Robert Orban
Greg Ogonowski

Maintaining Audio Quality in the Broadcast Facility

2003 Edition

orban

Table of Contents

This document is part number 95383.000.02

1525 Alvarado Street, San Leandro, CA 94577 USA
Phone: (1) 510/351-3500; Fax: (1) 510/351-0500; E-Mail: custserv@orban.com; Site: www.orban.com

Maintaining Audio Quality in the Broadcast Facility

By Robert Orban and Greg Ogonowski, Orban/CRL

Authors' Note

In 1999, we combined and revised two previous Orban papers on maintaining audio quality in the FM and AM plants, with a further revision occurring in 2003. In 2003, considerations for both AM and FM are essentially identical except at the transmitter because, with modern equipment, there is seldom reason to relax studio quality in AM plants. The text emphasizes FM (and, to a lesser extent, DAR) practice; differences applicable to AM have been edited into the FM text.

Introduction

Audio processors change certain characteristics of the original program material in the quest for positive benefits such as increased loudness, improved consistency, and absolute peak control.

The art of audio processing is based on the idea that such benefits can be achieved without allowing the listener to detect that anything has been changed. Successful audio processing performs the desired electrical modifications while presenting a result to the listener that, subjectively, sounds natural and realistic. This sounds impossible, but it is not.

Audio processing provides a few benefits that are often unappreciated by the radio or television listener. For example, the reduction of dynamic range caused by processing makes listening in noisy environments (particularly the car) much less difficult. In music having a wide dynamic range, soft passages are often lost completely in the presence of background noise. Few listeners listen in a perfectly quiet environment. If the volume is turned up, subsequent louder passages can be uncomfortably loud. In the automobile, dynamic range cannot exceed 20dB without causing these problems. Competent audio processing can reduce the dynamic range of the program without introducing objectionable side effects.

Further, broadcast program material typically comes from a rapidly changing variety of sources, most of which were not produced with any regard for the spectral balances of any other. Multiband limiting, when used properly, can automatically make the segues between sources much more consistent. Multiband limiting and consistency are vital to the station that wants to develop a characteristic audio signature and strong positive personality, just as feature films are produced to maintain a consistent look. Ultimately, it is all about the listener experience

Each broadcaster also has special operational considerations. First, good broadcast operators are hard to find, making artful automatic gain control essential for the correction of errors caused by distractions or lack of skill. Second, the regulatory authorities in most countries have little tolerance for excessive modulation, making peak limiting mandatory for signals destined for the regulated public airwaves.

OPTIMOD-FM, OPTIMOD-AM, OPTIMOD-DAB, OPTIMOD-TV, and OPTIMOD-PC have been conceived to meet the special problems and needs of broadcasters while delivering a quality product that most listeners consider highly pleasing. However, every electronic communication medium has technical limits that must be fully heeded if the most pleasing results are to be presented to the audience. For instance, the audio quality delivered by OPTIMOD is highly influenced by the quality of the audio presented to it. If the input audio is very clean, the signal after processing will probably sound excellent—even after heavy processing. Distortion of any kind in the input signal is likely to be exaggerated by processing and, if severe, can end up sounding offensive and unlistenable.

AM is limited by poor signal-to-noise ratio and by limited receiver audio bandwidth (typically 2-3 kHz). As delivered to the consumer, it can never be truly "high fidelity." Consequently, multiband audio processing for AM compresses dynamic range more severely than in typical FM practice. In addition, pre-emphasis (whether NRSC or more aggressive than NRSC) is required to ensure reasonably crisp, intelligible sound from typical AM radios. In AM, this is always provided in the audio processor and never in the transmitter.

Audio quality in TV viewing is usually limited by small speakers in the receivers, although the increasing popularity of DTV, HDTV and home theatre is changing this, increasing consumer demand for high audio quality. In everyday television viewing, it is important to avoid listener irritation by maintaining consistent subjective loudness from source to source. A CBS Loudness Controller or multi-band processing, both included in OPTIMOD-TV, can achieve this.

Netcasting, also known as webcasting, almost always requires low bit-rate codecs. Processing for such codecs should not use clipping limiting, and should instead use a look-ahead type limiter. OPTIMOD-DAB, OPTIMOD-HD FM, and OPTIMOD-PC provide the correct form of peak limiting for netcasting and other low bite rate services.

Achieving consistent state-of-the-art audio quality in broadcast is a challenging task. It begins with a professional attitude, considerable skill, patience, and an unshakable belief that quality is well worth having. This supplement provides some technical insights and tips on how to achieve immaculate audio, and keep it that way. Remember, successful audio processing results all starts at the source.

This publication is organized into four main parts:

1. **Recording media:** compact disc, CD-R and CR-RW, digital tape, magnetic disk and data compression, vinyl disk, phonograph equipment selection and maintenance, analog tape, tape recorder maintenance, recording alignment tapes and cart machine maintenance—see page 4.

2. **System considerations:** headroom, voice/music balance, and electronic quality—see page 22.

3. **The production studio:** choosing monitor loudspeakers, loudspeaker location and room acoustics, loudspeaker equalization, stereo enhancement, other production equipment, and production practices—see page 31.

4. **Equipment following OPTIMOD:** exciters, transmitters, and antennas—see page 37.

NOTE: Because the state of the art in audio technology is constantly advancing, it is important to know that this material was last revised in 2003. Our comments and recommendations obviously cannot take into account later developments. We have tried to anticipate technological trends when that seemed useful.

Part 1: Recording Media

Compact Disc

The compact disc (CD) is currently the primary source of most recorded music. With 16-bit resolution and 44.1 kHz sample rate, it represents the reference standard source quality for radio, although it may be superceded in the future by DVD-Audio, with 24-bit resolution and 96 kHz sample rate, or by SACD, which uses "bitstream" coding instead of the CD's PCM (Pulse Code Modulation). Because most audio is sourced at a 44.1 kHz sample rate, upsampling to 48 kHz does not improve audio quality. Further, many broadcast digital sources have received various forms of lossy data compression. While we had expected the black vinyl disk to be obsolete by this revision, it is still used in specialized applications like live "club-style" D.J. mixing.

Although CD technology is constantly improving, we believe that some general observations could be useful. In attempting to reproduce CDs with the highest possible quality, the industry has settled into technology using "delta-sigma" digital-to-analog converters (DACs) with extreme over-sampling. These converters use pulse width modulation or pulse-duration modulation techniques to achieve high accuracy. Instead of being dependent on the precise switching of voltages or currents to achieve accurate conversion, the new designs depend on precise timing, which is far easier to achieve in production.

Over-sampling simultaneously increases the theoretical signal-to-noise ratio and produces (prior to the reconstruction filter within the CD player) a signal that has no significant out-of-band power near the audio range. A simple, phase-linear analog filter can readily remove this power, ensuring the most accurate phase response through the system. We recommend that CD players used in broadcast employ technology of at least this quality. However, the engineer should be aware that these units might emit substantial amounts of supersonic noise, so that low-pass filtering in the transmission audio processor must be sufficient to reject this to prevent aliasing in digital transmission processors or STLs.

The radio station environment demands ruggedness, reliability, and quick cueing from audio source equipment. The CD player must also be chosen for its ability to track even dirty or scratched CDs with minimum audible artifacts, and on its ability to resist external vibration. There are dramatic differences between players in these areas! We suggest careful comparative tests between players using imperfect CDs to determine which players click, mute, skip, or otherwise mistrack. Striking the top and sides of the player with varying degrees of force while listening to the output can give a "feel" for the player's vibration resistance. Fortunately, some of the players with the best sound also track best. The depressing trade-off between quality and ruggedness that is inevitable in vinyl disk reproduction is unnecessary when CDs are used.

Reliability is not easy to assess without experience. The experience of your fellow broadcasters can be valuable here—ask around during local broadcast engineers' meetings. Be skeptical if examination of the "insides" of the machine reveals evidence of poor construction.

Cueing and interface to the rest of the station are uniquely important in broadcast. There are, at this writing, relatively few players that are specifically designed for broadcast use—players that can be cued by ear to the start of a desired selection, paused, and then started by a contact closure. The practical operation of the CD player in your studio should be carefully considered. Relatively few listeners will notice the finest sound, but all listeners will notice miscues, dead air, and other obvious embarrassments!

Some innovative designs that have already been introduced include jukebox-like CD players that can hold 100 or more CDs. These players feature musical selections that can be chosen through computer-controlled commands. An alternative design, which also tries to minimize CD damage caused by careless handling, places each CD in a protective plastic "caddy." The importance of handling CDs with care and keeping the playing surface clean cannot be over-emphasized. Contrary to initial marketing claims of invulnerability, CDs have proven to require handling comparable to that used with vinyl disks in order to avoid on-air disasters.

Except for those few CD players specifically designed for professional applications, CD players usually have unbalanced –10dBV outputs. In many cases, it is possible to interface such outputs directly to the console (by trimming input gains) without RFI or ground loop problems. If these problems do appear, several manufacturers produce low-cost –10dBV to +4dBu adapters for raising the output level of a CD player to professional standards.

Using a stand-alone CD player to source audio for a digital playout system is currently one of the most common ways to transfer CD audio to these systems. To achieve the best accuracy, use a digital interface between the CD player and the digital playout system. An alternative is to extract the digital audio from the CD using a computer and a program to "rip" the audio tracks to a digital file.

The primary advantage of computer ripping is speed. However, it is crucial to use the right hardware and software to achieve accuracy equivalent to that routinely found in a stand-alone CD player. A combination of an accurate extraction program (such as Exact Audio Copy or EAC) and a Plextor® CD drive (which implements hardware error correction) will yield exceptional results. Not all CD drives are capable of digital audio extraction and not all drives offer hardware error correction.

CD-R and CD-RW

The cost of CD-R (compact disk–recordable) has now dropped to the point where it is a very attractive solution as an on-air source and for archiving. The quality is equivalent to CD.

There are several dye formulations available, and manufacturers disagree on their archival life. However, it has been extrapolated that any competently manufactured CD-R should last at least 30 years if it is stored at moderate temperatures (below 75 degrees F) and away from very bright light like sunlight. On the other hand, these disks can literally be destroyed in a few hours if they are left in a locked automobile, exposed to direct sunlight.

CD-RW (compact disk–rewritable) is not a true random-access medium. You cannot randomly erase cuts and replace them because the cuts have to be unfragmented and sequential. However, you can erase blocks of cuts, always starting backwards with the last one previously recorded. You can then re-record over the space you have freed up.

The disadvantage of CD-RW is that most common CD payers cannot read them, unlike CD-R, which can be read by almost any conventional CD player if the disk has been "finalized" to record a final Table of Contents track on it. A finalized CD-R looks to any CD player like an ordinary CD. Once a CD-R has been finalized, no further material can be added to it even if the disk is not full. If a CD-R has not been finalized, it can only be played in a CD-R recorder, or in certain CD players that specifically support the playing of unfinalized CD-Rs.

Digital Tape

While DAT was originally designed as a consumer format, it has achieved substantial penetration into the broadcast environment. This 16-bit, 48 kHz format is theoretically capable of slightly higher quality than CD because of the higher sample rate. In the DAR environment, where 48 kHz-sample rate is typical, this improvement can be passed to the consumer. However, because the "sample rate" of the FM stereo system is 38 kHz, there is no benefit to the higher sampling rate by the time the sound is aired on FM.

The usual broadcast requirements for ruggedness, reliability, and quick cueing apply to most digital tape applications, and these requirements have proven to be quite difficult to meet in practice. The DAT format packs information on the tape far more tightly than do analog formats. This produces a proportional decrease in the durability of the data. To complicate matters, complete muting of the signal, rather than a momentary loss of level or high frequency content, as in the case of analog, accompanies a major digital dropout.

At this writing, there is still debate over the reliability and longevity of the tape. Some testers have reported deterioration after as little as 10 passes, while others have demonstrated almost 1000 passes without problems. Each demonstration of a tape surviving hundreds of passes shows that it is physically possible for R-DAT to be reliable and durable. Nevertheless, we therefore advise broadcasters not to trust the reliability of DAT tape for mastering or long-term storage. Always make a backup!

Because the cost of recordable CD blanks has dropped to the point where they are almost throwaway items, we advise using CD-R instead of DAT when long-term archivability is important.

Data Compression and Disks

Hard disk systems use sealed Winchester hard magnetic disks or optical disks (originally developed for mass storage in data processing) to store digitized audio. This technology has become increasingly popular as a delivery system for material to be aired. There are many manufacturers offering systems combining proprietary software with a bit of proprietary hardware and a great deal of off-the-shelf hardware. Provided that they are correctly administered and maintained, these systems are the best way to ensure high, consistent source quality in the broadcast facility because once a source is copied onto a hard drive, playout is consistent. There are no random cueing variations and the medium does not suffer from the same casual wear and tear as CDs. Of course, hard drives fail catastrophically from time to time, but RAID arrays can make a system immune to almost any such fault.

When one builds a music library on a digital delivery system, it is important to validate all audio sources. A track's being available on CD does not guarantee good audio quality.

Many original record labels are defunct and have transferred licenses to other major labels. It is usually safe to assume that the audio from the original record/CD label or authorized licensee is as good as it gets. Additionally, many of these major labels produce collections for other well-known marketing groups. All of these sources are usually acceptable. However, some obscure labels have acquired licenses from the original labels. These reissues should probably be avoided. Regardless of source, it is wise to use correct "hit" version of the original even if its audio quality is worse than alternative versions.

Many tracks, even from "desirable" labels, have been recently remastered and may sound quite different from the original transfer to CD. Some of the more recent remasterings may contain additional signal processing beyond simple tick and pop elimination. Because of the much-reviled advent of "hypercompression" in the mastering industry, newly remastered tracks should be validated very carefully, as the newer tracks may suffer from excessive digital limiting that reduces transient impact and punch. Therefore, the older, less-processed sources may stand up better to Optimod transmission processing.

It is beyond the scope of this document to discuss the mechanics of digital delivery systems, which relate more to ergonomics and reliability than to audio quality. However, two crucial issues are how the audio is input and output from the system, and whether the audio data is stored in uncompressed (linear PCM) form or using some sort of data compression.

Audio is usually input and output from these systems through sound cards. Please see the discussion on page 24 regarding sound cards and line-up levels.

There are two forms of compression—lossy, and lossless. Lossless compression provides an output that is bit-for-bit identical to its input. The best known of these systems for audio is MLP (Meridian Lossless Packing), which has been accepted for use with the DVD-Audio standard to increase its data carrying capacity by approximately 1.7x.

Lossy compression eliminates data that its designer has determined to be "irrelevant" to human perception, permitting the noise floor to rise instead in a very frequency-dependent way. This exploits the phenomenon of *psychoacoustic masking,* which means that quiet sounds coexisting with louder sounds will sometimes be drowned out by the louder sounds so that the quieter sounds are not heard at all. The closer in frequency a quiet sound is to a loud sound, the more efficiently the louder sound can mask it. There are also "temporal masking" laws having to do with the time relationship between the quieter and louder sounds.

A good psychoacoustic model that predicts whether or not an existing sound will be masked is complicated. The interested reader is referred to the various papers on perceptual coders that have appeared in the professional literature (mostly in the *Journal of* *the Audio Engineering Society* and in various AES Convention Preprints) since the late 1980s.

There are two general classes of lossy compression systems. The first is exemplified by ADPCM and APT-X®, which, while designed with full awareness of psychoacoustic laws, does not have a psychoacoustic model built into it. In exchange for this relative simplicity it has a very short delay time (less than 4ms), which is beneficial for applications requiring foldback monitoring, for example.

The second class contains built-in psychoacoustic models, which the encoder uses to determine what parts of the signal will be thrown away and how much the noise floor can be allowed to rise without its becoming audible. These codecs can achieve higher quality for a given bit rate than codecs of the first class at the expense of much larger time delays. Examples include the MPEG family of encoders, including Layer 2, Layer 3, and AAC. The Dolby® AC-2 and AC-3 codecs also fall in this category. The large time delays of these codecs make them unsuitable for any application where they are processing live microphone signals that are then fed back into the announcer's headphones. In these applications, it is sometimes possible to design the system to bypass the codec, feeding the undelayed or less delayed signal into the headphones.

Coding Technologies' "Spectral Band Replication" (SBR) process can be added to almost any codec. This system transmits only lower frequencies (for example, below 8 kHz) via the codec. The decoder at the receiver creates higher frequencies from the lower frequencies by a process similar to that used by "psychoacoustic exciters." A low-bandwidth signal in the compressed bit stream provides "hints" to modulate these created high frequencies so that they will match the original high frequencies as closely as possible.

There are two general applications for codecs in broadcasting — "contribution" and "transmission." A contribution-class codec is used in production. Accordingly, it must have high enough "mask to noise ratio" (that is, the headroom between the actual codec-induced noise level and the just-audible noise level) to allow its output to be processed and/or to be cascaded with other codecs without causing the codec-induced noise to become unmasked. A transmission-class codec, on the other hand, is the final codec used before the listener's receiver. Its main design goal is maximum bandwidth efficiency. Some codecs, like Layer 2, have been used for both applications at different bit rates (and Layer 2 continues to be used as the transmission codec in the Eureka-147 DAR system). However, assuming use of an MPEG codec, modern practice is to use Layer 2 for contribution only (at 256 kbps and higher), reserving transmission for AAC or AAC+SBR. There are many proprietary, non-MPEG codecs other than AC3 available, but these are beyond the scope of this document.

In 2003, the best overall quality for a given data rate in a transmission codec appears to be achieved by the MPEG AAC codec (at rates of 128 kbps or higher) and AAC+SBR (at rates below 128 kbps). The AAC codec is about 30% more efficient than MPEG1 Layer 3 and about twice as efficient as MPEG1 Layer 2. The AAC codec can achieve "transparency" (that is, listeners cannot audibly distinguish the codec's output from its input in a statistically significant way) at a stereo bit rate of 128 kb/sec, while the Layer 2 codec requires about 256 kb/sec for the same quality. The Layer 3 codec cannot achieve transparency at any bit rate, although its performance at 192 kbps and higher is still very good.

Lossy compression is one area where AM practice might diverge from FM and DAB practice. Because of the lower audio resolution of AM at the typical receiver, an AM station trying to economize on storage might want to use a lower data rate than an FM or DAR station. However, this is likely to be false economy if the owner of this library ever wants to use it on FM or DAR in the future. In general, increasing the quality reduces the likelihood that the library will cause problems in future.

Any library recorded for general-purpose applications should use at least 44.1 kHz-sample rate so that it is compatible with DAR systems having 20 kHz bandwidth. If the library will only be used on FM and AM, 32 kHz is adequate and will save considerable storage. However, given the rise of digital radio, we cannot recommend that any future-looking station use 32 kHz for storage.

At this writing, the cost of hard disks is declining so rapidly that there is progressively less argument for storing programming using lossy compression. Of course, either no compression or lossless compression will achieve the highest quality. (There should be no quality difference between these.) Cascading stages of lossy compression can cause noise and distortion to become unmasked. Multiband audio processing can also cause noise and distortion to become unmasked, because multiband processing "automatically re-equalizes" the program material so that the frequency balance is not the same as the frequency balance seen by the psychoacoustic model in the encoder.

Sony's MiniDisk format is a technology that combines data compression (Sony A-TRAC) and random-access disk storage. While not offering the same level of audio quality as CD-R or CD-RW, these disks are useful for field acquisition or other applications where open-reel or cassette tape had been previously used. They offer notably higher quality than the analog media they replace, along with convenient editing,.

Many facilities are receiving source material that has been previously processed through a lossy data reduction algorithm, whether from satellite, over landlines, or over the Internet. Sometimes, several encode/decode cycles will be cascaded before the material is finally aired. As stated above, all such algorithms operate by increasing the quantization noise in discrete frequency bands. If not psychoacoustically masked by the program material, this noise may be perceived as distortion, "gurgling," or other interference. Cascading several stages of such processing can raise the added quantization noise above the threshold of masking, such that it is heard.

In addition, at least one other mechanism can cause the noise to become audible at the radio. The multiband limiter in a broadcast station's transmission processor performs an "automatic equalization" function that can radically change the frequency balance of the program. This can cause noise that would otherwise have been masked to become unmasked because the psychoacoustic masking conditions under which the masking thresholds were originally computed have changed.

Accordingly, if you use lossy data reduction in the studio, you should use the highest data rate possible. This maximizes the headroom between the added noise and the threshold where it will be heard. In addition, you should minimize the number of encode and decode cycles, because each cycle moves the added noise closer to the threshold where the added noise is heard. This is particularly critical if the transmission medium itself (such as DAR, satellite broadcasting, or netcasting) uses lossy compression.

Vinyl Disk

Author's Note for the 2003 Edition:

The next sections devote considerable space to the vagaries of analog media—vinyl disk and analog tape—that are becoming less and less important in broadcast production. However, given that they are still in use, we have chosen to retain this material (with minor editing) in the current revision. Because these media are analog, they require far more tweaking and tender loving care than do the digital media discussed above. For this reason, the following sections are long and detailed.

Some radio programming still comes from phonograph records—either directly, or through dubs. Not only are some club DJs mixing directly on-air from vinyl, but also some old recordings have not been re-released on CD. This section discusses how to accurately retrieve as much information as possible from the grooves of any record.

Vinyl disk is capable of very high-quality audio reproduction. Consumer equipment manufacturers have developed high-fidelity cartridges, pick-up arms, turntables, and phono preamps of the highest quality. Unfortunately, much of this equipment has insufficient mechanical ruggedness for the pounding that it would typically receive in day-to-day broadcast operations.

There are only two reasonably high-quality cartridge lines currently made in the USA that are generally accepted to be sufficiently durable for professional use: the Stanton and the Shure professional series. Although rugged and reliable, these cartridges do not have the clean, transparent operation of the best high-fidelity cartridges. This phono cartridge dilemma is the prime argument for transferring all vinyl disk material to digital media in the production studio, and broadcasting only from digital media. In this way, it is possible (with care) to use state-of-the-art cartridges, arms, and turntables in the dubbing process, which should not require the mechanical ruggedness needed for on-air equipment. A current available state-of-the-art moving-magnet cartridge is the Shure V15VxMR, although this will not survive back-cueing and therefore should never be used live. Good high quality turntables and tonearms have become a bit scarce. However, the Technics SP-10 and its associated base (SH-10B3) and tonearm (EPA-B500/EPA-A250/EPA-A500) are very good choices for mastering vinyl to digital. This reduces the problem of record wear as well.

The following should be carefully considered when choosing and installing vinyl disk playback equipment:

1. **Align the cartridge with great care.**

   When viewed from the front, the stylus must be absolutely perpendicular to the disc, to sustain a good separation. The cartridge must be parallel to the headshell, to prevent a fixed tracking error. Overhang should be set as accurately as possible

±1/16-inch (0.16 cm), and the vertical tracking angle should be set at 20’ (by adjusting arm height).

2. **Adjust the tracking force correctly.**

   Usually, better sound results from tracking close to the maximum force recommended by the cartridge manufacturer. If the cartridge has a built-in brush, do not forget to compensate for it by adding more tracking force according to the manufacturer’s recommendations. Note that brushes usually make it impossible to “back-cue,” which should not be done when transferring to digital anyway.

3. **Adjust the anti-skating force correctly.**

   The accuracy of the anti-skating force calibration on many pick-up arms is questionable. The best way to adjust anti-skating force is to obtain a test record with an extremely high-level lateral cut (some IM test records are suitable). Connect the left channel output of the turntable preamp to the horizontal input of an oscilloscope and the fight channel preamp output to the vertical input. Operate the scope in the X/Y mode, such that a straight line at a 45-degree angle is visible. If the cartridge mistracks asymmetrically (indicating incorrect anti-skating compensation), then the scope trace will be “bent” at its ends. If this happens, adjust the anti-skating until the trace is a straight line (indicating symmetrical clipping).

   It is important to note that in live-disk operations, use of anti-skating compensation may increase the chance of the phono arm sticking in damaged grooves instead of jumping over the bad spots. Increasing tracking force by approximately 15% has the same effect on distortion as applying anti-skating compensation. This alternative is recommended in live-disk operations.

4. **Use a modern, direct-drive turntable.**

   None of the older types of professional broadcast turntables have low enough rumble to be inaudible on the air. These old puck-, belt-, or gear-driven turntables might as well be thrown away! Multiband audio processing can exaggerate rumble to extremely offensive levels.

5. **Mount the turntable properly.**

   Proper turntable mounting is crucial—an improperly mounted turntable can pick up footsteps or other building vibrations, as well as acoustic feedback from monitor speakers (which will cause muddiness and severe loss of definition). The turntable is best mounted on a vibration isolator placed on a non-resonant pedestal anchored as solidly as possible to the building (or, preferably, to a concrete slab). The turntable bases supplied by the turntable manufacturer are highly recommended.

6. **Use a properly adjusted, high-quality phono preamp.**

   Until recently, most professional phono preamps were seriously deficient compared to the best "high-end" consumer preamps. Fortunately, this situation has changed, and a small number of high-quality professional preamps are now available (mostly from small domestic manufacturers). A good preamp is characterized by extremely accurate RIAA equalization, high input overload point (better than 100mV at 1 kHz), low noise (optimized for the reactive source impedance of a real cartridge), low distortion (particularly CCIF difference-frequency IM), load resistance and capacitance that can be adjusted for a given cartridge and cable capacitance, and effective RFI suppression.

   After the preamp has been chosen and installed, the entire vinyl disk playback system should be checked with a reliable test record for compliance with the RIAA equalization curve. (If you wish to equalize the station's air sound to produce a certain "sound signature," the phono preamp is *not* the place to do it.) Some of the better preamps have adjustable equalizers to compensate for frequency response irregularities in phono cartridges. Since critical listeners can detect deviations of 0.5dB, ultra-accurate equalization of the entire cartridge/preamp *system* is most worthwhile.

   The load capacitance and resistance should be adjusted according to the cartridge manufacturer's recommendations, taking into account the capacitance of cables. If a separate equalizer control is not available, load capacitance and resistance may be trimmed to obtain the flattest frequency response. Failure to do this can result in frequency response errors as great as 10dB in the 10–15 kHz region! This is very often the reason many phono cartridge evaluations often produce colored results.

   The final step in adjusting the preamp is to accurately set the channel balance with a test record, and to set gain such that output clipping is avoided on any record. If you need to operate the preamp close to its maximum output level due to the system gain structure, then observe the output of the preamp with an oscilloscope, and play a loud passage. Set the gain so that at least 6dB peak headroom is left between the loudest part of the record and peak-clipping in the preamp.

7. **Routinely and regularly replace styli.**

   One of the most significant causes of distorted sound from vinyl disk reproduction is a worn phono stylus. Styli deteriorate sonically before any visible degradation can be detected even under a microscope, because the cause of the degradation is usually deterioration of the mechanical damping and centering system in the stylus (or actual bending of the stylus shank), rather than diamond wear. This deterioration is primarily caused by back-cueing, although rough handling will always make a stylus die before its time.

   Styli used in 24-hour service should be changed every two weeks as a matter of course—whatever the expense! DJs and the engineering staff should listen constantly for audible deterioration of on-air quality, and should be particularly sensitive to distortion caused by a defective stylus. *Immediately* replace a stylus when problems are detected. One engineer we know destroys old styli as soon as he replaces them so that he is not tempted to keep a stock of old, deteriorated, but usable-looking styli!

It is important to maintain a stock of new spare styli for emergencies, as well as for routine periodic replacement. There is no better example of false economy than waiting until styli fail before ordering new ones, or hanging onto worn-out styli until they literally collapse! Note also that smog- and smoke-laden air may seriously contaminate and damage shank mounting and damping material. Some care should be used to seal your stock of new styli to prevent such damage.

8. **Consider using impulse noise reduction to improve the sound of damaged records.**

   Several impulse noise reduction systems can effectively reduce the effects of ticks and pops in vinyl disk reproduction without significantly compromising audio quality. They are particularly useful in the production studio, where they can be optimized for each cut being transferred to other media.

   With the advent of "plug-in" signal processing architectures for both the PC and Mac platforms, DSP-based signal processing systems have become available at reasonable cost to remove tics, scratches, and noise from vinyl disk reproduction. In a paper like this, designed for reasonably long shelf life, we can make no specific recommendations because the performance of the individual plug-ins is likely to improve quickly. These plug-ins typically cost a few hundred dollars, making them affordable to any radio station.

   At the high end, the line of hardware-based processors made by CEDAR® in England has established itself as being the quality reference for this kind of processing. The CEDAR line is, however, *very* expensive by comparison to the plug-ins described above.

   The only serious rival to CEDAR at this writing is the Sonic Solutions No-Noise® system. This is available as part of the Sonic Solutions workstations for mastering applications.

Analog Tape

Despite its undeniable convenience, the tape cartridge (even at the current state of the art) is inferior to reel-to-reel in almost every performance aspect. Performance differences between cart and reel are readily measured, and include differences in frequency response, noise, high-frequency headroom, wow and flutter, and particularly azimuth and interchannel phasing stability.

Cassettes are sometimes promoted as a serious broadcast program source. We feel that cassettes' low speed, tiny track width, sensitivity to dirt and tape defects, and *substantial* high-frequency headroom limitations make such proposals totally impractical where consistent quality is demanded.

Sum and Difference Recording:

Because it is vital in stereo FM broadcast to maintain mono compatibility, sum and difference recording is preferred in either reel or cart operations. This means that the mono sum signal (L+R) is recorded on one track, and the stereo difference signal (L–R) is recorded on the other track. A matrix circuit restores L and R upon playback. In this system, interchannel phase errors cause frequency-dependent stereo-field localization errors rather than deterioration of the frequency response of the mono sum.

Because this technique tends to degrade signal-to-noise (L+R usually dominates, forcing the L–R track to be under-recorded, thereby losing up to 6dB of signal to-noise ratio), it is important to use a compander-type noise reduction system if sum-and-difference operation is employed.

Electronic Phase Correction

Several manufacturers have sold electronic phase correction devices that they claim eliminate the effects of interchannel phase shifts, although, to our knowledge, none of these is currently being manufactured.

One type of phase correction device measures the cross-correlation between the left and right channels, and then introduces interchannel delay to maximize the long-term correlation. This approach is effective for intensity stereo and pan-potted multitrack recordings (that is, for almost all pop music), but makes frequent mistakes on recordings made with "spaced array" microphone techniques (due to the normal phase shifts introduced by wide microphone spacing), and makes disastrous mistakes with material that has been processed by a stereo synthesizer.

Another type of phase correction device introduces a high frequency pilot tone amplitude modulated at a low-frequency into both the left and light channels. Although the accuracy of this approach is not affected by the nature of the program material, it does require pre-processing of the material (adding the pilot tone), and so may not be practical for stations with extensive libraries of existing, non-encoded material.

It is theoretically possible to use a combination of the cross-correlation and pilot tone phase correction techniques. The cross-correlation circuit should be first, followed by the pilot tone correction circuit. With such an approach, any mistakes made by the cross-correlation technique would be corrected by the pilot tone technique; older material without pilot tone encoding would usually be adequately corrected by cross-correlation. Encoding all synthesized stereo material with pilot tones would prevent embarrassing on-air errors.

Cheap Tape:

Cheap tape, whether reel or cart, is a temptation to be avoided. Cheap tape may suffer from any (or all) of the following problems:

- Sloppy slitting, causing the tape to weave across the heads or (if too wide) to slowly cut away your tape guides.

- Poor signal-to-noise ratio.
- Poor high-frequency response and/or high-frequency headroom.
- Inconsistency in sensitivity, bias requirements, or record equalization requirements from reel to reel (or even within a reel).
- Splices within a reel.
- Oxide shedding, causing severe tape machine cleaning and maintenance problems.
- Squealing due to inadequate lubrication.

High-end, name-brand tape is a good investment. It provides high initial quality, and guarantees that recordings will be resistant to wear and deterioration as they are played. Whatever your choice of tape, you should standardize on a single brand and type to assure consistency and to minimize tape machine alignment problems. Some of the most highly regarded tapes in 1990 use included Agfa PEM468, Ampex 406, Ampex 456, BASF SPR-50 LHL, EMI 861, Fuji type FB, Maxell UD-XL, TDK GX, Scotch (3M) 206, Scotch 250, Scotch 226, and Sony SLH1 1.

In 1999, the situation with analog tape manufacturing is changing rapidly. In the U.S., Quantegy has absorbed the 3M and Ampex lines. A similar consolidation appears to be occurring in Europe.

Tape Speed:

If all aspects of the disk-to-tape transfer receive proper care, then the difference in quality between 15ips (38cm/sec) and 7.5ips (19cm/sec) recording is easily audible. 15ips has far superior high-frequency headroom. The effects of drop-outs and tape irregularity are also reduced, and the effects of interchannel phase shifts are halved. However, a playback machine can deteriorate (due to oxide build-up on the heads or incorrect azimuth) far more severely at 15ips than at 7.5ips before an audible change occurs in audio quality.

Because of recording time limitations at 15ips, most stations operate at 7.5ips. (Many carts will not operate reliably at 15ips, because they are subject to jamming and other problems.) 7.5ips seems to be the lowest that is practical for use in day-to-day broadcast practice. While 3.75ips can produce good results under carefully controlled conditions, there are few operations that can keep playback machines well enough maintained to obtain consistent high quality 3.75ips playback on a daily basis. Use of 3.75ips also results in another jump in sensitivity to problems caused by bad tape, high-frequency saturation, and interchannel phase shift.

Noise Reduction:

In order to reduce or avoid tape hiss, we recommend using a compander-type (encode/decode) noise reduction system in all tape operations. Compander technology was greatly improved in the late 1980s, making it possible to record on analog reel-to-reel at 15ips with quality comparable to 17-bit digital. Even the quality of 7.5ips carts can be dramatically improved. We have evaluated and can enthusiastically recommend Dolby SR (Spectral Recording). Good results have been reported with Telcom C4 as well. dbx Type II noise reduction is also effective and has the advantages of economy, as well as freedom from mistracking due to level mismatches between record and playback.

Remember that to achieve accurate Dolby tracking, record and playback levels must be matched within 2dB. Dolby noise (for SR operations), or the Dolby tone (for Dolby A operations) should always be recorded at the head of all reel-to-reel tapes, and level-matching should be checked frequently. There should be no problem with level-matching if tape machines are aligned every week, as level standardization is part of this procedure. If a different type of tape is put in service, recording machines must be aligned to the new tape *immediately,* before any recordings are made.

In our opinion, all single-ended (dynamic noise filter) noise reduction systems can cause undesirable audible side-effects (principally program-dependent noise modulation) when used with music, and should *never* be used on-line. The best DSP-based systems can be very effective in the production studio (where they can be adjusted for each piece of program material), but even there they must be used carefully, with their operation constantly monitored by the station's "golden ears." Some possible applications include noise reduction of outside production work, and, when placed after the microphone preamp, reduction of ambient noise in the control room or production studio.

Tape Recorder Maintenance:

Regular maintenance of magnetic tape recorders is crucial to achieving consistently high-quality sound. Tape machine maintenance requires expertise and experience. The following points provide a basic guide to maintaining your tape recorder's performance.

1. **Clean heads and guides every four hours of operation.**

2. **Demagnetize heads as necessary.**

   Tradition has it that machines should be demagnetized every eight hours. In our experience, magnetization is usually not a problem in playback-only machines in fixed locations. A magnetometer with a ±5 gauss scale (available from R.B. Annis Co., Indianapolis, Indiana, USA) should be used to periodically check for permanent magnetization of heads and guides. You will find out how long it takes for *your* machines in *your* environment to pick up enough permanent magnetization to be harmful. You may well find that this never happens with playback machines. Recording machines should be watched much more carefully.

3. **Measure on-air tape machine performance weekly.**

   Because tape machine performance usually deteriorates gradually, measure the performance of an on-air machine weekly with standard test tapes. Take whatever corrective action is necessary if the machine is not meeting specifications. Test tapes are manufactured by laboratories such as Magnetic Reference Laboratory (MRL) (229 Polaris Ave. #4, Mountain View, California 94043, USA) and by Standard Tape Laboratory (STL) (26120 Eden Landing Rd. #5, Hayward, California 94545, USA).

4. **Measure flutter weekly.**

   Weekly maintenance should include measurement of flutter, using a flutter meter and high-quality test tape. Deterioration in flutter performance is often an early warning of possible mechanical failure. Spectrum analysis of the flutter can usually locate the flutter to a single rotating component whose rate of rotation corresponds to the major peak in the filter spectrum. Deterioration in flutter performance can, at very least, indicate that adjustment of reel tension, capstan tension, reel alignment, or other mechanical parameter is required.

5. **Measure frequency response and interchannel phase shifts weekly.**

   These measurements, which should be done with a high-quality alignment tape, can be expedited by the use of special swept frequency or pink noise tapes available from some manufacturers (like MRL). The results provide an early indication of loss of correct head azimuth, or of headwear. (The swept tapes are used with an oscilloscope; the pink noise tapes with a third-octave real time analyzer.)

   The head must be replaced or lapped if it becomes worn. Do not try to compensate by adjusting the playback equalizer. This will increase noise unacceptably, and will introduce frequency response irregularities because the equalizer cannot accurately compensate for the shape of the rolloff caused by a worn head.

6. **Record and maintain alignment properly.**

   Alignment tapes wear out. With wear, the output at 15 kHz may be reduced by several dB. If you have many tape machines to maintain, it is usually more economical to make your own "secondary standard" alignment tapes, and use these for weekly maintenance, while reserving your standard alignment tape for reference use. (See below.) However, a secondary standard tape is not suitable for critical azimuth adjustments. These should be made using the methods described above, employing a test tape recorded with a full-track head. Even if you happen to have an old full-track mono machine, getting the azimuth *exactly* right is not practical—use a standard commercial alignment tape for azimuth adjustments.

   The level accuracy of your secondary standard tape will deteriorate with use—check it frequently against your primary standard reference tape. Because ordinary wear does not affect the azimuth properties of the alignment tape, it should have a very long life if properly stored.

   Store all test tapes:

   - Tails out.
   - Under controlled tension.
   - In an environment with controlled temperature and humidity.
   - With neither edge of the tape touching the sides of the reel (this can only be

achieved if the tape is wound onto the storage reel at normal playback/record speeds, and *not* at fast-forward or rewind speed).

7. **Check playback alignment weekly.**

A) Coarsely adjust each recorder's azimuth by peaking the level of the 15 kHz tone on the alignment tape.

Make sure that you have found the *major* peak. There will be several minor peaks many dB down, but you will not encounter these unless the head is totally out of adjustment.

B) While playing back the alignment tape, adjust the recorder's reproduce equalizers for flat high-frequency response, and for low-frequency response that corresponds to the fringing table supplied with the standard alignment tape.

Fringing is due to playing a tape that was recorded full-track on a half track or quarter-track head. The fringing effect appears below 500Hz, and will ordinarily result in an apparent bass boost of 2-3dB at 100Hz.

Fine azimuth adjustment cannot be done correctly if the playback equalizers are not set for identical frequency response, since non-identical frequency response will also result in non-identical phase response.

C) Fine-adjust the recorder's azimuth.

This adjustment is ideally made with a full-track mono pink noise tape and a real-time analyzer. If this instrumentation is available, sum the two channels together, connect the sum to the real-time analyzer, and adjust the azimuth for maximum high-frequency response.

If you do not have a full-track recorder and real-time analyzer, you could either observe the mono sum of a swept-frequency tape and maximize its high-frequency response, or align the master recorder by ear. Adjust for the crispest sound while listening to the mono sum of the announcer's voice on the standard alignment tape (the azimuth on the announcer's voice will be just as accurate as the rest of the tape).

If the traditional Lissajous pattern is used, use *several* frequencies, and adjust for minimum differential phase at *all* frequencies. Using just one frequency (15 kHz, for example) can give incorrect results.

8. **Check record alignment weekly, and adjust as necessary.**

Set record head azimuth, bias, equalization, and calibrate meters according to the manufacturer's recommendations. We recommend that tape recorders be adjusted so that +4dBu (or your station's standard operating level) in and out corresponds to 0VU on the tape recorder's meters, to Dolby level, and to standard operating level. (This is ordinarily 250 nW/m for conventional tape and 315 nW/m for high output tape—refer to the tape manufacturer's specifications for recommended operating fluxivity.)

Current practice calls for adjusting bias with the "high frequency overbias" method (rather than with the prior standard "peak bias with 1.5-mil wavelength" method). To do this, record a 1.5-mil wavelength on tape (5 kHz at 7.5ips) and increase the bias until the maximum output is obtained from this tape. Then *further* increase the bias until the output has decreased by a fixed amount, usually 1.5 to 3dB (the correct amount of decrease is a function of both tape formulation and the width of the gap in the record head—consult the tape manufacturer's data sheet)

9. Follow the manufacturer's current recommendations

In addition to the steps listed above, most tape machines require periodic brake adjustments, reel holdback tension checks, and lubrication. With time, critical bearings will wear out in the motors and elsewhere (such failures are usually indicated by incorrect speed, increased flutter, and/or audible increases in the mechanical noise made by the tape recorder). Use only lubricants and parts specified by the manufacturer.

10. Keep the tape recorder and its environment clean.

Minimize the amount of dust, dirt, and even cigarette smoke that comes in contact with the precision mechanical parts. In addition to keeping dust away from the heads and guides, periodically clean the rest of the machine with a vacuum cleaner (in *suction* mode, please!), or with a soft, clean paintbrush. It helps to replace the filters in your ventilation system at least five times per year.

Recording Your Own Alignment Tapes

Recording a secondary standard alignment tape requires considerable care. We recommend you use the traditional series of discrete tones to make your secondary standard tapes.

A) Using a standard commercial alignment tape, very carefully align the playback section of the master recorder on which the homemade alignment tape will be recorded (see step 7 on page 18).

> *While aligning the master recorder,* write down the actual VU meter reading produced at each frequency on the spot-frequency standard alignment tape.

B) Subtract the compensation specified on the fringing table from the VU meter readings taken in step (A).

> Because you are recording in half-track stereo instead of full-track mono, you will use these compensated readings when you record your secondary standard tape.

C) Excite the record amplifier of the master recorder with pink noise, spot frequencies, or swept tones.

D) Adjust the azimuth of the master recorder's record head, by observing the mono sum from the playback head.

> Pink noise and a real-time analyzer are most effective for this.

If the traditional Lissajous pattern is used, use *several* frequencies, and adjust for minimum differential phase at *all* frequencies.

E) Set the master recorder's VU meter to monitor playback.

F) Record your secondary standard alignment tape on the aligned master recorder.

Use an audio oscillator to generate the spot frequencies. *Immediately* after each frequency is switched in, adjust the master tape recorder's record gain control until the VU meter reading matches the *compensated* meter readings calculated in step (B).

Your homemade tape should have an error of only 0.5dB or so if you have followed these instructions carefully.

Cartridge Machine Maintenance:

The above comments on tape recorder maintenance apply to cart machines as well. However, cart machines have further requirements for proper care—largely because much of the tape guidance system is located *within the cartridge,* and so is quite sensitive to variations in the construction of the individual carts.

1. Clean pressure rollers and guides frequently.

Because lubricated tape leaves lubricant on the pressure rollers and tape guides, frequent cleaning is important in achieving the lowest wow and flutter and in preventing possible can jams. Cleaning should be performed as often as experience proves necessary. Because of the nature of tape lubricant, it does *not* tend to deposit on head gaps, so head cleaning is rarely required.

2. Check head alignment frequently.

Even with the best maintenance, interchannel phase shifts in conventional cart machines will usually prove troublesome. In addition, different brands of cans will show significant differences in phase stability in a given brand of machine. Run tests on various brands of carts, and standardize on the one offering best phase stability.

3. Follow the manufacturer's maintenance and alignment instructions.

Because of the vast differences in design from manufacturer to manufacturer, it is difficult to provide advice that is more specific.

4. Consider upgrading the cart machine's electronics.

Many early (and some not-so-early) cart machines had completely inadequate electronics. The performance of these machines can be improved considerably by certain electronics modifications. Check the machine for the following:

A) record-amplifier headroom (be sure the amplifier can completely saturate the tape before it clips)

B) record amplifier noise and equalization (some record amplifiers can actually contribute enough noise to dominate the overall noise performance of the machine)

C) playback preamp noise and compliance with NAB equalization

D) power supply regulation, noise, and ripple

E) line amplifier headroom

F) record level meter alignment (to improve apparent signal-to-noise ratio at the expense of distortion, some meters are calibrated so that 0 corresponds to significantly more than 1% third-harmonic distortion!)

Probably the most common problem is inadequate record amplifier headroom. In many cases, it is possible to improve the situation by increasing the operating current in the final record-head driver transistor to a value close to its power dissipation limits. This is usually done by decreasing the value of emitter (and sometimes collector) resistors while observing the collector voltage to make sure that it stays at roughly half the power supply voltage under quiescent conditions, and adjusting the bias network as necessary if it does not.

Part 2: System Considerations

Headroom

Other than bad styli, the single most common cause of distorted air sound is probably clipping—intentional (in the audio processing chain) or unintentional (in the program chain). In order to achieve the maximum benefit from processing, there must be *no* clipping before the processor! The gain and overload point of *every* electronic component in the station must therefore be critically reviewed to make sure they are not causing clipping distortion or excessive noise.

In media with limited dynamic range (like magnetic tape), small amounts of peak clipping introduced to achieve optimal signal-to-noise ratio are acceptable. Nevertheless, there is no excuse for *any clipping at all* in the purely electronic part of the signal path, since good design readily achieves low noise and wide dynamic range.

Check the following components of a typical FM audio plant for operating level and headroom:

- Analog-to-digital converters
- Studio-to-transmitter link (land-line or microwave)
- Microphone preamps

- Console summing amplifiers
- Line amplifiers in consoles, tape recorders, etc.
- Distribution amplifiers (if used)
- Signal processing devices (such as equalizers)
- Specialized communications devices (including remote broadcast links and telephone interface devices)
- Phono preamps
- Tape and cart preamps
- Record amplifiers in tape machines

VU meters are worthless for checking peak levels. Even peak program meters (PPMs) are insufficiently fast to indicate clipping of momentary peaks (their integration time is approximately 10ms). While PPMs are excellent for monitoring operating levels where small amounts of peak clipping are acceptable, the peak signal path levels should be monitored with a *true* peak-reading meter or oscilloscope. Particularly, if they are monitoring pre-emphasized signals, PPMs can under-read the true peak levels by 5dB or more. Adjust gains so that peak clipping *never* occurs under any reasonable operating conditions (including sloppy gain riding by the operator).

For older equipment with very soft clipping characteristics, it may be impossible to see a well-defined clipping point on a scope. Or, worse, audible distortion may occur many dB below the apparent clip point. In such a case, the best thing to do is to determine the peak level that produces 1% THD, and to arbitrarily call *that* level the clipping level. Calibrate the scope to this 1% THD point, and then make headroom measurements,

Engineers should also be aware that certain system components (like microphone or phono preamps) have absolute *input* overload points. Difficulties often arise when gain controls are placed *after* early active stages, because the input stages can be overloaded without clipping the output. Many broadcast mic preamps are notorious for low input overload points, and can be easily clipped by high-output mics and/or screaming announcers. Similar problems can occur inside consoles if the console designer has poorly chosen gain structures and operating points, or if the “master” gain controls are operated with unusually large amounts of attenuation.

When operating with nominal line levels of +4 or +8dBu, the *absolute* clipping point of the line amplifier becomes critical, The headroom between nominal line level and the amplifier clipping point should be greater than 16dB. A line amplifier for a +4dBu line should, therefore, clip at +20dBu or above, and an amplifier for a +8dBu line should clip at +24dBu or above. IC-based equipment (which almost always clips at +20dBu or so unless transformer-coupled) is not suitable for use with +8dBu lines. +4dBu lines have become standard in the recording industry, and are preferred for all new studio construction (recording or broadcast) because of their compatibility with IC opamp operating levels.

The same headroom considerations that apply to analog also apply to many digital systems. The only digital systems that are essentially immune to such problems are those that use floating point numbers to compute and distribute the digital data. While floating point arithmetic is relatively common within digital signal processors and mixers, it is very uncommon in external distribution systems.

Even systems using floating-point representation are vulnerable to overload at the A/D converter. If digital recording is used in the plant, bear in mind that the overload point of digital audio recorders (unlike that of their analog counterparts) is abrupt and unforgiving. *Never* let a digital recording go “into the red”—this will almost assuredly add audible clipping distortion to the recording. Similarly, digital distribution using the usual AES3 connections has a very well defined clipping point—digital full-scale—and attempting to exceed this level will result in distortion that is even worse-sounding than analog clipping, because the clipping harmonics above one-half the sampling frequency will fold around this frequency, appearing as aliasing products.

Many systems use digital audio sound cards to provide a means of getting audio signal in and out of computers used to store, process, and play audio. However, not all sound cards have equal performance, even when using digital input and output. For example, a sound card may unexpectedly change the level applied to it. Not only can this destroy system level calibration, but gain can introduce clipping and loss can introduce truncation distortion unless the gain-scaled signal is correctly dithered. If the analog input is used, gain can also introduce clipping, and, in this case, loss can compromise the signal-to-noise ratio. Further, the A/D conversion can introduce nonlinear distortion and frequency response errors.

Level metering in sound cards is highly variable, with average, quasi-peak, and peak responses all common and often inadequately or incorrectly documented. This bears upon the question of line-up level. EBU R68 specifies reference level as –18dBfs, while SMPTE RP 155 specifies it as –20dBfs. If the sound card’s metering is accurate, it will be impossible to ensure compliance with the standards maintained within your facility. Many professional sound cards have adequate metering, while this is far less common on consumer sound cards. Further, consumer sound cards often cannot accommodate professional analog levels.

Many audio editing programs permit a sound file to be “normalized,” which amplifies or attenuates the level of the file to force the highest peak to reach 0 dBfs. This is unwise for several reasons. Peak levels have nothing to do with loudness, so normalized files are likely to have widely varying loudness levels depending on the typical peak-to-average ratio of the audio in the file. Also, if any processing occurs after the normalization process (such as equalization), one needs to ensure such processing does not clip the signal path. If the processing adds level, one must compensate by applying attenuation before the processing to avoid exceeding 0 dBfs, or by using floating point arithmetic. If attenuation is applied, one must use care to ensure that the attenuated signal remains adequately dithered (see page 27).

Voice/Music Balance

The VU meter is very deceptive when indicating the balance between voice and music. The most artistically pleasing balance between voice and music usually results from peaking voice 4–6dB *lower* than music on the console VU meter. If heavy processing is used, the difference between the voice and music levels may have to be increased. Following this practice will also help reduce the possibility of clipping voice, which is much more sensitive to clipping distortion than is most music.

If a PPM is used, voice and music should be peaked at roughly the same level. However, please note that what constitutes a correct "artistic balance" is highly subjective, and different listeners may disagree strongly. Each broadcasting organization has its own guidelines for operational practice in this area. So the suggestions above are exactly that: just suggestions.

It is sometimes difficult to train operators to maintain such a voice/music balance. However, this balance can easily be automated if the console has (or can be modified to have) separate summing amplifiers for live voice and music. Simply build a separate summing amplifier (using a single IC opamp) to drive the VU meter, and then sum the output of the voice-summing amplifier into the VU amplifier with greater gain than the output of the music-summing amplifier.

Electronic Quality

Assuming that the transmission does not use excessive lossy compression, DAR has the potential for transmitting the highest subjective quality to the consumer and requires the most care in maintaining audio quality in the transmission plant. This is because DAR does not use pre-emphasis and has a high signal-to-noise ratio that is essentially unaffected by reception conditions. The benefits of an all-digital plant using minimal (or no) lossy compression prior to transmission will be most appreciated in DAR service.

FM has four fundamental limitations that prevent it from ever becoming a transmission medium that is unconditionally satisfying to "golden-eared" audiophiles. These limitations must be considered when discussing the quality requirements for FM electronics. The problems in disk and tape reproduction discussed above are much more severe by comparison, and the subtle masking of basic FM transmission limitations is irrelevant to those discussions. AM quality at the typical receiver is far worse, and "golden ear" considerations are completely irrelevant because they will be masked by the limitations of the receivers and by atmospheric and man-made noise.

The four FM quality limitations are these:

A) **Multipath distortion**. In most locations, a certain amount of multipath is unavoidable, and this is exacerbated by the inability of many apartment-dwellers to use rotor-mounted directional antennas.

B) The FM stereo multiplex system has a "**sample rate**" of 38 kHz, so its bandwidth is theoretically limited to 19 kHz, and practically limited by the characteristics of "real-world" filters to between 15 and 17 kHz.

C) **Limited IF bandwidth** is necessary in receivers to eliminate adjacent and alternate channel interference. Its effect can be clearly heard by using a tuner with switch-selectable IF bandwidth. Most stations cannot be received in "wide" mode because of interference. But if the station is reasonably clean (well within the practical limitations of current broadcast practice) and free from multipath, then a clearly audible reduction in high-frequency "grit" is heard when switching from "normal" to "wide" mode.

D) Depending on the Region, FM uses either 50μs or 75μs **pre-emphasis**. This severely limits the power-handling capability and headroom at high frequencies and requires very artful transmission processing to achieve a bright sound typical of modern CDs. Even the best audio processors compromise the quality of the high frequencies by comparison to the quality of "flat" media like DAR.

These limitations have considerable significance in determining the cost-effectiveness of current broadcast design practice.

Most older broadcast electronic equipment (whether tube or transistor) is measurably and audibly inferior to modern equipment. This is primarily due to a design philosophy that stressed ruggedness and RFI immunity over distortion and noise, and to the excessive use of poor transformers. Frequency response was purposely rolled off at the extremes of the audio range to make the equipment more resistant to RFI. Cascading such equipment tends to increase both distortion and audible frequency response rolloffs to unacceptable levels.

Modern design practice emphasizes the use of high slew rate, low-noise, low-cost IC operational amplifiers such as the Signetics NE5534 family, the National LF351 family and the Texas Instruments TL070 family. When the highest quality is required, designers will choose premium-priced opamps from Analog Devices, Linear Technology and Burr Brown, or will use discrete class-A amplifiers. However, the 5532 and 5534 can provide *excellent* performance when used properly, and it is hard to justify the use of more expensive amplifiers except in specialized applications like microphone preamps, active filters, and composite line drivers. While some designers insist that only discrete designs can provide ultimate quality, the performance of the best of current ICs is so good that discrete designs are just not cost effective for broadcast applications—especially when the basic FM and DAB quality limitations are considered.

Capacitors have a subtle, but discernible effect upon sonic quality. Polar capacitors such as tantalums and aluminum electrolytics behave very differently from ideal capacitors. In particular, their very high dissipation factor and dielectric absorption can cause significant deterioration of complex musical waveforms. Ceramic capacitors have problems of similar severity. Polyester film capacitors can cause a similar, although less severe, effect when audio is passed through them. Accordingly, DC-coupling between stages is best (and easy with opamps operated from dual-positive and negative power supplies). Coupling capacitors should be used only when necessary (for example, to keep DC offsets out of faders to prevent "scratchiness"). If capacitors must be used, polystyrene, polypropylene, or polycarbonate film capacitors are preferred. However, if it is impractical to eliminate capacitors or to change capacitor types, do not be too concerned: it is probable that other quality-limiting factors will mask the capacitor-induced degradations.

Of course, the number of **transformers** in the audio path should be kept to an absolute minimum. However, transformers are sometimes the only practical way to break ground loops and/or eliminate RFI. If a transformer is necessary, use a high-quality device like those manufactured by Jensen Transformers, Inc., in North Hollywood, California, USA

(Phone (213) 876-0059, or FAX (818) 7634574).

In summary, the path to highest analog quality is that which is closest to a straight wire. More is not better; every device removed from the audio path will yield an improvement in clarity, transparency, and fidelity. Use only the minimum number of amplifiers, capacitors, and transformers. For example, never leave a line amplifier or compressor on-line in “test” mode because it seems too much trouble to remove it. Small stations often sound dramatically superior to their “big time” rivals because the small station has a simple audio path, while the big-budget station has put everything but the kitchen sink on-line. The more equipment the station has (or can afford), the more restraint and self-discipline it needs. Keep the audio path simple and clean! Every amplifier, resistor, capacitor, transformer, switch contact, patch-bay contact, etc., is a potential source of audio degradation. Corrosion of patch-bay contacts and switches can be especially troublesome, and the distortion caused by these problems is by no means subtle.

In **digital signal processing devices**, the lowest number of **bits per word** necessary to achieve professional quality is 24 bits. This is because there are a number of common DSP operations (like infinite-impulse-response filtering) that substantially increase the digital noise floor, and 24 bits allows enough headroom to accommodate this without audibly losing quality. (This assumes that the designer is sophisticated enough to use appropriate measures to control noise when particularly difficult filters are used.) If floating-point arithmetic is used, the lowest acceptable word length for professional quality is 32 bits (24-bit mantissa and 8-bit exponent; sometimes called “single-precision”).

In **digital distribution systems**, 20-bit words (120dB dynamic range) are usually adequate to represent the signal accurately. 20 bits can retain the full quality of a 16-bit source even after as much as 24dB attenuation by a mixer. There are almost no A/D converters that can achieve more than 20 bits of real accuracy, and many “24-bit” converters have accuracy *considerably* below the 20-bit level. “Marketing bits” in A/D converters are outrageously abused to deceive customers, and, if these A/D converters were consumer products, the Federal Trade Commission would doubtless quickly forbid such bogus claims.

There is considerable disagreement about the audible benefits (if any) of **raising the sample rate** above 44.1 kHz. To the author’s knowledge, as of 1999 there have been no rigorous tests of this that are double-blind and that adequately control for other variables, like the performance of the hardware as the sample rate is changed. (Assuming perfect hardware, it can be shown that this debate comes down entirely to the audibility of a given anti-aliasing filter design, as is discussed below.) Nevertheless, in a marketing-driven push, the record industry is attempting to change the consumer standard from 44.1 kHz to 96 kHz. Regardless of whether scientifically accurate testing eventually proves that this is audibly beneficial, it has no benefit in FM stereo because the sampling rate of FM stereo is 38 kHz, so the signal must eventually be lowpass-filtered to 17 kHz or less to prevent aliasing. It is beneficial in DAR, which typically has 20 kHz audio bandwidth, but offers no benefit at all in AM, whose bandwidth is no greater than 10 kHz in any country and is often 4.5 kHz.

Dither is random noise that is added to the signal at approximately the level of the least significant bit. It should be added to the analog signal before the A/D converter, and to any digital signal before its word length is shortened. Its purpose is to linearize the digital system by changing what is, in essence, “crossover distortion” into audibly innocuous random noise. Without dither, any signal falling below the level of the least significant bit will disappear altogether. Dither will randomly move this signal through the threshold of the LSB, rendering it audible (though noisy). Whenever any DSP operation is performed on the signal (particularly decreasing gain), the resulting signal must be re-dithered before the word length is truncated back to the length of the input words. Ordinarily, correct dither is added in the A/D stage of any competent commercial product performing the conversion. However, some products allow the user to turn the dither on or off when truncating the length of a word in the digital domain. If the user chooses to omit adding dither, this should be because the signal in question already contained enough dither noise to make it unnecessary to add more.

In the absence of “**noise shaping**,” the spectrum of the usual “triangular-probability-function (TPF)” dither is white (that is, each arithmetic frequency increment contains the same energy). However, noise shaping can change this noise spectrum to concentrate most of the dither energy into the frequency range where the ear is least sensitive. In practice, this means reducing the energy around 4 kHz and raising it above 9 kHz. Doing this can increase the effective resolution of a 16-bit system to almost 19 bits in the crucial midrange area, and is standard in CD mastering. There are many proprietary curves used by various manufacturers for noise shaping, and each has a slightly different sound.

It has been shown that passing noise shaped dither through most classes of signal processing and/or a D/A converter with non-monotonic behavior will destroy the advantages of the noise shaping by “filling in” the frequency areas where the original noise-shaped signal had little energy. The result is usually poorer than if no noise shaping had been used. For this reason, Orban has adopted a conservative approach to noise shaping, recommending so-called “first-order highpass” noise shaping and implementing this in Orban products that allow dither to be added to their digital output streams. First-order highpass noise shaping provides a substantial improvement in resolution over simple white TPF dither, but its total noise power is only 3dB higher than white TPF dither. Therefore, if it is passed through additional signal processing and/or an imperfect D/A converter, there will be little noise penalty by comparison to more aggressive noise shaping schemes.

One of the great benefits of the **digitization of the signal path** in broadcasting is this: Once in digital form, the signal is far less subject to subtle degradation than it would be if it were in analog form. Short of becoming entirely un-decodable, the worst that can happen to the signal is deterioration of noise-shaped dither, and/or added jitter. Jitter is a time-base error. The only jitter than cannot be removed from the signal is jitter that was added in the original analog-to-digital conversion process. *All* subsequent jitter can be completely removed in a sort of “time-base correction” operation, accurately recovering the original signal. The only limitation is the performance of the “time-base correction” circuitry, which requires sophisticated design to reduce added jitter below audibility. This “time-base correction” usually occurs in the digital input receiver, although further stages can be used downstream.

There are several pervasive myths regarding digital audio:

One myth is that **long reconstruction filters smear the transient response of digital audio**, and that there is therefore an advantage to using a reconstruction filter with a short impulse response, even if this means rolling off frequencies above 10 kHz. Several commercial high-end D-to-A converters operate on exactly this mistaken assumption.

This is one area of digital audio where intuition is particularly deceptive.

The sole purpose of a reconstruction filter is to fill in the missing pieces between the digital samples. These days, symmetrical finite-impulse-response filters are used for this task because they have no phase distortion. The output of such a filter is a weighted sum of the digital samples symmetrically surrounding the point being reconstructed. The more samples that are used, the better and more accurate the result, even if this means that the filter is very long.

It's easiest to justify this assertion in the frequency domain. Provided that the frequencies in the passband and the transition region of the original anti-aliasing filter are entirely within the passband of the reconstruction filter, then the reconstruction filter will act only as a delay line and will pass the audio without distortion. Of course, all practical reconstruction filters have slight frequency response ripples in their passbands, and these can affect the sound by making the amplitude response (but not the phase response) of the "delay line" slightly imperfect. But typically, these ripples are in the order of a few thousandths of a dB in high-quality equipment and are very unlikely to be audible.

The authors have proved this experimentally by simulating such a system and subtracting the output of the reconstruction filter from its input to determine what errors the reconstruction filter introduces. Of course, you have to add a time delay to the input to compensate for the reconstruction filter's delay. The source signal was random noise, applied to a very sharp filter that band-limited the white noise so that its energy was entirely within the passband of the reconstruction filter. We used a very high-quality linear-phase FIR reconstruction filter and ran the simulation in double-precision floating-point arithmetic. The resulting error signal was a minimum of 125 dB below full scale on a sample-by-sample basis, which was comparable to the stopband depth in the experimental reconstruction filter.

We therefore have the paradoxical result that, in a properly designed digital audio system, the frequency response of the system and its sound is determined by the anti-aliasing filter and not by the reconstruction filter. Provided that they are realized with high-precision arithmetic, longer reconstruction filters are always better.

This means that a rigorous way to test the assumption that high sample rates sound better than low sample rates is to set up a high-sample rate system. Then, without changing any other variable, introduce a filter in the digital domain with the same frequency response as the high-quality anti-aliasing filter that would be required for the lower sample rate. If you cannot detect the presence of this filter in a double-blind test, then you have just proved that the higher sample rate has no intrinsic audible advantage, because you can always make the reconstruction filter audibly transparent.

Another myth is that **digital audio cannot resolve time differences smaller than one sample period, and therefore damages the stereo image**.

People who believe this like to imagine an analog step moving in time between two sample points. They argue that there will be no change in the output of the A/D converter until the step crosses one sample point and therefore the time resolution is limited to one sample.

The problem with this argument is that there is no such thing as an infinite-risetime step function in the digital domain. To be properly represented, such a function has to first be applied to an anti-aliasing filter. This filter turns the step into an exponential ramp, which typically has equal pre- and post-ringing. This ramp can be moved far less than one sample period in time and still cause the sample points to change value.

In fact, assuming no jitter and correct dithering, the time resolution of a digital system is the same as an analog system having the same bandwidth and noise floor. Ultimately, the time resolution is determined by the sampling frequency and by the noise floor of the system. As you try to get finer and finer resolution, the measurements will become more and more uncertain due to dither noise. Finally, you will get to the point where noise obscures the signal and your measurement cannot get any finer. However, this point is orders of magnitude smaller in time than one sample period *and is the same as in an analog system.*

A final myth is that **upsampling digital audio to a higher sample frequency will increase audio quality or resolution**. In fact, the original recording at the original sample rate contains all of the information obtainable from that recording. The only thing that raising the sample frequency does is to add ultrasonic images of the original audio around the new sample frequency. In any correctly designed sample rate converter, these are reduced (but never entirely eliminated) by a filter following the upsampler. People who claim to hear differences between “upsampled” audio and the original are either imagining things or hearing coloration caused by the added image frequencies or the frequency response of the upsampler’s filter. They are *not* hearing a more accurate reproduction of the original recording.

This also applies to the **sample rate conversion** that often occurs in a digital facility. It is quite possible to create a sample rate converter whose filters are poor enough to make images audible. One should test any sample rate converter, hardware or software, intended for use in professional audio by converting the highest frequency sinewave in the bandpass of the audio being converted, which is typically about 0.45 times the sample frequency. Observe the output of the SRC on a spectrum analyzer or with software containing an FFT analyzer (like Adobe Audition or Cool Edit). In a professional-quality SRC, images will be at least 90 dB below the desired signal, and, in SRC’s designed to accommodate long word lengths (like 24 bit), images will often be –120 dB or lower.

And finally, some truisms regarding loudness and quality:

Every radio is equipped with a volume control, and every listener knows how to use it. If the listener has access to the volume control, he or she will adjust it to his or her preferred loudness. After said listener does this, the only thing left distinguishing the “sound” of the radio station is its texture, which will be either clean or degraded, depending on the source quality and the audio processing.

Any Program Director who boasts of his station’s $20,000 worth of “enhancement” equipment should be first taken to a physician who can clean the wax from his ears, then forced to swear that he is not under the influence of any suspicious substances, and finally placed gently but firmly in front of a high-quality monitor system for a demonstration of the degradation that $20,000 worth of “enhancement” causes! Always remember that less is more.

Part 3: The Production Studio

The role of the production studio varies widely from station to station. If used only for creation of spots, promos, IDs, etc., production studio quality is considerably less critical than it is where programming is transferred from disk to either tape or cart. Our discussion focuses on the latter case.

Choosing Monitor Loudspeakers

The loudspeakers are the single most important influence on studio quality. The production studio monitor system is the quality reference for all production work, and thus the air sound of the station. Achieving a monitor sound that can be relied upon requires considerable care in the choice of equipment and in its adjustment.

Loudspeakers should be chosen to complement room acoustics. In general, the space limitations in production studios dictate the use of bookshelf-sized speakers. You should assess the effect of equalization or other sweetening on small speakers to make sure that excessive bass or high-frequency boost has not been introduced. While such equalization errors can sound spectacular on big, wide-range speakers, it can make small speakers with limited frequency response and power-handling capacity sound terrible. The Auratone Model 5C Super Sound Cube has frequently been used as a small speaker reference. Although these speakers are no longer manufactured, they are often available on the used market. We recommend that every production studio be equipped with a pair of these speakers or something similar, and that they be regularly used to assure the production operator that his or her work will sound good on small table and car radios.

The primary monitor loudspeakers should be chosen for: high power-handling capacity low distortion high reliability and long-term stability controlled dispersion (omnidirectional speakers are *not* recommended) good tone burst response at all frequencies lack of cabinet diffraction

- relatively flat axial and omnidirectional frequency response from 40-15,000Hz
- physical alignment of drivers (when all drivers are excited simultaneously, the resulting waveforms should arrive at the listener's ears simultaneously, sometimes called “time alignment”).

There are a number of powered midfield monitors available from a large assortment of pro-audio companies, like JBL, Mackie, Genelec, Tannoy, and Alesis, among others. These speakers are very convenient to use because they have built-in power amplifiers and equalizers. Because they have been designed as a system, they are more likely to be accurate than random combinations of power amplifiers, equalizers, and passive loudspeakers. The principal influence on the accuracy of these powered speakers (particularly at low frequencies) is room acoustics and where the speakers are placed in the room. Some of these speakers allow the user to set the bass equalization to match the speaker's location. We believe that such speakers are a logical choice for main monitors in a broadcast production studio.

Loudspeaker Location and Room Acoustics

The bass response of the speakers is strongly affected by their location in the room. Bass is weakest when the speaker is mounted in free air, away from any walls; bass is most pronounced when the speaker is mounted in a comer. Comer mounting should be avoided because it tends to excite standing waves. The best location is probably against a wall at least 18 inches (45 cm) from any junction of walls. If the bass response is weak at this location because the speaker was designed for wall-junction mounting, it can be corrected by equalization (discussed below). It is important that the loudspeakers be located to avoid acoustic feedback into the turntable, because this can produce a severe loss of definition (a muddy sound).

Many successful monitoring environments have been designed according to the “LiveEnd/Dead-End” (LEDE™) concept invented by Don Davis of Synergistic Audio Concepts. Very briefly, LEDE-type environments control the time delay between the arrival of the direct sound at the listener’s ear and the arrival of the first reflections from the room or its furnishings. The delay is engineered to be about 20 milliseconds. This usually requires that the end of the room at which the speakers are mounted be treated with a sound-absorbing material like Sonex® so that essentially no reflections can occur between the speakers’ output and the walls they are mounted on or near. Listeners must sit far enough from any reflective surface to ensure that the difference between the distance from the speaker to the listener and the distance from the speaker to the reflective surface and back to the listener is at least 20 feet (6 meters). It is also desirable that the reflections delayed more than 20 milliseconds be well-diffused (that is, with no flutter echoes). Flutter echoes are usually caused by back-and-forth reflections between two parallel walls, and can often by treated by applying Sonex or other absorbing material to one wall. In addition, “quadratic residue diffusors” (manufactured by RPG Diffusor Systems, Inc.) can be added to the room to improve diffusion and to break up flutter echoes.

An excellent short introduction to the theory and practice of LEDE design is Don Davis’s article, “The LEDE Concept” in Audio Vol.71 (Aug. 1987): p.48-58. (For a more definitive discussion, see Don and Carolyn Davis, “The LEDE Concept for the Control of Acoustic and Psychoacoustic Parameters in Recording Control Rooms.” *J. Audio Eng. Soc.* Vol.28 (Sept. 1980): p.585-95.)

It should be noted that the LEDE technique is by no means the only way to create a good-sounding listening environment (although it is perhaps the best-documented, and has certainly achieved what must be described as a quasi-theological mystique amongst some of its proponents). Examples of other approaches are found in the August 1987 (vol. 29, no. 8), issue of Studio Sound, which focused on studio design.

Loudspeaker Equalization

The performance of any loudspeaker is *strongly* influenced by its mounting location and room acoustics. If room *acoustics are good,* the third-octave real-time analyzer provides an extremely useful means of measuring any frequency response problems intrinsic to the loudspeaker, and of partially indicating problems due to loudspeaker placement and room acoustics.

By their nature, the third-octave measurements combine the effects of direct and reflected sound. This may be misleading if room acoustics are unfavorable. Problems can include severe standing waves, a reverberation time which is not well-behaved as a function of frequency, an insufficient number of “normal modes” (Eigenmodes), lack of physical symmetry, and numerous problems which are discussed in more detail in books devoted to loudspeakers and loudspeaker equalization.

Time-Delay Spectrometry” (TDS) is a technique of measuring the loudspeaker/room interface that provides much more information about acoustic problems than does the third-octave real-time analyzer. TDS (which some sound contractors are licensed to practice) is primarily used for tuning recording studio control rooms, and for adjusting large sound reinforcement systems. The cost may be prohibitive for a small or medium-sized station, particularly if measurements reveal that acoustics can only be improved by major modifications to the mom. However, TDS measurements are highly useful in determining if LEDE criteria are met, and will usually suggest ways by which relatively inexpensive acoustic treatment (absorption and diffusion) can improve room acoustics.

With the advent of low-cost personal computers and sound cards, it is possible to buy economical software to do room analysis and tuning. Since the invention of TDS, a number of other techniques like MLSSA (Maximum-Length Sequence System Analyzer; http://mlssa.com) have been developed for measuring and tuning rooms with accuracy greater than that provided by traditional third-octave analyzers.

It is certainly true that room acoustics must be optimized as far as economically and physically possible *before* electronic equalization is applied to the monitor system. (If room acoustics and the monitor are good, equalization may not be necessary.)

Once room acoustic problems have been solved to whatever extent practical, make frequency response measurements to determine what equalization is required. A MLSSA analyzer, a TDS analyzer or a third-octave analyzer can be used for the measurements. To obtain meaningful results from the analyzer, the calibrated microphone that comes with the analyzer should be placed where the production engineer’s ears would ordinarily be located. If a third-octave analyzer is used, excite each loudspeaker in turn with pink noise while observing the acoustic response on the analyzer. If a MLSSA or TDS analyzer is used, follow the manufacturer’s instructions.

Place the analyzer test mic about 1m from the monitor speaker. Adjust the equalizer (see its operating manual for instructions) to obtain a real-time analyzer read-out that is flat to 5 kHz, and that rolls off at 3dB/octave thereafter. (A truly flat response is not employed in typical loudspeakers, and will make most recordings sound unnaturally bright and noisy.)

If the two channels of the equalizer must be adjusted differently to obtain the desired response from the left and right channels, suspect room acoustic problems or poorly matched loudspeakers. The match is easy to check: just physically substitute one loudspeaker for the other, and see if the analyzer reads the same. Move the microphone over a space of two feet or so while watching the analyzer to see how much the response changes. If the change is significant, then room acoustic problems or very poorly controlled loudspeaker dispersion is likely. If it is not possible to correct the acoustic problem or loudspeaker mismatch directly, you should at least measure the response at several positions and average the results. (Microphone multiplexers can automatically average the outputs of several microphones in a phase-insensitive way—they will help you equalize loudspeaker response properly.)

Although left and right equalizers can be adjusted differently below 200Hz, they should be set close to identically above 200Hz to preserve stereo imaging, even if this results in less than ideal curves as indicated by the third-octave analyzer. (This is a limitation of the third-octave analyzer, which cannot distinguish between direct sound, early reflections, and the reverberant field; stereo imaging is primarily determined by the direct sound.)

A few companies are now making DSP-based room equalizers that attempt to correct both the magnitude and phase of the overall frequency response in the room. (See, for example, http://moose.sofgry.com/SigTech/.) These can produce excellent results if the room is otherwise acoustically well behaved.

Finally, we note once again that the manufacturers of powered nearfield monitors have done much of the work for you. These monitors have built-in equalization, which will often be quite adequate even at low frequencies, provided that the monitor's equalizer can be set to complement the monitor's location in the room.

Stereo Enhancement

In contemporary broadcast audio processing, high value is placed on the loudness and impact of a station compared to its competition. OPTIMOD already has made a major contribution to competiveness. Orban's 222A Stereo Spatial Enhancer augments your station's spatial image to achieve a more dramatic and more listenable sound. Your stereo image will become magnified and intensified; your listeners will also perceive greater loudness, brightness, clarity, dynamics and depth. In use, the 222A detects and enhances the attack transients present in all stereo program material, while not processing other portions. Because the ear relies primarily on attack transients to determine the location of a sound source in the stereo image, this technique increases the apparent width of the stereo soundstage. Since only attack transients are affected, the average L–R energy is not significantly increased, so the 222A does not exacerbate multiple distortion.

Several of Orban's digital audio processors now incorporate the 222A algorithm in DSP.

Other Production Equipment

The preceding discussions of disk reproduction, tape, and electronic quality also apply to the production studio. Compact discs and DVD-Audio discs usually provide the highest quality. For cuts that must be taken from vinyl disk, it is preferable to use "high-end" consumer phono cartridges, arms, and turntables in production. Make sure that *one* person has responsibility for production quality and for preventing abuse of the record playing equipment. Having a single production director will also help achieve a consistent air sound—an important contribution to the "big-time" sound many stations want.

A new generation of low-cost all-digital mixers, made by companies like Soundcraft, Yamaha, Mackie, and Roland, provide the ability to automate mixes and to keep the signal in the digital domain throughout the production process. At the high end, Orban's Audicy digital workstation is oriented towards fast radio production. It combines a dedicated mixing control surface with no-delay RAM-based editing and high-quality built-in digital effects.

Although some people still swear by certain “classic” vacuum-tube power amplifiers (notably those manufactured by Marantz and McIntosh), the best choice for a monitor amplifier is probably a medium-power (100 watts or so per channel) solid-state amplifier with a good record of reliability in professional applications. We do not recommend using an amplifier that employs a magnetic field power supply or other such unusual technology, because these amplifiers literally chop cycles of the AC power line and tend to cause RFI problems.

Do not be tempted to dust off an old Gates or RCA power amplifier and place it in service because it saves you money. It is also usually unwise to use the monitor amplifiers built into most consoles.

Production Practices

The following represents our opinions on production practices. We are aware that some stations operate under substantially different philosophies. But we feel that the recommendations below are rational and offer a good guide to achieving consistently high quality.

1. **Do not apply general audio processing to dubs from commercial recordings in the production studio.**

   OPTIMOD provides all the processing necessary, and does so with a remarkable lack of audible side effects. Further compression is not only undesirable but is likely to be very audible. If the production compressor has a slow attack time (and therefore produces overshoots that can activate gain reduction in OPTIMOD), it will probably “fight” with OPTIMOD, ultimately yielding a substantially worse air sound than one might expect given the individual sounds of the two units.

   If it proves impossible to train production personnel to record with the correct levels, we recommend using the Orban 8200ST (an integrated gated leveler, compressor, high-frequency limiter, and peak clipper that is the successor to Orban’s 464A “Co-Operator”) to protect the production recorder from overload. When used for leveling only, the 8200ST does not affect short-term peak-to-average ratio of the audio, and so will not introduce unnatural artifacts into OPTIMOD processing.

   If production personnel control levels correctly, the 8200ST can be used as a safety limiter and high-frequency limiter by using only the 8200ST compressor function and adjusting its input gain so that broadband gain reduction never occurs when the console VU meters are peaking normally. With this set-up, only high frequencies will be controlled and high-frequency tape saturation will be prevented without adding unwanted broadband compression. (The 8200ST subtle broadband compressor will still prevent tape overload if the console output level is peaked too high.)

2. **Avoid excessive bass and treble boost.**

   Sub-standard recordings can be sweetened with equalization to achieve a tonal balance typical of the best currently produced recordings. However, excessive treble boost (to achieve a certain sound signature for the station) must be avoided if a tape speed of 7.5ips is used, because the tape is subject to high-frequency saturation due to the high-frequency boost applied by the recorder's equalization network. If production is recorded and played on-air digitally, there is no effective limit to the amount of HF boost you can apply in production. However be aware that large amounts of HF boost will stress your on-air AM or FM audio processor because it has to deal with pre-emphasis. We recommend using a modern CD typical of your program material as a reference for spectral balance. Very experienced engineers master major-label CDs using the best available processing and monitoring equipment, typically costing over $100,000 per room in a well-equipped mastering studio. The sound of major-label CDs represents an artful compromise between the demands of different types of playback systems and is designed to sound good on all of them. Mastering engineers do not make these compromises lightly. We believe it is very unwise for a radio station to significantly depart from the spectral balance typical of major-label CDs, because this almost certainly guarantees that there will be a class of receivers on which the station sounds terrible.

3. **Use a compander noise reduction system for all analog-taped material.**

   See the discussion of noise reduction on page 16.

4. **Pay particular attention to the maintenance of production studio equipment.**

   Even greater care than that employed in maintaining on-air equipment is necessary in the production studio, since quality loss here will appear on the air repeatedly. The production director should be acutely sensitized to audible quality degradation and should immediately inform the engineering staff of any problems detected by ear.

5. **Minimize motor noise.**

   To prevent motor noise from leaking into the production microphone, tape machines with noisy motors and computers with noisy fans and hard drives should be installed in alcoves under soffits, and surrounded by acoustic treatment. In the real world of budget limitations this is sometimes not possible, although sound-deadening treatment of small spaces is so inexpensive that there is little excuse for not doing it. But even in an untreated room, it is possible to use a directional microphone (with figure-eight configuration, for example) with the noisy machine placed on the microphone's "dead" axis. Choosing the frequency response of the microphone to avoid exaggerating low frequencies will help. In particularly difficult cases, a noise gate or expander can be used after the microphone preamp to shut off the microphone except during actual speech.

6. **Consider processing the microphone signal.**

   Audio processing can be applied to the microphone channel to give the sound more punch. Suitable equalization may include gentle low- and high-frequency boosts to crispen sound, aid intelligibility, and add a "big-time" quality to the announcer. But be careful not to use too much bass boost, because it can *degrade* intelligibility. Effects like telephone and transistor radio can be achieved with equalization, too

   The punch of production material can often be enhanced by tasteful application of compression to the microphone chain. However, avoid using an excessive amount of gain reduction and excessively fast release time. These cause room noise and announcer breath sounds to be exaggerated to grotesque levels (although this problem can be minimized if the compressor has a built-in expander or noise gate function).

Close-micing, which is customary in the production studio, can exaggerate voice sibilance. In addition, many women's voices are sibilant enough to cause unpleasant effects. High-frequency equalization and/or compression will further exaggerate sibilance. If you prefer an uncompressed sound for production work but still have a sibilance problem, then consider locating a dedicated de-esser *after* all other processing in the microphone chain.

Part 4: Equipment Following OPTIMOD

Some of the equipment following OPTIMOD in the transmission path can also affect quality. The STL, FM exciter, transmitter, and antenna can all have subtle, yet audible, effects.

STL

The availability of uncompressed digital STLs using RF signal paths has removed one of the major quality bottlenecks in the broadcast chain. These STLs use efficient modem-style modulation techniques to pass digitized signals with bit-for-bit accuracy. Provided that the user uses their digital inputs and outputs, and does not require them to do sample rate conversion (which can introduce overshoot if it a downward conversion that filters out signal energy), they are essentially transparent.

Uncompressed digital STLs using terrestrial lines (like T1s in the United States) also provide transparent quality and are equally recommended.

An older digital STL technology uses lossy compression. If the bit rate is sufficiently high, these can be quite audibly transparent. However, all such STLs introduce overshoot and are therefore unsuitable for passing processed audio that has been previously peak limited.

Analog microwave STLs provide far lower quality than either digital technology and are not recommended when high audio quality is desired. They are sometimes appropriate for AM, because receiver limitations will tend to mask quality limitations in the STL.

FM Exciter

Exciter technology has improved greatly in the last twenty years. The most important improvement has been the introduction of digitally synthesized exciters from several manufacturers. This technology uses no AFC loop, and can have frequency response to DC, if desired. It therefore has no problems with bounce or tilt to cause overshoot.

In conventional analog exciter technology, the major improvements have been lowered non-linear distortion in the modulated oscillator, and higher-performance Automatic Frequency Control (AFC) loops with better transient response and lower low-frequency distortion.

At this writing, the state-of-the-art in analog modulated oscillator distortion is approximately 0.02% THD at ±75 kHz deviation. (Distortion in digital exciters is typically 10 times lower than this.) In our opinion, if the THD of your exciter is less than 0.1%, it is probably adequate. If it is poorer than this (as many of the older technology exciters are), replacing your exciter will audibly improve sonic clarity and will also improve the performance of any subcarriers.

Even if the distortion of your modulated oscillator is sufficient, the performance of the AFC loop may not be. A high-performance exciter must have a dual time-constant AFC loop to achieve satisfactory low-frequency performance. If the AFC uses a compromise single time-constant, stereo separation and distortion will be compromised at low frequencies. Further, the exciter will probably not accurately reproduce the shape of the carefully peak- controlled OPTIMOD-FM output, introducing spurious peaks and reducing achievable loudness.

Even dual time-constant AFC loops may have problems. If the loop exhibits a peak in its frequency response at subsonic frequencies, it is likely to “bounce” and cause loss of peak control. (Composite STLs can have similar problems.)[1]

Digital exciters have none of these problems. However, a *properly designed* analog exciter can have good enough performance to limit overshoot due to tilt and bounce to less than 1% modulation. Therefore, either technology can provide excellent results.

FM Transmitter

The transmitter must be transparent to the modulated RF. If its amplifiers are narrowband (< 500 kHz at the -3dB points), it can significantly truncate the Bessel sidebands produced by the FM modulation process, introducing distortion. For best results, –3dB bandwidth should be at least 1MHz.

Narrowband amplifiers can also introduce synchronous FM. This can cause audible problems quite similar to multipath distortion, and can particularly damage SCAs. Synchronous FM should be *at least* –35dB below carrier level, with –40dB or better preferred.[2]

If the transmitter’s group delay is not constant with frequency, it can also introduce synchronous FM, even if the bandwidth is wide. Please note that the “Incidental FM” reading on most FM modulation monitors is heavily smoothed and de-emphasized, and cannot

---

[1] Co-author Greg Oqonowski, Orban’s Vice President of New Product Development, originally brought this to the industry’s attention. (www.indexcom.com). Ogonowski has developed modifications for several exciters and STLs that improve the transient response of their AFCs.

[2] Geoff Mendenhall of Harris has written an excellent practically-oriented paper on minimizing synchronous FM: G. Mendenhall, “Techniques for Measuring Synchronous FM Noise in FM Transmitters,” *Proc. 1987 Broadcast Engineering Conf., National Assoc. of Broadcasters, Las Vegas, NV, p*p.43-52 (Available from NAB Member Services)

be used to measure synchronous FM accurately. At least one device has appeared to do this accurately (Radio Design Labs' Amplitude Component Monitor Model ACM-1).

FM Antenna

Problems with antenna bandwidth and group delay can also cause synchronous FM, as can excessive VSWR, which causes reflections to occur between transmitter and antenna.

Perhaps the most severe antenna-induced problems relate to coverage pattern. Proper choice of the antenna and its correct installation can dramatically affect the amount of multipath distortion experienced by the listener. Multipath-induced degradations are far more severe than *any* of the other quality-degrading factors discussed in this paper. Minimization of received multipath is the single most important thing that the broadcast engineer can do to ensure high quality at the receiver.

AM Transmitter

We live in the golden age of AM transmitters. After 75 years of development, we finally have AM transmitters (using digital modulation technology) that are audibly transparent, even at high power levels. Previously, even the best high-power AM transmitters had a sound of their own, and all audibly degraded the quality of their inputs.

We recommend that any AM station that is serious about quality upgrade to such a transmitter. By comparison to any tube-type transmitter, not only is the quality audibly better on typical consumer receivers, but the transmitter will pay for itself with lower power bills.

AM Antenna

The benefits of a transmitter with a digital modulator will only be appreciated if it feeds an antenna with wideband, symmetrical impedance. A narrowband antenna not only audibly reduces the high frequency response heard at the receiver, but also can cause non-linear distortion in radios' envelope detectors if asymmetrical impedance has caused the upper and lower sidebands to become asymmetrical. Such antennas will not work for any of the AM IBOC systems proposed at this writing.

Correcting antennas with these problems is specialized work, usually requiring the services of a competent consulting engineer.

Summary

Maintaining a high level of on-air audio quality is a very difficult task, requiring constant dedication and a continuing cooperation between the programming and engineering departments.

With the constantly increasing quality of home receivers and stereo gear, the radio audience more and more easily perceives the results of such dedication and cooperation. One suspects that in the future, FM and DAR will have to deliver a state-of-the-art signal in order to compete successfully with the many other program sources vying for audience attention, including CD's, DVD's, videodiscs, digital audio, subscription television, direct satellite broadcast, DTV, streaming programming on the Internet, and who knows how many others!

The human ear is astonishingly sensitive; perceptive people are often amazed when they discover that they can detect rather subtle audio chain improvements on an inexpensive car radio. Conversely, the FM broadcast/reception system can exaggerate flaws in audio quality. Audio processors (even OPTIMOD) are especially prone to exaggerating such flaws.

In this discussion, we have tried to touch upon the basic issues and techniques underlying audio quality in radio operations, and to provide useful information for evaluating the cost-effectiveness of equipment or techniques that are proposed to improve audio quality. In particular, we concluded that today's high-quality IC opamps are ideally suited for use as amplification elements in broadcast, and that compromises in digital standards, computer sound cards, disk playback, and tape quality are all likely to be audible on the air. The all-digital signal path is probably the single most important quality improvement that a station can make, but the installing engineer must be aware of issues such as lossy compression (particularly when cascaded), word length, sample rate, headroom, jitter, and dither.

Following the suggestions presented here will result in better on-air audio quality—and that is a most important weapon in attracting and maintaining an audience that is routinely exposed to compact discs and other high-quality audio reproduction media. The future belongs to the quality-conscious.