

Bitová hloubka signálu

Zdroj: <http://www.audiozone.cz/recenze/stoparuv-pruvodce-digitalnim-zvukem-1-dil-t18487.html>

Co vlastně znamená „digitální“ a co všechno určuje bitová hloubka.

Digitální technika je dnes samozřejmostí. Je všude kolem nás a obzvlášť práci s videem a se zvukem si bez ní lze jen těžko představit, pokud uvažujeme běžnou praxi, extrémy vynechme. K jejímu rozšíření v praxi došlo docela rychle, přesto se zdá, že se s ní stále mnoho lidí zcela nesžilo a snaží se spíš jen přežít onu nevyhnutelnou spolupráci. Tento seriál by měl nejen přispět k ujasnění pojmů a principů digitální techniky a digitálního zpracování signálu, ale také posloužit jako přehled některých praktických dopadů vycházejících z těchto principů a vyvrátit některé mýty, které jsou, zdá se, mezi uživateli rozšířeny.

Digitální vs. Analogový

Tyto dva pojmy se v audio oblasti vyskytují prakticky neustále. Přesto myslím, že nebude na škodu si je vyjasnit. V praxi jsem se setkal například s představou, že „digitální = takové to novější a analogové = takové to starší“. Podle zaměření člověka je to pak také často zabarveno směrem „analogové = to starší a jediné pořádné, digitální = to nové strašně kazící zvuk“, což je názor velmi častý mezi hudebníky. Existuje i obrácená varianta, čili „analogové = to staré, překonané, digitální = to jediné správné, co je schopné přenést superkvalitní zvuk mé supernahrávky“.

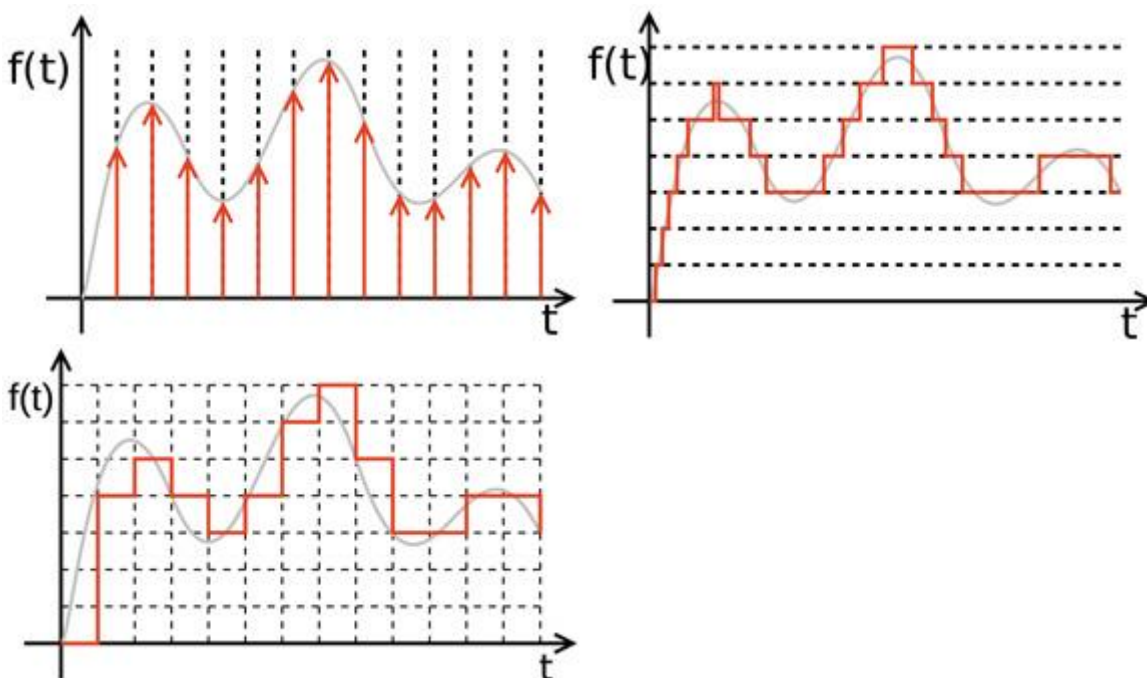
Akustické signály jsou spojité. Když je takový signál zaznamenán do analogového zařízení používajícího např. magnetický pásek jako záznamové médium, je převeden do jiné spojité veličiny, u magnetofonu do průběhu magnetizace pásku. Výraz digitální znamená doslova číslicový, hodnoty signálu jsou tedy uloženy ve formě čísel a to konečného počtu konečně velkých čísel. Z toho vyplývají zásadní vlastnosti, že signál v digitální podobě je tzv. vzorkovaný a kvantovaný – známe hodnoty pouze v daných časových okamžicích, tyto hodnoty jsou navíc z uzavřeného oboru. To je také příčinou dalších vlastností jako omezení frekvenčního rozsahu nebo kvantizační šum. Nejdůležitější je ale uvědomit si, že jsou to především zcela odlišné principy práce se signálem a ani jeden z nich není celkově lepší nebo horší, protože nabízejí odlišné možnosti a vlastnosti. Diskuze o vlivu digitalizace signálu na kvalitu jeho vjemu při následné reprodukci běží už dlouho a asi určitým způsobem bude probíhat neustále. Jsou tu ale například tyto dva nepopiratelné fakty:

1. Výměnou za vzorkování a kvantizaci úrovní nabízí digitální oblast možnosti stříhu a práce se záznamem, bez kterých bychom si dnešní praxi asi těžko dokázali představit (o možnosti kreativní práce se zvukem na téměř jakémkoli počítači - i na tom úplně malém co umí i telefonovat - ani nemluvě).

2. Poslech je v dnešní době poměrně drtivě prováděn z digitálních médií a navíc často na

nekvalitních zařízeních a pomocí formátů využívajících datovou kompresi. Takže ač je analogový záznam z principu kvalitnější a přesnější, koncovému uživateli je to často naprosto jedno.

Digitální svět ale nepřinesl jen zjednodušení, některé v analogu jednoduché záležitosti jsou v digitální oblasti řešeny poměrně složitě, aby se kvalitou mohly přiblížit svému analogovému předobrazu. To se týká třeba zkreslení a dynamických efektů, kde se nelinearity a časové chování elektrických obvodů musejí poměrně složitě matematicky modelovat. Jiné efekty jsou zase digitální už z principu (např. algoritmické dozvukové procesory jako legendární zařízení firmy Lexicon). Dalším příkladem je tzv. scrubbing záznamu pro nalezení přesného místa stříhu, kazu atp. Na studiovém magnetofonu jednoduše hýbeme s páskem jak je libo a celá věc „prostě funguje“ (s patřičným podladěním či nadladěním), protože časové rozlišení je téměř nekonečně jemné. Když ale to samé budeme dělat v audio aplikaci, máme k dispozici pouze stále stejný počet vzorků, tedy časových okamžiků, kdy hodnotu signálu známe. Zbytek je potřeba pro přehrání dopočítat, aby byl vytvořen souvislý zvuk bez nespojitostí, které by rušily (byly by to kazy vytvořené metodou pro hledání kazů, což práci moc neusnadňuje...). V tom, jak je scrubbing řešen, jsou mezi jednotlivými audio aplikacemi značné rozdíly. Dle mého názoru je tento nástroj velmi dobře řešen v Avid Pro Tools, kde skutečně připomíná práci s magnetofonem (jak mohu soudit ze svého letmého seznámení se s prací s analogovým záznamem). Naopak Steinberg Cubase v tomto směru dost pokulhává, ačkoli v poslední době zde byl scrubbing vylepšen. Před vylepšením byl ale nepříliš použitelný, stále je tedy prostor pro zdokonalení.



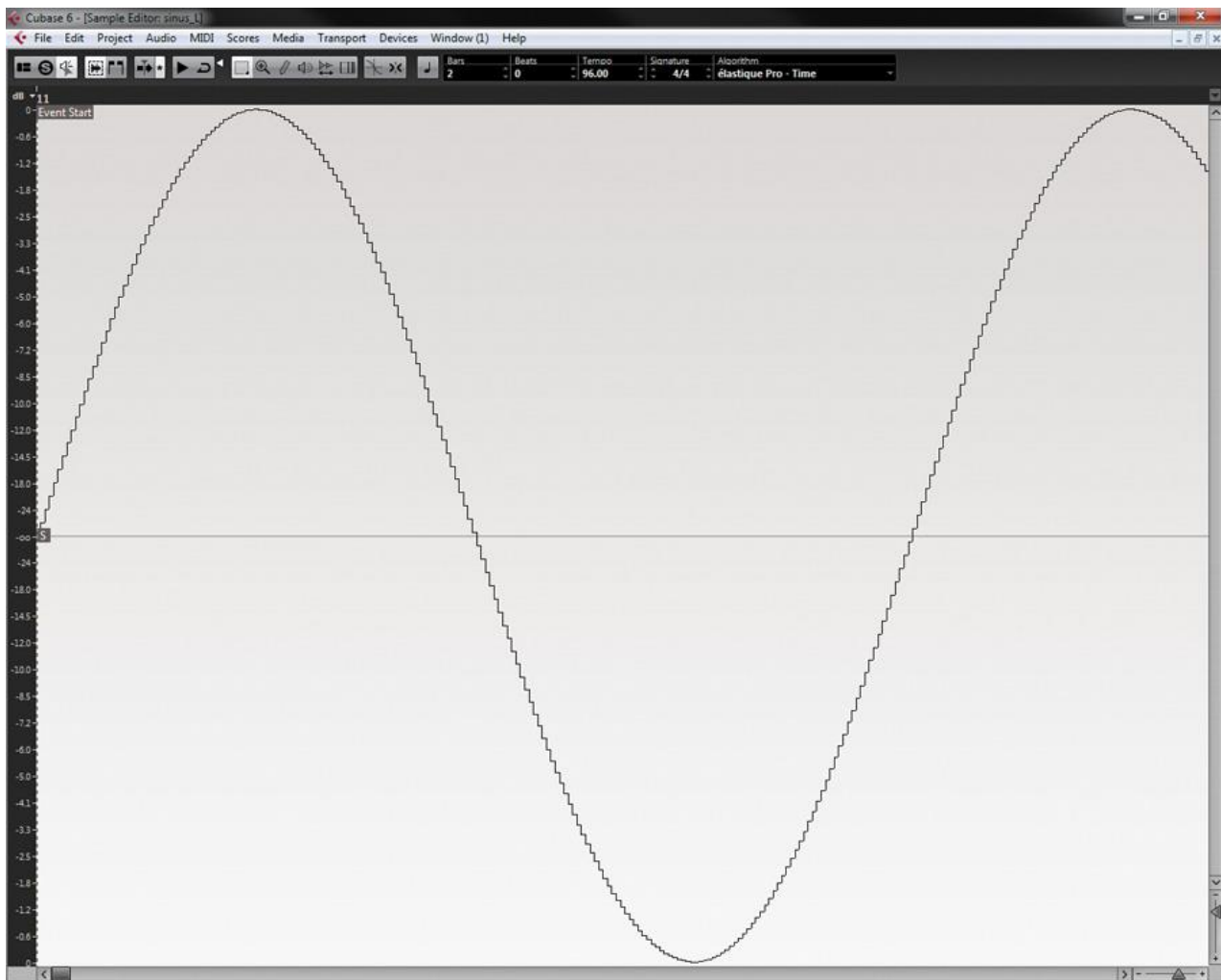
Horní dva obrázky zleva doprava: Vzorkovaný signál bez kvantizace úrovně, kvantizovaný signál bez vzorkování a Spodní obrázek: Výsledný digitální signál.

Základní pojmy digitálního systému

„Bity a kiloherty“ létají v debatách o zvuku všemi směry. Myslím ale, že nebude na škodu přece jen si vyjasnit, o co přesně se jedná. Jak již bylo řečeno, digitální signál je uložen ve formě konečného počtu konečně velkých čísel. S tím právě úzce souvisejí dva základní pojmy, bitová hloubka kvantizace a vzorkovací frekvence. Zjednodušeně řečeno je vzorkovací frekvence údajem o tom, jak často se převodník (např. na analogovém vstupu zvukové karty nebo kamery) podívá na momentální hodnotu signálu a bitová hloubka říká, jak přesně se tato hodnota zapíše. V tomto dílu našeho seriálu se zaměříme na bitovou hloubku neboli rozlišení či také délku slova signálu.

Bitová hloubka

Vezmeme-li signál v Audio CD kvalitě 16 bitů a 44,1 kHz, jeho bitová hloubka nám říká, že údaj o úrovni signálu je zaznamenán pomocí 16místného čísla ve dvojkové soustavě. Z toho vyplývá, že spojitý (čili mající nekonečný počet možných úrovní) akustický signál zaznamenaný tímto způsobem může v digitální podobě nabývat 65.536 různých hodnot (2^{16}). V praxi záznamu audio signálu se setkáme převážně s hodnotami 16 a 24 bitů. Jsou to formáty s pevnou řádovou čárkou, čili je jasně definováno nejvyšší číslo, které je formát schopen uchovat. To je vysvětlení, proč dochází k tvrdé limitaci neboli ořezu signálu při překročení této hodnoty cejchované 0 dBFS. FS značí anglický výraz Full Scale neboli „plný rozsah“ a jelikož je dB poměrná jednotka, záporné hodnoty na stupnici nám ukazují, o kolik je momentální hodnota menší než maximum. To je velká změna oproti praxi v čistě analogovém systému, protože tato jasná definice stropní hodnoty signálu přinesla úplně nový přístup k úrovním a práci s efektivní hodnotou signálu (RMS - Root Mean Square). Analogová nula zařízení je často spíše optimální hodnota RMS a špičkové úrovně nahrávek se mohou lišit v závislosti na žánru. Při přebuzení do kladných hodnot např. sběrnice hlavního mixu dochází v některých případech ke kvalitnímu a uchu příjemnému zkreslení, které se významně podílí na charakteru známém jako „analogový zvuk“. S digitálním zvukem se ale přešlo na srovnání špičkové úrovně a tím pádem se liší úrovně RMS, tedy subjektivní hlasitost. To eventuálně vedlo k současné špatné situaci známé jako Loudness War - válka hlasitostí.



Přibližně 1,5 periody sinusového signálu o frekvenci 220 Hz v kvalitě 16 bitů, 44,1 kHz.

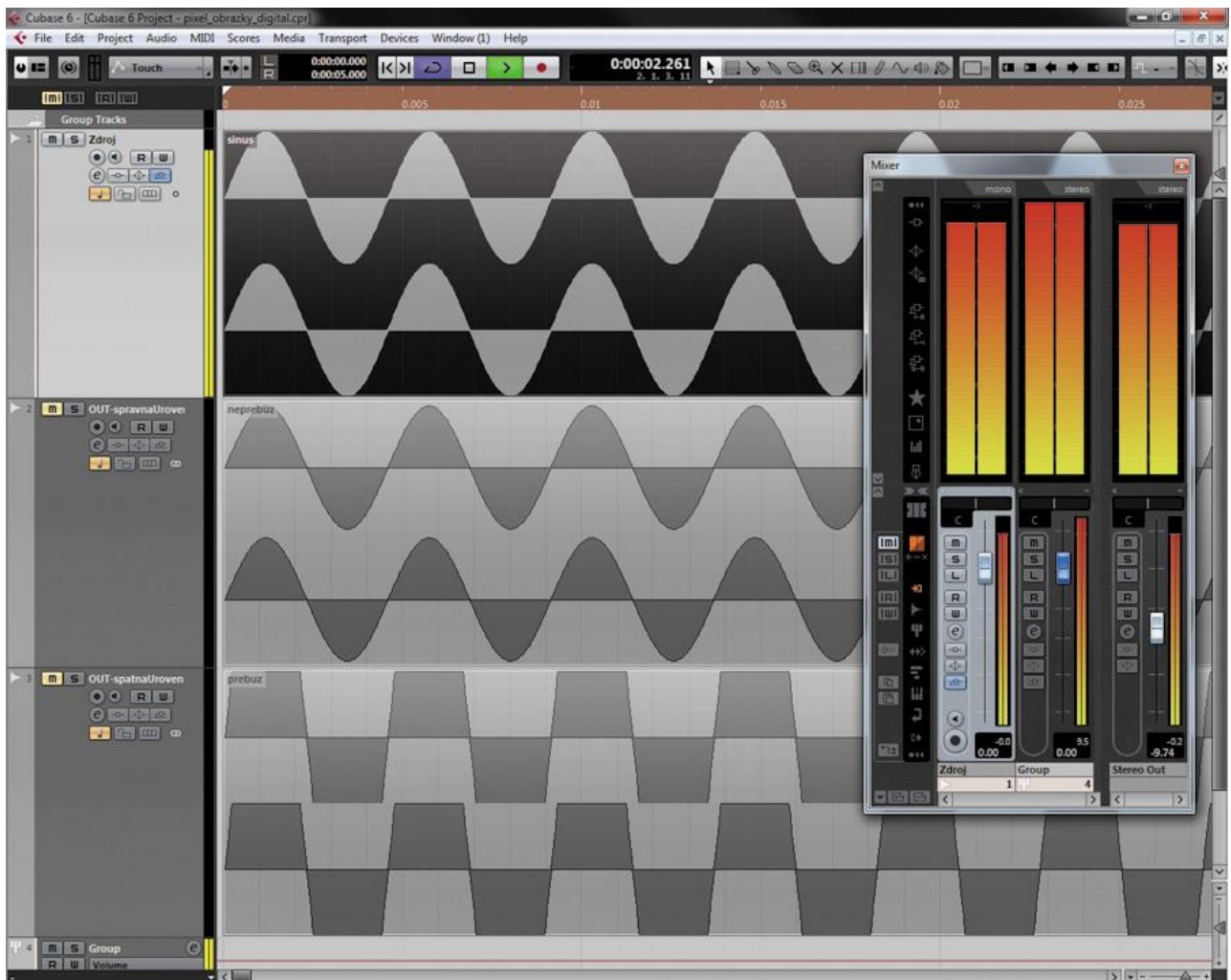
Není úplně obtížné narazit na názor, že 24bitové signály „mohou hrát o X dB hlasitěji“. To není pravda, 0 dBFS jakožto stropní hodnota stále odpovídá maximální hodnotě napětí v případě jak vstupu, tak i výstupu převodníku. Rozdíl si totiž musíme představit tak, jako kdybychom začali zapisovat např. o jedno desetinné místo navíc a získali tak další čísla mezi předtím sousedícími hodnotami. Místo nutnosti přeskočit z 0,8 rovnou na 0,9 máme najednou k dispozici řadu 0,81, 0,82 ... 0,89 0,9. 24bitové rozlišení tedy nabízí vyšší počet úrovní a proto přesnější převod, protože musíme zaokrouhlovat méně. Poměr vůči 16 bitům je konkrétně 2^{24} vs. $2^{16} = 16.777.216$ vs. 65.536. Z hlediska zvuku není lineární míra příliš podstatná, srovnajme si proto oba rozsahy v logaritmické míře v dB, poměr je pak 144 dB vs. 96 dB. Tato čísla také ukazují teoretický odstup od tzv. kvantizačního šumu (kvantizační chyby) a tím pádem teoretický dynamický rozsah. Kvantizační šum vzniká zaokrouhlováním úrovně signálu na nejbližší možnou hodnotu při převodu do digitální oblasti. Má velikost poloviny vzdálenosti mezi úrovněmi, zjednodušeně řečeno je to hodnota, kterou musíme „zahodit“, protože nemáme kolonku, kam bychom ji zapsali. V praxi je ale šum vstupního signálu téměř vždy výrazně větší než kvantizační šum 16bitového, natož 24bitového převodu. Mohlo by se tedy zdát, že větší dynamický rozsah jako takový je ta zásadní výhoda, kterou nám 24bitové rozlišení přináší. 144 dB je skutečně zajímavé číslo. Je ale dobré si uvědomit, že rozdíl mezi prahem slyšení a prahem bolesti na stupnici hladiny akustického tlaku je 130 dB. To ukazuje, že v praxi bude dynamický rozsah zaznamenávaného signálu nižší nebo ho bude nutné

redukovat, ať už ručně nebo pomocí dynamických procesorů. Nahrávání do digitálního záznamového zařízení je věčným soubojem o správně zvolené vstupní úrovni. Je jasné, že příliš vysoká úroveň přináší značné riziko přebuzení a nevratného (pouze retušovatelného) poškození signálu. Naopak příliš malá úroveň, tzv. podmodulování signálu, vede k nedostatečnému využití rozsahu převodníku. Tuto situaci si lze představit tak, že číslice vyšších řádů jsou stále na nule a nikdy se do nich nezapíše žádná hodnota, je to tedy jako bychom převáděli pomocí převodníku nižší bitové hloubky. Platí přibližně, že každý bit přináší 6 dB rozsahu. 24bitové rozlišení nám tak pro nahrávání umožňuje ponechání větší úrovně rezervy, aniž bychom potřebnou kvalitu nedostatečným využíváním rozsahu ztráceli a snižovali výrazně odstup signál-šum. Z rozdílu dynamických rozsahů vyplývá, že když bude nahrávaný signál dosahovat ve špičkách úrovně -48 dBFS při 24bitovém převodu, bude z hlediska využívaného počtu hladin (tedy přesnosti při kvantování) situace stejná jako na výstupu plně využitého 16bitového převodníku. Navíc máme ale stále 48 dB, které můžeme rozložit mezi rezervu pro nečekané nárůsty úrovně (protože např. živý záznam koncertu symfonického orchestru může přinést lečjaká překvapení) a dosažení přesnějšího záznamu díky většímu využití rozsahu převodníku.

Bitová hloubka při práci uvnitř počítačových audio aplikací

Mnoho z nás se setkalo se situací, kdy ukazatel úrovně příslušející stopě projektu ukazuje kladné hodnoty, vše svítí červeně a přesto nedochází ke zkreslení. To zdánlivě protičečí předcházejícím odstavcům. Při práci s více stopami je často třeba sčítat velké množství signálů a některé efekty vytvářejí značný úrovněový nárůst. Kdyby audio aplikace interně pracovaly ve stejném rozlišení jaké mají zdrojové audio soubory, byla by práce velmi nepohodlná a hlídání úrovní by bylo ještě komplikovanější než při nahrávání (bylo by nutné neustále snižovat všechny úrovně při přidání další stopy do mixu atp.). Navíc je třeba pokud možno nezvyšovat úroveň kvantizačního šumu, který by při sečtení většího množství signálů uvnitř stále stejného rozsahu mohl začít být patrný.

Jedním z možných řešení, starším a dobře kompatibilním s dedikovanými procesory (jako je Motorola 56k využívaný ve výpočetních kartách k Avid/Digidesign Pro Tools), je přidat nějaké ty bity navíc. Jak jsme si již řekli, přidání každého znamená zvýšení dynamického rozsahu o 6 dB. Systém Avid ProTools pracuje do verze 9 s interním rozlišením sběrnic 48 bitů, čili při využívání 24bitových souborů jakožto zdrojů je k dispozici dalších 144 dB na výpočty. U Pro Tools ale všechny audio stopy v projektu pracují stále s bitovou hloubkou odpovídající materiálu. Proto je potřeba úroveň přece jen hlídat, protože by mohlo dojít ke zkreslení při zápisu do audio souborů nebo na výstupní převodník. V Pro Tools proto mají všechny měřáky jako nejvyšší hodnotu nulu, což usnadňuje přehled o správném toku signálu.



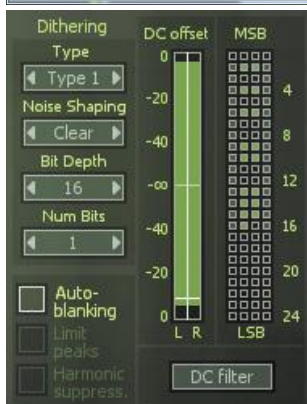
Ukázka přebuzení stopy v Steinberg Cubase a srovnání oříznutého a správného výstupního signálu.

Odlíšný přístup je zvolen ve většině čistě softwarových aplikací jako je např. Steinberg Cubase. Využívá se 32bitové rozlišení s plovoucí řádovou čárkou. 24 bitů - mantisa zůstává pro hodnotu čísla jako takovou a 8 bitů je exponent ukazující, v jakém řádu se zrovna pohybujeme. Pokud by mělo dojít k přetečení, posune se pouze řádová čárka. Tento formát zápisu čísel je v praxi velmi těžko přebuditelný, jelikož celkový dynamický rozsah je cca 1.500 dB. Přesto to není dokonalé řešení všech problémů. Při pohybu ve vysokých číslech ztrácíme přesnost (jelikož známe vždy jen 24 nejvyšších číslic a zbytek se ztrácí) a celkově jsou šumové vlastnosti o něco horší než u pevných rozsahů. Je to však ta nejlepší možnost, jakou v současné situaci máme a aktuální dedikované procesory ji využívají také (jako Analog Devices SHARC 21369 použitý u výpočtových UAD-2 karet firmy Universal Audio). Je také dobré si při hodnocení opět připomenout praktická fakta (na čem bude výsledek poslouchán, z jakého formátu atp.), protože to nejistoty ohledně ztrát kvality uvnitř audio systému často vyřeší. Tento číselný formát nalezneme také u technologie VST pro všechny operace (čili jak pro zpracování signálů, tak pro předávání hodnot parametrů mezi plug-iny a hostitelskou aplikací). Je to právě využití tohoto rozlišení, díky kterému je možné „přebuzovat“ audio stopy. Hodnota nula totiž odpovídá stále maximální hodnotě rozlišení s pevnou řádovou čárkou, což nám usnadňuje přehled o vztazích mezi vnitřními a vstupními a výstupními úrovněmi. Důležitá je volba bitové hloubky, když exportujeme zvukové soubory pro jejich další zpracování, ať už jde o finální mix nebo jednotlivé stopy. Pokud bychom totiž stopu, na které máme signál větší

než 0 dBFS, nechali uložit do audio souboru s menším rozlišením než 32 bitů s plovoucí čárkou, dojde k jeho ořezu. To je dobré si uvědomit právě při exportu stop, které jsou zapojeny někde uvnitř mixu a můžeme na nich mít vyšší úroveň, aniž bychom si toho všimli. Pokud pak uděláme export všech stop najednou, můžeme být později nepříjemně překvapeni. Je dobré si stopy pro další práci ukládat přímo ve formátu 32 bitů s plovoucí čárkou, abychom se vyhnuli zbytečným konverzím.

Převod bitových hloubek a dithering

Popisovat tuto problematiku by vydalo na samostatný článek, pro doplnění předcházejícího textu si alespoň stručně popíšme, o co se jedná. Přímý převod z nižšího rozlišení na vyšší nám moc užítku nepřinese, protože informaci nelze jen tak zpřesnit. Dojde jen k dopsání nul a zvětšení datového objemu, který soubor zabírá. Něco jiného je ukládat zpracovaný 16bitový zvuk do vyššího rozlišení, abychom co nejkvalitněji uchovali výsledky zpracování. Jak bylo popsáno výše, nejlépe se k tomu hodí formát použitý při výpočtech, čili 32-bit float.



Využití bitového rozsahu při snížení bitové hloubky na 16 bitů pomocí Izotope Ozone 5

Často ale potřebujeme rozlišení snížit, protože mezi výstupními formáty převažuje stále 16 bitů (Audio CD, kompresní formáty, zvuk na vysílacích digitálních kazetách pro video atp.). K přetečení nemůže dojít, pokud jsme si ohlíželi, že nepřekračujeme hodnotu 0 dBFS na výstupním kanálu nebo převádíme 24 bitů na 16, kde je tato hodnota shodná. Prakticky dochází k zaokrouhlení přesnějších hodnot na méně přesné, čili nám vzrůstá kvantizační chyba. Právě k minimalizování slyšitelných projevů tohoto jevu slouží tzv. dithering. Velmi stručně řečeno se jedná o dodatečný šum určitého charakteru, který má za úkol „přebít“ kvantizační chybu. Tento šum totiž ruší méně než kvantizační chyba, která může mít i charakter zvukových artefaktů. Toto vše se ale odehrává při velmi nízkých úrovních a není proto až tak třeba lámat si s tím hlavu. Tento problém je třeba ošetřit u velmi dynamického zvuku, kde se např. v doznívání tónů ve velmi tichých pasážích může něco projevit. Dither je obsažen ve velkém množství výstupních limiterů (např. Waves L1 nebo Sonnox Limiter), což je myslím dostačující, jen je potřeba zvolit správnou výstupní bitovou hloubku!