OPEN ACCESS Freely available online
V. R. Melchior, "High Resolution Audio: A History and Perspective"
J. Audio Eng. Soc., vol. 67, no. 5, pp. 246–257, (2019 May.).
DOI: https://doi.org/10.17743/jaes.2018.0056

High Resolution Audio: A History and Perspective

VICKI R. MELCHIOR, *AES Member*
(vmelchior@earthlink.net)

Consultant, Boston, MA, U.S.A.

Exploration of audio at digital resolutions higher than CD began in the late 1980s, arising from a wealth of interdependent sources including listening experiences, rapid technical advances, an appreciation of psychoacoustics, acoustic measurement, and an ethos that music recording should capture the full range of audible sound. High resolution formats developed and were incorporated into successive generations of distribution media from DVD, SACD, and Blu-ray to internet downloads and now to streaming. A continuing debate throughout has been whether, and especially why, higher resolutions should be audibly superior. This review covers the history of the period, focusing on the main directions that drove experimentation and development up to the present, and then seeks to explain the current view that, beyond dynamic range, the most likely technical sources differentiating the sound of digital formats are the filtering chains that are ubiquitous in traditional digital sampling and reconstruction of analog music sources.

0 INTRODUCTION

High resolution has had a 30-year history, arising in the late 1980s from a wealth of sources including listening experiences, rapid technical advances, an appreciation of psychoacoustics, acoustical measurements, and the ethos of music production. Although well established in both professional and high performance consumer audio, these formats nevertheless continue to generate controversy over what high resolution means, whether it is audible, and why it should be so. It is often assumed that the main parameters defining high resolution formats, the extension of sample rates and bit depth beyond those of the CD, imply the audibility of frequencies and dynamic range beyond the accepted limits of human hearing.

But formats cannot invariably guarantee greater transparency, and thus higher resolutions have always challenged our understanding of the contributions of signal processing and system design to the best achievable sonic performance and ultimately to the way the ear and brain receive and interpret sound.

This brief review considers the development of primary ideas about these formats beginning in the 1980s until the present and describes the coalescence of ideas, especially in the last ten years, into a modern view of what high resolution is most likely about. The context is that of format and media development, while the many concurrent developments in technology, signal processing, and studio practice can just be touched on given their wide scope. The discussion focuses on linear pulse code modulation since this remains overwhelmingly the dominant format although single bit methods (pulse density modulation) are considered in context.

1 EARLY DIGITAL

It is useful to consider this and later eras in terms of the themes that drove development: technology of course, and a growing comprehension of the role of psychoacoustics, but more specifically the philosophy of what recording should be and beliefs about dynamic range and frequency range.

Audio in the 1970s was based in analog technology, although digital methods for storage, recording, and transmission were developing. Experimentation in the early '70s with optical video recording coupled with work in digital broadcast and television [1–3] led to PCM audio recorders and then to the first music releases (as analog LPs) recorded entirely with digital recording systems, including those of Denon, Soundstream, 3M, and Decca-UK [4–6]. Recognition of the advantages of digital over analog led Sony and Philips to launch the Red Book standard in 1980 and the CD in 1982, and it was the period just after this that saw the real and rapid development of digital. Many of the techniques of today were developed during the 1980s including dither, noise shaping, rapid progression toward digital audio architectures including fixed and floating point, improvement to network architecture and bandwidths, improved converter linearity and bit capacity, and lossy compression codecs [7–13].

1.1 Dynamic Range

Dynamic range has always been easier to address than the extension of frequency, although in its digital form, it has received less formal audibility testing. Assessment of the dynamic ranges of music, rooms, and ears occurred at least as early as the 1940s including the work of Fletcher [14] and others. The CD standard of 1980, with a defined 16 bit (98 dB) capability, reflected the technology limits of the day. But it focused attention on the ethos of recording and reproduction of music, namely that systems should strive to reflect the dynamic range defined by the music itself and by the limits of human auditory capability. This is reflected in a well-known series of papers by Fielder [15–17] framing the issue as a ratio of peak instantaneous SPL measurable from live music at a favorable audience location to the just audible threshold for narrowband white noise added to program material and heard in quiet listening environments, studio as well as home. This noise value is meant to test the minimum audibility of recording equipment noise inherent in recordings. Fielder's ratios were 90–118 dB in 1982 and $\sim$122 dB (mono, stereo) in 1994 and provided an ideal to strive for. In digital terms, this implies a linear resolution of about 20–21 bits.

Neither in 1982 nor in 1994 did recording or reproduction equipment meet these dynamic range goals. But an interesting snapshot of technology limits by 1994 is provided in [17], indicating for example that although D/A's had achieved 20 bit input word lengths, converters generally were limited by non-linearity and noise to achieved ranges of around 106–110 dB (A/D) and 112 dB (D/A). Professional recording standards kept pace, with recording by 1994 generally done at 20 bits. The AES-EBU professional interface standard defined channel widths of 16–24 bits [18–20]. Additionally, more refined psychoacoustic models were used to evaluate bit depth by assessing the audibility of quantization noise and dither. Quantized and dithered signals have noise spectral densities (NSD) that can be modeled psychoacoustically as the equivalent noise level in the auditory filters, allowing a direct comparison of noise to the single tone hearing threshold. In [21], this method is used to compare NSD's from high acoustic gain signals of 120 dB SPL that had been dithered with appropriate triangular PDF dither and quantized to 16, 18, or 20 bits. The author argues that 19–20 bits are needed to render the noise floor inaudible, that is, below the single tone threshold, at all spectral frequencies.

1.2 Frequency Range

Frequency range has been the more contentious parameter leading, especially early on, to lengthy debates on why and whether frequencies above the accepted human audibility limit should be necessary in audio.

At the CD launch in 1982, beliefs of the time were that both frequency range and dynamic range should be defined by human auditory limits. The question of how to set digital standards for transmission and broadcast led to listening tests which, given the limits of equipment and testing at the time, concluded that very few people could detect differences in program material with content above 15 kHz [22–24]. This was despite recognition of the 20 kHz audibility limit measured for pure tones in [25], and that some engineers could detect the difference with cutoffs set at 18 kHz or 20 kHz [23]. Accordingly, early (1978) sampling rates were set at 32 kHz for broadcast (Nyquist at 16 kHz), with 32 kHz and 40–48 kHz adequate for consumer and pro audio respectively [26]. Among the psychoacoustic arguments offered by Plenge et al. [22] for 15 kHz were that the 5 kHz step from 15 kHz to 20 kHz only represents 1/3 octave out of 31 third-octaves, and that pitch and level differences aren't well developed in that range anyway!

Even in 1991, audio in common use usually extended no higher than 20 kHz with sampling rates of 32–44.1 kHz (consumer) or 48 kHz (pro); nevertheless professional gear including state of the art workstations and synthesizers were available at sampling rates of about 100 kHz [27] and DAT recorders at 96 kHz [27, 28]. While the higher bandwidth designs were sometimes manufactured for technical reasons including reduced distortion and better handling of aliases and intermodulation, Oohashi et al. [27] report that many recording engineers disagreed with the 20 kHz limit, believing instead that their wide bandwidth equipment distinctly affected perception. Oohashi points to problems in listening test methodology that might have biased the listening tests of the previous decade. Within the 1990s, recording formats of 88.2 kHz/24b and 96 kHz/24b were in active use, although not pervasively so [29]. The need for higher frequencies remained an open debate. The published literature of the era includes many reports saying that wider bandwidths sound better, at least some of which included informal ABX tests [27, 30–32], but the technical reasons why were debated, especially considering that transmission and storage bandwidths remained expensive [30, 32, 33]. Some authors argued that ultrasonic energy may somehow influence the perception of sound within the audible range [34, 35].

2 SURMOUNTING THE CD

With professional standards at 20 bits and the advantages of higher bit rates becoming apparent, the 16-bit limit of the well-entrenched CD became a barrier. A variety of processing techniques were studied with the goal of packing more of what could be recorded at 20 bits into 16. Noise-shaped dither, an error feedback technique allowing the spectral power of dither and quantization noise to be shifted upwards toward less acoustically sensitive frequencies, permitted a signal with (nearly) 20 bit dynamic range to be conveyed in 16 bits, and it could further be combined with emphasis, see for example [10, 36]. This was employed in commercial applications of the period including Sony's Super Bit Mapping [37]. The Pacific Microsonics HDCD codec [32] employed a buried data channel, as first developed by Gerzon and Craven [38], to expand 16 bits to about 19 on playback by re-expanding previously limited peaks and compressed low level signals [32]. Other approaches such as high frequency dither [39], that adds less noise than standard triangular PDF dither, were used in Apogee's UV22 process [40] and in HDCD.

The use of techniques such as noise shaping to extend dynamic range placed greater demands on converters however, since both A/D and D/A converters now needed to achieve 20 bit accuracy, including differential linearity, internal word size at least 20 bit, and noise floors (e.g., due to quantization) no higher than the 20th bit in order to preserve the advantage of noise shaping [36]. In fact the entire signal chain had to do so. Thus a range of techniques, both linear and non-linear, were developed to enhance converter performance, e.g., subtractive dither for linearization and attenuation of artifacts, and summation of several identical conversion channels to achieve non-coherent gains of 3 dB per summed pair [30, 41].

Likewise regarding sample rates, amidst the questioning of "why" higher rates should sound better, there was a growing engineering perception that the digital filtering chains used to downsample and upsample music in the transit to and from analog had an influence on sound. Specifically, it was widely thought that the very sharp anti-alias filters required at 44.1 kHz sampling by classical Shannon sampling theory had an audible negative effect. Also that the inclusion of a loop of downsampling (to 44.1 kHz/24b) and upsampling into an otherwise 88.2 kHz/24b chain produced an audible effect under careful test conditions [32]. Further, that relaxing the filter transition band from knife-edge to a more gradual roll-off may reduce sonic degradation [42]. An anti-alias filter that is flat to 20 kHz and rolls off gradually to the Nyquist frequency has a shorter time domain impulse response with less pre- and post-ring compared to a filter with a steep cut-off (Sec. 4.4) and is easier to design at higher sample rates than 44.1 kHz. The relationship of less ringing to reportedly better sound was noted in AES conference presentations of the 1990s [43] and filters of this type were designed into products for studio professionals [29, 42].

3 THE HIGH RESOLUTION ERA

DVD, the follow-on higher density optical disc after the CD, was marketed by Sony and Philips in 1995, and the DVD Forum was founded in 1997 [44]. From that flowed the modern array of DVD video and audio formats that were used firstly for video with multichannel sound and then for audio-only distribution incorporating high resolution formats. Appendix 1 summarizes the properties and formats for these including the later blue laser higher density discs, HD-DVD (2006) and Blu-ray from the Blu-ray Alliance (2006). These discs made possible multichannel audio of greater bit depth and much higher sample rates as well as video, metadata, and added features. But for the first time they also opened the significant question of how to use the available audio space efficiently.

That high resolution formats were incorporated into the higher density discs was the result of a steadily building demand during the 1990s. Besides listening perceptions, the reasons were the familiar ones of continuing technological advance of performance targets as the limits and distortions of hardware and signal processing were studied, an increased appreciation of the role of psychoacoustics, and the belief that recording and delivery should encompass the inherent range possible in music listening. In consequence, the earlier focus on accommodating to the CD gave way to the belief that music data recorded and processed correctly at higher sampling rates and 24 bits avoided the compromises of working around the 44.1 kHz/16 bit limit [17, 32, 45–47], and that the better sound available in the studio should be accessible to listeners.

The requirement for efficiency in accommodating video, metadata, and multichannel audio in a variety of formats led directly to a consideration of how to use the potential audio space on discs. An exploratory paper by Stuart in 2004 [47] attempts to restrict the over-specification of audio space by modeling the information capability of coding spaces against the expected noise and frequency limits of the hearing system, and the noise and frequency limits of recordings, concluding that a range of signal processing techniques might contain the bit growth of music releases.

3.1 Media, Formats, and the Expansion of High Resolution Data Rates

Higher capacity optical discs and their audio formats are listed in Appendix 1. Discs that focused on video distribution were very successful at least through around 2012 but, like all discs, are currently declining in favor of online distribution. Most soundtracks on video discs are compressed and lossy. The mandatory PCM tracks generally are 48 kHz since this has long been a recording standard for audio in cinema. The high resolution tracks that are available are varied in quality, notably on releases of older material that often simply upsampled the original 48 kHz data. Audio-only discs (e.g., DAD, DVD-A, SACD, Pure Audio Blu-ray, High Fidelity Pure Audio (Blu-ray)) have had quite limited commercial success that varies by region and with SACD the more successful over DVD-A and Blu-ray [48, 49]. Because of the high capacity and data rates of the Blu-ray disc, Blu-ray audio-only formats are especially well suited to multichannel audio at high sampling rates [50], but their uptake has been limited. This is due partly to lack of interest in the broad market and partly to the very rapid evolution of online digital distribution beginning in the late 1990s.

The Sony-Philips SACD format (1998) and discs based on the single-bit stream technology, DSD, are extensively characterized in the earlier literature [51–55]. DSD was conceived initially as a storage and distribution format for archived master tapes where the tapes needed little editing and where the DSD single-bit stream could be captured from an A/D sigma delta modulator, released on SACD, and then routed back to the D/A's sigma delta modulator without need for filtering [56]. But with SACD aspiring to replace the CD format, editing became essential, and the mathematical difficulties of doing anything but the simplest computations on a single-bit stream, obvious. This was addressed by converting to a multibit stream, termed

DSD wide (8 bit PCM at 2.8224 MHz), for processing while maintaining single-bit streams elsewhere [53]. More recently one of the early participants in SACD development, Merging Technologies, has developed a workstation that automates processing of DSD and SACD in a turnkey manner. Editing and signal processing are enabled using an intermediate PCM format, DXD, at 352.8 kHz/24b. This approach significantly simplifies handling of the formats in the studio.

The distribution of audio over the internet that began in the ‘90s relied on downloadable lossy compression formats that were suited to portable players and computers. The low bit rates of these formats enabled streaming at a time when transmission bandwidth as well as storage were limited and expensive. Introduction of the Apple iPod in late 2001 contributed, within a few years, to exponential growth in this sector and a major shift in music distribution [57, 48]. The MPEG-1 layer 3 (MP3) lossy codec and its successor, Advanced Audio Codec (AAC), together with later codecs [58, 59] continue to be successful in the broad market but have often been judged deficient by audiophiles when played on high quality audio systems.

Nevertheless, the advantages of downloading were clear and, as bandwidth and storage became more affordable in the 2000s, a new business model for high resolution distribution independent of the confines of discs, of the always contentious watermarking and digital rights management on discs, and of disc format specifications appealed broadly as a brave new world. Files at sample rates up to 192 kHz/24b began appearing about 2005. By 2012 they were widely available on websites from large aggregators down to small audiophile labels, orchestras, and artists. Downloads continue to be successful with audiophiles through 2019 despite a rapid switch to streaming in the broader market of lossy compressed music [60–62]. The formats available currently are shown in Appendix 2. Besides PCM, DSD was re-enabled as an encoding format for streaming. Support technologies have developed around it, for example a protocol to transfer DSD over USB by packing it into PCM frames [63]. Also newly appearing as release formats are several that had been exclusive to mastering: Merging Technologies’ DXD, double speed DSD (128 Fs) and quadruple speed DSD (256 Fs) [64]. Importantly, an enabling technology throughout has been the concurrent development of computer and server based audio that, in its diverse implementations, now frequently replaces disc players.

In several respects, consumer involvement has stimulated quality in online high resolution releases. One such impetus has been for the labeling of the provenance of recordings, as many releases labeled high resolution in earlier years were upsampled from CD back catalog and would not have benefitted from the signal processing advantages of high resolution (Sec. 4.4). A second is for stringent use of the highest quality source available in the remastering of older works.

File size growth is another aspect, possibly less justified, of the freedom from disc formats afforded by the internet. There has always been experimentation by designers with ever higher sampling and bit rates, and proposals for release formats have included up to 768 kHz LPCM, bit depths to 32, and octuple speed DSD (512 Fs). These are not based in general on engineering and psychoacoustic principles but instead on reports of incremental sonic improvement with each doubling of the sample rate. High data rates can make sense for internal engineering paths but the massive data rates involved, even with lossless compression, are problematic for downloading and storage, and entirely untenable for streaming.

A nascent migration of high resolution to streaming is at this time occurring. A few services worldwide stream losslessly encoded LPCM files up to 192 kHz/24b and DSD up to 256 Fs but the future for large file streaming, especially beyond stereo, is uncertain. A new proprietary coding technology, Master Quality Authenticated (MQA) addresses file size by hierarchically encoding bitrates from CD to 384 kHz in a single file, similar in size to CD alone. The encoded file can be streamed, downloaded, or recorded and is played back at CD or higher resolutions depending on the absence or presence of a proprietary decoder. The singular aspect of MQA is that it seeks to limit the overextension of high resolution bit rates by restricting the processed frequency and dynamic range space to a region defined by the spectra and noise levels of music data and hearing. General principles including the filtering and sampling approach are described in [65]; also see [66]. This approach is supported by many in the music industry and is currently available.

4 HIGH RESOLUTION SOUND, “WHY?”

Why higher resolutions should sound more transparent than CD has been debated for the past 30 years. High resolution formatting does not guarantee a perception of transparency, but music professionals with access to first generation data have widely reported subjectively better sound, and a meta-analysis of previously published listening tests comparing high resolution to CD found a clear, though small, audible difference that significantly increased when the listening tests included standard training [67]. This section considers four proposals for sonic differences, the third and especially fourth of which are the ones broadly accepted as likely.

4.1 Ultrasonic Frequencies

A frequent misconception is that high data sampling rates assume the audibility of frequencies above 20 kHz. Scientific study of ultrasonic frequencies continues, for example their role in bone conductive pathways and in the effects of airborne sound on brain waves measured by EEG. However, a role in normal audio listening has been rejected since about the late 1990s (due to lack of evidence) in favor of the ideas discussed in Secs. 4.4 and 4.3. Individuals able to hear above 20 kHz might experience subjective differences compared to those who don’t, however an ability to differentiate high resolution and CD data is reported, informally, by individuals whose measured limits are well below 20 kHz.

The 20 kHz limits of human auditory perception for pure tones transmitted by an airborne path were established from work on equal loudness contours [25]. The physiological basis for these limits, including the role of the outer and middle ear and cochlear processing, are summarized in [47, 68, 69]. Studies of bone-conducted ultrasound and EEG measurements are discussed in an audio context in [47, 65]. There also have been formal listening tests specifically addressing whether test tones and their harmonics in the ultrasound region are audible. Except in cases with very high amplitude stimuli [70, 71] where thresholds can be measured for some listeners, literature studies have shown negative results [72]. A recent study aimed primarily at the audibility of intermodulation distortion (IMD) also confirmed the non-audibility of the ultrasonic tone pairs used to generate the IMD [73].

4.2 Hardware

The question of whether hardware performance factors, possibly unidentified, as a function of sample rate selectively contribute to greater transparency at higher resolutions cannot be entirely eliminated.

Numerous advances of the last 15 years in the design of hardware and processing improve quality at all resolutions. A few, of many, examples: improvements to the modulators used in data conversion affecting timing jitter, bit depths (for headroom), dither availability, noise shaping and noise floors; improved asynchronous sample rate conversion (which involves separate clocks and conversion of rates that are not integer multiples); and improved digital interfaces and networks that isolate computer noise from sensitive DAC clocks, enabling better workstation monitoring as well as computer-based players. Converters currently list dynamic ranges up to ∼122 dB (A/D) and 126–130 dB (D/A), which can benefit 24b signals.

IMD originating from the interaction of system nonlinearities with tone pairs, where at least one of the pair originates in the ultrasonic or near ultrasonic region, has been questioned as a potential source of audible distortion with wider bandwidth waveforms. Recent studies have not found evidence supporting this idea [72, 73].

In principle, formal listening tests carried out comparatively across different system set-ups might help to resolve the question of hardware contributions [67]. At a minimum, design for auditory testing must be careful to exclude known sources of hardware distortion.

4.3 Dynamic Range

Assuming triangular PDF dither, the greater dynamic range of 24b words is clearly a factor in high resolution sound, and one that benefits from better dynamic range in modern converters and systems. But confirmation by testing in this area has been limited, and a number of considerations are to be noted: Dynamic range defined by bit depth may be less a factor for 24b music when additive noise in the recording determines range, see Figure 8 in [65]. Also, noise shaping can extend the dynamic range of 16b data to that of non noise-shaped 19–20 bit recordings. Lastly, the 144 dB dynamic range of 24 bits exceeds hearing limits and the reproduction capacity of consumer equipment, perhaps leaving room for a debate on the optimization of bit depth.

4.4 Filtering and the Time Domain

The chain of anti-alias and anti-imaging filters implemented in modern converters for sampling and reconstructing analog signals has long been suspected of being audible. That filters with shorter impulse responses, gentler transition bands, and less ring might sound better than steep brickwalls was suggested as early as 1984 [74], although several theories exist on the reasons.

4.4.1 Classic CD Era Filters

The classic Shannon sampling theory used throughout digital audio has been well described in texts and tutorials [75] (but also see [65, 66] for recent innovation) and will not be recounted here. A signal must obey some theoretical constraints, e.g., finite energy and continuity, but with those and correct bandlimiting, the bandlimited signal can, in principle, be perfectly reconstructed. The standard implementation of the theory in current data converters, comprising a gentle analog filter to attenuate RF noise, a modulator to convert analog to digital at 384 kHz/24b, and a series of (typically) three 2X half-band filters to decimate 384 kHz to 48 kHz, followed by the reverse path to reconstruct the analog waveform, is also considered elsewhere [65, 75]. Note that “2X” applied to any rate conversion filter means that the filter either doubles or halves the previous sample rate.

It is the set of three 2X filters, or 8x oversampling filters, both downsampling and upsampling, that are primarily at issue regarding sonic effects (although [65] also considers the decimation filter in the modulator). In both professional workstations and high performance playback, these filters are often custom designs while in virtually all other gear, they are carried out on separate chips of fixed design.

In the CD era, specifications were established for professional anti-alias and anti-imaging filters based in part on filter theory and in part on the audibility preferences of mastering engineers [29, 30, 32, 42]. Specifications are listed in Table 1. Steep brickwall filters are in line with sampling theory but also reflect the short transition space available between the audio band edge at 20 kHz and the Nyquist at 22.05 kHz or 24 kHz.

4.4.2 Filtering for High Resolution

The steep brick wall filters of Table 1 have time domain impulse responses (IR) approximating a sinc function and thus substantial pre- and post-ring extending well beyond the duration of the main peak. This is true for windowed sinc functions as well (windowing being a practical means of tapering the infinite sinc function for realizability). The filtered signal has similar ringing due to the steep bandlimit imposed by the filter. That this lengthened IR, well known from CD processing, appeared to influence sound led to sequential experimentation with a range of filters

Table 1. Typical specifications for mastering quality anti-alias filters at 44.1 kHz/16b.

— Brickwall cutoff close to the Nyquist frequency, approximating the Shannon ideal
— Flat passband with attenuation < 0.1 dB at 20 kHz. (> 0.1 dB deemed audible in mastering)
— Extremely low passband ripple (peak ripple 0.00001 dB for highest mastering quality)
— Stopband attenuation > 120 dB. (High stopband attenuations were adopted to suppress isolated high amplitude analog distortion present prior to sampling, cf., [32])

that reduce the amplitude of the ringing or at least of the pre-ringing:

1. Filters that are flat to 20 kHz but whose transition bands roll-off gradually from 20 kHz to the Nyquist frequency. Certain of these designs also extended the transition band to a point beyond Nyquist, allowing this post-Nyquist aliased region to fold back into the region above 20 kHz but not into the audio band itself. This allowed a more gradual roll-off even at 44.1 kHz that reduced the time extent of the 44.1 kHz filter by nearly half compared to a sharp brickwall [29, 30, 42, 76].
2. Minimum phase filters. Minimum phase low pass filters are causal without pre-ring but in consequence have increased post-ring that may be audible, although masked. Minimum phase filters also have phase errors that can be significant at 44.1 kHz [77, 78].
3. Apodizing filters. Apodizing filters, introduced by Peter Craven in 2003 [77, 79], are rate conversion filters whose transition bands descend to zero by a frequency about 90% of the Nyquist of the lowest sampling rate in the filter chain. The spectrum of the filter chain is a product (cascade) of the spectra of the individual filters present in the chain and is dominated by the steep brickwall of the lowest sample rate filter. Inclusion of an apodizing filter, whose roll-off is much gentler, filters out the steep roll-off as well as the ringing due to the entire chain. Implementation is typically minimum phase and is easiest to do at high sample rates where there is space above 20 kHz for the slow transition band. Apodizing filters are relatively common now in converters and players, although see [77] for caveats.
4. Minimum phase filters with reduced post-ringing. The reduced post ring of these filters is achieved by incorporating an extended gentle roll-off. Examples are available in current proprietary commercial designs.
5. Half-bands and brickwalls at high sample rates. The impulse response of any half-band or brickwall will ring, but the use of an identical filter design at each sample rate, with a transition band defined as a fixed percentage of the Nyquist, will result in progressively less ring with each rate doubling. This is due simply to the doubling of the transition width and halving of the time IR at the doubled sample rate. Designs of this type are typical of upsampling chips [43, 52].
6. Very short filters based on Finite Rate of Innovation (FRI) Sampling. A broader framework of sampling theory has developed in the last decade that includes Shannon theory as a subset. An innovative FRI-based design using a class of B-spline filters has appeared in the last few years that results in antialias filters with very short IR, minimal post-ring, and an ability to remove ringing from earlier filters used during production. This is described in [65] and included references; also see [66]. Adaptations are used in the commercially available MQA system.

Examples of the above are found in numerous current and past audio products. Sonically, they have been thought to improve transparency compared to brickwall filters and in some implementations gain strong press reviews, but published scientific listening tests beyond studio or lab testing have only recently begun [80]. Testing can explore the validity of these ideas and by inference, the role of filtering in sound. Further, listening tests typically involve single filters inserted into a chain, whereas the designs in 3 and 6 above have the ability to influence the entire cascade and thus the overall system response of the digital chain. The latter may be more important than previously recognized.

5 MODERN VIEWS OF TIME AND FILTER AUDIBILITY

The two ideas below, suggesting why filters with shorter IRs and reduced ring might sound more transparent, are each related to the time domain. The description given is based on signal properties but without assumptions about the ear's analytic capabilities.

Stuart and Craven provide a review of relevant auditory science and neuroscience as well as signal properties in [65].

5.1 Audibility of Ringing

It has been known since the classic paper of Lagadec and Stockham [74], based on digital measurements in the 1980s, that pre-echo peaks in the time response of filtered signals, which occur as the result of equal ripple across a filter's passband spectrum, are readily audible. Pre-echo audibility has long been cited in relation to the possible audibility of the ringing of brickwall filters, especially of pre-ring. The relative amplitude and spectral composition of pre-echo peaks versus those of pre-ring have been debated although without clear conclusions or definitive studies [76, 81]. The mechanisms for detecting pre- and post-ring at the ear's basilar membrane have remained a matter of hypothesis.

5.2 Blur

Blur is frequently misunderstood but refers to well understood engineering concepts. Blur implies the disturbance of the natural timing relationships in music signals by processes that smear adjacent areas together, cf., [65].

Music of course has time and frequency fine structure. Analysis has been undertaken in the study of acoustics, audio objects, speech, and similar research fields using methods such as short time Fourier transforms (STFTs), spectrograms, and wavelets that reveal much discernible fine structure, onsets, and local spectra. An interesting review by Woszczyk [82] recounts a range of interesting measurements, including of finely detailed and highly complex coherent structure in the decaying reverb of rooms. At its extreme, the complexity of detail probably exceeds the resolution of both recording equipment and of the ear itself, nevertheless it is clear that blurring due to audio processing and replay has the potential to impede natural timing resolution.

Two aspects of filtering contribute to blur:

Bandlimiting. In the classic Shannon sampling theory, an analog signal is first convolved with a sinc function and then sampled. The convolution has two aspects: bandlimiting of the signal and convolution of the bandlimited signal with the sinc function. It is the bandlimiting of the signal spectrum that creates time domain smear by eliminating higher frequencies of the signal and thus broadening and overlapping adjacent peaks in time. The promise of Shannon sampling is perfect reconstruction of the bandlimited signal, and for filters designed in accordance with theory, the convolution sequence preserves the time relationships of the bandlimited signal. A bandlimited impulse is perfectly reconstructed to an analog sinc function including ringing. A corollary is that if the signal is inherently bandlimited, e.g., by a spectral roll-off at 1/f (Figs. 5, 7 in [65]), its timing relationships will not be affected by ideal, spectrally flat filters whose cut-off frequency exceeds the signal's limit, however high frequency noise will be bandlimited[1].

Convolution with the impulse responses of analog components. A music signal interacts with analog equipment at both ends of the recording and playback chain. Mathematically this is a cascade of convolutions of the signal with the impulse responses of the various components. The overlap-add nature of these time domain convolutions results in a spreading and overlap of the detail in the signal, thus blurring its time relations. Of special significance is that when the sampling and reconstruction filters introduce time dispersion, for example ringing in the analog output, then the spreading due to convolution with the impulse responses of the other system components will be exacerbated [65].

[1] But note that the 1/f curve in Fig. 7 [65] is a plot of measured peak values in music spectra measured over long periods, whereas it is the instantaneous shape of the audio spectrum, which need not have the gradual roll-off of the 1/f curve, that determines the local frequency response.

6 CONCLUSIONS

Exploration of higher resolutions began within a few years of the appearance of the CD and have continued unabated in the years since. The drivers have always been an interdependent mix of listening results, technology evolution, psychoacoustic principles to evaluate the significance of noise and bandwidth to the auditory system, and a clear belief that reproduced music should capture the full range of audible sound. The reasons for the greater transparency ascribed to higher sample rates have been debated without consistent evidence but within the past 10 years have come into focus.

Scientific testing also has shown that high resolution formats are distinguishable from CD under formal test conditions, and that the percentage of detections improves significantly with training (i.e., with experience in listening).

Distribution of high resolution music has followed the evolution of media, appearing first in high density optical discs (e.g., DVD, DVD-A, SACD, Blu-ray, Blu-ray Audio), then in internet downloads and finally streaming. The audio-only versions of high density optical discs have not been highly successful despite offering an excellent delivery platform. But an important question that arose with discs is amplified with online file distribution: the need to understand and define the data space used for audio. Proposed distribution formats in recent years can require quite massive coding spaces based on small apparent sonic improvements but without a clear understanding of why, for example, 768 kHz/32b and its associated per-channel data rate of 24.6 Mbits/sec should be needed.

There have been two approaches to containing format size while maintaining high resolution. The first is to use signal processing techniques such as noise shaping, buried data channels, pre-emphasis, and limiting followed by regeneration of high peaks, to avoid extended bit depths and sample rates. The second, used in a current proprietary encoding system (MQA), is to pack data ranging from CD to high resolutions hierarchically into a limited word size while simultaneously limiting the data space to the bit depth and bandwidth defined by music and the auditory system. As modalities like streaming develop, especially with multichannel music, an identification, finally, of what is intrinsically important to sonic transparency is very desirable in order to avoid an unwieldy expansion of data rates.

Current thought regarding the relationship of sonic transparency to high resolution is that ultrasonic frequencies are not involved in normal music listening, that the effects are not due mainly to lower hardware distortion, and that, besides dynamic range, the most likely issues are the anti-alias and anti-imaging filtering chains used to implement classic Shannon sampling. The time behavior of filters is implicated because a broad group of design approaches that reduce time dispersion are reported to improve transparency compared to the sharp brickwalls typical of CDs.

Two outcomes of filter time dispersion are at issue: first, the potential audibility of the extended ringing caused by sharp bandlimiting of the signal, and second, audible blurring of the natural time elements of the signal due both to bandlimiting and to convolution of the music signal with the impulse responses of analog equipment during capture, production, and replay. If pre-ring is audible, its importance to the hearing system may lie in the lack of such non-causal signals in nature.

These filtering ideas are amenable to scientific testing but this has only recently begun.

7 REFERENCES

[1] H. B. Peek, "The Emergence of the Compact Disc," *IEEE Communications Magazine*, vol. 48, pp. 10–17 (2010 Jan.). https://doi.org/10.1109/MCOM.2010.5394021

[2] K. A. Immink, "The Compact Disc Story," *J. Audio Eng. Soc.*, vol. 46, pp. 458–465 (1998 May).

[3] https://www.sony.net/SonyInfo/CorporateInfo/History/SonyHistory/2-08.html

[4] https://en.wikipedia.org/wiki/Soundstream

[5] https://en.wikipedia.org/wiki/Digital_recording

[6] http://www.aes.org/aeshc/pdf/fine_dawn-of-digital.pdf

[7] J. Vanderkooy and S. P. Lipshitz, "Resolution Below the Least Significant Bit in Digital Systems with Dither," *J. Audio Eng. Soc.*, vol. 32, pp. 106–113 (1984 Mar.).

[8] J. Vanderkooy and S. P. Lipshitz, "Dither in Digital Audio," *J. Audio Eng. Soc.*, vol. 35, pp. 966–975 (1987 Dec.).

[9] M. A. Gerzon and P. G. Craven, "Optimal Noise Shaping and Dither of Digital Signals," presented at the *87th Convention of the Audio Engineering Society* (1989 Oct.), convention paper 2822.

[10] S. P. Lipshitz, J. Vanderkooy, and R. A. Wannamaker, "Minimally Audible Noise Shaping," *J. Audio Eng. Soc.*, vol. 39, pp. 836–852 (1991 Nov.).

[11] J. M. Snell, "Professional Real-Time Signal Processor for Synthesis, Sampling, Mixing, and Recording," presented at the *83rd Convention of the Audio Engineering Society* (1987 Oct.), convention paper 2508.

[12] J. M. Halbert and M. A. Shill, "An 18-Bit Digital-to-Analog Converter for High-Performance Digital Audio Applications," *J. Audio Eng. Soc.*, vol. 36, pp. 469–480 (1988 Jun.).

[13] K. Brandenburg, "Evaluation of Quality for Audio Encoding at Low Bit Rates," presented at the *82nd Convention of the Audio Engineering Society* (1987 Mar.), convention paper 2433.

[14] H. Fletcher, "Hearing, the Determining Factor for High Fidelity Transmission," *Proc. IRE*, vol. 30, pp. 266–277 (1942 Jun.). https://doi.org/10.1109/JRPROC.1942.230995

[15] L. Fielder, "Dynamic-Range Requirement for Subjectively Noise-Free Reproduction of Music," *J. Audio Eng. Soc.*, vol. 30, pp. 504–511 (1982 Jul./Aug.).

[16] L. Fielder, "Pre- and Post-Emphasis Techniques as Applied to Recording Systems," *J. Audio Eng. Soc*., vol. 33, pp. 649–658 (1985 Sep.).

[17] L. Fielder, "Dynamic Range Issues in the Modern Digital Audio Environment," presented at the *AES 9th UK Conference: Managing the Bit Budget* (1994 May), conference paper MBB-02.

[18] "Specification of the Digital Audio Interface," Tech 3250-E (2004). https://tech.ebu.ch/docs/tech/tech3250.pdf

[19] R. A. Finger, "AES3-1992: The Revised Two-Channel Digital Audio Interface," *J. Audio Eng. Soc.*, vol. 40, pp. 107–116 (1992 Mar.).

[20] J. Emmett, "Engineering Guidelines: The EBU/AES Digital Audio Interface" (1995 Sep.). https://tech.ebu.ch/docs/other/aes-ebu-eg.pdf

[21] J. R. Stuart, "Auditory Modeling Related to the Bit Budget," presented at the *AES 9th UK Conference*: *Managing the Bit Budget* (1994 May), conference paper MBB-18.

[22] G. H. Plenge, H. Jakubowski, and P. Schoene, "Which Bandwidth Is Necessary for Optimal Sound Transmission?" presented at the *62nd Convention of the Audio Engineering Society* (1979 Mar.), convention paper 1449.

[23] T. Muraoka, M. Iwahara, and Y. Yamada, "Examination of Audio Bandwidth Requirements for Optimum Sound Signal Transmission," *J. Audio Eng. Soc.*, vol. 29, pp. 2–9 (1981 Jan./Feb.).

[24] T. Muraoka, Y. Yamada and M. Yamazaki, "Sampling-Frequency Considerations in Digital Audio," *J. Audio Eng. Soc.*, vol. 26, pp. 252–257 (1978 Apr.).

[25] D. W. Robinson and R. S. Dadson, "A Re-Determination of the Equal Loudness Relations for Pure Tones," *Br. J. Appl. Phys.*, vol. 7, pp. 166–181 (1956). (Acoustics International Organization for Standardization (ISO) 226.) https://doi.org/10.1088/0508-3443/7/5/302

[26] F. A. Bellis and G. W. McNally, Letter to the Editor, *J. Audio Eng. Soc.*, vol. 26, p. 562 (1978 Jul./Aug.).

[27] T. Oohashi, E. Nishina, N. Kawai, Y. Fuwamoto and H. Imai, "High-Frequency Sound Above the Audible Range Affects Brain Electric Activity and Sound Perception," presented at the *91st Convention of the Audio Engineering Society* (1991 Oct.), convention paper 3207.

[28] www.datrecorders.co.uk/d9601.php

[29] V. Melchior, "Multiphase Filters for Sample Rate Conversion of High Resolution Audio," presented at the *105th Convention of the Audio Engineering Society* (1998 Sep.), convention paper 4854.

[30] M. Story, "High Resolution Converters—Characterising and Specifying their Performance," presented at the *AES 9th UK Conference: Managing the Bit Budget* (1994 May), conference paper MBB-04.

[31] J. R. Stuart, "The Next Generation Audio CD," presented at the *AES 11th UK Conference: Audio for New Media* (1996 Mar.), conference paper ANM-03.

[32] K. Johnson and M. Pflaumer, "Compatible Resolution Enhancement in Digital Audio Systems," presented at the *101st Convention of the Audio Engineering Society* (1996 Nov.), convention paper 4392.

[33] J. A. Moorer, "Breaking the Sound Barrier: Mastering at 96 kHz and Beyond," presented at the *101st Convention of the Audio Engineering Society* (1996 Nov.), convention paper 4357.

[34] S. Yoshikawa, S. Noge, M. Ohsu, S. Toyama, H. Yonogawa, and T. Yamamoto, "Sound Quality Evaluation of 96 kHz Sampling Digital Audio," presented at the *99th Convention of the Audio Engineering Society* (1995 Oct.), convention paper 4112.

[35] H. Sato, H. Yoshida, and M. Yasuchi, "Evaluation of Sensitivity of the Frequency Range Higher than the Audible Frequency Range," *Pioneer R&D*, vol. 7, pp. 10–16 (1997).

[36] J. R. Stuart and R. Wilson, "Dynamic Range Enhancement Using Noise-Shaping Dither Applied to Signals with and without Pre-Emphasis," presented at the *96th Convention of the Audio Engineering Society* (1994 Feb.), convention paper 3871.

[37] M. Akune, R. Heddle, and K. Akagiri, "Super Bit Mapping: Psychoacoustically Optimized Digital Recording," presented at the *93rd Convention of the Audio Engineering Society* (1992 Oct.), convention paper 3371.

[38] M. A. Gerzon and P. G. Craven, "A High-rate Buried Data Channel for Audio CD," presented at the *94th Convention of the Audio Engineering Society* (1993 Mar.), convention paper 3551.

[39] B. Blesser and B. Locanthi, "The Application of Narrow-Band Dither Operating at the Nyquist Frequency in Digital Systems to Provide Improved Signal-to-Noise Ratio over Conventional Dithering," *J. Audio Eng. Soc.*, vol. 35, pp. 446–454 (1987 Jun.).

[40] J. Goodwin, B. Jackson, and P. Levin, "Distilling the Essence of High-Resolution Digital Audio," presented at the *AES 9th UK Conference: Managing the Bit Budget* (1994 May), conference paper MBB-08.

[41] J. Macarthur, "DSP Techniques for Improving A/D Converter Performance," presented at the *AES 9th UK Conference: Managing the Bit Budget* (1994 May), conference paper MBB-03.

[42] M. Story, "A Suggested Explanation for (Some of) the Audible Differences Between High Sample Rate and Conventional Sample Rate Audio Material," http:/www.cirlinca.com/Include/aes97ny.pdf (1997 Sep.).

[43] B. Katz, "24-Bit/96-kHz Mastering: Where Do We Go From Here?," presented at the *103rd Convention of the Audio Engineering Society* (1997 Sep.), Workshop W-12.

[44] http://www.dvdforum.org/about-mission.htm

[45] L. Fielder, "Human Auditory Capabilities and their Consequences on Digital Audio Converter Design," presented at the *AES 7th International Conference: Audio in Digital Times* (1989 May), conference paper 7-010.

[46] J. A. Moorer, "48-bit Integer Processing Beats 32-bit Floating Point for Professional Audio Applications," presented at the *107th Convention of the Audio Engineering Society* (1999 Sep.), convention paper 5038.

[47] J. R. Stuart, "Coding for High-Resolution Audio Systems," *J. Audio Eng. Soc.*, vol. 52, pp. 117–144 (2004 Mar.).

[48] G. Margolis, "Why Super Audio CD Failed" (2016 Dec.). https://audiophilereview.com/cd-dac-digital/why-super-audio-cd-failed.html

[49] https://www.theguardian.com/technology/2007/aug/02/guardianweeklytechnologysection.digitalmusic

[50] F. Rumsey, "Blu-ray or Downloads for HD Audio Delivery?" *J. Audio Eng. Soc.*, vol. 59, pp. 149–154 (2011 Mar.).

[51] D. Reefman and E. Janssen, "One-Bit Audio: An Overview," *J. Audio Eng. Soc.*, vol. 52, pp. 166–189 (2004 Mar.).

[52] P. Nuijten and D. Reefman, "Why Direct Stream Digital (DSD) Is the Best Choice as a Digital Audio Format," presented at the *110th Convention of the Audio Engineering Society* (2001 May), convention paper 5396.

[53] P. Thorpe, N. Bentall, G. Cook, C. Gerard, C. Sleight, M. Smith, and P. Eastty, "DSD-Wide. A Practical Implementation for Professional Audio," presented at the *110th Convention of the Audio Engineering Society* (2001 May), convention paper 5377.

[54] H. Takahashi and A. Nishio, "Investigation of Practical 1-Bit Delta-Sigma Conversion for Professional Audio Applications," presented at the *110th Convention of the Audio Engineering Society* (2001 May), convention paper 5392.

[55] S. Lipshitz and J. Vanderkooy, "Why 1-Bit Sigma-Delta Conversion Is Unsuitable for High-Quality Applications," presented at the *110th Convention of the Audio Engineering Society* (2001 May), convention paper 5395.

[56] "The Advantages of DXD for SACD," reprinted from *Resolution* (2004 Jul./Aug.) http://www.lindberg.no/norsk/artikler/004.pdf

[57] https://en.wikipedia.org/wiki/IPod

[58] https://en.wikipedia.org/wiki/Advanced_Audio_Coding

[59] https://en.wikipedia.org/wiki/MP3

[60] https://medium.com/@RIAA/the-state-of-music-mid-way-through-2017-7e90cad298f9

[61] http://www.riaa.com/wp-content/uploads/2018/03/RIAA-Year-End-2017-News-and-Notes.pdf

[62] http://www.ifpi.org/news/IFPI-GLOBAL-MUSIC-REPORT-2017

[63] http://dsd-guide.com/dop-open-standard#.WqWHpLaZO7Y

[64] http://www.aes.org/technical/documentDownloads.cfm?docID=507

[65] J. R. Stuart and P. G. Craven, "A Hierarchical Approach for Audio Capture, Archive and Distribution," *J. Audio Eng. Soc.*, vol. 67 (2019 May). https://doi.org/10.17743/jaes.2018.0062

[66] J. Angus, "Modern Sampling: A Tutorial," *J. Audio Eng. Soc.*, vol. 67 (2019 May). https://doi.org/10.17743/jaes.2019.0006

[67] J. D. Reiss, "A Meta-Analysis of High Resolution Audio Perceptual Evaluation," *J. Audio Eng. Soc.*, vol. 64, pp. 364–379 (2016 Jun.). https://doi.org/10.17743/jaes.2016.0015

[68] P.A. Fuchs (ed.), *The Oxford Handbook of Auditory Science: The Ear* (Oxford University Press, 2010).

[69] J. Schnupp, I. Nelken, and A. King, *Auditory Neuroscience: Making Sense of Sound* (MIT Press, 2012).

[70] K. Ashihara, K. Kurakata, T. Mizunami, and K. Matsushita, "Hearing Threshold for Pure Tones Above

20 kHz," *Acoust. Sci. & Technology*, vol. 27, pp. 12–19 (2006 Jan.).

[71] K. R. Henry and G. A. Fast, "Ultrahigh-Frequency Auditory Thresholds in Young Adults: Reliable Responses Up to 24 kHz with a Quasi-Free Field Technique," *Audiology*, vol. 23, no. 5, pp. 477–489 (1984).

[72] K. Ashihara and S. Kiryu, "Detection Threshold for Tones Above 22 kHz," presented at the *110th Convention of the Audio Engineering Society* (2001 May), convention paper 5401.

[73] J. R. Stuart, R. Hollinshead, and M. Capp, "Is High-Frequency Intermodulation Distortion a Significant Factor in High Resolution Audio?" *J. Audio Eng. Soc.*, vol. 67 (2019 May). https://doi.org/10.17743/jaes.2018.0060

[74] R. Lagadec and T. G. Stockham, "Dispersive Models for A-to-D and D-to-A Conversion Systems," presented at the *75th Convention of the Audio Engineering Society* (1984 Mar.), convention paper 2097.

[75] S. P. Lipshitz and J. Vanderkooy, "Pulse-Code Modulation—An Overview," *J. Audio Eng. Soc.*, vol. 52, pp. 200–215 (2004 Mar.).

[76] J. Dunn, "Anti-Alias and Anti-Image Filtering: The Benefits of 96 kHz Sampling Rate Formats for Those Who Cannot Hear Above 20 kHz," presented at the *104th Convention of the Audio Engineering Society* (1998 May), convention paper 4734.

[77] P. G. Craven, "Antialias Filters and System Transient Response at High Sample Rates," *J. Audio Eng. Soc.* vol. 52, pp. 216–242 (2004 Mar.).

[78] P. Lesso and A. MaGrath, "An Ultra High Performance DAC with Controlled Time Domain Response," presented at the *119th Convention of the Audio Engineering Society* (2005 Oct.), convention paper 6577.

[79] P. G. Craven, "Controlled Pre-Response Antialias Filters for Use at 96 kHz and 192 kHz," presented at the *114th Convention of the Audio Engineering Society* (2003 Mar.), convention paper 5822.

[80] H. M. Jackson, M. D. Capp, and J. R. Stuart, "The Audibility of Typical Digital Audio Filters in a High-Fidelity Playback System," presented at the *137th Convention of the Audio Engineering Society* (2014 Oct.), convention paper 9174.

[81] M. A. Gerzon, "Why do Equalizers Sound Different?" Studio Sound, pp. 58–65 (1990).

[82] W. Woszczyk, "Physical and Perceptual Considerations for High-Resolution Audio," presented at the *115th Convention of the Audio Engineering Society*, (2003 Oct.), convention paper 5931.

Appendix 1. Major High Resolution Disc Formats.

Release Type	Date	Audio Formats	Sample Rate kHz	Bits	Channels
DVD-Video	1996	PCM, uncompressed	48, 96	16,24	2–8
		DTS (optional)	48, 96	Max 24	2,2.1,5,5.1,6.1
		Lossy formats: AC3, MP2			
Variants:					
DAD (Digital Audio Disc) - audio-only DVD for later DVD players		PCM	96	24	2
DVD-Audio	2000	PCM uncompressed	44.1,48,88.2,96,176.4,192	16,20,24	Various, max 5.1
		PCM, MLP lossless Compression	44.1,48,88.2,96,176.4,192	16,20,24	Various, max 5.1
Double-sided Variants:					
DualDisc	2004	Side a: DVD-(A or V)			
		Side b: non-Red Book 44.1kHz/16b			
DVDplus		Side a: DVD-(A or V)			
		Side b: Red Book CD			
HDAD		Side a: DVD-A			
		Side b: DVD-V or CD			
SACD	2000				
Single Layer disc		DSD	2822.4 (64 Fs)	1	Max 6
Hybrid disc (dual layer)		DSD	2822.4 (64 Fs)	1	Max 6
		CD (Red Book)	44.1	16	2
Dual Layer disc		DSD (both layers)	2822.4 (64 Fs)	1	Max 6
HD-DVD (blue laser disc)	2006, ended 2008	PCM uncompressed	44.1–192	Max 24	Max 8; 2@192 kHz
		Dolby True HD lossless compression	Max 192	Max 24	Max 8
		DTS-HD MasterAudio lossless compr, (opt.)			Max 8
		Lossy formats:			
		DTS-HD High Res			Max 71
		Dolby Dig. Plus,			Max 7.1
		AC3, DTS			Max 7.1,7.1
Variant:					
CBHD (China only)	2007				
Blu-ray (blue laser disc)	2006	PCM uncompressed	48,96,192	16,20,24	Max 8
		Dolby True HD lossless compr, (opt.)	48,96,192	16-24	Max 8
		DTS-HD MasterAudio lossless compr, (opt.)	48,96,192	16-24	Max 8
		Lossy formats:			
		DTS-HD High Res(opt.)	Max 96	Max 24	Max 7.1
		Dolby Dig. Plus (opt.)	48	16-24	Max 7.1
		AC3,DTS	48,48	16-24,16-24	Max 5.1,5.1
Ultra HD Blu-ray	2016	Allows immersive, object-based audio			
Audio-only Variants:					
Pure Audio Blu-ray			96,192	24	2–9.1
High Fidelity Pure Audio			96,192	24	2–5.1

Appendix 2. Current high resolution download and streaming formats

Download Formats		Availability
96 kHz, 24b	FLAC, ALAC compressed or WAV, AIFF	Common
176.4/24, 192/24	FLAC, ALAC compressed or WAV, AIFF	Common
352.8 kHz, 24b (DXD)		Rare, growing
384 kHz/ 24b		
384 kHz/ 32b		
DSD (64 Fs)		Common
DSD 2X (128 Fs)		Common
DSD 4X (256 Fs)		Rare, growing
Streaming Formats		
44.1 kHz, 16b	FLAC compression, typically	Majority
up to 192 kHz/24b	FLAC compression, typically	Available
MQA (hierarchically encoded)	FLAC compression	Available
DSD (64 Fs)		Rare
DSD (128 Fs)		Rare
DSD (256 Fs)		Rare

THE AUTHOR

Vicki R. Melchior

Vicki Melchior studied biophysics, receiving a M.Phil. and Ph.D. in the field from Yale University. Postdoctoral research at Brandeis University was on the molecular structure of nerve myelin based on x-ray, neutron beam, and optical diffraction studies. Joining Massachusetts Institute of Technology's Lincoln Laboratory in 1982, she spent the next twelve years at Lincoln and Raytheon Missile Systems, primarily doing systems analysis and the design and test of digital and analog processing for airborne radar systems.

Throughout, lifelong interests (and varying involvements) in the playing, teaching, and formal study of music meant continuing activity in small ensemble performance, some of it dabbling and some serious, especially of Baroque and Renaissance periods.

Vicki attended the Boston Audio Engineering Society's section meetings for a number of years while working in avionics, opting in 1995, at the beginning of the high resolution era, to join Sonic Solutions to work on the signal processing and software for Sonic's professional audio mastering workstation. Subsequent work after Sonic Solutions included four years as principal engineer at National Semiconductor developing algorithms and low level DSP as well as audio/video integration for a family of DVD VLSI chips, and then, continuing to the present, a consultancy in audio DSP that focuses on high quality audio algorithm development, room correction, IC design, and audio/video processing for professional and consumer audio.

Vicki has served as vice-chair, then chair, of the AES technical committee on High Resolution Audio since its inception in 2000. She is a member of AES technical council, DSP technical committee, and is an Associate Technical Editor. She is also an unabashed audiophile and symphony subscriber.