

**DATA STORAGE TUTORIAL
ON
MAGNETIC & OPTICAL RECORDING**

**Institute for
Information
Storage
Technology**

**Santa Clara University
March 23 - 27, 1992**

INSTITUTE for INFORMATION STORAGE TECHNOLOGY
Santa Clara University

TUTORIAL on
MAGNETIC & OPTICAL DATA STORAGE

March 23 - 27, 1992

*** Schedule and Syllabus ***

Monday, March 23

- 9 - 12 AM **MAGNETIC MATERIALS** ... **G. Bate**
Introduction to Magnetism
Intrinsic properties of ferromagnetics; spontaneous magnetization,
Curie Temperature, magnetocrystalline anisotropy, magnetostriction
Extrinsic properties; coercivity, remanence coercivity, permeability,
remanent magnetization, squareness
Hysteresis loops
Magnetic properties of particles
Magnetic properties of thin magnetic films, longitudinal and perpendicular
Non-magnetic properties
- 1 - 5 PM **MAGNETICS AND CIRCUITS** ... **J. Monson**
Physical laws
Unit systems
Magnetic scalar potential
Magnetic circuits
Gap fringe fields

Tuesday, March 24

- 9 - 12 AM **MAGNETIC RECORDING** ... **J. Monson**
Recording medium characteristics
Self demagnetization
Head fields
Longitudinal recording
Demagnetization effects
- 1 - 5 PM **READ PROCESS** ... **J. Monson**
Reciprocity between writing and reading
Readback pulse shapes
Wavelength effects:
gap loss
spacing loss
transition spreading
thickness loss
Magneto-resistive read heads

Wednesday, March 25

- 9 - 12 AM **READBACK CHANNELS** ... **A. Hoagland**
Linear block diagram
Pulse crowding, bit shift, inter-symbol interference
Equalization
Channel modulation codes:
characterization
example codes: NRZ, MFM, 2-7

1-5 PM **DRIVE DESIGN ISSUES** ... **A. Hoagland**
 Head slider arm assemblies
 Actuators, rotary & linear
 Servo systems
 dedicated and imbedded
 Track mis-registration and off-track performance
 Soft error rate
 sources of error
 window margin analysis
 Zone bit recording
 ECC
 Error detection and correction concepts
 Coding examples:
 Hamming , Reed-Solomon
 Burst corrections by interleaving
 Code redundancy

Thursday, March 26

9 - 12 AM **DATA STORAGE DEVICES** ... **A. Hoagland**
 Form factor and packaging
 Capacity, access time, data rate
 Response time
 Product features
 Disk arrays: reliability, availability, fault tolerance
 Tape units: longitudinal and rotary drives

1 - 5 PM **OPTICAL RECORDING, MEDIA AND SYSTEMS** ... **G. Bate**
 An introduction to WORM and reversible systems
 Thermo-magnetic recording
 Optics: conventional and holographic
 Lasers
 Magneto-optic drives
 Read-only optical disks, CD, CD-ROM
 Curie-point and compensation point writing

Friday, March 27

9 - 12 AM **INTERFACES** ... **A. Hospodor**
 Overview of interfaces
 SCSI -2
STORAGE ARCHITECTURE
 Evolution
 Optimizing Data Transfers
 Disk Arrays
 Disk Buffers and Caches

1 - 5 PM **FUTURE TRENDS** ... **Faculty Panel**
 Hybrid magnetic and optical systems
 Laser development
 Integrated optic heads
 Phase change, dye laser materials
 Semiconductor memories: SRAM, DRAM, ferroelectrics, etc.
 Track density limits
 Perpendicular recording
 Disk arrays
 High capacity flexible disks
 Helical-recording drives for data recording
 Magnetic Recording in 2001

TABLE OF CONTENTS

Monday, March 23rd

Magnetic Materials, *G. Bate* . . . Section 1

Magnetics and Circuits, *J. Monson* . . . Section 2

Tuesday, March 24th

Magnetic Recording, *J. Monson* . . . Section 3

Read Process, *J. Monson* . . . Section 4

Wednesday, March 25th

Readback Channels, *A. Hoagland* . . . Section 5

Drive Design Issues, *A. Hoagland* . . . Section 6

Thursday, March 26th

Data Storage Devices, *A. Hoagland* . . . Section 7

Optical Recording, Media and Systems, *G. Bate* . . . Section 8

Friday, March 27th

Storage Interfaces and Architectures, *A. Hospodor* . . . Section 9

Future Trends, *Faculty Panel* . . . Section 10

MAGNETISM and MAGNETIC MATERIALS

Geoffrey Bate

**Institute for Information Storage Technology
Santa Clara University**

Magnetism and Magnetic Materials

Geoffrey Bate
Institute for Information Storage Technology
Santa Clara University

Contents

- Origin of magnetism; electrons in motion
- review of atomic structure; quantum numbers, exclusion principle
- diamagnetism, paramagnetism, antiferromagnetism,
ferromagnetism, ferrimagnetism
- origin of the spontaneous magnetic moment: exchange interaction
- B, H, and M; units in magnetism
- intrinsic magnetic properties; magnetization, Curie temperature,
magneto-crystalline anisotropy
- extrinsic magnetic properties; coercivity, remanence-coercivity,
remanence, permeability
- hysteresis loop - domains and domain walls
 - single-domain particles
- particles for recording: iron oxide
 - cobalt-modified iron oxide
 - chromium dioxide
 - metal
 - barium ferrite
- particle interactions - Preisach Diagram
- switching-field distribution
- time dependence of magnetic properties
- particles versus thin films
- future developments.

UNITS IN MAGNETISM

<u>Quantity</u>	<u>Symbol</u>	<u>cgs units</u>	<u>x factor</u>	<u>= SI units</u>
		$B = H + 4\pi M$		$B = \mu_0 (H + M)$
• magnetic flux density	B	gauss (G)	$\times 10^{-4}$	= Tesla (T), Webers/m ²
• magnetic flux	$\oint \vec{B} \cdot d\vec{A}$	maxwell (Mx) (G.cm ²)	$\times 10^{-8}$	= Webers
• magnetic potential diff. (magneto-motive force)	\mathcal{U}	gilbert (Gb)	$\times \frac{10}{4\pi}$	= ampere (A)
• magnetic field strength	H	oersted (Oe.)	$\times \frac{10^3}{4\pi}$	= A/m
• magnetization (per volume)	M	emu/cc	$\times 10^3$	= A/m
• " (per mass)	α	emu/g	$\times 1$	= A.m ² /kg
• magnetic moment	m	emu	$\times 10^{-3}$	= A.m ²
• susceptibility (vol.)	χ	dimensionless	$\times 4\pi$	= dimensionless
• " (mass)	κ	"	$\times 4\pi$	= "
• permeability (vacuum)	μ_0	"	$\times 4\pi \cdot 10^{-7}$	= Wb/A.m
• " (material)	μ	"	$\times 4\pi \cdot 10^{-7}$	= Wb/A.m
• Bohr magneton	μ_B	$= 0.927 \times 10^{-20}$ erg/Oe.	$\times 10^{-3}$	= A.m ²
• demagnetizing factor	N	dimensionless	$\times 1/4\pi$	= dimensionless

- * Magnetism comes from moving electric charges, usually electrons
- * The movement can be
 - orbital around the nucleus
 - "spin"
- * In technologically important materials, spin is more important than orbital motion

- * An atom may be permanently magnetic --> PARAMAGNETIC MATERIALS
- " " may not " " " " --> DIAMAGNETIC MATERIALS

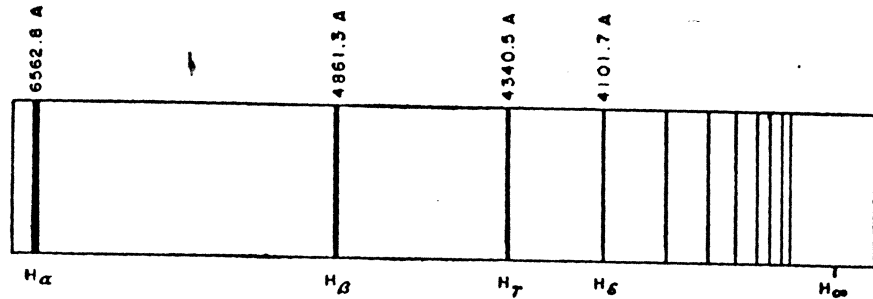
(N.B. all atoms are **DIAMAGNETIC** but some (those having a permanent magnetic moment) are **PARAMAGNETIC** as well and the paramagnetism is ~ 100 greater than diamagnetism and overwhelms it.

- * Paramagnetics are magnetized in the direction of the applied magnetic field
- * Diamagnetics " " " " " " opposite to " " "

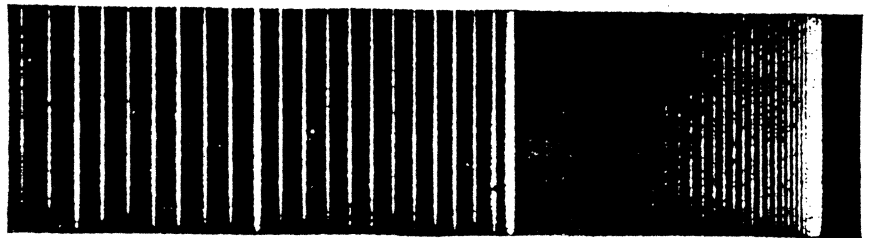
- * The total magnetization of para- or diamagnetic samples is zero if there is no applied field

- * **PARAMAGNETISM** is a necessary condition but not a sufficient condition for technological application of the material.

ATOMIC SPECTRA



(a)



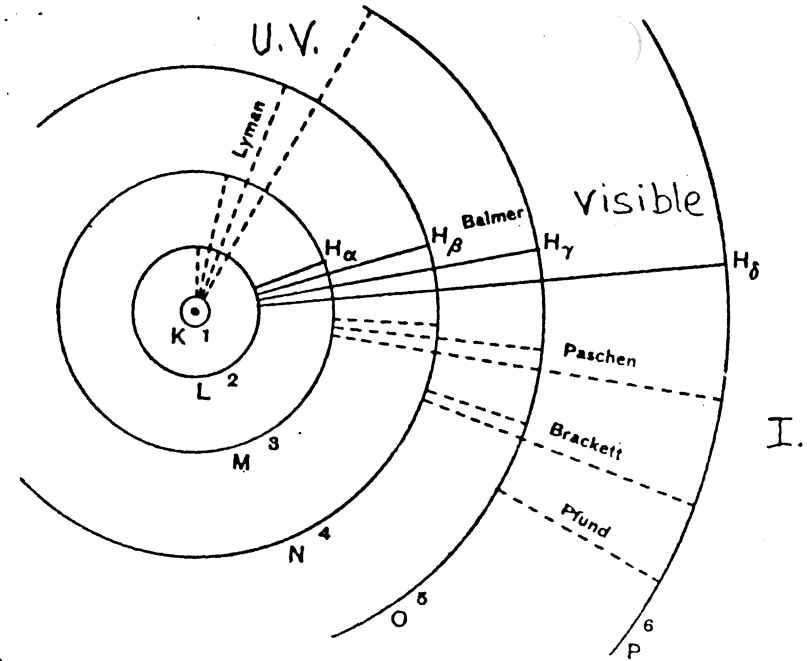
3871.4 A

3883.4 A

1885 Balmer

$$\nu' = \frac{1}{\lambda} = 109,677 \left(\frac{1}{4} - \frac{1}{n^2} \right) \text{ cm}^{-1}$$

$$\frac{2\pi^2 m e^4}{h^3 c} = \text{Rydberg's constant}$$



Lyman

$$\nu' = R \left(\frac{1}{1^2} - \frac{1}{n^2} \right); n = 2, 3, \dots, \infty$$

Balmer

$$\nu' = R \left(\frac{1}{2^2} - \frac{1}{n^2} \right); n = 3, 4, \dots, \infty$$

Paschen

$$\nu' = R \left(\frac{1}{3^2} - \frac{1}{n^2} \right); n = 4, 5, \dots, \infty$$

Brackett

$$\nu' = R \left(\frac{1}{4^2} - \frac{1}{n^2} \right); n = 5, 6, \dots, \infty$$

Pfund

$$\nu' = R \left(\frac{1}{5^2} - \frac{1}{n^2} \right); n = 6, 7, \dots, \infty$$

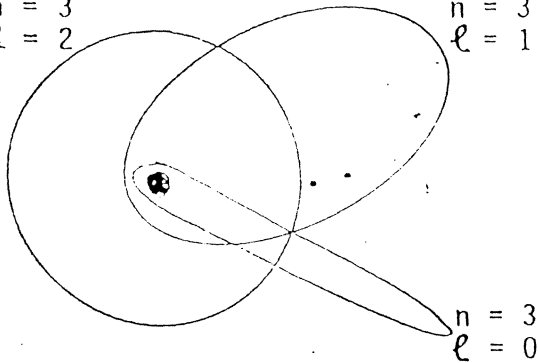
THE FOUR QUANTUM NUMBERS (Pauli)

1) PRINCIPAL QUANTUM NUMBER, n

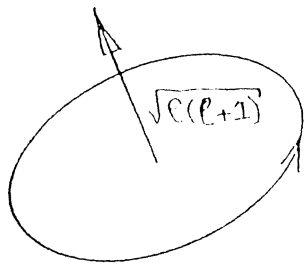
$$mvr = \frac{nh}{2\pi}$$

2) AZIMUTHAL QUANTUM NUMBER, l

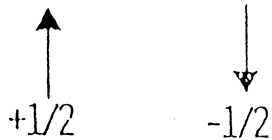
$$\begin{matrix} n = 3 & n = 3 \\ l = 2 & l = 1 \end{matrix}$$



3) MAGNETIC QUANTUM NUMBER, m_l



4) SPIN QUANTUM NUMBER



Bohr Magneton, μ_B

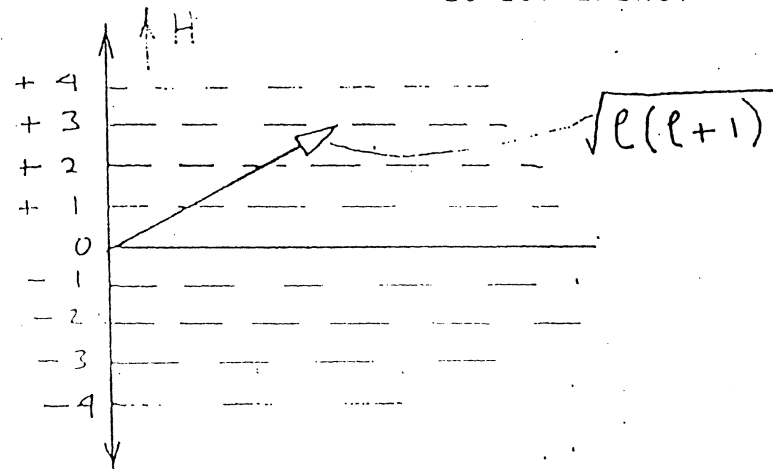
:- fundamental unit of magnetic moment

$$:- \frac{eh}{4\pi m} = 0.925 \times 10^{-20} \text{ e.m.u.}$$

POSSIBLE VALUES

$$l \leq n - 1$$

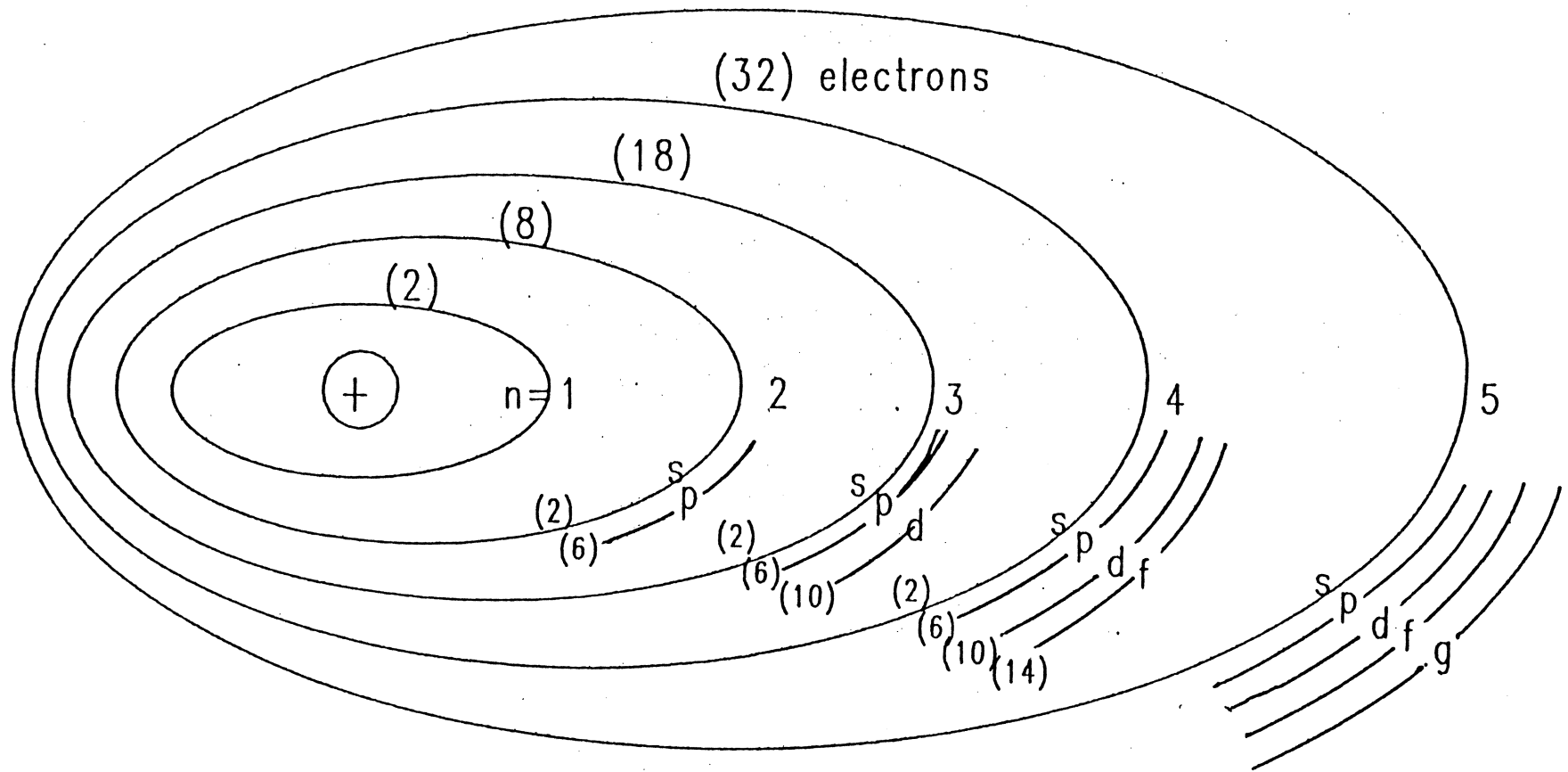
MOST ELLIPTICAL ORBITS = MOST TIGHTLY BOUND ELECTRON
= LOWEST ENERGY



possible values

$$+l, +(l-1), +(l-2), \dots, +1, 0, -1, \dots, -l$$

Atomic Orbitals



$$l = n-1, n-2, n-3, \dots, 0$$

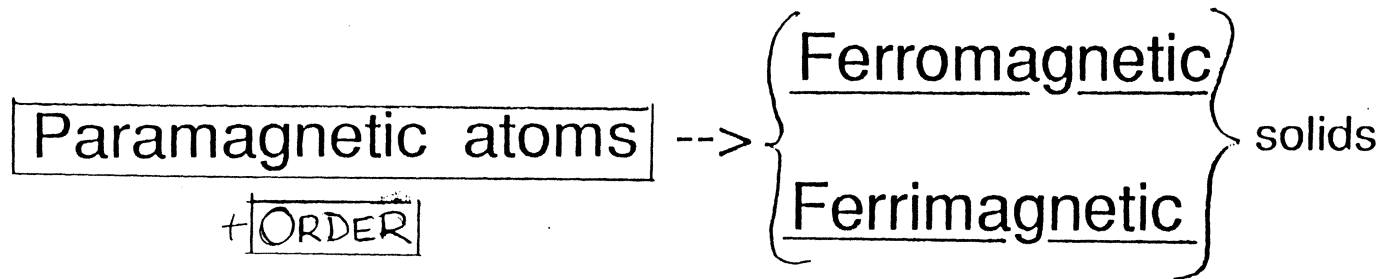
PERIODIC TABLE, MASS SUSCEPTIBILITIES

	K 1s	L 2s 2p	M 3s 3p 3d	N 4s 4p 4d 4f	O 5s 5p 5d	P 6s 6p 6d	Q 7s	Ground term	$\chi \times 10^6$
H 1	1							1S_1	-1.97
He 2	2							1S_0	-0.47
Li 3	2	1						$^2S_{1/2}$	+0.50
Be 4	2	2						1S_0	-1.00
B 5	2	2	1					$^2P_{1/2}$	-0.09
C 6	2	2	2					3P_0	-0.49
N 7	2	2	3					$^4S_{3/2}$	-0.8
O 8	2	2	4					3P_2	+106.2
F 9	2	2	5					$^4P_{3/2}$	-
Ne 10	2	2	6					1S_0	-0.88
Na 11			1					$^2S_{1/2}$	+0.51
Mg 12			2					1S_0	+0.55
Al 13			2	1				$^2P_{1/2}$	+0.05
Si 14		10 Ne core	2	2				3P_0	-0.18
P 15			2	3				$^4S_{3/2}$	-0.90
S 16			2	4				3P_2	-0.49
Cl 17			2	5				$^3P_{2,1}$	-0.57
A 18	2	2	6					1S_0	-0.48
K 19				1				$^2S_{1/2}$	+0.52
Ca 20				2				1S_0	+1.10
Sc 21				2				$^3D_{3,1}$	-
Ti 22				2				3F_4	+1.25
V 23				3				$^4F_{3,1}$	+1.4
Cr 24				5	1			5S_4	+8.08
Mn 25				5	2			6S_5	+11.8
Fe 26				6	2			5D_4	Ferro.
Co 27				7	2			5F_4	"
Ni 28				8	2			5F_4	"
Cu ⁺	2	2	6					1S_0	-
Cu 29				1				$^2S_{1/2}$	-0.086
Zn 30				2				1S_0	-0.157
Ga 31				2	1			$^2P_{1/2}$	-0.24
Ge 32				2	2			3P_0	-0.12
As 33				2	3			$^4S_{3/2}$	-0.91
Se 34				2	4			3P_2	-0.92
Br 35				2	5			$^3P_{2,1}$	-0.80
Kr 36	2	2	6					1S_0	-0.85
Rb 37								$^2S_{1/2}$	+0.21
Sr 38								1S_0	-0.20
Y 39				1				$^2D_{3/2}$	+5.0
Zr 40				2				3F_4	-0.45
Nb 41				4	1			$^4D_{3/2}$	+1.5
Mo 42				5	1			5S_4	+0.04
Nb 43				5	2			-	-
Ru 44				7	1			7F_4	+0.50
Rh 45				8	1			$^8F_{7/2}$	+1.11
Pd 46	2	2	6					1S_0	+5.4

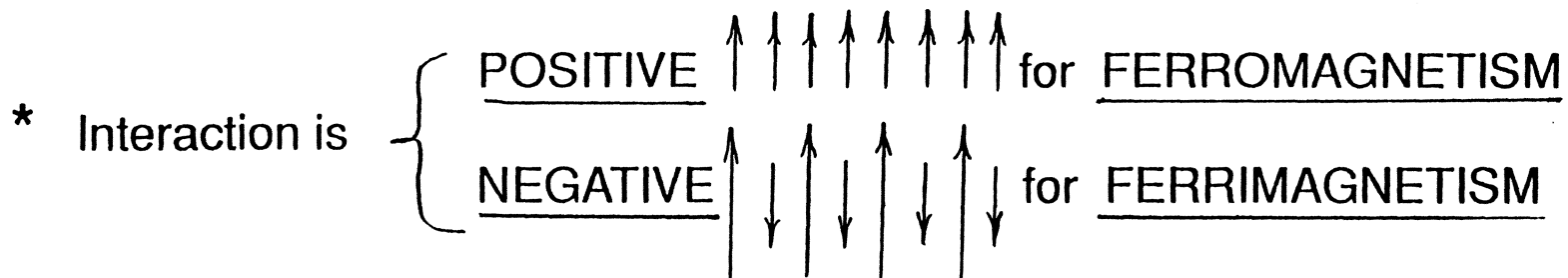
M. *

R. E.

	K 1s	L 2s 2p	M 3s 3p 3d	N 4s 4p 4d 4f	O 5s 5p 5d	P 6s 6p 6d	Q 7s	Ground term	$\chi \times 10^6$
Ag 47					1			$^2S_{1/2}$	-0.20
Cd 48					2			1S_0	-0.18
In 49					2	1		$^2P_{1/2}$	-0.11
Sn 50					2	2		3P_0	-0.25
Sb 51					2	3		$^4S_{3/2}$	-0.87
Te 52					2	4		3P_2	-0.81
I 53					2	5		$^3P_{2,1}$	-0.56
Xe 54	2	2	6	10				1S_0	-0.84
Cs 55								$^2S_{1/2}$	-0.22
Ba 56								1S_0	+0.9
La 57								$^2D_{3/2}$	+1.04
Ce 58								$^3H_{5/2}$	+15.0
Pr 59								-	+25.0
Nd 60								-	+86.0
Sm 62								-	-
Eu 63								7F_4	-
Gd 64								$^6D_{7/2}$	+22.0
Tb 65								-	Ferro.
Dy 66								-	-
Ho 67								-	-
Er 68								-	-
Tm 69								-	-
Yb 70								-	-
Lu ⁺⁺	2	2	6	10	14			-	-
Lu 71								$^2D_{3/2}$	-
Hf 72								3F_4	-
Ta 73								$^3F_{4,1}$	+0.57
W 74								3D_3	+0.28
Re 75								5S_4	-
Os 76								3D_3	+0.05
Ir 77								$^3D_{3,1}$	+0.15
Pt 78								3D_3	+1.10
Au ⁺	2	2	6	10	14			1S_0	-
Au 79								$^2S_{1/2}$	-0.15
Hg 80								1S_0	-0.166
Tl 81								$^2P_{1/2}$	-0.24
Pb 82								3P_0	-0.12
Bi 83								$^4S_{3/2}$	-1.85
Po 84								3P_2	-
At 85								-	-
Rn 86	2	2	6	10	14			1S_0	-
Fr 87								-	-
Ra 88								1S_0	-
Ac 89								$^3D_{3,1}$	-
Th 90								3F_4	+0.11
Pa 91								$^3F_{4,1}$	+2.0
U 92								3D_3	+1.78



- * Paramagnetism is a necessary (but not a sufficient) condition for spontaneous and large magnetic moments
- * In addition to atomic paramagnetism there must also be a strong EXCHANGE INTERACTION between adjacent atoms



S.I units: $B = \mu_0(H + M)$

C.G.S. units ;

$B = H + 4\pi M$

$\frac{B}{H} = 1 + 4\pi \frac{M}{H}$

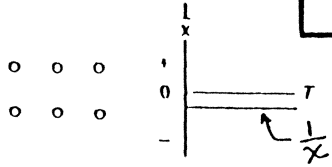
$\mu = 1 + 4\pi K$

$\left[\frac{K}{\rho} = \chi \right]$; $K = \text{vol. suscep}^y$
 $\chi = \text{mass suscep}^y$

ALL MATTER

Diamagnetic

e.g. Au

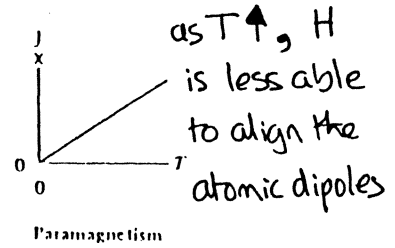
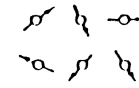


Diamagnetism

negative moment (diamag) is indep. of T

Paramagnetic

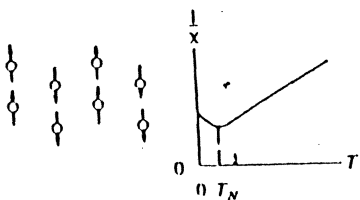
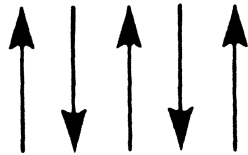
e.g. Mn



Paramagnetism

Antiferromagnetic

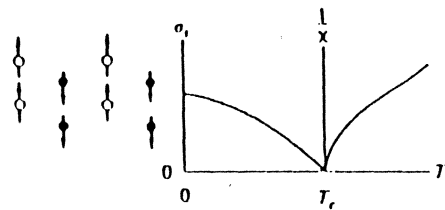
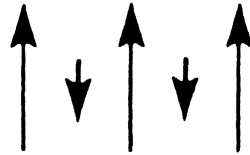
e.g. MnO_2



Antiferromagnetism

Ferrimagnetic

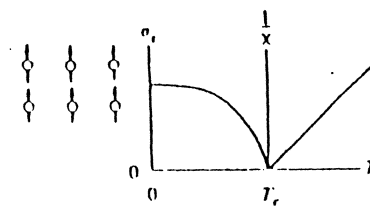
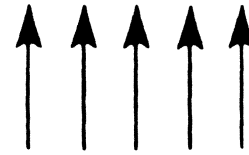
e.g. $\gamma-Fe_2O_3$



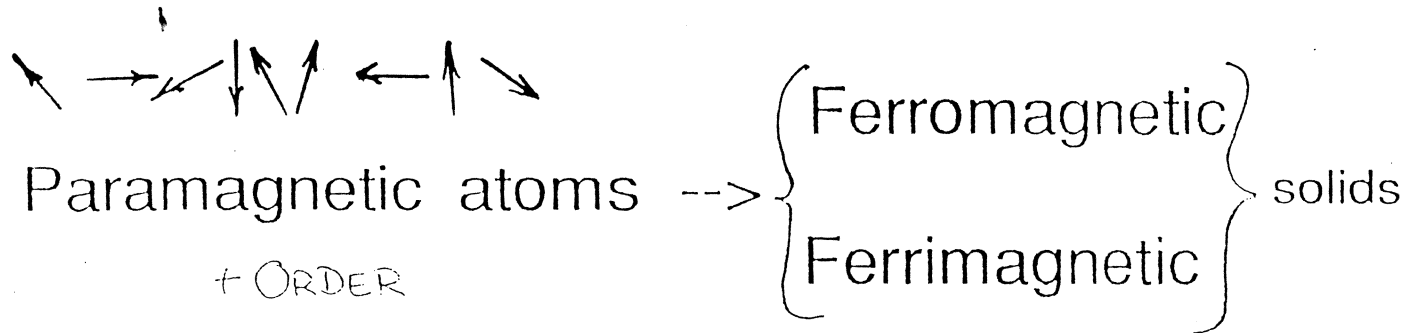
Ferrimagnetism

Ferromagnetic

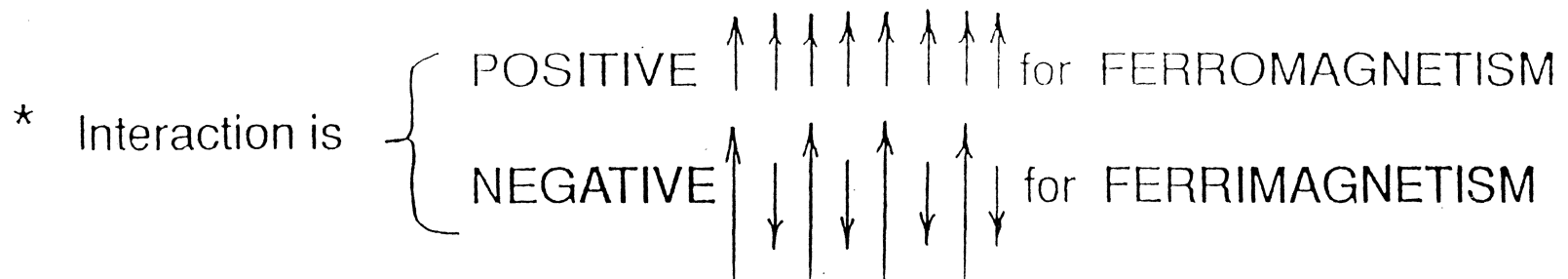
e.g. Fe



Ferromagnetism

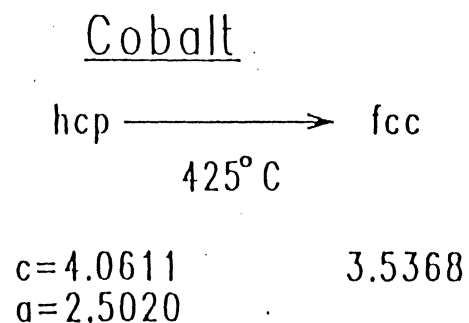
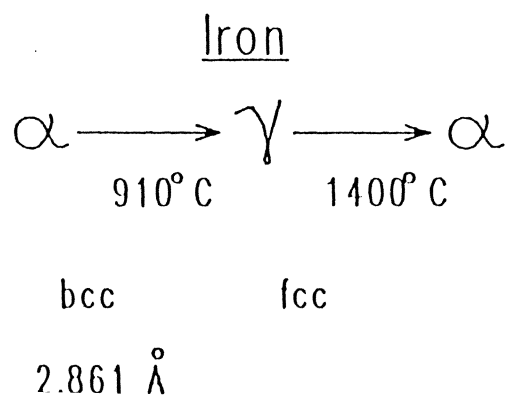


- * Paramagnetism is a necessary (but not a sufficient) condition for spontaneous and large magnetic moments
- * In addition to atomic paramagnetism there must also be a strong EXCHANGE INTERACTION between adjacent atoms



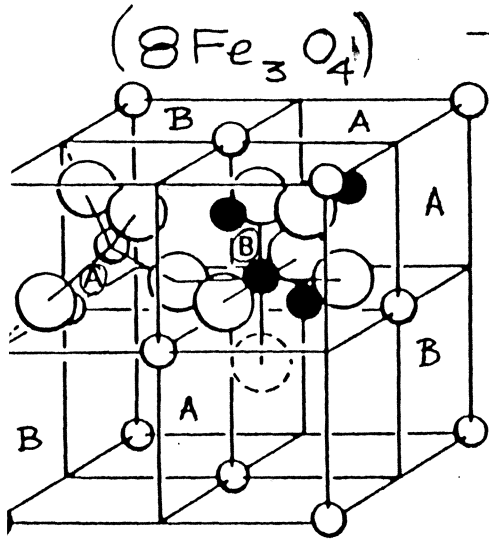
Elements of First Transition Series

	<u>Fe</u>	<u>Co</u>	<u>Ni</u>
density, g/cm ³	7.874	8.78–8.85	8.90
Curie point, °C	770	1131	358
melting point, °C	1539	1495	1455
σ _{293 K} , emu/g	217.75	161	54.39
σ _{0° K} , emu/g	221.89	162.5	57.50
M _s , emu/cm ³	1714	1422	484.1

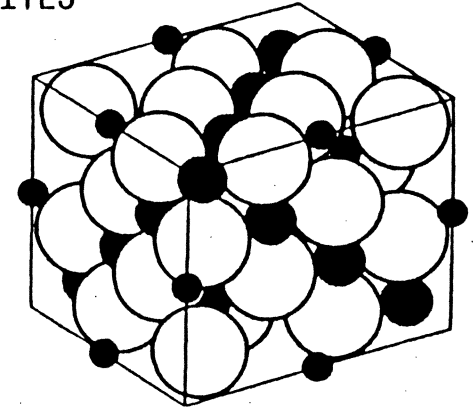
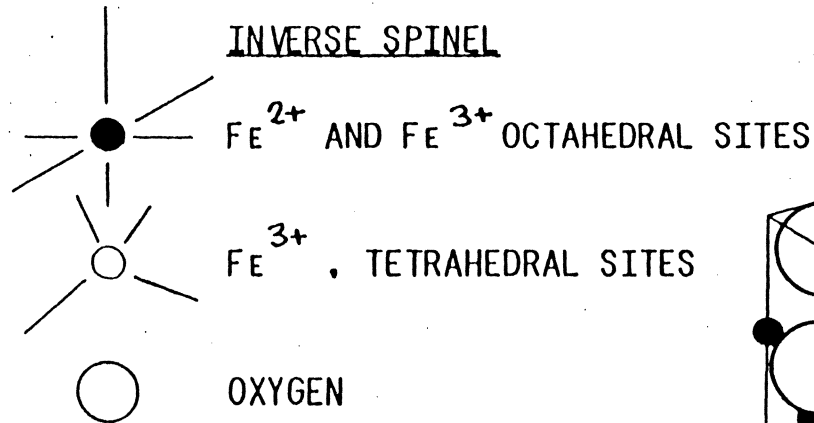


<u>ELEMENT</u>	<u>CURIE TEMPERATURE K</u>
COBALT	1393
IRON	1043
NICKEL	691
GADOLINIUM	289
TERBIUM	218
DYSPROSIUM	85
THULIUM	22
HOLMIUM	20
ERBIUM	20

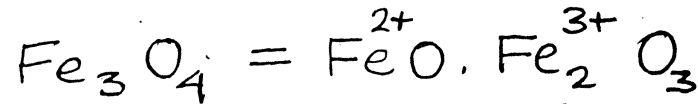
INTRINSIC MAGNETIC PROPERTIES OF Fe_3O_4 AND $\gamma-Fe_2O_3$



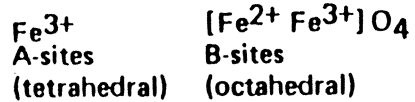
SPINEL STRUCTURE



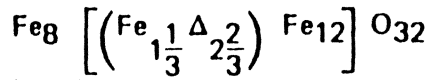
OCTAHEDRAL SITES = $\boxed{16}$, TETRAHEDRAL SITES = $(8 \times \frac{1}{8}) + (6 \times \frac{1}{2}) + 4 = \boxed{8}$



Magnetite – inverse spinel



γ – Ferric Oxide – same oxygen sublattice
 – vacancies arranged on a tetragonal superlattice ($c/a = 3$)



is the formula for $\frac{1}{3}$ unit cell.

	Fe_3O_4	$\gamma-Fe_2O_3$
Moment per molecule	4β	2.5β
σ_{so}	97 emu/g	82 emu/g
σ_s	87 emu/g	74 emu/g
Density	5.197 g/cc	5.074 g/cc
I_s	453 emu/cc	400 emu/cc
Curie temperature	575°C	590°C
Anisotropy constant		
	$K_1 = -1.10 \times 10^5$ erg/cc	$= -4.64 \times 10^4$ erg/cc
	<111> easy	<110> easy
	<100> hard	<100> hard

INTRINSIC PROPERTIES: $\gamma\text{-Fe}_2\text{O}_3, \text{Fe}_3\text{O}_4$

SATURATION MAGNETIZATION

Fe_3O_4

"INVERSE SPINEL"

A SITES

Fe^{3+}
 \longrightarrow
 $5\mu_B$

B SITES

Fe^{2+} Fe^{3+}
 \longleftarrow \longleftarrow
 $4\mu_B$ $5\mu_B$

★
 \swarrow
 0
 4

NET MOMENT = $4\mu_B$ PER FORMULA UNIT, " Fe_3O_4 "

$$\sigma_0 = \frac{N\mu_B}{M} n_B \quad ; M = 231.6 \quad \therefore \sigma_0 = 97 \text{ EMU/G} \quad (\text{EXPT} = 94 \text{ EMU/G})$$

$\gamma\text{-Fe}_2\text{O}_3$

Fe^{3+}_8
 \longrightarrow
 $(8 \times 5)\mu_B$

$(\text{Fe}^{3+}_{4/3} \Delta 8/3)$
 \longleftarrow
 $(4/3 \times 5)\mu_B$

Fe^{3+}_{12}
 \longleftarrow
 $(12 \times 5)\mu_B$

O_{32}

NET MOMENT = $(16/3 \times 5)\mu_B$ per Fe_2O_3

MOMENT PER FORMULA UNIT = $3/32 \times 16/3 \times 5\mu_B = 2.5\mu_B$

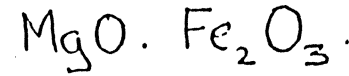
OR 1.25μ / IRON ATOM: EXPT = 1.18

$\sigma_0 = 82 \text{ EMU/G}$

$\sigma_0 = 74 \text{ EMU/G}$; EXPERIMENT
 R.T.

FERRIMAGNETISM

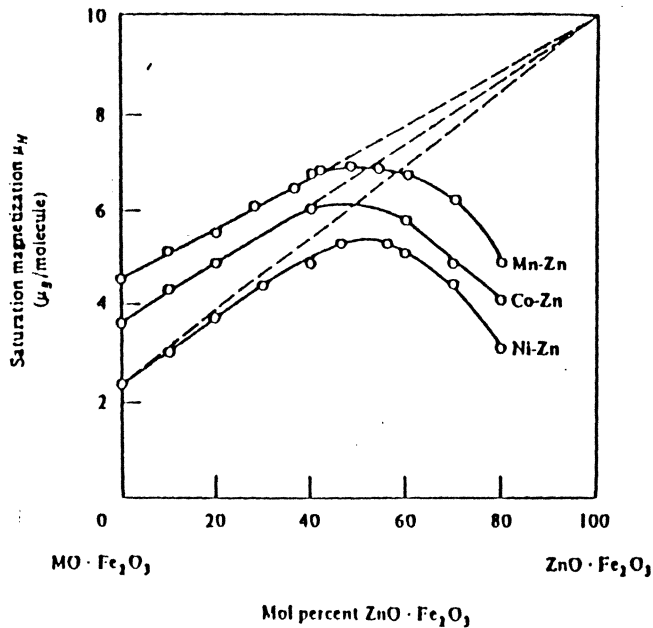
Substance	Lattice parameter a (Å)	Density (g/cm ³)	0 K		20° C		T_c (°C)
			σ_0 (emu/g)	M_0 (emu/cm ³)	σ_s (emu/g)	M_s (emu/cm ³)	
MnO · Fe ₂ O ₃	8.50	5.00	112	560	80	400	300
FeO · Fe ₂ O ₃	8.39	5.24	98	510	92	480	585
CoO · Fe ₂ O ₃	8.36	5.29	90	475	80	425	520
NiO · Fe ₂ O ₃	8.34	5.38	56	300	50	270	585
CuO · Fe ₂ O ₃	8.37*	5.41	30	160	25	135	455
MgO · Fe ₂ O ₃	8.36	4.52	31	140	27	120	440
BaO · 6Fe ₂ O ₃	$a = 5.88$ $c = 23.2$	5.28	100	530	72	380	450
Fe	2.87	7.87	222	1747	218	1714	770



(pure inverse $\Rightarrow n_{\mu_B} = 0$)

0.1 of Mg²⁺ } A sites
 0.9 of Fe³⁺ } 0.9 × 5 = 4.5 μ_B

0.9 of Mg²⁺ } B sites
 1.1 of Fe³⁺ } 1.1 × 5 = 5.5 μ_B



Example	Substance	Structure	Tetrahedral A sites	Octahedral B sites	Net moment (μ _B molecule)
1	NiO · Fe ₂ O ₃	Inverse	Fe ³⁺ 5 ←	Ni ²⁺ , Fe ³⁺ 2 5 ← ←	2
2	ZnO · Fe ₂ O ₃	Normal	Zn ²⁺ 0	Fe ³⁺ , Fe ³⁺ 5 5 ← →	0
3	MgO · Fe ₂ O ₃	Mostly inverse	Mg ²⁺ , Fe ³⁺ 0 4.5 →	Mg ²⁺ , Fe ³⁺ 0 5.5 ←	1
4	0.9 NiO · Fe ₂ O ₃	Inverse	Fe ³⁺ 4.5 →	Ni ²⁺ Fe ³⁺ 1.8 4.5 ← ←	2.8
	0.1 ZnO · Fe ₂ O ₃	Normal	Zn ²⁺ 0	Fe ³⁺ Fe ³⁺ 0.5 0.5 ← ←	
			4.5 →	7.3 ←	

TECHNOLOGICALLY IMPORTANT MAGNETIC MATERIALS

TWO KEY ATTRIBUTES:

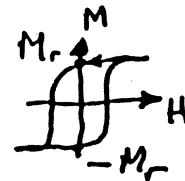
(and large)

1) SPONTANEOUS MAGNETIC MOMENT

2) MAGNETIC ANISOTROPY

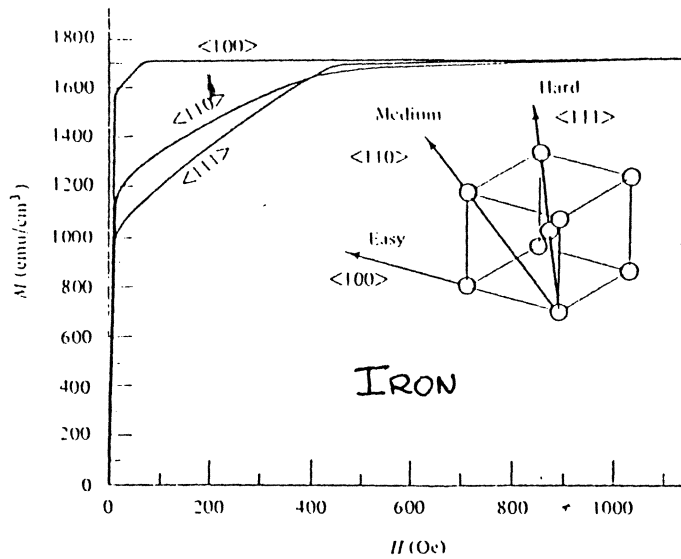
& MAGNETIC HETEROGENEITY

hysteresis

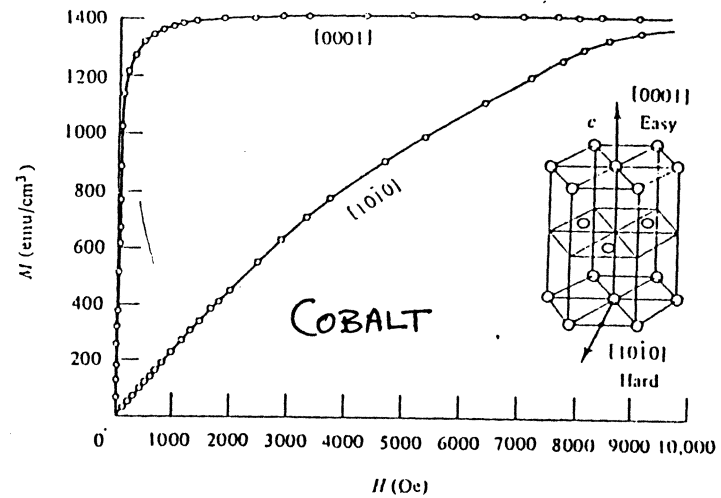


→ these are small for soft-magnetic materials
large for hard-magnetic materials

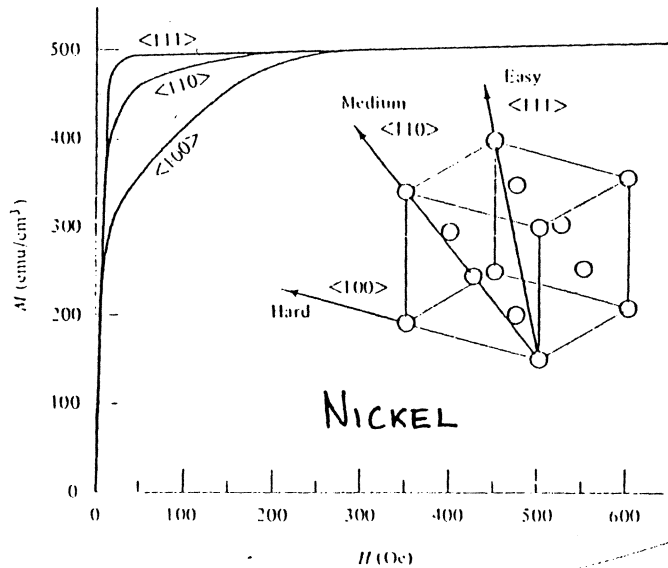
MAGNETO-CRYSTALLINE ANISOTROPY



(a)



Magnetization curves for a single crystal of cobalt (by Kaya [7.3]).



SPIN - ORBIT COUPLING is the origin of magneto-crystalline anisotropy. Orbital motion of electrons experiences inhomogeneous electric fields from the crystal lattice.

∴ expect $|K|$ to be LARGE when crystal symmetry is LOW

e.g. Co, MnBi, $|K| \sim 10^7 \text{ erg./cc.}$ but Fe, Ni, $|K| \sim 10^5 \text{ erg./cc}$

MAGNETOCRYSTALLINE ANISOTROPY

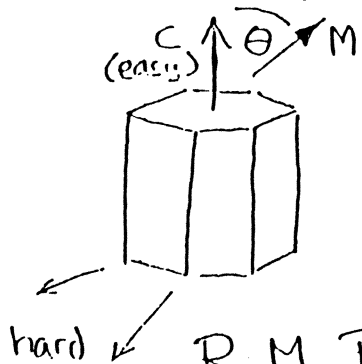
relationship between K_1, K_2 and EASY directions

Directions of Easy, Medium, and Hard Magnetization in a Cubic Crystal
(from Bozorth [G.4])

Cubic

	[Fe]			[Ni]		
K_1	+	+	+	-	-	-
K_2	+∞ to -9K ₁ /4	-9K ₁ /4 to -9K ₁	-9K ₁ to -∞	-∞ to 9 K ₁ /4	9 K ₁ /4 to 9 K ₁	9 K ₁ to +∞
Easy	$\langle 100 \rangle$	$\langle 100 \rangle$	$\langle 111 \rangle$	$\langle 111 \rangle$	$\langle 110 \rangle$	$\langle 110 \rangle$
Medium	$\langle 110 \rangle$	$\langle 111 \rangle$	$\langle 100 \rangle$	$\langle 110 \rangle$	$\langle 111 \rangle$	$\langle 100 \rangle$
Hard	$\langle 111 \rangle$	$\langle 110 \rangle$	$\langle 110 \rangle$	$\langle 100 \rangle$	$\langle 100 \rangle$	$\langle 111 \rangle$

UNIAXIAL



$$E = K'_0 + K'_1 \cos^2 \theta + K'_2 \cos^4 \theta$$

$$K'_1 + K'_2 > 0$$

axis (eg. [00.1], [100])

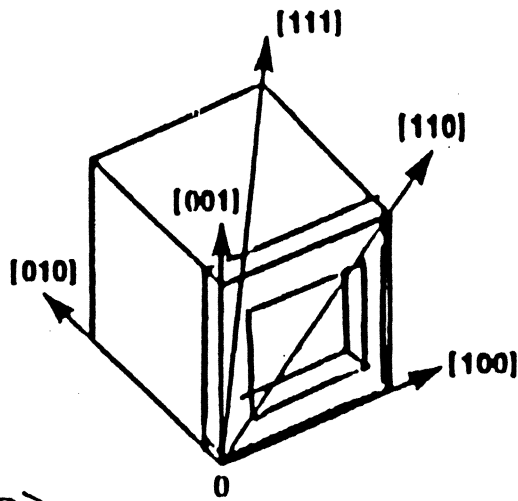
$$K'_1 + K'_2 < 0$$

basal plane (eg. [10.0])

hard ↙ R.M. Bozorth, Phys. Rev. 50, 1076-81 (1936)

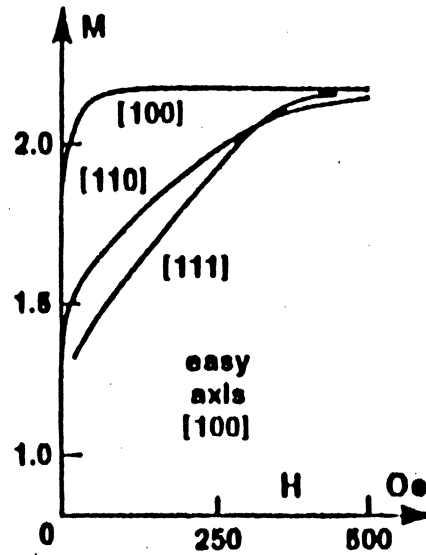
MAGNETO-CRYSTALLINE ANISOTROPY

Iron And Nickel



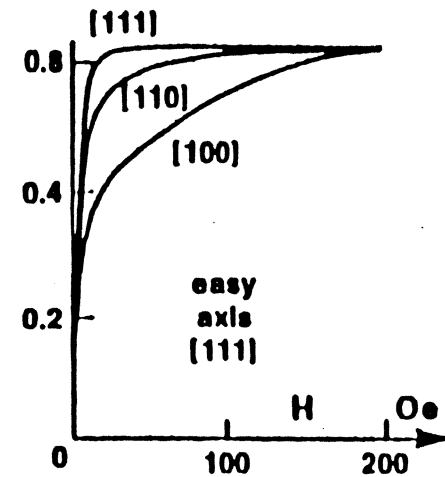
$\langle 100 \rangle$
 = $[100], [010], [001]$

IRON



$K_1 = 4.8 \times 10^5 \text{ erg/cc}$
 $K_2 = 5 \times 10^4$

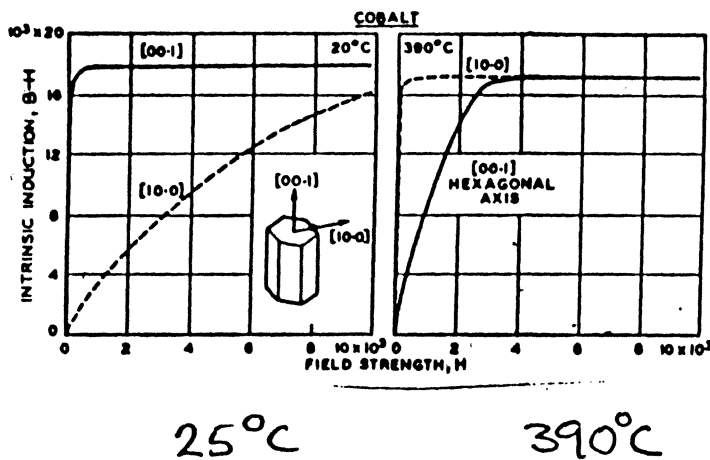
NICKEL



$K_1 = 4.5 \times 10^4$
 $K_2 = 2.34 \times 10^4$

magneto-crystalline energy = $K_0 + K_1 (\alpha_1^2 \alpha_2^2 + \alpha_2^2 \alpha_3^2 + \alpha_3^2 \alpha_1^2) + K_2 (\alpha_1^2 \alpha_2^2 \alpha_3^2) + \dots$

MAGNETO-CRYSTALLINE ANISOTROPY OF COBALT



$$B = H + 4\pi M$$

$$B - H = 4\pi M$$

$$E_K = K_1 \sin^2 \phi + K_2 \sin^4 \phi$$

$$K_1 + K_2 > 0$$

[00.1] easy (hex. axis)

$$K_1 + K_2 < 0$$

[10.0] easy (basal plane)

Cobalt

$$K_1 = +3.98 \times 10^6 \text{ erg/cc}$$

$$K_2 = +1.98 \times 10^6 \text{ erg/cc}$$

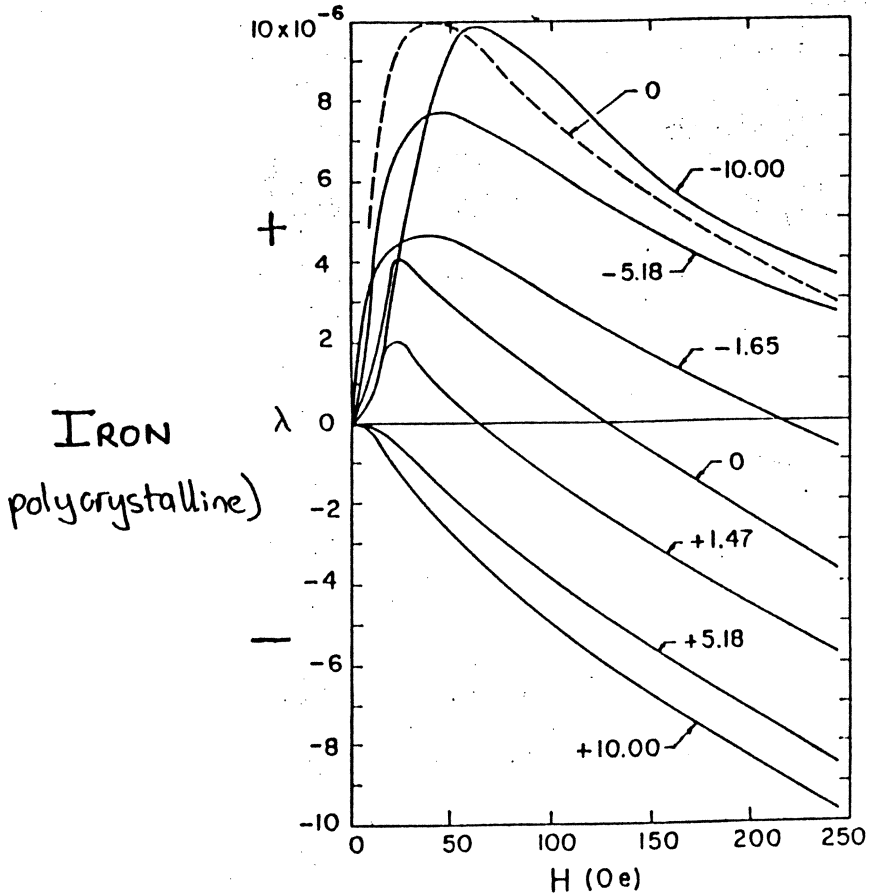
Origin of Magneto-crystalline anisotropy is Spin-Orbit interaction.

MAGNETOSTRICTION

• positive λ_s as well as $-\lambda_s$

{ INCREASING magnetization with TENSION, sample expands when magnetized
 }

 { DECREASING magnetization with TENSION, sample CONTRACTS when MAGNETIZED
 }

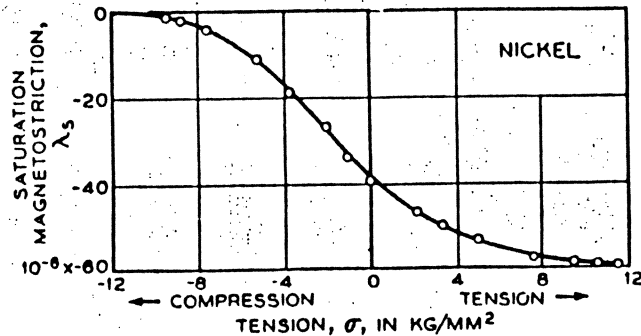
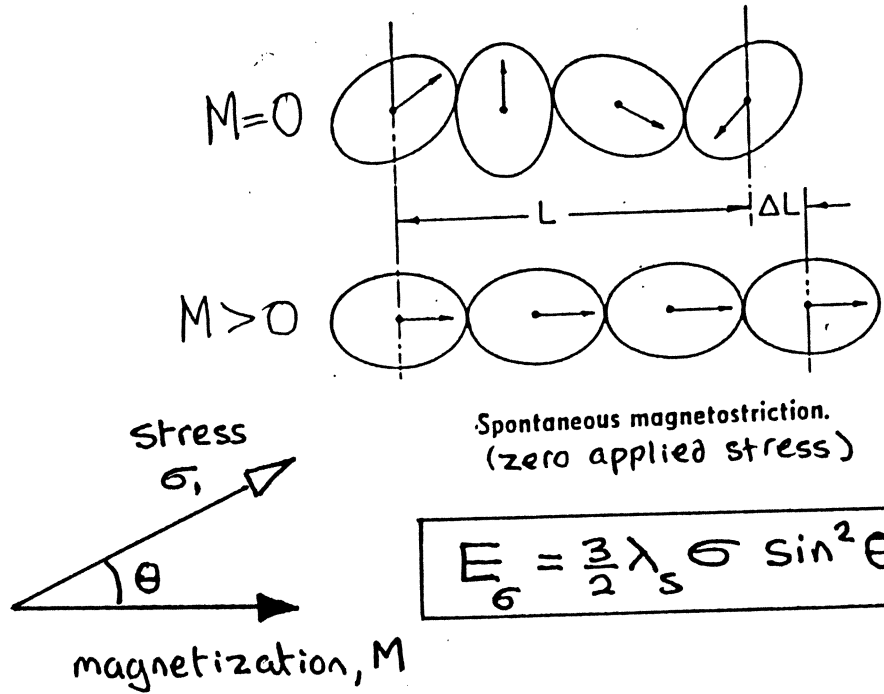


parameter is applied stress in 1000 lb/in².

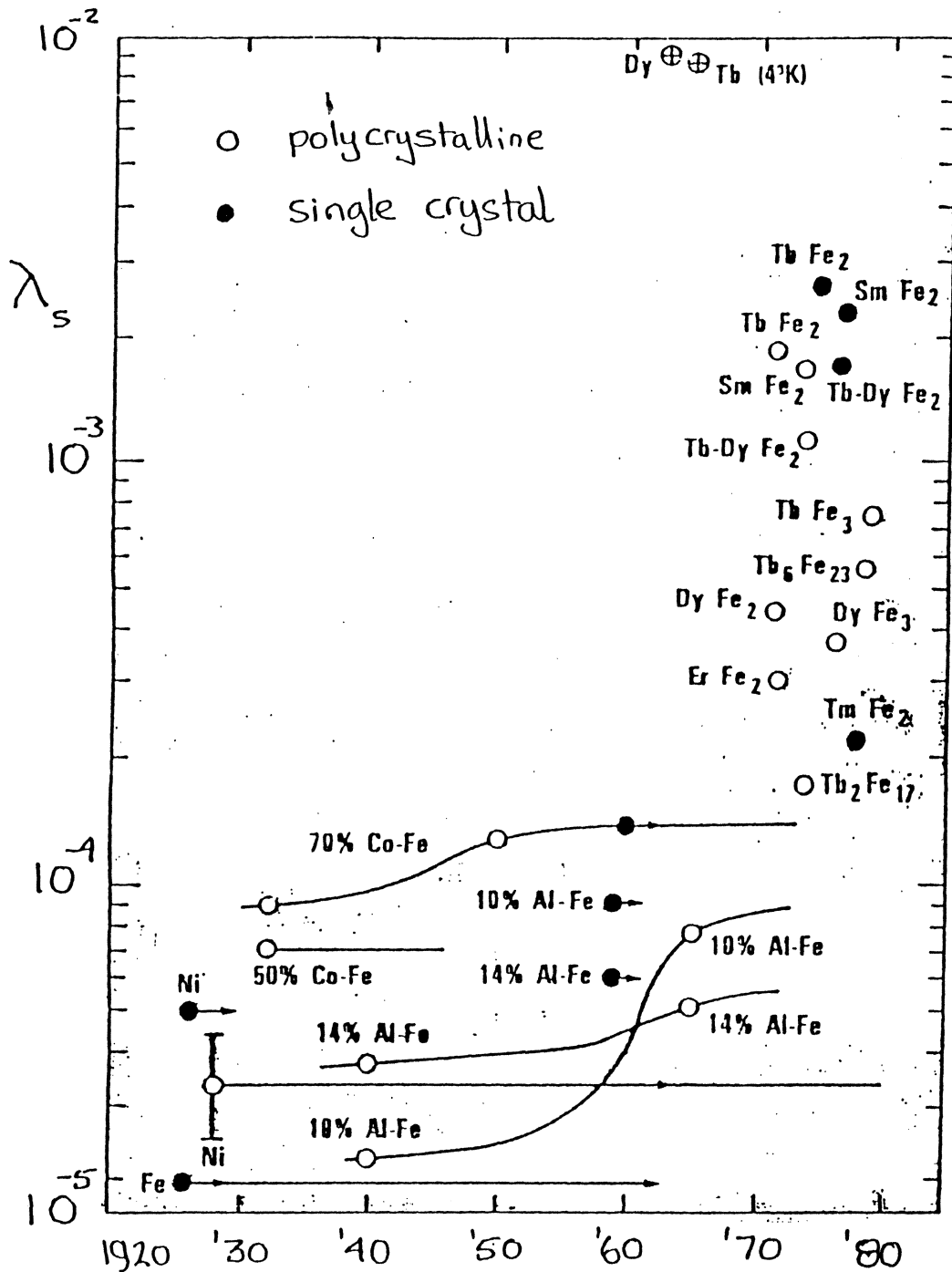
+ tensile
 - compressive

$\pm \lambda_s$ depends on

- material, eg. Fe, Co, Ni etc
- stress, tensile or compressive, magnitude
- applied field



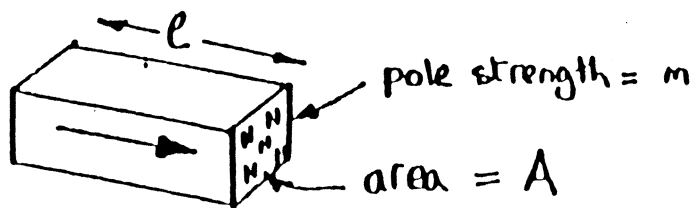
MAGNETOSTRICTION



magnetostrictive materials used to generate SONAR pulses

← [this class of materials (RE-TM) also has large magneto-optic effects]

\vec{B} , \vec{H} , and \vec{M}



magnetic moment = $m l$

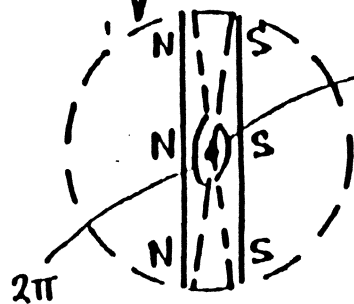
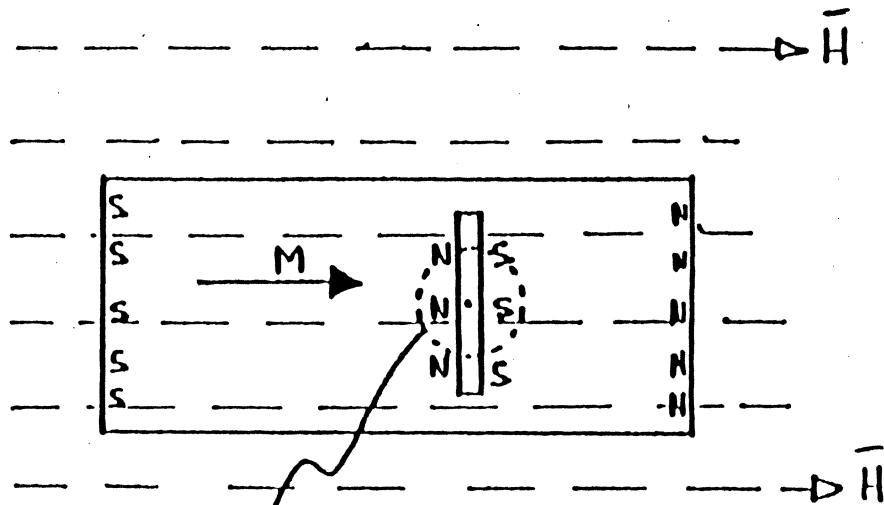
volume = $A l$

mag. mom. / volume = $\vec{M} = \frac{m}{A}$

$\text{div } \vec{B} = 0$

normal compⁿ \vec{B} is continuous

tangential " \vec{H} " "



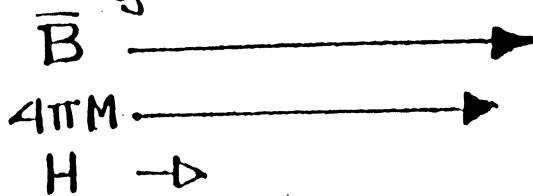
Magⁿ gives additional

Field = $2\pi \frac{m}{A} + 2\pi \frac{m}{A}$

= $4\pi \vec{M}$

$\vec{B} = \vec{H} + 4\pi \vec{M}$

gauss oersteds emu/cc



In SI Units

$\vec{B} = \mu_0 (\vec{H} + \vec{M})$

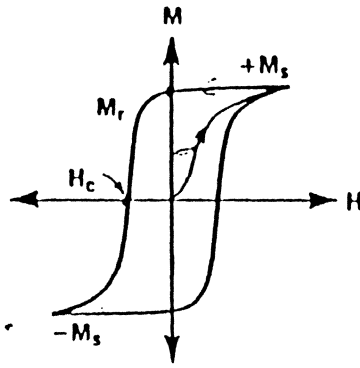
MAGNETIC PROPERTIES

Extrinsic

remanent magnetization M_r

coercivity H_c, H_r

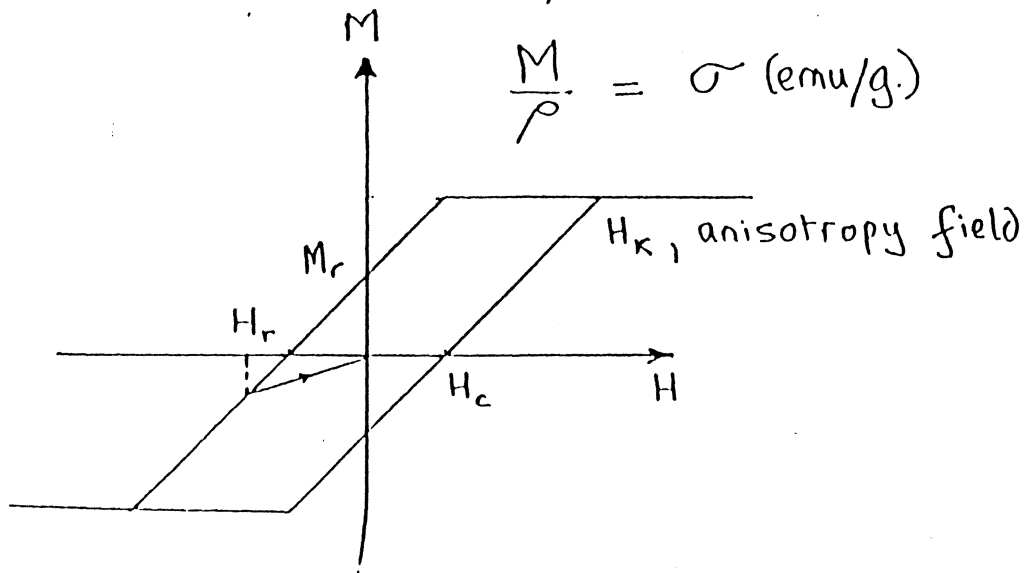
permeability μ



$$B = H + 4\pi M$$

$$\mu = 1 + 4\pi \kappa$$

$$\frac{M}{\rho} = \sigma \text{ (emu/g.)}$$



EXTRINSIC PROPERTIES also depend on:

- the type, number, arrangement of the atoms

- and the temperature

but also depend on-

the shape and size of the sample

and its previous history

$${}_{(B)}H_c < {}_{(M)}H_c$$

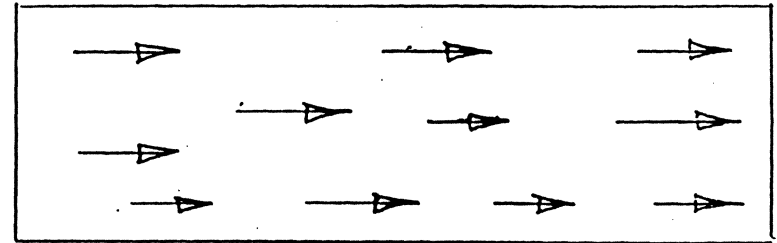
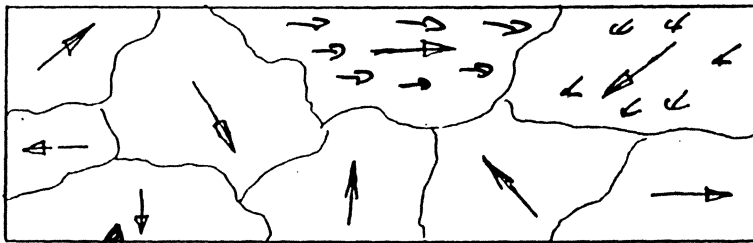
$$\frac{\kappa}{\rho} = \chi \text{ mass susceptibility}$$

1905 PROBLEMS IN FERROMAGNETISM

- 1) A PIECE OF IRON CAN BE COMPLETELY UNMAGNETIZED
- 2) YET IT CAN BE MAGNETIZED TO SATURATION BY A FIELD OF A FEW OERSTEDS

SOLUTION: PIERRE WEISS, 1905

DOMAINS: REGIONS MAGNETIZED TO SATURATION (DEPENDING ON TEMPERATURE)



$$H = 0$$

$$M = 0$$

$$H \rightarrow$$

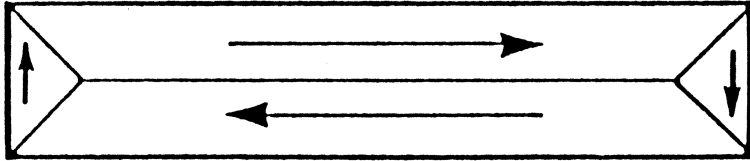
$$M = M_s$$

MAGNETIC DOMAINS

postulated to explain why a piece of iron

$$H = 0$$

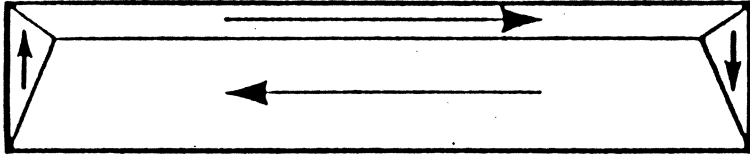
$$M = 0$$



—may not be magnetized

$$H \leftarrow$$

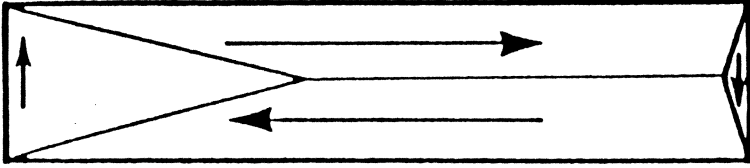
$$M \leftarrow$$



—may then be fully magnetized in small fields

$$H \uparrow$$

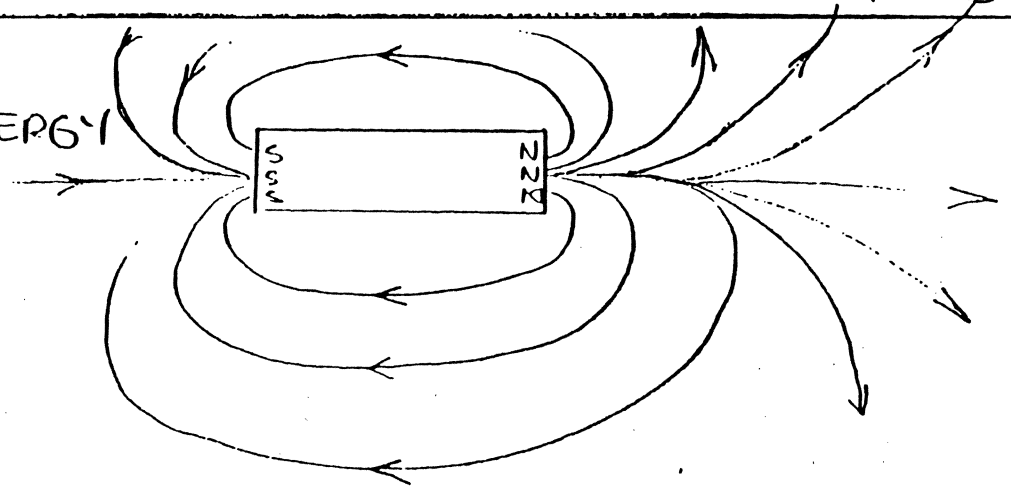
$$M \uparrow$$



DRIVING FORCE FOR THE ESTABLISHMENT OF MAGNETIC DOMAINS

- IS MAGNETOSTATIC ENERGY

$$= \int_{\text{Vol}} \frac{\mu H^2}{8\pi} dv$$



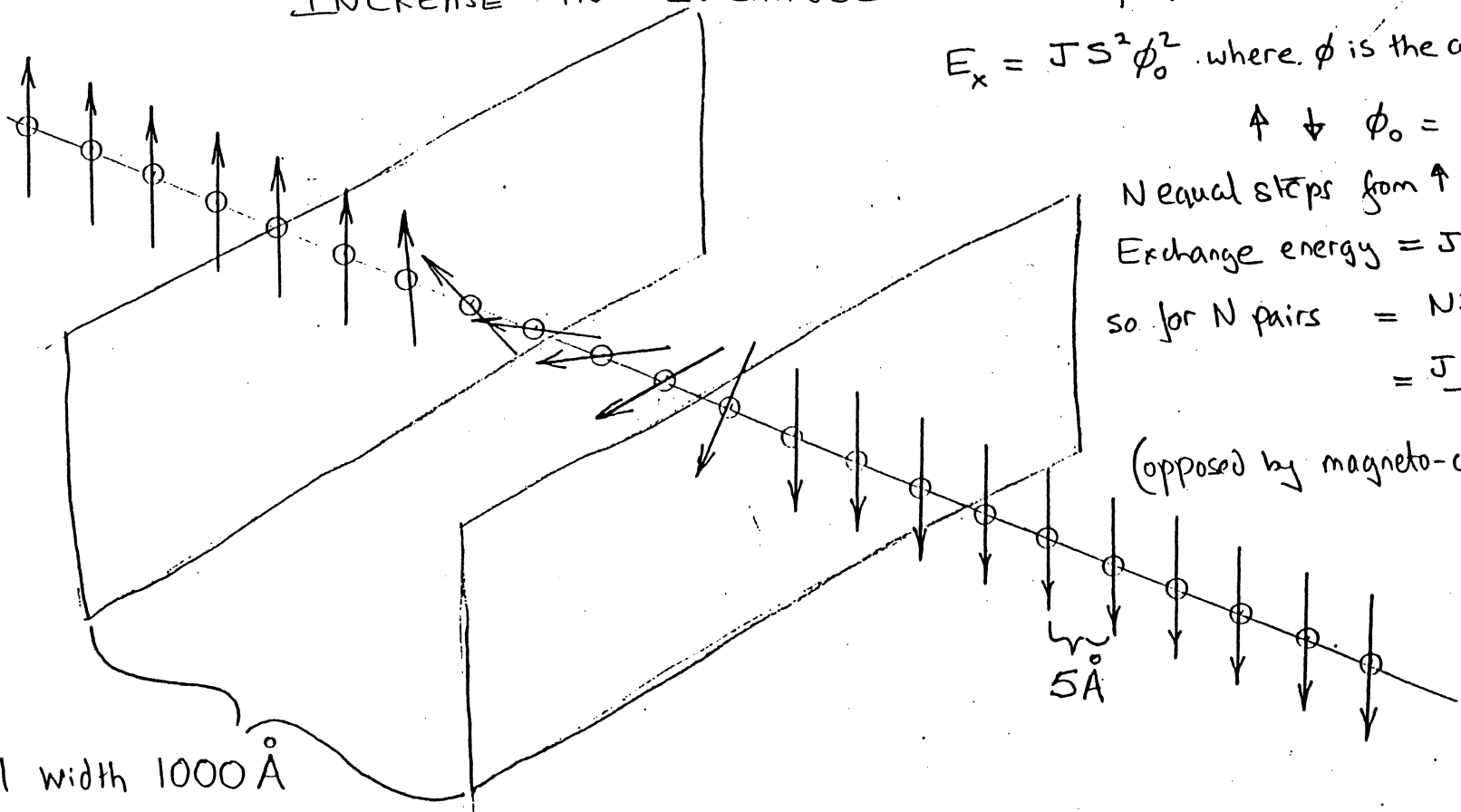
- in order to reduce magnetostatic energy the body's magnetization forms domains in which the magnetization is constant (depending only on the material and the temperature)

- a limit to the number of domains occurs because domain walls are regions of high magnetocrystalline energy.

division into domains occurs until the addition of one more wall would cost MORE MAGNETOCRYSTALLINE ENERGY than it would save in MAGNETOSTATIC ENERGY

WHAT LIMITS THE NUMBER OF DOMAINS?

DECREASE IN MAGNETOSTATIC ENERGY IS BALANCED BY THE INCREASE IN DOMAIN WALL ENERGY AND THE INCREASE IN EXCHANGE ENERGY, E_x



wall width 1000 Å

BLOCH WALLS

$E_x = JS^2 \phi_0^2$ where ϕ is the angle between spins.

$\uparrow \downarrow \phi_0 = 180^\circ$

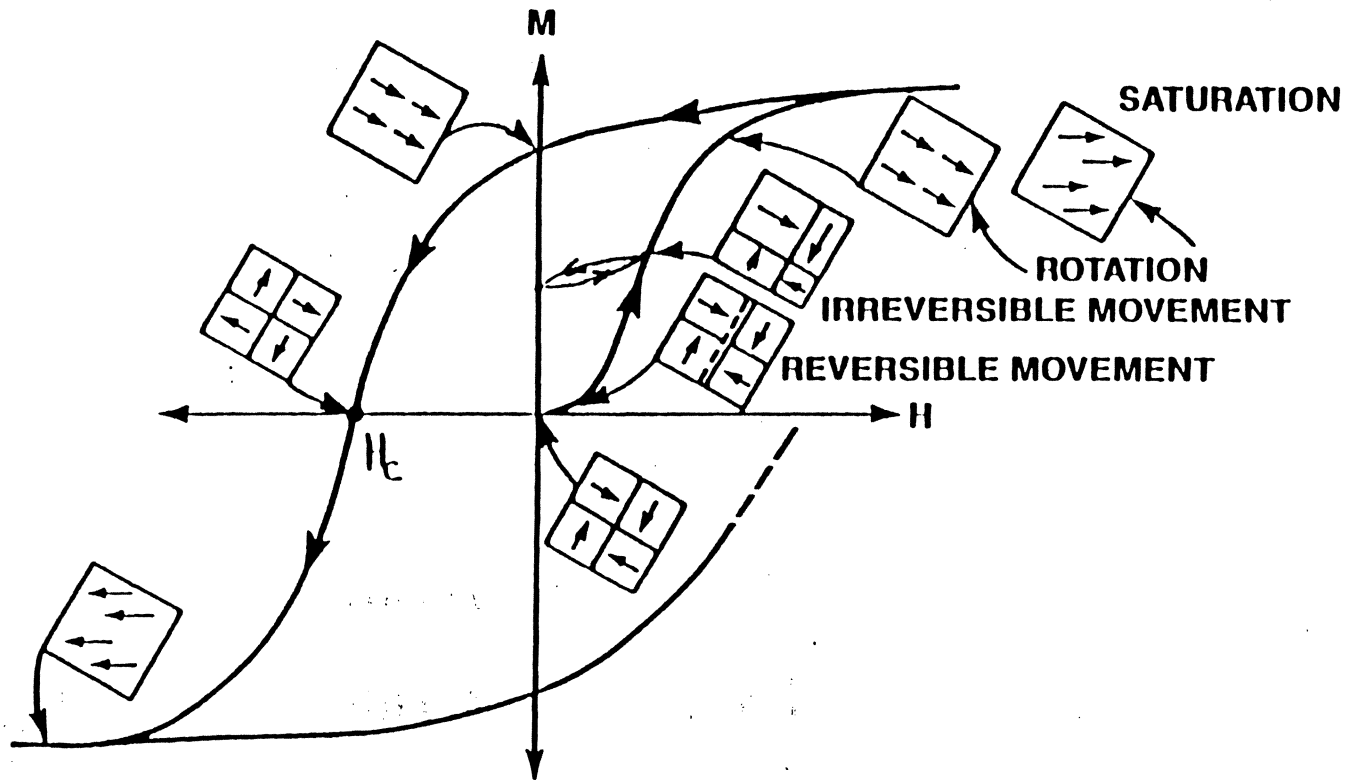
unequal steps from \uparrow to \downarrow
Exchange energy = $JS^2 \left(\frac{\phi_0}{N}\right)^2$ per pair

so for N pairs = $NJS^2 \left(\frac{\phi_0}{N}\right)^2$
= $\frac{JS^2 \phi_0^2}{N}$

(opposed by magneto-crystalline energy)

$254 \text{ \AA} = \frac{1}{1,000,000}$ "

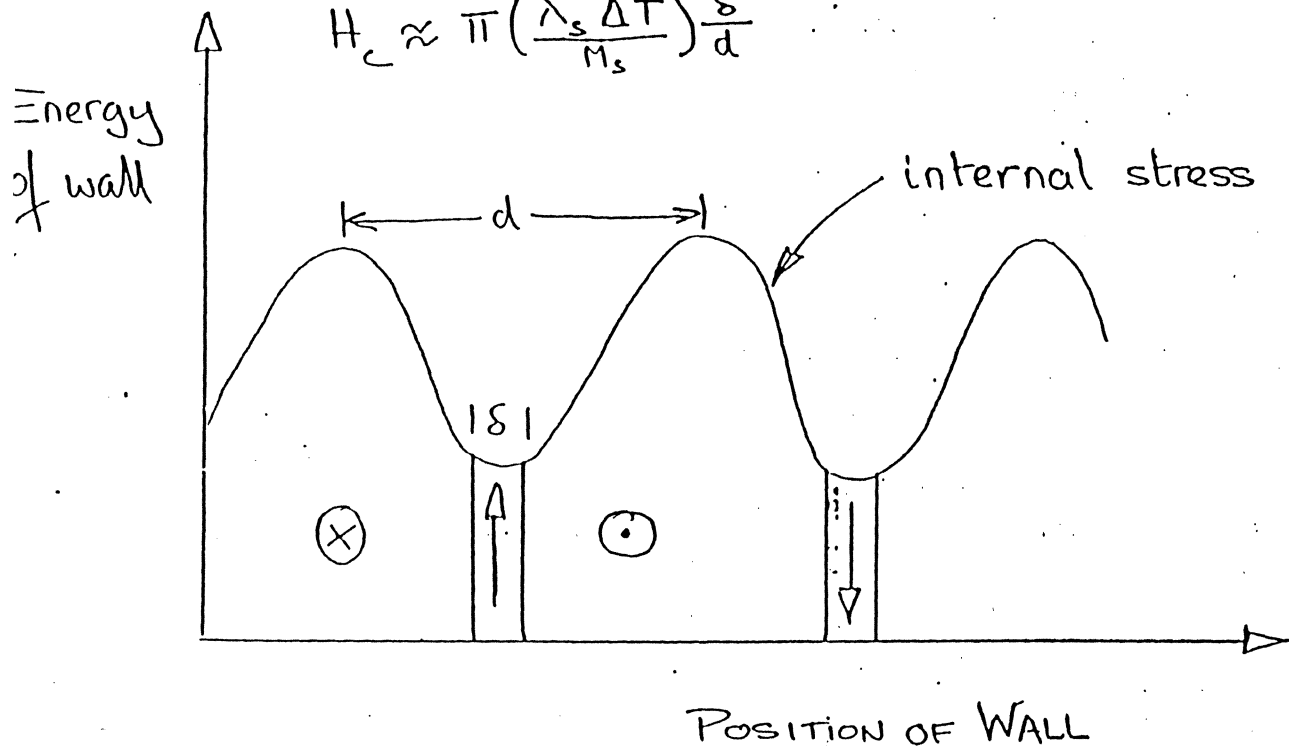
DOMAINS AND THE HYSTERESIS LOOP



ORIGINS OF COERCIVITY

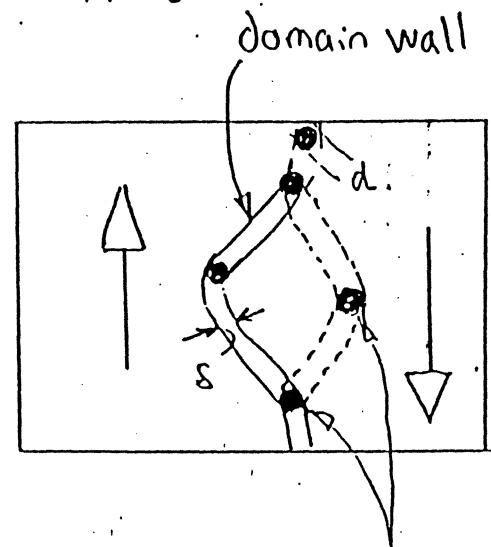
1) Internal stresses (Becker, 1932-9)

$$H_c \approx \pi \left(\frac{\lambda_s \Delta T}{M_s} \right) \frac{\delta}{d}$$



2) non-magnetic impurities (Kersten 1943)

$$H_c = \frac{k_i \delta}{M d} \propto \frac{2}{3}$$

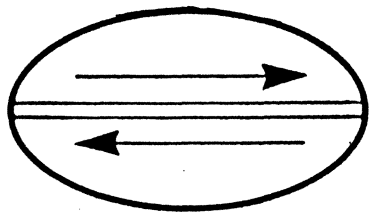


Impurities of volume concⁿ ∝

DOMAIN WALL MOVEMENT IS THE ORIGIN OF COERCIVITY IN SOFT MAGNETIC MATERIALS and it is found experimentally that work hardening and impurities increase the coercivity

Properties of Head Materials

	<u>Permalloy</u>	<u>Sendust</u>	<u>Ferrite</u>	<u>Ferrite</u>	<u>Amorphous Glass</u>
composition:	4Mo78Ni 17Fe	85Fe10Si5AL	(Fe,Mn,Zn) ₃ O ₄	(Fe,Ni,Zn) ₃ O ₄	Co ₇₀ Fe ₅ Si ₁₅ B ₁₀
B _s , gauss	10,000	10,500	5,000	3,300	14,000
μ _i (1KHz)	10,000	10,000	5,000	5,000	10,000
resistivity, Ωcm.	55x10 ⁻⁶	80x10 ⁻⁶	5	10 ⁵	134x10 ⁻⁶



← 5 μ →

Multi-Domain

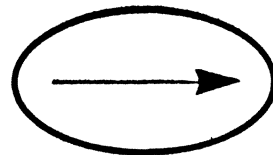
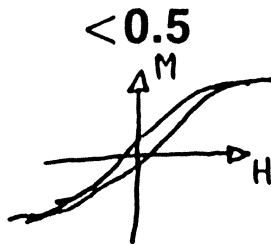
coercivity:

Hc < 100 Oe

Ms

400 – 1700 emu/cc

$\frac{Mr}{Ms}$



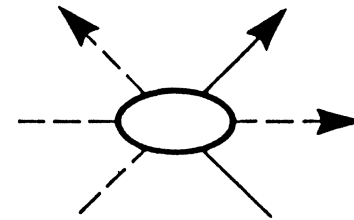
← < 1 μ →

Single Domain

Hc = 100 – 3000

400 – 1700

0.5 < $\frac{Mr}{Ms}$ < 1.0

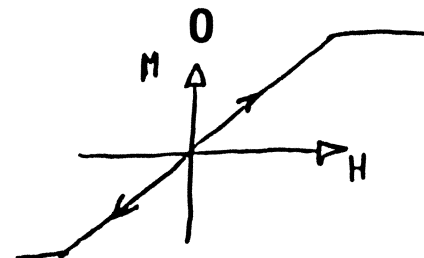


< 0.02 μ
→ ←

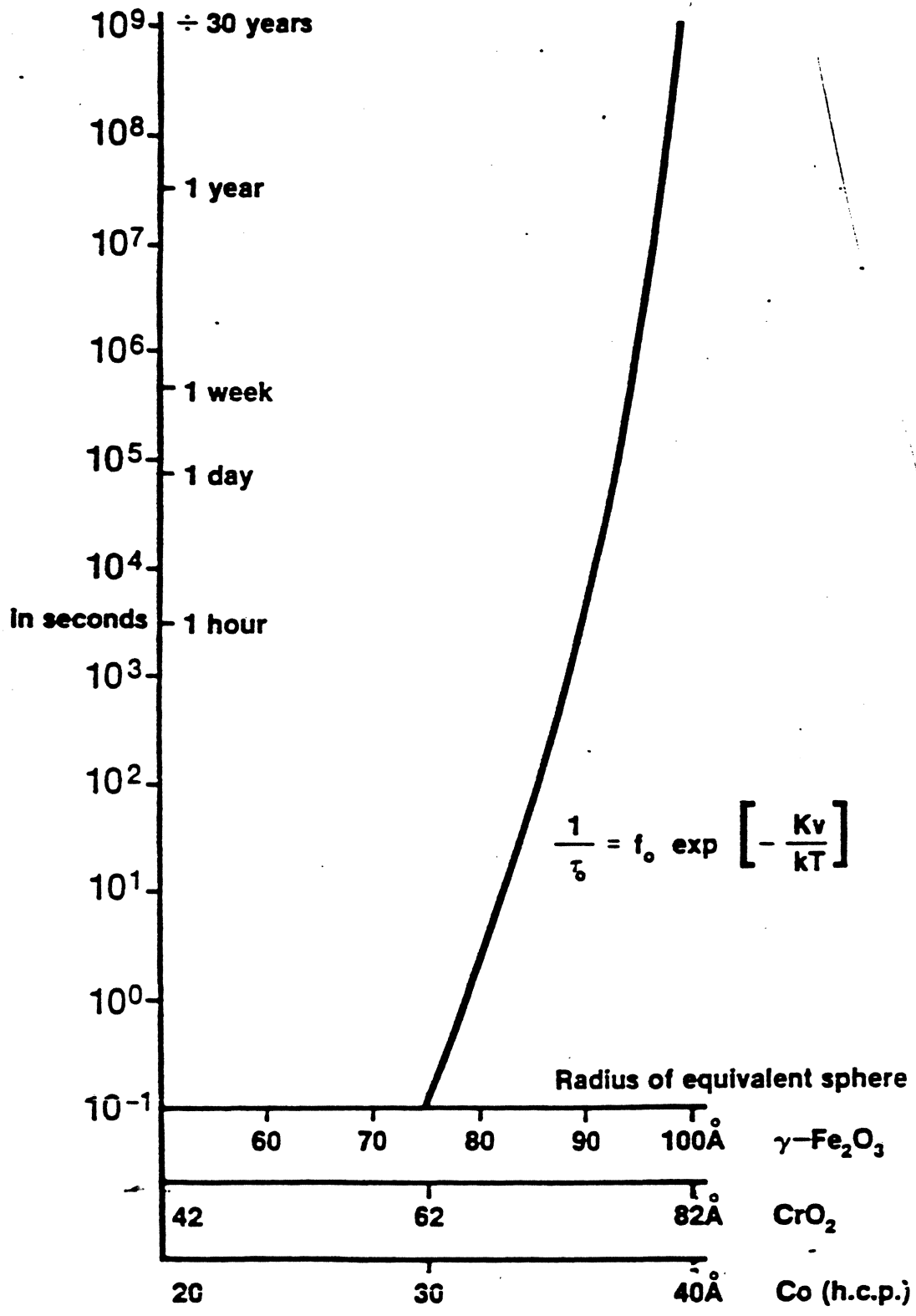
Superparamagnetic

Hc = 0

400 – 1700

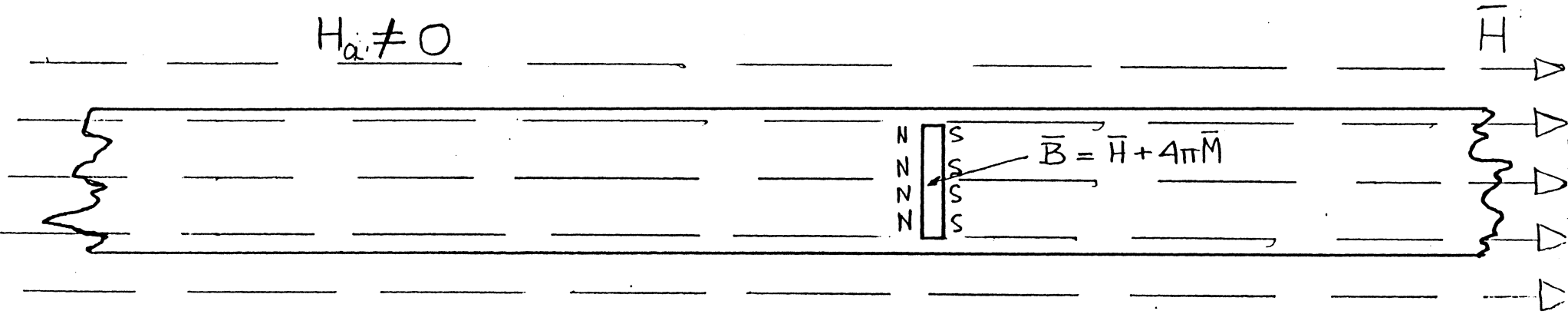


RELAXATION TIME FOR ACICULAR PARTICLES

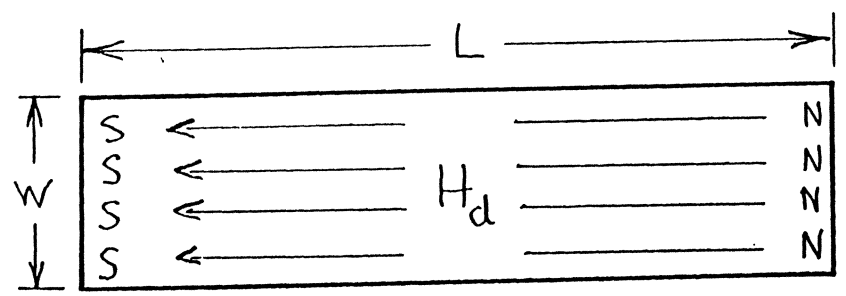


MAGNETIZING AND DEMAGNETIZING FIELDS

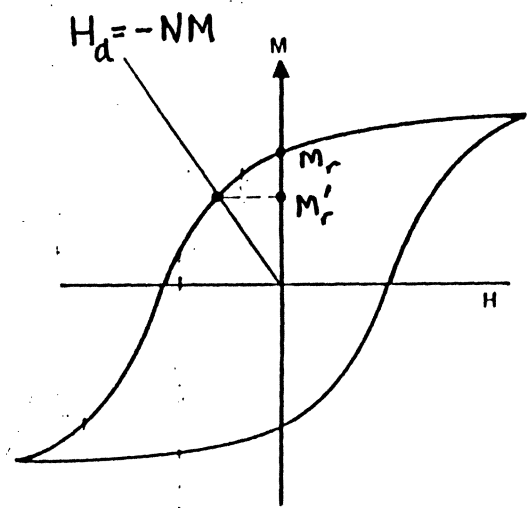
INFINITE BAR IN ZERO APPLIED FIELD, $H_a = 0 = M = H_d$



FINITE LENGTH OF MATERIAL

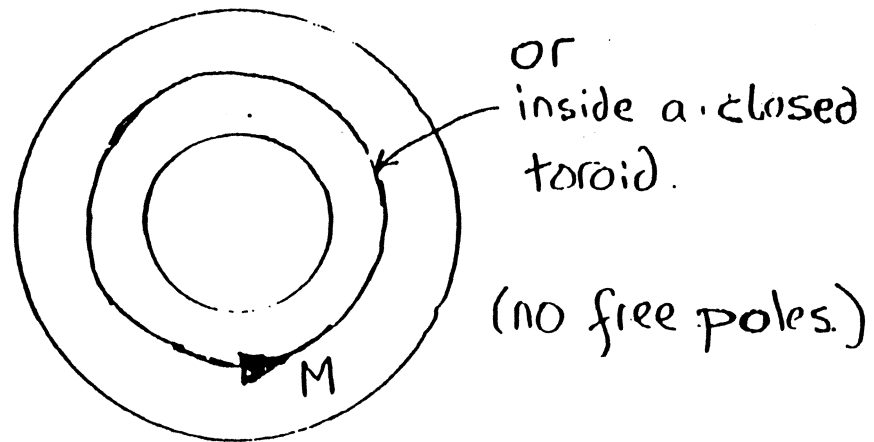


$$H_d = -N\left(\frac{W}{L}\right)M$$

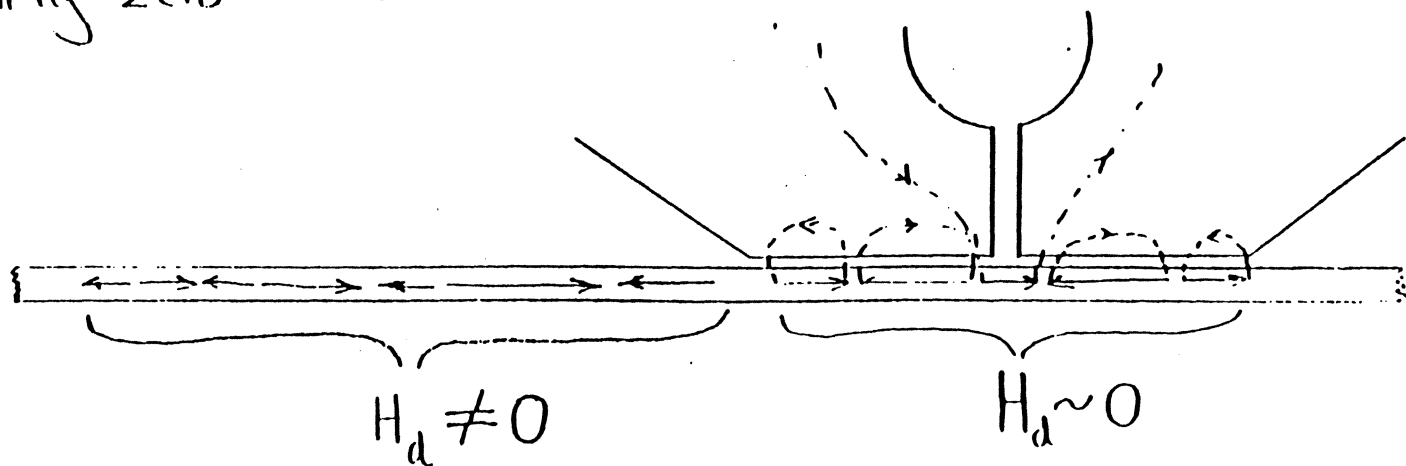


Demagnetizing Fields, H_d

- are equal to zero .
- at the center of a long bar



- are nearly zero in a recorded medium in contact with a head

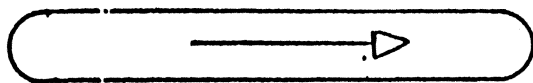


SHAPE ANISOTROPY

$$\Theta = 0^\circ$$

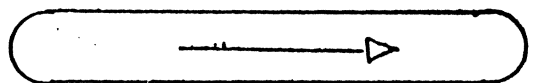
← Easy Axis →

Field
→
 $H = +3,000 \text{ Oe}$



$$M = +M_s$$

$H = 0$



$$M = +M_s = M_r$$

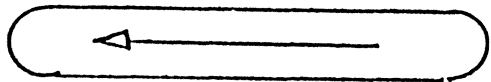
←
 $H = -2,400 \text{ Oe}$



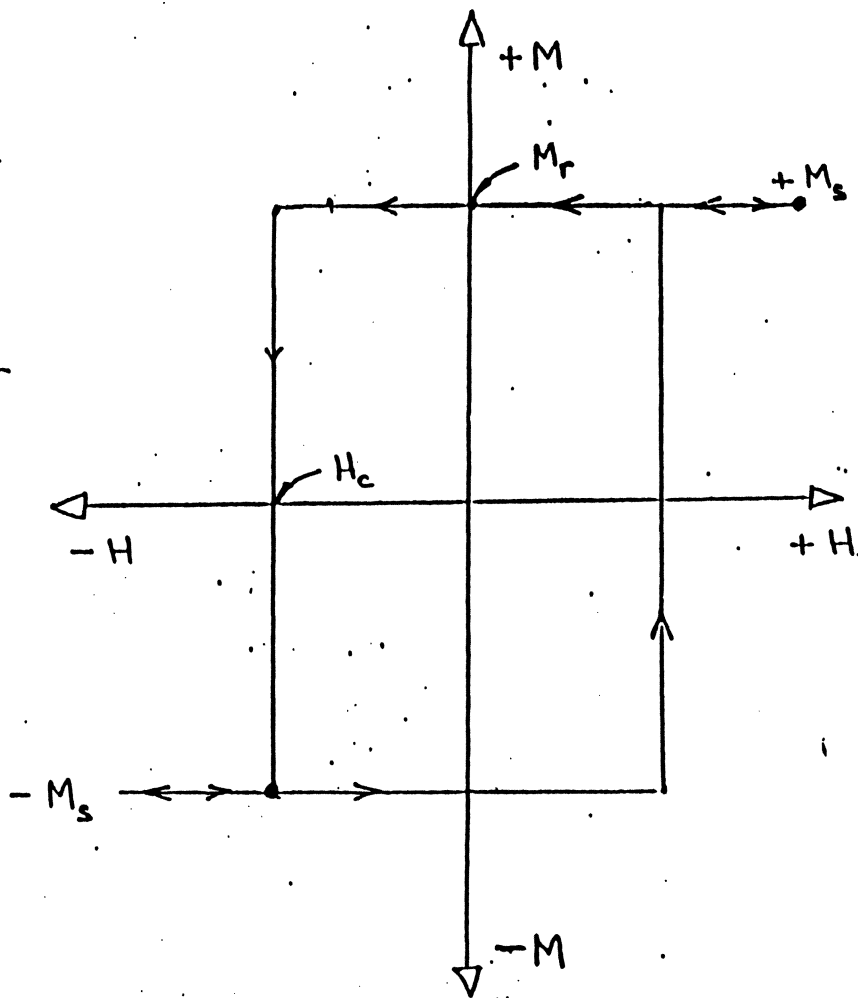
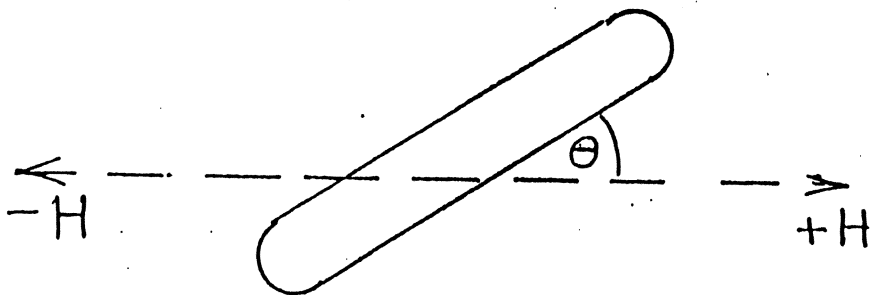
$$M = 0$$

$$H_c = 2,400 \text{ Oe.}$$

←
 $H = -3,000 \text{ Oe.}$



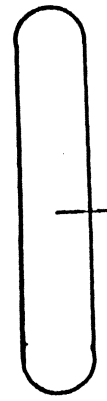
$$M = -M_s$$



SHAPE ANISOTROPY

$$\theta = 90^\circ$$

$H = +3,000 \text{ Oe}$



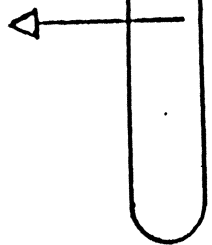
↑ Easy axis
↓
→ $M = M_s$

$H = 0$

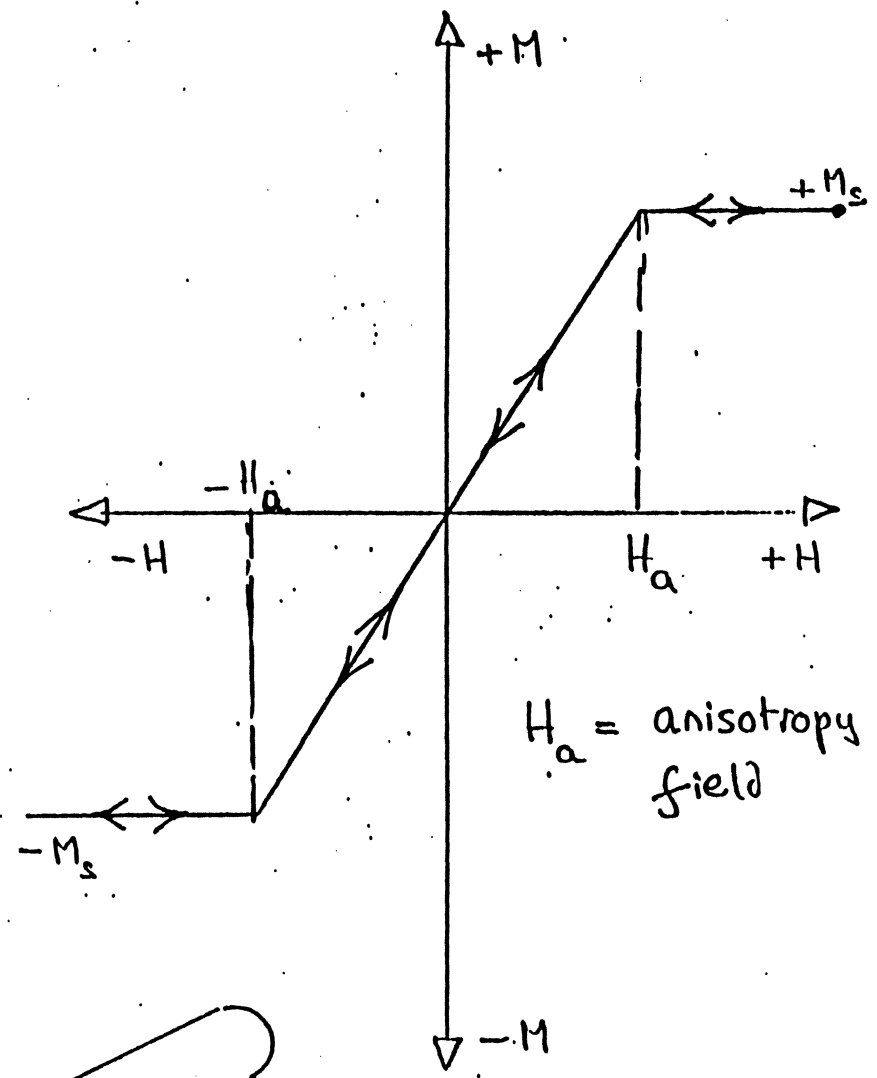
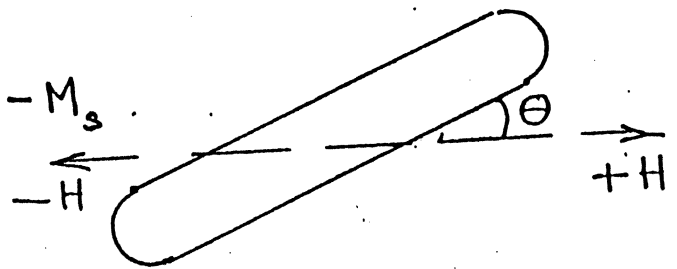


$M = 0 = M_r$

$H = -3,000 \text{ Oe}$



$M = -M_s$

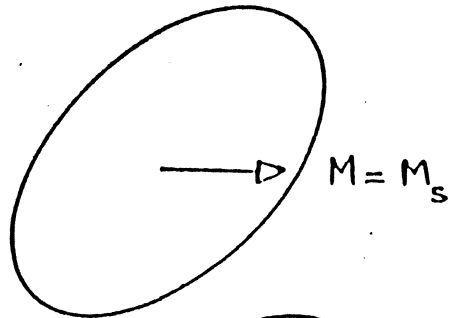


SHAPE ANISOTROPY

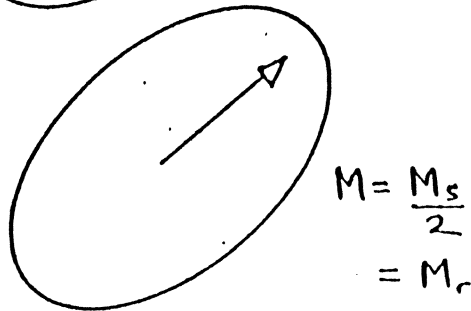
$$\theta = 45^\circ$$

Easy axis

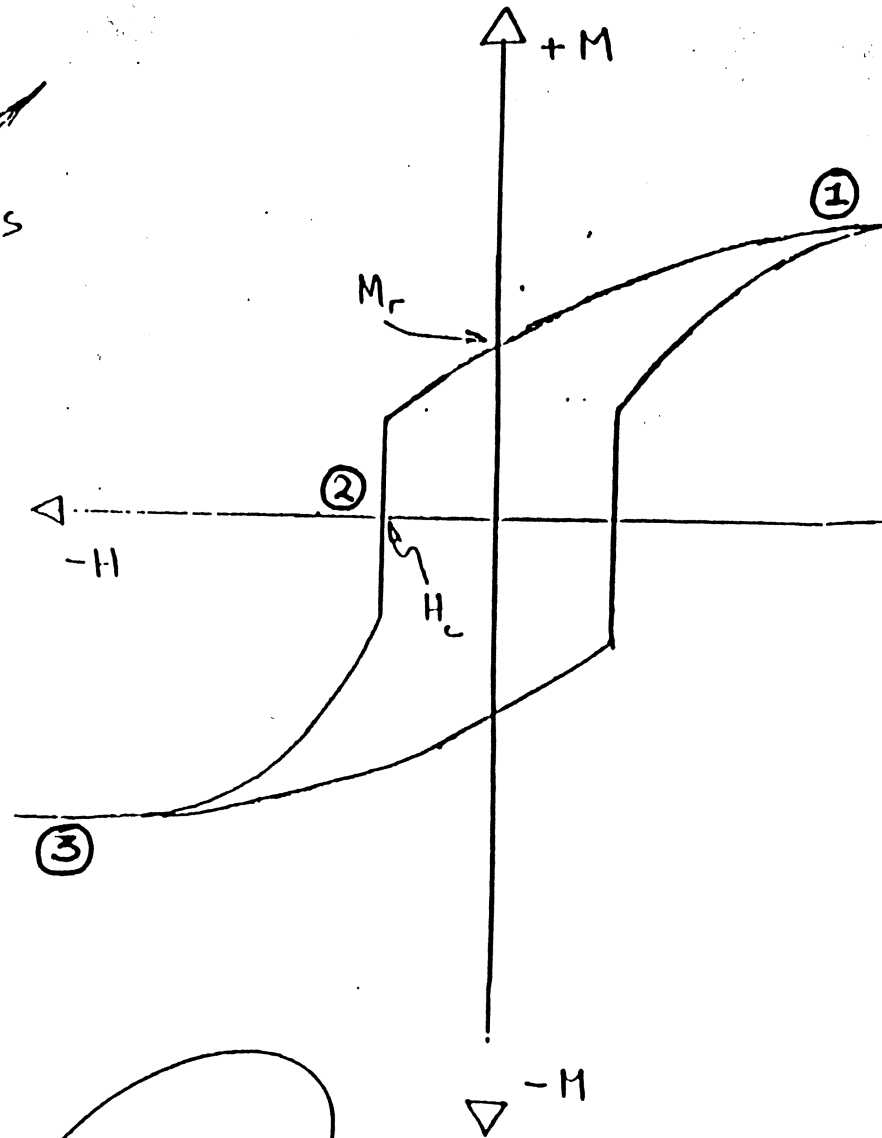
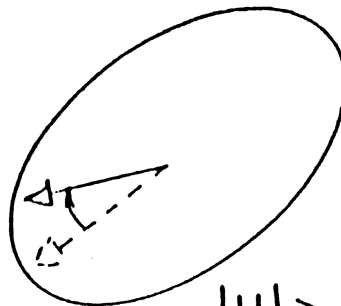
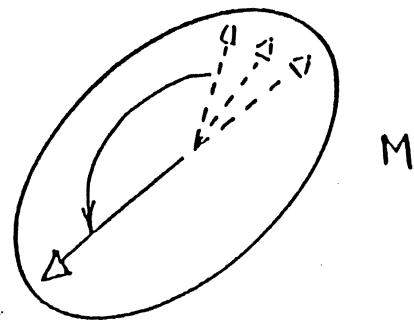
① $H = +3000 \text{ Oe}$



② $H = 0$



③ $H = H_c$



Néel 1947

Stoner and Wohlfart 1947

SINGLE DOMAIN PARTICLES

Anisotropy

Maximum Coercivity

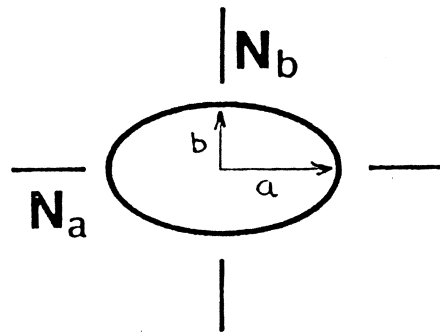
Crystalline

$$H_c = \frac{2 k_1}{M_s}$$

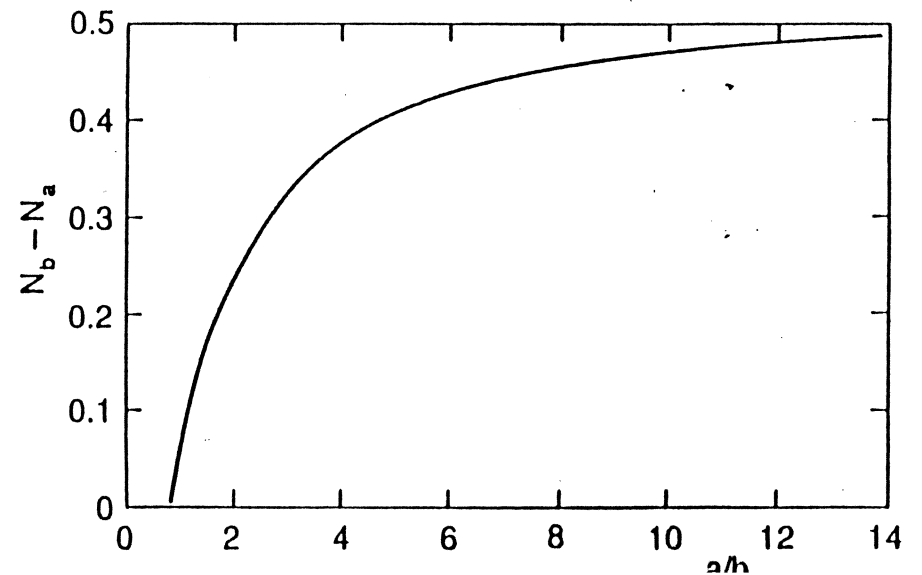
Strain

$$H_c = \frac{3 \lambda_s T}{M_s}$$

Shape



$$H_c = (N_b - N_a) M_s$$

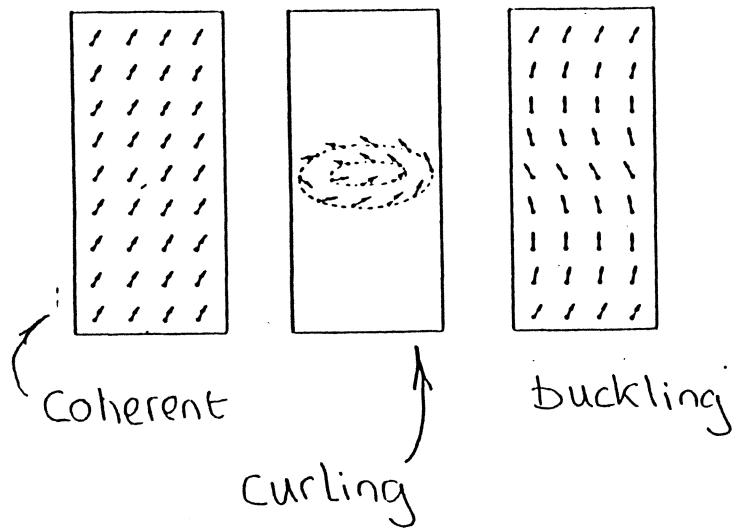


SINGLE DOMAIN PARTICLES

Maximum Coercivity

	Iron	Cobalt	Nickel	$\gamma\text{-Fe}_2\text{O}_3$
Crystalline	250	3,000	70	230
Strain	300	300	2,000	<10
Shape (10:1)	5,300	4,400	1,550	2,450

INCOHERENT MAGNETIZATION REVERSAL

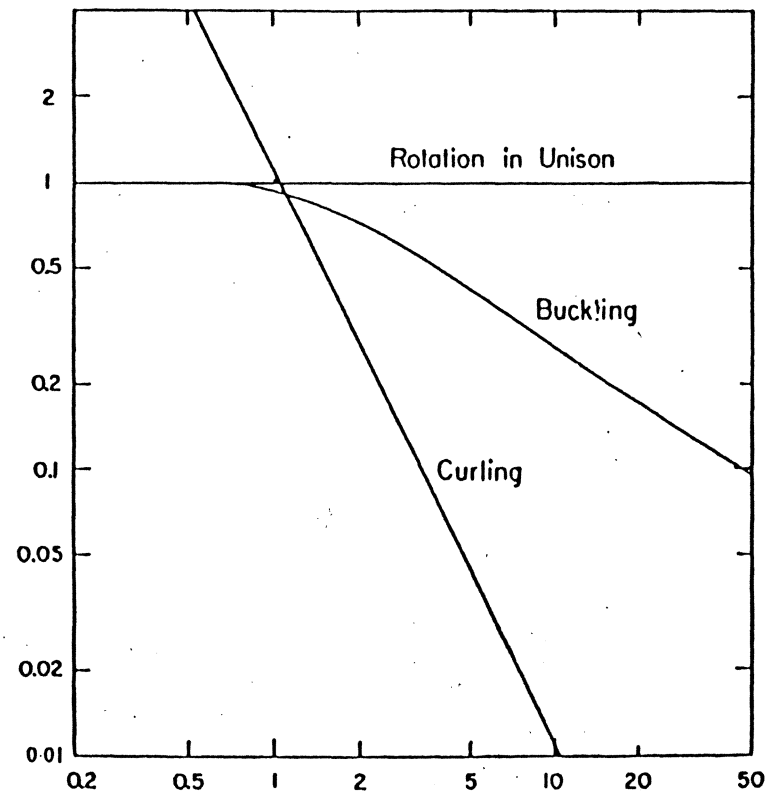


infinite cylinders

INCOHERENT REVERSAL

- H_c values closer to experiment
- H_c depend on particle dimensions below the critical dimensions for single-domain behavior.

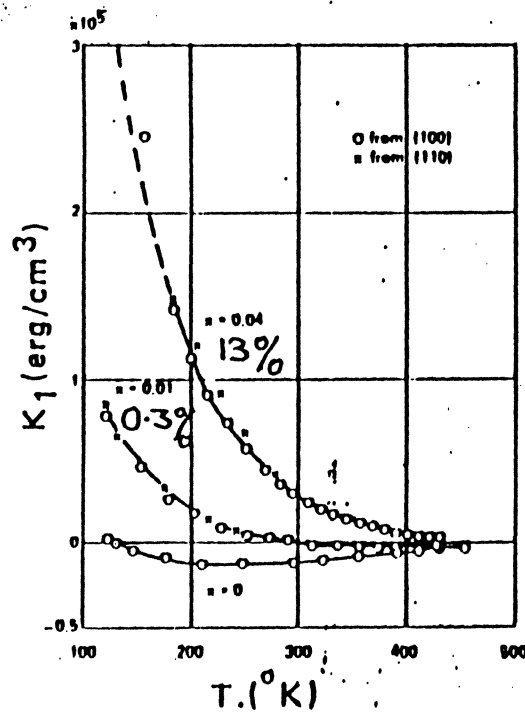
$$\frac{H_c}{2\pi M_s}$$



$$S = \frac{\text{cylinder radius}}{\text{cylinder length}}$$

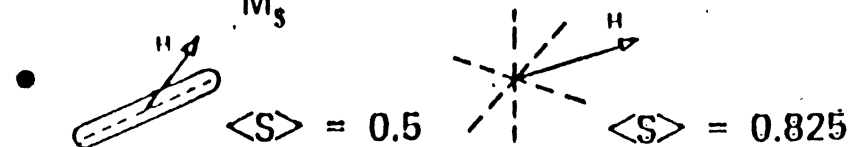
Cobalt-Modified Iron Oxides

Cobalt substitution

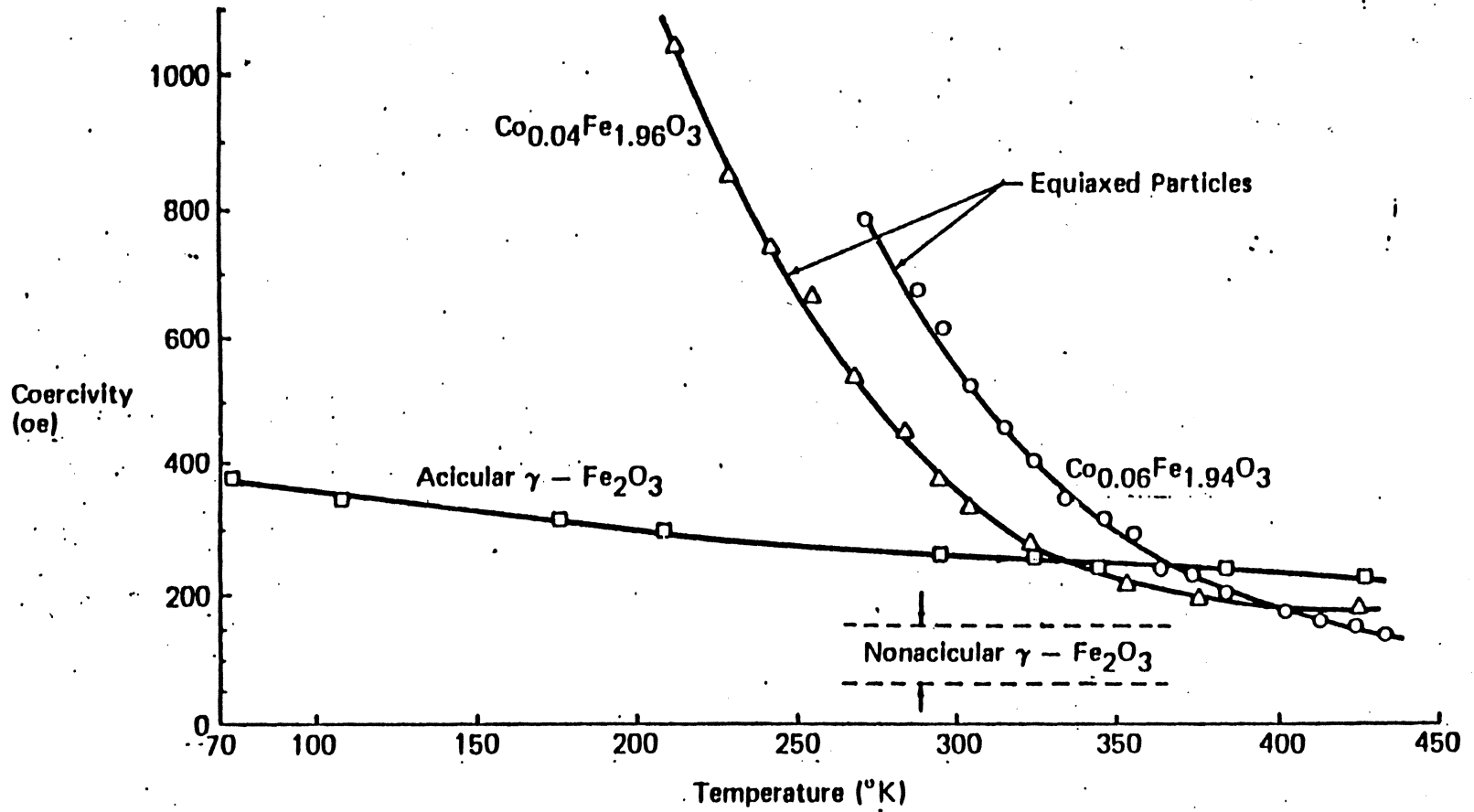


$\text{Co}_x\text{Fe}_{3-x}\text{O}_4$
Bickford et al. (1957)

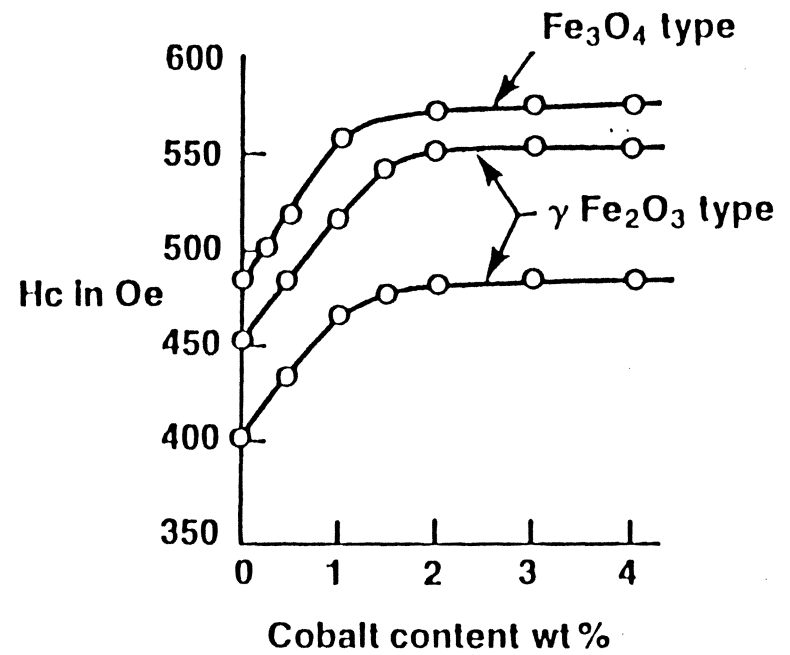
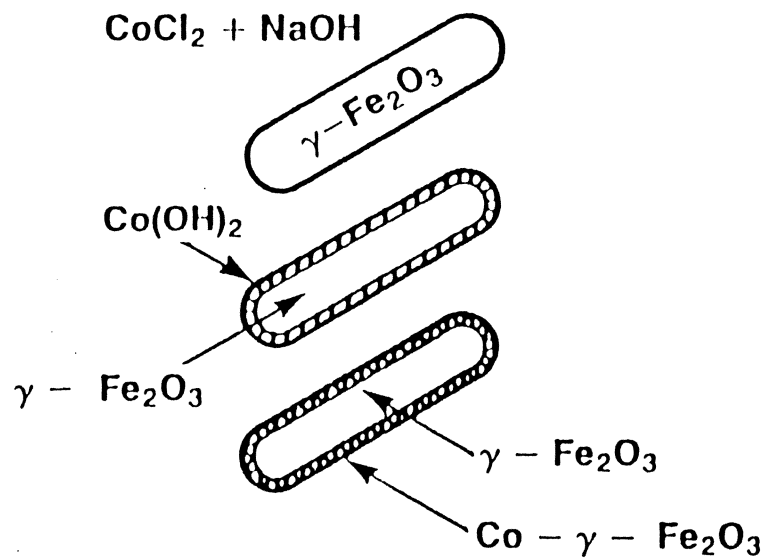
- $H_c = 0.64 \frac{K_1}{M_s}$, temperature sensitive



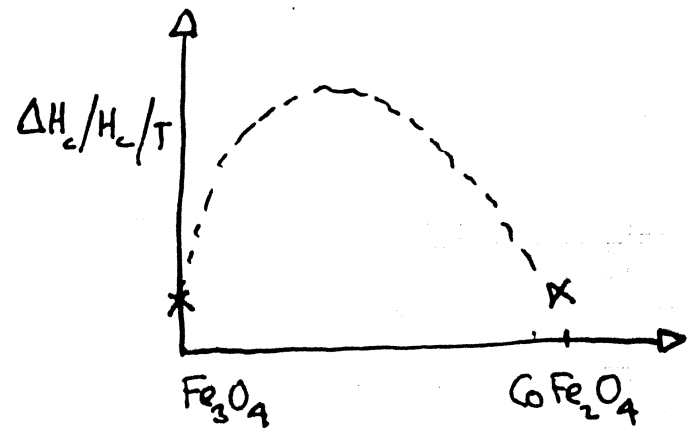
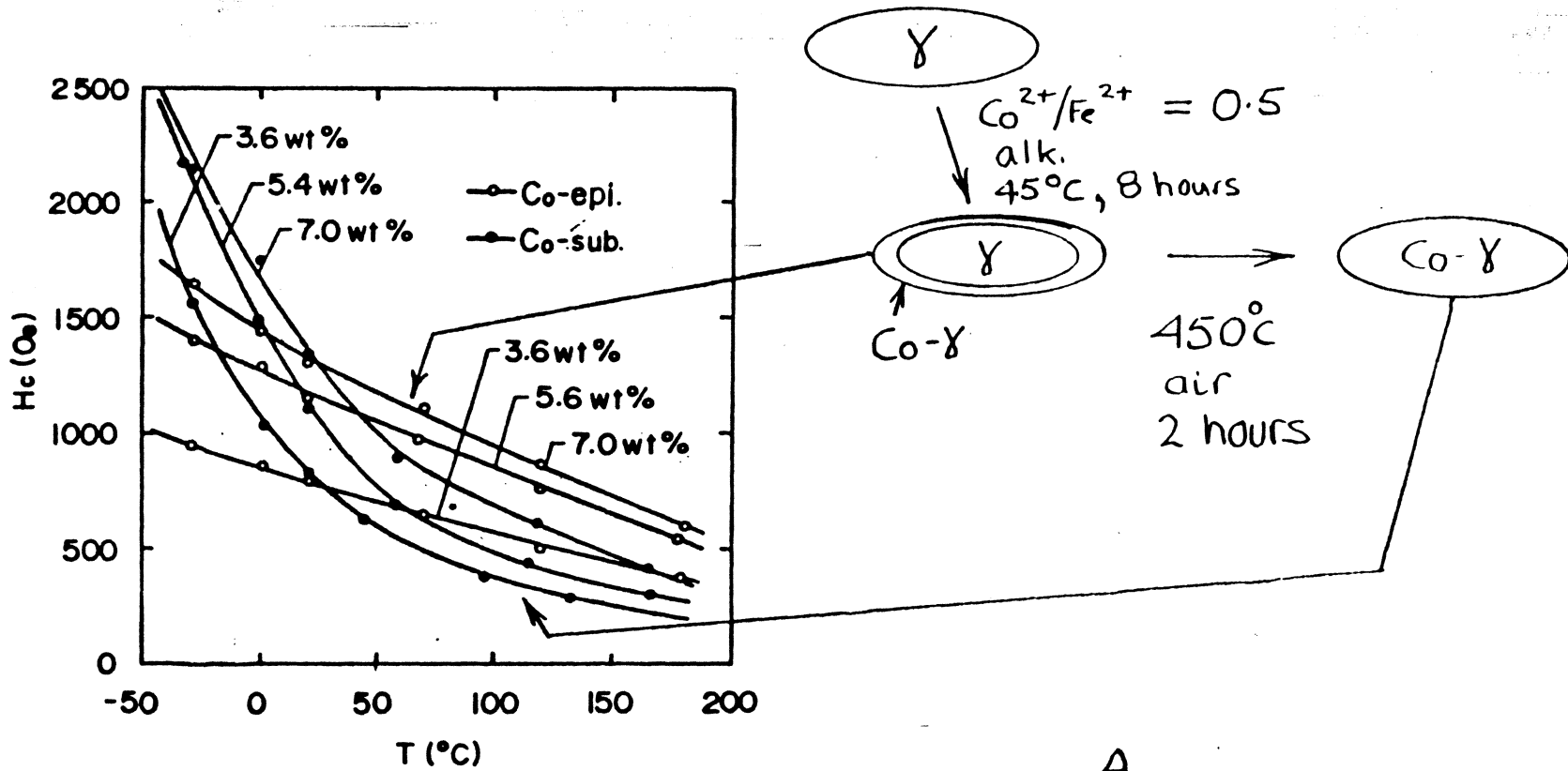
- Remanence time and stress dependent



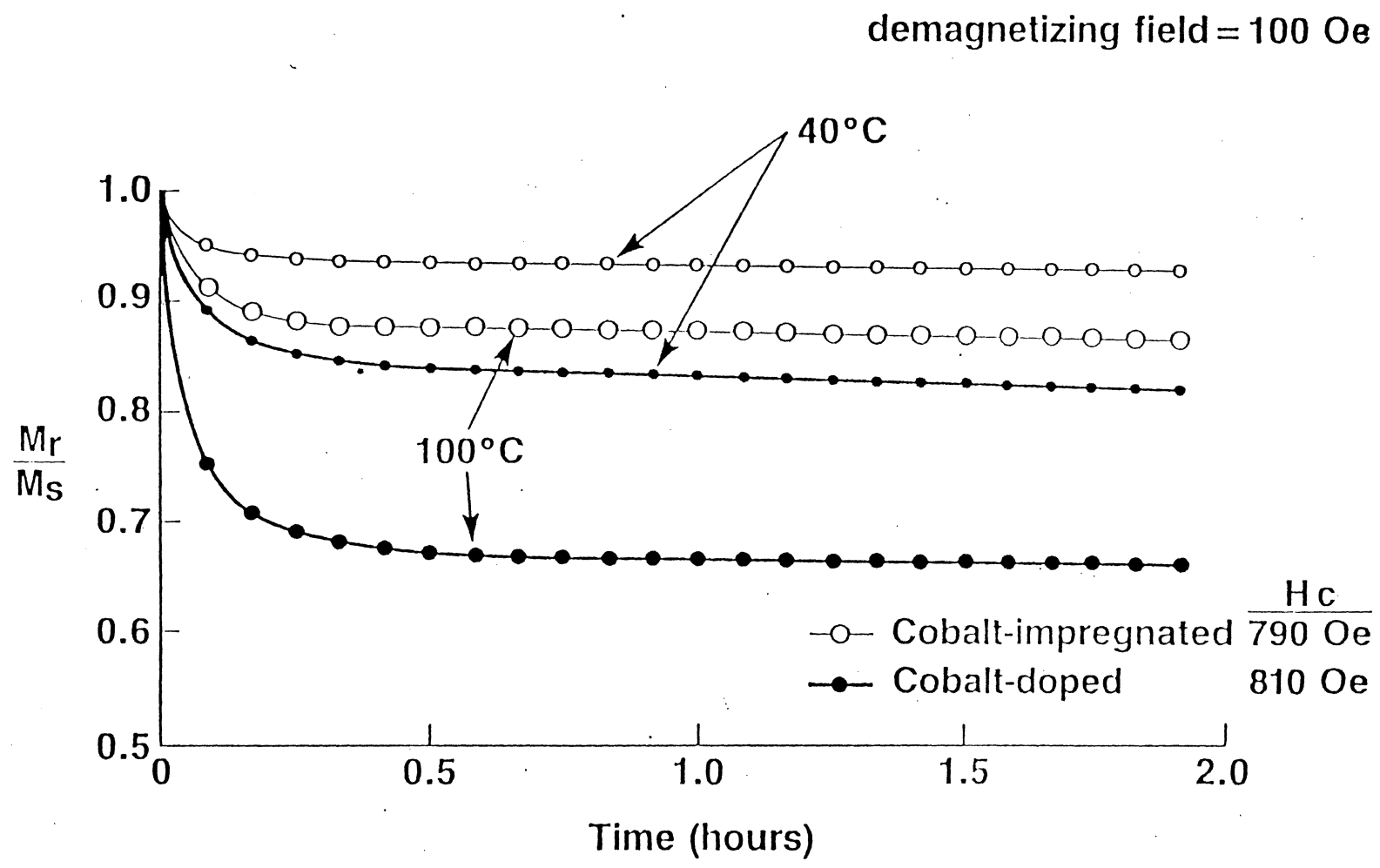
COBALT-IMPREGNATED IRON OXIDE



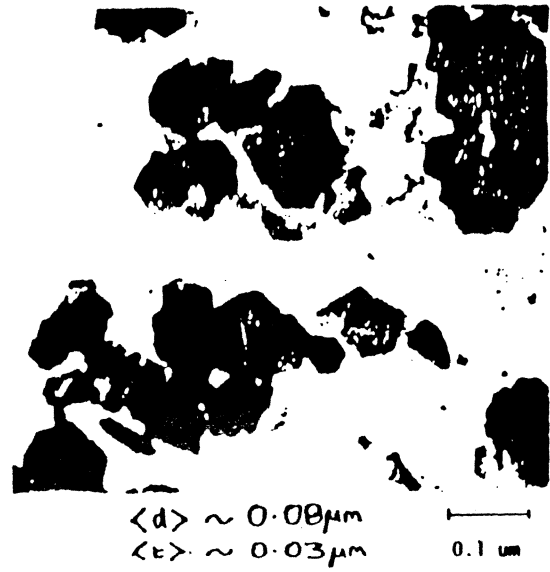
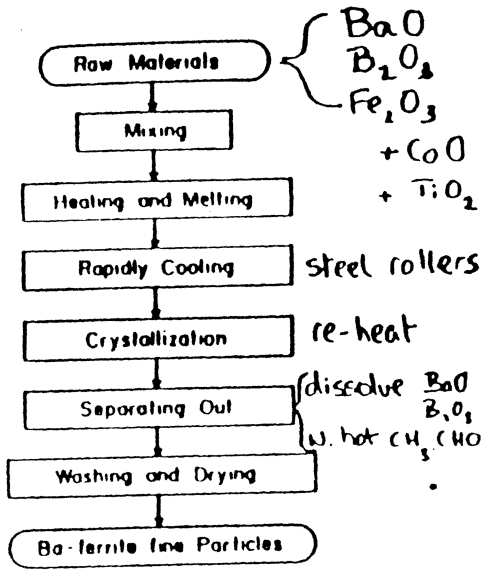
COBALT - MODIFIED γ - Fe_2O_3 PARTICLES



(Kishimoto, Amemiya, Hayama.)



PREPARATION AND PROPERTIES OF BARIUM FERRITE PARTICLES

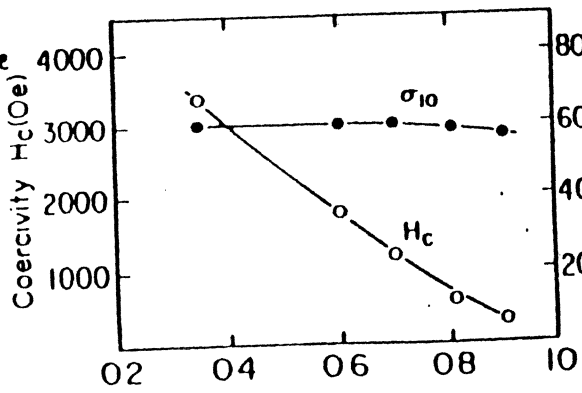


Block diagram showing the Ba-ferrite particle preparation process using the glass crystallization method. \rightarrow Uniform nucleation, \rightarrow small, well-separated, narrow size range

Typical Ba-ferrite fine particles characteristics

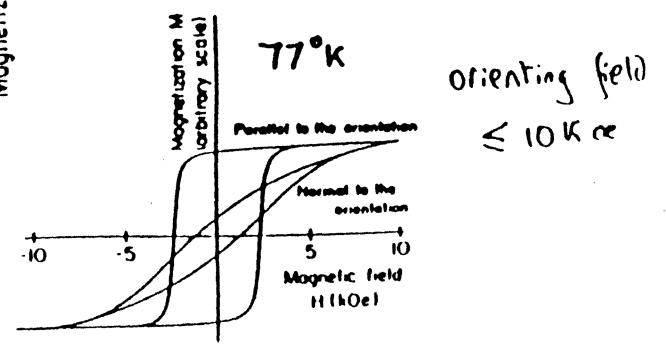
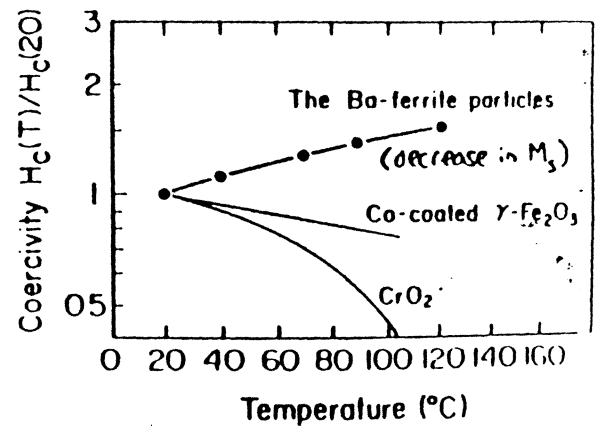
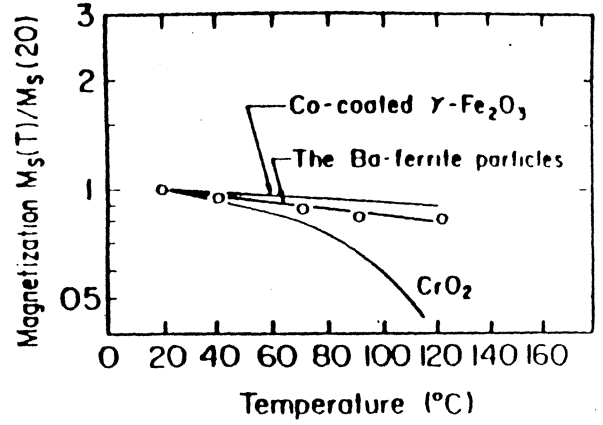
Particle size	0.08 x 0.03 μm
Density	5.25 g/ml
Specific surface area	22 m ² /g
Coercivity	900 Oe
Saturation magnetization	58 emu/g
Curie temperature	350°C
Squareness ratio	0.94
θ_c	320°C

- Co, Ti allow control of H_c w/o changes in d, t.
- easy axis // c-axis.



$$H_c = 0.48 \left(\frac{2K}{M_s} - NM_s \right)$$

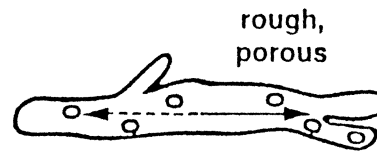
20°C, x = 0.8 \rightarrow 6K α 4K α



MATERIALS FOR VERY HIGH DENSITY RECORDING

VHD PARTICLES

Conventional Gamma-Iron Oxide
(Co-impregnated)
Example: HDX particles



RECORDING MODE

Longitudinal

"Rice Grain" Gamma-Iron Oxide
(Co-impregnated)



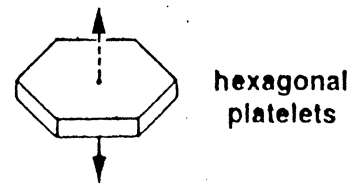
Longitudinal

Isotropic Gamma-Iron Oxide
(Co-doped)



Longitudinal/Perpendicular

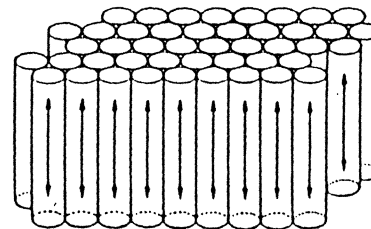
Barium Ferrite



Perpendicular

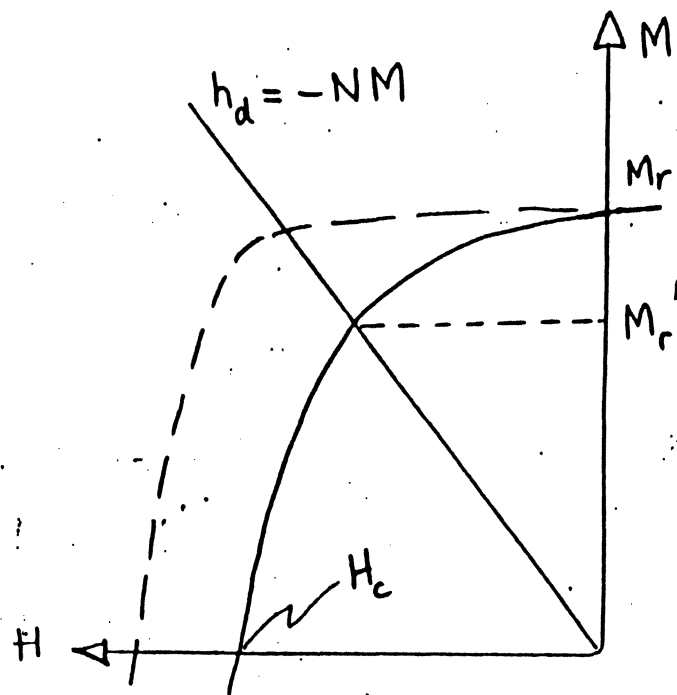
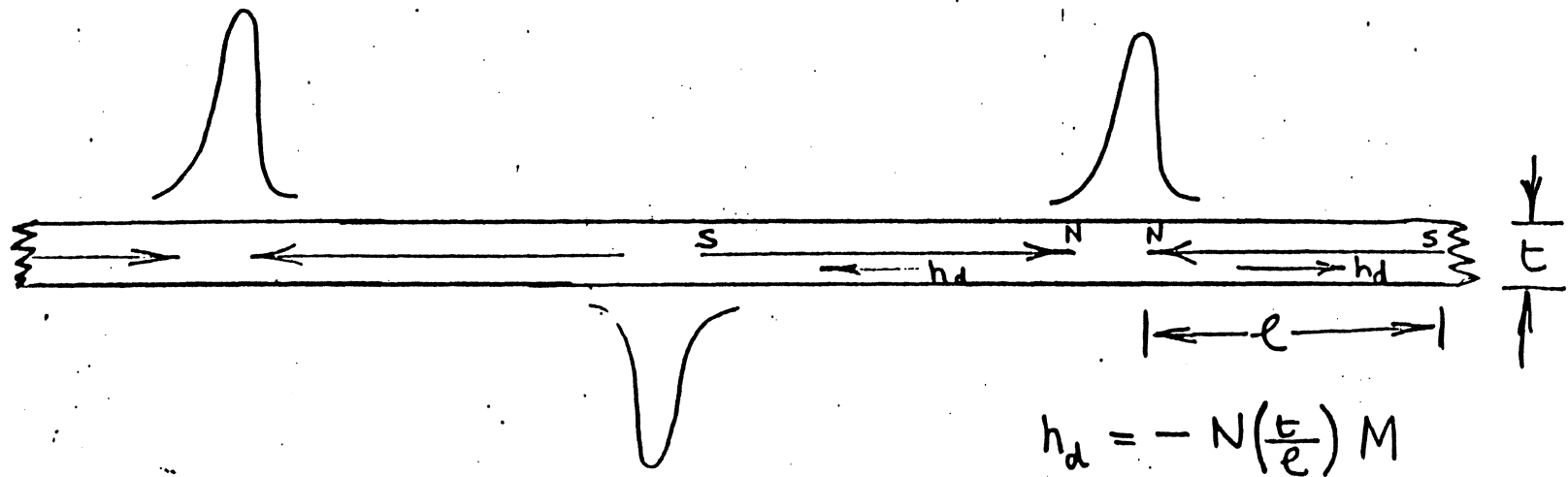
VHD THIN FILMS

Cobalt - Chrome
(sputtered)



Perpendicular

ADVANTAGE OF HIGH COERCIVITY.



increasing the coercivity
reduces the demagnetization:

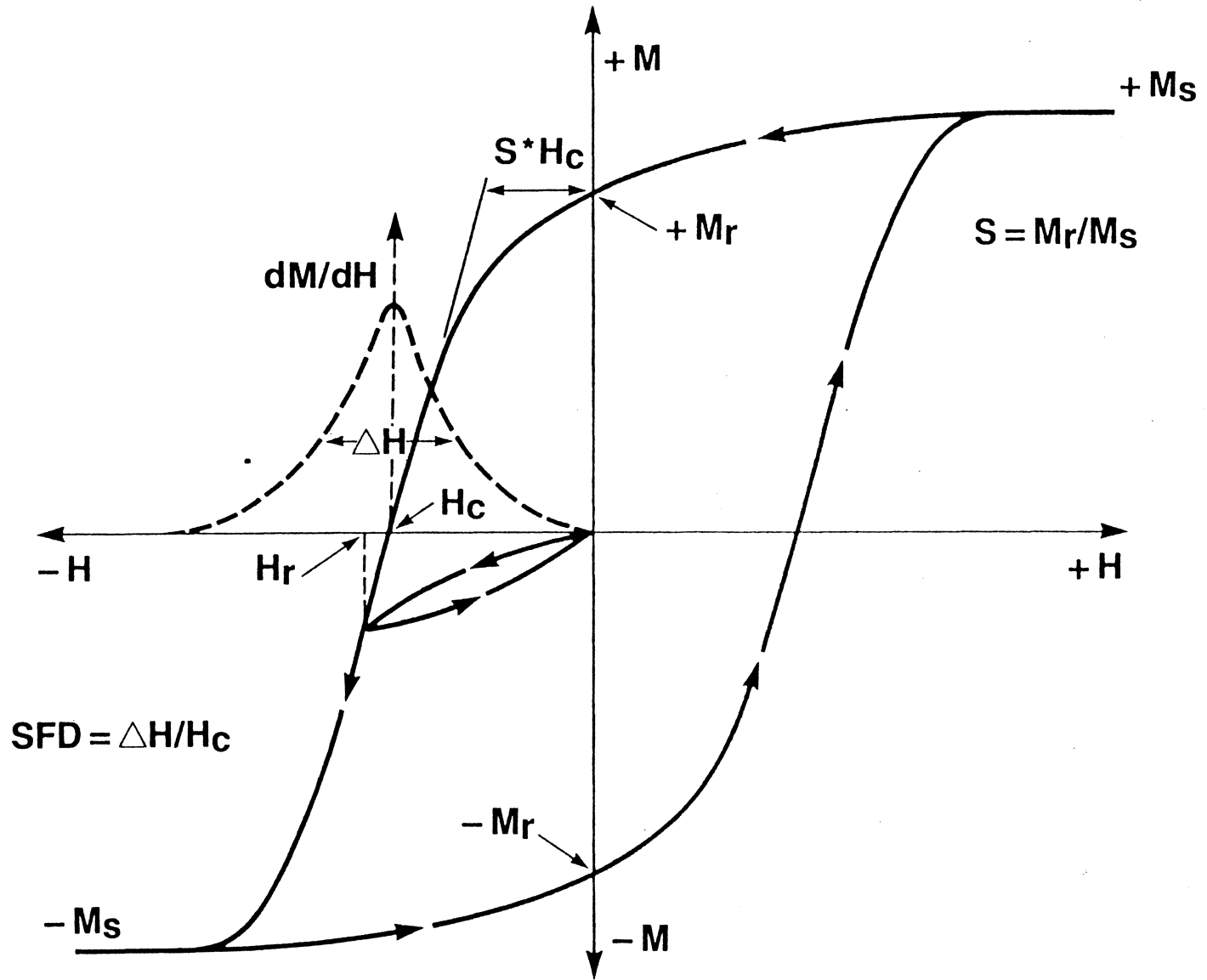
but

the higher the coercivity the greater
the field needed from the writing head

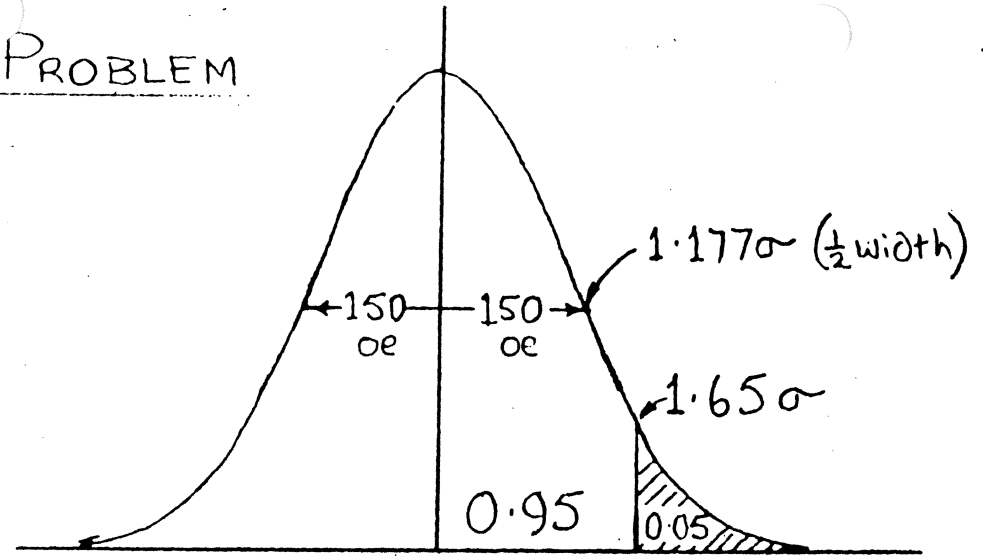
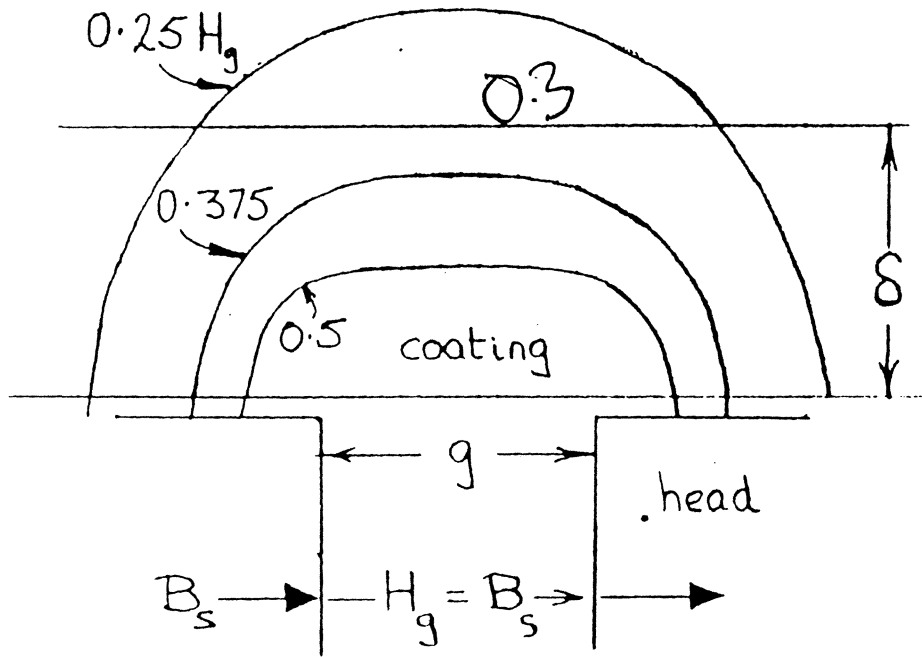
The maximum usable coercivity is
determined by the M_s of the material
from which the head is made.

HARD AND SOFT MAGNETIC MATERIALS

SOFT	<u>High Ms</u>	<u>Low Hc</u>	<u>Low Mr</u>	<u>High μ</u>
Fe	1700 emu / cc	1 Oe	< 500	20,000
80 Ni 20 Fe	660	0.1	< 300	50,000
Mn Zn Ferrite	400	0.02	< 200	5,000
Co ₇₀ Fe ₅ Si ₁₅ B ₁₀	530	0.1	< 250	10,000
HARD	<u>High Ms</u>	<u>High Hc</u>	<u>High Mr</u>	<u>Tc</u>
Particles γ -Fe ₂ O ₃	400	250 - 450	200 - 300	115 - 126
" CrO ₂	400	450 - 600	300	120
" Fe	870 - 1100	1,100 - 1,500	435 - 550	768
" BaO.6Fe ₂ O ₃	238 - 370	800 - 3,000	143 - 260	320
Alloys SmCo ₅	875	40,000	690	720
" Sm ₂ Co ₁₇	1,000	17,000	875	920
" Fe ₁₄ B Nd ₂	1,020	12,000	980	310



THE OVERWRITE PROBLEM



$H_c = 1000$ $H_r = 1050 oe$ $1260 oe$

Assumptions: • Normal Distribution.

• S.F.D. = 0.3

• coating thickness, $\delta =$ head gap, g .

Result

$0.3 H_g = 1260$

$H_g = 4,200$

Overwrite test:

- 1) write 1F, measure amplitude
- 2) overwrite with 2F
- 3) read remaining 1F amplitude

Overwrite criterion: $-26dB$

$\frac{\text{remaining 1F}}{\text{original 1F}} \leq 0.05$

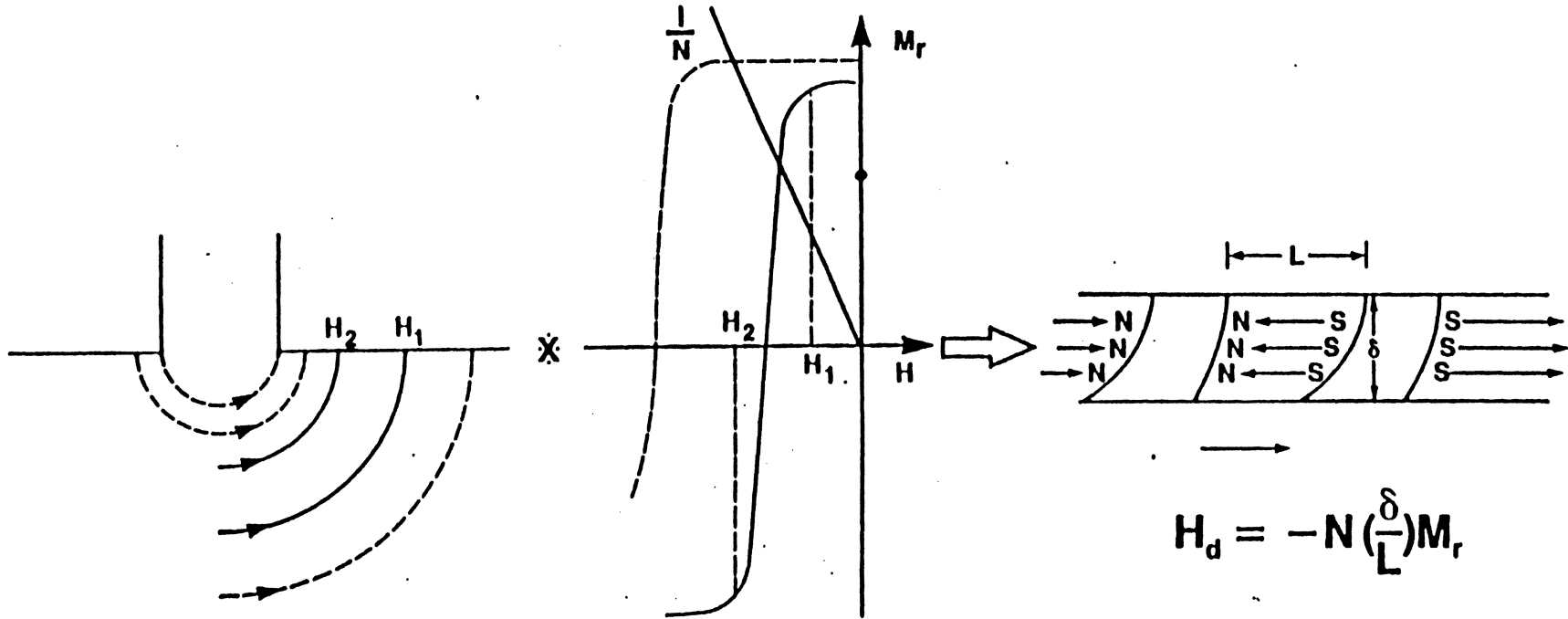
Ferrite	B_s
Ni Zn	4,000
Mn Zn	5,000

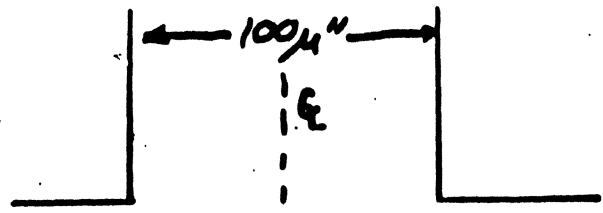
writing

$$\frac{dH}{dx} \times \frac{dM}{dH} \Rightarrow \frac{dM}{dx} \xrightarrow{\text{storage.}} \frac{d\phi}{dx} = \frac{dt}{dx} \cdot \frac{d\phi}{dt}$$

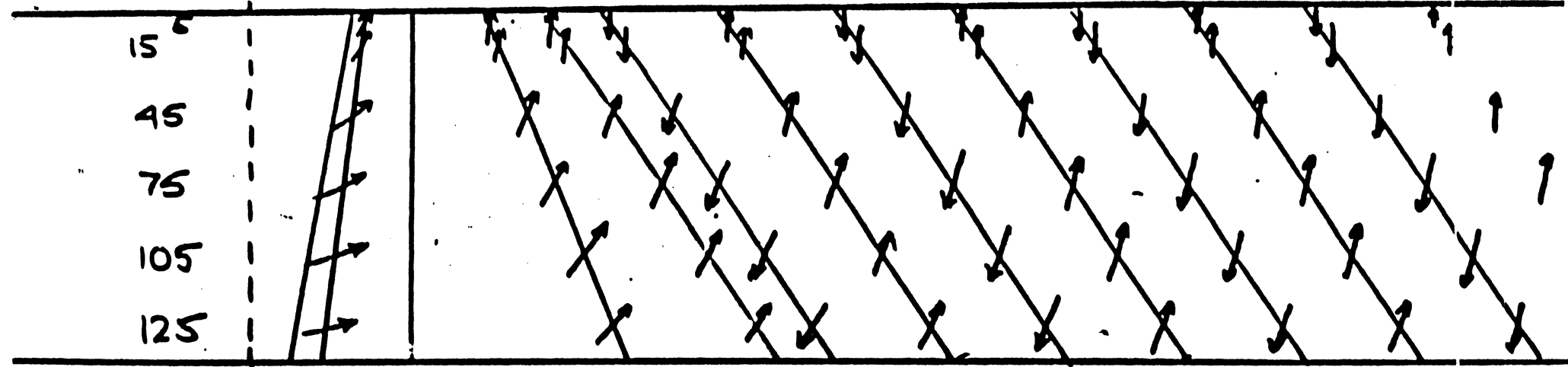
$$= \left(\frac{1}{\text{vel}}\right) \nabla_{\text{out.}}$$

Writing Process





20,000 f.c.i



H₂

H_r

H₁

CrO₂

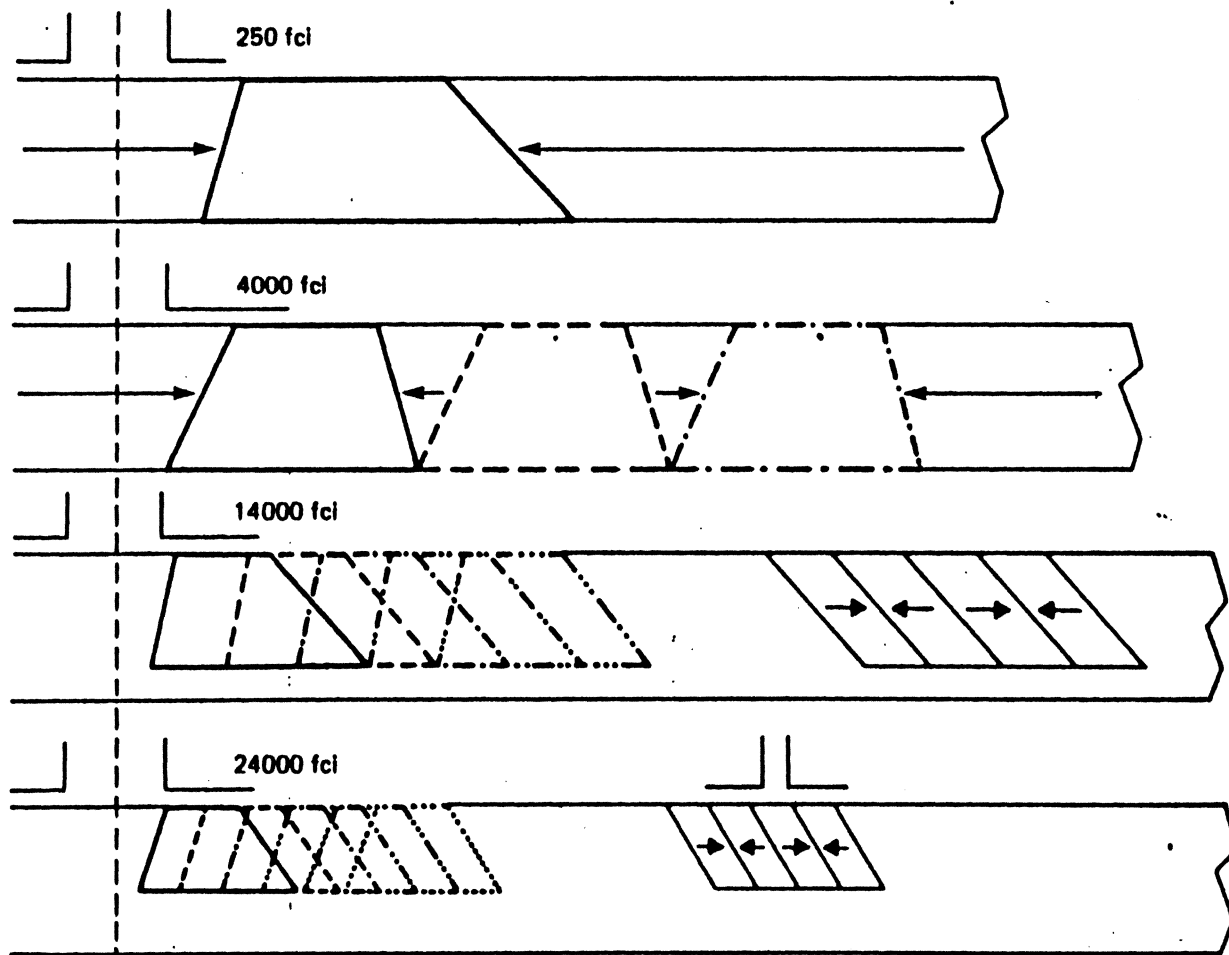
849

490

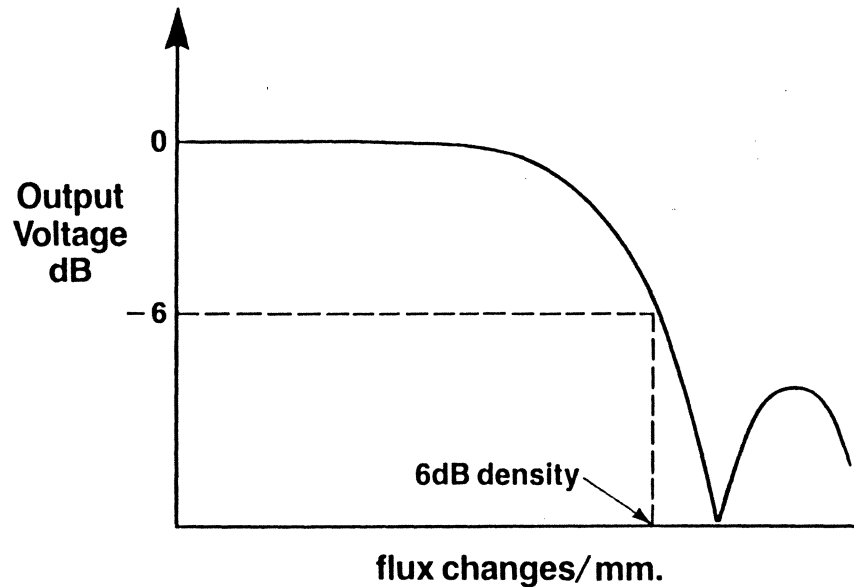
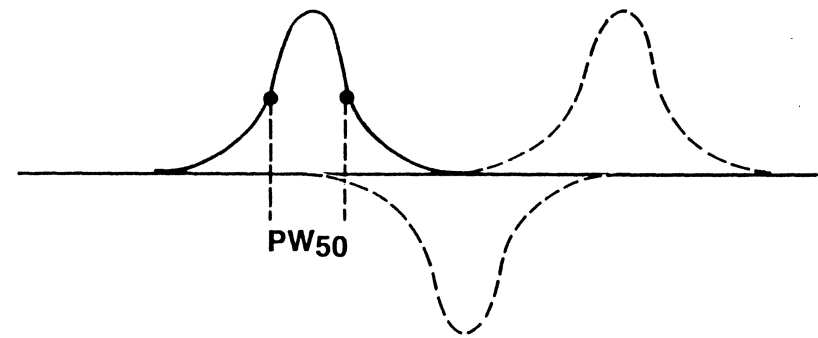
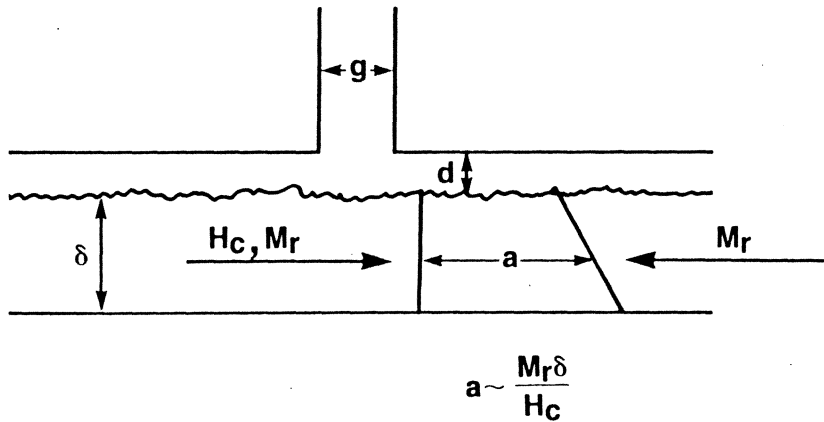
234

RECORDED TRANSITIONSTape: CrO₂, 150 μ"

Head gap = 100 μ"



Linear Density: Longitudinal Recording

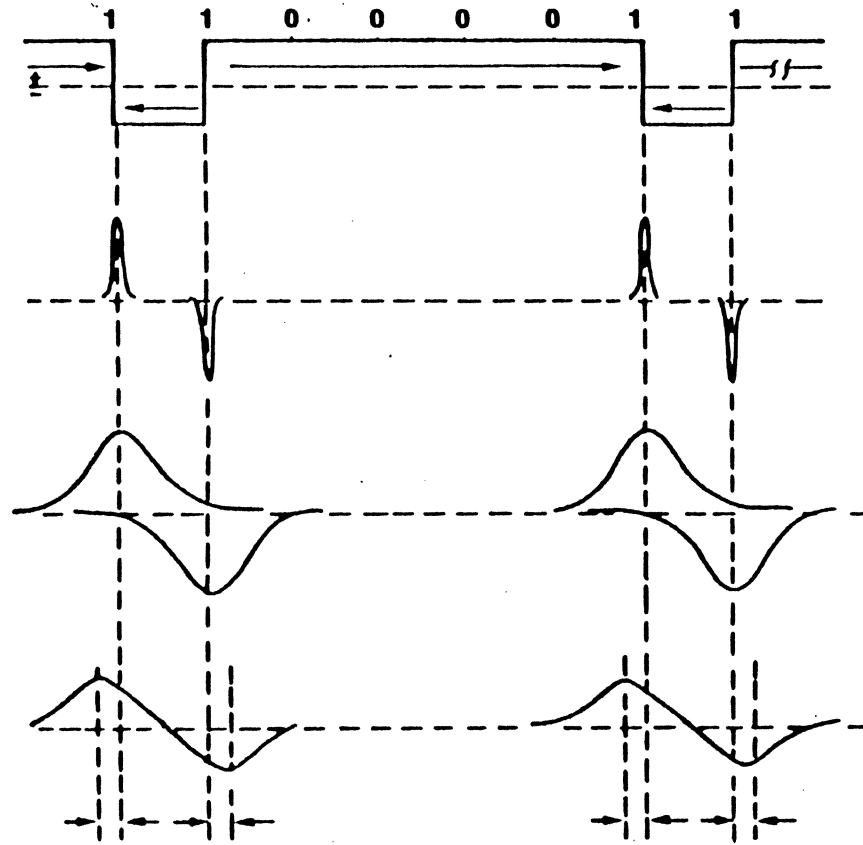


$$6\text{dB density} \approx \frac{1.4}{PW_{50}}$$

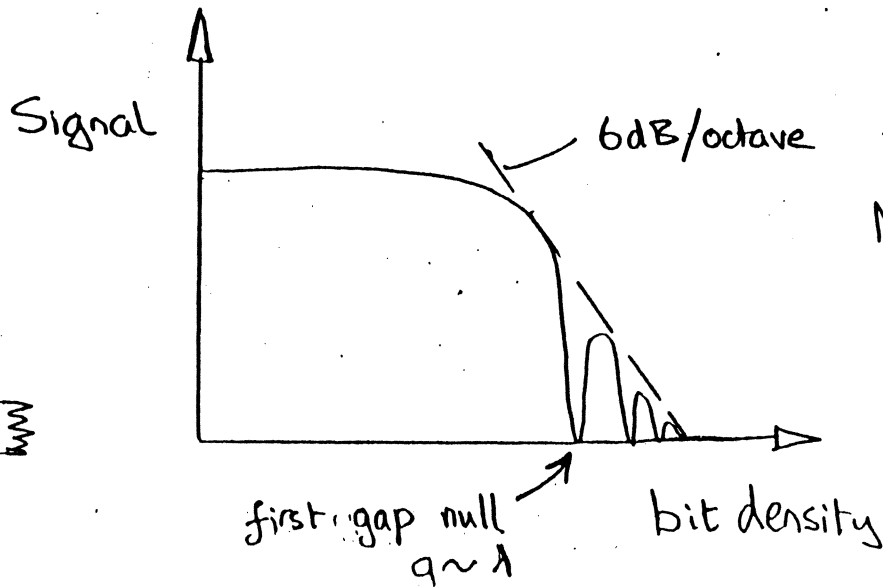
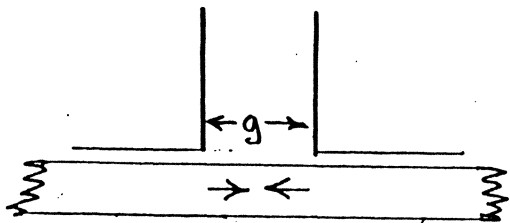
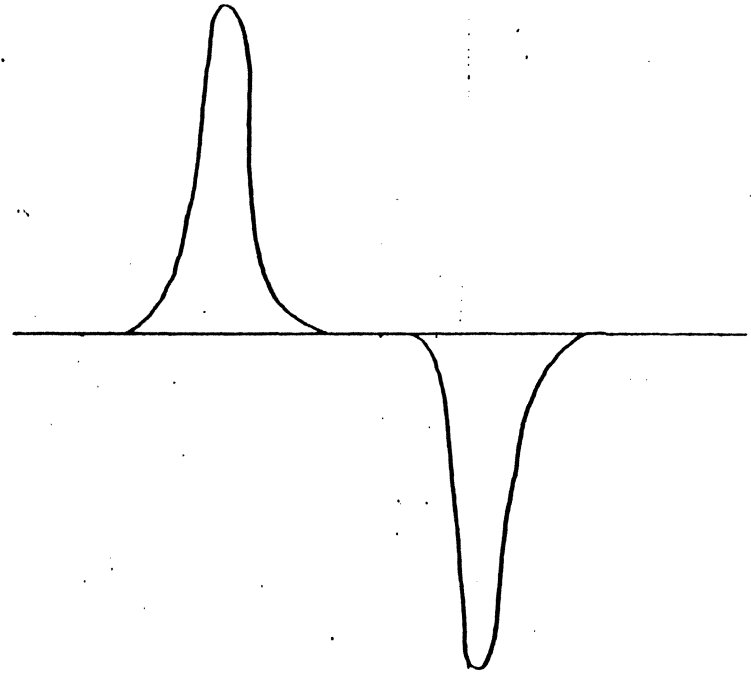
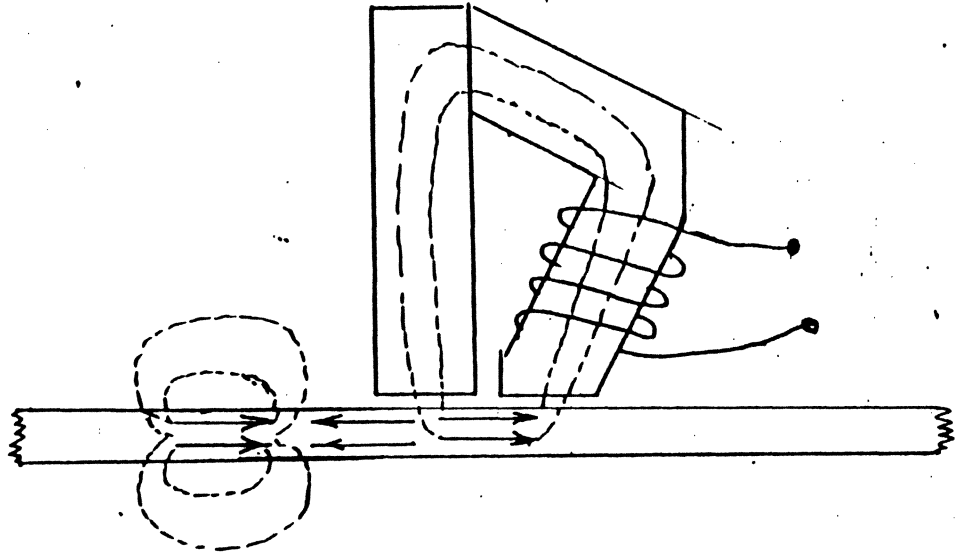
$$= \frac{1.4}{[g^2 + 4(d+a+\delta)(d+a)]^{1/2}}$$

PEAK SHIFT

NRZI

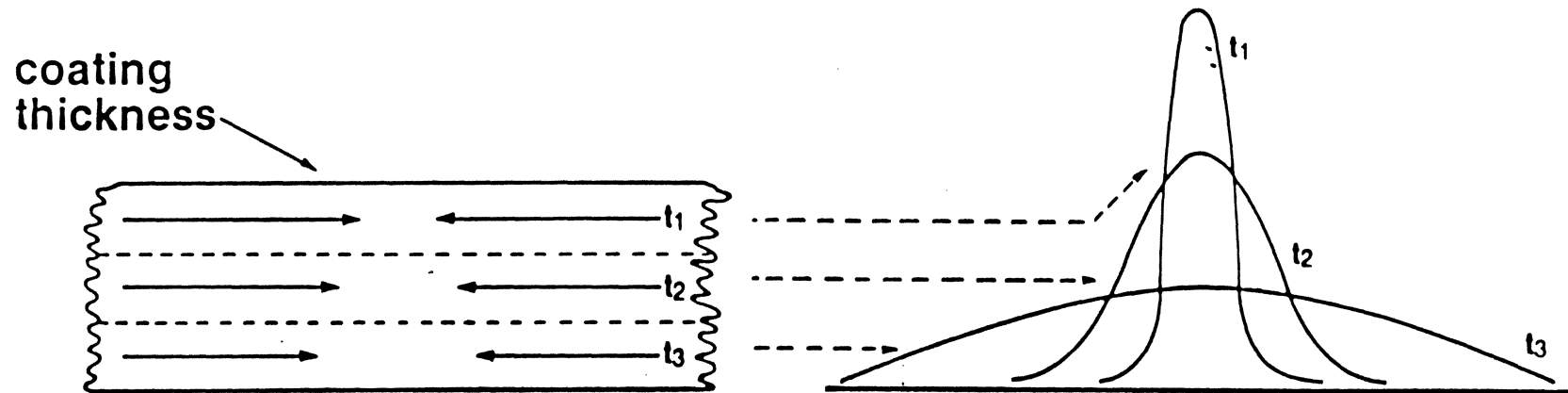


THE READING PROCESS



More exact formula
 $g = 0.88\lambda$

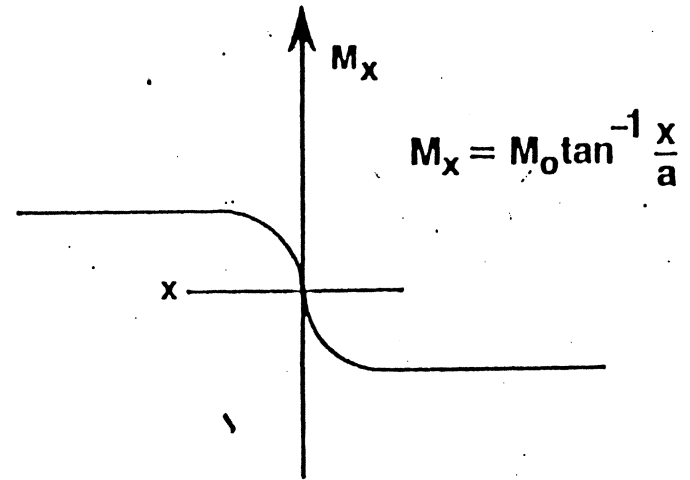
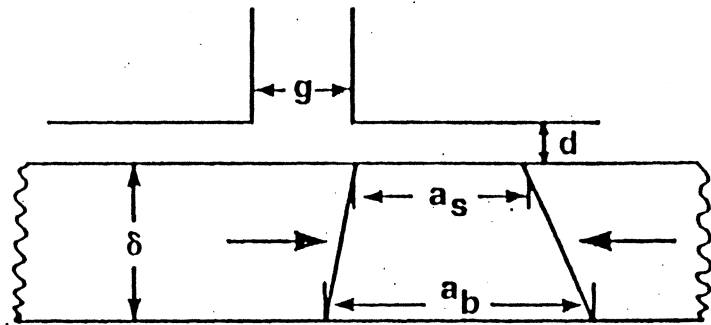
Coating Thickness



(after Hoagland 1963)

G. DATE
VERBATIM

Signal Amplitude



N = NUMBER OF TURNS

W = TRACK WIDTH

v = HEAD-MEDIUM VELOCITY

E = HEAD EFFICIENCY

$K = 2\pi/\lambda$

V = HEAD OUTPUT

$$\underbrace{\frac{V}{NWvE}}_{\text{NORMALIZED OUTPUT}} = \underbrace{\frac{B_r \delta}{\delta + a_b - a_s}}_{\text{MAGNETIC PROPERTIES}} \underbrace{\left(1 - e^{-K(a\delta + a_b - a_s)}\right)}_{\text{THICKNESS}} \underbrace{e^{-K(a_s + d)}}_{\text{SEPARATION}} \times \underbrace{\frac{\sin \pi g/\lambda}{\pi g/\lambda}}_{\text{GAP}}$$

Particulate Media

- wide variety of magnetic properties
- magnetic properties from particles
- mechanical properties from binder
- high coating speeds ($\leq 130\text{m/min}$)
- wide rolls ($\leq 1.3\text{m}$)
- uniform, reproducible properties
- yields high \rightarrow low cost

Thin Film Media

- thin coatings ($< 0.25\mu\text{m}$)
- magnetic properties not diluted by binder
- magnetic properties varied by changing deposition conditions.
- easier to coat on metallic than polymer substrates.

Magnetic Materials;
Unanswered Questions and Unsolved Problems

1. Coercivity in particles -- Fe_2O_3 - Fe_3O_4
 - cobalt iron oxides
 - versus coercivity in coatings
2. High coercivity, high M_r/M_s , isotropy in continuous thin films.
3. Surface effects on particles e.g. Fe_2O_3 - Fe_3O_4 and phosphate chains.
4. Role of particle interactions.
5. Time effects: short term and long term.
6. Orientation and disorientation.
7. Erasure.
8. Relation between magnetic properties and recording performance.

MAGNETISM and MAGNETIC MATERIALS

References:

1. B.D. Cullity, Introduction to Magnetic Materials, Addison Wesley, 1972
2. G. Bate, "Particulate Magnetic Materials", IEEE Proceedings: 74, Nov. 1986, pp 1513-25 (Invited paper)

T.C. Arnoldussen, "Thin-Film Recording Media", IEEE Proceedings:74, Nov. 1986, pp 1526-39 (Invited paper)
3. C.D. Mee, E.D. Daniel, Magnetic Recording, Vol. I: Technology, E. Koster, Chapt. 3, "Recording Media", McGraw-Hill, (1987)
4. S. Chikazumi, Physics of Magnetism (1978)
5. A.H. Morrish, Physical Principles of Magnetism, Wiley, (1965)
6. R.M. Bozorth, Ferromagnetism, Van Nostrand, (1951)
7. G. Bate, "Recording Materials" Chapt. 7, Ferromagnetic Materials, Vol 2, ed. E.P. Wohlfarth, North-Holland, Amsterdam, (1980)
8. Committee on Magnetic Materials, Magnetic Materials, National Academy Press, Washington D.C., (1985)

Chapter 1

MAGNETIC FORCE

The most fundamental aspect of magnetic fields is that forces and torques act on magnets and currents. The phenomenon of force and torque interactions between magnetic poles was investigated by Coulomb and Cavendish at the end of the 18th century. Oersted's discovery of the magnetic effect of electric currents led Ampere and Faraday to invent the first motors in the 1820's. The experimental relations of forces between currents were determined and reported by Ampere. The results of those experiments are the basis of the notion of a magnetic force field and its effect on magnetic poles and on electric currents. Magnetic force field, magnetic pole strength, and magnetic moment are quantities which are defined in this chapter.

1.1 CURRENT FORCES

Two parallel wires have a force of attraction which is proportional to the product of the two electric currents and inversely proportional to the separation of the wires. Figure 1.1 depicts two such currents flowing out of the paper. The force on the right-hand current, exerted over a length L_2 is:

$$F = K \frac{I_1 I_2 L_2}{d} \quad (1.1)$$

where K is a constant depending on the unit system. For currents flowing in opposite directions the force is repulsion, and when currents flow perpendicular to one another they exert no force on each other. Ampere demonstrated these effects and Maxwell formulated them by developing the notions of field theory. To deal with this *action at a distance*, Maxwell conceived of a force field which would act on a current dependent upon the current direction as

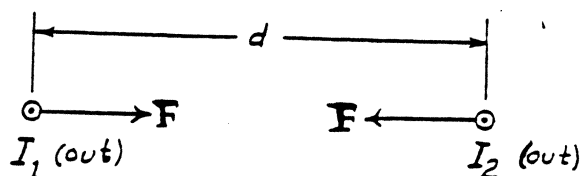


Figure 1.1: current forces

well as the length of a conductor which carries the current.

Taking Ampere's experimental results, concisely stated in equation 1.1, and predicating a force field which also deals with antiparallel currents and perpendicular currents was not a simple task. To do this Maxwell invented vectors. Briefly, to account for Ampere's observations, the field from the first current had to be perpendicular to the second, and the force per unit length on the second current had to be a cross-product of the first current's field and the second current.

\mathbf{B}_F is the symbol which is used here to designate the force field, a vector. The force field arising from current flowing in a long straight wire, as I_1 in figure 1.1, is perpendicular to both the current and to the radius vector from the wire to the point where the field is acting. This field is constant at a particular radius, directed in the right-hand sense around current I_1 : Pointing the right-hand thumb in the direction of the source current (I_1), the curled fingers indicate the direction of the resulting field (counter-clockwise).

In the MKS unit system, the force field acting on current 2 because of current 1 is $\mathbf{B}_{F_{21}}$:

$$\mathbf{B}_{F_{21}} = \frac{\mu_0 I_1}{2\pi r_{12}} (\mathbf{1}_{I_1} \times \mathbf{1}_{r_{12}}) \quad (1.2)$$

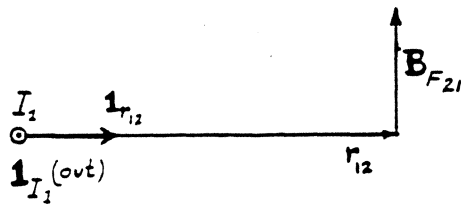


Figure 1.2: magnetic force field

where μ_0 is the permeability of free space, $\mu_0 = 4\pi \times 10^{-7}$, and the boldfaced **1** indicates a unit vector. The unit vector $\mathbf{1}_{I_1}$ is in the direction of the current I_1 , and the unit vector $\mathbf{1}_{r_{12}}$ is the radial direction from current I_1 toward the field point. The direction of the force field is then the vector cross-product of the direction of the source current and the direction from the source current to the field point.

The units of the force field is newtons per amp-meter, so the the constant K of equation 1.1 can be seen to be $\mu_0/2\pi$ in the MKS unit system. generalizing from this example, the element of force on an element of length dl carrying a current I in a magnetic force field \mathbf{B}_F [Lorentz's equation] is:

$$d\mathbf{F} = I dl \times \mathbf{B}_F \quad (1.3)$$

1.2 MAGNETS

Magnets are attracted to iron products such as refrigerator doors. Using a pair of magnets, you can easily demonstrate that they can attract and repel one another. This is explained by magnetic *poles*: a magnet has two poles of opposite *polarity*, and with a pair of magnets the poles that are alike repel one another while unlike poles attract each other.

Directions on the earth's surface are found with a compass, which points toward the geographical north pole, or more accurately toward the *magnetic pole* in the north. The compass is a needle magnet on a pivot. The end of the needle which points to the north is defined as the magnet's *north pole*. This leads us to conclude that the earth's magnetic pole in the north is actually a south pole, as it attracts the compass's north pole.

Attraction and repulsion are forces which act on pairs of magnets: Two north poles repel each other equally, and two south poles repel each other equally. The north pole of one magnet

and the south pole of another attract each other equally. The north and south poles of the same magnet exert forces, but they can only cause deformation of the magnet, not cause motion. For a two magnet system, the result is two forces acting on each pole of each magnet. See figure 1.3 where the pole forces are shown acting equally and oppositely on each pair of poles as single vectors, with summed forces indicated on each pole. The individual forces are inversely proportional to the square of the separation.

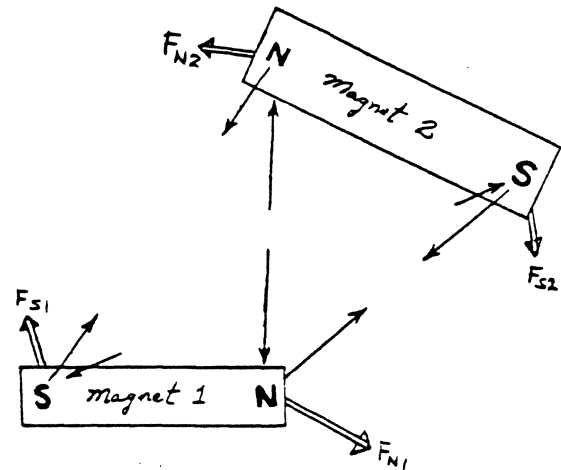


Figure 1.3: magnet forces

Suppose two identical magnets have pole strengths P and lengths l . Placed on a line with north poles adjacent and separated by a distance d as in figure 1.4. Coulomb's experiments showed that the forces between two like magnetic poles is a repulsion proportional to the product of the pole strengths and inversely proportional to the separation between the poles.

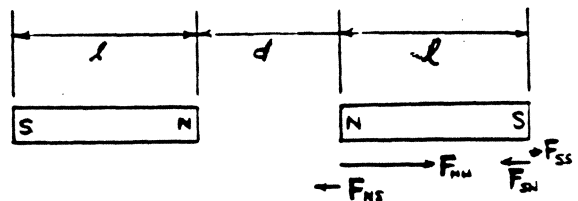


Figure 1.4

Then, using Coulomb's Law, the forces indicated are:

$$\begin{aligned} F_{NN} &= kP^2/d^2 \\ F_{SN} = F_{NS} &= -kP^2/(d+l)^2 \\ F_{SS} &= kP^2/(d+2l)^2 \end{aligned}$$

The total force (to the right) on the right-hand magnet can be calculated from this expression:

$$\begin{aligned} F &= F_{NN} + F_{SN} + F_{NS} + F_{SS} \\ &= k \frac{P^2}{d^2} - 2k \frac{P^2}{(d+l)^2} + k \frac{P^2}{(d+2l)^2} \end{aligned}$$

and the force on the left-hand magnet is equal but to the left. The constant k depends on the unit system and the definition of the pole strength P . The notion of force acting at a distance from the source of the force is called a field, and in particular a force field. Here with two magnets we find such a phenomenon. Each magnet experiences force because of the other magnet. Then each magnet experiences the other as a source of a force field. Each pole can be visualized as having the force acting on it as the product of the force field at that point and its own pole strength P . Because this is a magnetic force, it is of the same nature as the field caused by an electric current, \mathbf{B}_F . From figure 1.4, the north pole of the right hand magnet would experience a force field of to the right of magnitude:

$$B_F = kP/d^2 - kP/(d+l)^2$$

where the first term on the right-hand side is due to the left-hand magnet's north pole, and the second is due to the left-hand magnet's south pole.

Thus the field due to an isolated magnetic pole is directed from that pole toward the point where the field is being calculated. The field is proportional to the pole strength and inversely proportional to the square of the distance from the pole to the point. We define the vector from the source (pole) to the field point as \mathbf{r}_{sf} and its magnitude as r_{sf} . The vector field at the field point is \mathbf{B}_F :

$$\mathbf{B}_F = k \frac{P_s \mathbf{r}_{sf}}{r_{sf}^3} = k \frac{P_s \mathbf{1}_{sf}}{r_{sf}^2} \quad (1.4)$$

and the unit vector is then seen to be:

$$\mathbf{1}_{sf} = \mathbf{r}_{sf}/|r_{sf}|$$

For a group of poles, the field is calculated by vectorially adding the effects of each.

The force field due to a magnet acts on other magnets as well as on electric currents. This requires pole strength be defined to result in

correctly calculated forces on current carrying conductors. Dimensionally the force field \mathbf{B}_F is in newtons per amp-meter in eq. 1.2, and must be in units of newtons per unit pole in equation 1.4. Thus it is necessary that the P be in units of amp-meter. From eq. 1.4 it is necessary that the units of kP be henrys x amps to result in a force field in Teslas or Webers/m². Then k must be henries per meter, and so must contain μ_0 which has the correct units.

1.3 MAGNETS AND CURRENT LOOPS

The field generated by a long straight wire carrying a current is quite different from that of a magnet. In order to obtain the correct constants in the expression for the force field due to a magnetic pole, we must find an equivalency with current. The equivalency is between the *far field* of a magnet and that of a loop of current. The constant k is determined by comparing the fields generated by each. We begin with the field due to a magnet.

1.3.1 FIELD OF A MAGNET

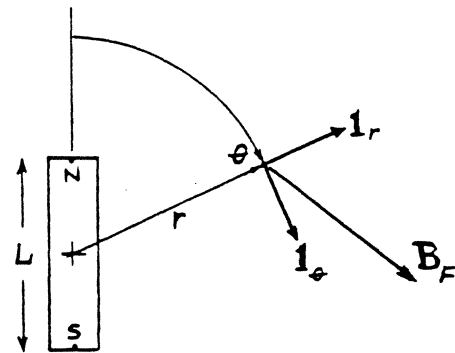


Figure 1.5: magnet field

A bar magnet consists of a *north pole* $+P$ (a positive magnetic charge) and a *south pole* $-P$ (negative magnetic charge) separated by a length L . To determine the magnetic field caused by a bar magnet or "dipole", we use a form of Coulomb's law. The coordinate system most convenient for this situation is the spherical system which uses the radial vector r , measured from the origin - which will be the middle of the bar magnet; an azimuthal angle, ϕ , which is measured from an arbitrary line

through the center of the magnet, but perpendicular to the axis joining the two poles; and a polar angle θ , measured from the half of the polar axis passing through the north pole of the magnet. The coordinates for our analysis is shown in fig. 1.5. the north pole is on the polar axis at $r = L/2, \theta = 0$ and the south pole is at $r = L/2, \theta = 180^\circ$. The application of Coulomb's Law, obtains the following expression for the field at (r, θ, ϕ) :

$$\mathbf{B}_F = \frac{kP\mathbf{1}_r(r - \frac{L}{2} \cos \theta)}{(r^2 + \frac{L^2}{4} - rL \cos \theta)^{3/2}} - \frac{kP\mathbf{1}_r(r + \frac{L}{2} \cos \theta)}{(r^2 + \frac{L^2}{4} + rL \cos \theta)^{3/2}} + \frac{kP\mathbf{1}_\theta(\frac{L}{2} \sin \theta)}{(r^2 + \frac{L^2}{4} - rL \cos \theta)^{3/2}} + \frac{kP\mathbf{1}_\theta(\frac{L}{2} \sin \theta)}{(r^2 + \frac{L^2}{4} + rL \cos \theta)^{3/2}}$$

The far field approximation is that $r \gg L$. Then, making the approximation that

$$(1 + x)^{-n/2} \approx 1 - nx/2$$

we have the far field expression:

$$\mathbf{B}_F = kPL\left\{\frac{2 \cos \theta}{r^3} \mathbf{1}_r + \frac{\sin \theta}{r^3} \mathbf{1}_\theta\right\} \quad (1.5)$$

1.3.2 FIELD OF A CURRENT LOOP

The field of a loop of current and the field of a bar magnet become identical when the distance from the source becomes large with respect to the dimensions of the source, i.e. L for the bar magnet ($r \gg L$) and the radius for the current loop. The analysis of the current loop,

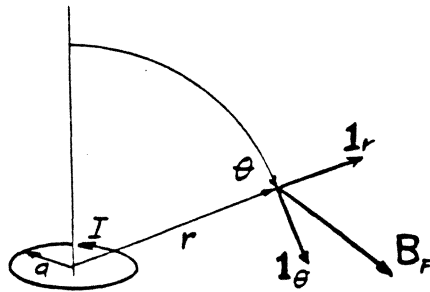


Figure 1.6: current loop field

which is centered at the origin of the spherical coordinate system, with the current on a

circle around the polar axis circulating in the right-hand sense, as indicated in figure 1.6, is most readily solved using the vector magnetic potential formulation of the Biot-Savart Law, and then taking the curl to obtain the force field. This is a common example in many text books on electromagnetic fields and is beyond the scope of this course. The results, when the distance from the source r is much greater than a , the radius of the current loop, the force field is found to be:

$$\mathbf{B}_F = \mu_0 \pi a^2 I \left\{ \frac{2 \cos \theta}{4\pi r^3} \mathbf{1}_r + \frac{\sin \theta}{4\pi r^3} \mathbf{1}_\theta \right\} \quad (1.6)$$

which again is a far-field expression.

1.3.3 MAGNETIC MOMENT

A comparison of equations 1.5 and 1.6 shows that equivalency is between (1) the product of the bar magnet's pole strength and pole separation, PL , and (2) the current loops product of current and area of the loop, $\pi r^2 I$. These terms are both called the *magnetic moment*:

MAGNETIC MOMENTS: $\pi a^2 I$ equivalent to PL

The magnetic moment is a vector, which for a bar-magnet points in the direction from the south pole ($-P$) toward the north pole ($+P$), while for the current loop the direction is found from the right hand rule applied to the current: curl the fingers of the right hand in the direction of the current, and the extended thumb points in the direction of the magnetic moment. This is equivalent to crossing the radius direction with the direction of current. The magnetic moment is symbolized as \mathbf{m} , and is defined as:

$$\text{BAR MAGNET: } \mathbf{m} = PL\mathbf{1}_n \quad (1.7)$$

$$\text{CURRENT LOOP: } \mathbf{m} = IA\mathbf{1}_{r \times I} \quad (1.8)$$

Again comparing equations 1.5 and 1.6, it is seen that the constant k for magnetic pole forces and fields in coulomb's law equations must be :

$$k = \frac{\mu_0}{4\pi} \text{ IN THE MKS SYSTEM}$$

Then the magnetic force field due to an isolated magnetic pole, for the MKS unit system (from equation 1.4) is:

$$\mathbf{B}_F = \mu_0 \frac{P_s \mathbf{1}_{s,f}}{4\pi r_{s,f}^2} \quad (1.9)$$

1.4 TORQUE ON A MAGNETIC MOMENT

Calculation of torque on a magnetic dipole can be done using either a magnet with Coulomb's Law, or a current loop with equation 1.3. In either case, the torque is found to be:

$$\text{TORQUE } \mathbf{T} = \mathbf{m} \times \mathbf{B}_F \quad (1.10)$$

Exercise 1.1: derive equation 1.10 for a bar magnet.

Exercise 1.2: Derive equation 1.2 for a square current loop with an area a^2 and a current I .

1.5 ELECTROMAGNETS

The term *magnet* refers to *permanent magnets* that continue the same behavior over a long period of time. On the other hand, it is possible to magnetize materials such as iron and steel so that they become temporary magnets. An iron rod can be magnetized by applying an electric current to a coil of wire wound around it. It will be strongly magnetic as long as the current continues, but will lose most of its magnetism when the current is stopped. Such magnets are called *electromagnets* and their strength is proportional to the current. After the current has stopped they may have a small remnant magnetism. These effects can be measured through forces acting on an electromagnet in the vicinity of another magnet.

1.6 SUMMARY

The definitions given here for magnetic pole strength and for magnetic moment have been made so the force on a pole and the torque on the magnetic moment are both calculated from the force field \mathbf{B}_F . The units for magnetic pole strength are amp-meters and the units for magnetic moment are amp-meters². The only unit system employed has been the MKS unit system. The units of the force field are directly determined as from the force equations

The magnetic force field field, \mathbf{B}_F , has been defined for the following cases:

1. Where it arises from a current in a long

straight wire (Ampere's Law, from eq. 1.2):

$$\mathbf{B}_{F_{21}} = \frac{\mu_0 I_1}{2\pi r_{12}} \mathbf{1}_{I_1 \times r_{12}}$$

2. from an isolated magnetic pole (Coulomb's Law, from eq. 1.7):

$$\mathbf{B}_F = \mu_0 \frac{P_p \mathbf{1}_{p,f}}{4\pi r_{p,f}^2}$$

3. Where it is the far-field arising from a magnetic dipole (from eq. 1.5 and 1.6):

$$\mathbf{B}_F = \frac{\mu_0 \mathbf{m}}{4\pi r^3} \{2\mathbf{1}_r \cos \theta + \mathbf{1}_\theta \sin \theta\} \quad (1.11)$$

The force on an element of electric current is part of the Lorentz force equation as given in eq. 1.3:

$$d\mathbf{F} = I d\mathbf{l} \times \mathbf{B}_F$$

The torque on a magnetic dipole moment \mathbf{m} is given in eq. 1.10:

$$\mathbf{T} = \mathbf{m} \times \mathbf{B}_F$$

Chapter 2

MAGNETIC INDUCTION

2.1 EXPERIMENTAL BASIS

The statement that like poles of magnets repel each other and unlike poles attract can be proven from the fact that it is possible to find three poles which mutually repel each other, but it is not possible to find three poles which mutually attract each other.

From the earliest times magnets were known for their ability to attract (but not repel) certain metals, most notably iron. This shows that a magnetic pole induces an unlike pole in the closest part of nearby iron. As magnetic poles occur only in pairs (i.e. magnetic monopoles have not been found), then a magnetic pole will also induce a like pole in the most remote part of the nearby iron. As these induced poles are equal in strength, the result will be an attraction of the iron to the closer pole of the magnet.

Oersted's discovery of the magnetic effect of electric currents eventually led Faraday to investigate whether magnetic fields could cause or induce electric current in a conductor. He found this could be done either by changing the current that created the field or by moving a conductor relative to a magnet or other field source. Henry had independently discovered the effect of the changing current. Faraday carefully described his experiments, including all the failures. This record keeping proved invaluable to Maxwell when he set out to understand electromagnetism.

It was in 1831 that Faraday discovered that electric current was induced in a wire which formed a closed circuit when the current in a nearby wire was changed. Faraday saw the induced current was the response to an *electromotive force* (abbreviated *emf*). Ohm's discovery of the law of resistance in 1827 provided the key

to understanding the relation between current and *emf*.

Faraday attributed the induced *emf* to a flux arising from the magnetic field of the current in the nearby wire. He envisioned the induced *emf* being caused by changing this flux. The induced current was proportional to the *emf*, by Ohm's Law.

The *emf* is equal to the negative time rate of change of the magnetic "flux":

$$emf = -\frac{d\phi}{dt} \quad (2.1)$$

The symbol ϕ (*phi*) indicates the quantity magnetic flux. The *emf* has units of electric potential, volts in the MKS unit system. Thus the units of the magnetic flux ϕ is volt-seconds, which is also called webers.

Equation 2.1 is the most fundamental form of Faraday's Law of electromagnetic induction and serves as the definition for magnetic flux.

2.2 MAGNETIC FLUX DENSITY

Magnetic flux is a *flow* concept which involves no tangible material. However it is useful to think of magnetic flux as analogous to the flow of water, which has a particular direction at a particular point. At such a point the direction and speed of the water, combined with the density of the water, can be used to define a flow density of the water. This flow density is a vector, and could also be called the flux density of the water at that point.

The flow of a stream of water is usually defined as the rate at which water passes through an imaginary plane or other surface. This is also the average flow density or flux density

multiplied by the area of the imaginary surface, provided that the surface is perpendicular to the flow density at every point. This can be expressed as a surface integral, which involves the scalar product (dot product) of two vectors: the flux density and the surface element's normal.

Magnetic flux density is similarly defined in terms of magnetic flux. Using the symbol \mathbf{B} for the vector magnetic flux density, the flux through a surface S is

$$\phi = \iint_S \mathbf{B} \cdot \mathbf{1}_n dS \quad (2.2)$$

where $\mathbf{1}_n$ is the unit vector normal to the element of surface area dS .

The MKS units of magnetic flux density are seen to be volt-seconds per square meter, or webers per square meter which are also called teslas.

Magnetic flux density \mathbf{B} is a vector quantity which is defined at a so-called macroscopic point. A macroscopic point is not the same as a mathematical point, but should be thought of as a very small volume or surface area, yet large enough to accommodate a large number of atoms.

Faraday's law of electromagnetic induction can then be written in terms of the flux density as:

$$emf = -\frac{\partial}{\partial t} \iint_S \mathbf{B} \cdot \mathbf{1}_n dS \quad (2.3)$$

This can further be expressed using the circulation of the electric field intensity \mathbf{E} to replace the emf , i.e. the line integral of the electric field intensity, \mathbf{E} , around the boundary of the surface S in the counter-clockwise direction is the emf .

$$emf = \oint_{C_S} \mathbf{E} \cdot d\mathbf{l} = -\frac{\partial}{\partial t} \iint_S \mathbf{B} \cdot \mathbf{1}_n dS \quad (2.4)$$

where the curve C_S is the boundary of the surface S .

2.3 FLUX INTO A VOLUME

In equation 2.4 the circulation of \mathbf{E} is the emf around any closed path, and the surface over which \mathbf{B} is integrated is any surface which is bounded by the circulation path. This leads

to an important law, which was formulated by Gauss together with a more general theorem of vector calculus.

Consider a square box as shown in fig. 2.1. We take the boundary of side A as the path for the circulation of \mathbf{E} , which encloses the surface A . The circulation is $-d\phi/dt$ passing through surface A . Notice also that the path of circulation also is the boundary of a surface composed of the box excluding side A . The circulation of \mathbf{E} is also $-d\phi/dt$ through that complex surface. The conclusion to be drawn here is that the time rate of change of the flux into a volume is identical to time rate of change of the flux out of the volume: there is no net time rate of change of flux into (or out of) a volume. Thus, in the absence of magnetic monopoles, the flux into a volume is equal to the flux out.

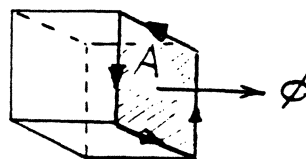


figure 2.1

It is useful to shrink the volume until it approaches the macroscopic point in size. In doing this, it is a straight-forward process to show that the divergence of the flux density is zero.

$$\text{div } \mathbf{B} = 0 \quad (2.5)$$

exercise 2.1: Using cartesian coordinates, and taking limits as the dimensions of the box in figure 2.1 shrink to zero, using the definition of the partial differential and the fact that the net flux into the box is zero, demonstrate that the divergence of \mathbf{B} is zero.

Equation 2.5 tells us that magnetic flux is continuous, having no sources or sinks (in the absence of monopoles). This requires that magnetic flux "closes" on itself, and for that reason is called solenoidal.

2.4 MOTIONAL EMF

A conductor moving in a magnetic field is a relativistic situation, and is equally viewed as a magnetic field moving with respect to the conductor. The charges within the conductor have a velocity relative to the magnetic field, which results in an effective electric field acting on the charges. If the charges are free to move, as in certain closed circuits, they will participate in a current flow. The force per unit charge is, by definition, the electric field intensity. The entire situation is summed up in the Lorentz force law:

$$\mathbf{F} = Q(\mathbf{E} + \mathbf{v} \times \mathbf{B}_F) \quad (2.6)$$

here Q is the charge, \mathbf{E} is the electric field intensity, \mathbf{v} is the relative velocity of the charge with respect to the magnetic force field \mathbf{B}_F .

In a situation where there is a segment of wire moving in a magnetic field, and there is no provision for current flow, then an equilibrium condition will arise according to equation 2.6, where the net force will be zero, so that the electrostatic field arising from free charge distribution within the wire and the electromagnetic field due to the relative velocity of the charges in the magnetic field are balanced. In that case, the electric field must be:

$$\mathbf{E} = -\mathbf{v} \times \mathbf{B}_F \quad (2.7)$$

The emf generated over a length of conductor is found by the line integral of the electric field over the length:

$$emf = - \int (\mathbf{v} \times \mathbf{B}_F) \cdot d\mathbf{l} \quad (2.8)$$

Equation 2.8 is fundamental to electrical power generation, and provides the relation between *emf* and motion of a conductor in a magnetic force field. In principle such a force field can be calculated from known current distributions and the resulting *emf* generated on a conductor moving in the force field also calculated. Such a calculation will agree with an experimental measurement, showing that the force field is directly involved with the motional generation of *emf*.

A loop of conductor can be placed in this calculated force field and the flux of the force field passing through the loop determined by integration. This flux will be a linear function

of the current. Thus the flux of the force field can be varied by varying the current. In doing so, the time rate of change of the flux due to the force field is found experimentally to be equal to the measured *emf* on such a loop. So, for the case of a gaseous media the magnetic force field \mathbf{B}_F and the magnetic flux density \mathbf{B} are found to be equal.

The magnetic force field \mathbf{B}_F acts on electric currents and on magnetic poles, but the results are measurable only in a non-solid environment. On the other hand, magnetic flux in a solid material is measurable as the *emf* in a conductor wrapped around the solid can be measured. As seen in equation 2.5, the divergence of the magnetic flux density is zero, which means that the normal component of magnetic flux is continuous in crossing from one medium to another.

Except for ferrofluids, a non-solid environment is also non-magnetic. In non-magnetic environments experiments show that the magnetic force field and the magnetic flux density are identical. Inside solid magnetic environments measurements of forces and torques are impossible. However, one experiment in a ferrofluid indicates that the magnetic force field and the magnetic flux density are different in a magnetic material (environment)¹. This experiment appears to support Maxwell's understanding of electromagnetic forces and *emf*. Maxwell seems to have espoused the notion that the magnetic force field depends only on currents and magnetic poles. Some later workers have found this an unacceptable idea, and have insisted that the medium or environment plays a key role. These dissident workers have succeeded in establishing their opinion in the most commonly used version of the MKS unit system, which is known as the Sommerfeld System. This will be discussed in some detail at a later point.

In order to follow Maxwell's understanding, we first establish the concept of the magnetic field intensity, \mathbf{H} , which is exactly proportional to the force field \mathbf{B}_F .

¹Experimental demonstration that the couple on a bar magnet depends on H , not B , R.W. Whitworth and H.V. Stopes-Roe, *Nature*, v 234, pp31-33, Nov 1971.

2.5 MAGNETIC FIELD INTENSITY

The magnetic field intensity \mathbf{H} is a vector, and, following Maxwell's ideas, is calculated from the combination of current distribution and magnetic pole distribution. It is proportional to the vector we have been calling the force field, with the definition:

$$\mathbf{B}_F = \mu_0 \mathbf{H} \quad (2.9)$$

The general expression for the calculation of \mathbf{H} is obtained by modifying the Ampere Law and Coulomb Law expressions of chapter 1 (equations 1.2 and 1.7). The current and its direction are combined in the form of current density enclosed in a volume element as indicated in figure 2.2, which then requires an integration over all current density. This results in an expression which is usually called the law of *Biot and Savart*:

$$\mathbf{H}_m = \iiint_{Vol} \frac{\mathbf{J} \times \mathbf{1}_{sf}}{4\pi r_{sf}^2} dVol \quad (2.10)$$

where \mathbf{J} is the current density (source point), r_{sf} is the distance from the source point to the field point (the field point is that where the field is being calculated), and $\mathbf{1}_{sf}$ is the unit vector from the source point toward the field point. Because magnetic materials are typically magnetized using electric currents, this part of the magnetic field \mathbf{H}_m will be called the *magnetizing field*.

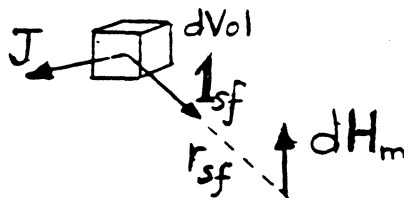


fig. 2.2 Biot-Savart field source element

Coulomb's Law is modified to deal with pole density rather than isolated poles, which also results in a contribution to the field intensity requiring integration over all pole density:

$$\mathbf{H}_d = \iiint_{Vol} \frac{\rho_m \mathbf{1}_{sf}}{4\pi r_{sf}^2} dVol \quad (2.11)$$

where ρ_m is the magnetic pole density (source point), $\mathbf{1}_{sf}$ is again the unit vector from the source point to the field point, and r_{sf} is the distance from the source point to the field point. Because inside a magnetic material this part of the field is in opposition to the magnetization, \mathbf{H}_d will be called the *demagnetizing field*.

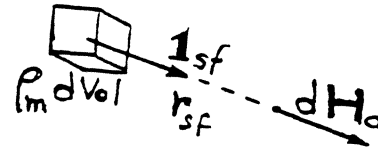


fig. 2.3 pole density field source element

Then the complete expression for magnetic field intensity is the sum of the *magnetizing field* and the *demagnetizing field* $\mathbf{H}_m + \mathbf{H}_d$:

$$\mathbf{H} = \iiint_{Vol} \frac{\mathbf{J} \times \mathbf{1}_{sf}}{4\pi r_{sf}^2} dVol + \iiint_{Vol} \frac{\rho_m \mathbf{1}_{sf}}{4\pi r_{sf}^2} dVol \quad (2.12)$$

2.6 SUMMARY

Faraday's law of electromagnetic induction is:

$$emf = -\frac{d\phi}{dt}$$

In the MKS unit system, magnetic flux is in webers while magnetic flux density is in teslas.

The divergence of the magnetic flux density is zero: $\text{div } \mathbf{B} = 0$

It has been shown that for the case of a gaseous media the magnetic force field \mathbf{B}_F and the magnetic flux density \mathbf{B} are equal.

As magnetic forces have been shown to be independent of the medium (as was envisioned by Maxwell), the magnetic field intensity is defined as proportional to the quantity we have called the force field, as in equation 2.9: $\mathbf{B}_F = \mu_0 \mathbf{H}$.

Chapter 3

MAGNETIZATION AND POLE DENSITY

3.1 INTRODUCTION

The magnetic moment of a dipole, arising either from a bar magnet or a current loop has been defined in chapter 1, and found to have MKS units of amp-meters². All atoms have magnetic moments which are attributed to the orbiting electrons and to the spinning of those orbiting electrons. In most materials the average magnetic moment tends to align with an applied magnetic field. This behavior is called *paramagnetism*. In some materials, the magnetic moments tend to align in opposition to an applied magnetic field. This behavior is called *diamagnetism*. Both paramagnetism and diamagnetism are weak effects.

Some materials have strong *magnetic* effects. To some degree, all such materials exhibit *spontaneous* magnetization, the property observed in permanent magnets, in that they behave magnetically in the absence of an applied magnetic field. These materials are broadly classified as magnetic. Other terms such as *ferromagnetic*, *ferrimagnetic* and *anti-ferromagnetic*, are used describe specific materials, but we will use two different classifications of magnetic materials, one of which is *soft*, and the other as *hard*.

The atomic magnetic moments in *soft magnetic* materials tend to align with an applied magnetic field rather easily. Also, *soft magnetic* materials have little remnant magnetization when the field is removed. Such materials are used in magnetic recording heads.

Hard magnetic materials are more difficult to magnetize, but retain the alignment of magnetic moments after the magnetizing field is removed. Such materials are used in magnetic recording materials.

3.2 FARADAY VOLTAGE AND FLUX MEASUREMENT

The most common method of measuring magnetic flux is via the Faraday Law relation of equation 2.1:

$$emf = -\frac{d\phi}{dt}$$

which involves the measurement of the voltage generated in a loop of wire. The ends of the wire are not quite touching, so that the voltage is the instantaneous value of the time rate of change of the flux passing through the loop. To obtain the flux, that voltage is electronically integrated as indicated in figure 3.1.

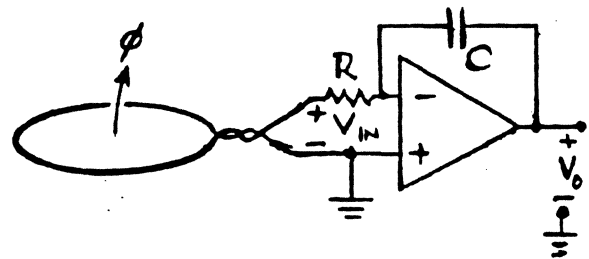


figure 3.1

The magnetic flux density is a quantity which is closely related to the magnetization of a magnetic material, and the flux density is approximated by dividing the flux by the area of the loop (which actually gives the average flux density):

$$B_{avg} = -\frac{1}{A} \int emf dt \quad (3.1)$$

In some cases, in order to increase the voltage available, several turns of wire are formed into

a coil of N turns. The resulting voltage is then seen to be:

$$v_{coil} = -emf = N \frac{d\phi}{dt} \quad (3.2)$$

Then the average flux density can be calculated in terms of the voltage from the N -turn coil:

$$B_{avg} = \frac{1}{A} \int \frac{v_{coil}}{N} dt \quad (3.3)$$

3.3 MAGNETIZATION

Magnetization is defined as the average magnetic moment per unit volume at a macroscopic point. Recall that a macroscopic point is a vector point, one which can contain a large number of atoms. The *magnetization* is a vector quantity, which is symbolized with \mathbf{M} in the MKS unit system:

$$\mathbf{M} = \lim_{Vol \rightarrow 0} \frac{\sum \mathbf{m}_i}{Vol} \text{ [amps/m]} \quad (3.4)$$

where the units of the individual magnetic moments (\mathbf{m}_i) are amp-meters².

The units of magnetization, \mathbf{M} , are here defined to be the same as magnetic field intensity \mathbf{H} , namely amps per meter. Within a magnetic material, as can be demonstrated experimentally using Faraday's law, the vector sum of the magnetization and the magnetic field intensity determine the flux density:

$$\mathbf{B} = \mu_o(\mathbf{M} + \mathbf{H})$$

It is recalled that \mathbf{H} consists of two parts, the magnetizing field due to currents, and the demagnetizing field due to magnetic poles: $\mathbf{H} = \mathbf{H}_m + \mathbf{H}_d$. Then we can write:

$$\mathbf{B} = \mu_o(\mathbf{M} + \mathbf{H}_m + \mathbf{H}_d) \text{ [teslas]} \quad (3.5)$$

The demagnetizing field arises from magnetic moments within the magnetic material. This can be understood from the properties of the magnetic flux density, \mathbf{B} , which must be continuous on a macroscopic field basis.

3.4 MAGNETIC POLE DENSITY

A fundamental decision that Maxwell made in the formulation of electromagnetism was in the

definition of magnetic flux density. This is expressed as in eq. 2.5:

$$\text{div } \mathbf{B} = 0$$

Applying this to equation 3.5, we obtain that:

$$\text{div}(\mu_o(\mathbf{M} + \mathbf{H}_m + \mathbf{H}_d)) = 0$$

where μ_o is a constant. In the quasi-static situation (i.e. no macroscopic eddy currents) the divergence of the magnetizing field (which has its sources outside of the magnetic material), is zero in the magnetic material., thus:

$$\text{div } \mathbf{M} + \text{div } \mathbf{H}_d = 0 \quad (3.6)$$

We can use this result to calculate the demagnetizing field arising from a known distribution of magnetization. To develop this ability, we first integrate equation 3.6 over a volume, separating the two parts:

$$\iiint_{Vol} \text{div } \mathbf{H}_d dVol = - \iiint_{Vol} \text{div } \mathbf{M} dVol$$

The divergence theorem can be applied to the left side of this equation to obtain:

$$\iiint_{Vol} \text{div } \mathbf{H}_d dVol = \iint_{S_v} \mathbf{H}_d \cdot \mathbf{1}_n dS_v$$

where S_v is the surface of the volume and $\mathbf{1}_n$ is the outward pointing normal to the surface. Then, substituting this result into the previous equation:

$$\iint_{S_v} \mathbf{H}_d \cdot \mathbf{1}_n dS_v = - \iiint_{Vol} \text{div } \mathbf{M} dVol \quad (3.7)$$

Which is a form of Gauss's law. This shows that the divergence of the magnetization plays the part of a pole density:

$$\rho_m = -\text{div } \mathbf{M} \quad (3.8)$$

where ρ_m is the magnetic pole density. Then equation 3.7 can be written:

$$\iint_{S_v} \mathbf{H}_d \cdot \mathbf{1}_n dS_v = \iiint_{Vol} \rho_m dVol \quad (3.9)$$

We can further develop this by taking the case where there is simply a small volume,

ΔVol , and we take the surface for integration as a sphere centered on that small volume. Then on the surface the flux due to the small volume is perpendicular to the surface, and of the same value at any point on the surface. The surface integral then becomes simply $4\pi r^2 H_d$, with a vector direction from the small volume to the surface point. Then equation 3.9 can be re-written to take in all pole density as:

$$H_d = \iiint_{Vol} \frac{\rho_m \mathbf{1}_{sf}}{4\pi r_{sf}^2} dVol \quad (3.10)$$

which matches the demagnetizing field portion of equation 2.12.

3.5 FIELD MEASUREMENT

A fundamental relation for magnetic materials is the B-H or M-H characteristic which is typically a hysteresis loop. Measurements for a B-H loop typically use the integration scheme describe with figure 3.1 to determine the flux. The measurement of the magnetic field in a magnetic material is impractical, because force measurements are not generally possible. The strategy employed is to calculate the magnetizing field from the current configuration, and either calculate the demagnetizing field or else use a sample configuration where the demagnetizing field is negligible, the latter practice being preferred. Here we consider three current configurations to give known field values. In each of these systems, a Faraday voltage pickup coil will have mutual inductive linkage with the magnetizing field source, so that such a voltage will be induced by a varying magnetizing field even in the absence of a magnetic material sample. This effect requires cancellation by some instrumentation means, or correction by computation.

3.5.1 THE HELMHOLTZ COIL PAIR

A Helmholtz coil pair is a system of two identical coils, each having N turns and carrying the same current I . The coils are on a common axis separated by the radius, r . This results in an axial field which is quite uniform on the axis at the midpoint between the two coils. This is illustrated in figure 3.2.

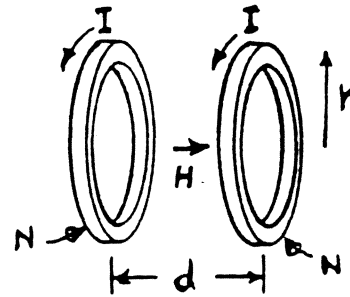


figure 3.2

This system is used for small samples of material which require relatively small magnetizing fields. The field at the center of the system is:

$$H_{axis} = \frac{r^2 NI}{[r^2 + d^2/4]^{3/2}} = 0.7155 \frac{NI}{r} \text{ [a/m]} \quad (3.11)$$

The region of uniformity is a volume of about 10% of the volume of the system.

3.5.2 THE LONG SOLENOID

A cylindrically wound coil of wire is called a solenoid. For a long, uniformly wound solenoid of N turns and length L meters, carrying a current I amps, the field at the center of the solenoid is:

$$H_{solenoid} = \frac{NI}{L} \text{ [a/m]} \quad (3.12)$$

The solenoid field, excluding the ends, is quite uniform. The end regions excluded are lengths of about a diameter. Such systems can be used with relatively large samples of magnetic material which do not require very large magnetizing fields.

3.5.3 THE ELECTROMAGNET

For materials requiring very large magnetizing fields, electromagnets are commonly used. The electromagnet is a structure of (magnetically) soft iron, constructed so that there is a gap where a uniform field is produced. The gap has two cylindrical poles with flat faces separated by a distance of g meters. The structure is energized by current I flowing in coils of N total turns. An electromagnet is sketched in figure

3.3. At distances of at least the gap-length (g) from the pole edges, the field is quite uniform:

$$H_{gap} = \frac{NI}{g} \text{ [a/m]} \quad (3.13)$$

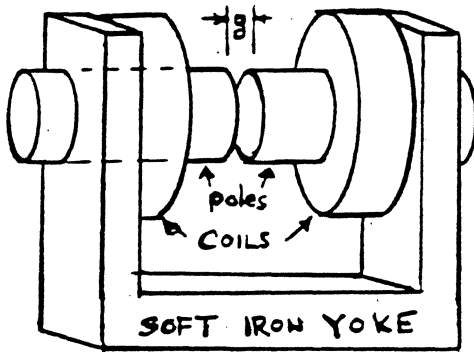


figure 3.3

3.6 MKS UNITS

The treatment of magnetic moments is different in two competing MKS unit systems. These two systems are further described below. The first has magnetization (M) in the same units as H , as has been done here. The second uses a variable called Intensity of magnetization (I) and its units are the same as B . The defining equations are:

$$\text{Sommerfeld: } \mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}) \quad (3.14)$$

$$\text{Kennelly: } \mathbf{B} = \mu_0\mathbf{H} + \mathbf{I} \quad (3.15)$$

In *linear* magnetic materials, also called soft magnetic materials, the magnetization is proportional to H for fields less than some particular level:

$$\mathbf{M} = \chi\mathbf{H}$$

for example, and χ (greek chi) is called the magnetic susceptibility. For such materials, the permeability is μ :

$$\mu = \mu_0\mu_r = \mu_0(1 + \chi)$$

and μ_r is the relative permeability.

The two competing MKS unit systems also differ in that the magnetic moment for a current loop of current i and area A is defined differently for each:

$$\text{Sommerfeld: } \mathbf{m} = iA\mathbf{1}_n$$

$$\text{Kennelly: } \mathbf{j} = \mu_0 i A \mathbf{1}_n$$

The definition of torque on a magnetic moment is therefore different in the two systems. Sommerfeld has the torque proportional to the magnetic flux density B , whereas we have defined it as proportional to the force field $\mu_0 H$ which is in accordance with the Kennelly system.

3.7 UNIT SYSTEMS

A variety of unit systems have been developed, and several survive in use today. The *System International, S.I.*, which is an MKS system, was devised by G. Giorgi in 1902. Electrical Engineering, has long been defined in this unit system.

Two competing MKS systems are in use for magnetic units. They both use the same definitions and units for magnetic field intensity, flux density and permeability: Magnetic field intensity: H [a/m], and Magnetic flux density: B [tesla]. And $\mu_0 = 4\pi \times 10^{-7}$ [henry/m]

The system most commonly used in text books on electromagnetic fields was invented by A. Sommerfeld, and was adopted by one of the most successful texts in the middle 20th century, *Electromagnetic Theory* by J.A. Stratton. Later authors tended to follow this text, so that the Sommerfeld system has become a de facto standard. The units used in the *Sommerfeld System* are:

$$\begin{aligned} \text{Magnetic moment: } \mathbf{m} & \text{ [ampere-m}^2\text{]} \\ \text{Magnetization intensity: } \mathbf{M} & \text{ [ampere/m]} \end{aligned}$$

The defining equations are:

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$$

$$\mathbf{T} = \mathbf{m} \times \mathbf{B}$$

The competing MKS system is the Giorgi-Kennelly system, generally following Giorgi's original MKS system, but was so strongly supported by A.E. Kennelly that it is known as the Kennelly System. The units used in the *Kennelly System* are:

$$\begin{aligned} \text{Magnetic moment: } \mathbf{j} & \text{ [weber-m]} \\ \text{Magnetization intensity: } \mathbf{I} & \text{ [weber/m}^2\text{]} \end{aligned}$$

The defining equations are:

$$\mathbf{B} = \mu_0 \mathbf{H} + \mathbf{I}$$

$$\mathbf{T} = \mathbf{j} \times \mathbf{H}$$

Unfortunately there is an inconsistency in the two systems. The symbols \mathbf{j} for magnetic moment and \mathbf{I} for magnetic intensity are employed in the Kennelly system and are used here to pursue the inconsistency as follows:

As \mathbf{B} and \mathbf{H} are the same in both units, the first defining equations give that:

$$\mu_0(\mathbf{H} + \mathbf{M}) = \mu_0 \mathbf{H} + \mathbf{I} \quad (3.16)$$

so that $\mathbf{I} = \mu_0 \mathbf{M}$

The magnetic moment of a volume is defined in both systems as the volume integral of the magnetization. Thus, with a particular volume, assuming uniform magnetization within the volume, the magnetic moment in each system (the same volume, only the unit systems are different):

$$\text{For Sommerfeld: } \mathbf{m} = \mathbf{M} \text{ vol}$$

$$\text{For Kennelly: } \mathbf{j} = \mathbf{I} \text{ vol}$$

In the presence of an applied field, the torque must be independent of the unit system:

$$\text{Sommerfeld: } \mathbf{T} = \mathbf{m} \times \mathbf{B} = (\mathbf{M} \times \mathbf{B}) \text{ vol}$$

$$\text{Kennelly: } \mathbf{T} = \mathbf{j} \times \mathbf{H} = (\mathbf{I} \times \mathbf{H}) \text{ vol}$$

Then, as $\mathbf{B} = \mu_0 \mu_r \mathbf{H}$ in both systems, we have that $\mathbf{I} = \mu_0 \mu_r \mathbf{M}$

This differs from the result from equation 3.16: $\mathbf{I} = \mu_0 \mathbf{M}$. These differ by the relative permeability of the media, and cannot both be correct. The problem is to determine where the error lies.

In the paper, *Experimental demonstration that the Couple on a Bar Magnet depends on H, not B*, in Nature, vol 234, November 1971, pp 31-33, R.W. Whitworth and H.V. Stopes-Roe of the University of Birmingham in England, reported the results of an experiment, which apparently proved the error was in the Sommerfeld system. This error is in the expression for the torque on the magnetic moment, so that the magnetization intensity relation is $\mathbf{I} = \mu_0 \mathbf{M}$, and the torque must be corrected in the Sommerfeld system:

$$\text{Sommerfeld correction: } \mathbf{T} = \mathbf{m} \times \mu_0 \mathbf{H}$$

The other surviving metric unit system is a cgs system called *electromagnetic units* (emu), which was one of the incomplete systems introduced by J.C. Maxwell in 1863. Yet another system which he introduced in 1863 was also an incomplete cgs system named the *electrostatic unit system* (esu) which is no longer used. The primary use of the emu system is in magnetics and magnetic materials, primarily because the International Electrotechnical Commission (IEC) has not yet set S.I. standards magnetic moment and magnetization. The emu system has the following magnetic units which are in common use:

Magnetic field intensity: \mathbf{H} [oersteds, oe]

Magnetic flux density: \mathbf{B} [gauss, g]

Magnetic moment: \mathbf{m} [ergs/oersted, or emu]

Magnetization intensity: \mathbf{M} [emu/cm³]

Specific magnetization: σ [emu/gram]

The defining equations are:

$$\mathbf{B} = \mathbf{H} + 4\pi \mathbf{M}$$

$$\mathbf{T} = \mathbf{m} \times \mathbf{H}$$

Thus it can be seen that the permeability of free space is unity in the emu system: $\mu_0 = 1$.

3.8 MAGNETIC UNIT CONVERSIONS

Magnetic Field Intensity: \mathbf{H}
1 oersted = $250/\pi$ amperes/m

Magnetic Flux Density: \mathbf{B}
10,000 Gauss = 1 tesla

Magnetic Intensity: \mathbf{I} (Kennelly)
 $4\pi \times 10^4$ emu/cm³ = 1 tesla

Magnetization: \mathbf{M} (Sommerfeld)
 4π emu/cm³ = $(250/\pi)$ a/m

Chapter 4

LINEAR MAGNETIC MATERIALS

4.1 LINEAR CHARACTERISTICS

Magnetic materials exhibit a hysteresis loop in a form similar to that shown in figure 4.1. The usual parameters are the coercive force H_c , the saturation flux density B_s , and the remnant magnetization B_r . A fourth parameter, the saturation field H_s , is added for our convenience here. A so-called *soft magnetic material* has B_r that is a small fraction of B_s . When its operation is limited to the unsaturated region ($|\mathbf{H}| < H_s$) the characteristic can be modeled mathematically as a linear device with reasonable accuracy. In this operating range, ignoring the hysteresis for the moment, the relation between the magnetic field intensity, \mathbf{H} , and the magnetic flux density, \mathbf{B} , is reasonably approximated by the linear relation: $\mathbf{B} = \mu\mathbf{H}$. (The hysteresis results in a power loss, which will be dealt with at a later point.) Referring to figure 4.1, the permeability μ is the slope of the unsaturated region and is approximated by B_r/H_c .

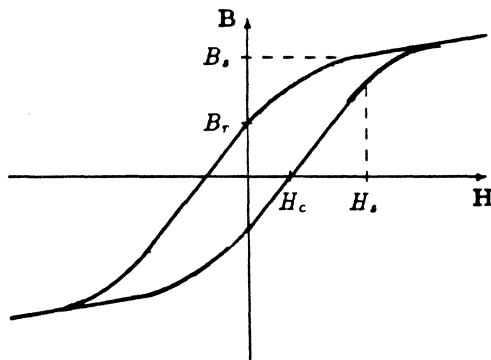


figure 4.1 magnetic characteristic

A more general way to express the relation of \mathbf{B} and \mathbf{H} is to include the magnetization \mathbf{M} . Here the Sommerfeld unit system is employed because it avoids μ_0 in relating \mathbf{M} and \mathbf{H} : $\mathbf{B} =$

$\mu_0(\mathbf{H} + \mathbf{M})$. For a linear material, $\mathbf{M} = \chi\mathbf{H}$, where χ is a constant. Because the divergence of \mathbf{B} is zero, and

$$\mathbf{B} = \mu_0(1 + \chi)\mathbf{H} = \mu\mathbf{H}$$

the divergence of \mathbf{H} is also zero inside of a linear material. In general

$$\text{div}(\mathbf{H} + \mathbf{M}) = 0 \text{ so } \text{div}\mathbf{H} = -\text{div}\mathbf{M}$$

4.2 SURFACE POLARITY

Inside of a linear material the divergence \mathbf{H} is zero and the divergence of \mathbf{M} is zero. At the surface of the material, the interface between the magnetic material and a non-magnetic region, there may well be a discontinuity of the magnetization. This occurs whenever there is a component of magnetization perpendicular to the surface. The discontinuity arises because the normal component of magnetization exists inside the magnetic material, but not in the exterior nonmagnetic region. In such a case there will also be a divergence of the magnetic field \mathbf{H} at the surface.

One viewpoint is to consider the continuity of the magnetic flux density, which cannot change as it crosses the surface separating the magnetic material from the nonmagnetic region. Designating the normal components as H_{ni} and M_n just inside the magnetic material and H_{no} just outside or in the nonmagnetic region, we have the relation that

$$B_n = \mu_0(H_{ni} + M_n) = \mu_0 H_{no}$$

so that $H_{no} - H_{ni} = M_n$.

The normal component of magnetization at the surface is a source of magnetic field intensity, and is experienced in the form of *magnetic*

poles. The magnetization is a volume density of magnetic moments, and its normal component at a surface results in a surface pole density. This is entirely analogous to a surface electric charge density on an electric conductor. The mathematical treatment of Gauss' Law for electric charge density is applicable here. For an element of surface area dS , with a normal component of magnetization M_n , the field element at a point p will be:

$$d\mathbf{H} = \mathbf{1}_{s,p} \frac{M_n dS}{4\pi r_{sp}^2} \quad (4.1)$$

where $\mathbf{1}_{s,p}$ is a unit vector directed from the surface element toward the point p , and r_{sp} is the corresponding distance. In order to determine the total field at the point p it will be necessary to perform an integration over the entire surface.

As we are first interested in the effects quite close to the surface, it is reasonable to take the surface as an infinite plane. In performing such an integration it is perhaps easiest to set up an element of surface which is made up of a ring centered below the point p , from which only a normal component of field will result, and which, when integrated over the infinite plane, will yield a result that $H_{no} = M_n/2$.

4.3 DEMAGNETIZING FIELD

Recalling that $H_{no} - H_{ni} = M_n$, it is seen that with $H_{no} = M_n/2$, it is necessary that $H_{ni} = -M_n/2$, so that the surface poles cause fields in both the nonmagnetic region and in the magnetic material. The interior field is in the opposite direction to the normal component of magnetization, and for that reason is called a demagnetizing field, \mathbf{H}_d .

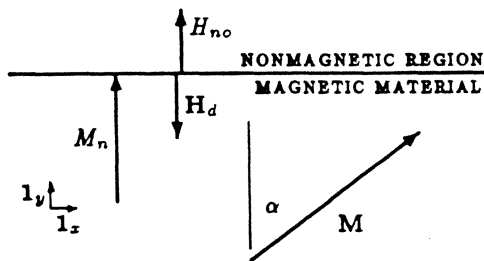


figure 4.2 Demagnetization

The magnetization is magnetic moment per

unit volume, and as such is a spatial average over a macroscopic point. Each individual moment within a macroscopic point is subject to torque due to the local magnetic field, which act to align the magnetic moment with the field. Thus, the demagnetizing field caused by a magnetization component normal to a surface tends to cause the normal component of magnetization to decrease. This can be seen in figure 4.2 where the magnetization is shown at an angle α to the surface normal, together with the normal component and the resulting demagnetizing field \mathbf{H}_d . The torque on a magnetic moment is the vectorcross-product, so the torque per unit volume on the magnetization is similar: $\mu_o \mathbf{M} \times \mathbf{H}$. The cartesian coordinate system shown corresponds with: $\mathbf{M} = \mathbf{1}_x M \sin \alpha + \mathbf{1}_y M \cos \alpha$, and $\mathbf{H}_d = -\mathbf{1}_y (M/2) \cos \alpha$, so the torque is:

$$\begin{aligned} \mathbf{T} &= \mu_o \mathbf{M} \times \mathbf{H}_d = -\mathbf{1}_z \frac{M^2}{2} \sin \alpha \cos \alpha \\ &= -\mathbf{1}_z \frac{M^2}{4} \sin 2\alpha \end{aligned}$$

It is seen that the torque is maximum at an angle of 45° , and has minima at zero and 90° . The minimum torque at $\alpha = 0$ can be shown to be an unstable equilibrium point (as any disturbance in that angle will tend to increase the angle), and the minimum at $\alpha = 90^\circ$ is a stable equilibrium for the magnetization. Thus, in the absence of any other torques on the magnetization, it would assume an orientation parallel to the surface.

4.3.1 DEMAGNETIZATION WITH AN APPLIED FIELD

For the case where there is an applied field acting on the magnetic material, we can break that field into a component tangent to the surface and a component normal to the surface: $\mathbf{H}_A = \mathbf{1}_t H_t + \mathbf{1}_n H_n$ where $H_t = H_A \sin \theta$ and $H_n = H_A \cos \theta$ where θ is the angle the applied field makes with the normal to the surface as indicated in figure 4.3.

The total interior field is the vector sum of the applied field and the resulting demagnetizing field: $\mathbf{H} = \mathbf{H}_A + \mathbf{H}_d$. The magnetization will lie in the direction of \mathbf{H} and will be proportional: $\mathbf{M} = \chi \mathbf{H}$. As was previously developed,

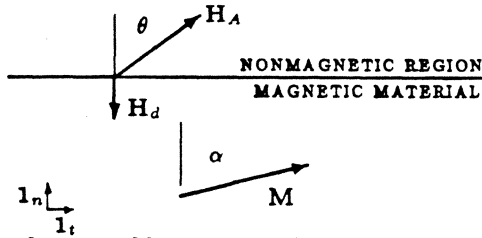


figure 4.3 Magnetization Near a Boundary

the demagnetizing field is $H_d = -\frac{1}{2}1_n M_n$. Then the total field can be written:

$$\mathbf{H} = 1_t H_t + 1_n (H_n - \frac{1}{2} M_n)$$

As $M_n = \chi(H_n - \frac{1}{2} M_n)$,

$$\mathbf{H} = 1_t H_A \sin \theta + 1_n H_A (1 - \frac{\chi}{2 + \chi}) \cos \theta$$

As M lies parallel to H , the angle the magnetization makes with the surface normal is found:

$$\tan \alpha = \frac{2 + \chi}{2} \tan \theta$$

For a susceptibility of $\chi = 100$, which is low for typical magnetic materials, for an applied field at 1° , the magnetization would be at an angle of 89 degrees to the normal.

For the following cases of magnetic circuits without nonmagnetic gaps, it is a reasonable approximation to assume the magnetization near a surface will be parallel to the surface.

4.3.2 DEMAGNETIZATION AT A GAP

For magnetic circuits with nonmagnetic gaps it is necessary to reconsider the foregoing material. Here we consider a relatively thin nonmagnetic region sandwiched between two magnetic regions which have similar magnetic properties.

The surfaces considered have large dimensions compared with thickness of the gap, and effects due to surface poles near the outer edges of the gap are considered far enough away to be negligible. The region being investigated is represented in figure 4.4. The nonmagnetic region has a thickness g , and the magnetization within the two magnetic regions is assumed to be the same (or at least the components normal to the surface are both the same).

As the magnetic material and the nonmagnetic material are both linear, the principle of

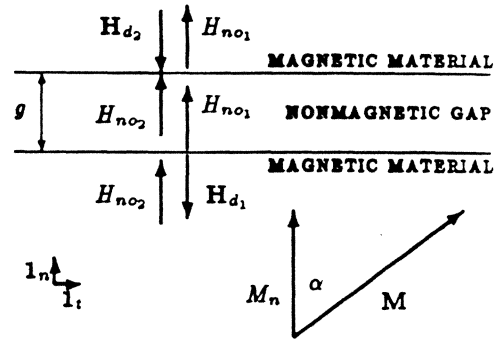


figure 4.4 Magnetization Near a Gap

superposition is applicable. The surface between the gap and the lower magnetic material has a positive surface pole density because the normal component of magnetization is directed toward the surface. This will result in normal field components pointing away from the surface, and having a magnitude of $\frac{1}{2} M_n$, indicated as H_{no1} (pointing upward, extending through the gap and through the upper magnetic material as well) and H_{d1} (pointing downward in the lower magnetic material).

The surface between the gap and the upper magnetic material has a negative surface pole density because the normal component of magnetization is directed away from the surface. This causes normal field components pointing toward that surface, and having a magnitude of $\frac{1}{2} M_n$, indicated as H_{no2} (pointing upward, but extending downward through the gap and into the lower magnetic material as well) and H_{d2} (pointing downward but extending upward through the upper magnetic material).

Applying superposition, it is seen that these normal field components add within the nonmagnetic material, and subtract within the magnetic materials, so there is no demagnetizing field associated with the gap.

The foregoing is an approximation, but is reasonably accurate in gap regions at least two gap-lengths ($2g$) away from any lateral edges of the gap.

4.4 SCALAR MAGNETIC POTENTIAL

For linear operation of soft magnetic materials with negligible hysteresis, the relation between

the magnetic flux density, \mathbf{B} and the magnetic field intensity, \mathbf{H} , is linear, and the divergence of \mathbf{B} is always zero, so that the divergence of \mathbf{H} within a linear magnetic material is also zero. Any vector whose divergence is zero can be expressed as the gradient of a scalar potential. Thus it is useful to express the magnetic field intensity within a linear magnetic material as the gradient of a magnetic scalar potential which we shall refer to as Ψ , which in the MKS unit system has the dimension of Amperes. Its gradient then has the dimensions of Amps per meter, and the relation is:

$$\mathbf{H} = -\text{grad } \Psi \quad (4.2)$$

As the divergence of \mathbf{H} is zero in the linear magnetic material, it is also true that $\text{div grad } \Psi = 0$, which is Laplace's Equation. The operation $\text{div grad } \Psi$ is called the Laplacian of Ψ , and is written in shorthand as $\nabla^2 \Psi$. This result is that the Laplacian of the magnetic scalar potential is zero. In cartesian coordinates it is written:

$$\nabla^2 \Psi = \frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} = 0 \quad (4.3)$$

EXAMPLE 4.4.1.

A block of linear magnetic material has a square cross-sectional area and a length ℓ , which is along the x-axis so that one end is at $x = 0$ and the other at $x = \ell$. Assuming the scalar magnetic potential is $100x$ within the block:

- (a) Determine the potentials at the two ends of the block.
- (b) Determine whether Laplace's Equation is satisfied within the block, and if so
- (c) Determine the magnetic field intensity inside the block.

SOLUTION OF EXAMPLE 4.4.1.

- (a) With $\Psi = 100x$, the potential at $x = 0$ is zero, and at $x = \ell$, $\Psi = 100\ell$.
- (b) With $\Psi = 100x$, the first partial derivative with respect to x is a constant, so the second partial derivative with respect to x is zero. As no other derivatives exist, Laplace's Equation is satisfied.
- (c) The gradient is:

$$\begin{aligned} \text{grad } \Psi &= \mathbf{1}_x \frac{\partial \Psi}{\partial x} + \mathbf{1}_y \frac{\partial \Psi}{\partial y} + \mathbf{1}_z \frac{\partial \Psi}{\partial z} \\ &= 100 \mathbf{1}_x \end{aligned}$$

Here it is seen that the field within the bar is $\mathbf{H} = -\text{grad } \Psi = -100\mathbf{1}_x$, which is perpendicular to the equipotential ends of the bar, and directed from the higher to the lower potential.

END OF EXAMPLE 4.4.1.

Laplace's Equation in circular cylindrical coordinates is written:

$$\nabla^2 \Psi = \frac{\partial^2 \Psi}{\partial \rho^2} + \frac{1}{\rho^2} \frac{\partial^2 \Psi}{\partial \phi^2} + \frac{\partial^2 \Psi}{\partial z^2} = 0 \quad (4.4)$$

EXAMPLE 4.4.2

A uniform ring of magnetic material is centered on the z-axis at the origin. It is assumed that the magnetic field is due to a current along the z-axis which will result in the field being circumferential within the ring (and elsewhere as a matter of fact). Prove that the equipotential surfaces will be planes radiating from the z-axis.

SOLUTION OF EXAMPLE 4.4.2

The easiest approach is to assume the answer and demonstrate that the field is circumferential. To do this, it is necessary to recognize that a plane radiating from the z-axis will include a particular radius vector, $\rho \mathbf{1}_\rho$, at a particular angle, ϕ . It is assumed that the equipotential surfaces are a linear function of ϕ alone: $\Psi = K_0 + K_1 \phi$. The Laplacian of this potential is obviously zero, and we determine the field from its negative gradient.

In the cylindrical coordinate system the gradient is:

$$\text{grad } \Psi = \mathbf{1}_\rho \frac{\partial \Psi}{\partial \rho} + \frac{1}{\rho} \mathbf{1}_\phi \frac{\partial \Psi}{\partial \phi} + \mathbf{1}_z \frac{\partial \Psi}{\partial z} \quad (4.5)$$

so the field is calculated:

$$\mathbf{H} = -\mathbf{1}_\phi \frac{1}{\rho} \frac{\partial (K_0 + K_1 \phi)}{\partial \phi} = -\mathbf{1}_\phi \frac{K_1}{\rho}$$

and it is seen to be only circumferential, which agrees with the assumption. Again it is seen that the equipotentials are perpendicular to the field vectors, and the field is directed in the direction of decreasing potential.

END OF EXAMPLE 4.4.2

At any point in a linear material the magnetic field, \mathbf{H} and the equipotential surface passing

through that point are perpendicular, as a result of the relation $\mathbf{H} = -\text{grad } \Psi$. This can be written on a small scale as $H\Delta\ell = -\Delta\Psi$, with the understanding that H and $\Delta\ell$ are aligned, so that this is equivalent to a vector dot product. This relation is illustrated in figure 4.5.

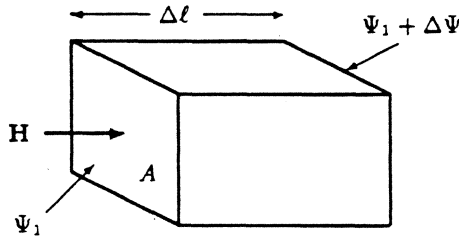


figure 4.5 magnetic element

The magnetic element in figure 4.5 is sufficiently small that there is no significant variation of \mathbf{H} over the area A , and the lateral boundaries are parallel to \mathbf{H} . As $\mathbf{B} = \mu\mathbf{H}$, the foregoing statements also apply to \mathbf{B} . From there it is a simple step to find that the element flux, $\Delta\Phi$, which is $\mathbf{B} \cdot \mathbf{A} = \mu H A$ enters the element through the surface marked A at potential Ψ_1 and leaves the element through the other end, which end is indicated by $\Psi_1 + \Delta\Psi$. Then the relation between the potential difference and the flux can be expressed as:

$$H\Delta\ell = \frac{\Delta\ell}{\mu A} \Delta\Phi = -\Delta\Psi \quad (4.6)$$

Then the relation of the flux, which enters the magnetic element at the higher equipotential surface and exits at the lower equipotential surface, is geometrically related to the drop in equipotential. It is to be noted that the flux does not pass through the lateral surfaces because those surfaces are defined as parallel to the field \mathbf{H} .

The relation of the element of flux to the drop in magnetic scalar potential is geometrical. This relation is often expressed in terms of the parameter reluctance (\mathcal{R}) or its inverse, permeance (\mathcal{P}). For the magnetic element this can be written as:

$$\mathcal{R} = \frac{1}{\mathcal{P}} = \frac{\Delta\Psi}{\Delta\Phi} = \frac{\Delta\ell}{\mu A} \quad (4.7)$$

4.5 MAGNETIC CIRCUITS

The concept of the magnetic circuit is useful in dealing with high permeability linear magnetic materials. It is based on Ampere's Law of Circuital Magnetism, and upon the concept of the continuity of magnetic flux. Ampere's Law states that the magnetomotive force, \mathcal{F} , around a closed path is equal to the current enclosed by the path. This is a definition based on the circulation integral of \mathbf{H} :

$$\mathcal{F}_c = \Sigma i_{\text{enclosed by } c} = \oint_c \mathbf{H} \cdot d\mathbf{l} \quad (4.8)$$

which is the counterpart of $\mathbf{H} = -\text{grad } \Psi$. The circulation integral can be broken into as many segments as necessary. For amenable geometries, it is possible to replace the integrations by a series of products of field and length on paths where the field is uniform:

$$\oint_c \mathbf{H} \cdot d\mathbf{l} = H_1\Delta l_1 + H_2\Delta l_2 + \dots + H_n\Delta l_n$$

It is convenient to conceive of each of the terms $H_j\Delta l_j$ drops of magnetomotive force between two equipotential surfaces, that is, mathematical surfaces which each have a constant magnetic scalar potential. The length of path between the equipotential surfaces must be short enough that the field intensity, \mathbf{H} , is reasonably constant along the path, and the path must be aligned with the direction of \mathbf{H} . When these conditions are met, we can call the resulting segment of the magnetic circuit a leg.

The continuity of magnetic flux is another way of saying the divergence of the magnetic flux density, \mathbf{B} , is zero. This requires that whatever flux enters a volume must also leave the volume: $\text{flux}_{\text{out}} = \text{flux}_{\text{in}}$

Then the properties of a leg are as follows:

- A leg is a length of magnetic material between two equipotential surfaces.
- The magnetization within a leg is parallel to the lateral surfaces of the leg.
- The magnetic flux which passes through one equipotential surface in the leg is the same as that passing through any other equipotential surface within the leg. Thus, the same flux which enters one end of the

leg exits at the other. That is to say that there is negligible leakage of flux from the sides of the leg.

The *leg* of a magnetic circuit is similar to a *magnetic element* with the difference that the field H need not be uniform over the equipotential surfaces of the leg. When considering a *segment* of a magnetic structure as a candidate for a *leg*, it is to be remembered that the magnetization prefers to align parallel to a lateral surface, which is no difficulty when the segment is straight. However, when a lateral surface is curved or contains a corner, it is necessary that the magnetization somehow conform to the surface.

The way that the magnetization and magnetic field conform to the lateral surfaces is by the creation of just enough magnetic *surface pole* density to cause the internal field and the magnetization to nearly conform to the lateral surfaces. By not quite conforming to the surface, there will be a normal component of magnetization, which gives rise to the surface pole density as discussed in section 4.3. Within the *linear* magnetic material M and H are colinear (i.e. they have identical orientations at any internal point). For this reason, the assumption of no leakage flux from a magnetic leg is only an approximation. There must be sufficient leakage flux to cause the magnetization to conform to the lateral surfaces.

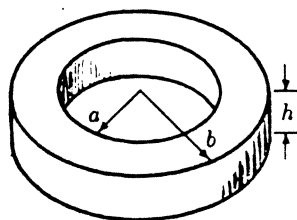


figure 4.6 Toroid Core

The simplest example of a leg is taken from the toroid having N turns of wire threaded through the aperture, so the magnetic field within the toroid is entirely in the circumferential direction, and can be expressed as the circumferential component:

$$H_\phi = \frac{Ni}{2\pi\rho} \quad (4.9)$$

The toroid core has an inner radius of a and an outer radius of b , with a thickness of h , as shown in fig. 4.6.

It is assumed that there is a total mmf of Ni amperes threading the aperture of the core, so the total potential drop around the toroid is also Ni .

An arbitrary section of the toroid can be taken as a leg, as indicated in figure 4.7. The equipotential surfaces indicated as Ψ_1 and Ψ_2 are radial surfaces which would pass through the axis of symmetry of the toroid if extended that far. Using the cylindrical coordinate system, the potential on an arbitrary equipotential surface can be written as $\Psi(\phi) = K_0 + K_1\phi$, where ϕ is the coordinate angle, not to be confused with the flux Φ .

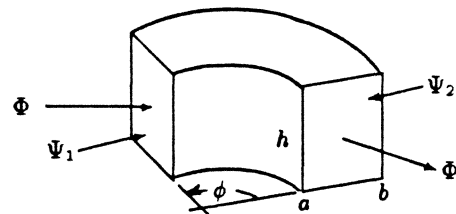


figure 4.7 magnetic circuit leg

It can be shown that with potentials as given above, the magnetic field is perpendicular to the potentials and thus circumferential in direction, as is to be expected from Ampere's Law of Circuital Magnetism.

Chapter 5

RECORDING HEADS AND FRINGE FIELDS

5.1 INTRODUCTION

In magnetic recording, a strong, localized field is needed to change the direction of magnetization in semi-hard (nonlinear) magnetic materials. A piece of copper wire could provide enough field when *soft* materials such as iron wire were used for recording in the early times, and as Valdemar Poulsen used in his original invention in 1898.

A current carrying wire produces a field somewhat less than the *long straight wire* field used to explain Ampere's law. For the long straight wire carrying a current I , the field is the current divided by 4π divided by the distance from the axis of the wire. American Wire Gage (AWG) copper size #40 has a radius of 0.0508 mm, which means that the maximum field (at the surface of the wire) is about 3000 A/m for 1 ampere of current flowing in the wire. The current capacity of this wire size is about 20 mA, so the maximum practical field it could provide is about 60 A/m, or about 0.8 oersteds. AWG #30 copper wire has a radius of .152 mm and has a current capacity of about 100 mA, to provide a maximum field of about 100 A/m or 1.3 oe. at the surface of the wire. These are strong enough to magnetize iron, but iron is not a practical recording material, because it is magnetically soft and can easily be demagnetized. Practical recording materials require fields of from 300 to 1500 oe. which cannot be achieved with a simple copper wire.

A magnetic circuit can be used to amplify the effective current by the number of turns, and can concentrate the flux in a very local region. However, to provide a recording field, the flux

must be released from the core and directed into the recording media. The simplest method of releasing the flux from its conduit is to put a nonmagnetic gap into the circuit, breaking it open and allowing the flux to spill out.

Magnetic flux is continuous, and is contained by magnetic materials which behave as conduits of flux, bending the flux gently around curves in the surface. The mechanism for the shaping of the field to fit the shape of the magnetic material is by surface poles, which act as further sources of magnetic fields. Surface poles are caused by the abrupt change of magnetization at the interface between a magnetic material, which has magnetization, and the nonmagnetic material which does not. This is most readily seen by studying the divergence of the magnetic flux density, which is zero:

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}) \quad (5.1)$$

As the divergence of \mathbf{B} is zero, the divergence of $\mathbf{H} + \mathbf{M}$ must also be zero. This means that: $\text{div } \mathbf{H} = -\text{div } \mathbf{M}$. The divergence of magnetization is therefore a source of magnetic field.

If there is a magnetic pole density ρ_m , then we can use Gauss' Law and Gauss's Theorem to establish that with $\text{div } \mathbf{H} = \rho_m$, the field due to such a pole density is related through Gauss's Theorem as:

$$\iint_{S_{\text{vol}}} \mathbf{H} \cdot \mathbf{1}_n dS = \iiint_{\text{vol}} (\text{div } \mathbf{H}) dV \quad (5.2)$$

Which is Gauss's Theorem, which can be applied to any conservative vector field. This is to be understood as follows: The normal component of the magnetic field intensity is integrated

over the entire surface of a volume, and the net result is the same as if the divergence were integrated over the volume bounded by that surface. Thus S_{vol} of eq. 5.2 is the surface of vol on the right side. Using the fact that the divergence of \mathbf{H} is ρ_m , eq. 5.2 can be written as:

$$\iint_{S_{vol}} \mathbf{H} \cdot \mathbf{1}_n dS = \iiint_{vol} \rho_m dVol \quad (5.3)$$

Equation 5.3 can be further manipulated to give the algorithm for calculating an element of \mathbf{H} at a field point due to an element of charge density at a source point. This is:

$$d\mathbf{H}_\rho = \frac{\rho_m \mathbf{1}_{sf}}{4\pi r^2} dVol \quad (5.4)$$

where r is the distance from the source point to the field point, and $\mathbf{1}_{sf}$ is the unit vector pointing from the source point toward the field point. The subscript of ρ is used on \mathbf{H}_ρ to indicate that this is a component of \mathbf{H} that is due to magnetic pole density. This magnetic pole density is due to the divergence of magnetization:

$$\rho_m = -\text{div } \mathbf{M} \quad (5.5)$$

As was previously discussed, a linear magnetic material has $\mathbf{M} = \chi\mathbf{H}$, so that there is no divergence of \mathbf{M} within linear materials, but can occur at the surfaces. This is the primary reason that magnetic circuits have the ability to contain flux. When the magnetization would otherwise have a component perpendicular to a surface, surface poles are created by the discontinuity of magnetization which are then sources of \mathbf{H} , which create a torque on the magnetization which tends to make the magnetization parallel to the surface.

The flux containment of magnetic circuits provides another benefit, as the fields due to the surface poles will compensate for lack of symmetry in the current configurations.

5.2 RELUCTANCE OF NONMAGNETIC GAPS IN MAGNETIC CIRCUITS

Gaps in magnetic circuits are of great technological interest, because magnetic circuits are able to contain and concentrate magnetic flux.

In electrical machinery very strong fields are obtained in air gaps and are used to provide the emf for generators and torque for motors.

The analysis of the fringe field around a gap will be undertaken later, we first must understand the effects of gaps on the magnetic circuit itself. From extensive experience in the measurement of reluctance of air gaps in electrical machines, where gaps which are short when compared with the dimensions of the so-called pole-faces, the reluctance is found from the length of the gap, g , and an adjusted area of the faces. The pole-faces are surfaces so designed that they would have been equipotentials had the gap not been introduced.

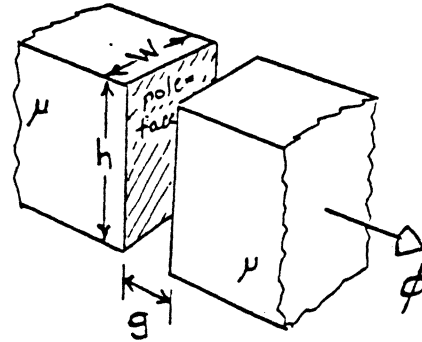


fig. 5.1. Gap pole faces.

For a rectangular cross-sectioned magnetic circuit as shown in figure 5.1, they have dimensions of width w and throat height h , so the pole-face area is hw . The gap length (between the pole faces) is g . The adjustment necessary to reasonably approximate the reluctance is to increase the effective area by adding the gap length to both w and h before calculating the effective area of the pole faces. Thus, following from equation 4.7 in the previous chapter, the reluctance of this gap is:

$$\mathcal{R}_{gap} = \frac{g}{\mu_o(h+g)(w+g)} \quad (5.6)$$

In ferrite recording heads, to obtain very flat pole faces, the head is made in two parts, to allow the two pole faces can be machined and polished. These two halves of the head are then put together to create a magnetic circuit having two gaps, one called the front gap, and the other the back gap. A simplified head geometry is shown in figure 5.2

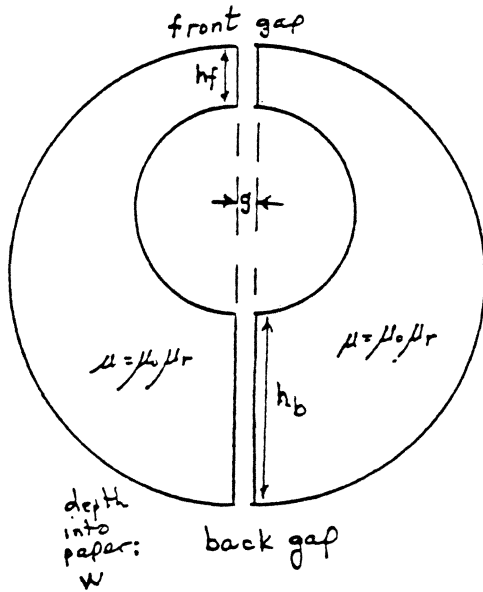


fig. 5.2. simplified head geometry.

For this particular example, for purposes of illustration, the height of the back gap (h_b) is 3.25 times the height of the front gap (h_f). The reluctance of the front gap is:

$$\mathcal{R}_{fg} = \frac{g}{\mu_0 w h_f (1 + g/h_f)(1 + g/w)}$$

and for the back gap:

$$\mathcal{R}_{bg} = \frac{g}{\mu_0 w h_b (1 + g/h_b)(1 + g/w)}$$

The total reluctance of the magnetic circuit including the nonmagnetic gaps (typically called air-gaps, although the material is not usually air) is:

$$\mathcal{R}_{total} = \mathcal{R}_{core} + \mathcal{R}_{fg} + \mathcal{R}_{bg} \quad (5.7)$$

Then it is necessary to determine the reluctance of the magnetic material portion of the circuit, referred to as the core.

5.2.1 MAGNETIC CORE RELUCTANCE

To deal with the magnetic material in the magnetic circuit of figure 5.2, we eliminate the air gap for the moment, to generate a form as shown in figure 5.3. The gaps are in series with this part, and so will be added in later.

A non-mathematical technique for the determination of the relation between the MMF and

flux is useful in certain geometries where the thickness is constant. This makes use of the reluctance relation given in eq. 4.7, which is repeated here for convenience:

$$\mathcal{R} = \frac{1}{\mathcal{P}} = \frac{\Delta\Psi}{\Delta\Phi} = \frac{\Delta l}{\mu A}$$

For a rectangular leg, the cross-sectional area is the product of its width and height, $A = w \times h$.

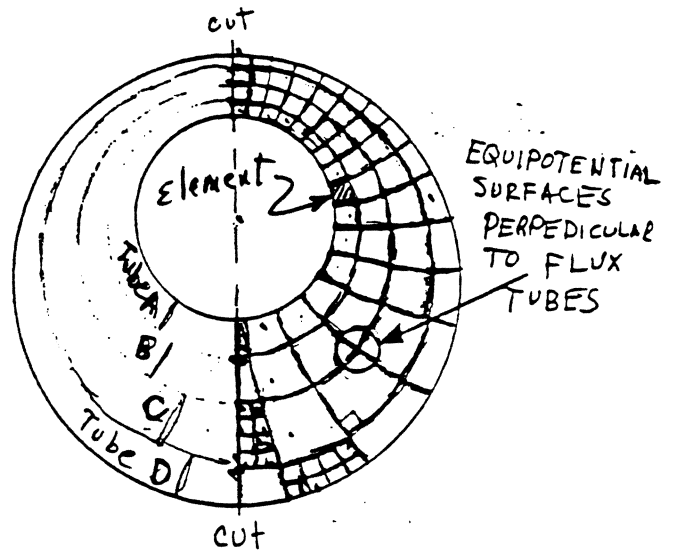


fig. 5.3. circuit with gap removed.

The magnetic circuit can be understood to be composed of tubes of flux, which close upon themselves in enclosing the MMF. The potential drop around each flux-tube circuit is equal to the MMF. The element of figure 5.3 is part of one tube of flux, indicated as tube A. No flux leaves a tube, though the tube may change in cross-sectional area in making the circuit. The quantity which is the same throughout the circuit of a tube is the element's flux $\Delta\Phi$. The potential drop from one end of the element to the other is:

$$\Delta\Psi = -\frac{\Delta l}{\mu h \Delta h} \Delta\Phi \quad (5.8)$$

where the negative sign indicates that the potential is actually dropping in the direction of the flux.

The magnetic circuit geometry is subdivided into equipotential surfaces and a number of flux

tubes which are perpendicular to the equipotential surfaces. The thickness of the circuit is uniform at w , and elements of a tube are constructed so that the length to width ratio (Δl to Δh) is 1.0 on all elements. In doing this, the reluctance of each element is the same as any other. Thus the total reluctance of the tube is simply the number of cells times the reluctance per element:

$$\mathcal{R}_{tube} = n_c \frac{1}{\mu w}$$

If the flux tubes are chosen so that they each contain the same flux, then the total flux will be the number of tubes (n_t) multiplied by the flux in one tube. As reluctance is mmf divided by flux, the total reluctance will then be the reluctance of one tube divided by the number of tubes:

$$\mathcal{R}_{core} = \frac{n_c}{n_t} \frac{1}{\mu w} \quad (5.9)$$

In figure 5.3 the inner and outer surfaces are circles, and they are the inner and outer surfaces of the total flux tube. We can divide this into smaller tubes by using other circular shapes as indicated on the right hand side, where the inner tube was laid down first by drawing a circle through a point on the top cut a quarter of the way from the inner to the outer surface, and a point on the bottom cut a quarter of the way from the inner to the outer surface.

Next, the cells were drawn in on the inner flux tube, making them as nearly square as possible, keeping the rule that flux tubes and equipotential surfaces must be perpendicular to one another.

The next flux tube was constructed by extending the equipotentials of the first tube, and then fitting the best circle to cause those extensions and the circle to make the most nearly square cells possible. In doing this we assure that the flux tubes contain the same flux.

The third tube was constructed using the same technique as the second.

A full fourth tube would not fit in, so the equipotentials were extended to the outer surface. Then one of the oblong cells was divided up into the best squares to determine how wide it was. Because there are four squares by two squares in this oblong, the width is one half the length, so the outer tube contains only half the

flux of the others. This gives a total number of tubes as $n_t = 3.5$.

In counting the number of cells in a tube, on the half drawn there are 15 and a fraction. To determine that fraction, a convenient shortened cell is divided into the best squares and we find it is 4 wide by 2 long, so its length is a half a cell. Thus there are 15.5 cells in half the tube circuit for a total of 31 cells in a tube of flux.

The total reluctance of the magnetic material portion of the circuit is then:

$$\mathcal{R}_{circuit} = \frac{31}{3.5} \frac{1}{\mu w}$$

A representative recording head has $h_{fg} \approx w \approx 25g$, (with $h_{bg} = 3.25h_{fg}$). Taking these values in this example, we have that:

$$\begin{aligned} R_{tot} &= \frac{1}{\mu_o w} \left(\frac{31}{3.5\mu_r} + \frac{g}{1.025h_{fg}} + \frac{g}{1.025h_{bg}} \right) \\ &= \frac{1}{\mu_o w} \left(\frac{8.86}{\mu_r} + 0.0474 \right) \end{aligned}$$

where μ_r is the relative permeability of the core material. Practical head materials have relative permeabilities from about 1000 to 10,000. At the low end, we can see from eq. 5.9, the core will contribute about 16% of the total reluctance, and at the high end it would contribute about 2%. With a typical relative permeability of 5000, the core contributes about 4% of the total reluctance. The use of more massive cores can reduce this to 1 or 2% of the total reluctance.

5.3 HEAD EFFICIENCY

The purpose of the recording head is to provide a strong local magnetic field. This field is the fringe field of the front gap, which we would like to have as large as possible. The potential drop across the front gap divided by the gap length will be the field at the geometrical center of the pole faces, and this will be the same over most of the gap, beginning to drop off at about one gap-length distance from the pole edges.

The total potential drop around the magnetic circuit is the MMF due to currents threading the aperture. The potential drop of each part of the circuit will be the flux multiplied by the reluctance of that part. In order to maximize

the potential drop of the front gap, we would like to make the reluctance of the front gap large compared with the other reluctances in the circuit. Eq. 5.7 shows the reluctances of the core, the front gap and the back gap. The relative permeability and size of the core determines its reluctance. The throat height of the back gap is the only factor which makes it different from the front gap in the simple head of figure 5.3. To minimize the effect of the back gap, it is necessary to make h_{bg} much larger than h_{fg} . Typically the ratio of h_{bg}/h_{fg} is more than 10:1, and as much as 20:1. Thus, to make a very efficient head, the back gap throat height h_{bg} must be much greater than h_{fg} , and the core material must have a very high relative permeability.

The head efficiency η is the ratio of the front gap reluctance to the total reluctance of the magnetic circuit:

$$\eta = \frac{100\mathcal{R}_{fg}}{\mathcal{R}_{fg} + \mathcal{R}_{bg} + \mathcal{R}_{core}} \quad (5.10)$$

The deep gap field in the front gap is the potential drop across the gap divided by the gap length: $H_g = \Psi_{gap}/g$, and the potential drop across the front gap is the MMF multiplied by the fractional head efficiency:

$$\Psi_{gap} = gH_g = \frac{\eta}{100} MMF \quad (5.11)$$

5.4 KARLQVIST'S FRINGE FIELD

In our example with a relative permeability of 5000, we found the core had only 4% of the potential drop. With higher permeabilities the potential drop around the core becomes even less. It is a short step to assume infinite permeability of the core, and thereby make the core parts of the magnetic circuit equipotential volumes. In practice the potential along the core changes very slowly as compared with the rate of change across the gaps. In addition, the length of the gaps is small compared with the length of the pole tips for ferrite heads. The model used by O. Karlqvist was infinitely long pole pieces separated by g , each with infinite width w and infinite throat height h_f . The geometry of his model is shown in figure 5.4. The left pole is assumed to be at a potential of $gH_g/2$, and the right pole at $-gH_g/2$, making

the midplane of the gap to be the zero equipotential, and is taken as the plane $x = 0$ (with the positive direction to the right in figure 5.4). The outer surface of the core is taken to be the plane $y = 0$, ignoring any curvature, and the positive y -direction is outward from the core (up in figure 5.4). The other coordinate is z , which has its positive direction out of the plane of the paper. The plane $z = 0$ is half-way across the head width (w). We shall carry out the analysis on the plane $z = 0$, assuming no variations in the z -direction. The problem then reduces to one of 2-dimensional potential theory.

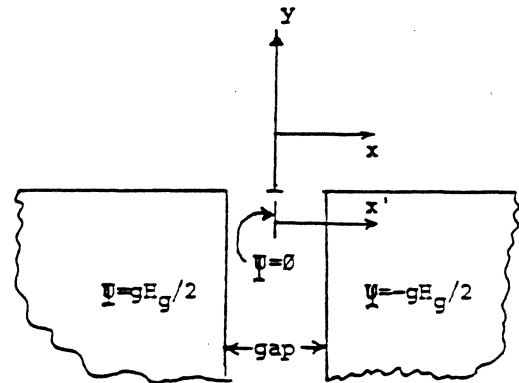


fig. 5.4. Karlqvist head geometry.

In 2-dimensional potential theory, the potential in the $y > 0$ half-plane can be calculated if the potential on the $y = 0$ axis (x -axis) is known. Except for along the top of the gap, called the gap region, the potential is known. Deep within the gap, the potential can be found to be $\Psi = -H_g x'$. For this analysis we assume that same potential distribution holds along $y' = 0$, in the gap region, from $x' = -g/2$ to $x' = g/2$. The potential for $y > 0$ is computed from:

$$\Psi(x, y) = \frac{1}{\pi} \int_{x'=-\infty}^{\infty} \Psi(x', 0) \frac{y dx'}{(x-x')^2 + y^2} \quad (5.12)$$

where (x, y) is the field point, and $(x', 0)$ is the source point and the integration is over the source point on x' from negative infinity to positive infinity. The integration is performed and

the result is:

$$\Psi(x, y) = \frac{H_g}{\pi} \left\{ \begin{aligned} &(x - \frac{g}{2}) \arctan \frac{x - \frac{g}{2}}{y} \\ &- (x + \frac{g}{2}) \arctan \frac{x + \frac{g}{2}}{y} \\ &+ \frac{y}{2} \ln \frac{(x + \frac{g}{2})^2 + y^2}{(x - \frac{g}{2})^2 + y^2} \end{aligned} \right\} \quad (5.13)$$

The field components, H_x and H_y , are found by taking the negative gradient of Ψ , which is a rather tedious process, but which eventually results in the following:

KARLQVIST FRINGE FIELD EQUATIONS

$$H_x = \frac{H_g}{\pi} \left\{ \begin{aligned} &\arctan \frac{x - \frac{g}{2}}{y} \\ &- \arctan \frac{x + \frac{g}{2}}{y} \end{aligned} \right\} \quad (5.14)$$

$$H_y = \frac{H_g}{2\pi} \ln \frac{(x - \frac{g}{2})^2 + y^2}{(x + \frac{g}{2})^2 + y^2} \quad (5.15)$$

In these equations, lengths are commonly normalized to half the gap length $g/2$, and upper case variables are used for normalized lengths: $X = 2x/g$, and $Y = 2y/g$.

The field components are normalized to H_g , and lower case h's are used: $h_x = H_x/H_g$ and $h_y = H_y/H_g$. Then these equations in normalized form are :

NORMALIZED KARLQVIST EQUATIONS

$$h_x = \frac{1}{\pi} \left(\arctan \frac{X + 1}{Y} - \arctan \frac{X - 1}{Y} \right) \quad (5.16)$$

$$h_y = \frac{1}{2\pi} \ln \frac{(X - 1)^2 + Y^2}{(X + 1)^2 + Y^2} \quad (5.17)$$

These expressions give components which have constant values which are circles in the z -plane, (actually cylinders parallel to the z -axis in three dimensions).

For the perpendicular component $h_y = \text{constant}$, the cylinders are centered on the $y = 0$ plane, and have field directions upward for negative values of x and downward for positive values of x as indicated in figure 5.5. Cylinders with equal radii have the same magnitude of constant field.

For the perpendicular component $h_y = \text{constant}$:

cylinder axis is line: $X = -\coth \pi h_y$, $Y = 0$
 cylinder radius of $R^2 = \text{csch}^2 \pi h_y$

The x -component is called the longitudinal component because the relative motion between

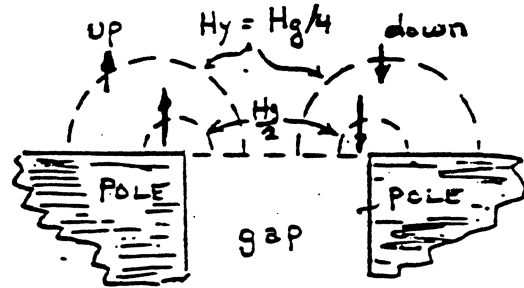


fig. 5.5 Karlqvist cylinders of H_y .

the head and medium is in the x -direction. The cylinders for the longitudinal component $h_x = \text{constant}$ are shown in figure 5.6. These cylinders are tangent to the corners of the polepieces.

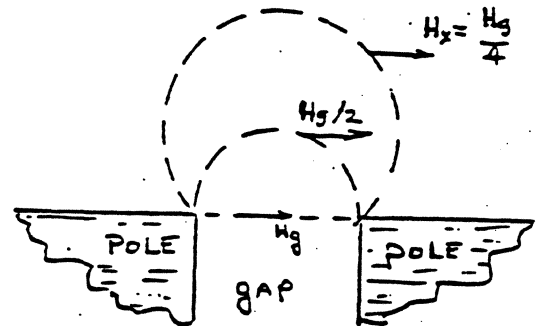


fig. 5.6. Karlqvist cylinders of H_x .

For the longitudinal component $h_x = \text{constant}$:
 cylinder axis is line: $X = 0$, $Y = \cot \pi h_x$
 cylinder radius of $R^2 = \text{csc}^2 \pi h_x$

We are primarily interested in the longitudinal component because most recording media is essentially longitudinal, i.e. where the recorded magnetic moments are essentially in the x -direction. There may yet be interest in perpendicular media, where the magnetization is in the y -direction.

The Karlqvist Head function is an approximation, but has been found to be quite adequate for many purposes. In some cases a simpler model is useful in understanding the complex process of magnetic recording.

5.5 FAR-FIELD MODEL

When the recording medium is at least one

gap length away from the gap, the gap-length can be reasonably ignored in calculating the field. The Karlqvist's result is not easily modified to obtain this result, so one must begin with eq. 5.12 and set the x -axis potentials to $gH_g/2$ from negative infinity to $x = 0$, and $-gH_g/2$ from $x = 0$ to positive infinity. This eliminates one integration and yields:

$$\Psi = \frac{gH_g}{\pi} \arctan \frac{y}{x}$$

Then, taking negative partial derivatives to obtain the field components we obtain that:

FAR FIELD EQUATIONS

$$H_x = \frac{gH_g}{\pi} \frac{y}{x^2 + y^2} \tag{5.18}$$

$$H_y = -\frac{gH_g}{\pi} \frac{x}{x^2 + y^2} \tag{5.19}$$

The locus of points where H_x is constant, from eq. 5.18 is where

$$x^2 + y^2 - \frac{gH_g}{\pi H_x} y = 0$$

As the equation of a circle is $(x - x_c)^2 + (y - y_c)^2 = r^2$, where the point (x_c, y_c) is the center of the circle, and r is the radius, equation 5.18 can be put into the same form by adding the term $(gH_g/2\pi H_x)^2$ to both sides (completing the square). Thus the locus of points where H_x is of constant magnitude is a cylinder with its axis through the point $(0, gH_g/2\pi H_x)$ and parallel to the z -axis and having a radius of $(g|H_g|/2\pi|H_x|)$. This is shown in figure 5.7.

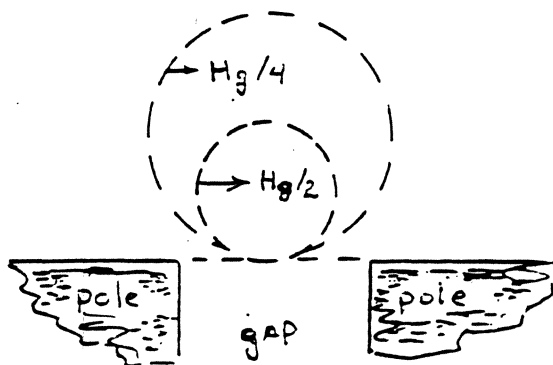


fig. 5.7 Far field H_x cylinders.

Similarly, equation 5.19 can be put into the circular form for constant H_y , with a cylindrical locus with the axis through the point $(-gH_g/2\pi H_y, 0)$ and parallel to the z -axis with a radius of $(g|H_g|/2\pi|H_y|)$, as indicated in figure 5.8.

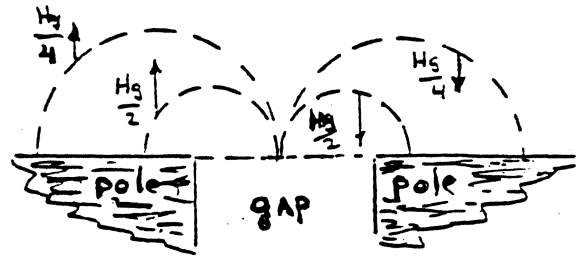


fig. 5.8. Far field H_y cylinders.

Chapter 6

THE RECORDING PROCESS

The recording process is complex, involving head fields, media characteristics and the demagnetizing fields which arise because of the magnetic record. The demagnetizing effects alter the recorded pattern, complicating the process to a point where it is nearly incomprehensible to the student. To simplify the procedure of learning about recording, the demagnetizing field effects are postponed to the next chapter. Thus the development in this chapter results in an idealization of recording which is to be taken as the limiting case. Actual recording can only be worse than what we find in this chapter.

6.1 CHARACTERISTICS OF RECORDING MEDIA.

The magnetic materials employed for recording are generally classified as semi-hard magnetic materials. The linear materials we have discussed for magnetic circuits are soft magnetic materials, in that they are well characterized by permeability. Permanent magnets are hard materials, and are characterized by an energy product, which will not be covered in this course.

Recording materials are classified primarily by their coercive force, H_c , and a squareness ratio. Figure 6.1 shows a typical saturation $M - H$ characteristic of this type of material, with the coercive force H_c , saturation magnetization M_s , and saturation remnant magnetization M_{rs} indicated. The squareness ratio is M_{rs}/M_s . For this characteristic can be either static, where the field is increased an increment and the magnetization measured at that field in a vibrating sample magnetometer (VSM), and then the field incremented and the measure-

ment repeated, or in a dynamic state, using a $B - H$ loop under the assumption that $M = B/\mu_0$.

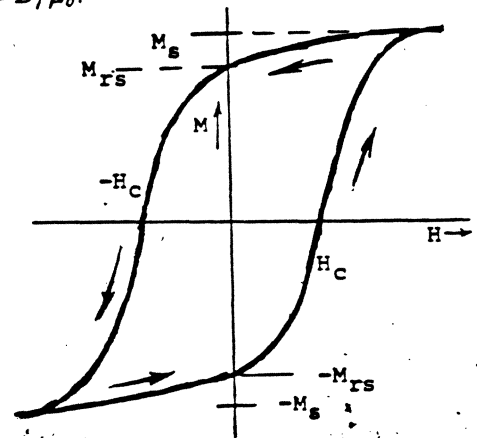


Figure 6.1. Saturation hysteresis loop for a magnetic recording medium.

The saturation $M - H$ characteristic, or $M - H$ loop as it is usually called, is of some value for recording, however it only approximates the information needed to study the recording process. In the recording process the recording medium moves past the head, experiencing a field while it is near the head, and then passes beyond the head field to a region of only ambient fields. Thus it receives a field and then relaxes to a zero field condition with whatever remnant magnetization it received from the recording field. Thus a characteristic for recording materials in which the remnant state which occurs after a field has been applied and then removed is of greater interest.

Data and video recording use a technique which is called non-return-to-zero (NRZ) or saturation recording. This is largely used in instrumentation recording and in digital audio recording as well. In this technique, information is

stored in the form of reversals of the magnetization from one direction to the other, i.e. from the positive longitudinal direction to the negative longitudinal direction (the x -direction with respect to the head coordinates in figure 5.4). These reversals of magnetization are commonly called flux changes or flux reversals, and information is coded using these flux changes. In this process recording is accomplished by driving the head winding current first in one direction and then reversing its direction. For each direction the current must be sufficient to saturate the local recording medium. So recording takes place by reversing the direction of a saturating field so as to reverse the direction of magnetization, so that the recording medium will experience a full saturating field except during the short time interval required to reverse the current direction. For this reason, the sort of an $M - H$ characteristic needed is one where the material begins at one saturation level and is driven with a certain field in the reversed direction. The characteristic of interest in recording is then a quasi-static one, which is shown on the following page in figure 6.3, and is called the remnant magnetization characteristic. This characteristic is not a loop as would be obtained with the dynamic $M - H$ characteristic, but is a two branched curve, one branch at negative remnance curving upward to the point where the field just gives full remnant magnetization. The other begins at positive remnance and curves downward to the negative field that just reaches negative remnance. These two branches do not intersect.

The remnant magnetization characteristic is obtained by a series of measurements, each beginning by driving the magnetization firmly to negative saturation. Then a positive field is applied for a short period of time and then reduced to zero, so that the magnetization returns to an intermediate remnant state. The measurement system required is a vibrating sample magnetometer (VSM). The applied field is measured when applied, and the remnant magnetization is then measured after the field is reduced to zero. This process is indicated in figure 6.2 where the dynamic saturation $M - H$ characteristic of figure 6.1 is shown as a dashed curve. The remnant magnetization is plotted against the field for a series of these tests, using increasing values of field for each succeeding

test, until the maximum remnant magnetization M_{rs} is obtained.

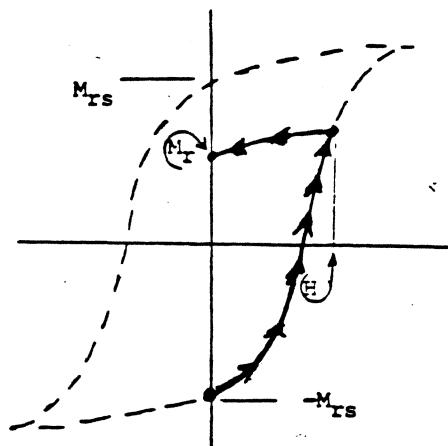


Figure 6.2. Generation of the remnant magnetization characteristic

The parameters of interest to us from the remnant magnetization characteristic are the the saturation remnant magnetization M_{rs} , the threshold field H_T and the switching or saturation field H_S , as indicated in figure 6.3. The threshold field H_T is the amount of field necessary to change magnetization from $-M_{rs}$ to $-0.9M_{rs}$, and the saturation field H_S is the field necessary to change the magnetization from $-M_{rs}$ to $+0.9M_{rs}$.

An ideal recording medium would have $H_S = H_T$

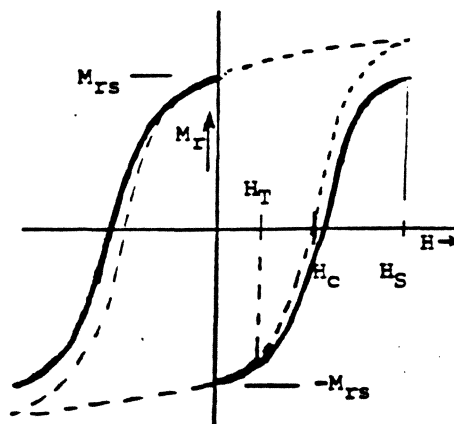


Figure 6.3. Remnant magnetization characteristic

6.2 LONGITUDINAL RECORDING.

At the present time, most magnetic recording is *in-plane*, that is the recording medium is in the form of tape or a disk which has a thin layer of recording medium coated or deposited on the surface, and the recording takes place essentially with the magnetization lying in the plane of the medium. For our purposes the principal component of field which affects in-plane magnetization is the longitudinal component of the head field (x -component in equations 5.14 and 5.16), so we shall concentrate on the action there. While there are experimental systems which use materials which magnetize only in directions normal to the plane, and as such utilize the perpendicular (y -direction in equation 4.15) field, we shall not consider perpendicular recording at this time.

Figure 6.4 is a schematic representation of the head-medium interface, with parameters of head-medium separation d , media thickness δ and gap length g . The recording media is attached to the substrate which is moving to the right past the gap with a velocity v , which is constant. The head has a width w perpendicular to the paper, and it is assumed there is no field variation across the width. For most purposes we shall use normalized values for the parameters d , δ and v , normalized to the half gap length $g/2$.

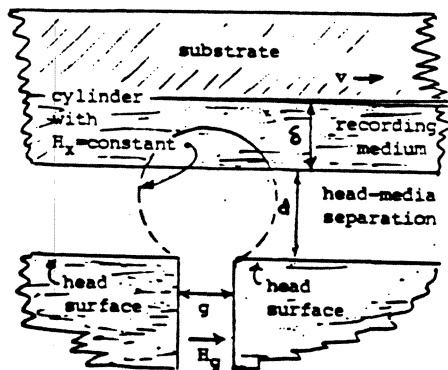


Figure 6.4. Head-medium interface

NORMALIZED PARAMETERS
head-media separation: $D = 2d/g$

medium thickness: $\Delta = 2\delta/g$

velocity: $V = 2v/g$

For saturation recording, the applied field must be sufficient to saturate the medium directly above the gap. We consider the cylinder of constant H_x (as discussed in chapter 5) which passes through that point, and adjust the gap field H_g so that the constant value of H_x on that locus is H_S . This insures that all points closer to the gap are saturated, as all points within that cylinder experience a field greater than H_S , neglecting the demagnetizing field, as illustrated in figure 5.6. The highest reach of a cylinder is at $x=0$, and the height that the $H_x = H_S$ cylinder must reach is $d + \delta$. Then setting $x = 0$, $y = d + \delta$, and $H_x = H_S$, we can solve for the minimum value of H_g to insure saturation recording. In this case we wish to use regular rather than normalized units, as we are setting H_g , the field used for normalization. Applying equation 4.14a, we obtain:

$$H_g = \frac{\pi H_S}{2 \arctan[g/2(d + \delta)]} \quad (6.1)$$

It is seen that the distance parameters in equation 6.1 become normalized in the process, so that normalized parameters can almost always be used. In general the field at the back of the medium ($X = 0$, $Y = D + \Delta$) should exceed H_S , so that equation 6.1 is the minimum for saturation recording, and

$$H_g \geq \frac{\pi H_S}{2 \arctan 1/(D + \Delta)} \quad (6.2)$$

All the magnetization within the constant H_x cylinder of $H_x = H_S$ is guaranteed to switch to the direction of the gap field. That is because all of the interior of a cylinder of constant H_x has H_x greater than that on the cylinder. Thus, when equation 6.2 is satisfied it is known that all media magnetization within the cylinder is in the direction of the gap field.

Let us assume that the gap field satisfies equation 6.2, and that the field has been on for some time. The medium is moving to the right with a velocity v , so that for some distance to the right of the gap the media is saturated in the direction of H_g . In figure 6.5 the situation is shown where the gap field is in the negative x -direction and has been in that direction for a

long enough time so that the the magnetization to the right of the cylinder marked $-H_S$, indicating that it is the cylinder of $H_x = -H_S$, is switched to the negative direction as indicated. Further to the left we see a cylinder marked $-H_T$, indicating it to be the cylinder for $H_x = -H_T$, which is just enough to begin switching some of the magnetization to the negative x -direction, if the magnetization at that cylinder were known to be saturated in the positive x -direction. The region between these two cylinders would be capable of being partially switched to the negative x -direction, however we know nothing of the previous history there, and so will leave everything to the left of the $H_x = H_S$ with a question mark.

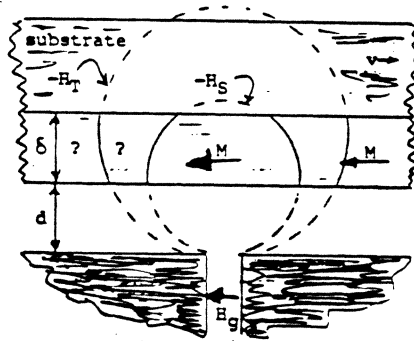


Figure 6.5. Condition before head current reversal

With the condition of figure 6.5 stabilized, the gap field is instantaneously reversed by reversing the head current. The new cylinders in figure 6.6 are seen to overlap the previously saturated region in the medium near the recording gap. Here the previous condition of the media to the right of the gap is well known, it was previously saturated to the left. Thus we can predict that the region to the right of the gap between the cylinders marked H_S and H_T are partially reversed, so that region can be considered to be a transition region between the earlier situation and the present condition.

It can be seen that the transition length is determined by the two material parameters H_S and H_T , together with the normalized (to the half gap-length) parameters D and Δ . The cylinder radii of the the loci of $H_x = H_S$ and

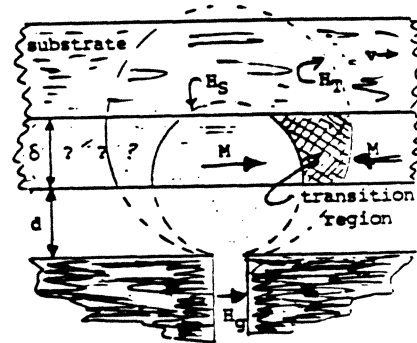


Figure 6.6. Condition at the time of head current reversal

$H_x = H_T$ have the ratio (see equation 5.14):

$$\frac{R_T}{R_S} = \frac{\sin \pi h_S}{\sin \pi h_T} \approx \frac{H_S}{H_T} \quad (6.3)$$

The approximation is within 5% when h_S is less than 0.55.

6.2.1 Example

For a particular recording system, the normalized parameters are $D = 1$ and $\Delta = 5$. The recording media has $H_T = 280$ oe. and $H_S = 580$ oe. We will set the gap field according to equation 6.2, and then determine the transition length at both the top and bottom of the medium. From equation 6.2 we obtain that H_g must be 5,517 oe., then $h_S = 0.330$ and $h_T = 0.159$. Then from equation 6.3, $R_T/R_S = 2.059$, while $H_S/H_T = 2.071$.

In order to solve for X from equation 4.15a, note that it can be manipulated to the form:

$$h_x = \frac{1}{\pi} \arctan \frac{2Y}{X^2 + Y^2 - 1} \quad (6.4)$$

so that

$$X^2 = 1 - Y^2 + 2Y \cot \pi h_x \quad (6.5)$$

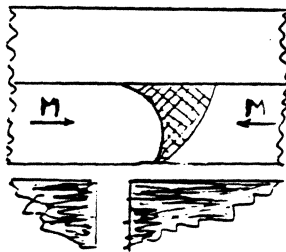
Then at $Y = D = 1$, $X_S = 2.416$, and $X_T = 3.532$, so the normalized transition width at the bottom of the medium is 1.115, which corresponds to 0.558g.

At $Y = D + \Delta = 6$, $X_S = 0$, and $X_T = 6.311$, which is also the transition length at the top of the medium, 3.165 g.

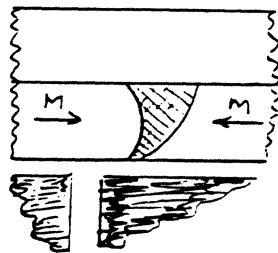
6.2.2 Example

Using the same system as example 6.2.1, but increasing H_g by 20% to 6,620 oe. h_S becomes 0.275 and h_T becomes 0.133. Then at $Y = D$ we obtain $X_S = 2.662$ and $X_T = 3.866$ for a normalized transition length of 1.204 or 0.602g. At $Y = D + \Delta = 6$, $X_S = 2.744$ and $X_T = 7.395$ for a normalized transition length at the top of the medium of 4.651 or 2.326g.

end of example



(a)



(b)

Figure 6.7. Transition regions for examples: (a) 6.2.1, (b) 6.2.2

The transition regions for these two examples are shown in figure 6.7 as the shaded areas. While they are quite similar in shapes, the second can be argued to be better, in that it is more concentrated than the first. On the other hand, it may be argued that first is superior because the transition length nearest the head is shorter. As other considerations, such as demagnetization and the reproduction process, are of importance, it is not completely accurate to base an optimum on either of these criteria. However, experiments find an optimum recording current in the range of these

two values, *i.e.* between placing the $H_x = H_S$ cylinder so that it reaches beyond the back of the medium, but with the axis of the cylinder not beyond the middle of the medium.

From the exercises 6.2.1 and 6.2.2 it can be seen that the ratio of H_S to H_T is of great importance in determining the transition length. To illustrate this we shall consider the same geometry with a different H_T .

6.2.3 Example

The medium has $H_S = 580$ and $H_T = 350$ oe. Here we set the center of the $H_x = H_S$ cylinder axis at $Y = D + \Delta/2$, *i.e.* the middle of the medium, requiring that H_g be 6,547 oe (very similar to example 6.2.2) and find $h_S = 0.278$ and $h_T = 0.168$. Then at $Y = D$ we have $X_S = 2.647$ and $X_T = 3.434$, for a normalized transition length of 0.789, or 0.395g.

At $Y = D + \Delta$ we find $X_S = 2.655$ and $X_T = 5.980$ for a normalized transition length of 3.325 or 1.662g.

end of example

Comparing the results of examples 6.2.2 and 6.2.3, we find with a 20% reduction of H_S/H_T the bottom transition width is decreased 34% and the top by 29%.

It should be clear that the ideal recording medium would have H_T very nearly H_S .

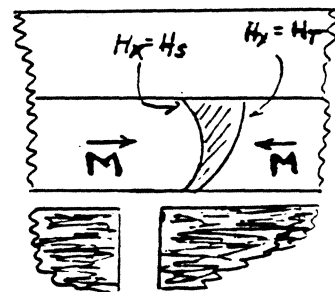


Figure 6.8. Example 6.2.3

6.3 DYNAMIC SWITCHING

In the previous section we considered transitions for the case where the head current reversed instantaneously. That could also be a reasonable approximation whenever the media moves only slightly during the time that it takes

for the head field to be reversed. In this section we shall examine the situation where the media moves appreciably during the time the head field is reversing.

Here we shall make use of the fact that the gap field changes from one saturation direction to the other in some continuous way, necessarily because of the energy stored in the magnetic field, which cannot change instantaneously. This can be thought of as the effect of the inductance of the head winding. Referring to figure 6.9, we will examine the field reversal process from the medium's point of view. The medium sees the head traveling past to the left, carrying its field with it. We begin our examination when the head just begins to switch its field from the negative x -direction to positive. The right-hand circle in figure 6.9(a) indicates the constant $H_x = -H_S$ cylinder at that time. As time goes on the head moves to the left while the field is first reduced from its original negative x -direction magnitude to zero, and then reversed until reaching the same magnitude but in the positive x -direction. The center point in figure 6.9(a) is the point where the field has reached zero. The left-hand circle indicates the fully reversed $H_x = +H_S$ cylinder.

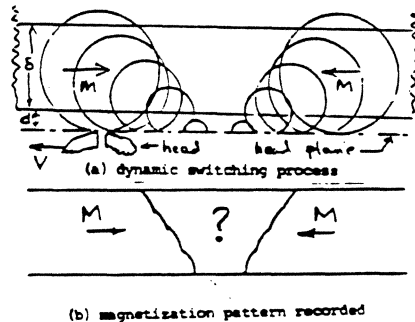


Figure 6.9

Figure 6.9(a) indicates the size of the $H_x = H_S$ cylinder for a discrete number of time samples, as to show more than that would further obscure the point at hand. Figure 6.9(b) shows the resulting recorded pattern of magnetization after the head passed out of the picture. Here again we see the pattern for the discrete samples of figure 6.9(a). Each portion of the boundaries of this discrete pattern is an arc of a circle corresponding to a segment of the $H_x = H_S$ cylinder for a particular gap field and time instant. However, with the field changing in a smooth manner, any locus of constant H_x is

also changing continuously. Thus there would be a continuum of circles in figure 6.9(a), resulting in a smooth curves for the boundaries of figure 6.9(b).

The medium is moving in the $+x$ -direction with constant velocity v . Viewed from the medium reference frame, the head is moving to the left with that velocity. When the switching begins, the region of the medium which is initially in the $H_x = -H_S$ cylinder is magnetized in the original direction, and the region immediately to the left has been partially switched to to that direction. As the field falls from the original direction and the head moves to the left, any point that has been saturated in the negative x -direction will remain in that state unless and until it experiences a reversed field of $+H_T$. In figure 6.9 we can see that the middle region of the medium, marked with a question mark, did not receive a field of H_S in either direction during this switching process. The previous recording history is unknown, so there is no way of determining the magnetization in that region. Data recording writes new data over the old record, so that the presence of uncontrolled regions such as the center region of figure 6.9 must be avoided. If they are allowed to occur, they will bring excessive noise to the play-back signal.

To eliminate noisy regions such as the center parts of figure 6.9, we must insure that the pattern generated by the growing $H_x = H_S$ cylinder overlaps the pattern of the previously shrunk $H_x = -H_S$ pattern. That is to say, in the process of flux reversal the new direction of magnetization must overlap and reverse the direction of part of the magnetization just previously recorded. (This is sometimes called over-write, but that term has another meaning which causes some confusion.) During this field reversal process it is necessary to assure that each point experience a field of magnitude H_S at least once. Then those points which originally had experienced a $-H_S$ field will be switched to some extent when they experience a maximum reversing field greater than $+H_T$, and will be fully switched if they experience a maximum reversing field of $+H_S$ or more. This will insure uniform transition patterns and subsequently uniform play-back signal waveforms. A particular point in this active region which has previously experienced a field of $-H_S$ and

then experiences a reversing field of at least H_T , its final remnant state (demagnetization being ignored) is determined by the remnant magnetization characteristic and the peak of reversed field experienced.

Here we examine a point which is experiencing its highest reversing field. It will next move to a lower field and be finished with the reversal process. Referring again to figure 6.9, as the $H_x = -H_S$ cylinder shrinks it leaves a boundary where the magnetization to its right is saturated in the negative x -direction, and the magnetization to its left is undefined. Similarly as the $H_x = +H_S$ cylinder grows it also establishes a boundary where the magnetization is saturated to its left and undefined to the right. We wish to remove the undefined region by overlapping the $H_x = +H_S$ boundary obliterating the $H_x = -H_S$ boundary, to establish a predictable transition region.

The boundary is a sequence of points which have experienced a field of H_S as their peak reversing field. In two dimensions the cylinders appear as circles, and each point on a boundary obtained from a unique circle. Two nearby boundary points have come from two circles, one occurring after the other. If we place these two points very near to each other, it should be clear that a line drawn between them will be tangent to both circles. Figure 6.10 shows the geometry of an expanding circle moving to the left (as seen from the medium). In a time dt , the circle moves a distance vdt to the left. The boundary will be the loci of all such points as the switching proceeds.

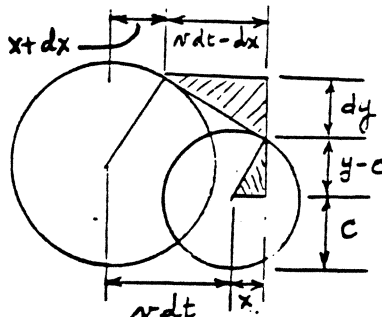


Figure 6.10. construction for the dynamics of switching

The parameters used here are the radius r , the center of the circle at $y = c$, and the velocity v . The results obtained can be applied to either the Karlqvist or the far-field head model. From

similar triangles we can obtain that:

$$\frac{y-c}{x} = vdt - dx \text{ or}$$

$$x dx + (y-c)dy = xvdt \quad (6.6)$$

The equation for the circle is: $x^2 + (y-c)^2 = r^2$, which we differentiate to obtain:

$$2x dx + 2(y-c)(dy - dc) = 2r dr \text{ or}$$

$$x dx + (y-c)dy = ydc + r dr - cd \quad (6.7)$$

As the left hand members of equations 6.6 and 6.7 are the same, the right-hand sides are equal, so that:

$$x = \frac{1}{v} \left(y \frac{dc}{dt} + r \frac{dr}{dt} - c \frac{dc}{dt} \right) \quad (6.8)$$

As r and c are not independent, it is worth while to discuss two cases. The simpler is the far-field approximation, so it is examined first.

6.4 DYNAMIC SWITCHING FOR THE FAR FIELD CASE

For this case $r = c$, so that equation 6.8 reduces to:

$$x = \frac{y}{v} \frac{dr}{dt} \quad (6.9)$$

Substituting equation 6.8 into the equation for the circle we obtain that:

$$y = \frac{2r}{1 + \left(\frac{1}{v} \frac{dr}{dt} \right)^2} \quad (6.10)$$

from which we can substitute into equation 6.8 to obtain x . However, this development has been in the head coordinate system, and we need the result in the media coordinates which are moving to the right of the head system with a velocity v . We use the subscript m to designate the medium coordinates, so that $y_m = y$ and $x_m = x - vt$, so that:

$$y_m = \frac{2r}{1 + \left(\frac{1}{v} \frac{dr}{dt} \right)^2} \quad (6.11)$$

$$x_m = -vt + \frac{1}{v} \frac{dR}{dt} \quad (6.12)$$

From the end of chapter 5 we have that $r = gH_g/2\pi H_x$, and for this example we look first at $H_x = H_S$, and secondly at $H_x = H_T$ to

locate the transition region. For $H_x = H_S$, we find that

$$r_S = \frac{gH_g}{2\pi H_S} \quad (6.13)$$

$$\frac{dr_S}{dt} = \frac{g}{2\pi H_S} \frac{dH_g}{dt} \quad (6.14)$$

and for $H_x = H_T$, we find

$$r_T = \frac{gH_g}{2\pi H_T} \quad (6.15)$$

$$\frac{dr_T}{dt} = \frac{g}{2\pi H_T} \frac{dH_g}{dt} \quad (6.16)$$

6.4.1 Far Field Example

The simplest case to consider is where the field changes from its negative saturation value to positive saturation as a linear function of time: $H_g = mvt$ so the loci where a particular value of H_x has been the largest reversing field experienced has radius varying as $r = kvtR$ and $dr/dt = kvR$, where $k = mg/2\pi H_x$. From equation 6.12 we obtain that:

$$vt = \frac{y_m}{2} \left(k + \frac{1}{k} \right) \quad (6.17)$$

substituting equation 6.17 into 6.12:

$$x_m = y_m \left(k - \frac{1}{k} \right) \quad (6.18)$$

which is the equation of a straight line in the medium coordinate system.

We recall that the recording of predictable transition regions requires that the boundary locus of the reversed $H_x = +H_S$ pattern must overlap the previous $H_x = -H_S$ pattern. we examine equation 6.18 for the situation of $H_x = -H_S$ and find that, with $k_S = mg/2\pi H_S$:

$$x_m = -\frac{y_m}{2} \left(k_S - \frac{1}{k_S} \right) \quad (6.19)$$

and that for $H_x = +H_S$:

$$x_m = \frac{y_m}{2} \left(k_S - \frac{1}{k_S} \right) \quad (6.20)$$

so that as the field begins reversal with the $H_x = -H_S$ pattern a straight line of a particular slope, and as the field reverses the $H_x = +H_S$ pattern is also a straight line with the

same magnitude of slope, but of the opposite sign.

The situation where the $H_x = +H_S$ line just touches the $H_x = -H_S$ line is where the two slopes are infinite, i.e. when x_m is zero: $k_S = 1/k_S$, or $k_S^2 = 1$. Which requires that the field must change with a minimum slope of $m = 2\pi H_S/g$ or that

$$H_g = \frac{2\pi H_S v}{g} t \quad (6.21)$$

It is of interest to note that, while this development has taken place with un-normalized units, the normalization will yield that the minimum slope is given by:

$$\frac{H_g}{H_S} = \pi V t \quad (6.22)$$

where $V = 2v/g$

While a linearly changing field is only an approximation of practical field switching, this example serves as a guideline to the maximum field switching time for predictable transition regions.

6.4.2 Dynamic Switching for the Karlqvist Case

Using the results obtained in equation 6.8:

$$x = \frac{1}{v} \left(y \frac{dc}{dt} + r \frac{dr}{dt} - c \frac{dc}{dt} \right)$$

For the Karlqvist fringe field approximation we found the radius and center of the constant H_x on page 26 in normalized form: $R = \csc h_x$ and $C = \cot h_x$. Then equation 6.8 can be written in normalized form:

$$X = \frac{1}{V} \left(Y \frac{dC}{dt} + R \frac{dR}{dt} - C \frac{dC}{dt} \right) \quad (6.23)$$

It can be readily shown that:

$$R \frac{dR}{dt} = C \frac{dC}{dt}$$

Thus we can use equation 6.9 in this case as well. Here it is rewritten in normalized form:

$$\text{KARLQUIST: } X = \frac{Y}{V} \frac{dR}{dt} \quad (6.24)$$

Substituting equation 6.24 into the normalized circle equation and using the facts that Y

must always be positive and $R^2 = C^2 + 1$, we obtain the following expression for Y :

$$\text{KARLQUIST: } Y = \frac{C + \sqrt{R^2 + \left(\frac{1}{V} \frac{dR}{dt}\right)^2}}{1 + \left(\frac{1}{V} \frac{dR}{dt}\right)^2} \quad (6.25)$$

With $C = \csc \pi h_x$ and $R = \cot \pi h_x$ we have

$$\frac{dR}{dt} = \frac{\pi H_x}{H_g^2} \left(\csc^2 \frac{\pi H_x}{H_g} \right) \frac{dH_g}{dt}$$

There is presently no known closed form solution for the boundaries of the transition region for the Karlqvist fringe field case, but numerical solutions have been performed which show similar behavior to the far-field case, except of course in the region very near the gap. In that region the linearly changing field indicates that a greater slope is needed as the field passes through zero on its reversal, and less at the beginning and end of the switching interval. Fortunately these are the usual and natural properties of magnetic fields in the transient state.

Chapter 7

DEMAGNETIZATION IN RECORDING

In chapter 6 we saw how a transition region would occur in the absence of demagnetizing fields. The transition is a region of changing magnetization, which means that there it is a region of finite divergence of magnetization, which in turn means it is a region containing magnetic pole density. The pole density is the source of demagnetizing field intensity H_d , which acts to spread out the transition as much as possible.

The hysteresis characteristic of the recording medium limits this spreading of the transition region. For any particular level of magnetization, a minimum field is necessary to change the magnetization. If the combination of the applied field (magnetizing field) and the demagnetizing field exceed that minimum, the magnetization will change.

While the demagnetizing field can change the transition region after recording has been completed (the applied field no longer present), the principle effect occurs during the writing process where the total field is the sum of the applied field and the demagnetizing field. This will result in a transition region shape that is altered from those predicted in chapter 6. The transition it will be wider (in the x -direction), and may even be shifted depending on the distances to adjacent transitions.

Analysis of the recording process, including the demagnetizing effects of the recorded magnetization, is an iterative problem requiring a large high speed computer to carry out *self consistent* calculations. This is often done using finite element mathematics, a process beyond the scope of this course. Insight is obtained only by using simplifications to allow the nature or characteristics of the process to be expressed in analytic form. The general rule is to sim-

plify until only the barest necessary complexity remains and so obtain approximate relations. After that, complicate as much as is practical to gain understanding of secondary effects. In this way we can also obtain results useful in system design and analysis.

7.1 THE UNIFORM TRANSITION AND ITS DEMAGNETIZING FIELD

The demagnetizing field is due to the divergence of the magnetization, and can be calculated using equation 3.10, which is written here in rectangular coordinates:

$$H_d = \iiint \frac{\rho_m \mathbf{l}_{s,f}}{4\pi r_{s,f}^2} dx_s dy_s dz_s, \quad (7.1)$$

where

$$\mathbf{r}_{s,f} = \mathbf{l}_x(x_f - x_s) + \mathbf{l}_y(y_f - y_s) + \mathbf{l}_z(z_f - z_s)$$

$$r_{s,f} = \sqrt{(x_f - x_s)^2 + (y_f - y_s)^2 + (z_f - z_s)^2}$$

and $\mathbf{l}_{s,f} = \mathbf{r}_{s,f}/r_{s,f}$.

The demagnetizing field is to be calculated at the middle of the recorded track, where $z_f = 0$. This requires integration of equation 7.1 over all space, which is effectively the space of the recorded track in the magnetic medium. To begin with, the magnetization distribution is not specified, but some specific examples are considered after some preliminaries where the y and z effects are taken into account by some simplifying assumptions.

In the spirit of simplicity the following assumptions are made:

1. assume the magnetization is entirely in the x -direction, so the divergence of \mathbf{M} is simplified to dM_x/dx_s .
2. As the x -components are of primary interest, the y and z components in $\mathbf{r}_{s,j}$ are eliminated.
3. Assume there is no variations in the z direction and that the track width is much greater than any other dimensions, and calculations are made at $z_j = 0$. The integral limits on the z go to infinity, and with no z dependency of the magnetization, the integration¹ over z is carried out to obtain equation 7.2, which is the x component of the demagnetizing field H_{xd} :

$$H_{xd} = \iint \frac{\rho_m}{2\pi} \frac{(x_j - x_s) dx_s dy_s}{(x_j - x_s)^2 + (y_j - y_s)^2} \quad (7.2)$$

7.2 ABRUPT OR STEP TRANSITION

The simplest imaginable transition is one which has the magnetization reversing in a zero-width transition, which can be written as a step function for the case where the magnetization is negative to the left of the transition (*i.e.* the magnetization points in the minus x -direction) and reverses to positive to the right of the transition, as indicated in figure 7.1.

$$M_x(x_s) = M_{rs} [2u(x_s - x_o) - 1]$$

where x_o is the location of the transition plane in source space.

The pole density is found from the negative of dM_x/dx_s :

$$\rho(x_s) = -2M_{rs} \delta(x_s - x_o) \quad (7.3)$$

where $\delta(x_s - x_o)$ is a mathematical delta function and is distinguished from the medium thickness, δ , by the fact that it always has an argument [*i.e.* $(x_s - x_o)$] whereas the medium thickness always occurs as an entity [*i.e.* as δ without decorations].

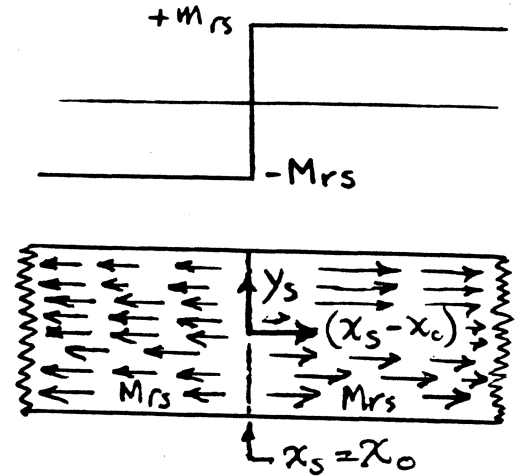


Figure 7.1. Abrupt magnetization transition

Then substituting the result of equation 7.3 into equation 7.2 we obtain:

$$H_{xd} = - \iint \frac{2M_{rs}}{2\pi} \frac{\delta(x_j - x_s) dx_s dy_s}{(x_j - x_s)^2 + (y_j - y_s)^2}$$

The infinite integral over x_s is a special type of integral involving the delta function, which is called the sifting function. This is evaluated by replacing the integration by the value of the integrand at the point where the delta function is non-zero², *i.e.* where $x_s = x_o$. This results in:

$$H_{xd} = - \int_{-\delta/2}^{\delta/2} \frac{M_{rs}}{\pi} \frac{(x_j - x_o) dy_s}{(x_j - x_o)^2 + (y_j - y_s)^2} \quad (7.4)$$

The final step is integration over the thickness of the medium, using equation 2 of appendix A to obtain:

$$H_{xd} = - \frac{M_{rs}}{\pi} \left\{ \tan^{-1} \frac{\delta/2 - y_j}{x_j - x_o} + \tan^{-1} \frac{\delta/2 + y_j}{x_j - x_o} \right\} \quad (7.5)$$

This function has its maximum magnitude at $x_j - x_o = 0$:

$$|H_{dx(max)}| = M_{rs} \text{ for } (-\delta/2 < y_j < \delta/2)$$

At the points $y_j = \pm\delta/2$ the maximum magnitude is only half this value.

¹see appendix A, equation 1

²see appendix A section A.2

The maximum of the function is found at the $y_f = 0$ to be:

$$H_{xd} = -2 \frac{M_{rs}}{\pi} \tan^{-1} \frac{\delta/2}{x_f - x_o} \quad (7.6)$$

At $y = \pm\delta/2$ which includes the near side of the recording media, where the transitions seen in chapter 6 were the narrowest. The demagnetizing field at $y_f =$ is:

$$H_{xd} = -\frac{M_{rs}}{\pi} \tan^{-1} \frac{\delta}{x_f - x_o} \quad (7.7)$$

which has its maximum magnitude at $x_f = x_o$:

$$|H_{dx(max)}| = \frac{M_{rs}}{2} \text{ at } (x_o, -\delta/2)$$

These demagnetizing field functions at the middle and surface of the media are shown in figure 7.2 with two intermediate heights. It can be seen that the surface of the media is a degenerate case, as all other points on the transition have a demagnetizing field equal to M_{rs} .

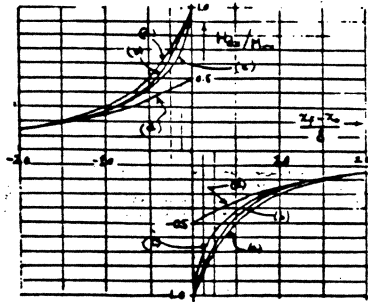


Figure 7.2. Abrupt transition demagnetization
(a) at $y = 0$, (b) at $y = \pm\delta/4$.
(c) at $y = \pm 3\delta/8$, (d) at $y = \pm\delta/2$.

These results provide upper limits to the demagnetizing fields that are possible in a magnetic medium with a transition involving the reversal of magnetization of the level M_{rs} . The maximum possible demagnetizing field is equal to the maximum magnetization.

7.3 ARCTANGENT TRANSITION

The most useful magnetization distribution for an isolated transition has been an arctangent. This model of the transition is written as:

$$M_x(x_s) = \frac{2M_{rs}}{\pi} \tan^{-1} \frac{x_s - x_o}{a} \quad (7.8)$$

where a is called the *a-parameter*. Here x_o is again the arbitrary center for the transition in the medium coordinate system, the point where $M_x = 0$. At the point where $x_s - x_o = \pm a$ we find that $M_x = \pm 0.5M_{rs}$. At extreme distances from x_o the magnetization approaches the appropriate saturated value. The derivative is taken:

$$\frac{dM_x}{dx_s} = \frac{2M_{rs}}{\pi} \frac{a}{(x_s - x_o)^2 + a^2} = -\rho_m \quad (7.9)$$

which can now be inserted into equation 7.2 for integration, where the range of x_s becomes $-\infty \leq x_s \leq \infty$.

$$H_{xd} = -\frac{1}{2\pi} \int_{-\delta/2}^{\delta/2} \int_{-\infty}^{\infty} \frac{2M_{rs}}{\pi} \frac{a}{(x_s - x_o)^2 + a^2} \times \frac{(x_f - x_s) dx_s}{(x_f - x_s)^2 + (y_f - y_s)^2} dy_s \quad (7.10)$$

This is the form of the convolution integral, as explained in Appendix A³. The method of evaluating such an infinite integral uses the Fourier Transform as outlined in Appendix A, section 4. Following that procedure, the following steps are made:

$$f_1(x_s) = \frac{a}{(x_s - x_o)^2 + a^2}$$

Using identity I4 in table A1, and transform T6 in table A2, it is found that:

$$F_1(k) = \pi e^{-jx_o k} e^{ak}$$

Next is identified:

$$f_2(u - z) \Rightarrow f_2(x_f - x_s)$$

$$f_2(x_f - x_s) = \frac{(x_f - x_s)}{(x_f - x_s)^2 + (y_f - y_s)^2}$$

so that

$$f_2(x_s) = \frac{(x_s)}{x_s^2 + (y_f - y_s)^2}$$

Because of the nature of the integral (equation 7.10) it is not clear whether the term $y_f - y_s$ or the term $y_s - y_f$ should be used. Keeping in mind that the y -variable of integration is actually y_s , equation 7.10 can be written in either

³Appendix A, section 3

of two ways, which ordinarily makes no difference. However in this procedure it does make a difference:

$$H_{xd} = -\frac{M_{rs}}{\pi^2} \int_{-\delta/2}^{\delta/2} \int_{-\infty}^{\infty} \frac{a}{(x_s - x_o)^2 + a^2} \times \frac{(x_j - x_s) dx_s d(y_s - y_j)}{(x_j - x_s)^2 + (y_s - y_j)^2} \quad (7.11)$$

So that using transform T5 from table A2, we obtain

$$F_2(k) = -j\pi e^{(y_s - y_j)k} \text{ for } d(y_s - y_j) \quad (7.12)$$

Alternatively equation 7.10 can be written as:

$$H_{xd} = \frac{M_{rs}}{\pi^2} \int_{-\delta/2}^{\delta/2} \int_{-\infty}^{\infty} \frac{a}{(x_s - x_o)^2 + a^2} \times \frac{(x_j - x_s) dx_s d(y_j - y_s)}{(x_j - x_s)^2 + (y_j - y_s)^2} \quad (7.13)$$

which results in a transform using T5 again:

$$F_2(k) = j\pi e^{(y_j - y_s)k} \text{ for } d(y_j - y_s) \quad (7.14)$$

When the entire evaluation procedure is completed, it is necessary that the resulting demagnetizing field be symmetrical with respect to y . As neither equation 7.12 nor equation 7.14 result in such symmetry, it is necessary to use both to achieve that symmetry. Therefore the transformed relation becomes:

$$H_{xd} = \frac{M_{rs}}{2\pi^2} \int_{-\delta/2}^{\delta/2} \pi e^{-jx_o k} e^{ak} \times \left[\begin{array}{l} j\pi e^{(y_j - y_s)k} d(y_j - y_s) \\ -j\pi e^{(y_s - y_j)k} d(y_j - y_s) \end{array} \right] \quad (7.15)$$

This can be manipulated to obtain:

$$H_{xd} = \frac{M_{rs}}{2\pi} \int_{-\delta/2}^{\delta/2} e^{-jx_o k} \times \left[\begin{array}{l} j\pi e^{(a+y_j-y_s)k} d(y_j - y_s) \\ -j\pi e^{(a+y_s-y_j)k} d(y_j - y_s) \end{array} \right] \quad (7.16)$$

Using transform T5 of Table A2 and identity I4 from table A1, the inverse transform is obtained yielding:

$$H_{xd} = \frac{M_{rs}}{2\pi} \int_{-\delta/2}^{\delta/2} \frac{(x_j - x_o) d(a + y_j - y_s)}{(x_j - x_o)^2 + (a + y_j - y_s)^2} - \frac{M_{rs}}{2\pi} \int_{-\delta/2}^{\delta/2} \frac{(x_j - x_o) d(a + y_s - y_j)}{(x_j - x_o)^2 + (a + y_s - y_j)^2}$$

This is evaluated using equation 2 of appendix A:

$$H_{xd} = \frac{M_{rs}}{2\pi} \left[\begin{array}{l} \tan^{-1} \frac{a+y_j-\delta/2}{x_j-x_o} \\ -\tan^{-1} \frac{a+y_j+\delta/2}{x_j-x_o} \end{array} \right] - \frac{M_{rs}}{2\pi} \left[\begin{array}{l} \tan^{-1} \frac{a+\delta/2-y_j}{x_j-x_o} \\ -\tan^{-1} \frac{a-\delta/2-y_j}{x_j-x_o} \end{array} \right] \quad (7.17)$$

which can be regrouped to obtain:

$$H_{xd} = \frac{M_{rs}}{2\pi} \left[\begin{array}{l} \tan^{-1} \frac{a-\delta/2+y_j}{x_j-x_o} \\ +\tan^{-1} \frac{a-\delta/2-y_j}{x_j-x_o} \end{array} \right] - \frac{M_{rs}}{2\pi} \left[\begin{array}{l} \tan^{-1} \frac{a+\delta/2-y_j}{x_j-x_o} \\ +\tan^{-1} \frac{a+\delta/2+y_j}{x_j-x_o} \end{array} \right] \quad (7.18)$$

Equation 7.18 has the proper symmetry with respect to both x_j and y_j , and in the case of $a \rightarrow 0$ the demagnetizing field also approaches that of the abrupt transition (equation 7.5). The highest field occurs on the plane $y = 0$, where the field is:

$$H_{xd} = \frac{M_{rs}}{\pi} \left[\tan^{-1} \frac{a-\delta/2}{x_j-x_o} - \tan^{-1} \frac{a+\delta/2}{x_j-x_o} \right]$$

This function has its maxima at $(x_j - x_o)^2 = (a - \delta/2)(a + \delta/2)$, which has the value of:

$$H_{dx(max)} = \frac{M_{rs}}{\pi} \left[\begin{array}{l} \tan^{-1} \sqrt{\frac{a-\delta/2}{a+\delta/2}} \\ -\tan^{-1} \sqrt{\frac{a+\delta/2}{a-\delta/2}} \end{array} \right]$$

But this can be simplified using the identity:

$$\tan(A \pm B) = \frac{\tan A \pm \tan B}{1 \mp \tan A \tan B}$$

to obtain

$$H_{dx(max)} = \frac{M_{rs}}{\pi} \tan^{-1} \frac{\delta}{\sqrt{4a^2 - \delta^2}} \quad (7.19)$$

The demagnetization field is shown in figure 7.3 for an a parameter of 1.5 δ .

For a situation where the demagnetizing field alone set the transition, the a parameter would be adjusted so that the maximum field would be approximately equal to the coercive force H_c . This is calculated from equation 7.19:

$$a \approx \frac{\delta/2}{\sin(\pi H_c / M_{rs})} \quad (7.20)$$

This can be seen to be a lower limit on the a parameter, although effects due to the high permeability of the head could reduce this length somewhat.

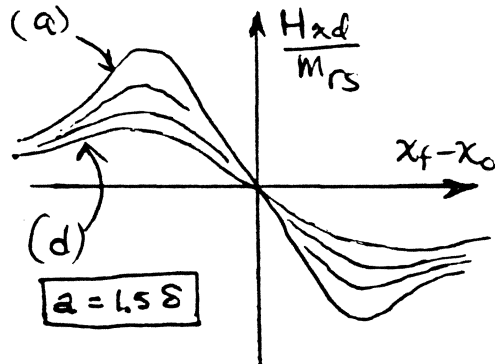


Figure 7.3. Arctangent transition demagnetization
 (a) at $y_f = 0$, (b) at $y = \pm\delta/4$.
 (c) at $y = \pm 3\delta/8$, (d) at $y = \pm\delta/4$.

7.4 MATERIAL CHARACTERISTICS

The arctangent transition is described by the *a parameter*, for which a minimum value can be estimated by setting the maximum demagnetizing field $H_{xd(max)}$ to H_c as in eq. 7.20. This may be optimistic, because the actual magnetization distribution is determined by applied head field $H_{\pm a}$ and the magnetization characteristic in conjunction with the demagnetizing field.

An accurate solution to the magnetization distribution is an iterative process which requires numerical techniques and a computer. However, useful information on the *a parameter* can be obtained by some simplifications and intuitive application of known relations.

For convenience, it is assumed that the *arctangent* transition being written is in the opposite direction from what was discussed previously, so the transition is being recorded with a positive gap field (in the positive x -direction) over a previously saturated region of at negative remanance. Then the resulting transition will be the negative of equation 7.8, with a value of $\mp M_{rs}/2$ at $(x_s - x_0) = \pm a$.

When the recorded transition is established and the transition has moved far away from the head, the middle of the transition will have no demagnetizing field. However, during the reversal process, the field at that point will have exceeded the coercive force (H_c) and at that time the magnetization and field will be at the point indicated in figure 7.4 as point I on the dynamic hysteresis curve. When the applied field decreases because of movement away from the head, the applied field relaxes to zero along

the path indicated by χ_r .

In the dynamic process of writing the transition, the slope of the M - H characteristic, referring to figure 7.4, is:

$$\frac{\partial M}{\partial H} = \frac{M_{rs}}{H_c - H_1} = \frac{M_{rs}}{H_c(1 - S^*)}$$

so that $S^* = H_1/H_c \leq 1$.

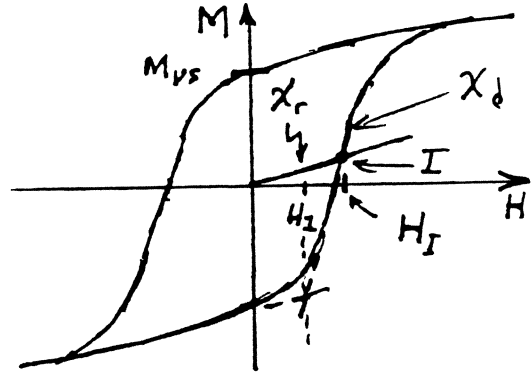


Figure 7.4 Magnetization characteristic

Again referring to figure 7.4, the middle of the transition ($x_s = x_0$) will occur at the point where the net field has been H_1 during the recording process. The relaxation path in figure 7.4 is from point I to the origin, which has the equation: $M = \chi_r H$

The dynamic characteristic in this condition can be described as a straight line:

$$M = \chi_d(H - H_c) \text{ where:}$$

$$\chi_d = \frac{\partial M}{\partial H} = \frac{M_{rs}}{H_c(1 - S^*)} \quad (7.21)$$

From these relations we obtain that:

$$H_1 = \frac{\chi_d H_c}{\chi_d - \chi_r} \quad (7.22)$$

7.5 FIELD GRADIENT

During the recording process the characteristic of the material and the total magnetic field interacts so that a particular magnetization distribution results in a demagnetizing field which alters the total magnetic field, resulting in a changed distribution. This can be treated by a series of calculations of the magnetization distribution, which then gives rise to a demagnetizing field which adds to the field, so that a new distribution must be calculated. This is an iterative process which is called *the self-consistent*

calculation method, and consumes much computer time. Here we simplify by recognizing that the derivative of the magnetization must follow the following condition:

$$\frac{dM_x}{dx} = \frac{\partial M_x}{\partial H} \frac{dH}{dx} \quad (7.23)$$

$$\text{but } H = H_{x_a} + H_{x_d}$$

where H_{x_d} is the negative of equation 7.18, and:

$$H_{x_a} = \frac{H_g}{\pi} \left\{ \begin{array}{l} \tan^{-1} \frac{x+g/2}{y} \\ -\tan^{-1} \frac{x-g/2}{y} \end{array} \right\} \quad (7.24)$$

From the negative of equation 7.9 at the middle of the transition ($x_s = x_o$) gives a slope of:

$$\frac{dM}{dx} = -\frac{2M_{rs}}{\pi a}$$

At $x_s = x_o$ the demagnetizing field is zero, so the applied field must be set to H_I . However it is desirable to make the slope of the applied field, dH_{x_d}/dx as large as possible so as to obtain as small an a parameter as possible.

Equation 7.24 is first differentiated, and then the second derivative obtained and set to zero to find the point of the maximum slope, which is the maximum of dH_{x_a}/dx :

$$\frac{dH_{x_a}}{dx} = \frac{H_g}{\pi} \left[\frac{\frac{y}{(x+g/2)^2+y^2}}{-\frac{y}{(x-g/2)^2+y^2}} \right] \quad (7.25)$$

Taking the second derivative with respect to x and setting the result to zero, the following expression is obtained:

$$\frac{x+g/2}{(x+g/2)^2+y^2} = \frac{x-g/2}{(x-g/2)^2+y^2}$$

solving this for the position of maximum applied field slope:

$$\left[\frac{x}{g/2} \right]^2 = \frac{2}{3} \sqrt{\left[\frac{y}{g/2} \right]^4 + \left[\frac{y}{g/2} \right]^2} + 1 - \frac{1}{3} \sqrt{\left[\frac{y}{g/2} \right]^2} - 1 \quad (7.26)$$

which will become the center of the transition, x_o

At the center of the transition the field is also to be set to H_I as given in equation 7.22. From this the value of the gap field can be determined

depending on which point in the media is selected to receive H_I .

The calculations are to be done at the middle of the recording media, which in the media coordinates is at $y_f = y_s = 0$, and in the head coordinates is at $y = d + \delta/2$. Then taking the negative of equation 7.9 at $x_s = x_o$ for the left side of equation 7.23, using equation 7.21 to obtain $\partial M/\partial H$, dH_{x_a}/dx from equation 7.25 at $x = x_o$ and $y = d + \delta/2$, and finally obtaining $\partial H_{x_d}/\partial x$ from the derivative of equation 7.18 with $y_s = 0$ and $x_s = x_o$ to obtain the following relation:

$$-\frac{2M_{rs}}{\pi a} = \frac{M_{rs}}{H_c(1-S^*)} \times \left\{ \begin{array}{l} \frac{H_g}{\pi} \left[\frac{\frac{d+\delta/2}{(x_o+g/2)^2+(d+\delta/2)^2}}{-\frac{d+\delta/2}{(x_o-g/2)^2+(d+\delta/2)^2}} \right] \\ -\frac{M_{rs}}{\pi} \left[\frac{1}{a+\delta/2} - \frac{1}{a-\delta/2} \right] \end{array} \right\} \quad (7.27)$$

where H_g is adjusted according to equations 7.22 and 7.24 to give:

$$H_I = \frac{\chi_d H_c}{\chi_d - \chi_r} = \frac{H_g}{\pi} \left\{ \begin{array}{l} \tan^{-1} \frac{x_o+g/2}{d+\delta/2} \\ -\tan^{-1} \frac{x_o-g/2}{d+\delta/2} \end{array} \right\} \quad (7.28)$$

Equation 7.27 can be evaluated for any specific geometry. The reader is encouraged to evaluate several examples.

A INTEGRALS AND TRANSFORMS

A.1 STANDARD INTEGRALS

The integrals here can be found in most tables of integrals. They can be verified by differentiation.

$$\int \frac{du}{(u^2 + v^2)^{3/2}} = \frac{1}{v^2} \frac{u}{u^2 + v^2} \quad (1)$$

$$\int \frac{du}{u^2 + v^2} = \frac{1}{v} \arctan \frac{u}{v} \quad (2)$$

$$\int \frac{u du}{u^2 + v^2} = \frac{1}{2} \ln(u^2 + v^2) \quad (3)$$

A.2 THE SIFTING FUNCTION

The delta function, $\delta(u - b)$, is also known as the *sifting function* because it has the property of selecting out the value of a function at the point $u = b$ whenever it multiplies the function and they are integrated over any interval that includes the point b .

$$\int_a^b \delta(u - b) f(u) du = f(b) \text{ for } a \leq b \leq c \quad (4)$$

A.3 THE CONVOLUTION INTEGRAL

The form of the convolution integral is as follows:

$$f_1 * f_2 = \int_{-\infty}^{\infty} f_1(x) f_2(a - x) dx \quad (5)$$

This form of integral occurs in the analysis of demagnetizing fields and of play-back voltages in magnetic recording systems. Two interesting forms of the convolution integral are:

$$f(x_f) = \int_{-\infty}^{\infty} f_1(x_s) f_2(x_f - x_s) dx_s$$

$$f(vt) = \int_{-\infty}^{\infty} f_1(x) f_2(vt - x) dx$$

In principle these are readily evaluated using Fourier Transforms. However, these particular functions (and others of possible interest in magnetic recording) are not usually found in most tables of transformations. For that reason, a special table of such transforms are developed in the following.

A.4 FOURIER TRANSFORMS

Transform pairs for various functions are derived using the fundamental definition of the transform and inverse transform. The term $1/2\pi$ must appear either on the transform or inverse transform, or be split between them as $1/\sqrt{2\pi}$, and the placement is not consistent from one table of transforms to the next, so it is important to note its placement when using a particular table of transform pairs. Here we place it in the definition of the transform of $F(k)$ to $f(x)$, which is usually called the *Inverse Transform*:

$$f(x) = \frac{1}{2\pi} \int_{k=-\infty}^{\infty} F(k) e^{jxk} dk \quad (6)$$

where $f(x)$ is a function of x , $F(k)$ is its Fourier transformation into k -space, and $j = \sqrt{-1}$, so that we are dealing with complex variables in k -space.

The transformation from x -space to k -space is usually called the *Transform*:

$$F(k) = \int_{x=-\infty}^{\infty} f(x) e^{-jxk} dx \quad (7)$$

The derivations and proofs of the properties of Fourier transform pairs is beyond the scope of this work. Here some of those properties are provided as a set of fundamental transformation identities, which are given in the table A1.

TABLE A1 TRANSFORM IDENTITIES

I1	$f(z) \Leftrightarrow F(k)$
I2	$Af(z) \Leftrightarrow AF(k)$
I3	$f_1(z) + f_2(z) \Leftrightarrow F_1(k) + F_2(k)$
I4	$f(z + c) \Leftrightarrow e^{jck} F(k)$
I5	$\frac{df(z)}{dz} \Leftrightarrow jk F(k)$
I6	$\int_{-\infty}^z f(u) du \Leftrightarrow \frac{1}{jk} F(k)$
I7	$\int_{-\infty}^{\infty} f_1(x) f_2(u - x) dx \Leftrightarrow F_1(k) F_2(k)$

The procedure in evaluating convolution integrals is to begin with identity I7 which has the convolution integral on the left side. We must next find the transformations of $f_1(x)$ and

$f_2(x)$: $F_1(k)$ and $F_2(k)$ respectively using table A2, the table of transform pairs.

The form $f_2(u-x)$ is essential, i.e. the function must be written so that the integration variable occurs in a subtractive relation to the independent variable of the function being evaluated. Thus f_2 must be a function of $vt-x$ or x_f-x_s . However, when finding F_2 , it is to be found from $f_2(x)$ or $f_2(x_s)$.

The next step is to manipulate the product $F_1(k)F_2(k)$ to the form of some function of k on the right hand side of table A2, which then allows the evaluation of the function of vt or x_f .

TABLE A2 FOURIER TRANSFORM PAIRS

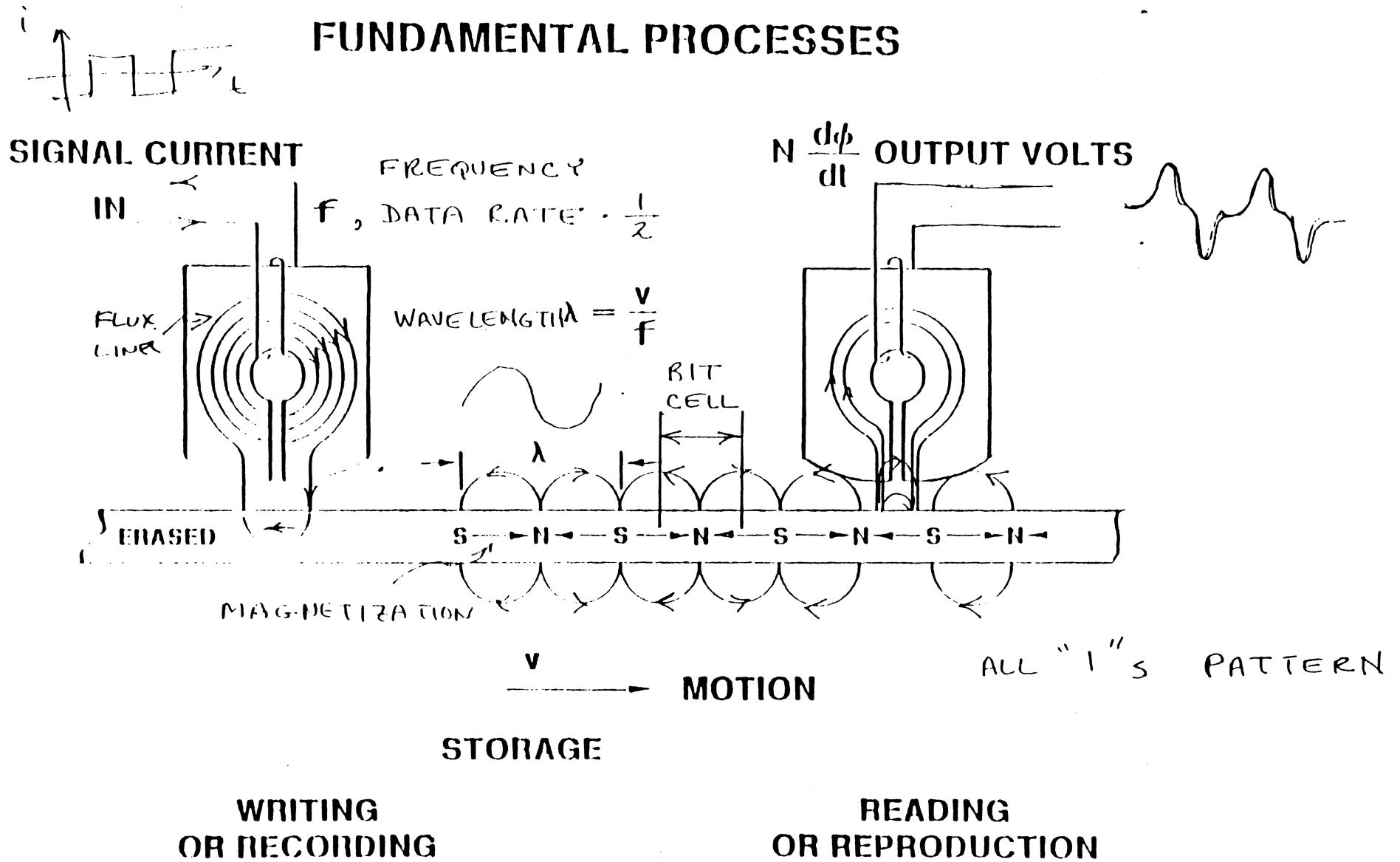
	$f(z)$	\Leftrightarrow	$F(k)$
T1.	$\delta(z+a)$	\Leftrightarrow	e^{jak}
T2.	$u(z+a)$	\Leftrightarrow	$-\frac{1}{jk}e^{jak}$
T3.	$\sin \beta z$	\Leftrightarrow	$j\pi[\delta(k+\beta) - \delta(k-\beta)]$
T4.	$\cos \beta z$	\Leftrightarrow	$\pi[\delta(k+\beta) + \delta(k-\beta)]$
T5.	$\frac{z}{z^2+u^2}$	\Leftrightarrow	$j\pi e^{juk}$
T6.	$\frac{u}{z^2+u^2}$	\Leftrightarrow	πe^{juk}
T7.	$\arctan \frac{z}{u}$	\Leftrightarrow	$\frac{\pi}{jk}e^{juk}$
T8.	$\frac{1}{2} \ln(z^2+u^2)$	\Leftrightarrow	$\frac{\pi}{k}e^{juk}$
T9.	e^{jaz}	\Leftrightarrow	$2\pi\delta(k-a)$

READ PROCESS

=====**IIST**=====

J. Monson
6/00

FUNDAMENTAL PROCESSES

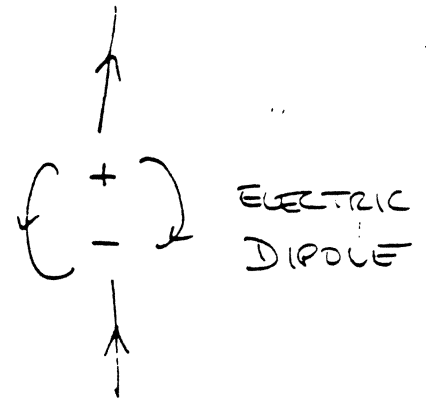


MAGNETIC CHARGE

SMALL MAGNET :



ANALOGOUS TO



ELECTRIC DIPOLE

N POLES \rightarrow + MAGNETIC CHARGE

S POLES \rightarrow - MAGNETIC CHARGE

DIPOLE MOMENT $\begin{matrix} +q \\ \uparrow l \\ -q \end{matrix} \Rightarrow \uparrow \vec{p}$ VECTOR

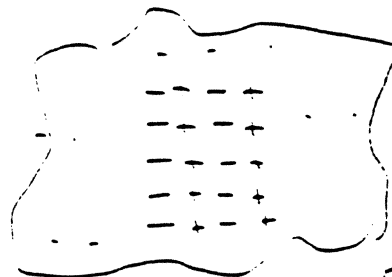
\vec{p} has magnitude $q \cdot l$

direction from $-q$ to $+q$

MAGNETIZATION :

$\vec{M} \equiv$ dipole moment per unit volume

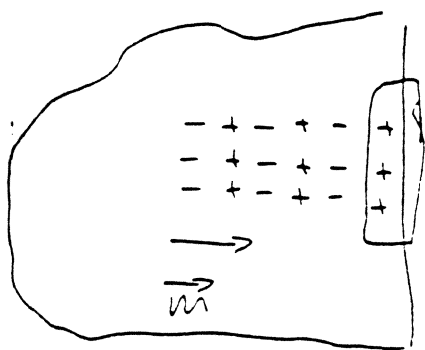
UNIFORMLY MAGNETIZED REGION



DIPOLLES
LINED UP
CHARGES
CANCEL AT
EACH LAYER

• NO NET CHARGE ANYWHERE

SURFACE BETWEEN MAGNETIZED REGION AND AIR



MAGNETIC MATERIAL

This charge layer not cancelled, giving sheet of + charge at surface

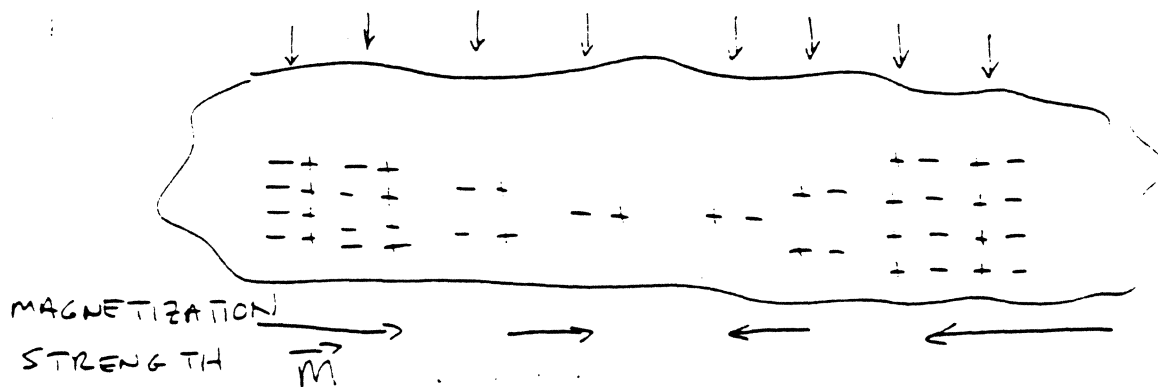
$$\vec{M} = 0$$

AIR

SURFACE CHARGE DENSITY $\sigma = \vec{M} \cdot \vec{n}$

$\vec{n} \equiv$ unit vector, out ward normal

NON-UNIFORMLY MAGNETIZED REGION



MAGNETIZATION STRENGTH \vec{M}

NET CHARGE

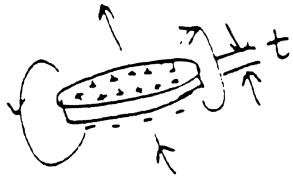


VOLUME CHARGE DENSITY

$$\rho = -\frac{\partial M_x}{\partial x} = -\vec{\nabla} \cdot \vec{M}$$

MAGNETIC SHELL

is a thin shell with uniform dipole moment per unit area.



Area ΔA
thickness t

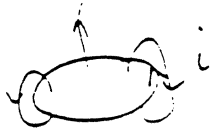
shell strength is

\vec{M}_A , dipole moment / unit area

in terms of volume moment, \vec{M}

$$\vec{M}_A = \vec{M} \cdot t$$

EQUIVALENCE TO WIRE LOOP CARRYING CURRENT

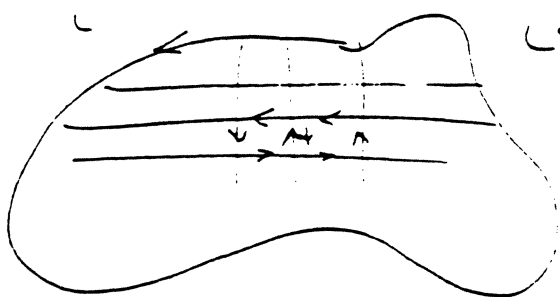


Area ΔA

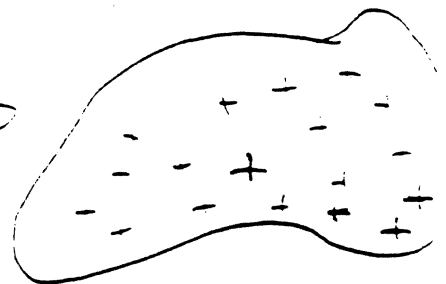
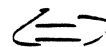
Also produces dipole field
Equivalent moment is $i \Delta A$

- $M_A \equiv i$

VALID FOR ANY SHAPE WIRE LOOP



LOOP



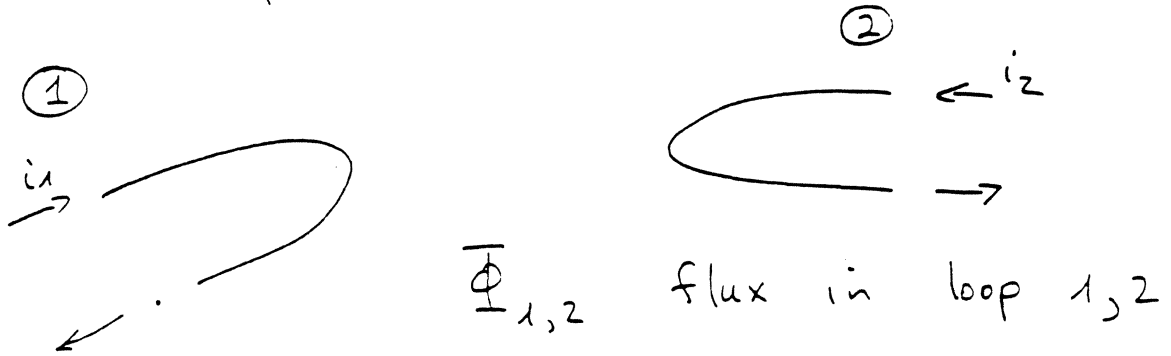
SHELL

SUBDIVIDE AND SUPERPOSE

ALL INTERIOR CURRENTS CANCEL

RECIPROCALITY THEOREM

Two loops of wire



Flux $\propto i$; $\Phi = L i$

$L \equiv$ inductance

$$\Phi_1 = L_{11} i_1 + L_{12} i_2$$

$$\Phi_2 = L_{21} i_1 + L_{22} i_2$$

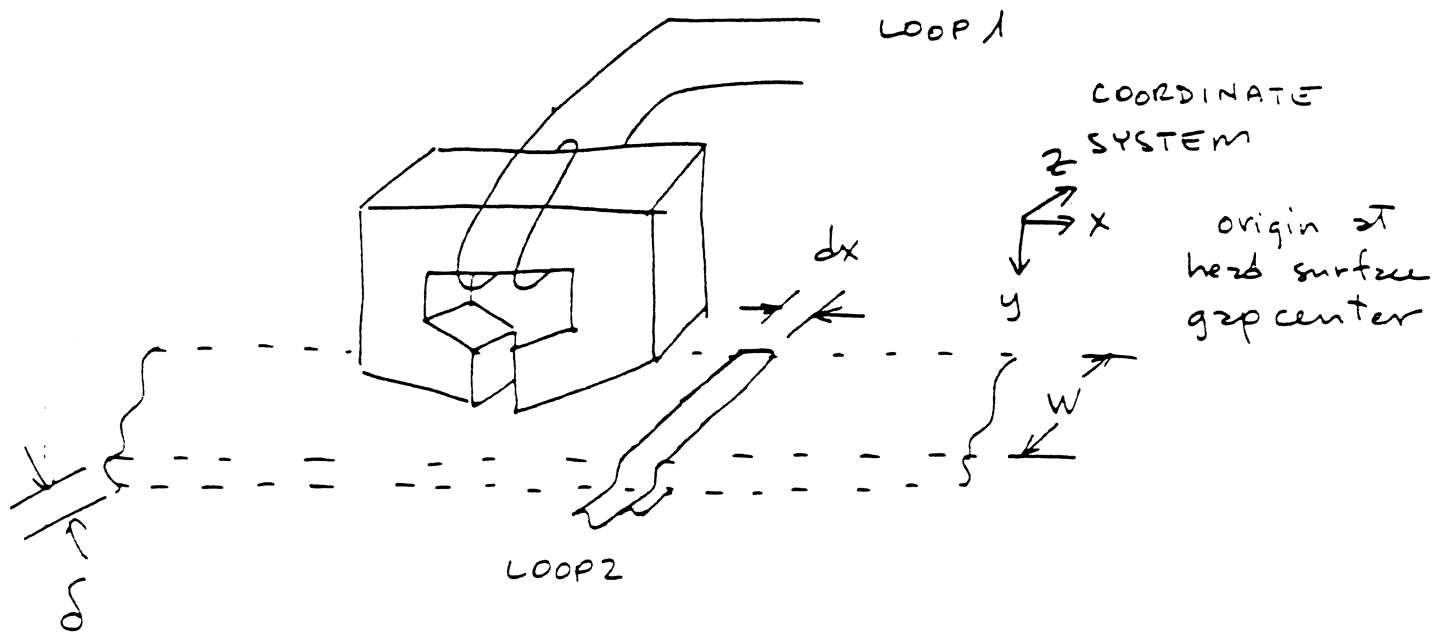
By reciprocity, $L_{21} = L_{12}$

(linear, passive network w/o isolators, gyrators, etc)

In recording, feedback process is linear, signal levels are small.

Also, it's passive - no external power sources are involved.

READ BACK



- By knowing head field, can find $\bar{\Phi}_2$ caused by current in LOOP 1
 - By reciprocity, this is also $\bar{\Phi}_1$ caused by current in LOOP 2
 - BUT, current in LOOP 2 is equivalent to magnetization in the medium by magnetic shell equivalence.
- ∴ can find $\bar{\Phi}_1$ in head coil caused by magnetization pattern on medium
this is the readback process.

Assume δ is small. Let loop 2 be located at coordinates x, y .

Flux in loop 2 caused by 1 Ampere-turn in the head coil is:

$$d\bar{\Phi}_{2x} = \underbrace{\delta W}_{\text{Loop 2 Area}} \underbrace{\mu_0 \frac{H_x(x,y)}{1A}}_{\text{flux density}} \cdot \underbrace{1A}_{\text{Head sensitivity function}}$$

L_{21}

By reciprocity, head flux ($\bar{\Phi}_1$) is

$$d\bar{\Phi}_{1x} = \underbrace{\delta W \mu_0 \frac{H_x(x,y)}{1A}}_{L_{12} = L_{21}} \cdot \text{Loop 2 current}$$

Loop 2 current is magnetic shell equivalent of $M_x(x,y) dx$

$$d\bar{\Phi}_{1x} = \delta W \mu_0 \frac{H_x(x,y)}{1A} \cdot M_x(x,y) dx$$

- Know H_x from Karlqvist or other calculations
- Assume standard magnetization patterns, e.g. sine waves or transition shapes
- Solve for Φ_H by integrating w/respect to x .
- Output voltage is

$$e = -N \frac{d\Phi_H}{dt}$$

$$= -Nv \frac{d\Phi_H}{dx}$$

Same result holds for \perp magnetization.

$$d\bar{\Phi}_y = \mu_0 H_y(x, y) dx W$$

\perp loop 2 magnetization equivalent to current
is $M_y(x, y) \delta$

then
$$d\bar{\Phi}_{Hy} = \mu_0 M_y(x, y) H_y(x, y) dx W \delta$$

Because of linearity,

\perp + \parallel contributions add

each magnetization chunk at a given
 x adds

• Total Flux

$$\bar{\Phi}_H = \mu_0 W \delta \int_{-\infty}^{\infty} \vec{M}(x, y) \cdot \vec{H}(x, y) dx$$

\vec{H} is for unit MMF across head
sometimes called head sensitivity function

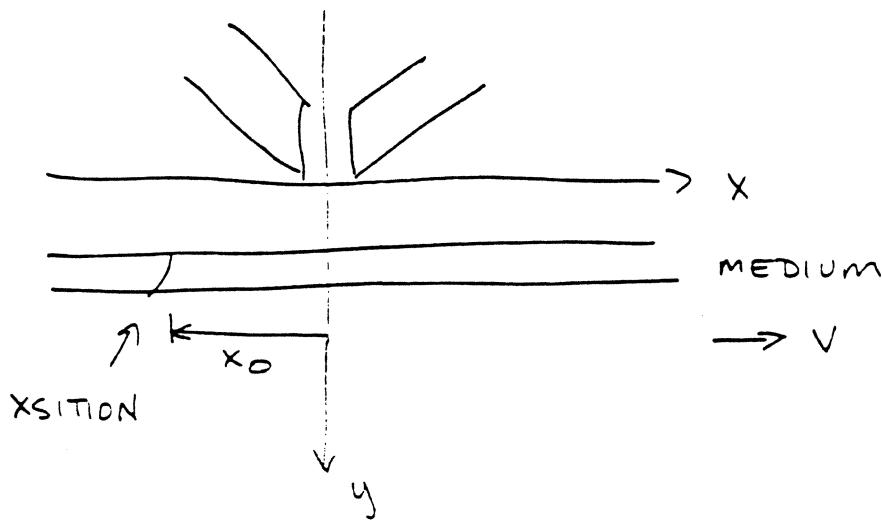
To account for moving medium :-

• Let x, y coord. system be fixed to head

• Give medium its own x_0, y_0 coordinate system
with origin fixed to it.

• $y_0 = 0$ (no y displacement)

$x_0 = vt$ (medium moving w/
constant velocity)



$$\vec{M}(x, y) \Rightarrow \vec{M}(x - x_0, y)$$

$$\Phi_H(x_0) = \mu_0 W S \int_{-\infty}^{\infty} \vec{M}(x - x_0, y) \cdot \vec{H}(x, y) dx$$

Can get voltage from Faraday's Law

$$e(x_0) = e(vt) = N \frac{d\Phi_H}{dt} \quad (\text{ignoring signs})$$

$$\frac{d\Phi_H}{dt} = \frac{d\Phi_H(x_0)}{dx_0} \cdot \frac{dx_0}{dt} \quad ; \quad \frac{dx_0}{dt} = v$$

$$e = Nv \frac{d\Phi_H(x_0)}{dx_0}$$

$$e = \eta \mu_0 W S N v \int_{-\infty}^{\infty} \frac{\partial}{\partial x_0} \vec{M}(x - x_0, y) \cdot \vec{H}(x, y) dx$$

$\eta \equiv$ head efficiency

EXAMPLE !

- Longitudinal medium, $M_y \Rightarrow 0$
- Narrow gap Karlqvist head

$$H_x = \frac{H_g g}{\pi} \frac{y}{x^2 + y^2}$$

- $y = d$
- $M_x(x)$ is a step function
 $\Rightarrow \frac{2M_x}{2x}$ is impulse function

or δ -function with "area" $2 M_R$

- Thin medium, $\delta/d \leq 0.25$

Then,

$$e(x_0) = \eta \mu_0 W \delta N V (2 M_R) \int_{-\infty}^{\infty} \delta(x-x_0) \frac{d}{\pi} \frac{dx}{x^2 + d^2}$$

Using "sifting" process of δ -function,

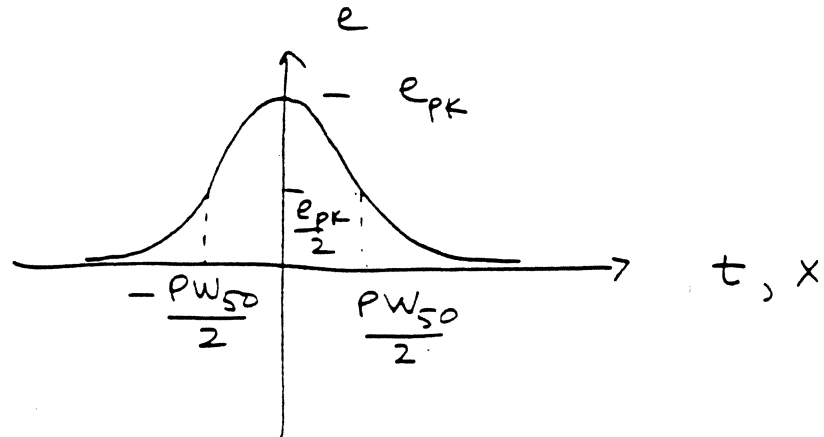
$$e(x_0) = \frac{\mu_0 W \delta N V}{\pi} 2 M_R \frac{d}{x_0^2 + d^2}$$

OUTPUT PULSE LOOKS LIKE HEAD FIELD!

Note! Don't confuse $\delta \equiv$ medium thickness

with δ -function \equiv impulse function

PULSE SHAPES



Lorentzian:

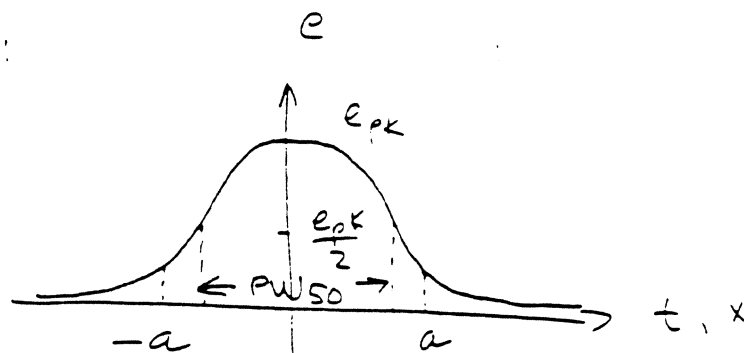
$$e = \frac{e_{pk}}{1 + \left(\frac{2x}{PW_{50}}\right)^2}$$

For thin medium \rightarrow spacing d

$$PW_{50} = 2d$$

Get Lorentzian pulse e from step transition
 + narrow gap head.

Gaussian:

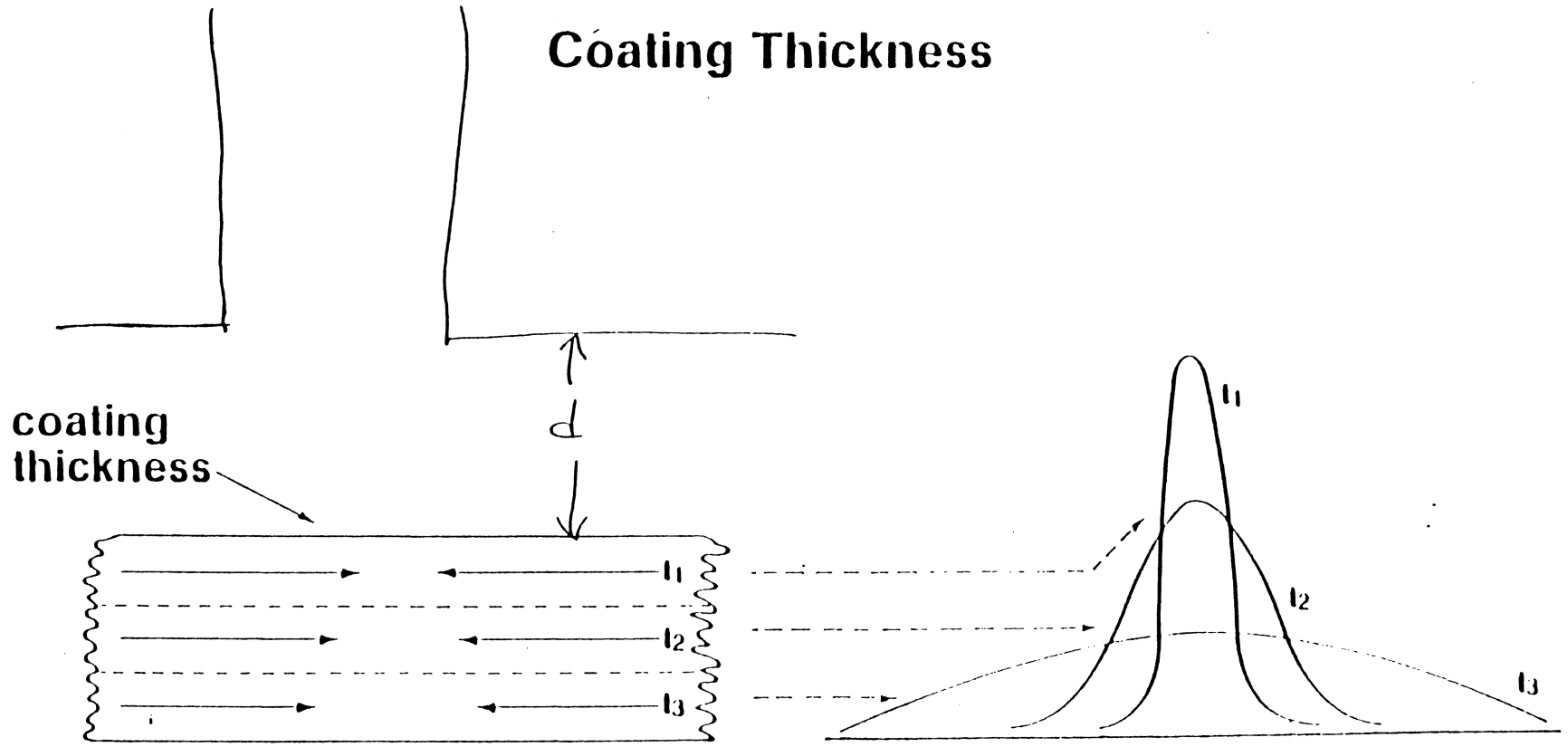


$$e = e_{pk} e^{-\left(\frac{x}{a}\right)^2}$$

$$\frac{PW_{50}}{2} = 0.83 a$$

EFFECT OF Coating Thickness

27



(after Hoagland 1 3)

BY SUPERPOSITION, CAN
INTEGRATE THROUGH THICKNESS!

$$e(x_0) = \mu_0 W N V \int_d^{d+\delta} \int_{-\infty}^{\infty} \frac{2}{2x_0} \vec{M}(x-x_0, y) \cdot \vec{H}(x, y) dx$$

Integral can be messy to evaluate.

Good approximation to take

$$PW_{50} = 2 [d(d+\delta)]^{1/2} \quad (\text{step position})$$

GAP LENGTH:

When gap length is included
in the Karlqvist head expressions,
integrals are again messy to
evaluate, but

$$PW_{50} = 2 \left[\left(\frac{a}{2}\right)^2 + d(d+\delta) \right]^{1/2} \quad (\text{step position})$$

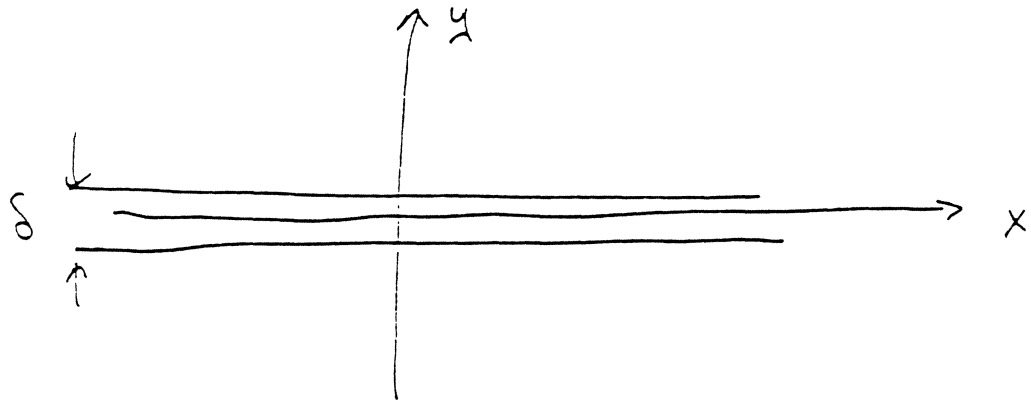
TRANSITION SPREADING:

$$M_x(x) = \frac{2Mr}{\pi} \tan^{-1} \left(\frac{x}{a} \right)$$

$$PW_{50} = 2 \left[\left(\frac{a}{2}\right)^2 + (a+d)(a-d+\delta) \right]^{1/2}$$

DIRECT CALCULATION OF FLUX FROM MAGNETIZATION (R.L. WALLACE)

Field from sinusoidally magnetized medium -

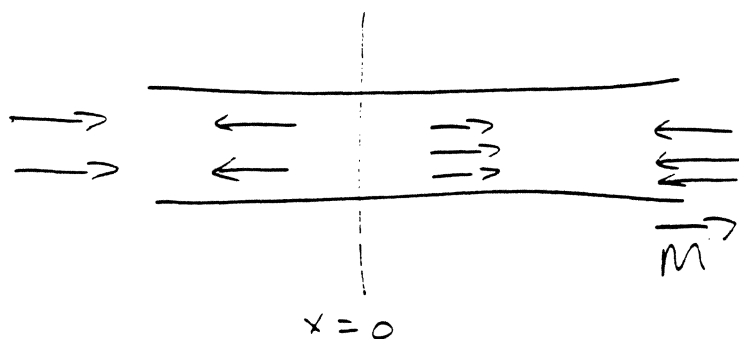


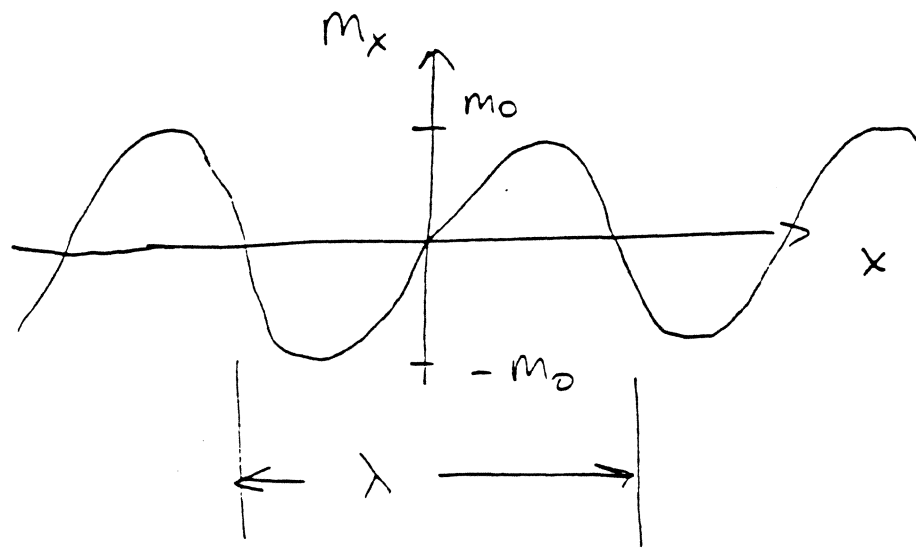
M has x component only (longitudinal record)

$$M_x = M_0 \sin kx$$

k is wave number $\equiv \frac{2\pi}{\lambda}$

λ is wave length

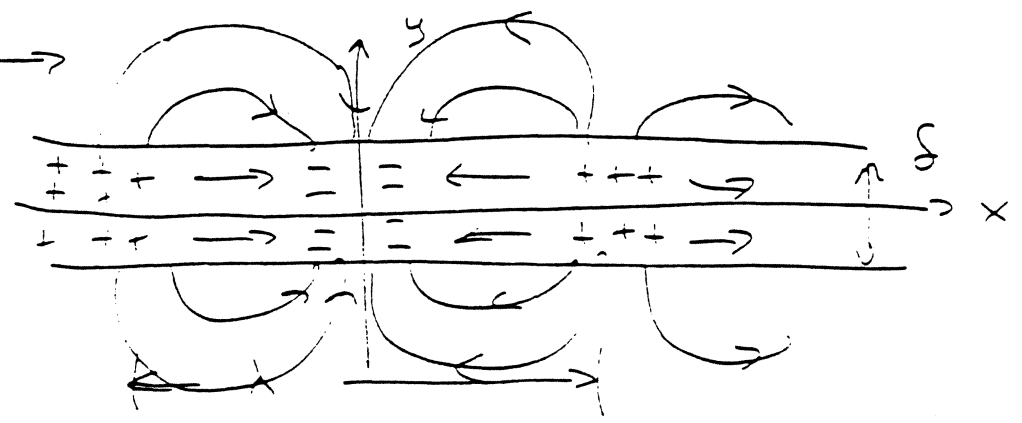




We have volume charge density

$$\rho_m = -\frac{2M_x}{\lambda} = -KM_0 \cos Kx$$

Magnetic field



Charge produces field as shown.

For $y > \delta/2$

$$H_x = -4\pi M_0 \sin Kx \sinh \frac{K\delta}{2} e^{-Ky}$$

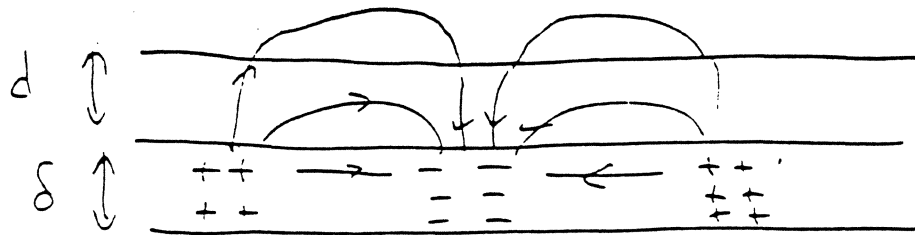
(cgs units) $H_y = -4\pi M_0 \cos Kx \sinh \frac{K\delta}{2} e^{-Ky}$

FIELDS DECAY EXPONENTIALLY WITH Ky !

Presence of head perturbs field distribution.

flux increased here

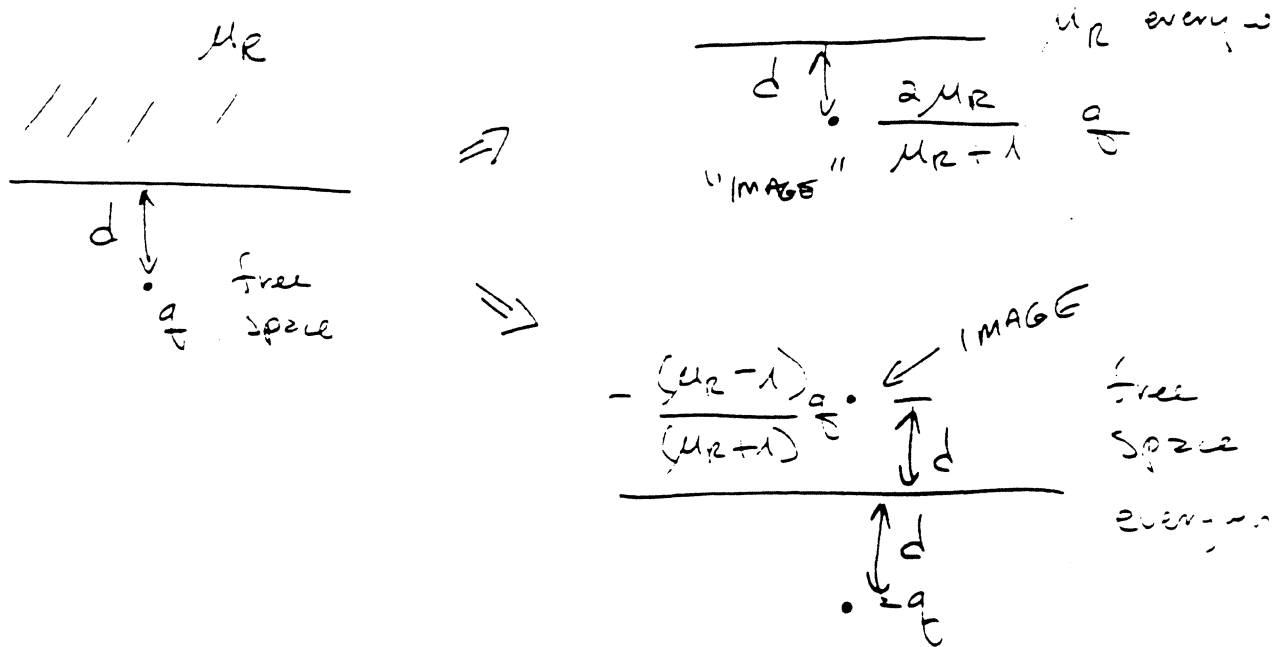
high permeability head



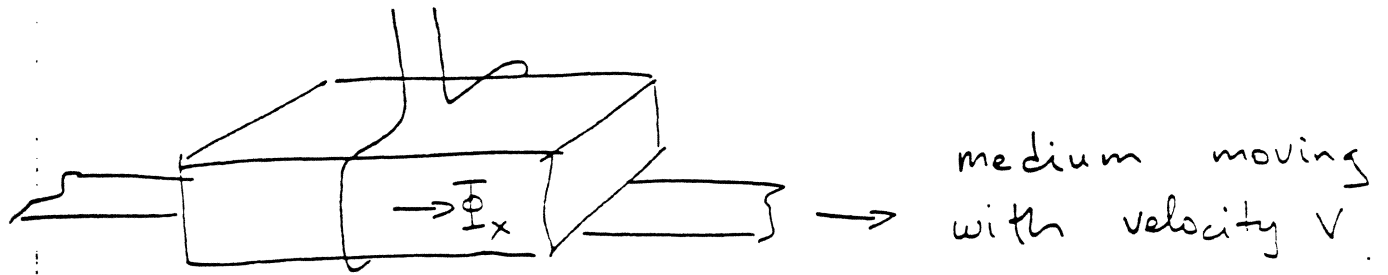
flux reduced here

high permeability region analogous to high dielectric region in electrostatics.

METHOD OF IMAGES



READ BACK VOLTAGE



coil senses $\frac{d\Phi_x}{dt}$ produced by changing H_x in head

- Solve for H_x in head using images

$$M_0 \Rightarrow \frac{2\mu_r}{\mu_r + 1} M_0$$

- Integrate H_x from $y = d + \frac{\delta}{2}$ to ∞ to get flux $\vec{\Phi}_x$

- Let $x = vt$; $k = \frac{\omega}{v}$

- Calculate $e = N \frac{d\Phi}{dt}$ (ignoring sign)

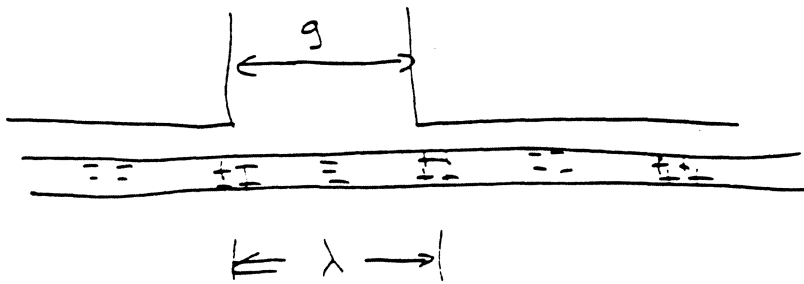
$$\vec{\Phi}_x(t) = - \frac{2\mu_r}{\mu_r + 1} 2\pi \delta W M_0 \sin \omega t \cdot \frac{(1 - e^{-k\delta})}{k\delta} e^{-kd}$$

$$e(t) = N\omega \frac{2\mu R}{\mu R + 1} \cdot 2\pi \delta W M_0 \cos \omega t.$$

$$\frac{(1 - e^{-k\delta})}{k\delta} e^{-kd}$$

GAP EFFECT:

Adding a gap introduces an aperture or windowing effect.



null at $\lambda = g$

gap effect is
$$\frac{\sin\left(\frac{k g_{\text{eff}}}{2}\right)}{\frac{k g_{\text{eff}}}{2}}$$

$g_{\text{eff}} \equiv$ effective gap length

$\sim 1.14 g$ from "exact" field

calculation as opposed to Kralqvist

WAVELENGTH ANALYSIS SEPARATES VARIOUS LOSS EFFECTS:

SPACING LOSS e^{-Kd} $(-55 \text{ dB}/\lambda)$

THICKNESS LOSS $\frac{(1 - e^{-K\delta})}{K\delta}$

Long wavelengths:

T.L. \Rightarrow

All of λ , medium thickness is contributing to rezobzok

short wavelengths:

T.L. \Rightarrow

$\frac{1}{K\delta}$, less + less of thickness contribute

GAP LOSS $\frac{\sin\left(\frac{K g_{\text{eff}}}{2}\right)}{\frac{K g_{\text{eff}}}{2}}$

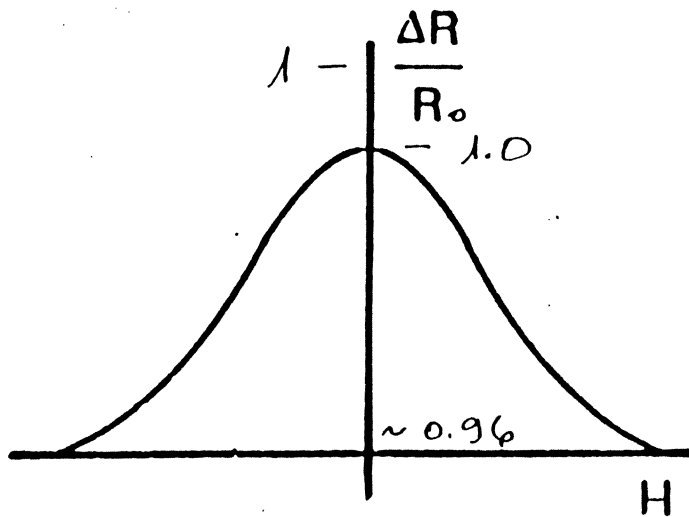
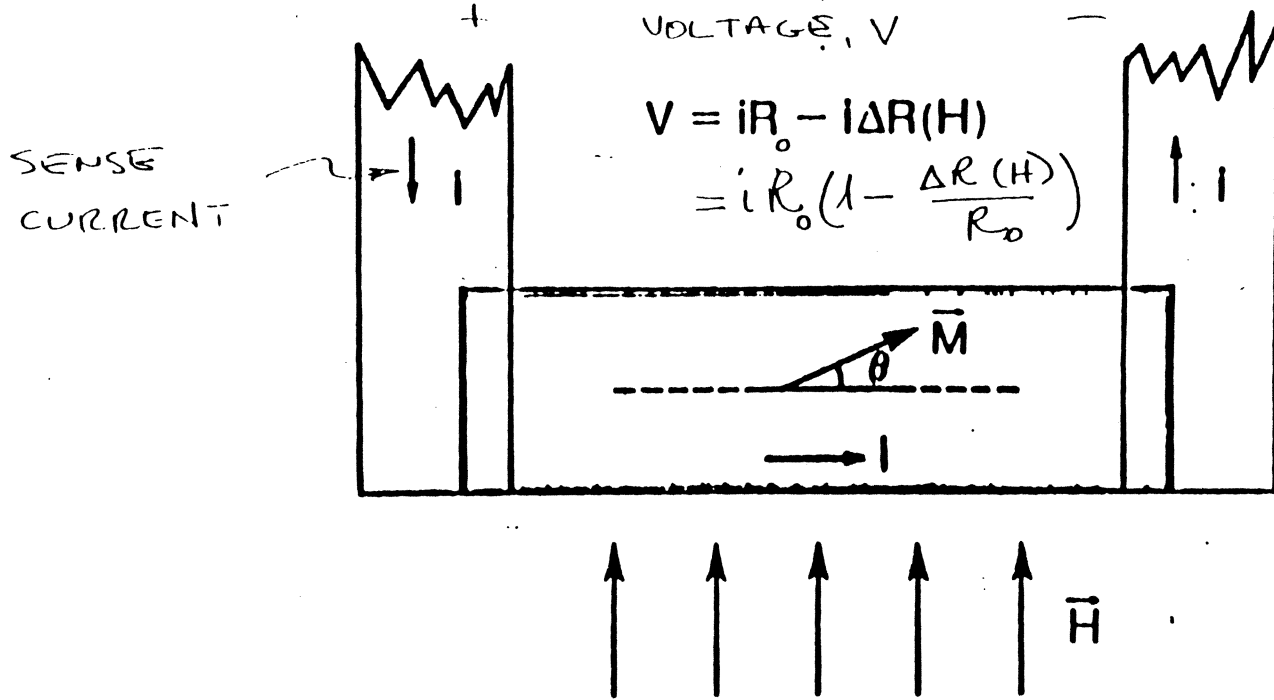
TRANSITION SPREADING LOSS

Wallace analysis assumes same amplitude for any wavelength.

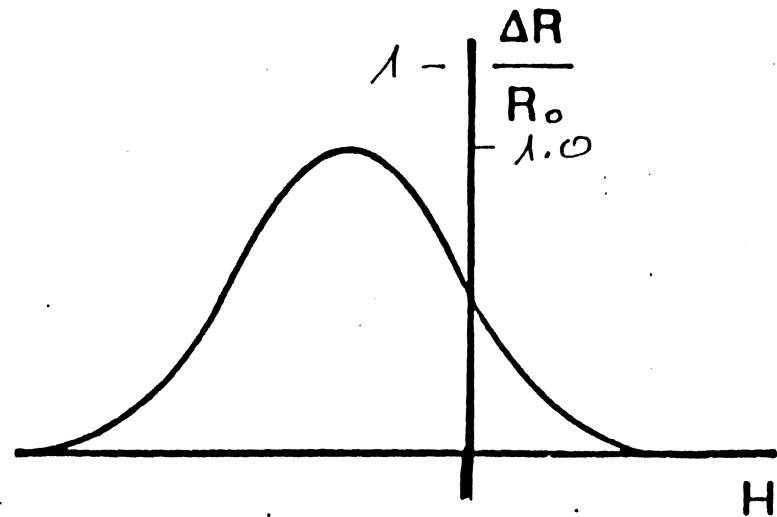
For real xsition, short wavelengths are attenuated.

e^{-Ka} $a \equiv$ transition spreading parameter

Magnetoresistive (MR) Sensor



unbiased response



biased response

- FILM IS ANISOTROPIC

- R MAX WHEN $M \parallel \hat{i}$

- M rotates in presence of H

- $R \approx R_0 - \Delta R \sin^2 \Theta$

$$\sin^2 \Theta \approx \left(\frac{H}{H_0} \right)^2 \quad ; \quad H_0 = \frac{2K}{M_s}$$

K is anisotropy constant

- $R \approx R_0 \left(1 - \frac{\Delta R}{R_0} \left(\frac{H}{H_0} \right)^2 \right)$

- Linearize by applying bias

$$H = H_{\text{bias}} + H_{\text{signal}}$$

- Then, $R \approx R_0 \left(1 - \frac{\Delta R}{R_0} \left(\frac{H_{\text{bias}}}{H_0} \right)^2 - \frac{\Delta R}{R_0} \cdot 2 \frac{H_{\text{sig}} H_{\text{bi}}}{H_0^2} \right)$

- $\frac{\Delta R}{R_0} \approx 2-4\%$, but still very sensitive.

- velocity independent output

- many geometries, biasing methods

MR HEADS

Advantages

- Large signal output
- Velocity independent signal amplitude
- Good linear resolution with shielded MR sensor

Basic Issues

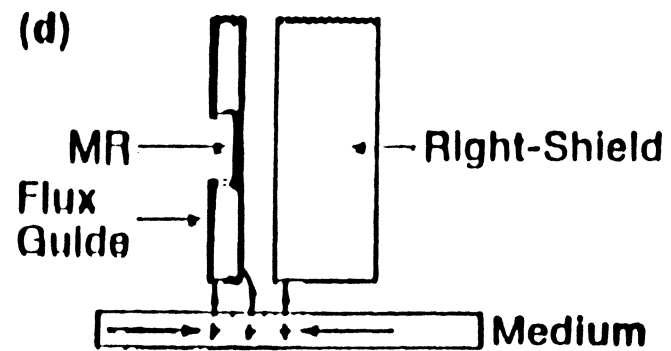
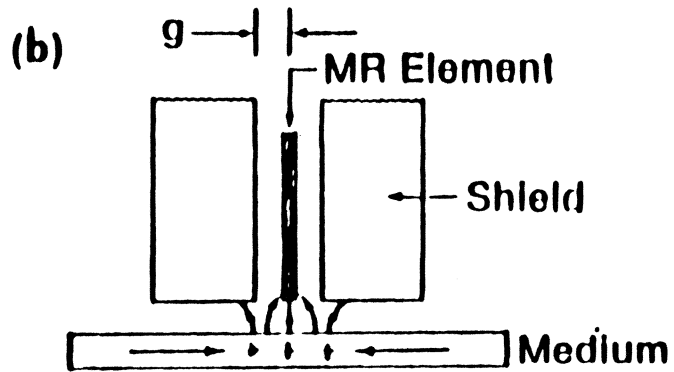
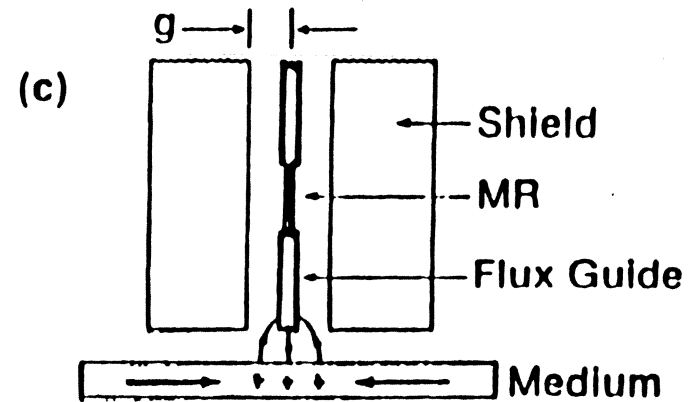
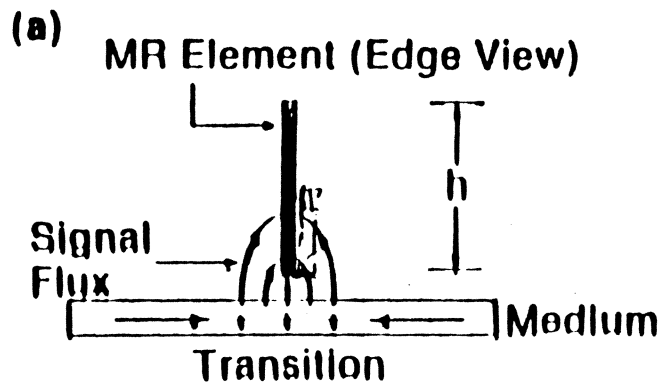
- Suppression of ^hEarkausen noise
- Linearization of MR response

Operating Magnetic Environments

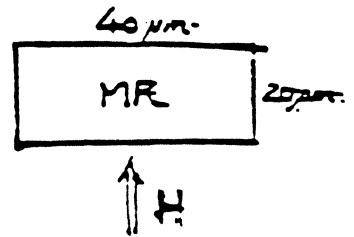
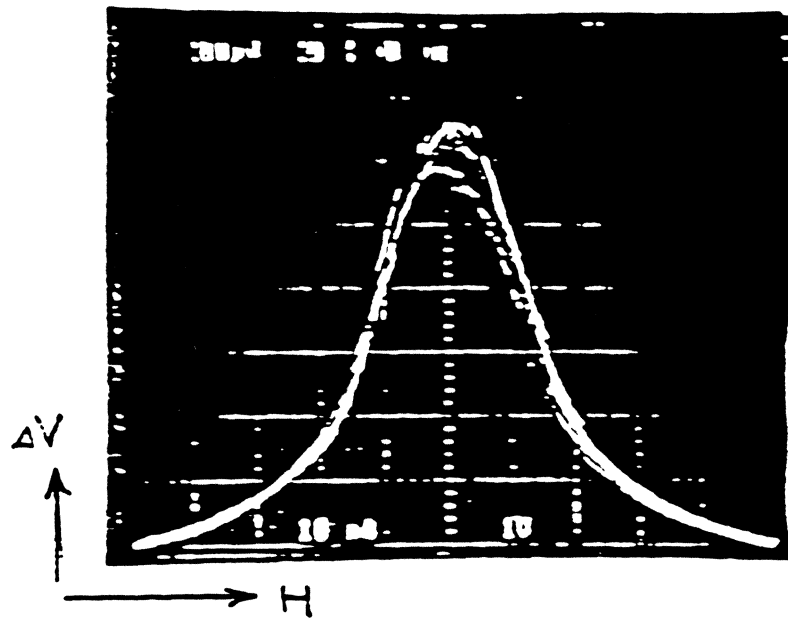
Imposed by requirements of:

- Linear Resolution:
 - Presence of soft-magnetic shields
- Write Capabilities:
 - Presence of inductive write element in the vicinity
- Trackwidth Resolution:
 - Certain sizes and geometries of MR sensor might be more preferable

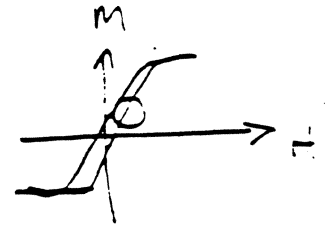
from Ching Tsang, IBM



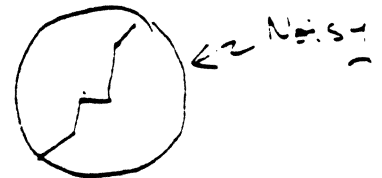
SHIELDING METHODS FOR HIGH
LINEAR RESOLUTION



The BARKHAUSEN NOISE Problem



• Domain activities in small MR elements



• Relation between domain activities & noise in MR response

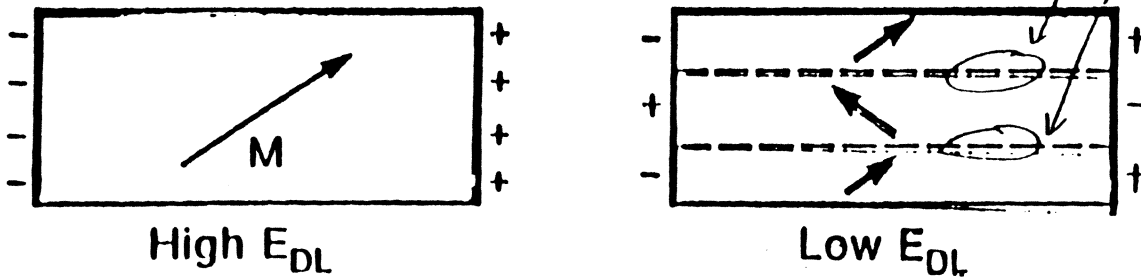
• Origin of domains

From Ching Tsang, 1977

Two Possible Causes for Domain Formation:

- Dispersion of the Anisotropy Easy-Axis
- Longitudinal Demagnetization Effect

no magnetic charge at these boundaries.

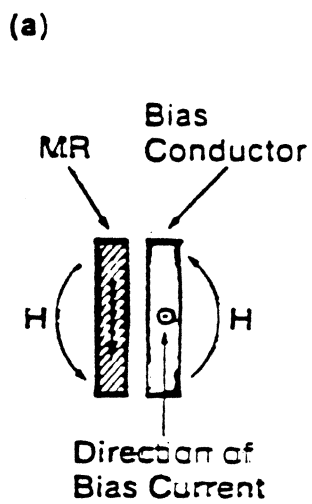


$E_{DL} \equiv$ Energy from longitudinal demagnetizing fields,

from Ching Tsang, IBM

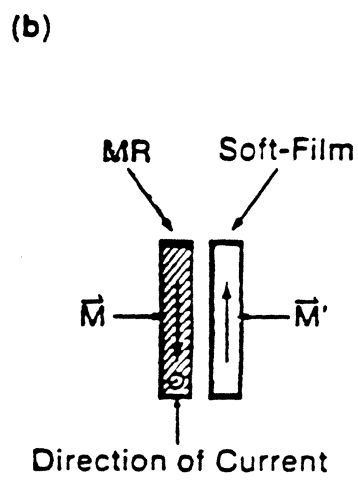
42-000-000

Techniques for Generating Transverse Bias Field:



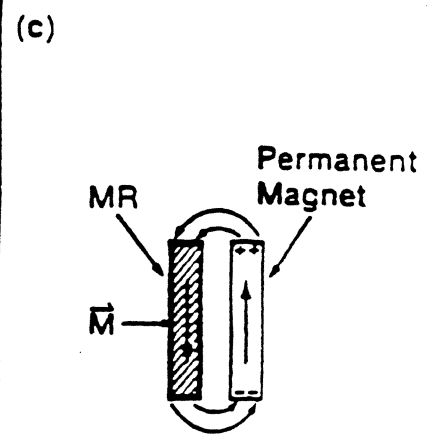
Shunt Biasing

- Weak effects
- Electrical shunting
- Non-uniform bias profile
- Enhancements



Soft-Film Biasing

- Strong effects permeability limited
- Thin spacer for uniform bias profile
- Selection of soft-film material important
- Saturated soft-film operations



Permanent Magnet Biasing

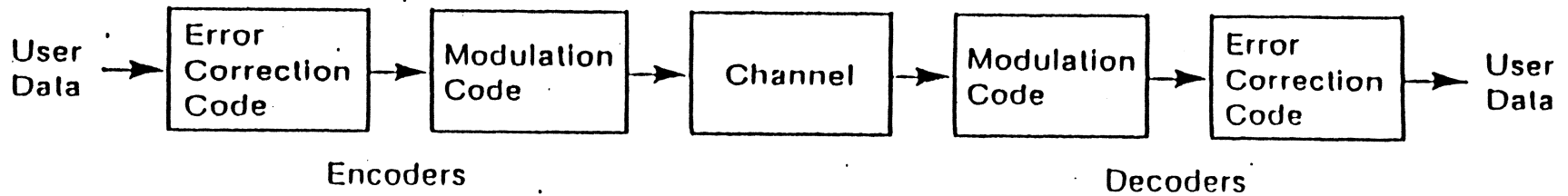
- Selection of permanent magnet film important
- Good bias profile

MAGNETIC RECORDING CHANNEL

IIST

A. Hoagland 3/90

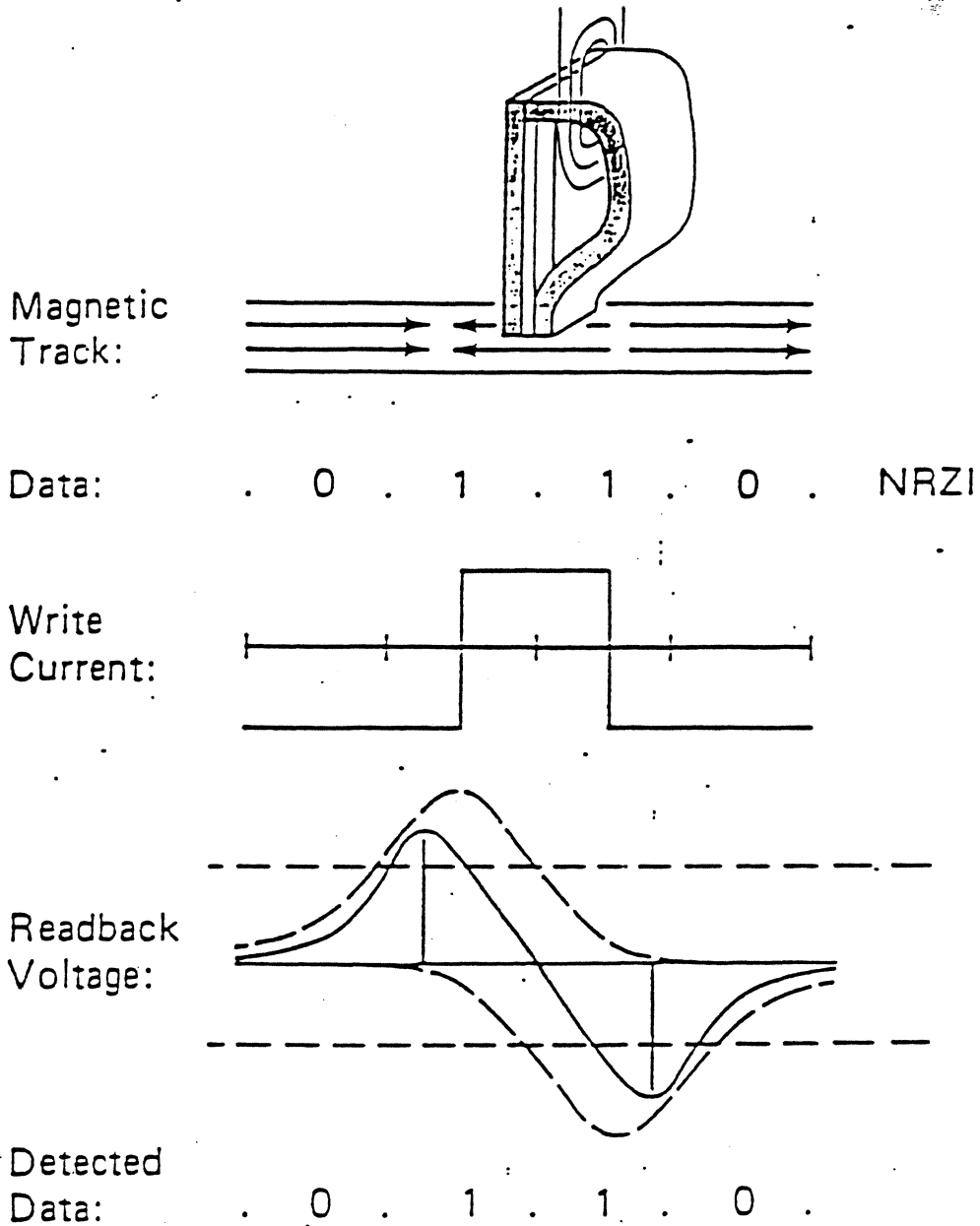
Configuration of Codes



Modulation Code: Matches recorded signal characteristics to channel bandwidth, detection method, read/write electronics, timing and tracking servo requirements

Error Correction Code: Detects and corrects data detection errors

Digital Magnetic Recording Channel



Causes of bit detection errors

- Random noise
 - Intersymbols Interference
 - Loss of clock synchronization
- } Bit pattern related

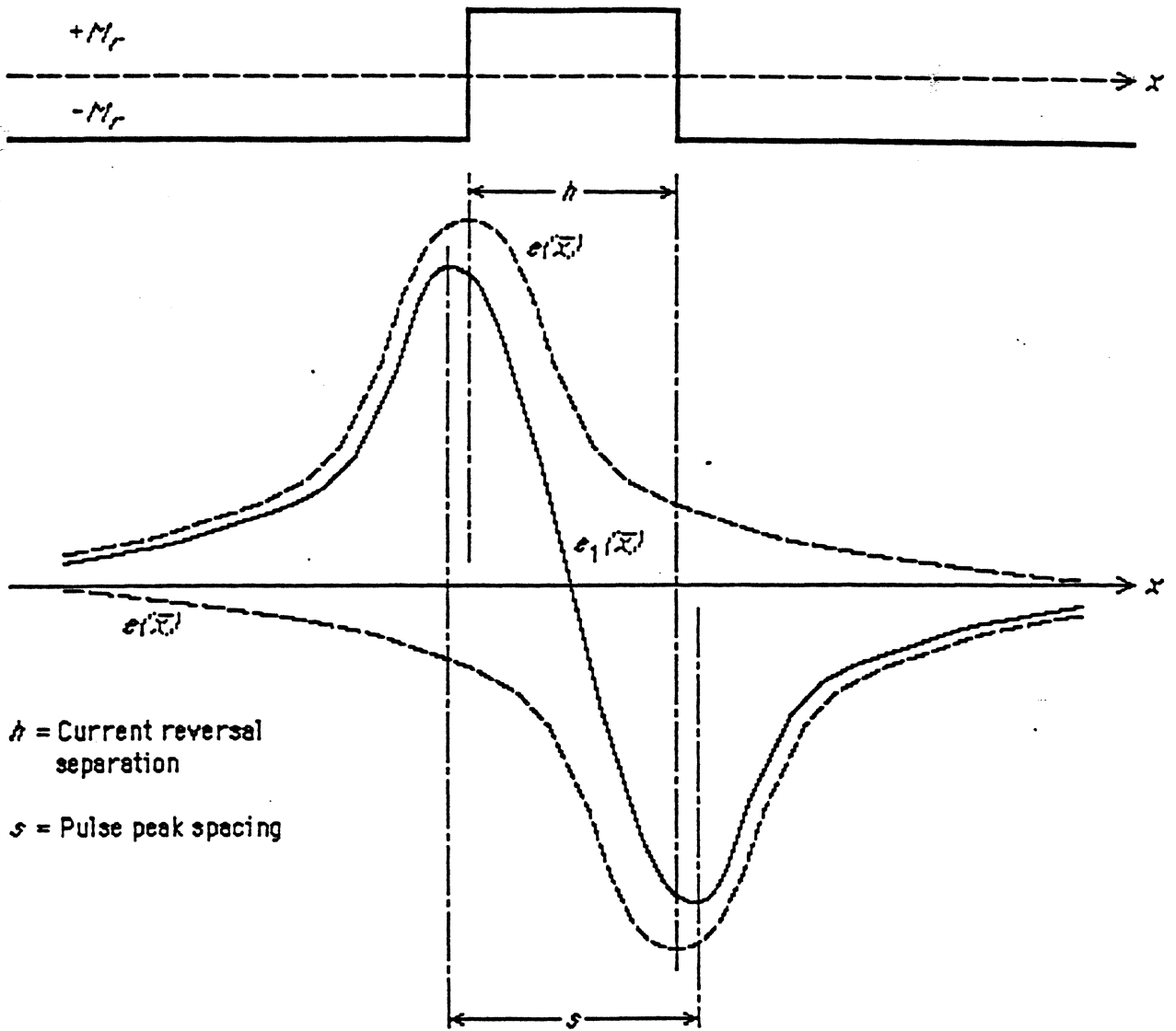
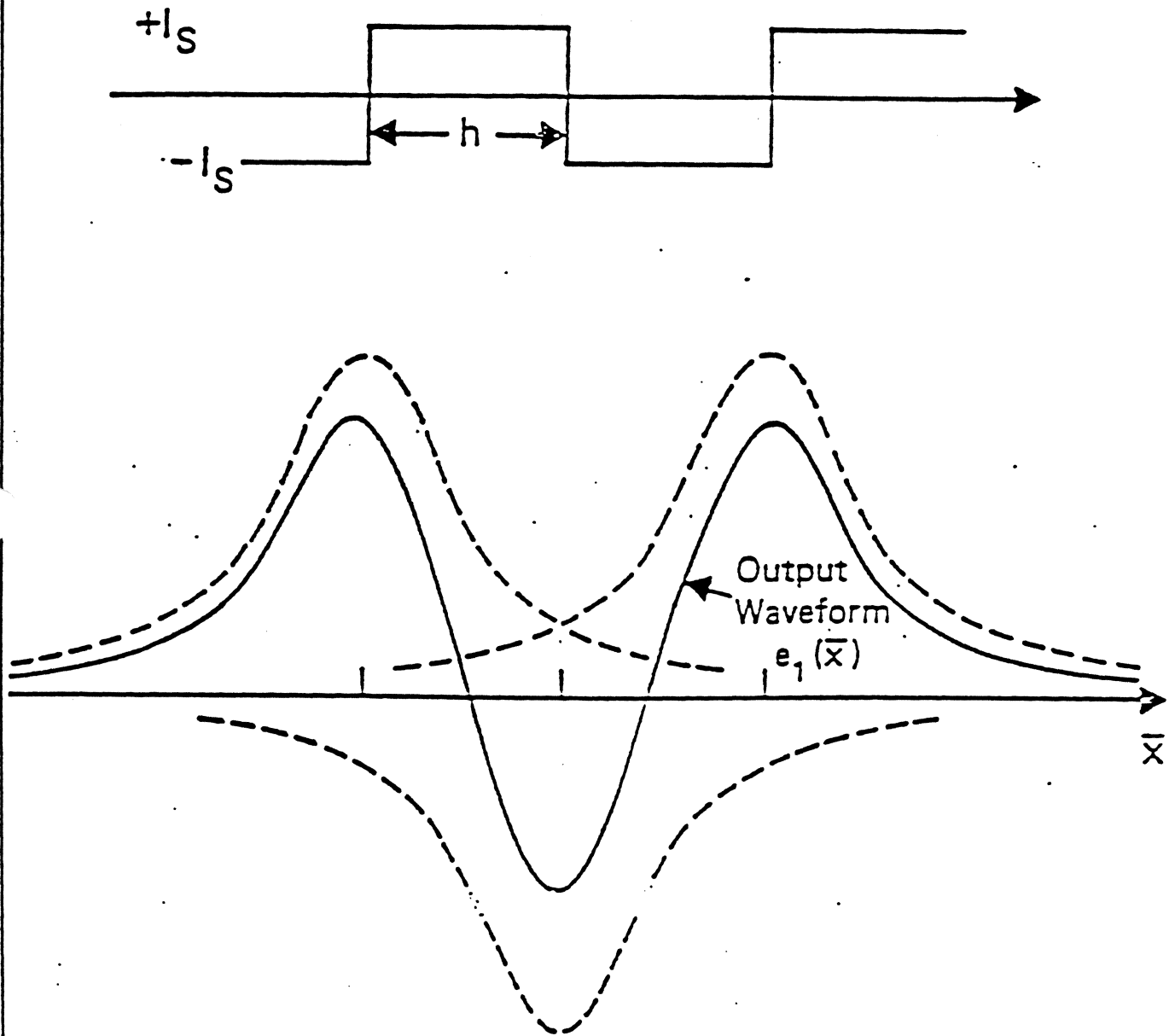


FIGURE 7.7

Pulse Crowding



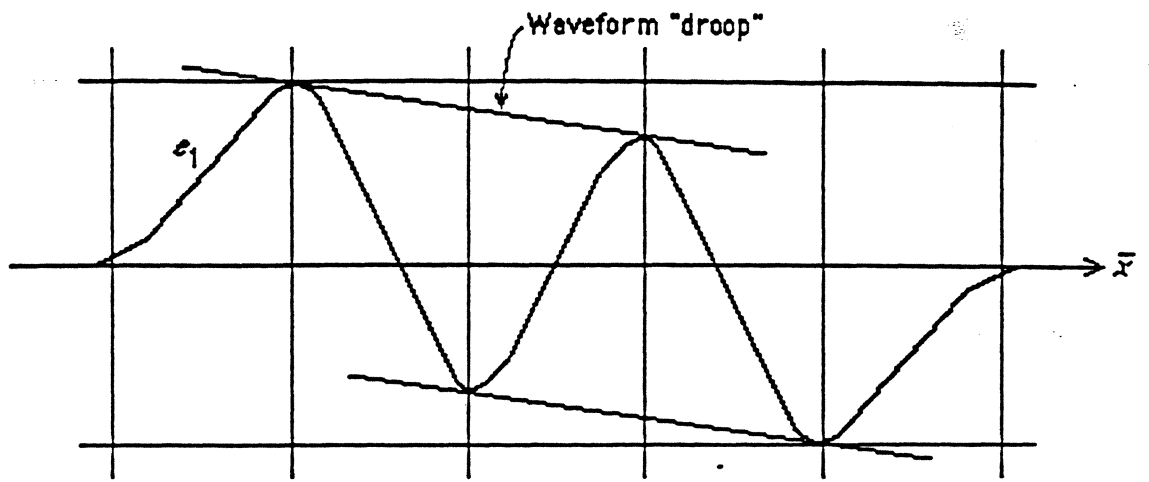


FIGURE 7 10

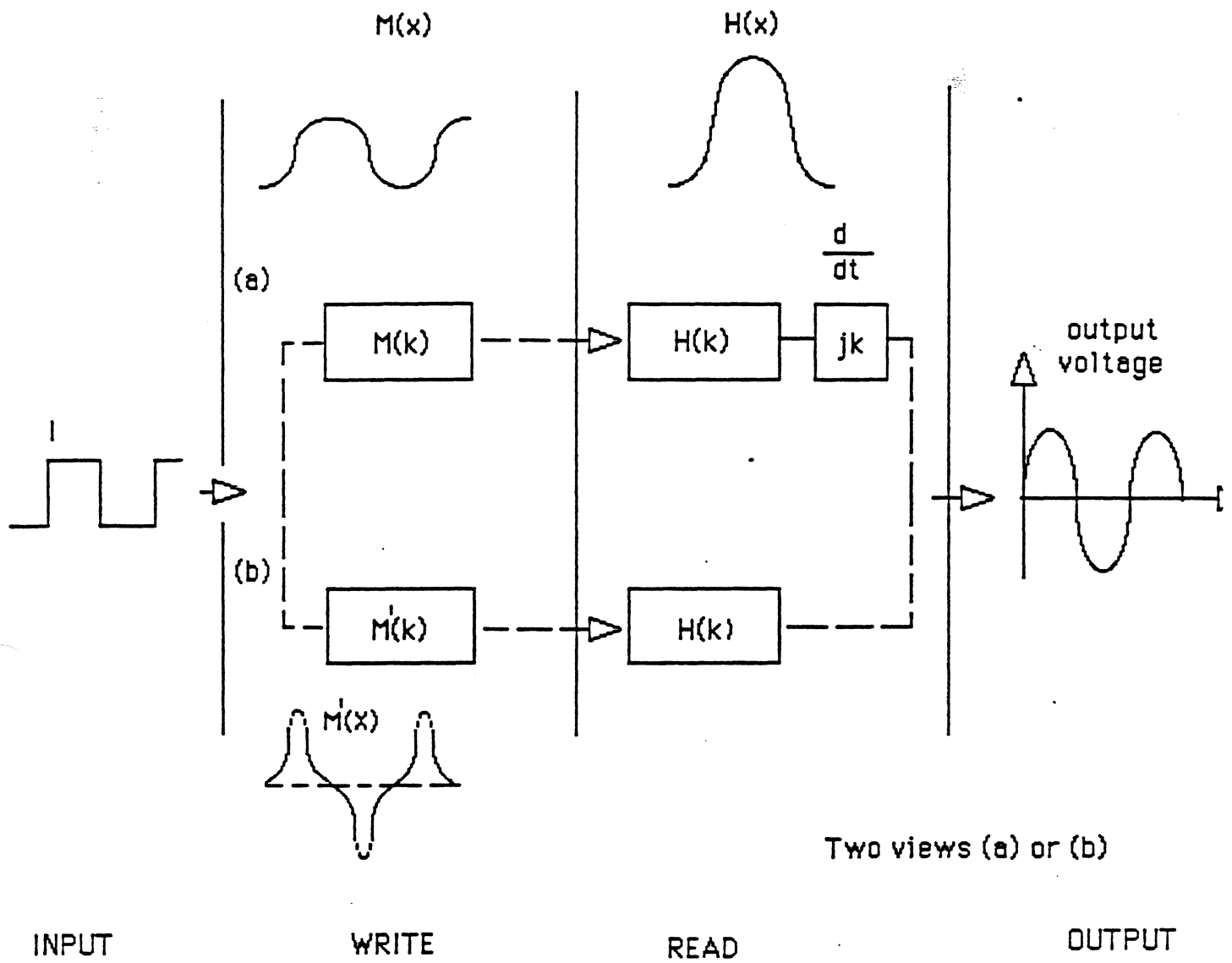


FIGURE 7.4

Recording channel wavelength response (transfer function)

$$M'_X(k) = jkM_X(k)$$

jk represents the differentiation on readback

and

$$H_X(k,d) = \left\{ \frac{1-e^{-k\delta}}{k} \right\} \cdot e^{-kd} \cdot \frac{\sin(kg/2)}{kg/2}$$

thickness spacing gap
factor loss loss

or

$$E(k) = jk\mu_0 N v W \delta \left\{ \frac{1-e^{-k\delta}}{k} \right\} \cdot e^{-kd} \cdot \frac{\sin(kg/2)}{kg/2} \cdot M_X(k)$$

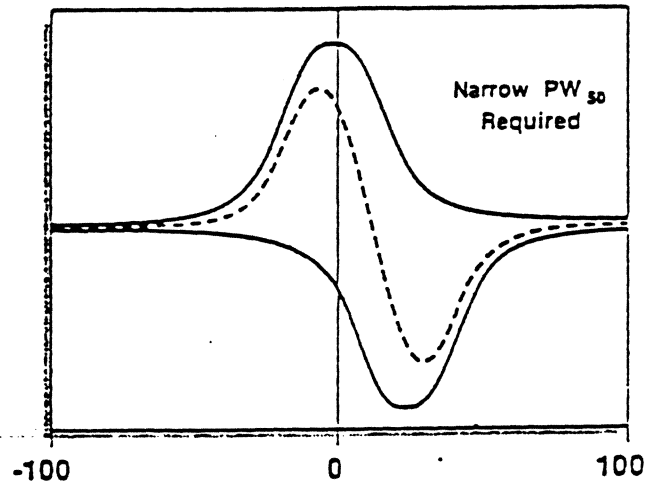
N = number turns

v = velocity

W = track width

δ = medium thickness

HIGH DENSITY RECORDING



$$PW_{50} = \sqrt{g^2 + 4(a+d)(a+d+\delta)}$$

g = gap length

a = transition parameter

d = flying height

δ = media thickness

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

H_c = media coercivity

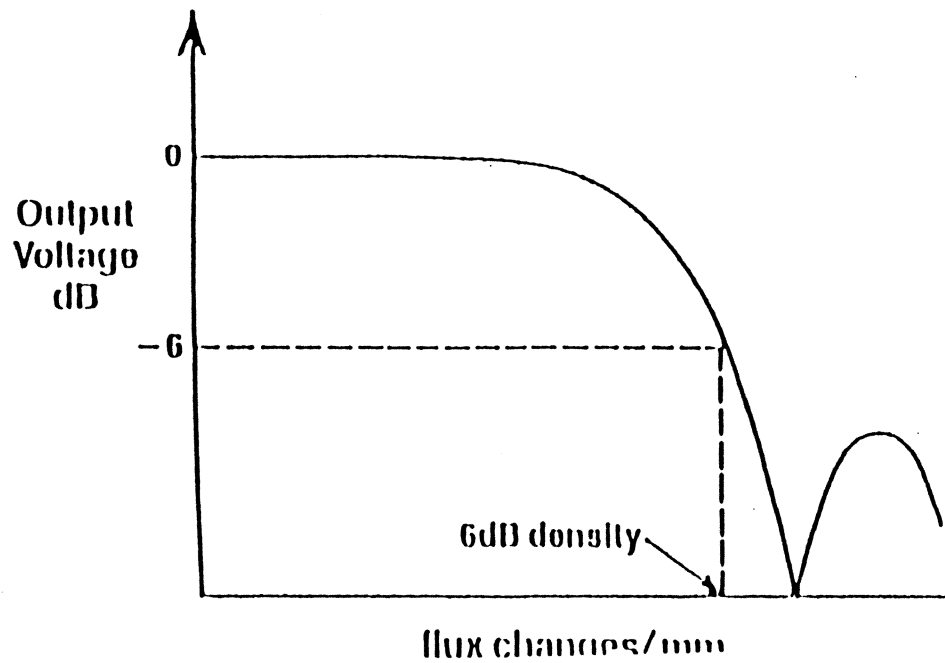
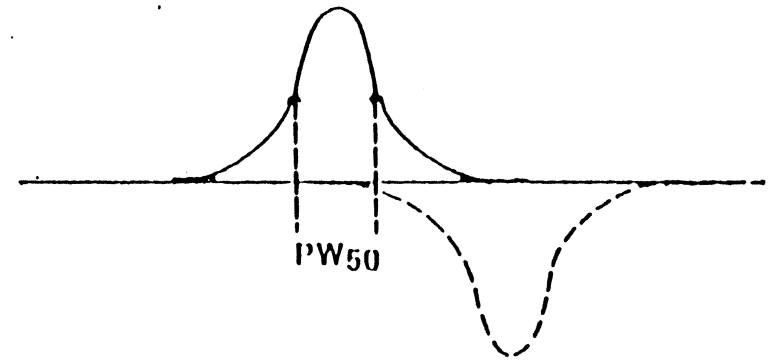
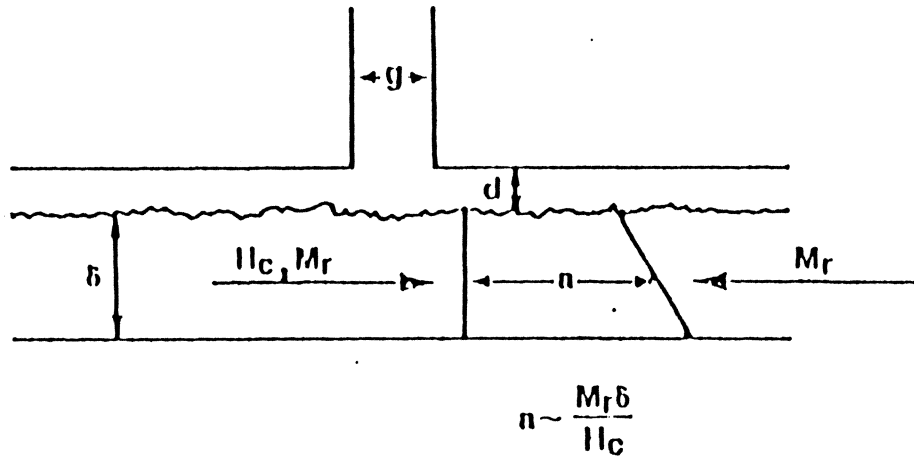
M_r = media remanent flux density

Source: Williams - Comstock (6)

MIG - Sutton

Applied Magnetics Corporation

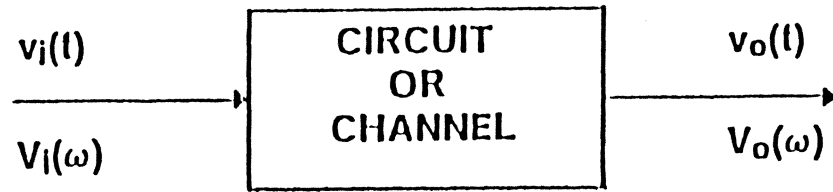
Linear Density: Longitudinal Recording



$$6\text{dB density} \approx \frac{1.4}{\text{PW}_{50}}$$

$$= \frac{1.4}{[g^2 + 4(d + n + \delta)(d + n)]^{1/2}}$$

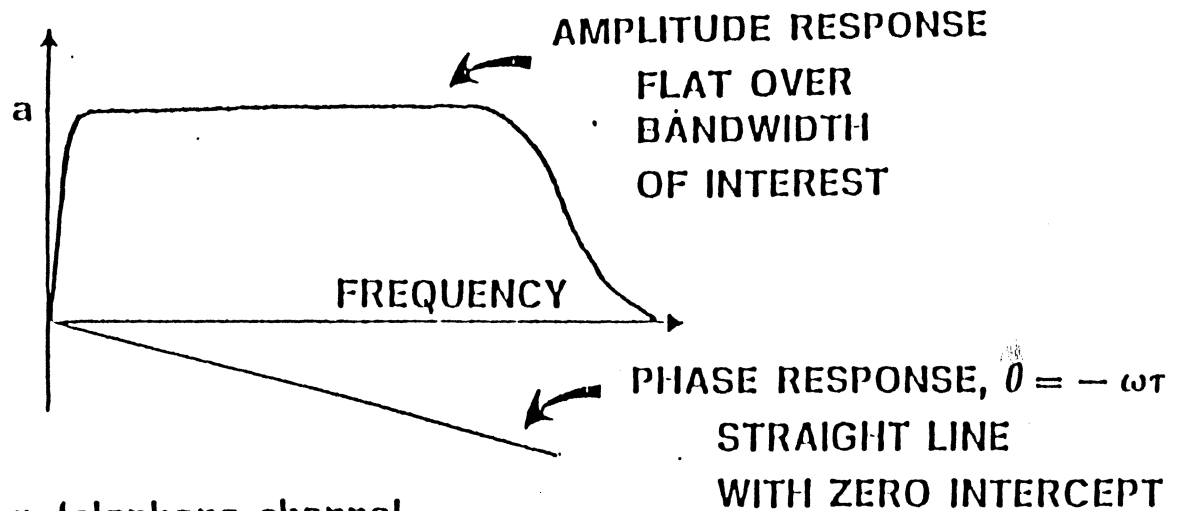
DISTORTIONLESS REPRODUCTION



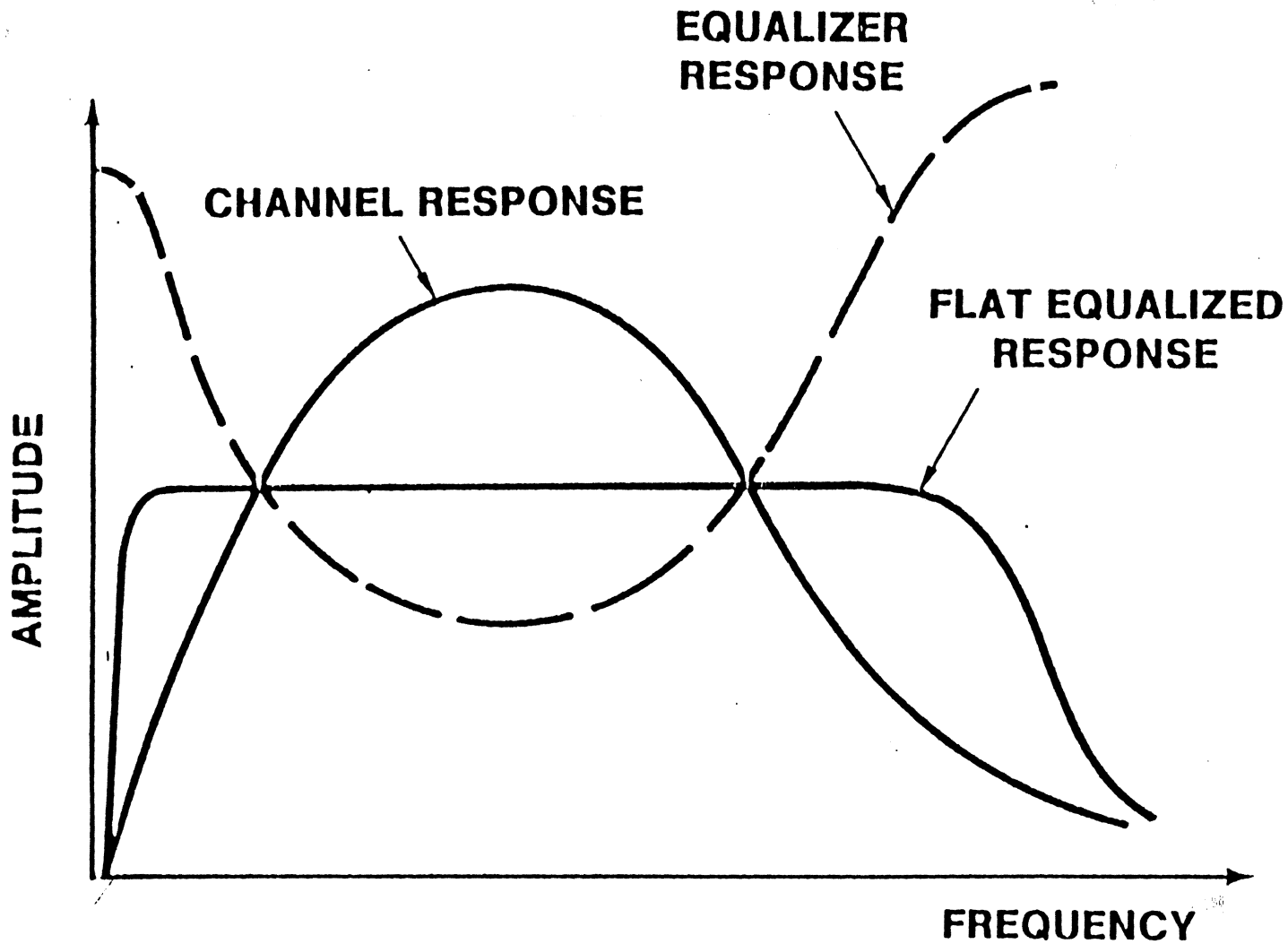
TRANSFER FUNCTION

$$H(\omega) = V_o(\omega)/V_i(\omega) \quad (\text{Independent of signal})$$

FOR A
DISTORTIONLESS
CHANNEL



e.g., Audio amplifier, telephone channel



**EQUALIZATION OF CHANNEL RESPONSE
(AUDIO RECORDERS, INSTRUMENTATION
RECORDERS)**

Saturation Recording

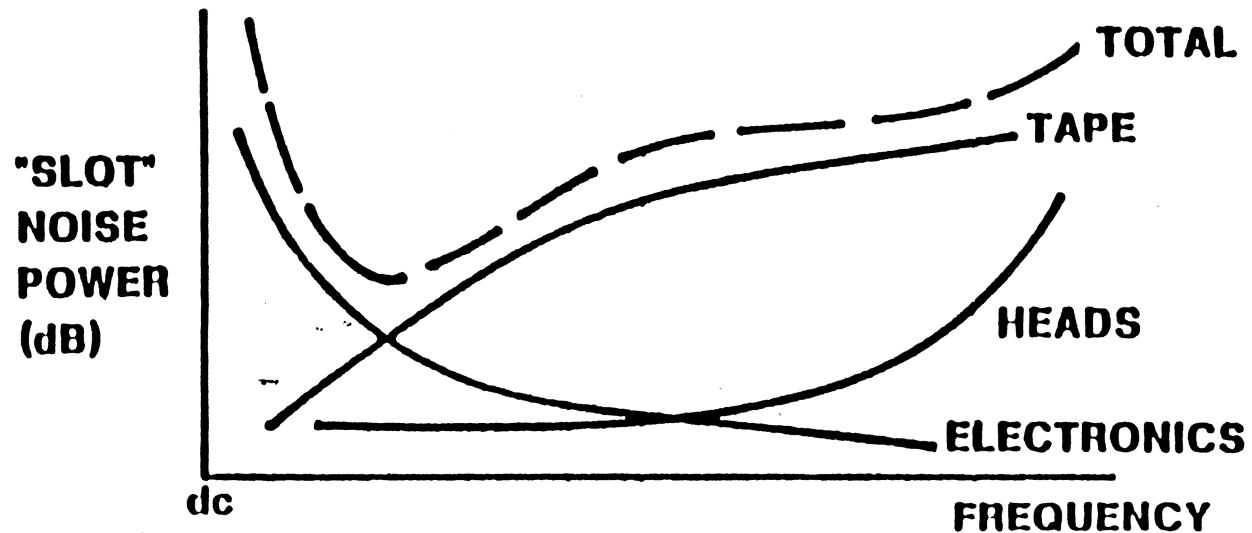
- For Track Density Change

- Signal $\propto \frac{1}{\text{tpi}}$

- Media Noise $\propto \frac{1}{\sqrt{\text{tpi}}}$

or SNR decreases 3db as double tpi

THREE MAIN SOURCES OF NOISE



**ORIGIN OF NOISES IS ALWAYS SOME
UNCERTAINTY ABOUT A PHENOMENON:**

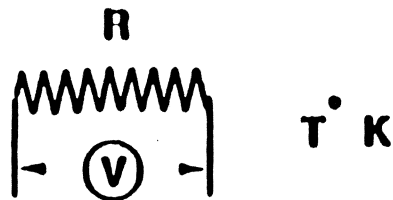
**ELECTRONICS: ELECTRON (NOISE) STATISTICS IN FIRST
STAGE REPRODUCE PRE-AMPLIFIER**

RESISTORS: ELECTRON STATISTICS

HEADS: DOMAIN WALL/ANGLE STATISTICS

**TAPES: TAPE PARTICLE POSITION OR ORIENTATION
STATISTICS**

RESISTORS
 / JOHNSON NOISE
 \ NYQUIST THEOREM



$$V = \sqrt{4kTR \Delta f} \text{ VOLTS}$$

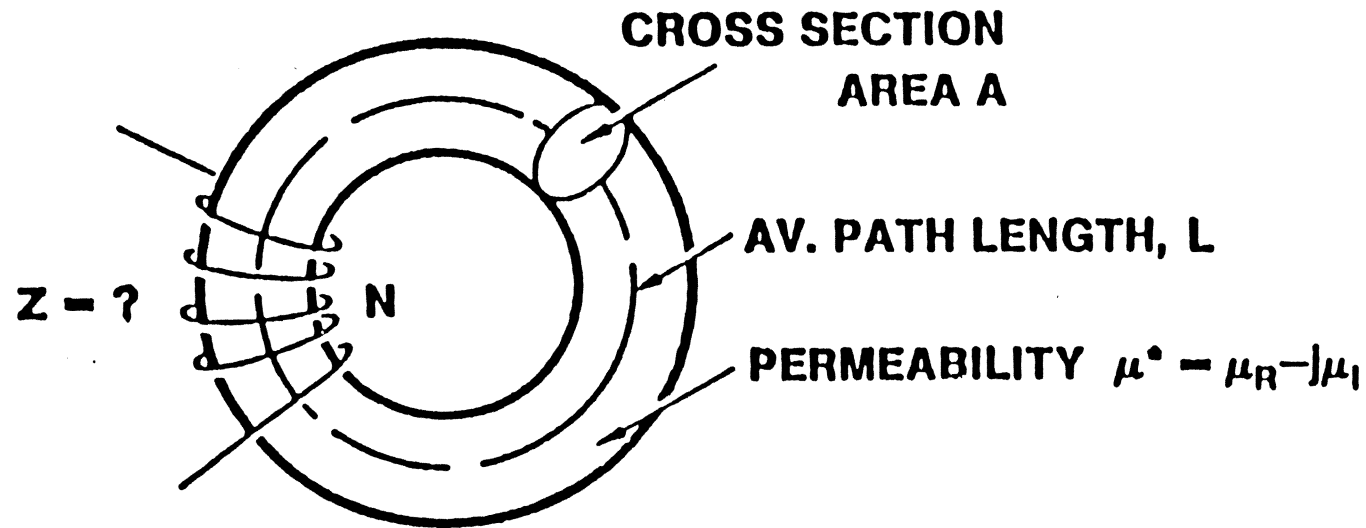
k = BOLTZMAN'S CONSTANT
(1.38 10⁻²³ JOULES/DEGREE)

T = ABSOLUTE TEMPERATURE

R = RESISTANCE IN OHMS

Δ f = BANDWIDTH OF MEASUREMENT
 (< HERZ
 CPS)

ELECTRICAL IMPEDANCE OF A TOROID



$$\text{INDUCTANCE, } L^* = (0.4 \pi 10^{-8}) N^2 A \frac{\mu^*}{L}$$

$$\text{IMPEDANCE} = j\omega L^* \approx j\omega \frac{N^2 A}{L} \mu_R + \omega \frac{N^2 A}{L} \mu_I$$

PURE INDUCTANCE

STORES ENERGY

\therefore NOISELESS

RESISTANCE

DISSIPATES ENERGY

\therefore NOISY

IF PARTICLE:

POSITIONS ARE RANDOM

DIRECTIONS OF MAGNETIZATION ARE RANDOM

**THEN, TO FIND THE TOTAL NOISE POWER, WE
SIMPLY ADD THE POWERS**

IF E_1 AND E_2 ARE TWO SIGNALS,

$$\overline{(E_1 + E_2)^2} = \overline{E_1^2} + \overline{E_2^2} + 2\overline{E_1 E_2}$$

IF E_1 AND E_2 ARE RANDOM $\overline{E_1 E_2} = 0$

**UNCORRELATED
INCOHERENT**

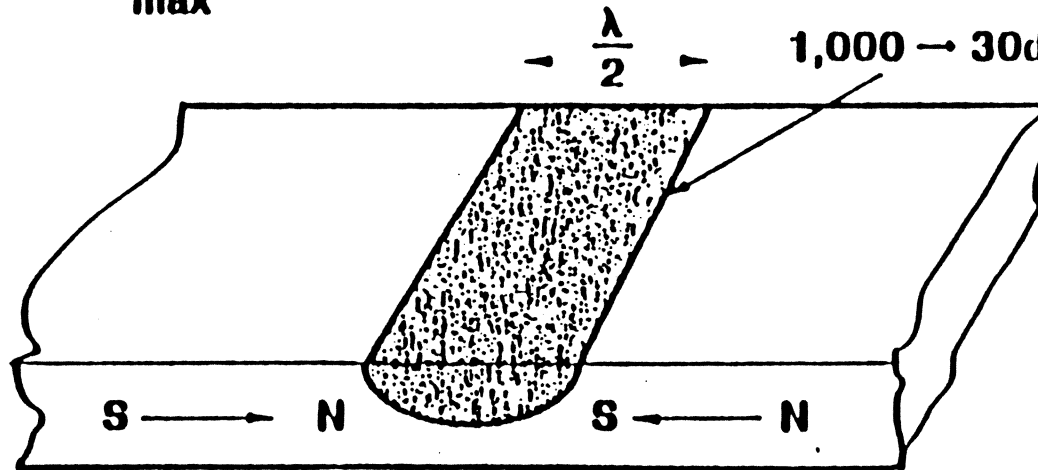
**THE NOISE POWER IS THE SUM OF THE SQUARES
(NOT THE SQUARE OF THE SUM)**

WHY IS IT THAT:

$$\text{SNR} \propto \frac{nW}{k_{\max}^2} \propto nW\lambda_{\min}^2 \quad ???$$

10,000 → 40dB

1,000 → 30dB



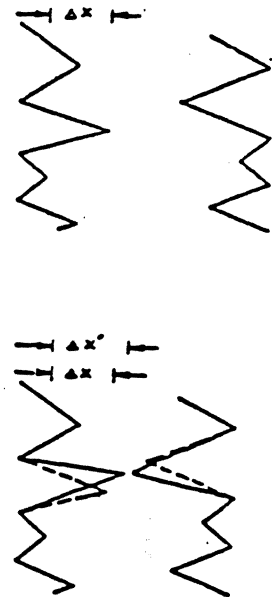
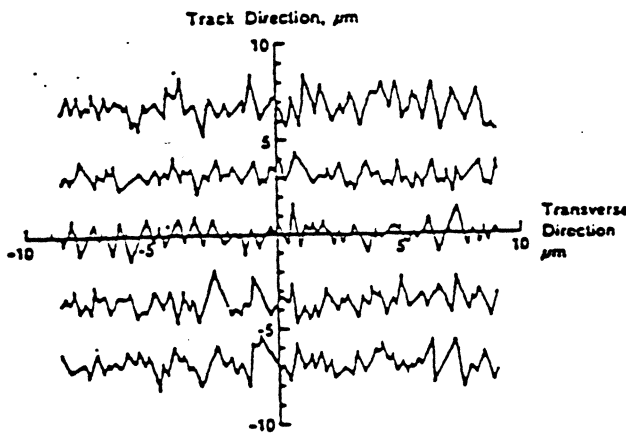
THE REPRODUCE HEAD ONLY SENSES A VOLUME OF TAPE, $1/2$ A WAVELENGTH LONG AND ABOUT $1/5 \lambda$ DEEP, AND THE TRACKWIDTH WIDE:

$$\text{VOLUME} \propto W\lambda^2$$

∴ ALWAYS BETTER TO REDUCE W THAN REDUCE λ !!

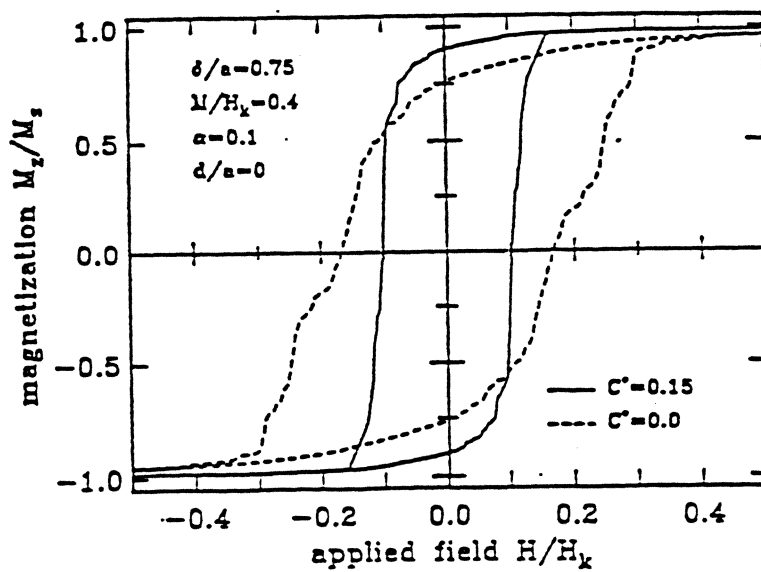
MEDIA NOISE

- Noise in thin metallic media is associated *mainly* with written transitions.
- Position jitter and amplitude variations contribute to total noise.
- In general, maximum noise occurs at a density which is related to the transition width ($\lambda_c \approx 4a$, Belk et al).
- Two regimes, (i) noise initially increases linearly with density followed by (ii) "supra-linear" increase as transitions crowd together.

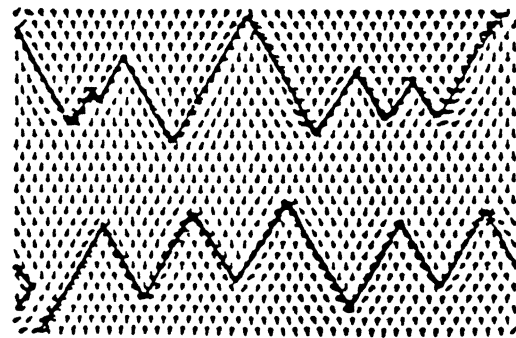
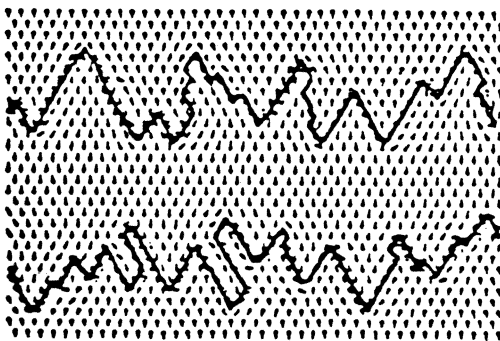


CONTROLLING MEDIA NOISE (2)

- Reducing the coupling between grains reduces the coercive squareness of the hysteresis loop, and decreases the irregularity in the transition boundary.



(Zhu and Bertram, J. Appl. Phys. 63 (1988))



EFFECT OF REDUCED EXCHANGE COUPLING

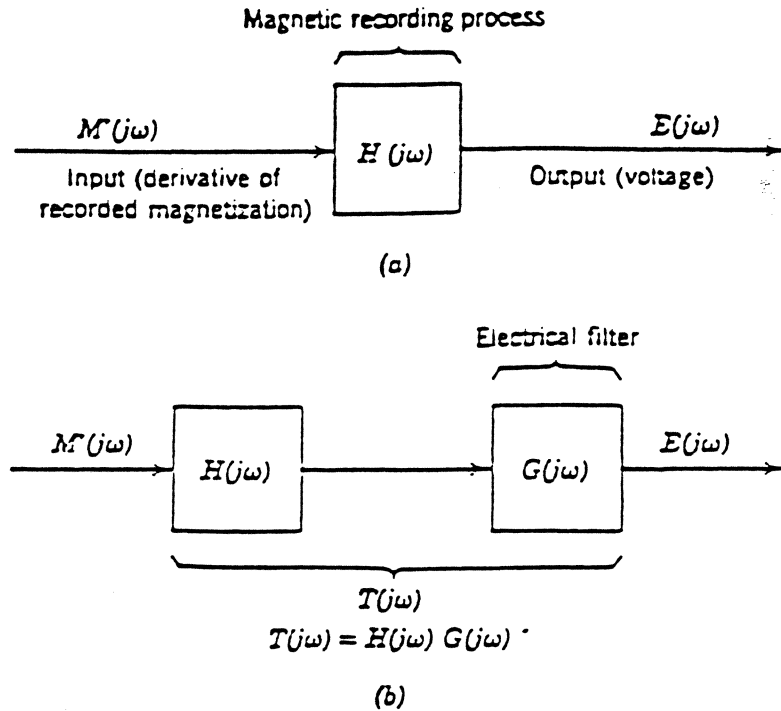


Fig. 6.20. Input-output transfer function block diagrams.

been defining $H(z)$ for $-\infty < z < +\infty$, and since in our analogy \bar{z} (or z) corresponds to "time," we see that our impulse response transfer function provides an anticipatory output.

Now we can modify the input-output relation shown in Fig. 6.20a by insertion of an electrical filter at the head output. If the voltage at the input terminals of this filter is not disturbed by its presence, we can redraw the input-output transfer relation, as shown in Fig. 6.20b, where $G(j\omega)$ represents the frequency characteristic of the electrical filter. This situation will apply, for H_z represents the transfer function of the magnetic recording process, and the electrical parameters of the magnetic head can be viewed as incorporated into the design of the inserted filter. The input $M'(z)$ is then related to the output voltage $e(z)$ by the following transfer function:

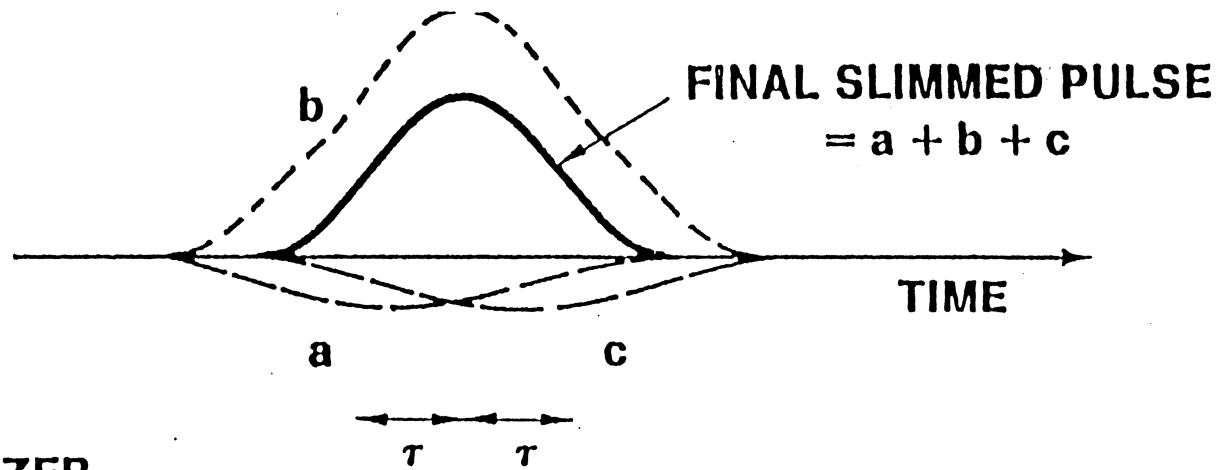
$$T(j\omega) = H(j\omega)G(j\omega) \quad (6.26)$$

Theoretically, if $G(j\omega)$ could be chosen such that

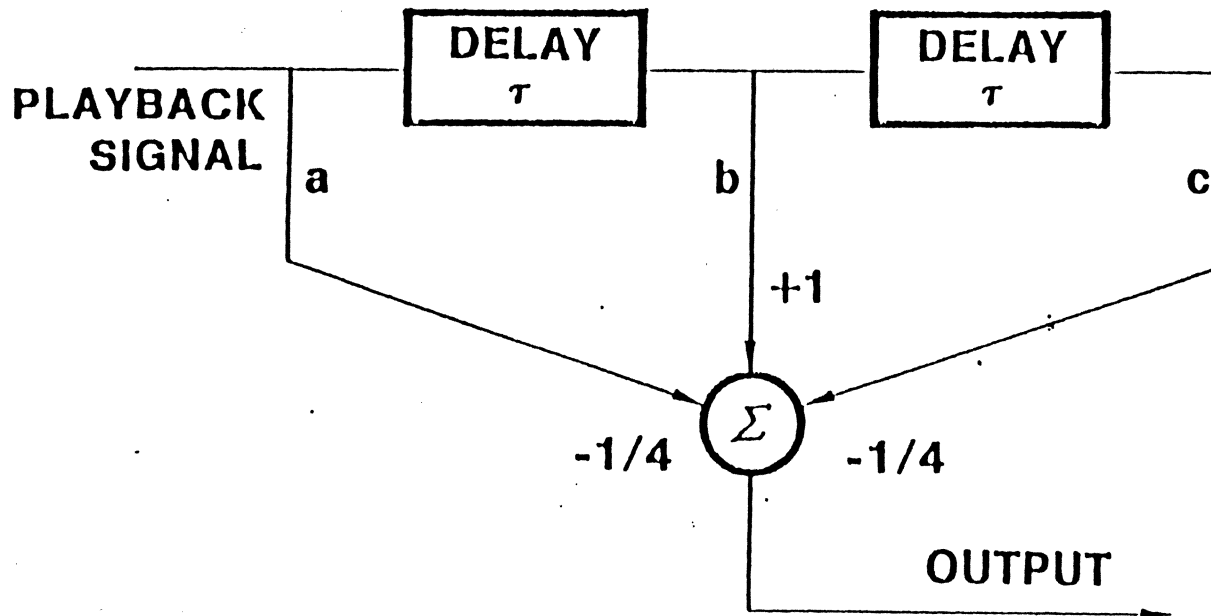
$$H(j\omega)G(j\omega) \approx 1.0$$

over the read signal bandwidth, then $E(j\omega)$ would be identically equal to $M'(j\omega)$ or

$$e(\bar{z}) = M'(\bar{z})$$



COSINE EQUALIZER



PULSE SLIMMING

RLL Code Classification (after Fransascek)

for instance MFM:

d	k	m	n	r
1	3	1	2	-

Number of word lengths
(Variable word length codes
only)

Number of code cell per bit
(=number of possible flux
reversal positions per bits)

maximum number of bits
required for coding at
any one time

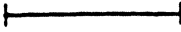
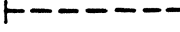


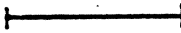

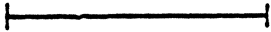

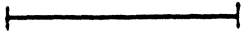
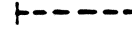
m/n: ratio data to code bits

maximum number of code cells
between flux transitions

minimum number of code cells
between flux transitions

Source: N. Mackintosh

Comparison of RLL (d,k) Constraints

Data:	 
	. 1 . 1 . 0 . 1 .
1/2(0,1)	 
	. 1 . 1 . 1 . 1 . 0 . 1 . 1 . 1 .
1/2(1,3)	 
	. 1 . 0 . 1 . 0 . 0 . 0 . 1 . 0 .
1/2(2,7):	 
	. 1 . 0 . 0 . 1 . 0 . 0 . 0 . 0 .
2/3(1,7):	 
	. 1 . 0 . 1 . 0 . 0 . 0 .

$(0,1) - (1,3) - \begin{matrix} (2,7) \\ (1,7) \end{matrix} \left\{ \begin{array}{l} \text{Less pulse crowding } \text{———} \\ \text{Same/larger window } \text{- - - - -} \end{array} \right.$

(2,7) vs (1,7) Involves tradeoffs

(2,7) vs MFM (1,3)

Fixed User Bit Density

MFM . 1 . 0 . 1 . $T_{min} = 2$

(2,7) . 1 . 0 . 0 . 1 . $T_{min} = 3$

(2,7) reduces intersymbol interference

Fixed Minimum Transition Spacing T_{min}

MFM . 1 . 0 . 1 . Density Ratio
 $2 \times 1/2 = 1 \text{ bit}/T_{min}$

(2,7) . 1 . 0 . 0 . 1 . Density Ratio
 $3 \times 1/2 = 1.5 \text{ bit}/T_{min}$

(2,7) increases density ratio by 50%
with decreased detection window

CHANNEL CODE PARAMETERS

T = DATA BIT TIME INTERVAL

m = # OF DATA BITS IN GROUP

n = # OF 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 = CLOCKING WINDOW

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

DR = DATA DENSITY/TRANSITION DENSITY

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

LOW FREQ ALL ONES/HIGH FREQ ALL ONES

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

Practical Example: Rate 2/3, (1,7) Code

<u>Data</u>	<u>Code</u>
00	101
01	100
10	001
<u>11</u>	<u>010</u>

Basic Coding Table

<u>Data</u>	<u>Code</u>
00.00	101.000
00.01	100.000
10.00	001.000
<u>10.01</u>	<u>010.000</u>

Substituting Coding Table

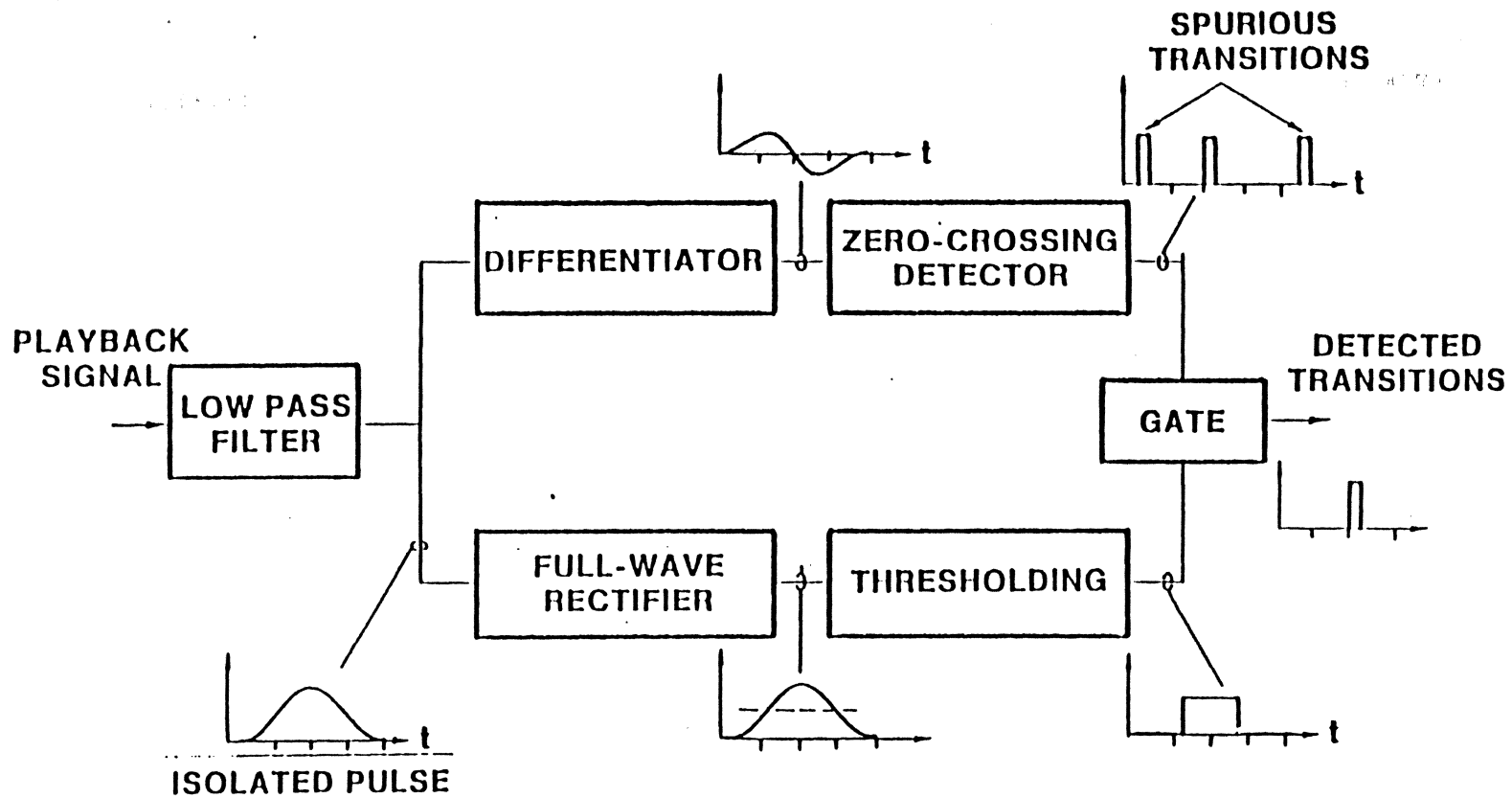
	M	N	rate	d	k	T_{min}	T_{max}	DR	W	CLOCK	DSV
NRZ	1	1	1	0	∞	T	∞	1	T	1/T	∞
DOUBLE FREQUENCY	1	2	1/2	0	1	T/2	T	1/2	T/2	2/T	T
MILLER MFM	1	2	1/2	1	3	T	2T	1	T/2	2/T	∞
ZM	1	2	1/2	1	3	T	2T	1	T/2	2/T	3T/2
MILLER ²	1	2	1/2	1	5	T	3T	1	T/2	2/T	3T/2
3PM	3	6	1/2	2	7	3T/2	4T	3/2	T/2	2/T	∞
2,7 RLL	2 3 4	4 6 8	1/2	2	7	3T/2	4T	3/2	T/2	2/T	∞
1,7 RLL	2 4	3 6	2/3	1	7	4T/3	16T/3	4/3	2T/3	3/2T	∞

M N rate d k T_{min} T_{max}

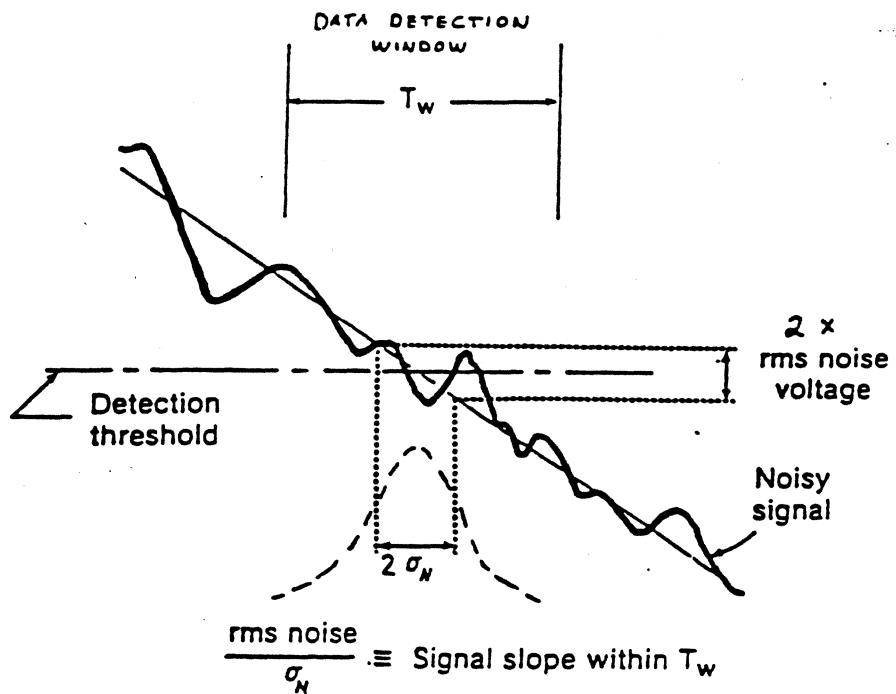
CHANNEL-CODE COMPARISON

69

SIGNAL DETECTION



GATED PEAK DETECTION



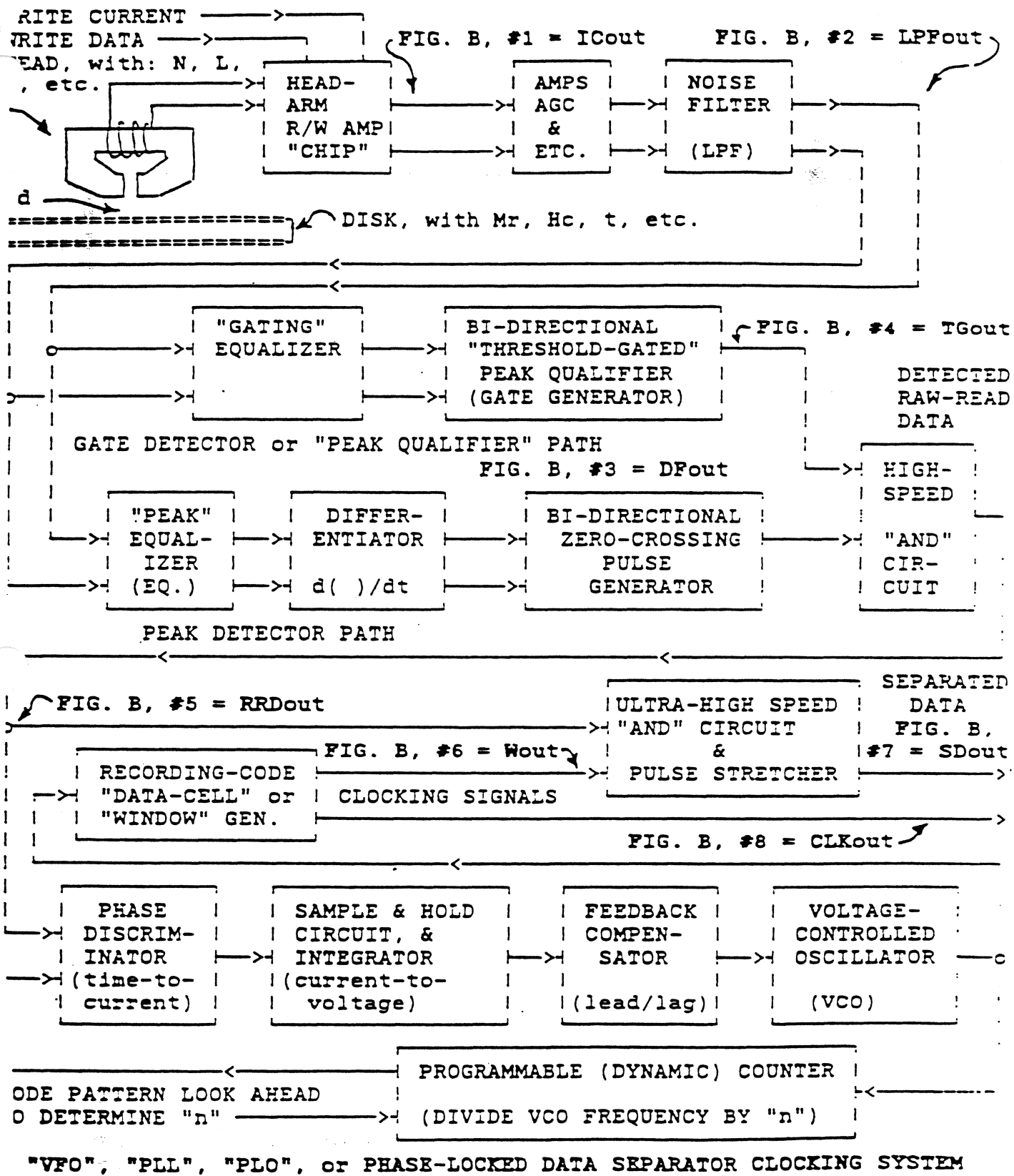
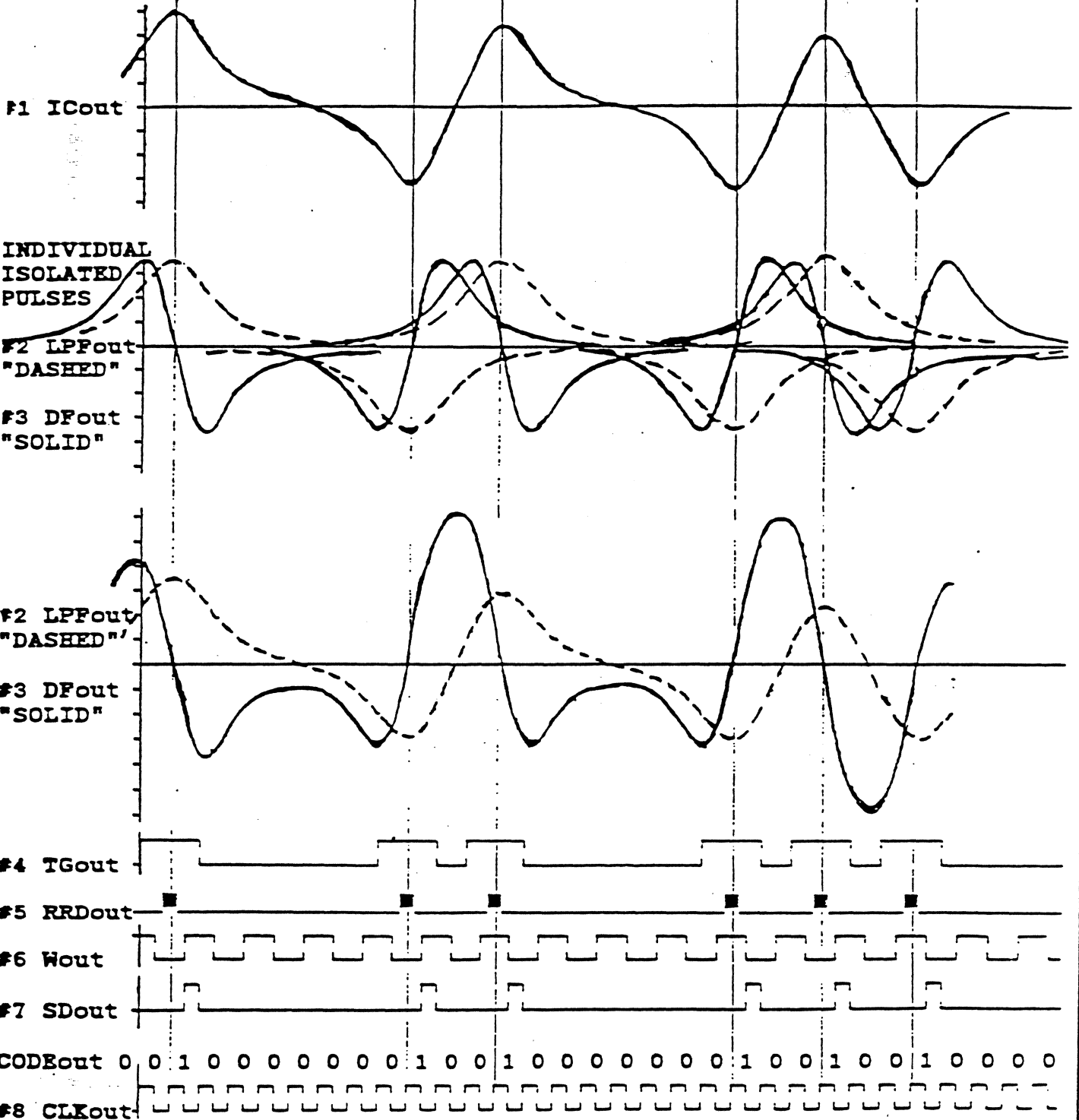


FIGURE A. TYPICAL HIGH-PERFORMANCE DISK DRIVE RECORDING CHANNEL

F. J. SORDELLO 1988

READBACK TRANSITIONS for 2, 7, RLL

CODE -> 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 1 0 0 0 0



READBACK SIGNALS AT VARIOUS STAGES WITHIN READ CHANNEL.
 (All intrinsic circuit delays have been eliminated for clarity. For 24 Mbit/sec = 3 Mbyte/sec, one "Wout" cell (1/2 complete cycle of "Wout") = 20.833 nanosecond.

F. J. SORDELLO 1988

$$\text{BIT SHIFT} = \frac{\% \text{ AMPLITUDE OF UNWANTED ADDITIVE SIGNAL}}{\text{SLOPE OF } D_{\text{fout}} \text{ SIGNAL NEAR "ZERO-CROSSING" POINT}} = \text{nsec}$$

IT IS AS SIMPLE AS THAT! ANY UNWANTED SIGNAL THAT "LIFTS" OR "LOWERS" THE DESIRED RECORDED INFORMATION BEARING D_{fout} SIGNAL WILL CAUSE THE "ZERO-CROSSING" OF THE COMPOSITE D_{fout} SIGNAL TO BE SHIFTED IN TIME FROM THE IDEAL "ZERO-CROSSING" TIME. THE TIME WIDTH OF THE DATA TRANSITION-CLOCKING WINDOW, W_{out} , DETERMINES THE TOTAL AMOUNT OF "LIFTING" OR "LOWERING" THAT CAN BE TOLERATED BEFORE THE COMPOSITE D_{fout} SIGNAL HAS ITS "ZERO-CROSSING" SHIFTED IN TIME BEYOND A PARTICULAR TRANSITION OR "BIT" CELL. FIGURE D SHOWS THAT THE 24 MEGABIT/SECOND RECORDING SYSTEM DESCRIBED CAN TOLERATE COMPOSITE D_{fout} "ZERO-CROSSING" TIME-SHIFTS ("BIT SHIFTS") UP TO $\pm (20.8 \text{ nsec})/2 = \pm 10.4 \text{ nsec}$ BEFORE OCCURRING OUTSIDE OF THE TIME BOUNDARIES OF W_{out} . FIGURE D ALSO SHOWS THAT THE $\pm 10.4 \text{ nsec}$ BIT SHIFT WILL BE CAUSED BY THE ADDITION OF UNWANTED SIGNALS EQUAL TO APPROXIMATELY $\pm 80\%$ OF THE BASE-TO-PEAK VALUE OF THE INFORMATION BEARING D_{fout} SIGNAL.

THE ABOVE NUMBERS ARE ASSOCIATED WITH A SINGLE ISOLATED READBACK TRANSITION. THE AMPLITUDE OF THE BASE-TO-PEAK VALUE OF D_{fout} CONTROLS THE ABSOLUTE AMOUNT OF "BIT SHIFT" THAT OCCURS FOR A GIVEN AMOUNT OF UNWANTED ADDITIVE SIGNAL. WHEN PORTIONS OF ISOLATED D_{fout} PULSES COMBINE DUE TO THEIR PROXIMITY, THE BASE-TO-PEAK VALUE OF THE NET RESULTING D_{fout} SIGNAL MAY INCREASE BY AS MUCH AS 50%, YIELDING LESS "BIT SHIFT" FOR A GIVEN AMOUNT OF UNWANTED ADDITIVE SIGNAL. SEE FIGURE , ON PAGE 3, AND ESTIMATE THE POSITIVE PORTIONS OF THE D_{fout} SIGNAL CORRESPONDING TO THE SECOND, THIRD, FOURTH, AND FIFTH TRANSITIONS; AND THE NEGATIVE PORTIONS OF THE FIFTH AND SIXTH TRANSITIONS. THEN, COMPARE THE ESTIMATED PORTIONS STATED ABOVE WITH THE NEGATIVE PORTIONS OF THE D_{fout} SIGNAL CORRESPONDING TO THE FIRST, SECOND, THIRD, AND FOURTH TRANSITIONS, AS WELL AS THE POSITIVE PORTIONS OF THE FIRST AND SIXTH TRANSITIONS.

THERE ARE OTHER WAYS THAT THE "ZERO CROSSING" CAN OCCUR SHIFTED. MOST PROMINENT IS BY WRITING THE ZERO MAGNETIZATION POINT IN THE MEDIUM IN AN UNWANTED, SHIFTED LOCATION. HEAD/MEDIUM COMBINATIONS WITH POOR OVERWRITE CAPABILITIES, COMBINED WITH FINITE WRITE CURRENT AND/OR HEAD GENERATED WRITE FLUX RISETIMES AND PREVIOUS RECORDED HISTORY IN THE MEDIUM MAY CAUSE TWO OR THREE NANoseconds BIT SHIFT.

THE MAJOR CAUSES OF DIGITAL MAGNETIC RECORDING CHANNEL BIT-SHIFT ARE THE FOLLOWING:

1. NOISE INDUCED BIT-SHIFT (NIB).
2. PATTERN INDUCED BIT-SHIFT (PIB).
3. OVERWRITE INDUCED BIT-SHIFT (OWIB).

F. J. SORDELLO 1958

LIST

4. ASYMMETRY INDUCED BIT-SHIFT (ASIB).
5. SYSTEM NOISE INDUCED BIT-SHIFT (SNIB).
6. MINOR MEDIA DEFECTS THAT PASS DEFECT SCANNING TESTING BECAUSE THE MINOR DEFECTS ARE WITHIN SPECIFICATIONS. (MMD)
7. ADJACENT TRACK INDUCED BIT-SHIFT (ATIB).

NIB IS CAUSED BY MEDIA, HEAD, AND ELECTRONIC CIRCUIT NOISE SOURCES. THE INSTANTANEOUS AMPLITUDE OF THIS NOISE IS COMMONLY ASSUMED TO BE DEPENDENT UPON THE NOISE SOURCE'S RMS VALUE, MAGNIFIED BY THE SIGMAS ASSOCIATED WITH THE PROBABILITIES OF GAUSSIAN DISTRIBUTION OF CHANCE. IF THE NOISE IS ASSUMED TO BE GAUSSIAN, IT LENDS ITSELF TO THE FOLLOWING PROBABILITIES:

RMS VALUE = SIGMA

SIZE OF NOISE "SPIKE" IN TERMS OF NOISE RMS VALUE	PROBABILITY OF OCCURRENCE OF THAT NOISE "SPIKE"
0	1/1
1	1/3.2
2	1/21.74
3	1/370.4
4	1/15,625
5	1/1,724,137.9
6	1/500,000,000
6.47 ~ 6.5	1/10,000,000,000
7	1/400,000,000,000
8	1/80,000,000,000,000,000

FIGURE D SHOWS THAT IF THE 24 MEGABIT/SECOND EXAMPLE SYSTEM HAD SNR EQUAL TO 26db (BASE-TO-PEAK SIGNAL DIVIDED BY RMS NOISE) AT THE OUTPUT OF THE DIFFERENTIATOR, THE COMPOSITE "ZERO CROSSING" WOULD SHIFT IN TIME WITH AN RMS JITTER EQUAL TO:

26db = 20:1; THE RMS NOISE IS 5% OF THE BASE-TO-PEAK D_{fout} .

HENCE, THE "ZERO-CROSSING" WILL SHIFT: $(5\%)/(10\%/nsec) = 0.5 nsec. (RMS)$

* NORMALLY, SNR IS SPECIFIED AT THE HEAD'S OUTPUT OR THE HEAD'S IC PREAMP'S OUTPUT. IN ORDER TO OBTAIN THE SNR AT THE DIFFERENTIATOR'S OUTPUT, THE RMS NOISE OR THE SPECIFIED NOISE SPECTRA CHARACTERISTICS ARE OBTAINED FROM THE SPECIFIED SNR TEST CONDITIONS, THEN POWER SPECTRA MODIFIED BY THE INVOLVED READ CHANNEL BLOCKS SUCH AS AGC AMPLIFIER, NOISE FILTER, "PEAK" EQUALIZER, AND, OF COURSE, THE DIFFERENTIATOR. THE RMS NOISE PRESENT AT THE OUTPUT OF THE DIFFERENTIATOR OBTAINED CAN BE USED EITHER TO CALCULATE SNR OR DIRECTLY TO PREDICT NIB.

F. J. SORDELLO 1988

**BIT-SHIFT
ON TRACK**

	(@ ANY TIME	@ $\times 10^{10}$ (6.5*sigma)
NIB	0.71	3.25
PIB	2.4	2.4
OWIB I	0.8	0.8
OWIB II	1 -> 3	1 -> 3
ASIB	0.5 -> 2	0.5 -> 2
SNIB	0.1 -> ?	0.1 -> ?
ATIB	~0	~0
	<hr/>	<hr/>
	5.51	8.05
ERROR RATE	10^{10}	10^{10}
MARGIN	4.91	2.37

**BIT-SHIFT
ON TRACK and DURING MINOR DEFECT**

	(DFout slope divided by 0.7) @ ANY TIME	@ $\times 10^{10}$ (6.5*sigma)
	1.01	4.64
	3.43	3.43
	0.8	0.8
	1 -> 3	1 -> 3
	0.5 -> 2	0.5 -> 2
	0.14 -> ?	0.14 -> ?
	~0	~0
	<hr/>	<hr/>
	6.88	9.71
ERROR RATE	10^{10}	10^{10}
MARGIN	3.54	0.71

**BIT-SHIFT
OFF TRACK**

	(DFout slope divided by 0.75) (@ ANY TIME	@ $\times 10^{10}$ (6.5*sigma)
NIB	0.95	4.33
PIB	2.4	2.4
OWIB I	0.8	0.8
OWIB II	1 -> 3	1 -> 3
ASIB	0.5 -> 2	0.5 -> 2
SNIB	0.13 -> ?	0.13 -> ?
ATIB	1.5	1.5
	<hr/>	<hr/>
	7.28	10.66
ERROR RATE	10^{10}	10^9
MARGIN	3.14	0

**BIT-SHIFT
OFF TRACK and DURING MINOR DEFECT**

	(DFout slope divided by 0.7*0.75) @ ANY TIME	@ $\times 10^{10}$ (6.5*sigma)
	1.35	6.19
	3.43	3.43
	0.8	0.8
	1 -> 3	1 -> 3
	0.5 -> 2	0.5 -> 2
	0.19 -> ?	0.19 -> ?
	2.14	2.14
	<hr/>	<hr/>
	9.41	14.25
ERROR RATE	10^{10}	10
MARGIN	1.01	0.

F. J. SORDELLO 1988

i i s t

A. Hoagland 3/90

Visualizing the shifts in detection that occur due to noise and Peak Shift as a *Statistical Distribution* within the Data Window is the first step in the modeling process. As shown in Figure 2, the perspective gained by this model yields an immediate observation, *the error rate of a drive is the total area of the distribution falling outside the Data Window.*

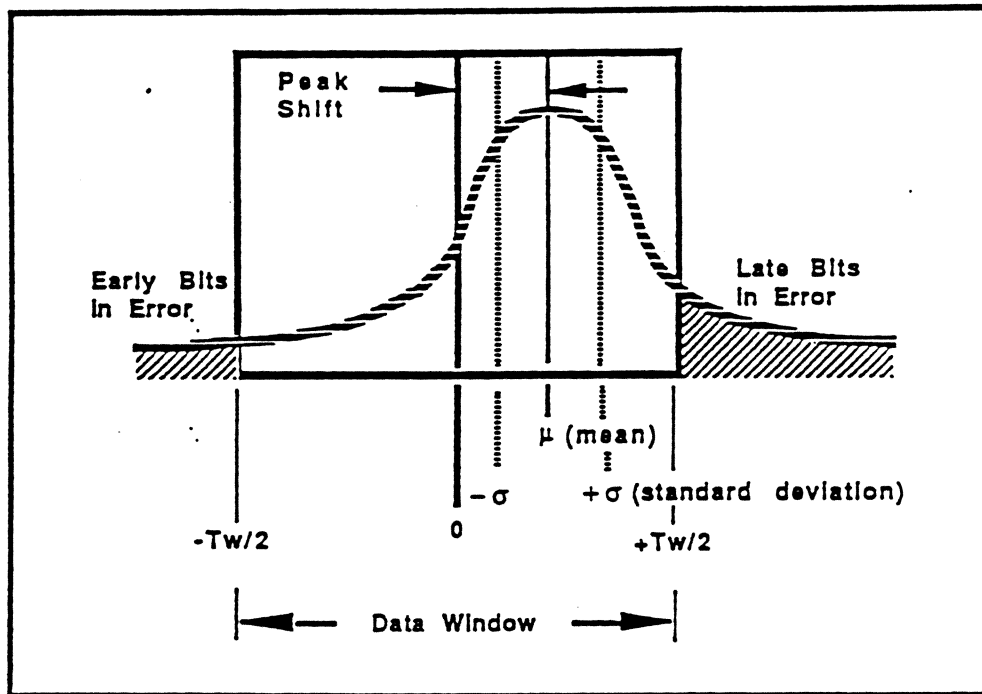


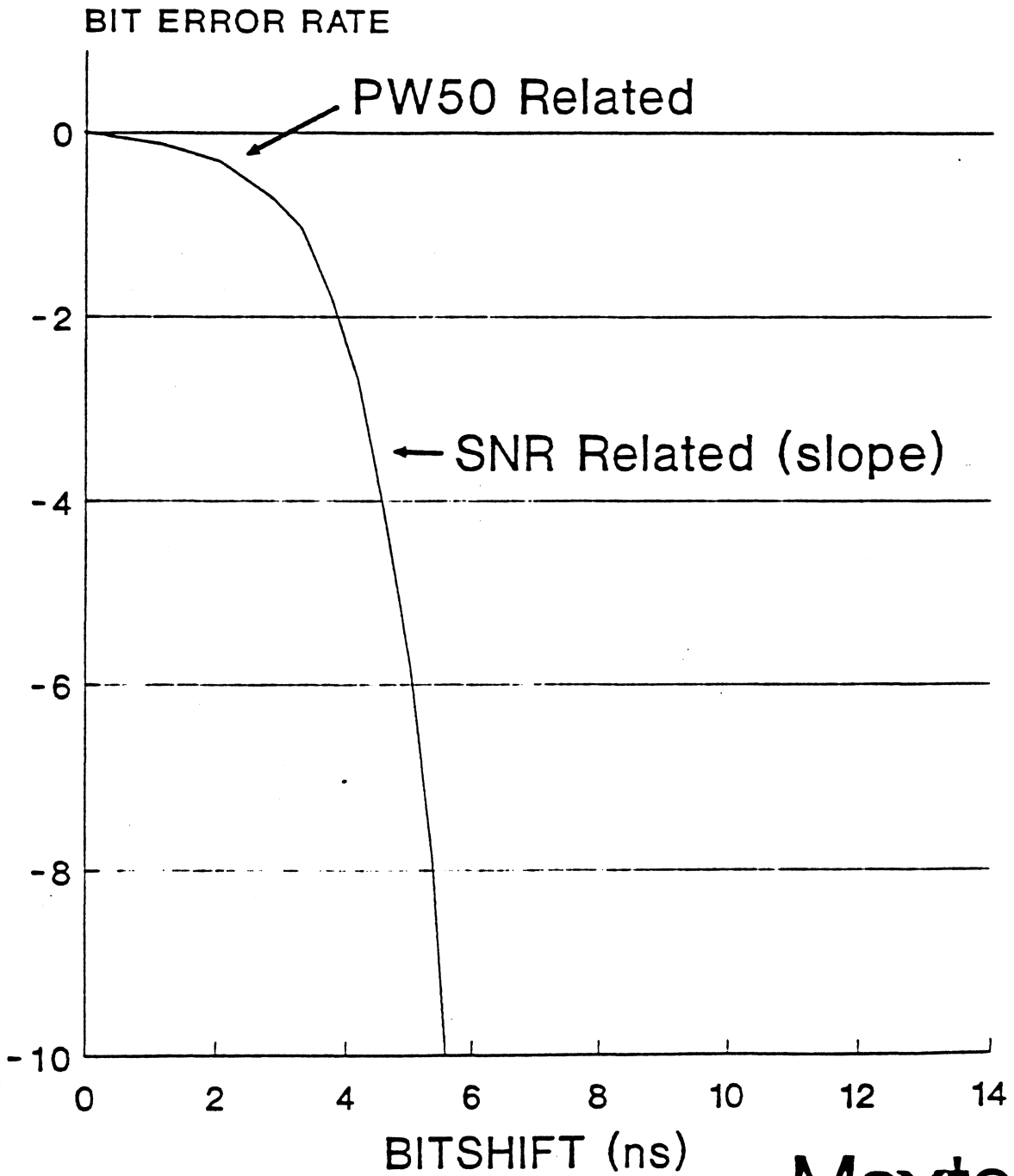
FIGURE 2

STATISTICAL DISTRIBUTION

The next step in the modeling process is the choice of statistics. Assuming that the shifts in peak detection conform to a *Gaussian* model, makes the task of comparing noise and Peak Shift to the Data Window simple. According to the Gaussian model, the area outside the Data Window is directly related to the ratio of Data window width to the Mean (μ) and Standard Deviation (σ) of the distribution. This is referred to as the *intrinsic error rate* † of the system.

† Katz and Campbell, "Effect of Bitshift on Error Rate in Magnetic Recording", IEEE Transactions on Magnetics, Vol. Mag. 15, No. 3, pp 1050-1053, May 1979

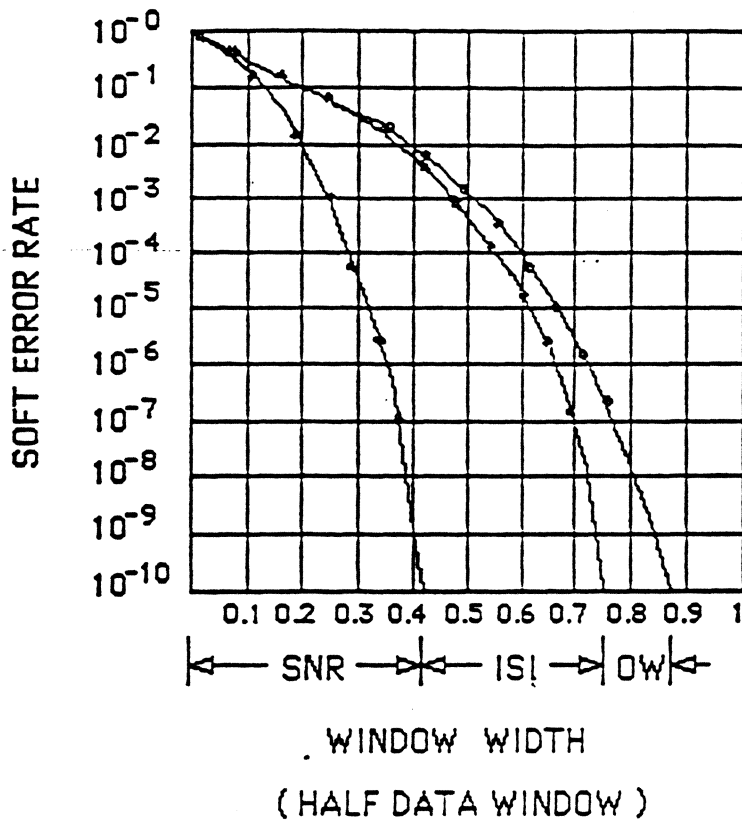
Bitshift Basics

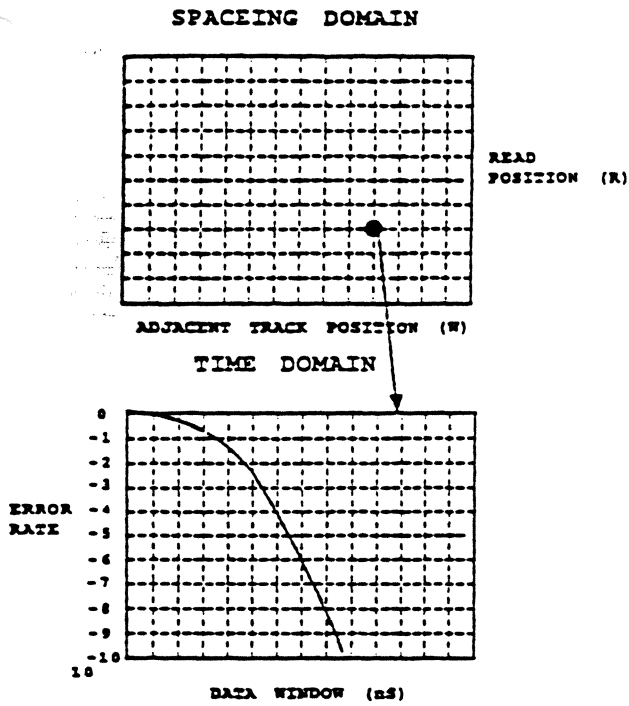


Maxtor

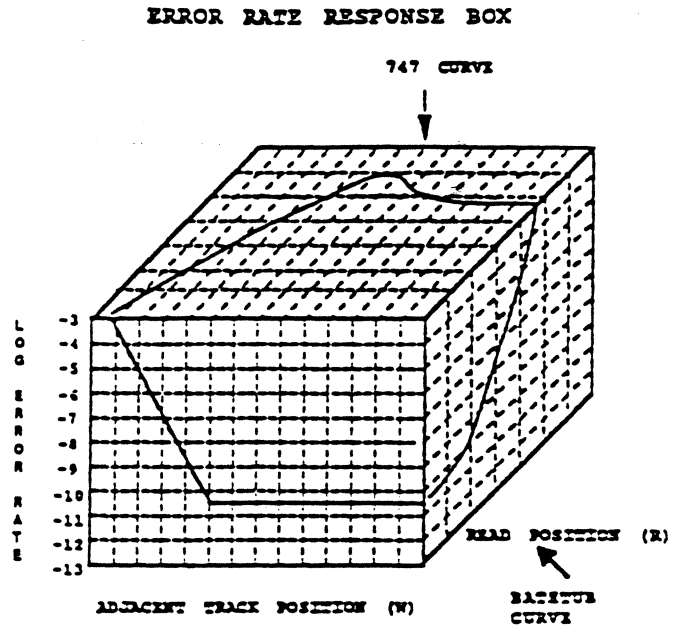
IIST

A. Hoagland 3/90





(a)



(b)

Figure 1.

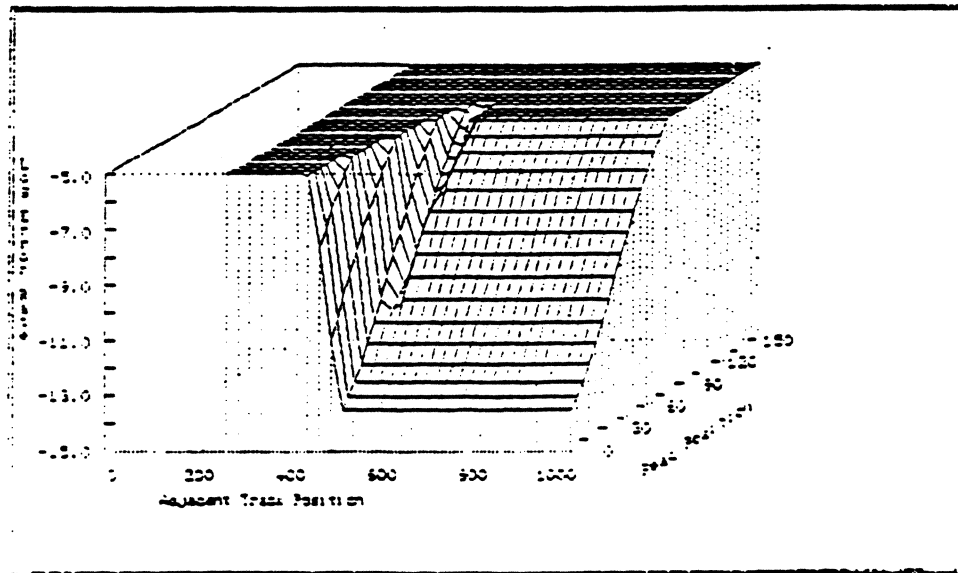


Figure 2.

PRML DETECTION

(PARTIAL RESPONSE CHANNEL, MAXIMUM LIKELYHOOD DETECTION)

CONCEPTS

- ✱ SAMPLE THE SIGNAL WAVEFORM ONCE PER BIT PERIOD
- ✱ DELIBERATELY ALLOW PULSES TO INTERFERE
- ✱ EXAMINE SEQUENCE OF SAMPLES FOR DECISION

PRINCIPLE

Match waveform against a library of all possible waveforms and select the one that gives the minimum sum of squares difference between the sample and the stored waveforms.

Application of Probabilistic Decoding to Digital Magnetic Recording Systems

Abstract: A digital magnetic recording system is viewed in this paper as a linear system that inherently includes a correlative level encoder. This encoder can be regarded as a linear finite-state machine like a convolutional encoder. The maximum likelihood decoding method recently devised by Viterbi to decode convolutional codes is then applied to digital magnetic recording systems. The decoding algorithm and its implementation are discussed in detail.

Expressions for the decoding error probability are obtained and confirmed by computer simulations. It is shown that a significant improvement in the performance with respect to other methods is achievable by the maximum likelihood decoding method. For example, under the Gaussian noise assumption the proposed technique can reduce raw error rates in the 10^{-2} to 10^{-4} range by a factor of 50 to 300. These results indicate that the maximum likelihood decoding method gains as much as 2.5 dB in signal-to-noise ratio over the conventional bit-by-bit detection method.

1. Introduction

In an earlier paper [1] it was shown that a digital magnetic recording channel can be viewed as a partial-response channel. The partial-response signalling or the correlative level coding is a technique recently developed by Lender [2], Kretzmer [3], van Gerwen [4] and by others in data communication systems, in which a controlled amount of intersymbol interference is intentionally introduced to improve the information rate [5]. In a digital magnetic recording system, on the other hand, the differentiation operation inherent in the read-back process generates, in effect, a correlative level coded sequence. Since the representation of a digital magnetic recording channel in terms of its equivalent partial-response channel is essential to the development of the present paper, a brief review of some earlier results [1, 6, 7] is given.

In the ordinary digital magnetic recording system saturation recording is performed, i.e., two stable states of magnetization represent binary data to be stored. Let $\{a_k\}$ represent an information sequence of "0" and "1" to be recorded on the magnetic surface. The magnetization pattern $m(t)$ recorded by the NRZ (Non-Return-to-Zero) method is representable as

$$m(t) = \sum_{k=0}^{\infty} (2a_k - 1)u(t - kT) - 1(-t), \quad (1)$$

where $u(t)$ is a rectangular pulse of duration T seconds:

$$u(t) = \begin{cases} 1, & 0 \leq t \leq T \\ 0, & \text{elsewhere,} \end{cases} \quad (2)$$

and $1(t)$ is a unit step function:

$$1(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0. \end{cases} \quad (3)$$

Here the amplitude of $m(t)$ is normalized by its saturation levels, i.e., $+1$ and -1 represent two saturation levels corresponding to "1" and "0" of the sequence $\{a_k\}$. We assume in Eq. (1) that $m(t) = -1$ for $t < 0$, i.e., the magnetic surface has been magnetized to the -1 saturation level before the arrival of data stream $\{a_k\}$.

In the read-back process the relationship between the output voltage $e(t)$ and magnetization pattern $m(t)$ is given by

$$e(t) = \left[\frac{d}{dt} m(t) \right] * h(t), \quad (4)$$

where $*$ means convolution and $h(t)$ represents the magnetic head field distribution characterized by the response due to a unit step function in $m(t)$. Figure 1 illustrates waveforms at various stages in the NRZ recording method.

On substituting Eq. (1) into (4) we obtain

$$\begin{aligned} e(t) &= h(t) * \left[\sum_{k=0}^{\infty} (2a_k - 1) \{ \delta(t - kT) \right. \\ &\quad \left. - \delta(t - kT - T) \} + \delta(t) \right] \\ &= 2 \sum_{k=0}^{\infty} x_k h(t - kT), \end{aligned} \quad (5)$$

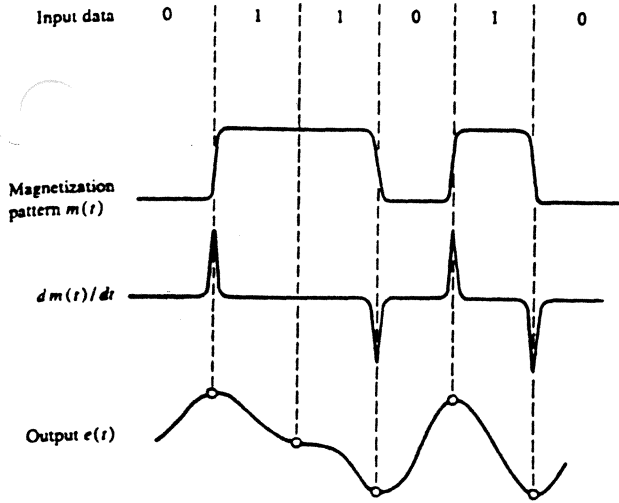


Figure 1 Waveforms at various stages in the NRZ recording system.

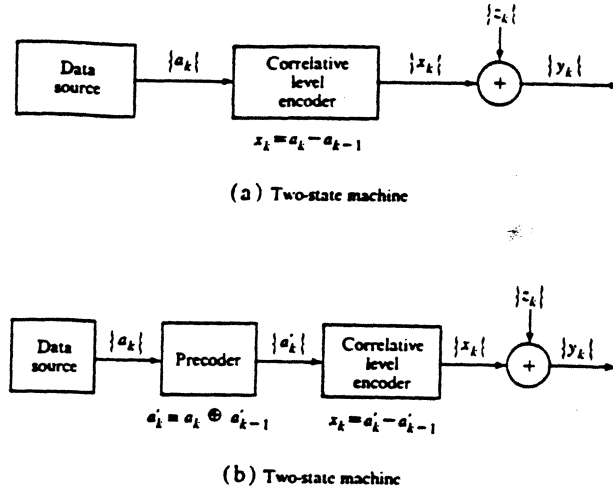


Figure 2 Discrete system representations of (a) the NRZ recording system and (b) the NRZI recording system.

where

$$x_k = \begin{cases} a_k - a_{k-1}, & k \geq 1 \\ a_0, & k = 0. \end{cases} \quad (6)$$

As seen from Eq. (6), the sequence $\{x_k\}$ is a three-level sequence of -1 's, 0 's and $+1$'s. Unlike the situation in data communication systems, the sequence $\{x_k\}$ per se is not generated nor clearly observed in any part of the recording system. What we actually observe is $e(t)$, a linear function of the sequence $\{x_k\}$ as shown in Eq. (5). In other words, we consider for analytical convenience that the magnetic recording channel contains some imaginary correlative level encoder as a part of the system. If $\{a_k\}$ takes on "1" and "0" equally likely and is independent from bit-to-bit, the sequence $\{x_k\}$ possesses the following statistical properties:

$$\Pr \{x_k = 0\} = \frac{1}{2}, \Pr \{x_k = -1\} = \Pr \{x_k = +1\} = \frac{1}{4} \quad (7)$$

and

$$E[x_k x_l] = \begin{cases} \frac{1}{2}, & k = l \\ -\frac{1}{4}, & |k - l| = 1 \\ 0, & \text{elsewhere.} \end{cases} \quad (8)$$

Equation (8) shows that adjacent digits are highly correlated and hence $\{x_k\}$ is called a correlative level coded sequence. In other words $\{x_k\}$ is a sequence that contains redundancy.

Let $e(t)$ be passed into a linear filter $f(t)$, the output of which is denoted by $r(t)$:

$$r(t) = e(t) * f(t). \quad (9)$$

If the total response function $g(t) = h(t) * f(t)$ satisfies the condition

$$g(kT) = \delta_{k,0}, \quad k = 0, \pm 1, \pm 2, \dots, \quad (10)$$

then the sampled value of the filter output is

$$r(kT) = 2x_k. \quad (11)$$

Equation (10) is satisfied if the filter $f(t)$ includes an equalizer so that the effect of intersymbol interference is removed. However, the sampled voltage cannot be exactly equal to $2x_k$ because of the noise and the residue of intersymbol interference. Therefore, what we actually observe at a sampling instant is represented by the following random variable y_k :

$$y_k = x_k + z_k, \quad (12)$$

where z_k represents the total disturbance.

From Eqs. (6) and (12) we obtain the block diagram of Fig. 2(a), which is a linear discrete system representation of a magnetic recording system. Here $\{a_k\}$ is a sequence of "1" and "0", $\{x_k\}$ is a sequence of -1 's, 0 's and $+1$'s, whereas $\{y_k\}$ is a random sequence that may take on any real number. In earlier papers [1, 6, 7] we described a decision scheme that quantizes $\{y_k\}$ into a three-level sequence $\{q_k\}$. The data sequence $\{a_k\}$ can be estimated on the basis of this "hard" decision output $\{q_k\}$ by solving Eq. (6). However, an erroneous decision in $\{q_k\}$ would result in the propagation of error in decoding the data sequence $\{a_k\}$. To avoid such error propagation, $\{a_k\}$ is transformed into another binary sequence $\{a'_k\}$ by the following relation before being passed into a correlative level encoder:

$$a'_k = a_k \oplus a'_{k-1} \text{ mod } 2 \quad (13a)$$

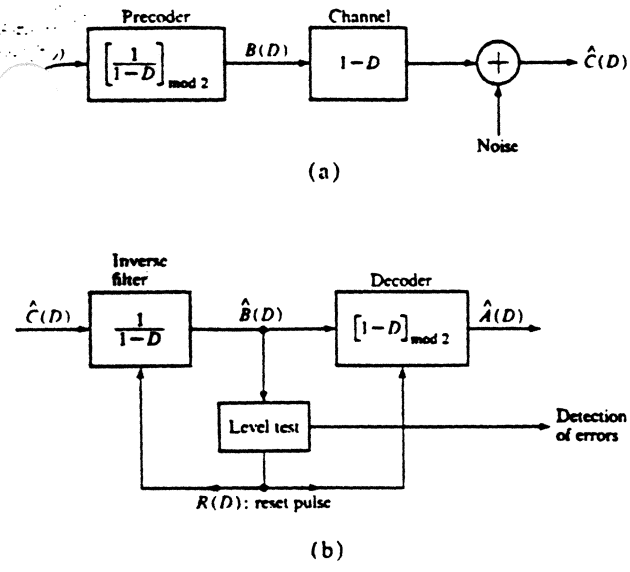


Figure 3 (a) A representation of the NRZI recording system as a partial-response system; (b) error detection and decoding method in the NRZI recording system.

where

$$c_k = \begin{cases} b_k - b_{k-1}, & k \geq 1 \\ b_0, & k = 0. \end{cases} \quad (9)$$

The response function $h(t)$ virtually satisfies the following condition at sampling instants,

$$h(nT) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0, \end{cases} \quad (10)$$

then the sampled value of the output voltage is

$$e(nT) = 2c_n, \quad (11)$$

which is a three-level sequence, as is clear from Eq. (9). Thus, the recording channel can be regarded as a partial-response channel with a discrete transfer function

$$G_0(D) = 1 - D. \quad (12)$$

From our earlier discussion on partial-response channels, the corresponding precoder here should perform a transformation that is "linear" in the binary sense, and has a transfer function $[1/(1 - D)]_{\text{mod } 2}$. More specifically, the precoder input $\{a_k\}$ is related to its output $\{b_k\}$ by

$$\begin{aligned} b_k &\equiv a_k + b_{k-1} \pmod{2} \\ &\equiv \sum_{i=0}^k a_i \pmod{2}, \end{aligned} \quad (13)$$

with usually assumed to be zero.

It is easy to see that the precoder obtained above by viewing the magnetic channel as a partial-response channel performs exactly the function of an NRZI encoder [Fig. 3(a)]. That is, a symbol in the precoded binary sequence $\{b_k\}$ is always reversed from the preceding symbol when a ONE in the binary input sequence $\{a_k\}$ is to be recorded. This new observation, nevertheless, enables one to employ the general detection method developed for partial-response channels and to take full advantage of the inherent redundancy in the three-level channel output.

A possible structure of the error detection method based on the theorem in Section 2 is diagrammatically shown by Fig. 3(b). Here the rectifier or "mod 2 detector" in the conventional system is replaced by the combination of the inverse filter with transfer function $(1 - D)^{-1}$ and the decoder with transfer function $[1 - D]_{\text{mod } 2}$.

It should be pointed out that, in this special case, the error criterion is equivalent to the observation that voltage pulses read from a saturation recording system must alternate in polarity. However, it is reassuring to know that all detectable errors can be detected this way. The special form used to implement the error detection logic may also be of some practical interest.

4. A new high-density recording scheme

The function $h(t)$ of Eq. (1) is the output voltage response to an impulse in $d[m(t)]/dt$; i.e., to a step change in the direction of recording medium saturation. It is known that a Gaussian characteristic, $h(t) = a \exp(-bt^2)$, can generally provide a good fit to experimentally obtained characteristic voltage pulses.¹ It is seen from Fig. 4(a) that although the pulse shape $h(t)$ has no overshoot, it is not suitable for high-density recording, since a bit interval T of a fairly large value must be chosen to avoid excessive intersymbol interference.

The pulse in Fig. 4(c) is the same as that of Fig. 4(b). However, the sampling rate is increased by 50 percent; i.e., sampling is done at every T' second, where $T' = \frac{2}{3}T$. If we define the new time axis t' by

$$t' = t + \frac{1}{2}T' \quad (14)$$

and a function $f(t')$ by

$$f(t') = h(t + \frac{1}{2}T'), \quad (15)$$

then, as can be seen from Fig. 4(c)

$$f(nT') = \begin{cases} 1, & n = 0, 1 \\ \approx 0, & n \neq 0, 1, \end{cases} \quad (16)$$

where the values at $n = 0, 1$ are normalized. Hence, the transfer function is characterized by

$$F(D) = \sum_{n=-\infty}^{\infty} f(nT')D^n = 1 + D. \quad (17)$$

impulse response to dM/dt

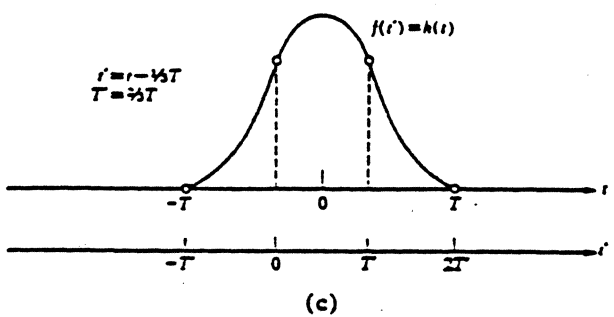
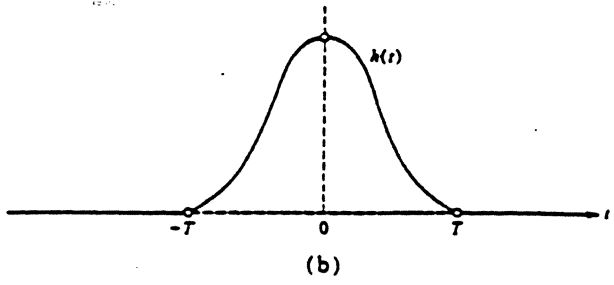
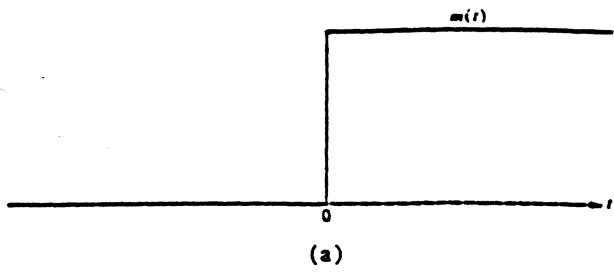


Figure 4 The principle of a higher-density recording system.

Figure 5 (a) Interleaved NRZI recording method; (b) error detection capability added to the system.

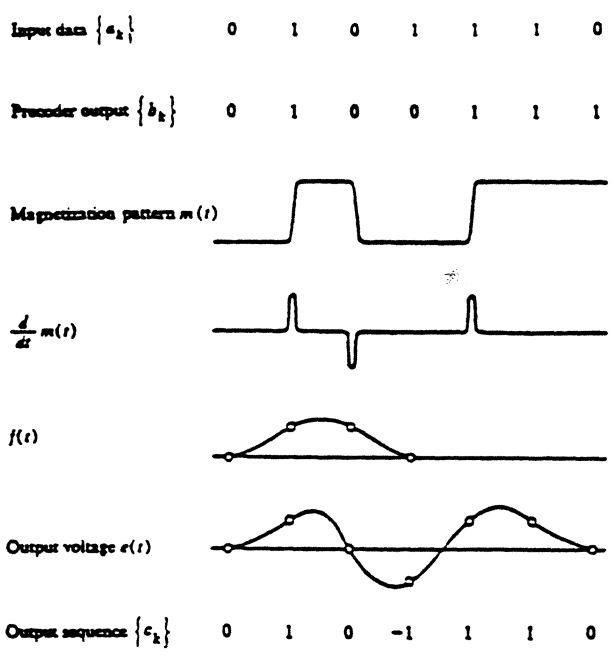
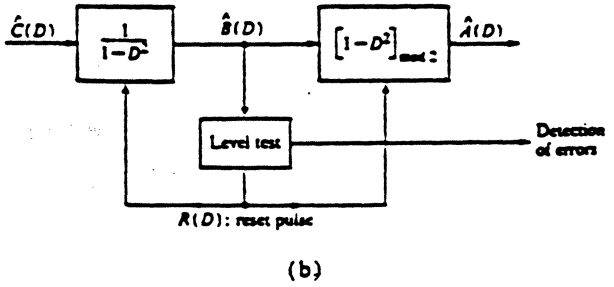
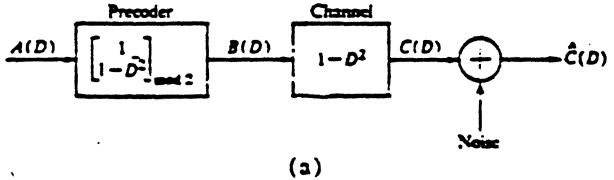


Figure 6 Waveforms at various stages in the Interleaved NRZI recording system.

Then the transfer function of the whole system is, from Eqs. (12) and (17),

$$G_1(D) = G_0(D) F(D) = 1 - D^2. \quad (18)$$

The corresponding precoder has a transfer function $[1/(1 - D^2)]_{\text{mod } 2}$. The input sequence $\{a_k\}$ and output sequence $\{b_k\}$ are related by

$$b_k = a_k + b_{k-2} \pmod{2}. \quad (19)$$

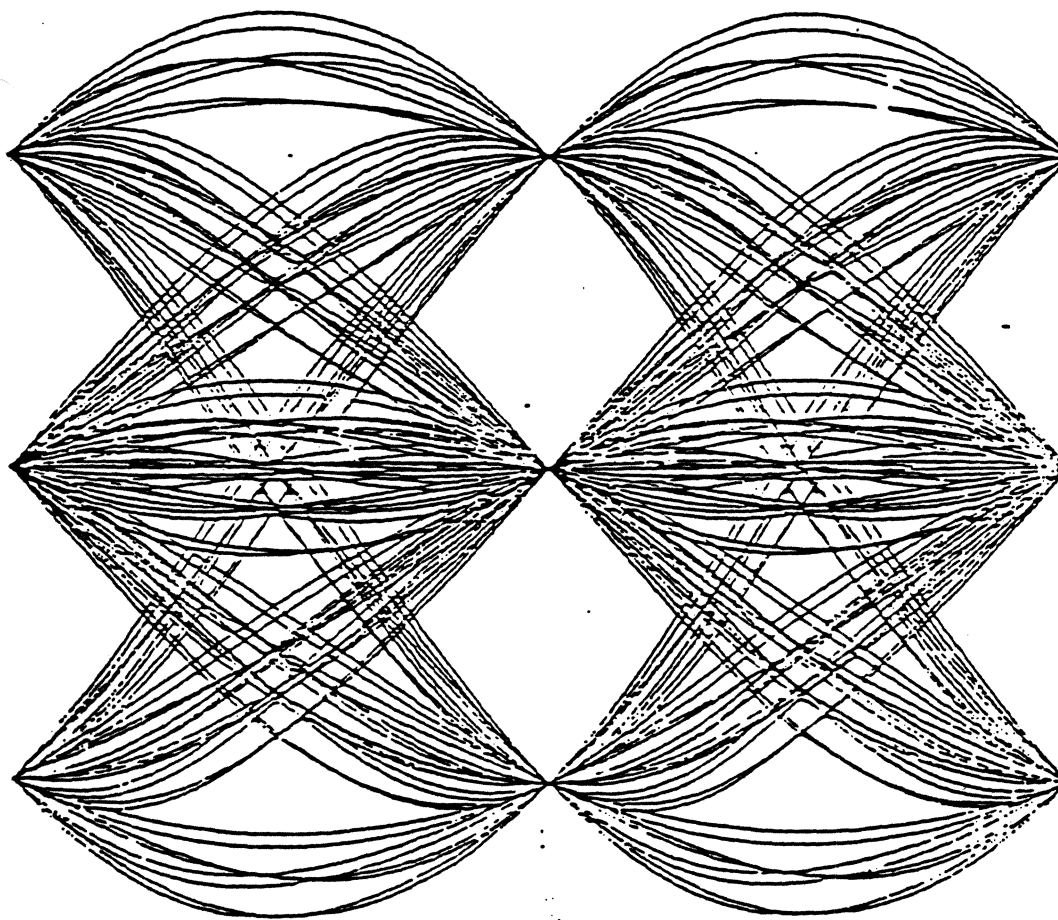
We shall call this new system the "Interleaved NRZI" recording system. The block diagram of the system is shown in Fig. 5. Fig. 6 shows the actual waveforms observed at various stages in the Interleaved NRZI system.

The Interleaved NRZI system described above requires the channel response to satisfy Eq. (16). In practice, however, such desired channel response can not always be maintained (especially if the sampling rate is pushed higher), and, as a result, system performance may deteriorate. A common remedy is the introduction of "shaping" or "equalization" in the frequency characteristic of the channel. A possible frequency characteristic $H(\omega)$ [the Fourier transform of the channel response function $h(t)$] assumes a cosine shape in the magnitude, as shown in Fig. 7, while possessing a linear phase characteristic. Equalization or channel shaping can also be realized in the time domain. Automatic or adaptive equalization developed in data transmission systems^{1,9} can thus be applied to the magnetic recording channel. However, this subject is beyond the scope of the present paper.

EYE DIAGRAM

F0M

CLASS IV PARTIAL RESPONSE



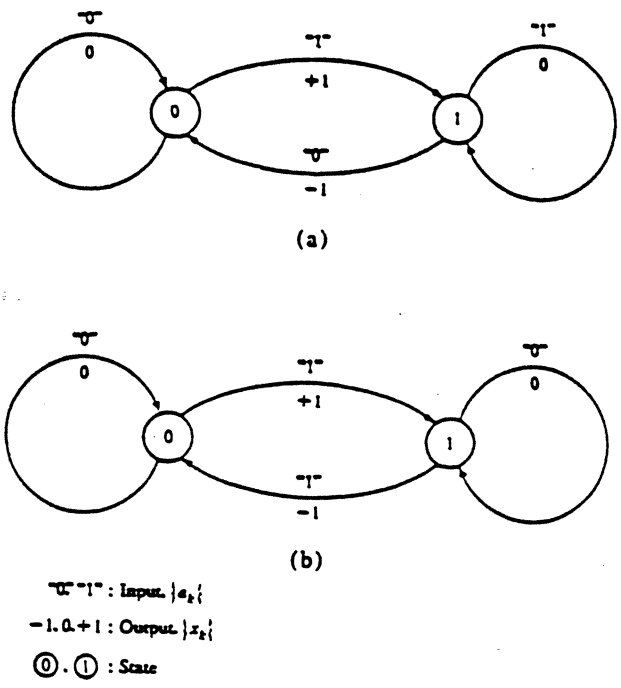
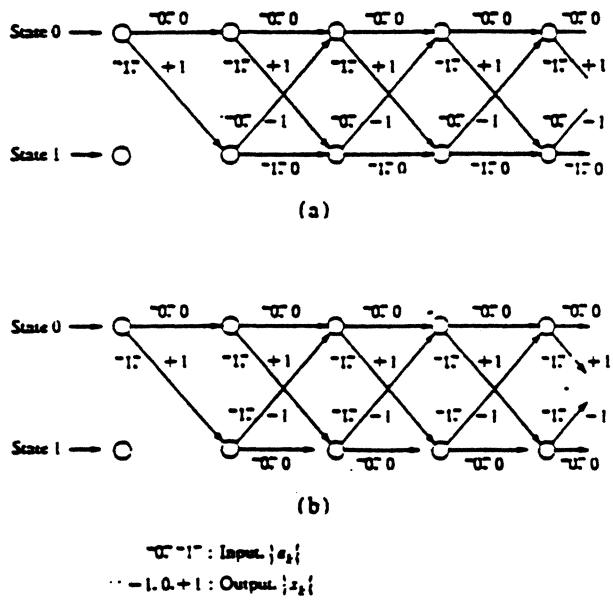


Figure 3 State transition diagrams of (a) the NRZ recording system and (b) the NRZI recording system.

Figure 4 Trellis picture representations of the state transition in (a) the NRZ recording system and (b) the NRZI recording system.



and $a'_{-1} = 0$. (13b)

This transformation is usually called precoding in data communication systems [2]. It has been shown [1] that the so-called NRZI (Non-Return-to-Zero-Inverse) recording method is equivalent to a precoding operation followed

by the NRZ recording method. Figure 2(b) is an equivalent discrete system representation of the NRZI system.

An algebraic method of error detection proposed in References 1, 6 and 7 makes full use of the inherent redundancy of the three-level sequence $\{x_k\}$. This algebraic approach has been further extended to the case in which the receiver makes a "soft" decision, i.e., the number of quantization levels is increased from three to five or seven including ambiguity levels [6, 7].

The present paper describes a completely different approach to decoding the magnetic recording output. This decoding method is a very simple scheme to realize the maximum likelihood decoding (MLD) rule and is a probabilistic decoding scheme rather than the algebraic one discussed earlier. It will be shown that a significant improvement in the performance is obtainable by the proposed decoding scheme.

2. Maximum likelihood decoding

In 1967 Viterbi [8] devised a new nonsequential decoding algorithm for convolutional codes. Forney [9] showed that this algorithm is in fact the maximum likelihood decoding rule. Omura [10] discussed the algorithm in a state-space context and showed its equivalence to the dynamic programming.

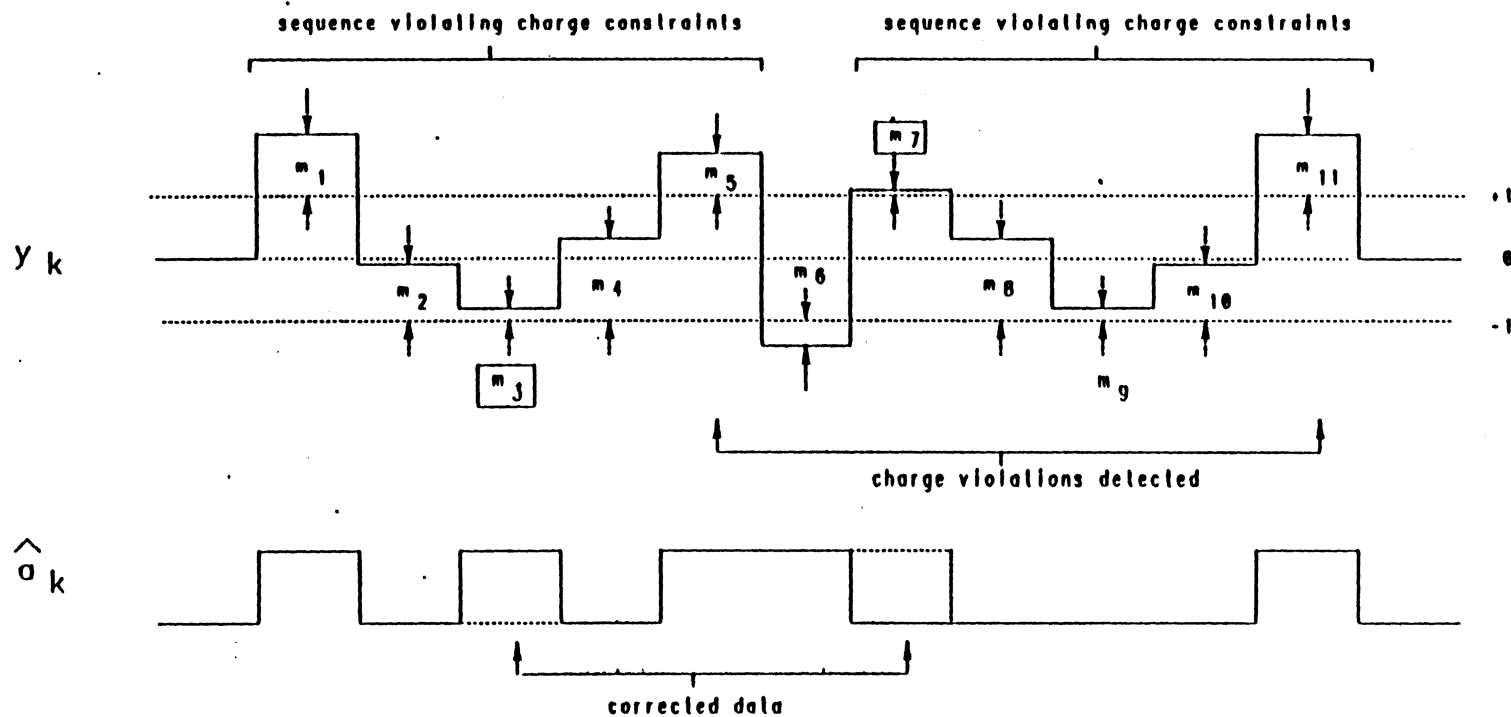
A correlative level encoder can be viewed as a simple type of linear finite-state machine over the real number field as opposed to a Galois field over which a convolutional encoder is defined [11]. Now that we know the equivalence between a magnetic recording channel and a correlative level encoder, it is not difficult to show that the Viterbi decoding rule is applicable to our problem.

We define s_k the state of the imaginary correlative level encoder by the latest encoder input, i.e., $s_k = a_k$ in the NRZ recording system and $s_k = a'_k$ in the NRZI recording system. A precoder defined by Eq. (13) is also a two-state machine; hence we can combine the precoder and correlative encoder in the state representation of the NRZI system. Figures 3(a) and (b) show the state transition diagrams of the NRZ and NRZI recording systems, respectively, where 1 and 0 in small circles represent two possible states. Each time the machine receives a new bit, "1" or "0", a state transition takes place depending on the input and the current state. Numbers -1, 0 or +1 attached to arrows represent the encoder output $\{x_k\}$. Although the diagram of Fig. 3 completely describes our system, the description in terms of the trellis picture introduced by Forney [9] will provide a better understanding of the decoding rule to be discussed.

The trellis picture of Fig. 4 shows the transition of the encoder state as a function of time t . Here the input "1" or "0" and the corresponding output -1, 0 or +1 are attached to each branch connecting two states. Starting from $s_0 = 0$ the encoder follows a particular path according

A SUBOPTIMUM DICODE DECODER

1. Estimate data by threshold decoding
2. On detection of a charge violation, invert the bit with the least detection margin in the direction which removes the violation.



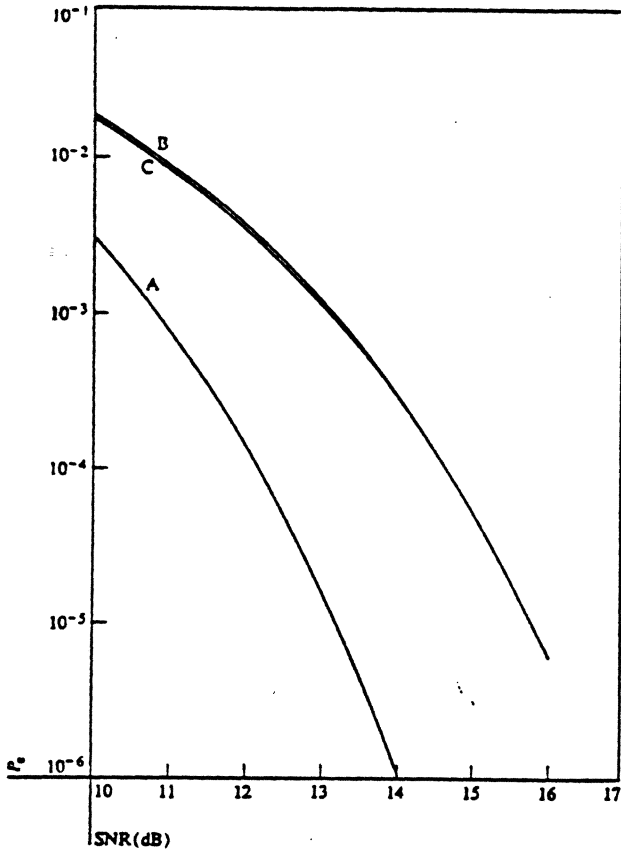


Figure 7 Probability of error vs SNR: (A) the maximum likelihood decoding. (B) the bit-by-bit detection with thresholds at $-A/2$ and $A/2$. (C) the bit-by-bit detection with the optimum thresholds.

the channel symbol x_i takes on A , 0 , and $-A$ with probabilities, $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{1}{4}$, respectively. The optimum thresholds are t_{opt} and $-t_{opt}$, where t_{opt} is the solution of the equation

$$\frac{1}{2}\phi(t_{opt}/\sigma) = \frac{1}{4}\phi[(t_{opt} - A)/\sigma], \quad (40)$$

which yields

$$t_{opt} = \frac{A}{2} + \frac{\sigma^2}{A} \ln 2. \quad (41)$$

Thus the bit error rate with these optimum threshold values is given by

$$P_{e, BIT} = \Phi\left(-\frac{d}{\sqrt{2}} - \frac{\sqrt{2}}{d} \ln 2\right) + \frac{1}{2} \Phi\left(-\frac{d}{\sqrt{2}} + \frac{\sqrt{2}}{d} \ln 2\right). \quad (42)$$

Curves B and C in Fig. 7 are plots of Eqs. (39) and (41), respectively. Their difference is fairly small. We can see, however, a substantial difference between Curves A and C. For example, at SNR = 13 dB $P_{e, BIT} = 1.1 \times 10^{-3}$ whereas $P_{e, MLD} = 1.8 \times 10^{-5}$, i.e., improvement by a factor of 70. The performance improvement is even higher

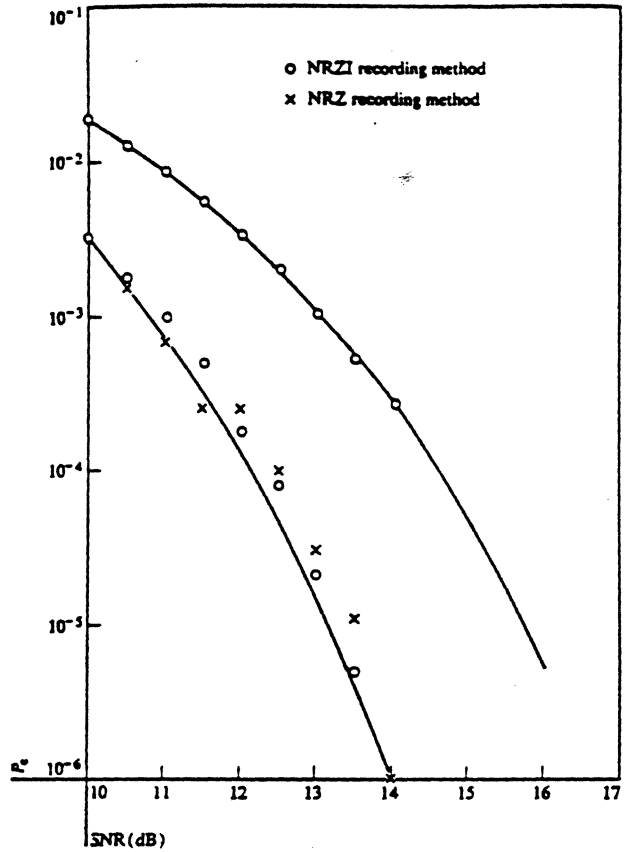


Figure 8 Simulation results of the maximum likelihood decoding method and the bit-by-bit detection method.

for a higher SNR: the decrease in the error probability by a factor of several hundred is possible beyond SNR = 14 dB. In terms of SNR, the maximum likelihood decoding method gains as much as 2.5 dB in the range of raw error rate 10^{-3} to 10^{-4} .

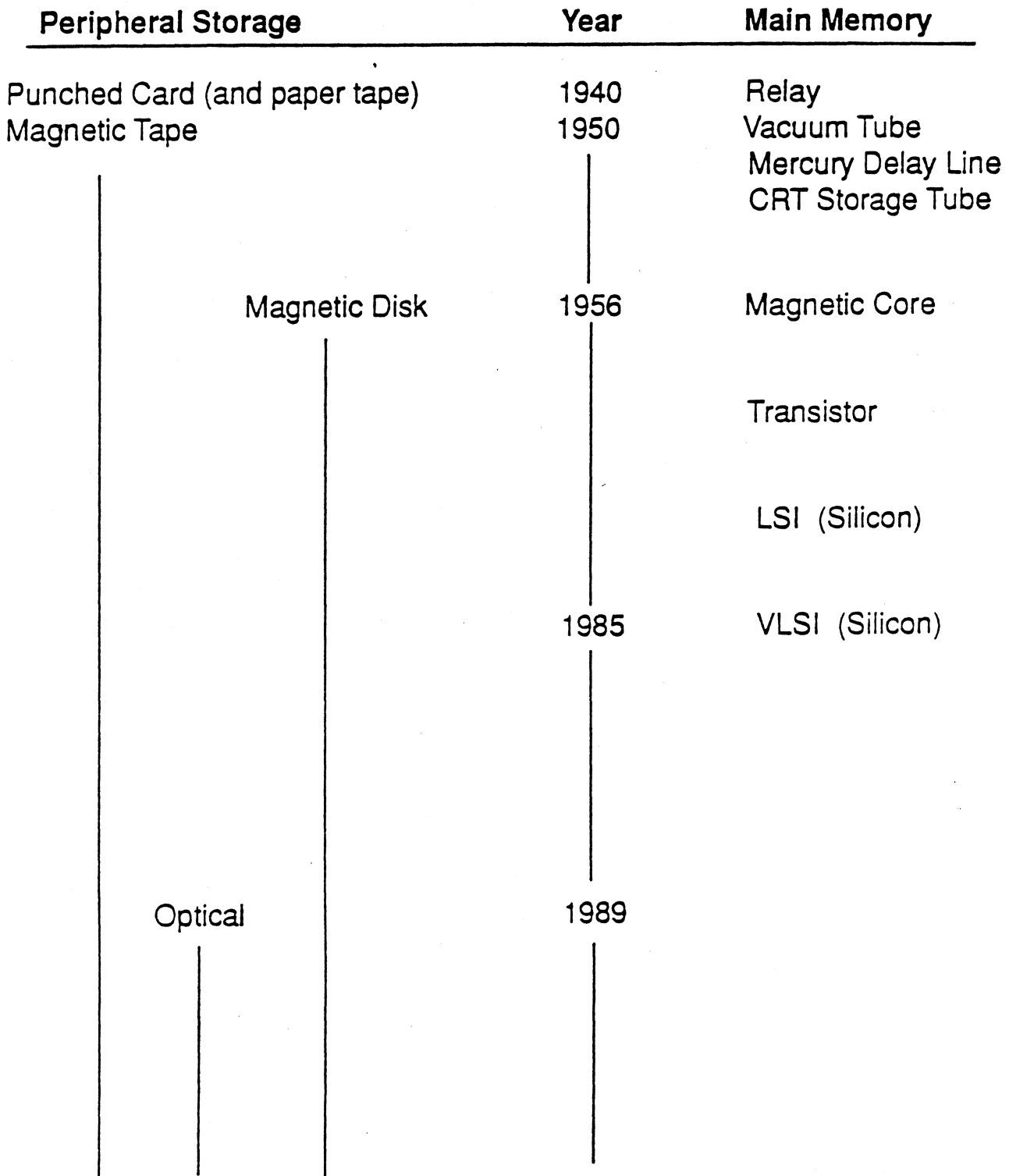
Now we shall report the computer simulation results and confirm the analytical results obtained above. The discrete channel models of Figs. 2(a) and (b) are assumed. The data sequence $\{a_i\}$ was generated through the random number generator program, and the noise sequence $\{z_i\}$ was generated by transforming the random variable with a uniform distribution through a polynomial approximation formula of the mapping $\Phi^{-1}(\cdot)$ [12]. It will be worth mentioning here that most existing subroutine programs under the name "Gaussian Random Generator" are not appropriate to this type of simulation, in which a high accuracy is required at the tail of the probability density distribution.

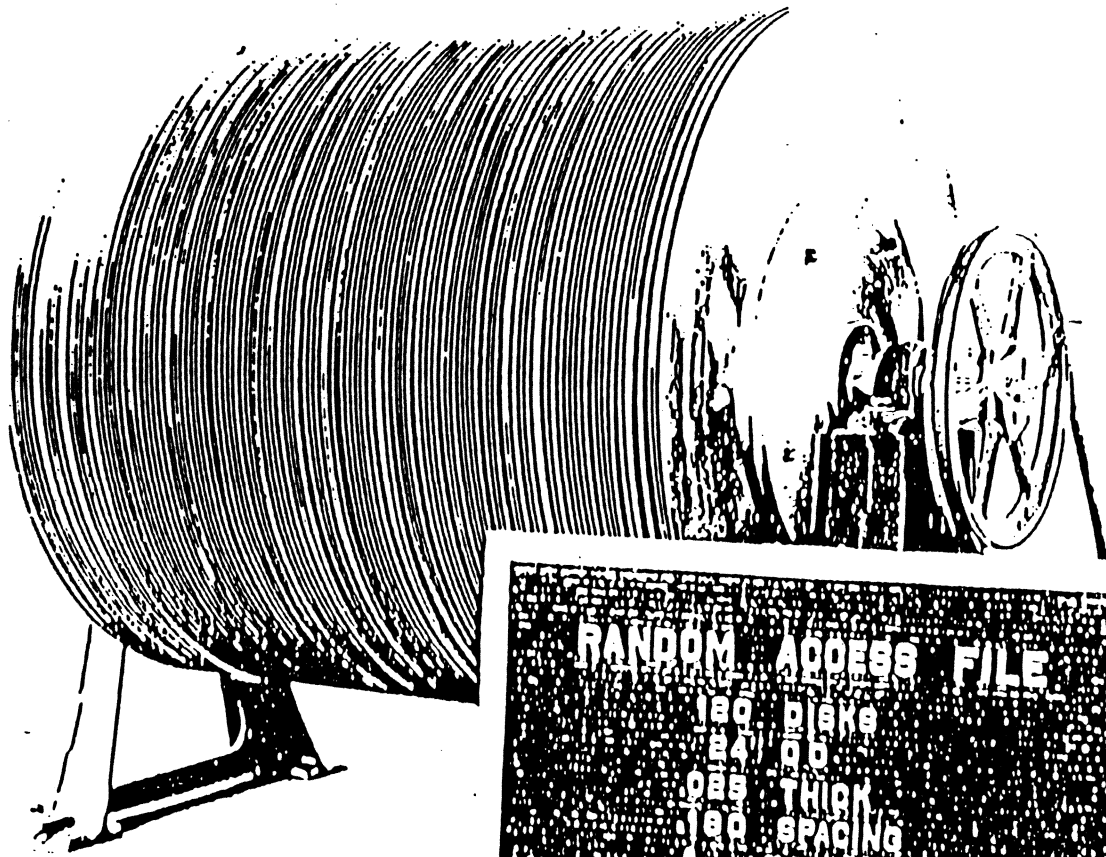
The simulation results are plotted in Fig. 8, where decoding error rates for the NRZ and NRZI recording methods are marked by X and O, respectively. The size N of data is 10^5 for SNR = 10 to 11.5 dB, and 10^6 for

DISK DRIVE DESIGN

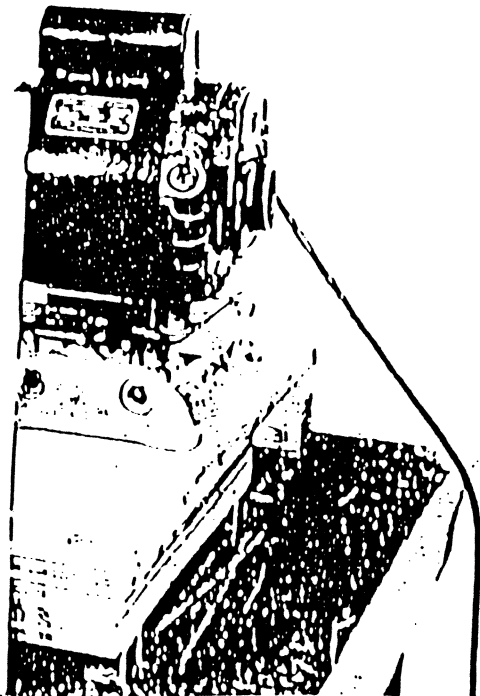
IIST

EVOLUTION of MEMORY and STORAGE TECHNOLOGY





RANDOM ACCESS FILE
180 DISKS
124 DD
088 THICK
180 SPACING
APR 29, 1953
PHOTO NO. 143



IIST

A.Hoanland 390

HEAD-MEDIUM INTERFACE

- How close is contact - or how do you define zero?
(from laws of scaling are relentlessly driven towards zero spacing)

On Tape:

Assume contact and set limit on number of passes

On Disk:

Assume no contact and expect unlimited number of passes

- Wear versus time

Recording System Scaling

Nanometers

100000

10000

1000

100

1965 1970 1975 1980 1985 1990 1995 2000

General Availability

Track Pitch

Head Gap

Head-Disk Spacing
Medium Thickness

Bit Cell

Bit Cell Length

Head Gap

Medium Thickness

Head/Disk Spacing

2314

3330

3340

3360

3370

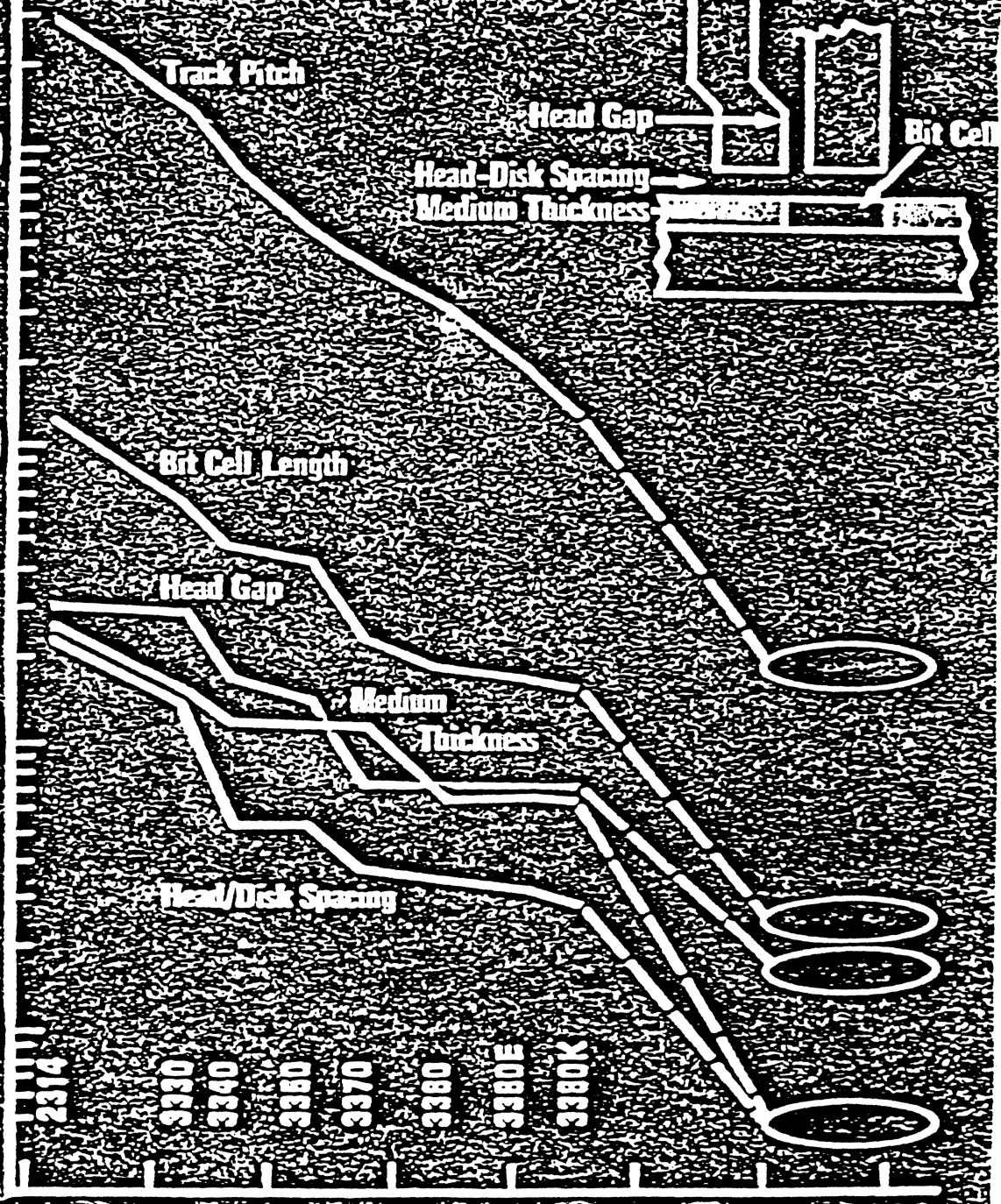
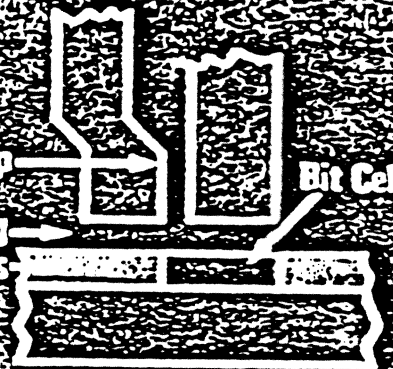
3380

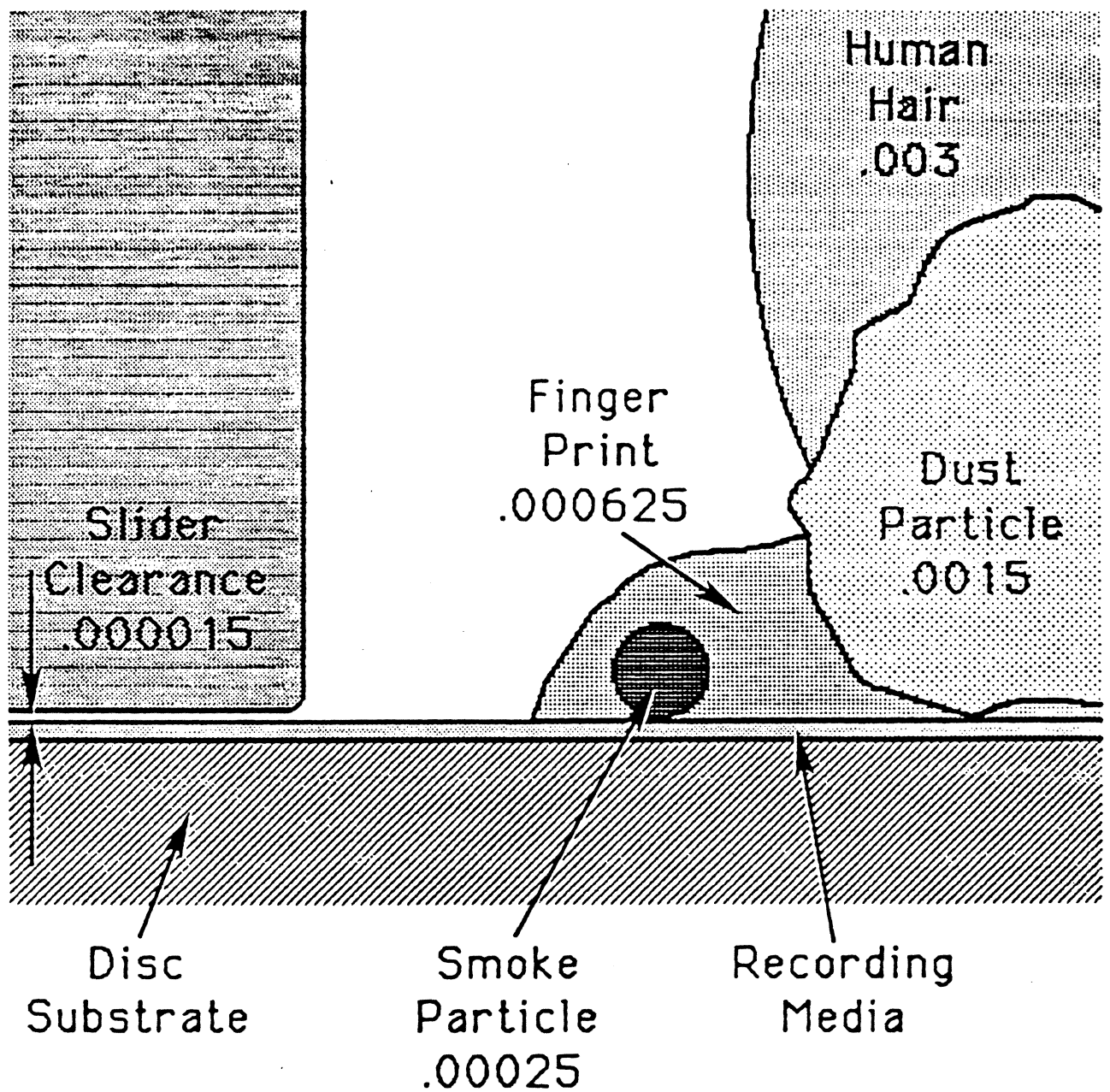
3380E

3380K

IIST

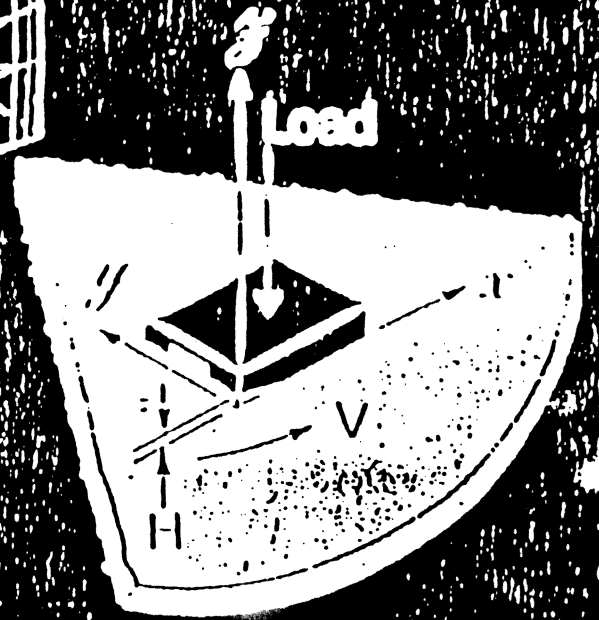
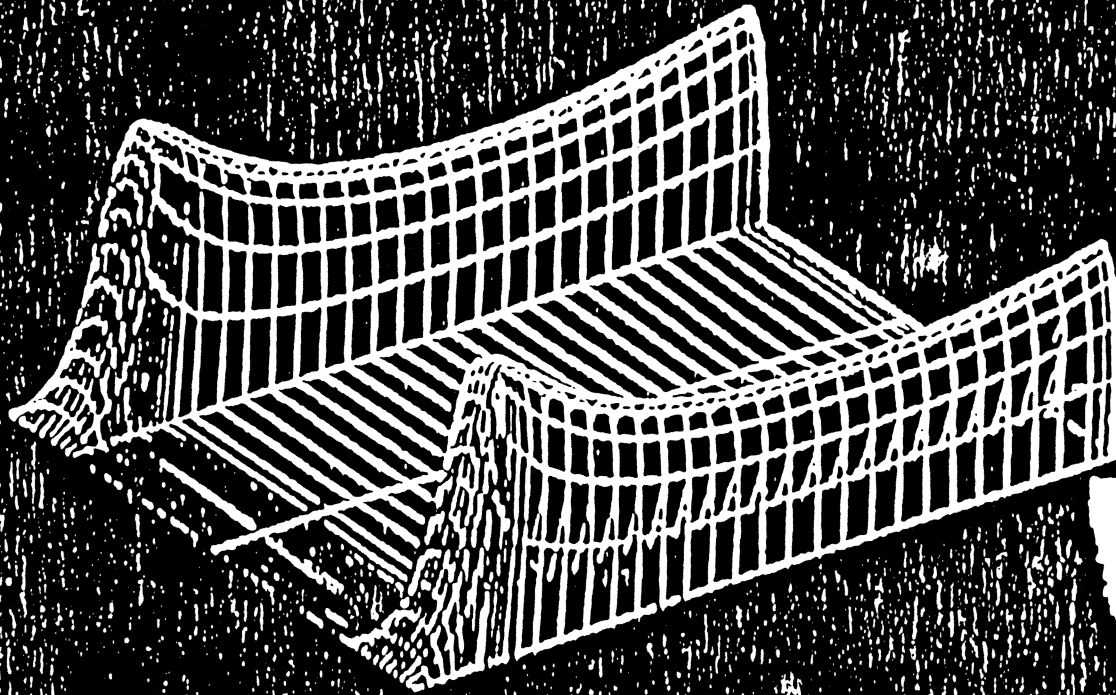
A. Hoagland 3.99





Cleanliness Requirements:
Contamination vs. Flying Height

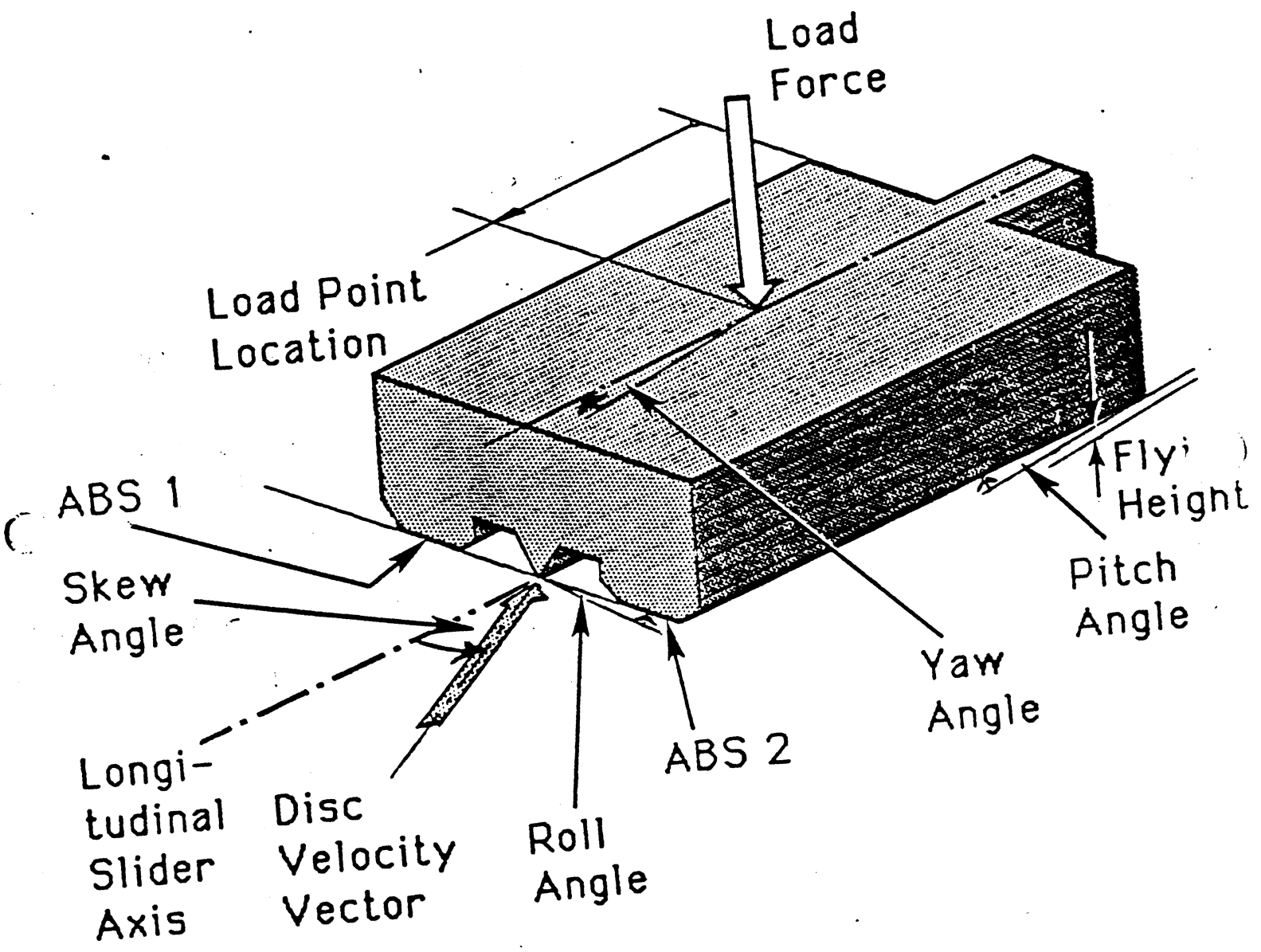
$$\nabla \cdot [H^3 P \nabla P + 6\lambda_a P_a H^2 \nabla P - 6\mu V P H] = 12\mu \frac{\partial(PH)}{\partial t}$$



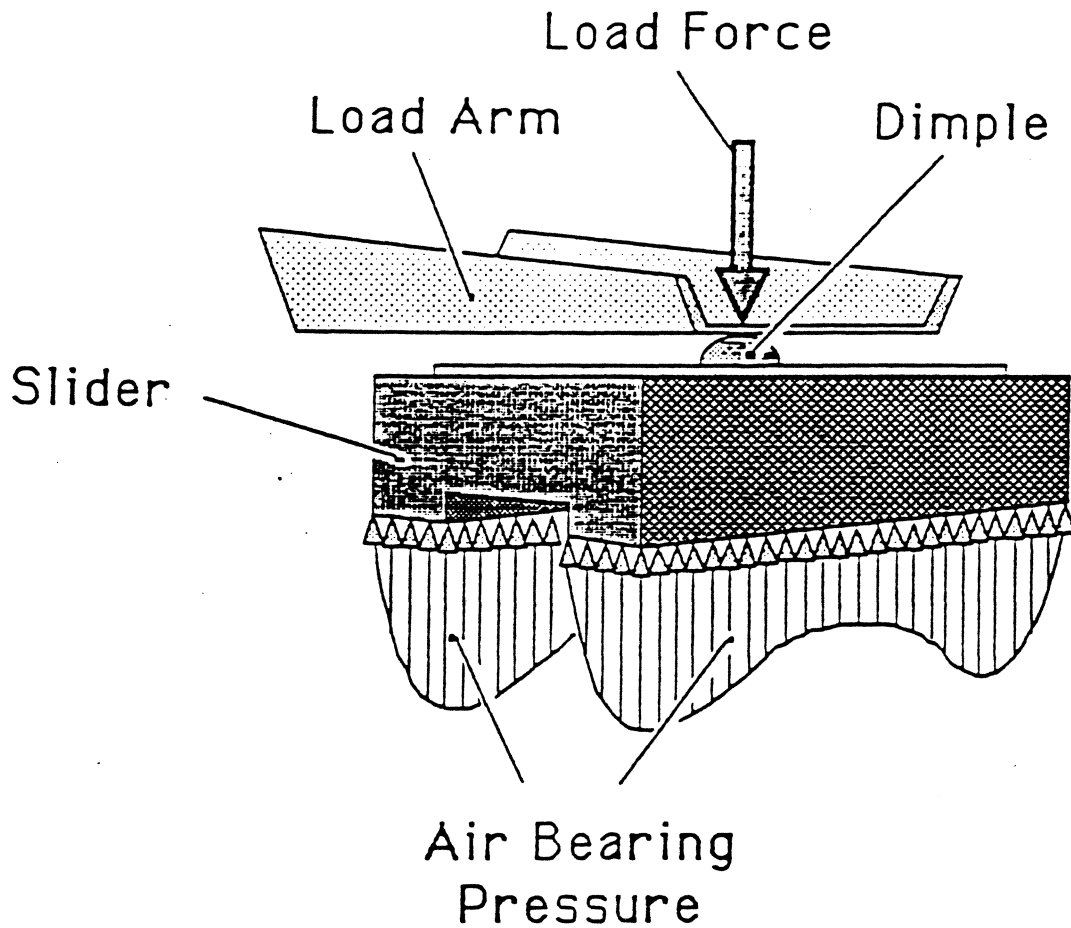
IIST

A. Hoagland 3/90

Flight Attitude - Definitions



3370 Type Pivot



HG Ctk I-Pivot

Saturation Recording

- For Track Density Change

- Signal $\propto \frac{1}{\text{tpi}}$

- Media Noise $\propto \frac{1}{\sqrt{\text{tpi}}}$

or SNR decreases 3db as double tpi

Track Width

- (W = Track width) (s = signal output)
- $s \propto W e^{-\pi d(fci)}$
- $tpi \propto \frac{1}{W}$
- track density *versus* linear density

$$\text{bits/in}^2 \propto (fci)(tpi) \propto (fci) \times s \times e^{\pi d(fci)}$$

holding s constant, maximize bits/in^2 in terms of bpi and tpi

$$\frac{\partial(\text{bits/in}^2)}{\partial fci} = 0 \quad \text{or} \quad fci = \frac{1}{\pi d}$$

(ignoring noise) operate where a further increase in linear density causes a greater signal loss than an increase in track density would.

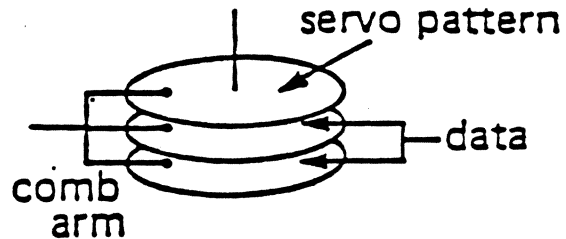
HEAD POSITIONING

Open Loop

low cost

Servo

- dedicated



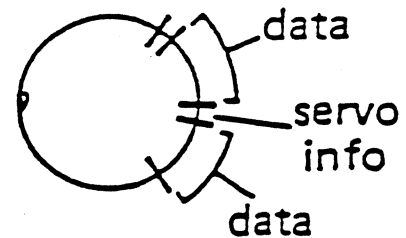
Issue - cascaded mechanical & thermal tolerances

- imbedded

(a) Sector

Issues - open loop when R/W

- real estate
- slew rate
- tolerance to disk defects

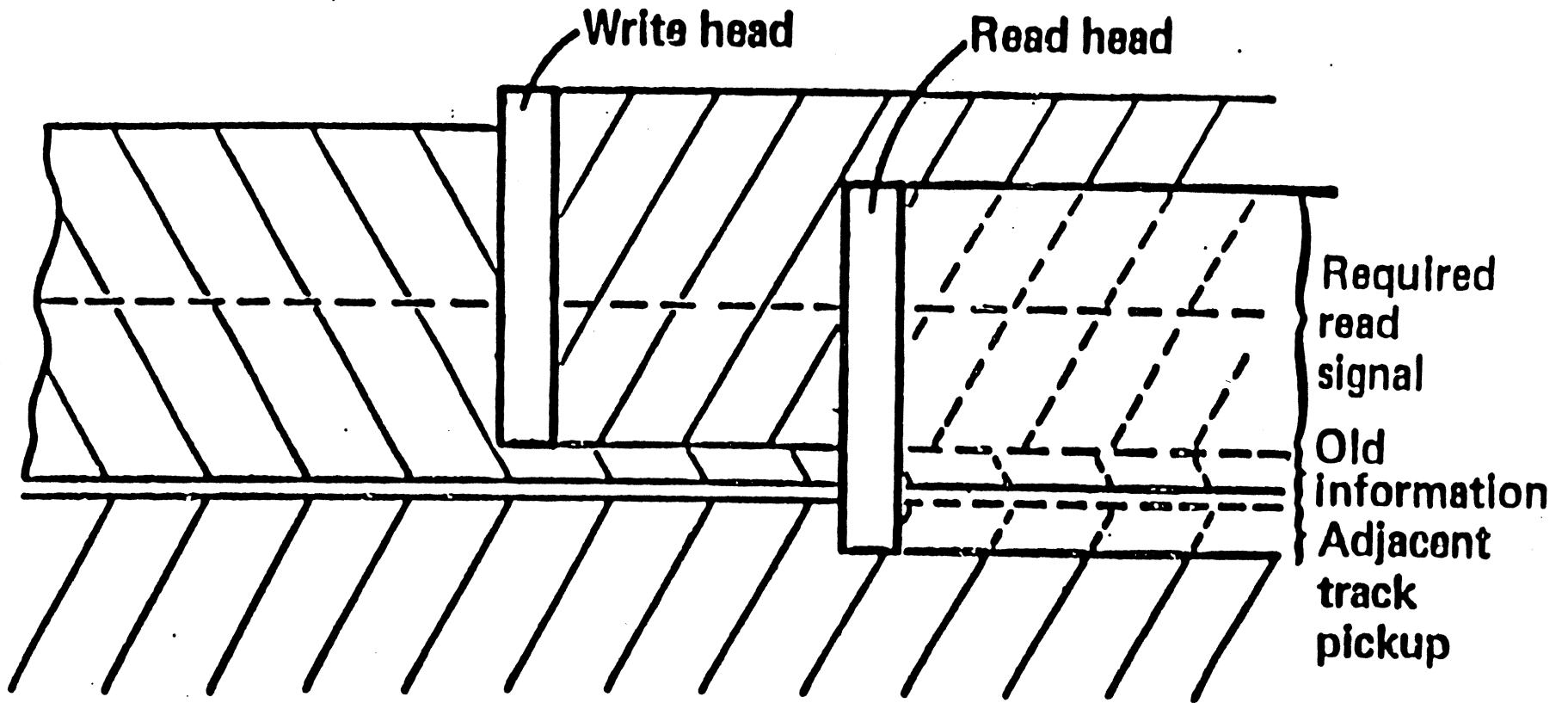


(b) Continuous

Issues - source of PES (position error signal)

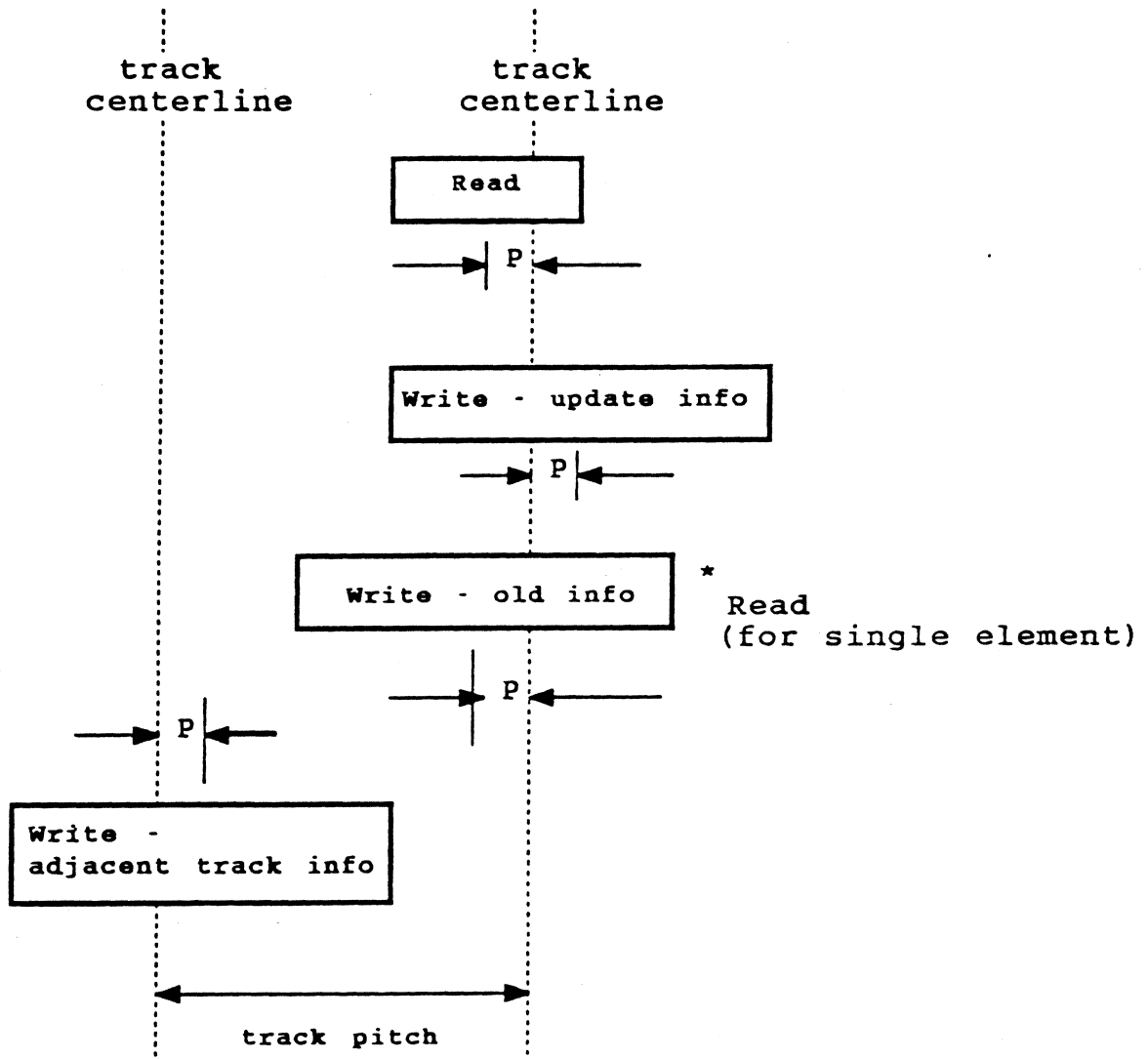
- disk fabrication
- sensor
- servo while write

Read head pickup of old/adjacent track information



Write-wide Read-narrow Head Design

Worst case analysis



W = write head width R = read head width

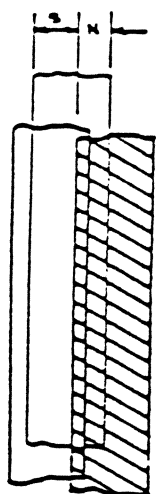
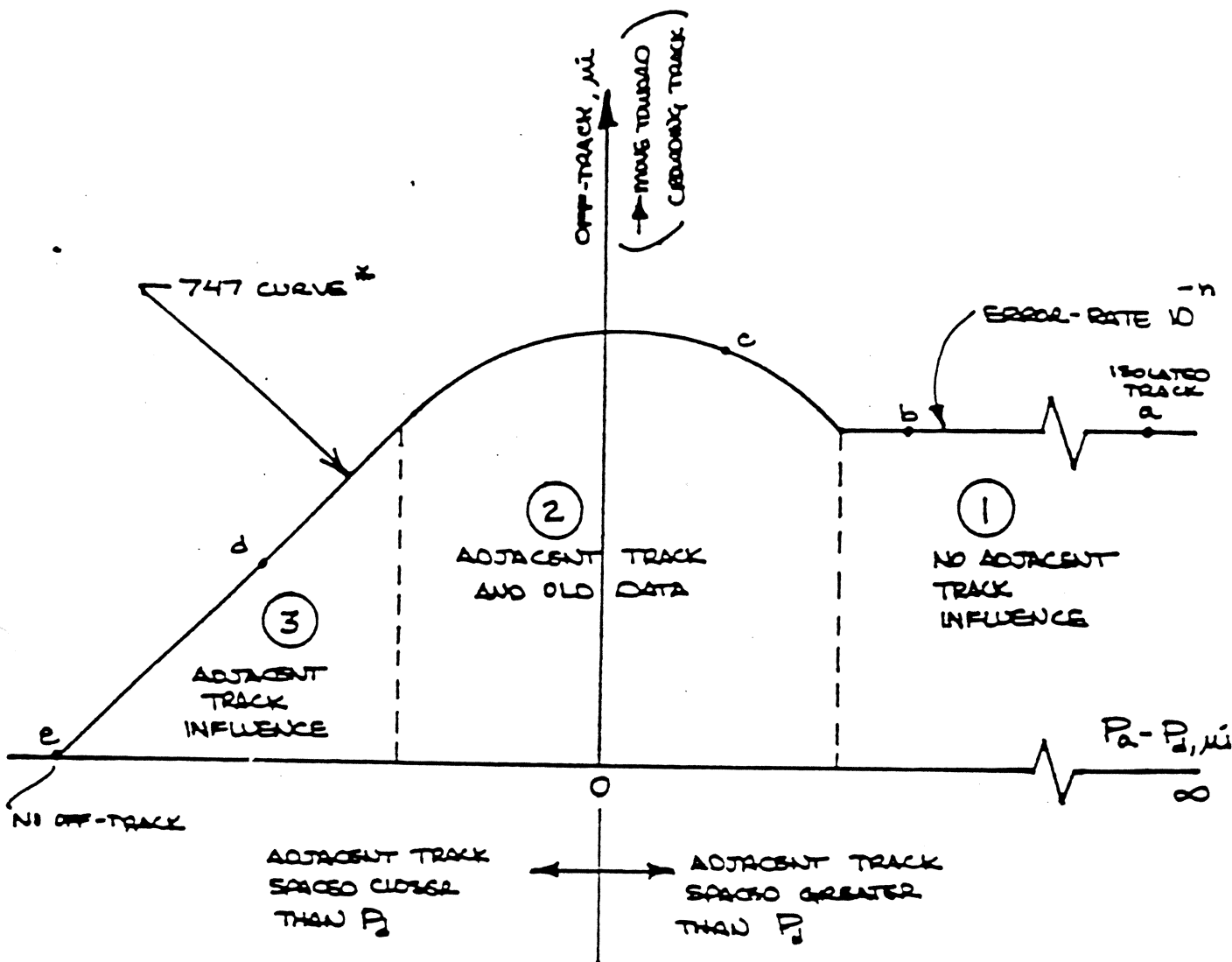
$$W = R + 4P$$

$$\text{track pitch} = W \text{ or } tpi = 1/W$$

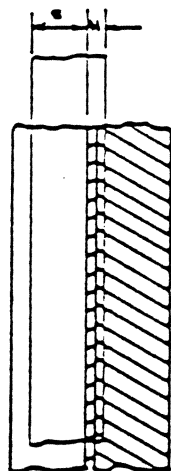
Single element: Signal $\propto W - 2P$

$$\text{Noise } \propto \sqrt{N_{\text{old}}^2 + N_{\text{adj}}^2}$$

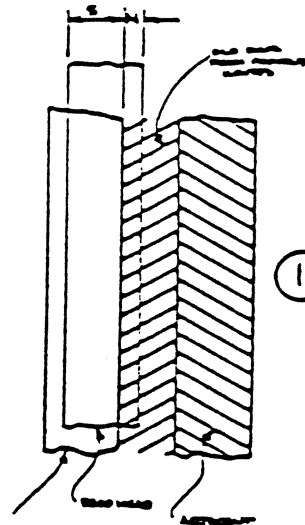
747 OTC CURVE



3



2



1

IIST

HEAD POSITIONING

Two requirements on head positioning

- Track following
- Track seeking

Track following

- Sources of track position variation
 - Spindle bearing runout
 - disk deflections
- Mechanical disturbances
 - windage
 - carriage bearing friction
 - vibration and shock

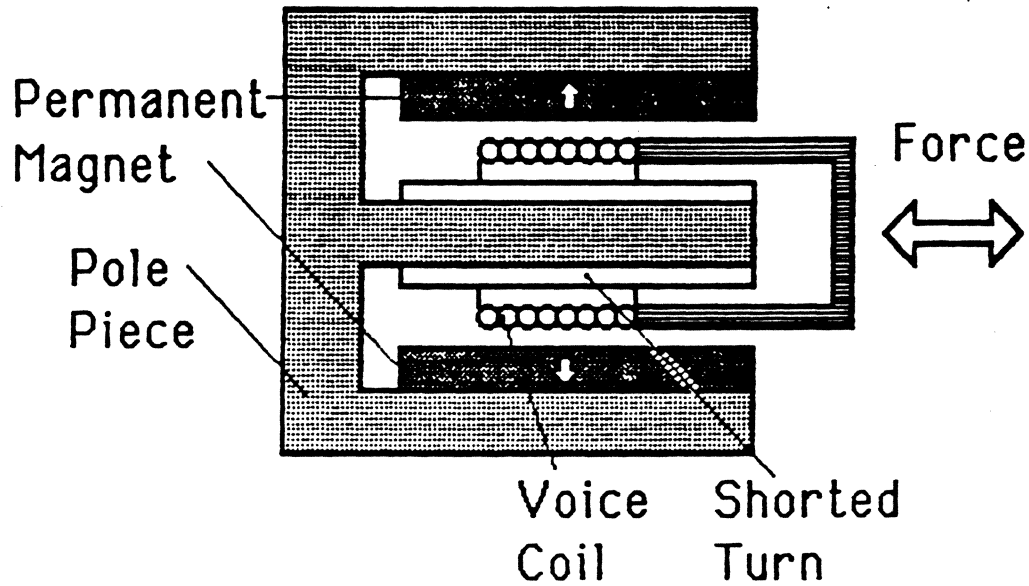
Track seeking

- For minimum access time
 - Use maximum power
 - max accel then max decel
 - bang-bang servo
 - design involves choosing switch point

IMPORTANCE OF SERVO SYSTEM

- ACCESS TIME
- TRACK DENSITY
- THERMAL ROBUSTNESS
- RELIABILITY

Voice Coil Actuator I



$$F = B l i \quad (9.1)$$

$$= K i \quad (9.2)$$

F: Force (Newton)

B: Flux density in gap (Tesla)

i: Current (Ampere)

l: Active length of coil (Meter)

(e.g. for cylindrical geometry:

$l = \pi \times \text{coil diameter} \times \text{Number of turns}$)

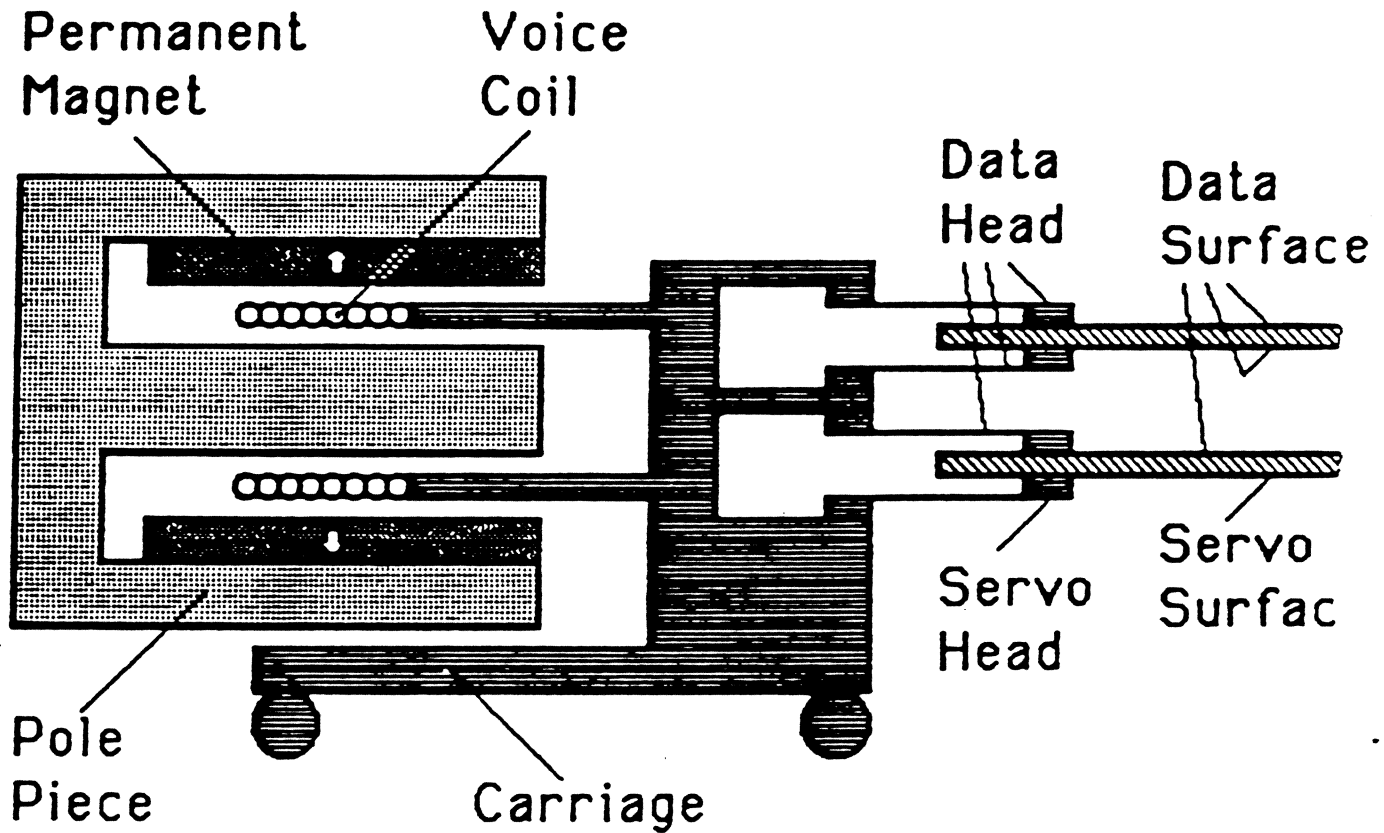
K: Force factor

Source: A. Wagner

IIST

A. Hoagland 3/90

Closed Loop Actuator - Principle



The carriage is actuated by a linear ("voice coil") motor. The carriage position is sensed from a magnetic servo pattern. Most commonly, the servo pattern is on a dedicated surface, (servo surface) and read by a servo head. The magnetic servo pattern is prerecorded during drive manufacturing.

Voice Coil Actuator III Force and Acceleration

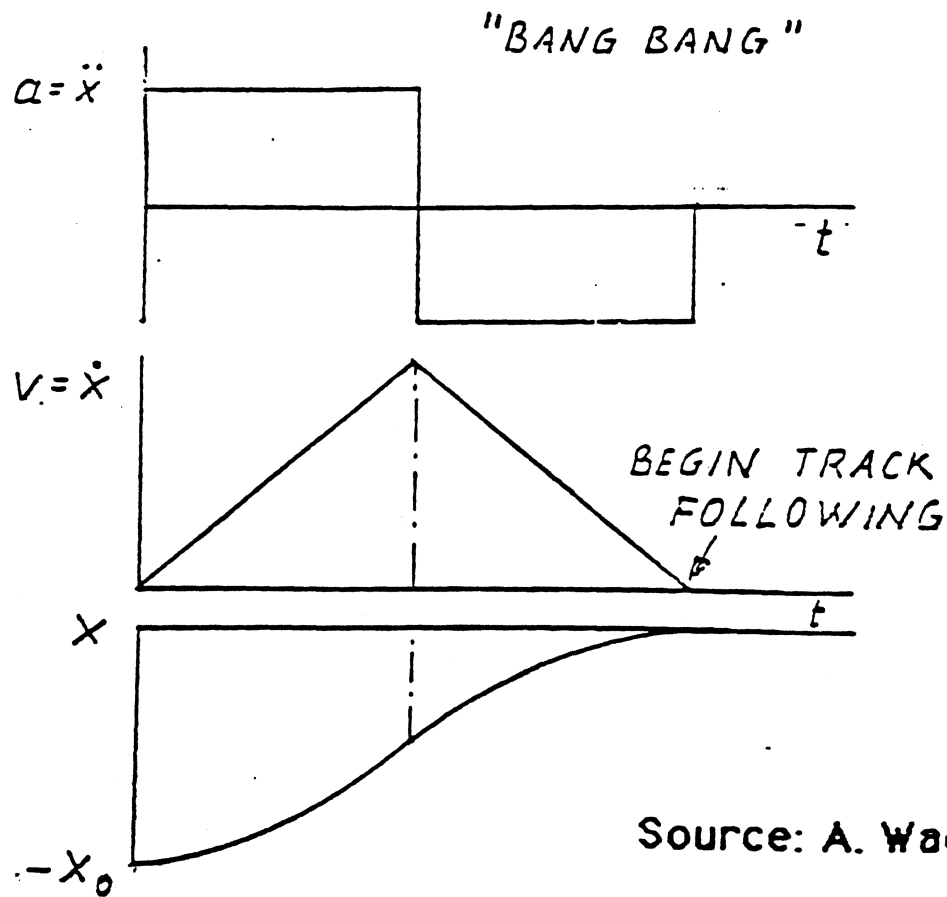
110

$$F = m a \quad (9.3)$$

$$K i = m a \quad (9.4)$$

$$i = \frac{V_0}{R_1 + R_2} \quad (9.5)$$

$$a = \frac{K i}{M} = \frac{K}{M} \frac{V_0}{R_1 + R_2} \quad (9.6)$$



Source: A. Wagner

HEAD POSITIONING

Equations - where mass alone considered

$$a = F/M$$

a = acceleration, F = Force and M = mass of actuator assembly.

For voice coil actuator

$$F = KI$$

I = maximum current available from power supply. $K = Bl$.
(B equal magnetic flux density and l = coil length).

$$a = \frac{K}{M}I$$

and

$$v = \frac{KI}{M}t \quad \text{or} \quad t = \frac{M}{KI}v$$

v = actuator velocity and t = time.

Then

$$x = \frac{KI t^2}{M} = \frac{1}{2} \frac{M}{KI} v^2$$

Let X_s = stroke length and T_{sk} = seek time

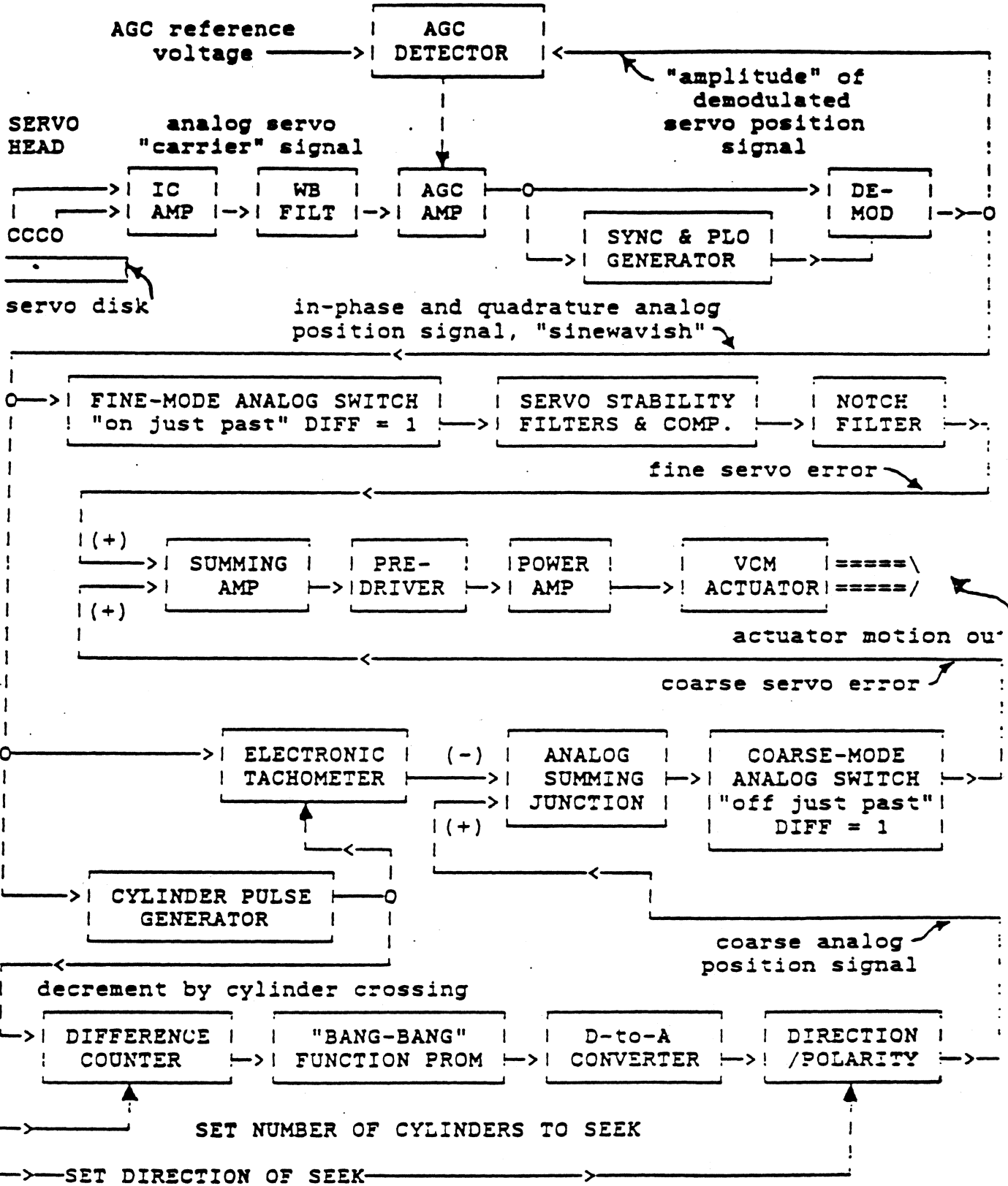
Then

$$X_s \propto \frac{KI}{M} T_{sk}^2 \quad \text{or} \quad I \propto \frac{M X_s}{K T_{sk}^2}$$

As Power needed $\propto I^2$

$$\text{Power} \propto \left(\frac{M}{K}\right)^2 \frac{X_s^2}{T_{sk}^4}$$

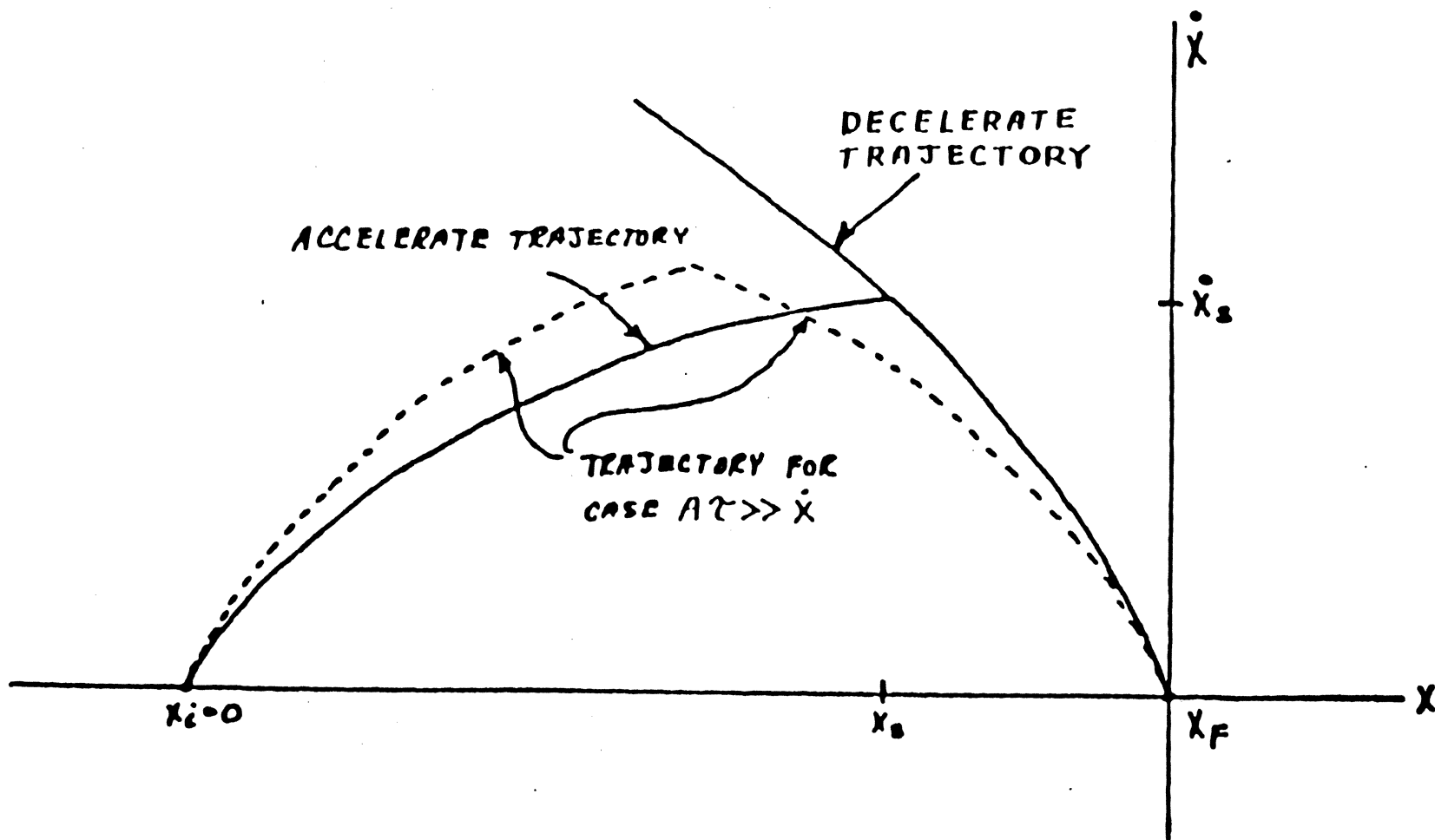
HIGH PERFORMANCE HEAD POSITIONING SERVO SYSTEM (HPHPSS) BLOCK DIAGRAM



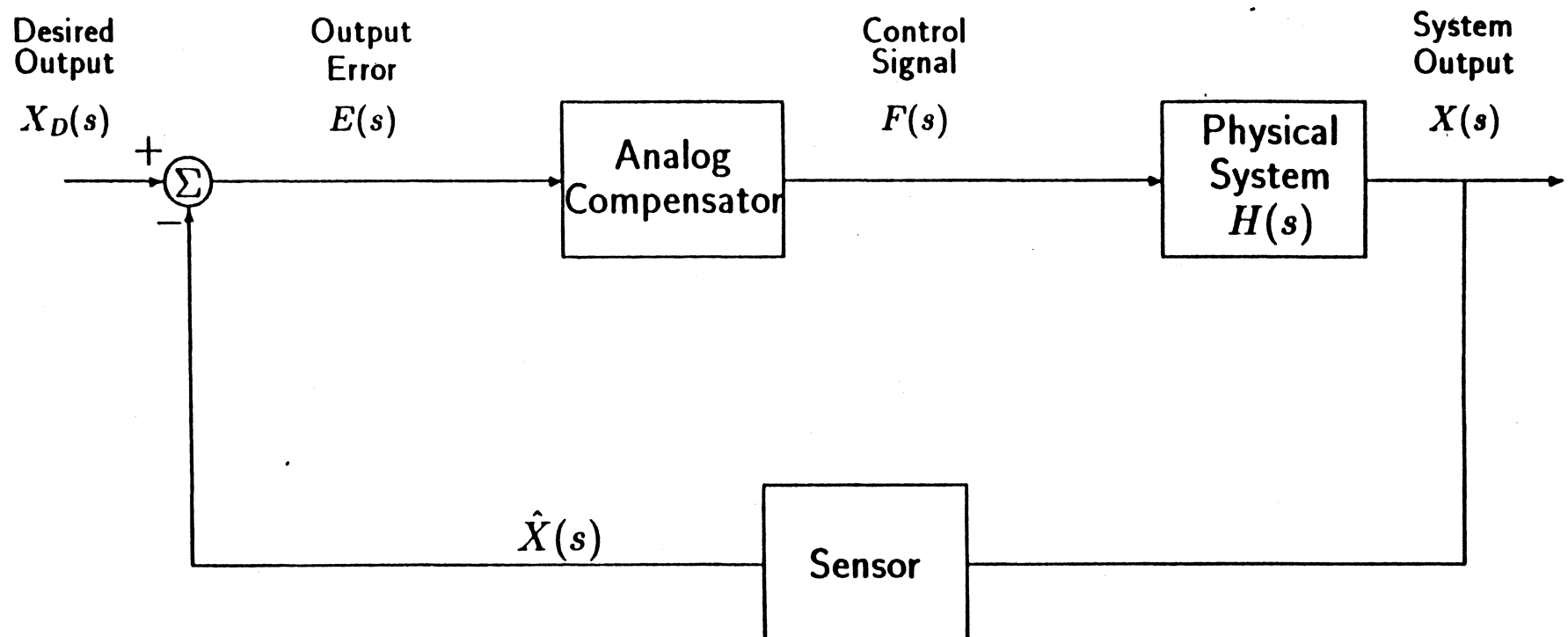
F. J. SORDELLO 1988

IIST

A. Hoagland 3/90



Structure of a Control System



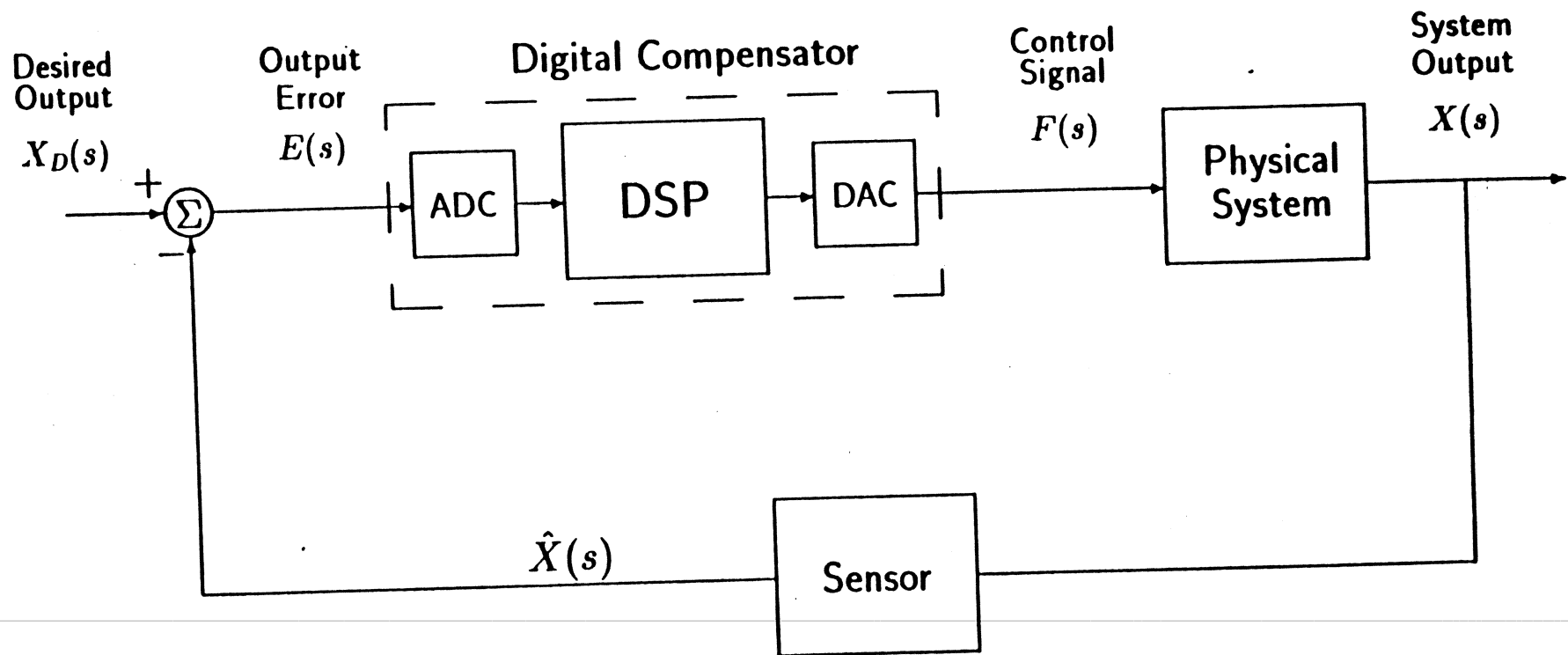
Analog Compensators

Advantages

- Cheap parts
- Control Theory is well known

Disadvantages

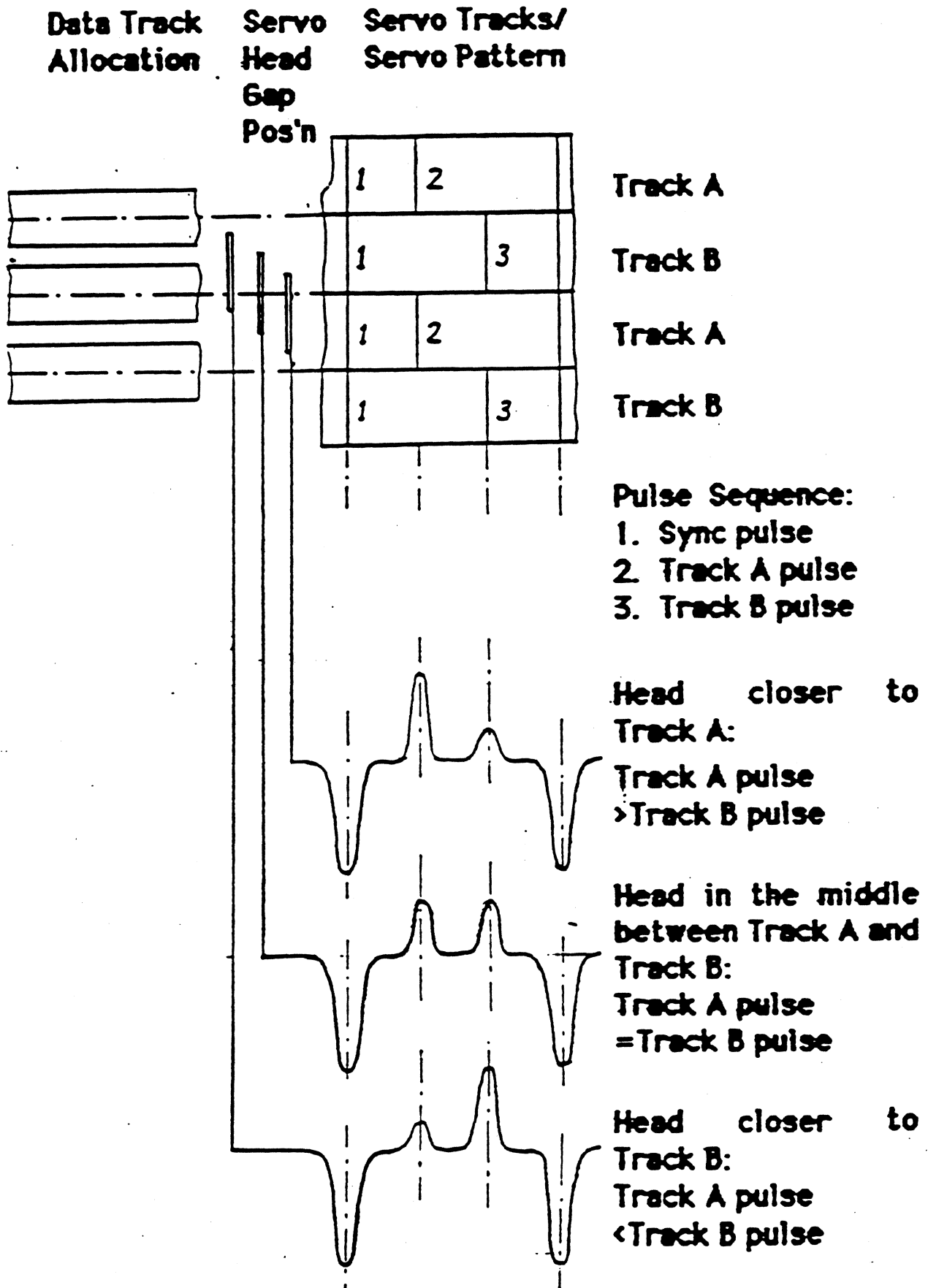
- Difficult to modify
- Components change with time
- Circuits cannot adapt to changes

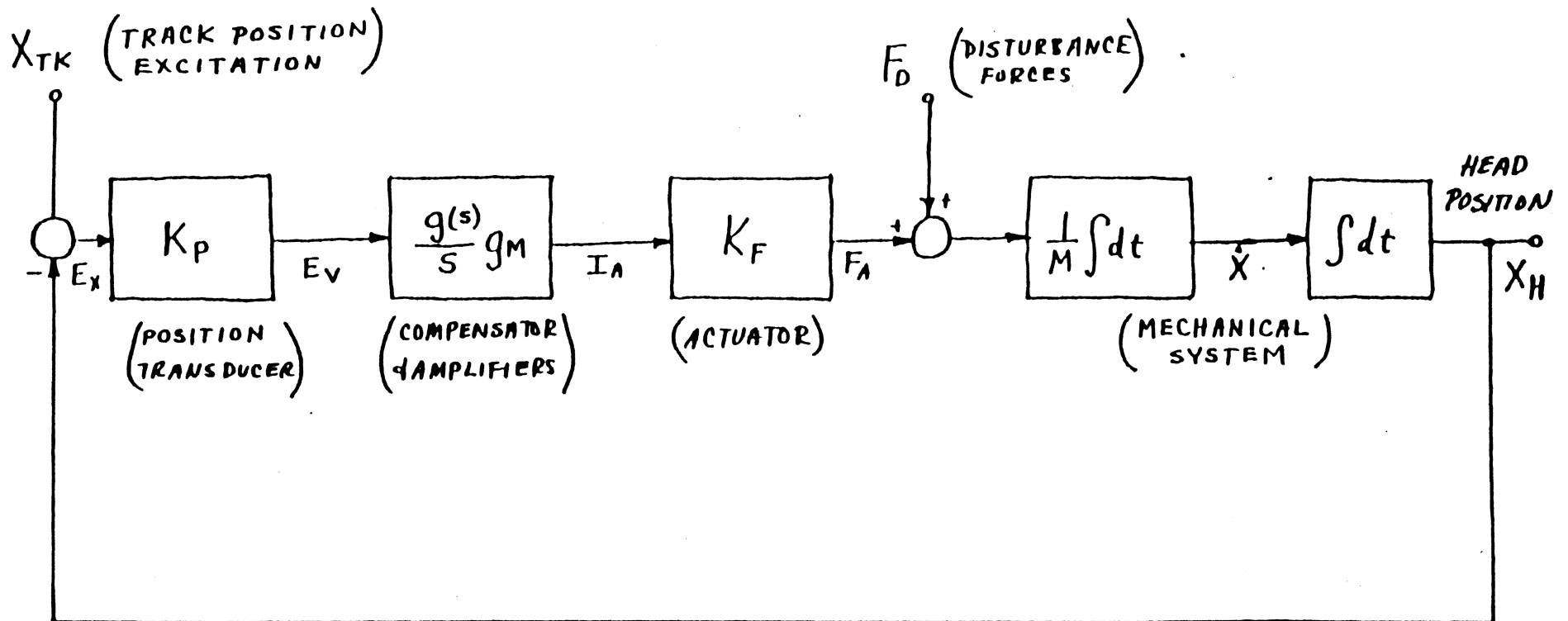


Advantages of Digital Control

- **Greater Flexibility**
 - Rapid prototyping
 - built in diagnostics
 - dynamic decision making
- **Filter Characteristics do change with time or temperature**
- **Learning Ability**
 - On-line identification
 - Adaptive control

"Tribit" Servo Pattern

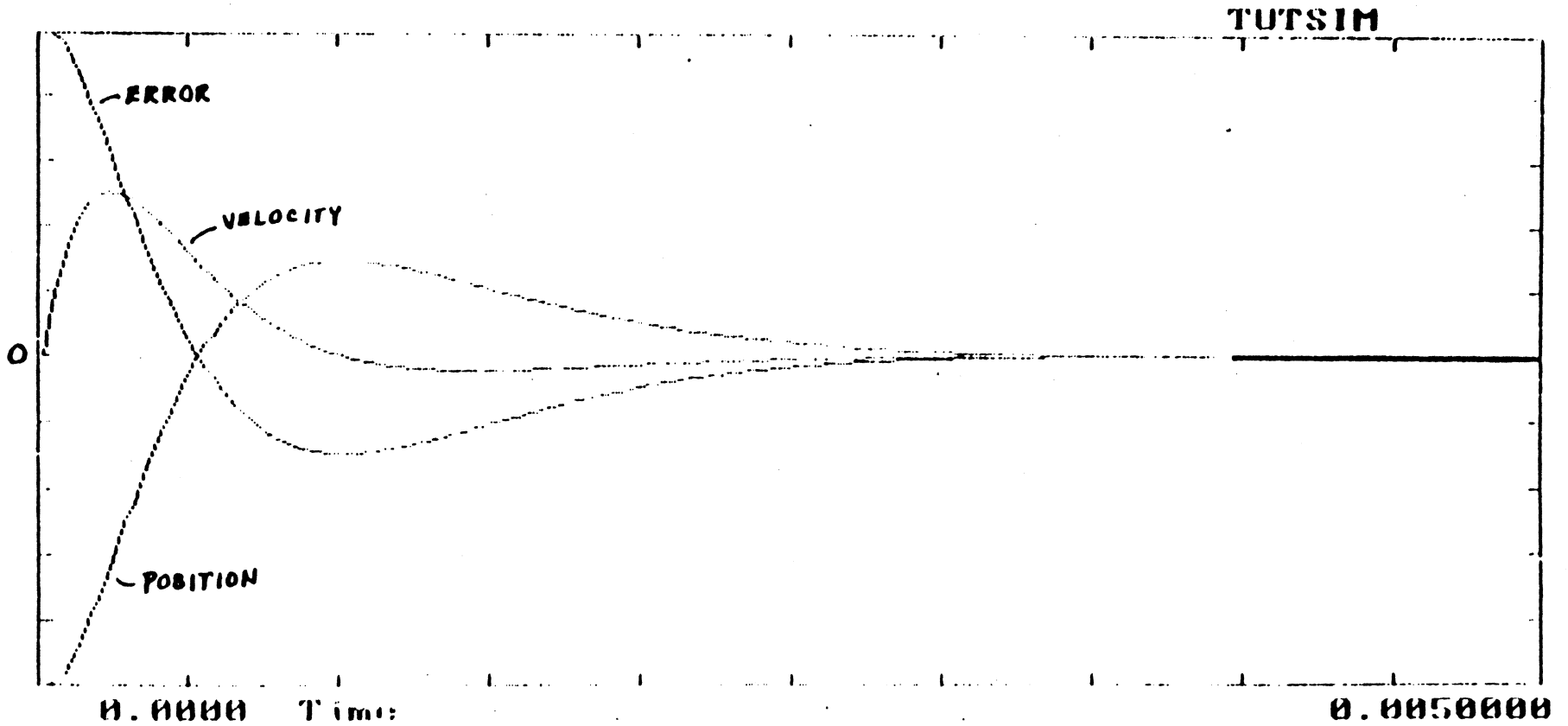


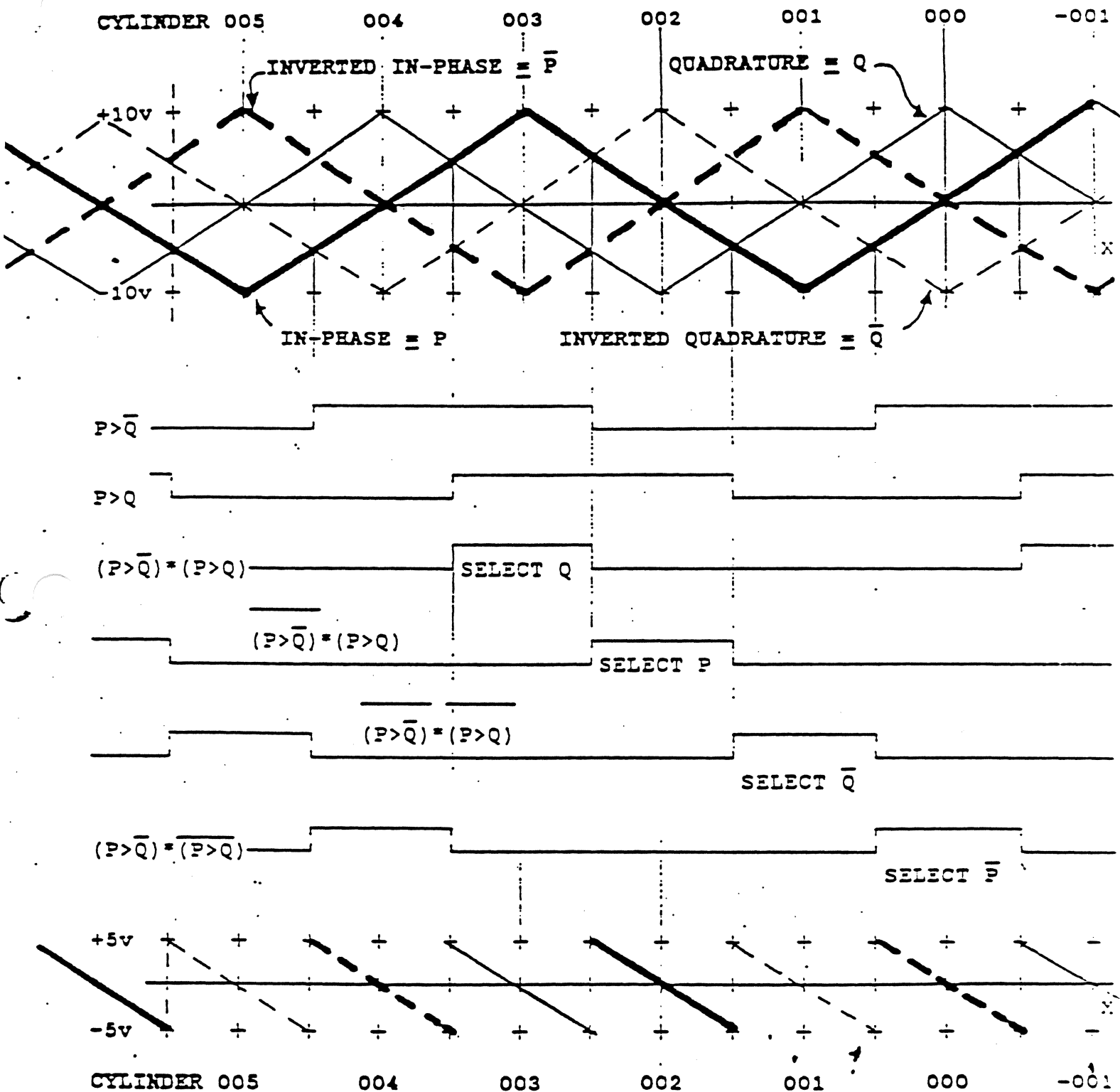


TRACK FOLLOWING SERVO

12% overshoot 28%
settling time ≈ 2.3 ms.

Transient response
 $x_H(t) = \mathcal{L}^{-1} \left[\frac{1}{s} \cdot T(s) \right]$





POSITION TRANSDUCER "IN-PHASE", "INVERTED IN-PHASE", "QUADRATURE", AND "INVERTED QUADRATURE" SIGNALS VERSUS RADIAL DISK POSITION OR CYLINDER LOCATION; LINEAR REGION SELECTION LOGIC AND CYLINDER COUNTING IN COARSE MODE; AND, SELECTED LINEAR REGION FOR FINE MODE.

ENCODER AND WEDGE/TRACK

+ PROS;

- SIMPLE SERVO WRITER
- LOW OVERALL COST
- LOW OVERHEAD
- TRANSPARENT TO CONTROLLER - FREE FORM FORMATTING

+ CONS;

- LOW BANDWIDTH FEEDBACK FROM DISK (ONCE AROUND)
- CORRECTS FOR LONG TERM EFFECTS ONLY
- HIGHER INERTIA OF ARM STACK (ACCESS TIME CONSIDERATION)
- LIMITED TPI UPPER LIMIT

DEDICATED SURFACE

+ PROS;

- FEEDBACK FROM DISK STACK
- WIDE SERVO BANDWIDTH
- CORRECTS RUNOUT
- BETTER SHOCK AND VIBRATION IMMUNITY
- WELL DOCUMENTED APPROACH
- MORE PRECISE SERVOWRITING
- TRANSPARENT TO CONTROLLER - FREE FORM FORMATTING

+ CONS;

- FEEDBACK FROM SERVO HEAD ONLY
- NOT EFFECTIVE FOR DIFFERENTIAL DISK SLIP
- SERVO TO DATA SURFACE RELATIONSHIP DESIGNED IN (COST OF MECHANICS)
- EXPENSIVE SERVOWRITERS
- % OF OVERHEAD INCREASES WITH SMALLER FORM FACTORS

EMBEDDED

+ PROS;

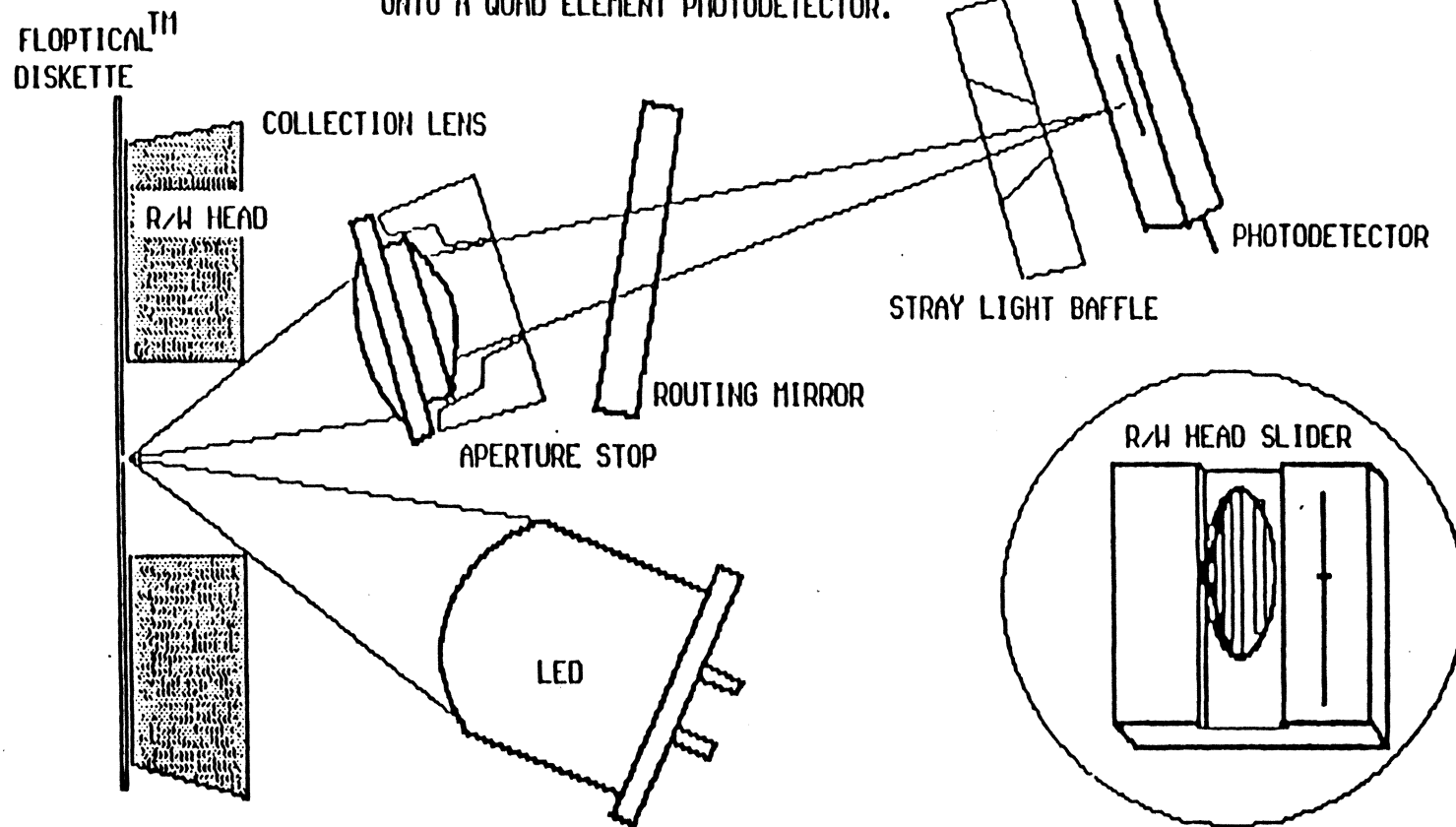
- FEEDBACK FROM DATA HEADS
- REMOVES THERMAL TILT VARIATIONS
- LOWER OVERHEAD

+ CONS;

- "NON ZERO" HEAD SWITCH TIME
- LOWER BANDWIDTH SERVO SYSTEM
- HARD SECTORED - LESS TRACK FORMAT FLEXIBILITY
- INCOMPATIBLE WITH DEVICE LEVEL INTERFACES
- *SETTLING TIME*

OPTICAL SCHEMATIC

AN I.R. LED LIGHT SOURCE IS USED TO REFLECT THE IMAGE OF THE SERVO TRACKS THROUGH AN ASPHERICAL COLLECTION LENS ONTO A QUAD ELEMENT PHOTODETECTOR.



FLEXIBLE DISK DRIVES

Brier Technology fields a first: a 20M-byte, 3½-inch flexible drive

Mike Seither, Senior Editor

San Jose start-up Brier Technology Inc. hopes to turn a few heads at the COMDEX/Fall show in Las Vegas next month when it introduces a 20M-byte, 3½-inch flexible disk drive. That's more than double the storage of the newest breed of high-capacity 5¼-inch flexible disk drives.

With an average access time of 35 msec, Brier's BT 3020 high-capacity product also rivals the performance of comparable-size Winchester disks and offers a removable media to boot. This is something that is attractive to OEMs and system integrators working in the engineering, scientific and military segments, where workstations are shared and the best way to protect one's data is to remove it.

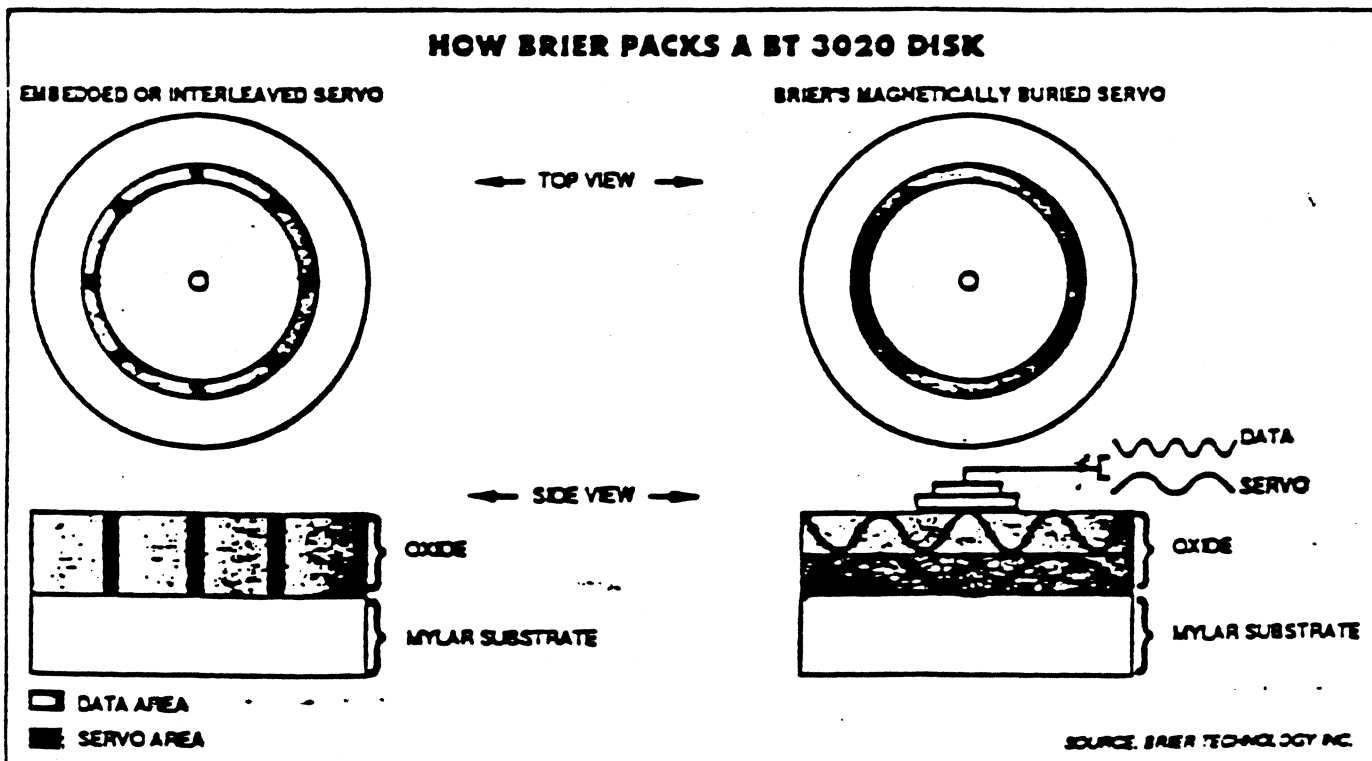
Brier's first product uses a proprietary system that embeds continuous servo information into the magnetic oxide coating of a standard 3½-inch high-density disk. The company claims that it has been able to divide the disk's oxide media into two logical layers. Servo information in the deeper layer, recorded at a very low frequency, is used to keep the read/write head on track. The upper layer is dedicated solely to user data and is written at a much higher frequency than the servo information.

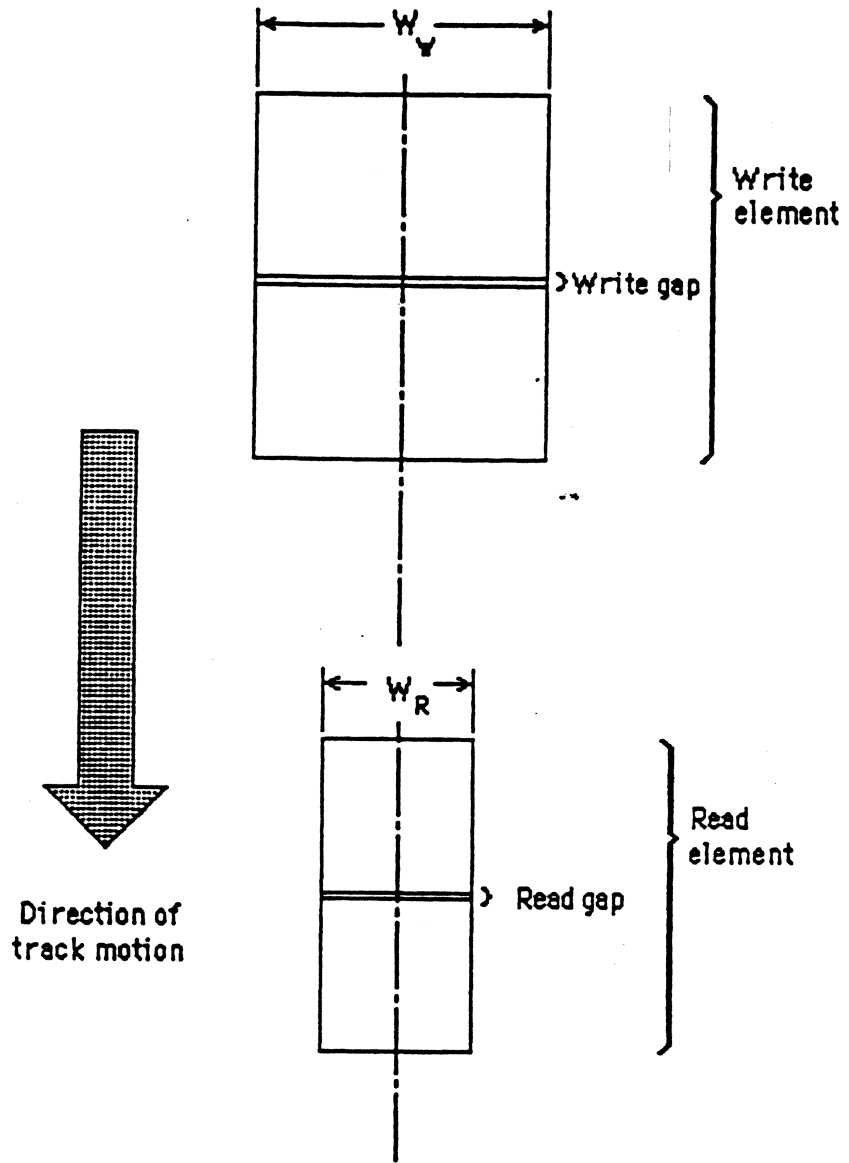
By comparison, most vendors selling 5¼-inch disk drives in the range of 4M bytes to 10M bytes embed servo information directly on the data surface. As a result, up to half the total data area is given up to servo.

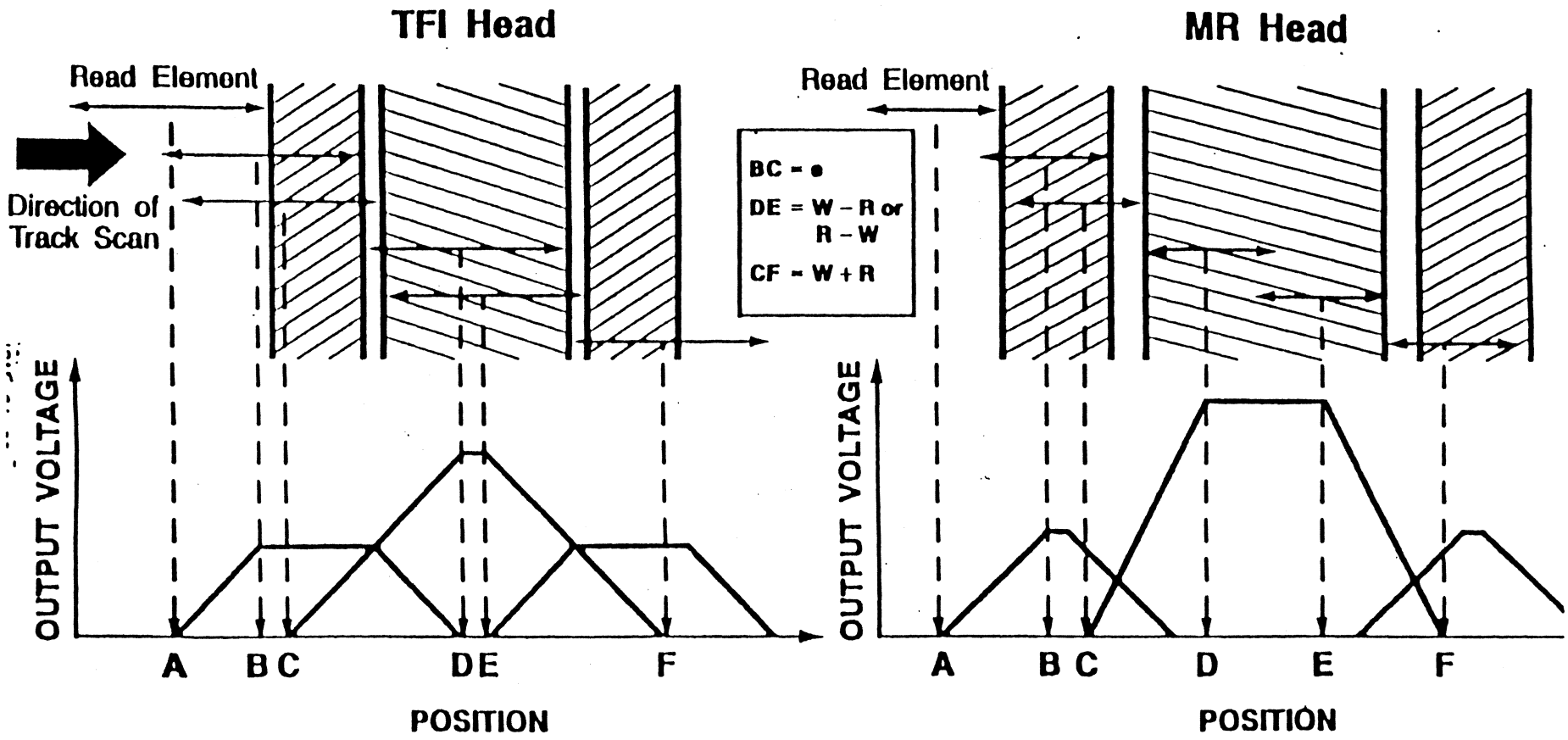
Thanks to this buried-servo system,

Brier has been able to overcome the biggest technical hurdle—track density. The BT 3020 has 777 tracks per inch, compared with the 96 tpi on a regular 5¼-inch disk. By increasing the bit density and using advanced

Embedded servos (left) typically use part of the data surface to the entire depth of the oxide media to store read/write head tracking information. Brier horizontally divides the oxide into two logical layers (right). Prerecorded, continuous servo information is written at a low frequency onto the lower. Data is written over the entire upper layer at a high frequency. Brier's drive electronics sends data one way and head tracking information another.







DISCRETE TRACKS - POSSIBILITIES FOR HIGH-DENSITY MAGNETIC DISKS?

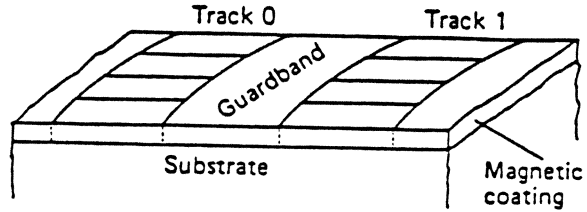
Ian L. Sanders and Steven E. Lambert

PROCEEDINGS

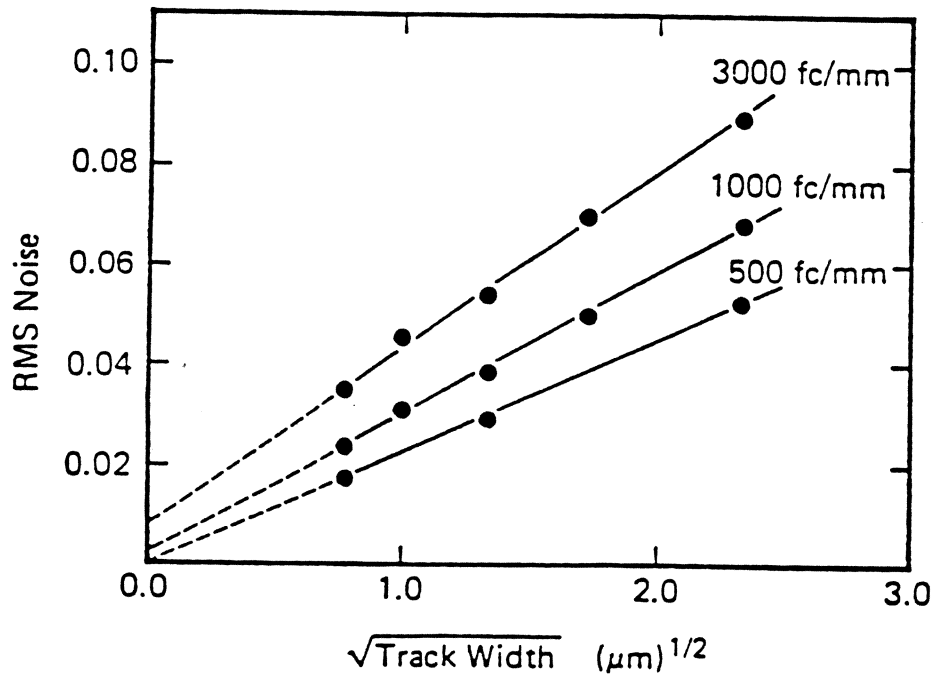
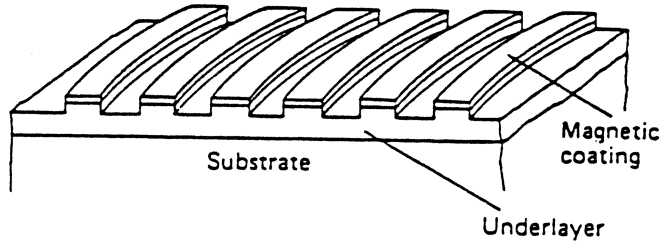
VLSI and Computer Peripherals

3rd Annual
European
Computer Conference

Continuous



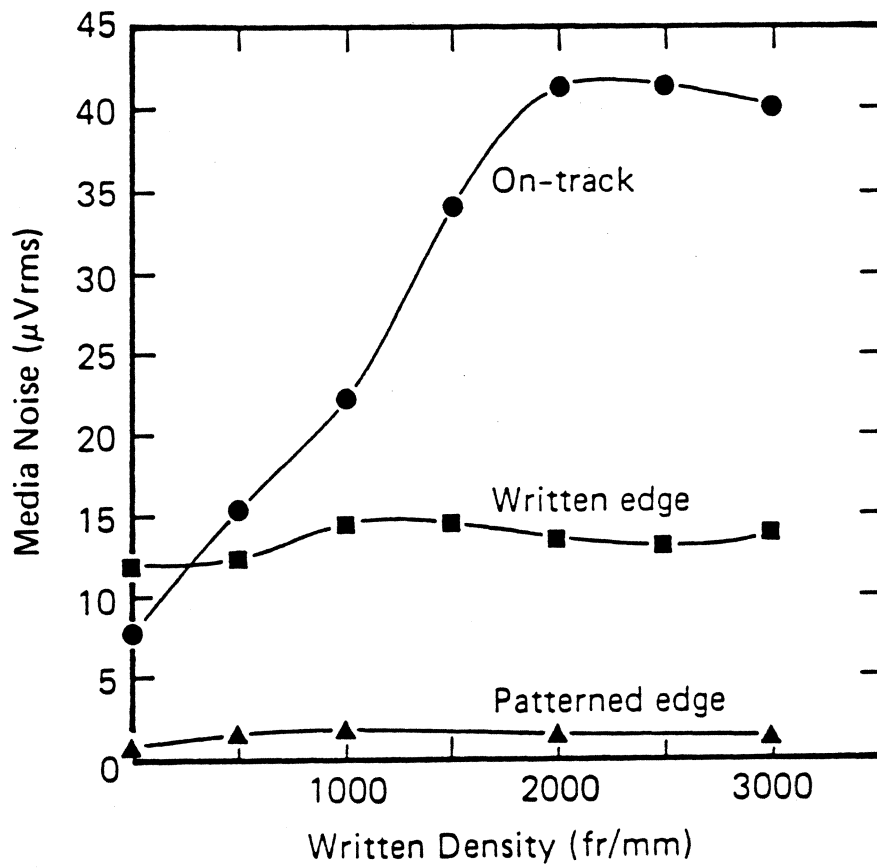
Discrete



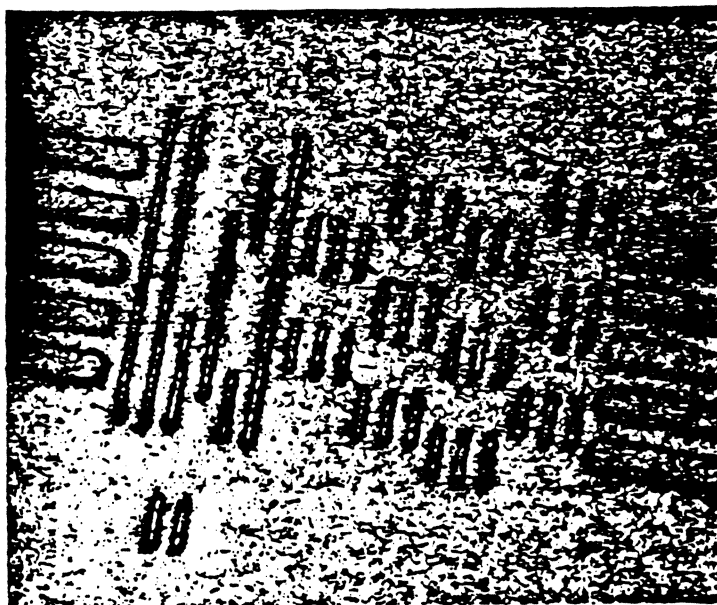
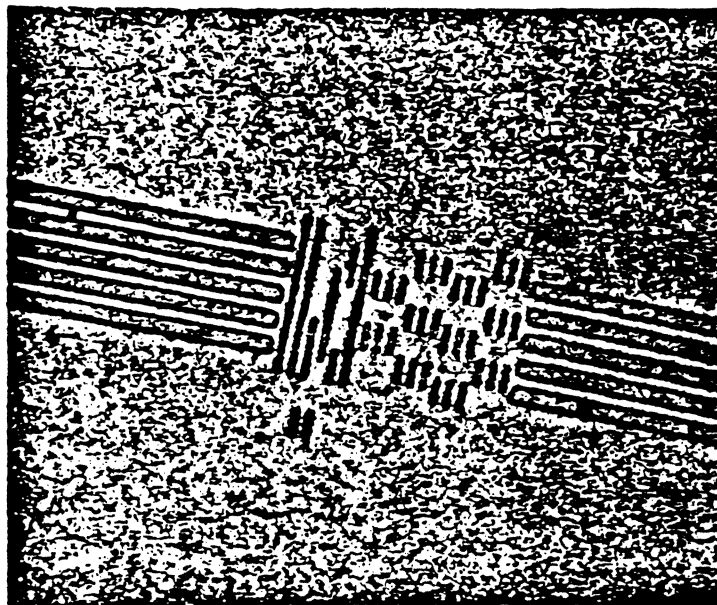
IIST

A. Hoagland 3/90

Edge Noise Reduction in Discrete Tracks



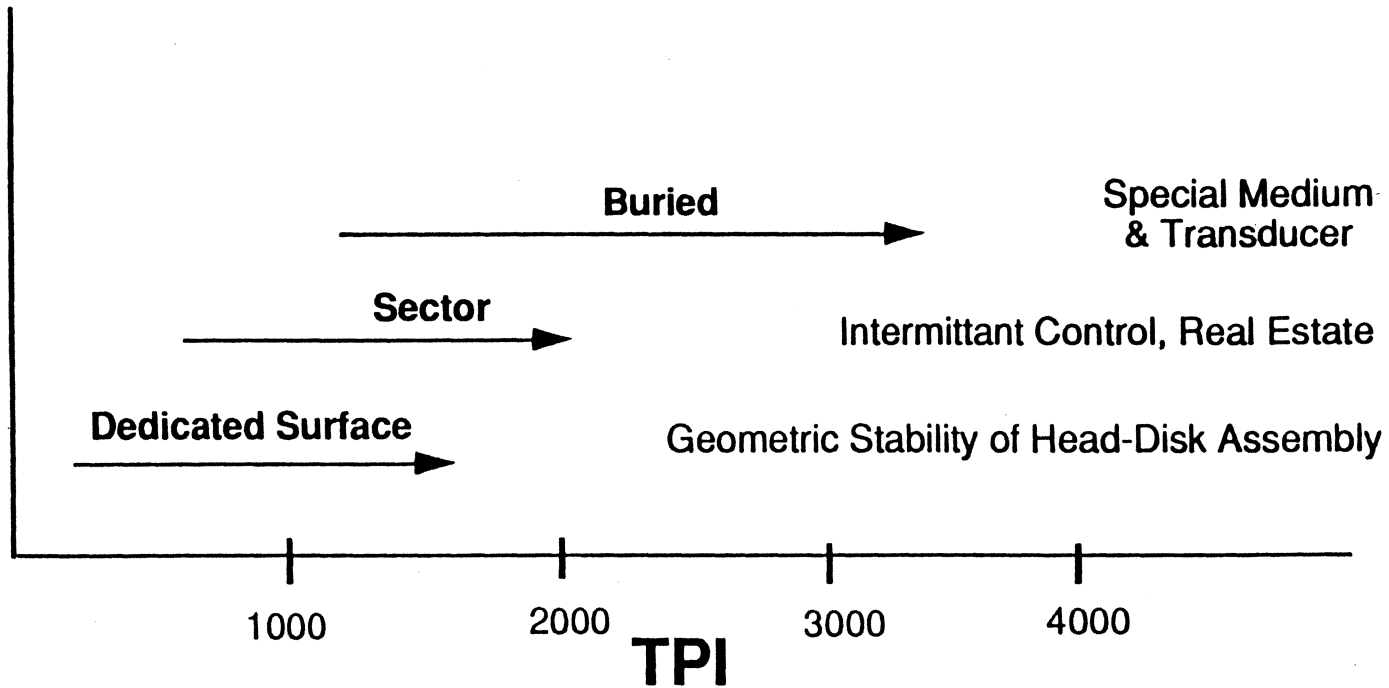
LITHOGRAPHICALLY DEFINED SECTOR-SERVO



IIST

A. Hoagland 3/90

Track Density in Magnetic Storage



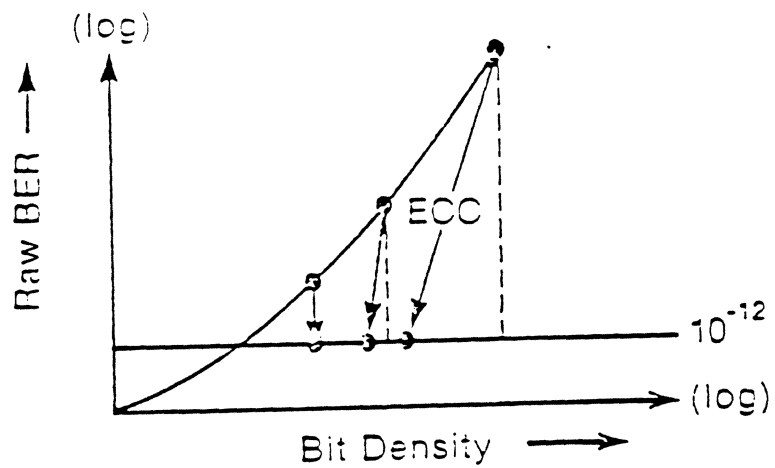
DATA STORAGE DEVICES

IIST

A.S. Hoagland '90

Error Control and Correction

- ECC Overhead Increases with Bit Density



- For Density $> 10^9$ bpi²
Raw BER 10^{-4} - 10^{-6}
⇒ ECC Overhead 10-20%

The error probability p , equal to the BER, is also called the transition probability, where transition refers to a "1" changing into a "0", and vice versa. The probability of no errors is $1-p$, and n independent symbols are therefore received error free with a probability of $(1-p)^n$. We can summarize:

$$P(\text{correct}) = (1-p)^n$$

$$P(1 \text{ error}) = p(1-p)^{n-1}$$

$$P(2 \text{ errors}) = p^2(1-p)^{n-2}$$

+++++

$$P(b \text{ errors}) = p^b(1-p)^{n-b} \quad (25.6)$$

+++++

The actual probability will be greater, since one (or more) errors can result in different error patterns:

ONE error: E X X X X
 - or X E X X X
 - or X X E X X
 - or X X X E X
 - or X X X X E .

For a n -dimensional word there are n different patterns for a single error, and the true error probability is therefore

$$P(1 \text{ error}) = n \cdot p(1-p)^{n-1}$$

Two errors will result in $n \cdot (n-1)/2$ patterns:

E E X X X	X E X E X
E X E X X	X E X X E
E X X E X	X X E E X
E X X X E	X X E X E
X E E X X	X X X E E

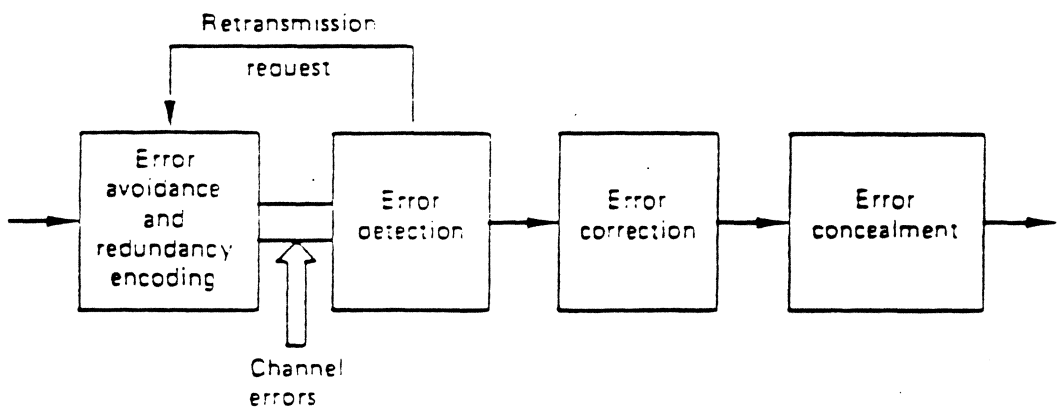


Figure 7.1 The major processes in an error-handling system.

Recording density (RD) is the number of user bits which can be recorded per unit length of track (e.g., inch). For our purposes it is most convenient to take this unit length to be T_{min} , the minimum recorded spot size. (This quantity is a measure of the R/W channel bandwidth and is usually difficult to increase.) That is,

$$RD = \text{number of user bits in interval } T_{min}$$

$$= (m/n) \cdot \text{number of code bits in interval } T_{min}$$

If the interval T_{min} is divided into $d+1$ detection windows of length T_w ; in other words, there are $d+1$ code bits in an interval of length T_{min} . Thus,

$$RD = \frac{m}{n} \cdot (d+1)$$

ERRORS ON MAGNETIC DISKS

- Surfaces of both rigid and flexible disks have few defects and very little dirt. Also, signal-to-noise ratio is typically very high (e.g., 35db).
- Disks are formatted and certified before use. During certification a test pattern is written at each sector and then read back. If it contains any errors, the sector is discarded. (Several spare sectors are provided.)
- As a result of the above two facts, errors are rare. Typical bit error rates are in the range of 10^{-10} to 10^{-12} , giving sector error rates of 10^{-7} to 10^{-9} .
- Errors are caused by:
 - dirt
 - head skipping
 - marginal oxide areas
 - noise
 - mis-tracking
 - demodulator timing jitter
 - intersymbol interference
- A typical error event is caused by misreading a single flux change. There are three ways in which this can occur:
 - Drop in - reading a flux change where there is none.
 - Drop out - missing a flux change.
 - Shift - reading a flux change in the wrong timing interval.
- Errors are extended by the modulation code decoder. Result of typical error event is a short burst of errors (2-10 bits).
- Most errors are soft in that they do not repeat on multiple reads.

ISC 4

HARD VS. SOFT SECTORS

		Sector Type	
		Hard (factory sectored, servo tracking, high density)	Soft (user sectored, open-loop tracking, low density)
Media Type	flexible (mylar)	High-performance floppies (e.g., Bernoulli Box), old 8" floppies	floppies
	rigid (aluminum)	High-performance Winchester	PC Winchesters

Hard-sectored disks are formatted very precisely at the factory. They are capable of accurate tracking and thus achieve high recording density. This high density, in turn, makes system more sensitive to noise, ISI and timing jitter.

ERROR CORRECTION METHODS FOR MAGNETIC DISKS

1. Detect-and-Reread
 Error detection code only. Reread until correct read achieved. (System parameters like tracking offset or demodulator timing offset may be adjusted on different reads.) Code must have large number of check bits to achieve moderate reliability. Disk performance poor.
2. Reread-and-Correct
 Fire or Computer-generated (single-burst correcting) code. Same as Detect-and-Reread, except that after a fixed number of unsuccessful reads, correction is performed. Correction is typically performed by CPU with location and pattern provided by controller. Better performance.
3. Correct-and-Reread
 Error correction is performed in controller. If a detectable but uncorrectable pattern occurs, reread is invoked. Typical code: Interleaved RS2. Since rereading is rare, best performance.

131

Binary System:

Addition

	0	1
0	0	1
1	1	0

Module 2
Multiplication

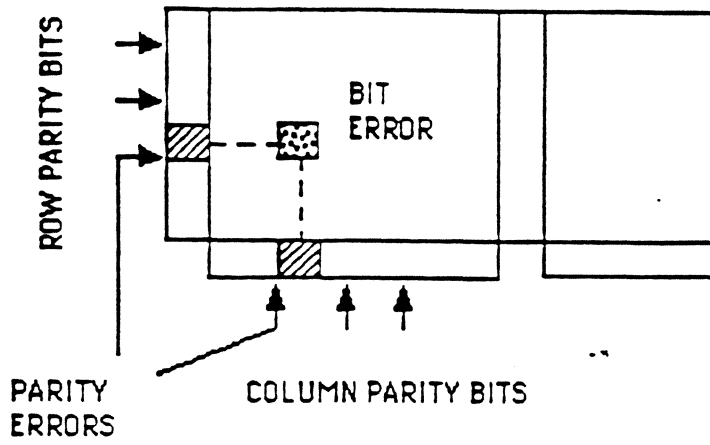
	0	1
0	0	0
1	0	1

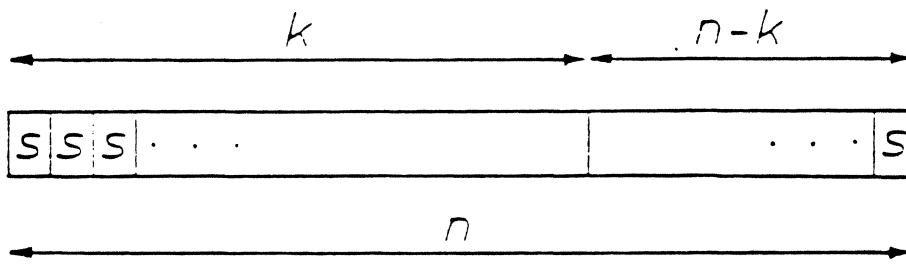
Vector

transmitted word (t)	1011001
received word (r)	1001001
error pattern (e)	0010000

$$t = r + e$$

thus once error vector determined can add to received word to correct and get transmitted word.





A code word of length n consists of an information block of k symbols and a parity block of $n - k$ symbols; each symbol comprises s bits. The number of possible words of n symbols is 2^{ns} . The parity bits are fixed for each combination of the ks information bits in accordance with established encoding rules. The number of code words is thus 2^{ks} . It follows that the fraction $2^{(k-n)s}$ of the number of possible words consists of code words.

BLOCK CODES - BASIC CONCEPTS

code length (n) - number of bits per word (all words have same length)

number of information bits (k) - transmission of a code word conveys a k-bit message, where $k < n$.

number of parity check bits - $n - k$

code rate or efficiency (R) - k/n . Since $k < n$, it follows that $R < 1$.

systematic code - block code in which the first k bits of each code word are identical to the k-bit message

encoding - associating a code word with a k-bit message

decoding - associating a k-bit message with a received word

ENCODING TABLE

This is a simple list of messages and their associated code words. In principle, encoding can be accomplished by looking up message in table, then selecting corresponding code word.

Example In a system with four messages and binary code words of length 5, one possible encoding table is as shown.

messages	code words
1	11000
2	00110
3	10011
4	01101

CONSTRUCTION OF DECODING TABLE

In a decoding table, all 2^n possible received words are grouped into 2^k columns, with each column headed by a code word. Important point: entries in a column are selected to be the words most likely to be received when the corresponding code word is transmitted.

Example: $n = 5, k = 2$

Message	1	2	3	4
Code Words	11000	00110	10011	01101
	11001	00111	10010	01100
	11010	00100	10001	01111
Other Received Words	11100	00010	10111	01001
	10000	01110	11011	00101
	01000	10110	00011	11101
	11110	00000	10101	01011
	01010	10100	00001	11111

Note: Since there are 2^n entries in the table and 2^k columns, there must be $2^n / 2^k = 2^{n-k}$ rows.

USE OF DECODING TABLE

Case 1: Error Correction Only: Received word is located in the table and is decoded into the code word at the head of its column. The decoder outputs the corresponding message.

Case 2: Error Detection Only: The decoder determines if received word is above or below the solid line. If it is above, then it is a code word and the corresponding message is outputted. If it is below, an error has been detected.

Case 3: Combined Error Detection and Correction: The decoder determines if received word is above or below the dashed line. If it is above then it is decoded into the code word at the head of the column. If it is below an error has been detected.

NEAREST NEIGHBOR DECODING

Errors are normally unlikely; for errors which occur at random, the probability of a multiple error pattern decreases as the weight increases. In other words, few errors are more likely than many.

A good error-correcting code is one which successfully corrects patterns having few errors and (i.e., likely ones) and (necessarily) falls into high-weight patterns.

The decoding rule which is most appropriate for random errors is nearest neighbor decoding. Here the received word is decoded into the code word which is nearest to it (in the sense of Hamming distance). Thus, the decoder always guesses that few errors, rather than many occurred.

HAMMING WEIGHT AND DISTANCE

(Hamming) weight of a vector v is

$$w(v) = \text{number of non-zero components in } v$$

(Hamming) distance between two vectors v_1 and v_2

$$d(v_1, v_2) = \text{number of places in which } v_1 \text{ and } v_2 \text{ differ}$$

$$= w(v_1 - v_2)^*$$

$$= w(v_1 + v_2)^* \text{ in the binary case}$$

minimum (Hamming) distance of a code is the distance between the two closest code words.

that is, $d_{\min} = \text{smallest value of } d(v_i, v_j) \text{ where } i \neq j.$

Hamming distance is primarily useful in dealing with random errors.

*Addition (or subtraction) of vectors is always performed by adding (or subtracting) corresponding terms.

ERROR CORRECTION CAPABILITIES ARE DETERMINED BY d_{\min}

Error Correction:

$$\begin{matrix} \text{Case 1: } d_{\min} = 2t + 1 \\ \text{Case 2: } d_{\min} = 2t + 2 \end{matrix} \left. \vphantom{\begin{matrix} \text{Case 1: } d_{\min} = 2t + 1 \\ \text{Case 2: } d_{\min} = 2t + 2 \end{matrix}} \right\} \text{guaranteed error-correction capability of code} = t$$

Proof: let c_t = transmitted code word

c_n = neighboring code word distance d_{\min} away

v = received word at distance $e \leq t$ from c_t

by triangle inequality

$$d(c_t, c_n) \leq d(c_t, v) + d(v, c_n)$$

so

$$\begin{aligned} d(v, c_n) &\geq d(c_t, c_n) - d(c_t, v) \\ &\geq d_{\min} - e \\ &\geq 2t + 1 - t = t + 1 \end{aligned}$$

Hence c_n is farther away and so nearest-neighbor decoding corrects any error pattern of weight $\leq t$.

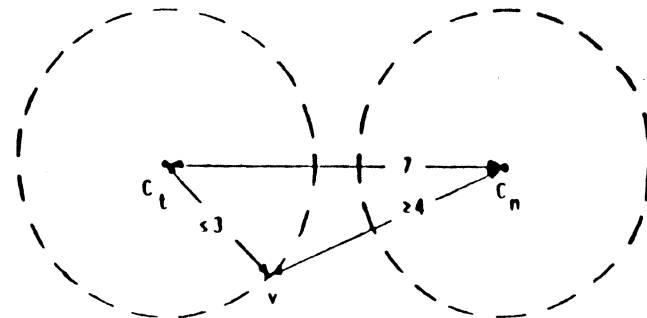
t	d_{\min}
1	3 or 4
2	5 or 6
3	7 or 8
4	9 or 10
\vdots	\vdots

NEAREST NEIGHBOR DECODING AND d_{\min}

c_t = transmitted code word

c_n = code word at distance $d_{\min} = 7$

v = received vector



A code with minimum distance

$$d_{\min}$$

can correct any error pattern of

$$\frac{d_{\min} - 1}{2}$$

or fewer errors in a codeword.

ERROR CORRECTION AND DETECTION

Any error pattern having $\leq e_c$ errors can be corrected and any pattern having $\leq e_d$ errors ($e_d \geq e_c$) can be detected provided:

$$e_c + e_d = d - 1$$

Proof: Since $2e_c < d$, error correction works. For detection to work the received word must be further than e_c from every other code word.

$$\begin{aligned} d(v, c_n) &\geq d(c_t, c_n) - d(c_t, v) \\ &\geq d - (d - 1 - e_c) \\ &\geq e_c + 1 \end{aligned}$$

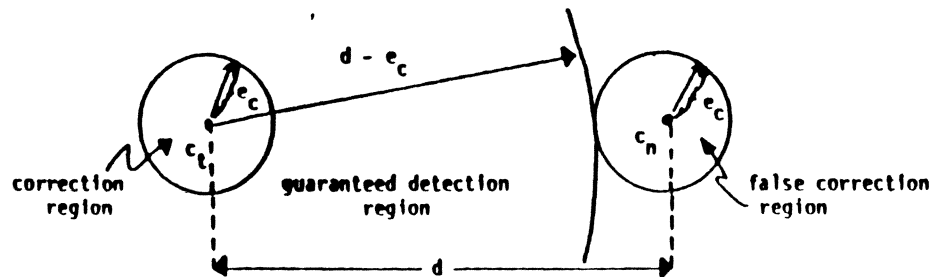
Thus, detection works too.

A SPECIAL CASE

If $e_c = 0$ we have the result:

A code with minimum distance d has error detecting capability $e_d = d - 1$.

Since error correction capability is e_c , each code word is centered in a sphere of radius e_c . These spheres are at least distance d apart.



If fewer than $d - e_c$ errors occur, the received word cannot "reach" the sphere around c_n . Hence any error patterns of weight $d - e_c - 1$ or less will be detected so the error-detecting ability of the code is

$$e_d = d - e_c - 1$$

CORRECTION VS. DETECTION

At design time the designer decides how much distance will be used for error correction and how much for detection. Once the decoder is built, this decision is frozen.

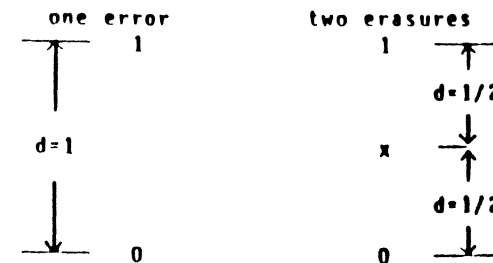
Example: Suppose $d = 7$; then the following choices are possible.

e_c	e_d
0	6
1	5
2	4
3	3

CORRECTION OF ERRORS AND ERASURES

Sometimes the read-head output signal is quantized into three rather than just two regions. The third output is then seen by the decoder as an erasure.

We treat erasures as having Hamming weight $1/2$. It takes two erasures (but only one error) to confuse a transmitted 1 with a 0.



Correction of either errors or erasures is normally done by finding the code word which has the smallest distance to the received word; that is, nearest neighbor decoding.

ERROR CORRECTION

Hamming code (single error correction).

Need check word that can point to single bit in error.

n = word length

k = # information bits

$n - k$ = # check bits

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

i.e.; n possible error locations

+ error free condition

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

or: $n = 2^{(n-k)} - 1$

<u>$n-k$</u>	<u>n</u>	<u>k</u>
2	3	1
3	7	4
4	15	11

TABLE 3.4-2 *Encoding a 4-Bit Message and Locating Error*

<i>Encode</i>							Position
1	2	3	4	5	6	7	
-	-	1	-	0	1	1	Message
0	1	1	0	0	1	1	Encode
		x					Error
0	1	0	0	0	1	1	Receive

Locate error

Check 1	1	3	5	7	
	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	=	3 → position of error

Correct (add 1 into the error position)

	1						Correct error
0	1	1	0	0	1	1	Corrected message

Code Sample after R. Hamming:

Error correction code for four binary digits. We will need three parity checks, i.e. the message including parity bits is seven bits long. If there is no error, the syndrome will be 000, if there is an error, the syndrome gives position of error.

Example message: 1101. For simplicity reasons, the parity bits are interlaced with the data. Dashes symbolize positions, where parity bits are to go.

Bit Position	7	6	5	4	3	2	1
Data	1	1	0	-	1	-	-

1st parity check:

Parity at position
1,3,5,7.

Parity bit position 1: "0" 1 1 0 - 1 - 0

2nd parity check:

Parity at position
2,3,6,7.

Parity bit position 2: "1" 1 1 0 - 1 1 0

3rd parity check:

Parity at position
4,5,6,7.

Parity bit position 4: "0" 1 1 0 0 1 1 0

Parity Check:

Bit Position	7	6	5	4	3	2	1
Message	1	1	0	0	1	1	0
Syndrome:				0		0	0

- no error

Error: Bit 4 changed into 1

Message	1	1	0	1	1	1	0
Syndrome:				1		0	0

- 4 decimal:

Position 4 is position of error. To correct message, reverse bit 4.

Parity check of corrected message:

Message	1	1	0	0	1	1	0
Syndrome				0		0	0

The syndrome of the correct message is 000 ~ message correct.

Note: The Hamming error correction code has been chosen to demonstrate the principle of error correction.

More sophisticated codes are used in disk drives.

THE REED-SOLOMON CODES

- These are cyclic codes whose symbols are binary m -tuples rather than bits.
- length = $N = 2^m - 1$ symbols or $n = m(2^m - 1)$ bits
- If use $N - K$ check symbols ($m(N - K)$ check bits), then code can correct $t = \lfloor (N - K)/2 \rfloor$ randomly located symbol errors. That is, $d = N - K + 1$.
- For t -error-correction, need $2t$ check symbols so $N - K = 2t$
- In a t -error-correcting RS code, the following error patterns are correctable:
 - 1 burst of total length $b_1 = (t - 1)m + 1$ bits
 - 2 bursts of total length $b_2 = (t - 3)m + 3$ bits
 - ⋮
 - ⋮
 - ⋮
 - 1 bursts of total length $b_t = (t - 2t + 1)m + 2t - 1$ bits
- Decoding is only slightly more complex than for BCH codes. In addition to the locations of the errors, their values (m -tuples) must also be determined.

EXAMPLE

For $m = 8$ we can construct the following RS codes (among others)

(n, k)	error corr. ability t	code rate	b_1	b_2
(2040, 1976)	4	.97	25	11
(2040, 1912)	8	.94	57	43
(2040, 1784)	16	.87	121	107

where b_1 is the total length of 1 correctable bursts.

Note: n , k , and b_1 are expressed as numbers of bits.

THE GENERATOR POLYNOMIAL OF AN RS CODE

For an RS code of length $2^m - 1$, symbols are binary m -tuples, i.e., elements of $GF(2^m)$. The generator polynomial of an RS code with minimum distance d has the form

$$g(x) = (x + \alpha)(x + \alpha^2) \cdots (x + \alpha^{d-1})$$

where α^1 is an element of $GF(2^m)$.

Actually, any set of $d-1$ consecutive powers of α can be used. Thus, we could also write

$$q(x) = (x + \alpha^n)(x + \alpha) \cdots (x + \alpha^{d-2})$$

An RS (N, K) code has $d = N - K + 1$.

EXAMPLE OF AN RS CODE

To construct a 3-error-correcting RS code of length 15 with 4-bit symbols, a total of $d - 1 = 6$ check symbols are necessary. The generator polynomial of this (15,9) code is

$$g(x) = (x + \alpha)(x + \alpha^2)(x + \alpha^3)(x + \alpha^4)(x + \alpha^5)(x + \alpha^6) \\ = x^6 + \alpha^{10}x^5 + \alpha^{14}x^4 + \alpha^4x^3 + \alpha^6x^2 + \alpha^9x + \alpha^6$$

Viewed as a binary code, this is a (60,36) code which is capable of correcting any error bursts which are confined to three 4-bit symbols.

As in the binary case, encoding can be accomplished by dividing by $g(x)$.

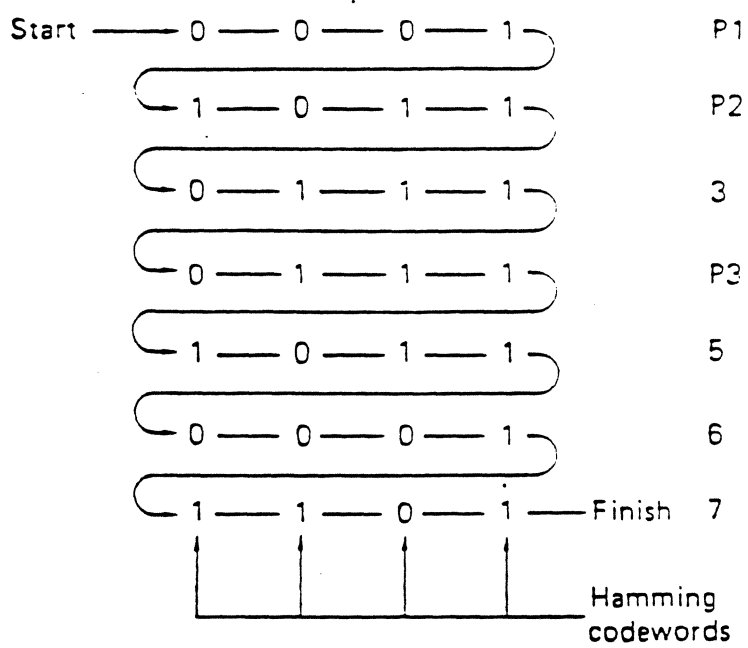
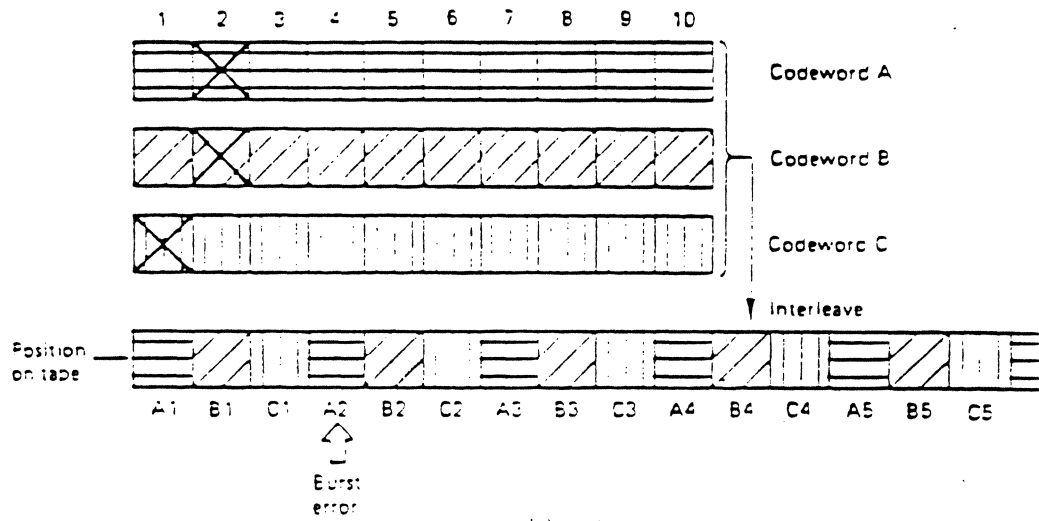
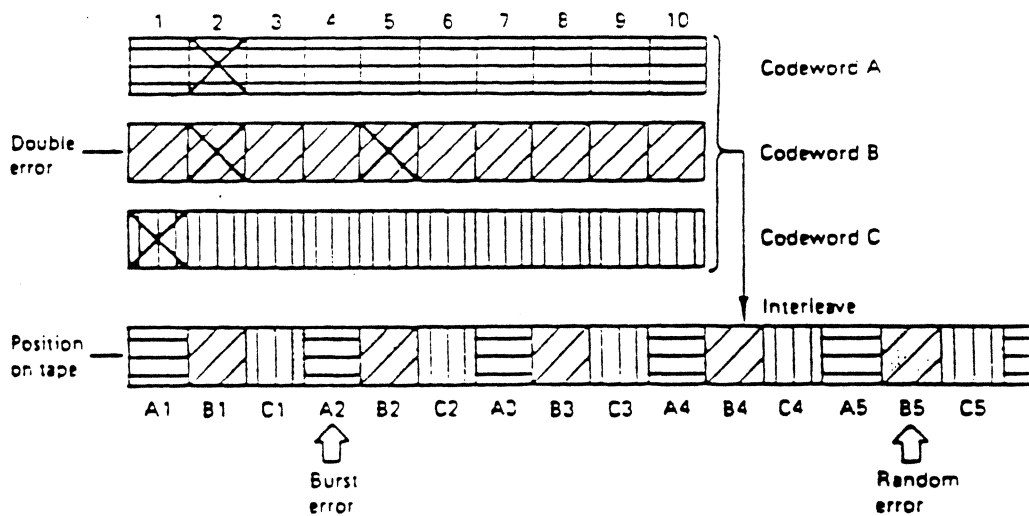


Figure 7.8 The vertical columns of this diagram are all codewords generated by the matrix of Figure 7.6, which can correct a single-bit error. If these words are recorded in the order shown, a burst error of up to four bits will result in one single-bit error in each codeword, which is correctable. Interleave requires memory, and causes delay. Deinterleave requires the same.

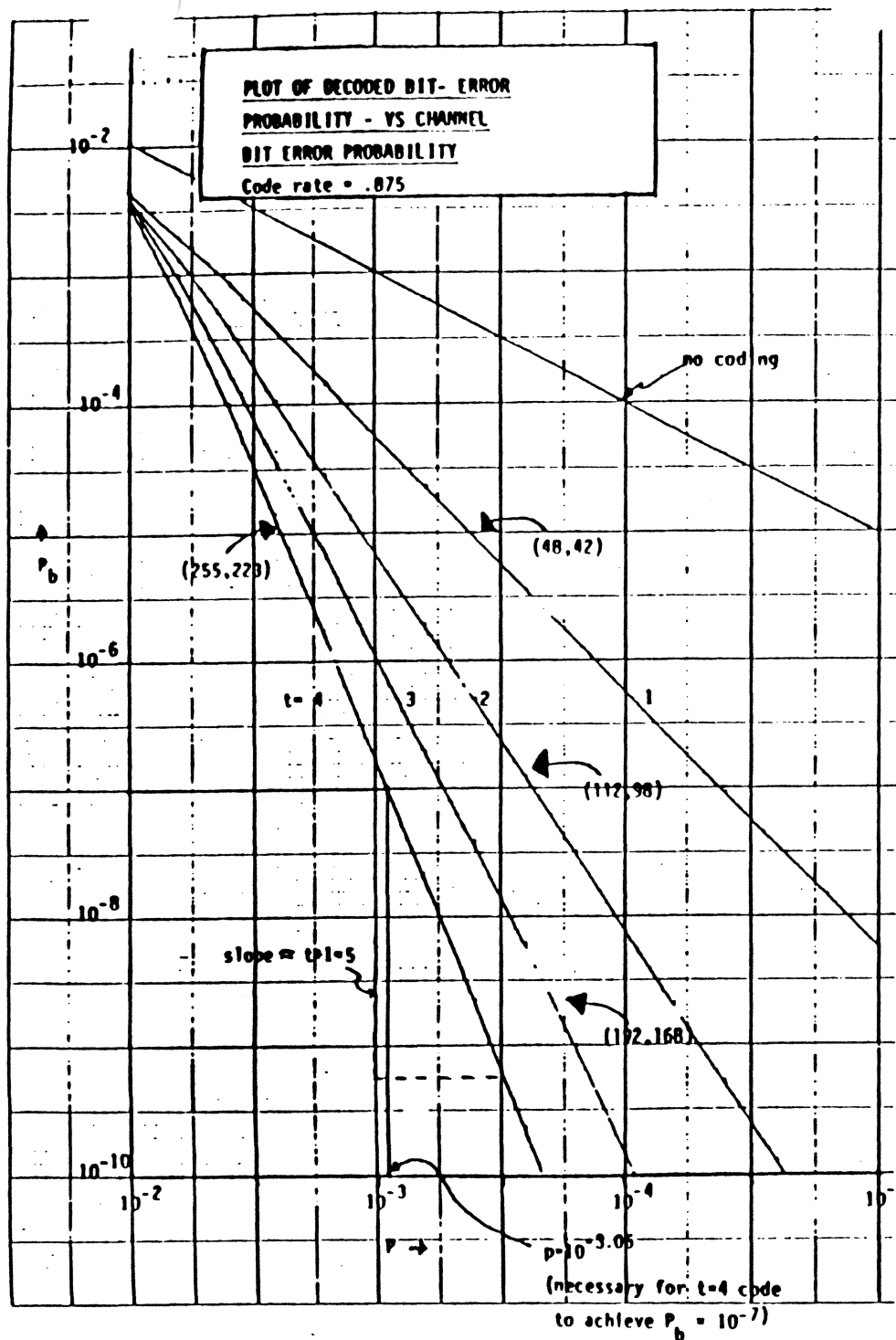


(a)



(b)

Figure 7.27 A1 (a), interleave controls the size of burst errors in individual code words, but at (b) the system falls down when a random error occurs adjacent to a burst.



CODING GAIN

To achieve a desired bit error probability (e.g. $p = 10^{-7}$) an uncoded system requires a certain value of E_b/N_0 (e.g. 11.3 db).

A coded system can achieve the same value of error probability at a lower value of E_b/N_0 . The difference between these two values of E_b/N_0 is referred to as the gain of the code.

Example: To achieve a bit error probability of 10^{-7} , the 4-error correcting (255,223) code requires a channel bit error rate of $p = 10^{-3.06}$. This value of p , in turn, requires a signal-to-noise ratio of $E_b/N_0 = 7.5$ db. Hence

Coding Gain = 11.3 - 7.5 = 3.8db

CODING GAIN OF VARIOUS CODES

$(P_b = 10^{-7})$

R \ t	.875	.75	.5
1	1.8db	1.7 1.3	
2	2.8	2.7 2.6	2.3
3	3.4 3.3	3.3 3.2	2.7 2.4
4	3.8	3.7 3.6	

*Convolutional codes are self-orthogonal majority - logic decodable codes.

ECC IN DISK STORAGE DEVICES

YEAR SHIP	DEVICE	DENSITY		DATA CODE	ECC CAPABILITY	
		bpi	tpi			
1957	350	100	20	NRZI	PARITY CHECK	
196X	13XX	520	50	NRZI	CRC CHECK	
	23XX	1020	50	FM	MULTIPLE ERROR DETECTION	
197X	33XX	2200	100		4000- 15000	200- 800
					<u>MODIFIED FIRE CODE</u>	
					CORRECT./DETECT. = OF CHECK BYTES	
1971	3330	4040	192	MFM	11 BITS / 22 BITS 7 / RECORD	
1973	3340	5636	300	MFM	3 BITS / 11 BITS 6 / RECORD	
1976	3350	6425	475	MFM	4 BITS / 10 BITS 6 / RECORD	
					<u>MODIFIED REED-SOLOMON CODE</u>	
1979	3370	12134	635	(2,7)	9 BITS / 17 BITS 9 / BLOCK	
1980	3375	12134	800	(2,7)	9 BITS / 17 BITS 12 / RECORD	
1981	3380	15200	800	(2,7)	17 BITS / 33 BITS 12 / RECORD	

ECC IN TAPE STORAGE PRODUCTS

YEAR SHIP	DEVICE	DENSITY	ECC CAPABILITY
195X	72X	100-556 bpi	PARITY TRACK & LRC
1952	<u>729</u>	800 bpi NRZI 9 TRACKS/1/2"	PARITY TRACK & CRC ONE TRACK CORRECTION ON REREAD
1956	<u>24XX</u>	1600 bpi PE 9 TRACKS/1/2"	PARITY TRACK & PE ENCODING (CRC) ONE TRACK CORRECTION ON-THE-FLY
1973	<u>3420</u>	6250 bpi GCR 9 TRACKS/1/2"	GCR GROUP CODED RECORDING (GCR) PARITY TRACK & 1 CHECK BYTE/7 BYTES TWO TRACK CORRECTION ON-THE-FLY
1975	3850 <u>MSS</u>	6888 bpi ZM 67 tpi	INTERLEAVED SUBFIELD CODE WORD = 15 SECTIONS. SECTION = 16 BYTES REDUNDANCY 2 SECTIONS/15 SECTIONS ON-THE-FLY CORRECTION OF 2 SECTIONS
198X	<u>3480</u>	-22000 bpi (0.3) 18 TRACKS/1/2"	MULTI-TRACK CORRECTION ADAPTIVE CROSS PARITY CODE

DISK ARRAYS

MIRRORING

2N

(Also called shadowing or dual copy)

- Complete duplication of data, i.e., disks paired
- Read from disk with shortest access (have twice as many paths to data)
- Twice the number of writes

Used in transaction processing where fault tolerance essential

Notes: Disk drives indicate when they become inoperative
Hot spares used for availability and minimum MTTFDL

PARITY DISKS

N+1

Byte interleaved:

- Array viewed as single logical disk
- Parallel readout for high data rate - striping (Synchronized spindles)

Block interleaved:

- Second write to update parity (therefore rotate parity among disks to minimize delays)

Improving Reliability with Redundancy

- Add redundant drives to handle failures

Redundant

Array of

Inexpensive

Disks

- Redundancy offers 2 advantages:
 - 1) Data not lost: Reconstruct data onto new disk
 - 2) Continuous operation in presence of failure

- Several RAID organizations
 - Mirroring / Shadowing
 - ECC
 - Parity

Types of Disk Failures

Transient

⇒ retry or ECC modification

Media

⇒ ECC modification, factory remapping,
read after write verification

Electrical

⇒ end-to-end checksum, complex interface
protocol, internal buffer & path ECC

Mechanical

⇒ hope you can scavenge data

Disk Lifetime Distribution



- exponential lifetime good first guess
⇒ desparately seeking better data
- also assume independent, but ...
support hardware
bad batches
human operators, maintenance

Lifetime Parameters

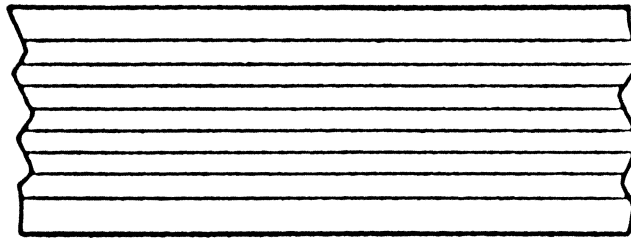
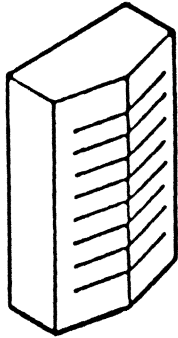
- temperature
- age
- power cycles
- seek events
- read/write events

Effects of Arrays on Failures

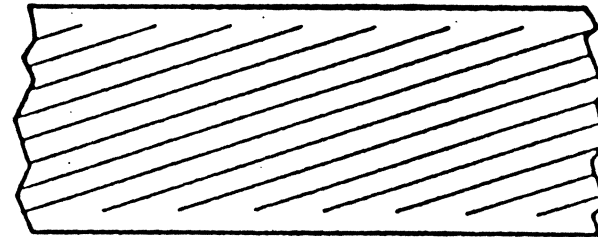
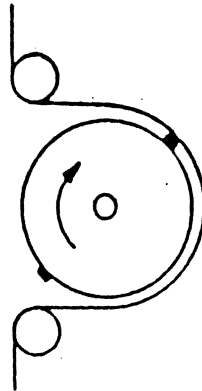
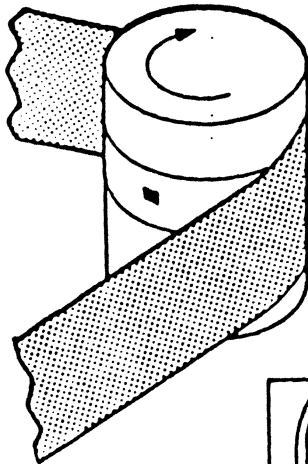
- higher data rates
⇒ higher misdetection/mis correction rates
⇒ better error detection
- striped files
⇒ more files exposed to each failure
⇒ encourages failure tolerance
- more disks
⇒ more frequent failures
⇒ automatic isolation and easy repair
- larger total capacity
⇒ more to backup
⇒ reduce frequency of backups

TYPES OF MAGNETIC RECORDING FORMATS

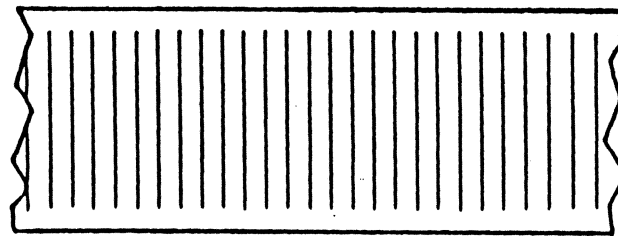
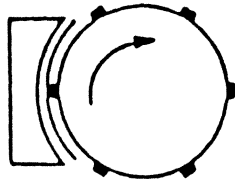
LONGITUDINAL



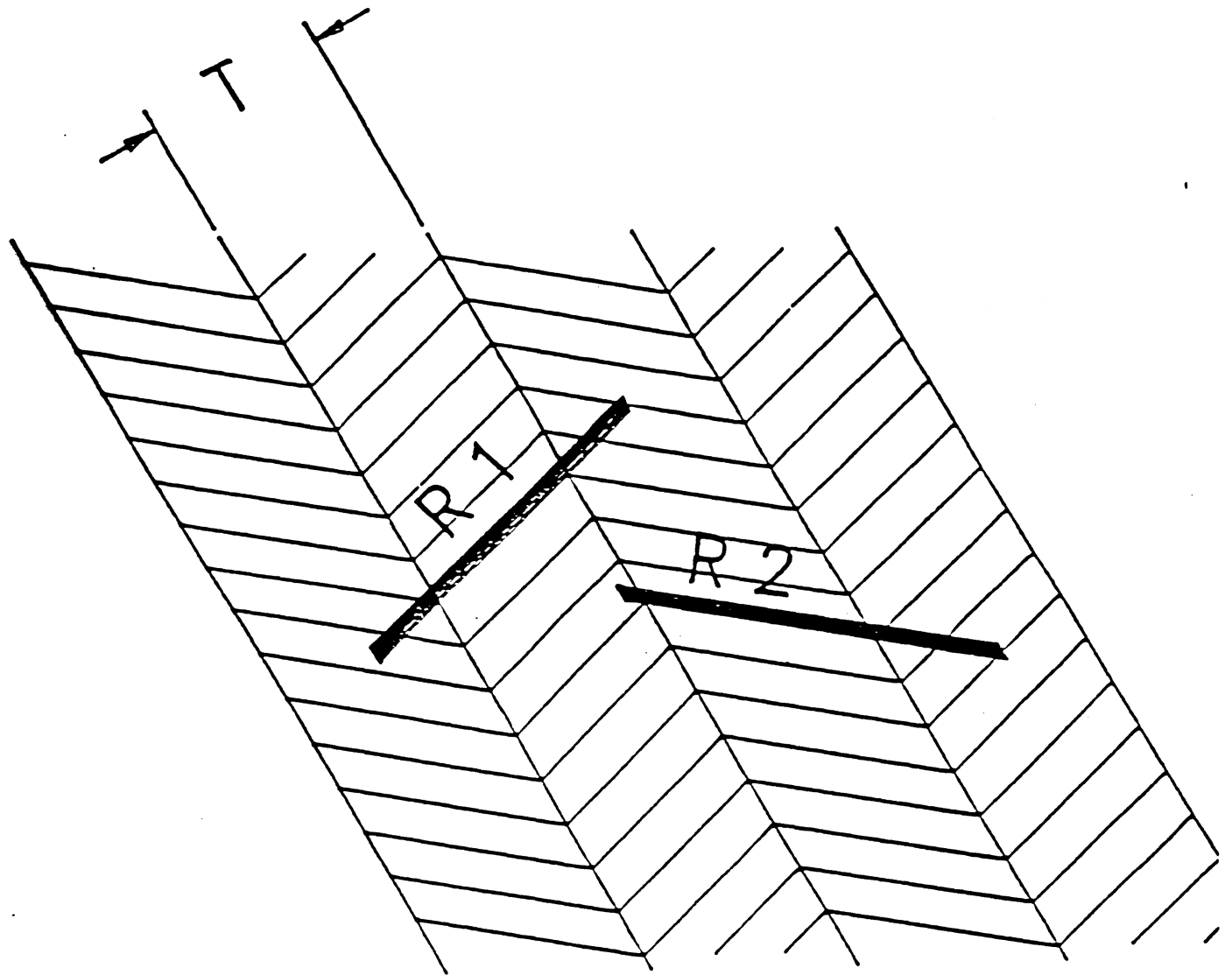
HELICAL



TRANSVERSE



AMPEX

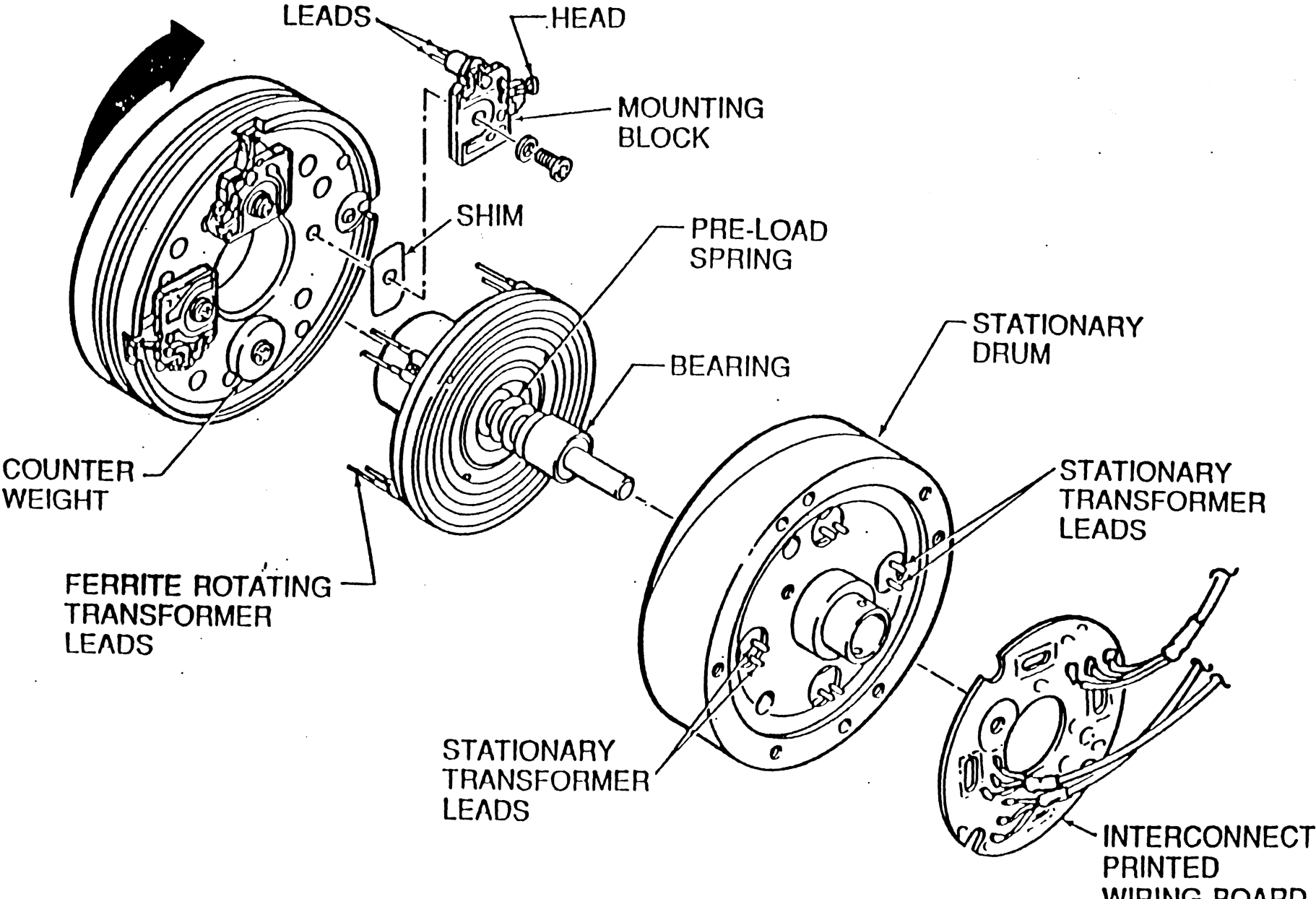


$$R1, R2 < 3T$$

ALTERNATE AZIMUTH
READ

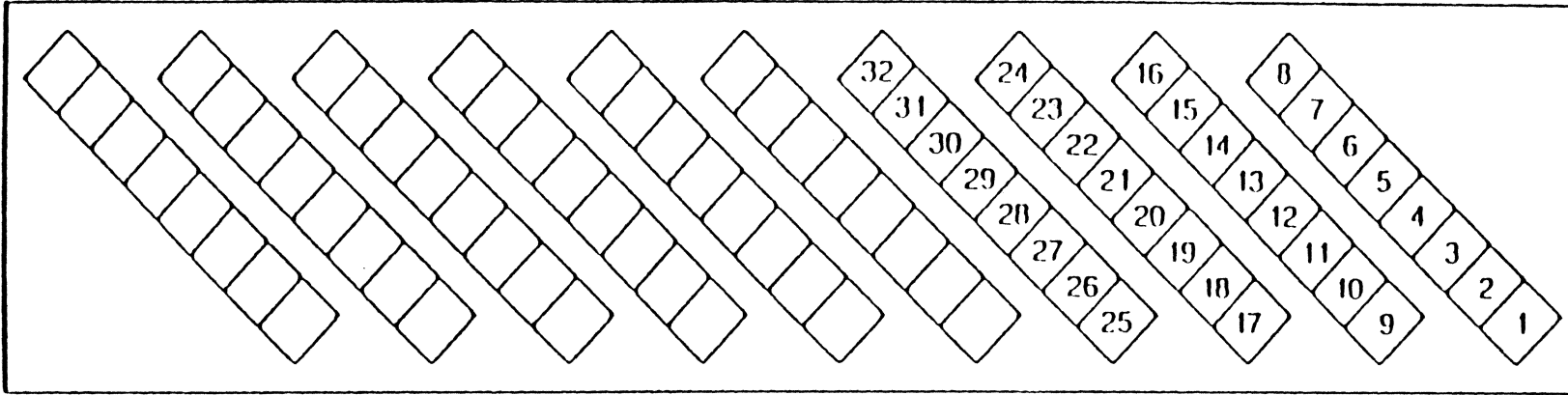
IIST

HELICAL SCAN HEAD ASSEMBLY

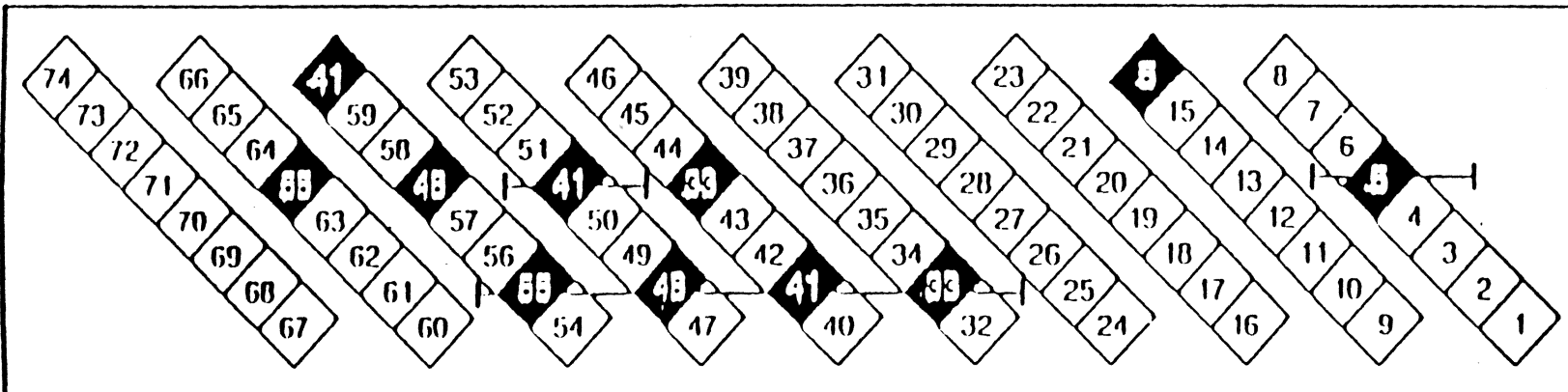


Read After Write

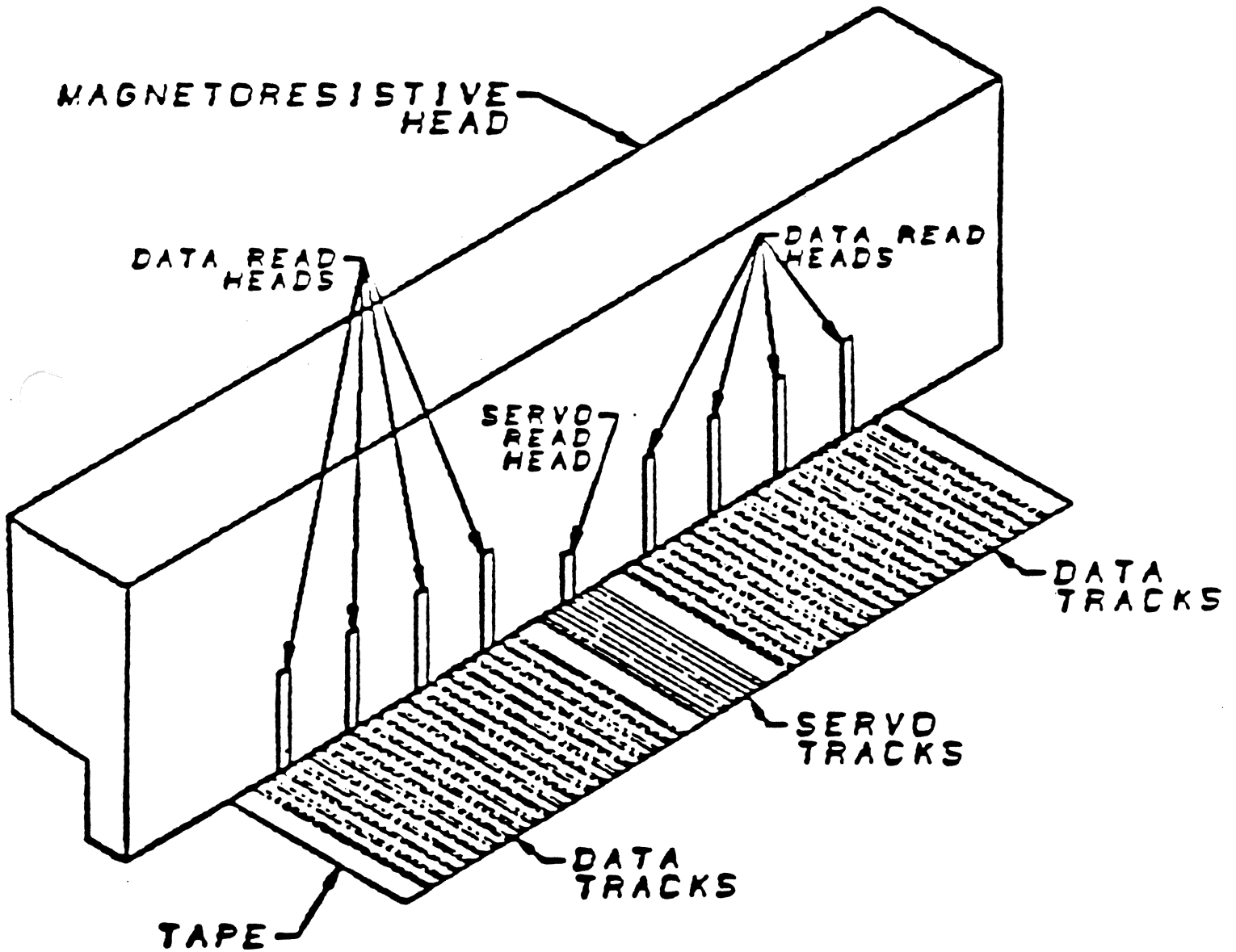
Normal Sequencing

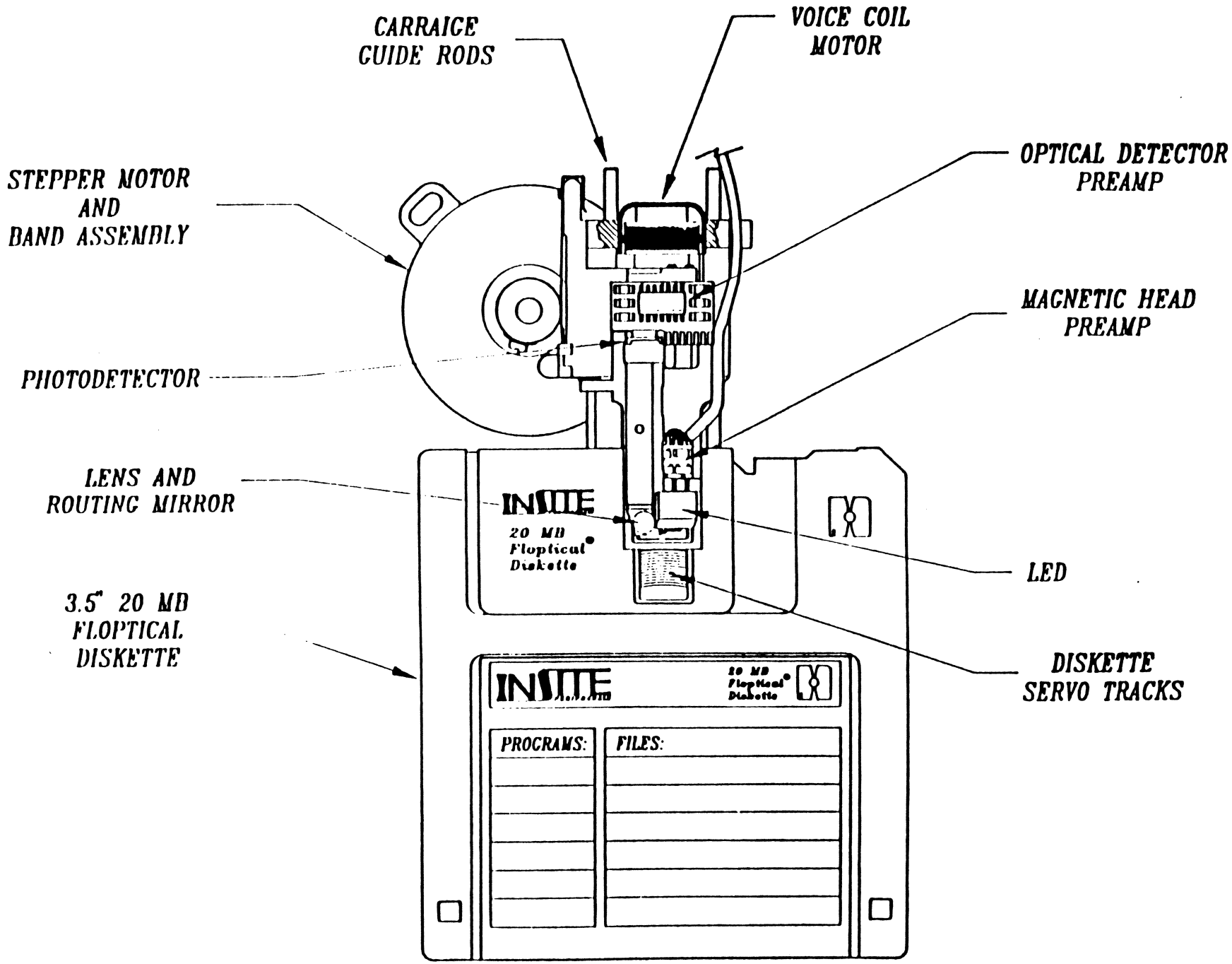


Rewrite Sequencing



MCR HEAD/TAPE FORMAT





OPTICAL RECORDING

Geoffrey Bate

IIST "Data Storage Tutorial"
May 24, 1991

Optical Data Storage

TABLE OF CONTENTS

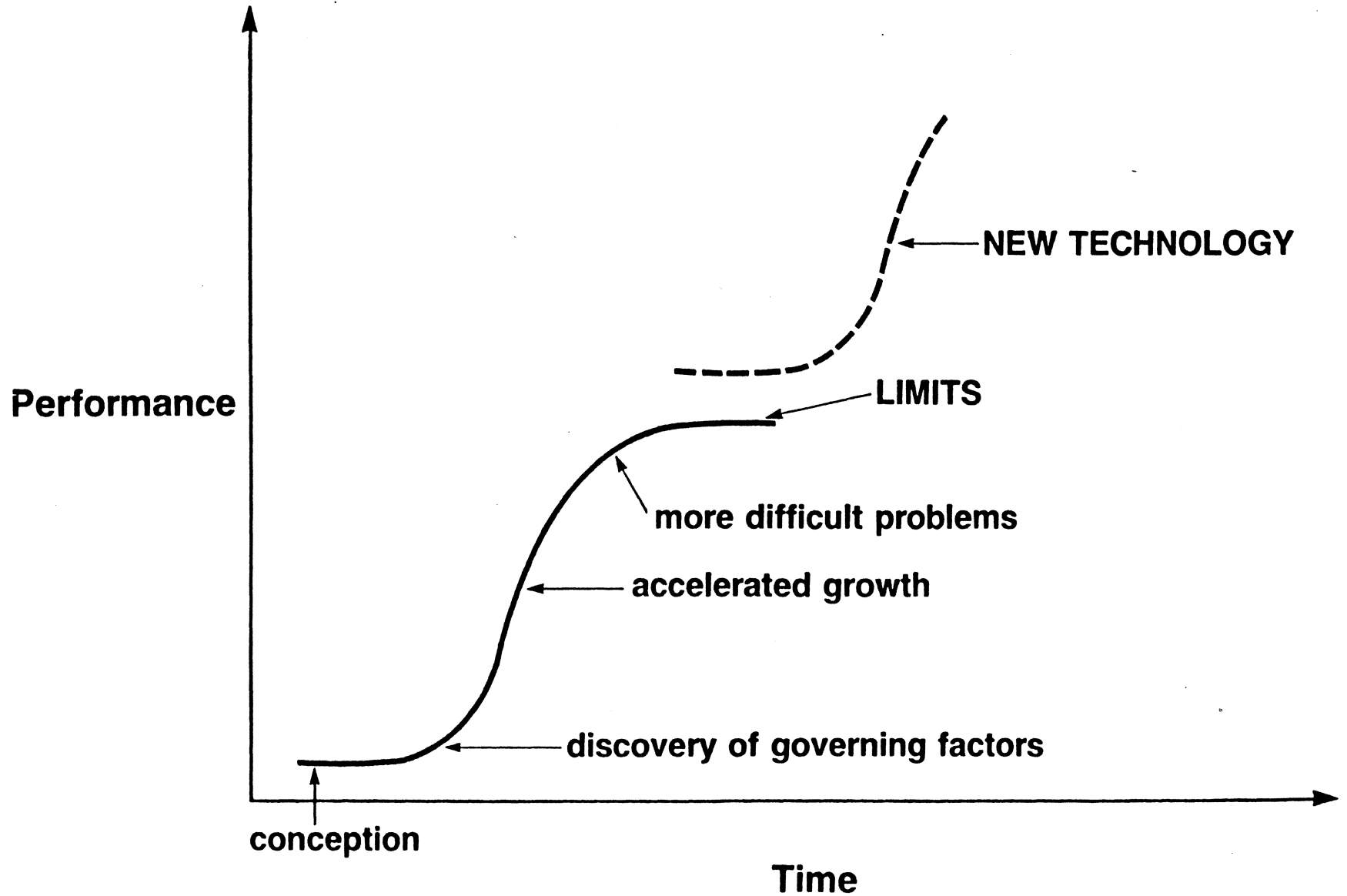
- * COMPARISON OF OPTICAL AND MAGNETIC RECORDING
- * OPTICAL COMPONENTS
 - laser diodes
 - lenses, beam-shaping prisms, beam splitters, holographic lenses
- * AUTOMATIC TRACKING AND FOCUSING OF THE BEAM
- * READ-ONLY DISKS: CD, CD-ROM
- * RECORDABLE OPTICAL DISKS
- * WRITE-ONCE DISKS
- * REVERSIBLE DISKS:
 - phase-change
 - dye-polymer
 - thermo-magneto-optic (TMO)
- * CURIE-POINT AND COMPENSATION-POINT RECORDING
- * ORIGIN OF MAGNETO-OPTIC EFFECTS
- * ENHANCEMENT OF KERR MAGNETO-OPTIC ROTATION
- * PROPERTIES OF BARE EARTH - TRANSITION METAL AMORPHOUS FILMS
- * SPUTTERING OF TMO FILMS
- * DESIRABLE PROPERTIES OF TMO FILMS
- * FACTORS LIMITING THE SIZE OF WRITTEN BITS
- * COMMERCIAL TMO DISKS AND DRIVES
- * FLEXIBLE OPTICAL DISKS AND TAPES
- * ISSUES

MAGNETIC RECORDING

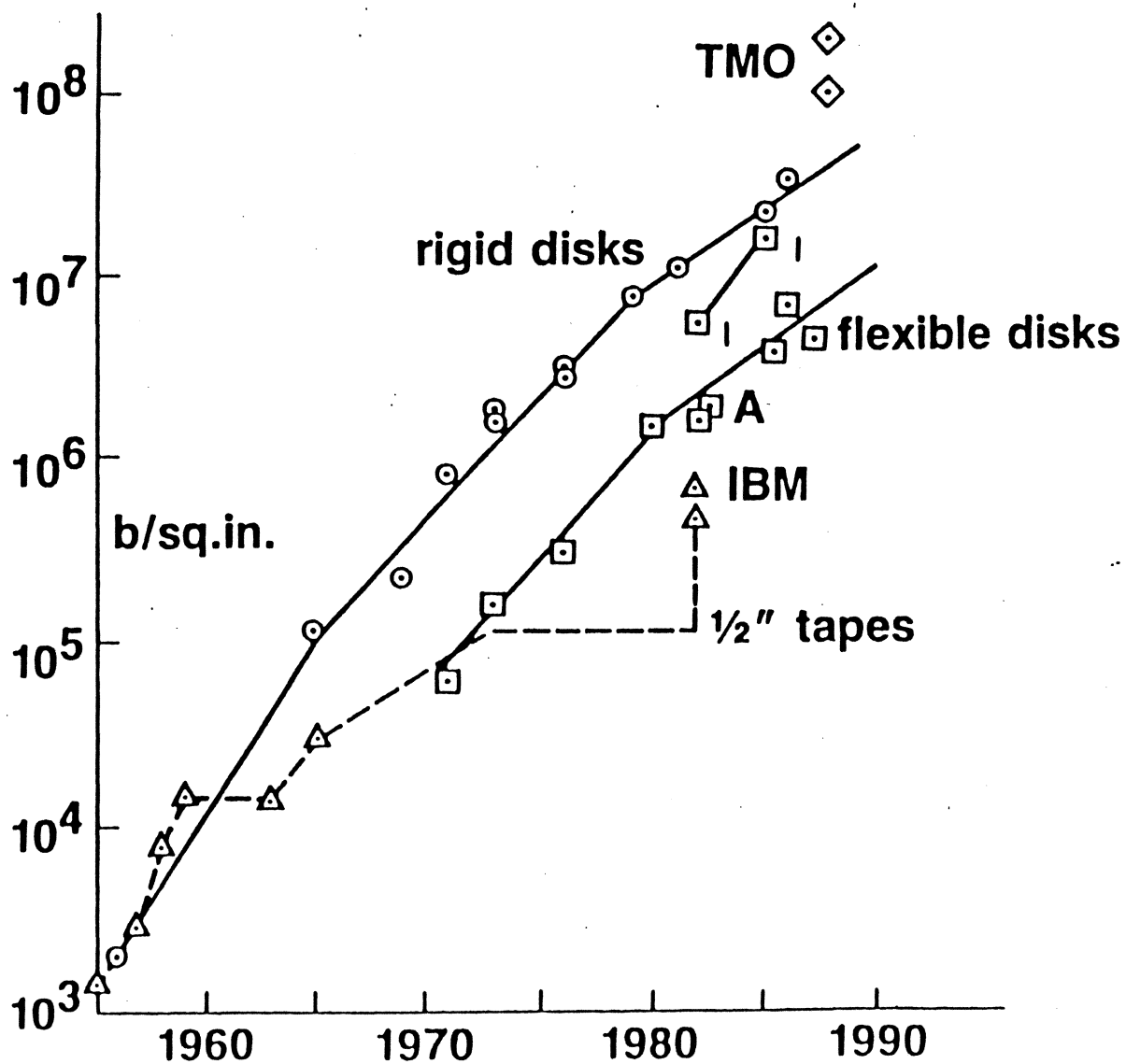
- SIMPLE - NO CRITICAL COMBINATION OF INGREDIENTS NEEDED
- RELIABLE
- INEXPENSIVE MEDIA
- NO PROCESSING NEEDED
- (REMOVABLE)
- INFORMATION STORED IN A STABLE STATE OF THE MEDIUM
- UPDATABLE
- INFINITELY ERASABLE AND REVERSIBLE

G. BATE

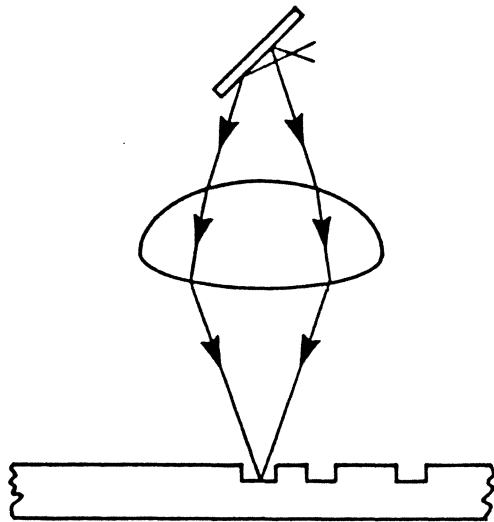
Technology Development



Areal Density versus Time

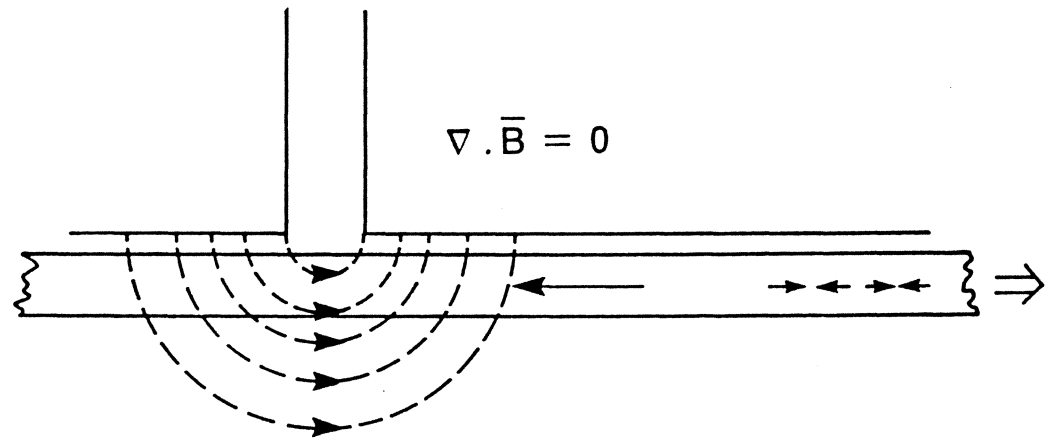


OPTICAL

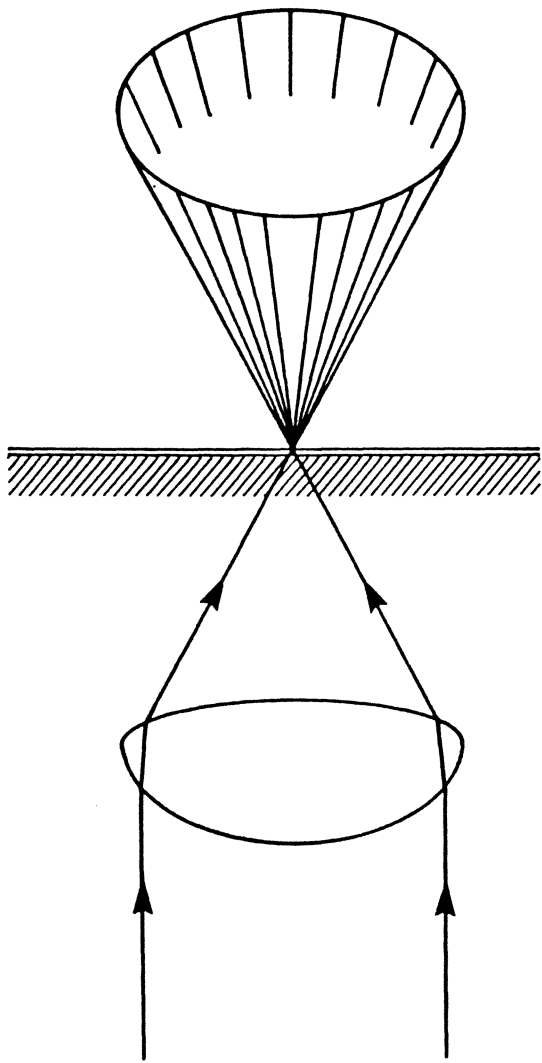


■ Large Separation

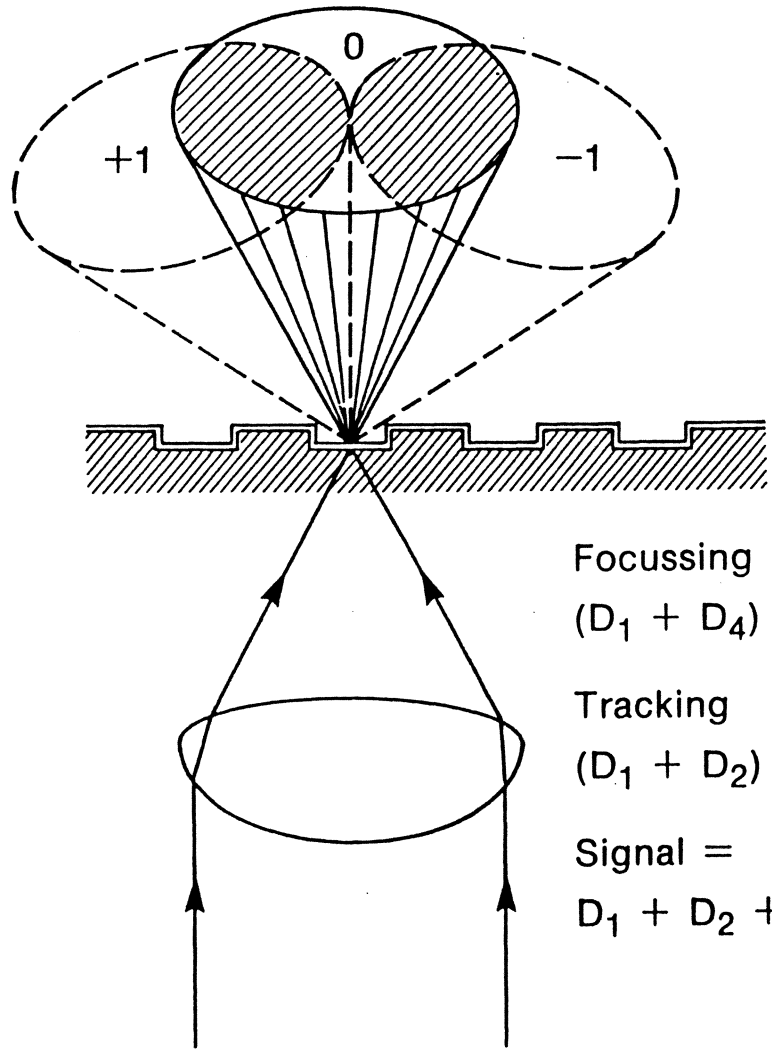
MAGNETIC



■ Very Small Head-to-Disk Separation



D_1 D_2 D_3 D_4



Focussing

$$(D_1 + D_4) - (D_2 + D_3) = 0$$

Tracking

$$(D_1 + D_2) - (D_3 + D_4) = 0$$

Signal =

$$D_1 + D_2 + D_3 + D_4$$

OPTICAL vs. MAGNETIC RECORDING

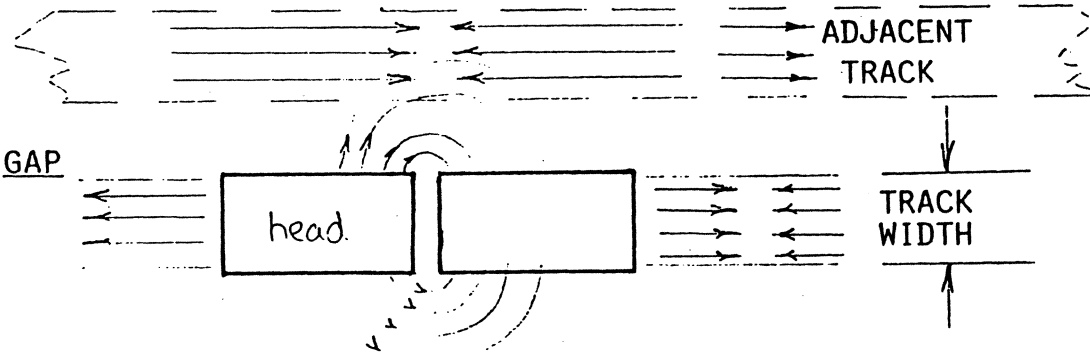
ADDITIONAL DIFFERENCES

- MAGNETIC RECORDINGS ARE ALWAYS "SWITCHED ON"
(OPTICAL RECORDINGS - INACTIVE UNTIL SWITCHED ON; LIKE DNA
- ONLY THE ILLUMINATED BIT CONTRIBUTES TO THE SIGNAL)

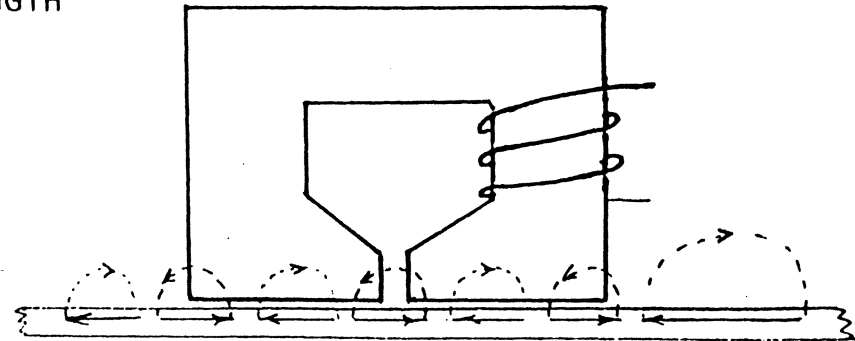
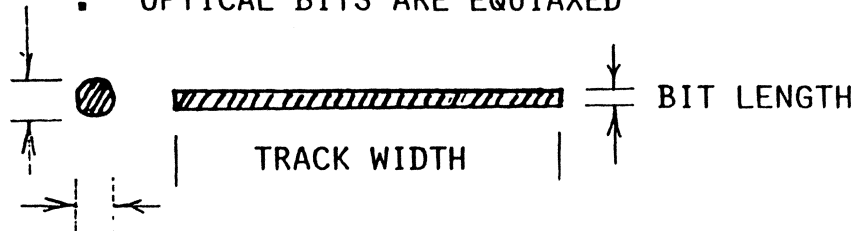
- WHOLE RECORDED AREA OF MEDIUM EMITS FLUX THAT CAN BE PICKED UP BY THE READING HEAD.

- MAGNETIC HEAD READS SIDWAYS
AS WELL AS DOWNWARDS

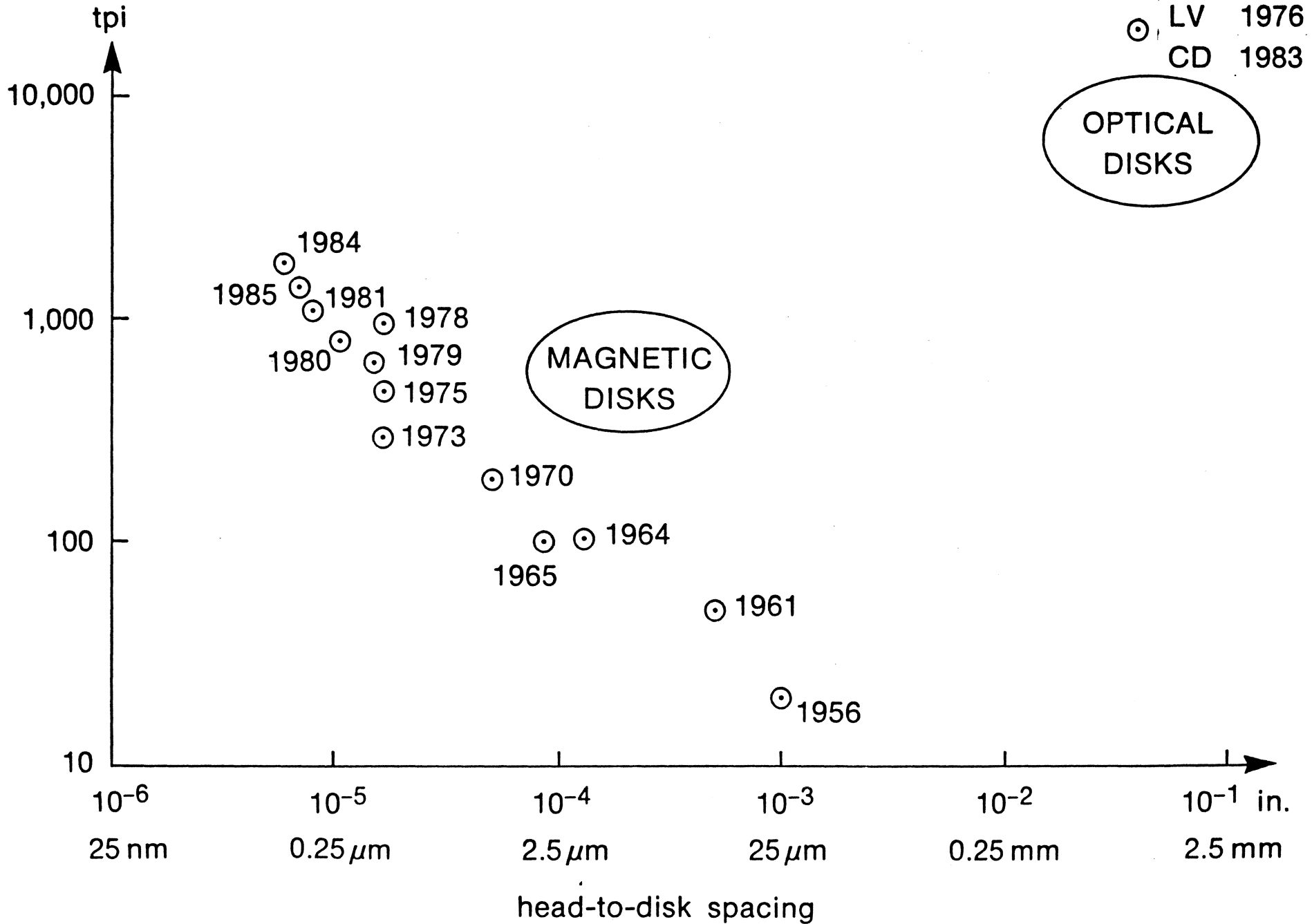
- SENSITIVE AREA NOT LIMITED TO GAP



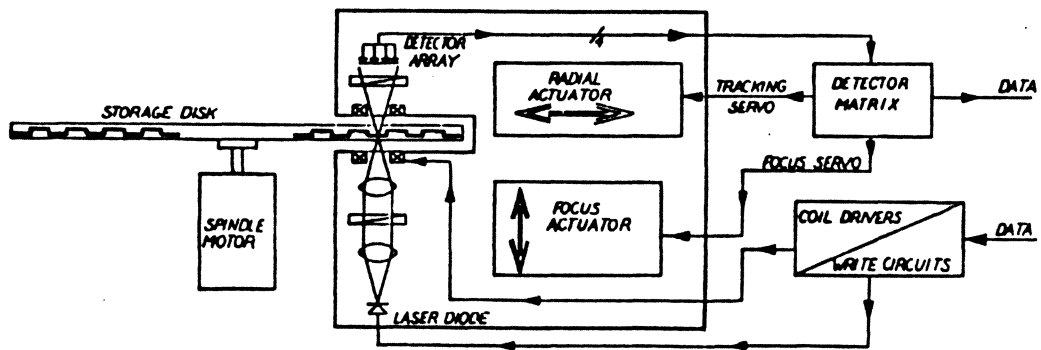
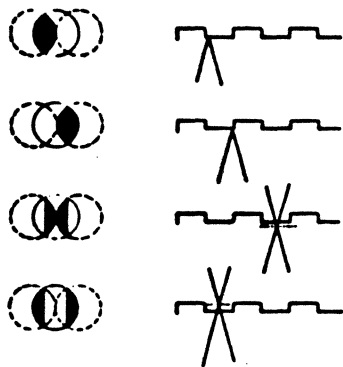
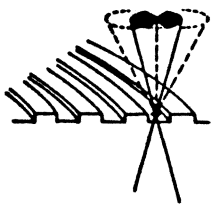
- OPTICAL BITS ARE EQUIAXED



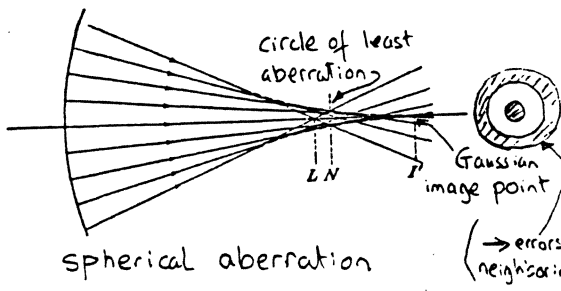
TRACK DENSITIES OF OPTICAL AND MAGNETIC DISKS VERSUS HEAD-TO-DISK SPACING



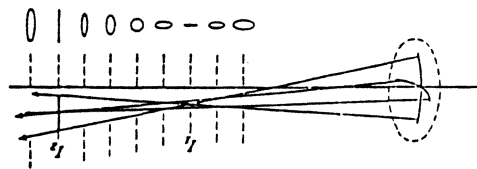
MAGNETO-OPTICAL DRIVE



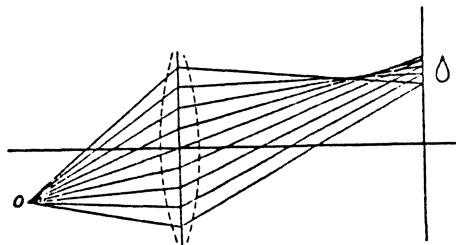
LENSES : problems



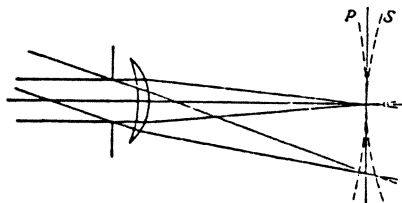
spherical aberration



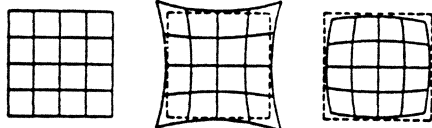
astigmatism (off-axis)



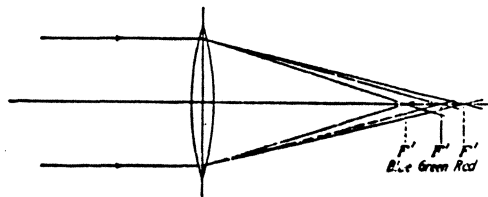
coma (off axis)



curvature of image field (off-axis)



pincushion barrel distortion (off-axis)

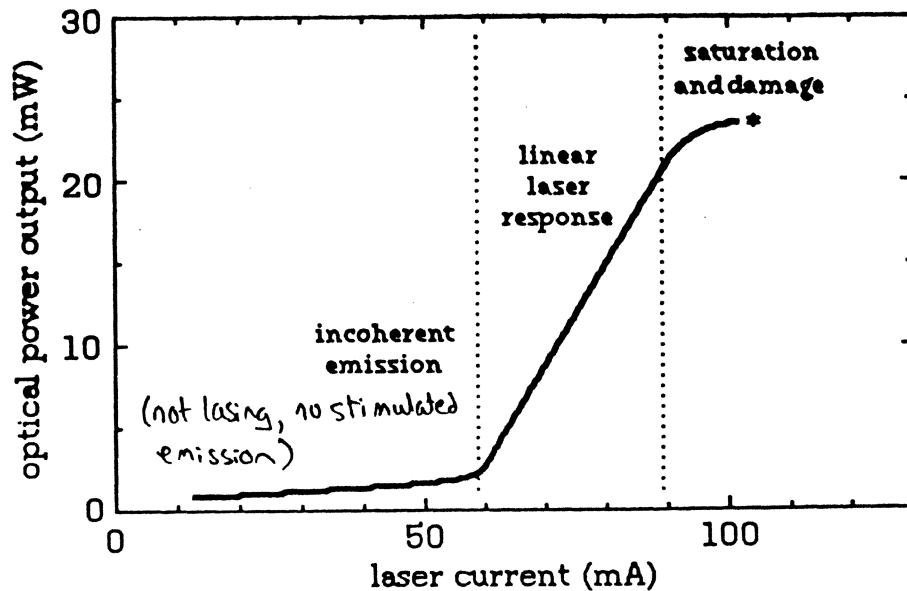
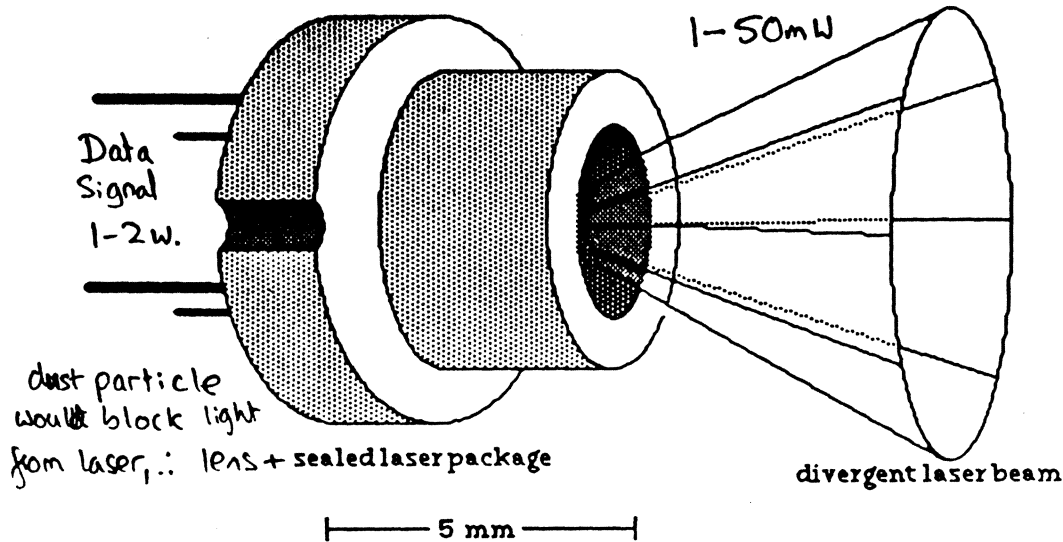


chromatic aberration

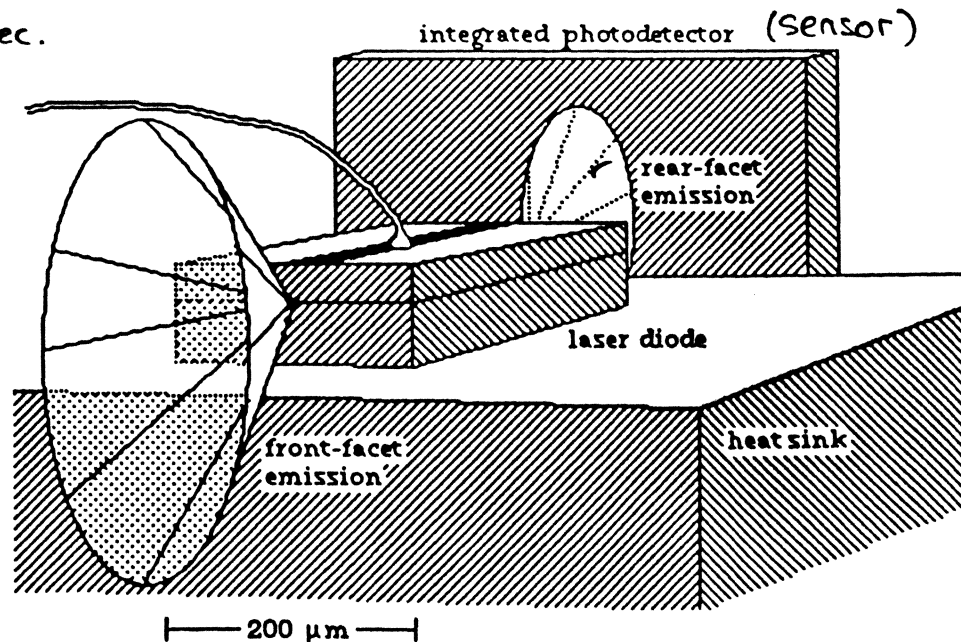
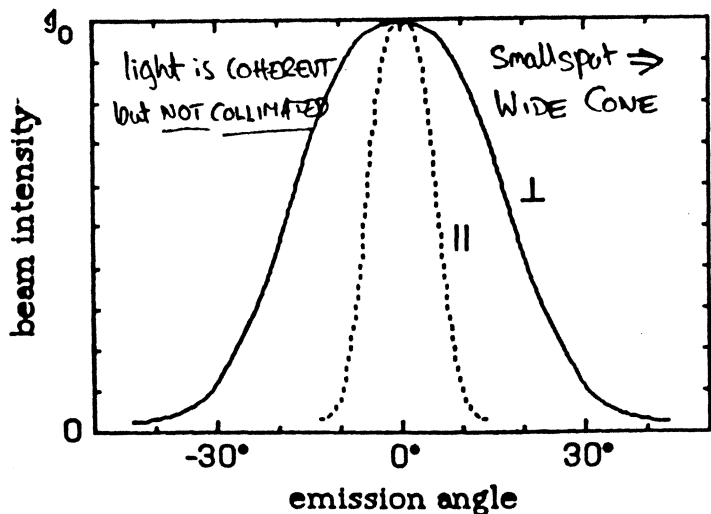
LASER DIODES 1962 - used.

1970 - CW.

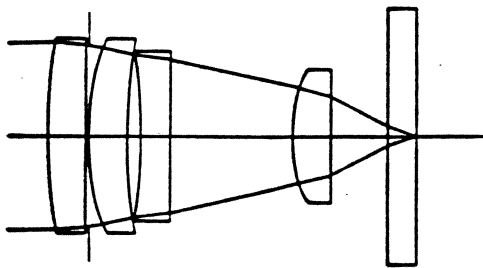
1980 - used in fiber-optic communication
- in optical disks



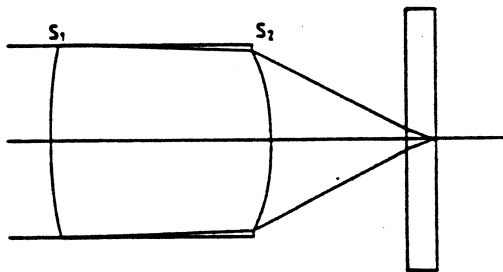
input power : 100 - 200 mW.
directly modulatable; data rates 1-3 MBytes/sec.



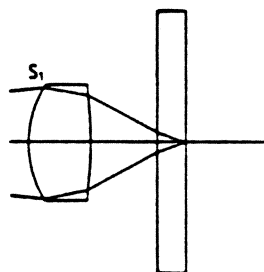
Lifetime $\geq 10,000$ hr. (laser diodes are the most sensitive and delicate of semiconducting device - must be limited to prevent burn-out. O.K. to exceed nominal 20 mW up to 30 mW for nano s



Focal distance 8 mm
 NA 0.45
 Field diameter 1.0 mm (OPD at 400 μm from the centre is 0.052 λ)
 Field curvature 6 μm at the rim
 Diameter 7.5 mm
 Length 11 mm
 Weight < 2 g (mount included)

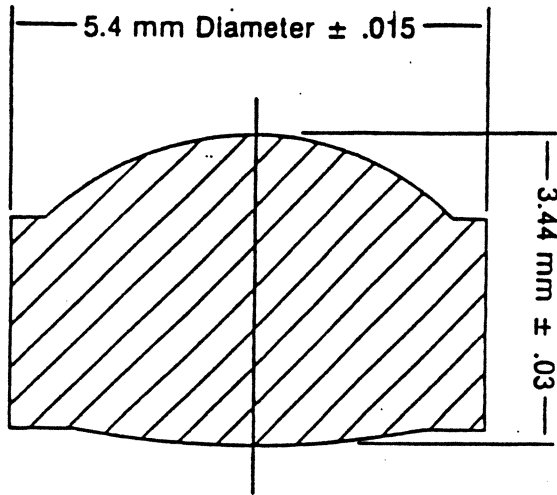


Focal distance 8 mm
 NA 0.45
 Field diameter 2.0 mm (OPD at 800 μm from the centre is 0.047 λ)
 Field curvature 24 μm at the rim
 Asphericity (S_1) $\pm 25 \mu\text{m}$ (with respect to the best fit sphere)
 Asphericity (S_2) $\pm 2 \mu\text{m}$
 Diameter 7.5 mm
 Thickness 8.5 mm
 Weight < 2 g (mount included)



Focal distance 3.7 mm
 NA (object side) 0.10
 NA (image side) 0.45
 Field diameter 0.2 mm (OPD at 80 μm from the centre is 0.045 λ)
 Field curvature 2 μm at the rim
 Asphericity (S_1) $\pm 11 \mu\text{m}$ (with respect to the best fit sphere)
 Diameter 4.1 mm
 Thickness 3.2 mm
 Optical throw 26.0 mm
 Weight < 0.25 g (mount included)

Coming Molded Glass Aspheric Lens

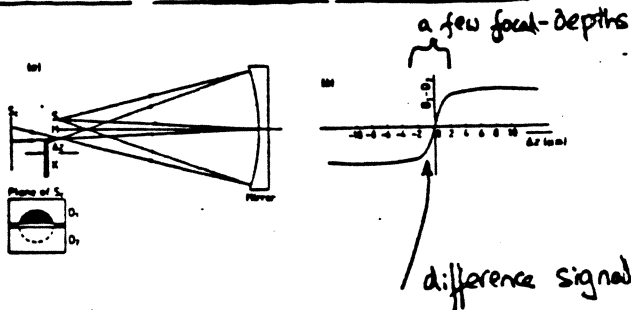


Diameter: less than 6 mm
 Operating wavelength: 780 nm
 Effective focal length: 4.47 mm
 Entrance pupil diameter: 4.2 mm
 Magnification: infinite conjugate
 Working distance: 1.85 mm
 Field of view: 0.150 mm
 Mass: 0.24 g
 RMS wavefront aberration: less than 0.05 λ at full field

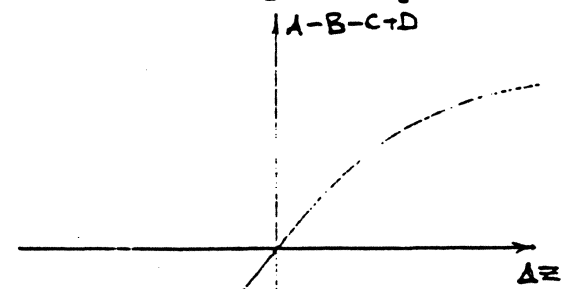
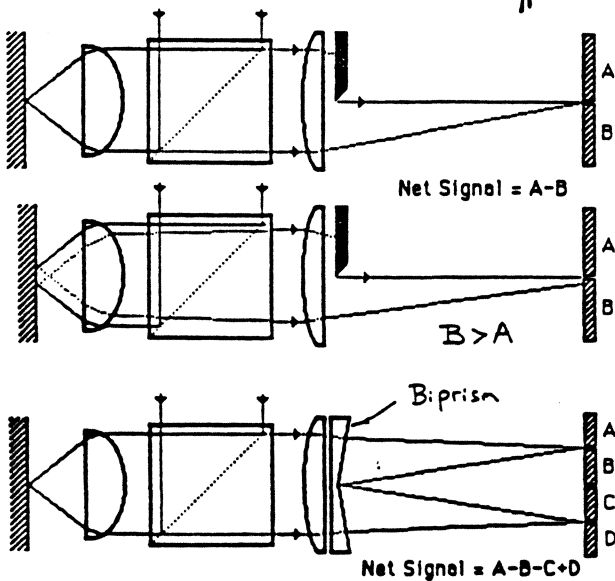
$$\theta = 48.62^\circ$$

$$NA = \sin \theta = 0.75$$

FOCUS SENSING METHODS



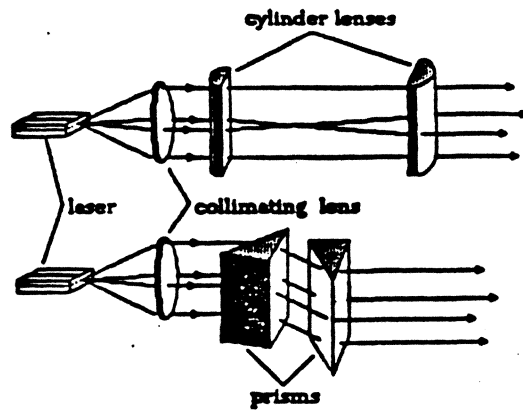
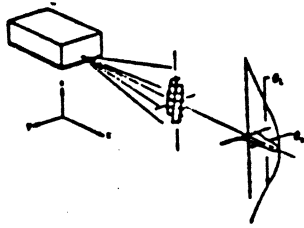
Detectors D_1 and D_2 are equally illuminated only when the knife-edge is at the focal point of the mirror, i.e. $\Delta z = 0$



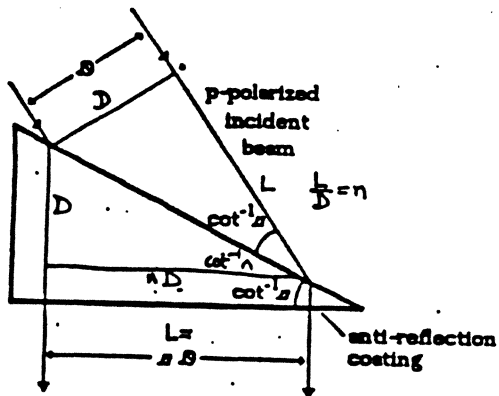
BIPRISM VARIATION

BEAM SHAPING

- want circular not elliptical beams
- $L \because$ lenses are circular
- \therefore use ANAMORPHIC SYSTEMS.



Prism Details

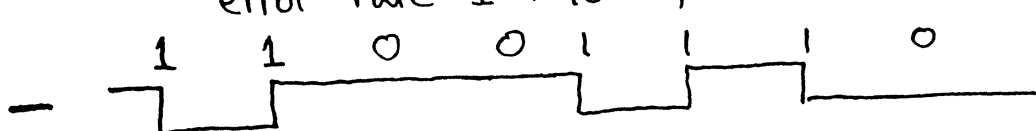


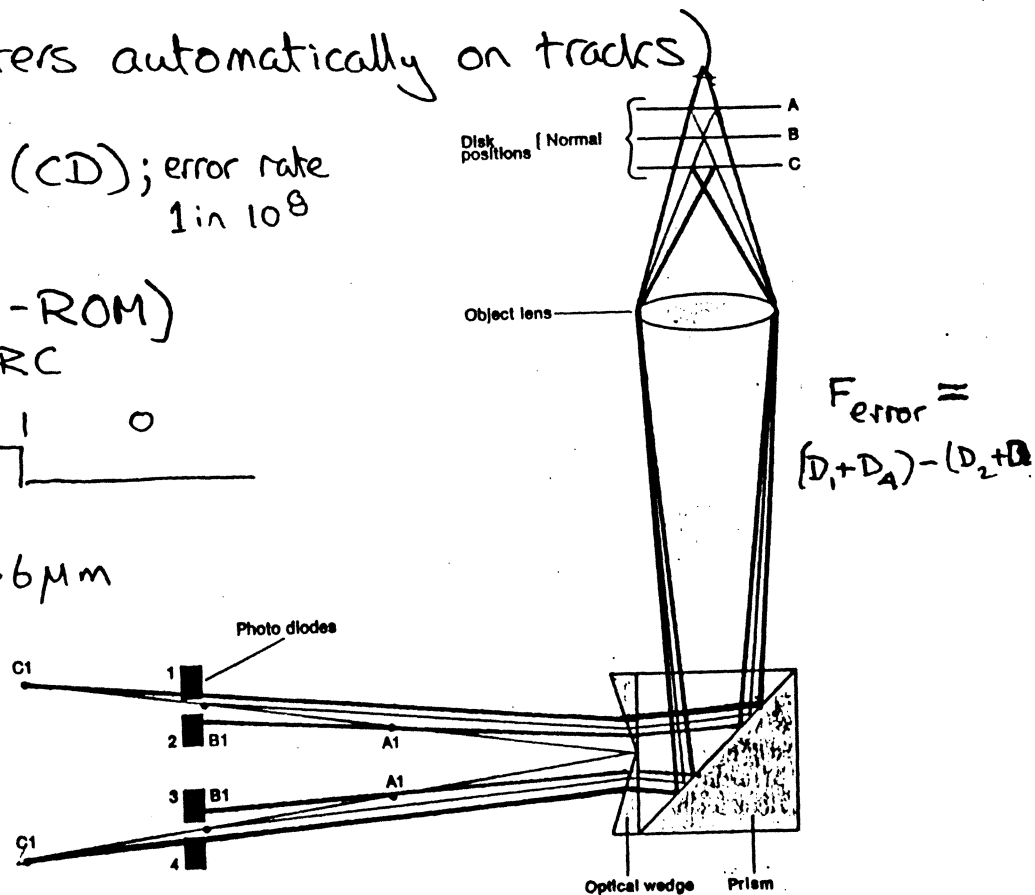
- light incident on prism at Brewster's Angle, all light is transmitted
- result is an expansion of the beam, in the direction shown, by a factor of n
- using two prisms, the beam emerges parallel to its original path.

LASERS: Future

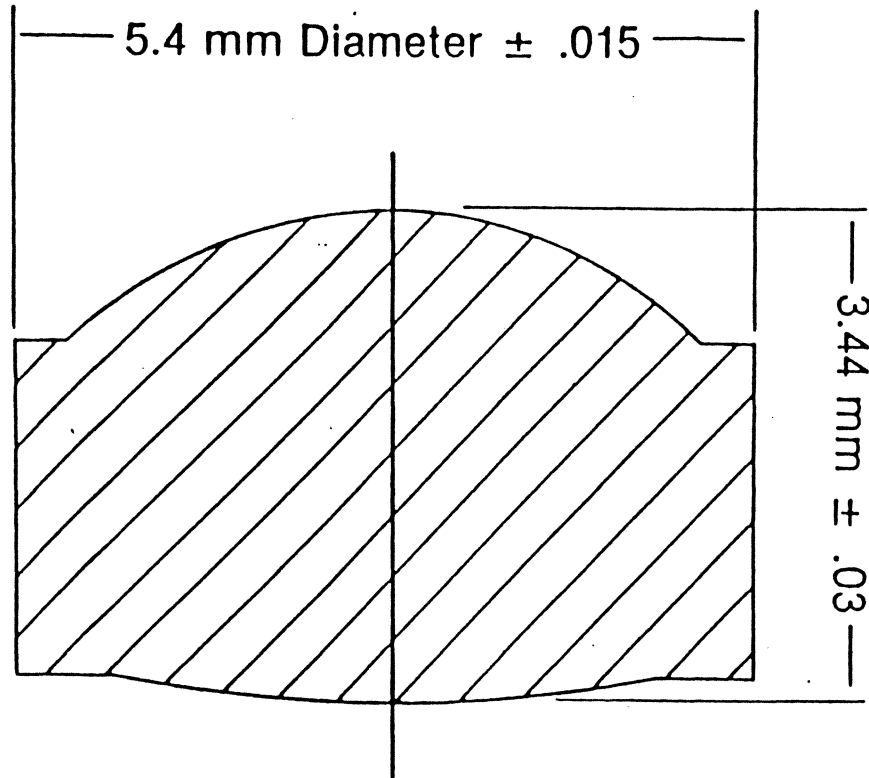
- Extending and improving current technology should allow the reliable, CW powers to be increased from 20-30 mw. to 50 mw.
- If non-absorbing mirrors prove practical, it may be possible to make single-element laser diodes that operate reliably at 100 mw.
- Phase-locked arrays have potential to exceed 100 mw. but have Formidable Problems.
 - arrays not yet demonstrated -
 - Diffraction - Limited,
 - Stable, Single-Beam in C.W. operation
 - long-term reliability is unknown (dissipation of heat more difficult than in single lasers)
 - failure of one laser could alter beam's pattern.
- New Ideas include:-
 - Single laser diode with etched convex mirrors acting as unstable resonator.
 - Coupling of individual lasers via external cavity so that they operate as a single source.

COMPACT DISK CD, 120mm.

- after molding, evaporate a thin film of Al on top of the pitted surface. Must stay FLAT to $\pm 0.4\mu\text{m}$ at $T \leq 55^\circ\text{C}$ surface.
- apply label (on the other side)
- apply protective lacquer
- punch hole (punch centers automatically on tracks)
- 70 mins. recorded music (CD); error rate $1 \text{ in } 10^8$
- 500 - 600 MB data (CD-ROM)
error rate $1 \text{ in } 10^{15}$, CIRC
- 
- track width = $0.6\mu\text{m}$, pitch = $1.6\mu\text{m}$
625. L/mm ; 15,875 tpi
- $\text{b/mm} = 1,378$; 35,000 bpi
- (3, 11) code



Corning Molded Glass Aspheric Lens



Diameter:	less than 6 mm
Operating wavelength:	780 nm
Effective focal length:	4.47 mm
Entrance pupil diameter:	4.2 mm
Magnification:	infinite conjugate
Working distance:	1.85 mm
Field of view:	0.150 mm
Mass:	0.24 g
RMS wavefront aberration:	less than 0.05λ at full field

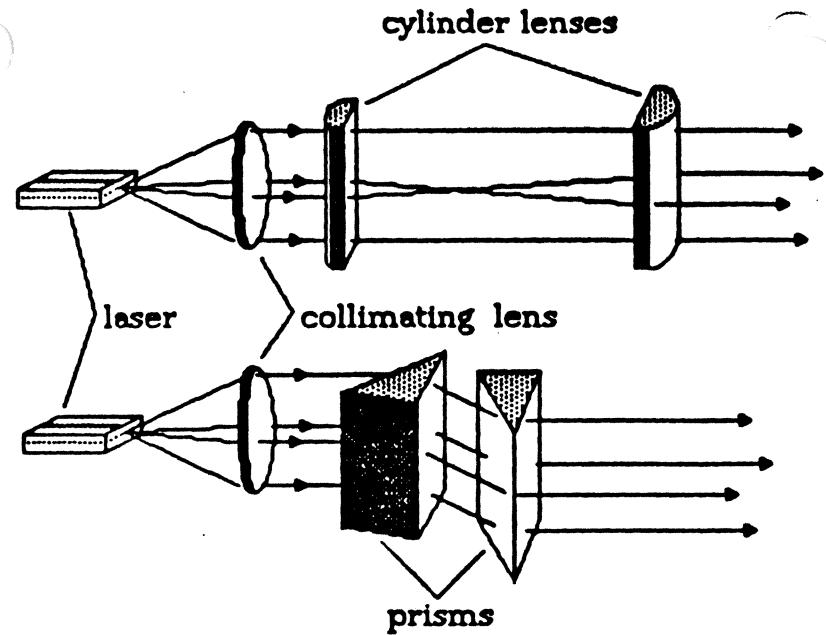
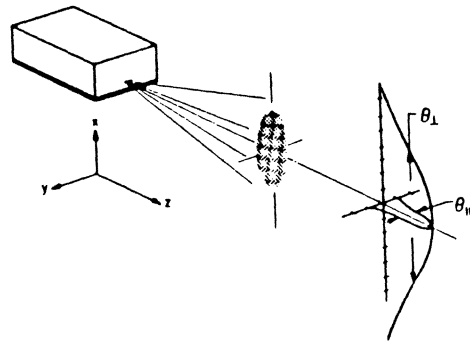
$$\theta = 48.62^\circ$$

$$NA = \sin \theta = 0.75$$

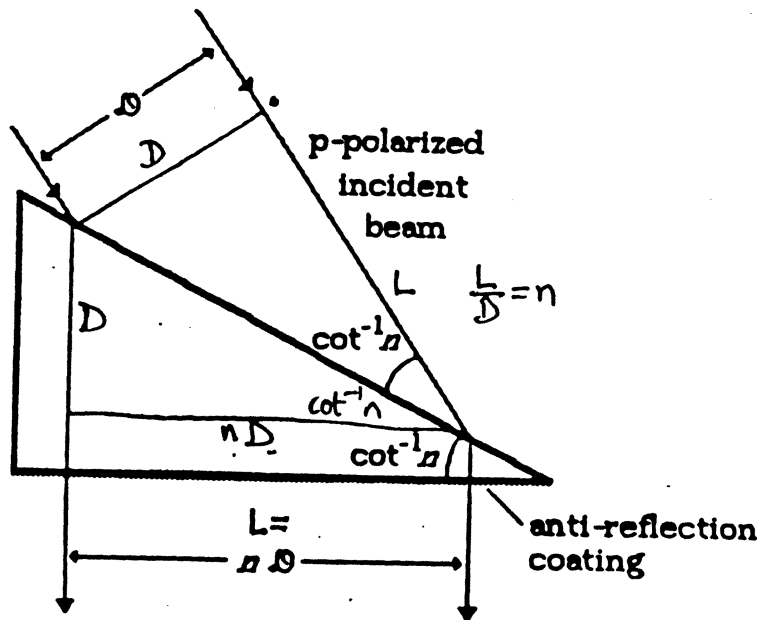
BLM SHAPING

- want circular not elliptical beams
 \because lenses are circular

\therefore use ANAMORPHIC SYSTEMS.

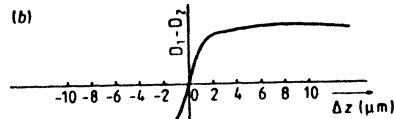
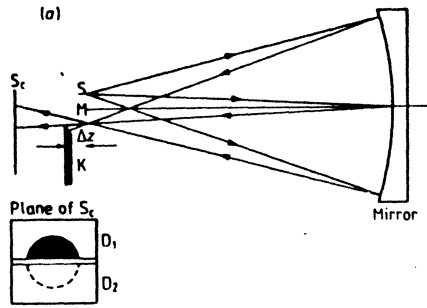


Prism Details



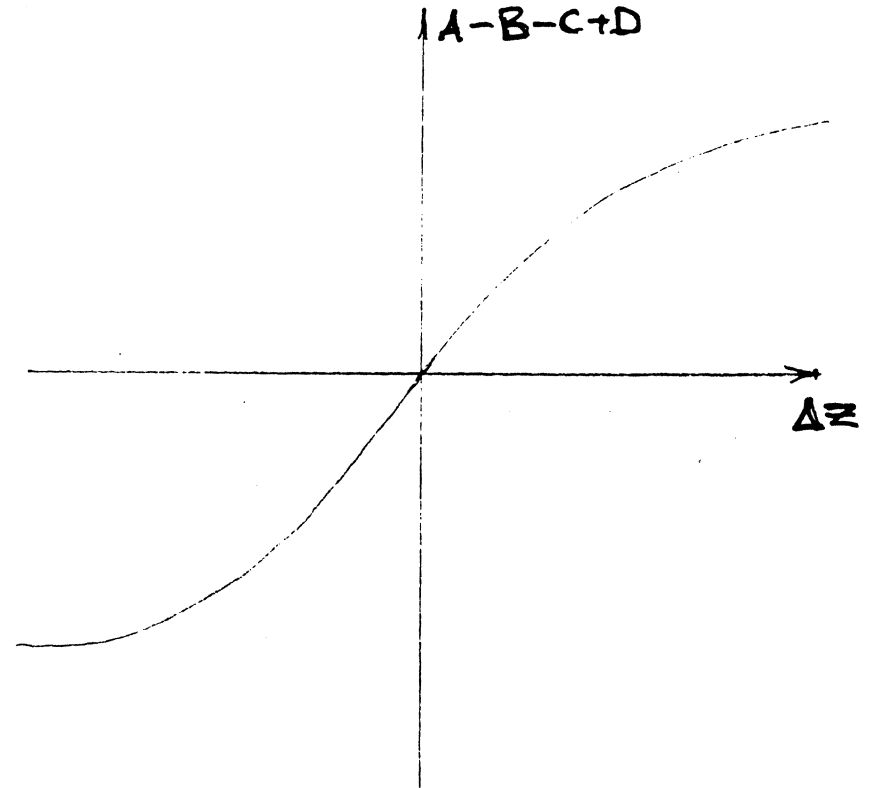
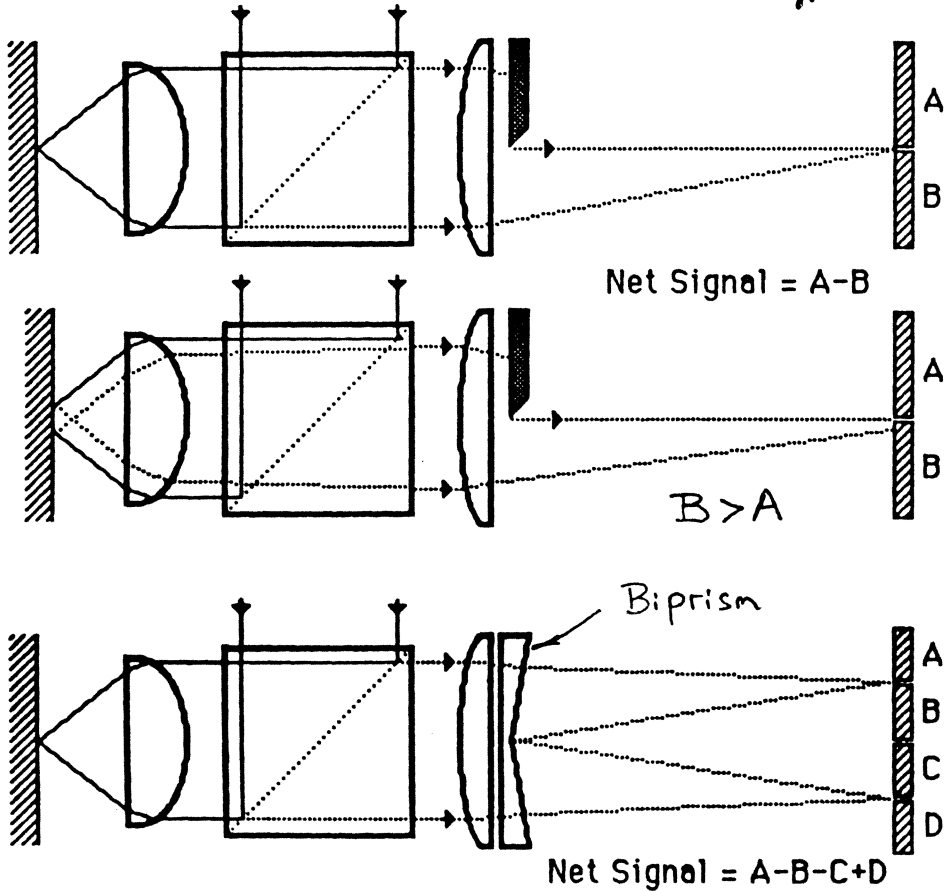
- light incident on prism at Brewster's Angle, all light is transmitted
- result is an expansion of the beam, in the direction shown, by a factor of n
- using two prisms, the beam emerges parallel to its original path.

FOCUS SENSING METHODS



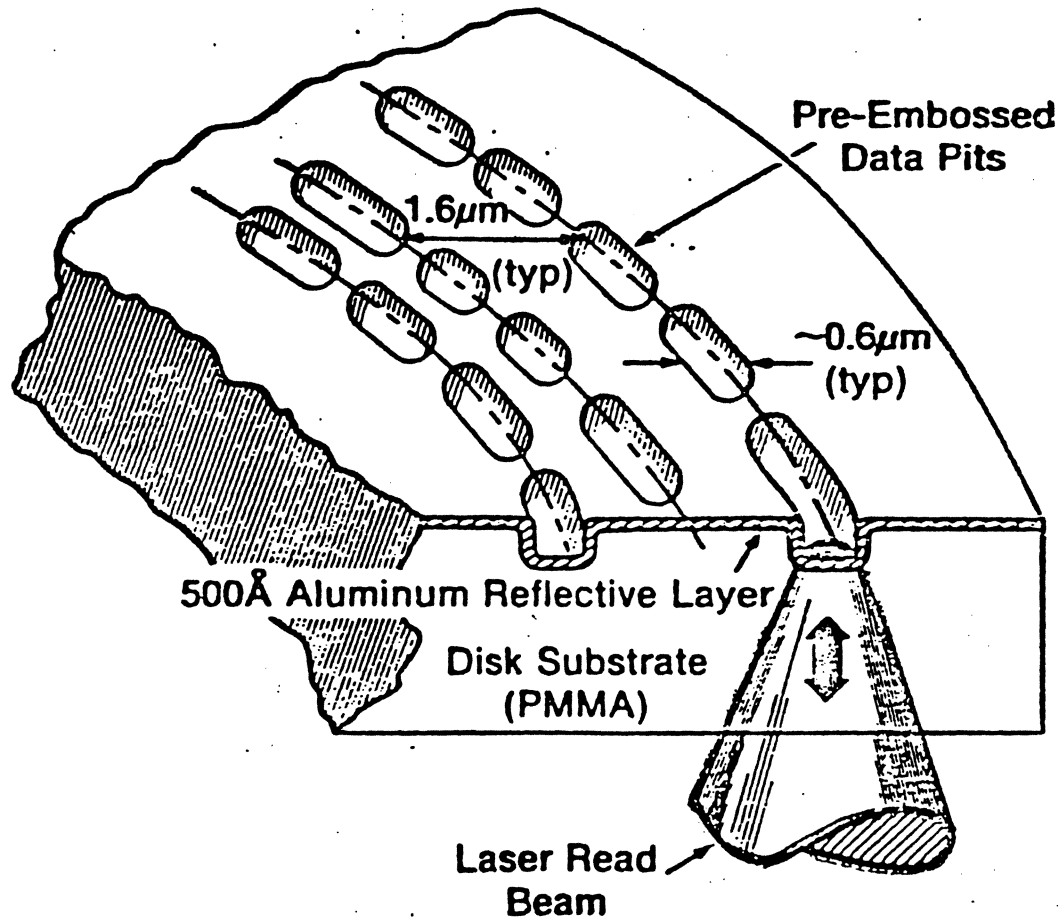
Detectors D_1 and D_2 are equally illuminated only when the knife-edge is at the focal point of the mirror, i.e. $\Delta z = 0$

difference signal is a non-linear function of Δz .



BIPRISM VARIATION

FREAD- ONLY OPTICAL DISK



- spot size larger than pit width
- pit edges are detected.

1981 LASERVIDEO, LV

- 30cm. diameter PMMA disk

- $\frac{1}{2}$ hr. video

1983 COMPACT AUDIO DISK, CD

- 120mm. diam. POLYCARBONATE

- \approx 75 mins., audio

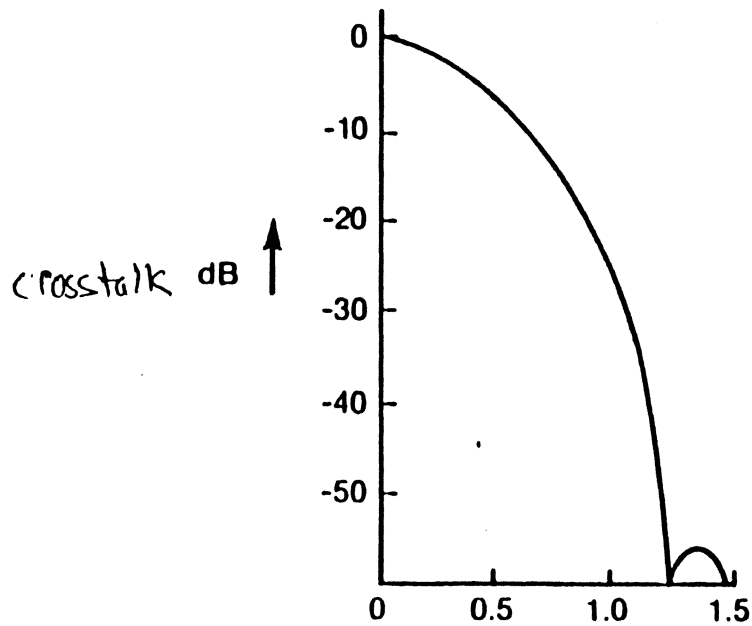
CD-ROM

- 120mm diam. PC

- 550 MB of data

5.25" O-ROM

Track Density - Optical Storage



- To Limit Crosstalk

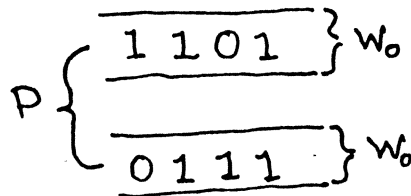
$$\frac{P}{\omega_0} \sim 1.5$$

- $\lambda = 820\text{nm}$

NA	TPI
0.5	18K
0.65	24K

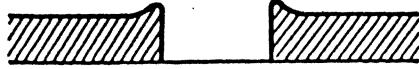
- Residual error limited by:
 - centering error
 - spindle runout
 - gain-bandwidth of servo

$$\frac{P}{\omega_0} = \frac{\text{pitch}}{\text{track width}}$$



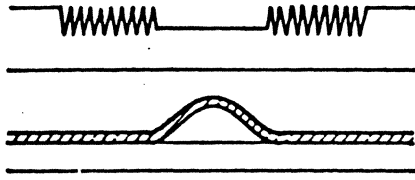
Typical Optical Recording Media and Their Writing Mechanism.

1. Ablation



Te-alloys, organic dyes,
Ag in gelatin

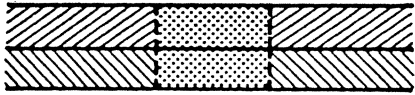
2. Topography Change



Textured Ge, Si

Au (or Pt) alloys/organic polymer

3. Chemical Reaction



Rh/Si, Se/Bi

4. Particle Coalescence



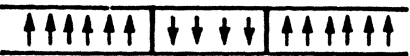
Discontinuous Au

5. Phase Change



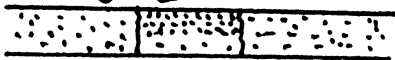
Chalcogenide, TeO_x

6. Magneto-Optic



Mn-Bi, Tb-Fe, Ga-Tb-Fe

7. Segregation



TeO_x , Au/ TeO_2

Basic Read/Write Medium Requirements

1. High laser writing sensitivity

-Compatible with diode laser recording at 1-3MB/s

2. Well defined threshold for recording

-To permit unlimited ($>10^8$) non-destructive reading

3. High resolution and regular mark geometry

- To achieve adequate raw BER at high density

4. High contrast between recorded mark and unrecorded track

-To maximize read signal amplitude(SNR)

5. Long lifetime

-To maintain low raw BER ($<1/10K$) for at least 10 years

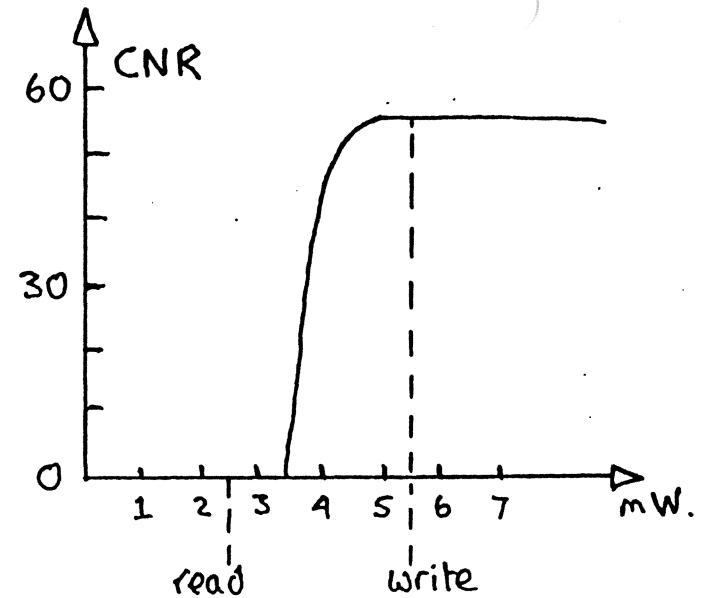
6. No intermediate processing, postable (data can be added later)

-To permit data verification during recording

7. Comparable with formattable substrate

-For track servo, sector ID, etc.
- To certify unwritten media

8. Compatible with low cost, high volume manufacturing process



Laser Power (mW)

- specified by laser manuff-

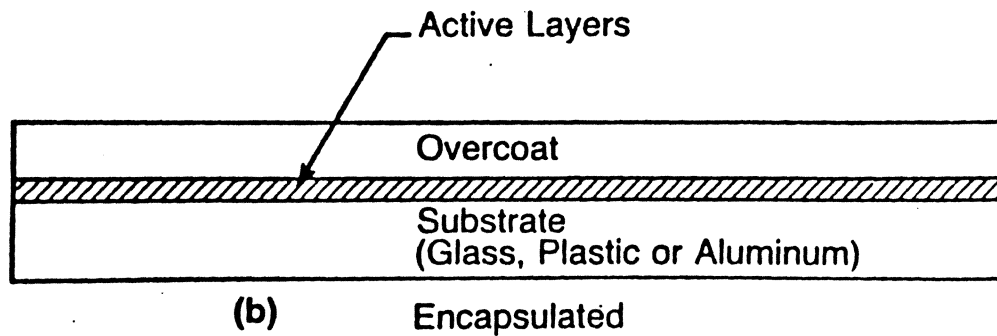
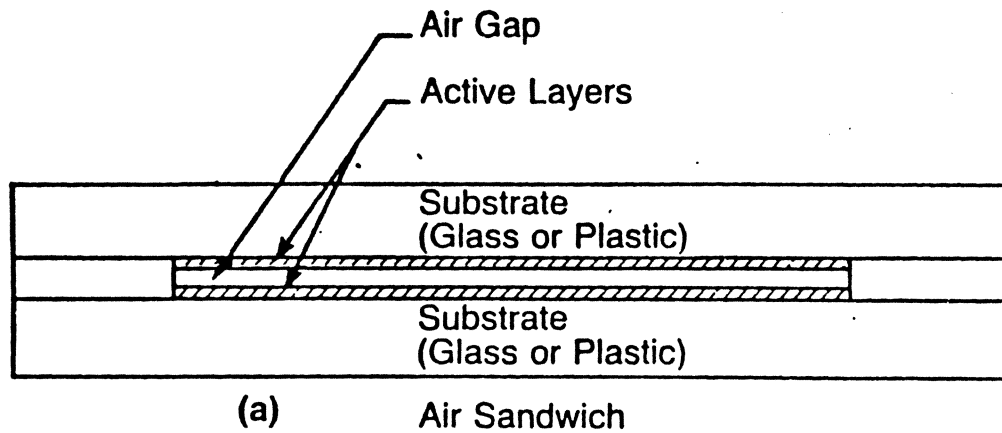
Writing Power Density (mW)

- depends on laser power
and on optics

Writing Energy nJ

- depends on laser power
- optics
- rotatⁿ speed

OPTICAL DISK STRUCTURES



Air Sandwich

used with :

- ablation
- bump forming
- topography change

Encapsulated

used with :

- read-only
- chemical reaction
- discontinuous gold
- thermo-magneto-optic

REVERSIBLE OPTICAL DATA STORAGE

◦ Advantages

- Lower cost of media
- Compatible with magnetic recording
- Multi-function drives (read-only, write-once, reversible)

◦ Disadvantage

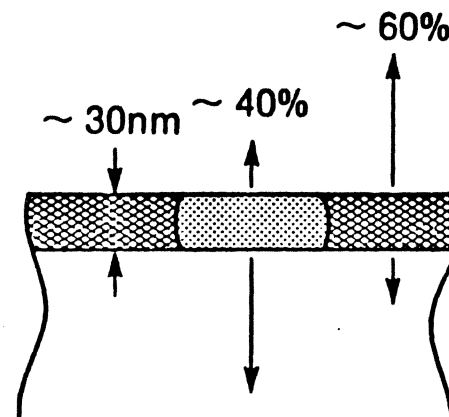
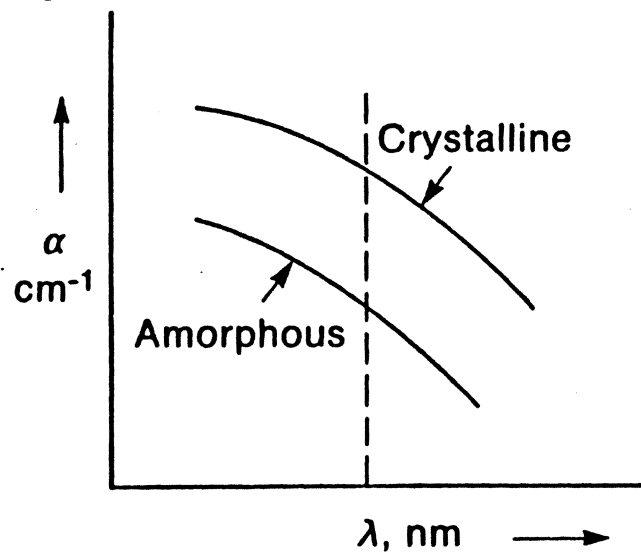
- Erasure is not easily detectable

◦ Issues

- Limited erasability ($\sim 10^3$ cycles) may be acceptable for storage of images, audio, but not for data
- Erasure should be fast so that writing data rates match reading data rates

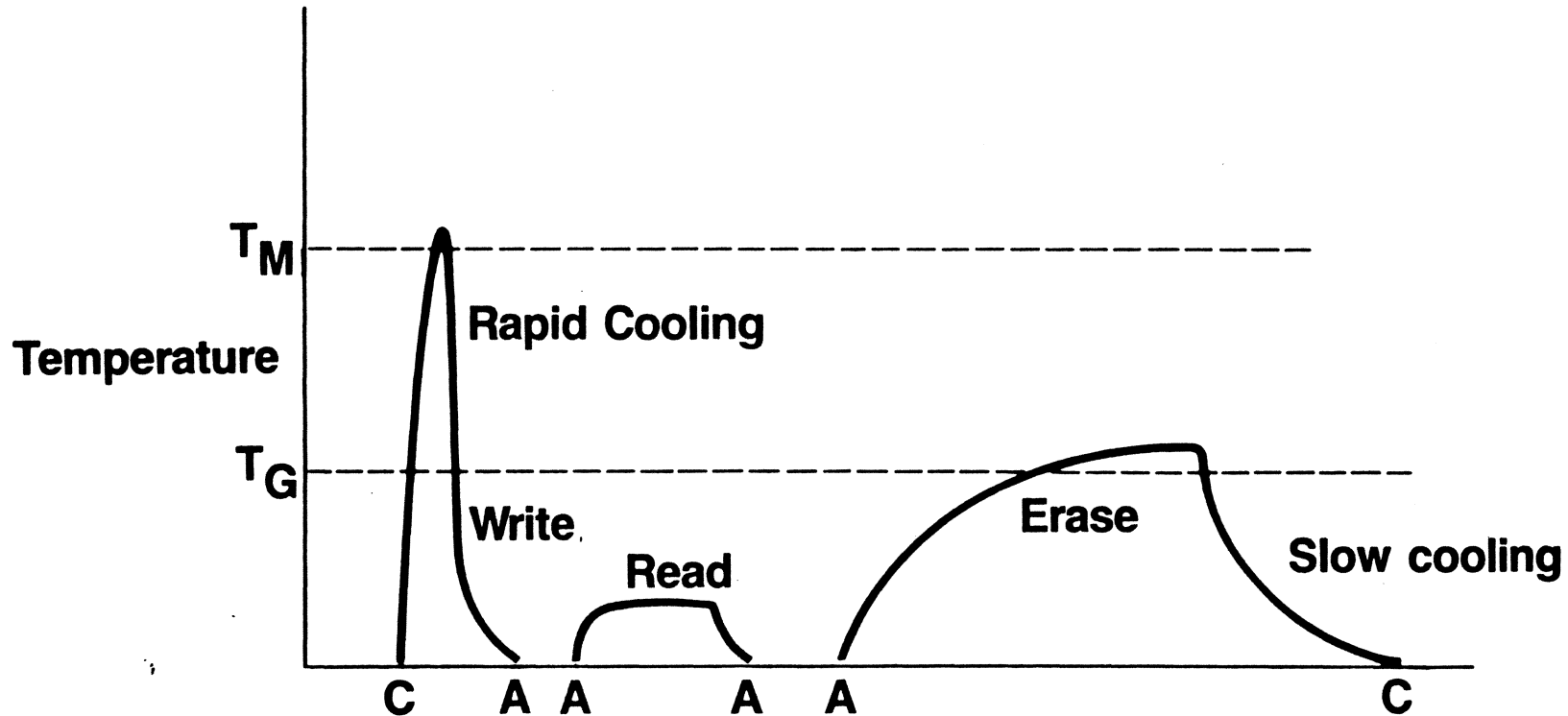
Amorphous to Crystalline Phase Change in Te-Based Materials

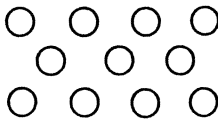
eg TeAsSe, TeGe, Te



Reverse-Mode Writing:
Amorphised marks along
a polycrystalline track

Phase-Change Recording



C = crystalline phase  **high reflectivity, low transmissivity**

A = amorphous phase  **low reflectivity, high transmissivity**

Ideal Optical Storage Material

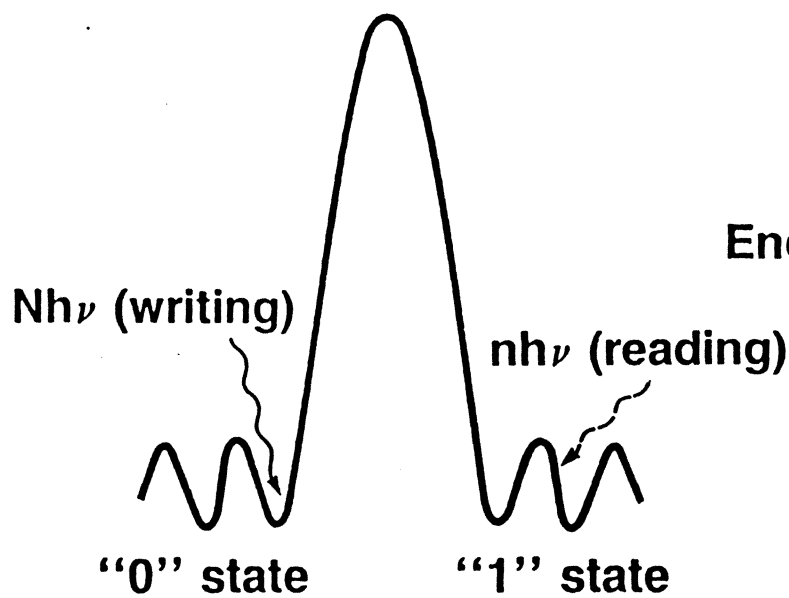
- Need a material: –writable with low-power laser
 - high reading SENSITIVITY
 - but STABLE with respect to TIME
 - TEMPERATURE
 - HUMIDITY
 - STRAY FIELDS
- Medium must also be STABLE and REVERSIBLE after millions of write-read-write cycles

Answer: Thermo-Magneto-Optics, TMO

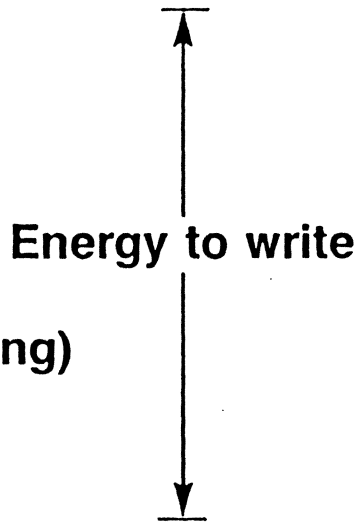
Writing and Reading Energies

Write-Once Media

Reversible Media



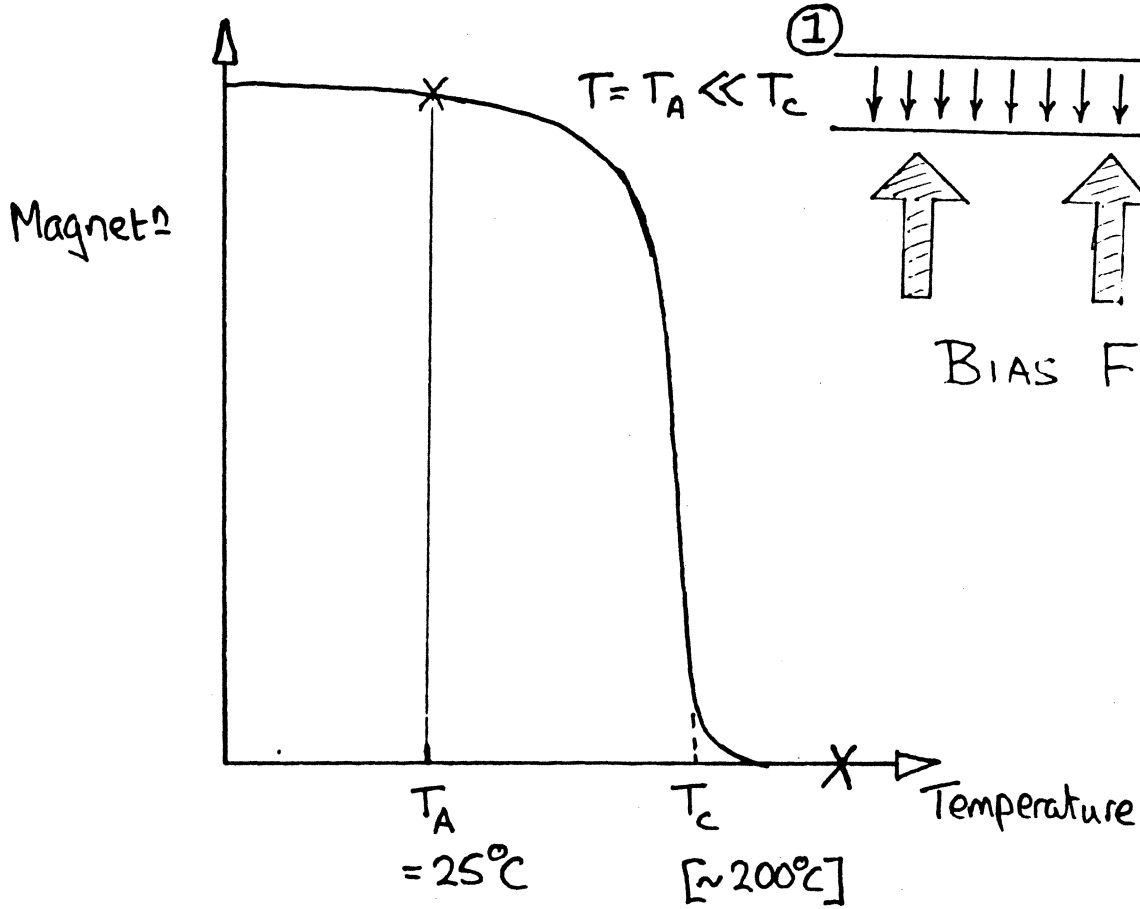
e.g. Hole-burning



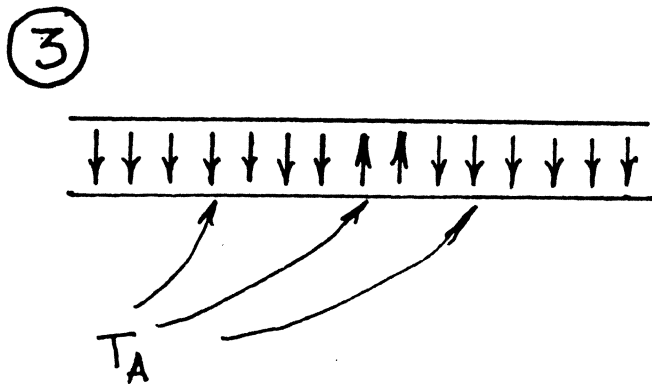
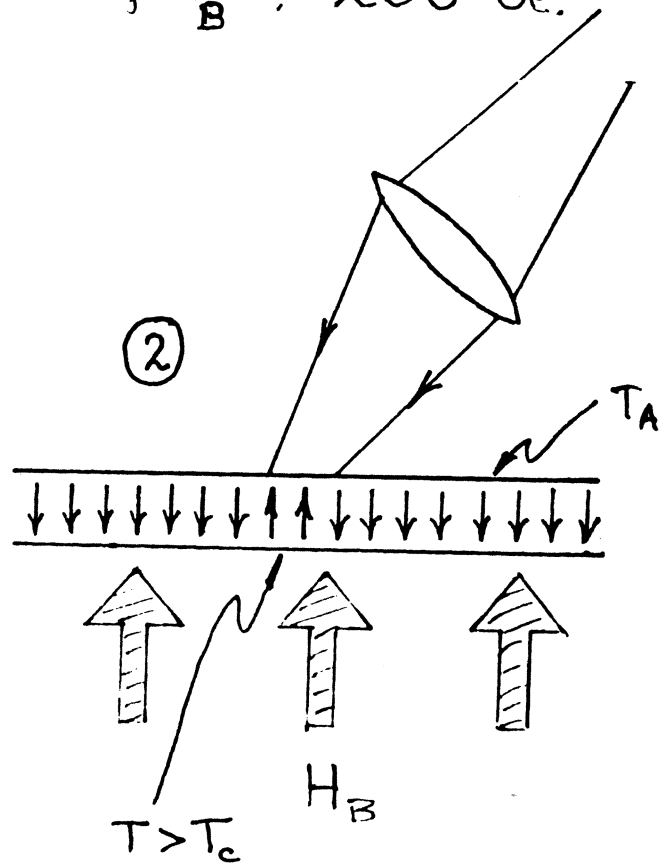
e.g. Thermo-Magneto-Optic

THERMO - MAGNETIC RECORDING

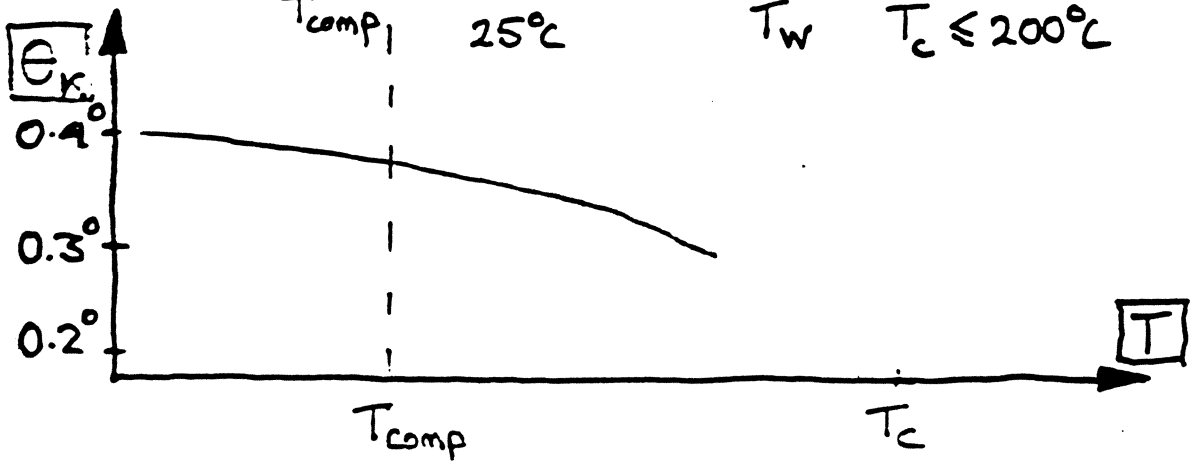
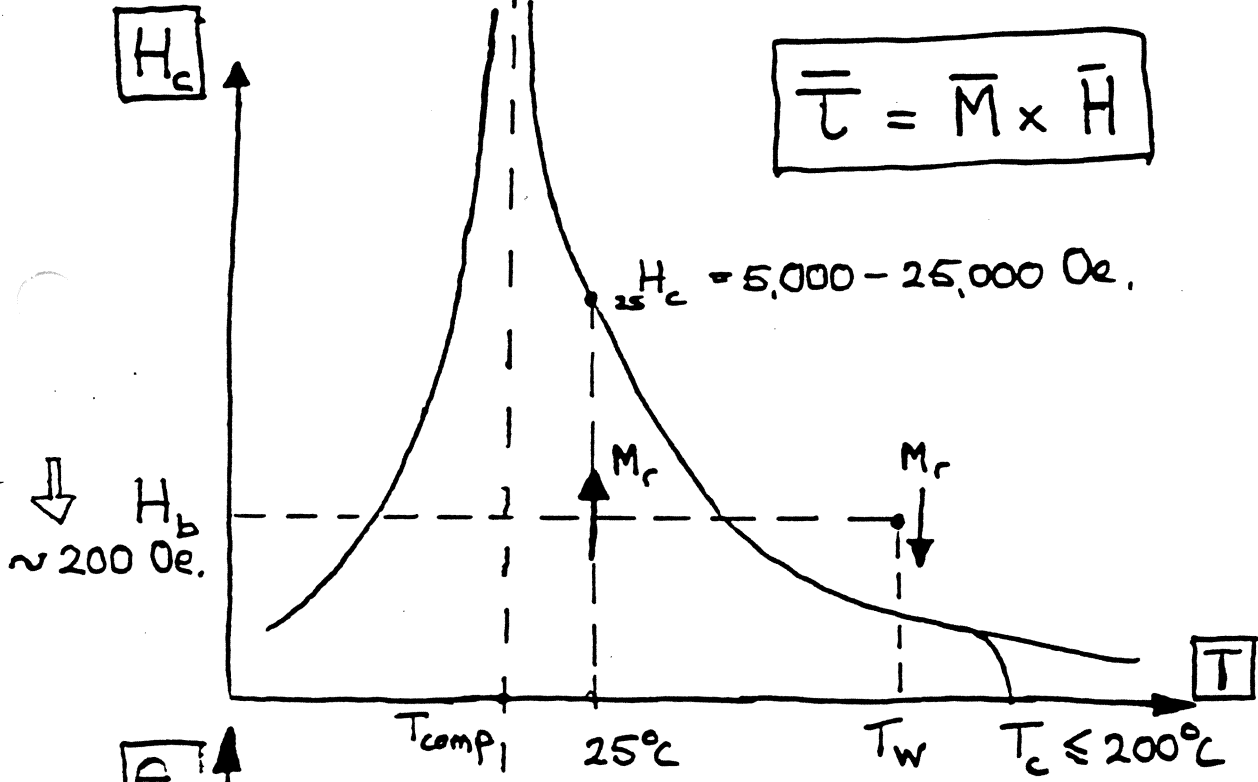
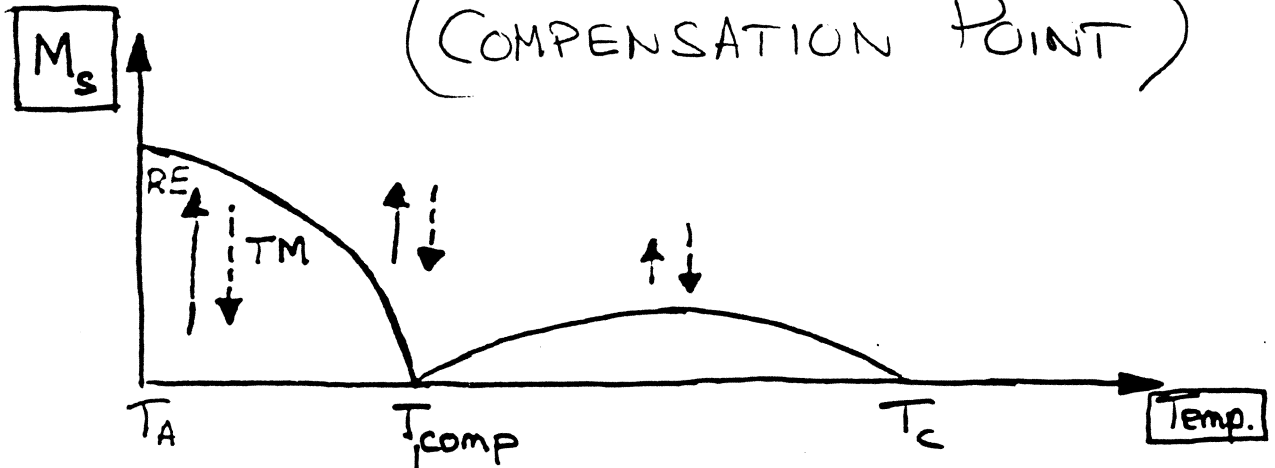
CURIE POINT



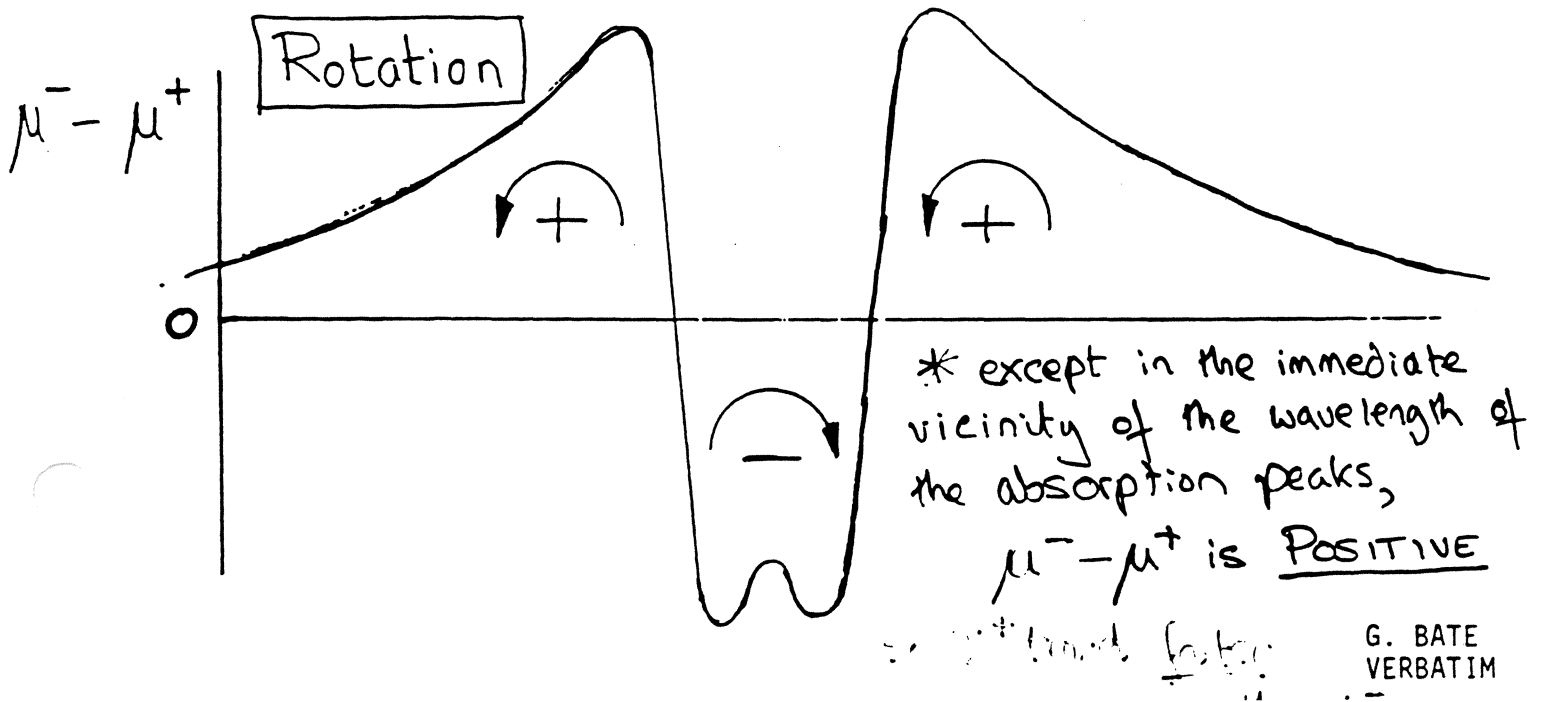
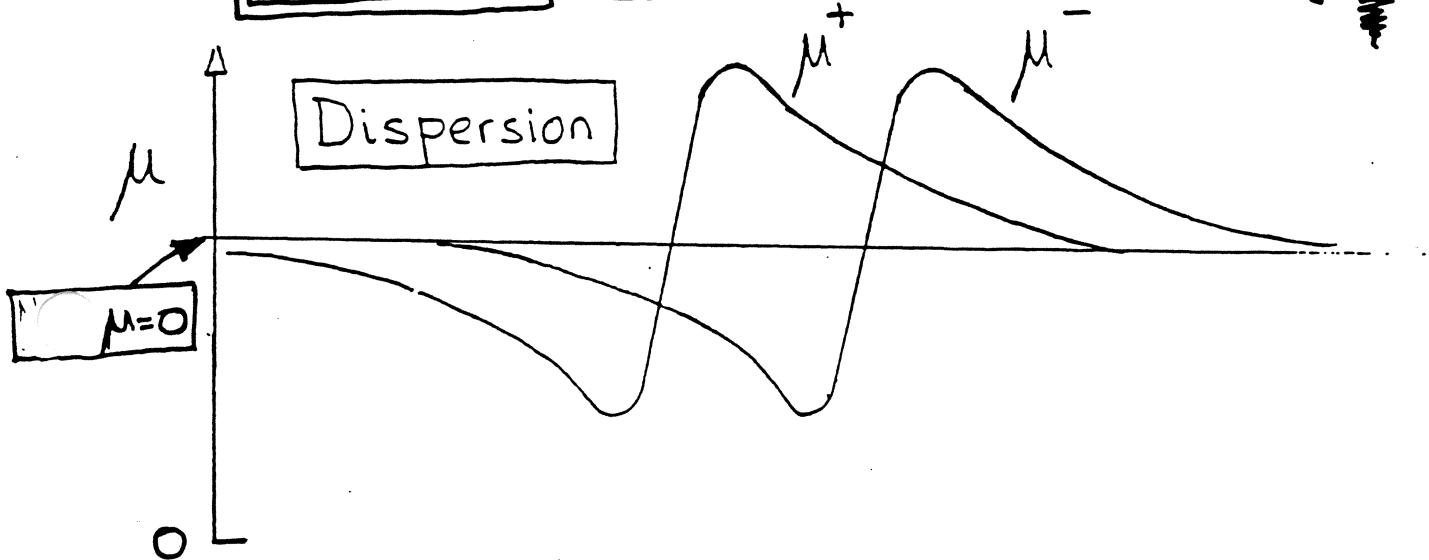
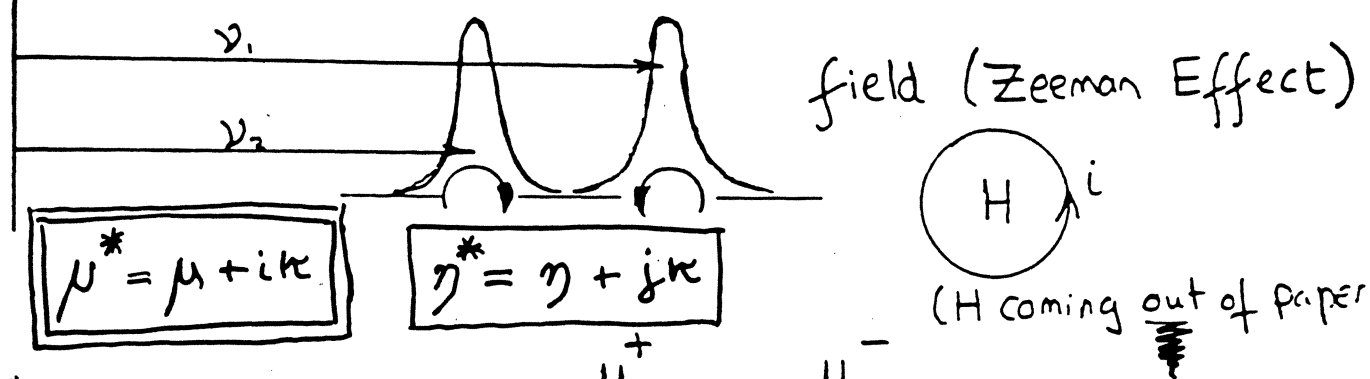
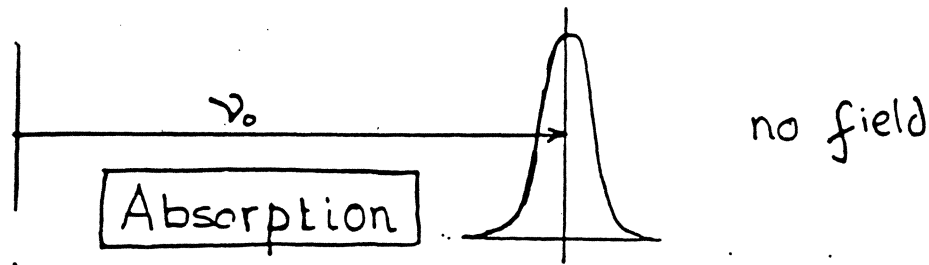
BIAS FIELD, $H_B \doteq 250 \text{ Oe.}$



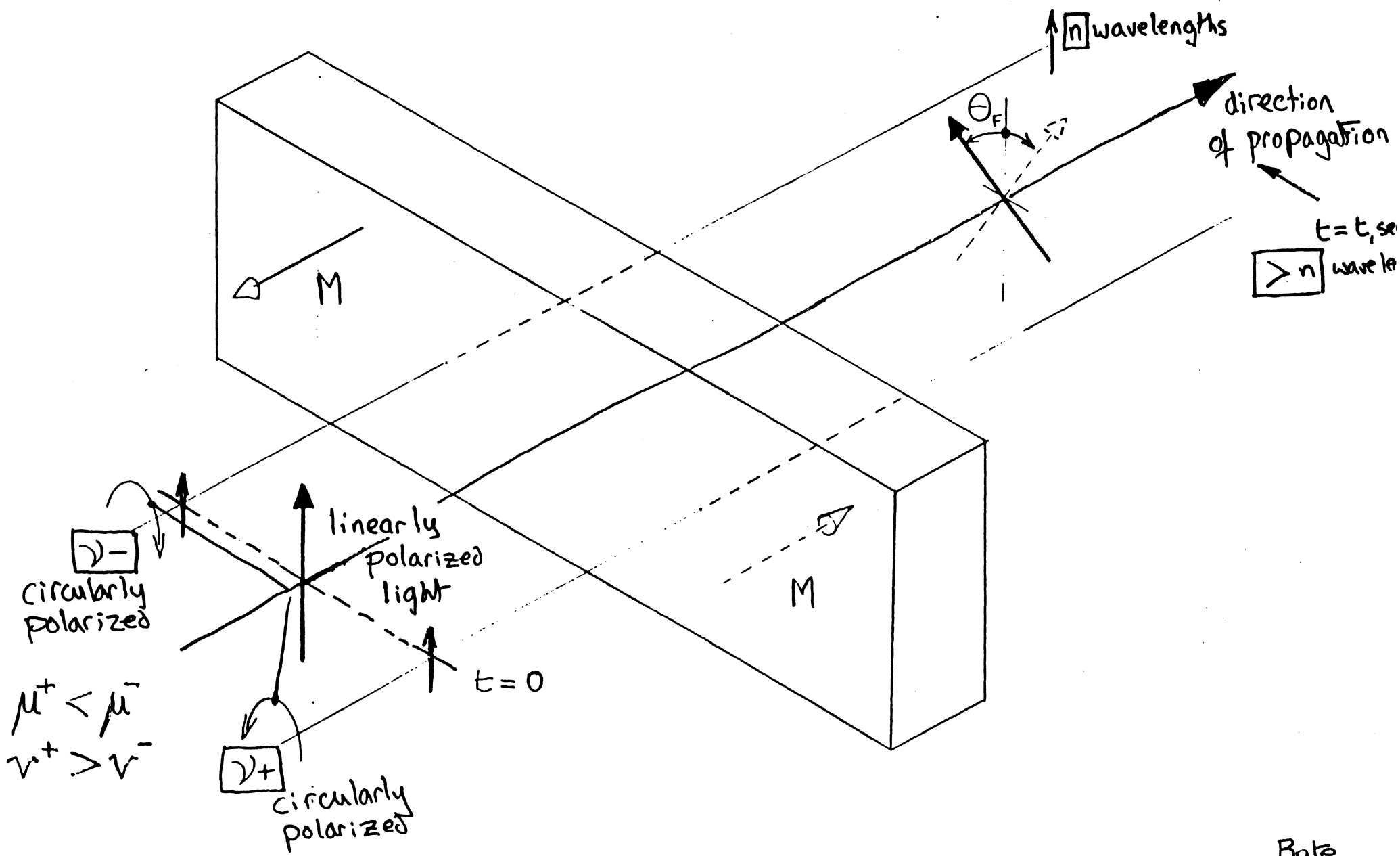
TMO RECORDING (COMPENSATION POINT)



FARADAY EFFECT (1845)



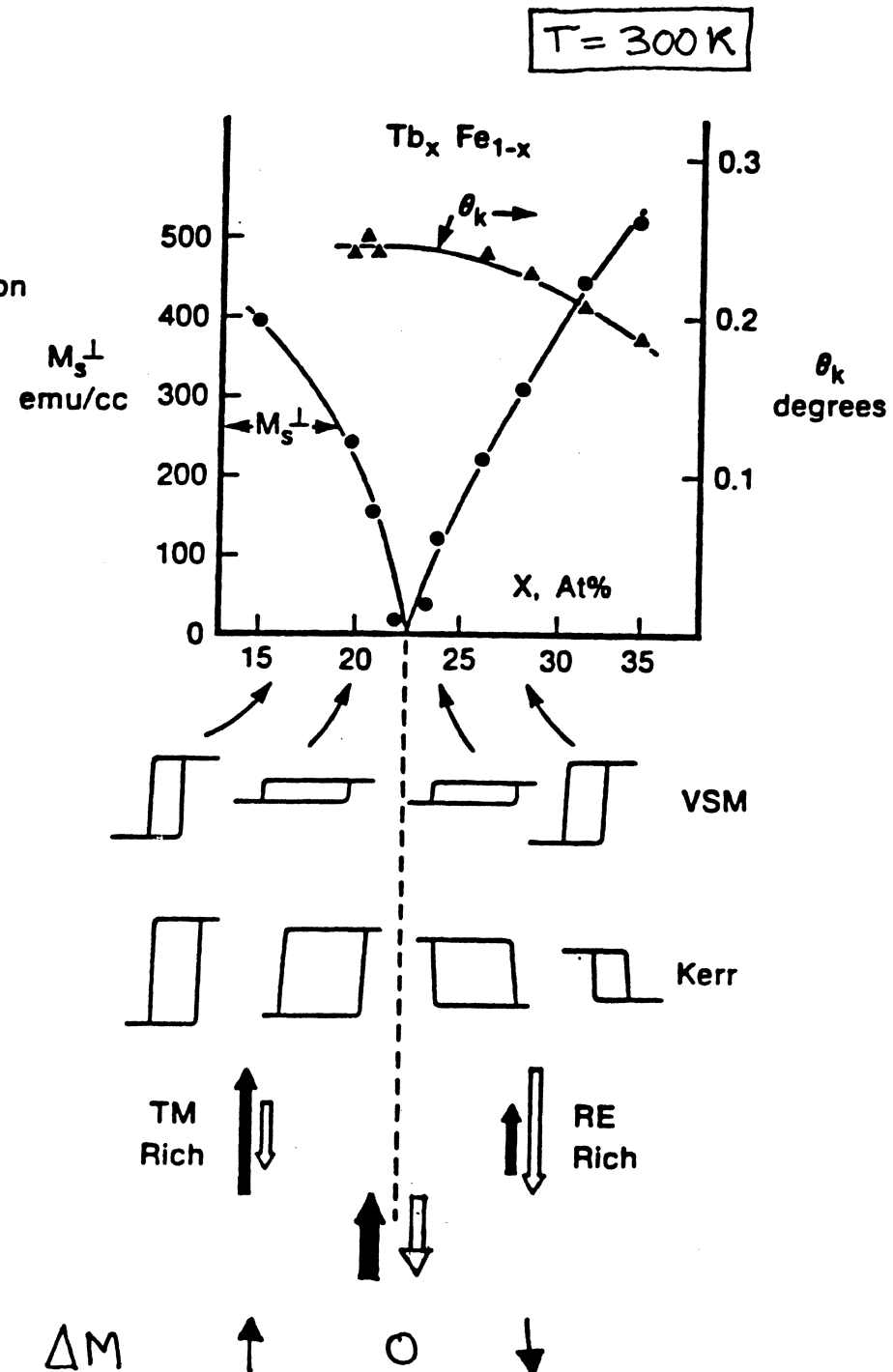
FARADAY EFFECT



Properties of Rare-Earth, Transition-Metal Alloys

- Kerr Rotation Determined by TM Magnetization
- θ_k Ranges 0.25° - 0.5° Depending on Composition

Material	Curie Temp (K)	Compensation Temp (K)	Coercive Energy Density (10^4 J/m ³)	Anisotropy Constant (10^4 J/m ³)	Kerr Rotation (deg)
	T_c	T_{comp}	$M_s H_c$	K_u	θ_k
Tb ₂₃ Fe ₇₇	400	300	4	30	0.23
Tb ₂₁ Co ₇₉	--	300	4	14	0.33
Gd ₂₆ Fe ₇₄	480	300	0.02	2.5	0.29
Gd ₂₁ Co ₇₉	--	300	0.02	1	0.33
Tb ₂₂ Fe ₆₆ Co ₁₂	~500	<300	8	10	0.38
Gd ₂₂ Tb ₄ Fe ₇₄	450	300	1.5	4	0.30
Gd ₁₆ Tb ₆ Co ₇₈	--	280	1.1	4	0.32

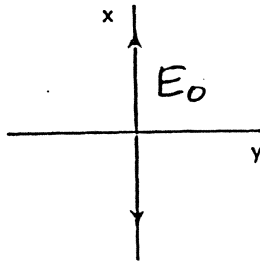


T. M.O. Materials : Kerr Rotation, Θ_k

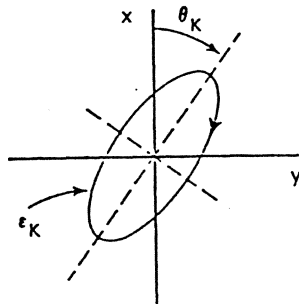
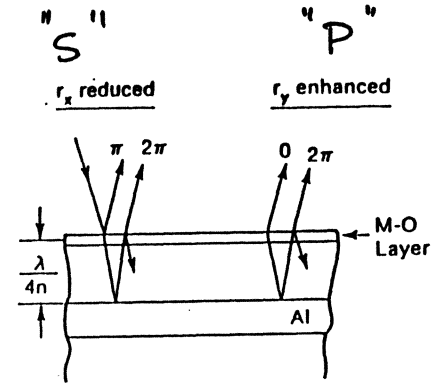
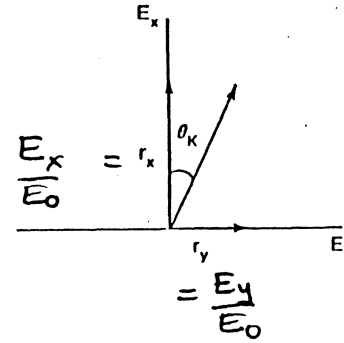
<u>Year</u>	<u>Material</u>	Θ_k	<u>Wavelength</u>	<u>Temperature</u>
1958	MnBi	0.7°	visible	25° C
1985	GdTeFeCo	0.45°	780 nm	25
1987	PtMnSb	1.9°	780	25

1986	CeSb _{0.75} Te _{0.25}	3°	1.2 μ m	2 K
1986	CeSb	14°	2.4	2

Enhancement of Kerr Rotation



Incident:
Plane polarized

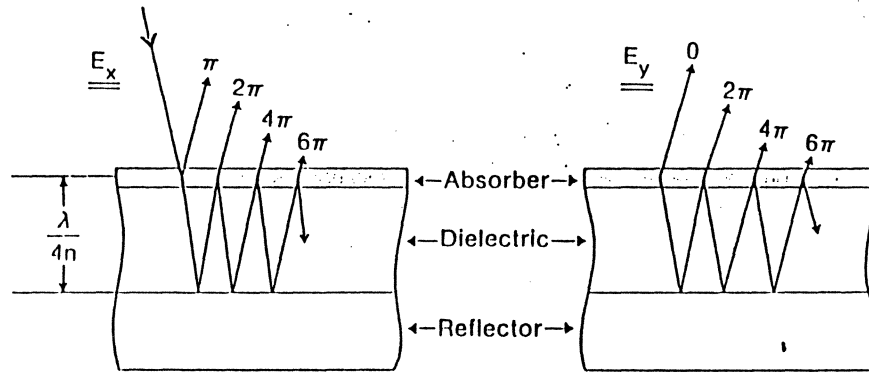


Reflected:
Elliptically polarized

	θ_K°
Single layer	0.27
Bilayer	0.5
Trilayer	0.75
Quadri-layer	1.7

GdTbFe, $\lambda = 800 \text{ nm}$
Ohta et al, Sharp (1983)

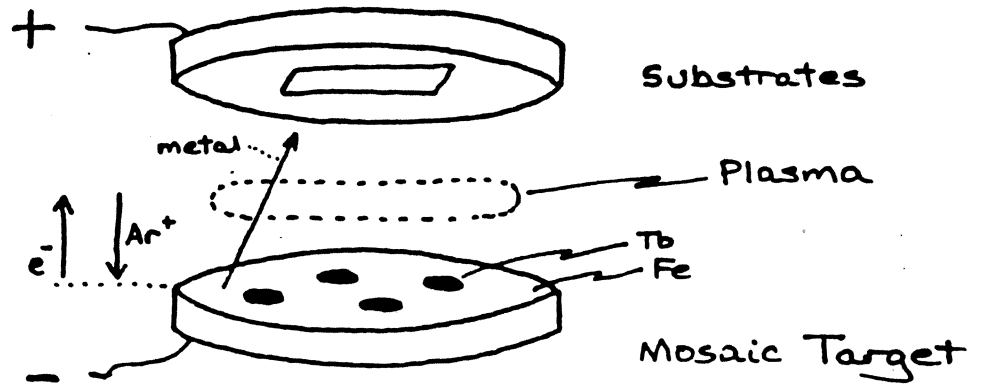
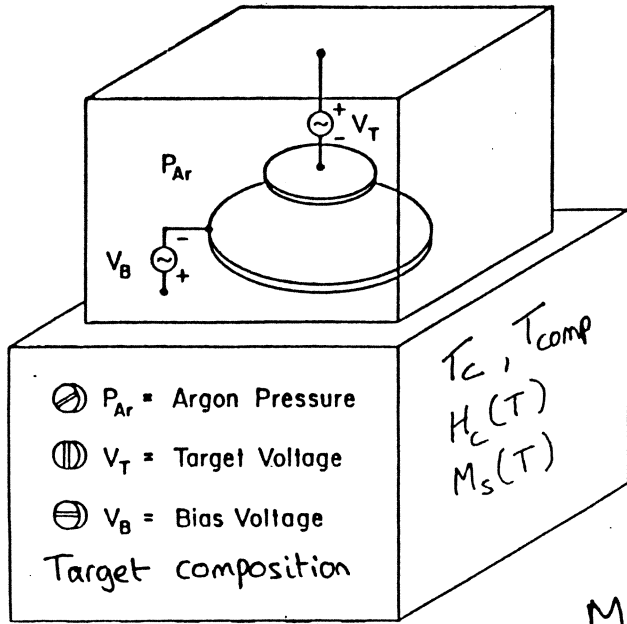
Magneto-optic Anti-reflection Trilayer



Destructive Interference
 $\bar{r}_x = 0$
High Absorption of Record-Beam Energy

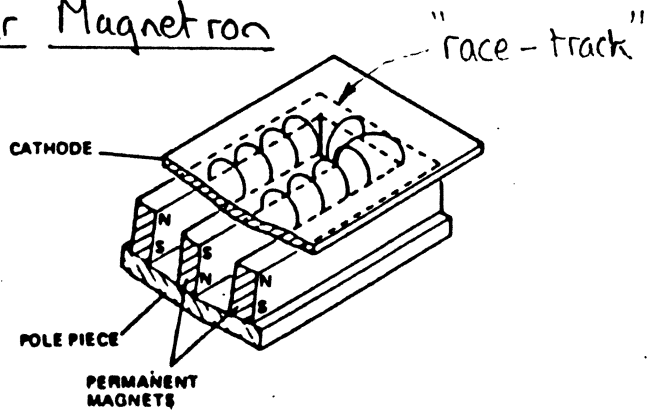
Constructive Interference
 \bar{r}_y Enhanced
Increased Playback SNR

SPUTTERING

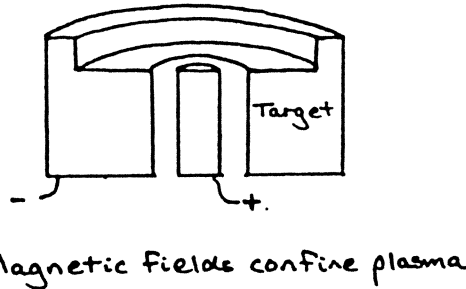


Magnetron Sputtering used to increase deposition rates

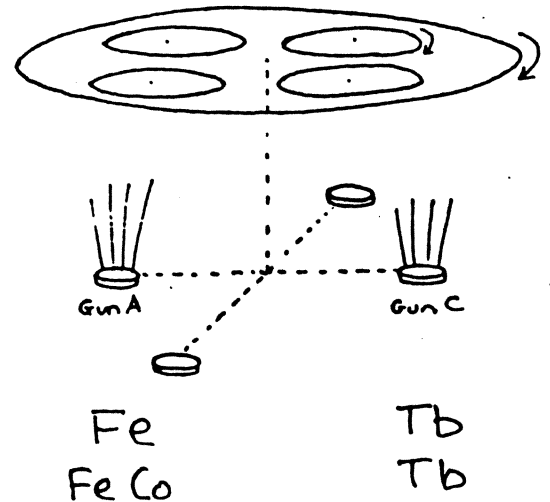
Planar Magnetron



Co-sputtering S-gun



Alloy targets:

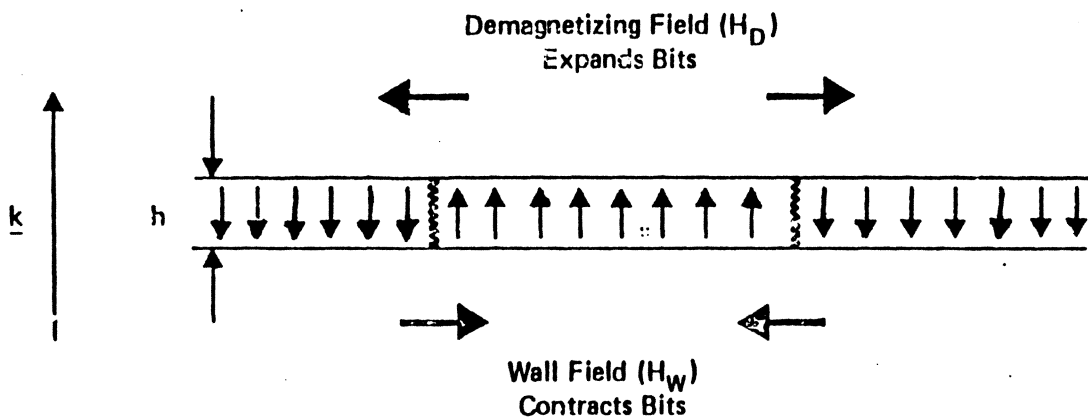


Overcoats may be sputtered directly from compound targets; Al_2O_3 , Si, SiO_2 , ZnS

Properties of Interest in T.M.O. Films

<u>Writing</u>	<u>Storage</u>	<u>Reading</u>
T_c, T_{comp}, M_s $\alpha(\lambda), R(\lambda)$ $\left. \frac{dH_c}{dT} \right) T_{comp}$ thickness specific heat therm. conductivity (film and substrate) spot diameter	$H_c, M_s (20^\circ C)$ $\Delta H_c, \Delta M_s (x, T)$ $K_u > 2\pi M_s^2$ magnetostatic vs. wall energy corrosion	$\mu^* = n - ik (\text{Temp.})$ $[\alpha(\lambda), R(\lambda)]$ thickness $\theta_F, \theta_K, \frac{\theta_F}{\alpha}, \theta_K R$ T_{max} grain structure interference layers

Domain Stability



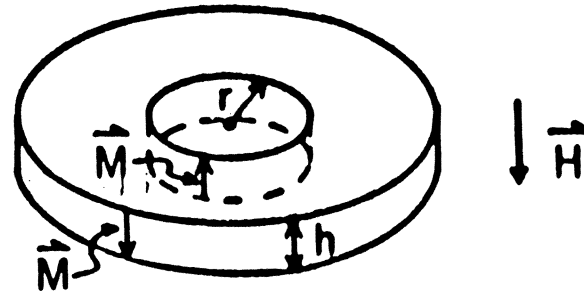
$$\underline{H_D} \doteq \frac{4\pi M_s h}{h + \frac{3}{4}d} \cdot \underline{k}$$

Bit Diameter = d

$$\underline{H_W} = -\frac{\sigma_w}{M_s d} \cdot \underline{k}$$

- Stability Criterion: $H_{EFF} = |H_D - H_W| < H_c$
- Minimum Stable Domain: $d_{min} = \frac{\sigma}{MH_c}$

DOMAIN STABILITY



$$\begin{array}{cccc}
 \text{(Normalized)} & \text{(Force)} & \text{(Wall Surface)} & \text{(Demagnetizing)} & \text{(Applied)} \\
 & \text{Area)} & \text{Tension)} & \text{Field)} & \text{Field)} \\
 \frac{\Delta E}{\Delta r} \cdot \left(\frac{1}{2\pi r h} \right) & = & -\frac{\sigma}{r} & + 4\pi M^2 \cdot \frac{h}{r} F\left(\frac{2r}{h}\right) & - 2M_s H
 \end{array}$$

FOR COERCIVE STABILITY:

$$\frac{\Delta E}{\Delta r} \cdot \left(\frac{1}{2\pi r h} \right) > -2MH_c$$

OR:

$$r_{\min} = \frac{\sigma}{2M(H_c - H)} \quad (\text{Small } M, \text{ thin films})$$

TMO Drives

Company	Disk Diameter	Capacity (2 - sides)	Access Time	RPM	Data Rate
MAXTOR	5.25 "	1 GB 650 MB	35 ms	2,200	10 Mb/sec.
SONY	5.25 "	650	90	2,400	7.4
CANNON	5.25 "	512	80	3,000	9.1
OLYMPUS/RICOH	5.25 "	650	44	1,800	5.6
SHARP	5.25 "	650	67	2,400	7.4
H - P	5.25 "	650	95	2,400	5.1 5.6
VERBATIM	3.25 "	64 MB (1 - side)	30	2,400	2.1

OPTICAL DRIVE TECHNOLOGY TRENDS:

- G.E. has demonstrated an 8-track laser device
- frequency - variable lasers are being developed for use with multiple-layer films
- seek-time now < 20 ms.
- rotational speed $\rightarrow 3,600$ rpm
- higher-level of electronic integration

1995

data rate > 3 MB/sec.

access time < 15 msec on 3.5 " drives
 < 20 msec on 5.25 " drives

capacity > 400 MB on 3.5 " disk
 > 2 GB on 5.25 " disk

cost $< \$1$ /MB

Data Storage in 2000—Trends in Data Storage Technologies

MARK H. KRYDER

MAGNETIC DISK DRIVE IN 2000

Capacity: 10 Gbytes

Data Rate: 100 Mbytes/s

* Eight 3.5-in disks
125 000 bits/in
7000 tracks/in
0.7-Gbyte/side
Zone Bit Recording

6 Mbytes/s—channel
at 5000 r/min
• 1 head/surface •

Non-removable disk

MAGNETOOPTIC DISK DRIVE IN 2000

Capacity: 10 Gbytes

Data Rate: 100 Mbytes/s

* One 3.5-in disk
115 500 bits/in
0.33- μ m spot
1.5 bits/transition
63 500 tracks/in
0.33- μ m track
0.07- μ m guard band
5-Gbyte/side
Zone Bit Recording

6 Mbytes/s—channel
at 5000 r/min
* 8 lasers/head *
1 head/surface

removable disk

ISSUES:

- Reversible versus write-once or read-only
- Data rate, overwrite by magnetic field modulation - need fast erasibility
- Bit density limited by λ and N.A.
- Magnetization determined by R.E. and T.M.
- Magneto-optic effects determined by T.M.
- Optical heads - integrated?, short λ lasers, holographic lenses
- CAV vs. CLV vs. banded CAV
- Removability, vs. multiple disks/drive
- Access time (head mass)
- Continuous vs. sample servo
- Multi-functional drives
- Hierarchy of storage must be transparent to users
- Media cost
- Media life expectancy, reversibility
- Media backward compatibility

TMO versus Conventional Magnetic Recording

TMO:

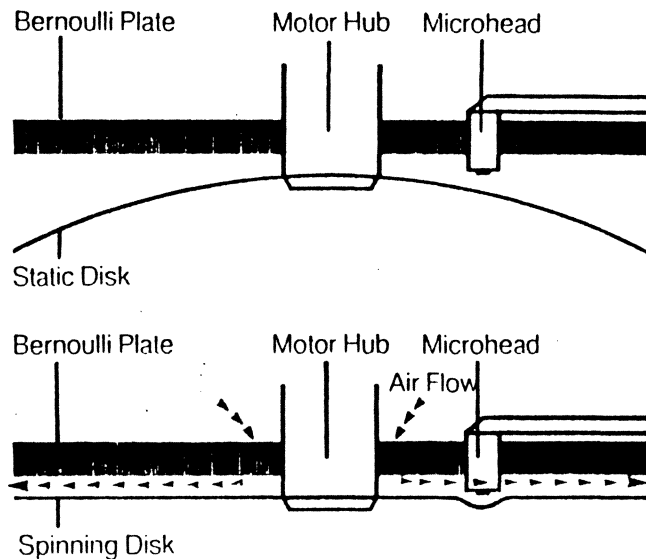
- much greater head–disk separation
- no risk of head crashes
- removable disks
- tpi x 10 higher
- lens can handle multiple signals
- bits are passive until interrogated

MAGNETIC RECORDING:

- mature, established tech.
- problems are known
- standards established
- small, light heads
(→ faster access times)
- multiple disks/spindle
- high data rates
- direct overwrite
- stable media

G. Bate

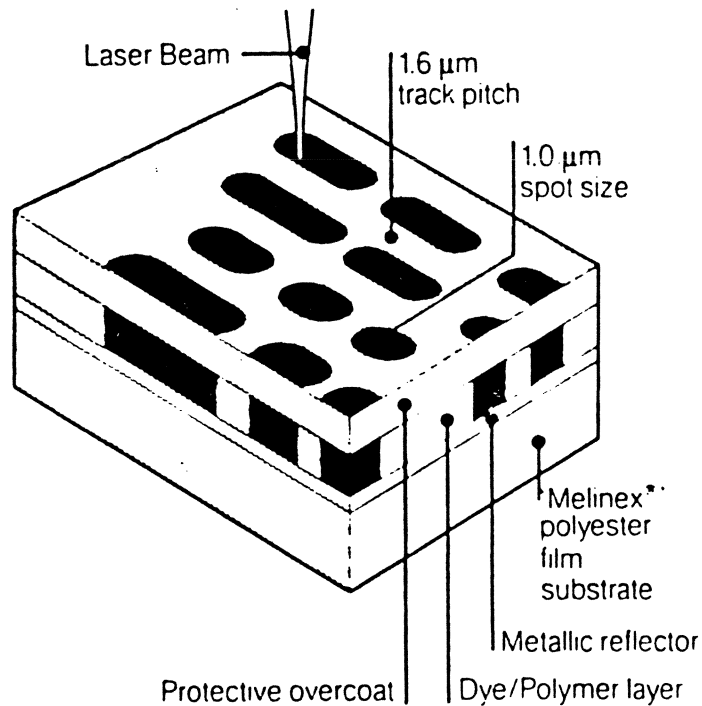
FLEXIBLE OPTICAL DISK SYSTEMS



- Bernoulli principle stabilizes disk - head separation.
 - No dynamic focussing required
 - Small, light weight head, high N.A.
- Two disks can be spun in one cartridge

- 130 mm disk
- 2-disks/cartridge, 1 GB
- 40 ms. average access time
- 1.5 MByte/sec data rate
- $\frac{1}{2}$ height

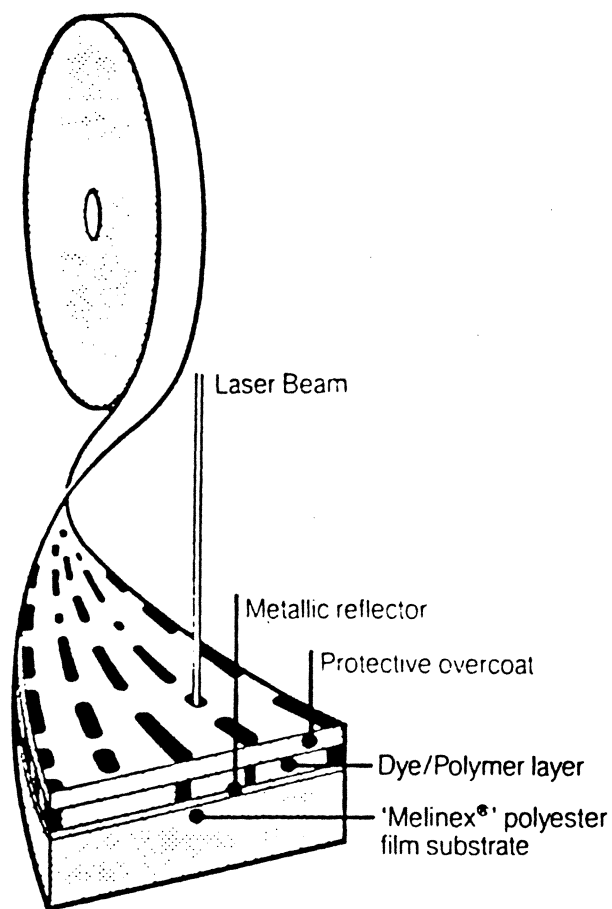
ICI DIGITAL PAPER



- PET substrate, 25-75 μm thick
- I.R. absorbing, dye polymer, $\begin{cases} 780 \\ 830 \text{nm} \end{cases}$
- preformatting of data & tracks by embossing dye-polymer layer
- writing by pit-forming in dye-polymer layer. 1 μm pits
- life expectancy: 15 years

- recording rate of 10 MHz using a 10mW laser
- 10 1/2" reel of half-inch tape holds 600 GB (cf. magnetic tape 180 MB)

ICI 1012 OPTICAL TAPE



PRODUCT DESCRIPTION

TAPE PARAMETERS

Length	880m
Width	35mm
Thickness	0.075mm
Elastic Modulus	4500N/mm
Substrate	'Melinex®' polyester film

REEL PARAMETERS

Reel I.D.	76mm
Reel O.D.	317mm
Flange Spacing	37mm
Flange Thickness	2mm
Reel Composition	Chemcor

STORAGE AND OPERATING REQUIREMENTS

Long Term Storage

Optimum Temperature	18 deg C
Maximum Temperature	65 deg C
Minimum Temperature	5 deg C
Maximum RH	93%

Operating

Optimum Temperature	18 deg C
Maximum Temperature	32 deg C
Minimum Temperature	16 deg C
Maximum RH	80%

READ/WRITE CHARACTERISTICS

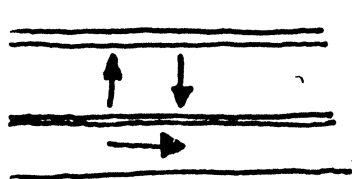
Lasers	Power: 10 milliwatt	Wavelength: 830nm
CNR	55dB using 830nm laser at 7 Meters per second	
Raw BER	1×10^{-5} to 1×10^{-6}	

COMPATABILITY

Creo 1003 Optical Tape Drive and equivalents

METHODS OF STABILIZING TMO FILMS

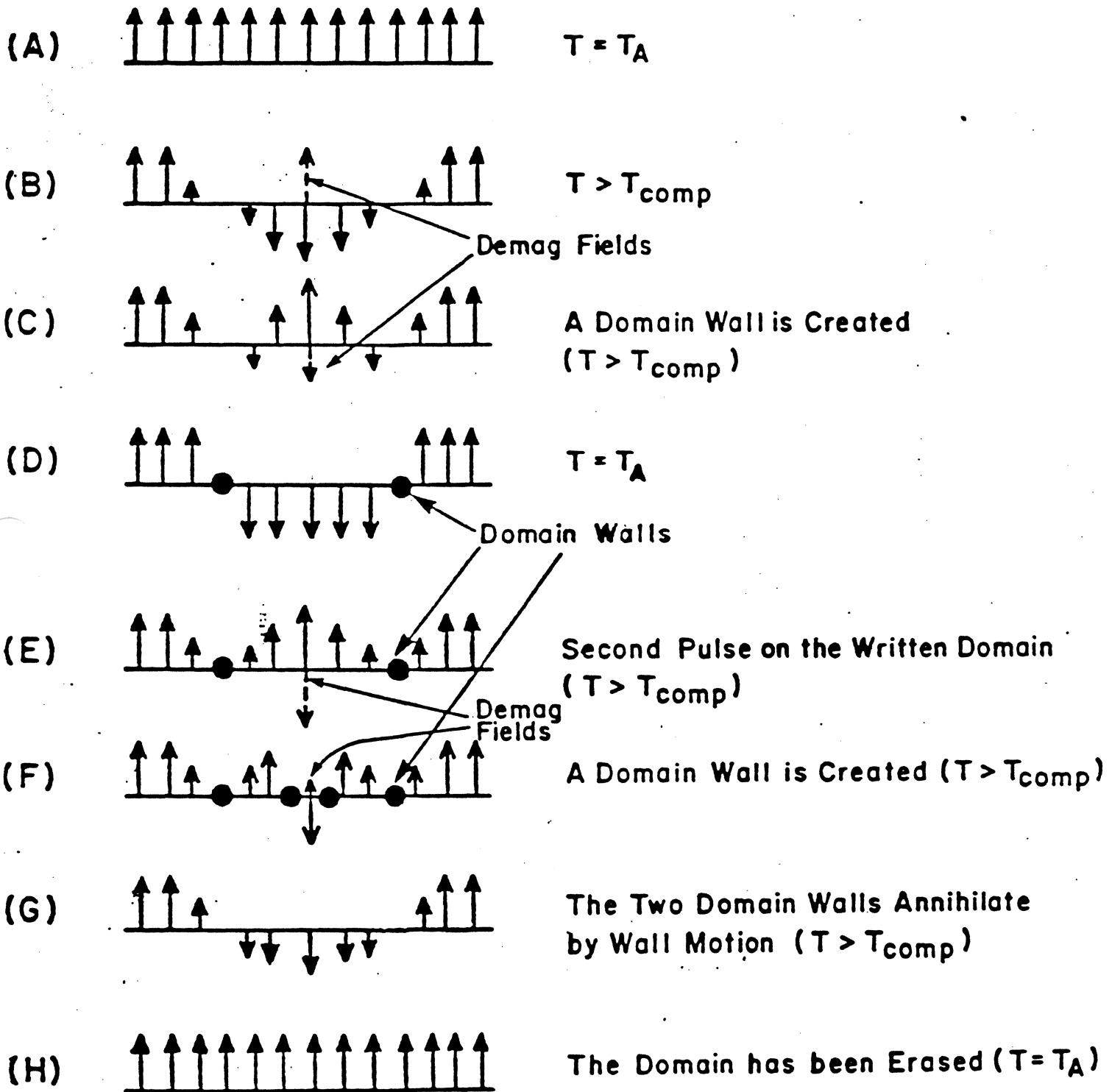
- addition of 1-8 at.% of $\boxed{\text{Pt}}$, Cr, Al
- undercoating and overcoating eg. SiO_2 , $\boxed{\text{AlN}}$
 \uparrow peels
- addition of 80 at.% Tb to SiO_2
 $(\eta = 1.9)$
 $\theta_K \quad 0.4^\circ \rightarrow 0.6^\circ$
 $\left[\text{or } \text{Al}_2\text{O}_3 - \text{Ta}_2\text{O}_5 \right]$



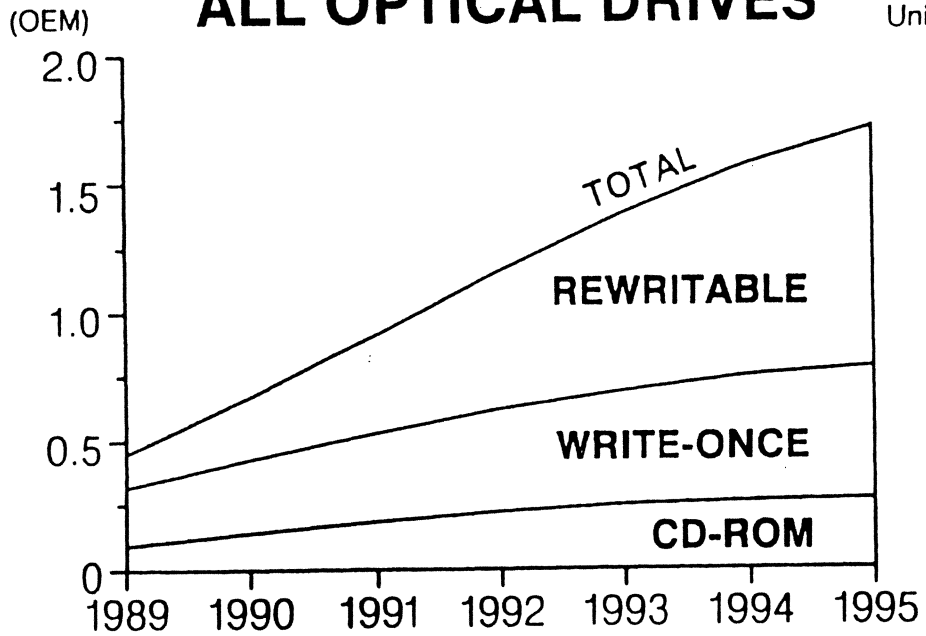
Chromium (absorber)

Garnet : high θ_F , low H_c , $\perp K_u$

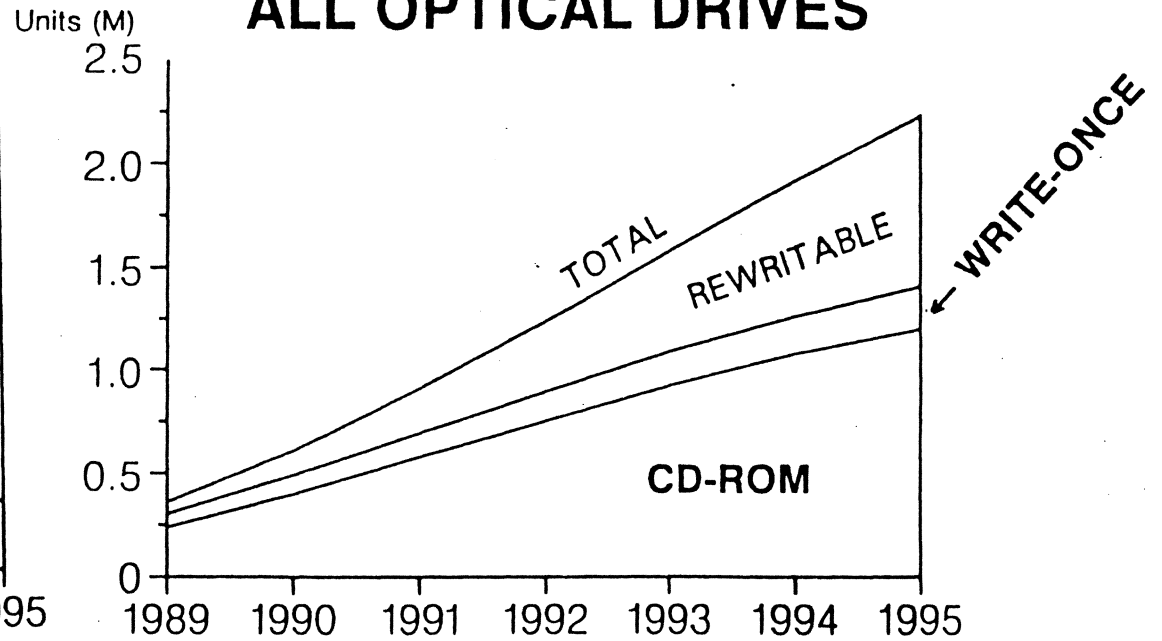
Metal : low θ_K , high H_c , $\parallel K_u$



WORLDWIDE REVENUE ALL OPTICAL DRIVES



WORLDWIDE SHIPMENTS ALL OPTICAL DRIVES



Storage Interfaces and Architectures

SCSI-2

- Overview
- New Features

Storage Architecture

- Evolution
- Optimizing Data Transfers
- Disk Arrays
- Disk Buffers and Caches

Storage Interfaces and Architectures

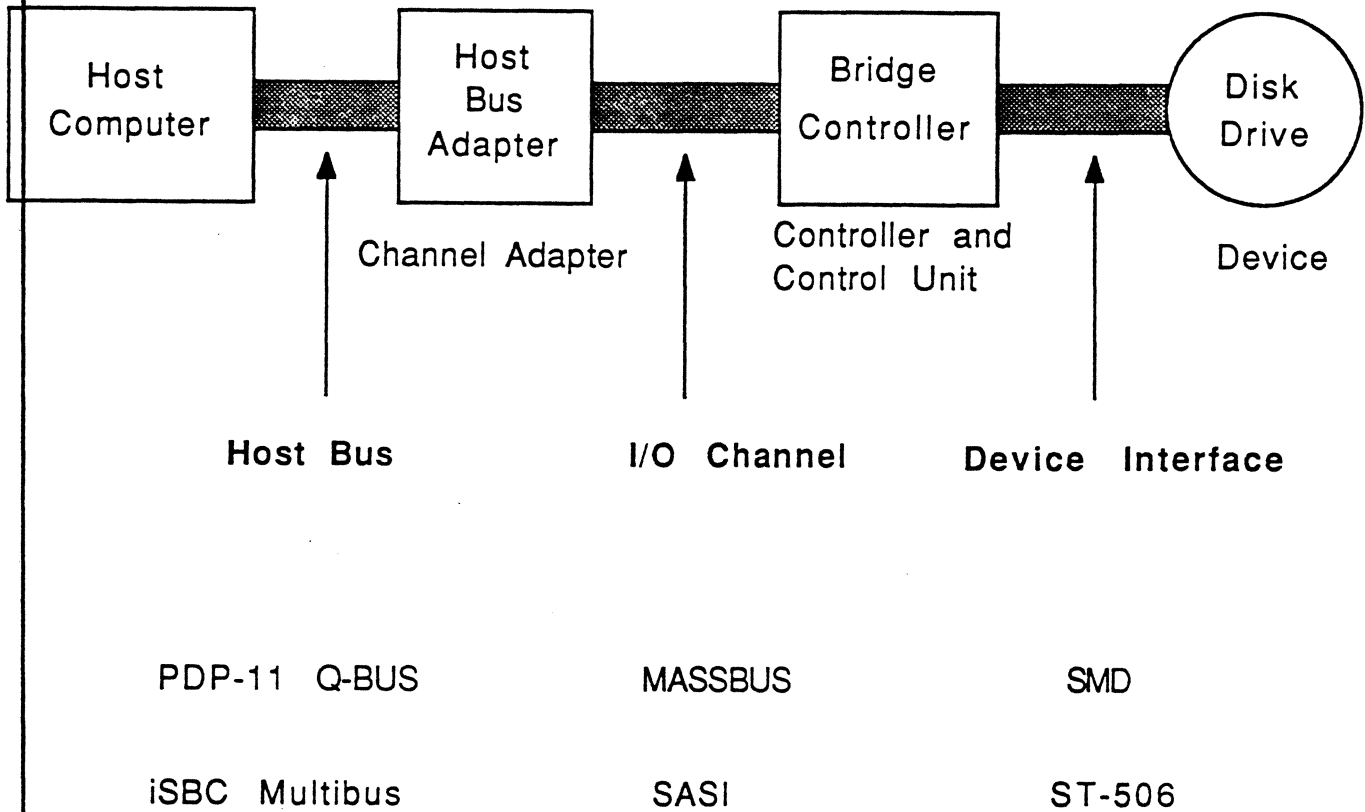
SCSI-2

- Overview
- New Features

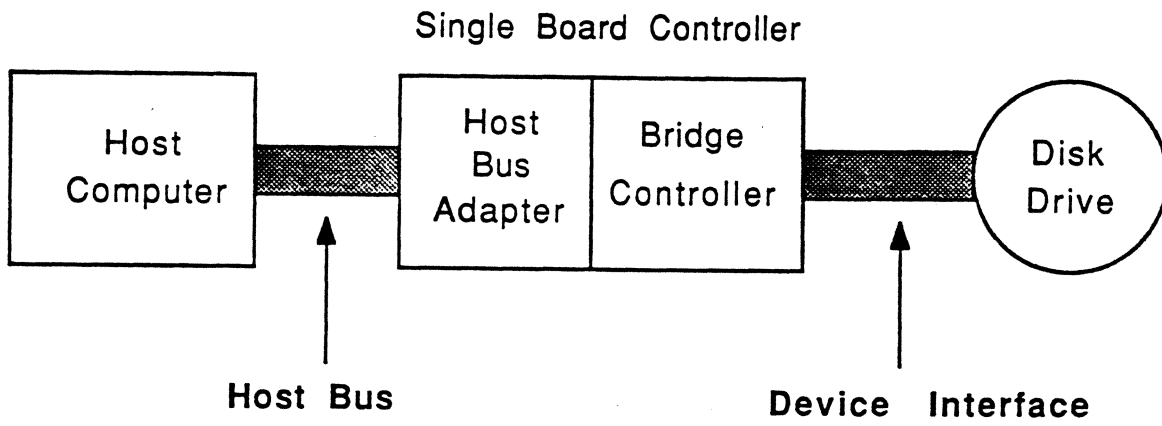
Storage Architecture

- Evolution
- Optimizing Data Transfers
- Disk Arrays
- Disk Buffers and Caches

Storage Architecture (1975)

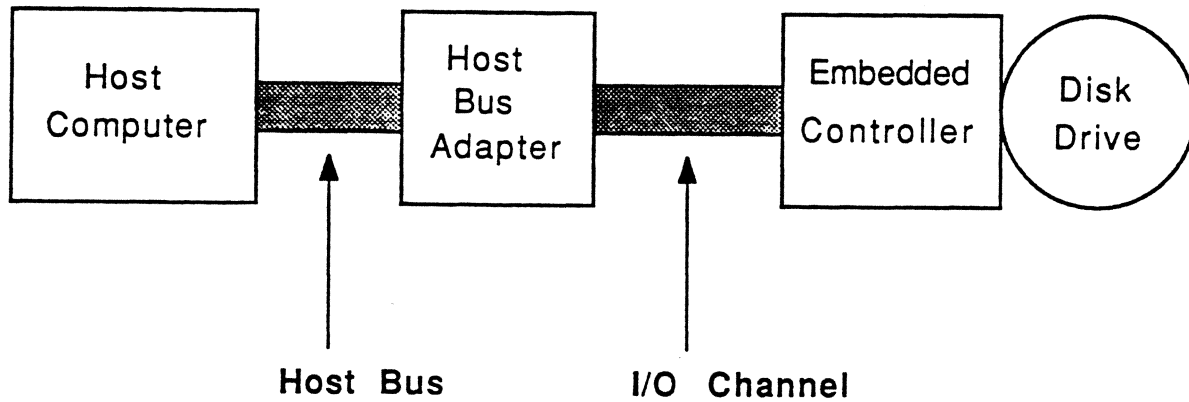


Storage Architecture (1980)



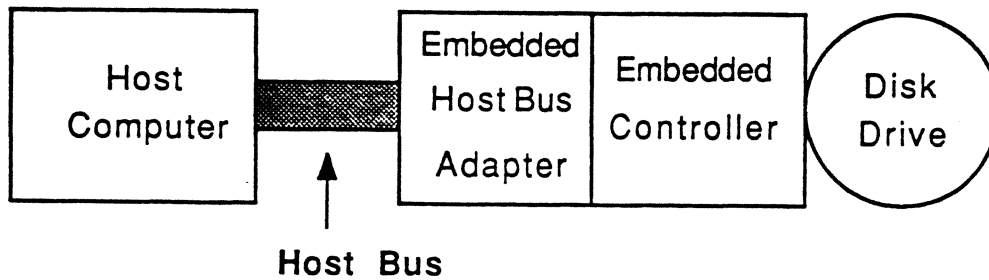
	Host Bus	Device Interface
LSI-11	Q-Bus	SMD
iSBC	Multibus	ST-506
IBM PC	PC BUS	ST-506
IBM PS/2	Micro Channel	ESDI

Storage Architecture (1985)



PC	PC-AT Bus	SCSI
PS/2	Micro Channel	SCSI
Sun	VME bus	SCSI
Vax	BI bus	SCSI
IBM	4300	370 oemi IPI-3 HPPI

Storage Architecture (1990)



- Put the HBA and controller on the Disk Drive
- Lower cost, higher reliability
- Most popular System Bus is IBM PC-AT (ISA) bus
- Motherboard manufacturers now including IDE connector with signals and power for attachment

ST-506

ESDI

SMD-E

Are Device Level Interfaces Dead ?

- Existing storage architectures require DLI drives for expansion
- Most vendors continue to sell DLI drives
- Customers not designing DLI drives into new systems
- Primary emphasis away from DLI drives

Storage Interfaces and Architectures

SCSI-2

- Overview
- New Features

Storage Architecture

- Evolution
- Optimizing Data Transfers
- Disk Arrays
- Disk Buffers and Caches

SCSI Vocabulary

Initiator - Host, Host DMA channel, HBA

Target - Controller/Control Unit and Device

CDB - Command Descriptor Block

LUN - Logical Unit Number

LBA Logical Block Address

HBA Host Bus Adapter

SCSI People

Andy Hospodor
IIST Santa Clara Univ (408) 554-6853

John Lohmeyer
SCSI-2 Chairman (316) 636-8703

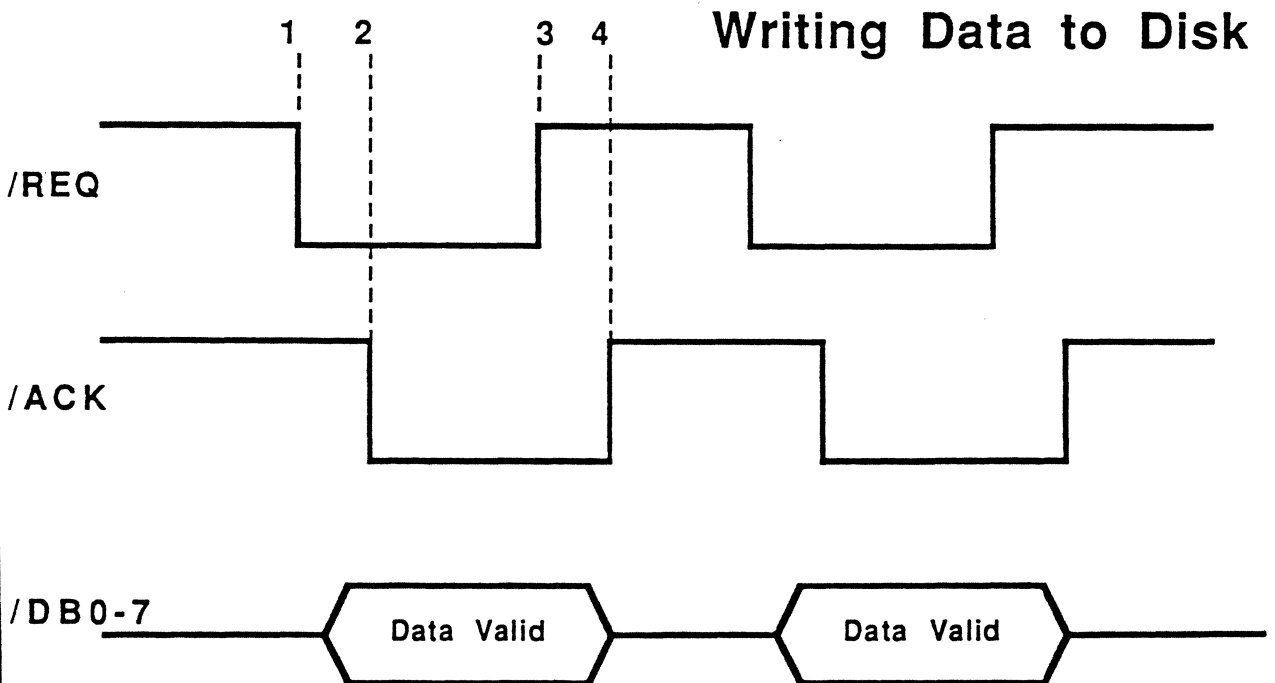
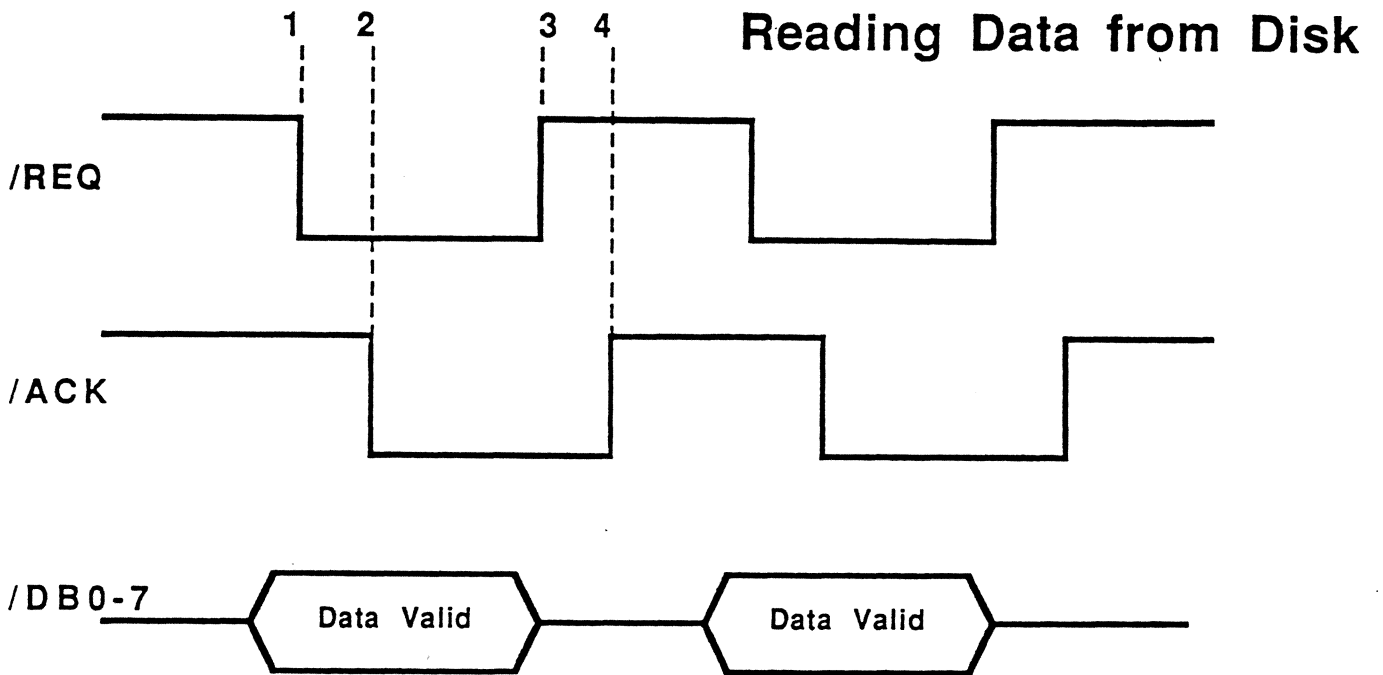
SCSI Bulletin Board
Download the latest spec (316) 636-8700

SCSI Phases

	SCSI Control Lines				
	SEL	BSY	C/D	I/O	MSG
Bus Free	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Select	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Command	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data In (Read from Disk)	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Out (Write to Disk)	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Status	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Message	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>

11ST

SCSI Phases



IIST

SCSI Req-Ack Handshake

1. Target Requests a Transfer
2. Initiator Acknowledges Transfer
3. Target Removes /REQ
4. Initiator Removes /ACK

What Made SCSI take off

1. Industry moved away from ST-506. Customers wanted data separators on the disk drives, leading to integration of the Read/Write Channel.
2. Microprocessor technology allowed a logical interface, error correction, and defect mgt.
3. Custom VLSI technology allowed electronics to fit on the disk drive. Controller, Servo, and Motor speed electronics integrated in 3-5 ICs.
4. Embedded controllers allow a generic interface to the host, while provided an optimal disk interface.

Hey, What About IPI ?

IPI-2 is a device level interface

IPI-3 is an I/O channel or system level interface

- Currently two domestic manufacturers shipping IPI-2 products. Only one IPI-3 product shipping
- Many manufacturers announcing IPI versions of new products
- Chip sets are sole sourced and priced prohibitively
- "If it doesn't look like a washing machine...
and doesn't run on 220 volts...
then it isn't a high performance solution"

SCSI-2

Standard Expected:1990

- 1. New Mandatory Commands**

Separate command sets for: Disk, Tape, WORM, CD-ROM

- 2. New Messages**

- 3. Command Queuing**

- 4. Faster Data Transfer Rates**

SCSI Command Descriptor Block

6 Byte Command

OPCODE	
LUN	LBA
LBA	
LBA	
XFER len	
Control	

10 Byte Command

OPCODE	0	
LUN	resrv	1
LBA		2
LBA		3
LBA		4
LBA		5
reserved		6
XFER len		7
XFER len		8
Control		9

IIST
A Standard 10/06

OPCODE - operation to be performed

LUN - Logical Unit Number
=0 for embedded SCSI drives

LBA - Logical Block Address

XFER len - Number of blocks to be transferred

SCSI Command Descriptor Block Examples

6 Byte Command

0 8	
LUN	1 4
C 3	
6 3	
0 1	
Control	

Read one block
starting at block 14C363

10 Byte Command

2 A		0
LUN	resrv	1
0 1		2
2 3		3
4 5		4
6 F		5
reserved		6
8 0		7
0 0		8
Control		9

Write 8000h blocks
starting at block 0123456F

Ten Byte commands address larger devices and offer longer transfer counts, making them more suitable for backup operations.

Mandatory Messages

SCSI (1986)

Command Complete

SCSI-2 (1990)

Abort

Bus Device Reset

Command Complete

Identify

Initiator Detected Error

Message Parity Error

Message Reject

No Operation

IIST

Mandatory Commands

SCSI (1986)

Test Unit Ready

Request Sense

Format

Read

Write

SCSI-2 (1990)

Inquiry

Test Unit Ready

Send Diagnostic

Request Sense

Format

Read

Write

Reserve

Release

IIIST

Command Queuing

1. Untagged queuing allows single commands from multiple initiators to be queued
2. Tagged queuing allows multiple commands from one or more initiators to be queued. Each command has an associated Queue Tag
 - a. Commands with Ordered Queue Tags must be executed in the order received. (FIFO)
 - b. Commands with Head of Queue Tags are placed first in queue, to be executed next. (LIFO)
 - c. Commands with Simple Tags are executed in an order determined by the target.

Advantage: **Seek Optimization**

Disadvantage: **Indefinite Postponement**

LIST

Data Transfer Rates

1.	Asynchronous SCSI	2.0 MBytes/s
2.	Synchronous SCSI	5.0 MBytes/s
3.	Fast SCSI-2 *	10.0 MBytes/s
4.	Wide SCSI-2 *	40.0 MBytes/s

* - Requires Differential Drivers on SCSI cable
Maximum data transfer rate also limited by:

Noise

Cable Impedance

Termination

New SCSI-2 Mode Select Pages

- Caching Page
- Notch and Partition Page
- Read/Write Error Recovery Page
- Verify Error Recovery Page

SCSI-2

Reconfigurability

- Medical Imaging System
- Database system search/sort/query
- Workstation used for I.C. layout
- Verify Error Recovery Page

Application software can use Mode Select to optimize caching strategy for specific job.

Storage Interfaces and Architectures

SCSI-2

- Overview
- New Features

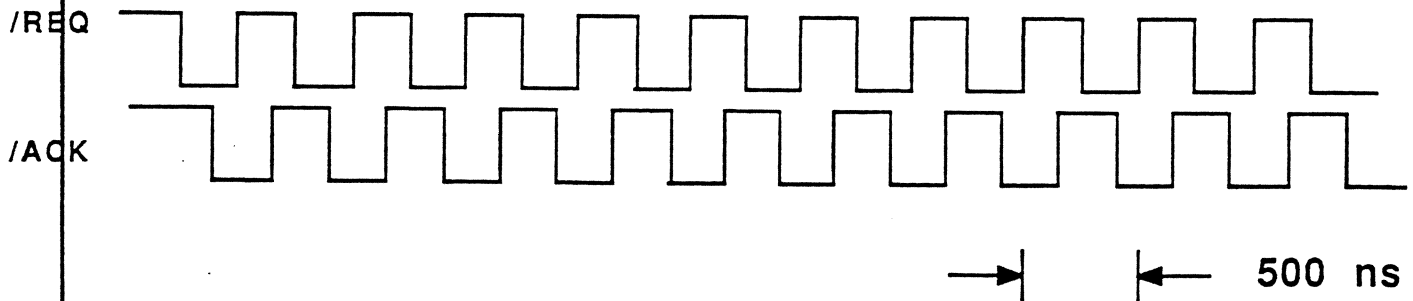
Storage Architecture

- Evolution
- **Optimizing Data Transfers**
- Disk Arrays
- Disk Buffers and Caches

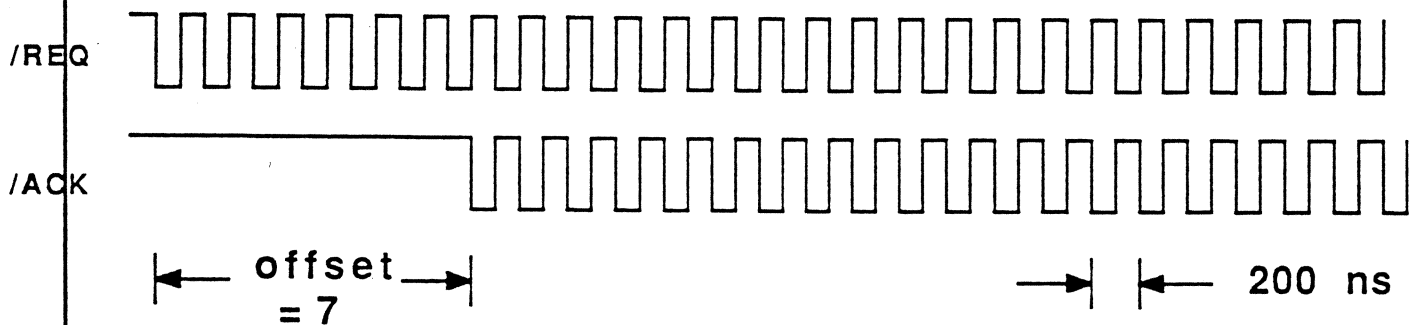
Data Transfer Rates

Comparison of SCSI and SCSI-2

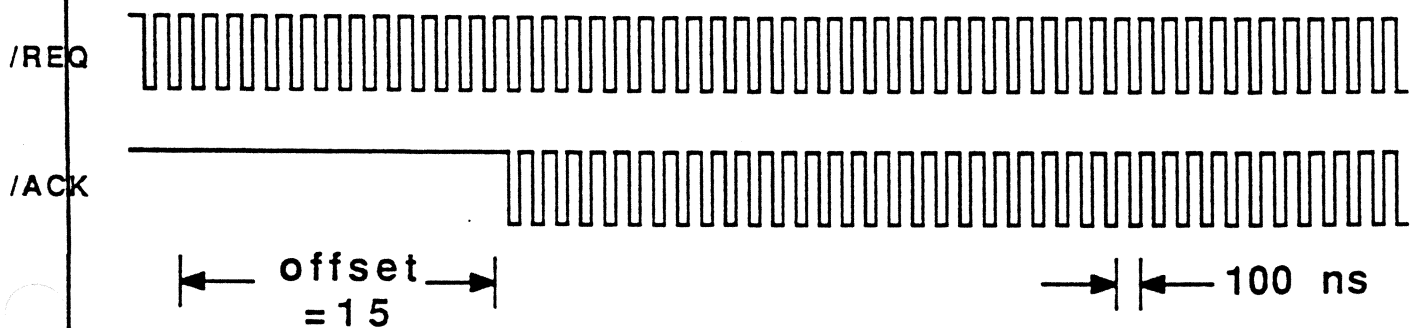
Asynchronous SCSI



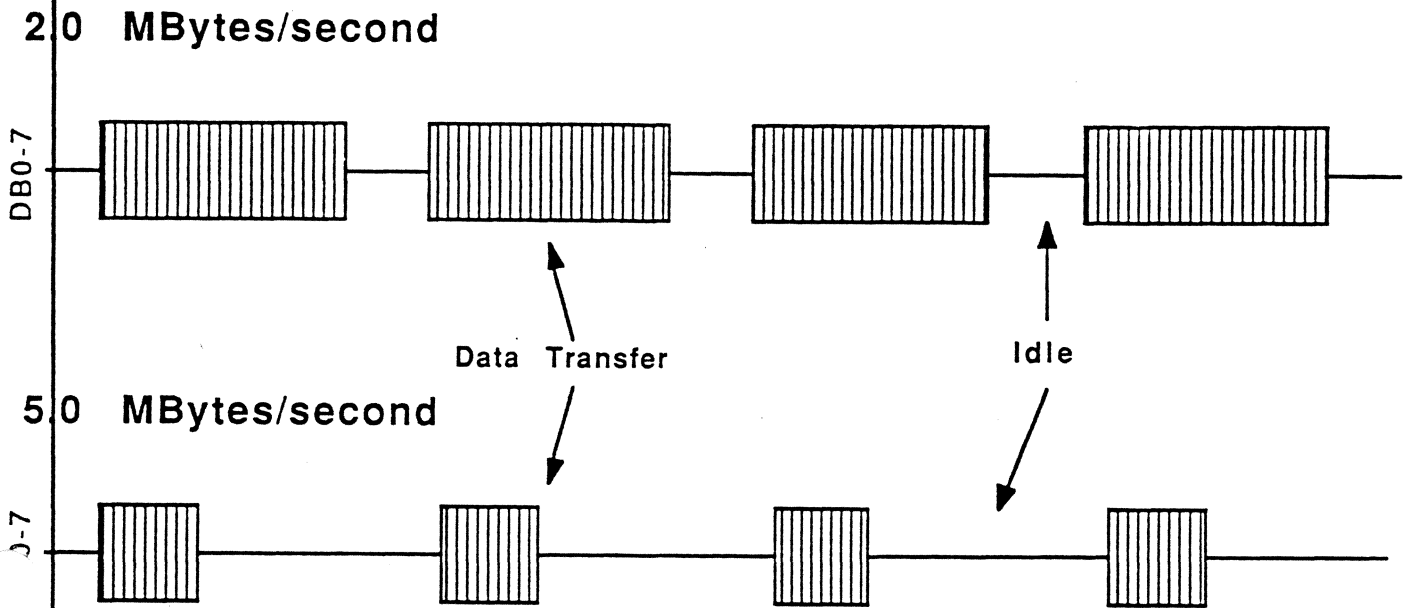
Synchronous SCSI



Fast Synchronous SCSI-2



Thread Management and Data Path Management



Shorter periods of high activity with longer idle periods

Why not relinquish the bus during these idle periods ?

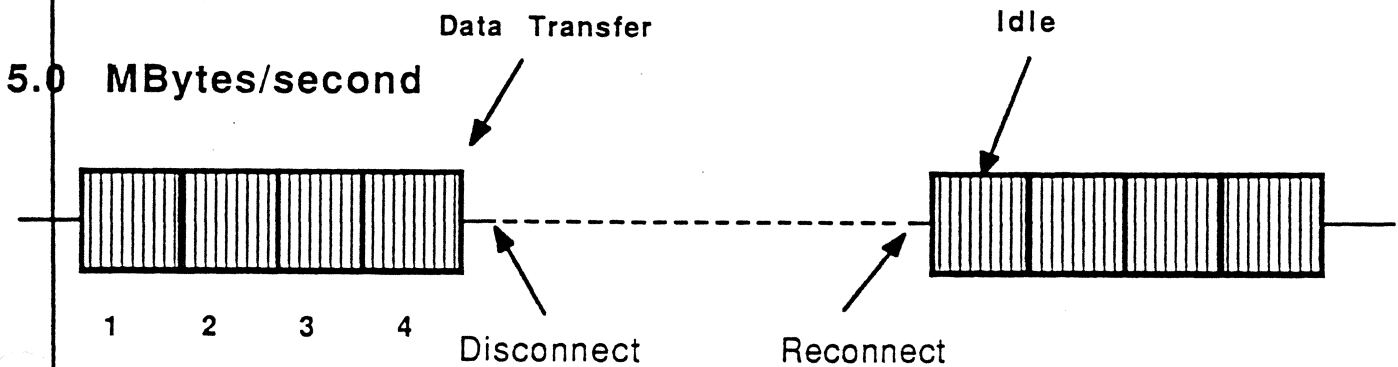
1. Target must send disconnect message to initiator 200
2. Initiator must send ID message to Target 200
3. Target must reconnect 100
4. Initiator must restore pointers & resume 300

Need 800 microseconds between blocks to use disconnect/reconnect

Efficient

Thread Management and Data Path Management

IIST
A. Hosnodor 12/90



Large data transfers may be packed together, allowing efficient disconnect/reconnect. In this example 4 blocks are transferred during each burst, then the SCSI bus is relinquished.

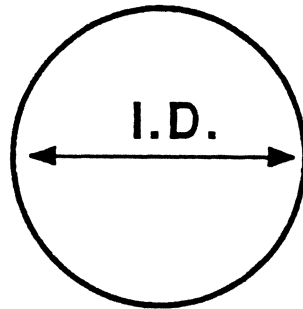
A winchester disk with 1 block = 512 bytes, and 32 blocks/track, 3600 rpm

Each block would take $16.67/32 = 520$ microseconds to read from disk

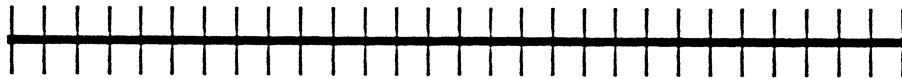
A group of 4 blocks would take 2080 microseconds to read from disk
and would take 409 microseconds to burst over SCSI

Leaving about 1500 microseconds to disconnect/reconnect

Notched Recording

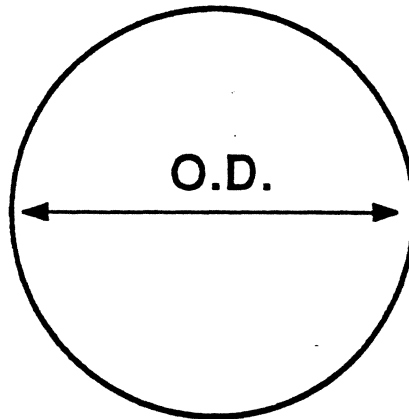


Inner Track

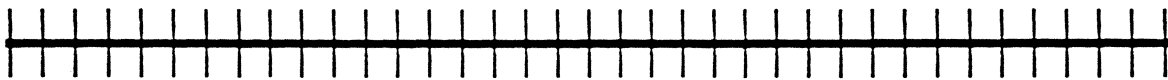


28 Sectors per Track

3600 RPM, 7.5 Mb/s



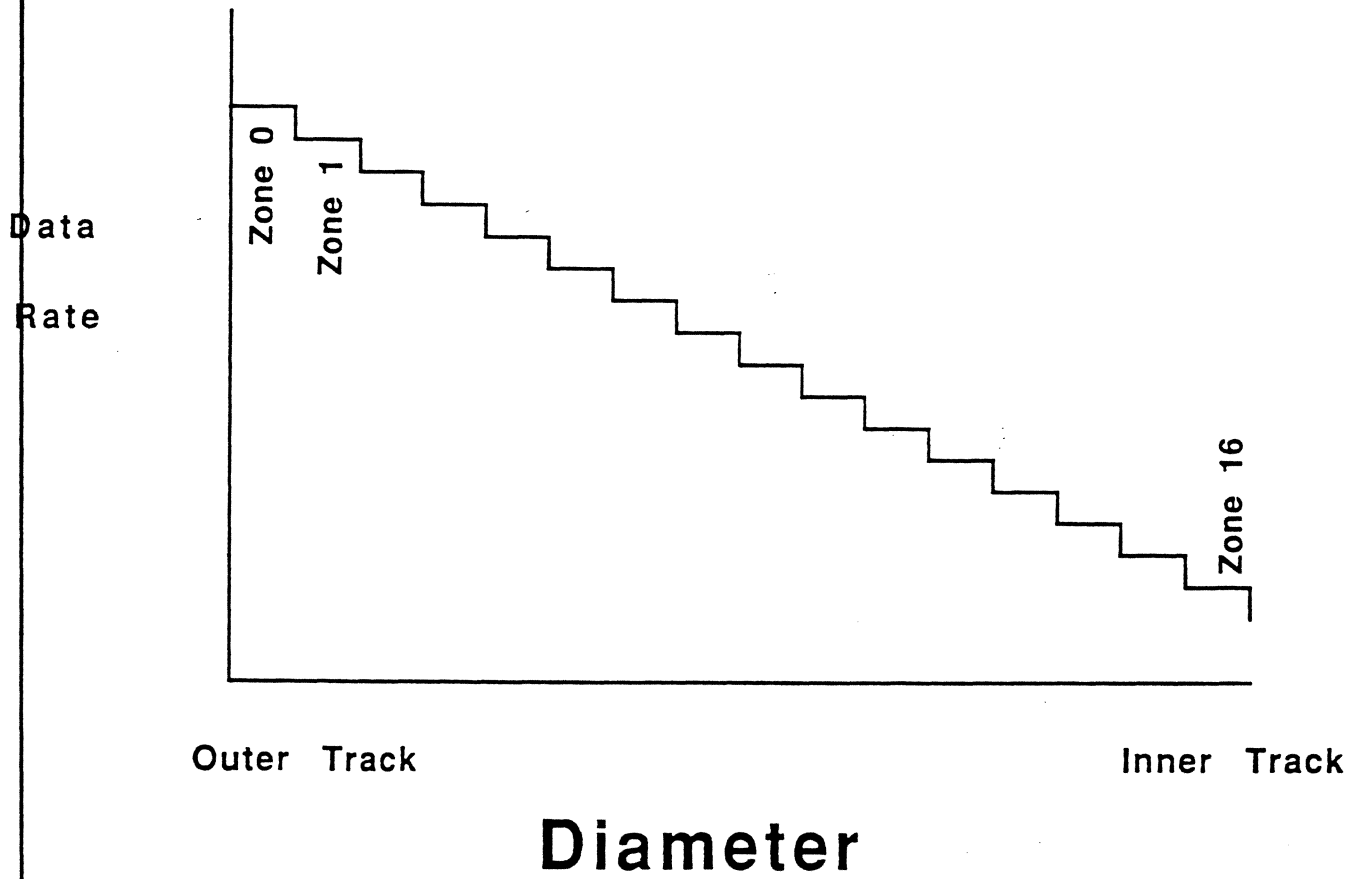
Outer Track



36 Sectors per Track

3600 RPM, 9.6 Mb/s

Notched Recording Data Rates



IIST

Storage Interfaces and Architectures

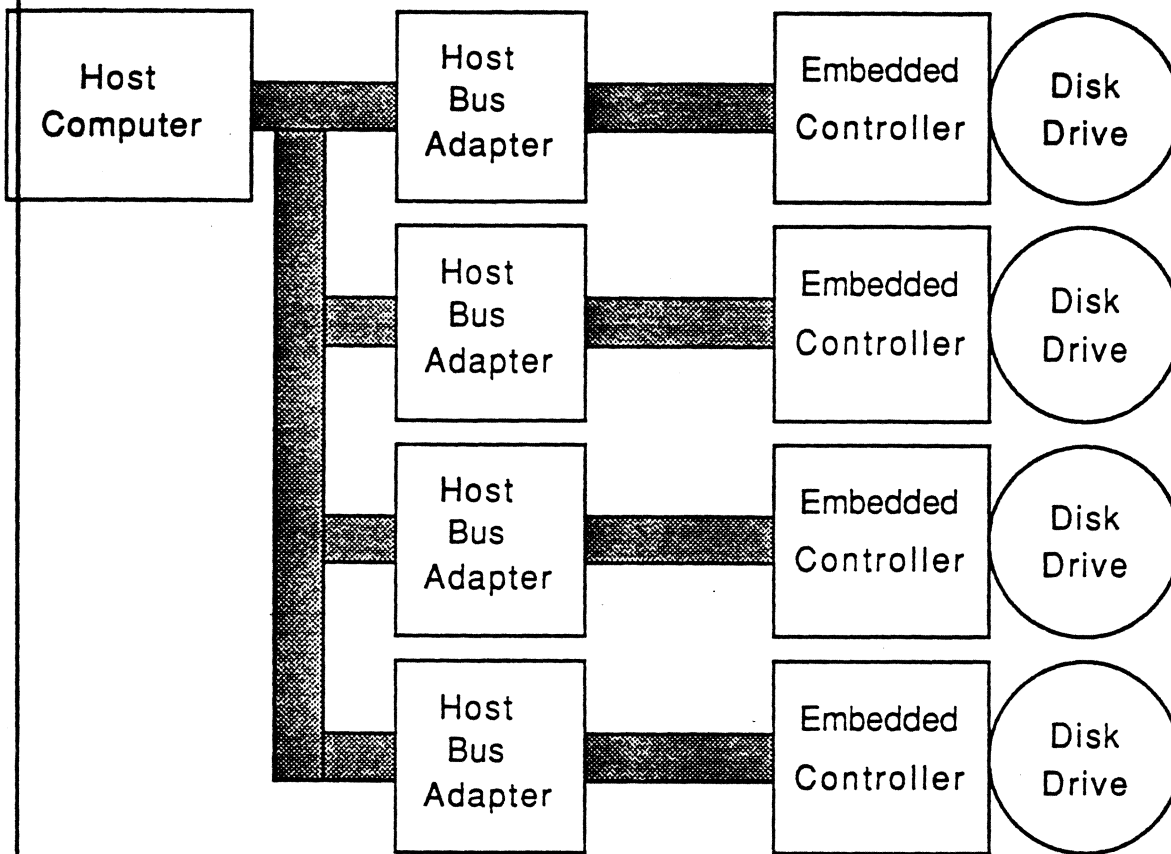
SCSI-2

- Overview
- New Features

Storage Architecture

- Evolution
- Optimizing Data Transfers
- **Disk Arrays**
- Disk Buffers and Caches

Striped Disks



- **Increases Throughput performance**
- **Files are broken up and spread out**
- **Least reliable way to store data may be implemented with mirroring**

Storage Interfaces and Architectures

SCSI-2

- Overview
- New Features

Storage Architecture

- Evolution
- Optimizing Data Transfers
- Disk Arrays
- **Disk Buffers and Caches**

Where Should the Cache go ?

1. On the Disk Read/Write channel
2. On the Disk Controller
3. On the Host Bus Adapter
4. In Main Memory

Semiconductor Memory

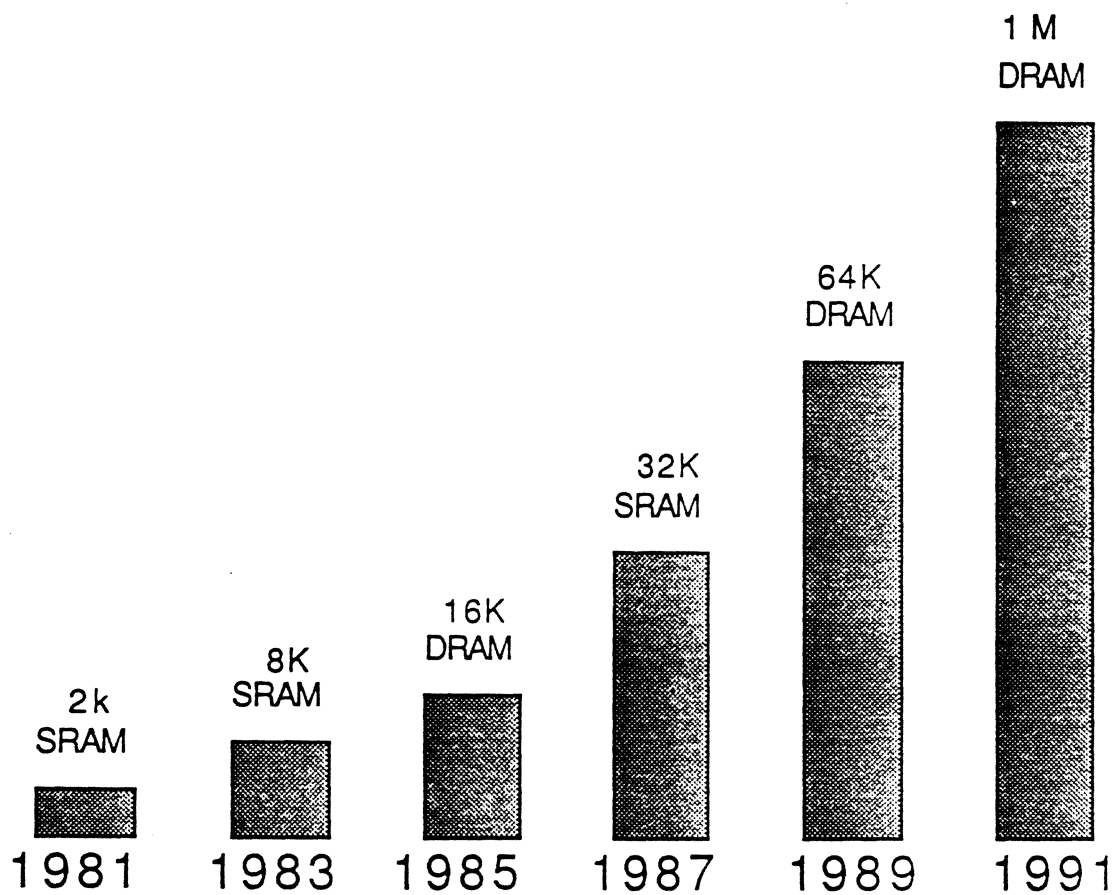
On Disk Drives

- Minimum RAM data buffer - commodity drives
- Caching RAM data buffer - high end drives
- All RAM data buffer, no disk - Solid State Disk

1. How much buffer does your system require ?

2. Buffer Size versus Performance

Disk Drive Buffer Sizes



Low end: 2k-8K Static Ram Buffers

High end: 1Mx9 DRAM SIP array caching buffer

Cache Performance

The access time equation:

$$t_{\text{access}} = t_{\text{overhead}} + t_{\text{seek}} + t_{\text{rot_latency}} + t_{\text{transfer}}$$

In the ideal case:

$$t_{\text{access}} = t_{\text{overhead}} + t_{\text{cache_overhead}} + t_{\text{transfer}}$$

Let $p(h)$ = probability of a cache hit

$$p(m) = 1 - p(h)$$

$$t_{\text{access}} = t_{\text{overhead}} + t_{\text{seek}} + t_{\text{rot_latency}} + t_{\text{transfer}} + p(h) [t_{\text{cache_overhead}} - t_{\text{seek}} - t_{\text{rot_latency}}]$$

Design question:

What is the minimum $p(h)$ required to break even ?
or How much Cache Overhead can be tolerated ?

Tuning a Cache Buffer

Cache Mode Select Page

Enabling Fast Write where appropriate

Minimum prefetch length

Maximum prefetch length

Disconnect-Reconnect Page

Buffer Full/Empty Ratios

Disconnect/Connect time limits

Maximum length of data bursts

Application software has the ability to control:

What data goes into the cache

How and when the data comes out of the cache

Predictions

1990 - 2000

1. 5.25" read/write channels will expand to envelope SCSI synchronous data rate (5.0 MByte/sec)
2. Hardwired decoding of SCSI commands on chip sets
3. Cache buffers using 1Mbx9 and 4Mbx9 DRAM arrays
4. Integration of SSD & Mag Disk = Mixed media devices
5. Multiple head/slider - Paired Read/Write channels
6. Disk arrays - for those who can afford them
7. Key performance issue: Rotational Latency
Higher Rotation Rates (4800, 5400 RPM)
Ability to read sectors out of sequence

TRENDS IN DATA STORAGE

A.S. HOAGLAND
GEOFFREY BATE

IIST

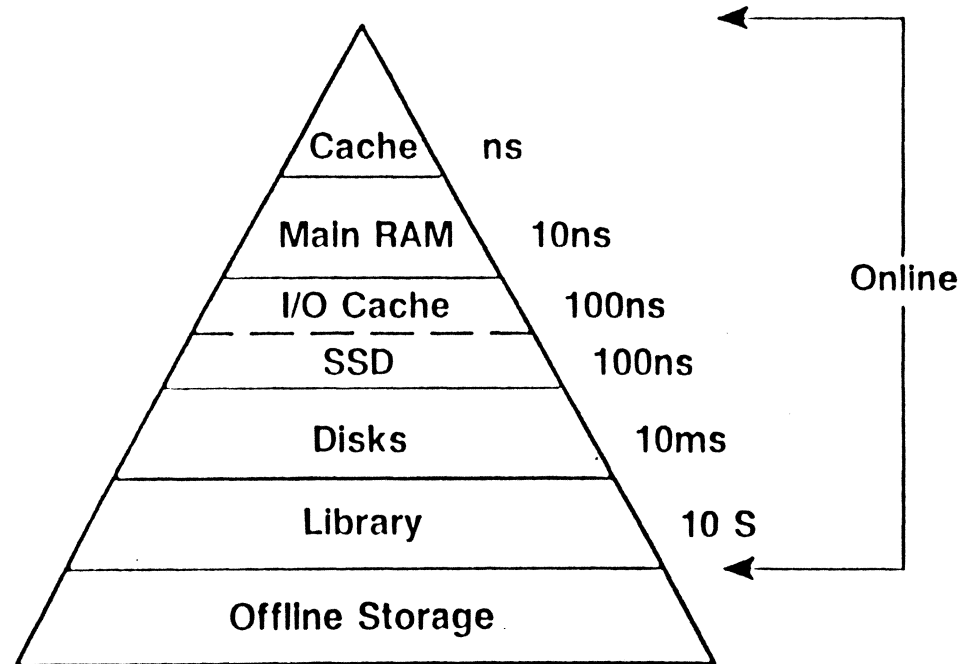
ASH & GB 6/89

Trends in Data Storage

TABLE OF CONTENTS

- * ADVANCES IN EXISTING TECHNOLOGIES
- * PROBLEMS WITH OPTICAL RECORDING
 - magnetic recording resurgent
- * TRENDS IN OPTICAL RECORDING
 - integrated heads
 - short-wavelength lasers
 - high Kerr-rotation media
 - multi-layer films
- * OPTICAL AND MAGNETIC RECORDING DENSITY LIMITS
- * COMPETITION FROM SEMICONDUCTORS
- * STORAGE HIERARCHY IN 2001
- * COMPUTER MEMORY EVOLUTION
- * BIOLOGICAL STORAGE

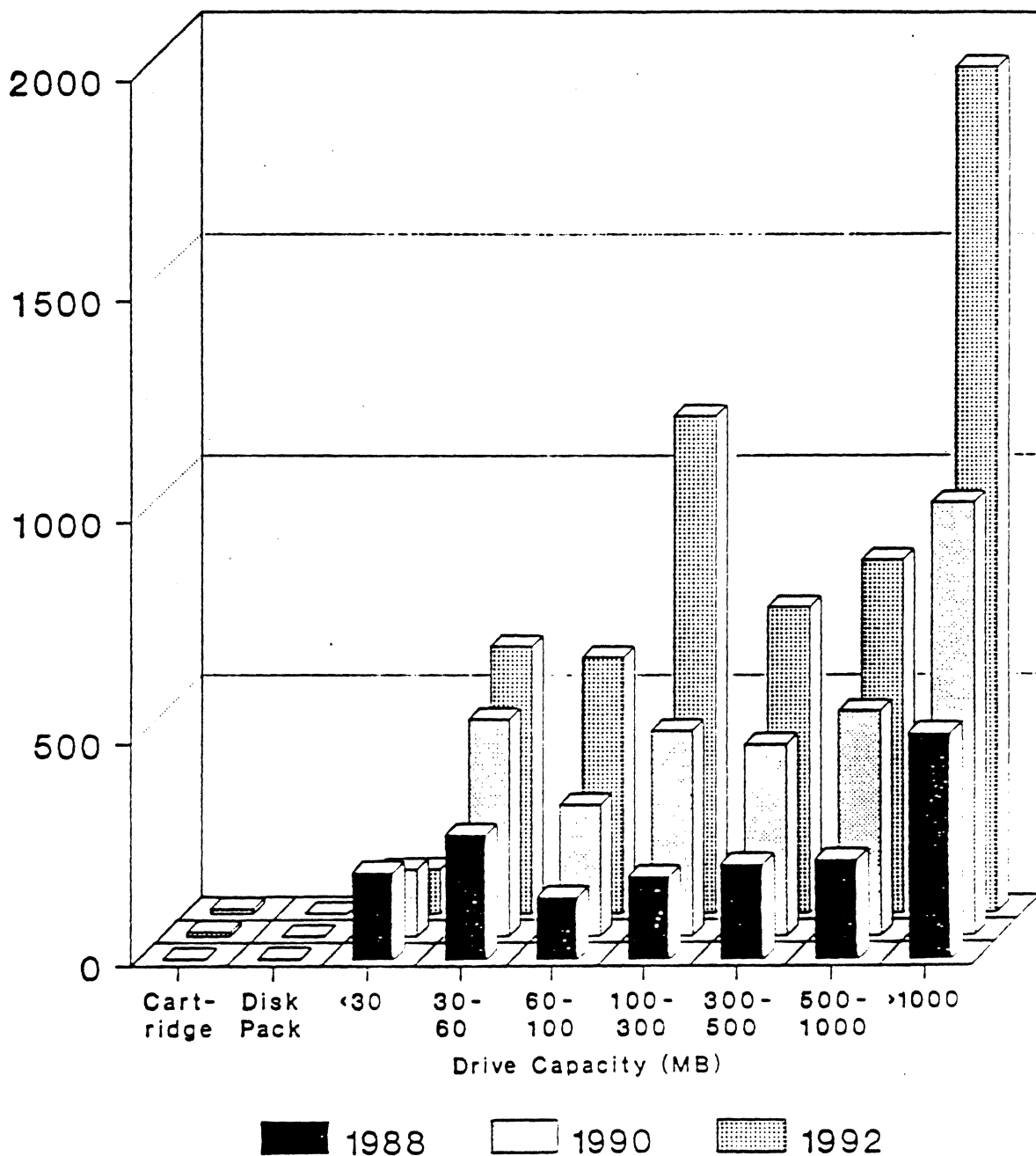
STORAGE HIERARCHY - CURRENT ARCHITECTURES



Constraint: Storage Technology

CAPACITY SHIPMENT SUMMARY

Worldwide Shipments in Terabytes

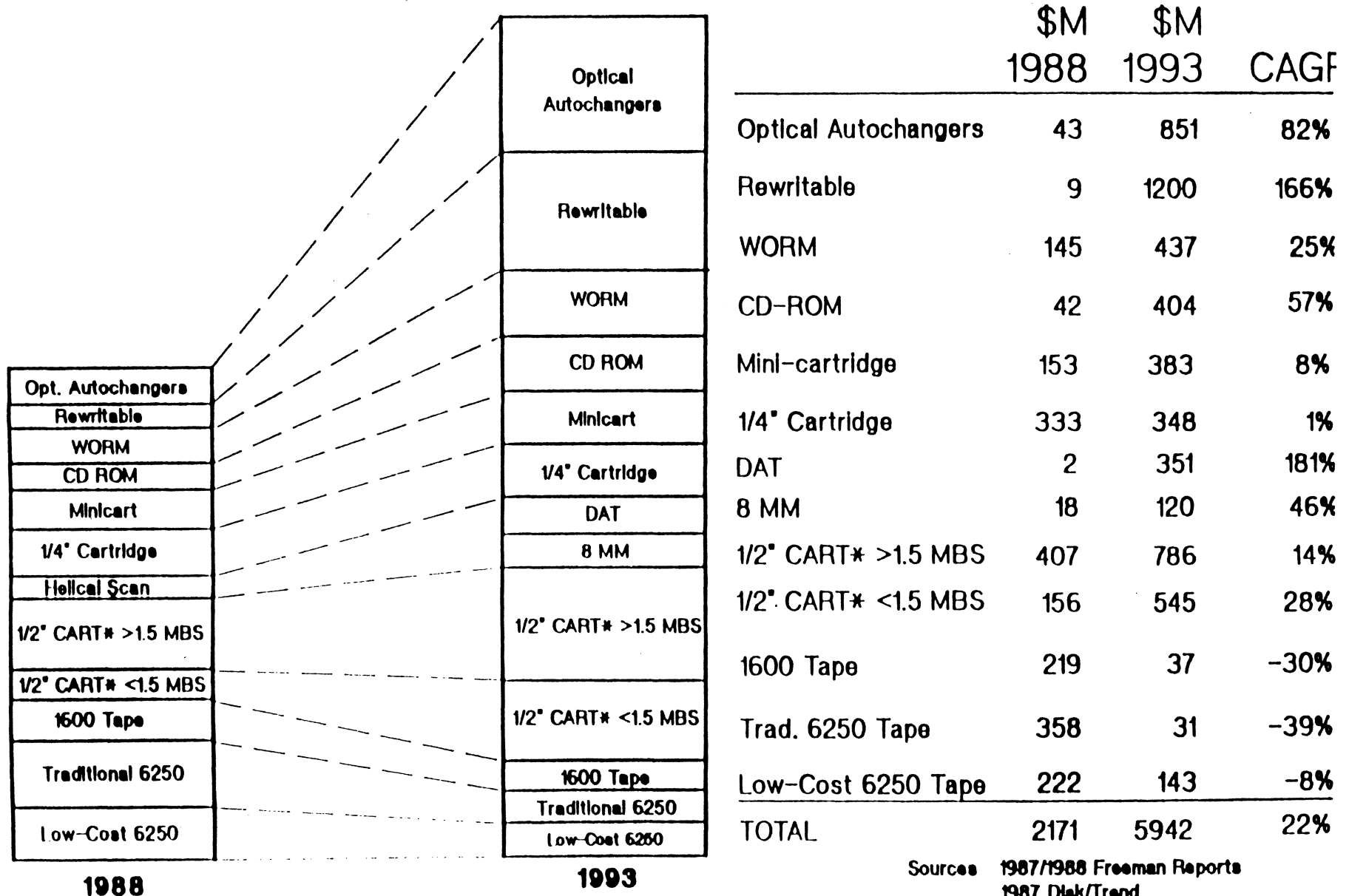


1989 DISK/TREND REPORT

IIST

A. Hoagland 3/90

Secondary Storage Market Growth Estimates (In OEM \$)



1988
Growth

* Dominated by 3480 cartridges

Source: 1987/1988 Freeman Reports
1987 Diak/Trend
1987 BIS Macintosh

Industry Structure:

LIMITED VERTICAL INTEGRATION . . . still
COMPONENT VENDORS OWN CRITICAL TECHNOLOGIES

DRIVE MANUFACTURERS

Today: Consolidation / Partnerships

Non-Captive -- boxes

Captive -- storage systems

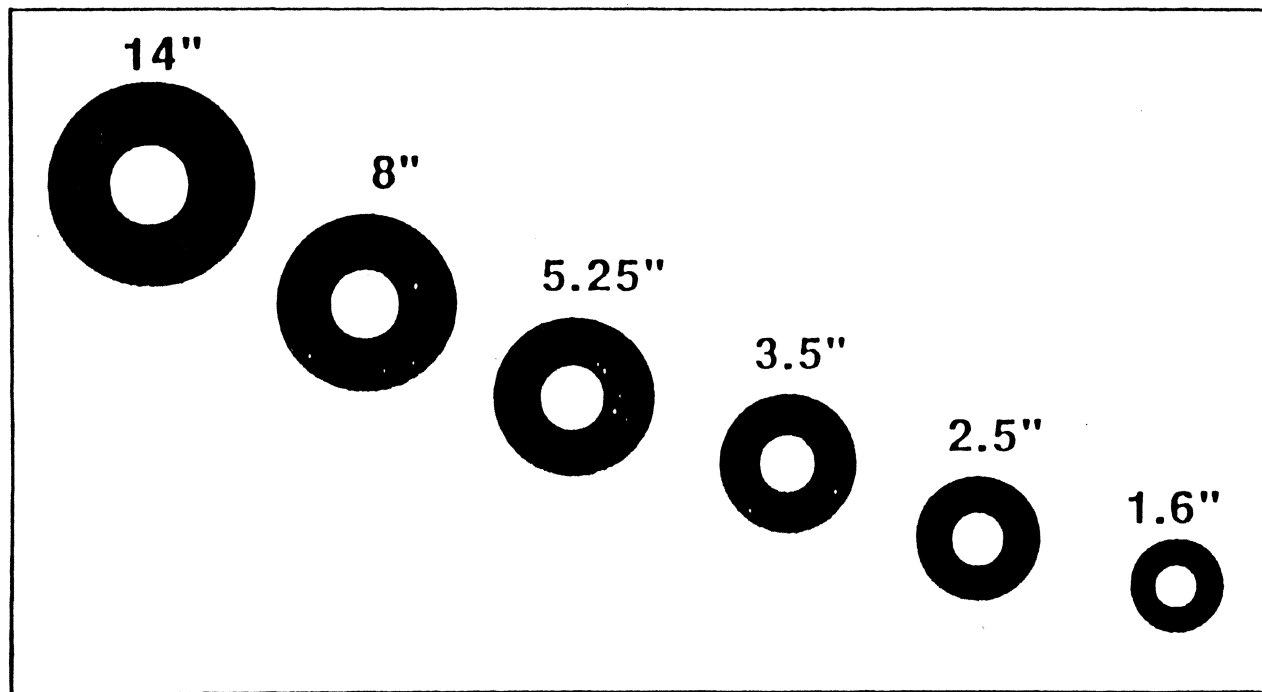
TRENDS: R & D . . . U.S.A.
 MANUFACTURING . . . OFF-SHORE
 HEADS & MEDIA . . . JAPAN

IIST

A.S. HOAGLAND
FMR

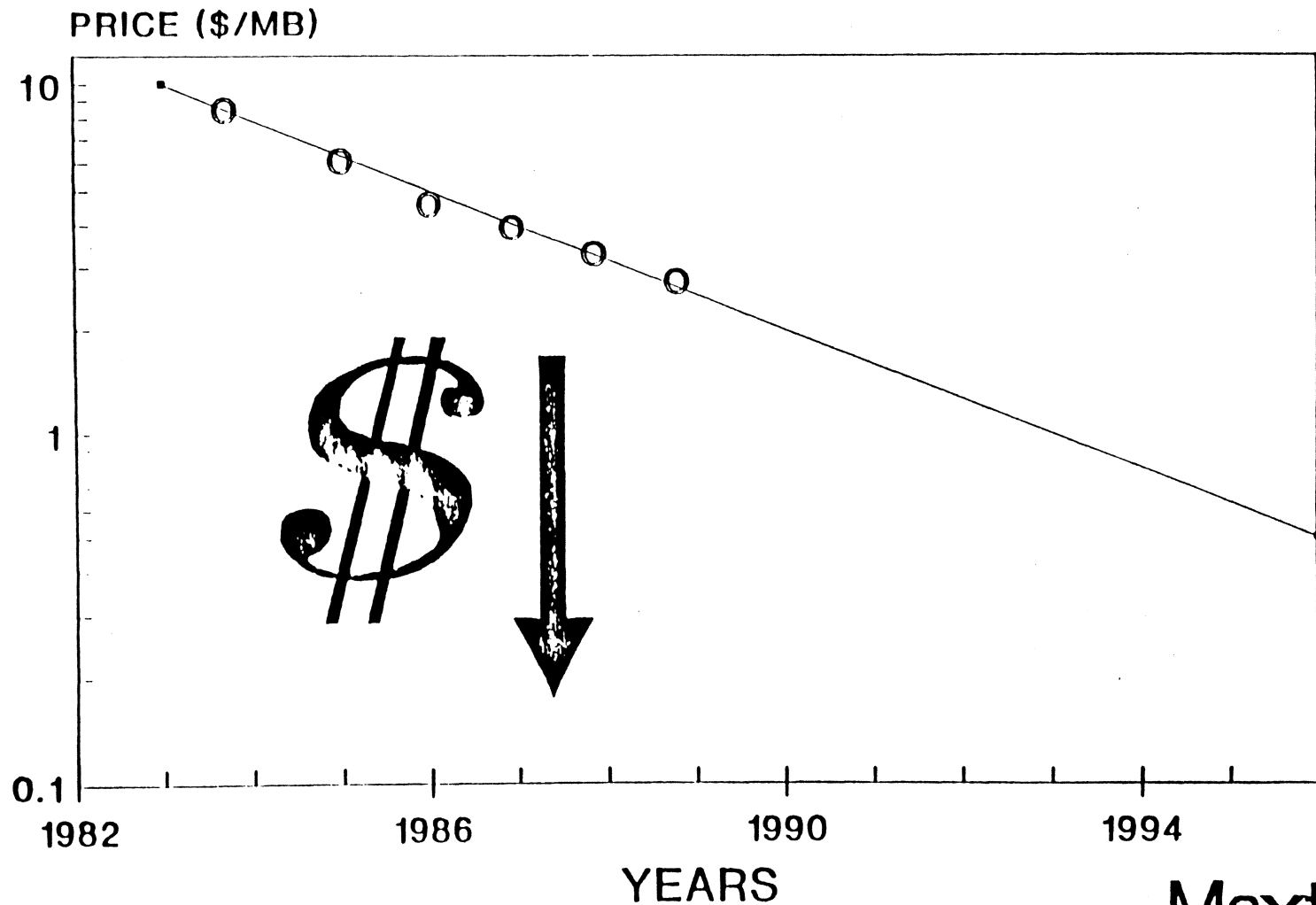
Disk Size Trends

Size



Time →

OEM Price per Megabyte



Maxtor

IIST

Abstract 222

State of the Art:

<u>Disk</u>	OXIDE	to	THIN FILM (startups)
	5.25"	to	3.5" (Apple)
	OPEN LOOP	to	SERVO
	DEDICATED	to	IMBEDDED
	ST506	to	SCSI
	MTBF 5,000 - 10,000 hrs.	to	50,000 hrs.

<u>Tape</u>	LONGITUDINAL	to	ROTARY
-------------	--------------	----	--------

<u>Optical</u>	TALK	to	DELIVERIES
----------------	------	----	------------

Example of change:	5.25" disk	5 MB	to	1.2 GB
		85 ms	to	11 msec.
		5 MBits/sec.	to	2+ MBytes/sec.
		\$20-30/MB	to	\$2-3/MB

IIST

Technology Dynamics:

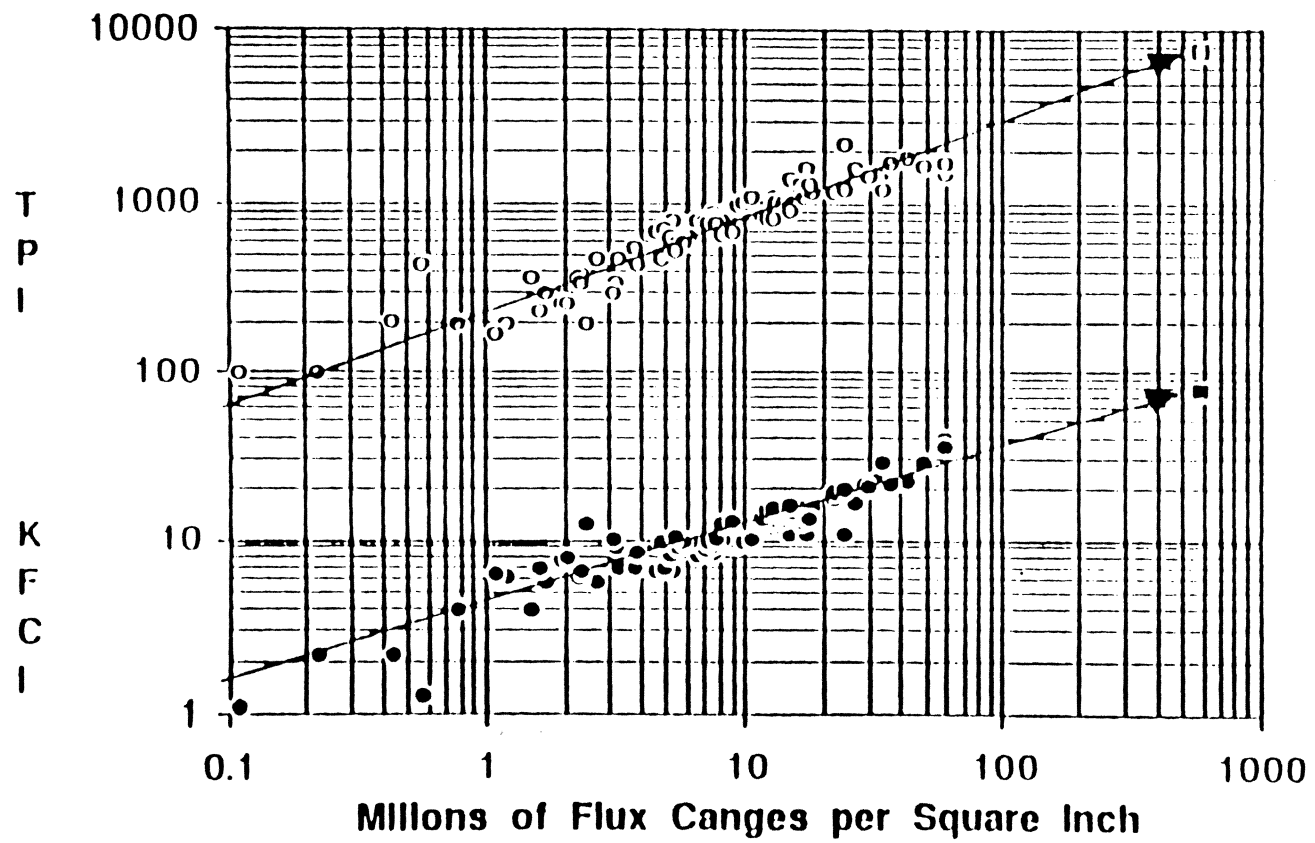
MAINSTREAM

Thin Film Disk (startups)
MIG heads (Japan)

UNDERWAY

Glass Disk	... Areal (NSG)
ZBR	... Imprimis
Rotary Drive	... Exabyte
2.5" Hard Disk	... Prairetek
2" Drive	... Sony
Optical Servoing	... Insite
MR Head	... PCI
Horizontal Head	... Leti
MET	... Japan, Inc
Disk Arrays	... Everyone !!

IIST



- KFCI
- TPI
- Est KFCI
- Est TPI

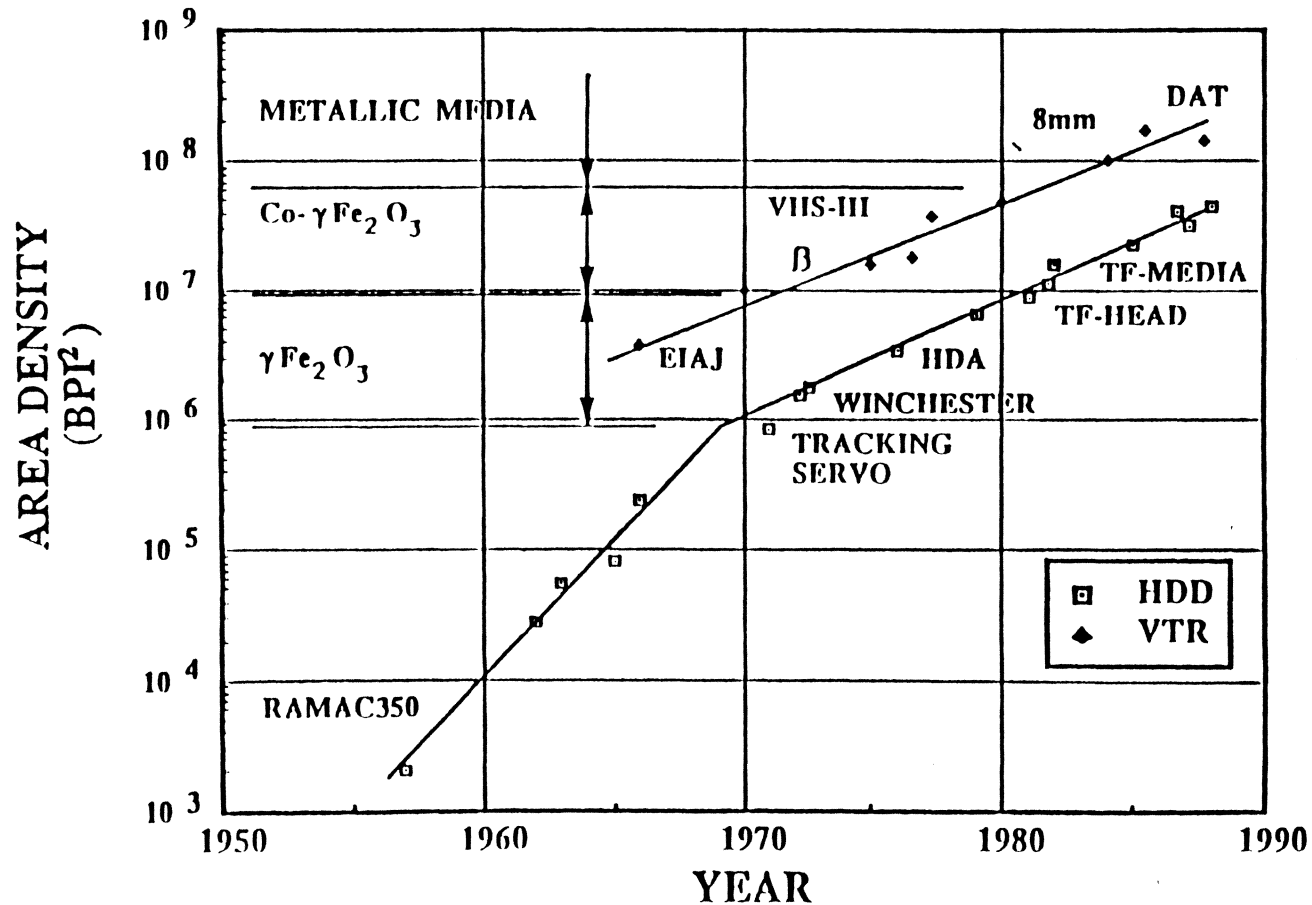
Sony Corporate Research Labs

Institute for Information Storage Technology

Workshop V October 11-14, 1989

Presented by Senri Miyaoka

YEARLY ADVANCE of RECORDING DENSITY

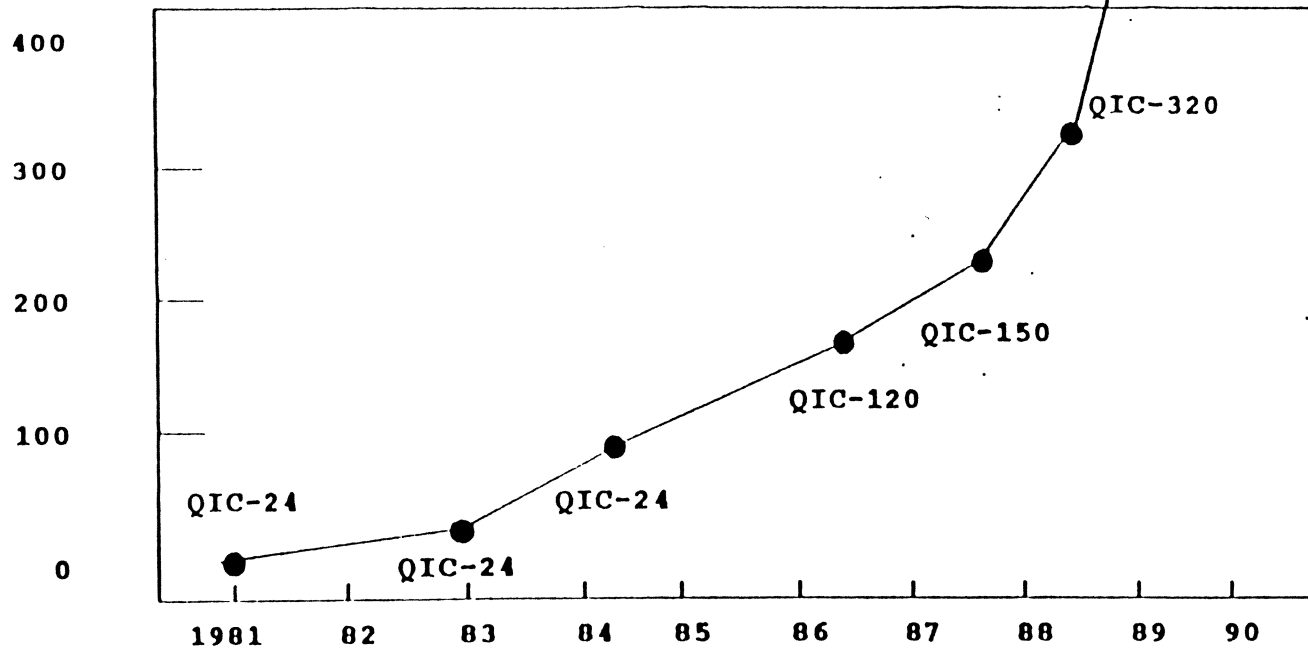


HELICAL SCAN RECORDING

	<u>8mm</u>	<u>4mm</u>	<u>VHS (1/2")</u>
Drum Diameter	40mm	30mm	62mm
Wrap Angle	221°	90°	180°
Track Width	25 micron	20 micron	32 micron
Gap Depth	15 micron	15 micron	30 micron
Head Tape Speed (ips)	148	124	230
FCI	77,000*	76,200	41,700
Azimuth Angle	±10°	±20°	±6°
Tape Speed (ips)	.56	.32	1.31
*Hi8	103,800		
*Exabyte	54,000		

CAPACITY INCREASE
(DC-600 CARTRIDGE)

CAPACITY (MB)



REMOVABLE HARD DISK DRIVES

Evolution of Technology and Capacity Examples

Drive	Capacity/Pack	# of disks	TPI	BPI
IBM 1131	1.024 MB	1	100	1100
Burroughs 9480	4.68 MB	1	100	4400
Ampex DM-442	3.125 MB	1	200	2200
IBM 3340	34.9 MB	2	300	5636
CII D120	10 MB	1	500	4750
CDC Lark	8.35 MB	1	237	10161
DMA	6.75 MB	1	454	8617
SQ306	6.38 MB	1	435	12000
SQ312	12.76 MB	1	741	12608
SQ555	54.78 MB	1	1086	23642

FK 3/9/90

IIST

A. Hoagland 3/90

Heads & Media -- Long Term

As **Track Density** Limits Get More Intensive Focus

- Will require "pre-formatted" substrates (a la optical disk)
- Disks in turn will depend upon multi-element transducers
 - To servo during read and write
 - To separately optimize read and write elements
 - To permit erase option to improve overwrite
- Will accelerate move to thin film head devices

Technology for the Mid and Late 1990's

Media: Sputtered Thin Co-Alloy Films
Glass Substrates
Possibly With Discrete Tracks

Heads: Integrated Magnetoresistive
(MR) Read, Inductive Write Heads

Source: Applied Magnetics Corp.

- The User
 - Lowest cost/bit
 - Read/Write and nonvolatile
 - Competitive marketplace-large choice of product offerings

- The Manufacturer
 - Multi-billion dollar industry
 - Established production processes
 - Demand for capacity increases faster than storage density from technological progress—growing market

- R&D
 - Areal density still far from ultimate limits
 - Fairly well understood phenomenon
 - Reversible process—inherently stable

Other Directions

High Velocity "Contact" Recording

Vertical Recording

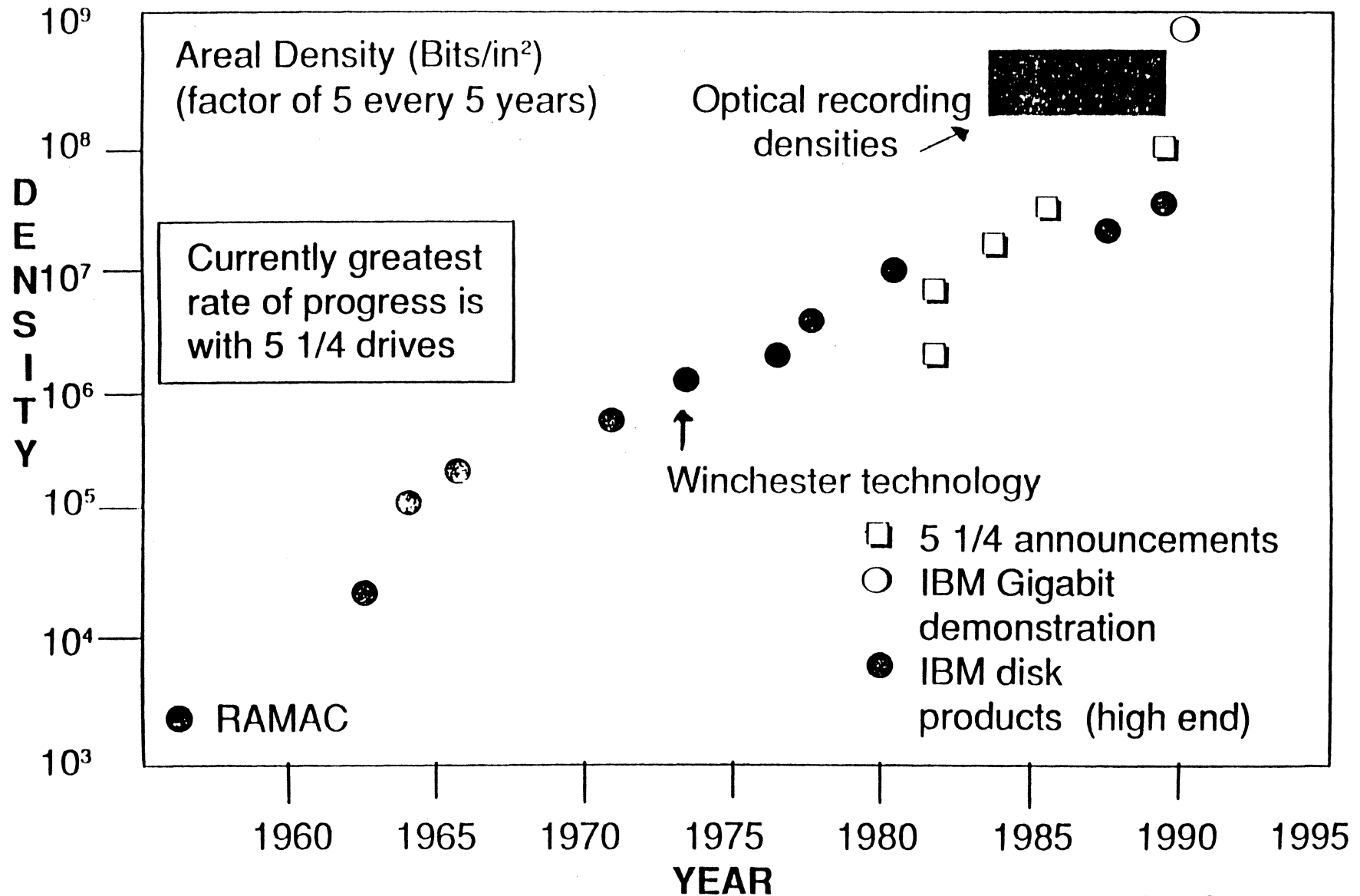
Is History Passing It By?

- Need high readback resolution to leverage very narrow transition
- Thus potential depends on very small separation

Best opportunity is then flexible disk
(ultimately could involve pole type write head)

- Major alternative is magneto-optical
i.e., high density and removability with no contact
- Vertical recording now seen as only evolutionary extension
-- are still far from limits on conventional magnetic recording

Magnetic Disk Storage Density



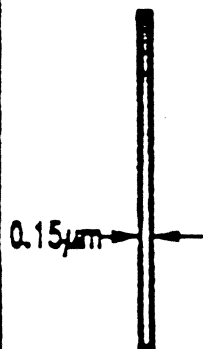
Gigabit Demonstration

BH Cell Size

3380K

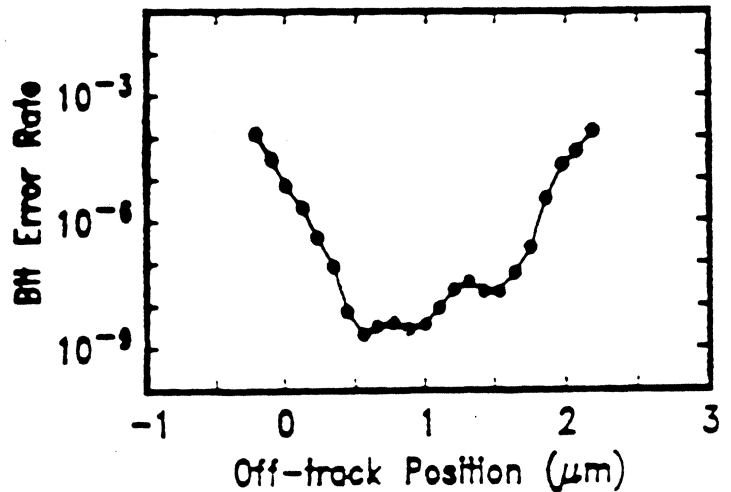


Gigabit Demo



- Scaled MR Head
- Low Noise Film Disk
- Laboratory Demonstration

Error Rate



36 Mbit/in²
 17.2 Kbpi
 2098 tpi
 0.26 μm
 0.6 μm

1180 Mbit/in²
 158 Kbpi
 7470 tpi
 0.04 μm flying height
 0.2 μm head gap

LIST WORKSHOP V PRODUCT PROJECTIONS
 First customer shipments: 4th Quarter 1991 or sooner

PARAMETERS	III EMD 5.25" DISK	III EMD 3.5" DISK	SINGLE 3.5" DISK	SINGLE SIB-3.5"	OPTICAL 5.25" FRAS.	OPTICAL 3.5" FRAS.	III CAPAC. FLOPPY	III PERF FLOPPY	CART TAPE 1/4" LOW	CART TAPE 1/4" MID	CART TAPE 1/2"	HELICAL SCAN-OMM	HELICAL SCAN-DAT	SEMI- CONDUCTOR
Areal Density Mb/in ²	160	160	160	>=75	500	500	40-45	70-100	6.2	12.0	1.4	75	114	30-000
BPI x 000	65	65	65	50	31.8	31.8	N/S	40	52	80	38	57	61	5.5 - 28
TPI	2500	2500	2500	>=1500	18.2	18.2	N/S	1.6-2.4	.12	.16	.036	1640	1870	5.5 - 28
On-Line Capacity-MB:F	2400	1200	160	40+	750	200	40	100-160	1350	2000	480	5000	1300	4 - 12
Average Seck Time-ms	10	12	15-17	20	25	22.5	50	15-20	37500	43000	N/S	40000	20000	.1 - .2
Average Latency-ms	5.6	5.6	8.3	9 +/- 2	12.5	12.5	30	8.3-10	--	--	--	--	--	--
Transfer Rate MB/s	6.0	3.0	2.5-3.0	2.0	1.28	0.83	0.4	1.25-2.0	1.5	3.0	6.0	0.5	0.183	4 - 32
Capacity/ Volume-MB/in ³	15	30	9	4	52	98	24	40	100	280	46	1400	540	6.7 - 20
MIBF - Khrs	250	250	100	100	20	20	N/S	60	50	50	N/S	30	10	Very high
OEM Unit Drive Cost-\$	3000	1500	300	200	1600	600	125	300-500	550	650	20000	2300	800	256 - 800**
OEM Drive Cost/MI--\$/MI	1.35	1.25	2.00	4 - 5	2.13	3.00	6.25	3 - 4	0.41	0.24	42	0.46	0.62	64 - 200
Media Unit Cost-\$ (E.U.)	--	--	--	--	90	25	10	40 - 60	25	35	7.0	8.0	8.0	--
Media Cost \$ per MI (E.U.)	--	--	--	--	* 0.06	* 0.062	0.25	0.35-0.4	0.010	0.018	0.015	0.0016	0.0061	--
Power (Watts)			< 1.5	< 1.0										

Comments FS form factor. 8 disks. TF media & heads. 1.625" form factor. 8 disks. TF media & heads. 1.625" form factor. 2.5" single disk. * For 2 sided media Capacity is for 1 side. 3.5" 3.5" Bernoulli, 2 disks. QIC 1350 3480 DRAM configured as a disk drive, Non-removable. **Sub-ass'y

Notes: Areal density in Mbits/square inch.
 On-line capacity: U- unformatted, F- formatted
 N/S - Not specified, -- = Not appropriate
 Media pricing at end user levels

ULTIMATE LIMITS

- Magnetic Recording
 - Minimum Domain Size
 - Medium Signal to Noise Ratio

- Spacing Dominant Parameter

- Optics
 - Limits due to Wavelength of Light

TECHNOLOGY LIMITS

MANUFACTURING PROCESSES KEY

- Medium quality
- Head medium interface

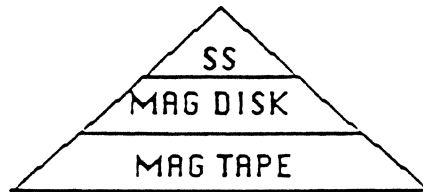
PRODUCT SIGNIFICANCE

- Capacity demand
40% CGR
- Technological progress
26% CGR

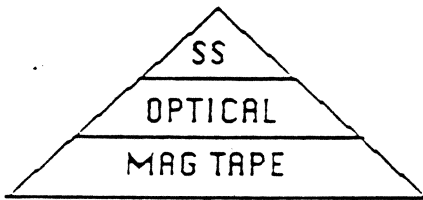
SCENARIOS...

VISIONS OF THE FUTURE

TODAY



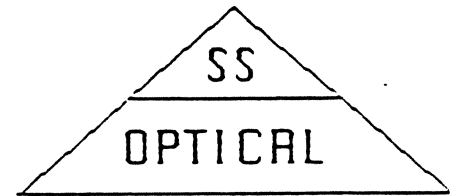
THE DISASTER



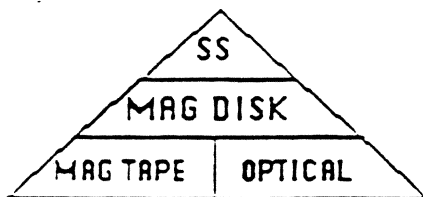
THE HOPE



THE INCONCEIVABLE



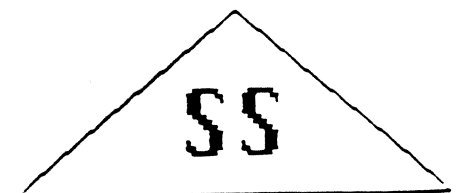
THE REALITY



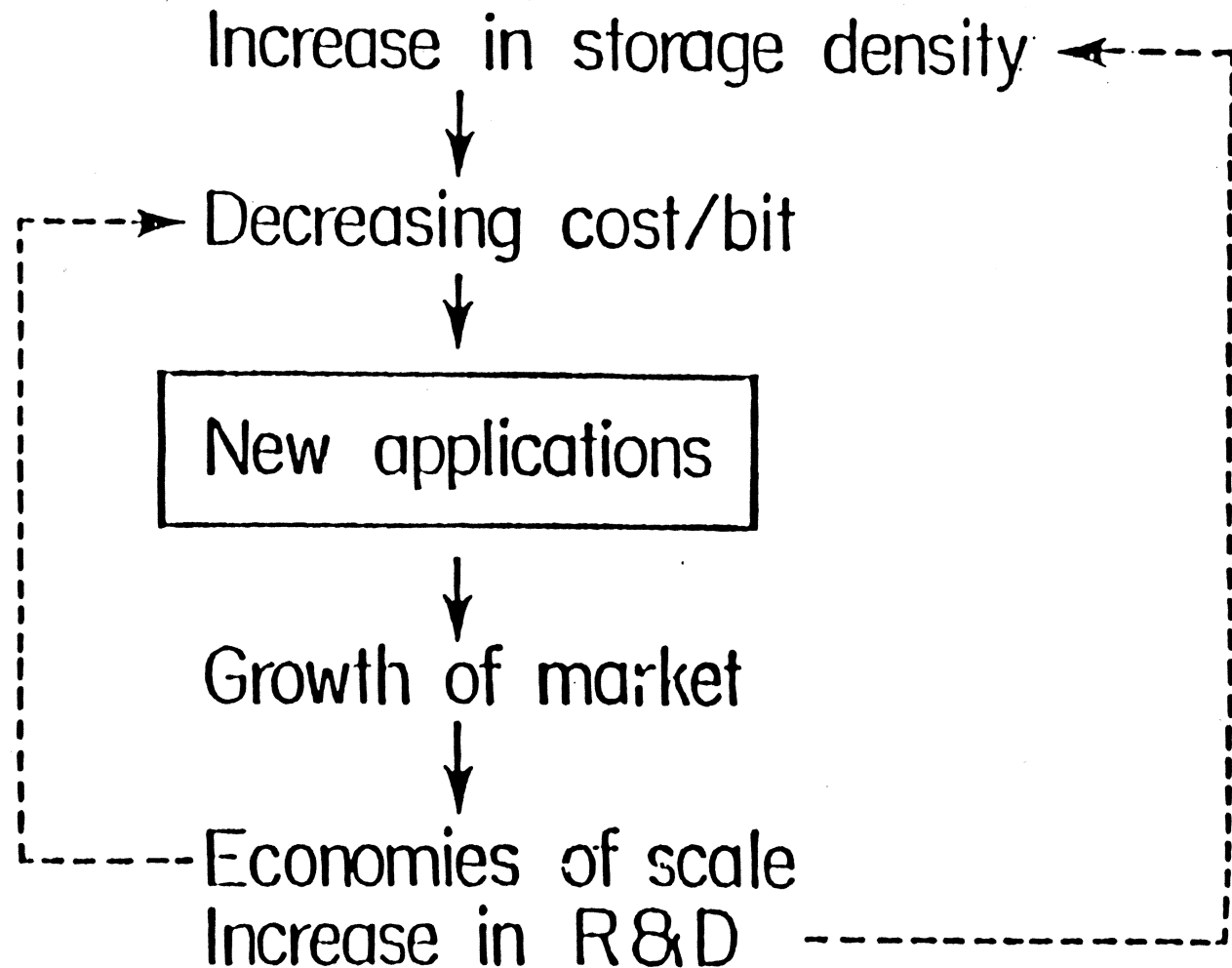
THE FEAR



THE DREAM



Mass storage



Alternative Storage Technologies

GEOFFREY BATE

INSTITUTE for INFORMATION STORAGE TECHNOLOGY
Santa Clara University

CEI, September 1991

High-Performance Data Storage Systems

- 50% of the total system cost is in the storage system
- Bank of America, San Francisco, (4 large computers, IBM 3090)

- 1 GB of main storage
- 1 GB of extended storage
- 416 MB of solid-state disk
- 96 MB of disk cache (within controller)
- 750 GB of magnetic disk
- 2000 x 200 MB tape cartridges/day

-
- 24 hour operation: 12 hour on-line; 12 hour batch
 - 20 - 30% growth in capacity
 - cost/MB and MB/cu.ft. are key metrics

R. Katz

Problems with Optical Recording

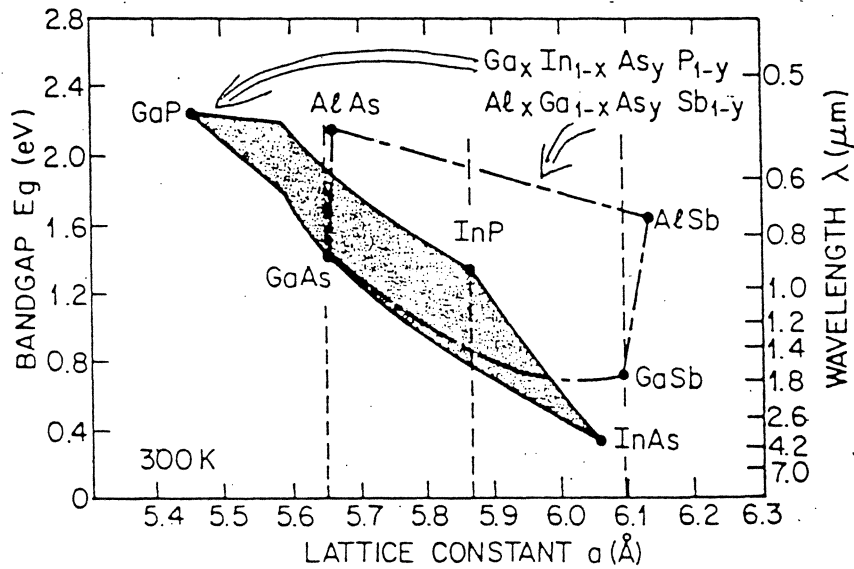
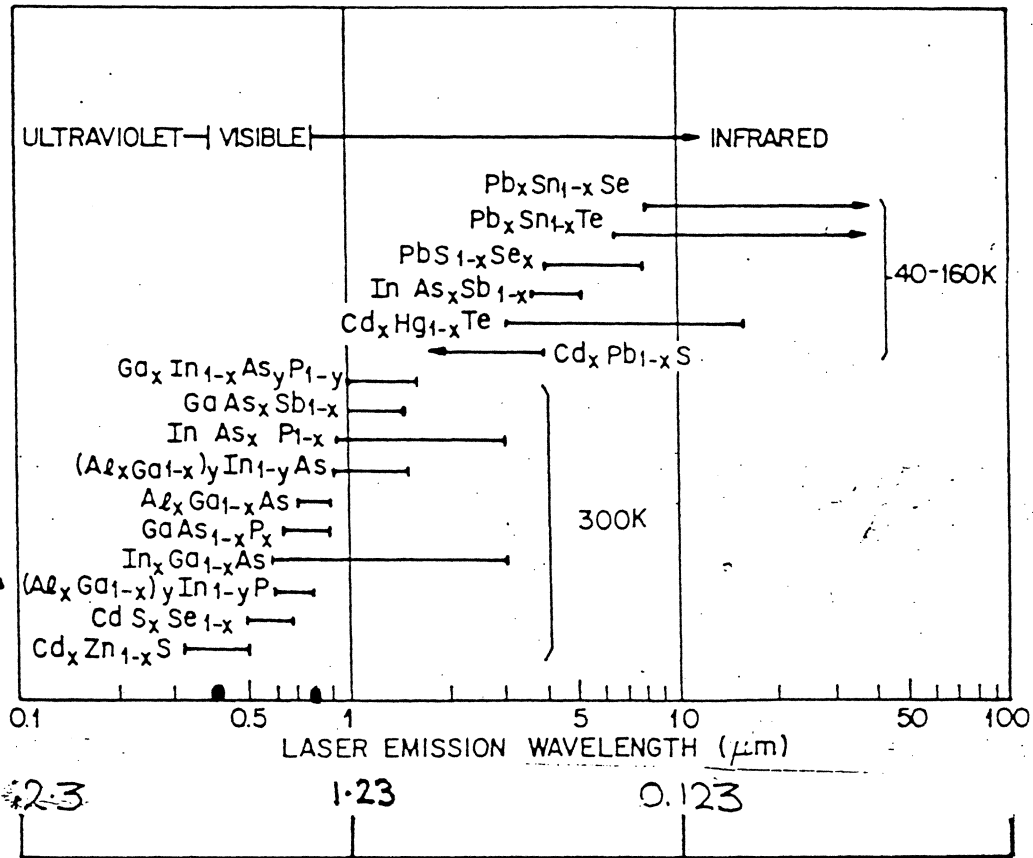
Present

- long-term corrosion resistance, archival storage
- overwrite data rate
- access time
- standards
- writing by thermal effects rather than photon effects

Future

- high-power, 50 - 100 mw long-life, cheap laser diodes
- laser diodes emitting in blue (only gain factor of 2)
- multi-layer films

Semiconductor Lasers



Energy bandgap and lattice constant for two III-V solid solutions.

TMO Drives : Trends

Access Time: split hds. with galvo. mirror < 25 ms
head / surface
caching controller (some applications) < 10 ms

Data Rate: ≥ 10 Mb/s

- higher rpm $\geq 3,600$ rpm
- multichannel; laser array heads
- one head / disk surface
- pulse width modulation

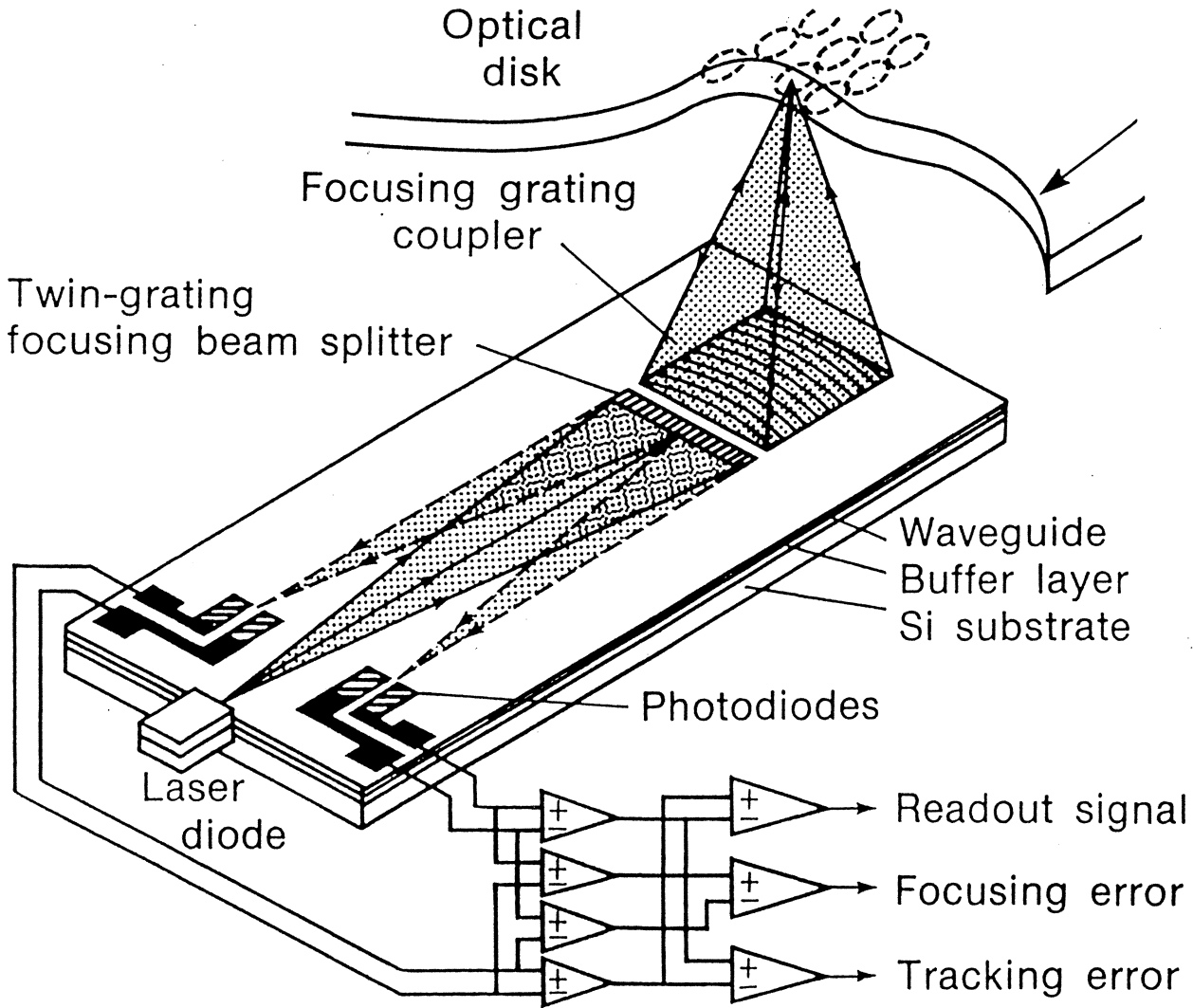
Direct Overwrite

Higher Capacity

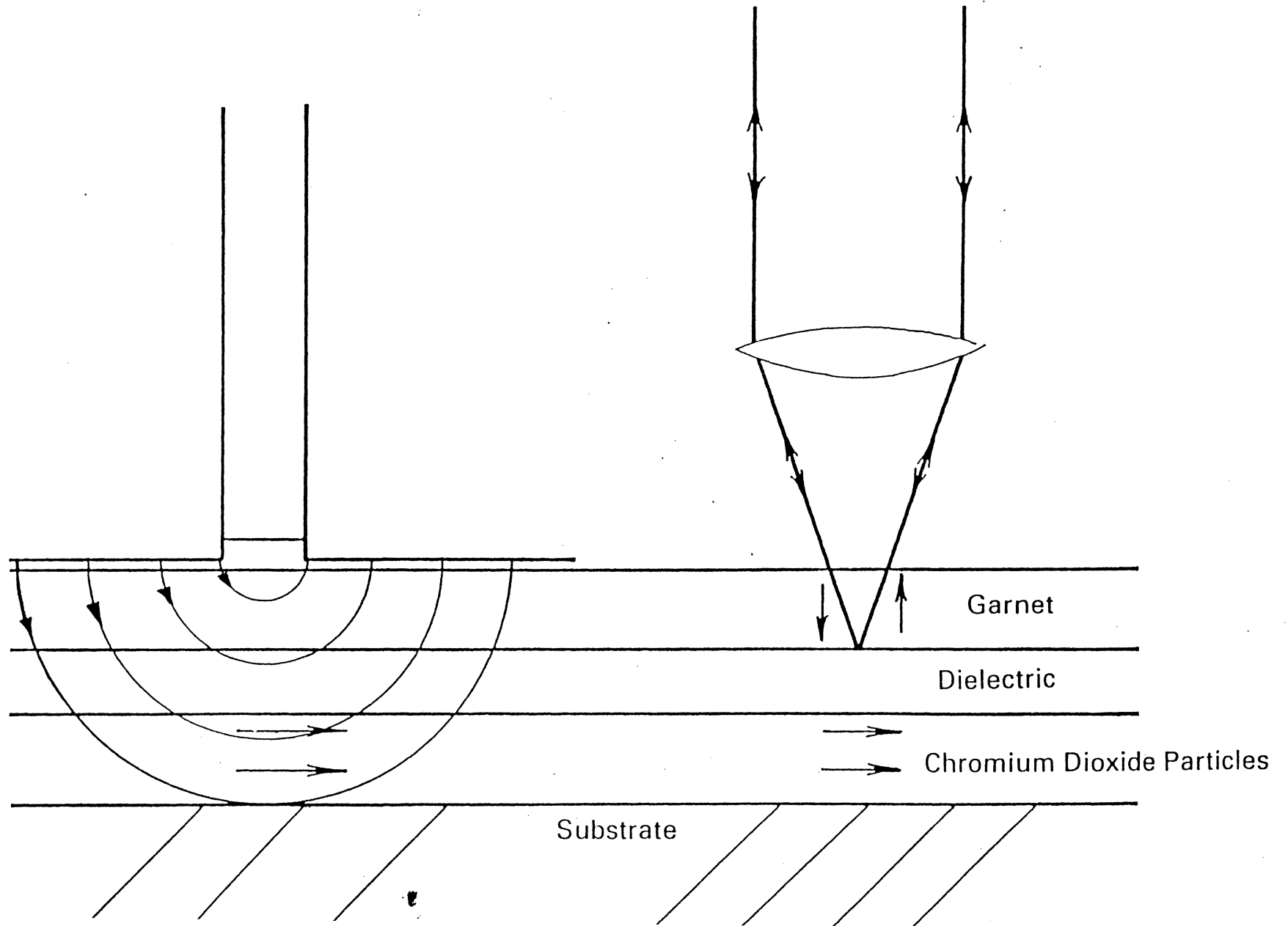
New Form Factors : 1/2 - height, 5.25"

New Substrate Materials for $> 3,600$ rpm

Monolithic Integrated Optical Head



MAGNETIC RECORDING MAGNETO - OPTIC READING



MAGNETIC RECORDING RESURGENT

- * 1987, Iwasaki et al, 680,000 bpi on CoCr
- * 1987, Lambert et al, 0.5 μm wide, discrete tracks
- * 1987, R-DAT; digital recording of audio signals, 58,000 bpi x 2000 tpi, helically on 8 mm tape (1.3GB)
- * 1987, Exabyte; digital recording of data
2.3 GBytes on standard 8 mm video tape cartridge
- * 1987, StorageTek; 4400 Automated Storage System
using IBM 3480, 1/2" data tape cartridges
19.2 Terabytes on 96,000 cartridges
- * 1988, Insite Peripherals; flexible disk with optical tracking
1,250 tpi, 24,145 bpi, 21.6 MB on 3.5" diameter disk
- * 1989, Opticard, 50 MB, arced tracks
- * 1989, Domain Tech., PrairieTek, 2.5" plated disks
- * 1989, Superconducting Super Collider, 1 GB/sec.
- * 1990, IBM laboratory demonstration of 1.2×10^9 bits/in²

IBM Gigabit Demonstration

158,000 bpi }
7,400 tpi } 1.180 x 10⁹ bits / in²

Flying height 0.04 μm (1.6 microinch)

Head gap 0.20 μm (8 microinches)

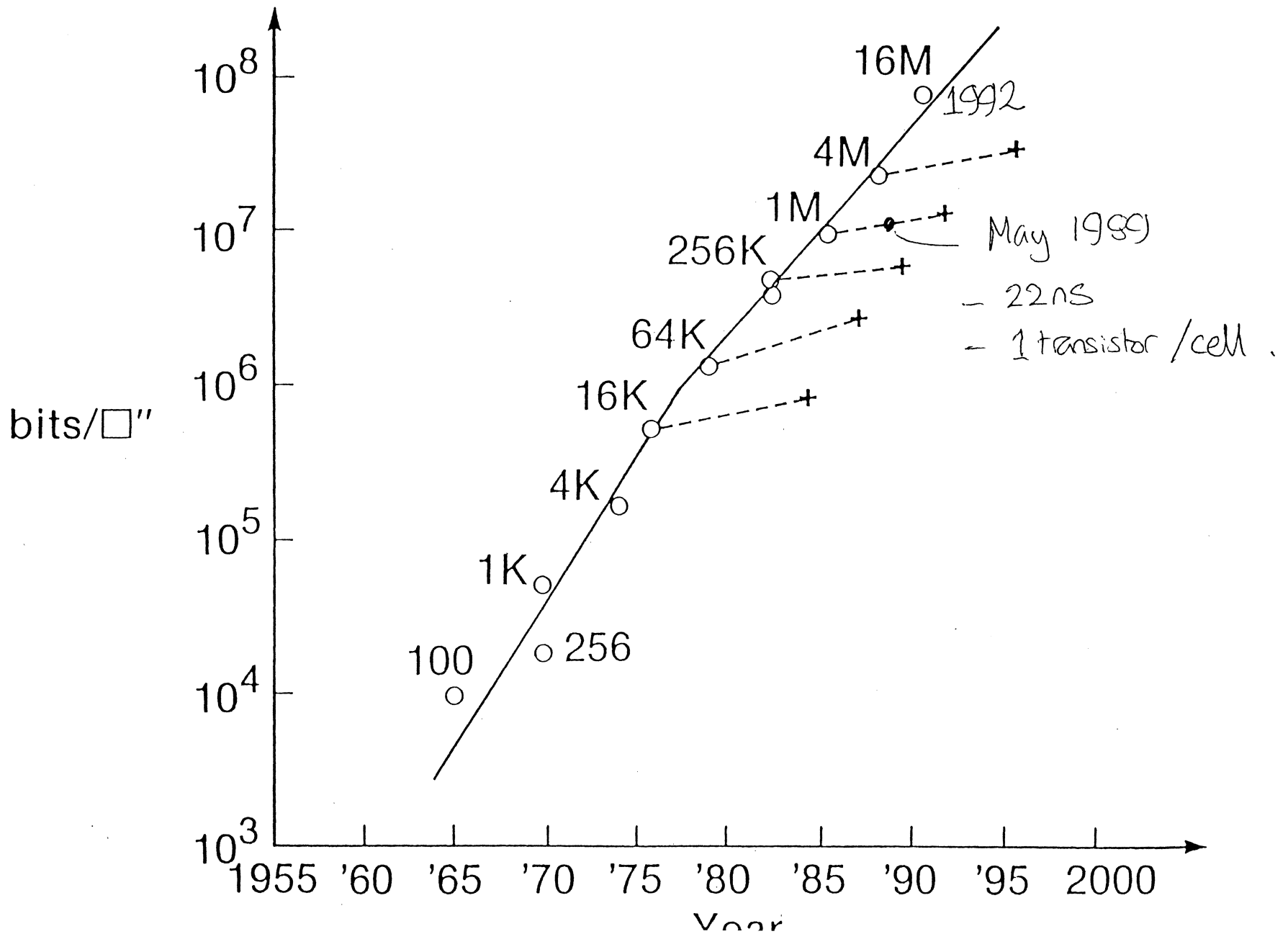
Medium -- sputtered thin film
-- 5 layers

Head -- magneto-resistive read
-- "giant M/R" material ?

Alternative Storage Technologies

- Semiconductor RAM
- Ferroelectric RAM
- Magnetoresistive RAM
- Vertical Bloch Line Memory
- Holographic Memory
- Spectral Hole-Burning Memory
- Ultra-violet Write -- Infra-red Read
- Two-photon Recording In 3-dimensions
- DNA

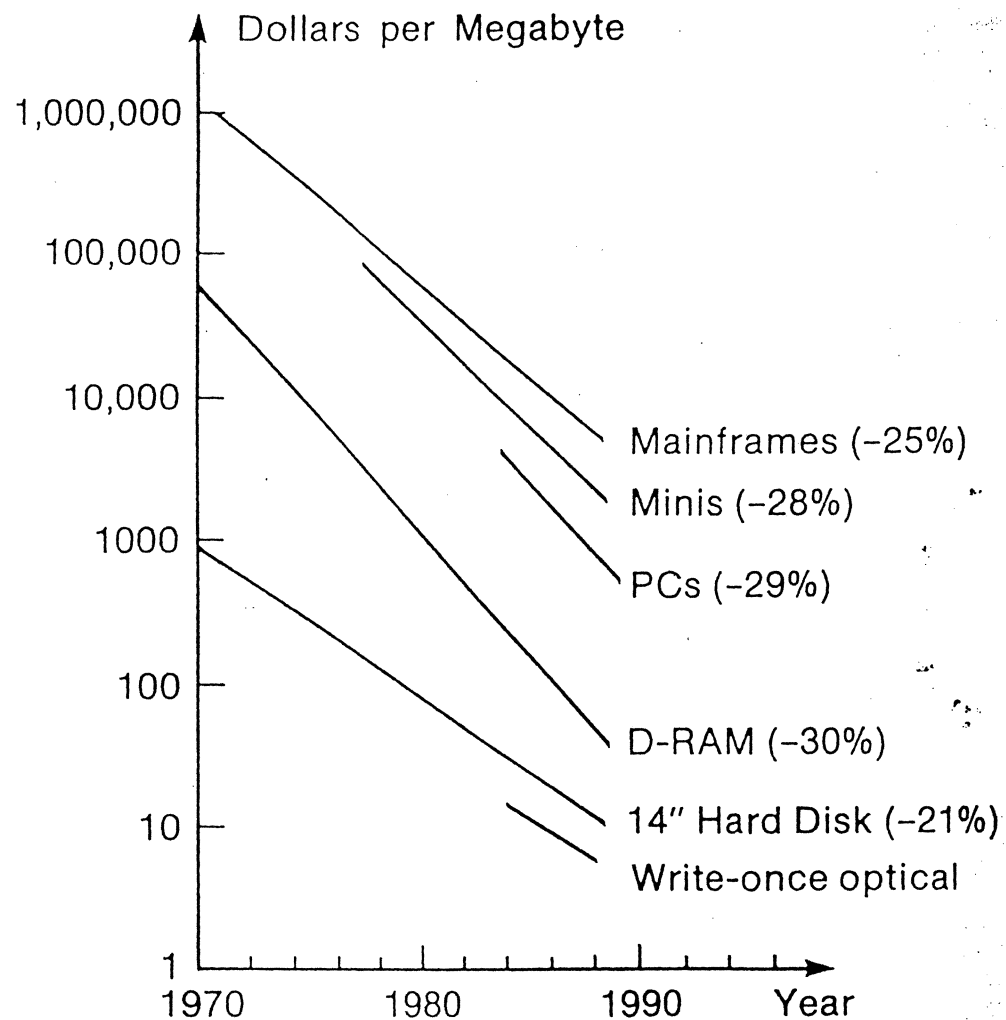
D-RAM: Bits Per Square Inch versus Time



D-RAMS versus Disks

D-RAM cost reduction factors:

- higher resolution processes
- lower defect densities
- larger wafers
- fewer components/bit
- multilayer processing



SOLID STATE MEMORY — D-RAM

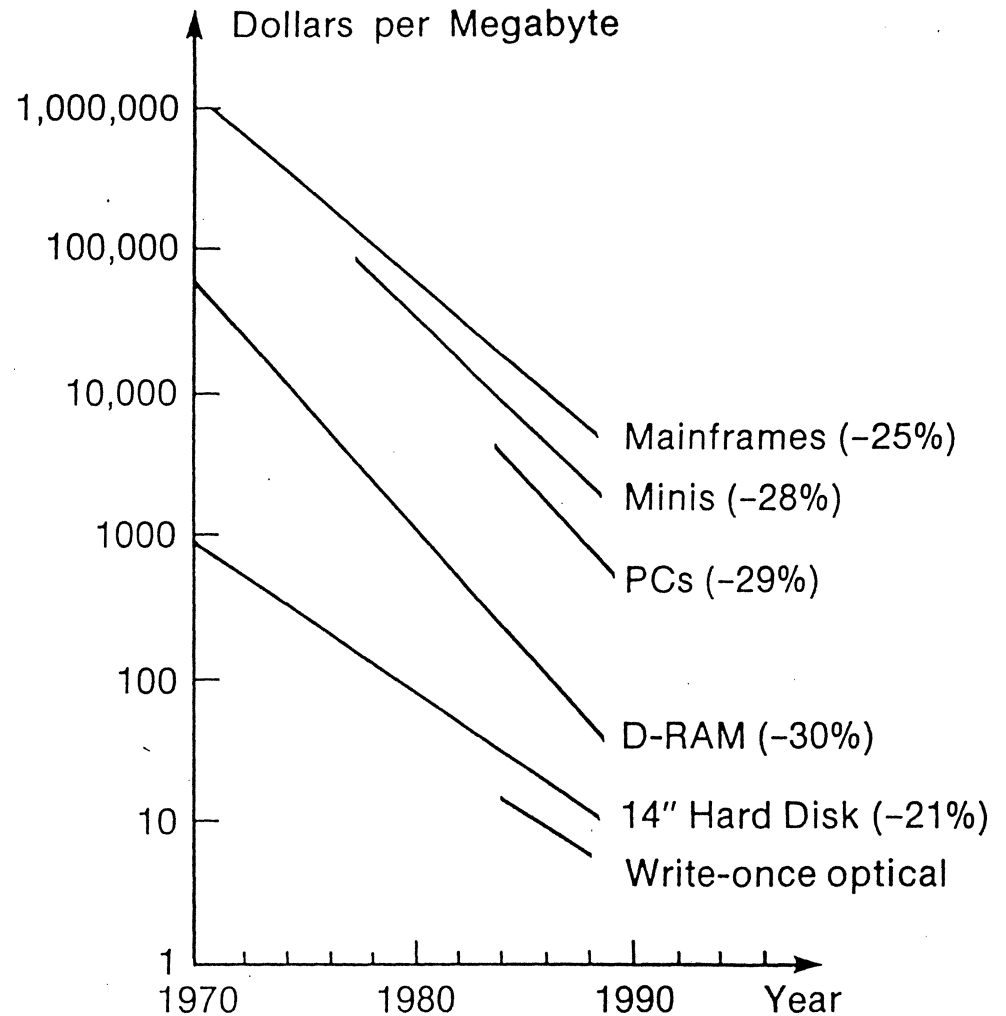
Year of Introduction	Capacity (bits)	Minimum Line-width
1976	16 K	5 - 7 μm
1981	64 K	3.0
1983	256 K	2.0
1985	1 M	1.3
1987	4 M	0.7 - 0.8
1990	16 M	0.4 - 0.5
1993	64 M	0.2 - 0.3
1996	256 M	~ 0.1

Wicat

D-RAMS versus Disks

D-RAM cost reduction factors:

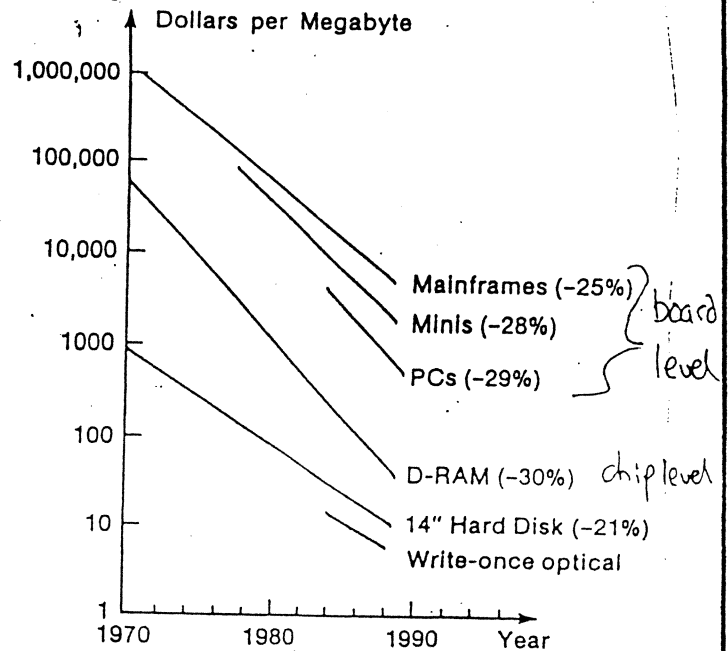
- higher resolution processes
- lower defect densities
- larger wafers
- fewer components/bit
- multilayer processing



D-RAMS versus Disks

D-RAM cost reduction factors:

- higher resolution processes
- lower defect densities
- larger wafers
- fewer components/bit
- multilayer processing



Ura, H-P

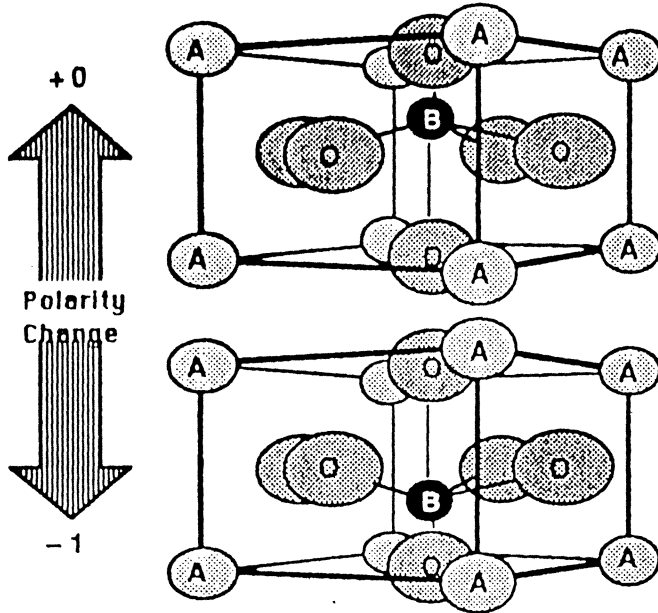
SOLID STATE MEMORY — D-RAM

Year of Introduction	Capacity (bits)	Minimum Line-width
1976	16 K	5 - 7 μm
1981	64 K	3.0
1983	256 K	2.0
1985	1 M	1.3
1987	4 M	0.7 - 0.8
1990	16 M	0.4 - 0.5
1993	64 M	0.2 - 0.3
1996	256 M	~ 0.1

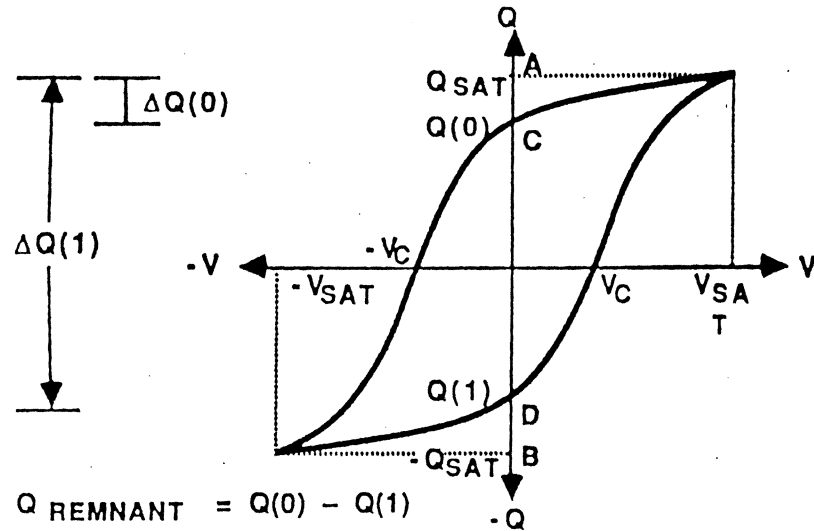
Wicat

FERROELECTRIC MEMORY TECHNOLOGY

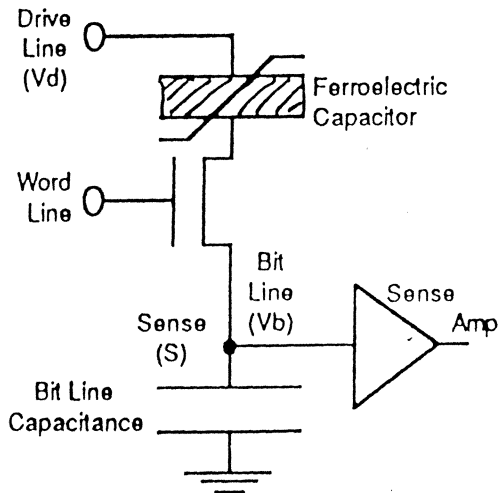
Ferroelectric Polarization



Hysteresis of a Ferroelectric Capacitor

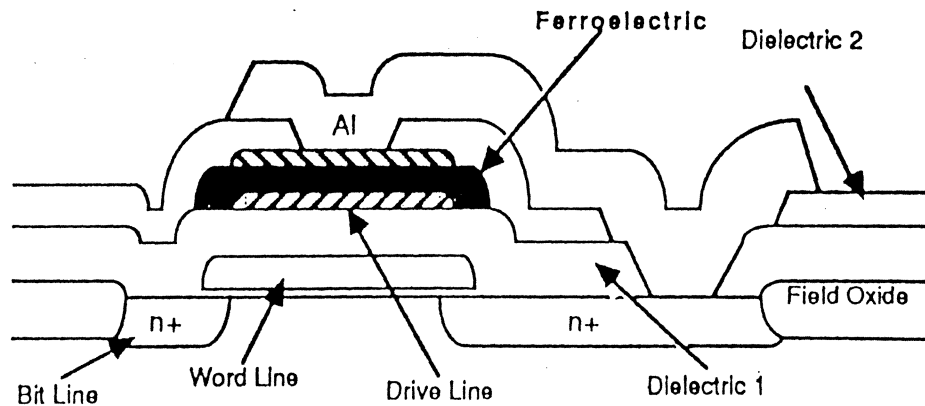


Ferroelectric Memory Schematic



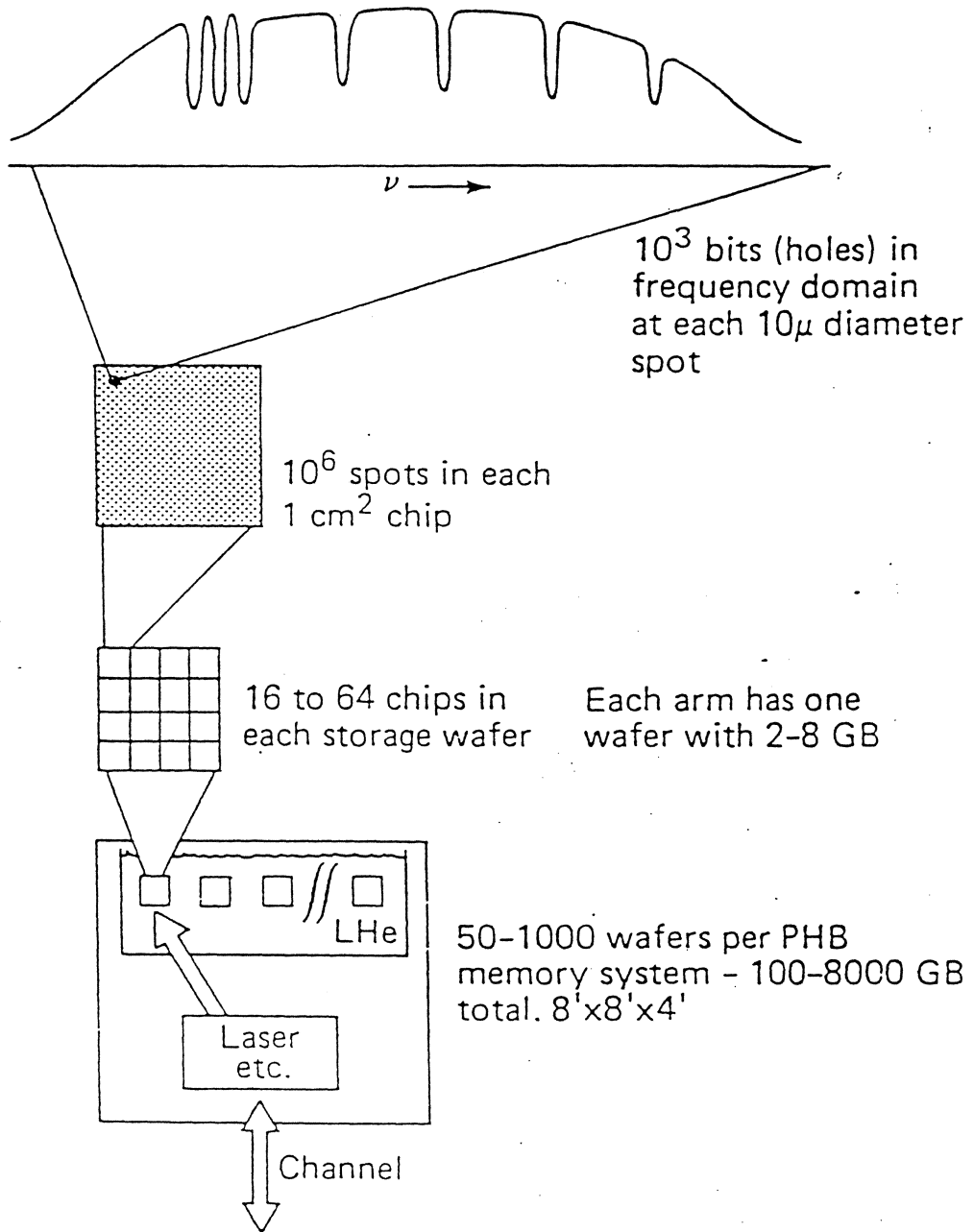
Patents Pending

Cross-Section of Ferroelectric Memory Cell



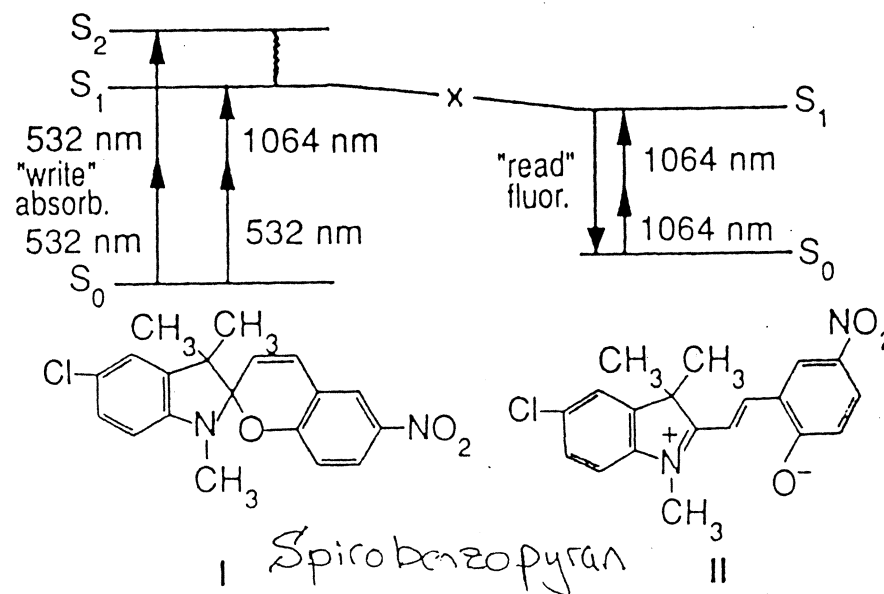
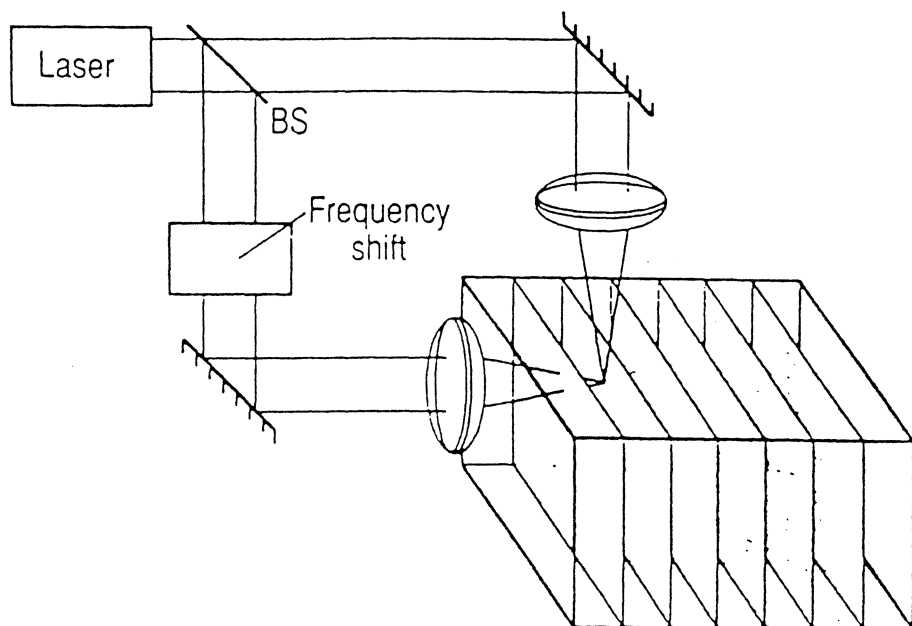
U.S. Patent No. 4,759,823
and other Patents Pending

Spectral Hole-Burning



Three-Dimensional Optical Storage Memory

DIMITRI A. PARTHENOPOULOS AND PETER M. RENTZEPIS



Schematic diagram of a 3-D optical memory based on a two-photon process. Replacing lenses with holographic gratings allows parallel addressing. BS, beam splitter.

- random access, parallel addressing
- writing, (reading) in nsecs.
- no moving parts
- no cross-talk between adjacent bits

$$\lambda = 532 \times 10^{-9} \text{ m}$$

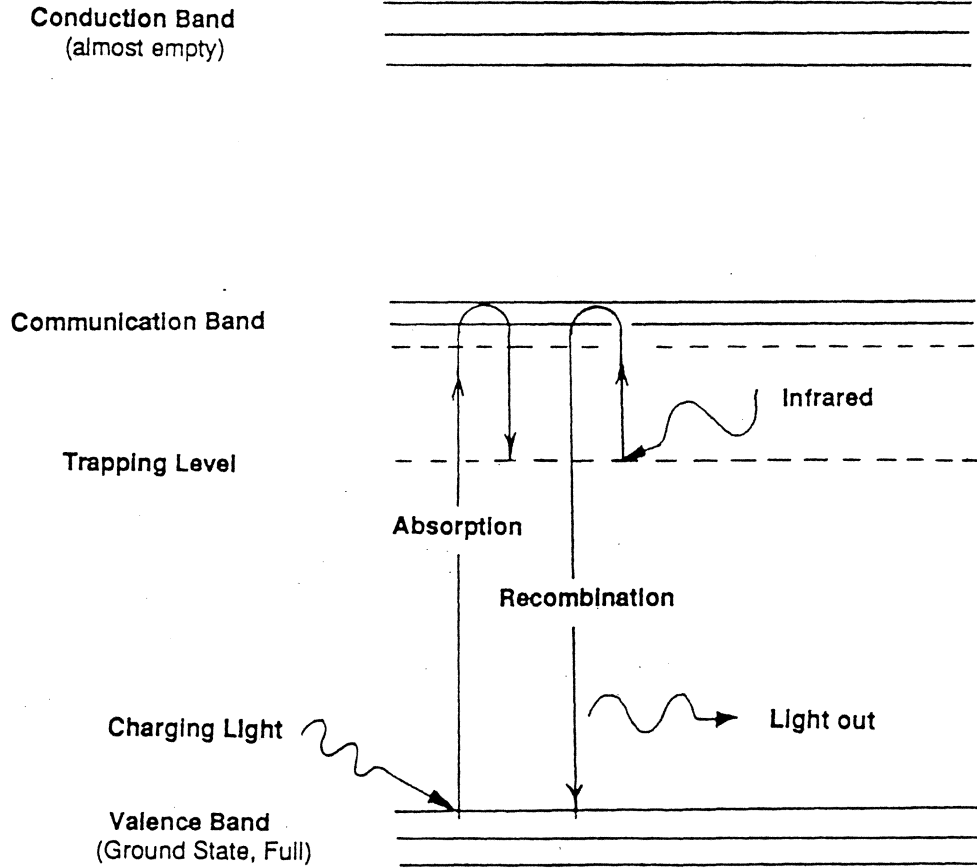
$$\Rightarrow 3.533 \times 10^8 \text{ b/cm}^2$$

$$3.533 \times 10^{12} \text{ b/m}^2$$

$$830 \text{ GB/cc}$$

A New Erasable Optical Memory

Joseph Lindmayer
Optix Corp., Rockville, Maryland

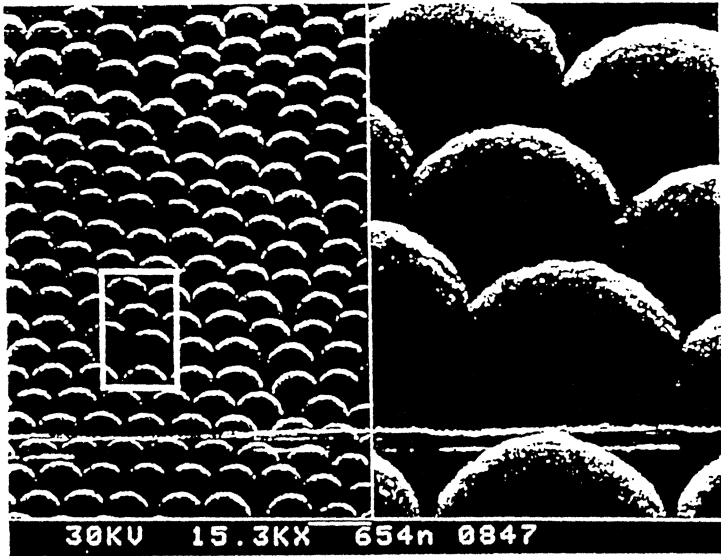
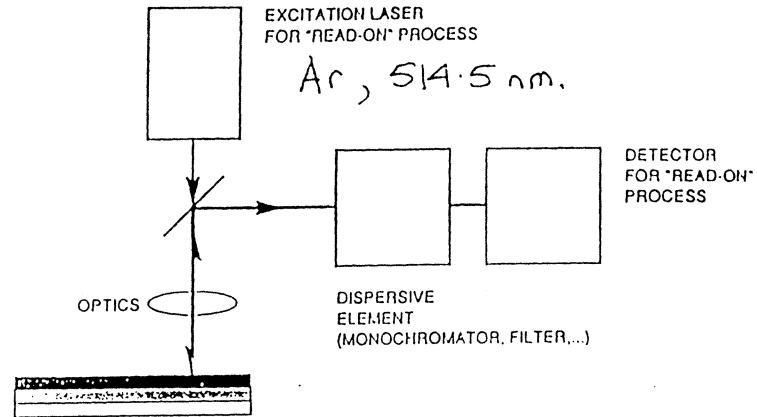
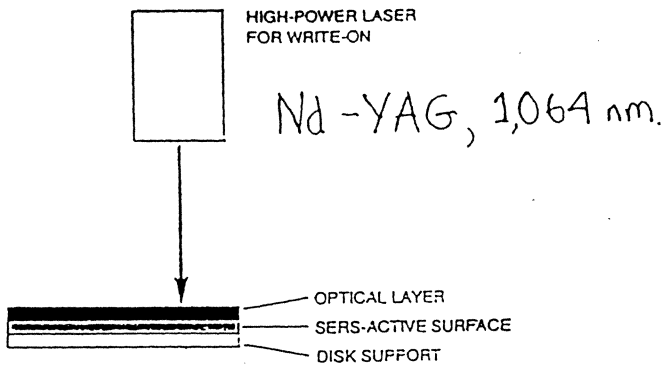


Surface-Enhanced Raman Optical Data Storage System

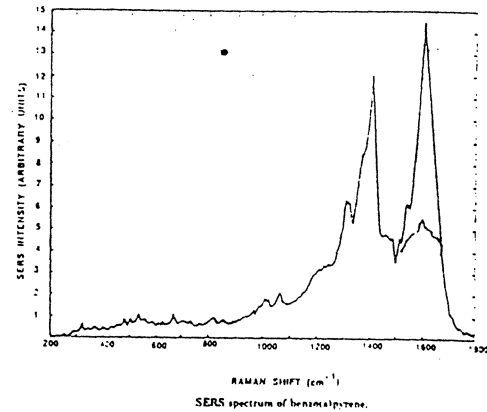
"SERODS"

Tuan Vo-Dinh

Oakridge N. L.

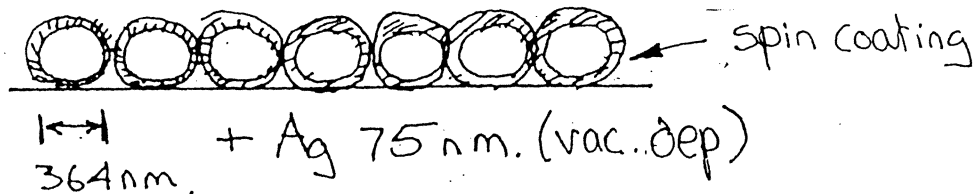


SEM photograph of silver-coated microsphere substrate.

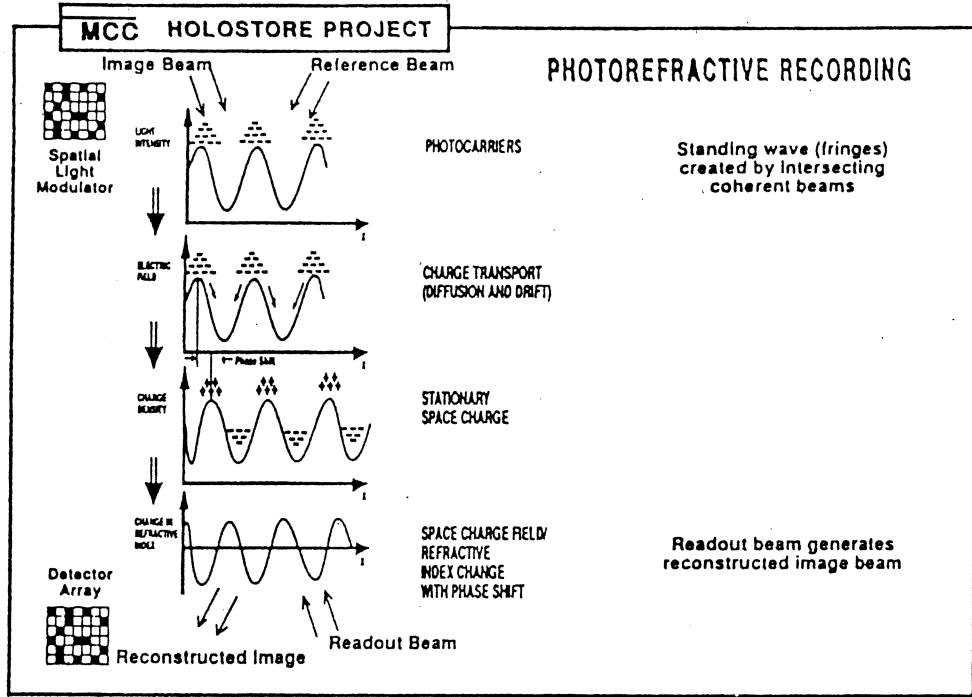


- 19,500 pre-writing

- 5,000 counts after-writing



MCC HOLOSTORE PROJECT

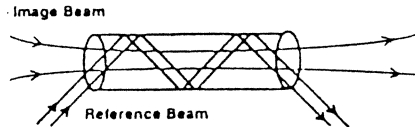
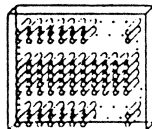


MCC HOLOSTORE PROJECT

MCC PATENTED INNOVATIONS

ARRAY OF CRYSTALLITES (FIBERS)

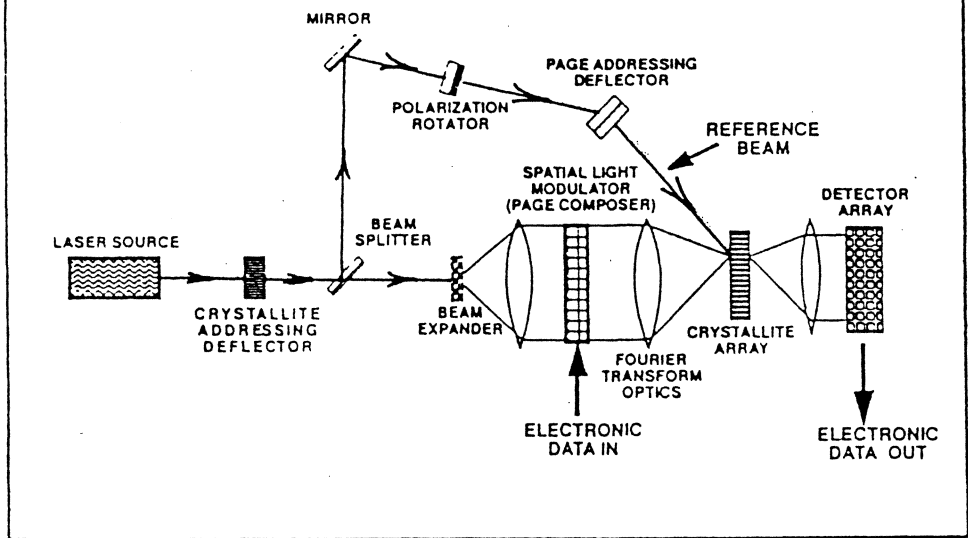
- Scalable storage capacity (larger array)
- Larger capacity (more pages per stack)
- Low crosstalk
- Lends itself to low cost production



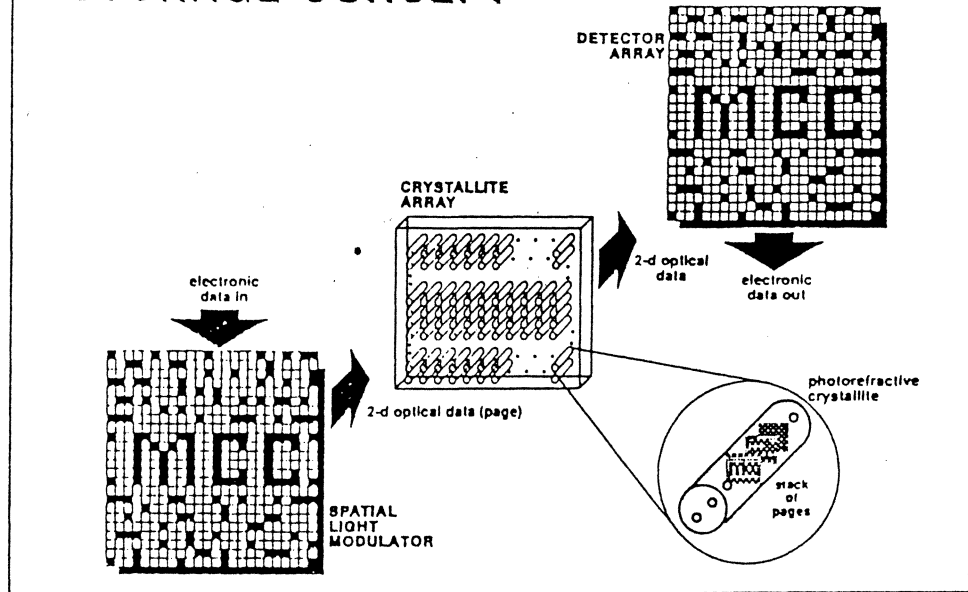
NON-DESTRUCTIVE READOUT TECHNIQUE

- Allows prolonged readout in photorefractive material
- Billions of reads without signal / noise degradation
- May result in archival storage

OPTICAL CONFIGURATION

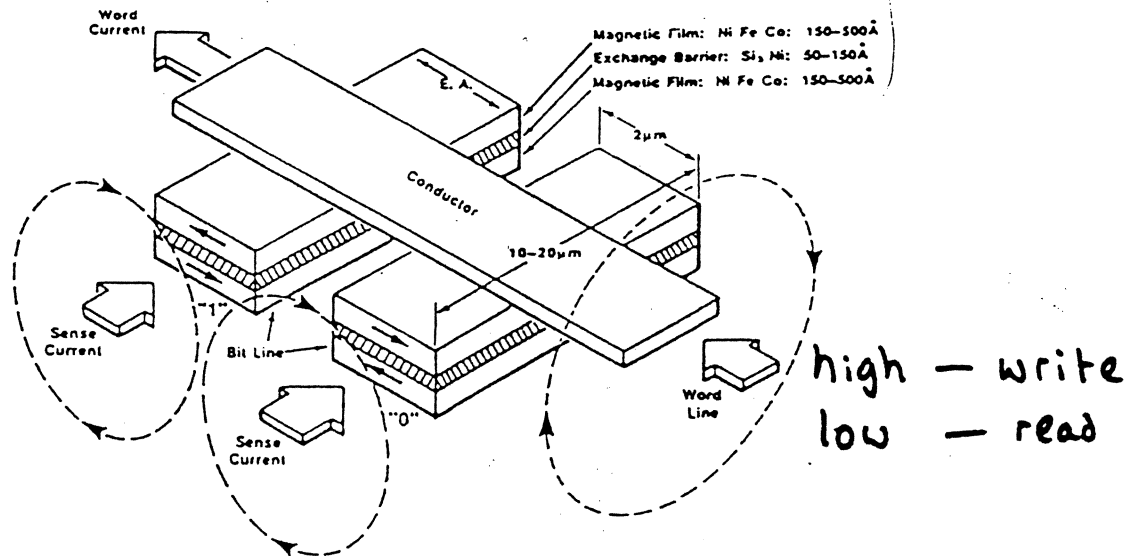
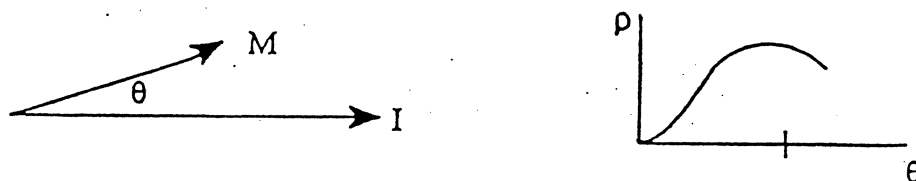


STORAGE CONCEPT



Magnetoresistive RAM

Anisotropic magnetoresistance:



- high density
1 Mb @ 1.5 μm x 5 μm cell demonstrated
- high speed
- radiation hard.
- IC compatible

Ni Fe-Co, non-magnetostrictive, $H_K \sim 10-30 \text{ Oe}$.

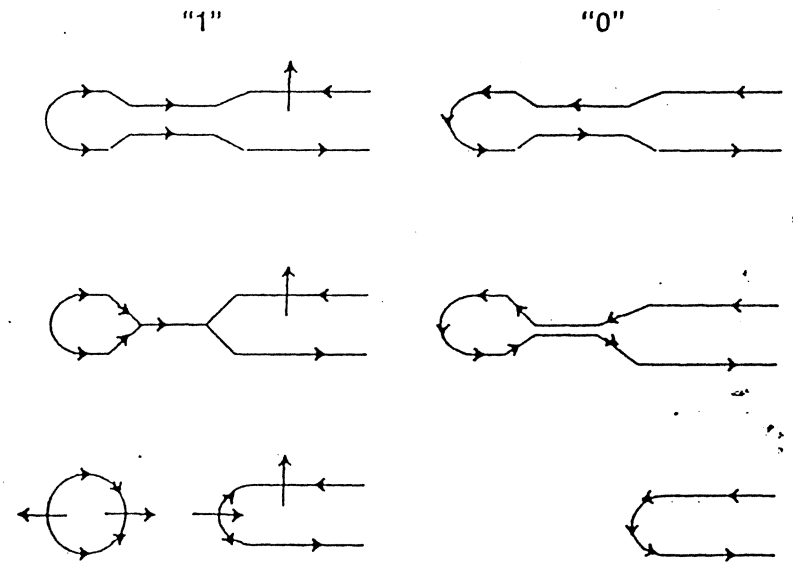
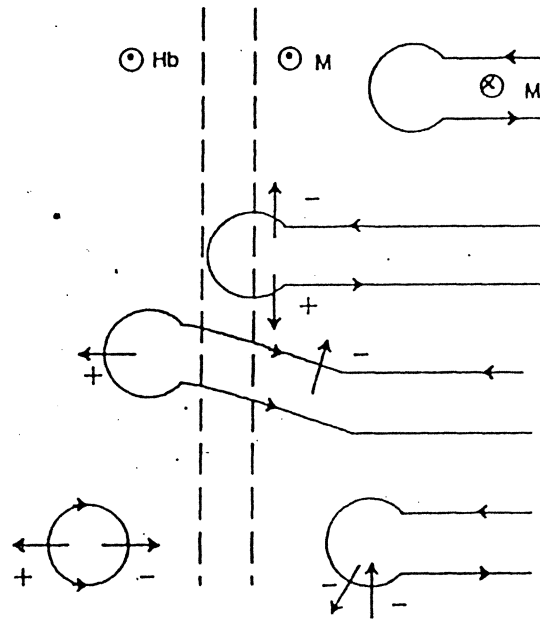
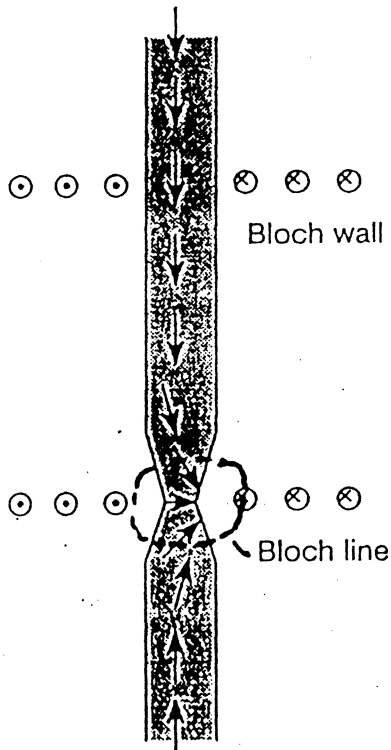
$$\frac{\Delta\rho}{\rho} \sim 3\% \quad (\text{thickness} > 300 \text{ \AA})$$

$$\sim 2\% \quad (\text{ " } \sim 150 \text{ \AA})$$

Vertical Bloch Line Memory

Write ONE

READ



bubble

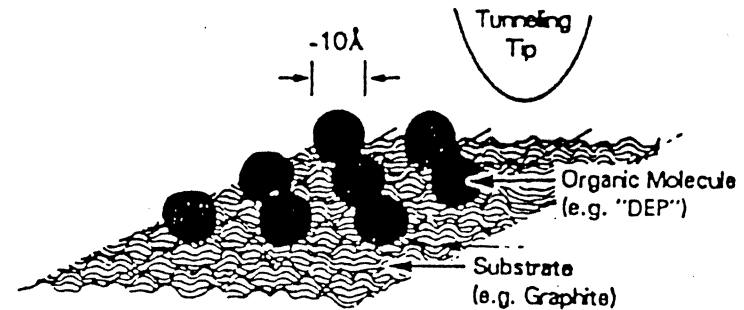
no bubble

- 1,600 Mbits/in²
- Current access

STM for Storage?

Starting Assumptions

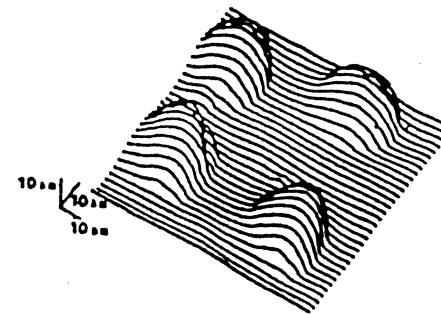
- STM can read on atomic level
- Potential to write on atomic scale
 - 25Å - 250Å bit size
 - 10^{12} - 10^{14} bits/sq.in.



- Possibility of 10^{12} - 10^{14} bits/sq. in.
- Demonstrated Read/Write/Erase Feasibility
 - Voltage Pulse to Write and Erase
 - Tunneling Current to Read

Storage Requirements

	<u>STM</u>	<u>Current DASD</u>
• Areal Density-Mb/sq.in.	10^6 - 10^8	62
• Data Rate		
Writing speed	100ns	30ns
Reading speed	?	30ns
• Access Rate	?	12ms
• Signal-to-Noise	?	30dB
• Reliability		
lifetime	?	7 years
error rate	?	10^{-12}
• Margins		
Disk Roughness	?	100Å
Head Dimensions	?	10%



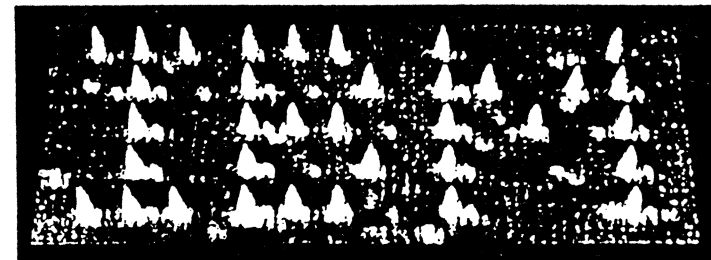
Induce topographical features
on metallic glass

($Rh_{25}Zr_{75}$)

350Å features induced (melting?)

$\tau \sim 1s$

16 nm



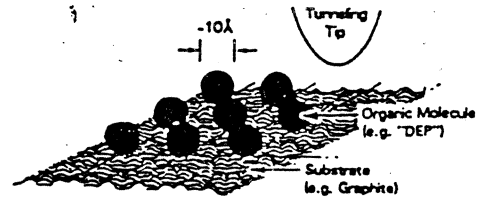
XENON ATOMS spell out the name of the researchers' corporate benefactor.

on Ni at 4K

STM for Storage?

Starting Assumptions

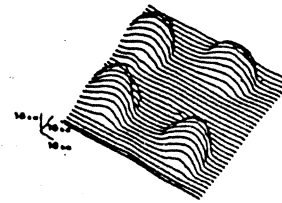
- STM can read on atomic level
- Potential to write on atomic scale
 - 25Å - 250Å bit size
 - 10^{12} - 10^{14} bits/sq.in.



- Possibility of 10^{12} - 10^{14} bits/sq. in.
- Demonstrated Read/Write/Erase Feasibility
 - Voltage Pulse to Write and Erase
 - Tunneling Current to Read

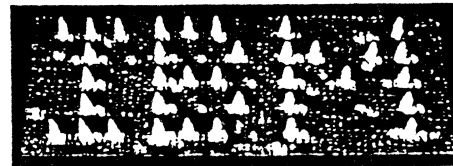
Storage Requirements

	STM	Current DASD
• Areal Density-Mb/sq.in.	10^6 - 10^8	62
• Data Rate		
Writing speed	100ns	30ns
Reading speed	?	30ns
• Access Rate	?	12ms
• Signal-to-Noise	?	30dB
• Reliability		
lifetime	?	7 years
error rate	?	10^{-12}
• Margins		
Disk Roughness	?	100Å
Head Dimensions	?	10%



Induce topographical features on metallic glass ($Rh_{25}Zr_{75}$)

350Å features induced (melting?)
 $\tau \sim 1s$



XENON ATOMS spell out the name of the researchers' corporate benefactor.

Types of Limits

<u>Limit</u>	<u>Concerned With</u>	<u>Determined By</u>
FUNDAMENTAL	What is possible	Laws of Nature
TECHNOLOGICAL	What is practical	Man's Ingenuity
ECONOMIC	What is profitable	Costs and Markets
SUBJECTIVE	What is acceptable	Personal Preferences