Level and Distortion in digital broadcasting

Thomas Lund
TC Electronic A/S

CD mastering – the most seasoned digital audio discipline – has turned into a loudness war rather than a quest for getting decent audio quality out of a potentially well-sounding media. Maximum loudness is also becoming a goal in itself in new movies, so that film operators are having to turn down the replay level to avoid complaints from the audience.

In general, when audio normalization is based on peak-level detection, material with narrow dynamic range ends up the loudest. CD production not only relies on a peak-level measure (i.e. measurement scheme), it relies on a particularly bad and simplistic one, allowing massive amounts of distortion to be generated downstream of the studio in data-reduction systems and consumer equipment.

The purpose of this article is to justify and recommend more fitting ways of measuring and controlling the audio level in digital broadcasting than looking at isolated samples or quasi-peak levels. The new ITU-R BS.1770 standard, specifying long-term loudness and peak-level detection, is evaluated and a *centre of gravity* approach to loudness control is suggested. Metadata associated with Dolby AC3 is shown to be insufficient at tackling the level and distortion issues across broadcast platforms, while legitimate control practices may be derived more cheaply and without ambiguity using statistical descriptors and real-time metering derived from BS.1770.

Listener requirements

According to a study that evaluated algorithms to measure perceived loudness, consumers were found to have a distinct Dynamic Range Tolerance (DRT) specific to their listening environment *(see Fig 1)*.

The DRT is defined as a Preferred Average window with a certain peak-level headroom above it. The average level has to be kept within certain boundaries in order to maintain speech intelligibility, and to avoid music or effects from getting annoyingly loud or soft. It was noted that listeners object more often when the dynamic range is wide, rather than when it is narrow.

Experienced audio engineers instinctively target a certain DRT profile when mixing but, because level normalization in broadcast and music production is based on peak-level measures, low dynamic range signatures end up the loudest as shown in *Fig. 2.*

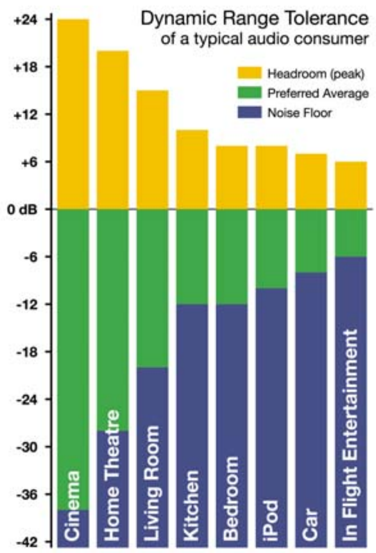

Figure 1
Dynamic Range Tolerance for consumers under different listening conditions

Engineers therefore learn to “move right” in the diagram, going for ever-decreasing dynamic range. The music industry is far to the right by now, beyond “In Flight Entertainment” in the illustration.

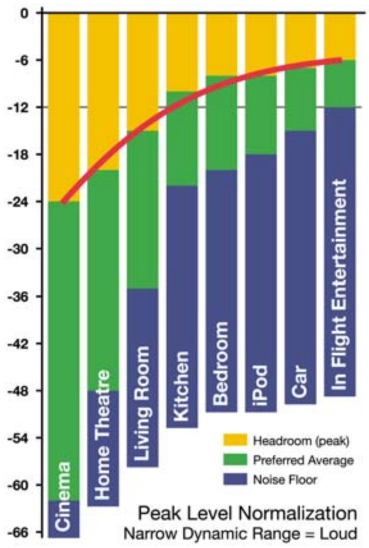

Figure 2
Peak level normalization means that narrow dynamic range material gets loud

Digital peak level

When the first digital mass media – the CD – was introduced, analogue tapes were generally used for production. During mastering, the sound was passed through analogue processing and eventually converted to digital, where the level was read sample-by-sample straight out of the AD converter. Under such circumstances, sample detection was a reliable indicator of peak level.

Today, with digital processors manipulating the samples without any respect for the sampling theorem, the situation is very different, but the way we measure the level has remained the same. Sample detection has even spread to other areas, such as broadcast and film – all now relying on a simplistic and easily-fooled peak-level measurement.

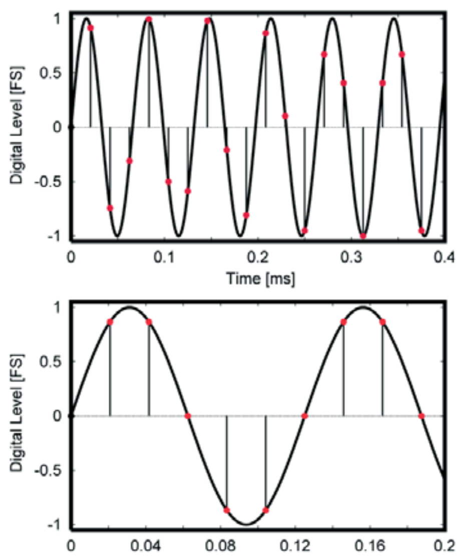

Figure 3
Digital (red dots) vs. Analogue level (black line)

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

Put differently, sine waves may need a DA conversion headroom of 3 dB for distortion-free reproduction in a linear system such as CD, while other signals can be created in the digital domain where a headroom of 6 dB or more is needed for reconstruction.

A square wave can be constructed from individual cosine components as shown in *Fig 4*. For each higher harmonic added, the peak magnitude decreases slightly. The effect of omitting the third harmonic of a (1, 3, 5) harmonic series is also illustrated. Note how the peak value is increased substantially. This kind of peaking happens when clipping is performed in the digital domain and a band-stop filter is later applied.

In this article, the resulting reconstructed or re-sampled 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.

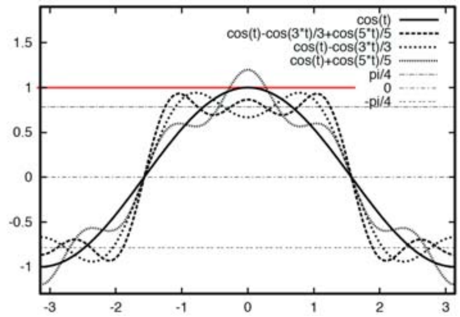

Figure 4
When digital clipping is used, the peak level increases with downstream low-pass filtering

Distortion

Distortion is the price we pay for trying to set the listener’s level control through the use of compression, limiting and clipping. However, only some of the deterioration may currently be recognized in the studio; namely the dynamic distortion. Excessive level doesn't show on the studio meters, and isn't taken into account before the signal goes through a reproduction chain.

Therefore, an additional penalty is added to an already compromised signal: unpredictable reproduction due to exhausted headroom in DA converters, sample-rate converters and data-reduction systems.

In previous papers [1][2][3][4], it has been shown how the 0 dBFS+ level generates massive amounts of distortion in CD players and data-reduction systems, and with the Sample Meter way of detecting the level, the offending signals are not recognized. A typical way of generating the 0 dBFS+ level is by *clipping* in the digital domain and afterwards attenuating the signal by a fraction of a dB, so the abuse remains undetected by a Sample Meter.

A digital square wave with its steep slopes, sharp edges and flat top does not fulfil the sampling theorem as explained in *Fig. 4.* The results are therefore ***aliasing*** – a perceptually unpleasant artefact – and, if the clipping happens close to Full Scale, ***peaks*** of 0 dBFS+. Overly fast dynamics processors can also generate various amounts of alias distortion, so a familiar name – such as a compressor or limiter – is no guarantee against invisible pollution of the signal in ways that analogue processors would not have suffered.

Distortion in CD players

While testing a selection of older and newer CD players, we didn't find a single one that doesn't significantly distort when subjected to 0 dBFS+ signals *(see Fig. 5)*. When the intrinsic level reaches +3 dBFS, 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 distortion will linger for a period of time after a peak has occurred.

For a presumably linear system like CD, we used a simple subtractive method to *listen* to these artefacts *(Fig. 6).*

You may listen to examples of these artefacts by following this link: **http://www.tcelectronic.com/EBUTechnicalReview**

There is a good reason why music lovers favour original CD releases rather than re-mastered ones. The recent detrimental use of limiting, clipping and loudness optimization on CD re-releases outweigh the positive effect of all our better converters and high-resolution processors combined. Think about it: mixes captured with non-over-sampled 14-bit converters, brick-wall analogue filters and a L/R timing offset of one sample (Sony F1) sound better than new pop tracks.

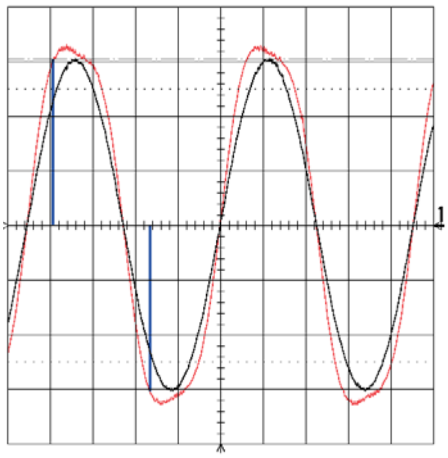

Figure 5
Sine waves reproduced by a NAD512 CD player, analogue out measured with LeCroy 9350A:
Black curve: Intrinsic level = 0 dBFS
Red curve: Intrinsic level = +3 dBFS
Blue line: Sample position for red curve (0 dBFS)

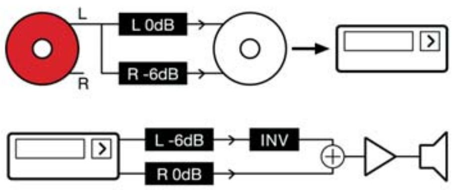

Figure 6
Listening for headroom problems in CD players:
Commercial CD is red, test CD is white;
CD player (DUT) is shown as a square box

Distortion in data reduction

Most new audio delivery systems, including digital broadcasting, make use of perceptually-based data-reduction systems. Inside these codecs, the signal is filtered and quantized, often with quite narrowband filters, so they would likely be sensitive to 0 dBFS+ level if special precautions were not taken in the design.

In order to investigate the influence of different encoding settings, various combinations of coding algorithms, data rates and coding modes have also been tested. The test signals used were excerpts from some typical off-the-shelf CDs known to contain 0 dBFS+ level. It was noticed that the peak level indeed 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. MP3 encodings at 128 kbit/s should for instance stay below –5 dBFS sample peak in order not to get exhausted frequently.

The evaluation of data-reduction codecs was recently complemented by M/S listening tests. Because data reduction isn't a linear process, the level-testing procedure used with CD players can't be directly applied. The method used instead is shown in *Fig. 7.*

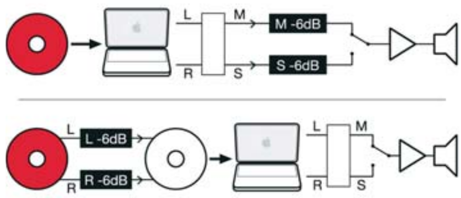

Figure 7
AAC and MP3 codec M and S listening tests
Upper drawing shows normal level encoding
Lower drawing shows attenuated encoding

It was evident how MP3 and AAC tracks at 128 kbit/s sacrifice imaging in general, and they clip audibly when subjected to 0 dBFS+ level. Data-reduction listening examples are also available if you follow the link given above for CD player distortion examples.

It will be apparent that iTunes encoding at default settings ought not to be described as stereo. "Mono with side effects" would be more appropriate. The author really only has two things against his iPod: the sound is too compromised at its default data rate, which is the same used for Music Store downloads; and iTunes attaches an "Explicit Language" warning to Always Look On the Bright Side of Life.

In conclusion, the default settings in Apple iTunes and other low-bitrate music systems lead to image distortion at any level, combined with glitches and extra distortion when the level is pushed into the 0 dBFS+ zone. *If* the consequence is early listener fatigue, *then* pop CDs mastered within the past 10 years have long-term survival odds stacked against them.

Music engineers should therefore listen through an encode-decode signal path when mastering, like film engineers have been doing for years with AC3. Data reduction isn't a free lunch, and it won't be for digital broadcast either. The behaviour of data-reduction systems should give reasons for concern, because broadcast stations typically rip music CDs and transfer them *data reduced* to a server *on entry*, thereby ending up with audio containing distortion in their archives.

A brief history of "ever louder"

1900-1950 – Mechanical recording and reproduction led to favouring Brass sections on popular recordings, and later, massive string sections.

1940-1970 – Juke boxes had fixed gain, so varigroove 45rpm recordings resulted in maximum impact.

1950 to date – US AM radio stations have a limited statutory ERP, but found that they could extend their effective coverage area (and thus advertising income) by increasing the apparent modulation using signal processing.

2010 predictions – Digital Radio has no need to restrict the dynamic range at all, and FSD will always be FSD. However, music from the 1950s onwards is very popular and so, in order to quickly recognise such a themed Radio station, software versions of the old signal processors are used.
Madonna produces an album download consisting entirely of square-waves, but processing downstream is found to actually reduce the loudness.

Dr John Emmett,
Chairman of P/AGA
(EBU Advisory Group on Production Audio)

Distortion in sample-rate conversion

All types of real-time re-samplers are prone to add distortion to 0 dBFS+ signals, unless they incorporate limiters – which none of the currently-available ASRC (Audio SRC) chips do. See the example in *Fig. 8.*

Studios and broadcast stations often employ sample-rate converters or asynchronous routers. If distortion is to be avoided, CDs need to be attenuated before rate conversion or routing. This wouldn't have been so needed, had the music industry followed the EBU R89 digital-delivery specification, calling for sample peaks below –3 dBFS.

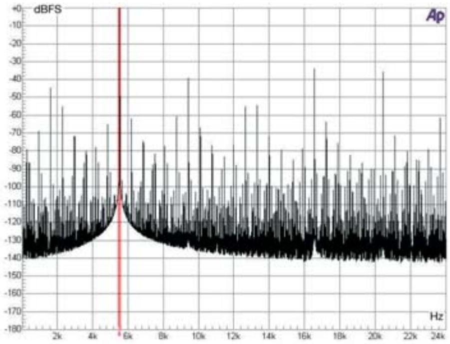

Figure 8
Sine wave at +0.7 dBFS through SRC.
CS8420 re-sampling from 44.1 to 48 kHz.
Note the tone at 5 5 kHz and the distortion products

Level meters

The insufficient level control in music CD production causes distortion to develop several places downstream of the studio. Unfortunately, there is still a trend towards more and more level maximization and digital-domain clipping on new CDs and commercials [1][2][3][4].

All pro and consumer equipment should therefore be able to process and handle 0 dBFS+ audio in a rational way. At the broadcast station, a better way of looking at level is an important part of the solution. It won't make the alias distortion go away, but at least level jumps and extra distortion may be avoided.

Sample based meters are cheap to implement and are currently widely used, but tell little about loudness, and are easy to fool. Max. sample detection is the general rule in digital mixers and DAWs. The side effect of using such a simplistic measure has become clear over the last decade, and CD music production stands as a monument to its deficiency. Sample-based peak meters require a headroom of at least 3 dB in order to prevent downstream distortion [3][4].

Quasi-peak level meters conforming to IEC 268-18 are also peak oriented, and therefore favour low-dynamic-range material when used unconsciously for normalization. The headroom needed to stay clear of distortion is 8-9 dB. However, in recent investigations we found this type of meter to be less of an open invitation to clipping than sample-peak meters.

The only type of standard-level instrument that does not display some sort of peak level is the *VU meter*. Though developed for another era, this kind of meter is arguably better at presenting an audio segment's centre of gravity. However, a VU meter is not perceptually optimized, or ideal for looking at audio with markedly different dynamic-range signatures.

Abbreviations

AAC	(MPEG) Advanced Audio Coding	**DRT**	Dynamic Range Tolerance
AD	Analogue-to-Digital	**FSD**	Full-Scale Deflection
ASRC	Asynchronous Sample-Rate Converter	**LU**	Loudness Unit
CD	A music album on a CD-A carrier at 16-bit 41.1 kHz	**MP3**	Formerly known as "MPEG Layer II"
		PPM	Peak Programme Meter
BLV	Between Listener Variability	**SLM**	Standard Loudness Measure
CLM	Consistency Loudness Measure	**SPL**	Sound Pressure Level
DA	Digital-to-Analogue	**SRC**	Sample-Rate Converter
DAW	Digital Audio Workstation	**VU**	(Audio) Volume Units
DRC	Dynamic Range Control	**WLV**	Within Listener Variability

Loudness meters and ITU-R BS.1770

Unlike electrical level, *loudness* is subjective and listeners weigh its most important factors – SPL, Frequency contents and Duration – differently. In search of an "objective" loudness measure, a certain Between Listener Variability (BLV) and Within Listener Variability (WLV) must be accepted, meaning that even loudness assessments by the same person are only consistent to some extent, and depends on the time of day, his/her mood etc. BLV adds further to the blur, when sex, culture, age etc. are introduced as variables. In the real world, unknown reproduction systems, of course, add even more blur.

Because of the variations, a generic loudness measure is only meaningful when it is based on large subjective reference tests and solid statistics. Together with McGill University in Montreal, TC Electronic has undertaken extensive loudness model investigations and evaluations *(see Fig. 9).* Each thin blue line represents hundreds of human judgments of a particular audio segment; for instance speech, guitar, yodelling, pop music, a battle scene, etc. If

Loudness algorithms

A total of ten commercially developed monophonic loudness meters/algorithms were submitted by seven different proponents for evaluation at the Audio Perception Lab of the Communications Research Centre, Canada.

In addition, Soulodre contributed two additional basic loudness algorithms to serve as a performance baseline. These two objective measures consisted of a simple frequency-weighting function, followed by an RMS measurement block.

One of the two measures, Leq(RLB) uses a high-pass frequency weighting curve referred to as the revised low-frequency B-curve (RLB). The other measure, Leq (Equivalent Sound Level), is simply an unweighted RMS measure.

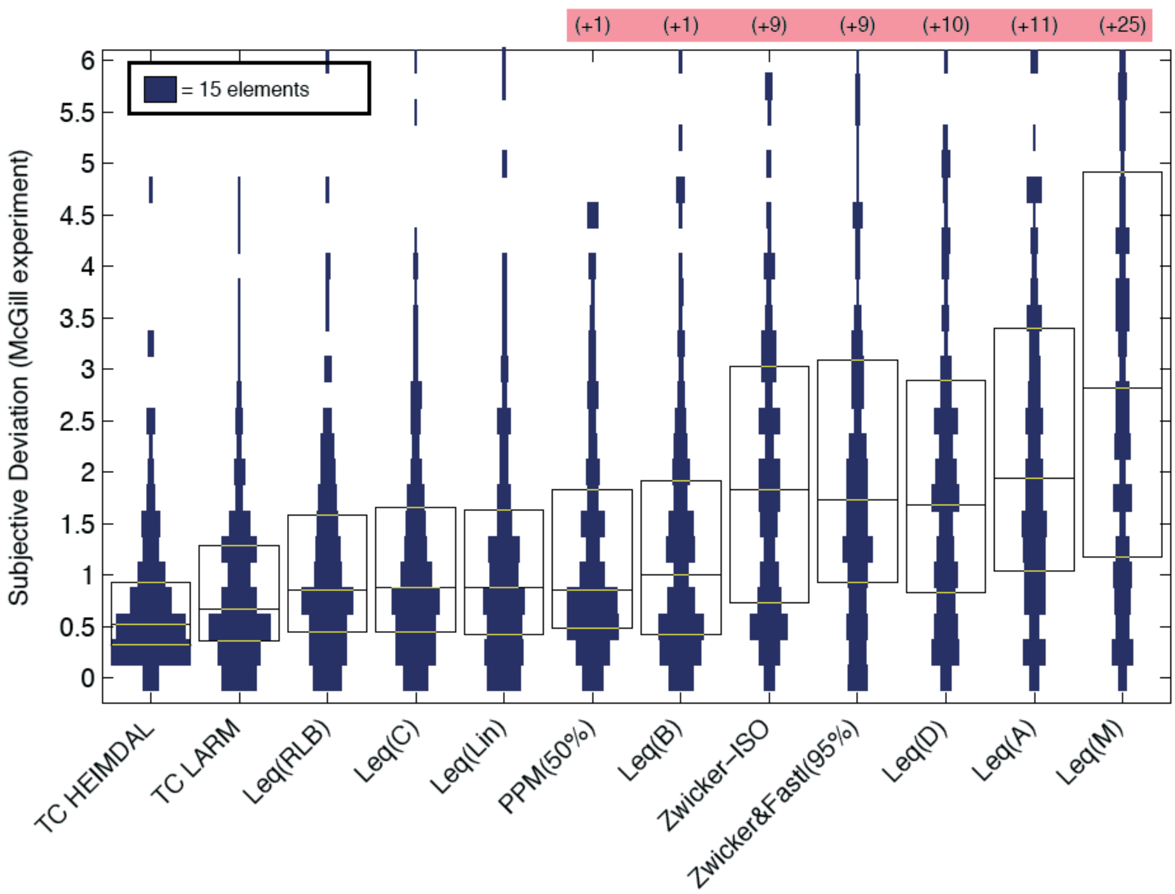

Figure 9
Evaluation of different Loudness Models using a wide range of broadcast audio material [5].
Loudness models to the left are in better agreement with human listeners than models to the right of the chart.
A red marking above the diagram indicates if a particular loudness model was ever more than 6 dB off target. The number inside the red marking denotes how many audio segments were misjudged by this amount.

many of the blue lines occur at the top of the illustration, that particular loudness model is not doing a very good job.

The results denounce a couple of Leq measures – namely A and M weighted – as being trustworthy generic loudness measures. In fact, a quasi-peak meter showed better judgement of loudness than Leq(A) or Leq(M). Even used just for speech, Leq(A) is a poor choice [5][6][7], and it performs worse on music and effects. The mediocre performance of Leq(M) might be a reason why cinema replay levels are currently out of control.

An appropriate choice for a low complexity, generic measurement algorithm has been labelled Leq(RLB). Though it better describes loudness than a quasi-peak meter [5], its performance against normal VU or slow VU has not yet been systematically tested.

In 2006, ITU-R Working Party 6J drafted a new loudness and peak-level measure, BS.1770 [8]. Concerns have been raised about the loudness part not being robust enough, because it will obviously be exploited where possible. However, with homogenous mono material, Leq(RLB) has been verified in independent studies to be a relatively accurate measure [5].

It therefore seemed justified to use Leq(RLB) as a *baseline* measure for long-term loudness, as long as room for improvement was built into the standard. However, despite being less verified, a stereo and multichannel annex using a revised weighting filter, R2LB, made it into the final standard. The multichannel extension, especially, should be used with great caution [2] until it has been properly verified. Different weighting curves typically used with Leq measuring are shown in *Fig. 10.*

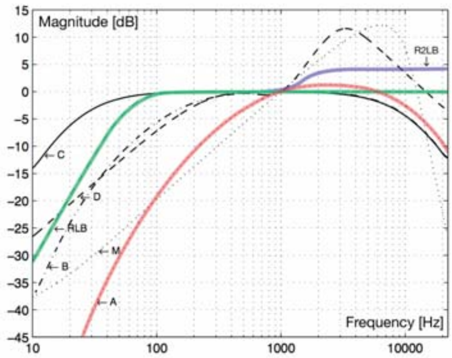

Figure 10
Weighting filters used with Leq measures
A weighting: Red, RLB: Green, R2LB: Blue

The other aspect of BS.1770, the algorithm to measure *true-peak*, is built on more solid ground. Inconsistent peak meter readings, unexpected overloads, distortion in data-reduced delivery and conversion etc. has been extensively described [1][2][3][4], so, in liaison with AES SC-02-01, an over-sampled true-peak level measure is included with BS.1770. Depending on the over-sample ratio, different maximum under-read ratios can be anticipated, e.g. 0.7 dB at four times over-sampling.

In conclusion, BS.1770 is an honourable attempt at specifying loudness and peak level separately, instead of the simplistic (sample peak) and mixed-up measures (quasi-peak) in use today.

BS.1770-compliant metering

The BS.1770 measure may be presented to the user on a traditional real-time display with certain rise-and-fall times to be specified*.* The ITU-R BS.1771 draft covers this type of meter [9]. However, loudness control is not just a matter of absolute limits. Overly loud auditory events don't have to last for long before we react or get annoyed.

TC Electronic has conducted listening tests to design a precise loudness model suited for both short-term and long-term measurements on speech, music and effects [5][6]. We were concerned that describing the level variations of an entire programme using just one number was an over-simplification, and would not provide enough information about its broadcast suitability.

To control loudness developments consistently over time, we believe the most transparent method would be to have the loudness *history* visualized already in production. A mixing engineer or a journalist should be able to identify long-term as well as short-term loudness developments.

An example of a meter showing instant loudness, history and long-term descriptors is shown in *Fig. 11.* The OBS indicator (top left) was used in our concept study to show a "possible problem": for instance, a channel completely dead, phase anomalies etc. It's a prompt to make the user try another type of view such as stereo or 5.1 bargraph meter.

The loudness history can be set at, for instance, one revolution per minute. The round display distinguishes itself from a normal PPM or VU meter, making a point that the measure is also different. Its angular reading, like a watch, means that the numbers need not be visible. Reference loudness is at 12 o'clock and can be seen on even a small picture, for instance a superimposed image.

If only a long-term average number is displayed, local soft or loud events remain undetected, and real-time work is jeopardized. If only short-term loudness is displayed, a programme's Standard Loudness Measure (i.e. its centre of gravity) is unknown and therefore how well it fits with other programmes, and across a variety of broadcast platforms.

For broadcast programming meant to be distributed over a number of platforms, it is fundamental to define its centre of gravity. With this centre point well defined, it is simple to transcode a given programme to any platform with a minimum of processing.

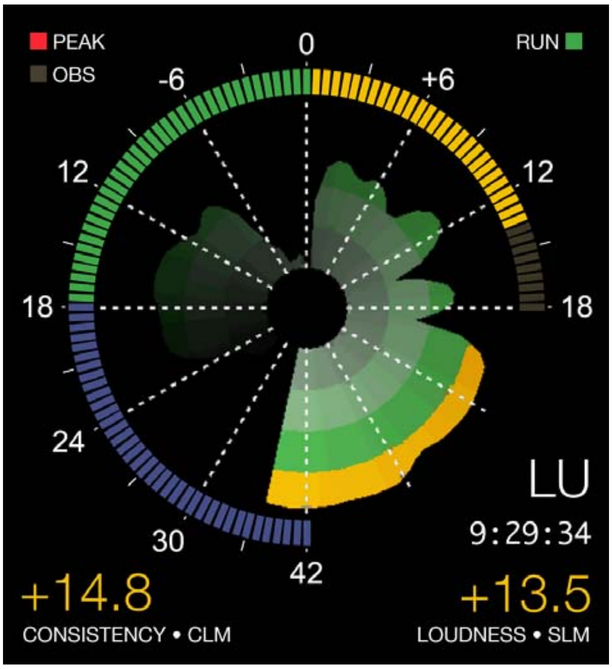

Figure 11
Real-time Loudness meter showing Current Loudness in the outer ring, History in the "radar view" and Statistical descriptors at the bottom

Delivery specifications

It has been suggested that programming should be referenced to the level of its dialogue, which to some extent works for film. However, this has bad consequences in broadcast, where mixing aesthetics between programmes may vary significantly, where dialogue not always take centre stage, where any type of sound may be disturbing, and where the consumer Dynamic Range Tolerance is lower. The sound of a phone ringing in a commercial, John Frusciante's guitar, or a fighting scene in *Pirates of the Caribbean* can all make some people reach for the remote, and should naturally have an influence on the loudness of a programme. Even if a station carries only news, documentaries or drama, it will still have accompanying sounds that can be annoying.

BS.1770 is an open standard for measuring the peak level and loudness. It may be used to enable level offsets (long-term loudness) to correspond with real-time measuring and correction (short-term loudness) across programme transitions, and across multiple broadcast platforms such as HD, SD, IP and iCast. For this to work, however, consistency has to be established between long-term and short-term corrections.

When all programming hits master control at the same Standard Loudness Measure (SLM), on-line correction for the various platforms can be centred around this value, and be as gentle, transparent and foreseeable as possible *(see Figs 12-14).* Content with different dynamic-range signatures can be seamlessly mixed this way. Production, live and external content should be aimed at the HDTV dynamic-range signature, which is a little wider than what is used for today's analogue TV delivery. The HDTV signature is automatically narrowed during transmission to fit other broadcast platforms in a predictable way which is also transparent to a production engineer.

Despite the improved loudness consistency enabled by using a centre of gravity anchor, HDTV should not be aimed at a wider dynamic range than requested by most consumers. For some stations, it may even be advantageous to use the SDTV dynamic-range signature for all HD delivery, with the possible exception of film. Long-term loudness normalization (level offsets) can be taken

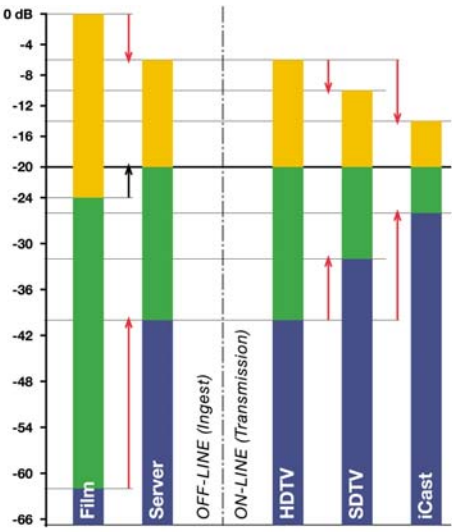

Figure 12
Film to Broadcast
Black arrows: Level offset
Red arrows: Dynamics processing

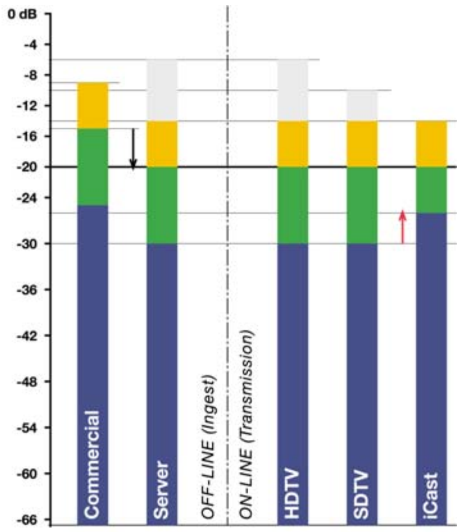

Figure 13
Commercial to Broadcast

care of during ingest or inside a file server. Under the same off-line conditions, relevant statistical information other than the SLM may be derived.

Studies of dialogue from broadcasts, films, music, commercials and sound effects have led to the conclusion that at least one more telling parameter should be used for programme-delivery specification, for instance the *Consistency Loudness Measure* (CLM). CLM is a long-term statistical measure also rooted in BS.1770. It indicates intrinsic loudness variations within a programme. A combination of the SLM and the CLM is a superior broadcast-suitability predictor than a single number such as, for instance, Dialnorm.

In the example of *Fig. 11*, the source is a hot pop track from CD, Madonna's *Hung Up*. The current loudness is at +14 LU (outer ring), the history is almost as loud (SLM=+13.5), and the consistency history shows very little variation (CLM=+14.8). It should be noted how the SLM and CLM numbers are directly operational. In the example, *Hung Up* would be broadcast-fit if offset by -13.5 dB. In this case, with a high positive CLM value, no further dynamics processing is needed to transmit *Hung Up* to any broadcast platform.

Film would typically have a negative CLM ... production material on target should read around "0" ... while a commercial, like in *Fig. 13*, would often have a positive CLM – but less extreme than hyper-compressed pop music such as *Hung Up*.

Metadata and end-listener level control

In DTV using Dolby AC3, extra information may be sent alongside the audio. Such information is known as "metadata" and is added before transmission at the broadcast station. AC3 metadata allows three end-listener level-control parameters to be set:

- **Dialnorm** adjusts the receiver's level control. The closer this setting gets to 0 dBFS, the lower the reproduction level.
- **Line-mode DRC** enables dynamic range restrictions with a wideband boost being given to low levels, and compression being applied to high levels.
- **RF-mode DRC** does the same with additional level boosts and limiting meant to be compatible with analogue TV.

The DRC settings specify a dynamic range reduction profile, with names such as “None”, “Speech”, “Music Light”, “Film Standard” etc.

The hope that decoders deployed inside consumer equipment would be able to restrict dynamic range appropriately at the end-listener has not been fulfilled, because AC3 is far from able to fill the gap between cinema and iPod. With its wideband design, pumping and other artefacts already become notable at boost or cut ratios of 6 dB [2], with much more regulation being indicated *(see Fig. 1).*

Metadata only get used if they provide clear advantages without downsides. When benefits are not obvious, the extra work and equipment needed to create metadata, the extra latency, and the potential compatibility issues they will pose over time, work against the concept. It’s no wonder why broadcasters are seeking more effective methods to control loudness than basing a station on part of a solution for just one platform.

To use AC3 metadata as the main level and range control, actually has further downsides. It is unpredictable how a consumer has his/her receiver set up, and the reproduction level can become a mess when metadata is missing or wrong. Acknowledging these problems, Dolby has introduced a loudness control solution, Dolby Volume, for manufacturers of consumer equipment. Dolby Volume is single-ended and doesn’t require metadata to function. If its complexity is high enough, it may completely disregard metadata and not worry about whether they are correct or not.

Single-ended consumer control of loudness has been a long time coming, but should be welcomed. Apple’s relatively simple solution in iTunes was among the first to offer an answer better than peak-level normalization to the general public. With Dolby Volume, and other technologies on the horizon, we can finally hope to rebuke the loudness war in music and film production.

With regard to broadcast, however, intangible consumer processing cannot be relied on. Metadata is one layer of extra unpredictability; single-ended consumer processing is another. Audio should therefore be adequately preconditioned at the station, and transmitted with fixed metadata to keep uncertainties at a minimum. Fortunately, AC3 can work well without stations having to go through the trouble of using more of its metadata extension than to signal changes between stereo and 5.1.

Best practice

Based on experiences from broadcasters around the world, consistent audio is best assured when aiming HDTV transmissions at nearly the same dynamic-range signature as SDTV. The dynamic range should be only slightly wider, see *Fig 14*, with other platforms being fed and suitably processed from the HDTV stream. The widened dynamic range is made possible by centring all programming around a long-term loudness measure derived from ITU-R BS.1770 rather than the varying degrees of peak normalization used in broadcasting today.

During *ingest* or inside the *server*, programming is offset using the long-term Standard Loudness Measure. If the Loudness Consistency of ingested material is not high enough, dynamics processing is applied to comply with the HDTV dynamic range signature. Access to a BS.1770-based loudness meter should also be provided in *production and editing*. The new loudness measure has the advantage of being understandable not only to audio experts, but to video editors, journalists and other non-specialists as well.

In *master control* and *transmission*, dynamic-range conditioning for the different platforms takes place. The HD dynamic-range

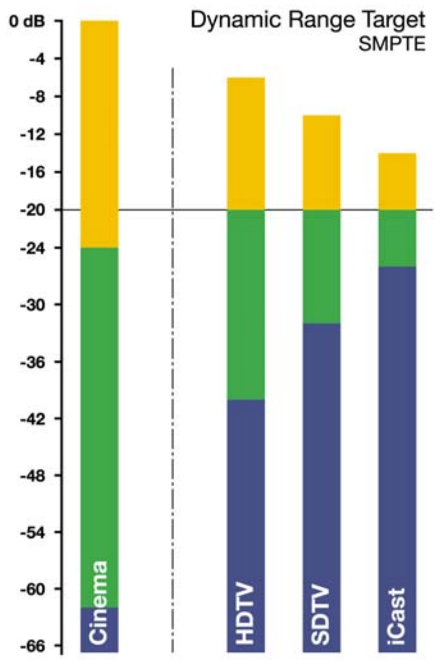

Figure 14
Suggested target dynamic range for different broadcast platforms. The Loudness target of –20 dB is used as an example

signature has already been targeted during production and ingest, so processing for this platform only plays a role when errors have been made at previous stages. Audio conditioning for other platforms is performed automatically *(Figs 12-14).*

For transmission where metadata is required, e.g. with Dolby AC3, the best practice is to keep Dialnorm, Line-mode DRC and RF-mode DRC fixed at certain values. With the gently widened dynamic range suggested here, Dialnorm should be set between –20 and –26 dBFS. A lower setting may generate more loudness at the end listener, but also with more wideband processing taking place. Therefore, it's a sign of inadequate upstream level and/or processing, if the Dialnorm number has to be lowered to keep loudness aligned with other stations. The DRC parameters too are set in a way not asking the impossible of the decode processing. The most predictable results are obtained with Line-mode DRC disabled (setting "None"). RF-mode DRC should also be disabled, or set to one of the gentle profiles, "Music Light" or "Film Light".

A content provider *delivery specification* should as a minimum describe the required SLM, and the max peak level tolerated. Peak detection is based on over-sampling, and is typically set 10 to 14 dB above the SLM. If only a sample peak measure is available, which is the rule today, digital ingest and files transfer may be louder and more distorted than expected [4]. It is helpful for a content provider to also know a CLM target in order to understand how much processing can be expected during delivery to various platforms.

Conclusions

The article has described an increasing level-maximization problem in music, commercial and film production – which causes distortion to be developed downstream of the studio, and also leads to unexpected level jumps in broadcasts. It has been demonstrated how a sample peak meter does not display the level reliably, and how peak-level normalization in general makes low dynamic-range material appear loud.

A peak measure is therefore a poor guideline in digital broadcasting, where content with different dynamic-range signatures is mixed, and more intelligent ways of measuring and controlling levels should be realized. In essence, a programme's *centre of gravity* should be used as a guideline for level offsets, plus a dependable peak level measure to stay clear of distortion. It is important that the centre of gravity calculation takes all audio content into account – speech, music and effects – and that the peak measure is based on intrinsic level rather than a sample-by-sample assessment.

The new ITU-R BS.1770 standard may eventually fulfil both criteria. Its Leq(RLB) long-term measure for mono signals was verified in 2003-04 but, later, stereo and multi-channel extensions using a revised weighting, Leq(R2LB), have not been confirmed in independent studies. At this point, the 5.1 annex should therefore only be considered an early draft.

A centre of gravity number alone, however, cannot be regarded as a "loudness meter", and therefore cannot express a programme's broadcast suitability. The number must be complemented by at least one more descriptor, namely a consistency measure; and, for use in mixing and live situations, compliant real-time metering. Neither of these essential factors were investigated in the standardization process, even though they were clearly part of the original ITU question.

Consequently, TC Electronic and McGill University, have carried out additional listening tests and experiments to extract short-term

Thomas Lund started as a musician, drifting into audio engineering while studying medicine at the University of Aarhus in Denmark. For the past 10 years, he has been with TC Electronic in Denmark, where he currently holds the title of Program Manager.

Mr Lund has written papers for Nordic organizations, the AES, SMPTE, NAB and BCA. He has participated in peer review boards and taken part in various international audio standardization activities.

functionality from the BS.1770 measure, and to derive statistical descriptors more meaningful than Dialnorm from the results.

A multifaceted solution has been described, where a combination of a real-time loudness meter and statistical descriptors may be used to streamline content delivery, ingest, production and transmission across various platforms. With: (i) long-term adjustments being applicable inside a file server, (ii) a meter being readable by a person who is not an audio expert and (iii) an automatic trickle-down routine from HD to SD to IPTV being available during transmission ... these solutions call for less time being spent per audio stream at the station. During delivery, fixed metadata may be used to keep consumer uncertainties and station workload at a minimum, while digital transmission using the AC3 format is improved thanks to light being automatically shed on its blind angles.

The extension of BS.1770 together with the procedures described in this article could help put an end to the digital production loudness war, and hopefully make the CD format its last casualty.

References

[1] S. Nielsen and T. Lund: **Overload in Signal Conversion**
AES 23 conference, Copenhagen, Denmark, 2003.

[2] T. Lund: **Control of Loudness in Digital TV**
NAB BEC, Las Vegas, United States, April 2006.

[3] T. Lund: **Distortion to The People**
Tonmeistertagung 23, Leipzig, Germany, 2004. Paper A05.

[4] T. Lund: **Stop Counting Samples**
AES 121 convention, San Francisco, United States, October 2006. Preprint 6972.

[5] Skovenborg & Nielsen: **Evaluation of Different Loudness Models with Music and Speech**
AES 117 convention, SFO, United States, 2004. Preprint 6234.

[6] Skovenborg, Quesnel & Nielsen: **Loudness Assessment of Music and Speech**
AES 116 convention, Berlin, Germany, 2004. Preprint 6143.

[7] Brian C.J. Moore, Brian R. Glasberg and Michael A. Stone: **Why Are Commercials so Loud? – Perception and Modeling of Amplitude-Compressed Speech**
JAES, no. 12, December 2003.

[8] ITU-R, Draft BS.1770: **Algorithms to measure audio programme loudness and true-peak level**
Geneva, Switzerland, April 2006.

[9] ITU-R, Draft BS.1771: **Requirements for loudness and true-peak indicating meters**
Geneva, Switzerland, April 2006.