



OPERATING EUROVISION AND EURORADIO

## TECH 3343

# PRAKTYCZNE WSKAZÓWKI DOTYCZĄCE PRODUKCJI I IMPLEMENTACJI ZGODNYCH Z ZALECENIEM EBU R 128

INFORMACJE DODATKOWE DLA ZALECENIA R 128

STATUS: WERSJA 1.0

THIS INFORMAL TRANSLATION OF TECH 3343 INTO POLISH HAS BEEN KINDLY PROVIDED BY POLISH RADIO. PLEASE NOTE THAT THE OFFICIAL AND DEFINITIVE VERSION OF TECH 3343 IS THE EBU ENGLISH VERSION. PLEASE REPORT ANY ERRORS YOU NOTICE IN THIS TRANSLATION TO [TECH@EBU.CH](mailto:TECH@EBU.CH)

Genewa  
Luty 2011



EBU – TECH 3342

Praktyczne wskazówki  
dotyczące produkcji i implementacji  
zgodnych z Zaleceniem EBU R 128



**Informacje dodatkowe dla EBU R 128**

**Status: Wersja 1.0**

Genewa  
Luty 2011



## Notacja zgodności

Niniejszy dokument zawiera zarówno tekst o charakterze **normatywnym** jak i **informacyjnym**.

Cały tekst ma postać normatywną poza tekstem we Wprowadzeniu, każdym paragrafem jednoznacznie oznakowanym jako „informacyjny” lub poszczególnymi paragrafami zaczynającymi się od słowa „Uwaga:”.

Tekst **Normatywny** opisuje elementy nieodzowne lub obowiązkowe. Zawiera kluczowe słowa zgodności „musi”, „powinien” lub „może”, zdefiniowane poniżej:

„Musi” oraz „nie musi”: wskazują wymagania, które muszą być ściśle spełnione i odnośnie których, w celu zgodności z dokumentem, nie zezwala się na jakiegokolwiek odchyłki.

„Powinien” oraz nie „powinien”: wskazują że, spośród kilku możliwości, jedna z nich jest zalecana jako szczególnie odpowiednia, bez wzmiankowania lub wykluczania pozostałych.

LUB wskazuje określony kierunek działania, który jest preferowany, ale niekoniecznie wymagany.

LUB wskazuje że (w formie negatywnej) konkretna możliwość lub kierunek działania nie jest zalecana, ale nie jest też niedozwolona.

„Może” oraz „nie może” wskazuje kierunek działania dozwolony w granicach podanych w tym dokumencie.

Słowo **Domyślny** określa obowiązkowe (w zdaniach zawierających „musi”) lub zalecane (w zdaniach zawierających słowo „powinien”) presety (ustawienia wstępne) które mogą opcjonalnie być zastąpione działaniem użytkownika lub uzupełnione innymi opcjami w zastosowaniach zaawansowanych. Obowiązkowe pozycje domyślne muszą być obsługiwane. Obsługa zalecanych domyślnych jest preferowana, ale niekoniecznie wymagana.

Tekst **informacyjny** jest potencjalnie pomocny użytkownikowi, ale nie jest niezbędny i może być usunięty, zmieniony lub dodany podczas edycji bez wpływu na tekst normatywny. Tekst informacyjny nie zawiera żadnych słów kluczowych zgodności.

Implementacja zgodna jest taką, która zawiera wszystkie obowiązkowe warunki („musi”) oraz, jeśli nastąpi wprowadzenie, zawiera wszystkie opisane tutaj zalecane warunki („powinien”). Implementacja zgodna nie wymaga włączenia opcjonalnych warunków („może”) i nie musi ich wprowadzać tak, jak zostały opisane.



## Spis treści

<b>1.</b>	<b>Wprowadzenie</b>	<b>7</b>
<b>2.</b>	<b>EBU R 128, ITU-R BS.1770</b>	<b>9</b>
2.1	Głośność programu	11
2.2	Zakres głośności	12
2.3	Poziom rzeczywisty szczytowy (TPL) i Maksymalny Dozwolony TPL	13
2.4	Logo R 128	15
<b>3.</b>	<b>Ogólna koncepcja normalizacji głośności</b>	<b>15</b>
3.1	Wartości szczytowe a głośność	15
3.2	Normalizacja sygnału a metadane	16
3.3	Używanie parametru o nazwie Zakres Głośności	19
3.4	Wspinając się na Rzeczywisty Szczyt	20
<b>4.</b>	<b>Strategie dla normalizacji głośności</b>	<b>21</b>
4.1	Produkcja i post-produkcja	21
4.2	Pomiar głośności w produkcji i post-produkcji	23
4.3	„Gotowi, ustawieni (poziomy), START!”	24
4.4	Zakres głośności dla produkcji i post-produkcji	26
<b>5.</b>	<b>Co należy mierzyć w produkcji i post-produkcji?</b>	<b>28</b>
5.1	Normalizacja niezależna od sygnału a normalizacja „zakotwiczona”	28
5.2	Kanał Efektów Niskoczęstotliwościowych (LFE)	29
<b>6.</b>	<b>Produkcja i emisja materiałów opartych na plikach</b>	<b>30</b>
6.1	Bloki tworzące obróbkę	31
6.2	Ogólne strategie regulacji poziomu głośności - obróbka	32
<b>7.</b>	<b>Metadane</b>	<b>34</b>
7.1	Metadane Głośności Programu	35
7.2	Metadane Kontroli Zakrest Dynamikii	35
7.3	Współczynniki downmiksingu (zgrania do wersji stereo)	37
<b>8.</b>	<b>Regulacja sygnałów w świetle normalizacji głośności</b>	<b>38</b>
8.1	Regulacja sygnału i poziomu	38
8.2	Poziom odstuchu	39
<b>9.</b>	<b>Implementacja i migracja</b>	<b>40</b>
9.1	Ogólne porady dotyczące Migracji oraz implementacji	40
9.2	10 punktów Działań dla migracji i implementacji	41

<b>10.</b>	<b>Problemy związane z rodzajem programu .....</b>	<b>41</b>
10.1	Reklamy i Trailery .....	42
10.2	Muzyka .....	43
<b>11.</b>	<b>Źródła .....</b>	<b>44</b>

## **Podziękowania**

Aczkolwiek dokument ten jest wynikiem intensywnej współpracy w ramach istniejącej w EBU grupy PLOUD, ale to jej szefowi właśnie, Florianowi Camererowi zawdzięczam, że dopiero po jego wielotygodniowej a nawet wielomiesięcznej pracy i wysiłku włożonym w opracowanie, zestawianie, wzbogacanie oraz oczyszczanie tego tekstu, mógł się on pojawić jako publikacja.

## **Dedykacja**

Dokument ten jest dedykowany dwóm wielkim inżynierom elektroakustykom, ich nazwiska to Gerhard Stoll i Gerhard Steinke.



## Praktyczne wskazówki dotyczące produkcji i implementacji zgodnych z Zaleceniem EBU R 128

<i>Komitet EBU</i>	<i>Pierwsze wydanie</i>	<i>Poprawiony</i>	<i>Ponownie wydany</i>
Komitet Techniczny	2011		

**Słowa kluczowe:** Audio, Głośność, normalizacja, produkcja, implementacja.

### 1. Wprowadzenie

Niniejszy dokument opisuje w praktycznych szczegółach jedną z najważniejszych zmian w historii dźwięku w środowisku nadawczym - zmianę podejścia dotyczącego regulacji poziomów z normalizacji wartości szczytowej do **normalizacji głośności**. Można bez przesady stwierdzić, że **pomiar głośności** oraz **normalizacja głośności** oznaczają **prawdziwą rewolucję regulacji poziomu** dźwięku. Zmiana ta jest istotna w obliczu problemu, który stał się głównym źródłem poirytowania widzów i radiosłuchaczy na całym świecie. Problemu skoków poziomu dźwięku w przerwach pomiędzy programami, pomiędzy samymi programami oraz pomiędzy kanałami (*patrz przypis 2 do definicji terminu „program”*).

Podejście i sposób regulacji głośności wpływa na wszystkie etapy toru sygnału audio, od produkcji do dystrybucji i transmisji. Tak więc celem ostatecznym jest takie zharmonizowanie poziomów głośności dźwięku, aby uzyskać jednakowy uniwersalny poziom głośności korzystny dla słuchacza.



**Normalizacja głośności to  
prawdziwa rewolucja w regulacji  
poziomu dźwięku!**

Należy tutaj od razu podkreślić, że **nie oznacza to**, że poziom głośności powinien być przez cały czas niezmienny i jednolity w *obrębie programu*. Wręcz przeciwnie! Normalizacja głośności powinna sprawić, że **średnia głośność w całym programie** będzie taka sama we wszystkich programach, natomiast w danym programie poziom głośności może się oczywiście zmieniać, odpowiednio do potrzeb natury artystycznej czy technicznej. Przy nowym (rzeczywistym) poziomie szczytowym oraz (w większości przypadków) niższym średnim poziomie głośności, możliwy do uzyskania zakres dynamiki (lub raczej „zakres głośności”, *patrz punkt 2.2*) będzie tak naprawdę większy niż ten, jaki uzyskujemy stosując w praktyce miksowania w studiach aktualną normalizację względem wartości szczytowych.

Podstawą koncepcji normalizacji głośności jest połączenie **Zalecenia Technicznego EBU** o nazwie **R 128 „Normalizacja głośności i dopuszczalny poziom maksymalny sygnałów audio”** [1] oraz **Zalecenia ITU-R BS.1770 „Algorytmy pomiaru głośności programów audio oraz rzeczywistego szczytowego poziomu audio”** [2].



## Zalecenia R128 oraz ITU-R BS.1770 stanowią podstawę. Informacje szczegółowe są podane w czterech dodatkowych dokumentach EBU

Oprócz R 128, grupa EBU PLOUD opublikowała cztery inne dokumenty:

- Dokument Techniczny EBU 3341 „*Pomiar głośności: pomiar typu ‘Tryb EBU’ uzupełniający normalizację głośności zgodnie z EBU R 128*” [3]
- Dokument Techniczny EBU 3342 „*Zakres głośności: deskryptor uzupełniający normalizację głośności zgodnie z EBU R 128*” [4]
- Dokument Techniczny EBU 3343 „*Praktyczne wskazówki dotyczące produkcji i implementacji zgodnych z EBU R 128*” [ten dokument] oraz
- Dokument Techniczny EBU 3344 „*Praktyczne wskazówki dla systemów dystrybucji zgodnych z EBU R 128*” [5]

Dokumenty techniczne dotyczące „*Pomiarów głośności*” oraz parametru „*Zakresu głośności*” również odgrywają ważną rolę w praktyce wprowadzania normalizacji głośności. Będą one przedstawione i będziemy się do nich odwoływać w odpowiednich rozdziałach tego dokumentu.

Cykl ten zamyka dokument zatytułowany „*Wskazówki dotyczące dystrybucji*” omawiający wszystkie aspekty normalizacji głośności w procesie dystrybucji sygnałów audio oraz krytyczne powiązania pomiędzy produkcją programu a jego końcowym odbiorcą - klientem. Ponieważ jest to dokument bardzo szczegółowy, nie będzie tutaj dokładnie przedstawiony a tylko okazjonalnie wymieniany.

Na początku tych „*Praktycznych wskazówek*” wprowadzono najważniejsze elementy dokumentów EBU R 128 oraz ITU-R BS.1770, po nich omówiono ogólną koncepcję oraz filozofię normalizacji głośności. Następnie przedstawiono strategie głośności przeznaczone dla produkcji oraz post-produkcji (pomiar, miksowanie, metadane itd.), potem zaś organizację pracy opartej na plikach, dotyczy to zagadnień wprowadzania danych, odczytu-emisji oraz archiwizacji (pomiar, normalizacja automatyczna, metadane itd.).

W oddzielnym rozdziale opisano bardziej szczegółowo **Metadane**. Omówiono tam regulację sygnałów audio oraz studyjne **poziomy odsłuch**, oraz podano **praktyczne porady** przy **przechodzeniu** do produkcji opartej na normalizowaniu głośności (implementacja oraz migracja). **Zagadnienia dotyczące specyficznych materiałów**, takich jak reklamy czy zwiastuny (trailery) jak również programy muzyczny są omówione w rozdziale ostatnim.

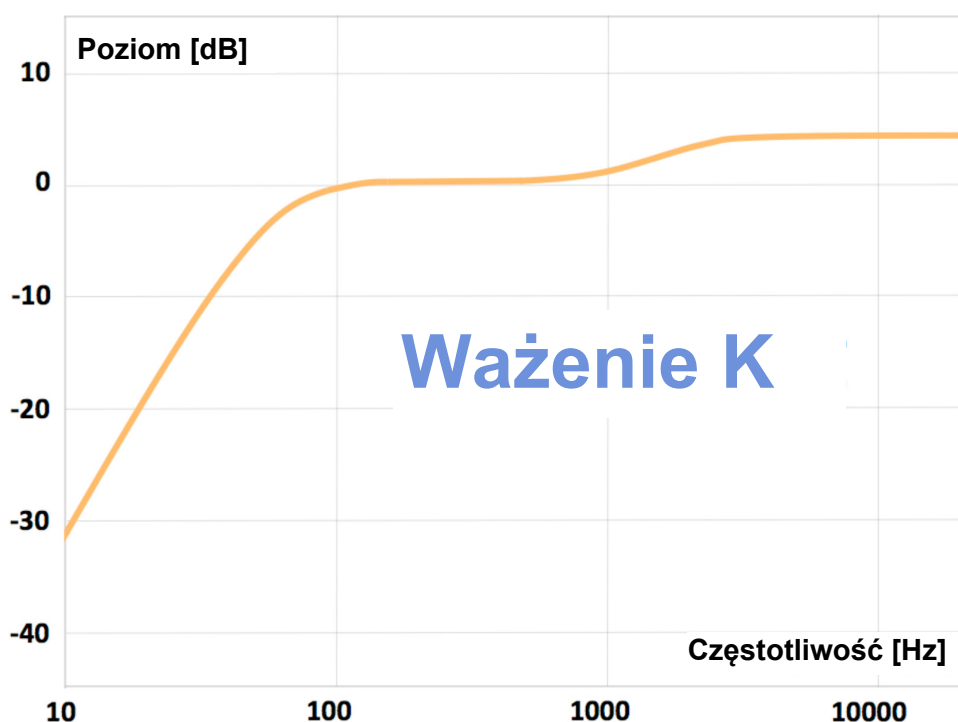
Te Praktyczne Wskazówki stanowią „**dokument żywy**”, który w miarę upływu czasu i gromadzenia doświadczeń przez nadawców będzie uzupełniany o kolejne informacje i wskazówki przydatne dla tej fundamentalnej zmiany sposobu traktowania i zapewniania właściwych proporcji sygnałów audio.

Prosimy pamiętać, że wiele dokumentów normalizacyjnych jest co pewien czas poprawianych, dotyczy to również i tej pozycji. Usilnie zalecamy więc zapoznanie się z najnowszymi wersjami takich dokumentów.

## 2. EBU R 128, ITU-R BS.1770

Dokument EBU R 128 ustanawia przewidywalną i dobrze zdefiniowaną metodę pomiaru głośności w programach typu wiadomości, sport, reklamy, teatr, muzyka, promocje, film itd. w poszczególnych miejscach toru u nadawcy, pomaga zatem profesjonalistom w tworzeniu solidnych specyfikacji dotyczących wprowadzania, produkcji, odtwarzania oraz dystrybucji materiałów na wielu platformach. R 128 jest oparty całkowicie na standardach otwartych i ma na celu zharmonizowanie sposobu produkcji i pomiarów dźwięku w skali międzynarodowej.

Podstawą R 128 jest Zalecenie ITU-R BS.1770, będące rezultatem wyłożonych prac ITU (Międzynarodowego Związku Telekomunikacyjnego). Celem tego standardu było ustanowienie uzgodnionego otwartego algorytmu dla pomiarów głośności oraz rzeczywistych poziomów szczytowych programu. Jest to standard pewny a jego zaletą jest prostota wprowadzania. Pokróćce - definiuje on krzywą filtra realizującego operację „Ważenia K” (zmodyfikowany filtr górnoprzepustowy drugiego rzędu, patrz Rysunek 1), filtr ten stanowi podstawę dla dopasowywania naszych nieodłącznie subiektywnych wrażeń do obiektywnych wyników pomiaru.



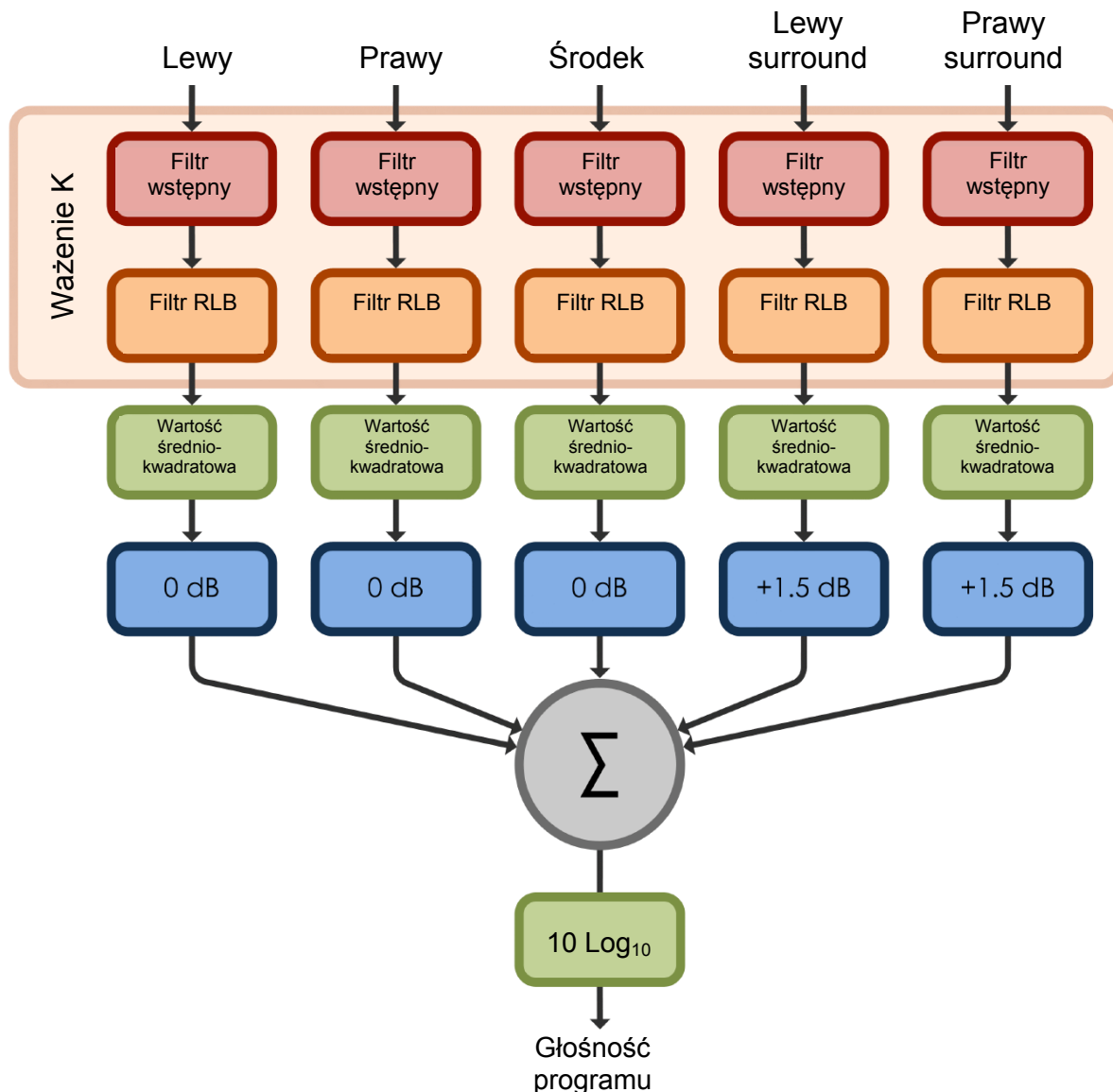
Rysunek 1: Charakterystyka filtra realizującego „Ważenie K” w pomiarze głośności

Taka krzywa ważąca jest wprowadzana do wszystkich kanałów (poza kanałem efektów niskoczęstotliwościowych „LFE”, który aktualnie jest wykluczany z pomiarów, patrz uwaga poniżej). Następnie zostaje obliczona sumaryczna wartość średniokwadratowa (przy różnych współczynnikach wzmocnienia dla kanałów z przodu oraz kanałów otaczających, patrz Rysunek 2) i tak otrzymany wynik jest prezentowany jako „LKFS” (Loudness, K-Weighting, referenced to digital Full Scale - głośność, ważona K, odniesiona do pełnej skali), lub „LUFS”<sup>1</sup> (Loudness Units, referenced to digital Full Scale - jednostki głośności odniesione do pełnej skali). Dla pomiarów względnych są używane jednostki głośności „LU” (Loudness Units), przy czym 1 LU odpowiada 1 dB.

<sup>1</sup> EBU zaleca używanie jednostek „LUFS” (tak jak to określono w Dokumentie Technicznym EBU 3341). „LUFS” jest równoważna „LKFS” i eliminuje niespójność pomiędzy dokumentami ITU-R BS.1770 a ITU-R BS.1771. „LUFS” spełnia również wymagania międzynarodowej normy ISO 80000-8 [6] dotyczącej nazewnictwa.

### Kanał efektów niskoczęstotliwościowych Low Frequency Effects (LFE)

Zgodnie z ITU-R BS.1770, kanał LFE - Low Frequency Effects (to ta jedynka „1” w nazwie konfiguracji dźwięku otaczającego „5.1”), występujący w wielokanałowym sygnale audio, aktualnie nie jest brany pod uwagę przy pomiarach głośności. Możliwe są więc sytuacje niewłaściwego zasilania tego kanału nieprawidłowym sygnałem o zbyt wysokim poziomie. Trwają prace badawcze w celu oszacowania subiektywnego wpływu, jaki sygnał LFE wywiera na odczuwanie głośności, jak również właściwego sposobu wprowadzenia tego kanału do obiektywnych pomiarów głośności.



Rysunek 2: Obróbka i sumowanie sygnałów w kanałach zgodnie z ITU-R BS.1770

O ile dokument BS.1770 definiuje samą metodę pomiaru, to dokument R 128 rozszerza to definiując dokładnie specyficzny „Poziom Docelowy” dla normalizacji głośności jak również metodę **bramkowania**, poprawiającą dopasowywanie głośności tych programów, które zawierają dłuższe chwile cisy lub odizolowaną mowę. Opracowanie EBU było potrzebne po to, aby sprostać potrzebom twórców programów, ze szczególnym uwzględnieniem zapewnienia środków do pomiarów miksów końcowych (a nie tylko jednego elementu programu, takiego jak mowa czy muzyka) oraz zakresu głośności programu.

W tym celu, EBU wprowadziło trzy nowe parametry:

- Głośność programu
- Zakres głośności
- Rzeczywisty poziom szczytowy

## 2.1 Głośność programu

Głośność programu opisuje długoczasową i uśrednioną głośność liczoną za czas trwania programu<sup>2</sup>. Parametr ma postać jednej liczby (w jednostkach LUFs, z jednym miejscem po przecinku) wskazującej „jaka jest średnia głośność danego programu”. Pomiar jest wykonywany miernikiem zgodnym z zaleceniem ITU-R BS.1770 uzupełnionym o funkcję **bramkowania**. Bramka służy tutaj do wstrzymywania pomiaru głośności, gdy sygnał opada poniżej ustalonej wartości progowej. Bez funkcji bramkowania, programy z dłuższymi odcinkami ciszy lub zawierające ciche dźwięki w tle, miałyby zaniżoną wartość poziomu głośności. Takie programy, po ich późniejszej normalizacji, byłyby za głośne.

Po serii badań EBU uzgodniło dla bramki wartość progową równą **-8 LU** (Loudness Units - jednostki głośności, 1 LU  $\equiv$  1 dB) ustawianą względem pomiaru niebramkowanego LUFs, długość bloku ustalono na 400 ms (dalsze szczegóły -patrz [7]). Aktualnie jest to przedmiotem poprawek dokumentu ITU-R BS.1770. Badanie te potwierdziły również (poza innymi wnioskami) wybór **Poziomu Docelowego**, do którego ma być normalizowany każdy sygnał audio, poziom ten wynosi:

**-23.0 LUFs (przy progu bramki równym -8 LU)**



**- 23 LUFs to nowe centrum  
świata regulacji poziomów  
audio !!!**

### ***Dlaczego akurat -23 LUFs?***

Badania i pomiary poziomu głośności sygnałów rzeczywistych stacji nadawczych wykazały, że średni poziom głośności ma wartość około **-20 LUFs** (z wieloma wyjątkami...). Drugi argument w tej dyskusji dostarczyła ITU, w postaci dokumentu ITU-R BS.1864 „*Działania praktyczne dotyczące głośności stosowane w międzynarodowej wymianie programów telewizji cyfrowej*” [8]. W dokumencie BS.1864, zalecono Poziom Docelowy równy **-24 LUFs**, ale bez funkcji bramkowania. Nieformalne testy przeprowadzone przez członków grupy EBU PLOUD wykazały, że różnice w wynikach pomiarów programów o małym do średniego poziomie głośności bez bramki lub z bramką ustawioną na poziomie względnym **-8 LU** wynoszą około 0 - 1 LU. Tak więc wartość **-23 LUFs** z bramką jest w wielu przypadkach niemal równoważna z wartością **-24 LUFs** bez bramkowania. Wartość **-23 LUFs** (z bramką) została więc uznana za **najniższy możliwy poziom głośności programu** dzięki czemu przejście z poziomu średniego równego **-20 LUFs** stało się mniejszym wyzwaniem.

<sup>2</sup> Termin „program” oznacza również reklamę, materiał promocyjny itd. Dla jasności - reklamy i pozycje podobne, umieszczane na zewnątrz i wewnątrz emisji materiału, który generalnie uważa się za „program” są traktowane jako programy na własnych prawach (również pojedyncze reklamy w danym bloku są oddzielnymi programami); łatwiej jest wówczas integrować takie pozycje z programami dłuższymi. Ewidentnie twórcy każdego z typów tak rozumianych programów nie wiedzą, jaki inny materiał będzie im towarzyszyć a zatem każdy typ programu powinien być rozważany oddzielnie. W tym dokumencie, termin **program** oznacza program ukończony przez produkcję a nie kombinację programu, przerywników oraz reklam, które docierają do odbiornika widza lub słuchacza w całym czasie emisji programu.

Ponieważ sądzono, iż wartość -20 LUFS nie zapewni wystarczającego zapasu dynamiki, podjęto decyzję o przyjęciu wartości **-23.0 LUFS**.

Odchyłki rzędu  $\pm 1.0$  LU są do przyjęcia w tych programach, w których dokładna normalizacja do Poziomu Docelowego równego **-23.0 LUFS** nie jest praktycznie możliwa (programy typu „na żywo” lub programy o zbyt krótkim czasie). W przypadkach, gdy poziomy poszczególnych sygnałów programu są zbyt często *nieprzewidywalne*, gdy program zawiera tylko elementy tworzące tło (na przykład podkład muzyczny przy prognozie pogody) lub gdy programy są specjalnie bardzo ciche, poziom głośności może wypaść poza granicami tolerancji. Jednak liczba takich przypadków powinna wciąż się zmniejszać.

Ostatnio (wrzesień 2010), EBU zasugerowała ITU dołączenie do treści dokumentu BS.1770 metody z bramkowaniem względnym. Na kolejnym spotkaniu ITU ta sugestia została przyjęta, choć z nieco niższym poziomem progowym bramki wynoszącym **-10 LU** poniżej poziomu głośności niebramkowanej. Zgodnie z testami PLOUD, wyniki uzyskane przy bramce o progu względnym **-10 LU** różnią się tylko marginalnie od wyników dla wartości **-8 LU**. Tak więc, po opublikowaniu kolejnej poprawionej wersji ITU-R BS.1770 z dołączoną funkcją bramki **-10 LU**, funkcja ta będzie również dołączona do dokumentu EBU R 128 i dokumentów towarzyszących, zwłaszcza Dokumentu Technicznego EBU 3341.

*Szczegóły realizacji funkcji bramkowania są niezbędne tylko dla producentów sprzętu. Dla użytkownika różnice poszczególnych wykonań są małe, niemniej jednak zaleca się użytkownikom aktualizację ich urządzeń w celu zachowania spójności pomiarów.*

## 2.2 Zakres głośności

Innym ważnym tematem jest zakres głośności, jaki byłby odpowiedni dla *wszystkich* programów (o ile nie przekraczają one tolerowanego zakresu głośności podczas odsłuchu w domu). Parametr o nazwie **zakres głośności (Loudness Range (LRA))** określa (w jednostkach LU) rozrzut wartości w pomiarach głośności programu. Parametr jest oparty na statystycznym rozkładzie głośności w programie, wykluczając w ten sposób wartości ekstremalne. Dlatego, dla przykładu, pojedynczy strzał z broni palnej nie jest w stanie wpłynąć na wynik obliczenia LRA. Zalecenie EBU R 128 nie podaje maksymalnej dozwolonej wartości LRA, ponieważ zależy ona od takich czynników jak próg tolerancji u średniego słuchacza danej stacji, rozkład typów programów obecnych stacji itd. Natomiast Zalecenie R 128 usilnie **zachęca do korzystania z parametru LRA** przy określaniu, czy konieczna jest obróbka dynamiki sygnału audio oraz podczas dopasowywania takiego sygnału do wymagań konkretnego kanału transmisji lub platformy. Więcej szczegółów odnośnie parametru LRA można znaleźć w Dokumencie Technicznym EBU 3342.

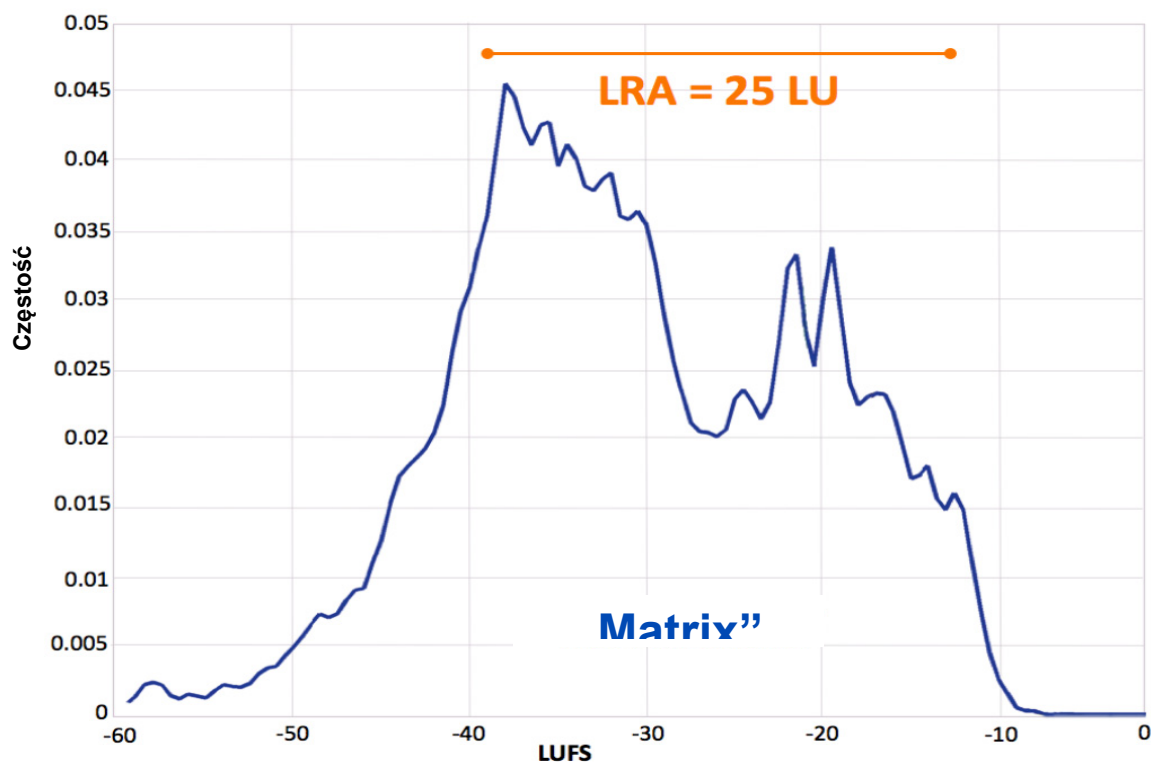


**Zakres głośności jest parametrem ogólnym, pomagającym w podejmowaniu decyzji o stosowaniu kompresji dynamiki.**

Pierwsze doświadczenia pochodzące od nadawców sugerują przyjęcie maksymalnej wartości LRA równej około **20 LU** w materiałach o dużej dynamice, takich jak filmy akcji czy muzyka klasyczna. Większość programów nigdy nie będzie korzystała z tak dużej wartości LRA lub raczej nie będzie w stanie jej osiągnąć!



Dla programów bardzo krótkich (<30 s) takich jak reklamy, promocje czy trailery, lepszym sposobem kontroli własności dynamicznych sygnału będzie ustanowienie limitu wartości **maksymalnych** dla Poziomu głośności **Chwilowego** lub **Krótkotrwałego**<sup>3</sup> tworząc w ten sposób niejako „drugą linię obrony” (patrz punkty 7 i 10).



Rysunek 3: Zakres głośności (LRA) jako wynik statystycznego rozkładu poziomów głośności

Na rysunku 3 pokazano rozkład głośności oraz wartość LRA w ścieżce dźwiękowej filmu Matrix, podana tu wartość 25 LU jest prawdopodobnie zbyt wielkim wyzwaniem dla większości domowych salonów...

### 2.3 Poziom rzeczywisty szczytowy (TPL) i Maksymalny dozwolony TPL

W Europie najbardziej rozpowszechnionym miernikiem poziomu był (i jest wciąż bardzo popularny) miernik wartości quasi-szczytowej (Quasi Peak Programme Meter (QPPM, o czasie uśredniania = 10 ms). Po przejściu do obróbki sygnałów cyfrowych pojawiły się mierniki wartości szczytowej próbek (sample peak meters). Miernik QPPM nie jest w stanie pokazywać krótkich szczytów sygnału (<<10 ms), co wynika z jego konstrukcji, podobnie miernik szczytów próbek może nie pokazywać rzeczywistych poziomów szczytowych.

Obróbka cyfrowa lub kodowanie stratne mogą wprowadzać **szczyty międzypróbekowe**, wykraczające poza wskazywany poziom szczytów próbek. Dla nadawców ważne jest wskazywanie rzeczywistych poziomów sygnałów obecnych na różnych platformach i mających różne częstotliwości próbkowania. Taki miernik powinien wskazywać obcinanie sygnału, nawet jeśli jego wartości szczytowe leżą pomiędzy próbkami. Dzięki temu można **przewidzieć i uniknąć zniekształceń**, jakie mogą się pojawić w następnych elementach toru takich jak przetworniki D-A, konwertery częstotliwości próbkowania lub powszechnie teraz używane kodeki. Miernik szczytów próbek nie może tego robić a zatem jest niewystarczający dla potrzeb współczesnych nadawców (patrz Lund, Th.: 'Stop counting samples' [9]).

<sup>3</sup> Maksymalny poziom głośności chwilowej (Maximum Momentary Loudness Level (Max ML)) jest największą wartością (mierzoną w LUFS) poziomu głośności chwilowej sygnału audio (przy czasie uśredniania 400 ms). Maksymalny poziom głośności krótkotrwałej (Maximum Short-term Loudness Level (Max SL)) jest najwyższą wartością (w LUFS) poziomu głośności krótkotrwałej (przy czasie uśredniania 3 s).

**Poziom rzeczywisty szczytowy (true peak level)** wskazuje maksymalną (dodatnią lub ujemną) wartość sygnału w ciągłej dziedzinie czasu, wartość ta może być większa niż największa wartość próbki w dziedzinie momentów czasu w których są pobierane próbki. Za pomocą miernika rzeczywistych szczytów z nadpróbkowaniem, zgodnego z ITU-R BS.1770, można teraz wykrywać takie prawdziwe szczyty (symbol jednostki, zgodny z ITU-R BS.1770, ma postać: **dBTP** - decyBel odniesiony do pełnej skali cyfrowej mierzony za pomocą miernika rzeczywistych szczytów). Dokładność miernika zależy od częstotliwości nadpróbkowania. Niezbędne jest tylko pozostawienie zapasu 1 dB poniżej poziomu 0 dBFS, aby wciąż można było radzić sobie z potencjalnym zaniżeniem wskazania o mniej więcej 0,5 dB (dla miernika rzeczywistych szczytów z 4-krotnym nadpróbkowaniem, przy podstawowej częstotliwości próbkowania równej 48 kHz).



**Mierniki wartości szczytowej z nadpróbkowaniem umożliwiają prawidłową ocenę rzeczywistej wartości szczytowej sygnałów audio. Mierniki mierzące szczyty próbek - nie.**

Zatem **Maksymalny dozwolony poziom rzeczywistego szczytu (Maximum Permitted True Peak Level)** zalecany w dokumencie R 128 wynosi:

**-1 dBTP**

Stosuje się tę wartość w środowisku produkcyjnym dla powszechnie używanych liniowych sygnałów audio. Należy zauważyć, że niektóre miejsca toru, takie jak analogowe nadajniki retransmisyjne oraz powszechnie używane kodeki redukujące strumień danych, wymagają mniejszego poziomu rzeczywistej wartości szczytowej. Dokument EBU zatytułowany „*Distribution Guidelines*” (*Wskazówki dotyczące dystrybucji*) (Dokument Techniczny EBU 3344) zawiera wyczerpujące omówienie tego zagadnienia.

### **Podsumowanie Zalecenia R 128 EBU**

- Sygnał audio określają parametry „**Głośność Programu**”, „**Zakres Głośności**” oraz „**Maksymalny poziom rzeczywistej wartości szczytowej**”;
- **Poziom Głośności Programu** powinien być znormalizowany względem wartości **-23.0 LUFS**;
- Tolerancja pomiaru zazwyczaj jest równa  $\pm 1.0$  LU dla programów, w których praktycznie nie można osiągnąć dokładnej normalizacji;
- Pomiary powinny być wykonywane miernikiem zgodnym z wymaganiami podanymi w dokumentach **BS.1770 ITU-R** oraz **Dokument Techniczny 3341 EBU** („Tryb EBU” - również definiującymi metodę bramkowania);
- Parametr **Zakres Głośności** powinien być pomocny w podejmowaniu decyzji odnośnie konieczności zastosowania kompresji dynamiki (zależnie od rodzaju materiału, docelowych słuchaczy oraz platformy transmisji);
- **Maksymalny Dozwolony Poziom Rzeczywistych Szczytów** w produkcji wynosi **-1 dBTP**;
- **Metadane** głośności powinny być ustawiane tak, aby wskazywały wartość **-23 LUFS** (dla programów, które zostały znormalizowane do takiego poziomu a co jest zalecane); metadane głośności powinny zawsze wskazywać **prawidłową** wartość głośności programu, nawet jeśli z jakiegoś powodu program może nie być znormalizowany do -23 LUFS.



## 2.4 Logo R 128

EBU wprowadziło oficjalne logo dla Zalecenia R 128, logo składa się z liczb 1, 2 oraz 8 - które tworzą szczęśliwą i roześmianą buzię:

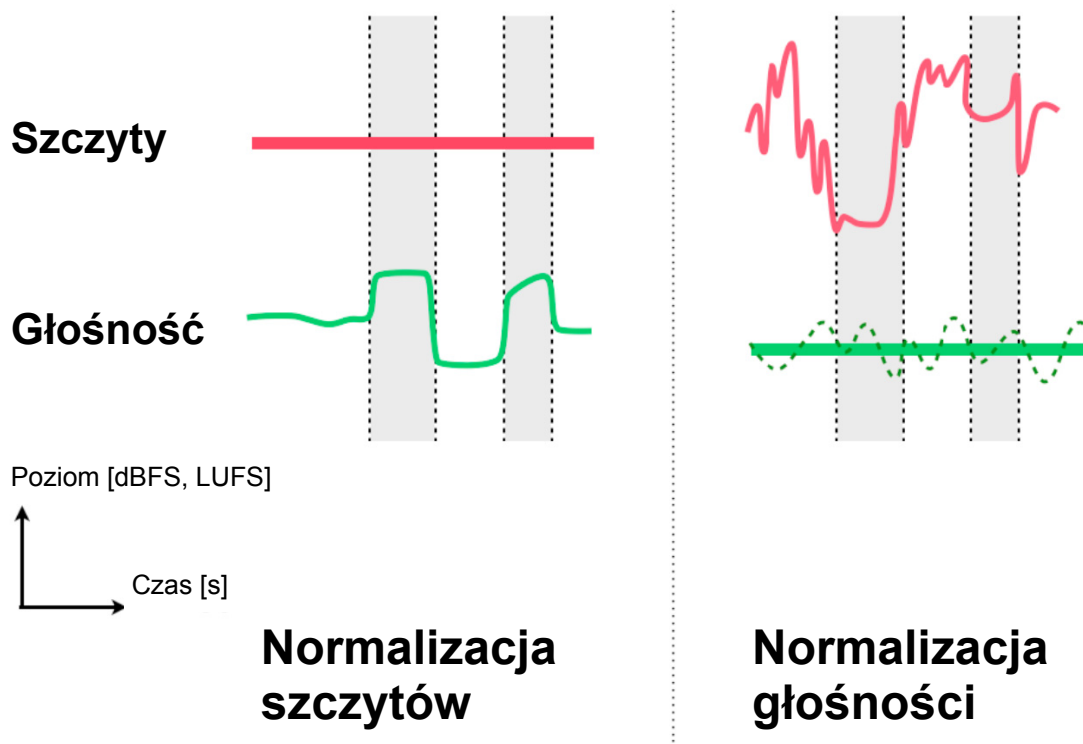


Takie logo może być używane (pod pewnymi warunkami) przez producentów w celu sygnalizowania zgodności urządzenia z pomiarem typu „Tryb EBU”.

## 3. Ogólna koncepcja normalizacji głośności

### 3.1 Wartości szczytowe a głośność

Rozpowszechniona szeroko koncepcja regulacji poziomu audio polegająca na *normalizacji szczytów* i odniesienia ich do dozwolonego poziomu maksymalnego (Permitted Maximum Level (PML, na przykład -9 dBFS), doprowadziła do ujednoczenia poziomów szczytowych programów, ale przy bardzo zmiennym poziomie ich głośności. Rzeczywiste zmiany tego poziomu zależą od stopnia kompresji dynamiki sygnału. Dla kontrastu - **normalizacja głośności zapewnia wyrównaną średnią głośność programów zawierających szczyty o zmiennej wartości, zależnie od treści programu jak i wymagań artystycznych czy technicznych (patrz Rysunek 4). Jeśli zakres głośności programu mieści się w dozwolonym przedziale tolerancji, to słuchacz może cieszyć się ujednoczonym poziomem średniej głośności we wszystkich programach, eliminując w ten sposób częste korzystanie z pilota do regulacji głośności.**



Rysunek 4: Normalizacja poziomu szczytowego a normalizacja poziomu głośności dla kilku programów

I znowu, to wszystko **NIE** oznacza, że w *samej programie* poziom głośności musi być stały, przeciwnie! Również **NIE** oznacza to, że *poszczególne składniki programu* (na przykład premiksy, miksy źródłowe, wersje typu muzyka i efekty lub odizolowana ścieżka komentarza) muszą mieć ten sam poziom głośności! Zmienna głośność to przecież narzędzie artysty a koncepcja normalizacji głośności, zgodnie z Zaleceniem R 128 tak naprawdę **zachęca do bardziej dynamicznego miksowania!** Normalizowana jest tu **średnia, scałkowana głośność całego programu.**

### 3.2 Normalizacja sygnału a Metadane

Zasadniczo istnieją dwa sposoby realizacji normalizacji głośności dla użytkownika - jeden to rzeczywista **normalizacja sygnału audio** jako takiego, wówczas programy stają się jednakowo głośne z definicji - drugi to **skorzystanie z metadanych głośności**, które opisują głośność danego programu. W tym ostatnim przypadku, rzeczywiste poziomy uśrednionej głośności programu nie muszą być sprowadzane do wartości znormalizowanej i mogą ulegać nawet znacznym zmianom w każdym programie. Dla posiadaczy najnowszych urządzeń oznacza to, że normalizacja może być wykonywana u słuchacza/widza dzięki wykorzystaniu opisujących głośność wartości metadanych które zmieniają wzmocnienie dźwięku w programach i wyrównują poziomy do jednego poziomu odstuchu.



**Jednolitą głośność można osiągnąć normalizując sygnał audio lub wykorzystując metadane głośności.**

W ramach zasad regulacji poziomów głośności zgodnych z EBU R 128 zachęca się do korzystania z *pierwszego* rozwiązania, które ma podane niżej zalety:

- prostota oraz
- potencjalne podwyższenie jakości sygnału audio.

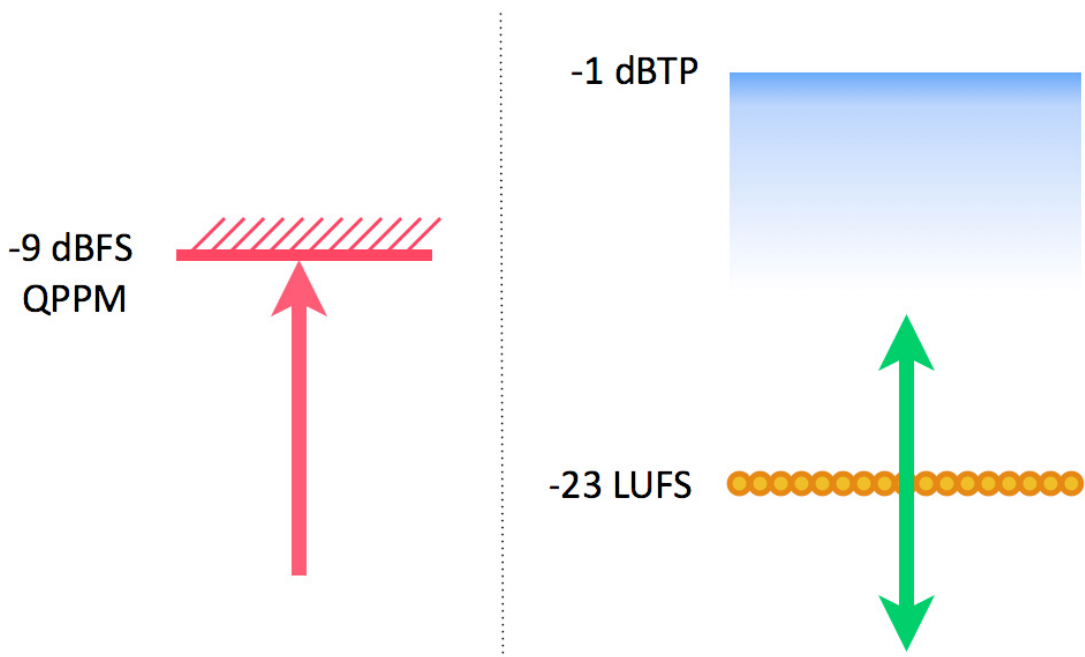
Drugie rozwiązanie nie jest zabronione (patrz również „*Distribution Guidelines*” /wskazówki odnośnie dystrybucji/ - dokument EBU o nazwie Dokument Techniczny 3344), zaletą jest tutaj posługiwanie się tylko jedną liczbą (-23 LUFS), co ma ogromne znaczenie w rozpowszechnianiu koncepcji regulacji poziomów głośności, ponieważ łatwo jest to zrozumieć oraz stosować. Aktywna normalizacja poziomu źródła w pewien sposób „karze” sygnały zbyt mocno skompromowane, zatem automatycznie zachęca osoby zajmujące się produkcją do stosowania innych, bardziej dynamicznych i twórczych sposobów wpływu na ich program. Innymi słowy, rzeczywista zmiana *techniczna* poziomu sygnału audio w wyniku **aktywnej normalizacji do poziomu -23 LUFS** ma bezpośredni wpływ na proces *artystyczny* i to - wpływ dodatni! Produkcja jest dzięki temu uwolniona od udziału w „wojnie o głośność” - niefortunnym i rozpowszechnionym rezultacie koncepcji normalizacji szczytów.

Tak czy inaczej, należy tu stwierdzić, że obie wspomniane metody mogą się nawzajem uzupełniać, nie można na nie patrzeć jak na metody przeciwne lub też traktować w kategoriach „czarne lub białe” realizacje tego samego zagadnienia. Obie koncepcje są częścią dokumentu R 128 - z uwagi jednak na podane wyżej zalety, zaleca się wykonywanie procesu **normalizacji sygnałów audio**.



**Zaleca się normalizację głośności sygnału audio podczas produkcji z uwagi na prostotę tego procesu oraz potencjalne podwyższenie jakości.**

Działanie w kierunku wprowadzenia ujednoczonego poziomu głośności oznacza przyjęcie **całkowicie nowej** koncepcji miksowania, regulacji poziomów i generalnie pracy z dźwiękiem. W przypadku limitera szczytów sygnału, ustawionego na dopuszczalny poziom maksymalny (zazwyczaj jest to -9 dBFS, mierzone miernikiem QPPM, quasi-szczytowym), zapewniającego tym samym coś w rodzaju „*bezpiecznego pułapu*”, dzięki któremu niezależnie od tego, jak duży szczyt się pojawi, sygnał zawsze będzie sprowadzony do „*prawidłowego*” poziomu maksymalnego. Natomiast skorzystanie z zasad regulacji poziomu głośności bardziej przypomina swobodne „*unoszenie się w przestrzeni, pod otwartym niebem*” (patrz Rysunek 5).



Rysunek 5: Normalizacja poziomu oparta na wartości quasi-szczytowej („bezpieczny pułap”) a normalizacja poziomu głośności

Przy pomiarze i normalizacji głośności bezpieczny znika bezpieczny pułap ograniczający wartości szczytowe sygnału. Niektórych to może wręcz przerażać, ponieważ poprzednio nie trzeba było śledzić tak starannie poziomu, co było w pewien sposób „wygodne” - umieszczony na końcu toru limiter sprawiał, że poziom był zawsze skutecznie „poskramiany” na wyjściu. Efektem ubocznym takiej metody było podniesienie głośności, ponieważ bardziej skomplikowane procesory w trosce o zmniejszenie niekorzystnych efektów limitowania wprowadzały kompresję dynamiki.

Z drugiej strony regulacja poziomu głośności zachęca do wykorzystywania z pewnością najlepszego przyrządu pomiarowego jakim jest ludzkie ucho. To z kolei implikuje bardziej uważne miksowanie i poprawia jakość oferty dźwiękowej. Doświadczenia kilku członków EBU wykazały, że praca oparta na regulacji głośności ma charakter **uwalniający** i satysfakcjonujący. Znika walka o to „kto jest najgłośniejszy”, obniżają się poziomy ogólne, co w połączeniu z wyższym maksymalnym dozwolonym poziomem rzeczywistej wartości szczytowej (-1 dBTP), zapewnia potencjalnie **bardziej dynamiczne miksy**, charakteryzujące się **większym ujednoczeniem głośności** w danym programie. Kompresja dynamiki wraca do roli narzędzia artystycznego a przestaje być orężem walki o to, aby było możliwie najgłośniejsze - a zatem **wzrasta jakość oferty przekazu audio!**

Sprowadzenie z powrotem na właściwe tory miksowania za pomocą uszu to jak ulga po długim wyczekiwaniu. Osoba miksująca jest teraz zachęcana do miksowania z użyciem tylko własnych uszu (kolejny wpływ mierzenia głośności) - po ustawieniu poziomów i ustalonej głośności odsłuchu w studio (patrz punkt 8).



**Regulacja poziomu głośności zachęca do miksowania „tylko za pomocą uszu” - po ustawieniu poziomów oraz ustalonej głośności odsłuchu w studio.**

W dalszych etapach procesu produkcji nadawca musi sprostać konieczności normalizacji poziomu różnych materiałów pochodzących z różnych miejsc. Zwłaszcza podczas okresu przejściowego wciąż będziemy mieć

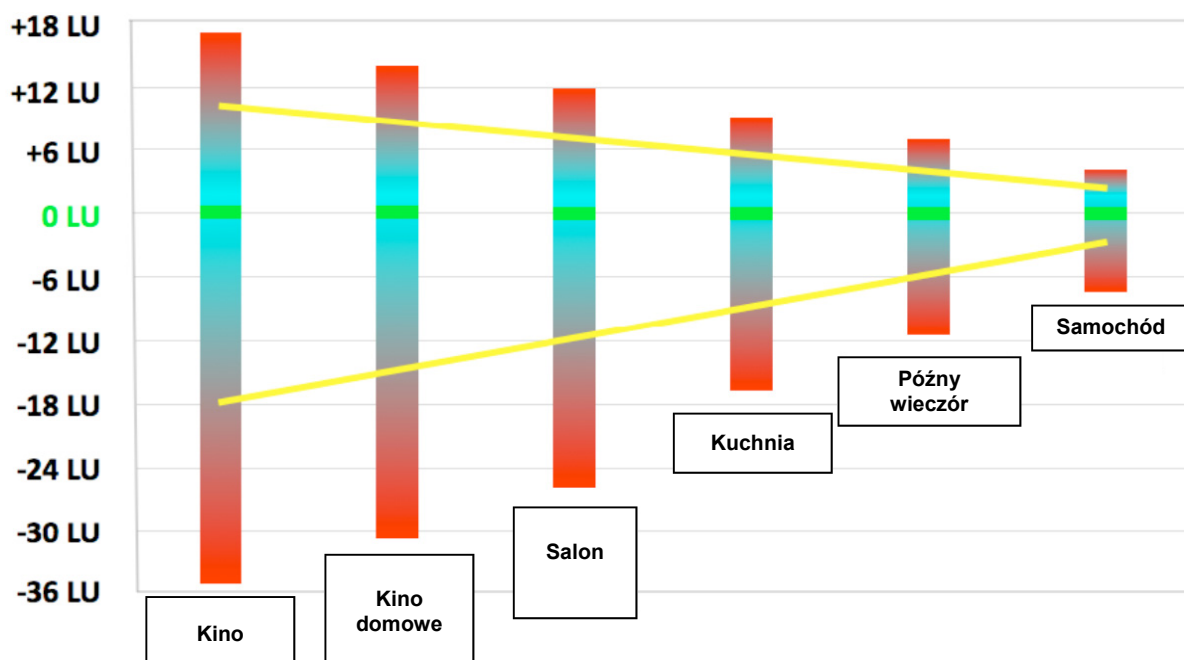
do czynienia z wieloma programami, które nie były jeszcze poddane procesowi normalizacji głośności. Należy opracować odpowiednie strategie odnośnie tych programów, takiej jak bezpośrednia **automatyczna normalizacja** tuż po wprowadzeniu do serwera odczytu lub zainstalowanie na wyjściu głównym toru urządzenia zabezpieczającego, zdolnego do regulacji głośności, będącego w stanie na przykład obsłużyć łącza na żywo, których materiał nie jest tworzony przy poziomie docelowym **-23 LUFS**.

Tego typu zagadnienia zostaną jeszcze dokładniej omówione w dalszych punktach (4.3, 6).

### 3.3 Używanie parametru o nazwie Zakres Głośności

Po raz pierwszy można teraz określać ilościowo dynamikę programu. W przeszłości było to często „zgadywanką dla wykształconych”, dla doświadczonego personelu audio, decydującego o tym, czy dany program będzie pasował do progów tolerancji głośności odbiorców, dla których jest przeznaczony. Dysponując **Zakresem Głośności - Loudness Range (LRA)** - mamy teraz tę zgadywankę z głowy. Uzyskany na końcu okresu pomiaru (zazwyczaj na końcu programu) liczbowy wynik umożliwia mikserowi/operatorowi zdecydować, czy konieczna jest dalsza obróbka dynamiki.

Ważne jest zrozumienie tego, że nie jest możliwe określenie jednej maksymalnej wartości LRA dla wszystkich nadawców i wszystkich programów. **Konkretna wartość maksymalna LRA** zależy od rodzaju (rodzajów) programu (kanały tematyczne o bardzo jednolitej treści, takie jak wiadomości, na pewno nie będą musiały mieć równie wysokiej wartości LRA z jakiej korzystają nadawcy publiczni prezentujący bardzo zróżnicowane programy, na przykład koncerty muzyki klasycznej). Maksymalna wartość LRA może zależeć od platformy dystrybucji, na przykład platformy mobilne oraz od miejsca w którym program jest odtwarzany (*patrz Rysunek 6; odległość pomiędzy żółtymi liniami wskazuje przykładowe różne wartości zakresu głośności*). Wybierając maksymalną wartość LRA dla specyficznych programów nadawca kieruje się środowiskiem odsłuchowe, wiekiem docelowej grupy odbiorców, rodzajem „strefy wygodnego odsłuchu” u odbiorcy i innymi parametrami. **Koncepcja kontroli zakresu głośności** zaczyna się od ogólnie akceptowanej maksymalnej wartości zakresu głośności, zgodnej z zasadami opisanymi powyżej a następnie dostosowywania tej wartości w procesie produkcji w taki sposób, aby spełnić wymagania techniczne poszczególnych platform dystrybucji i środowisk odsłuchowych.



Rysunek 6: Przykłady różnych zakresów głośności w zależności od środowiska odsłuchowego

Za określenie wartości LRA dla programu odpowiedzialna jest osoba miksująca i to na etapie produkcji programu, osobie tej zaleca się przestrzegać podanych tu zasad. W środowiskach, gdzie nie jest możliwa kontrola człowieka i jego interwencja, pomiar LRA pomaga w doborze odpowiednich presetów procesora dynamiki, który został skonfigurowany odpowiednio do typu programu i sygnału audio.

Niemniej jednak zaleca się dążyć do sytuacji, w których to inżynier dźwięku ma wpływ na wartość LRA, zgodną z wymaganiami nadawcy, ponieważ takie rozwiązanie potencjalnie podnosi jakość przekazu dźwiękowego.

Aby uwzględnić zapotrzebowanie na różne wartości zakresów głośności, Zalecenie EBU R 128 nie zawiera maksymalnej dopuszczalnej wartości LRA, zamiast tego usilnie zachęca do korzystania z parametru Zakresu Głośności podczas oceny potencjalnych potrzeb obróbki dynamiki danego sygnału zgodnie z różnymi kryteriami wymienionymi uprzednio.

Zakres Głośności jest również przydatnym **wskaźnikiem** możliwych procesów *redukcji* dynamiki, jakie miały miejsce w torze sygnałowym, procesów wykonanych celowo lub przypadkowo. Jeśli na przykład wartość LRA programu po jego przejściu przez procesor jest niższa niż przed przejściem, to znaczy, że wykonano proces redukcji dynamiki.

### 3.4      *Wspinając się na Rzeczywisty Szczyt*

Trzeci parametr, zalecany przez R 128, dotyczy **maksymalnego poziomu rzeczywistej wartości szczytowej** sygnału audio. Pominąwszy na chwilę reguły normalizacji wartości szczytowych, musimy pamiętać, że w celu uniknięcia przesterowań i zniekształceń wciąż jest **bardzo ważne mierzenie i kontrola szczytów w sygnale programu** a zwłaszcza w zakresie szczytów maksymalnych.

Miernik głośności zgodny z „trybem EBU” (patrz Dokument Techniczny 3341 EBU) wykonuje również pomiar i pokazuje poziomy rzeczywistych wartości szczytowych programu. Limitery zabezpieczające przed przemodulowaniem muszą pracować w trybie *rzeczywistych szczytów (true-peak mode)* i muszą być doregulowywane do odpowiedniego maksymalnego dopuszczalnego poziomu rzeczywistego szczytu, zarówno w produkcji, jak również na wyjściu regulatora master, na stacji dystrybucyjnej oraz w nadajniku. W dokumencie („Wskazówki odnośnie dystrybucji” - „Distribution Guidelines” EBU, Dokument Techniczny 3344) oprócz wartości maksymalnego poziomu rzeczywistego szczytu dla standardowych sygnałów PCM stosowanych podczas produkcji (-1 dBTP) poddano dalsze sugestie odnośnie różnych zastosowań.

## 4. Strategie dla Normalizacji Głośności

### 4.1 Produkcja i post-produkcja

Zastosowanie normalizacji głośności w tych obszarach oferuje dwie możliwości: pierwsza to zachowanie aktualnych metod regulacji poziomów i wykonanie potem dodatkowego przesunięcia poziomu, druga to zmiana dotychczasowych nawyków regulacji poziomów w kierunku **sterowania głośnością** oraz jej **normalizacji** uzupełnianej potem nieznaczną lub zerową korekcją poziomu (Rysunek 7).

Pierwsze podejście jest bardziej odpowiednie na wczesnych etapach przechodzenia i być może przyda się zwłaszcza tym, którzy pracują z **programami na żywo**. Zachowane zostaną używane do tej pory mierniki, limityery oraz praktyki miksowania a korekcja poziomu będzie wykonywana na wyjściu konsoly (po miernikach głównych) w celu osiągnięcia **Docelowego Poziomu** głośności równego **-23 LUFS**.



**Rysunek 7:** Dwie zasadnicze metody pracy zapewniające ujednoczenie głośności w produkcji i post-produkcji

Miernik głośności jest umieszczony za układem przesuwania poziomu w celu dokładnego pokazania korekcji (co na początku jest wciąż rodzajem zgadywanki). Dobrym pomysłem jest zawsze zastosowanie **miernika głośności** pracującego **równolegle** z miernikiem konwencjonalnym. Umożliwia to zebranie przydatnych doświadczeń przed prawdziwym zanurkowaniem do głębokiej wody świata regulacji poziomu. Co więcej, zastosowanie miernika głośności w celu pomiaru starych programów tego samego typu daje dobre wskazówki odnośnie miejsc wpływających na poziom.

Dla programów wykończanych w **post-produkcji** wymagane przesunięcie poziomu, stosowane w podejściu **1**, jest łatwe do wykonania. Po pomiarze całego programu w jednym przejściu można dokładnie wyznaczyć niezbędną zmianę wzmocnienia, co w dzisiejszym, „opartym na plikach” świecie, jest operacją szybką i łatwą.

Oczywiście w przypadku **programów na żywo** wyzwaniem jest (jeśli nie kwestią szczęścia) dokładne wyznaczenie Poziomu Docelowego. Dlatego dla programów, dla których dokładna normalizacja do Poziomu Docelowego wynoszącego **-23 LUFS** nie jest praktycznie możliwa (poza programami na żywo są to na przykład te, które mają nadzwyczaj krótki czas) - akceptowalna jest odchyłka rzędu **±1.0 LU**. Wczesne doświadczenia NDR (Nadawca w północnych Niemczech), ORF (Nadawca w Austrii) oraz RTBF (Nadawca w Belgii - oddział francuskojęzyczny) wykazały, że w przypadku miksów na żywo naprawdę jest możliwe ich umieszczenie wewnątrz okna tolerancji **±1 LU** dopuszczalnego przez EBU R 128.





**Dopuszczalne są odchyłki  $\pm 1$  LU względem poziomu docelowego (-23 LUFS). Oczekuje się, że w specjalnych przypadkach, pojawią się wyjątki od tej reguły.**

W tych przypadkach, gdzie poziomy poszczególnych sygnałów programu są w większości *nieprzewidywalne* lub program zawiera tylko elementy tła (na przykład podkład muzyczny dla prognozy pogody) lub wymagania dramaturgii programu wymuszają poziom głośności niższy niż pożądany poziom docelowy, taka tolerancja może okazać się za wąska. Przewiduje się więc, że w takich właśnie przypadkach, uśredniony poziom głośności może jednak wypaść poza granicami podanymi w Dokumencie R 128.

W **rozwiązaniu 1 regulacji poziomu** (zachowanie aktualnych praktyk regulacji poziomu) prawdopodobnie niemal we wszystkich przypadkach wymagana zmiana wzmocnienia będzie *ujemna* (tłumienie). Dlatego dodatkowy etap redukcji zakresu dynamiki i/lub ograniczenia maksymalnego poziomu rzeczywistej wartości szczytowej będzie zazwyczaj zbędny. Potencjalne tłumienie, konieczne w ogromnej większości przypadków, jest również powodem, dla którego opisane w punkcie 3.2 rozwiązanie oparte o metadane nie jest zalecane dla rozwiązania 1.

**Rozwiązanie 2 regulacji poziomu** (przejsięcie od razu na normalizację głośności) jest rozwiązaniem **zalecanym** w tym dokumencie (Praktyczne Wskazówki). W tym przypadku, po początkowych pomiarach i okresie testowania starych programów oraz po instalacji miernika głośności równoległe z miernikiem (zazwyczaj QPPM) do tej pory używanym, zalety regulacji poziomu głośności mówią same za siebie. Możliwy teraz większy zakres dynamiki świetnie przydaje się przy dźwiękach tłumów ludzi, na przykład w programach sportowych, co zwiększa oddziaływanie imprezy sportowej na widzów i słuchaczy. Komentarze studyjne, które często mają skompresowaną dynamikę wynikającą z potrzeb artystycznych (i dlatego niższy jest stosunek głośności do wartości szczytowej) - będą teraz bardziej wyrównane i dopasowane do lepszego, dynamicznego dźwięku zarejestrowanego w miejscu imprezy itd.

Na następnych stronach zostanie omówiony wpływ korzystania z mierników głośności w produkcji i post-produkcji.



## 4.2 Pomiar głośności w produkcji i post-produkcji

W „Trybie EBU” miernik głośności, zdefiniowany w Dokumencie Technicznym 3341 EBU, może mieć 3 różne skale czasowe:

- Głośność Chwilowa (**Momentary Loudness**, skrótowo „**M**”) - okno czasowe: **400 ms**
- Głośność Krótkotrwała (**Short-term Loudness**, skrótowo „**S**”) - okno czasowe: **3 s**
- Głośność uśredniona (**Integrated Loudness**, skrótowo „**I**”) - od „startu” do „stopu”

Okna czasowe **M** oraz **S**<sup>4</sup> są przeznaczone do stosowania w natychmiastowych regulacjach poziomu oraz miksach sygnałów audio. Początkowe ustawienie poziomu może być wykonane najlepiej za pomocą miernika głośności chwilowej, poprzez regulację poziomu kluczowych lub bazowych elementów (takich jak głos, muzyka lub efekty dźwiękowe) w taki sposób, aby wypadły w pobliżu Poziomu Docelowego równego **-23 LUFS**. Oczywiście osoba miksująca musi wiedzieć w każdej chwili, jaka jest głośność rzeczywistego sygnału, co właśnie jest głównym powodem stosowania pomiarów chwilowych i krótkoterminowych.

Wskutek niespójności pomiędzy dokumentami ITU-R BS.1770 a ITU-R BS.1771, w Dokumencie Technicznym 3341 EBU zasugerowało inną konwencję nazewnictwa, zgodną z normą ISO 80000-8:

- Symbol dla „Poziomu Głośności, Ważonego K” („Loudness Level, K-weighted”) powinien mieć postać „**L<sub>K</sub>**”.
- Symbol jednostki „**LUFS**” wskazuje wartość **L<sub>K</sub>** odniesioną do pełnej skali cyfrowej.
- Symbol jednostki „**LU**” wskazuje wartość **L<sub>K</sub>** bez bezpośredniej bezwzględnej wartości odniesienia a zatem opisuje różnice poziomów głośności.

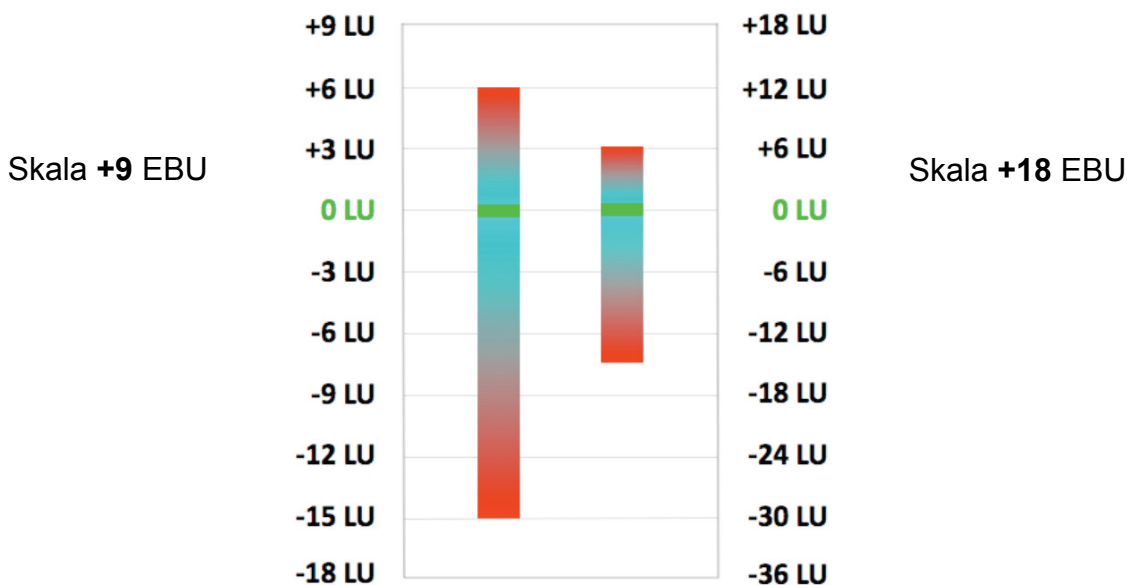
Ostatecznie nie podano specyfikacji szczegółów dotyczących interfejsów graficznych lub interfejsów użytkownika dla miernika głośności spełniającego wymagania „Trybu EBU”, niemniej jednak zdefiniowano dwie skale: „Skalę +9 EBU” („**EBU +9 Scale**”), która powinna nadawać się dla większości programów oraz „Skalę +18 EBU” („**EBU +18 Scale**”), która może być potrzebna dla programów o dużej rozpiętości wartości LRA. Obie skale mogą pokazywać względny poziom głośności w jednostkach **LU**, lub poziom bezwzględny w jednostkach **LUFS**. Przy czym „**0 LU**” w „trybie EBU” jest równe poziomowi docelowemu **-23 LUFS**. Producenci mierników w grupie PLOUD zgodzili się na wprowadzenie zestawu parametrów dla „Trybu EBU”, co sprawi, że odczyty ich mierników będą ujednolicone.



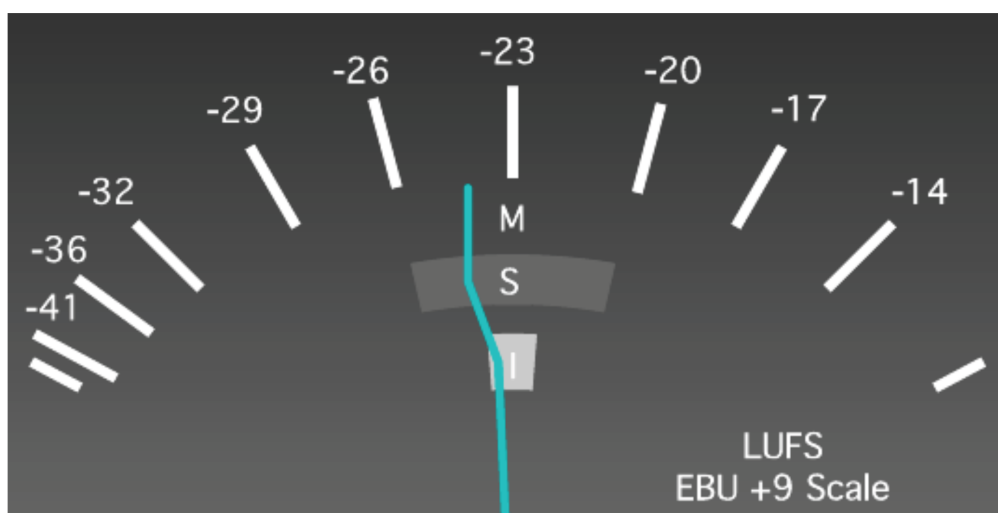
**W „Trybie EBU” miernika głośności 0 LU jest równe -23 LUFS.**

Wielu innych producentów również przyjęło „Tryb EBU” lub właśnie ten tryb przyjmuje. Na Rysunku 8 pokazano schematycznie wygląd skali miernika (prążki typu bar-graph) z dwoma skalami trybów EBU, na Rysunku 9 pokazano, jak mógłby wyglądać miernik oparty na oprogramowaniu wskazujący wyniki za pomocą „strzałki”.

<sup>4</sup> **M** oraz **S** są od dawna powszechnie stosowane w stereofonii i oznaczają odpowiednio środek, **Mid** oraz strona (boki) - **Side**. Aby dla wartości chwilowej (**Momentary**) oraz krótkoterminowej (**Short-term**) oznaczenia czasów integracji były inne, można stosować postaci **ML<sub>K</sub>** oraz **SL<sub>K</sub>** (jak również **IL<sub>K</sub>**). Symbol **L<sub>K</sub>** oznacza poziom głośności, ważony K (**Loudness Level, K-weighted**) i jest zgodny z międzynarodową normą ISO 80000-8 dotyczącą nazewnictwa.



Rysunek 8: Schematyczna prezentacja dwóch skal głośności (tutaj w LU) opisanych w Dokumencie Technicznym 3341 EBU



Rysunek 9: Schematyczna prezentacja proponowanego miernika głośności ze „zgiętą strzałką”

### 4.3 „Gotowi, ustawieni (poziomy), START!”

Zaleca się na początku **ustawiać poziomy** z zachowaniem pewnej **ostrożności** - psychologicznie łatwiej jest stopniowo zwiększać poziom uśrednianej głośności podczas miksu niż go zmniejszać. Ma na to wpływ również funkcja bramkowania: dla zmniejszenia poziomu średniego potrzebne są większe zmiany poziomu niż dla jego zwiększenia. Nieznaczny wzrost poziomu w trakcie trwania programu jest również bardziej naturalny z punktu widzenia jego dramaturgii. Taka od początku defensywna strategia zapewnia inżynierowi dźwięku swobodę manewru w przypadku nieoczekiwanych czy nieprzewidywalnych sygnałów i zdarzeń.

Po ustawieniu poziomów poszczególnych sygnałów i wprowadzeniu ustalonego wzmocnienia w torze monitora odsłuchu (patrz punkt 8), inżynier dźwięku może przejść do **miksowania tylko za pomocą** uszu. Obserwacja chwilowego lub krótkoterminowego poziomu głośności oraz od czasu do czasu popatrzenie na wartość głośności uśrednionej powinny upewnić operatora, że miks zmierza prawidłowo do Poziomu Docelowego. Dzięki obserwacji wartości liczbowej 'l' z dokładnością do jednego miejsca po przecinku lub graficznej postaci tej wartości w podobnej rozdzielczości, można przewidzieć w **jakim kierunku zmierza sygnał** i podjąć stosowne środki. Należy to robić płynnie, ponieważ zbyt drastyczne zmiany w większości przypadków nie są zadowalające artystycznie.

Przy **Maksymalnym Dozwolonym Poziomie Rzeczywistej Wartości Szczytowej** równym **-1 dBTP**, mniej prawdopodobne jest pojawienie się zjawiska „uderzenia w ścianę” (czyli limitera zabezpieczającego pracującego przy poziomie -9 dBFS). Rozsądne i świadome wykorzystanie limitera sprawia, że ten mechanizm „*otwierania pokrywy*” połączony z normalizacją głośności do poziomu -23 LUFS daje **bardziej dynamiczne miksy**, o mniejszych szkodliwych efektach ubocznych kompresji dynamiki typu „pompowanie” a zatem **wzrasta jakość dźwięku!** Twórcy programów, którzy dawniej faworyzowali miksy dynamiczne są teraz uwolnieni od potencjalnych kompromisów, ponieważ ich program brzmi bardziej miękko niż program z większą kompresją. Po zastosowaniu normalizacji głośności taki kompromis znika. Nareszcie!

*Elementy miksu, najważniejsze dla subiektywnego wrażenia jednolitej głośności są zwane dźwiękami „przedniego planu” - wokal, głos, muzyka lub najważniejsze efekty dźwiękowe. Poszczególne składniki dźwiękowe wykazują duże różnice pomiędzy poziomem głośności a poziomem wartości szczytowych. Na przykład „brzęk” dwóch kieliszków podczas spełniania toastu ma wysoki poziom szczytowy, ale zupełnie niski poziom głośności. Z drugiej strony, dynamicznie skompresowany i twardy „riff” gitary rockowej ma poziom głośności niemal równy poziomowi szczytowemu! Jeśli teraz te dwa sygnały zestawisz według ich szczytów to riff gitarowy będzie dużo głośniejszy niż brzęk kieliszków. Ten przykład tylko ilustruje samą zasadę, a NIE oznacza, że takie dwa sygnały muszą być koniecznie miksowane w oparciu o ich głośność! Poziom poszczególnych elementów czy składników (takich jak pre-miksy, miksy-źródłowe, miksy tylko muzyki czy ścieżka komentarza) w miksie finalnym jest kwestią decyzji artystycznej, ale pomiar głośności może pomóc osobie miksującej oferując jej wizualne wskazania, które tak naprawdę pokazują to, co osoba miksująca po prostu słyszy!*

Wracając do pomiarów, po dojściu do końca programu możliwe są dwa scenariusze:

- dojście dokładnie do poziomu docelowego (-23.0 LUFS) lub
- dojście do wartości mniejszej lub większej niż poziom docelowy

Dla **produkcji na żywo** bardziej prawdopodobnym jest scenariusz drugi. Jeśli rzeczywisty poziom głośności wypada w granicach akceptowalnej tolerancji  $\pm 1.0$  LU, to nie trzeba już nic więcej robić. Jeśli wskutek wyjątkowo nieprzewidywalnej natury programu lub rzadkiego pojawiania się elementów „przedniego planu” poziom wypada poza tymi granicami, jest to wciąż do przyjęcia z ogólnego punktu widzenia produkcji (jak to podano wcześniej). W dalszej części toru można stosować pewne środki zaradcze, w celu „ujarzmienia” takich przypadków, mogą to być **procesory głośności**, które stopniowo regulują poziom uśrednionej głośności takich programów w dyskretny sposób, pełniąc rolę swego rodzaju „zabezpieczenia głośności”. Efekt taki powinien być uzyskiwany przy odpowiednio wolnym czasie reakcji, tak aby właściwa wewnętrzna dynamika sygnału nie została naruszona. Gdy chodzi o poszczególne presety takich procesorów głośności lub miejsca ich zainstalowania w torze sygnałowym pomocne będzie **rozdzielanie** programów na żywo od programów opartych na plikach. Procesor może być potrzebny tylko dla programów na żywo, jeśli obróbka i przepływ programów opartych na plikach spełnia już całkowicie wymagania EBU R 128. Jeśli procesor dynamiki i głośności znajduje się na wyjściu **reżyserni głównej**, to powinna istnieć możliwość **omijania** tego miejsca przez programy zgodne z R 128. Oczekuje się, że takie omijanie stanie tym częstsze im więcej programów będzie znormalizowanych do wartości docelowej. Ostatecznym i zalecanym celem jest **normalizacja każdego sygnału audio jako takiego**.

W obszarze **post-produkcji** jest bardziej prawdopodobne osiągnięcie poziomu docelowego z powodu samej natury wykonywanych tutaj operacji - większe możliwości ponownych mikсів czy ich zmiany a zatem i zmiany poziomu głośności. Ponadto w post-produkcji zazwyczaj istnieje wystarczająco dużo czasu, aby wykonać pełny pomiar uśredniony całego programu po jego dojeździe do końca, jak również wykonać **korekcję wzmocnienia**. W środowisku opartym na plikach taka korekcja może być wprowadzona znacznie szybciej niż w czasie rzeczywistym. Takie sytuacje mogą często występować tam, gdzie miksy w post-produkcji wykonuje się „*jak na żywo*” na przykład, bezpośrednio na taśmie, gdzie pojawia się raptem kilka błędów wykonawcy lub ich w ogóle nie ma (w przypadku miksu komentarza). Do tej kategorii należy również przykładowy proces **kopiowania taśmy 1:1** z korekcją głośności wykonywaną „w locie”. Takie sytuacje są więc podobne to tych, jakie istnieją w produkcji na żywo i powinny być traktowane podobnie.

Zwłaszcza w **fazie przechodzenia** do normalizacji głośności takie uprzednio wspomniane procesory głośności, umieszczone w dalszych miejscach toru, będą na pewno pomocne podczas wprowadzania regulacji poziomu głośności oraz wyłapywaniu możliwych odchyłek. Nadawcy (oraz inżynierowie miksu) powinni dążyć do tego, aby takie procesory pracowały jak najmniej w miarę zwiększania ilości materiałów o głośności mieszczącej się w akceptowalnych granicach tolerancji. Dokładny scenariusz przejścia, jego harmonogram oraz plan wprowadzania będą oczywiście różne dla każdego nadawcy (patrz punkt 9). W oczekiwaniu na takie działania należy pamiętać, że w interesie odbiorców, przejście do normalizacji głośności powinno nastąpić w odpowiednim momencie z uwagi na istotne korzyści, jakie taka normalizacja przynosi słuchaczom.

#### 4.4 Zakres głośności dla produkcji i post-produkcji

Stosowanie normalizacji głośności bezpośrednio implikuje kontrolę **Zakresu Głośności (LRA)**, ponieważ występuje tu możliwość stosowania większej dynamiki. Jest to ważne dla zapewnienia właściwego sygnału dla wybranej grupy słuchaczy oraz toru dystrybucji. O ile w produkcji i post-produkcji można tworzyć miks „ogólny” (o względnie wysokiej wartości LRA i Poziomie Maksymalnej Dozwolonej Rzeczywistej Wartości Szczytowej równej -1 dBTP), to poszczególne platformy mogą wymagać niższych wartości LRA oraz niższego poziomu maksymalnej dozwolonej rzeczywistej wartości szczytowej (zachowując jednak poziom głośności programu -23 LUFS). System określony Zaleceniem R 128 uznaje takie podejście ogólne, zakładające późniejszą obróbkę dopasowującą sygnał do poszczególnych środowisk oraz platform, obróbkę wykonywaną w kolejnych miejscach toru.

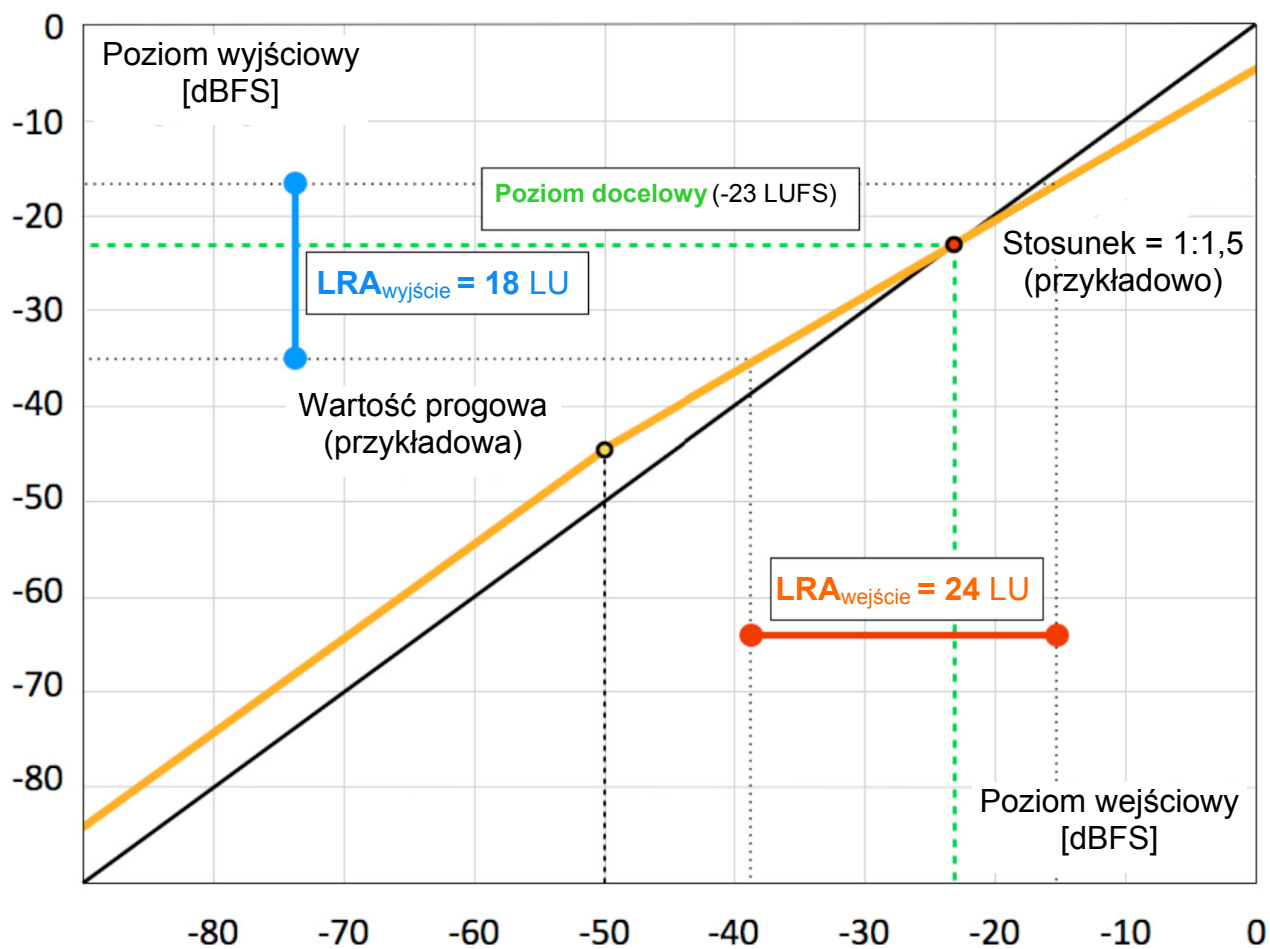
Dysponując parametrem Zakresu Głośności można teraz systematycznie określać odpowiednie środki dla uzyskania *potencjalnej kompresji dynamiki* programu w celu jego dopasowania do przedziału tolerancji słuchaczy lub do platformy dystrybucji sygnału. W praktyce, zadowalające wyniki może zapewniać **niskopoziomowa kompresja całego materiału** (patrz przykład na Rysunku 10): niska wartość progu (< -40 dBFS) oraz umiarkowany stosunek kompresji (1:1,2 - 1:1,5), z długim czasem odpuszczania (>0.5 s) zapewnią tutaj jednolitą kompresję w całym zakresie wartości sygnału.

#### Kompresja LRA (przykład):



- niska wartość progowa (< -40 dBFS)
- mały stosunek kompresji (1:1,2—1:1,5)
- dobór odpowiedniej wartości wzmocnienia wyrównującego

Zależnie od początkowego poziomu głośności można tu równoległe przejść do Poziomu Docelowego -23 LUFS dzięki ustawieniu odpowiedniego wzmocnienia wyrównującego na wyjściu kompresora.

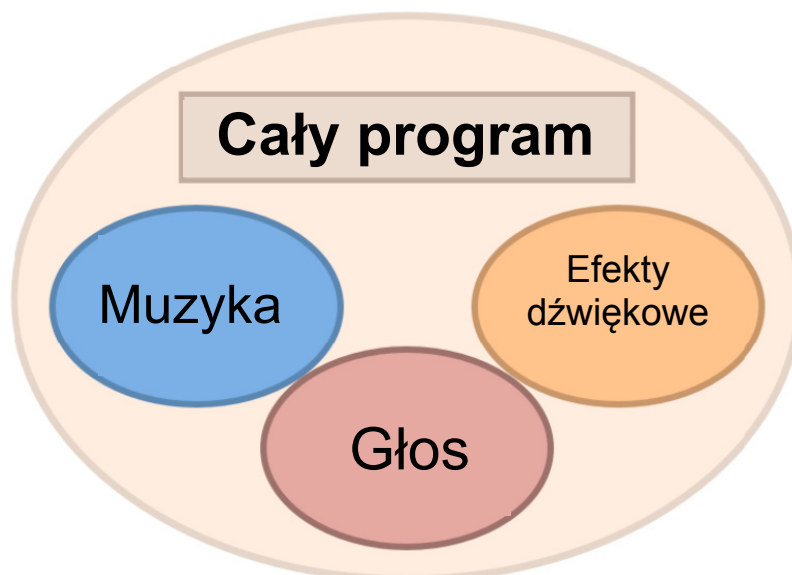


Rysunek 10: przykład kształtowania Zakresu Głośności (LRA) za pomocą kompresora o niskiej wartości progowej (-50 dBFS) i umiarkowanym stosunku kompresji (1:1,5)

## 5. Co należy mierzyć w produkcji i post-produkcji?

### 5.1 Normalizacja niezależna od sygnału a normalizacja „zakotwiczona”

Dokument EBU R 128 zaleca pomiar całego programu, niezależnie od typu poszczególnych sygnałów takich jak głos, muzyka lub efekty dźwiękowe (patrz Rysunek 11). Uważa się, że jest to najbardziej powszechna praktyka, stosowana w przeważającej większości programów:



Rysunek 11: Elementy programu

Dla programów o coraz szerszym zakresie głośności (około  $>12$  LU), można dla normalizowania głośności, wykorzystać opcjonalnie tak zwany **sygnał kotwiczący**, korzystając tym samym z czegoś w rodzaju *metody indywidualnego bramkowania*. To powinien być sygnał, który producent lub inżynier dźwięku chce uczynić reprezentatywnym dla średniej głośności programu, taki jak mowa lub śpiew, konkretny fragment muzyki grany *mezzoforte* czy powtarzana sekwencja efektu dźwiękowego ważna dla dramaturgii programu itd.

Należy jednak podkreślić, że wybór sygnału kotwiczącego jest **procesem aktywnym**, wymagającym udziału doświadczonego operatora. Takie podejście powinno być tylko brane pod uwagę dopiero wtedy, gdy operatorzy i inżynierowie dźwięku zaznajomią się i zaakceptują koncepcję normalizacji głośności. Takie właściwie stosowane podejście może, dzięki wybraniu sygnału kotwiczącego, pomóc w dokładnym dobieraniu głośności w programach o dużym zakresie głośności.

Istnieje również automatyczny pomiar jednego specyficznego sygnału kotwiczącego, jest to tak zwany „*Dialogue Intelligence*”, algorytm opracowany i należący do firmy Dolby Laboratories. Algorytm opiera się na sygnale mowy - powszechnym i ważnym sygnale w środowisku nadawczym. Algorytm wykrywa, czy w programie występuje mowa a następnie mierzy głośność tylko w odcinkach programu zawierających mowę. Dla programów o wąskim zakresie zmian głośności różnica pomiędzy pomiarem ograniczonym do mowy a pomiarem wykonywanym w całym programie jest mała, zazwyczaj mniejsza od 1 LU. Dla programów o szerokim zakresie głośności, takich jak filmy, ta różnica staje się potencjalnie większa, czasem przekracza wartość 4 LU. Automatyczna detekcja sygnału kotwiczącego ma pomóc znaleźć element, który powinien stanowić Poziom Docelowy. Podobnie jak w przypadku każdego algorytmu, wykrywającego specyficzne sygnały spośród kompletnego i złożonego miksu, algorytm dyskryminacji mowy też można oszukać - albo sygnałem o widmie podobnym do widma mowy (na przykład niektóre instrumenty dęte lub skrzypce solo) lub sygnałem mowy, który nie mieści się w granicach progu dyskryminacji (na przykład niektóre dialekty językowe). Z kolei przy powtarzaniu pomiaru głośności programów, w których sygnały kotwiczące bez przerwy oscylują wokół progu dyskryminacji, wyniki poszczególnych pomiarów mogą przyjmować znacząco różne wartości.

Dla programów **krótkich**, takich jak reklamy, promocje, zwiastuny i podobne, (automatyczna) normalizacja mowy prawdopodobnie nie zapewni zadowalających rezultatów zwłaszcza, gdy uwzględnimy mające nastąpić w przyszłości zwiększenie zakresu głośności oraz rozbudowane koncepcje dramaturgiczne. W takich przypadkach, większość zaleceń międzynarodowych (wliczając i to) zgadza się na mierzenie „wszystkiego” wszelkimi możliwymi środkami.

Tak czy inaczej, nadawcy muszą mieć świadomość, że zwłaszcza w środowisku opartym na plikach, gdzie w większości przypadków występować będzie pomiar automatyczny całych programów, niezależnie od typu sygnału (mowa, muzyka, efekty dźwiękowe), prawdopodobnie trzeba będzie opracować inną strategię obróbki programów opartych na normalizacji zakotwiczonej.

#### **Podsumowanie:**

Właśnie z powodu wspomnianych wyżej niepewności oraz faktu, że mowa reprezentuje tylko część całego programu (choć bardzo ważną i powszechną) dokument R 128 zaleca mierzenie „wszystkiego” - to znaczy całego programu, niezależnie od typu sygnału (takiego jak głos, muzyka czy efekty dźwiękowe).

Potwierdzają to podane poniżej obserwacje:

- Różnice pomiędzy pomiarem „wszystkiego” a pomiarem sygnału kotwiczącego (takiego jak głos, muzyka lub efekty dźwiękowe) jest mała dla programów o wąskim Zakresie Głośności;
- Różnica pomiędzy pomiarem „wszystkiego” a „kotwicy” zależy wyraźnie od treści programu, ale sądzi się, że będzie większa jeśli zwiększy się Zakres Głośności;
- Automatyczne wykrywanie sygnału kotwiczącego może być prawidłowe dla większości programów, ale może być też „oszukane” przez sygnały podobne lub może w ogóle nie zadziałać, nie dostarczając zatem stuprocentowo zgodnych wyników;
- Środowiska oparte na plikach wymagają takich metod pomiaru, które można stosować w odniesieniu do każdej możliwej treści i które zapewnią „wystarczająco dobre” wyniki dla wszystkich programów;
- Wybór sygnału kotwiczącego wymaga wkładu pracy ze strony doświadczonego operatora lub zastosowania algorytmu do dyskryminacji, taki algorytm z kolei może dawać niepewne wyniki, co omówiono powyżej.

Normalizacja zakotwiczona może dawać lepsze rezultaty dla materiałów o szerokim zakresie LRA. Jest to jednak zadanie wymagające wiedzy i doświadczenia a zatem czasu i pieniędzy a przy stosowaniu automatycznej dyskryminacji nie ma stuprocentowej pewności. Specjalne środki są konieczne, gdy materiał poddany obróbce z wykorzystaniem zakotwiczenia zostaje wprowadzony do systemu normalizacji działającego na serwerach plików, pomijającego procesy automatyczne. **W celu uwzględnienia tych wszystkich problemów R 128 zaleca, nawet przy materiałach o szerokim zakresie LRA, pomiar całego programu z wszystkimi jego elementami zamiast pomiaru zakotwiczonego.**

## **5.2      Kanał Efektów Niskoczęstotliwościowych (LFE)**

Jak to podano w opisie ITU-R BS.1770 (patrz punkt 2), kanał LFE jest aktualnie **wyłączony** z pomiarów. Jednym z powodów są dość powszechne wątpliwości odnośnie tego rozwiązania występujące wśród konsumentów oraz inżynierów audio jak również występujące różnice w sposobie wysterowania tego kanału (wzmocnienie +10 dB w paśmie roboczym). Pominięcie kanału LFE podczas pomiaru głośności może prowadzić do jego nieprawidłowego wykorzystywania. Podjęcie decyzji, czy i w jaki sposób można by było dołączyć sygnał LFE, wymaga dalszych badań i praktycznych doświadczeń. Jedno rozwiązanie całkowicie eliminujące wszystkie potencjalne problemy z sygnałem LFE polega na nie używaniu go w ogóle, jeśli nie ma potrzeby dodatkowego wzmocnienia w regionie niskich częstotliwości.

## 6. Produkcja i emisja materiałów opartych na plikach

W sytuacji, gdy środowisko nadawcze przechodzi na pracę opartą na plikach, ważne jest i tutaj pełne przyjęcie koncepcji normalizacji głośności. Podstawowa zasada pozostaje ta sama: zaleca się **normalizację głośności** oraz **kontrolę dynamiki sygnału audio**, zwłaszcza dla nowych materiałów. Ponieważ integralną częścią systemów opartych na plikach są **Metadane**, opisano tu również rozwiązania, które polegają bardziej właśnie na Metadanych (punkt 7).

Plik zawierający sygnały audio może pojawić się u nadawcy dzięki procesowi wprowadzania danych (ingest), wskutek transferu z zewnętrznego serwera lub pobraniu z archiwum również opartego na plikach.

Dla istniejących programów (treść archiwalna) istnieją zasadniczo cztery opcje realizacji normalizacji głośności:

- Rzeczywista zmiana poziomu głośności na „docelowy” we **wszystkich plikach audio**
- Zmiana poziomu głośności tylko „na żądanie”
- Wykorzystanie wyniku pomiaru poziomu głośności w celu **regulacji poziomu przy emisji bez zmiany oryginalnego poziomu głośności**
- **Przeniesienie prawidłowych metadanych** głośności do odbiorcy, u którego zostanie wówczas wykonana normalizacja głośności

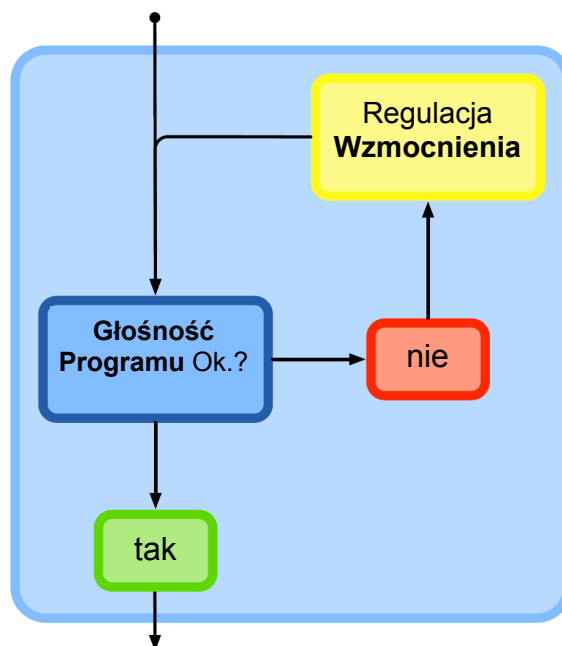
To, które z tych rozwiązań zostanie wybrane, zależy od takich czynników jak specyficzna infrastruktura, organizacja pracy, zarządzanie zasobami medialnymi, dostępność odpowiednich urządzeń, źródła finansowania, czas itd.

Na samym początku „drogi życiowej” pliku powinny być wykonane pomiary, które dostarczą wartości trzech charakterystycznych parametrów dźwięku określonych w EBU R 128 - **Poziomu Głośności Programu**, **Zakresu Głośności** oraz **Poziomu maksymalnego rzeczywistej wartości szczytowej** (dla bardzo krótkich materiałów można również zmierzyć i zapisać wartości chwilowego i krótkoterminowego poziomu maksymalnego (<30 s, patrz punkt 7)). Zależnie od wyników tego pomiaru oraz zastosowanej potem metody normalizacji głośności oraz zgodności z akceptowalnym Zakresem Głośności, zostaje uruchomiony odpowiedni schemat obróbki, składający się z poszczególnych bloków lub „głównych zadań”. Na następnych stronach zostanie omówiony taki schemat obróbki, ilustrowany za pomocą rysunków pokazujących poszczególne bloki zadań.



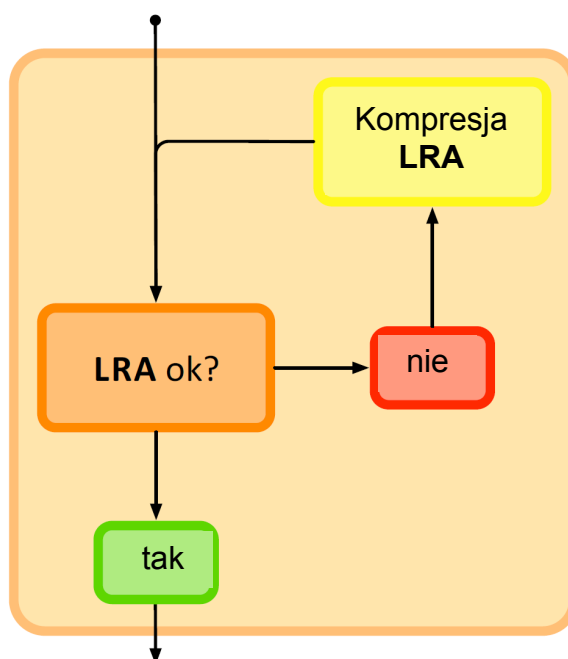
## 6.1 Bloki tworzące obróbkę

Obróbka Poziomu Głośności Programu (Rysunek 12)



Rysunek 12: Blok obróbki głośności programu

Obróbka Zakresu Głośności (Rysunek 13)



Rysunek 13: Blok obróbki Zakresu Głośności

Obróbka Maksymalnego Poziomu Rzeczywistej Wartości Szczytowej (Rysunek 14)



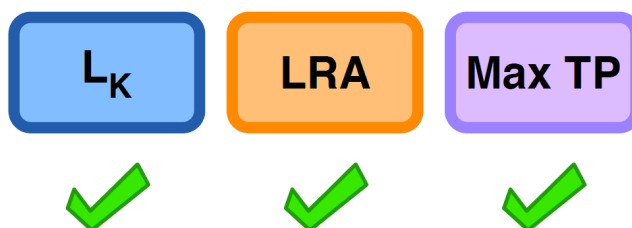
Rysunek 14: Blok obróbki maksymalnego poziomu rzeczywistej wartości szczytowej

### 6.2 Ogólne Strategie regulacji poziomu głośności - obróbka

Trzy opisane powyżej podstawowe bloki tworzą rdzeń każdego procesu kontroli jakości dowolnego pliku odnośnie parametrów technicznych materiału audio. Na początku każdego możliwego do wykonania procesu obróbki zostają zmierzone trzy parametry - Poziom Głośności Programu ( $L_K$ ), Zakres Głośności (LRA) oraz Poziom Maksymalnej Rzeczywistej Wartości Szczytowej (Max TP). Wyniki tych pierwszych pomiarów określają dalsze postępowanie.

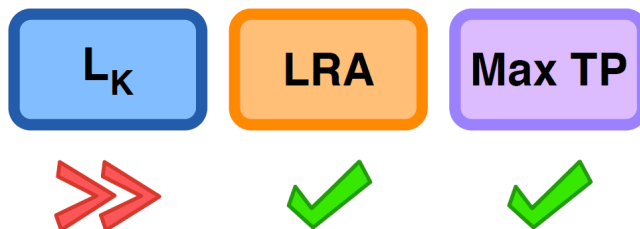
Możliwych jest kilka różnych scenariuszy:

- a) Wszystkie trzy parametry są w porządku.



Jest to oczywiście wynik pomiarów idealny: Poziom Głośności Programu wynosi -23 LUFs, Zakres Głośności mieści się w określonych przez nadawcę granicach (zależnych od rodzaju programu i/lub platformy dystrybucji programu) a Maksymalny Poziom Rzeczywistej Wartości Szczytu jest równy lub mniejszy od podanej wartości maksymalnej odpowiedniej dla określonego systemu dystrybucji.

- b) Poziom Głośności Programu jest większy niż -23 LUFS.

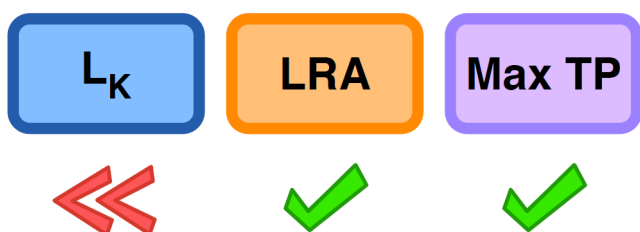


Problem ten rozwiązuje prosta zmiana wzmacnienia (zmniejszenie poziomu) tak, aby:

$$\text{Wzmacnienie (dB)} = L_{K\text{docelowy}} - L_{K\text{zmierzony}}$$

(Przykład: zmierzona wartość  $L_K$  wynosi -19.4 LUFS; Poziom Docelowy jest równy -23 LUFS; wymagane wzmacnienie równa się  $[-23 - (-19.4)] = -3.6$  dB. Maksymalna wartość TP jest naturalnie zmniejszona tak samo jak  $L_K$ .)

- c) Poziom Głośności Programu jest mniejszy niż -23 LUFS.



Po zastosowaniu dodatniej zmiany wzmacnienia należy ponownie obliczyć wartość Maksymalnego Poziomu Rzeczywistej Wartości Szczytowej (początkowo zmierzony Max TP + zmiana wzmacnienia = wynikowy Max TP) - ponieważ wartość ta może znaleźć się powyżej dozwolonego progu. Jeśli nowa wartość Max TP rzeczywiście przekracza ten próg, musi być wykonane **Limitowanie według rzeczywistej wartości szczytowej** pokazane na rysunku Bloku Obróbki Rzeczywistych Wartości Szczytowych. Inne rozwiązanie, stosowane wówczas gdy takie limitowanie nie jest możliwe lub pożądane (lub potencjalnie niebezpieczne) polega na pozostawieniu wartości  $L_K$  na początkowym, niższym poziomie i wprowadzeniu **odpowiedniej wartości metadanych głośności** (niższej niż -23, odzwierciedlającej początkowy poziom głośności). Do tego potrzeba jednak w pełni funkcjonalnego systemu obsługującego i przynoszącego metadane (np. Dolby Digital lub MPEG-4).

W obu scenariuszach b oraz c, można wprowadzić zwykłą wartość wzmacnienia, zapisaną jako metadane połączoną z potencjalnym dalszym limitowaniem, jeśli po zwiększeniu wzmacnienia zostanie przekroczona wartość Max TP, (scenariusz c). Taka wartość wzmacnienia może kontrolować poziom pliku podczas emisji tak aby poziom ten osiągnął wartość -23 LUFS.

- d) Poziom Głośności Programu jest niższy niż -23 LUFS a Zakres Głośności jest szerszy niż wewnętrzna tolerancja dla danego typu programu lub kanału dystrybucji.



Poziom Głośności Programu może być poddany takiej obróbce jak w punkcie c. Zakres Głośności zostaje poddany obróbce (Blok obróbki Zakresu Głośności) a zatem potencjalnie redukuje wartość Max TP. Mimo iż po zwiększeniu wzmacnienia  $L_K$  wartość Max TP może przekroczyć dozwolony limit, obróbka Max TP może nie być konieczna, bo zmniejszyła się wartość LRA. Konieczne jest zatem obliczenie Max TP podczas procesu redukcji LRA.

- e) Zakres Głośności jest szerszy niż granice tolerancji dla danego typu programu lub kanału dystrybucji.



Tak jak podano w punkcie 4.4, kompresor z ustawioną niską wartością progu i bardzo umiarkowanym stosunkiem kompresji może zawęzić zakres głośności (Blok tworzenia Zakresu Głośności). W przypadku plików, zaletą jest tu ich automatyczna obróbka z „docelową wartością LRA”. Alternatywne rozwiązanie polega na tym, że wynik pomiaru LRA może uruchomić w dalszym miejscu toru kompresor dynamiki z presetami o wartościach podobnych do wymienionych w punkcie 4.4. Ponieważ wartość Max TP może tylko się obniżyć, tak więc nie ma tu potrzeby dodatkowej obróbki rzeczywistych szczytów.

- f) Przekroczony jest Maksymalny Dozwolony Poziom Rzeczywistych Szczytów.



Przekroczenie poziomu Max TP w danym systemie dystrybucji wiąże się z ryzykiem **zniekształceń**, jakie mogą pojawić się dalej w torze (np. w przetworniku C-A, konwerterze częstotliwości próbkowania, kodeku redukującym strumień danych). W celu obniżenia wartości Max TP stosowana jest obróbka limitowania rzeczywistych szczytów programu zgodna z rysunkiem bloku obróbki maksymalnego poziomu rzeczywistej wartości szczytowej. To, czy pojawią się wówczas istotne zmiany w Głośności Programu, zależy od liczby i rozmiarów zmienionych szczytów.

Wszelkie inne kombinacje wyników pomiarów początkowych wartości  $L_K$ , LRA oraz Max TP odbywają się w procesach omówionych do tej pory w powyższych scenariuszach.

## 7. Metadane

Jak to podano w punkcie 3.2, normalizacja głośności może być uzyskana albo w wyniku **normalizacji sygnału audio** (metoda zalecana) lub za pomocą **Metadanych** w których jest zapisana wartość rzeczywistego poziomu głośności. W tym przypadku, przejście do Poziomu Docelowego może być wykonane albo w mikserze emisji audio, podczas przenoszenia pliku audio do serwera emisji, dzięki wybraniu odpowiednich presetów dla podłączonego dalej procesora dynamiki lub też bezpośrednio u słuchacza w domu za pomocą regulacji poziomu odstuchu.

Metadane ogólnie mogą być *aktywne* (potencjalnie zmieniające sygnał audio) lub *deskryptywne* (dostarczające informacji o sygnale, takich jak jego format, prawa autorskie itd.). Naturalną konsekwencją prac grupy PLOUD oraz publikacji EBU R 128 i pomocniczych dokumentów rdzeń metadanych głośności w plikach audio tworzą trzy główne parametry: **Głośność Programu (Programme Loudness)**, **Zakres Głośności (Loudness Range)** oraz **Maksymalny Poziom Rzeczywistego Szczytu (Maximum True Peak Level)**. Trwają prace nad włączeniem tych parametrów do nagłówka (pozycja Broadcast Extension (BEXT)) danych formatu Broadcast Wave File (**BWF**) (*szczegółowy opis formatu BWF - patrz pozycje [10], [11] i [12]*). Co więcej, powinny tam również być zapisywane wartości **Maksymalnego Chwilowego Poziomu Głośności (Maximum Momentary Loudness Level)** oraz **Maksymalnego Krótkotrwałego Poziomu Głośności (Maximum Short-term Loudness Level)**, ponieważ te parametry są pomocne podczas kontroli dynamiki bardzo krótkich odcinków (<30 s; patrz również punkt 10). Zamierza się też Metadane Głośności umieścić w słowniku SMPTE wraz z potencjalnymi ulepszeniami takimi jak „Profile Głośności”, omawiającymi na przykład, różne presetu obróbki stosowane w pracujących w dalszych miejscach toru procesorach głośności.

Parametry Metadanych stosowane w istniejących systemach i najważniejsze dla głośności to:

- głośność programu
- słowa kontrolne zakresu dynamiki
- współczynniki miksu (zgrywania)

Na przykład w metadanych systemu Dolby AC-3 parametry te mają nazwę *dialnorm* (dialogue normalization - normalizacja dialogu), *dynrng* (dynamic range - zakres dynamiki) oraz *Centre/Surround Downmix Level* (poziom miksowania środka w stosunku do kanałów otaczających). Parametr *dialnorm* wiernie opisuje głośność całego programu z wszystkimi jego elementami takimi jak głos, muzyka lub efekty dźwiękowe (program tylko muzyczny ma też swoją wartość „*dialnorm*”). Może to być nieco mylące, ale powodem takiego rozwiązania jest oparcie systemu Dolby na normalizacji zgodnej z kotwiczącym sygnałem dialogów.

## 7.1 Metadane Głośności Programu

Zgodnie z decyzją normalizowania poziomu sygnału audio w produkcji do wartości **-23 LUFS**, odpowiadający temu parametr Metadanych powinien oczywiście również być ustawiony tak, aby zapewniał wskazanie **-23 LUFS**, jeśli dany program został znormalizowany do Poziomu Docelowego. Dlatego, po rozpowszechnieniu normalizacji źródłowych sygnałów audio parametr Głośności Programu w Metadanych będzie parametrem **statycznym**.

Sytuacje wyjątkowe, kiedy może być używana wartość inna niż **-23** to:

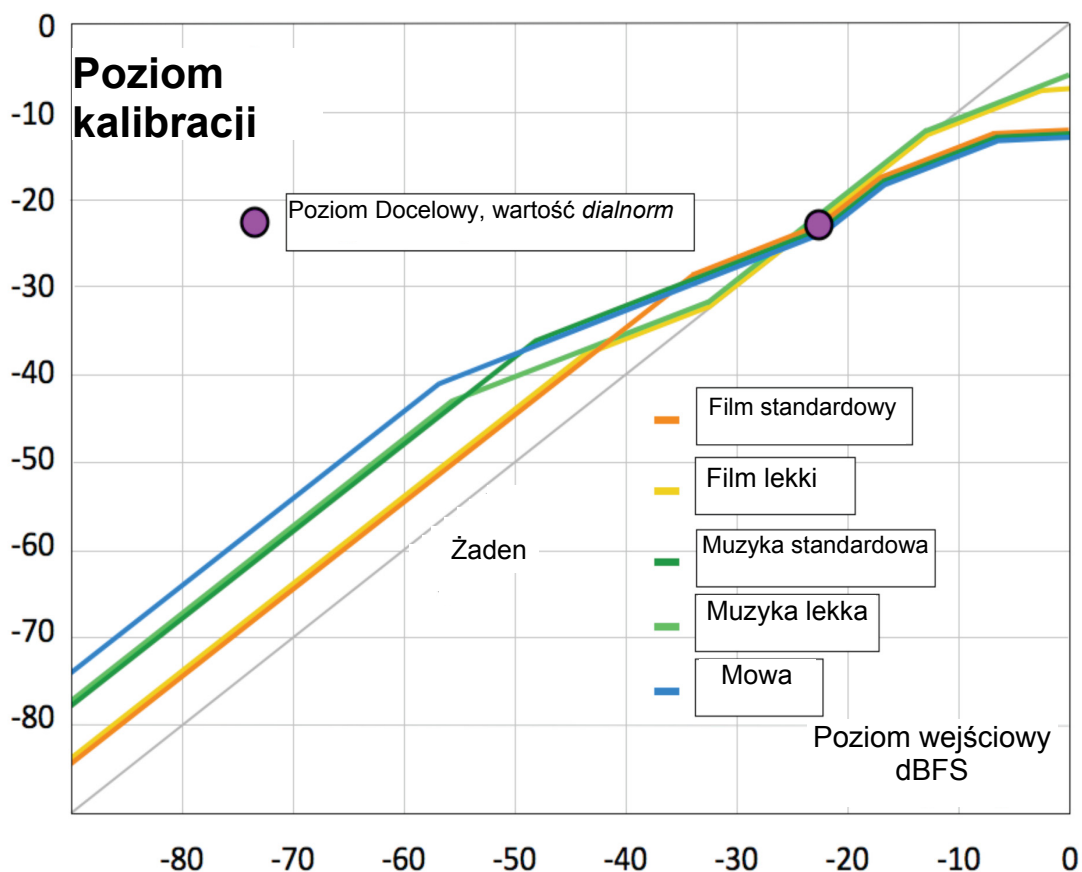
- Program nie mieści się w przedziale określonym wartościami **-23 LUFS** oraz **-1 dBTP**. Może to się pojawić przy bardzo dynamicznych filmach oraz u nadawców, którzy chcą nadawać swój program z takim dużym stosunkiem głośności do wartości szczytowych;
- Programy archiwalne, w których mogą okazać się niemożliwe takie zmiany, które sprowadzą je do poziomu docelowego systemu R 128;
- Zewnętrzne programy na żywo mające różne poziomy głośności i metadane.
- Jest wykorzystywany w pełni funkcjonalny system dostarczający i korzystający z metadanych w całym torze sygnałowym. Implikuje to wierne przenoszenie metadanych głośności do urządzeń domowych słuchacza.

We wszystkich tych sytuacjach, za pomocą wszelkich możliwych środków, powinny być ustawiane **prawidłowe** wartości metadanych dla Głośności Programu, zmierzone miernikiem zgodnym z „trybem EBU”. Systemy dystrybucyjne jak również urządzenia kina domowego powinny dalej korzystać z takich metadanych (patrz *Dok. Tech. 3344 EBU*).

## 7.2 Metadane Kontroli Zakresu Dynamiki

Normalizacja głośności może być wprowadzona w sygnale źródłowym lub poprzez metadane, to samo dotyczy obróbki zakresu dynamiki. W środowisku Metadanych informacja o kompresji zakresu dynamiki jest wysyłana jako część strumienia danych w formie *słów-wzmocnienia* (*gain-words*). W kinie domowym u odbiorcy taka informacja jest wykorzystywana do zmniejszenia zakresu dynamiki sygnału, albo domyślnie albo po aktywacji przez użytkownika. Sterowanie zakresem dynamiki poprzez metadane nie jest tym samym, co zapewniają wysublimowane procesory dynamiki, ale zapewnia swego rodzaju „plaster opatrunkowy” w sytuacjach, gdy użytkownik życzy sobie wyraźnie mniejszego zakresu dynamiki.

Wracając raz jeszcze do systemu Dolby Digital, istnieje tu **6 presetów kompresji** które zmuszają koder do generowania różnych słów sterujących wzmacnieniem, wysyłanych w strumieniu bitów do dekodera w sprzęcie użytkownika, są to: Film Standard, Film Light, Music Standard, Music Light, Speech oraz None (film standardowy, lekki, muzyka standardowa, lekka, mowa oraz „żaden”). Presety te wprowadzają mniejszą lub większą kompresję skoncentrowaną wokół wartości *dialnorm*, co jest kolejnym powodem konieczności prawidłowego ustawienia wartości tego parametru w metadanych (patrz *Rysunek 15* pokazujący krzywe kompresji skupione dookoła wartości **-23 LUFS**).



Rysunek 15: Krzywe ogólne kompresji Zakresu Dynamiki w systemie AC-3

W Dolby Digital istnieją dwa profile kompresji: „*Line mode*” (tryb liniowy) oraz „*RF mode*” (tryb RF). Dla każdego z nich można wybrać oddzielny preset kompresji.

W programach kształtowanych według reguł systemu R 128 i jego koncepcji normalizacji sygnału audio do wartości -23 LUFS, jak również wykorzystujących parametr Zakres Głośności dla ich dowolnej możliwej obróbki, można tu skorzystać z presetu „None” (żaden). Może to być zwłaszcza przydatne w trybie „*Line mode*” jak również, domyślnie w trybie „*RF mode*”.

Sterowanie Zakresem Głośności poprzez rzeczywistą obróbkę sygnału audio podczas transmisji to ogólnie kierowanie problemem wyżej. Jednak w przypadku specyficznych programów nadawca może wybrać dla systemów w trybie RF profil *gentle* (łagodny) (aby uniknąć zbyt aktywnie działającego zabezpieczenia przed przesterowaniem) przy zachowaniu wyboru „None” dla systemów w trybie Line. Nadawcy, którzy chcą w ich wewnętrznych procesach korzystać z profili innych niż „None” muszą być świadomi, że taka funkcjonalność nie zawsze może być niezawodnie wprowadzona do urządzeń słuchaczy. Producentom i dystrybutorom sprzętu zaleca się dopilnowanie tego, aby produkowane przez nich urządzenia spełniały wymagania podane w Dokumencie Technicznym 3344 EBU („*Distribution Guidelines*”).

### 7.3 Współczynniki downmixu (zgrania do wersji stereo)

Parametry w metadanych (znów przykładowo wzięte z Dolby Digital) są oczywiście tylko stosowane dla sygnałów dźwięku otaczającego - sterują wzmacnieniem (w dB) kanału środkowego oraz kanałów otaczających przy ich wmixowywaniu do kanałów Lewy przedni i Prawy przedni w celu otrzymania sygnału stereofonicznego, 2-kanałowego. Głośność dwukanałowego sygnału stereo powstałego wskutek automatycznego mixu za pomocą metadanych zależy od:

- rzeczywistych współczynników mixu jako takich (+3/+1.5/0/-1.5/-3/-4.5/-6/-∞)
- treści programu w kanale środkowym i kanałach otaczających oraz
- potencjalnego limitowania zabezpieczającego przed przesterowaniami

W tak miksowanym sygnale należy uważać, aby **uniknąć przesterowań**. Można to uzyskać dzięki procesorowi dynamiki. Należy unikać statycznego skalowania (redukcji poziomu ogółem) ponieważ to systematycznie wprowadza różnice głośności pomiędzy wersją zmiksowaną do 2-kanałów stereo a początkowym sygnałem dookólnym. Rozwiązaniem może być skalowanie dynamiczne.

Współczynniki zgrania (downmixu) występujące w systemie Dolby-Digital są zarządzane przez dwa profile zgrania. Na początku, gdy istniał tylko jeden profil, parametry miały charakter bardziej zgrubny - były to wartości -3/-4.5/-6 dB dla środka oraz -3/-6/-∞ dB dla kanałów otaczających. Obecnie informacja Extended Bitstream Information (Extended BSI) dodaje dokładniejsze wartości przejściowe, wymienione powyżej (dla przypomnienia: współczynniki downmixu podane w standardzie DVB TS 101 154 również oferują tę samą rozdzielczość jak Dolby Digital Extended BSI). Nadawcy powinni wiedzieć o tym, że nie każde urządzenie odtwarzające jest w stanie dostarczyć słuchaczowi zamierzonych wrażeń, jeśli wykorzystywane będą parametry Extended BSI, ponieważ starsze dekodery mogą nie być w stanie wyodrębnić tej informacji i powrócić do mniej licznych a bardziej zgrubnych współczynników profilu 1.

W przypadku braku metadanych zgrania (lub danych niepewnych), dobrym początkiem może być uwzględnienie współczynników opisanych w dokumencie ITU-R BS.775-2 [13]:

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

Należy tu zauważyć, że podczas pomiaru głośności zgodnie z ITU-R BS.1770, kanały otaczające są ważone wartością +1.5 dB<sup>5</sup>. Po automatycznym zgraniu takie ważenie nie jest wykonywane, wynikiem mixu jest tylko frontalny sygnał 2-kanałowego stereo (Lewy i Prawy Przód). Programy z dużą ilością treści w kanałach otaczających będą w związku z tym wykazywać potencjalnie większe różnice głośności pomiędzy miksem surroundowym a downmiksem stereo niż programy o bardziej „konserwatywnym” wykorzystaniu kanałów przestrzennych.

<sup>5</sup> Współczynnik wagi +1.5 dB dla sygnałów otaczających, stosowany podczas pomiaru głośności zgodnie z ITU-R BS.1770, nie powinien być mylony z rzeczywistym wzmacnieniem +3 dB sygnałów surroundowych stosowanym w kinach! W kinie, dwa oddzielne kanały otaczające są ustawione o 3 dB niżej niż kanały przednie po to, aby ich łączny poziom był równy poziomowi kanału przedniego. Powodem takiego rozwiązania jest konieczność kompatybilności z filmami Mono-Surround (kodowane matrycowo „Dolby Stereo” ma tylko (w ograniczonym paśmie) sygnał typu mono-surround), kiedy oba kanały otaczające mogłyby dostać sygnały identyczne. Dla oddzielnych miksów audio wielokanałowego („5.1” etc.) sygnały surround w miksie finalnym są więc o 3 dB „gorętsze”, ponieważ inżynier miksujący kompensuje w ten sposób o 3 dB niższe ustawienie poziomów kanałów otaczających. Jeśli jest nadawany miks kinowy, to różnica 3 dB musi być skompensowana (należy również stosować inne parametry, takiej jak Zakres Głośności).

O ile wzmacnienie +3 dB kanałów otaczających ma powód czysto techniczny, to wzmacnienie +1.5 dB sygnałów otaczających podczas pomiaru ich głośności ma powody psychoakustyczne. Ludzki słuch odbiera sygnały o tym samym poziomie ciśnienia akustycznego dochodzące z tyłu jako głośniejsze od tych, przychodzących z przodu. Przyrząd pomiarowy nic o tym „nie wie”, więc potrzebuje takiej zmiany wzmacnienia.

W żadnym wypadku nie ma gwarancji, że metadane dostarczone z plikiem *zewnętrznym* (lub innym medium) są prawidłowe. Metadane Głośności Programu wskazujące wartość -27 (fabryczna wartość domyślna dla parametru *dialnorm* w systemie Dolby-Digital) lub -31 (najniższa możliwa wartość w tym systemie) powinny szczególnie zwrócić naszą uwagę, ponieważ są tu szanse, że takie metadane nie były prawidłowo przygotowane lub były niewłaściwie interpretowane w programie, który brzmiał (znacznie) głośniej gdy był odtwarzany u słuchacza w domu.


Zaleca się więc **odrzucać metadanych głośności oraz kontroli zakresu dynamiki** w przypadku źródeł zewnętrznych (poza tymi, kiedy możemy w pełni zaufać takiemu materiałowi). Współczynniki downmixu mogą przechodzić przez w pełni funkcjonalny system metadanych. Cały proces pomiaru trzech głównych parametrów audio musi być wykonany od nowa. Tylko to może zapewnić prawidłową późniejszą obróbkę. Dla celów wewnętrznych metadane mogą być lepiej kontrolowane.

## 8. Regulacja sygnałów w świetle normalizacji głośności

### 8.1 Regulacja sygnału i poziomu

Sygnał kalibracyjny u nadawcy jest sinusoidą o częstotliwości 1 kHz, używaną do technicznego strojenia toru urządzeń pracujących z programami zawierającymi dźwięk. W systemach cyfrowych **poziom** takiego sygnału kalibracyjnego wynosi **18 dB** poniżej maksymalnego poziomu kodowania, niezależnie od całkowitej liczby dysponowanych bitów (-18 dBFS). Przejście na normalizację głośności **NIE** zmienia tej sytuacji, ponieważ kalibracja nie oznacza tu obowiązkowego uwzględnienia wskazywania lub pomiaru głośności.

Dlatego kalibracja poziomów podczas wymiany programów dźwiękowych może być wykonywana tak, jak do tej pory, przy użyciu sinusoidy o częstotliwości 1 kHz i o poziomie -18 dBFS.

 **Regulacja poziomu dla wymiany programów z dźwiękiem nie musi być zmieniana. Stosuje się tu, jak dotychczas, sinusoidę o częstotliwości 1 kHz o poziomie - 18 dBFS.**

Zostało to podane w Zaleceniu R 68 EBU [14]. W tym samym dokumencie wciąż znajduje się „Maksymalny Dopuszczalny Poziom”, w postaci zdefiniowanej w Zaleceniu ITU-R BS.645-2 [15]; uwzględnienie zmiany na „Maksymalny Dopuszczalny Poziom Rzeczywistej Wartości Szczytowej” (-1 dBTP dla ogólnych produkcji PCM) który różni się od zalecanego poziomu -9 dBFS podanego w ITU-R BS.645 (ponieważ metodyka oparta na QPPM /wartości quasi-szczytowej/ stała się przestarzała) oznacza, że powinny być sprawdzone i zmienione odpowiednie punkty w dokumentach EBU R 68 - 2000 oraz ITU-R BS.645 (jak również w dokumentach, które odnoszą się do definicji „Dopuszczalnego Poziomu Maksymalnego” i znajdują się w tych Zaleceniach).

Poziom kalibracyjny równy **-18 dBFS** (ton 1 kHz) jest odczytywany jako **-18 LUFS** na mierniku głośności ze skalą bezwzględną (lub +5 LU na skali względnej trybu EBU), o ile taki ton 1 kHz jest obecny (w jednakowej fazie) w obu kanałach lewym i prawym sygnału stereo lub surroundowego. Jeśli ton 1 kHz jest obecny tylko w pojedynczym kanale przednim, miernik głośności będzie wskazywał wartość -21 LUFS (lub +2 LU na skali względnej).





**Sinusoida stereo o częstotliwości 1 kHz i poziomie -18 dBFS daje na mierniku głośności w trybie EBU odczyt bezwzględny -18 LUFS (lub +5 LU względny).**

## 8.2 Poziom odstuchu

Innym zagadnieniem jest **Poziom Odstuchu** w systemie odtwarzającym audio. W odpowiednim dokumencie, Dokument Techniczny 3276-E EBU „*Warunki odstuchu dla oceny materiału programu dźwiękowego*” (oraz w Dodatku 1, rozszerzającym ten temat do dźwięku wielokanałowego), dla regulacji poziomu w jednym głośniku przyjęto podane niżej wzory [16]:

- (1)  $L_{LISTref} = 85 - 10\log_2 dB_A$  (dla stereo 2-kanałowego)
- (2)  $L_{LISTref} = 96 dB_C$ , odniesiony do sygnału cyfrowego o poziomie pełnej skali (dla audio wielokanałowego do 5.1)

Aby to osiągnąć, należy wykorzystać sygnał zawierający szum o jednakowej mocy przypadającej na oktawę częstotliwości i pokrywający albo cały zakres częstotliwości (równanie (1)) albo zakres częstotliwości od 500 Hz do 2 kHz (równanie (2)). Pomiary powinny być wykonane przy średnim poziomie sygnału równym **poziomowi kalibracji**, który jest tu zdefiniowany jako leżący o 18 dB poniżej pełnej skali cyfrowej. W takich warunkach wzmocnienia w głośnikach powinny być regulowane tak, aby dojść do **Poziomu Odniesienia Odstuchu** ( $L_{LISTref}$ ) równego  $85 - 3 = 82 dB_A$  poziomu ciśnienia akustycznego dźwięku (SPL) na jeden głośnik w systemach 2-kanałowego stereo oraz  $96 - 18 = 78 dB_C$  SPL w systemach wielokanałowych. Pomiary powinny być wykonywane w miejscu odniesienia dla odstuchu za pomocą miernika poziomu dźwięku z wolnym czasem reakcji i z ważeniem według krzywej A dla 2-kanałowego stereo oraz krzywej C dla dźwięku wielokanałowego.

Może to nieco dziwić, te różne liczby i różne sygnały szumowe oraz różne ważenie częstotliwościowe dla miernika poziomu dźwięku. Ale te właśnie różnice kompensują w pewien sposób i zapewniają podobny poziom odstuchu zarówno w systemach 2-kanałowego stereo jak i systemach wielokanałowych.

### Podsumowanie:

Dla stereo 2-kanałowego:	$L_{LISTref} = 82 dB_A$ SPL na głośnik	(przy szumie w paśmie 20 Hz - 20 kHz o jednakowej mocy na oktawę i poziomie -18 dBFS rms)
Dla wielokanałowego 5.1:	$L_{LISTref} = 78 dB_C$ SPL na głośnik	(przy szumie w paśmie 500Hz - 2 kHz o jednakowej mocy na oktawę i poziomie -18 dBFS rms)

Oczekuje się, że po wprowadzeniu Zalecenia EBU R-128 poziom średni programów dźwiękowych *będzie niższy*. Obniżenie poziomu może sięgać aż do 3 LU (w przypadkach ekstremalnych nawet więcej). To zaś czyni prawdopodobnym odpowiedni **wzrost** poziomu odstuchu w systemach odtwarzających. Jak to wspomniano powyżej, *Poziom Kalibracji* nie musi ulec podobnej zmianie, ponieważ ta ważna procedura kalibracji nadal zapewnia rozsądny podział wzmocnienia oraz wysoki stosunek sygnału do szumu w kanale odczytu. Jeśli w przyszłości przyjmie się powszechnie zwiększenie poziomu odstuchu wskutek obniżenia poziomu średniego - odpowiednie dokumenty zostaną przejrzane ponownie.



**Ponieważ poziom –23 LUFS jest około 3 LU niższy od aktualnego średniego poziomu programu - można rozważyć odpowiednie podniesienie poziomu sygnałów kierowanych do odsłuchu.**

## 9. Implementacja i migracja

Jasne jest, że taka fundamentalna zmiana sposobu pomiaru, pokazywania oraz obróbki sygnałów audio oraz jej wpływ na wszystkie etapy produkcji, dystrybucji, archiwizacji oraz transmisji audio nie zajdzie nagle, za jednym „pstryknięciem”. Każdy nadawca czy zespół ludzi zajmujących się audio musi znaleźć swój indywidualny sposób na wprowadzenie tej zmiany, instalację odpowiednich urządzeń, przeszkolenie pracowników - umożliwiające wjazd na drogę wiodącą do głośnościowego raj! Niemniej jednak można podać tu kilka pewnych informacji, które przydadzą się każdemu. Są one omówione w podanych niżej punktach.

### 9.1 Ogólne porady dotyczące Migracji oraz Implementacji

- Załóż **wewnętrzną grupę ds. głośności** w celu dyskusowania podstawowych implikacji i strategii przekonujących zarząd, twórców programów oraz kolegów.
- **Zacznij od zaraz** - nie czekaj aż wszystko będzie na miejscu i zrobią to inni, nie staraj się też być doskonałym już od samego początku.
- Zanim zaczniesz cokolwiek zmieniać, Twoje kierownictwo musi się zgodzić na tę zmianę i jej wszystkie konsekwencje. Uzyskaj od głównego szefa **pisemną zgodę** lub dokument typu „wezwanie do działania”.
- Dostarcz **mierniki głośności** kluczowym osobom w produkcji. Pozwól im je wypróbować, uzyskać pierwsze doświadczenia i wiedzę o zaletach i swobodzie, jakie przynosi metodyka regulacji głośności tak, aby ci ludzie stali się opiniotwórczy dla swych kolegów.
- **Badaj rynek** pomiarów głośności i zarządzania głośnością w celu określenia co najbardziej pasuje do twojego środowiska.
- Określ **kluczowe obszary**, w których powinny pojawić się działania w kierunku regulacji głośności. Potencjalni kandydaci to: studia produkcyjne, stanowiska post-produkcji, wozy transmisyjne, kontrola jakości.
- Bądź przygotowany na różne przeciwności („to zawsze robiliśmy w ten sposób”, „tego nigdy nie robiliśmy w ten sposób”, „kim ty jesteś, żeby mówić nam, co i jak mamy robić”). **Cierpliwość** oraz pokazywanie praktycznych przykładów zawsze się opłaci. Stań się miejscowym guru w sprawach normalizacji głośności („umiara - prostota - naturalność”).
- **Daj każdemu czas na przystosowanie się**. Chociaż słuchacze czekają na to rozwiązanie od dziesięcioleci, nie twórz dodatkowych problemów próbując działać za dużo i za szybko.
- Rozwiązania dla schematów pracy opartych na plikach są wciąż rzadkością (Luty 2011). Obserwuj rynek i żądaj nowych rozwiązań od dostawców.
- Wykorzystaj tę fundamentalną zmianę jako możliwość generalnej **dyskusji o jakości dźwięku** oraz rozwijaniu „**brzmienia korporacyjnego**” obejmującego na przykład zrozumiałość mowy, proporcje pomiędzy mową a muzyką oraz oczywiście normalizację głośności programów.
- **Używaj uszu i uwierz w swoje uszy!** To są najlepsze mierniki głośności. Śmieję się z tych, którzy stosują reguły regulacji głośności tak, jakby nic innego poza nimi nigdy nie istniało.

## 9.2 10 Punktów Działań dla migracji i implementacji

- 🔊 Załóż wewnętrzną grupę do spraw głośności.
- 🔊 Nie czekaj, zacznij od zaraz.
- 🔊 Uzyskaj pisemną zgodę od szefostwa.
- 🔊 Zapewnij kluczowym osobom mierniki głośności.
- 🔊 Badaj rynek urządzeń do pomiarów głośności.
- 🔊 Wybierz kluczowe obszary, które należy zacząć zmieniać.
- 🔊 Będą „schody”. Bądź cierpliwy, daj czas...
- 🔊 Wykorzystaj tę chwilę do dyskusji o jakości audio.
- 🔊 Używaj i wierz swoim uszom.
- 🔊 Zostań miejscowym guru od głośności.

## 10. Problemy związane z rodzajem programu

Koncepcja podana w Zaleceniu EBU R 128 skupia się na normalizacji każdego programu do jednego Poziomu Docelowego (-23 LUFS). Są dwa powody, dlaczego to rozwiązanie nie może być idealnym:

- Nieobiektywny pomiar głośności nie może być idealny
- Zawsze będą istniały indywidualne preferencje

Z tego powodu, idealne rozwiązanie jest *ogólnie* niemożliwe, ponieważ musi być dla każdego inne. Mówiąc o metodyce EBU R 128 ważne jest zrozumieć, że celem tutaj nie jest wyrównanie głośności w oparciu o poziom rzeczywisty ciśnienia akustycznego specyficznego sygnału audio, ale zapewnienie większości słuchaczy satysfakcjonujących wrażeń odsłuchowych przy programach różnego typu.

Przejawia się to na przykład tym, że kwartet smyczkowy Schuberta będzie miał taki sam poziom uśrednionej głośności jak symfonia Mahlera, a mianowicie -23 LUFS. Mimo iż to nie odzwierciedla warunków rzeczywistych, to sprawia, że obie te pozycje będą dopasowane pod względem głośności do szerokiej gamy programów sąsiadujących i taki jest właśnie cel **faworyzowania jednej liczby**.

Ponieważ ten dokument oparty jest na zbiorze pewnych doświadczeń, może to kogoś zachęcić do rozważenia ponownej definicji opisanego tu zbioru reguł, gdy normalizacja głośności będzie już szeroko rozpowszechniona. Słuchacze zaakceptują to, jeśli poziom głośności znajdzie się w ramach tak zwanej „strefy wygody” wynoszącej około 8-9 LU, przy asymetrycznym rozkładzie głośności (na przykład +3 LU/-5 LU). Wówczas nawet w tych przypadkach, w których algorytm obiektywnej głośności nie zawsze zapewni wynik idealny, program będzie najprawdopodobniej wciąż znajdował się we wspomnianej strefie wygody. Nadawcy powinni również pamiętać o tym, że słuchacze mogą w dalszym ciągu regulować poziom głośności za pomocą pilotów zdalnego sterowania, w celu ostatecznego dostosowania do własnych gustów.

EBU zachęca do **normalizacji do jednego poziomu docelowego** mimo potencjalnych udoskonaleń, czy zmian, przeznaczonych dla poszczególnych typów programu. Dopuszczenie zbyt wielu różnych wersji (lub nawet tylko kilku) może być, na samym początku, zbyt trudne dla systemu do jednakowej głośności średniej. Oczywiście zachodzi obawa, że takie zmiany będą ukierunkowane w stronę zwiększenia głośności.

Ustawienie Głośności Programu **niżej** niż Poziom Docelowy jest nieco innym problemem. Jako „studium przypadku” zostaną dalej opisane dwa typy programu, które w pewnych konkretnych okolicznościach mogą zostać poddane specyficznej obróbce (również w odniesieniu do Maksymalnego Poziomu Głośności), są to reklamy i trailery oraz programy muzyczne.

## 10.1 Reklamy i Trailery

Program tego typu jest prawdopodobnie najczęściej wymieniany wówczas, gdy chodzi o **udrękę słuchaczy** a zatem jest główną przyczyną aktualnych problemów z głośnością. W niektórych krajach, w celu opanowania tych problemów zaczęto ostatnio nawet stosować instrumenty prawne, są to Wielka Brytania (przepisy BCAP - Broadcast Committee of Advertising Practice - komitet nadawców do spraw reklam) oraz USA (przepisy CALM - Commercial Advertisement Loudness Mitigation - ograniczenia głośności reklam). Jest naprawdę ważne, aby system normalizacji głośności oparty na dokumencie **R 128 EBU** stanowił zbiór skutecznych narzędzi do realizacji tego zadania - należy zapobiegać nadużyciom. W świecie znormalizowanej głośności do kontroli dynamiki reklam, atakującymi nagłymi i zbyt dużymi różnicami głośności (za głośne „skompensowanie” dłuższego odcinka o niskim poziomie leżącym tuż powyżej progu zadziałania bramki), nie nadaje się niestety parametr Zakres Głośności (LRA), ponieważ jego wyliczenie jest oparte na krótkoterminowych wartościach głośności (odstęp 3 sekund). Dlatego dla bardzo krótkich pozycji istnieje za mało punktów danych dla uzyskania istotnej wartości LRA. Parametr Zakres Głośności nie jest temu winny, ponieważ nigdy nie był przeznaczony dla tego celu.

Alternatywą może być tu wykorzystanie **Poziomu maksymalnego chwilowej głośności** (Max ML - 400 ms) i/lub **Maksymalnego Poziomu Krótkoterminowego** (Max SL - 3 s). Parametry te mogą być skutecznie wykorzystane w celu ograniczenia szczytów głośności pozycji bardzo krótkich (<30 s). Pierwsze doświadczenia członków grupy PLOUD wskazały wartość około **+8 LU** (-15 LUFS) jako możliwą wartość graniczną dla parametru Max ML oraz **+3 LU** (-20 LUFS) dla Max SL. W każdym przypadku, oba te parametry (Max ML i Max SL) są częścią sugerowanego rozszerzenia metadanych formatu Broadcast Wave File Format (BWF). Członków EBU zachęca się do prób wykorzystania indywidualnych limitów Max ML lub Max SL dla pozycji krótkich i raportowania o wynikach.



**Aby zapobiec za dużej głośności w pozycjach bardzo krótkich (<30 s) można wykorzystać limit dla wartości Maksymalnej Głośności Chwilowej i Maksymalnej Głośności Krótkoterminowej.**

Dla programów składających się tylko z tła lub pożądanym przez twórcę dźwięków o niskim poziomie, można zastosować poziom głośności **niższy** niż Poziom Docelowy. Zgadza się to z byłą i aktualną praktyką limitowania poziomów maksymalnych szczytów (obecnie: poziomu głośności), ale nie umieszcza całego materiału w w pobliżu tego maksimum. Dźwięk o specjalnie ustalonym niskim poziomie pozwala na stosowanie **kontrastu**, jednego z najbardziej podstawowych narzędzi tworzenia w każdym rodzaju sztuki. Krótki czas trwania reklam lub trailerów, skutecznie wykorzystujących takie narzędzie dramaturgii nie powinien mieć jakiegokolwiek wpływu na całodzienną, długoterminową średnią głośność programów danej stacji.

Programy przeznaczone do emisji z poziomem niższym niż Poziom Docelowy wymagają specjalnej uwagi, należy zadbać o to, aby przechodziły automatyczny proces normalizacji bez szkody. Powinny naprawdę być wyjątkiem a nie regułą.

Ostatecznie odpowiedzialność za wszystkie te przypadki i decyzje ponoszą odpowiednio producent, reżyser lub inni twórcy.

## 10.2 Muzyka

Doświadczenia pasjonatów słuchających często muzyki sugerują, że pewnych programów, zawierających głównie muzykę - i tę o dużych zmianach głośności, jak muzyka klasyczna i tę z wyższym poziomem kompresji dynamiki stosowanej jako środek wyrazu artystycznego (koncerty rockowe), słucha się z wyższym poziomem głośności (średnio do +2-3 LU) niż programów innych. Powodem tego może być fakt, że w rzeczywistych warunkach występują w takiej muzyce wyraźnie wyższe poziomy ciśnienia akustycznego (fortissimo orkiestry symfonicznej czy gra kapeli rockowej dodatkowo nagłośnionej) oraz fakt, że dla tego typu muzyki nie istnieje problem proporcji „dźwięków przedniego planu” a „dźwięków tła” - wszystko co słychać to dźwięki pierwszoplanowe.

Ale jak już wspomniano, możliwe zróżnicowanie poziomu docelowego dla każdego z tych programów może spowodować więcej szkody niż pożytku przywracając tendencję do zwiększania głośności w celu odróżnienia od innych typów programów. Opierając się na tym samym rozumowaniu, jak w przypadku reklam czy trailerów, **nie zaleca się** normalizacji do innego (tutaj: wyższego) poziomu docelowego. Słuchacze wciąż mogą korzystać z pilotów w swoich urządzeniach w celu regulacji (zwiększenia) poziomu głośności odpowiednio do ich gustów. Konsekwentnie zostanie wówczas zmieniona głośność programów towarzyszących, takich jak reklamy czy zwiastuny. Przewiduje się, że nie powinno to usuwać tych programów ze wspomnianej uprzednio „strefy wygody”.

## 11. Literatura

- [1] Zalecenie Techniczne R 128 EBU “*Loudness normalisation and permitted maximum level of audio signals*” (*normalizacja głośności i dozwolony maksymalny poziom sygnałów audio*) (2010)
- [2] ITU-R BS.1770 “*Algorithms to measure audio programme loudness and true-peak audio level*” (*algorytmy pomiaru głośności programu audio oraz poziomu rzeczywistej wartości szczytowej audio*) (2006-2007)
- [3] EBU Tech Doc 3341 “*Loudness Metering: ‘EBU Mode’ metering to supplement loudness normalisation in accordance with EBU R 128*” (*Pomiar głośności: Pomiar w ‘Trybie EBU’ uzupełniający normalizację głośności zgodnie z R 128 EBU*) (2010)
- [4] EBU Tech Doc 3342 “*Loudness Range: A descriptor to supplement loudness normalisation in accordance with EBU R 128*” (*Zakres Głośności: deskryptor dopełniający normalizację głośności zgodnie z R 128 EBU*) (2010)
- [5] EBU Tech Doc 3344 “*Practical Guidelines for Distribution Systems in accordance with EBU R 128*” (*Praktyczne Wskazówki dla Systemów Dystrybucji zgodnie z R 128 EBU*) (2011)
- [6] ISO 80000-8: “*Quantities and Units - Part 8: Acoustic*” (*Wielkości i Jednostki - Część 8: Akustyka*)
- [7] Grimm E., Skovenborg E. & Spikofski G. “*Determining an Optimal Gated Loudness Measurement for TV Sound Normalization?*” (*Wyznaczanie Pomiaru Optymalnej Bramkowanej Głośności dla Normalizacji Dźwięku w telewizji*), Referat AES Convention Paper N° 8154, 128<sup>th</sup> AES Convention, May 2010
- [8] ITU-R BS.1864 “*Operational practices for loudness in the international exchange of digital television programmes*” (*Praktyki operacyjne odnośnie głośności w międzynarodowej wymianie programów telewizji cyfrowej*) (2010)
- [9] Lund, Th. “*Stop counting samples*” (*Przestańmy liczyć próbki*), AES paper N° 6972, 121<sup>st</sup> AES Convention, October 2006
- [10] Zalecenie Techniczne R 85 EBU: “*Use of the Broadcast Wave Format for the Exchange of Audio Data Files*” (*Wykorzystanie formatu BWF do wymiany plików danych audio*) (2004)
- [11] Zalecenie Techniczne R 111 EBU: “*Multi-channel Use of the BWF Audio File Format (MBWF)*” (*Wykorzystanie formatu BWF (MBWF) w wielokanałowym audio*) (2007)
- [12] EBU Tech Doc 3306: “*MBWF/RF64: An extended File Format for Audio*” (*MBWF/RF64: Rozszerzony format plików audio*) (2009)
- [13] ITU-R BS.775-2 “*Multichannel stereophonic sound system with and without accompanying picture*” (*System wielokanałowego dźwięku stereofonicznego z i bez towarzyszącego mu obrazu*) (2006)
- [14] Zalecenie Techniczne R 68 EBU: “*Alignment level in digital audio production equipment and in digital audio recorders*” (*Poziom kalibracyjny w cyfrowych urządzeniach produkcji audio i w cyfrowych rejestratorach audio*) (revision 2000)
- [15] ITU-R BS.645-2 “*Test signals and metering to be used on international sound programme connections*” (*Sygnaty testowe oraz pomiary przewidziane do użytku w wymianie międzynarodowej programów dźwiękowych*) (1992)
- [16] EBU Tech Doc 3276-E (+ supplement 1) “*Listening conditions for the assessment of sound programme material*” (*Warunki odsłuchu dla oceny materiału programów dźwiękowych*) (1998, 2004 - supplement 1)