

# Syntéza audio signálů

## Aditivní syntéza symfonického orchestru a akordeonu

*Bedřich Smetana - Vltava*

*3 oktávy durové stupnice*

*Johann C. F. Fischer - Preludium a fuga G dur*

*Bedřich Smetana - Jiřinková polka*

*Václav Král*

2016

<b>Teoretická část</b>	<b>2</b>
Midi toolbox	2
Aditivní syntéza	2
Stereofonní zvuk	2
<b>Praktická část</b>	<b>3</b>
Odečtení amplitud harmonických frekvencí	3
Tvorba obálky	4
Syntéza audio signálu	7
<b>Reference</b>	<b>9</b>

## Teoretická část

### Midi toolbox

Midi je standard elektronické komunikace mezi hudebními nástroji a počítačem. Přenos je sériový. Informace o hraném tónu jsou přenášeny jako koeficienty - např. číslo nástroje, frekvence, délka tónu, síla zvuku, zda jde o legato, atd.

Midi také umožňuje uložení notového zápisu a následnou syntézu. V tomto projektu je použit upravený Midi toolbox pro Matlab (autor Ken Schutt). Tento toolbox umí mimo jiné načíst notový zápis ve formátu midi a syntetizovat z něj hudební nahrávku. Toto dovolí soustředit se v tomto projektu čistě na syntézu zvuku jednotlivých hudebních nástrojů, aniž by bylo třeba zabývat se vytvářením notových zápisů.

### Aditivní syntéza

Při aditivní syntéze vezmeme zvuk nástroje a převedeme ho do spektrální oblasti. Takto můžeme získat rozložení harmonických složek a jejich velikosti. Budeme-li uvažovat harmonický signál, můžeme převádět signály z časové oblasti do spektrální a zpět beze ztrát. Použitím dat získaných ze spektra původního signálu můžeme vytvořit nový signál tak, že sečteme jednotlivé vážené násobky požadované frekvence. Získáme tak zvukový signál se stejnou barvou, jakou by měl originální zvuk. Vjem zvuku ovlivňuje i tvar obálky signálu, kterým je také potřeba se zabývat.

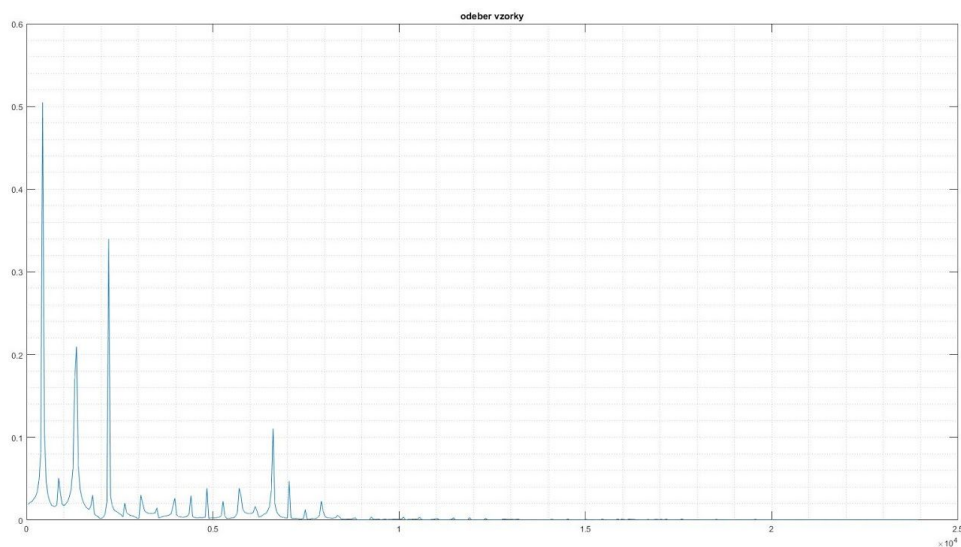
### Stereofonní zvuk

Mezi hlavní způsoby určení pozice zdroje zvuku patří meziušní rozdíly v hlasitosti (ILD) a čase příjmu (ITD). Pokud máme monofonní zvuk, můžeme se znalostí polohy posluchače a nástroje simulovat stereofonní zvuk. Časové zpoždění příjmu signálu na jednotlivých uších se dá vypočítat přes Pythagorovu větu, zatímco rozdíly v hlasitosti lze odhadnout.

## Praktická část

### Odečtení amplitud harmonických frekvencí

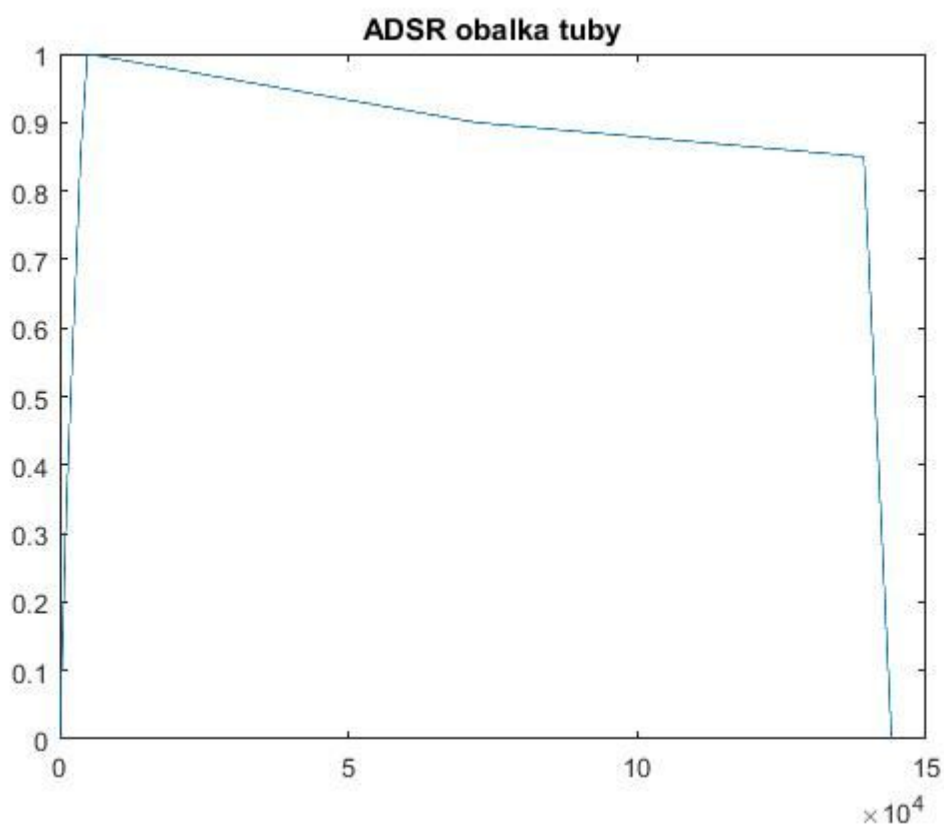
Použil jsem zvukové nahrávky různých nástrojů volně dostupné na stránkách freesound.org pro získání barvy zvuku jednotlivých nástrojů symfonického orchestru a nahrávku vlastního akordeonu. Po načtení do Matlabu ve skriptu *analyza\_main* byl signál oříznut na několik period, kde je předpokládán rozvoj všech harmonických, a předán funkci *analyza*. Tato funkce provede rychlou Fourierovu transformaci signálu do spektra s oříznutím. Skript poté pokračuje grafickým vybráním velikostí jednotlivých harmonických z grafu. Pro každý nástroj je třeba zvolit jiný počet harmonických. Po vybrání všech hodnot se velikosti harmonických *ampl* uloží do souboru označeného číslem nástroje ve složce *data*.



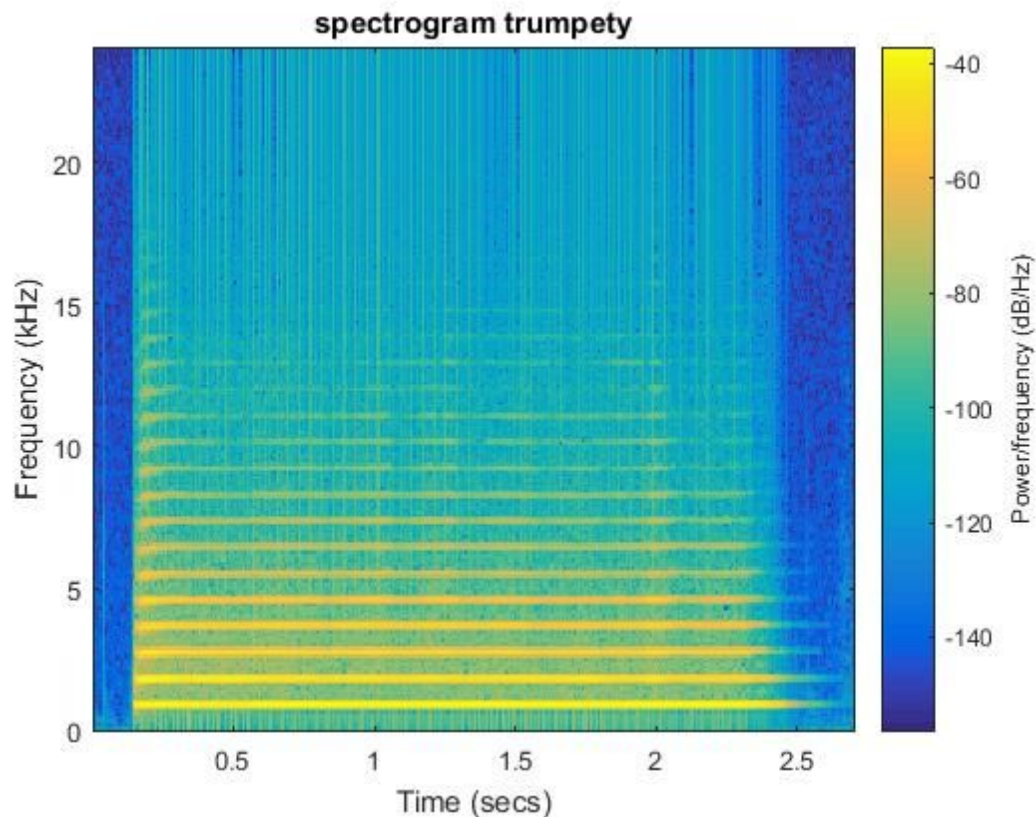
## Tvorba obálky

Pro napodobení zvuku je důležitá i jeho obálka. Obálka perkusních signálů se dá popsat exponenciálou, u neperkusních signálů se používá obálka typu ADSR.

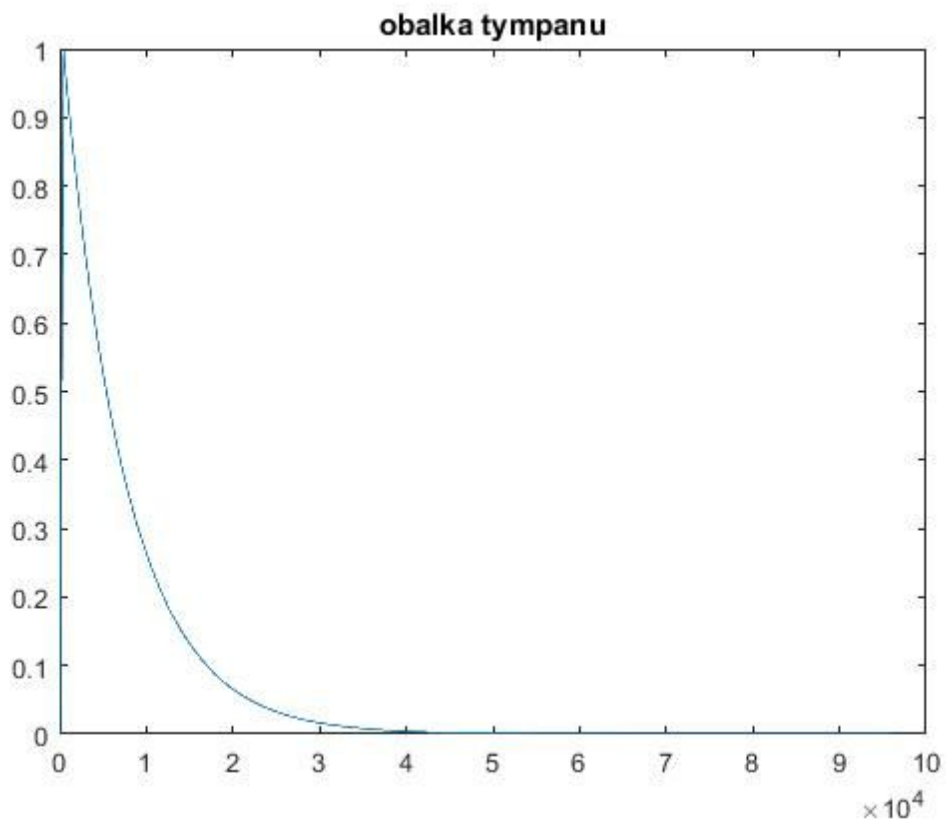
U obálky typu ADSR je potřeba zvolit významné body v obálce, kterými se zajistí tvar obálky podobný původnímu signálu. Poté je lineární interpolací dopočten obálkový signál, kterým se vynásobí signál generovaný pro získání správné barvy zvuku. Délka A, D a R se typicky nemění v závislosti na délce signálu (popisují náběh zvuku, resp. dozívání), zatímco délka S (držení) přímo odpovídá délce tónu. Z toho také vyplývá, že není možné signál nekonečně zkracovat. Proto jsou všechny skripty pro generování tónů opatřeny minimální délkou tónu pro daný hudební nástroj.



U neperkusních hudebních nástrojů je časté, že jednotlivé harmonické nezačínají hrát současně, ale čím vyšší je harmonická, tím déle začne znít a tím dříve skončí znít. Pro tyto nástroje byla nejdříve odečtena obálka původního signálu, resp. základní harmonické, z časové reprezentace signálu. Poté je ze spektrogramu odhadnuto zpoždění nástupu dalších harmonických, resp. jejich poklesu. Pro jednoduchost se používá násobení hodnot náběhu první harmonické konstantou závislou na čísle harmonické. Doba náběhu a doznívání se zkracuje do té doby, dokud se jejich součet nerovná délce signálu, poté mají zbývající vyšší harmonické všechny stejnou obálku.

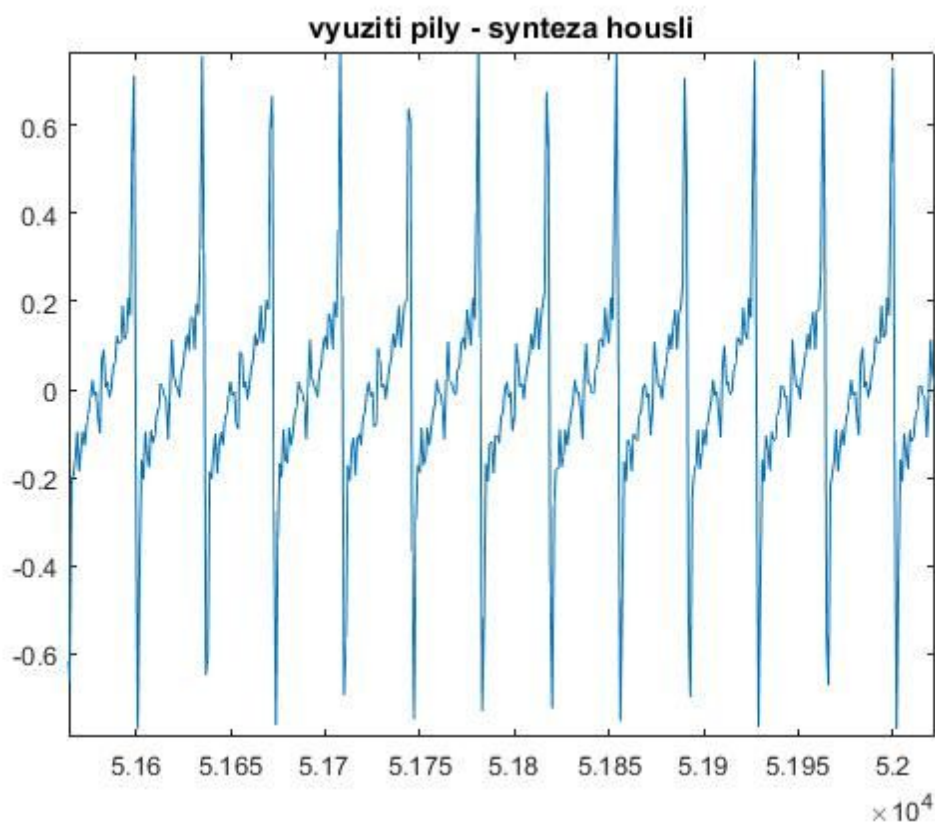


Pro perkusní nástroje je využita kombinace ADSR a exponenciální obálky. Pro náběh signálu je použita obálka typu ADSR s konstantní délkou jako pro neperkusní signály. Exponenciální obálka je použita pro pokles, resp. doznění. Jsou použity 2 exponenciály, aby bylo možné modelovat rychlejší spád obálky pro pokles signálu a zároveň docílit určité doby doznívání. Parametry  $\tau$  jsou voleny tak, aby obálka co nejdříve napodobovala tvar původního signálu. Stejně jako u obálky typu ADSR je třeba i zvuků perkusních nástrojů dodržet určitou délku trvání signálu, aby mohl i např. krátký požadovaný signál správně doznít (např. harfa doznívá velmi dlouhou dobu).



## Syntéza audio signálu

Ke generování zvuku je použit jak v aditivní syntéze typický sinusový signál, tak i signál obdélníkový a pilový. Tyto signály jsou generovány funkcí *rect.m*, resp. *saw.m*. Z matematických popisů tvorby těchto signálů (nekonečná suma sinů s rostoucí frekvencí) je vytvořen skript, který zajistí omezení aliasingu. Obdélníkové signály zajistí zajímavější zvuk pro dechové nástroje, zatímco pilovitý signál lépe popisuje smyčcové nástroje. Protože takto generované zvuky obsahují harmonické až do poloviny vzorkovacího kmitočtu, jsou omezeny filtrem typu dolní propust na 15 kHz. Nejedná se navíc o přesné signály typu obdélník/pila, protože amplitudy vyšších harmonických jsou slabší než podle předpisu. Toto zajistí bohatší barvu zvuku, aniž by byl zvuk moc agresivní.



Podívejme se nyní na syntézu obecného nástroje. Ve skriptu *synth.m* je pro každý nástroj nejdříve vyhodnocena délka signálu a případně zajištěna minimální délka pro správný zvuk. Pokud je splněna podmínka pro frekvenční rozsah daného nástroje, pokračuje se syntézou zvuku, jinak je výstupem prázdný signál. Následuje cyklus s výpočtem signálu pro harmonické složky. Je vypočteno případné tremolo (intenzita závisí na harmonické, frekvence podle zdrojového zvuku). Následně je vypočtena obálka pro danou harmonickou. Pokračuje se generováním harmonické sinem, popřípadě pilou nebo obdélníkovým signálem. Takto generovaný signál je vynásoben obálkou a tremolem. Dále je třeba zajistit nulovou střední složku, normalizaci signálu a vynásobení chtěnou amplitudou. Každá jedna syntetizovaná nota se předá nadřazené funkci, kde je přičtena k výstupnímu signálu. Před uložením se výstupní signál omezí pozvolnou dolní propustí na 15 kHz pro omezení rušivých složek způsobených použitím pilových a obdélníkových signálů.

Pro syntézu houslí a čela je využit pilovitý průběh zajištěný funkcí *saw.m*. Navíc je použito vibrato, jehož intenzita je závislá na čísle harmonické. Při syntéze dechových nástrojů se



využívá sinusový nebo obdélníkový průběh tak, aby zněl daný nástroj přirozeně. Žestový soubor obsahuje trubku, tubu a lesní roh, kde každý nástroj má omezený frekvenční rozsah podle svých fyzikálních možností. U tympánů je základní zvuk generovaný aditivní syntézou (jako u zbytku), navíc je použita i subtraktivní syntéza pro lepší napodobení barvy zvuku. Jelikož se jedná o perkusní nástroj, je použita ADSR+exponenciální obálka. Zvuk harfy a drnknutí do houslí je syntetizován pomocí funkce *saw\_piz.m*, resp. *saw\_harp.m*, které zajišťují pilovitý průběh a zároveň vylepšenou obálku ADSR+exponenciála.

Funkce *synth.m* umožňuje i tvorbu stereofonní nahrávky. Aby se nemusely upravovat vnitřní funkce midi toolboxu, stereofonní zvuk se syntetizuje manuálně, kdy ve funkci *synth.m* lze zvolit syntézu mono/levý kanál/pravý kanál. Při volbě mono proběhne základní syntéza, kdy amplituda daného tónu je řízena pouze amplitudou předepsanou v notovém zápisu. Pokud zvolíme syntézu levého nebo pravého kanálu, je situace složitější. Byla zvolena pozice posluchače 10 metrů před orchestrem a velikost orchestru je půlkruh o poloměru také 10 metrů. Ve funkci *itd\_stereo.m* je pro jednotlivé nástroje vypočteno ITD, které je následně ve funkci *synth.m* přepočteno na počet vzorků zpoždění. Následně je kanál pro stranu, kde je nástroj blíže, opatřen nulami na konci signálu a druhý kanál nulami na začátku signálu. Počet nul odpovídá zpoždění ITD. Následně je normalizovaný signál pro každý kanál přenásobený amplitudou určenou v notovém zápisu a odhadnutou amplitudou pro ILD.



Nakonec se jednotlivě vypočtené nahrávky spojí do stereofonní nahrávky. Navíc je možné do nahrávky přidat konvoluční reverb. Pro orchestr přinesla nejlepší výsledek impulsní odezva Vídeňského sálu Musikverein, pro menší počet nástrojů je tento prostor ovšem moc veliký, ostatní skladby jsou tedy opatřeny reverbem "velké dobře tlumené místnosti". Jako poslední úprava je přičten šum s odstupem 70 dB, který by měl simulovat přítomnost posluchačů v sálu.

Hudební stupnice je syntetizována ve skriptu *stupnice.m*. Je zde vytvořen vektor frekvencí a vektor hudebních nástrojů. Následně je volána přímo funkce *synth.m* a jednotlivé zvuky jsou spojeny do nahrávky *stupnice\_stereo.wav*.

Na stejném principu funguje skript *fuga.m*, kde je syntetizována nahrávka *fuga.wav*. Použil jsem notový zápis z mé loňské semestrální práce, ale syntéza akordeonu je implementována nově, a to stejným způsobem jako ostatní nástroje v této práci. Pro využití midi toolboxu pro syntézu akordeonu byla navíc syntetizována skladba Jiřinková polka od Bedřicha Smetany, kterou jsem také na akordeon hrál v ZUŠ.

## Reference

<http://sami.feld.cvut.cz/syn/%C3%9A%20vod.pdf> (Midi)

<http://kenschutte.com/midi> (Midi toolbox)

<https://freesound.org> (volně dostupné nahrávky hudebních nástrojů)

<http://www.phy.mtu.edu/~suits/notefreqs.html> (frekvencní rozsah hudebních nástrojů)

<http://legacy.spa.aalto.fi/projects/poririrs/> (impulsní odezva finské koncertní síně)

<http://www.voxengo.com/impulses/> (balíček odezev různých prostorů)