

Praktické pokyny pro výrobu a implementaci podle EBU R 128

Doplňující informace k EBU R 128

Status: verze 2.0

Ženeva
srpen 2011

Označování shody

Tento dokument obsahuje jak **normativní**, tak **informativní** text.

Veškerý text je normativní kromě úvodu, odstavců výslovně označených jako „informativní“ nebo jednotlivých odstavců, které začínají označením „Poznámka:“.

Normativní text popisuje nutné nebo povinné prvky. Obsahuje klíčová slova „bude“ („musí“), „měl by“ nebo „může“, definovaná následovně:

„Bude“ a „nebude“ („musí“ a „nesmí“): označují požadavky, které se musí dodržovat přesně, od nichž není přípustná žádná odchylka, pokud mají být požadavky tohoto dokumentu dodrženy.

„Měl by“ a „neměl by“: označují, že z několika možností je jedna doporučena jako zvláště vhodná, aniž by se zmiňovaly či vylučovaly ostatní NEBO naznačují, že určitý postup je upřednostňovaný, ale nevyžaduje se povinně NEBO udávají (v záporném tvaru), že určitá možnost nebo postup se nedoporučuje, ale rovněž nezakazuje.

„Může“ a „nemusí“ označují přípustný postup v rámci limitů tohoto dokumentu.

Výchozí nastavení znamená povinné (ve větách vyjadřujících nutnost) nebo doporučené (ve větách vyjadřujících doporučení) nastavení, které může být volitelně přepsáno uživatelem nebo nahrazeno jinými možnostmi v pokročilých aplikacích. Povinná výchozí nastavení se musí dodržovat. Dodržování doporučených výchozích nastavení se upřednostňuje, ale nevyžaduje jako povinné.

Informativní text je potenciálně užitečný pro uživatele, ale není nenahraditelný a lze jej redakčně odstranit, změnit nebo doplnit, aniž by to poškodilo normativní text. Informativní text neobsahuje žádná klíčová slova pro shodu.

Řádná implementace je taková, která zahrnuje všechna povinná ustanovení („musí“) a v případě realizace i všechna doporučená ustanovení („měl by“) podle popisu. K tomu, aby implementace byla řádná, se nemusí realizovat volitelná ustanovení („může“) a tato ustanovení se nemusí realizovat tak, jak jsou popsána.

Obsah

| | |
|---|----|
| 1. Úvod | 5 |
| 2. EBU R 128, ITU-R BS.1770 | 7 |
| 2.1 Hlasitost pořadu | 9 |
| 2.2 Rozsah hlasitosti | 10 |
| 2.3 Skutečná špičková úroveň (TPL), maximální přípustná TPL | 11 |
| 2.4 Logo R 128 | 13 |
| 3. Obecná koncepce normalizace hlasitosti | 13 |
| 3.1 Špička versus hlasitost | 13 |
| 3.2 Normalizace signálu versus metadata | 14 |
| 3.3 Používání parametru rozsah hlasitosti | 17 |
| 3.4 Výstup na skutečnou špičku | 18 |
| 4. Strategie pro normalizaci hlasitosti | 19 |
| 4.1 Produkce (výroba), postprodukce | 19 |
| 4.2 Měření hlasitosti v produkci a postprodukci | 21 |
| 4.3 „Připravit se, nastavit (úroveň), START!“ | 22 |
| 4.4 Rozsah hlasitosti pro produkci a postprodukci | 24 |
| 5. Co měřit v produkci a postprodukci | 26 |
| 5.1 Normalizace nezávislá na signálu versus normalizace založená na kotvě | 26 |
| 5.2 Kanál nízkofrekvenčních efektů (LFE) | 27 |
| 6. Bezpásková produkce a vysílání | 28 |
| 6.1 Stavební bloky | 29 |
| 6.2 Obecné strategie vyrovnávání hlasitosti - zpracování | 30 |
| 7. Metadata | 32 |
| 7.1 Metadata hlasitosti pořadu | 33 |
| 7.2 Metadata regulace dynamického rozsahu | 33 |
| 7.3 Koeficienty pro mix na menší počet kanálů | 35 |
| 8. Vyrovnávání signálů ve světle normalizace hlasitosti | 36 |
| 8.1 Vyrovnávací signál a úroveň | 36 |
| 8.2 Poslechová úroveň | 37 |
| 9. Implementace a migrace | 38 |
| 9.1 Všeobecná doporučení k migraci a implementaci | 38 |
| 9.2 10 akčních bodů pro migraci a implementaci | 39 |
| 10. Specifické problémy jednotlivých žánrů | 39 |
| 10.1 Reklamy a upoutávky | 40 |
| 10.2 Hudba | 41 |
| 11. Literatura | 42 |

Poděkování

Ačkoliv je tento dokument výsledkem rozsáhlé týmové práce v rámci skupiny PLOUD EBU, tím, kdo jako první sepsal, shromáždil, obohatil a pročistil text do podoby k uveřejnění za mnoho a mnoho týdnů a měsíců úsilí byl trpělivý předseda této skupiny, Florian Camerer.

Věnování

Tento dokument je věnovaný dvěma skvělým zvukovým technikům, Gerhardu Stollovi a Gerhardu Steinkemu.

Praktické pokyny k výrobě a implementaci podle EBU R 128

| <i>Komise EBU</i> | <i>První vydání</i> | <i>Revize</i> | <i>Nové vydání</i> |
|-------------------|---------------------|---------------|--------------------|
| Technická komise | 2011 | srpen 2011 | |

Klíčová slova: zvuk, hlasitost, normalizace, výroba, implementace.

1. Úvod

Tento dokument v praktických detailech popisuje jednu z nezásadnějších změn v historii zvukového vysílání; změnu modelu vyrovnávání úrovně z normalizace špiček na **normalizaci hlasitosti**. Nelze ani dostatečně zdůraznit, že **měření hlasitosti a normalizace hlasitosti** znamená **skutečnou revoluci ve vyrovnávání zvukové úrovně**. Tato změna je důležitá kvůli problému, který se stal hlavním zdrojem rozladění televizních diváků a rozhlasových posluchačů po celém světě; tj. kvůli skokům v úrovni hlasitosti při přerušování pořadů, mezi pořady a mezi programy („pořad“ je definován v poznámce pod čarou 2).

Model vyrovnávání hlasitosti postihuje všechny fáze vysílaného zvukového signálu od výroby po distribuci a přenos. Konečným cílem je tak harmonizovat úrovně hlasitosti zvuku, aby se dosáhlo **stejněměrné univerzální úrovně hlasitosti** ve prospěch posluchače.

Normalizace hlasitosti je skutečnou revolucí ve vyrovnávání úrovně zvuku

Hned je třeba zdůraznit, že to **neznámá**, že úroveň hlasitosti bude neustále konstantní a rovnoměrná v rámci pořadu, právě naopak! Normalizace hlasitosti zajistí, že **průměrná** hlasitost **celého** pořadu bude stejná pro všechny pořady; v rámci pořadu může úroveň hlasitosti samozřejmě kolísat podle uměleckých a technických potřeb. Při nové (skutečné) špičkové úrovni a (ve většině případů) nižší průměrné úrovni hlasitosti bude dynamický rozsah (nebo spíše „rozsah hlasitosti“, viz § 2.2) ve skutečnosti vyšší než při současné normalizaci špiček a směřovacích postupech ve vysílání.

Základem koncepce normalizace hlasitosti je kombinace **Technického doporučení EBU R 128** „Normalizace hlasitosti a maximální povolená úroveň zvukových signálů“ [1] a **Doporučení ITU-R BS.1770** „Algoritmy k měření hlasitosti zvukových pořadů a skutečné špičkové úrovně zvuku“ [2].

Dokumenty R 128 a ITU-R BS.1779 představují základ. Čtyři další technické dokumenty EBU uvádějí podrobnosti.

Vedle dokumentu R 128 uveřejnila skupina PLOUD EBU čtyři další dokumenty:

- **EBU Tech Doc 3341** „Měření hlasitosti: měření v „režimu EBU“ na doplnění normalizace hlasitosti v souladu s EBU R 128“ [3]
- **EBU Tech Doc 3342** „Rozsah hlasitosti: parametr na doplnění normalizace hlasitosti v souladu s EBU R 128“ [4]
- **EBU Tech Doc 3343** „Praktické pokyny k výrobě a realizaci podle EBU R 128“ [tento dokument] a
- **EBU Tech Doc 3344** „Praktické pokyny pro distribuční systémy podle EBU R 128“ [5]

Důležitou roli v praktické implementaci normalizace hlasitosti hrají i technické dokumenty o „měření hlasitosti“ a o parametru „rozsah hlasitosti“. Ty budou také představeny a budou na ně uvedeny odkazy v příslušných kapitolách.

Kruh uzavírají „Pokyny k distribuci“, které se zabývají všemi aspekty normalizace pro distribuci zvukových signálů a řeší kritická spojení mezi výrobou a konečným příjemcem, spotřebitelem. Protože se jedná o velmi podrobný dokument, nebudeme ho zde představovat, kromě občasných odkazů.

Na začátku těchto *Praktických pokynů* představíme klíčové části dokumentů **EBU R 128** a **ITU-R BS.1770**, potom bude následovat **všeobecná koncepce a filozofie normalizace hlasitosti**. Dále se dokument podívá na **strategie hlasitosti** ve **výrobě** a **postprodukci** (měření, mixování, metadata, atd.) a na bezpáskové pracovní toky, tj. problémy s načítáním, vysláním a archivací (měření, automatizovaná normalizace, metadata, atd.).

Samostatná kapitola se podívá podrobněji na **metadata**. Probírá se **vyrovnání** zvukových signálů a **úroveň poslechu** ve studiu a jsou uvedeny **praktické rady** pro **přechod** na výrobu s normalizovanou hlasitostí (implementaci a migraci). **Specifické žánrové problémy** související s reklamami a upoutávkami i hudebními pořady řeší poslední kapitola.

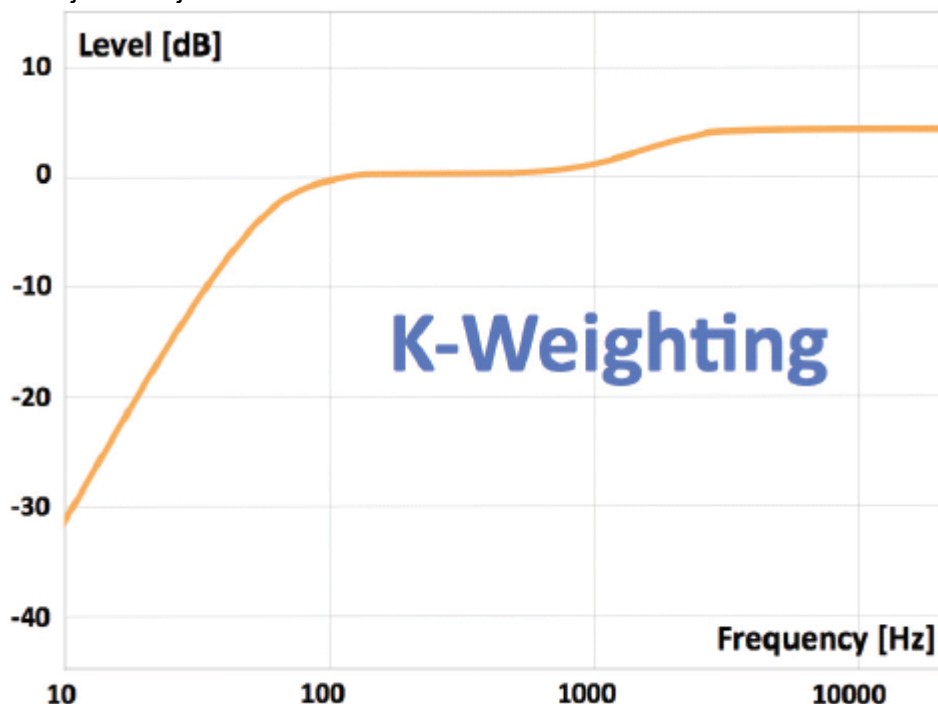
Tyto praktické pokyny jsou koncipované jako „**živý dokument**“, do něž se časem dostanou zkušenosti vysílatelů, které poskytnou další informace a rady k této zásadní změně způsobu zpracování a vyvažování zvukových signálů.

Prosím, vezměte v úvahu, že mnoho standardních dokumentů čas od času prochází revizemi, což se týká i tohoto dokumentu. Silně se doporučuje, abyste kontrolovali, zda máte poslední verzi.

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

EBU R 128 zavádí předvídatelnou a dobře definovanou metodu měření úrovně hlasitosti pro zpravodajství, sport, reklamy, televizní hry, hudbu, upoutávky, filmy, atd. v celém vysílacím řetězci a tím pomáhá profesionálům vytvořit robustní technické specifikace pro vnímání, výrobu, přehrávání a distribuci na mnoha různých platformách. Dokument R 128 je založen výlučně na otevřených standardech a snaží se harmonizovat způsob, kterým vyrábíme a měříme zvuk v mezinárodním měřítku.

Základem pro R 128 je norma **ITU-R BS.1770**, výsledek rozsáhlé práce Mezinárodní telekomunikační unie. Účelem této normy bylo zavést dohodnutý otevřený algoritmus pro měření hlasitosti a skutečných špičkových úrovní pořadů. Jedná se o robustní normu, která má výhodu jednoduché implementace. Ve stručnosti definuje křivku filtru „**váhování K**“ (upravený hornopropustný filtr druhého řádu, viz obr. 1), jež představuje základ pro sladění inherentně subjektivního dojmu s objektivním měřením.



Level = úroveň
K-Weighting = váhování K
Frequency = frekvence

Obr. 1: Křivka filtru „váhování K“ pro měření hlasitosti

Křivka vážení se aplikuje na všechny kanály (kromě kanálu nízkofrekvenčních efektů (LFE), který je v současnosti vyřazen z měření, viz níže), vypočítá se celková střední kvadratická hodnota úrovně (s různými faktory zesílení pro přední a prostorové kanály; viz obr. 2) a výsledek se zobrazí jako „LKFS“ (hlasitost (L) vážená K vztahovaná k digitální plné stupnici (FS)) nebo „LUFS“¹ (jednotka hlasitosti vztahovaná k digitální plné stupnici). Pro relativní měření se používá jednotka hlasitosti (LU), kde se 1 LU rovná 1 dB.

¹ EBU doporučuje používat jednotku „**LUFS**“ (jak je definována v Tech. dok. EBU 3341). „**LUFS**“ je ekvivalentem „**LKFS**“ a překonává nekonzistentnost mezi ITU-R BS.1770 a ITU-R BS.1771. „**LUFS**“ také vyhovuje mezinárodní terminologické normě ISO 80000-8 [6].

Kanál nízkofrekvenčních efektů (LFE)

Kanál nízkofrekvenčních efektů (kanál “.1” v “5.1”) vícekanálového zvukového signálu se v současné době nezohledňuje při měření hlasitosti podle ITU-R BS.1770. To může vést ke zneužívání LFE zbytečně vysokými úrovněmi signálu. Probíhající průzkumy se snaží vyhodnotit subjektivní vliv, který má LFE na vnímání hlasitosti a vhodný způsob zařazení tohoto kanálu do objektivního měření hlasitosti.

| | | | | | |
|----------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Vážení K | Levý | Pravý | Střed | Levý prostorový | Pravý prostorový |
| | Pre filtr | Pre filtr | Pre filtr | Pre filtr | Pre filtr |
| | Filtr RLB | Filtr RLB | Filtr RLB | Filtr RLB | Filtr RLB |
| | Střední kvadratická hodnota | Střední kvadratická hodnota | Střední kvadratická hodnota | Střední kvadratická hodnota | Střední kvadratická hodnota |
| | 0 dB | 0 dB | 0 dB | + 1,5 dB | + 1,5 dB |
| | Σ | | | | |
| | | | 10LOG_{10} | | |
| | | | Hlasitost pořadu | | |

Obr. 2: Zpracování kanálů a sumarizace v ITU-R BS.1770

Zatímco BS.1770 definuje metodu měření, dokument **R 128** ji rozšířil tím, že skutečně definoval konkrétní „**cílovou úroveň**“ pro normalizaci hlasitosti a **hradlovací** metodu ke zlepšení přizpůsobení hlasitosti pořadů, které obsahují delší pasáže ticha nebo izolované promluvy. Vývoj EBU byl potřeba k tomu, aby se vyšlo vstříc výrobcům pořadů s obzvláštním ohledem na zajištění prostředku měření kompletních zvukových mixů (spíše než jednoho komponentu jako je řeč nebo hudba) a rozsahu hlasitosti pořadu.

K tomuto účelu EBU zavedla tři nové parametry:

- **Hlasitost pořadu**
- **Rozsah hlasitosti**
- **Skutečná špičková úroveň**

2.1 Hlasitost pořadu

Hlasitost pořadu popisuje dlouhodobou **integrovanou** hlasitost po dobu trvání pořadu². Tento parametr se skládá z jednoho čísla (v LUFs s jedním desetinným místem), které udává, „jak hlasitý je pořad v průměru“. Tato hlasitost se měří pomocí měřiče, jenž vyhovuje ITU-R BS.1770 a obsahuje **hradlovací** funkci. Toto hradlo slouží k tomu, že přeruší měření hlasitosti, když signál poklesne pod určitou prahovou úroveň. Bez této hradlovací funkce by pořady s dlouhými pasážemi ticha, zvukem nebo hlukem nízké úrovně na pozadí získaly příliš nízkou hodnotu úrovně integrované hlasitosti. Následně po normalizaci by takové pořady byly příliš hlasité.

Po sérii testů bylo EBU odsouhlaseno hradlo o hodnotě -8 LU (jednotka hlasitosti; 1 LU \equiv 1 dB) vzhledem k měření LUFs bez hradla (další podrobnosti, viz [7]). Následně byl do ITU odeslán návrh na zahrnutí metody relativního hradlování do BS.1770. V ITU-R BS.1770-2 je teď hradlo využívající relativní prahovou hodnotu součástí měřicího algoritmu pro integrovanou hlasitost, i když s mírně nižší prahovou hodnotou **-10 LU**. Tuto prahovou hodnotu přebírá EBU R 128 a všechny související dokumenty.

(O podrobnosti funkce hradlování se potřebují zajímat pouze výrobci. Z pohledu uživatele je rozdíl malý, ale doporučuje se, aby uživatelé obnovovali svoje zařízení, aby byla zajištěna konzistentnost měření.)

Vedle jiných poznatků i poslechové testy potvrdily volbu **cílové úrovně**, na níž se bude každý zvukový signál normalizovat, a to:

-23,0 LUFs (hradlo -10 rel)

-23 LUFs je novým středem vesmíru vyrovnávání úrovní zvuku!!!

Proč -23 LUFs?

Průzkumy a měření úrovně hlasitosti skutečných vysílacích stanic ukázaly průměrnou úroveň hlasitosti přibližně **-20 LUFs** (s mnoha výjimkami...). Další vstupní údaj do diskuse přišel z ITU ve formě dokumentu ITU-R BS.1864 „*Provozní praxe v oblasti hlasitosti při mezinárodní výměně digitálních televizních pořadů*“ [8]. V BS.1864 se doporučuje cílová úroveň **-24 LUFs**, i když se nepředpokládá zařazení funkce hradla do BS.1770. Neformální testy provedené členy skupiny

² Pojem „**pořad**“ zahrnuje také reklamy, upoutávky, apod. Aby se předešlo pochybnostem, reklamy, které se obvykle umísťují mimo vysílací čas a do vysílacího času toho, co se obecně považuje za „pořad“, se považují za specifické pořady samy o sobě (rovněž jednotlivé reklamy v reklamním bloku jsou samostatné pořady); tím se zjednodušuje jejich integrace s delšími pořady. Výrobci obou typů pořadu evidentně nemohou vědět, s čím se jejich pořad bude střídat, a tak se každý typ musí posuzovat zvlášť. V tomto dokumentu se pojem „**pořad**“ vztahuje na pořad tak, jak je dokončený z výroby, a ne na kombinaci pořadů, upoutávek a reklam, která se dostane do přijímače diváka nebo posluchače v průběhu celého vysílacího času pořadu.,

PLOUD EBU ukázaly, že rozdíl v měření hlasitosti s relativním hradlem -10 LU a bez něj u pořadů s malým až středním rozsahem hlasitosti činí 0 – 1 LU. **-23 LUFS s** hradlem se tedy v mnoha případech téměř rovná **-24 LUFS bez** hradla. **-23 LUFS** (s hradlem) se tedy začalo považovat za *nejnižší možnou úroveň hlasitosti pořadu*, aniž by se přechod z průměrné úrovně -20 LUFS ještě více ztěžoval. Protože panoval názor, že -20 LUFS neposkytne dostatečnou rezervu pro dynamický obsah, bylo přijato rozhodnutí, že se hodnota ustálí na **-23.0 LUFS**.

Odchylka **±1,0 LU** je přijatelná pro pořady, kde není přesná normalizace na cílovou úroveň **-23,0 LUFS** prakticky dosažitelná (jako jsou živé pořady nebo pořady, které mají mimořádně krátký obrat). V případech, kdy jsou úrovně jednotlivých signálů pořadu do velké míry *nepředvídatelné*, kde se pořad skládá jen ze zvuků pozadí (např. hudebního doprovodu pro předpověď počasí), nebo kde jsou pořady úmyslně namixované na nižší úroveň, může ležet úroveň hlasitosti pořadu mimo toleranci. Takové případy by ovšem měly být stále vzácnější.

2.2 Rozsah hlasitosti

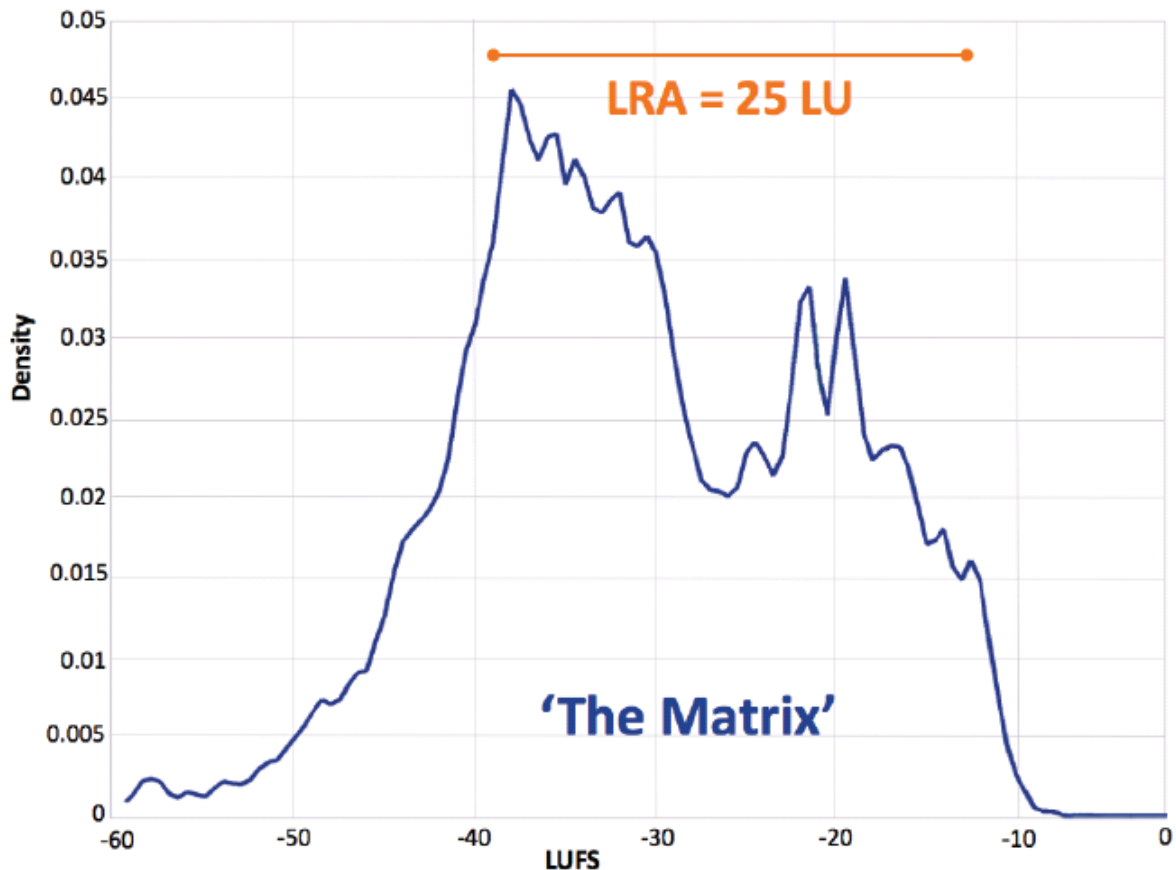
Dalším hlavním tématem byl rozsah hlasitosti, který bude potřeba k pokrytí *všech* pořadů (za předpokladu, že nepřekračují přípustný rozsah hlasitosti pro domácí poslech). Parametr **rozsah hlasitosti (LRA)** kvantifikuje (v LU) kolísání výsledků měření hlasitosti pořadu. Je založen na statistickém rozložení hlasitosti v rámci pořadu, čímž vylučuje extrémny. Proto např. jeden výstřel z pistole nezkrusí výsledek výpočtu LRA. Doporučení EBU R 128 nestanovuje maximální přípustnou hodnotu LRA, protože závisí na faktorech, jako je rozsah tolerance průměrného posluchače stanice, rozložení žánrů stanice, apod. Nicméně R 128 silně **podporuje používání LRA** k určení, zda je potřeba zpracování dynamiky zvukového signálu, aby byl signál v souladu s požadavky konkrétního přenosového kanálu nebo platformy. Další podrobnosti o LRA najdete v Technickém dokumentu EBU 3342.

Rozsah hlasitosti je obecný parametr, který pomáhá při rozhodnutí, zda je potřeba komprese dynamiky.

První zkušenosti vysílacích stanic naznačují, že maximální hodnota LRA je přibližně **20 LU** pro vysoce dynamický materiál, jako jsou akční filmy nebo vážná hudba. Většina pořadů nikdy nebude potřebovat plně využívat tak vysokou hodnotu LRA a vlastně jí ani nedosáhne!

Pro velmi krátké pořady (<30 s), jako jsou reklamy nebo upoutávky, může nastavení limitu pro **maximální hodnoty okamžité** nebo **krátkodobé** úrovně hlasitosti³ poskytnout lepší způsob kontroly dynamických vlastností jako určitá „druhá obranná linie“ (viz § 7, § 10).

³ **Maximální okamžitá úroveň hlasitosti** (Max ML) je nejvyšší hodnota (v LUFS) okamžité úrovně hlasitosti zvukového signálu (doba integrace 400 ms). **Maximální krátkodobá úroveň hlasitosti** (Max SL) je nejvyšší hodnota (v LUFS) krátkodobé úrovně hlasitosti zvukového signálu (doba integrace 3 s).



Density = hustota

Obr. 3: Rozsah hlasitosti (LRA) jako výsledek statistického rozložení úrovní hlasitosti

Obr. 3 ukazuje rozložení hlasitosti a LRA filmu „Matrix“; hodnota 25 LU bude pravděpodobně náročná pro většinu obyčasných pokojů ...

2.3 Skutečná špičková úroveň (TPL), maximální přípustná TPL

V Evropě byl nejrozšířenějším měřicím přístrojem (a dosud ve velkém měřítku je) **Quasi Peak Programme Meter (QPPM)** měřič kvazišpičkové úrovně; doba integrace = 10 ms). S přechodem na digitální zpracování signálu se objevily **měřiče špiček vzorků**. Zatímco QPPM nemůže zobrazovat krátké špičky (<<10 ms) vzhledem ke své konstrukci, ani měření špiček vzorků nemusí odhalit skutečnou špičkovou úroveň představovanou digitálním signálem.

Digitální zpracování nebo ztrátové kódování může mít za následek, že **špičky jsou mezi vzorky**, a překračují úroveň vzorků uváděnou přístrojem. Ve vysílání je důležité mít spolehlivé měření úrovně na různých platformách a pro různé vzorkovací frekvence. Tento měřič by měl signalizovat ořezání, i když bude špička ležet mezi vzorky, aby bylo možné **předvídat a eliminovat** zkreslení, k němuž může docházet v následných převodnicích digitálního na analogový signál, převodnicích vzorkovací frekvence nebo u běžně používaných kodeků. To měřič špiček vzorků nedokáže a je proto nevhodný pro použití v moderním vysílání (viz *Lund, Th.: „Přestaňte počítat vzorky“* [9]).

Skutečná špičková úroveň udává maximální (pozitivní nebo negativní) hodnotu signálu v kontinuální časové doméně: tato hodnota může být vyšší než nejvyšší hodnota vzorku v doméně časově omezených vzorků. Pomocí převzorkování v měřicích skutečných špiček vyhovujících požadavkům ITU-R BS.1770 je teď možné tyto skutečné špičky (symbol jednotky podle ITU-R BS.1770: **dBTP – deciBel** vztažený k digitální plné stupnici měřený měřičem skutečných špiček

(True Peak)) detekovat. Přesnost závisí na frekvenci převzorkování. Pouze je potřeba nechat rezervu 1 dB pod dBFS, aby zbyl prostor pro potenciální chybu spočívající v zobrazení nižší než skutečné hodnoty asi o 0,5 dB (u měřiče skutečných špiček s 4x převzorkováním; základní vzorkovací frekvence: 48 kHz).

Měřiče špiček s převzorkováním poskytují dobrý odhad skutečných špiček zvukového signálu. Měřiče špiček vzorků ne.

V důsledku toho bude **maximální přípustná skutečná špičková úroveň** doporučená v R 128:

-1 dBTP

Tuto hodnotu je možné aplikovat ve výrobním prostředí na obecné lineární zvukové signály. Pozor, že některé části řetězce, jako jsou analogoví vysílatelé přebírající signál a uživatelé běžně používaných kodeků na redukci dat, vyžadují nižší skutečnou špičkovou úroveň. Tímto tématem se komplexně zabývají „*Pokyny k distribuci*“ EBU (Tech. dok. EBU 3344).

Shrnutí dokumentu EBU R 128

- Zvukový signál charakterizují parametry „**hlasitost pořadu**“, „**rozsah hlasitosti**“ a **maximální skutečná špičková úroveň**;
- **Úroveň hlasitosti pořadu** bude normalizovaná na **-23.0 LUFS**;
- Tolerance je obecně **±1,0 LU** u pořadů, kde přesná normalizace není prakticky dosažitelná;
- Měření se bude provádět pomocí měřiče vyhovujícího požadavkům **ITU-R BS.1770-2** a **Tech. Dok. EBU 3341** („Režim EBU“ – mimo jiné shrnujících hradlovou metodu popsanou v BS.1770-2);
- Parametr **rozsah hlasitosti** se bude používat jako pomocné měřítko pro rozhodnutí, zda je potřeba dynamická komprese (v závislosti na žánru, cílovém obecnstvu a přenosové platformě);
- **Maximální přípustná skutečná špičková úroveň** ve výrobě je **-1 dBTP**;
- **Metadata** hlasitosti se nastaví tak, aby udávala **-23 LUFS** (u pořadů, které byly normalizovány na tuto úroveň podle doporučení); metadata hlasitosti musí vždy udávat **správnou** hodnotu pro hlasitost pořadu, i kdyby z jakéhokoli důvodu pořad na -23 LUFS normalizován nebyl.

2.4 Logo R 128

EBU představila oficiální logo R 128, které se skládá z číslic 1, 2 a 8 – tvořících šťastnou, usměvavou tvář:

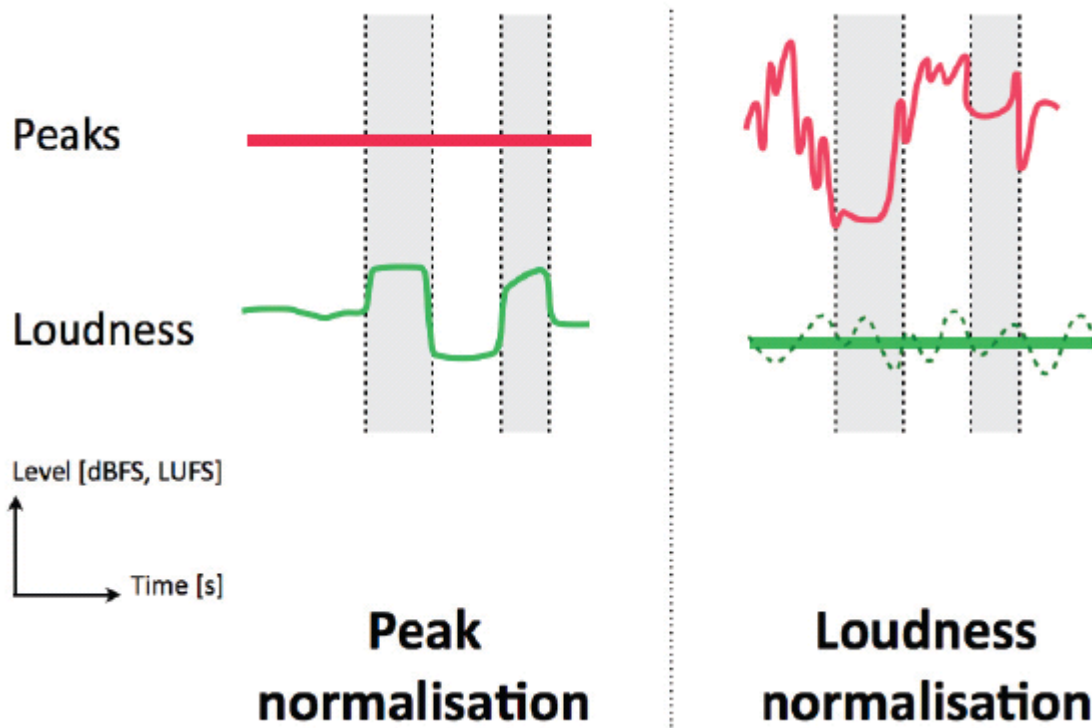


Logo mohou používat (při splnění určitých předpokladů) výrobci k doložení shody s „režimem EBU“.

3. Obecná koncepce normalizace hlasitosti

3.1 Špička versus hlasitost

Stále silně rozšířená koncepce vyrovnávání zvuku spočívající v *normalizaci špiček* s odkazem na přípustnou maximální úroveň (PML; např. -9 dBFS) zajistila stejnoměrné špičkové úrovně pořadů, ale úrovně hlasitosti zůstaly silně variabilní. Vlastní variabilita závisí na stupni dynamické komprese signálu. Naopak **normalizace hlasitosti** dosahuje **stejněměrné průměrné hlasitosti pořadů**, přičemž špičky kolísají podle obsahu a podle uměleckých a technických potřeb (viz obr. 4). Za předpokladu, že rozsah hlasitosti pořadu bude v rámci přípustné tolerance, **získá posluchač stejnoměrnou průměrnou úroveň hlasitosti u všech pořadů, takže již nebude muset často upravovat hlasitost pomocí dálkového ovládání.**



Peaks = špičky
 Loudness = hlasitost
 Level = úroveň
 Time = čas
 Peak normalisation = normalizace špiček
 Loudness normalisation = normalizace hlasitosti

Obr. 4: Normalizace špičkové úrovně versus normalizace úrovně hlasitosti série pořadů

Opět to **NEZNAMENÁ** že úroveň hlasitosti musí být konstantní *v rámci* jednoho pořadu, právě naopak! Také to **NEZNAMENÁ**, že *jednotlivé* komponenty pořadu (např. pre mixy nebo mixy svazků kanálů a verze hudby s efekty nebo izolovaná hlasová stopa) musí být všechny na stejné úrovni hlasitosti!. Změny hlasitosti jsou umělecký nástroj, takže koncepce normalizace hlasitosti podle R 128 ve skutečnosti **podporuje dynamičtější mixování!** Normalizuje se **průměrná, integrovaná hlasitost celého** pořadu.

3.2 Normalizace signálu versus metadata

Existují v zásadě *dva způsoby* jak dosáhnout normalizace pro spotřebitele: prvním z nich je **normalizace** vlastního **zvukového signálu**, aby pořady byly stejně hlasité svým provedením – druhý způsob spočívá v **používání metadat hlasitosti**, které popisují, jak hlasitý pořad je. U druhého způsobu se skutečné průměrné úrovně hlasitosti pořadů nemusí měnit na normalizovanou hodnotu a přesto se mohou hodně lišit pořad od pořadu. Při použití moderního vybavení se normalizace může provést na spotřebitelském konci za použití individuálních hodnot metadat hlasitosti, pomocí nichž se upraví zesílení pořadů na stejnou úroveň přehrávání.

Stejnou hlasitost lze dosáhnout normalizací zvukového signálu nebo použitím metadat hlasitosti.

V rámci principu vyrovnávání hlasitosti EBU R 128 se podporuje *první* řešení vzhledem k následujícím výhodám:

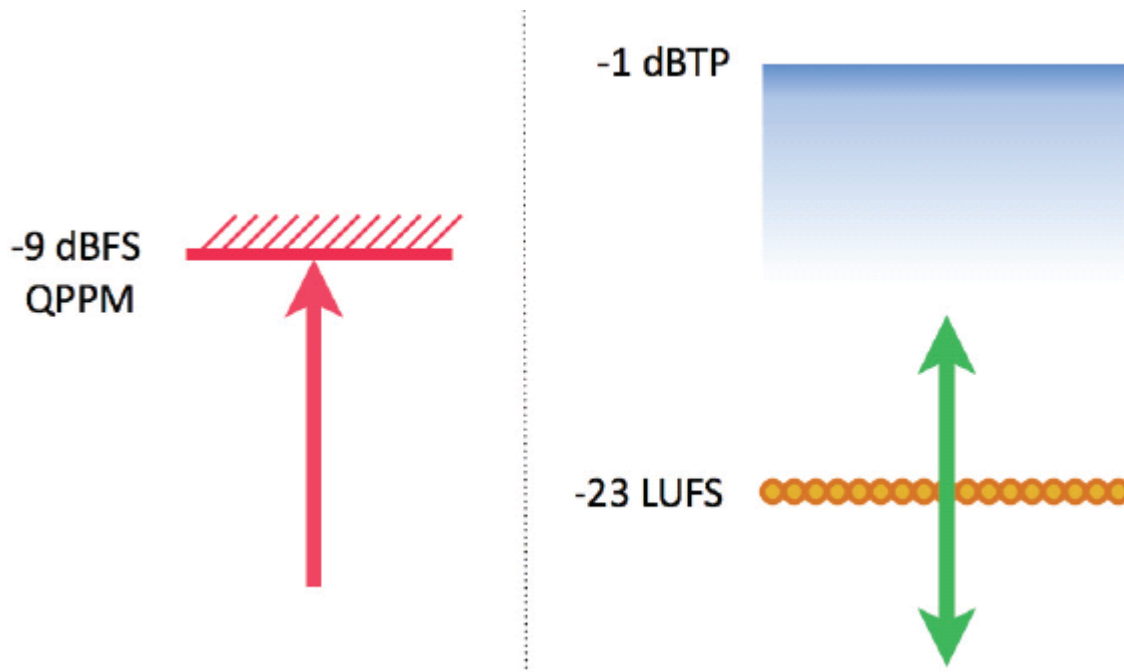
- jednoduchosti a
- potenciálnímu zvýšení kvality zvukového signálu.

Druhé řešení není zakázané (viz též dokument „*Pokyny k distribuci*“, Tech. dok. EBU 3344), ale mít jedno jediné číslo (**-23 LUFS**) znamená velkou sílu při šíření koncepce vyrovnávání hlasitosti, protože je snadno pochopitelné a dobře se s ním pracuje. A aktivní normalizace zdroje svým způsobem „trestá“ překomprimované signály a tím automaticky nabádá pracovníky ve výrobě, aby přemýšleli o jiných, dynamičtějších a kreativnějších způsobech jak zapůsobit svým pořadem. Jinými slovy vlastní *technická* změna úrovně zvukového signálu **aktivní normalizací na -23 LUFS** má přímé důsledky pro *umělecký* proces – a to pozitivní! Výrobní straně se tak může ulevit od vedení „hlasitostní války“ – což je nešťastný a široce rozšířený výsledek principu normalizace špiček.

Nicméně je nutno uvést, že obě metody se mohou doplňovat, nemělo by se na ně pohlížet jako na soupeřící metody ani je vnímat černobílým pohledem na stejný problém. Oba přístupy jsou součástí R 128 – ale vzhledem k výše uvedeným výhodám se doporučuje **normalizace zvukového signálu**.

Normalizace zvukového signálu se doporučuje v oblasti výroby vzhledem k její jednoduchosti a potenciálu zlepšování kvality.

Snaha o jednotnou úroveň hlasitosti představuje **celou novou koncepci** mixování, vyrovnávání, obecně práce se zvukem. Zatímco omezovač špiček nastavený na přípustnou maximální úroveň (obvykle -9 dBFS, měřenou pomocí QPPM) poskytoval určitý „*bezpečnostní strop*“, kde bez ohledu na to, jak jste zabrali, měli jste vždy tu „správnou“ maximální úroveň, model vyrovnávání hlasitosti spíše připomíná „*vznášení se v prostoru s otevřenou oblohou nahoře*“ (viz obr. 5).



Obr. 5: Normalizace kvazišpičkové úrovně („bezpečnostní strop“) versus normalizace úrovně hlasitosti

S normalizací a měřením hlasitosti je bezpečnostní strop pryč. To může někoho děsit, protože svým způsobem bylo „pohodlné“, že jste nemuseli sledovat úrovně tak bedlivě – omezovač na konci řetězce vždy zajistil zkrocení vašeho výstupu. Ale vedlejším účinkem bylo, že úrovně hlasitosti šly nahoru, protože propracovanější procesory vykazovaly méně artefaktů dynamické komprese a omezování.

Vyrovňávání hlasitosti na druhou stranu podporuje používání zdaleka nejlepšího měřicího zařízení: **ucha**. To s sebou nese pozornější mixování a zkvalitňování zvuku. Zkušenosti několika členů EBU ukázaly, že práce s modelem hlasitosti je **osvobozující** a uspokojující. Boj o to, „kdo bude silnější“ je pryč, celkové úrovně se snižují a v kombinaci s vyšší přípustnou skutečnou špičkovou úrovní (**-1 dBTP**) to znamená potenciálně **dynamičtější mixy s konzistentnější hlasitostí** v rámci pořadu. Dynamická komprese se opět stala uměleckým nástrojem a ne zbraní dávající vyšší hlasitost, **kvalita zvuku se zvyšuje!**

Návrat „**mixování podle poslechu**“ je vítanou úlevou, která měla přijít už dáno. Zvukový technik je teď motivován, aby mixoval jen podle poslechu (další účinek měření hlasitosti) – po nastavení úrovní při pevné hlasitosti monitoru (viz § 8).

Vyrovňávání hlasitosti motivuje k mixování „pouze podle poslechu**“ - po nastavení úrovní na monitoru s pevnou hlasitostí.**

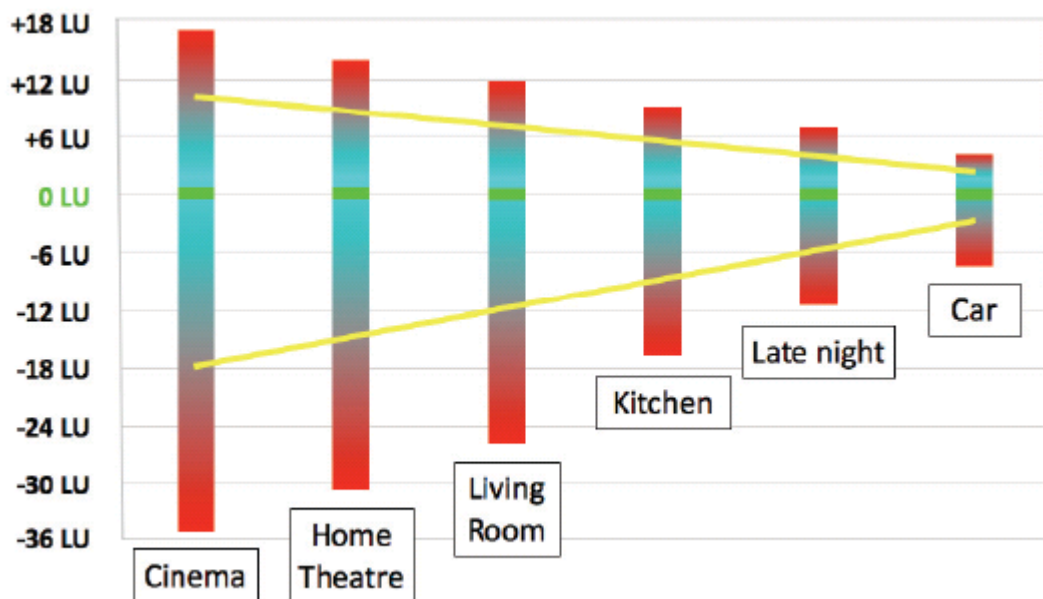
Na dalším článku v řetězci po výrobě je vysílatel konfrontován s potřebou normalizovat rozmanitý obsah pocházející z různých míst. Zvláště během přechodového období bude ještě existovat mnoho pořadů, které dosud nebudou mít normalizovanou hlasitost. Pro tyto pořady se musí vyvinout strategie, jako je **automatizovaná normalizace** přímo po načtení do přehrávacího serveru nebo instalace bezpečnostního zařízení pro regulaci hlasitosti na výstupu hlavní řídicí místnosti, aby se mohly upravit např. živé přenosy, jež nejsou vyrobeny s použitím cílové úrovně-**23 LUFS**.

Tyto a podobné záležitosti budou s dalšími podrobnostmi probrány později (§ 4.3, § 6).

3.3 Používání parametru rozsah hlasitosti

Pomocí parametru **rozsah hlasitosti (LRA)** je teď možné *kvantifikovat* dynamiku pořadu. V minulosti se muselo o tom, zda pořad zapadne do okna tolerance hlasitosti cílových posluchačů, rozhodovat formou „poučeného odhadu“ zkušených zvukových techniků. Díky používání rozsahu hlasitosti odhadování skončilo – po skončení intervalu měření (obvykle celého pořadu) stačí mixážnímu technikovi / obsluhujícímu pracovníkovi jediné číslo k rozhodnutí, zda je potřeba další dynamické zpracování.

Důležité je pochopit, že je nemožné definovat jednu maximální hodnotu LRA pro všechny vysílací společnosti a všechny pořady. **Individuální maximální LRA** závisí na žánru (žánrech) (tématické kanály s velmi jednotvárným obsahem, jako jsou třeba zpravodajské kanály, jistě nebudou mít tak vysoký maximální LRA jako veřejnoprávní televize s velkou rozmanitostí žánrů např. pro pořady, jako je koncert vážné hudby). Maximální hodnota LRA může být také odlišná pro distribuční platformy, jako je mobilní vysílání, a pro různá prostředí přehrávání (viz obr. 6; vzdálenost mezi žlutými čarami udává příklady různých hodnot rozsahu hlasitosti). Volbu maximálních hodnot LRA stanice pro specifické pořady ovlivňuje průměrné poslechové prostředí, věk cílového publika, „zóna poslechového komfortu“ spotřebitelů a další parametry. **Model řízení rozsahu hlasitosti** vychází z obecné maximální přijímané hodnoty rozsahu hlasitosti podle výše uvedených zásad a postupně tuto hodnotu upravuje tak, aby vyhovovala technickým požadavkům jednotlivých distribučních platform a prostředí přehrávání.



Kino Domácí kino Obývací pokoj Kuchyň Pozdě v noci Auto

Obr. 6: Příklady různých rozsahů hlasitosti podle poslechového prostředí

Za stanovení hodnoty LRA pořadu odpovídá zvukový technik ve fázi výroby pořadu; doporučuje se, aby zvukový technik dodržoval tento princip. V prostředích, kde není ovládání nebo zásah člověkem možný, nebo se nepoužívá, může měření LRA vést k používání vhodných předvoleb dynamického procesoru, který byl nakonfigurován pro konkrétní žánr zvukového signálu. Doporučuje se snažit se o situaci, kdy zvukový technik může ovlivnit LRA podle požadavků vysílatele, protože tím se potenciálně zvyšuje kvalita zvuku.

V důsledku potřeby různých hodnot rozsahu hlasitosti **EBU R 128** *neobsahuje* maximální přípustnou hodnotu LRA, ale místo toho silně doporučuje **používat** parametr **rozsahu hlasitosti** k vyhodnocování potenciální potřeby zpracování dynamického rozsahu podle různých výše uvedených kritérií.

Rozsah hlasitosti je také užitečný **indikátor** potenciálních *procesů omezování dynamiky* v signálovém řetězci, které se provádějí účelově nebo nechtěně. Když bude hodnota LRA pořadu např. potom, co prošel řetězcem zpracování, nižší než původně, došlo k takovému omezení.

3.4 Výstup na skutečnou špičku

Třetí opatření, které R 128 doporučuje, se týká **maximální skutečné špičkové úrovně** zvukového signálu. I když byl model normalizace špičkové úrovně opuštěn, samozřejmě je **nadále důležité měřit a upravovat špičky pořadu** a obzvláště jeho maximální špičku, aby nedocházelo k přehlčení a zkreslení.

Měřič hlasitosti s „**režimem EBU**“ (viz *Tech. dok. EBU 3341*) je také vybaven funkcí měření a zobrazení skutečných špičkových úrovní pořadu. Bezpečnostní omezovače předcházející přemodulování budou muset fungovat v *režimu skutečných špiček* a budou se muset nastavit na příslušnou maximální skutečnou špičkovou úroveň jak ve výrobě, tak na výstupu hlavního ovládní, na konci distribuce a na vysílači. Vedle maximální skutečné špičkové úrovně pro obecné signály PCM ve výrobě (-1 dBTP) uvádí Tech. dokument EBU 3344 („*Pokyny k distribuci*“) další doporučení pro různé aplikace.

4. Strategie pro normalizaci hlasitosti

4.1 Produkce (výroba), postprodukce

Přístup k normalizaci hlasitosti v těchto oblastech nabízí dvě možnosti: první z nich spočívá v zachování současných vyrovnávacích postupů a posunutí úrovně následně a druhý spočívá v přechodu na **řízení hlasitosti** a **normalizaci**, přičemž následně už není potřeba posouvat úroveň nebo stačí jen malý posun (obr. 7).

První přístup je vhodnější pro rané fáze přechodu a je možná zvláště užitečný pro ty, kteří pracují s **živými přenosy**. Stávající měřiče, omezovače a mixovací postupy se zachovávají a na výstupu mixážního pultu (po hlavních měřičích) se provede posun úrovně, aby bylo dosaženo **cílové úrovně -23 LUFS**.

| 1 | 2 |
|-----------------------------------|--|
| Zachování vyrovnávacích zvyklostí | Přechod na normalizaci úrovně hlasitosti |
| Provedení posunu | Žádný posun není potřeba |

Obr. 7: Dvě základní pracovní metody k dosažení stejnoměrné hlasitosti v produkci a postprodukci

Po posuvu úrovně se umístí měřič hlasitosti, aby technici znali přesnou velikost posunu (který je zpočátku stále trochu založen na odhadu). V každém případě je zapojení **měřiče hlasitosti paralelně** s konvenčním měřičem dobrý nápad. Tímto způsobem lze získat zkušenosti, než se člověk skutečně ponoří do světa vyrovnávání hlasitosti. Dále může měření starších pořadů stejného žánru pomocí měřiče hlasitosti poskytnout dobrou orientaci, kde se nacházejí příslušné úrovně.

U pořadů, které se dokončují v **postprodukci**, se posun úrovně z přístupu **1** provádí snadno. Při změření celého pořadu najednou je možné přesně určit potřebný posun zesílení a v dnešním bezpáskovém světě je výpočet zesílení velmi rychlou a snadnou záležitostí.

Samozřejmě u **živých pořadů** je dosažení cílové úrovně náročné (pokud to není přímo otázka štěstí). Proto je u pořadů, kde není přesná normalizace na cílovou úroveň -23 LUFS prakticky dosažitelná (vedle živých pořadů se jedná např. o pořady s velmi krátkým obratem) přípustná odchylka $\pm 1,0$ LU. První zkušenosti v NDR (Severoněmecká televize), ORF (Rakouská televize) a RTBF (Belgická televize – francouzská část) ukázaly, že je rozhodně možné, aby se živé mixy vešly do tolerance ± 1 LU, kterou připouští EBU R 128.

Je přípustná tolerance ± 1 LU od cílové úrovně (-23 LUFS). Předpokládají se výjimky ve speciálních případech.

V případech, kde jsou úrovně jednotlivých signálů pořadu do velké míry *nepředvídatelné*, kde je pořad úmyslně sestaven pouze z prvků pozadí (např. hudebního doprovodu u předpovědi počasí) nebo kde je na základě dramaturgického záměru úroveň hlasitosti výrazně nižší, než by byla žádoucí cílová úroveň, může být tato tolerance příliš těsná. Proto se v těchto případech předpokládá, že integrovaná úroveň hlasitosti může ležet mimo toleranci stanovenou v R 128.

U **řešení vyrovnávání 1** (zachování současných vyrovnávacích postupů) je pravděpodobné, že téměř ve všech případech bude potřebný posun zesílení *negativní* (utlumení). Z toho důvodu nebude další krok snižování dynamického rozsahu a/nebo omezování maximální skutečné špičkové úrovně obvykle potřeba. Potenciální utlumení ve velké většině případů je také důvodem pro to, že řešení s metadaty popsané v § 3.2 se pro tento přístup nedoporučuje.

Řešení vyrovnávání 2 (přímý přechod na normalizaci hlasitosti) je tím řešením, které tyto Praktické pokyny **doporučují**. Opět platí, že po počátečním období měření a testování starších pořadů a instalaci měřiče hlasitosti paralelně se stávajícím měřičem (obvykle QPPM) budou výhody modelu měření hlasitosti mluvit samy za sebe. Větší možný dynamický rozsah bude vítaným bonusem např. pro hluk davu u sportovních pořadů, čímž se umocní působení zápasu na diváky a posluchače. Studiové komentáře přes původní zvuk, které jsou často dynamicky komprimované z uměleckých důvodů (a kde bude proto poměr hlasitosti ke špičce nižší) budou lépe vyvážené s dynamičtější originálním zvukovým záznamem, atd., atd.

V dalším textu probereme vliv práce s měřičem hlasitosti v produkci a postprodukci.

4.2 Měření hlasitosti v produkci a postprodukci

Měřič hlasitosti s režimem EBU⁴ tak, jak je definovaný v Tech. dok. EBU 3341 nabízí 3 odlišné časové stupnice:

- **Okamžitá** hlasitost (zkratka „M“) – časové okno: **400 ms**
- **Krátkodobá** hlasitost (zkratka „S“) – časové okno: **3 s**
- **Integrovaná** hlasitost (zkratka „I“) – od „začátku“ do „konce“

Časová okna M a S⁴ jsou určena k okamžitému vyrovnávání a mixování zvukových signálů. Počáteční nastavení úrovně se může nejlépe provést pomocí měřiče okamžité hlasitosti, kdy se upraví úroveň klíčových nebo kotvících prvků (jako je hlas, hudba nebo zvukové efekty) zhruba na cílovou úroveň **-23 LUFS**. Zvukový technik samozřejmě musí kdykoli vědět, jak je zvukový signál silný, a to je hlavním účelem okamžitého a krátkodobého měření (obě jsou *nehradlovaná*).

Z důvodu nesouladu mezi ITU-R BS.1770 a ITU-R BS.1771, navrhuje Tech. dokument EBU 3341 odlišnou konvenci označení v souladu s normou ISO 80000-8:

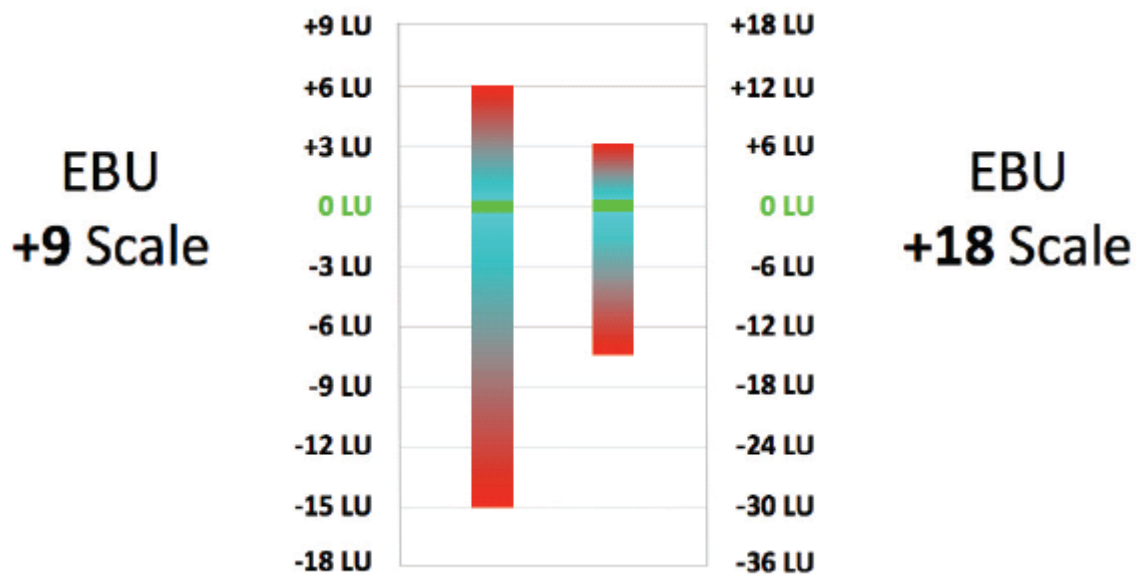
- Symbol „úroveň hlasitosti s vážením K“ by měl být „L_K“.
- Jednotka „LUFS“ udává hodnotu L_K s odkazem na digitální plnou stupnici.
- Jednotka „LU“ udává L_K bez přímého absolutního odkazu, a tudíž popisuje rozdíly v úrovni hlasitosti.

Případné podrobnosti grafického nebo uživatelského rozhraní měřiče hlasitosti s „režimem EBU“ nebyly úmyslně stanoveny; nicméně byly definovány dvě stupnice: „**stupnice EBU +9**“, která by měla být vhodná pro většinu pořadů a „**stupnice EBU +18**“, která může být potřeba pro pořady s širokým LRA. Obě stupnice mohou zobrazovat buď relativní úroveň hlasitosti v LU nebo absolutní úroveň v LUFS. „**0 LU**“ v „režimu EBU“ se rovná cílové úrovni **-23 LUFS**. Výrobci měřičů ve skupině PLOUD se dohodli, že zavedou soubor parametrů „režimu EBU“, aby bylo zajištěno, že údaje jejich měřičů budou konzistentní.

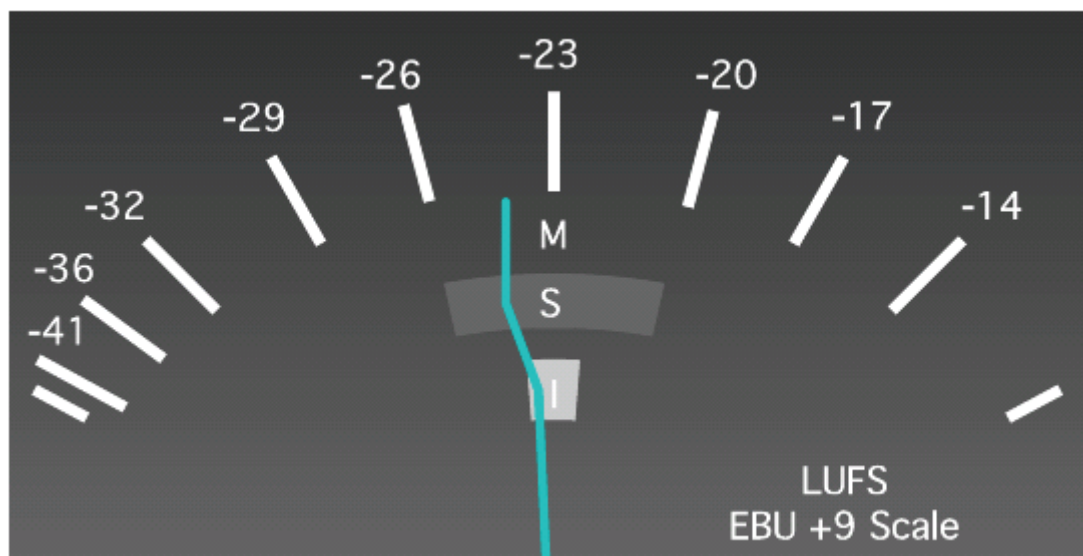
U měřiče hlasitosti s „režimem EBU“ se **0 LU rovná -23 LUFS**.

„Režim EBU“ převzalo mnoho dalších výrobců, nebo se na to připravují. Na obr. 8 je schematické zobrazení měřiče s indikačními pruhy se dvěma stupnicemi režimu EBU. Obrázek 9 ukazuje, jak by mohl vypadat „ručkový“ softwarový měřič.

⁴ „M“ a „S“ jsou zkratky, které se běžně používají ve stereofonii pro „střed“ a „stranu“. Aby se rozlišil jejich význam pro „okamžité“ a „krátkodobé“ integrační časové intervaly, mohou se používat verze „ML_K“ a SL_K (i „IL_K“). „L_K“ znamená „úroveň hlasitosti s vážením K a vyhovuje požadavkům mezinárodní terminologické normy ISO 80000-8.



Obr. 8: Schematické znázornění dvou stupnic hlasitosti (zde v LU) podle popisu v Tech. dokumentu 3341



Obr. 9: Schematické znázornění emulovaného měřič hlasitosti s „zahnutou ručkou“

4.3 „Připravit se, nastavit (úrovně), START!“

Doporučuje se na začátku **nastavovat úrovně** trochu **obezřetně**, protože je psychologicky snazší postupně zvyšovat integrovanou úroveň hlasitosti během mixování, než ji snižovat. Vliv má také funkce hradla: ke snížení průměrné úrovně jsou potřeba větší změny úrovně než k jejímu snížení. A také mírný nárůst v průběhu pořadu bývá obvykle dramaturgicky přirozenější – a „defenzivní“ strategie na začátku dává technikovi prostor k manévrování v případě nepředvídaných signálů nebo událostí.

Jakmile budou nastavené úrovně jednotlivých signálů a provedeno pevné nastavení zesílení monitoru (viz § 8), může zvukový technik přejít na **mixování pouze podle poslechu**. Sledování okamžité nebo krátkodobé úrovně hlasitosti a občasný pohled na hodnotu integrované úrovně hlasitosti by mělo zajistit dostatečné potvrzení toho, že je mix na správné cestě směrem k cílové úrovni. Pomocí numerického zobrazení hodnoty „I“ s přesností na jedno desetinné místo nebo grafického zobrazení se stejným rozlišením je možné předvídat **trendy** a přijímat vhodná opatření. To by se mělo provádět plynule, protože příliš drastické změny jsou ve většině případů umělecky neuspokojivé.

Vzhledem k **maximální přípustné úrovni špičkového signálu -1 dBTP** je teď méně pravděpodobné, že dojde k jevu „naražení do zdi“ (čímž se rozumí bezpečnostní omezovač zasahující na úrovni -9 dBFS). Toto „*otevření víka*“ použité racionálně a s jasným záměrem, společně s normalizací hlasitosti -23 LUFS s sebou přináší **dynamičtější mixy** s méně artefakty dynamické komprese, jako je pumpování, což znamená **celkové zvýšení kvality zvuku!** Výrobci pořadů, kteří dávali v minulosti přednost dynamickým mixům teď již nemusí dělat potenciální kompromisy, protože by jejich pořad zněl měkčeji než více komprimované pořady. Díky normalizaci hlasitosti je tento kompromis pryč. Konečně!

Prvky mixu, které jsou nejdůležitější pro subjektivní dojem stejnoměrné hlasitosti, jsou tzv. zvuky „na popředí“ – jako je hlas, hudba nebo klíčové zvukové efekty. Jednotlivé zvukové prvky skutečně mají silně variabilní rozdíly mezi svou úrovní hlasitosti a špičkovou úrovní. Např. „cinknutí“ dvou sklenic při přípitku má vysokou špičkovou úroveň, ale poměrně nízkou úroveň hlasitosti. Na druhou stranu hardrockový kytarový riff s dynamickou kompresí má téměř stejnou úroveň hlasitosti jako špičkovou úroveň! Když se tyto dva signály vyrovnají podle špiček, bude kytarový riff mnohem hlasitější než cinknutí sklenic. Účelem tohoto příkladu je ilustrovat koncepci, NEZNAMENÁ to, že by se tyto dva signály musely nezbytně mixovat se stejnou hlasitostí! Úroveň jednotlivých prvků a komponentů (jako jsou pre-mixy nebo kmenové mixy a pouze hudební mix a stopa hlasového komentáře) v mixu je přirozeně uměleckým rozhodnutím, ale měření hlasitosti může poskytnout zvukovému technikovi užitečnou vizuální zpětnou vazbu, která vlastně ukazuje to, co slyší!

Když se vrátíme k měření, na **konci pořadu** jsou **dva scénáře**:

- přesně jsme trefili cílovou úroveň (-23.0 LUFS) nebo
- odchýlili jsme se od cílové úrovně v jednom nebo druhém směru

Pochopitelně u **živých pořadů** bude druhý scénář pravděpodobnější. Když bude skutečná úroveň hlasitosti v přijatelné toleranci $\pm 1,0$ LU, není potřeba podnikat žádné další kroky. Když bude úroveň ležet mimo tuto toleranci kvůli zvláště nepředvídatelné povaze pořadu nebo vzácnému výskytu prvků na popředí, je to i tak přijatelné z hlediska obecné produkce (jak bylo uvedeno výše). Mohou být učiněna opatření při dalším zpracování za účelem „zkrocení“ těchto případů ve formě **procesorů hlasitosti**, které pozvolna upraví integrovanou úroveň hlasitosti takových pořadů nerušivým způsobem a mohou fungovat jako určitá „*záchranná síť hlasitosti*“. Takového působení musí být dosaženo použitím adekvátně pomalého reakčního času, aby nedošlo k narušení vnitřní dynamiky produkce. Co se týká individuálního přednastavení takového procesoru hlasitosti nebo místa v signálním řetězci, kde je procesor nainstalovaný, mělo by být možné **rozlišení** mezi živými a souborovými pořady. Procesor může být potřeba pouze pro živé pořady, pokud pracovní tok pořadů v souborech již plně vyhovuje požadavkům EBU R 128. Pokud je následně zařazený procesor dynamiky a hlasitosti umístěn na výstupu **hlavní řídicí místnosti**, mělo by být možné jej **obejít** v případě pořadů splňujících požadavky R 128. Předpokládá se, že takové obcházení se stane normálním pracovním postupem s tím, jak bude stále více pořadů „přesně nastaveno“, protože konečným doporučeným cílem je zpracovávat a **normalizovat samotný zvukový signál**.

V oblasti **postprodukce** je pravděpodobnější, že se trejí cílová úroveň, vzhledem k samotné povaze pracovního procesu s více příležitostmi předělávat a měnit mix, a tedy i úroveň hlasitosti. Navíc obvykle bývá dost času na provedení kompletního integrovaného měření celého pořadu po jeho dokončení i následné **korekce zesílení**. V produkčním bezpáskovém prostředí se tato korekce může provádět mnohem rychleji než v reálném čase. Běžně mohou nastat situace, kdy se mixy v postprodukční oblasti provádějí „jako živé“, tj. přímo na pásku s málo chybami, pokud vůbec, od profesionálních mluvčích (v případě mixování hlasového komentáře). Do této kategorie také spadá např. proces **kopírování na pásku 1:1** s úpravou hlasitosti „za chodu“. Tyto situace jsou potom podobné živé produkci a je třeba k nim tak přistupovat.

Obzvláště v **přechodové fázi** k normalizaci hlasitosti budou výše uvedené procesory hlasitosti zařazené po produkční fázi nepochybně užitečné pro vysílací subjekty v rámci procesu přizpůsobování se systému vyrovnávání hlasitosti a pro zachycení případných odchylných pořadů. Cílem vysílatele (a také zvukových techniků) by mělo být nasazování těchto procesorů co nejméně s tím, jak bude integrovaná úroveň hlasitosti pořadů stále více ležet v rámci přijatelné tolerance hlasitosti. Přesný scénář přechodu, časový horizont a plán implementace budou samozřejmě individuálně odlišné pro každý vysílací subjekt (viz § 9). I když se toto předpokládá, měl by se v zájmu spotřebitele přechod na normalizaci hlasitosti provést v řádné lhůtě, protože prospěch posluchače má zásadní důležitost.

4.4 Rozsah hlasitosti pro produkci a postprodukci

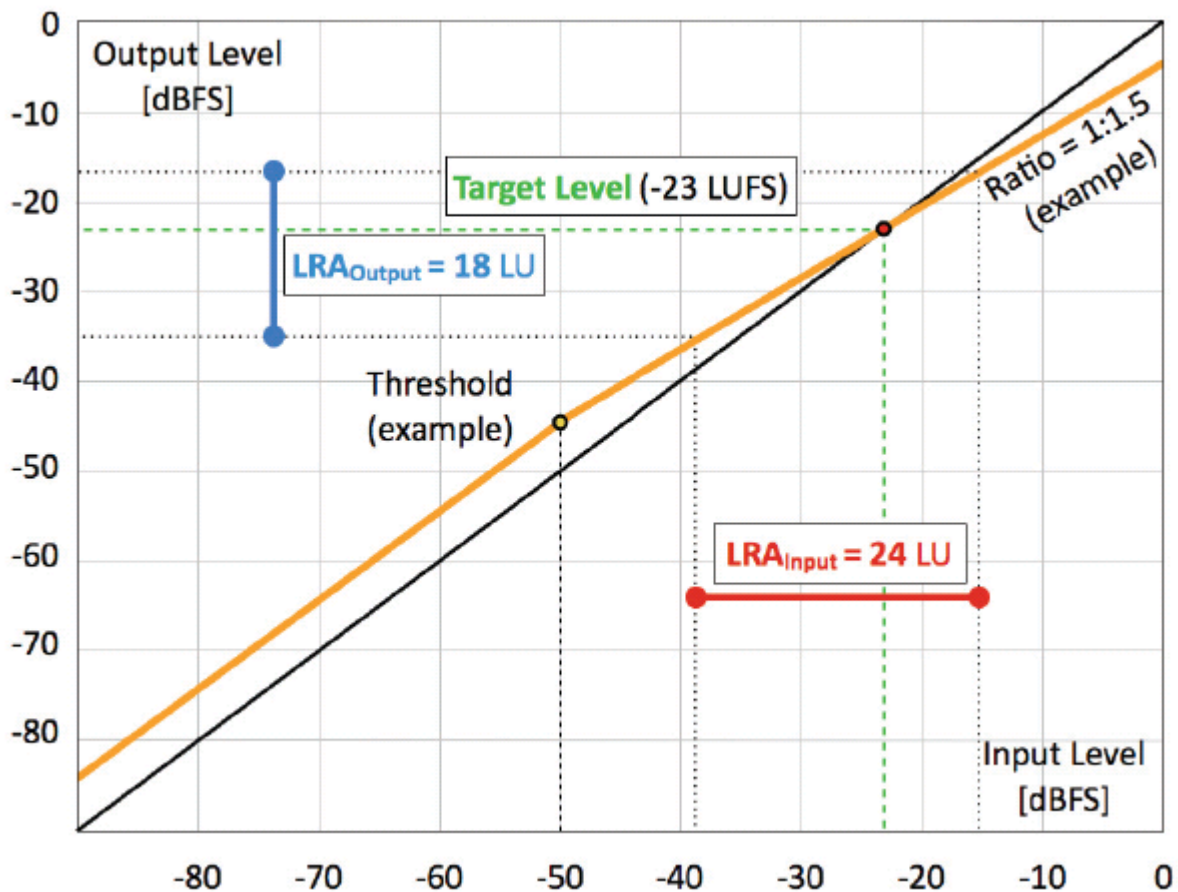
Práce s normalizací hlasitosti hned od začátku také implikuje regulování **rozsahu hlasitosti** (LRA) vzhledem k rozšíření dynamických možností. To je důležité pro dosažení odpovídajícího signálu pro cílové obecnstvo a distribuční řetězec. Zatímco v produkci a postprodukci se může vytvořit „obecný“ mix (s relativně vysokou hodnotou RLA a maximální přípustnou skutečnou špičkovou úrovní -1 dBTP), mohou různé platformy vyžadovat nižší hodnotu RLA a nižší maximální přípustnou špičkovou úroveň (při dodržení úrovně programové hlasitosti na -23 LUFS). **Systém podle R 128 uznává tento obecný přístup s dalším zpracováním v dalších fázích řetězce za účelem přizpůsobení signálu jednotlivým prostředím a platformám.**

Pomocí parametru rozsah hlasitosti je teď systematicky možné stanovovat vhodná měřítka pro *potenciální dynamickou kompresi* pořadu tak, aby se vešel do mezí tolerance publika nebo distribuční platformy. V praxi může **celková komprese nízké úrovně** přinášet uspokojivé výsledky (viz obr. 10 jako příklad): nízká prahová hodnota (< -40 dBFS) a mírný kompresní poměr (1:1,2 – 1:1,5) společně s dlouhou dobou uvolnění (>1 s) zajišťují stejnou úroveň komprese celého rozsahu signálu.

Kompresse LRA (příklad):

- nízká prahová hodnota (< -40 dBFS)
- nízký poměr (1:1,2 - 1:1,5)
- odpovídající úprava **kompensačního (make-up) zesílení**

V závislosti na původní úrovni hlasitosti se může posun na cílovou úroveň **-23 LUFS** provést paralelně pomocí úpravy kompenzačního zesílení kompresoru.



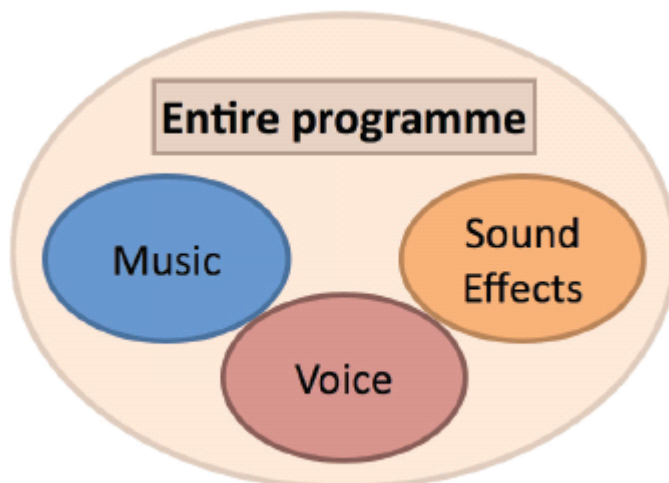
Output level = výstupní úroveň
 Target level = cílová úroveň
 Ratio = poměr
 Example = příklad
 LRA output = LRA (výstupní)
 Treshold = prahová hodnota
 LRA input = LRA (vstupní)
 Input level = vstupní úroveň

Obr. 10: Příklad zpracování rozsahu hlasitosti (LRA) kompresorem s nízkou prahovou hodnotou (-50 dBFS) a mírným kompresním poměrem (1:1.5)

5. Co měřit v produkci a postprodukci

5.1 Normalizace nezávislá na signálu versus normalizace založená na kotvě

EBU R 128 doporučuje měřit **celý pořad**, nezávisle na jednotlivých typech signálu, jako je hlas, hudba nebo zvukové efekty (viz obr. 11). To se považuje za nejobecněji použitelnou praxi pro drtivou většinu pořadů:



Entire programme = celý pořad; Music = hudba; Voice = hlas; Sound Effects = zvukové efekty

Obr. 11: Prvky pořadu

Pro prvky s výrazně širokým rozsahem hlasitosti (>12 LU, přibližně) je možné alternativně používat tzv. **kotevní signál** pro normalizaci hlasitosti, tedy jinými slovy aplikovat *metodu individuálního hradlování*. Toto by měl být signál, který je podle výrobce nebo zvukového technika reprezentativní pro průměrnou hlasitost pořadu, jako např. řeč nebo zpěv, určitá část hudebního pořadu v *mezzoforte*, konzistentně používaný a dramaturgicky důležitý zvukový efekt, apod.

Musí se ale zdůraznit, že volba kotevního signálu je **aktivní proces** vyžadující vstup od zkušeného pracovníka. O tomto přístupu by se mělo uvažovat až poté, co se produkční pracovníci a zvukoví technici dokonale seznámí s koncepcí normalizace hlasitosti. Když se tento postup provádí dobře, může pomoci jemně vyladit hlasitost pořadů s širokým rozsahem hlasitosti podle zvoleného kotevního signálu.

Také existuje automatické měření jednoho specifického kotevního signálu ve formě „*Dialogové inteligence*“, což je firemní algoritmus společnosti Dolby Laboratories, jenž předpokládá, že řeč je běžný a důležitý signál ve vysílání. Tento algoritmus detekuje, zda se v pořadu vyskytuje řeč, a když je aktivovaný, měří pouze hlasitost během řečových intervalů. U pořadů s úzkým rozsahem hlasitosti je rozdíl mezi měřením omezeným na řeč a měřením prováděným u celého pořadu malý, obvykle <1 LU. U pořadů s širokým rozsahem hlasitosti, jako jsou akční filmy, se tento rozdíl potenciálně zvětšuje a někdy překračuje dokonce 4 LU. Automatická detekce kotevního signálu má pomoci identifikovat, co má být na cílové úrovni. Jako každý algoritmus detekce specifických signálů z kompletního a komplexního mixu je možné rozlišování řeči obelstít – buď signály, které silně připomínají spektrální strukturu řeči (např. určitými dřevěnými dechovými nástroji nebo sólovými houslemi) nebo řečovými signály, které jsou příliš daleko od rozlišovacího prahu (např. určité dialekty jazyka). U pořadů, kde se tyto kotevní signály konzistentně pohybují kolem rozlišovacího prahu může měření hlasitosti rovněž výrazně kolísat, když se měření provádí opakovaně.

U **krátkých** pořadů, jako jsou reklamy, upoutávky a propagační pořady, bude (automatická) normalizace řeči pravděpodobně poskytovat neuspokojivé výsledky ve světle budoucího nárůstu rozsahu hlasitosti a potenciálně zdokonalených dramaturgických koncepcí. V takových případech se většina mezinárodních doporučení (včetně tohoto) shoduje na měření „**celého**“ pořadu za všech okolností.

V každém případě si vysílatelé musí být vědomi toho, že obzvláště v bezpáskovém prostředí, kde se u většiny druhů obsahu bude celý pořad nezávisle na typu signálu (řeč, hudba, zvukové efekty) měřit automaticky, může být potřeba vytvořit jinou strategii pro práci s pořady založené na normalizaci pomocí kotvy.

Shrnutí:

Kvůli těmto nejistotám a skutečnosti, že řeč představuje pouze jednu část celého pořadu (i když velmi důležitou a běžnou), R 128 doporučuje měřit „**vše**“ – tj. celý pořad nezávisle na typu signálu (jako je hlas, hudba nebo zvukové efekty).

Tento přístup podporují i následující zjištění:

- Rozdíl mezi měřením „všeho“ a měřením kotevního signálu (jako je hlas, hudba nebo zvukové efekty) je malý u pořadů s velmi úzkým rozsahem hlasitosti;
- Rozdíl mezi měřením „všeho“ a „kotevního signálu“ silně závisí na obsahu pořadu, ale lze předpokládat, že bude větší, když bude rozsah hlasitosti větší;
- Automatické rozlišování kotevního signálu může fungovat dobře u většiny pořadů, ale některé signály je mohou obelstít, nebo se nemusí vůbec spustit, takže neposkytuje 100% konzistentní výsledky;
- Bepásková prostředí potřebují princip měření, který bude aplikovatelný na 100% obsahu a který bude poskytovat výsledky, jež budou „dostatečně dobré“ pro všechny pořady;
- Určení kotevního signálu vyžaduje zadání od zkušeného technika nebo rozlišovací algoritmus; takový algoritmus může podléhat určitým výše uvedeným nejistotám.

Normalizace podle kotevního signálu může nabízet lepší výsledky u materiálů s širokým LRA. Jedná se nicméně o úkol, který vyžaduje odbornost, a tedy čas a peníze, a když se používá automatické rozlišení, takový algoritmus nemůže být 100% spolehlivý. Když obsah upravený podle kotevního signálu vstupuje do normalizačních systémů na bezpáskových serverech a vyžaduje tedy obcházení zavedených automatických procesů, je třeba přijímat speciální opatření. **Jakožto největší společný jmenovatel R 128 doporučuje měřit celý pořad se všemi jeho prvky namísto kotevního signálu, a to i u materiálů s širokým LRA.**

5.2 Kanál nízkofrekvenčních efektů (LFE)

Jak je uvedeno v popisu ITU-R BS.1770 (viz § 2), je kanál LFE v současnosti **vyloučen** z měření. Jedním z důvodů je rozšířená nejistota spotřebitelů a zvukových techniků a rozdíly v kompenzaci tohoto kanálu u různých zařízení (pásmové zesílení +10 dB). Vypuštění kanálu LFE během měření hlasitosti by mohlo způsobit jeho zneužívání. Další zkoumání této záležitosti a praktické zkušenosti budou potřeba k tomu, aby se rozhodlo, zda a jak by bylo možné kanál LFE do měření zařadit. Jedním řešením jak kompletně eliminovat všechny potenciální problémy se signálem LFE je nepoužívat ho vůbec, např. když nebude potřeba zvláštní rezerva v oblasti nízkých basů.

6. Bezpásková produkce a vysílání

Protože svět vysílání přechází na pracovní bezpáskové procesy, je důležité, aby i v této oblasti byla koncepce normalizace hlasitosti plně přijata. Základní princip zůstává stejný: doporučuje se **normalizace hlasitosti a dynamická regulace zvukového signálu**, obzvláště pro nový obsah. Nicméně protože nedílnou součástí souborových systémů jsou **metadata**, řešení, která se opírají o metadata, jsou také uvedena (§ 7).

Vysílaný soubor, který obsahuje zvukové signály, může pocházet z procesu načtení, z přenosu z externího serveru nebo ze souborového archivu.

Pro stávající pořady (archivní obsah) existují v zásadě čtyři možnosti dosažení normalizace hlasitosti:

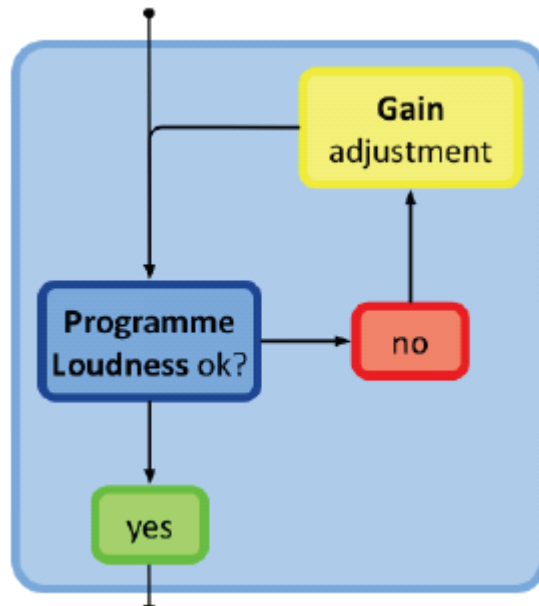
- Skutečná **změna** úrovně hlasitosti **všech zvukových souborů** na „cílovou hodnotu“
- **Změna** úrovně hlasitosti pouze „**podle potřeby**“;
- Použití výsledku měření úrovně hlasitosti k **úpravě vysílací úrovně**, aniž by se měnila původní úroveň hlasitosti;
- **Přenos správných metadat** hlasitosti ke spotřebiteli, kde se provádí normalizace.

To, které řešení se ve skutečnosti zvolí, závisí na faktorech, jako je specifická infrastruktura, pracovní procesy, správa mediálního obsahu, dostupnost vhodných zařízení, finanční zdroje, čas, atd.

Na samém začátku životnosti souboru v zařízení je potřeba provést měření, jež poskytnou hodnoty **úrovně hlasitosti pořadu, rozsahu hlasitosti a maximální skutečné špičkové úrovně** – tří charakteristických zvukových parametrů, jež jsou definované v EBU R 128 (pro velmi krátké materiály (<30 s, viz § 7) se může měřit a ukládat také maximální okamžité a maximální krátkodobá úroveň hlasitosti). Na základě výsledků těchto měření a následné metody dosažení normalizace hlasitosti a dodržení přijatelného rozsahu hlasitosti se vypracuje schéma zpracování, které se bude skládat ze „stavebních bloků“ nebo „klíčových úkolů“. Tento pracovní proces teď prozkoumáme dopodrobna z pomoci obecných vývojových diagramů.

6.1 Stavební bloky

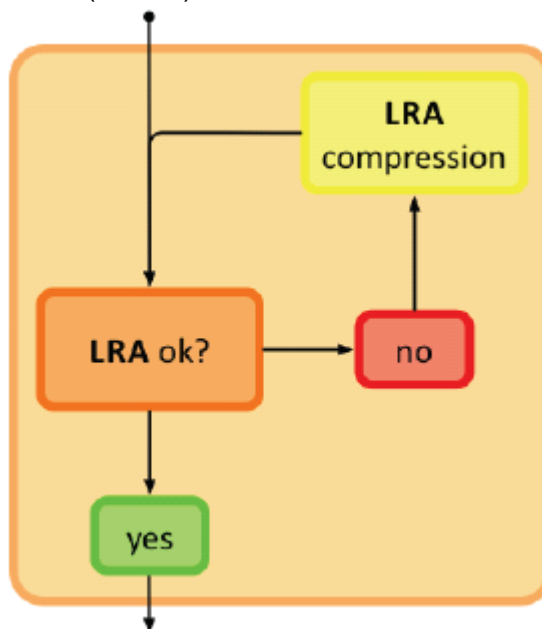
Zpracování úrovně hlasitosti pořadu (obr. 12)



Programme Loudness ok? = hlasitost pořadu OK?
Gain adjustment = úprava zesílení
no= ne; yes = ano

Obr. 12: Blok zpracování hlasitosti pořadu

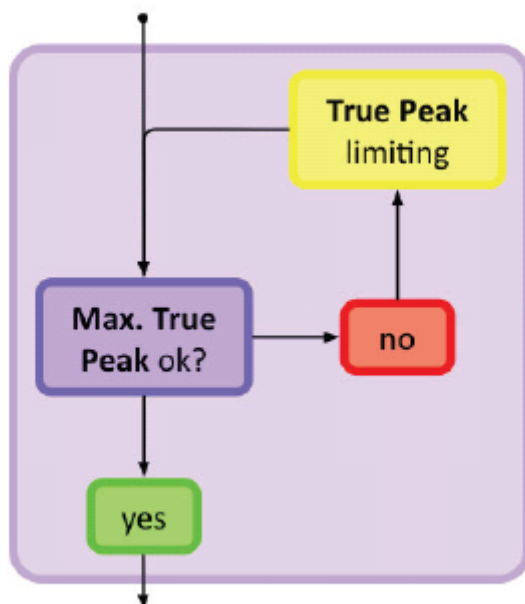
Zpracování rozsahu hlasitosti (obr. 13)



LRA compression = komprese LRA

Obr. 13: Blok zpracování rozsahu hlasitosti

Zpracování maximální skutečné špičkové úrovně (obr. 14)



Max. True Peak ok? = max. skutečná špička OK?
True Peak limiting = omezení skutečné špičky

Obr. 14: Blok zpracování max. úrovně skutečných špiček

6.2 Obecné strategie vyrovnávání hlasitosti - zpracování

Tři výše uvedené základní stavební bloky jsou jádrem každého procesu kontroly kvality s ohledem na technické parametry zvukového obsahu. **Na začátku každého potenciálního zpracování se změří hodnoty úrovně programové hlasitosti (L_K), rozsahu hlasitosti (LRA) a maximální skutečné špičkové úrovně (Max TP).** Výsledek těchto výchozích měření určuje následné zpracování.

Může nastat několik různých scénářů:

a) Všechny tři parametry jsou OK.



Toto je samozřejmě ideální výsledek měření: **úroveň programové hlasitosti je -23.0 LUFS, rozsah hlasitosti** je v rámci předepsaných limitů vysílatele (v závislosti na žánru a/nebo distribuční platformě) a **maximální skutečná špičková úroveň** se rovná stanovené maximální hodnotě určeného distribučního systému nebo je nižší.

b) Úroveň hlasitosti pořadu je vyšší než -23.0 LUFS.



Tento problém se řeší jednoduše úpravou zesílení (snížením úrovně):

$$\text{Zesílení (dB)} = L_K \text{ cílová} - L_K \text{ naměřená}$$

(Příklad: naměřená L_K je -19,4 LUFS; cílová úroveň je -23,0 LUFS; potřebné zesílení je $[-23,0 - (-19,4)] = -3,6$ dB. Max. TP se přirozeně sníží o stejnou hodnotu jako L_K .)

c) Úroveň programové hlasitosti je nižší než -23,0 LUFS.



Po aplikaci pozitivní kompenzace zesílení je potřeba přepočítat maximální skutečnou špičkovou úroveň (Max TP) (původně naměřená Max TP + kompenzace zesílení = výsledná Max TP), protože může potenciálně ležet nad přípustným limitem. Když bude nová hodnota Max TP skutečně ležet nad přípustným limitem, musí se provést **omezení skutečné špičky** podle stavebního bloku zpracování skutečné špičky. Dalším řešením, které je použitelné, když takové omezení skutečné špičky nebude možné nebo žádoucí (nebo potenciálně příliš závažné), je nechat L_K na původní nižší úrovni a aplikovat **odpovídající nastavení metadat hlasitosti** (nižší než -23, odrážející původní úroveň hlasitosti). To vyžaduje, aby byl k dispozici plně funkční systém, který podporuje a přenáší metadata (např. Dolby Digital nebo MPEG-4).

U obou scénářů **b** a **c** se může použít jednoduchá **hodnota zesílení** uložená jako metadata s možností následného omezení, pokud by došlo k překročení Max TP po aplikaci pozitivní kompenzace zesílení (scénář c). Pomocí této hodnoty zesílení se může provést regulace úrovně vysílání souboru tak, aby se dosáhlo -23 LUFS.

d) Úroveň hlasitosti pořadu je nižší než -23.0 LUFS a rozsah hlasitosti je širší, než je interní tolerance pro žánr nebo distribuční kanál.



Úroveň hlasitosti pořadu lze upravit podle výše uvedeného scénáře c. Rozsah hlasitosti se bude upravovat (stavební blok LRA), což může potenciálně snížit hodnotu Max TP. Ačkoliv hodnota Max TP mohla překročit přípustný limit, když se na L_K uplatnila pozitivní kompenzace zesílení, nemusí být nutné provádět úpravu Max TP díky snížení LRA. Proto je v rámci procesu snižování LRA potřeba provést výpočet Max TP.

e) Rozsah hlasitosti je širší, než je tolerance pro žánr nebo distribuční kanál.



Jak je uvedeno v § 4.4, může se k zúžení LRA použít kompresor s nízkou prahovou hodnotou a velmi mírným poměrem (stavební blok rozsahu hlasitosti). Pro soubory jsou výhodné automatické procesy s „cílovým LRA“. Alternativně může výsledek měření LRA aktivovat přednastavení kompresoru dynamiky v další fázi zpracování s parametry, jež jsou podobné těm uvedeným v § 4.4. Max TP se může pouze snižovat, takže zde není žádný potenciál pro případnou úpravu skutečné špičkové hodnoty.

f) Je překročena maximální přípustná skutečná špičková úroveň.



Překročení úrovně Max TP příslušného distribučního systému s sebou nese riziko **zkreslení** dále v řetězci (např. v převodníku D/A, převodníku vzorkovací frekvence nebo kodeku pro snižování bitového toku). Podle stavebního bloku Max TP se provede omezení skutečné špičkové úrovně za účelem snížení Max TP. Zda v důsledku toho dojde k významné změně programové hlasitosti, závisí na počtu a velikosti omezovaných špiček.

Jakékoli jiné kombinace výsledků výchozího měření LK, LRA a Max TP pokryjí procesy, jež již byly představeny ve výše uvedených scénářích.

7. Metadata

Jak již bylo uvedeno § 3.2, normalizace hlasitosti je možné dosáhnout buď **normalizací zvukového signálu** (doporučená metoda) nebo uložení skutečné úrovně hlasitosti **za použití metadat**. V tomto druhém případě se může posun na cílovou úroveň provést buď během přenosu zvukového souboru do vysílacího serveru, ve vysílacím mixážním pultu, použitím odpovídající předvolby dále zařazeného procesoru dynamiky nebo přímo na straně spotřebitele úpravou úrovně přehrávání.

Metadata mohou být obecně *aktivní* (potenciálně měnící zvukový signál) nebo *popisná* (poskytující informace o signálu, jako je formát, autorská práva, atd.). Přírozeným důsledkem práce ve skupině PLOUD a uveřejnění dokumentu EBU R 128 a jeho podpůrných dokumentů je, že jádro metadat hlasitosti ve zvukových souborech budou tvořit tři hlavní parametry, tj. **hlasitost pořadu, rozsah hlasitosti a maximální skutečná špičková úroveň**. Tyto tři parametry jsou již zařazeny do hlavičky (rozšiřující blok Broadcast Extension (BEXT)) formátu Broadcast Wave File (**BWF**) (viz *Tech. dokument 3285 [10]*; *podrobný popis BWF - viz [11], [12] a [13]*). Dále je třeba ukládat hodnoty **maximální okamžité úrovně hlasitosti a maximální krátkodobé úrovně hlasitosti**, protože tyto parametry jsou užitečné pro regulaci dynamiky velmi krátkých materiálů (<30 s; viz také § 10). Také se plánuje, že metadata budou zařazena do slovníku SMPTE s potenciálními vylepšeními, jako jsou „profily hlasitosti“, jež budou např. řešit různé předvolby pro zpracování v dále zařazených procesorech hlasitosti.

Parametry metadat ve stávajících systémech, jež mají primárně vztah k hlasitosti, jsou:

- hlasitost pořadu
- slova pro regulaci dynamického rozsahu
- koeficienty směšování na menší počet kanálů

Např. v systému metadat Dolby AC-3 se tyto parametry nazývají *dialnorm* (normalizace dialogů), *dynrng* (dynamický rozsah) a *Centre/Surround Downmix Level* (úroveň mixu středových/prostorových kanálů na nižší počet kanálů). Parametr *dialnorm* objektivně popisuje hlasitost celého pořadu se všemi jeho prvky, jako je hlas, hudba nebo zvukové efekty (i jen hudební pořad má hodnotu „*dialnorm*“). To se může zdát matoucí: důvodem je zaměření systému Dolby na normalizaci podle dialogu kotevního signálu.

7.1 Metadata hlasitosti pořadu

Vzhledem k důrazu na **normalizaci zvukového signálu** ve výrobě na **-23 LUFS** musí být relevantní parametr metadat přirozeně také nastaven tak, aby udával **-23 LUFS** za předpokladu, že byl pořad normalizován na cílovou úroveň. V důsledku toho bude po provedení normalizace zdrojových zvukových signálů v širokém měřítku parametr metadat hlasitosti pořadu **stický**.

Výjimky, kde se může používat jiná hodnota než -23, jsou:

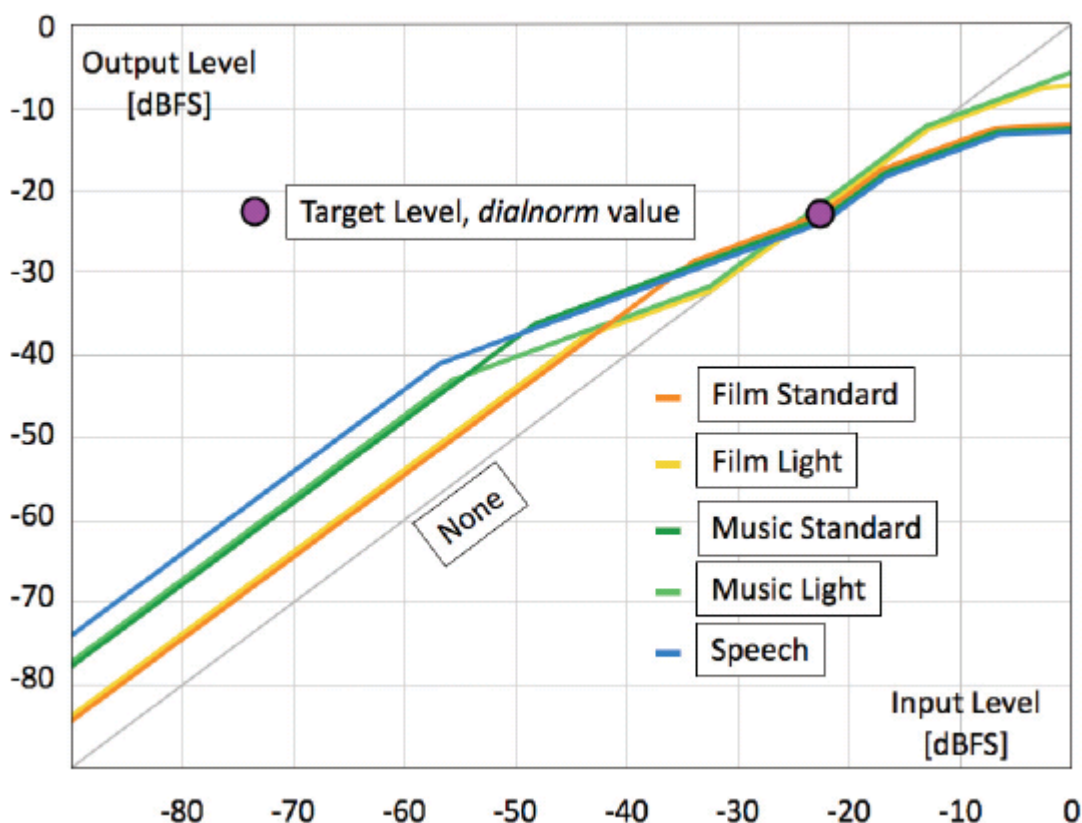
- Pořad nezapadá do okna, jež poskytuje -23 LUFS a -1 dBTP. K tomu může docházet zejména u velmi dynamických celovečerních filmů a u vysílatelů, kteří chtějí přenášet tyto pořady s tak velkým poměrem hlasitosti k špičkové úrovni;
- Starší pořady z archivu, které není možné včas upravit tak, aby splňovaly cílovou úroveň požadovanou R 128;
- Externí živé pořady mohou být poskytovány s odlišnými úrovněmi hlasitosti a metadaty.
- Plně funkční systém poskytování a používání metadat v celém řetězci signálu je již zaveden. To předpokládá věrný přenos metadat hlasitosti do domácích zařízení spotřebitelů.

Ve všech těchto situacích je rozhodně nutné nastavit **správnou** hodnotu metadat pro hlasitost pořadu, měřenou měřičem s „režimem EBU“. Distribuční systémy a zařízení domácích kin zvládají tuto situaci v následných fázích signálového řetězce (viz *Tech. dok. EBU 3344*).

7.2 Metadata regulace dynamického rozsahu

Tak, jako je možné provádět normalizaci hlasitosti u zdrojového zvukového signálu nebo pomocí metadat, totéž platí pro zpracování dynamického rozsahu. V prostředí metadat se informace o kompresi datového rozsahu odesílají jako součást datového proudu ve formě *slov pro zesílení*. V zařízení domácího kina spotřebitele se tyto informace používají ke snižování dynamického rozsahu buď ve výchozím nastavení nebo po aktivaci uživatelem. Regulace dynamického rozsahu pomocí metadat není srovnatelná s propracovaným procesorem dynamiky, ale poskytuje určitou „náplast“ pro situace, kdy spotřebitel požaduje výrazně nižší dynamický rozsah.

Když opět odkážeme na systém Dolby Digital, existuje *6 předvoleb komprese*, na jejichž základě kodér generuje různá slova pro regulaci zesílení, která se posílají v bitovém proudu do dekodéru spotřebitele: Film Standardní, Film Lehký, Hudba Standardní, Hudba Lehká, Řeč a Žádná. Tyto předvolby zajišťují větší nebo menší kompresi kolem hodnoty *dialnorm*, což je další důvod pro to, aby se tento parametr metadat nastavoval (*křivky komprese kolem -23 LUFS ilustruje obr. 15*).



Output level = výstupní úroveň; Target level, dialnorm value = cílová úroveň, hodnota dialnorm; None = žádná; Film Standard = Film Standardní; Film Light = Film Lehký; Music Standard = Hudba Standardní; Music Light = Hudba Lehká; Speech = řeč; Input Level = vstupní úroveň

Obr. 15: Obecné křivky komprese dynamického rozsahu systému AC-3

V systému Dolby Digital existují dva profily komprese: „Linkový režim“ a „RF režim“. Pro každý z nich je možné zvolit samostatnou předvolbu komprese.

V rámci systému **R 128** a jeho koncepcí normalizace zvukového signálu na **-23 LUFS** a používání parametru rozsah hlasitosti k rozhodnutí o možné úpravě se může používat předvolba „**Žádná**“. To může být použitelné zvláště na „*Lineární režim*“ a v základním nastavení i na „*RF režim*“.

Regulace rozsahu hlasitosti vlastním zpracováním zvukového signálu při přenosu obecně přesouvá tento problém blíže ke zdroji. Nicméně u specifických pořadů může vysílatel zvolit *jemný* profil pro systémy s RF režimem (aby nedocházelo k příliš aktivní ochraně před zahlcováním), nicméně pro systémy s linkovým režimem může zvolit předvolbu „*Žádná*“. Vysílatelé, kteří potřebují jiné profily než „*Žádná komprese*“ na podporu svých interních pracovních procesů, si musí být vědomi toho, že tato funkce se nemusí vždy spolehlivě realizovat na zařízeních jejich posluchačů. **Doporučuje se, aby výrobci a distribuční společnosti dbali na to, aby se zařízení vyráběla v souladu s Tech. dokumentem EBU 3344 („Pokyny k distribuci“).**

7.3 Koeficienty pro mix na menší počet kanálů

Tyto parametry metadat (opět jako příklad zde pro Dolby Digital) jsou přirozeně použitelné pouze na prostorové zvukové signály, přičemž regulují zesílení (v dB) středového kanálu a prostorových kanálů (surround), když se směšují do levého předního a pravého předního kanálu při odvozování 2kanalového stereofonního signálu. Hlasitost 2kanalového stereo signálu, který je výsledkem automatického směšování za použití metadat závisí na:

- samotných skutečných koeficientech směšování (+3/+1.5/0/-1.5/-3/-4.5/-6/-∞)
- obsahu pořadu ve středovém kanálu a prostorových kanálech a
- potenciálním bezpečnostním omezení bránícím přebuzení

Je třeba dbát na to, aby **nedocházelo k přebuzení** signálu smíchaného do méně kanálů. To lze zajistit předsazením procesoru dynamiky. Je třeba vyhnout se statickému škálování (celkovému snižování úrovně), protože systematicky zavádí rozdíly v hlasitosti mezi výsledným 2kanalovým stereofonním mixem a původním signálem prostorového zvuku. Řešení může nabízet dynamické škálování.

Směšovací koeficienty možné v rámci systému Dolby-Digital se řídí dvěma směšovacími profily. Původně, když existoval pouze jeden profil, byly parametry hrubší, s hodnotami -3/-4.5/-6 dB pro středový kanál a -3/-6/-∞ dB pro prostorové kanály. V současnosti poskytují rozšířené informace bitového toku (Extended BSI) výše uvedené jemnější mezikroky (u první odrážky; i směšovací koeficienty DVB TS 101 154 nabízejí stejné rozlišení jako Extended BSI Dolby Digital). Vysílatelé by si měli být vědomi skutečnosti, že ne všechna reprodukční zařízení dokážou předat zamýšlený 2kanalový zážitek v případě použití Extended BSI, protože starší dekodéry nemusí být schopny tyto informace extrahovat a mohou sklouznout k menšímu počtu hrubších koeficientů profilu 1.

V případě chybějících nebo nespolehlivých metadat pro směšování na menší počet kanálů je dobrým výchozím krokem podívat se na koeficienty popsané v ITU-R BS.775-2 [14]:

L, R front: **0** dB (levý, pravý přední)
C, LS, RS: **-3** dB (středový, levý, pravý prostorový)

Také znovu zdůrazňujeme, že prostorové kanály mají při měření hlasitosti podle ITU-R BS.1770 váhu +1,5 dB⁵. Po automatickém smíchání na stereo se tato váha neuplatňuje, protože výsledkem je pouze přední dvoukanalové stereo (levý a pravý přední kanál). Pořady s velkým množstvím obsahu v prostorových kanálech v důsledku toho vykazují potenciálně větší odchylky hlasitosti prostorového mixu oproti 2kanalovému stereofonnímu mixu s „konzervativnějším“ používáním prostorových kanálů.

⁵ **Koeficient váhy +1.5 dB** pro prostorové signály při **měření hlasitosti** podle ITU-R BS.1770 se nesmí zaměňovat se skutečným zesílením **+3 dB** pro prostorové signály v **kině!** V kině mají dva samostatné prostorové kanály nastavenou úroveň o 3 dB níže než přední kanály, aby se jejich kombinovaná úroveň rovnala jednomu přednímu kanálu. Důvodem pro to je kompatibilita s filmy s mono prostorovým zvukem (matricově kódované „Dolby Stereo“ má pouze (pásmově omezený) monofonní prostorový signál), kde by oba kanály dostaly identický signál. U oddělených vícekanalových zvukových mixů („5.1“, atd.) jsou proto prostorové signály v konečném mixu o 3 dB „silnější“, protože zvukový technik kompenzuje o 3 dB nižší nastavení prostorových kanálů. Když se vysílá mix pro kino, musí se tento rozdíl o 3 dB kompenzovat (ostatní parametry jako rozsah hlasitosti je třeba rovněž upravit).

Zatímco +3dB zesílení prostorových kanálů má čistě technický důvod, zesílení +1,5 dB pro prostorové kanály při měření hlasitosti má důvody psychoakustické. Lidé vnímají zvuky přicházející zezadu jako hlasitější než signály zepředu se stejnou úrovní zvukového tlaku. Měřicí zařízení nemá mozek, a tak potřebuje tento faktor zesílení.

V každém případě neexistuje žádná záruka, že metadata dodaná s *externím* souborem (nebo jiným médiem) budou správná. Metadata hlasitosti pořadu udávající -27 (výchozí tovární hodnota pro *dialnorm* v systému Dolby-Digital) nebo -31 (nejnižší možná hodnota v tomto systému) pravděpodobně vyvolají speciální pozornost, protože existuje možnost, že metadata buď nikdo nekontroloval nebo byla zneužita k tomu, aby se pořad jevil jako (mnohem) hlasitější při přehrávání na straně spotřebitele.

Proto se doporučuje **nepoužívat metadata regulace hlasitosti a dynamického rozsahu** externích zdrojů (pokud zdroj není naprosto důvěryhodný). Koeficienty pro směšování na méně kanálů mohly projít plně funkčním metadatovým systémem. Celý proces měření tří hlavních zvukových parametrů je třeba provést od začátku. Jedině tím bude zaručeno správné následné zpracování. Metadata pro interní účely je možné lépe kontrolovat.

8. Vyrovnávání signálů ve světle normalizace hlasitosti

8.1 Vyrovnávací signál a úroveň

Vyrovnávací signál ve vysílání se skládá ze sinusového signálu na frekvenci 1 kHz, který se používá k technickému vyrovnávání připojení zvukového pořadu. V digitálních systémech je **úroveň** takového vyrovnávacího signálu **18 dB** pod maximální úrovní kódování bez ohledu na celkový počet dostupných bitů (**-18 dBFS**). Při přechodu na normalizaci hlasitosti se tento přístup **NEZMĚNÍ**, protože z vyrovnávání nevyplývá povinné připojení k měření hlasitosti.

Proto je možné vyrovnávání pro výměnu zvukových pořadů provádět jako obvykle pomocí sinusového signálu 1 kHz na úrovni **-18 dBFS**.

Vyrovnávací úroveň pro výměnu zvukových pořadů se nemusí měnit. Používejte 1 kHz sinusovou křivku na úrovni -18 dBFS jako obvykle

To je uvedeno v Doporučení EBU R 68 [15]. V tomtéž dokumentu se ještě zmiňuje „Povolená maximální úroveň“ tak, jak je definovaná v Doporučení ITU-R BS.645-2 [16]; S přechodem na „Maximální povolenou skutečnou špičkovou úroveň“ (-1 dBTP pro obecné produkce PCM), jež se liší od doporučené hodnoty -9 dBFS v ITU-R BS.645 (protože scénář měření pomocí QPPM zastarává), dochází k tomu, že relevantní kapitoly EBU R 68 – 2000 a ITU-R BS.645 (a dokumentů, které odkazují na definici „Povolené maximální úrovně“ v těchto doporučeních) budou potenciálně potřebovat revizi.

Vyrovnávací úroveň **-18 dBFS** (1kHz tón) se bude zobrazovat jako **-18 LUFS** na měřiči hlasitosti s absolutní stupnicí (nebo +5 LU na relativní stupnici režimu EBU) za předpokladu, že tento 1kHz tón bude přítomen (ve fázi) jak na levém, tak na pravém kanálu stereofonního nebo prostorového zvukového signálu. Když se bude 1kHz tón používat pouze u jednoho předního kanálu, bude měřič hlasitosti udávat -21 LUFS (nebo +2 LU na relativní stupnici).

Stereofonní sinusový signál 1kHz na -18 dBFS bude zobrazovat absolutní hodnotu -18 LUFS (relativní +5 LU) na měřiči hlasitosti s režimem EBU

8.2 Poslechová úroveň

Jiným tématem je **poslechová úroveň** u zvukového reprodukčního systému. V relevantním dokumentu Tech. dok. EBU 3276-E "Poslechové podmínky pro hodnocení materiálu zvukového pořadu" (a Dodatku 1, který jej rozšiřuje na vícekanálový zvuk) se používají následující rovnice k úpravě hladiny jednoho reproduktoru [17]:

- (1) $L_{LISTref} = 85 - 10\log_2 \text{ dBA}$ (pro 2kanálové stereo)
- (2) $L_{LISTref} = 96 \text{ dB}_C$, vztaženo na úroveň signálu digitální plné stupnice (pro vícekanálové audio do 5.1)

K tomuto účelu je třeba používat signál, který se skládá z šumu se stejnou energií na oktávu a pokrývající buď celý frekvenční rozsah (rovnice (1)) nebo frekvenční rozsah od **500 Hz do 2 kHz** (rovnice (2)). Měření by se měla ve skutečnosti provádět na střední úrovni signálu, jež se rovná **vyrovňovací úrovni**, která je zde definována jako 18 dB pod digitální plnou stupnicí. Za těchto podmínek by se mělo zesílení reproduktoru upravit tak, aby bylo dosaženo **referenční poslechové úrovně** ($L_{LISTref}$) $85 - 3 =$ úroveň zvukového tlaku (SPL) **82 dB_A (SPL)** na jeden reproduktor u 2kanálových stereo systémů a $SPL\ 96 - 18 =$ **78 dB_C SPL** na jeden reproduktor u vícekanálových systémů. Měření by se mělo provádět v referenční poslechové pozici za použití A váženého měřiče zvukové úrovně pro 2kanálové stereo a C váženého měřiče pro vícekanálový zvuk.

To je pravda trochu matoucí vzhledem k odlišným číslům, odlišným šumovým signálům a odlišnému vážení měřiče úrovně zvuku. Ale tyto rozdíly svým způsobem vlastně zajišťují kompenzaci a vytvářejí podobnou poslechovou úroveň jak pro 2kanálové stereo, tak pro vícekanálové systémy.

Shrnutí:

| | | |
|------------------------|--|--|
| Pro 2kanálové stereo : | $L_{LISTref} = 82 \text{ dB}_A \text{ SPL}$ na reproduktor | (za použití 20 Hz – 20 kHz šumu se stejnou energií na oktávu při -18 dBFS rms) |
| Pro 5.1 MCA: | $L_{LISTref} = 78 \text{ dB}_C \text{ SPL}$ na reproduktor | (za použití 500Hz – 2 kHz šumu se stejnou energií na oktávu při -18 dBFS rms) |

Zneužívání normalizace špiček vedlo k podstatně nižším poslechovým úrovním, takže s normalizací hlasitosti může dojít k tomu, že se teď skutečně budou používat doporučené úrovně. Také se předpokládá, že průměrná úroveň zvukových pořadů bude po zavedení EBU R 128 do praxe *nižší*. Snížení úrovně může být v řádu až 3 LU (v extrémních případech ještě více). V důsledku toho se zdá, že se pravděpodobně **zvýší** stávající úroveň monitorů reprodukčních systémů. Jak bylo uvedeno výše, *vyrovňovací úroveň* se nemusí adekvátně měnit, protože proces vyrovnávání je nadále platný, aby byla zajištěna racionální struktura zesílení a vysoký poměr signálu k šumu v reprodukčním řetězci. Pokud dojde v budoucnosti k široké dohodě o zvýšení úrovně do monitorů kvůli nižší průměrné úrovni, budou příslušné dokumenty znovu přezkoumány.

Protože hodnota -23 LUFT je asi o 3 LU nižší, než je dnešní průměrná programová úroveň, je možné uvažovat o adekvátním zvýšení úrovně přiváděné do monitorů.

9. Implementace a migrace

Je evidentní, že taková zásadní změna měření a zpracování zvukových signálů, která ovlivňuje všechny fáze výroby, distribuce, archivace a přenosu zvukových pořadů, se neprovede přes noc stisknutím vypínače. Každý vysílatel a producent zvukových materiálů si musí najít svůj vlastní způsob jak tuto změnu realizovat, jak nainstalovat odpovídající zařízení, proškolit personál a nastoupit na cestu do ráje hlasitosti! Nicméně můžeme uvést několik konstant, které budou platit pro všechny. Ty jsou uvedeny v následujících kapitolách.

9.1 Všeobecná doporučení k migraci a implementaci

- Ustanovte **interní skupinu pro hlasitosti**, která prodiskutuje základní souvislosti a strategie jak přesvědčit vedení, tvůrce pořadů a vaše kolegy.
- **Začněte hned** – nečekejte, až bude všechno na místě a všichni ostatní už to budou mít hotové, a nesnažte se být dokonalí od samého začátku.
- Než budete moci začít něco dělat, musí s touto změnou a jejími důsledky souhlasit vedení. Získejte **písemnou dohodu** nebo „výzvu k realizaci“ od generálního ředitele.
- Svým klíčovým výrobním pracovníkům poskytněte **měřiče hlasitosti**. Nechte je si s nimi pohrát, získat první zkušenosti a zjistit výhody a úlevy modelu měření hlasitosti, aby mohli být určovateli názorů pro svoje kolegy.
- **Proveďte si průzkum trhu** měření a regulace hlasitosti, abyste zjistili, co bude nejvhodnější pro vaše prostředí.
- Určete **klíčové oblasti**, v nichž by měla práce na hlasitosti začít. Potenciálními kandidáty jsou: produkční studia, postprodukční prostory, přenosové vozy, oddělení kontroly kvality.
- Buďte si vědomi toho, že narazíte na překážky („vždycky se to dělalo takhle“, „nikdy se to nedělalo takhle“, „kdo jsi, že nám říkáš, že to máme dělat takhle“). **Trpělivost** a předvádění praktických příkladů se vyplatí. Staňte se zenovým mistrem normalizace hlasitosti ve své firmě („sebeovládání – jednoduchost – přirozenost“).
- **Dejte všem čas na přizpůsobení**. Ačkoliv publikum čeká na řešení už několik desetiletí, nevytvářejte další problémy tím, že budete příliš spěchat.
- Řešení pro **pracovní bezpáskové procesy** jsou zatím vzácná (únor 2011). Sledujte trh a žádejte řešení od prodejců.
- Využijte této zásadní změny jako příležitosti pro všeobecnou **diskusi o kvalitě zvuku** a vývoji „**firemního zvuku**“, který bude např. zahrnovat srozumitelnost řeči, vyvážení řeči a hudby a samozřejmě normalizaci hlasitosti pořadů.
- **Používejte uši a důvěřujte jim!** Jsou to nejlepší měřiče hlasitosti. Usmívejte se, když uvidíte, jak vaši kolegové pracují s modelem hlasitosti, jako kdyby nic jiného nikdy neexistovalo.

9.2 10 akčních bodů pro migraci a implementaci

**Založte interní skupinu pro hlasitost.
Nečekejte, začněte hned.
Získejte písemný souhlas vedení.
Dejte svým klíčovým pracovníkům měřiče hlasitosti.
Sledujte trh s měřicími a regulačními zařízeními.
Zvolte klíčové oblasti, které se mají změnit jako první.
Objeví se překážky.
Buďte trpěliví, ponechte čas.
Využijte změnu k diskusi o kvalitě zvuku.
Používejte uši a důvěřujte jim.
Staňte se zenovým mistrem hlasitosti své firmy.**

10. Specifické problémy jednotlivých žánrů

Koncepce **EBU R 128** se soustřeďuje na normalizaci hlasitosti každého pořadu na jednu cílovou úroveň (**-23 LUFS**). Existují dva důvody, proč to nemůže být dokonalé řešení:

- Žádné objektivní měření hlasitosti nemůže být nikdy dokonalé
- Vždy budou existovat individuální preference

To znamená, že dokonalé řešení *obecně* není možné, protože vnímání hlasitosti se liší od člověka k člověku a závisí na faktorech, jako je věk, pohlaví, nálada, atd. V rámci EBU R 128 je důležité pochopit, že účelem není dosáhnout rovnováhy hlasitosti na základě skutečné úrovně zvukového tlaku konkrétního zvukového signálu, ale místo toho poskytnout uspokojivý poslechový dojem pro rozmanitou směs žánrů většině posluchačů.

Z toho vyplývá např. to, že Schubertův smyčcový kvartet bude mít stejnou integrovanou úroveň hlasitosti jako Mahlerova symfonie, a to -23 LUFS. I když to neodpovídá realitě, budou tyto programové položky zapadat do široké škály po sobě následujících pořadů a to je účel **prosazování jediného čísla**.

Protože tento dokument má sloužit jako zásobník zkušeností, člověk by byl v pokušení uvažovat o zdokonalení tohoto modelu, až se normalizace hlasitosti rozšíří. Nicméně posluchači berou jako přijatelné, když úroveň hlasitosti pořadů leží v rámci tzv. „**komfortní zóny**“ přibližně **8-9 LU**, přičemž rozložení je asymetrické (např. +3 LU/-5 LU). V případech, kdy objektivní algoritmus hlasitosti ne vždy poskytuje dokonalý výsledek, bude pořad s největší pravděpodobností i přesto ležet uvnitř této komfortní zóny. Vysílatelé by také měli pamatovat na to, že posluchači mohou i nadále upravovat úroveň hlasitosti dálkovým ovládním podle toho, co se jim líbí a co ne.

EBU podporuje **normalizaci na jedinou cílovou úroveň** navzdory možnosti jemné úpravy pro jednotlivé žánry. Když se připustí příliš mnoho odchylek (nebo dokonce jen pár), systém stejné průměrné hlasitosti se může zpochybnit od samého začátku. Přirozeně existuje obava, že by odchylky směřovaly hlasitějším směrem.

Vykazování úrovně hlasitosti programu **nižší**, než je cílová úroveň, je poněkud odlišné téma. Jako „případová studie“ se teď zkoumají dva žánry, kde může být za určitých okolností vhodný specifický přístup (totéž platí i pro maximální úroveň hlasitosti): **reklamy a upoutávky a hudební pořady**.

10.1 Reklamy a upoutávky

Tento typ pořadu je rozhodně nejčastěji zmiňovaným v souvislosti s **rozladěním posluchačů**, a je tedy hlavně zodpovědný za problémy s hlasitostí, s nimiž se teď setkáváme. Ve Velké Británii (Pravidla BCAP – Komise pro praxi při vysílání reklam) a v USA (zákon CALM – Zmírnění hlasitosti reklam) i v Polsku a v Itálii byla dokonce nedávno zavedena *legislativní opatření* ke zkrácení tohoto žánru. Je určitě zásadní, aby systém normalizace hlasitosti založený na **EBU R 128** poskytoval efektivní sadu nástrojů pro tento úkol – zneužívání se musí bránit. Pro kontrolu nad dynamikou reklamy ve světě s normalizovanou hlasitostí, kde hrozí nebezpečí náhlých příliš vysokých rozdílů v hlasitosti (příliš hlasité „rozuzlení“ po delším intervalu nízkourovňových signálů těsně nad prahovou hodnotou hradla) není parametr rozsah hlasitosti (LRA) vhodný, protože výpočet je založen na krátkodobých hodnotách hlasitosti (3s interval). V důsledku toho je u velmi krátkých položek k dispozici příliš málo datových bodů, aby se z nich odvodilo smysluplné číslo pro LRA. Za tuto skutečnost parametr rozsah hlasitosti nemůže, protože nikdy nebyl určen pro tento účel.

Alternativu lze najít v používání **maximální okamžité úrovně hlasitosti** (Max ML – 400 ms) a/nebo **maximální krátkodobé úrovně hlasitosti** (Max SL – 3 s). Obzvláště u velmi krátkých programových položek (<30 s) je možné tyto parametry účinně používat k omezování špiček hlasitosti. První zkušenosti členů PLOUD naznačily hodnotu zhruba **+8 LU** (-15 LUFs) jako možný limit pro Max ML a **+3 LU** (-20 LUFs) pro Max SL. V každém případě jsou oba parametry součástí rozšíření metadat souborového formátu Broadcast Wave (BWF) (*viz Tech. dok. EBU 3285, verze 2.0, 2011 [10]*). Doporučuje se, aby členové EBU používali individuální limity Max ML nebo Max SL pro krátké položky a hlásili svoje poznatky.

Limit pro maximální okamžitou a maximální krátkodobou hlasitost se může používat jako ochrana před zneužíváním velmi krátkých položek (< 30 s).

U pořadů tohoto žánru, které obsahují pouze pozadí nebo kreativně využívají zvuků na nízké úrovni, se může používat **nižší** úroveň hlasitosti než cílová úroveň. To je v souladu s minulou a současnou praxí omezit maximální špičkové úrovně (teď: úrovně hlasitosti), ale nevyžadovat, že by veškerý obsah musel být na stanoveném maximu. Úmyslně nízká úroveň zvuku vytváří **kontrast** a to je jeden z nezákladnějších kreativních nástrojů v každé umělecké formě. Krátké trvání reklam nebo upoutávek, které by mohly efektivně využívat tento dramaturgický nástroj, znamená, že to pravděpodobně nemůže mít žádný vliv na každodenní dlouhodobou průměrnou hlasitost dané stanice.

Pořady určené k vysílání při nižší než cílové úrovni vyžadují speciální pozornost, aby bylo zajištěno, že projdou procesy automatické normalizace nepoškozené. Měly by skutečně představovat výjimku, ne pravidlo.

Konečná odpovědnost za všechny tyto případy a rozhodnutí leží na producentovi, režisérovi, respektive ostatním kreativním personálu.

10.2 Hudba

Zkušenosti vášnivých posluchačů hudby naznačují, že určité pořady, které obsahují většinou hudbu - buď s širokým rozsahem hlasitosti v případě vážné hudby nebo hudbu s vysokým stupněm dynamické komprese jako uměleckého prvku např. u rockového koncertu - mají tendenci být poslouchány při vyšší úrovni hlasitosti (až v průměru o **+2-3 LU**) než jiné žánry. Důvodem pro to může být výrazně vyšší potenciální úroveň zvukového tlaku ve skutečnosti (fortissimo symfonického orchestru, rocková skupina s výkonnou aparaturou) a skutečnost, že u hudby neexistují žádné „zvuky v popředí“ a „zvuky na pozadí“ – v popředí je vše.

Ale jak jsme uvedli výše, potenciální odlišení cílové úrovně pro tyto pořady může způsobit větší škodu spočívající v otevření zadních vrátek pro zvýšení hlasitosti oproti ostatním místo toho, aby se situace výrazně zlepšila. Na základě stejné argumentace jako u reklam a upoutávek se normalizace na odlišnou (=vyšší) cílovou úroveň **nedoporučuje**. Publikum si může nadále upravovat (zvyšovat) úroveň hlasitosti ve svém reprodukčním prostředí dálkovým ovládním podle svého vkusu. Následující pořady jako reklamy a upoutávky budou mít důsledku toho hlasitost také posunutou. Předpokládá se, že toto by nemělo vytlačit tyto pořady z komfortní zóny.

11. Literatura

- [1] EBU Technical Recommendation R 128 '*Loudness normalisation and permitted maximum level of audio signals*' (2010, revize 2011) (Normalizace hlasitosti a maximální povolená úroveň zvukových signálů)
- [2] ITU-R BS.1770 '*Algorithms to measure audio programme loudness and true-peak audio level*' (2006-2007; revize 2011) (Algoritmy k měření hlasitosti zvukových pořadů a skutečné špičkové úrovně zvuku)
- [3] EBU Tech Doc 3341 '*Loudness Metering: 'EBU Mode' metering to supplement loudness normalisation in accordance with EBU R 128*' (revision 2011) (Měření hlasitosti: Měření v „režimu EBU“ na doplnění normalizace hlasitosti v souladu s EBU R 128)
- [4] EBU Tech Doc 3342 '*Loudness Range: A measure to supplement loudness normalisation in accordance with EBU R 128*' (revize 2011) (Rozsah hlasitosti: Parametr na doplnění normalizace hlasitosti v souladu s EBU R 128)
- [5] EBU Tech Doc 3344 '*Practical Guidelines for Distribution Systems in accordance with EBU R 128*' (2011) (Praktické pokyny pro distribuční systémy podle EBU R 128)
- [6] ISO 80000-8: „*Veličiny a jednotky – Část 8: Akustika*“
- [7] Grimm E., Skovenborg E. & Spikofski G. '*Determining an Optimal Gated Loudness Measurement for TV Sound Normalization*', Příspěvek Konvence AES č. 8154, 128. Konvence AES, květen 2010 (Určení optimálního měření hradlované hlasitosti pro normalizaci TV zvuku“)
- [8] ITU-R BS.1864 '*Operational practices for loudness in the international exchange of digital television programmes*' (2010) (Provozní praxe pro hlasitost při mezinárodní výměně digitálních televizních pořadů)
- [9] Lund, Th. '*Stop counting samples*' (Přestaňte počítat vzorky), Příspěvek AES č. 6972, 121. Konvence AES, říjen 2006
- [10] EBU Tech Doc 3285: '*Specification of the Broadcast Wave Format (BWF) – A format for audio data files in broadcasting*' (verze 2.0, 2011) (Popis formátu Broadcast Wave (BWF) - formátu pro zvukové datové soubory ve vysílání)
- [11] EBU Technical Recommendation R 85: '*Use of the Broadcast Wave Format for the Exchange of Audio Data Files*' (2004) (Používání formátu Broadcast Wave pro výměnu zvukových datových souborů)
- [12] EBU Technical Recommendation R 111: '*Multi-channel Use of the BWF Audio File Format (MBWF)*' (2007) (Vícekanálové používání formátu zvukových souborů BWF (MBWF))
- [13] EBU Tech Doc 3306: '*MBWF/RF64: An extended File Format for Audio*' (2009) (Rozšířený souborový formát pro zvuk)
- [14] ITU-R BS.775-2 '*Multichannel stereophonic sound system with and without accompanying picture*' (2006) (Vícekanálový stereofonní zvukový systém s doprovodným obrazem a bez něj)
- [15] EBU Technical Recommendation R 68: '*Alignment level in digital audio production equipment and in digital audio recorders*' (revize 2000) (Úroveň vyrovnání v digitálním zvukovém produkčním zařízení a digitálním rekordérech zvuku)
- [16] ITU-R BS.645-2 '*Test signals and metering to be used on international sound programme connections*' (1992) (Testovací signály a měření určené pro mezinárodní připojení zvukových pořadů)
- [17] EBU Tech Doc 3276-E (+ supplement (dodatek) 1) '*Listening conditions for the assessment of sound programme material*' (1998, 2004 – dodatek 1) (Poslechové podmínky pro hodnocení materiálu zvukových pořadů)