CONSUMER VIDEO TAPE DUPLICATION TECHNIQUES

A TUTORIAL

by Jim Brown
SOUND ENGINEERING ASSOCIATES
CHICAGO, ILL.

CONSULTANTS TO

BELL AND HOWELL/COLUMBIA PICTURES VIDEO SERVICES

ABSTRACT

This tutorial paper begins with a survey of the audio parameters of VHS, Beta, and 8 mm videocassette formats, and describes the audio technology necessary for successful operation of a duplication facility. Topics discussed are mastering, duplication, distribution, routing, and signal processing equipment, system design, standards of operation, program interchange, and quality control.

INTRODUCTION

Consumer videocassette duplication offers a complex, but not totally unique set of challenges to the audio engineer. Home recorders and players are sold for a small fraction of the cost of professional video recorders. Tape formulations are optimized for video, with major compromises being made in audio performance. Tape speeds are low, placing serious limitations on what can be achieved with noise, headroom, and distortion. New systems which allow multiplexing of the audio with the higher speed, higher bandwidth video signal offer significantly improved performance but add new pitfalls. A brief overview of the audio recording systems in these formats is fundamental to an understanding of the problems and the techniques used to solve them.

FORMATS

The two videocassette formats (VHS and Beta) which currently dominate the consumer market will be discussed in detail, but the principles will apply to others (including the 8mm format) as well. Both operate on 1/2 inch magnetic tape, and at very low tape speeds. Helical scanning by rotating video heads allows high writing speeds for video, yielding resolutions of 240 lines with 45 dB signal to noise in both formats. Relatively narrow audio tracks accessed by stationary heads are used, limiting audio bandwidth to about 7 KHz for Beta and 10 KHz for VHS. Signal to noise (referred to 3% THD) is on the order of 48 dB without noise reduction. Although both formats allow for recording at several speeds, virtually all commercial software is recorded at a single speed (one of the higher ones) for each format. A summary of typical system and performance specifications appears as Figure 1a.

AUDIO ON VIDEO RECORDERS

Analog magnetic recording is limited by the distortion produced from peak levels that can be recorded on tape as compared to the noise of biased, unmodulated tape. This signal to noise ratio is improved by greater track widths and high energy-low noise tape formulations. Signal to noise is approximately proportional to the square roots of both track width and tape speed; and distortion rises rapidly with increasing signal level. Thin oxide coatings required for good video performance cause increased distortion at low frequencies. Operating at slow tape speeds makes recording of high frequencies difficult. While this limitation can be partially overcome by pre-emphasis, the price of increased distortion and reduced signal levels at high frequencies must be paid. Head alignment, bias, and equalization all have major impact on amplitude and phase response, as well as the noise and distortion performance. Many video recorders (even professional ones) provide no adjustments for some or all of these parameters, and those adjustments which are provided are often inadequate.

Audio performance of videotape recorders thus suffers from three fundamental limitations. First, narrow track widths are used. Compared to a professional 2-track audio recorder having a track width of 1.9 mm, the C-format recorder suffers a 3.8 dB disadvantage. Second, videotape is very thinly coated, with formulations optimized for scanning by rotating video heads. The flux level for 3% THD for audio recording on videotape is currently 320 nWb/m, 10 dB lower than the 1040 nWb/m formulations currently used in master quality audio recording. Third, recording speeds are lower than for audio-only recording. B or C-format recorders, which operates at 9.6 inches/sec, suffer a 2 dB disadvantage compared to a professional 2-track audio recorder at 15 ips. The combined theoretical disadvantage of the C-format machine is 16 dB for equal levels of noise, distortion, bandwidth and high frequency headroom as compared to a 2-track quarter inch professional audio recorder. The same logic puts the linear tracks of the VHS video cassette at a 4.5 dB disadvantage compared to the 1 7/8 ips, 1/8 inch audio cassette, assuming tape of equal quality when used for audio. Naturally, the compromises made between those various performance parameters are made differently for each system.

VHS

The VHS format (ANSI Type H), was announced by JVC in 1976. Embraced by the vast majority of consumer manufacturers, VHS now holds a 4:1 lead in hardware and software sales over its only serious competitor, the Beta format. The VHS format allows for a total audio track width of 1.0 mm. It may be recorded monaurally or split into a stereo pair of 0.35 mm each, with a guard band of 0.3 mm between them. There is no guard band between audio track two and the tape edge, making the format very sensitive to tape width uniformity (slitting) problems. Recording pre-emphasis and playback de-emphasis are employed, using time constants of 120 and 3180 microseconds. Commercially available tapes are recorded at 33.35 mm/sec (1.313 ips). Because these tracks are recorded by stationary heads which produce tracks parallel to the tape path, they are referred to as "linear" or "longitudinal" tracks. Dolby B Type (the orginal consumer type) noise reduction is available on many consumer machines as well as on most duplicators.

BETA

The Beta format (ANSI Type G) is supported by only a few manufacturers, but one of them is the giant Sony Corporation, which introduced Beta I in 1975. Now used only in government and industrial applications, Beta I provides a stereo track of 0.35 mm with a 0.3 mm guard band between tracks and 0.09 mm between the track tape edge. A level sensitive compansion system with pre-emphasis, called Beta Noise Reduction, is used. When no noise reduction is used, pre-emphasis is 3180 and 50 microseconds. Reference level is 125 nWb/m.

Consumer Beta machines support the Beta II (20 mm/sec) and Beta III (13.3 mm/sec) formats. Virtually all linear audio is recorded monaurally on a 1.05 mm track, with all duplicated software produced in Beta II (20 mm/sec, 0.79 ips). Pre-emphasis and de-emphasis of 175 and 3180 microseconds is used at this speed. Reference level is 100 nWb/m for both Beta II and III. Some consumer machines will play tapes recorded in the Beta I format but have no record capability at that speed.

A very limited volume of industrial duplication is done for a special variation of the Beta II format. Used primarily by the airline industry for in-flight movies, these specially manufactured machines have stereo linear audio tracks with Beta Noise Reduction. There is no home market.

COMPANSION

Compansion is a generic audio process, whereby audio is compressed when recorded and expanded when played back. The compression and expansion processes are carefully matched to each other, so that under ideal conditions, neither compression nor expansion results from the total process. Low level audio is recorded at higher than normal levels, and high level audio is recorded at lower than normal levels. Noise within the compansion loop (i.e. between the compression and expansion stages) is reduced in relation to the signal, and distortion caused by overmodulation is prevented. Compansion is widely used in satellite communications, professional tape recording, and the new BTSC stereo TV transmission system. Dolby, dbx, and Telcom are examples of different approaches to compansion.

The proper operation of a compansion system depends on careful matching of the compression (encoding) and expansion (decoding) systems. Any non-linearity within the compansion loop causes mismatching. Possible non-linearities are frequency response errors and amplitude distortion. Any components of the signal fed to the decoder which were not in the encoded input (hum and noise, for example) also cause mistracking. The audible sound of this mistracking is most often breathing, usually with a noise "fur" surrounding the audio.

BETA HI-FI

The audible mediocrity of the Beta II format quickly pushed Sony to develop Beta Hi-Fi Audio. The audio modulates FM carriers which coexist with the video signal in the video recording system. A 2:1 compansion process (similar in concept to dbx noise reduction) with RMS detection and low and high frequency pre-emphasis is used to optimize performance. The manufacturer claims a "consumer weighted" signal to noise ratio of 80 dB with harmonic distortion of less than 1% below clip. Audio bandwidth was extended to 20 KHz.

VHSHD

The VHS forces quickly followed suit with their own modulation scheme. The VHS HD system modulates the tape differently from the Beta system, using separate heads for HD audio, but uses similar compansion with RMS detection and pre-emphasis. Audio bandwidth is 20 KHz, signal to noise is 80 dB, and harmonic distortion is less than 1% below clip. The audio side of consumer video, it appeared, was finally capable of even better performance than the video. Because Beta Hi-Fi appeared first, the VHS HD system is often referred to in the generic sense as a "Hi-Fi" type of format. That practice will be followed in this paper, with the understanding that references to "Hi-Fi" formats include the VHS HD system. Beta Hi-Fi recorders for the European PAL standard use the VHS depth modulation scheme because the recorded video spectrum does not allow room for the insertion of the Sony Hi-Fi carriers.

8 MM VIDEOCASSETTE

The latest arrival on the scene is the 8 mm videocassette. This system boasts video quality which is roughly comparable to VHS and a monaural Hi-Fi audio track. While the format allows for longitudinal audio as an option, no manufacturer currently supports it. Stereo audio is provided by the addition of a separate PCM encoder/decoder. The audio is 8-bit encoded, companded, and FM modulated into the video passband. A 31.5KHz sampling rate is used. Claimed audio performance specifications are comparable those of VHS HD. Consumer duplication is now taking place in this format.

PROBLEMS WITH THE HI-FI FORMATS

Like most systems of this complexity and low cost, there are some problems. The Hi-Fi formats are sensitive to poor alignment of the video recording system (or, in the case of VHS HD, the HD audio heads). Less than ideal mechanical alignment of these heads causes interruptions in the video signal at the vertical frequency (60 Hz). The video signal is not seriously affected until mis-alignment is large, since the head switching occurs when the scan is off the screen at the bottom of its trace (just before the vertical sync interval).

The audio signal, however, being continuous, will contain noise as a result of imperfect head switching. This noise appears as a harsh buzz (30 Hz and its harmonics) and is present at the input of the expansion system upon playback. Although its audibility is reduced by the compansion system and masked by complex program material, the buzz also causes errors in the decoding of the companded audio, particularly when audio levels are low and alignment is poor. At its worst, the buzz will be heard modulating relatively low level audio. The resulting signal to noise ratios can often be 10-20 dB worse than the wonderful 80 dB numbers quoted by machine manufacturers.

Another problem with the Hi-Fi formats is inconsistency in the maintenance of standards by the video industry. There are minor variations in the pre-emphasis between VHS machines of different manufacturers, and there are major differences in the pre-emphasis used in different generations of Beta Hi-Fi machines of the same manufacturer (Sony). Literally millions of tapes made on early Beta Hi-Fi mastering recorders exhibit very noticeable breathing (i.e. compansion decoding errors) and frequency response errors when played back on machines of later vintage.

There are no published specifications of which the author is aware that attempt to make the consumer or duplicator aware of these inconsistencies, nor is there any specification to define this compansion process. Instead, a new signal to noise weighting system called "consumer weighting" was invented to obscure the problems of Hi-Fi audio modulation schemes which legitimate specifications would clearly expose. A new level of specsmanship has clearly been reached! Figure 1b summarizes typical measured performance specifications of selected Hi-Fi formats.

THE CONSUMER

The videocassette consumer is a relatively unsophisticated one with respect to audio. More than 95% listen to audio from the linear tracks of non-Hi-Fi recorders through the four inch speaker built into their table model television receiver, many of which face backwards. High and low frequency response of the receiver is often limited and the speaker is rarely driven by more than a five watt output stage. Figures 2a and 2b are the measured responses of the audio systems of two "good" table model consumer television receivers.

On the other hand, a small, but increasing number listen through component high fidelity systems and expect the superior performance they have paid well for. There are several systems on the market which allow the consumer to decode the Dolby Surround encoding present in the Dolby Stereo release print and feed it to a home implementation of a four channel (left, right, center, surround) playback system. Good amplitude and phase response is critical to the proper operation of those systems. The signal processing and mixing demands dictated by these radically different ways of listening are somewhat conflicting.

THE DUPLICATION PROCESS

At its current state of the art, the duplication of videocassettes (or more correctly, the mass transfer of program material to the videocassette format) consists of one of two basic processes. In the first, a one-inch B or C format professional videotape recorder plays back a master at normal viewing speed to an array of thousands of duplicator machines. This process is called in-cassette duplication. The duplicator machines involved are usually heavy duty, high specification consumer or industrial machines.

In the second method, large "pancake" reels of videotape on dedicated recording machines records up to 48 hours of programs which are then loaded into the videocassette. This recording process also occurs in real time, and uses bulk loading techniques which are common to the audio cassette industry and which will not be discussed here. Although this paper will largely discuss the first process, the parallels to a pancake operation should be readily apparent.

Two other methods of duplication currently exist, but have not gained widespead use because of practical problems with each. In the first, an anti-hysteretic process makes use of a very high coercivity (2000 Oersted) master to do what is essentially contact printing to a normal coercivity slave tape in the presence of a magnetic field. The Sony Sprinter and the Panasonic VTP system are examples of this process. It has not gained widespread use primarily because of the difficulty of generating masters, and the master's short operating life (500-1000 copies).

In the second, thermal magnetic duplication, a high coercivity master is also used. The slave tape is heated by a laser to the temperature where it has lost its retentivity, and then cooled in contact with the master and a glass prismatic pressure roller. Duplication speeds of up to 100:1 are possible. This experimental process is a joint venture of Bell and Howell and Dupont, and is not currently in production.

THE DUPLICATION FACILITY

The duplication facility consists of six basic parts. The head end consists of the professional recorders used to play back masters, signal processing for control of audio levels, and routing systems to assign programs to banks of duplicator machines. The distribution system consists of distribution amplifiers and cabling to interconnect the thousands of duplicator machines which make up the production line. Three quality control functions, (the evaluation of incoming masters, tape stock and/or cassette housings, and outgoing product) complete the functions.

THE MASTER PLAYBACK MACHINES

Most entertainment program material for commercial release is mastered in the video medium on one inch B or C format professional recorders. These formats operate at about 243 mm/sec (9.6 ips), with two audio tracks of 0.8 mm width. Reference level is 100 nWb/m. Signal to noise is 58 dB referred to 3% THD, which occurs at 320 nWb/m on the tape formulations currently in use. The distortion performance of these formats is four times higher than the Hi-Fi formats within 15 dB of clip, and increases at low frequencies because of the very thin oxide coatings used.

Dolby A Type Noise Reduction is normally used, improving signal to noise ratio to about 70 dB (Unwtd). Until its adoption as part of the BTSC system, dbx noise reduction had not attained widespread use in the video industry. This is primarily because early video recorders had considerable video to audio crosstalk and audio heads with poor frequency response, both of which caused excessive mistracking of the decoders. Although this would most likely not be a problem with modern professional video recorders, Dolby Type A has become firmly established as a standard and is unlikely to become displaced. Dolby A type noise reduction cards are available to replace the audio cards of many C-format recorders. The audio performance of some older machines which are limited by their record electronics is improved by use of these cards (in addition to the improvement to be expected from Dolby encoding). B-format recorders include Dolby A type processing as standard equipment.

Proper audio head alignment and equalization of the playback machine is critical to good audio performance. It is also necessary that these parameters match those of the mastering recorder. Dolby and dbx noise reduction systems will mistrack if playback levels or frequency response do not match those of the mastering recorder. Matching is aided by level and frequency response test signals recorded on the master tape.

DUPLICATOR MACHINES

In the early days of videocassette duplication (which was only a few years ago) the duplicator machines were simply selected production from the highest quality consumer machine available. At this time, however, cassette duplication machines are usually dedicated machines manufactured specifically for duplication use. Most have test points to speed up routine alignment, separate inputs for the Hi-Fi and linear tracks, and provisions for interface to automation. Most are fairly well supported by their manufacturers, who have a vested interest in keeping their real customer, the consumer, supplied with high quality software.

Input signal levels and impedances of available duplicators vary somewhat, but are usually on the order of 1 volt into 47Kohms unbalanced for 100% modulation of the audio track. Audio metering systems and their calibration are more unpredictable. Some are quasi-peak reading, and a few are even pre-emphasized. Each duplicator machine must be individually aligned for input level, head alignment, bias, equalization, and meter calibration.

Alignment tapes for the consumer formats are of generally poor quality, difficult to obtain, and expensive when available. There are no professional quality audio alignment tapes currently available for the Beta consumer formats. The VHS format is only somewhat better supported in this regard. This presents a major problem in the setup and maintenance of these machines, and the duplicator is left to his own devices.

ROUTING SYSTEMS

Most facilities will need several master playback machines, with the capability of running several programs simultaneously to diferent groups of duplicator machines. This allows more efficient use of the facility with low quantity duplication runs. Conventional video and audio routing switchers are used. It is often necessary to route from one master machine to another. A routing switcher capable of a 32X32 stereo matrix is appropriate for most facilities, but it is always wise to purchase one which will allow for expansion. Provision for control by an automation system is also a desireable feature.

Our facilities are designed so that each bank of approximately two hundred duplication machines is fed from a dedicated output of the routing switcher. With this arrangement, very small runs can be accommodated. A single stereo 1X8 DA located at each bank feeds the machines for that bank. While a 1X6 DA would suffice, the 1X8 is used because of operational and packaging features designed into that particular manufacturer's unit. Among these features are the provision for metering, and a stereo package with self contained power supply.

In addition to the general audio specifications summarized in Figure 4, routing switchers should conform to a crosstalk specification of 95 dB between any input and any output. Many routing switchers terminate audio inputs with 600 ohm resistors, and add 600 ohm resistance in series with outputs. While this is good engineering practice for video, it is not for audio. Input and output impedances should conform to the specifications for all audio systems equipment -- 50K (bridging) inputs and 60 ohm (low impedance) outputs. Caution is advised, however. A routing switcher designed for 600 ohm input termination may not meet its crosstalk specification (or have other problems, like oscillation) if its terminations are removed. This doesn't mean you should not run unterminated, but it does mean that you should use a routing switcher designed to be run that way.

Another aspect of signal routing is that there is sometimes a need to route different audio tracks of the master tape to the duplication line. This need occurs because not all masters conform to the same standard useage of audio tracks. Any of the three audio tracks from the master tape needs to be assignable to any audio track of the duplication line. Examples of situations that might create this need are multiple language tracks and industrial program duplication. Hardware to perform this task will often need to be built in house.

INTERCONNECTION OF AUDIO PROCESSING, ROUTING, AND DISTRIBUTION EQUIPMENT

It is no longer good engineering practice (and has not been since vacuum tube days) to build out and terminate audio system components in most installations. Figure 5b is an example of that obsolete method of interconnection. Modern input stages are high impedance and balanced, whether transformer isolated or differential. Modern output stages are themselves very low impedance (on the order of a fraction of an ohm) with a series build out resistance of about 60 ohms. This series resistance isolates the usually capacitive reactance of the line from the output stage, eliminating frequency response anomolies and circuit instability that might result if the stage saw a reactive load.

Figure 5a illustrates this method of interconnection. Since this output stage works from a low impedance source into a bridging (i. e., high impedance) load, the series resistance does not cause any voltage drop either from a resistive termination (flat attenuation) or from a reactive load (high frequency rolloff). As compared to a terminated system, this reduces the cost of individual components, simplifies system design, and allows longer lines to be driven without high frequency loss and phase shift. Cost reduction is achieved because no output voltage is lost in the resistive divider made up of series reistance and its termination (6 dB); and because many bridging loads may often be driven from a single output stage; and because most well designed output stages have somewhat better headroom when driving a bridging load (i. e., they do not have to supply power, only voltage).

In Figure 5b, the older "Power matched system", resistance (called "build-out resistance") is added to the output stage to bring it up to a 600 ohm source impedance, and a 620 ohm resistor is added to the input terminals to make it a 600 ohm load. 6 dB of power and voltage is lost in the resulting divider, and the system -3 dB bandwidth is reduced to 16.7 KHz when driving the same 1000 ft of cable. Audio cables do not become transmission lines until their lengths exceed 2000 feet, and the characteristic impedance of most in common use is much closer to 200 ohms than 600 ohms.

In Figure 5a, a typical output stage is connected directly to an input stage. All of the voltage developed in the output stage is delivered to the input of the following stage, and the stray capacity of the interconnecting cable is "swamped out" by the very low 60 ohm impedance of the output stage. In our example, the system bandwidth (-3dB point) is 83 Khz.

Some audio technicians inexperienced in RF circuits discovered that the 600 ohm resistance across the input terminals reduced the system's susceptibility to radio frequency interference (RFI). A small capacitor is much more effective in supressing RFI, and has no effect on the audio. The 220 pf capacitor is chosen if the interfering signal is VHF; a considerably larger value (0.02 uf) might be appropriate if the interfering signal is an AM broadcast signal. Neither value will degrade the audio, and both will be considerable more effective than the 600 ohm resistance in reducing RFI.

Most video facilities are set up for operation at +8 dBm, primarily because most video recorder audio outputs are set up that way. Operation at +4 dBm is more cost effective, and is recommended. The additional 4 dB of signal drive capability requires more expensive output stages and higher voltage power supplies in audio equipment and is not needed to meet system performance specifications. A DA or line driver capable of +27 dBm costs more than one capable of +23 dBm.

THE AUDIO DISTRIBUTION SYSTEM

Distribution of signals for in-cassette videocassette duplication differs from broadcast and studio systems primarily in that the duplicator machine input is unbalanced, bridging, and consumer level (approximately -10 dB re 0.775v). The video input is also unbalanced. The cost of modifying the inputs to make them balanced is prohibitive. The best practice is to locate a distribution amplifier with each bank of machines that it serves. A single 1X6 DA can easily feed several hundred machines.

Because the duplication machine inputs are unbalanced, transformers are used on distribution amplifier outputs for ground isolation. If the DA uses differential outputs, transformers must be added externally. Attenuation of 14 or 18 dB is used to reduce the +4 or +8 dBm output to the -10 dBu required. Since high impedance pads are used, more than 60 duplicators could be driven from a single output stage designed for a 600 ohm load. If half that number were driven, headroom and distortion performance of most distribution amplifiers would be improved by several dB. In our facilities, the latter practice is followed, with a typical DA feeding 15-30 machines per output. Since DA inputs are also bridging, many may be fed from a single line driver at the head end.

The distribution system should be designed for performance specifications at least an order of magnitude better than those of the best Hi-Fi formats. This allows for improvements in machine performance as well as minimizing degradation of the signal by the distribution. An appropriate design specification for the distribution system as a whole is 95 dB signal to noise (unweighted, referred to the RMS value of peaks); 0.05% THD at peak level and 0.025% below peak level; and ± 0.1 dB between 20 and 20,000 Hz.

In practice a major component of distortion can be the audio transformers used for isolation of ground loops and for input and output of system amplifiers. Jensen transformers are recommended where transformers must be used. High quality electronically balanced inputs are desireable for distribution amplifiers and in the head end. Care must be taken that true balanced inputs are used. Many of the circuits passed off as electronically balanced inputs have unequal input impedances which can cause problems in complex installations such as these. A typical distribution system is illustrated in Figure 6.

DISTRIBUTION OF AC POWER; GROUNDING CONSIDERATIONS

A major implication of the unbalanced audio inputs to the duplicator machines is that AC power distribution and grounding is critical. It is very desireable that all duplicators be run off a single phase of the AC power system. If three phase power must be used, care must be taken to minimize ground currents by balancing the load between phases at all times. Given the dynamic nature of the operation of such a duplication system, single phase power to the machines is by far the best option.

Good grounding practice is essential in either method. Isolated ground wiring must be used, with each AC outlet group's green wire pulled back to the main power panel and conduits isolated from system grounds. A study of grounding techniques is recommended before designing the power distribution system.

THE PROGRAM MATERIAL

Program material to be duplicated ranges from feature films and music videos to educational and industrial training materials. Feature films are usually produced for the motion picture exhibition facility. Sound tracks for music videos are usually produced by audio recording studios whose experience is primarily the phonograph record market. While they are generally well skilled in mastering for home playback, it is not uncommon for their masters to be supplied mixed for a compact disc format, under the mistaken impression that the videocassette format has that kind of dynamic range. Most other program material is produced specifically for the video medium and originates from a very wide variety of sources. These producers share the same limited knowledge of good audio techniques as the television industry in general.

MASTERING STANDARDS

Mastering of feature films is done in this country almost exclusively in Hollywood by video houses which specialize in that process. The output of these houses varies from excellent to less than excellent. Masters of other program material come from a wide variety of sources and from vastly different producers of program material. The better Hollywood mastering houses need little guidance in the production of a videocassette master, particularly when the source is a feature film. Other mastering houses and program producers may need considerable education, depending on their understanding of audio and the videocassette medium.

It has become standard practice to supply masters on one-inch B or C format videotape, with audio recorded on the stereo tracks. Dolby A type encoding has become the de facto standard, and is recommended. Much of the non-entertainment program material will come in on a variety of formats. It will generally be most efficient to convert these to a one-inch format for duplication. It is much better to get the master supplied on one-inch if that can be done with a minimum number of generations. In any event, careful control of levels and the use of noise reduction are recomended to minimize degradation of audio quality in the transfer.

TEST SIGNALS

Audio level and frequency response test signals are critical to proper playback of the master and should be included. Desireable test signals are Dolby reference signals, pink noise, and no modulation (for signal to noise checks). In the absence of pink noise, a reasonable number of tones of discrete frequencies or a standard sweep are needed. There is presently considerable inconsistency in the use of audio test signals, and usually only Dolby reference or other level setting tones are provided, although a number of studios do supply pink noise on their masters.

Since it is the purpose of the duplication process to duplicate the original program producer's original program material as faithfully as possible, frequency response test signals which originated at the production house are also desirable. This allows for quality control of the entire distribution process from program production to mastering to finished duplicate, and aids in maintaining the precise phase response required by the Dolby Surround decoders being marketed for the home environment. Similarly, level references from that stage are useful, both because they remove errors by the mastering house from the process, and because they represent the judgments of the original audio mix engineer.

SIGNAL LEVELS

At this point a brief digression is in order. Standard practice in the recording and broadcast industries is to read levels using both VU and peak reading meters. Peak to average levels vary from one house and engineer to another, but in general, program peaks are on the order of 10-12 dB above the levels indicated on standard VU meters. Level calibration tones, are normally recorded at 10 dB or more below tape saturation.

The film industry observes a different standard. Dolby level on film is established at 50% modulation of the optical track, defined by SMPTE as 8 dB below clip, or "optical clash" (50% was -6 dB when I went to school). Dolby tone is recorded at 185 nWb/m on the magnetic full-coat optical print master (a few studios use 320 nW/m), from which master transfers of feature films are normally made. This places Dolby tone 12 dB below tape saturation, but it means that signal peaks on tape will be 6 (or 8) dB below saturation. This level is well maintained within the film industry, since it is tied to the physical limitations of the optical print as well as to the Dolby level standard.

The video duplication facility is faced with the problem of receiving motion picture masters on one standard and music video and industrial masters on another. Film audio standards carefully keeps peaks less than 6 dB above level calibration tones. The remainder of the audio industry allows peaks 10-14 dB above tone. It is important to keep sight of these differing standards when discussing audio levels.

There are three ways to deal with the problem. The first is to set a standard for the ratio of peak level to tone for masters and adhere to it. The second is to monitor program material at the incoming master evaluation station, note the highest peaks, and set playback levels appropriately. The third is to use AGC in a signal processing system to set program levels. The last method will perform some compression of program dynamic range, but is otherwise effective. The second method is the best, because it does no automatic gain riding. It is the method used in the mastering of records. The first method carries with it the problem of deciding which standard to embrace as correct and then convincing the opposite camp that 60 years of standard practice is wrong. That is not a job I would volunteer for.

SIGNAL PROCESSING AND DYNAMIC RANGE

To understand the signal processing requirements of the videocassette duplication facility, it is necessary to understand the dynamic range of the videocassette formats and the home listening environment. As was seen in the discussion of the formats, realistic dynamic signal to noise ratios for videocassette formats are 75 dB for VHS Hi-Fi, 60 dB for Beta Hi-Fi, 48 dB for linear tracks, and 58 dB for VHS linear tracks with Dolby B. Audio program material is often presented to the video duplicator which has dynamic range of up to 80 dB. Only one of the formats, VHS Hi-Fi, has the capability of handling it without problems.

An ideal home listening room might have a noise level of 40 dB and enough audio horsepower to produce 110 dB peaks at the listening position. This is 70 dB of dynamic range, and is about 20 dB better than most consumers can (or want to) do. Movie theatres have similar limitations. Tom Holman of Lucasfilm went so far as to add shaped noise to his very quiet mixing theatres monitor systems so that his mix engineers didn't mix soft program material at too low a level!

Signal processing is necessary in videocassette duplication for the same reason it is necessary in broadcasting and record mastering. It is necessary to protect the recording medium from overloading on signal peaks which exceed system headroom, while at the same time optimizing signal to noise ratio. It is necessary because not all programs are properly mastered. It is necessary because of the possibility of operator error in the setting of master playback levels. Since the linear and Hi-Fi formats have much different headroom limitations, their signal processing requirements are different.

The optimum system would have two audio masters running for each program, one mixed for Hi-Fi formats and one for linear tracks. This is clearly not an economical solution. One solution that is practical, however, is to utilize a master mixed for the Hi-Fi tracks and process it separately for the linear tracks. This system requires only one audio master tape, but does require that the duplicator machine allow separate inputs for the Hi-Fi and linear tracks. Most dedicated duplicator machines now have this capability. It also requires that two distribution systems be maintained for the duplicator line, one for the linear tracks and one for the Hi-Fi tracks. Such a system also affects the number of outputs required of the routing system.

While VHS duplicator machines have the capability to encode the linear tracks with Dolby B type noise reduction, the quality of these encoders leaves something to be desired. If Dolby encoding of the linear tracks is to be performed, it is best done using the Dolby 330 professional Dolby B type encoder in the head end following the processing provided for the linear tracks.

An example of an appropriate signal processing system is illustrated in Figure 7. The linear tracks receive two kinds of signal processing. The first is an integrated AGC and compression system which sets the averge program level. The second is pre-emphasized peak limiting to insure that high frequency program peaks do not overmodulate the highly pre-emphasized (120-175 usec) track.

A properly mixed master tape should be capable of being transferred directly to VHS Hi-Fi with no additional processing. Here, all that is generally required is peak limiting to protect tracks from overmodulation caused by improperly peak limited masters. (Remember that not all masters come from good mastering houses). The same integrated AGC and compression system used for the linear tracks is used here, but adjusted so that properly mastered program material falls below its threshold and is not processed.

DIGITAL AUDIO

Digital recording of program audio is used in the mastering of some programs where the audio quality available in the master and the budget of the duplication client permit. Digital recording is used to get around the limitations of the quality of the audio tracks of the one-inch video recorders used for mastering. The one-inch video recorder has four times the distortion and, with Dolby A, 8 dB more noise than the best Hi-Fi format. Its frequency response is also somewhat poorer. It thus degrades the program audio by increasing the distortion and noise and reducing the flatness of frequency and phase response.

The digital recorder improves the mastering process by not adding measureable noise or harmonic distortion, and by not degrading the frequency response. The mastering engineer must not, however, make the mistake of mixing or processing audio for the wider dynamic range of the digital medium, forgetting that the final product is a 60 dB dynamic range video cassette! He or she must also not forget that upwards of 90% of consumbers will be listening to audio coming from their TV speaker, and with audio from the linear tracks of VHS or Beta recorders.

While there are many suitable digital formats, digital processors which encode stereo audio into NTSC video have found the most widespread industry acceptance for digital audio mastering. Of these, the Sony PCM 1610 system and the dbx 700 are most widely used. NTSC encoding systems are used primarily because video houses can much more economically free up a spare video machine on which to master audio than they could purchase a dedicated digital audio machine that would be used a relatively small portion of the time. Synchronization between the video and audio recorders is accomplished by recording SMPTE drop frame time code on auxiliary tracks of both machines. Our facilities use the Sony PCM 1610 and the Cypher Digital (formerly BTX) synchronization system. One major advantage of that system over other SMPTE synchronization systems is that it allows considerable automation of the duplication process by the synchronizer controller.

The PCM 1610 performs 16 bit encoding at a 44.056 KHz sampling rate. This sampling rate is chosen because it is most easily synchronized by an audio house not having its own video sync generator. The dbx system uses companded predictive delta modulation at a rate of 700 Kbits per second. The two systems are not compatible except by decoding and re-encoding into the other format. The digital data, encoded as NTSC video, is recorded or transmitted via any stable video recording or transmission format, (i.e., any video signal path). The advantage of one video recording format over another is freedom from dropouts and standardization between the mastering and duplication house - i.e., both houses need machines available to play and synchronize the tapes.

EVALUATION OF INCOMING MASTERS

The wide variety of sources of program material and the variations in quality and practices of the industry make the evaluation of incoming masters an important function. One person is employed full time in each of our facilities whose sole function is incoming master evaluation. That person is responsible for inspection for both audio and video problems. Key items are levels, phase (signal polarity), stereo perspective, channel balance, proper Dolby encoding of both channels, excessive dynamic range or excessive compression, overall quality, dropouts, and audible distortion and noise. Peak and average signal levels are checked in relation to the reference levels supplied by the mastering house. Each incoming master is monitored for audio and video problems for 100% of its duration.

The listening, as well as the viewing, skills of this person need to be carefully developed. The master evaluation operator needs ear training in perception of intermodulation and harmonic distortion, unencoded and undecoded Dolby, noise, frequency response deficiencies, drop outs, excessive compression, excessive dynamic range, and lip sync. When uncorrectable problems are discovered, the master should be rejected and a replacement obtained. Other than for minor corrections of level and channel balance, the duplication facility should not be expected to modify the audio content of program masters.

Lip sync errors are easily generated in the duplication and transfer process and need to be minimized. These errors build up when video time base correctors, frame store systems, and audio layback machines are used. Each use cascades an additional frame (33 ms) or two of video signal delay, which, if not tracked by a corresponding amount of audio delay can add to an annoying degree of error. Some time base correctors and frame synchronizers provide this correction. A high quality audio delay is available from Lexicon that will allow precise correction of large amounts of offset to be dialed in by the operator.

TAPE STOCK FOR DUPLICATION USE

Tape stock for duplication may be purchased in bulk and loaded into cassettes by the duplication facility or purchased already loaded. Depending on the volume of business, the length of program material to be recorded, and corporate resouces either or both may be adviseable.

Many manufacturers supply tape for videocassette use. Each of these has a slightly different set of audio performance characteristics which may or may not be obvious to the purchasing agent. Many may be perfectly acceptable for video but require different audio level or bias for good audio performance. Overbias will cause the recording to sound dull, while underbias will produce bright, distorted recordings. A tape with high sensitivity may be too easily overloaded, while one with too low sensitivity will produce recordings which are weak and noisy. While these audio parameters do not affect the Hi-Fi audio tracks, they can wreak havoc with the linear tracks (which 95% of consumers listen to). A good practice is to give your purchasing agent a choice of several tape stocks which will meet video performance specs and which will perform well with the same levels of bias and audio drive. This will allow him to locate multiple sources for tape which meet your needs and keep the production line running smoothly.

Incoming tape stock needs to be checked for quality control. Factors which affect audio performance are smooth edge slitting, bias, senstivity, and freedom from dropouts. Once a tape type has been determined to be satisfactory for audio purposes, each incoming batch needs to be checked to make sure it is within tolerances. It is common for slitting, bias and sensitivity to vary from batch to batch. Slitting problems will show up as poor playback of channel 2 linear VHS audio, because this track is adjacent to the tape edge. The tests for bias and sensitivity are frequency response, distortion, and noise level. Variations in bias will shift frequency response, noise, and distortion. Since the linear tape speeds are very low, a sensitivity check at 1 kHz will be very useful in determining bias compatibility. In general, if a batch of tape meets video performance standards, it will perform within reasonable tolerances for audio.

QUALITY CONTROL OF THE FINISHED CASSETTE

Test signals recorded on the tape can allow fast and accurate checks on performance of both the total system and individual duplicator machines. If each finished tape is numerically coded with the number of the machine on which it was recorded, the pass number, and other pertinent data, problems can be easily located and corrected when rejects occur. The most useful test signals are pink noise, no modulation, and a single frequency midrange tone (1 kHz is fine). From these frequency response, channel balance, phase and head alignment, distortion, and system noise can quickly be measured. Listening to audio at random spots within the program completes the test.

NEW DEVELOPMENTS

Two newly introduced audio products may have impact on the interchange of audio program material for the videotape medium. At the 1986 Montreaux AES, Dolby Laboratories introduced its SR Noise Reduction system, boasting noise performance meeting or exceeding digital specifications. Although no reduction of non-linear distortion is offered by the new system, system distortion could be reduced to about the level of that on the Hi-Fi tracks by reducing operating levels on the 1-inch videotape by about 6 dB. Considering the noise reduction offered by the new system, this may be a viable alternative. Electronics capable of processing for the new system will be available this summer. Cost of the hardware to support the new system is about one third more than existing Dolby encoding/decoding product. The new modules are plug-in compatible with most existing Dolby frames, but, due to circuit complexity, are not expected to be available to fit inside videotape recorders in the forseeable future.

At the 1986 NAB Convention in Dallas, Sony Corporation announced a new digital encoding sytem for its 1-inch C-format videotape recorders. The system is said to add two 16-bit, 44.056, 44.1, or 48 KHz encoded audio tracks to the format without interference with the existing audio or video channels. Cost of the hardware to support the new system is high (approximately $100,000 per complete machine), but the new system will allow one videotape to contain all of the following signals simultaneously:

1 - NTSC Video, as existing
3 - Linear audio channels, as existing
2 - Digital audio channels

ACKNOWLEDGEMENT

The author would like to thank John Laberdie of Matrix Video; Gene Reich, John Sergeant, Du Wayne Stephens, and Hal Miller of Bell and Howell Columbia Pictures Video Services; and David Robinson of Dolby Laboratories for their sharing of information for this work; and Jim Addie of WFMT, Inc. for his work performing measurements to verify system performance.

REFERENCES

Greg Badger, Richard Allen, "THE AUDIO SIDE OF THE LASER VIDEODISC", Paper presented at the 72nd AES Convention, October 1982, Preprint # 1935

Jim Brown, "TERMINATION, IMPEDANCE MATCHING, AND THE MAXIMUM POWER TRANSFER THEOREM IN AUDIO SYSTEMS", Tech Topic Volume 10, Number 1, Synergetic Audio Concepts, San Juan Capistrano, CA., 1982

Allan Burdick, "A CLEAN AUDIO INSTALLATION GUIDE", Broadcast Engineering, March, 1986

F. Alton Everest, "INSTRUCTION IN CRITICAL LISTENING", Paper presented at the 72nd AES Convention, October 1982, Preprint # 1890

Richard Hess (ABC-TV), "VOLTAGE TRANSMISSION FOR AUDIO SYSTEMS", Paper presented at the 67th AES Convention, October 1980, Preprint # 1708

Tomlinson Holman, "AUDIO PRODUCTION TECHNIQUES FOR MAJOR MOTION PICTURES", Workshop Presented at 74th Convention of AES, October, 1983.

Stephen Juhlstrom, "A HIGH PERFORMANCE SURROUND SOUND PROCESS FOR HOME VIDEO", Paper presented at the 79th AES Convention, October 1985, Preprint # 2300

Charles Repka, "A SOUND MIXER'S GUIDE TO VIDEODISC", SMPTE Journal, September, 1984

Michael Rettinger, "SOUND TRANSMISSION THROUGH PERFORATED SCREENS", SMPTE Journal, December, 1982

Yuma Shiraishi, "HISTORY OF HOME VIDEOTAPE RECORDER DEVELOPMENT", SMPTE Journal, December, 1985

SMPTE Engineering Guideline EG-9-1985 "AUDIO RECORDING REFERENCE LEVEL FOR POST PRODUCTION OF MOTION PICTURE RELATED MATERIAL" SMPTE Journal, December, 1985

APPLICABLE STANDARDS

TYPE B VIDEOTAPE AUDIO:
SMPTE C98.16M-1980
SMPTE C98.17M-1980
SMPTE RP 107-1982 (Reference tapes)

TYPE C VIDEOTAPE AUDIO:
SMPTE 20M-1985
SMPTE V98.19M-1983
SMPTE RP 99-1983 (Reference tapes)

Beta Videocassette:
SMPTE RP-119-1984
SMPTE V98.34M-1984

VHS Videocassette:
SMPTE RP 112-1983
SMPTE V98.32-1983

STANDARDS AND PERFORMANCE OF SELECTED AUDIO RECORDING FORMATS

	(LINEAR TRACKS) VHS (SP)	BETA (II)	8 mm	C-format	1/4 inch 2-trk	AUDIO CASSETTE
Tape speed (in/sec)	1.313	0.79	0.565	9.6	15	1.875
Track width (Mono) (mm)	1	1.05	0.6	--	--	1.52
(Stereo)	0.3	--	--	0.8	1.91	0.6
Pre-emph (usec)	3180, 120	3180, 175	n/a	3180, 15	3180, 50	1590, 120/50
Ref level (nWb/m)	100	100	--	100	320	200
3% THD (nWb/m)	100	100		320	1040	450
S/N (Unwtd) (dB re 3% THD)	48	48		58	67	54
S/N (w/Dolby) (Unwtd, re 3%)	54	--		70 (B Dolby)	79	58 (B Dolby)
Coercivity (Oe)	600	600	800	600	320	360

Notes:
1) Audio cassette is standard Phillips cassette using high quality consumer tape.
2) 1/4 inch, 2 trk is professional mastering recorder, premium tape.
3) C-format machine uses current best quality videotape.

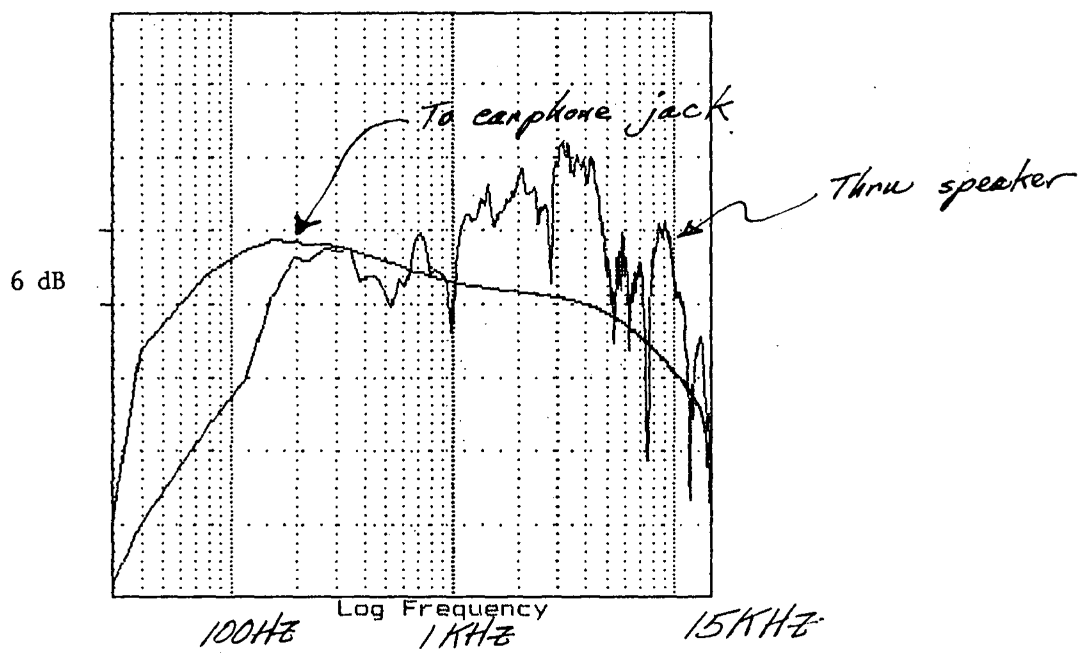
FIGURE 1a

TYPICAL SYSTEM AND PERFORMANCE SPECIFICATIONS OF HI-FI VIDEOCASSETTE FORMATS

	VHS HD	BETA Hi-Fi	8 mm Hi-Fi	8 mm PCM	C-format (Linear)	2-trk 1/4 inch
Video writing speed (in/sec)	275	150	150		1008	--
FM Carriers (Mhz)	1.3 1.7	1.38 1.53 1.68 1.83	1.5		--	--
Deviation (Khz)	150	75	100		--	--
Audio Bandwidth	20-20KHz	20-20KHz	20-20KHz	20-15KHz	30-17KHz	30-20KHz
THD+N at ref level	0.25%	0.5%			0.6%	0.3%
THD+N below clip	0.5%	1%			3.5%	3%
s/n re clip (Unwtd)	77	60			58 (70 w/A Dolby)	70 (79 w/A Dolby)

FIGURE 1b

Mag. vs Hz (EFC) of SONY KV-1920D Rcvr, thru SONY SL-HF300 Modulator
By JIM BROWN
On 10/3/85
At S. E. A. LAB



Vertical: 6dB/div

Horizontal: Auto 30.00Hz to 14999.70Hz
Log freq axis (2.7decades)

Resolution: 7.7385E+00 Meter & 44.47Hz

Time of test: 86 microseconds, 2.9594E-02 Meter

Sweep Rate & Bandwidth: 1977.45Hz/Sec & 44.47Hz

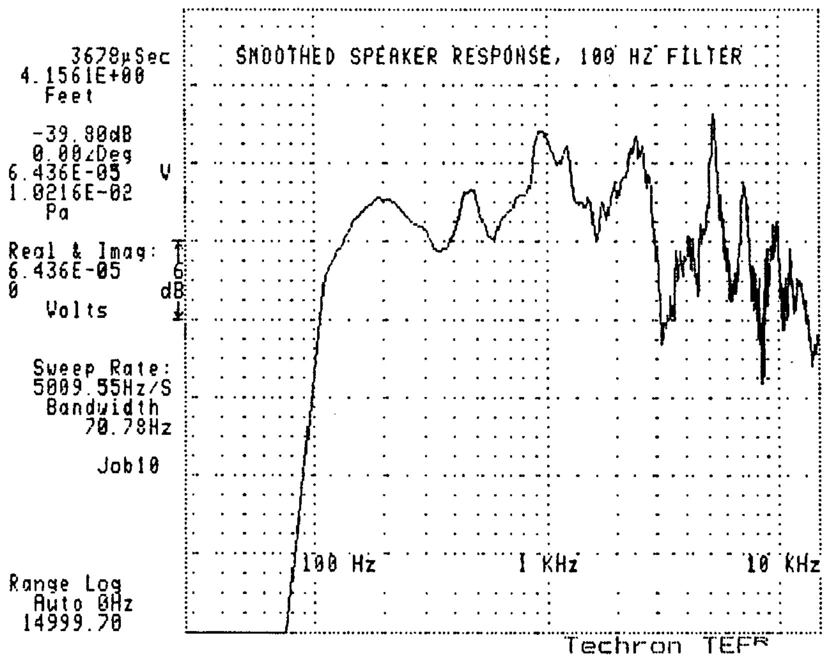
Input configuration: Channel 1 Non-inverting
with 0dB of input gain & 12dB of IF gain.

Remarks:

Audio response of a popular 1978 vintage table model television receiver in the $500 range. The test signal was fed to the audio input of a consumer VCR, which was used as an RF modulator for the receiver. The small loudspeaker faces forward. The measurement was made by Time Delay Spectometry, using the Techron TEF analyzer.

FIGURE 2a

Mag. vs Hz (EFC) of JVC 2690 LEFT SPEAKER
By JIM BROWN
On 3/22/86
At SOUND ENGINEERING ASSOCIATES LAB



Vertical: 6dB/div

Horizontal: Auto 0.00Hz to 14999.70Hz
Log freq axis (2.7decades)

Resolution: 1.5967E+01 Feet & 7.0778E+01Hz

Time of test: 3678 microseconds, 4.1561E+00 Feet

Sweep Rate & Bandwidth: 5009.55Hz/Sec & 7.0778E+01Hz

Input configuration: Balanced
with 60dB of input gain & 15dB of IF gain.

Audio response of a premium quality 1986 table model receiver/monitor which sells for about $ 1,000. It has video and audio inputs, a BTSC stereo decoder, two five watt amplifiers, stereo headphone jacks, audio outputs, and provision for the connection of external speakers. Its two built in speakers face sideways, and come equipped with reflectors adjustable to an approximate 45° angle to direct the sound toward the viewer. Response was measured at a normal viewing angle, with the generator fed to the audio input.

FIGURE 2b

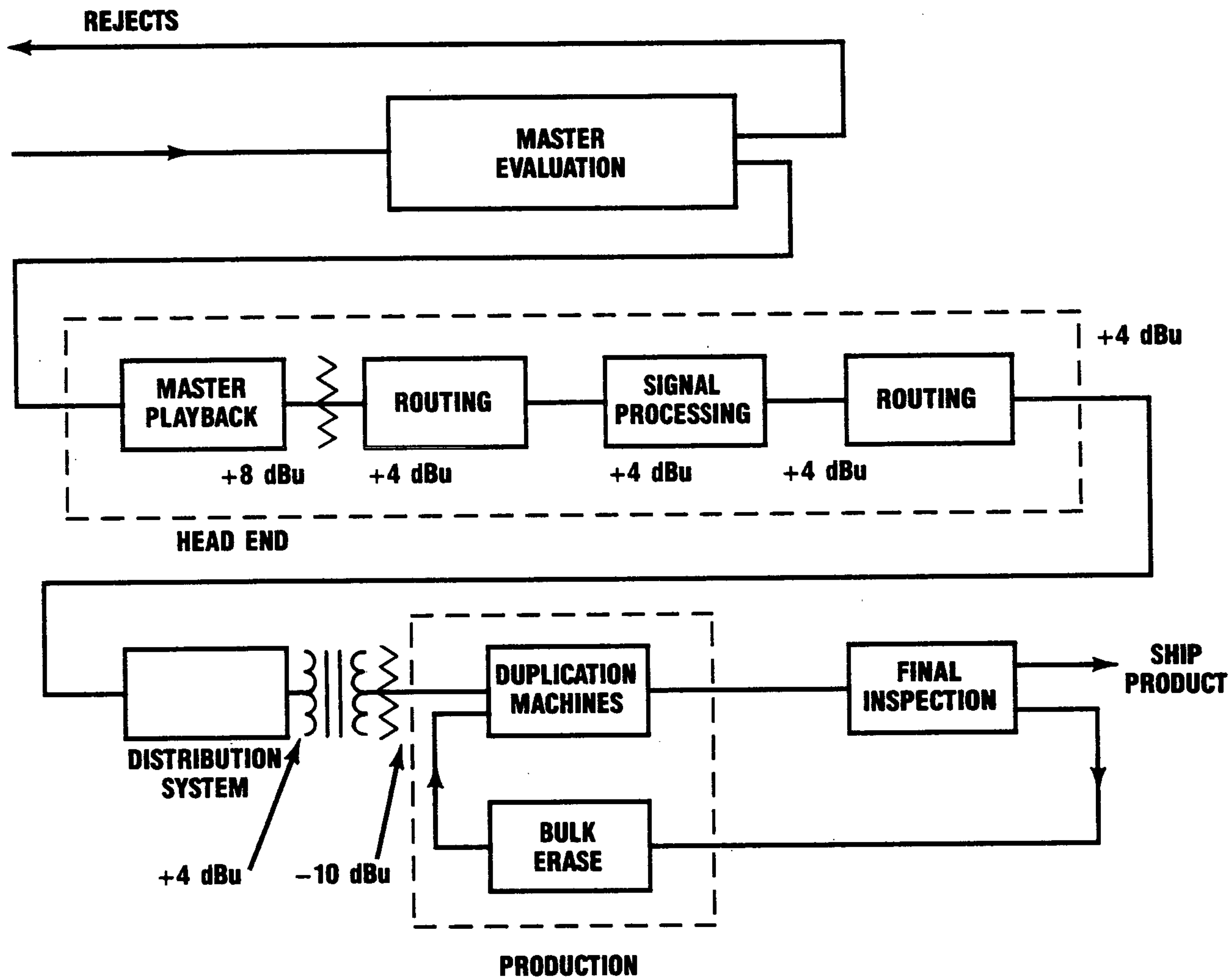


Figure 3 SIMPLIFIED PROCESS FLOW AND SIGNAL LEVELS IN A TYPICAL DUPLICATION FACILITY

SPECIFICATIONS FOR AUDIO EQUIPMENT TO BE USED IN HEAD END (0 dBu = 0.775v)

Maximum output level: +25 dBm into 600 ohms
+27 dBu unterminated
Distortion: < 0.01% THD below +24 dBu unterminated
+22 dBm into 600 ohms
< 0.05% THD at maximum output level
Hum and Noise: -88 dBm
Output impedance: < 60 ohms, balanced
Input Impedance: > 10K ohms, balanced
Maximum Input level: +27 dBu
Frequency response: ±0.1 dB, 20-20,000 Hz

SPECIFICATIONS FOR COMPONENTS TO BE USED IN DISTRIBUTION SYSTEM

Maximum output level: +25 dBm into 600 ohms
Distortion: < 0.01% THD below +24 dBu unterminated
+22 dBm into 600 ohms
< 0.05% THD at maximum output level
Hum and Noise: -85 dBm
Output impedance: < 60 ohms, balanced
Input Impedance: > 50K ohms, balanced
Maximum Input level: +25 dBu
Frequency response: ±0.1 dB, 20-20,000 Hz

SPECIFICATIONS FOR TOTAL AUDIO CHAIN

(Measured at duplicator machine inputs, system fed from master playback recorder outputs).

Overall system voltage gain: -14 or -18 dB, with loss taken at duplicator machine inputs.
Operating level (tone): -10 dBu
Maximum output level: +7 dBu
Distortion: < 0.05% @ +4 dBu
< 0.025% @ -10 dBu
Hum and noise: < -90 dBu (25 uv rms)
Frequency response: ±0.25 dB, 20-20,000 Hz
Phase response: ± 5°, 50-8,000 Hz between L & R channels
Polarity: Absolute polarity should be maintained through the system

FIGURE 4

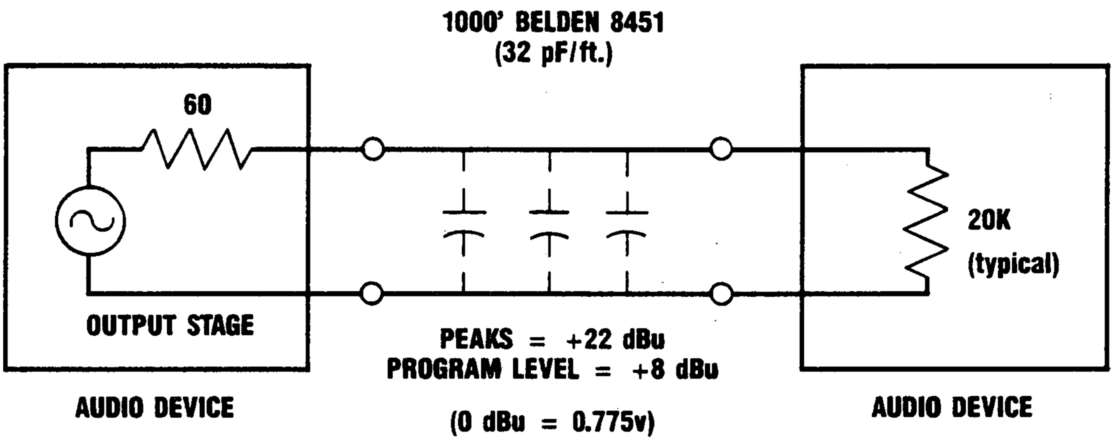


Figure 5a: CONSTANT VOLTAGE DISTRIBUTION

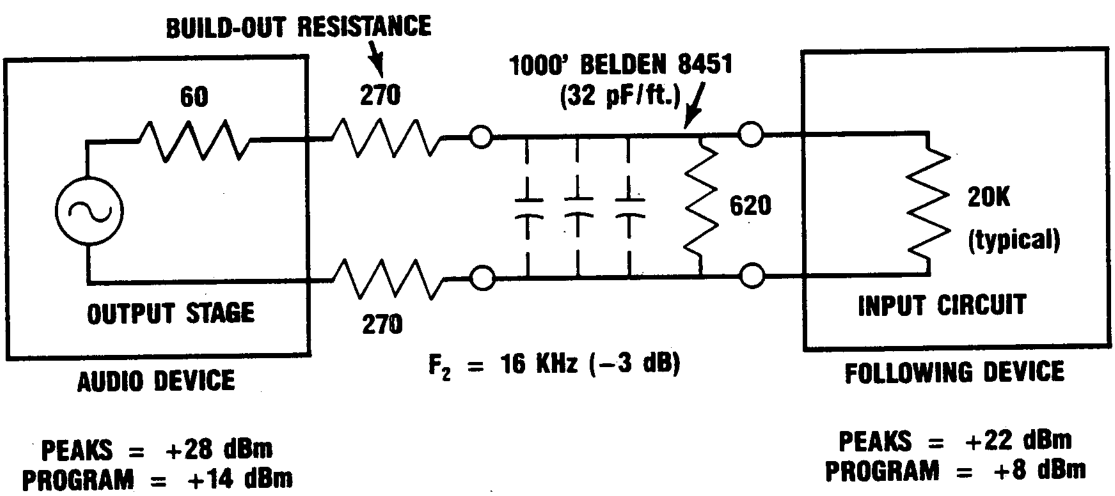


Figure 5b: POWER MATCHED DISTRIBUTION

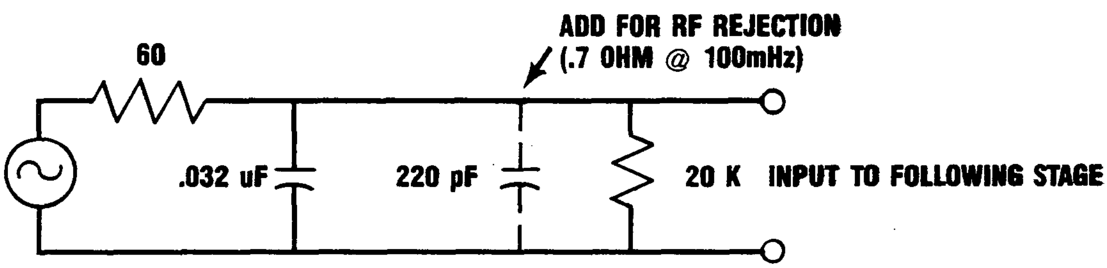


Figure 5c: EQUIVALENT CIRCUIT OF 5a

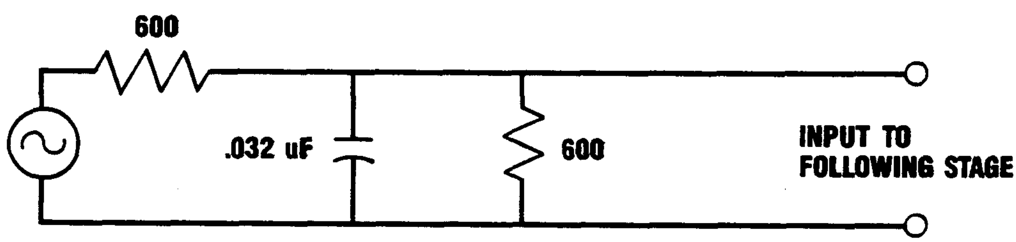


Figure 5d: EQUIVALENT CIRCUIT OF 5b

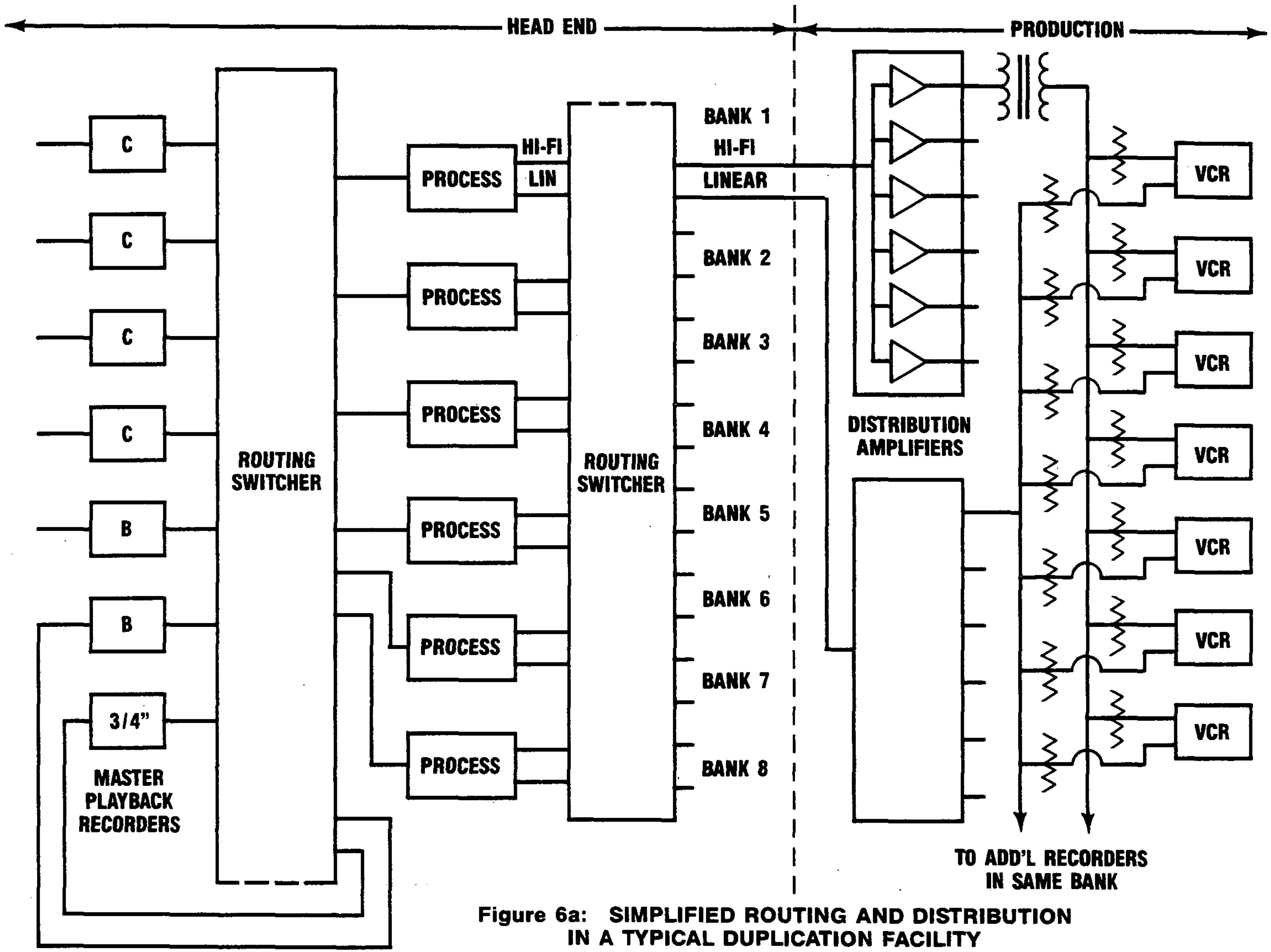


Figure 6a: SIMPLIFIED ROUTING AND DISTRIBUTION IN A TYPICAL DUPLICATION FACILITY (EACH LINE REPRESENTS A STEREO AUDIO PATH)

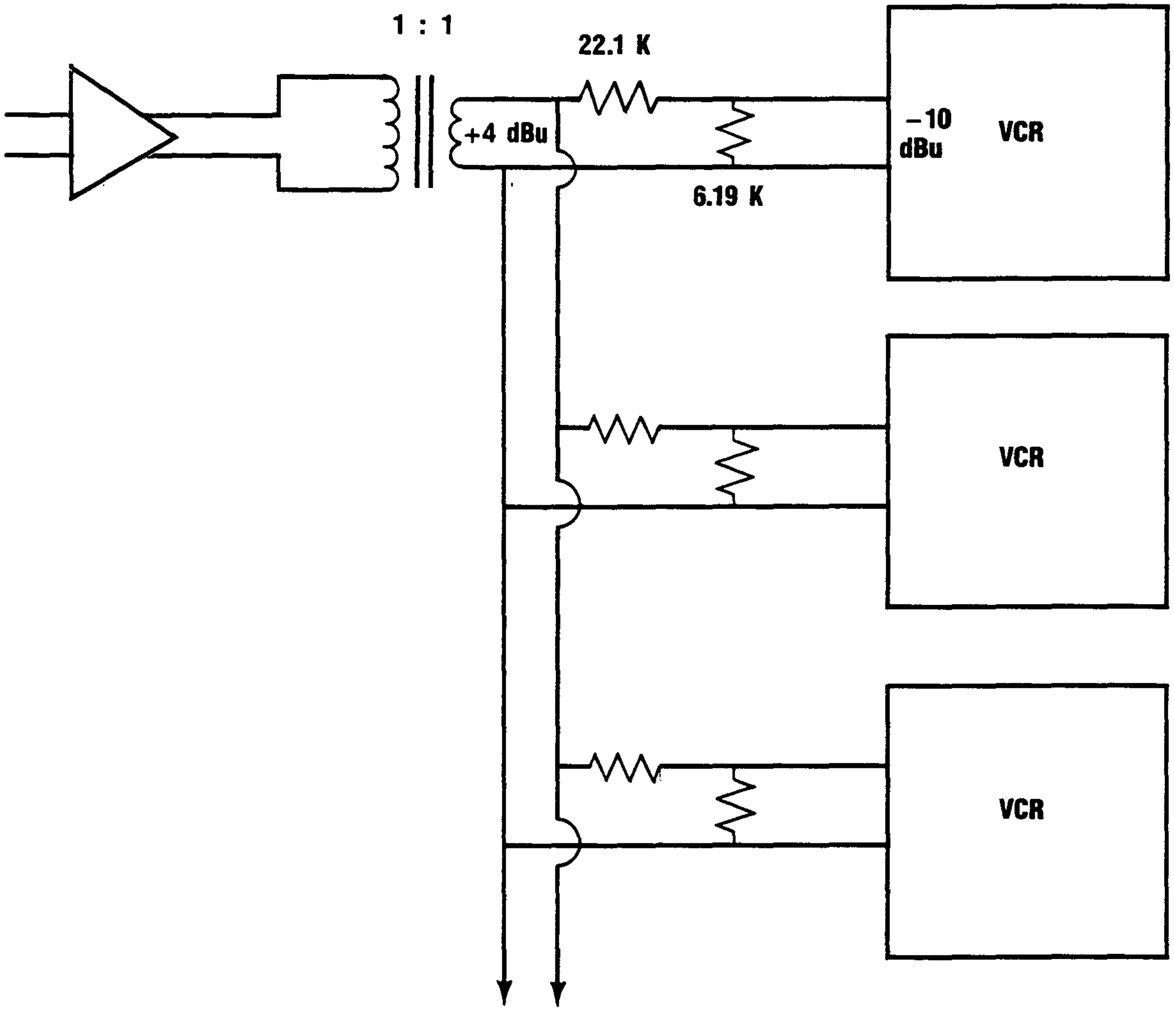


Figure 6b: DETAIL OF DISTRIBUTION OUTPUT ATTENUATION AND SIGNAL LEVELS

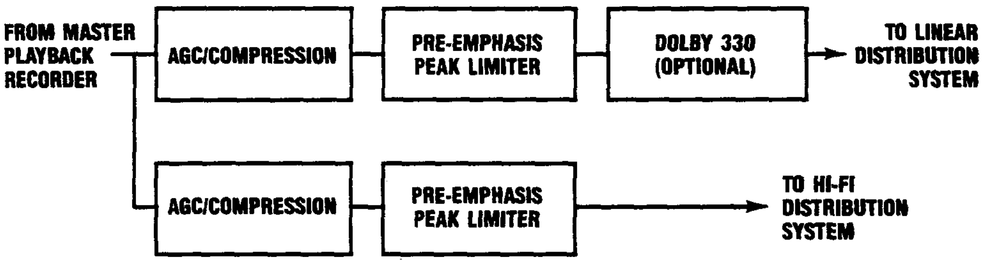


Figure 7: A TYPICAL SIGNAL PROCESSING SYSTEM, SHOWING SEPARATE PROCESSING FOR LINEAR AND HI-FI TRACKS AND A PROFESSIONAL DOLBY TYPE B ENCODER AT THE HEAD END

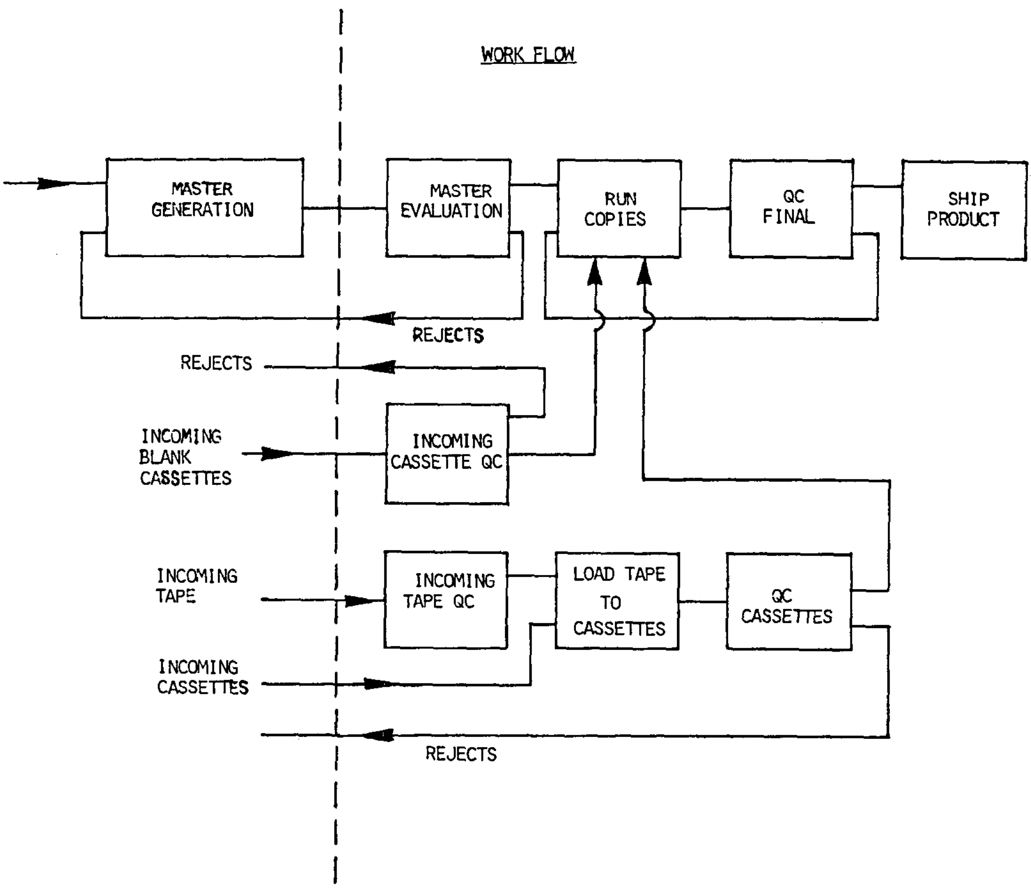


FIGURE 8