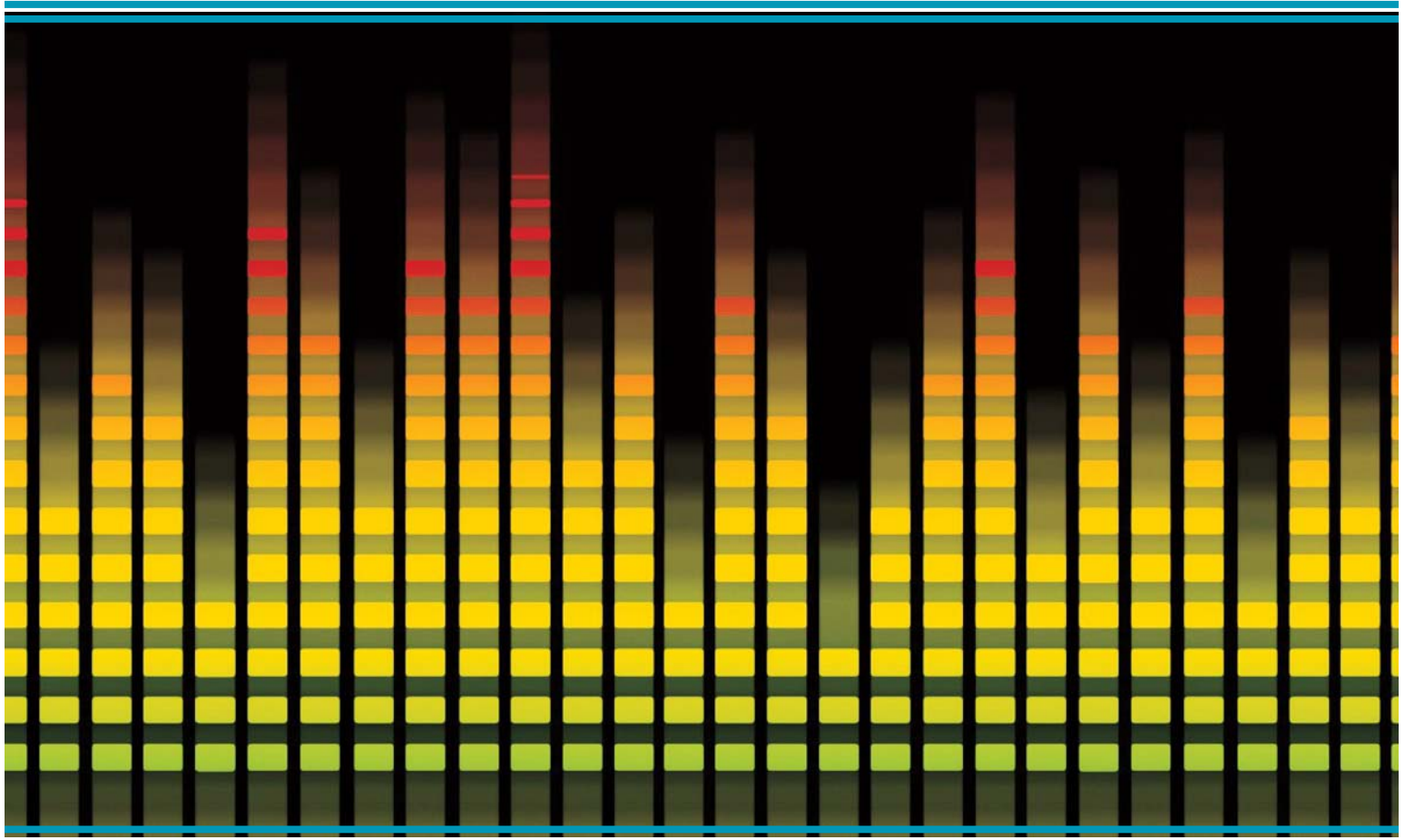


# TVTechnology

## GUIDE TO AUDIO LOUDNESS

SECOND EDITION



*Sponsored by*



A Special  
Supplement to

**TVTechnology**  
BROADCAST • PRODUCTION • BROADBAND • SATELLITE • MOBILE

**NewBay  
Media**

# THE INTERNATIONAL APPROACH TO LOUDNESS

N. America and Europe see things differently

by Lon Neumann

*Lon Neumann, consulting engineer with Neumann Technologies and a specialist in audio for DTV, considers the challenges facing U.S. broadcasters in advance of the Dec. 13 CALM Act deadline.*

The good news is that things are getting better. In many cases, formerly objectionable commercials have been brought into compliance. Things are far less annoying now than they were previously. However, there is still work to be done. In my experience, there is still a great lack of understanding afoot. Without understanding, there is not a great likelihood of proper compliance. There is also some likelihood of unfortunate unintended consequences.

There are forces at work that encourage using the easiest, least expensive, set-and-forget solutions. Such solutions may have the unintended consequence of returning us to the dark days of NTSC audio. Let's hope that's not our collective destiny. That certainly has not been the intent of the framers of our approach—quite the contrary. The intent is to deliver DTV audio with quality equivalent to cinema audio.

## TWO DIFFERENT APPROACHES

Part of the problem currently is that now two distinctly different approaches to the general problem of loudness management in DTV are in practice in the world—the North American approach and the European approach. They are significantly different.

Here in the U.S., we cite good research to support our assertion that listeners judge the general loudness of programs by the loudness of normally spoken dialogue. From the beginning of our DTV standard (ATSC A/53), there has been the explicit precept that the metadata parameter known as dialnorm must tell the truth about the level of the dialogue. That precept has had the power of law since then, even if it was never enforced until recently.

Now the Europeans have decided to go a different route. First, in many instances (most instances?), they are not necessarily encoding audio with the AC-3 codec and therefore have no metadata, and thus have no dialnorm for controlling the playout volume at the listener's home receiver. Furthermore, they have decided that their mea-

surements of loudness will average together all the soundtrack elements, not just dialogue.

The Europeans have gone on to develop a new loudness measurement algorithm that includes a system of level-gating, as outlined in the EBU R128 document as its attendant four technical documents. This system provides for excluding portions of silent and low level content from the measurement of loudness. Most people agree that it would be wrong to include the measurement of silence in an overall average measurement of loudness. Clearly, that would skew the results towards an undesirable result.

Having two distinctly different approaches to the problem has tended to confuse people. If the European approach is the most recent development, wouldn't it make sense that it's the most highly developed and most up-to-date? Not necessarily.

This backstory is now also complicated by a couple of developments here stateside—Annexes J and K to the ATSC A/85 document. These annexes were late additions to the standard that address the situations of loudness management without AC-3 encoding and also the loudness measurement of short-form content, such as commercials and other interstitial content.

Annex K addresses the situation where, in the case of codecs other than AC-3 (usually MPEG 1 Layer 2 and AAC), there is no dialnorm to control playout volume at consumers' receivers. The guidance here is for the operator to work towards a target loudness with long-form content (-24 LKFS is recommended). Then the short-form content must be at the same loudness.

Annex J now provides the explicit guidance that the measurement of loudness in short-form content be for the duration of the content. In other words, it is not intended to isolate the dialogue as the anchor element.

This has presented a confusing situation to the industry. There's now one approach for long-form content and a different approach for short-form content. For long-form, the task is still to measure the loudness of dialogue



Lon Neumann

as the anchor element. For short-form, the task is to average the measurement of all content for the duration.

To its credit, Dolby (and others) has proposed an elegant solution. In my experience, it has long been the case that Dolby's Dialogue Intelligence has provided the most consistent results for measuring loudness of dialogue in an automated workflow. This automated system applies seven different tests to audio content and reliably finds the portions of content that are normal dialogue. The result thus derived is then used to control the loudness measurement by dialogue-gating, rather than level-gating as per the EBU.

It's all very elegant and works very well. The only problem has been that it was proprietary to Dolby and thus was not included as a requirement in the ATSC standard. So, Dolby has decided that, for the betterment of the Industry, Dialogue Intelligence must be released to the world free of any royalty charges. That is now the case.

Now developers everywhere can implement Dialogue Intelligence without royalties due to Dolby. I, for one, strongly encourage all to do so. There is no longer any good reason to not implement Dialogue Intelligence. In the ATSC world, the directive will remain that it is the loudness of dialogue that needs to be encoded into dialnorm. In an automated workflow, there is no better way to do that than with Dialogue Intelligence.

But what about that short-form content? We're now told that we must average our measurements over the duration of the content. Thus, Dialogue Intelligence will not necessarily help us there. What to do?

Well, maybe this is a situation where the European approach of level-gating (as spelled out in ITU-R BS.1770-2) may actually help us. Recent discussions within the ATSC suggest that maybe a hybrid approach would work best. There are currently ongoing discussions that suggest that perhaps the best approach would be to use dialogue-gating for long-form measurements and in turn using level-gating for the short-form measurements. Thus, we would have the best of both worlds. Dolby has provided just such guidance for the automated workflow in the new User Guide for the Dialogue Intelligence Reference Code. This is available free of charge online at the Dolby website.

## COOPERATING THROUGHOUT THE PRODUCTION CHAIN

In the FCC's recent ruling about the enforcement of the CALM Act, they stress the importance of cooperation at the different stages along the path from production to transmission. Consistent with that is the new paradigm of "Safe Harbor" that is granted to operators when passing through content that is certified by the upstream provider to be fully compliant with A/85. This is a new concept. The production community has no history with providing such certification. Furthermore, in my experience, there is not a lot of understanding of the details of A/85 within the production community. In the past, when it's a "wrap," the content is

delivered and it's left to the downstream entities to deal with the A/85 properties: measurement, metadata authoring/management, dynamic range control and loudness management in general. Now there is a need to rethink that.

There are different cultures at work here. Cinema is not TV. They are two different worlds. As long as they stay separate, that's not a particular problem. But they come together when movies are played on TV. Then there can be a culture clash. There are now entirely different standards for loudspeaker monitoring in cinema and TV. There are different monitoring levels (85 dB SPL vs. 78 dB SPL). There are different configurations for the monitoring geometry. There are different calibrations for the surround channels. There are all the metadata parameters that are specific to TV.

The point here is that movies really need to be prepared specifically for TV. It's not enough to simply send the theatrical mix to air. Ideally, the program should be mixed specifically for TV. If not that, at a minimum, it should at least be remastered for TV. It needs to be monitored in the "ITU Circle" (BS.775), in a small space and at 78 dB SPL. Ideally, it

**Movies really need to be prepared specifically for TV. It's not enough to simply send the theatrical mix to air. Ideally, the program should be mixed specifically for TV.**

should include the AC-3 encode, where the DRC (dynamic range control) is chosen by the production team, and all the selected metadata effects should be monitored to ensure that all the creative intent remains intact.

In general, the portion of A/85 dedicated to the monitoring environment is the least understood. This is unfortunate. The benefits of standardizing the monitoring environment have been well demonstrated in cinema. It needs to be the same in broadcast. If operators are given consistently calibrated monitors, they very quickly become adept at knowing by merely listening when things are too loud or too soft. Consistency of monitoring is the key. There is a standard. Monitors need to be calibrated to the standard.

This is especially valuable in live TV. In live TV, the first line of defense must be the ears of the mixers. It will still be important to supplement that with proper loudness measurement gear. But, in the live environment, the mixers will mostly use their ears with just occasional glances at the loudness meter to confirm they're on target. Monitoring is a key component in this mix. That's why the ATSC saw fit to include the topic as one of the four main concepts in A/85—along with Loudness Measurement, Metadata Management and Dynamic Range Control. It matters. It's important.