

AKADEMIE MÚZICKÝCH UMĚNÍ V PRAZE
HUDEBNÍ A TANEČNÍ FAKULTA

Hudební umění

Zvuková tvorba

DIPLOMOVÁ PRÁCE

**NORMALIZACE HLASITOSTI ZVUKOVÉHO SIGNÁLU
V DIGITÁLNÍCH MÉDIÍCH**

se zaměřením na internetová audio a audiovizuální média

BcA. Filip Beneš

Vedoucí práce:	doc. MgA. Ing. Ondřej Urban, Ph.D.
Oponent práce:	PhDr. Aleš Dvořák
Datum obhajoby:	6. září 2021
Přidělovaný akademický titul:	MgA.

Praha, 2021

ACADEMY OF PERFORMING ARTS IN PRAGUE
MUSIC AND DANCE FACULTY

Art of Music

Sound production

MASTER THESIS

**LOUDNESS NORMALIZATION OF SOUND SIGNAL
IN DIGITAL MEDIA**

focusing on audio and audiovisual internet media

BcA. Filip Beneš

Thesis supervisor:	doc. MgA. Ing. Ondřej Urban, Ph.D.
Opponent:	PhDr. Aleš Dvořák
Date of defence:	6 th September 2021
Academic degree conferred:	MgA.

Prague, 2021

Prohlášení

Prohlašuji, že jsem diplomovou práci na téma „Normalizace hlasitosti zvukového signálu v digitálních médiích“ vypracoval(a) samostatně pod odborným vedením vedoucího práce a s použitím uvedené literatury a pramenů.

V Praze dne

.....

podpis diplomanta

Upozornění

Využití a společenské uplatnění výsledků diplomové práce nebo jakékoliv nakládání s nimi je možné pouze na základě licenční smlouvy, tj. souhlasu autora a AMU v Praze.

Abstrakt

Práce se zabývá problematikou normalizace hlasitosti a její implementace v digitálních a především internetových médiích. Shrnuje současné poznatky o měření hlasitosti a zabývá se dopodrobna nejrozšířenějším systémem pro měření hlasitosti podle ITU-R BS.1770. Rozebírá strategie normalizace a zkoumá příčiny nejednotných hlasitostních standardů mezi médii. Popisuje způsoby normalizace hlasitosti vybraných internetových služeb a hlasitostní standardy některých služeb byly ověřeny experimentálně. Práce také nabízí vhled do důsledků problematiky v práci zvukových mistrů.

Abstract

The thesis deals with the loudness normalisation and its implementation in digital media and especially in internet media. It summarizes current knowledge of loudness measurement and deals with the most common loudness measurement system according to ITU-R BS.1770 in detail. It analyses loudness normalisation strategies and investigates the causes of inconsistent loudness standards across media. It describes particular loudness normalisation implementations of selected internet media services and some loudness standards had been verified experimentally. The thesis also offers insight into the consequences of the problematics in sound engineers' work.

Obsah

Seznam příloh.....	1
Seznam použitého označování a zkratk.....	2
Úvod	3
1 Hlasitost a válka hlasitostí	5
1.1 Vjem hlasitosti.....	5
1.2 Hlasitost signálu	6
1.3 Válka hlasitostí	6
2 Měření úrovně signálu	9
2.1 Analogový signál.....	9
2.1.1 VU metr	9
2.1.2 Peak programme meter	9
2.2 Digitální signál.....	10
2.2.1 Peak meter	10
2.2.2 RMS meter	10
3 Systém jednotek LUFS.....	11
3.1 Algoritmus pro výpočet hlasitosti dle ITU-R	11
3.1.1 Výpočet hodnoty hlasitosti	11
3.1.2 Dialogue gating	14
3.1.3 True peak	15
3.1.4 Momentary loudness	16
3.1.5 Short-term loudness	17
3.2 Jednotky dle EBU	17
3.2.1 Loudness range	18
3.2.2 Další analytické hodnoty	19
3.3 Měřiče hlasitosti	19
4 Normalizace hlasitosti z obecného hlediska.....	21
4.1 Principy a strategie normalizace hlasitosti	23
4.1.1 Maximální integrovaná hlasitost.....	23

4.1.2	Cílová hodnota integrované hlasitosti (target level)	24
4.2	Doporučené hodnoty dle EBU R 128 pro vysílání	25
4.3	Normalizace hlasitosti pomocí metadat	26
4.4	Film	26
4.5	Mimoevropské vysílací normy.....	27
4.5.1	ATSC A/85 v USA	27
5	Normalizace hlasitosti v internetových audio a audiovizuálních médiích	29
5.1	Platformy pro audio streaming.....	30
5.1.1	Normalizace dle skladby versus alba	30
5.1.2	Spotify	31
5.1.3	Apple Hudba a Podcasty (bývalé iTunes)	32
5.2	Video portály a video streaming.....	33
5.2.1	YouTube.....	33
5.2.2	Streaming filmů a seriálů.....	35
5.3	Rozhlasové a televizní stanice distribuované přes internet	36
5.4	Webové portály bez normalizace.....	36
6	Důsledky pro práci zvukového mistra.....	38
6.1	Omezení hlasitostního algoritmu	38
6.2	Normalizace z pohledu tvůrce.....	39
6.2.1	Čtení hlasitostních měřičů.....	39
6.2.2	Kalibrace poslechu.....	41
	Diskuze	42
	Závěr.....	45
	Seznam literatury.....	47

Seznam příloh

Příloha 1: Slovník pojmů kolem normalizace hlasitosti.....	50
Příloha 2: Řetězec pro měření výstupu neznámých aplikací.....	51

Seznam použitého označování a zkratk

AAC – Advanced Audio Coding – ztrátový kodek

AC-3 – ztrátový kodek společnosti Dolby

AES – Audio Engineering Society

BWF – Broadcast Wave Format – rozšíření formátu Microsoft WAV

ATSC – Advanced Television Systems Committee – regulátor TV vysílání v USA

A/D – převodník signálu z analogové do digitální podoby

D/A – převodník signálu z digitální do analogové podoby

DAW – Digital Audio Workstation – software pro zpracování zvuku

dB – decibel

dBFS – decibel vztažený k nejvyšší hodnotě digitálního záznamu

dBTP – skutečná špička (true peak) v dBFS

EBU – European Broadcasting Union – Evropská vysílací unie

Gain – zisk signálu vyjádřený v dB

HPF – high-pass filter – horní propust

IIR – infinite impulse response – označení filtru s nekonečnou impulzní odezvou

ITU – International Telecommunication Union – Mezinárodní telekomunikační unie

ITU-R – radiokomunikační skupina v rámci ITU

L, R, C, LFE, Ls, Rs – označení jednotlivých kanálů ve vícekanálovém systému

L_K – integrovaná hlasitost

LKFS – jednotka integrované hlasitosti totožná s LUFS

LPF – low-pass filter – dolní propust

LUFS – jednotka integrované hlasitosti totožná s LKFS

Master – finální podoba nahrávky určená k distribuci

Mastering – proces tvorby masteru, závěrečná část postprodukce

PL – programme loudness – integrovaná hlasitost

Streaming – v kontextu této práce distribuce mediálního obsahu přes internet

TP – true peak – skutečná špička

Úvod

Při práci se zvukem dříve nebo později tvůrce narazí na otázku výsledné hlasitosti díla. Jak „nahlas“ to bude znít jinde? Nebude moje práce v porovnání s ostatním audio materiálem hrát příliš slabě nebo naopak příliš silně? Porovnávání s „ostatním audio materiálem“, ale pomoci nemusí, protože se těžko hledá obecná shoda. A tuto zkušenost s nekonzistentností výsledků naší práce bohužel mají i běžní uživatelé a oprávněně si pak stěžují. To všechno jsou i mé motivace, proč jsem se do tématu normalizace hlasitosti ponořil hlouběji. I když stojím na té „výrobní straně“ procesu, tak jsem v problematice před psaním této práce plně zorientovaný nebyl a i jako prostý uživatel jsem se cítil frustrovaný z rozdílů v hlasitostech napříč médii. Nejen odbornému čtenáři tak touto prací nabízím vhled do problematiky.

Práce si proto klade za cíl prozkoumat a shrnout současné znalosti i současnou praxi normalizování hlasitosti od původních norem a doporučení. Jelikož se internet stal univerzální mediální platformou současnosti i budoucnosti, tak se na něj tato práce o to více zaměřuje a zkoumá jak jednotliví poskytovatelé audio a audiovizuálního obsahu s normalizací hlasitosti nakládají.

Nejpokročilejší je v současnosti (2021) normalizace hlasitosti v rámci vysílacích platform, kdy například v Evropě platí norma *EBU R 128 (1)* a v USA norma *ATSC A/85 (2)*. Tyto normy se opírají o systém pro měření hlasitosti navržený mezinárodní vysílací unií (3), který se neformálně souhrnně označuje jako „Loudness unit“, někdy též LUFS anebo ITU a kterému se podařilo vnést aspoň částečný řád do hlasitostní válkou poznamenané praxe práce se zvukem.

Stejný systém výpočtu integrovaly do svých aplikací některé internetové streamingové platformy, ale ani mezi nimi, a zvláště pak mezi nimi a světem terestriálního vysílání, zatím není jednotna v jeho implementaci, což vytváří zmatek nejen na straně uživatelů, ale i tvůrců obsahu. Nabízí se i otázka, jak docílit (a zda je to vůbec možné) jednotného audio světa, kde aspoň na té odborné straně nebudou kolem hlasitosti stávající otazníky.

V kapitole 1 shrnuji akustické a psychoakustické znalosti důležité v kontextu této práce. Je zde i popsán fenomén války hlasitostí, která zavedení normalizace hlasitosti nepřímo vynutila.

V kapitole 2 se zabývám existujícími typy měřících zařízení vyjma systému loudness unit. Tomu je věnována samostatná kapitola 3, ve které jsou jednotky detailně rozebrány po technické stránce.

Kapitola 4 řeší obecné principy a strategie normalizace hlasitosti a představuje evropskou a americkou normu pro terestriální vysílání.

Na kapitolu 4 tematicky navazuje kapitola 5, kde se zabývám normalizací hlasitosti na internetu. Stránky a služby, které plné specifikace neposkytují, byly aspoň přeměřeny pomocí (převážně) softwarového testovacího řetězce, jehož princip je vysvětlen v příloze 2.

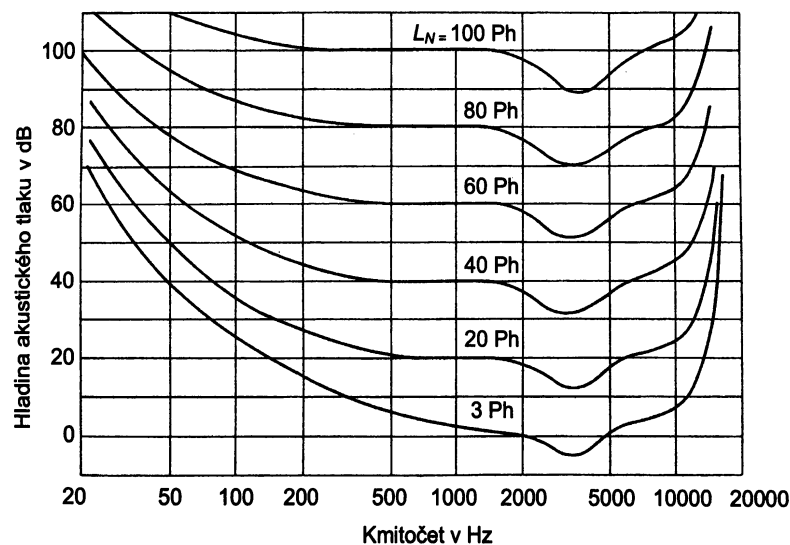
V kapitole 6 se zabývám důsledky pro praxi, jelikož celý systém normalizace je stále relativně nový a zatím není využit jeho potenciál.

1 Hlasitost a válka hlasitostí

1.1 Vjem hlasitosti

Hlasitostí se rozumí „vlastnost sluchového vnímání, podle které lze uspořádat zvuky na stupnici v rozsahu od tichých po hlasité“ (4). Objektívni veličinou, která nejbližší odpovídá hlasitosti, je hladina akustického tlaku (SPL), ale vjem hlasitosti je závislý i na frekvenčním složení a délce trvání zvuku. (5) (6)

Čisté tóny o stejné intenzitě vnímá sluchový orgán s rozdílnou hlasitostí v závislosti na frekvenci. Tuto závislost popisují křivky stejné hlasitosti (dle tvůrců různě uváděné jako Kingsburyho, Fletcher-Munsonovy, Zwicker-Feldtkellerovy), které popisují hladiny, na kterých mají čisté tóny stejnou hlasitost (obrázek 1.1). Hodnoty jednotlivých hladin jsou vztažené k čistému tónu o frekvenci 1 kHz, respektive pro čistý tón o frekvenci 1 kHz platí, že se jeho hladina akustického tlaku (v dB SPL) rovná jeho hladině hlasitosti, jejíž jednotkou je fón (Ph). (5) (6)



Obrázek 1.1: Křivky stejné hlasitosti znázorňující frekvenční závislost vjemu hlasitosti. (5)

Stejně jako mají na hlasitost vliv ostatní subjektivní vlastnosti zvuku, tak naopak hlasitost ovlivňuje je, zejména pak výšku a barvu. (5) (6)

Z nestejně vzdálenosti křivek stejné hlasitosti (obrázek 1.1) vyplývá důležitá psychoakustická skutečnost, že při srovnávání dvou zvuků, jejichž objektivní vlastnosti jsou až na hladinu akustického tlaku shodné, bude silnější z nich považovaný posluchačem

za lepší a teoreticky barevně bohatší, což je způsobeno vlivem vyšší vnímané hlasitosti okrajových pásem spektra. (7)

1.2 Hlasitost signálu

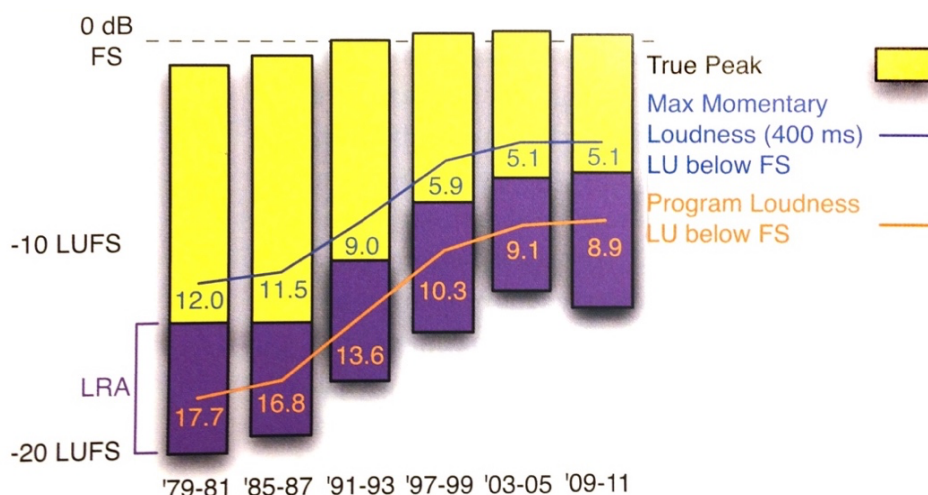
Jelikož se tato práce věnuje normalizaci hlasitosti v digitálních médiích, je třeba na začátek užití termínu hlasitost zasadit do kontextu. Dle definice je hlasitost vjem a spojení hlasitost signálu tak není přesné. Správněji by mělo být řečeno „hlasitost, která danému signálu odpovídá“, ale ani to stále nevyhovuje, protože bez znalosti citlivosti reprodukční aparatury nelze přisoudit digitálnímu signálu hladinu hlasitosti. Vzhledem k zažitým konvencím bude i tak spojení „*hlasitost signálu/nahrávky/média*“ v této práci užíváno, a to ve významu energetické veličiny signálu úměrné vnímané hlasitosti, kterou jsme schopni alespoň přibližně měřit pomocí systému ITU-R BS.1770 popsaného dále v práci.

1.3 Válka hlasitostí

Válka hlasitostí (*loudness war/race*) je pojmenování pro neustálé zvyšování hlasitosti nahrávek za cenu výrazných dynamických úprav. Do určité míry provázela hudební, převážně populární průmysl od vzniku záznamu zvuku, ale výrazně se urychlil s přechodem na digitální audio nosiče, zejména CD.

Motivací pro tento trend je snaha být hlasitější než ostatní, protože při porovnání dvou až na úroveň signálu identických nahrávek je hlasitější z nich subjektivně hodnocená jako zvukově kvalitnější. Rozhodnutí posluchače o koupi nahrávky tím může být (třeba jen podvědomě) ovlivněno, což motivovalo nahrávací společnosti ke stále vyšším hlasitostem. Hlasitostní válka zasáhla i rádiové stanice, které se předháněly (a v rámci FM vysílání se bohužel některé stále ještě předhánějí) v množství provedených dynamických úprav, aby se při přeladování rádia posluchač rozhodl pro hlasitější z nich. Stejně prostředky využívaly před zavedením normalizace v televizním prostředí i reklamní příspěvky pro upoutání pozornosti. (7)

Před přechodem k digitálním médiím existovala válka hlasitostí jen v omezené podobě, protože analogovému zařízení hrozilo vlivem signálu s příliš vysokou úrovní dokonce poškození. S nástupem digitálních médií v čele s CD, které měly přinést hlavně velký dynamický rozsah, došlo postupně v populární hudbě vlivem války hlasitostí



Obrázek 1.2: Postupné zvyšování průměrných hodnot hlasitosti (oranžově) na přelomu tisíciletí, ale za snižování dostupného headroomu (žlutě) dynamickými úpravami signálu. (7)

i k opaku. Na obrázku 1.2 je znázorněn trend ve zvyšování hlasitosti a snižování headroomu. (7)

Na extrémní podobě války hlasitostí se podílelo několik faktorů. Přechod do digitálního zpracování signálu přinesl zdokonalení techniky umožňující pokročilejší, ale zároveň i agresivnější metody komprese a limitace signálu. Dalším faktorem byla ztráta reference, kterou do té doby bylo běžné normování signálu (a s ním v určité toleranci i hlasitosti) na 0 VU. Novější digitální měřiče ukazují špičky signálu, které ale neodpovídají hlasitosti. Dynamicky upravený signál s nižší špičkou je pak možné zesílit (normalizovat podle špičky / peak normalisation) a se započtením již zmíněných marketingových důvodů nastalo s počátkem hlasitostní války postupné posouvání hranic výše. Při porovnávání CD ze začátku devadesátých let s těmi po roce 2000 hrozí posluchači až absurdní skoky v úrovních hlasitosti. (7)

Populární hudební průmysl se dostal do začarovaného kruhu, kdy album, které nebylo takto dynamicky zkomprimováno, bylo v porovnání s ostatními nevýrazné a tvůrci tak čelili tlaku, aby album upravili obdobně. Snímky již vydané tvořily předobraz tvorby nové, čímž se celý problém jen umocňoval. Dobrovolně bylo velmi těžké z tohoto začarovaného kruhu vystoupit. (7)

Dynamické úpravy snímků ale dosáhly svého maxima, kdy takto upravené snímky přestaly být atraktivní pro posluchače i umělce. Dynamický rozsah některých nových populárních snímků byl nižší než rozsah voskových válečků. (7) V hlasitosti už nebylo kam stoupat, ale vrátit se bylo obtížné.

V roce 2006 vznikla první verze dokumentu BS.1770 a algoritmu pro výpočet hlasitosti a v letech 2009/10 vznikly první verze dokumentů upravujících hlasitostní normalizaci v sítích EBU a ATSC. CD sice nebyla konzistentní mezi sebou, ale delší poslech více než jednoho (většinou hlasitostně konzistentního nosiče) byl technicky složitý a pro přehrávání z jiného alba bylo nutné CD vyměnit. MP3 přehrávače a hudební streaming umožnil přehrávání napříč alby, kde skoky v hlasitosti mezi skladbami byly obtěžující, což hudební streamovací služby motivovalo k zavedení normalizace hlasitosti. Hlasitostní válka tak postupně začala ustupovat. (7)

2 Měření úrovně signálu

Vizualizace zvuku představuje důležitou zpětnou vazbu pro kohokoliv, kdo s ním pracuje. Měřič schopný zobrazit „aktuální hodnotu“ pak zvukovému mistrovi slouží ke dvěma účelům. Ve vztahu k technickým parametrům záznamu či přenosového média a ve vztahu právě k hlasitosti. V následující části budou představeny nejvýznamnější způsoby měření. Nejnovějšímu systému LKFS/LUFS je věnována samostatná kapitola 3.

2.1 Analogový signál

2.1.1 VU metr

S VU metrem (volume¹ unit meter) se je možné ještě stále setkat na analogových zařízeních nebo v přenesené podobě na digitálních, která analogové chování emulují. Jedná se o analogový ručičkový měřič. Proud protékající cívkou v magnetickém poli pohybuje ručičkou, která je brzděna pružinou. Hmotnost ručičky a tuhost pružiny způsobuje, že ručička není schopna reagovat na příliš rychlé změny, zejména pak impulzivní signály. Zároveň příliš malé proudy nejsou schopny ručičku rozpohybovat, a tak VU metr nabízí pouze přibližné integrované hodnoty (s časem integrace přibližně 300 ms) kolem referenční hodnoty. VU metr vytvářel z hlediska hlasitosti normalizované prostředí (byť nedokonalé). (7) (8) (9)

2.1.2 Peak programme meter

Peak programme meter (PPM) je měřič, který dokáže měřit přibližnou špičku analogového signálu (quasi-peak). PPM se dočkal mnoha verzí a mnoha stupnic v závislosti na vysílacích normách v konkrétních zemích, ale i na základě konkrétních vysílacích společností (například BBC PPM). Je typický rychlou reakcí na špičku (integrační čas přibližně 5–10 ms) a dlouhým doběhovým časem ukazatele (1–3 vteřiny). Na rozdíl od VU byl převeden i do digitální podoby. (9) (10) Před zavedením normy R 128 se používal k normalizaci úrovně při vysílání. (11)

¹ Angličtina používá pro neobdobné označení hlasitosti slovo *volume*, které ale v přesném významu znamená *objem*, v přeneseném *množství něčeho* (44) a pro odborné užití je velmi zavádějící.

2.2 Digitální signál

2.2.1 Peak meter

Na rozdíl od analogového signálu má digitální signál jednoznačnou a nepřekročitelnou horní hranici 0 dBFS, za kterou již signál není možné kvantovat a může být slyšitelně zkreslený. Znat špičku signálu má tedy své opodstatnění. Měřič špičky (peak meter) se pak v závislosti na výrobci liší v použité stupnici (lineární/logaritmická), dolní hranici stupnice a v použitých měřicích časech (v angličtině označovaných ballistics). Někdy jsou tyto stupnice doplněny i o měřič RMS (následující kapitola) nebo jinak integrující metriku. (12)

Měření podle špičky signálu nicméně nevypovídá nic o jeho hlasitosti. (1) (2) (3) (7) (9) (12) (13) (14) (15)

2.2.2 RMS meter

Root (of the) mean (of the) square (RMS) je efektivní hodnota signálu. Více než jako samostatný měřič se používá jako sekundární stupnice k peak metru. Při použití integračních časů 300 ms RMS měřič slouží podobně jako VU metr. (12) Podobně jako VU je sice k odpovídající hlasitosti blíže než peak metr, ale stále nedostatečně, protože obě metody měření nijak nereflektují vlastnosti lidského sluchu, zejména křivky stejné hlasitosti. (16)

3 Systém jednotek LUFS

V roce 2006 představila mezinárodní telekomunikační unie (ITU) ve svém doporučení ITU-R BS.1770 algoritmus pro měření hlasitosti. Na něm je založen systém jednotek a pozdějších doporučení EBU i ATSC a využívá se pro normalizaci hlasitosti ve vysílacích řetězcích i v některých internetových platformách.

Jednotka, kterou toto doporučení představilo se nazývá *loudness, K-weighted, relative to full scale* (LKFS), tedy ve volném překladu „hlasitost vážená pomocí K-filtru vzhledem k plnému rozsahu“. Nicméně EBU tuto jednotku ve svých normách a doporučeních přejmenovala na *loudness unit, relative to full scale* (LUFS), tedy „hlasitostní jednotka vzhledem k plnému rozsahu“. LKFS a LUFS tedy znamenají to samé. (1) (13) V této práci bude nadále užíváno označení LUFS.

V této práci budou uváděny i české ekvivalenty anglické terminologie podle překladů příslušných norem Rady pro rozhlasové a televizní vysílání. (17)

3.1 Algoritmus pro výpočet hlasitosti dle ITU-R

Doporučení ITU-R BS.1770 popisuje algoritmus výpočtu hlasitosti L_K s jednotkou LUFS. Výsledkem algoritmu je decibelová hodnota, která odpovídá hlasitosti signálu. Zároveň se zde definuje výpočet skutečné špičky (*true peak*, zkráceně *TP*) s jednotkou *dBTP*, která poskytuje přesnější informace o špičce signálu po převodu do analogové domény než maximální špička (*sample peak*).

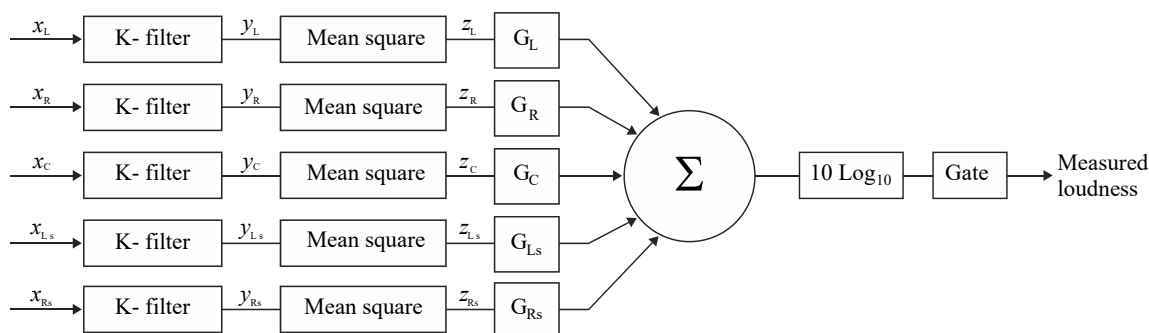
3.1.1 Výpočet hodnoty hlasitosti

G. A. Souldre provedl poslechový test, ve kterém zjišťoval přesnost několika algoritmů pro výpočet hlasitosti monofonního signálu. (18) Jako nejpresnější vyšel algoritmus $Leq(RLB)^2$, ze kterého výpočet dle ITU-R BS.1770 vychází.

Na obrázku 3.1 je znázorněno blokové schéma hlasitostního algoritmu podle ITU-R. První krok je frekvenční váhování tzv. *K-filtrem*, což je souhrnné označení pro dva filtry,

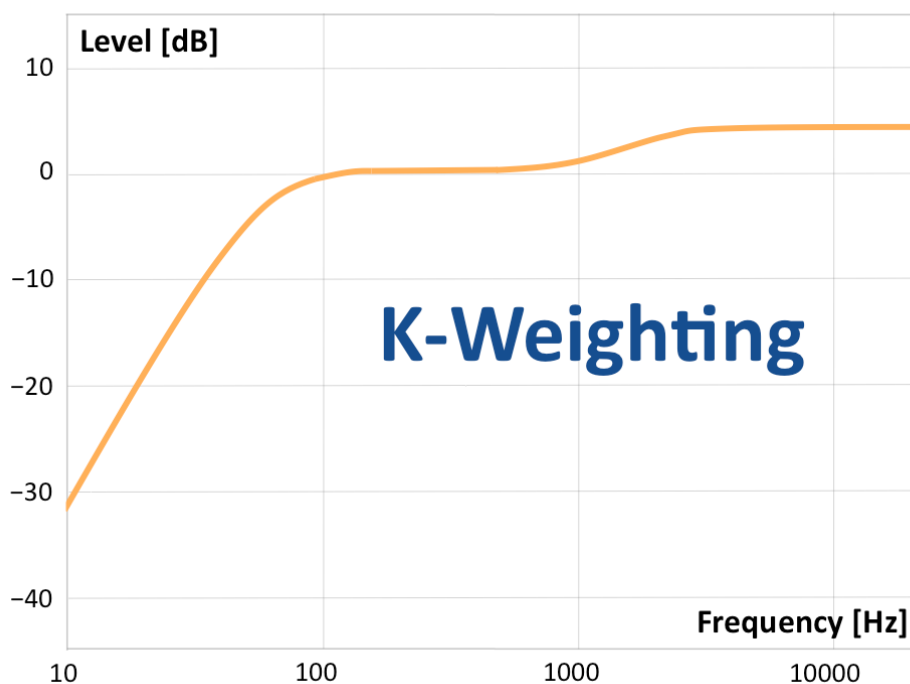
² **Leq** – equivalent sound level; výpočet hlasitosti na základě energie signálu vůči referenčnímu signálu vyjádřený v dB. (18)

RLB – značí váhování pomocí revised low frequency **B**-curve filtru (HPF). Více v (18) a (3).



Obrázek 3.1: Zjednodušené blokové schéma vícekanálového algoritmu hlasitosti. (3)

kteří mají za úkol přiblížit signál k psychoakustickému modelu percepcie hlasitosti. Prvním je shelving filtr, který kompenzuje vliv stínu hlavy, a druhým je *RLB filtr*², který v měření potlačuje velmi nízké frekvence (jeho zpřesňující vliv na měření prokázal právě Souloudre (18) ve své práci). Souhrnná charakteristika těchto filtrů je znázorněna na obrázku 3.2.



Obrázek 3.2: Souhrnná křivka K-filtru využívaná k měření hlasitosti. (13)

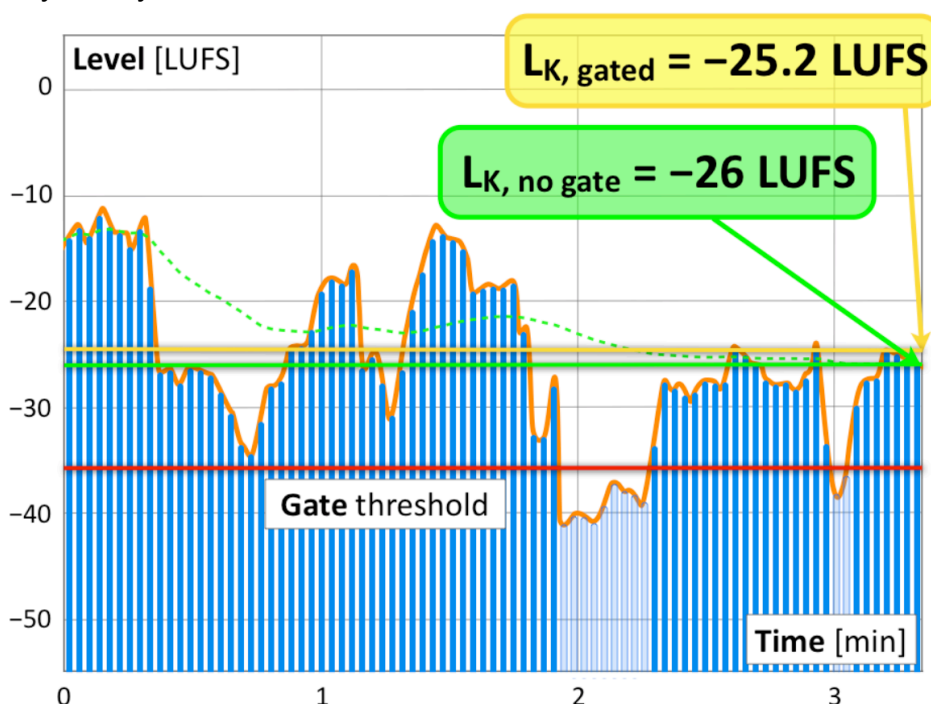
Signál je rozdělen do bloků o délce 400 ms s překryvem minimálně 75 %. Hlasitost jednoho bloku se počítá jako součet středních výkonů filtrovaných signálů jednotlivých kanálů, kterým jsou přiřazeny váhy v závislosti na typu kanálu (obrázek 3.1). LFE kanál je výslovně vyřazen, ale ITU-R připouští jeho možné zohlednění v budoucnosti. (3)

Hlasitost (L_K) bloku o délce T se vypočte podle následujícího vzorce:

$$L_K = -0,691 + 10 \log_{10} \sum_i G_i \cdot \frac{1}{T} \int_0^T y_i^2 dt$$

kde i je index kanálu (L, R, C, Ls nebo Rs), G_i koeficient příslušného kanálu (například pro L a R je to vždy 1) a y_i je signál v kanálu i (správněji $y_i(t)$ jako funkce času). (3)

Hodnoty jednotlivých bloků jsou průměrovány, ale vyřazeny z průměrování (*gating*) jsou ty bloky, jejichž hlasitost je menší než -70 LUFS (*absolutní práh*), a ty, jejichž hlasitost je o 10 dB nižší než průběžná hodnota integrované hlasitosti (*relativní práh*)³. Tento krok brání vlivu ticha a velmi tichých pasáží na celkovou hlasitost – obrázek 3.3. Kdyby tomu tak nebylo, tak by dva signály, lišící se pouze délkou ticha v pauze, vykázaly rozdílnou hlasitost L_K , což neodpovídá tomu, jak hlasitost vnímáme. Nekompletní bloky jsou vyřazeny z měření.



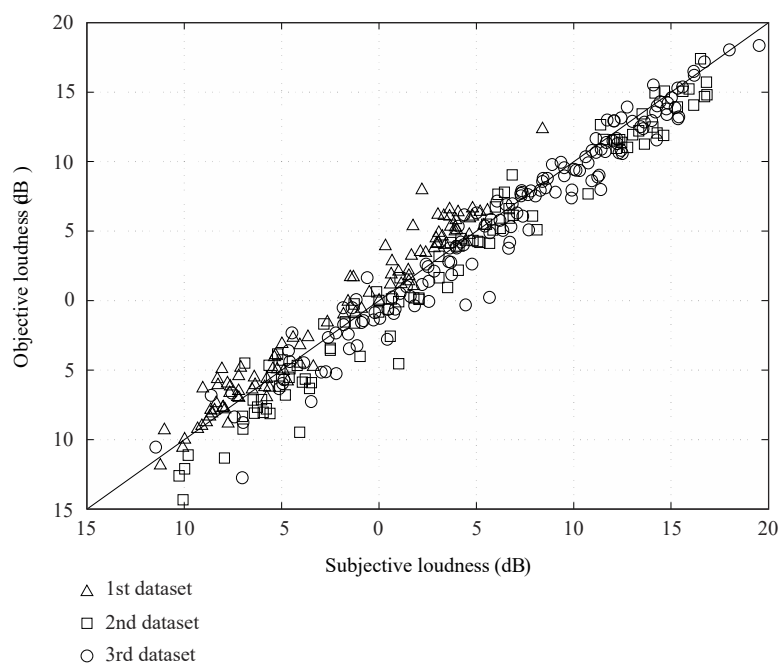
Obrázek 3.3: Význam gatingu na hlasitost signálu. Bloky pod relativním prahem (červeně vyznačený) jsou při výpočtu $L_{K, \text{gated}}$ vyřazeny z průměrování. (13)

Vypočtená hodnota hlasitosti L_K měřeného signálu se nazývá *integrovaná hlasitost* (též *hlasitost pořadu* (17); *integrated/programme loudness*) a její jednotkou je LUFS

³ Prvních několik bloků využívá pouze absolutní práh, protože aplikovat relativní práh je možné až po zjištění hodnoty hlasitosti.

(*loudness unit relative to full scale*). LUFS je decibelová hodnota⁴, takže zůstává kompatibilní se zaběhnutou praxí. Zesílením signálu o 1 dB, vzroste hlasitost o 1 LU. (3)
(13)

ITU podrobila tento algoritmus poslechovému testu, ve kterém respondenti porovnávali hlasitost vzorku a reference. Na obrázku 3.4 je souhrnný výsledek všech 336 použitých vzorků. Koeficient korelace mezi vypočtenou hlasitostí L_K a objektivní hlasitostí byl 0,977, což potvrdilo účinnost algoritmu. (3)



Obrázek 3.4: Výsledky poslechového testu (na ose y je hlasitost L_K a na ose x subjektivní hlasitost dle respondentů). (3)

3.1.2 Dialogue gating

Zejména ve filmovém prostředí a americké televizní praxi se používá měření hlasitosti podle mluveného slova (*dialogue normalisation*). Pro tyto účely existuje nepůvodní úprava algoritmu ITU-R BS.1770, která gating relativní k integrované hlasitosti (popsaný na straně 13) nahradí automatickým rozpoznáváním řečového signálu a ostatní prvky do integrace nejsou zahrnuty. Algoritmus *Dialogue Intelligence*TM vyvinula firma Dolby a jeho kód je otevřený například pro výrobce měřičů hlasitosti. (2) (19)

⁴ Teoreticky bychom ji mohli označovat například $dB(LUFS)$ jako *hladinu integrované hlasitosti*.

Je důležité zdůraznit, že při dialogue gatingu je vypočtená hodnota hlasitosti odlišná od té podle BS.1770, ale bohužel označená stejnou jednotkou LUFS/LKFS s dodatkem „*dialogue gated*“ nebo jen „DIA“. Příklad tohoto rozdílu je na obrázku 3.5.

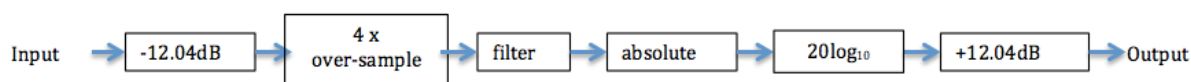


Obrázek 3.5: Stejný materiál (v tomto případě díl seriálu) může zobrazit výrazně odlišnou hodnotu hlasitosti při použití upraveného algoritmu BS.1770 s gatingem podle dialogů (vlevo), ve srovnání s relativním gatingem (vpravo).

3.1.3 True peak

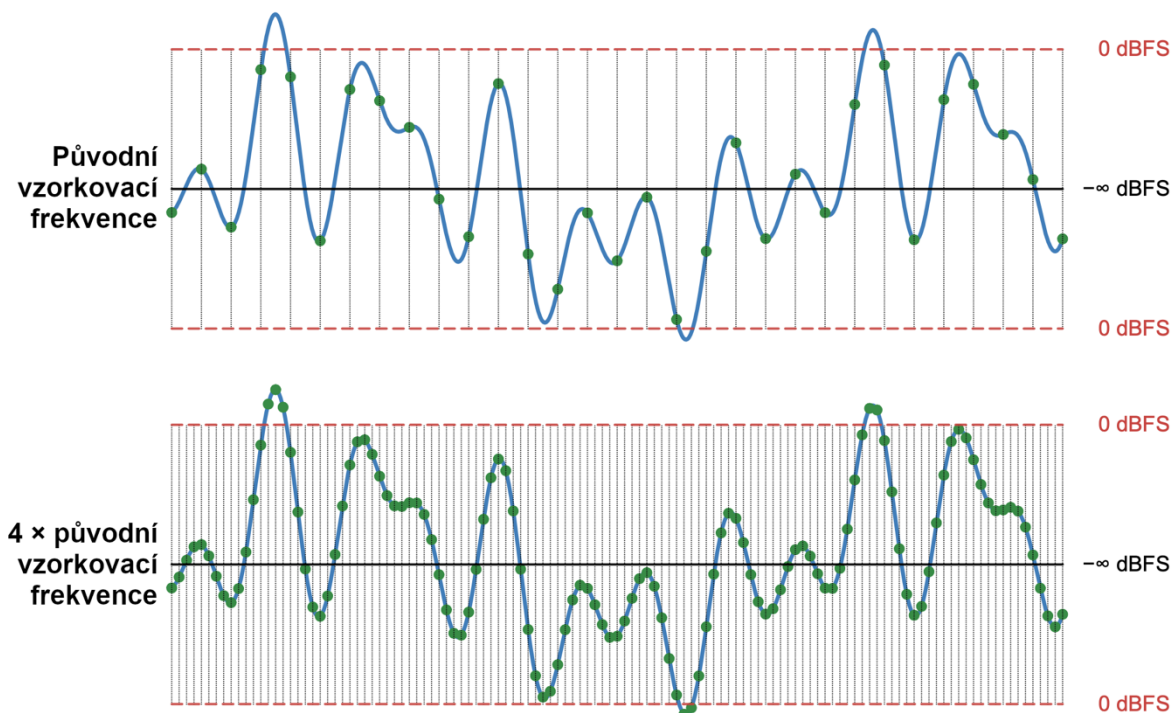
Digitální hardware i software většinou zobrazuje maximální špičku signálu v dBFS (sample peak). Tato hodnota ale nemusí odpovídat skutečné maximální úrovni signálu po D/A převodu. Špičky signálu, které sice v digitální doméně nepřesahují 0 dBFS, mohou být po převodu do analogové podoby zkresleny.

Na základě této skutečnosti zavádí ITU-R BS.1770 parametr *true peak*, neboli „maximální skutečnou špička“ (17), zkráceně *TP* s jednotkou *dBTP*. Na obrázku 3.6 je zjednodušené schéma výpočtu jeho hodnoty. Signál je třeba převzorkovat do násobně vyšší vzorkovací frekvence (oversampling). Pro vzorkovací frekvence do 48 kHz ITU doporučuje čtyřnásobek, pro vyšší frekvence dvojnásobek původní vzorkovací



Obrázek 3.6: Blokové schéma algoritmu pro výpočet hodnoty true peak. Jelikož zkoumáme hodnoty vyšší než maximální, je třeba operaci provádět ve formátu float. (3)

frekvence. (3) Nejvyšší absolutní hodnoty vzorků jsou poté dostatečně blízko odpovídající hodnotě signálu po D/A převodu a algoritmus je tak schopen odhalit skrytý clipping mezi vzorky tak, jak to znázorňuje obrázek 3.7. V horní části je samplovaný úsek, a ačkoliv jsou všechny vzorky pod maximální hodnotou 0 dBFS, signál po převedení do analogové domény dosahuje i hodnot vyšších, kterých ale nemusí dosáhnout D/A převodník a může dojít ke zkreslení. (3) Takový signál by vykázal kladnou hodnotu TP, kterou by odhalil právě oversampling naznačený v dolní části obrázku 3.7.

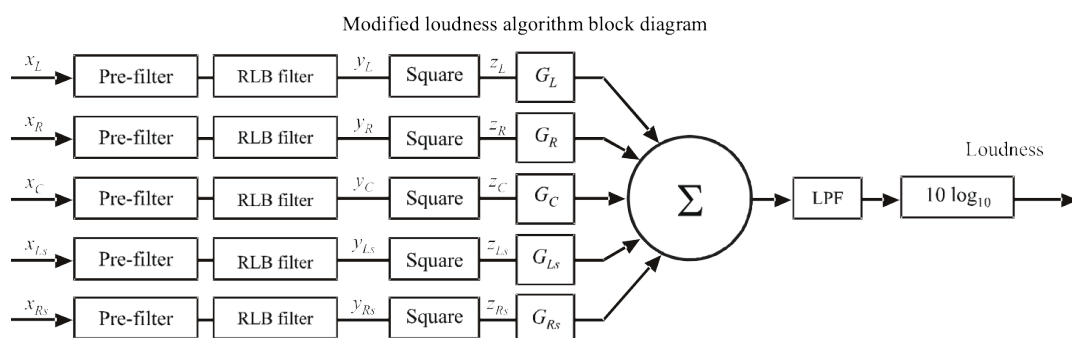


Obrázek 3.7: Odhalení skrytých špiček pomocí oversamplingu.

3.1.4 Momentary loudness

Momentary loudness (okamžitá hlasitost) se vypočte pomocí upraveného algoritmu BS.1770 na obrázku 3.8. Po sečtení středních výkonů projdou hodnoty skrz IIR dolní propust (v analogové doméně odpovídající pasivnímu integračnímu článku), která zohledňuje vliv délky podnětu na vjem hlasitosti. Pro uvědomění hlasitosti potřebujeme více než 100 ms (20) od začátku podnětu, a naopak vjem přetrvává i po jeho skončení. Impulzivní signály a skoky v hlasitosti by bez této úpravy mohly vykazovat vyšší hlasitost neodpovídající vjemu. Měřič by ale i při 100ms integraci byl příliš rychlý a tím pádem nečitelný, a proto se používá hodnota 400 ms. Pro okamžitou i krátkodobou hlasitost (kapitola 3.1.5) se gating nepoužívá, protože by bylo matoucí, kdyby v tichu

a tichých pasážích měřič nezobrazoval signál. (16) (21) Okamžitá hlasitost představuje určitou analogii k VU metru. (7) (16)



Obrázek 3.8: Upravený algoritmus pro výpočet okamžité hlasitosti. Integrace se provádí pomocí LPF. (16)

3.1.5 Short-term loudness

Short-term loudness (krátkodobá hlasitost) se počítá podobně jako integrovaná hlasitost. Využívá stejný výpočet pro jednotlivé bloky, ale neaplikuje gating a místo integrování celého signálu se zabývá vždy časovým intervalem 3 vteřiny. (Některé hlasitostní měřiče ji nazývají *sliding loudness* a umožňují nastavit integrační čas.) Minimální obnovovací frekvence této hodnoty je 10 Hz. (21) Short-term loudness představuje okamžitý náhled na signál z hlediska jeho příspěvku k integrované hlasitosti. Zároveň je důležitou metrikou pro posuzování krátkých pořadů, kde může být normalizace pomocí integrované hlasitosti zavádějící (pro nedostatek dat). (13)

3.2 Jednotky dle EBU

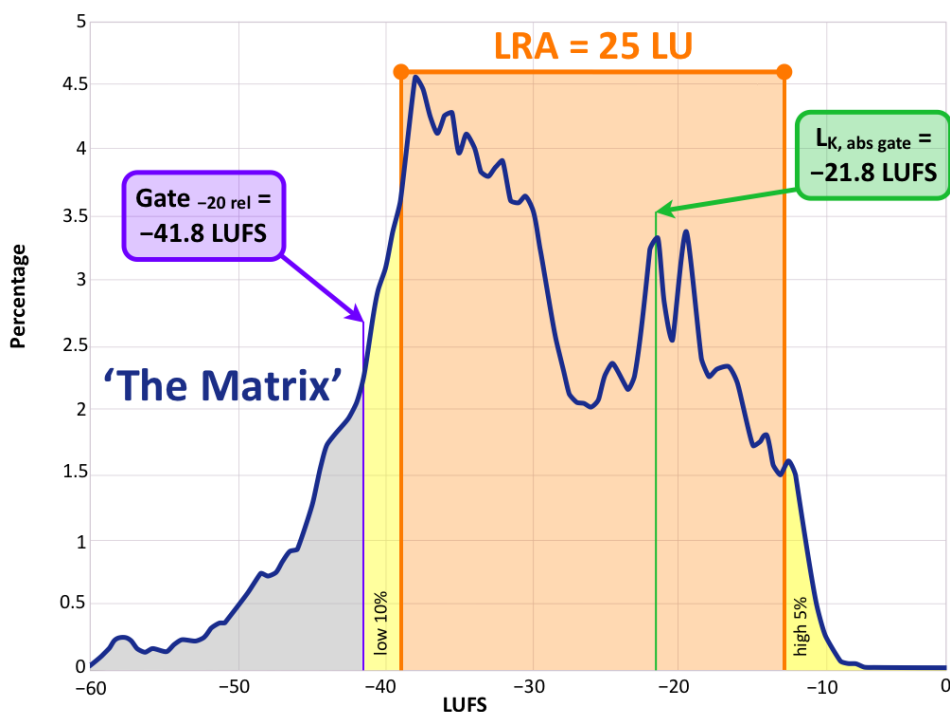
Na úsilí ITU-R navázala European Broadcasting Union (EBU) zřízením pracovní skupiny *PLoud*. Vznikla tak norma *EBU R 128*, která společně s dodatky a doprovodnými normami doporučuje hodnoty hlasitosti v rozhlasovém a televizním vysílání a k systému dle ITU-R BS.1770 přidává další jednotky. Více k obsahu normy R 128 v kapitole 4.2. Zpětně ovlivnila i definici integrované hlasitosti zavedením relativního prahu (doplňného do původní normy od BS.1770-2). (22)

Norma v návaznosti na ITU-R BS.1771 zavádí jednotku *LU (loudness unit)* ve dvou kontextech. V kontextu relativním se jedná o vyjádření rozdílu dvou absolutních hodnot LU. Například ± 1 LU kolem konkrétní hodnoty nebo při vyjádření, že hodnoty -26 LUFS a -18 LUFS se liší o 8 LU. V obou případech LU představuje decibel. V absolutním kontextu se za 0 LU považuje vždy doporučená hodnota dle příslušné normy

(například v systému EBU 0 LU = -23 LUFS). (21) (23) Název *loudness unit* tak navazuje na historickou jednotku VU neboli volume unit.

3.2.1 Loudness range

Hodnota rozsahu hlasitosti (*loudness range*, zkráceně *LRA*) je analytická jednotka, která odpovídá dynamickému rozsahu měřeného signálu. Vypočte se na základě statistického rozdělení naměřených hodnot hlasitosti. Hodnoty hlasitosti jsou integrovány po blocích délky 3 sekund s minimálně 2,9 s překryvy. Následuje kaskádový gating, kdy je nejprve aplikován práh absolutní a všechny bloky pod -70 LUFS jsou vyřazeny, poté jsou vyřazeny všechny bloky pod relativním prahem -20 LU vůči integrované hlasitosti a teprve z těchto hodnot je stanoven rozdíl mezi 10. a 95. percentilem (10 % dolních hodnot a 5 % horních je záměrně ignorováno). Princip je graficky znázorněn pomocí příkladu na obrázku 3.9. Tento postup byl zvolen s ohledem na signály vyskytující se v praxi a s ohledem na vlastnosti lidského sluchu. Je tak zaručeno, že hodnota LRA není zkreslená impulsivními zvuky nebo z opačné strany dozvuky či například šumem sálu v pauzách. Norma Tech 3342 nedefinuje úplně přesný způsob výpočtu (zejména pak řešení překryvů bloků), takže různé měřiče hlasitosti mohou zobrazovat



Obrázek 3.9: Graf znázorňuje četnost konkrétních hodnot krátkodobých hlasitostí původního znění filmu Matrix. Fialově je znázorněný relativní práh a oranžově oblast mezi 10. a 95. percentilem. Rozdíl mezi příslušnými krajními hodnotami představuje LRA. (13)

různé hodnoty. EBU pouze po výrobcích požaduje zobrazení stejných hodnot (v určité toleranci) u předdefinovaných testovacích signálů. (13) (24)

Hodnota LRA je spíše orientační analytická jednotka poskytující tvůrcům a operátorům vysílání zpětnou vazbu o dynamických vlastnostech mixu. Vhodná je pro porovnávání pořadů, ale její hodnoty by měly být čteny a vnímány s větší tolerancí než hodnoty ostatních jednotek v systému LUFS, což i normy EBU přímo doporučují. Zároveň nedoporučují stanovování striktních limitů LRA. (13) (24) Z popsaneho způsobu výpočtu zároveň vyplývá, že čím kratší signál, tím menší vypovídající hodnotu LRA má. (24)

3.2.2 Další analytické hodnoty

Headroom je (při využití nových definic dle ITU/EBU) rozdíl mezi nejvyšší dosažitelnou hodnotou média a integrovanou hlasitostí konkrétního audio materiálu. Například nahrávka s integrovanou hlasitostí -31 LUFS má headroom 31 dB vůči celé digitální škále, ale v systému EBU s nejvyšší povolenou hodnotou -1 dBTP má headroom 30 dB. (7)

Peak to loudness ratio (PLR) zavedl masteringový zvukový mistr Bob Katz. Je to rozdíl mezi true peak hodnotou nahrávky a její integrovanou hlasitostí. Jednotka představuje měřítko mikrodynamiky a je vhodná pro posuzování množství dynamických úprav, zejména pak limitace u hlasitostní válkou silně poznamenaných snímků. (7)

3.3 Měřiče hlasitosti

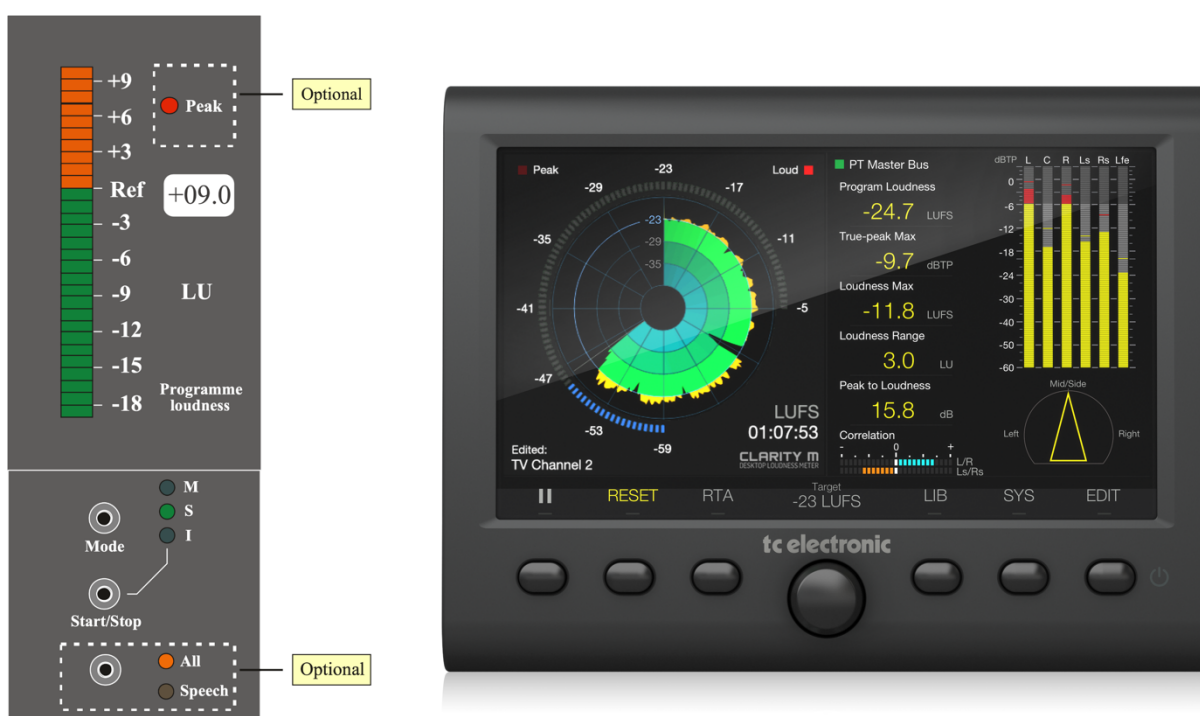
Měření hlasitosti může probíhat na konkrétních audio souborech, ale i v reálném čase. ITU i EBU vydaly své požadavky na podobu měřičů hlasitosti určených právě pro měření v reálném čase, které jsou zásadní při vysílání.

Měřič hlasitosti by měl obsahovat 3 stupnice, které nemusí být zobrazovány naráz: Stupnici *M* (*momentary*), *S* (*short-term*) a *I* (*integrated*), jejichž principy byly popsány v dřívějších kapitolách. Měl by být vybaven tlačítkem pro ovládání integrační doby „start/stop“ (dle ITU) nebo „pause“ (dle EBU) a hlavně také tlačítkem „reset“ pro uvedení měřiče do počátečního stavu. (21) (23)

Měřicí stupnice může být relativní vůči standardu (LU) nebo absolutní (LUFS). Podle ITU může být případně vybaven i volbou gatingu. (21)

EBU na měřiče doplňuje ještě ukazatel hodnoty loudness range (LRA) a zavádí termín „EBU mode“ pro měřiče, které naměří stejné hodnoty jako testovací soubor nahrávek a kde 0 LU = -23 LUFS. (23)

Dosud zmíněné vlastnosti jsou požadavky norem a doporučení, ale v praxi do podoby měřičů zasáhli jednotliví výrobci (obrázek 3.10). Měřiče vzniklé v prostředí normy ATSC A/85 mohou být doplněny dalšími funkcemi jako například *dialogue gating*.



Obrázek 3.10: Srovnání měřičů hlasitosti tak, jak jej navrhuje ITU-R BS.1771 (vlevo); a příklad reálného výrobku od firmy TC Eletronic Clarity M (vpravo). (21) (43)

4 Normalizace hlasitosti z obecného hlediska

Normalizování hlasitosti zvukového signálu je úprava jednotlivých pořadů zesílením nebo zeslabením tak, aby jejich hlasitost splňovala kritéria daná odpovídající normou nebo interními předpisy poskytovatele audio nebo video materiálu.

Smyslem normalizace hlasitosti je konzistentnost média. Přílišné skoky v hlasitosti mezi jednotlivými pořady⁵/soubory vytváří nepohodlí pro posluchače a v případě uměleckých žánrů mohou škodit celkovému dojmu. Příkladem může být hlasité reklamní sdělení navazující na film končící nízkou hlasitostí – v atmosféře případných doznívajících emocí z filmu nás skok v hlasitosti může nepřírozeně vytrhnout. Nejrozšířenější je aktuálně normalizace hlasitosti v rozhlasovém a televizním vysílacím řetězci odkud většina norem a doporučení pochází.

Jak vyplývá z předchozích kapitol nejlepším měříčem a normalizátorem hlasitosti je člověk a jeho sluchový vjem. Zatím žádný existující algoritmus nedokáže posoudit hlasitost s bezchybnou přesností tak, jako lidské ucho nebo přesněji poslechová skupina. Co ale algoritmus dokáže, je přiblížit se k určité toleranci. V případě algoritmů pro výpočet LUFS tuto toleranci naznačuje již dříve zmíněný obrázek 3.4.

Při volbě cílové hodnoty (*target level*) pro normalizaci hlasitosti proti sobě stojí dva (často protichůdné) zájmy, které vychází z poslechových podmínek posluchače. Prvním z nich je zachování audio kvality a tím druhým je přizpůsobení poslechu hlučnému prostředí, který je motivací pro zvyšování hlasitostního standardu. Jeho zvýšením je totiž zároveň i snížen dostupný headroom a více než hlasitost se poté normalizuje dynamický rozsah. Rozpor těchto podmínek je vyjádřen v tabulce 1.

Audio materiál s nízkým dynamickým rozsahem (kterým nemusí být jen výrazně zkomprimovaný popový hit, ale například i nahrávka sólového cembala či zpravodajská relace) bude ve většině poslechových podmínek ve výhodě oproti materiálu s vysokým dynamickým rozsahem (jako jsou například nahrávky symfonické hudby nebo filmy). Druhou výhodou pro takový materiál je nevyužitý headroom, který umožňuje zesílení signálu při normalizaci podle špiček, která by se dala v kontextu války hlasitostí s nadsázkou považovat za „doping“. Při normalizaci podle špičky signálu pak může být

⁵ Pořady podle definice EBU, tedy jakýkoliv ucelený a samostatný audio nebo audiovizuální úsek. (1)

takovýto materiál s nízkým dynamickým rozsahem (a především nízkým PLR) výrazně hlasitější než ostatní.

	Hlučné prostředí (např. poslech v dopravním prostředku)	Klidné prostředí (například domácí kino)
Audio materiál s nízkým dynamickým rozsahem	Poslech možný	Poslech možný
Audio materiál se sníženým, původně vysokým, dynamickým rozsahem	Poslech možný	Poslech možný, ale vlivem snížení dynamického rozsahu může dojít ke ztrátě kvality
Audio materiál s vysokým dynamickým rozsahem	Poslech nepříjemný až nemožný	Poslech možný

Tabulka 1: Rozpor poslechových podmínek a dynamického rozsahu audio materiálu.

Argumentem pro normalizaci podle špiček je často technický přístup spočívající v potřebě využití maximálního rozsahu cílového média (určité dědictví po práci s analogovými médii). I kdybychom se omezili jen na 16bitový záznam, tak cílené zanechání 12 dB nevyužitého headroomu stále znamená plně využitelných 78 dB hodnot od kvantizačního šumu. Obsah s nízkým nebo sníženým dynamickým rozsahem takový prostor ale ze své podstaty nepotřebuje a normalizace podle špičky pak slouží primárně k zesílení signálu. Oproti tomu audio obsah s velkým dynamickým rozsahem velmi pravděpodobně nebude mít tak vysokou integrovanou hlasitost a je přirozenější ve špičkách tento rozsah až k nejvyšším hodnotám pod 0 dBFS využít. V současnosti ale většina nových záznamů určená pro nejkvalitnější reprodukci má 24bitovou hloubku, která ve stejném srovnání při zanechání špiček na -12 dBFS nabízí 126 dB odstup od kvantizačního šumu, což je rozsah hodnot překračující hranice lidského sluchu a výrazně překračující hranice i kvalitně odhlučněných poslechových místností. (7)

Normalizace hlasitosti tak pomohla ukončit, nebo aspoň výrazně utlmit, nejen válku hlasitostí mezi jednotlivými televizními či rozhlasovými stanicemi, ale i mezi hudebními mastery. Jakýkoliv zvuk, který se snaží ostatní „překřičet“ za cenu vysoké míry dynamických úprav (a nízkého PLR), bude v normalizovaném prostředí zeslaben, jeho headroom zůstane nevyužitý, což snižuje míru motivace pro výrobu obdobných masterů (zbylý headroom je lepší využít pro špičky signálu místo jejich bezdůvodné

limitace). Zároveň také normalizace hlasitosti přispívá ke zvýšení posлуhačského komfortu vyšší konzistentností médií. Zavedením vysílacího limitu pro maximální krátkodobou hlasitost (short-term) se dostaly pod kontrolu reklamní spoty a upoutávky, které využívaly podobně „nekalých“ prostředků jako již zmíněné dynamicky výrazně upravené mastery. (7) (13)

V současnosti se však mediálním centrem, kde se kombinují zvuky z různých audio a audiovizuálních zdrojů, staly naše počítače, mobilní telefony či domácí kina samy o sobě, kde si program sestavujeme sami. Podrobněji se normalizaci hlasitosti právě na internetu věnuji v kapitole 5.

4.1 Principy a strategie normalizace hlasitosti

Požadavky cílových médií na hlasitost audio masteru se většinou skládají ze dvou parametrů. Jedním upravující technickou stránku obvykle pomocí maximální povolené hodnoty dBTP. Jejím omezením se zabraňuje přebuzení analogového výstupu v důsledku skrytých špiček. Smyslem maximálních hodnot nižších než 0 dBTP (většinou -1 dBTP) je ponechat bezpečnou rezervu pro průchod ztrátovým kodekem (mp3, aac...) nebo převzorkováním, kde se vlivem filtrací může nejvyšší hodnota signálu nechtěně zvýšit. Tím druhým nejčastějším parametrem, který se věnuje právě hlasitosti, je hodnota integrované hlasitosti, kterou musí master nebo pořad mít.

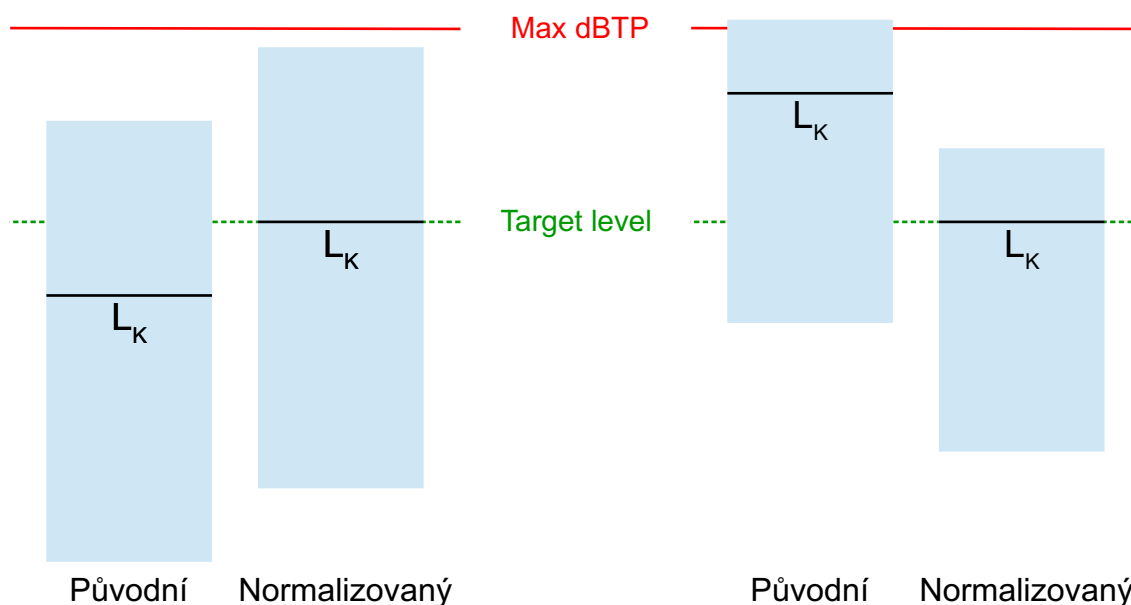
Samotná normalizace se provádí buď přímo na konkrétním obsahu, nebo pomocí metadat. (13)

4.1.1 Maximální integrovaná hlasitost

Stanovením maximální hodnoty integrované hlasitosti se vytváří bariéra vůči příliš hlasitým pořadům, zatímco ty slabší jsou ponechány beze změny (pokud je není potřeba zeslabit kvůli příliš vysoké hodnotě TP). Jediným zásahem do audia tak může být pouze jeho zeslabení. Smyslem tohoto přístupu je především stanovení limitu pro ty nejvíce dynamicky upravené mastery, jejichž hodnoty integrované hlasitosti se šplhají i nad -10 LUFS a v přímém porovnání s ostatním audio materiálem vytváří nepříjemné skoky v hlasitosti. V médiích, která maximální integrovanou hlasitost používají (v současnosti například server YouTube), se tato hodnota díky tvůrcům stává neoficiálním target levellem z podobných, víceméně marketingových, důvodů, které rozpoutaly válku hlasitostí.

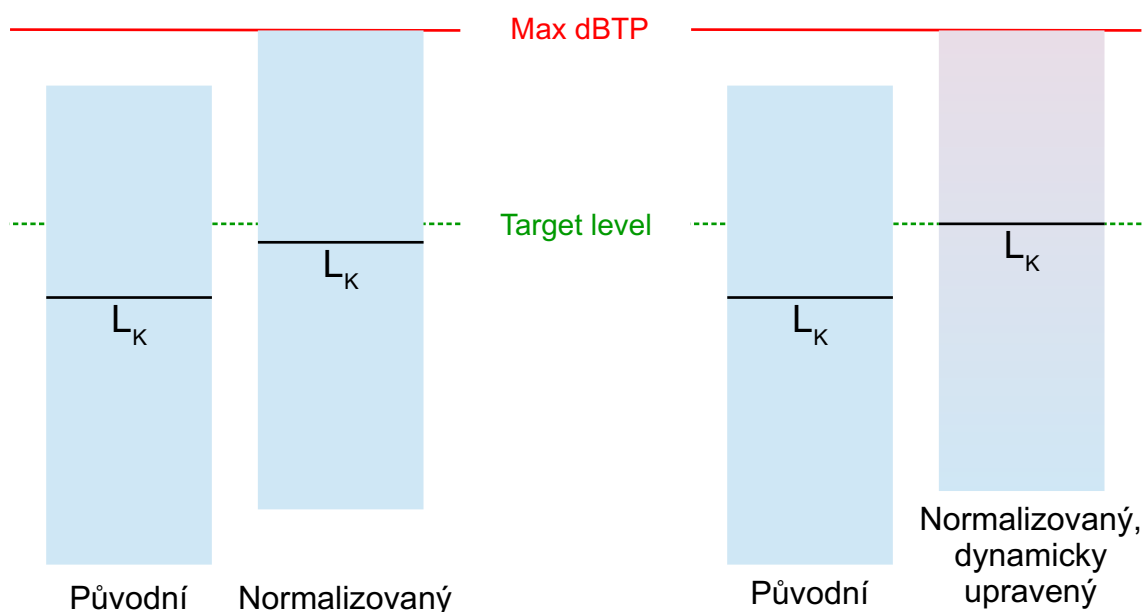
4.1.2 Cílová hodnota integrované hlasitosti (target level)

Druhou možností, používanou například v rozhlasovém a televizním vysílání je stanovit konkrétní cílovou hodnotu integrované hlasitosti, kterou musí splnit ideálně veškerý obsah⁶. Pořady silnější jsou zeslabeny na cílovou hodnotu a ty slabší jsou zesíleny, pokud to dovoluje jejich headroom. Pokud to nedovoluje, tak jsou dvě možnosti, z nichž si nejhodnější strategii volí každé médium samo. Jednou možností je audio ponechat beze změny a tou druhou je signál dynamicky upravit (ručně, či automaticky) tak aby požadavky splňoval. Zmíněné možnosti znázorňují obrázky 4.1 a 4.2.



Obrázek 4.1: Normalizace hlasitosti aplikací gainu na celý program v případě dostatečného headroomu

⁶ EBU v dokumentu Tech 3344 definuje metodiku pro vysílání, podle které stanovuje podmínky pro ověření dodržování standardu skrz tzv. *service loudness*, kdy se celodenní audio záznam vysílacího proudu rozdělí do bloků po hodinách, z nichž se vyberou ke zprůměrování pouze ty nejhlasilnější v rozsahu ± 1 LU. Tím se předchází zneužívání dlouhodobé integrace, kdy by v lukrativních vysílacích časech byla hlasitost záměrně vyšší a po zbytek dne naopak nižší. (40)



Obrázek 4.2: Možnosti normalizace velmi dynamického obsahu. Vlevo bezzásahová strategie, vpravo strategie preferující jednotnou hlasitost nebo dynamický rozsah.

4.2 Doporučené hodnoty dle EBU R 128 pro vysílání

Norma R 128 používá pojem *programme*, jemuž v českém jazyce přesněji odpovídá označení *pořad*, kterého se nadále budu držet. Pořad definuje jako: „Individuální, samostatný audiovizuální nebo pouze zvukový materiál určený k prezentaci v rozhlasu, televizi nebo jiných elektronických médiích.“ (17) Mezi pořady tak podle této definice patří například i reklamy nebo upoutávky. (1)

EBU pro všechny pořady zavádí jednotný hlasitostní standard (cílovou hodnotu; target level) jako (1) (13):

$$-23 \pm 0,5 \text{ LUFS}$$

a maximální skutečnou špičku signálu jako:

$$-1 \pm 0,3 \text{ dBTP}$$

V případě velmi dynamického obsahu norma připouští i nižší hodnoty s podmínkou, že je tato skutečnost dostatečně označena a je zabráněno nechtěné kompenzaci hlasitosti a z toho plynoucího zkreslení špiček. Pro živě vysílaný obsah, kde není možné předpovědět hlasitost pořadu, je tolerance kolem cílové hodnoty ± 1 LU. (1)

Při použití EBU mode (více na str. 19 a 41) se $0 \text{ LU} = -23 \text{ LUFS}$.

Z důvodu nespolehlivosti měření integrované hlasitosti na krátkých časových úsecích, doporučuje EBU pro posuzování krátkých pořadů (například upoutávek a reklam) používat i parametr maximální krátkodobé hlasitosti s konkrétní hodnotou -18 LUFS (5 LU na relativní stupnici). (25)

Normu přijali někteří regulátoři terestriálního vysílání v Evropě a je tak závazná například i pro všechny digitálně vysílající rozhlasové a televizní stanice v ČR. (26)

4.3 Normalizace hlasitosti pomocí metadat

Normalizace hlasitosti může být provedena dvěma základními způsoby. Buď přímo během vzniku audio materiálu, nebo pomocí metadat o hlasitosti. Audio (nebo audiovizuální) soubor tak s sebou v takovém případě nese informaci o své hlasitosti, podle které se nastaví následný řetězec tak, aby byla dodržena příslušná norma. (13)

Metadata jsou výhodná pro archivní obsah (například velkých médií), jehož převod na jednotnou hlasitost by vyžadoval velké úsilí a prostředky. Další výhodou může být nezávislost na cílové hodnotě hlasitosti a možnost nabídnout uživateli více módů dle poslechových podmínek (například Spotify, kapitola 5.1.2).

Ve specifikacích BWF metadat je tento parametr označen jako *LoudnessValue*, v systému metadat Dolby AC-3 jde o *dialnorm* (původně pro normalizaci podle dialogů; více v kapitole 4.5.1). (2) (27)

Pro případy záměrně nižší hlasitosti (například slabší věta symfonie) EBU doporučuje použití parametrů „loudness offset“ nebo „low loudness flag“, aby nedošlo k nechtěné normalizaci. (13) Parametr „album normalisation“ (více v kapitole 5.1.1) by i ve vysílání dával smysl.

Metadata byla autoritami doporučena i streamingovým službám, a to ideálně tak, aby byla zaslána až do koncového zařízení společně s audio/video streamem, kde se výsledný posun určí až po sečtení výstupní úrovně aplikace a posunu v důsledku normalizace, čímž se předchází například dvojité ztrátové kompresi. (14) (15)

4.4 Film

Filmový zvuk obecně patří k nejdynamičtějšimu zvukovému materiálu. Hlasitost je v současnosti upravena stejnou kalibrací kinosálů a postprodukčních pracovišť (například systém Dolby se stupni podle dynamického rozsahu poslechu), takže na rozdíl od

jiných žánrů mají tvůrci přesnější představu, jak hlasitě bude jejich dílo znít. Nicméně při distribuci filmů do domácího prostředí nastává problém s velkým dynamickým rozsahem, který nemusí vyhovovat domácímu prostředí. (13) (28)

Hlasitost dialogů, tedy toho nejpodstatnějšího, co by divák z filmu měl slyšet, může být řádově nižší než integrovaná hlasitost celého filmu vlivem dalších zvukových prvků filmu/pořadu. Z toho důvodu se v některých částech filmového a televizního prostředí používá i normalizace hlasitosti podle dialogů, která předchází nepoměrům v hlasitosti obdobných prvků mezi pořady a je vhodnější pro pořady s vyšším dynamickým rozsahem obsahující dialogy, které jsou referencí. Provádí se označením určitého úseku jako „*anchor element*“, většinou středně hlasitého mluveného slova, podle kterého se nastaví hlasitost celého pořadu. Druhou možností je upravený algoritmus BS.1770 s gatingem podle dialogů (kapitola 3.1.2). (2) (13)

4.5 Mimoevropské vysílací normy

ITU-R doporučuje pro mezinárodní výměnu rozhlasových a televizních pořadů používat -24 LUFS, TP zatím nespecifikuje. (29)

4.5.1 ATSC A/85 v USA

Odlišný přístup než EBU zvolil americký regulátor televizního vysílání ATSC v normě A/85. Cílovou hlasitost celého vysílání upravuje na:

$$-24 \pm 2 \text{ LUFS}^7$$

a maximálně -2 dBTP. (2)

System ponechává velkou část rozhodnutí na zvukových mistrech a hlasitost jednotlivých pořadů se upravuje pomocí parametru dialnorm. Dialnorm (dialog normalisation) je součástí metadat systému Dolby AC-3, který udává hlasitost dialogů pořadu. Jeho hodnota může být zjištěna automatickým měřením podle dialogů, ale doporučené je měření pomocí „*anchor element*“, tedy „kotvy“, kdy se mluvené slovo prohlásí za hlasitostní referenci, ke které se zbytek pořadu vztahuje. Oproti jiným žánrům to v televizním zvuku dává smysl, protože pokud akční film s množstvím hlasitých efektů a střelby porovnáme s přírodovědeckým dokumentem pouze na základě integrované hlasitosti

⁷ A/85 používá označení LKFS; LU a LUFS je označení používané v EBU, znamenající to samé a v rámci této práce preferované.

celého pořadu, tak hlasitost jednotlivých prvků nemusí odpovídat tomu, co by očekával divák. Tato „kotva“ se změří samostatně a podle této hodnoty se jednotlivé pořady tvořící vysílací proud srovnávají k sobě. Pořady, které neobsahují dialogy se změří podle BS.1770 celé a výsledek se prohlásí přímo za dialnorm. Naopak krátké pořady (jako například reklamy) podle „kotvy“ měřeny být nesmí a je nutné použít BS.1770. (2)

Tím, že dialnorm nemusí odpovídat hlasitosti pořadu, může celý proud kolísat, a proto mají televizní stanice relativně velkou toleranci kolem standardu (± 2 LU), kterou by neměly zneužívat a kterou norma prohlašuje za komfortní zónu pro posluchače. (2)

Použitý kodek AC-3 umožňuje posílat metadata na konec řetězce, kde se teprve provádí normalizace a kde je technicky možné poskytnout uživateli více módů normalizace v koncovém zařízení. (2)

5 Normalizace hlasitosti v internetových audio a audiovizuálních médiích

Pod pojem internetové audio a audiovizuální média jsou v této práci zahrnuty všechny internetové platformy, které poskytují videa nebo zvuk online. Například streamingové služby, video portály, hudební a podcastové aplikace, sociální sítě, zpravodajské portály – jedná se tedy o rozmanitý rozsah různých zdrojů, kde samotný výčet možností je problematický. Tyto zdroje se sbíhají v koncovém přehrávacím zařízení posluchače, který si „na vyžádání“⁸ reguluje obsah.

Mezi poskytovateli či médii, ale i v rámci jednoho poskytovatele může docházet ke skokům v hlasitosti, které jsou v tom lepším případě pouze obtěžující, v tom horším může být prudký skok v hlasitosti až bolestivý (například se sluchátky v uších).

Dalším specifikem internetových médií je prolínání mezi profesionálním a uživatelským obsahem. Například na platformě YouTube nebo Facebook se tak může objevit profesionálně vyrobený stream koncertu vedle videa natočeného na mobilní telefon.

Jelikož norma R 128 je závazná jen pro terestriální vysílání, tak internetové platformy buď normu nemají vůbec, nebo si systém kolem jednotek LUFS implementovali po svém a samy si zvolily svůj hlasitostní standard.

V následujícím výčtu se pokusím prozkoumat normy vybraných internetových médií. Cílem není vytvářet přehled technických požadavků jednotlivých služeb vůči tvůrcům, ale spíše tyto požadavky, jednotlivé hodnoty a jejich čtení zasadit do kontextu hlasitostní normalizace. Informace jsou aktuální v době vzniku této práce na jaře 2021, ale vzhledem k rychlému vývoji na internetu nemusí být platné už na podzim téhož roku. Budoucího čtenáře-tvůrce tak upozorňuji, aby na možnou neaktuálnost norem jednotlivých platform bral zřetel a konkrétní hodnoty si případně pro účely své práce ověřil.

⁸ Z anglického „*on demand*“; například SVOD (subscription video on demand, příkladem může být služba Netflix), kdy má uživatel za pravidelný poplatek neomezený přístup k rozsáhlé mediální knihovně a sám si „sestavuje program“. (42) Stejným způsobem obohacují své služby i „tradiční“ vysílací platformy (například iVysílání České televize, či mujrozhlas.cz Českého rozhlasu).

5.1 Platformy pro audio streaming

5.1.1 Normalizace dle skladby versus alba

Při zavádění automatizované normalizace nastává problém s kontextem, ve kterém se případně skladby streamují. V rámci hudebního streamingu je možné přehrávání v nahodilém pořadí napříč všemi dostupnými skladbami (*shuffle*) nebo přehrávání celých alb. Skladby v albu mají většinou rozdílné hlasitosti a jejich poměry jsou tvůrčím záměrem. Při normalizaci jednotlivých skladeb uvnitř alba by se tyto poměry narušily, což by mohlo poničit jeho koncept, kdy jsou zesilovány části, které mají být záměrně slabší a naopak. Proto se rozlišuje normalizace podle tracku a podle alba (*track/album normalisation*). (13) (15)

Pro normalizaci podle celého alba pak existují dvě možnosti provedení:

- Hlasitost je stanovena na základě integrované hlasitosti celého alba jakožto jednolitého programu (což odpovídá i definici dle EBU). Tuto možnost využívá například Spotify a Apple Music. (30)
- Hlasitost se stanoví podle nejhlasitější skladby alba, která se srovná na maximální povolenou hodnotu a všechny skladby své poměry odvodí podle ní. Tuto možnost využívá například služba Tidal. (15)

Poslední jmenovaná možnost se opírá o výzkum Eelco Grimma, který stojí za doporučením nastavení normalizace pro audio streamingovou službu Tidal. Grimm analyzoval hlasitost celé dostupné hudební knihovny (asi 4,2 milionu alb) a doporučil právě normalizaci alba podle nejsilnější skladby. Poté zkoumal, zda uživatelé preferují při „shuffle módu“, tedy nahodilém přehrávání skladeb mezi alby, normalizaci podle tracku nebo podle alba, a 80 % uživatelů preferovalo zachování normalizace podle alba i mimo jeho kontext. Myšlenka normalizovat alba podle nejhlasitější skladby má silnou oporu v přirozené vlastnosti většiny alb snažit se využít dynamický rozsah. Srovnáním toho, co alba považují za „stav nahlas“, se ukotví hlasitost i ostatním skladbám a skladby, které jsou záměrně slabší pak nejsou zbytečně kompenzovány. Zároveň se tím snižuje integrovaná hlasitost celé služby, protože standardu dosahují pouze nejsilnější skladby, což motivuje tvůrce k zachování headroomu. (15)

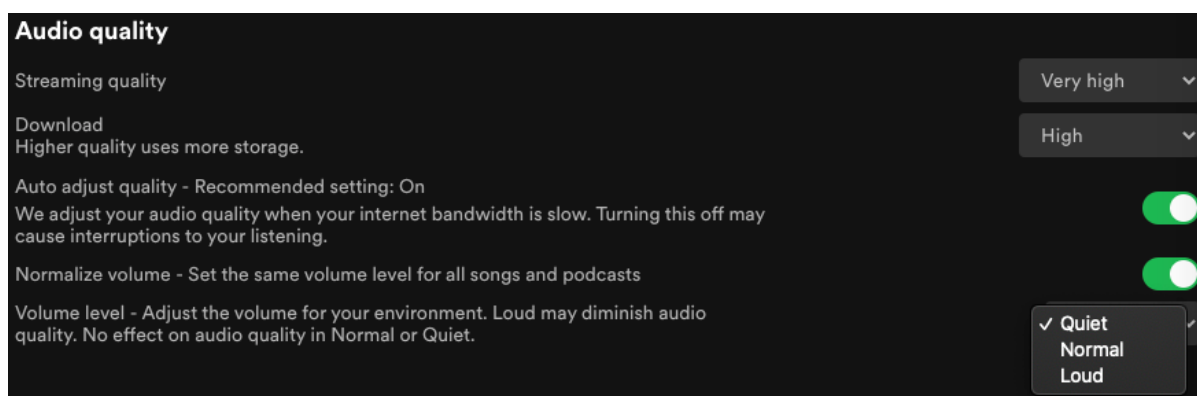
5.1.2 Spotify

Služba Spotify je nejpoužívanější placenou hudební streamingovou platformou na světě a po YouTube je druhou nejpoužívanější internetovou hudební platformou. (31)

Od konce roku 2020 služba přestala využívat pro normalizaci hlasitosti systém ReplayGain a přešla na systém podle ITU-R BS.1770. (32) Ve výchozím stavu pak normalizuje na -14 LUFS podle alb a kladný gain aplikuje pouze pokud to headroom dovolí. Skladby tedy dynamicky neupravuje (na obrázku 4.1 možnost vlevo). (30)

Po tvůrcích požaduje pouze maximální hodnoty integrované hlasitosti -14 LUFS a maximálně -1 dBTP. Pro mastery hlasitější než -14 LUFS doporučuje maximálně -2 dBTP. (30)

Pro prémiové uživatele nabízí v nastavení aplikace 3 módy normalizace dle poslecho- vých podmínek (obrázek 5.1). „**Normální**“ (Normal) je výše popsané výchozí nastave- ní, „**Potichu**“ (Quiet), kdy normalizuje na -23 LUFS (také bez dynamických zásahů) a „**Hlasité**“ (Loud), kde je audio automaticky normalizováno „za každou cenu“ na cílo- vou hodnotu -11 LUFS a vše s nižší hlasitostí je kompenzováno pozitivním gainem. Do cesty je v takovém případě zařazen limiter a veškerý obsah s větším headroomem je poznamenán limitací. Nastavení limiteru je -1 dB threshold, 5 ms attack a 100 ms release, což vysoce dynamickému obsahu (například klasické hudbě) může slyšitelně uškodit. Limiter reaguje pouze na sample peak, nikoliv true peak, takže může docházet i k přebuzení výstupu. Volba tohoto režimu je ale zcela na uživateli. (30)



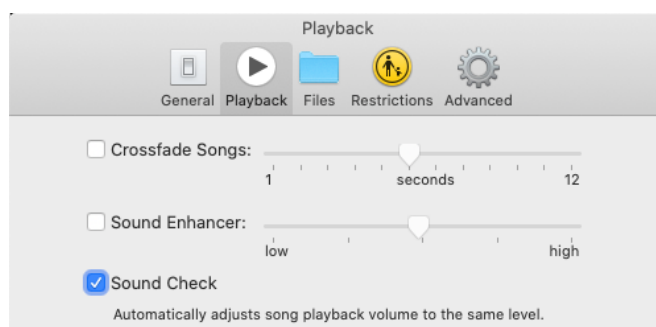
Obrázek 5.1: Nastavení přehrávání v aplikaci Spotify.

Při vypnutí normalizace hlasitosti se skladby přehrávají se svou původní hlasitostí a to i bez kontroly maximálního TP, což potvrzuje, že služba Spotify využívá pro

normalizaci metadata. Webové rozhraní a aplikace třetích stran, přes které služba Spotify může fungovat, normalizaci hlasitosti zatím nepodporují. (30)

5.1.3 Apple Hudba a Podcasty (bývalé iTunes)

Firma Apple využívá k normalizaci hlasitosti ve svých aplikacích svůj proprietární nástroj Sound check, který je ale ve výchozím stavu vypnutý. Velmi pravděpodobně využívá algoritmus dle ITU nebo jemu příbuzný, ale bohužel (jako o mnohých technických podrobnostech firmy Apple) se o něm není možné dozvědět více. V aktuálních dokumentech, které jsou určené tvůrcům obsahu doporučuje pro mastery hlasitost $-16 \text{ LUFS} \pm 1 \text{ LU}$ a maximálně -1 dBTP (33) (34). Jiné zdroje, jako například Katz, normalizaci v prostředí bývalého iTunes přisuzují hodnotu kolem $-16,5 \text{ LUFS}$. (7) Sound check normalizuje skladby dle alba jako celku. (33)



Obrázek 5.2: Jediné existující nastavení funkce Sound check v nastavení aplikace Apple Music.

Pomocí metody, jejíž princip je detailněji popsán v příloze 2, bylo provedeno měření výstupu aplikace Apple Music, kterým byl používaný hlasitostní standard ověřen. Pro ověření byla záměrně vybrána alba, která svou masteringovou estetikou ilustrují válku hlasitostí. Jednalo se o kompilační album „*Number One*“ od Michaela Jacksona z roku 2003 a album „*A Different Kind of Truth*“ od skupiny Van Halen z roku 2012. Jelikož bez normalizace dosahují hodnot nad -10 LUFS s pozitivními hodnotami TP, při normalizaci jsou tato alba pravděpodobně ve všech službách ztlumena na nejvyšší povolenou hlasitost konkrétní služby.

Se zapnutou funkcí Sound check obě tato alba shodně dosáhla $-16,5 \text{ LUFS}$ při měření podle metodiky popsané v příloze 2. Samostatné singly podobně masterovaných nahrávek (například *Living on a Prayer* od Rickiho Martina) dosáhly hodnot maximálně $-16,2 \text{ LUFS}$.

Pozitivní gain, podobně jako u služby Spotify, aplikuje Sound check pouze tehdy, když je TP snímku nižší než maximální povolený TP (-1 dBTP) a nejvýše do velikosti tohoto rozdílu, což bylo opět ověřeno metodou dle přílohy 2.

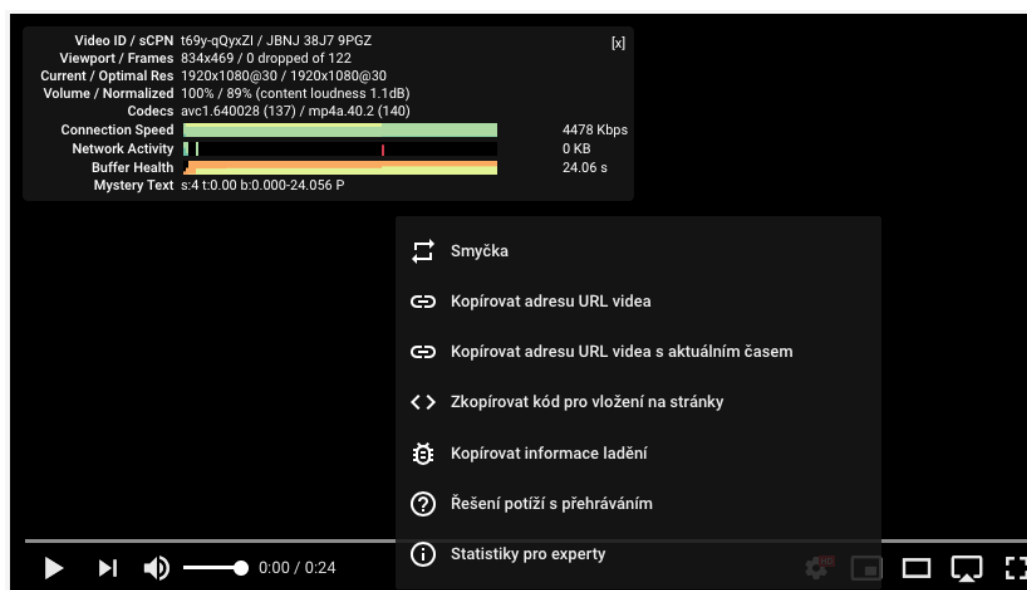
Rozsah mezi -16,5 až -16 LUFS lze prohlásit za target level funkce Sound check, kterého ale není dosahováno za cenu dynamických úprav, clippingu, či překročení -1 dBTP a nahrávkám není snižována kvalita v důsledku normalizace.

5.2 Video portály a video streaming

5.2.1 YouTube

Server YouTube je nejpopulárnější video platformou (po serveru Google, pod který spadá, druhou nejpopulárnější internetovou platformou vůbec) (35) a zahrnuje tak komplexní škálu audio materiálu v rámci jedné platformy. Na YouTube se setkává profesionálně vytvořený obsah s uživatelským/amatérským, lze tam vysílat přímé přenosy a skrz platformu YouTube Music funguje i jako hudební streamingová služba.

YouTube, respektive Google, aktuálně detailní informace o hlasitostní normalizaci ne nabízí. Ve videích je ale možné zobrazit *Statistiky pro experty / Stats for nerds* (obrázek 5.3), kde je údaj *Content loudness* (hlasitost obsahu). Tento údaj potvrzuje domněnky, že YouTube používá systém ITU-R BS.1770 s maximální povolenou hodnotou integrované hlasitosti -14 LUFS s použitím relativní stupnice vůči standardu (content



Obrázek 5.3: Umístění údaje content loudness v rámci YouTube přehrávače. Hodnota 1,1 dB (LU) znamená, že audio má integrovanou hlasitost -12,9 LUFS.

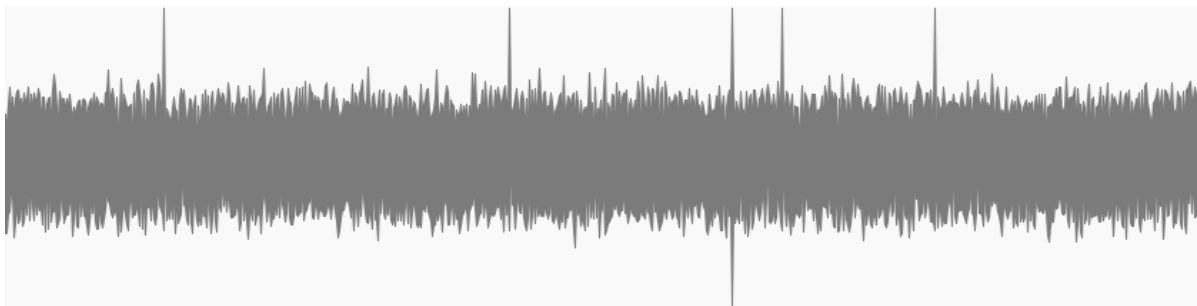
loudness 0 dB = 0 LU = -14 LUFS). V případě překročení nulové hodnoty je aplikován gain s opačnou hodnotou. Oproti službě Spotify ale YouTube neaplikuje kladný gain ani nenabízí jakékoliv nastavení normalizace. Jelikož YouTube neposkytuje své přesné specifikace, tak byly tyto údaje ověřeny pomocí uploadu testovacích video souborů o délce 24 s. Jednalo se o různové šumy a některé ukázky byly záměrně upraveny tak, aby v nich byl clipping s TP vyšším než 0 dBTP. Ukázky byly vždy renderovány do formátu *Quicktime MOV* s audiem ponechaným ve formátu *WAV*. Pro přesnost jsou v tabulce 2 uvedeny hodnoty měřené přímo z renderovaných souborů před uploadem.

	Původní soubor		YouTube metriky	Měření výstupu	
	L _K (LUFS)	TP (dBTP)	Content / L _K (LU/(LUFS))	L _K (LUFS)	TP (dBTP)
Ukázka 1	-14,4	-4,1	-0,4 (-14,4)	-14,3	-4,0
Ukázka 2	-25,4	-15,3	-11,4 (-25,4)	-25,3	-15,3
Ukázka 3	-12,9	-2,7	1,1 (-12,9)	-14,0	-3,6
Ukázka 4	-14,4	5,1	-0,4 (-14,4)	-14,5	0,6
Ukázka 5	-14,4	7,4	-0,4 (-14,4)	-14,5	1,0
Ukázka 6	-14	1	0 (-14)	-14,4	0,8

Tabulka 2: Výsledky měření normalizace hlasitosti na YouTube.

Naměřené hodnoty potvrzují, že YouTube normalizuje na -14 LUFS a integrovaná hlasitost odchozích souborů přesně souhlasí s relativní metrikou YouTube. Rozdíl mezi těmito hodnotami a hodnotami naměřenými na výstupu pomocí metodiky popsané v příloze 2 je dán odchylkou měřicího řetězce (vnitřní audio sběrnice, internetový prohlížeč, konkrétní měřič hlasitosti...).

Překvapením jsou hodnoty true peak. Ukázky 4, 5 a 6 vycházely z ukázky 1, ale byly záměrně upraveny tak, aby překračovaly hodnoty 0 dBTP ať už přímo s clippingem (4 a 5) nebo se špičkou těsně pod 0 dBFS pomocí limitace (ukázka 6). YouTube ale tyto ukázky normalizoval stejně jako ukázku 1, pravděpodobně pouze na základě integrované hlasitosti a soubory tak při přehrávání vykazovaly kladné hodnoty TP. Špičky dosahující hodnot 0 dBFS byly pouze potlačeny kodekem AAC (na straně YouTube), který je chráněn proti nechtěnému přebuzení při enkódování a dekodování pomocí limitace (7), která byla slyšitelná, ale která TP nesnížila pod 0 dBTP.



Obrázek 5.4: Waveforma ukázky 6 se záměrně vytvořenými impulzy s peak hodnotou -0.1 dBFS, ale true peak hodnotou $+1$ dBTP.

YouTube nepřipouští při playbacku vyšší hodnoty integrované hlasitosti než -14 LUFS. Maximální hodnotu TP pravděpodobně nevyužívá a nechává to na odpovědnosti tvůrců obsahu. Jelikož je normalizované každé video zvlášť, tak v případě obsahu, který na sebe navazuje, může při překročení -14 LUFS u jednoho z videí dojít k zásahu do původně zamýšlených poměrů – YouTube tedy nedisponuje normalizací podle alba nebo nějakou její obdobou, ale vzhledem k tomu, že videa nikdy nejsou zesilována, tak pokud žádné z videí z jedné série nepřekročí -14 LUFS, bude obsah přehráván dle záměrů tvůrce.

5.2.2 Streaming filmů a seriálů

Podobně jako klasická hudba, tak i filmy a seriály při streamingu přes internet naráží na rozdíl oproti relativně vysokým hodnotám integrované hlasitosti, na které normalizují okolní platformy. V kapitole 4.5.1 byl popsán princip normalizace v americkém televizním prostředí a podobně jako tam, se při distribuci filmů využívá normalizace pomocí dialogu nebo „ukotvení“.

Například streamingová služba Netflix od tvůrců vyžaduje -27 ± 2 LKFS *dialogue gated*. Je potřeba upozornit, že *LKFS dialogue gated* není to samé jako LKFS/LUFS,

a proto zde bylo mimořádně využito označení jednotky dle ITU. (Příkladem tohoto rozdílu může být obrázek 3.5.) Hodnoty hlasitosti váhované podle mluveného slova a integrované hlasitosti jsou si nejbližší, pokud je mluvené slovo jediným prvkem pořadu. Proto nelze hlasitostní standardy Netflixu a podobných služeb jednoduše porovnávat s hudebním a vysílacím světem. (36)

Tvůrci seriálů a filmů, ale i tvůrci audio/audiovizuálních děl klasické hudby jsou v nevýhodě při střetu s praxí okolo -14 LUFS (například YouTube a Spotify), ale adaptovat se na tuto hodnotu pro ně může být problém kvůli nedostatečnému dynamickému rozsahu.

5.3 Rozhlasové a televizní stanice distribuované přes internet

Webové portály rozhlasových a televizních stanic představují pro distributory výzvu (zejména pro ty, kteří plný dynamický rozsah využívají), protože se jedná o příklad střetu prostředí o různých hlasitostních standardech. Služby, které postupují v duchu normy R 128 se při internetovém využití audio materiálu původně určeného pro vysílání mohou ocitnout v nevýhodě a například vůči YouTube či Spotify být o 9 dB slabší.

EBU v dodatku k normě R 128 (1) doporučuje, aby materiál původně vyrobený pro vysílání, ale poskytovaný přes internet nebo přímo pro něj vyráběný, byl i nadále zpracováván v souladu s R 128 a Tech 3343, tedy při -23 LUFS. Pro účely internetového streamingu doporučuje normalizovat na -18 LUFS a v případě potřeby provádět dynamické úpravy buď na straně poskytovatele, nebo pomocí metadat v koncovém zařízení posluchače. Při normalizaci v koncovém zařízení je výhodou případné uživatelské nastavení podle podmínek poslechu. (37)

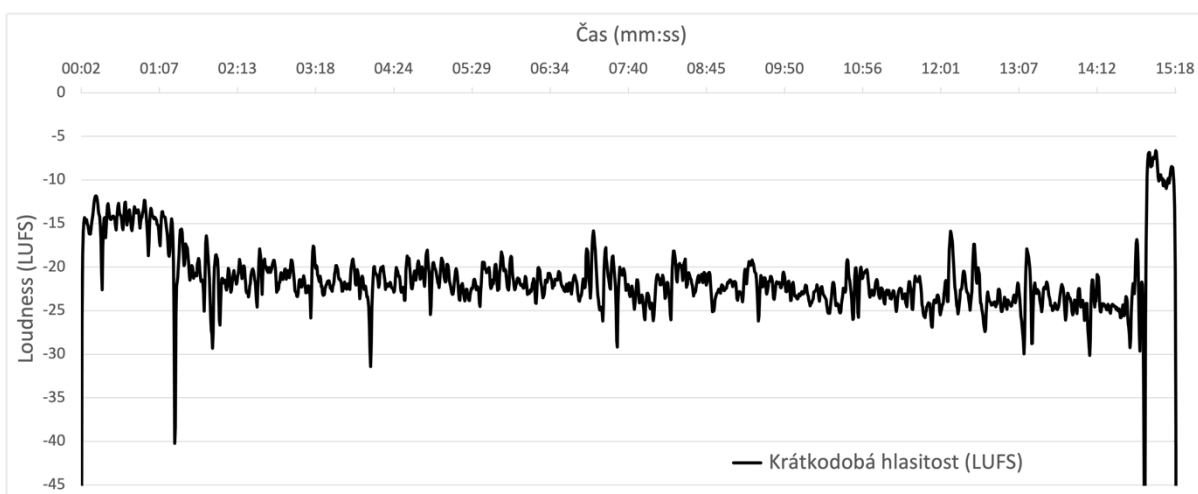
5.4 Webové portály bez normalizace

Pomocí stejných (a případně dalších upravených) testovacích souborů jako při testování služby YouTube (kapitola 5.2.1) byly testovány i další služby, nicméně ani náznak normalizace hlasitosti u nich nalezen nebyl a audio soubory v rozsahu -26 – (-7) LUFS po uploadu zůstaly v rozsahu $\pm 0,5$ LU od své původní hodnoty.

Sociální síť **Facebook** audio nenormalizuje a je zcela bez ochrany proti skrytým špičkám. Stejně tak síť **Twitter** a **Instagram**.

Video portál **Vimeo** je určen pro sdílení video obsahu ve vysoké kvalitě. Do audio obsahu nijak nezasahuje (kromě ztrátové komprese) a normalizaci hlasitosti ani kontrolu skutečné špičky taktéž nepodporuje.

Pro menší webové platformy, kde je multimediální obsah pouze doplňkem, není normalizace hlasitosti bohužel prioritou. Typickým příkladem jsou zpravodajské platformy, jako například **iDnes** (idnes.cz) nebo **Aktuálně** (aktualne.cz). Problém ovšem nastává, když jsou videa ohraničená samostatnými reklamními příspěvky, protože mezi obsahem a reklamou může docházet k výrazným skokům v hlasitosti. Jeden takový příklad je na obrázku 5.5.



Obrázek 5.5: Graf průběhu krátkodobé hlasitosti v čase u zpravodajského pořadu DVTV v rámci přehrávače na portálu aktualne.cz. Dva reklamní příspěvky před začátkem a jeden extrémně hlasitý po skončení.

6 Důsledky pro práci zvukového mistra

V této kapitole bych se rád zaměřil na to, co normalizace znamená pro tvůrce obsahu. Pro další informace, užitečné nejen v prostředí EBU, doporučuji dokument *Tech 3343*. (13) Pro účely zjišťování konkrétních norem (obzvláště ve světě hudebního zvuku) slouží pluginy jako například *Loudness penalty*, který dokáže přibližně napovědět, jakou hlasitost bude mít v konkrétních službách zpracovávaný audio materiál. (38)

6.1 Omezení hlasitostního algoritmu

Algoritmus pro výpočet L_K je ve své podstatě jen řetězec skládající se z filtru, rozdělení do bloků, výpočtu výkonu, selekce bloků (gating) a následného součtu se zohledněním váhy jednotlivých kanálů. Vstupnímu signálu přiřadí číslo. Ačkoliv je relativně přesný, tak nemůže nahradit úsudek člověka a lidské ucho zatím stále zůstává nejpřesnějším měřičem hlasitosti.

To, proč se zatím nepoužívají komplexnější a teoreticky i přesnější algoritmy než ITU-R BS.1770, je v jeho stávající účinnosti vzhledem k výsledkům a nízkým nárokům na výpočetní výkon. Při výpočtu skutečné špičky sice probíhá převzorkování, ale pro měřící účely stačí nejjednodušší metody.

Integrovaná hlasitost je méně přesná pro velké dynamické rozsahy, protože vlivem relativního gatingu (práh -10 LU pod integrovanou hlasitostí) mohou být nejnižší pásáže vyřazeny z měření, ačkoliv jsou plnohodnotnou součástí programu.

Výpočet je relevantní pro komplexní signály, při měření čistých (sinusových) tónů je zavádějící. (3) To je třeba mít na paměti při testování úrovní, kde je sinusový signál o frekvenci 1 kHz zažitou praxí při předávání signálu. Algoritmus byl navržen tak, aby pro sinusový signál o frekvenci 1 kHz ve stereu zobrazoval stejnou hodnotu jako na peak metrech (například -18 dBFS odpovídá -18 LUFS). Signál o stejné úrovni pouze v jednom z L/C/R kanálů zobrazí o 3 LU méně v důsledku váhování, což ale odpovídá i vjemu hlasitosti. Vlivem K-filtru (obrázek 3.2 na straně 12) ale tato skutečnost neplatí pro všechny sinusové frekvence stejně. (13) (23)

6.2 Normalizace z pohledu tvůrce

Smyslem celého popisovaného systému je vrátit řád odvětví, které se vymklo kontrole. Zejména hlasitostní válka v hudebním průmyslu bude i nadále odstrašujícím memem.

V prostředích s fixní hodnotou hlasitosti (jako například EBU) existují dvě cesty, jak vyhovět cílové hlasitosti. Tou první je zachovat stávající pracovní zvyklosti („workflow“) a normalizaci hlasitosti provést až jako poslední krok celého procesu. V případě audio obsahu s nízkým dynamickým rozsahem (zpravodajství, populární hudba) je tento krok většinou bez problému. Tento postup je výrazně jednodušší při tvorbě postprodukovaného obsahu než v případě živého vysílání. (13)

Druhou možností je plně se adaptovat na hlasitostní měřiče a nad hlasitostí tak mít kontrolu od začátku do konce procesu. Ukotvení cílové hladiny hlasitosti neznámá, že všechny hlasitostní měřiče budou mít přesně tuto hodnotu a výsledek naší práce bude mít konstantní hlasitost – naopak; společně s kalibrací poslechu nás vybízí k posuzování hlasitosti a dynamiky na základě skutečného vjemu. (7) (13)

Jednoduchou a poměrně účinnou metodou částečné adaptace v prostředí, kde již existuje referenční poslechová hladina, může být přeměření již vyhotovených snímků/přehledů a upravení poslechové hlasitosti tak, aby snímky následující byly přirozeně adaptovány na požadovanou normu. (Pokud například většina vyprodukovaných snímků měla hlasitost přibližně -21 LUFS, pro potřeby EBU stačí zesílit poslech o 2 dB). (13)

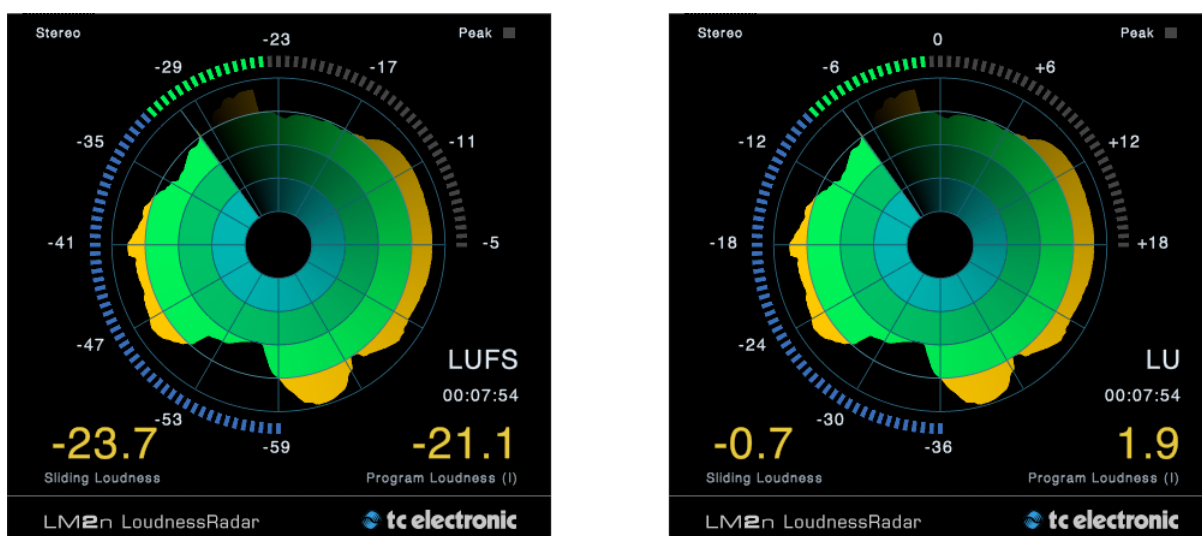
6.2.1 Čtení hlasitostních měřičů

Jak již bylo zmíněno, tak měření (ani normalizování) podle špičky digitálního signálu neříká nic o hlasitosti a peak metr by měl sloužit primárně k nastavení technických parametrů média (zejména jako ochrana před přebuzením nebo extrémním podmodulováním jednotlivých stop) a je nejdůležitější při záznamu zvuku. Z tohoto důvodu většina hlasitostních měřičů (které slouží primárně pro mix a mastering), nabízí pouze jediný „technický“ parametr, a to maximální skutečnou špičku (*max true peak*).

Pro posuzování aktuální hlasitosti slouží dvě stupnice bez gatingu, a to okamžitá (*M; momentary*) a krátkodobá (*S; short-term*). Záleží na kontextu práce, ale i na preferencích zvukového mistra podle které stupnice se mu lépe pracuje. **M** poskytuje informaci o okamžitém stavu a **S** více o vnímané hlasitosti a trendu. Stupnice **I** (*integrated/programme loudness*) je relevantní až v delším časovém horizontu a při

postprodukci může být dokonce zavádějící (příkladem může být opakované přehrávání jednoho místa z celku a podobně). Při kalibraci poslechu slouží spíše k ověření, že mix nebo probíhající přenos odpovídají požadavkům. (7) (13)

Další užitečnou volbou hlasitostních měřičů, která většinou není ve výchozím stavu zapnutá, je relativní stupnice. EBU mode nabízí dvě relativní stupnice *EBU +9 scale* a *EBU +18 scale*. 0 LU pak znamená „target“ (-23 LUFS v EBU) a příslušné číslo označuje nejvyšší zobrazitelnou hodnotu nad 0 LU. Volba rozsahu závisí na očekávaném dynamickém rozsahu. (7) (13) (23) Tento režim je výhodný v okamžitém vyjádření o kolik je mix potřeba kompenzovat. Pokud například na konci zkoušky před vysíláním je integrovaná hlasitost 1,9 LU, je zřejmé, že pokud očekáváme během samotného vysílání podobnou hlasitost, je třeba celý mix přibližně o tolik dB zeslabit. Další výhodou relativního nastavení může být rychlost čtení. Při živém vysílání musí být měřič čitelný i při velmi letmém pohledu, příklad tohoto rozdílu je na obrázku 6.1. V případě metrů s vertikálními stupnicemi je rozdíl ještě markantnější.



Obrázek 6.1: Porovnání zobrazení stejné situace s rozdílem v použitých stupnicích u pluginu TC Electronic LM2n. Vlevo absolutní stupnice s LUFS a vpravo relativní s LU.

6.2.2 Kalibrace poslechu

Ovládání hlasitosti poslechu by mělo být také relativní a ideálně s kalibrační hodnotou označenou jako „0“. Na myšlence kalibrovaného poslechu byl postaven už *K-system* Boba Katze, který byl jedním z prvních pokusů o normalizaci hlasitosti a zastavení války hlasitostí na straně tvůrců. Myšlenkou kalibrace se inspiroval z filmové praxe. (39) Konkrétní kalibrační hodnoty, na které by měla být režie kalibrována, závisí na velikosti místnosti a zvolené metodice, pracovním prostředí a produkovaném žánru. (2) (7) (13) (39) EBU svoji doporučenou metodiku popisuje v dokumentu Tech 3343 v kapitole 8 (13); ATSC přímo v A/85 v kapitole 10 (2).

Diskuze

Systém algoritmů a norem pro normalizaci hlasitosti vyvinutý ITU a EBU je navržen velmi robustně, nicméně plnohodnotně stále ještě implementován není. Jeho největší slabinou je zatím stále nejednotný standard integrované hlasitosti. Pro ilustraci nabízím tabulku 3, kde jsou shrnuty standardy služeb, které normalizují (aktuální na jaře 2021):

Poskytovatel či systém norem	Cílová hodnota L_K	Maximální hodnota L_K	Další specifikace
EBU (R 128)	-23 LUFS	–	Max short-term -18 LUFS
ATSC (A/85)	-24 LUFS	–	Volnější přístup
YouTube	–	-14 LUFS	
Spotify	-11 LUFS pouze v režimu loud	-14 LUFS / -23 LUFS	3 módy podle podmínek; normalizace dle alba
Apple Music		-16,5 LUFS ⁹	Nutné zapnout v nastavení; norm. dle alba
Tidal	-14 LUFS ¹⁰	–	
Netflix	-27 LKFS dial. ¹¹	–	
AES TD1004	-20–(-16) LUFS	-16 LUFS	(doporučení pro internet)

Tabulka 3: Stručný souhrn hlasitostních standardů různých služeb. Podrobněji v kapitolách 4 a 5.

I z uvedených hodnot je patrná obdoba konfliktu mezi dynamickým rozsahem posluchových podmínek, který je rozebrán na straně 21.

Audio Engineering Society vydala dokument *TD1004*, ve kterém doporučuje internetovým platformám, aby nenastavovaly maximální hodnotu integrované hlasitosti

⁹ Přibližná hodnota.

¹⁰ Hodnota integrované hlasitosti nejsilnější skladby alba.

¹¹ Algoritmus s gatingem dle dialogů, více v kapitole 3.1.2 na straně 10.

výše než na -16 LUFS. S ohledem na mobilní zařízení naopak nedoporučuje volit nižší hodnoty než -20 LUFS pro jejich případně nedostatečný výstupní výkon. (14)

Právě mobilní zařízení zatím značně limitují snahy o zavedení jednotného standardu. V roce 2013 byla přijata norma *CENELEC EN 50332*, která zavazuje výrobce k limitaci maximálního výstupního výkonu přenosných přehrávačů (v současnosti převážně mobilních telefonů a tabletů) z důvodu ochrany sluchu uživatelů. Tento limit byl bohužel definován absolutně vůči testovacím signálům, a tak audio materiál o hlasitosti nižší než -23 LUFS nemusí být při nejvyšším zesílení slyšet nebo je uživatel upozorňován na příliš vysokou hlasitost v rozporu se skutečností. (14) Dle mého názoru, ovlivnila nastavování těchto hodnot pravděpodobně tehdy ještě probíhající hlasitostní válka. V roce 2017 byla norma upravena tak, aby uživatele upozorňovala s přihlédnutím ke skutečným energetickým vlastnostem signálu, nicméně zřejmě ještě nějakou dobu potrvá, než bude norma plně implementována. Z tohoto důvodu zatím sice AES nedoporučuje pro internetové platformy nižší hodnoty než -20 LUFS, ale zároveň připouští, že do budoucna bude vhodnější přejít na nižší standard. (14)

Dobrym příkladem směru vývoje jsou módy normalizace hlasitosti na základě poslechových podmínek uživatele ve službě Spotify (detailněji popsáno v kapitole 5.1.2). Ačkoli je hlasitý mód v současnosti nastaven poněkud nemilosrdně k dynamickému obsahu, tak je řešením pro hlučná prostředí. Pro uživatele, kteří poslouchají hudbu v autě nebo v letadle, je prioritou přenos informace a dodatečné dynamické úpravy jsou tak přijatelná cena, která ještě může být snížena zkvalitněním dynamických úprav v aplikaci například lepšími limity.

Volba takových poslechových podmínek ale nemůže omezovat ostatní uživatele (a zpětně i tvůrce) v dostupném dynamickém rozsahu. Pro výrobu a distribuci by měl být hlasitostní standard do budoucna nižší pro zachování dostatečného headroomu (a tím pádem i dostupného prostoru pro dynamiku). V extrémních poslechových podmínkách nebo v zařízeních s malým výkonem by pak případná automatická úprava signálu měla probíhat až na konci řetězce, a to nejlépe pomocí metadat. Problémy dynamického rozsahu a hlasitosti jsou bohužel stále směřovány do jednoho a pomocí hlasitosti se řeší problém přizpůsobování dynamického rozsahu.

Dalším významným prvkem jsou nastavení aplikací. Má-li být normalizace hlasitosti úspěšná, musí být zapnutá ve výchozím nastavení, což doporučuje i Grimm (15). Kolísání hlasitosti napříč skladbami/pořady není problém uživatelů,

z nichž velkou část pravděpodobně ani nenapadne nastavení otevřít, a nemůžeme po nich chtít, aby se v problému orientovali.

Dalším nepříjemným problémem je nejednotnost úrovní na straně techniky určené pro domácnosti, kde se setkávají různé zdroje, jako například domácí kina, televizory nebo počítače a kde se stávající rozdíly mezi hlasitostmi služeb můžou ještě prohloubit. Skupina PLOud proto vydala dokument *Tech 3344*, který doporučuje výrobcům těchto zařízení do budoucna přejít na jednotný standard a pro zpětnou kompatibilitu umožnit uživatelům specifikovat připojená zařízení, podle nichž se aspoň přibližně nastaví úrovně. (40) (41)

Na rozdíl od rozhlasového a televizního vysílání, které díky své technické povaze musí být regulováno, je internet často o krok napřed před legislativou, která ho mnohdy ani dosáhnout nedokáže. Kontroverzní téma regulace internetu je daleko za hranicemi této práce, každopádně v případě problematiky normalizace hlasitostí je těžké přesvědčit velké firmy, aby udělaly krok do neznáma a snížily své hlasitostní standardy ve prospěch dynamičtějšího obsahu. Neexistuje žádné globální ministerstvo nebo ředitelství internetu, které by dokázalo něco takového nařídit všem a dodržování vymáhat, a tak bohužel není zaručeno, že nepříjde nová masově úspěšnější služba, která normalizovat nebude a vtáhne tvůrce do nové hlasitostní války.

Závěr

V kapitole 3 byl detailně prozkoumán základní algoritmus pro výpočet integrované hlasitosti signálu dle ITU-R BS.1770 i s jeho rozšířeními a v kapitole 6 jsem se zabýval jeho limitacemi. Algoritmus nemůže nahradit úsudek člověka, ale podává uspokojivé výsledky umožňující relativně přesně mezi sebou porovnávat signály v přípustné toleranci. Zároveň není příliš náročný na výpočetní výkon, což (vzhledem ke vzrůstajícímu výkonu i těch nejmenších zařízení) nebrání jeho užití na úplném konci řetězce.

Dále jsem se v kapitolách 4 a 5 zabýval existujícími principy normalizace hlasitosti, a to jak ve vysílání, tak na internetu. Překvapením pro mě bylo, že je systém pracovaný do detailu a připravený k použití všude, kde o něj bude zájem. Autority jako EBU, ITU a AES problematiku reflektují a je zřejmá snaha o sjednocení a vývoj.

Ve vysílání se praxe normalizace hlasitosti osvědčila a zvláště v Evropě EBU nabízí mnoho pomocných dokumentů, které mohou usnadnit přechod na jednotné (a při dodržení všech doporučení i jednodušší) prostředí uvnitř vysílacích společností. Americké televizní prostředí řídicí se normou ATSC A/85 je příkladem odlišného přístupu, který nechává větší důvěru a zodpovědnost na tvůrcích a který může být inspirací pro řešení hlasitosti u internetových portálů orientujících se na filmy a seriály.

Největší internetové služby již sice hlasitost normalizují a brání tak úspěšně hlasitostní válce, ale jejich cílové/maximální hodnoty hlasitosti jsou stále relativně vysoké a omezují tak dynamičtější žánry. Některé služby jako například Apple Music nemají normalizaci hlasitosti ve výchozím nastavení zapnutou a nepoučený uživatel se tak o ní vůbec nemusí dozvědět. V případě, kdy služba neprovádí dynamické úpravy a nehrozí tak ztráta kvality, je normalizace hlasitosti spíše benefitem a není důvod ji vypínat ve výchozím stavu.

V kapitole 5 bylo zároveň provedeno několik měření, které prokázaly očekávané nastavení internetových platform. Překvapením byla absence kontroly skutečné špičky na serveru YouTube a absence jakékoliv normalizace hlasitosti na sociálních sítích Facebook, Instagram a Twitter.

Sjednocení zatím brání především přenosná zařízení, která mohou mít slabší výstupní výkon a hodnoty integrované hlasitosti pod -20 LUFS pro ně mohou představovat problém. Největší potíž ale vidím ve směřování problému hlasitosti a problému dynamického rozsahu. Tato práce se zabývala hlasitostí a dynamický rozsah řešila spíše

okrajově, ale rozhodně by na ni šlo navázat právě zkoumáním problematiky vysokých dynamických rozsahů a jejich případných úprav v koncových zařízeních pro podmínky se sníženým headroomem, či případně zkoumáním standardizace rozsahů nebo předávání metadat s LRA. Standardizace dynamického rozsahu do horších poslechových podmínek se stále řeší zvýšením hlasitostního standardu, místo snížení maximální špičky, které by zajistilo základní srovnatelnost napříč audio světem. Stále je důležité znovu opakovat, že snížit hlasitostní standard neznamena automaticky zvyšovat dynamický rozsah. Normalizace hlasitosti nijak neomezuje výrobu dynamicky jakkoliv upravených mixů.

Obtížněji řešitelný zůstává do budoucna problém rozdílů mezi médii, které normalizují podle integrované hlasitosti, a těmi, které využívají normalizaci podle dialogů (například Netflix).

V kapitole 6 jsem se zaměřil na důsledky pro praxi nás zvukových mistrů. Kapitola shrnuje strategie přechodu do „hlasitostně normalizovaného světa“ a snaží se nabídnout vhled do základů práce v něm, protože vzhledem k různým implementacím v různých platformách má normalizace hlasitosti různé podoby, používá různý jazyk a může být nejasná i samotným tvůrcům obsahu. K řešení problému mohou tvůrci přispět minimálně stejnou měrou jako distributoři.

Seznam literatury

1. **European Broadcasting Union.** *Loudness Normalisation and Permitted Maximum Level of Audio Signals.* Ženeva, Švýcarsko: EBU, 2020. EBU R128-2020.
2. **Advanced Television Systems Committee.** *ATSC Recommended Practice: Techniques for Establishing and Maintaining Audio Loudness for Digital Television.* Washington, D.C.: Advanced Television Systems Committee, 2013. A/85:2013.
3. **International Telecommunication Union.** *Algorithms to measure audio programme loudness and true-peak audio level.* Ženeva, Švýcarsko: ITU, 2015. ITU-R BS.1770-4.
4. **Český normalizační institut.** *Akustika – Terminologie.* Praha: Český normalizační institut, 2003. ČSN 01 1600.
5. **Melka, Alois.** *Základy experimentální psychoakustiky.* Prague: Akademie múzických umění v Praze, 2005. ISBN 80-7331-043-0.
6. **Syrový, Václav.** *Hudební akustika.* Praha: Akademie múzických umění v Praze, 2013. ISBN 978-80-7331-297-8.
7. **Katz, Bob.** *Mastering Audio, Third Edition: The Art and the Science.* Burlington: Focal Press, 2015. ISBN 978-0-240-81896-2.
8. **H. A. Chinn, D. K. Gannett, and R. M. Morris.** *A New Standard Volume Indicator and Reference Level.* New York: Institute of Radio Engineers, 1940. Proceedings of the I.R.E., Vol. 28.
9. **International Telecommunication Union.** *Audio levels and loudness.* Ženeva: ITU, 2014. ITU-R BS.2054-4.
10. **Robjohns, Hugh.** What's the difference between PPM and VU meters? *Sound on sound.* [Online] SOS Publications Group, červenec 2013. [Citace: 18. červen 2021.] <https://www.soundonsound.com/sound-advice/q-whats-difference-between-ppm-and-vu-meters>.
11. **Grimm, E.M., Everdingen, R. Van a Schöpping, M. J. L. C.** *Toward a Recommendation for a European Standard of Peak and LKFS Loudness Levels.* místo neznámé: SMPTE, 2010. 10.5594/J11396.
12. **Thornton, Mike.** Understanding The Advanced Metering In Pro Tools - Expert Tutorial. *Pro Tools Expert.* [Online] 6. ledna 2021. [Citace: 18. červen 2021.] <https://www.pro-tools-expert.com/home-page/2018/5/4/tutorial-understanding-the-advanced-metering-in-pro-tools>.
13. **European Broadcasting Union.** *Guidelines for Production of Programmes in accordance with R 128.* EBU. Ženeva, Švýcarsko: EBU, 2016. Tech 3343-2016.
14. **AES Technical Committee on Transmission and Broadcasting.** *Recommendation for Loudness of Audio Streaming and Network File Playback.* New York: Audio Engineering Society, Inc., 2015. AES TD1004.1.15-10.
15. **Grimm, Eelco.** *Recommendation for loudness normalization by Music Streaming Services.* [Document] Utrecht: University of the Arts Utrecht, 2017.
16. **International Telecommunication Union.** *Short-term loudness metering.* Ženeva: ITU, 2008. ITU-R BS.2103-1.

17. **European Broadcasting Union.** *Normalizace hlasitosti a maximální povolená úroveň zvukových signálů. (překlad)* Ženeva/Praha: EBU, Rada pro rozhlasové a televizní vysílání, 2011. EBU R128-2011.
18. **Soulodre, Gilbert A.** *Evaluation of Objective Loudness Meters.* Berlín: Audio Engineering Society, 2004. Convention Paper 6161.
19. **Dolby Laboratories, Inc.** Dolby.io – Voice, Audio and Video API's for Communications and Media Transformations. *Understanding Loudness.* [Online] Dolby Laboratories, Inc. [Citace: 17. června 2021.] <https://dolby.io/developers/media-processing/guides/understanding-loudness>.
20. **Hugo Fastl, Eberhard Zwicker.** *Psychoacoustics: Facts and Models.* Berlin: Springer, 2006. ISBN 978-3-540-68888-4.
21. **International Telecommunication Union.** *Requirements for loudness and true-peak indicating meters.* Ženeva: ITU, 2012. ITU-R BS.1771-1.
22. **TC Electronic.** TC Electronic | Broadcast Standards. *TC Electronic.* [Online] Music Tribe Global Brands Ltd., 2021. [Citace: 14. dubna 2021.] <https://www.tcelectronic.com/broadcast-standards.html>.
23. **European Broadcasting Union.** *'EBU MODE' METERING TO SUPPLEMENT EBU R 128 LOUDNESS NORMALISATION.* Ženeva: European Broadcasting Union, 2016. Tech 3341-2016.
24. **European Broadcasting Union.** *Loudness range: A measure to supplement Loudness normalisation in accordance with EBU R 128.* Ženeva, Švýcarsko: EBU, 2016. Tech 3342-2016.
25. **European Broadcasting Union.** *Loudness Parameters for Short-form Content.* Ženeva: EBU, 2020. EBU R128 s1-2020.
26. **Česká republika.** Vyhláška o některých charakteristikách zvukové složky reklam, teleshoppingu a označení sponzora v televizním vysílání a o způsobu měření hlasitosti zvukové složky reklam, teleshoppingu a označení sponzora v televizním vysílání. *Předpis 122/2013 Sb.* Praha: Rada pro rozhlasové a televizní vysílání, 2013. Sv. Částka 55.
27. **European Broadcasting Union.** *Specification of the Broadcast Wave Format (BWF).* Geneva: EBU, 2011. EBU Tech 3285.
28. **Thornton, Mike.** Loudness And Dynamics In Cinema Sound - Part 1. *Pro Tools Expert.* [Online] 5. července 2017. [Citace: 16. června 2021.] <https://www.pro-tools-expert.com/home-page/2017/6/21/loudness-and-dynamics-in-cinema-sound>.
29. **International Telecommunication Union.** *Operational practices for loudness in the international exchange of digital television programmes.* Ženeva: ITU, 2016. ITU-R BS.1864-0.
30. **Spotify AB.** Loudness normalization. *Spotify for Artists.* [Online] Spotify AB. [Citace: 24. května 2021.] <https://artists.spotify.com/help/article/loudness-normalization>.
31. **Armstrong, Martin.** The world's most popular music streaming services. *Statista.* [Online] 13. února 2020. [Citace: 25. dubna 2021.] <https://www.statista.com/chart/20826/music-streaming-services-with-most-subscribers-global-fipp/>.

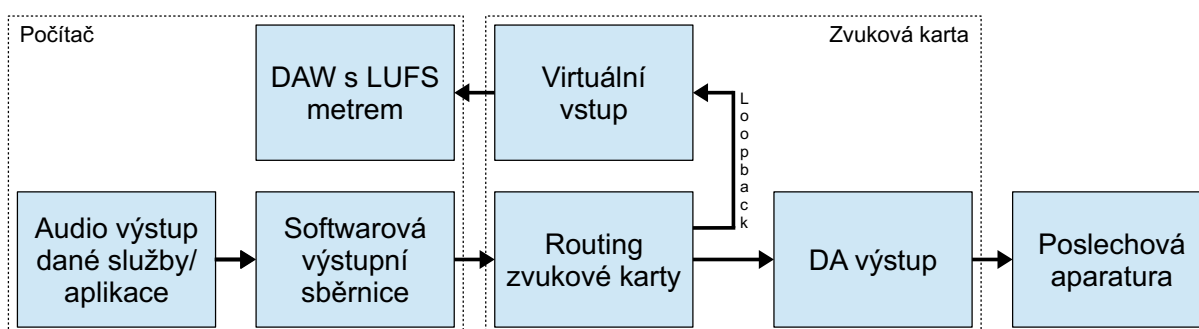
32. **Stewart, Ian.** Mastering for Streaming Platforms: Normalization, LUFS, and Loudness Myths Demystified. *iZotope.com*. [Online] iZotope, Inc, 20. května 2021. [Citace: 10. června 2021.] <https://www.izotope.com/en/learn/mastering-for-streaming-platforms.html>.
33. **Apple Computer, Inc.** Apple Digital Masters: Music as the Artist and Sound Engineer Intended. *Apple*. [Online] duben 2021. [Citace: 13. června 2021.] <https://www.apple.com/itunes/docs/apple-digital-masters.pdf>.
34. **Apple, Inc.** Audio requirements. *Apple Podcasts for Creators*. [Online] [Cited: 13. června 2021.] <https://podcasters.apple.com/support/audio-requirements>.
35. **Alexa Internet, Inc.** youtube.com Competitive Analysis, Marketing Mix and Traffic - Alexa. *Keyword Research, Competitor Analysis, & Website Ranking | Alexa*. [Online] Alexa Internet, Inc., 13. června 2021. [Citace: 13. června 2021.] <https://www.alexa.com/siteinfo/youtube.com>.
36. **Thornton, Mike.** Has Netflix Turned The Clock Back 10 Years Or Is Their New Loudness Delivery Spec A Stroke Of Genius? *Pro Tools Expert*. [Online] 30. srpna 2018. [Citace: 20. června 2021.] <https://www.pro-tools-expert.com/home-page/2018/8/23/has-netflix-turned-the-clock-back-10-years-or-is-their-new-loudness-delivery-spec-a-stroke-of-genius>.
37. **European Broadcasting Union.** *Loudness in Streaming*. Ženeva: EBU, 2020. EBU R128 s2-2020.
38. **MeterPlugs.** Loudness Penalty. *Loudness Penalty*. [Online] MeterPlugs, 2021. [Citace: 20. června 2021.] <https://www.loudnesspenalty.com>.
39. **Katz, Bob.** *An Integrated Approach to Metering, Monitoring, and Levelling Practices*. Los Angeles: Digital Domain, Inc., 1999.
40. **European Broadcasting Union.** *Guidelines for distribution & reproduction in accordance with EBU R 128*. Ženeva: EBU, 2021. Tech 3344-2016.
41. *Loudness – don't forget the distribution chain!* **Van Everdingen, Richard.** Ženeva: EBU, 2012, Sv. Q3. ISSN: 1609-1469.
42. **Techopedia™.** Subscription Video on Demand (SVoD). *Techopedia*. [Online] Techopedia™. [Citace: 25. května 2021.] <https://www.techopedia.com/definition/29272/subscription-video-on-demand-svod>.
43. **TC Electronic.** CLARITY M. *tc electronic*. [Online] [Citace: 17. června 2021.] <https://www.tcelectronic.com/product.html?modelCode=P0D9L>.
44. **Cambridge University Press.** Cambridge Dictionary. *VOLUME | meaning in the Cambridge Dictionary*. [Online] Cambridge University Press, 2021. [Citace: 19. června 2021.] <https://dictionary.cambridge.org/dictionary/english/volume>.

Příloha 1: Slovník pojmů kolem normalizace hlasitosti

- Album normalisation** – normalizace hlasitosti na základě hlasitosti celého alba (s. 30)
- Anchor element** – vybraný úsek (většinou mluveného slova), podle kterého se nastaví („ukotví“) hlasitost celého pořadu (str. 27)
- ATSC A/85** – závazná norma pro televizní vysílání v USA (str. 27)
- Dialnorm** – parametr kodeku AC-3 obsahující hodnotu hlasitosti „anchor elementu“ nebo případně integrovanou hlasitost jako takovou. (str. 27)
- Dialogue gating** – gating podle mluveného slova upravující algoritmus BS.1770; v některých prostředích používáný pro posuzování hlasitosti (str. 14)
- EBU R 128** – norma upravující target level pro vysílání v Evropě (str. 25)
- Headroom** – rozdíl mezi nejvyšší dosažitelnou hodnotou a integrovanou hlasitostí (str. 19)
- Integrated loudness (L_K)** – integrovaná hlasitost, integrovaná v celé délce programu nebo měření (str. 11)
- ITU-R BS.1770** – norma definující výpočet integrované hlasitosti (str. 11)
- K-filter/K-weighted** – filtr předřazený $Leq()$ algoritmu při výpočtu LUFS (str. 11)
- $Leq()$** – algoritmus, na jehož základě funguje BS.1770 (str. 11)
- LKFS** – označení jednotky LUFS v dokumentech ITU
- Loudness range (LRA)** – jednotka makrodynamiky popisující dynamický rozsah audio materiálu (str. 18)
- LU** – loudness unit; 1. relativní jednotka v systému LUFS; 2. relativní stupnice v systému R 128, kde 0 LU = -23 LUFS (str. 17, 19)
- LUFS** – Loudness Unit Relative To Full Scale; hlasitostní jednotka vypočtená algoritmem BS.1770 (str. 11)
- Momentary loudness** – okamžitá hlasitost; integrační čas 400 ms (str. 16)
- Peak programme meter** – zastaralý měřič používáný převážně ve vysílání (str. 9)
- Peak to Loudness Ratio (PLR)** – přibližné měřítko mikrodynamiky audio materiálu; $PLR = |TP-PL|$ (str. 19)
- Programme loudness (PL)** – integrated loudness (integrovaná hlasitost)
- Short-term loudness** – krátkodobá hlasitost; integrační čas 3 s (str. 17)
- Sound check** – proprietární nástroj firmy Apple k normalizaci hlasitosti (str. 32)
- Speech gating** – dialogue gating
- Target level** – cílová hodnota hlasitosti při normalizaci (str. 24)
- Track normalisation** – normalizace jednotlivých skladeb bez ohledu na kontext v rámci alba (str. 30)
- True peak (TP)** – skutečná špička signálu, tedy včetně skrytých špiček (str. 15)

Příloha 2: Řetězec pro měření výstupu neznámých aplikací

Pro ověřování hlasitosti neznámých softwarových audio výstupů byl použit řetězec popsaný na obrázku P2.1 používající zvukovou kartu s funkcí *loopback*. Tím je možné poslat digitální signál z výstupu zpět na vstup, což umožňuje měřit výstup softwarové audio sběrnice. Pro měření bylo použito VST pluginu *iZotope Insight*. Vzorkovací frekvence zvukové karty byla vždy 44,1 kHz.



Obrázek P2.1: Řetězec pro měření výstupu neznámých aplikací.



Obrázek P2.2: Měřič hlasitosti iZotope Insight.