

STANDARDS AND INFORMATION DOCUMENTS

AES77-2023



STANDARDS

AES Recommended Practice Loudness Guidelines for Internet Audio Streaming and On-Demand Distribution

Users of this standard are encouraged to determine if they are using the latest printing incorporating all current amendments and editorial corrections. Information on the latest status, edition, and printing of a standard can be found at: <http://www.aes.org/standards>

AUDIO ENGINEERING SOCIETY, INC.
132 East 43rd St., Suite 405, New York, NY 10017. US.



The AES Standards Committee is the organization responsible for the standards program of the Audio Engineering Society. It publishes technical standards, information documents and technical reports. Working groups and task groups with a fully international membership are engaged in writing standards covering fields that include topics of specific relevance to professional audio. Membership of any AES standards working group is open to all individuals who are materially and directly affected by the documents that may be issued under the scope of that working group.

Complete information, including working group scopes and project status is available at <http://www.aes.org/standards>. Enquiries may be addressed to standards@aes.org

The AES Standards Committee is supported in part by those listed below who, as Standards Sustainers, make significant financial contribution to its operation.



audio-technica



WEISS



This list is current as of
2022/2/18

AES Recommended Practice Loudness Guidelines for Internet Audio Streaming and On-Demand Distribution

Published by
Audio Engineering Society, Inc.
Copyright ©2023 by the Audio Engineering Society

Abstract

This AES Recommended Practice (RP) provides comprehensive recommendations establishing and implementing an effective Distribution Loudness for streaming and on-demand audio file playback. It is intended for use by distributors of Internet audio streams and on-demand audio files.

Implementing these guidelines will provide consistent Loudness and appropriate playback loudness range, will reduce audio quality degradation from excessive limiting, will preserve the original artistic intent, and will improve the listening experience. This document does not recommend Loudness Range (LRA), device target playback loudness or device dynamic range.

An AES standard implies a consensus of those directly and materially affected by its scope and provisions and is intended as a guide to aid the manufacturer, the consumer, and the general public. The existence of an AES standard does not in any respect preclude anyone, whether or not he or she has approved the document, from manufacturing, marketing, purchasing, or using products, processes, or procedures not in agreement with the standard. Prior to approval, all parties were provided opportunities to comment or object to any provision. Attention is drawn to the possibility that some of the elements of this AES standard or information document may be the subject of patent rights. AES shall not be held responsible for identifying any or all such patents. Approval does not assume any liability to any patent owner, nor does it assume any obligation whatever to parties adopting the standards document. Recipients of this document are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation. This document is subject to periodic review and users are cautioned to obtain the latest edition.

Audio Engineering Society Inc. 697 Third Ave, Suite 405. New York, NY 10017, US.

www.aes.org/standards standards@aes.org

2023-07-11 printing

 **Documents**
Document Sharing Hub

Contents

0 Introduction	4
0.1 General	4
0.2 Patents	4
0.3 Documentation conventions	4
1 Scope	4
2 Normative references	5
3 Terms, definitions and abbreviations	6
4 Introduction	11
5 Recommendations	11
5.1 General	11
5.2 Assorted content when speech is measurable	11
5.3 Assorted content when speech is not measurable	11
5.4 Music content which is Track-normalized	12
5.5 Music content which Album-loudest track normalized	12
5.6 Interstitial Content	12
5.7 Virtual Assistant messages	13
5.8 Loudness recommendation summary	13
Annex A (Informative) Normalization Principles	14
A.1 Loudness and Normalization	14
A.2 No Separable Speech Anchor	15
A.3 Normalization Workflow	15
A.4 Music Streaming Services	17
Annex B (Informative): Loudness and Dynamic Range Control Metadata	18
B.1 Understanding Metadata	18
B.2 Examples of the Benefits of Loudness and DRC Metadata	18
Annex C (Informative): Technical Notes	19
C.1 Interstitial Content	19
C.2 Peak Control	19
C.3 Mono/Stereo Downmix	21
C.4 Single-Channel (Monophonic) Production, Distribution and Playback	21
C.5 Loudness, System Sounds and the Virtual Assistant	23
C.6 Portable Media Players and Hearing Loss	24
Annex D Informative References	25
Annex E Bibliography	27

Foreword

This foreword is not part of the *Loudness Guidelines for Internet Audio Streaming and On-Demand Distribution*.

The Streaming Loudness subcommittee was formed in 2018 to study issues related to streaming audio Loudness variations. Its goal is to develop comprehensive recommendations, providing effective guidelines for managing audio Loudness of audio streaming and on-demand delivery files.

This group consists of volunteer members with expertise and/or interest in the creation, distribution and emission of professional audio. AES Standards membership is open to all stakeholders with a material interest in its work, regardless of AES membership status.

The members of the writing group that developed this document in draft are:

This draft standard was developed by writing group within task group SC-02-12-Q led by Bob Katz and David Bialik with the following members: David Bialik, Rob Byers, Jim Coursey, Eelco Grimm, Bob Katz, John Kean, Scott Norcross, Robert Orban, Shawn Singh, Jim Starzynski, Alessandro Travaglini,

Additional contributions were provided by Ian Shepherd and Greg Ogonowski.

Morten Lave
Chair, Working Group SC-02-12
2023/04/28

Note on normative language

In AES standards documents, sentences containing the word “shall” are requirements for compliance with the document. Sentences containing the verb “should” are strong suggestions (recommendations). Sentences giving permission use the verb “may”. Sentences expressing a possibility use the verb “can”.

AES Recommended Practice

Loudness Guidelines for Internet Audio Streaming and On-Demand Distribution

0 Introduction

0.1 General

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 Television, or On-Demand Video). Guidelines for that material are covered in other industry recommendations and standards (e.g., AES71-2018) [1].

These recommendations provide a necessary step in an evolutionary process (see Section 4, Introduction) 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 Annex A have significantly improved the listener experience; forthcoming metadata encoded streams and compliant de-vices 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.

0.2 Patents

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. AES shall not be held responsible for identifying any or all such patent rights.

0.3 Documentation conventions

Capitalization denotes the term is defined in Section 3, Terms, definitions and abbreviations.

1 Scope

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 [2] 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

Note that this document does not recommend Loudness Range (LRA), device target playback loudness or device dynamic range.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ITU-R BS.1770, *Algorithms to measure audio program loudness and true-peak audio level*, International Telecommunications Union, Geneva, Switzerland, <https://www.itu.int/rec/R-REC-BS.1770/en>

ITU-R BS.1771-1, *Requirements for loudness and true-peak indicating meters*, January 2012, International Telecommunications Union, Geneva, Switzerland, <https://www.itu.int/rec/R-REC-BS.1771-1-201201-I/en>

3 Terms, definitions and abbreviations

For the purposes of this document, the following terms, definitions, and abbreviations apply.

3.1

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.

3.2

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 Dialog Integrated Loudness.

3.3

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.

3.4

dBFS

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 as defined in AES17 [3] . Sometimes written as dB FS.

3.5

dBTP

decibels True Peak

The true peak amplitude, relative to 100% full-scale, expressed in dB (per ITU-R BS.1770 Annex 2) Sometimes written as dB TP.

3.6

Dialog Integrated Loudness

The 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. Also called Dialog Loudness or Speech Loudness.

3.7

Distribution

As used in this document, a Distributor's act of sending content to end users.

3.8

Distribution Loudness

The intended and/or actual Integrated Loudness or Dialog Integrated Loudness of a given piece of distributed content or an entire stream, depending on the normalization method employed. 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 reduce confusion.

3.9

Distributor

As used in this document, the entity in the delivery chain tasked with normalization of content loudness before streaming or download by end users. This entity performs Distribution, and is typically a branded streaming service, on-demand content provider, or “online radio station.”

3.10

Downward Normalization

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

3.11

Dynamic Range

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

3.12

Dynamic Range Control (DRC)

The process of continually adjusting audio signal level to control the loudness difference between loud and soft passages.. 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.

3.13

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 payout.

3.14

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 assembled from elements that have different genres.

3.15

Genre

A name, such as “speech,” “popular music,” “classical music,” etc., characterizing Long-form Content that has a homogenous texture and style.

3.16

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.

3.17

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.

3.18

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.