EBU
OPERATING EUROVISION AND EURORADIO

TECH 3343

GUIDELINES FOR PRODUCTION OF PROGRAMMES IN ACCORDANCE WITH EBU R 128

Geneva
November 2023

Document History

EBU Committee	TC	
Drafting Group	PLOUD	
First published	February 2011	
Revised	August 2011	V2; relative gate changed from −8 to −10 LU
	January 2016	V3; major re-write, several modifications due to extensive experience of EBU members
	November 2023	V4; substantial re-write, some deletions, § 9 almost new (genre- and platform specific issues); incorporation of the latest R 128 supplement 4 ('Loudness normalisation of Cinematic Content') as well as s2 and s3 ('Loudness in Streaming' & 'Loudness in Radio')

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

Acknowledgement

EBU technical publications are the work of experts from EBU Members, Associate Members and third parties consisting of international standards bodies, industry partners, academic institutions and independent consultants.

Their contribution to EBU technical publications is a very generous act by the individuals concerned and by their employers. The EBU appreciates their efforts and thanks them most sincerely.

This document has been produced with the assistance of the following entities:

EBU Members: Florian Camerer (ORF), Christian Hartmann (IRT), Askan Siegfried (NDR), Andrew Mason (BBC R&D)

Third parties: Richard van Everdingen (Delta-Sigma-Consultancy), Thomas Lund (Genelec), Michael Kratschmer (Fraunhofer IIS), Scott Norcross (Dolby),

EBU Project Manager: Paola Sunna

Contents

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

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[1] 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.

Loudness normalisation is a true audio levelling revolution!

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**)**.

[1] '*Levelling*' is used in this document as a generic term in the sense of 'choosing a level'. It is not meant as a dynamics process.

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

- **EBU R 128 s1** *'Loudness Parameters for Short-form Content (adverts, promos etc.)'*; supplement 1 to EBU R 128 [3]
- **EBU R 128 s2** *'Loudness in Streaming'*; supplement 2 to EBU R 128 [4]
- **EBU R 128 s3** *'Loudness in Radio'*; supplement 3 to EBU R 128 [5]
- **EBU R 128 s4** *'Loudness Normalisation of Cinematic Content'*; supplement 4 to EBU R 128 [6]
- **EBU Tech Doc 3341** *'Loudness Metering: 'EBU Mode' metering to supplement loudness normalisation in accordance with EBU R 128'* [7]
- **EBU Tech Doc 3342** *'Loudness Range: A descriptor to supplement loudness normalisation in accordance with EBU R 128'* [8]
- **EBU Tech Doc 3343** *'Guidelines for Production of Programmes in accordance with EBU R 128'* (this document)
- **EBU Tech Doc 3344** *'Guidelines for Distribution and Reproduction in accordance with EBU R 128'* [9] and
- **EBU Tech Doc 3401** *'Guidelines for Radio production and distribution in accordance with EBU R 128'* [10]

EBU R 128 and ITU-R BS.1770 are the basis. Nine more EBU loudness publications provide details.

The four supplements provide guidance on specific use cases and programme types. They will be introduced in the appropriate sections. For supplement 3, an accompanying Tech Doc offers guidelines for Radio production and distribution, going into more detail and also covering use cases of EBU members.

The Technical Documents about '*Loudness Metering*' and 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*' cover all aspects of loudness normalisation for the distribution of audio signals and address 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 at **archival** issues (normalisation options).

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 (*supplement 1*) as well as music/radio programmes (*supplement 3*) and movies/cinematic content (*supplement 4*) will be addressed in a dedicated chapter (§ 9). Also, Loudness in Streaming (*supplement 2*) will briefly be covered for completeness' sake, despite belonging to the distribution realm.

These Guidelines are meant to be a '**living document**', where, over time, experiences of broadcasters will find their 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 anymore.

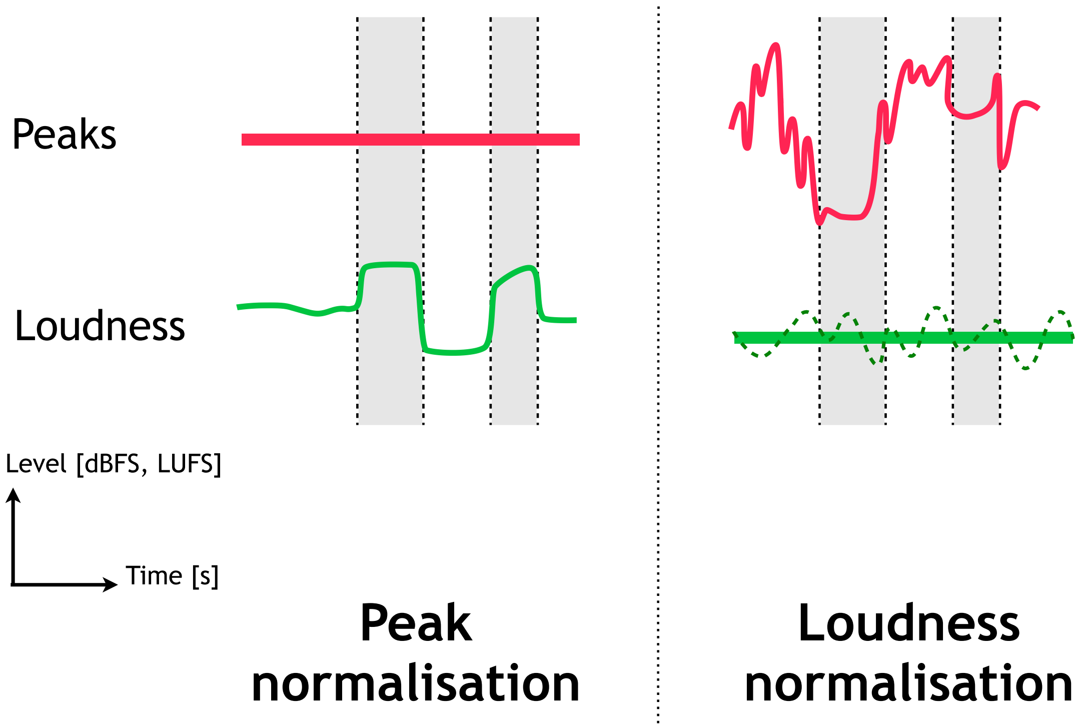


Figure 1: Peak normalisation vs. Loudness normalisation of a series of programmes

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) all have 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 individual ("agile") **Loudness Metadata** (see § 6) which reflects the actual average programme loudness. 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 source 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 is recommended in production 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

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

Having one number has great strength in spreading the loudness concept, as it is easy to understand and act upon. And the active normalisation of the source 'punishes' overly 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**).

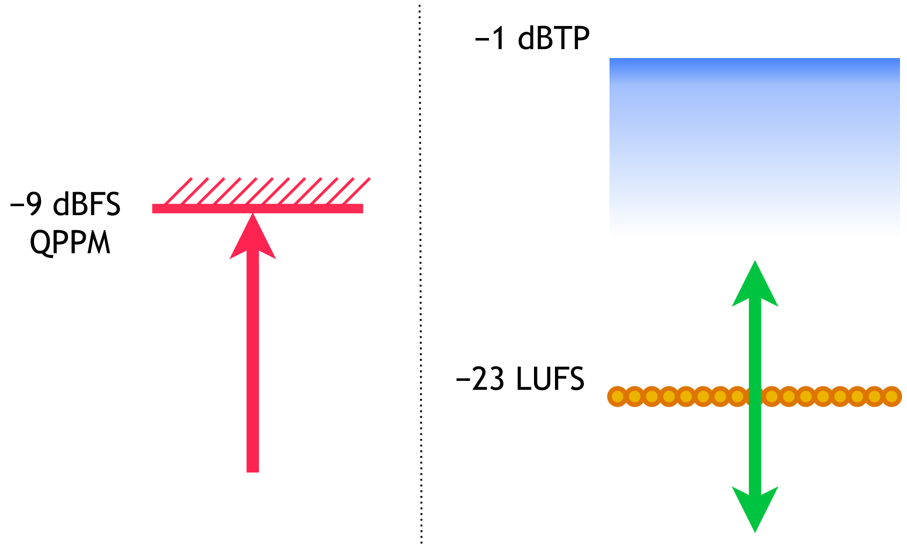


Figure 2: Legacy Peak normalisation vs. Loudness normalisation

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 a loudness competition started, fuelled by ever more sophisticated dynamics processors.

Loudness levelling, on the other hand, encourages the use of by far the best metering device: **the ears**. This implies more alert mixing and fosters audio quality. Extensive experience of 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 to mix 'only by ear' - after setting levels and a fixed monitor gain.

2.4 Loudness processors

Downstream of production, the broadcaster may still be confronted with the need to normalise diverse content originating from different places. Especially during the early phase of transitioning to the loudness paradigm there will be programmes that are not yet loudness normalised. Strategies for these programmes have to be developed, such as **automated normalisation** directly after ingesting to a playout server or the installation of a safety loudness regulation device (Loudness Processor) at the output of Master Control. 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 should 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

Before undertaking the first loudness-based mix, it is highly advantageous to **measure past programmes of the same genre** and generally mixes from the past **with a loudness meter** to determine the average programme loudness level in the peak normalisation world. Typically, the result will be somewhat higher than the Loudness Target Level of −23 LUFS (the average loudness level for European public broadcasters in the peak days was approximately −20 LUFS). Such a difference should be followed by a complementary change in the listening level (see § 8.2), so that the average acoustic level while mixing stays the same! For example, if one typically mixed around −20 LUFS, then the listening level should be *increased* by 3 dB. 'Tricking' oneself with an adapted listening level to reach the new Loudness Reference Level is the best and fastest way to success!

Adapting the listening level is the best 'trick' to quickly reach the Loudness Target Level of −23 LUFS!

It is **highly recommended** to immediately start using a **loudness meter** and mix towards the new reference. 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 boxes on page 13).

In **post-production** any deviation from the Target Loudness Level after a programme mix is finished is easy to correct. Measuring the whole programme, the necessary static gain offset can be determined exactly, and in today's file-based world a gain calculation is a very quick and easy operation. Typically, offline loudness meters perform both the measurement as well as the correction. 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 ±0.2 LU around the Target Level of −23 LUFS is acceptable.**[2]

Especially for **live programmes** it is challenging (if not a matter of luck) to achieve Target Level at the end of the mix. Therefore, **a tolerance of ±1.0 LU around the Target Level of −23 LUFS is acceptable for such programmes** (in addition to live programmes, for example, ones which have an exceedingly short turnaround).

A tolerance of ±0.2 LU around −23 LUFS is accepted for loudness workflows. For live programmes, the tolerance is ±1.0 LU.

[2] The tolerance of ±0.2 LU around the Target Level of −23 LUFS has been introduced in EBU R 128 revision 3 (2020).

In cases where a programme deliberately consists of only background elements (for example, the music bed for a weather programme) or due to specific dramaturgical reasons an **intentionally lower** loudness level is possible. This is specified in R 128 (Programme Loudness **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 intentionally 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)]:

- **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 intentionally 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 of several broadcasters since the transition to loudness metering has shown that for initial 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** (that means, a bit lower than −23 LUFS), as background sounds will only add to the Programme Loudness Level. Furthermore, it is 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 foreground signals are set, and a fixed monitor gain has been established (see § 8.2), the audio engineer can **mix only by ear**. Checking the Momentary or Short-term Loudness Level and an occasional glance at the Integrated Loudness Level should provide 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 reaction 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**, less dynamic-compression artefacts such as 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 most important elements of a mix for a uniform subjective loudness impression are so-called '**foreground sounds'** — such as voice, title music or key sound effects. Sound elements may 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 low loudness level. On the other hand, a compressed rock guitar riff has a loudness level that is almost the same as its peak level! If those two signals are peak-aligned, the guitar riff will be much louder than the clink of the glasses. This example is meant to illustrate; it does NOT mean that those two signals should be mixed at equal loudness! The level of individual segments 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 achieved 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 **±0.2 LU** (or **±1.0 LU** for live and similar programmes), then no further action is needed. In post-production, if the level lies outside the tolerance, a simple **corrective static 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** which 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** 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)** (see § 11.2.2) quantifies the **loudness variation** within 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. Loudness Range is a useful **tool** especially for production to judge the long-term dynamic properties of a mix. LRA is also a useful **indicator** of potential *dynamics reduction processes* in the 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.

Working with loudness normalisation right away thus implies observing 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 should be created, 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 potential processing only downstream to tailor the signal to individual environments and platforms**.

It is important to understand that it is impossible and not useful to define a strict limit for LRA for all broadcasters and all programmes. LRA should not be a brick wall parameter for delivery specifications of programmes. Nevertheless, guiding upper 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 a certain LRA value for specific programming.

In any case, no parameter and 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 may 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 consecutive audio elements 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 heterogenous situation regarding Loudness Range, **EBU R 128** does *not* include a maximum permitted LRA value. A broadcaster has to decide if guidelines for upper limits of LRA for different distribution paths are to be used. These should be treated with care, as there exist programmes (for example, in the 'art house' film genre) which exhibit large values of LRA despite being suitable for being broadcasted. Essentially, an experienced sound engineer should check such cases.

3.4 *Climbing the True Peak*

The third measure recommended by R 128 concerns the **Maximum True-Peak Level** (*see also ITU-R BS.1770 and § 11.2.3*) 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. The True-Peak Level for generic PCM signals in production **shall not exceed −1 dBTP**; further values for different applications and distribution systems are given in EBU Tech Doc 3344 ('*Distribution Guidelines*').

The True-Peak Level in production shall not exceed −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 often unpredictable nature of the genre. 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 (for example, at −24 LUFS), so that unexpected crowd noise has 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 so-called 'comfort zone'.

The same thinking applies if the commentary temporal "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; see § 11.1.1 for an explanation of the gate), 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 perceived equal, the mixer/broadcaster could qualify the second event as a 'special circumstance' and have its Integrated Loudness Level deliberately lower than −23 LUFS.

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) is/are levelled around −23 LUFS.

3.5.2 Show

In Entertainment 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 apparent anchor signal: the host, the presenter(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 presenter(s)! Consequently, it may be more advantageous to balance the **audience** around Target Level – and have the presenter 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 presenter 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 slightly *increase* the listening level / monitor gain (1-2 dB) for such a programme. This usually avoids being carried away too easily with the excitement and having to deal with a final Loudness Level way above the tolerance window.

4. What to Measure in Production and Post-Production

4.1 Signal-Independent vs. Anchor-Based Normalisation

EBU R 128 generally recommends measuring the **whole programme**, independent of individual elements or components such as voice, music or sound effects (see **Figure 3** for a possible programme structure). This is considered to be the most generic practice for the vast majority of programmes.

EBU R 128 generally recommends measuring the whole programme.

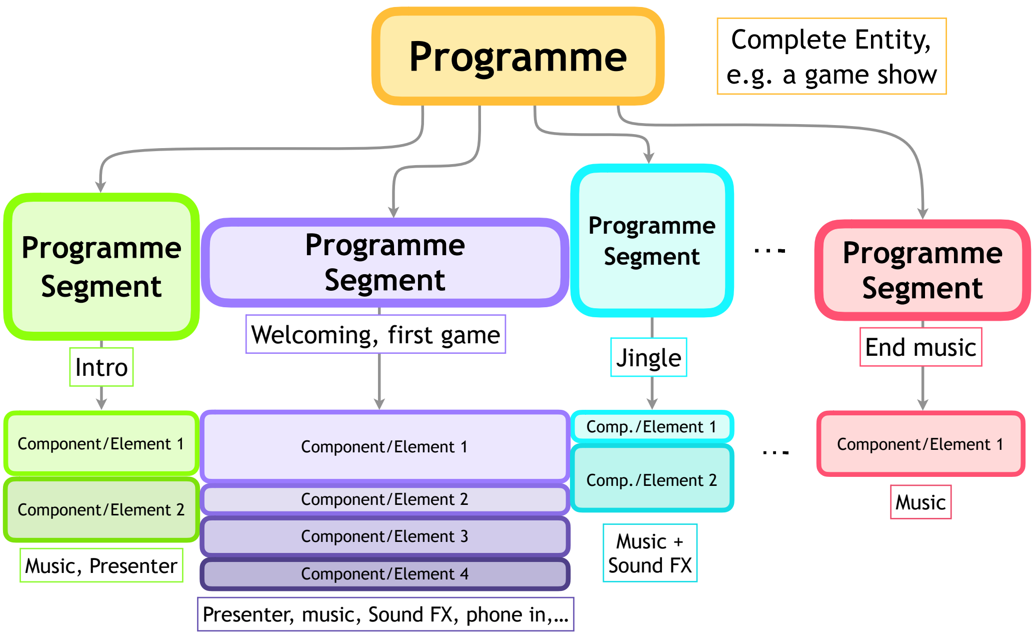


Figure 3: Programme structural hierarchy with its *Segments* and *Components/Elements*

For programmes with a very wide Loudness Range (>20 LU, approximately) and/or with a significant difference between Programme Loudness and Dialogue Loudness (>5 LU) one may optionally use a so-called **anchor signal** for loudness normalisation. 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.** This is anticipated and still within the spirit of R 128 (see § 9.4.3). Anchor-based normalisation can particularly help to fine-tune the loudness balance between wide loudness range programmes and advertisement breaks.

In any case, broadcasters have to be aware that, especially in a file-based environment, a different workflow might have to be established to treat programmes with anchor-based normalisation.

For **Cinematic Content** with speech as a main component (for example, feature films, drama productions or TV series) a **more elaborate strategy** may be used, using the measurement of additional parameters such as Dialogue Loudness and its difference to Programme Loudness (**Loudness-to-Dialogue Ratio** or **LDR**) to shape the dynamics processes accordingly. This is the subject of R 128 supplement 4 [6] and is covered in detail in § 9.4.

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 of a 5.1 Surround Sound mix is **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. For the time being, no such cases have been reported. 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. This is typically the case for the majority of broadcast content with the notable exception of mainly action movies.

5. Archival Content

For archival programmes there are basically **four options** to achieve loudness normalisation:

1) actually **changing** the loudness level of **all audio programme files** to be ‘on target’
2) **changing** the loudness level only ‘**on demand**’ for a temporary playout file
3) using the result of a loudness level measurement to automatically **adjust** the **playout level** without changing the original loudness level
4) **transporting** the **correct** individual (‘agile’) 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 during playout and are therefore the **preferred solutions** because they typically result in more headroom than the fourth option (see § 2.2, 2.3). Archive experts usually oppose the first solution as they want to keep the “essence” in its original shape and format. Solution 2 and 3 are widespread and often do not necessitate any further dynamic treatment of the content, because the average loudness level before the switch to loudness normalisation typically was higher than −23 LUFS. The Metadata solution is dependent on the correct value of the Programme Loudness parameter. In the past, this has been a challenge for many broadcasters as well as the source of abuse. Nevertheless, current audio codecs include Loudness and Dynamic Range Control (DRC) Metadata by default. Playback of content can be tailored to different devices, reproduction environments and listening situations. A closer look at the topic can be found immediately below in § 6.

The final choice depends on factors such as specific infrastructure, workflows, media asset management, availability of suitable equipment, financial resources, time etc.

6. Metadata

As mentioned 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 altering 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* [11], [12] and [13]). Furthermore, the values for the **Maximum Momentary Loudness Level** (MML) as well as the **Maximum Short-term Loudness Level** (MSL) may be stored, as especially MSL is helpful for controlling the dynamics of short-form content (typically <30 s; see also § 9.1). In January 2014 the EBU has published an open metadata standard, the **Audio Definition Model** [14], later adopted by the ITU as ITU-R BS.2076 [15], 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 such as '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 *Centre/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 dialogue as the anchor signal), the Metadata-parameter *dialnorm* actually describes the loudness of the dialogue, but only in that case (albeit a common one, especially on DVDs).

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**. Consequently, the Programme Loudness Metadata parameter will be **static**. In any case, Programme Loudness Metadata should always indicate the actual Programme Loudness (see § 9.4.4 for details regarding Cinematic Content).

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

- 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.

6.1.1 Intentionally Lower Programme Loudness, Loudness Offset

For programmes with an **intentionally lower Programme Loudness Level** (which consist, 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 such as individual movements of a symphony or short sound design production elements (for example, 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 or programme element. Also, 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 (DRC) Metadata

Just as loudness normalisation can be performed at the source audio signal or via Metadata, the same applies to dynamic processing. In the Metadata environment, DRC information typically generated during encoding is sent as part of the datastream in the form of *gain-words* and subsequent *profiles*. In the equipment of the consumer, this information is applied to potentially adapt the dynamic properties 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. For a more detailed description of actual methods, readers are encouraged to check out § 9.4.5.

The control of Loudness Range through actual processing of the audio signal at transmission is generally shifting the issue upstream. Broadcasters using DRC to perform dynamic adaptation downstream 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 multichannel audio 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 (typically from a 5.1 channel source signal). The loudness of such 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 are 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 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 and/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** (**L**eft **o**nly/**R**ight **o**nly) which directly combines the channels according to the downmix coefficients, and **Lt/Rt** (**L**eft **t**otal/**R**ight **t**otal) which applies ±90° phase shifting to a Mono-sum of the surround channels in order to achieve better compatibility with matrix-surround playback systems (for example, Dolby ProLogic).

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.1770[3], 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 by default!
Another systematic loudness difference arises when a lot of *"divergence"* is used for signals in the three front channels. Divergence controls the spill from, for example, a Centre channel signal to Left and Right, allowing a continuous change from a hard Centre signal (no divergence) to a signal being equally reproduced by all three front speakers (full divergence). In the latter case, the Stereo-downmix of such a Surround-Sound mix can have an up to 3 LU higher loudness level!
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 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").

[3] The +1.5 dB weighting coefficient for the surround signals in a loudness measurement according to ITU-R BS.1770 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 the weighting coefficient.

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

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.

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 [16]:

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. Relevant are:

- 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 ideally constantly monitors 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 programme's audio path. 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 [17]. In the same document the "Permitted Maximum Level (PML)" is still mentioned, as defined in ITU-R BS.645-2 [18]; 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 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. Due to the frequency-weighting curve of the loudness measurement algorithm ("K-weighting") using a loudness meter for alignment purposes is a potential source for errors. The EBU therefore recommends using a **peak-meter for alignment**.

The EBU recommends using a peak-meter for audio level alignment.

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** (**$L_{LISTref}$**) 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) [19] 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** 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 dB_C 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 [20], have to be employed. The (mono) 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 [21].

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 **LFE** (Low Frequency Effects channel) and the **subwoofer** need a separate calibration. Subwoofer alignment is a delicate matter and outside the scope of this document. Nevertheless, the following *guidance* is given: 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** dB_C 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 from the EBU Technical website (tech.ebu.ch/loudness).

The average loudness level of programmes according to EBU R 128 is typically *lower* (approximately 3 LU 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 highly dynamic audio signals. **Nevertheless, raising the monitor gain further to compensate for extremely dynamic audio signals such as, 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 establish a consistent, repeatable listening level to develop an "inner loudness reference".**

A consistent listening level is vital to develop 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- and Platform-specific Issues

The concept of **EBU R 128** emphasizes 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-world sound pressure level of a specific source, 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. 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/−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 individual genres. Allowing variations may challenge the system of equal average loudness from the onset. Naturally, the fear is that variations would be biased to the louder side. Exhibiting a Programme Loudness Level **intentionally 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 can be appropriate. Specifically, four areas are now investigated: **Commercials** (Advertisements) and **Trailers**, **Streaming**, **Radio** and generally **music** programmes as well as **Cinematic Content** (movies, drama productions etc.).

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 <60 s) there are too few data points to derive a meaningful number for LRA.

An alternative can be found in using the **Maximum Short-term Loudness Level** (MSL)[4].

[4] The **Maximum Short-term Loudness Level** is the highest value (in LUFS) of an audio signal's Short-term Loudness Level (integration time 3 s).

The EBU PLOUD group has published a dedicated **supplement** to R 128 (**R 128 s1**) that deals with the usage of MSL 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**
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 limit for MSL for short-form content.

For short-form content (< 60 s), a limit of −18 LUFS (+5 LU) for Maximum Short-term Loudness is recommended.

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 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, intentionally lower-loudness programmes will also pass unharmed. § 6.1.1 provides more details.

9.2 Streaming

Streaming of broadcast content (live programmes as well as file content) has become an increasingly important distribution method for EBU members. On streaming platforms, broadcasters are competing with a vast array of major content providers such as Facebook, Apple, Amazon, Netflix, Google or Disney as well as music services such as Spotify, Apple Music and Tidal or audio-only services such as podcasts. Typically, these services use a higher Target Loudness Level than specified in Broadcast Loudness standards. Consequently, the EBU PLOUD group has published a dedicated **supplement** to R 128 (**R 128 s2**) that gives guidance how to potentially adapt compliant programmes to streaming service providers.

Especially for mobile devices, European broadcasters should be aware of updated hazard standards by CENELEC and IEC ([22], [23]). Instead of setting a gain limit, sound level is now controlled by the energy of the content. Treating streaming or on-demand delivery differently than regular broadcasting is therefore no longer needed. It is hoped that broadcasters and major content providers anticipate this development and relax their Target Loudness Level accordingly.

There generally exist **two major use cases** for streaming:

- situations with sufficiently high playback gain and headroom as well as low background noise to faithfully reproduce dynamic programmes (for example, devices such as TV receivers, Smart Speakers and Home Theatre Equipment);
- and situations with limited playback gain and headroom and/or higher background noise (for example, devices such as smartphones or Personal Music Players);

In the first case, programmes should be streamed **unchanged**, that is at **−23.0 LUFS** (see **Figure 4**):

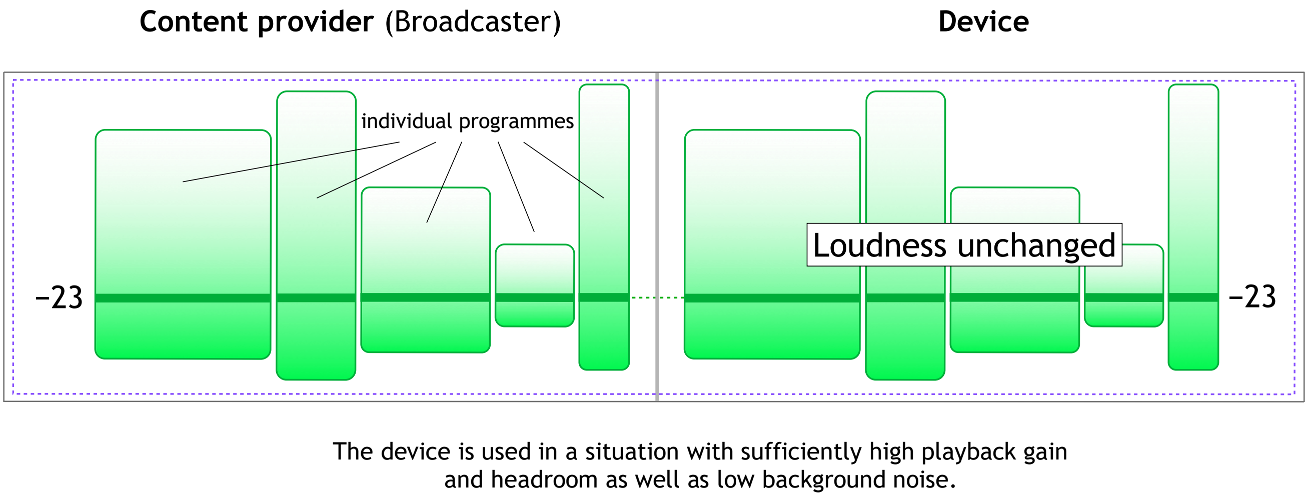


Figure 4: Normalisation Scheme for a stream at −23 LUFS with unchanged Loudness Level

In the second case, if the broadcaster wants to be in control of the quality of the potential dynamic treatment, the interim value for the *Distribution Loudness Level* may be raised and should be in the range of **−20.0** to **−16.0 LUFS** (see **Figure 5**):

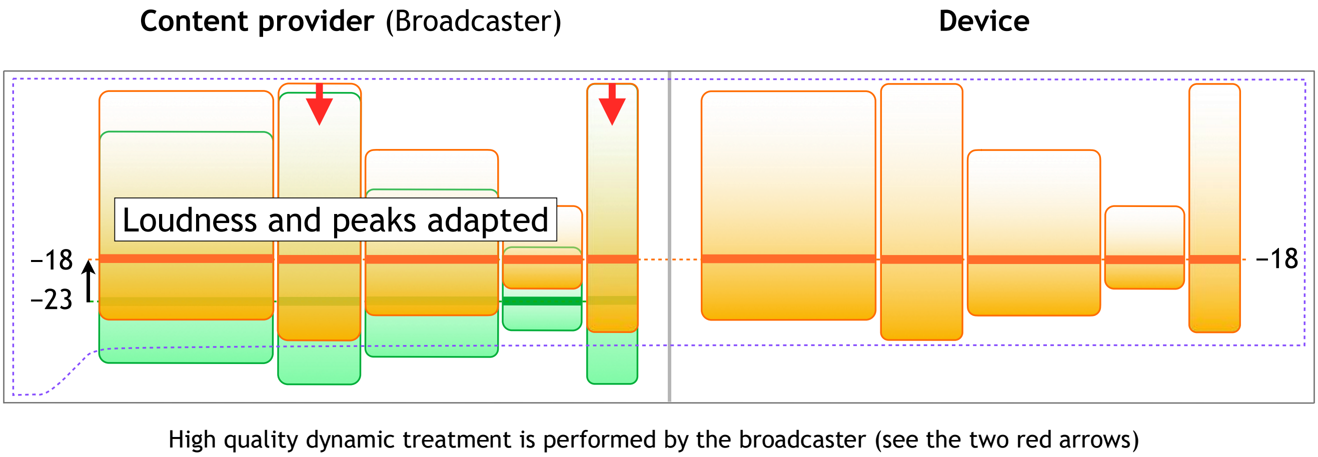
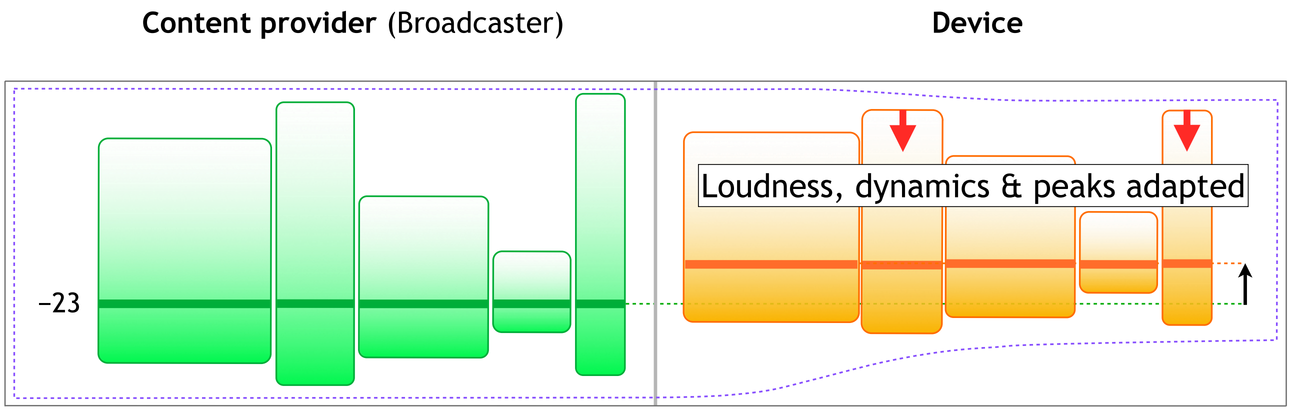


Figure 5: Normalisation Scheme for a stream at −23 LUFS with adaptation of the Loudness Level to, for example, −18 LUFS by the broadcaster prior to streaming

There is an alternative approach for the second case (situations with limited playback gain and headroom and/or higher background noise). If a broadcaster provides its own interface (app on a mobile device, web-browser) including loudness metadata, content may be streamed unchanged (that is, at −23.0 LUFS) with subsequent adaptation of the Target Level and the dynamic properties **in the device** (see **Figure 6**):



The device is used in a situation with limited playback gain and headroom and/or higher background noise;
High quality True-peak limiting is performed where needed (see the two red arrows)

Figure 6: Normalisation Scheme for a stream at −23 LUFS with adapted Loudness Level and dynamics processing (where applicable) in the reproduction situation

9.3 Radio, Music

9.3.1 Radio

After the introduction and successful adoption of recommendation **EBU R 128** in the TV realm, the uptake in **Radio** has been slower. In order to foster this transition, the EBU felt the need to give specific guidance in the form of another dedicated **supplement** to R 128 (**R 128 s3**) as well as an entire Tech Document (**Tech 3401** - Guidelines for Radio production and distribution in accordance with EBU R 128). The reader is encouraged to look up these publications for more details.

The ***key concept*** is the clear differentiation between the Radio *production* and the *distribution* side.

Loudness in Radio is facilitated by a clear distinction between production and distribution.

Despite different workflow details and programme genres, there is no general difference between loudness normalisation in Radio production and TV production. Thus, the same principles apply. Loudness measurements are performed according to **ITU-R BS.1770**, on the complete signal. A **common Loudness Level of −23 LUFS** enables straight forward exchange of content between Radio, TV and Online in the production stage.

For Loudness normalisation in *production*, the same principles apply for Radio and TV.

During ***production*** the aim is to craft a generic, accurately controlled Radio programme with open dynamics, agnostic of the distribution channel. If signal compression is used, it is applied due to artistic reasons, fitting the content. The Peak-to-Loudness Ratio (PLR), a measure of micro-dynamics, is appropriate for the content and *not* restricted in advance for a certain distribution platform.

During ***distribution*** the programme may be adapted according to the specific requirements of different distribution channels. This includes the choice of a *Distribution Loudness Level* as well as dynamic processing to comply with specifications such as legal requirements for FM or different loudness targets for Web-platforms.

Frequently, ***Metadata*** links production and distribution, supporting, for example, automated workflows or audio processing in the consumer device.

Figure 7 illustrates the clear distinction between *production* and *distribution* in Radio:

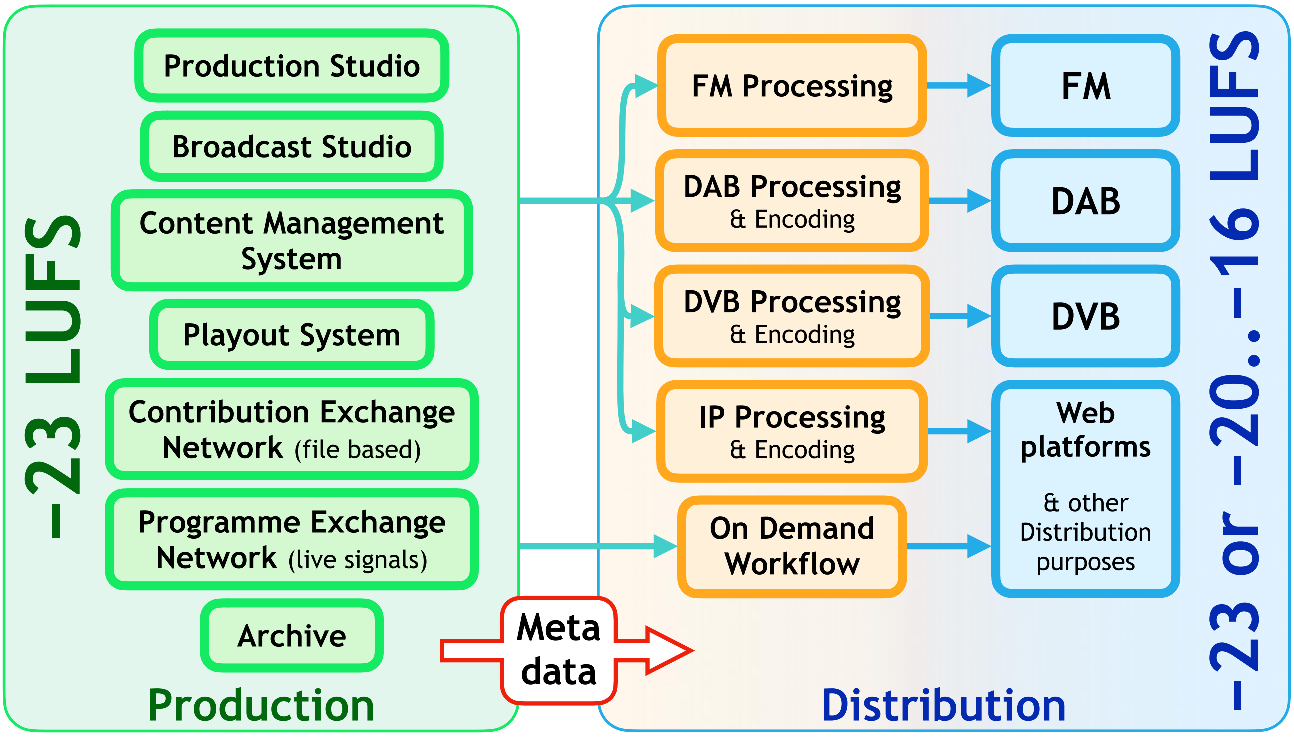


Figure 7: Two-stage structure of a Radio station workflow

To underline the benefits of **loudness normalisation**, two different levelling schemes of a Radio station using the *FM distribution* path are illustrated below. In **Figure 8**, note how applying loudness normalisation in *production* reduces the degree of compression and peak limiting applied. The Distribution Loudness Level is also −23.0 LUFS, the alignment to nominal FM carrier deviation is adapted (R 128 1). In **Figure 9**, the same sequence of programmes is all over the place regarding loudness levels which is tamed through severe compression and limiting. The Distribution Loudness Level and the alignment stay unchanged (Legacy 2). Both cases result in the same loudness level on the receiver side. Note the difference in the dynamic properties: R 128 1 shows a more open sound and a perfectly consistent loudness level.

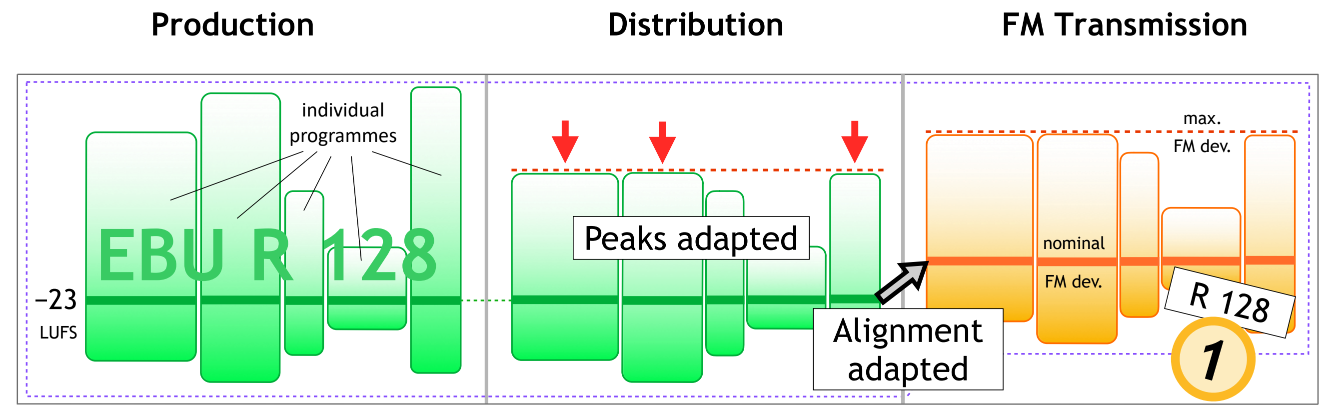


Figure 8: Levelling Scheme for a Radio station using R 128 compliant loudness normalisation at **−23 LUFS** in production and distribution with adapted alignment for FM (R 128 1). Mild peak limiting and/or dynamics processing is performed where needed (see the red arrows)

,

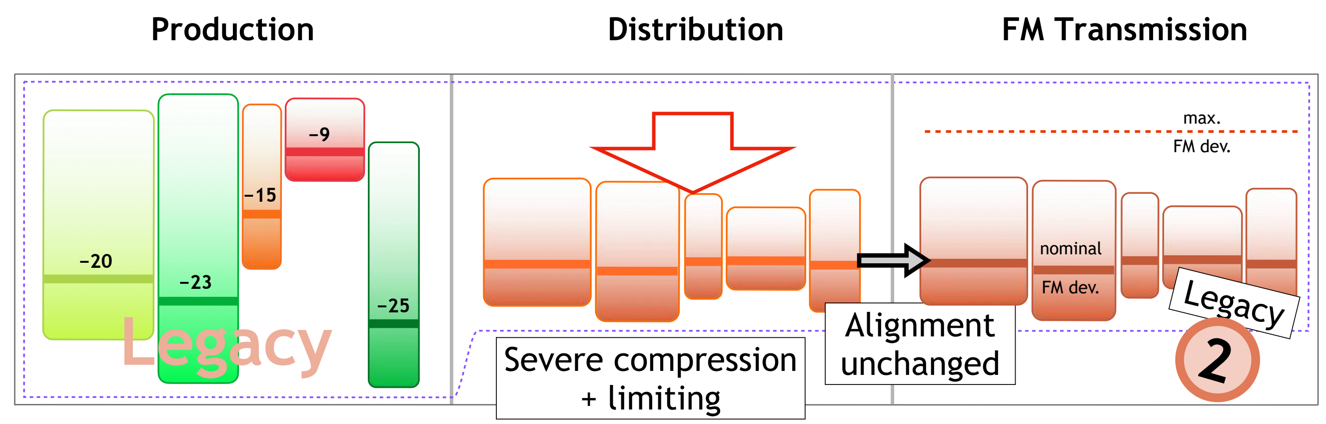


Figure 9: Levelling Scheme for a Radio station using no normalisation in production with legacy processing and alignment for FM. The resulting average loudness is the same as in Fig.1. Severe peak limiting and compression lead to a less dynamic, more squashed sound (Legacy 2).

9.3.2 Music, Loudness Offset

The experience of passionate music listeners suggests that certain programmes that contain mostly or only music, either with a wide Loudness Range such as classical music or with a higher degree of dynamic compression as an artistic property such as 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 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**. Exceptions for Radio are described in Tech 3401. In any case, the audience can still use their remote control to adjust (increase) the loudness level in their reproduction environment to their taste. Following programmes such as 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** such as Personal Music Players ('Track normalisation' vs. 'Album normalisation') and will also be of relevance for Radio programmes consisting of many individual elements. Those elements 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. See **EBU Tech 3401** for more specific guidelines for Loudness in Radio.

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

9.4 *Cinematic Content*

Because of frequently very dynamic mixes, **Cinematic Content** with speech as a main component (for example, feature films, drama productions or TV series) arguably presents the **biggest challenge** to integrate seamlessly into loudness-balanced programming. In several cases, the original mix is not suitable for a typical domestic listening environment. The difference between Programme Loudness (PL) and Dialogue Loudness (DL) – the **Loudness-to-Dialogue Ratio** or LDR – can be excessive, sometimes more than 15 LU! If such content 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/or too loud sound effects. The issue becomes especially noticeable when movies are interrupted by commercial blocks, also known as 'ad breaks': even if both the movie and the commercials are R-128 compliant, the difference in Dialogue Loudness can cause significant perceived transition jumps (see **Figure 10**).

Two particular questions regarding Cinematic Content are: what is the definition of "dialogue" and its level in this context? And: is there an unambiguous "average" and/or "normal" dialogue loudness? One of the few documents which give a hint is ITU-R BS.1864 [24] where *"normal" dialogue is that which is spoken in a normal level of speaking voice, that is, not shouting or whispering*. But explaining "normal" by itself ("normal level of speaking") is not entirely helpful. Furthermore, dialogue in Cinematic Content often has a wide range of levels which can lead to a significant difference between the "average" and the "normal" loudness of dialogue, whatever those terms actually mean. Imagine, for example, an action movie with a lot of shouting or a romantic film with predominantly whispering dialogue parts.

To offer a way out of the terminology uncertainty, this document introduces the terms **"wide-range"** and **"narrow-range"** Dialogue Loudness. *Wide-range* DL corresponds to the mean value of most or all of the spoken voice parts. *Narrow-range* DL is equivalent to the level of speech that is "not whispering, not shouting, not singing, and with no competing music or sound effects".

In order to get a better understanding of the situation in general and the range of PL and DL values (and subsequently, their difference LDR) in particular, the chairman of the PLOUD group conducted an extensive analysis of 44 feature films. This content came from ORF's database, was dubbed in German and **already dynamically adapted** by two experienced sound engineers without prior knowledge of the work of PLOUD regarding this topic.

The first goal was to compare different methods trying to measure *wide-range* Dialogue Loudness – manually and automatically. The manual method separated most of the speech parts by editing the Centre channel of a particular film, just eliminating very loud shouting and very soft whispering. The automatic method applied Dolby's Dialogue Intelligence (DI) algorithm. Loudness was measured according to ITU-R BS.1770, although DI does not apply the level gating. The resulting *mean difference* for the 44 dynamically adapted films was small (0.4 LU). The mean **Loudness-to-Dialogue Ratio** was for both methods below 4 LU.

Relying on the help of two PLOUD members, the analysis was extended to evaluate *narrow-range* DL. The speech separation was performed only manually, due to the lack of a suitable algorithm. Using the already separated parts from the first analysis, speech corresponding to the narrow-range DL-characteristics was isolated. One member constrained this process even more, taking samples only from the beginning, middle and end of the respective movie ("spot-checking"). As expected, the level of Dialogue Loudness using this method was lower than for wide-range DL. LDR was then obviously higher, but still ≤ 5 LU. As a consequence of the dynamic adaptation of the analysed films, the mean difference between narrow-range and wide-range DL-based LDR was just 1.1 LU, which is well within the range of audible acceptability. The chosen separation method is thus not so critical.

These results, as well as the observation that the LDR of the ORF-adapted films was 5 LU or lower, formed the basis for the recommended maximum value for LDR in **supplement 4 to R 128 'Loudness for Cinematic Content'**.

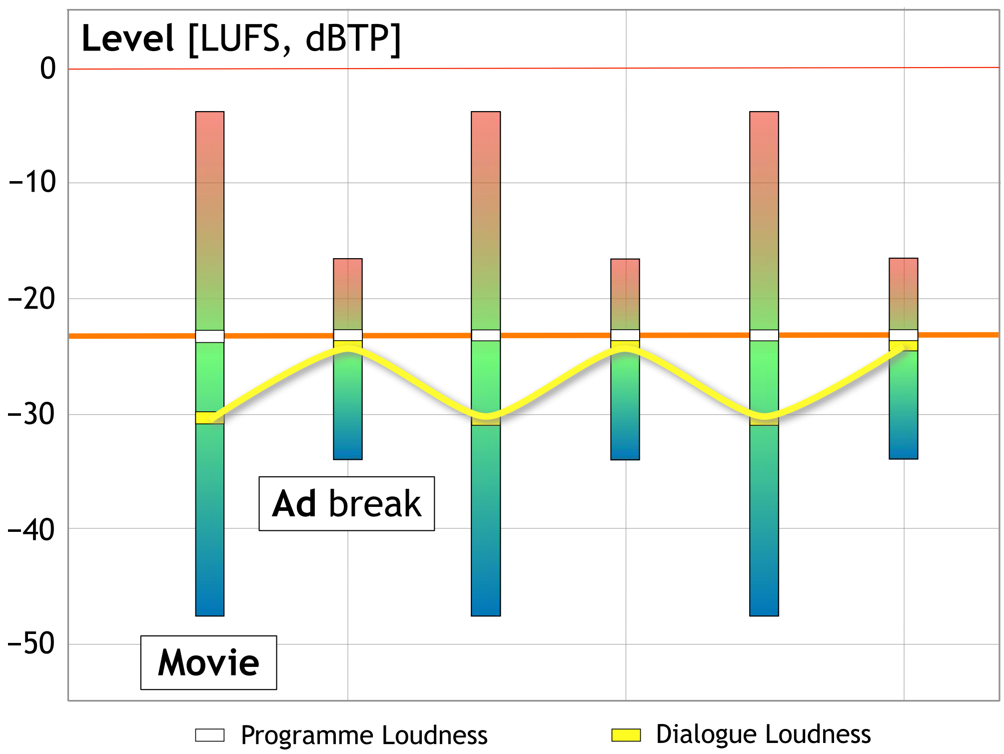


Figure 10: Jumps in Dialogue Loudness Levels between a movie and an Ad break

The recommendations in **EBU R 128 s4** help to quantify the issue and give guidance on the potential dynamic treatment and loudness normalisation of Cinematic Content. The recommended dynamic adaptation mainly concerns the **Loudness-to-Dialogue Ratio** (LDR).

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, such as clarity of pronunciation, choice of recording location, microphone technique etc.

The workflow to prepare Cinematic Content that needs adaptation for broadcasting or streaming and the subsequent normalisation can be somewhat more involved as choices need to be made regarding the measurement of Dialogue Loudness, the adaptation process targeting LDR and the actual normalisation scheme.

9.4.1 Measurement of the Loudness-to-Dialogue Ratio (LDR)

Reflecting the different *Dialogue Loudness* flavours used in the extensive feature film analysis mentioned above, a broadcaster should choose between either **manual** or **automatic** dialogue separation with subsequent loudness measurement:

- *Manual separation* naturally leads to a less consistent result as it depends on the experience and diligence of individuals. Analysing the whole film to extract wide-range dialogue manually is an exhausting, time-consuming process. Therefore, *spot-checking* narrow-range dialogue is more feasible and should be performed in the following fashion:
    - choose three on-screen dialogue sequences from the beginning, the middle and the end of the film (omit off-screen narration, radio voices and similar signals);
    - the selections should not contain music and/or competing sound effects;
    - each of the three sequences should be about 30 seconds long;
    - measure the individual loudness of the three sequences with ITU-R BS.1770;
    - take the average of the three loudness values to obtain *narrow-range Dialogue Loudness*

- *Automatic separation* gives more consistency and repeatability but is dependent on the separation algorithm. It is anticipated that algorithms using Artificial Intelligence (AI) can lead to increasingly precise results, reflecting the actual *wide-range Dialogue Loudness*.

A broadcaster should specify which separation and measurement method to use. This is applicable for internal processes and for external providers if the broadcaster already demands compliant adaptation of challenging Cinematic Content. If in-house assessment and adaptation is performed by default, a broadcaster may accept or even encourage original mixes without any platform- or provider-specific processing in advance. In that case, no specification other than a high-quality mix is required.

As the measurement of *Programme Loudness* according to ITU-R BS.1770 is straightforward, the value for the Loudness-to-Dialogue Ratio is easy to calculate: it is the difference between Programme Loudness and Dialogue Loudness (LDR = PL − DL).

9.4.2 Adaptation process

EBU R 128 s4 specifies that the LDR value of adapted content should not exceed 5 LU. If after the calculation of LDR this is already the case (LDR ≤ 5 LU), no dynamic processing needs to be performed. Depending on the normalisation scheme, True-peak limiting may be applicable, though.

If LDR exceeds 5 LU, dynamic processing should be applied. The challenge is to adapt the dynamic properties while fully preserving the intentions and workflow of EBU R 128. At the same time, the quality and experience of the original soundtrack should be preserved as much as possible. Again, the broadcaster has the choice between **manual** and **automatic** processes:

- *Manual processing* relies on experienced audio engineers, using tools such as compression, limiting and gain riding of the whole mix or individual channels to control the dynamic properties and bring LDR below or equal to 5 LU. If separate stem mixes are available, results can be optimized due to the dialogue track not being "contaminated" with other signal types such as sound effects or music;
- *Automatic processing* again relies on the processing algorithm. Offline and iterative algorithms can leave deliberate loudness jumps intact and can therefore provide a more unobtrusive adaptation result with fewer artefacts. Although sophisticated methods are available to fine-tune the dynamics processing on the fly, algorithms in iterative offline workflows are preferred over real-time or "live" processors which typically react without knowledge of the coming audio signal or its characteristics so far.

A broadcaster should specify which adaptation method to use. This is applicable for internal processes and for external providers if the broadcaster already demands adaptation of challenging Cinematic Content to be compliant with its specific workflow and Quality Control (QC) procedure. If in-house adaptation is performed by default, a broadcaster may accept or even encourage original mixes without any platform- or provider-specific processing in advance. In that case, no specification other than a high-quality mix is required.

In the following table the main parameters for dynamically adapted Cinematic Content are listed:

Loudness-to-Dialogue Ratio (LDR)	**≤ 5 LU**
Maximum True-Peak Level	**−1 dBTP** (linear audio)

9.4.3 Normalisation schemes

Similar to the measurement of LDR and the dynamics adaptation process, the subsequent **normalisation** of Cinematic Content offers a choice for the broadcaster:

- *Normalisation of Programme Loudness (PL) to −23.0 LUFS:* this is the default scheme in accordance with the established solution described in EBU R 128. It is also the recommended scheme if Dialogue Loudness is very close to Programme Loudness. Nevertheless, even with the Loudness-to Dialogue Ratio below or equal to 5 LU, the Dialogue Loudness Level can be as low as −28 LUFS. Especially for broadcasters with Cinematic Content being interrupted by Ad breaks (check again Figure 10 above), this normalisation scheme may lead to noticeable loudness jumps of speech. In such a scenario, the second scheme may be chosen, which is:
- *Normalisation of Dialogue Loudness (DL)* ***up to*** *−23.0 LUFS:* this is so-called "**anchor-based**" normalisation with dialogue as the anchor signal. It is important to note that the choice for the level of DL is one out of a *range* of possible levels: with the Target Level of −23.0 LUFS as defined in EBU R 128 and a maximum LDR of 5 LU, Dialogue Loudness can effectively sit on one level in the range of −28.0 ≤ DL ≤ −23.0 LUFS. In order to better balance the average speech level of adapted Cinematic Content with speech during Ad breaks, DL will probably sit closer to or at −23.0 LUFS for such a use case.

A broadcaster should specify which normalisation scheme to use. The choice is dependent on internal workflows, legacy normalisation schemes, metadata- or non-metadata-reliant solutions and other factors. Broadcasters should be aware of the fact that anchor-based normalisation was part of the EBU R 128 paradigm since the beginning, it has always been described in this Tech Document as a potential scheme for highly dynamic content with speech as a main component. In case a broadcaster chooses anchor-based normalisation, the *range* for the choice of the single value for Dialogue Loudness described above should be anticipated.

Anchor-based normalisation using DL typically leads to a Programme Loudness Level higher than −23.0 LUFS (with LDR ≤ 5 LU, it can maximally reach −18.0 LUFS). This results in less headroom than normalisation to Programme Loudness. True-peak limiting may have to be applied accordingly.

To illustrate the typical situation of quantifying, adapting and normalising Cinematic Content, the graph and corresponding explanations in EBU R 128 supplement 4 are reproduced here for clarification purposes (**Figure 11**):

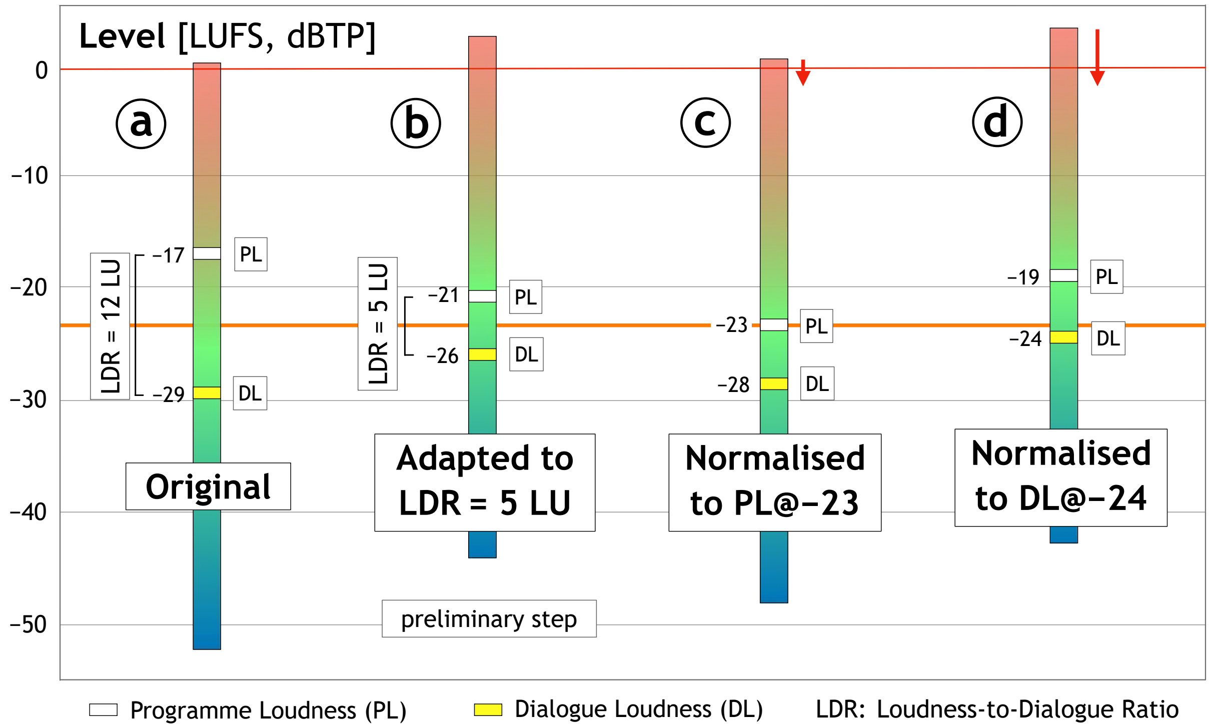


Figure 11: Illustration of the two normalisation schemes (PL vs. DL), after adaptation of the Loudness-to-Dialogue Ratio (see details below)

Case **a**: Original movie mix; in this example, PL = −17 LUFS, DL = −29 LUFS (thus, LDR = 12 LU). If this programme is normalised with PL to −23 LUFS (6 dB attenuation), DL would be at −35 LUFS, 12 LU *lower* than the Target Level! If anchor normalisation to −23 LUFS using DL is performed (+6 dB gain), PL would be at −11 LUFS, 12 LU *higher* than the Target Level (with 12 dB True-peak limiting necessary)! Both solutions are not satisfactory;

Case **b**: Example of dynamic adaptation of the whole programme to bring the Loudness-to-Dialogue Ratio (LDR) to 5 LU: after adaptation the value of DL is −26 LUFS and PL is −21 LUFS. This is a preliminary step which can result in higher peaks than in the original signal, depending on the adaptation process;

Case **c**: The adapted programme is normalised to **PL = −23 LUFS** (−2 dB gain) which is the *default normalisation* scheme according to EBU R 128. The level of DL at transmission is −28 LUFS. Headroom is slightly increased, but True-Peak Limiting may still be necessary depending on the degree of LDR-adaptation (case b) and on the distribution path;

Case **d**: The adapted programme is *in this example* normalised to **DL = −24 LUFS** (+2 dB gain) which is *anchor-based normalisation*. The level of PL at transmission is −19 LUFS. Dialogue Loudness close to or at the Target Level can provide a better balance to the level of speech, especially in the presence of advertisement breaks. Headroom is slightly decreased, so True-peak limiting will typically be necessary for most distribution paths. −24 LUFS is a value for DL anticipating the **range** of possible levels. DL can be normalised to −23 LUFS if wanted, but it does not need to be. It can, for example, be normalised to −25 LUFS, if a broadcaster wants to ensure that the resulting PL never exceeds −20 LUFS. A broadcaster can choose the appropriate level according to its workflow, consumer preferences, taste and other criteria.

9.4.4 Metadata for Cinematic Content

In EBU R 128 s4 three scenarios of normalisation with regards to metadata are presented in a hierarchical fashion: normalisation without loudness metadata — normalisation with one loudness metadata parameter — and normalisation with two principal loudness metadata parameters. The focus is on anchor-based normalisation, but the concept holds true in a generic way also for normalisation of Programme Loudness:

- *Systems without loudness metadata (see § 9.4.3 above):* **default normalisation** is achieved by applying a static gain to the whole programme so that **Programme Loudness** is at the Target Level of −23.0 LUFS. **Anchor-based normalisation** is achieved by applying a static gain to the whole programme so that **Dialogue Loudness** is at the desired Target Level (***up to*** −23.0 LUFS). Subsequent peak limiting may be applicable;
- *Systems with one loudness metadata parameter:* **default normalisation** is achieved by setting this parameter to the **Programme Loudness** value. According to EBU R 128 the normalisation process is recommended already for the production stage. Consequently, the single loudness metadata parameter should be set to indicate −23. **Anchor-based normalisation** is achieved by setting this parameter to the **Dialogue Loudness** value. For compliance with EBU R 128 the single chosen value should lie in the range of −28 ≤ DL ≤ −23;
- *Systems with two loudness metadata parameters:* when both metadata fields are available (for example, in the Audio Definition Model [15]), Loudness Metadata should indicate Programme Loudness and Dialogue Loudness. **Default normalisation** is achieved by applying the **PL value** to achieve the desired Target Level (−23.0 LUFS for compliance with EBU R 128). **Anchor-based normalisation** is achieved by applying the **DL value** to achieve the desired Target Level. When both loudness metadata parameters (PL and DL) are actually present, providers define whether normalisation to PL or anchor-based normalisation to DL should be used.

9.4.5 Dynamic Range Control at the consumer’s side

Dynamic Range Control (DRC) in codecs is traditionally realised by a real-time process in the encoder that generates DRC metadata for transmission to the consumer’s decoder. Alternatively, DRC metadata can be also generated by an offline process before encoding that is based on a pre-analysis of the complete programme. In the consumer‘s decoder, the transmitted DRC metadata is applied to the untouched mix dependent on the specific playback configuration (see also [25]).

EBU R 128 s4 encourages an offline process in which the programme can be extensively analysed and the signal itself iteratively processed prior to transmission. This can lead to a significant increase in the audio quality of the dynamic processing. Furthermore, EBU R 128 s4 recommends that if the Loudness-to-Dialogue Ratio is lower than 5 LU, the mix should be left untouched and the process becomes transparent. In general, this cannot be achieved by using a traditional, non-guided real-time DRC, thus potentially compromising audio quality.

Anchor-based normalisation for Cinematic Content in some cases has the consequence of a more symmetrical processing regarding automatic Dynamic Range Control in, for example, the legacy Dolby AC-3 system at the consumer’s side (the process for MPEG4 legacy DRC in AAC is similar). The DRC process centres around the **dialnorm** value. If the dialnorm value corresponds to Dialogue Loudness, 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 speech level might lie at or below 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. Time constants applied to the gain-control signal reduce frequent toggling. For the more aggressive DRC

profiles *Film Standard* (the default profile), *Music Standard* and *Speech*, dialnorm corresponds to the lower edge of the unity part anyway. DRC will consequently be often applied to programmes with considerable variations, regardless if the dialnorm value indicates Dialogue Loudness or Programme Loudness.
Dynamic adaptation according to EBU R 128 s4 reduces the impact of DRC processing. *To further minimize processing artifacts when using legacy codecs, a moderate DRC profile or no DRC processing at all should be chosen, as in the situation of programmes with LDR ≤ 5 LU.*

For legacy codecs, a moderate DRC profile or no DRC processing at all should be chosen.

Similar to AC-3/E-AC-3, AC-4 also uses the dialnorm value for loudness normalization and serves as the reference level for the DRC processing in the decoder. The DRC in AC-4 includes additional features such as the carriage of gain or compression profile metadata allowing for custom DRC to be defined.

In MPEG codecs such as xHE-AAC and MPEG-H Audio, it is possible to use an offline process for the generation of Dynamic Range Control metadata before the actual bitstream encoding. This process is usually based on an extensive analysis of the complete original mix upfront. Unlike traditional real-time DRC, this approach allows for an iterative optimization of the intended Dynamic Range Control behavior in the consumer's decoder. If the target criteria are met and no dynamic adaptation is needed, the original mix is left untouched independent of whether programme- or anchor-based normalization is used.

Broadcasters should consider extended loudness and dynamic processing features in their choice of transmission codecs.

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 flip 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 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.
- 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. Keep an eye on the market and demand action 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 (and all other EBU loudness documents) is **ITU-R BS.1770**, the result of extensive work by the International Telecommunication Union. The purpose of this recommendation was to establish an open source algorithm for the measurement of electrical **loudness** and the **true peak levels** of audio signals, with the benefit of a simple implementation. In brief, it defines a "**K-weighting**" filter curve (a modified second-order high-pass filter, see **Figure 12**), which forms the basis for matching an inherently subjective impression with an objective measurement.

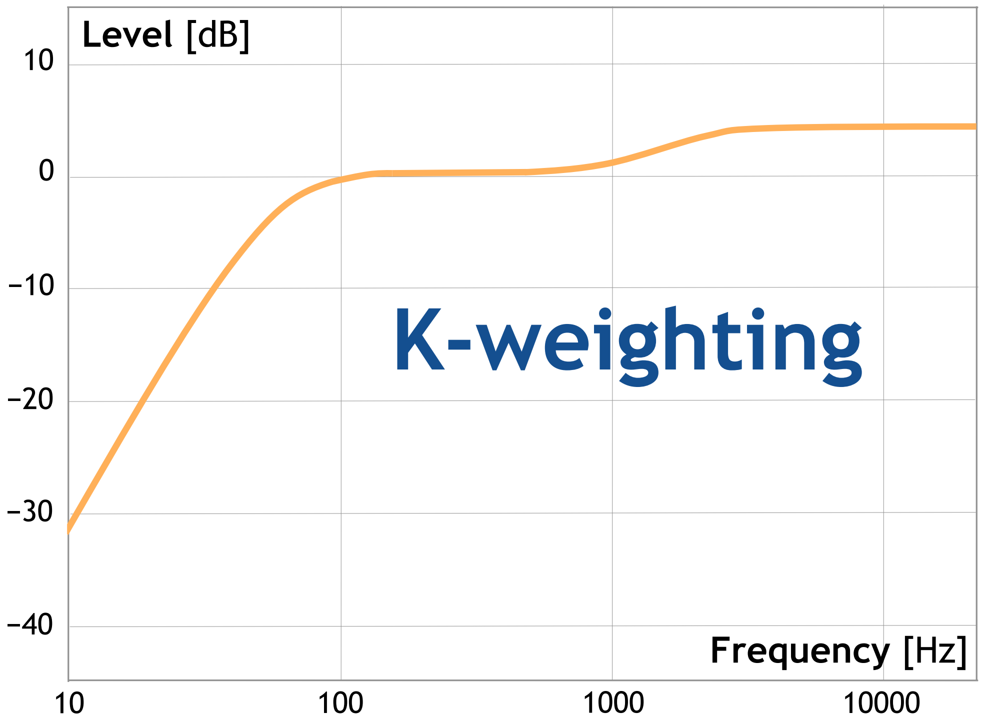


Figure 12: "K-weighting" filter curve for loudness measurement

The K-weighting curve is applied to all channels. The total mean square level is calculated (with different gain factors for the front and surround channels of a 5.1 signal (see **Figure 13**) and the result is displayed as "LKFS" (Loudness, K-Weighting, referenced to digital Full Scale), or "LUFS"[5] (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 considered for the loudness measurement according to ITU-R BS.1770. This may lead to abuse of the LFE with unnecessary high signal levels. Inclusion of the LFE in ITU-R BS.1770 is still pending.

[5] 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 [26].

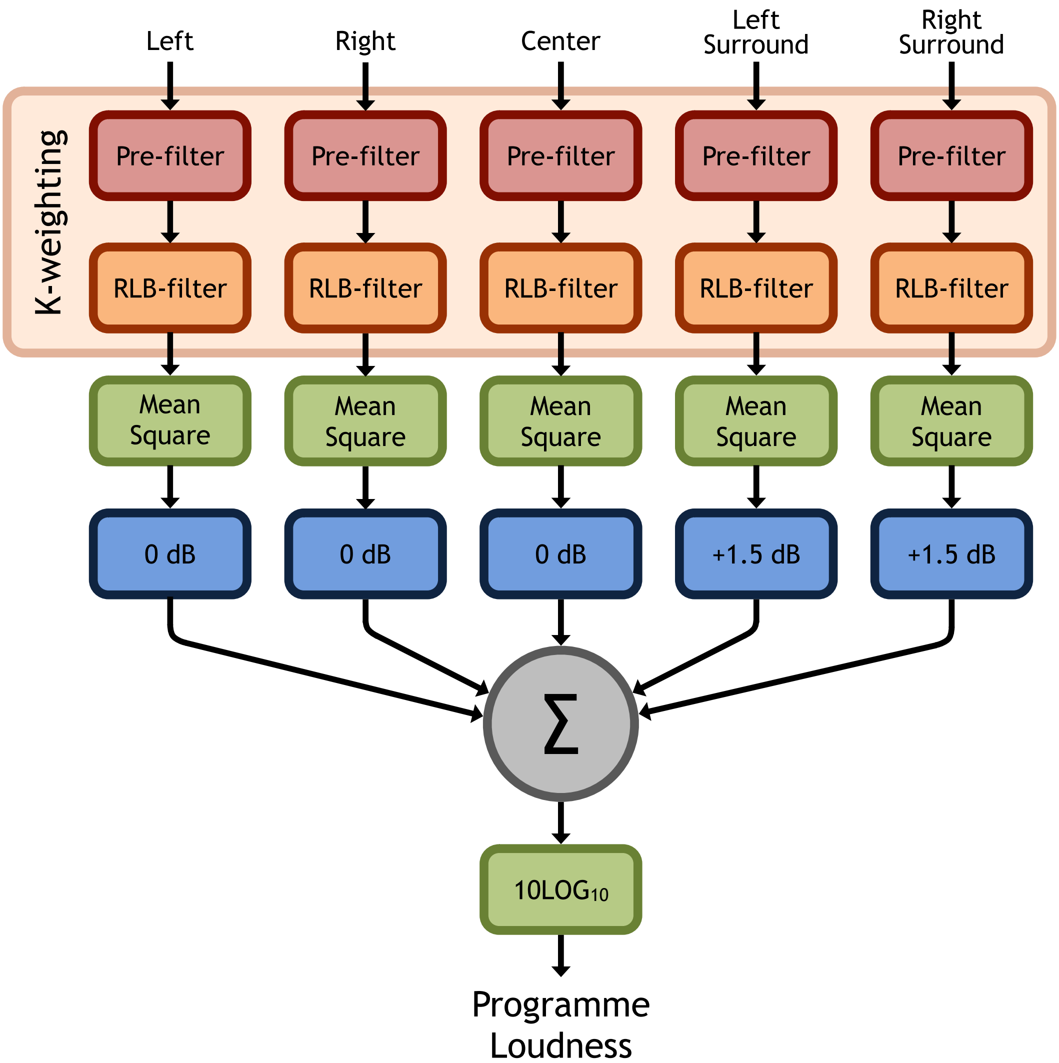


Figure 13: Channel processing and summation in ITU-R BS.1770

11.1.1 Gating

The **gating** method developed by the EBU PLOUD group is part of **ITU-R BS.1770** (since revision 2), and thus a **worldwide recommendation**. This refinement 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:

- An absolute 'silence' gating threshold at **−70 LUFS** for the computation of the absolute-gated loudness level;
- A relative gating threshold, **10 LU** below the absolute-gated loudness level;
- 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 14 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 in light blue) — and then the remaining loudness levels are averaged, giving the gated result (−25.2 LUFS).

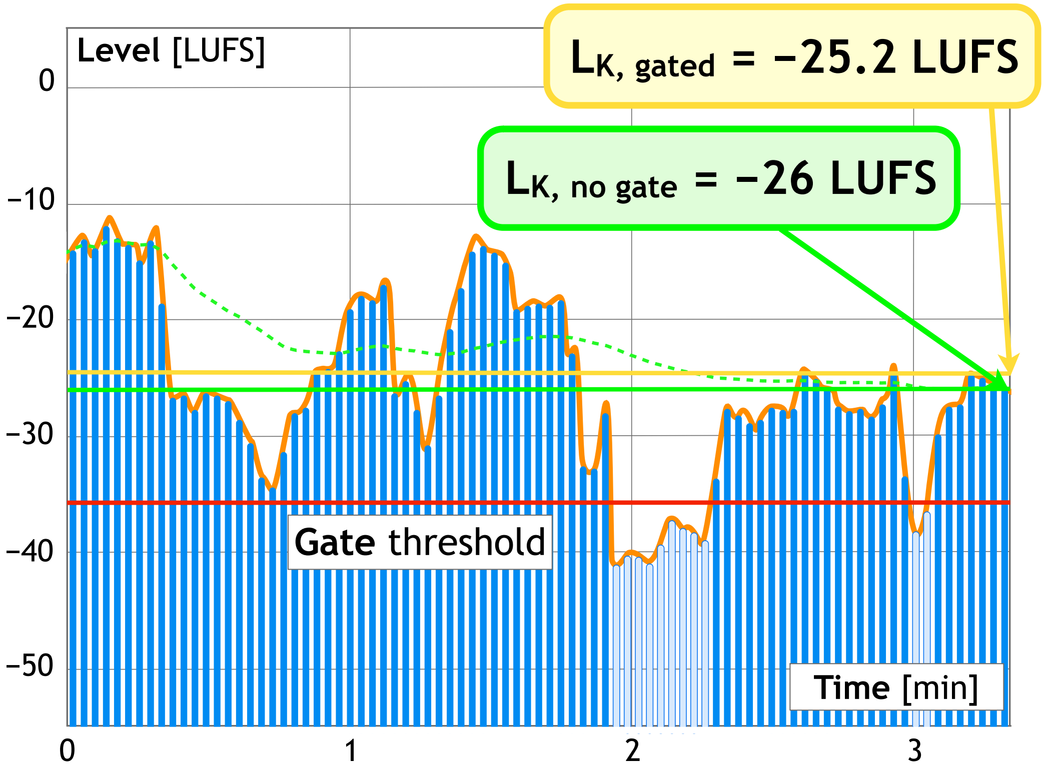


Figure 14: Explanation of the relative gating measurement

In revision 4 of ITU-R BS.1770 the loudness measurement algorithm has been expanded to Advanced Audio systems according to ITU-R BS.2051 [27]. Basically, any additional channel gets the same treatment as the ones covered so far, albeit with a slight refinement concerning the angular spread and elevation of channels with a gain factor of +1.5 dB:

Elevation (φ)	**Azimuth (θ)**		
	|θ| < 60°	**60° ≦ |θ| ≦ 120°**	**120° < |θ| ≦ 180°**
|φ| < 30°	0 dB	+1.5 dB	0 dB
else	0 dB		

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 recommending a specific '**Target Level**' for loudness normalisation, introduces the parameter **'Loudness Range'** (LRA) and provides a value for the **'Maximum True Peak Level'** that shall not be exceeded. 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 programme[6]. 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) and EBU Tech Doc 3341.

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

−23.0 LUFS

In order to avoid rejection of programmes due to accumulation of metering tolerances, **a general measurement tolerance of ±0.2 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 such as 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 loudness within a programme, thereby excluding the extremes. 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

[6] The term '**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.

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** as a tool for production, 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 can be found in EBU Tech Doc 3342. See also § 3.3 for further information.

Figure 15 shows the loudness distribution and the subsequent statistical derivation of LRA for the movie 'The Matrix'; 25 LU is probably challenging for most living rooms...

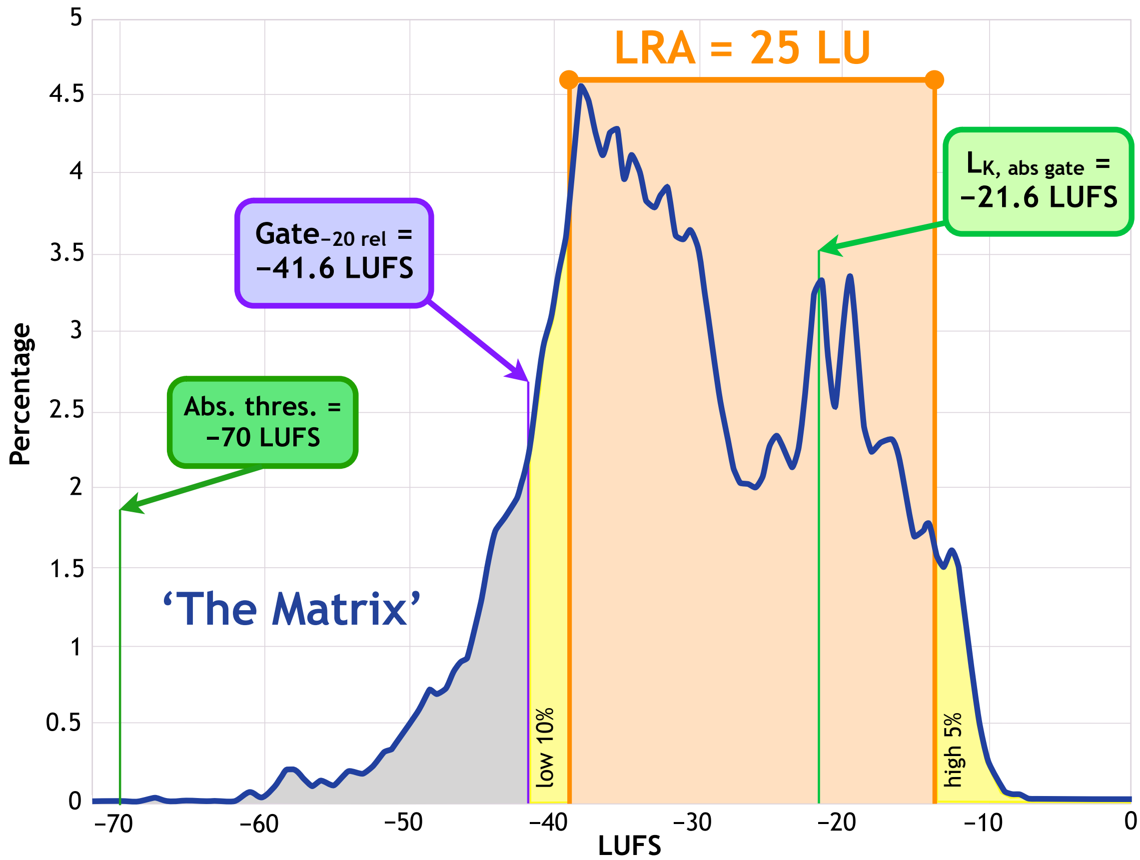


Figure 15: Loudness Range (LRA) as a result of the statistical distribution of loudness levels

For short-form programmes (typically <60 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. Limiting the **maximum** value of the **Short-term Loudness Level** 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 still is) the **Q**uasi **P**eak **P**rogramme **M**eter (**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 [28].

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: dB**TP** – **d**eci**B**el referenced to digital Full Scale using a **T**rue-**P**eak 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). The measurement tolerance is **±0.3 dB** (for signals with a bandwidth limited to 20 kHz). 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 in R 128 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**;
- 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** is mainly a tool for production and can be used to help decide if dynamic compression is needed (dependent on genre, target audience and transmission platform);
- The **Maximum True-Peak** Level in production shall not exceed **−1 dBTP**. For data reduction systems (like MPEG 1L2 and Dolby AC-3) the limit is **−2 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' Loudness Metering according to Tech 3341 (see below).

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 windows[7] 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 [29], 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 intentionally 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.

[7] '**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.

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

Many more manufacturers have adopted 'EBU Mode' in their metering equipment. 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 recommendation for loudness metering.

12. References

[1] **EBU R 128** *'Loudness Normalisation and Permitted Maximum Level of audio signals'*

[2] **ITU-R BS.1770** *'Algorithms to measure audio programme loudness and true-peak audio level'*

[3] **EBU R 128 s1** *'Loudness Parameters for Short-form Content (adverts, promos etc.)'*; supplement 1 to EBU R 128

[4] **EBU R 128 s2** *'Loudness for Streaming'*; supplement 2 to EBU R 128

[5] **EBU R 128 s3** *'Loudness in Radio'*; supplement 3 to EBU R 128

[6] **EBU R 128 s4** *'Loudness normalisation for Cinematic Content'*; supplement 4 to EBU R 128

[7] **EBU Tech 3341** *'Loudness Metering: 'EBU Mode' Metering to supplement Loudness Normalisation in accordance with EBU R 128'*

[8] **EBU Tech 3342** *'Loudness Range: A Descriptor to supplement Loudness Normalisation in accordance with EBU R 128'*

[9] **EBU Tech 3344** *'Guidelines for Distribution and Reproduction in accordance with EBU R 128'*

[10] **EBU Tech 3401** *'Guidelines for Radio production and distribution in accordance with EBU R 128'*

[11] **EBU R 85** *'Use of the Broadcast Wave Format for the Exchange of Audio Data Files'*

[12] **EBU R 111** *'Multi-channel Use of the BWF Audio File Format (MBWF)'*

[13] **EBU Tech 3306** *'MBWF/RF64: An extended File Format for Audio'*

[14] **EBU Tech 3364** *'Audio Definition Model — Metadata Specification'*

[15] **ITU-R BS.2076** *'Audio definition model'*

[16] **ITU-R BS.775** *'Multichannel stereophonic sound system with and without accompanying picture'*

[17] **EBU R 68** *'Alignment level in digital audio production equipment and in digital audio recorders'*

[18] **ITU-R BS.645** *'Test signals and metering to be used on international sound programme connections'*

[19] **EBU Tech 3276 -E** *'Listening conditions for the assessment of sound programme material'* (1998, 2004 — supplement 1)

[20] **IEC 61260-1** *'Electroacoustics — Octave-band and fractional-octave-band filters — Part 1: Specifications'*

[21] **IEC 61672-1** *'Electroacoustics — Sound level meters — Part 1: Specifications'*

[22] **CENELEC EN 50332-3** *'Sound system equipment: headphones and earphones associated with personal music players - maximum sound pressure level measurement methodology - Part 3: measurement method for sound dose management'*

[23] **IEC 62368-1** *'Audio/video, information and communication technology equipment - Part 1: Safety requirements'*

[24] **ITU-R BS.1864** *'Operational practices for loudness in the international exchange of digital television programmes'*

[25] **CTA-2075** *'Loudness Standard for Over the Top Television and Online Video Distribution for Mobile and Fixed Devices'*

[26] **ISO 80000-8** *'Quantities and Units — Part 8: Acoustics'*

[27] **ITU-R BS.2051** *'Advanced sound system for programme production'*

[28] **Lund, Th.** *'Stop counting samples'*, AES paper N°6972, 121st AES Convention, October 2006

[29] **ITU-R BS.1771** *'Requirements for loudness and true-peak indicating meters'*