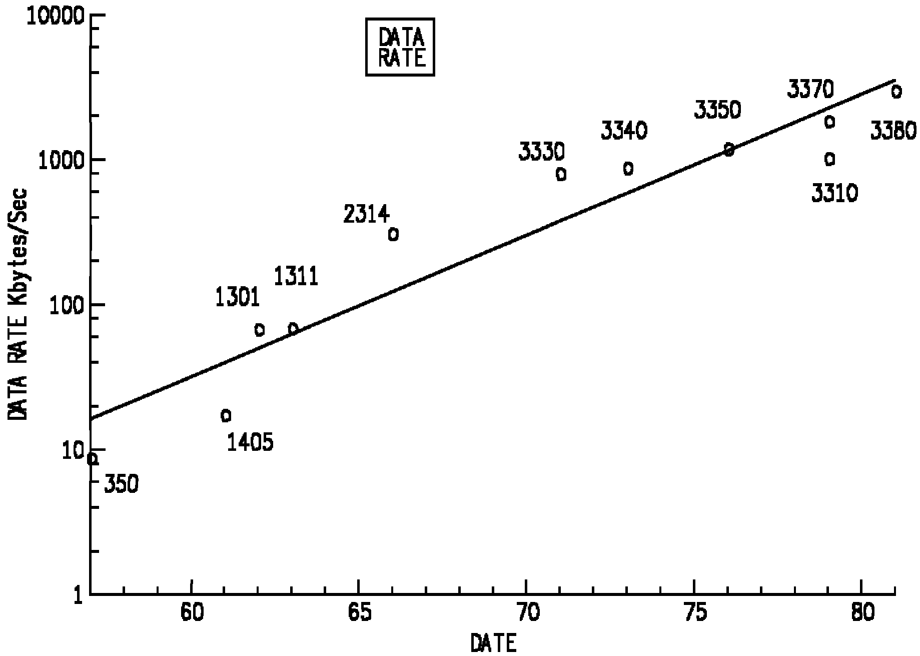


Figure 3. Data rate against time

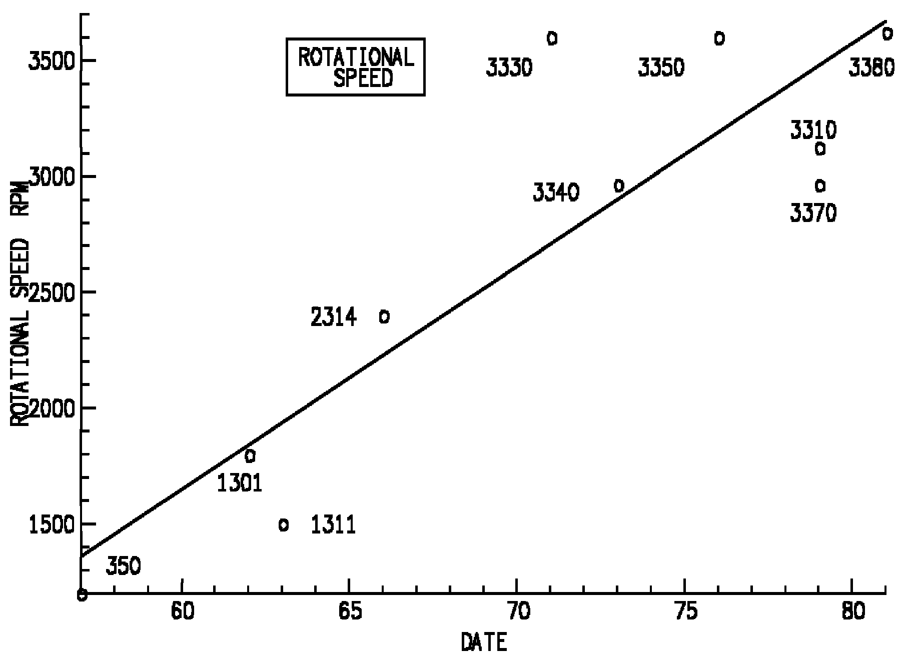


Figure 4. Rotational speed against time

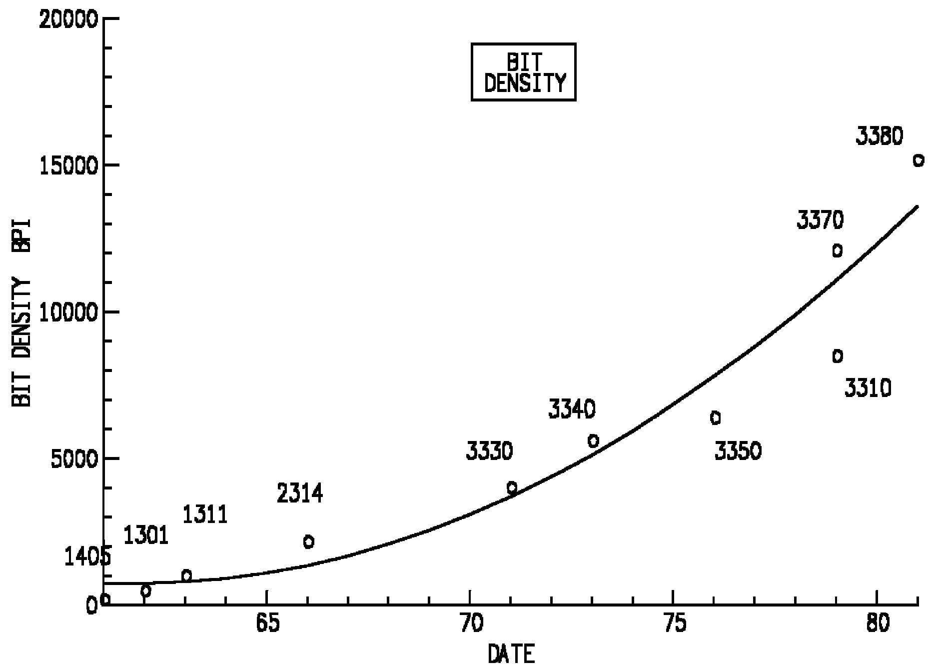


Figure 5. Bit density against time

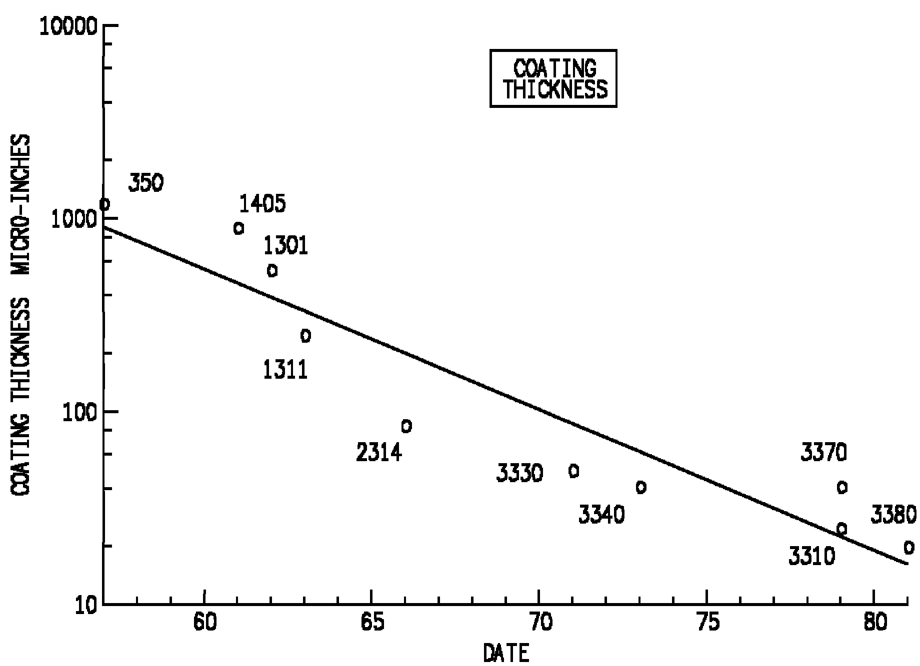


Figure 6. Coating thickness against time

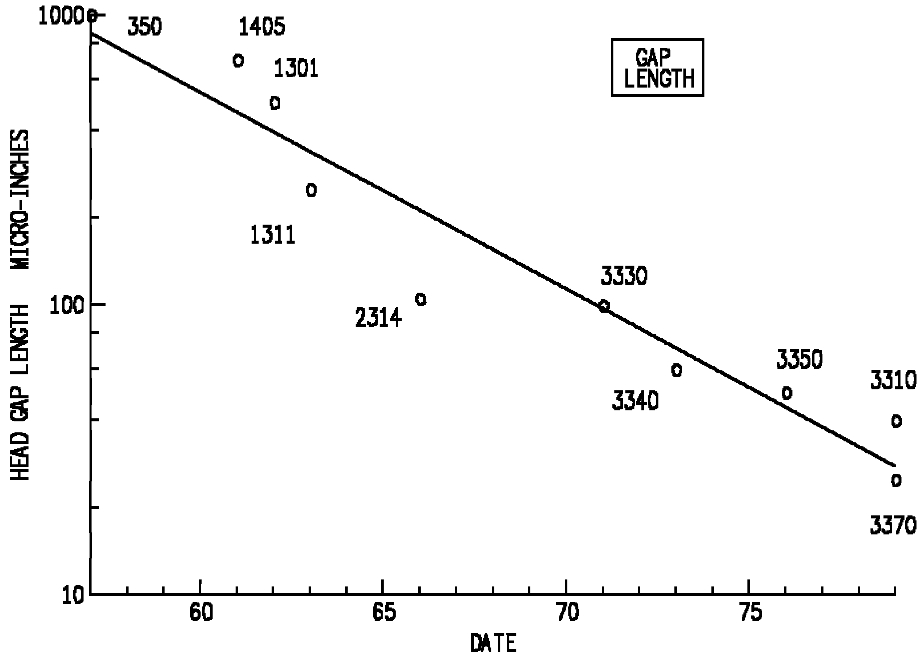


Figure 7. Gap length against time

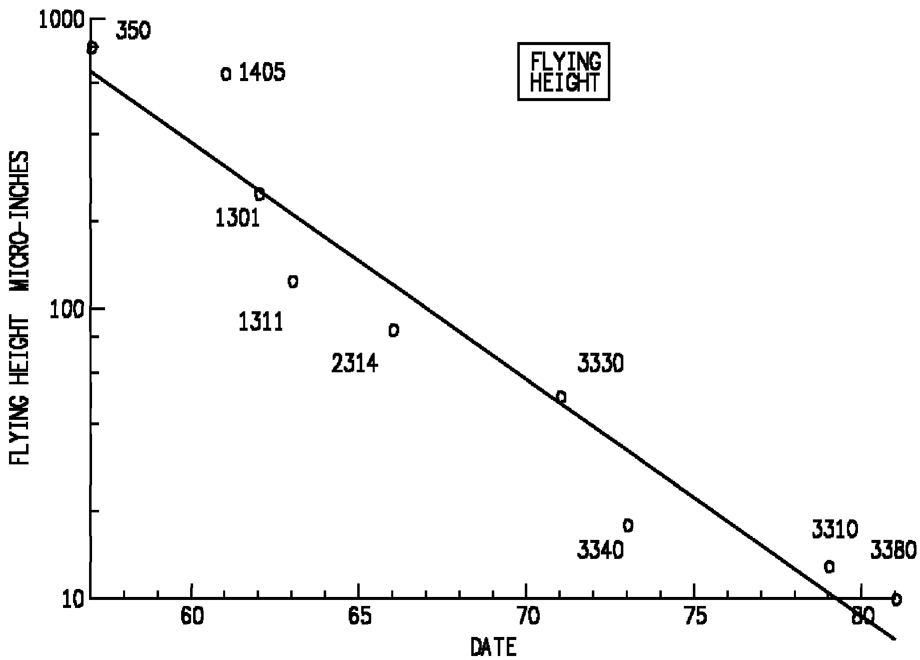
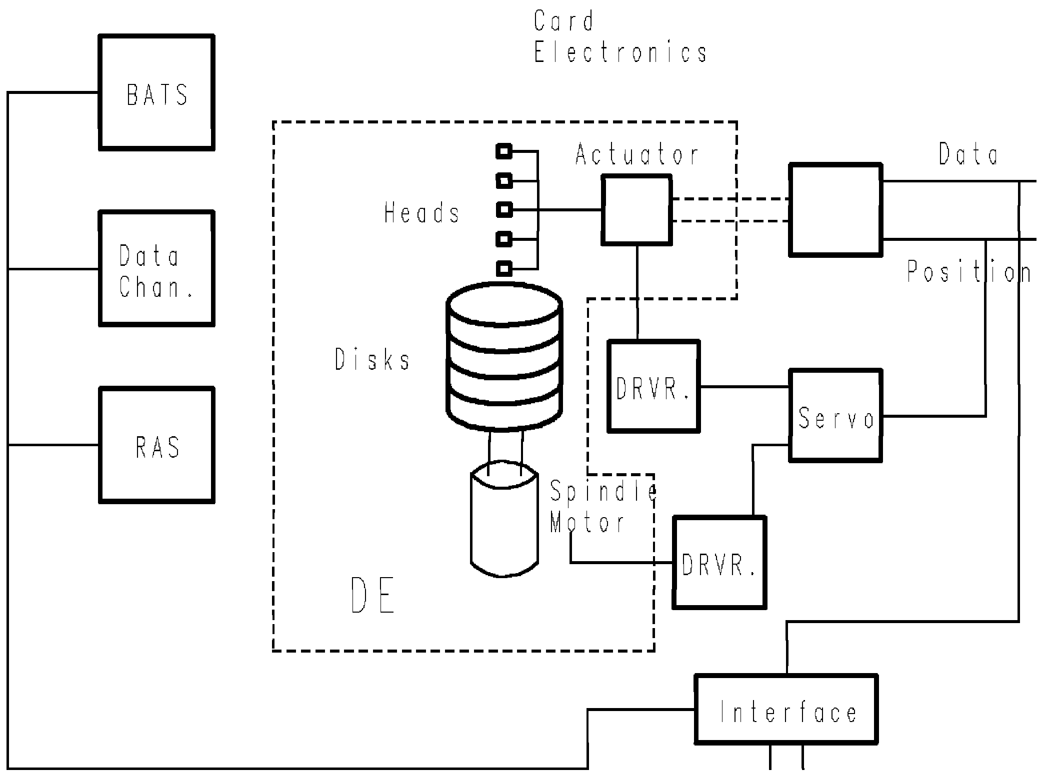


Figure 8. Flying height against time

Disc File Design

A disc file is made up of several sections as listed below -

- A number of discs. The number of discs is to allow for the required storage of the whole drive.
- Flying heads to read the magnetic information on the discs
- A Mechanical Actuator to move the heads across a disc to be able to read the magnetic information written on a selected disc
- A Spindle Motor to rotate the discs
- Electronics to do the following -
  - Write the data required by the customer onto the discs
  - Read the written data which is the customer's saved data
  - Detect the data written by the Servo Writer which will produce an error signal, which defines how far, and in which direction, the head is from the centre line of the written track
  - Drive the actuator under control of a servo loop to achieve the following -
    - Move to a desired track where the required data is stored
    - Follow the achieved track with the minimum of positional error
  - Communicate with the host system, which includes this disc drive, to pass commands and data to and from the drive
  - Provide Service information to assist in maintenance of the disc drive
  - Drive the Spindle Motor, under servo control, to ensure very accurate speed control of the discs



Simplified Diagram of a Disk File

Explanation of acronyms

Acronym	Description
AE	Arm Electronics
ALE	Address Latch Enable
APAR	A Problem Activity Report
BATS	Basic Assurance Tests
CRM	Customer Removable Media/Module
CRU	Customer Replaceable Unit
CSM	Command State Machine
DASD	Direct Access Storage Device
DATS	DASD Advanced Tracking System
DMA	Direct Memory Address
DRP	Data Recovery Procedure
DSP	Digital Signal Processor
DVT	Device Verification Test
ECC	Error Corection Code
EMC	ElectroMagnetic Compatibility (Radio Frequency interference)
ERP	Error Recovery Procedure
ESP	Early Ship Plan
EVT	Engineering Verification Test
FCL	File Control Logic(Glue logic)/Language
FCM	File Control Microprocessor
FIFO	First In First Out
FRU	Field Replceable Unit
FSM	Finite Series Machine
GA	General Availability
HDA	Head/Disc Assembly
ID	Inner Diameter (of the disc)
IOP	Input/Output Processor
IPC	Interface Protocol Circiut
LSM	Link State Machine
LUN	Logical Unit Number
MVS	Multiple Virtual Systems
MVT	Manufacturing Verification Test

MSW Machine Status Word

NRM Non-Repeatable Misregistration
(Distance of the head from the desired track which changes with each revolution.)

PMA Program Memory Address

OD Outer Diameter (of the disc)

POH Power On Hours

POR Plan Of Record

PLD Power Line Disturbance

PRML Partial Response Maximum Likelyhood

(A new Data Channel method for extracting data, designed in Rochester)

RA Repair Action

RAS Reliability Availability and Servicability

RM Repeatable Misregistration

SFC SERDES Formatter Circuit

SCSI Small Computer System Interface

SID Servo ID

The name of the timing mark found on the servo disc following which the servo information is written.

SLM Serial Link Macro

TMR Track MisRegistration
(ie. how far the head is from the centre of the track to be written or read.)

TTPE Track to Track Phase Error

WWM Write to Write Misregistration

Discs

The discs are aluminium discs which are either coated with a magnetic film, which is held in place with a binder, or are plated with a magnetic film by electrolysis or by deposition.

The magnetic surfaces of the discs are used to store the data by reversing the direction of magnetisation. These reversals are the digital bits of the stored data.

Heads

In earlier days a head was made as follows - The air bearing was made from ceramic and had its flying surface ground into a spherical shape. The magnetic head was made from laminations of iron material which was glued into a cut out in the flying surface of the head. A problem with this head was that the magnetic laminations would protude from the ceramics flying surface by amounts which varied with humidity. The problem whith this was that the height of the writing and reading gap varies with humdity and thus the resolution (ie. 1F to 2F ratio) varied also. This is like having the head flyng height varying with humidity.

This head also had 2 lamination sections by each side of the head, parallel to the written tracks. This was used to create a gap between the tracks which was called a tunnel erase. Having created this gap between the tracks, the track misregistration could be much larger than is now acceptable, this is because the head does not see information from the adjacent track until it has passed over the tunnel erased section. The need for this greater TMR was because the earlier machines which used tunnel erase were machines which were mechanically detented at each track, ie. they did not follow the track by reading position information from it. A detail of this head is shown in Figure 9 on page 11.

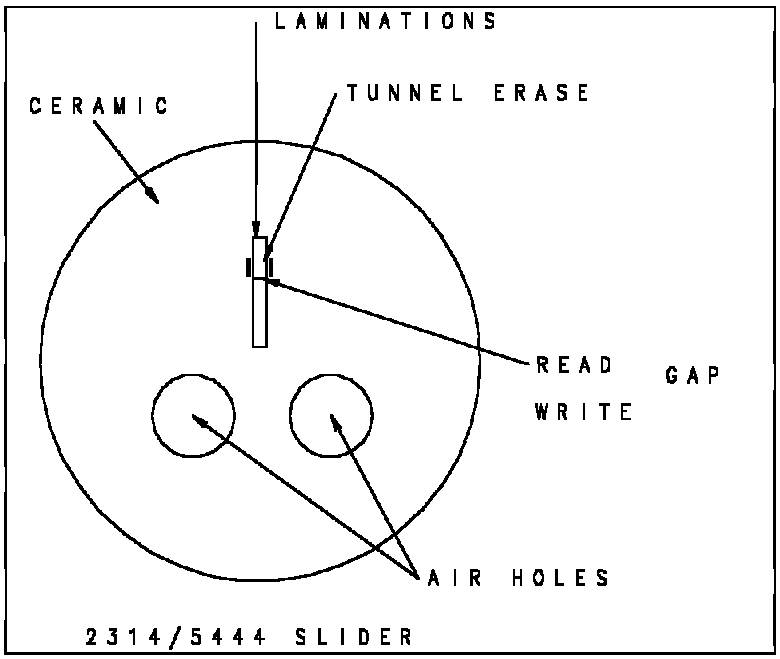


Figure 9. Detail of the magnetic part of a Ceramic flying head.

The ceramic slider head and the first of the ferrite heads is shown in Figure 10 on page 12. This ceramic flying head was used in the 2314 designed and made in San Jose and in 5444 designed in Hursley and made in Havant.

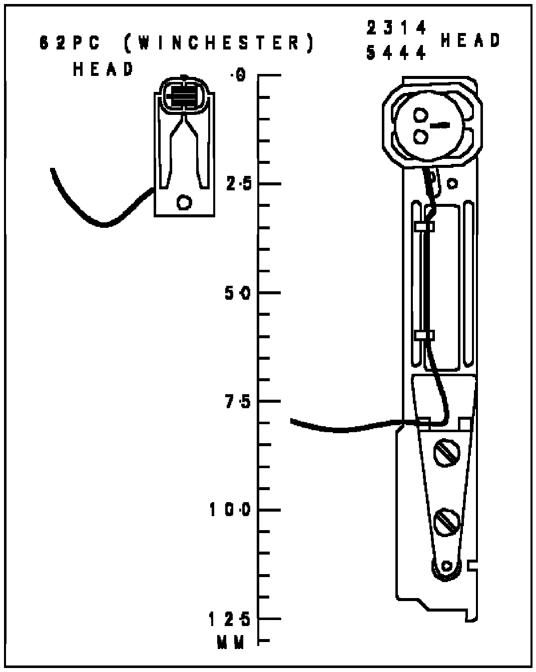


Figure 10. Comparison in size between the Ceramic and the Ferrite heads.

At present there are 2 types of heads which are used to detect the transitions of magnetisation on the disc surface -

- Ferrite Heads
- Thin Film Heads

In the future there will be magneto-restrictive heads.

Ferrite Heads

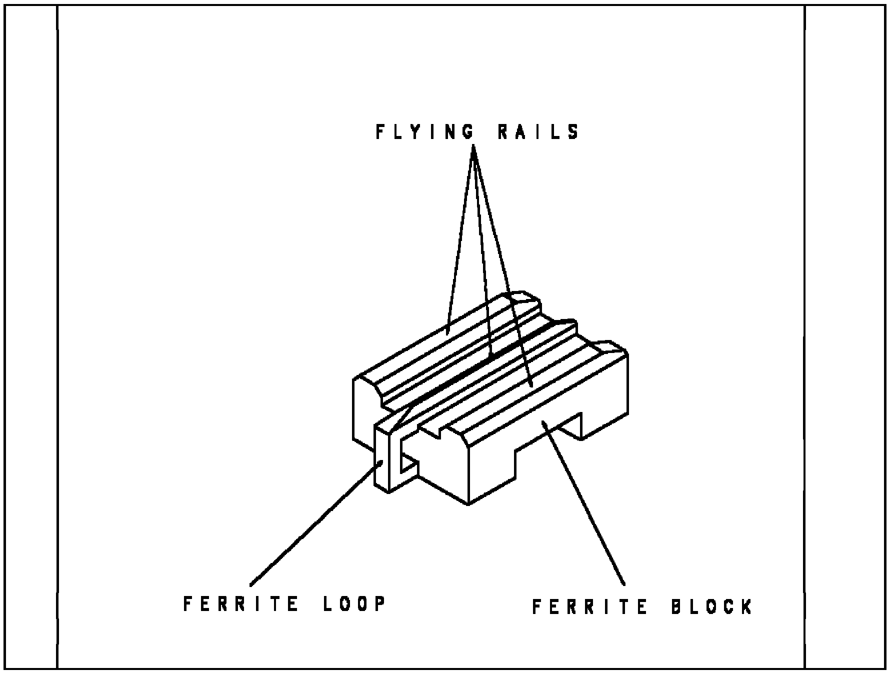


Figure 11. Example of a ferrite head

A Ferrite head is made up of -

- A solid block of ferrite
- 2 or 3 flying rails
- A loop of ferrite at the back end of the major block of ferrite above
- A coil which is wound around the loop

See Figure 11 on page 13.

The head glides over the surface of the disc using the upward pressure from the air slides. The slides are made very narrow to achieve a very small distance between the back of the head and the magnetic surface. The three rail version of these heads are called Winchester heads by the industry. The 2 rail version was first used on the 3330 from San Jose and the 3 rail version used by various products from San Jose and by Hursley on the 62GV (Gulliver) the 62PC (Piccolo) and the 62SW (Swallow) from Hursley. There is a very narrow gap between the loop and the main body of ferrite. As the head glides over the disc it experiences the magnetic field from the disc changing in polarity. The coil wound around the back end of the head detects this change and produces a voltage proportional to the rate of change of this linked flux. Therefore the peak of the voltage picked up by the head is when the head passes over a change in polarity of the magnetisation on the surface of the disc.

This procedure is enlarged below. See "Basic Principles of Magnetic Recording" on page 17 later.

Thin Film Heads

A Thin Film Head is made up as follows -

- A piece of ceramic (the major part of the head) which includes 2 flying rails which are chamfered at one end to cause the head to fly above the surface of the disc.
- Magnetic head(s) deposited on the trailing end of the flying rails.

This type of head is detailed in Figure 12 on page 14.

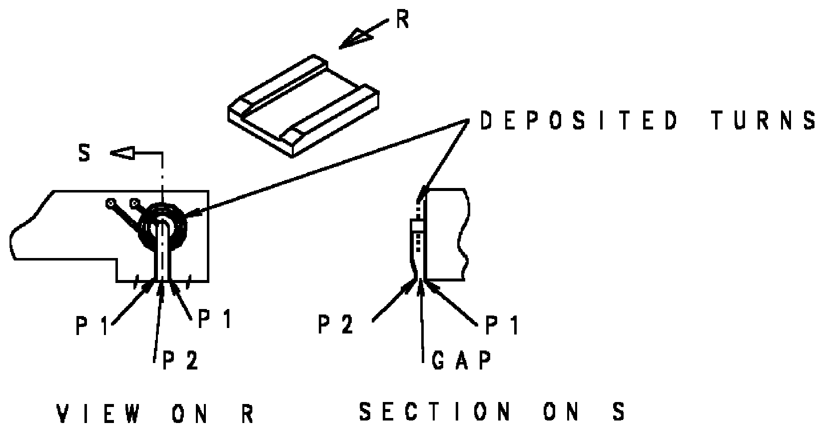


Figure 12. Example of a Thin Film Head

Architecture

There have in the past been two architectures used in disc files.

- Fixed Block Architecture (FBA)

  This has always been used where the data fields are interupted by servo information (Sector Marks), as there is a fixed continuous distance between interuptions. This has often been called a sectorised format.

  .in 0

- Count Key Data (CKD)

  This has always been used where variable record lengths are required.

```
                 SINGLE TRANSITION          DOUBLE TRANSITION (DIBIT)
                       ┌──────                    ┌─────────┐
                       │                          │         │
 Write Current         │                          │         │
                       │                          │         │
                 ______┘                    ______┘         └______

                 >>>>>>|<<<<<<              >>>>>>|<<<<<<<<<|>>>>>>
Magnetisation    >>>>>>|<<<<<<              >>>>>>|<<<<<<<<<|>>>>>>
                 >>>>>>|<<<<<<              >>>>>>|<<<<<<<<<|>>>>>>

                       *

                      * *                         *
 Read-back                                       * *
 Signal              *   *                      *
                    *     *                    *
                  *         *                *       *
               *               *         *
 *********                          *                  *                *
                                                                    *
                                                         *       *
                                                               *
                                                              *
                                                           * *
                                                            *
```

!!!!Join up the dots to get the picture!!!!

Figure 13. Recorded Pulse Responses

Introduction

Magnetic recording involves storing on and retrieving information from a magnetic coating on tape, disk and, in the distant past, a rotating drum. The information is written by switching current in a recording head which is in contact or close proximity to the magnetic surface. This causes the magnetic particles (domains) in the coating to be oriented in one of two directions, depending on the direction of the current in the head, as shown in Figure 13 on page 17.

The information is retrieved normally using the same recording head. As the head passes over the points in the coating where the current was switched (called transitions), the direction of magnetization changes, and the change of magnetic flux induces a small current in the winding of the head, which can be detected, and the data extracted by the recording channel.

Competition in the sales of Disk Files is quite fierce, and this means that designers try to squeeze as much information as possible on each square inch of recording surface. As the 'density' of recording increases, the net effect is that the transitions cannot be read back by the head totally independently of each other. They start to suffer from what is called 'intersymbol interference', where the signal induced in the head by one transition is affected by the adjacent ones. This effect limits the performance of traditional recording channels at high bit densities, but the impact of it can be reduced by various techniques.

Intersymbol Interference

The response of a recording head to both a single and a pair of transitions (dibit) is shown in Figure 13 on page 17. The voltage induced in the head is proportional to the rate of change of flux, and so approximates the differential of the original recording current. As you can see, in the case of the dibit, the shape of the read-back signal for each of the two transitions has been affected by the presence of the other. The 'peaks' have been somewhat flattened by the tails of the pulses. Where the dibit is at its shortest, around the inner radius of the disk, the position of the peaks can also be affected. This phenomenon is called 'non-linear bitshift', or just 'bitshift'. of the signal where the transitions are close together reduces the immunity of the detector to the presence of 'noise' on the signal, and so increases the error rate.

Coding and Detection

Coding

Since the early days of disk file design the encoding techniques have been through many variations. The following is a list of some of the methods -

- Double Frequency Recording
- MFM
- PRML (Partial Response Maximum Likelihood)

Double Frequency Recording

This method has transitions written onto the disc which include regular transitions between every data transition. These extra transitions are used as the clocking information which enables the data to be extracted from the combined stream very simply.

This coding uses a frequency range of 1F to 2F, where 1F is the frequency of the bit rate directly from the disk.

MFM

This is a method to increase the stored data by not including the clock pulses used in the Double Frequency Recording and by insisting that transitions are written even when the data is all zeroes.

Explanation - Consider the track as divided up into cells, there are two relevant places in each cell, the centre and the edge of the cell. The rules for putting down transitions are as follows - The transitions must not be

closer than one cell
further apart than two cells

These restrictions limit the bandwidth of the read back data to a ratio of 2 to 1.
1/2F to 1F.
limits on the position of transitions it is still possible to extract the data and to produce a clock from the readback data. It requires a Phased Locked Loop (PLL) and a fair amount of logic.

Both of the two preceeding encoding techniques require the use of a Peak Detection channel. This is one which will detect the point in time where the analog readback signal passes through a maximum or a minimum.

Partial Response Equalization.: One of the problems remaining after Equalization has been applied to a channel using a peak-detection scheme is that the over-shoots and ringing can cause shifting of the adjacent peaks. The PRML Channel uses a scheme of detection which doesnt rely on the position of the peaks as such, but relies on taking samples of the instantaneous amplitude of the read-back signal at the recorded bit-rate.

In order to eliminate the effect of interference on the sampling scheme, the most obvious technique would be to equalise the read-back signal such that, for a given transition, the samples would be all zero except for the one at t=0. Refer to Figure 14 on page 20. In this case, the sample at 't=nT' for a stream of transitions would be just the value for the 'n'th transition, the values for all adjacent transitions being zero at that time.

This technique suffers from certain disadvantages however, the main one being that the filter required to achieve this boosts the high-frequency response of the channel, resulting in a degradation in Signal-to-Noise performance. A better approach is not to attempt to eliminate the interference altogether, but to equalise in such a way as to make it predictable, and hence able to be compensated for, whilst lowering the bandwidth to retain acceptable S/N performance. This type of equalization is called 'Partial Response'.

Partial Response Class IV: The PRML Channel uses 'Partial Response Class IV' equalization (PRIV for short). The read-back signal is filtered such that the single transition response has the same value at times t=0 and T, and is zero at other sample times. As you can see in Figure 14 on page 20, the resultant pulse is wider, and so requires a lower band-width to handle it.

The effect of PRIV equalization on a string of transitions is shown in Figure 15 on page 21. The read-back signal is just the sum of all the individual responses. It is interesting to note a few properties of the PRIV signal:.

- The samples have only 3 ideal 'normalised' values, +1 0 -1.
- The Channel band-width required is only half the recorded bit rate.
- The maximum continuous recorded frequency to be read back is one quarter of the recorded bit-rate.
- When the write current is switched at the bit rate, the resultant signal is virtually null, the samples being all zero. This appears the same to the channel as a 'DC erase', and is often referred to as an 'AC erase'.

The next problem is how to recover the original data as written. As you can see from Figure 15 on page 21, the samples dont seem to be simply related to the original data. To explain this, we first consider the effect of writing an isolated '1' in a long string of zeroes. Referring to Figure 16 on page 22, the resulting samples comprise a string of zeroes containing the sequence '+1 0 -1'.

Now the relationship can be seen. If you take the Write current (Iw) signal, delay it by '2T', subtract it from the original, then normalise it, you arrive at the read-back samples (S). For example, applying this to the data in Figure 15 on page 21:

```
                                     _______
Write Current                        |
                               ______|


                               >>>>>>|<<<<<<
Magnetisation                  >>>>>>|<<<<<<
                               >>>>>>|<<<<<<

                                     *
                                    * *

 Read-back                         *   *
 Signal
 Equalised
                                 *       *
                      *                             *           *
          ...*.....*.....*.....*...........*.....*.....*.....*.....
                *           *                 *           *

           -4T   -3T   -2T    -T     0     T    2T    3T    4T

                                       * *
                                     *     *
 Read-back                          *       *
 Signal
 P.R.IV Equalised
                                 *             *
                      *                                   *
          ...*.....*.....*.....*.................*.....*.....*.....*.
                *           *                       *           *

           -4T   -3T   -2T    -T     0     T    2T    3T    4T
```

Figure 14. Partial Response Class IV Equalization

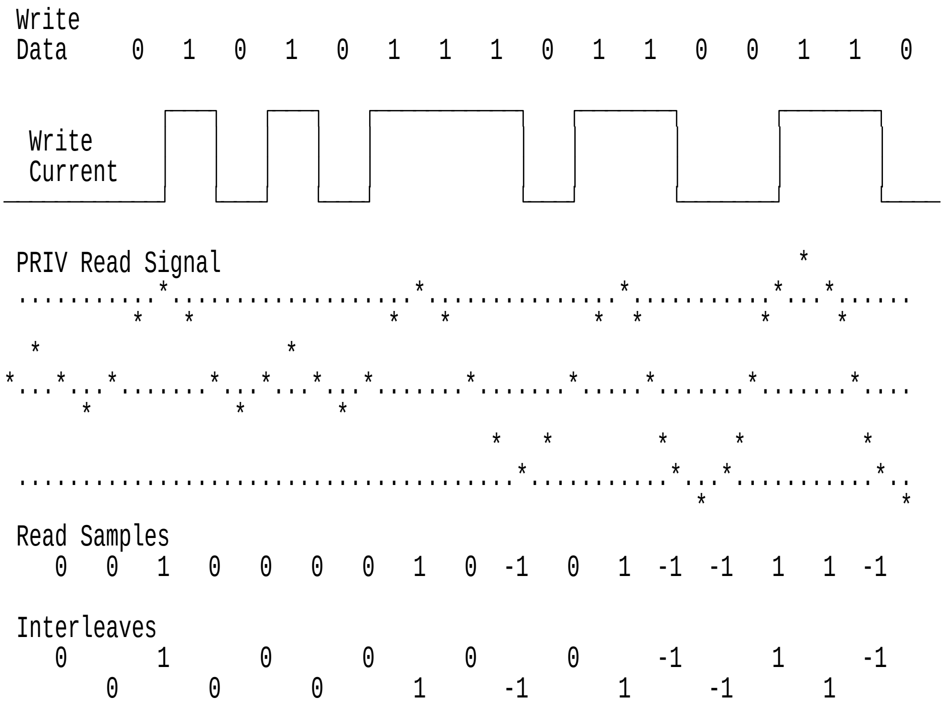


Figure 15. Partial Response Class IV Read-back signal

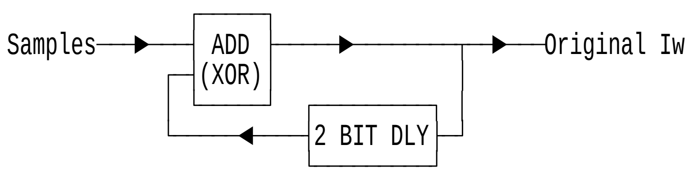
```
   Write Data  0  0  1  0  1  0  1  1  1  0  1  1  0  0  1  1  0
 Minus delayed twice  0  0 -1  0 -1  0 -1 -1 -1  0 -1 -1  0  0 -1 -1  0
              ----------------------------------------------------------
  Samples      ..... 1  0  0  0  0  1  0 -1  0  1 -1 -1  1  1 -1 ......
```

This relationship can be expressed as:

```
S = Iw(1 - D**2)      where 'D' is the 'delay' operator.
```

This expression can be manipulated mathematically, so to recover the original Write current:

```
Iw = S/(1 - D**2)    which is realisable as the following:
```



Pre-coding: In practice, the read-back samples are not perfect, due to the presence of noise in the read-back signal. This noise would propogate around the above 2-bit delay loop probably resulting in unacceptable S/N performance. In the PRML Channel, the compensating function, in this case called 'Precoding', is carried out logically prior to writing, so that it is not affected by the noise. The normalised moduli of the Read-back samples then become identical to the original data before Precoding.

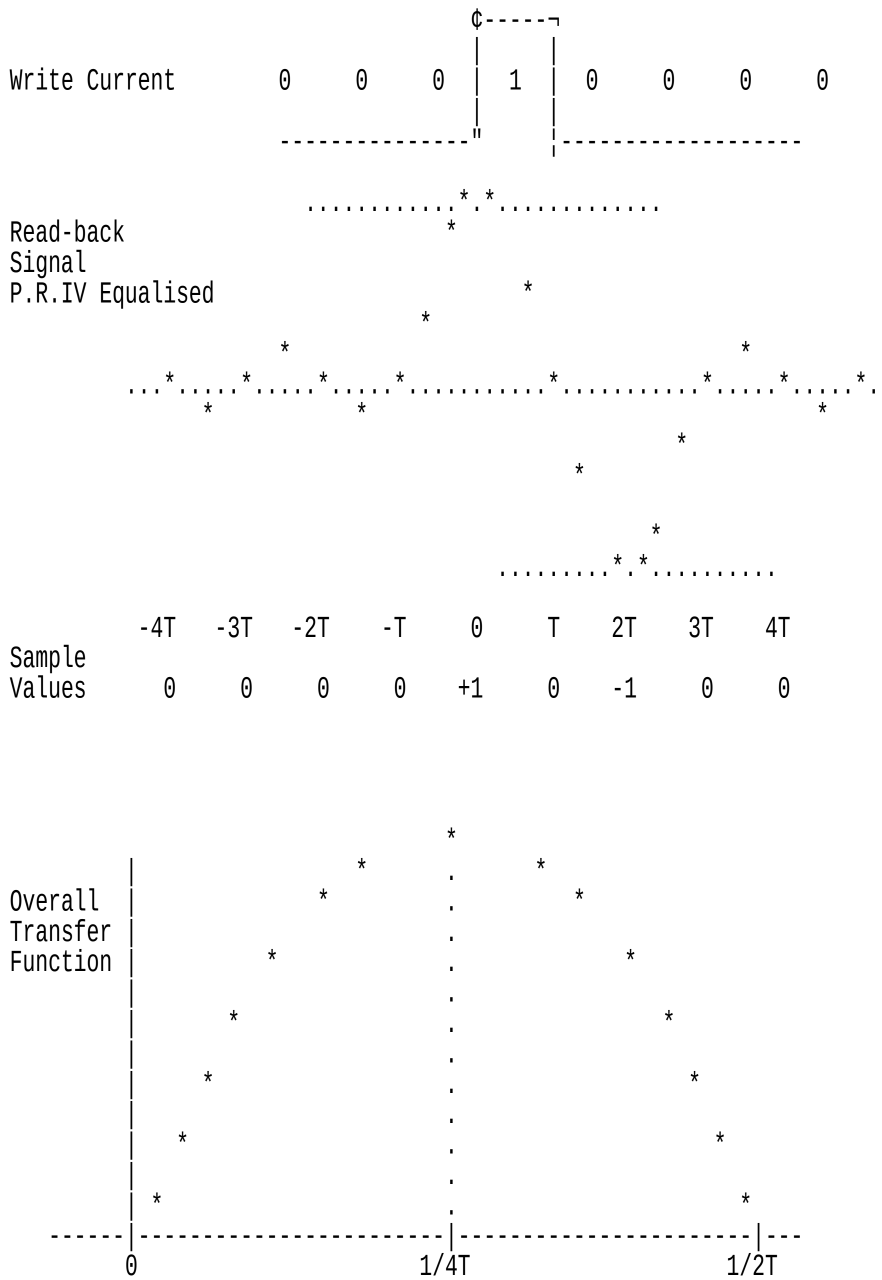


Figure 16. Partial Response Class IV Dibit Response

Maximum Likelihood Sequence Estimation.

Interleaved Samples

The M.L. process makes use of one of the many properties of the sampled PRIV data. By taking all the 'odd' samples as one string and all the 'even' samples as the other, you obtain the two 'interleaves' of the data. Because the PRIV transfer function (1-D**2) involves a 2-bit delay, the precoding and interference operate independantly on each interleave with no interaction between the two.

Referring to the example in Figure 15 on page 21, it can be shown that all the 'ones' on a given interleave will always alternate in polarity. This is called a 'pseudo-ternary constraint', and leads to the possibility that information about the 'real' value of a given sample in the presence of noise is available in the rest of the sequence of samples on the same interleave.

The rules which the samples must follow can be summarised as:

1. For all the samples:
   - There cannot be more than 4 consecutive zeroes (0,4/4 code limit)
2. For each interleave:
   - There cannot be more than 4 consecutive zeroes (0,4/4 code limit)
   - A 'one' must be the opposite polarity to the previous 'one'

Mean-Square Error Criterion

This is a method of finding the one sequence of 'perfect' samples which is 'most likely' to be represented by the string of samples taken by the channel in the presence of noise, bearing in mind that the above rules must apply. The criterion states quite simply that -

> The data sequence which has the minimum Mean-Squared Error (MSE) from the observed sequence is the one 'most likely' to have been used to create it.

To be fully correct, this algorithm depends on the 'noise' in the read-back samples being perfectly additive, uncorrelated and Gaussian. This is not completely true in the case of magnetic recording, so the algorithm is something of an approximation.

If you want to see a reasonably simple proof of this, and a number of the other algorithms that will be described in this section, they are contained in The 'Rosier' Technical Report, referenced in the front of this document.

The Viterbi Algorithm

Considering the hundreds or thousands of samples which would constitute a complete record, its clearly not feasible to implement a process to carry out the millions of MSE calculations to arrive at the 'most-likely' sequence. Originally for use in Telecommunications, A J Viterbi developed an algorithm for arriving at the ML sequence without all the complicated calculations, which can equally well be applied to magnetic recording.

Actuator

There are 2 types of Actuator -

- Linear
- Rotary

Linear Actuators

These are of various designs -

- Rack (of a Rack and Pinion).
- Pneumatic
- Leadscrew
- Taught Band (This has a dc motor to move the actuator via a taught band attached to a hub on the motor and to the linear actuator which travels on linear ways.)
- Direct Drive via a Voice Coil Motor attached to the actuator which travels on linear ways.

Rotary Actuators

There are several designs of Rotary Actuators -

- Direct drive via a cog on the Actuator
- The Actuator has a Voice Coil Motor directly attached

Actuator to achieve rotary motion about the spindle of the Actuator.

Actuator Servo

The servo system following this applies to both the rotary and linear actuators detailed above.

The Servo system is the method which has been used, in recent times, to position the read/write head over a given track and to follow that track with great precision. Previously the actuators had been positioned in an open loop fashion, ie. with no feedback to a controlling system. There are 2 aspects to the servo -

- Access to a demanded track
- Following a track

Access

In earlier days the access was achieved in an open loop fashion, ie. with no feedback about the position of the head with respect to the data on the disc. The following list is of various disk drives with an explanation of their access method.

Open Loop Systems

- Pneumatic actator. (2314) This had a grating attached to the actuator which was used to count track crossing pulses. When the appropriate number had been crossed the Pneumatics was turned off and the data was expected to be under the head.
- Friction Drive Leadscrew. (5444) This had a cotton-reel-like part which was constantly rotated by a drive belt which also turned the discs. The actuator was driven by a leadscrew. Mounted on the end of the leadscrew, which was at right angles to the cotton-reel-like part, was a flexible stainless steel disc. This flexible part was made to contact the cotton reel with 2 clutches. When one clutch was energised it presses one side (diameter) of the flexible disc against one edge of the cotton-reel. This make the leadscrew rotate in one direction eg. clockwise. When the other clutch is energised it presses the other side (diameter) of the flexible disc against the other end of the cotton-reel. This makes the actuator rotate in the opposite direction (anti-clockwise). Either of these clutches make the actuator move in or out on the leadscrew.

  The leadscrew has a cogged wheel mounted on it, which is detected with a magnetic pickup, to count the cogs, which are equivalent to tracks on the magnetic disc. Pawls are then inserted into the cogged wheel to detent the actuator on what it believes to be the required track.
- Stepping Motor on Leadscrew. (5444 mod2) This has a stepping motor mounted directly on the end of a leadscrew. Every position where the stepping motor detents by itself will cause the actuator, mounted on it, to move by 1 track. By having optical detectors on the leadscrew one is able to control the speed of accessing to be much faster than the preceeding friction drive.
- Stepping Motor connected via Taught band. (Other people) A solid wheel is mounted on the end of the stepping motor. The wheel is attached to the actuator by 2 stainless steel bands which are fastened to the wheel in opposite directions. The other ends of the bands are fastened to the actuator in opposite direction to enable the bands to pull the actuator in either direction. The access time of this system is limited by the cogging time of the stepping motor.

Closed Loop System

Direct Drive via a Voice Coil Motor (VCM).

The Voice Coil is attached directly to the actuator which moves on slide ways. These ways can be made of stainless steel or ceramic compounds, providing that the way has very little friction and does not produce wear particles. By having the VCM directly mounted onto the actuator, one is able to have position information, available from the read heads, to feedback in a servo loop to control the position of the actuator. The amount of control is set by the bandwidth of the servo loop. This bandwidth is limited by the mechanics which suffers from resonances.

Servo Systems

Track Following

Only the closed loop system described before is able to perform a closed loop function for following the written track. All of the limitations, described above, restrict the bandwidth of the track following servo. This will be explained below.

Continuous Systems

When Closed Loop control systems were first used, they were made from analogue components, which had a large purchased tolerance. This fact meant that it was very difficult to construct a very tight loop which was required for accurate track following. The analogue components were/are expensive particularly if low tolerance components were/are specified.

To get a low data error rate when reading the data from a track the head must be kept very close to the centre of the track. There are many forces trying to move the head from the correct position -

- Vibration from outside
- Wind from the surface of the disc
- Stiffness of any cable attached to the actuator

To achieve this position, a servo loop with a high gain is required. This high gain loop has to be closed around the actuator and heads which therefore, according to Newton, is a loop which includes two integrators (force to velocity and velocity to position). These two integrators mean that the phase of the system would be 180˚ of lag. This means that the control loop has to be stabilised so that it does not oscillate at all times. This stabilisation is achieved by including two networks in the loop; 1. A phase lead network 2. A phase lag network. These are used to advance the phase of the loop so that it has less than 180˚ of phase lag at the frequency where the loop has unity gain. This frequency is called the bandwidth of the loop. The Bode response of a loop stabilsed as above is shown below.

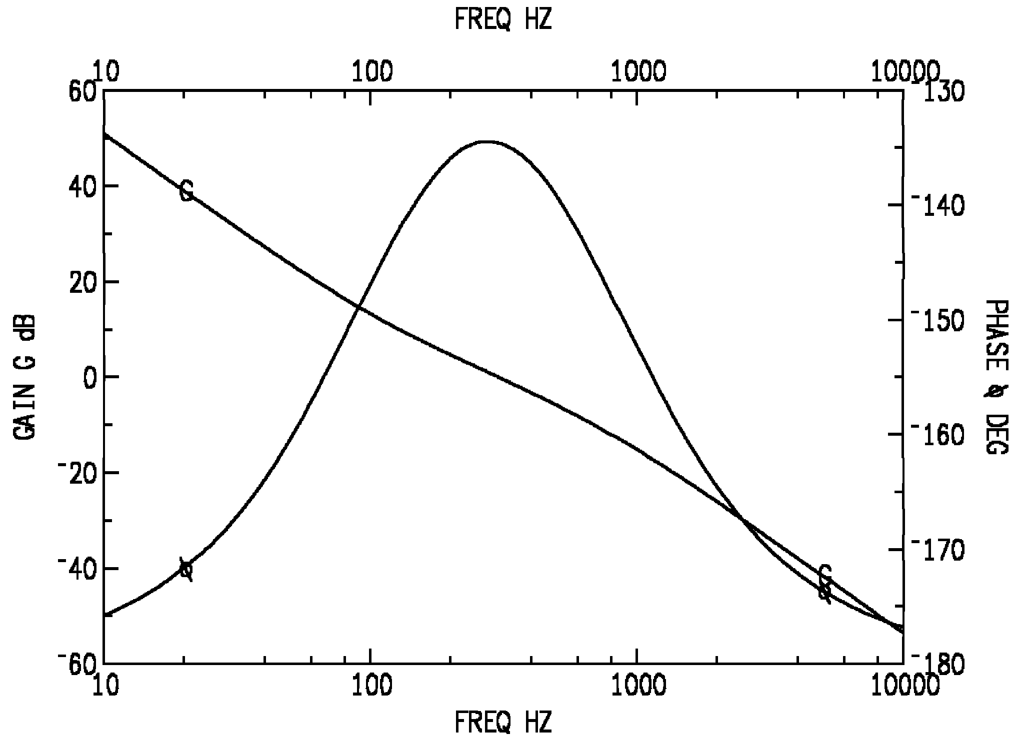


Figure 17. Open loop reponse of a continuous system.

This results in a closed loop response Figure 18 on page 31.

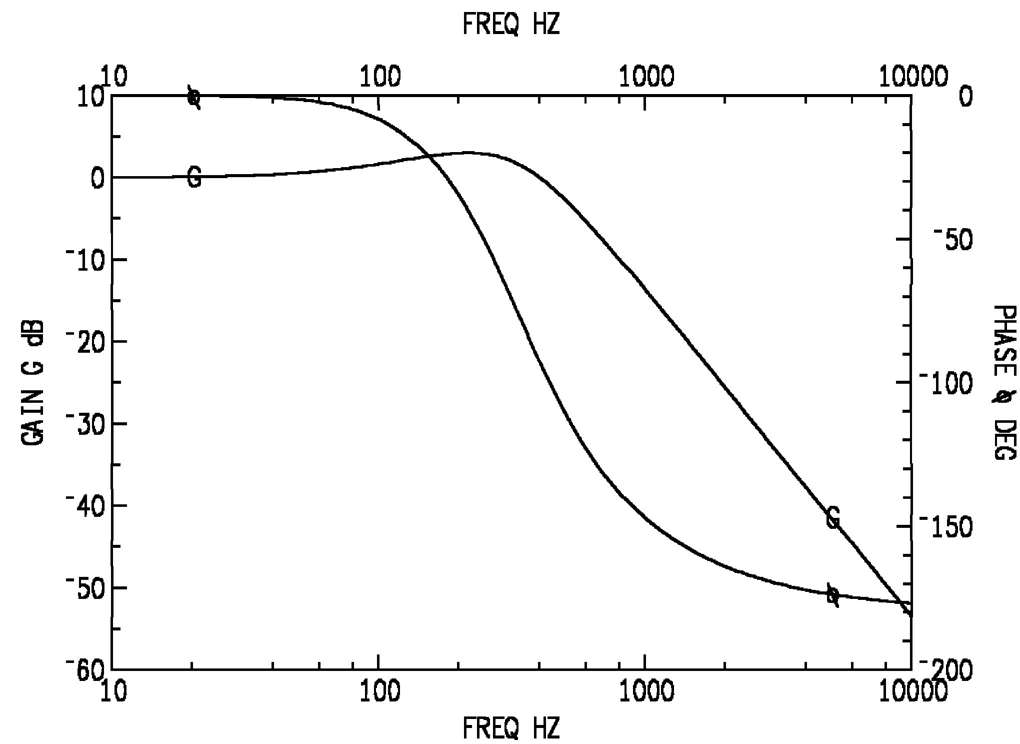


Figure 18. Closed loop response of a continuous system.

This was ok as the closed loop frequency response did not have too much of a gain rise. This amount of gain rise was directly affected by the gain of the open loop system. The gain rise was directly affected by some of the values of the analogue components. To chose these with a very low tolerance was expensive.

The effect of the gain rise is that there are certain frequencies where the following of the centre of the desired track is not as good as is required by the error rate of the write or readback process. This is shown below as a plot of the error function. When this goes above zero db the loop is amplifying the effects which are trying to move the actuator from the centre of the track to be written or read. A plot of the error function follows.

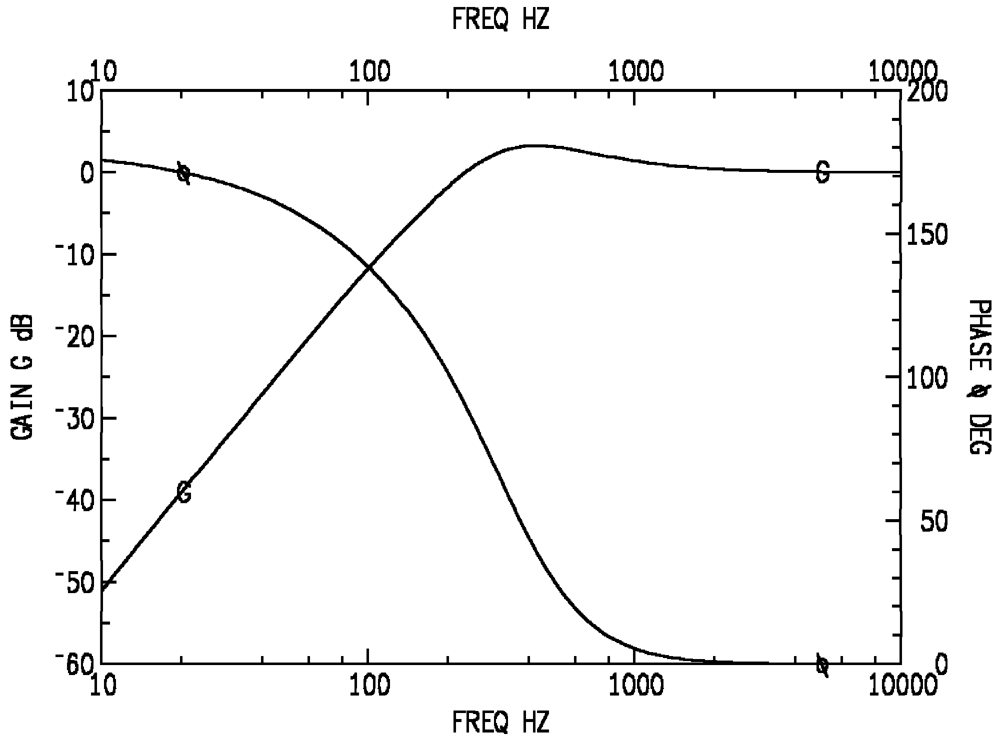


Figure 19. Error frequency response of a continuous system.

Sampled Systems

Before Digital Signal Processors were available a method was used which was a specialised piece of logic hardware which processed the control algorithms under commands from a low performance microproccesor.

This system consisted of only 2 major extra sections -

- The main microproccesor
- The algorithm unit

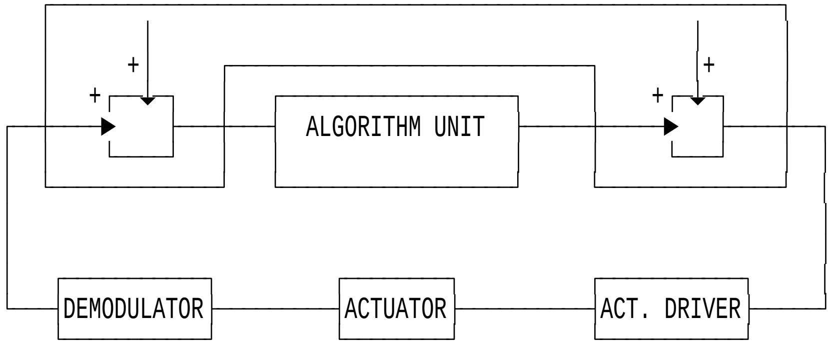


Figure 20. Block diagram of system using an Algorithm Unit.

This system enabled the track following and the accessing to use the same loop. The disadvantage is that the accessing has to be at a defined rate, (usually slower than that which is obtainable with seperate loops). This method was only used for a short period and has now been replaced by the use of DSPs.

The action of the arrangement in Figure 20 on page 33 is as follows -

1. Track following. During this time the output from the demodulator was the best version of the position of the head with respect to the centre of the track. This position information was corrupted by noise in the readback process. If the track was not circular, its shape could be estimated and this estimate could be fed into the left summing junction as a correction to the position error signal (PES). The input to the algorithm unit was therefore a quieter version of the PES, which means that the algorithm unit has to create less output. Whilst track following the extra input to the summing junction on the right is zero.

2. Accessing. During this action both of the summing junctions are used. Whilst accessing you have to know where you should be at each sector time. This position is input to the left junction so that only an error from this hoped for position is what is input into the algorithm unit. To make as access happen an open loop (feedforward) current is applied to the actuator via the right summing junction. This means that the servo loop should have to supply input only when there is an error between the expected positions and the actual positions.

This use of a loop which is common to both track following and accessing means that there is no step change when changing between these modes. As a result of this lack of change the settling onto a desired track is faster. This method is not currently being used on Hursley product, but this last point is being used on the latest products from San Jose.

As of now DSPs are available on the open market. These have great advantage as the loop thus created has far less tolerances included in it. Also the power of the DSP is available to do functions which were not possible with the analogue loop; also one can make changes in the code of the DSP to take some account of late changes in the mechanics.

There are some differences in the control loop when a sampled DSP is used. The 3 preceeding diagrams of Open loop, Closed loop and Error frequency responses are modified as shown in the 3 following diagrams -

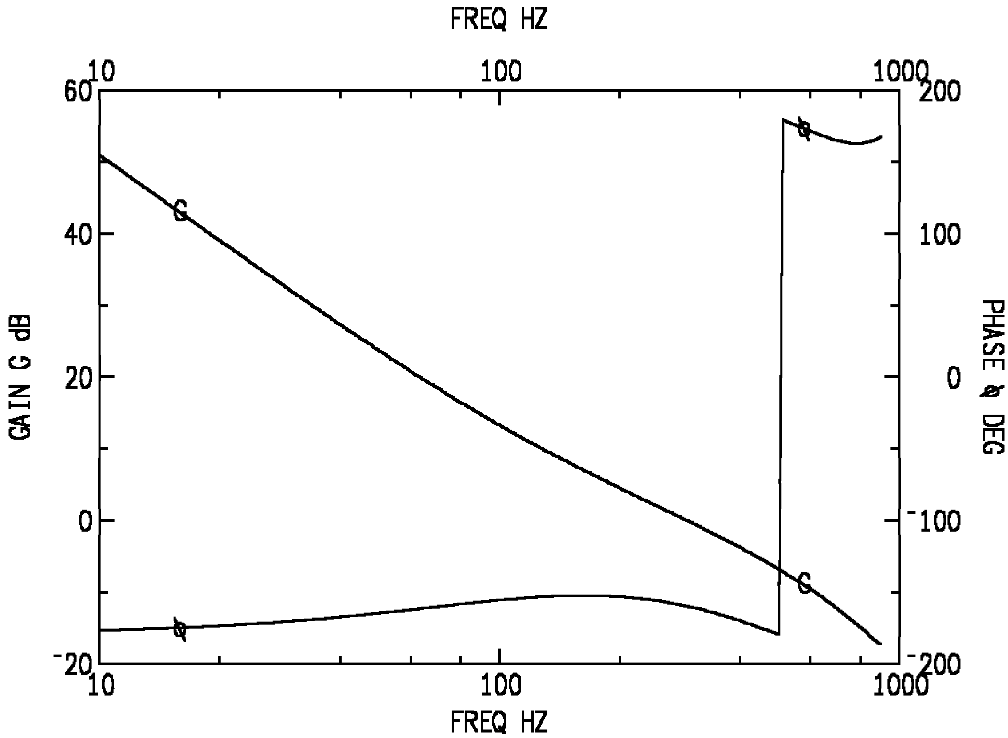


Figure 21. Open loop response of the same system sampled at 2kHz.

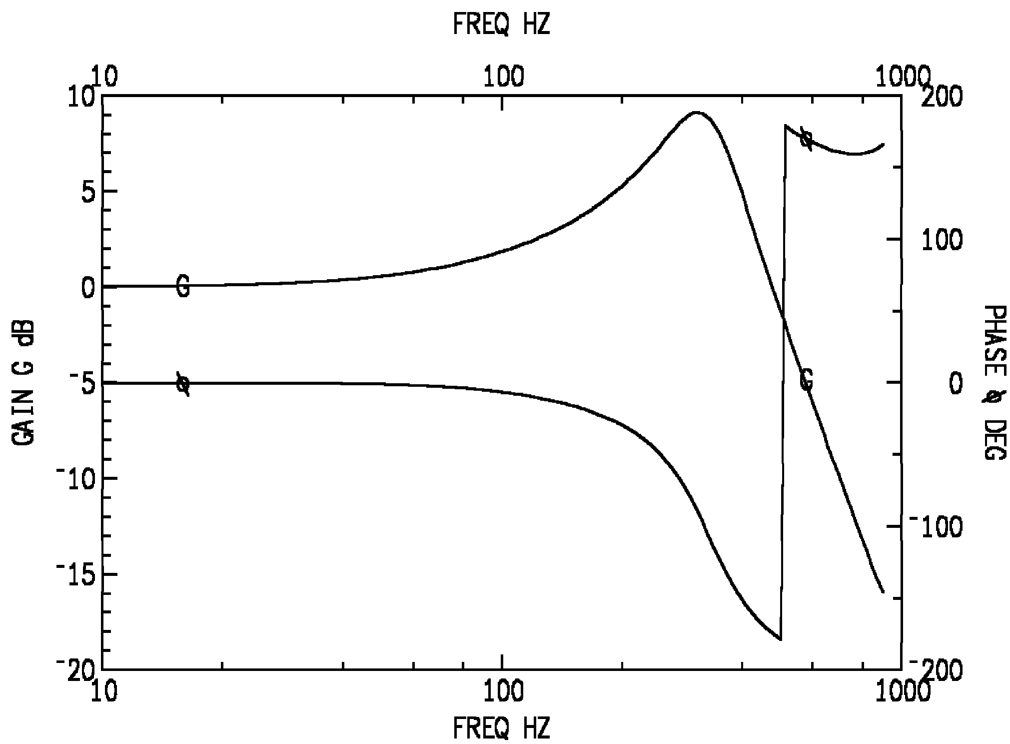


Figure 22. Closed loop response of the same system sampled at 2kHz.

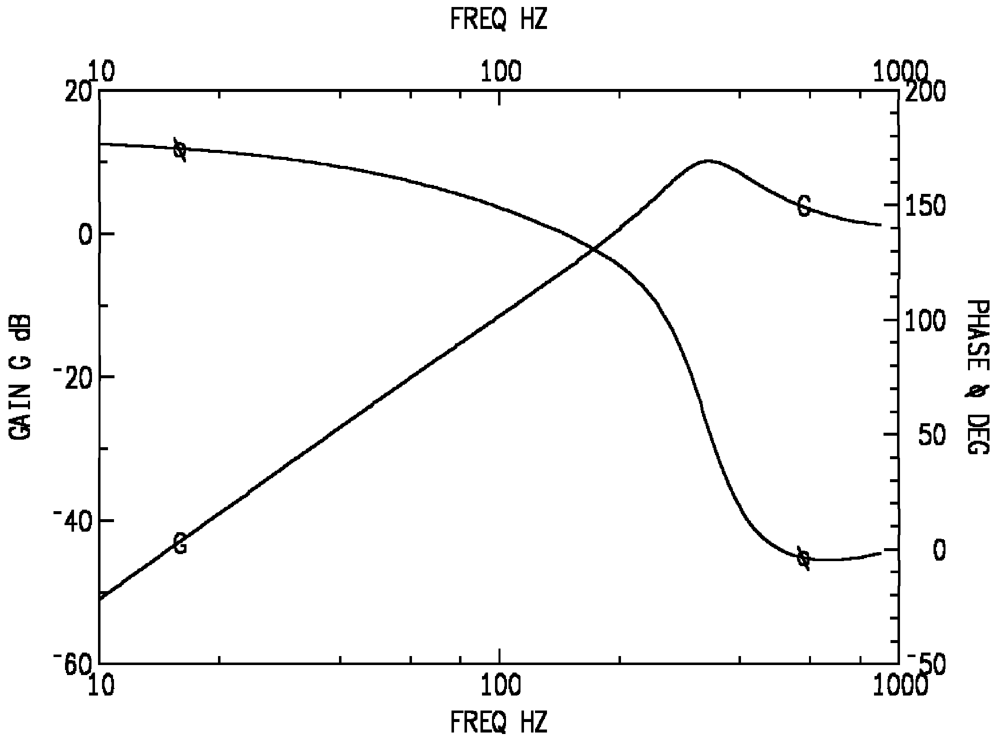


Figure 23. Error response of the same system sampled at 2kHz.

As can be seen from the 3 preceeding plot there is too much gain rise in both of the last two plots. This would give rise to bad following of the track centre line. This is overcome by having more compensation at the centre frequencies.

The compensation of a sampled system in a DSP is called a controller. The controller operates on several of the last inputs to the control loop and several of the last outputs. It multiplies each of these values by a predetermined set of coefficient and by adding all of these multiplied values arrives at a value to be applied to the actuator.

Spindle Motor

The Spindle Motor have been of several types during the last 2 decades -

1. Mains AC induction

   This type was like a small washing machine motor driven with a flat belt.

2. AC induction

   This was a specially designed motor which is made as part of the disk spindle. The motor has a start and a run winding. The run winding is driven from a transistor inverter. At start, the start winding is driven from the mains and the run winding is driven at mains frequency. When the motor has had time to achieve nearly synchronous speed, the start winding is switched off and the run winding frequency is slowly ramped up from the mains frequency to full speed running frequency. The motor is stopped rapidly by passing DC through the run winding.

3. Brushless DC

   Most people drive these motors by having Hall effect devices sensing the rotational position of the spindle and driving the windings through transistors. Speed control is acheived by using the information from the Hall effect devices and then controlling the drive to the transistors.

Spindle Motor Servo

This is only used with the last motor type. Normally it is only a speed control loop but it could be a rotational position loop.

In the speed control loop the output from the hall effects can be used as an input to a phase locked loop (PLL). The output from this loop is used to modulate the transistors which drive the motor. The accuracy of speed control is dependent on the gain that can be achieved in the loop, which include the PLL and the motor. The Servo system used to control the motor can be the same as that used to contol the position of the actuator with the difference that there is only one integrator in the loop to be controlled.

Attachment

Many types of attachment between the disk drives and the using system are or have been used. Some of these are listed following on from here.

1. Native Attachment.

   This means that a disk drive is directly connected to the using system without any standard interface. This was the method used by every file developed in Hursley up to and including 62SW.

2. An IBM standard interface eg. DFCI

   This was the method used on 9335 (Kestrel).

3. An industry standard interface eg. SCSI

1. The fist of these was a special interface which only was able to transmit the commands required by the drive to enable it to access, read and write, and send diagnostic information back in response to a specific request from the using system.
2. The second was an IBM version of an industry standard (DFCI is an IBM version of IPI3).

Some of the commands on the interface are like -

- Test that attached drive is ready for operation
- Access to a desired track
- Select a particular head
- Read or write a give number of bytes from/to the accessed track using the selected head
- Request sense information from the drive
- Format the drive. On the native attached drives this could have been by using the previous 3 or 4 command types. On the industry standard interface this is performed as a seperate command.

On Industry standard interfaces there might also be commands as follows -

- Inquire about the capacity of the selected drive
- Start or stop the drive
- Inquire about the status of a drive
- Write data to the drive and verify that the written data can be recovered correctly
- Reassign block (sectors) of data. This might be done if the using system had decided that a particular block (sector) was liable to be unable to store data reliably.
- Download microcode to the drive to enable the drive to perform other operations
- Change the parameters of, for example, the ERP procedures
- Inquire about the parameters of, for example, the ERP procedures