Technical Document

AESTD1008.1.21-8 (supersedes TD1004)

Recommendations for Loudness of Internet Audio Streaming and On-Demand Distribution

August 30, 2021

Audio Engineering Society, Inc.
132 East 43rd Street, Suite 406
New York
NY 10017
USA
Tel: +1 212 661 8528. Email : hq@aes.org. Internet : http://www.aes.org

The Audio Engineering Society's Technical Council and its Technical Committees respond to the interests of the membership by providing technical information at an appropriate level via conferences, conventions, workshops, and publications. They work on developing tutorial information of practical use to the members and concentrate on tracking and reporting the very latest advances in technology and applications. This activity is under the direction of the AES Technical Council and its Committees.

The Technical Council and its first Technical Committees were founded by the Audio Engineering Society in 1979, and standing rules covering their activities were established in 1986, with the intention of defining and consolidating the technical leadership of the Society for the benefit of the membership. The Technical Council consists of the officers of the Technical Council, the chairs of the Technical Committees, the editor of the Journal, and as ex-officio members without vote, the other officers of the Society.

TECHNICAL COUNCIL
Steve Hutt - Co-Chair
Juergen Herre - Co-Chair
Michael Kelly - Vice-Chair

TECHNICAL COMMITTEES
Acoustics and Sound Reinforcement
Archiving, Restoration and Digital Libraries
Audio for Cinema
Audio for Games
Audio Forensics
Audio for Telecommunications
Automotive Audio
Broadcast and Online Delivery
Coding of Audio Signals
Fiber Optics for Audio
Hearing and Hearing Loss Prevention
High-Resolution Audio
Loudspeakers and Headphones
Machined Learning and Artificial Intelligence
Microphones and Applications
Network Audio Systems
Perception and Subjective Evaluation of Audio Signals
Recording Technology and Practices
Semantic Audio Analysis
Signal Processing
Spatial Audio

www.aes.org/technical

The goal of AES Technical Documents is not to state a policy or support for a technical method, device or principle, but to provide information that represents the collective knowledge of a group of experts. Neither AES nor any of its Committees or members shall be responsible for any consequences resulting from the use of information contained in this publication.

AES Technical Committee on Broadcasting and Online Delivery

Technical Document

Recommendations for Loudness of Internet Audio Streaming and On-Demand Distribution

2021–8–30

Prepared by the Study Group on Streaming Audio Loudness
of the Technical Committee for Broadcast & Online Delivery

Co-chairs: Bob Katz and David Bialik

Writing group

John Kean, Editor
David Bialik
Rob Byers
Jim Coursey
Eelco Grimm
Bob Katz
John Kean
Scott Norcross
Robert Orban
Shawn Singh
Jim Starzynski
Alessandro Travaglini

Other Contributors

Ian Shepherd
Greg Ogonowski

TABLE OF CONTENTS

Recommendations for Loudness of Internet Audio Streaming and On-Demand Distribution

1 Introduction

Internet Audio Streaming and on-demand file playback have become major sources of media delivery, affecting the ways that audio is recorded, mixed, post-produced and delivered. Excessive loudness compromises quality, inconsistent loudness annoys listeners. To resolve these issues, this document provides recommendations for establishing and implementing an effective Distribution Loudness for streaming and on-demand audio file playback.

This Document:

- Is intended for use by distributors of Internet audio streams and on-demand audio files.
- Does not provide recommendations for content production. However, content creators and producers will find it essential to their work.
- Is not intended for sound-with-picture content (Over-The-Top, or On-Demand Video). Guidelines for that material are covered in other industry recommendations and standards (e.g., AES71-2018).

These recommendations provide a necessary step in an evolutionary process (see Section 5: Normalization Principles) to accommodate the inadequate maximum gain and limited metadata capability of some current and older playback devices. Establishing this process addresses current needs while anticipating increased dynamic range capability and lower Distribution Loudness. Current metadata-enabled music services and their applications described in Section 4D have significantly improved the listener experience; forthcoming metadata encoded streams and compliant devices can accommodate more extensive improvements. These future advancements make up step two of the evolutionary process, supporting a revision to this document at the appropriate time.

DOCUMENT CONVENTIONS

Underlined text denotes an active link to another section of this document, often to a definition in Annex A or a reference in the following annexes.

LUFS and LKFS are identical units of measurement as specified in ITU-R BS.1770. This document uses LUFS.

2 Primary Goals

This document recommends appropriate loudness for streaming and on-demand file playback content to:

- Optimize distribution and the listener experience
- Recognize the *evolutionary process* by recommending a Distribution Loudness that is well-suited for current fixed and mobile listening, while creating awareness for loudness management using metadata encoded in streams and in future ANSI/CTA-2075 devices
- Recommend a consistent real-time Distribution Loudness for streams
- Ensure loudness consistency of on-demand files and online streams from different sources

- Provide loudness consistency within a specific online stream composed of different long-form content and interstitials, which will alleviate loudness jumps when interstitial content (ads, promos, PSA's) is inserted
- Prevent excessive peak limiting or other processing from degrading perceived audio quality
- Avoid loudness wars
- Encourage the use of audio metadata

3 Evolutionary Process

The recommendations in this document recognize the ANSI/CTA-2075 standard, which supports Loudness Management, including Dynamic Range Control (DRC) and Metadata for fixed and mobile devices. Compliance with that standard will optimize the amount of available gain in a device and offer effective use of DRC. This allows listeners to manually or automatically adapt their sound to the listening environment by modifying the dynamic range.

As ANSI/CTA-2075 adoption proceeds, it is recommended to observe a near-term and long-term approach. While many things such as changes to production practices can be applied now, it is expected that full device adoption of these new standards may require years as consumer hardware and/or software are upgraded. Nevertheless, many compliant devices are already available, which distributors can support by utilizing tandem streams: one without metadata for legacy playback devices, and a second one with metadata to provide user-DRC features. Because many productions will continue to be available for years to come, they can be "future-proofed" by being produced and archived at a loudness of -24 LUFS or lower and then remastered for distribution to current devices by applying linear gain followed by peak limiting if needed.

It is recommended that as the audio capabilities of devices advance and loudness metadata becomes more widely supported, industry experts and stakeholders reconvene to further revise this document by lowering its Distribution Loudness recommendations by 6 LU. This will harmonize this document with others such as EBU R 128, ATSC A/85, ANSI/CTA-2075 and AES71-2018, which recommend -23 to -24 LUFS. This outcome will align audio-only and audio-with-video content to a universal loudness value, benefiting content creators, distributors and all listeners.

4 Recommendations

For the following recommendations, Distribution Loudness is not to be targeted to the upper tolerance. See Section 5 Normalization Principles for a discussion on the lower tolerance and Normalization practices.

Dialog Integrated Loudness measurement is recommended, as indicated in Table 1. When not possible or applicable, Integrated Loudness measurement is recommended.

For all content, it is recommended that the Maximum True Peak level not exceed -1 dBTP at the codec input of lossy-encoded streams. See Section 7B: Evolutionary Process - Peak Control for further explanations.

Table 1

Content		Distribution Loudness (LUFS)	Upper Tolerance (LU)	Loudness Measurement Method
Assorted [1]	Speech is measurable [2]	-18	+1	Dialog Integrated Loudness [3]
	Speech is not measurable See Section 5B	-18	+2 format-specific see Table 2	Integrated loudness
Music [4]	Track-normalized [5]	-16	+0.2	Integrated loudness
	Album-loudest track (e.g., on-demand music services, Section 5D)	-14 [6]	+0.2	Integrated loudness
Interstitial [7]		-18	+0.2	Integrated loudness
Virtual Assistant		-18	n/a [8]	Integrated loudness of assistant's voice preceding volume control

[1] This applies to radio-style streams, musical concert performances, podcasts containing speech, music and/or effects elements.
[2] This applies when speech, music and effects elements are produced with reference to a speech anchor.
[3] See Dialog Integrated Loudness.
[4] Instrumental or vocal music with or without spoken elements and applies to automatically-assembled streams.
[5] -16 applies when practicing Track Normalization; See Music Streaming Services.
[6] -14 applies to the loudest track of an album when practicing Album Normalization.
[7] This applies to Interstitial Content such as commercial advertising, public service announcements, promotional material, political advertising and the like.
[8] Due to the nature of the digitally-generated voice, including alerts, specifying an upper tolerance does not apply. See Loudness, System Sounds and the Virtual Assistant.

5 Normalization Principles

A. Loudness and Normalization

ITU-R BS.1770 defines a method for measuring Integrated Loudness of audio: a frequency-weighted level-gated measurement of average power over an interval of time. BS.1770 Loudness is an electric signal measurement relative to digital full scale, not an acoustic measurement. Absolute loudness is measured in LUFS, loudness units relative to full scale. Relative loudness with respect to a reference loudness is measured in LU, loudness units. A unit of loudness difference (LU) is equivalent to a decibel (dB). The measurement is frequency-weighted to approximate the sensitivity of the ear to different frequencies, and is level-gated to emphasize the parts of the audio contributing most to the sensation of loudness. EBU - TECH 3341 and ITU-R BS.1771 provide additional information about loudness measurement.

Loudness Normalization adjusts the loudness of content to match a desired Distribution Loudness by applying uniform attenuation or gain. This reduces annoying loudness jumps and the incentive to produce loud content.

The Loudness Normalization Process

When applying normalization, the loudness of the original audio might be above or below the desired Distribution Loudness. If it is above, normalization attenuates the audio to the desired value. This process is commonly called downward normalization. If it is below, normalization applies positive gain. This is commonly called upward normalization and may require peak limiting. (See Figure 1)

When upward normalization would cause clipping, peak limiting is required to protect the channel, but can alter the sound of the content. At their discretion, distributors may decide to preserve more of the dynamics of the audio by not fully normalizing it. This is illustrated in Figure 1, labeled “Partially Normalized.” To prevent vulnerable players from sounding too quiet even at maximum gain, it is recommended to keep the Integrated Loudness of content above -20 LUFS. See in Section 5C: Wide Dynamic Range Content for exceptions.

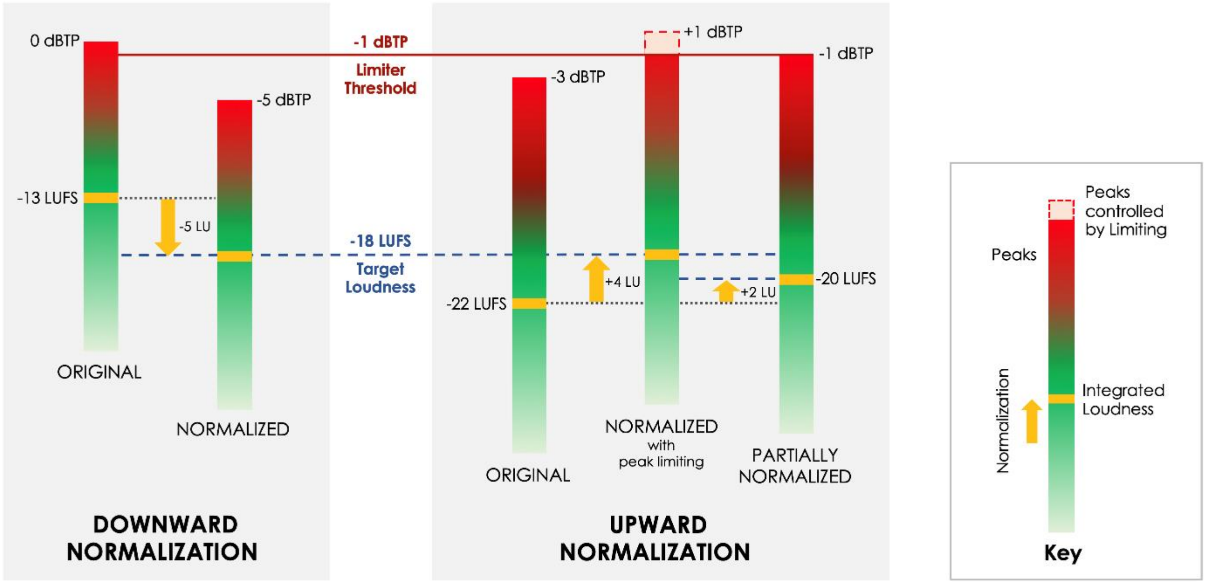

Figure 1 - Processes for Downward and Upward Loudness Normalization

Speech vs. Music

Numerous independent tests of the ITU-R BS.1770 algorithm with human listeners indicate that it is among the best metrics for normalizing a wide variety of long-form content, including broadcast programs and music tracks. Studies have concluded that adjusting the speech portions of audio content to a consistent loudness leads to greater listener satisfaction [1, 2, 3]. However, formal tests with listening panels showed that speech normalized to the same BS.1770 Integrated Loudness as music is typically perceived 2 to 3 dB louder than the music [4, 5]. Therefore, if operationally workable, the listener experience can be improved by normalizing music 2 or 3 LU higher than speech.

Table 1 recommends that speech within streams be normalized to -18 LUFS to accommodate the gain of current player devices. Therefore, it is additionally recommended that music be normalized to an average of -16 LUFS in operations where music and speech are separately normalized and played out automatically. Music normalization can be implemented through either Album Normalization or Track Normalization. Using Table 1 typically achieves the most perceptually pleasing balance between speech and music.

B. No Separable Speech Anchor

When a stream contains speech that is normalized to -18 LUFS and music that is track-normalized to -16 LUFS, as in Table 1, its Integrated Loudness will typically measure between -18 LUFS and -16 LUFS, which encompasses the +2 LU tolerance in Table 1. This depends on the relative proportion of speech and music in the stream. It is impractical for some providers, such as commercial radio stations using radio-style online audio processing, to control speech and music separately, so these providers must instead choose a single Distribution Integrated Loudness for the entire stream. Table 2 shows examples of radio-style, podcast, etc. formats and an appropriate integrated loudness for each that will interoperate well with streams where speech and music are treated separately according to Table 1.

Table 2

Format	Distribution Integrated Loudness
News/Talk	-18 LUFS
Pop music	-16 LUFS
Mixed format	-17 LUFS
Sports	-17 LUFS
Drama	-18 LUFS

Providers may further refine and customize these values depending on the proportion of speech and music in their specific streams. A useful formula is:

$$\text{Distribution Integrated Loudness} = -16 - \left[2 \times \left(\frac{SpeechPercentage}{100}\right)\right] \text{ (LUFS)}$$

Because approximately 1 LU is considered a just noticeable loudness difference, this formula will still provide satisfactory results when the percentage of speech is roughly approximated [3].

C. Normalization Workflow

Figure 2 illustrates an example of a loudness normalization workflow from the receipt of audio content to the point of Internet distribution. See Table 1 and Table 2 for tolerances and details.

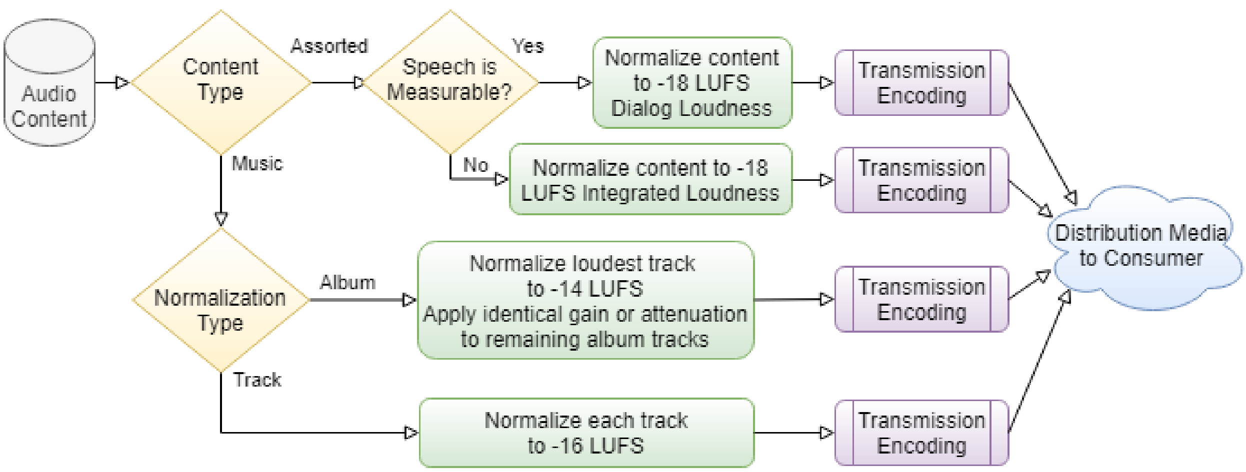

Figure 2 – Example of content delivery workflow, assorted content

Wide Dynamic Range Content

Classical music and fine arts styles have unique requirements due to their large peak-to-loudness ratio and wide dynamic range. For classical music, it is suggested that musical content be album-normalized and that speech content be 3 LU to 6 LU below the loudest parts of the music, depending on the music's style and dynamic range. The provider must choose whether to allow upward normalization, which may compromise quality, or to choose a lower Distribution Loudness to avoid peak limiting.

For example, album-normalized classical music streams with a loudest track at -18 LUFS and speech normalized at -21 LUFS would typically not be upward normalized to avoid peak limiting.

Concerns with Lowered Distribution Loudness

The following may apply to wide dynamic range content with Distribution Loudness lower than this document's recommendations:

- Content with lower loudness may sound too quiet compared to content complying with Table 1. Listeners switching from content complying with Table 1 may need to adjust their volume controls.
- Interstitials (including direct ad insertion) with loudness complying to Table 1 may be too loud and should be attenuated to achieve consistent loudness.
- Lowered loudness can cause content to be unsatisfyingly quiet on some player devices with insufficient gain, particularly when environmental noise is present.
- Dynamic range compression may be needed to allow the quietest passages in the content to be audible. Upward compression is the preferred approach for raising low level passages. When the evolutionary process is complete, using DRC metadata will help address this problem.

D. Music Streaming Services

Music Streaming Services stream separate music and speech tracks on demand. These services commonly store tracks in their originally mastered form and usually operate in a *closed ecosystem* using metadata supported by a dedicated player application (see Section 3: Evolutionary Process). To achieve the intended loudness for listeners, it is recommended that audio-specific metadata indicating the loudness value of each track be provided with the audio content. This metadata is used in the distribution or playback system to normalize loudness.

Music Normalization Techniques

Two methods of music normalization are currently practiced (see Table 1):

Track normalization: Normalize each track to the "track normalized" loudness specified in Table 1.

Album normalization: Normalize the loudest track of the album to the loudness of the "album-loudest track" specified in Table 1. Then apply the same normalization gain or attenuation to the remaining tracks. This preserves the relative loudness between the tracks and maintains the artist's intent.

When operationally feasible it is strongly recommended to use album normalization, even when playing tracks outside the album context, including shuffle play [6]. This is because the quieter tracks of all albums have been mastered to sound correct relative to louder tracks. Their loudness is intentionally lower. Research indicates that listeners prefer to hear tracks at their intended relative level, even outside the album context [7]. An additional advantage of Album Normalization is that different genres become more compatible and play at appropriate loudness: different music genres often can be seamlessly played together in the same playlist. This is because regardless of genre, the mastering engineer produces the quiet tracks of all albums in a similar esthetic proportion to the loudest track.

Except for Wide Dynamic Range Content, it is recommended that distributors normalize the loudest track of an album to -14 LUFS, per Table 1. This will typically result in an integrated album loudness of -16 LUFS, since the loudest tracks are typically 2 LU louder than the average loudness of their albums. This procedure keeps the majority of quiet tracks above -20 LUFS. Although it is recommended that Album Normalization be the default, if album normalization is not possible, it is recommended to normalize all tracks to -16 LUFS.

High resolution content

When high resolution content is intended to provide an audiophile-quality experience, it is recommended to:

- Provide bit-accurate transfer by using a lossless codec and not normalizing before the encoder
- Distribute loudness metadata that supports Loudness Normalization in the player using its designer's preferred gain-scaling method

6 Loudness and Dynamic Range Control Metadata

Metadata is encoded information that describes the essence of digital content streams and media files, including audio, video and photographic images. This document provides limited information about metadata in this first step of the evolutionary process (see Section 3: Evolutionary Process). Some information is included here because there are providers already using metadata and its usage will increase in the future.

There are multiple types of metadata. Audio-related metadata describes the content. Control-related metadata specifies how the audio should be played on a device. Dynamic range control (DRC) and loudness metadata are examples of metadata that can control compliant players, enabling loudness management for various playback scenarios. ANSI/CTA-2075 documents use this metadata. Codecs supporting such metadata include:

- Extended HE-AAC and MPEG-H with MPEG-D DRC metadata
- HE-AAC with MPEG-4 AAC Metadata
- AC-3, E-AC-3, AC-4
- DTS-HD and DTS-UHD

Some online music services implement alternative types of loudness metadata, which may use the ID3 metadata container.

Examples of the Benefits of Loudness and DRC Metadata

- *Loudness metadata* is used by compliant players to normalize content loudness preceding the user's volume control. By allowing the playback device control of loudness normalization, distributors can store and/or distribute content with its original loudness and dynamics. The loudness metadata must be set properly for the distribution system being used so that the recommended Distribution Loudness values (or relative values) in Table 1 are achieved in playback. ANSI/CTA-2075 compliant players will apply peak limiting if needed to prevent their output amplifiers from clipping.
- *Dynamic Range Control metadata* is included with the content at encoding without altering the original audio. It allows a player to optionally reduce the dynamic range of the content it receives, accommodating the capabilities of the playback device, noise environment and/or the listener's choice. Listeners in noisier conditions can set a limited dynamic range for improved audibility, while listeners in quiet environments can choose to enjoy the content's full dynamic range.

As of this writing, loudness metadata is already in use by various services, some using custom metadata systems to implement Track Normalization/Album Normalization and the speech/music loudness offsets indicated in Table 1. Loudness and DRC metadata will be further adopted in an evolutionary process as the numbers of ANSI/CTA-2075-compliant consumer mobile and fixed devices increase. Codec and metadata developers are encouraged to include information in their application notes illustrating how to use the available Loudness metadata to implement this document's recommendations.

7 Technical Notes

A. Interstitial Content

Audiences dislike annoying loudness variations that force them to adjust the volume control when poorly matched interstitial content plays adjacent to their favorite shows and music [7]. Additionally, government regulation sometimes mandates managed loudness normalization. This document leverages already-proven loudness normalization practices. Its recommendations will align the loudness of all content while retaining good sound quality of the material. Implementing these practices can satisfy the needs of content creators, listeners and possible governmental concerns.

B. Peak Control

Peaks generally do not affect a loudness measurement, though they do affect perceived sound quality. A recording with high peak to loudness ratio (PLR) is often perceived as clearer and less fatiguing than one that has been excessively peak-limited. In this discussion, dBTP refers to peak levels measured using a true peak meter.

BS.1770-compliant true-peak measurement requires the audio signal to be sampled at 192 kHz or higher to approximate the true-peak level following D/A conversion or sample rate conversion. If the original sample rate is less than 192 kHz, this requires upsampling. True-peak meters typically have an error of less than 0.6 dB, assuming an ideal D/A converter with a linear-phase reconstruction filter.

Peak limiting may be required under at least four circumstances, which may apply simultaneously when:

- Incoming audio level must be increased to achieve the desired Distribution Loudness
- Peak level must be lowered to accommodate overshoots induced by lossy encoding
- Audio is applied to a filter
- Audio is applied to a sample rate converter

Some of this may occur downstream in the transmission chain, outside the content distributor's control, requiring the threshold of the content provider's peak limiter to be set based on an estimate of the amount of peak overshoot that downstream elements introduce.

If peak limiting is required to meet this document's recommended Distribution Loudness, the distributor must trade off possible limiter-induced audio degradation against the possibility that reducing the loudness of such material will cause one or more of the problems set forth in Section 4C: Normalization Workflow - Wide Dynamic Range Content.

When peak limiting is required, use a true-peak limiter, which anticipates and controls the peak level after the player device's digital-to-analog converter and reconstruction filter.

Sources of Peak Overshoot

Codecs

Safety limiting should take into account the peak signal level anticipated at the decoder's output, which may be higher than the peak level at the encoder's input. High-rate (e.g., 256 kbps) coders may work satisfactorily with as little as -0.5 dBTP for the limiting threshold. Typically, peak overshoot increases as the bit rate decreases, so the limiting threshold may need to be reduced below the recommended -1.0 dBTP.

Filters

Lowpass, highpass, and bandpass filtering can all add overshoots by removing program energy and if not linear-phase, by introducing group delay distortion. Hence, filtering (if used) should ideally occur before the peak limiter. While it may seem counterintuitive that removing program energy can increase peak levels, consider a square wave. Its fundamental is a sinewave whose peak level is 2.1 dB higher than the flat-top peak level of the square wave. Lowpass-filtering all the harmonics from the square wave will thus increase its peak level by 2.1 dB.

If the signal path after the peak limiter has a high pass characteristic (as do most analog paths), the −3 dB frequency must be below 0.15 Hz to prevent the path from introducing more than 0.1 dB of overshoot.

Sample Rate Converters

Sample rate conversion (SRC) may be upward SRC, downward SRC, or a sequence of both operations. A useful conceptual model of SRC is digital-to-analog conversion followed by resampling the resulting analog signal, which will require an anti-aliasing filter if the target sample rate is lower than the source sample rate.

Except for trivial cases (such as downsampling by a factor of two when the input spectrum is known to be band-limited to half of the output sample rate), SRC requires a specialized form of filtering that is closely related to the filtering required for analog-to-digital and digital-to-analog conversion.

SRC does not usually preserve the values of the input samples. Using the conceptual model described above, it is clear that the problem of controlling peak levels during SRC is the same as controlling the peak levels at the output of a D/A converter. These can also be solved with a true-peak limiter before the SRC. However, there are two caveats:

- Filter-induced group delay distortion (which minimum-phase filters can introduce) will modify the shape of the audio waveform by changing the time relationship between energy at different frequencies. The true-peak limiter will then be unable to predict the resulting peak level. Hence, the SRC's filtering must have constant group delay (linear phase) in its passband and transition region, which is readily and commonly achieved.
- SRC must not remove input program energy. Downward SRC can remove program energy and thus induce overshoot (as discussed in the section on Filters above). True-peak limiting prior to the SRC cannot prevent this overshoot; in this case, placing a true-peak limiter *after* the SRC is the only way to reliably prevent peak overshoot. However, this is often outside the content distributor's control.

Guidelines for Peak Limiter Use and Setup

- To protect the consumer D/A converter from clipping and to accommodate possible downstream SRC, always include a true-peak limiter as the last element before the codec, even if it is lossless. This is because older content that was mastered prior to the advent of true-peak limiting may have true-peak overshoots baked in.
- Set the threshold of the limiter to accommodate anticipated downstream overshoot sources, including the codec and any downward SRC. Codec and downward SRC overshoots can be

additive. Because the detailed mechanisms are complex, the ideal guide to setting the limiter threshold is actual true-peak measurement of the overall signal path from the peak limiter's output to the consumer's D/A converter, combined with listening tests.

A maximum of 1 dB of limiting prior to final encoding is recommended as a starting point. Do this only if frequently-occurring peaks cause audible distortion or artifacts at the decoder. Use the ear as the final judge, because limiting more than about 1 dB may produce more audible artifacts than simply letting the audio material clip on occasional transients. However, when judging, beware of Players' Handling of True-Peak Overloads**,** discussed below. Peak limiting, not clipping, may be occurring in the player. Different player designs can sound different when handling peak overshoots, so it is wise to audition several types of players.

- If practical, specify a Distribution Loudness low enough to prevent potential clipping or limiting by the distributor or the player. However, if this would cause the Distribution Loudness to fall below the recommendations in Table 1, take into account the caveats regarding "Lowered Distribution Loudness" in the "Wide Dynamic Range Content" section above.

Players' Handling of True-Peak Overloads

When deciding whether to allow some true-peak overload at the player, be aware of several aspects of player design:

- Some players prevent clipping by including peak limiters. Each limiter design has its own sound, making audio quality unpredictable when content distributors allow true-peak overshoots and target several types of players.
- Many mobile devices have built-in peak limiters that prevent the power amplifiers in battery-operated mobile devices from clipping. Hence, content with a high peak-to-average ratio (such as a −23 LUFS classical music stream) may activate these peak limiters, with unpredictable subjective results.
- Some players have decoders that use fixed-point arithmetic. These decoders can clip internally unless designed with sufficient headroom to accommodate true-peak overshoots.
- Some common operating systems (like Microsoft Windows®, Vista® or higher) have built-in peak limiters that produce gain reduction on codec overshoots that would otherwise cause clipping in a float-to-fixed-point conversion following the (floating point) decoder. Consequently, failure to provide sufficient headroom for codec overshoots on the transmission side can produce as much as 3 dB of gain reduction in the decode-side peak limiter.

C. Mono/Stereo Downmix

When considering how loudness control applied to a source affects the loudness of a downmix to fewer channels (for example, stereo to mono), it is important to understand the difference between acoustic and electrical summation of program elements in stereo or multi-channel mixes. In most listening rooms, loudness of a given element in the mix is best represented as the sum of the power produced by that element in each loudspeaker because room acoustics tend to randomize the relative phase between the elements at the listening position. Following this principle, the BS.1770 algorithm uses RMS channel summation.

For example, when heard in stereo, an element present equally in the left and right channels of a stereo recording will sound about 3 dB louder than either channel by itself. On the other hand, when the stereo channels are electrically summed to mono, elements add arithmetically, so in this example the element will be 6 dB louder in the mono mix and the relative balance of the elements in the mono sum will not be the same as the stereo; the balance will change according to the element's placement in the stereo panorama. This means that loudness control applied to a stereo mix can have errors as large as 3 LU when heard in mono.

A partial solution to this dilemma in stereo-to-mono downmixing is to pass the stereo source through a 90-degree phase difference network (mathematically termed a "Hilbert transformer") in the player device before summation to mono. Any in-phase elements common to the left and right channels will add to produce the same loudness balance between the elements as in the original stereo, so loudness control applied to the stereo will also work correctly in mono. This is best done by using a special FIR (finite-impulse-response) digital filter, where one channel is delayed by the filter's group delay and the other channel is passed through the filter. Compared to using a pair of "allpass" filters (each with non-constant group delay) for the left and right channels, this technique minimizes phase distortion, yet the filter's impulse response will still somewhat smear transient elements in the mix. The tradeoff in this technique is thus preserving subjective balances in the stereo mix against preserving the integrity of transients. Additionally, player device developers and manufacturers should be aware that the Hilbert Transform process is likely to increase peak levels and the player's signal path must have enough headroom to accommodate this.

D. Single-Channel (Monophonic) Production, Distribution and Playback

Single-channel (mono) content can be a source of loudness inconsistencies, depending on how it is accommodated in the player or device. This issue affects distributors, player software developers, and device manufacturers.

Considerations for distributors

When content is distributed in a single-channel mode, the player device needs to ensure that its loudness is properly aligned with content distributed in multi-channel mode (usually stereo). When the single-channel content is reproduced in multi-channel, the device plays the content through both stereo channels and may play through all front surround channels. This increases loudness. For example, two-channel reproduction will measure 3 LU higher when both channels carry the same content.

To prevent content from sounding too loud, modern devices and player software rely on channel-mode metadata to apply the correct attenuation. It is recommended that distributors include such metadata in their single-channel audio streams so that players respond properly. The measured multi-channel loudness will then be equal to the measured loudness of the original single-channel.

Distributors transcoding single-channel content into a multi-channel format prior to distribution should compensate for the increased loudness caused by the conversion. This can be done by measuring the content loudness after the conversion and using the result, either in loudness metadata or by loudness normalization of the content via a gain change.

Considerations for player developers and device manufacturers

Developers of players and device manufacturers can help remedy single-channel loudness issues. Though it is beyond the scope of this document to make specific recommendations for players and devices, some general best practices follow:

Multi-channel reproduction of single-channel content

When the single-channel content is reproduced in multi-channel, the device plays the content through both stereo channels and may play through all front surround channels. This increases its loudness.

Players can compensate by responding to the channel-mode metadata in both single-channel and multi-channel streams. The measured multi-channel loudness should then be equal to the measured loudness of the original single-channel. If the metadata is not handled by the decoder, the general approach is to reduce the level of the decoder output by 3.01 dB before passing it to the left and right channel outputs (stereo) or 4.77 dB before passing it to the left, center and right channel outputs (surround). See Figure 3.

Single speaker reproduction (e.g., smart speakers) of single-channel and multi-channel content

Properly reproducing content through a single speaker (such as a smart speaker) is more complicated, requiring developers to consider both the channel mode of the content (single-channel or multi-channel) and Mono/Stereo Downmix issues.

For example, consider a single-channel smart speaker that must consistently reproduce:

- Correctly signaled single-channel content
- Single-channel content distributed in stereo form, where the distributor duplicated the content on two channels before measuring its loudness and distributing it, and
- True stereo content

Because electrically summing two identical (fully-correlated) channels increases loudness by 6 LU, these scenarios will produce different outcomes when downmixed for use on a single speaker reproducer.

- In scenario 1, assume that the measured loudness of the single-channel has a fixed relationship to the acoustically reproduced loudness from the single speaker. Let this acoustic loudness be a reference for comparing the loudness produced by scenarios 2 and 3.
- In scenario 2, the measured loudness of the content agrees with the acoustic loudness produced by multi-channel playback, but due to the full correlation, the content will sound 3 LU too loud when reproduced through the single speaker.
- In scenario 3, content centered in the stereo panorama and reproduced through a single speaker again sounds 3 LU too loud (due to full correlation) compared to left-only or right-only content. This alters the original balance between the mix elements (as explained in Mono/Stereo Downmix).

There are at least two approaches for addressing these scenarios in the player. Both approaches require the player to respond to channel-mode metadata.

- An approach that optimizes all scenarios at the expense of complexity is to first duplicate correctly signaled single-channel content to both channels as if the device were playing in stereo. As described above, this typically requires attenuating such content by 3 dB, which the decoder may do automatically. All content is thus in two-channel form prior to downmix. If it were downmixed with no further processing, the mono part of all content would be up to 3 LU louder than the true stereo part of the content (scenario 3). To remedy this, phase-shift one audio channel by 90 degrees with respect to the other before downmixing, as described in Mono/Stereo Downmix. As explained there, this approach preserves the intended loudness balance of elements within true stereo content and makes the loudness of the downmix track the loudness of the stereo content.
- The simplest approach is to apply 3 dB of extra gain to content signaled as single-channel before it is transmitted to the single speaker reproducer. As explained, the loudness of downmixed true stereo material can be up to 3 LU soft in this case.

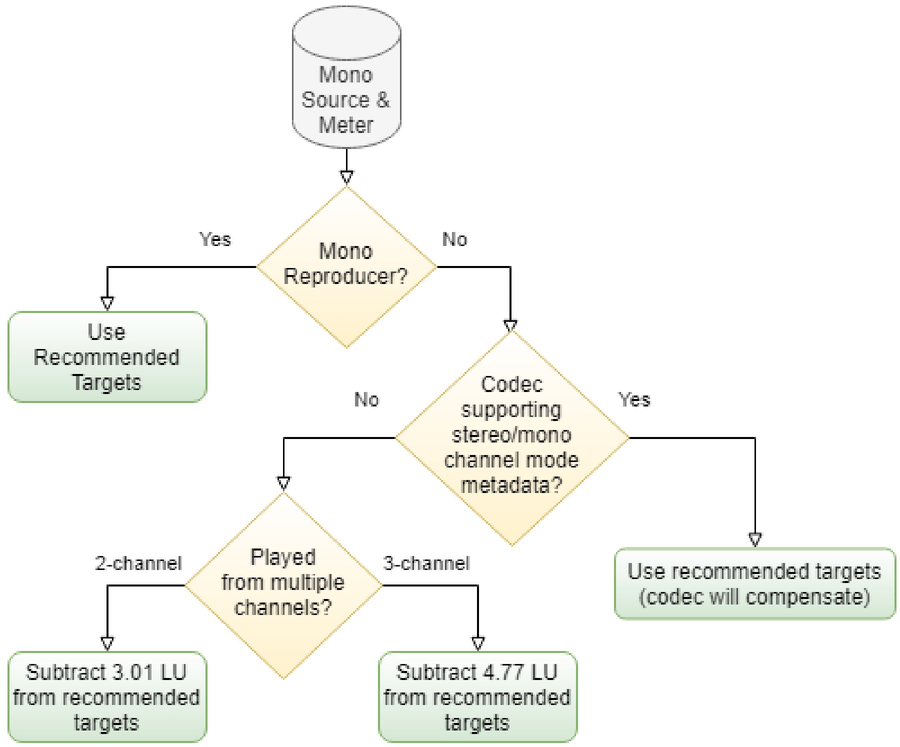

Figure 3 - Adjustment procedures when single-channel material is played out of 1, 2 or 3 speakers

E. Loudness, System Sounds and the Virtual Assistant

To achieve a pleasing listening experience, application developers should strive to match the loudness of virtual assistants and other system-generated sounds to surrounding audio content whenever appropriate. Notwithstanding, AES71 recognizes that additional factors can influence loudness needs, such as: distinguishing incoming “rings” over content and background noise, localizing sound to picture, matching the user’s desired conversational loudness, maintaining intelligibility, meeting accessibility needs, and delivering emergency information (e.g., Wireless Emergency Alerts-WEA), etc. An increase (or in certain cases reduction) to the loudness of virtual assistants and other system-generated sounds with respect to other audio content is expected.

F. Portable Media Players and Hearing Loss

If Portable Media Players (PMPs) are played too loudly for too long, they will cause hearing loss. The Scientific Committee on Emerging and Newly Identified Health Risks estimates that between 2.5 and 10 million people in the EU are at risk of developing early hearing loss as a result of listening to PMPs [8]. Europe was the first region to implement regulations to protect the hearing of PMP users: any PMP and headphone sold in Europe must comply with CENELEC EN 50332. In the original 2013 version, compliance was achieved by limiting the maximum gain of the device. This successfully reduced the maximum SPL from PMPs sold across Europe, but prevented quieter audio content from being played loudly enough to be heard under demanding listening conditions. A second adverse effect was that content louder than the EN 50332 Reference Test Signal still played at dangerously loud levels.

CENELEC TC108X/WG3 acknowledged these problems and addressed them in the 2018 revision EN 50332-3. ITU has adopted their method in ITU-T H.870. These documents recommend implementing a sound dose meter in the device that estimates the actual SPL that the user receives. Devices can then have enough gain to play -24 LUFS content at satisfying acoustic levels, while at the same time protecting PMP users' hearing. ANSI/CTA 2075 encourages device manufacturers to implement the EN 50332-3 / ITU-T H.870 recommendations.

Until EN 50332-3-compliant players are ubiquitous, European PMPs may have insufficient gain to allow satisfactory playing of material with Distribution Loudness below −20 LUFS. The recommendations in this document are written with this transition phase in mind.

Annex A
Terms, Definitions and Abbreviations

Album Normalization: A method of Loudness Normalization applied to entire mastered albums. It preserves the relative loudness balance between tracks, respecting the artists' intent. Album normalization consists of measuring the Integrated Loudness of each track, then calculating the gain or attenuation necessary to normalize the loudest track to a specified Distribution Loudness and applying that gain or attenuation equally to all tracks.

Anchor: Within content, the genre (often speech) of the parts used to calculate loudness normalization. The resulting normalization gain or loss is then applied uniformly to the content. This process conveys the anchor elements at the listener's preferred loudness while preserving the source's original dynamics. For example: "We normalized this file to -18 LUFS using a speech anchor." See Speech Loudness.

BS.1770 Loudness: An electrical measurement using the ITU-R BS.1770 algorithm. This estimates the perceived loudness of content, referenced to 0 dBFS, in a unit of "LKFS." The acronyms LUFS (Loudness Units with respect to digital Full Scale) and LKFS (K-weighted Loudness with respect to digital Full Scale) are interchangeable and specify identical values. The algorithm consists of a K-weighting filter feeding a level-gated RMS detector followed by a logarithmic converter scaled such that a loudness change of 1 LU corresponds to a level change of 1 dB. The first stage of the filter is a high frequency 4dB shelving boost that accounts for the acoustical effects of the head. The second stage is a low-frequency roll off based on "B-weighting," which roughly approximates the 70 phon equal-loudness contour (see ISO 226:2003).

dBFS; dB FS; (decibels full-scale): RMS amplitude of digital samples expressed as a level in decibels relative to a sine wave whose peak level is 100% full-scale and whose frequency is asynchronous with the sample frequency. [per AES17]

dBTP; dB TP; (decibels True Peak): decibels, True Peak relative to 100% full-scale (per ITU-R BS.1770 Annex 2)

Dialog Integrated Loudness (also known as Speech Loudness or Dialog Loudness): See Speech Loudness

Distribution Loudness: The intended and/or actual Integrated Loudness or Speech Loudness of a given piece of distributed content or an entire stream, depending on the normalization method employed (see Table 1 and Table 2). The recommendations in this document apply to content distribution and not content creation. Therefore, this document uses the specific term "Distribution Loudness" instead of the generic "Target Loudness" where needed to avoid confusing content creators.

Downward Normalization: Applying a uniform amount of attenuation to a piece of content to bring its Integrated Loudness down to the desired Distribution Loudness.

Dynamic Range: As used in this document, the difference in Loudness between the loudest and softest passages of content, excluding silence.

Dynamic Range Control (DRC): The process of continually adjusting audio signal level to control the loudness difference between loud and soft passages. This can help overcome noise in the listening environment, meet the dynamic range capability of the playback equipment, or both. The term "dynamic range control" typically implies that the codec's encoder generates the gain control signal as metadata and that the player device's decoder acts on this metadata. While this can also be described as "dynamic range compression," that term is more commonly used when a single device (like the player) generates the gain control signal from an internal algorithm applied to the incoming audio and does not exploit DRC metadata.

File: Storage of audio content as digital media on a computer or server. A file may contain a single element like a song or D.J. announcement, or may contain a complete audio asset such as a radio show, podcast, music album, etc. These files may support the online transfer of these audio assets to the listener for on-demand playout.

Format: As used in this document, a name such as "contemporary hit radio" or "news" that describes the type or style of given long-form content. A stream often has only one format, but some netcasters transmit different formats at different times ("mixed-format" streams).

Formats are often built from elements that are related to each other and presented sequentially, but that have different genres, such as music, studio-quality DJ announcements, and noisy, limited-bandwidth speech from a helicopter-based traffic reporter.

Genre: A name, such as "speech," "popular music," "classical music," etc., characterizing a long-form element that has a homogenous texture and style.

Headroom: As used in this document, the difference between the maximum level of the signal and the maximum level that can be represented in, or conveyed by, the system.

High resolution content: As used in this document, any content that is linearly or losslessly encoded. As used generally, losslessly encoded content that exceeds CD sample rate and/or bit depth.

Integrated Loudness: The average loudness between two points in time, measured electrically using the BS.1770 loudness algorithm with gating. To estimate the Integrated Loudness of a continuous stream, it is necessary to choose a sliding rectangular integration time window long enough to prevent the measurement from being disturbed by short-term dynamics in the content.

ID3 Tagging (ID3 Metadata): A type of metadata container used to store information about the audio; this information is embedded within the audio file. Use of an ID3 tag allows a file to include relevant information such as Now Playing information (for example: artist name, track title, album, track number and genre) and can be used to access other databases. ID3 metadata currently supports loudness management and can be extended further.

Interstitial content; Interstitial: Short form content — commercial advertising, commercial, promotional or public service-related material. The typical duration is less than approximately two to three minutes.

LKFS: Loudness K-weighted with respect to digital Full Scale. Identical to LUFS: except for nomenclature. See **BS.1770 Loudness**

Long-form Content: "Show or program-related material or essence. The typical duration is greater than approximately two to three minutes." [per ATSC A/85] Examples include radio shows, concerts, dramas and musical albums.

Loudness: A subjective metric of the perceived audio level of a program or other piece of content. No simple physical property like amplitude, fully describes it. Instead, it is determined in a complex way by the content's amplitude and spectrum, and how these vary over time. When all else is held equal, it is approximately proportional to amplitude within the acoustic loudness range that audiences prefer when actively listening to content, which is approximately 65 to 80 phon. This is the range over which BS.1770 is most applicable.

Loudness Management: A process that adjusts the Integrated Loudness of content or loudness metadata to provide loudness consistency. A distributor can implement this by using Loudness

Normalization. Loudness management can also include Dynamic Range Control (DRC). The distributor can implement this by irreversibly applying dynamic range compression or alternatively, by including DRC and loudness metadata with the original content. The latter preserves the content and allows the consumer and/or device to set the amount of dynamic range compression.

Loudness Normalization: As used in this document, a process that applies a fixed gain or attenuation to content so that its loudness matches a specified Distribution Loudness. Unless the content already meets this goal, this process requires modifying the audio data. This can also be performed non-destructively by applying metadata.

LU (Loudness Units): A unit of measure defined in ITU-R BS.1771 as the difference between two ITU-R BS.1770 measurements and/or specified target or reference loudness values. A loudness change of 1 LU corresponds to a level change of 1 dB. For example, "this file's Integrated Loudness is 3 LU above the target, so applying 3 dB of attenuation puts the file on target."

LUFS: Loudness Units with respect to digital Full Scale. Identical to LKFS except for nomenclature. See BS.1770 Loudness

Maximum True Peak: The largest True Peak in the content under consideration. This indicates whether the signal reaches amplitudes that could overload the following stages of an audio system.

Metadata; Audio Metadata: Non-audio data included in audio files used to identify, label and present audio content. Metadata includes information such as content loudness, artist, genre, label, song titles, album name and track numbers.

Normalization: See Loudness Normalization. Normalization can also refer to Peak Normalization, which provides poor loudness consistency and is not recommended.

On-Demand File Playback: Playout or streaming of a File through interconnected audio facilities, usually between the point of distribution and the consumer (listener).

Over-the-Top-Television (OTT): To deliver video content via Internet streaming, VOD, pay TV, IPTV and download via IP mechanisms.

Parallel compression: A dynamic range compression technique used in sound recording and mixing to implement upward compression. It is accomplished by mixing an unprocessed audio signal with a compressed version of the same signal. (If the compressor has enough delay to cause audible comb-filtering of the sum, the unprocessed signal must be delayed to match it.) As the unprocessed signal becomes louder, the compressor's gain reduction increases and it contributes progressively less to the mixed output. Therefore, parallel compression amplifies quiet sounds while preserving most of the dynamics of loud sounds. This is dissimilar to conventional inline compression, which compresses the loudest sounds downward. Parallel compression is particularly useful for classical music, allowing quiet passages to be heard clearly in noisy environments while preserving the impact of loud passages.

Peak Normalization: The practice of adjusting the peak level of audio content to (usually) full scale. This produces widely-varying loudness levels. This practice has been cited as the original cause of loudness wars in broadcast and music production.

PLR; Peak to Loudness Ratio: The ratio of Maximum True Peak level of an audio content segment to its integrated BS.1770 Loudness.

Podcast: Episodic audio file accessed via an on-demand system or download for local file playback, available through a subscription.

Short-Form Content: See Interstitial Content

Speech Loudness (also known as Dialog Integrated Loudness): The BS.1770 Integrated Loudness of pure spoken voice content that is not mixed with other elements such as a music bed. The measurement requires isolating speech segments of content either manually, or by measuring the dialogue stem, or with an automatic speech detection tool.

Streaming: A continuous transmission over a network (typically the Internet) that consists of pieces of content presented sequentially.

Streamer: A content provider offering a streaming service to customers.

Target Loudness: The intended Integrated Loudness or Speech Loudness of a given piece of content or an entire stream, depending on the normalization method employed (see Table 1 and Table 2). "Target Loudness" is a generic term. The recommendations in this document apply to content distribution and not content creation. Therefore, this document instead uses the more specific term "Distribution Loudness" where needed to avoid confusing content creators.

Track Normalization: A method of Loudness Normalization applied independently to each music track, so each track plays at the same BS.1770 Loudness. This may conflict with the creative intent of the artists. (See Album Normalization.)

True Peak Level: Amplitude of peaks (positive or negative) of a digitized signal in the continuous time domain after the signal has been applied to a mathematically ideal digital-to-analog converter and reconstruction filter. True Peak Level is expressed in units of dBTP: decibels with respect to the peak level produced by a digitized continuous sinewave signal whose maximum sample level is digital full-scale and whose frequency is asynchronous with the sample frequency. True Peak Level can be estimated in the digital domain by upsampling the signal to at least 192 kHz and measuring the peak sample values. Upsampling to higher than 192 kHz will reduce the estimation error. See ITU-R BS.1770 Annex 2, particularly the Table in Appendix 1 to Annex 2, which shows the maximum estimation error as a function of the upsampling ratio.

True Peaks (aka "intersample peaks") can be substantially higher than the level of digital samples engendering them and can exceed 0 dBTP. This is common in older content mastered before the advent of true-peak-aware limiters.

Upward compression: Upward compression amplifies low-loudness audio while mostly preserving the dynamics of high-loudness audio. It can be achieved by appropriately shaping the input/output gain curve of the compressor so that the amplification is highest at low input levels and approaches 0 dB at high input levels. It can also be achieved by mixing the output of a downward compressor with its input, a technique known as parallel compression.

Upward normalization: Applying a uniform amount of positive gain to a piece of content to bring its Integrated Loudness up to the desired Distribution Loudness. It may require peak limiting to prevent peak levels from exceeding 0 dBTP.

Annex B
Related Loudness Standards and Recommendations

AES71-2018, "Recommended Practice Loudness Guidelines for Over-the-Top Television and Online Video Distribution," July 2018. https://www.aes.org/publications/standards/search.cfm?docID=107

ANSI/CTA-2075, "Loudness Standard for Over-the-Top Television and Online Video Distribution for Mobile and Fixed Devices," https://shop.cta.tech/products/loudness-standard-for-over-the-top-television

ATSC A/85, "ATSC Recommended Practice: Techniques for Establishing and Maintaining Audio Loudness for Digital Television," http://atsc.org/wp-content/uploads/2015/03/Techniques-for-establishing-and-maintaining-audio-loudness.pdf

CENELEC - EN 50332-2, "Sound system equipment: Headphones and earphones associated with personal music players - Maximum sound pressure level measurement methodology - Part 1: General method for 'one package equipment'", 2013, https://standards.globalspec.com/std/1641364/EN%2050332-2

CENELEC - EN 50332-3:2017, "Sound system equipment: headphones and earphones associated with personal music players - Maximum sound pressure level measurement methodology - Part 3: Measurement method for sound dose management," 2018, https://www.en-standard.eu/csn-en-50332-3-sound-system-equipment-headphones-and-earphones-associated-with-personal-music-players-maximum-sound-pressure-level-measurement-methodology-part-3-measurement-method-for-sound-dose-management/

EBU R 128, "Loudness Normalisation and Permitted Maximum Level of Audio Signals," August 2020, https://tech.ebu.ch/docs/r/r128.pdf

EBU R 128 s1, "Loudness Parameters for Short-Form Content (Adverts, Promos, etc.)," August 2020, https://tech.ebu.ch/docs/r/r128s1.pdf

EBU R 128 s2, "Loudness in Streaming," August 2020, https://tech.ebu.ch/docs/r/r128s2.pdf

EBU - Tech 3341, "Loudness Metering. 'EBU Mode' Metering to supplement loudness normalization in accordance with EBU R 128," https://tech.ebu.ch/docs/tech/tech3341.pdf

EBU - Tech 3342, "Loudness Range: A measure to supplement loudness normalisation in accordance with EBU R 128," https://tech.ebu.ch/docs/tech/tech3342.pdf

EBU - Tech 3343 version 3.0, "Practical guidelines for Production and Implementation in accordance with EBU R 128," January 2016, https://tech.ebu.ch/docs/tech/tech3343.pdf

EBU - Tech 3344 version 2.1, "Practical guidelines for distribution systems in accordance with EBU R 128," July 2016, https://tech.ebu.ch/docs/tech/tech3344.pdf

ITU-R BS.1771-1, "Requirements for loudness and true-peak indicating meters," January 2012, https://www.itu.int/rec/R-REC-BS.1771-1-201201-I/en

ITU-R BS.1770, "Algorithms to measure audio Program Loudness and true-peak audio level," https://www.itu.int/rec/R-REC-BS.1770/en

ITU-T H.870 - SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS, "E-health multimedia systems, services and applications – Safe listening,"2018, https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-H.870-201808-I!!PDF-E&type=items

Annex C
References

[1] J. Riedmiller, et.al. "Intelligent Program Loudness Measurement and Control: What Satisfies Listeners?" *presented at the 115th AES Convention* (October 2003), paper 5900, https://www.aes.org/e-lib/browse.cfm?elib=12415

[2] E. Benjamin, "Comparison of Objective Measures of Loudness Using Audio Program Material," paper 5703, *presented at the 113th AES Convention* (October 2002), paper 5703, https://www.aes.org/e-lib/browse.cfm?elib=11300

[3] K. Persons, et.al "Speech Levels in Various Noise Environments," U.S. Environmental Protection Agency (revised December 2005), document ID EPA/600/1-77/025, https://cfpub.epa.gov/si/si_public_record_Report.cfm?Lab=ORD&dirEntryID=45786

[4] F. Begnert, H. Ekman, J. Berg, "Difference between the EBU R-128 Meter Recommendation and Human Subjective Loudness Perception," *presented at the 131st AES Convention* (October 2011), paper 8489, http://www.aes.org/e-lib/browse.cfm?elib=16015

[5] S. Norcross, "Using ITU-R BS.1770 to Measure the Loudness of Music versus Dialog-based Content," *presented at the 149th AES Convention* (2020 October), paper 10448, https://www.aes.org/e-lib/browse.cfm?elib=20985

[6] E. Grimm, "Analyzing Loudness Aspects of 4.2 million Music Albums in Search of an Optimal Loudness Target for Music Streaming," *presented at the 147th AES Convention* (October 2019), paper 10268, https://www.aes.org/e-lib/browse.cfm?elib=20641

[7] J. Riedmiller, S. Lyman, C. Robinson, "Intelligent Program Loudness Measurement and Control: What Satisfies Listeners?" *presented at the AES 115th Convention* (October 2003), paper 10268, https://www.aes.org/e-lib/browse.cfm?elib=20641

[8] T. Lund, "Prevention of Hearing Loss from the Use of Personal Music Players" *presented at the 59th International Conference: Music Induced Hearing Disorders* (June 2015), paper 6-1, http://www.aes.org/e-lib/browse.cfm?elib=17796

Annex D
Bibliography

"Adjusting Anchor Loudness," *online article by Apple Inc.* (2020), https://developer.apple.com/documentation/http_live_streaming/adjusting_anchor_loudness (Apple Developer article provides a model to improve accuracy of speech-gated loudness measurement when speech activity is low.)

C. Bertin, "Standards for safe listening devices: situation analysis," *World Health Organization* (2015), https://www.itu.int/en/ITU-T/Workshops-and-Seminars/safelistening/Documents/Standards_for_safe_listening_devices_situation_analysis_report.pdf, (Compiled in collaboration with the WHO and ITU.)

R. Byers, "The Audio Producers Guide to Loudness," *online article by Transom.org* (August 2021), https://transom.org/2021/the-audio-producers-guide-to-loudness/ (Discussion of best practices for loudness measurement and normalization in audio storytelling production.)

R. Byers, "Podcasting Basics: Loudness for Podcasts vs. Radio," *online article by Transom.org* (May 2016), https://transom.org/2016/podcasting-basics-part-5-loudness-podcasts-vs-radio/ (Discussion of differences in using loudness measurement and normalization in radio and podcast production.)

J. Kean, E. Sheffield, "Study of Audio Loudness Range for Consumers in Various Listening Modes and Ambient Noise Levels," *presented at the NAB Broadcast Engineering Conference* (April 2015), https://www.aes.org/technical/documentDownloads.cfm?docID=523 (Investigation of dynamic range desired by listeners in range of environmental noise conditions with speakers and headphones.)

J. Starzynski, "Digital Television Audio Loudness Management", *National Association of Broadcasters Engineering Handbook, 11th Edition* (2018), pp. 767-790

J. Kean "Audio Signal Analysis", *National Association of Broadcasters Engineering Handbook, 11th Edition* (2018), pp. 1767-1794

R. Orban, "Transmission Audio Processing", *National Association of Broadcasters Engineering Handbook, 11th Edition* (2018), pp. 1059–1085, also "Measuring and Controlling Loudness" pp. 1066-1067, "Processing for Surround and Digital Television" pp. 1078–1080, and "Online Processing for Digital Television" pp. 1080-1081

G. Ogonowski, "Normalized Audio and 0 dBFS+ Exposure," *online article of Modulation Index, LLC* (October 2020), https://www.indexcom.com/tech/zerodbfsplus/

"ContentDepot: PRSS Audio Loudness Standard," *listing of articles for the Public Radio Satellite System* (2014–2015), https://prss.org/loudness (Various guidelines and measurement tool for NPR and public radio producers and engineers.)

E. Skovenborg, S. Nielsen, "Evaluation of Different Loudness Models with Music and Speech Material," paper 6234, 117th *presented at the AES Convention* (October 2004), https://www.aes.org/e-lib/browse.cfm?elib=12891

E. Skovenborg, T. Lund, "Loudness Descriptors to Characterize Programs and Music Tracks," *presented at the 125th AES Convention* (October 2008), paper 7514, http://www.aes.org/e-lib/browse.cfm?elib=14666 (Authors discuss subjects' loudness range tolerance: 50% of subjects react to +4/-6 LU systematic changes and 95% of subjects to do so for +6/-8 LU.)