Digital Distortion in CD's and DVD's:

The Consequences of Traditional Digital Peak Meters

A white paper by Nika Aldrich for Trillium Lane Labs

Introduction

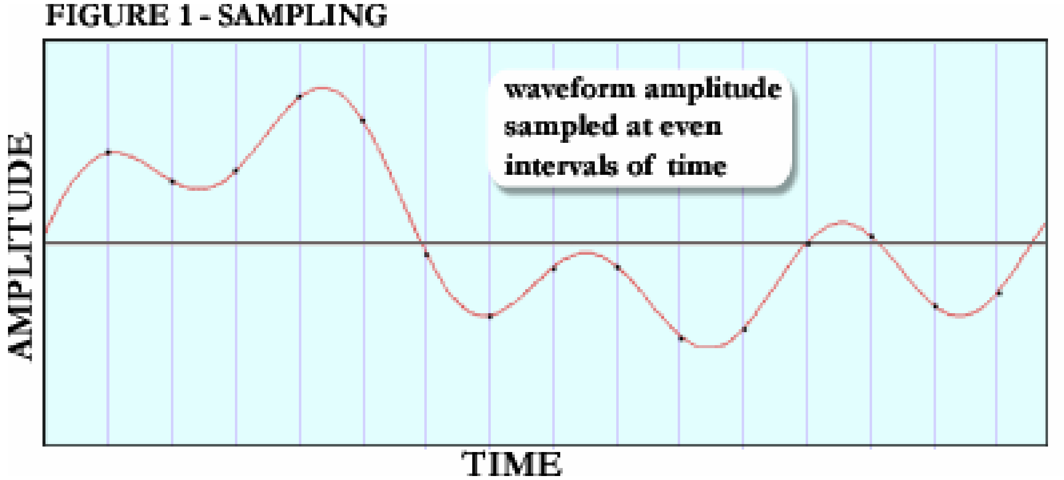
The advent of the Red Book CD standard in the early 1980's was heralded as a breakthrough in consumer audio quality. Many discerning listeners however, claim that modern CD's have actually decreased in audio quality. This white paper explores the impact of industry demands for the loudest possible mixes and the resulting effect on digital audio quality as one possible reason for the perceived decrease in quality in modern CD's.

History

The digital sampling theory was first proposed in 1928 in a paper by H. Nyquist entitled "Certain Topics in Telegraph Transmission Theory" [Nyquist 1928] The theory was mathematically proven by mathematician Claude Shannon in his paper, "Communication in the Presence of Noise" [Shannon 1949]. According to Shannon's theorem:

"Theorem 1: If a function f(t) contains no frequencies higher than W cps [cycles per second, or "Hz"], it is completely determined by giving its ordinates at a series of points spaced 1/2W seconds apart."

FIGURE 1 - SAMPLING



The significance of this statement is often overlooked. Shannon tells us that the *entire* waveform, in proper amplitude, frequency, and phase can be recreated (as he says, "completely determined") with only sampling points given at greater than half the highest frequency to be sampled. A key observation is that the entire waveform is represented by the sampling points, but a reconstruction process still needs to occur in order to recreate the waveform represented. One cannot simply 'connect the dots' between sample points and yield the original waveform. Shannon says that the waveform (function) can be "completely determined" but does not describe that process – he simply alludes that it can be done.

FIGURE 2A - A COMPLEX WAVEFORM

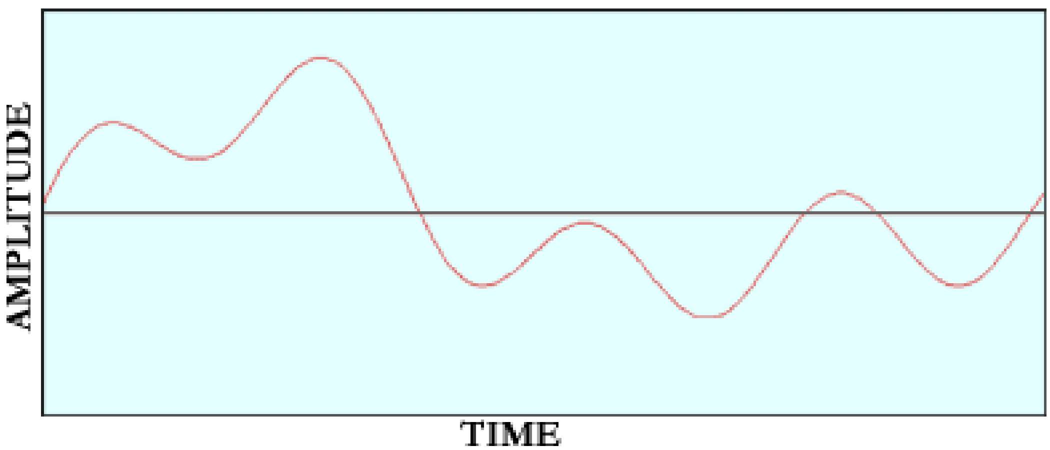


FIGURE 2B - WAVEFORM IN FIGURE 2A, SAMPLED

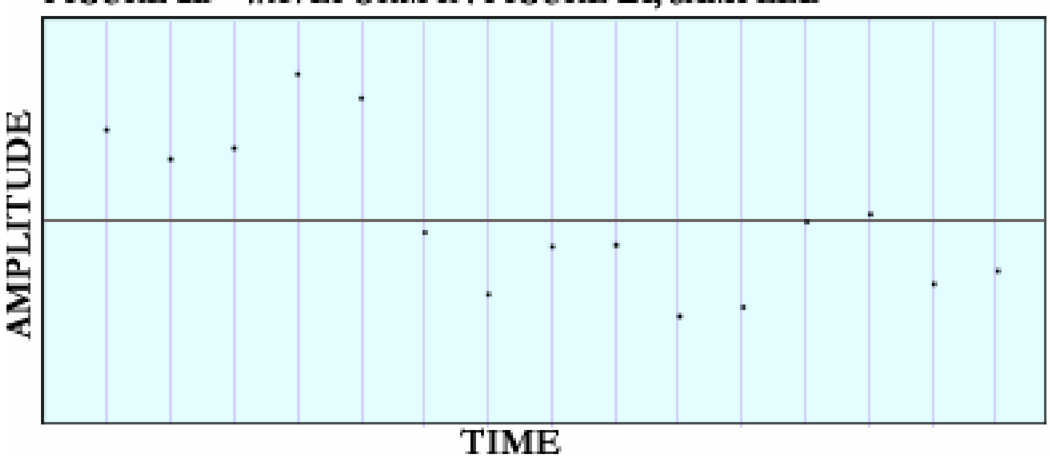


FIGURE 2C - WAVEFORM AS REPRESENTED IN DAW

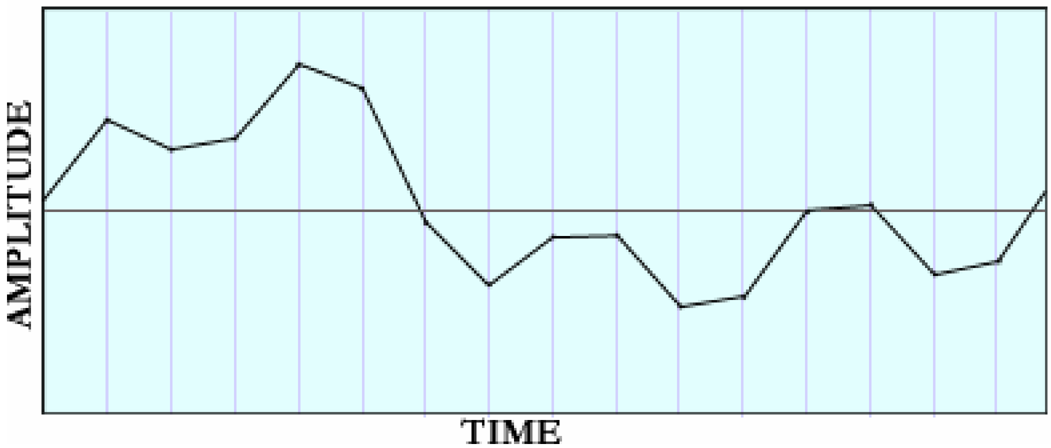
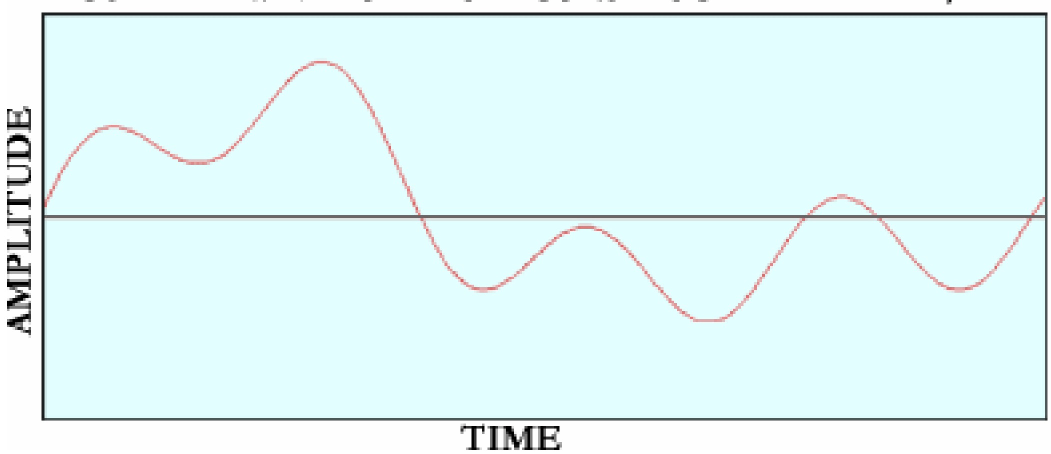


FIGURE 2D - WAVEFORM AS RECONSTRUCTED AT THE D/A



The process of recreating the original waveform involves a filter called a reconstruction filter. This filter removes all content above the Nyquist frequency (half the sample rate). The range below the Nyquist frequency defines the 'legal' range of allowed frequencies as frequencies in this range can be accurately reproduced. All frequencies above the Nyquist frequency do not adhere to Nyquist or Shannon's theorems regarding allowable frequencies, cannot be reproduced and are therefore considered illegal frequencies. Because of mathematical realities observed by Fourier in the 1800's, and subsequently by Shannon in 1948, when a waveform has all frequencies removed above the Nyquist frequency, the resulting waveform will be the original waveform that was sampled.

This process is significantly more involved than simply 'connecting the dots' between sample points. Today it involves extremely sophisticated means of reconstructing the waveform, using filters that are highly complex mathematical systems utilizing 'oversampling,' 'upsampling,' 'linear phase, equiripple FIR' designs and much more. The result is that today's digital to analog converters get closer to the original than ever before, making music played on systems today as accurate as possible. Even today's inexpensive components such as off-the-shelf CD players have drastically improved filters, and thus better reconstruction abilities than in years past.

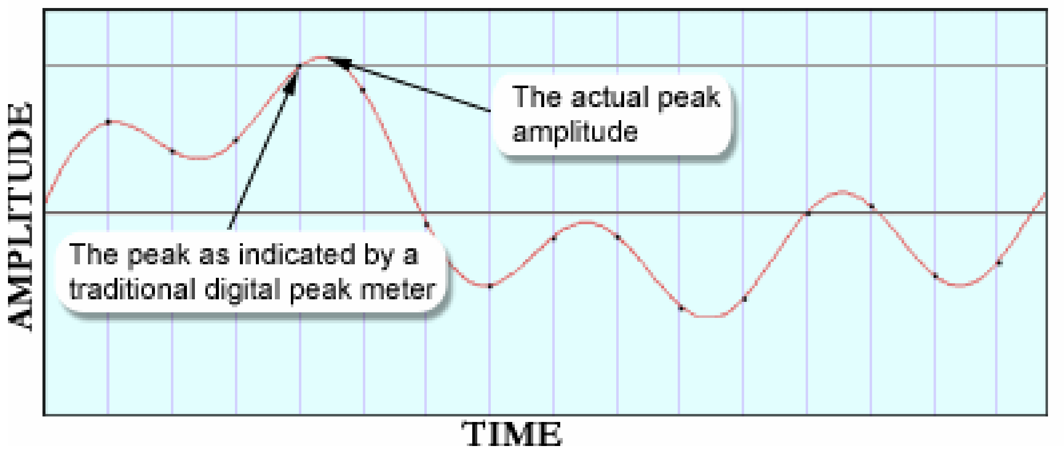
Application

Most contemporary audio recording is done with Digital Audio Workstations (DAWs), although digital mixing systems in the form of outboard digital mixers are also very popular. To the user, these digital systems appear similar to traditional audio tools, and are designed order to emulate the operation of a conventional analog recording system.

One familiar analog tool that has been bought over to the digital realm is a 'peak meter' that tells the amplitude of the waveform's peaks. In the analog realm, peak signal was an indicator that would tell the audio engineer when the peak signal level was getting too high. A peak signal in analog recording would cause the tape to saturate, creating distortion. In an analog system however, this type of distortion was often deliberately engineered into tracks in order to achieve a certain sound. In the digital realm this type of meter is important and more vital, because if the amplitude of a waveform exceeds the top of the measurable scale (full scale, or 'full code') the signal will 'clip' causing unwanted and unpleasant distortion rather than the traditional distorted sound of analog. This digital clipping occurs because the waveform is 'lopped off' and the data is changed. When the waveform is reconstructed it cannot be accurately done so in order to represent the original waveform. Instead, it has a significant amount of inharmonic distortion caused by aliasing. For this reason, digital recording has a maximum level at which signals can be recorded. Anything exceeding this level (full scale) has undesirable consequences.

The method used for computing the peak value inside the system however, is not particularly accurate. DAW and digital mixer manufacturers typically take the amplitude of the *samples* and use these as the basis for the peak meter. The problem with this approach is easily identified: the *samples themselves* do not represent the peak value of the waveform. The waveform is only complete after the reconstruction process. Until this process has been completed, the waveform is inaccurately represented by the samples. This is the reason that in most DAWs the waveform is represented on the screen as a 'dot to dot' connection between sample points. They do not undergo the reconstruction process inside the system, so all that can be represented is the sample points, and for the sake of visual ease, they connect the dots between them with straight lines. They save the reconstruction process for the digital to analog converters and show the user inaccurate information instead.

FIGURE 3 - INTERSAMPLE PEAKS



The consequence of the way in which DAW's treat waveforms is that the meter inside the DAW or other digital mixers inevitably shows inaccurate information. It is virtually a mathematical certainty that the *waveform* will actually *exceed the amplitude of the samples* in any sampling system. The samples themselves only *represent* a waveform, It is important to understand that the amplitude of the waveform will invariably exceed the sample values.

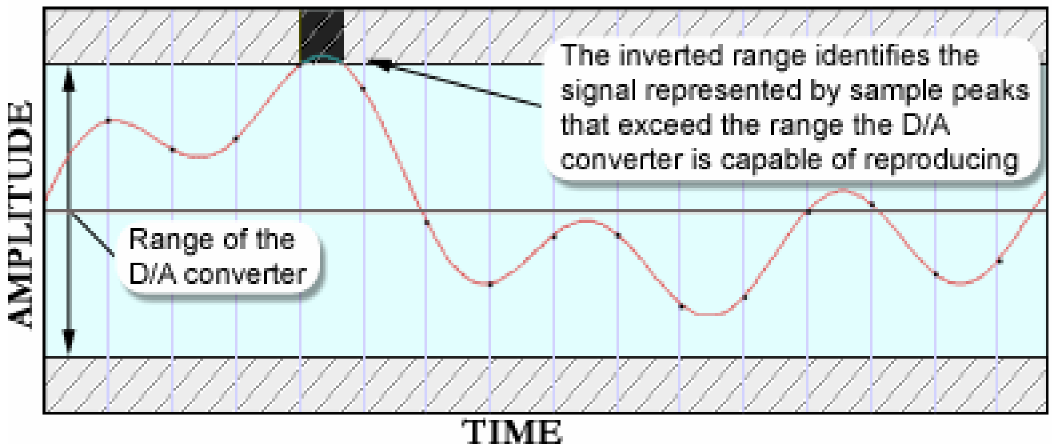
Manifestation

One may ask why this poses a problem. For various reasons, mostly having to do with marketing demands and industry trends, recordings made and mastered in today's recording environment are mixed and mastered as 'hot' as is possible, pushing the levels up to the highest tolerable amount, supposedly just short of clipping. Sophisticated digital tools allow music to be highly compressed, then recompressed, compressed even more so with multi-band compressors, limited, normalized, and maximized to get the audio to play as loud as possible out of a consumer's system. Hence, it is very common for popular music CDs to be full of digital samples that are at, or nearly at full scale.

The problem is realized in that while going through these digital gyrations and utilizing digital tools to amplify the signal as much as possible, both during mixing and during mastering, the 'peak value' of the sample points is closely watched to ensure that it does not get to full scale. Since, the peak meters in said DAW and digital mixing systems are inaccurate, and do not actually indicate the peak values of the resulting waveform, the result is that while the *samples themselves* do not exceed full scale, and are carefully monitored to insure this, the resulting *waveforms* ***represented*** *by the samples may exceed full scale throughout any standard CD!*

While the digital mixing system is not clipping the music or distorting the music, the digital to analog converters that have the task of recreating the audio through digital reconstruction filters are clipping repeatedly throughout most CDs on the market. The result is that most CDs and DVDs end up distorting with regularity when they are asked to reconstruct and play back audio that *appears* to be completely 'legal' because not a single *sample* actually clipped.

FIGURE 4 - D/A CONVERTER RANGE



In a recent paper [Nielsen 2003], seven consumer CD players were subjected to tests designed to analyze their ability to reproduce and reconstruct signal levels above full scale (0dBFS). All of the players experienced difficultly dealing with signal levels this high, further showing that, while all of the samples can be legal, the level can still be hotter than is legal the result being that a CD player can be unable to reproduce the audio accurately.

It is nearly certain that this constant barrage of distortion that we, the consumers, are hearing on compressed and mastered CDs contributes to the 'digital harshness' still reported by the more sensitive audiophiles in the music industry. According to industry insiders, not a single off-the-shelf digital to analog converter chip made today can accurately pass a signal wherein the samples are under full scale but the waveform that they represent exceeds full scale. Only a few high end converters in the professional market can do this. This means that the preponderance of consumer (and professional audio) playback equipment is not designed to deal with these 'hotter than full scale' signal levels.

Monitoring in most mastering studios is typically performed using high end digital converters. Consequentially, audio and mastering engineers are often putting out music that cannot be accurately recreated by consumer playback equipment. In some cases the reconstruction sounds 'perfect' to the mastering engineer – because the engineer's equipment actually *can* reproduce the waveforms properly.

Potential solutions

There are several potential solutions to this problem. On the reproduction side of the equation, one solution would be to equip the consumers' home and car playback equipment with modified digital to analog conversion so that it can actually recreate illegal waveforms that exceed full scale at their peaks. The magnitude of this chore is not only unrealistic, but unnecessary as well. In truth, the waveforms we are asking the machines to recreate are illegal and contain information that exceeds the bounds specified by the systems being used. Rather than change the equipment, a more appropriate solution can be found in merely changing the content.

On the content creation side, there are also alternatives for audio and mastering engineers. Changing the content requires not putting out mastered music that exceeds full scale. This means putting out music that, even though the *samples themselves* hardly ever exceed full scale, does not have the *waveforms* exceeding full scale.

One alternative for engineers is simply to turn down the level by a fixed amount at the mastering stage, to ensure the waveform will not clip when reconstructed. This is an imperfect solution for two reasons. First, it sacrifices potentially unused dynamic range and second, it is unlikely to be acceptable to clients given current industry demands for the loudest possible mixes.

The second is to monitor the *reconstructed* waveform for clipping at the final mix and mastering stages and make appropriate adjustments without sacrificing overall level or dynamic range. This requires a digital mixing and mastering system that has peak meters that simulate the reconstruction filters used in digital to analog converters throughout the professional audio and consumer industries. This task is difficult without appropriate metering tools that allow mixing and mastering engineers to know what the *actual* signal level of their music is while they work on it, rather than just the level of the samples used to represent it.

Impact on mixing and mastering

The inevitable result is that, in order to comply with the actual legal range of the digital audio system, mixes will have to be reduced in amplitude so that when a specific waveform exceeds the sample values there is enough headroom for it to be reconstructed below full scale. Studies have shown that waveforms can exceed full scale (considering the reconstruction filters on most digital to analog converters) by more than 6dB. This means that the peak amplitude of the actual waveform might be more than twice as high in amplitude as the highest sample value. This is only likely to happen when music is heavily compressed however, and most music will practically require less than 6dB of headroom above the highest sample value to ensure accurate reproduction.

In the end the mastering engineer will be left with two choices if they want the music to be in compliance. They will either need to compress or limit the music less, so that the RMS value of the music is lower, lowering the 'perceived' level of the music in order to keep it legal, meaning that the music will take advantage of the increase in headroom to have more dynamic range. The other solution is that they can continue to compress the music and limit it as hard, but will have to turn it down to give room for the peaks that occur above the sample levels. Either way, the result of compliance will require a lowering of music that ends up in delivery formats to a return to the legal range.

FIGURE 5A - A WAVEFROM USING MAXIMUM DIGITAL RANGE

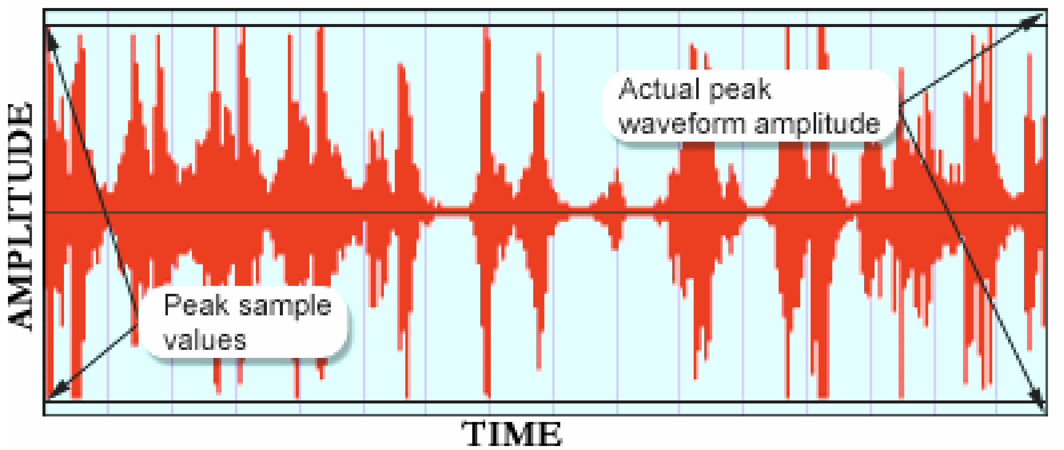
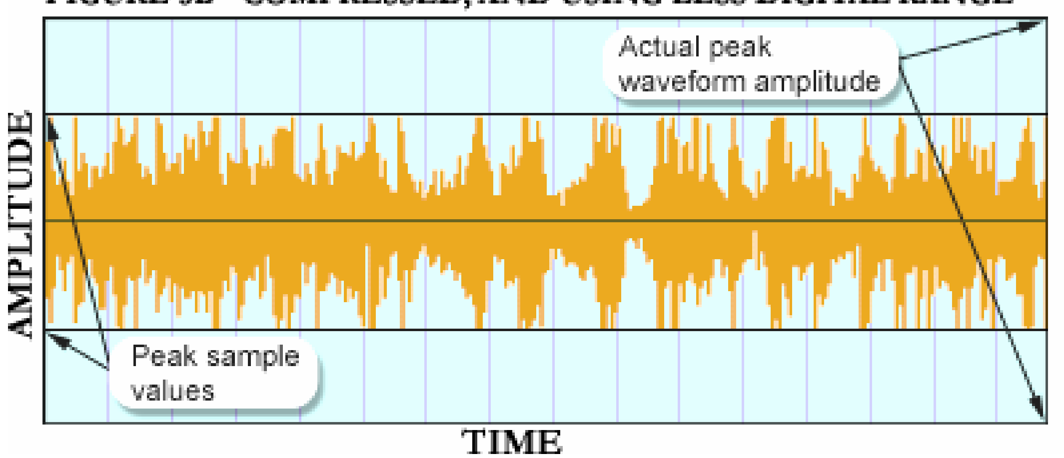


FIGURE 5B - COMPRESSED, AND USING LESS DIGITAL RANGE



It is worth noting that Sony's new SACD format includes measures that prevent the music from ever clipping in the way described. Mastering engineers who work on SACD releases have observed the notion that heavily compressing the audio inevitably results in the need to 'turn down' the overall level on the disk. Left with the choice of compressing the disk and turning it down, or simply leaving it the opportunity to 'breathe' with some headroom, most mastering engineers are mastering to SACD disks differently than they have been to DVD's and CDs. Many professionals in the audio industry are claiming that audio that has been remastered for the SACD format 'breathes more,' 'has more life,' and 'doesn't have the digital harshness' of the CD counterparts.

Conclusion

The mastering or mixing engineer that first starts using an oversampled peak meter capable of representing the audio waveforms may at first be frustrated that it is difficult to get their final result as loud as they could otherwise. This is only partially true.

Of course, since many popular music CD's have been clipping consumer digital to analog converters, accommodating those systems will inevitably require lowering the level in some method or another, resulting in a quieter final product in some capacity, although likely by only a few decibels. Where this is not necessarily true, however, is that the PCM system and the CD both allow for the representation of illegal waveforms such that it is not a *requirement* to lower the level just because the mastering engineer is endowed with tools that show him that he is allowing a distorted result to be reproduced. The red book format for CDs and the DVD specs both allow for this illegal content, and the mastering engineer is still allowed to put out releases that meet the *spec* while allowing consumers' players to distort. At least with an oversampled peak meter the engineer will be able to know that the music is clipping, by how much, and where. With this knowledge the engineer can then decide with complete information whether or not to accommodate the legal range of digital audio on a PCM sampled system.

The consumer has continued to complain that CDs and DVD's sound 'harsh,' the mastering engineers have argued that peak meters continue to be inaccurate, and everyone continues to demand better sounding mixes. Utilizing an oversampled peak meter in the digital audio studio that represents the reconstruction filters in digital to analog converters is the first step toward an improvement in audio quality in music releases.

References and further reading

Aldrich, Nika. Digital Audio Explained For the Audio Engineer. San Francisco: Backbeat Books, 2004.

Banquer, Dan, Dick Pierce, Herbie Robinson, et al. "Intersample Peaking." Pro Audio Mailing List. 21 December, 2002 - 31 December, 2002.

Nielsen, Soren and Thomas Lund. "Level Control in Digital Mastering." Preprint 5019, 107th AES Convention. Denmark, 1999.

Nielsen, Soren and Thomas Lund. "0dBFS+ Levels in Digital Mastering." TC Electronic: Risskov, Denmark. 17 July, 2003. <http://www.tcelectronic.com/media/Level_paper_AES109.pdf>.

Nyquist, Henry. "Certain Topics in Telegraph Transmission Theory." Transactions of the AIEE. Vol. 47 (April 1928): 617-644.

Shannon, Claude E. "Communication in the Presence of Noise." Proceedings of the IRE. Vol. 37 (January 1949): 10-21.