Telestream
Whitepaper

# Audio Loudness – A Guide to CALM Act Compliance

"the CALM Act is intended to spare viewers the annoyance of constantly adjusting the volume on their TVs to compensate for the significantly higher audio level of commercials"

## Contents



## Introduction

For many years, viewers around the world have complained about big differences in volume between commercials and the surrounding programming. But it's only in recent years that lawmakers around the world - spurred by public pressure and newly refined techniques for quantifying apparent loudness - have taken concrete action to address the issue with enforceable regulation. In the United States, that action is the Commercial Advertisement Loudness Mitigation (CALM) Act of 2010, which came into effect December 13, 2012. The Canadian CRTC also adopted the underlying specifications, with enforcement starting in September 2012. For European countries, the EBU (European Broadcasting Union) issued the R128 loudness recommendations in 2010. Compliance with these standards is mandatory in each region.

Like its counterparts in other countries, the CALM Act is intended to spare viewers the annoyance of constantly adjusting the volume on their TVs to compensate for the significantly higher audio level of commercials. The law holds broadcasters responsible for ensuring volume consistency across all program components. Those failing to meet this responsibility may incur significant financial penalties imposed by the Federal Communications Commission (FCC).

While the law is simple enough in concept, the devil is in the details. In particular, three main aspects of the issue must be well understood in order to comply intelligently and efficiently with the requirements of the new law:

- How does the law envision that apparent loudness will be quantified and compared (e.g. A/85:2013, BS 1770-1, EBU R128-2014, and Dolby Dialog Detection)?
- What are the most effective loudness control techniques available to ensure compliance with the law?
- What are the most efficient approaches to integrating loudness control into the workflow of media enterprises - such as broadcasters, cable MSOs, and satellite providers - that handle a significant quantity of video content?

By providing an overview of the issues above, this document offers guidance on coping most effectively with the technical aspects of complying with the CALM Act.

## The CALM Act and ATSC A/85: 2013

The CALM act directed the FCC to introduce regulations requiring any broadcast station, cable operator, or other multichannel video programming distributor (MVPD) to control the loudness of the commercial advertisements that accompany their programming.

The law mandates that the application of loudness control shall conform to the recommended practice developed by the Advanced Television Systems Committee (ATSC) and codified as RP A/85, which is the ATSC's Recommended Practice Techniques for Establishing and Maintaining Audio Loudness for Digital Television. A/85 was originally approved in 2009, and subsequently updated in 2011 and again in 2013. The current version is therefore A/85:2013.

Because the CALM Act mandates conformance with A/85, understanding compliance begins with understanding the ATSC's recommended practices.

The relevant portions of A/85:2013 deal with these main issues:

- Loudness measurement - what techniques are to be employed to determine the loudness of a given clip (commercial or programming)? What is to be measured, and how?
- Loudness adjustment - if a given commercial clip does not have the desired loudness, how is the loudness of the commercial best adjusted?
- True peak - what is the impact of loudness correction on the maximum level of the program?

## Defining the Anchor Element

The starting point for understanding A/85:2013 loudness measurements is to define what is to be measured. A/85:2013 recommends that loudness measurement be done only on the "Anchor Element" of the audio, which is defined as the perceptual loudness reference point of the content.

According to A/85:2013, in most programming, most of the time, the perceptual loudness reference point is the dialog. This reflects the fact that people are generally more sensitive to loudness of speech than to the loudness of other elements in the audio. Because speech is critical to our understanding what is happening on screen, it's more annoying to be unable to hear speech clearly (because it's too quiet, for example) than to be unable to clearly hear background music or sound effects.

A/85:2013 also allows for an element other than dialog, such as music, to serve as the Anchor Element for a loudness measurement if that element is deemed more appropriate in the context of a particular piece of content. In such a situation, the Anchor Element shall be the element that "a reasonable viewer would focus on when setting their volume control."

## Loudness Measurement Techniques

The human ear perceives loudness as a combination of sound pressure and the dynamics of the sound. Short sudden peaks of sound level sound much louder than the same high levels when heard continuously for longer periods. The sensitivity of the human ear changes as a function of frequency, so loudness is also related to the frequency of the sound.

"the technique used to measure loudness must correspond to human perception of what is and is not loud"

Loudness is therefore a combination of audio level, dynamics and frequency, and there are many variations of algorithms that seek to combine these to match the physiology of the human ear. A commonly used family of curves in audio loudness or noise measurement is known as A-weighting, relating to the measurement of sound pressure level.

Loudness is also relative to the surrounding ambient sound level or reference level. Someone speaking at normal levels in your room at home may be perfectly audible but speaking at the same level in the middle of a rock concert would be inaudible. Dolby uses a 'Dialog Norm' reference level in their AC3 audio compression which is a measure of the A-weighted average level of dialog within a presentation against a normal speaking level. It ranges in integer values from 31, where decoder gain remains at unity, to a value of 1, where decoder gain is reduced by 30 dB.

To be valuable in combating the problem of inconsistent loudness, the technique used to measure loudness must correspond to human perception of what is and is not loud. It turns out that this is largely context dependent, and not as simple as measuring sonic energy at a given instant in time.

Loudness measurements have traditionally been based on the VU (volume unit) and Peak Level meters, technologies that originated in the analog audio world and were carried over into the digital domain. Peak Level indicates the moment of highest voltage, which translates into the greatest sound pressure level (SPL) when the program is reproduced via a loudspeaker (see Figure 1).

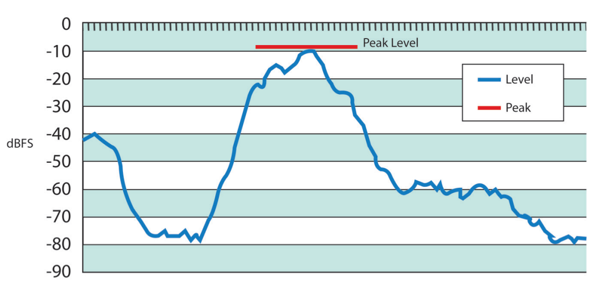



Figure 1 – Peak level measurement.

While peak level is a useful measurement for preventing distortion in electronic circuits (e.g. clipping), its utility in predicting the human perception of loudness varies greatly depending on its relationship to the average level of the program.

The closer the average level is to the peak, the louder the overall program will seem to be. Conversely, isolated peaks in an otherwise quiet program (low average level) don't create the perception of overall loudness. Thus, as illustrated in Figure 1, if the peak level is much higher than the average level, then overall volume adjustments based on peak may make some sections of the content unacceptably quiet.

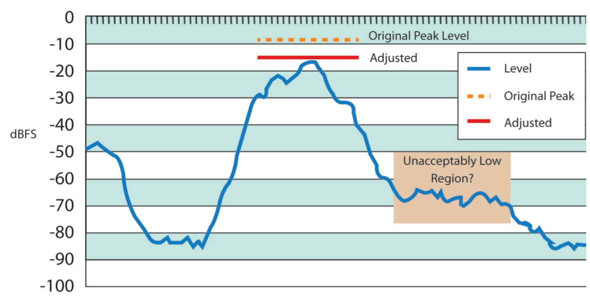
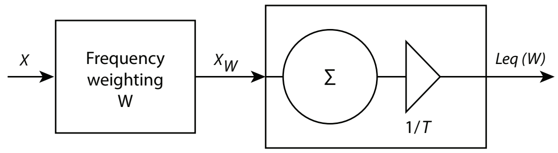



Figure 2 – Adjusting overall level based on peak level may make some parts of the program too quiet.

## A new approach to loudness measurement

Recognizing the problems inherent in a simple peak-based assessment of program level, the International-Telecommunications Union (ITU) developed an alternative specification for loudness measurement. ITU-R BS.1770 was first published in 2006 and subsequently revised as 1770-1 (2007) and, more recently, 1770-2 (2011). The Calm Act refers to A/85, and A/85:2013 specifies BS.1770 (specifically referencing BS.1770-1) as the source of its loudness measurement techniques (1770-2 did not exist at the time A/85 was finalized). So BS.1770-1 currently serves as the yardstick by which U.S. television programming will be evaluated for CALM Act compliance.

ITU-R BS.1770 specifies 'Algorithms to measure audio program loudness and true-peak audio level' and states that 'for the purpose of program exchange, it is essential to have a single recommended algorithm for objective estimation of subjective loudness'. The measurement system specified is shown in Figure 3. BS.1770 recommends the Leq(RLB) measurement algorithm, where Leq(W) the frequency weighted sound level measure, xw is the signal at the output of the weighting filter, xRef is the reference level, and T is the length of the audio sequence.




$$Leq\,(W) = 10\,\log_{10}\left[\frac{1}{T}\int_0^T \frac{x_W^2}{x_{Ref}^2}\,dt\right] \qquad \text{dB}$$

Figure 3 – BS.1770 measurement system.

Subjective testing was carried out by the Audio Perception Lab of the Communications Research Center, Canada, using program materials from television and radio broadcasts from around the world, as well as from CDs and DVDs. The sequences included music, television and movie dramas, sporting events, news broadcasts, sound effects and commercials. The reference signal used was a level of 60 dBA, a level found to be a typical listening level for television viewing in actual homes. However, this is for a single mono channel. In multichannel loudness measurement, the loudness of each of the individual audio channels is measured independently by the Leq(RLB) algorithm before taking the mean square and summing them together. Pre-filtering is applied to each channel prior to the measure. The system for multichannel measurement is shown in Figure 4.

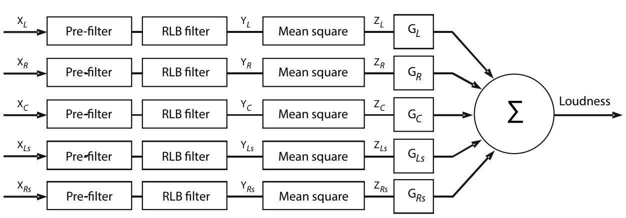



Figure 4: BS.1770 measurement system for multichannel audio.

The weighting applied to each channel depends on the number and positioning of channels. Two channels of stereo can be combined with the same weighting, but with surround sound the different channels are weighted as follows:

| Channel | Weighting, $G_i$ |
|---|---|
| Left ($G_L$) | 1.0 (0dB) |
| Right ($G_R$) | 1.0 (0dB) |
| Centre ($G_C$) | 1.0 (0dB) |
| Left surround ($G_{Ls}$) | 1.41 (~ +1.5 dB) |
| Right surround ($G_{Rs}$) | 1.41 (~ +1.5 dB) |

The drawback of BS.1770 as originally conceived is that it measures average loudness over the entire length of content. This may be fine if the loudness is fairly consistent over time. If not, a quiet section of content may, as illustrated in Figure 5, bias the average level so that it measures as acceptable despite having some sections that are unacceptably loud.

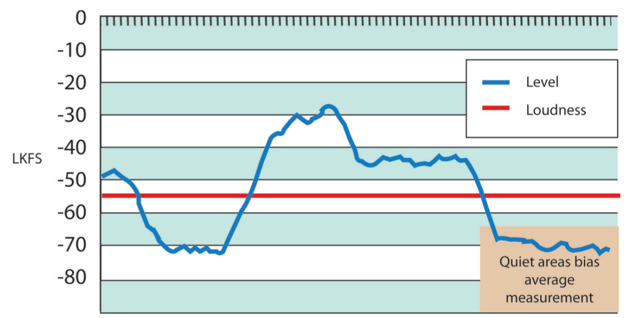



Figure 5 – Average loudness measurement.

## Measuring a dialog sample

The downside of the average measurement technique described in BS.1770-1 helps explain why the A/85:2013 concept of the Anchor Element is so important in obtaining valid measurement results. While BS.1770-1 envisions that a measurement will be taken for the full duration of the content, A/85:2013 recognizes that in practice there are situations that may be either difficult or misleading (as shown previously in Figure 5). So it allows instead the measurement of a representative sample of the Anchor Element, and presents guidelines for choosing that sample in different situations (live content, finished long-form content, short-form content, file-based content, etc.).

In practice, implementation of Anchor Element measurements involves identifying - either manually or using automated "dialog detection" or "speech detection" techniques – those areas of the program where dialog is predominant. The BS.1770-1 measurement is then applied only to sections of content that contain dialog. This generally removes measurement bias, as dialog is generally not very quiet, and it also acknowledges the fact that people are particularly sensitive to dialog levels.

Useful as it is, dialog detection has some potential weaknesses. Dialog detection algorithms vary in their accuracy, and while algorithms from industry-leading companies such as Dolby are remarkably accurate, there is currently no completely foolproof method of automated dialog detection. Further, while the vast majority of content has a significant amount of dialog, applying dialog-based measurements to content without much dialog may not give a good indication of subjective loudness. A fallback plan is generally required for such content.

## Gated Measurements

Measuring a dialog sample is not the only way to avoid the potentially misleading results that can come from measuring across an entire program. An alternative approach is to use a measurement gate. First put forward as part of the European Broadcast Union's R128 standard, this technique was later added to BS.1770-2. Since A/85:2013 has not been revised to reference 1770-2, it is currently unknown whether measurements made using this technique are considered by the FCC to be consistent with the requirements of the CALM Act.

Gated measurement works by analyzing the loudness of the audio in short sections. The loudness value of a given section will count toward the overall loudness value of the program only if that section measures above a certain threshold (the "gate" value). The gate effectively excludes quiet periods from the final measurement. Figure 4 shows a gated measurement applied to the same signal as the ungated measurement from Figure 6.

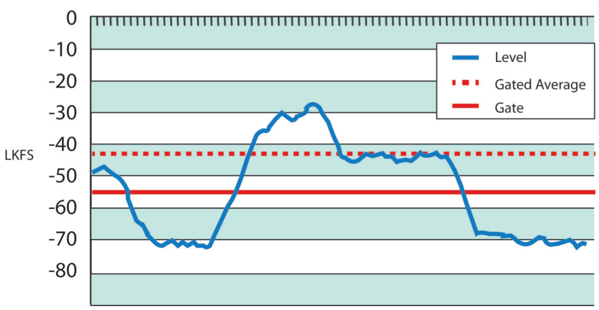



Figure 6 – Gated Loudness measurement.

There remain on-going discussions on whether gated measurements or dialog measurements are more effective in producing loudness measurements that track the perceptions of the typical TV watcher. Neither gated measurement nor dialog measurement is explicitly mandated by A/85:2013, but since A/85:2013 envisions that dialog will normally be the Anchor Element it could be argued that the dialog method more closely reflects the intent of the current standard. There are many in the world of broadcast audio, however, who feel that automated dialog detection is particularly inaccurate or inappropriate for short form content, such as the 30 second commercials that represent the majority of content covered by the CALM act. It is quite likely that, at some point, use of gated measurement will be recommended as best practice for such content.

The ideal may be a hybrid of both approaches, with dialog measurement being used for content that has a high proportion of dialog, and gated measurement for content that doesn't have much dialog. Because of this uncertainty, the wisest course for now may be to avoid getting locked into a loudness measurement system that supports only one, but not both, of these approaches.

## Loudness correction with dialnorm

With the measurement techniques described above, it's possible to identify materials that are likely non-compliant with the CALM Act and to quantify the amount of loudness correction needed. To understand how the correction itself is best applied requires familiarity with "dialnorm". Dialnorm is a metadata parameter used in a number of audio compression schemes, including the Dolby Digital (AC-3) codec that is part of the ATSC specifications for broadcast television in the U.S.

> "...dialnorm is a metadata parameter used in a number of audio compression schemes... dialnorm can be used to standardize program output to a consistent level of dialog"

Carried in the metadata associated with each compressed audio stream, the dialnorm value represents the loudness of the dialog in that stream expressed in LKFS as measured with BS.1770 techniques. Because every Dolby Digital decoder is equipped with the ability to adjust the audio output level based on the dialnorm of the content being decoded, dialnorm can be used to standardize audio output to a consistent level.

To see how this works in practice, consider the transition from a television program to a commercial. If the program's measured loudness dialnorm value is -24 and the commercial's is -21, the commercial is 3 LKFS louder than the program. The audio decoder corrects this by enforcing the longform dialnorm value to the commercial content and thus attenuates gain by 3dB for the duration of the commercial, removing the attenuation when the commercial ends and the program resumes. The net effect would be to make the apparent loudness of the commercial dialog consistent with that of the program.

The concept of dialnorm can be applied in different ways depending on the situation:

- Gain-based loudness correction, also referred to as Fixed dialnorm, involves the enterprise - network, station, or cable operator - settling on a standard dialnorm loudness target and then adjusting the gain of each program and commercial so that its loudness measures at the target value. Fixed dialnorm is the only option for audio signals that do not carry dialnorm metadata, and it removes the requirement for level adjustments by the decoder at the receiving end. The drawback of this approach is that it requires a broadcaster to analyze every piece of content, and to correct any piece whose dialnorm falls outside a predefined Target Loudness standard (typically within 2 dB of -24 LKFS). At a very minimum this correction process requires decoding, adjusting, and re-encoding the non-standard audio content.
- Metadata-based loudness correction, also referred to as Agile dialnorm, is the strategy of measuring the dialnorm on content and simply putting the correct dialnorm value in the metadata of the audio, relying on the decoder at the receiving end to adjust the volume accordingly. This approach has the advantage of allowing loudness to be corrected without decoding and re-encoding the audio, but it assumes that all receivers have the ability to adjust level based on dialnorm.

## True Peak adjustment

Dialnorm provides a valuable framework for loudness-correction, but to be complete an effective loudness control scheme must account for the impact of gain adjustments on other aspects of the adjusted audio stream, including the stream's absolute maximum amplitude, which is referred to as "true peak." Referencing Annex 2 of BS.1770, A/85:2013 describes true peak as being measured in dB TP, meaning decibels relative to full-scale (the absolute maximum possible amplitude).

If positive gain is applied to a stream whose true peak is already close to the maximum possible value, the result may be clipping (overload), introducing audible distortion into the audio. A/85:2013 recommends a target true peak of -2 dB TP for interchanged audio so that headroom is available to apply some downstream processing without clipping.

Ensuring that loudness-corrected audio complies with True Peak guidelines requires measuring the True Peak of the program and calculating the effect of loudness correction on that peak. A couple of different strategies are available for dealing with situations in which loudness correction would make the true peak too high:

- Reduce the amount of gain applied by loudness correction (or adjust the dialnorm value) such that true peak does not exceed the specified limit. The downside of this approach is that the loudness-corrected content will be quieter than it should be to achieve full loudness correction.

- Apply a peak limiting algorithm to reduce the peak without significantly affecting the overall loudness of the content. This is typically the preferred approach, but in some types of program peak limiting can result in noticeable audio artifacts (e.g. "pumping").

In addition to loudness and True Peak, A/85:2013 also concerns itself with a number of related issues including dynamic range control, setup and calibration. These topics are complex and fall beyond the scope of this document.

## Workflows for loudness control

At this point it should be evident that the CALM Act, while yielding obvious benefit for the content consumer, places an important new responsibility on the content provider, which is to ensure that content is delivered to the consumer with the correct audio loudness and metadata values. While the failure to do so may result in significant penalties, the burden of compliance need not be onerous.

The extent to which compliance is problematic for a given enterprise depends largely on the workflow employed in readying program and commercials for delivery to viewers:

- Facilities relying solely on real-time baseband SDI-based infrastructure will find CALM Act compliance the most difficult. Expensive new hardware will be required at some point in the signal chain to provide the loudness regulation capability. Additionally, because real-time devices must work in a single pass, they are unable to analyze an entire piece of content before making level corrections. As a result, audio quality may be compromised in order to achieve regulatory compliance.

- Facilities already utilizing a file-based workflow may well find CALM Act compliance relatively easy and painless. A typical file-based infrastructure already incorporates workflow automation and transcoding systems that route incoming assets for re-wrapping or re-encoding as needed for distribution and archiving. Assuming that such automated processes have been well implemented, adding loudness measurement and correction to the workflow is a relatively simple matter.

In a well-designed file-based workflow, conformance to A/85:2013 would typically require the addition of only one step, which is the analysis of program loudness (LKFS) and true peak (dB TP). In many settings the workflow may already include some form of audio analysis, in which case the existing analysis methodology need only be conformed to A/85:2013's recommended practices.

The favored method would be to perform the analysis step on content as it is first delivered to the facility, so that metadata from the analysis is available to downstream processes. That way, if correction is required it can be applied in an existing downstream workflow step. This approach allows loudness regulation to be added to an existing workflow with little or no additional cost in terms of processing time or additional workflow steps. Once the workflow changes have been made, the process continues to work as seamlessly as before.

## Integrated workflow solutions

Telestream has long believed that tight, flexible integration between process steps is the key to maximizing the speed and resource-efficiency of content processing solutions. In a file-based loudness regulation system that is not designed for tight integration, of which there are a number on the market, the overall efficiency of the workflow can suffer as a result of this loose coupling.

A vendor that specializes in audio, for example, may provide excellent loudness regulation but may not provide the necessary format support for video codecs and container formats. This can lead to files having to be re-wrapped or transcoded into a format that the audio correction software can deal with, then transcoded again into the intended delivery format.

A tightly integrated, file-based environment makes it possible to achieve far better quality in loudness corrected content. Integration facilitates the handoff of analysis metadata and allows the application of different loudness correction techniques simultaneously as a given piece of content is repurposed for multiple delivery platforms (e.g. broadcast, web delivery, satellite network distribution). Ideally, a file-based system should also be able to choose between measurement methods (dialog detection or gating) based on the proportion of dialog in a piece of content.

Finally, and equally important, such a system can provide logging of data about the analysis and correction performed on every asset that passes through. Such logging could prove invaluable if there were a question about compliance with regulations.

The benefits of using a well-designed, tightly integrated file-based workflow automation system for loudness correction are enough to warrant the introduction such a system into a facility for the first time in response to impending CALM Act enforcement. The cost of a server and software license will frequently compare favorably with the cost of a less-effective hardware solution. The Telestream Vantage platform, for example, combines analysis, metadata transfer, transcoding, and audio level correction into a highly integrated file-based workflow - providing the ideal environment for broadcasters, cable operators, and other multichannel video programming distributors to address CALM Act requirements. Both gated and dialog measurements are included, allowing customers to choose the methodology that they feel is best suited to meeting their compliance needs.

Vantage is one of the few solutions on the market today that can provide CALM Act compliance in a tightly integrated solution with a choice of loudness measurement options.

## To Learn More

Visit www.telestream.net or call Telestream at 1.530.470.1300

www.telestream.net | info@telestream.net | tel +1 530 470 1300