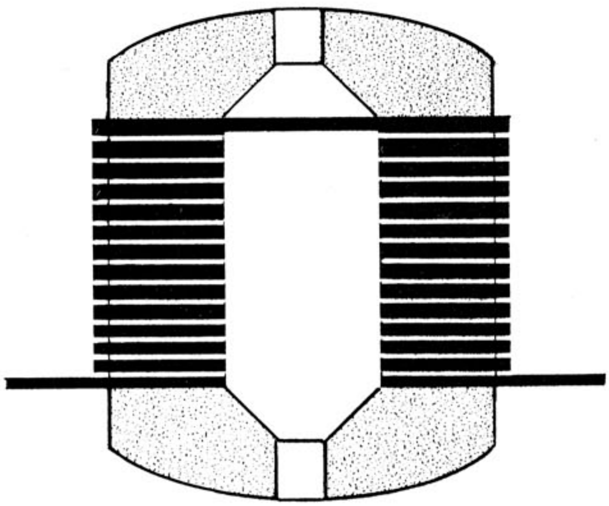
MAGNETIC RECORDING THEORY FOR INSTRUMENTATION



Ampex Corporation

TRAINING DEPARTMENT

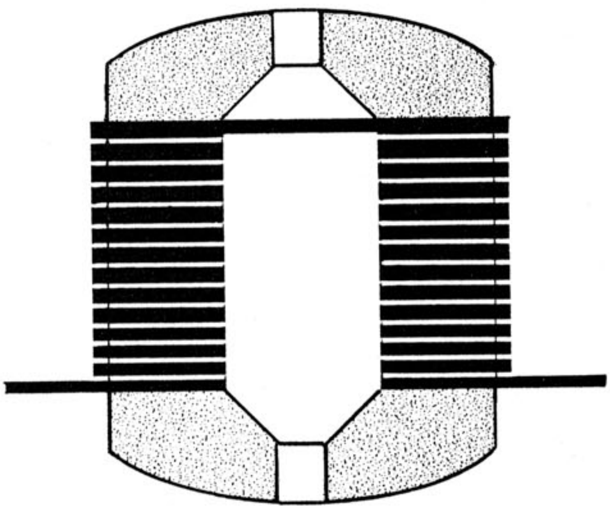
MAGNETIC RECORDING THEORY FOR INSTRUMENTATION

by

Charles E. Lowman

and

George J. Angerbauer



AMPEX TRAINING DEPARTMENT

PREFACE

This book has been specifically prepared to be used as the basis of a course on instrumentation recording theory. This course is normally given in conjunction with the training programs on Ampex equipment.

An attempt has been made to assemble under one cover all the basic information necessary to understand the fundamental principles of magnetic tape recording, and recording techniques.

It is assumed that the reader has a technical background and an acquaintance with electronic terminology, but has little or no experience with magnetic recording and equipment.

Grateful acknowledgement is extended to the members of the Training Department who, in one way or another, have assisted in the preparation of this book. These include the Department Manager, Mark Sheldon; Instructors Robert Byers, Carl Meyer, Don Smith; and the secretaries whose untiring efforts in typing this manuscript have made our task much easier.

Charles E. Lowman
Geo. J. Angerbauer

CONTENTS

Contents

Contents

INDEX OF ILLUSTRATIONS

Index of Illustrations

Index of Illustrations

CHAPTER I

INTRODUCTION

WHY MAGNETIC TAPE RECORDERS?

Nearly everyone is familiar with the widespread acceptance and usage of magnetic tape recorders for audio purposes--the recording of speech and music for entertainment, training and broadcasting. Even more dramatic in its recent growth and development, although less widely understood and appreciated, have been the advances made in magnetic tape recording for instrumentation purposes. This phase of recording concerns itself with measurement, data storage, data analysis, research and industrial control.

The most exciting thing about the instrumentation recording field is its continuing growth at an ever-expanding rate. There are two principal reasons for the rapid strides being made in applying magnetic recording for instrumentation purposes:

1. The need and demand for making very large quantities of measurements at very fast rates; and for reducing this data rapidly to a form which will allow it to be used efficiently and effectively.

2. The inherent advantages of the magnetic recording process itself. In some cases, it permits the attainment of results which cannot be achieved in any other way.

THE NEED FOR MORE RAPID, EFFICIENT AND EFFECTIVE DATA ACQUISITION AND REDUCTION:

To illustrate the first point, consider the problems involved in designing a modern airplane. The flight test is the traditional method used for evaluating the performance of an airplane and gathering the information needed to prove out its basic design. In the early days of flight tests, the pilot would read the instruments on his control panel and note his observations on a pad strapped to his knee. As design techniques were refined, test engineers required more extensive data and introduced more automatic methods of data acquisition to free the pilot for his primary mission of flying the plane. Measurements were brought to dials mounted on a panel which was photographed by a motion-picture camera. Graphic recorders, using pen-and-ink and oscillographic traces on photographic film, permitted continuous recording of high-frequency information, such as vibrations and flutter. Many of these techniques are still in use today and serve a useful purpose. But the demand for more and more measurements and a more rapid reduction of these measurements to usable physical quantities upon completion of the flight, requires the use of more modern techniques. This has been one of the major factors behind the recent growth of magnetic tape development.

To illustrate the magnitude of the problem, consider this fact: The number of measurements which have been taken on a single test flight of a B-52 bomber exceeds all of the measurements which were taken in all the flight tests that were ever performed on the old B-17. One jet transport designer today has a test program which involves taking one million measurements during a single test flight. If these measurements were to be recorded on photo panels, or on oscillographs, several weeks of concentrated effort, involving large numbers of persons, would be required to reduce the data by manual means, before the design engineer would have the significant numbers he needs to either correct or improve the design. In the accelerated plane and missile programs which we face today, time lags of this magnitude cannot be tolerated.

INHERENT CAPABILITIES OF MAGNETIC TAPE RECORDING:

The second principal reason for the rapid growth of tape for measurement and analysis has been the inherent capabilities of tape itself as a recording medium. Listed below are some of these advantages which will be amplified later:

1. Wide frequency range. Magnetic tape permits the recording of information from dc up to megacycles.

2. A very wide dynamic range of recording, in excess of 50 decibels, permitting accurate and linear recording from full-scale level (100%) down to 1/3 of 1% full scale.

3. Magnetic tape has low inherent distortion characteristics. When overload occurs, it occurs rather gracefully, as contrasted with a galvanometer or other mechanical devices.

4. The signal information is preserved in its electrical form, so that the original event can be recreated at any future time. This lends itself, of course, to automatic reduction of the data.

5. Recordings made on tape are available for immediate play-back, with no time lost in photographic processing.

6. The economic advantage that the tape itself is reusable, since it can be erased many times.

7. Tape can be played back thousands of times, which permits extracting every bit of useful information from the recording.

8. Tape provides facility for multiple-channel recording. Thousands of channels of information may be recorded simultaneously,

using various multiplexing techniques. Very accurate time and phase relationships can be maintained between these simultaneous signal channels.

9. Very high-density storage may be provided by tape, with up to several million data points contained on a single reel $10\frac{1}{2}$ inches in diameter.

10. Tape provides something which no other medium provides--the ability to alter the time base. This permits events to be recreated on playback either faster or slower than they actually occurred, with resulting multiplication or division of all frequencies involved.

MILESTONES IN THE EVOLUTION OF AMPEX MAGNETIC TAPE RECORDERS:

1900 FIRST MAGNETIC RECORDER patented in the United States. Having developed, exhibited, and patented a magnetic wire recorder in Europe, Valdemar Poulsen was granted a United States patent on his Telegraphone. Output was very low with poor signal-to-noise ratio.

1906 FIRST D C BIAS patented in the United States. Poulsen and Pedersen had found that pre-magnetizing the wire in their Telegraphone recorder with a direct current would yield increased output but with a low signal-to-noise ratio.

1927 FIRST A C BIAS patented in the United States. During their work on magnetic recording techniques in a U.S. Navy laboratory Carlson and Carpenter had discovered that alternating current bias was in every way superior to direct current pre-magnetization.

1935 FIRST MAGNETIC TAPE RECORDER introduced in Europe. The first magnetic tape recorder was introduced using iron oxide coating impregnated in non-metallic plastic film or paper base. All previous models of magnetic recorders had used steel wire, steel ribbon, or plated metallic ribbon.

1948 UNITED STATES FIRST MAGNETIC TAPE RECORDER (AMPEX MODEL 200). Magnetic recorder production in the U. S. prior to this year had utilized steel wire or coated metal ribbon.

1950 INSTRUMENTATION RECORDERS INTRODUCED (AMPEX MODEL 300). Custom adaptations of professional audio tape recorders to instrumentation requirements.

1956 100% INSTRUMENTATION TAPE RECORDER (AMPEX MODEL FR-100). Conceived, designed and built expressly for instrumentation requirements. Modular signal electronics permitted simple plug-in modules to suit any instrumentation application.

1957 COMPUTER ORIENTED TAPE HANDLER (AMPEX MODEL FR-200). Provided a superior storage capacity and a more rapid transfer of data than punched cards or punched tape.

1957 VIDEOTAPE RECORDER (AMPEX MODEL VR-1000). Provided immediate record/reproduce of a video camera output or composite video signal.

1960 BROAD BAND INSTRUMENTATION TAPE RECORDER (AMPEX MODELS FR-700 and AR-300). Provides recording/reproducing of two channels of 10 cycle to 4 mc information.

INTRODUCTION TO BASIC ELEMENTS OF A MAGNETIC TAPE RECORDER:

To record a signal on tape it is first necessary to change the input voltage into a current whose frequency and amplitude correspond to the signal to be recorded. This current produces a varying flux in the head. When the tape passes over the head a varying magnetic pattern is recorded, which corresponds to the original signal. Refer to Figure 1.1 for the basic elements of a simple magnetic recorder.

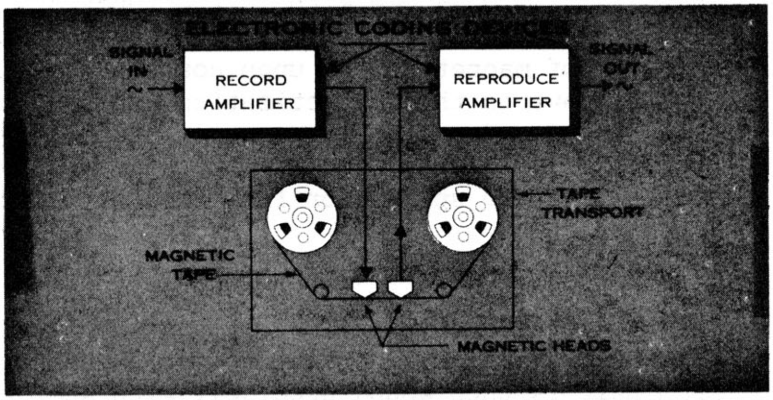


Figure 1.1 BASIC MAGNETIC TAPE RECORDER

To reproduce the previously recorded signal, the tape must be passed over the reproduce head. The field pattern induces a small voltage across the reproduce head. The function of the reproduce amplifier is to amplify this voltage to a level and form equal to the original recorded signal.

The tape transports mechanism's basic functions are:

1. To draw the tape past the recording and reproducing heads at a constant and accurate rate of speed.

2. Store sufficient quantities of tape for the amount of record/reproduce time required.

In the following chapters each of these basic elements will be treated in more detail.

CHAPTER II

MAGNETIC HEADS

EVOLUTION IN MAGNETIC HEADS:

Somewhat analagous to the weak-link and strong-chain maxim is the fact that a magnetic instrumentation recorder isn't one iota better than its head. In fact, it is widely acknowledged that the head is the "heart" of the system. Without near-flawless head performance, the value of the system is in jeopardy.

Put simply, the head is a tiny device with triple functions - all vital. First, it converts information, in the form of encoded electrical signals, into a pattern of magnetization upon coated tape - the storage medium. Second, it reconverts this magnetization pattern into electrical signals, to be decoded and put to an array of practical use. Finally, it obliterates old information, wipes it forever from the tape, leaving the slate clean, so to speak, for infinite further use.

IF THE HEAD IS THE "HEART", WHAT MAKES IT BEAT?

Fundamentally, the magnetic head is made up of two identical core halves, built of thin laminations of a very special magnetic alloy (Figure 2.1). Each half is then wound precisely with an identical number of turns (the exact number depends on which of the head's three functions is intended), and assembled with non-magnetic separators at front and rear, with a minuscule gap remaining at the front, or bottom - which is the business end of the works. Here's why:

Magnetic tape consists of an acetate or Mylar* plastic base coated uniformly with extremely fine particles of magnetic oxide material. Signal current from the amplifier flows through the head windings and around the core halves, producing magnetic flux. As the tape transport draws the tape across the head, the gap is shunted, and the flux path is completed. The tape moves across the gap at a linear velocity; each magnetic oxide particle, then, is magnetized permanently to a degree proportional to flux flowing in the head gap at that instant. (Figure 2.2)

The head gap is equally critical in the reproduction phase of the system. When the now-magnetized tape is drawn across the gap, the portion of tape in actual contact with the gap bridges the head's magnetic core, causing flux to flow through the core. (Figure 2.3)

* A Du Pont Trademark

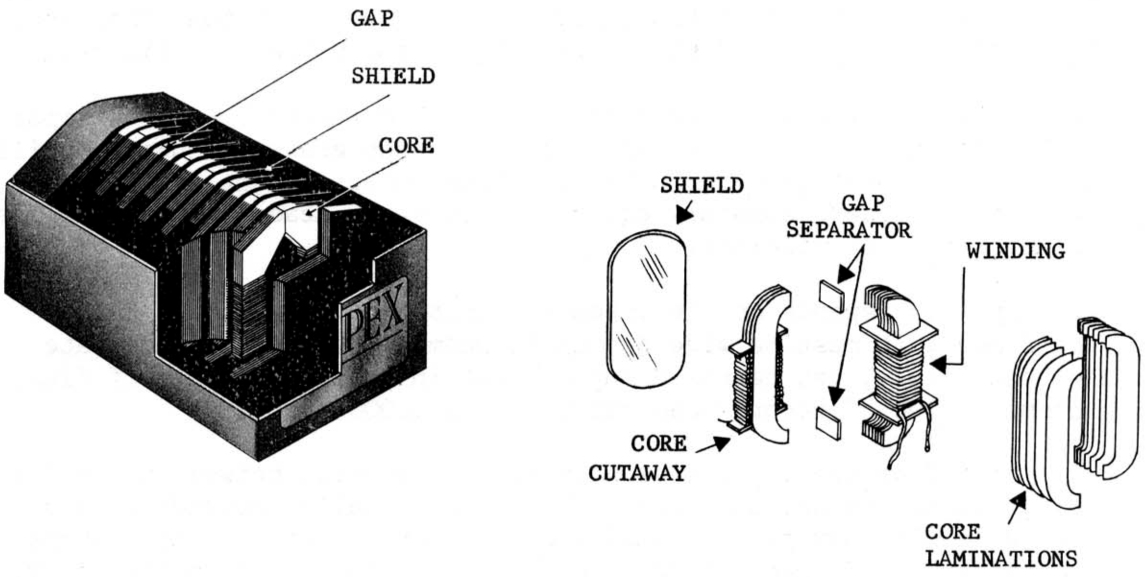


Figure 2-1 Construction and Component Parts of Magnetic Head

RECORD HEAD

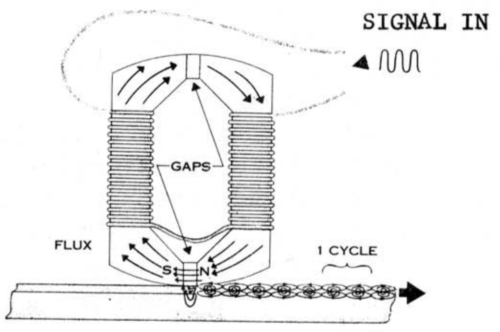


Figure 2-2 Magnetization of Tape's Oxide Particles Depends on Flux Flow in Gap at Instant Each Crosses Trailing Edge

REPRODUCE HEAD

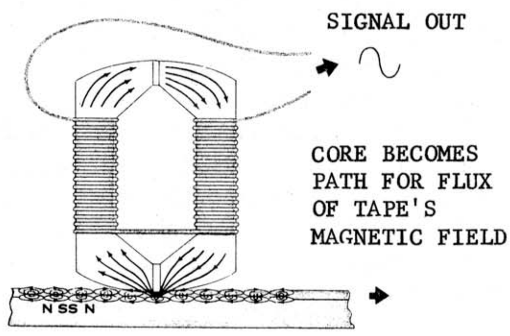


Figure 2-3 Reproduction of Signal From Magnetic Tape

The magnetization pattern of the portion of the tape actually spanned by the gap at a given instant determines the strength of the flux, and ultimately, the nature of the reproduced signal. In actual practice, of course, the reproduced signal is a perfect duplicate of the original.

The function of the erase head gap is not so critical. What is important is that the magnetized tape pass through enough erase frequency oscillations (in the head gap) to insure complete erasure. The only instrumentation recorders commonly equipped with erase heads are the familiar tail-chasing loop machines.

Head-gap size depends on the intended function of the head. For a record head, it must be wide enough to permit the flux to penetrate the tape deeply, yet narrow enough to obtain sharp gradients of flux. A common value of record-head gap width is 0.0005 inch.

For a reproduce head, gap size must be a compromise between upper-frequency limit, dynamic range, and head life. Theoretically, reproduce-head voltage is directly proportional to frequency. Voltage output drops at a constant 6 db-per-octave rate for each octave drop in frequency. But at the high-frequency end of the spectrum, gap size becomes important. Since average value of tape magnetization is zero when wave length equals gap size, as wave length approaches gap width, the reproduce-head output wanes rapidly. (Figure 2.4) While there is no way to banish this phenomenon, there are several ways to live with it gracefully. The reproduce-head gap can be narrowed, or tape speed can be hiked. But with either alternative, there is a compromise. If gap size is reduced, output voltage slacks. And as tape speed goes up, accelerated head wear is inevitable. Good head design strikes the best balance between these conflicting parameters.

Current instrumentation needs have created a heavy demand for multiple-track heads, in which two to fourteen tracks are aligned in a single stack. Such a head is composed of several cores, each built as described above. These are moulded together in plastic to form a head stack. But each track actually is a separate head in itself. Many problems arise in designing head stacks. Maximum tape widths available restrict track width - but wide tracks provide the best signal-noise ratio. Available tape widths also require close between-track spacing. This condition, though, leads to the undesirable effect of crosstalk, i.e., capacitive, inductive and magnetic coupling between adjacent tracks. Inter-track shielding curbs crosstalk to some degree.

To keep crosstalk at a minimum, analog-recorder heads are usually staggered, or interleaved. This arrangement permits close track-to-track spacing on the tape, but because tracks are separated individually, rather than in-line, there is less crosstalk.

Extremely close mechanical tolerances are crucial to optimum multiple-head performance. Among the dimensions demanding tightest control is that between gap center lines of the two interleaved stacks. At Ampex, this is held to ±0.0005 inch. Two other vital mechanical factors are gap scatter and azimuth. Gap scatter denotes alignment of track gap-center lines within the stack. Deviation from the linear alignment is held to less than 100 microinches. Gap azimuth - the perpendicularity of gap center line with tape surface - is maintained within plus or minus a single minute of arc. (Figure 2.5)

These tolerances pertain to accurate gap alignment in multitrack heads. However, in some cases it is desirable to vary this alignment. A variable head arrangement, featuring a ±5/8" track advance or retard, allows controlled time delays or advances up to 40 milliseconds at a 30 ips tape speed. With this arrangement, auto-correlation and cross-correlation techniques can be applied to analysis of signals obscured by background noise.

The tolerances under discussion, it should be emphasized, are Ampex standards and are attained not on a model-shop basis but on full production runs. They reflect both painstaking assembly and inspection, and inherent Ampex head-design characteristics. All heads are 100 per cent inspected; each has an inspection record which corresponds to its serial number.

To consider exacting mechanical tolerances, per se, is meaningless. What is significant is their overwhelming effect on electrical tolerances. The relationship is very nearly a direct one. Excessive mechanical deviations, whatever their nature, can be expected to generate electronic performance errors of similar magnitude.

Exacting mechanical tolerances held at Ampex permit these enviable electrical tolerances - or performance standards:

1. Plus or minus 2 db deviation from head current level at 1000 cps (60 ips tape speed).

2. Frequency response within plus or minus 3 db (60 ips tape speed).

The genesis of recording heads was an all metal affair. Known as a sandwich type head, its mu-metal laminations were encased in a non-magnetic alloy. (Figure 2.6)

The need for greater precision dictated the switch to the cast epoxy head. First, the method of stacking each sandwich channel individually in the all metal head was prone to error. Perpendicularity of tracks - to avoid frequency loss, and accurate alignment of gap center lines - to avoid excessive gap scatter, were tough to achieve, even when the separate channels in the stack were positioned microscopically.

The Ampex cast epoxy head was actually a new concept in head design. Each core half was cast separately. Channels were made in a multiple operation which assured their uniformity. Then they were fused with epoxy resin to the separately cast core halves. Finally, the halves were lapped mechanically with optical precision. The finished product had no peer in track perpendicularity and gap-center alignment. (Figure 2.7)

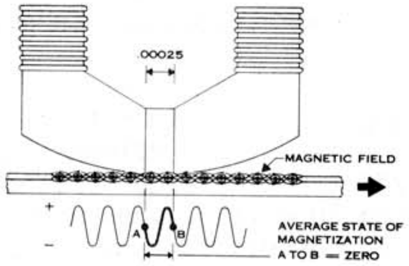


Figure 2-4. Effect of Wave Length and Gap Size on Reproduction Output

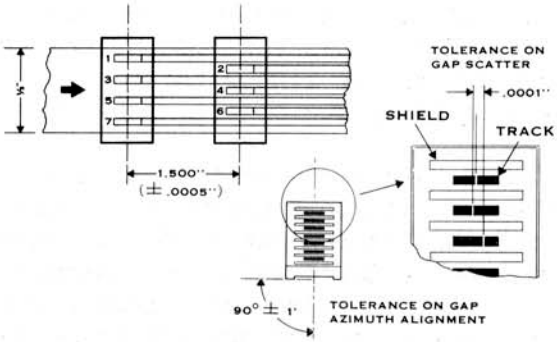


Figure 2-5. Ampex Mechanical Tolerances

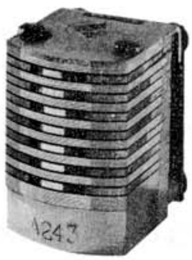


Figure 2-6. All Metal Head -- Original Ampex Design

Figure 2-7. Cast Epoxy Head Featured Separately Cast Core Halves

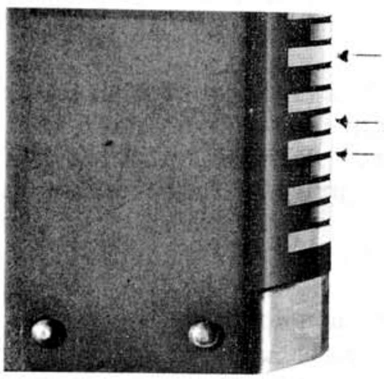


Figure 2-8. Head Surfaces Isolated in Newest Design

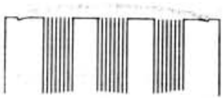


Figure 2-9. Tendency of Tape to Groove Earlier Types of Heads

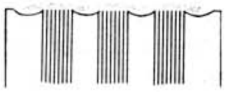


Figure 2-10. Advantages of Current Ampex Design

The next transition to the current Ampex head design was prompted primarily by the need for more perfect tape-to-head contact in instrumentation systems. Signal losses, however infrequent, are intolerable in instrumentation. Any one of several thousand pulses may be crucial in a complex mass of information.

Cutting away the special plastic insert which sheathes each head surface reduced the area which the tape must contact; (Figure 2.8) consequently, the tape is able to "ride" the head gap much more closely, providing superior signal reproduction.

This gain in tape-head contact actually stems from higher unit pressure where hold back tension is the same. The improvement is on the order of 10 - 15 per cent. Of course, hand-in-hand with higher tape pressure goes increased head wear; however, this increase is so slight, according to Ampex head engineering specialists, that it is a negligible factor when considering the gains in signal reproduction. Actually, these specialists state, the magnetic head, whatever its construction, will have a life at least commensurate with major overhaul time on other components in the system. For example, head life should be on the order of 5,000 hours on either an Ampex FR-100 half-inch machine with 8 oz holdback tension, or an FR-100 one-inch machine with 16 oz holdback tension.

Two other barriers to better signal retention were alleviated in the new design. For one, there was the ever present problem of oxide particle buildup on the head surface. Because tape coatings - though much improved - are still imperfect, small particles tend to loosen and accumulate on the head surface. Obviously, as this accumulation mounts, the tape surface will be lifted away from the head gap, causing appreciable signal loss.

The new Ampex head will not eliminate such oxide buildup, which is a factor of tape coating materials and methods, and can be expected to persist until tape technology is refined considerably. However, the new Ampex head design does provide natural recesses in which oxide particles, once loosened from the tape surface, can collect without causing either signal loss or tape and head surface damage. Periodic head cleaning removes the deposits.

Also resolved was the tendency of tape to "groove" earlier types of heads. This means, simply, that edges of any tape would tend, in time, to cut grooves in the plastic or metal inserts shielding the head surfaces (Figure 2.9 and 2.10). Because tape widths - like tape surfaces - are prone to deviations, a fresh tape (although the same specified width) might not always precisely fit the grooves carved by its predecessor. The upshot often was signal loss, stemming from wrinkling or skew which lifted tape surface from head gap.

Since the plastic filler separating the cores has been cut away there is nothing left for the tape edges to groove, and minute deviations in tape width are, therefore, immaterial.

The demand for ever increasing data storage capability has raised the upper recording frequency limits to the 4 megacycle region and above. To successfully record signals at these frequencies it is expedient that a head-to-tape velocity of approximately 1300 ips be employed. Conventional techniques using longitudinal recording no longer suffice for these requirements. The wavelength of the recorded signal will be directly proportional to tape speed and inversely proportional to the frequency of the recorded signal.

$$\lambda = \frac{s}{f}$$

λ = wavelength on tape (inches)

s = tape speed (inches/second)

f = frequency (cycles/second)

To achieve this higher head-to-tape speed rotating heads are utilized. (Figure 2.11)

The rotating heads sweep transversely across the tape while the tape longitudinal velocity is just fast enough to properly space the tracks. Four heads are mounted in quadrature around the drum. A concave guide holds the tape against the rotating heads.

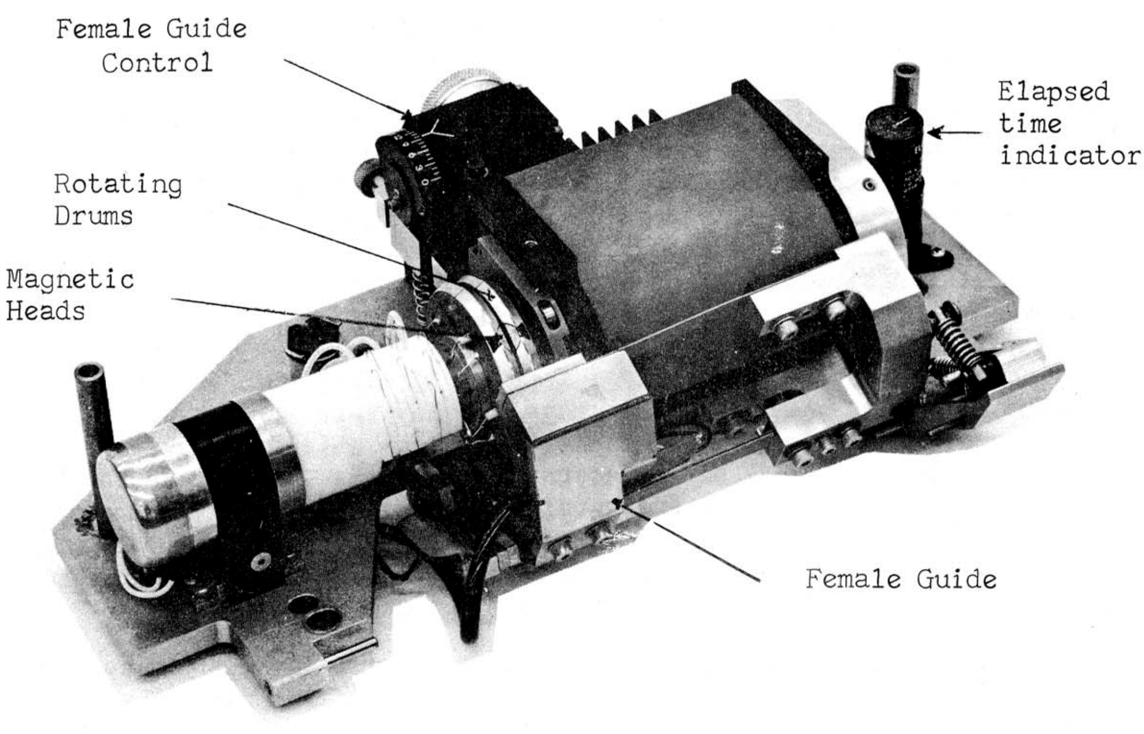


Figure 2.11 WIDEBAND ROTARY HEAD ASSEMBLY

TYPICAL HEAD CHARACTERISTICS

	Analog	Digital	Rotary
Wire size (record)	42	42	42
Wire size (reproduce	46	42	42
Number turns (record)	60	40	42
Number turns (reproduce)	300	160	42
Track width (record)	50 mils	.032"	10 mils
Track width (reproduce)	50 mils	.024"	10 mils
Gap width, front (record)	.0005"	.00025"	115μ"
Gap width, rear (record)	.0005"	None	---
Gap width, front (reproduce)	.00025"/.00008"	.00025"	115μ"
Gap width, rear (reproduce)	.00025"/.00008"	None	---
Record current *	18-49 ma	75 ma	22 ma
Reproduce voltage *	50μv min.	12.5 mv	80μv
Track spacing 14 tracks	.070"	.070"	.0156"
Gap - gap spacing	1.5" $\pm$ .0005"	.390" or 1.375"	---
Gap azimuth	$90^o \pm 1'$	$90^o \pm 1'$	---
Gap scatter	100μ"	100μ"	---
Point of recording	Trailing edge	Trailing edge	Trailing edge
Point of reproducing	Across gap	Across gap	Across gap

*Peak values

CHAPTER III

MAGNETIC TAPE

INTRODUCTION

The universally accepted medium for magnetic recording is a tape consisting of a plastic base with a coating of minute ferrous oxide particles.

TAPE BACKING

The tape backing consists of thin sheets of acetate or Mylar* film. Thicknesses commonly employed are 1.5 mil, 1.0 mil and 0.5 mil.

CHARACTERISTICS OF TAPE BACKING MATERIALS
(Based on 1.5 mil, .25" wide tape)

	Acetate	Mylar*
Economy	Less expensive	More expensive
Uniformity of thickness	Excellent	Good
Tensile strength	5.6 lbs	11 lbs
Tear strength	4 grams	25 grams
Mildew resistance	Low	High
Fungus resistance	Low	High
Coefficient of thermal expansion Per 1^{o} Change of temperature (70^{o} - 120^{o} F)	30×10^{-6} in./in.	15×10^{-6} in./in.
Coefficient of humidity expansion Per 1% Change of Relative Humidity (20 - 92% R.H.)	150×10^{-6} in./in.	11×10^{-6} in./in.
Width tolerance	$.248 {}^{+.000}_{-.004}$	$.248 {}^{+.000}_{-.004}$
Length tolerance	-0 +30'	-0 +30'

*T.M. DuPont

TAPE COATING

It should be remembered that the base material is merely a support for the magnetic layer. The magnetic layer (coating) may be classed under two headings:

1. Magnetic oxide
2. Binder material

MAGNETIC MATERIAL

The most suitable magnetic material for tape coating is the acicular form of gamma ferric oxide ($Fe_2O_3\gamma$). It will have a particle length of .2 to .8 microns (1 micron = 10^{-6} meters or 39 μ inches) and a width of 1/2 to 1/6 of its length; a specific gravity of 4.7 and a coercivity of approximately 260 oersteds. This is obtained from the basic raw material which is the alpha form of Fe_2O_3.

The manufacturing process to obtain the end product is as follows:

1. Raw material - yellow ferrite ($Fe_2O_3 \cdot H_2O$) is dehydrated to αFe_2O_3. This is red, non-magnetic oxide.

2. Addition of hydrogen or carbon monoxide plus heat converts the αFe_2O_3 to Fe_3O_4 (black, magnetic ferrous ferric oxide).

3. Application of oxygen and heat (under closely controlled conditions) will produce the final gamma form of Fe_2O_3.

During the manufacture of the various oxides, extreme care is taken to produce a crystal of a particular size and shape. The crystal formation is due to the dehydration process. This process encourages the crystals to adhere, forming small masses. These masses are then separated into individual particles in a ball mill. Care is taken during the balling that no grinding takes place in order to prevent reduction of the particle size.

During the ball milling process the binders are added, which consist of:

1. Binder - cement that holds particles onto the backing.

2. Plasticizer - gives flexibility to binder.

3. Wetting agent - used to keep the Fe_2O_3 particles from re-combining.

4. Lubricant - prevents the binder from adhering to the next layer of tape when wound on a reel.

5. Resin - aid in dispersion of Fe_2O_3 particles and toughen coating.

6. Solvents - enchances the binding between the coating and backing.

7. Anti-bloom agent - prevents powder-like residue from forming on tape.

The ratio of oxide to binder generally is 60 to 40 percent. The particle size and uniformity determines to some extent the surface roughness of the tape coating which is one of the contributing factors in high frequency resolution.

Figure 3-1. Sketch of Electron Microscope Presentation of Magnetic Oxide Particles, Showing Typical Size and Shape.

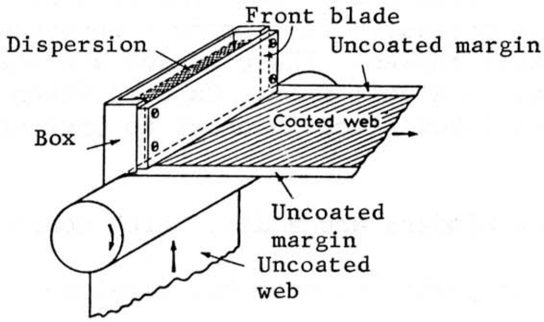


Figure 3-2. Knife Coating Process

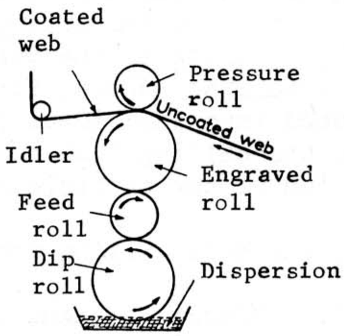


Figure 3-3. Rotagravure Coating Process

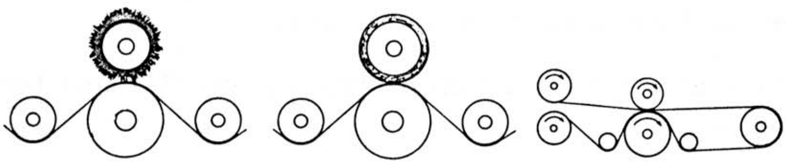


Figure 3-4. Methods for Polishing Magnetic Tape.
(a) High Speed Brushes (b) Buffing Wheel (c) Self-rubbing Action

COATING AND POLISHING TECHNIQUES:

There are three principle methods of applying the coating to the backing. These are:

1. Knife or blade - coating spread on tape.

2. Rotogravure - coating applied to an engraved roller which deposits small beads of coating on backing which settle to a uniform thickness.

3. Reverse rollers - controlled amount of coating poured on roller and deposited on backing.

In each case the thickness of the coating is closely maintained. Common thicknesses are .41 and .46 mils.

After backing is coated the tape is then polished to remove the small needle-like projections caused by the particle shape. The four principle methods of accomplishing this are:

1. High speed brushes of nylon or horsehair.

2. Buffing wheel.

3. Self-rubbing action.

4. Ferrosheening.

In the first three methods the protuberances are either ground down or pulled out. In the latter method these protuberances are flattened into the surface.

The final step in the manufacturing process is the slitting of the tape into specific widths.

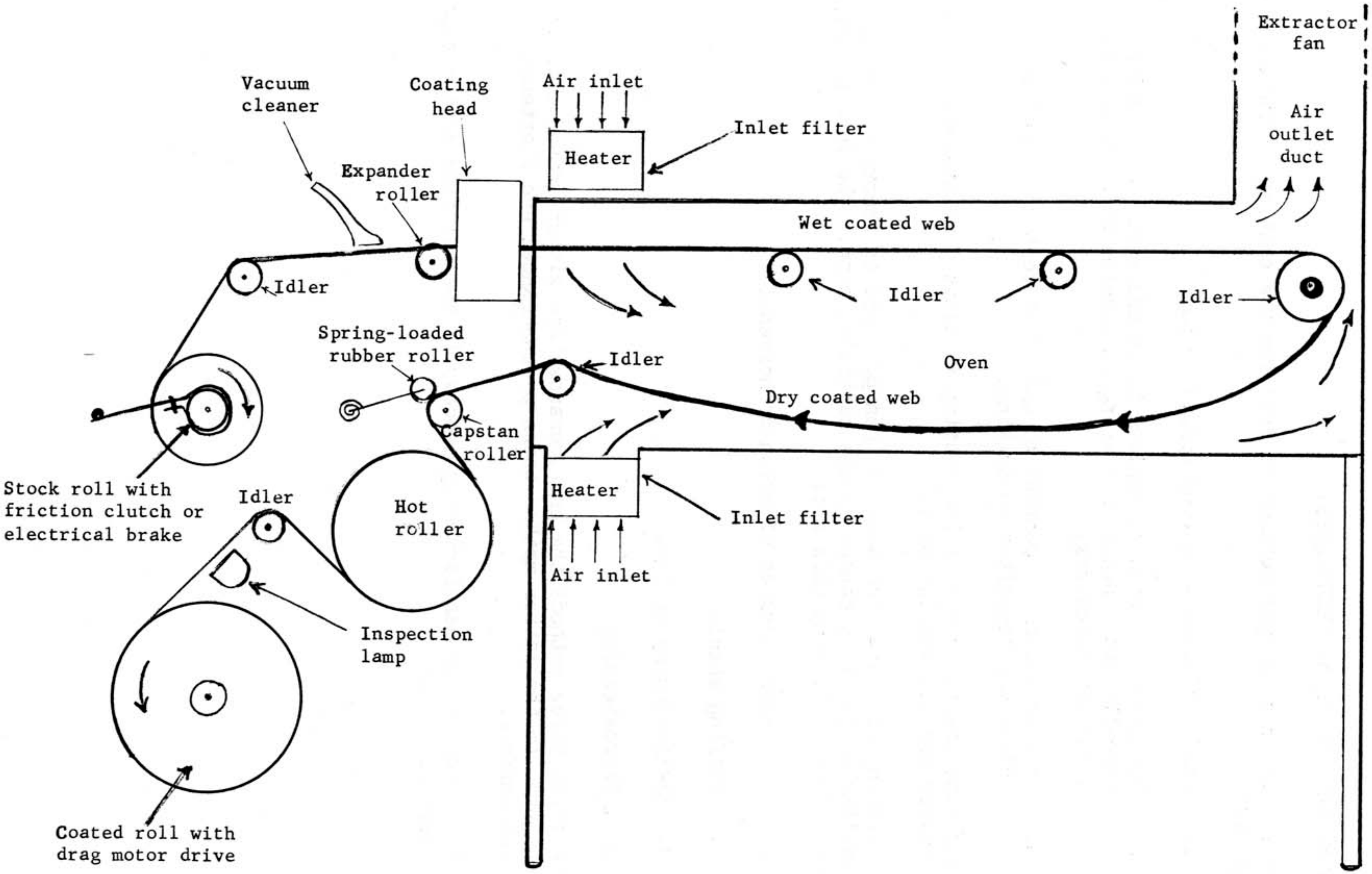


Figure 3-5. Typical Coater

CHAPTER IV

THEORY OF MAGNETISM:

In understanding the fundamental principles of magnetism, we must consider the modern theories of spinning electrons in discrete orbits and the group behavior of molecular masses called domains.

Recently physicists have explained the mechanism of the atom and included in their conjectures the number of electrons in the atoms, their orbital placements, and direction of spin. This backs to a great extent Ampere's theories of molecular magnets of over 100 years ago.

It may, therefore, be concluded that:

1. All Ferromagnetic materials are crystalline structures.

2. Each elemental crystal has a unique and ordered arrangement of atoms.

3. There are uncompensated spinning electrons in each atom, and the spin of the electron around its own axis has exactly the same effect as the flow of current in a coil.

4. Each atom in the crystal unit is a miniature magnet.

From these conclusions we can then expect that these miniature magnets will influence each other, just as a group of bar magnets would if they were similarly arranged.

It is because of this intracellular interaction that some of the atoms will be more readily influenced by one directional magnetic field than another.

THE WEISS THEORY OF DOMAINS:

The Weiss theory assumes that the ferromagnetic materials are made up of large numbers of elemental volumes, all of which are magnetized to saturation due to the spinning of uncompensated electrons. These elemental volumes are called domains and are so randomly oriented in regard to their direction of magnetizm that the material as a whole appears unmagnetized. The numerous exchange forces between atoms serve to align the axis of adjacent spinning electrons. Although these alignments usually extend over appreciable atomic areas, the actual volume of alignment is limited. These limited volumes are called domains. Thus each domain is spontaneously magnetized to saturation in some direction.

Actual changes which take place in the material being magnetized consist of two types, namely:

1. Changes in volumes of some domains at the expense of others

2. Changes in direction of magnetization of domains.

The directional changes can be either reversible or nonreversible. Reversible changes are those which automatically disappear when the originating magnetizing force is removed. Nonreversible changes are those which require the application of a magnetizing force in the opposite or near opposite direction to erase them.

Domains are envisaged as being bounded by partitions, called Bloch walls. The actual positions of these walls are assumed to be associated with points of strain. It is believed that when subjected to a magnetizing force, the walls of the domain, whose directions of magnetization roughly approximate that of the magnetizing force, move outward at the expense of the other domains with widely divergent directions of magnetization. (Figure 4.1)

A. The application of very weak magnetizing forces, corresponding to the magnetization from the origin up to the instep of the magnetization curve, are associated with changes in the Bloch walls. The removal of the magnetizing forces cause the Bloch walls to return to their original position.

B. Stronger forces, which raise the magnetization up the relatively straight portion of the magnetization curve between the instep and the knee, cause changes in orientation from one preferred direction of magnetization to another that is nearer to the direction of the magnetizing force. There will be relatively small change in induction when the magnetizing force is removed. The change that has taken place is irreversible and the application of a substantial magnetic force in the reverse direction is required to return the material to something approaching its original condition.

C. A very strong magnetizing force takes the magnetization over the knee of the curve into the region of saturation.

The direction of magnetization is now being rotated toward that of the magnetizing force under conditions of strain which is immediately relieved as soon as magnetizing force is reduced. Magnetization of this kind is reversible, as indicated by the substantial drop to B_r value when the magnetizing force is removed.

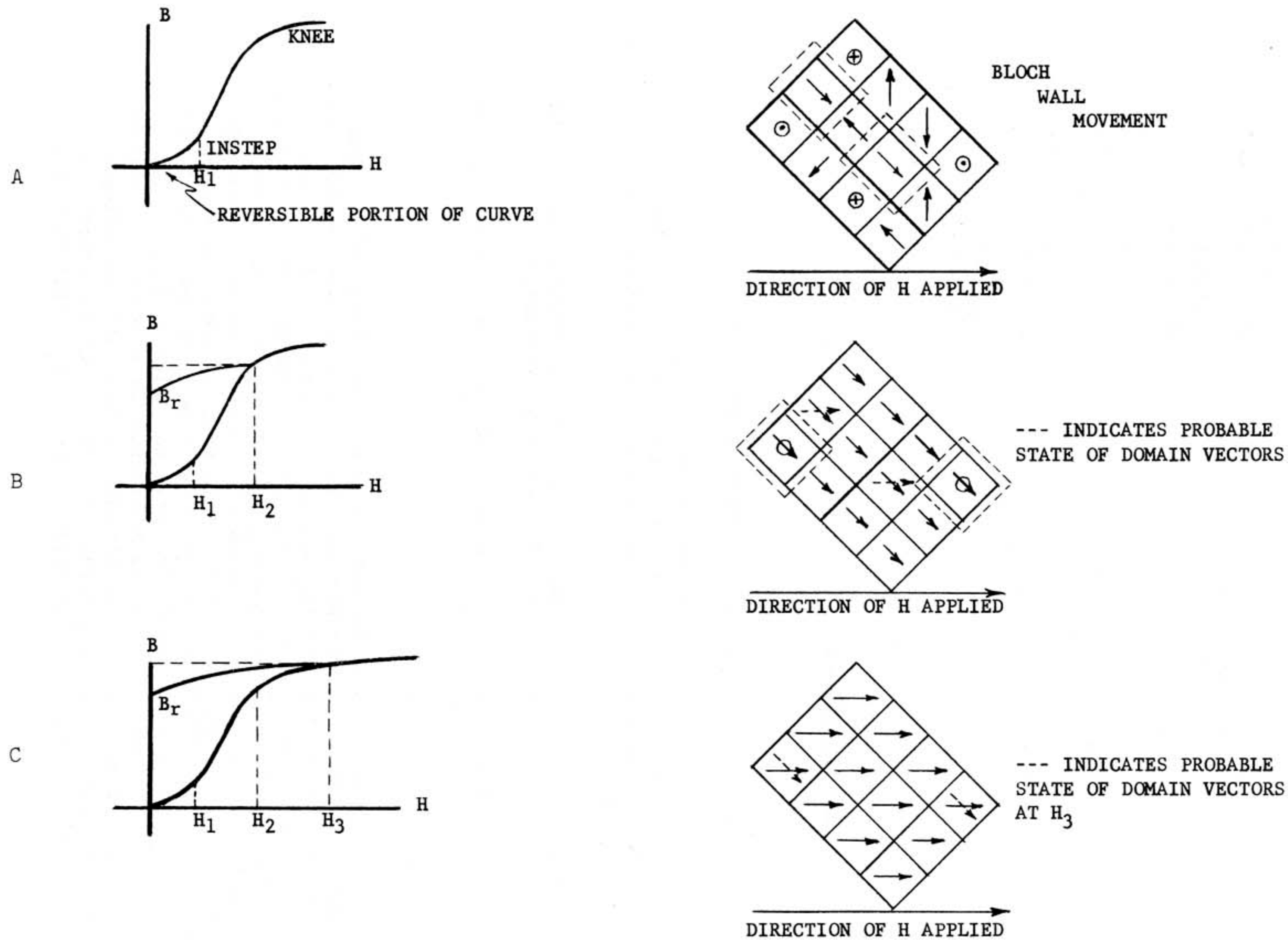
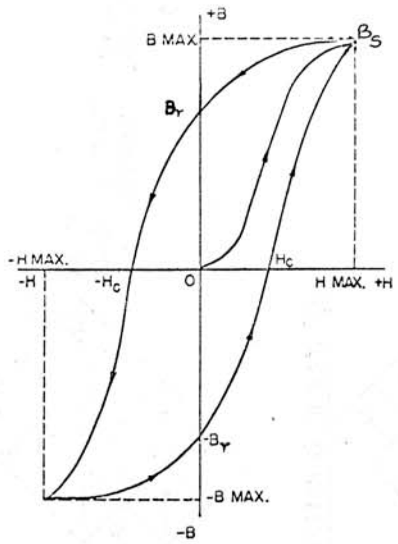


Figure 4-1 EFFECTS OF MAGNETIZING FORCE ON DOMAINS

HYSTERESIS LOOP AND MAGNETIZATION CURVE:



B_S Saturation Inductance

B_r Residual Inductance

H_c Coercive Force

H Magnetizing force in Oersteds

B Induction in Kilogauss

Figure 4.2 Typical Hysteresis Loop

The hysteresis loop may be considered as nothing more than a visual indication of the lagging of magnetic induction (flux-,B) behind the magnetizing force that produces the magnetism in the material (H).

The horizontal axis represents the value of magnetizing force (H) in oersteds per unit length applied - positive and negative.

The vertical axis represents the amount of flux per unit area (B) in gauss induced in the material.

THE HYSTERESIS LOOP:

Measurement begins with the material completely demagnetized. The magnetizing force is increased to a positive value beyond which a further increase in the magnetizing force produces no further increase in magnetism of the material. This registers a curve from 0 to B_S (saturation induction). Then the magnetizing force is reduced from H_{max} to 0. However, a certain amount of induced magnetism (depending upon its remanence) remains in the material and is represented by the intersection of the return curve from B_S to the vertical axis at B_r (residual induction). The numerical value of B_r is called the retentivity of the material and represents the peak residual magnetism that the material will retain without demagnetizing influence. From this point, the magnetizing force is changed in direction and increased to a negative saturation. Repetition of this magnetization and demagnetization process will result in the establishment of the complete symmetrical loop about the point 0 - the static hysteresis loop.

The value H_c to 0 is known as the coercive force. This is a measurement of the force required to reduce the induced magnetism to zero. The coercivity of the tape is a measure of the magnetizing field necessary to record or erase signals.

Non-linearity which is associated with the hysteresis loop is of great importance in magnetic recording. It may be represented by the B_r - H curve which shows the residual magnetism (remanent magnetism) against magnetizing force. This curve is found by applying a small H and measuring the B_r after the H has been removed. Material is then demagnetized and the measurement taken again at a higher value of H. The relationship is non-linear everywhere but conspicuously so at high and low values of H. (Figure 4.3)

FAMILY OF HYSTERESIS LOOPS:

When the maximum field strength required to obtain the normal hysteresis loop is diminished in a series of successive steps, a family of hysteresis loops results. (Figure 4.4) The loop shapes for small fields are characteristically double-convex lens. When the field is increased, the loop will widen until it assumes its saturated shape.

THE COMPOSITE HYSTERESIS LOOP

When two alternating fields are superimposed simultaneously on a ferromagnetic substance, a complex set of curves is established. Figure 4.5 indicates a combination of a 10,000 cycle per second signal which traces the larger or major loop upon which is superimposed a high frequency bias of 335,000 cps, which traces a series of minor loops. This configuration will produce the effect of a collapsed major loop passing through the center of the minor loops. An increase of high frequency bias will tend to collapse the major loop into a relatively straight line, thus high frequency bias tends to remove the non-linear characteristic instep and knee formation of the normal B_r-H curve. (Figure 4.3).

ERASURE

When a magnetic material is placed in a gradually diminishing cyclic state a family of successively smaller and smaller loops will be formed until demagnetization (erasure) is reached. This is the process used when degaussing a tape. (Figure 4.6)

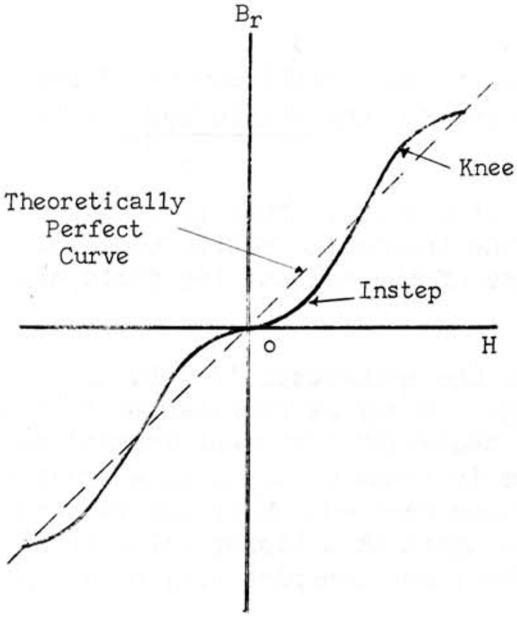


Figure 4.3 Typical B_r-H Curve

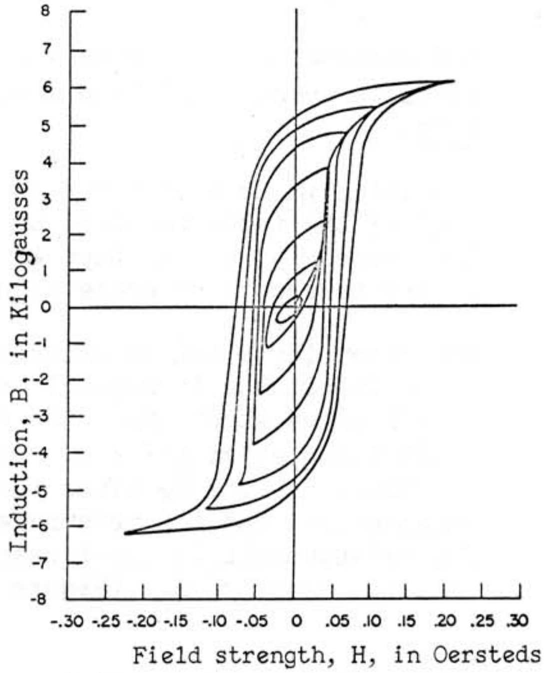


Figure 4.4 Family of Hysteresis Loops

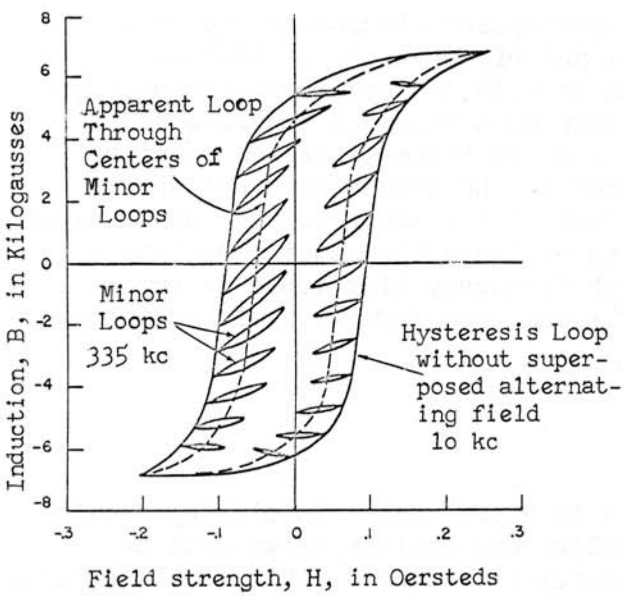


Figure 4.5 Composite Dynamic Hysteresis Loops

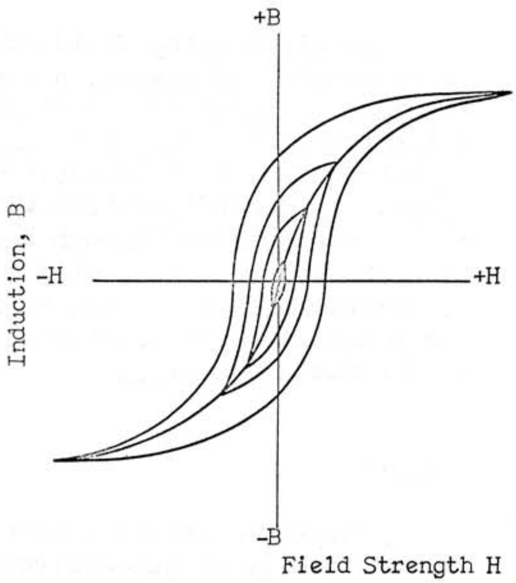


Figure 4.6 The Transduction Curve (Demagnetization)

CHAPTER V

THE RECORDING AND REPRODUCING PROCESSES

INTRODUCTION:

The universally accepted medium for magnetic recording is a tape consisting of a plastic base, with a coating of minute ferrous oxide particles dispersed in a synthetic resin binder. Expressed in accordance with the domain theory, the coating contains a ferromagnetic material comprising a multplicity of regions, called "domains", each magnetized to saturation in some direction. The direction of magnetization, in the absence of strain and magnetic field, is determined by the orientation of the crystal structure. The domains in the unmagnetized, or neutral state of the material are randomly distributed in a way such that the resultant magnetic flux external to the tape- referred to as "surface induction"- approaches zero. Application of a magnetizing field changes the direction, but not the magnitude of magnetization of the domains.

THE RECORDING PROCESS:

Conventional magnetic heads, whether used for recording or reproducing, are basically ring-shaped cores of high-permeability material, with a high reluctance gap across which the tape is passed. The presence of tape in the region of the recording head gap causes penetration of the coating material by magnetic lines of induction, in the form of leakage flux extending across the gap between the pole pieces of the core. The decreased reluctance of the ferromagnetic coating with respect to that of the gap distorts the field, causing a degree of flux concentration in the tape.

Figure 5.1 illustrates the approximate distribution of leakage in a cross-sectional view of tape coating and pole piece tips having typical dimensions. The edges of the gap are shown as slightly rounded because a truly sharp edge is impossible to attain in practice. However, even though a sharp edge could be produced, the extremely high flux density existing at the sharp corner could saturate the core material and create the effect of a radius.

The magnetizing field intensity produced in the tape by a recording head is proportional to the magnitude of the current flowing in the head coil windings. Under the influence of a low flux density, a small number of randomly-aligned magnetic domains in the tape coating are forced into alignment. A high flux density will cause a large number of domains to be forced into alignment with the flux pattern in the region of the gap. Because of hysteresis and demagnetization effects, the intensity of magnetization normally is not directly proportional to the magnetizing field intensity.

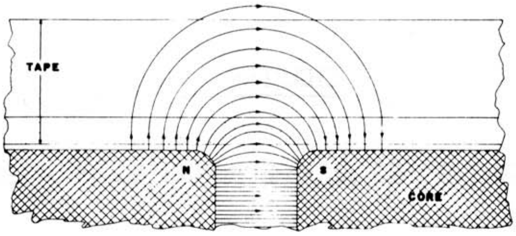


Figure 5-1 Leakage Flux Distribution in Recording-Head Gap Region

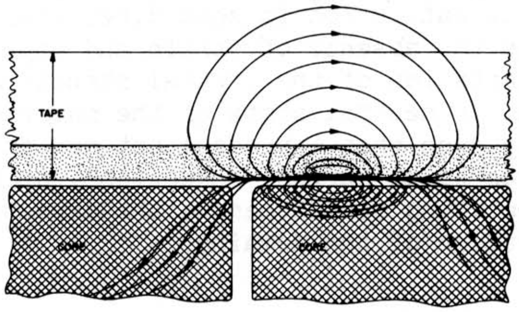


Figure 5-2 Flux Distribution of Recorded Tape in Reproducing-Head Gap Region

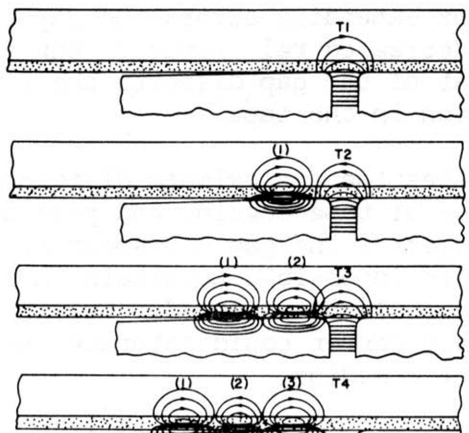


Figure 5-3 Simplified Synthesis of Recording Process

If the tape is moved linearly past the gap, with a sufficiently high velocity, the domain pattern will be altered to produce at any point a net surface induction having a magnitude and direction that is a function of the magnetizing field intensity that existed at the instant the same point left the region of influence of the gap. For that reason, and because the field is concentrated at the edges of the gap, the trailing edge is the most effective region.

THE REPRODUCTION PROCESS:

The field existing around a magnetized section of tape brought into the scanning gap region of a reproducing head will be distorted by the presence of the high-permeability core. In a typical field distribution, as shown in Figure 5.2, the flux follows the inter-pole path of least reluctance. If the high-reluctance gap is in the path, the flux is forced to travel through the high-permeability core, thereby linking the coil windings.

In accordance with Faraday's law of magnetic induction, a voltage proportional to the rate of change of flux will be induced in the coil. Theoretically, for a given value of flux density, doubling the rate of change of flux by doubling the frequency will double the voltage. This ideal characteristic usually is expressed as a voltage increase at the rate of 6 db per octave. As a practical matter, the characteristic is modified in the high-frequency region if the recorded wavelength is short compared with the length of the gap, and in the low-frequency region if the recorded wavelength is long with respect to the length of the head.

RECORDING DEMAGNETIZATION (PARTIAL ERASURE):

A simplified graphical synthesis of the recording process is shown in Figure 5.3. Magnetically neutral recording tape is moved across the gap while pulses of current are passed through the coil of the recording head. The duration of each pulse is short with respect to the time required for an element of tape to move through the gap field, but long enough to provide the energy required to saturate the tape.

An initial pulse produces tape magnetization as shown at T1. After an elapsed time sufficient for the magnetized region tc move outside the influence of the gap, a second pulse having a polarity opposite to that of the first creates the field shown at T2. Before the second magnetized region has moved sufficiently far to be unaffected by the magnetizing field of the gap, a third pulse produces a field as shown at T3.

It may be seen that the second and third field overlap. Because they are in opposing directions, the effect is that of disturbance of the pattern recorded at T2. The resultant magnetization is shown at T4. As is apparent, the second pattern has been partially erased. In the extreme case, if the tape had not been moved at all before it was subjected to the reversed field, the direction of magnetization would have been reversed completely. This erasure phenomenon, sometimes referred to as "recording demagnetization", obviously is related to wavelength. It is, therefore, one of the limiting factors with respect to bandwidth.

In summing up the record and reproduce process, it can be shown mathematically that the record current at any instant equals:

RECORDING:

$$i = I \sin (wt) \quad \text{Eq. 1}$$

where t = seconds
I = maximum value
i = instantaneous value
$w = 2\pi f$
f = recording frequency

The instantaneous magnetizing force H will be proportional to the instantaneous current i. This will mean that the corresponding values of remanent flux ϕ_r and remanent induction Br will be proportional to the value of H. ϕ_r can now be substituted in the above formula for i.

$$\phi_r = K\, I \sin (wt) \quad \text{Eq. 2}$$

In order to represent ϕ_r along the tape it will be necessary to change time to distance along the H axis represented by X and consider wavelength instead of frequency. This relationship is shown below.

$$t = \frac{x}{s}$$

x = distance in inches
s = tape speed ips
λ = wavelength

$$f = \frac{s}{\lambda}$$

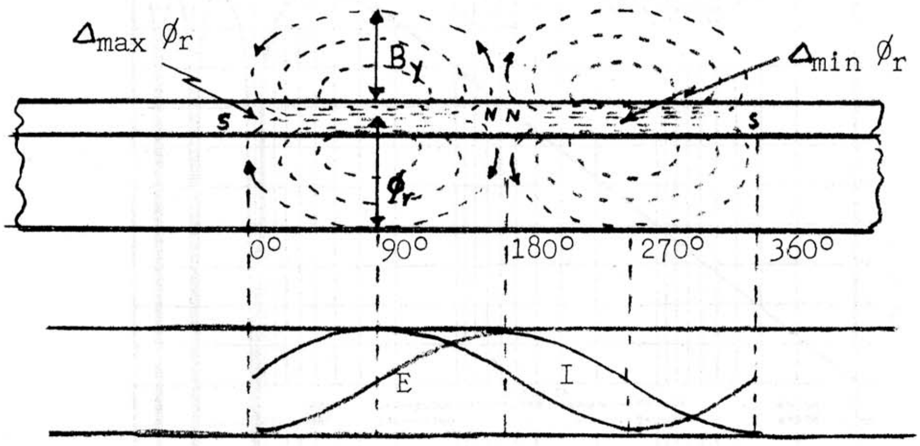
Therefore $ft = \dfrac{x}{\lambda}$

Substituting $\dfrac{x}{\lambda}$ for ft in the (wt) of Equation 2

$$\phi_r = K\, I \sin \left(2\pi \frac{x}{\lambda}\right) \quad \text{Eq. 3}$$

REPRODUCING:

The voltage induced in the reproduce head is dependent upon the flux lines which emerge from the surface of the tape and pass through the reproduce core and NOT the total flux ϕ_r in the tape. The emerging flux is designated as B_y and will be proportional to and vary with ϕ_r. At the center of the magnetic field inside the tape ϕ_r is maximum and its rate-of-change is zero. At the ends of the magnetic field the ϕ_r is zero but the rate-of-change is maximum.



Phase Shifting Between Record and Reproduce Processes

Figure 5.4

Thus, $$B_y = K \frac{\Delta \phi_r}{\Delta t} \qquad \text{Eq. 4}$$

By substituting for $\frac{\phi_r}{t}$, $$B_y = K\, I\, \frac{2\pi}{\lambda} \cos\left(\frac{2\pi X}{\lambda}\right) \qquad \text{Eq. 5}$$

Since the instantaneous voltage developed in the reproduce head is proportional to the number of lines cut in unit time, then

$$e = K\, B_y\, s = K\, I\, s\, \frac{2\pi}{\lambda} \cos \frac{2\pi x}{\lambda} \qquad \text{Eq. 6}$$

In more conventional form

$$e = K\, I\, f \cos (wt) \qquad \text{Eq. 7}$$

These formulae have shown that

1. The output voltage is proportional to record current.

2. Playback frequency is the same as record frequency (if tape speed the same).

3. The change from sine to cosine terms indicates a phase difference of 90° between output voltage and recording current for corresponding points along the tape.

4. Output voltage is proportional to frequency.

5. For a constant level of recording current surface induction and output voltage increase linearly with frequency (where head losses, etc., are taken into account).

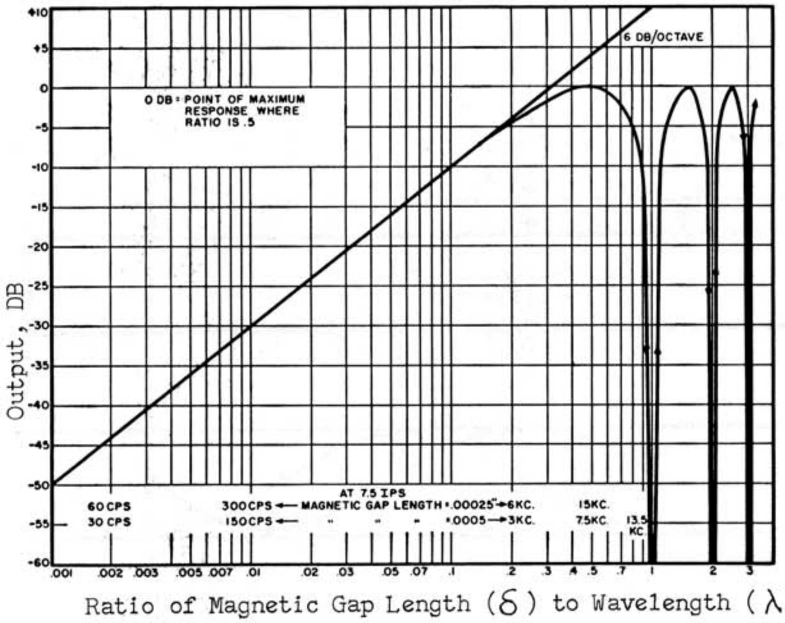


Figure 5.5 Theoretical Response of Perfect Playback Head Maximum Output Occurs When Magnetic Gap Length Equals One-half the Wavelength. 20 Log Sin $(180^o \times \frac{\delta}{\lambda})$

BIAS AND DIRECT RECORDING:

In the mathematical considerations just made, for convenience the B_r - H curve was assumed to be linear. This is not true.

Therefore, when deciding what area of Br - H curve to use for recording purposes there are two apparent choices:

1. In immediate area of origin of Br - H curve, and
2. In portion of curve between the instep and knee.

In Figure 5.6 a signal is placed so that the operating point is at the origin of the B_r - H curve.

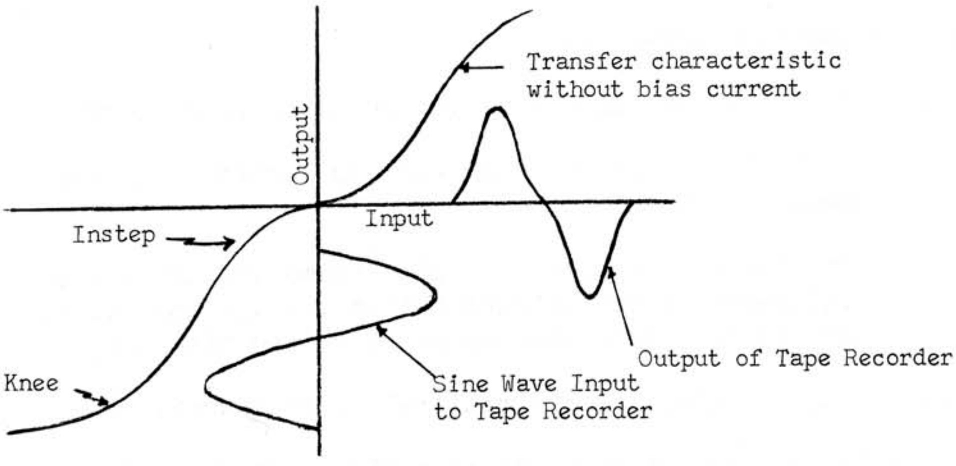


Figure 5.6 Input - Output Transfer Characteristic Without Bias

This is where curve is linear but where a restriction on maximum operating level invariably results in poor signal/noise ratio. The output is rich in odd harmonic distortion.

The second portion of the curve as shown in Figure 5.7 between A and B most nearly approaches a straight line. Thus, if a fixed dc bias was used (fixed magnetizing force) to bring the operating point to midway between A and B, the signals would be much improved.

A fixed dc bias invariably introduces noise and even order distortion. The answer to this problem was to introduce an ac bias.

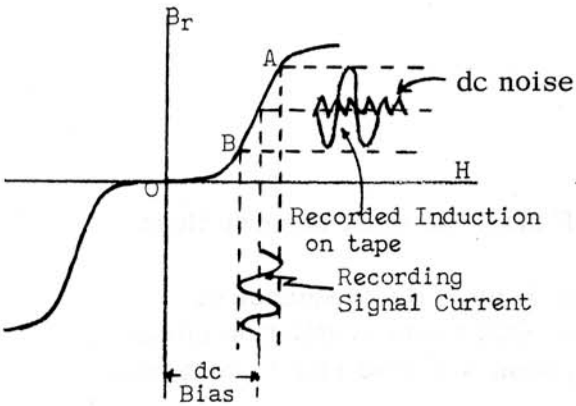


Figure 5.7 Application of DC Bias

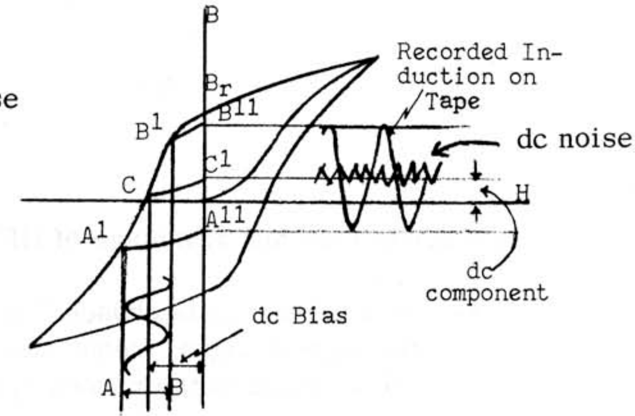


Figure 5.8 Alternate Method of Applying DC Bias. Tape Premagnetized to Saturation before Signal Applied with DC Bias.

The ac bias can be thought of as providing the same function as water color painting. The water is used as a carrier for the paint. When the painting dries no water is left, yet the painted picture remains.

With ac bias the recording signals are superimposed on an alternating magnetizing field, the amplitude and frequency of which is much greater than that of the signal.

The results show a vast improvement in every respect. (Figure 5.9) The sensitivity is increased,as also the maximum undistorted output level. The absence of dc components in magnetizing force assures a low noise level.

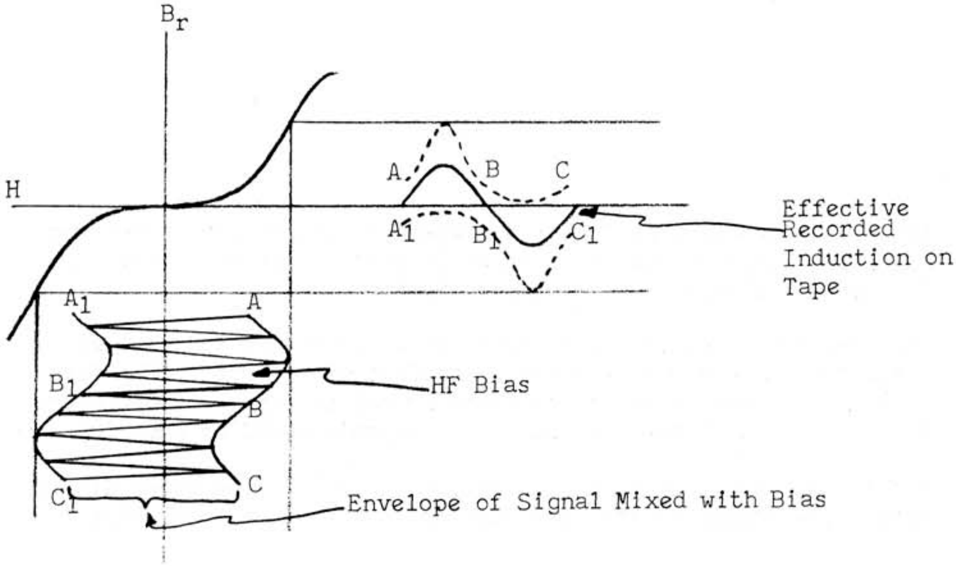


Figure 5.9 Application of AC Bias

In order to take full advantage of HF bias, we must observe that:

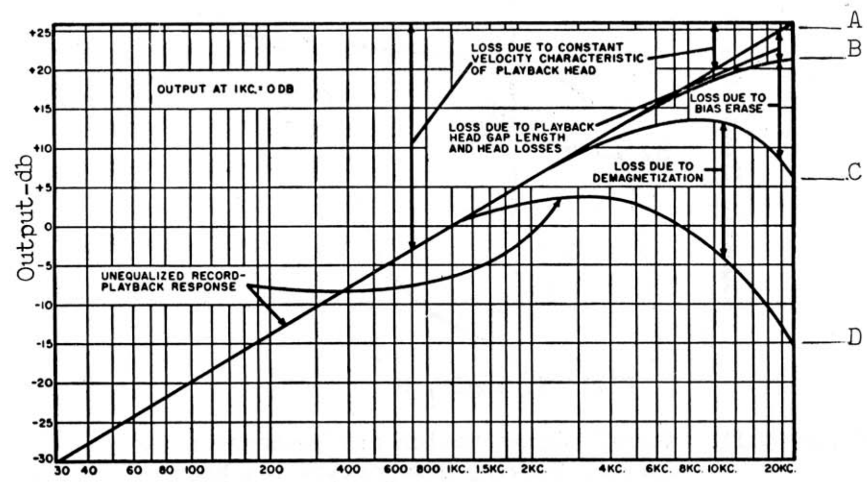
1. Bias frequency is chosen 3 to 5 times the frequency of the highest signal frequency. Otherwise modulation effects taking place in the record system will give rise to beat notes.

2. No even harmonics shall be permitted in the HF bias waveform. Even harmonics will make positive and negative peaks of bias current unequalized, thus leaving a dc component of magnetization on tape, causing noise.

3. Optimum bias current selected. One which represents the best balance of low distortion, extended high frequency response, and high output.

RECORDING/REPRODUCING LOSSES:

The basic reproducing characteristic is a linear one rising 6db/octave. In Figure 5.10, Curve A shows the ideal curve, and B, the typical curve of the conventional recorder. The drop-off of the D curve is a combination of tape and head losses, without any equalization being introduced into the system. These losses are discussed below.



Speed of Tape 7.5 IPS

Figure 5.10

DEMAGNETIZATION:

When a material is magnetized by exposure to an applied field, poles generated by magnetization create a new field in opposition to the applied field. The true field, acting upon any point, is the vector sum of the applied and the opposing fields. Because the resultant field is less than the applied field, the opposing field is referred to as a DEMAGNETIZING FIELD and the net effect is called DEMAGNETIZATION.

The poles of a very long bar magnet are separated so far that the demagnetizing field is negligible and the resultant field strength is determined primarily by the intensity of magnetization. As the magnet is shortened, the demagnetizing field becomes increasingly effective in reducing the magnetic induction. Very short magnets are almost completely demagnetized, although it is obvious that some magnetic induction must be present to create the magnetizing field.

If a short rod of magnetically-neutral ferro-magnetic material is placed in a magnetic field, and the field strength is increased from zero to a high intensity in one direction and then reduced to zero. The variation of magnetic induction with respect to field strength will be typically as shown in Figure 5.11.

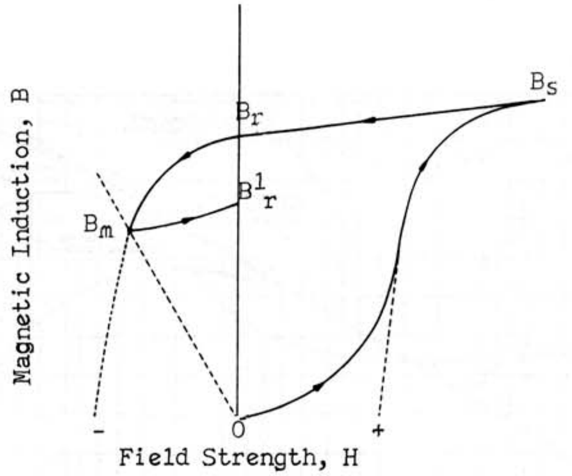


Figure 5.11 Normal Magnetization and Hysteresis Curves, Showing Demagnetization and Partial Recovery

Commencing at the origin, the induction will increase along the NORMAL MAGNETIZATION CURVE until the point of SATURATION INDUCTION B_s, is reached. As the applied field strength is reduced, the induction will decrease slowly, crossing the neutral field-strength axis at the point of RESIDUAL INDUCTION, B_r. At that point the applied field will be equal and in opposition to the demagnetizing field. The point of MAXIMUM INDUCTION, B_m, will be reached after the applied field reaches zero. The negative-sign field strength corresponding to the point of maximum induction is the DEMAGNETIZING-FIELD STRENGTH.

If the demagnetizing field is removed by the presence of a high-permeability shunt, such as when magnetized tape comes in contact with the pole pieces of the magnetic head during reproduction, the induction is increased along the path between B_m and $B1_r$. The loss of induction represented by the distance between B_r and $B1_r$ is irrecoverable.

The angle existing betwen the vertical line through the origin and the dashed "shearing" line, extending from the origin through B_m, is a function of the shape factor of the magnet. The angle approaches zero as the length-to-diameterratio is increased. The shearing line, therefore, becomes vertical for very long magnets, and the point of residual induction, B_r, becomes coincident with that of maximum induction, B_m. With very short magnets, the irrecoverable loss of induction becomes excessive.

Demagnetization becomes an important factor in determining the bandwidth of a magnetic tape recorder when the wavelength of the signal recorded on the tape becomes less than approximately 20 mils. A partial recovery of the short-wavelength loss caused by this effect occurs in the reproduction process.

LINEARITY:

An alternating signal current flowing in the winding of the recording head will create an alternating magnetizing field. The direction of the field will be determined by the direction of current flow, and the field intensity will be proportional to the magnitude of the current. The resulting magnetic induction produced in the tape is not, however, a linear function of the current because of hysteresis in the coating material. The transfer characteristic can be linearized by mixing an ac biasing current with the signal current.

The variation of magnetic induction with respect to field strength, starting with the medium in a neutral state, is along the normal magnetization curve between the origin and B_s in Figure 5.11. The inflection resulting from hysteresis effects can be removed by the process of <u>anhysteretic magnetization</u>.

Anhysteretic, or <u>ideal</u> magnetization, is effected by superimposing on the magnetizing field an ac field varying symmetrically to saturation in alternate directions; then gradually reducing the amplitude of the ac field to zero. The relationship between a normal and an ideal magnetization curve is shown in Figure 5.12.

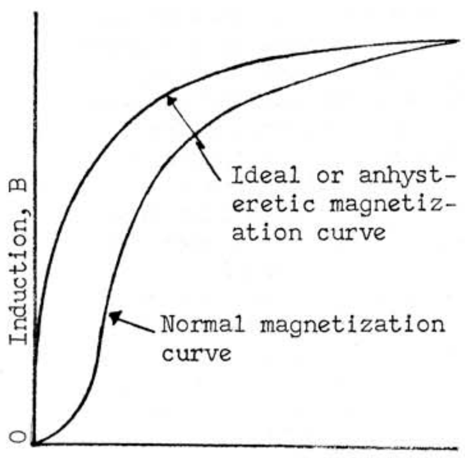


Figure 5.12 Typical Normal and Ideal Magnetization Curves

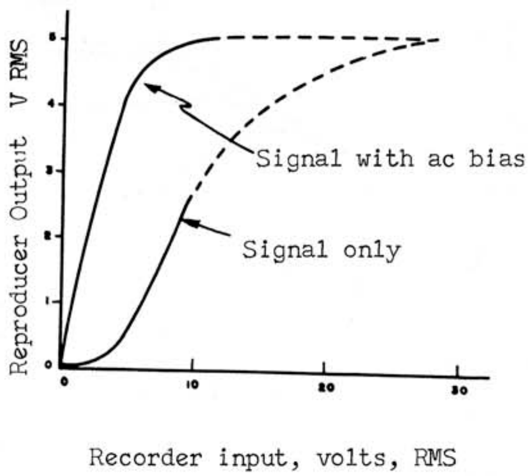


Figure 5. 13 Voltage-Transfer Characteristics of a Magnetic Tape Recorder, with and without AC Bias. (3 mil Wavelength)

Figure 5.13 represents the transfer characteristics of a magnetic tape recorder operating with and without ac bias. It is obvious by comparison of Figures 5.12 and 5.13 that a close correlation exists between anhysteretic magnetization and the linearizing effects of ac bias in magnetic recording. The fundamental difference is that the magnetizing field remains constant during the decay of the ac field in the ideal case, whereas, both fields decay simultaneously in the recording process. Experimental evidence has shown the results to be similar, except that the magnetization obtained in the recording process is always less than that obtained by the anhysteretic method.

The shape of the amplitude-transfer characteristic obtained with ac bias is such that the harmonic distortion introduced is odd-order, with the third harmonic predominant. The long wavelength signal amplitude resulting in the generation of one percent third harmonic distortion is commonly chosen as the normal operating level, and will be used as a reference level for the purpose of this discussion.

The bandwidth-limiting effects of demagnetization and of erasure, or recording demagnetization, are evident in Figure 5.14, which shows the amplitude-transfer characteristics of a magnetic tape recorder for four recorded wavelengths typically encountered in practice. The decreasing maximum output values, as the wavelength is shortened, are representative of the lowered values of residual induction resulting from demagnetization. Recording demagnetization is indicated in the negative slope regions of the 1-mil and 1/2-mil curves.

The 60-mil curve is representative of long wavelength recordings, which approach the value of maximum induction for the tape. Tape saturation is somewhat indeterminate because of the gradual approach to the limiting value. A practical and convenient point of reference for "saturation" is the operating level that results in a 6 db compression, or departure from linearity. It may be seen, by applying that convention to Figure 5.14, that the useful upper limit of dynamic range decreases with wavelengths less than the order of 10 mils. As shown, an increase in recorder input level actually causes a decrease in output level beyond the maxima at short wavelengths.

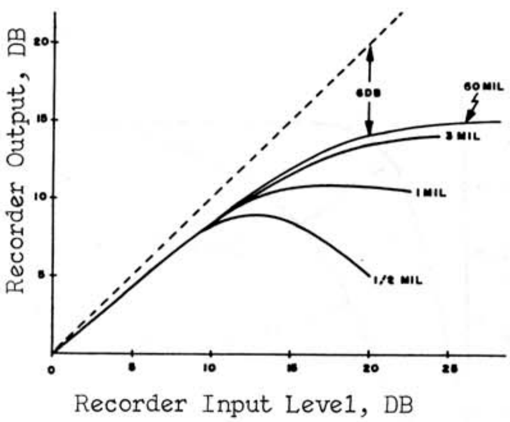
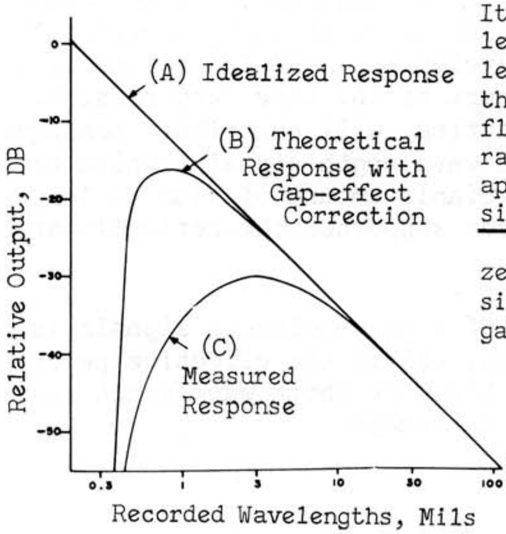


Figure 5.14 Amplitude-Transfer Characteristics of a Magnetic Tape Recorder for Four Typical Wavelengths. (Reference-Normal Level)

BASIC REPRODUCTION BANDWIDTH LIMITATIONS:

Bandwidth limitations imposed upon the system by the recording process affect only the short wavelength signals. Because the reproduction process depends upon the rate of change of the flux linking the coil windings of the reproducing head, any factors which change the ratio of flux in the tape to that linking the coil will affect the reproduced voltage.

Dimensional factors related to wavelength that modify the idealized 6 db per octave characteristic include the gap effect, spacing loss, and the effective length of the tape contacting region of the reproducing head pole pieces. The thickness of medium also is a determining factor.



It is well known that as the wavelength approaches the effective length of the gap, the ratio of the flux linking the coil to the flux in the tape decreases. The rate of decrease is described approximately by the periodic function $\frac{\sin x}{x}$. This function reduces to zero when the wavelength of the signal is equal to the length of the gap.

Figure 5.15 Idealized and Theoretical Wavelength Response Compared with Measured Response of a Typical Reproducing Head. (Effective Gap length, 0.35 mils)

Figure 5.15 shows the wavelength response of a typical reproducing head (C) compared with the idealized 6 db per octave characteristic (A) and the theoretical curve (B), resulting from modification of the idealized characteristic by the gap effect. Curve (C) represents the relative output resulting from constant current recording at normal level. The difference between curves (B) and (C) is the net effect of all losses other than those caused by the gap effect.

SPACING LOSS:

It has been demonstrated theoretically and experimentally that the introduction of a space between the reproducing head and the tape decreases the reproduced signal output proportionally, by approximately 55 db per wavelength of separation.

The seriousness of this form of loss can be seen when it is realized that a separation of only 50 microinches at a wavelength of 1/2 mil will result in a signal loss of 5.5 db. The inherent surface roughness of tape - the average particle size being 25 microinches - and the accumulation of foreign material on the head contribute to this loss.

THICKNESS LOSS:

Thickness of the medium is another controlling factor in the wavelength response. The spacing loss introduced at short wavelengths is much greater than that at long wavelengths. It is reasonable to conclude, therefore, that only the part of the tape very close to the head, near the surface of the coating, will contribute measurably to the reproduced amplitude of short wavelength signals, while the entire coating will furnish an appreciable amount of flux at long wavelengths. This conclusion has been supported theoretically and by an experimental evidence.

Practically, the amplitude response of long wavelength signals is proportional to the coating thickness, within the effective penetration limit of the recording field, while the very short wavelength response is substantially independent of the thickness.

LONG WAVELENGTH LOSSES:

As the recorded wavelength approaches the length of the head, the scanning gap becomes a decreasing factor in determining the ratio of flux in the tape to that linking the coil. Instead, leakage flux paths develop between the tape and the pole pieces. The effect is comparable to the introduction of a new gap having a length equal to that of the head, and the resulting periodic variations in the amplitude response are usually referred to as "head bumps". Contributions of the gap in the transition region interfere constructively and destructively with the leakage flux. Because they occur in the useful part of the wavelength response, particularly at the higher tape speeds, head bumps are a determining factor of the low frequency limit.

When the wavelength becomes greater than the length of the head, the reluctance of the flux return path increases because it includes the increasingly large airgaps between the tape and the pole pieces. The amplitude response falls rapidly thereafter, decreasing at the rate of 18 db per octave. Practically, this effect usually occurs at a lower frequency than is useful from a signal-to-noise ratio standpoint.

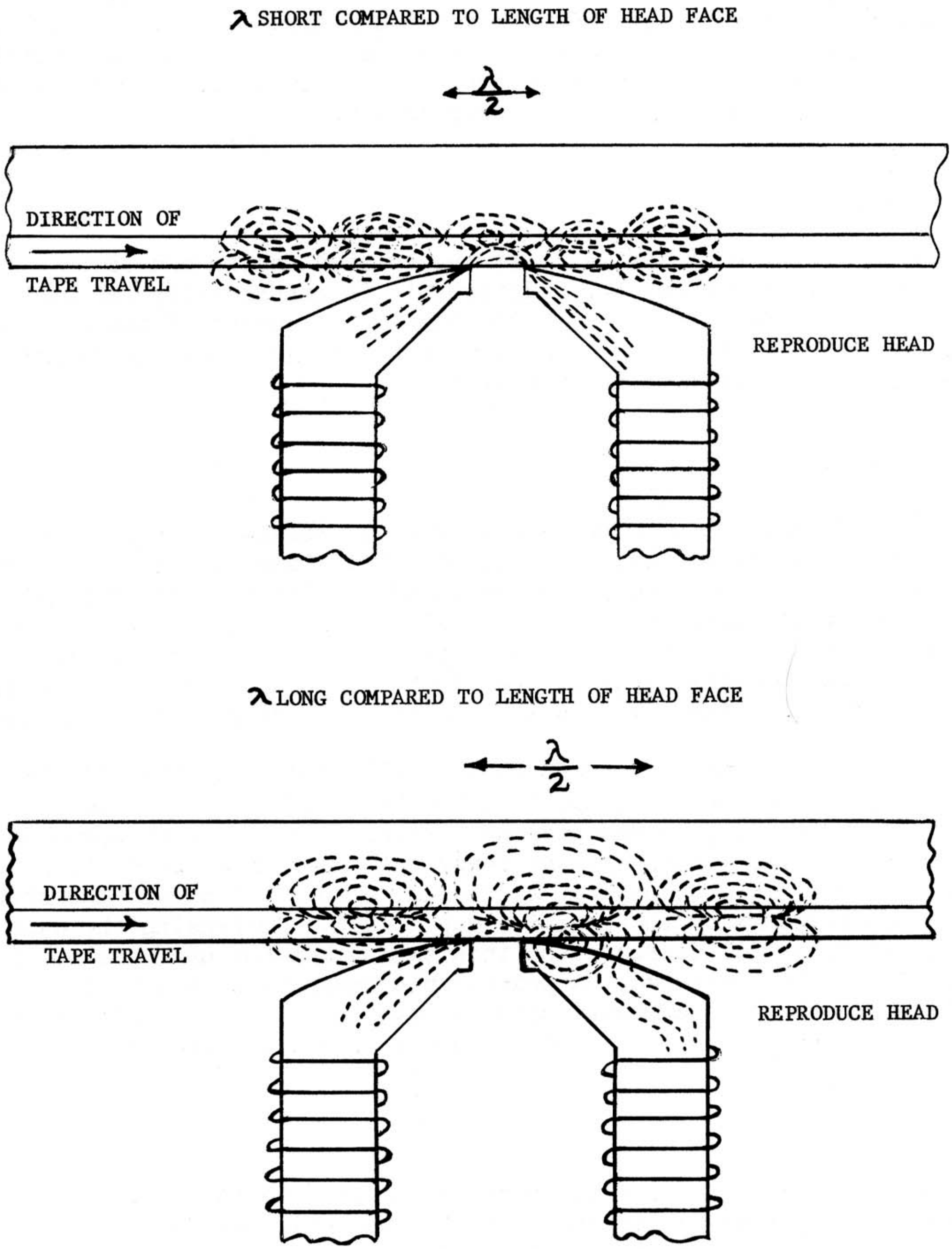


Fig. 5-16. Long Wavelength Losses

OTHER LOSSES:

The losses related to wavelength include those caused by demagnetization, recording demagnetization, gap effect, spacing, thickness, and the long wavelength phenomena. Losses related to frequency are also present, to a degree determined by the construction and physical characteristics of the components involved. Eddy current and hysteresis losses introduce frequency-sensitive effects in both the recording and the reproducing heads. These losses are relatively small and stable in nature, and can either be neglected or easily compensated.

DYNAMIC RANGE:

The dynamic range of a system is always related to the bandwidth. In a magnetic tape recorder, the relationship is particularly important because the dynamic range is not the same in all parts of the pass band. It is important to the user of such a device to know how the dynamic range varies with respect to frequency.

SIGNAL LEVEL:

The design of the reproducing head is closely related to the bandwidth requirements of the system. The reproduced voltage is a function of the wavelength of the recorded signal, and is directly proportional to the number of turns in the coil winding. It is desirable, of course, to recover the signal at as high a voltage level as is practicable, to reduce the susceptibility of the input circuits to noise currents. This suggests that the coil should consist of as many turns as possible.

The limit to the number of turns in the coil is usually reached when the inductance becomes so large that the head, with its associated circuit capacitances, becomes resonant within the frequency pass band. Increasing the gap length will decrease the inductance, permitting the addition of more turns, but will simultaneously decrease the short wavelength response by the modified gap effect. Decreasing the depth of the gap also will decrease the inductance, but the useful life of the head will be shortened. A practicable compromise results in excellent performance at wavelengths as short as 0.5 mil, which corresponds to a frequency of 2 kc per inch per second of tape velocity.

NOISE:

The IRE Standard definition of noise, as applied to sound recording and reproducing systems, is "any output power which tends to interfere with the utilization of the applied signals, except for output signals which consist of harmonics and subharmonics of the input signals, intermodulation products, and flutter or wow".

System noise is "the noise output which arises within or is generated by the system or any of its components, including the medium".

Equipment noise: It is noteworthy that system noise includes not only the equipment noise output, which is contributed by the elements of the system - hum, tube noise, etc. - but also "noise that can be specifically ascribed to the medium". Usually, the equipment noise in a well designed magnetic tape recorder is sufficiently low in magnitude so that the system noise is determined primarily by that of the medium.

The amplitude of the signal voltage at the terminals of the reproducing head is of the order of 50 uv at low frequencies normally encountered in magnetic recording practice. For effective utilization of signals of that level, careful design and construction techniques must be observed to minimize the equivalent input noise of the reproducing amplifier. As an example, at the 60 cps power frequency, and equivalent input noise of approximately 0.5 uv must be realized if a 40 db signal-to-noise ratio is required.

Medium noise: The coating of modern magnetic tape is a heterogeneous mixture of needle-shaped magnetic particles having an average length of 0.7 micron - approximately 25 microinches - and a shape factor of roughly 7:1. Nearly one-half of the total volume of the coating is occupied by the binder material. A normal reproducing head scanning gap bears a dimensional relationship to the coating such that six to ten particles, lying end to end, are scanned in the longitudinal direction.

If the distribution of magnetic material within the binder is not uniform, the particles tend to become arranged in clumps, and the result is an irregular surface as well as a coating having non-uniform magnetic properties. Medium noise is generally ascribed to these properties and to other physical and dimensional characteristics of the medium.

Random orientation of the magnetic domains is fundamental to the neutral state of magnetization. The net field produced in the gap region of the reproducing head is the resultant of the contributions of many domains. Because of the random domain orientation and non-uniform distribution of the particles, the net field varies in a random manner, and a noise voltage will be induced in the coil of the reproducing head when the medium is in motion.

It has been stated, and the statement supported by experimental evidence, that the noise spectra of oxide tapes show a form that is approximately the same as the head response, leading to the conclusion that magnetic tape noise is white noise if the tape is in the fully-demagnetized condition.

Modulation noise: A phenomenon which occurs in magnetic recording - although it is not unique to the process - is the generation of modulation noise. By definition, it exists only in the presence of a signal and is a function of the instantaneous amplitude of the recorded signal.

The exact mechanisms involved in the generation of modulation noise are the subjects of many theories and some disagreement. It is generally agreed, however, that the prinicipal sources can be divided into three categories:

> Amplitude modulation of the signal caused by variations in the uniformity of the physical and magnetic properties of the medium;
>
> Amplitude modulation of the signal caused by variations in spacing between the medium and the head; and
>
> Frequency modulation of the signal caused by velocity changes of the medium.

It has been shown that the amplitude of the reproduced signal at long wavelengths is proportional to the thickness of the medium; i.e., a five percent variation in thickness will cause a five percent change in amplitude. If the coating is 0.5 mil thick, a tolerance of five percent would permit a deviation of only 25 microinches - the length of one particle of oxide. Surface roughness of the base material has a marked effect on the total thickness of the coating. For example, the amplitude modulation generated by modern plastic-base tapes is inherently much lower than that produced by the early paper-base tapes.

Uniformity of the magnetic properties is essential to uniform amplitude response, a factor that requires all particles to have similar characteristics, in addition to the requirement that the dispersion must be uniform.

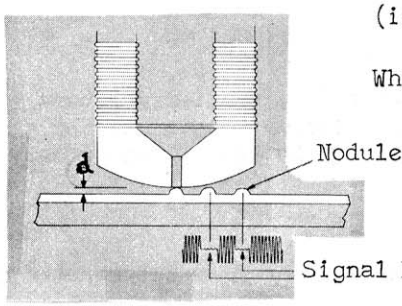
Modulation noise energy generated by variations in uniformity of the physical and magnetic properties of the medium is primarily confined to the low-frequency part of the spectrum, as might be expected by consideration of the relationship of the modulation process to the phenomenon of thickness loss.

Amplitude modulation of the signal by variations in the spacing between the medium and the head is the direct result of spacing loss. The noise energy is predominantly of a high-frequency nature because, for a given separation distance, the attenuation of short wave length signals is much greater than that of long wavelength signals.

Pressure between the tape and the head surface must be sufficient to maintain intimate contact under dynamic conditions. Intimate contact prescribes that both surfaces must be extremely smooth. Any degree of roughness of either surface will produce an effective separation, with consequent attenuation of high-frequency signals. Fortunately, the head surface is maintained in a state of high polish by the abrasion of the tape. Polishing techniques have been used to produce tapes having a greatly reduced surface roughness. More intimate contact with the head has been exhibited by some of these media, as evidenced by improved response at very short wavelengths. It is expected that the high-frequency modulation noise generated by the use of polished tapes will be lower in amplitude by a comparable degree.

The presence of dust or other foreign material, either in the coating or carried on the surface, will separate the tape from the head.

If the separation is great enough, the attenuation will be of such magnitude that a signal "dropout" will be produced.



Drop in Playback Level (in decibels) $\frac{55d}{\lambda}$ db

Where: d = Separation of tape from head

λ = Recorded wavelength on tape

Figure 5.17 Effect of Tape-Surface Imperfections

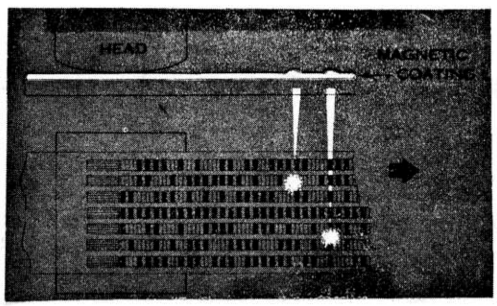


Figure 5.18 Effect of Drop-outs in Digital Recording

Frequency modulation of the signal by velocity variations of the medium is a source of noise commonly referred to as flutter. Modulation can occur in the recording process(in which case the generated sidebands are recorded on the tape), in the reproducing process (during which they are generated by the reproducing system), or in both processes.

SIGNAL-TO-NOISE RATIO:

A comparison of the system noise spectrum with the normal level signal output of the reproducing head, plotted with respect to frequency, provides an overall picture of the finite limits of the useful frequency pass band. The limits of a typical, unequalized, ac biased system, operation at a tape speed of 30 ips, are illustrated in Figure 5.19.

As shown, a wide band frequency response can be obtained before the signal-to-noise ratio starts to deteriorate. The amplitude transfer characteristic of the reproducing amplifier can be modified as required to produce the desired overall frequency response. The fundamental relationship of signal-to-noise will, of course, be unchanged by equalization.

CROSSTALK:

Interchannel signal interference in a multi-track tape recorder is a wavelength effect that is closely related to the physical spacing of the recorded tracks on the tape. Two primary sources contribute to crosstalk:

Inductive coupling of the leakage flux between the gaps of adjacent heads, and

Coupling of leakage flux from the signals recorded on one track into the head of another track.

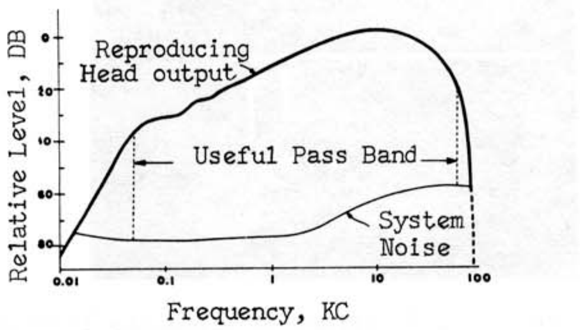


Figure 5.19 Unequalized System Noise and Reproducing-Head Output vs Frequency. Normal-Level, Constant-Current Recording at 30 IPS

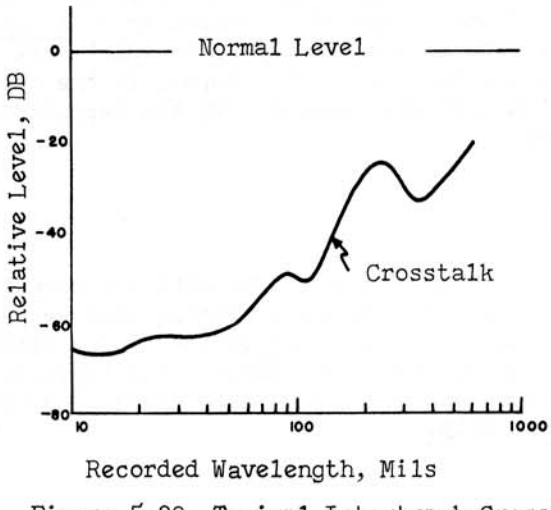


Figure 5.20 Typical Intertrack Crosstalk Between 50-Mil Tracks Spaced 70-Mils Center-to-Center

The degree of inductive coupling is a function of the effectiveness of the inter-track shielding, the physical spacing between tracks, and the signal current. This source of coupling can be minimized by carefully applied shielding techniques, so that tape crosstalk becomes the limiting factor.

Coupling of leakage flux from the tape is a function of the physical separation of the tracks and of the recorded signal level. It can be compared to the effect that causes the generation of "head bumps" at long wavelengths.

The most commonly used multitrack configuration is a series of 50-mil tracks separated 70 mils, center-to-center. As a usual practice, two stacks of heads are employed, with adjacent track heads placed in alternate stacks to minimize inductive coupling. With that arrangement, the crosstalk interference with respect to wavelength is typically as shown in Figure 5.20. In this example, the interfering signal was recorded at normal level, which level resulted in generation of a third harmonic content of one percent, with ac bias.

Either decreasing the intertrack spacing or placing all heads in line in one stack tends to degrade performance by increasing crosstalk.

TIME BASE DISTORTION:

Faithful reproduction of signals includes preservation of the time base of the original event. The reproduction must be synchronous in time with the recording. Deviation from synchronism with respect to time results in the accumulation of a time-base distortion called "time displacement error". A type of modulation noise, commonly known as "flutter", is a function of the rate of change of the time displacement.

In the strict interpretation, both are frequency modulation phenomena which behave in accordance with the classical theory. From a practical standpoint, however, the effects upon the reproduced signal can be generalized as follows:

> A recorded sinusoidal signal having a given frequency will be reproduced as a different frequency, without FM noise, if the system introduces time displacement error linearly with respect to time. An example would be the situation in which the recording and reproducing velocities were different, but uniform. If the system introduces flutter, and the net time displacement error is equal to zero, the signal will be reproduced at the correct frequency, but as the center frequency of an FM wave, accompanied by modulation noise components in the form of sidebands.

CHAPTER VI

THE TAPE TRANSPORT

INTRODUCTION:

Recording techniques have been known to mankind as far back as the Stone Age. At that time, stones were the recording media and manpower was the mechanism used to make recordings. During the past thousands of years various recording devices have been developed, perfected and used. In the twentieth century, scientists and engineers are perfecting the most modern recording techniques of all, that of "magnetic tape" recording. This new development has been applied to three major fields:

Audio ---
Magnetic recording and playback of audio frequencies in the range from 40 to 15,000 cps;

Video ---
Magnetic recording of sound and television signals to at least 4 mc;

Instrumentation or Data ---
Magnetic recording, storage, and reproduction of data of various phenomena from dc to 4 mc and beyond.

All three types of magnetic tape recording need a somewhat similar mechanism to transport the recording medium -- magnetic tape -- at a constant velocity across magnetic heads, which record or read-out magnetic information through various electronic devices. So it may be seen that a magnetic tape recording system consists of four major entities: a tape transport mechanism; electronic assemblies; magnetic heads; and the magnetic tape. All components are of equal importance, and when combined will result in a magnetic tape recorder/reproducer.

INSTRUMENTATION MAGNETIC TAPE RECORDING:

Magnetic tape recorders for instrumentation use are being developed in increasing numbers and can be found all over the world. Their military and commercial applications include data acquisition and storage, data analysis and reduction, machine and process programming and dynamic simulation. Numerous tape recorders have been developed for the above applications. The most typical are as follows:

1. General Laboratory Tape Recorder

   The "work horse" of magnetic tape recorders. In general, it is used in laboratories (however, many recorders are also adaptable for limited mobile use) to record scientific data in the range from dc to 500,000 cps. Maximum versatility for varying test requirements and a wide range of tape speed selection are major criteria.

2. Flight or Mobile Tape Recorder

   Compact and of light weight, designed for data recording under conditions encountered in airborne, shipboard and vehicular use. It must operate under conditions of severe shock, vibration and acceleration.

3. Digital Tape Recorder

   Typical characteristic is the starting and stopping of magnetic tape in an extremely short time, as fast as 2 milliseconds. It is used primarily for input-output and storage of data for digital computers or in related processing fields.

4. Seismic Recorders

   An important instrument for new oil field discoveries by obtaining optimum seismograms. For the short recording time, approximately 5 seconds required for each shot, a drum recorder with as many as 26 recording channels is frequently used.

5. Missile Recorder

   The midget of the family of instrumentation tape recorders. Very small, an extreme rigidity to shock, vibration and acceleration are the main requirements of such a recorder.

6. Loop Tape Recorder

   Used in continuous loop operation for repetitive examination of recorded data. Tape loops are adjustable to different lengths and of various delay times.

All these tape recorders have been designed within the past few years and are being used at various installations all over the world.

THE IDEAL TAPE TRANSPORT:

The tape transport moves tape over the heads. An ideal transport would move tape at perfectly constant speed, completely free from instantaneous and long-term speed changes. Every linear element of the tape would be moved always at the same speed as every other linear element and always at exactly the same position, both crosswise and vertically. Each time the same piece of tape were moved across the heads by the transport the time intervals between any sets of points on the tape would be exactly the same as before; and in order to achieve this exact time interval repetition (in spite of dimensional changes in the tape) the transport would automatically make suitable corrections. Thus the output timing would be identical to the input signal timing when the tape was recorded.

In this ideal machine tape tension at the heads would be perfectly constant from end to end of each reel and from pass to pass of entire reels. The tape to head contact would be identically the same at all times. Further-more, the transport would prevent deposition of oxide from tape to head, or automatically keep oxide removed.

This ideal tape transport would accelerate the tape instantaneously from standstill to constant specified speed and it would stop the tape instantaneously on command. Thus it would effectively use every linear inch of the tape.

This ideal transport would be capable of complete control at the transport or remote from it, either manually or in response to instructions from other machine sources. It would infallibly do what the operator desired in spite of miscellaneous aberrations in manipulating its controls. It would not stretch tape, break tape, nor wear its surface or edges; and while maintaining constant tension at the head it would program the tension onto the take-up reel in such fashion that the recorded reel could withstand wide temperature changes during shipment or storage without damage to the tape. It would do all these things throughout wide ranges of supply voltage and frequency. The transport would run indefinitely at any temperature, at any altitude and with any pollution of the atmosphere, and it would do this not only without operator attention but in spite of attention from incompetent operators!

Even the best tape transports fall short of these ideal specifications. There is still opportunity for improvement, but further incremental gain in any respect is becoming progressively more difficult and costly to achieve. The time has come when instrumentation tape recorder users can save themselves both cost and trouble by helping identify the optimum combination of system requirements and tape recorder capabilities. There have been many times when earlier and better user understanding would have enabled him to get better results at lower cost.

THE TAPE TRANSPORT MECHANISM:

1. GENERAL

The main purpose of a tape transport mechanism is to move magnetic tape at a constant linear velocity across magnetic heads with the least amount of disturbance to the tape.

The effort in designing a mechanism must be directed toward instantaneous and long term speed control of the recording medium. The instantaneous speed variations are known as flutter and their values are measured in percent peak-to-peak. Long term speed variations are introduced by inaccuracy of the drive system and are expressed in percent variations from the absolute selected speed. The problems of rotating a surface at uniform speed are generally recognized by engineers. A logical development of each major assembly, combined with numerous tests and a careful selection of components will result in a good instrumentation tape transport mechanism with reliable performance. The tape transport mechanisms may differ from each other in appearance, size, and operation, but basically the design principle is the same. The most important criteria of a tape transport mechanism are:

Reliability and Dependability

Reliability can be achieved only by keeping the configuration as simple as possible. In most instances, this is a very difficult part of the design. Careful testing of materials and parts under all desired conditions will aid in the achievement of reliability.

High Performance

The performance of an instrumentation tape transport mechanism is measured by the ability to transport the tape at the desired speed with no introduction of variables in the signal to be recorded and no disturbance of the tape.

Standardization and Compatibility With Other Equipment

During the past few years of instrumentation tape transport mechanism development, standards have been established for tape speed, tape width, reel size, track spacing, and other variables. Private industry and Government agencies are aware of the urgent need of a better form of standardization (which may come in the near future).

Low "Down Time" and Maintenance Cost

By designing a tape transport mechanism with access from the front or top of its installation and by using easily removable subassemblies, the "down time" and maintenance cost is minimized.

Flexibility for Modifications and Freedom From Obsolescence

The requirements of a user may change over a period of time so that modification may be desirable. By designing the tape transport mechanism with modular construction, completely independent major sub-assemblies, modifications can easily be accomplished. The analysis of design theories and problems in magnetic tape transport mechanism may be arbitrarily divided into five major independent sub-assemblies: a precision mounting plate assembly; a drive assembly; a reel or tape storage assembly; a control assembly; and a frame assembly.

2. PRECISION MOUNTING PLATE ASSEMBLY

This assembly contains all components which are closely associated with the direct driving of the tape across the heads. They are, capstan and pinch roller assembly, turnaround idler, tape guides, and magnetic heads which must be located perpendicular to a reference surface. The assembly can be a plate or casting of stable material with a reference plane established by milling or grinding a surface flat and parallel to extremely close tolerances. Such a reference surface holds the accumulation of tolerances to a minimum, and assures proper longitudinal guiding in relation to the tracks of the magnetic head. Thus, uniformity will be retained from mechanism to mechanism, and interchangeability of tape is guaranteed.

Tape Drive Systems

Four basic tape drive assemblies are most commonly used on various instrumentation tape transports.

In the open-loop drive (See Fig. 6-1), the tape passes over a small inertia idler, where it is stabilized before passing over the heads. The long unsupported length of tape is pulled at a constant speed by the capstan in the direction of capstan rotation when the single pinch roller is engaged. It is applicable where no great accuracy is required and is also simple in design and operation. The most serious defect is high-frequency flutter introduced by the long unsupported length of tape. If the distance of the unsupported tape is approximately 6 inches, the resonant frequency will be between 2000 and 3500 cps, which can be very objectionable in a wide band FM recording system.

The zero-loop tape drive (See Fig. 6-2) is an approach which makes the tape loop as close as possible to zero by feeding the tape directly around the capstan. The magnetic heads are arranged to contact the tape on the capstan, and the unsupported length of tape is reduced to a minimum. The heads are spring loaded against the tape and compensate for any tolerance variations. The tape can be clamped against the rotating capstan by pinch rollers or by vacuum applied through the capstan, thus holding the tape firmly to the surface of the capstan. Such arrangements are limited for various head mounting combinations; they are however, being used very successfully for special applications.

Most instrumentation tape transport mechanisms have adopted the closed-loop drive (See Fig. 6-3). A similar method has been used by the motion picture industry for many years. The unsupported tape is kept as short as physically possible in order to eliminate any vertical displacement across the heads. This tape drive consists of a constant speed capstan with pinch rollers clamping the tape to both sides. The record and reproduce heads are mounted on opposite sides of the tape. It eliminates practically all tape tension variations in supply and takeup systems. The tape between the unsupported points is very short and, therefore, the displacement is kept to a minimum, resulting in a great improvement in high frequency flutter. Any looseness of tape during the start will be removed by a continuous, minute slip on the capstan. The study of a closed-loop drive is best approached with a high speed camera.

A video type drive system is shown in Figure 6-4. By this method a relatively high speed is attained between heads and tape by rotating the heads at high velocity across the tape. The tape is pulled by the capstan at just the right speed to permit close

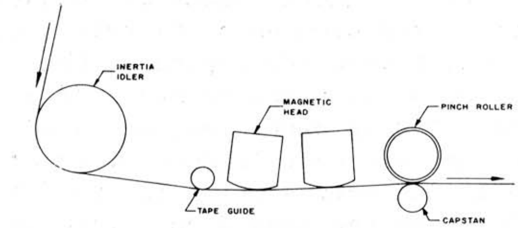


Fig. 6-1. Open-Loop Drive System

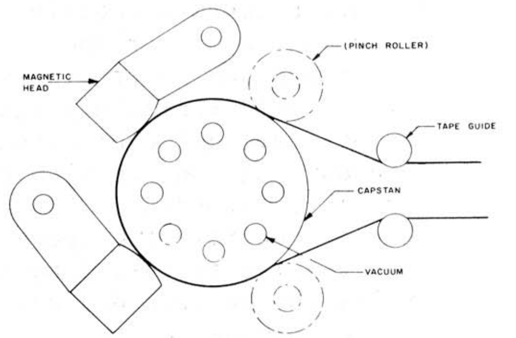


Fig. 6-2. Zero-Loop Tape Drive System

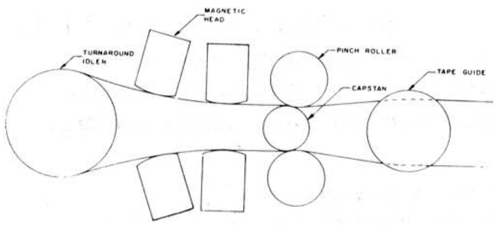


Fig. 6-3. Closed-Loop Tape Drive System

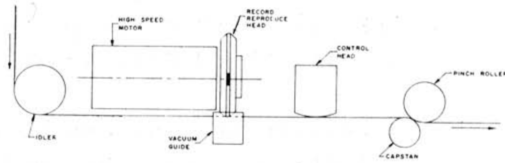


Fig. 6-4. Video Type Drive System

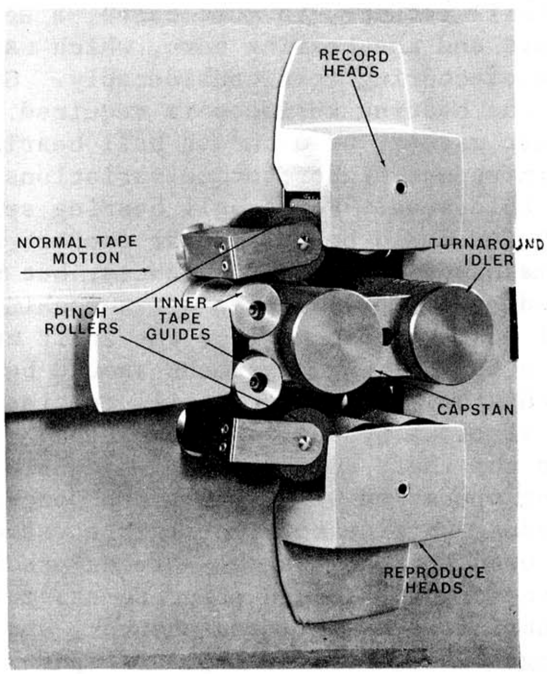


Fig. 6-5. Capstan and Pinch Roller Assembly

spacing of the transverse recorded tracks yet avoid any interference between successive tracks. Four magnetic heads are mounted in quadrature around the periphery of a drum which is driven at high speed and is precisely synchronized with the capstan. The tape is held intimately against the rotating heads by means of a concave tape guide or vacuum system. During operation, one head is always in contact with the tape. A reference or control track signal is recorded on the tape and during reproduction is used to control the relative speed of the capstan and head drum. Through such a drive system, frequencies into the region of several megacycles can be recorded and reproduced at reasonable tape speeds on tape of nominal width.

Capstan and Pinch Roller Assembly

This assembly is mounted directly on the precision mounting plate and is one of the most vital parts of the transport mechanism. A typical capstan and pinch roller assembly is shown in Figure 6-5. Proper capstan design is a challenge to good engineering.

The capstan bearing can be designed as a sleeve, ball, or combination sleeve-ball bearing with no radial or axial play. The main objective in the design of a capstan is to keep eccentricity as low as practicable and the torque variations to a minimum. Both parameters will introduce flutter directly with no possibility of filtering any undesired frequencies.

Sleeve bearing will require, in most cases, a selective fit between the shaft and the bearing bore, which naturally will increase the manufacturing cost considerably. Continuous lubrication of the bearing surfaces is required. For axial loading, a thrust washer, or a thrust ball bearing, must be used, and may introduce higher torque variations than are desired or may be tolerated. For a ball bearing selection, the most logical choice would be an angular contact bearing. The parts must be machined to close tolerances, but no selections will be required. To avoid warpage after machining, a very stable material must be selected for the shaft and housing. In most cases, a stabilizing treatment should be employed after rough machining. But, there still remains the danger of introducing stresses by the last operation -- grinding. The bearing fits on the shaft and housing. Adequate cooling during the grinding operation will reduce the concentration of stresses and reduce warpage to a negligible value. A dust-free room should be used for the assembly of the precision bearings and all parts thoroughly cleaned prior to assembly. The bearings are pre-lubricated by the manufacturer, and no lubrication is needed during the life of the tape transport. To eliminate eccentricity of the capstan shaft, a final grind with the capstan assembled will reduce the total run-out to approximately 0.00002" to 0.00008" T.I.R. The selection of the capstan diameter must be a compromise between the desired flywheel inertia, which will increase with a smaller diameter, and the easier machinability and greater stability achieved by using a greater diameter.

The capstan flywheel can be mounted directly or coupled through a mechanical filter unit to the capstan; the choice depends on the selection of the drive assembly. A mechanical filter unit will not be essential when a pure drive assembly is designed, and the disturbance frequency kept to a minimum.

Pinch rollers have two major functions -- to clamp the tape against the turning capstan -- and to isolate any tension variations in the tape caused by the supply and take-up units. The rollers are usually rubber-coated pulleys mounted on bearings; the thickness and the diameter of the rubber depends upon the physical size and arrangement of the pinch roller. The best isolation results, and the amount of pinch required to meter the tape without slipping on the capstan, are most easily established by extensive tests. It is not advisable to select a rubber with a durometer hardness of less than 50 nor higher than 90. Close tolerances are required to engage the pinch roller parallel to the capstan, or the tape will be forced away from the established line to the heads, with resultant poor frequency response and non-interchangeability of tapes from one tape transport to another. A self-equalizing pinch roller engagement mechanism will assure equal pressure to the pinch roller on the capstan, if a closed loop tape drive is used. The mechanism may be actuated by a solenoid or spring force.

Turnaround Idler

Another important component of a closed loop tape drive is the turnaround idler, which will reverse the direction of the tape approximately 180 degrees. (The same order of accuracy is required -- as for the capstan.) Should a flutter peak be caused by the natural frequency of the turnaround idler, it can be changed to a more acceptable band width by changing the inertia of the idler. A typical problem of turnaround idler disturbance experienced during engineering is shown in Fig. 6-6.

The three flutter peaks are disturbing natural frequencies introduced by turnaround idlers of various inertia. The flutter was taken at half-octave band widths at a tape speed of 60 ips. It can be observed that by shifting the natural frequency of the idler to a more acceptable band the flutter peak can be moved outside the frequency band of interest to the user. A typical cumulative flutter spectrum is shown in Fig. 6-7. By designing a very stable turnaround idler, such disturbance frequencies can be eliminated.

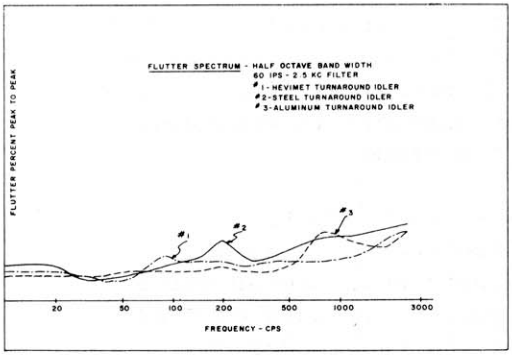


Fig. 6-6. Turnaround Idler Disturbance

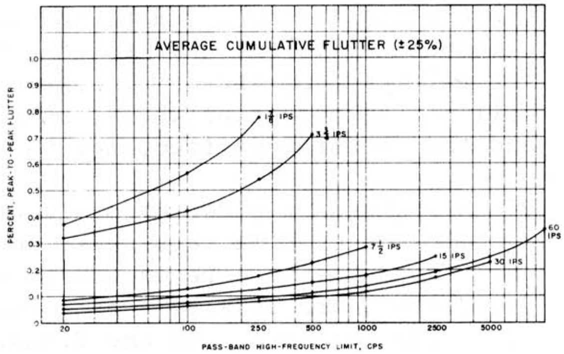


Fig. 6-7. Cumulative Flutter Curve

Guides

Guides locate the magnetic tape with respect to elevation of the heads. Two different types of guides are used in general -- either stationary or roller guides. The wrap-around of the tape around stationary guides should not exceed 60°, or the sliding friction and the power required to overcome this friction will increase excessively and will result in undesired tension variations. To eliminate the abrasive action of the tape, a stationary tape guide must be hardened to approximately Rockwell C-55. To assure the best guiding of the tape across the heads, the main guide should be placed as close as possible to the capstan. Adding guides inside a closed loop is not advisable and will cause ineffective guiding. Once the tape is clamped tightly against the capstan, by the pinch roller, no guiding in the loop is possible. The tape would be forced against the flanges of the guides and the edges of the tape might be damaged. Only precision-mounted components will assure perfect guiding inside the closed loop.

Roller guides may have sleeve or ball bearings, and the housings can be made of aluminum because no abrasive action occurs. By decreasing the wrap of the tape below 45°, the guiding of the tape will be less effective. It is not advisable to locate two roller guides closely together because the tape will tend to climb on the flanges of the guides, and may break if the guides are not perfectly parallel to each other.

3. DRIVE ASSEMBLY

The drive assembly is the connecting link between the drive source -- in general, a hysteresis synchronous motor -- and the capstan flywheel. Various methods of reduction drives have been applied, and some successfully used -- belt drives: (flat, round, timing or "V" belts, etc.) -- gear: (spur, helical, worm, bevel, or others) -- friction drives: (rubber-coated pulley against a steel pulley, and others). Speed accuracies of $\pm$ .1% to $\pm$ .3% are required for most instrumentation tape transport mechanisms. It would be unwise to design, knowingly, a drive system with great variations in speed, and record data under such conditions. Elaborate servo control and electronic compensation systems to correct for tape speed errors introduced during recording may have to be used for accurate reproduction of the information. It may be called the "negative approach" in designing a drive assembly for a tape transport.

Gears and timing belts (toothed rubber belts) require a mechanical filter unit between the reduction drive and the capstan. Such a filter unit must be tuned to a lower natural frequency than the lowest rotational frequency of the reduction drive. Round belts, "V" belts, or "O" rings are not suitable for accurate speed translation. By using such devices there is no accurate method possible to establish the neutral axis of pulley and belt within the limits required for an instrumentation tape transport mechanism. Rubber rim drives have been used very successfully in instrumentation tape recorders in the past. Unfortunately, the effective diameter of the rubber-coated pulleys must be established by a trial method in the assembly of the tape transport. Flat belts have been used most successfully. They will transfer ratios from one pulley to another within the accuracy required for an instrumentation tape transport mechanism. Unfortunately very few manufacturers make belts of uniform thickness (variations not to exceed .001 inch). A few years ago, some machines very successfully utilized nylon garment straps sewed together to the proper length. By cutting a crown on the pulley, the belt will be guided in the center of the rim. The amount of belt tension and effective ratio must be established very carefully by tests. Tension variations will change the ratio from one shaft to another.

If a drive assembly is developed as a completely independent unit, as shown in Fig. 6-8, any other assembly may be designed and installed, if modifications are required, with the least amount of cost to the user. (Speeds of a standard drive assembly usually range from 1 7/8 to 60 ips in a speed ratio of 1:2:4:8:16:32). A modification is shown in Fig. 6-9, in which a 120 in./sec. search speed has been designed into a standard drive assembly. A second capstan is placed on the precision mounting plate and is driven by a coaxial shaft placed through the standard drive pulley. A completely different drive assembly, with two speeds is also available. Rapid switching from one speed to the other in this assembly is accomplished by a built-in magnetic clutch. Modified drive assemblies are continuously being designed and incorporated in standard tape transport mechanisms. These modifications demonstrate how the design of independent major sub-assemblies can make a mechanism flexible to meet various requirements.

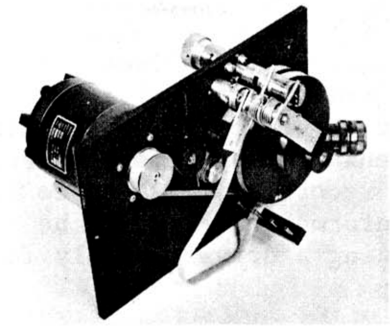


Fig. 6-8. Drive Assembly

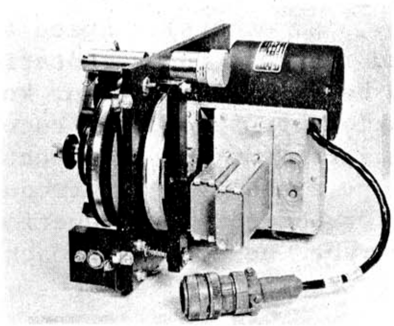


Fig. 6-9. Drive Assembly With Two Capstans

4. REEL OR TAPE STORAGE ASSEMBLY

These assemblies store the magnetic tape in a suitable manner on either side of the capstan on reels or other means of storage. They also supply a constant hold-back tension to the tape before entering the drive system, and take up the tape metered out by the capstan on the other side. To accomplish such a task, various methods have been developed in the past, shown in Figs. 6-10 through 6-14.

Figure 6-10 shows a simple container, with tensions supplied by the weight of the tape. It is applicable when a limited amount of tape is used; for example, computer-type recorders.

Figures 6-11 and 6-12 show a method in which tape is stored on reels with supply and hold-back tensions applied by vacuum. Vacuum switches or photo-electric cells control the amount of tape in the columns through a series of resistors which control the power to the torque motors. These types of tape storage assemblies are used for high speed computer tape transport mechanisms with starting times as low as 2 milliseconds.

Figure 6-13 shows sensing arms for controlling the tape tension in a very simple manner. The arm follows the radius of the tape reel, and transfers this movement directly to a brake drum; or it may be connected to a potentiometer which controls the power to the torque motors. Such a system is only an approximation of a constant hold-back tension system, because the sensing rate is too slow to follow reel eccentricities, etc. Sensing arms do not allow simple modifications of tape transport mechanisms for different reel sizes.

Fig. 6-14 illustrates a servo-controlled tape tensioning assembly. Such a servo can be designed mechanically, electronically, or as a combination of the two. It will supply constant hold-back tension under all conditions within very small increments.

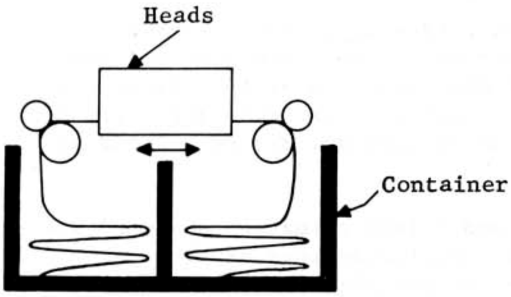


Fig. 6-10. Simple Tape Storage Container

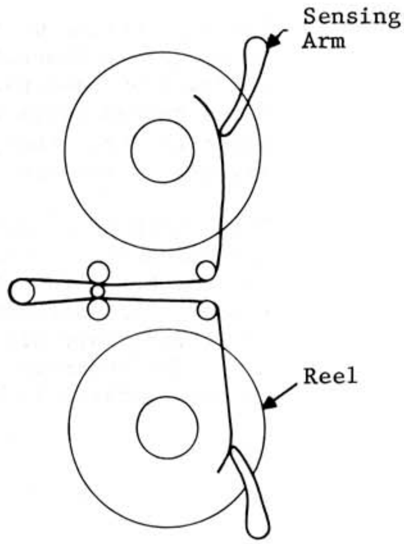


Fig. 6-13. Simple Tape Sensing System

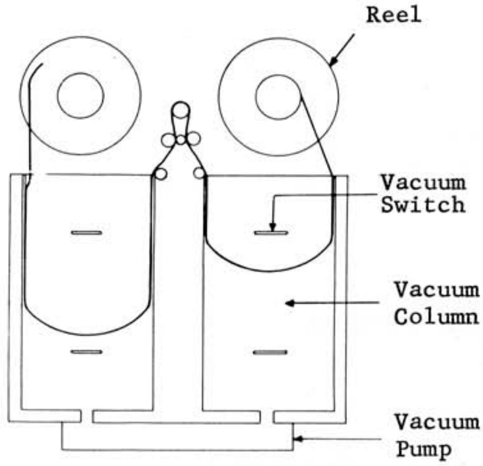


Fig. 6-11. Tape Storage Container Using Vacuum

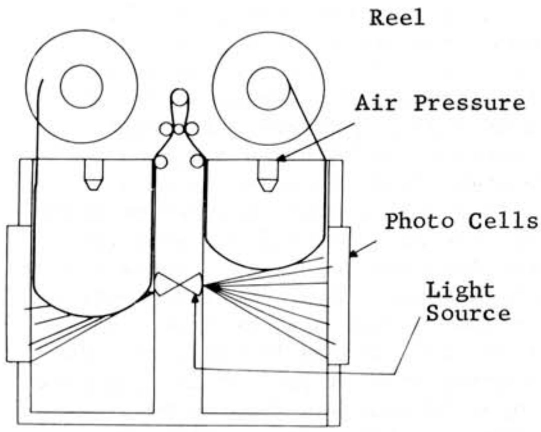


Fig. 6-12. Tape Storage Container Using Vacuum and Photoelectric Sensing

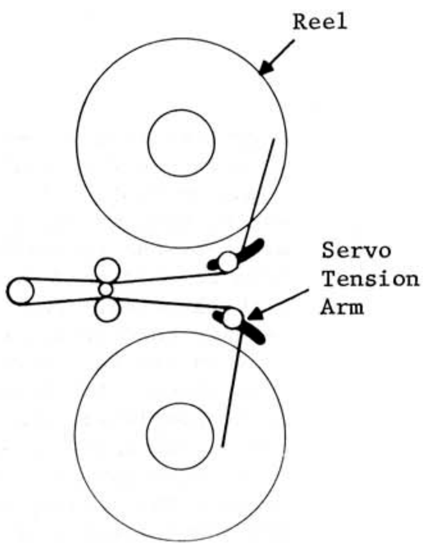


Fig. 6-14. Servo Tape Sensing System

The most common method of storing tape is by means of reels. Standards in dimension have been established and are, in general, followed by most tape recorder manufacturers in the United States. Non-standard reels will cause problems in interchangeability of tapes to other tape transport mechanisms and are very disadvantageous to the user.

Tape width of ¼-inch, ½-inch, and 1-inch primarily are used for instrumentation data recording. Reel diameters are 10½-inch and 14-inch with a tendency to use the 14-inch reel more frequently. By increasing the tape width and reel diameter, the rate of feeding tape off and on reels becomes more and more of a problem. The increase of inertia vs. tape width and reel diameter is demonstrated in Fig. 6-15.

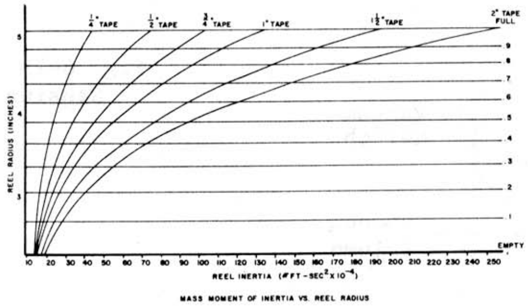


Fig. 6-15. Mass Moment of Inertia vs. Reel Radius

A typical reel assembly with a mechanical servo tension control is shown in Fig. 6-16. All components are mounted on a flat, aluminum plate. A turntable is pressed directly on the motor shaft, eliminating a special bearing design. The outside diameter of the turntable is used as a brake surface. A brake-band is stretched around this surface, and kept in place by a brake housing. One end of the brake-band is connected to a release-band strap; the other, to a bell crank. A pivot-mounted tension arm is connected to the release strap and actuates the brake-band. Two brake solenoids are used solely to release the brake completely during some operational modes. A safety switch can be contacted by the tension arm to shut off the tape transport mechanism in the event of any malfunction in the tape. To relieve the braking surface of heavy duty, a part of the required hold-back tension is supplied by a tension and damping housing. For a better understanding of the operation, assume that the

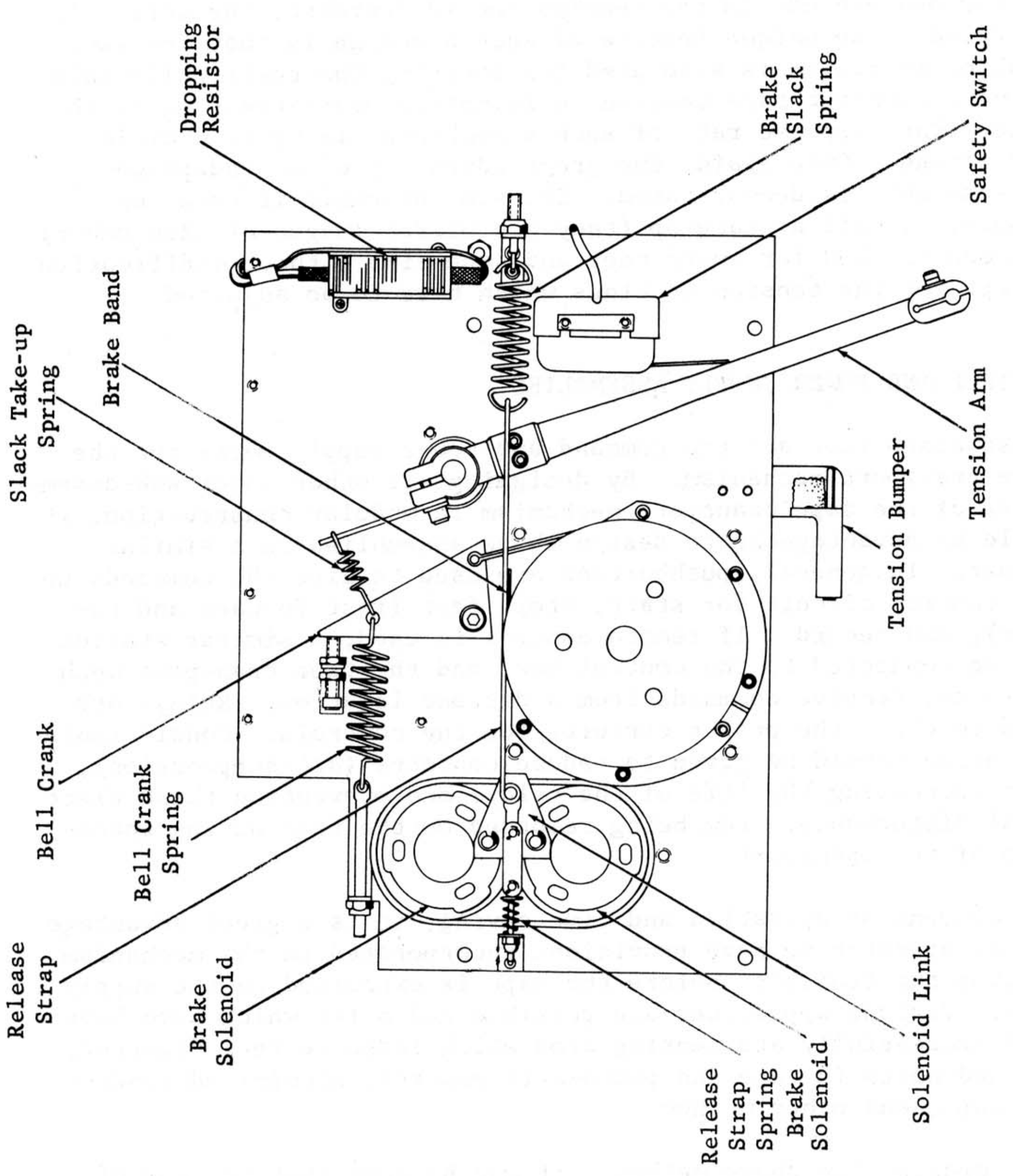


Fig. 6-16. Typical Reel Assembly With Mechanical Servo Tension Control

assembly is installed in the tape transport mechanism and the machine is operating in "normal drive". The capstan will pull tape from the supply reel over a tape guide attached to the end of the tension arm, and meter the tape to the take-up reel. The arm will move against the tensioner assembly and simultaneously partially release the brake. The arm will settle to a position of pre-adjusted tension. Should the tension on the arm increase by a small amount, the braking force will be decreased by release of the brake-band. If the tension should decrease, the action is reversed. The unique feature of such a design is that the same braking mechanism is also used for stopping the reels while maintaining constant tape tension, eliminating over-stressing of the tape. The response rate of such a mechanism is up to 4 cycles per second. Once again, the great advantage of an independent sub-assembly is demonstrated. The same assembly is used for supply, as well as take-up (they are mirror images of each other) and can be used for every reel and tape size without modification, except for the tension settings which have to be adjusted.

CONTROL AND POWER SUPPLY ASSEMBLIES

These assemblies are the command and power supply units for the tape transport mechanism. By designing all other major sub-assemblies of the tape transport mechanism in modular construction, it would be advantageous to design these assemblies in a similar manner. In general, pushbuttons are used to give the commands to the control circuit for start, stop, fast (fast forward and rewind), and record. If remote control is used, a similar station can be connected to the control box, and the tape transport mechanism can receive commands from a distant location. Relays are used to close the proper circuits for the controls. Considerable attention should be given to reduce contacts (arc-suppression), thus increasing the life of the relays and preventing these electrical disturbances from being recorded on the tape during operation of the transport.

For convenient operation and time saving, it is a great advantage to the operator to have provisions incorporated in the mechanism to stop the transport before the tape is exhausted on the supply reel. Various approaches are possible and a few which have been used successfully are sensing arms which indicate reel diameter, punched holes in tape and photo-cell control, aluminized conductive tape, and other methods.

By comparing the above methods, it may be seen that the use of aluminized conductive tape is a very simple method of stopping a tape transport at either end of the reel, or at any intermittent position, by providing an electrical connection between a pair of contacts. This conductive tape can be spliced at any point on the reel. Two contacts may easily be established by using existing tape guides and insulating one of them. This feature can also be used to close a contact of an external relay to control other equipment. By using modular construction, additional control units can be added and inter-connected by cables. Because of the nature of this equipment, standard AN connectors should be used as they are readily available and safe (no contacts exposed) in handling.

The power supply must distribute power to all sources, and usually ac and dc power are required for motors and solenoids. By fusing the ac and dc voltages, the mechanism will be protected if an electrical failure should occur. The power for the drive motor may be obtained from a 60-cycle line, or a precision frequency source. Rectifiers will supply dc power if the input voltage is ac. 115 volts ac 60 cps is generally used as primary power for laboratory tape transport mechanism.

6. FRAME ASSEMBLY

The frame assembly combines all major sub-assemblies into one integrated mechanism. Misaligning of components attached to the frame must be avoided by designing the frame in a rigid and stable manner. For easy accessibility, the entire tape transport mechanism can be hinge-mounted, or designed with a quickly-removable cover. Either of these methods will enable the operator to reach all components from the front, and will decrease the problem of locating a suitable place for its use in laboratories, trucks, airplanes, etc. The front, or top, of a tape transport should be protected by adequate covers or doors to prevent the accumulation of dust, or other foreign particles from the magnetic heads and tape.

QUALITY OF A TAPE TRANSPORT MECHANISM:

The quality of a tape transport mechanism is best checked by flutter tests and by its capability in handling magnetic tape as gently as possible. Any imperfections such as eccentricity of rotating drive components, torque variations in bearings, or tape tension variations will be indicated as flutter.

THE IMP WITH THREE NAMES

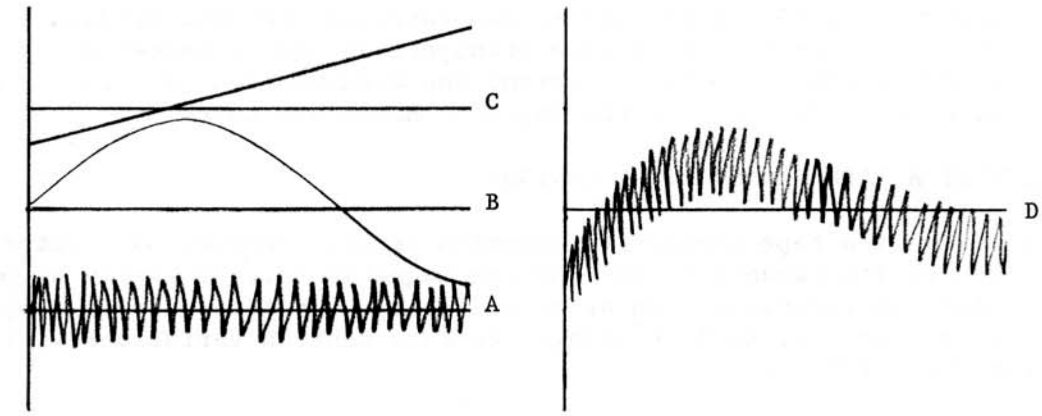
FLUTTER, WOW AND DRIFT:

The wealth of work and planning going into a modern tape recorder is fantastic. So are the results. Yet, as with all good things, there's a liability -- an elusive imp called Flutter. His dishonorable history goes back to the early days of sound recording when he manifested himself as a surging sound from record players afflicted with cyclic speed errors. This "wow, wow, wow" was only one of his guises. Higher frequency errors caused a fluttery tremolo and audiophiles soon coined the term "flutter and wow" to characterize the imp. He's still with us, but research is lessening his depredations.

DEFINITIONS:

Generally, the human ear can only distinguish the unpleasant sounds due to speed variations below 200 cps (cycles per second) and of a strength above 0.2% of the amplitude of the reproduced sound. Thus, in early recording, flutter and wow only meant the audible results of speed error.

When magnetically taped music entered the home, both terms retained their original meanings. With the coming of tape to instrumentation, where frequencies ouside the ear's range are used, flutter and wow began to mean something else. Now they no longer identify the audible result of speed error, but apply to the error itself. Flutter generally denotes speed variations occurring at frequencies above 10 cps; wow includes those between 0.1 and 10 cps; a new term, "drift", tags those with frequencies below 0.1 cps. All refer to changing tape velocity as it passes the record or reproduce heads. In addition, lengthening or shortening of the tape, lateral tape motions and variations in the intimacy of head-tape contact result in data errors which are charged to flutter, wow or drift. Obviously, for good recording, these deviations must be eliminated or compensated for -- a tricky process, since all of the effects of Flutter, Wow and Drift result in tape speed errors -- the more correct definition for these deviations would be "Instantaneous tape speed errors".



A Flutter B Wow C Drift D Combined Effects of A, B and C

Fig. 6-17. Relationship Between Flutter, Wow and Drift

THE IMP'S MISDEEDS:

Flutter, wow and drift (hereafter lumped under the term "flutter", and more correctly called "instantaneous tape speed errors") are enemies of instrumentation recording because they introduce errors. Events occur with an exact time interval between them and tape monitoring these events will show other than the true intervals if it moves past the record or reproduce heads at varying speeds. This is called "time-base distortion" and is intolerable in precision instrumentation work.

Also, a sudden spurt -- or slowing down -- causes an unwanted increase or decrease in the reproduced signal frequency. Amplitude of the signal suffers, too, since it is a function of the rate at which the magnetic flux from the tape changes as it passes the head. Thus, flutter causes misleading increases or decreases in signal strength. In addition, the imp plays a special kind of hob with high signal frequencies. Suddenly accelerated tape ups the reproduced frequency of a previously recorded signal. But in recording, acceleration causes the recorded wavelength on the tape to become longer, causing a lowered frequency on reproduction.

Flutter has different effects on the various types of magnetic recording because each type uses a different method of coding data:

Direct Recording. Direct application of the data signal to the tape with no intermediate coding makes it subject to all the adverse effects mentioned above.

FM-Carrier Recording. Here, flutter itself modulates the carrier and is added to the data to appear as output noise or "hash". Time-base distortion also occurs.

Pulse Duration Modulation Recording. Effects vary, depending upon the flutter frequency and number of multiplexed data channels. If the error is of a frequency such that it occurs simultaneously with each signal pulse, a gross mistake in output signal value is possible along with time alteration.

Digital Recording. In this method of recording, information is contained in the pulse-patterns on the tape. Flutter does not affect signal accuracy, but the reproduced signals must appear within narrow time limits. Due to flutter, the signals may get out of their time slots and this requires extending the limits. This lowers the "packing density" of the recorder, which means that fewer characters per inch can be written on the tape.

CAUSES OF FLUTTER, WOW AND DRIFT:

Erratic tape movement and tensions are caused by two types of recorder components:

Rotating Parts

These include the rotating idlers, flywheels, shafts and the capstan. With these parts, eccentricity or some other cause of irregular torque is one troublemaker. It not only gives the tape uneven motion but applys longitudinal tension oscillations.

Another bugaboo -- resonant frequency -- is determined by the mass and resilience of each part. Should that part be influenced by some chance vibration near its own resonant frequency, it suffers cyclic speed variations of increasing force which affect tape movement. Many things excite this resonance and providing against them is an engineer's nightmare.

The "noise" of imperfect or dirty bearings and external vibrations transmitted through the transport frame are among them. Even "stiction" due to tape movement over non-rotating parts can be the culprit.

Part of the defense against these troubles is keeping the resonant frequency such that it will not be excited by ambient frequencies. This is a feat in itself, because the tape, drive system and reels all have many possible speeds, and other turning components have at least two! The problem is compounded by reels and turn-tables having varying instantaneous velocities and the fact that the mass of a reel keeps changing from the beginning of a run to the end. In short, the assembled tape transport is a collection of rotating parts, each operating at any of a number of speeds which must be kept as far as possible from the point of resonance.

Fixed Parts

These are the heads, tape guides, etc., over which the tape slides to create "stiction". This causes longitudinal oscillations in the tape, their frequency depending upon the length and mass of tape undergoing the excitation.

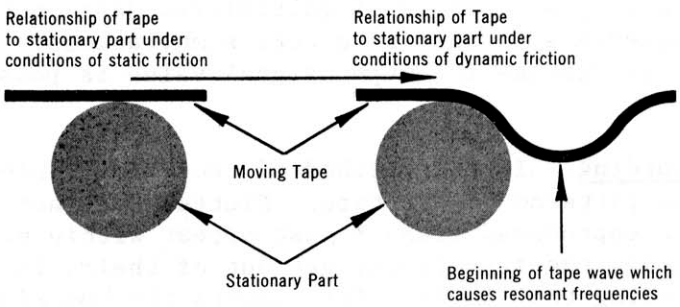


Fig. 6-18. Transition From Static to Dynamic Friction Showing Resonant Frequencies That Can be Created

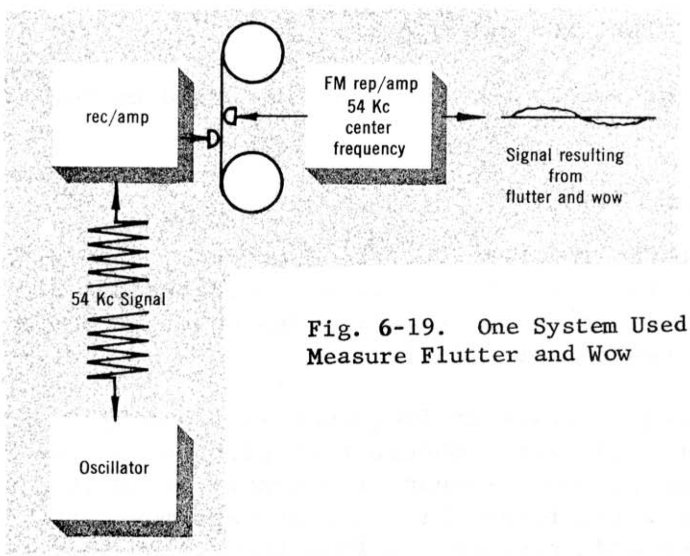


Fig. 6-19. One System Used to Measure Flutter and Wow

TAKING THE MEASURE OF THE IMP:

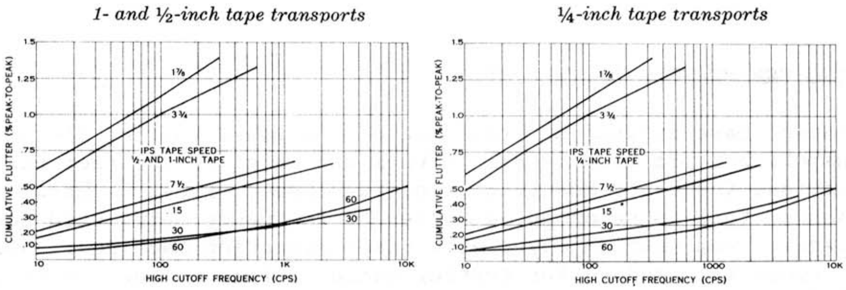
Engineers have evolved numerous methods of measuring the flutter in a transport but practically all of them are based on the fact that speed variations frequency-modulate a recorded tone. First, a known steady-frequency tone is recorded, a different frequency being used for audio recorders from that employed with instrumentation machines. 3000 cycles per second is standard for testing audio equipment, but the frequencies used for instrumentation work vary according to the bandwidth being used and the tape speed. For instance, one Ampex Instrumentation system employs 54,000 cps when evaluating wide-band systems at 60 inches-per-second tape speed.

Various precision instruments are used to produce these constant frequencies for recording on the magnetic tape. If (ideally) there were no flutter during this recording process, each cycle of the tone as impressed on the tape would have the same length and the frequency would not vary. However, cycle length increases and frequency falls whenever there is a tape speedup; slowdowns cause shorter wavelengths and raised frequencies.

When the recorded signal is reproduced, the flutter error (from both the record and reproduce phases) appears as pitch variations in the originally steady tone. So, the varying tone played back is actually a frequency-modulated signal, the original tone being the carrier signal whose frequency is changed by the flutter. Demodulation of the reproduced signal leaves the flutter-induced frequency error as a residual which is actually an electrical signal directly proportional to the speed variations. Because upon demodulation frequency change becomes a proportional amplitude change, flutter may be measured in terms of amplitude (volts). This fact permits the troublemaker to be visibly demonstrated by means of an oscilloscope or one of the other instruments mentioned below, measured and even graphed.

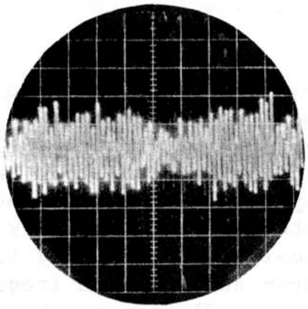
Both the oscilloscope and the oscillograph are used for "visualizing" and measuring the hobgoblin. Each shows "peak-to-peak" amplitudes and the oscilloscope can be calibrated in terms of percent flutter as explained below. For permanent records, the face of this instrument may be photographed as illustrated. The oscillograph, of course, pens or photographs the flutter spikes on moving paper.

When it is desired to make a flutter graph such as the one shown, Fig. 6-20, the oscilloscope is used, Fig. 6-21. Usually, such a chart depicts the amplitudes of the flutter components having frequencies between dc and about 10,000 cps. They are expressed in cumulative peak-to-peak values, which means that the flutter signal is measured at its peak amplitude value, then expressed as a percentage of the maximum input voltage required to provide a fully modulated FM signal on the tape. The error signals are measured through a filter which excludes all frequencies above a certain point, this point being variable from about 10 cps to 10,000 cps. To begin taking readings, the filter passband is set at dc through 10 cps, and all error components in this range are evaluated as they appear on the 'scope which displays the algebraic sum of the components present in the passband.



These graphs show maximum cumulative flutter as exhibited by any tape-width or reel-size transport of the FR-100A series. These figures are the usual conservative specifications provided for Ampex equipment, and were purposely not made from selected or peaked-up units.

Fig. 6-20. Cumulative Flutter of the Ampex FR-100A Series



Picture of "peak-to-peak" flutter as it appears on the face of an oscilloscope. Upper spikes might represent tape speedups during playback and the consequent higher frequencies; lower spikes, speed reductions and reduced frequencies.

Fig. 6-21.

Keeping the lower level at dc, the upper cutoff limit is progressively increased, with measurements made at each step. Finally, a series of readings for frequencies up to 10,000 cps exists, so that each point on the graph shows the flutter at that particular cutoff frequency PLUS the accumulated sum from all the lower frequency ranges.

In addition to the oscilloscope and the oscillograph, several other indicating devices are used to evaluate the extent of the sprite's mischief. Because the type of device employed can have important effects on the accuracy of the measurements, confusion is sometimes encountered due to the difficulty of correlating and comparing the readings obtained on different instruments for the same flutter component. Depending upon the character of the ac-flutter signal, the error introduced by a given type of indicator may be surprisingly high -- and different from the error resulting from another type of instrument. For example, three popular instruments found in many laboratories are the Cathode Ray Oscilloscope (CRO), the Thermo-couple Ammeter and the AC Vacuum Tube Voltmeter (VTVM).

Any of these can be and are used to measure alternating current values, but they must be applied carefully because they respond differently to currents having different wave shapes. This fact is of great importance because flutter rarely generates a signal having a simple periodic waveform. Instead, it is commonly composed of many periodic waves, depending on the mechanical complexity and excellence of the tape-transport mechanism. Unfortunately, it is only when the wave has a simple sinusoidal shape that each of these three typical indicators will give the same reading as the others.

It is customary to describe the magnitude of an alternating current by its "effective" value or heating capability, which is called the "Root-Mean Square Value" or simply the rms. Most ac-measuring instruments, as described below, are calibrated to read the rms value of a sine wave, which is the form most frequently encountered.

Cathode Ray Oscilloscope

This instrument produces an indication by an electron beam that causes a bright spot to appear on a viewing screen. The spot can be moved up or down by changing the voltage applied to the input terminals of the 'scope. Because the deflection of the spot is exactly proportional to the voltage, the spot traces a path appearing as a bright line having a length exactly proportional to the peak-to-peak voltage. The sensitivity can be easily adjusted so that, for example, an inch-long path traced on the screen might represent a 2-volt peak-to-peak wave.

Because the deflection is caused by the total voltage applied to its input, the CRO is a "summing" device -- it performs simple arithmetic by adding the values of all applied signals at any given instant and produces a deflection proportional to their sum. For that reason, it has no inherent "waveform error", but gives true peak-to-peak values whether the waveform is simple or complex.

Thermocouple Ammeter

This type of device indicates the effective value of current by actually measuring the heat produced by current flowing through the instrument. It responds to, and reads out in terms of, the rms value, irrespective of the current waveform.

AC Vacuum Tube Voltmeters

Most common types of VTVM's contain amplifiers followed by rectifiers to convert the applied ac to dc for operating a coil-and-magnet D'Arsonval-type indicator. The dc rectifier output is proportional to the average value of the applied ac therefore the reading is proportional to the average current rather than to its effective heating or rms value.

Because there is an exact relationship between the average and the effective values of a sine wave, both the thermocouple and rectifier types of instruments can be calibrated to give identical readings for the same input -- so long as the input waveshape is sinusoidal, i.e., single frequency.

Although flutter signals are actually much more complex than those shown in the illustrated comparison of the readings obtained with the above indicator devices, three different waveshapes have been chosen to point up the magnitude of possible "waveform error" and its misleading results. The waves shown represent the fm-discriminator output caused by flutter, with the system adjusted so that 1 volt represents 1% flutter.

As shown in Fig. 6-22A in the comparison, all three instruments indicate as expected in the case of the simple sine wave:

The oscilloscope trace extends to 2 inches, indicating a 2% peak-to-peak flutter.

The thermocouple instrument responds to the 0.707-volt rms value of the 2-volt sine wave, indicating 0.707% rms flutter.

Responding to the average value of the wave, the VTVM correctly indicates 0.707% flutter -- as it was calibrated to do.

So, simple sine wave evaluation causes no difficulty. But the usual flutter signals are not simple, and complex-wave measurement can be a problem!!

For example, the simple sine-wave flutter signal could be modified by the addition of a second sine wave exactly three times as high in frequency as the original and with one-half its amplitude. This is called a "50% third-harmonic distortion". Depending on the relative timing or phasing of the two waves, the combined or summed waveshape could be that shown at either B or C of Figure 6-22 in the illustrated comparison. Note in B. that each of the three instruments has reacted to the complex wave to a different degree:

Flutter indicated on the oscilloscope screen increased only 8%, from 2% to 2.16% peak-to-peak.

The thermocouple device has responded to the increased heating effect of the complex wave over that of the simple wave by a figure of 11%, now reading 0.78% rms flutter.

With the averaging VTVM equipment, the complex wave has caused a 16% increase in the reading to 0.82% rms flutter!

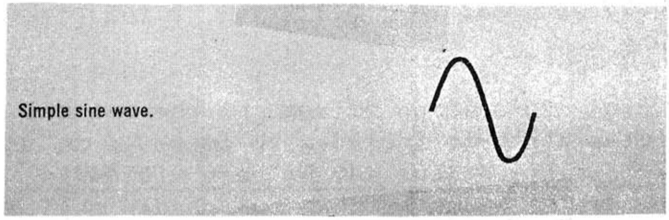
To be sure, the differences occurring in Figure 6-22B are small and probably not particularly serious, but look at what can occur if the phase relationship between the same flutter components changes as in Figure 6-22C.

Peak-to-peak amplitude on the oscilloscope screen has increased to 3%.

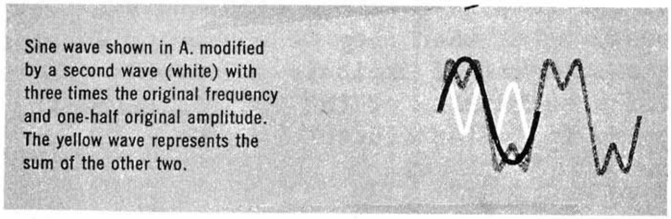
The rms value shown with the thermocouple instrument is unchanged from Figure 6-22B -- still 0.78% rms flutter.

A decrease to 0.64 rms flutter reading is obtained with the VTVM because of the lower average level!

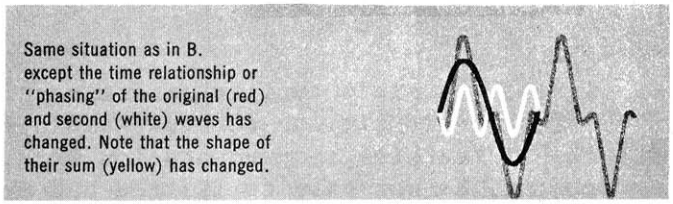
A



B



C



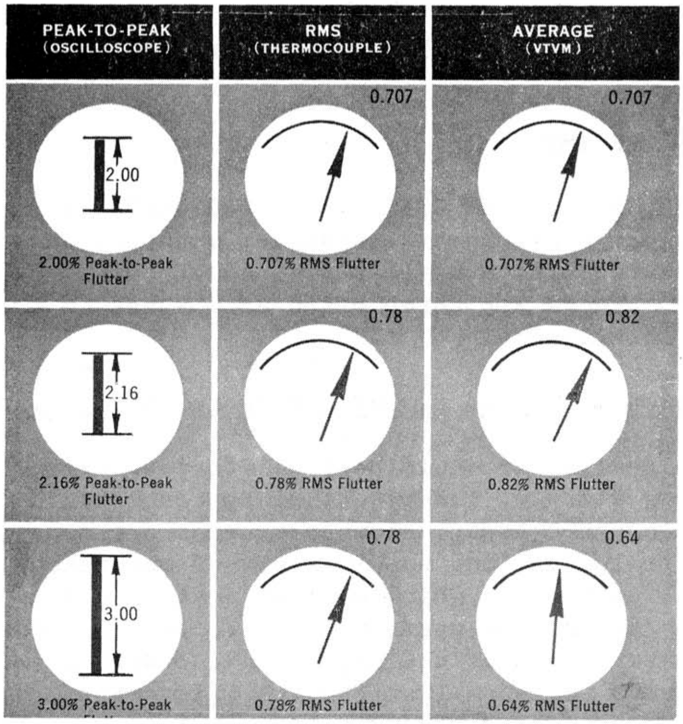


Fig. 6-22. Comparison of the Readings Obtained for the Same Flutter Components Using Different Indicating Instruments

Actually the discrepancy between readings made by different methods can -- and probably does -- get much worse. Even in the simple cases cited, the flutter error having a true rms value of 0.78% could be indicated variously as 0.64% rms, 0.82% rms, 2.16% peak-to-peak or 3.00% peak-to-peak!

The above simplified explanation of some of the methods used to measure speed eccentricities has been included to point up an important fact: Published flutter specifications may be very misleading unless the method of measurement is known so that the reader may interpret the quoted figures correctly. This is again dramatized in the accompanying comparison of the results obtained with rms and peak-to-peak methods of flutter measurement when they were used to chart the identical flutter components in terms of decibels. Decibels (db) are the units used for logarithmic expression of the ratio between the data signal and the system noise including flutter.

$$db = 20 \log_{10} \frac{E_1}{E_2}$$

where E_1 is the data signal and E_2 is the combined system noise and flutter.

Generally, flutter signals contain a variety of frequencies and amplitudes so that the combination is similar to broadband system noise and often confused with it. In earlier instrumentation work when the imp was larger than at present, he was easy to isolate and measure. But today's improved transports often hold him in the range of the system noise where he is actually hidden. Because the two are often indistinguishable except by highly specialized techniques, flutter is usually measured only to the level of the noise inherent in the system.

HOW TO LIVE WITH THE IMP:

This section might be titled "Eliminating Flutter", except that it can't be done. Though he can't be ejected, modern transport engineering and manufacturing plus ever-improving electronics make it possible to live with him. There are three basic approaches to this peaceful -- but watchful -- coexistence.

Mechanically Improved Transports Through Better Design and Manufacturing

Considering only design, it would be ideal to just eliminate rotating parts! This being a dream, the practical approach is to isolate them so they can't influence the tape. Because a transport consists basically of the heads, tape guides, one supply reel and one takeup reel, these are the essential units with which we must deal. Such a basic transport is shown (Fig. 6-24) and was used on early recorders. For instrumentation, however, interaction of the supply takeup systems with the tape in the head area causes too much flutter with this configuration. Isolating elements are therefore added to negate the influence of the two systems. When applied to the basic transport, they constitute an Open-Loop Tape Drive (Fig. 6-24B). To counteract the influence of the takeup system, a

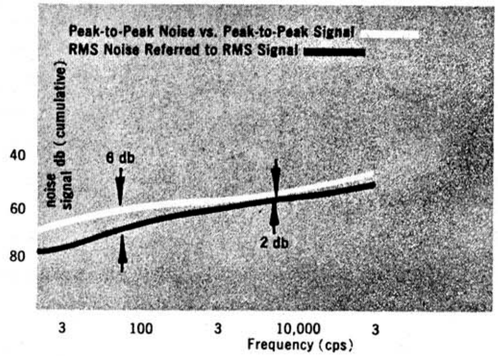
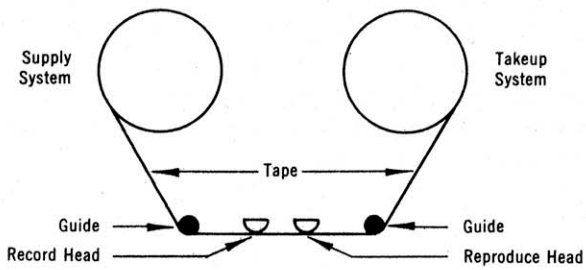
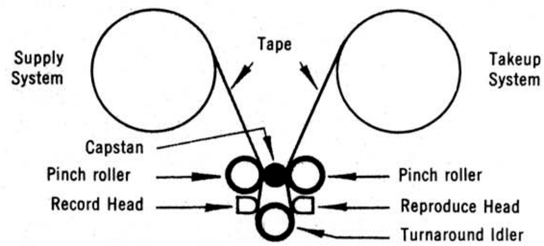


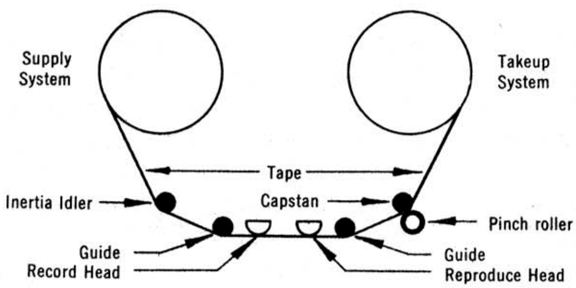
Fig. 6-23. Peak-to-Peak vs. RMS Noise



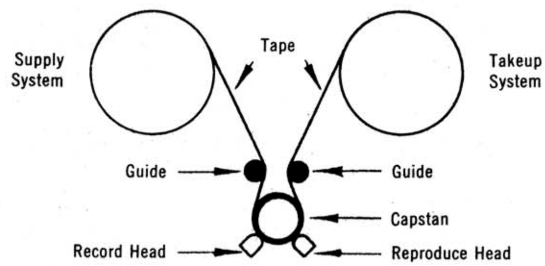
A. Basic Magnetic Tape Transport



C. Closed-Loop Tape Drive



B. Open-Loop Tape Drive



D. Zero-Loop Tape Drive

Figure 6-24

tape-driving capstan and a capstan idler are used. Supply system flutter is controlled by a tape-driven inertia idler equipped with a flywheel to counteract erratic tape velocities. The inertia idler in turn introduces another problem -- because of its weight, a measurable amount of time is required to bring the idler up to tape speed.

To evade this slower starting and give shorter lengths of unsupported tape, a Closed-Loop Tape Drive was evolved wherein the tape wraps around a "turn-around" idler and is clamped against the capstan before and after passing the heads (Fig. 6-24C). The idler is extremely light for fast starting and the effect of the double tape-to-capstan clamping inhibits the imp to a considerable degree. Because the unsupported lengths of tape are short, any "scrape" flutter is of a very high frequency.

The ultimate in isolation is the Zero-Loop Tape Drive, (Fig. 6-24D). Here, the tape is pressed against the heads by the capstan itself so that idlers are not needed. So -- why aren't all recorders built this way? The answer is the high cost of the meticulous machining required on the capstan.

The work of the designer, no matter how good, becomes worthless if the manufacturing process is not held to the highest standards. This means one thing -- CARE. Care in selecting and handling components, the use of perfect bearings and precision motors, constant checking of rotating parts for concentricity, and a vigilant attitude toward anything and everything that might give aid or comfort to the enemy.

Electronic Flutter Compensation ---

This type of flutter correction depends on the generation of coherent -- identical in the time domain -- errors in all channels. Because of this identity in time, noise generated by flutter in an unmodulated channel, called a "reference" channel, can be used to reduce or cancel the noise generated in the data channels and super-imposed on the data signals themselves.

Because electronic flutter compensation corrects frequency aberrations resulting from velocity irregularities, it combats the elusive troublemaker only in FM recording, the only one of the four basic recording methods in which data amplitude is encoded in terms of frequency.

The compensating signal is derived from a continuous, single-frequency "carrier" like that described earlier to measure flutter. This constant frequency is recorded on a reserved "reference" channel on the tape and is not modulated with data signals. This monitoring signal is used in either of two basic ways to compensate for flutter in the data channels:

Signal Subtraction. Here, reference-channel flutter is "inverted" and added to the demodulated data-channel output to cancel flutter.

Pulse-Area Control. In this case the reference correction is applied to the complex data-channel signal before it is demodulated.

With the signal-subtraction technique, the nominal carrier frequency of the reference signal can be set anywhere in the pass band. If it is made identical to the center frequency of the data channel, the "coherent" noise appearing on the unmodulated data channel can be almost completely cancelled. But if the mean frequency of the data channel is offset, say by a slowly changing signal, the amplitude of the reference noise will change. This is because the amplitude of flutter noise is proportional to the carrier frequency -- the relationship is a constant percentage -- and 1% of 70,000 cps is more than 1% of 50,000 cps! In other words, subtraction compensation is truely effective at only one predetermined level or carrier frequency. However, it is quite popular because of the simplicity with which it can be accomplished.

Pulse-area control, or pulse-width control, as it is also known, is an effective way of correcting data amplitude errors due to the imp. With this system, the reference-channel carrier must be removed from the reference-channel output signal before it is fed to the data discriminators. Unfortunately, the filter for the carrier introduces time delays that differ between high-and low-frequency signals. That is, the frequency determines the time required for a signal to pass through the filter. As a result, the high-frequency reference signals arrive at the data discriminator too late to do their work but the low-frequency errors can be very effectively eliminated.

Because most of the flutter disturbances generated by tape-transport mechanisms lie in the low-frequency region, the superior effectiveness of the pulse-area control technique in this region led to its adoption for compensating Ampex Instrumentation recorders.

Servo Tape-Speed Control Systems

As mentioned earlier, tape is involved in several ways in the creation of flutter. Two of these are speed irregularities during recording and/or reproduction and shrinking or elongation of the tape. Several electro-mechanical regulating or "servo" systems are used to combat these conditions:

Precision Frequency Supply

A precise 60-cycle signal originating at the frequency standard is amplified to drive the capstan motor and tape at an exact rate. This system only regulates tape speed; it does not correct errors caused by speed or tape-dimension changes. It improves the characteristics of record, reproduce or record/reproduce machines.

Record Servo System

This is used with record-mode machines when tapes for servo-controlled reproduction are being made. The system is identical to the Precision Frequency Supply with the addition of a Control-Track Oscillator-Modulator unit. This employs a 60-cycle reference frequency to modulate the control track signal before it is recorded on one tape track. When used with the Reproduce Servo System, the Record System helps correct both tape-speed and tape dimension errors.

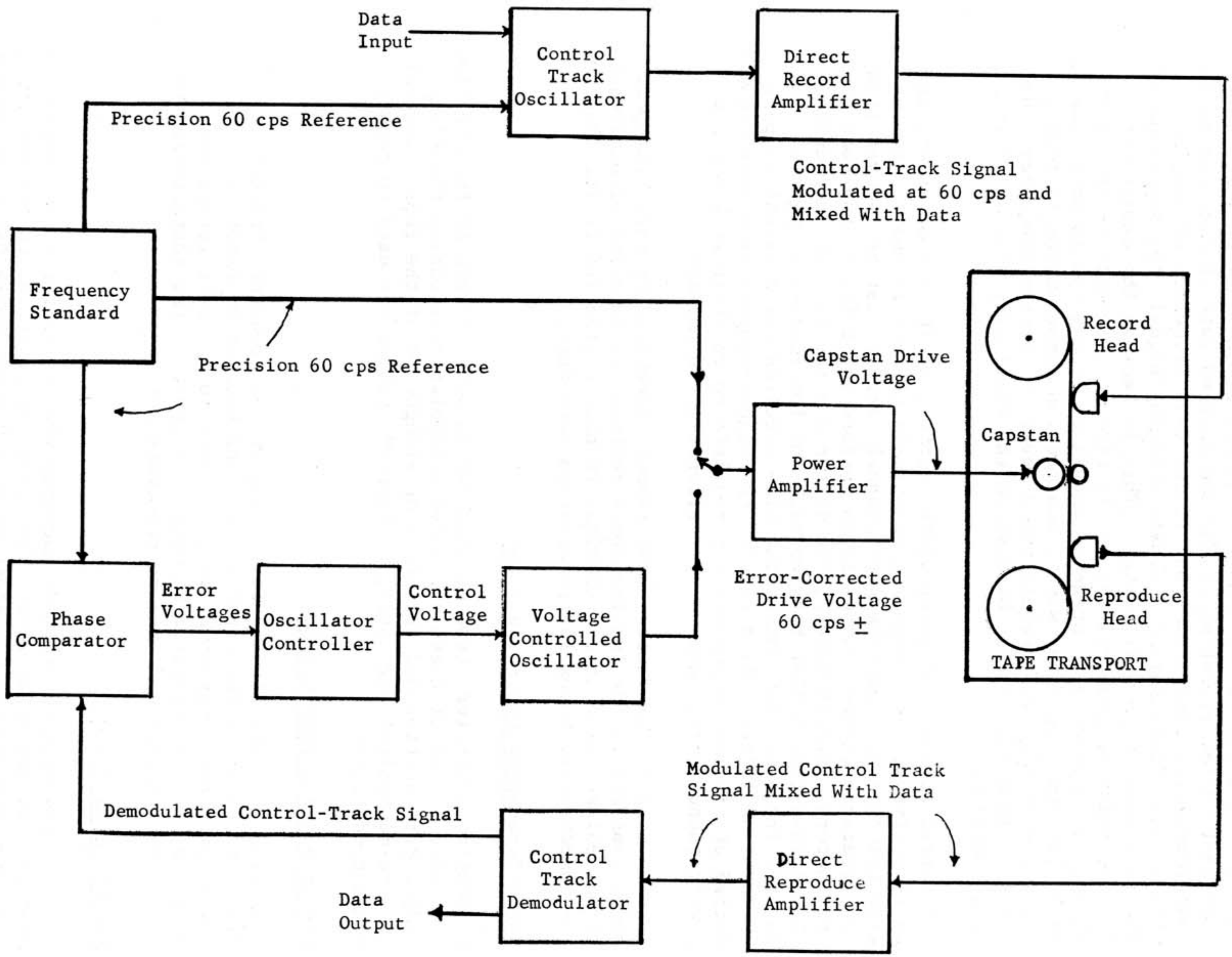


Fig. 6-25. Direct Record/Reproduce Servo System

Reproduce Servo System

Used with reproduce-mode machines. The control track generated by the Record Servo System is demodulated and compared with the 60-cycle reference to give an error voltage proportional to the phase difference. This voltage deviates the voltage-controlled oscillator from its nominal 60-cycle output. The resulting signal is amplified and applied to power the capstan-drive motor. The degree of frequency deviation from the nominal is such that the capstan speed increases or decreases to correct precisely any tape-speed or tape-length irregularities.

Record/Reproduce Servo Systems (Fig. 6-25)

For record/reproduce machines. Here, the frequency standard and power amplifier, along with their power supplies, do double duty by serving both the record and reproduce systems of the transport.

A BLACK DAY IN THE OFFING:

The rascal, alias "Wow," alias "Drift," finds himself engaged in fighting a losing war with the magnetic tape recording industry.

The attack on Flutter is waged on four fronts: Improved mechanical design of tape transports; better manufacturing techniques and controls; electronic compensation when Flutter can not be completely defeated; and as a further harassment, electromechanical control of tape speed.

CHAPTER VII

RECORD/REPRODUCE ELECTRONICS

There are several distinctly different recording processes in common use - each of which requires a different form of electronic encoding and decoding of the signal information. Some of the more common recording processes will be described and explained. A graphic illustration of some of the differences existing between the various recording methods are shown in Figure 7-1.

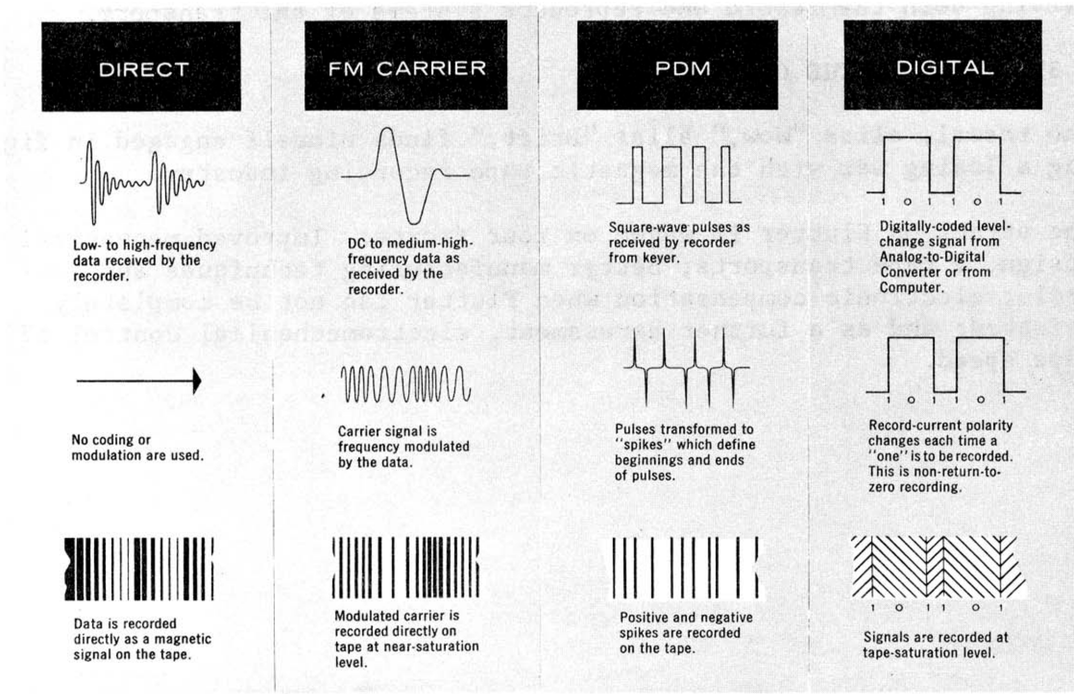


Figure 7-1

DIRECT RECORD/REPRODUCE

Normal analog signals are recorded on tape by the direct record method. Here the information lies in the magnetization level from point to point along the tape. High Frequency bias current, of frequency approximately three to five times the highest signal frequency to be recorded, is usually added to the signal at the record head. This bias current has the effect of linearizing the magnetization curve of the oxide. Direct recording utilizes the maximum bandwidth capability of the recorder but is limited in accuracy by all extraneous phenomena causing amplitude changes. The data signal to be recorded is amplified, mixed with bias and presented directly to the recording head as a varying electric current, i.

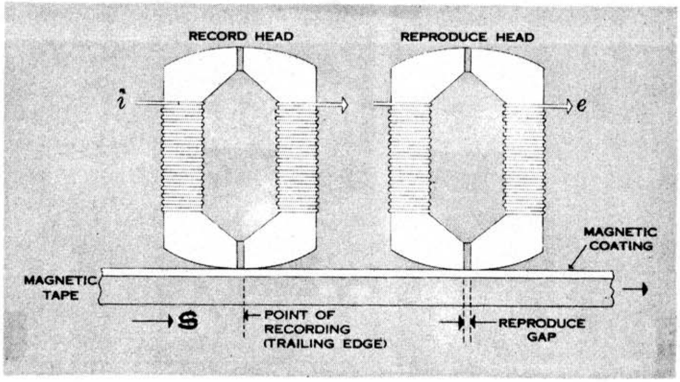


Fig. 7-2. The magnetic tape recording and reproducing process

The recording takes place at the trailing edge of the record head gap. The signal current flowing through the winding which surrounds the magnetic core produces a magnetic flux ϕ, whose magnitude will be proportional to the recording current.

A wavelength of recorded signal along the tape will occur for each complete alternation of the input electrical signal. This wavelength will be directly proportional to the tape speed and inversely proportional to the frequency of the recorded signal.

$$\lambda = \frac{S}{f}$$

λ = wavelength on tape (inches)
S = tape speed (inches/sec)
f = frequency (cycles/sec) of electrical signal

During playback, the magnetized surface of the tape passes the gap of a reproduce head, which is similar in construction to the record head. The portion of tape in contact with the gap is bridged by the magnetic core of the reproduce head, which causes magnetic lines of flux to flow through the core. The magnitude of this flux is a function of the average state of magnetization on that portion of the tape actually spanned by the gap at any given instant. As the tape passes by the reproduce gap, the amount of flux through the core varies with the varying state of magnetization on the tape and causes a voltage to be generated in the winding linking the core. It is important to note that the voltage generated is proportional not to the magnitude of the flux -- but to its rate of change. Therefore, the playback voltage is dependent upon the frequency and for constant-current recording (recording current constant at all frequencies) the output voltage will vary in direct proportion to frequency.

Fig. 7-3 illustrates the effect of frequency on the recording and reproduce characteristics and calls attention to two facts. The first is that the reproduce (decoding) amplifier must have a frequency response characteristic which is the inverse of the reproduce head characteristic, in order that an overall flat frequency response can be obtained. This is referred to as "playback equalization" and is illustrated in Fig. 7-4.

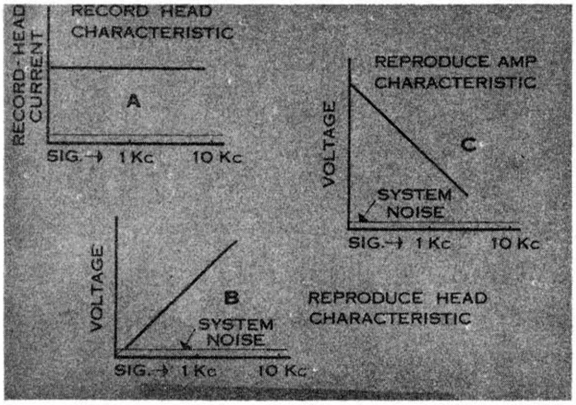


Fig. 7-3. Effect of Frequency on Recording and Reproducing Characteristics

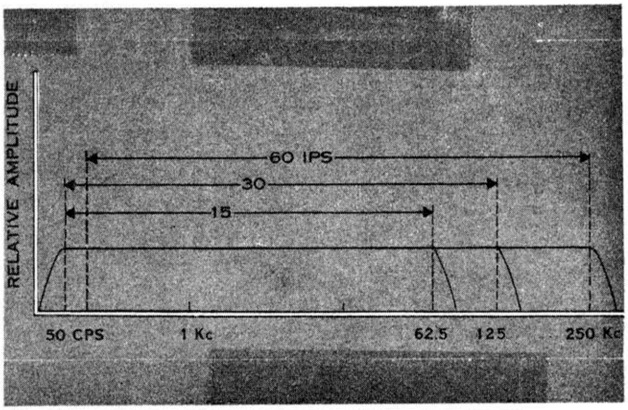


Figure 7-6. Direct-Record Frequency Response

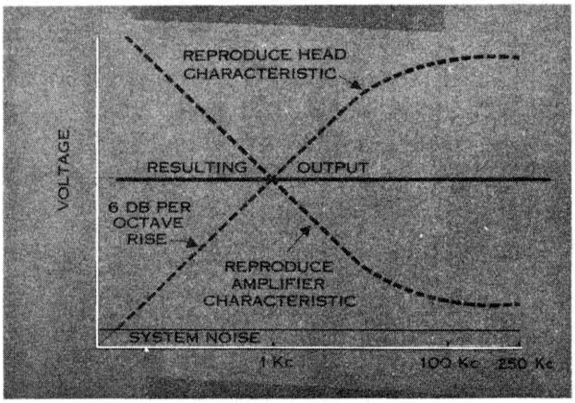


Fig. 7-4. Reproduce Amplifier Output

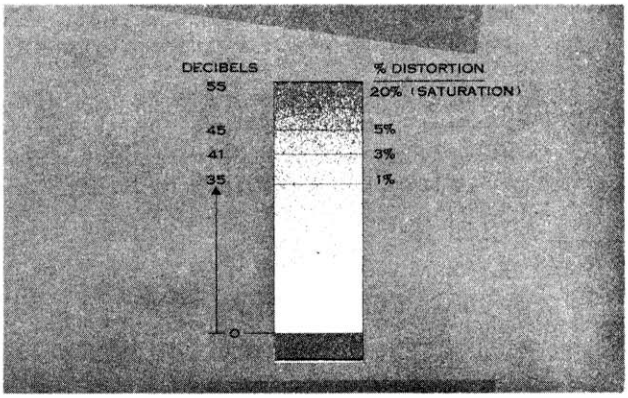


Fig. 7-7. Dynamic Range

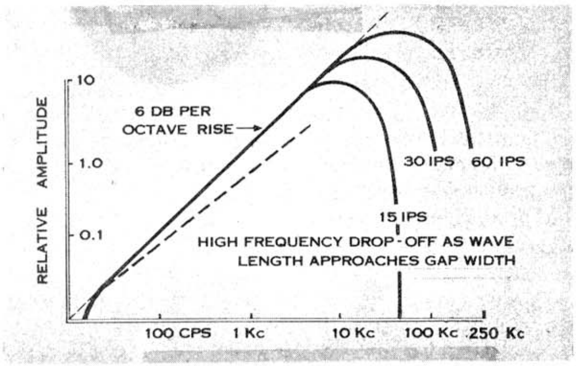


Fig. 7-5. Reproduce Head Output

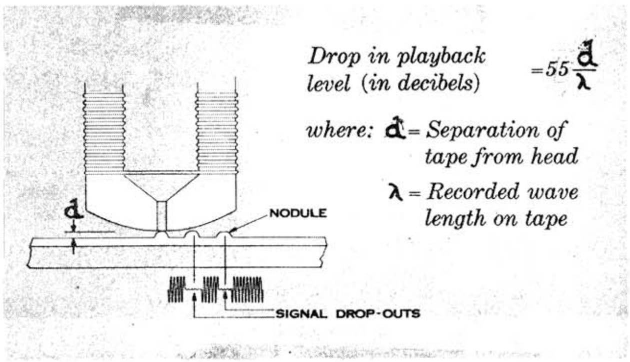


Fig. 7-8. Effect of Tape-Surface Imperfections

The second fact observed is that as the frequency goes lower and lower, the output voltage from the reproduce head decreases until it approaches the inherent noise level of the system. At this point, it is impossible to recover the signal by equalization. This condition leads to one of the principal limitations of the Direct Recording Process. Refer to Fig. 7-5.

The frequency response of a typical direct record system is 250,000 cycles/second at 60 inches/second or 4000 sine-wave cycles/inch of tape. This latter number can be taken as a figure of merit for the Direct Recording System and can be used as a basis of comparison with other recording systems. Refer to Fig. 7-6.

DYNAMIC RANGE (signal-to-noise ratio): The dynamic range is the ratio of the maximum signal which can be recorded (at a given level of distortion) to the minimum signal which can be recorded (determined by the inherent noise level of the system). As the recording level is increased into the nonlinear magnetic characteristic approaching saturation, the distortion increases correspondingly. The bar graph in Fig. 7-7 illustrates some typical values for dynamic range expressed in decibels.

AMPLITUDE INSTABILITY: Another limitation of the Direct Recording Process is Amplitude Instability. This is a condition brought about by causes external to the recorder itself, namely, the surface condition of the magnetic tape medium. The effect manifests itself by instantaneous lapses or reduction in signal level, which are commonly referred to as "drop-outs". The surface defects of tape will create signal drop-outs as shown in Fig. 7-8.

The "drop-out" effect is relatively unimportant for audio recording, because the ear tends to integrate variations in signal level which occur instantaneously and, thus, is insensitive to them. For instrumentation recording, on the other hand, this effect might be intolerable, as for example if it is required to preserve accurately the waveshape of a transient phenomenon.

AUDIO RECORDERS: Before leaving the Direct Recording Process, it would be well to mention the inherent dangers of using an audio recorder for instrumentation purposes. An audio recorder may be considered to be a special case of the Direct Recording Process. The characteristics of the record and reproduce amplifiers are modified to conform to the particular characteristics of speech- and music-type signals. It has been established that the energy content in speech and music signals is not uniformly distributed over the range of signal frequencies. For this reason, pre-equalization circuits are incorporated in the record amplifier which pre-emphasize some portions of the frequency spectrum (the extreme low and high ends). These are the frequencies at which the energy content of audio signals is low. By raising their level, it is possible to approach a constant -flux recording situation on the tape at all frequencies. In this way benefits can be achieved in signal-to-noise ratio without sacrifice in distortion. Of course, the inverse frequency-response characteristic must be introduced in the reproduce amplifier, in the form of postequalization, to counteract the effect of the pre-equalization and produce a final output signal which is a replica of the original input signal.

The danger in using an audio recorder for instrumentation recording is that the instrumentation-type signal does not in general have the peculiar spectral energy distribution characteristics of speech or music. The result is that the pre-emphasis in the record amplifier could result in serious distortion of the high and low frequencies. This could only be overcome by reducing the recording level by a considerable amount -- with resulting deterioration of the signal-to-noise ratio of the recording.

APPLICATIONS OF DIRECT-RECORD PROCESS:

The advantages of the Direct Recording Process is that it has the widest frequency spectrum for a particular tape speed. Using practical tape speeds, signals can be recorded over a continuous range of frequencies from 50 cps to 250,000 cps. Other advantages are its wide dynamic range and its ability to handle moderate overloads gracefully, without sudden or drastic increases in distortion. The major applications for the Direct-Record process are:

1. The recording of signals where the significant information is contained in the relation between frequency and amplitude on a logarithmic basis. Examples are in the measurement and subsequent spectrum analysis of noise and underwater-sound signals.

2. The recording of voice commentary on one of the tracks of a multitrack recorder for the purpose of logging and identification.

3. Multiplexing a number of signals simultaneously on one track by assigning to each channel of signal information a separate portion of the wide frequency spectrum which is available in the Direct Recording Process.

Inherent limitations in the Direct Recording Process have been circumvented by various modulation schemes. Prominent are frequency modulation (FM), pulse duration modulation (PDM) and pulse code modulation (PCM).

CHAPTER VIII

FREQUENCY MODULATION RECORDING PROCESS

This method employs a carrier frequency which is frequency modulated by the signal to be recorded. Thus, a particular frequency is selected as the center frequency corresponding to zero input signal. A dc signal of positive polarity would deviate the carrier frequency a given percentage in one direction. A dc signal of negative polarity would deviate the center carrier frequency an equal percentage in the opposite direction. An ac input signal would deviate the carrier alternately on both sides of center frequency, at a rate equal to the frequency of the input signal. Thus, all information presented to the tape is preserved in the frequency domain and normal amplitude instabilities will have little or no effect on the recording.

Figure 8-1 illustrates an elementary block diagram of the electronic coding employed in the Frequency Modulation Recording Process and illustrates the relation between the input signal and the signal presented to the recording head. On playback, the signal is demodulated and fed through a low-pass filter which removes the carrier and other unwanted frequencies generated in the modulation process.

The first widespread application of the FM recording technique was in frequency-division multiplexing where a number of individual carrier frequencies (F_1, F_2, F_3---) are each modulated by a separate input signal (f_1, f_2, f_3---) as illustrated in Figure 8-2.

The resulting multiplicity of signals is then mixed linearly and the composite signal recorded using the Direct Recording Process. Thus, the wide bandwidth and the linearity of the Direct Recording Process is used to permit the simultaneous recording of many channels of signal information on one track of tape.

It is important to note that the FM recording process makes very stringent demands on the ability of the tape transport to move tape across the heads at a precisely uniform speed. Any speed variations introduced into the tape at its point of contact with the heads will cause an unwanted modulation of the carrier frequency and result in system noise. This is the limiting factor in the dynamic range and accuracy of the FM system.

A more detailed explanation of frequency modulation follows:

1. The amount of deviation is controlled by the amplitude of the signal input.

2. The rate of deviation is controlled by the frequency of the deviating signal.

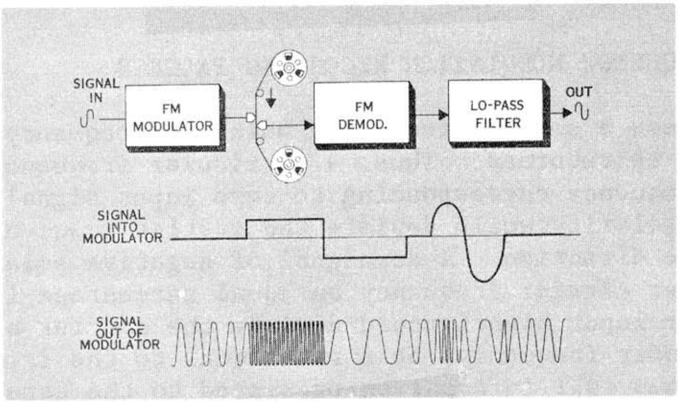


Fig. 8-1. Basic FM System

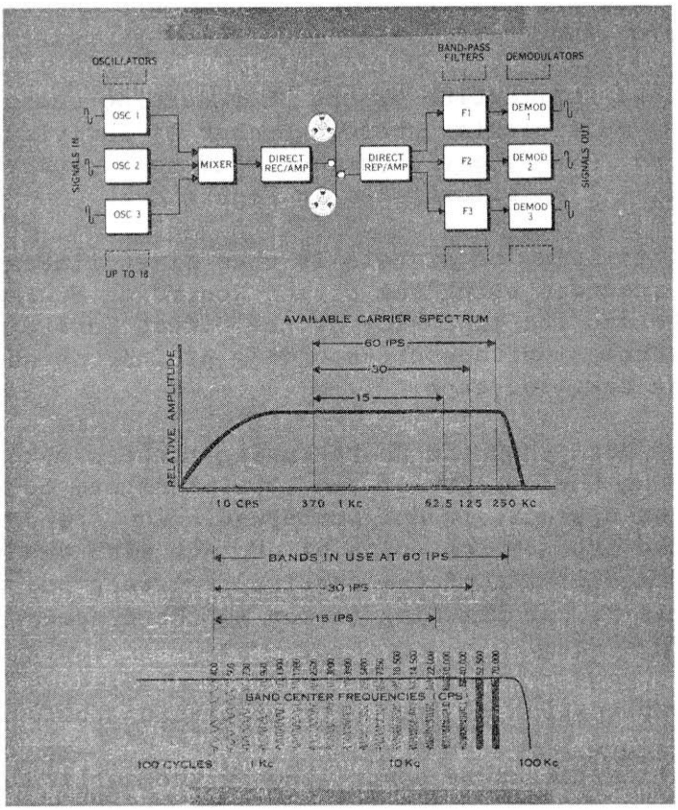


Fig. 8-2. Basic Frequency-Multiplexing Narrow-Deviation FM System

3. The concept of the frequency modulation as applied to telemetry and other forms of communication systems, and that used in magnetic tape recording is similar.

For a single modulating frequency there is an infinite number of sidebands separated from the carrier frequency by the modulation frequency. Thus, the frequency band specified by the signal is not clearcut and cannot be defined precisely as it is in amplitude modulation. The spectrum of the FM signal is considerably broader than that of an AM signal. It is wider than 2 x df. df is defined as the amount of deviation, i.e., the change of frequency from center carrier frequency. As the modulation index becomes large, the sideband spectrum approaches more closely ±100% deviation. For very small modulation indexes, the spectrum width will approach that of the center carrier frequency.

In most transmission systems, it is assumed that a flat amplitude frequency characteristic over the requisite bandwidth is all that is necessary. It is true so long as a system has a normal phase-frequency characteristic corresponding to its flat amplitude-frequency characteristic. In other words, the phase characteristic is linear over the required bandwidth. There are systems, and magnetic tape is one of these, in which the normal relationship between amplitude-frequency and phase-frequency characteristics does not hold. In these cases, we must consider carefully the effect of the system on the modulated carrier signal.

A study of the phaser diagrams of an amplitude modulated signal and that of an FM modulated signal would prove that:

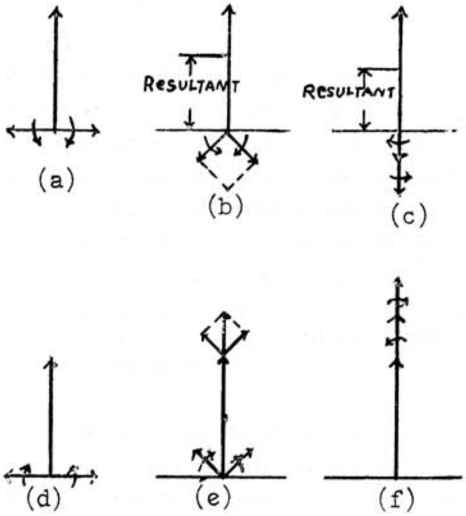
1. The amplitude modulation resultant phaser remained in a fixed position, indicating no phase or frequency change, but varied in amplitude. Fig. 8-3.

2. The frequency modulation resultant phaser swings back and forth in phase, indicating frequency change, with little change in amplitude. Fig. 8-4.

3. The modulation index represents the maximum phase deviation of the frequency modulated signal.

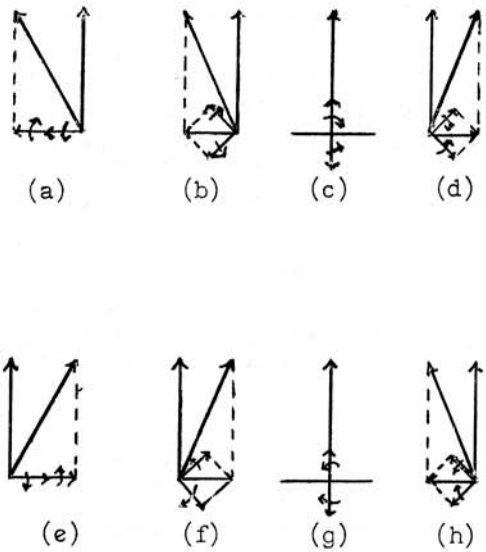
A study of the Bessel functions would show that the sideband spectrum of the signal is not contained within the deviation range. The center carrier frequency, the amplitude of which is not fixed, is dependent upon the modulation index, where the modulation index

$$= \frac{\text{carrier frequency deviation}}{\text{modulating frequency}} = \frac{df}{f_1}$$



Amplitude Modulation Vector Diagrams

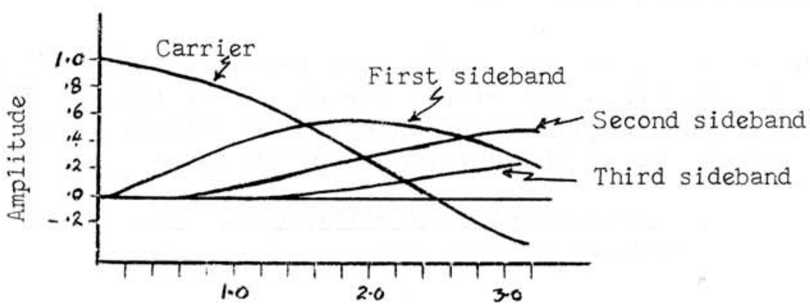
Fig. 8-3



Frequency Modulation Vector Diagrams

Fig. 8-4

It would indicate that the energy in each of the sidebands is equivalent to the (Bessel functions)^2. Although the total energy of the envelope will remain unchanged, or constant, the instantaneous energies of the center carrier frequency, and the resulting sideband frequencies will vary. Since the process of frequency modulation removes energy from the carrier and distributes it to the sidebands, the number of sidebands that are necessary will obviously depend upon the modulation index. The separation of these sidebands from the center carrier frequency is a function of the modulating frequency. The bandwidth occupied by the signal is the product of the total number of necessary sidebands and the modulating frequency. It is a complex function of the frequency deviation and modulating frequency, and is not determined by either one alone. (Fig. 8-5).



Modulation Index

Fig. 8-5

When the modulation index $\frac{df}{f_1}$ is such that f_1 is the maximum modulating frequency, and, therefore, the modulation index would be $\frac{df}{f_{1_{max}}}$ this will be called the **DEVIATION RATIO.** This important parameter determines to a large degree the susceptibility of the system to noise and spurious signals other than noise. The deviation ratio for the FM broadcasting industry is approximately 5 (df = 75kc, $f_{1_{max}}$ = 15kc). This value is normally used in FM/FM telemetry subcarriers. In wideband FM recording, deviation ratios in the range of 1-2 are more common. Ampex has used, in the FR-100 series, a deviation ratio of slightly more than 2 (df = 21.6kc, $f_{1_{max}}$ = 10kc). In the FR-600 equipments, the same deviation ratio is used.

The primary consideration for the choice of center carrier frequency is such that the range from maximum to minimum deviation will fall within the bandwidth limitations of the recorder.

In addition to the above the center carrier frequency is chosen to be of such a value that time base expansion and contraction is possible over the full range of tape speeds of the recorder/reproducer.

The **PERCENTAGE DEVIATION** $\frac{100\ df}{f_o}$ is also a parameter of major importance in both telemetry and tape recording. Telemetry subcarriers deviations of 7½ and 15% are standard, while in wideband recording systems, deviations will range from 20 to 75%. Ampex uses a 40% deviation as standard.

It is important to note that FM recording processes make very stringent demands upon the ability of the tape transport to move tape across the heads at a precise and uniform speed. Any speed variations introduced onto the tape at its point of contact with the heads will cause unwanted modulation of the center carrier frequency and result in system noise. This is a limiting factor in the dynamic range and accuracy of the FM systems.

When using telemetered or other types of record signals, in a tape system, the recording amplifier, head-tape system, the playback amplifier, the bandpass filter and limiter-amplifiers which preceed the demodulators must be included as limiting factors for percentage of deviation. Where a deviation percentage of $\pm 7\frac{1}{2}\%$ is used, and where the system introduces a 1% deviation due to flutter and wow in the transport, this ratio of deviation percentage to the frequency error due to flutter and wow can be represented as the signal to noise ratio. In the stated case, the noise signal would be $\frac{1}{7.5} \times 100 = 13.3\%$.

Effectively, this means that any errors introduced by the instantaneous tape speed changes (flutter and wow) would be multiplied 13.3 times.

For wideband FM recording, the transmission system includes the recording circuitry following the FM oscillator, the head tape system, and the limiter-amplifiers between the playback head and demodulator.

Now, with a $\pm 40\%$ deviation and a 1% deviation, due to tape speed error, (flutter and wow) this would appear as $\frac{1 \times 100}{40} = 2.5\%$ noise signal. Thus, wideband FM recording is an improvement in the signal-to-noise ratio better than five times that of the narrow band system.

The ideal transmission system which includes the tape transport has a flat amplitude-frequency characteristic, a zero or linear phase-frequency characteristic over the bandwidth required by the signal. Where we can tolerate a small amount of phase, or frequency modulation, introduced into an AM signal a non-linear phase response introduced into an FM signal will cause distortion of the modulating signal and appear as noise in the system's output.

When a wider signal bandwidth is required and a reduction in signal-to-noise can be tolerated, this can be obtained by reducing the amount of deviation. By reducing the deviation by one-half we may increase the data bandwidth by two.

FM recorders are generally designed for lv in and lv out, rms, but recently developed low-level FM input circuitry affords such improvement in sensitivity and stability that full deviation is accomplished for an input voltage of only 2 mv.

Performance of the low level system is fully as good as that of conventional high level systems in all respects and superior in certain others. The advantage lies in the record amplifier. Previous record amplifier linearity is typically within $\pm 0.2\%$ from best zero based straight line whereas the low level record is linear to within $\pm 0.05\%$ on the same basis. One of the most notable features of this new low level FM record amplifier is its low center frequency drift. Over ambient temperature rise from $25^{\circ}C$ to $50^{\circ}C$ it drifts no more than 0.8%. Its long-term drift at room temperature is 0.2% for one week compared with 1% for 2 hours in conventional high level amplifiers.

In view of the fact that transducers and the recorder may be some distance apart there may be ground potential differences between them. Such potentials can be serious in a recorder having 2 mv rms full-scale input sensitivity; and special arrangements are required to reduce the effect of this common mode voltage.

The preferred arrangement is to ground the shield of the signal lines at the transducer and to connect this shield at the recorder to a metallic guard enclosing the shielded electronics. In a typical conventional system wherein the common mode voltage might cause an amplifier input error as high as 4 volts rms at 60 cps, this guard can reduce the error voltage to 15 micro micro-volts -- small even compared with the 2 millivolt input signal.

APPLICATIONS OF THE FREQUENCY MODULATION PROCESS:

The advantages of the Frequency Modulation Recording Process are: (1) its ability to record low frequencies down to dc; (2) its freedom from the effects of tape drop-outs; and (3) its excellent phase-shift versus frequency characteristics with the attendant ability of accurately preserving the waveform of a recorded signal. The disadvantages of the FM process are: (1) less efficient utilization of the tape, requiring approximately ten times the tape speed for a given upper-frequency limit; (2) the additional complexity of the electronic circuitry requiring modulators, demodulators, and low-pass filters; and (3) the requirement for a tape transport which is engineered and manufactured to high standards of precision.

The major applications for the FM Recording Process are:

1. The recording of low-frequency signal information, such as vibrations, noises, and under-water sounds for spectrum analysis.

2. The recording of transient phenomena, such as shock, blast and ignition, where accuracy of wave shape is important. Short transients having rise times as short as 60 microseconds can be resolved with this process.

3. Changes in time base permitting a speed-up or slow-down of a given event; and permitting the frequency components of a given signal to be scaled up or down by large factors, up to 1000 times or more.

PULSE DURATION MODULATION RECORDING PROCESS

A frequency-division multiplexing technique has been discussed for recording a number of signal channels on a single recorder track by sharing the available frequency spectrum among the signal channels. There is a second technique to accomplish a similar result, in which time can be shared between a number of channels of signal information. This technique is called time-division multiplexing and requires an instantaneous sampling of a number of signal channels on a sequential basis.

When recording a sine wave as in Figure 9-1, it is normal to think of a continuous recording of each instantaneous value of the wave. It is possible, however, to sample the sine wave at uniformly spaced discrete intervals; record only the instantaneous values at the time of sampling; and then reconstruct the original sine wave on playback by passing the discontinuous readings through an appropriate filter. An accurate reproduction of a sine wave can be made using as few as six samples per sine-wave cycle shown as points A through F in Figure 9-1. This technique is, of course, equally valid for non-sinusoidal signals, provided the sampling rate is at least six times the highest significant frequency component of the non-sinusoidal wave.

If a data signal is being sampled at discrete intervals, it is possible to use the time between these sampling intervals for the purpose of sampling other data signals. This is most conveniently accomplished using a rotating commutator, as shown in Figure 9-3, wherein the outputs of a number of transducers are being sampled in sequence, once per revolution of the commutator. Figure 9-3 is a simplified block diagram of the entire PDM recording system and should be referred to throughout the following discussion.

Two of the contacts on the commutator are usually reserved for frame reference to allow synchronization with the commutator used on playback. Two of the other contacts can be used for calibration signals, one corresponding to full-scale voltage, and the other corresponding to zero voltage. Thus, the system permits continuous calibration (once each revolution of the commutator). The remaining sets of contacts are available for connection to the outputs of various transducers (T_1, T_2, T_3---).

The standard system established by the Research and Development Board for telemetry purposes is based on the use of 900 samples per second. This permits some flexibility in system design, where we can employ a 30-contact commutator rotating at 30 revolutions per second; or a 45-contact commutator rotating at 20 rps; or a 90-contact commutator rotating at 10 rps. Using the first of these combinations, 26 channels

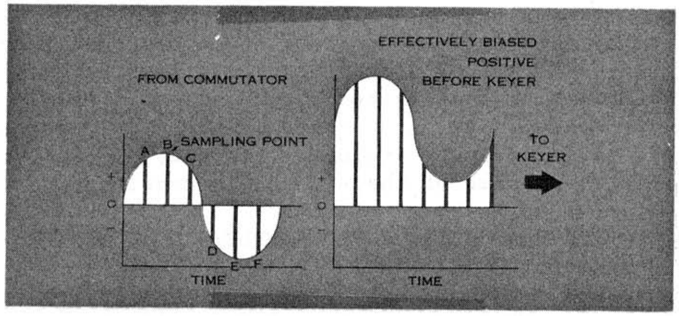


Fig. 9-1. Sampling a Signal in PDM

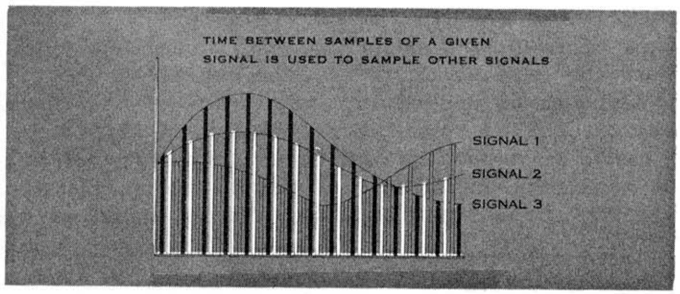


Fig. 9-2. PDM Sampling Technique

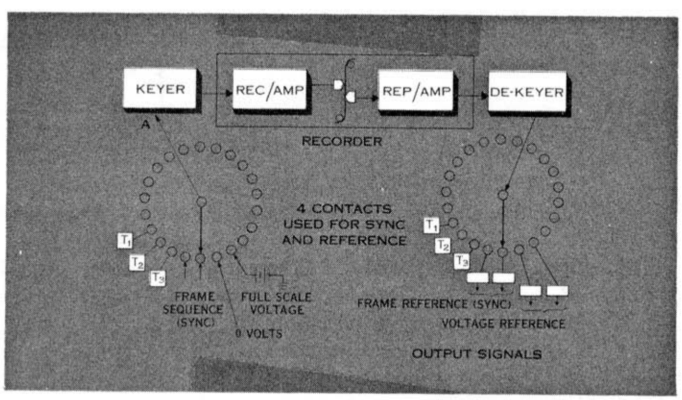


Fig. 9-3. PDM Recording

of information can be recorded (subtracting the 4 contacts used for frame reference and calibration) each having an upper-frequency limit of 5 cps (30 revolutions or samples per second, where 6 samples determine a sine-wave cycle). By using other combinations, we can record more channels having a lower frequency content, or vice versa. The following table shows some of these possible combinations.

POSSIBLE COMBINATIONS USING A 900 SAMPLE/SECOND SYSTEM

Commutator		Signal Information	
No. of Contacts	RPS	No. of Channels	Upper Freq. Limit
30	30	26	5 cps
45	20	41	3
90	10	86	1.5

Figure 9-4

Additional flexibility is possible by paralleling 2 or more contacts and connecting them to the same transducer. (Figure 9-5). This will increase the number of samples per second (and consequently the upper frequency limit) at the expense of fewer data channels, where the higher frequency response is needed for only a few of the channels. Examine the signal out of the commutator (point A in Figure 9-3). It can be seen that there is a sequence of very-short-duration pulses of varying amplitude occurring at intervals of once every 1100 microseconds (the reciprocal of 900 samples per second). These appear as shown in Figure 9-6A.

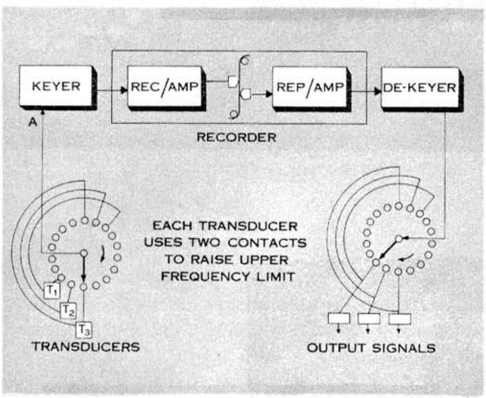


Fig. 9-5. PDM Recording

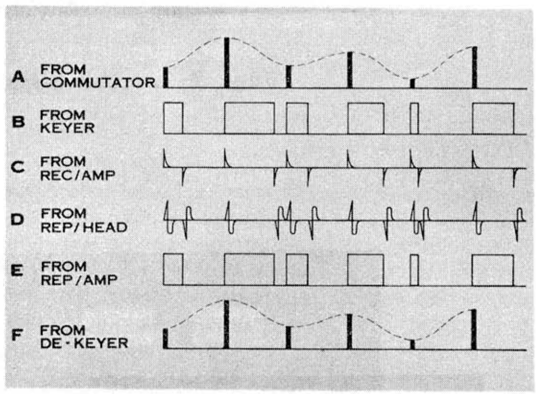


Fig. 9-6. Signals at Various Points In PDM

For the reasons discussed in the Direct Recording Process, the inherent amplitude instability of magnetic tape would prevent accurate and dependable recording of these varying-amplitude signals. So, the signals are passed through a "keyer" which converts them from varying-amplitude signals into constant-amplitude signals of varying pulse-width or pulse-duration, as shown in Figure 9-6B. A definite pulse width can be assigned to each value of input signal amplitude within the 1100 microsecond total interval available between samples. The following pulse durations are normally used:

Input Signal Voltage	Pulse Duration
0	90 microseconds
5 volts (full-scale)	660 microseconds

Any intermediate voltage between zero and full-scale would be represented by a definite pulse duration somewhere between these two limiting values.

Since the only significant information in the signals from the keyer Figure 9-6B, is contained in the time at which each pulse begins and ends, the record amplifier (encoder) sharply differentiates these pulses, presenting to the record head a positive "spike" corresponding to the beginning of a pulse and a negative "spike" corresponding to the end of a pulse. This is shown in Figure 9-6C. Here, too, it is important to use a tape transport having a minimum of instantaneous tape-speed variation. But flutter and wow is less critical in this process because only the integrated speed errors occurring between the instant of pulse-start and pulse-stop would introduce an error in the recorded data.

The output from the reproduce head consists of a differentiation of the recorded "spike", since, as was explained before, the reproduce head responds to rate-of-change of magnetization on the tape. The wave form out of the reproduce head is shown in Figure 9-6D, where the point of axis crossing of the reproduced wave represents the instants of pulse-start and stop. The reproduce amplifier (decoder) contains a multivibrator which recreates the original pulses from the output of the reproduce head, resulting in the signal in Figure 9-6E, which is a replica of Figure 9-6B, the original varying-width pulse.

These pulses are then fed through a "de-keyer", where they are converted back into varying-amplitude pulses of short duration, as shown in Figure 9-6F, which in turn is a replica of the original output of the commutator - Figure 9-6A. The final operation is to feed this output into another commutator where decommutation occurs and the original data channels are separated out into their individual filters. These filters serve to reconstruct and deliver the original data signals.

APPLICATIONS OF PULSE DURATION MODULATION RECORDING PROCESS:

The chief advantage of the Pulse-Duration Modulation Recording Process is its ability to record a large number of simultaneous channels of information. Using a 10 rps by 90-contact commutator, it is possible to record 86 channels of information on one track of tape, or 1200 channels of information on a 14-track recorder. Other advantages are the high accuracy (better than 1% overall) made possible by the self-calibrating feature; and the inherently high signal-to-noise ratio, resulting from the narrow signal-frequency bandwidths involved.

The disadvantage of the PDM process is the limited frequency response of each channel; the less efficient utilization of the tape (one-quarter that of the FM process -- giving a comparative figure of merit of 40 sine-wave cycles per inch of tape); and the increased complexity of the auxiliary electronic equipment, such as commutators, keying amplifiers and filters.

The applications of the PDM recording process are all those which involve a multiplicity of signal channels having relatively low-frequency content. Examples are flight testing and engine testing, where the information to be recorded is derived from a large number of transducers, such as thermocouples and strain gauges.

CHAPTER X

AN INTRODUCTION TO DIGITAL RECORDING

The language and techniques of digital recording are starting to stabilize. The recent boom in the computer business has forced engineers to standardize terms, formats, and equipment. To those who grew up on analog equipment, words like transfer rate, bit packing density, and access time, can prove confusing, even meaningless. Here is an explanation of some of the techniques and terminology used by digital tape users.

Magnetic recording devices have become an indispensable element in modern digital computing equipment. While other mediums, punched paper tape and cards, have been used with success, they have fallen far behind magnetic tape in the two most important criteria:

1. Transfer rate, the speed with which information can be put into or obtained from the storage medium, and

2. Packing density, the amount of information that can be stored in a given volume of the medium. A 2400-foot roll of one-half-inch-wide magnetic tape can contain 1.5×10^7 characters. 180,000 punched cards, or 150,000 feet of punched paper tape would be required for the same amount of information.

THE TERMS:

Digital data processing employs either the coded binary system or an **alphanumeric** code or modifications of either. In computer use the fundamental unit of the system is the bit, which is indicated by a pulse. The value of the bit may represent either a one or a zero condition (hence the term binary). The value "one" or "zero" would be indicated by either the polarity of the pulse or by its presence or absence. The binary system itself is a numbering system based on a base of two rather than ten as is used in our everyday decimal system. Figure 10-1 shows how numbers from zero through nine can be indicated with the binary system by a combination of the values 1, 2, 4, or 8. This two-state system can perform arithmetical operations with much less electronic circuitry than if a decimal system were used.

The **alphanumeric** code uses coded bits in certain configurations to represent numbers, letters of the alphabet and symbols.

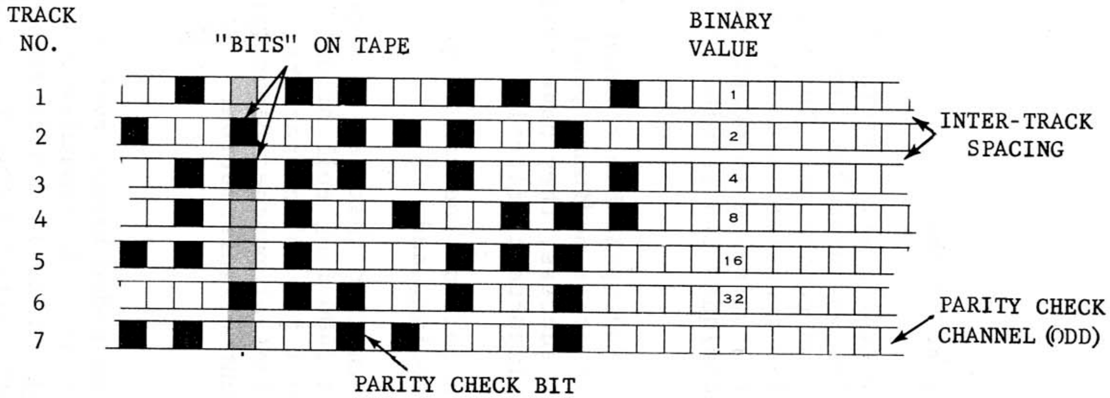
The properties of magnetic tape lend themselves readily to a two-state system. Analog recording is accomplished by magnetizing tape at an infinite number of levels, whereas digital recording magnetizes the tape in either of two directions and two levels only, either positive or negative. The tape is recorded to saturation in either direction.

CHARACTER	BINARY CODE	EQUIVALENT
0	0000	0+0+0+0
1	0001	0+0+0+1
2	0010	0+0+2+0
3	0011	0+0+2+1
4	0100	0+4+0+0
5	0101	0+4+0+1
6	0110	0+4+2+0
7	0111	0+4+2+1
8	1000	8+0+0+0
9	1001	8+0+0+1

560

0101 0110 0000

Figure 10-1. Binary representation of the numbers 0 through 9. The position of the 1's (positive pulses) in the four-bit binary code indicates the presence of the respective binary digit; 8, 4, 2, or 1. These, added, restore the code to the original form. The number 560 is shown in both forms.



0 = 0

1 = 2

2 = 4

0 = 0

0 = 0

1 = 32

38

Figure 10-2. A drawing of a section of tape recorded in 7-track format. The tape-transport moves tape past the read head one character at a time instead of in a continuous motion. The computer "reads" each character by adding the binary digits represented by the dark squares. The binary value of each track is shown at the right, and the binary and numerical value of one character is shown at the bottom.

GETTING IT ON THE TAPE:
There are several basic formats for writing ones and zeros. Figure 10-3 shows the series 0 1 1 0 1 1 1 recorded by each of the following formats.

<u>Return to Bias, or RB</u>, considers saturation in one direction as zero, and saturation in the other direction as 1.

<u>Return to Zero, or RZ</u>, returns the tape to the demagnetized state between each bit. In contrast to the other formats, RZ recording produces a recorded pulse for each bit. This form of recording has the advantage of carrying its own clocking, since there is a pulse for each and every bit.

<u>Non-return to Zero, or NRZ</u>, changes the state of magnetization each time a 1 bit occurs.

<u>Non-return to Zero, or NRZ</u>, comes in two varieties: NRZ(C) meaning "change", and NRZ(M), meaning "mark". NRZ(M) changes the state of magnetization of the tape each time a 1 occurs, but not when a zero occurs. NRZ(C) changes state whenever a 1 is followed by a zero, or a zero by a 1.

Remember that Figure 10-3 shows the bits on <u>one</u> track of a several-track system, and each pulse shown represents one bit of a character, and that the other bits for a particular character are contained on the other tracks.

To use NRZ recording an external clock system is required; in other words, we must look at the tape at regular intervals to determine whether a "zero" or a "one" should be read. Since conventional heads read only a change of magnetic flux, no signal from the head is

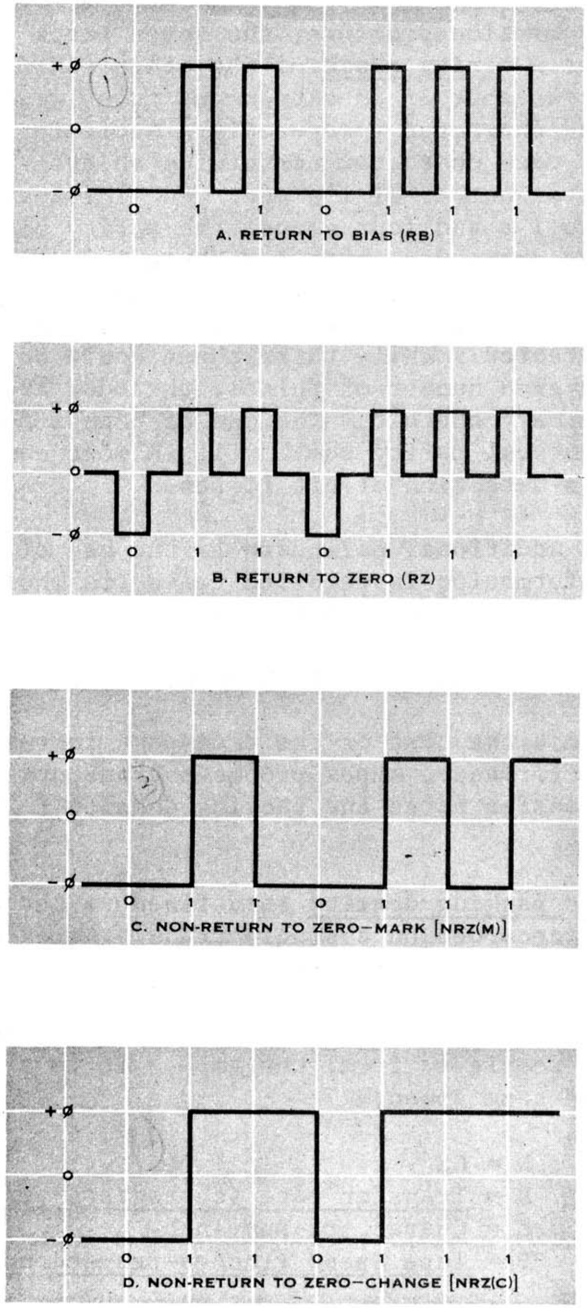


Fig. 10-3. The series 0 1 1 0 1 1 1 recorded in the five most common computer formats. The differences in these formats (explained in the text) makes it necessary to use electronic translators when feeding one format into a computer designed for another.

interpreted as a zero in NRZ(M).

TO FIND AN ERROR:

As mentioned before, the seven track format uses the seventh track for a parity check of the other six. The loss of a pulse, or the occurrence of an extraneous pulse will cause the computer to read the character incorrectly. A bit is recorded on the parity track in each character containing an even number of 1's. This makes the total number of 1's odd. For instance, if a character consisted of two 1's and four zeros, the parity bit would be recorded as a 1; if the character contained three 1's, the parity bit would be recorded as a zero. By totaling the character 1's and parity 1's during readout of the character, the loss or addition of one pulse can be detected. While this scheme could be fooled by the loss or gain of an even number of pulses, the odds are in favor of the loss or gain of only one bit. The use of both a horizontal parity check and a vertical parity check will in most cases completely eliminate the non-detection of bit losses.

An additional safeguard is the use of "redundancy" in which the same information is recorded twice (in whole or in part) on parallel but separate tracks on the tape.

TRANSFER RATE AND PACKING DENSITY:

Since the cost of the transport increases directly as a function of performance, Ampex produces transports with low, medium, and high transfer rates and packing density to best serve each computer design.

Bit packing density is defined as the closeness with which bits may be recorded on a single track. Ampex's newer digital transports can record 555.5 bits-per-inch.

Data-transfer rate is the rate at which information can be placed on, or retrieved from, the tape. It is a function of bit packing density and tape speed.

R = CS
R = Transfer rate (characters-per-second)
C = Characters-per-inch
S = Tape Speed (inches-per-second)

200 characters-per-inch at 60 inches-per-second will produce a transfer rate of 12,000 characters-per-second. Until recently, this was a typical transfer rate. Because of this low rate, computers have been "tape-limited"; the speed of operation of the computer was limited by transfer rate of the tape recorder. Present computers operate at transfer rates of 60 kc and higher, by using 150 ips and one-inch tape, two characters may be recorded across the tape which will yield a 60 kc transfer rate.

BUFFERING:

In a sense, the buffering system of a computer buys time. It is the electronic analogy of "Hurry Up and Wait". Defined, a buffer is a short-term storage for data. For instance, in spite of the best efforts of design engineers, not all the bits in a character pass over the reproduce head gap at the same instant. Although the difference is in microseconds, the computer must store each bit in a buffer until they all arrive -- in order to read the character without error. Another buffer system overcomes the relatively slow transfer rate of tape (compared to the computation rate). A command is given by the computer tape transport to read a certain portion of the tape. The tape transport or the tape control center remembers this information and starts a search for the desired information.

Meanwhile, the computer is free to perform other operations which it may need to do. When the tape unit has found the desired address, the information is read into the input buffer where it is stored until the computer decides it needs it. When the computer is ready, it interrogates the buffer and transfers the information into an internal memory device for instant access. When data is to be transferred back to tape, it is fed to an output buffer for storage until the tape transport decides where to put the information. Eight years ago, input-output buffers were not used, and the entire computer had to stop and wait for the tape transport.

Buffer size, record length, and tape-transport speed accuracy are inter-related, and this causes some interesting problems. One computer performs an operation called copying -- transferring data from a file on one tape to a file on another tape. This is not a duplication process because the computer acts on the data between tapes. Mechanical tolerances and tape dimensional changes can cause a speed difference between the tape transports, so that the information might come out of the first transport faster than the second transport could write it.

This computer inserts a buffer between the two transports but nominally keeps the buffer only half full. If the transfer rate is high or low, the buffer will slowly fill or empty, but feed the data to the second transport at a proper and constant rate.

ABOUT STORAGE TECHNIQUES:

Earlier it was implied that not everyone used the same technique for distributing pulses along the tape. There are, in fact, many schools of thought on storage techniques.

Bits are added to make characters, characters then make words. The characters can be grouped to form either a fixed word or a variable word. A fixed word always contains the same number of characters. In one computer, this number is twelve. Sixty words form a record.

Another computer uses words of variable length with an indefinite number of characters in each word. The variable-length words can be written serially along the tape to form a record which can be some fixed length, say 100 words. The record does not occupy any fixed distance on the tape, but must contain the 100 words. Other systems have variable word lengths, and variable record lengths. In these systems, words are formed of any number of characters, and any number of words can form a record.

Each of these methods has advantages and disadvantages. The fixed word, fixed record system allows rewriting over an existing record without danger of harming adjacent records. And, this system uses less tape because the distance over which a record is written is long compared to the total length of tape used for inter-record gaps. Also, fixed word length makes design of the computer and its programming simpler, since a standard size piece of information is always used.

Variable word, variable record formats have the advantage in applications like payroll records. Because employee's names, salaries, taxes, and insurance premiums vary greatly in length, a rigid word-record system could pose a serious systems problem.

If, however, many records are quite short, a large amount of tape can be wasted in inter-record gaps. Short records are also wasteful from the standpoint of start-stop times required between records. To avoid this problem, variable word/variable-record-length systems allow the lumping of many records together to form one long composite record which can be read in a single pass. The same technique is sometimes used in fixed word/fixed-record-length formats. The procedure is called "packing" but requires more complicated programs.

Figure 10-4 shows the waveshape of a NRZ(M) bit as it travels through the write amplifier, the heads, and the read amplifier. Ampex digital electronics can accept positive or negative pulses, or a flip-flop pulse configuration. In the write amplifier, the input signal triggers a flip-flop circuit which switches the current flow through the write head. The write-head voltage is shown in B, and resultant head current is shown in C. The flux change in the tape is sensed by the read head, and the $d\phi/dt$ characteristics of this head produce the voltage shown in D. In the read amplifier, the signal is amplified to produce the waveshape in E. This 2 volt peak-to-peak signal can be translated by Ampex electronics into any of the waveforms shown in F.

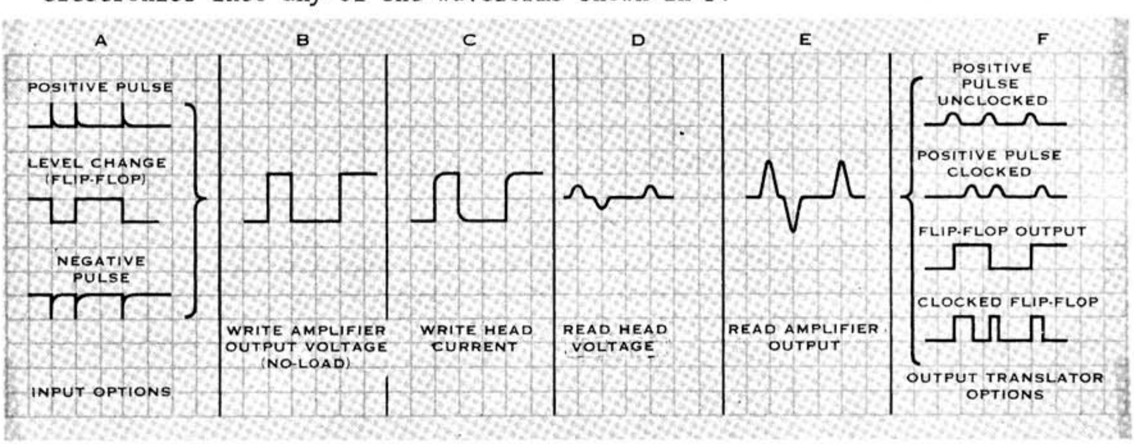


Fig. 10-4. Signal Waveshapes From Input to Output Through Ampex Digital Electronics

In some ways the digital process is a simpler one than those previously discussed and presents fewer design problems. For example, the record (write) and reproduce (read) amplifiers can be quite elemental. The speed stability of the transport is not as important, since relatively large amounts of flutter and wow can be tolerated without affecting the recording accuracy.

There are other problems, however, which become more important in digital recording. Sensitivity to tape drop-out errors is the most obvious one. Since all information is contained in the presence or absence of pulses on playback, the loss of pulses or the generation of spurious pulses caused by the tape imperfections cannot be tolerated. For this reason, special precautions are taken in the manufacture, inspection and selection of tape intended for digital recording. This does not completely solve the problem, however. As discussed under Direct Recording Process, tape drop-outs become most critical at short wavelengths, those approaching the size of the gap in the reproduce head. This indicates a practical minimum duration for pulses, an obvious factor in limiting the pulse packing density -- the number of pulses per inch of tape.

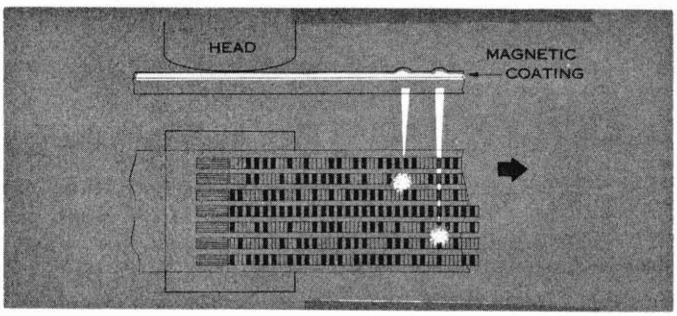


Fig. 10-5. Effect of Drop-outs in Digital Recording

Part of this same problem is the one of maintaining excellent head-to-tape contact to minimize the drop-out effect. This requires an extremely fine finish on the surface of the head, adequate tape pressure, and a minimum tendency for the head to collect oxide particles from the surface of the tape. It goes without saying that cleanliness must be maintained in the environment and handling of a digital recorder and its tape. (Figure 10-5.)

A second problem of the greatest importance to high packing densities is that of tape skew. This is any tendency for the center line of the tape to depart from a perpendicular to the line of record and reproduce head gaps. The reason for this is that the digits (or bits) making up a given number or character, are usually recorded in parallel fashion, all at the same time, each on a different track of the tape.

Figure 10-6 illustrates in exaggerated form the effect of tape skew in which the top and bottom head tracks would be reading bits from different characters at any given instant, instead of reading all the bits from a single character at one time. The preventative is to maintain excellent tape guiding in the design of the transport. Skew is perhaps the most important single factor in pulse-packing density, since the closer the pulses are packed together, the more important will become the possibility of errors due to tape skew.

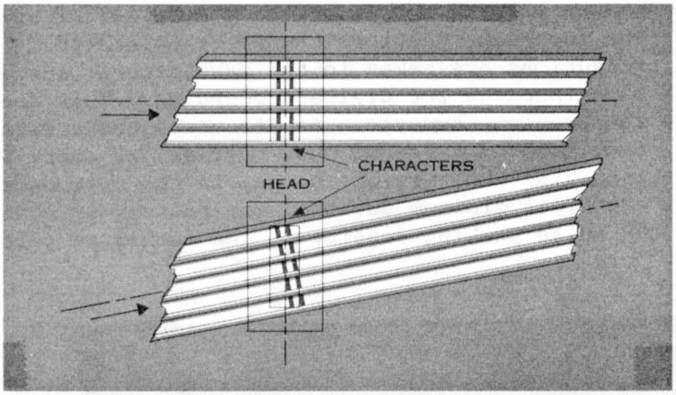


Fig. 10-6. Error Caused by Tape Skew

If the concept of the figure of merit for efficiency of tape utilization be applied to digital recording, it would be found that the tape utilization is much lower than that of any other type of recording process. For example:

1. Assuming a sampling rate of six data samples per sine wave (or highest frequency component of signal to be recorded)

2. Assuming a 1% accuracy requires 7-bits per data sample

3. Using a pulse-packing density of 200-bits per inch, then

$$\frac{1000 \text{ (bits/inch)}}{6 \text{ (data samples/sine wave) x 7 (bits/sample)}} = 24 \text{ sine wave/inch}$$

Summarizing and comparing this result with that from other recording processes, it can be seen that:

Recording Process	Tape Utilization (sine wave cycles/inch)
Direct	4000
Frequency-Modulation	333
Pulse-Duration Modulation	40
Digital	24

In spite of the fact that the tape utilization is lower, digital recording does preserve the recorded data in the language of the digital computer. Digital output can be fed directly into a computer where such operations as calibration, transducer linearization, scale factor corrections, etc. can be applied readily to the data. When large quantities of data must be taken, and when the original data is in analog form, prior to the conversion from analog to digital, it is customary to "edit" the data.

This is sometimes done using quick-look graphic techniques. From a quick visual inspection, it is possible to screen and determine those selected portions of the data which have significance. These are then converted from the analog to the digital form and recorded for further refinement and computation.

Where the transducers have the capability of measuring to degrees of accuracy beyond that of the other recording processes, and where it is required to preserve this accuracy, the original information is recorded digitally. This requires either of two approaches. The first is the use of transducers whose output is in digital form, of which many types are now commercially available. The second is to digitize the information from the transducers before recording it. This requires considerable electronic equipment (analog-to-digital converters) to be associated with the recorder; and it creates a weight, size, and complexity problem under certain conditions, such as when recording data in flight.

APPLICATIONS OF THE DIGITAL RECORDING PROCESS:

Summarizing the advantages of the Digital Recording Process:

1. Inherent capability of extremely high orders of accuracies.

2. Recording relatively insensitive to tape transport speed instabilities.

3. Simple record and reproduce electronic circuitry.

4. Output information is in the proper form for feeding directly into digital computers.

The relative disadvantages of the Digital Recording Process can be summarized as follows:

1. Poor tape economy -- 1/14th that of the FM process.

2. Data must be digitized at the source, or special digital transducers must be employed.

3. Reliability extremely dependent on tape quality, requiring redundancy and parity-check features.

The applications of digital recording are primarily for the processing of edited data, involving digital-computer techniques; and as an input device, output device, and internal storage for digital computers.

CHAPTER XI

PULSE CODE MODULATION (PCM)

KINDS OF MODULATION:

The conveying of information from one point to another need not necessarily take the form of a carrier that is a continuously varying wave. A train of pulses may be used. This type of carrier may be modulated in many ways.

1. Pulse Amplitude Modulation (PAM). In this type of modulation, the amplitude of each pulse will represent the amplitude of the modulating wave at the time of the pulse. There are two kinds of PAM.

   A. Uni-directional, and

   B. Bi-directional (Figure 11-1 B and C).

2. Pulse Duration Modulation (PDM), or Pulse Width Modulation (PWM). The duration or the width of the pulse is varied in accordance with the amplitude of the modulating signal. (Figure 11-1 D).

3. Pulse Position Modulation (PPM). The position of each pulse is varied with respect to some timing mark. (Figure 11-1 E).

4. Pulse Frequency Modulation (PFM). The pulse repetition rate is varied. (Figure 11-1 F).

5. Pulse Delta Modulation (PΔM). The individual pulse or groups of pulses will represent the difference between the amplitude of the modulating wave at one sampling point and the next. The differences in amplitude can be used to modulate the duration or position. (Figure 11-1 G).

6. Pulse Amplitude Delta Modulation (PAΔM). Like Pulse Delta Modulation, differences in amplitude can be used to modulate the duration or position. Any of the pulse trains can be used to modulate an rf carrier, either amplitude, frequency or phase wise. (Figure 11-1 H).

7. Pulse Code Modulation (PCM). One of the most promising methods of pulse modulation, but radically different from all of the other pulse systems. In this type of modulation, amplitudes are represented by groups of pulses. Each group is a code symbol that represents the amplitude of the modulating wave. (Figure 11-1 I).

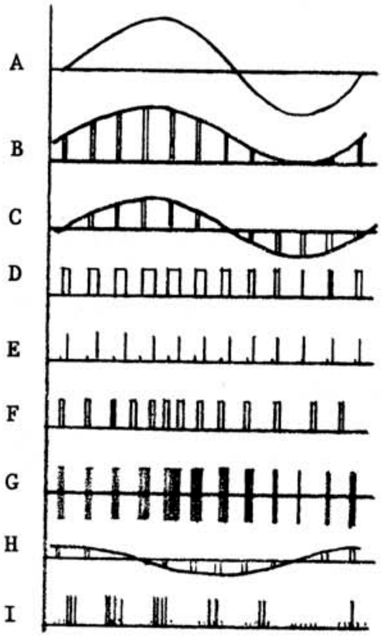


Figure 11-1 Kinds of Pulse Modulation

TIME DIVISION MULTIPLEX:

Any of the varieties of pulse modulation can be multiplexed. This will mean that the modulating signal will not be continuously represented but will be sampled periodically. It is necessary that any such sampling be at a rate which is in excess of twice the highest frequency component. For example, for frequencies up to 100 kc's, it will be necessary to sample at a rate of at least 200 kc. The width of the pulse may be as short as it is practically possible provided the amplifier bandwidth can be made wide enough to reproduce these pulses without excessive distortion. For instance, if a pulse has a length of 1 microsecond and is repeated at a 10 kc rate there will be 99 microseconds between pulses. Under these conditions there would be room between the successive pulses for other trains of pulses. Consider that an allowance of 10 microseconds for each pulse be made so that the pulses can be either duration or position modulated. Ten channels of information could be transmitted under these conditions. The mixing of a number of signals in this fashion is called "time division multiplexing".

Time Division Multiplex is extensively used in telemetry, and is being seriously considered for the simultaneous transmission of hundreds of voice channels on a single wideband radio relay circuit such as is presently being used for television.

PULSE CODE MODULATION (PCM):

PCM, a new type of pulse system, represents a major contribution to the communication art. In PCM continuous time function is sampled and the height (magnitude) of each sample is rounded off to the nearest value of those which will be permitted to be transmitted. Unike PAM, PDM or PPM, which are continuously variable, PCM samples are quantized into discreet steps. Thus only a finite number of values will be permitted and a particular value can be transmitted as a code group. This code group is related to the height (magnitude) of the quantized sample. For simplicity of detection and instrumentation, digital coding of PCM signals is the most popular. Digitizing requires that two levels be transmitted only -- zeros or ones. These levels correspond to the carrier being off or on. When decoding a PCM pattern, each code pattern is identified and caused to produce a voltage which is proportional to the original quantized sample. It is from this succession of samples that the original wave will be approximated. The transformation of an electrical signal to code pulses is represented in Figure 11-2. Since the actual signal value appears with both positive and negative values (Figure 11-2 B) with a zero average, a bias is added to simplify the coding problem. This way the signal will assume only positive values as shown in Figure 11-2 C.

Compression of the wave may be used, as is done in the case illustrated (Figure 11-2 C), to reduce the large peaks and increase the lower values. When such a system is put into effect, fewer numbers of quantizing levels are required for a given accuracy. This will also reduce the channel bandwidth requirements.

In decoding the signal is expanded in the reverse and is thus restored to its average value.

In our example (Figure 11-2) samples are taken every 100 μ seconds and converted into a 6-pulse binary code. The solid lines represent a pulse and a dotted line the absence of a pulse. Thus in code group C of Figure 11-2 D the code represents 0 + 2 + 4 + 8 + 16 + 32 = 62 volts. If an "n" digit code is used then 2^n values can be transmitted.

The transmitted pulses in a PCM system normally occur at a uniform rate in order to minimize the bandwidth requirements. For example, there are:

1. M signals to be multiplexed.

2. N pulses in each code group (6-pulse binary code).

3. Sampling occurs at 2 times the maximum frequency of each message = $2f_{max}$.

then it will follow that the time between pulses is ${}_2 NMf_{max}$. and

$$\text{minimum bandwidth} = NMf_{max}.$$

It is important to note here, that the bandwidth is proportional to the number of pulses per code group. This will mean that the bandwidth is related to the accuracy with which the signal may be recovered.

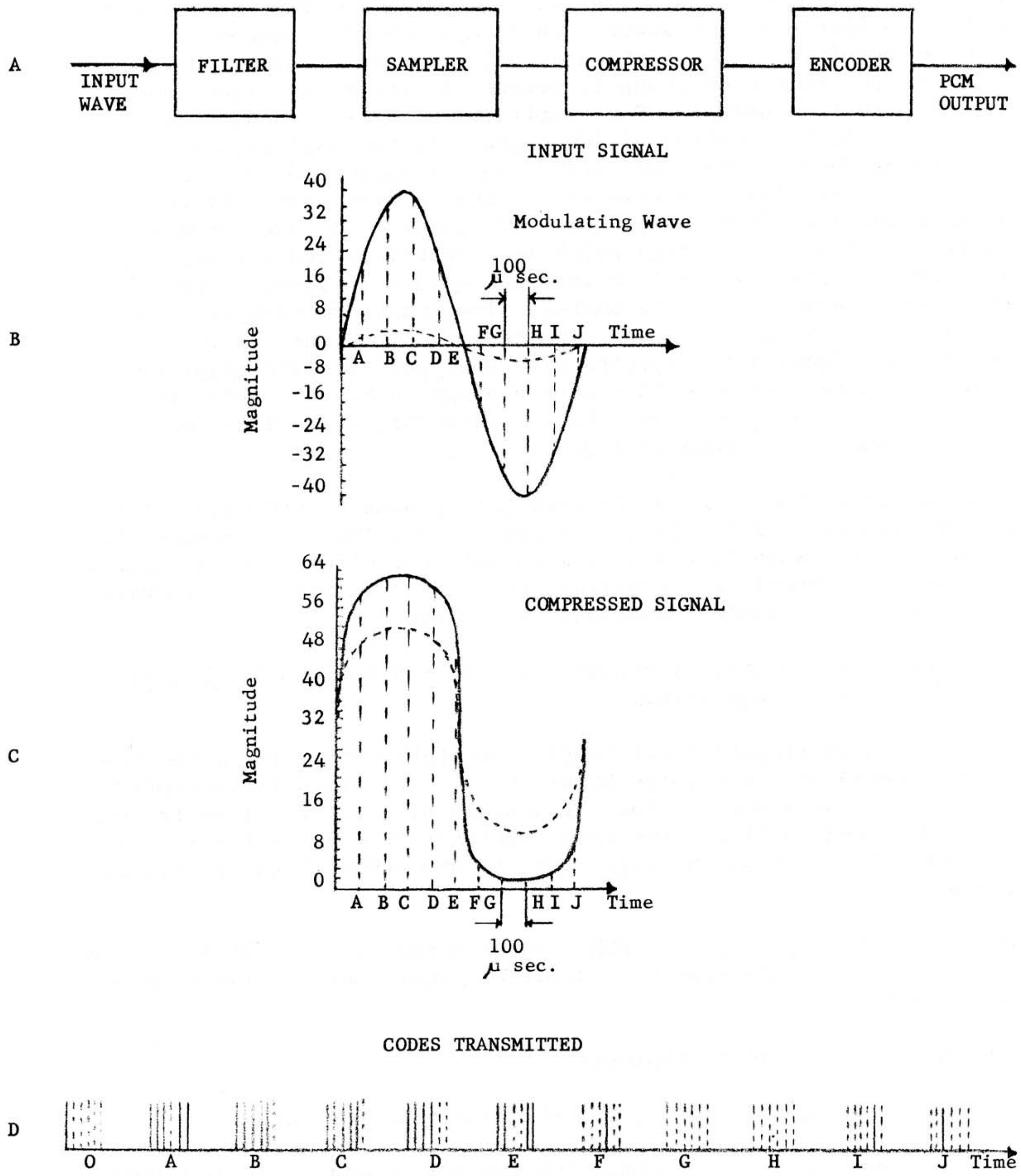


Figure 11-2

Pulse Code Modulation Transforming A Sine Wave Input To A Sequence Of Code Pulses

PCM has two very outstanding advantages:

1. It is free from noise and interference, and

2. The signals can be re-generated repeatedly without introduction of significant distortion.

The development of the modern PCM system stems from two factors:

1. Recognition of the fact that quantized samples will closely approximate the exact outline of a continuously varied wave.

2. The availability of improved techniques whereby quantized samples can be generated and decoded into complex codes at the high speeds required by our present-day wideband systems of communication. It is an attractive means of communication, since it places minimum restrictions on the kinds of messages that can be transmitted. It

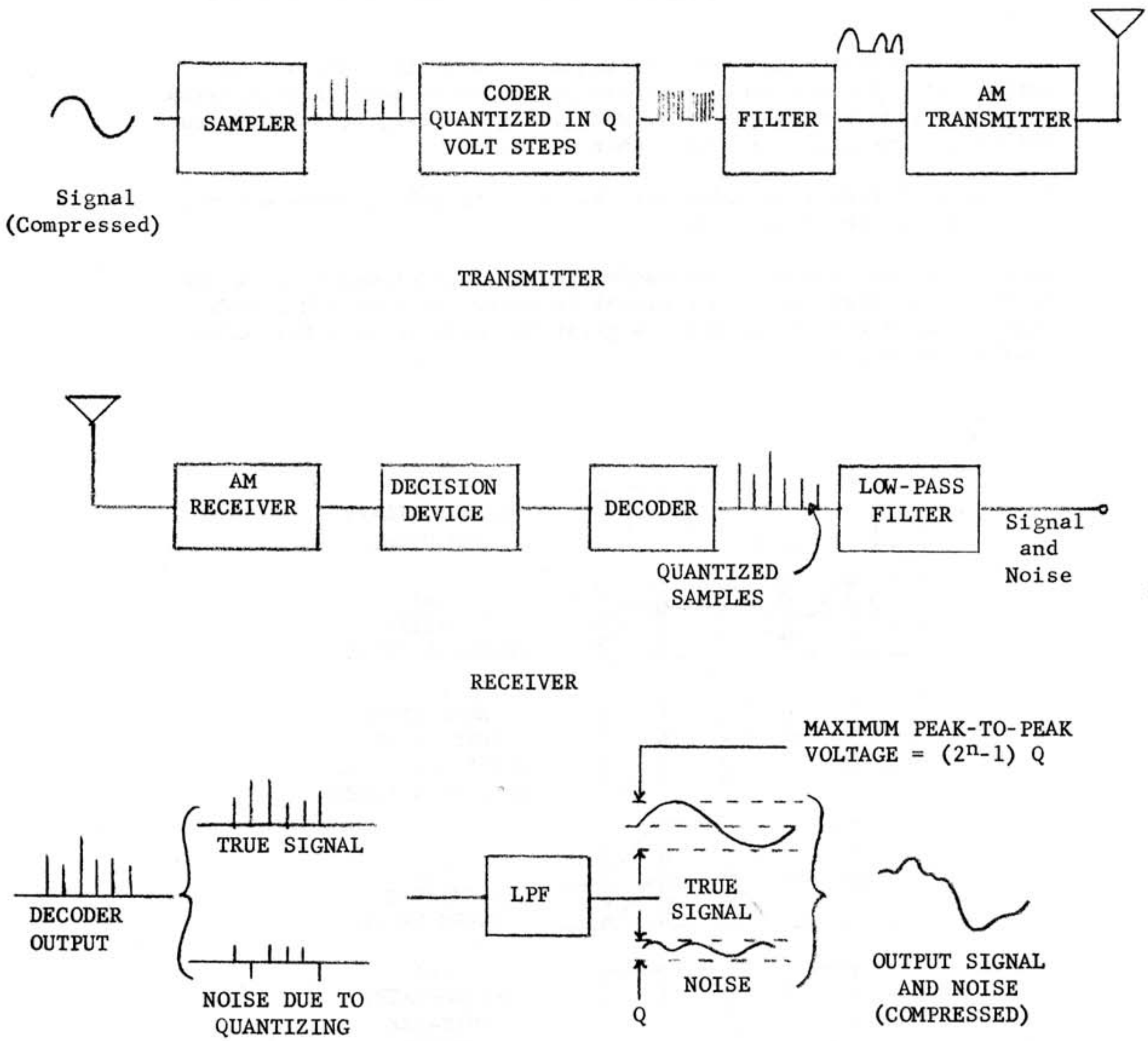


Figure 11-3 A PCM/AM System Showing Noise Due To Quantizing

has a signal to noise ratio that is substantially independent of the number of repeaters provided that the noise in each repetitive step adds up to less than one-half of the quantum step value.

SIGNAL TO NOISE RATIO OF PCM:

Neglecting noise on the channel due to transmission (for very seldom does this cause a pulse to be lost or misinterpreted by the receiver), the only noise source that can be considered is that due to the original quantization of the signal.

This quantizing noise is introduced by the encoder and could be as high as one-half a given quantizing level.

Referring to Figure 11-3 the signal samples do not have any error associated with them. Quantization error is introduced at the encoder.

Thus the receiver could have two signals fed to it. One a true sample value and the other a sample value due to quantization error. These two receiver inputs will, when fed to filter, produce two output waves, one signal and the other noise.

The ratio of signal to noise will be equal to $2^{2n}-1$, where n = number of pulses per code group.

When it is recalled that the bandwidth is proportional to n, it may be concluded that the output signal to noise ratio of PCM increases exponentially with bandwidth - a great improvement on other pulse modulation systems.

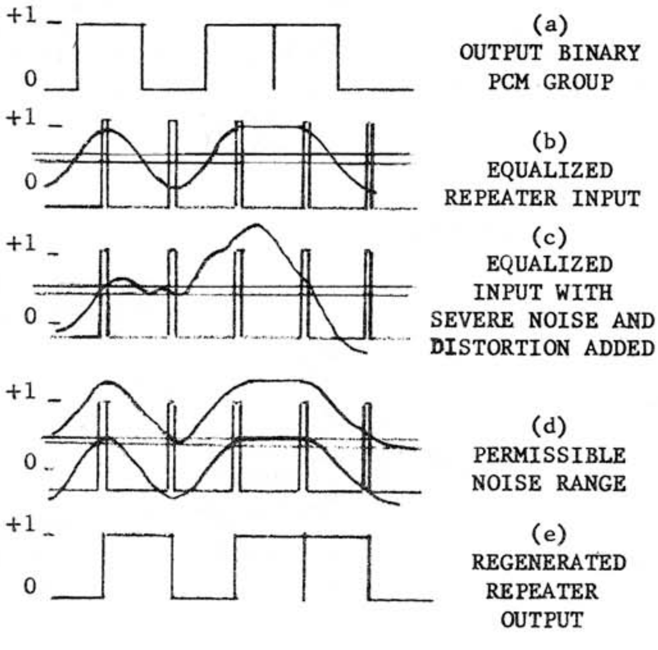


Figure 11-4 Input of a Coded Signal to a Receiver or Repeater

In summarizing the effects of the noise induced by the encoder during the quantizing process, refer to figure 11-4. The binary coded group which was fed to the AM transmitter was transmitted as simple pulses of known shape, amplitude and spacing. Either a pulse or a space must appear in each of the time slots. The receiver or repeater is designed to re-generate the signals; that is to say, they will make a completely new set of pulses from the received input. To do this, a simple "yes" or "no" decision must be made once each interval of time.

The signal sequence as shown in Figure 11-4 A could appear at the input of the receiver as Figure 11-4 B. Whenever the signal exceeds the double horizontal line at the half-amplitude level, it will be assumed that a pulse is present. Whenever the signal amplitude is below the horizontal lines it will be assumed that a space is present. The space between the double horizontal lines is a region in which the decision circuits cannot decide "yes" or "no". The result could be "maybe". This being the case, the narrower this "maybe" region is, the better the receiver or repeater will be.

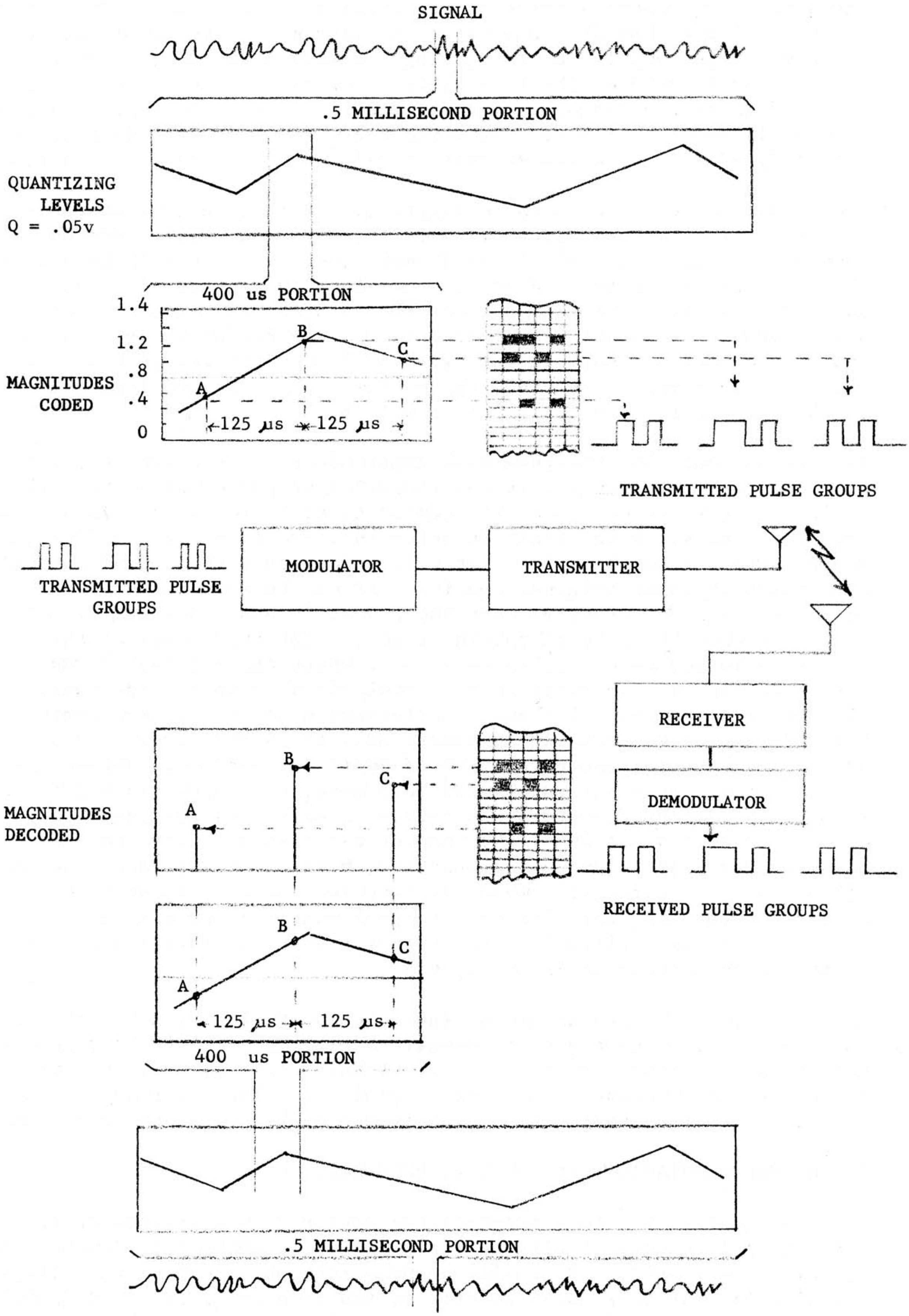
We must not only be concerned with amplitude re-generation, but also the time re-generation. Time re-generation is indicated by the narrow gating pulse which occurs at the center of each time slot. This gating pulse restricts the final decision to only those signals (levels) which occur during the pulse. It will assure that the signal pulses will maintain their original spacing. It can be said then that a pulse will only be produced when the signal is above half-amplitude at a time when the gate is passing signal. The importance of this procedure is indicated in Figure 11-4 C, where the original signals were received with an interfering signal added to them. The re-generation process reduces noise and interference effects, since although the first pulse was reduced to almost half the normal base-to-peak amplitude, it still exceeded the half-amplitude threshold during the gating pulse, and a new pulse will be generated as shown in Figure 11-4 E. Practical repeaters and receivers have been designed which will operate within 1 db of the theoretical 6 db signal-to-impulse noise margin. It is an unwise practise, however, to operate with so close a margin, especially when the interfering noise is ac in character; for in this case, the peak-to-peak interference can approach the peak-to-peak amplitude of the signal. This is indicated by the cross-shaded portion of Figure 11-4 D.

The combating of interference as indicated above is not obtained without some cost, since a greater bandwidth will be required to transmit information in binary form. As part of this cost, an inherent distortion known as quantizing noise is produced. This, already explained, is due to the transmission of a finite number of signal amplitudes.

PULSE CODE MODULATION AND THE TAPE RECORDER:

When a signal is fed to a transmitting device from a transducer or other signal source, it will be first filtered, sampled at discreet intervals, compressed, and finally encoded into pulse code groups (Figure 11-2 A). At this point it will be fed to a modulator, and finally transmitted as a series of sine waves similar in form to that shown in Figure 11-4 B. It might be noted at this time that should a tape re-

Figure 11-5 SIMPLIFIED DRAWING OF THE CODING AND DECODING PROCESSES



cording be required of the signal it would have been fed to a tape recorder from the encoder at the same time as it was being fed to the modulator. This would be in digital form and possibly in a 6-digit binary code as shown in Figure 11-5.

Similarly, when the pulse groups are received at the receiving end these groups will be recorded on a Direct Recorder, such as the Ampex FR-600 with Direct Electronics, read from this transport, and passed to a correlation device, from there into a format control buffer to be read out into a digital recorder such as the Ampex FR-300, or straight to a digital computer. Should the analog form of the signal be required in place of the digital form, the signal would have been read from the Direct Recorder into a filter, sampler, de-compressor, and decoder. Its final form would have been the original input signal from the transducer or other signal generating device. A simplified illustration of the system of coding and decoding is illustrated in Figure 11-5.

Since pulse code modulation is a digital system which depends upon the presence or absence of pulses to convey information, the pulses are usually recorded on a tape transport with head current greater than that required to saturate the oxides. Recent investigations into the recording of pulses at saturation level have shown that data can be lost in the decision-making logic which is necessary prior to recording. In addition, at high packing densities, self-erasure of the tape tends to weaken the recording process. Packing densities up to a level of 200 changes-per-inch of tape may best be handled by saturation techniques (digital). When the requirement for packing densities is in the neighborhood of 6,000 per inch, however, this requirement will be best handled by the direct recording process. There is a gray area between these two densities in which there is a cross-over point not precisely identified but probably in the neighborhood of 800 to 1,000 level of changes-per-inch.

The inter-track time displacement error which is introduced by all recording equipments is detrimental to pulse recording reliability. Therefore, the time correlation capability of the equipment must of necessity affect the choice of coding technique. When track-to-track correlation is good, NRZ coding with clock track is practical. When correlation is <u>very</u> good, parallel PCM is practical. When there is inadequate correlation, serial recording of a self-clocking code is used.

In tests recently held, serial PCM recorded at saturation and non-saturation have disclosed that both methods are reliable, but that direct recording permitted standard electronics to be used with less cross-talk. Both methods recorded at 1300-bits-per-inch yield error rates less than 2 in 10^5. Using two tracks the error rate can be divided in half. Investigations have also indicated that unidirectional pulse techniques are an advantage since they reduce the need for logic circuits after the reproduce electronics.

Where an application requires 7 tracks, 2 of which being serial PCM and the balance being for direct recording of analog information, the use of non-saturated recording of PCM will permit the same type of amplifier to be used for each of the 7 tracks. This provides excellent interchangeability.

When using the Ampex FR-600 with 500 kc electronics, tests have demonstrated that packing densities up to 10,000 flux-changes per inch can be obtained with excellent results.

Where a program requires a packing density of 6,000 changes per inch, a standard 250 kc FR-600 with direct record electronics can be used. Through the use of redundant recording on two adjacent channels, a reliability of 1 in 10^5 errors can be achieved.

One of the most critical considerations in any PCM program is the tape medium itself when it is subjected to large environmental changes. During tests, the use of a reproduce head provided with means for azimuth adjustment and the provision of individually adjustable delay lines have to great extent overcome some of the problems introduced into a PCM system by inter-track time displacement error (ITDE), and even poorly recorded tapes could be read with the ITDE kept well within limits for computer inputs.