Magnetic Disk Drive Technology Short Course

March 22-23, 1999

Presented by:

IIST

Santa Clara University
School of Engineering
Santa Clara, CA 95053

Phone: (408) 554-6853 * Fax: (408) 554-7841 *
iist@iist.scu.edu * http://www.iist.scu.edu

AfterWords

A century of magnetic recording

Just as the invention of the printing press in the 15th century forever changed how we communicate, the development of digital magnetic recording in the 20th century has profoundly affected how we record, store, and disseminate information. Financial services, airline reservations, and perhaps most significantly, the Internet are supported by huge on-line real-time databases.

The year 1998 marked the 100th anniversary of the magnetic storage industry. And while its birth took place in Denmark, there is no better place to celebrate the impact this industry has had than in the heartland of computer data storage – Silicon Valley.

Leading the celebration at Santa Clara University was SCU's Institute for Information Storage Technology (IIST) and the Center for Science, Technology, and Society (CSTS). Together they hosted a December conference on Magnetic Recording and Information Storage: Technological Milestones and Future Outlook. IBM, Fujitsu Computer Products of America, Adaptec, and IDEMA co-sponsored the conference, which more than 130 information technology professionals attended.

Kicking off the day's activities, CSTS Director Jim Koch put the recording Industry's accomplishments into perspective by noting that in 1855 it cost five cents to send one word from Philadelphia to St. Louis by telegraph. Today it costs four-and-a-half cents to store one-and-a-half million bytes of information that can be sent anywhere in the world with a click of a mouse – virtually for free.

Presenters from the industry assessed the impact of digital magnetic recording and direct-access storage on information processing applications, including the Internet, the disk-drive industry, and future technology.

Keynote luncheon speaker Al Shugart, founder of Seagate Technology, recounted the early days of the industry, which were often more low than high tech. He recalled that on the first disk drive, the RAMAC, the disks were coated by pouring iron oxide paint from a Dixie cup onto a spinning platter.

Al Hoagland, IIST director, reviewed the paradigm shift in magnetic recording to direct access data storage brought about by the advent of the electronic digital computer in the 1940s, when a critical need developed for rapid access to digital data. Until then, the focus was on analog sound recording. While early computers relied on punched cards and paper tape for data storage, magnetic recording quickly became recognized as the best technology to meet the storage needs of computers. And as the personal computer emerged commonplace in both homes and offices around the world, the need for memory and storage soon became insatiable.

Magnetic recording is effectively replacing paper for recorded data and e-mail. Today, the technology is advancing at its most rapid rate ever, making even more data-intensive applications possible, such as digitally recorded images replacing photographic film.

Innovations in every aspect of magnetic disk drives have driven up storage density at a phenomenal rate. Since 1991 areal density has advanced at a compound annual growth rate of 60 percent per year. And industry experts believe that by taking different approaches to scaling magnetic recording, hard disk-drive technology should be extendible by another factor of 100, ensuring the industry's dominance well into the new millennium.

TIMELINE

1898	Danish engineer Valdemar Poulsen invents the Telegraphone, the first telephone answering machines employing an electromagnet moving along a length of piano wire.
1920	Between 1920 and 1929, various inventors create steel tape and wire recording devices such as dictation machines and radio studio recorders.
1928	Austrian inventor Fritz Pfleumer creates first magnetic tape by gluing pulverized iron particles to a strip of paper.
1933	The Magnetophone, using cellulose magnetic tape, invented in Germany.
1947	Singer Bing Crosby contracts with Ampex Corp. to market its broadcast-quality audio tape recorder.
1948	UC-Berkeley Computer Project creates first magnetic drum for storing binary computer data with capacity of 800 bits/in2.
1951	UNIVAC ships first computer using magnetic tape storage system.
1957	IBM releases RAMAC, first commercial disk drive, storing 5 megabytes of data and featuring a pressurized air-bearing head.
1961	IBM Disk Drive launched with 50 megabyte capacity – prototype of future generations of disk drives, having one head per surface and using flying heads.
1973	Floppy disk introduced.
1980	Seagate Technology launches first 5.25-inch hard drive for desktop computers.
1983	3.5-inch hard drive introduced.
1991	Hard drive shrinks to 1.8 inches.
1998	Disk drives the size of a quarter are made available that capture 320 megabytes of data. Areal density is projected to continue to increase at a 60 percent compound annual growth rate.

Magnetic Disk Drive Technology

A. S. Hoagland
IIST
Santa Clara University

Course outline

– Day one

- Disk storage devices
  - Historical perspective
  - Disk drive components
  - Disk performance factors
- Magnetic materials and fields
- Magnetic recording process
  - Fundamentals of magnetic recording
    - Writing, transitions, "a" parameter, NLTS
    - readback, signal response, ISI
  - Pulse and frequency/wavelength characterization
- Magnetic heads and media
  - Write heads
  - MR and GMR read heads
  - Thin film media

– Day two

- Magnetic recording channel
  - SNR, channel codes, equalization, detection techniques (PRML, etc.)
    - Error rates
- Storage density
  - Head positioning and tracking
  - Track density
    - Off track, TMR, 747 curve
  - Magnetic data recording and ECC
    - RAID systems
- Status and Future trends
  - Disk drive technology
  - Alternative technologies
  - Applications/industry changes

Magnetic Disk Drive Technology

Disk Storage Devices

Introduction

- First section
  - Disk drive historical evolution
  - General configuration aspects of disk drives
  - Storage performance measures and relation to linear and track densities
  - General trends

Historical evolution

- 1946-1953
  - Setting the stage
  - Available technology
- 1953-1961
  - Sorting out design approaches for on-line storage, **Ramac**
- 1962-1979
  - IBM defines the market
    - Era of reverse engineering
    - Alternative approaches fall by wayside
- 1980-1991
  - Introduction of PC's
    - Emergence of low end drives
- 1991
  - Driven by technology advances

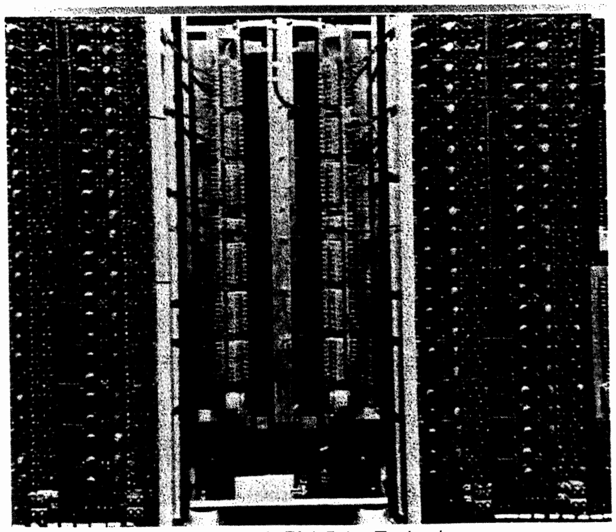
The Cause - The Arrival of The Electronic Digital Computer

- **The Eniac Computer - developed in 1940's**
  - **Memory --> vacuum tube**
  - **Input/Output media --> paper tape, punched cards**
- **Memory**
  - **High speed, high availability, relatively small capacity**
- **Input/Output media**
  - **Slow data transfer, low cost and open ended capacity**

Magnetic Drum Memory Features

- Write, read and update in-place a small block of binary encoded (NRZ) data
- Continuous availability and reliability of stored data
- Short access time (milliseconds) to any record
  - High head/medium relative velocity
    - Non-contact head/medium separation
    - High data rate
- Non-volatile

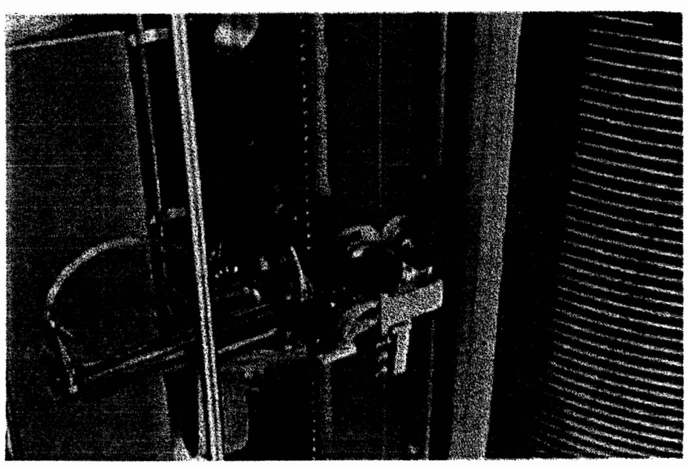
UC Magnetic Drum Memory



Ramac

- Announced 1956, shipped 1957: Capacity 5MB
- 50 - 24 inch disks
- bpi 100, tpi 20: storage density (max) 2000 b/in^2
- Features
  - Pressurized air bearing
  - R,Z detent positioning of heads
  - particulate magnetic oxide spin coating
  - "Delta" head with wide-erase
  - self clocking with fixed data rate
- Application
  - Integrated accounting system

305 RAMAC Disk Drive



RAMAC

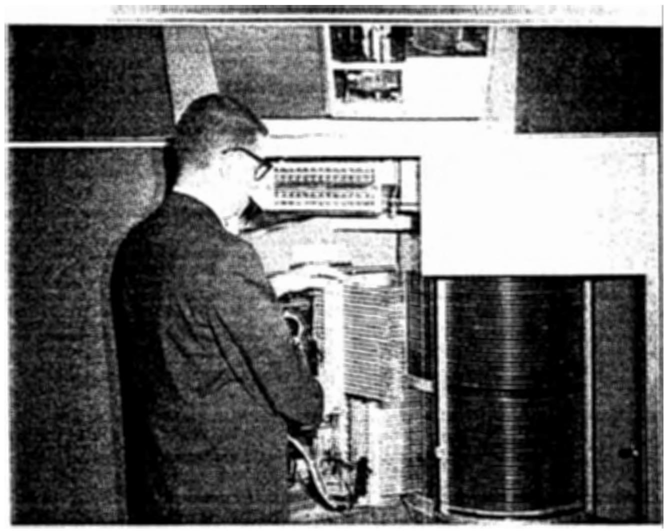
The IBM 350 RAMAC Disk file

IBM 1301 and 1311

- 1301 announced 1961
- First commercial disk drive to use "flying" heads and one head per recording surface
- bpi 500, tpi 50
- Still used open loop head positioning
- (Al Shugart was engineering manager)

  - 1311 announced in 1962
  - First removable disk pack drive
  - 14 inch disks
  - 2MB/disk pack

Al Shugart, Chief Engineer



ASH Magnetic Disk Drive Technology IIST

First Drive with Flying Heads & Servo Track Positioning

ASH Magnetic Disk Drive Technology IIST

IBM 3340/3350

- 3340 shipped in 1973
- Features
  - contact start-stop
    - low mass head
    - lubed disk
  - 2 heads/surface
- Capacity 1.7 Mbits/in^2
- Replaceable HDA module (incl.. heads)
- 3350 followed soon after with fixed HDA
  - volume leader
  - Source of term “Winchester” Technology

Early 1980’s

- The magnetic “bubble” bursts
- Personal computers come on the horizon
  - Low end disk drives emerge, initially based on Winchester technology
- Disk diameter goes to 5 1/4 size
  - Lack of existing products leads to intense competition with many suppliers
- 1982-1984
  - Industrial consortium leads to University centers in data storage
  - Magnetic recording finally recognized as critical technology

Trends in storage density

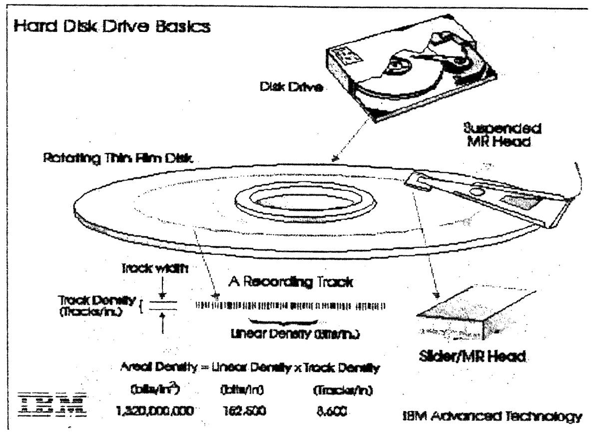
- From 1956 to 1990 density increased by a CGR of approximately 32%
- From 1991 density has been increasing at a 60% CGR
- Change driven by:
  - technology advances and market demand
  - consolidation to few major industry players
  - semiconductor advances

Disk diameters

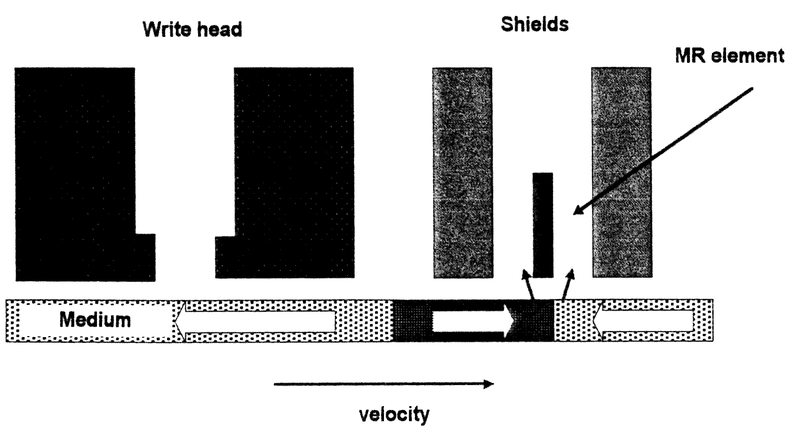
Disk Diameter	Relative Capacity
8	64
5.25	27.56
3.5	12.25
2.5	6.25
1.8	3.24
1.3	1.69

If storage density doubled then for same capacity could move to smaller diameter by factor of 1/sqrt(2)

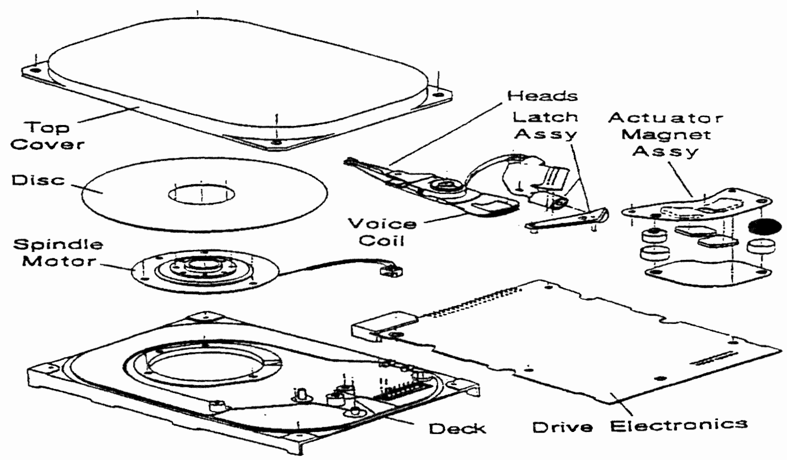
Disk drive basics



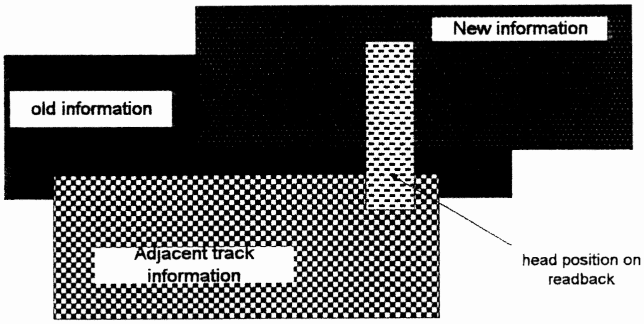
Magnetic head structure



Disk drive configuration



Direct overwrite



The requirement that individual data blocks can be overwritten with new data leads to signal distortion and noise in the desired signal from the concurrent sensing of the previously written data and adjacent track information, arising from tracking tolerances.

Magnetic vs physical separation

- Magnetic spacing - distance from magnetic head pole faces to midpoint of magnetic medium
- Physical separation - distance from slider low point to lubricant layer
    - ***Differences arise from***
- Disk texture
    - lower stiction by reducing slider-surface contact area
- Disk overcoat
    - protect magnetic layer mechanically, corrosion resistance
- Lubricant
    - reduce wear, static friction
- Pole tip recession from slider bearing surface

Data rate influence

$$\text{Capacity/track} = \text{bpi} \otimes (\pi \text{D}) = \text{R}\left(\frac{60}{\text{RPM}}\right)$$

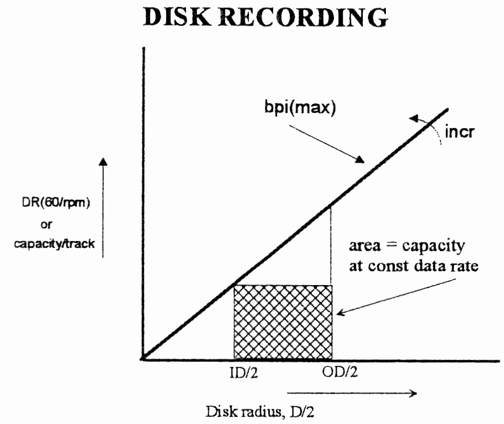
Where Data rate =R

$$\text{The total disk capacity} = (\text{capacity/track})(\text{tpi})\left(\frac{\text{OD - ID}}{2}\right)$$

For a constant data rate

$$\text{The capacity/track} = \text{bpi}_{\text{max}} \otimes (\pi \otimes ID)$$

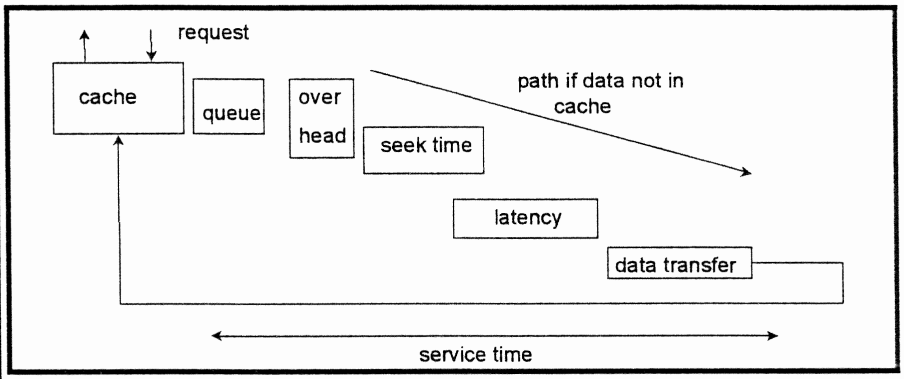
Magnetic disk recording



Data rate

- Data rate varies directly with linear density, track radius and with rpm.
- Latency varies inversely with rpm.
- To accommodate recording codes the channel rate is greater than the user data rate and can be significantly greater.
- More sophisticated ECC requires higher data rates to maintain the same throughput.
- With a 60% CGR in storage density the linear density will increase at least at a 20% CGR. Increases in rpm and with the same or larger disk diameters data rate needs in increase at approximately a 40% CGR

Access time of disk drive



- hit ratio
- read ahead
- write back vs write through

Density limit concerns

- Nonvolatile characteristic of medium
  - Thermal stability of medium magnetization
    - particle volume
      - thermal energy barrier and demagnetizing fields
    - limits on increasing H_c from attainable saturation magnetization for write head
    - Patterned bit cells
- Increasing data rate
  - Switching time of write head field
    - laminated head structures
      - suppress eddy current effects
  - Recording is pattern dependent
    - time constant of write current and head field responses

Thermal instability

$$p_r = f_o e^{-(E_B/KT)}$$

p_r is the probability of a grain switching (crossing the energy barrier), f_o is approx. 10^9, E_B is the height of the energy barrier and KT gives the thermal energy available.

In the absence of demagnetizing fields (which lower the energy barrier) $E_B = k_u v$, where k_u is the anisotropy energy density of the particle and v is the particle volume.

The time constant of random switching of bits due to thermal effects is $1/p_r$.

Drive failure rate

- Assume constant failure rate
- POH = power on hours per year. (For full usage is 8760 hrs/yr)
- MTBF = mean time between failures

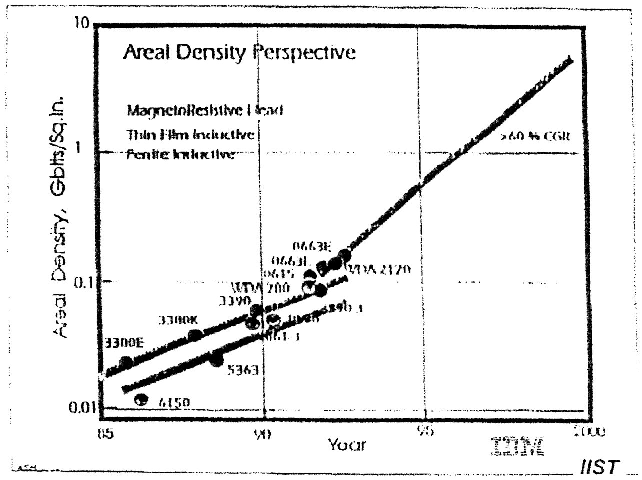
failure rate = **POH**

The number of failed drives is the product of the failure rate times the relevant population.

- Example: MTBF =500,000 hrs. Failure rate = 8760/500.000 = .0175.
- For a 10 million drive population the number of failures/year = 175,000 or 3,370 per week.

Today MTBF ranges from 250,000 to one million hours

Original Forecast of 60% CGR



Trends in storage density

- From 1956 to 1990 areal density increased by a CGR of 32%
- From 1991 areal density has been increasing at a 60% CGR
  - **This rate equates to a factor of 10 increase in storage density every 5 years.**
- This increase in the rate of technology progress is attributed to:
  - **Technology advances**
    - **Magnetoresistive heads**
    - **capability for much lower flying heights**
  - **Market demand**
  - **Consolidation within data storage industry**
  - **Advances in semiconductors**

Magnetic Disk Drive Technology

Magnetic Materials and Fields

Magnetic materials and fields outline

- Magnetism and magnetic phenomena
  - **Hard and soft magnetic material properties**
- Fundamental magnetic field quantities
- Basic equations governing magnetic fields and their interaction with magnetic materials and currents.
  - **Relationship to M-H loops**

Magnetism

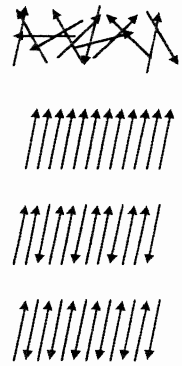
Coupling between adjacent atomic moments

Non-existent:
Paramagnetism

Parallel:
Ferromagnetism
e.g. Fe, Co, Ni

Partially antiparallel:
Ferrimagnetism
e.g. MnZn, NiZn

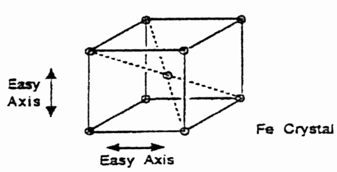
Antiparallel:
Antiferromagnetism



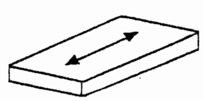
Curie Temperature:
Temperature at which magnetic material becomes nonmagnetic due to thermal disordering forces.

Magnetic anisotropy

Single crystals show anisotropic magnetism curves, i.e. it is easier to magnetize them along some crystal directions (easy axes) than others.



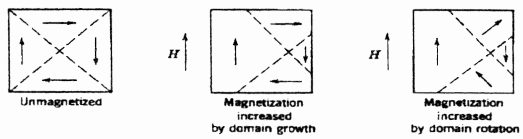
Shape



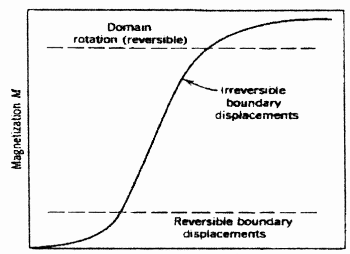
Longitudinal axis is the preferred axis for easy magnetization

ASH Magnetic Disk Drive Technology IIST

The magnetization process



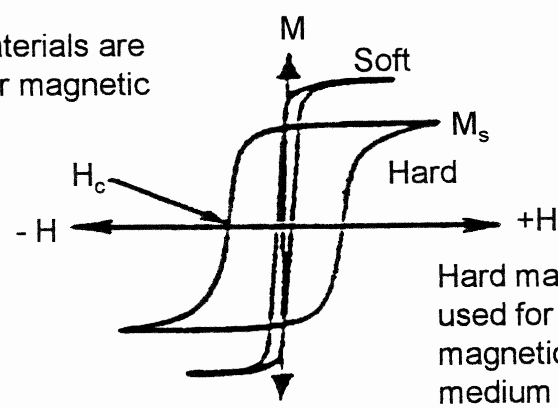
Fundamental magnetization process



Typical magnetization curve

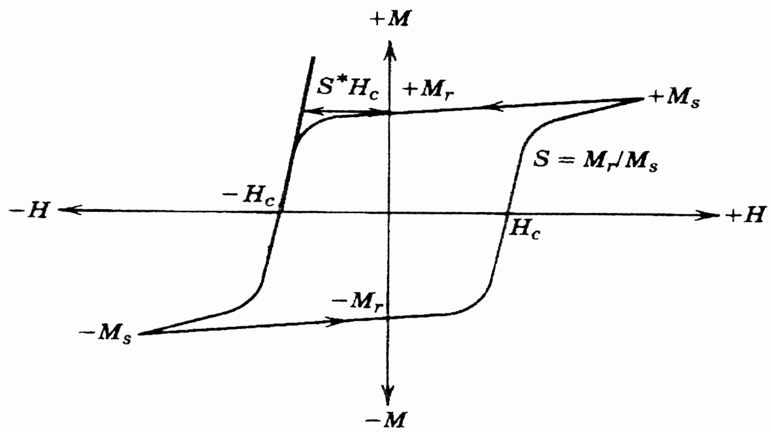
Hard and soft magnetic materials

Soft materials are used for magnetic heads



Hard materials are used for the magnetic storage medium

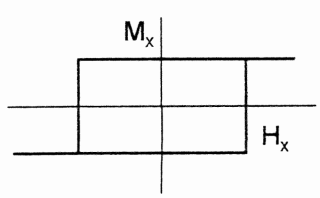
M-H loop with parameters



ASH | Magnetic Disk Drive Technology | IIST

M-H loop, easy axis

$$U_a = K\sin^2(\theta)$$
$$T_a = 2K\sin(\theta)\cos(\theta)$$
$$T_m = MH_x\sin(\theta)$$
for $\theta \approx 0,\ T_m = T_a$ **at** $H_x = H_k = \frac{2K}{M}$



ASH | Magnetic Disk Drive Technology | IIST

M-H loop, Hard axis

For applied field in hard direction

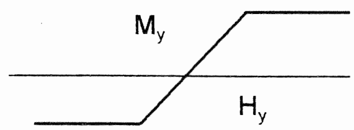
$$T_a = 2K\sin(\theta)\cos(\theta)$$

$$T_m = M H_y \cos(\theta)$$

$$\text{for } T_m = T_a$$

$$\sin(\theta) = \frac{M H_y}{2K} = \frac{H_y}{H_k}$$

$$\text{and } M_y = M\sin(\theta) = \left(\frac{M^2}{2K}\right) H_y$$



Magnetic units

Quantity	Symbol	cgs units	multiplier	SI units
		$B=H+4\pi M$	>>	$B=\mu_0[H+M]$
flux density	B	Gauss	1E-4	Telsa
flux	Φ	Maxwell	1E8	Webers
field strength	H	Oersted	1E3/4π	A/m
Magnetization	M	emu/cc	1E3	A/m
moment	m	emu	1E-3	$A{\cdot}m^2$
permeability (vacuum)	μ_0	dimensionless	4π 1E-7	Wb/A·m

Magnetostatics basic equations (1)

SI units

$$B = \mu_o [H + M]$$

M is only defined within a magnetized material, outside B and H are the same except for a constant factor,

$$\oint B \bullet dS = 0$$

The net flux out of any volume is zero (i.e.., there are no sources or sinks)

Therefore flux lines are continuous. Consequently.

$$\oint H \bullet dS = -\oint M \bullet dS$$

Thus, a magnetic field, H, will arise from an equivalent "magnetic charge" associated with a net change in total magnetization within a given volume.

Magnetostatics basic equations (2)

$$\oint H \bullet dl = NI$$

The integral around a closed path of the component of H in that direction will equal the ampere-turns encircled. This result is independent of the particular path chosen.

These are equations for the force on a current carrying conductor and the torque on a magnetized region due to an external magnetic field.

A linear medium with a high permeability provides a linear relationship between the magnetization and an applied field.

Faraday's Law

$$e = -\overline{Nd\varphi}\ . \quad \phi = A \cdot B$$

Magnetostatics basic equations (3)

As $\oint_S B \bullet ds = 0$

$B_{n1} = B_{n2}$ **(normal components at boundary)**

As $\oint H \bullet dl = NI$, **then when** $NI = 0$

$\int_a^b H \bullet dl = \Phi_b - \Phi_a$ **where** Φ **is a scalar magnetic potential**

and

$H_{1T} = H_{2T}$

(tangential components at boundary)

Field maps

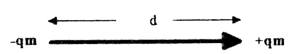
$$\frac{\Delta\Psi}{\Delta\Phi} = \frac{B\Delta s}{H\Delta n} = \mu_0\left[\frac{\Delta s}{\Delta n}\right]$$ **where Δs and Δn are scaling factors**

Self-demagnetizing fields

Magnetic dipole

$$m = qm \cdot d$$

$$qm = \sigma \cdot A$$



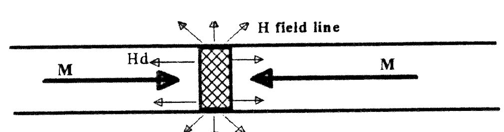
A = cross sectional area

$$M = \frac{m}{v} = \frac{qm \cdot d}{A \cdot d} = \frac{qm}{A}$$

If M varies with x the charge density is given by:

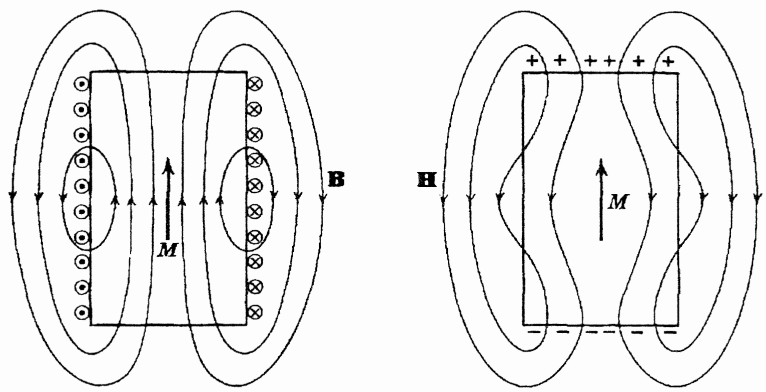
$$\rho(x) = -\frac{dM}{dx}$$

Recorded transition



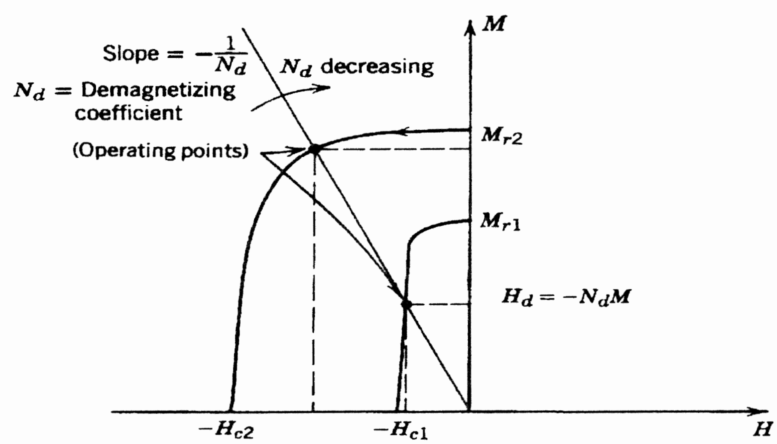
Hd shown is an example of self-demagnetizing field.

B and H fields of permanent magnet



ASH Magnetic Disk Drive Technology IIST

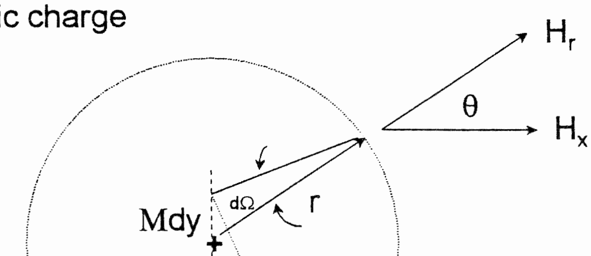
M-H curve and demagnetizing factor



ASH Magnetic Disk Drive Technology IIST

Magnetic field from magnetic line charge

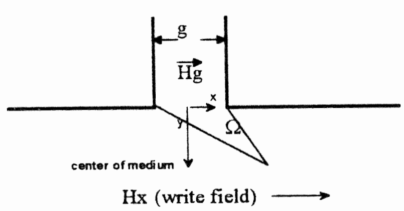
Mdy=magnetic charge



$$H_r = \frac{Mdy}{2\pi r}$$

$$H_x = H_r \cos\theta = \frac{M}{2\pi}\left(\frac{dy\cos\theta}{r}\right) = \frac{M}{2\pi}d\Omega$$

Magnetic write head



$H_x \cong H_g\left(\frac{\Omega}{\pi}\right)$ where $\Omega = \tan^{-1}\left(\frac{x+g/2}{y}\right) - \tan^{-1}\left(\frac{x-g/2}{y}\right)$

$H_g \cdot g + H_m \cdot l_m = NI$ For a high permeability core with a short path length

$H_g = \varepsilon\left(\frac{NI}{g}\right)$ where ε is defined as the head efficiency and is approximately 80%

Thermal instability

$$p_r = f_o e^{-(E_B/KT)}$$

p_r is the probability of a grain switching (crossing the energy barrier), f_o is approx. 10^9, E_B is the height of the energy barrier and KT gives the thermal energy available.

In the absence of demagnetizing fields (which lower the energy barrier) $E_B = k_u v$, where k_u is the anisotropy energy density of the particle and v is the particle volume.

The time constant of random switching of bits due to thermal effects is $1/p_r$.

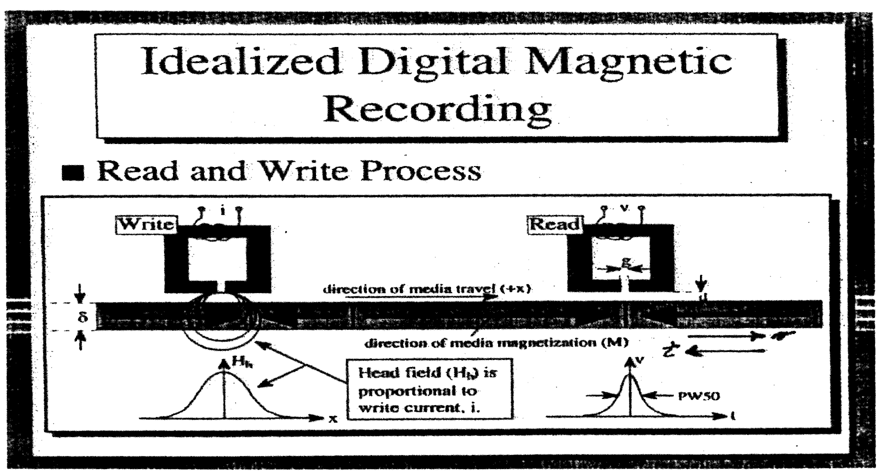
Density limit concerns

- Nonvolatile characteristic of medium
  - Thermal stability of medium magnetization
    - particle volume
      - thermal energy barrier and demagnetizing fields
    - limits on increasing H_c from attainable saturation magnetization for write head
    - Patterned bit cells
- Increasing data rate
  - Switching time of write head field
    - laminated head structures
      - suppress eddy current effects
  - Recording is pattern dependent
    - time constant of write current and head field responses

Magnetic Disk Drive Technology

Magnetic Recording Process

Digital magnetic recording



Magnetic recording process outline

- Recording of digital data
  - Modes of recording
- Basic attributes
  - Role of head and medium parameters in writing and readback
    - design relationships
- Write process
  - transitions and and their interactions
- Read process
  - transition response and wavelength (and frequency) characterization
    - Viewed as data channel

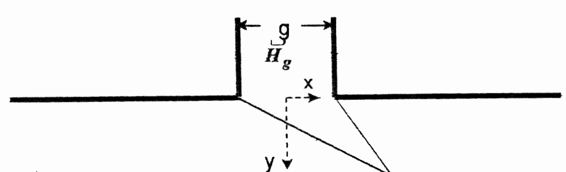
Attributes of digital magnetic recording (1)

- Magnetic fields scale geometrically
- Distance and time only become related through velocity
- A transition is associated with a change in saturation direction
  - **A transition gives rise to an output pulse whose polarity depends on the direction of magnetization change. (thus, the output pulses inherently alternate in polarity)**

Attributes of digital magnetic recording (2)

- Track density is a function of head width and head positioning and tracking
- Sense signal proportional to read track width
- Low head/medium spacing key to high density
- Ratio of bit to track density historically has essentially varied between 10 and 20
  - Air bearing head sliders give much better spacing control than actuator servo systems do for head positioning
- Areal density set by operational SNR

Magnetic head write field

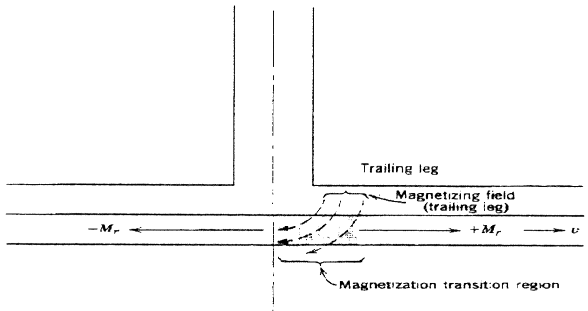


$$H_x \simeq \frac{H_g}{\ }\,\Omega$$

$$\Omega = \tan^{-1}\left(\frac{x+g/2}{\ }\right) - \tan^{-1}\left(\frac{x-g/2}{\ }\right)$$

$H_g = \frac{\varepsilon NI}{\ }$ where ε = head efficiency

Magnetization saturation reversal



The transition is displaced from the gap center, being formed under the trailing pole leg. Typically, the maximum fringing field at the medium location under the gap center is 2 to 2 1/2 times the medium coercive force, H_c .

Magnetic arctan transition

The following uses cgs-emu units, i.e., B = H + 4oM

$$M(x) = \frac{2}{\pi} M_r \arctan\left(\frac{x}{a}\right)$$

$\rho_m = -\frac{dM}{dx} = -\frac{2M_r}{\pi}\left(\frac{a}{x^2 + a^2}\right)$ **, where ρ_m is the volume magnetic charge density**

$H_d(x) = \frac{4\delta\, x}{x^2 + a^2} M_r$ **and** $H_d(\max) = \frac{2\delta}{a} M_r$**, and occurs at** $x = \pm a$

It is necessary that

$H_d(\max) \le H_c$ **, where H_c is the medium coercive force.**

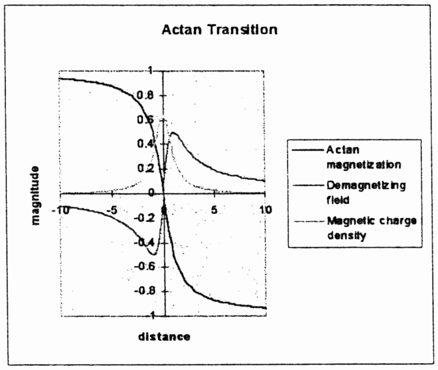
Thus, the minimum value for the transition parameter, *a*,

(based on the medium parameters) is

$$a = \frac{2M_r\delta}{H_c}$$

The magnetic charge distribution is given by the the derivative of the magnetization and is Lorentzian in shape.

Demag field and mag charge density arctan transition



The graphs are based on M_r =1, a = 1 and δ = 0.25

Williams-Comstock model

Considers M-H loop, write head field gradient and the medium properties together in determining the arctan *a* parameter.

Slope model:

$$\frac{dM}{dx} = \frac{dM}{dH}\left(\frac{dH_h}{dx} + \frac{dH_d}{dx}\right)$$

where this expression is evaluated at the transition center, i.e., x=0.

For an actan transition

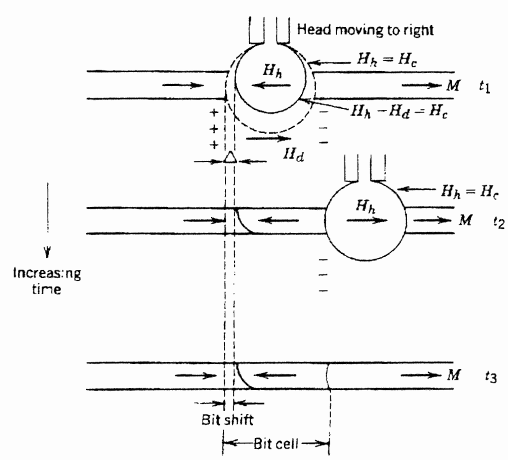
$$\frac{dM}{dx} = \frac{2M_r}{\pi a} \text{ and } \frac{dH_d}{dx} = \frac{M_r \delta}{\pi a^2}$$

The other two terms are the head field gradient at $H_x = H_c$ and the slope of the M-H loop at $H = H_c$. The minimum value of *a* is given by the medium expression alone (in the absence of the head). The loop shape and head field gradient establish how much larger *a* may be.

Non linear transition shifts (OW and NLTS)

- Cause - demagnetizing fields from adjacent transitions either increasing or decreasing the head field, H_h, changing the location where the combined field is equal to Hc, which is where the transition is written.
- **Hard or late bit transition** (arises when **overwriting** a recorded pattern)
  - Named from fact that when when the head writes a transition it may also creates a transition at the leading pole side of the gap, depending on the current state of the existing magnetization. (When this occurs the head field will also be reversing the magnetization under the gap.) The leading edge transition creates a demagnetization field opposing the applied head field, shifting the transition being written closer to the gap center. This shift means this "hard" transition will appear later in time than otherwise would be the case on readback.
  - Unpredictable as dependent on the existing state of the magnetization of medium in the head gap region

OW transition shift



Non linear transition shift (NLTS)

- Arises from preceding transition (or transitions)
  - If a recorded transition is immediately followed by another transition, the first one will create a demagnetizing field in the same direction as the new head field, thus tending to shift the point where the total field is equal to Hc. The effect is to shift the second transition closer to the previous one, or further from the head gap center.
  - This shift becomes of significance at higher densities where transitions become closer and closer together.
  - Since the incoming bit sequence being written is known precompensation (adjusting the timing of current reversals) can be used to control the spacing of transitions.

Readback principle of reciprocity

A small current loop is equivalent to a magnetic dipole where

$$I*(W\delta)=M_x*(W\delta)dx \text{ or } I=M_x dx$$

W is the track width and δ is the medium thickness, the product being the cross-sectional area of the medium.

The principle of reciprocity states that the coupling between two current loops is given by a "mutual inductance" term, i.e.,

$$\phi_h = kI_m \quad \text{and} \quad \phi_m = kI_h$$

Since $k = \frac{\phi_m}{I_h} = (W\delta)H_x$

Therefore

$$\phi_h = W\delta H_x M_x dx$$

or

$$\Phi_h(vt) = \int \phi_h(vt)dx = W\delta \int M(x-vt)H_x dx$$

Linearity is assumed and in this case benefits from the existence of the non-magnetic gap.

Finally

$$e(vt) = N\left(\frac{d\Phi_h}{dt}\right) = Nv\left(\frac{d\Phi_h}{dx}\right) = NvW\delta \int M'_x(x-vt)H_x dx$$

ASH Magnetic Disk Drive Technology IIST

Head field and readback signal

A step function in magnetization corresponds to an ideal transition or an "impulse" type input for the magnetic head channel.

For a *step function* change in M_x,

$$e(vt) = e_x(vt) \propto vH_x(x)$$

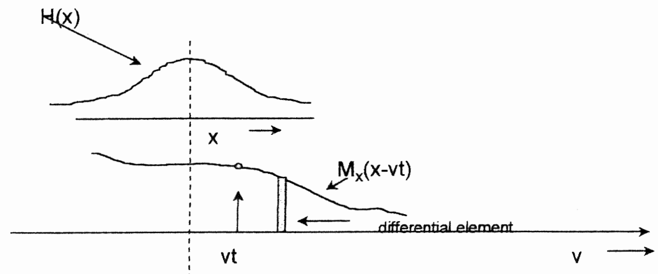
Likewise foor a step function change in M_y,

$$e(vt) = e_y(vt) \propto vH_y(x)$$

When both components are present,

$$e(vt) \propto v\left(\Delta M_x \cdot H_x + \Delta M_y \cdot H_y\right)$$

Readback superposition integral



$$e(vt) = \int M_x'(x - vt)H_x(x)dx$$

The medium is passing the head from left to right or the magnetization as a function of time at the gap center is

or in time the magnetization is sensed from from right to left.

Distance and time relationships

$x = vt,$ and $f = v / \lambda$

where λ is the recorded wavelength

and f is the frequency.

The wavenumber $k = \frac{2\pi}{\lambda} = \frac{\omega}{v}$

or $f = k\left(\frac{v}{2\pi}\right)$

in terms of transition density

the flux changes per inch (fci) is given by

$fci = \frac{2}{\lambda}$ or $k = \pi fci$

Readback
MR vs. Inductive head

For an inductive head we have seen for longitudinal recording that

$$e \quad \alpha \quad \int M'_x H_x dx$$

or the voltage is proportional to the gradient of magnetization (summed over the sensing region).

The MR head senses the flux exiting the medium and entering the stripe (which is normal to the surface). At a transition the external magnetic field normal to the medium is Hy which arises from changes in Mx.

$$H_y \quad \alpha \quad \frac{dM_x}{dx}$$

Therefore the flux in the stripe is proportional to the gradient of medium magnetization in the same manner as the inductive head.

PW50 equation (1)

From the Karlqvist expression we find

$Pw50 = \sqrt{g^2 + 4d^2}$

where g is the gap length and d is the spacing from the head.

Now, for an "zero" gap head the x component of the fringing field is given by

$H_x = C * \left(\frac{d}{x^2 + d^2} \right)$, where C is a constant. (This expression is a Lorentzian pulse.)

Integrating over the medium thickness, δ, we get the expression

$d(d+\delta)$ as the term to replace d^2.

PW50 expression (2)

The "infinitesimal" gap head has a Lorentzian waveform for H_x and the wavelength dependence of signal on spacing for this pulse shape is given by e^{-kd}. The Fourier transform of an arctan function is:

$M(k) \;\alpha\; \frac{e^{-ka}}{k}$ **and** $M'(k) \;\alpha\; e^{-ka}$. **(** $\frac{dM}{dx}$ **is also Lorentizan).**

Now, these two terms can be combined, i.e., $e^{-kd}e^{-ka} = e^{-k(d+a)}$ **to show the influence of these two parameters.**

Since, *d* and *a* have the same channel response function it is possible to replace *d* by *(d + a)* in the expression for PW50. Note, however, that *d* is associated with the readback process and *a* with the write process. Making all these substitutions we get the following equation for PW50.

$$PW50 = \sqrt{g^2 + 4(d+a)(d+a+\delta)}$$

Output from magnetic transition

$$e(\bar{x}) = \frac{e_o}{1 + \left(\frac{x}{PW50/2}\right)^2}$$

$$PW50 = \sqrt{g^2 + 4(d+a)(d+a+\delta)}$$

$$a \approx \frac{2M_r\delta}{H_c}$$

$$\int e\,dt = N \int d\varphi = N\left[(\varphi_\infty) - (\varphi_{-\infty})\right]$$

Channel wavelength characterization (1)

Geometrical factors

head gap	$\left[\frac{\sin\left(\frac{kg}{2}\right)}{\frac{kg}{2}}\right]$
spacing	e^{-kd}
magnetic thickness	$\left[\frac{1-e^{-k\delta}}{k}\right]$

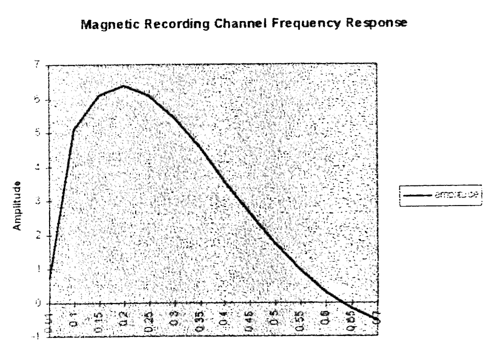
Channel wavelength characterization (2)

Now sensor output arises from $\left[M'(k) = kM(x)\right]$ or giving the factor k

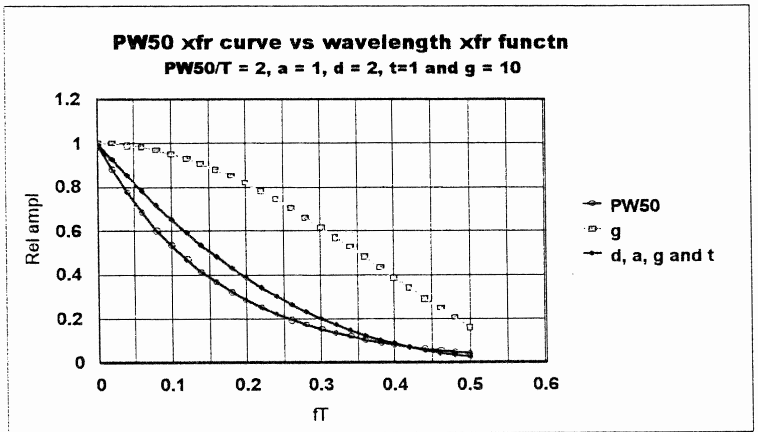
$$E(k)\ \alpha\ k\left[e^{-kd}\right]\left[\frac{1-e^{-k\delta}}{k}\right]\left[\frac{\sin\left(\frac{kg}{2}\right)}{\frac{kg}{2}}\right]$$

Magnetic recording transfer function

- Zero at dc
- 6 dB rise initially
- Rapid attenuation due to spacing
- Gap null



Lorentzian compared to Karqvist



d = spacing, a = arctan parm, g = gap and t = medium thickness

Views of recording channel

- MR channels are power limited in that the write field is limited by the head saturation magnetization and the medium magnetization by M_r.
- The basic channel frequency response has:
  - A zero at dc.
  - Rises initially at 6 dB/octave then drops rapidly due to spacing loss.
  - A zero at the gap null
- The Lorentzian PW50 pulse model essentially treats all loss factors in terms of an "equivalent" spacing equal to PW50/2.
- PW50/T measures channel usage efficiency and today is in the range of 2
  - You will find T is sometimes defined as: (a) user bit period, (2) channel bit period, (3) minimum transition spacing period

NRZ and NRZI modes of recording

NRZ

- A binary "1" is associated with one direction of saturation and a "0" with the opposite
  - Both a 1 and a 0 are uniquely defined
  - An output signal only is available when there is a change in the binary sequence
  - The signal polarity is defined by the order of the binary sequence change
    - A missing or extra readback pulse will propagate errors in subsequent bits

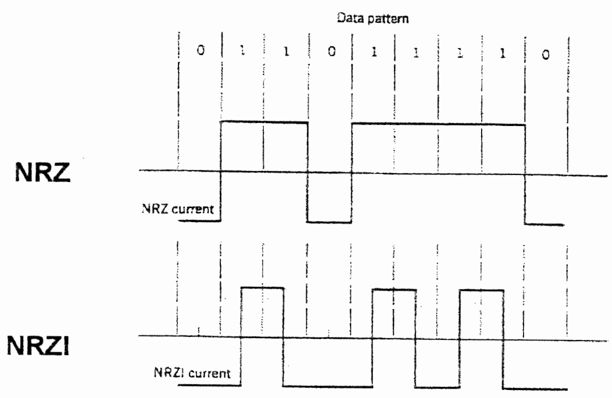
NRZI

- A "1" is defined as a transition and a "0" as the absence of a transition
  - A "1" is associated with an output pulse of either polarity
  - A missing or extra pulse affects only a specific bit

Current reversals

- The fundamental input signals are current reversals
- The recording channel is usually characterized by a step function response since it has no true impulse response.
- A current reversal writes a transition or narrow zone of magnetic charge proportional to the spatial derivative of the magnetization
- Linear density increases gives rise to ISI (or pulse crowding), partial erasure and NLTS due to the closer physical spacing of transitions
  - ISI is a linear and predictable consequence of pulse superposition

NRZ and NRZI current for same data pattern



Channel response and transition spacing

While the readback process is essentially linear, starting from a known magnetization pattern. The write process is not linear (in terms of the relation of current to magnetization) particularly as linear density is increased. In writing a dibit pattern, for example, the readback dipulse departs more and more from that predicted from the superposition of two isolated pulses.

This non-linear behavior arises from the influence of the preceding transition on the total field when writing a transition.

Channel performance can be characterized by how small a PW50/T can be achieved, providing a measure of the linear density obtained from a given "quality" of the channel.

Precompensation is one way to mitigate this departure from linearity.

Magnetic Disk Drive Technology

Magnetic heads and media

Magnetic heads outline

- Inductive write heads
- Thin film inductive read heads
  - Design aspects
- MR heads
  - Principle of operation
    - Sensitivity, linearity
    - Noise, thermal spikes, etc.
  - Write-wide read-narrow feature
- GMR devices
    - Basic phenomenon
    - Spin valve implementation

Read head technology progression

- Ferrite or thin film inductive heads
  - Based on Faraday's Law
    - rate of change of flux through coil on core
- Magnetoresistive or AMR heads
  - uses resistance change arising from resistive anisotropy with direction of magnetic field.
- GMR/Spin valve heads
  - Spin dependent electron scattering with magnetization directions in multilayer films
    - Greatly increased ratio of $\Delta R/R$ leads to much higher sensitivity

Write heads

- **Inductive magnetic heads are always used for writing**
  - **A higher B_{sat} head material allows the use of higher coercivity media.**
    - **For overwrite the write field at the medium is approximately 2 1/2 times H_c at gap center.**
    - **magnetic core designs need to assure that the core body and gap pole tips do not saturate.**
  - **<u>The head flux switching time is a major barrier to increasing data rates.</u>**

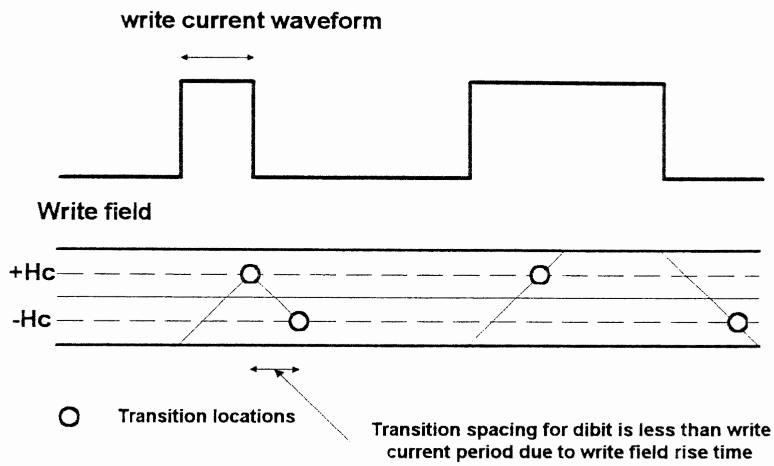
Features of separate write head

- Use write-wide read-narrow to deal with track registration problems
- Can use larger gap for writing
  - Write gap length limited by desire for high write field gradient
- Can use fewer turns as not concerned with a readback signal amplitude
- Greater write current possible for high data rates
  - No read after write issues

Head magnetic materials

Material	Saturation flux density in Gauss
MnZn Ferrite	5000
Permalloy (NiFe)	10,000
Sputtered Sendust (AlFeSi)	10,000
Sputtered Amorphous CoZrX	15,000

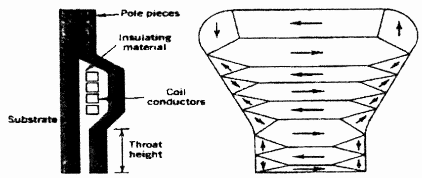
Dibit transition shift due to write field time constant



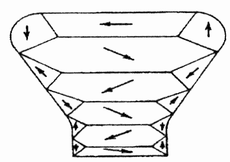
Inductive read heads

- **On readback, the magnetic fields from the medium are so low that saturation of the head material is not an issue. However the effectiveness of the magnetic coupling between the head coil and the medium is important to achieve a suitable readback amplitude, requiring a high permeability and a suitable ratio of gap length to spacing.**
- **Signal is proportional to N, the number of turns, but the head inductance is proportional to N^2.**
  - **Head frequency response is less than that of MR heads.**

Thin film head



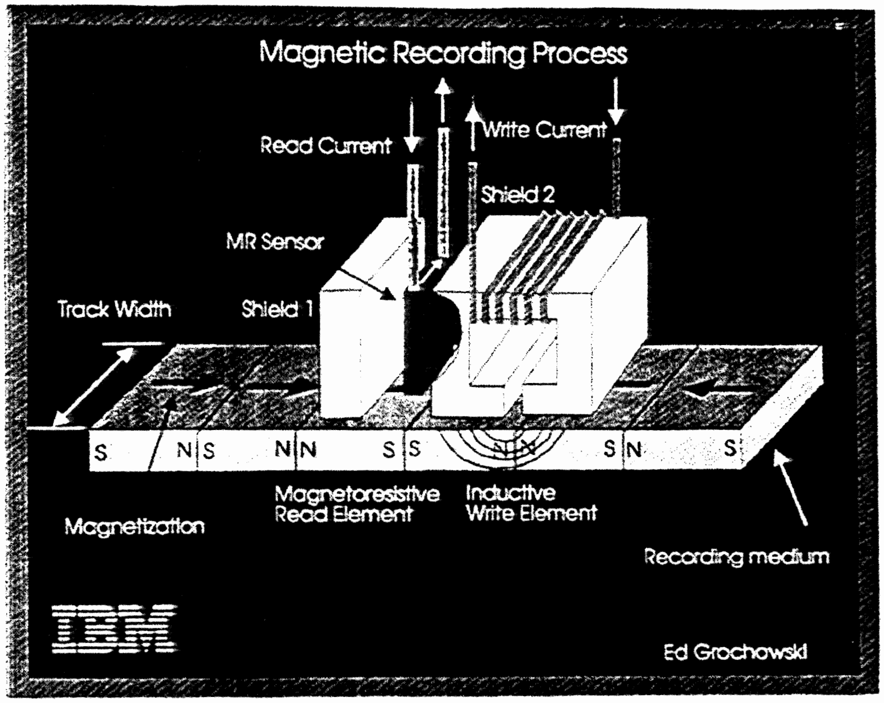
(a) Cross section of head *(b) Pole piece domains, no applied field*

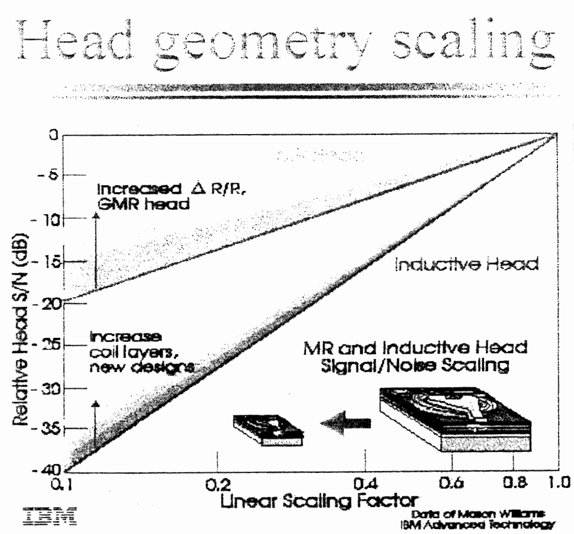


(c) Pole piece domains, peak write current

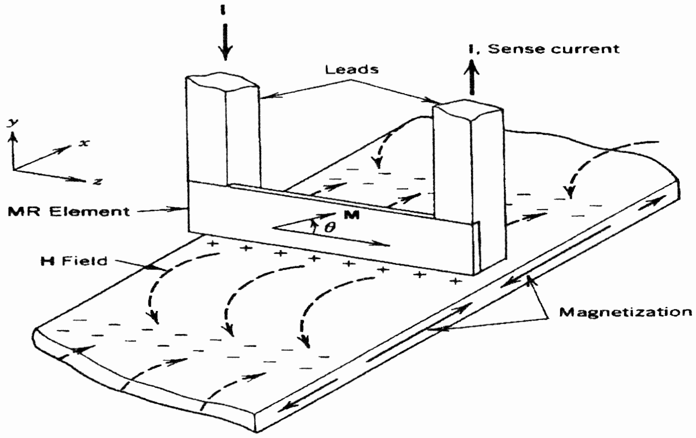
MR read/inductive write heads

- **Advantages**
  - **Large signal/unit trackwidth**
  - **Velocity independent**
  - **Large bandwidths possible**
  - **Separately optimized read and write functions**
    - **Write-wide read-narrow**
    - **Isolated pulse shape with no undershoots**
    - **Write head requires few turns**
- **Disadvantages**
  - **Thin film active read element exposed at ABS**
    - **Thermal asperities**
    - **Head/medium voltage gradient**
    - **Corrosion, smearing**


Magnetic Recording Process
Read Current
Write Current
Shield 2
MR Sensor
Track Width
Shield 1
S N S N N S S N S N S
Magnetoresistive Read Element
Inductive Write Element
Magnetization
Recording medium
IBM
Ed Grochowski


Head geometry scaling
Increased Δ R/R, GMR head
Inductive Head
Increase coil layers, new designs
MR and Inductive Head Signal/Noise Scaling
Relative Head S/N (dB)
0
-5
-10
-15
-20
-25
-30
-35
-40
0.1
0.2
0.4
0.6
0.8
1.0
Linear Scaling Factor
IBM
Data of Mason Williams IBM Advanced Technology

Magnetoresistive sensor



AMR heads

Anisotropic magnetoresistance.

The resistivity of an MR stripe is given by

$\rho = \rho_0 + \Delta\rho \cos^2\theta$

where ρ_0 is the resistivity when the magnetization is perpendicular to the direction of current flow.

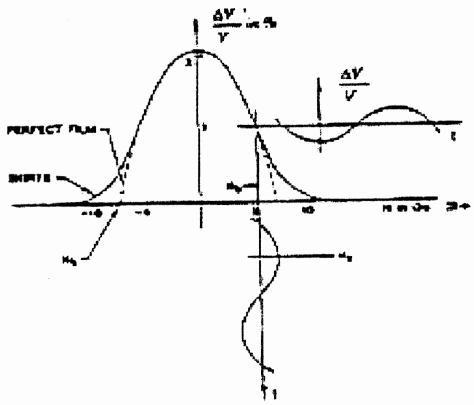
θ is the angle between the direction of magnetization and direction of the current.

$\frac{\Delta\rho}{\rho}$ **is in the range of 2% to 3% for permalloy.**

Seagate MR Head Technology

RESPONSE LINEARIZATION

To maximize the linear response of an MR sensor, the magnetization is biased. The optimal bias angle is about $\phi_o = 45°$.

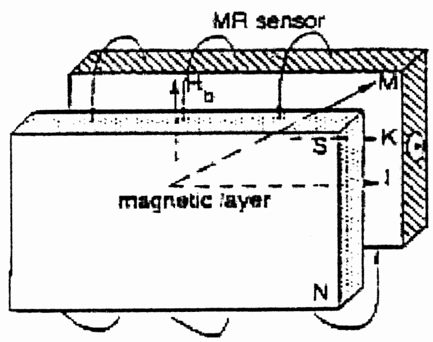


Seagate

Seagate MR Head Technology

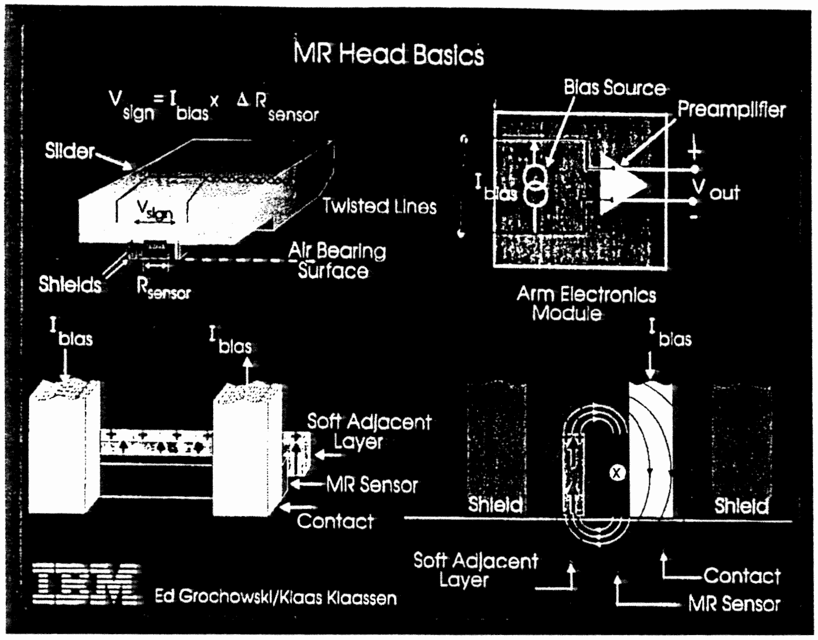
BIASING TECHNIQUES

- Rotate M relative to easy axis



- permanent magnet
- external field
- asymmetric gap location
- soft adjacent layer
- exchange
- shunt biasing

Seagate



MR Head Sense Signal

- MR head signal
  - Varies linearly with track width
  - Varies inversely with sensor height
  - Approximately linearly with disk $M_r t$

Use

$\Delta R_{MR} / R_{MR}$ **detection, for example** $e_s = \frac{\Delta R}{R} E_{bias}$

Then the output signal is insensitive to variations in

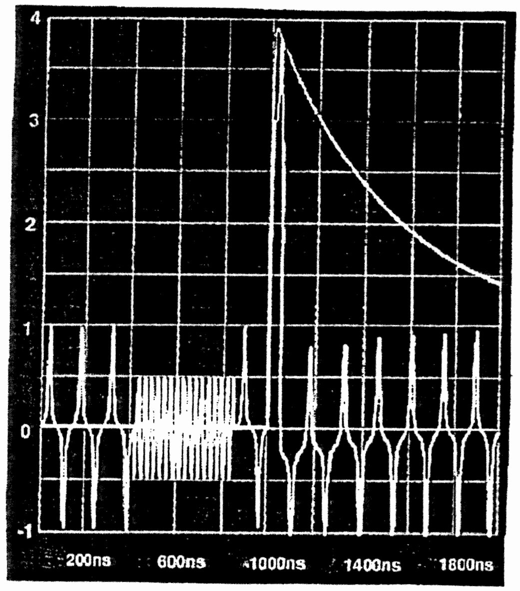
That is, voltage bias and voltage sensing
or current bias and current sensing

TA Base Line Restoration

Detect base line variation, subtract from signal

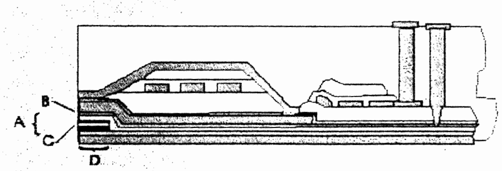
Asperity Reduction Circuit (ARC)

Subtractive restoration also provides restored TA



MR head specifications

Advanced MR Head Designs



		1 Gbit/in² Head	3 Gbits/in² Head
A	Total Read Gap	0.25 µm	0.20 µm
B	Sensor/Shield Spacing	< 1200 Å	< 1000 Å
	Read Trackwidth	2 µm	1.1 µm
C	MR Layer	150 Å	120 Å
D	Sensor Height	1.0 µm	0.6 µm
	Flying Height	1.5 µ-in	1.5 µ-in

IBM

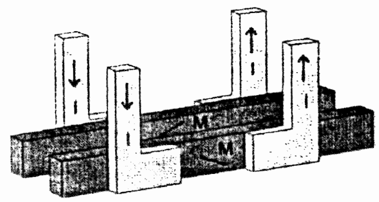
Data of Robert Fontana
IBM Advanced Technology

Dual-stripe head

In the dual stripe head, two insulated elements are biased in opposite directions by parallel currents and sensed differentially. The elements are between shields not shown here.

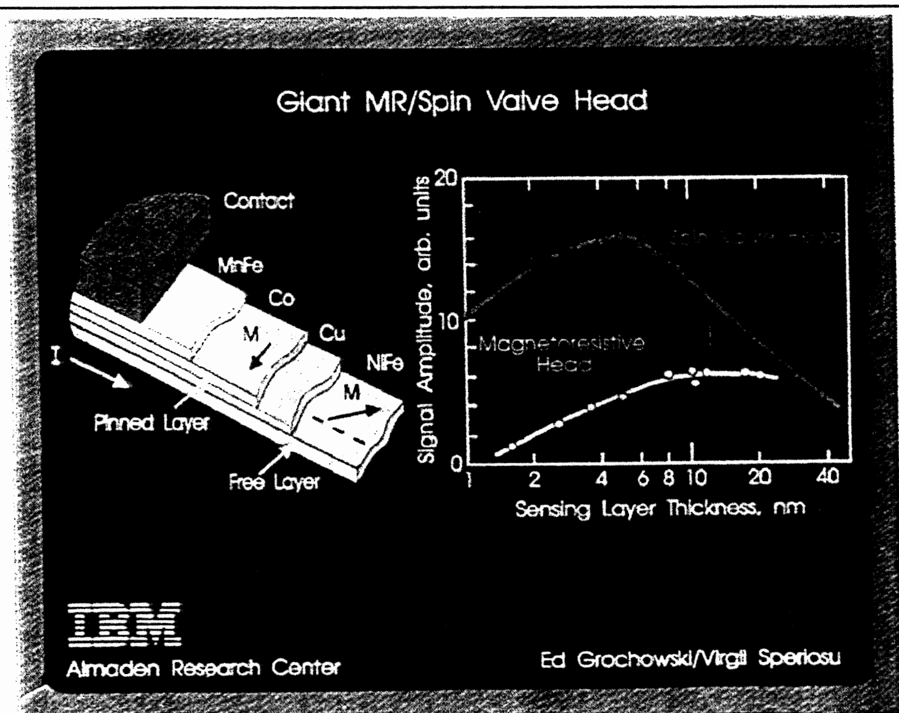
Advantages

Common mode thermal rejection
Symmetric track profile and signal
No signal lost in shunting

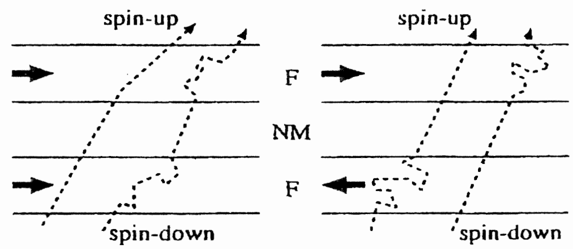


Potential problems

Shorting of elements spaced 20% of s-s gap
Alignment tolerances
Matching of MR elements



GMR -spin valve model



$$\rho\uparrow\uparrow = \frac{\rho\uparrow\rho\downarrow}{\rho\uparrow+\rho\downarrow} \qquad \rho\uparrow\downarrow = \frac{\rho\uparrow+\rho\downarrow}{4}$$

$$\Delta\rho = \rho\uparrow\downarrow - \rho\uparrow\uparrow = \frac{(\rho\uparrow-\rho\downarrow)^2}{4(\rho\uparrow+\rho\downarrow)} > 0 \quad \therefore \rho\uparrow\downarrow > \rho\uparrow\uparrow \text{ and } \frac{\Delta\rho}{\rho\uparrow\downarrow} = \left(\frac{\rho\uparrow-\rho\downarrow}{\rho\uparrow+\rho\downarrow}\right)^2$$

$\frac{\Delta\rho}{\rho}$ **can be much larger than for the standard AMR head.**

Spin valve MR head

Pinned layer magnetization is normal to medium with a thickness in the range of 3.5 nm.

Pinning layer is anti-ferromagnetic material exchange coupled to the pinned layer.

Free layer is longitudinally biased with a thickness in the range of 6 nm and separated from pinned layer by a non-magnetic conducting spacer, e.g., Cu with a thickness in the range of 2.5 nm. Then:

$$\frac{\Delta R}{R} \propto M_s \sin(\Theta) \propto H_y$$

Thus, the spin valve offers better linearity than the AMR head.

GMR advantages

- Higher signal ranging up to 5 times standard AMR (anisotropic magnetoresistive) head
  - This directly translates into narrower tracks or a much higher tpi potential
- Readback transfer function is linear
- The complexity and cost do not differ greatly from the conventional MR head
- Like the MR head this increased sensitivity places an increasing importance on low noise media

MR/GMR heads re disk medium

- A very smooth surface is necessary to minimize the occurrence of thermal spikes
- The head/medium interface must minimize the chance of electrostatic discharges
- The $M_r t$ of the medium must be reduced to be compatible with any reduction in MR stripe thickness
- To fully take advantage of the sensitivity of MR heads very low noise media are required

Magnetic disks

- Media layers
  - Magnetic versus physical spacing
- Head/disk interface
  - tribology considerations
- High density magnetic requirements
  - medium noise
    - domain or particle size
      - thermal stability limits
  - Mr and Hc
- Relation to head technology

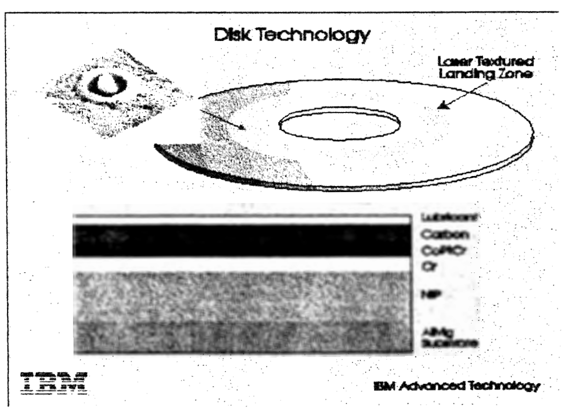
Head/media interface

- Minimizing head-medium spacing as near to zero as possible is the route to higher and higher storage density
  - linear density benefits from the laws of scaling
  - track density benefits from a reduction in side fringing fields
- Flying versus "contact recording"
- While disk recording has always involved some intermittent contact the requirement is that performance degradation due to wear is unacceptable.
  - Tape recording can use contact to improve performance since the number of hours of usage on a particular medium is small and head replacement feasible.
- Head, medium surfaces and slider design are critical.
  - Proximity, pseudo, virtual, etc. indicate the difficulty in defining the limit to low flying recording without wear or performance deterioration

Magnetic spacing vs. physical separation

- Magnetic spacing - the distance from magnetic head poles to midpoint of magnetic medium
- Physical separation - distance from the low point of the slider to the lubricant layer
- Differences arise from:
  - Disk texture
    - Reduces slider-surface contact area and hence stiction
  - Disk overcoat
    - necessary to protect magnetic layer mechanically, corrosion resistance
  - Disk lubricant
    - reduce wear, static friction
  - Pole tip recession (and head location on slider)

Disk technology



Tribological Regimes

- Stiction
  - Related to force required to overcome on startup the stiction force that develops from the slider resting on the medium surface for some period.
- Frictional wear
  - Periods of startup and shut down where the low speeds do not support an air bearing (or flying).
- Flying
  - An air bearing exists such that only have intermittent contacts caused by medium asperities, contaminants or mechanical resonance's.

Wear and head/medium interface

Archard's wear law

$$V = \frac{kF_c s}{H}$$

where:

V = **total wear volume**

F_c = **normal contact force**

s = **total sliding distance**

H = **hardness of wearing surface**

k = **wear coefficent**

This relationship argues for a lightly loaded low mass slider and a hard disk overcoat.

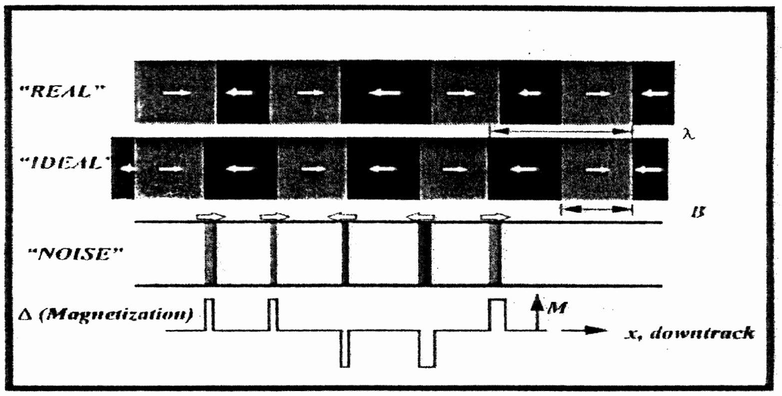
Magnetic medium properties

- Directions
  - To go to higher and higher linear densities the transition "a" parameter must decrease and therefore the medium coercive force must continually increase and the M_rt product decrease.
  - The increased sensitivity of MR heads also requires a lower M_rt in order to operate in the linear region of the transfer function of the head.
  - Since the sensitivity to noise sources on the disk increases to the same degree as the signal, lower noise media become essential. Smaller, isolated particles (or fine grain structure) are needed.
  - A thinner medium reduces the effective magnetic spacing as well as the flux from the medium surface.
- For 10 gigabit/in^2: $H_c \cong 3000$ oe, $M_rt \cong 0.6*10^{-3}$ emu/cm^2
- Grain size $\cong 10$ nm for $SNR_p > 20$ dB

Media "noise"

- Grain structure of media and "particle" volume density
  - SNR_{power} is proportional to the number of particles in a bit cell
- Transition jitter
  - noise increases with linear density
- Partial Erasure
  - Transition integrity destroyed over part of track
    - shows up when transition separation only 3 or 5 times the transition parameter "a".
- Medium defects
- The "noise" signals from the medium are "amplified" to the same degree as is the recorded "signal".
  - Thus, the greater sensitivity of the MR head places greater demands on achieving low noise media.

Transition jitter: magnetization noise



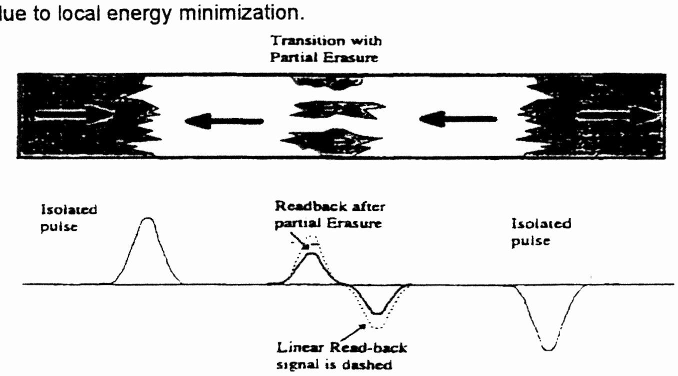
Seagate

Partial erasure

Results from:

1. pulse broadening with closer transition spacing

2. elimination of cross track regions of opposite magnetization due to local energy minimization.



SNR and particle density

n = number particles per unit volume
v = volume of bit cell
S = output signal, N = medium noise

✣e = signal from individual particle

$$S^2 \propto (nv\,✣e)^2$$

$$N^2 \propto nv(✣e)^2$$

or, $SNR_p \propto nv$. the number of particles in a bit cell

==

SNR vs tpi (W= track width)

$$S \propto W,\ N^2 \propto W\ .$$

therefore

$$SNR_v \propto \sqrt{W} \propto \frac{1}{\sqrt{tpi}}$$

Magnetic media capabilities

The magnetic material limits are set by superparamagnetism , that is when the magnetic particles in a continuous film become so small that they become thermally unstable.

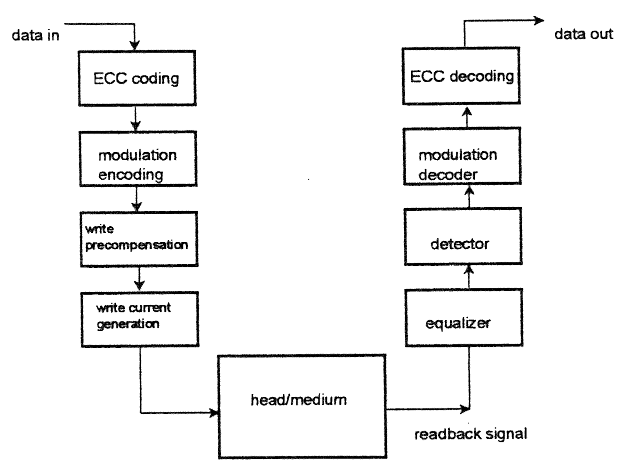
This phenomenon as a limiting factor is seen to arise for the particle sizes required to record in the range of 100 gigabits/inch^2.

Today the ability to correctly resolve (write and read) smaller and smaller bits in the magnetic medium poses the major challenges to continuing progress.

Magnetic Disk drive Technology

Magnetic Recording Channel

Magnetic recording channel block diagram



MR channel outline

- Signal, noise and BER
- RLL modulation codes
- Peak detection
  - Equalization, detection and window margin
- Digital signal processing
  - NRZ/NRZI sampling detection
- Partial response recording
  - PR4
    - Equalization, precoder, Veterbi detector
    - Error events and error rate
  - Extended PR4 systems
    - EPR4

Signals, noise, distortion and interference

- Signal
  - Waveform ("ideal") that hope to obtain
- Noise
  - unpredictable random perturbations
    - thermal (Johnson) noise
    - magnetic domain instabilities
    - medium granularity
- Distortion
  - Linear: ISI, frequency bandwidth limitations
  - Non-linear: transition shifts, partial erasure
- Interference
  - Adjacent track and old information
  - EMI

Head and preamplifier noise

- Johnson or thermal noise
  - Arises from real part (resistive) of head impedance and pre-amplifier

$$n_v = \sqrt{4kTR\Delta f}$$ **rms noise voltage**

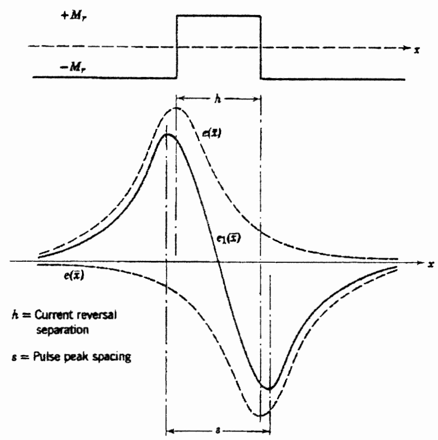
Commonly expressed in nano-volts per root Hz since the actual magnitude of the noise depends on the bandwidth of the channel.

The stripe resistance of an MR head is 20 to 40 ohms. The pre-amplifier noise contribution is less but may be of the same order of magnitude.

Media "noise"

- Grain structure of media and "particle" volume density
  - SNR_{power} is proportional to the number of particles in a bit cell
- Transition jitter
  - noise increases with linear density
- Partial Erasure
  - Transition integrity destroyed over part of track
    - shows up when transition separation only 3 or 5 times the transition parameter "a".
- Medium defects
- The "noise" signals from the medium are "amplified" to the same degree as is the recorded "signal".
  - Thus, the greater sensitivity of the MR head places greater demands on achieving low noise media.

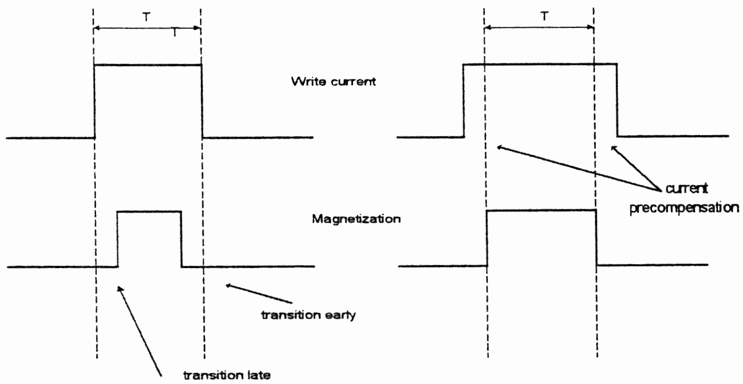
ISI with dipulse signal (distortion)



Channel non-linearities

- Hard/easy bit transitions
  - Shift is dependent on particular pattern being **overwriten**
  - Source is demagnetizing field from temporary transitions created at leading edge pole of head
  - Leads to shifts in "ideal" transition locations when overwrite data
  - Not a function of transition density
- NLTS
  - Transition shifts arising from preceding transition(s) that have just been written
  - Since bit sequence being written is known can use precompensation to make timing adjustments of write current to reduce
  - More and more significant as density increased
- Partial erasure
  - Output signal amplitude reduced from broadening of transition
  - Output signal amplitude reduced from loss of some track width contribution to transition signal

NLTS and write precompensation



Late transitions due to overwrite, early transitions due to immediately previously written transition NLTS).

BER and SNR

- It is important to know the tradeoffs between SNR and the raw bit error rate (BER).
  - Normally random noise (or a Gaussian distribution) is assumed
- The relation between BER and SNR obviously depends on the definition of SNR - for which there is not any generally accepted agreement
  - Typically the signal is defined as the zero to peak output voltage
- An increase in SNR can be taken to increase density while preserving the previously available SNR or to reduce the bit error rate. The BER vs SNR curve shows the major improvement in error rate available from a relatively small increase in SNR
- **However, this relationship gives no information on how SNR varies with increasing density**

SNR and BER relationship (1)

$$p(e_n) = \frac{e^{-\left(\frac{e_n^2}{2\sigma^2}\right)}}{\sqrt{2\pi}\sigma}$$

$$BER = P(|e_n| \geq K) = erfc\left(\frac{K}{\sqrt{2}\sigma}\right)$$

K is a detection threshold and $P(e_n)$ is the cumulative probability that the noise voltage exceeds the threshold (an error results whenever a pulse is transmitted and not detected or no pulse is transmitted but a pulse is detected).

SNR and BER relationship (2)

$$SNR = \frac{E_{0-pk}}{\sigma}, \ then \quad K = \frac{E_{0-pk}}{2}$$

$$BER = erfc\left(\frac{SNR}{2\sqrt{2}}\right) = 2Q\left(\frac{SNR}{2}\right)$$

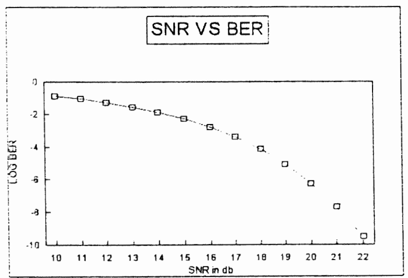
where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{z^2}{2}} dz$$

SNR and BER relationship (3)

For reasonable values of SNR we can approximate the BER as follows:

$$BER \approx \frac{e^{-z^2}}{\sqrt{\pi}z} \quad \text{where} \quad z = \left(\frac{SNR}{2\sqrt{2}}\right)$$



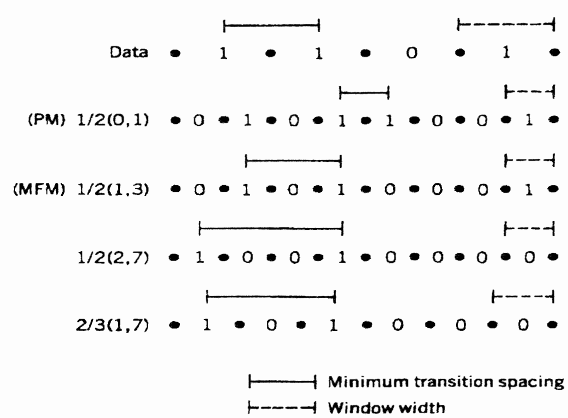
Pulse detection and channel codes

- **Peak detection**
  - Provides more accurate clocking of transition locations than straight amplitude detection
    - is less sensitive to signal waveform variations arising from changes in recording conditions than would be straight amplitude detection.
- **Amplitude sensing** - lends itself to digitizing signal for processing
- Pulse detection has as key sources of errors
  - Intersymbol interference
  - loss of clock synchronization with data
    - ISI and clocking limitations can be mitigated by channel codes
  - noise

RLL codes

- NRZI recording (a 1 bit corresponds to a transition)
  - m = number of user bits in group
  - Number of possible binary sequences = 2^m
  - Convert m bits to n code bits where n > m
  - Possible number of code sequences = 2^n
  - As $2^n > 2^m$ can choose bit sequences that meet certain restrictions
- To minimize ISI, place a bound on the minimum number of zeros between to ones.
- To assure reliable self clocking, place a bound on the maximum number of zeros between two ones.

View of RLL code constraints



RLL channel code parameters

T = user data bit period
m = # of data bits in group
n = # of channel code bits in group
$\frac{m}{n}$ = code rate
d = minimum # of zeros between ones
k = maximum # of zeros between ones
T_{min} = minimum time between transitions

$$T_{min} = \frac{m(d+1)}{n}T$$

T_{max} = maximum time between transitions

$$T_{max} = \frac{m(k+1)}{n}T$$

W = channel clock window

$$W = \frac{m}{n}T$$

$$\frac{bpi}{fci_{max}} = \frac{T_{min}}{T} = \frac{m(d+1)}{n}$$

$$\frac{fci_{max}}{fci_{min}} = \frac{k+1}{d+1}$$

1,7 code

Coding rules

Basic encoding table:

Data words	Code words
0	101
1	100
10	1
11	10

Substitution Encoding table if the d = 1 constraint is violated.

Data words	Code words
00.00	101.000
00.01	100.000
10.00	001.000
10.01	010.000

Code comparisons

Channel	PD	PRML
Code: (m/n)(d,k)	2/3(1,7)	8/9(0,4,4)
Code rate	2/3	8/9
Density ratio (bpi/fci_max)	4/3	8/9
Channel rate/data rate	3/2	9/8
Window	(2/3)T	(8/9)T
fci_max/fci_min	8/3	5

Peak detection

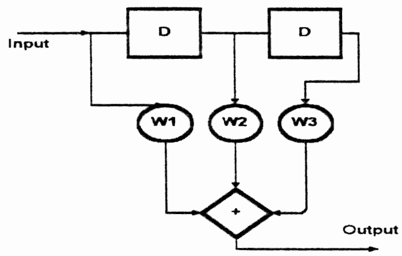
The standard method to detect digital data recorded on magnetic disks from the beginning.

Prior magnetic recording devices such as drums had a head per track and one track was used for clocking. In this situation amplitude threshold detection worked well,

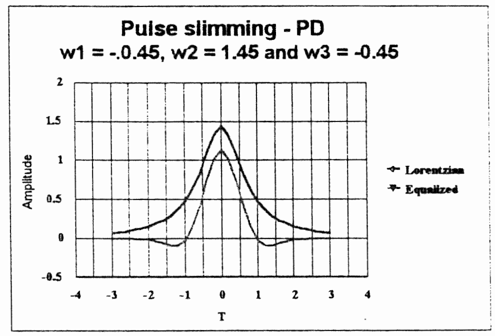
In disk recording, where one head covers many tracks and self clocking in necessary for timing, the advantages in peak sensing of more immunity from amplitude variations and robustness with self clocking made it the technique of choice for 35 years.

Peak detection is well matched for a PW50/T ratio of 1 or somewhat more. (T = channel bit period)

PD equalizer



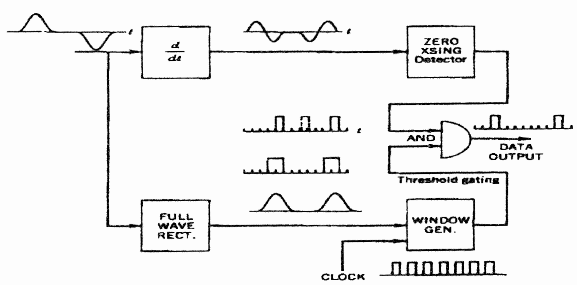
Tapped delay line equalizer



$$E_o(j\omega) = \left[w_1 e^{+j\omega T} + w_2 + w_3 e^{-j\omega T}\right] E_i(j\omega)$$

(Time origin associated with center pulse)

Peak detector

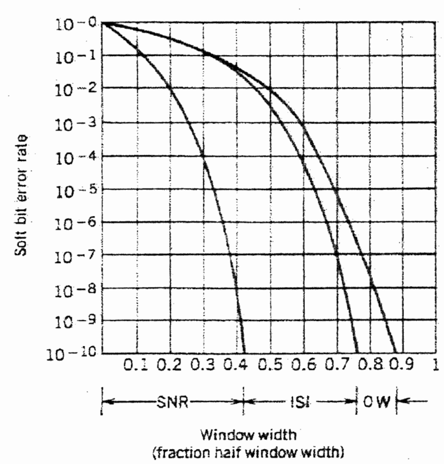


Note that amplitude thresholding is used and that pulse polarity information is ignored.

PD precompensation

- The goal in write precompensation for peak detection is to minimize or eliminate peak shift due to ISI so that the peak of the pulses will occur in the center of the clocking window, generated by a phase locked loop to implement self clocking.
- The dibit pattern is the worst as regards ISI where it was shown that the peaks spread apart as the pulses come closer together..
  - Writing the second transition of a dibit earlier can mitigate this effect and maintain the two peaks properly separated. A particular set of binary patterns provide the rules for shifting of current reversal times.
  - Since the peak amplitudes also decrease as density is increased the continuing distortion of the waveform with linear density puts a limit on the use of PD.

Margin testing with PD channels



Digital signal processing

- Advantages
  - Implementation of complex algorithms
  - Solutions based on adaptation
  - Testability
  - Reproducibility of performance
  - Control of accuracy
- Concerns
  - A/D conversion functions
  - Synchronization issues
  - Power-speed tradeoffs

Sampled data system basics

- Use amplitude samples to obtain more waveform information per transition than provided by PD.
  - Can cope with ISI, which is results from linear superposition
    - **Requires a linear channel**
- Use sequential detection (base decisions on previous samples as well as present one).
- Use maximum likelihood decision making based on samples.
- Use partial response choice which closely matches the actual magnetic recording channel, thereby allowing reasonable degree of equalization.

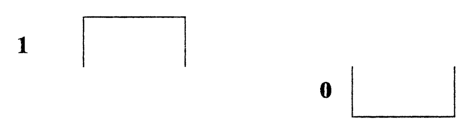
NRZI & NRZ

•Error propagation

•NRZI does not lead to error propagation since each bit is identified by the presence or absence of a output pulse unlike NRZ.

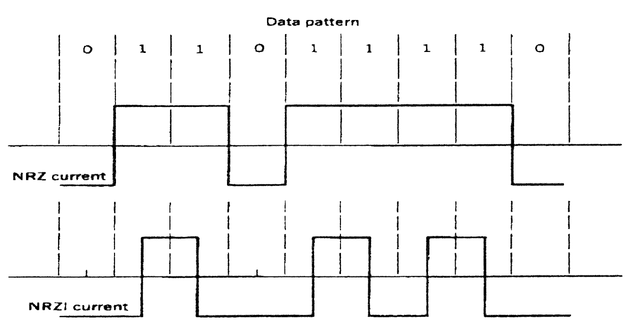
NRZ identifies the stored information by the state of magnetization and the polarity of an output pulse is uniquely identifies the direction of a change in the binary sequence. Thus, NRZ provides a three level signal that in a more fundamental manner characterizes the channel. Alternating pulses are inherent and their polarity is meaningful.

NRZ "1" and "0" current inputs can be defined as follows:



The pulse width is equal to the channel bit period. To preserve the advantages of NRZI and take advantage of the benefits of NRZ we can use a "precoder" to convert NRZI data to NRZ patterns.

NRZ and NRZI binary coding



NRZ signal waveform

Let $h(t)$ be the "characteristic" pulse response of the channel.

Then the output signal is:

$e(t) = \sum_k c_k h(t - kT)$ where T is the channel bit period.

If b_k is the NRZ bit sequence then

$c_k = b_k - b_{k-1}$

That is, a pulse of the appropriate polarity arises each time there is a change in the state of medium magnetization.

Let D represent the channel bit period delay operator.

Then $Db_k = b_{k-1}$

or $c_k = (1 - D)b_k$

In the frequency domain $D = e^{-j\omega T}$

Role of precoder

- The input data is NRZI which does not give rise to error propagation (call this bit sequence a_k).
- Precode the NRZI bit stream into NRZ, which provides a specific relationship between transition polarities and the direction of current reversals. (call the NRZ bit sequence b_k)
- Then the expected output samples, c_k, will relate each transition with the appropriate polarity. For PR4 these samples will be the same as the NRZI input sequence except polarity will be included.
- This additional information leads to sequence detection where the decision procedure will never recognize as valid sample sequences which do not reflect an alternating pulse train.

Precoding (NRZI to NRZ)

Let a_k = NRZI bit sequence, b_k = NRZ write current

Then $c_k = b_k - b_{k-1} = (1-D)b_k$ and is a three level signal (-1 , 0, +1)

Now since an NRZI "1" corresponds to a transition, then

$a_k = b_k \oplus b_{k-1} = (1 \oplus D)b_k$ For this case we can write $a_k \oplus b_{k-1} = b_k \oplus b_{k-1} \oplus b_{k-1}$

but as $b_{k-1} \oplus b_{k-1} = 0$, we get

$b_k = a_k \oplus b_{k-1}$ For PRML or interleaved NRZI $b_k = a_k \oplus b_{k-2}$

In general for a partial response channel given by $C(D)$, $b_k = \left(\frac{a_k}{C(D)} \right)$ mod 2.

We then find c_k provides NRZI type sequences but including polarity information.

Channel transfer function

The channel transfer function for a sampled system (NRZ) is :

$(1-D)H(D)G(D)$

where $(1-D)$ represents the fundamental differentiating character of the magnetic recording channel

$H(D)$ is the sampled isolated pulse response of the head

and

$G(D)$ is any channel equalization used

For the $(1-D)$ channel $H(D)G(D) = 1$ and $c_k = b_k - b_{k-1}$

Note that for a (1-D) channel there is no ISI at sample times, but the ratio of PW50/T is in the range of 1.0.

NRZI to NRZ translation

Let a_k = *NRZI* **data sequence**

Let b_k = *NRZ* **write current waveform corresponding to the NRZI data**

Then,

$$c_k = b_k - b_{k-1} = (1 - D)b_k$$

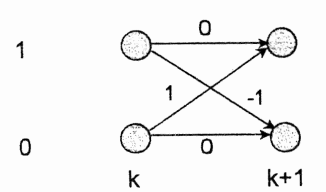
For c_k to correspond to the NRZI sequence, including polarity information

$$b_k = \left(\frac{a_k}{1 \oplus D}\right) \text{ or}$$

$b_k \oplus b_{k-1} = a_k$ **and therefore** $b_k = a_k \oplus b_{k-1}$

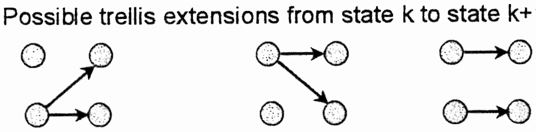
		k	k + 1	k + 2	k + 3	k + 4	k + 5	k + 6	k + 7
NRZI	a_k	1	0	1	1	0	0	0	1
NRZ	$b_k = a_k \oplus b_{k-1}$	1	1	0	1	1	1	1	0
	$c_k = b_k - b_{k-1}$ 1	1	0	-1	1	0	0	0	-1

NRZ state diagram and trellis extensions



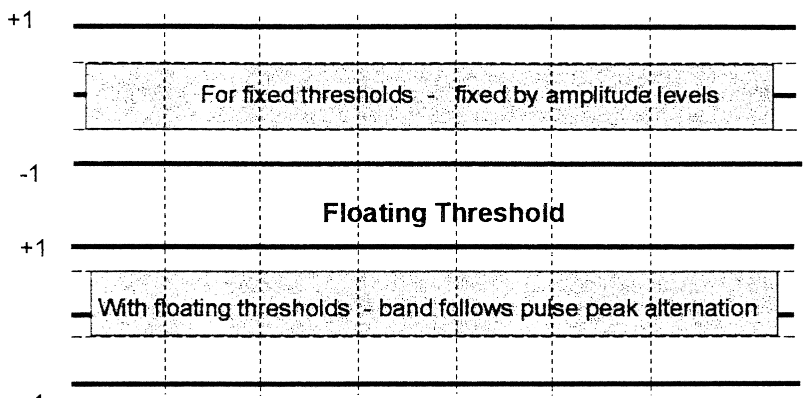
State diagram for a PRML interleave

Possible trellis extensions from state k to state k+1



Current state = 0 Current state = 1 Current state = unknown

Threshold detection (3 levels)

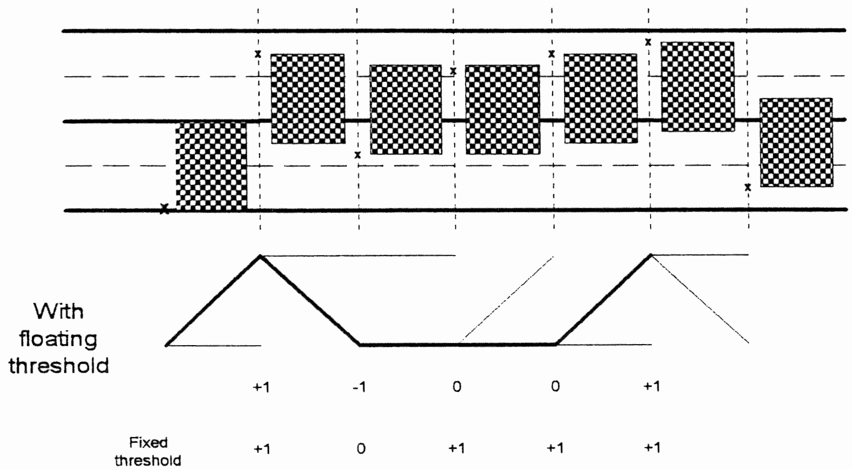


The floating threshold implements sequence detection based on the fundamental property that transition pulse polarities alternate.

Alternating transitions and floating threshold detection

- If a readback sample is above the threshold band (whose width is equal to that for a fixed threshold), this sample is provisionally considered to represent a positive transition and its value is used to reset the upper edge of the floating band.
- If the next sample is more positive the upper band edge is again adjusted upward to that level.
- If the following sample falls within the band it is not considered to arise from a transition and the threshold band location remains the same.
- If a following sample is sufficiently negative to be below the band, this sample is considered to represent a negative transition and the lower band edge is adjusted downward to this value. This action confirms the earlier positive sample as actually representing a positive transition.

Sequence detection



Advantages of floating threshold

- The threshold band automatically determines the most likely sequence of bits.
- It rejects any bit sequences which violate the inherent requirement that transitions generate pulses of alternating polarity.
- It chooses the most likely of the two acceptable sample sequences: I.e., 0 0 versus +1 -1
  - The floating threshold band width represents will choose the most likely sequence based on the probability of random noise corrupting the samples.

Error events - floating threshold

- **An error will occur when the sum of the noise associated with two successive samples is greater than 1, where the three normalized target levels are 1, 0, and -1.**
- **The sample sequence "peaking finding" method assures the decoded signals will not violate the fundamental criterion that a correct output pulse sequence must alternate.**
- **A reasonable SNR is assumed such that there will never be a noise input that will corrupt a valid pulse into one of the opposite polarity.**

PRML

- Uses partial response equalization which is compatible with ISI and allows more samples per transition to be used.
- Takes advantage of alternating polarity of transition pulses to provide sequential detection to improve decision making.
- Assumes a linear channel, that is that noise is not a function of the input signal.
- To the degree linearity and noise do not meet these requirements the main advantages are through DSP and data rate.

PR4 channel (1)

- (1-D)(1+D) = 1- D^2 channel. n = 1
  - H(D)G(D) = 1 + D
    - the ideal samples from a transition are 0 1 1 0
    - An NRZ "1" or the dibit response is 0 1 0 -1 0
      - That is, $(1-D)(b_k + b_{k-1}) = b_k - b_{k-2}$
- From before, $c_k = (1 - D)[H(D)G(D)]b_k = b_k - b_{k-2}$
- Thus, since the even and odd samples are mutually exclusive sets the channel can be viewed as **interleaved** NRZI.
- The precoding step generates a write current which generates samples in each of the even and odd bit streams which are three level and therefore carry pulse amplitude and polarity information similar to the situation with the simple NRZ channel.

PR4 channel (2)

In the frequency domain

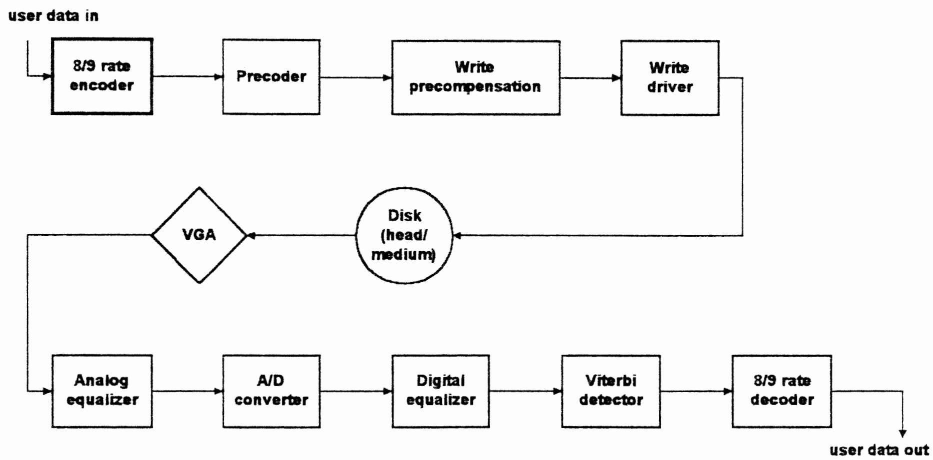
$$1 - D^2 = 1 - e^{-2j\omega T} = 2je^{-j\omega T}(\sin \omega T)$$

Thus the frequency function is a sine wave with

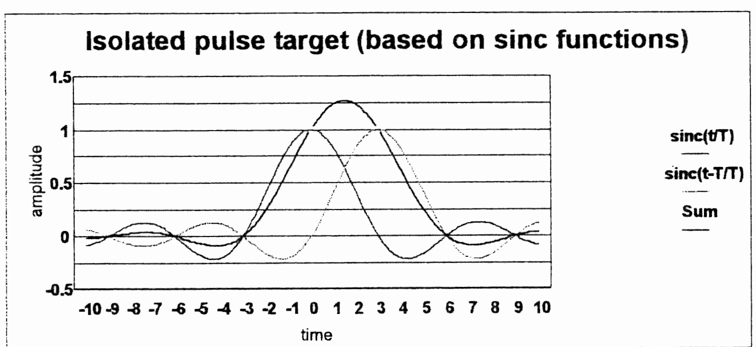
a zero at dc and at $f = \frac{1}{2T}$, **the Nyquist frequency. (The sampling rate is 1 / T)**

PR4 consists of two interleaved sequences since the odd and even samples are independent; the output sample c_k depends on only one previous magnetic state and with PR4 PW50/T is near 2. The detector can operate on each of the interleaved sequences concurrently at half the channel data rate.

PRML channel block diagram

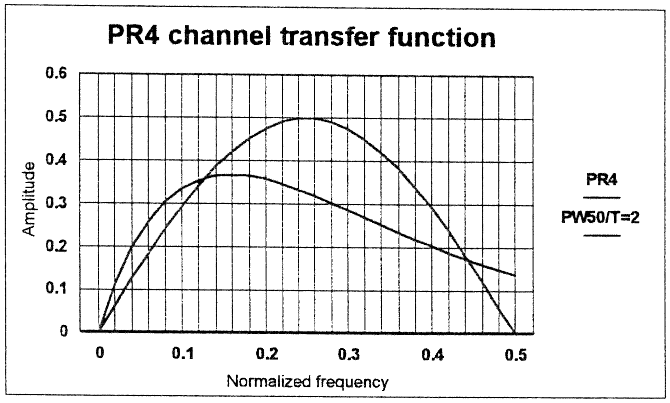


Target pulse responses



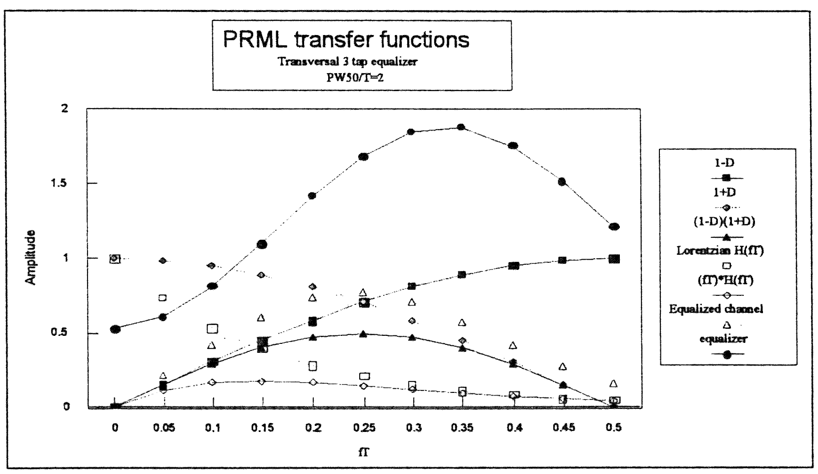
T = 3 (for PRML). Time origin is peak of first snc pulse

PR4 equalization

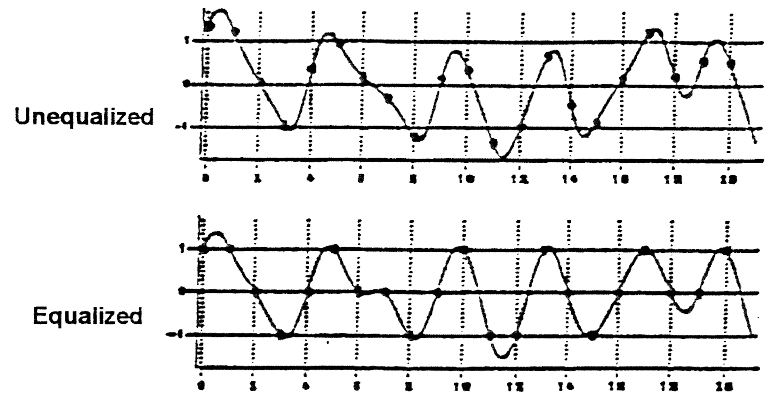


Channel equalization (analog and digital) modify the unequalized channel response function to the PR4 target function shown.

PR4 channel transfer blocks



Example of PRML equalization



Sampling of unequalized compared to equalized waveforms

PR4 summary

The PR4 channel function is

$$1 - D^2 = 1 - e^{-2j\omega T} = e^{-j\omega T}\left(e^{j\omega T} - e^{-j\omega T}\right) = 2je^{-j\omega T}(\sin\omega T)$$

Since $D^2 = 2D$ the data stream consists of two interleaved sequences with polynomial (1 - 2D).

The NRZI input, precoded NRZ, and ideal sample sequence is shown in the table below.

	k	k+1	k+2	k+3	k+4	k+5	k+6	k+7	k+8
a_k	1	1	0	0	1	1	0	1	0
$b_k = a_k \oplus b_{k-2}$	1	1	1	1	0	0	0	1	0
$c_k = b_k - b_{k-2}$	1	1	0	0	-1	-1	0	1	0

Note that for both the even and odd sequences the target three level outputs have alternating sample pulse outputs and provide a positive or negative value for every transition.

Viterbi detector for PRML

- The interleaved nature of PR4 leads to a relatively simple detector and it **only need be understood** in terms of either the even or odd samples.
  - The maximum likelihood procedure incorporates, through the trellis NRZ state diagram, the fact that output pulses must alternate in polarity, Thus, it is not possible to have two successive positive sample pulses that could be interpreted as two NRZI "1" signals. A second positive pulse requires that first the detector switch its state.
- The most common code is 8/9(0,4,4). The value of 4 for the maximum run of zero's truncates the path length the detector must handle for each interleave and provides suitable self-clocking
- The detector is based on a reasonably good SNR, thus only three extensions for state transitions in the Trellis are allowed.

ML detection

- A trellis diagram is used to track the possible state paths, defined by state changes with time. With NRZ the states represent the two possible directions of saturation magnetization in the medium. State changes are represented by line segments between successive states. In the absence of noise the possible inputs, for example, are: 0 (no change) +1 (state 0 to state 1) and -1 (state 1 to state 0). I practice the input readback samples are used to trace paths while keeping track of their likelihood of occurrence. That is, given the output samples what is the most likely sequence of binary data that was recorded at the input.
- Whenever two paths merge the most likely is continued and the other terminated. This procedure leads to a process which provides an output binary sequence that "most likely" matches the input.

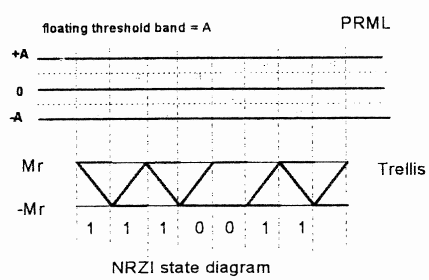
Viterbi detector

NRZI input	1	1	1	1	1	1	1
readback samples	+1	-1	+1	-1	+1	-1	+1
Only acceptable sequences			+1	-1	+1		
			0	0	+1		
			0	+1	0		
			+1	0	0		

A choice of any acceptable sequence, other than the correct one requires that two samples be in error.

Sample sequences showing bit error types

For an all ones pattern it can be seen from the sample sequence below that a single bit error can be corrected but a two bit error pattern cannot. However, given the sample sequence the detector certainly chose the most likely bit sequence.



Error events and PRML code

The predominate PRML error event can be defined by the following condition:

$n_1 + n_2 > 1$ **, where the ideal sample values are +1, 0 and -1.**

Here, the noise makes the input target sequence a less likely path than another acceptable sequence.

PRML 8/9(0,4,4) code

- High rate code, interleaved - advantageous in terms of channel electronics
- k = 4 for both even, odd and global sequences
  - Forces merging and decisions by Viterbi detector
  - Assures clock timing
- d = 0 as ISI handled by partial response

PRML and error rate

For simple threshold detection of pulses (magnitude 1) the distance between levels is 1 and the threshold setting for detection would be 1/2. Then, for a distance of 1 between detection states we have

$ber \ \alpha \ Q\left(\frac{1}{2\sigma}\right)$ where σ is the standard deviation of the noise.

For PRML the minimum error distance between states distant 1 is

$\sqrt{1^2+1^2}=\sqrt{2}$

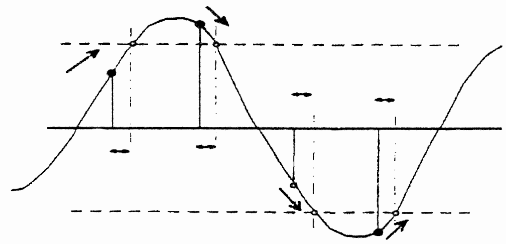
Therefore $ber \ \alpha \ Q\left(\frac{\sqrt{2}}{2\sigma}\right)$

ML detection provides the equivalent of a $\sqrt{2}$

increase in the noise threshold for a PR4 system or a 3dB effective increase in SNR. This result can also be viewed as a consequence of the fact that an error event in PRML requires that there be two bits in error.

Phase and gain recovery

PRML, being a sampled amplitude detection system, requires accurate control of the sampling clock times as well as gain control since the decision making procedure is dependent on the relative amplitudes of the samples.



Phase update direction >>>

The change between successive sample values, illustrated, provides the information needed for phase correction.

PRML vs PD

For low error rates the probability of a single bit in error in terms of signal to noise ratio for a long binary sequence can be approximated as:

$$p(1) \propto e^{-(SNR_1)^2}$$

While for two bits in error

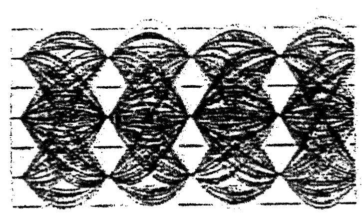
$$p(2) \propto p(1)^2 \propto e^{-(\sqrt{2}SNR_1)^2}$$

Thus, PRML on this basis offers a gain of 3 dB in SNR. This gain can be traded for improved reliability or increased density.

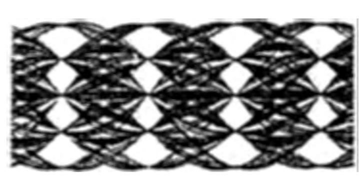
Extended partial response systems

- Pulse shapes
  - Well matched to recording channel characteristics
  - Minimum bandwidth signals, well-defined
  - Transition response samples are binomial coefficients:
    - n = 0   1   (Dicode)
    - n = 1   1 1   (PR4)
    - n =2   1 2 1   (EPR4)
    - n = 3   1 3 3 1   (E^2PR4)
- Veterbi detection
  - Trellis states: 2^{n+1}
  - Two likelihood functions per state
  - Multiplier-free soft-decision implementation

Eye diagrams



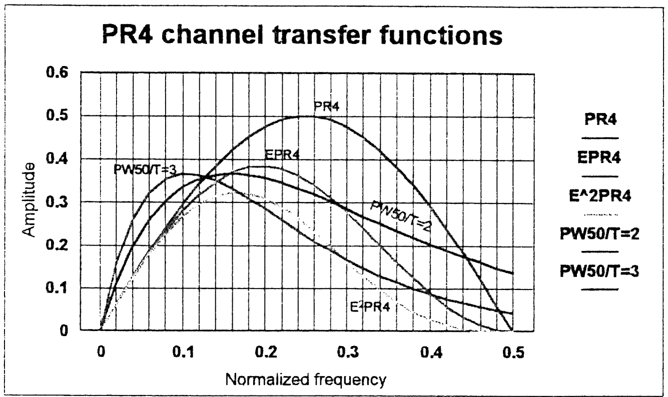
Eye diagram for PR4 Channel: 3 output levels



Eye diagram for EPR4 channel: 5 output levels

ASH Magnetic Disk drive Technology IIST

Transfer functions
PR4 extensions



EPR4

$H(D)G(D) = (1+D)^2 = 1+2D+D^2$

$C(D) = (1-D)(1+2D+D^2) = 1 + D - D^2 - D^3$

or $c_k = b_k + b_{k-1} - b_{k-2} - b_{k-3}$

the ideal transition sample sequence 0 1 2 1 0 0

the ideal dibit sample sequence is 0 1 1 -1 -1 0

Since each sample depends on the current direction of the magnetization plus the directions of the three previous periods of magnetization the EPR4 state diagram will have 8 levels. Each channel bit period the current state can change to a new state along the two paths defined by the direction of the upcoming magnetization.

Trellis for EPR4 where maximum sample mormalized to 1.0.

EPR4 polynomial

$$H(D)G(D) = (1+D)^2 = 1 + 2D + D^2$$
$$C(D) = (1-D)(1+D-D^2) = 1 + D - D^2 - D^3$$
or $$c_k = b_k + b_{k-1} - b_{k-2} - b_{k-3}$$

Since each sample depends on the current state of magnetization plus the three previous states of magnetization, the EPR4 state diagram will have 8 levels

Transition samples	0	1	2	1	0	0
Dipulse samples	0	1	1	-1	-1	0

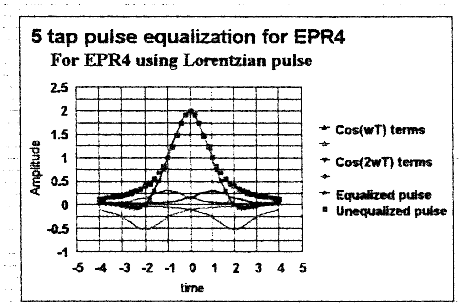
EPR4 channel equalization

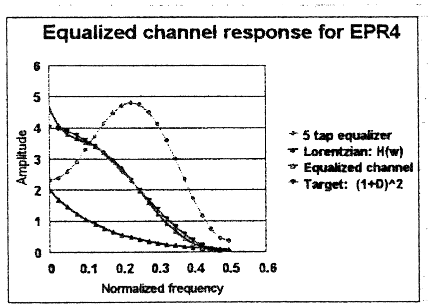
Filter design for EPR4

Matrix solution for tap weights

PW50/T = 2	2	1	3.4	0.2	0.12	0
PW50 = 2	1	2	1	0.4	0.2	1
T = 1	0.4	1	2	1	0.4	2
PW50/2 = 1	0.2	0.4	1	2	1	1
	0.12	0.2	2.4	1	2	0

0.6705375	-0.359033	0.0570342	-0.021195	-0.005138	-0.26616
-0.359033	0.8627389	-0.389734	0.06882	-0.021195	0.152091
0.0570342	-0.389734	0.8669202	-0.389734	0.0570342	0.954373
-0.021195	0.06882	-0.389734	0.8627389	-0.359033	0.152091
-0.005138	-0.021195	0.0570342	-0.359033	0.6705375	-0.26616





EPR4 (precoder)

$$C(D) = (1-D)(1+D)^2 = 1 + D - D^2 - D^3$$

$$or \quad c_k = b_k + b_{k-1} - b_{k-2} - b_{k-3}$$

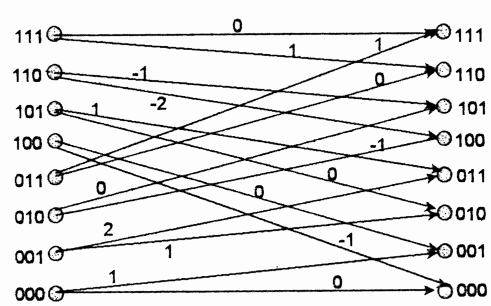
$$and \quad b_k = \left(\frac{1}{1 \oplus D \oplus D^2 \oplus D^3}\right) a_k \qquad (\text{mod } 2)$$

$$Then, \quad b_k = a_k \oplus b_{k-1} \oplus b_{k-2} \oplus b_{k-3}$$

a_k	0	0	1	1	0	0	0	0	1	1	1	0	0	0	1	0	1	0	0
b_k	0	0	1	0	1	0	1	0	0	0	1	1	0	0	0	0	1	1	0
c_k	0	0	1	1	0	0	0	0	-1	-1	1	2	0	-2	-1	0	1	2	0

Note that $c_k = a_k \pmod 2$

Trellis for EPR4



Each channel bit period the current state can change to a new state along the two paths defined by the two options for the next state of magnetization

As with PRML the bit sequence is determined by the surviving path based on maximum likelihood.

Summary

- Track width and SNR will decrease
- Efficient signaling allows intersymbol interference
- Detection by maximum-likelihood sequence estimation
- Write precompensation alleviates nonlinear transition shifts
- Adaptive equalization improves performance when conditions vary
- Trellis codes allow reliable detection even with low SNR

Magnetic Disk Drive Technology

Storage Density

Storage Density outline

- Head positioning
  - Servo methods
  - Track seeking
  - Track following
- Track density
  - Head mis-registration
    - Side reading and writing
  - Old information and adjacent track interference
    - 747 curve
- Magnetic data recording and ECC
  - RAID systems

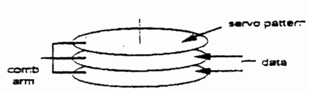
Servo system

- Head positioning involves:
  - Track seeking system
    - minimize time to seek and settle on track
  - Track following system
    - cope with non-repeatable runout, mechanical vibration & shock, etc.
- The servo system is an Important element in determining
  - Seek time
  - track density
  - overall reliability

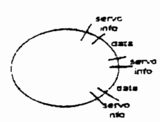
Head-medium track following

Head Positioning

Open Loop
- Low Cost

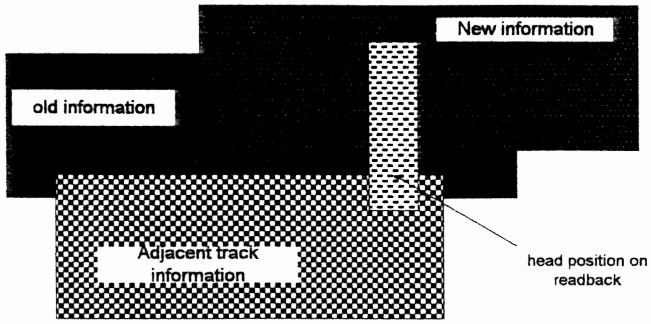


Servo
- Dedicated
  - Issue - cascaded mechanical & thermal tolerances
- Embedded
  - » Sector
    - Issues - Open loop when R/W
    - real estate
    - slew rate
    - tolerance to disk defects



  - » Continuous
    - Issues - Source of PES (Position Error Signal)
    - disk fabrication
    - sensor
    - servo while write

Direct overwrite

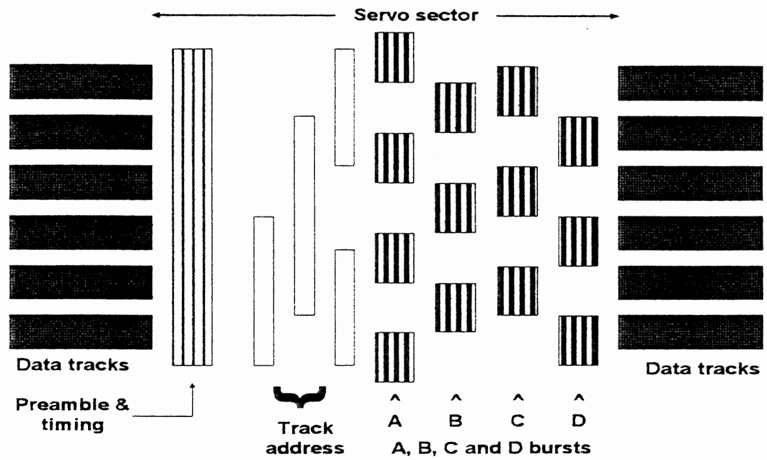


The requirement that individual data blocks can be overwritten with new data leads to signal distortion and noise in the desired signal from the concurrent sensing of the previously written data and adjacent track information, arising from tracking tolerances.

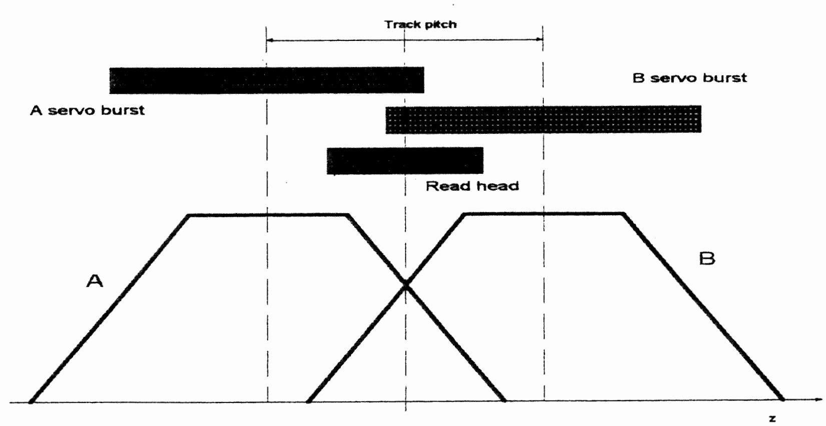
Servo pattern writing

- Need accurate radial position increments of $T_p/2$ (T_p=track pitch) with precise track phase coherence so that the sections of the radial transitions written on successive passes are read as single transition.
  - Use laser interferometer (through slot in cover) with retro-reflector temporarily mounted on the drive head suspension block and a clock head, also temporarily inserted through port in side of HDA, flying at the extreme outer diameter of the disk to provide timing information.
- Sector track address encoded in Gray code.

Sector servo
Normal and quadrature servo

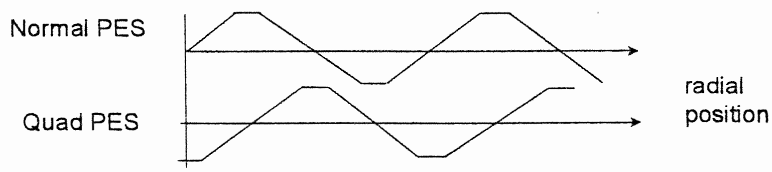


Servo PES signal generation



$$PES = \frac{A - B}{[illegible]}$$

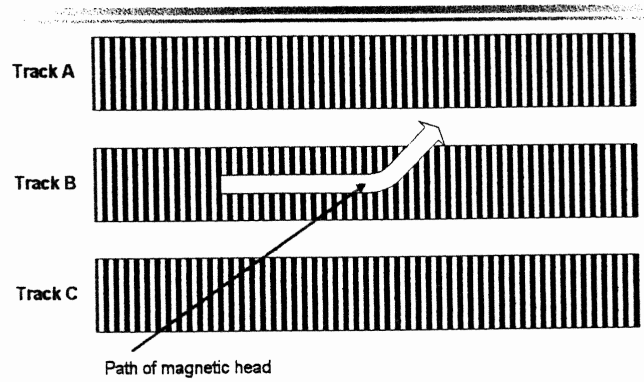
Normal and quad PES



While the A and B bursts define the track location there is no way from their signals to accurately determine the head position when it is the flat portion of the transfer function (about half way between tracks). To resolve this situation two additional bursts are added (C and D) offset by a radial displacement of 1/2 track pitch.

Since the C-D PES is shifted 90 degrees from that of the A-B (in terms of radial position it is usually referred to as the quadrature PES. By using both outputs continuous position data with continuous derivatives for velocity determination are always available for track seeking.

Sensor for dealing with shock

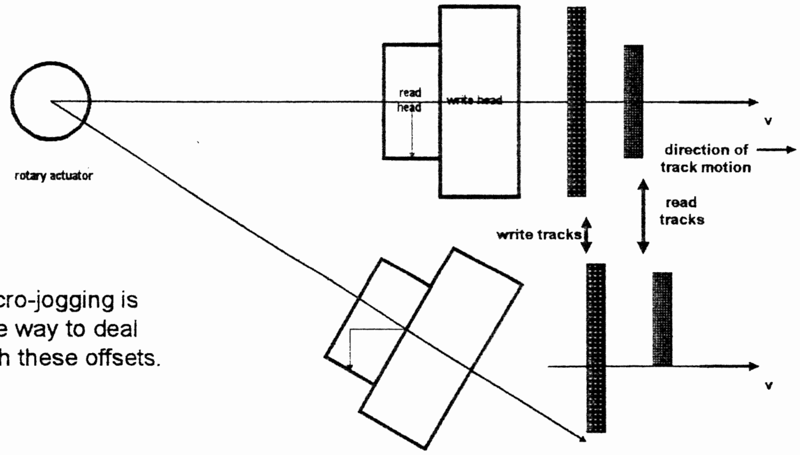


If head goes off track during write then data on A nonrecoverable.

If during read operation initiate re-read.

Solution: Accelerometer to shut off write gate before head off-track.

Skew from swing arm actuator

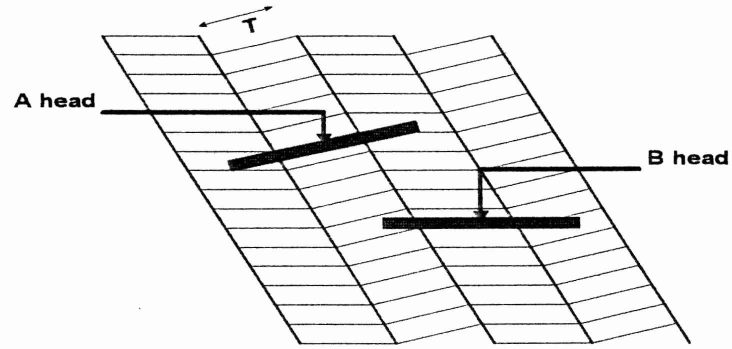


Micro-jogging is one way to deal with these offsets.

Helical scan recording

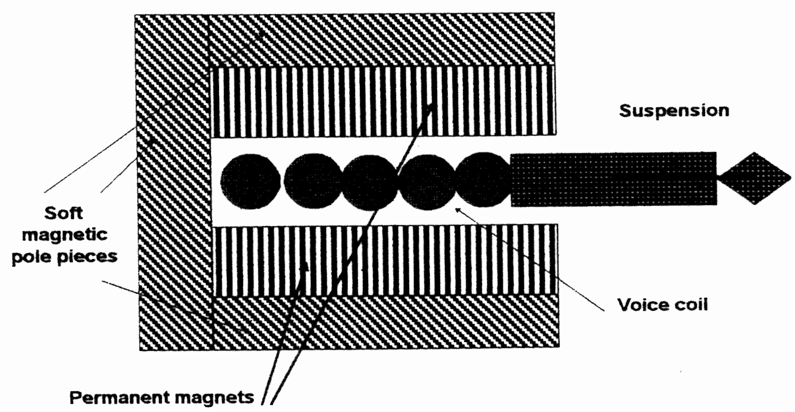
- Adjacent tracks recorded with alternate azimuth angles
- No guard band allows greater areal density
- Requires one pass writing
- Write heads can be wider than track width to allow larger tracking error
- Easily implemented with 2 or 4 head system

Adjacent track skew



A and B head track widths less than 3T

Voice coil actuator



F = Bli

Track seeking

A minimum time seek is based on using maximum available power (a so-called "bang-bang" servo system). Here, a single switch point is used to go from maximum acceleration to maximum deceleration. (In practice longer seeks will be velocity limited.) For a second order system (i.e., considering mass only) a square root trajectory is obtained.

$$F = M\frac{d^2x}{dt^2}$$

and $t = \sqrt{2\left(\frac{M}{F}\right)x}$

This latter expression gives the relationship of seek time to seek distance

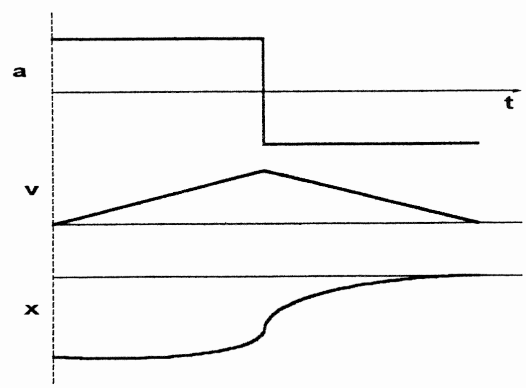
Track seeking system

a = acceleration

v = velocity

x = position

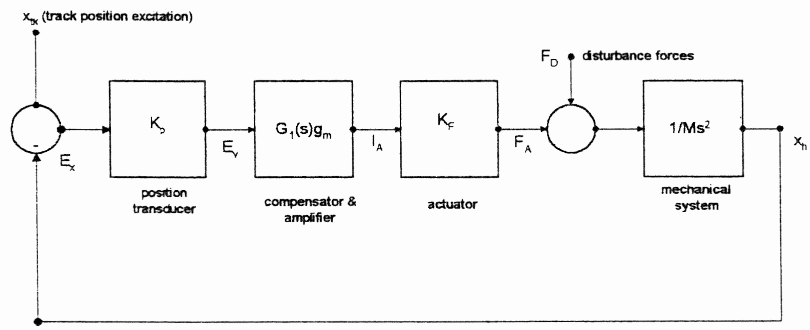
t = time



Bang - Bang positioning characteristic.

Based on time-optimal response for given power available

Track following servo



Overall SNR for channel

S = desired signal from track being read

Noise sources

σ^{d} = rms disk noise

σ^{h} = rms head noise

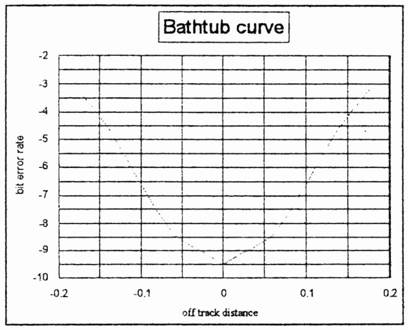
σ^{a} = rms amplifier noise

σ^{ol} = rms noise from old information

σ_{adj} = rms noise from adjacent track

$$SNR = \frac{S}{\sqrt{\left(\sigma_d^2 + \sigma_h^2 + \sigma_a^2 + \sigma_{ol}^2 + \sigma_{adj}^2\right)}}$$

Bathtub curve

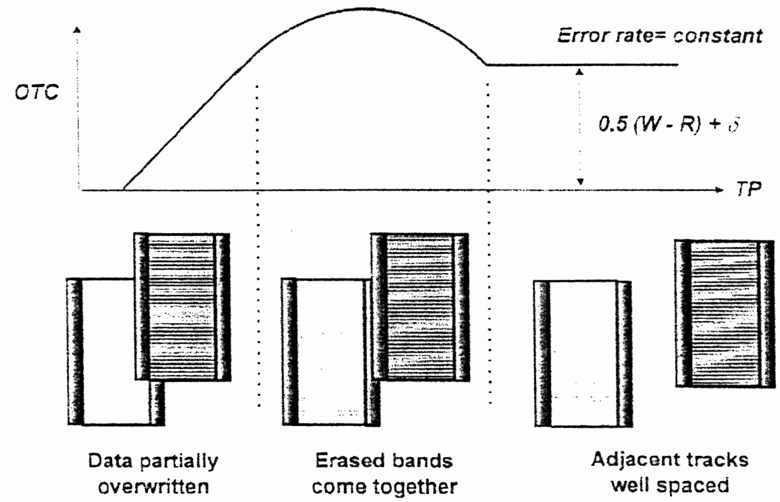


Gives raw error rate as a function of read head track offset.

747 Curve

- The 747 curve shows there can be an optimum track pitch for a given head width and specified bit error rate in terms of maximizing head offset.
- The reason the allowable head offset evidences a "hump" as the effective track density in increased can be explained as a combination of two factors:
  - Initially the side erase bands of the head lead to a reduction in "old" information as the tracks come into proximity and become closer.
  - When the track "squeeze" reaches the stage where the head senses a combination of 'old" and adjacent track information, since these two sources are not highly correlated, their total "noise" contribution is less than from the same head width positioned over either one alone.
    - side erase bands are advantageous in terms of track density!

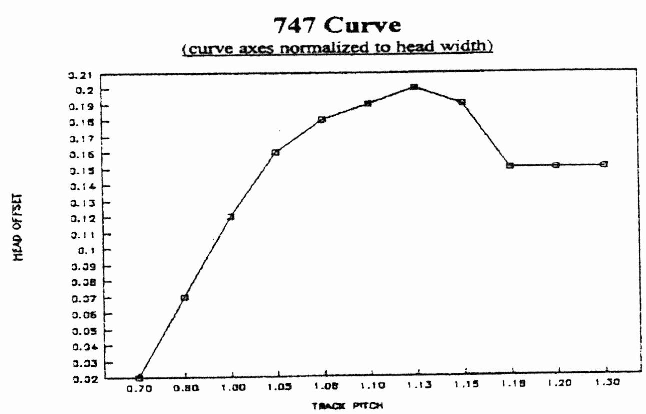
Main Features of 747 Curve



R. S. Beach
Advanced Concepts
San Jose

Seagate
THE DATA TECHNOLOGY COMPANY

747 curve showing % off track



Error detection and correction

- Error correcting codes permit the accurate reconstruction of data that is corrupted due to noise and media defects.
  - Encoding
    - Adding redundancy by appending parity symbols. Correction power is determined by the level of redundancy added.
  - Decoding
    - Generating syndromes identifying error locations and correction values.
  - Capability
    - If the errors are within the correction capabilities of the code the data can always be reconstructed,
    - If an error event exceeds the correction capabilities but not the detection guarantees, the data can always be flagged as unusable.
    - If an error event exceeds the correction and detection capabilities the error would be undetected by the code. The probability of failing to detect errors can be reduced to an arbitrarily low level with the addition of sufficient redundancy.

Requirements on EDAC

- BER requirements (today) before EDAC
    - raw error rate
  - On-track BER < 1E-9
  - Recording channel BER < 1E-6
- Disk drive data integrity requirements
  - Frequency of reread due to error < 1E-10
  - Frequency of unrecoverable data due to error < 1E-15
  - Frequency of not detecting corrupted data < 1E-21

Error types

- Error in word : unknowns, error location and error value
- Hard or permanent errors
  - disk defects
    - On surface allow for additional tracks
    - within tracks allow for additional sector
    - map out defective regions
- Soft errors
  - Non repeatable as arise from noise
- Burst errors
  - Group of bits taken as all in error
- Erasures
  - One or more bits in a group which cannot be sensed
    - example: dropouts

Value of parity or check bits

Let p be the probability of a bit in error. For a codeword of length n:

$$p_r \approx np(1-p)^{n-1}$$

p_r is probability of the codeword in error. If the codeword includes single error correction then

$$p_r \approx \frac{n(n-1)}{2} p^2 (1-p)^{n-2}$$

Example. For Hamming single error correction code for 4 message bits needs 3 parity bits. For p = 1E-5

$$p_r = 7 \cdot 10^{-5} \quad \text{without ECC}$$

$$p_r = 21 \cdot 10^{-10} \quad \text{with single error correction}$$

For short codewords clearly the code is inefficient since the level of redundancy is quite high.

Error correction codes

A codeword more generally can consist of symbols, bits, bytes, etc.

The Reed-Solomon codes operate on symbols. A typical symbol size is one byte or eight bits.

The number of possible codewords using n symbols is $\mathbf{2^{ns}}$

For k message symbols the number of codewords is $\mathbf{2^{ks}}$

Thus, the fraction $\mathbf{2^{(k-n)s}}$ of the possible words are codewords.

On-the-fly correction capabilities implies the error decoding process can proceed essentially in real time.

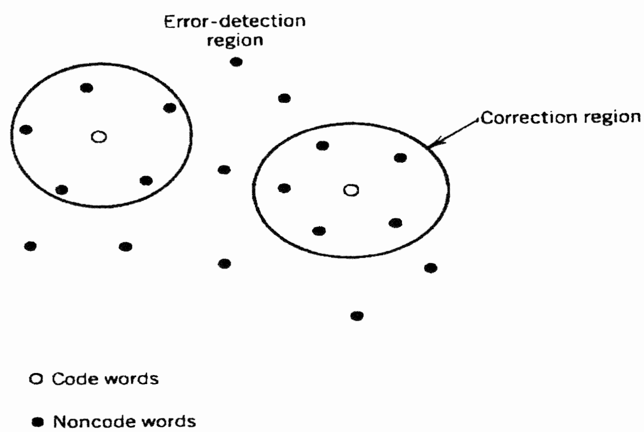
Block codes

- Code length (n) - number of bits in word (all words are of same length and a subset are codewords, i.e., identified with messages)
- Information or message bits (k) - k < n
- Parity check bits - (n-k)
- Code rate of efficiency (R) - k/n (R < 1)
- Systematic code - block code in which parity bits are appended to message bits.
- Encoding - establishing a <u>codeword</u> for a k bit message
- Decoding - associating a k bit message with a specific codeword
- Nearest neighbor decoding
  - Correctable words are associated with a given codeword on the basis of a nearest neighbor relationship, i.e., errors are rare.

Hamming code

- A word of length n can be viewed as an n dimensional vector.
- Hamming weight: number of non-zero components in a binary word
- Hamming distance: number of bits in which two word vectors differ.
- The minimum Hamming distance of a code is the minimum distance between the two closest codewords.
  - A code with a minimum distance d_{min} can correct any error pattern of $(d_{min} -1)/2$ or fewer errors in a codeword.

Spatial view of codewords



Single error correction

n = word length

k = number of bits per message

n-k = number of check bits

The n-k parity check bits have to indicate n+1 conditions:

n possible error locations plus the error free status

Thus, $2^{(n-k)} = n+1$

n-k	n	k
2	3	1
3	7	4
4	15	11

Encoding a 4 bit message and locating error

Encode

1	2	3	4	5	6	7	**Bit position**
-	-	1	-	0	1	1	**Message**
0	1	1	0	0	1	1	**Encoded**
		x					**Error**
0	1	0	0	0	1	1	**Received**

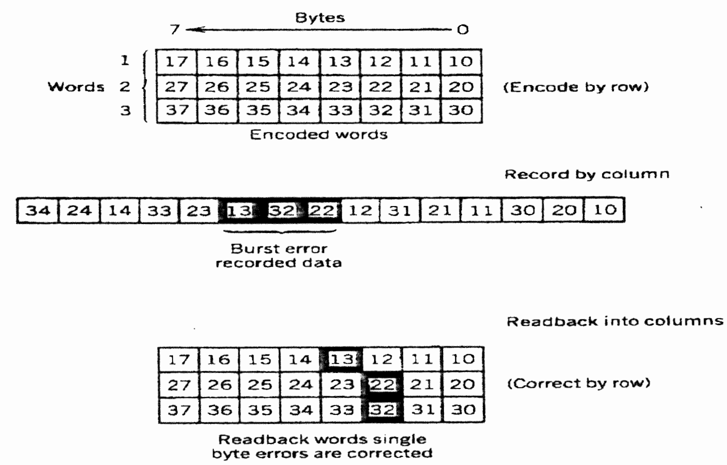
Locate error

Check 1	1	3	5	7		**record**
	0	0	0	1	**Fails**	1
Check 2	2	3	6	7		
	1	0	1	1	**Fails**	1
Check 3	4	5	6	7		
	0	0	1	1	**Correct**	0
Syndrome	0	1	1		**error position**	3

Correction - add error vector to received word

Error vector	0	0	1	0	0	0	0
Received word	0	1	0	0	0	1	1
Word after correc	0	1	1	0	0	1	1

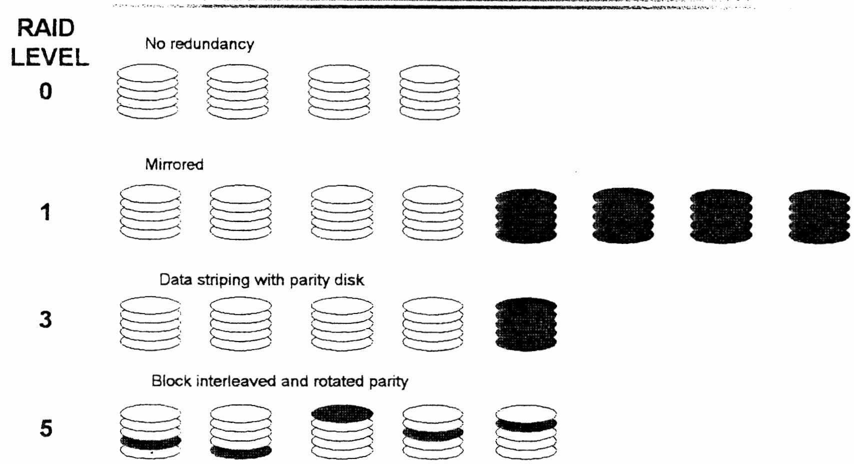
Interleaving for burst error correction



RAID storage systems Definitions

- Disk farm - assembly of attached disk drives
- RAID - Redundant array of independent disks
- Disk array - Storage subsystem using RAID technology
- Fault tolerant storage system
  - system that continues service after storage system controller, power or drive failure
- High Availability
  - continues service after drive failure

RAID levels - disk drive usage



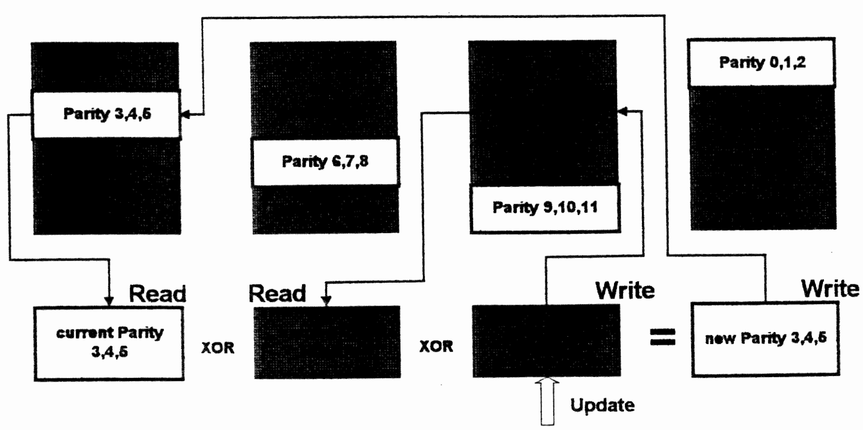
Raid classifications

- 0. Data striped over disks - no redundancy. Parallel data transfer
- 1. Mirrored disks
- 2. Dedicated parity disk
- 3. Dedicated parity disk with word or byte-wide data striping. Provides parallel data transfer
- 4. Dedicated parity disk with block data striping
- 5. Data plus parity, striped across all disks in array

RAID level comparisons

- RAID 1
  - Requires twice the storage capacity
    - can read concurrently from both drives
    - need to write same records on each drive
- RAID 3
  - Can continue at same performance level even if one disk fails. Attractive for parallel transfer
- RAID 5
  - Takes four disk operations for one write: reading old data and parity then writing new data and parity
  - Allows multiple read and write requests

Write operation, Raid 5



RAID features

- Level 1.
  - High I/O rate for small data blocks
- Level 3.
  - High bandwidth for large data files
    - minimal cost in disk drives
- Level 5.
  - High I/O rate for small data blocks

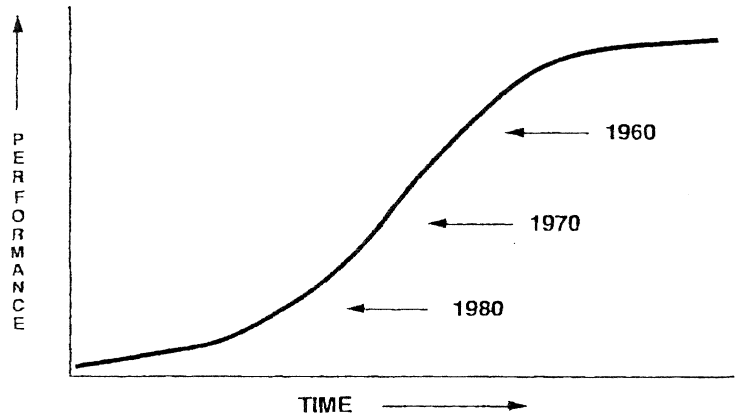
Magnetic Disk Drive Technology

Status and Future Trends

Storage Technology
Status and Future Trends

- Disk Drives
  - Current status and technological opportunities
  - Ultimate limits
  - Future trends
- Alternate Storage Technologies
  - Optical storage
  - Flash memory
  - Removable storage devices and applications
- Outlook and data storage industry perspectives

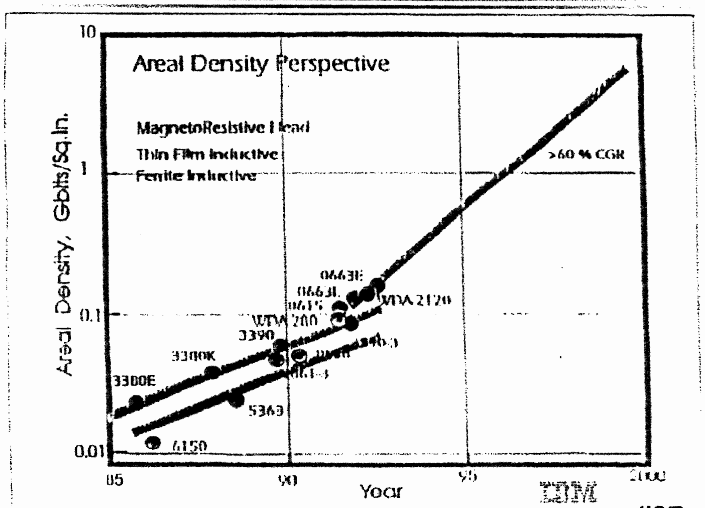
Changing perceptions with time



Data storage demand

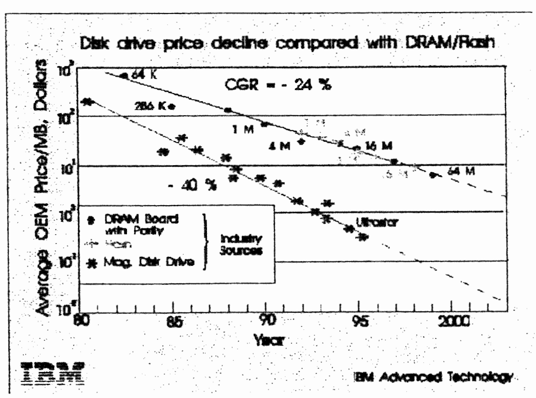
- While magnetic disk storage technology is increasing at a 60% CGR, the growth in data storage is increasing at a 100% CGR.
- Thus, while drive shipments are increasing dramatically, the opportunities for alternative technologies have never been better.
  - Removable storage in all forms represents a rapidly growing market.
- The Internet, multimedia applications, digital cameras, HDTV, home theaters, etc., all are placing huge demands on the realization of very low cost storage.
- For example -
  - 2 hrs of HDTV = 1.2 terabytes of data at 160 MB/sec (uncompressed)
    - *Clearly compression algorithms, ECC and recording codes will become of more relevance to the design of storage devices.*

Historic projection - 60% CGR



IIST

Cost of memory and storage



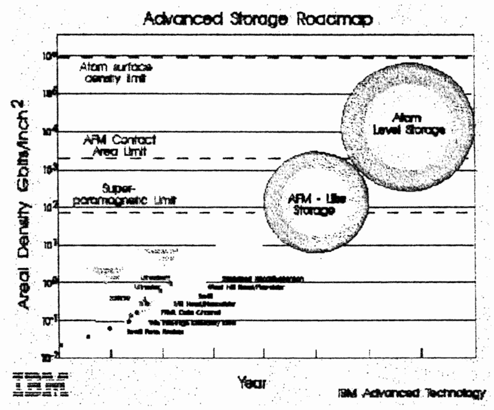
Technological opportunities ? (1)

- Heads
  - GMR heads
    - spin valve
    - multi-film GMR devices
  - Magnetic heads with integrated electronics
  - Parallel head structures for disk drives
- "Contact" recording
  - In range of 10 nm effective spacing
  - Glass or other super smooth substrates
- Magnetic media
  - Coercivities 3000 oe and above
  - $M_r t$ less than 0.4 milli-emu/cm^2
  - Patterned servo information
    - Patterned bit cells

Technological potential ? (2)

- Multi-stage actuators for fast, high precision head positioning
  - MEMS technology
  - Multiple actuators per disk for increased thruput
- One chip implementation of all disk drive electronics
- Multi-element read sensors to deal with track misregistration and adjacent track interference
- Fluid and air bearing spindles
- One inch disk drives that mount on a motherboard like memory chips
  - Use of RAID concepts with an array of small disks in personal computers
- On drive "processors" to allow local disk management of data access

Physical limits



ASH | Magnetic Disk Drive Technology | IIST

Ultimate magnetic recording limits

- Superparamagnetism
  - Sets smallest magnetic particle size that can still retain a preset magnetic state at room temperature for an extended period of time.
  - This limit has been approximately placed in the range of 5nm.
  - Given a one particle thick medium the above leads to a limit less than 100 gigabits per square inch, assuming 100 particles per bit cell will give an adequate SNR. (excludes notion of individually fabricated bit cells)
- Today limiting factors are related to spacing and tracking with techniques that offer high data rates.
- As long as state of the art is a factor on 10 or more from the ultimate limit magnetic disk storage will remain a moving target that will not be seriously challenged in its traditional role

ASH | Magnetic Disk Drive Technology | IIST

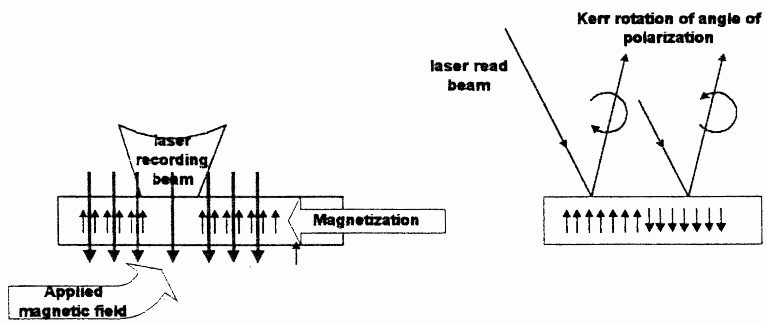
CGR analysis of density trends

PRODUCT	RAMAC	IBM 1301	IBM 1311	IBM 2314	IBM 3330	IBM 3350	IBM 3380	INDUSTRY	**AT 37%**	**AT 60%**
YR SHIPPED	1957	1962	1963	1966	1971	1976	19811	1991	**2001**	**2001**
MAX STORAGE DENSITY	2.0E+03	2.6E+04	5.1E+04	2.2E+05	7.8E+05	3.1E+06	1.3E+07	1.3E+08	**3.0E+09**	**1.4E+10**
STORAGE DENSITY CGR		67%	96%	63%	29%	32%	32%	26%		
BPI	100	520	1,025	2,200	4,040	6,425	15,200	48,000	**270,000**	**4.10E+05**
TPI	20	50	50	100	192	478	820	2,700	**11,000**	**2.50E+04**
	STORAGE DENSIT		**LINEAR DENSITY**		**TRACK DENSITY**					
AVERAGE CGR TO 1991	**37%**		**19%**		**15%**					
PROJECTED 1991 TO 2001	**60%**		**24%**		**25%**					
ASSUMING 60% CGR										

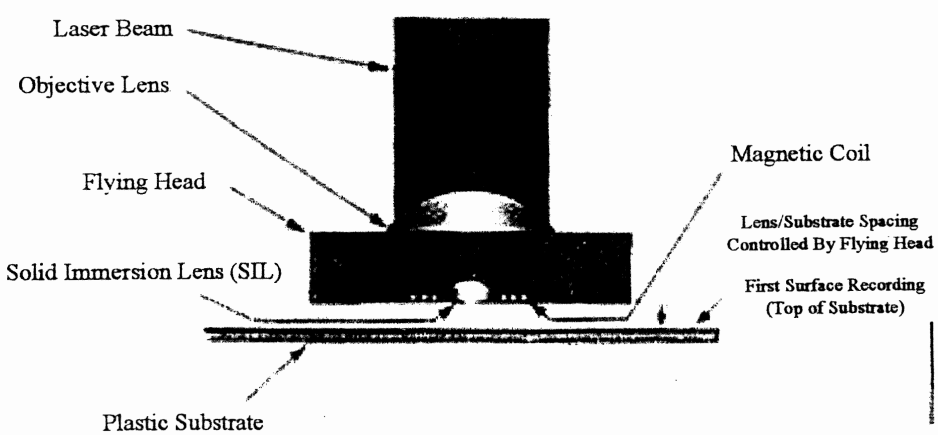
Optical storage attributes

- Optical disk
  - high capacity in small form factor
  - removable storage media
  - direct access
- Multi-function
  - Read only - stamped
  - Write recordable and read/write becoming available
- CD ROM now accepted for program distribution and multi-media
- Removable disk feature attractive for higher performance automated libraries
- Convergence for entertainment and computer needs make DVD a unique product opportunity.

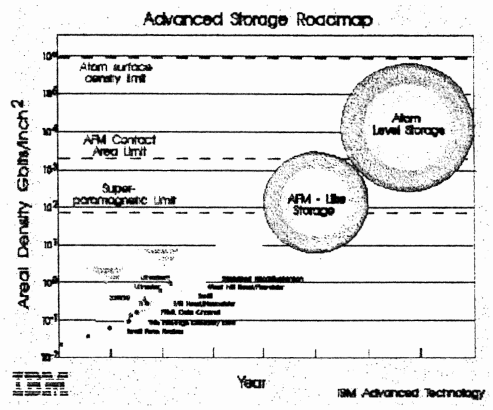
Magneto-optic recording



Near Field Recording Architecture



Physical limits



Ultimate magnetic recording limits

- Superparamagnetism
  - Sets smallest magnetic particle size that can still retain a preset magnetic state at room temperature for an extended period of time.
  - This limit has been approximately placed in the range of 5nm.
  - Given a one particle thick medium the above leads to a limit less than 100 gigabits per square inch, assuming 100 particles per bit cell will give an adequate SNR. (excludes notion of individually fabricated bit cells)
- Today limiting factors are related to spacing and tracking with techniques that offer high data rates.
- As long as state of the art is a factor on 10 or more from the ultimate limit magnetic disk storage will remain a moving target that will not be seriously challenged in its traditional role

CGR analysis of density trends

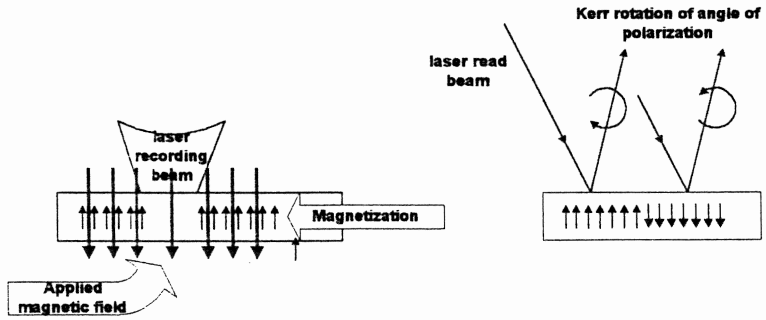
PRODUCT	RAMAC	IBM 1301	IBM 1311	IBM 2314	IBM 3330	IBM 3350	IBM 3380	INDUSTRY	**AT 37%**	**AT 60%**
YR SHIPPED	1957	1962	1963	1966	1971	1976	19811	1991	**2001**	**2001**
MAX STORAGE DENSITY	2.0E+03	2.6E+04	5.1E+04	2.2E+05	7.8E+05	3.1E+06	1.3E+07	1.3E+08	**3.0E+09**	**1.4E+10**
STORAGE DENSITY CGR		67%	96%	63%	29%	32%	32%	26%		
BPI	100	520	1,025	2,200	4,040	6,425	15,200	48,000	**270,000**	**4.10E+05**
TPI	20	50	50	100	192	478	820	2,700	**11,000**	**2.50E+04**

	STORAGE DENSIT	**LINEAR DENSITY**	**TRACK DENSITY**
AVERAGE CGR TO 1991	**37%**	**19%**	**15%**
PROJECTED 1991 TO 2001 ASSUMING 60% CGR	**60%**	**24%**	**25%**

Optical storage attributes

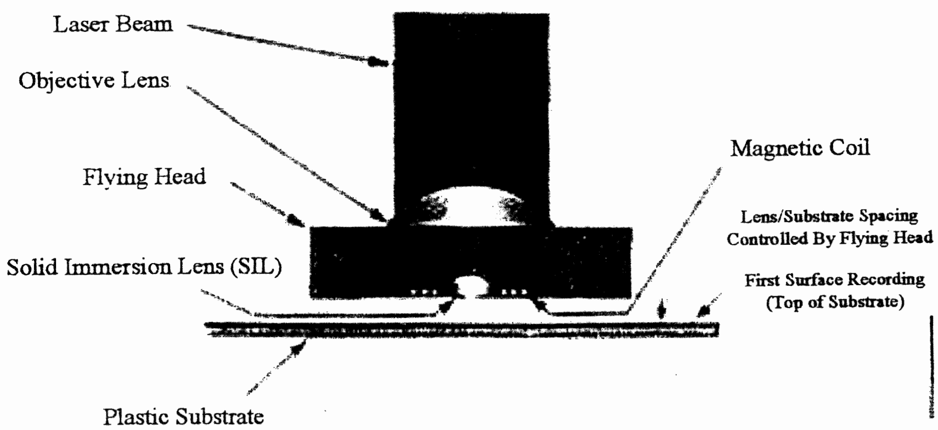
- Optical disk
  - high capacity in small form factor
  - removable storage media
  - direct access
- Multi-function
  - Read only - stamped
  - Write recordable and read/write becoming available
- CD ROM now accepted for program distribution and multi-media
- Removable disk feature attractive for higher performance automated libraries
- Convergence for entertainment and computer needs make DVD a unique product opportunity.

Magneto-optic recording



ASH Magnetic Disk Drive Technology IIST

Near Field Recording Architecture



IIST

Near field recording

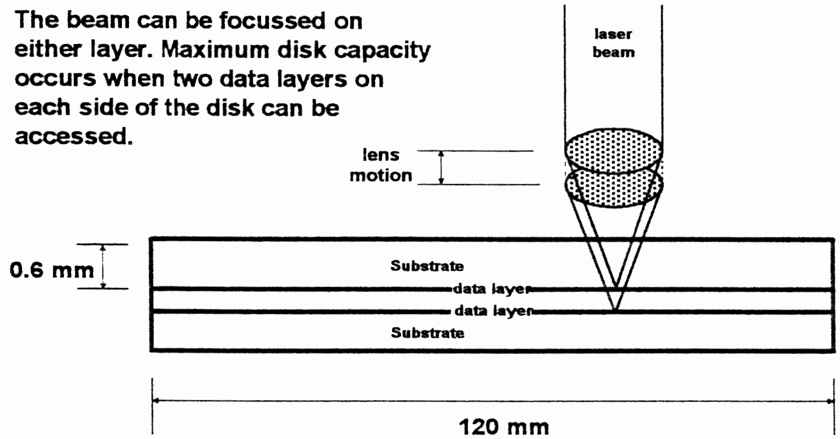
- Technology being pursued by TeraStor
- Optical flying head with solid immersion lens
  - Spacing 4 to 6 microinches, much greater than magnetic disk
    - Focus control provided by air bearing as in magnetic disk
- Based on magneto-optic storage medium
- Laser, magnetic coil and lens system all integrated into slider
- Laser beam at surface less than wavelength of light
  - SIL effectively increases NA, reducing beam size
- Projecting approximately 10 to 20 times the areal density of hard magnetic disks
- Initial implementation to offer removable disk
- TeraStor to license technology

DVD format

	DVD Format (one side)	Compact Disc
Capacity	4.7 GB x2	0.68 MB
Online capacity	4.7GB / 8.5GB	0.68 MB
Laser wavelength	635-650 nm	780 nm
Numerical aperature	0.6	0.45
Track density	34 Ktpil	16 Ktpi
Bit density	96 Kbpi	43 Kbpi
Areal density	3.3 gb/sq. in.	0.69 Gb/sq in.
Recording band	14 - 58 mm	25 - 57 mm
Reference velocity	3.27 m/s	1.2 - 4.8 m/s
Data Rate	10 mb/s	1.2 -4.8 Mb/s

Two layer DVD disk

The beam can be focussed on either layer. Maximum disk capacity occurs when two data layers on each side of the disk can be accessed.



DVD specifications

- Capacity requirement set by digital video application
  - 135 min movie
  - picture quality equal or better than laser disk
  - 3 sound tracks, Dolby AC3 5.1 at 384 Kb/s each
  - 10 subtitles at 10 Kb/s each
- Double sided construction
  - Up to two recorded layers per side
  - symmetrical design reduces warp
- Tilt tolerance is a critical parameter that must be met
- Based on phase change media
  - readout uses reflectivity change

Flash drive

- Positive features
  - Low power and low energy consumption
    - less attractive with high ratio of writes to reads
  - Rugged, less sensitivity to shock
  - Small form factors: Compact and PC cards
  - Fits lower capacity applications, e.g., digital cameras
- Limitations
  - High cost per megabyte
    - not as attractive for removable low cost storage.
    - write process slow compared to recordable disk
- Basic box cost of basic disk drive (even for low capacities) opens opportunity for flash memory in the 2 to 40 MB range today

Flash media management

- Write management
  - linear writing
    - Data transfer set by cycle speed and unit block size
  - tag for update/overwrite, block overwrite
  - background cleanup
- Wear leveling
  - Limited number of overwrite cycles leads to moving data around to distribute activity uniformly over all storage cells
- Able to store more than one bit per cell by proper thresholding

Removable and secondary storage

- **Zip and Jaz (Iomega)**
- **Super-disk (Imation)**
- **Super-floppy (Sony)**
- **CD recordable and rewriteable**
- **CD/DVD**
- **M-O storage devices**

Role of magnetic tape

Second hard disk

IBM micro-drive

Storage of data on network

Industry perspectives (open discussion)

- Drive assembly, component manufacturers and R&D
- Disk drive companies
  - consolidation
  - vertical integration
  - hi end versus low end
- Competition
  - US versus foreign
- Component suppliers and the food chain
- Research level and technology advances
- Market size and growth