Audio Engineering Society

Convention Paper

Presented at the 121st Convention
2006 October 5–8 San Francisco, CA, USA

Stop Counting Samples

Thomas Lund

TC Electronic A/S, Risskov, 8240, Denmark
thomasl@tcelectronic.com

ABSTRACT

Level restriction in digital music production has traditionally been based on measuring the value of individual samples. Where sample counting may have been appropriate in the early days of digital, previous work has revealed how dynamics processors now exploit our archaic measurement principles to an extent where significant distortion can be expected to develop downstream of the studio in perceptual codecs, DA and Sample Rate converters.

The paper suggests that production methods in combination with simplistic level assessment is responsible not only for more distortion and listener fatigue, but also for level jumps where digital interfacing or file transferring is used, e.g. at a broadcast station. Improved working practices and measurement methods are suggested.

1. INTRODUCTION

When CD was introduced, analog tapes were typically used for production. During mastering, the sound was passed through analog processing, and eventually converted to digital, where the level was read fresh out of the AD converter. Today, production procedures have changed dramatically, and data reduced delivery is more the rule than the exception, but the way we measure level has remained the same. The old CD level control method has even spread to other production areas, such as broadcast, post and film.

The purpose of this paper is to justify and recommend more fitting ways to measure and control level in production and mastering than looking at isolated samples. Audio engineers should realize that aiming only at max samples and max absolute loudness has its price of unpredictable and distorted reproduction.

This paper describes the consequences of current leveling techniques, unconscious digital clipping in general, and clipping when the signal is close to Full Scale. The topics discussed are relevant to professional production and mastering engineers in music, broadcast and film.

1.1. Definition of Terms

Even the simplest of waveforms, the sine wave, can be constructed in ways which cause analog peaks not to align with digital peaks representing the same signal, see *Fig 1*. The analog level of a sine wave at fs/6 (*Fig 1-2*) can be up to 1.25 dB above the peak level in the digital domain, while at fs/4 the discrepancy can be up to 3 dB.

Put differently, sine waves can need a DA conversion headroom of 3 dB for distortion-free reproduction, but other signals can be created in the digital domain (for instance square waves or pseudo-random MLS sequences), where a headroom of 6 dB or even more would be needed for reconstruction. A specific DA converter can be targeted with its worse case signal, and require ridiculously large amounts of headroom.

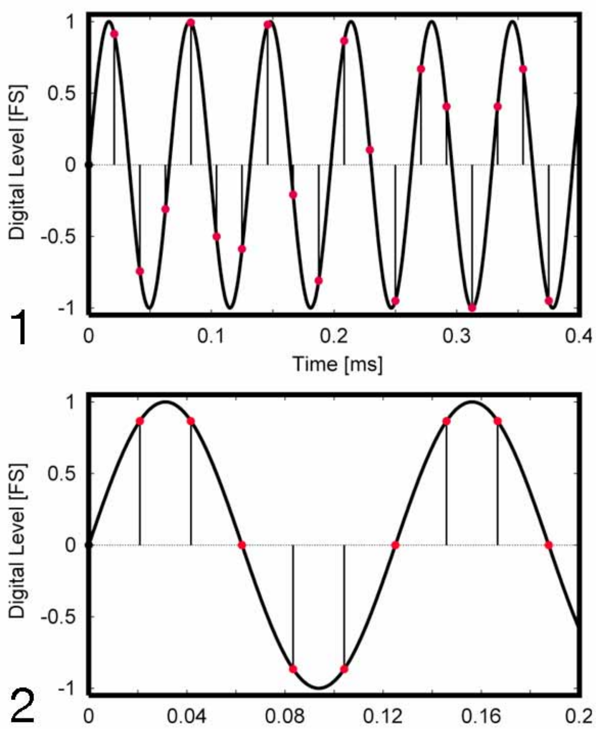

Fig 1. Digital (dots) vs. Analog (line) level.

In this paper, the resulting reconstructed or resampled true-peak level will be called *intrinsic level*, and when it is above Full Scale (with ideal reconstruction), it will be referred to as 0 dBFS+.

A digital level meter showing the max sample level will be called a Digital *Sample* Meter, while a meter showing intrinsic level will be called a Digital *Signal* Meter.

Though sample synchronous sine waves are rarely used in audio production, previous studies have proved them useful for testing a signal-path for non-linearities when intrinsic level exceeds 0 dBFS.

1.2. The Creation of 0 dBFS+ Level

As trivial as it may sound, the basic cause of overload in digital audio is the lack of waveform preservation in various conversion processes. There are several possible causes for this waveform change [1, 2], which can cause production of 0dBFS+ results when samples get in the vicinity of Full Scale.

1. The very nature of sampled signal representation
2. Change of phase (example *Fig 1*)
3. Change of bandwidth
4. Non-linear distortion such as clipping

Removing part of the spectral content of a broadband signal is likely to cause a peak level increase - despite the reduced energy in the resulting signal. A very simple example of this is a low frequency limitation, i.e. a high pass filter, which is present at most analog interfaces in order to avoid DC offset. Filtering a low frequency square wave with just a first order high pass filter will result in a peak level increase of up to 6 dB. Sometimes, such filters occur also within the digital domain. It may be argued that some of the peak level increase is caused by the phase change rather than the gain change, but still the peak level rises when applying a high pass filter to this type of signal.

For a given filter it is possible to construct a test signal which will excite the filter in a worst-case fashion with respect to achieving a maximum output peak level, i.e. to create an impulse-like output. As such test signals would have to be adapted to each specific filter they have not been take into consideration here.

A more probable extreme signal is a pseudorandom sequence, alternating between plus and minus full scale at (pseudo-)random intervals. The MLS sequences often used for acoustical measurements are of this type [1].

Constructing a square wave in the digital domain from a harmonic series will be limited by the Nyquist frequency, and thus the square wave will be rounded and oscillating. An alternative way of generating square waves in the digital domain is by *clipping*. The resulting signal is not necessarily a pure square wave but it typically has a flat top and sharp edges produced by saturation logic. Overly fast dynamics processors can also generate this kind of distortion, so a familiar name, such as compressor or limiter, is no guarantee against invisible pollution of the signal in ways analog processors would not.

The effects of truncation of the Fourier series in order to represent a square wave in a finite bandwidth system is sometimes described as Gibb's phenomenon [3]. The phenomenon plays an important role in windowed FIR filter design where the resulting ripple (oscillations) in the passband and stopband frequency response is a design parameter.

It should be noted that digitally clipped signals don't respect the sampling theorem, and therefore produce variable amounts of alias distortion *on top of* the overload trouble they may cause, if the clipping happens close to Full Scale.

2. DIGITAL LEVEL MEASUREMENT

The principle behind measuring level in CD production is as old as the media itself. Level was and is measured purely on a peak sample by sample basis. The only thing to be concerned about is not to hit 0 dBFS with too many samples in a row.

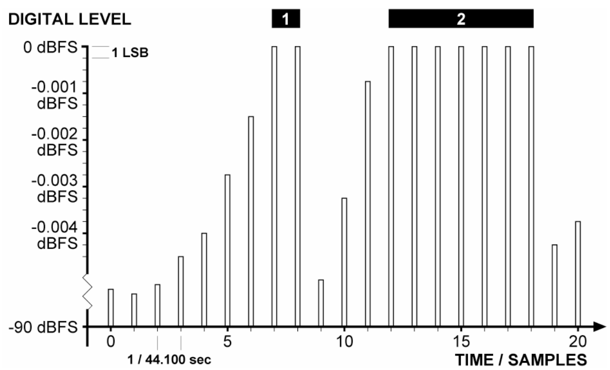

Fig 2. Example of consecutive samples on a CD. Note magnification of Full Scale area.

Fig 2 shows samples encoded to 0 dBFS. Event 1 would typically not be considered an over, while Event 2 might cause rejection by some CD mastering plants.

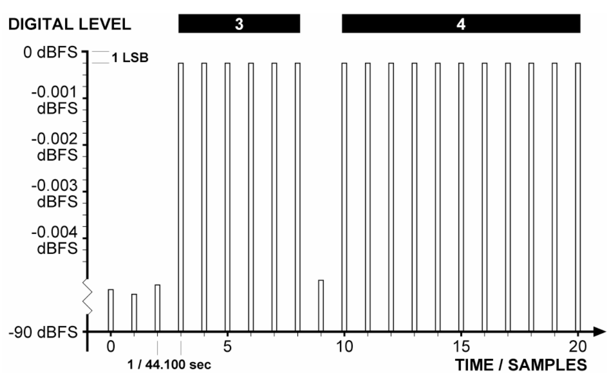

Fig 3. The samples of Fig 2 with 1 dB of boost, minus 1 LSB.

Fig 3 shows the signal from Fig 2 subjected to a 1 dB boost in level, and subsequent attenuation of 1 LSB. Neither event 3 or 4 will be detected as an over at the mastering plant, even though Fig 3 of course is more distorted than Fig 2. The clipping is not detected because 0 dBFS is never reached.

When CD was released to the public in 1982, analog 1/4 or 1/2 inch tapes were typically used for music mixing. During mastering the signal was passed through analog processing, and eventually converted to digital. The level and the AD converter headroom was read in the digital domain. Back then, the consecutive sample count method for detecting overloads made sense, because all signals fulfilled the sample theorem fresh out of the AD converter. The conversion process, with its associated low-pass filters, ensured the validity of the digital signal, and was the reason why the rudimentary sample counting way of measuring level worked in the first place.

While the Sony PCM codecs 1610/1630, DMU-30 meters and DTA-2000 analyzers were good instruments, and perfectly suited for the era they were designed for, the level control principles they utilize don't take digital processing into account, and therefore are not adequate anymore.

Nowadays, production methods are different. Audio is digital when it arrives at the mastering studio, and there is no guarantee the signal doesn't already contain out of band components from clipping or misbehaved upstream digital processors or workstations.

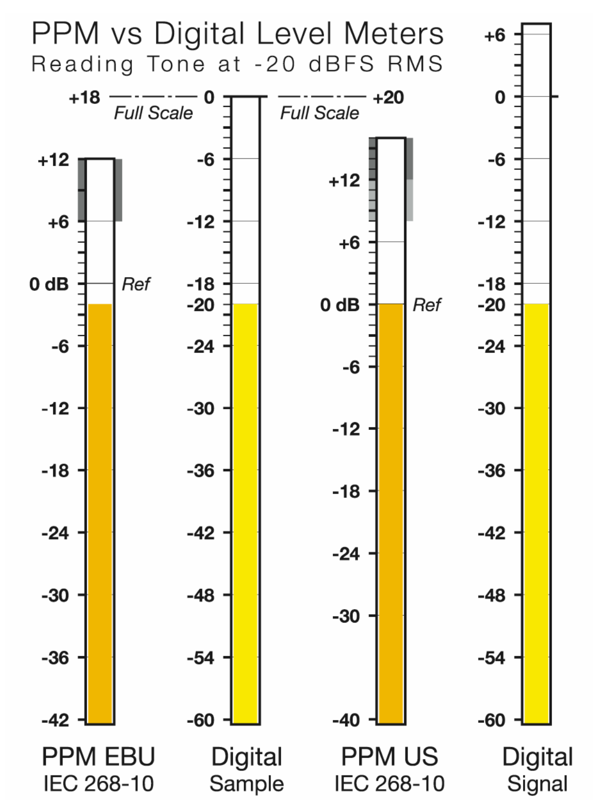

Fig 4. Level meters measuring a reference tone at –20 dBFS RMS.

Maybe the mastering engineer even applies an additional arsenal of digital weaponry. If so, she does it blindly, because there are no meters to indicate how polluted the end result will be. Certainly, sample counting for level measurement is of no use anymore, because elementary rules of digital audio are not observed - but what are the alternatives?

Fig 4 shows that digital Sample and Signal level meters may both be calibrated for a meaningful indication of a reference tone, typically at -18 or -20 dBFS RMS depending on local standards [4, 5, 6, 7].

Fig 5 shows the reproduction demands put on studio equipment when dealing with commercial CDs. Even professional equipment may run out of voltage or current drive capabilities in their analog circuitry if peak level is dancing well above +24 dBu.

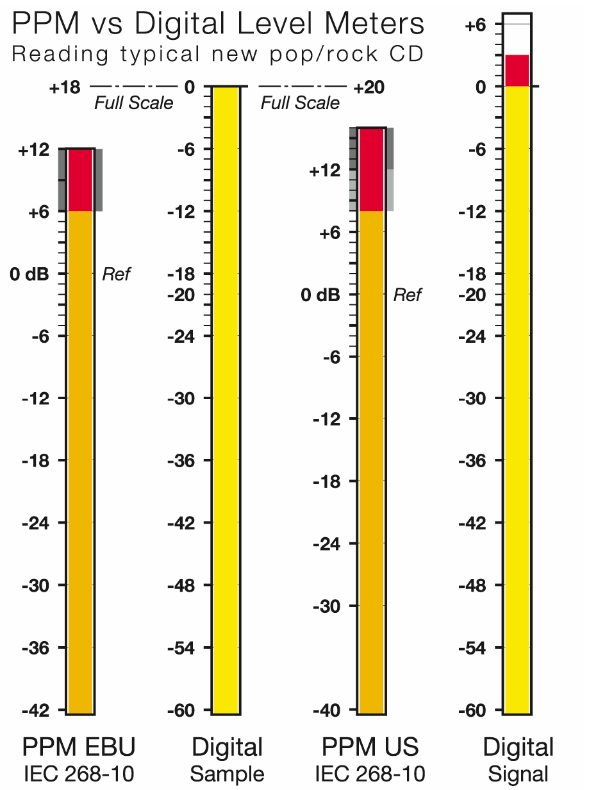

Fig 5. Typical hot CD level reading.
Note 0 dBFS+ indication on the Signal meter.

2.1. A new Peak Level Measure

In the new report from AES SC-02-01, supporting ITU loudness and peak level meter investigations [8], an algorithm for estimating the true-peak level within a single channel linear PCM audio signal is described like this: "True-peak level is the maximum (positive or negative) value of the signal waveform in the continuous time domain; this value may be higher than the largest sample value in the 48 kHz time-sampled domain.

The algorithm provides an estimate for the signal as it is, and, optionally, as it would be in the event that some downstream equipment were to remove the DC component of the signal. Optional mild high frequency pre-emphasis in the peak measurement signal path can enable the algorithm to report a higher peak level for high-frequency signals than is actually the case. The purpose for this is that the phase shifts of subsequent signal processing stages (such as Nyquist filters) could cause growth of high frequency signal peaks, and in some applications this feature could be useful to provide further protection from downstream clipping."

At the 121st AES convention in San Francisco, we can celebrate the release of this much awaited Digital Signal Meter specification draft. The next sections contain information about why it is needed.

2.2. Loudness Control in Broadcast

Current broadcast level standards are of little help to prevent multiple audio segments from ending up with very different apparent loudness at the consumer. Commercials, music CDs, digitally ingested material and file transfers are typically the most ambiguous to handle, and often turn out louder than other sources.

Part of the consistency problem, obviously, is that programming currently is judged by a peak level measure, thereby making material with a low dynamic range appear louder. We have witnessed precisely the same level mishap in pop CD production.

Broadcast levels can be set more consistently using a perceptually based measurement system, a loudness meter [9], in combination with peak level detection to avoid electrical overloads.

In search of a realtime loudness measure which reacts as quickly as the listener does, it has proven useful to adopt a sample/signal ratio into the loudness equation, in order to better control digitally ingested and file transferred material.

Tricky or "bit-stacked" material is quickly identified this way, and may be subjected to an extra stringent loudness inspection or level attenuation.

3. EXPERIMENTS AND RESULTS

In this section, hot level distortion problems are sought quantified through various measurement and listening experiments.

3.1. DA Conversion

The output of a D to A converter normally contains a lowpass filter in order to remove the images of the baseband spectrum occurring periodically around each integer multiple of the sampling frequency.

In previous experiments, we haven't found a single professional or consumer CD player that doesn't significantly distort when subjected to intrinsic level above 0 dBFS [10, 11]. A typical example is shown in *Fig* 6. Note the apparent distortion of the perfectly legal +3 dBFS sine wave, which in the case of this particular player amounts to >10%, a typical value for the players investigated.

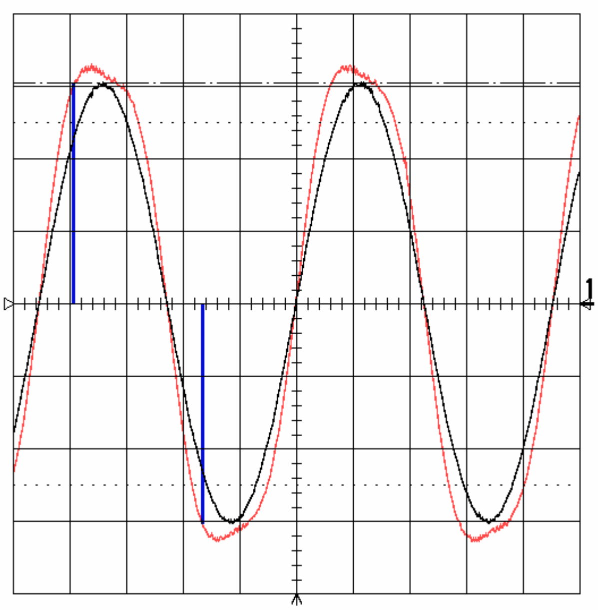

Fig 6. Sine waves reproduced by NAD512 CD player analog out measured with LeCroy 9350A.
Black curve: Intrinsic level = 0 dBFS
Red curve: Intrinsic level = +3 dBFS
Blue line: Sample position (0 dBFS)

Obviously, a D to A converter needs headroom in both the over-sampling filter, and in the subsequent analog stages, including adequate voltage swing and current drive capabilities.

3.1.1. DA Conversion Listening Tests

For a presumably linear system like CD, a simple subtractive method was developed to *listen* to the artifacts described. To get a better idea of the size of the hot level problem, we found it important to verify errors sonically, rather than relying solely on constructed signals, and FFT measurements. Hearing the artifacts, and referencing them to the music signal, was an ear-opener.

In vivo measurements and error signal listening was performed using the setup of *Fig* 8, [2]. The left channel of a hot commercial CD is copied to the left channel of a new test CD. The same signal is copied and time aligned to the right channel also, but shifted 1 bit down in level (approximately 6 dB). The right channel of the commercial CD was not used.

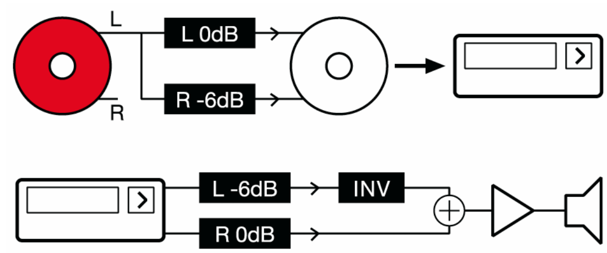

Fig 8. Listening for 0 dBFS+ headroom problems.
Commercial CD is red, test CD is white.
CD player (DUT) shown as square box.

The outputs of the DUT were level calibrated using fully correlated and phase aligned tone and noise tracks at -6 dBFS and -12 dBFS on the CD. Before summing the channels, a phase reversal and enough gain to completely cancel the output was applied to the left channel. When replaying the realworld music material, nonlinear discrepancies between the channels can be readily heard and measured.

It is assumed that headroom exhaustion problems on the left output is generally responsible for this outcome. It was also evident that some CD players were worse than others. Many devices exhibit a prolonging effect every time 0 dBFS+ is hit even briefly, thereby making a short transient overload worse than it otherwise might have been. The cause for the sustained distortion (often 150-600 ms) is believed to be analog circuitry latch-up and/or recursive filters.

The subtractive tests have been a useful supplement to measurements when assessing the prolonging effects of short 0 dBFS+ peaks, but offer less help in directly determining the possible listener fatigue consequences.

3.2. Sample Rate Conversion

Sample rate conversion (SRC) is almost exclusively integrated with other types of processing, such as in a hardware processor or in an audio editing software package. In the hardware case, a special purpose chip is often used instead of the general DSP(s).

Two of the most common SRC chips have been tested (Analog Devices AD1890 and Crystal CS8420), both of which are used e.g. in some equipment from TC. A typical example of a measurement is shown in *Fig 7.* A sine wave with an intrinsic level of +0.7 dBFS and sampled at 44.1 kHz was fed to a sample rate converter producing results at 48 kHz.

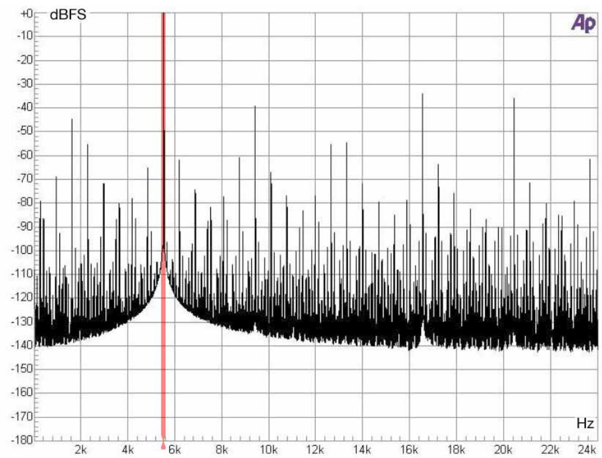

Fig 7. Sine wave at +0.7 dBFS through SRC.
CS8420 re-sampling from 44.1 to 48 kHz.
Note tone at 5.5 kHz and distortion products.

The output spectrum clearly shows the many harmonic and aliasing distortion products. The input frequency, if clipped at its original sample, would not generate aliasing, but as the rate is changed the integer relationship between signal and sampling frequency is no longer maintained, so the aliasing products occur all over the spectrum instead of just at the harmonics of the input frequency.

When SRC chips are part of designs where attenuation is not possible on the input side, which is most often the case because dual domain processing is costly, the described distortion cannot be avoided.

3.2.1. SRC Listening Tests

The same procedure followed to monitor artifacts in the DA conversion process was used with the above mentioned SRC chips.

On most audio material, "only" short glitches were heard when hot CD's were passed through. The distortion signal was typically 16-30 dB below music peak level, and contrary to the DA conversion findings, no prolonging effects were observed.

Some of the recent music releases marked "cnt" in *Table* 2, however, *did* produce more or less continuous error signals around –28 dBFS.

3.3. Digital Processing

Digital filters, pitch change and time stretch algorithms in many processors and DAW plug-ins have also been found prone to produce alias distortion when subjected to intrinsic level above 0 dBFS [12]. The same is true for broadcast processors utilizing so-called phase-rotation in order to make the audio as symmetrical (and loud) as possible.

Typical signal/distortion ratios where found to be -32 dB at +0.7 dBFS, -25 dB at +1.25 dBFS, and -18 dB when the input was +3 dBFS.

3.4. Data Reduction

Some audio systems change bandwidth dynamically. This makes it difficult to predict the exact change of waveform. In particular, perceptual coding systems - such as DTS, AC3, MP3, MP4 etc. - involve quite narrowband filters to change bandwidth and image width depending on the allocation of bits to different frequency regions, and to the correlation between the channels, [2].

	Codec	**Mode**	**Avg. data rate per ch**	**Max peak re. -6 dB**
1	MP3	Stereo HQ	160 kbps	+1.7 dB
2	MP3	Stereo HQ	180 kbps	+2.3 dB
3	MP3	Int-st HQ	64 kbps	+5.3 dB
4	MP3	Int-st fast	64 kbps	+3.0 dB
5	MP3	Int-st HQ	48 kbps	+4.7 dB
6	AAC	ITunes def	64 kbps	+3.3 dB
7	DTS	6 ch	206 kbps	+0.6 dB

Table 1. Peak levels observed in hot CD excerpts when subjected to various codecs and data rates.
MP3 is an abbreviation of MPEG-1 Layer 3.
AAC is default in Apple iTunes for import.

In order to investigate the influence of different encoding settings, various combinations of coding algorithms, data rates and coding modes were tested. Especially the encoding data rate was expected to influence the effective bandwidth of the encoded signal, as varying bandwidth is a relatively unobtrusive way of saving bits.

The test signals used were 12 excerpts from contemporary CDs, signals that are partially clipped and show levels well in excess of 0 dBFS on a Digital Signal Meter. The encoded signals were dual mono with a level reduction of 6 dB in the right channel.

The channel with full scale input signal was clipped more or less frequently. The half level channel was used to measure the size of the overshoots occurring. Results of this measurement are shown in *Table* 1.

It should be noticed that the peak level rises when applying perceptual codecs to 0 dBFS+ audio signals. Only the highest peak value occurring across the 12 audio excerpts is listed here. As seen in the table quite high overshoots can occur, depending on the coding scheme and its parameters. The size of the overshoots corresponds quite well with the encoded data rate, in that the lower data rates generate higher output peak values than the higher data rates.

Because AC3 - one of the commonly used codecs in some multichannel and broadcast applications - is not well publicly described, and was not available in a known good digital in, digital out version, it was not possible to test it in these experiments. AC3, however, is fundamentally two generations older than AAC, and conceptualized when 0 dBFS+ signals were non-existent. With typical data rates around 64-80 kbps per channel, I therefore find this particular codec likely to be equal or probably worse to the performance of MP3 in this respect.

In the real world, the findings reported in this section may not be typical. The amount of distortion generated by data reduction codecs could be underestimated, because tests were based on correlated audio. Bit allocation might get more stretched if a mixture of hot and decorrelated signals are used.

The behavior of data reduction systems should give particular reason for concern in broadcast, because stations typically rip music CD's and transfer them data reduced to a server *during ingest*, thereby potentially ending up with distorted audio in their archives.

3.4.1. Data Reduction Listening Tests

Due to the non-linear nature of perceptual coding it cannot be guaranteed that the number of bits allocated (if any) to various frequency bands of the two channels is identical. Thus, constructing a difference signal between the decoded left and right channels may not be as informative with respect to overload behavior as in the case of digital to analog converters and sample rate converters.

Therefore, for the testing of a data reduction codec, the artifacts produced by the subtraction test described in section 3.1.1. were compared against the artifacts produced when the encoded signals were attenuated by 6 and 12 dB.

With MP3 at the iTunes standard data reduction setting (128 kbps stereo, labeled "good quality"), the level of the artifacts more than scaled with the input signal level, and, more importantly, the subtraction result revealed clicks and glitches when the input signal was hot, but none when the level was 6 dB lower.

From the above it is clear that perceptually based data rate reduction coding schemes are just as critical as other digital conversion with respect to handling of 0 dBFS+ input signals. Codecs basically add another layer of uncertainty to what can be considered safe levels and tolerable amounts of digital clipping.

3.5. CD Overload Frequency

As described in the previous sections, 0 dBFS+ intrinsic level can be readily generated using artificial signals. To determine if hot signals challenging the headroom of a downstream signal path are becoming more or less frequent, a number of commercial pop and rock CDs with a spread in release dates were investigated.

	Track	Artist	Yr	H1	H2
1	Candy Shop	50 Cents	05	6	0
2	Don't Cha	P'cat Dolls	05	cnt	15
3	Incomplete	B S Boys	05	2	0
4	Hung Up	Madonna	05	cnt	20
5	It's Like That	M Carey	05	cnt	12
6	Believe Me	Fort Minor	05	cnt	2
7	Pon de Replay	Rihanna	05	cnt	0
8	Do Something	B Spears	04	18	6
9	Bad Day	D Powter	04	cnt	8
10	Fight For Your…	Beastie Bs	04*	10	0
11	Since U Been…	K Clarkson	04	cnt	10
12	Lonely No More	R Thomas	04	8	6
13	Lonely	Akon	03	20	12
14	Clap Back	Ja Rule	03	12	12
15	Work It	Missy Elliott	02	20	16

16	La Fiesta De…	A Valdez	02	2	0
17	Lose Yourself	Eminem	02	cnt	20
18	Time of My Life	Macy Gray	02	16	8
19	Don't Know Why	Norah Jones	02	0	0
20	Who's That Girl	Eve	01	cnt	4
21	Family Affair	Mary J Blige	01	20	12
22	Loved Enough	L Cohen	01	2	0
23	Don't Stop	Anastacia	01	cnt	15
24	Played Alive	Safri Duo	01	cnt	16
25	The Call	B S Boys	00	cnt	18
26	Larger Than Life	B S Boys	99	cnt	20
27	Livin' la Vida Loca	R Martin	99	12	5
28	Razor Tongue	DJ Mendez	99	17	9
29	I Got a Girl	Bega	99	cnt	3
30	Let's Get Loud	J Lopez	99	cnt	10
31	Smooth	Santana	99	20	15
32	Oye Como Va	Santana	99*	0	0
33	Avalon	Roxy Music	99*	5	0
34	Believe	Cher	98	10	4
35	Miami	Will Smith	97	17	9
36	That Don't Impr…	S Twain	98	3	0
37	Vissa Har Det	Bo Kaspers	98	1	0
38	Block Rockin…	Chem Bros	97	8	5
39	El Cuarte de Tula	B Vista	97	0	0
40	Dimples	JL Hooker	97	0	0
41	Bla Bla Bla	OK Hustlers	96	3	0
42	Bob Yu Did Yu Job	J Cliff	96	6	1
43	Where It's At	Beck	96	1	0
44	Wannabe	Spice Girls	96	5	0
45	The Only Thing…	B Adams	96	2	0
46	We'll be Together	Sting	94	1	0
47	Off the Ground	McCartney	93	1	0
48	I've Been to M'phis	Lyle Lovett	92	0	0
49	Good Stuff	B52's	92	5	0
50	Gloria's Eyes	Springsteen	92	0	0
51	Mysterious Ways	U2	91	0	0
52	S'thing to Talk…	Bonnie Raitt	91	0	0
53	Black or White	M Jackson	91	0	0
54	End of Innosence	Don Henley	89	0	0
55	Dirty Blvd	Lou Reed	88	0	0
56	Nick of Time	Bonnie Raitt	89	0	0
57	Living in America	J Brown	86	0	0
58	Graceland	Paul Simon	86	0	0
59	Two Tribes	FGTH	84	1	0
60	She Took Off My…	D Lindley	81	0	0
61	Little Sister	Ry Cooder	79	0	0

Table 2. Pop/Rock CD level over time.
Yr: Release year. * denotes remastering year.
H1: Number of <+1 dBFS incidents per 10 seconds.
H2: Number of >+1 dBFS incidents per 10 seconds.
"cnt" means continuous (>25 incidents per 10 seconds).

The CDs have not been randomly chosen. Besides from being pop or rock songs, they reflect the CD collections, and therefore to some extent the musical preferences, of me and my children.

The overload frequency clearly has been going up since around 1996. A few signs suggest it has leveled out over the last few years. Backstreet Boys have improved (level-wise), most new releases are only on the verge of overloading, while, in this investigation, Madonna and Pussycat Dolls are still mastered Y2k style.

4. CONCLUSION

Consequences of the hot level cocktail - digital clipping, deficient processing, uneducated users, and an outdated level control measure - have been demonstrated.

It is clear that current consumer, pro music, film and broadcast equipment has not been designed to handle the overly hot signals now found on most pop/rock CD releases, and it should be evident that this is *not* an audiophile issue comparable to differences in speaker wire, flavors of dither, 44.1 versus 96 kHz sampling, DSD etc.

Much music delivery today relies on data reduction at low bit-rates, but these codecs are the most prone to generate level induced distortion at the consumer. Early listener fatigue could also be a consequence; but more listening tests are needed to determine if this actually is the case when clipped material is encoded, and/or when 0 dBFS+ level is encoded.

Regardless if consumers are listening to CD or to data reduced music, they get more than they (don't) pay for: A fair amount of distortion. Our musical heritage has been badly and irreversibly affected by the described deterioration for years now, so it is about time to turn the tide and start using more intelligent routines, tools and level measurement criteria. If we believe audio quality makes a difference, and is not just an excuse for selling new gear, the audience should have a chance of getting a non-distorted experience.

Even if headroom got built into new devices it would take a long time before clean reproduction could be taken for granted. Therefore, it is suggested to correct the abuse of the digital domain and CD format on the production side by defining an updated level measure which is less easily fooled than Sample Level meters. The ambition should be to get rid of *most* 0 dBFS+ level in a standardized way, and to better detect digital clipping no matter what level it is carried out at.

The AES 121 convention is historic in this respect. Audio professionals should take a good look at the new peak level meter report from SC-02-01. A better meter, however, is still only a diagnostic tool. It has to be complemented by more informed users and better working procedures. Some suggestions can be found in the Appendix.

5. APPENDIX

What can be done level-wise to improve the audio quality in digital film and music production?

DDD production is typically not as forgiving as mixed analog and digital studios. With the exit of analog multitracks, pro audio lost its source specific level-frequency censorship inherent with emphasis recordings. With the introduction of digital dynamics processors and unconscious full scale level normalization, instead we started exploiting the top extreme of a level scale, which was never designed for this purpose.

Better consumer control over loudness level is imminent, so the absolute level advantage overly hot CDs and commercials may have now will soon be history, while their distortion is here to stay.

To reduce hot level generated distortion, we should look at level from a more educated point of view than sample by sample peak, and the whole production process with numerous semi-pro and generic computer devices should be inspected carefully [13].

5.1. Recording and Mixing

Digital clipping can happen several places during production: Inside an audio workstation (typically on the mix-buss or in the plug-ins), deliberately in various types of dynamics processors (yes, also inside a TC Finalizer or other types of digital limiters), or even inherited from digital sample libraries. Clipped kick-drums or snares may be chosen for creative reasons, but if the sample is brought close to full scale during mix or mastering, unpredictable results will occur.

Clipping in the digital domain should be practiced with caution - and for artistic reasons only by an engineer who likes alias distortion and knows what she is doing. As long as clipped audio stays clear of full scale in the final mix by a substantial margin (at least 6 dB), it won't "explode" as an accident waiting to happen until the CD meets a broadcast station or a consumer player. Lower level clipped audio will probably stay the way it is heard during the mix,, unless low bitrate codecs are employed downstream, but how does the engineer avoid making productions that reproduce with totally unpredictable results?

For a mixing engineer, advice is simple: If you mix to digital, don't peak higher than -3 dBFS on a Sample meter. Alternatively, mix to analog.

If you're often working with samples that never leave the digital domain before they are brought into a mix take the time to go through your sample library, and check at least the favorites. Normalize samples to Full Scale, and observe how a Digital Signal Meter reacts. The samples that generate 0 dBFS+ peaks could do with a caution sign in the library notes. Remember that frequent hits of the no-go-zone can put downstream equipment in a permanent state of distortion, so pay extra attention to loud and often repeated sounds in the library (kick, snare, clap, tambourine etc).

The same tests could be performed on processors and plug-ins for compression or limiting. Screen your frequently used tools, and put the offending ones in the black book. Sometimes it will be a question of not pushing a piece of gear too hard, so by checking you will be informed of its safe operating area.

5.2. Mastering

For a music or film mastering engineer, the situation is more complicated. First, check incoming mixes on an Digital Signal Meter (as opposed to a Sample Meter, the standard mastering tool). If frequent 0 dBFS+ peaks are already present, notify the mixing engineer that the sound is overly hot, and that reproduction is ambiguous. If a new mix is out of the question, at least get rid of those unpredictable peaks before doing anything more, including sample rate conversion. Several methods can be used or combined in the cleaning process:

1) Attenuate the signal in the digital domain.

Most consumer CD players and broadcast processors will survive when the signal is lowered by 3 dB, while data reduction codecs (MP3, AAC, DTS, AC3 etc.) may require up to 5 dB attenuation or more, depending on the data rate.

2) Make a detour to the analog domain.

The DA converter used should have 0 dBFS+ headroom, or its input should be attenuated by 3 dB in the digital domain. Analog processing can be added if required, and the signal made digital again. Afterwards, audio can be normalized safely to 0 dBFS.

3) Use an oversampled limiter.

Select a model that doesn't contaminate audio again when doing its job, and which leaves the signal untouched, unless 0 dBFS+ peaks are encountered.

Finally, when setting the level of the master, keep an eye on more cultivated measures than a Sample Meter. Signal Meters and histograms are valid visual tools, from which you can also learn how level has been set on other albums. *Fig* 7 shows an example of a histogram which can visualize peak distribution across an entire program or music track.

An album where level has been adjusted to perfection, peaking to 0 dBFS even on an over-sampled meter, and showing a natural distribution on a histogram, is Bonnie Raitt's "Luck of The Draw" from 1991, but of course there are many other good examples created in times, where they didn't have the measurement tools available today.

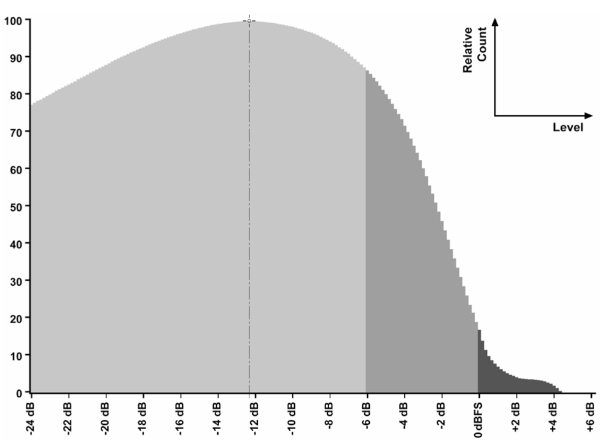

Fig 8. Histogram showing intrinsic level.

Activity above 0 dBFS is a sign of danger.
Useful information can also be extracted from the steepness of the curve, and where its peak is positioned.

Keep in mind that it probably won't be long before valid realtime loudness control is built into even consumer equipment to dynamically correct loudness audio material regardless of its origin. I believe that today's overly hot CDs will be punished by ending up weak but still distorted, while pristine recordings will keep on shining.

Please remember that louder is not better. Consistency is what counts, while sounds trying to grab our attention by being loud feel obtrusive and get deselected. If the listener wants it louder, she will turn up the volume control, but there is no distortion control to turn down once material leaves the mastering studio. Therefore, always judge little dynamics processing at a certain SPL against more processing at the same SPL.

5.3. Calibrated SPL

Calibrated speakers help restoring the ears of the audio engineer as an important tool regardless of the speaker format, and additionally defines a level at which a mix is spectrally well balanced.

Therefore, consider to start using an integrated meter/monitor solution like Bob Katz' "K-meter" [14]. Make sure, though, to install a *Signal* meter rather than a *Sample* meter.

Loudness calibrated speakers are helpful in all types of digital production, but for a large production complex, like a broadcast station or post facility, it is even more important to be able to listen at the same apparent loudness across small and large studios.

6. ACKNOWLEDGEMENTS

More than a decade ago, I had the privilege of discussing a then dawning digital level problem with the late Julian Dunn. Julian was a great source of inspiration on this audio issue as well as so many others.

My colleague, Soren H Nielsen, too has complemented and challenged my perceptual viewing angles, and provided plenty of hard evidence as well as research inputs for this and previous co-written publications.

7. REFERENCES

[1] Julian Dunn, Digital Filter Overshoot and Headroom, DA Tech paper for Audio Precision, June 2000.

[2] Soren H. Nielsen & Thomas Lund: Overload in Signal Conversion. Proceedings of the AES 23 conference, Copenhagen, 2003.

[3] Allan V. Oppenheim and Ronald W. Schafer. Digital Signal Processing. Prentice Hall, Englewood Cliffs, NJ, 1975. ISBN 0-13-214635-5.

[4] IEC: IEC 268-10, Peak program level meters, Second edition, 1991

[5] IRT: Pflichtenheft 3/6, Aussteurungsmesser, Second edition, 1998.

[6] IEC: IEC 268-18, Peak program level meters - Digital audio peak level meter, First edition, 1995.

[7] ITU-R Recommendation BS.645-2: Test signals and metering to be used on international sound-programme connections, 1992.

[8] ITU-R, WP6P: Audio Metering Characteristics Suitable for the Use in Digital Sound Production, Geneva, 2000.

[9] Esben Skovenborg, René Quesnel & Soren H. Nielsen: Evaluation of Different Loudness Models with Music and Speech Material. Proceedings of the 117th AES convention, San Francisco 2004. Preprint 6234.

[10] Soren H. Nielsen & Thomas Lund: Level Control in Digital Mastering. Proceedings of the 107th AES convention, New York 1999. Preprint 5019.

[11] Soren H. Nielsen & Thomas Lund: 0dBFS+ Level in Digital Mastering. Proceedings of the 109th AES convention, Los Angeles, 2000. Preprint 5251.

[12] Thomas Lund: Headroom in Digital Pro Audio Equipment. Proceedings of the 111th AES convention, New York, 2001.

[13] Thomas Lund: Distortion to The People, Articles for Resolution Magazine, March and May, 2004.

[14] Bob Katz: An Integrated Approach to Metering, Monitoring and Level Practices. JAES, no. 9, September 2000.