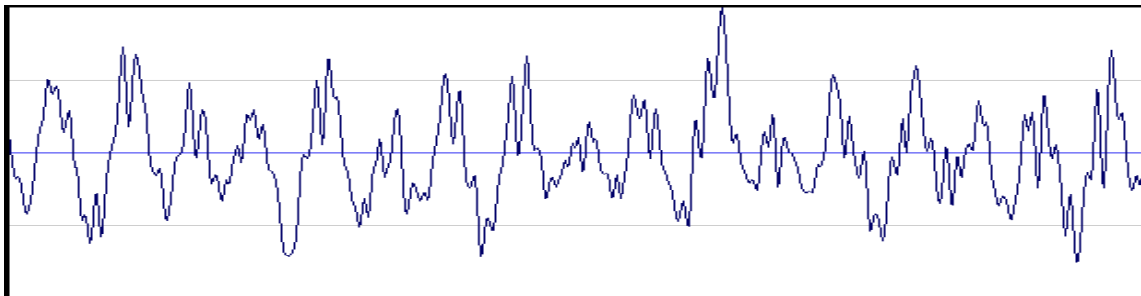


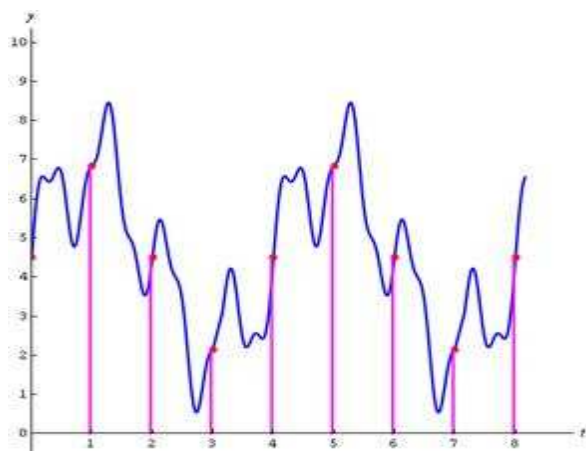
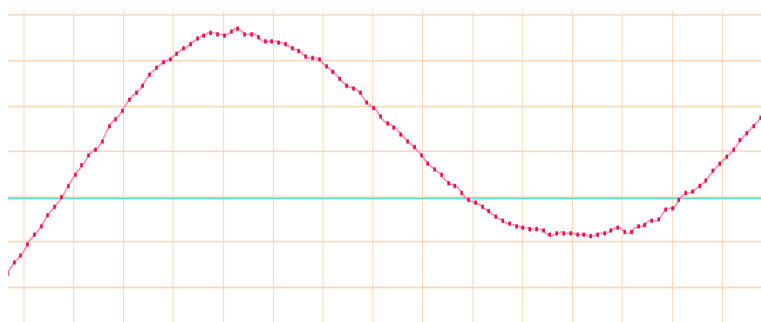
Digitální zpracování audio signálu – CD kvalita

(Materiál pro potřeby výuky předmětu KMT/UE – Pavel Kratochvíl, kratinek@kmt.zcu.cz)

Audio-signal je zásadně analogová veličina. Časový průběh akustické výchylky (případně akustického tlaku) je zachycen na následujícím obrázku. Shodný je i časový průběh napětí na výstupu mikrofону, kterým je zvuk snímán.



Jeho převod do digitální podoby je založen na vzorkování = měření a záznam velikosti napětí na mikrofónu v pravidelných časových intervalech. Na obrázku vpravo je vidět, že po digitalizaci je zdánlivě analogový průběh složen z dostatečně husté posloupnosti diskretních bodů.

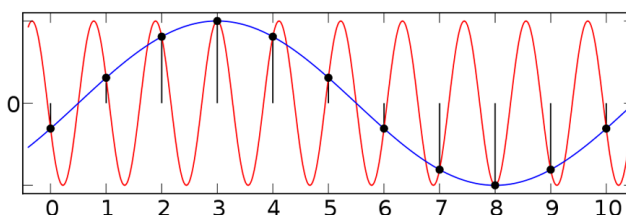


Aby nebyl záznam zkreslen, je nutné zvolit dostatečně vysokou vzorkovací frekvenci. Na obrázku vlevo je zřejmé, že při takto nízké vzorkovací frekvenci se z naměřených dat již nepodaří zrekonstruovat původní zvuk.

Zdroj: <http://fyzika.jreichl.com/main.article/view/1355-digitalizace-analogoveho-signalu>

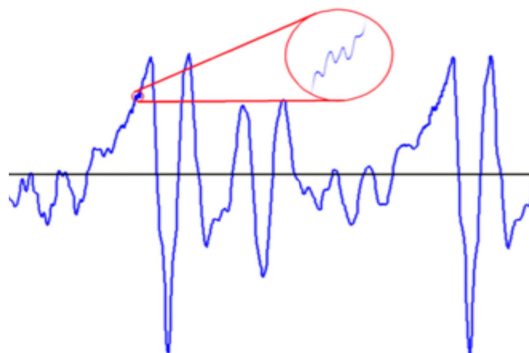
O velikosti vzorkovací frekvence hovoří Shannon-Kotelnikův teorém. Z něho plyne, že na každou periodu kódovaného signálu musí připadnout alespoň dva vzorky.

Při použití nižší vzorkovací frekvence hrozí znehodnocení původního (červeného) signálu (obr. vpravo). Výsledkem je zcela jiný (modrý) signál. Toto je také princip chyby zvané aliasing (viz dále).



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

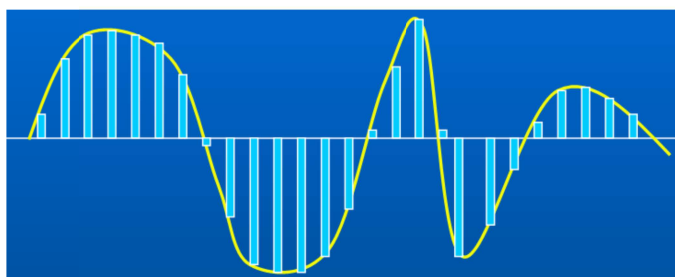
Výsledek však nemusí být tak fatální. Na obrázku vpravo je zachycen průběh zvuku, jehož frekvenci bychom mohli určit z přechmitů mezi kladnými a zápornými hodnotami. Při větším zvětšení však vidíme, že zvuk obsahuje i vyšší harmonické frekvence (červený detail). Vzorkovací frekvenci je nutné volit i s ohledem na tyto složky. Při volbě nízké vzorkovací frekvence, by byly tyto vyšší harmonické složky odstraněny.



Zdroj: <https://cs.wikipedia.org/wiki/Vzorkov%C3%A1n%C3%AD>

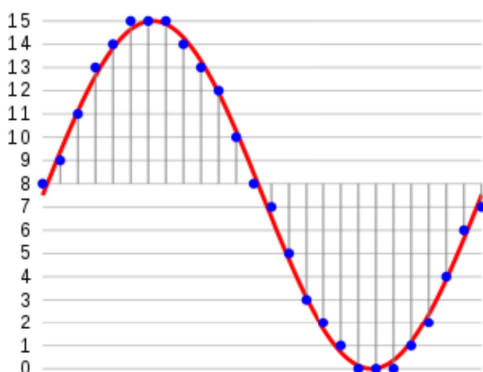
Protože je horní hranice frekvenčního rozsahu lidského ucha 20 kHz, je třeba při digitalizaci zvuku použít vzorkovací frekvenci alespoň 40 kHz (CD používá 44,1 kHz; DVD až 192 kHz).

Navzorkováním signálu jsme teprve na půlce cesty k digitalizaci zvuku. Dostali jsme posloupnost analogových čísel (libovolné velikosti), vyjadřujících akustickou výchylku v pravidelných časových intervalech. (Tato podoba signálu je označována jako PAM – pulzně amplitudová modulace).



Nyní je ještě třeba převést jednotlivé analogové vzorky do digitální podoby – každý vzorek vyjádříme pomocí binárního čísla. (Tato podoba signálu je nazývána PCM – pulzně kódovaná modulace).

Je zřejmé, že musí docházet k zaokrouhlování hodnot vzorků, protože v binárním kódu máme omezený počet hodnot (podle počtu použitých bitů). Na obrázku vlevo je použito čtyřbitové kódování (každý vzorek vyjádříme 4-bitovým číslem) – máme k dispozici 16 hodnot (2^4).



Zaokrouhlováním dochází ke zkreslení signálu – vzniká šum.

Záznam v CD kvalitě vychází z požadavku odstupu šumu od signálu 96 dB. Toho je dosaženo použitím 16-ti bitového kódování. K zakódování každého vzorku tedy máme k dispozici 65536 hladin (2^{16}).

První bit udává, zda se jedná o kladnou či zápornou hodnotu. Na každou polaritu pak zbývá 15 bitů - 32768 hladin (2^{15}).

Zdroj: https://cs.wikipedia.org/wiki/Pulzn%C4%9B_k%C3%B3dov%C3%A1n%C3%AD_modulace

Nyní máme posloupnost šestnáctibitových čísel představující surová digitální audio data. Ta je možné zapsat na CD v podobě audio CD, nebo na disk počítače – většinou v podobě souboru wav.

Možné chyby při digitalizaci audio signálu:

- přemodulovanost: musím zajistit, aby největší amplituda celého záznamu odpovídala maximálně 2^{15} - jakmile by byla větší dojde k „seřiznutí“ signálu
- granulační šum: ve velmi tichých pasážích (-60 dB) mám k dispozici jen 2^5 bitů, tj. 16 úrovní - na hlasitosti se objevuje jisté schodování – granulace
- aliasing: pokud se v akustickém signálu objeví frekvence f větší než vzorkovací f_{vz} , objeví se to při zpětném převodu jako frekvence $f-f_{vz}$ - ve zvuku se objevují pazvuky
(audio signál běžně obsahuje i vyšší frekvence než je schopný lidský sluch zachytit - původně neslyšitelná frekvence se po zakódování a dekódování (při nesplnění požadavku Shannon-Kotělnikovova teorému) projeví jako slyšitelný pazvuk. Při kódování je proto nutné odfiltrovat neslyšitelné frekvence)

Dekódování PCM

Dekódování digitálního signálu PCM je principiálně velmi snadné. Stačí v pravidelných časových intervalech posílat jednotlivé vzorky do D/A převodníku – výsledkem je původní časový průběh analogového signálu.

(následuje ještě vyhlazení velmi ostrým filtrem typu dolní propust na 20 kHz – nejnáročnější část prvních CD přehrávačů)

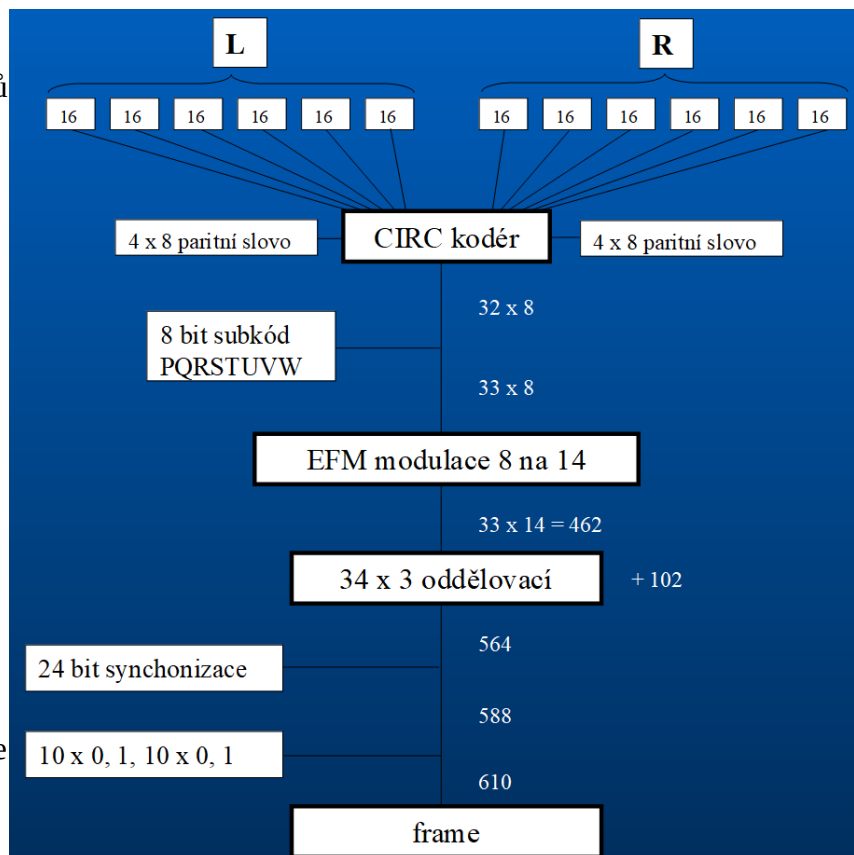
Záznam na CD

Šestnáctibitové PCM vzorky není zdaleka to jediné, co je zapsáno na CD. Vzorky jsou uspořádány do jakýchsi balíčků (rámců), aby bylo zajištěno jejich správné přečtení¹; je doplněn opravný kód a přídatné textové a obrazové informace. Struktura jednoho rámce je zachycena na následujícím obrázku.

Každý rámec obsahuje 12 vzorků (6 pro levý a 6 pro pravý kanál).

Poté jsou ke vzorkům přidány bity opravného kódu pro opravy špatně přečtených dat (čtyři 8-bitová slova pro každý kanál). Je použito Reed-Solomonovo kódování – postupné kontrolní součty datových bitů, na jejichž základě lze doplnit chybějící nebo opravit špatně přečtené bity.

Dále může být do rámce přidán subkód – název skladby, text, obrázek... Do každého rámce je přidáno jedno 8-bitové slovo. Z jednu sekundu je přehráno více než 7000 rámců, což je dostatek místa pro doplňkové informace.



Pro správné čtení dat je třeba, aby vypálené „důlky“ v povrchu CD nebyly ani příliš dlouhé ani příliš krátké. Z tohoto důvodu jsou data překódována z 8 na 14-bitová slova (modulace EFM). Tím je zajištěno, aby mezi dvěma 1 byly minimálně dvě a maximálně deset 0.

Následuje přidání několika bitů pro oddělení jednotlivých vzorků.

Pak je přidáno 24 bitů pro regulaci otáček CD.

Nakonec jsou přidány bity oddělující od sebe jednotlivé rámce.

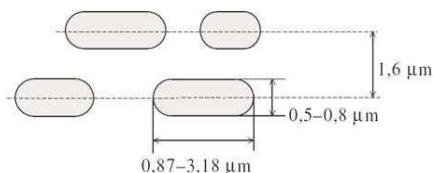
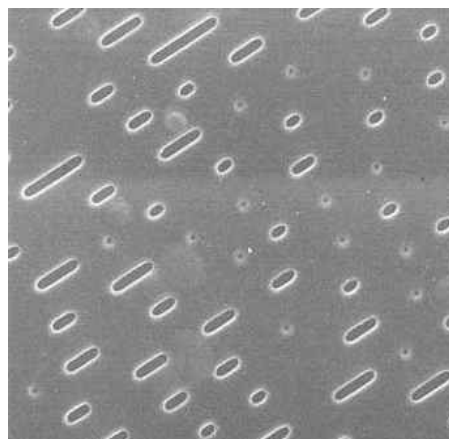
Jak je vidět - z původních 192 bitů surových audio dat (dvanáct 16-bitových vzorků), rámec obsahuje 610 bitů výsledných dat vypalovaných na CD.

Pozn: Reed-Solomonův opravný kód umožňuje dokonalou resuscitaci až 2500 po sobě jdoucích bitů (2,4 mm stopy).

Audio CD však umožňuje i opravu delších poškozených úseků – až 12000 poškozených bitů (8,5 mm stopy) je možné interpolačně dopočítat – není to již dokonalá resuscitace, ale na poslechu je toto „odhadnutí dat“ neznatelné.

¹ Představte si, že by vlivem škrábance na CD byl při čtení vynechán jeden bit a nahrazen následujícím bitem z dalšího vzorku. Všechny další vzorky už by byly přečteny špatně.

Posloupnost jedniček a nul je na CD vypálen do spirály podobně jako záznam na gramofonové desce. Začíná ale uprostřed a postupuje k okraji. Délka spirály je zhruba 6 km. Detail záznamu je zachycen na obrázku vpravo. Hustota záznamu je uprostřed i na okraji stejná, proto se úhlová rychlost otáčení CD musí od středu k okraji zmenšovat. Obvodová rychlost otáčení (v místě čtení) je zhruba 1,2 až 1,4 ms^{-1} (1cm záznamu je méně než 0,01 s). Rozměry vypálených důlků tzv. pitů jsou zachyceny na obrázku dole.



T3		833 nm	1001
T4		1111 nm	10001
T5		1388 nm	100001
T6		1666 nm	1000001
T7		1944 nm	10000001
T8		2221 nm	100000001
T9		2499 nm	1000000001
T10		2777 nm	10000000001
T11		3054 nm	100000000001

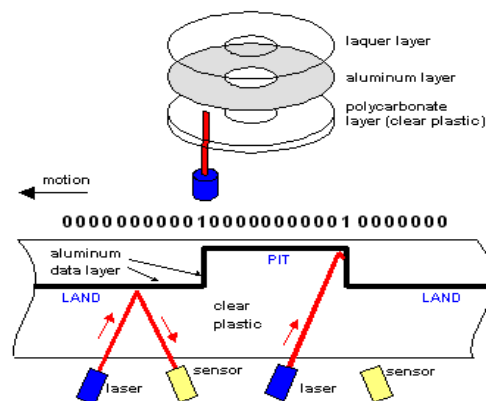
Note: this is at 1.2 m/sec,
with a channel bit size of 277.662 nm

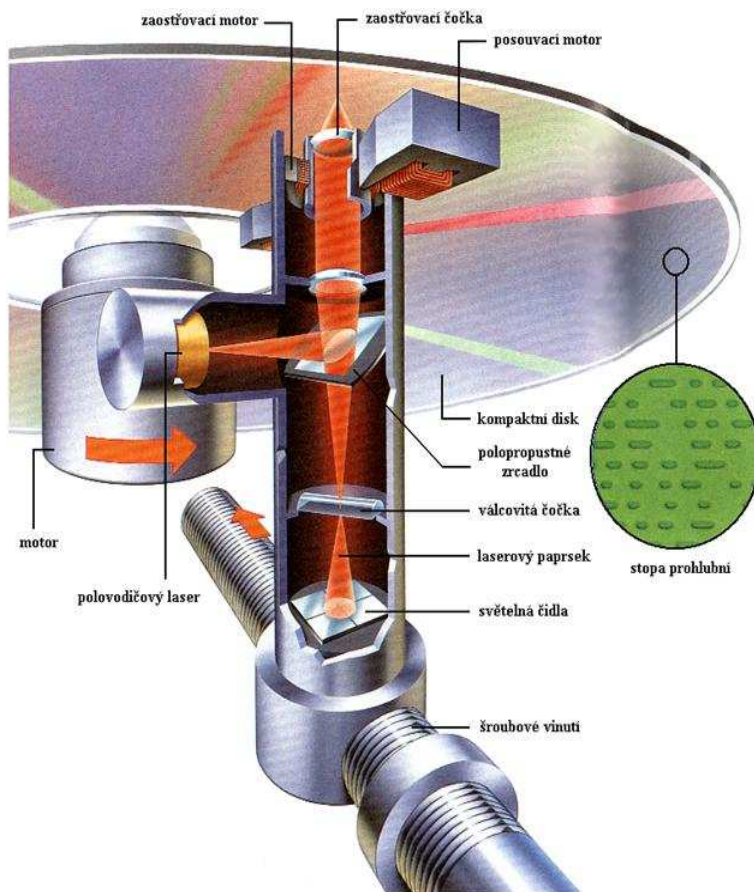
Aby laser snímající záznam neztratil stopu, je vzdálenost mezi pity (i délka pitů) omezena – max. 10 znaků. Kvůli omezení chybovosti čtení však nesmí být pity ani příliš krátké (stejně tak i mezery mezi pity) – minimální rozměr pitu nebo mezery musí odpovídat dvěma znakům. To je zajištěno EFM modulací z 8 na 14 bitů (viz. předchozí strana).

Čtení dat funguje na principu různé odrazivosti povrchu CD. Dalo by se očekávat, že vypálený pit představuje logickou 1, mezera logickou 0 – tak to ale není.

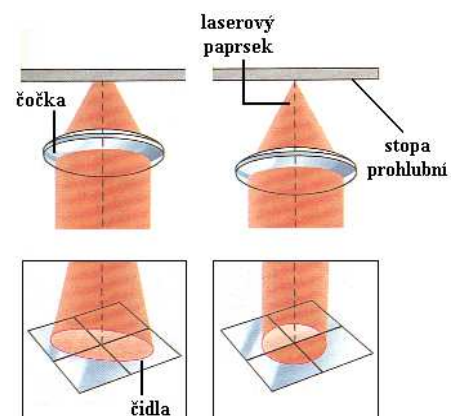
Úroveň 1 představuje přechod mezi pitem a mezerou vše ostatní znamená 0 (obr. vpravo).

From Computer Desktop Encyclopedia
© 1998 The Computer Language Co. Inc.





Ke snímání záznamu je používám polovodičový laser. Jeho paprsek dopadá na polopropustné zrcátko a je odražen k povrchu CD. Optikou je zaostřen na průměr $1\ \mu\text{m}$. Od povrchu CD se paprsek odrazí a postupuje zpět skrz polopropustné zrcátko do vyhodnocovacího čidla. Čidlo kromě snímaných dat vyhodnocuje i správné zaostření laseru. Čidlo je složeno ze čtyřech částí – laser je zaostřen tak, aby byly všechny části osvětleny stejnoměrně.



V dnešní době již CéDéčka pomalu mizí v propadlišti dějin. K poslechu hudby je používán především komprimovaný formát MP3. Jeho obrovskou výhodou je zhruba desetinásobně menší velikost souboru. Velikou nevýhodou je však skutečnost, že používá ztrátovou kompresi – mnoho dat je za účelem zmenšení souboru vypuštěno, což má mnohdy velký vliv na kvalitu poslechu – nejedná se již o plnohodnotnou nahrávku. Z tohoto důvodu je stále hojně využíván formát wav, který obsahuje plnohodnotná data. Využívá stejný princip kódování jako audio CD a má i velmi podobnou strukturu.

V současnosti je stále více používám formát FLAC. Využívá tzv. bezztrátovou kompresi dat – data zůstávají plnohodnotná, pouze je ušetřeno místo jejich vhodným přeuspořádáním (lze ho přirovnat k datovým kompresním metodám zip, rar). Tento formát nepřináší příliš velkou úsporu místa (FLAC zabírá zhruba 60% prostoru původního souboru wav), přináší však stále plnohodnotný zvuk. Využívá se proto v internetové distribuci hudby a zvuku. Díky jeho relativně rychlému zpracování je využíván také v oblasti streamovaných médií.