Lately, there has been a surge in interest in what has been dubbed "upconversion" of digital audio to higher sampling frequencies and bit-rates. This is fed by the availability of 24-bit / 96 kHz audio, recorded on DVD-video discs. It is also fed by some companies' promotion of their products as doing "upconversion" of the signal, making analogies with upconversion of video signals by excellent products like the Snell & Wilcox Interpolator.

This paper will try to dispel some of the myths surrounding these "upconversion" technologies.

There is a finite amount of information contained in any linear (PCM) digital recording. At best, assuming perfect conditions throughout the recording chain, this quantity of information is dictated by the sampling rate (how many samples of the original analog waveform are taken per second) and the sampling depth (the precision, measured in bits, with which these samples are measured). The standard rate and depth for the Compact Disc is 44,100 samples per second and 16 bits per sample.

Sonic improvements can be realized by increasing either the sample rate or the sample depth, as long as there is no significant limiting factor in the analog path before the signal is converted to digital. [Obviously, a "96/24" (96,000 samples per second at 24 bits per sample) recording using a Fisher-Price microphone will not sound significantly different than the 44/16 version of the same thing.] However, once you have captured everything that a given sample rate and depth can capture, hold, any leftover information is lost forever.

On playback, the trick is to preserve every iota of the information contained within the digital recording in order to recreate as much of the original performance as the recording managed to capture. hold, any leftover information is lost forever.

When Sony and Philips first introduced Compact Disc in 1982, the partners themselves disagreed about how to best accomplish this goal of preserving the information on playback.

Given the limitations of 1982 technology, Sony felt it best to use 44.1 kHz, 16-bit digital to analog converters. After all, these would exactly reverse the process that had been conducted to get the recording onto the CD in the first place. omplish this goal of preserving the information on playback, Philips elected to use 14-bit converters running at a higher frequency—specifically, 176.4 kHz, which is four times faster than 44.1 kHz. In order to do this, Philips used a technique called "oversampling," whereby three additional samples are interpolated between each original sample from the recording itself. This process is done using various mathematical techniques, but creates intermediate values between the "real" or original samples. The result is similar to a connect-the-dots child's drawing that now has more dots, spaced more closely together. Specifically, it contained four times as many dots and became known as "4x oversampling."

Mathematically, both the Sony and the Philips approaches in 1982 were identical. Each contained exactly the same amount of information—Philips had simply re-packaged the information into a different form, using more samples with less depth per sample. The advantage Philips had was that their approach allowed them to use better-sounding analog circuitry after the signal had been converted back to analog.

This takes a moment to explain. When digital audio is converted back to analog, it does not recreate the continuous, smooth waveform that was originally sampled. Instead, the waveform is only a reasonable facsimile of the original, but with small "staircase steps" all along the length of the curve. These sharp edges create large bands of distortion centered on multiples of the conversion frequency, with a bandwidth equal to plus-and-minus the range of frequencies recorded. In 44.1 kHz audio recordings, these bands extend out ± 20,000 Hz around 44,100 Hz and its multiples (88.2 kHz, 132.3 kHz, 176.4 kHz, etc.)

Note that the first of these bands of distortion reaches as low as 24.1 kHz (44.1 kHz minus the 20kHz bandwidth of audio). This is precariously close to what we hear, and it needs to be filtered out effectively in order to avoid serious sonic compromises like IM distortion, overloaded tweeters, and power amps having fits over large-signal ultrasonic information they were never designed to amplify.

For this reason, Sony had to include what are called "brick-wall filters" in their first-generation CD players: ninth- to eleventh-order analog filters that added complexity, cost and significant audio compromises (terrible phase shift, in-band ripple, and other problems).
By contrast, the Philips approach was designed to fool the digital to analog converter into thinking that it was working with a 176.4 kHz signal. Even though there was no additional information contained in the signal, the fact that it was being converted at 176.4 kHz meant that the bands of severe distortion were centered on multiples of 176.4 kHz (and they were still only ±20 kHz wide). Thus the lowest edge of this band of distortion reached down only to 156.4 kHz—nowhere near the audio band, and easily handled by a simpler, better-sounding analog filter.

In the long term, the Philips approach won out, of course. Every CD player since approximately 1984 has used oversampling to ensure that a simpler, better-sounding output filter could be used.

Of course, in relatively short order, DACs capable of reproducing all 16 bits at higher frequencies (4x, 8x, even 16x 44.1 kHz) became available. Then 18-bit DACs, 20-bit DACs, and even 24-bit DACs became available, all capable of running at extremely high rates with excellent accuracy. Even the cheapest DACs today will outshine the finest technology had to offer in 1982, and by a wide margin.

So what’s the point?

The point is that “upconverting” digital audio to higher rates for conversion has been around, and in fact has been the overwhelmingly dominant way of doing things since somewhere between 1982—1984, depending on whom you ask. Historically, it has been called “oversampling.” But whether you call it “oversampling,” “upconversion,” “upsampling,” or anything else, it remains the same. It is a method by which the originally sampled data is converted to a higher rate for the purpose of better-quality conversion to analog.