

# Vzorkovací frekvence

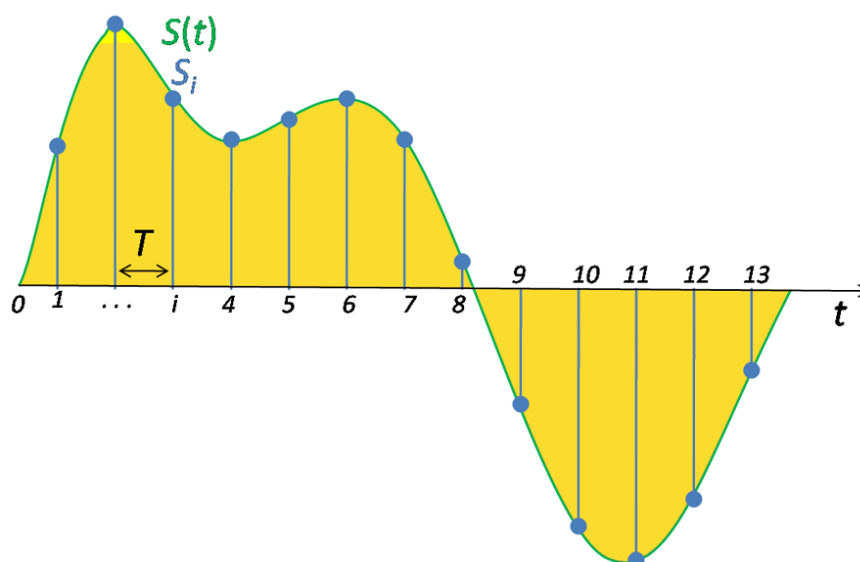
---

Zdroj: <http://www.audiozone.cz/recenze/stoparuv-pruvodce-digitalnim-zvukem-2-dil-t18556.html>

V tomto dílu se podíváme na problematiku spojenou s druhým ze základních parametrů digitálního signálu a to vzorkovací frekvencí. Zatímco bitová hloubka nám říká, kolik číslic má informace o úrovni signálu, vzorkovací frekvence udává, kolik těchto čísel za sekundu signálu známe. Udává se v hertzech (Hz), stejně jako frekvence signálu jako taková. Pro přesnost si doplníme, že tato jednotka udává počet zopakování nějakého děje za sekundu, fyzikální rozměr má  $s^{-1}$ , tedy doslova „za sekundu“. V našem případě je to např. počet okamžiků, kdy si vzorkovač запиše okamžitou hodnotu při AD převodu, nebo určení, kolik čísel za sekundu bude „přehráno“ při zpětném DA převodu. Převrácená hodnota vzorkovací frekvence je tzv. vzorkovací perioda čili údaj v sekundách informující nás o tom, jak daleko jsou od sebe dva sousední vzorky na časové ose.

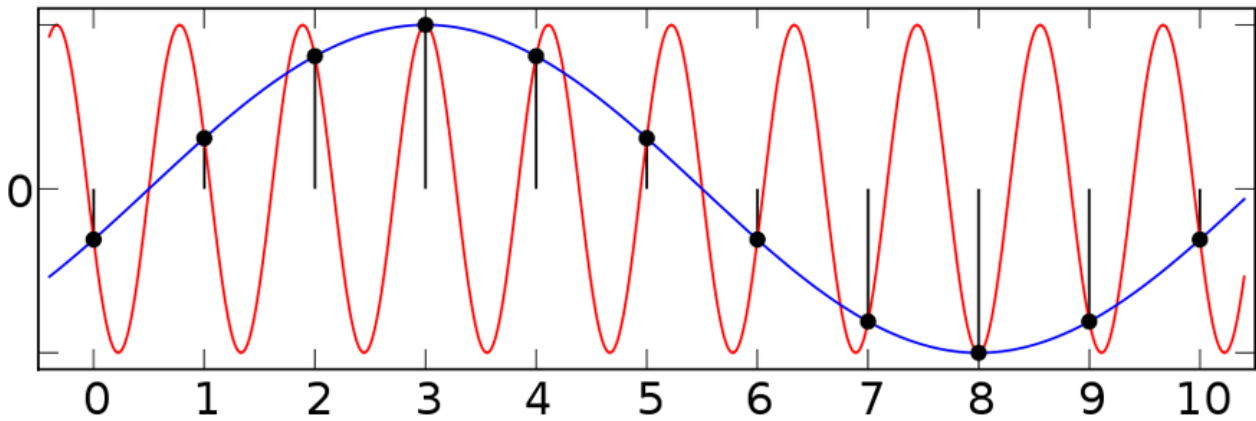
## **Vztah vzorkovací frekvence a spektra signálu**

Představme si analogový (tedy spojitý) signál. Když ho vzorkujeme, je to jako když začneme na jeho průběh v pravidelné vzdálenosti (dané vzorkovací periodou) umisťovat značky a jen ty si uložíme. Jako bychom si vybírali jen určitou část informace. Při zpětném převodu naopak mezi tyto značky opět kreslíme průběh signálu - podobně jako v dětských omalovánkách, kde se spojují očíslovaná místa, a vznikne obrázek. Je důležité si uvědomit, že když bude značek moc málo, nebudou si obrázek a předloha, ze které byly značky vytvořeny, moc podobné. V případě, že bychom takto kreslili siluetu horského masivu, stačí nám na vystižení velkých rozdílů (hluboká údolí, vysoké kopce) velmi málo uložených informací, když ale budeme chtít být schopni reprodukovat i detaily skalních útvarů, zjistíme, že bychom potřebovali bodů mnohem víc. Přesně to se děje i ve zvuku. Pomalé změny odpovídající přítomnosti nízkofrekvenčního obsahu, tedy např. basové oblasti, není problém v digitální podobě uchovat. Jinak tomu ale bude s vysokofrekvenčním obsahem (míněno z hlediska lidského sluchu).



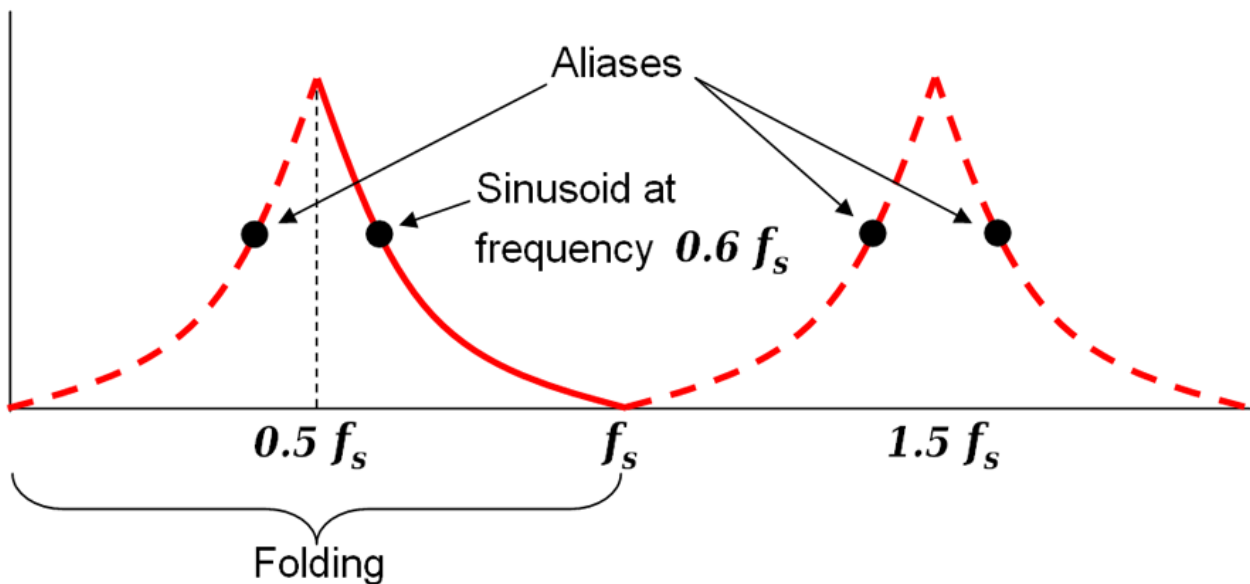
*Ilustrace vzorkování obecného signálu bez kvantizace úrovně.*

Vezměme si stále nejčastěji používanou vzorkovací frekvenci v hudbě, tedy  $44.100 \text{ Hz} = 44,1 \text{ kHz}$ , jejíž vzorkovací perioda (vzdálenost mezi vzorky) bude  $22,7 \text{ } \mu\text{s}$  (mikrosekund). Když budeme vzorkovat sinusový signál o frekvenci  $100 \text{ Hz}$ , jehož perioda je  $0,01\text{s}$ , budeme mít na její popsání cca 441 vzorků. Je logické, že pro frekvenci  $1.000 \text{ Hz}$  to bude cca 44 vzorků, pro  $10.000 \text{ Hz}$  přibližně čtyři vzorky. Pro  $20.000 \text{ Hz}$ , hodnotu uváděnou jako teoretický strop rozsahu lidského sluchu, půjde již pouze o dva vzorky, což evidentně není mnoho. Tento fakt je zjednodušeným vysvětlením původu nejdůležitější vlastnosti vzorkování pro uživatele - omezení spektra vzorkovaného signálu. Pokud budeme chtít být přesnější, omezení je důsledkem toho, že při převodu do digitální oblasti dochází k tzv. aliasingu. Vzorkovací teorém praví, že „Přesná rekonstrukce spojitého, frekvenčně omezeného signálu z jeho vzorků je možná tehdy, pokud byl vzorkován frekvencí alespoň dvakrát vyšší, než je maximální frekvence rekonstruovaného signálu.“ Z letmých výpočtů na předcházejících řádcích vyplývá, že při dalším zvyšování frekvence vzorkovaného signálu nám velmi záhy vzorky „dojdou“. Statisticky nebude zbývat ani jeden vzorek na periodu signálu. Je tedy jasné, že sinusoidu takové frekvence můžeme za daných okolností těžko správně rekonstruovat. Navíc se nám ale z uchovaných informací jeví jako frekvence jiná, odtud tedy termín aliasing.



*Vznik aliasingu - obě sinusoidy při vzorkování danou frekvencí dávají stejné vzorky.*

V případě vzorkování frekvencí 44.100 Hz jsme schopni uchovat složky do frekvence 22.050 Hz, čili poloviční. Vyšší složky by se nám tzv. zrcadlily, kdy pomyslné zrcadlo je umístěno na frekvenční ose v místě poloviny vzorkovací frekvence a odráží nám informace o složkách za zrcadlem zpátky mezi správně vzorkované frekvence. Tento „nepořádek“ dopadne na hodnotu rozdílu vzorkovací frekvence a dané frekvence, např. informace získaná za těchto podmínek vzorkováním signálu o frekvenci 34.100 Hz dopadne ve spektru na místo vyhrazené pro 10.000 Hz. To samozřejmě natropí nenávratné škody, protože nebudeme schopni rozpoznat, co patří k jaké frekvenci signálu. Řešení je relativně prosté, je třeba zařídit, aby za zrcadlem nic nebylo a nepoškodilo nám to naše hezky schované informace. Toho je dosaženo aplikací tzv. antialiasingového filtru, což je filtr typu dolní propust. Když si představíme, jak vypadá křivka ekvalizéru s tímto filtrem, dojde nám, že to není zdaleka tak jednoduchý úkol jak se může zdát, protože máme požadavek na vysokou strmost filtru a zároveň co nejnižší útlum v okolí kritické frekvence (poloviny vzorkovací frekvence). V tomto článku se tímto problémem nebudeme dále zabývat, uživatel audio techniky ho může považovat za vyřešený v převodníku, ale je dobré být si vědom, že existuje, skýtá mnohá úskalí a že způsob jeho vyřešení představuje jeden z rozdílů mezi levnými a dražšími převodníky.



Zrcadlení ve spektru signálu - bod 0,6 vzorkovací frekvence má ještě 3 „alias“.

### Používané audio formáty a vzorkovací frekvence

Velká většina digitálního audio hardwaru umožňuje pracovat hned v několika různých vzorkovacích frekvencích. Pojdme si v rychlosti projít ty nejpoužívanější.

**8.000 Hz** - vzorkovací frekvence používaná v klasické digitální telefonii. Vychází z toho, že těžiště informací v lidském hlase se nachází ve spektru do 4 kHz, dnešní komunikační formáty (např. VOIP) používají i vyšší v závislosti na použitém kodeku.

**44.100 Hz** - stále asi nejběžnější hodnota využívaná v CD-Digital Audio a u většiny hudby uložené ve formátu MP3. Původně zvolena firmou Sony kvůli relativně snadné implementaci do modifikovaného existujícího video hardwaru.

**48.000 Hz** - vzorkovací frekvence využívaná pro záznam zvuku do video hardwaru, používá se také v digitální TV, DVD i v moderních video rozhraních jako je SDI.

**96.000 Hz** - vzorkování využívané v DVD-Audio, Blu-ray a HD DVD, běžně se využívá i při nahrávání.

**192.000 Hz** - opět formáty s vysokým rozlišením, často nejvyšší frekvence, na které je ještě schopna pracovat většina profesionálního audio hardwaru a softwaru.

## Co znamená vzorkovací frekvence pro uživatele

Jakou frekvenci tedy zvolit, když zakládáme nový projekt? Je potřeba zvážit několik hledisek. Čím vyšší vzorkovací frekvenci v projektu zvolíme, tím vyšší bude frekvence nejvyšší zaznamenané složky signálu, což vyplývá z kapitoly o aliasingu. Pro 192 kHz to bude tedy až 96 kHz. To je skutečně úžasné číslo. Zároveň je ale nejvyšší frekvence vnímaná lidským sluchem 20 kHz a to v extrémním případě, prakticky je to číslo spíš ještě nižší. Narážíme tady na určitý nepoměr. Pokud nepatříme mezi těch pár magicky nadaných lidí, kteří na první poslech poznají rozdíl např. mezi 192 kHz a 96 kHz vzorkovaným signálem (samozřejmě za předpokladu propojení všeho superdrahými superkabely, jinak to přece „nehraje“), v čem je tedy ta výhoda, kterou nám vyšší vzorkování přináší? Projeví se především při dalším zpracování. Máme k dispozici více informací o signálu a proto jsme ho schopni mnohem více a lépe měnit a transformovat. Nejjasněji je tento rozdíl vidět na přeladování, změnách časové délky, opravách (např. dokreslování signálu „tužkou“), odšumování atp. Proto se například může vyplatit zaznamenávat si i ruchy a zvuky v 96 kHz, pokud s nimi plánujeme značně kreativně pracovat. Dnes nám totiž tuto možnost nabízejí i kapesní přehrávače jako např. Zoom H4n. Můžeme si tak usnadnit postprodukci, nebo ji naopak zkomplikovat...

Samozřejmě i zde platí něco za něco, takže za kvalitnější a větší možnosti při zpracování platíme nárůstem objemu dat. Jelikož minuta 24bitového stereo signálu vzorkovaného 48 kHz zabere cca 16,5 MB, je dvojnásobek nebo čtyřnásobek docela znát. Samozřejmě je to mnohem menší problém než před několika lety vzhledem k neustálému zlepšování kvality, rychlosti a velikosti datových úložišť. Vzrůstá ovšem také výpočetní náročnost při práci s takovými signály. Máme prostě mnohem víc informací, které je potřeba zpracovat. Opět to není neřešitelný problém, současná výkonnost počítačů umožňuje poměrně pohodlnou práci i při využití 192 kHz. Jen je dobré si tento fakt uvědomit a přizpůsobit mu způsob práce, který často může být z hlediska zohledňování náročnosti prováděných operací dost bezstarostný vzhledem k tomu, že při práci v 44,1 nebo 48 kHz zvládne i průměrný počítač opravdu velké množství paralelně běžících výpočtů. Pokud ale budeme dělat rekonstrukční práci (např. převod a čištění zvuku z kazety nebo desky) ve 192 kHz, narazíme využitím komplexních nástrojů jako jsou odšumovače a declickery (Izotope RX2, sada nástrojů X firmy Waves apod.) na výkonnostní strop počítače velmi rychle.

Také je dobré v případě pomýšlení na využití vyšších vzorkovacích frekvencí pořádně prohlédnout technické specifikace. Některá zařízení nenabízejí v těchto módech veškeré své možnosti. Např. převodníky umožňují převod jen omezeného počtu kanálů, populární rozhraní ADAT funguje při 96 kHz s polovičním počtem kanálů, při 192 kHz vůbec, vícestopé rekordéry neumožňují záznam do všech stop atp. Nárůst datového toku totiž často naráží třeba na limity interního procesoru zařízení, který obstarává směrování signálů, komunikačních kanálů (USB) nebo úložišť (paměťové karty). Samozřejmě je mnohý dnešní hardware na dostatečné úrovni pro téměř jakoukoli aplikaci v audio, ale hlavně u levnějších produktů je potřeba přesvědčit se o jejich schopnostech, přece jen u nich nelze očekávat, že jsou v nich nejnovější a nejvýkonnější komponenty.

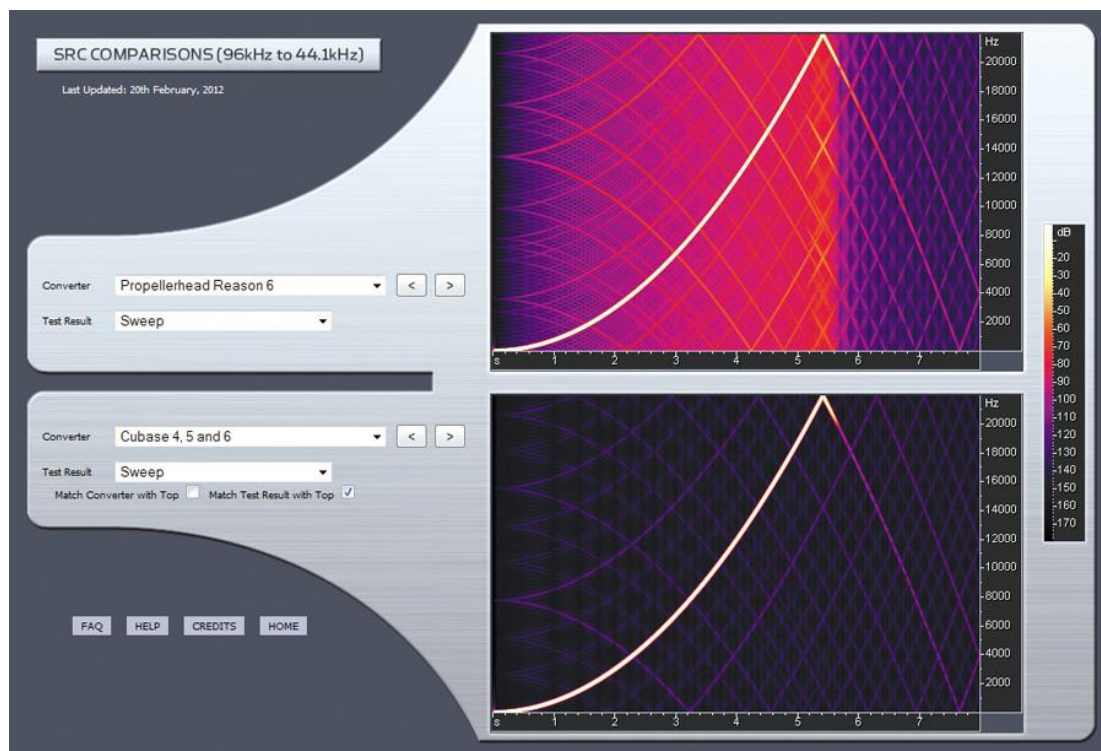
## Změny vzorkovací frekvence

Zatím jsme uvažovali pouze o situaci, kdy zvolíme nějakou vzorkovací frekvenci, zaznameneáme signál a pak s ním dále pracujeme. V praxi je ale často potřeba vzorkovací frekvenci záznamu změnit tak, aby všechny zdroje v projektu byly správně přehrávány. Tuto funkci mají dnes všechny audio aplikace a zdá se samozřejmá a jednoduchá. Jedná se o tzv. resampling nebo také sampling rate conversion (SRC) - stačí kliknout na „Ano, chci převést“ při importování materiálu do projektu a je to. Až tak jednoduché a ideální to ale není.

Když převzorkováváme čili zvyšujeme vzorkovací frekvenci, provádíme tzv. interpolaci, tedy „vymýšlíme“ vzorky přidané mezi ty stávající. Při podvzorkování provádíme naopak tzv. decimaci, kdy některé vzorky mažeme, například každý druhý, a to podle poměru frekvencí, mezi kterými převádíme. Když si představíme navzorkovaný sinusový signál, kde nám vzorky perfektně zapadají na významná místa signálu (vrcholy, nulové body atp.), zdá se všechno jednoduché, prostě dokreslíme další kuličky a je to. V praxi ale máme obecný komplexní signál. Co když navíc budeme převádět signál nahraný v 96 kHz na cílové médium, které bude velmi pravděpodobně vzorkováno 44,1 kHz (CD-DA, MP3)? Poměr těchto frekvencí je nesoudělný (nedává při dělení celé číslo), takže je zřejmé, že bude velmi těžké skutečně správně signál podvzorkovat a že se nějakým těm chybám nevyhneme. Z tohoto důvodu jsou převodníky většinou schopny pracovat v násobcích 44,1 kHz (např. MOTU Traveler má volby 44,1, 48, 88,2, 96, 176,4 a 192 kHz).

Přesto to ale nevyřeší veškeré problémy. Slušnou paseku může nadělat nedobře ošetřený aliasing (původní spektrum je totiž širší než to nové a je potřeba ho vhodně omezit), mohou vzniknout různé artefakty, časové problémy (pre-echa, zpoždění) atp. V praxi hudebního zvuku se proto zvýšená vzorkovací frekvence velmi často nepoužívá, resp. zohledňuje se cílové médium. V nahrávkách vážné hudby se téměř vždy pracuje s 44,1 kHz, pokud není záměr vydat produkt na některém HD médiu. Pokud je převod nevyhnutelný, mnoho profesionálů volí cestu „přes analog“ a raději dají do cesty signálu dvojnásobný převod na kvalitních převodnicích a signál si v cílové vzorkovací frekvenci znovu nahrají. Na stránkách [src.infinetwave.ca](http://src.infinetwave.ca) je k nahlédnutí test právě problematického převodu 96 kHz - 44,1 kHz. Jak je vidět na příkladu převedení signálu se stoupající frekvencí v čase, mají jak Steinberg Cubase, tak Propellerhead Reason docela problémy, Reason dokonce nezanedbatelné.

Na obrázku níže vidíme srovnání spektrogramů výstupního signálu. V ideálním případě bychom měli vidět pouze silnou stoupající křivku a ta by měla skončit v bodě, kde se dotýká horní hrany grafu, vše ostatní jsou chyby převodu. U Cubase je úroveň těchto nově vytvořených složek v řádu cca -110 dB, u Reasonu je to ale i pouhých cca 55 dB! Některé softwary jsou na tom ale mnohem lépe, často jsou to ty, které se na SRC specializují, např. Izotope 64bit SRC.



*Srovnání konverze 96 kHz - 44,1 kHz v programech Steinberg Cubase 6 a Propellerhead Reason 6.*

## Převzorkování při zpracování signálu

Některé plug-iny používají vnitřní převzorkování, pracují tedy s vyšší vzorkovací frekvencí nezávisle na tom, jak je nastaven hostitelský systém. Za všechny zmiňme ekvalizéry, které tento postup využívají k lepší kontrole chování na vysokých frekvencích (včetně samotné hranice frekvenčního rozsahu signálu), nebo nelineární efekty (zkreslení, dynamické procesory), kde převzorkováním zabráníme aliasingu nově vzniklých frekvenčních složek a můžeme upravený signál správně ošetřit před odesláním zpět do hostitele.

Možná překvapivě lze převzorkování využít také při měření úrovně. Digitální přebuzení je totiž definováno tak, že tři po sobě jdoucí vzorky mají hodnotu 0 dBFS, což je maximální možné číslo. V takovém případě usuzujeme, že došlo ke zkreslení. Problém ale nastává v současné stále vyhrocené situaci, kdy je velmi nadměrně používáno tzv. peak limiting neboli omezování špiček limiterem. Signál se často permanentně dotýká stropní hodnoty, ale přesto není indikováno přebuzení, protože nedochází ke splnění zmíněné podmínky. Při převodu do analogu musíme dostat spojitý signál a navíc v analogu neexistují skoky jako v digitální doméně, strmost analogového signálu je vždy konečná. Velmi často se pak stane, že dojde k překmitu přes maximální výstupní napětí převodníku a tedy ke zkreslení na analogové straně právě vlivem velmi časté přítomnosti maximální hodnoty v digitální oblasti. Takže ačkoli se spokojeně díváme na digitální měřák, který ukazuje, že je vše v pořádku, ve skutečnosti tomu tak není. Řešení je právě v použití měřáku s převzorkováním, který je schopen tuto „mezivzorková“ zkreslení (anglicky intersample clipping) detekovat, protože si pomocí převzorkování dopočítává chování signálu jakoby na půli cesty do analogového světa. Tuto schopnost mají dnes i leckteré limity (např. Sonnox Limiter



a jeho funkce „Recon Meter“ ukazující překmity nad zvolenou úroveň limitace) nebo dedikované měřicí nástroje (např. freewarový X-ISM od Solid State Logic mají dvě signalizace, pro digitální a analogový clipping).



*Pouze SSL X-ISM a Sonnox Limiter (na horní části měřáku vpravo) ukazují mezivzorkové zkreslení, iZotope Ozone 5 a interní měřák Cubase 6 ukazuje pouze špičku 0dBFS.*

## CD-DA kvalitu, nebo vysoké rozlišení?

V oblasti zvuku existuje mnoho mýtů a polopravd, které mohou stát zbytečně mnoho peněz. Jeden z nich se týká i vysokého rozlišení. Spousta lidí skálopevně tvrdí, co všechno nepoznají a jak je to všechno jasné. Slyšel jsem o cellistovi, který si svou hru nahrává na svůj Zoom H4n zásadně v 96 kHz, protože rozdíl je prý drastický. Používá k tomu integrované mikrofony, těžko říci, na čem výsledek poslouchá. Je velkou otázkou, jak se chovají mikrofony v tomto zařízení v oblasti nad 20 kHz, jak se v této oblasti chovají reprodukcni zařízení atd. Vášniví hifisté ve svých debatách často argumentují testy, které provedli, vesměs doma. Jenže je test ve stylu „dal jsem tam A, pak B a bylo to úplně o něčem jiném, jestli to neslyšíte, tak se nemáme o čem bavit, protože jste hlší“ skutečně průkazný? Samozřejmě není. Opravdový test vypadá jinak. Jeden takový provedli pod hlavičkou Boston Audio Society pánové Brad Meyer a David Moran a jeho výsledky publikovali v září 2007 v žurnálu Audio Engineering Society (AES). Během celého roku testovali skupinu 60 lidí, kteří byli vybráni z řad audio profesionálů, hifistů a studentů renomovaných zvukařských škol. Všichni byli audiometrickým vyšetřením testováni na slyšení zvuku v oblasti nad 15 kHz. Bylo provedeno celkem 554 srovnávacích testů. Srovnání probíhalo mezi SACD nebo DVD-Audio ve vysokém rozlišení reprodukováném na naprosto špičkovém řetězci a tím samým signálem, který prošel AD-DA převodem přes CD rekordér značky HHB, čili z digitálního hlediska byla jeho kvalita snížena na CD-DigitalAudio standard. Následně byla zesilovačem v cestě CD signálu (nemohl tedy „kazit“ SACD signál) dorovnána hlasitost mezi signály. Testovaní lidé měli poznat, který signál je



který. Navíc v náhodných chvílích, když chtěli mezi signály přepnout, k žádné změně nedošlo (tzv. A-B-X testování). A jak to celé dopadlo? Celkový výsledek byl 49,82 % správně vyhodnocených srovnání. Což je z pravděpodobnostního hlediska obdobné, jako když házíme mincí a podle toho, co nám na ní padne, se rozhodujeme...

Myslím, že v oblasti zvukové techniky a technologie se dá ušetřit hodně peněz, když člověk zbytečně mýtům nepropadá. Moje rada na závěr tedy zní, popřemýšlejte již na začátku projektu o svých záměrech a podle toho volte vzorkovací frekvenci. Vyhněte se zbytečným změnám, takže pokud to, co vám vyšší rozlišení nabízí, nehodláte využít (např. při sound designu, zásadních editacích délky, ladění, čištění atp.), volte vzorkovací frekvenci cílového formátu.