Specifying Audio for HD

THOMAS LUND

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

ABSTRACT

The numbers of broadcast channels and platforms are going up, while the number of viewers remains the same. Digitization was expected to help streamlining audio delivery, but this has not happened yet. Confusion about peak level, loudness, end-listener requirements, formats, and the generation of metadata, has made digital broadcast an obstacle rather than the simplification needed.

This paper shows how broadcasters can take advantage of the new ITU recommendation BS.1770 to cut down on the workloads, and improve the audio quality.

Important aspects of BS.1770 are described, including novel statistical descriptors that can be derived from it. The techniques can be used to generate more precise specification documents for content providers, and to help automating ingest, production and transmission.

The paper is targeted to broadcast, music and film industry professionals.

DYNAMIC RANGE TOLERANCE

Even though DTV has the potential to carry wide dynamic range audio, this aspect is not important to the general consumer [10]. What matters most is speech intelligibility and consistency of loudness.

Consumers have a defined Dynamic Range Tolerance, DRT, specific to their listening situation, see *Fig 1*. The DRT is defined as a Preferred Average window with a specific 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 sound effects from getting annoyingly loud or soft. In situations with significant background noise, it may not even be possible to get a wide dynamic range message across - be it music or spoken - without reproduction distortion getting added, or damaging the listener's ears.

It should be noted that TV listeners more often object against audio when the dynamic range is too wide, than when it is too restricted. The only reproduction situation where a wide dynamic range is a recognized benefit to the general public is in a cinema.

Therefore, it is a main concern for the broadcaster to get speech intelligibility and consistency of loudness catered for not only on HDTV, but across all platforms.

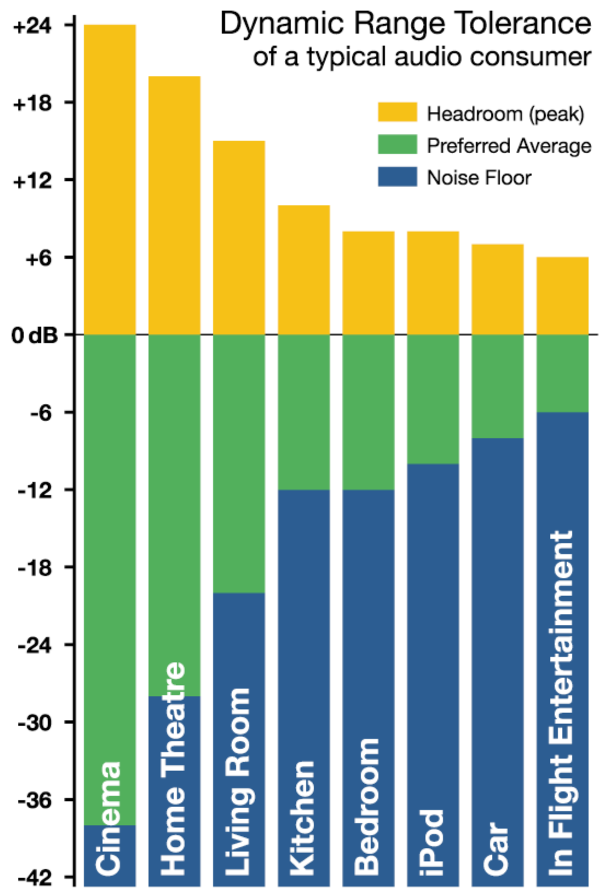

Fig 1. DRT for consumers under different listening situations.

MULTI-PLATFORM BROADCAST

The most predictable, best sounding and least costly way to simulcast on multiple platforms is to inflict a certain dynamic range signature upon each of them. Using the consumer DRT as a tool, different platforms can be designed to target distinct listening situations associated with each of them.

By means of the new ITU BS.1770 standard, this has become less ambiguous than before because a Standard

Loudness Measure for each program can be defined. Using this "center of gravity", it becomes easier to transcode a given program from platform to platform.

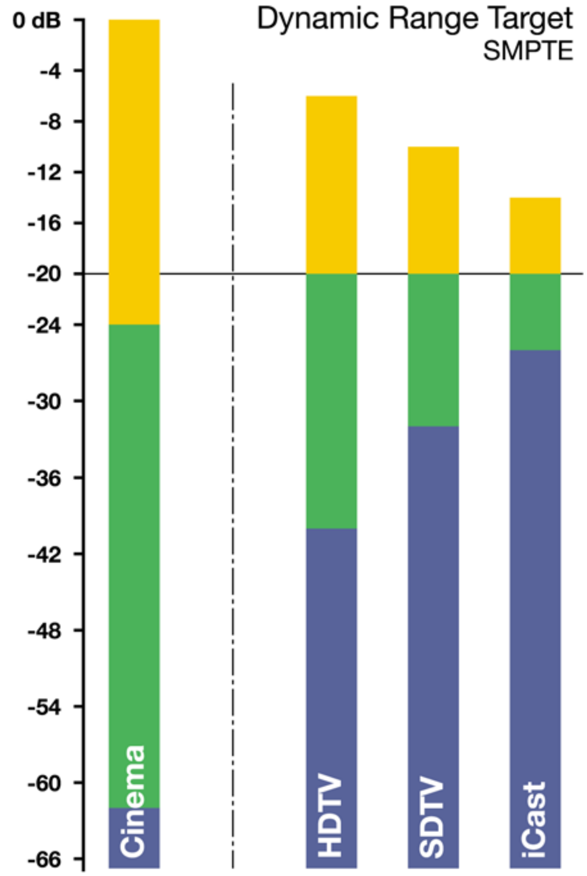

Fig 2. Target Dynamic Range for different broadcast platforms. "iCast" denotes any personal, mobile platform. IPTV would normally target a signature between SDTV and iCast.

When translating Fig 1 to a digital scale, and looking at the dynamic range target for broadcast platforms representing different listening environments, the results come out as shown in *Fig. 3*. These are the dynamic range signatures a station should aim at to give its listeners what they need with regard to dynamic range.

Because DTV features a rudimentary, consumer based dynamic range control, it can be justified to allow it a bit wider dynamic range than analog TV. It should be noted, however, that some stations won't risk a message to get lost because of unpredictable consumer adjustments, and therefore treat audio for ATV and DTV precisely the same.

One could also apply a cinema-centric rather than a broadcast-oriented angle, and lower the average level. Imagine the three platform bars in *Fig 2* shifted down 4 dB or more. However, this would be inconsistent with normal production procedures and archives, and the extra headroom would not give the station's audience any advantages, ref. *Fig 1*.

LEVEL AND LOUDNESS

When level normalization at the station is based on a peak level measure, it favors low dynamic range signatures as shown in *Fig 3*.

Quasi-peak level meters used by many broadcasters have this effect. They tell little about loudness, and also require a headroom above the maximum permitted level in order to stay clear of distortion [11, 12]. Using IEC 268-18 type meters, the headroom needed is typically 8-9 dB.

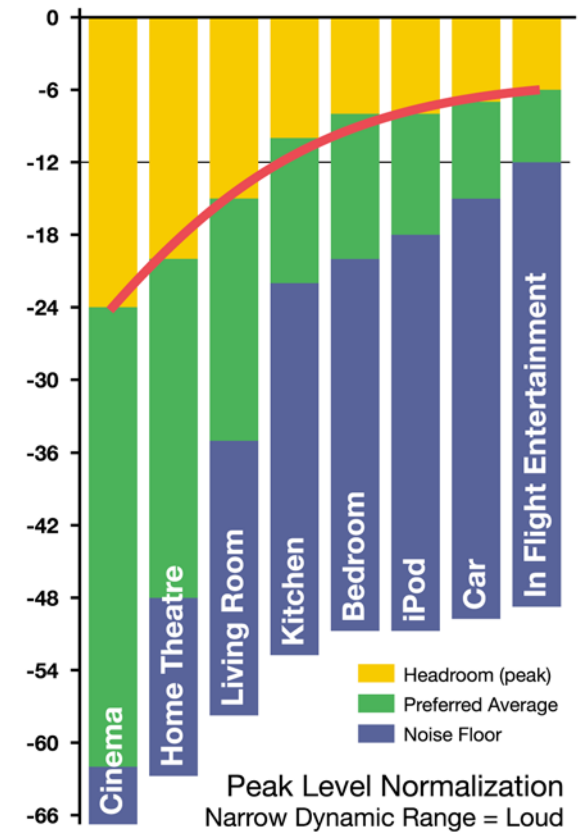

Fig 3. The effect of Peak Level Normalization.
The average level is represented by the red line, and correlates with the difference in loudness between the signatures.
Narrow dynamic range material end up louder.
This is what happened to CD.

Sample based meters are also widely used, but tell even less about loudness. Max sample detection is cheap from an implementation point of view, and therefore the general rule in digital mixers and DAWs. Unfortunately, the side effect of using such a simplistic measure has become terribly clear over the last decade, and CD music production stands as a monument over its deficiency.

Sample based peak meters require that a headroom of at least 3 dB is maintained in order to prevent distortion and listener fatigue [1-5].

The only type of standard level instrument that does not display some sort of peak level is the *VU meter*, [13]. Though developed for another era, this kind of meter is arguably better at presenting an audio segment's *center of gravity*. However, a VU meter is not perceptually optimized, or ideal for looking at audio with markedly different dynamic range signatures.

Instead of a peak meter which isn't really peak (a sample meter), a quasi-peak meter which is neither peak nor average, and an average meter which isn't optimum average (VU), the ITU set out to define an updated standard describing the issues one by one [14] and introducing the perceptual measure, loudness.

Unlike electrical level, *loudness* is subjective, and listeners weigh the most important factors differently:

- Sound pressure level
- Frequency contents
- Duration

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, her mood, the degree of attention etc. BLV adds further to the blur, when sex, culture, age etc. are introduced as variables.

Because of the variations, a generic loudness measure is only meaningful when it is based on large subjective reference tests and solid statistics. Like other technical definitions, the results also have to be repeatable.

Together with McGill University in Montreal, TC Electronic has undertaken extensive loudness model investigations and evaluations [8, 9, 10]. The results denounce a couple of Leq measures, namely A and M weighted, as 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 pick [7, 9], and it performs worse on music and effects.

An appropriate choice for a low complexity, generic measurement algorithm has been labeled Leq(RLB) [17]. Though it better describes loudness than a quasi-peak meter [9], its performance against normal VU or slow VU has not yet been systematically tested.

It should be noted that the idea of a perceptually based level calculation is not new. An aging, but respectable measure such as "CBS Loudness", is still being used with success for automated level control [6]. This model has served as a de facto reference for objective loudness measurement, in the broadcast community for decades.

ITU BS.1770

In 2006, ITU-R Working Party 6J drafted a new loudness and peak level measure, BS.1770 [15].

It has been debated if the loudness part of the standard is robust enough. As a global loudness reference, 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, taking its simplicity into account [9, 10, 17].

It seems justified to use Leq(RLB) as a *baseline* measure for long-term loudness, as long as room for improvement is built into the standard.

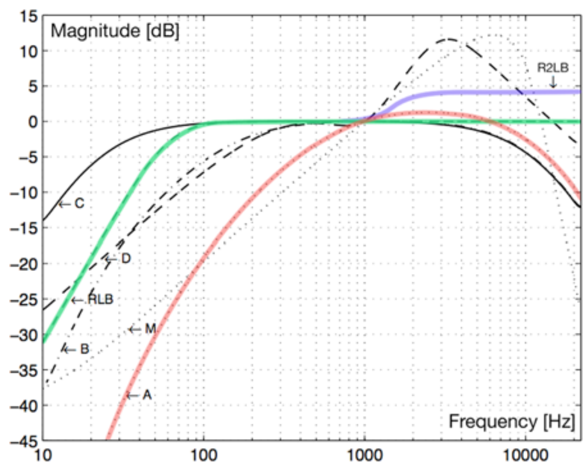

Fig 4. Weighting filters used in combination with Leq measures. A: Red, RLB: Green, R2LB: Blue (follows RLB below 500 Hz).

Note that a stereo and multichannel annex using a revised weighting filter, R2LB [18], is less verified. Especially the multichannel extension should be used with caution [10]. The different weighting curves are shown in *fig 4*.

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 [1-5].

In liaison with AES SC-02-01, an over-sampled true-peak level measure has been specified. Depending on the measurement over-sample ratio, different under-read ratios can be expected. For instance, up to 0.7 dB at four times over-sampling, but better than the 3 dB uncertainty with a sample based measure. This improved peak level measure now being included in BS.1770 will hopefully make its way back to the music industry, and help put an end to the damaging of our musical heritage due to overly hot level.

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

Though combined peak and loudness measures already exist, the time is right for an international standardization. Analog level is not an issue anymore, so one of the historically difficult variables rooted in hardware has disappeared, and the dynamic range is wider. It is now more a question of how the average and maximum level is used across a variety of platforms.

BS.1770 COMPLIANT METERING

The BS.1770 measure may be presented to the user in a traditional, realtime way with certain rise and fall times to be specified, see *Fig 5*. The ITU-R BS.1771 draft is this type of meter [19].

Meter 1 uses negative LU numbers like a typical digital sample meter. NABA has suggested this method [20], at least until a Reference Loudness has been agreed upon. The other meters have a "0" Reference point like proposed by Australia [19].

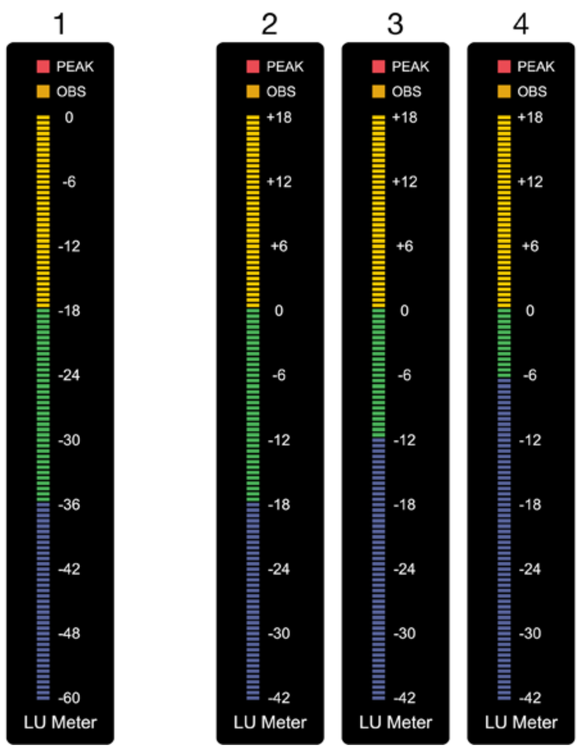

Fig 5. Loudness meters using a traditional, linear display, and color coded in accordance with Fig 2.
1: HDTV meter using negative only LU numbers.
2: HDTV meter with 0 LU at Reference Loudness.
3: SDTV meter with 0 LU at Reference Loudness.
4: iCast meter with 0 LU at Reference Loudness.

Note that the same LU meter display is expected to be valid across a number of audio formats. This raises some issues about mono, stereo and 5.1 summing, and reference tone displaying that were still open for discussion at the time of writing.

Informative combined loudness and peak level meters of course already exist, for instance the ones from Dorroughs. BS.1770 "just" offers a standardized way of measuring both numbers, and the means to tie short-term and long-term loudness consistently together.

BS.1770 AND LOUDNESS HISTORY

Loudness control is not just a matter of absolute limits. A change in loudness is by evolutionary default meant to grab our attention. Therefore, alerting auditory events don't have to last very long before we react.

TC and McGill University have conducted extensive listening tests to design a precise loudness model suited for both short-term and long-term measurements, speech, music and effects [8, 9].

We were concerned that describing the level variations of an entire program using just one number was an over-simplification, and would not provide enough information about its broadcast suitability, see *Fig 6*.

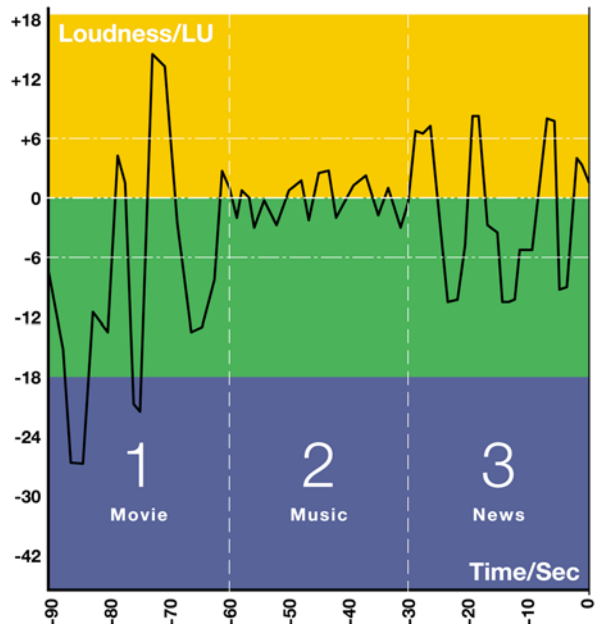

Fig 6. A one-number-per-program approach to Loudness is an over-simplification. Audio segment 1, 2 and 3 may be normalized and produce the same Standard Loudness number, but their profiles over time are clearly different.

To control loudness developments consistently over time, the most effective method is probably to have the loudness history visualized from production onwards. A mixing engineer or a journalist should be able to identify long-term as well as short-term loudness developments.

If only short-term loudness is displayed, a program's Standard Loudness Measure, its "center of gravity", is unknown, and therefore how well it fits with other programs, and across a variety of broadcast platforms.

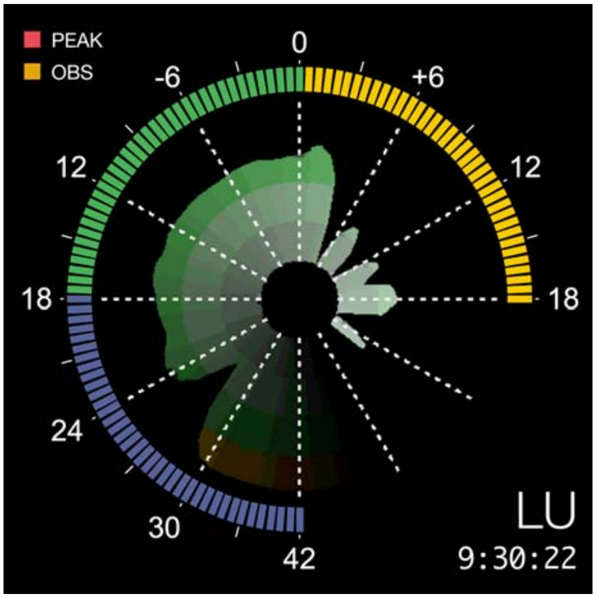

Fig 7. Loudness meter with loudness history in the "radar view" and color coded in accordance with the HDTV targets of Fig 2.

The loudness history aspect is not addressed in the BS.1771 draft meter, but an example with such virtues is shown in *Fig 7-8*. The loudness history can be set at, for instance, one revolution per minute.

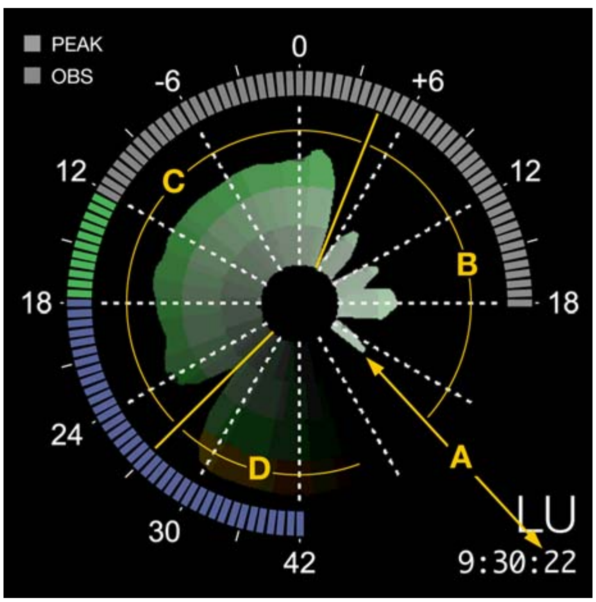

Fig 8. Loudness Meter explained. "Radar head", A, is moving clockwise, one revolution (e.g.) per minute. The outer ring shows current loudness at around –12 VU.
A: Current time, 22 seconds past the minute.
B: Current segment (Film): Low Consistency and soft.
C: Rock finished 17 sec ago. High Consistency and loud.
D: Pop finished 45 sec ago. High Consistency and very loud.

The round display distinguishes itself from a normal PPM or VU meter, making a point that the measure is also different. Its angular reading means that the numbers need not be visible, thereby enabling condensed views like the superimposed one shown in *Fig 9*. Reference loudness is at 12 o' clock, and can be identified from a small picture and at a distance.

SPEECH, MUSIC AND EFFECTS

For broadcast programming meant to be distributed over a number of platforms, it is fundamental to define its Standard Loudness, or "center of gravity". With this center point well defined, it is simple to transcode a given program to any platform.

It has been suggested to reference programming to the level of its dialog, which to some extent works for film. However, this has bad consequences in broadcast, where mixing esthetics between programs may vary significantly, where dialog not always take center 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 grab the remote, and should naturally have an influence on the loudness of a program. Even if a station is news only, documentaries or drama, it will still have accompanying sounds that can be annoying.

Fig 9. Loudness meter superimposed onto a TV picture. The outer ring shows short-term loudness, the arrow shows long-term loudness. For both indicators,12 o' clock represents Reference loudness. The red center indicator would normally be off. It signifies notable channel differences, and electrical low or high level.

The most compelling reason to identify and treat speech differently from other sources would be to guarantee its intelligibility, not to anchor the entire program around it. Loudness is just one parameter when determining the intelligibility of dialog.

The right production techniques in combination with platform specific preconditioning, however, can ensure speech intelligibility, cross-platform fitness, and the elimination of unacceptable level fluctuations. All in accordance with BS.1770, and without additional time being spent at the station.

DELIVERY SPECS: STANDARD LOUDNESS

BS.1770 is an open standard for measuring peak level and loudness. It may be used to have level offsets (long-term loudness) correspond with realtime measuring and correction (short-term loudness) across program transitions, and across multiple broadcast platforms such as HD, SD, IP and iCast. It is necessary, though, to establish a consistency between off-line level offsets and on-line correction.

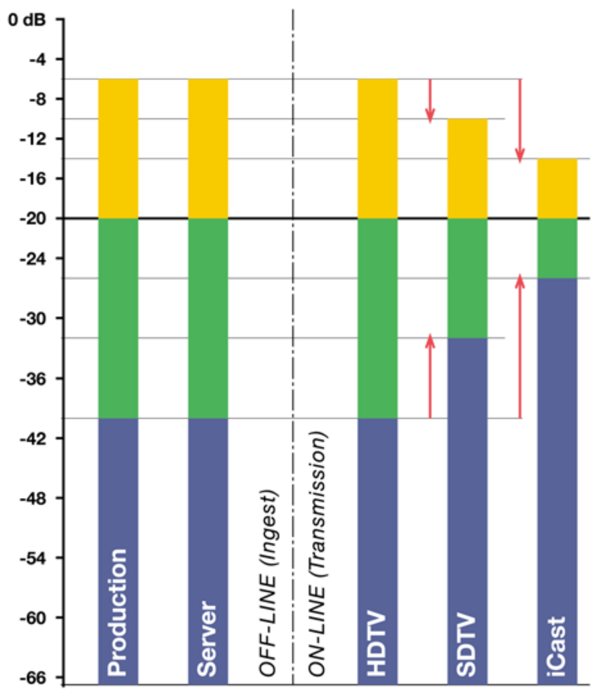

Fig 10. Level handling from Production to Broadcast.
All platforms use the same nominal SLM.
Red arrows: Dynamics processing.

When all programming hits master control at the same Standard Loudness Measure (SLM), on-line correction for the various platforms can be centered around this value, and be as gentle, transparent and foreseeable as possible, see *Fig 10-13*. 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 analog TV delivery. The HDTV signature is automatically narrowed during transmission to fit other broadcast platforms.

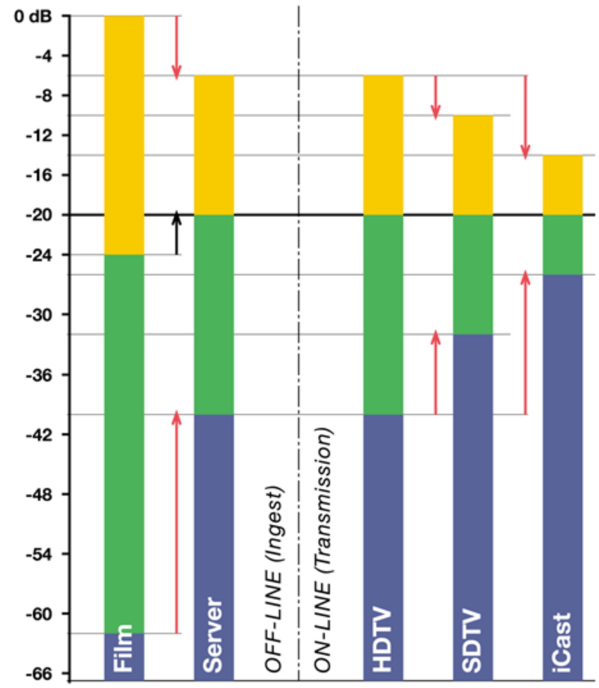

Fig 11. Level transcoding from Film to Broadcast.
First align SLM, then process as indicated.
Black arrows: Level offset.
Red arrows: Dynamics processing.

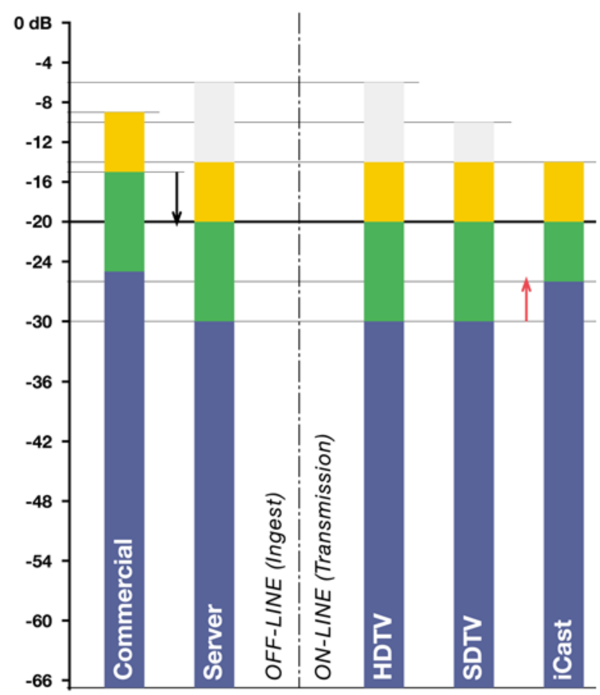

Fig 12. Level transcoding from Commercial to Broadcast.
First align SLM, then process as indicated.
Black arrows: Level offset.
Red arrows: Dynamics processing.
Gray marking: Unused headroom.

Despite the improved loudness consistency enabled through BS.1770 compliance, it is important not to aim at a wider dynamic range for HDTV than requested by most consumers, ref *Fig 1*. 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.

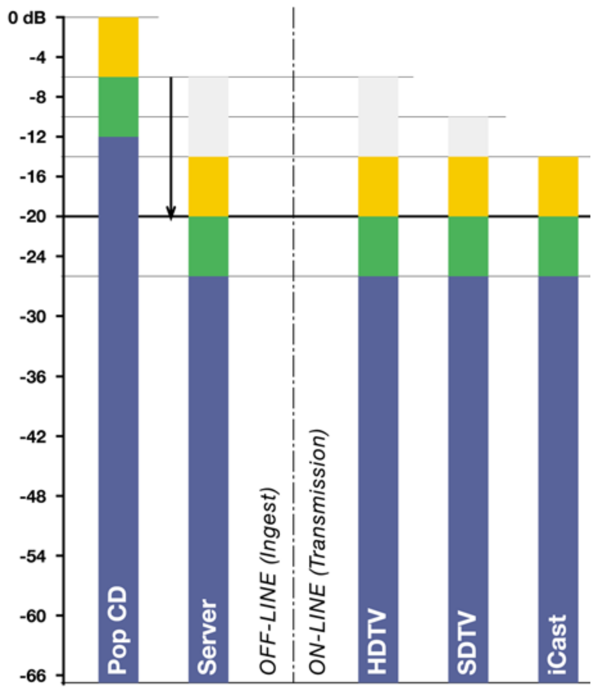

Fig 13. Level transcoding from CD Levels to Broadcast.
First align SLM. More processing is rarely required.
Black arrows: Level offset.
Gray marking: Unused headroom.

Long-term loudness normalization (level offsets) should therefore be taken care of during ingest or inside a file server. Under the same off-line conditions, relevant information besides from a program's SLM may also be derived.

DELIVERY SPECS: LOUDNESS CONSISTENCY

Studies of dialog from broadcast and film, music, commercials and effect sounds have led to the conclusion that at least one more telling parameter should be used for program delivery specification, namely the Consistency Loudness Measure (CLM).

CLM is a long-term statistical measure also rooted in BS.1770. It has been designed to indicate intrinsic loudness variations inside a program, and relate them to the HDTV range signature. A combination of the Standard Loudness Measure, SLM, and the Consistency Loudness Measure, CLM, is a superior broadcast suitability predictor to a single number such as, for instance, Dialnorm.

CLM	Description
Below -12	Very wide dynamic range, e.g. a movie. Significant processing to be expected for HD delivery.
-12 to -6	Wide dynamic range. Notable processing to be expected for HD delivery.
-6 to -2	Extended dynamic range. Moderate processing to be expected for HD delivery.
-2 to +2	On target for HD. Little or no processing to be expected for HD delivery.
+6 to +12	More uniform than needed. Little or no processing to be expected for SD delivery.
Above +12	Very narrow dynamic range, e.g. a pop CD. On target for iCast.

Table 1. The Consistency Loudness Measure, CLM.

In the example of *Fig 14*, 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). The radar view reveals that the previous segment, a battle scene from the movie Pearl Harbor, ended 15 seconds ago. Pearl Harbor shows lower loudness and less consistency, and SLM/CLM descriptors at +6.4/+1.1 confirm this.

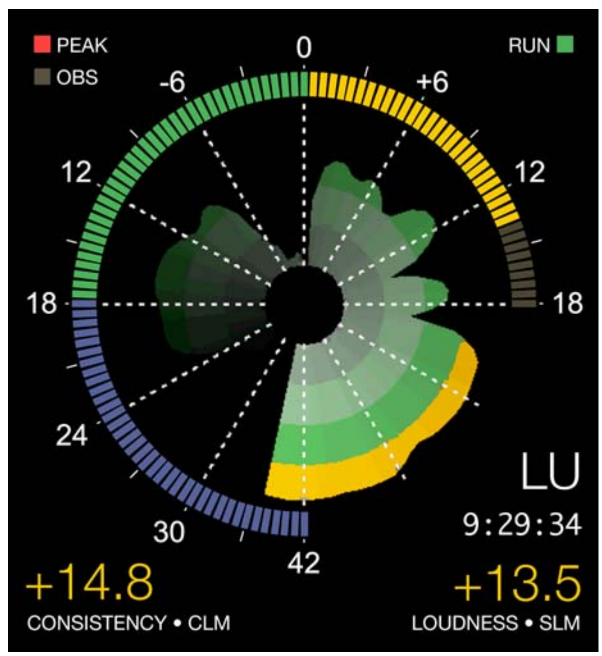

Fig 14. LU meter showing Realtime Loudness, Loudness History and Long-term Loudness SLM and CLM descriptors.

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, while the scene in Pearl Harbor needs a correction of -6.4 dB. In both cases with positive CLM values, no further dynamics processing is indicated for HD. When all of Pearl Harbor is logged instead of just the battle scene, SLM shows -4.8 and CLM -11.5, suggesting an average offset of +4.8 dB plus substantial dynamics range processing before broadcast, like the curve in *Fig 15*.

METADATA

In DTV using Dolby AC3 data reduction, provision has been made to include extra information on top of the audio itself. Such information is known as "metadata", and added before transmission at the station. AC3 metadata allows three end-listener level control parameters to be set.

Dialnorm adjusts the receiver's level control in order to keep dialog at a constant level. The closer this setting gets to 0 dBFS, the lower the reproduction level. Line-mode DRC enables dynamic range restriction with wideband boost of low level, and compression of high level. RF-mode DRC does the same with additional level boost and limiting meant to be compatible with analog TV. The DRC settings specify a dynamic range reduction profile, with names such as "none", "speech", "music light", "film standard" etc.

The hope that AC3 decoders deployed inside consumer equipment would thus be able to restrict dynamic range appropriately at the listener has not been fulfilled, because AC3 is far from able to fill the gap between cinema and iPod. With a wideband design like this, pumping and other artefacts already become quite notable at boost or cut ratios of 6 dB [10], 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, and the potential compatibility issues they may pose over time, naturally work against the concept. It's no wonder why broadcasters around the world are seeking more effective methods to control loudness than basing a station on part of a solution for one platform.

Fortunately, AC3 can work well without stations having to go through the trouble of using more of its metadata extension than changes between stereo and 5.1.

END-LISTENER DYNAMIC RANGE CONTROL

Using AC3 metadata as the main level and range control has other well known downsides than the ones described before. It is unpredictable how a consumer has her receiver set, and reproduction level becomes a mess when metadata is missing or wrong.

Acknowledging these problems, Dolby has introduced a chipped loudness control solution, Dolby Volume, to 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 if they are correct or not.

Single-ended consumer control of loudness has been a long time coming [4], but should be welcomed. Apple's relatively simple solution in iTunes was the first to offer a solution better than peak level normalization to the general public. With Dolby Volume, and other solutions to come, 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.

BEST PRACTICE

Based on experiences from broadcasters around the world, consistent audio and satisfied listeners are best assured when aiming HDTV transmission on nearly the same dynamic range signature as SDTV. For best results, the dynamic range should be only slightly wider, ref. *Fig 2*, with other platforms being fed and suitably processed off the HDTV stream.

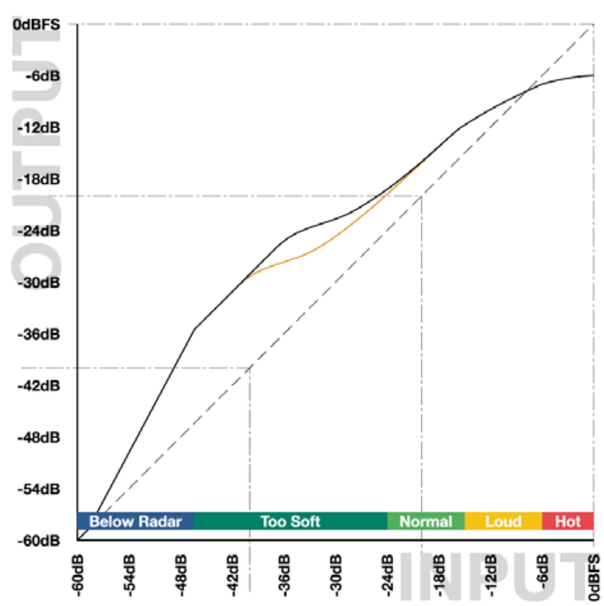

Fig 15. Example of dynamic range re-mapping of a 5.1 feature film to HD broadcast.
Black curve: Center channel.
Orange curve: L, R, Ls, Rs.

The widened dynamic range is made possible by centering all programming around a long-term loudness measure derived from ITU-R BS.1770 rather than the varying degrees of peak normalization used in broadcast today.

During *ingest* or inside the *server*, programming is normalized 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, example *Fig 15*. To precondition very wide dynamic range movies at ingest, it may be indicated to use a "telecine approach", transcoding sensitive scenes one at a time.

Access to a BS.1770 based loudness meter should also be provided in *production and editing*. A combination of a suitable realtime display, a history view and long-term descriptors has been shown in *Fig 14*. This type of display presents a target loudness and consistency measure during mixing, and informs about how much downstream processing will be applied for the production to fit various broadcast platforms. The new loudness measure also 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. HD has already been targeted upstream, so processing on this platform only plays a role when errors have been made at previous stages. Audio conditioning for other platforms is performed automatically, ref *Fig 10-13*.

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 -24 dBFS. A lower setting may generate more loudness at the end listener, but also 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 decoding processing. The best and most predictable audio 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 such as "Music Light" or "Film Light".

A content provider *delivery specification* describes the required SLM and max peak level. The peak level detection uses 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 [5]. It may also be helpful for a content provider to know a CLM target in order to understand what processing can be expected when the program is delivered to various platforms.

CONCLUSION

The paper has described how the open ITU-R BS.1770 standard can be used to streamline the flow of audio at the station by aligning content based on loudness, and implementing an automated trickle-down routine from HD to SD to IP to iCast. The "center of gravity" alignment also enables a station to mix content with different dynamic range signatures seamlessly, and without annoying level fluctuation.

Speech intelligibility and consistency of loudness are still the responsibilities of a broadcaster across all delivery platforms. Intangible consumer processing can not be relied on when dealing with such fundamental issues, and the Dynamic Range Tolerance of a consumer has to be respected.

A realtime loudness meter and powerful statistical descriptors have therefore been derived from BS.1770, and may be employed to automatically ensure loudness consistency within and between programs. The new SLM and CLM descriptors are also effective at specifying content, and make an AC3 based HD delivery system perform better by shedding light on its blind angles. Furthermore, less time needs being spent on audio issues at the station, and the tools proposed can be used effectively inside a file server, or by a person who is not an audio expert.

Over the next years, BS.1770 will have to prove itself against the currently used level measurement and control techniques at the broadcaster. The advantages of a separate peak and average measure will soon become clear, regardless if users in different parts of the world stick to their traditional reference levels or not.

In conclusion, BS.1770 is more about applying the same "center of gravity" level measure everywhere than necessarily having the most advanced loudness measure available from the start. A respectable and standardized average measure is better than any peak level normalization, and can help streamlining production, program interchange and transmission to various platforms with or without the use of metadata.

BS.1770 and the procedures described in this paper could also help put an end to the digital production loudness war, and maybe make the CD format its last casualty.

REFERENCES

[1] Julian Dunn: Measurement Techniques for Digital Audio, Digital Filter Overshoot and Headroom. Tech paper for Audio Precision. Beaverton, OR, 2001.

[2] Nielsen, S. & Lund, T.: Level Control in Digital Mastering. Proceedings of the AES 107 convention. New York, NY, 1999. Preprint 5019.

[3] Nielsen, S. & Lund, T.: Overload in Signal Conversion. Proceedings of the AES 23 conference. Copenhagen, 2003.

[4] Lund, T.: Distortion to The People. Proceedings of the Tonmeistertagung 23. Leipzig 2004. Paper A05.

[5] Lund, T: Stop Counting Samples. Proceedings of the AES 121 convention. San Francisco, CA, October 2006. Preprint 6972.

[6] Jones, B.L. & Torick, E.L.: A New Loudness Indicator for Use in Broadcasting. Proceedings of the AES 71 Convention. Montreux, 1982.

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

[8] Skovenborg, Quesnel & Nielsen: Loudness Assessment of Music and Speech. Proceedings of the AES 116 convention. Berlin, 2004. Preprint 6143.

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

[10] Lund, T: Control of Loudness in Digital TV. Proceedings of NAB BEC. Las Vegas, NV, April 2006.

[11] IRT: ARD Pflichtenheft 3/6, Aussteurungsmesser, Second edition. München, 1998.

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

[13] IEC: IEC 268-17, Standard volume indicators. First edition. Geneva, 1990.

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

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

[16] Kurozumi, K & Hoshi, H. Realtime Loudness Meter for Digital Broadcast. AES 11th regional Convention. Tokyo, 2003.

[17] Norcross, S & Soulodre, G: Objective Measures of Loudness. Proceedings of the AES 115 convention. New York, NY, 2003. Preprint 5896.

[18] Norcross, S & Soulodre, G: Stereo and Multichannel Loudness Perception and Metering. Proceedings of the AES 119 convention. New York, NY, 2005. Preprint 6618.

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

[20] NABA presenting North American viewpoints to the ITU on question R-2/6: Proposed specification for audio programme loudness. Geneva, August 2006.