Distortion to the People

What if somebody told you that many new pop CD releases are so distorted that 16 -bit/44.1kHz is complete overkill? Read on to find out how music is being mistreated and to learn that this is not an audiophile issue comparable to differences in speaker wire, flavours of dither, or 192kHz/DSD. TC Electronic's THOMAS LUND describes the fundamentals of the hot level problem.

BECAUSE CD MASTERING incorporates no level restrictions, besides an ancient rule of how many consecutive samples are allowed at full scale, it has turned into a loudness war rather than a quest for getting decent audio quality out of a potentially good sounding media.

The mastering engineer now has a number of secret weapons to squeeze the last 0.1dB of loudness out of any song or movie. The loudness war, as we shall see, has now become so furious that equipment downstream of mastering is not able to pass or reproduce the audio material without adding significant distortion to it. If we allowed weapons inspectors into our studios, they would find weapons of mass destruction.

Engineers feeling pushed to drive the limiters harder than they think is necessary should keep the distortion examples outlined in this article to hand. Demonstrating the consequences of overly hot level through listening is more efficient than showing distortion measures in numbers, where the point easily can get lost in academic subtleties.

The principle behind measuring level in CD production is as old as the media itself. When the format was released to the public in 1982, master tapes in the pro world were analogue 1/2-inch, so the way of measuring and restricting level was designed for a completely different world than that of today. Level was and is measured purely on a peak sample by sample basis. The highest possible level in a digital encoding is called 0dBFS (or Full Scale or FSD) and the only thing to be concerned about is not to hit that ceiling with too many samples in a row.

Note magnification of the Full Scale area of Figure 1.
Event 1: 0dBFS hit of 2 consecutive samples.
Event 2: 0dBFS hit of 7 consecutive samples.

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

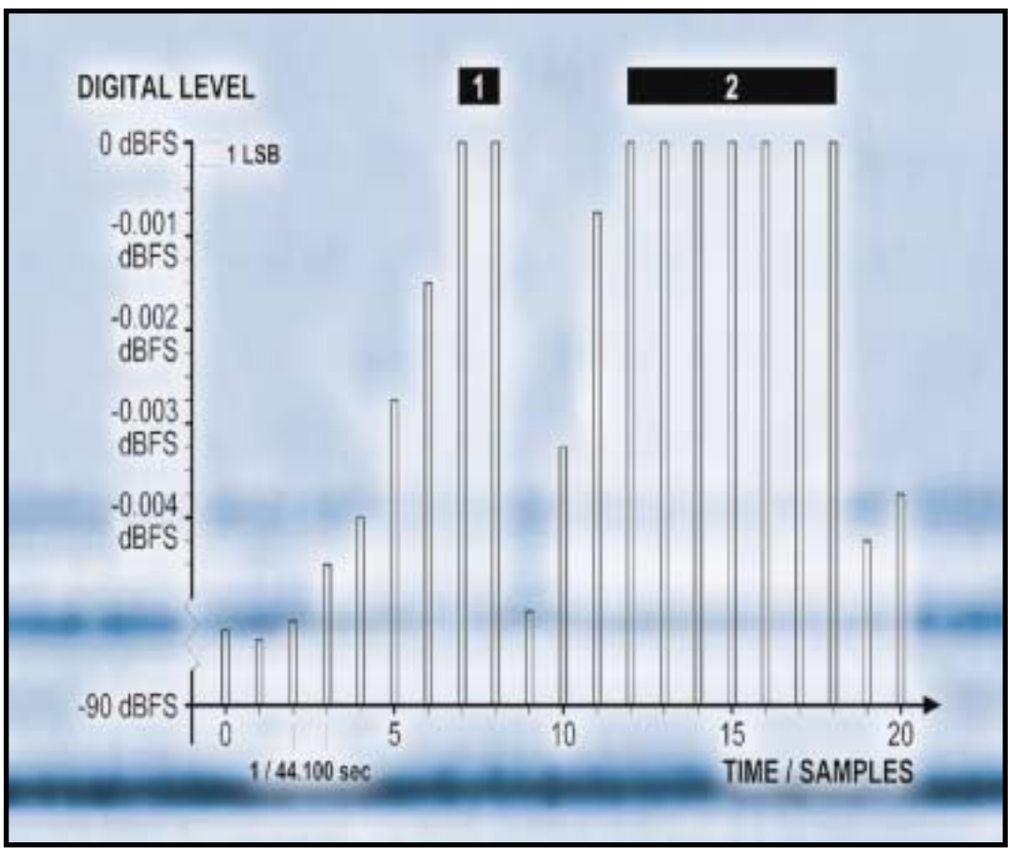

Figure 1. Example of consecutive samples on a CD.

Figure 2 shows the samples of Figure 1 subjected to a 1dB boost in level, and subsequent attenuation of 1-bit.
Event 3: 0dBFS-1 LSB hit of 6 consecutive samples.
Event 4: 0dBFS-1 LSB hit of 11 consecutive samples.

Neither event 3 or 4 will be detected as an over at the mastering plant, even though Figure 2, of course, is more distorted than Figure 1. The clipping is not detected because 0dBFS is never reached.

This leads to a discussion of how level and clipping should be measured, and if we are concerned with the digital or analogue outcome of a signal.

Even the simplest of waveforms, the sine wave, can be constructed in ways that cause analogue peaks not to align with digital peaks representing the same signal, see Figure 3. The analogue level of a sine wave at fs/6 (8kHz when sampling at 48kHz) can be up to 1.25dB above the peak level in the digital domain, while at fs/4 the discrepancy can be up to 3dB.

Dots represent digital level, curve represents analogue level.

Note: Digital peak level does not equal analogue peak level.

Put differently, sine waves can need a D-A conversion headroom of 3dB 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 or even 10dB would be needed for D-A reconstruction or sample rate conversion. In this article, inter-sample level above Full Scale will be referred to as '0dBFS+'.

If level in the digital domain is close to Full Scale, it's important to know if a reproduction chain has the headroom to accommodate all peaks. In other words, whether it has 0dBFS+ headroom or not.

In our own tests, we haven't found a single professional or consumer CD player that doesn't significantly distort when subjected to 0dBFS+ signals. When peaks reach +3dBFS, most players distort more than 10%. Many of them also display a prolonging effect: they latch-up, and take a little while to get out of distortion mode again, meaning that a period of time after a peak has occurred will also be distorted.

For a linear system like CD, we developed a simple subtractive method to listen to these artefacts and details and audio examples can be found at

http://www.tcelectronic.com/TechLibrary#LoudnessControlandMastering

The areas in which 0dBFS+ signals should cause concern at broadcast stations are in data reduction systems, 'phase rotators' and sample rate convertors.

Inside perceptually based data reduction systems the signal is filtered and quantised, often with quite narrowband filters. To investigate the influence of

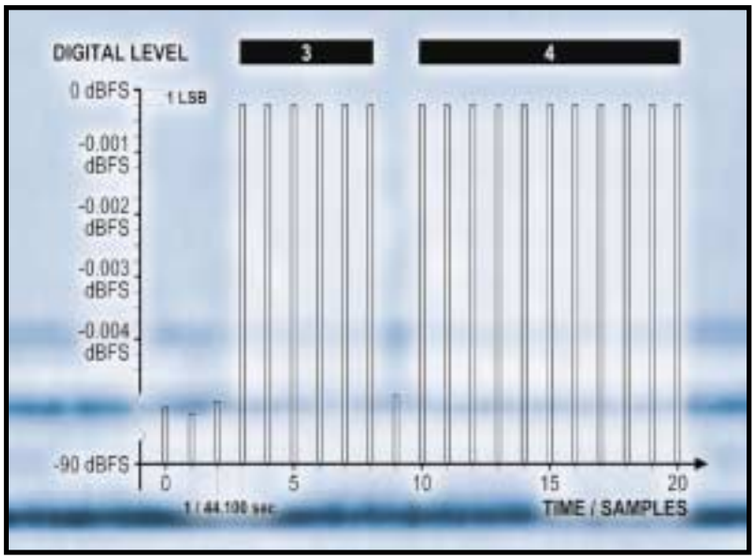

Figure 2. The samples of Figure 1 with 1dB of boost, minus 1 LSB.

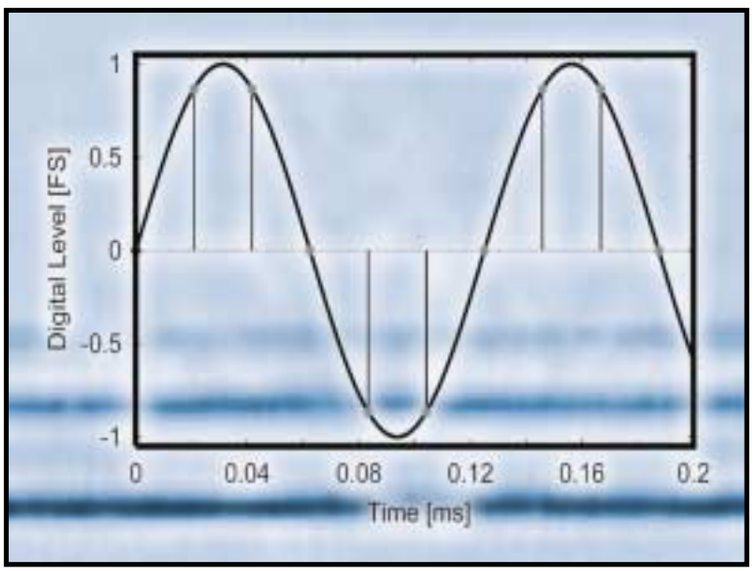

Figure 3. Digital Vs analogue level. 8kHz sine at 48kHz sampling.

different encoding settings, we tested various combinations of coding algorithms, data rates and coding modes. 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 excerpts from some of the CDs shown in Table 1. We noticed that the peak level rises when applying perceptual codecs to hot signals. Overshoots occur depending on the coding scheme and its parameters. The size of the overshoots corresponds well with the encoded data rate, in that the lower data rates generate higher output peak values than the higher data rates.

We concluded that perceptually based data rate reduction coding schemes are just as critical as other digital conversion methods with respect to handling of hot input signals.

The behaviour of data reduction systems should give particular reasons for concern because broadcast stations now typically rip music CDs and transfer them, data reduced, to a server on entry, thereby ending up with audio containing distortion in their archives.

When the archived material is aired, it is often passed through a so-called phase rotator to squeeze more average energy (RMS level) out of a medium with a fixed maximum peak value. This is done by smearing the signal and changing the phase over a broad frequency range, thus causing peaks to be spread out in time. This technique can be successful to decrease the crest factor of speech signals, but phase rotation may have the opposite effect with hot music. It will typically increase the peak value of signals with rather flat tops, which means more harm is done to the already once distorted archive material. Listening fatigue times two.

Finally, broadcast stations often employ asynchronous routers. All types of real-time sample rate convertors are prone to add distortion to 0dBFS+ signals, unless they incorporate limiters — and none of the currently available ASRC chips do so.

The clipping shown in Figure 2 generates square waves in the digital domain. A digital square wave with its steep slopes, sharp edges and flat top does not fulfill the sampling theorem. Such a waveshape in principle requires an infinite bandwidth, and the practical digital systems in use at present are not even of high bandwidth. The results are therefore aliasing — a perceptually unpleasant artefact — and, if the clipping happens close to Full Scale, 0dBFS+ peaks.

To determine if hot signals challenging the headroom of a downstream signal path are becoming more or less frequent, we investigated a number of commercial pop and rock CDs. Table 1 shows the list of tracks and the measurement results.

The table signifies a trend towards more and more level maximisation and digital domain clipping. 0dBFS+ signals are becoming commonplace, so all pro and consumer equipment should be able to process and pass them.

Many golden ears in the audio industry believe that the best sound in pop and rock music was produced between 1982 and 1995. Despite higher resolution in convertors and DSP, lower jitter and probably a better overall understanding of digital media, we seem to be on a declining rather than inclining sound quality slope even though people buying records and films may not be conscious of it.

So, if the public doesn't care, why should we? Because pride in our industry, craftsmanship and the conservation of musical talent tells us to be concerned. If we believe audio quality makes a difference, and it is not just an excuse for selling new gear, the audience should have the chance of getting a non-distorted experience.

There may even be a selfish motive for correcting the situation. Could there be a correlation between CD sound quality, record sales, and sales of good hi-fi going down? Could horrendous amounts of distortion in CD players, radio broadcast and data-reduced formats such as MP3 be one of the reasons?

We believe that current production practices generate increased distortion, and there is a reason to believe that early listening fatigue can be one of the consequences. We ought to give this somewhat technical explanation of what is happening commercially to pop music a chance, and stop producing music that detonates as rubbish the moment it hits a broadcast station or a consumer reproduction device.

Music pushed above a certain limit, exhibits this behaviour, but in mastering studios the effect can neither be seen, nor heard, because we are stuck with ancient rules of counting samples. While the Sony PCM 1610/1630, DMU-30 meters and DTA-2000 analysers were good designs, and perfectly suited for the era they were designed for, the level control principles they use don't take digital processing into account, and therefore are not adequate anymore. Maybe the biggest advantage of SACD is an update of that procedure, rather than the format's improved audio bandwidth and resolution.

My next article about 0dBFS+ signals will describe the measures that can be applied in production and mastering studios to avoid downstream distortion.

Until then: stop counting samples. ■

	Track	Artist	Year	Max Digi	Hot Spots 1	Hot Spots 2	Rating
1	Lose Yourself	Eminem	2002	0.0	>25	>25	4
2	Time of My Life	Macy Gray	2002	0.0	16	8	3
3	La Fiesta De Amadito	Amadito Valdez	2002	0.0	2	0	1
4	Don't Stop	Anastacia	2001	0.0	>25	15	4
5	Played Alive	Safri Duo	2001	0.0	>25	16	3
6	The Call	Backstreet Boys	2000	0.0	>25	18	4
7	Livin' la Vida Loca	Ricky Martin	1999	0.0	12	5	3
8	Razor Tongue	DJ Mendez	1999	0.0	17	9	4
9	I Got a Girl	Lou Bega	1999	0.0	>25	3	4
10	Let's Get Loud	Jennifer Lopez	1999	0.0	>25	10	4
11	Smooth	Santana	1999	0.0	20	15	4
12	Oye Como Va	Santana	1970/99	0.0	0	0	1
13	Avalon	Roxy Music	1982/99	0.0	5	0	2
14	Believe	Cher	1998	0.0	10	4	2
15	Miami	Will Smith	1997	0.0	17	9	3
16	That Don't Impress Me...	Shania Twain	1998	0.0	3	0	2
17	Vissa Har Det	Bo Kaspers	1998	0.0	1	0	1
18	Block Rockin' Beats	Chemical Bros.	1997	0.0	8	5	2
19	El Cuarte de Tula	Buena Vista SC	1997	-0.2	0	0	1
20	Dimples	John L. Hooker	1997	0.0	0	0	1
21	Bla Bla Bla	Ostkyst Hustlers	1996	0.0	3	0	2
22	Bob Yu Did Yu Job	Jimmy Cliff	1996	0.0	6	1	2
23	Where It's At	Beck	1996	0.0	1	0	1
24	Wannabe	Spice Girls	1996	0.0	5	0	2
25	The Only Thing...	Bryan Adams	1996	0.0	2	0	2
26	We'll be Together	Sting	1994	-0.2	1	0	1
27	Off the Ground	Paul McCartney	1993	0.0	1	0	1
28	I've Been to Memphis	Lyle Lovett	1992	-0.9	0	0	1
29	Good Stuff	B52's	1992	0.0	5	0	2
30	Gloria's Eyes	B. Springsteen	1992	0.0	0	0	1
31	Mysterious Ways	U2	1991	-0.1	0	0	1
32	Something to Talk About	Bonnie Raitt	1991	-0.9	0	0	1
33	Black or White	Michael Jackson	1991	-0.2	0	0	1
34	The End of the Innocence	Don Henley	1989	-2.2	0	0	1
35	Dirty Blvd	Lou Reed	1988	-0.2	0	0	1
36	Nick of Time	Bonnie Raitt	1989	-2.1	0	0	1
37	Living in America	James Brown	1986	-2.6	0	0	1
38	Graceland	Paul Simon	1986	-3.4	0	0	0
39	Two Tribes	Frankie Goes...	1984	-0.7	1	0	1
40	She Took Off My Romeos	David Lindley	1981	-1.9	0	0	1
41	Little Sister	Ry Cooder	1979	-8.7	0	0	0

Table 1. Measure of encoded peak level on pop and rock CDs.

Explantations to Table 1

The column 'Max Digi' shows the maximum repeated digital peak level as encoded on the CD. 'Hot Spots 1' shows an estimated typical number of occurrences of level between 0dBFS and +1dBFS per 10 seconds. 'Hot Spots 2' shows an estimated typical number of occurrences with level exceeding +1dBFS per 10 seconds.

The column 'Rating' is an evaluation of the level practices used in mastering, and indicates the track's susceptibility to reproduction distortion:

0 Low level. Not full use of CD dynamic range. No problems besides reduced S/N ratio during reproduction.

1 Well aligned level. Fully used CD dynamic range, and no problems to be expected during reproduction.

2 Level to the hot side. Probably no reproduction distortion.

3 Hot level. Distortion when passed through data reduction systems, sample rate convertors and CD players to be expected.

4 Very hot level. Obvious distortion when passed through data reduction systems, sample rate convertors and CD players to be expected.