Audio Engineering Society

Convention e-Brief 318

Presented at the 142nd Convention
2017 May 20–23 Berlin, Germany

Loudness Management in the Blu-ray Disc ecosystem in the context of today's playback environments

Andreas Ehret[1], Sripal Mehta[2], and Mike Ward[2]

[1] *Dolby Laboratories, Inc., Deutschherrnstr. 15-19, 90429 Nürnberg, Germany*

[2] *Dolby Laboratories, Inc., 1275 Brannan Street, San Francisco, USA*

Correspondence should be addressed to Andreas Ehret (andreas.ehret@dolby.com)

ABSTRACT

Loudness management within the Blu-ray Disc ecosystem has historically been less of a priority than in other media playback ecosystems. Instead, the industry has focused on delivering the highest fidelity, and full dynamic range audio. As a result, the measured loudness of the content on Blu-ray Disc is generally not accurately indicated in the audio bitstreams carried on Blu-ray discs. However, as more use-cases emerge to connect Blu-ray Disc players to playback environments with limited dynamic range reproduction capabilities (such as TVs or Sound bars), loudness management is becoming more important to ensure optimal playback for these new device types. This brief explains the value of loudness management in the Blu-ray Disc ecosystem to address new playback environments, and gives example workflows for correctly setting loudness values in audio bitstreams delivered on Blu-ray Disc.

1 Introduction

For a consumer, selecting a single audio playback level accurately in a consumer playback device is increasingly difficult in today's environment, with an ever-increasing number of input sources, formats, use-cases, and output configuration options. One approach to help solve this problem, has been the introduction of metadata alongside the digital audio data that describes the loudness of the audio content. However, if the value of this metadata is not set accurately, the intended purpose of managing loudness to achieve an optimal consumer experience cannot be realized and may lead to a compromised listening experience. With the emergence of new audio playback devices with constrained output level and dynamic range capabilities, loudness management for Blu-ray Disc audio has become increasingly important to ensure a high-quality consumer playback experience.

2 Loudness Management

The capabilities of audio playback systems vary widely - e.g. a full-range home-theater system versus a mass-market consumer flat-panel television. When loudness is set correctly consumers will have a consistent experience from content to content, source to source. The purpose of loudness management is to optimally map the reference level and dynamic range of the digital audio signal to the reproduction capabilities of the playback system.

To achieve this, the audio level is adjusted so that the average level (commonly known as the dialogue level) is aligned to a common playback reference level, and for devices or playback use cases that require a limited dynamic range, the dynamic range of the original audio is also reduced – see Figure 1.

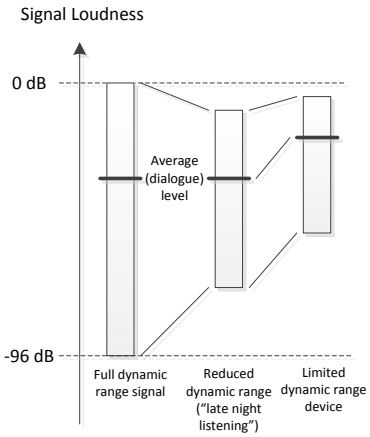

Figure 1: Matching Dynamic Range of the Signal and Playback System

If the loudness management system is correctly configured, it benefits both content producers and consumers. A content producer can focus on creating the best sounding mix given the capabilities of the target audio format, and a consumer will experience audio playback at a consistent loudness level with best possible audio quality within the constraints of the playback device. For consumers who may choose to play back audio at reduced dynamic range (e.g. late night listening), correctly configured loudness management ensures that loud sounds are properly attenuated and dialogue is still audible in quiet scenes.

3 Dialogue Normalization

For loudness management to work as described above, a consistent playback reference level needs to be defined. Humans anchor the perceived loudness around the average level of the dialog within a program. With the assumption that across the range of program material, the dialogue level can be used as an anchor point, the concept of dialogue normalization (dialnorm) has been introduced. The dialnorm parameter identifies the average loudness level of a program allowing an audio decoder to scale the audio to a consistent output level across multiple programs. As illustrated in Figure 2, the playback level of the content is adjusted so that the dialogue level matches a consistent output reference level (See Figure 2).

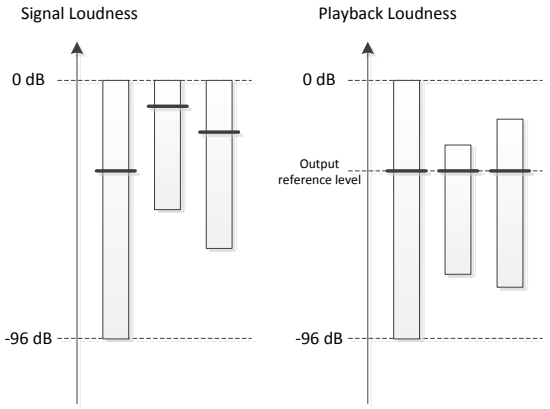

Figure 2: Using dialogue normalization to align varying audio loudness to a common reference playback level.

Dialogue normalization is typically applied during decoding of the audio, based on the value of the dialogue normalization metadata parameter carried in the encoded audio bit-stream. This ensures that scaling of the audio only occurs during decoding, allowing the effects of dialogue normalization to be bypassed in certain use cases. E.g. if the output reference level matches exactly to the dialogue normalization, no levelling needs to be applied.

The value of the dialogue normalization parameter is an input to the audio bitstream encoder. For Blu-ray Disc encoding products, this is typically a manual process. For example, a movie soundtrack has an average loudness level of -23 LKFS. The dialogue normalization parameter input of the encoder is set to -23. To level the average loudness of the movie to the reference playback level of -31 LKFS, 8 dB of attenuation is applied during decoding of the movie soundtrack. If a movie has an average loudness of -31 LKFS, then the dialogue normalization parameter is set to -31, and no levelling of the audio during decoding is required.

4 Dynamic Range Compression based on Dialogue Normalization

Leveling the audio to a common playback reference level ensures consistent playback across different pieces of content and does not restrict the dynamic range of the audio. To achieve this, dynamic range compression (DRC) data is calculated by the audio

encoder and is delivered in the encoded bitstream as time-varying gain words. These gain words are applied during decoding to both reduce the level of loud sounds and boost the level of quiet sounds. This permits all elements of the audio presentation to be heard for the chosen listening environment. To achieve proper dynamic range compression, the dialnorm parameter is input to the audio encoder to be used as the null point of the compressor calculating the DRC gain words for the audio bitstream. If the value of the dialogue normalization parameter is incorrect, the DRC gain words in the audio bitstream may specify incorrect levels of boost and more importantly cut values to be applied during decoding when DRC is active in the playback device.

5 Loudness in Blu-ray Disc audio productions

In the Blu-ray Disc authoring workflow, loudness can be considered at two points: real-time during the actual home theater mix process, as well as a file-based long-term measurement of the entire audio program after the audio master has been created.

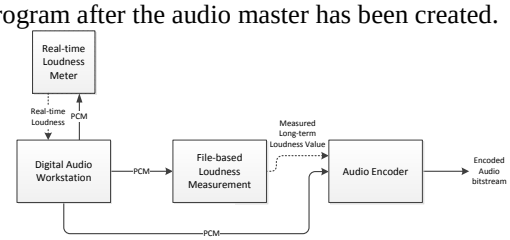

Figure 3: Blu-ray Disc Loudness Workflow

After a movie is mixed on a cinematic mix stage for the cinema environment, it is often brought to a home-theater mix stage to allow a mixer to edit the content and adjust dynamic range for the home theater environment. During the home-theater mix process, loudness can be measured in real time via real-time loudness meter plugins for digital audio workstations. These plugins provide the mixer with short-term loudness measurements to give the mixer a general sense of the loudness of the program or the section that is being re-mixed.

Once the sound mixer has finished re-mixing the content, a digital audio workstation will export a multichannel wav file or multiple mono wav files, one for each channel of a multi-channel mix. In order to obtain the loudness for the entire program, the entire program audio must be measured by a non-real-time, file-based loudness meter to calculate the average program loudness. The resultant loudness value can be fed into the audio encoding software to correctly set the average program loudness within the final audio bitstream.

With the introduction of new immersive audio formats, tools, and content, sound mixers have been using the extended creative capabilities now available to improve sound mixes. The effect is that there is even greater dynamic range in content than in the past – mixers are now mixing full-dynamic-range sounds into surround and height positions within an audio mix. Hence, to accurately convey loudness metadata is even more important.

6 Playback Environment Landscape

For Blu-ray Disc production, it has been common industry practice to set a constant dialogue normalization parameter of -31dB. This value for the dialogue normalization parameter matches the “line-mode” output reference level of an audio decoder typically used in Audio/Video Receiver or Soundbar playback device implementations. This value means that during decoding, no level adjustment will be applied, with the intent that the consumer is hearing the sound mix at the same mix level as intended by the audio mixer. Apparently, setting loudness metadata appropriately is not being considered consistently in today’s Blu-ray encoding workflows, leading to variations in loudness across Blu-ray titles.

Historically, while loudness between different Blu-ray Discs may not have been consistent, this did not greatly affect the consumer experience. If program loudness was consistently set across all content on a single disc (e.g. between trailers and the main movie), the consumer could set the volume level on their playback device once and could experience consistent loudness.

However, with Blu-ray Disc players now embedded in video game consoles, consumers can quickly switch between disc-based, streaming-based, and

game content. If the loudness across these content sources is not consistent, the consumer would need to frequently adjust the volume control as they switch between sources. Additionally, if a consumer switched to a different playback device such as a set-top box or digital media adapter, the consumer would need to further adjust the volume level of their playback system.

Additionally, while Blu-ray Disc audio was targeted at consumers with premium audio/visual reproduction capabilities in the home (e.g. multichannel loudspeaker systems), consumers today commonly connect their Blu-ray Disc player directly to their television using HDMI. Due to the physical limitations of the loudspeakers on televisions and the need to align with broadcast loudness practices, televisions are typically configured to decode audio with heavy dynamic range compression (e.g. RF mode in AC-3) to meet the reduced dynamic range capabilities of the smaller drivers in televisions.

7 User Experience Impact

When heavy DRC is applied to content with improper loudness metadata values, audible artifacts such as level pumping and over-compression of the audio, even dialogue audio, may be experienced, resulting in poor audio quality. An example of over-compression effects due to an incorrect dialogue normalization value is shown in Figure 4.

In this case, the measured average loudness of the audio is -22 LKFS but the dialogue normalization value specified at encoding time and present in the bitstream is -31. The result is that the null area of the heavy dynamic range compression is 9 dB too low causing the audio to experience near-constant cut gain values of 10 dB with some cut gains as large as 20 dB. In addition, none of the quieter audio elements are boosted as desired. The audible result is an overly-compressed audio mix with a significant risk of level pumping as the dynamic range compressor is active way too often.

Televisions now commonly support streaming and broadcast content services where content loudness management is common. Consumers may be able to access the same content from these sources that they have available on Blu-ray Disc. Despite Blu-ray Disc being arguably the premium and highest quality format for delivery of this content, if audio loudness

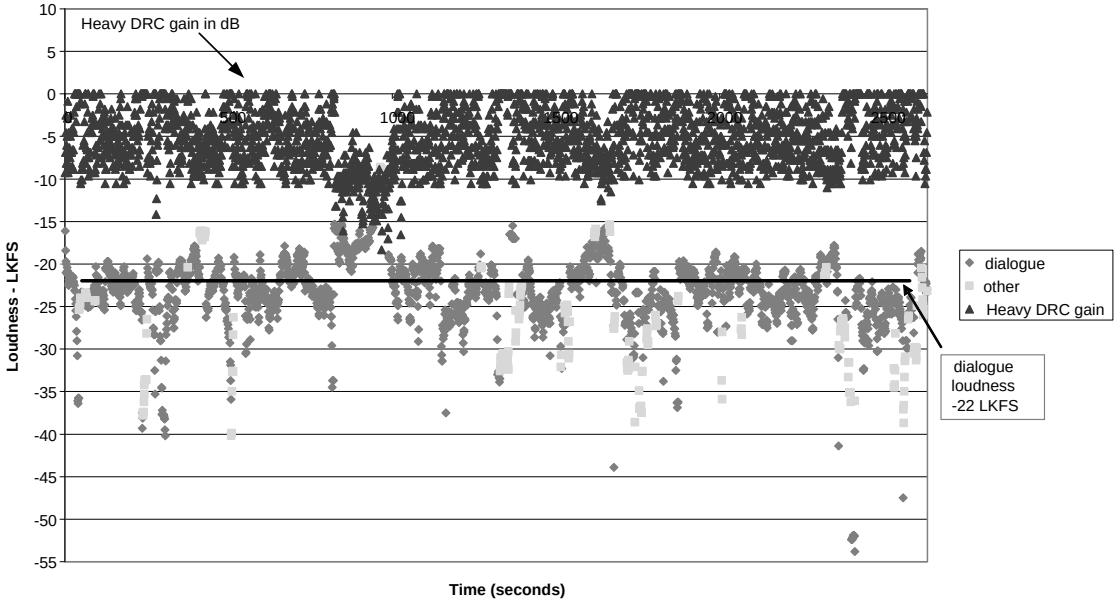

Figure 4: Excessive heavy DRC cut gains due to incorrect dialogue normalization value, resulting in over-compression

is not correctly managed, the experience when listening to the Blu-ray Disc audio may be worse than when listening to the same content from a broadcast or online source due to the over-compression artifacts previously described.

8 Implications

According to our investigations, dialogue normalization loudness values (according to [2] with dialog gating) on Blu-ray Discs today typically vary between -19 and -30 LKFS – whereas in all those cases the dialogue normalization parameter of the audio bitstreams on the disc had been set to -31 dB. Table 1 shows the dialnorm parameters for several example commercial Blu-ray Discs and the actual measured content dialogue level:

Content	Bitstream Value	Measured Value
Asset 1	-31 LKFS	-23 LKFS
Asset 2	-31 LKFS	-19 LKFS
Asset 3	-31 LKFS	-27 LKFS
Asset 4	-31 LKFS	-20 LKFS
Asset 5	-31 LKFS	-30 LKFS

Table 1: Measured loudness for Blu-Ray Discs

This investigation has revealed that loudness level inconsistencies between the dialogue normalization parameter and the actual loudness of the audio content of up to 12 dB can be observed, which can lead to a compromised listener experience as described above.

If the dialogue normalization parameter were set correctly, the audio playback level would be reduced to align with the playback loudness of other content. Additionally, if dynamic range compression is enabled on the playback device, proper dynamic range processing would be applied preserving optimum audio quality.

9 Best Practice Recommendations

Loudness and dynamic range should be given good consideration during the mastering and encoding process. Loudness should always be measured on the finalized asset before audio encoding to determine program loudness and properly set dialogue normalization parameter in the audio bitstream. To ensure that dynamic range is handled properly, the dynamic range profile(s) should also be set to ensure that the decoder employs the desired amount of compression in normal (e.g. Line mode) and heavy (e.g. RF mode) compression. Finally, checking the final encoded audio on a constrained dynamic range device will ensure that the desired audio levels and quality are achieved.

10 Conclusions

Loudness management in a consumer playback device can only work as intended if loudness metadata is set accurately. This has not been a priority of the Blu-ray content industry today. However, with today's consumer's home setups and playback environments it would seem increasingly important to incorporate loudness measurement in the Blu-ray production workflow in order to ensure an optimal audio listening experience for the consumer.

References

[1] Jeffrey C. Riedmiller, Steve Lyman, and Charles Robinson, *Intelligent Program Loudness Measurement and Control: What Satisfies Listeners?*, AES 115th Convention (2003)

[2] ITU-R BS.1770-4, *Algorithms to measure audio programme loudness and true-peak audio level*, International Tele-communications Union, Geneva (2015)

[3] ATSC A/85, *Techniques for Establishing and Maintaining Audio Loudness for Digital Television*, Advanced Television Systems Committee (2013)

[4] EBU Technical Recommendation R128, *Loudness Normalisation and Permitted Maximum Level of Audio Signals*, European Broadcasting Union (2015)

[5] Scott G. Norcross, Sachin Nanda, and Zack Cohen, *ITU--R BS.1770 based Loudness for Immersive Audio*, AES 140th Convention (2016)