



Goldmund White Paper

DOES HIGH-RESOLUTION AUDIO SOUND BETTER?

Audiophiles, audio professionals and scientists tend to disagree about whether high-resolution digital audio sounds better than standard, CD-quality audio. This paper examines the technical and psychoacoustic issues to find an answer.

With the recent advent of high-resolution downloads, audiophiles have enthusiastically embraced high-resolution digital audio. Digital-to-analog converters (DACs) are now expected to include high-resolution capability, and are often judged by the maximum digital resolution and file formats they support. High-resolution audio technology has been a standard and expected feature in professional digital audio equipment for more than a decade.

However, audio researchers and scientists have been less eager to embrace high-resolution audio. They point out technical problems with the various formats. They question whether the theoretical advantages of high-resolution audio can have any benefit given the limits of human hearing and the typical listening environment. They point to the fact that the difference between standard-resolution and high-resolution audio is only rarely detectable in controlled, blind listening tests.

Which side is correct? Or is the truth somewhere in between? This paper will examine the issues and come to a satisfactory conclusion.

Potential Advantages of High-Resolution Audio

Proponents of high-resolution audio cite two advantages for the technology: greater bandwidth and greater dynamic range (or less noise).

Greater Bandwidth: The greater bandwidth of high-resolution audio comes from higher sampling rates. Standard-resolution pulse-code modulation (PCM) digital audio on CDs has a sampling rate of 44.1 kilohertz, which yields a frequency response of approximately 20 kHz. High-resolution PCM audio typically uses sample rates of 96 or 192 kHz, for frequency response of slightly less than 48 and 96 kHz, respectively. Direct Stream Digital (DSD) high-resolution audio in its standard version use a sample rate of 2.8 megahertz; it achieves a frequency response of approximately 100 kHz, although its noise and resolution performance at high frequencies is poor.

The highest frequency humans can hear is about 20 kHz; this number decreases with age, especially in males. What, then, is the advantage of capturing higher frequencies?

Improved Timing Resolution: Some of the proponents of high-resolution audio cite the timing resolution of the human ears. Sound almost always arrives at the left and right ears at slightly different times. Humans can detect timing differences between the right and left ears as small as 15 microseconds; to capture this information, a sampling frequency of 66.7 kHz is required. Omitting frequencies above 20 kHz eliminates these subtle timing differences, which could, at least in theory, affect the perception of sound.

Reduced Filtering Demands: In PCM digital audio, the signal is filtered during recording to eliminate aliasing (interference with audio band signals by spurious high-frequency energy), and during playback to reconstruct a smooth sound wave from the “stair step” quantized output of a digital-to-analog converter. Typically, a very steep filter slope (36 to 48 dB per octave) is used so that frequencies below the sampling frequency are affected as little as possible, while frequencies above the sampling frequency are strongly attenuated. The steepness of the filter introduces phase shift and ripples in the device’s frequency response, producing measurable effects at frequencies near the cutoff frequency of the filter.

With high-resolution PCM audio, because the sampling rate is higher, the cutoff frequency of the anti-aliasing and reconstruction filters can also be higher, and a gentler filter slope such as 12 dB per octave can be used. The sampling rate of standard DSD is roughly 2.8 MHz, so its filters can be set to frequencies very far beyond the audio range. Thus, any audible effects of these filters would be greatly reduced or eliminated entirely.

Greater Dynamic Range, Lower Noise: The greater dynamic range and lower noise of high-resolution audio come from its increased amplitude resolution. Standard-resolution PCM digital audio uses 16-bit sampling, which can record 65,536 different signal levels, for a theoretical dynamic range 96 dB (6 dB per bit). This means that the inherent noise level is -96 dB lower than the highest level of sound the system can record. High-resolution PCM recordings using 24-bit sampling can capture 16,777,216 different signal levels, for a dynamic range of 144 dB. Thus, the noise falls to -144 dB below the maximum level, although even the best amplifiers and digital-to-analog converters have a noise floor of about -120 dB. (DSD does not offer a clear advantage in this area because its dynamic range varies with frequency, from better than -100 dB in the bass to as little as -6 dB at ultrasonic frequencies.)

Audiophiles have cited various benefits of the extra amplitude resolution of high-resolution audio, including a greater sense of detail and dynamics, quieter backgrounds and a more natural sound overall. Professional engineers record using high-resolution formats partly because it makes setting recording levels less critical; they can set the level conservatively to avoid exceeding the maximum signal level (i.e., clipping) while still maintaining ample dynamic range. Professionals also cite the greater precision of high-resolution audio as reducing artifacts sometimes encountered in digital post-processing of audio.

Potential Disadvantages of High-Resolution Audio

We have seen much discussion in consumer and professional audio publications of the potential advantages of high-end audio, but little discussion of the potential technical disadvantages – disadvantages that are likely more audible than an increase in dynamic range or frequency response.

Increased Distortion: The main audible flaw of high-resolution audio involves intermodulation distortion (IMD), the effect created when two audio tones interfere with each other. With IMD,

sum and difference tones are created, usually at frequencies that are not harmonically related to the two original tones. IMD occurs in all audio equipment to some degree, but decades of design evolution have reduced it to insignificance.

However, IMD is a more common problem at ultrasonic frequencies. Equipment not designed to reproduce such high frequencies – including many amplifiers and most of the high-frequency drivers (tweeters) used in today’s speakers – may produce substantial IMD if forced to operate at frequencies they were not designed to handle. Unfortunately, the effects of IMD are not limited to high frequencies.

For example, if a high-resolution recording contains tones at 28 and 30 kHz, a speaker or amplifier that is prone to high-frequency IMD will reproduce (or attempt to reproduce) not only the 28 and 30 kHz tones, but also the sum and difference tones. The difference tone – 30,000 minus 28,000 – will occur at 2 kHz, right in the middle of the frequency range in which the human ear is most sensitive. Thus, the assumption that extending an audio system’s high-frequency capability will always be beneficial is incorrect.

A 2001 paper titled “Detection of Threshold for Tones Above 22 kHz,” by researchers working in Japan’s National Institute of Advanced Industrial Science and Technology, confirmed this thesis. The researchers used test signals that combined a 2 kHz tone played with and without ultrasonic harmonics. When a single speaker was used to reproduce the sound, IMD occurring in the playback system allowed listeners to detect the presence of the ultrasonic harmonics. When a second speaker/amplifier system was used to reproduce the ultrasonic harmonics, and the original speaker/amplifier system reproduced only the 2 kHz tone, the listeners could not detect the ultrasonic tones. This suggests that while the IMD caused by the ultrasonic tones was audible, the ultrasonic tones were not, even though they were recorded at the same level as the 2 kHz tone.

Potential Reduction in Equipment Lifespan: Typical tweeters start to reach their breakup modes – the frequencies at which their physical components behave in a non-linear manner – at frequencies between 25 and 30 kHz. When breakup modes occur, the tweeter diaphragm (dome) distorts out of its original shape, creating wave patterns in the formerly smooth diaphragm. Constantly distorting the diaphragm by exciting these breakup modes can result in physical fatigue of the diaphragm and other mechanical components of the driver, causing distortion and possible failure of the driver.

Most mass-market amplifier and preamplifier circuits filter out ultrasonic frequencies in order to avoid oscillation, a state in which the circuit spontaneously generates high-amplitude, high-frequency tones and quickly burns itself out. However, this filtering is not total or perfect. As any amplifier technician can probably attest from personal experience, forcing an audio circuit to reproduce high frequencies at high levels often causes failure of the electrical components in the circuit. This may not be a problem with the best high-end amplifiers and preamps, because many of them are specifically designed to handle ultrasonic frequencies. But it can be a problem with lower-quality components connected to a high-end system, or with lower-quality systems elsewhere in a home.

Does High-Resolution Audio Really Outperform CD Audio?

Although the purported technical benefits of high-resolution audio are often touted in marketing materials and audio publications, these benefits are often already achieved in standard-resolution systems.

Oversampling to Reduce Filtering Demands: The benefits of sampling at frequencies higher than 44.1 kHz can be and are achieved by using oversampling in standard-definition analog-to-digital and digital-to-analog conversion.

Almost all currently available analog-to-digital converter (ADC) and digital-to-analog converter (DAC) chips oversample at high rates, typically 96 kHz or higher. With oversampling, the anti-aliasing filter in the ADC and the reconstruction filter in the DAC can employ higher frequencies and/or more gradual slopes. If the audio is stored in standard resolution, the extra samples are simply discarded – yet the benefits of the higher filter frequency and/or more gradual filter slopes remain, while the problem of IMD at ultrasonic frequencies is eliminated.

Dynamic Range Beyond 96 dB: While it is generally accepted that recording with greater bit depth allows quieter sounds to be captured, 16 bits is actually enough to capture any sound that can be reproduced through even the best audio equipment.

The noise floor of a digital audio system is usually cited as minus 6 dB per bit – i.e., a 16-bit digital audio system has a noise floor 96 dB below the maximum recordable signal level. However, the common perception in audio is that the noise floor represents an impermeable barrier beneath which nothing can be heard, and this is incorrect.

In PCM digital audio recording, the system does not simply shut off when the signal level falls below -96 dB. Dither – a small amount of noise – is added to the signal so that the system does not exclude information below the theoretical minimum signal level. Typically, dither is limited to higher frequencies where it is not audible. The result can further be improved through noise-shaping techniques that shift the inherent noise of digital audio systems to ultrasonic frequencies. The result is that the noise floor of a 16-bit system within the audible frequency range can be as good as -110 dB or even -120 dB in practice.

Dynamic Resolution Not Increased: It is widely accepted that high-resolution audio's increased bit depth results in more precise definition of audio levels, and thus better rendition of so-called "microdynamics." But again, this is not the case. The increased bit depth merely allows greater dynamic range: higher maximum and minimum recordable signal levels. For example, a kick drum sample that is +46.324 dB louder than a note from a flute will be +46.324 dB louder whether a 16-bit or 24-bit system is used. Adding dynamic range beyond 16 bits is analogous to adding wider range to a radar used to measure the speed of cars; a radar that can measure speeds from 0.01 to 100,000 kilometers per hour is no more useful in this application than one that can measure speeds from 1 to 1,000 kilometers per hour.

There is one application in which higher bit depth can have benefits: in professional audio recording and editing. Audio often goes through many generations of processing during recording, editing and post-production, and each step can add noise. By using 24-bit processing, added noise is minimized. When the audio production is finished, it can then be converted to 16 bits for distribution.

Using bit depths greater than 16 bits in consumer audio systems introduces no ill effects, it merely wastes space, requiring greater amounts of storage and longer download times.

Are the Advantages of High-Resolution Audio Even Detectable?

Even if one embraces the purported benefits of high-resolution audio while disregarding the potential downsides, it is difficult and perhaps even impossible to make the case that high-resolution audio yields any perceptible benefit under even the best listening conditions.

Limited Dynamic Range of the Listening Environment: Large, professionally installed home theater systems are calibrated to a reference sound pressure level (SPL) of 105 dB for each main channel. However, most people consider this level too loud for comfortable listening. (For reference, a jackhammer produces approximately 100 dB SPL at 1 meter.) Even most home theater enthusiasts listen at levels about -6 dB lower. Rarely are two-channel audio systems played at levels exceeding 100 dB.

On the opposite end of the dynamic range, the best professional recording studios have a noise floor around 30 dB SPL. For a professionally installed, acoustically isolated and treated home theater, the noise floor might be about 40 dB SPL. For a living room, the noise floor is typically about 50 dB SPL – or even louder if the living room is open to other parts of the house.

Thus, even when considering an extremely high listening level of 110 dB, in an acoustically isolated listening room with a 40 dB SPL noise floor, the dynamic range of the listening environment is only 70 dB. This is well within the commonly assumed 96 dB dynamic range of standard resolution digital audio.

Ultrasonic Frequencies are Undetectable by Ear: Research has failed to show a significant audible improvement gained through increased sampling rates. One of the few studies that has shown a statistically significant ability of listeners to perceive the effects of increasing sampling rate beyond 44.1 kHz – “Sampling Rate Discrimination: 44.1 kHz vs. 88.2 kHz,” a 2014 paper by researchers affiliated with McGill University in Canada – concluded that such differences are “very subtle and difficult to detect.” It is important to note that this test did not evaluate the listeners’ *preference* for the higher sample rate, only whether or not they could distinguish it from the lower rate – which they could on only one of the five musical selections in the test.

The reason why the extended frequency response of high-resolution recordings is so difficult to perceive is that the human ear simply does not possess the physical means of detecting it. The ear detects sounds using a series of hair cells that run along the basilar membrane. The position of each hair cell on the membrane tunes it to receive a certain narrow frequency range of sound. Above the highest tuning frequency of the hair cells, no sound is detected.

Thus, the response of the human ear does not gradually taper off into infinity, as is commonly assumed. It does decline as the frequency of sound approaches 20 kHz, but if the sound is at a frequency higher than that of the highest-tuned hair cell, no sound is detected at all. This is much like the limits of human eyesight. We can detect a certain range of light frequencies, but light in the infrared and ultraviolet frequencies is entirely undetectable through our eyes.

Does the DSD Format Offer Any Advantages Over PCM?

Some audiophiles, audio equipment manufacturers and recording professionals believe that above and beyond DSD’s extended frequency response compared with 16/44.1 PCM, it offers inherently better and more natural sound than PCM. However, other than anecdotal

reports gathered from sighted listening tests, there has been little, if any, evidence presented to support this claim.

A 2004 paper titled “DVD-Audio vs. SACD: Perceptual Discrimination of Digital Audio Coding Formats” compared 2.8 MHz DSD with 24/176.4 PCM in a blind listening test. This test used musical performance recordings made in both formats using microphones with extended high-frequency response rated to 40 or 50 kHz. Both formats were fed with analog signals directly from the microphones, with no mixing. The tests were conducted with 110 listeners. Of 2,900 comparisons, there were 1,454 correct choices and 1,446 incorrect ones – about the same results as flipping a coin. The authors noted, “These people, for the most part, were well accustomed to critical listening on a professional level, but they found that they could not even begin to recognize any sonic differences....”

Furthermore, while DSD recorders are now available even at prices as low as US\$500, all but the simplest DSD recordings must be converted to a PCM format for editing, then back to DSD for distribution. This conversion eliminates any of DSD’s theoretical or purported advantages.

Conclusion

While many music labels, audio equipment manufacturers and consumers have touted the benefits of high-resolution audio, there is as yet no significant scientific evidence that it is beneficial for use in consumer audio devices. There is, however, some evidence that high-resolution audio may in some circumstances result in reduced fidelity compared with standard-resolution audio.

Expanding bit depth from 16 to 24 bits does no harm in consumer applications, but it wastes storage and transmission space without delivering any real benefit. Increasing sample rate from 44.1 to 96 kHz or higher also delivers no real benefit, and can actually reduce fidelity.

Of course, there many audio enthusiasts and professionals dispute these contentions, but we know of no scientific evidence that supports their views.

Written by the Goldmund Acoustic Laboratory in collaboration with industry expert Mr. Brent Butterworth.