

EBU

OPERATING EUROVISION AND EURORADIO

TECH 3343

GUIDELINES FOR PRODUCTION OF PROGRAMMES IN ACCORDANCE WITH EBU R 128



VERSION 3.0

Geneva
January 2016

Conformance Notation

This document contains both **normative** text and **informative** text.

All text is normative except for that in the Introduction, any section explicitly labelled as 'Informative' or individual paragraphs which start with 'Note:'

Normative text describes indispensable or mandatory elements. It contains the conformance keywords 'shall', 'should' or 'may', defined as follows:

- 'Shall' and 'shall not': Indicate requirements to be followed strictly and from which no deviation is permitted in order to conform to the document.
- 'Should' and 'should not': Indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others.
OR indicate that a certain course of action is preferred but not necessarily required.
OR indicate that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.
- 'May' and 'need not': Indicate a course of action permissible within the limits of the document.

Default identifies mandatory (in phrases containing "shall") or recommended (in phrases containing "should") presets that can, optionally, be overwritten by user action or supplemented with other options in advanced applications. Mandatory defaults must be supported. The support of recommended defaults is preferred, but not necessarily required.

Informative text is potentially helpful to the user, but it is not indispensable and it does not affect the normative text. Informative text does not contain any conformance keywords.

A conformant implementation is one that includes all mandatory provisions ('shall') and, if implemented, all recommended provisions ('should') as described. A conformant implementation need not implement optional provisions ('may') and need not implement them as described.

Contents

1. Introduction	7
2. General Concept of Loudness Normalisation	9
2.1 Peak vs. Loudness	9
2.2 Normalisation of the Signal vs. Metadata	10
2.3 Target Level, new mixing concept	10
2.4 Loudness processors	12
3. Strategies for Loudness Levelling	13
3.1 Basic Mixing Approach	13
3.2 Loudness Metering for Production and Post-Production	14
3.3 Loudness Range	16
3.4 Climbing the True Peak	18
3.5 Advanced Live Mixing Strategies	19
3.5.1 Sports	19
3.5.2 Show	19
4. What to Measure in Production and Post-Production	20
4.1 Signal-Independent vs. Anchor-Based Normalisation	20
4.2 Low Frequency Effects (LFE) Channel	21
5. File-Based Playout and Archives	21
5.1 Loudness Levelling Strategies – Processing	22
5.2 Archival Content	23
6. Metadata	24
6.1 Programme Loudness Metadata	24
6.1.1 Deliberately Lower Programme Loudness, Loudness Offset	25
6.2 Dynamic Range Control Metadata	26
6.3 Downmix Coefficients	26
7. Surround Sound vs. Stereo – Downmix and Upmix issues	26
7.1 Downmix	26
7.2 Upmix	28
8. Alignment of Signals and Listening Level	28
8.1 Electrical Alignment Signal and Level	28
8.2 Acoustical Alignment, Listening Level	29
9. Genre Specific Issues	30
9.1 Commercials (Advertisements) and Trailers	31
9.2 Feature Films (Movies)	32
9.3 Music	35
10. Transition strategy	35
11. Appendices	37
11.1 Appendix 1: ITU-R BS.1770	37
11.1.1 Gating	38
11.2 Appendix 2: EBU R 128	39

11.2.1	Programme Loudness	40
11.2.2	Loudness Range	40
11.2.3	True Peak Level (TPL), Maximum Permitted TPL	41
11.2.4	R 128 Logo	43
11.3	Appendix 3: Loudness Metering with 'EBU Mode'	43
11.4	Appendix 4: DRC (Dynamic Range Control) Presets for Dolby Digital	44
12.	References	45

Acknowledgements

Although this document is the result of much collaborative work within the EBU's PLOUD group, it is the long-suffering chairman of this group, Florian Camerer, who has written, collated, enriched and distilled the text into its publication form over many, many weeks and months of effort.

Dedication

This document is dedicated to two great audio engineers, Gerhard Stoll and Gerhard Steinke.

Guidelines for Production of Programmes in accordance with EBU R 128

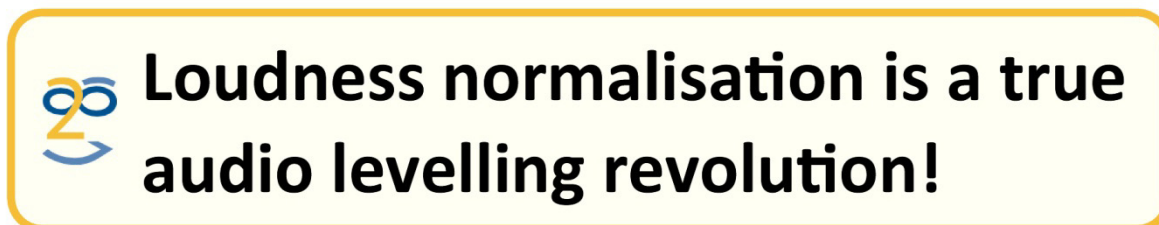
<i>EBU Committee</i>	<i>First Issued</i>	<i>Revised</i>	<i>Re-issued</i>
TC	2011	2016	

Keywords: Audio, Loudness, normalisation, production, implementation.

1. Introduction

This document describes in practical detail one of the most fundamental changes in the history of audio in broadcasting: the change of the levelling paradigm from peak normalisation to **loudness normalisation**. It cannot be emphasized enough that **loudness metering** and **loudness normalisation** signify a *true audio levelling revolution*. This change is vital because of the problem which has become a major source of irritation for television and radio audiences around the world; that of the jump in audio levels at the breaks in programmes, between programmes and between channels.

The loudness-levelling paradigm affects all stages of an audio broadcast signal, from production to distribution and transmission. Thus, the ultimate goal is to harmonise average audio loudness levels to achieve an **equal universal loudness level** for the benefit of the listener.



It must be emphasised right away that this does NOT mean that the loudness level shall be all the time constant and uniform *within* a programme, on the contrary! Loudness normalisation shall ensure that the average loudness of the whole programme is the same for all programmes; within a programme the loudness level can of course vary according to artistic and technical needs. With a new (true) peak level and the (for most cases) lower average loudness level the potential differences between the loud and soft parts of a mix (or the 'Loudness Range'; see § [3.3](#)) can actually be significantly greater than with peak normalisation and peak mixing practices in broadcasting.

The basis of the concept of loudness normalisation is a combination of EBU Technical Recommendation R 128 'Loudness normalisation and permitted maximum level of audio signals' [\[1\]](#) and Recommendation ITU-R BS.1770 'Algorithms to measure audio programme loudness and true-peak audio level' [\[2\]](#). Both documents are explained in detail in Appendices 1+2 (§ [11.1](#), [11.2](#)).

In addition to R 128, the EBU PLOUD group has published five other documents:

- EBU R 128 s1 'Loudness parameters for short-form content (adverts, promos etc.)', Supplement 1 to EBU R 128 [\[3\]](#)
- EBU Tech Doc 3341 'Loudness Metering: 'EBU Mode' metering to supplement loudness normalisation in accordance with EBU R 128' [\[4\]](#)
- EBU Tech Doc 3342 'Loudness Range: A descriptor to supplement loudness normalisation in accordance with EBU R 128' [\[5\]](#)
- EBU Tech Doc 3343 'Guidelines for Production of Programmes in accordance with EBU R 128' (this document) and
- EBU Tech Doc 3344 'Guidelines for Distribution and Reproduction in accordance with EBU R 128' [\[6\]](#)



EBU R 128 and ITU-R BS.1770 are the basis. Five more EBU Technical Documents provide details.

The Technical Documents about '*Loudness Metering*' and about the parameter '*Loudness Range*' also play an important role for the practical implementation of loudness normalisation. They will be explained in **Appendices** as well and referred to in the relevant sections (Appendices 2+3 (§ [11.3](#), [11.2.2](#))).

The '*Distribution Guidelines*' close the circle, covering all aspects of loudness normalisation for the distribution of audio signals and addressing the critical links between production and the final recipient, the consumer. As this is a very detailed document in itself it will not be covered here except for the occasional reference.

At the beginning of these '*Guidelines for Production of Programmes*' the **general concept and philosophy of loudness normalisation** will be introduced. The document will then look at **loudness strategies for production and post-production** (metering, mixing, Metadata, etc.), and for **file-based workflows**, that is, ingest, playout and archiving issues (metering, automated measurement and normalisation, Metadata etc.).

Separate chapters will look at the parameter **Loudness Range (LRA)** and **Metadata** in more detail. Electro-acoustical **alignment** of audio signals and studio **listening levels** are discussed, and **practical advice** is given for the **transition** to loudness-normalised production (implementation and migration). **Genre-specific issues** regarding commercials (advertisements) and trailers as well as movies and music programmes will be addressed in a dedicated chapter (§ [9](#)).

These Guidelines are meant to be a '**living document**', where, over time, experiences of broadcasters will find its way into the document, providing additional information and guidance for this fundamental change of the way audio signals are treated and balanced to each other.

Please note that many standards documents are subject to revision from time to time, including this one. You are strongly advised to check for the latest versions.

2. General Concept of Loudness Normalisation

2.1 Peak vs. Loudness

The audio levelling concept of *peak normalisation* with reference to a Permitted Maximum Level (PML; for example, -9 dBFS), has led to uniform peak levels of programmes, but widely varying loudness levels. The actual variation is dependent on the programme itself, as well as the degree of dynamic compression of the signal.

In contrast, *loudness normalisation* achieves equal average loudness of programmes with the peaks varying depending on the content as well as on the artistic and technical needs (see Figure 1). The listener can enjoy a uniform average loudness level across all programmes, thus not having to use the remote control for frequent volume adjustments any more.

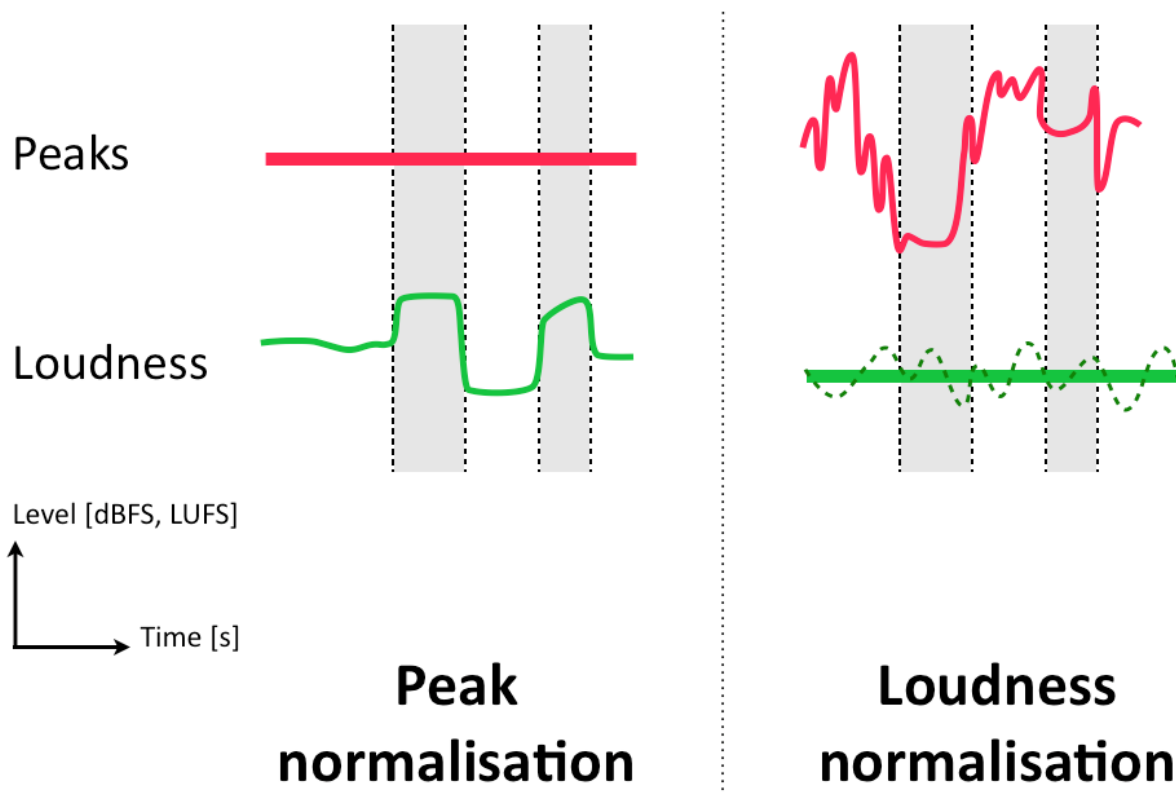


Figure 1: Peak level normalisation vs. Loudness level normalisation of a series of programmes

Again, this does NOT mean that *within* a programme the loudness level has to be constant, on the contrary! It also does NOT mean that *individual components* of a programme (for example, pre-mixes or stem-mixes, a Music & Effects version or an isolated voice-over track) have all to be at the same loudness level! Loudness variation is an artistic tool, and the concept of loudness normalisation according to R 128 actually encourages more dynamic mixing! It is the average, integrated loudness of the whole programme that is normalised.

2.2 Normalisation of the Signal vs. Metadata

There are basically *two ways* to achieve loudness normalisation for the consumer: one is the actual normalisation of the audio signal itself, so that the programmes are equally loud on average by design – the other method is with the use of Loudness Metadata (see § 6) that describe how loud a programme is. For the latter, the actual average programme loudness levels don't need to be changed to a normalised value and can still vary a lot from programme to programme. For those with up-to-date equipment, the normalisation can be performed at the consumer's end using the individual Loudness Metadata values to gain-range the programmes to the same replay level.



Equal loudness can be achieved by normalising the audio signal or by using loudness metadata.

Within the EBU R 128 loudness levelling paradigm the *first* solution – loudness normalisation of the programme itself – is recommended due to the following advantages:

- Simplicity and
- Potential quality gain of the audio signal (see § 2.3 'new mixing concept').



Loudness normalisation of the audio signal in production is recommended because of simplicity and potential quality gain.

2.3 Target Level, new mixing concept

EBU R 128 defines the new Reference Loudness Level (the so-called 'Target Level') as:

-23.0 LUFS (± 0.5 LU)



-23 LUFS is the Target Loudness Level for EBU R 128 compliance.

Having one single number has great strength in spreading the loudness-levelling concept, as it is easy to understand and act upon. And the active normalisation of the source in a way 'punishes' over compressed signals and thus automatically encourages more dynamic and creative ways to make an impact. In other words, the actual *technical* change of the audio signal level through

active normalisation to -23 LUFS has direct influence on the *artistic* process – and in a positive way! The production side is thus relieved from fighting the ‘loudness war’ – an unfortunate result of the peak-normalisation paradigm.

Working towards a common loudness level signifies a **whole new concept** of mixing, of levelling, of generally working with audio. Whereas a peak limiter set to the Permitted Maximum Level (usually -9 dBFS, measured with a QPPM (Quasi Peak Programme Meter)) provided a sort of ‘*safety ceiling*’ where, no matter how hard you hit it, it always ensured the ‘correct’ maximum level, the loudness levelling paradigm more resembles ‘*floating in space, with the open sky above*’ (see Figure 2).

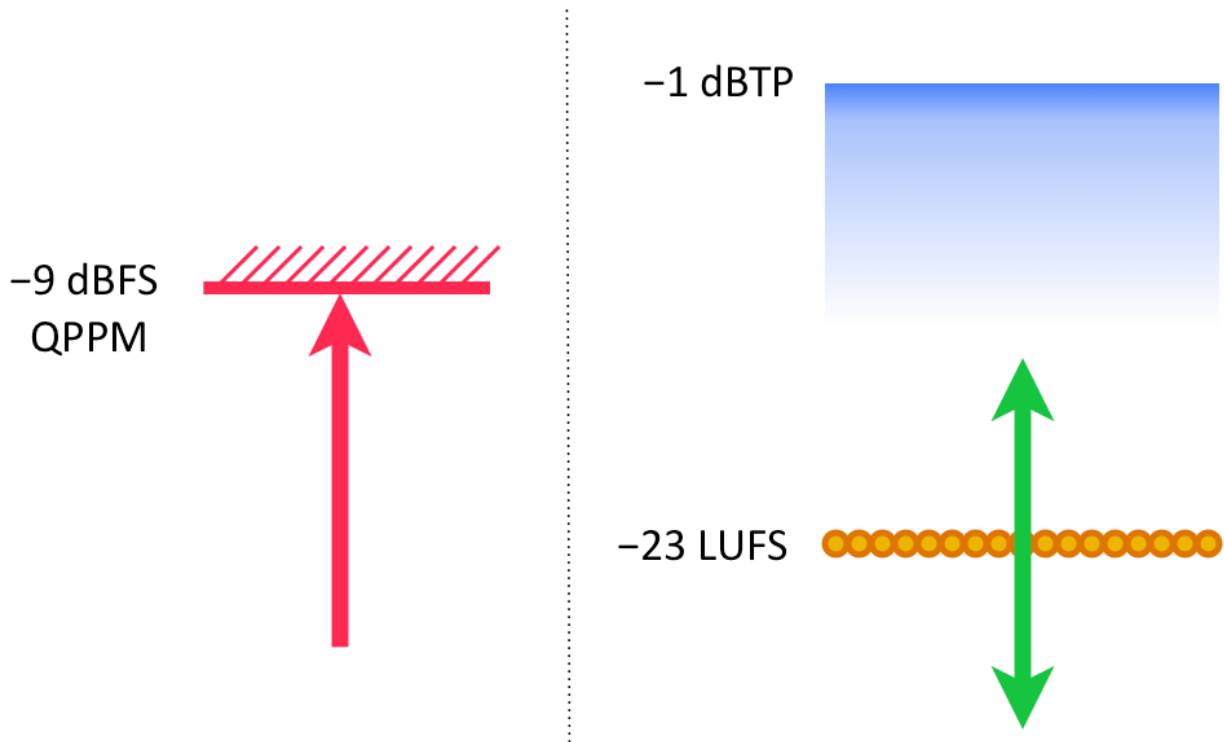


Figure 2: Quasi-Peak Level normalisation (‘safety ceiling’) vs. Loudness Level normalisation

With loudness normalisation and metering, the low safety ceiling is gone. This might be intimidating for some, as it was in a way ‘comfortable’ that one didn’t have to listen so attentively – the limiter at the end of the chain ensured that your output was always tamed. But the side effect was that loudness levels went up, the peak normalisation paradigm got abused and started a loudness competition, fuelled by ever more sophisticated dynamics processors.

Loudness levelling, on the other hand, encourages the use of by far the best metering device: **the ear**. This implies more alert mixing and fosters audio quality. Experience of several EBU members has shown that working with the loudness paradigm is **liberating** and satisfactory. The fight for ‘Who is the loudest?’ is gone, overall levels go down, and this in combination with a higher Maximum Permitted True Peak Level (-1 dBTP for linear audio; see also § [11.2.3](#) in Appendix 2 for more details) results in potentially **more dynamic mixes** with **greater loudness consistency** within the programme. Dynamic compression is again an artistic tool and not a loudness weapon – the **audio quality increases!**

Putting ‘**mixing by ear**’ back on track is a welcome relief and long overdue. The mixer is now encouraged to mix by ear alone (another effect of loudness metering) – after setting basic loudness levels for ‘**anchor signals**’ (audio signals in the foreground such as a narrator, the opening music etc.), using a fixed monitor gain (see § [8.2](#)).



Loudness levelling encourages mixing 'only by ear' – after setting basic levels, using a fixed monitor gain.

2.4 Loudness processors

Downstream of production, the broadcaster is confronted with the need to normalise diverse content originating from different places. Especially during the transition period there will still be programmes that are not yet loudness normalised. Strategies for these programmes have to be developed, like **automated normalisation** directly after ingest to a playout server or the installation of a safety loudness regulation device (Loudness Processor) at the output of Master Control to be able to handle, for example, live feeds that are not yet produced to the Target Level of **-23 LUFS**. The use of a Loudness Processor is a delicate matter, though. If the Processor operates more severely on the louder parts of a programme, a lower Programme Loudness Level may be the result, thus effectively “denormalising” the content. Consequently, **the broadcast system shall signal to the Loudness Processor when loudness-compliant content is played** (it is assumed that the mix is appropriate for broadcast). The processor should then switch to **Bypass Mode** or to a preset that only applies **safety true peak limiting**. Such signalling may be performed via GPIO or control data network systems.



For compliant material, a Loudness Processor shall be switched to BYPASS.

A Loudness Processor might also be used for live production, in the spirit of “harmonizing the source”. With appropriate settings such a processor can aid the mix engineer to actually tame some of the unpredictably loud parts of a live programme. Generally, care has to be taken not to create “loudness sausage” through overly aggressive processing, thus ruining the original intention of increasing the contrast and dynamics, which enhances the excitement of a mix.

Don't produce loudness sausage!



3. Strategies for Loudness Levelling

3.1 Basic Mixing Approach

Approaching Loudness Levelling in production offers two possibilities: the first is to immediately change the levelling habit to **loudness mixing** and **normalisation** with no or only a small level shift needed afterwards (if you don't land exactly on Target Level), and the second is to keep current (peak) levelling practices with a definite need for a subsequent level shift (Figure 3).

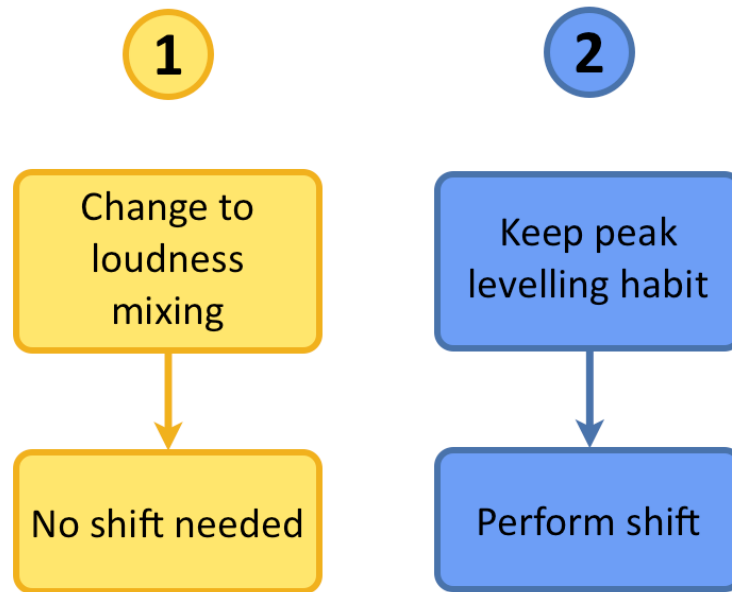


Figure 3: Two principal working methods to achieve uniform loudness in production and post-production


Levelling solution 1 (changing to loudness mixing and metering right away) is the one that is recommended in these Production Guidelines. After an initial measurement and testing period using a loudness meter to measure past programmes of the same genre, one has good guidance as to where the levels sit and how much level difference generally is to be expected and needs taking care of. Such a level difference should be accompanied by a complementary change in the listening level (see § 8.2), so that the average acoustic level while mixing stays the same! The advantages of the loudness-levelling paradigm then speak for themselves. The greater headroom will be a welcome bonus for crowd noise, for example, of sports programmes, enhancing the impact of a game for the viewers and listeners. Studio voice-overs that are often dynamically compressed due to artistic reasons (and where therefore the peak-to-loudness ratio will be lower) will be better balanced with more dynamic original location recordings etc. (see also info box on page 16).

Levelling solution 2 is more relevant for the early stages of the transition, and it may initially be more suitable to those who work on **live programmes**. The existing meters, limiters and mixing practices are retained and a level shift is done at the output of the console to achieve the loudness **Target Level** of **-23 LUFS**. A loudness meter is placed after the level shift to enable the engineers to understand the exact amount of shift (which initially is still a bit of guesswork). Using a loudness meter in parallel with a conventional meter is in any case a good idea to facilitate the transition. In this way experience can be gained before actually diving into the loudness-levelling world. As mentioned before, it is good practice to measure past mixes to get an idea how much average loudness difference needs to be compensated for. Adjusting the listening level (for example, if your average loudness level used to be **-20 LUFS**, increase the listening level by 3 dB) is the best trick to easily land close to Target Level, retaining the familiar listening level.

When keeping current levelling practices it is likely that the necessary gain shift will be *negative* (attenuation). Therefore an additional step of reducing the mixing dynamics and/or limiting the Maximum True Peak Level is usually **not** necessary. As a first creative step towards the new levelling paradigm, the thresholds and ratios of compressors and limiters can be already loosened a bit to explore the new dynamic possibilities.

For programmes that are finished in **post-production** the necessary level shift for any approach is easy to perform. Measuring the whole programme (off-line or in real-time), the necessary gain offset can be determined exactly, and in today's file-based world a gain calculation is a very quick and easy operation. Consequently, the Target Level of -23 LUFS can be achieved precisely. Nevertheless, in order to avoid a rejection of programmes due to accumulation of metering tolerances, a general tolerance of ± 0.5 LU around the Target Level of -23 LUFS is acceptable.¹

Of course, for **live programmes** it is challenging (if not a matter of luck) to achieve Target Level. Therefore, a deviation of ± 1.0 LU is acceptable for those programmes where a normalisation to the Target Level of -23 LUFS (± 0.5 LU) is not achievable practically (in addition to live programmes, for example, ones which have an exceedingly short turnaround). Experience at several broadcasters having made the transition to loudness levelling has shown that it is certainly possible for live mixes to fall within the ± 1 LU window permitted by EBU R 128.



A tolerance of ± 0.5 LU around target level (-23 LUFS) is generally accepted. For live programmes, the tolerance is ± 1 LU.

In cases where the levels of a programme's individual signals are to a large extent *unpredictable*, where a programme deliberately consists of only background elements (for example, the music bed for a weather programme) or where the dramaturgical intention of a programme makes a loudness level particularly lower than Target Level desirable, this tolerance may be too tight. It is therefore anticipated for such cases that the integrated loudness level may lie outside the tolerance specified in R 128 (being lower than -23 LUFS).

Furthermore, for programmes, which are part of a **dedicated sequence** (like, for example, the tracks of a music album or the movements of a symphony), the loudness levels may also deviate from the Target Level more than what lies within the accepted tolerance. In order that such programmes are able to pass an automated loudness workflow without getting unintentionally normalised to the Target Level one might use dedicated Metadata. Handling Metadata for such cases with deliberately different (lower) loudness levels will be covered in § [6.1.1](#) + [9](#).

In what follows, the impact of working with a **loudness meter** in production and post-production will be examined.

3.2 Loudness Metering for Production and Post-Production

An 'EBU Mode' loudness meter as defined in EBU Tech Doc 3341 offers 3 distinct time scales [see also Appendix 3 (§ [11.3](#))]:

¹ The general tolerance of ± 0.5 LU around the Target Level of -23 LUFS has been introduced in EBU R 128 revision 2 from June 2014.

- **Momentary Loudness** (abbreviated “M”) – time window: 400 ms
- **Short-term Loudness** (abbreviated “S”) – time window: 3 s
- **Integrated Loudness** (abbreviated “I”) – from ‘start’ to ‘stop’

‘EBU Mode’ also defines two scales: “EBU +9 Scale” which ought to be suitable for most programmes and “EBU +18 Scale” which may be needed for programmes with a wide Loudness Range. Both scales can either display the relative Loudness Level in LU, or the absolute one in LUFS. ‘0 LU’ in ‘EBU mode’ equals the Target Level of –23 LUFS. EBU Mode does not specify the graphical interface, so different practical solutions will be encountered.



In an ‘EBU Mode’ loudness meter, 0 LU equals –23 LUFS.

“Ready, Set (Levels), GO!”

Experience since the transition of several broadcasters to loudness metering has shown that for level setting, the **Short-term** integration window is especially useful when dealing with ‘**foreground signals**’, for example, a narrator’s voice. The 3-second-window nicely bridges most gaps between words and sentences, resulting in a stable and easy-to-read indication of the voice level. The **Momentary** Loudness Meter behaves more agile and thus provides more detail. It is up to the user to decide which of the two meters to use for basic levelling, the Momentary or the Short-term Meter – or even both. In general, it is advisable to set the levels of **foreground sounds** with a bit of **caution** initially (that means, a bit lower than –23 LUFS), as background sounds will only add to the Programme Loudness Level. Furthermore, it is psychologically easier to gradually increase the integrated loudness level during a mix (if needed) than to decrease it. Usually, a slight increase in the course of a programme is also dramaturgically more natural – and an initially “defensive” strategy leaves the engineer room to manoeuvre in case of unexpected or unpredictable signals and events.

Once levels of individual foreground signals are set, and a fixed monitor gain has been established (see § 8.2), the audio engineer can switch to **mixing only by ear**. Checking the Momentary or Short-term Loudness Level and an occasional glance at the value of the Integrated Loudness Level should give enough confirmation that the mix is on the right track towards Target Level. With a numerical readout of the ‘I’-value with one decimal point precision or a graphical display of similar resolution, **trends** can be anticipated and the appropriate measures taken. This should be performed in a smooth manner (only careful adjustments in fractions of a dB/LU, for example, of the principal foreground sound or the main fader), waiting for the corresponding result in the readout of the ‘I’-value), as too drastic changes will be artistically unsatisfactory and may result in ‘chasing’ the Target Level.

With the **Maximum Permitted True Peak Level** in production being –1 dBTP the phenomenon of ‘hitting the wall’ (meaning the former safety limiter usually operating at –9 dBFS) is now much less likely to occur. Used reasonably and with a clear intention, this ‘*opening of the lid*’ together with loudness normalisation to –23 LUFS results in potentially **more dynamic mixes**, in less dynamic-compression artefacts like pumping and thus in an **overall increase of audio quality!** Programme makers who favoured dynamic mixes in the past are now relieved from potential compromises because their programme would sound softer than more compressed ones. With loudness normalisation, this compromise is gone. At last!

The elements of a mix that are most important for a uniform subjective loudness impression are so-called 'foreground' sounds – like voice, title music or key sound effects. Individual sound elements do have a widely varying difference between their loudness level and their peak level (their 'Peak-to-Loudness Ratio' (PLR)). For example, the 'clink' of two glasses when toasting has a high peak level, but quite low loudness level. On the other hand, a dynamically compressed hard rock guitar riff has a loudness level that is almost the same as its peak level! If those two signals are aligned according to their peaks, the guitar riff will be much louder than the clink of the glasses. This example is meant to illustrate the concept; it does NOT mean that those two signals should be mixed at equal loudness! The level of individual elements and components (like pre-mixes or stem-mixes, a music-only mix or a voice-over track) in the mix is an artistic decision, naturally, but loudness metering can help the mixer with useful visual feedback that actually shows what he or she hears!

Coming back to metering, at the end of a programme there are two scenarios:

- Having hit Target Level (-23.0 LUFS) or
- Having missed Target Level in either direction

Understandably the second scenario will be more likely, also for post-produced programmes. If the actual loudness level is within the accepted tolerance of ± 0.5 LU (or ± 1.0 LU for live programmes), then no further action is needed. If the level lies outside the tolerance, this is still acceptable from a generic production standpoint (as mentioned earlier). In a post-production situation, a simple **gain calculation** will put the programme at Target Level. For live programmes not "on target", correction measures may be taken downstream in the form of **loudness processors** that gradually adjust the integrated loudness level of such programmes in an unobtrusive manner and can act as a sort of '*loudness safety net*'. This must be achieved in a way such that the inner dynamics of the production are not harmed. The processor may only be needed for live programmes if the workflow for file-based programmes is already fully compliant with EBU R 128. If a downstream dynamics and loudness processor is situated at the output of the **Master Control Room**, it should be able to be (automatically) **bypassed** for programmes compliant with R 128 and appropriate for broadcasting (see also § 2.4). This bypass situation is expected to become the normal way of working, the more programmes are "on target", as the recommended goal is to **normalise the audio signal at the source**.

Especially in the **transition phase** moving towards loudness normalisation such aforementioned loudness processors downstream may be helpful for broadcasters to adapt to the loudness levelling system and to catch possible outliers. **It should be the goal of the broadcaster (and also the mixing engineers) to have these processors work less and less and as little as possible, as the integrated loudness level and the dynamic properties of programmes are increasingly within the accepted tolerances. Ultimately, this will result in the elimination of a loudness processor altogether!**

3.3 Loudness Range

The measure **Loudness Range (LRA)** quantifies the **loudness variation** of a programme. In the past, it had to be 'educated guesswork' of experienced audio personnel to decide if a programme would fit into the loudness tolerance window of the intended audience. Using Loudness Range at the end of the measurement period (usually the whole programme), a single number helps the mixer/operator to decide if further dynamic treatment is necessary (For programmes with very wide mixing dynamics a different normalisation strategy as well as additional measures may be useful – for example, the average voice level, its distance to the overall programme loudness level or the maximum short-term loudness level. See § 9.2 + 9.3 for a detailed description of potential strategies for such programme (for example, feature films)).

Working with loudness normalisation right away thus implies observing and potentially controlling also Loudness Range as the dynamic possibilities are expanded. This is important to ensure an appropriate signal for the intended audience and distribution chain. Whereas in production and post-production a 'generic' mix may be created (with a relatively high LRA value and a Maximum Permitted True Peak Level of -1 dBTP), different platforms may need a lower LRA value and a lower Maximum Permitted True Peak Level (while keeping the Programme Loudness Level at -23 LUFS). The system within R 128 appreciates this generic approach with further processing downstream to tailor the signal to individual environments and platforms.

With the measure Loudness Range (as well as, where appropriate, additional measures like the average voice level, its distance to the overall programme loudness level or the maximum short-term loudness level – see above) it is now more systematically possible to determine appropriate strategies for *potential dynamic treatment* of a programme to fit it to the tolerance window of the audience or distribution platform. For dynamic programmes that consist mainly of *music*, overall **low-level compression** may lead to satisfactory results (see Figure 4 as an example): a low threshold (< -50 dBFS) and a moderate compression ratio (1:1.2 – 1:1.5) ensure uniform compression of the whole signal range. Dependent on the original loudness level, a shift to the Target Level of -23 LUFS may be performed in parallel through adjusting the make-up gain of the compressor accordingly.

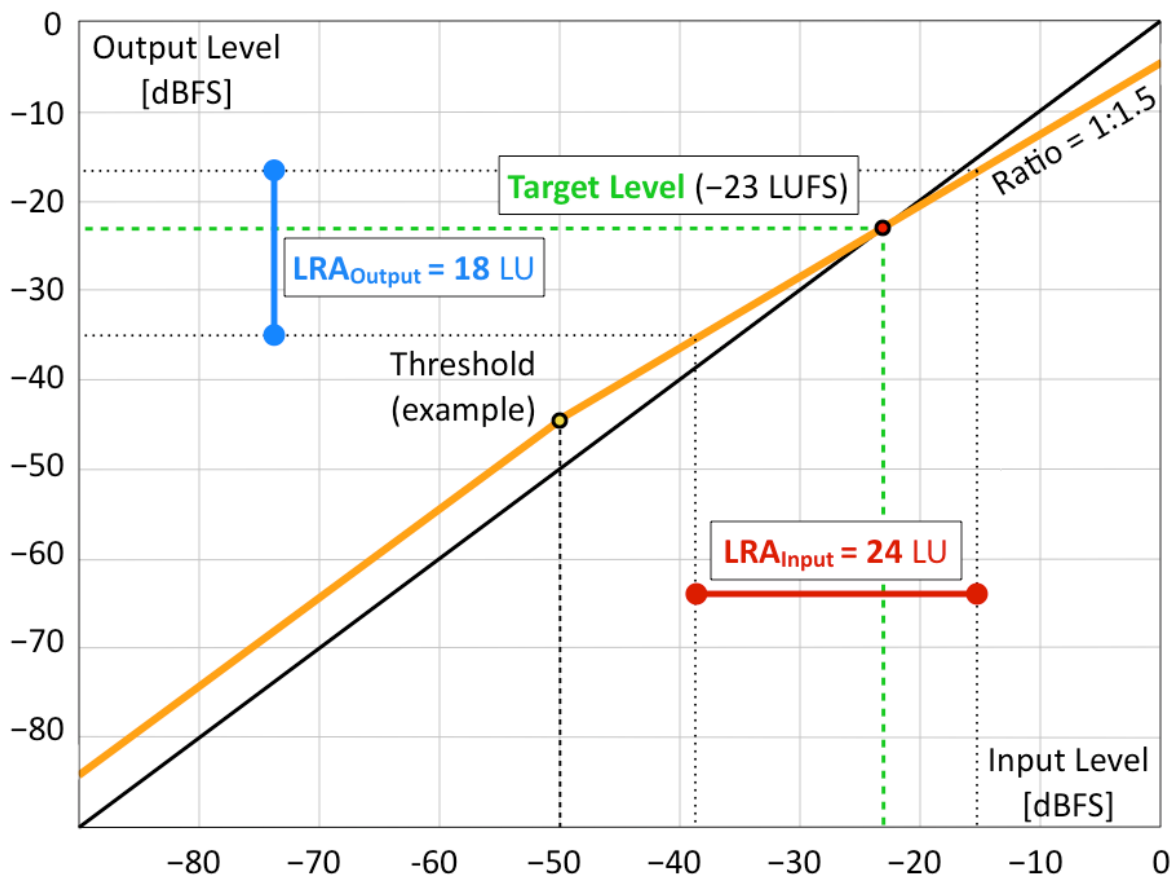


Figure 4: Example for processing of Loudness Range (LRA) with a compressor with a low threshold (-50 dBFS) and a moderate compression ratio (1:1.5)

A specific approach for the genre feature film is described in § [9.2](#).

It is important to understand that it is impossible to define one maximum value of LRA for all broadcasters and all programmes. Furthermore, LRA is very useful as a mixing tool and should not be a brickwall parameter for delivery specifications of programmes. Nevertheless, **recommended**

individual maximum values for LRA can provide a good dynamic framework for different formats (for example, 5.1 vs. 2.0), genres, distribution platforms as well as different replay environments. The average listening environment, age of the target audience, 'listening comfort zone' of the consumer and other parameters all influence the acceptance of an LRA value for specific programming. The **Loudness Range Control Paradigm** starts from a generic accepted maximum value of Loudness Range according to the principles described above and adapts this value downstream to comply with technical necessities of individual distribution platforms and replay environments.

In any case, no parameter and a corresponding maximum allowed value can guarantee a good mix! This is also true for Loudness Range. To judge the quality of a mix, experienced listeners have to evaluate the programme with their ears. LRA gives general guidance regarding the basic dynamic properties of a mix, it can furthermore be used to steer dynamic treatment in a loudness processor, and the development of LRA over time can be used to distinguish junctions of audio elements in a row where the start and end of these individual elements are not known. Loudness Range does have its use in describing a programme in more detail and/or instigate dedicated processing.

As a result of the need for different values of Loudness Range, **EBU R 128** does *not* include a maximum permitted LRA value, but instead encourages the use of the measure **Loudness Range** to evaluate the potential need for dynamic range processing according to the different criteria mentioned above. To give an **example**, some EBU broadcasters have chosen a maximum value of 20 LU for LRA for surround sound programmes, up to which no dynamic reduction is performed. For stereo programmes, the value some broadcasters have chosen is 15 LU. Other broadcasters might have chosen different values or none at all! It is important to note that any of these values can only give general guidance; **they should not be followed too strictly!** A certain **flexibility or tolerance** above these values should be allowed (for example, +2 LU), as LRA might not give all the necessary information to decide if and what kind of dynamics processing is needed.



If a maximum value for Loudness Range is used, this limit should not be too strict (allow up to 2 LU more).

Loudness Range is also a useful indicator of potential *dynamics reduction processes* in a signal chain, performed on purpose or accidentally. If the LRA value of a programme after it has passed through a processor chain is, for example, lower than it was originally, such a reduction process has occurred.

3.4 Climbing the True Peak

The third measure recommended by R 128 concerns the **Maximum True Peak Level** of an audio signal. Having abandoned the peak normalisation paradigm, it is of course **still vital to measure and control the peaks** of a programme, and especially its maximum peak to avoid overload and distortion.

A loudness meter compliant with 'EBU mode' (see *EBU Tech Doc 3341*) also features the measurement and display of the true peak levels of a programme. Safety limiters to avoid overmodulation will have to be able to work in *true-peak mode* and need to be adjusted to the appropriate Maximum Permitted True Peak Level, in production as well as at the output of master control, at the distribution head end and the transmitter site. Next to the Maximum Permitted True Peak Level for generic PCM signals in production (**-1 dBTP**), further values for different applications and distribution systems are given in EBU Tech Doc 3344 ('*Distribution Guidelines*').



The Maximum True Peak Level in production (linear audio) is -1 dBTP.

3.5 *Advanced Live Mixing Strategies*

3.5.1 Sports

Sports is arguably one of the more challenging genres as far as loudness levelling and normalisation is concerned. This is due to the sometimes-unpredictable nature of the event. A few goals in the last 15 minutes of a football match, for example, can boost the Integrated Loudness Level considerably, resulting in a value outside the tolerance specified in R 128 (-23.0 LUFS ± 1.0 LU). There is basically little that one can do about that, unless one is prepared to severely influence the dynamic properties of the signal (using a loudness processor at the output of the mixing desk or the Master Control room). In any case, it is advisable to have the voice loudness level(s) of the commentator(s) sit a bit below Target Level (at -24 LUFS, for example), so that unexpected crowd noise has 1 LU (Integrated Loudness!) more room to move. If such audience reactions don't happen, the average loudness level will then obviously be lower than Target Level, but usually still within the tolerance.

The same thinking applies if the commentary density of different programmes is varying. In one event, there might, for example, be two commentators who talk most of the time rather excitedly. In another event, there might be only one commentator who talks less frequently, and with a softer voice. If in the second event the crowd noise is above the gate threshold (but several LU below the Integrated Loudness Level), those parts will 'drag' the average loudness lower than -23 LUFS, when the commentator doesn't speak. In the generic sense of R 128 the second programme would have to be boosted in order to sit at Target Level. As a consequence, the commentary of the second programme would be perceived louder than the two commentators of the first programme. This is anticipated. To have the commentary level in both cases be equal, the mixer/broadcaster could qualify the second event as a 'special circumstance' and have its Integrated Loudness Level lower than -23 LUFS, effectively performing anchor-based normalisation.

For sports with a rather quiet atmosphere (for example, golf), the relative gate within the integrated measurement will eliminate most of the pauses of the commentary during the loudness calculation. Such a programme should consequently 'land' easily within the tolerance around -23 LUFS, if the commentator(s) are levelled around -23 LUFS.

3.5.2 Shows

In Entertainment Shows, for example Game or Music Shows, the predictability of the event is certainly higher than in Sports as there is a concept, a storyboard so to speak. What is similar is an obvious anchor signal: the host, the moderator(s). But also the **audience** always plays a vital part as it transports much of the emotion and excitement. Therefore the audience is as important a signal as the moderator(s)! Consequently, it may be more advantageous to balance the **audience** around Target Level - and have the moderator fly above and below. The exact choice of foreground or anchor sounds is dependent on the individual programme. For Music Shows, obviously the music is the most important signal and will mainly determine the Programme Loudness Level. The moderator will then probably sit below -23 LUFS, but this is fine, as long as the signal is still within the 'comfort zone' of the listener (about $+3/-5$ LU around Target Level).

If the Show to be mixed is expected to be exceptionally vivid and loud, one strategy could be to temporarily increase the monitor gain (1-2 dB). This usually avoids being carried away too easily with the excitement and having to deal with a final Loudness Level way above the tolerance.

4. What to Measure in Production and Post-Production

4.1 Signal-Independent vs. Anchor-Based Normalisation

EBU R 128 recommends measuring the **whole programme**, independent of individual signal types like voice, music or sound effects (see Figure 5). This is considered to be the most generally applicable practice for the vast majority of programmes:

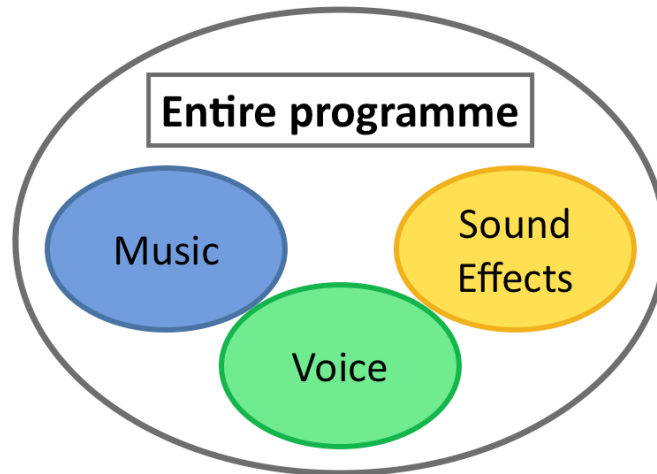


Figure 5: Elements of a programme

For programmes with a very wide Loudness Range (>20 LU, approximately) or with a significant difference between the Programme Loudness Level and the Voice Loudness Level (>3-5 LU, approximately) one may optionally use a so-called **anchor signal** for loudness normalisation, thus performing a *signal type gating method*, so to speak. This signal might be speech or a singing voice, or, for example, a certain part of a music programme in *mezzo forte*. Such an anchor signal will typically have a lower loudness level than Programme Loudness (PL). **Anchor-based loudness normalisation consequently leads to higher values of PL than the Target Level.** If special processing is applied (as described in § 9.2), this is anticipated and still within the spirit of R 128.

It must be emphasized, though, that choosing an anchor signal generally is an **active process** requiring input from an experienced operator. This approach should only be considered after operators and sound engineers have become very comfortable with the concept of loudness normalisation. Performed well, it may help to fine-tune the loudness normalisation of wide loudness range programmes according to the chosen anchor signal.

There also exists an automatic measurement of one specific anchor signal in the form of '*Dialogue Intelligence*', a proprietary algorithm of Dolby Laboratories, anticipating that speech is a common and important signal in broadcasting. The algorithm detects if speech is present in a programme and, when activated, only measures the loudness during the detected speech intervals. For programmes with a narrow loudness range the difference between a measurement restricted to speech and one performed on the whole programme is small, usually <1 LU. For programmes with a very wide Loudness Range, such as action movies, this difference gets potentially bigger, sometimes exceeding 8 LU or more! Automatic detection of an anchor signal is intended to help identify what should be at Target Level. Like any algorithm for detecting specific signals out of a complete and complex mix, speech discrimination can be tricked — either by signals closely resembling the spectral pattern of speech (for example, certain woodwind instruments or a solo violin) or by speech signals that are too far off the discrimination threshold (for example, certain language dialects). For programmes where these anchor signals are consistently moving around the discrimination threshold, the loudness measurement can also vary if the measurement is performed repeatedly. Furthermore, the level of speech can vary significantly over the duration of a programme, sometimes up to 20 LU! In such cases, "speech level" may be quite ambiguous [7].

For **short** programmes like commercials, advertisements, trailers and promotional items, (automatic) speech normalisation is likely to give unsatisfactory results in the light of the future increase of mixing dynamics and potentially enhanced dramaturgical concepts. In such cases, international recommendations (also this one) agree on measuring ‘all’ by all means.

In any case, broadcasters have to be aware that especially in a file-based environment, where for most content the whole programme independent of signal type (speech, music, sound effects) will be measured automatically, a different strategy might have to be established to treat programmes based on anchor normalisation.

To summarize:

It is because of these uncertainties and the fact that speech represents only one part of the whole programme (albeit a very important and common one) that R 128 generally recommends measuring ‘all’ – that is the whole programme, independent on the signal type (such as voice, music or sound effects).

This is supported by the following observations:

- The difference between measuring ‘all’ and measuring an anchor signal (such as voice, music or sound effects) is small for most programmes;
- The difference between ‘all’ and ‘anchor’ measurements depends strongly on the content of the programme, but can be expected to be bigger if the Loudness Range is bigger;
- Automatic anchor signal discrimination may perform well for a majority of programmes, but may be tricked by similar signals or may not trigger at all, thus not giving 100% consistent results;
- Identifying an anchor signal needs input from an experienced operator or a discrimination algorithm; such an algorithm may be subject to the potential uncertainties listed above.

Still, anchor normalisation can offer better results on wide LRA material. It is however a task requiring expertise, and if automatic discrimination is used, such an algorithm cannot be 100% reliable. Special measures need to be taken when anchor-adjusted content enters normalisation systems on file servers.

For **feature films** and other similarly dynamic programmes a **more elaborate strategy** may be used, using the measurement of additional parameters like Voice Loudness, its difference to Programme Loudness, the variation of the Voice Loudness Level or the Maximum Short-term Loudness to shape the dynamics processes accordingly. This will be covered in more detail in § [9.2](#).

4.2 Low Frequency Effects (LFE) Channel

As noted in the description of ITU-R BS.1770 (see Appendix 1 (§[11.1](#))), the LFE channel is currently **excluded** from the loudness measurement. One of the reasons is the widespread uncertainty of consumers and audio engineers as well as equipment implementation differences regarding the alignment of this channel (+10 dB in-band-gain). The omission of the LFE channel during the loudness measurement might cause its abuse. Further investigations of this matter and practical experience are needed to decide if and in which way the LFE signal might be included. One solution to completely avoid all potential issues with the LFE signal is not to use it at all (“5.0” Surround Sound) if there is no need for extra headroom in the low frequency region.

5. File-Based Playout and Archives

As the broadcast world is changing to file-based workflows, it is vital that the loudness normalisation concept is also fully embraced there. The basic principle stays the same: **loudness normalisation** and **dynamic control** of the **audio signal at the source** is recommended for the production of new content. Nevertheless, as **Metadata** is an integral part of archival systems, solutions that rely more on Metadata are described as well (§ [6](#)).

The origin of a broadcast file that contains audio signals can be via an ingest process, via transfer from an external server and from a file-based archive.

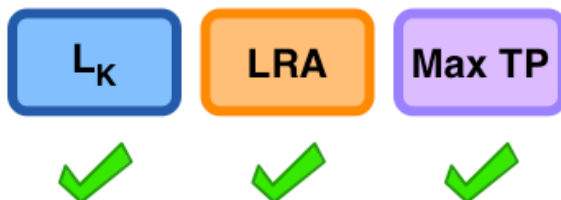
At the very beginning of a file's life inside a facility, measurements have to be made, providing the values for **Programme Loudness Level**, **Loudness Range** and **Maximum True Peak Level** - the three characteristic audio parameters defined in EBU R 128 (Maximum Short-term Loudness Level may be measured and stored too, especially for short form content (typically <30 s, see § 9.1)). Depending on the results of these measurements and the subsequent method to achieve loudness normalisation and compliance with the acceptable Loudness Range, different processing schemes may be executed. Some elements of the workflow will now be examined in more detail.

5.1 Loudness Levelling Strategies — Processing

At the beginning of any potential processing the three main parameters **Programme Loudness Level** (L_K), **Loudness Range** (LRA) and **Maximum True Peak Level** (Max TP) are measured. The result of this initial measurement determines the subsequent processing. The basic processing at the core of any file quality control process concerns the Programme Loudness Level.

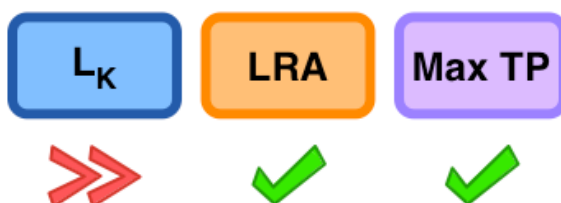
Several different scenarios are possible:

- a) All three parameters are OK.



This is obviously the ideal outcome of the measurement: the **Programme Loudness Level** is $-23 \text{ LUFS} \pm 0.5 \text{ LU}$, the **Loudness Range** is not exceeding the specified tolerance of the broadcaster (depending on the genre and/or the distribution platform) and the **Maximum True Peak Level** is equal to or below the specified maximum value for the designated distribution system.

- b) The Programme Loudness Level is higher than -23 LUFS .



A simple gain ranging (level reduction) operation solves that:

$$\text{Gain (dB)} = L_{K\text{Target}} - L_{K\text{measured}}$$

(Example: the measured L_K is -19.4 LUFS ; Target Level is -23 LUFS ; the necessary gain is $(-23 - (-19.4)) = -3.6 \text{ dB}$. Max TP is naturally reduced by the same amount as L_K .) This gain operation can either be performed directly on the audio file (options € and •) or during playback (option ,).

- c) The Programme Loudness Level is lower than -23 LUFS.



After applying a positive gain offset, the Maximum True Peak Level has to be recalculated (originally measured Max TP + gain offset = resulting Max TP). If the new Max TP exceeds the permitted limit, True Peak Limiting has to be performed.

For both scenarios b and c a simple gain value stored as Metadata may be used with potential subsequent limiting if Max TP is exceeded after a positive gain offset (scenario c). This gain value can control the payout level of the file so that -23 LUFS is reached.

EBU R 128 also allows programmes to be transmitted at a lower level than -23 LUFS. Such a programme has to have clear labelling so that it doesn't get normalized unintentionally (see § 6.1.1).

- d) The Programme Loudness Level is lower than -23 LUFS and Loudness Range is much wider than the internal tolerance for the genre or distribution channel.



This is most likely to be found in programmes like feature films or classical music. Optimal processing of the dynamic properties and a potentially different normalisation strategy are dependent on the content and may need further measures. Also, more sophisticated automatic processes with a 'target-LRA'-value or processing based on the difference of the Voice Level and the Programme Loudness Level (amongst other parameters) are increasingly coming to the market (September 2015). Dynamics processing should in any case be performed reasonably avoiding "sausage processing". Refer to § 9 for appropriate genre-specific strategies.

- e) The Maximum Permitted True Peak Level is exceeded.



Exceeding the Max TP level of the respective distribution system incurs a risk of **distortion** further downstream (in a D-to-A-converter, sample-rate converter or bitrate-reduction codec, for example). True-Peak Limiting should be applied to lower Max TP. Whether there is a significant change to Programme Loudness as a result of this depends on the number and size of the peaks that are affected.

5.2 Archival Content

For existing programmes (archival content) there are basically **four options** to achieve loudness normalisation:

- € actually changing the loudness level of all audio files to be 'on target'
- changing the loudness level only 'on demand'
- ,
- using the result of a loudness level measurement to adjust the playout level without changing the original loudness level
- f transporting the correct loudness Metadata to the consumer where normalisation is performed by appropriate home equipment

The first three options result in a Programme Loudness Level of -23 LUFS and are therefore the preferred solutions because they typically result in more headroom than the fourth option (see § 2.2, 2.3). The final choice among these three options depends on factors like specific infrastructure, workflows, media asset management, availability of suitable equipment, financial resources, time etc.

6. Metadata

As described in § 2.2, Loudness normalisation can be either achieved through normalisation of the audio signal (the recommended method) or by using Metadata to store the actual loudness level. For the latter, the shift to Target Level can be performed either during the transfer of the audio file to the playout server, in the playout audio mixer, through choosing the appropriate preset of a downstream dynamics processor or directly at the consumer end with an adjustment of the playback level.

Metadata generally can be *active* (potentially changing the audio signal) or *descriptive* (providing information about the signal, such as format, copyright etc.). As a natural consequence of the work within PLOUD and the publication of EBU R 128 and its supporting documents, the three main parameters Programme Loudness, Loudness Range and Maximum True Peak Level shall form the core of loudness Metadata in audio files. These parameters can be stored in the header (Broadcast Extension (BEXT) chunk) of the Broadcast Wave File (BWF) format (*for a detailed description of BWF, see [8], [9] and [10]*). Furthermore, the values for the Maximum Momentary Loudness Level as well as the Maximum Short-term Loudness Level may be stored, as these parameters are helpful for controlling the dynamics especially of short form content (typically <30 s; see also § 9.1). Recently (January 2014), the EBU has published an open metadata standard, the Audio Definition Model [11], to ensure compatibility across all systems. The 5 parameters mentioned above are an integral part of this model. Loudness Metadata is also intended to be included in the SMPTE dictionary with potential refinements like 'Loudness Profiles', addressing, for example, different processing presets of downstream loudness processors.

The Metadata parameters in existing systems that are of primary interest concerning loudness are:

- Programme Loudness
- Dynamic Range Control
- Downmix Coefficients

For example, in the Dolby AC-3 Metadata system, these parameters are called *dialnorm* (dialogue normalisation), *dynrng* (dynamic range) and *Center/Surround Downmix Level*. The parameter *dialnorm* genuinely describes the loudness of an entire programme with all its elements such as voice, music or sound effects (also a music-only programme has a 'dialnorm' value). This may seem confusing; the reason is the focus of the Dolby system on normalisation according to the anchor signal dialogue, due to its roots in film mixing. Thus, if anchor-based normalisation is performed (with speech or dialogue as the anchor signal), the Metadata-parameter *dialnorm* actually describes the loudness of the dialogue, but only in that case.

6.1 Programme Loudness Metadata

Following the emphasis on normalising the audio signal in production to -23 LUFS, the relevant Metadata parameter shall naturally also be set to indicate -23 LUFS, provided the programme has

been normalised to the Target Level. Consequently, after widespread normalisation of the source audio signals the Programme Loudness Metadata parameter will be **static**. In any case, Programme Loudness Metadata shall always indicate the actual Programme Loudness.

Exceptions where a different value than -23 may be used are:

- The programme does not fit into the window provided by -23 LUFS and -1 dBTP. This may occur mainly with very dynamic feature films and/or those with a significant difference between Voice Loudness (VL) and Programme Loudness (PL);
- Legacy programming from the archive may not be able to be adjusted in time to fulfil the Target Level system of R 128;
- External live programmes may be provided with different loudness levels and Metadata;
- A fully functional system of providing and using Metadata over the whole signal chain is already in place. This implies faithful transportation of Loudness Metadata to the consumer's home equipment.

In all these circumstances the **correct** Metadata value for Programme Loudness, measured with an 'EBU Mode' compliant meter, shall be set by all means. For the case of the first bullet point (feature films), special processing may be applied (reduction of the difference between VL and PL, limiting Maximum Short-term Loudness etc.) that may furthermore indicate the use of anchor-based loudness normalisation (using mainly Voice Loudness as the anchor signal). Consequently, the Metadata parameter for loudness normalisation on the consumer's side may reflect Voice Loudness for these programmes.

6.1.1 Deliberately Lower Programme Loudness, Loudness Offset

For programmes with a deliberately lower Programme Loudness Level (consisting, for example, mainly of background sounds), this shall be clearly indicated; Care must be taken that such lower-loudness programmes do not get compensated accidentally. This is also relevant for music programmes like the movements of a symphony or short sound design production elements in Radio, which may be played at a different level than Target Level (also higher) for dramaturgical purposes.

The recommended solution introduces another Metadata parameter called "**Loudness Offset**" which may be positive or negative (the default value is zero; it is recommended to use a 3-digit signed number with one digit after the decimal point). Metadata for Programme Loudness shall always indicate the actual Programme Loudness, and the Loudness Offset indicates if compensation to the Target Level is wanted (Loudness Offset = 0) or if the programme should be at a different loudness level than Target Level. The actual value of the Loudness Offset parameter thus indicates the distance from Target Level. Consequently, if the Target Level changes (for example, in a streaming situation with limited playback gain and headroom), this parameter ensures that the relative loudness levels stay the same.

Some broadcasters already use a "**Low Loudness Flag**" in addition to keeping programmes in their original form (including lower Programme Loudness Metadata if provided). Not normalising these programmes to -23 LUFS leaves the original headroom intact and prevents unnecessary true-peak limiting. This is a special case of "**Loudness Offset**" and is indicated by choosing the lowest possible value for this parameter (-99.9).

Positive values for "**Loudness Offset**" must be the **exception** and treated with **extra care**. The Maximum True-peak Level may lie outside the tolerance after boosting the programme. Also, the Programme Loudness may be too high and compromise listener comfort.

It is eventually the responsibility of the broadcaster to ensure the correct treatment of audio signals with an average loudness level deliberately different from Target Level.

6.2 Dynamic Range Control Metadata

Just as loudness normalisation can be performed at the source audio signal or via Metadata, the same applies to dynamic range processing. In the Metadata environment, dynamic range compression information is sent as part of the datastream in the form of *gain-words* and subsequent *profiles* (see Appendix 4 for details on the Dolby DRC profiles (§ 11.4)). In the Home Theatre Equipment of the consumer, this information is applied to reduce the dynamic range of the signal, either by default or after user activation. Dynamic range control through the use of Metadata is not comparable with a sophisticated dynamics processor, but it provides a 'band aid' for situations where the consumer wants a considerably lower loudness range.

The control of Loudness Range through actual processing of the audio signal at transmission is generally shifting the issue upstream. In the transmission system Dolby Digital, a broadcaster may choose a *gentle* profile for *RF mode* (to avoid too active overload protection) while still choosing 'None' for *Line mode*. Broadcasters that need other profiles than 'None' to support their internal workflow must be aware that this functionality may not always be implemented reliably in their listeners' equipment. **Manufacturers and distribution companies are advised to ensure that equipment is made in accordance with EBU Tech Doc 3344 ('Guidelines for Distribution and Reproduction').**

6.3 Downmix Coefficients

These Metadata parameters are obviously only applicable to surround sound signals, controlling the gain (in dB) of the centre channel and the surround channels when mixed to Left front and Right front to derive a 2-channel-stereo signal. The loudness of a 2.0-stereo signal, which is the result of a manual downmix or an automatic one using Metadata, is dependent on several factors, which will be described in § 7, together with general implications on loudness for surround sound and stereo programmes.

For *external* files (or other media), there can be no guarantee that the Metadata supplied are correct. Programme Loudness Metadata indicating -27 (the factory default for *dialnorm* in the Dolby-Digital system) or -31 (the lowest possible value in that system) are likely to raise special awareness, as chances are that Metadata have either not been looked at or been abused for the programme to appear (much) louder when replayed at the consumer's side.

It is therefore recommended to **discard loudness and dynamic range control Metadata** for external sources except where the source can be fully trusted. An entire measurement process of the three main audio parameters needs to be conducted afresh. Only this will ensure the correct subsequent processing. For internal purposes, Metadata can be better controlled.

7. Surround Sound vs. Stereo — Downmix and Upmix issues

7.1 Downmix

Downmixing of a 5.1- or 5.0-Surround Sound signal is performed regularly and in two ways: **during production** as a dedicated manual process to derive a 'custom' 2.0-stereo signal or **during transmission** in the home receiver of the consumer according to the Downmix-Metadata sent within the bitstream. There exist two different Downmix methods in a receiver: **Lo/Ro** (Left only/Right only) which directly combines the channels according to the downmix coefficients, and **Lt/Rt** (Left total/Right total) which applies $\pm 90^\circ$ phase shifting to a Mono-sum of the surround channels in order to achieve better compatibility with matrix-surround playback systems (for example, Dolby Surround).

The Programme Loudness Level (PL) of the resulting 2.0-stereo signal is dependent on:

- The chosen Downmixing method (Lo/Ro or Lt/Rt)
- The actual downmix coefficients themselves (+3/+1.5/0/-1.5/-3/-4.5/-6/- ¥)

- The programme content in the centre and surround channels
- The correlation between the channels and
- Potential safety-limiting to avoid overload

Ideally, a downmix operation should be **loudness-agnostic**. This is especially challenging for the Downmix method Lt/Rt, but in general for multichannel mixes with very active surround signals. The weighting of the two surround signals is +1.5 dB in the algorithm specified in ITU-R BS.1770¹, but the default downmix coefficient for these signals is -3 dB! As a result, the difference in loudness terms regarding the surround signals of the two mixes is 4.5 dB!! To achieve the same PL for the downmix signal consequently necessitates either a subsequent loudness measurement and potential static gain correction or a sophisticated real-time process that constantly analyses the development of the PL of the downmixed signal and performs careful and unobtrusive adjustments.

Mainly due to the artefacts of matrix surround systems, an Lt/Rt Downmix is even more unpredictable as far as the resulting Loudness Level is concerned. Together with the general sonic alterations, this is the reason why the EBU recommends the Downmix method Lo/Ro as the default setting for the relevant Metadata parameter (“Preferred Downmix Method”).



The Downmix method Lo/Ro (Left only/ Right only) is recommended as the default setting for relevant Metadata parameter.

Care should also be taken to **avoid overload** of the downmixed signal. This can be achieved with a dynamics processor upstream. Static scaling (overall level reduction) should be avoided, as it systematically introduces loudness differences between the 2-channel stereo downmix and the original surround sound signal. Dynamic scaling may offer a solution.

The downmix coefficients possible within, for example, the Dolby Digital system are governed by two downmix profiles. Initially, when there was only one profile, the parameters were coarser, with -3/-4.5/-6 dB for the Centre and -3/-6/-∞ dB for the Surround channels. Now, Extended Bitstream Information (Extended BSI) provides the finer intermediate steps listed above (in the second bullet point; also DVB TS 101 154 downmix coefficients offer the same resolution as the Dolby Digital Extended BSI). Broadcasters should be aware of the fact that not all reproduction equipment is able to deliver the intended downmix experience if Extended BSI is used, as legacy decoders may not be able to extract this information and would fall back to the fewer and coarser coefficients of profile 1.

¹ The +1.5 dB weighting coefficient for the surround signals in a loudness measurement according to ITU-R BS.1770 is not to be confused with the actual +3 dB gain for the surround signals for a cinema mix! In the cinema, the combined loudspeakers for the two surround channels are aligned 3 dB lower in level than the front channels (in a 5.1 system), so that their total level equals one front channel. The reason for this is compatibility with Mono-Surround movies (matrix-encoded ‘Dolby Stereo’ has only a (band-limited) mono-surround signal), where both surround channels would get the identical signal. For discrete multichannel audio mixes (5.1 etc.) the surround signals in the final mix are therefore 3 dB ‘hotter’, as the mixing engineer compensates for the 3-dB-lower alignment of the surround channels. If a cinema mix is broadcast, this 3-dB difference has to be compensated. This is often done while repurposing the feature film for broadcast.

Whereas the +3 dB gain for the surround channels has a purely technical reason, the +1.5 dB gain for the surround signals in the loudness measurement has psychoacoustic reasons. Humans appear to perceive direct sounds coming from the back louder than frontal ones with the same sound pressure level. A measurement device does not have a brain and thus needs this gain factor.

In the case of missing or unreliable downmix Metadata, a good starting point is to look at the coefficients described in ITU-R BS.775-2 [\[12\]](#):

L, R front:	0 dB
C, LS, RS:	-3 dB

7.2 Upmix

Upmixing a 2.0-stereo signal to 5.1- or 5.0-surround sound is becoming more popular due to the improved quality of upmixing algorithms and the interest of broadcasters to provide a spatially more homogenous listener experience. An upmix should never replace an original discrete multichannel audio mix, though. Similar loudness dependencies as for downmixing apply:

- The actual upmix algorithm (channel correlation, centre extraction, front-back level, etc.)
- The programme content and dispersion of signals on the stereo base and
- The downmix of the upmix and its loudness

There are upmix algorithms on the market that are able to downmix perfectly, resulting in exactly the original 2.0-stereo signal (if the recommended Downmix method Lo/Ro is chosen). For such cases, the PL of the downmix would be the same. The PL of the upmix, though, is a different matter. In a real-time installation, the upmix algorithm needs to constantly monitor the development of PL in order to be loudness-agnostic. For file-based operation, the resulting surround sound signal can be subsequently measured and gain corrected, just like the downmix case.

8. Alignment of Signals and Listening Level

8.1 Electrical Alignment Signal and Level

An Alignment Signal in broadcasting consists of a sine-wave signal at a frequency of typically 1 kHz, which is used to technically align a sound-programme connection. In digital systems the level of such an Alignment Signal is 18 dB below the maximum coding level, irrespective of the total number of bits available (-18 dBFS). The switch to loudness normalisation does NOT change this approach, as electrical alignment does not imply a mandatory connection to loudness metering or measurement.

Therefore, electrical alignment for sound-programme exchange can be performed as usual, with a sine-wave signal of 1 kHz at a level of -18 dBFS.



The alignment level for sound-programme exchange does not need to change. Use a 1 kHz sine wave at -18 dBFS as usual.

This is specified in EBU R 68-2000 [\[13\]](#). In the same document the “Permitted Maximum Level (PML)” is still mentioned, as defined in ITU-R BS.645-2 [\[14\]](#); with the change to the “Maximum Permitted True Peak Level” (-1 dBTP in production for generic PCM signals) the recommended PML of -9 dBFS in ITU-R BS.645 becomes obsolete. The relevant sections of EBU R 68-2000 and ITU-R BS.645 (as well as documents that refer to the definition of “Permitted Maximum Level” within

these recommendations) potentially need to be revised.

The alignment level of **-18 dBFS** (1 kHz tone) will read as **-18 LUFS** on a loudness meter with the absolute scale (or +5 LU on the relative EBU-mode scale), provided that the 1 kHz tone is present (in phase) on both the left and right channel of a stereo or surround sound signal. If the 1 kHz tone of **-18 dBFS** is used only in a single front channel, the loudness meter will read **-21 LUFS** (or +2 LU on the relative scale).



A stereo 1 kHz sine wave at -18 dBFS reads as -18 LUFS absolute (+5 LU relative) on an EBU mode loudness meter.

8.2 Acoustical Alignment, Listening Level

The Reference Listening Level of a loudspeaker reproduction setup characterizes the sensitivity of a playback system and generally describes the bias point (0 dBr) of the corresponding volume controller. It is used to set the reference gain for the repeatability of level adjustments between different listening and mixing sessions. Since the Reference Listening Level also represents a reasonable sound pressure level for home reproduction of audio signals with higher Loudness Range, most audio mixes should at least be validated at this level.

It has become apparent that the relevant measurement instructions according to EBU Tech Doc 3276-E-1998 (and Supplement 1-2004, extending it to Multichannel Sound) [15] are **not convenient** in the context of loudness normalisation and need to be revised. For that reason it is recommended to adjust every main loudspeaker to the Reference Listening Level ($L_{LISTref}$) as follows. This procedure results in reproduction levels comparable to EBU Tech Doc 3276 Supplement 1 and is applicable to every playback system with a channel configuration from 1.0 to 5.1.

Each of the main loudspeakers should be adjusted such that:

- $L_{LISTref} = 73 \text{ dB}_C \text{ SPL}$, using a 500-2000 Hz reference noise signal at **-23 LUFS**

To achieve this, a monophonic test signal of noise of equal energy per octave (pink noise) and covering a frequency range from **500 Hz** to **2000 Hz** should be used. For the generation of this signal, filters with a slew rate that at least meet the requirements of third octave band filtering according to IEC 61260 [16], have to be employed. The testing signal should be adjusted to a Programme Loudness Level of **-23 LUFS**. Under these conditions, the loudspeaker gains should be raised until a Reference Listening Level ($L_{LISTref}$) of **73 dB_C Sound Pressure Level** per loudspeaker is achieved. The measurements should be made at the reference listening position at ear height using a C-weighted slow response sound pressure level meter (RMS, slow) compliant with IEC 61672 [17].

The deviation between the levels of any two channels should not exceed **1 dB SPL**. For any form of stereophony the close matching of the front speakers is especially important. They should be adjusted, so that the difference between any two of them is less than **0.5 dB SPL**.

In a 5.1 surround sound system, the **subwoofer** and the **LFE** (Low Frequency Effects channel) need a separate calibration. Subwoofer alignment is a delicate matter and outside the scope of this document. A future revision of the relevant document EBU Tech 3276 will include this procedure.

If a surround sound signal contains an LFE-signal, it should be reproduced with **+10 dB gain** relative

to the same limited frequency band in a main channel (“In-band” gain). To test this, two different signals should be chosen (one for the LFE, one for a main channel) that can faithfully be reproduced by the subwoofer and by a main loudspeaker respectively. The bandwidth should be at least one octave and the energy should be the same for both signals. An example is pink noise from 60 to 120 Hz for the LFE and pink noise from 200 to 400 Hz for a main channel. The LFE channel gain (NOT the subwoofer gain! It is assumed that the subwoofer has already been calibrated.) should then be adjusted so that the Sound Pressure Level of the respective test signal is 10 dB higher than the respective signal on the main loudspeaker.

To summarize:

$L_{LISTref} = 73 \text{ dB}_C \text{ SPL per main loudspeaker}$	(Using a 500-2000 Hz monophonic noise of equal energy per octave at a Programme Loudness Level of -23 LUFS)
--	---

This Reference Listening Level should be used for an average-size audio mixing room between 125 m^3 and 250 m^3 . The testing signal 500-2000 Hz pink noise is available for download at the EBU Technical website (<https://tech.ebu.ch/loudness>).

The average level of sound programmes according to EBU R 128 is typically *lower* (approximately 3 LU on average for TV programmes) than programmes levelled in the “old” QPPM-world. The above-mentioned procedure ensures a high signal to noise ratio and pleasant monitoring level even for high dynamic audio signals. **Nevertheless, raising the monitor gain further to compensate for extremely dynamic audio signals like, for example, classical music is an appropriate step to ensure familiar listening levels.** A potential deviation from the reference listening level depends on the room size as well as the main purpose of the room. For example, in a master control room, the listening level is anticipated to be considerably lower than in a production studio where often low-level details of a mix must be qualified. **In any case, it is important to consistently keep the listening level to establish an “inner loudness reference”.**

No satisfactory method for measuring sound pressure levels produced by headphones can be recommended. The level should be adjusted in such a way that a perceived loudness equal to a reference sound field produced by loudspeakers is achieved.

9. Genre Specific Issues

The concept of EBU R 128 centres on the loudness normalisation of each programme to one single Target Level (-23 LUFS). There are two reasons why this cannot be a perfect solution:

- No objective loudness measurement can ever be perfect
- There will always be individual preferences

Thus, a perfect solution is *generally not possible* as it differs from person to person. Within the scope of EBU R 128 it is vital to understand that it is not intended to achieve a loudness balance based on the real sound pressure level of a specific audio signal, but instead to provide a satisfactory listening experience for a diverse mix of genres for the majority of listeners.

This may result, for example, in a Schubert string quartet having the very same integrated loudness level as a Mahler symphony, namely -23 LUFS. While this does not reflect reality, it makes these items fit into a wide array of adjacent programming, and that is the intention of **advocating one single number.**

As this document shall serve as a pool of experiences, one might be tempted to consider refining this paradigm, once loudness normalisation becomes widespread. But most listeners do accept if the loudness level of programmes lies within a so-called ‘comfort zone’ of about 8 LU around the Target Level, whereas the distribution is asymmetric (for example, +3 LU/-5 LU). In cases, where

the objective loudness algorithm does not always provide a perfect result, the programme may still lie within this comfort zone. Broadcasters should also bear in mind that the audience can still adjust the loudness level with their remote control, to accommodate likes and dislikes.

The EBU generally encourages the **normalisation to one single Target Level** despite a potential refinement for many individual genres. Allowing too many variations (or even only a few) may challenge the system of equal average loudness from the outset. Naturally, the fear is that variations would be biased to the louder side. Exhibiting a Programme Loudness Level **deliberately lower** than the Target Level is a different topic that has already been touched in § [6.1.1](#) and will be examined further below in § [9.1](#) and § [9.3](#).

Nevertheless, there are cases, where a deviation from the general scheme and additional specific treatment may be appropriate. Specifically, three genres are now investigated: **commercials** (advertisements) and **trailers**, **feature films** (movies) and **music programmes**.

9.1 **Commercials (Advertisements) and Trailers**

This type of programme was arguably the most frequently mentioned one as far as **listener annoyance** is concerned, and thus was mainly responsible for the loudness problems encountered in the past and sometimes still today. In the UK (BCAP rules – Broadcast Committee of Advertising Practice) and the USA (CALM Act – Commercial Advertisement Loudness Mitigation) even *legislation* has been put into place to tame this genre. It is certainly vital that the system of loudness normalisation based on **EBU R 128** provides an effective toolset for this task – abuse shall be prevented. To control the dynamics of a commercial in a loudness normalised world where there exists the danger of suddenly too high loudness differences (overly loud ‘pay-off’ after a longer period of low-level signals just above the gate threshold), the parameter Loudness Range (LRA) is not suited, as the calculation is based on the short-term loudness values (3 s interval). Therefore, for short-form content (up to 2 minutes duration, typically <30 s) there are too few data points to derive a meaningful number for LRA. The Loudness Range parameter is not to blame for this fact, as it was never intended for this purpose.



Loudness Range does not provide meaningful results for programmes shorter than typically 30 seconds.

An alternative can be found in using the **Maximum Short-term Loudness Level (Max SL)**. The EBU PLOUD group has published a dedicated **supplement** to R 128 that deals with the usage of Max SL for short-form content. This document is based on the experiences of PLOUD members and the necessity to clarify and harmonise the approach for commercials, adverts, promos etc.

The loudness parameters for short-form content are:

Programme Loudness	-23.0 LUFS ±0.5 LU
Maximum True Peak Level	-1 dBTP
Maximum Short-term Loudness	-18.0 LUFS (+5.0 LU on the relative scale)
Loudness Range	– (not applicable)

EBU members are encouraged to use these parameters and especially the individual limit for Max SL for short-form content.



A limit for Maximum Short-term Loudness (Max SL) should be used to control short-form content. Recommended value: Max SL ≤ +5 LU (= -18 LUFS)

For special programmes of this genre that consist mainly of background or creatively intended low-level sounds, a loudness level lower than Target Level may be used. This is anticipated in the supplement. Programmes destined for playout at lower than Target Level need special attention to ensure they pass automatic normalisation processes unharmed. They should really be the exception, not the rule. Such programmes need to be clearly identified (for example, with a 'Low Loudness Flag'), so that they will not be normalised to Target Level by accident. Another solution is to only perform negative gain corrections if programmes are being delivered off-target. Thus, deliberately lower-loudness programmes will also pass unharmed. § 6.1.1 provides more details.

9.2 Feature Films (Movies)

Because of the potentially very dynamic content, feature films arguably present the **biggest challenge** to integrate seamlessly into loudness-balanced programming. In many cases, the original mix for the cinema is not suitable for a typical domestic listening environment. Usually, the difference between the Programme Loudness Level and the Maximum Momentary or Short-term Loudness Level is too big for this reproduction case. Furthermore, the Programme Loudness Level often differs significantly from the Voice Loudness Level, sometimes more than 10 LU! This situation can be found in a wide range of films, but these characteristics can sometimes also be encountered in television drama series. If content like that is pushed through a regular programme loudness normalisation process according to EBU R 128, this may result in a too low level of speech and in too loud sound effects. Using anchor-based loudness normalisation might improve the former (speech level) but will make the latter (loud sound effects) even more annoying. The issue becomes more noticeable when movies are interrupted by commercial blocks ('Ad break'): both the movie and the commercials are R-128-compliant, but the Voice Level difference causes significant perceived transition jumps (see Figure 6).

Therefore a special dynamic treatment seems appropriate, taking into account the Programme Loudness Level, the Voice Loudness Level, the Maximum Short-term Level (and/or Maximum Momentary Loudness Level) as well as Loudness Range.

It is important to notice that the level of speech alone is not the only variable that determines the perceived balance of loudness normalisation – a parameter that is at least equally relevant is **speech intelligibility**. Intelligibility is dependent on several factors, like clarity of pronunciation, choice of recording location, microphone technique etc. For feature films, speech intelligibility is usually high due to the specific production methods (Automatic Dialogue Replacement (ADR) if the location sound is too compromised).

The challenge is to adapt cinema dynamics in a sophisticated and intelligent way while fully preserving the intentions and workflow of EBU R 128. This challenge is even tougher when such an

adaptation needs to be performed automatically.

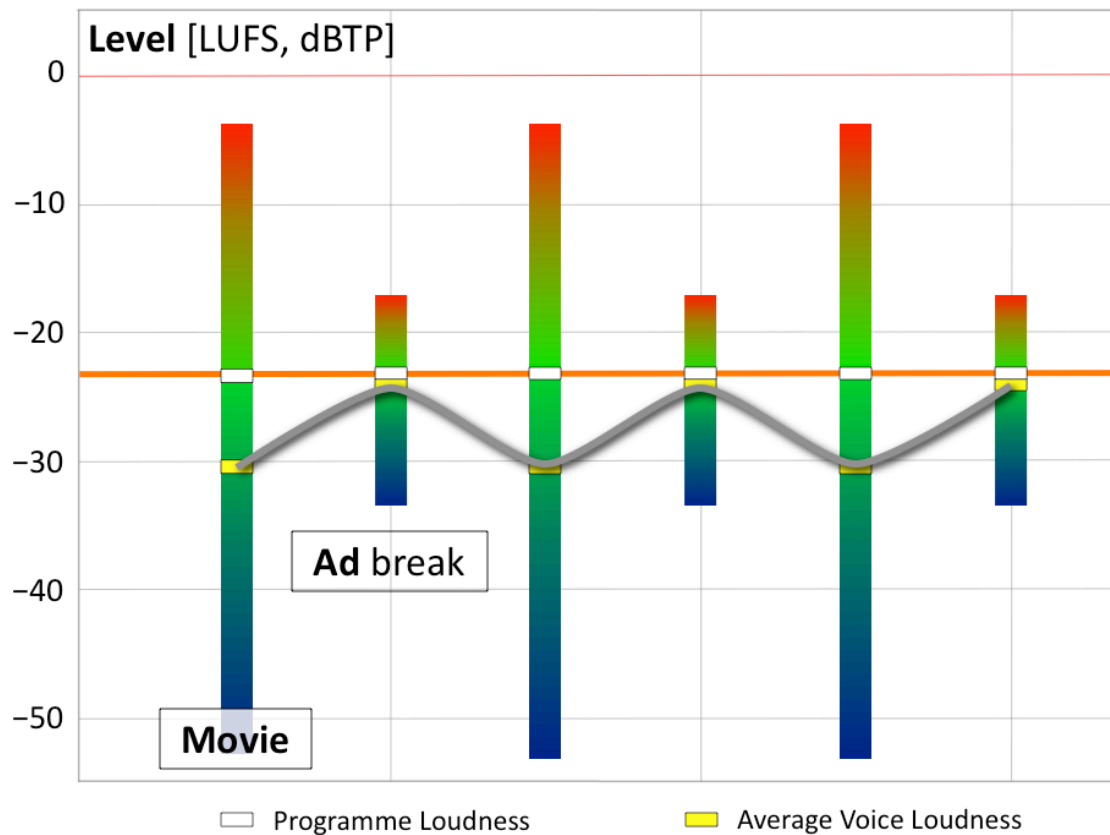


Figure 6: Jumps in Average Voice Levels between a movie and an Ad break

Summarising, the objectives for a controlled dynamic adjustment are:

- To preserve the quality and experience of the original sound track as much as possible
- To bring the 'average level' of speech closer to Target Level (it is assumed that speech intelligibility is high)
- To carefully adapt the dynamic properties of the mix to the reproduction environment
- To enable automating the process wherein analysis and processing strategy are combined

Analysis, extending the measurement

Next to the main measures according to EBU R 128 (Programme Loudness Level, Loudness Range and Maximum True Peak Level) the following measures may be performed:

- Average Voice Loudness (VL)
- Difference of Voice Loudness to Programme Loudness Level ($\Delta VL-PL$)
- Maximum Momentary Loudness (Max ML)
- Maximum Short-term Loudness (Max SL)

A difference larger than 4-5 LU between Programme Loudness and Voice Loudness is likely to be problematic. If furthermore the distance between the average Voice Loudness and the Maximum

Short-term Loudness (especially for 5.1 surround sound mixes) exceeds about 12-15 LU, this probably indicates a theatrical mix, which would benefit from dynamic processing prior to going through the R 128 loudness alignment. Also Loudness-Range values exceeding 20-23 LU (for 5.1 mixes) are a hint to a theatrical mix.

Processing

A file-based (offline) process may use a stepped and iterative approach wherein measurements are adaptively repeated. For real-time or 'live' processing, extensive analysis like mentioned above is not possible beforehand. Sophisticated methods have recently appeared (2015) to fine-tune the dynamics processing on the fly. It is important to notice that no single method can be the ultimate solution, it all depends on the framework wherein it is applied, its constraints, the individual preferences of the users/consumers and last, but not least, the content.

Anchor-based normalisation, Dynamic Range Control (DRC)

A broadcaster may decide to switch to anchor-based normalisation only for feature films, also if sophisticated processing like described above has been performed. Arguably, if Voice Loudness is very close or identical to Programme Loudness, it does not make a difference which normalisation method one uses. In such a case it is recommended to keep the Programme Loudness paradigm. If the difference between Voice Loudness and Programme Loudness was originally very high (for example > 10 LU), processing intended to heavily reduce $\Delta VL - PL$ may lead to unpleasant artefacts. Keeping $\Delta VL - PL$ still in the order of about 4 LU puts fewer burdens on the processing algorithm. Normalisation according to the Voice Loudness Level can be chosen, so that ad breaks (if existent; see Figure 6) do integrate more seamlessly. **The result will be a Programme Loudness Level higher than the Target Level of -23 LUFS!** Another solution is to still normalise to Programme Loudness but provide Loudness Metadata that corresponds to Voice Loudness. This is only an option for this specific genre (and similarly dynamic content)!

VL-based normalisation for movies (or Metadata corresponding to VL) in some cases has the consequence of a more symmetrical processing regarding automatic Dynamic Range Control in the Dolby AC-3 system at the consumer's side. The DRC process centres at the **dialnorm** value. If the dialnorm value corresponds to Voice Loudness in bespoke treated feature films, the dynamic processing of speech will be symmetrical for the moderate presets *Film Light* and *Music Light*. If the dialnorm value corresponds to Programme Loudness, the Voice Level might lie at the lower edge of the unity gain part of the compressor profile. This may result in frequent on-off dynamic processing of speech if the variation of such a speech signal is considerable. For the more severe DRC profiles *Film Standard* (the default profile), *Music Standard* and *Speech* dialnorm corresponds to the lower edge of the unity part anyway. Dynamic processing will consequently be often compromised, regardless if the dialnorm value indicates Voice Loudness or Programme Loudness (see Appendix 4 for Dolby DRC presets (§ [11.4](#))).

Example for adapting the dynamics of feature films for TV

The following parameter values show an example (this is no rule!) how automatic adaptation of feature films and similar content for broadcasting may be done:

Maximum Difference PL-VL	4-5 LU
Maximum Momentary Loudness	11-14 LU above PL
Maximum Short-term Loudness	7-10 LU above PL
Normalisation Paradigm	Anchor-based (for example, Voice Level); PL will be higher than -23 LUFS!

As an alternative to Max ML and Max SL, the parameter Loudness Range (LRA) can be used to quantify the potential dynamic processing. Bringing the difference between PL and VL within 3-5 LU may still be applicable, just as anchor-based normalisation using VL.

Manual control

A broadcaster can choose to let **experienced sound engineers** perform the adaptation of feature films for TV. This is arguably the most sophisticated method, as the result is immediately checked by ear. Many, if not all, of the relevant parameters mentioned above may assist the engineer in her task, so that the result is appropriate for the home listener. 'Manual' compression through dynamically adjusting the main fader as well as adjustment of the centre channel (in a 5.1 mix) to bring the Voice Level closer to the Programme Loudness Level are but two ways to achieve this.

9.3 Music

The experience of passionate music listeners suggests that certain programmes that contain mostly music, either with a wide Loudness Range like classical music or with a higher degree of dynamic compression as an artistic property like a rock concert, have the tendency to be listened to with a higher loudness level (up to +2-3 LU on average) than other genres. Reasons for that might be the significantly high potential sound pressure level in reality (fortissimo of a symphony orchestra, rock band with powerful public address system) and the fact that for music there do not exist 'foreground sounds' vs. 'background sounds' – everything is in the foreground.

But as mentioned above, a potential differentiation of the Target Level for these programmes may cause more harm in opening a backdoor to being again louder than the rest instead of improving the situation significantly. Based on the same reasoning as for commercials, advertisements and trailers, normalisation to a higher Target Level is **discouraged**. The audience can still use their remote control to adjust (increase) the loudness level in their reproduction environment to their taste. Following programmes like commercials or trailers will consequently be perceived louder too. It is anticipated that this should not push those programmes out of the comfort zone.

A **different issue** is the case of music programmes with a deliberately lower Loudness Level. This might be, for example, the slow movement of a classical symphony or a ballad on a music album. Such a programme is part of a more extended sequence of programmes, and when presented within this sequence, the individual relationship between the 'elementary' programmes should be preserved. If such a programme is played out of its sequence context, normalisation to the Target Level is performed as usual.

These two scenarios necessitate additional metadata to accomplish the dual-fold use. The original Loudness Metadata Information should store the actual Programme Loudness Level, and the additional metadata should indicate the deliberately lower/different reproduction level ('Loudness Offset', 'Low Loudness Flag') (see also § 6.1.1). This situation is very common for music playback on **mobile devices** like Personal Music Players ('Track normalisation' vs. 'Album normalisation') and will also be of relevance for radio programmes consisting of many individual items (which can be considered as programmes or programme elements). Those items might have a significantly different Loudness Level on purpose, but they will increasingly be played back automatically without intervention by a radio moderator (operating in Radio-DJ mode) or a sound engineer.



Programmes deviating from Target Level need additional metadata (Loudness Offset, Low Loudness Flag) for different normalisation cases.

10. Transition strategy

It is clear that such a fundamental change in the way audio signals are measured, metered and treated, and that affects all stages of audio production, distribution, archiving and transmission, is not done overnight with the flick of a switch. Every broadcaster and audio facility must find its

individual way to perform this change, to install the appropriate equipment, train the staff and get on the road to loudness nirvana! Nevertheless, a few constants can be stated that will be applicable for everyone:

- Establish an **internal loudness group** to discuss basic implications and a strategy to convince management, programme makers and your colleagues.
- **Start now** – don't wait until everything is in place and all the others have done it; there is no need to be perfect in the very beginning.
- Before you can do anything, management has to agree to this change and all its consequences. Get a **written agreement** or 'call for action' from the general director.
- Provide **loudness meters** to your key production personnel. Let them gain first experiences and learn the advantages and liberations of the loudness paradigm so they can be opinion leaders for their colleagues.
- **Survey the market** regarding loudness metering and loudness management to determine what is best suited for your environment.
- Determine the **key areas** in which loudness work should start. Potential candidates are: production studios, post-production suites, OB vans, QC (Quality Control) department and MCR (Master Control Room).
- Be aware that you will encounter obstacles ("It has always been that way", "It has never been that way", "Who are you to say we should do it that way"). **Patience** and demonstrating practical examples will pay off. Become your facility's Zen-master of loudness normalisation ("restraint – simplicity – naturalness").
- **Allow everybody time to adapt**. Although the audience has been waiting for a solution for decades, don't create more problems by trying to do too much too quickly.
- Solutions for **file-based workflows** are increasingly established (November 2015). Keep an eye on the market and demand solutions from vendors.
- Use this fundamental change as an opportunity for a general **discussion** about **audio quality** and the development of a '**corporate sound**', which includes, for example, speech intelligibility, the balance of speech vs. music and, of course, loudness normalisation of programmes.
- **Use and trust your ears!** They are the best loudness meters.

11. Appendices

11.1 Appendix 1: ITU-R BS.1770

The basis of EBU R 128 and thus EBU Tech Doc 3343 is ITU-R BS.1770, the result of extensive work by the International Telecommunication Union. The purpose of this standard was to establish an agreed open algorithm for the measurement of electrical loudness and the true peak levels of audio signals. It is a robust standard, which has the benefit of a simple implementation. In brief, it defines a “K-weighting” filter curve (a modified second-order high-pass filter, see Figure 7), which forms the basis for matching an inherently subjective impression with an objective measurement.

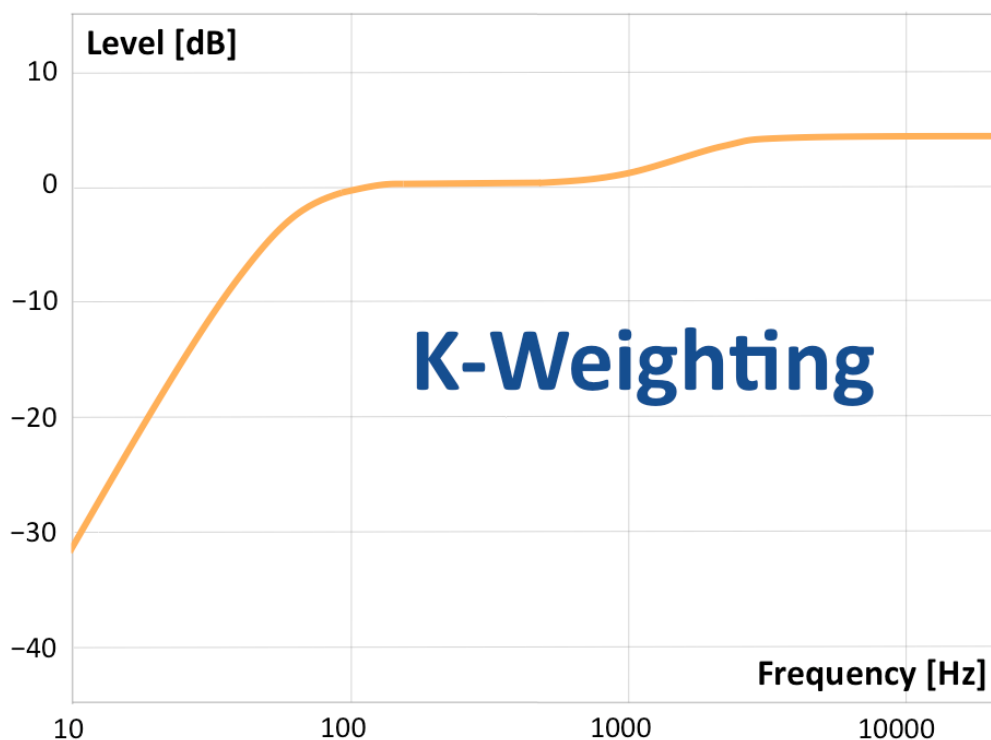


Figure 7: “K-Weighting” filter curve for loudness measurement

This weighting curve is applied to all channels except the Low-Frequency Effects Channel (LFE), which is currently discarded from the measurement (see below). The total mean square level is calculated (with different gain factors for the front and surround channels (see Figure 8) and the result is displayed as “LKFS” (Loudness, K-Weighting, referenced to digital Full Scale), or “LUFS”¹ (Loudness Units, referenced to digital Full Scale). For relative measurements, Loudness Units (LU) is used, where 1 LU is equivalent to 1 dB.

Low Frequency Effects (LFE) channel

The Low Frequency Effects channel (the “.1”-channel in “5.1”) of a multichannel audio signal is currently not taken into account for the loudness measurement according to ITU-R BS.1770. This may lead to abuse of the LFE with unnecessary high signal levels. Ongoing investigations try to evaluate the subjective effect the LFE has on the perception of loudness as well as the appropriate way to include it in the objective loudness measurement.

¹ The EBU recommends the use of ‘LUFS’ (as specified in EBU Tech Doc 3341). ‘LUFS’ is equivalent to ‘LKFS’ and overcomes an inconsistency between ITU-R BS.1770 and ITU-R BS.1771. ‘LUFS’ also complies with the international naming standard ISO 80000-8 [\[18\]](#).

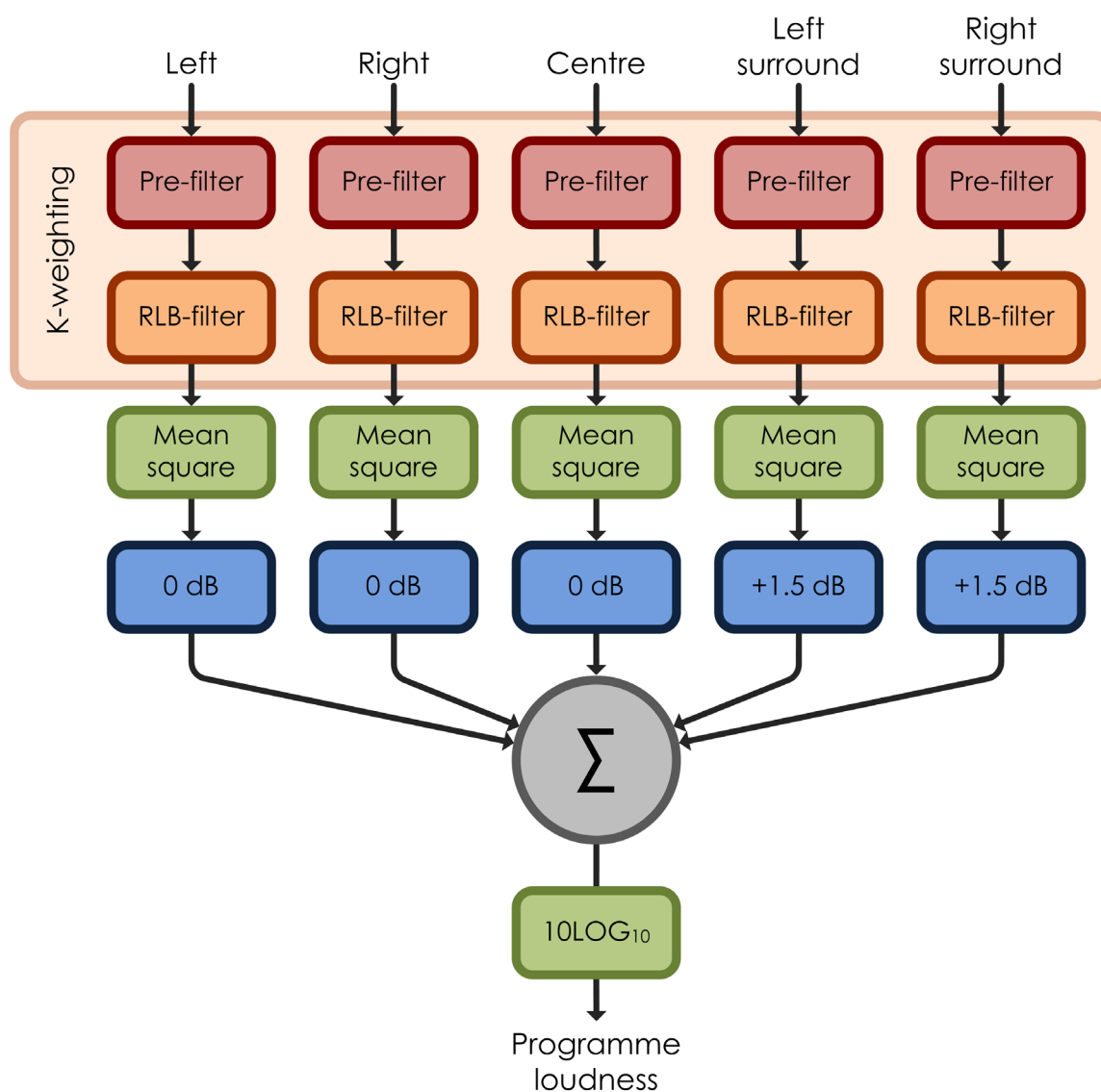


Figure 8: Channel processing and summation in ITU-R BS.1770

11.1.1 Gating

The **gating** method developed by the EBU PLOUD-group is now part of ITU-R BS.1770 (since revision 2), the **worldwide standard loudness algorithm**. This refinement method was introduced to improve the loudness matching of programmes with large dynamics and/or which contain longer periods of silence or isolated utterances. The gate serves to pause the loudness measurement when the signal drops below a certain threshold. Without this gating function, programmes with longer periods of silence, low-level background sounds or noise will get too low an integrated loudness reading. Such programmes would subsequently be too loud after normalisation.

Gating is only applied to the **Integrated Loudness Measurement** ('Programme Loudness' in R 128) and consists of the following elements:

1. An absolute 'silence' gating threshold at -70 LUFS for the computation of the absolute-gated loudness level;
2. A relative gating threshold, 10 LU below the absolute-gated loudness level;
3. The measurement input to which the gating threshold is applied is the loudness of the 400 ms blocks ('Momentary Loudness') with a constant overlap between consecutive gating blocks of 75% .

If the end of an integrated loudness measurement lies within a gating block, the incomplete gating block shall be discarded.

Note: The gating function excludes from the measurement those blocks of audio that are below a threshold. For the relative-threshold based gating function this requires the computation stages described above, as the threshold to be used is itself based on a measurement of loudness. In a live meter the integrated loudness has to be recalculated from the preceding (stored) loudness levels of the blocks from the time the measurement was started, by recalculating the threshold, then applying it to the stored values, every time the meter reading is updated.

Figure 9 shows how the relative gating function works: the green line shows the loudness measurement without gating ($L_K = -26$ LUFS); the red line is the gating threshold (in this case at -36 LUFS) which lies 10 LU below the ungated Programme Loudness Level; the loudness levels below the gating threshold are discarded (loudness blocks with blue border only) – and then the remaining loudness levels are averaged, giving the gated result (-25.2 LUFS).

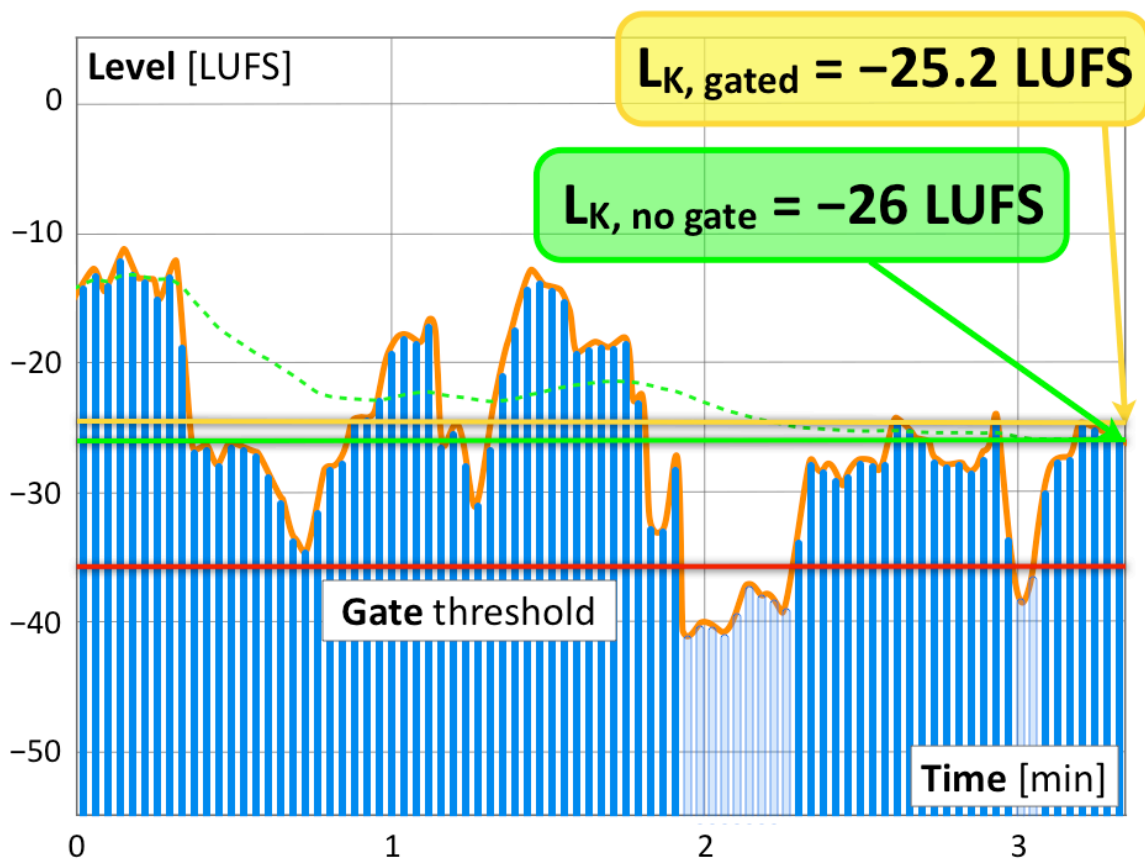


Figure 9: Explanation of the relative gating measurement

11.2 Appendix 2: EBU R 128

EBU R 128 establishes a predictable and well-defined method to measure the loudness level for news, sports, advertisements, drama, music, promotions, film etc. throughout the broadcast chain and thereby helps professionals to create robust specifications for ingest, production, play-out and distribution to a multitude of platforms. R 128 is based entirely on open standards and aims to harmonise the way we produce and measure audio internationally.

Whereas ITU-R BS.1770 defines the measurement method, R 128 extends it by actually defining a specific 'Target Level' for loudness normalisation, introduces the parameter 'Loudness Range' (LRA) and provides a limit for the 'Maximum Permitted True Peak Level'. The EBU's development was needed to accommodate the needs of programme makers, with particular regard to having a means to measure complete mixes (rather than just one component, such as speech or music) and the loudness range of the programme.

To repeat, EBU R 128 specifies three parameters:

- Programme Loudness
- Loudness Range
- Maximum Permitted True Peak Level

11.2.1 Programme Loudness

Programme Loudness describes the long-term **integrated** loudness over the duration of a programme¹. The parameter consists of one number (in LUFS, with one number after the decimal point), which indicates “how loud the programme is on average”. This is measured with a meter compliant with ITU-R BS.1770 (which includes the **gating** function).

The **Target Level** to which an audio signal will be normalised is:

-23.0 LUFS (± 0.5 LU)

In order to avoid rejection of programmes due to accumulation of metering tolerances, a **general tolerance of ± 0.5 LU around the Target Level of -23 LUFS is acceptable**. A deviation of ± 1.0 LU is acceptable for programmes where an exact normalisation to the Target Level of **-23.0 LUFS** is not achievable practically (such as live programmes or ones which have an exceedingly short turn-around). There are also cases where the Programme Loudness Level may deliberately lie outside the tolerance window (for example, if there are no foreground sounds or if the programme is part of a dedicated sequence like a music album). Programmes with a deliberately lower/different Target Level need special attention (see § [6.1.1](#), § [9.1](#) and § [9.3](#) for more details).

11.2.2 Loudness Range

Another major topic was the loudness range, which would be needed to accommodate the majority of programmes (provided that they don't exceed the tolerable loudness range for domestic listening). The **Loudness Range (LRA)** parameter quantifies (in LU) the variation of the loudness measurement of a programme. It is based on the statistical distribution of the Short-term (3s) loudness levels within a programme, thereby excluding the extremes (the low 10% and the high 5% of the distribution after applying a relative gate of -20 LU; an absolute gate at -70 LUFS is applied before the relative one). Therefore, for example, a single gunshot is not able to bias the outcome of the LRA computation. EBU Recommendation R 128 does not specify a maximum permitted LRA, as it is dependent on factors such as the tolerance window of the average listener to the station, the distribution of genres of the station etc. R 128 does, however, **encourage the use of LRA** to determine if dynamic treatment of an audio signal is needed and to match the signal with the requirements of a particular transmission channel or platform. More details about LRA may be found in EBU Tech Doc 3342.

Figure 10 shows the loudness distribution and LRA of the movie 'The Matrix'; 25 LU is probably challenging for most living rooms...

¹ The term 'a programme' is also used to mean an advertisement, a promotional item etc. For clarity, the advertisements etc. which are placed around and within the running time of what is generally considered to be 'a programme' are treated as programmes in their own right (also individual advertisements within a block are separate programmes); their integration with the longer programmes is thus made easier. Evidently, the makers of either type of programme can have no knowledge of what will be placed with it and so each type has to be considered separately. In this document, the term 'programme' refers to the programme as completed by Production and not the combination of the programme, interstitials, and advertisements that arrives at the viewer's or listener's receiver within the overall running time of the programme.

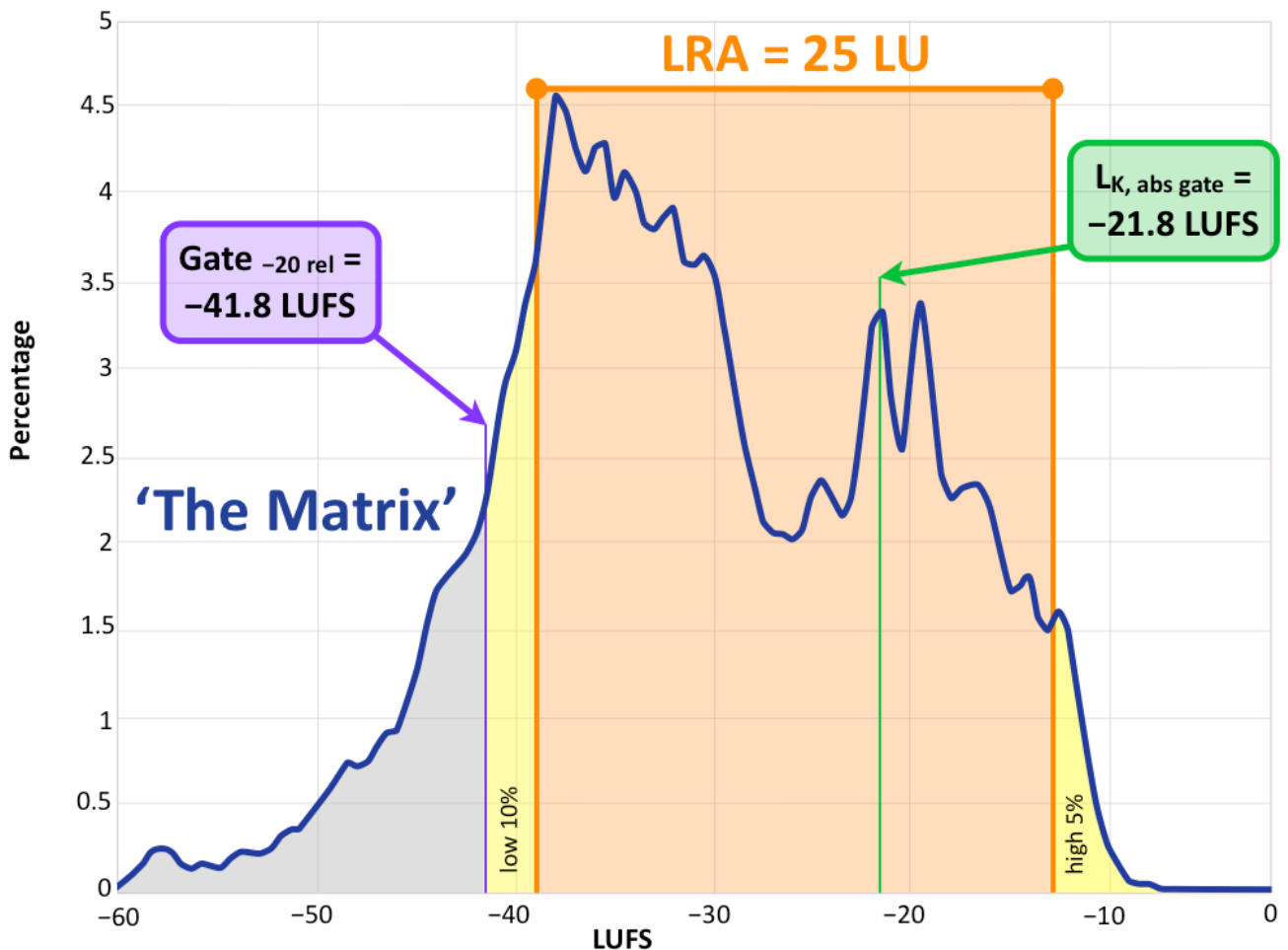


Figure 10: Loudness Range (LRA) as a result of the statistical distribution of loudness levels

For short-form programmes (typically <30 s) such as commercials, advertisements or trailers, there are too few data points to derive a meaningful result for LRA, as the calculation is based on the Short-term loudness levels (3 s window). A maximum or minimum number for Loudness Range shall therefore not be specified for such content. Setting a limit for the **maximum** permitted value of the **Short-term Loudness Level**¹ provides a better way to control the dynamic properties of short-form programmes (see § 9.1).

11.2.3 True Peak Level (TPL), Maximum Permitted TPL

In Europe, the most widespread metering device was (and in some countries, it still is) the Quasi Peak Programme Meter (QPPM; integration time = 10 ms). With the transition to digital signal processing, **sample peak** meters appeared. While a QPPM cannot display short peaks (<<10 ms) by design, also a sample peak measurement may not reveal the actual peak level represented by a digital signal either.

Digital processing or lossy coding can cause **inter-sample** peaks that exceed the indicated sample level. In broadcasting it is important to have a reliable indication of peak level across platforms and across sample rates. This meter should indicate clipping, especially if the peak lies in between samples, so that the **distortion** that can happen in subsequent Digital-to-Analogue converters, sample rate converters or commonly used codecs can be **predicted** and **avoided**. A sample peak meter cannot do that and is therefore insufficient for use in modern broadcasting [19].

¹ **Maximum Short-term Loudness Level (Max SL)** is the highest value (in LUFS) of an audio signal's Short-term Loudness Level (integration time 3 s).

The **true-peak level** indicates the maximum (positive or negative) value of the signal waveform in the continuous time domain; this value may be higher than the largest sample value in the time-sampled domain. With an oversampling true-peak meter compliant with ITU-R BS.1770, those true peaks (unit symbol according to ITU-R BS.1770: dBTP – deciBel referenced to digital Full Scale measured with a True-Peak meter) can now be detected. The accuracy depends on the oversampling frequency.



Oversampling Peak Meters provide a good estimate for the true peak of an audio signal. Sample Peak Meters don't.

It is only necessary to leave a headroom of 1 dB below 0 dBFS to still accommodate the potential under-read of about 0.5 dB (for a 4x oversampling true-peak meter; basic sample rate: 48 kHz).

The **Maximum Permitted True Peak Level** in production recommended in R 128 is consequently:

-1 dBTP

This is applicable to the production environment for generic linear audio signals (linear PCM, WAV). Note that some parts of the chain, such as analogue re-broadcasters and users of commonly used data reduction codecs require a lower True Peak Level. EBU Tech Doc 3344 contains comprehensive coverage of the topic. For two in Europe commonly used data reduction systems (MPEG1 Layer2 and Dolby AC-3) the recommended Maximum Permitted True Peak Level is:

-2 dBTP

This generally ensures that those codecs have appropriate headroom to perform data reduction without additional distortion. The actual Maximum True Peak Level chosen is dependent on the data reduction ratio.

Summary of EBU R 128

- The parameters 'Programme Loudness', 'Loudness Range' and 'Maximum True Peak Level' characterise an audio signal;
- The Programme Loudness Level shall be normalised to -23.0 LUFS ± 0.5 LU;
- The tolerance is ± 1.0 LU for programmes where normalisation within the general tolerance is not achievable practically (for example, live programmes);
- The measurement shall be done with a meter compliant with ITU-R BS.1770 and EBU Tech Doc 3341;
- The parameter **Loudness Range** can be used to help decide if dynamic compression is needed (dependent on genre, target audience and transmission platform);
- The **Maximum Permitted True Peak Level** in production (linear audio, PCM) is **-1 dBTP**;
- **Loudness Metadata** shall be set to indicate **-23 LUFS** (for programmes that have been normalised to that level, as is recommended); loudness Metadata shall always indicate the **correct value for Programme Loudness** even if for any reason a programme may not be normalised to -23 LUFS.

11.2.4 R 128 Logo

The EBU has introduced an official logo for R 128, comprised of the numbers 1, 2 and 8 – forming a happy, smiling face:



Manufacturers may use the logo (with certain prerequisites) to indicate compliance with ‘EBU Mode’.

11.3 Appendix 3: Loudness Metering with ‘EBU Mode’

An ‘EBU Mode’ loudness meter as defined in EBU Tech Doc 3341 offers 3 distinct time scales:

- **Momentary Loudness** (abbreviated “M”) – time window: 400 ms
- **Short-term Loudness** (abbreviated “S”) – time window: 3 s
- **Integrated Loudness** (abbreviated “I”) – from ‘start’ to ‘stop’


The M and S time windows¹ are intended to be used for the immediate levelling and mixing of audio signals. If he/she wants to, a mixer has to know at any time how loud the actual signal is, and that is the main purpose of the Momentary and Short-term measurement.

Due to an inconsistency between ITU-R BS.1770 and ITU-R BS.1771 [20], EBU Tech Doc 3341 suggests a different naming convention, complying with ISO 80000-8:

- The symbol for ‘Loudness Level, K-weighted’ should be ‘ L_K ’.
- The unit symbol ‘LUFS’ indicates the value of L_K with reference to digital full scale.
- The unit symbol ‘LU’ indicates L_K without a direct absolute reference and thus also describes loudness level differences.

Any graphical or user-interface details of a loudness meter complying with ‘EBU Mode’ have deliberately not been specified; nevertheless, two scales have been defined: “EBU +9 Scale” which ought to be suitable for most programmes and “EBU +18 Scale” which may be needed for programmes with a wide LRA. Both scales can either display the relative Loudness Level in LU, or the absolute one in LUFS. ‘0 LU’ in ‘EBU mode’ equals the Target Level of –23 LUFS. The meter manufacturers in the PLOUD Group have agreed to implement the ‘EBU Mode’ set of parameters to make sure their meters’ readings will be aligned.

¹ ‘M’ and ‘S’ are commonly used in stereophony for ‘Mid’ and ‘Side’. To distinguish the integration times ‘Momentary’ and ‘Short-term’, the versions ‘ ML_K ’ and ‘ SL_K ’ (as well as ‘ IL_K ’) may be used. ‘ L_K ’ stands for ‘Loudness Level, K-weighted’, and complies with the international naming standard ISO 80000-8.

 In an 'EBU Mode' loudness meter, 0 LU equals -23 LUFS.

Many more manufacturers have adopted 'EBU Mode' too, or are in the process of doing so. On a related note, the latest revision of ITU-R BS.1771 has also standardised Momentary and Short-term Loudness as well as the two metering scales (there is a difference in the calculation of Momentary Loudness, though). The essence of 'EBU Mode' is thus part of the international requirements for loudness metering.

11.4 Appendix 4: DRC (Dynamic Range Control) Presets for Dolby Digital

In the Dolby Digital system (AC-3), there are 6 compression presets that cause the encoder to generate different gain control words that are sent in the bitstream to the consumer's decoder: *Film Standard*, *Film Light*, *Music Standard*, *Music Light*, *Speech* and *None*. These presets result in more or less compression centred at the *dialnorm* value, one more reason to set this Metadata parameter correctly (see Figure 11 for the compression curves around -23 LUFS).

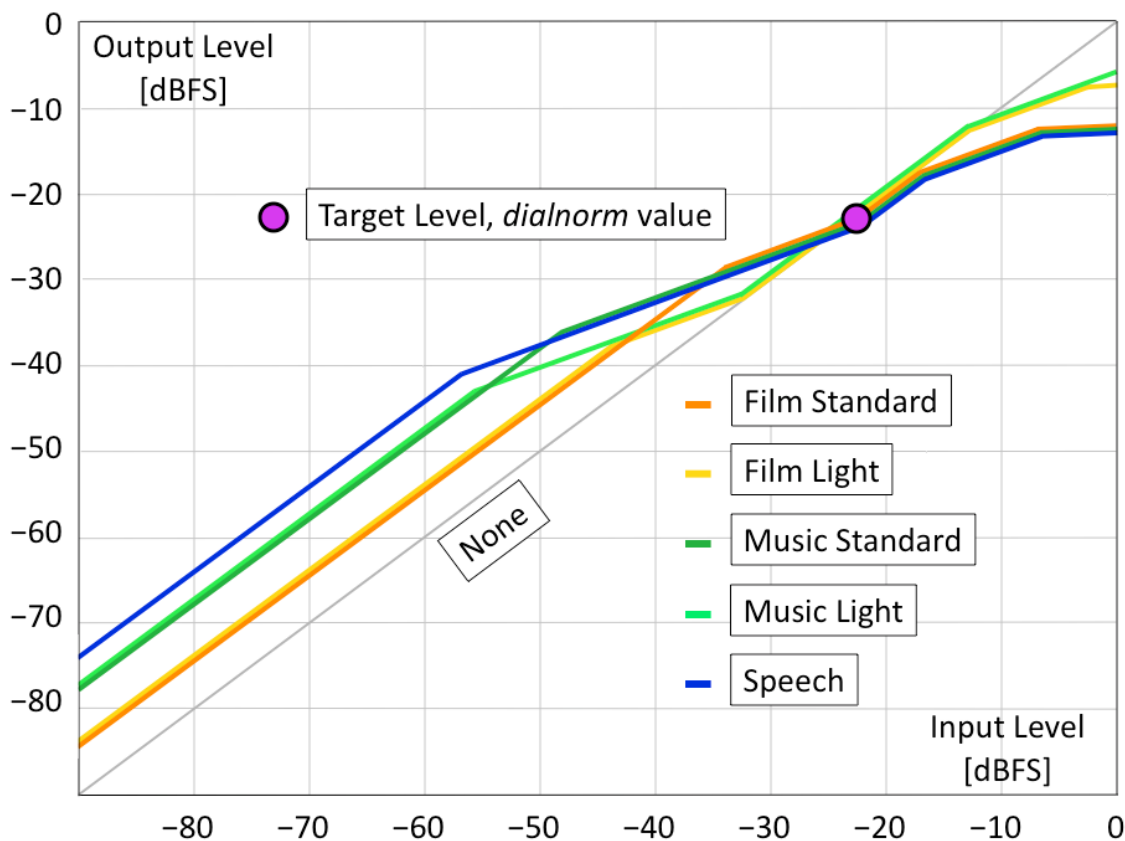


Figure 11: Dynamic Range Compression curves of the Dolby AC-3 system

Two compression profiles exist within Dolby Digital: '*Line mode*' and '*RF mode*'. For each, a separate compression preset can be chosen. See EBU Tech 3344 for more details.

Within the system of R 128 and its concept of normalising the audio signal to -23 LUFS (resulting in static loudness metadata which equals a fixed *dialnorm* value) as well as using the parameter Loudness Range to determine any potential processing, the preset '*None*' may be used. This may be applicable in particular for '*Line mode*' and also by default in '*RF mode*'.

12. References

- [1] [EBU Technical Recommendation R 128](#) 'Loudness Normalisation and Permitted Maximum Level of audio signals' (2014)
- [2] ITU-R BS.1770 'Algorithms to measure audio programme loudness and true-peak audio level' (2015)
- [3] [EBU Technical Recommendation R 128 s1](#) 'Loudness Parameters for Short-form Content (adverts, promos etc.)' (2016), Supplement 1 to EBU R 128
- [4] [EBU Tech Doc 3341](#) 'Loudness Metering: 'EBU Mode' Metering to supplement Loudness Normalisation in accordance with EBU R 128' (2016)
- [5] [EBU Tech Doc 3342](#) 'Loudness Range: A Descriptor to supplement Loudness Normalisation in accordance with EBU R 128' (2016)
- [6] [EBU Tech Doc 3344](#) 'Guidelines for Distribution and Reproduction in accordance with EBU R 128' (2011)
- [7] Skovenborg E. & Lund Th., 'Level-normalization of Feature Films using Loudness vs. Speech' in Proceedings of the 135th AES Convention, October 2013
- [8] [EBU Technical Recommendation R 85](#): 'Use of the Broadcast Wave Format for the Exchange of Audio Data Files' (2004)
- [9] [EBU Technical Recommendation R 111](#): 'Multi-channel Use of the BWF Audio File Format (MBWF)' (2007)
- [10] [EBU Tech Doc 3306](#): 'MBWF/RF64: An extended File Format for Audio' (2009)
- [11] [EBU Tech Doc 3364](#): 'Audio Definition Model – Metadata Specification' (2014)
- [12] ITU-R BS.775-2 'Multichannel stereophonic sound system with and without accompanying picture' (2006)
- [13] [EBU Technical Recommendation R 68](#): 'Alignment level in digital audio production equipment and in digital audio recorders' (revision 2000)
- [14] ITU-R BS.645-2 'Test signals and metering to be used on international sound programme connections' (1992)
- [15] [EBU Tech Doc 3276-E](#) (+ [supplement 1](#)) 'Listening conditions for the assessment of sound programme material' (1998, 2004 – supplement 1)
- [16] IEC 61260-1 'Electroacoustics – Octave-band and fractional-octave-band filters – Part 1: Specifications' (2014)
- [17] IEC 61672-1 'Electroacoustics – Sound level meters – Part 1: Specifications' (2013)
- [18] ISO 80000-8: 'Quantities and Units – Part 8: Acoustics'
- [19] Lund, Th. 'Stop counting samples', AES paper N° 6972, 121st AES Convention, October 2006
- [20] ITU-R BS.1771 'Requirements for loudness and true-peak indicating meters' (2012)