



LIFE BEYOND 20 KHZ

By David E. Blackmer

A critical examination
of the limitation of human
hearing will add a new
dimension to your
systems design strategy.

There is much controversy about how we might move forward towards higher quality reproduction of sound. The compact disc standard assumes that there is no useful information beyond 20 kHz and therefore includes a brick wall filter just above 20 kHz. Many listeners hear a great difference when 20 kHz-band limited audio signals are compared to wideband signals. A number of digital systems that sample audio signals at 96 kHz and above with up to 24 bits of quantization have been proposed.

Many engineers have been trained to believe that human hearing receives no meaningful input from frequency components above 20 kHz. I have read many irate letters from such engineers who insist that information above 20 kHz is clearly useless, and any attempt to include such information in studio signals is deceptive, wasteful and foolish. They assert further that any right-minded audio engineer

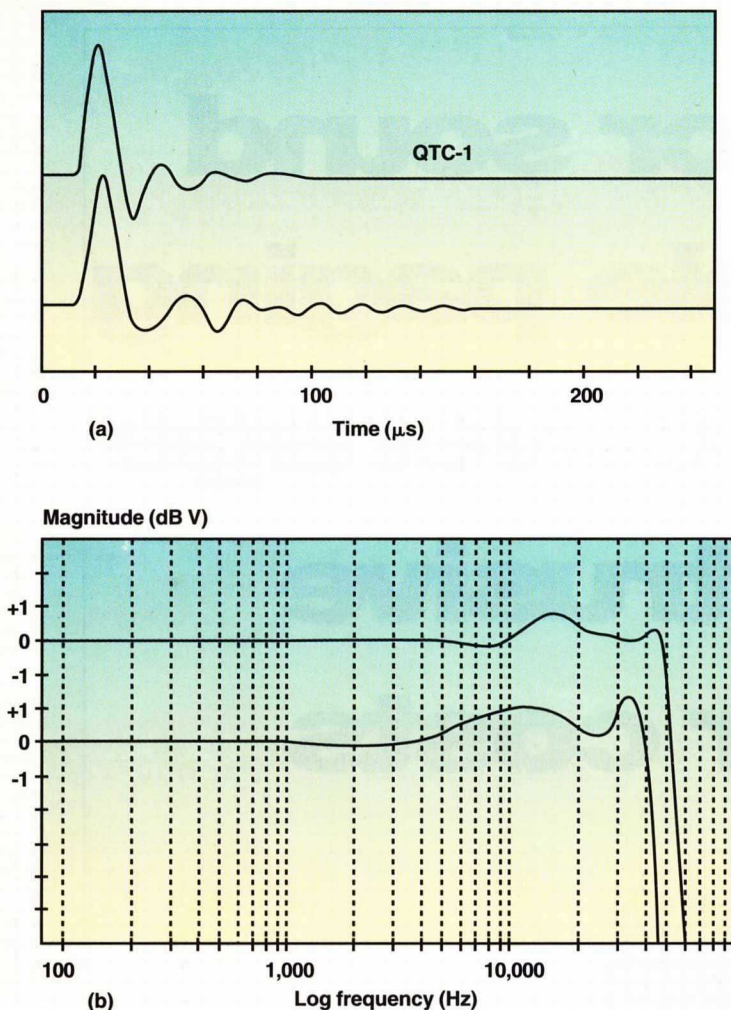


Figure 1. (a) Measured impulse responses of Earthworks' QTC-1 and another omnidirectional mic. (b) Measured frequency response for Earthworks' QTC-1 (top curve) and another omnidirectional mic.

should know that 20 kHz has been acknowledged as an absolute limitation for decades. Those of us who are convinced that there is critically important audio information to at least 40 kHz are viewed as misguided.

So what's going on? We must look at the mechanisms involved in hearing and attempt to understand them. Through that understanding we can then develop a model of the incredible capabilities of the transduction and analysis systems in human audition and thereby work towards better standards for audio system design.

When viewed from an evolutionary standpoint, human hearing has become what it is because it is a survival tool. The human auditory sense is effective at extracting every possible detail from the world around us so that we and our ancestors might avoid danger, find food, communicate, enjoy the sounds of nature and appreciate the beauty of music. Human hearing is generally, I believe, misunderstood to be primarily a frequency-analysis system. The prevalent model of human hearing presumes that auditory perception is based on the brain's interpretation of the outputs of a frequency-analysis system, which is essentially a wide dynamic range comb filter wherein the intensity of each frequency component is transmitted to the brain. This comb filter is certainly an important part of our sound-analysis system. Each frequency zone is tuned sharply with a negative mechanical resistance system.

Furthermore, the tuning Q of each filter element is adjusted in accordance with commands sent back to the cochlea by a series of pre-analysis centers (the cochlear nuclei) near the brain stem. A number of fast transmission rate nerve fibers connect the output of each hair cell to these cochlear nuclei. The human ability to interpret frequency information is amazing; however, clearly something is going on that cannot be explained entirely in terms of our ability to hear tones.

What started me on my quest to understand the capabilities of human hearing beyond 20 kHz was an incident in the late 1980s. I had just acquired an MLSSA system and was comparing the sound and response of a group of high-quality dome tweeters. The best of these had virtually identical frequency response to 20 kHz, yet they sounded different. When I looked closely at their response beyond 20 kHz they were visibly different. The metal dome tweeters had an irregular picket fence of peaks and valleys in their amplitude response above 20 kHz. The silk dome tweeters exhibited a smooth fall off above 20 kHz. The metal dome sounded harsh compared to the silk dome. Admittedly, I cannot hear tones even to 20 kHz, yet the difference was audible. Rather than denying what I clearly heard, I started looking for other explanations, and I have found surprising information hidden in the literature about human hearing.

The inner ear is a complex device with incredible details in its construction. Acoustical pressure waves are converted into nerve pulses in the inner ear, specifically in the cochlea, which is a liquid-filled spiral tube. The acoustic signal is received by the tympanic membrane where it is converted to mechanical forces that are transmitted to the oval window and then into the cochlea where the pressure waves pass along the basilar membrane. This basilar membrane is an acoustically active transmission device. Along the basilar membrane are rows of two different types of hair cells, usually referred to as inner and outer. The inner hair cells clearly relate to the frequency-analysis system described above. Only about 3,000 of the 15,000 hair cells on the basilar membrane are involved in transducing frequency information using the outputs of this travelling wave filter.

The outer hair cells clearly do something else, but what? There are about 12,000 outer hair cells arranged in three or four rows, four times as many as inner hair cells. Only about 20% of the total available nerve paths, however, connect them to the brain. Outer hair cells are interconnected by nerve fibers in a distributed network. This array seems to act as a waveform analyzer, a low-frequency transducer and a command center for the fast muscle fibers (actin) that amplify and sharpen the travelling waves passing along the basilar membrane, thereby producing the comb filter. It also has the ability to extract information and transmit it to the analysis centers in the olivary complex and then on to the cortex of the brain where conscious awareness of sonic patterns takes place. The information from the outer hair cells, which seems to be more related to waveform than frequency, is certainly correlated with the frequency domain and other information in the brain to produce the auditory sense.

Our auditory analysis system is extraordinarily sensitive to boundaries (any significant initial or final event or point of change). One result of this boundary-detection process is the heightened awareness of the initial sound in a complex series of sounds such as a reverberant sound field. This initial sound component is responsible for most of our sense of content, meaning and frequency balance in a complex signal. The human auditory system is evidently sensitive to impulse information imbedded in the tones. My

Continued on page 26.

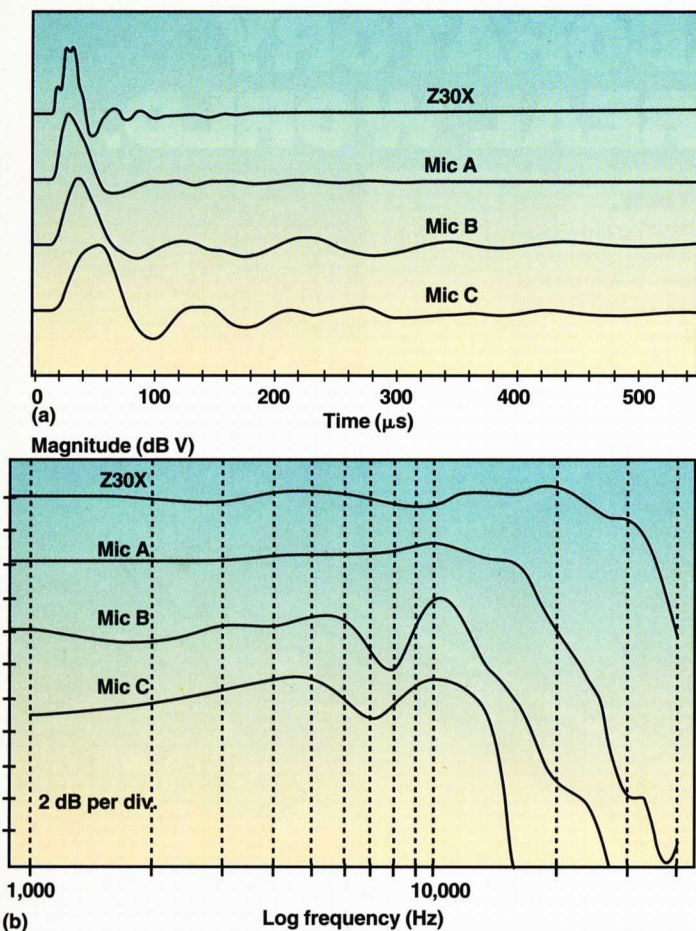


Figure 2. (a) Measured impulse responses of Earthworks' Z30X and three other omnidirectional mics. (b) Measured frequency responses of Earthworks' Z30X and three other omnidirectional mics.

Continued from page 22.

suspicion is that this sense is behind what is commonly referred to as air in the high-end literature. It probably also relates to what we think of as texture and timbre—that which gives each sound its distinctive individual character. Whatever we call it, impulse information is an important part of how we hear.

All output signals from the cochlea are transmitted on nerve fibers as pulse rate and pulse position modulated signals. These signals are used to transduce information about frequency, intensity, waveform, rate of change and time. The lower frequencies are transduced to nerve impulses in the auditory system in a surprising way. Hair cell output for the lower frequencies are transmitted primarily as groups of pulses that correspond strongly to the positive half of the acoustic pressure wave with few if any pulses being transmitted during the negative half of the pressure wave. Effectively, these nerve fibers transmit on the positive half wave only. This situation exists up to somewhat above 1 kHz with discernable half-wave peaks riding on top of the auditory nerve signal being clearly visible to at least 5 kHz. There is a sharp boundary at the beginning and end of each positive pressure pulse group, approximately at the central axis of the pressure wave. This pulse-group transduction with sharp boundaries at the axis is one of the important mechanisms that accounts for the time resolution of the human ear. In 1929, Von Békésy published a measurement of the human sound position acuity, which translates to a time resolution of better than 10 μs between the ears. Nordmark, in a 1976 article, concluded that the interaural resolution is

better than 2 μs; interaural time resolution at 250 Hz is said to be about 10 μs, which translates to better than 1° of phase at this frequency.

The human hearing system uses waveform and frequency to analyze signals. It is important to maintain accurate waveform up to the highest frequency region with accurate reproduction of details down to 5 μs to 10 μs. The accuracy of low-frequency details is equally important. We find many low-frequency sounds, such as drums, take on a remarkable strength and emotional impact when waveform is exactly reproduced. Please notice the exceptional drum sounds on The Dead Can Dance album *Into the Labyrinth*. The drum sound seems to have a very low fundamental, maybe about 20 Hz. We sampled the bitstream from this sound and found that the first positive waveform had twice the period of the subsequent 40 Hz waveform. Apparently, one half cycle of 20 Hz was enough to cause the entire sound to seem to have a 20 Hz fundamental.

The human auditory system, both inner and outer hair cells, can analyze hundreds of nearly simultaneous sound components, identifying the source location, frequency, time, intensity and transient events in each of these sounds simultaneously, and it can spatially map these sounds with awareness of each sound source, its position, character, timbre, loudness and all other identification labels that we can attach to sonic sources and events. I believe that this sound quality information includes waveform, embedded transient identification and high-frequency component identification to at least 40 kHz (even if you cannot hear these frequencies in isolated form). To meet the requirements of human auditory perception, a sound system must cover the frequency range of about 15 Hz to at least 40 kHz (some say 80 kHz or more) with more than 120 dB dynamic range to handle transient peaks properly and with a transient time accuracy of a few microseconds at high frequencies and 1° or 2° phase accuracy down to 30 Hz. This standard is beyond the capabilities of modern systems, but it is important that we understand the degradation of perceived sound quality resulting from compromises made in today's sound-delivery systems. The transducers are the most obvious problem areas, but the storage systems and all the electronics and interconnections are important, too.

Mics are the first link in the audio chain, translating the pressure waves in the air into electrical signals. Many of today's mics are not accurate, and few have accurate frequency response over the entire 15 Hz to 40 kHz range. In most mics, the active acoustic device is a diaphragm that receives the acoustical waves, and like a drum head, it will ring when struck. To make matters worse, the pickup capsule is usually housed in a cage with many internal resonances and reflections, further coloring the sound. Directional mics, because they achieve directionality by sampling the sound at multiple points, are by nature less accurate than omnidirectional mics. The ringing, reflections and multiple paths to the diaphragm add up to excess phase. These mics smear the signal in the time domain.

At Earthworks, we have learned after many measurements and careful listening that the true impulse response of mics is a better indicator of sound quality than frequency amplitude response. Mics with long and asymmetrical impulse performance will be more colored than those with short impulse tails. To illustrate this point, we have carefully recorded a variety of sources using two different omnidirectional mics (Earthworks QTC1 and another well known model), both of which have flat frequency response to 40 kHz within ±1 dB. (See Figure 1.) When played back on high-quality loudspeakers, the sounds of these two mics is quite different. When played back on loudspeakers with nearly perfect impulse and step response, the sounds of the two

S m a l l



Wonder

The miraculous DPA 4060 Miniature Microphone tops the bill in wireless systems for theatre and television. Not only does the 4060 offer outstanding audio performance under difficult conditions, it is also extremely robust in operation. Unique connection adapters ensure compatibility with a wide variety of VHF and UHF systems. A range of sensitivities encompass the vast majority of applications where high quality audio, near invisibility and lightness is required. Developed from many years of professional audio experience, the 4060 is just one of the high quality products from the renowned 4000 series - available now from DPA Microphones.

Series 4000 Microphones from DPA

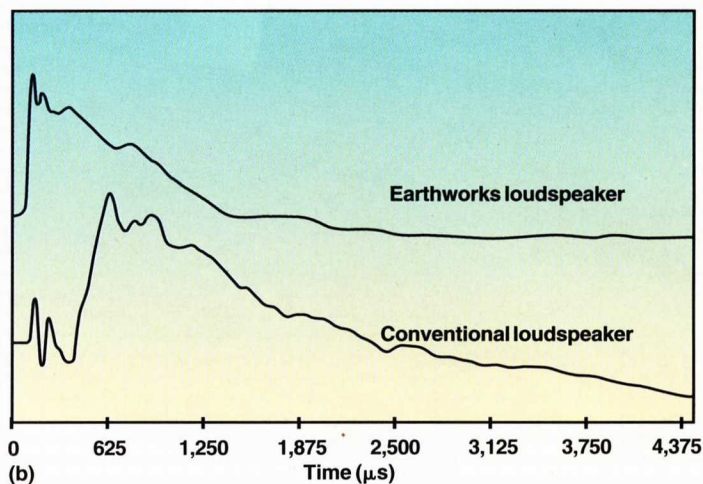
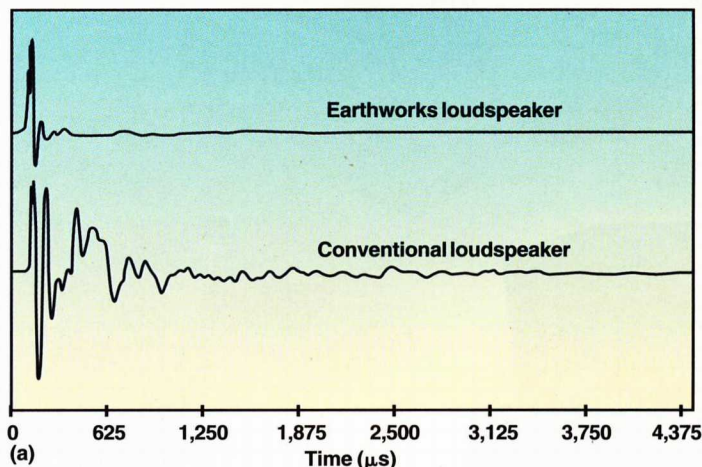
Hejrevang 11, 3450 Allerød, Denmark
T: +45 48142828 F: +45 48142700

NORTH AMERICAN DISTRIBUTORS:

TGI North America Inc. 300 Gage Ave,
Unit 1 Kitchener, Ontario N2M 2C8, Canada
T: 519 745 1158 F: 519 745 2364

www.dpamicrophones.com

DPA
MICROPHONES



Transfer function mag—dB V/V

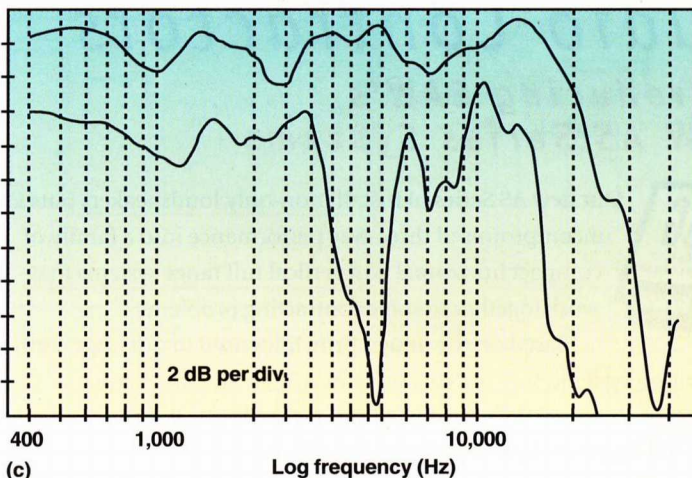


Figure 3. (a) Measured impulse responses of an Earthworks experimental loudspeaker and a conventional three-way loudspeaker. (b) Measured step responses of an Earthworks experimental loudspeaker and a conventional three-way loudspeaker. (c) Measured frequency responses for the Earthworks experimental loudspeaker (top curve) and a conventional three-way loudspeaker.

mics vary even more widely. The only meaningful and identifiable difference between these two mics is their impulse response. We have developed a system for deriving a mic's frequency

response from its impulse response. After numerous comparisons between the results of our impulse conversion and the results of the more common substitution method, we are convinced

of the validity of this as a primary standard. You will see several examples of this in Figure 2. Viewing the waveform as impulse response is better for interpreting high-frequency information. Low-frequency information is more easily understood from inspecting the step function response, which is the mathematical integral of impulse response. Both curves contain all information about frequency and time response within the limits imposed by the time window, the sampling processes and noise. The electronics in high-quality sound systems must also be exceptional. Distortion and transient intermodulation should be held to a few parts per million in each amplification stage, especially in systems with many amps in each chain. In the internal circuit design of audio amps, it is especially important to separate the signal reference point in each stage from the power supply return currents that are usually terribly nonlinear. Difference input circuits on each stage should extract the true signal from the previous stage in the amp. Any overall feedback must reference from the output terminals and compare directly to the input terminals to prevent admixture of ground grunge and crosstalk with the signal. Failure

Errors in reverberant sound are accepted as normal unless there are bad standing waves in the listening space.

to observe these rules results in transistor sound. Transistors can be used in a manner resulting in an arbitrarily low distortion, intermodulation, power supply noise coupling and whatever other errors we can name and can therefore deliver perceptual perfection in audio signal amplification. (I use perceptual perfection to mean a system or component so excellent that it has no error perceptible to the best human hearing.) My current design objective on amps is to have all harmonic distortion, including 19 kHz and 20 kHz twin tone intermodulation products, below one part per million and to have a weighted noise at least 130 dB below maximum sine wave output. I assume that a signal can go through many such amps in a system with no detectable degradation in signal quality.

Many audio signal sources have extremely high transient peaks, often as high as 20 dB above the level read on a volume indicator. It is important to have some adequate measurement tool in an audio amplification system to measure peaks and to determine that they are being handled appropriately. Many of the available peak reading meters do not read true instantaneous peak levels but respond to something closer to a 300 μ s to 1 ms averaged peak approximation. All system components, including power amps and loudspeakers, should be designed to reproduce the original peaks accurately. Recording systems truncate peaks beyond their capability. Analog tape recorders often have a smooth compression of peaks, which is often regarded as less damaging to the sound. Many recordists even like this peak clipping and use it intentionally. Most digital recorders have a brick-wall effect in which any excess peaks are squared off with disastrous effects on tweeters and listeners' ears.

Compressors and limiters are often used to reduce peaks that would otherwise be beyond the capability of the system. Such units with RMS level detectors usually sound better than those with average or quasi-peak detectors. Also, be careful to select signal processors for low distortion. If they are well designed, distortion will be low when no gain change is required. Distortion during compression will be almost entirely third harmonic distortion, which is not easily detected by the ear and is usually reasonably acceptable when audible. A look at the specifications of some of the highly rated, super-high-end, no-feedback, vacuum-tube power amps reveals how much distortion is acceptable (or even preferable) to some well heeled audiophiles. All connections between different parts of the electrical system must be designed to eliminate noise and signal errors due to power line ground currents, AC magnetic fields, RF pickup, crosstalk and dielectric absorption effects of poor wire insulation. This is critical.

Loudspeakers are the other end of the audio system in that they convert the electrical signal back into pressure waves in the air. Loudspeakers are usually less accurate than mics. Too many of our common sound systems are below the capabilities of today's technology. Listen to cinema sound, for example. Enormous improvement has been made in the delivery of high-quality digital sound to the theater, but cinema loudspeakers are almost always horn loudspeakers. Horn loudspeakers can be quite good except when they must also possess constant directivity, which is usually achieved by adding a sharp discontinuity in the internal horn profile. Such loudspeakers often have so many nearly equal level internal reflections that no current DSP system could adequately correct their sound.

To make matters worse, the powers that be have decreed that any good theater must be equalized, which all too often means placing one or more test mics at ear level in representative seats in the auditorium and adjusting for flat response with a $1/3$ octave EQ and matching analyzer. So what's wrong with this? When we listen in a reverberant space, we selectively give a strong preference to first

Imagine Design Build



RDL™

Radio Design Labs

Specialists In Practical Precision Engineering

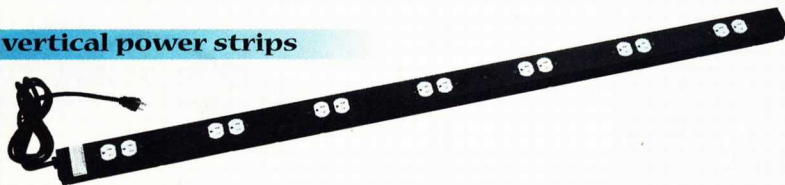
USA Phone (800) 281-2683 Europe Phone (+31) 20-6238 983
USA Fax (805) 684-9316 Europe Fax (+31) 20-6225 287
Internet Access: www.rdl.net

Circle (22) on Rapid Facts Card

See us at AES—Booth #1752

...another **POWERFUL** addition to the product line

vertical power strips



rackmount sequencing



Please request
our all-new
68 page
Master Catalog
for detailed
new product
information



Middle Atlantic Products, Inc.
POWER
North Corporate Drive, Riverdale NJ 07457
800-266-7225 voice • 800-392-3955 fax
973-839-1011 voice • 973-839-1976 fax

Circle (24) on Rapid Facts Card

CONTRACTOR SERIES Architecture in Sound



SURFACE MOUNT & FLUSH MOUNT SPEAKERS
Exceptional acoustics for ambient music and paging

Sound solutions for
retail outlets, bars, hotels, restaurants, health clubs, airports...



The Martin Experience

Martin Audio Limited (TGI N.A.), P.O. Box 44019, Kitchener, Ontario, Canada N2N 3G7
Telephone: (519) 578-0213 Facsimile: (519) 578-1575
www.martin-audio.com e-mail: info@martin-audio.com

Circle (25) on Rapid Facts Card

arrival in judging sound quality. The reverberant sound field arrives later and is perceived to be less important. It is therefore beneficial to achieve good frequency response in first arrival sound. Errors in reverberant sound are accepted as normal unless there are bad standing waves in the listening space, which must be corrected with appropriate physical acoustic treatment. You cannot solve room acoustics problems with filters. Time-windowed analyzers, such as MLSSA, SMAART or TEF, should be used for all tuning of listening spaces. If a time windowed-analyzer is not available, the measurement mic should be placed nearer to the active loudspeaker, optimally between one third and one half the distance from the loudspeaker to the mid-audience seat chosen as a reference point so that the direct loudspeaker sound dominates in the mic input signal.

There has been a dramatic improvement in the tools available to measure systems and components. Many manufacturers have improved their amps, recorders and transducers. Those of us involved in equipment design and/or sound system design must learn both to recognize and reward excellence from our suppliers and to deliver excellent sound to our customers. I suspect that truly excellent sound, perhaps even perceptual perfection, especially in large spaces, must await the development of a high-accuracy, high-power, direct-radiating 40 kHz tweeter system with inherently good impulse response integrated into a system that gives good impulse and step function response over the entire listening area.

I have heard that the Victor Talking Machine Company ran ads in the 1920s in which Enrico Caruso was quoted as saying that the Victrola was so good that its sound was indistinguishable from his own voice live. In the 1970s, Acoustic Research ran similar ads (with considerably more justification) about live vs. recorded string quartets. We have come a long way since then, but can we achieve perceptual perfection? As a point of reference, you should assemble a test system with both mics and loudspeakers having excellent impulse and step response, hence nearly perfect frequency response, together with low distortion amps. Test it as a sound-reinforcement system and/or studio monitoring system with both voice and music sources. You, the performers and the audience will be amazed at the result.

S&VC

Blackmer, inventor of the dbx VCA, is president of Earthworks Audio Products, Wilton, NH. He can be reached at 603-654-6427 or via Web site at www.earthwks.com.