Loudness Normalization:
The Future of File-Based Playback

Music Loudness Alliance[1]
http://music-loudness.com
Version 1.0, July 2012

Today, playback on portable devices dominates how music is enjoyed.[2] [3] A large portion of songs present on the average music player come from online music services, obtained either by per-song purchase or by direct streaming. The listener often enjoys this music in shuffle play mode, or via playlists.

Playing music this way poses some technical challenges. **First**, the sometimes tremendous differences in loudness between selections requires listeners to adjust volume. **Second**, the reduction in sound quality of music production over the years due to a loudness war. The art of dynamic contrast has almost been lost because of the weaknesses of current digital systems. **Third**, the potential damage to the ear caused by these loudness differences and a tendency towards higher playback levels in portable listening especially when using earbuds.

The Three Challenges

Loudness Differences

In digital audio, the maximum (peak) audio modulation has a hard ceiling that cannot be crossed. Digital audio tracks are routinely peak normalized. This results in tremendous loudness differences from track to track because the **peak level** of a signal is not representative of its subjective loudness. Rather, the listener perceives loudness according to the **average energy** of the signal. Because of the widespread practice of peak normalization, program producers apply severe compression, limiting and clipping techniques to the audio. This removes the original peaks and allows normalization to amplify the signal increasing its average energy. This has resulted in a loudness war with large loudness differences between newer and older material

---

[1] The Music Loudness Alliance is a group of technical and production experts led by Florian Camerer and consists of audio engineers Eelco Grimm, Kevin Gross, Bob Katz, Bob Ludwig and Thomas Lund. We can be contacted via Florian's email: florian.camerer@orf.at.

[2] Steve Guttenberg (September 2009). *Poll: Where do you listen to music?* CNet. http://www.webcitation.org/68yrk7e5h

[3] Reineke Reitsma (July 2010). *The Data Digest: Which Music Devices Do People Use?* Forrester. http://www.webcitation.org/68yrr5gO2

and different genres. When older recordings are included in a playlist with new material, the listener experiences noticeable jumps in loudness from track to track requiring frequent adjustments in playback level. The differences can be as large as 20 dB. The same problem occurs when different musical genres share a playlist. Portable device listening is therefore not the comfortable experience it could be, and computer playback exhibits some of the same problems.

Restoration of Sound Quality to our Recorded Legacy

In the practice commonly referred to as the "loudness war", many artists, recording engineers and record labels strive to make their recordings sound louder so they will stand out compared to others. The aggressive dynamic range compression used to produce loud recordings reduces peak-to-average-energy ratio. The effect has been that the important artistic and narrative tool of dynamic contrast has almost totally disappeared in modern music production.

The result of this pressure to be louder is that the steps of the production process, recording, mixing and mastering, produce masters that incorporate several generations of digital processing which can cumulate clipping and alias products. This distortion is exacerbated when the product is finally encoded to a lossy medium like AAC.[4] Cumulative distortion also leads to further significant distortion being added during distribution or playback.[5] This is fatiguing to the ear, which turns off some listeners and may even be the cause of reduced sales of contemporary music.[6] [7] This reduction in signal quality and dynamic range amounts to a removal of the very parts of the sound which make programs sound interesting.

By switching from peak normalization to loudness normalization as a default in playback media, producers who wish to mix and master programs with wide dynamic range and without distortion can do so without fear that their programs will not be heard as loudly as the 'competition'. Loudness normalization also permits older, more dynamic material to live alongside the newer, which will allow listeners to appreciate the sound qualities of more dynamic recordings and permit them to mix genres and recording styles.

Hearing Damage

High playback levels, whether achieved by accident, chosen per personal preference or to overcome ambient noise, are a potential source of hearing damage.[8] This is especially true for headphones and earbuds which, due to their close proximity to the eardrums, require relatively

---

[4] http://www.apple.com/itunes/mastered-for-itunes/

[5] Nielsen & Lund (May 2003), *Overload in Signal Conversion*. Audio Engineering Society.

[6] Greg Milner (2010). *Perfecting Sound Forever: An Aural History of Recorded Music*. Faber & Faber. ISBN 978-0865479388.

[7] Earl Vickers (May 2011). *The Loudness War: Do Louder, Hypercompressed Recordings Sell Better?* Audio Engineering Society. http://www.aes.org/e-lib/browse.cfm?elib=15934

[8] Scientific Committee on Emerging and Newly Identified Health Risks (2008), *Potential health risks of exposure to noise from personal music players and mobile phones including a music playing function*, European Commission. http://www.webcitation.org/68yrYudQh

little power to reach damaging levels. In the past some European countries have attempted to address hearing damage by legislating maximum peak output level for portable players. The net result is that it is difficult to enjoy old recordings or dynamic genres like classical music at suffi-cient loudness on these output-limited devices. Unfortunately this has increased pressure on mastering engineers to remove dynamic peaks from tracks in order to provide loud enough playback levels for the restricted peak level. Again, peak output level is not directly connected to perceived loudness. It is also not used as a predictor of hearing damage potential in internation-al law. Instead the integrated level over a certain period of time should be used.

An Integrated Solution

ITU Loudness Normalization

There is a solution for problems of inconsistent playback, the loudness war, hearing damage and the sound quality issues. This solution is founded in the massive adoption of file-based mu-sic consumption in all kinds of formats. All playback devices and music servers are effectively computers that may analyze the average perceptual energy of a file and adjust its playback lev-el accordingly. For international broadcasting, the ITU-R[9] BS.1770-2 standard for loudness measurement has recently been developed.[10] It defines the equivalent loudness of an audio signal as its **LUFS** level, meaning *Loudness Units relative to Full Scale*.[11] BS.1770-2 does a very good job in predicting subjective loudness.[12] Loudness normalization based upon BS.1770-2 is being rolled out worldwide for television broadcast.[13] Apple has successfully implemented loudness normalization in its **Sound Check** algorithm for iTunes and supported portable play-ers. A similar open system known as **ReplayGain**[14] is available for other players. The adoption of BS.1770-2 by these systems would be advantageous in the sense that music normalization would then be based on one international standard for loudness measurement.

ON by Default

Listener experience will generally improve when a loudness normalization algorithm is turned **ON by default**. This will also facilitate compliance with regulations to prevent hearing loss. **Loudness normalization ON by default** would also help to put an end to the "loudness war" in music production. In order for playback devices not to drop in loudness level compared to what listeners had been familiar with, we suggest a different form of system level control that we call **Normalized Level Control**.

---

[9] International Telecommunication Union, a United Nations agency, Radiocommunication Division

[10] ITU (March 2011) *Recommendation ITU-R BS.1770-2, Algorithms to measure audio programme loud-ness and true-peak audio level.* http://www.webcitation.org/68yrRmvrC

[11] A step of 1 LU (Loudness Unit) is equivalent to 1 dB.

[12] Paul Nygren (2009). *Achieving equal loudness between audio files*. KTH Royal Institute of Technology. http://www.webcitation.org/68ys6rwC0

[13] EBU R 128 in Europe. CALM act and ATSC A/85 in the U.S.

[14] http://www.replaygain.org

NORM-L (Normalized Level Control)

Typical loudness normalization solutions normalize playback material to a **fixed target** loudness level. A separate user-adjusted volume control sets playback level following normalization. This is a compromise: if the target level is too low, the maximum acoustical level will not be sufficient in battery-operated devices; if it is too high, normalization will be compromised or distortion introduced. **"NORM-L"** (Normalized Level control) is a method of resolving the shortcomings associated with a traditional fixed-target solution. The idea behind NORM-L is that upon playback **the listener's volume control sets the loudness target level** to which the files will be adjusted. Loudness normalization and volume control are integrated into one gain step. If this would lead to clipping of the file, the applied gain is restricted appropriately.[15] (See Appendix 1 for a detailed description of NORM-L).

Album Normalization

One important refinement to loudness normalization is **album normalization**. Although it is common for music nowadays to be bought as separate songs, most artists still release their songs in album format. The loudness of album tracks has been carefully balanced by the mastering engineer to optimize the artistic impact of the recordings. In a classical symphony recording, for example, individual movements have a distinct dynamic relationship to each other. If all tracks were normalized to the same target loudness these important aesthetic properties would be lost. Listeners commonly construct playlists from many different albums. In these cases, the loud and soft songs should be reproduced at the producer's intended relative level; the soft songs should not be brought up to the same loudness as the loud ones. (See Appendix 2 for further details.)

We propose that **album normalization be turned on as a default,** in order to satisfy the aesthetics of the artist and album producer and the majority of playback situations.[16]

Hearing Damage Protection

In Europe new safety requirements for A&V equipment have been published that prescribe mobile music players must show a warning to users when their hearing is in danger. By integrating these demands in NORM-L, automatic compliance to European law is obtained with the best possible user experience. (See Appendix 3 for a more detailed description and suggested solutions).

---

[15] Peak limiting or clipping are not acceptable practices for loudness normalization as they distort the sound quality. Gain should only be applied in normalization if the peak level would not exceed full scale.

[16] We believe album normalization best serves the artist's intent in most playback scenarios. One possible exception is for dance parties where it may be desirable to use per-track normalization so that ballads play as loudly as the loud songs regardless of the normal aesthetics. We suggest that deactivation of album normalization be set on a playlist basis, remaining in force by default for all other playlists.

Appendix 1: NORM-L (Normalized Level Control)

NORM-L analyzes a file's average loudness and stores this alongside the audio as *FileLUFS* metadata.[17] The file's peak level is also stored, as *FilePeak* metadata. The audio content of the file is not changed. NORM-L can be described algebraically as follows:

Gain = min (FaderPosition - FileLUFS, -FilePeak)

Where:

Gain is the setting applied to playback hardware in decibels.

FaderPosition is the physical position of the listener's volume control. The range of this control is from a *MaxFaderPosition* at the physical top, down to -infinity. In other words, if *MaxFaderPosition* is -13 dB, when the user's fader is at its physical maximum, the value applied to the calculation is -13 dB (see Appendix 3 for *MaxFaderPosition* recommendations).

FileLUFS is the loudness measurement of the file in LUFS units.

*FilePeak* is the maximum peak level of the file in decibels relative to digital full scale.

NORM-L can be described graphically as follows:
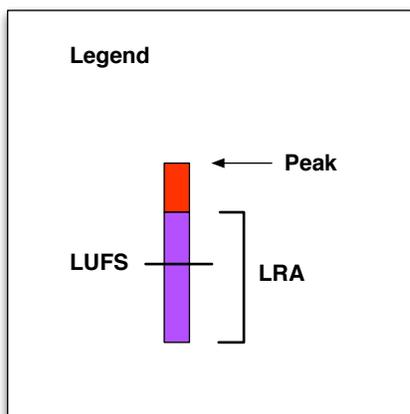
The recorded file has an average measured loudness (LUFS), indicated with a horizontal line, a maximum peak level (at the top of the red section), and a loudness range (LRA) (the purple segment) which is a measure for the macrodynamics of a recording, the difference between the average loud and soft parts.

---

[17] See Appendix 4 for a discussion of loudness analysis practices.

This figure illustrates the loudness distribution which may be found in three different genres:
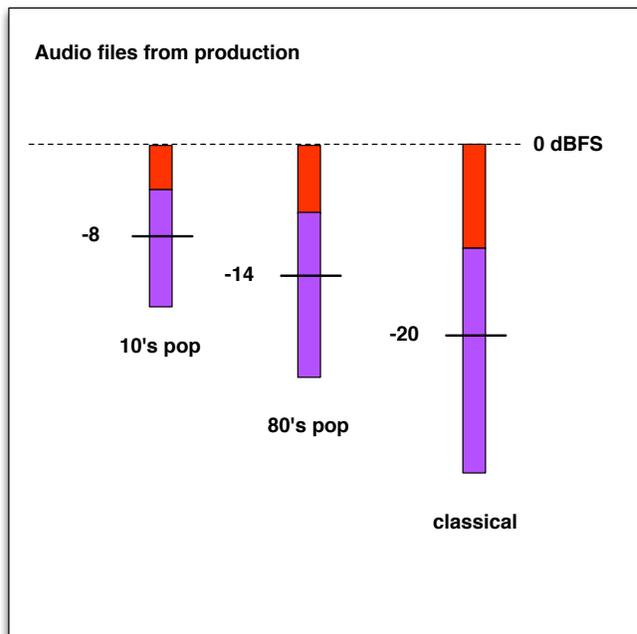

Because of the differences in the measured average loudness, it is obvious that playback of these three tracks in one sequence would lead to loudness jumps.

*Next, an example of how NORM-L solves the problem:*
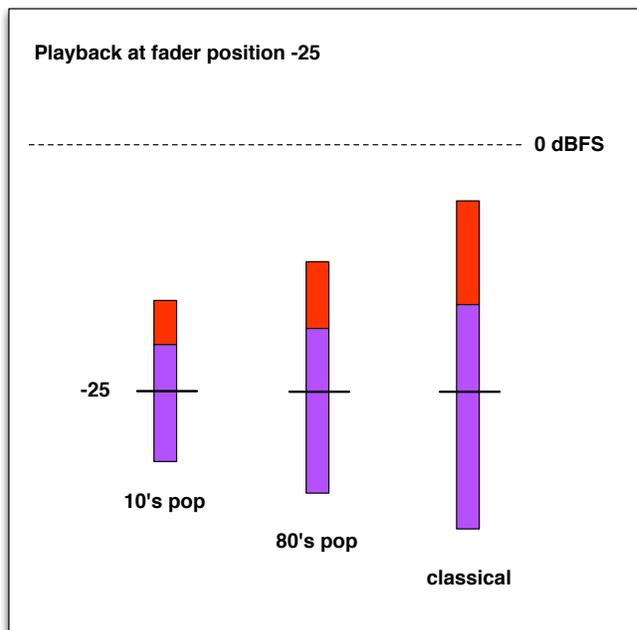
Now the three files play back at the same loudness. The first file's loudness level was *-8* LUFS and the NORM-L fader position is at *-25*, so this file will be attenuated by 17 dB at the moment of playback. Likewise, the *-20* LUFS classical file will be attenuated by 5 dB.

Now, we raise the level control to position -20:


Playback at fader position -20
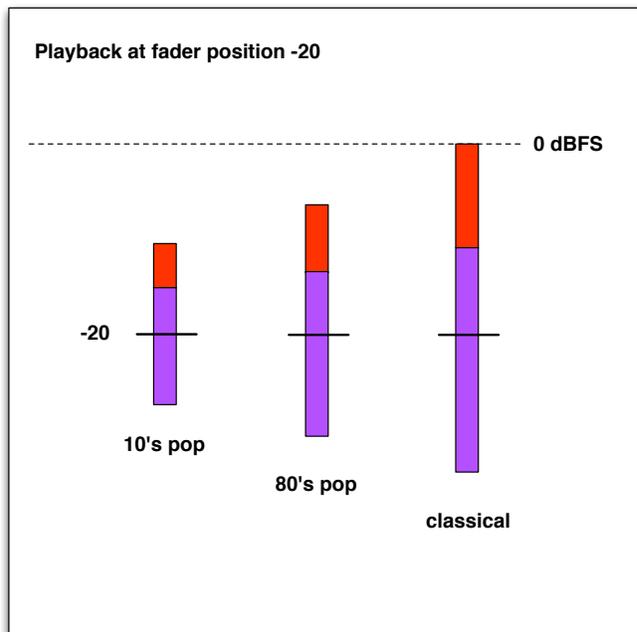
Even when turning up NORM-L by 5 dB, the classical material is peaking to the maximum level but does not clip. The other two files still have ample headroom.

However, now the fader position is set to -15:

Playback at fader position -15 without limiter
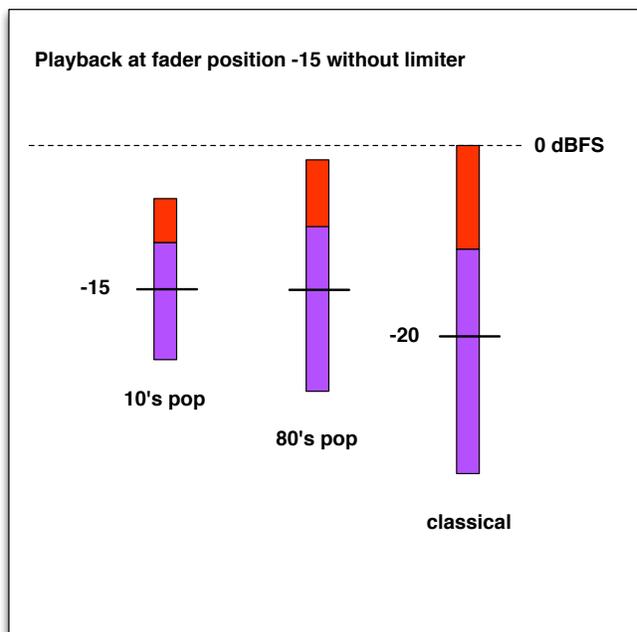
This setting would cause the classical music to clip, so NORM-L constrains its normalization to prevent this. In this case the file will be played back effectively at *-20* although the fader is set to *-15*. The classical music plays back 5 dB quieter than the other two examples, but is not clipped.

An alternative is to add a limiter to the playback device, as shown below:
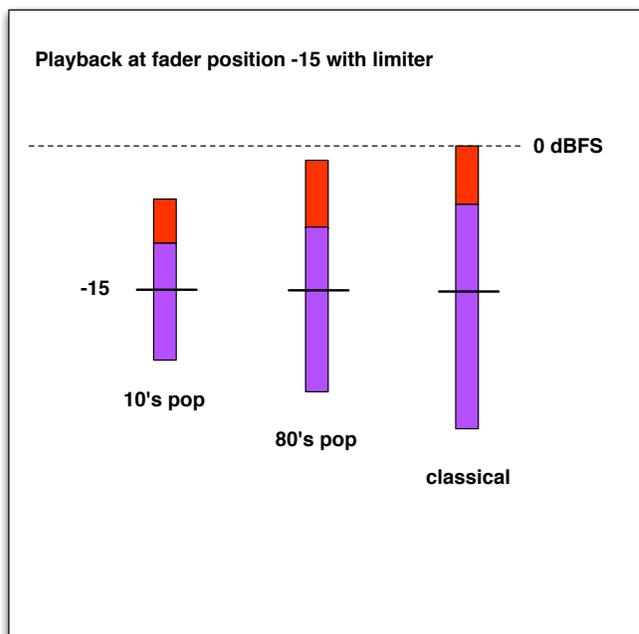
This allows the user to increase the loudness of dynamic tracks beyond the normal clipping level, but compromises sound quality as transients are removed by the limiter. The vast majority of recorded music is at an average measured loudness of -16 LUFS or higher, so only extremely dynamic material such as late-romantic symphonies, and rare pop material will encounter this clipping issue and only if the listener turns the level control up too far.

For the user, this new type of level control will behave in exactly the same manner he was used to. The only difference is that all songs will sound equal in loudness, regardless of the peak level of the recordings. ***The main advantage of NORM-L over fixed target systems, such as Sound Check and ReplayGain, is that normalization improves as the fader is lowered.***

Appendix 2: Album Normalization

All tracks from one album should receive the metadata value of the loudest track of the entire album, **AlbumLUFS**. When available, this *AlbumLUFS* should be used instead of *FileLUFS* metadata. When a quieter track from an album is played in sequence with other tracks, it will then still receive the intended lower loudness level. To determine the maximum gain, the **File-Peak** level is still used. Algebraically, the NORM-L formula becomes:

Gain = min (FaderPosition - AlbumLUFS, -FilePeak)

Appendix 3: MaxFaderPosition, Hearing Damage

In the context of our NORM-L proposal, we advise to limit the volume control of mobile music players to a certain "MaxFaderPosition". The same parameter can be used to limit the maximum acoustic level of a player and headphone combination as demanded by new safety standards in Europe. We differentiate between four situations.

a) Portable devices and other devices with sufficient headphone output level, not sold in the Euro zone.

For devices with sufficient output level,[18] we recommend a *MaxFaderPosition* of -13.[19] Well-designed players have more than sufficient analog output to allow a -13 *MaxFaderPosition*. A -13 max value provides effective normalization for the vast majority of music which is encountered today. Furthermore, most listeners will experience a minimal change or no level drop when normalization is introduced.[20] This will help ensure easy adoption and success of normalization. A higher *MaxFaderPosition* would offer an even smaller level drop, but this potentially leads to a large dead zone at the top of the volume control for files with low loudness. It would also lead to poor normalization when the user sets the player's volume control to maximum and the headphone output is connected to a line input feeding an external amplified speaker or a car system.

b) Lower cost mp3 players and other devices with lower output level and headroom, not sold in the Euro zone.

In this case we recommend the lowest possible value that still produces sufficient acoustical output through the included earbuds. Values higher than -13 provide inadequate normalization at higher fader settings. As an alternative, manufacturers should consider improving the headphone output capability of their players so as to provide adequate level and peak headroom.

c) Line or digital outputs and wireless connections on mobile players, media systems and personal computers.

When placing a mobile device in a docking station, the audio will often be played via a separate digital or analog line output. This output is connected to an amplifier which has its own volume control that functions as the main volume control for the sound system.[21] NORM-L has no ad-

---

[18] Devices with sufficient level are able to produce 100 dB SPL(A) or higher playing CENELEC's EN 50332 reference signal at their highest volume setting.

[19] The FaderPosition value of -13 is applied when the NORM-L fader is at the top of its range. Devices which display fader position in decibels may choose to label this point 0 dB to avoid user confusion.

[20] The -13 MaxFaderPosition recommendation produces a maximum output 3.5 dB louder than *Sound Check* and 1 dB louder than *ReplayGain*.

[21] When using wireless playback facilities such as Airplay, the main volume control is in the receiving device as well.

vantage here and we advise to use a fixed target level of, preferably, -23 LUFS (based on the EBU Tech Doc 3344[22]). Although this may seem like a low value, the connected amplifiers normally have more than sufficient gain to compensate and the advantage is that even most classical music will be properly loudness normalized without clipping. Another advantage is that when switching to modern AV systems that operate at the same target level, the user will not experience any loudness jump.

d) Hearing Loss Protection in the Euro zone.

For hearing loss prevention, international laws prescribe use of A-weighted intensity measurement and equivalent exposure over time (the **dose**). In Europe a CENELEC[23] working group, in consultation with the European Committee, has published a standard for portable music playback devices with their included earbuds.[24] The standard requires a safety warning message be given should intensity exceed 85 dB SPL A-weighted (dBA). The listener is required to actively confirm the message before he is allowed to play at higher levels and under no circumstance is playback above 100 dBA permitted.[25] Conventionally the 85 dBA limit is enforced by measuring in real-time the average energy over a 30 second window. As a result, loud passages in dynamic recordings (classical music, for instance) may unnecessarily trigger the warning. The CENELEC group was aware of this and allows that "if data is available of the average level of the whole song, the message may also be given in case the integrated average level exceeds 85 dBA."

While loudness normalization is being performed, hearing loss prevention can be accomplished by calculating a per-track *WarningFaderPosition* and *MaxFaderPosition* that take into account the file's A-weighted level. By measuring and storing *FileDBA* (the integrated average A-weighted level), in addition to the *FileLUFS* loudness level,[26] the same metadata mechanism used for loudness normalization can also be used to produce hearing damage warnings and operating restrictions that conform to EU law.[27] Using *FileDBA* instead of a conventional real-time measurement has several benefits: The user can be warned of excessive loudness at the

---

[22] EBU Tech Doc 3344 "Practical Guidelines for Distribution System in accordance with EBU R 128" http://www.webcitation.org/68yu4FvT8

[23] CENELEC is the European Committee for Electrotechnical Standardization and is responsible for standardization in the electrotechnical engineering field. CENELEC prepares voluntary standards, which help facilitate trade between countries, create new markets, cut compliance costs and support the development of a Single European Market.

[24] 2011 Amendment 12 to EN 60065 (safety requirements for A&V equipment) and EN 60950 (safety requirements for ICT equipment - incl mobile phones)

[25] Measured according to EN 50332 (Sound system equipment: Headphones and earphones associated with portable audio equipment. Maximum sound pressure level measurement methodology and limit considerations.)

[26] Using dBA for loudness normalization as well (instead of LUFS) is suboptimal: Gilbert Soulodre (2004-05-08), Evaluation of Objective Loudness Meters, AES 116th convention, preprint 6161.

[27] CENELEC requires the listener acknowledge a warning every 20 hours of play when playback exceeds 85 dBA and limits maximum output to 100 dBA.

beginning of a track instead of being interrupted in the middle. Also, the hearing damage potential of dynamic content like classical music is judged on a more long-term basis in line with hearing loss protection standards and law.

The per-file fader position at which the device must show a warning and the level at which the device limits its output can be described algebraically as follows:

> *WarningFaderPosition* = 85 - IECLevel + RefnoiseLUFS - FileDBA + RefnoiseDBA
> *MaxFaderPosition* = 100 - IECLevel + RefnoiseLUFS - FileDBA + RefnoiseDBA

Where:

> *WarningFaderPosition* is the fader position above which the player must display a warning in conformance with EN 60065.
> *MaxFaderPosition* is the physical maximum of the device's fader used for the duration of the file.
> *IECLevel* is the EN 50322 measured acoustical level of a portable device at its maximum gain (NORM-L in bypass) with its standard headphones in dB(A) SPL.
> *RefnoiseLUFS* is the measured loudness of EN 50322 single channel reference noise. A value of -13 LUFS should be used here.[28]
> *FileDBA* is the A-weighted level of the file's loudest channel.
> *RefnoiseDBA* is the A-weighted level of the EN 50322 reference noise. A value of -12.6 dBA should be used here.

For example, suppose a device can produce a maximum acoustic output of 104 dBA from factory earbuds when playing the reference noise. While playing a file whose loudest channel measures -14.6 dBA, *MaxFaderPosition* = 100 - 104 - 13 + 14.6 - 12.6 = -15 would need to be used to prevent output from exceeding the 100 dBA hearing loss protection limit.

Portable players may feature an equalizer function. If present, EN 50332 requires that this equalizer be adjusted in order to maximize the sound pressure level and that this setting be used to establish the 100 dBA limit. Because the impact of an equalizer on sound pressure level is content dependent, when the equalizer is engaged it's no longer possible to accurately determine the 85 dBA and 100 dBA thresholds based on *FileDBA*. To meet EN 50332 requirements, a system must account for the effect of EQ settings. This can be done by conservatively biasing *WarningFaderPosition* and *MaxFaderPosition* to ensure that in the presence of EQ, the thresholds are never exceeded. Alternatively manufacturers may choose to design the equalizer such that it will never boost the sound level at any frequency; to affect a boost everything else is cut. An additional advantage of the latter method of EQ is that the system cannot overload before the volume control.

By following this rule, the portable audio device automatically complies to the maximum acoustic level of 100 dBA as specified in EN 60065 and any additional available headroom is used to im-

---

[28] Since a mono EN 50322 reference noise is used, its mono level is -13 LUFS, the same if fed to one channel of a stereo system.

prove normalization effectiveness. Note that NORM-L in this case should not be defeatable by the user or the device would become illegal. In EU countries where device output so far has been limited to meet the law, old recordings and uncompressed genres like classical music can once again be played with adequate loudness.

Appendix 4: Loudness Analysis

When or where should the loudness analysis of the file take place? Ultimately this is a decision made by player manufacturers. There are at least four options:
1. by the record label or mastering house
2. at the point of sale (iTunes Store or other web store)
3. in the media server (iTunes in the context of Apple products for instance)
4. in the portable player itself.

Metadata from an unknown source cannot be trusted. So unless the source is secure (as with iTunes), we advise to let the portable player perform the analysis itself as it only has to be performed once. Battery power consumption may be a reason to perform loudness normalization of content outside the player. Again, this is a decision ultimately made by manufacturers.