Magnetic Recording: Analog Tape

Until recently, there was one dominant way of recording sounds so that they could be reproduced at a later time or in a different location: analog magnetic recording (of course, there were mechanical methods like phonograph records, but they could not be recorded easily). In fact, magnetic recording techniques are still the most common way of recording signals, but the encoding method is digital. Magnetic recording relies on the imposition of a magnetic field, derived from an electrical signal, on a magnetically susceptible medium that becomes magnetized. The magnetic medium employed in analog recording is magnetic tape: a thin plastic ribbon with randomly oriented microscopic magnetic particles glued to the surface. The record head magnetic field alters the magnetic polarization (not the physical orientation) of the tiny particles so that they align their magnetic domains with the imposed field: the stronger the imposed field, the more particles align their orientations with the field, until all of the particles are magnetized. The retained pattern of magnetization stores the representation of the signal. When the magnetized medium is moved past a read head, an electrical signal is produced by induction. Unfortunately, the process is inherently very non-linear, so the resulting playback signal is different from the original signal. Much of the circuitry employed in an analog tape recorder is necessary to undo the distortion introduced by the non-linear physics of the system.

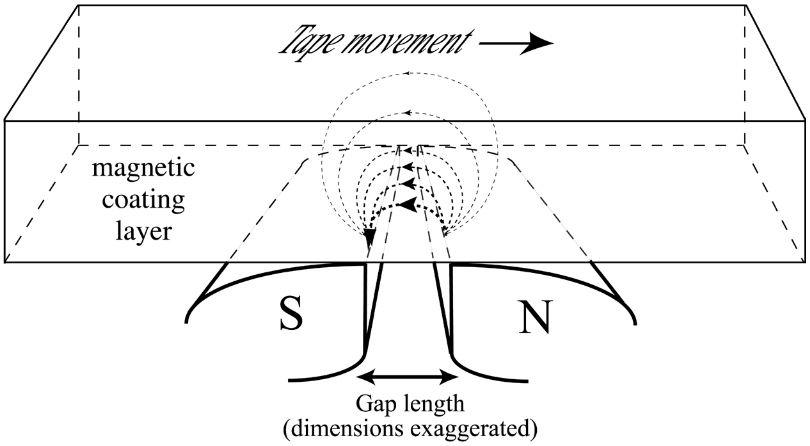
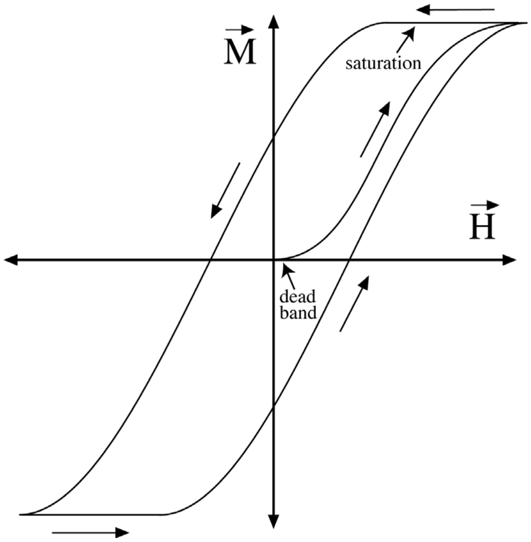


Figure 1: In a record head, magnetic flux flows through low-reluctance pathway in the tape's magnetic coating layer.

Magnetic tape Magnetic tape used for audio recording consists of a plastic ribbon onto which a layer of magnetic material is glued. The magnetic coating consists most commonly of a layer of finely ground iron oxide particles, but chromium dioxide and barium ferrite have also been used. Current production of magnetic tape uses polyester (polyethylene terephthalate) about 35 μ thick as the film substrate, a front surface magnetic coating of about 15 μ and a thin back coating to reduce friction and dissipate static electricity. The particles are acicular (needle-like) in shape and approximately 0.5 x 0.1 μ in size, but each particle may contain more than one domain. The individual magnetic domains that act as separate magnets vary in size, shape and orientation within the particles. For analog magnetic tape, the particles are oriented parallel to the surface of the tape since the system uses longitudinal magnetization to write the signal to tape. The magnetic particles are attached to the substrate with a binder, a kind of glue that fixes the particles to the tape and provides lubrication and a smooth surface that minimizes head wear. The chemical makeup of the binder is important, as it is the primary material susceptible to failure that can render the tape unplayable. It is important to understand that the particles themselves do not move with magnetization, only the magnetic polarization of the domains is changed by the applied field.

The recording process The record head converts an electrical signal into a magnetic field that can be used to create a pattern of magnetization in the tiny magnetic particles of the tape. The head consists of a toroidal core with a small air gap. A coil of wire is wound around the core, which is made of a magnetically permeable metal. Much like a transformer, the record head converts an electrical signal into a changing magnetic field. Since the gap in the head exhibits a high reluctance (the magnetic equivalent of resistance), the field flows through the lower reluctance of the magnetic tape as it moves past the gap. (See Figure 1) As the tape moves away from the gap, the magnetic flux (the magnetic equivalent of current) decreases with the inverse square of distance. At some distance from the gap, it is no longer strong enough to affect the magnetic particles on the tape and the magnetization pattern then present is retained. This means that the actual recording takes place at the so-called "trailing edge" of the gap, rather than over the entire gap length.



H=applied field   M=retained field

Figure 2: Hysteresis loop. The applied field magnetizes the material until no unmagnetized domains remain. When the applied field reverses, the material remains magnetized until the reversed field strength exceeds the coercivity of the material.

The process of recording information to magnetic polarizations involves the interaction of an imposed magnetic field with a magnetizable layer on the tape. The process is not inherently linear: as the imposed recording magnetic field increases from zero, it at first has no effect on the magnetic particles (See Figure 2.) As the field strength increases, it begins to magnetize some particles. For some amount of signal level increase, the magnetization left on the tape increases linearly. At high levels, there are fewer and fewer magnetic particles left to magnetize and the tape becomes saturated until no unmagnetized particles are left. When the imposed field changes direction, nothing happens to the magnetized material until the reversed applied field strength exceeds the coercivity of the material. Eventually, the reversed imposed field begins to reverse the magnetic polarization previously stored. This effect is called hysteresis: tape magnetic domains are not linearly changed by the imposed signal but "remember" their previous state until the reversed field is strong enough to re-magnetize them.

The overall magnetization of tape is a summation of the individual domain magnetic fields. Since there are so many domains, the total field appears to be continuous however it is really a population effect in which the individual domains are so numerous that the tape seems to be continuously magnetized. The goal of tape manufacturing is to maximize the apparent continuity of magnetic remanance. Variations in the random distribution of particles and domains result in the background noise inherent in the medium.

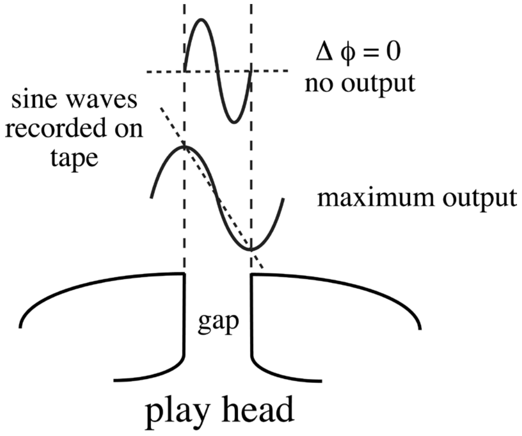


Figure 3: In the play head, the rate of change of the gap flux determines the output voltage. When the recorded signal wavelength equals the gap length, the average flux is 0 and there is no voltage output.

The playback process The reproduction process is conceptually the reverse of the recording process: as the magnetized tape is moved past the reproduce head gap, it's changing magnetic field induces a flux in the head. This flux then induces a current to flow in the coil of wire which is wrapped around the head core. Unlike the recording head, the length of the playback head gap is critically related to the ability of the head to reproduce the frequencies recorded on the tape. This is because the flux generated in the gap is gathered over the entire gap length rather than at the trailing edge as in the recording process. This means that a recorded signal with a wavelength just equal to the gap length produces no net flux, and is therefore not reproduced at all! (See Figure 3.) Hence, there is an inherent high-frequency roll-off as the signal wavelength approaches the gap length.

The equation relating the reproduced voltage (V) and the flux seen by the head is:

$$V = -N\frac{\Delta\Phi}{\Delta t} = -Nv\frac{\Delta\Phi}{\Delta x}$$

where Φ = gap flux , t = time , v= velocity, x = position, and N = number of turns of wire in the head. This shows that the output voltage is a function of the rate of change of flux on the tape. Since this rate of change increases with increasing frequency of the recorded signal, the output voltage increases with increasing frequency at constant amplitude. (It will also increase if the tape is pulled past the head faster.) This results in a 6 dB / octave high-pass filter being created. The low frequency roll-off created by the physics of the reproduce process requires that the reproduced signal be equalized with a 6 dB/octave low-pass filter in order to restore the original signal.

Definitions

recording field : the magnetic field produced by a record head when electric current is applied.
Units = A/m or Oe (oersteds): $1\ \text{A/m} = 4\pi \times 10^{-3}$ Oe

remanence : the amount of field left on tape by recording.
Units = A/m or Oe.

flux : magnetic equivalent of current.
Units = Wb (webers) or Mx (maxwells):$1\ \text{Wb} = 10^{8}$ Mx

flux density : measure of flux per unit area of magnetized material.
Units = T (tesla) = Wb/m^2 or G (gauss) :1 T = 10^4 G

retentivity : a measure of flux remaining after the magnetic field has been removed.
Units = same as flux density.

coercivity : magnetic field strength required to completetly demagnetize a material.
Units = same as recording field.

<u>Analog</u> <u>versus</u> <u>digital</u>: There are two approaches to recording audio information onto magnetic tape: analog and digital. In summary, analog magnetic recording uses the electronic signal to create a magnetic field whose strength is proportional to the voltage of the signal. This magnetic field is then used to magnetize a strip of moving plastic film onto which a coating of magnetic particles has been glued. As the tape moves past the record head, it magnetizes the particles on the tape, leaving a magnetic pattern as the tape moves away from the head. The recorded tape, when pulled past a playback head, reproduces the original electronic signal at the output of the reproduce head after corrective equalization. Analog recording needs to be a linear process where magnetic flux varies continuously in proportion to the electronic signal amplitude. In addition to the signal we wish to record, a high frequency sinusoidal signal called a bias must be mixed in before the signals are sent to the record head in order to linearize the process. The bias helps reduce the distortion of low-amplitude inputs that would otherwise distort the reproduced signal. We will discuss bias in depth below.

Digital magnetic recording is different in that the electronic signal is first measured electronically and converted into numbers which encode the signal voltage at a particular time interval, the sample rate. The numbers are processed and then written as binary data, which is recorded to the magnetic tape as magnetic polarizations of two distinct polarities which code for 1s and 0s. Because digital magnetic recording uses only two states, it does not require the use of bias current to reduce the zero-crossing distortion encountered in analog linear magnetic recording: only the saturation areas are used to encode the data. In order to achieve the high data density required by digital recording, the magnetic signal is recorded vertically (perpendicular to the tape) rather than longitudinally as it is for analog recording.

Digital recording requires plenty of signal conditioning in order to make the accuracy of the process acceptable. Since the digital data is recorded as single bits coded as magnetic polarizations, the code must somehow include markers indicating the start of a word and break up the data into frames which include not only the audio data but information used for error detection and correction. Digital magnetic recording is only feasable with extensive provisions for correcting the errors which are inevitable and would otherwise corrupt the data. The processing of signals necessary for conversion to digital representation is quite complicated and will be investigated at length later.

Analog Tape Recorders

The process of analog magnetic recording, while conceptually simple, is full of technical compromises due to the complicated physics of the process of transferring magnetic energy from the tape head to the tape and back again. The physics of recording is quite different from the physics of the playback process, thus a different set of trade-offs is presented by each. The overall process must allow the transfer of audio information with a minimum of distortion and added noise.

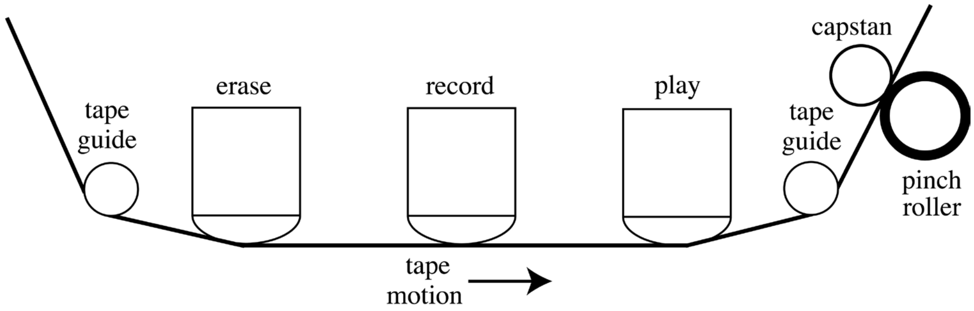


Figure 4: Typical tape head and transport assembly. Tape passes over erase head first, then record and finally the play head. The pinch roller presses the tape against the rotating capstan to maintain contact, ensuring constant tape speed.

The Record Head

As discussed previously, the actual recording occurs at the trailing edge of the record head as the imposed magnetic field strength decreases with the square of the distance from the gap. When the recording field strength falls below the coercivity of the tape, the signal is permanently recorded. This occurs at a finite distance from the gap. In fact, the effective gap length at the record head is about 14% larger than the physical gap length because the magnetic field spreads out in space. This spread can also result in crosstalk between adjacent channels when the record head is used for both recording and playback in the process of overdubbing.

Bias Current: The recording process is complicated by the need for a bias current to reduce the non-linear effects of magnetic tape's transfer function. The distortion is a result of the fact that the magnetic domains of the tape magnetic coating have a minimum applied field strength (threshold) below which they are not able to be magnetized, so low-amplitude audio signals cannot be transferred to tape magnetization. By mixing a high-frequency sine wave with the audio signal, the resultant signal amplitude keeps the recording head magnetic field magnitude large enough to avoid the dead-band region of the transfer curve (See Figure 5.) (The bias itself may be distorted by the dead-zone, but the analog signal is not affected.) The bias signal is not modulated by the audio signal, but it can still act as a carrier because the record field extends beyond the head gap and may be picked up by adjacent tracks in a multi-channel head used for tape monitoring while overdubbing. Since the bias signal is much higher in frequency than the audio signal and is mixed in linearly, it can be removed with a simple filter, the bias trap. Also, due to its high frequency, the bias signal is not efficiently read by the playback heads. The use of bias current dramatically reduces the distortion of the recording stage.

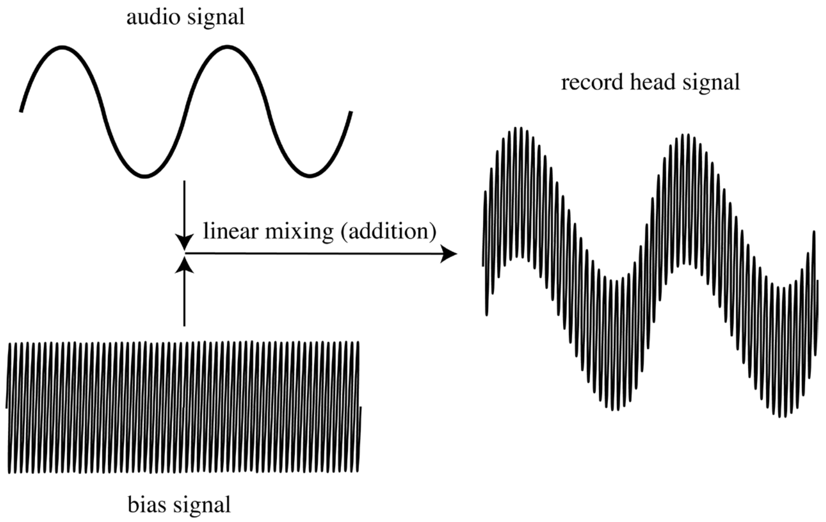


Figure 5: A bias signal is added to the audio signal before it is delivered to the record head gap. The bias is filtered out of the reproduce signal, restoring the original audio signal and reducing the distortion caused by the minimum amplitude necessary to magnetize the tape.

Several characteristics of the recording process are affected by the amount of bias used, including sensitivity, noise, frequency response and distortion. (See Figure 6.) Setting the bias level gives the engineer the ability to optimize whichever characteristic is most important to them. Each track of a multichannel recorder has an independent bias setting, so theoretically each track can be optimized for the particular signal being recorded. Different tape formulations have different curves, a factor reflected in the difference in the sound of the various tapes.

<u>Equalization</u>: In addition to the bias adjustment, recorders have a record equalization circuit that allows the input signal to be equalized for optimal frequency response. Equalization can be applied in both the record and playback circuitry. Generally, the playback equalization is adjusted using a calibration tape and the record equalization is then adjusted to provide the flattest playback. This is done for each track independently.

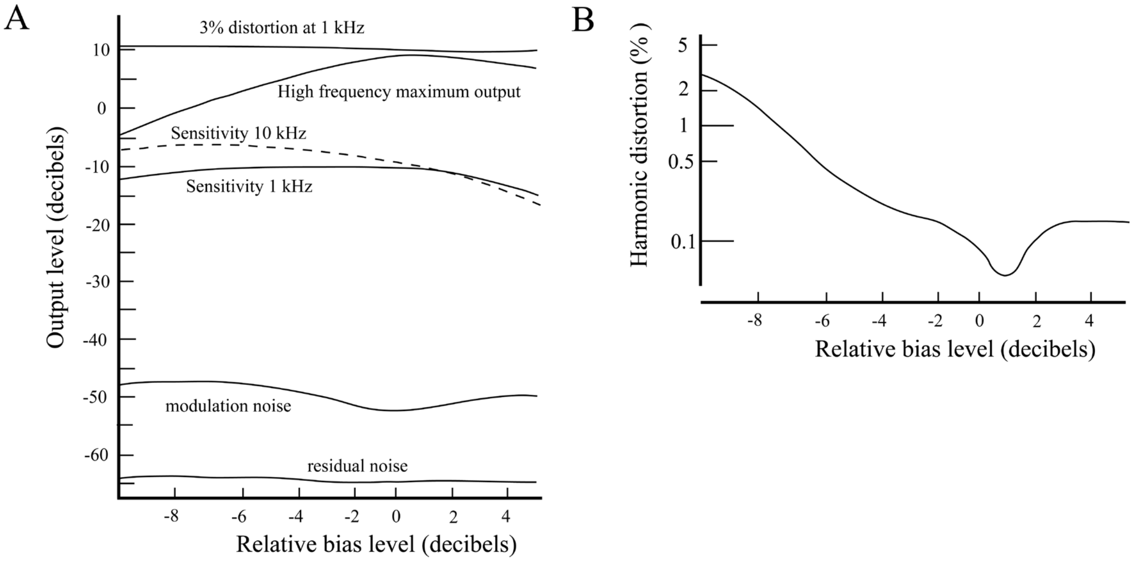


Figure 6: The effect of bias level on recording sensitivity, noise and distortion. The optimal bias for each consideration is different, so trade-offs are necessary. The curves differ for different tape formulations.

Tapes are usually recorded with sine tones of known frequency at the beginning of the tape so that it can be reproduced on any tape machine with correct frequency response adjustments by the playback equalization. This allows tapes made on one machine to properly play back on a different machine.

Self Erasure: One kind of high frequency loss that takes place at the record head is self-erasure. This phenomenon takes place as the magnetized tape leaves the gap area. A secondary (or phantom) gap is created between the record head and the newly-magnetized tape. This gap field tends to erase the high-frequencies on the tape. It is made worse if the head is magnetized, so we must be careful to demagnetize the heads regularly.

Recording Head Gaps: The length of the recording head gap is chosen to allow the magnetic field to optimally penetrate the magnetic coating: the gap is generally about equal to the thickness of the magnetic coating. (Record head gaps range from 2.5 to 12 microns.)

The Erase Head: A special record head is used to erase the tape before a new recording is made. The erase head has a larger gap length, since it must completely demagnetize the entire thickness of the tape to remove any remaining signal. The gap length used in erase heads can be 3 to 4 times the thickness of the magnetic layer. Pure bias current is used to remove any remanent magnetization from the tape and randomly magnetize the particles. (Erase head gaps range from 25 to 125 microns.) The tape passes over the erase head before it passes over the record head. (See Figure 3.)

The Reproduce Head

The playback or reproduce head reciprocally converts the magnetic pattern on tape to an electrical signal. As the tape moves over the play head gap, it induces a magnetic flux in the head, which is wrapped by a coil of wire. The tape must be moving because the electrical output of the head is proportional to the rate of change of the flux, not its instantaneous amplitude. The play head output is determined by the changing flux encountered by the head gap, so a sinusiodal pattern exactly the length of the gap produces no output since the total flux change sensed by the gap is zero (See Figure 3.) This sets an upper frequency limit for the system. This phenomenon creates a trade-off, since more flux is gathered by a larger gap but high frequency response is limited by a large gap.

The equation approximating the reproduce head output voltage for a sinusoidal magnetic signal is:

$$e(x) = -\mu_0 V w M_0 (H_g g/i)\, k\delta\, [e^{-kd}]\, [(1-e^{-k\delta})/k\delta]\, [\sin(kg/2)/(kg/2)] \cos(kx)$$

where: x = longitudinal position

$e(x)$ = voltage output from longitudinal magnetic recording
μ_0 = magnetic permeability of a vacuum
V = tape to head velocity
w = track width
M_0 = peak value of sine-wave magnetization
H_g = deep gap field
g = gap length
i = current in head coil
k = wavenumber ($= 2\pi / \lambda$) where λ = wavelength
δ = thickness of the magnetic medium
d = distance from tape to head

The terms in square brackets are loss terms relating to specific physical relationships which act to reduce the output voltage at the reproduce head:

Spacing Loss (decrease in output signal due to distance from tape to head):

$L_d = e^{-kd}$

Spacing loss increases exponentially with increasing distance as a ratio of the wavelength of the signal, thus high frequencies are more susceptible to dropouts. (Spacing loss can be expressed as 54.6 dB / wavelength distance, meaning that almost 60 dB of dropout is produced if the tape is separated from the head by the wavelength of the signal of interest [~19μ @20 kHz & 15 ips]!)

Gap Loss (decrease in output signal due to the length of the gap)

$L_g = \sin(kg/2)/(kg/2)$

Gap loss reflects the fact that the reproduce head responds to the average flux in the gap: therefore, if the wavelength of the signal just equals the gap length, there is no signal produced. At low frequencies, this also results in a series of peaks and dips in the frequency response known as "head bump".

Thickness Loss (decrease in output signal due to the thickness of the magnetic medium)

$L_d = (1-e^{-k\delta})/k\delta$

Thickness loss is not as severe as spacing loss, but it does indicate that the thickness of the magnetic layer is real and does affect the output signal to a measureable degree.

Reproduce Head Gaps: The gap length of the reproduce head must be optimized to allow the greatest signal recovery without compromising the highest frequency (shortest wavelength) that we need to preserve. Reproduce head gaps are the smallest of the three types of heads: the average output level is compromised somewhat to allow higher frequencies to be reproduced. The reproduce head gap length is more critical than are record head gap lengths. (Reproduce head gaps range from 1.5 to 6 microns.)

Equalization

Because the process of recording and reproducing signals on magnetic tape is a frequency-dependent process, equalization is used to restore flat frequency response within the limits imposed by the physical processes already discussed.

Since reproducing a tape recording is a differentiating process, there is a 6 dB/octave rise in the amplitude of the reproduced signal. Equalization can be applied during the record process and during the reproduce process to correct the resultant frequency response. It is desirable to boost frequencies during record, since this allows us to cut frequencies during playback, however we must be careful not to boost frequencies enough to cause tape saturation. Since the signal at playback contains noise and distortion produced by the recording process, filtering (cutting some frequencies) during playback provides a way of reducing the unwanted components: they are filtered out by the post-equalization.

There are several equalization standards, NAB and IEC being the most common. Recorder manufacturers specify which is intended for a particular machine and calibration tapes using that standard are used to align the recorder heads and electronics.

Head alignment

The geometric alignment of the head with the tape is a critical adjustment. The gaps must align perfectly with the previously recorded tracks in order to reproduce the recorded signals properly. All three dimensions need to be correctly aligned to optimize playback.

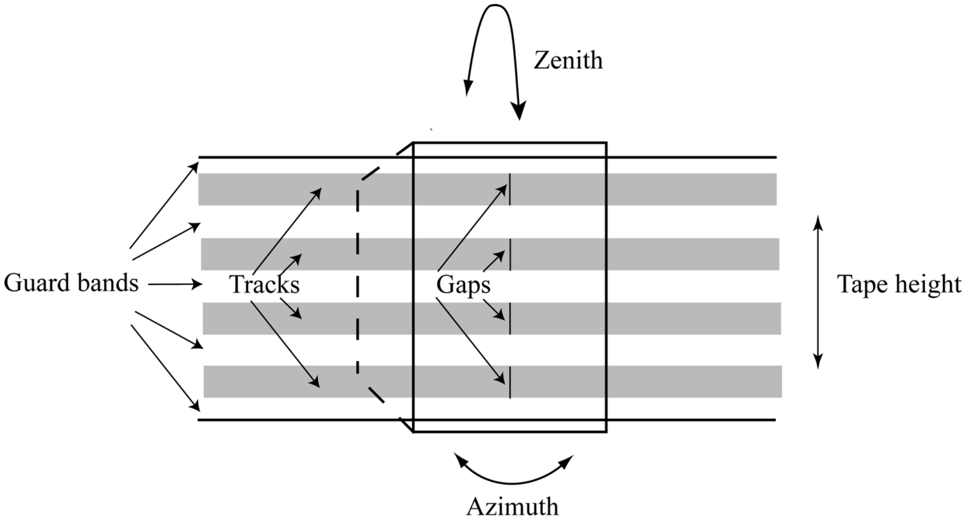


Figure 7: Tape head alignment. Zenith adjusts the face of the head to be parallel to the tape while azimuth adjusts the head gaps to be perpendicular to the tape travel. Height guarantees the tracks and gaps are aligned vertically.

Tape head alignment is established using a calibration tape with sine tones recorded across the entire width of the tape. First, the playback azimuth is adjusted using the calibration tape and the equalization is set to be flat. Then, tones are recorded and played back. The outermost tracks are fed to an X-Y oscilloscope where they can be monitored as the azimuth is adjusted: when correct the scope shows a 45° angle. Zenith and tape height are more complicated to adjust but rarely need re-adjustment. Azimuth is adjusted often, particularly when tapes recorded on one machine must be played back on another.

Tape transport

Because the voltage produced by the play head depends on the rate of change of the flux in the gap, the tape must be pulled past the head at a fixed and constant rate. This may sound simple, but as the tension of the tape against the heads must remain constant as it moves, the mechanical system used to regulate the motion of the tape must be capable of maintaining constant speed and tension regardless of changing conditions as the tape moves from the supply reel to the take-up reel. Early tape recorders used mechanical systems to keep tape moving. More recent tape machines use servo systems to drive the motors electronically. In many cases, a disc with an optical pattern of fine lines is used to measure the rate of rotation of the capstan motor, which pulls the tape over the heads. In many machines, a pinch roller is used to clamp the tape to the capstan while some newer systems use entirely electronic systems, balancing back tension on the supply and take-up reel motors. These electromechanical systems must be checked from time to time to prevent wow and flutter as well as speed variations over the length of each tape.

Mechanical resonances can develop as tape slides over guides intended to keep the tape at the correct height as it passes over the heads. Some machines use mechanical means to damp any vibration of the tape that might occur between the heads.

Tape speeds are standardized, with 15 inches/second (ips) and 30 ips in professional recorders and 7 1/2 and 3 3/4 ips for general purpose tape recorders. Cassettes run at 1 7/8 ips. Since playback is determined by the rate of change of the flux, faster tape speeds create higher output signal levels and give better high-frequency response. The frequency of the low frequency ripple known as head bump is also linked to tape speed. Head bump is a damped oscillation in the low frequency output created by the limited length of the head relative to the wavelength of the signal on tape. In a sense, it is a sampling error. Head bump can lead to rather dramatic changes in the low end frequency response of a recording and varies significantly with the physical design of the play head. It is likely part of the reason analog tape is often considered “punchy”.

Multitrack recorders

The most popular use for analog recorders is for multitrack recording. Multitrack recorders are capable of recording several separate audio tracks in synchrony and are able to independently record and play back each of the tracks. In order to do this, the machines must have some important features: each track must have independent erase and record functions and they must be able to play back through the record head to keep previously recorded outputs and new inputs synchronized.

In the standard analog recorder, there are three heads: erase, record, and play (See Figure 4.) Each is separated from its neighbor by an inch or more. Since the tape moves at a fixed rate, a patch of tape passes the erase head a fraction of a second before it passes the record head and then the play head. If we energize the erase and record heads simultaneously, the portion of tape between the erase and record head will not be erased. If we record a signal at the record head and play it from the play head, the playback is delayed by the time it takes the tape to move from the record head to the playback head. This introduces a time delay which would make playing (and recording) a new signal go out of synchrony with the tracks already recorded. To avoid this, we must use a record head that can also be used to play back signals already recorded on the other tracks.

Since the record signal with its bias tends to spread out in space somewhat like a radio wave, it tends to leak into adjacent playback gaps when we use the same physical head to record and play at the same time. We have some leakage audible as the price for synchronous monitoring. Of course, when we’re done recording, the problem is no longer a worry as we play everything without a recording signal.

The number of tracks that maybe recorded on a multitrack depends on the arrangement of the heads and the width of the tape. Standard widths for reel-to-reel systems go from 1/4” to 2”. It is possible to fit as many as 8 tracks on a 1/4” tape, but with poor fidelity. The standard 24-track 2” recorder corresponds to a track width of about 0.037” while a 1/4” 2-track mastering recorder has a track width of about 0.080”. A cassette has a track width of only about 0.021”. The width of the track will determine the maximum signal strength possible, but the quality of the head construction and electronics employed can also affect the sound quality.

So how well can an analog magnetic recorder perform? Often cited as among the best available recorders, the Ampex ATR-100 series provide a signal-to-noise ratio of 72 dB @ 30 ips and with a frequency response of 35-28 kHz +/- 2 dB. This is referenced to the input level that produces 3% harmonic distortion, a figure that would not be acceptable in a digital system. While this may sound like a lot of distortion, it can be surprisingly difficult to hear - loudspeakers often provide even more distortion while sounding acceptable.

Because of the inherent noise in the system, due both to the tape and the electronics, the addition of noise-reduction systems was popular with multitrack analog recorders. These systems worked by compressing the signal before sending it to tape and then expanding the signal on playback, sometimes by also pre-equalizing and post-equalizing as well. This has the effect of reducing noise added in the record/playback process without altering the signal, in theory at least. With the development of quieter, high-output tapes, the use of noise reduction may not be necessary, particularly for popular music recording. Now the big question is how long will analog tape be manufactured? With more and more engineers using computer-based recording systems, the days of analog tape may be numbered. Where there once was a wide choice of formulations, there are only a handful of magnetic tapes manufacturers left. But even in the era of digital audio, there is something considered desirable about the sound of analog magnetic recording: many plug-ins are available to simulate the sound of analog recorders.