



Recommendation for loudness normalization by Music Streaming Services

*Eelco Grimm, HKU University of the Arts Utrecht, The Netherlands,
eelco.grimm@hku.nl*

From the jukebox days in the 50's artists and record labels have pushed mastering engineers to make their songs louder than any other's. With the advent of fast digital peak limiters in the mid 90's, the pressure on mastering engineers has led to a dramatic loss of dynamics in pop music. This is known as the "Loudness War"¹. Even artists who do not feel an urge to compete in loudness are pulled into this race. For more than a decade sound engineers, musicians and music lovers have discussed how to end it. There was no solution possible when CD's and downloads were the main source of music consumption, because there is no central control on their level. But now Music Streaming has taken over² as the major source of music consumption, the loudness war can be ended if these services turn metadata driven loudness normalization on by default.

Loudness is subjective, but can be measured with the right tools. In 2006 the ITU has introduced a standard for measuring loudness in broadcasts, BS1770³. Since it is the only open standard for loudness

¹ https://en.wikipedia.org/wiki/Loudness_war

² <http://www.ifpi.org/facts-and-stats.php>

³ <https://www.itu.int/rec/R-REC-BS.1770/en>

measurements available and music producers demand that their songs be treated equally on all streaming services, we recommend to use BS1770 to measure loudness of music productions.

The main issue is to select the type of loudness normalization: track normalization or album normalization. With track normalization all tracks are made equally loud. With album normalization, just the loudest tracks of an album are made equally loud and the other tracks keep the relative level they had on their album. If one listens to an album, album normalization makes most sense. But when streaming, people do not just listen to albums as a whole but also to randomly picked tracks in a shuffled playlists. So the question is: does album normalization work for a shuffled playlist too?

The second question is: what should the target level be for mobile playback? The current generation of mobile devices such as smart phones have a limited headroom to comply with governmental hearing protection laws⁴ in the EU. Because of this, AES TD1004⁵ recommends to keep the loudness of programs above -20 LUFS in mobile devices.

In cooperation with TIDAL we did a large survey on 4.2 million albums of their catalog. A proposal was developed to use album normalization where the loudest track of each album is normalized to -14 LUFS and the other tracks are aligned to the relative level they have in the album. These levels will then also be used when tracks are played in a randomly shuffled playlist with other albums' tracks. This proposal was tested against track normalization in a shuffled playlist of 24 songs with 38 subjects. It turned out that 80% of the subjects preferred album normalization, even though the tracks were selected for an extreme difference in loudness, of up to 10 LU. Based upon this we could write the following recommendations.

⁴ Cenelec 2011 Amendment 12 to EN 60065 (safety requirements for A&V equipment) and EN 60950 (safety requirements for ICT equipment - incl mobile phones)

⁵ http://www.aes.org/technical/documents/AESTD1004_1_15_10.pdf

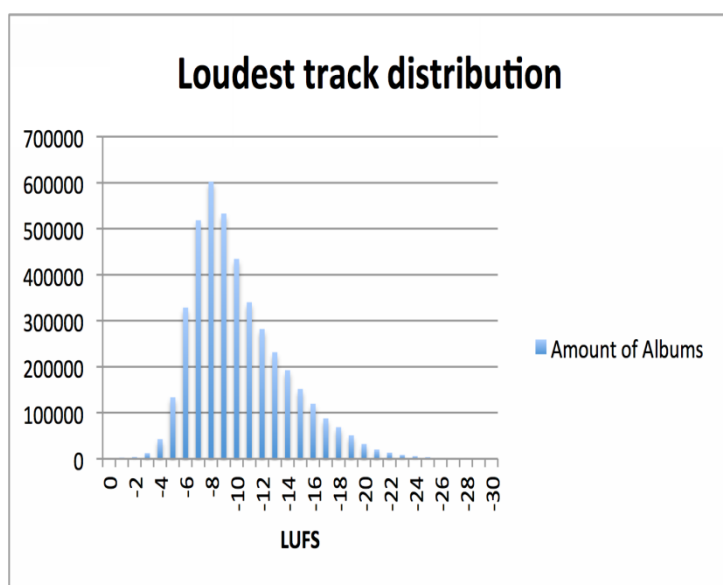
Recommendations to Music Streaming Services

1. Turn album normalization ON by default. With album normalization, the loudest tracks of all albums are adjusted to the same loudness during playback.
2. Use the industry standard ITU BS1770-4 measurement so that mastering engineers can predict the result of their work for all services at once.
3. In mobile devices, use a target level of -14 LUFS for the loudest track of an album.
4. In a living room setting, lower target levels such as -18 LUFS to -20 LUFS are preferred. The advantage is that albums of more dynamic genres such as classical and jazz will also become loudness normalized.
5. To avoid clipping, only attenuate tracks, never apply positive gain. If the loudest track of an album is softer than target level, all tracks of the album will play soft.
6. In the preferences, loudness normalization should optionally be turned off, for people who are concerned about automatic changes to the data, or for research purpose.
7. Although it seems attractive to add the third option of “track normalization” in the preference to satisfy all users, this can potentially confuse many people. It is not recommended.
8. Provide playback device manufacturers with the option to merge loudness normalization with playback gain, since this offers the highest possible loudness normalization and sound quality. In this case the loudness meta data information of a track is sent downstream all the way to the final gain stage of the system. The playback device manufacturer then merges the normalization data with the user set playback gain and applies just a single gain change. (Mark that if the system is floating point throughout, both gains can also be set independently). The advantage is that over a large range of playback levels, both very dynamic and compressed albums will be perfectly aligned in loudness, and there is no need to select a ‘target level’ anymore.

The research in more detail

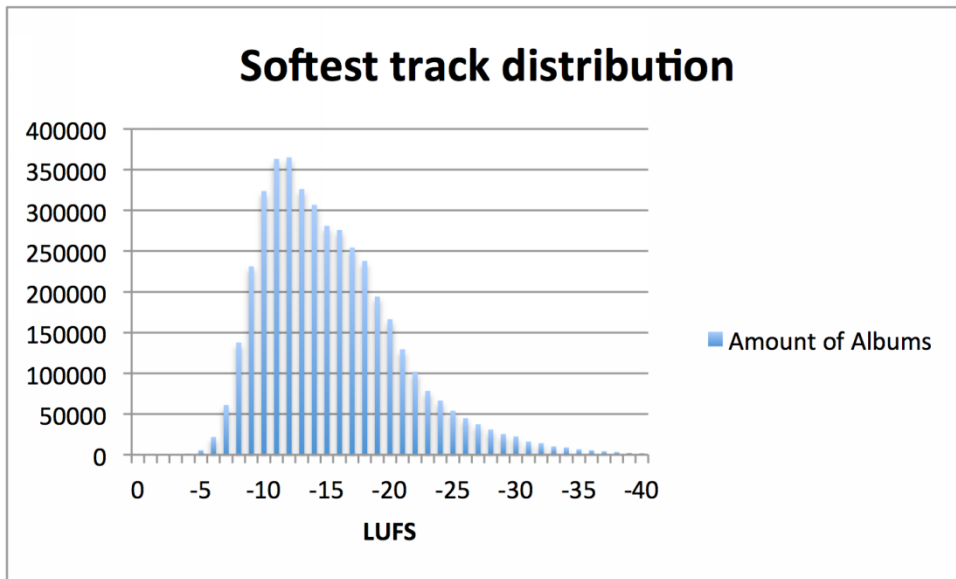
Analysis of the TIDAL database of 4.2 million albums showed that until the end of the 90's albums had a peak in the distribution at around -14 LUFS, while in the 00's and 10's that moved up to -8 LUFS. The loudness war is real.

Here is a graph that shows the distribution of the loudest track of all albums of the database:



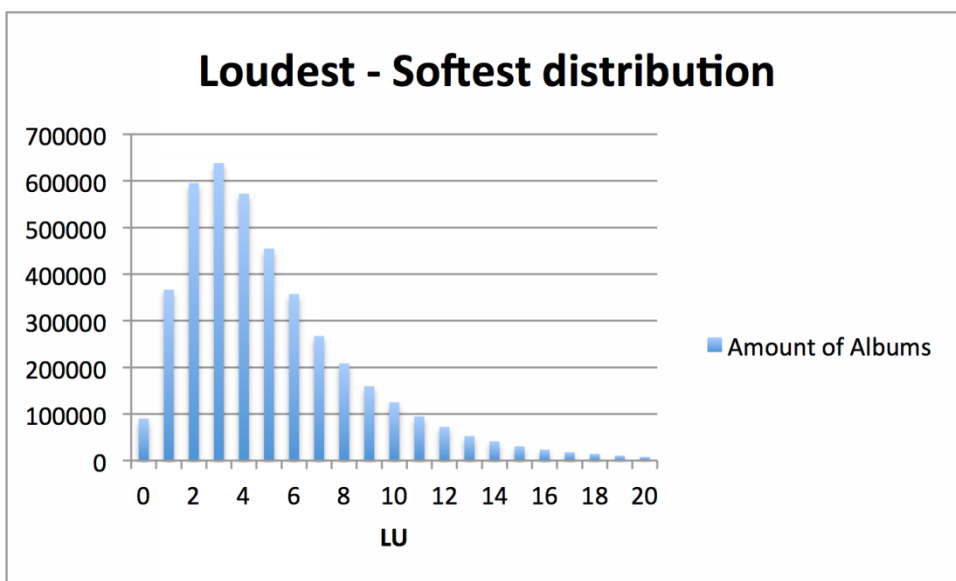
87% of all albums have a loudest track that is louder than -14 LUFS. Part of the remaining 13% was manually evaluated and most of these albums were either classical or jazz, or were made in the 80's. Many of these are -15 or -16 LUFS, so would play just 1 or 2 dB low if a target of -14 LUFS was chosen.

Here is a graph that shows the distribution of the softest track of all albums:



The peak in the distribution is broader than with the loudest tracks, and just 50% of all albums have their softest track louder than -14 LUFS. This means that if -14 LUFS would be chosen as a target level for *track* normalization, the softest tracks of 50% of all albums can not be normalized properly because they would clip.

Here is a graph that shows the distribution of the difference between the loudest and softest track of all albums:



In just 2% of all albums the softest track is of equal loudness as the loudest. That means that track normalization will change the artistic

loudness characteristics of 98% of all albums ever made. In 72% of the albums the softest track is 6 LU or less below the level of the loudest track. If album normalization is used and the loudest track is aligned to -14 LUFS, the softest tracks of these albums would still be in the range of the AES TD1004 recommendation of -16 to -20 LUFS.

A group of albums from the remaining 28% (with a difference of 7 LU or larger) was manually evaluated and it was found most of the albums were either classical or jazz. Among the pop albums in this group there were some quite recent ones. Apparently artists still love to add soft tracks to their albums sometimes, as an intermezzo. The question was if the low loudness of such tracks would still be appreciated when they are played outside the album context, in a shuffled playlist. We took the loudest and softest tracks from 12 albums with substantial differences of 7 to 10 LU. These were put in a randomly shuffled playlist of both track normalized and album normalized style, which were tested with 38 subjects.

The result was that 71% of the subjects preferred album normalization blindly. Another 9% said they would never accept track normalization if it was turned on by default and in that case they'd rather accept the level differences in the shuffled playlist. Which means 80% have a preference for album normalization in case normalization should be turned on by default. That percentage would likely be larger for albums with more typical loudness differences between tracks. The testing done here was on material with very large differences.

Regarding the optimal target level, normalizing the loudest track to -14 LUFS seems quite a good choice. If it is lowered to for instance -16 LUFS, a larger amount of 'loudest tracks' will be correctly normalized, but at the same time many more soft tracks will fall below the -20 LUFS lower limit that is recommended by AES TD1004. Conversely, if the target is raised to let's say -12 LUFS, a lot of albums made before the 2000's will lose normalization. At the moment Spotify and TIDAL have a track normalization target at ca. -14 LUFS. If they switch to album normalization at -14 LUFS, most

contemporary hit songs will remain at the same level since they are usually among the louder tracks of the albums.

In a stationary situation, such as in a living room, it makes sense to lower the target level since there are no headroom limitations in that case. Broadcasts are normalized to -23 LUFS (in Europe) or -24 LUFS (US), and by lowering the target of the loudest track to ca. -20 LUFS, a switch between music and broadcasts would show few loudness jumps. Additionally, music genres with large dynamics, such as classical music, will then be loudness normalized with pop music. Care should be taken to loudness normalize system sounds as well. It is recommended that mobile devices automatically switch to a lower loudness target if they stream to a stationary device, for instance via Airplay or Google Chromecast.

Additional notes about album normalization

1. A special property of album normalization is that an artist can still decide to choose track normalization for his or her album. For instance by releasing albums with just one track. Or by mastering an album in such a way that all tracks are equally loud. That means that album normalization offers the creative option for the artist to select track normalization. On the other hand, if track normalization were to be the standard there is no option to release an album with album normalization, except by putting all songs in one single track, which is highly inconvenient.

2. Album normalization comes in two types: normalize according to the average loudness of the whole album or according to the loudest track of the album. The ‘average album loudness’ variant is currently implemented in Apple iTunes and in ReplayGain. It unfortunately has a few drawbacks. Someone could master an album with many soft songs and just one loud song, which will then play louder than the loud songs from other albums. The loudness war would therefore continue. Another disadvantage is that the playback loudness of a certain track would only be known when the full album is finished

and measured as a whole. This is unacceptable for a mastering engineer. All in all, we recommend to never use the ‘average album loudness’ version but always the ‘loudest track’ version.

3. Some people have suggested to switch automatically between track normalization and album normalization, based upon the user’s behavior. Our research showed that 71% of the subjects had a blind preference to use album normalization in a shuffled playlist, so this automatic switch may not be the ideal solution. But apart from that it will be hard to implement such behavior by the music streaming service, because users may decide to play the full album after they have started a track. That track was then started at ‘track loudness’ level but the next track would have to be played at ‘album loudness’, which of course breaks the loudness sequence.

4. If one looks at the question “track or album normalization?” in a wider perspective it is clear that track normalization has an artistic problem because it deprives the artists from part of their artistic freedom in the creation of a music album. Using track normalization on classical music is obviously unacceptable to all users. It would be a bit cynical to have loudness normalization limit the artistic freedom of musical artists, since the aim of normalization is to stop the loudness war that currently limits that artistic freedom. For these reasons alone, any loudness normalization that is turned on by default would have to be album normalization.

Thanks to TIDAL, to Maurits Lamers of HKU and to the MLA⁶ group.

V1.2 June 6th 2017

⁶ <http://music-loudness.com>