HD Audio Delivery and Playback in TWS Earphones

White Paper

WP07

The widespread availability of High-Definition (HD) streaming music is spurring demand for HD-capable TWS earphones. This paper describes the technology behind HD music delivery and how audio designers can meet the growing need.

Multiple sources[1, 2] point to increasing demand for HD audio capable wireless earphones. These same sources show hearing personalization growing in popularity, reflecting the trend that people of all ages want the full HD experience even if they have typical age-related hearing degradation. These trends are driving the development of HD audio support at all stages of the music delivery chain.

Evolution of HD audio

The term HD audio (or high-res audio) does not have a strict technical definition. It is generally used to describe audio systems that support higher data rates than were available in the early years of such equipment's adoption. The term was first used to describe digital audio systems that could support higher data rates than the Compact Disk (CD) standard. This included alternative disk formats, and later files containing digital audio recordings. It has also been applied to audio streaming, and more recently to wireless headphones capable of delivering better than typical audio quality.

Improvements in recording and distribution of audio have increased the data rates available to mobile listeners. Wireless headphones lagged behind due to Bluetooth limitations. However, newer Bluetooth codecs make it possible for wireless headphones to deliver HD audio. This places new demands on the audio hardware including the drivers. This paper will explain technical terms associated with HD audio, describe the improvements in the audio chain from source to headphones, and how using balanced armature (BA) tweeters together with dynamic woofers delivers HD audio-quality sound in combination with ANC, occlusion reduction, and hearing personalization.

Technical background

Digital audio formats are most simply described by their sample rate and bit depth. Digital audio represents analog sounds by sampling the signal amplitude rapidly in time. The amplitude of each sample is measured and saved as a binary number. The number of samples per second is the sample rate. The size of the binary number used to describe the amplitude is the bit depth.

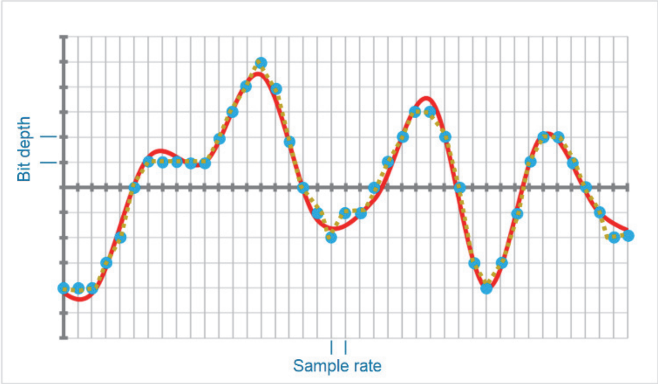
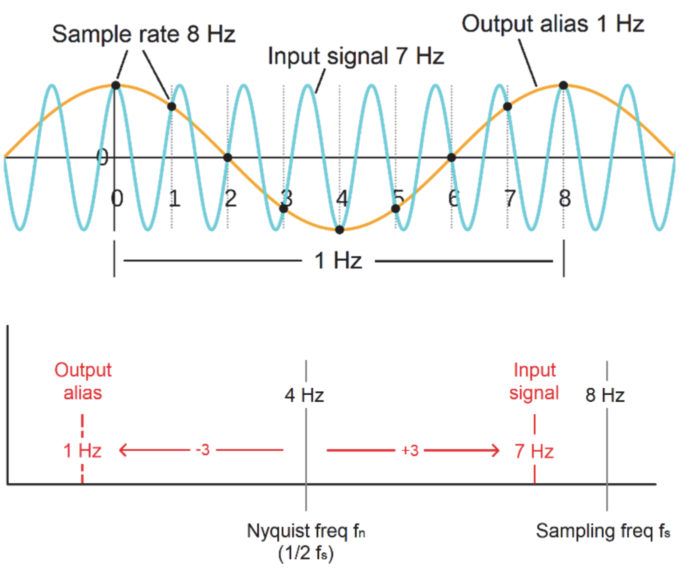


Figure 1. Sample rate and bit depth in digital audio.

Sample rate

Information theory credited to Nyquist (and also Shannon and Whittaker) states that we must sample a sine wave at least two times per cycle to accurately capture its information. Otherwise the reconstructed output will be an alias at the wrong frequency as illustrated by figures 2a[3] and 2b. The alias is a mirror image in frequency domain with the "Nyquist frequency" (1/2 the sampling rate) as the axis. For music, the sample rate must be at least 2x the highest frequency to be reproduced. Thus a 40 kHz sampling rate is required If the maximum audio frequency to be represented is 20 kHz. This requires a very sharp 20 kHz low-pass filter at the ADC. Any audio above 20 kHz which passes through the filter will be recreated as an audible alias when converted back to analog, reducing the fidelity of the music. Because no filter is perfect, in real world practice some margin is provided between the highest frequency to be represented and the sampling limit. In the compact disk format, the audio to be sampled is low-pass filtered to 20 kHz. The sample rate is set to 44.1 kHz, somewhat above the theoretical limit of 40 kHz, to provide the filter margin.



Figures 2a and b. Alias in time and frequency domains.

Bit depth

The size of the word used to describe a sample, or bit depth, determines how accurately each sample is digitized. Imagine describing how tall a person is by rounding to one digit. You could then say someone is two or three meters tall. A second digit would be better, allowing someone to be described as being 2.1 or 2.2 meters tall. This is still a bit crude, but one can keep adding digits until sufficient resolution is provided.

The same is true in digital audio. Each time an additional bit is used in the binary word, the amplitude can be described with twice the number of values. This in turn reduces the errors by a factor of two. While one would intuitively expect rounding the values to lead to distortion, it turns out that by applying a tiny bit of noise called dither, the errors (quantization errors) can be converted from distortion into noise. Therefore a good recording system has no distortion and a noise floor that goes down 6 dB with each added bit of depth. The ratio between the loudest undistorted sound and the noise floor is 6 * bit depth.

Since digital systems are built on multiples of 8 bits, digital audio uses multiples of 8 bits for the word size. Very early computer audio used just 8 bits. Having the noise floor only 48 dB below the loudest music was not very practical. The compact disk supports a bit depth of 16, providing a signal to noise ratio of 96 dB. This covers a very useful range. If the playback volume is set to a reasonable level, the remaining hiss between songs is lower than the background sounds in the listening environment. With elevated volumes there may be some audible hiss between songs or in quiet passages. Therefore an argument can be made for using higher bit depths to reduce that component of the total system noise.

Figure 3 shows the effects of varying the sample rate and bit depth. Increasing the sample rate increases the range of frequencies captured in the data. Increasing the bit depth increases the dynamic range. Dynamic range is the spread between the loudest and softest sounds.

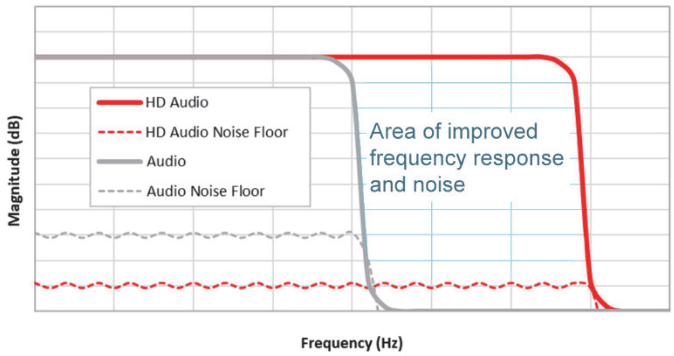


Figure 3. Effect of increasing sample rate and bit depth.

Compression

The combined bit rate of the compact disk is 16 bits/sample * 44,100 samples/second * 2 channels = 1,411 thousand bits per second (kbps). This data rate was too large to be practical with early digital audio players. It took too long to download a song and available memory was limited. Compression methods were developed to solve this problem.

Pattern recognition can be used to provide more concise descriptions of the patterns of ones and zeros in the data stream. This is the idea behind compressing data files in a computer. Some files can be compressed to less than 1/10 the original size, then later restored without loss.

Music files cannot be compressed by much with these methods. By using a linear prediction algorithm, some improvement can be gained. However, compression ratios still rarely exceed 2:1. Since there is no data lost, these are known as lossless encoders. Two popular examples are FLAC and Apple Lossless compression.

To achieve further reduction in the data rate, one can use psychoacoustic modeling of human hearing to minimize the audibility of the discarded data. It is quite common for one sound to hide another from human perception. (See figure 4.)[4] The overall data rate is reduced by encoding those hidden sounds at a lower rate or by discarding them altogether. While psychoacoustic compression methods have improved, even the best designed encoders will still cause some audible artifacts[5]. Lower compression rates offer fewer and milder audible artifacts at the expense of larger files and higher data rates. This is the science behind the movement toward HD audio streaming, made possible by the ability to transfer music at higher speeds.

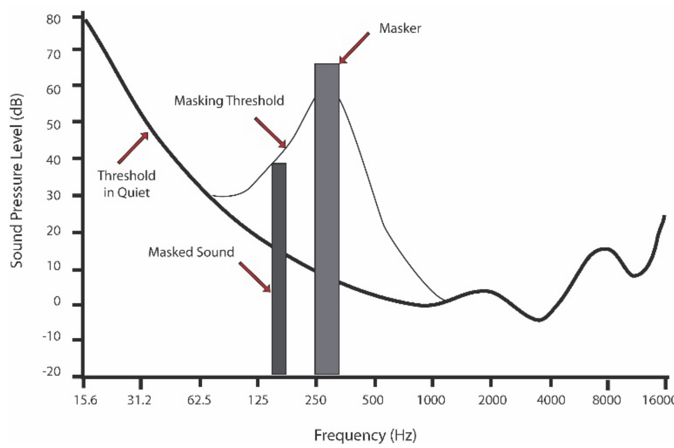


Figure 4. Loud sound masking perception of quieter sounds with similar frequencies.

Extended Bandwidth

The terms “HD audio” and “High resolution audio” are often used synonymously, but this can lead to confusion. “HD audio” is used to describe any audio with a higher data rate than used traditionally. The Japan Audio Society (JAS) licenses the use of the “Hi-Res Audio” logo for hardware that meets specific requirements, including the ability to reproduce audio up to 40 kHz. (The exact requirements can be obtained under license from the JAS).

HD audio is the result of any combination of increasing the sample rate, increasing the bit depth, and decreasing the compression rate. If the sample rate is increased beyond 44.1 kHz, the digitized audio will represent a closer approximation of the original signal which in turn helps maintain fidelity throughout the delivery chain. It also becomes possible to expand the bandwidth beyond 20 kHz. Not all HD audio formats and equipment have bandwidth beyond 20 kHz but some do, especially those targeting JAS certification.

The music delivery chain

To get HD audio to the listener's ears, every step in the pipeline must be of sufficient quality. These include the original song preparation process, the download or streaming service, the playback system, the connection to the headphones, and the headphone drivers.

Recording and delivery

Music recording equipment is widely available that allows storage with a variety of sample rates and bit depths, many of them exceeding the Compact Disk specification. Once the music is recorded it must be delivered to the user. File download services have supported HD audio for some time now. More recently, most popular streaming services have either announced or are already offering lossless CD quality as a baseline and a growing subset of their libraries with high bitrate HD, often without extra charge.

Music playing hardware

Today one may listen to downloaded or streamed audio using a variety of devices including dedicated music players, mobile phones, PCs, and more. HD audio playback is widely available in these devices. The additional cost to support HD audio is relatively low, with sample rates up to 192 kHz and word lengths of up to 32 bits routinely supported. In some cases the analog headphone output of these devices (if one exists at all) is limited to 20 kHz, and often does not have noise performance exceeding 16 bits. However, in these cases one can plug an external DAC into the digital port to get higher performance.

Headphone codec

The last step in the chain is the headphone. In a wireless headphone, a key limiting factor is the Bluetooth radio link. A Bluetooth codec (COder/DECoder) reduces the data rate prior to transmission to the headphone. To meet increasing demand for premium audio in wireless headphones and TWS, codecs with increasing data rates have been introduced. (Table 1.) In a properly designed headphone, these newer codecs can deliver a higher fidelity listening experience even if not completely lossless. Going further, Qualcomm recently announced their intent to support lossless transmission of CD quality audio via Bluetooth.[6]

Codec	Sponsor	Max Bit Depth	Sample Rate/ (Bandwidth kHz)	Max Bitrate (kbps)
COMPACT DISC DIGITAL AUDIO	Standard	16	44.1 / (21)	1411 (lossless)
LDAC	Sony	32	96 / (46)	990
LHDC	Savitech	24	96 / (46)	900
Qualcomm snapdragon SOUND	Qualcomm	24	96 / (46)	860
Qualcomm aptX Adaptive	Qualcomm	24	48 / (23)	420
SBC	Standard	16	48 / (23)	345
AAC	Standard	24	44.1 / (17)	256 (Apple) 320 (Android)

Table 1. Common Bluetooth codecs compared to CD audio.

Several of the new HD Bluetooth codecs support extended bandwidth. The headphone can then reproduce frequencies above 20 kHz. Figure 5 shows the response curve measurement of a commercially available TWS earphone. Significant output is visible up to 40 kHz.

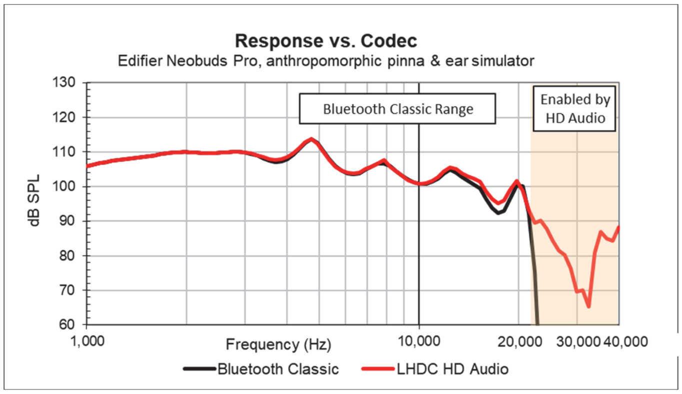


Figure 5. Comparison of bandwidth of Edifier Neobuds Pro TWS with SBC and hi-res codec.

HD TWS earphone audio design

Customers purchasing HD TWS earphones will have very high expectations for audio. Acoustic design can be especially challenging in TWS earphones due to the overall feature set and limited space available for the driver system. TWS earphones must fit comfortably within the ear and increasingly must offer ANC and other advanced features. This creates unique challenges for the acoustic designer.

The interaction between the needs of ANC and HD audio are particularly important. To deliver effective ANC in loud environments, the drivers must support high bass output with low distortion. Leaky or semi-open designs to reduce occlusion place further demands on the bass output. At the same time, HD audio playback requires extended treble output to 20kHz or even more, especially if seeking JAS Hi-Res Audio Wireless certification. Meeting both requirements simultaneously with a single dynamic speaker becomes increasingly difficult as speaker size is reduced. Yet modern TWS designs target small, comfortable form factors.

To deliver on these seemingly conflicting needs, many earphones use a separate dynamic woofer and BA tweeter in place of a single full range driver. This hybrid configuration provides a smoother and more extended high frequency response along with less need for electronic equalization that can increase power consumption and reduce headroom. The woofer design is focused on providing strong bass for music, ANC, and occlusion reduction, while the BA tweeter is optimized for reproducing clear and distinct treble to support HD playback.

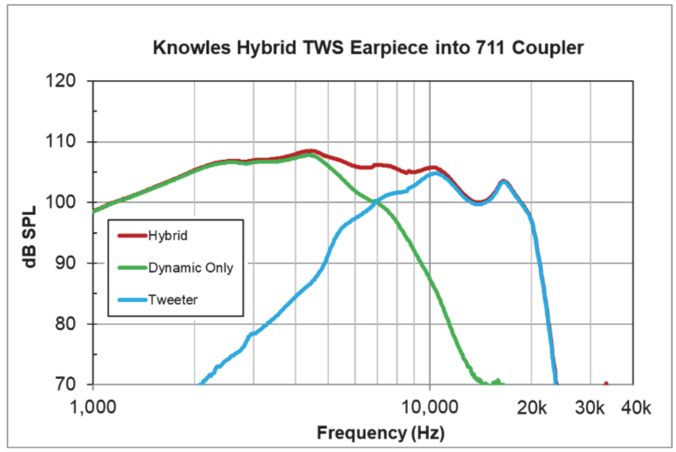


Figure 6. Frequency response of Knowles ref. design hybrid showing woofer, tweeter, and combined response.

BA tweeters also support driving high treble gain with low power consumption, making them ideal for earphones offering hearing personalization. Their sealed, closed-back design limits the possibility of acoustic feedback to the microphones when personalization is enabled.

Such a design also offers more freedom in the arrangement of drivers. It enables moving the woofer to a location or angle less directly aligned with the ear tip while still keeping the BA tweeter near the opening to minimize the inertia of the air trapped between the tweeter and ear tip. In this way maximum treble extension is achieved with more flexibility in the overall earphone shape, improving user comfort without loss of high frequency performance.

Using a BA tweeter provides several tools for adjusting the high frequency response. The acoustic features near the tweeter opening can be shaped to further refine the high frequency output. The crossover can be adjusted for smooth blending of the woofer and tweeter signals. The sensitivity of the tweeter can be adjusted by selecting a higher or lower coil impedance to get a better match to the woofer, and final shaping can be accomplished by DSP tuning. The high, extended output of the BA tweeter minimizes the need to add treble gain just to produce the desired baseline response.

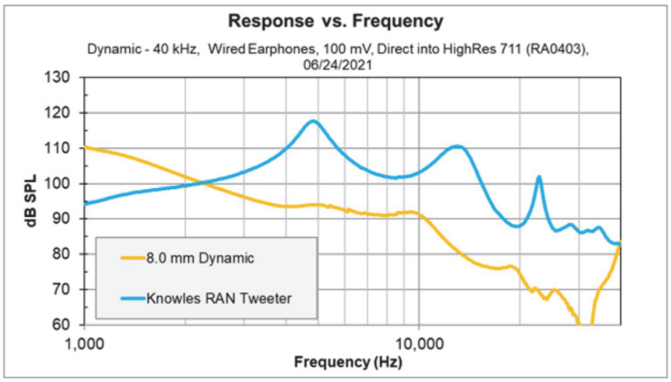


Figure 7. Treble response of Knowles RAN tweeter compared to typical 8mm dynamic speaker. The additional headroom provided by the tweeter reduces need for adding treble gain, especially when personalizing audio.

Since many Bluetooth SoCs have two outputs, woofer and tweeter can be driven by individual amplifiers for even more flexibility in shaping the response. The Bowers & Wilkins PI7[7] and Edifier Neobuds Pro[8] are two examples of TWS earphones using separate amplifiers and active crossovers.

One can also use the tweeter to provide response extending well above 20 kHz. Figure 7 compares the treble output of the Knowles BA tweeter model RAN to a typical 8mm dynamic speaker. The BA tweeter provides the increased treble output and extension necessary for HD audio, including the capability to support hearing personalization or enhancement.

Conclusion

Consumers are increasingly demanding HD audio quality in TWS earbuds. Delivering HD to the listener requires upgrading all of the stages in the delivery chain. Music streaming services, mobile phones, and Bluetooth codecs are in position to provide higher fidelity than ever before. Now is the time for designers of TWS earphones to take full advantage of these changes to meet consumer demand. To do so, one must also pay close attention to the driver system. A hybrid driver with dynamic woofer and BA tweeter is the optimum solution for HD audio TWS earphones.

1 Qualcomm, "2021 State of Sound Report," 5, 8, 10, 12, 18
2 Sonova "Virtual Investor & Analyst Day 2021 Presentation," 56
3 Adapted from Andrew Jarvis, CC BY-SA 4.0, via Wikimedia Commons
4 Daxx4434, Public domain, via Wikimedia Commons
5 Karlheinz Brandenburg, "MP3 and AAC Explained," *Fraunhofer Institute for Integrated Circuits FhG-IIS A*, 6-7
6 Qualcomm, "Qualcomm adds Bluetooth Lossless Audio Technology to Snapdragon Sound," *press release*, 1 Sept. 2021
7 Bowers & Wilkins website, PI7 page
8 Edifier website, Neobuds Pro page

More information on Knowles balanced armature drivers for hearables & music earphones can be found at **www.KnowlesPremiumSound.com**

ABOUT THE AUTHOR:
Thomas Miller is an Engineering Fellow with Knowles Corp's Hearing Health Tech division. He can be reached at +1 630 285 5882 or Tom.Miller@knowles.com.