

AKADEMIE MÚZICKÝCH UMĚNÍ V PRAZE
HUDEBNÍ A TANEČNÍ FAKULTA

Hudební umění
Zvuková tvorba

BAKALÁŘSKÁ PRÁCE

**VÍCESAMPLOVÁ GRANULÁRNÍ SYNTÉZA A
JEJÍ VÝSTUP**

David Pevný

Vedoucí práce : Ing. Milan Guštar, Ph.D.

Oponent práce: doc. MgA. Ing. Ondřej Urban, Ph.D.

Datum obhajoby:

Přidělovaný akademický titul: BcA.

Praha, 2022

ACADEMY OF PERFORMING ARTS IN PRAGUE

MUSIC AND DANCE FACULTY

Art of Music

Sound production

BACHELOR THESIS

**MULTI-SAMPLE GRANULAR SYNTHESIS
AND ITS OUTPUT**

David Pevný

Thesis supervisor : Ing. Milan Guštar, Ph.D.

Opponent: doc. MgA. Ing. Ondřej Urban, Ph.D.

Date of defence:

Academic degree conferred : BcA.

Prague, 2022

Prohlášení

Prohlašuji, že jsem bakalářskou práci na téma „Vícesamplová granulární syntéza a její výstup“ 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 jakákoliv nakládání s nimi je možné pouze na základě licenční smlouvy, tj. souhlasu autora a AMU v Praze.

Poděkování

Rád bych poděkoval Ing. Milanu Guštarovi, PhD. nejen za zajímavý nápad k zpracování, ale také za vstřícné a trpělivé konzultace.

Abstrakt

Cílem této práce je nejen důkladné teoretické zavedení a následná praktická realizace specifického a dosud jen krajně prozkoumaného problému v rámci granulárního způsobu zvukové syntézy - granulace s použitím více rozdílných vstupních samplů (zvukových souborů) - za účelem jejich kombinace v novou zvukovou barvu, ale také detailnější shrnutí granulárního přístupu ve zvukové tvorbě. Práce se zaměřuje jak na obecnou problematiku zvukových mikrostruktur, historii jejich využití a na výstup kreativní práce s nimi, tak na následnou praktickou aplikaci těchto poznatků na vícesamplový vstup.

Součástí práce je analýza výstupu vícesamplové granulace s přihlédnutím k jejímu kreativnímu využití.

Klíčová slova: zvuková syntéza, granulární syntéza, hudební programování

Abstract

The main goal of this thesis is not only a thorough theoretical introduction and subsequent practical implementation of a specific and so far only hardly described phenomenon in the scope of granular sound synthesis - granulation using multiple different input samples (sound files) for the purpose of their combination into a new sound colour, but also a more detailed summary of granular approach in sound production. This thesis focuses not only on general problematics of sound microstructures, history of their use and the output of creative work with them, but also on the subsequent practical application of this knowledge on a multi-sample input.

Part of the thesis consists of an analysis of the output of multi-sample granulation, with an focus on its creative implementation.

Keywords: sound synthesis, granular synthesis, music programming

Úvod	1
1 Obecné parametry zvuku	3
1.1 Objektivní fyzikální vlastnosti zvukového signálu	3
1.1.1 Amplituda	3
1.1.2 Akustický tlak	3
1.1.3 Frekvence	4
1.1.4 Frekvenční spektrum	4
1.1.5 Fáze	4
1.1.6 Časový průběh	5
1.2 Subjektivní vlastnosti zvukového signálu.....	5
1.2.1 Hlasitost	5
1.2.2 Výška tónu	6
1.2.3 Barva zvuku	6
2 Dělení zvukových signálů	8
2.1 Dělení signálů dle výškové rozpoznatelnosti	8
2.2 Dělení signálů dle frekvenčního průběhu	9
2.3 Dělení signálů dle trvání	10
2.3.1 Klasická měřítka trvání zvuku	10
2.3.2 Mikrozvuk	11
2.3.2.1 Psychoakustická specifika mikrozvuků	11
2.3.3 Samplové a subsamplové měřítko zvuku	11
3 Digitální reprezentace zvukového signálu	13
3.1 Pulzně kódová modulace.....	13
3.1.1 Vzorkování	13
3.1.2 Kvantování	14
3.2. Formát Waveform	15
4 Zvuková syntéza	16
4.1. Typy zvukové syntézy	16
5 Granulární syntéza: Charakteristika a historie	18
5.1 Historie mikrozvuku a granulace.....	18
6 Praktické aplikace granulární syntézy	22
6.1. Časově synchronní syntéza	22
6.2 Časově asynchronní syntéza	23
6.3. Další parametry granulární syntézy.....	24
6.3.1 Délka zrna	24

6.3.2 Obálka zrna	24
6.3.3 Nasycení granulárního proudu	24
6.3.4 Obsah zrna	25
6.3.5 Prostorové rozmístění	25
7 Zavedení vícesamplového vstupu do granulární syntézy	26
7.1 Současný stav poznání	26
7.2 Vícesamplovost jako způsob barevné syntézy	27
7.2.1 Volba vhodných parametrů pro praktickou realizaci	27
8 Popis vývoje vícesamplového granulárního syntezátoru v jazyce Python 3	31
8.1 Volba vhodného prostředí	31
8.2. Volba algoritmického přístupu, nástrojů a parametrů	31
8.2.1. Volba knihoven	32
8.2.2 Volba konkrétních parametrů digitálního audia	33
8.3. Popis kódu vlastního vícesamplového granulárního syntezátoru	34
8.3.1 Příprava souborů ke granulaci	35
8.3.2 Granulační proces: metoda granulator_engine	38
8.3.3 Zápis do výstupního souboru	40
9 Analýza výstupních dat programu	41
9.1 Čistě sinusový vstupní signál	41
9.1.1 Vliv délky zrna a prolínání na výstup	42
9.2. Vstup s komplexním spektrem	46
9.2.1 Obecný vliv granulace na tónový charakter výstupu	46
9.2.2 Vliv obálky vstupních signálů na granulární výstup	46
9.2.2.1 Kombinace signálů perkusního charakteru	47
9.2.3 Vliv poměrů délek zrn na barevný charakter výstupu	48
9.2.3.1 Experiment: změna poměru signálu	49
10 Možné druhy kreativní práce s vícesamplovou granulací	50
10.1 Vícesamplová granulace za účelem tvorby samplů	50
10.2 Vícesamplová granulace jako alternativa k mixu	51
10.3 Randomizace granulovaného materiálu	51
11 Závěr	53

Seznam zkratek

IFF - Interchange File Format

AIFF - Audio Interchange File Format

WAV - Waveform formát

FT - Fourierova transformace

FFT - Rychlá FT

PCM - Pulse-code modulation

GRM - Groupe de Recherches Musicale

WDR - Westdeutsche Rundfunk

DAW - Digital Audio Workstation

EBU - European Broadcasting Union

MAX/MSP - Prostředí pro grafické programování zvukových modulů

PD - Pure Data, prostředí pro grafické programování zvukových modulů

IDE - Integrated Development Environment

DBFS - Decibel Full Scale

Úvod

Tato práce se zabývá především obecnou problematikou tvorby a využití granulárního syntezátoru s důrazem na metodu vícesamplové granulace. Rozebírá nejprve základní zvukové pojmy a metody jejich zavedení a zpracování v digitální rovině, nutné pro uchopení problému. Dále se věnuje granulární syntéze obecně, jak z hlediska historického vývoje, tak z hlediska praktické aplikace. Práce se pak věnuje především zavedení problému vícesamplovosti do granulární syntézy, jak v teorii, tak na praktickém příkladě multisamplového syntezátoru vyvinutého pomocí jazyka Python 3. Výstup syntézy za pomoci více samplů je pak následně analyzován. Na základě této analýzy jsou pak na konci práce navržena několikera kompoziční východiska.

V prvních dvou kapitolách považuji za nezbytné uvést alespoň stručný přehled objektivních vlastností zvukového signálu, neb bez důkladného porozumění těmto pojmům není možné pochopit složitější aspekty práce se zvukem, kterými se tato práce zabývá.

Třetí kapitola se zabývá digitální (číslicovou) podobou zvuku, aby položila teoretický základ pro algoritmické procesy, které se zvukem v digitální podobě pracují.

Cílem čtvrté kapitoly je stručně popsat základní parametry zvukové syntézy a její typy, aby na základě tohoto poznání mohla být hlouběji rozebrána syntéza granulární.

V páté a šesté kapitole se práce věnuje obecné teorii granulární syntézy. Pátá kapitola probere stručně její krátkou historii a ukáže rozdílné přístupy skladatelů s tímto fenoménem pracujících.

Šestá kapitola pak pojedná o způsobech naplnění obecného rámce granulární syntézy v praxi. Probere problémy a úskalí, které z jednotlivých přístupů vycházejí a navrhne jejich řešení. S těmito problémy se bez výjimky potýká každý, kdo s granulární syntézou pracuje. Uvede také jednotlivé způsoby práce se zrny.

Cílem sedmé kapitoly je uvést vlastní přístup k obecnému rámci granulární syntézy, spočívající ve využití více vstupních dat ve formě předem nahraných samplů (elektronických i akustických) za účelem jejich kombinace (syntézy) v novou zvukovou barvu. Stručně probere podobné druhy implementace napříč (velmi krátkou) historií a pokusí se teoreticky zavést přístup vlastní, detailněji představit jeho specifika a nastínit možné formy praktického výstupu.

Osmá kapitola si klade za cíl popsat proces tvorby programu pro vícesamplovou granulární syntézu co nejstručněji tak, aby tento proces mohl být znalým čtenářem hypoteticky znovu rekonstruován. Popíše jak volbu praktického způsobu, jak k cíli

dojít, tak volbu konkrétních parametrů a neposledně popíše samotný proces programování spolu s praktickými problémy z něj vyvstalými.

Devátá kapitola analyzuje výstup vícesamplové granulace s různými druhy vstupních materiálů. K analýze používá především metody analýzy frekvenčního spektra a časového průběhu, prezentuje také subjektivní dojmy při poslechu výstupních souborů a jejich vazby na konkrétní vstupní soubory a parametry.

Na závěr práce bych rád jen rámcově nastínil několik možných kreativních přístupů, kterými se případný uživatel vícesamplového granulárního syntezátoru může vydat.

1 Obecné parametry zvuku

Definice zvuku se v literatuře nachází nepřehledné množství, nejobecněji lze však tento pojem uchopit jako důsledek kmitavého pohybu hmoty zachycený lidským uchem. Jedná se o pohyb tzv. akustických částic - tedy myšleného, homogenního souboru molekul pružného prostředí - které se v daném prostředí střídavě zhušťují a zředňují okolo rovnovážné polohy. Tím vzniká rychlá výchylka relativně statického barometrického tlaku, která se označuje jako tlak akustický. Tato výchylka se v pružném prostředí šíří dále ve směru daném zdrojem kmitání rychlostí zvuku, která je však odlišná od tzv. akustické rychlosti, jež vyjadřuje okamžitou rychlost akustických částic.

Lidské ucho vnímá však pouze signály nacházející se v tzv. sluchovém poli - poli ohraničujícím rozsah zvuků, které je lidské ucho schopno vnímat - ve frekvenční doméně je ohraničeno frekvenčním rozsahem sluchu (16 Hz - 20 kHz), v doméně hlasitosti pak zespoda prahem slyšení a zehora prahem bolesti. V časové doméně je pak sluchový vjem omezen tzv. mezním časovým prahem¹, tedy minimálním časem, kdy ucho stihne postřehnout změnu zvukového signálu. Pod tímto prahem není ucho schopné rozpoznat barvu, ani výšku zvuku.

Pokud se signál nachází ve sluchovém poli a zachytí ho lidské ucho, dojde uvnitř něj k rozkmitání vláskových buněk, které odešlou nervové vzruchy do mozku. Ten kmitání následně rozpozná jako slyšitelný zvuk.

1.1 Objektivní fyzikální vlastnosti zvukového signálu

Každý zvukový signál je vlněním a časový průběh každého vlnění se dá popsat souborem několika úzce souvisejících fyzikálních veličin:

1.1.1 Amplituda

Amplituda signálu odpovídá maximální výchylce signálu. V akustice se amplituda většinou značí A . U zvukového signálu se její hodnota dá měřit v dílčích jednotkách metru.

1.1.2 Akustický tlak

Akustický tlak je způsoben kmitavým pohybem zvuku. Jedná se o malou lokální změnu (barometrického) tlaku v místě působení zvuku

¹SYROVÝ, Václav. Hudební akustika. 2., dopl. vyd. V Praze: Akademie múzických umění, 2008. Akustická knihovna Zvukového studia Hudební fakulty AMU. ISBN 978-80-7331-127-8.

Pro vyjádření hladiny akustického tlaku a intenzity používáme logaritmickou jednotku decibel (dB). Hladina akustického tlaku je vyjádřena vztahem $L = 10 \log \frac{P}{P_{ref}}$. Na

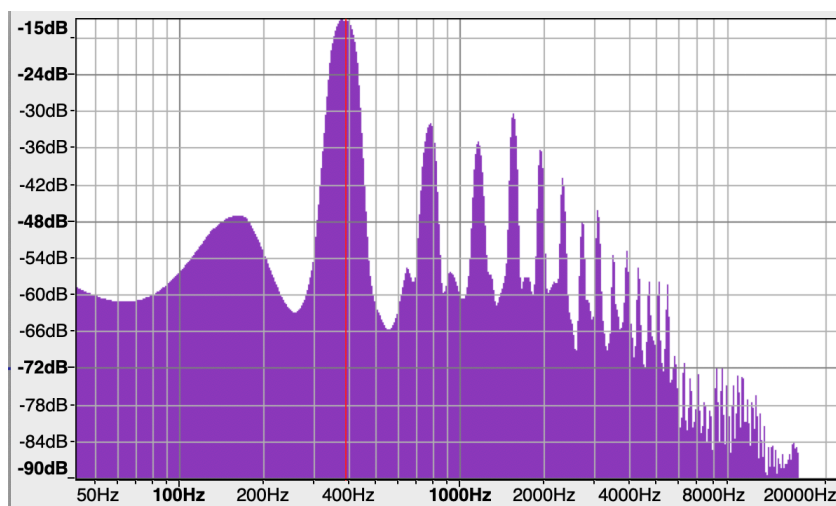
hladině akustického tlaku pak závisí subjektivní vjem hlasitosti, kde $p_{ref} = 2 \cdot 10^{-5}$ značí referenční hodnotu akustického tlaku vztahenou k sluchovému prahu hlasitosti.

1.1.3 Frekvence

Frekvence je veličina vlastní periodickým signálům - vyjadřuje počet kmitů signálu za vteřinu. Tato jednotka se značí písmenem f a můžeme jí vyjádřit pomocí vzorce $f = T^{-1}$, (kde T označuje délku jedné periody signálu). Její jednotkou je Hertz (Hz).

1.1.4 Frekvenční spektrum

Frekvenčním spektrem se nazývá souhrn všech frekvenčních složek, ze kterých se zvukový signál skládá.



Obr. 1.1.3.1: Spektrogram basklarinetového tónu a1

1.1.5 Fáze

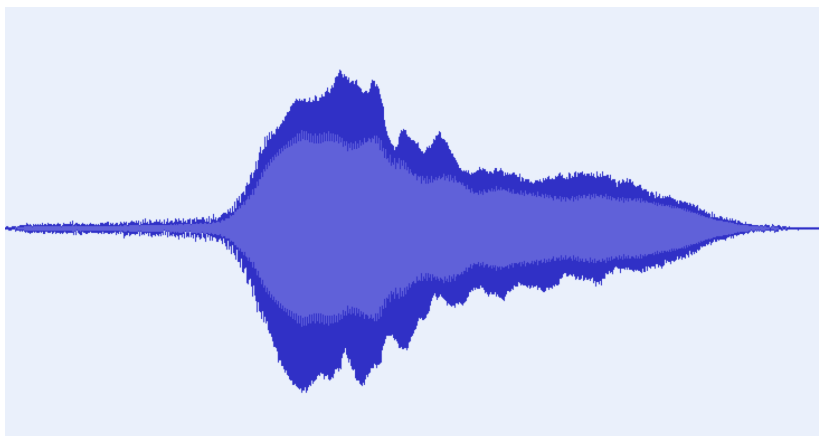
Fáze signálu, značená φ , je bezrozměrná veličina označující posun zvukového signálu oproti tzv. fázi počáteční φ_0 . Fáze se vzhledem ke kruhovému charakteru periodických kmitavých pohybů vyjadřuje úhlem ve stupních či radiánech, kde 360° (2π rad) značí počátek nové periody zvukové vlny. Zvukový signál nemusí vždy začínat s nulovou fází.

Fáze je veličina, která není zachytitelná sluchem. Projevuje se pouze tehdy, dochází-li k interferenci dvou signálů s rozdílnými fázemi - například při poslechu dvou identických signálů s fází opačnou, tedy posunutou o 180° , dojde k úplnému odečtení

signálů. Naopak při poslechu bez fázového posunu dojde k sečtení obou signálů a tudíž k dvojnásobné amplitudě signálu výstupního.

1.1.6 Časový průběh

Časový průběh signálu se dá vyjádřit dvourozměrným grafem, zobrazujícím závislost okamžité výchylky na čase. Toto znázornění se nazývá oscilogram. Z časového průběhu signálu lze pomocí analýzy zjistit všechny výše uvedené parametry v daném čase.



Obr. 1.1.4.1: Časový průběh basklarinetového tónu a1

1.2 Subjektivní vlastnosti zvukového signálu

Dle psychoakustického poznání mají výše uvedené fyzikální vlastnosti zvukového signálu (vyjma fáze, kterou - jak je výše uvedeno - sluchem nelze zachytit) příčinný vztah k jejich tzv. subjektivním reprezentacím²:

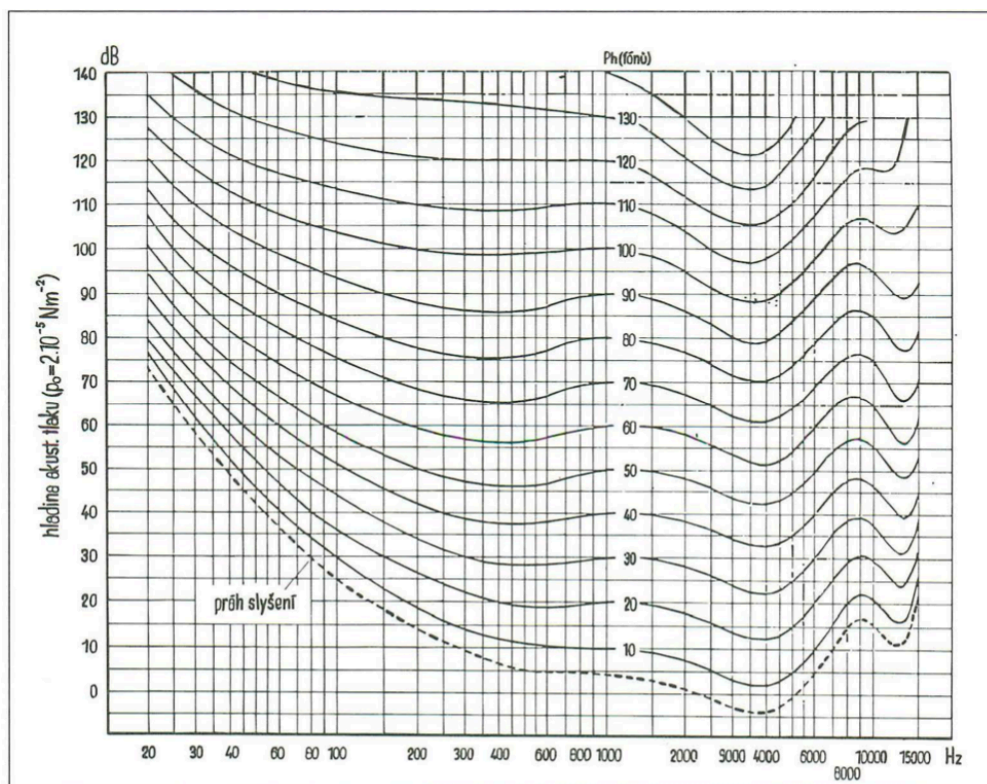
1.2.1 Hlasitost

Hlasitost je jednorozměrnou vlastností popsateľnou na ose silnější-slabší. Přesné určení hladiny hlasitosti sluchem je téměř nemožné. Hlasitost je subjektivním důsledkem amplitudy akustického tlaku, vyvolaného zvukovým signálem. Lidské ucho vnímá poměrně velký rozsah hodnot akustického tlaku, subjektivní hlasitost zvuku pak vnímá v závislosti na hodnotě akustického tlaku logaritmicky.

Subjektivní vjem hlasitosti však nezávisí na hladině akustického tlaku přímo; odvisí od výšky i barvy tónu dle tzv. křivek stejné hlasitosti, které udávají hladiny stejné subjektivní hlasitosti (Obr. 1.2.1.1). Tyto hladiny se proto nevyjadřují v deciblelech, nýbrž ve fónech (Ph). Tato jednotka je identická s hladinou akustického tlaku pouze

² SYROVÝ, Václav. *Hudební zvuk: příspěvek k teorii zvukové tvorby*. V Praze: Akademie múzických umění, 2009. Akustická knihovna Zvukového studia Hudební fakulty AMU. ISBN 978-80-7331-161-2.

na frekvenci 1kHz. Pro jiné frekvence tato shoda neplatí, jak je vyobrazeno na obr. 1.2.1.1.



Obr 1.2.1.1: Křivky stejné hlasitosti (převzato Syrový: Hudební akustika²)

1.2.2 Výška tónu

Výška tónu se zpravidla uvádí pro signály periodické a kvaziperiodické³. Je přímo odvislá od frekvence, narozdíl od ní se však jedná o subjektivní vjem.

Každý posluchač vnímá výšku tónu s přesností danou svými sluchovými schopnostmi. Schopnost určit přesnou frekvenci (název tónu, popř. jeho rozladění v centech) se nazývá absolutní sluch.

Výška je opět jednorozměrnou vlastností na ose vyšší-nižší.

1.2.3 Barva zvuku

Barva zvuku, někdy také zvaná tímbr (z fr. *timbre*), je oprotí výše zmíněným subjektivním vlastnostem zvuku vlastností komplexní, vícerozměrnou. Je odvislá od frekvenčního spektra signálu. Je ze všech subjektivních vlastností nejvíce závislá na

³ SYROVÝ, Václav. Hudební akustika. 2., dopl. vyd. V Praze: Akademie múzických umění, 2008. Akustická knihovna Zvukového studia Hudební fakulty AMU. ISBN 978-80-7331-127-8.

subjektivní interpretaci posluchače (na jeho slovní výpovědi) a vyjadřuje se různými pojmovými páry⁴ nepřímo závislými na komplexnějších akustických veličinách (např. zvuk tupý - ostrý, temný - světlý etc.).

Barva konkrétního zvuku je závislá nejen na relativně statickém spektrálním složení ustáleného tónu, ale především na jevech přechodných (transientních). Toto způsobuje zajímavý fenomén - při odstřižení náběhu tónu se často ztratí charakteristika daného tónu a je tak možné zaměnit tón dvou naprosto rozdílných zdrojů (např. housle a trubka).

⁴ SYROVÝ, Václav. *Hudební zvuk: příspěvek k teorii zvukové tvorby*. V Praze: Akademie múzických umění, 2009. Akustická knihovna Zvukového studia Hudební fakulty AMU. ISBN 978-80-7331-161-2.

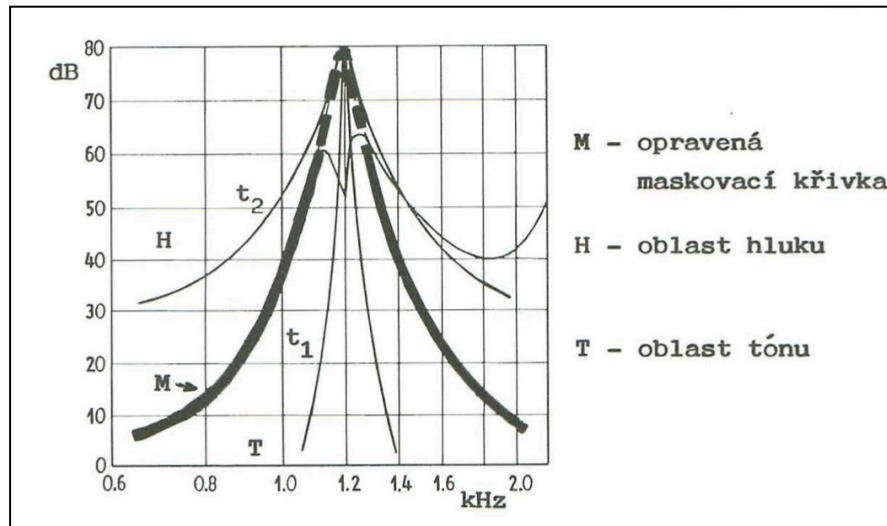
2 Dělení zvukových signálů

Zvukové signály můžeme dělit dle mnoha různých parametrů. Tato kapitola nabídne tři různé druhy dělení, jejichž znalost bude potřebná pro další výklad.

2.1 Dělení signálů dle výškové rozpoznatelnosti

Jak uvádí Syrový⁶, dle statického pojetí signálu se jako tónové jeví signály periodické, jako hlukové pak neperiodické. Tento způsob výkladu je však značně nepřesný, neb „...každý zvuk, tj. i hluk má svoji subjektivní výšku. Existuje pouze jediný, teoretický případ zvukového signálu, který by z pozice subjektivního vjemu výšky byl naprosto indiferentní, a to bílý šum „filtrovaný“ křivkou stejné hlasitosti pro odpovídající okamžitou hladinu hlasitosti.“⁵

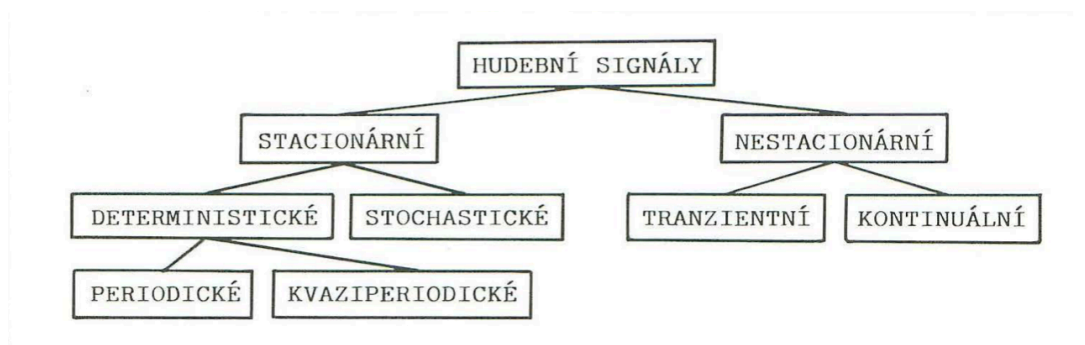
Oproti tomu Syrový staví dynamické pojetí signálu. Toto pojetí signálu pracuje s premisou, že pro přiřazení nekonečně tenké spektrální čáry (konkrétní frekvence) sinusovému tónu je nutné, aby tón trval nekonečně dlouho. Začne-li se tón zkracovat, jeho spektrum se začne rozšiřovat do stran a v při určité míře zkrácení začne postupně ztrácet svůj tónový charakter. Hranice mezi tónem a hlukem se nazývá *opravenou maskovací křivkou*, jak uvádí Syrový⁵ dle Korna a je specifická pro danou hladinu tónu.



Obr 2.1.1.: Hranice mezi tónem a hlukem dle opravené maskovací křivky, sinusový tón 1200 Hz (převzato Syrový: Hudební akustika⁶)

⁵ SYROVÝ, Václav. Hudební akustika. 2., dopl. vyd. V Praze: Akademie múzických umění, 2008. Akustická knihovna Zvukového studia Hudební fakulty AMU. ISBN 978-80-7331-127-8., s. 75

2.2 Dělení signálů dle frekvenčního průběhu



Obr 2.2.1: Dělení signálů dle frekvenčního průběhu
(převzato Syrový: Hudební akustika⁷)

Na úvod této podkapitoly je dobré poznamenat, že toto dělení dle frekvenčního průběhu, převzté ze Syrového⁷, je značným zjednodušením (idealisací) reality zvuku a slouží spíše k pochopení různých variací širokého spektra zvukových projevů. Žádný reálný signál nespadá dokonale pouze do jedné z níže uvedených kategorií,. Ač se může daný signál jedné z definic nejvíce blížit, vždy se v reálném světě bude jednat o kombinaci těchto prvků.

Všechny zvukové signály lze takto rámcově dělit na signály stálé a nestálé (stacionární a nestacionární).

Mezi signály stacionární patří signály *deterministické* a *stochastické*. *Deterministické* signály mají vlastnosti popsatelné matematickou funkcí⁶ narozdíl od *stochastických*, kterýchžto vlastnosti jsou *nahodilé* tj. popsatelné pouze statisticky. Příkladem takového signálu je bílý šum. Stochastické signály mají spojité spektrum. Deterministické signály se dále dělí na *periodické* a *kvaziperiodické*.

Periodické signály jsou takové signály, které jejichž spektrální složení tvoří frekvence dané celými násobky frekvence základní. *Kvaziperiodické* tuto premisu splňují jen částečně - některé složky mohou být mezi sebou v iracionálním poměru⁷. Typický kvaziperiodický průběh mají například zvony či gongy.

Signály nestacionární dělíme na *tranzientní* a *kontinuální*. *Tranzientní* signály mají charakter impulzu a podobně jako stochastické signály vykazují spojité spektrum. *Kontinuální* signály jsou z hlediska tohoto dělení nejsložitější a jsou tvořeny kombinací výše zmíněných druhů signálů. Téměř každý zvukový projev je defacto kontinuálním signálem - jako příklad může posloužit tón libovolného akustického

⁶ SYROVÝ, Václav. Hudební akustika. 2., dopl. vyd. V Praze: Akademie múzických umění, 2008. Akustická knihovna Zvukového studia Hudební fakulty AMU. ISBN 978-80-7331-127-8.

nástroje. Každé nasazení tónu vyvolá určitý tranzient - přechodovou složku, která je nestacionárního typu. Následuje relativně stacionární oblast tónu, než tón opět před svým koncem přejde do tranzientní oblasti.

2.3 Dělení signálů dle trvání

Vzhledem k cílům této práce nelze opomenout dělení zvukových signálů dle doby trvání.

Roads (1, s. 3-43) rozděluje zvukový čas do devíti různých časových měřítek (*scales*). Horní dvě měřítka se nazývají *nekonečné* a *supra scale*. *Nekonečné měřítko* je měřítko víceméně filosofické, ohraničující Roadsův temporální systém. *Supra scale* je měřítko, které do sebe zahrnuje nejdelší hudební jevy. Roads sem řadí například proces kompozice díla, či několikadenní hudební události.

Vzhledem k extrémně dlouhým dobám trvání těchto měřítek se však nemá cenu jimi v kontextu práce dále zabývat.

Další tři měřítka v pořadí odzhora jsou měřítka, na kterých se v drtivé většině případů (alespoň částečně) staví zvukové dílo.

2.3.1 Klasická měřítka trvání zvuku

Třetí měřítko odzhora Roads označuje *macro scale*. Toto měřítko do sebe zahrnuje zvukové vjemy trvajících minuty až hodiny, zpravidla jednotlivé hudební celky (skladby a jejich cykly či věty).

Čtvrté, *meso scale* obsahuje vjemy v řádu sekund až jednotek minut, jedná se zpravidla o jednotlivé motivy, fráze díla, či určitou hierarchii zvukových objektů⁷

Páté měřítko zvané *sound object scale*, tedy měřítko zvukových objektů, do sebe zahrnuje jednotlivé zvukové objekty - nejklašičtěji tóny, ale může se jednat i o různé hlukové a témbrové „objekty“, zpravidla s dobou trvání několika stovek milisekund až jednotek sekund.

Následující tři měřítka popisují jevy, které se nachází na hranici, či již za hranicí lidského vnímání. Tyto jevy byly až do druhé poloviny 20. století víceméně ignorovány, díky omezeným technickým možnostem, které neumožňovaly práci s takto krátkými zvukovými intervaly.

⁷ ROADS, Curtis. *Microsound*. Cambridge, Mass.: MIT Press, c2001. ISBN 0-262-18215-7.

2.3.2 Mikrozvuk

Šesté měřítko Roads nazývá *micro scale*. Signály nacházející se v tomto měřítku - tzv. *grainy* - jsou hlavním bodem zájmu této práce. Jedná se o signály v časovém rozsahu v řádu od asi 100ms (hranice percepce krátkých zvukových objektů⁸) do vyšších desetin milisekundy. Tyto zlomky zvuku jsou prakticky využívány především v procesu digitální granulární syntézy, ale lze je slyšet i v přírodě (např. při narušení struktury suché větve), či v hudebních nástrojích (rumba koule, rainstick).

2.3.2.1 Psychoakustická specifika mikrozvuků

Poslech těchto jednotlivých *grainů* je zpravidla ještě možný (většinou se jejich doba trvání nachází nad mezním časovým prahem sluchu), ale vzhledem k jejich hraniční délce zde již můžeme nalézt mnoho nepravidelností a psychoakustických zákonitostí neznámých ze zkoumání zvuku o běžných délkách (podobně jako u fyziky mikrosvětla).

Prvním takovým problémem je rozeznávání výšky tónu. Jak uvádí Roads (1, s. 24), Doughty a Garner dělí hranici rozeznání výšky tónu na dvě oblasti: Nad 1 kHz musí zvuk trvat minimálně 10 milisekund, aby byla rozeznána jeho výška. V oblastech pod 1 kHz obecně stačí 2-3 periody zvuku. Kratší zvukové signály ucho vnímá jako výškově neurčitě. Tento jev souvisí s výše uvedeným faktem, že pro zkracující se signál se rozšiřuje jeho spektrum. V extrémním případě (jednotkový Diracův impuls) by pak bylo spektrum dokonale spojitě, tj. dokonale hlukové.

Druhým problémem je vnímání hlasitosti. Pro zvuky kratší než 100 ms neplatí běžné zákonitosti pro vjem hlasitosti. Pro zvuky pod touto hranicí, kterou nazýváme *časový práh vjemu hlasitosti* platí, že pro stejný vjem hlasitosti u dvojnásobně krátkého segmentu je třeba signál zesílit o asi 3 dB⁸.

Dalším problémem je problém rozeznatelnosti jednotlivých mikrozvuků od sebe. Pro dva za sebou přehrané zvuky kratší než 100 ms platí, že již nejsou rozeznatelné jako dva oddělené signály, ale jistým způsobem se směšují. Do zhruba 50 ms lze ještě stále rozeznat, že se nejedná o jeden signál, ale pod touto hranicí již zvuky začínají v jeden splývat⁹.

2.3.3 Samplové a subsamplové měřítko zvuku

Sedmé měřítko dle Roadse - *sample scale* - se již nachází za hranicí sluchu. Je odvislé od vzorkovací frekvence digitálního záznamu. V případě nejběžnější vzorkovací frekvence (44,1 kHz) se tedy body zájmu tohoto měřítka pohybují v řádu setin milisekundy. Jak již bylo uvedeno výše, mezní časový práh lidského ucha činí

⁸ SYROVÝ, Václav. Hudební akustika. 2., dopl. vyd. V Praze: Akademie múzických umění, 2008. Akustická knihovna Zvukového studia Hudební fakulty AMU. ISBN 978-80-7331-127-8.

asi 2ms. Tvůrčí práce na této úrovni je prakticky nemožná a čítá pouze pár experimentálních pokusů, především v 70. letech 20. století.⁹

Osmé měřítko - *subsample scale* - pracuje s jevy nezachytitelnými digitálním samplováním zvuku. Jako slyšitelný příklad mohou sloužit například aliasingové relikty frekvencí dle Nyquistova teorému nezachytitelných (zvuky nad cca 22 kHz při vzorkovací frekvenci 44,1 kHz - viz další kapitola).

Poslední, deváté měřítko, *infinitesimální* se nachází již pouze v matematické teorii. Nejběžnější příklad infinitesimálně krátkého zvuku je tzv. Diracova delta funkce (Diracův impuls).

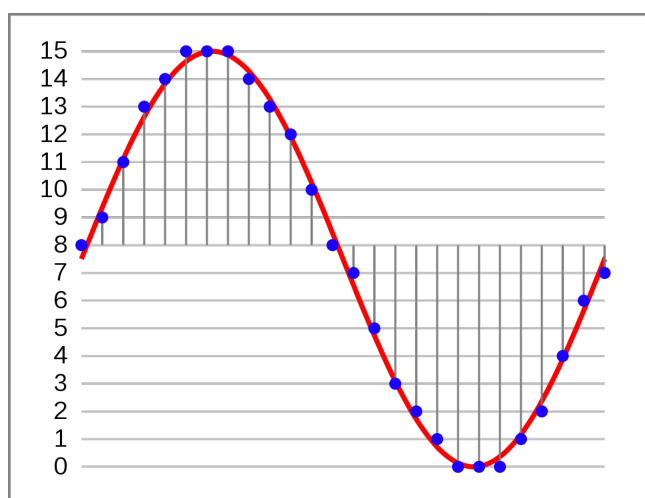
⁹ ROADS, Curtis. *Microsound*. Cambridge, Mass.: MIT Press, c2001. ISBN 0-262-18215-7.

3 Digitální reprezentace zvukového signálu

Forma zvuku, se kterou tato práce bude pracovat, bude forma digitálního záznamu. Tato kapitola stručně pojedná jak o o převodu zvuku mezi digitální a analogovou (reprodukovatelnou) rovinou, tak o obecných parametrech digitálního záznamu.

3.1 Pulzně kódová modulace

Metoda pulzně kódové modulace byla vynalezena v roce 1937 Alecem Reevesem. Jedná se o způsob převodu spojitého analogového signálu na digitální pomocí matematické diskretizace. Signál se pomocí A/D-D/A převodníků převádí z digitálního signálu na analogový a zpět. Tyto převodníky pracují s dvěma hlavními typy operace.



Obr 3.1.1: Pulzně kódová modulace: na horizontále diskrétní vzorky, na vertikále diskrétní kvanta

3.1.1 Vzorkování

Teorie vzorkování (neboli diskretizace signálu v čase) vychází z předpokladu, že každou spojitou funkci lze redukovat na množinu bodů na ní ležících. Vzorkováním signálu je myšleno zachycení okamžitých hodnot (bodů) signálu v daném čase v podobě tzv. vzorků neboli samplů (Obr 3.1.1). Čím častější je frekvence odebrání vzorku, tím přesněji výsledná množina vzorků odpovídá průběhu spojitého signálu. Tato frekvence se nazývá **vzorkovací frekvencí (sample rate)**.

Běžným problémem spojeným se vzorkováním signálu je tzv. **aliasing**. Aliasing souvisí s nemožností zachycení frekvencí nacházejících se nad tzv. Nyquistovou frekvencí (nejvyšší možná zachytitelná frekvence), která je definovaná jako polovina samplovací frekvence. Pokud zvuk překročí tuto frekvenci, objeví se v záznamu frekvence jiná, zrcadlená dle osy Nyquistovy frekvence do zachytitelného pásma. Tato frekvence není v původním signálu obsažená, čímž se signál zkresluje.

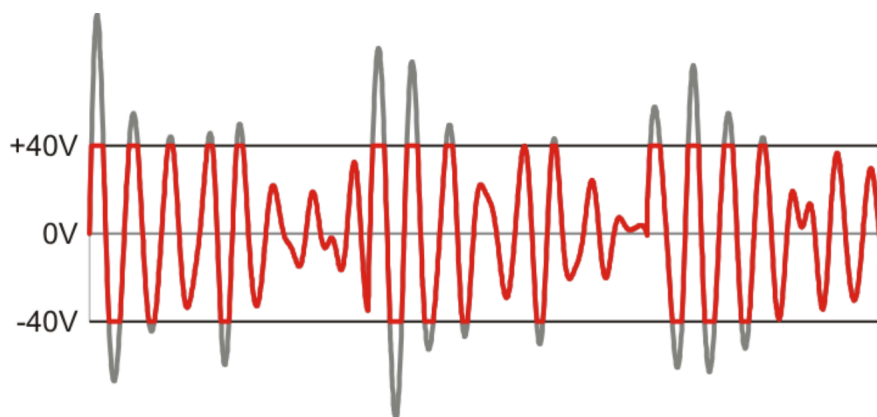
Aliasingu lze předejít jednak pomocí zvýšení vzorkovací frekvence tak, aby její polovina obsahovala frekvence všech zachycovaných signálů, tak pomocí tzv. aliasing filtru, který frekvence nad Nyquistovou před vstupem do převodníku odfiltruje.

3.1.2 Kvantování

Kvantováním (neboli diskretizací signálu v amplitudě) signálu se rozumí přiřazení číselné hodnoty na svislé ose ke konkrétnímu vzorku. Stále platí, že v digitální rovině není možné s neomezenou přesností zachytit spojitý signál, proto i přiřazení okamžité výchylky signálu na spojitě ose podléhá určité diskretizaci. Přesnost diskretizace závisí na **bitové hloubce** převodníku či programu. Bitová hloubka značí maximální počet bitů, které je možné použít k popisu hladin. Kvantizér převodníku přiřadí vzorku nejbližší bitovou hladinu.

Hlasitost digitálního audia se zpravidla měří jako odstup efektivní hodnoty signálu v decibelech od nejvyšší bitové hladiny. Nejvyšší bitová hladina má pak stanovenou hlasitost 0 dBFS (decibel full scale), ostatní hladiny záznamu jsou vždy v záporných hodnotách dBFS.

Mezi nejběžnější problémy spojené s kvantováním patří clipping, tedy ořez signálu z důvodu přebuzení a tzv. kvantizační šum.



Obr. 3.1.2.1: Clipping digitálního audia

Clipping, jak je výše zmíněno, vzniká při přebuzení vstupního signálu na převodníku buďto z důvodu špatně nastavené vstupní úrovně, či z důvodu příliš vysoké výstupní hlasitosti zdroje. Dochází zde ke zkreslení signálu prostým oříznutím hodnot nad maximální (poslední) hodnotou bitového čísla na jeho úroveň (obr. 3.1.2.1).

Kvantizační šum vzniká sdružením kvantizačních odchylek způsobených zaokrouhlením okamžité výchylky signálu ku nejbližší kvantizační úrovni. S rostoucí bitovou hloubkou kvantizační šum klesá.

3.2. Formát *Waveform*

Waveform formát (.wav), je jedním z nejběžnějších formátů pro práci s digitálním záznamem v počítači. Waveform je formátem bezztrátovým; to znamená, že obsahuje nijak nezkreslená data z PCM. Jeho datový obsah je víceméně shodný s RAW formátem (formátem obsahujícím pouze řetězec bitových čísel). Narozdíl od něj však obsahuje hlavičku, ve které jsou uloženy informace potřebné k dekodování signálu. Formát WAV využívá i praktická část této práce, neb se jedná o nejpřímější, nekomprimovanou reprezentaci signálu (spolu s velmi podobným formátem AIFF, používaným v počítačích Mac a dalšími formáty vycházejícími z formátu IFF). Proto je také jednoduché s WAV soubory algoritmicky pracovat. WAV, jak bylo výše naznačeno, se dělí na dvě části, tzv. subchunky: *fmt* (format) subchunk a *data* subchunk¹⁰.

Subchunk *fmt* WAV souboru obsahuje několik informací, které ovlivňují přehrání audiodat. Jedná především o počet kanálů, vzorkovací frekvenci a bitovou hloubku jednotlivého samplu.

Subchunk *data* obsahuje pak samotná audiodata. Tyto jsou pak ve formátu výše uvedeném. Dle počtu kanálů se pro jednotlivý sample uvádí odpovídající počet hodnot - ve stereo signálu je většinou dle normy levý kanál na prvním místě, pravý na druhém.

¹⁰ WAV - Waveform Audio File Format . *File Format Docs* [online] [cit. 9.3.2022]. Dostupné z: <https://docs.fileformat.com/audio/wav/>

4 Zvuková syntéza

Zvuková syntéza je fenomén, který je znám již od z daleké minulosti. Jako jeden příklad za všechny lze uvést varhany. Rozsáhle zvukové možnosti varhan jsou dány možností spojování různých druhů a velikostí píšťal za účelem tvorby kýžené zvukové barvy - jev srovnatelný s kombinováním oscilátorů v současných syntezátorech.

Rozmach důsledné kreativní práce s tímto jevem však souvisí až se zvukem elektroakustickým a elektronickým. První experimenty s elektronickým zvukem se počaly současně s rozšiřováním elektřiny jako takové, tedy na přelomu 19. a 20. století. Zpočátku se jednalo spíše o tzv. elektromechanické nástroje, tedy nástroje mechanické s přídatnými elektronickým součástkami, většinou snímači či zesilovači. Objevovaly se však i první pokusy o generátory zvuku, zpravidla pomocí rotačních generátorů elektromagnetického kmitání¹¹ (např. Telharmonium). Tyto první pokusy později nahradily čistě elektronické obvody, na jejichž bázi fungují analogové syntezátory dodnes.

Digitální syntéza pracuje oproti tomu nejčastěji na principu generování signálu pomocí matematických funkcí. Tento signál je až posléze převeden na signál elektrický, a to pomocí D/A převodníku.

4.1. Typy zvukové syntézy

Slovo „syntéza“ dle obecné slovníkové definice znamená : „*sjednocení jednotlivých částí, složení v celek*“¹². V případě zvuku se však pochopitelně jedná o problém specifičtější. Jak uvádí Syrový¹³, zvuková syntéza je: „*...změna vlastností signálu za účelem naplnění zvukového záměru, prakticky proces řízené generace [zvuku]*“.

Typem syntézy, který nejvíce odpovídá obecné slovníkové definici, je *aditivní syntéza*. Jedná se o nejzákladnější typ syntézy, postavený pouze na využití oscilátorů. Sinusové signály jednotlivých oscilátorů se v různých poměrech skládají za účelem vytvoření (nejčastěji harmonického, či kvaziharmonického) spektra. Tato syntéza je leckdy velmi náročná na provedení, neb často pracuje s velkým počtem oscilátorů.

¹¹ GUŠTAR, Milan. *Elektrofony: historie, principy, souvislosti*. Praha: Uvnitř, 2008. ISBN 978-80-239-8447-7.

¹² PETRÁČKOVÁ, Věra – KRAUS, Jiří. *Akademický slovník cizích slov: [A-Ž]*. Praha: Academia, 1997. ISBN 80-200-0607-9.

¹³ SYROVÝ, Václav. *Hudební signál a jeho syntéza* [online]. s. 169 [cit. 2022-02-24]. Dostupné z: https://www.ziva-hudba.info/files/2014/04/140403201437_pdf_0.pdf

Dalším typem syntézy je syntéza *subtraktivní*. Tento druh syntézy - oproti syntéze aditivní - pracuje naopak s redukcí frekvenčního spektra signálu pomocí filtrace - jednotlivé frekvenční složky se potlačují. Některé typy filtrů mohou však frekvenční složky i zvýrazňovat.

Běžným typem syntéz jsou též syntézy *modulační*. Obecný princip modulačních syntéz spočívá v průběžné změně vlastností signálu pomocí signálu jiného. Může se jednat buď o modulaci amplitudovou, frekvenční, spektrální či fázovou. Dalším typem (amplitudové) modulace je modulace kruhová, která vzniká násobením dvou signálů.

Při modulaci amplitudové a kruhové vzniká na výstupu *součtová a rozdílová složka*, jejichž frekvence odpovídají součtu a rozdílu frekvence původního a modulačního signálu. Při kruhové je původní frekvence zcela potlačena. S amplitudově-modulačním zkreslením se potýká například mnoho implementací granulární syntézy, což bude rozebráno v dalších kapitolách.

Tvarová syntéza pracuje s tvorbou signálu za pomoci jeho nelineárního tvarování. Signál je veden obvodem s nelineární charakteristikou, čímž se jeho průběh a tím i jeho zvuková podoba razantně mění.

Jako určitý druh syntézy se dá popsat také *počítačové modelování* akustických nástrojů. Tímto způsobem je možno dnes již téměř dokonale simulovat všechny fyzikální podmínky v daném nástroji pro přesnou simulaci jeho tónu.

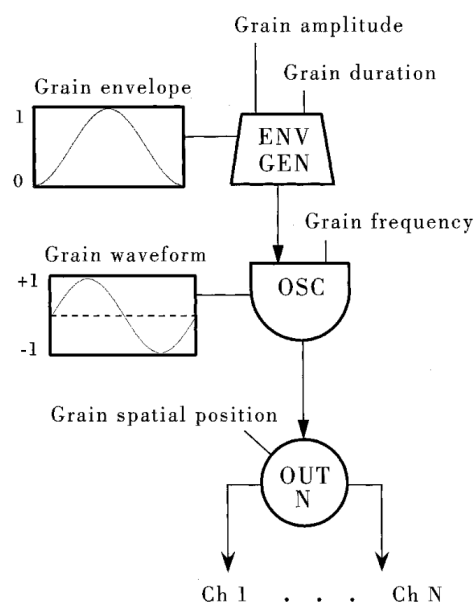
Mezi druhy zvukové syntézy se také zařazuje tzv. „*samplování*“, tedy práce s již přednahraným materiálem, např. za pomoci různých smyček, rytmů atd.

Zvláštním případem samplování je tzv. *granulární syntéza*. Ač je tento typ zvukové tvorby stále pouze prací s předem připraveným materiálem, jeho granulace, tedy rozdělení na dílčí části (časově mezi cca. 1 až 50 ms¹⁴) a následná práce s nimi je však tak zásadním zásahem do jeho struktury, že se dá téměř hovořit o „generování“ zcela nové barvy signálu z materiálu původního. Granulace je proces na hranici syntézy, resamplingu a efektu - konkrétní kvalita výstupu záleží jak na parametrech granulátoru (délka zrna, prolínání atd.), tak na vstupních signálech.

¹⁴ROADS, Curtis. *Microsound*. Cambridge, Mass.: MIT Press, c2001. ISBN 0-262-18215-7.

5 Granulární syntéza: Charakteristika a historie

Jak již bylo výše naznačeno, granulární syntézou se nazývá určitý druh samplování (resamplování) zvuku (může se jednat o komplexní zvuky; terénní nahrávky apod., ale i o prosté sinusové vlny) za využití extrémně krátkých časových úseků původního samplu tzv. zvukových zrn, (anglicky *grain*) seřazených za sebe. Samotné „rozstříhání“ a následní přehání zvuku by však samozřejmě nemělo na zvuk žádný vliv, proto se každé jednotlivé zrno upravuje dle předem zadaných parametrů. Nejběžnějším typem takovéto úpravy je přidání obálky, které z kontinuálního toku zvuku udělá jednotlivé „mikrozvuky“ s vlastním náběhem (attack) a doběhem (decay). Dále lze samozřejmě na jednotlivých zrnech ovlivňovat všechny parametry zmíněné v kapitole 1. Běžným parametrem granulace je též četnost, se kterou granulární syntezátor vypouští jednotlivá zrna v čase, rozmístění zrn po jednotlivých kanálech či promíchání pořadí zrn. Různé typy granulace dále detailněji rozebere další kapitola.



Obr 5.1: Schéma jednoduchého granulátoru sinusového průběhu s obálkou Gaussovy křivky. (Převzato Roads: Microsound)

5.1 Historie mikrozvuku a granulace

Granulární syntéza jako způsob zvukové tvorby je fenomén relativně nový (granulace zvuku do mikročástic vyžaduje techniku, která nebyla do druhé poloviny 20. století dostupná), tudíž historie samotné granulace není zatím příliš obsáhlá.

Kořeny vnímání zvuku jako celku složeného z jednotlivých částí lze hledat již v antické předsokratické tradici atomistů, kteří věřili, že jakákoliv hmota je složená z nedělitelných částic - atomů. Tato tradice prakticky započala dodnes trvající fyzikální debatu o dichotomii vlna/částice, kterou známe například z kvantové fyziky či optiky. Doba atomismu však neměla dlouhého trvání a na dlouho upadla do zapomnění, než jí na přelomu 16. a 17. století znovu vskřísili filosofové a přírodovědci René Descartes (1596-1650) a Pierre Gassendi (1592-1655). Na základě jejich práce pak holandský vědec Isák Beekman přišel s tzv. částicovou teorií zvuku (1616). Dle jeho teorie vibrující objekty narušovaly strukturu vzduchu kolem sebe čímž tvořily jakési malé částice, které pak vibrace předmětu odrážely do různých směrů. Spolu s pokrokem vědy a novými objevy však částicová teorie zvuku začala být neudržitelnou a postupně jí nahradila dodnes uznávaná teorie zvuku jako vlny.

Určité narušení tohoto trendu přemýšlení o zvuku přineslo až 20. století. V roce 1947 přišel maďaský fyzik a nobelista Dennis Gabor s teorií tzv. „akustických kvant“¹⁵. Gabor si byl vědom nedostatečnosti Fourierovy transformace (a obecně představy zvuku jako statické spojitě funkce) pro analýzu a fyzikální aplikaci:

The orthodox method starts with the assumption that the signal s is a function $s(t)$ of time t . This is a very misleading start. If we take it literally, it means that we have a rule of constructing an exact value of $s(t)$ to any instant of time t . Actually we are never in a position to do this. If there is a bandwidth W at our disposal, we cannot mark time any more exactly than by a time-width of the order $1/W$; hence we cannot talk physically of time elements smaller than $1/W$ ¹⁶.

Další problém, který Gabora vedl k formulaci jeho teorií, je již výše několikrát zmíněný problém nekonečné délky sinusového průběhu vycházejícího z praktického využití FT (proto, aby byla FT sinusového signálu bodová, musí být signál nekonečně dlouhý; viz kapitola 3). Jeho řešením bylo zvuk analyzovat v čase po malých „kvantech“. Ačkoliv Gabor tento svůj objev nikdy nevedl do elektronické praxe, sestrojil první (mechanický) optický „granulátor“ na základě filmového projektoru, jenž spojitý zvukový záznam zapsaný na filmovém pásu granuloval pomocí rotačního bubnu s mnoha průduchy. Své teze také popsal matematicky a položil tak svou tzv. Gaborovskou maticí, zobrazující zvuková kvanta, základ okénkové FT (short-time FT), používané k analýze průběhu signálu v reálném čase.

¹⁵ Sound In A Nutshell: Granular Synthesis. *Granular Synthesis: Resource Site* [online] [cit. 10.03.2022]. Dostupné z: <http://www.granularsynthesis.com/hthesis/gabor1.html>

¹⁶ GABOR, Dennis. „Lectures on communication theory.“ Technical Report 238, Research Laboratory of Electronics. Cambridge, Mass.: Massachusetts Institute of Technology, 1952.

Gaborovy teze o zvuku byly průlomové pro vývoj toho, co dnes nazýváme granulární syntézou. Jeho texty měly velký ohlas nejen mezi fyziky, ale i mezi zvukovými umělci, vědci a skladateli, sdruženými v pařížské interdisciplinární skupině Groupe de Recherches musicales (GRM). V rámci této skupiny se řecký experimentální skladatel Iannis Xenakis seznámil s fyzikem Abrahamem Molesem, pracujícím na bázi Gaborových výzkumů. Již před tím byl však fascinován fyzikálně-statistickým a mikroskopickým přístupem k hudbě. V roce 1958 zkomponoval skladbu *Concret PH₁*, která prozkoumávala pomalý vývoj hudby v rámci krátkých zvukových objektů, v případě této skladby vteřinu dlouhých částic nahrávky praskajícího ohně, O rok později se k této technice vrátil, tentokrát již za použití exaktnějších metod kompozice vycházejících z výzkumů jeho kolegy Molese. Výsledkem byla skladba *Analogique B*, (elektronická část kompozice *Analogique A-B* pro smyčce). Xenakis v této skladbě pracoval již pouze s čistými sinusovými vlnami, které (stále ještě za použití magnetického pásku) stříhal na malé části. Xenakis byl nejen otcem termínu „zrna zvuku“ - takto nazýval jednotlivá zvuková kvanta ve své skladbě - ale taktéž byl zjevně vůbec prvním umělcem, který se zrna kreativně pracoval ve smyslu granulární syntézy. Xenakis zrna organizoval pomocí stochastické matice, kterou taktéž využil pro kompozici instrumentální části *Analogique A*.

Vzhledem k omezením, které s sebou práce s páskem přinášela, byla práce na díle velmi zdlouhavá. Další problémy, které analogový přístup přinášel byla například nemožnost implementace příliš krátkých zrn či efektivního překrývání zrn za účelem eliminace modulačních artefaktů (např. při délce zrna 10ms modulované obálkou o tvaru Gaussovy křivky se na signál amplitudově namoduluje frekvence 100 Hz, čímž vzniká součtová a rozdílová složka, viz kapitola 4).

Nový úsvit granulární syntézy přišel (díky výše popsaným problémům) však až s rozšířením výpočetní techniky a s tím spojené digitalizace zvuku. Digitání zvuk umožnil práci na úrovni jednotlivých sampleů zvuku a také umožnil celý proces významně urychlit díky algoritmickému skládání jednotlivých zrn. Prvním, kdo tento způsob implementoval do praxe, byl americký skladatel a programátor Curtis Roads (1951)¹⁷.

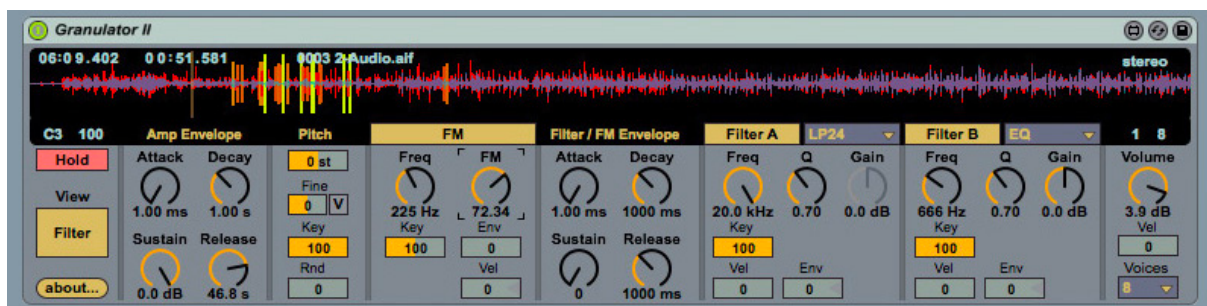
Roads významně rozšířil poznání v oblasti granulární syntézy, nejen svými praktickými implementacemi granulárního zvuku v podobě počítačových programů či hudebních komposic s granulárním materiálem, ale především svou velmi obsáhlou prací *Microsound*, která je i jedním z výchozích textů této práce. Tato práce zavádí granulární syntézu jako specifický typ přístupu ke zvuku, shrnuje její historii a řeší praktická a kompoziční využití granulace a mikrozvuku obecně. Práce je zajímavá nejen svým obsahem, ale také formou, která odraží postmoderní trend syntézy

¹⁷ ROADS, Curtis. *Microsound*. Cambridge, Mass.: MIT Press, c2001. ISBN 0-262-18215-7.

exaktní vědy a umění - práce používá terminologii striktně vědeckou, avšak spojuje vědecké termíny s termíny nevědeckými, převážně uměleckými a synestetickými termíny používanými pro popis zvuku, času atd. - čímž se stává práce velmi čtivou jak pro (elektro)akustiky a programátory, tak pro umělce a skladatele s těmito technologiemi pracujícími. Roads rozšířil práci s mikrozvukem o další typy částicové syntézy s různými granulárními objekty, které nazýval glissony (určitá mikroglissanda), grainlety (zrna o délce vázané na jejich frekvenci), trainlety (série impulsů) a pulsary (imulsní generátory vlnění). Proudů zrn nazývá oblaka (clouds) a k jejich třídění si dopomáhá meteorologickou taxonomií oblaků (stratus, cummulus atd.).

Dalším významným průkopníkem granulární syntézy je kanadský skladatel Barry Truax. Jeho největším přínosem byla první implementace granulární syntézy v reálném čase, realizovaná na konci 90. let v rámci jeho skladby *Beauty and the Beast*.

Dnes jsou již různé granulární programy pracující v reálném čase běžně implementovány jako zásuvné moduly mnoha virtuálních prostředí pro kreativní práci se zvukem.



Obr. 5.1.1.: Plugin Granulator II v DAW
Ableton 10

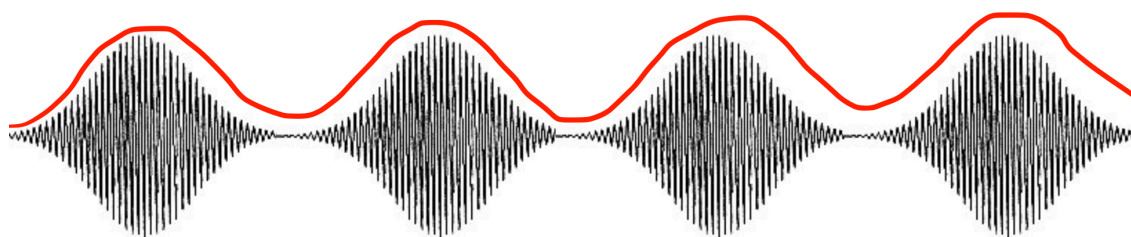
6 Praktické aplikace granulární syntézy

Při aplikaci granulárního procesu na data můžeme použít dva základní druhy přístupů: syntézu časově synchronní a syntézu časově asynchronní:

6.1. Časově synchronní syntéza

The spacing within the grains in granular synthesis can radically change the texture that is created. With synchronous granular synthesis the grains are all separated by the same amount of time, or with some linear relationship¹⁸.

Časově synchronní typ syntézy se - jak je z výše uvedeného citátu patrné - vyznačuje způsobem organizace zrn v pravidelných intervalech za sebou. Hlavním parametrem je zde hustota zrn. Mezi přehráním jednotlivých zrn může být buď určitý interval ticha, či mohou jít zrna spojitě za sebou v jakémsi řetězu, případně se (při větší hustotě přehrávání, než je délka jednotlivého zrna) navzájem překrývat.



Obr 6.1.1.: Amplitudově namodulovaný signál (červeně) při časově synchronním způsobu organizace zrn

Hlavním atributem (a svým způsobem i problémem) tohoto typu organizace zrn jsou amplitudově-modulační artefakty, které výsledném signálu produkují parazitní frekvence součtové a rozdílové (popř. celá postranní pásma) buď ve smyslu rytmickém (při malé hustotě zrn), či již slyšitelné frekvence (hustota zrn asi 30/s a více), přičemž frekvence a spektrum těchto parazitních signálů jsou ovlivněny jak samotnou „vnitřní“ frekvencí jednotlivého zrna, tak tvarem obálky zrn. Tyto součtové a rozdílové frekvence jsou patrnější u periodických a kvaziperiodických průběhů granulovaného materiálu. Možností, jak tento jev redukovat, se nabízí několik:

První možností by bylo narušení časové synchronicity zrn pomocí (nejčastěji) náhodného zpoždování jejich nástupu. Tím dojde k „rozmazání“ parazitních signálů

¹⁸ ROADS, Curtis. The Computer Music Tutorial. Mass.: MIT Press, 1996. ISBN 978-0262680820

směrem k šumovým spektrům¹⁹. Detailněji je tento problém rozebrán v podkapitole 6.2.

Dále je také možno jednotlivá zrna překrývat. Při zvolení vhodné obálky zrna tímto způsobem dojde k výraznému snížení hlasitosti postranních pásem, avšak pouze za cenu zkrácení vstupního materiálu o součet překývaných pásem.

6.2 Časově asynchronní syntéza



Obr 6.2.1.: Rozmístění zrn při časově asynchronní syntéze

Časově asynchronní syntézou je myšleno vypouštění jednotlivých zrn do proudu v nesynchronních, stochastickým algoritmem řízených intervalech, projevující se jako náhodná vzdálenosti mezi zrn (viz např. *Xenakis - Analogique B*, kap. 5). I zde lze řídit hustotu zrn, avšak stále pouze jako obecný rámec (např. rozsah povolených vzdáleností) stochasticky generovaných vzdáleností mezi zrn.

I při tomto druhu syntézy dochází k určitému druhu amplitudové modulace signálu. Vzhledem k stochastickému způsobu organizace zrn v čase však modulační signál nabývá šumového rázu. Subjektivní efekty tohoto typu modulace za dodržení stabilní délky zrna lze nalézt v tabulce 6.2.1. Délku zrn lze však také variovat, což bude detailněji rozebráno v další kapitole.

¹⁹ ROADS, Curtis. *Microsound*. Cambridge, Mass.: MIT Press, c2001. ISBN 0-262-18215-7.

Doba trvání zrna	Modulační frekvence	Výsledný subjektivní efekt
0,2 ms	5 kHz	Šumová disintegrace signálu
1 ms	1 kHz	Ztráta výšky tónu
10 ms	100 Hz	Třepetání, zurčení
50 ms	20 Hz	Formace stabilních výšek tónu
100 ms	10 Hz	Rychlé, aperiodické tremolo
200 ms	5 Hz	Aperiodické tremolo

Tabulka 6.2.1.: Subjektivní efekty aperiodické modulace v závislosti na délce zrn. (Převzato: Roads: Microsound (1, s. 101))

6.3. Další parametry granulární syntézy

6.3.1 Délka zrna

Výstup při stejných délkách zrna byl již dostatečně probrán v předchozí podkapitole. Délku zrna lze však také variovat. Při tom lze využít více způsobů: délky jednotlivých zrn lze stanovit *dle stochastického algoritmu*, při čemž u časově synchronní syntézy dochází k vjemu podobnému syntéze časově asynchronní. Dále lze také délky variovat *dle určité funkce v čase*. Délku zrna lze též *navázat na jeho vnitřní frekvenci*. Tento typ zrna Roads ve své knize nazývá *grainlet*.

6.3.2 Obálka zrna

Změna obálky zrna může zapříčinit změnu ve spektrálním složení výsledného signálu (např. Při strmějším náběhu či doběhu obálky vzniká bohatší spektrum). Obálky lze samořejmě také v čase variovat, ale opět za docílení více „šumového charakteru“ výstupu.

6.3.3 Nasycení granulárního proudu

Hustota zrn sama o sobě neprozrazuje nic o charakteru výstupu - je vázaná na délku zrna. Roads (1, s. 105) přichází za účelem analýzy subjektivně vnímané hustoty s tzv. *fill factorem*, který popisuje „nasycení“ proudu zrn a počítá se dle vzorce $n = \rho \cdot t$, kde n je nasycení, ρ hustota zrn [s^{-1}] a t čas jednotlivého zrna [s]. Při nasycení 1 jdou zrna v řetězu za sebou a čas mezi začátky zrn (převrácená hodnota hustoty) se rovná přesně délce jednoho zrna. Při větším nasycení než 1 se zrna již překrývají.

Jak bylo popsáno v kapitole 2, mikrozvuky jdoucí za sebou v rychlých sledech začnou v určitých intervalech sluchově splývat. Tato hranice se uvádí okolo 50 ms. Pro zrna kratší, než 50 ms tedy platí, že limitní vzdálenost mezi náběhy zrn, při které

zrna začnou splývat je asi 50 ms (delší zrna začnou splývat i při nižší vzdálenosti díky překrývání)

6.3.4 Obsah zrna

Obsahem zrna je zde myšlen jeho signálový obsah. Granulární proud může být vytvořen z jednoho či více druhů vstupního signálu. V nejjednodušším případě se jedná o jediný sinový signál dále zpracovaný na bázi výše uvedených parametrů. Může se samozřejmě také jednat o více různých signálů, které se ve své mikropodobě různě střídají. Podobného efektu, jako u kombinace více vstupních signálů se také dá dosáhnout použitím granulace pouze jednoho nestacionárního, kontinuálního signálu (např. orchestrální skladba s rychle se měnící melodií, harmonií i instrumentací).

Při využití komplexnějších signálů s obsahem vyšších frekvencí na spektru mohou granulaci snadno vznikat aliasingové jevy. Toto lze ošetřit pokud možno co nejvyšší vzorkovací frekvencí, či vstupní filtrací.

6.3.5 Prostorové rozmístění

Při stereofonním, či ideálně vícekanálovém výstupu se dá každému zrně (či jeho skupinám) přiřadit specifický výstupní kanál. Granulární proud se dá opět buď jednotlivým kanálům přiřazovat za pomoci funkce (např. kruhový pohyb u surround systémů), či za pomoci stochastického algoritmu. Stochastickým způsobem vzniká mezi jednotlivými kanály lokalizačně nejasný, prostorový vjem granulárního „oblaku“

Tímto způsobem lze i granulovat libovolný signál bez použití obálky, pouze přepínáním jednotlivých kanálů. (v monofonním výstupu by granulace nebyla slyšitelná).

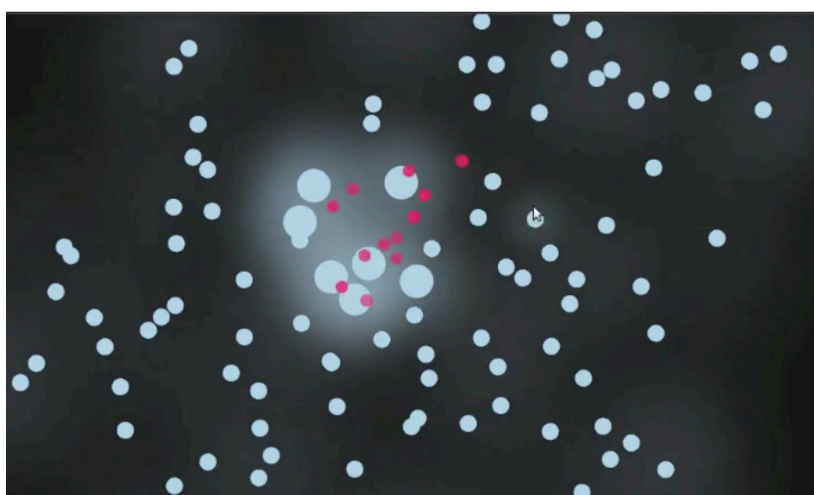
7 Zavedení vícesamplového vstupu do granulární syntézy

7.1 Současný stav poznání

Rané experimenty s granulární syntézou, které položily základ tohoto způsobu zvukové tvorby pracovaly téměř výhradně s jednoduchými sinusovými vlnami. Tento způsob tvorby je nejčastěji zmiňován i v Roadsově publikaci *Microsound*, kde jsou sinusové vlny většinou brány jako výchozí materiál pro granulaci. Granulaci přednahranych samplů či dokonce granulaci za použití více samplů se Roads téměř nevěnuje a tento druh práce zmiňuje víceméně jen ve výčtu svých komposic. Zavádí nicméně dichotomii mezi tzv. *monochromatickým a polychromatickým* proudem zrn, přičemž polychromatický (vícebarevný) proud obsahuje namíchaná zrna z rozdílných vstupních materiálů a tím dává výslednému zvuku barevnější výraz.

V současné době existuje několik projektů, pohybujících se na hraně toho, lze nazvat jako vícesamplová granulární syntéza. Prvním z nich je projekt Audiostellar.

Audiostellar je netradiční software/rozhraní zaměřený na kreativní práci se zvukovými bankami. Využívá nestandardního grafického rozhraní ve formě barevných „hvězdokup“ zvukových souborů automaticky tříděných dle barvy zvuku, z nichž každá hvězda představuje jeden zvukový soubor (sample). Pomocí tzv. particle (částicového) módu může uživatel vyvolat třepotající se zrna (na obrázku 7.1.1 červené tečky), která soubory (na obrázku reprezentované světlem modrými hvězdami) při dotyku spouští rychle za sebou. Tím vzniká jakýsi typ granulace s použitím více vstupních samplů. Uživatel může hustotu těchto červených zrn, rychlost jejich třepotání a další parametry kontrolovat pomocí zadávání dat v reálném čase.



Obr. 7.1.1.: Hvězdokupy v rozhraní Audiostellar

Dalším, velmi podobným, avšak značně rozsáhlejším projektem je open source projekt Fluid Corpus Manipulation při centru hudby a hudebních technologií na britské University of Huddersfield. Projekt sdružuje skladatele, programátory a mezioborové umělce za účelem implementace souboru algoritmů pro třídění a zpracování rozsáhlých zvukových bank, především pomocí umělé inteligence²⁰. Vzhledem k množství audia ve zvukových bankách se zde granulární (mikroskopický) přístup nabízí jako jedna z možností kreativní kombinace zvukových souborů v takto rozsáhlých kolekcích.

7.2 Vícesamplovost jako způsob barevné syntézy

Ani jeden z výše zmíněných přístupů vícesamplové granulace se však přímo nedotýká způsobu, který chci v této práci dále rozvést. Nejblíže je přístup Roadsův. Roads přemýšlí o míchání více vstupních samplů jako o míchání barev (polychromatický přístup). Tento přístup však již dále příliš nerozvádí, což je dle mého názoru chyba, neb se v něm (podobně jako v dalších přístupech, které Roads rozvádí více) skrývá velmi zajímavý potenciál pro zvukovou tvorbu. Tímto způsobem lze rámec granulární syntézy využít pro skutečně plnohodnotnou syntézu v pravém smyslu původního slova „syntéza“ jako proces kombinace více vstupních elementů za účelem vytvoření nové kvality.

Vstupní samplý mohou být samozřejmě mnoha druhů. Může se jednat o jednoduché průběhy některých syntezátorů či hudebních nástrojů, komplexnější perkusní (kvaziperiodické, hlukové) a impulsní signály či ruchové atmosféry různé topologie, spektrální bohatosti apod. Na základě této rozmanitosti je nasnadě navrhnout obecný přístup, který bude vyhovovat všem těmto druhům signálu.

7.2.1 Volba vhodných parametrů pro praktickou realizaci

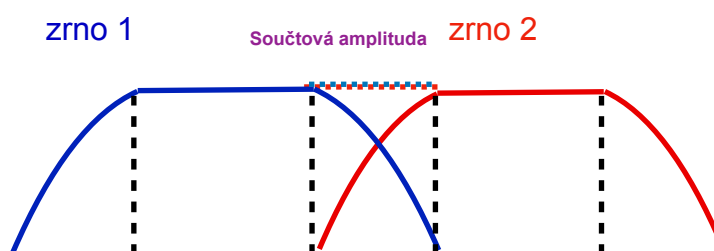
Jako první krok praktické části této práce bylo nutné vůbec zvolit způsob, jakým se dostat ke kýženému výsledku. Vzhledem k tomu, že se jedná o dosud pravděpodobně nezkoumaný problém, je třeba samozřejmě při teoretickém rozvrhu práce předpokládat s určitou variabilitou jeho vstupních parametrů ustálenou až na základě praktických experimentů. Proto bude rozvržení těchto parametrů pouze rámcové.

Vycházet musíme především z představy ideálního výsledku, kterého chceme dosáhnout. Jak je již výše zmíněno, cíl práce je především smíchání dvou a více samplů granulárním způsobem dohromady za účelem tvorby (syntézy) nové zvukové barvy postavené na základě barev těchto vstupních samplů. Z tohoto nám již vyplývá první směr, kterým se můžeme dále ubírat:

²⁰ Fluid Corpus Manipulation. *Fluid Corpus Manipulation* [online]. Copyright © [cit. 17.03.2022]. Dostupné z: <https://www.flucoma.org/>

Při konstrukci granulačního programu je dobré si nejprve stanovit, zdali chceme pracovat na bázi časově synchronní či asynchronní. V našem případě dává z mnoha důvodů smysl pracovat spíše na bázi časově synchronní. Časově asynchronní syntéza zavádí do výstupu určitou náhodnost, která se ve zvukové doméně projevuje více či méně proměnným šumem. Tento efekt je jistě nezanedbatelně efektní v jiných druzích kompozice a kreativní práce se zvukem obecně, avšak v našem případě konstrukce programu za účelem exaktní analýzy jeho výstupu nám vyhovuje spíše takový přístup, který nám dovolí se zvukem pracovat tak, abychom všechny parametry měli pod kontrolou. Distorzní šumová složka je navíc při takovéto kombinaci zvukových barev nežádoucí, neb přidává barvu vlastní a tím znemožňuje výstup brát čistě jako kombinaci barev vstupních.

S časově synchronní syntézou nám však vyvstávají modulační problémy uvedené v předchozí kapitole. Zavedení určité časové asynchronicity jsme již vyloučili, tudíž nám za účelem eliminace amplitudové modulace zbývá volit cestu překrývání jednotlivých zrn určitým druhem křížového prolínání (crossfade).

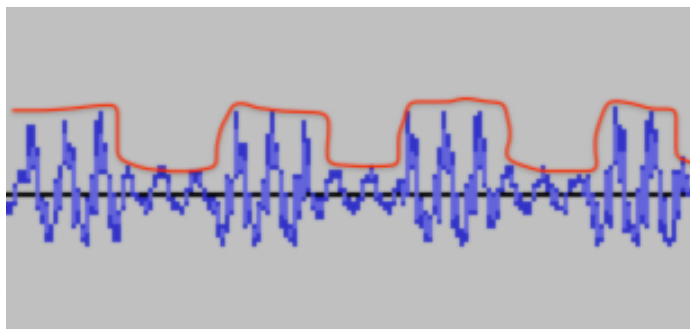


Obr 7.2.1.1.: Dokonalé křížové prolínání dvou stejně hlasitých zrn

S křížovým prolínáním souvisí i volba obálky. Ideální obálka pro dokonalé křížové prolínání je obálka o struktuře *fade in-původní amplituda-fade out*, přičemž *fade out* část libovolného zrna se sčítá s *fade in* částí zrna následujícího a je s ní délkou rovna (viz obr 7.2.1.1). Takto zůstane zachována původní amplituda obou signálů a ona parazitní amplitudová modulace způsobená náběhy a doběhy jednotlivých obálek na nich zmizí (ne však případná modulace způsobená rozdíly amplitud signálů, viz dále). Délka prolínání (déłky *fade in* a *fade out* částí na obálce) může být buď absolutní (stejná pro všechny déłky zrn) či relativní k déłce zrna. Za účelem zjednodušení procesu bude náš program využívat možnost první, neb *k účelu barevné syntézy je záhodno využít co nejkratší déłku zrna*. Délka prolínání se tedy nastaví adekvátně k nejkratším běžným déłkám zrn (desetiny ms), čímž se nepoměr mezi „prolínačkou“ a déłkou zrna srovná. Jak je dále patrné, tento přístup vyžaduje minimální déłku zrna proto, aby mohlo být prolínání dokonale provedeno. Tato minimální déłka zrna se pak odvíjí od součtu délek *fade in* a *fade out* částí.

Zde nám však již vyvstává několik problému spojených s tímto přístupem. Prvním z nich je zkrácení stopáže vstupních materiálů o součet trvání jednotlivých prolínání (při oné minimální délce zrna by zkrácení materiálu dosáhlo až 50%). Tento problém se však stává závažným pouze tehdy, granulujeme-li kontinuální (vyvíjející se) signály - například nějakou melodii - za účelem zavedení jakéhosi granulárního „efektu“. V našem případě však s granulací nechceme pracovat jako s efektem, ale jako s nástrojem pro barevnou syntézu na základě víceméně stacionárních vstupních signálů (např. tóny určitých hudebních nástrojů či statické zvukové objekty). V případě zkrácení stacionárního signálu prolínáním by u stacionárních signálů neměl vznikat téměř žádný rozdíl v kvalitě zvuku. V případě kvazistacionárních signálů, jako jsou již výše zmíněné tóny nástrojů bude v následné analýze výstupu našeho granulátoru zajímavé sledovat oblast tranzientní, na kterou tento „zkracovací“ problém bude mít zjevně největší dopad.

Jak bylo výše zmíněno, tento přístup prolínání signálů však nevyřeší případnou amplitudovou modulaci způsobenou rozdíly mezi amplitudami signálů jednotlivých samplů (viz obrázek 7.2.1.2). Proto je nutné hlasitosti jednotlivých samplů buďto předem normalizovat, či v ideálním případě algoritmu rovnou dodat samplý s co nejbližší hlasitostí.



*Obr.: 7.2.1.2: Amplitudová modulace
způsobená rozdílně hlasitými zrny
(dvousamplová granulace)*

Pro zachování frekvence a tím i barvy původních signálů je také nezbytně nutné se posouvat v granulovaných zvukových souborech (soubory granulujeme a přiřazujeme výstupnímu souboru postupně, např. 1-2-3 1-2-3 1-2-3 etc) v časové kontinuitě s celkem - tedy tak, že vždy při provedení jednoho „cyklu“ granulace se v znovuuchopeném souboru posuneme nikoliv jen o čas jednoho zrna tohoto souboru (čas přečtený v minulém cyklu), avšak rovnou o čas rovný součtu délek všech v minulém kole zapsaných zrn. Příklad: třisamplová granulace - bereme 2ms z prvního, 3ms z druhého a 5ms z třetího souboru. Po zapsání zrna třetího souboru následuje návrat na soubor první. V souboru se posuneme nikoliv o 2 ms (zde jsme minule skončili), avšak rovnou o 10 ms (2+3+5). Toto platí i pro soubory ostatní (součet cyklu

je vždy stejný). Kdybychom takto nepostupovali a v následujícím souboru skočili opět na místo, kde jsme v minulém cyklu skončili, změnilo by se spektrální složení signálu roztáhnutím vlny v čase. (např. kdyby zrno sinusového průběhu trvalo půl periody a následovala dvě zrna šumová, následující půlperiody by se zapsala s časovým zpožděním rovným součtu délek šumových zrn, čímž by do spektra přibyla další složka o nižším kmitočtu).

Naposledně je také nutné stanovit, zda budeme pracovat v reálném čase, či ne. Pro účely zpětné analýzy volíme způsob práce mimo reálný čas, tedy způsob na ose: *nahrání zvukového souboru do paměti programu->granulární algoritmus provedený na tomto souboru->zápis výstupu do nového souboru*. Výstup jistě v reálném čase být zpracován může (přímé odesílání výstupu do reproduktoru), avšak tento přístup není pro účely zpětné analýzy ideální.

Máme tedy již nastíněn způsob konstrukce algoritmu. Nyní se musíme ještě krátce zamyslet nad druhy vstupních signálů, které chceme algoritmu zadávat.

Jak již bylo výše nastíněno, pro tvorbu nové, dostatečně stacionární barvy signálu je zákonitě nutné mít dostatečně stacionární vstup. Takový vstup může být periodický (kvaziperiodický) či stochastický, případně někde mezi oběma přístupy (perkusi zvuky, hlukové atmosféry).

8 Popis vývoje vícesamplového granulárního syntezátoru v jazyce Python 3

8.1 Volba vhodného prostředí

Než bylo vůbec možné začít s prací, bylo pochopitelně nutné zvolit vhodný nástroj k tvorbě granulárního programu. V oblasti algoritmické práce s audiem se nám nabízí několik možností, kterými se lze vydat:

Jednou z možností jsou programovací prostředí přímo navržená k práci s audiem - mezi nimi můžeme zmínit například platformy Pure Data, MAX/MSP či CSound či SuperCollider. Zatímco Pure Data a MAX/MSP jsou grafická prostředí určená převážně k algoritmické práci v reálním čase, CSound a SuperCollider jsou naproti tomu postavená na rozhraní textovém - procedurálním (podobně jako další klasické programovací jazyky), čímž se hodí spíše k práci mimo reálný čas. Pracovat v reálném čase však z výše uvedených důvodů nepotřebujeme, proto volíme raději procedurální přístup, který umožní při složitějších konstrukcích větší přehlednost, než grafický. Nabízí se nám tedy prostředí typu CSound a SuperCollider. Jejich specifika pro práci s audiem může být jistě přínosem pro spoustu projektů, avšak za cenu poněkud netradiční syntaxe oproti klasickým programovacím jazykům (C, Java etc.) V našem případě pouhé mechanické manipulace s jednotlivými vzorky audia však není nutné provádět složitější operace s audiem a tónovými výškami, proto jsem (i vzhledem k mým programátorským zkušenostem z klasických, obecných jazyků) zvolil raději cestu práce s obecným jazykem, který je mé zkušenosti bližší.

Z obecných jazyků jsem zvolil jazyk *Python 3*. Python kombinuje logiku a možnosti procedurálních víceúrovňových jazyků s jednoduchostí práce s ready-made strukturami ve formě rozsáhlých externích i interních knihoven. Komunita uživatelů Pythonu je rozsáhlá a Python proto obsahuje nepřehledné množství snadno implementovatelných knihoven pro práci téměř s jakýmkoliv druhem dat, včetně množství knihoven určených pro práci s audiem. V našem projektu bude práce na úrovni audia (jak bylo výše uvedeno) pouze základní, tudíž je jazyk Python vzhledem ke své jednoduchosti, přehlednosti a i vzhledem k mým dosavadním zkušenostem ideálním prostředkem pro dosažení vytýčeného cíle.

8.2. Volba algoritmického přístupu, nástrojů a parametrů

Jazyk *Python 3* je jazykem multiparadigmatickým (nabízejícím práci v různých programovacích paradigmatech). Je tedy nasnadě před začátkem práce takové paradigma zvolit. Vzhledem k typu problému se bude jednat o jednoduchý program s několika navázanými funkcemi (čtení audia, granulace, zápis), proto jsem zvolil

přístup zapouzdření souboru hlavních funkcí programu do třídy, tedy přístup objektově orientovaný.

Dále je nutné zvolit vývojové prostředí. Python je samozřejmě možné psát v textovém editoru a následně interpretovat v příkazové řádce, avšak za účelem kontroly chyb, strukturace kódu a obecného urychlení práce je nasnadě využít nějaké z dostupných IDE. Jako jedno z nejmodernějších a nejoblíbenějších bezplatných IDE s nejvíce funkcionalitami se nabízí IDE *PyCharm CE* české firmy JetBrains, které jsem také z výše uvedených důvodů pro práci nakonec zvolil.

8.2.1. Volba knihoven

Python samotný neumožňuje efektivní čtení audia. Za tímto účelem však lze importovat příslušnou knihovnu, která nám práci zjednoduší. Těchto knihoven samozřejmě existuje na internetu mnoho (*SciPy.wave*, *PyAudio* aj.), avšak i interní soubor knihoven Python Standard Library obsahuje knihovnu s tímto zaměřením (byť pouze se základními funkcemi) a tou je knihovna *wave*, určená pro čtení a zápis souborů ve formátu *waveform*. Tato knihovna obsahuje tři třídy - *wave* a z ní vycházející *wave_read* a *wave_write*²¹. Třída *wave* slouží (pomocí funkce *wave.open*) k otevření souboru. Třídy *wave_read* a *wave_write* dále pracují s objektem vzniklými na základě volání metody *wave.open*. Tyto třídy mají pouze několik základních funkcí pro čtení a zápis do hlavičky souboru za účelem stanovení, či zjištění parametrů souboru (vzorkovací frekvence, bitová hloubka, počet kanálů etc.) a pro čtení a zápis samotných audiodat. Pro naše účely však takto jednoduchá knihovna bohatě stačí.

Další důležitou knihovnou pro naší práci bude knihovna *NumPy*. Jedná se běžně užívanou, velmi rozsáhlou knihovnu matematických funkcí určenou především pro práci s vektory a maticemi²². Tato knihovna umožňuje např. sčítat či násobit více řetězců dat mezi sebou. Vzhledem k tomu, že digitální audio je defacto posloupností za sebou řazených zvukových vzorků, tato knihovna se bude jistě hodit například při sčítání dvou zrn audia v oblasti křížového prolínání (viz dále).

Další využitou knihovnou je všestranná knihovna pro zpracování audia *Pydub*. Z této knihovny využívám metody *dbfs* a *apply_gain* k normalizaci hladiny vstupních audiosouborů.

²¹ *wave* — Read and write WAV files — Python 3.10.2 documentation. 302 *Found* [online]. Copyright © [cit. 28.01.2022]. Dostupné z: <https://docs.python.org/3/library/wave.html>

²² NumPy [online]. Copyright © 2022 NumPy. All rights reserved. [cit. 30.03.2022]. Dostupné z: <https://numpy.org/>

Poslední dvě knihovny, které budou k našemu účelu potřeba budou interní knihovny *struct* a *os*. Knihovna *struct* bude využita k převedení audiodat zakódovaných po jednotlivých bytech (*struct.unpack*) a knihovna *os* k výpisu vstupních souborů do funkce *main* a ke smazání normalizovaných souborů po dokončení granulace.

8.2.2 Volba konkrétních parametrů digitálního audia

Dále je třeba stanovit jednotné parametry digitálního zpracování audia. Mezi tyto parametry počítám počet kanálů, vzorkovací frekvenci a bitovou hloubku.

Počet kanálů byl dle běžných standardů zvolen jako 2 kanály, tedy stereofonní vstup a výstup. V našem případě granulace za účelem tvorby nové barvy by jistě postačil vstup monofonní (stranová informace je pro nás v tomto případě nedůležitá), avšak pro kompatibilitu s rozličnými externími soubory, které mohou být ve stereofonním formátu jsem zvolil přístup stereo kompatibility, neb jejich monofonizace by mohla přinést nežádoucí fázové artefakty ovlivňující výslednou barvu (hřebenový filtr). Proto jsem se (v rámci zjednodušení práce) rozhodl program psát výhradně pro dvoukanalové vstupní soubory - případný mono vstup vypíše chybovou hlášku a program ukončí. Jistě by bylo možné napsat jednoduchý skript pro stereofonizaci mono vstupu (zdvojením jednoho kanálu), ale počítám s tím, že případný uživatel programu je (v případě potřeby práce s mono soubory) tento elementární úkol schopen provést sám buď přímo v Pythonu, či v některém z nepřeberného množství programů určených k úpravě audia.

Výchozí vzorkovací frekvenci programu lze nastavit přímo v programu. Přitom platí, že vstupní soubory musí dodržet tuto nastavenou vzorkovací frekvenci, jinak dojde k změnám v rychlosti přehrávání souboru a tím i ke změně spektra. Proto v programu před začátkem proběhne kontrola vzorkovací frekvence a při její neshodě s výchozí program vypíše upozornění, granulaci však stále provede, neb z teoretického hlediska tomuto procesu nic nebrání, ač dojde k výše zmíněné změně spektra.

V rámci našeho programu nastavím pro jednoduchost práce a širokou kompatibilitu v základu normalizovanou frekvenci 44,1 kHz. Stále však platí, že vzorkovací frekvence je jednoduše nastavitelná v proměnné třídy *granulator*.

Výchozí bitová hloubka výstupu byla pro naši práci nastavena na 16-bit, ale podobně, jako vzorkovací frekvence, se dá ručně nastavit na jinou přímo v metodě *granulator.write*. Zde však již není nutné dodávat vstupní soubory o stejné bitové hloubce - program automaticky provádí přepočítání bitových hladin všech souborů na float s hodnotami v rozmezí -1 (nejnižší bitová hladina) až 1 (nejvyšší) dle vzorce *aktuální bitová hladina/maximální bitová hladina*. S těmito daty se nadále pracuje a

na konci programu dojde opět opačnou operací k jejich zpětnému přepočítání na bitové hladiny. Samotný popis funkce je uveden v další kapitole.

Pro normalizaci hlasitosti všech audio souborů byla určena dle standardu EBU²³ hladina prozatím -18 dBFS.

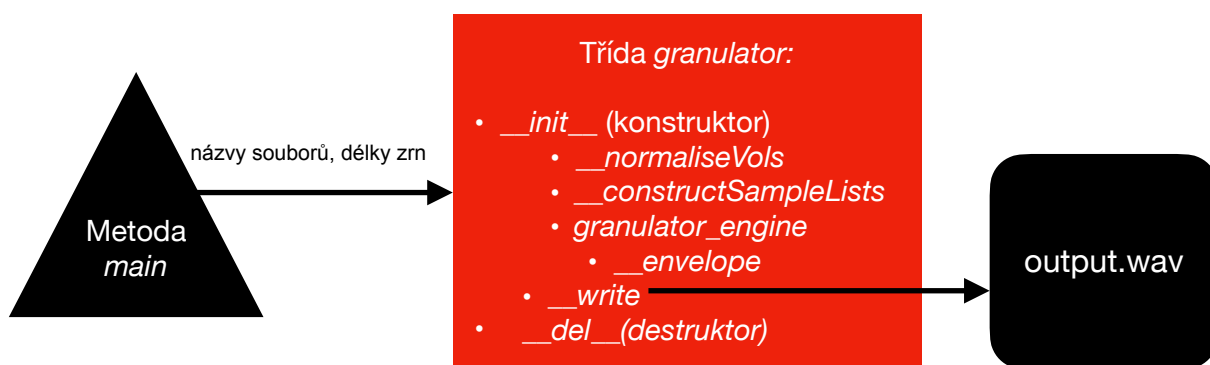
Délku prolínání signálu jsem prozatím stanovil na 30 vzorků, z čehož vyplývá minimální délka zrna 60 vzorků (asi 1,36 ms při vzorkovací frekvenci 44,1 kHz). Tato délka byla stanovena z důvodu umožnění práce na dolní hranici granulární časové osy. Délka prolínání se dá též jednoduše změnit v kódu programu.

8.3. Popis kódu vlastního vícesamplového granulárního syntezátoru

Před hlubším ponořením do problému je dobré nastínit základní strukturu programu.

Samotný program obsahuje pouze dva soubory: hlavní soubor *main.py* a třídu *granulator.py*. Obsahuje také složku *audiofiles*, kam uživatel vloží vstupní soubory pro granulaci. Ovládání programu je prozatím ponecháno v rámci rozhraní příkazové řádky. Kód programu jsem opatřil poznámkami v anglickém jazyce.

Jak již bylo výše uvedeno, při programování jsem zvolil metodu zapouzdření do jedné třídy, zvané *granulator*. Instance této třídy s dvěma parametry (seznam souborů a seznam korespondujících délek zrn v samplech) se volá v hlavní funkci *main*, která celý program provádí.



Obrázek 8.3.1.: Schéma postupu programu

Funkce *main* obsahuje pouze několik funkcionalit. Jako první vypíše všechny dostupné soubory ve složce *audiofiles*. Následuje skript *read_input*, ve kterém

²³ Practical guidelines for Production and Implementation in accordance with EBU R 128 [online]. Copyright ©S [cit. 01.04.2022]. Dostupné z: <https://tech.ebu.ch/docs/tech/tech3343-v2.pdf>

uživatel vybere libovolný počet z nich pro granulaci zadáním jejich názvů oddělených čárkou a následně pro každý vybraný soubor uvede délku zrna. Tímto lze docílit vícesamplové granulace v různých poměrech (např. 1:2, 2:3:4 etc.). Poté je s těmito parametry (názvy souborů a délky zrn) volána instance třídy *granulator*. Po zavolání destrukturu třídy *granulator* se program ukončí závěrečnou hláškou.

Třída *granulator* má pouze jedinou funkcionalitu - a to vícesamplovou granulaci na výše uvedeném způsobu. Pracuje tedy sama o sobě a veškerá její funkcionality je provedena rovnou v konstrukturu. Po provedení konstrukturu se instance okamžitě ukončí destruktorem, který vypíše hlášku o délce výstupu v sekundách.

Konstruktor samotný volá několik metod za sebou. Jako první je třeba normalizovat úroveň vstupních souborů. Pak se soubory pro další práci převedou na seznam za sebou jdoucích bitových hladin. Následně je možné soubory granulovat a seřazovat jejich zrna za sebou do nového, výstupního seznamu. Nakonec je nutné výstupní seznam opět převést na audio soubor.

8.3.1 Příprava souborů ke granulaci

Nejdříve tedy konstruktor (za účelem eliminace v minulé kapitole uvedených modulačních artefaktů) volá metodu `__normaliseVols`. Tato metoda využívá externí knihovny *Pydub* za účelem normalizace hlasitosti na úroveň -18 dBFS. Vstupní soubory normalizuje a uloží pod jiným názvem. Tyto soubory jsou v zápětí otevřeny pomocí metody `wave.open` jako objekt třídy `wave_read` a uloženy do seznamu. Tento seznam se předává konstrukturu, který ho ukládá do proměnné `normalisedFiles`:

```

def __normaliseVols(self, wavfiles):
    wavfilesNormalized = []
    for wavfile in wavfiles:
        sound = AudioSegment.from_file('audiofiles/'+ wavfile, "wav")
        change_in_dBFS = -18.0 - sound.dBFS
        wavfileNormalized = sound.apply_gain(change_in_dBFS)
        name = 'audiofiles/'+ wavfile + 'normalised.wav'
        wavfileNormalized.export(name, format="wav")
        wavfilesNormalized.append(wave.open(name, 'rb'))
    return wavfilesNormalized

```

Spolu s tímto úkolem konstruktor do proměnné `grainSizes` uloží vstupní seznam délek zrn v audio vzorcích zadaných uživatelem ve funkci `main`.

Následuje funkce `_constructSampleLists`, která převádí audio data vrácená metodou `readframes` objektu `wave_read` z každého souboru zvlášť postupně po vzorcích nejprve na dvojici bitových hladin (stereo vzorek ve formátu 2x integer - celé číslo) pomocí funkce `struct.unpack`. Tuto dvojici následně vydělí patřičným počtem bitových hladin (v závislosti na bitové hloubce souboru zjištěné metodou `getparams[1]`) za účelem získání univerzálního, na bitové hloubce souboru nezávislého seznamu čísel s plovoucí desetinnou čárkou (float) přepočítaného do rozmezí (-1,1). Tuto přepočítanou dvojici uloží do lokální proměnné `sampleList`. Takto program projde celý soubor a na konci uloží dílčí `sampleList` do třídní proměnné `self.sampleLists`. Toto se provede postupně se všemi vstupními soubory, čímž získáme seznam seznamů dvojic bitových hladin přepočítaných do typu float v rozmezí (-1,1).

Jak lze vidět níže, program pracuje pouze s formáty 8-bit, 16-bit a 24-bit. Formát 32-bit byl z důvodu jeho široké nekompatibility s externím audio softwarem vyřazen.

```

def __constructSampleLists(self):
    for i in range(len(self.normalisedFiles)):
        print 'audio file', i+1, 'frame count',
self.normalisedFiles[i].getnframes()

        sampleList = [] #for every file we annulate the sample list
which is then saved to sampleLists variable (before moving to a next file)

        for sc in range(self.normalisedFiles[i].getnframes()):
            frame = self.normalisedFiles[i].readframes(1)
            data = struct.unpack('<hh', frame)
            if (self.normalisedFiles[i].getparams()[1] == 1):
                t1 = tuple(float(ti) / 127 for ti in data) #8bit
            elif (self.normalisedFiles[i].getparams()[1] == 2):
                t1 = tuple(float(ti) / 32768 for ti in data) #16bit
            elif (self.normalisedFiles[i].getparams()[1] == 3):
                t1 = tuple(float(ti) / 8388607 for ti in data) #24bit
            else:
                print 'Unsupported bit
depth',self.normalisedFiles[i].getparams()[1] * 8,'bit in file', i+1

            sampleList.append(t1)
self.sampleLists.append(sampleList)

```

Jelikož máme soubory nyní rozřazené do seznamů v proměnné *self.sampleLists*, můžeme zjistit poslední parametr potřebný pro samotnou granulaci, a tou je délka výstupního souboru. Pomocí délky výstupního souboru zjišťujeme, kde se v granulaci nacházíme a můžeme tím proces granulace před vyčerpáním materiálu včas ukončit a přejít k zápisu výstupního souboru.

Výstupní soubor nesmí být totiž z logiky věci uvedené na konci kapitoly 7.2.1. delší, než nejkratší vstupní soubor: pokud bychom se dostali na konec nejkratšího souboru a chtěli dále pokračovat v granulaci, nebyl by na část z ní již materiál a program by vypsal chybu. Je tedy nasnadě zvolit jako maximální délku souboru právě onu délku souboru nejkratšího. Se zavedením prolínání jednotlivých zrn a tím způsobeným zkrácením celého signálu se však situace komplikuje. Je třeba nejprve sečíst délky samotných prolínání a ty pak od délky nejkratšího souboru ještě odečíst. Za účelem tohoto byla na začátku třídy granulátor zadefinována konstanta *self.CFlen*, obsahující délku křížového prolínání ve vzorcích. Tato konstanta se pronásobí s počtem zrn (bez jedné, poslední zrno prolínáčku nemá) nejkratšího souboru (metoda *__getTotalNumOfGrains*) a následně se celé číslo odečte od délky nejkratšího souboru (ve vzorcích) a výsledek se uloží do proměnné *self.outputSampleCount*.

```

self.outputSampleCount = len(min(self.sampleLists, key=len)) -
((self.__getTotalNumOfGrains() - 1) * self.CFlen)

```

Nyní je vše připraveno pro samotnou granulaci, konstruktor tedy volá metodu *granulator_engine*.

8.3.2 Granulační proces: metoda *granulator_engine*

Metoda *granulator_engine* je nejrozsáhlejší metodou programu a je zodpovědná za samotný proces vícesamplové granulace. Na začátku metody se definují dvě provázané lokální proměnné určené pro měření zbývajících vzorků audia pro granulaci - *sampsElapsed* s hodnotou 0 a *sampsRemaining* s hodnotou *self.outputSampleCount*. Včasné ukončení granulace zajišťuje cyklus *while* s podmínkou *sampsElapsed < self.outputSampleCount*.

Při dodržení této podmínky metoda cyklicky prochází zadané soubory. Pro každý soubor prvně vytvoří prázdný seznam *grain*, do kterého se následně budou ukládat vzorky jednotlivého zrna. Před granulací program ještě zkontroluje, že do konce zbývá více vzorků, než je zadaný počet vzorků zrna (délka zrna) daného souboru. V případě, že je toto splněno, nejsme na konci souboru a můžeme pokračovat dále (část *else* podmínky).

V každém souboru se nejprve program přesune na určené místo (viz kapitola 7) dle proměnné *sampsElapsed* a následně se jednotlivé vzorky od tohoto místa čtou a postupně zapisují do seznamu *grain* až do vyčerpání zadané délky zrna.

```
for sc in range(self.grainSizes[filenum]):  
    grain.append(self.sampleLists[filenum[sc+sampsElapsed]])
```

Následně se hotovému zrně pomocí metody *__envelope* přiřadí obálka.

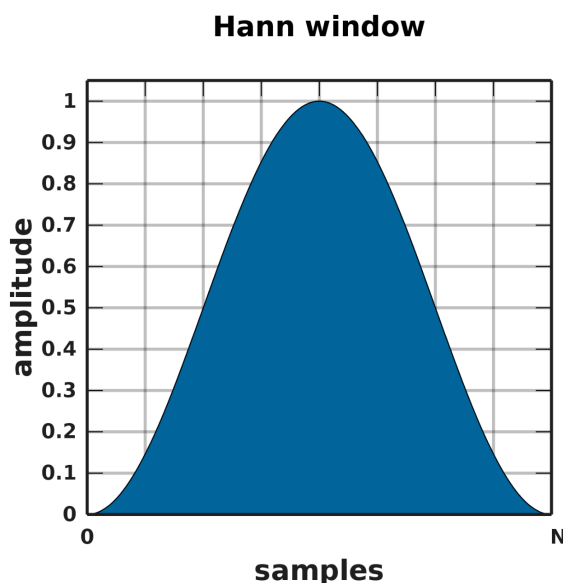
Obálka je zvolena ve tvaru zhruba uvedeném na Obr. 7.2.1.1. Tohoto tvaru bylo docíleno vnořením plochy stejné maximální hlasitosti mezi nelineární náběh a doběh tzv. Hanningova okénka. Toto běžné okénko užívané pro Fourierovu analýzu jsem zvolil jednak proto, aby byla eliminována linearita náběhu, která by mohla kvůli svým ostrým hranám zvyšovat podíl spojitého spektra na výstupu, jednak pro to, že se jedná o jednu z funkcí knihovny NumPy, která vrací diskrétní množinu čísel s průběhem Hanningova okénka o délce stanovené uživatelem. Délka Hanningova okénka je konstantní a rovná se dvojnásobku stanovené délky křížového prolínání, délka plochy stejné maximální hlasitosti se pak rovná délce zrna bez okrajových prolínaček (délka celého Hanningova okénka):

```
hanning = np.hanning(self.CFlen * 2) #a fixed crossfade length  
hanning = hanning.tolist()  
fadeIn = hanning[:len(hanning)//2] #we use a divided hanning envelope for  
fade in and fade out  
fadeOut = hanning[len(hanning)//2:]
```

```
sustain = [1] * (grainsize - len(hanning)) #length of sustain = total  
length - length of fades
```

```
env = np.array(fadeIn + sustain + fadeOut)
```

Okénko se na zrno následně navlékne pomocí pronásobení bitových hladin jednotlivých vzorků s hodnotami obálky uložené v proměnné *env*. Toto se musí však dít pro každý kanál sterea zvlášť, proto nejprve zrno rozdělíme na levý a pravý kanál. Tyto kanály pak pomocí knihovny NumPy násobíme zvlášť každý proměnnou *env*. Oba kanály poté opět sesadíme dohromady a vracíme zpět do metody *granulator_engine*.



Obr.: 8.3.2.1 Hanningovo okénko

Nyní je již zrno připravené k připojení do výstupní proměnné *self.outputList*. Nelze však opomenout prolnutí zrna se zrnem předchozím. Nejprve však program pomocí podmínky *if not self.outputList* (výstupní seznam je prázdný) zjistí, nacházíme-li se na začátku souboru. Pokud ano, prolnutí se neprovede, neb zrno stojí jako první a stačí ho rovnou do výstupní proměnné zapsat. Pokud ne, je již třeba prolínání provést.

Prolnutí se provede ve dvou krocích. V prvním se opět pomocí NumPy sečte konec (doběh) obálky posledního zrna v proměnné *self.outputList* s náběhem obálky zrna připojovaného. V druhém kroku se připojí zbytek zrna. Počet zapsaných vzorků se přičte k celkovému počtu do proměnné *sampsElapsed*, která následně promění i obsah proměnné *sampsRemaining*.

V případě, že se již blížíme konci souboru však může nastat výše nastíněná situace - délka dalšího zrna (ve vzorcích) k zapsání již převyšuje zbývající počet vzorků daný proměnnou *sampsRemaining*. V tomto případě se granulace ukončí a zbývající vzorky jsou vyřazeny.

8.3.3 Zápis do výstupního souboru

Po návratu z metody `granular_engine` se dostáváme zpět do konstruktoru, který následně volá metodu `__write`.

Tato metoda nejprve zadefinuje parametry výstupního audia a následně soubor s těmito parametry vytvoří pomocí funkce `wave.open` jako objekt `wave_write`.

Dále již metoda slouží jako inverzní metoda k metodě `__constructSampleLists` - převádí náš seznam floatů postupně po vzorcích zpět na seznam integerů bitových hladin a tyto pak opět pomocí funkce `struct.pack` převádí na audiodata v binární formě, která nakonec pomocí metody `writeframesraw` objektu `wave_write` zapisuje do výstupního souboru.

```
def __write(self):
    outputFile = wave.open('output.wav', 'wb')
    params = (2, 2, 44100, self.outputSampleCount, "NONE", "none")
    outputFile.setparams(params)
    for i in range(len(self.outputList)):
        s11 = tuple(float(ti) * 32768 for ti in self.outputList[i])
        wd = struct.pack('<hh', int(s11[0]), int(s11[1]))
        outputFile.writeframesraw(wd)
```

Nyní jsme došli na konec konstruktoru. Zde se ihned volá destruktorka `__del__`, který smaže dočasné, hlasitostně normalizované soubory a instanci třídy `granulator` ukončí.

Tím se dostáváme na konci funkce `main` a tudíž i na konec celého programu.

9 Analýza výstupních dat programu

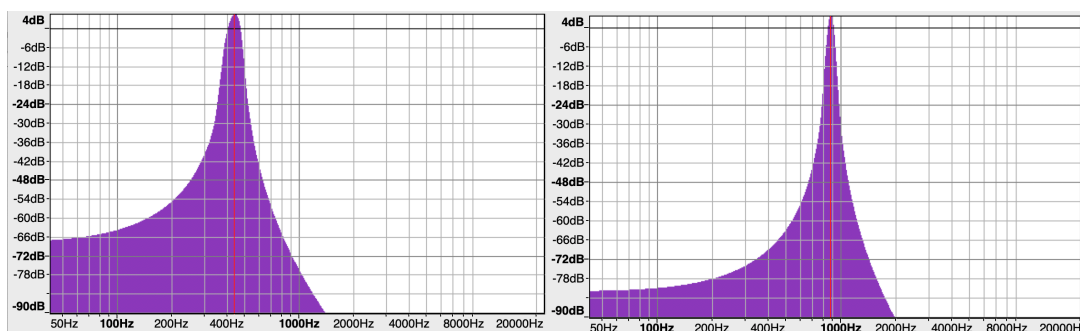
Vzhledem k tomu, že vícesamplový přístup v granulární syntéze je víceméně neprozkoumaným fenoménem, je dobré jeho výstup také teoreticky analyzovat, čímž lze přispět k odbornému zavedení tohoto fenoménu do povědomí. Na základě této analýzy bude pak také možné i lépe porozumět jevům, které vznikají při kreativní práci s vícesamplovou granulací.

Za účelem analýzy výstupu je však nutné dodat vstupní materiál určité kvality. Před pokusem o analýzu granulární kombinace více složitějších signálů je však nasnadě začít se signály elementárními, tedy sinusovými - neb na jejich jednoduché struktuře lze nejlépe vypořádat změny provedené granulací. Výsledky analýzy sinusových průběhů na různých parametrech nám posléze pomohou ujasnit, co se děje při zadávání vstupu se složitějšími průběhy.

9.1 Čistě sinusový vstupní signál

Jako základ naší analýzy výstupu jsem zvolil dva sinusové signály vzdálené o jednu oktávu (440 Hz a 880 Hz). Tyto signály o délce 30 sekund jsem vygeneroval v programu Audacity. Poměr zrn jsem zvolil pro jednoduchost prozatím 1:1 - Pro účely analýzy barevných variací se zatím budeme soustředit spíše na další granulační parametry - délku zrna a délku prolínání.

Spektrogram byl za účelem analýzy vykreslen v programu Audacity za použití rychlé Fourierovy transformace²⁴ s Hanningovým okénkem.



Obr 9.1.1: Výsledek FFT pro oba zvolené
sinové signály (zleva doprava 440, 880 Hz)

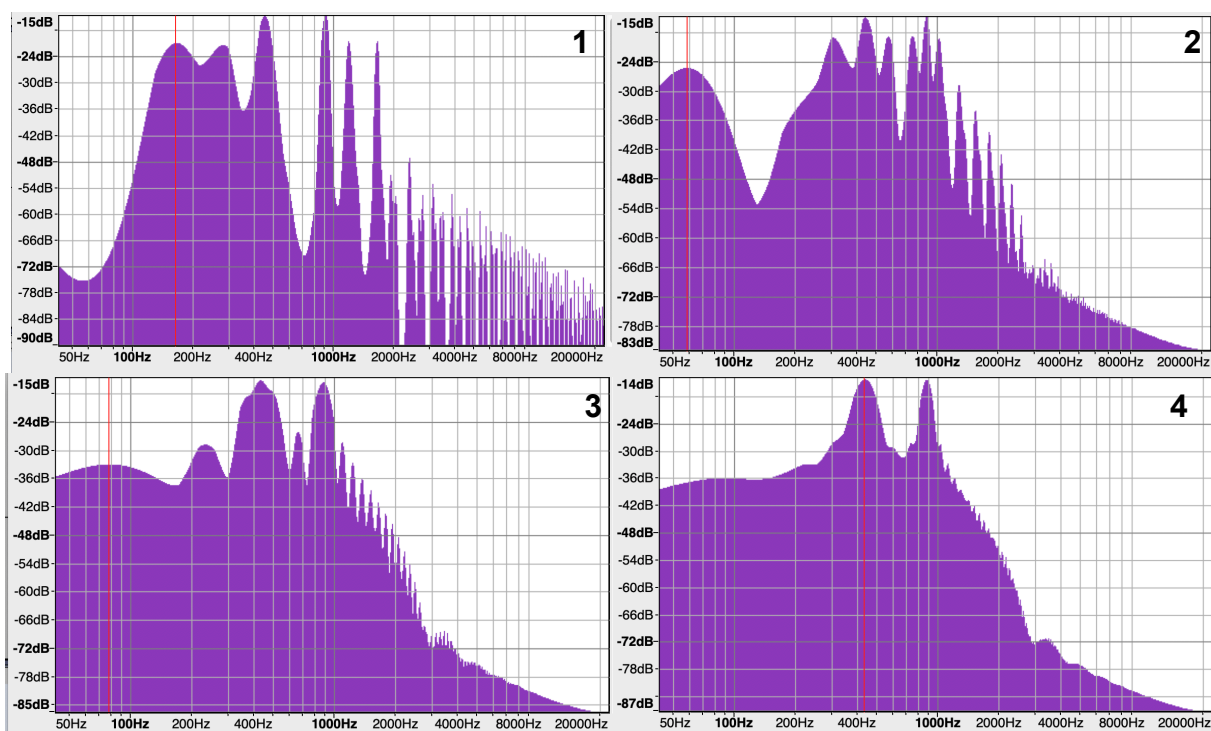
²⁴ Plot Spectrum - Audacity Manual. *Audacity Manual* [online]. Dostupné z: https://manual.audacityteam.org/man/plot_spectrum.html

9.1.1 Vliv délky zrna a prolínání na výstup

Délka prolínání byla prozatím ponechána původní, vzorkovací frekvence byla ponechána na 44,1 kHz.

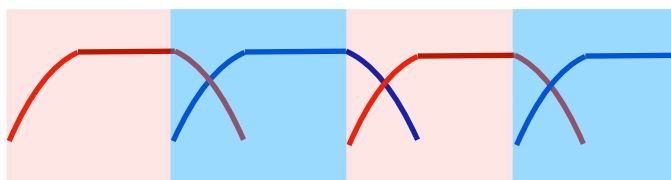
Programu jsem zadal dva výše zmíněné sinusové průběhy a postupně zvyšovat délku zrna. Začal jsem nejmenší možnou délkou tak, aby si výsledné zrno zachovalo alespoň jeden vzorek plné amplitudy signálu (neovlivněný prolínacím okénkem), která činí 61 (2x délka prolínání + 1). Pokračoval jsem s délkami zrna 100 (1 perioda signálu 440 Hz), 250 (2,5 periody) až do výše 500 vzorků (5 period).

První věc, kterou lze z výsledků vidět je ta, že s rostoucí délkou zrna roste poměr spojitého spektra. Při délce zrna 500 vzorků (22,05 ms při dané vzorkovací frekvenci) je již také patrný jev popsáný v kapitole o mikrozvuku, tedy určitá diskontinuita zvukové barvy (která s dále zvyšující se délkou zrna přechází až ve vjem dvou střídavě za sebou jdoucích signálů). U délky zrna 61 vzorků lze pozorovat kromě (oproti zbytku) nedokonale spojitého spektra také určité ostré neharmonicity (kromě násobků základních frekvencí jsou zde patrné i špičky, které se těmto násobkům nerovnají). Při poslechu tohoto výstupu se tento jev projevuje „zvonovitostí“ zvuku, tedy neharmonickým projevem typickým pro některé idiofonické nástroje. Tento jev jednotlivých neharmonických špiček s rostoucí délkou zrna ustupuje ve prospěch spojitého spektra, neharmonicity jako takové se tím však nezabavujeme.



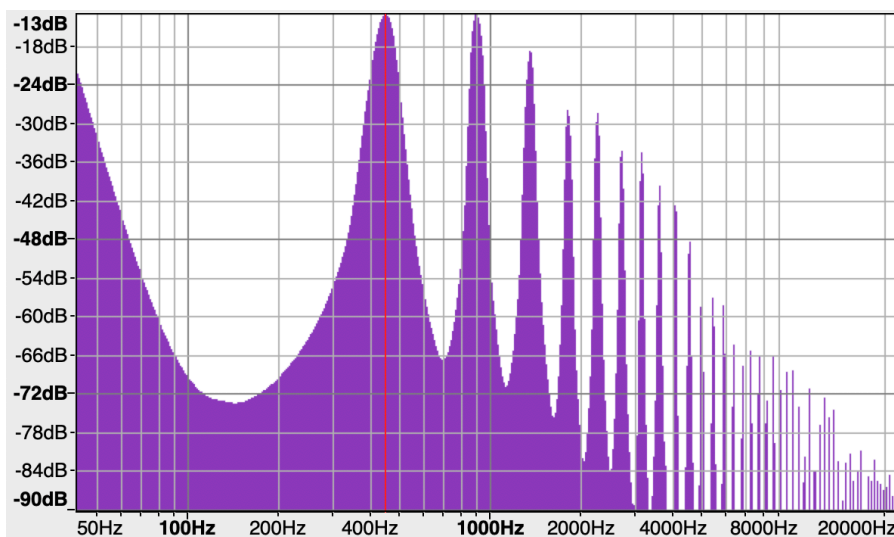
Obr 9.1.1.1: Spektrogramy pro délky zrn 61, 100, 250 a 500 vzorků

Neharmonické frekvence jsou zde s největší pravděpodobností tvořeny periodickým střídáním zrn po určité vzdálenosti (vnitřní periodičita granulace), které (i přes normalizaci hlasitosti a prolínání) působí jako oscilátor. Správnou volbou této periodičity v závislosti na vstupních datech pak lze tuto periodičitu synchronizovat s frekvencí vstupního materiálu (pokud je materiál tónový a má stejnou, či celočíselně násobnou výšku jako vstupní materiál), čímž se ve výstupu výše zmíněné neharmonicity eliminují.



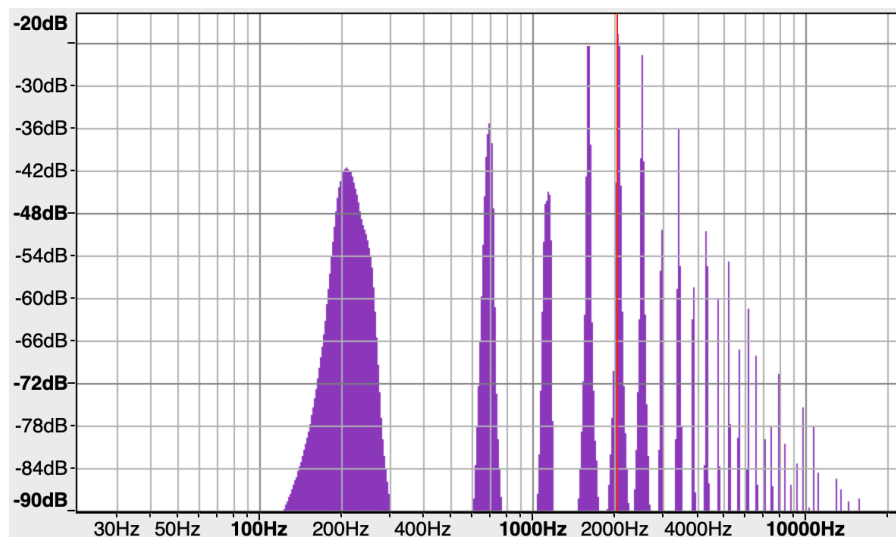
Obr 9.1.1.2 Periodičita opakování obálky

Jako důkaz k tomu může sloužit následující experiment: Ve výše zmíněné granulaci dvou sinusových průběhů (440 Hz a 880 Hz při délce zrna 61 samplů) synchronizujeme periodičitu zrn s délkou periody signálu 880 Hz (asi 1,13 ms, což odpovídá v tomto případě asi 50 vzorkům) tak, aby délka zrna bez jednoho prolínání (periodičita opakování obálky, viz obr, 9.1.1.2) této periodě odpovídala. Tohoto lze dosáhnout změnou délky prolínání tak, aby celková délka zrna bez jednoho prolínání byla oněch 50 vzorků. V našem případě se tedy bude jednat o délku prolínání 11 vzorků ($61 - 11 = 50$). Neharmonicita signálu tímto téměř zmizí a dostáváme harmonické spektrum s výrazně ubývajícími amplitudami harmonických nad 3. harmonickou (součet frekvencí vstupních souborů, viz obrázek 9.1.1.3). Při změně jednoho z obou parametrů tento stav zaniká ve prospěch neharmonicit se základní frekvencí rovnou převrácené hodnotě délky výše zmíněné periody. Spodní maximum na obrázku 9.1.1.3 značí záznam (okolo 2 Hz), vzniklé zaokrouhlením na základě nemožnosti přesně pomocí diskrétních vzorků zachytit periodu signálu 880 Hz.



*Obr 9.1.1.3: Spektrogram pro délku zrna 61
vzorků a délku prolínání 10 vzorků*

Tento experiment nám však ještě nic neříká o obecnosti využití tohoto přístupu, ani o důvodu vzniku těchto dalších harmonických frekvencí. Pravdivost tohoto experimentu a především důvod vzniku harmonického spektra lze jednodušeji zjistit, necháme-li výstupní parametry (délku zrna a prolínání) stejné, jako výše, avšak jako vstupní materiál pro granulární kombinaci dodáme jednu konkrétní, na vnitřní periodicitě nezávislou frekvenci a ticho. Pro tento příklad jsem zvolil frekvenci 2 kHz. Výsledný spektrogram (obr. 9.1.1.4) ukazuje, že výše zmíněná frekvence 880 Hz, způsobená rozdílem délky zrna a délky prolínání se zde projevuje jako vzdálenost mezi jednotlivými špičkami, vycházejícími z maxim utvořených okolo vstupní frekvence 2 kHz. Tyto maxima (cca 1560 a 2440 Hz) jsou součtovou a rozdílovou složkou amplitudové modulace o frekvenci 440 Hz (dvojnásobná délka periodicity jednoho zrna - v případě granulace dvou složek do stejně dlouhých zrn se jedná o stále se opakující periodicitu jedné sekvence zrn = *zrno 1* -> *zrno 2*). Nutno dodat, že toto již neplatí pro tří- a vícesamplovou granulaci, ani pro granulaci s rozdílným poměrem zrn, neb zde není součet délek jedné sekvence zrn dvojnásobkem periodicity jednoho zrna.



*Obr. 9.1.1.4: Spektrogram granulační směsi
sinusového signálu 2 kHz a ticha*

Ve výše uvedeném případě (obr. 9.1.1.3) tedy ono harmonické spektrum vzniklo tím, že spodní pásmo amplitudové modulace o frekvenci 440 Hz na signálu 880 Hz se překrylo s původním signálem 440 Hz a další maxima s jeho lichými harmonickými frekvencemi. Lze také vidět, proč se při delší délce zrna zvyšuje poměr spojitého spektra - délka periody se zvětšuje, tudíž modulační frekvence je nižší a tím jsou vznikající postranní pásma a jejich celočíselné násobky a podíly blíže u sebe, jak ostatně lze vidět i z obrázku 9.1.1.1.

Z experimentů se sinovými signály tedy vyplývá jedno - pokud chceme vícesamplovou granulární syntézu využít za účelem čistého směšování barvy dvou tónových průběhů, je nasnadě algoritmu zadávat materiály o stejné tónové výšce a vstupní parametry nastavit tak, aby vnitřní periodičita odpovídala periodicitě tónové výšky vstupu. Rozladění dvou signálů o stejném tónu - například z důvodu nepřesné intonace hráče či rozdílného ladění nástroje - však může na výstupu způsobovat výraznou modulaci typu záněhů. Na toto je třeba vždy pamatovat a případně zvolit kompromis vstupu rozdílných tónových výšek za cenu následného vzniku neharmonicit.

Nakonec nutno však také podotknout, že jistá neharmonicitu hraje v granulární syntéze důležitou estetickou roli a tuto roli není možné opomíjet. Proto je toto naše zjištění spíše než pro praktickou aplikaci důležité pro lepší porozumění jevům vznikajícím ve výstupu v závislosti na vstupních parametrech. V multisamplové granulační praxi za účelem kreativní zvukové tvorby také jistě budeme raději chtít mít možnost libovolně operovat s poměry a délkami zrn (za účelem nastavení poměrů vstupních souborů ve výstupu), či libovolně využívat vstup dvou a více

souborů, což hledání výše zmíněného ekvilibria ve většině případů stejně znemožňuje, či alespoň výrazně ztěžuje.

9.2. Vstup s komplexním spektrem

Po analýze jednoduchého sinusového vstupu můžeme přejít k analýze vstupu komplexního - sestávajícího z více frekvenčních složek. Takovýto vstup může být jakýkoliv - od signálů tónových, přes neharmonické až hlukové perkusivní zvuky až po hlukové atmosféry.

9.2.1 Obecný vliv granulace na tónový charakter výstupu

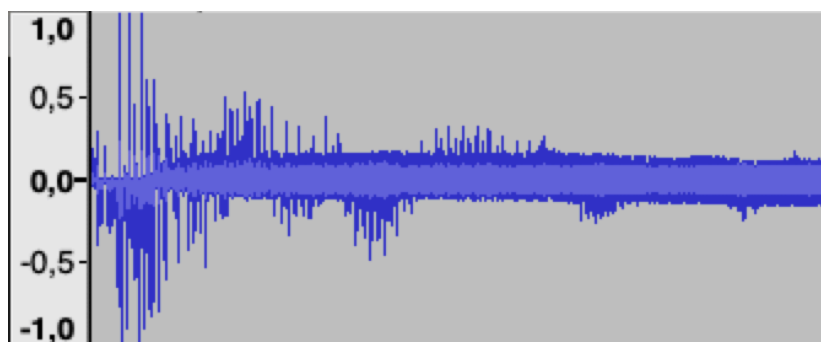
U signálů s bohatším harmonickým spektrem platí pro každou harmonickou frekvenci zvláště to, co jsme popsali v kapitole minulé pro jednoduchý sinusový signál. Granulární syntézou proto vznikají signály se značně bohatým spektrem. Tato spektrální bohatost při vícesamplovém vstupu samozřejmě ještě vzrůstá. Tónový charakter vstupu lze za speciálních okolností uvedených v předchozí podkapitole na výstupu částečně zachovat, avšak jistým neharmonicitám v signálu se u reálného vstupu nikdy nevyhneme.

I u signálů hlukových a perkusních se granulace projevuje kvazitónově - spektrálním zabarvením způsobeným zvýšeným podílem frekvencí způsobených periodicitami v granulaci ve výsledném spojitém (hlukovém) spektru. Eliminace tohoto zabarvení je možná pouze určitou randomizací (náhodným měněním) délek zrn, či alespoň volbou odlišných, harmonicky nezávislých délek zrn (period) pro každý vstupní soubor. Další možností je volit délku zrna tak velkou, aby převrácená hodnota trvání periodicit v signálu (frekvence) byla již pod spodní hranicí slyšitelného pásma.

9.2.2 Vliv obálky vstupních signálů na granulární výstup

Narozdíl od signálů sinusových mají běžně se vyskytující (např. v hudebních nástrojích) signály určitou obálku, sestávající (obecně řečeno) z náběhu, průběhu a doběhu. Charakter signálu je často z velké části určen jeho náběhem (viz kapitola 1), při kterém i v tónových signálech vznikají tranzientní, netónové jevy s komplexním spektrem. Experimenty s programem bylo zjištěno, že charakter náběhu vstupních signálů značně určuje charakter výstupu.

Vstup signálů s perkusním náběhem (tón klavíru, kytary, úder do bubnu etc.) kombinovaný se signály s hladším náběhem (např. dechové nástroje) se i při velmi malém poměru (viz obrázek 9.1.2.1) ve výsledné granulaci projeví perkusním charakterem výstupu. Na tento fakt je dobré pamatovat pokaždé, granulujeme-li více signálů, mezi nimiž se nachází alespoň jeden perkusivní signál. Tento signál významně ovlivní charakter výstupu směrem k perkusivnímu charakteru.

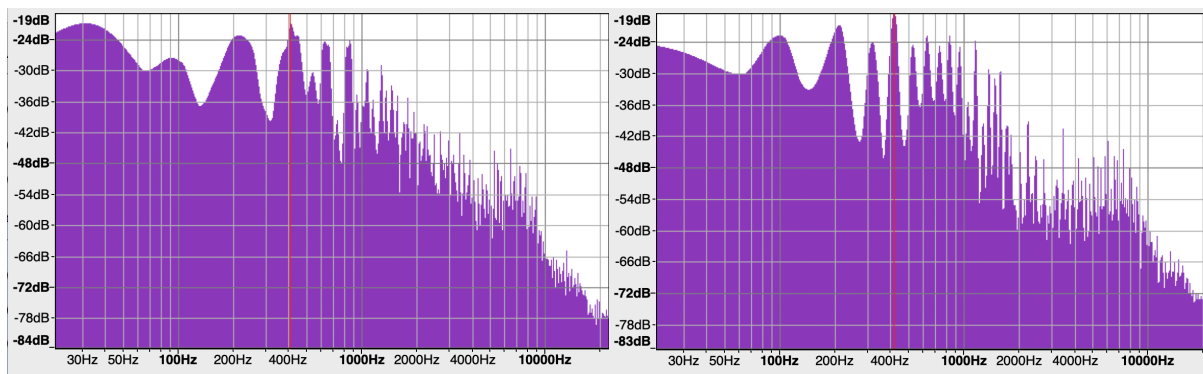


Obr. 9.1.2.1: Časový průběh granulární směsi tónů a1 klavíru a klarinetu při granulačním poměru 10:100 vzorků

Pokud se chceme věnovat čistě barevným aspektům zvuku, je dobré pro vícesamplovou granulaci zvolit signály s podobným náběhem, neb ve výsledném signálu vždy nakonec bude přetrvávat charakter signálu s náběhem výraznějším (viz předchozí odstavec).

9.2.2.1 Kombinace signálů perkusního charakteru

Při granulární kombinaci dvou signálů s výrazným perkusním náběhem jsem během experimentů zjistil, že díky významnému rozdílu amplitud v oblasti náběhu a zbytku signálu vzniká při normalizaci nepoměr, který způsobuje v oblasti náběhu přebuzení audia. Přebuzení vzniká v důsledku součtu rychlých neperiodických změn v náběhu. Díky tomuto zjištění jsem se nakonec rozhodl normalizaci hlasitosti upravit z -18 na -18 dBFS mínus dvojnásobek počtu vsupních souborů (pro každý další soubor se úroveň normalizace sníží o 2 dB), což zajišťuje dostatečnou rezervu pro tento typ vstupu. Dostatečnou ochranu proti přebuzení zajistí také delší obálka, která má mírnější náběh. Delší obálka také zajistí menší poměr spojitého spektra (především na vysokých frekvencích), čímž dojde k určitému vyhlazení signálu a redukci distorsní, šumové složky vzniklé rychlým prolínáním (sčítáním) hlukových tranzientních jevů obou signálů (viz obr. 9.1.2.1.1).



Obr. 9.1.2.1.1: Úbytek spojitého spektra a neharmonicit při zvýšení délky prolínání z 11 (vpravo) na 30 vzorků (vlevo) u granulární kombinace tónu a1 kytary a klavíru

Nutno však opět podotknout, že pro kreativní práci s výstupem však lze využít naopak krátkou obálku a přebuzení jako vlastní estetickou kvalitu.

9.2.3 Vliv poměrů délek zrn na barevný charakter výstupu

Poslední, dosud opomíjený vstupní parametr, který je granulátoru možné zadat, je poměr mezi délkami jednotlivých zrn, zadatelný pomocí přiřazení různých délek zrna každému souboru zvlášť dle zvoleného poměru. Tento parametr byl zaveden především pro možnost zvyšovat, či snižovat poměr barvy zvuku jednoho ze vstupních souborů ve výstupu a je parametrem pro kreativní využití asi nejzajímavějším. Vzhledem k tomu, že se jedná o asi nejvíce subjektivní fenomén, budu zde vycházet především z vlastní poslechové praxe.

Během experimentů jsem zjistil, že při kombinaci více signálů s rozdílnými spektry signál s bohatším spektrem výrazně proniká do výstupu s podobnou razancí, jako tomu bylo u signálů s více perkusivním náběhem, tedy i při malém poměrovém zastoupení. Jako příklad uvedu tón trubky a klarinetu. Spektrálně bohatý tón trubky (při detailním pohledu na vlnění připomínající pilový průběh) v kombinaci se spektrálně ochuzeným tónem klarinetu (charakteristický zvuk klarinetu vzniká nepřítomností sudých harmonických frekvencí) tvoří na výstupu i při vysokých poměrech zvuku klarinetu ku zvuku trubky zvuk s bohatým frekvenčním spektrem, ve kterém se výsledná klarinetová barva ztrácí, avšak charakteristická barva trubky ne. Toto je způsobeno nejen rozdílem v bohatosti spekter, ale především samotným charakterem granulace. Jak bylo patrné na granulaci sinových průběhů, granulace způsobuje značnou spektrální bohatost sama o sobě a tím nedovolí dokonalé zachování barev charakteristických pro signály typu klarinetu či jiné signály charakteristické nepřítomností některých harmonických frekvencí.

9.2.3.1 Experiment: změna poměru signálu

Zajímavé je taktéž od nuly zvyšovat poměr určitého signálu ku druhému a pozorovat, při jakém minimálním poměru se již začne v onom signálu projevovat signál přidávaný a jak se při dalším zvyšování podílu v poměru jeho kvalita ve výsledné směsi projevuje. Pro tento experiment jsem zvolil dva tóny a1 houslí a trubky, mimo jiné díky jejich podobné, víceméně rovné obálce. Jako výchozí tón jsem zvolil tón houslí, ke kterému postupně přidávám razantnější tón trubky. Jako délku prolínání jsem zvolil 10 vzorků, aby bylo možno docílit malé délky zrna - v tomto případě je minimum 20 vzorků, na které nastavuji statickou délku zrna pro zvuk trubky.

Jako výchozí poměr jsem zvolil poměr 1:50 (20 ku 1000 vzorků). Při tomto poměru se tón trubky projevuje pouze jako určité tiché praskání na pozadí jinak čistého houslového tónu.

Pokračoval jsem s dvojnásobkem poměru trubky, tedy poměrem 1:25. Při tomto poměru je již praskání hlasitější a při detailním poslechu lze slyšet určitou jemnou modulaci na rovném tónu houslí.

Při dalším zdvojnásobení poměru (1:12,5) je již modulace patrnější a projevuje se jako určité zkreslení signálu podobné přebuzení analogových zesilovačů. Praskání se dále zesiluje.

Při poměru 1:8 praskání přechází již v určitý periodický praskavý signál očividně modulující již značně zkreslený signál houslí.

Při dalším zvýšení poměru trubky na 1:6 je již patrná zcela nová, distorsní zvuková barva, stále však na základě houslového tónu.

Signál trubky se začne vynořovat při poměru 1:3 a při postupném snižování poměru houslí se začne její charakter stále více projevovat.

Subjektivní vjem rovného poměru (subjektivně 1:1) slyším však již okolo poměru 1:2. Při reálném poměru 1:1 je pak patrnější průraznější charakter trubky. Toto je způsobeno pravděpodobně tím, co jsem popsal v úvodu této podkapitoly (9.2.3). Výše uvedené poměry a jejich odpovídající efekt se proto u rozdílných druhů vstupu značně liší.

10 Možné druhy kreativní práce s vícesamplovou granulací

Vícesamplová granulární syntéza nabízí vícero nových, dosud nezkoumaných možností kompozice a zvukové tvorby; její hlavní těžiště však tkví - jak již bylo v této práci mnohokrát zmíněno - především v tvorbě nových zvukových barev vzniklých kombinací vstupního materiálu a specifického spektra vzniklého granulací (a případným prolínáním). Možnostmi, jak pracovat s výstupní barvou je více - nabízí se přístup se statickými či proměnnými (na základě předem definovaných funkcí) vstupními parametry (délka zrna, délka prolínání), či kombinace těchto přístupů při práci v reálném čase. V této kapitole se pokusím alespoň okrajově nastínit pár možných využití tohoto dosud nezkoumaného syntetického přístupu.

Výchozím způsobem granulace výše popsaného programu je granulace mimo reálný čas za statických vstupních parametrů. Jednoduchou implementací funkce délky zrna v čase ohraničeném začátkem a koncem granulace (ve výše popsaném programu je postup časovou osou reprezentován proměnnou *sampsElapsed*) lze však program upravit i pro přístup dynamický. Práce v reálném čase lze také v Pythonu docílit, avšak za tímto účelem lépe poslouží některé z výše zmíněných algoritmických prostředí určených pro práci v reálném čase (Pure Data, MAX/MSP etc.)

10.1 Vícesamplová granulace za účelem tvorby samplů

Druh práce, který využívám v předchozí kapitole je granulace za využití krátkých vstupních samplů (jednotlivých tónů a krátkých zvukových objektů v řádech od setin sekundy až po několik sekund). Využití tohoto typu granulace má smysl především za účelem tvorby nového zvukového objektu (samplu), který lze využít pro následné samplování (práci se samplem).

Pro tento způsob granulace se nejvíce nabízí přístup statický, tedy za stabilních vstupních parametrů. Granulární struktura výstupu se v čase neproměňuje - délky a poměry zrn a prolínaček zůstávají v čase stejné, čímž vzniká krátký statický zvukový objekt působící jako poměrná granulární směs vstupních souborů. Nelze však opomenout ani přístup dynamický - např. rychlá změna poměrů zrn na takto krátkém časovém úseku může vyvolat nečekané barevné proměny. Zde však opět nutno poznamenat, že plynulost proměny délky zrna v čase přímo souvisí se zvolenou vzorkovací frekvencí - pro tento typ práce by tedy bylo nasnadě zvolit pokud možno co nejvyšší vzorkovací frekvenci.

Zajímavou kombinací obou v minulém odstavci zmíněných přístupů by mohla být například syntéza jednotlivých tónů různých výšek (např. rovnoměrně temperované

ladění), přičemž poměry délek zrn by se odvíjely od výšky tónu (např. pro vyšší tóny použít větší poměr perkusivního zvuku etc.).

10.2 Vícesamplová granulace jako alternativa k mixu

Při vícesamplové granulaci delších zvukových útvarů se nám nabízí mnoho zajímavých přístupů, jak s takovým materiálem pracovat, především za využití dynamické proměny vstupních parametrů - nejčastěji délek zrn, avšak lze alterovat i délku prolínání - např. za účelem určitého „zdrsnění“ přechodu mezi jednotlivými zrny.

Tento způsob lze zajímavě využít například v předem připravených kompozicích s větším množstvím stop (např. mix více hudebních nástrojů či ploch) jako určitý druh dynamického mixu. Velký potenciál pro kreativní práci a performanci nabízí v tomto případě granulace v reálném čase, která uživateli granulárního programu umožňuje na živo granulárně míchat vstup z několika zdrojů najednou a tím docílit jakési alternativy k mixu běžnému (prováděnému změnou poměrů hlasitosti). Nutno podotknout, že tato operace bude v reálném čase značně náročná na výpočetní výkon.

Největšího efektu lze tímto způsobem dosáhnout pravděpodobně při relativně statickém vstupu (zvukové plochy, dlouhé tóny). Při proměnném vstupu (například vyvíjející se melodie) nevynikne dokonale ono granulární splývání barev dohromady a výstup může (zvláště při krátkých dobách prolínání a dlouhých délkách zrn) působit pouze jako více nespojitých granulací dohromady.

10.3 Randomizace granulovaného materiálu

Za pozornost stojí také přístup randomizační, tedy přístup náhodného promíchání sekvence zrn různých obsahů (výstupu vícesamplové granulace) do jakési „zrnné směsi“, tedy náhodné posloupnosti. Tímto způsobem, využitelným nejlépe u relativně statických signálů (u vyvíjejících se signálů dostáváme mnohem větší podíl šumu) lze docílit syntézy od vstupu radikálně odlišné barvy s velkým podílem spojitého spektra.

Při tomto přístupu naopak spojitě spektrum stoupá, využíváme-li příliš krátkých délek zrn (při hypotetický nejnižší délce zrna - 1 vzorek audia - by pak došlo k úplné destrukci vstupního signálu směrem k bílému šumu). Aby nebyl ztracen tónový charakter, je nutné dbát na poučku uvedenou v druhé kapitole; tedy že pro tónový vjem signálu výšky pod 1 kHz nám stačí 2-3 periody signálu. U vstupního signálu o frekvenci 440 Hz musíme tedy při vzorkovací frekvenci 44,1 kHz nastavit délku zrna na alespoň 300 vzorků (3 periody), abychom měli jistotu, že tónový charakter na výstupu nezanikne. Z hlediska délek zrn se zde proto pohybujeme v mnohem vyšších číslech, než u předchozích, prolínacích přístupů.

Tento způsob, jak jsem již výše uvedl se hodí spíše pro statictější signály s rovnou obálkou. Při zadání perkusivních signálů do tohoto typu granulace se vysoké amplitudy náběhu náhodně rozprostřou do výstupního souboru, čímž na výstupu vytvoří občasné náhodné špičky amplitudy, čímž se do granulovaného signálu dostává určitý náhodný „rytmus“.

11 Závěr

V této práci jsem se pokusil zavést nový typ zvukové syntézy granulárního typu - granulární syntézy na bázi vícesamplového vstupu. Teoretické zavedení, které tvoří zhruba první polovinu práce nám poskytlo cenné informace nutné nejen k realizaci granulárního programu a následné analýze jeho výstupu, ale taktéž k návrhu kreativních způsobů práce s tímto druhem syntézy.

Vícesamplová granulární syntéza je užitečným nástrojem k tvorbě nových zvukových barev na základě vstupu ve formě více zvukových souborů. Bylo zjištěno, že její výstupní zvuková barva obsahuje složky všech vstupních signálů; oproti prostému míšení těchto signálů však získává navíc ještě typický granulární charakter způsobený nejen rychlým (subpercepčním) střídáním barev signálů (zrn), ale taktéž modulačními frekvencemi vznikajícími na základě tohoto střídání (daného délkou zrn a délkou prolínání), které modulují signály vstupní. Tím vzniká neotřelý a harmonicky bohatý výstup přímo odvozený ze vstupních signálů a jejich barev.

Senam použité literatury

- 1) Analog-to-digital converter - Wikipedia. [online] [cit. 10.03.2022]. Dostupné z: https://en.wikipedia.org/wiki/Analog-to-digital_converter
- 2) AudioStellar. *AudioStellar* [online]. Dostupné z: <https://audiostellar.xyz/>
- 3) BIANCHINI, Riccardo. *Virtual Sound - Sound Synthesis and Signal Processing - Theory and Practice with Csound*. London: Contemponet, 2008.
- 4) Fluid Corpus Manipulation. *Fluid Corpus Manipulation* [online]. Copyright © [cit. 17.03.2022]. Dostupné z: <https://www.flucoma.org/>
- 5) GABOR, Dennis. „Lectures on communication theory." Technical Report 238, Research Laboratory of Electronics. Cambridge, Massachusetts: Massachusetts Institute of Technology, 1952.
- 6) Granular Synthesis. *Simon Fraser University* [online] [cit. 28.01.2022]. Dostupné z: <https://www.sfu.ca/~truax/gran.html>
- 7) GUŠTAR, Milan. *Elektrofony: historie, principy, souvislosti*. Praha: Uvnitř, 2008. ISBN 978-80- 239-8447-7.
- 8) JANOUŠEK, Ivo. *ABC akustiky pro hudební praxi*. Praha: Supraphon, 1979. Příručky ABC.
- 9) PETRÁČKOVÁ, Věra – KRAUS, Jiří. *Akademický slovník cizích slov: [A-Ž]*. Praha: Academia, 1997. ISBN 80-200-0607-9.
- 10) ROADS, Curtis. *The Computer Music Tutorial*. Mass.: MIT Press, 1996. ISBN 978-0262680820
- 11) ROADS, Curtis. *Microsound*. Cambridge, Mass.: MIT Press, 2001. ISBN 0-262-18215-7.
- 12) SCHAEFFER, Pierre: Akuzmatika. In: Ch. Cox, D. Warner (ed.), I. Koska, P. Zagar (přel.) *AUDIOKULTÚRA*. Texty o modernej hudbe. Hudobné centrum, Bratislava 2013.
- 13) Sound In A Nutshell: Granular Synthesis. *Granular Synthesis: Resource Site* [online] [cit. 10.03.2022]. Dostupné z: <http://www.granularsynthesis.com/hthesis/gabor1.html>
- 14) STOCKHAUSEN, Karlheinz. *Texte zur Musik*. Dieter Schnebel(ed.); Kolín n. R.: Verlag M. DuMont Schauberg, 1971
- 15) SYROVÝ, Václav – GUŠTAR, Milan. *Malý slovník základních pojmů z hudební akustiky a hudební elektroniky*. 2. vyd. V Praze: Akademie múzických umění, 2012. ISBN 978-80-7331- 237-4.
- 16) SYROVÝ, Václav. *Hudební akustika*. 2., dopl. vyd. V Praze: Akademie múzických umění, 2008. Akustická knihovna Zvukového studia Hudební fakulty AMU. ISBN 978-80-7331-127-8.
- 17) SYROVÝ, Václav. *Hudební signál a jeho syntéza* [online]. s. 169 [cit. 2022-02-24]. Dostupné z: https://www.ziva-hudba.info/files/2014/04/140403201437_pdf_0.pdf
- 18) SYROVÝ, Václav. *Hudební zvuk: příspěvek k teorii zvukové tvorby*. V Praze: Akademie múzických umění, 2009. Akustická knihovna Zvukového studia Hudební fakulty AMU. ISBN 978-80-7331-161-2.
- 19) VANIŠ, Jan. *Metody modulační syntézy zvuku*. Praha, 2020. Bakalářská práce. AMU. Vedoucí práce Ing. Milan Guštar, Ph.D.
- 20) WAV - Waveform Audio File Format . *File Format Docs* [online] [cit. 9.3.2022]. Dostupné z: <https://docs.fileformat.com/audio/wav/>

- 21) wave — Read and write WAV files — Python 3.10.2 documentation. 302
Found [online]. Copyright © [cit. 28.01.2022]. Dostupné z: <https://docs.python.org/3/library/wave.html>
- 22) Python 3.10.4 Documentation [online]. Copyright © [cit. 30.03.2022]. Dostupné z: <https://docs.python.org/3/>
- 23) NumPy [online]. Copyright © 2022 NumPy. All rights reserved. [cit. 30.03.2022].
Dostupné z: <https://numpy.org/>
- 24) Practical guidelines for Production and Implementation in accordance with EBU R 128 [online]. Copyright ©S [cit. 01.04.2022]. Dostupné z: <https://tech.ebu.ch/docs/tech/tech3343-v2.pdf>
- 25) Plot Spectrum - Audacity Manual. *Audacity Manual* [online]. Dostupné z: https://manual.audacityteam.org/man/plot_spectrum.html