CHAPTER 18

High Sample Rates: Is This Where It's At?

I. Introduction

Now that we've cured the wordlength blues—it's time to tackle the sample rate issue. Whatever the eventual real benefits for the professional and the consumer, the current relentless drive for higher sample rates is certainly very lucrative for the hardware manufacturers. Clearly, engineers who must regularly replace their expensive high-resolution processors to keep up with the Joneses will spend big dollars.

I've been working with higher sample rates for several years,* but after some experiments that I will relate below, I have concluded that most hardware design engineers are having trouble seeing the forest for the trees. I think that a fresh look at how A/Ds and D/As are designed may reduce the need for extreme sample rates!

A great number of engineers think that the reason higher sample rate recordings sound better is because they permit reproduction of extreme high frequencies. They point out the *open, warm, extended* sound of these recordings as evidence for this contention.[1] However, most objective evidence shows that higher bandwidth is not the reason for the superior reproduction; remember that **the additional frequencies that are recordable by higher sample rates are inaudible.** But if we can't hear these frequencies, then why are we inventing expensive processors and wasting so much bandwidth and hard disc space? And how can 50-year-old ears detect differences between 44.1 kHz and 96 kHz and even 192 kHz sample rates, even though most of us can't hear much above 15 kHz?

* I was the recording engineer for the world's first 96 kHz/24 bit audio-only DVD.

I believe the answer lies in the design of digital **low-pass filters**, which are part of the requirements of digital audio. Digital filters are used in **oversampling A/D and D/A converters** and in **sample rate converters**. Digital filters employ complex mathematics, which is expensive to implement and so, cheaper filters have to include greater quality tradeoffs, such as lowered calculation resolution, ripple in the passband, or potential for aliasing.

One type of filter has a sharp cutoff; the consequences of sharp filtering include time-smearing of the audio, possible short (millisecond) echos which are caused by amplitude response ripples in the passband frequency response (20 Hz-20 kHz), even ripples as small as 0.1 dB. Moving the filter cutoff frequency to 48 kHz (for 96 kHz SR) relaxes the filtering requirement and makes it easier to engineer filters with less ripple in the passband and less phase shift near the upper frequency limit.

"The filters in a typical compact disc player or in the converter chips used in most of today's gear are mathematically compromised."

Oversampling

One of the biggest improvements in digital audio technology came in the late 80's, with the popularization of oversampling technology by DBX's Bob Adams, in a high-quality, 128x oversampling 18-bit oversampling A/D. An **oversampling A/D** converter has a front end which typically operates at 64 or 128 times the base sample rate and produces 1-bit to 5-bit words in delta-sigma format,[2] depending on the model. In other words, for 44.1 kHz operation, the input of a 128X converter actually operates at 5.6448 MHz! Oversampling takes the converter's noise, spreads it around a wider frequency spectrum, and shapes it, moving much of the noise above the audible frequency range. In addition, when it is digitally downsampled to the base rate at the output of the converter, some of the higher frequency noise is filtered out, to yield as much as 120 dB or even better signal-to-noise ratio within a 20 kHz bandwidth.

The downsampling is accomplished with a digital circuit called a **decimator,** which is a form of divider or sample rate converter, and which must contain a filter at half the sample rate to eliminate aliases, requiring a 22.05 kHz cutoff at a 44.1 kHz SR. This filter must be designed without compromise or it will affect the sound. Some manufacturers concentrate on transient response, others on phase response, ripple, linearity, or freedom from aliasing. But all of these characteristics are important, and getting it right is expensive—precision construction requires more math, and math requires labor and parts (size of the integrated circuit die). **Thus, the filters in a typical compact disc player or in the converter chips used in most of today's gear are mathematically compromised.**

On the D/A (output) side, at low sample rates, sharp anti-imaging filters are required to retain frequency response to 20 kHz. It is impractical (probably impossible) to build a sharp analog filter with the required characteristics, so instead an **oversampling or upsampling** digital filter multiplies the base sample rate up 2x to 8x or more, moving artifacts and distortion above the audible band. The higher sample rate permits using a gentle, uncompromised analog filter. But the typical digital filters used in the inexpensive chips have poor performance. To minimize the effect of these concessions, the most progressive high-end D/A manufacturers add an additional upsampling filter of their own design, in front of the DAC chip. The additional filter reduces the error contribution of the chip's own filter, in essence because the internal DAC's filter does not have to work as hard. Internally, these advanced DACs are always operating at 88.2 or 96 kHz regardless of the incoming rate. At the double sampling rates, the supplementary filter is disabled. The supplementary filter would be unnecessary if the manufacturers of the converter chips used higher quality filters in the first place.

An Upsampling Experience

Audiophiles, and some professionals, have been experimenting with digital upsampling boxes which are placed in front of D/A converters. In some cases they report greatly improved sound. Although the improvement may be real, in my opinion they can be attributed to the various digital filter combinations, not to bandwidth or frequency response or (especially) the sample rate itself. Remember that all original 44.1 kHz SR recordings are already filtered, so they cannot contain information above about 20 kHz. An upsampler cannot "manufacture" any new frequency information that wasn't there in the first place.

I've compared the sound of upsamplers versus DACs working alone. Sometimes I hear an improvement, sometimes a degradation, sometimes the sound quality is the same either way. Sometimes the sound gets brighter despite a ruler-flat frequency response, which can probably be attributed to some form of phase or intermodulation distortion in the digital filter. **Sonic differences have come down to mathematics in this new digital audio world**.

The Ultimate Listening Test: Is It The Filtering or the Bandwidth?†

In December 1996, I performed a listening test, with the collaboration of members of the Pro Audio maillist. The idea was to develop a test that would eliminate all variables except bandwidth, with a constant sample rate, filter design, DAC, and constant jitter. The question we wanted to answer was this: Does high sample rate audio sound better because of increased bandwidth, or because of less-intrusive filtering?

"The issues of the audibility of bandwidth and the audibility of artifacts caused by limiting bandwidth must be treated separately. Blurring these issues can only lead to endless arguments." —Bob Olhsson*

The test we devised was to create a filtering program that takes a 96 kHz recording, and compare the effect on it of two different bandwidth filters. The volunteer design team consisted of Ernst Parth (filter code), Matthew

* From the Mastering Engineer's Webboard.

† I previously published some of this information in *Audiomedia* Magazine; we publish the full story in this book.

Xavier Mora (shell), Rusty Scott (filter design), and Bob Katz (coordinator and beta tester). We created a digital audio filtering program with two impeccably-designed filters which are mathematically identical, except that one cuts off at 20 kHz and the other at 40 kHz. The filters are double-precision dithered, FIR linear phase, 255-tap, with >110db stopband attenuation, and <.01 dB passband ripple.

MYTH: *Upsampling makes audio sound better by creating more points between the samples, so the waveform will be less jagged.*

After the filter program was designed, I took a 96 kHz SR orchestral recording, filtered it and brought it back into a Sonic Solutions DAW for the comparison. I expected to hear radical differences between the 20 kHz and 40 kHz filtered material. But I could not! Next, I compared the 20 kHz filtered against "no filter" (of course, the material has already passed through two steep 48 kHz filters in the A/D/A). Again, I could not hear a difference! The intention was to listen double-blind; but even sighted, 10 additional listeners who took part in the tests (one at a time) heard no difference between the 20 kHz digital filter and no filter. And if no one can hear a difference sighted, why proceed to a blind test?

I tried different types of musical material, including a close-miked recording I made of castanets (which have considerable ultrasonic information), but there was still no audible difference. I then created a test which put 20 kHz filtered material into one channel of my Stax electrostatic headphones, and the time-aligned wide-bandwidth material into the other channel. I was not able to detect any image shift, image widening or narrowing—there was always a perfect mono center at all frequencies in the headphones! This must be a pretty darn good filter!

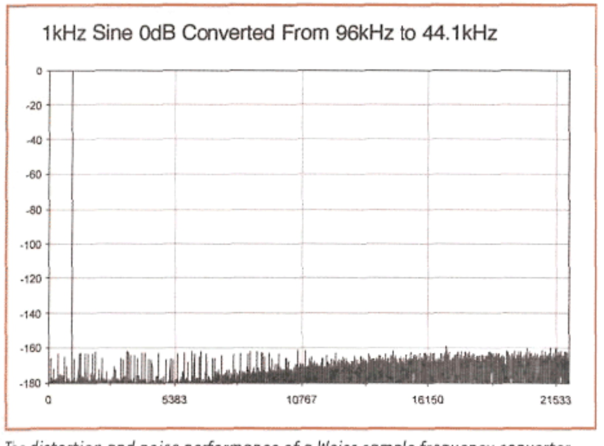
As a last resort, I went back to the list and asked maillist participant Robert Bristow Johnston to design a special "dirty" filter with 0.5 dB ripple in the passband. Finally, with the dirty filter, I was able to hear a difference... this dirty filter added a boxy quality that resembles the sound of some of the cheaper 44.1 k CD players we all know.

This 1996 test seems to show that a "perfect 20 kHz filter" can be designed, but at what cost? Also note that as this test was conducted in the context of a 96 kHz sample rate, the artifacts of two other 48 kHz steep filters already in use may have obscured or masked the effect of the filter under test. Since I conducted my test, several others have tried this filtering program, and most have reached the same conclusion: the filter is inaudible. One maillist participant, Eelco Grimm, a Netherlands-based writer and engineer, performed the test and reported that there were no audible differences using the Sonic Solutions system, yet he and a colleague were able to pick out differences between filtered and non-filtered blind using an Augan workstation. He did not compare the sound of the 20 kHz versus 40 kHz filters, so we are not sure if he's hearing the filter or the bandwidth, but I believe he was hearing the filter, which must not be ideally-designed. I believe the reason he did not hear the differences on the Sonic system is perhaps its jitter was high enough to mask the other differences, which must be very subtle indeed!

Regardless of whether Eelco's group did reliably hear the bandwidth differences, it should be clear by now that the so-called "dramatic" differences people hear between sample rate systems are not likely to be due to bandwidth, but probably to the filter design itself. Ironically, it was necessary to make a high sample rate recording in order to prove that high sample rates may not be necessary.

As I mentioned, 44.1 kHz reproduction has improved considerably in recent DACs employing add-on high-quality upsampling filters. The next figure illustrates Weiss's THD measurement of their SFC, showing that its filter has textbook-perfect distortion and noise performance.

Why can't more manufacturers introduce filters of this quality into their converter chips? The evidence all indicates that it will be a lot less expensive for end-users if the manufacturers of converter chips upgrade the filtering software in their chip sets instead of directing us to this mad, expensive sample rate and format war. Objective experiments must be performed using state-of-the-art digital filters to determine what is the lowest practical sample rate which can be used without audible compromise.



The distortion and noise performance of a Weiss sample frequency converter.

It's A Matter of Time!

Let's be logical: since the human ear cannot hear above (nominally) 20 kHz, then any artifacts we are hearing must be in the audible band. It is well-known that low-Q parametric and shelving filters sound better than high Q; it's not a stretch to conclude this is also true for low-pass filters. Audio researcher Jim Johnston,* who knows as much about the time-domain response of the ear as anyone, has shown that steep low-pass filters create pre-echos which the ear interprets as a loss of transient response, obscuring the sharpness or clarity of the sound.

The pre-echo length is the inverse of the transition bandwith, so a sharp filter with a 500 Hz transition would create a 2 ms. pre-echo. Steep filtering and its attendant transient degradation is probably a reason why 44.1 kHz SR sounds less clear than 96K. Likewise, the increased clarity and purity of 1-bit recordings is probably due to their use of gentle filters rather than some mumbo-jumbo about the "magic" of 1-bit. Jim has experimentally calculated that the minimum sample rate which would support a Nyquist filter gentle enough to elude the ear would be 50 kHz.[3] I suggest that manufacturers and engineers must test as soon as possible the audibility of gentle low-pass filters, at the more common sample rate of 96 kHz. It would

* In correspondence. JJ is the inventor of the science of perceptual coding, which led to coding developments such as MP3, Atrac, etc.

be trivial to build a 96 kHz SR A/D/A system with the gentlest possible filter that's flat at 20 kHz and removes aliasing at 48 kHz, but no current chip manufacturer has done so. This system can be compared against the analog source, and against the competing DSD recording system. If the gentle-filtered PCM wins or sounds as good, it would be the triumph of psychoacoustic research over empirical design. Still, if it can be shown that good-sounding DSD at the consumer end is cheaper to implement than good-sounding gentle-filtered PCM reproduction, it is cheaper for us to record and process with gentle-filtered PCM and finally convert to DSD for the consumer (this is how most 1-bit DACs operate anyway).

I firmly believe that some minimal sample rate (perhaps 96 kHz) will be all that is necessary if PCM-converters are redesigned with psychoacoustically-correct filters (hopefully inexpensively). For the benefit of the myriads of consumers and professionals, we need to make a cost-analysis of the whole picture instead of racing towards bankruptcy.

The Advantages of Remastering 16/44.1 Recordings at Higher Rates

Researchers such as J. Andrew Moorer of Sonic Solutions, and Mike Story of dCS have demonstrated theoretical improvements from working at a higher sampling rate. Moorer pointed out that post-production processing, such as filtering, equalization, and compression, will result in less distortion in the audible band, as the errors are spread over twice the bandwidth—and half of that bandwidth is above 20 kHz.[4] Measurements discussed in Chapter 16 confirmed these conclusions. In addition, as we've seen above, if after processing the destination is DVD-A or SACD, then the master can be left at the higher sample rate and wordlength, avoiding another generation of sound-veiling 16-bit dither and yet another sharp filter at the end of the process. Thus, consumers should not scoff at DVDs which have been digitally remastered from original 16-bit/44.1K sources. They will be getting real, audiophile-quality sonic value in their remasters.

---

1 Other engineers who do not fully understand the nature of PCM argue that the higher sampling rate sounds better because it would seem to create a more accurate 20 kHz sine wave, as there are more "dots to connect" to describe the wave. But this is erroneous; while there are more "dots," in reality only 2 samples are necessary to describe an undistorted 20 kHz sine wave; the low-pass filtering smooths out the waveform and eliminates all the glitches.

2 DSD, also known as 1-bit or **Direct Stream Digital**, a trademark of Sony and Philips is the format of the SACD and employs a form of Delta-Sigma modulation. Delta-Sigma modulation is the very dense native coding format of the first stage of modern-day oversampling converters, about 2.8 Megabits per second, as opposed to 44.1 kHz/16-bit PCM, **Pulse Code Modulation**, which runs at about 1.4 Megabits per second. When you study the block diagram of a record-reproduce chain, the significant difference between using DSD format and PCM is that PCM requires a steep Nyquist filter at half the sampling rate (about 20 kHz with 44.1 kHz SR).

3 This is based on the length of the shortest *organic* filter in the human ear, and Jim Johnston notes that the 50 kHz number nicely matches the original work with antialaising filters done by Tom Stockham for the Soundstream project.

4 Julian Dunn (in correspondence) clarifies: A 3 dB reduction in distortion results because the error products are spread amongst twice the bandwidth. This is true for **uncorrelated** quantization errors which fall evenly throughout the frequency range from dc to fs/2. And does not work for distortion products which will correlate with the signal. Jim Johnston (in correspondence) indicates that processing at higher rates is **required** for any non-linear processing, such as compression. These non-linear processes produce new frequency components, some at higher frequencies. A high enough sampling rate avoids aliasing of these new frequency components (see Cranesong and Weiss FFTs in Chapter 16).