

# VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ  
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

FACULTY OF INFORMATION TECHNOLOGY  
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

## ČASOVĚ-FREKVENČNÍ ANALÝZA SIGNÁLU

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. RADOVAN KOVAČEV

BRNO 2012



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ  
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ  
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

FACULTY OF INFORMATION TECHNOLOGY  
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

# ČASOVĚ-FREKVENČNÍ ANALÝZA SIGNÁLU

TIME-FREQUENCY SIGNAL ANALYSIS

DIPLOMOVÁ PRÁCE  
MASTER'S THESIS

AUTOR PRÁCE  
AUTHOR

Bc. RADOVAN KOVAČEV

VEDOUCÍ PRÁCE  
SUPERVISOR

Ing. DAVID BAŘINA

BRNO 2012

## Zadání diplomové práce

Řešitel: **Kovačev Radovan, Bc.**

Obor: Počítačová grafika a multimédia

Téma: **Časově-frekvenční analýza signálu**  
**Time-Frequency Signal Analysis**

Kategorie: Zpracování řeči a přirozeného jazyka

Pokyny:

1. Seznamte se s metodami časově-frekvenční analýzy signálu.
2. Podrobně popište analýzu pomocí krátkodobé Fourierovy transformace a vlnkových transformací. Popište vybrané aplikace této analýzy.
3. Vybranou aplikaci implementujte, srovnajte na ní zmíněné transformace. Vyberte buď odhad okamžité frekvence, frekvenční modulaci nebo změnu délky trvání zvukové nahrávky.
4. Zhodnoťte dosažené výsledky, diskutujte o vhodnosti použití obou metod.

Literatura:

- MALLAT, Stéphane. A Wavelet Tour of Signal Processing : The Sparse Way. With contributions from Gabriel Peyré. 3rd edition. [s.l.] : Academic Press, c2009. xx, 805 s. ISBN 9780123743701.

Při obhajobě semestrální části diplomového projektu je požadováno:

- Bez požadavků.

Podrobné závazné pokyny pro vypracování diplomové práce naleznete na adrese

<http://www.fit.vutbr.cz/info/szz/>

Technická zpráva diplomové práce musí obsahovat formulaci cíle, charakteristiku současného stavu, teoretická a odborná východiska řešených problémů a specifikaci etap, které byly vyřešeny v rámci ročníkového a semestrálního projektu (30 až 40% celkového rozsahu technické zprávy).

Student odevzdá v jednom výtisku technickou zprávu a v elektronické podobě zdrojový text technické zprávy, úplnou programovou dokumentaci a zdrojové texty programů. Informace v elektronické podobě budou uloženy na standardním nepřepisovatelném paměťovém médiu (CD-R, DVD-R, apod.), které bude vloženo do písemné zprávy tak, aby nemohlo dojít k jeho ztrátě při běžné manipulaci.

Vedoucí: **Bařina David, Ing.**, UPGM FIT VUT

Datum zadání: 19. září 2011

Datum odevzdání: 23. května 2012

**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**  
Fakulta informačních technologií  
Ústav počítačové grafiky a multimédií  
612 00 Brno, Božetěchova 2



doc. Dr. Ing. Jan Černocký  
vedoucí ústavu

## Abstrakt

Ústřední předmět této práce představuje časově-frekvenční analýza signálu. Úmyslem je poskytnout nejdřív základní teorii se zaměřením na spojitou vlnkovou transformaci, kde se kromě jiného provádí taky srovnání jejich klíčových vlastností s příbuznou krátkodobou Fourierovou transformací. Následně na to navazuje demonstrace účelu analýzy pomocí implementace jednoduchého praktického příkladu. Konkrétním cílem je vytvoření fázového vokodéru, který řeší problém změny délky trvání zvukové nahrávky a tonální posun. Na tomhle místě je pak podrobně popsán princip jeho fungování, návrh, postup při psaní programu, výstupy a dosažené výsledky.

## Abstract

The main subject of this work represents the time-frequency signal analysis. Firstly, it intends to provide the most essential theoretical background with focus on the continuous wavelet transform, where also a comparison of the key features with its close relative the short-time Fourier transform is performed. Afterwards, there follows a demonstration of the purpose with a practical example. The particular aim is to create a phase vocoder solution for modifying the length of a sound record duration and pitch shifting. Here, in this place, the functional principles, design, procedure of assembling, outputs and achieved results are well documented.

## Klíčová slova

spojitá vlnková transformace, krátkodobá Fourierova transformace, fázový vokodér

## Keywords

continuous wavelet transform, short-time Fourier transform, phase vocoder

## Citace

Radovan Kovačev: Časově-frekvenční analýza signálu, diplomová práce, Brno, FIT VUT v Brně, 2012



# Časově-frekvenční analýza signálu

## Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením pana Ing. Davida Bařiny.

.....  
Radovan Kovačev  
18. dubna 2012

## Poděkování

Tímhle bych chtěl poděkovat panu Ing. Davidovi Bařinovi za odborné vedení, konzultace a rady, které mi poskytoval během vzniku této práce. Dále bych rád vzdal díky svoji rodině a přátelům za podporu.

© Radovan Kovačev, 2012.

*Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů.*

# Obsah

<b>1</b>	<b>Úvod</b>	<b>3</b>
1.1	Štruktúrne členenie . . . . .	4
<b>2</b>	<b>Matematické pozadie</b>	<b>5</b>
2.1	Vektorový priestor . . . . .	5
2.2	Hilbertovský priestor . . . . .	6
2.3	Integrálna transformácia . . . . .	6
2.4	Fourierova transformácia . . . . .	7
2.5	Konvolúcia a vzájomná korelácia . . . . .	7
<b>3</b>	<b>Spojité vlnková transformácia</b>	<b>8</b>
3.1	Vlnka . . . . .	8
3.2	Požiadavky pre vlnku . . . . .	9
3.3	Energetické spektrum vlnky . . . . .	10
3.4	Transformácia . . . . .	11
3.5	Rozpoznávanie podobných štruktúr . . . . .	14
3.6	Detekcia hrán . . . . .	15
3.7	Inverzná vlnková transformácia . . . . .	16
3.8	Energia signálu na základe vlniek a výkonové spektrum . . . . .	16
3.9	Vlnková transformácia Fourierovou . . . . .	18
3.10	Komplexná vlnka . . . . .	19
3.11	Časovo-frekvenčná rovina a Heisenbergove boxy . . . . .	20
3.12	Praktický výpočet spojitej vlnkovej transformácie . . . . .	22
<b>4</b>	<b>Rýchla Fourierova transformácia</b>	<b>24</b>
4.1	Význam a použitie . . . . .	24
4.2	Princíp . . . . .	24
4.3	Algoritmus . . . . .	26
4.4	Inverzná transformácia . . . . .	27
<b>5</b>	<b>Fázový vokodér</b>	<b>28</b>
5.1	Prehľad . . . . .	28
5.2	Princíp . . . . .	29
5.3	Analýza . . . . .	29
5.4	Spracovanie . . . . .	29
5.5	Syntéza . . . . .	30

<b>6 Implementácia</b>	<b>31</b>
6.1 Návrh architektúry programu . . . . .	31
6.2 Graf hustoty energie . . . . .	32
6.3 Prostriedky . . . . .	33
6.4 Zložitosť a rýchlosť . . . . .	34
<b>7 Výsledky</b>	<b>35</b>
7.1 Spektrogramy . . . . .	35
7.2 Princíp neurčitosti . . . . .	36
7.3 Porovnanie rýchlostí módov . . . . .	37
7.4 Zmena trvania a tónu nahrávky . . . . .	38
<b>8 Záver</b>	<b>40</b>

# Kapitola 1

## Úvod

Signál dnes predstavuje veľmi často skloňovaný pojem s neodmysliteľným významom v oblasti informačných technológií. Ide totiž o zjednocujúcu abstrakciu pre nositeľ informácie a prostriedok komunikácie. V užšom slova zmysle sa signálom často chápe závislosť nejakej (fyzikálnej, biologickej, ekonomickej, ..) veličiny od času, priestoru alebo iného kontinua. Značná pozornosť sa v praxi kladie hlavne na spojité signály realizované energiou elektro-magnetického žiarenia (röntgen, svetlo, rádiové vlny, ..) a mechanického (ultra, infra, počuteľné zvuky a iné) vlnenia, ktoré sa zvyknú prevádzať na elektrické impulzy, reprezentujúce diskkrétne binárne dáta, nad ktorými operujú číslicové stroje - počítače. Tie umelo vyprodukované a vyslané majú zvyčajne vopred zavedenú interpretáciu. Ak má však signál pôvod prirodzený (zachytáva určitý aspekt javu v prírode), alebo jeho štruktúra nie je úplne jasná, obsiahnutá informácia sa stáva objektom skúmania.

Pri pozorovaní špecifických vlastností nestacionárnych signálov v čase je v niektorých prípadoch za potreby vykonať ich časovo-frekvenčnú analýzu. Ako už možno predpokladať, kombinuje časový a frekvenčný pohľad. Ide tak o vyjadrenie vývoja frekvenčného zloženia. Populárnymi nástrojmi tejto metódy sú vlnkové a krátkodobé Fourierove transformácie. Ich prirodzenou schopnosťou je lokalizácia a určenie rozsahu zmien. Vo všeobecnosti sa používajú k spektrálnej modulácii, odhadu okamžitej frekvencie, odstráneniu šumu, detekcii príznakov, kompresii a urýchleniu algoritmov maticových operácií. Do širšieho spektra ich aplikácií patrí napríklad geofyzika (analýza seizmických signálov), povaha tekutín (turbulentné toky, atmosferické deje), pozorovanie vibrácií (odhaľovanie závad strojov), nedeštruktívne testovanie (detekcia prasklín), medicína (rozpoznávanie QRS komplexov v signáloch EKG, evokovaných potenciálov v EEG, sledovanie korelácií v sekvenciách DNA), ekonomika (rozbor burzových indexov), štúdium plazmy, astronómia a mnohé ďalšie [4].

Časovo-frekvenčná analýza nachádza svoje uplatnenie aj v oblasti počítačovej produkcie multimédií, kde bežne vzniká potreba pre synchronizáciu. Moderné prístupy k hudobnej tvorbe zahŕňajú rozmanité manipulácie so zvukovými vzorkami vrátane zmeny dĺžky trvania so zachovaním tónu. Štúdiový inžinieri v produkcii sa denne stretávajú s problémom modifikácie tempa melodickéj frázy alebo rytmickej slučky bicích, ktoré už nie je možné nanovo zaznamenať. Inokedy je zas požadovaná transpozícia nahrávky alebo resyntéza v reálnom čase. Pri živých vystúpeniach sa obvykle stáva, že je nutné podporiť výkon speváka, ktorý už nedosiahne požadované tóny, zvýšením rozsahu (tonálnym posunom) alebo automatickým zarovnávaním do tóniny. V hudobnom priemysle sa ako svetoznámy prostriedok k tomuto účelu stalo zariadenie Auto-Tune, založené na princípe fázového vokodéru [22].

## 1.1 Štruktúrne členenie

Osnova tejto práce je systematicky rozvrhnutá nasledovne.

Prvá kapitola sa zameriava na veľmi stručné zhrnutie zloženiu matematických základov týkajúcich sa signálov. Rieši zhruba problematiku vektorového, Hilbertovho priestoru, všeobecnej integrálnej a Fourierovej transformácie, konvolúcie a vzájomnej korelácie.

V druhej kapitole sa pozornosť kladie na vysvetlenie fungovania metódy spojitej vlnkovej transformácie. Zaoberá sa hlavne vlnkou, jej vlastnosťami a energetickými aspektami analýzy. V rámci tohto výkladu je okrem iného spomenutý aj zastrešujúci koncept časovo-frekvenčnej roviny a krátkodobá Fourierova transformácia.

Tretia kapitola má za účel oboznámiť čitateľa s prakticky nepostrádateľným algoritmom rýchlej Fourierovej transformácie, značne urýchleným spôsobom počítania diskkrétnej Fourierovej transformácie, ktorý sa ďalej v praktickej časti používa.

V štvrtej kapitole je úsilie o priblíženie princípu fázového vokodéru, jednoduchého rozšírenia krátkodobej Fourierovej transformácie a aplikácie časovo-frekvenčnej analýzy, slúžiacej k zmene dĺžky trvania zvukovej nahrávky a hudobnej transpozície.

Piata kapitola pojednáva o podrobnostiach uskutočnenej implementácie spomínanej aplikácie. Obsahuje schému návrhu vykonávania, vstupno-výstupné špecifikácie a taktiež popis triku pre výpočet dvoch reálnych diskrétnych Fourierových transformácií pomocou jednej komplexnej.

V šiestej kapitole sú odprezentované výsledky testovania fázového vokodéru. Mimo iného sa predkladá experimentálny dôkaz lineárnej časovej zložitosti a deštruktívnosti operácie.

Posledná, siedma kapitola, je venovaná zhrnutiu celej práce. Záverečné zhodnotenie je rozdelené na zimný a letný semester s dôrazom na praktickú časť. Na tomto mieste sú tiež uvedené možnosti pre smer ďalšieho potenciálneho pokračovania.

## Kapitola 2

# Matematické pozadie

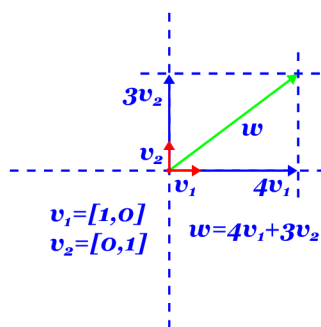
Ako už samotný názov kapitoly napovedá, reč bude o matematických základoch potrebných k pochopeniu ďalšieho výkladu časovo-frekvenčnej analýzy. Poskytnutá teória je však aplikovaná v oblasti spracovania signálov vo veľkej miere všeobecne.

### 2.1 Vektorový priestor

Jedným z predmetov štúdia lineárnej algebry je pojem vektorového priestoru  $V$ . Jedná sa o pomenovanie špeciálneho prípadu algebraickej štruktúry s vlastnosťami telesa, spolu s množinou, ktorú generuje. Nosič pozostáva z vektorov rovnakého rozmeru, ktorý určuje počet dimenzií priestoru, a možno na ne uplatniť dve operácie. Prvou je násobenie vektora skalárom, druhou súčet vektorov. Dohromady tak umožňujú budovať nové prvky metódou známou pod názvom lineárna kombinácia nasledovne:

$$w = \sum_i a_i \cdot v_i \quad (2.1)$$

kde  $w$  predstavuje generovaný vektor a koeficienty  $a_i$  váhy konštrukčných vektorov  $v_i$ . Vektorový priestor je preto známy aj ako lineárny [9]. Na obrázku 2.1 je znázornená lineárna kombinácia v dvojrozmernom vektorom priestore.



Obrázok 2.1: Lineárna kombinácia: 2-dimenzionálny vektorový priestor.

Elementy obsiahnuté v nosnej množine bývajú zvyčajne lineárne nezávislé. Znamená to, že odobratím jedného vektora z tejto množiny ho nemožno lineárnou kombináciou zo

zvyšných zostrojít. Takáto množina sa nazýva bázou. Kanonickú bázu predstavujú stĺpce jednotkovej matice [3].

## 2.2 Hilbertovský priestor

Hilbertovským priestorom sa rozumie úplný unitárny vektorový priestor. Skutočnosť, že je na ňom definovaný skalárny súčin:

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C} \quad (2.2)$$

umožňuje ortogonálnu projekciu do podpriestorov [41].

Z hľadiska spojitých signálov je dôležitý špeciálny prípad, Lebesgueov priestor  $L^2(\mathbb{R})$ , kde skalárny súčin má formu:

$$\langle u, v \rangle = \int u(t)v^*(t) dt \quad (2.3)$$

Jeho elementmi sú všetky spojité signály v čase [7].

Diskrétné signály zas zastupuje priestor  $\ell^2(\mathbb{Z})$ , pre ktorý sa skalárny súčin definuje ako:

$$\langle u, v \rangle = \sum_i u_i v_i^* \quad (2.4)$$

Systém vektorov  $\{b_i\}$ , ktorý spĺňa podmienku:

$$\langle b_i, b_j \rangle = \delta(i - j) \quad (2.5)$$

sa označuje ako ortogonálny. Vektory takéhoto systému sú vzájomne na seba kolmé a ich skalárny súčin je rovný nule. Ak majú navyše každú veľkosť rovnú jednej, hovoríme o ortonormálnom systéme. Pre všetky Hilbertove priestory sa dajú pomocou projekcií zostrojít ortonormálne bázy [38].

## 2.3 Integrovaná transformácia

Ľubovoľnú transformáciu  $T$  funkcie  $f(t)$  v tvare:

$$T(n) = \int f(t)\psi(t, n) dt \quad (2.6)$$

kde  $\psi$  je jej jadrom, voláme integrovanou. V zjednotení týmto typom sú projekcie v Hilbertových priestoroch, Fourierova, vlnková transformácia a ďalšie [24].

Inverzná transformácia je definovaná ako:

$$f(t) = \int T(n)\psi^{-1}(n, t) dn \quad (2.7)$$

Inverzné jadro  $\psi^{-1}$  nemusí nutne existovať [4].

## 2.4 Fourierova transformácia

Populárnou metódou frekvenčnej analýzy stacionárnych signálov je Fourierova transformácia. Poskytuje informáciu o miere výskytu frekvenčných zložiek v globále, pretože nie je schopná ich lokalizovať. Z časovej do frekvenčnej domény a súčasne naspäť tak prevádza iba periodické signály. Definuje sa:

$$\hat{s}(f) = \int_{-\infty}^{\infty} s(t)e^{-i2\pi ft} dt \quad (2.8)$$

Funkcia  $\hat{s}$  sa zvykne nazývať spektrom signálu [20].

Inverzná Fourierova transformácia je daná:

$$s(t) = \int_{-\infty}^{\infty} \hat{s}(f)e^{i2\pi ft} df \quad (2.9)$$

## 2.5 Konvolúcia a vzájomná korelácia

Operátor konvolúcie počíta výstup lineárne a časovo invariantného systému [26]. Operuje nad jeho odozvou a vstupným spojitým signálom takto:

$$(f * g)(t) = \int f(\tau)g(t - \tau) d\tau = \int f(t - \tau)g(\tau) d\tau \quad (2.10)$$

Ďalšou významnou operáciou je vzájomná korelácia. Tá vyjadruje podobnosť dvoch signálov pri prechode jedného druhým. Definuje sa nasledovne:

$$(f \star g)(t) = \int f^*(\tau)g(t + \tau) d\tau = \int f^*(\tau - t)g(\tau) d\tau \quad (2.11)$$

Vzájomnú koreláciu možno vhodne nahradiť konvolúciou a naopak [34].



## Kapitola 3

# Spojité vlnková transformácia

Predmetom pojednávania v tejto kapitole bude spojité vlnková transformácia (angl. continuous wavelet transform, skr. CWT). Zo všetkého najskôr sa čitateľ oboznámi s pojmom vlnky a jej kritickými vlastnosťami. Ďalej sa bude môcť zapodievať skúmaním toho, akú úlohu v transformácii zohráva a čo všetko dokáže. Bude vysvetlené ako funguje rozpoznávanie štruktúry signálu. Analyzujú sa energetické aspekty transformácie. Zvážia sa výhody komplexnej vlnky. Potom sa štúdium orientuje na objasnenie Heisenbergovho princípu neurčitosti z pohľadu signálov. Porovnaná bude časovo-frekvenčná rovina vlnkovej a krátkodobej Fourierovej transformácie a napokon sa kapitola uzatvorí zamyslením nad riešeniami niektorých prekážok pri praktickom výpočte.

### 3.1 Vlnka

Vlnka predstavuje fundamentálny koncept, na ktorom je vybudovaná teória vlnkovej transformácie. Už samotný názov transformácie napovedá o jej dôležitosti. Jedná sa v podstate o lokálnu funkciu  $\psi$  kmitavého charakteru, ktorá spĺňa špeciálne matematické kritéria [5].

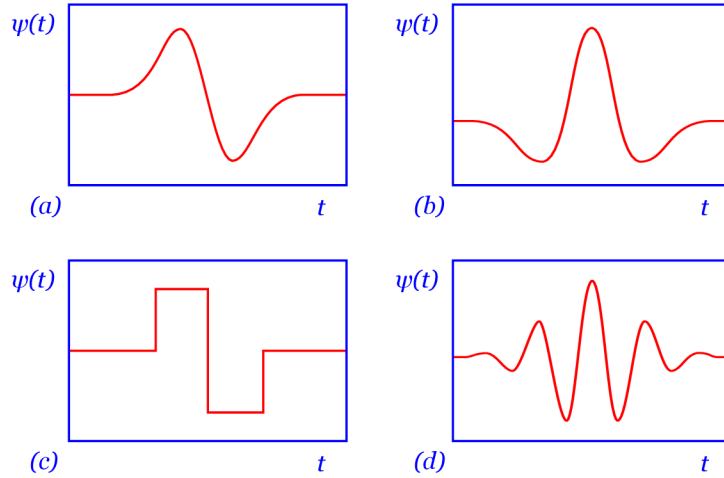
V tejto práci sa výklad teórie obmedzí na analýzu jednorozmerných signálov v čase. Bude sa teda uvažovať o vlnkách tohto typu. V priebehu transformácie možno s vlnkou narábať dvoma spôsobmi. Posúvať ju pozdĺž horizontálnej časovej osi a meniť jej mierku. Zmenou mierky sa chápe rovnomerné rozťahnutie a stlačenie vlnky. Tieto dve manipulácie majú za následok, že transformovaný signál je reprezentovaný v doméne mierky a posunutia.

Na obrázku 3.1 je znázornených pár v praxi bežne používaných vlniek. Na niektoré z nich sa bude dať pozrieť detailnejšie neskôr. Mimo tejto vzorky existuje ešte veľké množstvo vlniek špecializovaných pre príslušné aplikácie. Voľba tej najvhodnejšej je závislá na povahe signálu a požiadavkách na analýzu javu, ktorá je skúmaná.

V ďalšom texte je pozornosť viac sústredená na konkrétnu vlnku známu ako Mexický klobúk, ktorá je veľmi vhodná pri ilustrácii vlnkovej transformácie. Ide v podstate o druhú deriváciu funkcie normálneho rozdelenia pravdepodobnosti. Za vlnky však možno nasadiť všetky derivácie Gaussovej distribúcie. Prvá a druhá derivácia Gausiánu sú zachytené na obrázku 3.1. Derivácie vyšších rádov sú už menej bežné. Mexický klobúk sa definuje ako:

$$\psi(t) = (1 - t^2)e^{-t^2/2} \quad (3.1)$$

a vo vyššom detaile je vykreslený na obrázku 3.2. Vlnka popísaná rovnicou (3.1) je známa ako *materská*. Je to základná forma od ktorej sa manipulovaním odvodzujú ďalšie verzie.



Obrázok 3.1: Vlnky: a) prvá derivácia Gaussovej funkcie b) Mexický klobúk c) Haarova a d) Morletova/Gáborova (reálna časť)

## 3.2 Požiadavky pre vlnku

K tomu, aby sa mohla považovať funkcia za vlnku a bola plne použiteľná pri transformácii, je nevyhnutné aby spĺňala nasledujúce matematické kritériá [21]:

1. Energia vlnky musí byť konečná:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \quad (3.2)$$

2. Nech  $\hat{\psi}(f)$  je Fourierovou transformáciou  $\psi(t)$ , je ďalej potrebné aby platilo nasledovné:

$$C_g = \int_0^{\infty} \frac{|\hat{\psi}(f)|^2}{f} df < \infty \quad (3.3)$$

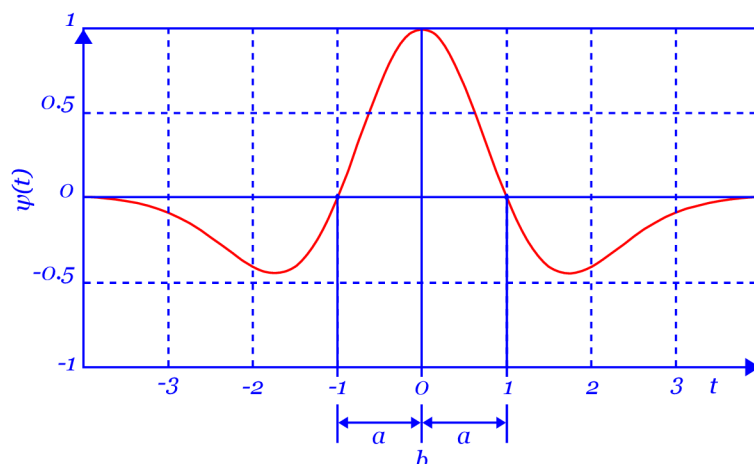
Uvedená relácia je známa ako podmienka prípustnosti, kde  $C_g$  sa nazýva jej konštantou. Hodnota  $C_g$  je závislá na zvolenej vlnke a pre Mexický klobúk podľa definície (3.1) sa rovná  $\pi$ . Z uvedeného vzťahu ďalej vyplýva, že  $\hat{\psi}$  neobsahuje jednosmernú zložku, čiže:

$$\hat{\psi}(0) = 0 \quad (3.4)$$

čo ďalej implikuje nulový priemer funkcie:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (3.5)$$

3. Pre komplexnú vlnku sa dodatočne vyžaduje, aby jej Fourierova transformácia bola súčasne reálna a nulová pre negatívne frekvencie.



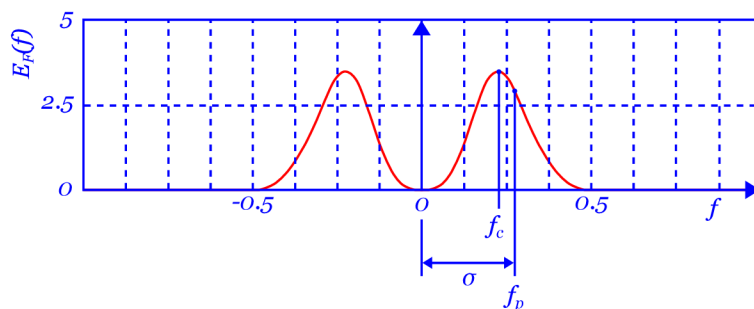
Obrázok 3.2: Materská vlnka: Mexický klobúk.

Okrem nutných podmienok pre konštrukciu vlnky existujú ešte užitočné konvencie. Patrí medzi ne napríklad jednotková normalizácia plochy vlnky:

$$\|\psi\| = 1 \quad (3.6)$$

### 3.3 Energetické spektrum vlnky

Významným zdrojom popisu vlastností vlnky je jej energetické spektrum. Poskytuje charakteristiku pre *pásmový* filter, ktorým sa vlnka splnením podmienky prípustnosti stáva. Znamená to, že pri konvolúcii so signálom prepúšťa iba frekvencie v rámci konečného intervalu, *pásmu*.



Obrázok 3.3: Fourierovo energetické spektrum pre Mexický klobúk.

Umocnením absolútnej hodnoty Fourierovej transformácie vlnky sa vypočíta jej Fourierovo energetické spektrum. Pre Mexický klobúk je dané ako:

$$E_F(f) = |\hat{\psi}(f)|^2 = 32\pi^5 f^4 e^{-4\pi^2 f^2} \quad (3.7)$$

a obrázok 3.3 znázorňuje graf tejto závislosti. Ešte pre poznamenanie, existujú aj energetické spektra získané inými metódami ako Fourierovou.

Z energetického spektra sa dokázu vyčítať charakteristické frekvencie vlnky, ktoré sú za potreby k tomu, aby bol pozorovateľ schopný určiť, aký kmitočet sa pri analýze vyšetruje. Dominantnú frekvenciu  $f_p$  určuje špičková hodnota spektra. Druhou reprezentatčnou frekvenciou je  $f_c$ , *stred pásma*. Definovaná je takto:

$$f_c = \sqrt{\frac{\int_0^\infty f^2 |\hat{\psi}(f)|^2 df}{\int_0^\infty |\hat{\psi}(f)|^2 df}} \quad (3.8)$$

a v podstate sa jedná o smerodajnú odchýlku spektra od jeho stredu. V prípade vlnky Mexický klobúk sa  $f_c$  rovná  $\sqrt{5/2}/2\pi$  a  $f_p = \pm \sqrt{2}/2\pi$ . Zmenou mierky v procese vlnkovej transformácie sa energetické spektrum vlnky mení a dôsledkom toho aj jej charakteristická frekvencia.

Z prvej podmienky pre vlnku (3.2) je zrejmé, že celková energia Mexického klobúka je konečná a daná výrazom:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 dt = \int_{-\infty}^{\infty} [(1-t^2)e^{-t^2/2}]^2 dt = \frac{3}{4}\sqrt{\pi} \quad (3.9)$$

Hodnotu tejto energie však možno vypočítať aj ako integrál energetického spektra. Pre konkrétne zvolenú vlnku je to:

$$E = \int_{-\infty}^{\infty} |\hat{\psi}(f)|^2 df = \int_{-\infty}^{\infty} 32\pi^5 f^4 e^{-4\pi^2 f^2} df = \frac{3}{4}\sqrt{\pi} \quad (3.10)$$

a platí tak:

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt = \int_{-\infty}^{\infty} |\hat{\psi}(f)|^2 df \quad (3.11)$$

Podľa tvrdenia Parsevalovho teorému možno presne tento výsledok predpokladať pre akúkoľvek funkciu [1].

Vlnky sa v praxi zvyknú normalizovať na jednotkovú energiu. V prípade normalizácie Mexického klobúka je zo vzťahu (3.9) zrejmé, že sa jeho pôvodná definícia upraví vydelením  $(3\sqrt{\pi}/4)^{1/2}$ . Tak možno dostať:

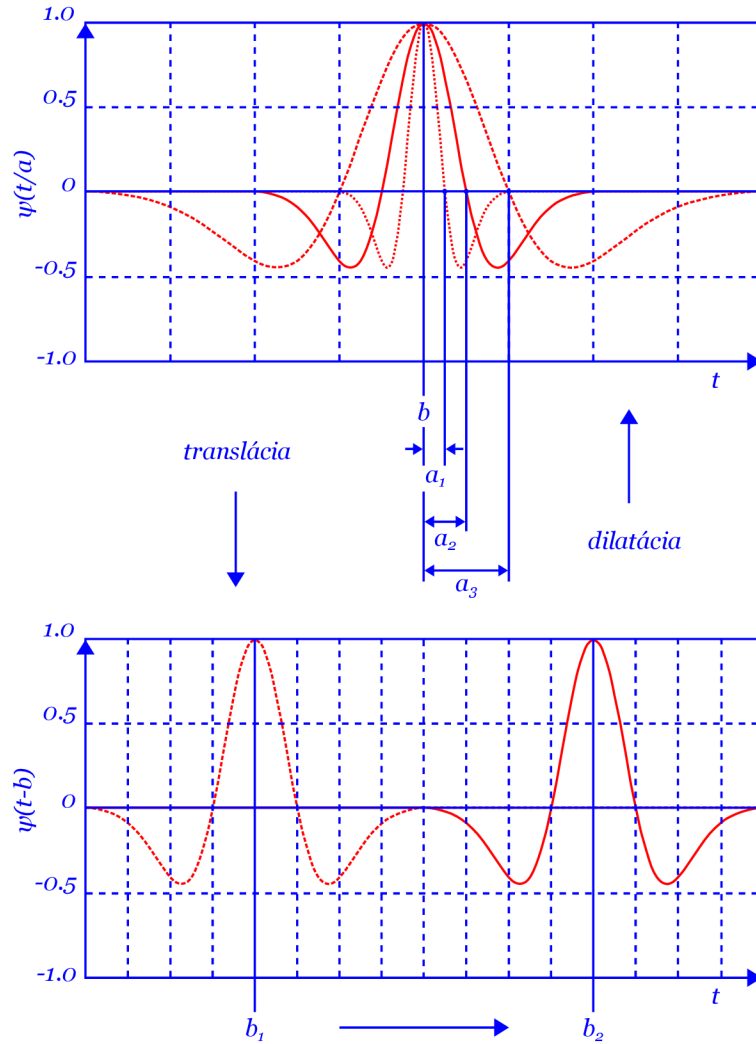
$$\psi(t) = \frac{2}{\sqrt{3}\sqrt[4]{\pi}}(1-t^2)e^{-t^2/2} \quad (3.12)$$

Pri použití normalizovanej verzie je ešte nutné vykonať úpravu hodnoty konštanty prípustnosti  $C_g$  z  $\pi$  na  $4\sqrt{\pi}/3$ . Vo zvyšku práce sa bude ďalej objasňovanie opierať o prvú, nenormalizovanú verziu Mexického klobúka podľa definície (3.1).

### 3.4 Transformácia

Zvolením požadovanej vlnky sa analýza dostáva do fázy samotnej transformácie. V jej priebehu sa však bude od funkcie  $\psi$  vyžadovať určitá miera flexibility. Tú možno dosiahnuť umožnením dvoch manipulácií. Ako už bolo povedané, ide o dilatáciu a transláciu. Pre pripomenutie, dilatácia spôsobuje proporcionálne roztiahnutie alebo stlačenie vlnky a translácia ovláda jej lokalizáciu v čase. Rozsahy týchto manipulácií možno vyjadriť číselne ako hodnoty dilatačného  $a$  a translačného parametra  $b$ . Prostredníctvom ich zmeny sa zmienené manipulácie vykonávajú [15].

Pre ilustráciu, veľkosť dilatačného parametra pre Mexický klobúk má jednoduchú interpretáciu, predstavuje vzdialenosť stredu vlnky od priesečníku vodorovnej osi. Vrchná časť obrázku 3.4 znázorňuje vplyv dilatácie na Mexický klobúk. Translačný parameter  $b$  udáva posunutie od centra materskej vlnky. Translácia Mexického klobúku je zachytená na spodnej časti tohto obrázku.



Obrázok 3.4: Dilatácia a translácia Mexického klobúku. Dilatácia (hore): Plná kontúra reprezentuje materskú vlnku, čiarkovaná jej roztiahnutú verziu na dvojnásobok a bodkovaná stlačenie na polovičnú šírku. Translácia (dole): Vlnka prechádza z bodu  $b_1$  do  $b_2$ .

Definícia vlnky Mexický klobúk podľa rovnice (3.1) sa dá teda rozšíriť o dilatačný a translačný parameter. Získa sa tak flexibilná vlnka  $\psi[(t-b)/a]$ , kde:

$$\psi\left(\frac{t-b}{a}\right) = \left[1 - \left(\frac{t-b}{a}\right)^2\right] e^{-\frac{1}{2}[(t-b)/a]^2} \quad (3.13)$$

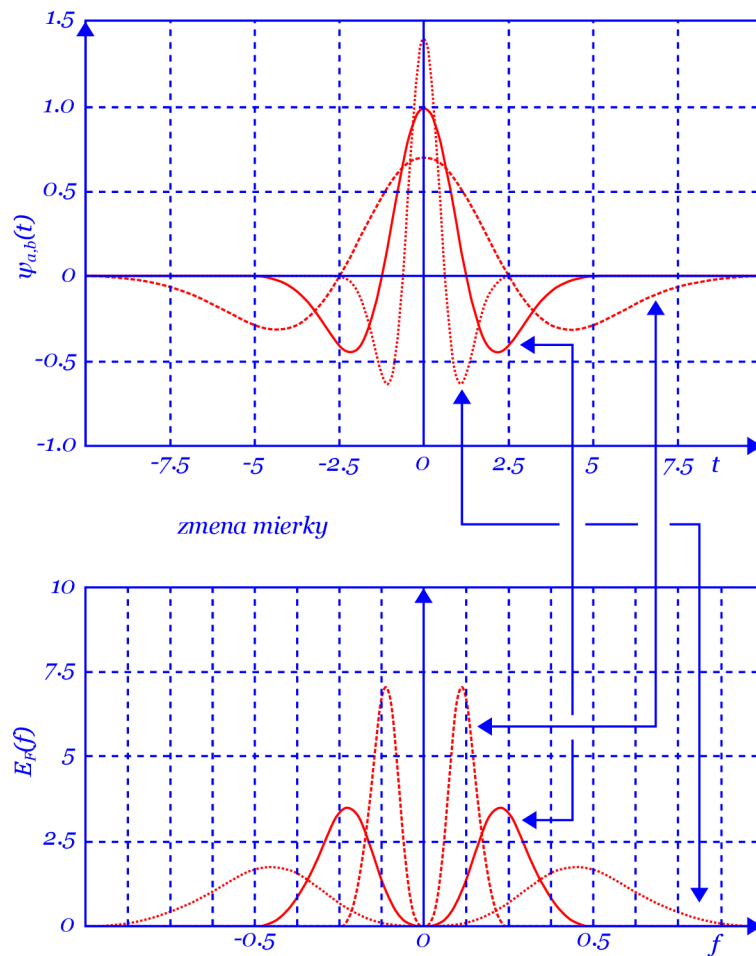
Materská vlnka Mexického klobúku má tieto parametre nastavené na  $a = 1$  a  $b = 0$ . Obor

$a$  a  $b$  spolu tvoria dvojrozmerný neortogonálny priestor, ktorý predstavuje doménu funkcie vlnkovej transformácie.

Pomocou parametrizovanej verzie vlnky sa počíta vlnková transformácia ako:

$$T(a, b) = w(a) \int_{-\infty}^{\infty} x(t) \psi^* \left( \frac{t-b}{a} \right) dt \quad (3.14)$$

kde  $w(a)$  je váhová funkcia. Hviezdička indikuje komplexne združenú funkciu vlnky. Pre reálne funkcie ako Mexický klobúk je však táto informácia irelevantá.  $w(a)$  sa zvyčajne volí ako  $1/\sqrt{a}$  z dôvodu zachovania rovnako veľkej energie pre rozťahnuté a stlačené vlnky. Váhovú funkciu však možno nastaviť aj inak v závislosti od aplikácie. Tento výklad sa však bude držať konvencie. Dilatáciu vlnky v prípade takejto normalizácie môžno nazvať aj zmenou mierky, pretože má analogický efekt k približovaniu v geografických mapách [14].



Obrázok 3.5: Zmena mierky Mexického klobúku (hore) a jeho energetického spektra (dole).

Zmena mierky Mexického klobúku a jej efekt na energetické spektrum sú ilustrované na obrázku 3.5. Rozťahnutím vlnky sa spektrum stláča a naopak. Je to jasný dôsledok toho, že predlžovaním časovej periódy frekvencia klesá. Dilatačný parameter  $a$  je nepriamo úmerný frekvencii vlnky.

Transformáciu energeticky bežne normalizovanou vlnkou možno zapísať aj takto:

$$T(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi^* \left( \frac{t-b}{a} \right) dt \quad (3.15)$$

Rovnica sa priamo interpretuje ako súčin signálu a parametrizovanej vlnky integrovaný cez celkové trvanie. Z viac abstraktného pohľadu možno vlnkovú transformáciu tiež chápať ako cross-koreláciu signálu s vlnkami rôznych hodnôt parametru dilatácie a translácie.

Energeticky normalizovaná vlnka sa zvykne zapisovať v kompaktnom tvare nasledovne:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left( \frac{t-b}{a} \right) \quad (3.16)$$

To umožňuje zapísať transformačný integrál ako:

$$T(a, b) = \int_{-\infty}^{\infty} x(t) \psi_{a,b}^*(t) dt \quad (3.17)$$

Definíciu vlnkovej transformácie možno skrátiť ešte viac na:

$$T(a, b) = \langle x, \psi_{a,b} \rangle \quad (3.18)$$

Transformáciu vlnkami nazývajú matematickým mikroskopom, prostriedkom pre dekompozíciu signálu alebo metódou prevodu do popisnejšej formy [17]. Čo všetko sa skrýva za jednoduchým vzorcom je možné sa dočítať v nasledujúcej časti.

### 3.5 Rozpoznávanie podobných štruktúr

Matematický vzorec vlnkovej transformácie je už známy. Jej fungovanie sa teraz pre lepšiu predstavu môže vizualizovať.

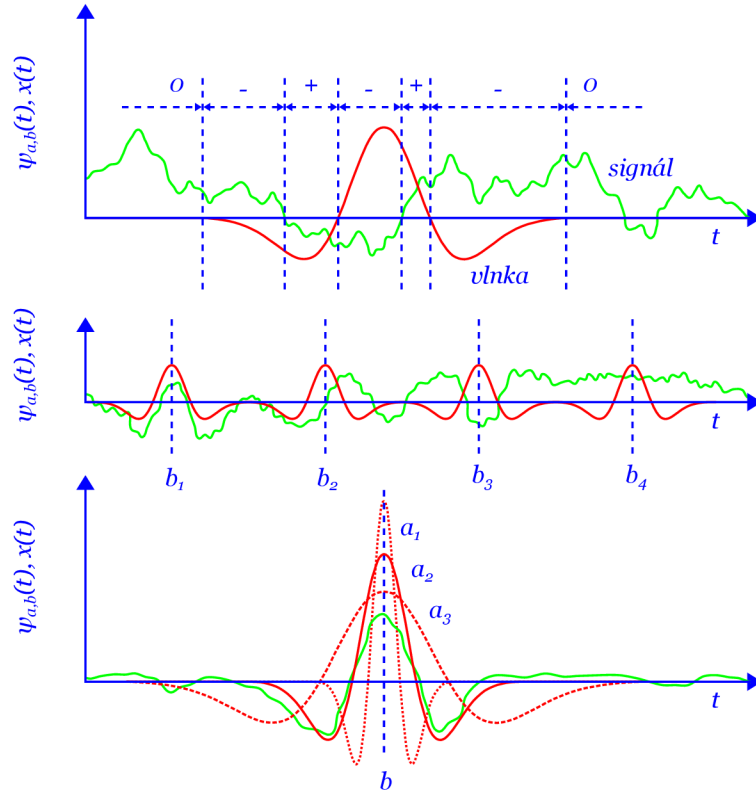
Nech sa pozornosť chvíľu sústreďí na hornú časť obrázka 3.6. Je na nej ilustrovaný výpočet jedného bodu roviny  $a, b$ . Ako sa dá postrehnúť, oblasti pozitívneho, žiadneho a negatívneho príspevku do finálnej hodnoty určujú prechody signálu a vlnky nulou. Oblasť, v ktorej má vlnka a signál zhodnú polaritu, prispieva kladne. V prípade opačnej polarity záporne. Nulový priebeh vlnky alebo signálu má za následok neutrálny príspevok. Výstižne povedané, jedná sa o integrálnu transformáciu skalárneho súčinu.

Vlnku v štyroch rôznych polohách  $b_1, b_2, b_3$  a  $b_4$  vidieť v strede obrázka. Pri prvej translácii  $b_1$  má vlnka s príslušným úsekom signálu relatívne veľkú podobnosť, čo sa týka tvaru aj prechodov cez nulu. To sa prejaví vysokou hodnotou vlnkovej transformácie v danom bode. V polohe  $b_2$  nie sú signál a vlnka vo fáze. Pozitívne a negatívne príspevky sa rušia. To má za následok nulový súčin. Pozícia  $b_3$  ilustruje opačnú fázu, pre ktorú možno násobením signálu a vlnky dostať vysoký záporný výsledok. V poslednej lokácii  $b_4$  je časť signálu približne konštantná s malými výkyvmi, čo sa odzrkadlí ako hodnota blízka nule. Pomocou translácie sa týmto spôsobom lokalizuje daná frekvencia v čase. Prechod vlnky jednej mierky po intervale parametra  $b$  je tak vzájomnou koreláciou.

V spodnej tretine ilustrácie je napokon vyobrazený vplyv mierky na skalárny súčin so signálom. Jedine vlnka v mierke  $a_2$  je signálu podobná a dáva pozitívny výsledok. V dilatáciách  $a_1$  a  $a_3$  vlnky reprezentujú frekvencie, ktoré sa v danom intervale signálu nevyskytujú vo významnej miere a preto je transformácia v týchto bodoch blízka nule.

Spojité vlnková transformácia sa počíta pre vhodne zvolené intervaly parametrov  $a$  a  $b$ . Výpočet spojitaj roviny teoreticky prebieha v nekonečne malých krokoch, čo je ale prakticky

nemožné realizovať. Pre získanie výsledku počítačom treba zvoliť dostatočnú presnosť pre iteráciu tak, aby sa vzniknuté chyby mohli pokladať za zanedbateľné. Takáto pseudospojitá transformácia bude reprezentovaná maticou [30].



Obrázok 3.6: Vizualizácia výpočtu vlnkovej transformácie.

### 3.6 Detekcia hrán

Identifikácia prudkých zmien a nespojitostí signálu je ďalšou užitočnou vlastnosťou vlnkovej transformácie.

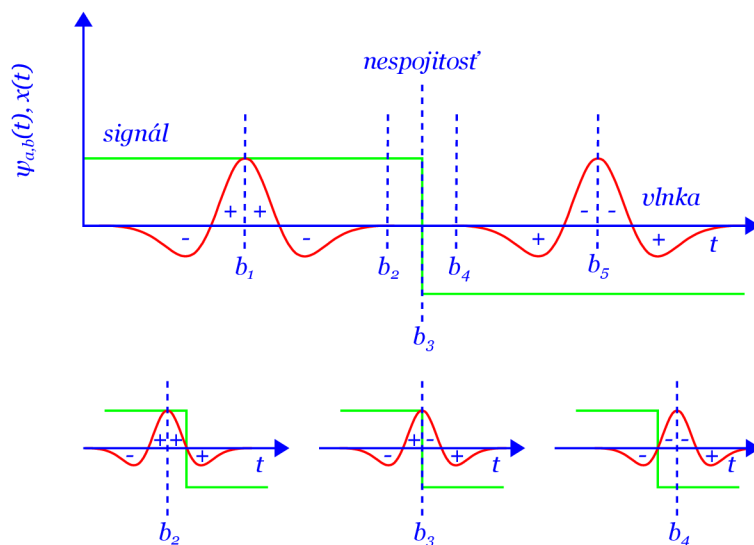
Triviálny príklad pre demonštráciu takejto situácie je vykreslený na obrázku 3.7, v ktorom sa k tomuto účelu využíva konštantný kladný signál so skokom do opačnej polaroty a Mexický klobúk. Bude na ňom vysvetlený efekt prechodu vlnky hranou a jej indikácia. Veľkosť vlnkovej transformácie sa uvažuje v kľúčových bodoch  $b_1$ ,  $b_2$ ,  $b_3$ ,  $b_4$  a  $b_5$ .

V lokáciách ako  $b_1$ , pred vstupom do nespojitosti, je hodnota transformácie nulová. Začiatok penetrácie v polohe  $b_2$  poskytuje pozitívny výsledok, pretože prevažujú kladné príspevky symetrickej vlnky. Mexický klobúk centrovany na hranu v posunutí  $b_3$  dáva v súčine so signálom nulu. Kladné a záporné zložky sa vyrovnávajú. Opúšťaním prieniku v translácii  $b_4$  je hodnota záporná. V nule sa ustáli po odchode v bode  $b_5$ .

Hrana nutne nemusí meniť polaritu signálu, takže vo vlnkovej transformácii nie je vždy rozpoznaná prechodom cez nulu. Dôležitý je tvar grafu vlnkovej transformácie v okolí hrany, na ktorej rozbor možno vo výstupnom grafe uplatniť numerické metódy pre odhad gradientu



a iné. Na rozdiel od vlniek ako Mexický klobúk, stredovo súmerne, teda osovo antisymetrické vlnky detekujú hranu lokálnym extrémom [33].



Obrázok 3.7: Detekcia hrany vlnkou.

### 3.7 Inverzná vlnková transformácia

Definícia inverznej vlnkovej transformácie je nasledovná:

$$x(t) = \frac{1}{C_g} \int_{-\infty}^{\infty} \int_0^{\infty} T(a, b) \psi_{a,b}(t) \frac{dad b}{a^2} \quad (3.19)$$

Umožňuje rekonštrukciu (syntézu) pôvodného signálu z vlnkovej transformácie a vlnky použitej pri analýze.

Ak sa výpočet obmedzí iba na určité rozpatie mierok  $a$ , tak že ostatné sa nebudú uvažovať, vykoná sa tým základná filtrácia pôvodného signálu. Prahovaním mierky je tak možné zbaviť sa neželaného šumu. Dokonalejšou metódou ako dosiahnuť tento cieľ je *modulus maxima* [25].

### 3.8 Energia signálu na základe vlniek a výkonové spektrum

Pre celkovú energiu  $E$  obsiahnutú v signále  $x(t)$  platí:

$$E = \int_{-\infty}^{\infty} |x(t)|^2 dt = \|x(t)\|^2 \quad (3.20)$$

Táto rovnica dáva zmysel hlavne pre signály s obmedzenou veľkosťou energie.

Relatívny energetický príspevok vlnkovej transformácie sa v danom bode získava ako vlnková hustota energie:

$$E(a, b) = |T(a, b)|^2 \quad (3.21)$$

Vykreslením  $E(a, b)$  možno dostať *škálogram*. Ide o analógiu pre *spektrogram* krátkodobej Fourierovej transformácie. Škálogramom sa nazýva v podstate každá funkcia, ktorá vznikne vynásobením  $|T(a, b)|^2$  nejakou konštantou [19].

Celkovú energiu možno pomocou škálogramu a konštanty prípustnosti vyjadriť aj nasledovným spôsobom:

$$E = \frac{1}{C_g} \int_{-\infty}^{\infty} \int_0^{\infty} |T(a, b)|^2 \frac{da}{a^2} db \quad (3.22)$$

Relatívny príspevok vlniek v mierke  $a$  k celkovej energii je zas daný distribúciou energie podľa mierky:

$$E(a) = \frac{1}{C_g} \int_{-\infty}^{\infty} |T(a, b)|^2 db \quad (3.23)$$

Ak chce pozorovateľ porovnať Fourierovo energetické spektrum  $E_F(f)$  s tým, ktoré je získané vlnkovou transformáciou  $E_W(f)$ , bude potrebné naň previesť distribúciu energie podľa mierky  $E(a)$ . K tomu je nutné poznať charakteristickú frekvenciu vlnky. Ako už bolo skôr povedané, najčastejšie ju reprezentuje stred pásma výkonového spektra. Na základne nej možno mierku prepočítať na frekvenciu takto:

$$f = \frac{f_c}{a} \quad (3.24)$$

kde  $f_c$  je konštantou a mierka s frekvenciou sú, ako vidieť, v nepriamej úmere.

Totálna energia signálu sa ďalej dá alternatívne v závislosti od  $E(a)$  zapísať aj nasledovne:

$$E = \int_0^{\infty} E(a) \frac{da}{a^2} \quad (3.25)$$

Podľa predošlých dvoch vzťahov sa potom získava nové vyjadrenie:

$$E = \int_0^{\infty} E_W(f) df \quad (3.26)$$

kde platí, že  $E_W(f) = E(a)/f_c$  a  $da/a^2 = -df/f_c$ .

Keď sa nahradí v transformácii mierka frekvenciou, možno dostať  $T(f, b)$  a na tejto báze zodpovedajúcu energiu:

$$E = \frac{1}{C_g f_c} \int_{-\infty}^{\infty} \int_0^{\infty} |T(f, b)|^2 df db \quad (3.27)$$

Pri predpoklade  $E(f, b) = (|T(f, b)|^2)/(C_g f_c)$  bude zas obdržané:

$$E = \int_{-\infty}^{\infty} \int_0^{\infty} E(f, b) df db \quad (3.28)$$

Lepším povšimnutím si vzorcov sa dá dospieť k tomu, že škálogram v závislosti od frekvencie pokrýva svojou plochou celkovú energiu na rozdiel od toho, ktorý je vykreslený podľa mierky. Škálogramy sú zvyčajne zobrazované v logaritmickú stupnici a graf  $E(f, b)$  je v jednoduchosti iba posunutý a obrátený k  $E(a, b)$ .

Ďalším dôležitým prostriedkom popisu signálu je výkonové spektrum. Jedná sa vlastne o energetické spektrum v časovej perióde. To poskytuje priemernú hodnotu energie za jednotku času. Napríklad, pre signál dĺžky  $\tau$  sú Fourierove a vlnkové spektrum:

$$P_F(f) = \frac{1}{\tau} E_F(f) \quad (3.29)$$

$$P_W(f) = \frac{1}{\tau} E_W(f) = \frac{1}{\tau f_c C_g} \int_0^\tau |T(f, b)|^2 db \quad (3.30)$$

Kôli frekvenčnej distribúcii cez všetky vlnky je výsledné výkonové spektrum trochu odlišné od Fourierovej verzii. Tvar vlnky a jej charakteristická frekvencia sa teda do spektra premieta [11].

K výpočtu dominantných mierok vlnky, ktoré najviac ovplyvnia spektrum slúži rozptyl vlnky:

$$\sigma^2(a) = \frac{1}{\tau} \int_0^\tau |T(a, b)|^2 db \quad (3.31)$$

### 3.9 Vlnková transformácia Fourierovou

Ako už možno vedieť, vlnková transformácia sa počíta ako konvolúcia alebo vzájomná korelácia signálu s vlnkou. Platí to však aj pre ich Fourierové transformácie [18]. Vlnková transformácia sa pomocou Fourierovej, signálu  $\hat{x}(f)$  a vlnky  $\hat{\psi}_{a,b}(f)$ , získa ako:

$$T(a, b) = \int_{-\infty}^{\infty} \hat{x}(f) \hat{\psi}_{a,b}^*(f) df \quad (3.32)$$

Fourierova transformácia parametrizovanej vlnky je:

$$\hat{\psi}_{a,b}(f) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) e^{-i(2\pi f)t} dt \quad (3.33)$$

Substitúciou  $\acute{t} = (t-b)/a$ , teda  $dt = a d\acute{t}$  sa dostane:

$$\hat{\psi}_{a,b}(f) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi(\acute{t}) e^{-i(2\pi f)(a\acute{t}+b)} a d\acute{t} \quad (3.34)$$

Ďalšou úpravou vyjde:

$$\hat{\psi}_{a,b}(f) = \sqrt{a} e^{-i(2\pi f)b} \int_{-\infty}^{\infty} \psi(t) e^{-i(2\pi a f)t} dt \quad (3.35)$$

z čoho sa dá vyjadriť táto definícia:

$$\hat{\psi}_{a,b}(f) = \sqrt{a} \hat{\psi}(af) e^{-i(2\pi f)b} \quad (3.36)$$

Komplexne združená Fourierova transformácia vlnky je potom:

$$\hat{\psi}_{a,b}^*(f) = \sqrt{a} \hat{\psi}^*(af) e^{i(2\pi f)b} \quad (3.37)$$

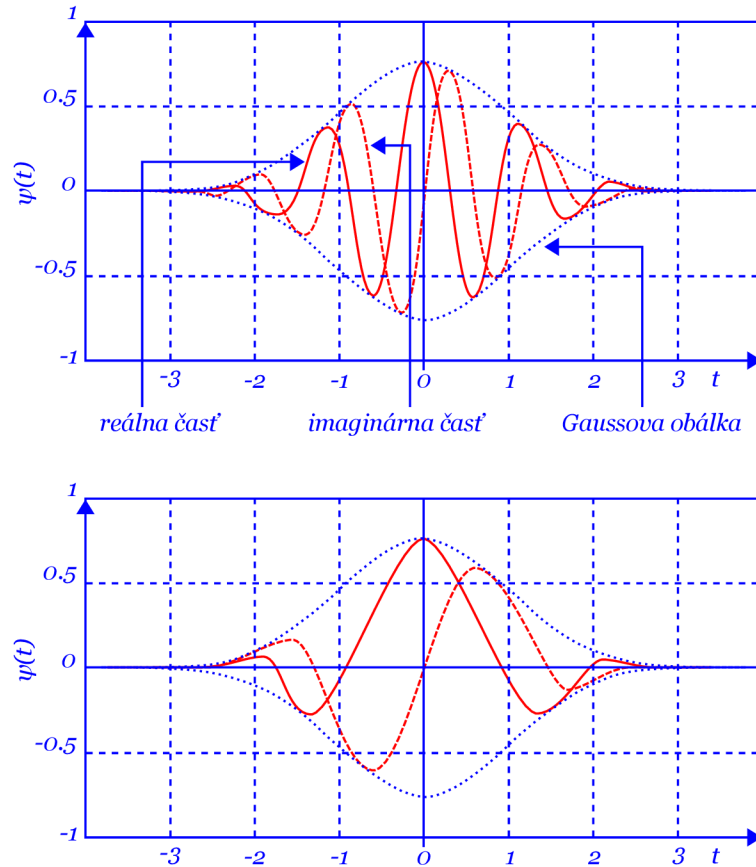
Rovnica pre vlnkovú transformáciu tak môže byť rozšírená na:

$$T(a, b) = \sqrt{a} \int_{-\infty}^{\infty} \hat{x}(f) \hat{\psi}^*(af) e^{i(2\pi f)b} df \quad (3.38)$$

Tak bol získaný vzťah vo forme inverznej Fourierovej transformácie. Tento výsledok je užitočný z dôvodu možnosti zapojenia algoritmu rýchlej Fourierovej transformácie do výpočtu.

### 3.10 Komplexná vlnka

Výhodou komplexnej vlnky oproti reálnej je jej schopnosť oddeliť fázové komponenty od amplitúdových [2]. Aj vďaka tomu dostala prívlastok analytická. Jej Fourierova transformácia je pre záporné frekvencie nulová. Z Mexického klobúku sa teda jednoducho vytvorí komplexná vlnka, tak že inverzná Fourierova transformácia sa uskutoční na upravenom spektre. Pre demonštráciu komplexných vlastností bude však uprednostnená *Morletova vlnka*.



Obrázok 3.8: Morletova vlnka.

Definícia Morletovej vlnky znie:

$$\psi(t) = \pi^{-1/4} (e^{i2\pi f_0 t} - e^{-(2\pi f_0)^2/2}) e^{-t^2/2} \quad (3.39)$$

kde  $f_0$  je centrálnou frekvenciou materskej vlnky. V jednoduchom poňatí určuje frekvencia  $f_0$  počet efektívnych oscilácií v rámci vlnky. Druhý člen v zátvorke je takzvaný korekčný. Upravuje to, aby mala vlnka nulový priemer. Pre dostatočne vysoké hodnoty  $f_0 \gg 0$  môže byť ignorovaný a Morletova vlnka zapísaná v jednoduchšom tvare:

$$\psi(t) = \pi^{-1/4} e^{i2\pi f_0 t} e^{-t^2/2} \quad (3.40)$$

Z tejto definície je pre odborníka na prvý pohľad jasné, že ide o normalizovanú komplexnú

exponenciálu v Gaussovej obálke. Zjednodušený tvar vlnky je už teoreticky neprípustný, no prakticky použiteľný s minimálnou chybou.

Reálna a imaginárna zložka Morletovej vlnky je zachytená na obrázku 3.8, v hornej časti s vyššou a v dolnej s nižšou centrálnou frekvenciou. Je evidentné, že reálna a imaginárna časť sa líšia vo fáze o štvrtinu periódy.

Podobne ako pre Mexický klobúk, parametrizovaná definícia vlnky bude skonštruovaná substitúciou  $(t - b)/a$  za  $t$  takto:

$$\psi\left(\frac{t-b}{a}\right) = \frac{1}{\pi^{1/4}} e^{i2\pi f_0[(t-b)/a]} e^{-\frac{1}{2}[(t-b)/a]^2} \quad (3.41)$$

Kedže bod vlnkovej transformácie Morletovou vlnkou  $T(a, b)$  je komplexné číslo, bude sa dať zapísať pomocou fázy  $\phi(a, b)$  a modulu  $|T(a, b)|$ . Fáza nadobúda hodnoty z intervalu  $(-\pi, \pi)$ . Maximálna transformácia indikuje nulová fáza a minimá  $\pi$  alebo  $-\pi$  [6].

Fourierova transformácia Morletovej vlnky je ďalej daná:

$$\hat{\psi}(f) = \pi^{1/4} \sqrt{2} e^{-\frac{1}{2}(2\pi f - 2\pi f_0)^2} \quad (3.42)$$

Podobá sa na Gaussovu distribúciu centrovanú na  $f_0$ , ktorá sa bežne považuje aj za charakteristickú frekvenciu Morletovej vlnky.

Energetické spektrum je napokon získané ako:

$$|\hat{\psi}(f)|^2 = 2\pi^{1/2} e^{-(2\pi f - 2\pi f_0)^2} \quad (3.43)$$

### 3.11 Časovo-frekvenčná rovina a Heisenbergove boxy

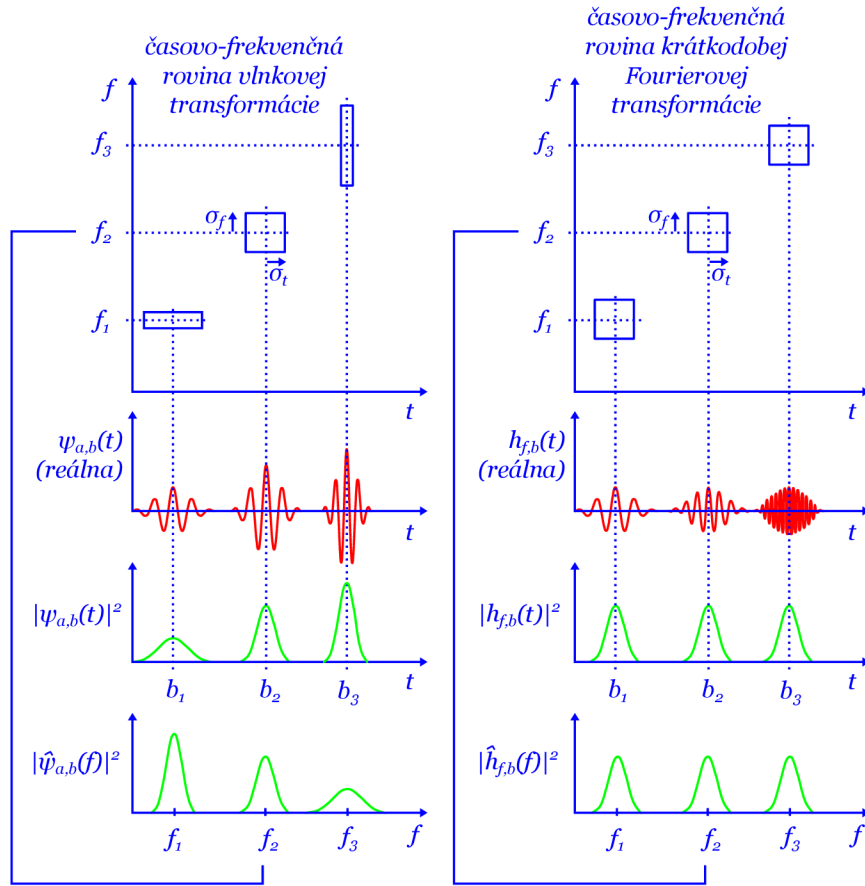
Analyzovať signál v čase a frekvencii zároveň možno aj inak ako vlnkovou transformáciou. Pred vynájdením tejto sofistikovanej metódy sa k tomuto účelu používalo v pricípe veľmi jednoduché rozšírenie Fourierovej transformácie, známe pod pomenovaním krátkodobá Fourierova transformácia [35] (angl. short-time Fourier Transform, skr. STFT).

V nasledujúcich pár riadkoch textu bude pozornosť venovaná porovnaniu časovo-frekvenčných charakteristík týchto dvoch metód. Uvažovať sa budú špecifické prípady. Vlnková transformácia bude demonštrovaná s Morletovou vlnkou. Pre krátkodobú Fourierovú sa zvolí Gáborová transformácia. Funkcie analyzačných jadier oboch využívajú Gaussovo okno a sú si na prvý pohľad podobné, ako sa ďalej uvidí. Pri transformácii Morletovou vlnkou sa však mení mierka Gaussovej obálky a exponenciály súčane, kde pri Gáborovej transformácii sa ponecháva dĺžka okna konštantná, škálovaná je iba sinusoida a to je podstatný rozdiel. Časové trvanie vlnky sa teda mení a tým aj jej presnosť časovej lokalizácie.

Pre vlnky krátkeho trvania, respektívne vysokej frekvencie, sa frekvenčná distribúcia rozširuje. Naopak pre vlnky dlhého trvania, v respektíve nízkej frekvencie sa zužuje. Toto je ilustrované na ľavej polovici obrázku 3.9, na ktorý je vhodné na chvíľu upriamiť pozornosť.

Uprostred obrázka sú znázornené tri rôzne Morletove vlnky. V dolnej časti sú vykreslené hustoty energií týchto vlniek v závislosti od času  $|\psi_{a,b}(t)|^2$  a od frekvencie  $|\hat{\psi}_{a,b}(f)|^2$ . Na obrázku možno ďalej vidieť, že pri stlačení v čase je vlnka zložená z vyšších frekvencií s väčším rozpätím a pri roztiahnutí je situácia opačná. Rozpätie  $|\psi_{a,b}(t)|^2$  a  $|\hat{\psi}_{a,b}(f)|^2$  možno číselne kvantifikovať pomocou  $\delta_t$  a  $\delta_f$ , smerodajných odchýliek od ich osí súmernosti.

Časové a frekvenčné rozpätia vlniek sa dajú zakresliť do časovo-frekvenčnej roviny ako boxy s dĺžkou  $2\delta_t$  a výškou  $2\delta_f$  a sú znázornené v hornej časti obrázku. Tieto boxy sú známe ako Heisenbergove podľa Heisenbergovho princípu neurčitosti, ktorý vymedzuje minimálny



Obrázok 3.9: Heisenbergove boxy pre vlnkovú a krátkodobú Fourierovu transformáciu.

obsah plochy obdĺžnika, čo pokrývajú. Presnejšie povedané, súčin  $\delta_t \delta_f$  musí byť väčší alebo rovný  $1/4\pi$ . Plocha boxu je tak prinajmenšom  $1/\pi$ . Pre Morletovu vlnku je presne toľko, pretože Gaussova distribúcia poskytuje optimálny tvar okna.

Heisenbergov princíp neurčitosti predstavuje pri analýze signálov problém rozlíšenia. Dobré časové aj frekvenčné rozlíšenie sa vzájomne vylučujú. Inými slovami, čím presnejšie časové určenie (menšie  $\delta_t$ ), tým menej presné spektrálne určenie (vyššie  $\delta_f$ ) a opačne. V čase sa dobre lokalizujú vysoké spektrálne komponenty a vo frekvenčnej oblasti nízke [29].

Ešte pre poznamenanie, vertikálnu pozíciu Heisenbergovho boxu tvaru štvorca v časovo-frekvenčnej rovine udáva centrálna frekvencia materskej Morletovej vlnky  $f_0$ .

Lahkou modifikáciou Fourierovej transformácie ju možno rozšíriť na nástroj časovo-frekvenčnej lokalizácie, krátkodobú Fourierovu transformáciu. Tá zapája do pôvodného procesu okno  $h(t - b)$ , ktoré slúži ako obálka k vymaskovaniu intervalu konečnej dĺžky z komplexnej exponenciály  $e^{-i2\pi ft}$ . Definová je ako:

$$F(f, b) = \int_{-\infty}^{\infty} x(t)h(t - b)e^{-i2\pi ft} dt \quad (3.44)$$

a vďaka oknu známa aj pod názvom oknová Fourierová transformácia [28].

Okná existujú rôzne, Hanningovo, Hammingovo, kosínusové, Kaiserovo alebo Gaussovo.

Ďalej bude uvažovaná demonštrácia oknovej Fourierovej transformácie na Gaussovom okne. Táto transformácia dostala pomenovanie Gáborová. Má tvar:

$$h(t) = \frac{1}{\sqrt{\sigma\pi^{1/4}}} e^{-\frac{1}{2}(t^2/\sigma^2)} \quad (3.45)$$

kde  $\sigma$  udáva šírku Gaussovho okna na časovej osi.

Kombinácia okna s komplexou exponenciálou vytvára časovo-frekvenčný atóm:

$$h_{f,b}(t) = h(t-b)e^{i2\pi ft} \quad (3.46)$$

Konvolúcia komplexne združenej funkcie k tomuto atómu so signálom  $x(t)$  sa prejaví ako časovo-frekvenčná dekompozícia. V detaile teda pracuje Gáborová transformácia takto:

$$F(f,b) = \int_{-\infty}^{\infty} x(t)h_{f,b}^*(t)dt = \int_{-\infty}^{\infty} x(t)\frac{1}{\sqrt{\sigma\pi^{1/4}}} e^{-\frac{1}{2}[(t-b)^2/\sigma^2]} e^{-i2\pi ft} dt \quad (3.47)$$

kde  $f$  je interná frekvencia atómu transformácie, ktorá sa mení v rámci okna pevnej dĺžky lokalizovanom v bode  $b$ . Tvar má veľmi podobný transformácii Morletovou vlnkou:

$$T(a,b) = \int_{-\infty}^{\infty} x(t)\psi_{a,b}^*(t)dt = \int_{-\infty}^{\infty} x(t)\frac{1}{\sqrt{a\pi^{1/4}}} e^{-\frac{1}{2}[(t-b)^2/a^2]} e^{-i2\pi(f_0/a)(t-b)} dt \quad (3.48)$$

Kľúčová rozdielnosť medzi danými transformáciami by mala byť teraz jasná. Ďalší menej závažný rozdiel spočíva v tom, že komplexná exponenciála Morletovej vlnky je viazaná a centrovaná na okno. Pri Gáborovej transformácii je vždy pevne ukotvená v počiatku súradnej sústavy a okno sa po nej kľže [42].

Časovo-frekvenčná rovina Gáborovej transformácie je na pravej polovici obrázku 3.9. Boxy majú pre tri rôzne interné frekvencie s tou istou dĺžkou okna rovnaké rozmery a nemusia byť nutne štvorcové. Platí to pre všetky tvary okna, nie len pre Gaussovo. Rovnako ako pri Morletovej transformácii, aj v tomto prípade majú optimálne Gaussove okná za následok minimálny obsah plochy škatuliek.

### 3.12 Praktický výpočet spojitej vlnkovej transformácie

Pri praktickom použití matematického nástroja pre vyšetrovanie fyzikálneho javu treba vopred zvážiť problémy, ktoré môžu nastať. Ani vlnková transformácia v tomto zmysle nie je výnimkou. Skutočne obdržané výsledky sú vzhľadom na teóriu obmedzené, čo pramení z mnohých príčin. Sú nimi napríklad disktrétna povaha dát, konečné rozlíšenie a rozsah alebo limitácie numerického výpočtu na stroji [36].

Prvoplánovo by sa spojitá vlnková transformácia počítala prostou diskretizáciou jej definície. Integrál by bol nahradený sumou na základe vzorkovacieho intervalu  $\Delta t$ . Parameter  $a$  by sa krokoval logaritmicky a  $b$  lineárne. Takýto prístup je však veľmi ťažkopádny. Omnoho efektívnejší je výpočet pomocou rýchlej Fourierovej transformácie. Podľa nedávnych zistení sa podarilo nájsť dokonca ešte rýchlejší algoritmus ( $O(N)$ ), konkrétne pre filtrovanie Morletovou vlnkou. Ide o rekurzívny Gáborov filter. Ten počíta konvolúciu jej rozseknutím na súčet dvoch IIR (celý názov: Infinite Impulse Response, slovensky: sústava s nekonečnou impulzovou odozvou) filtrov. Postup je relatívne nový a využíva polynomiálnu aproximáciu vlnky [40].

Ešte pre ujasnenie. Čitateľ sa môže stretnúť s pojmom disktrétna vlnková transformácia (angl. Discrete Wavelet Transform, skr. DWT), čo ale nezodpovedá diskretizovanej spojitej.

DWT na rozdiel od CWT umožňuje existenciu ortogonálnej bázy pre dokonalú rekonštrukciu pôvodného signálu.

Experimentálne signály v praxi sú konečné a často je vyžadovaná transformácia v celom rozsahu dostupných dát. Pri praktickej analýze možno pozorovať, že približovaním sa vlnky k okrajom signálu konečného rozsahu začína presahovať do neznámej oblasti. Vzniká tak *hraničný efekt*. Konce signálu majú nespojitú povahu a správajú sa ako hrany. Na to ako si poradiť s touto nepríjemnosťou už bolo vyvinutých veľa spôsobov. Patria medzi ne predĺženie signálu z oboch strán nulou alebo koncovou hodnotou, pomalý útlm, periodizácia, zrkadlenie, konvolúcia oblúkovým oknom, napasovanie polynomiálnou extrapoláciou a precyklenie vlnky. Aplikovaním hociktorej z týchto metód by si mal byť pozorovateľ vedomý, že sa k okrajom signálu dodali syntetické informácie [39].

Vo vykreslení vlnkovej transformácie je zvyčajne potrebné vedieť zvýrazniť vlastnosti týkajúce sa zvolenej mierky  $a$ . Uskutoční sa to zmenou váhovej funkcie  $w(a)$  z obvyklej hodnoty  $1/\sqrt{a}$  na inú podľa potreby. Tu si treba zas uvedomiť, že zvýraznenie jednej oblasti má za následok potlačenie výraznosti ostatných. Ďalším prostriedkom zvýraznenia je zmena lineárnej stupnice mierky na logaritmickú. Tu však tiež treba riešiť ťažkosti, avšak matematického charakteru vyplývajúce z definície logaritmu.



## Kapitola 4

# Rýchla Fourierova transformácia

Účelom tejto kapitoly je vysvetliť ako pracuje rýchla Fourierova transformácia (anglicky Fast Fourier Transform, v skratke FFT).

### 4.1 Význam a použitie

FFT predstavuje urýchľovaciu techniku výpočtu diskretnej Fourierovej transformácie (DFT), digitálnej obdoby spojitej verzie. Jedná sa o jeden z najdôležitejších a najpoužívanejších algoritmov v oblasti spracovania signálov. Pomocou FFT možno optimalizovať a zefektívniť ďalšie výpočty, medzi ktoré patrí aj diskretizovaná spojitá vlnková transformácia [37]. Existuje viacero variánt FFT, preto bude ďalej uvažovaná iba najpopulárnejšia alternatíva Cooley-Tukey.

### 4.2 Princíp

Pre pripomenutie, Fourierovu transformáciu spojitých funkcií v komplexnej doméne (všeobecnejšie, s odhliadnutím od veličín času a frekvencie analógových signálov) vypočítame ako:

$$F(u) = \int_{-\infty}^{\infty} f(x)e^{-i2\pi ux} dx \quad (4.1)$$

Prevodom spojitej funkcie vzorkovaním na dĺžku  $N$  nastáva prechod k jej diskretnej podobe:

$$F_n = \sum_{k=0}^{N-1} f_k e^{-i\frac{2\pi}{N}kn} \quad (4.2)$$

Pre DFT je tak vstupom postupnosť  $f_0, f_1, \dots, f_{N-1}$  a výsledkom  $F_0, F_1, \dots, F_{N-1}$ . Priamočiarou úvahou možno dospieť k tomu, že výpočet sa realizuje v dvojúrovňovej slučke. Pre každý z  $N$  prvkov výstupu sa spočíta  $N$  sčítancov. Časová zložitosť takéhoto postupu je neprekvapujúco stupňa  $O(N^2)$ . Vďaka určitým vlastnostiam transformácie však možno zredukovať počet operácií na stupeň  $O(N \log_2 N)$  [8]. V nasledujúcom príklade sa demonštruje urýchľujúci trik.

Nech  $N = 8$ , DFT je potom:

$$F_n = f_0 e^{-i\frac{2\pi}{8}0n} + f_1 e^{-i\frac{2\pi}{8}1n} + f_2 e^{-i\frac{2\pi}{8}2n} + f_3 e^{-i\frac{2\pi}{8}3n} + f_4 e^{-i\frac{2\pi}{8}4n} + f_5 e^{-i\frac{2\pi}{8}5n} + f_6 e^{-i\frac{2\pi}{8}6n} + f_7 e^{-i\frac{2\pi}{8}7n} \quad (4.3)$$

Výraz je možné upraviť na dva podobné súčty vytriedením párných a nepárných členov vyňatím spoločného násobiteľa takto:

$$F_n = \left[ f_0 + f_2 e^{-i\frac{2\pi}{8}2n} + f_4 e^{-i\frac{2\pi}{8}4n} + f_6 e^{-i\frac{2\pi}{8}6n} \right] + \left[ f_1 + f_3 e^{-i\frac{2\pi}{8}2n} + f_5 e^{-i\frac{2\pi}{8}4n} + f_7 e^{-i\frac{2\pi}{8}6n} \right] e^{-i\frac{2\pi}{8}n} \quad (4.4)$$

Ďalšou separáciou súm sa obdrží:

$$F_n = \left[ \left( f_0 + f_4 e^{-i\frac{2\pi}{8}4n} \right) + \left( f_2 + f_6 e^{-i\frac{2\pi}{8}4n} \right) e^{-i\frac{2\pi}{8}2n} \right] + \left[ \left( f_1 + f_5 e^{-i\frac{2\pi}{8}4n} \right) + \left( f_3 + f_7 e^{-i\frac{2\pi}{8}4n} \right) e^{-i\frac{2\pi}{8}2n} \right] e^{-i\frac{2\pi}{8}n} \quad (4.5)$$

Takýmto spôsobom združovania sa súčty rozvrstvi do troch ( $\log_2 8$ ) úrovní, ktoré sú vo vzorci rozlíšené typom zátvoriek. Druhý člen na danej úrovni je nasobený vždy tým istým činiteľom.

Formula sa dá ďalej zjednodušiť vykrátením exponentov nasledovne:

$$F_n = \left[ \left( f_0 + f_4 e^{-i\pi n} \right) + \left( f_2 + f_6 e^{-i\pi n} \right) e^{-i\frac{\pi}{2}n} \right] + \left[ \left( f_1 + f_5 e^{-i\pi n} \right) + \left( f_3 + f_7 e^{-i\pi n} \right) e^{-i\frac{\pi}{2}n} \right] e^{-i\frac{\pi}{4}n} \quad (4.6)$$

Najdôležitejším spozorovaním, ktoré vedie k urýchleniu, je periodicitu komplexných exponenciál:

$$e^{i(\phi+2\pi)} = e^{i\phi} \quad (4.7)$$

Pre exponenciály v najhlbších (jednoduchých) zátvorkách sa prvá najkratšia perióda získa pri  $n = 2$ . Znamená to, že na tejto úrovni vnorenia možno dosiahnuť rovnakého výsledku pre  $n = 0, 2, 4, 6$  a  $n = 1, 3, 5, 7$ . Je teda za potreby vykonať  $4 \times 2 = 8$  sčítaní. Keďže  $n = 1, 3, 5, 7$  korešponduje s  $n = 0, 2, 4, 6$  v polovine periódy (s rozdielom  $\pi$ ), exponenciálny činiteľ je rovnaký, avšak s opačným znamienkom:

$$e^{i(\phi+\pi)} = -e^{i\phi} \quad (4.8)$$

V skutočnosti je jeho hodnota pre párne  $n$  rovná 1 a pre nepárne  $-1$ .

Komplexné exponenciály v hranatých zátvorkách s periodicitou  $n = 4$  produkujú rovnaké súčty pre tieto páry:  $n = 0, 4$ ;  $n = 1, 5$ ;  $n = 2, 6$  a  $n = 3, 7$ . Znamená to, že pri strednej úrovni zátvorkovania je nutné počítať  $2 \times 4 = 8$  súm. Dôsledkom toho, že vzdialenosť  $n$  a  $n + 2$  je  $\pi$ , možno druhú polovicu súčtov získať opačnou hodnotou. Pre  $n = 0, 4$  je tak činiteľ rovný 1, pre  $n = 2, 6$  je  $-1$ , pre  $n = 1, 5$  zas  $-i$  a pre  $n = 3, 7$  je  $i$ .

V najvrchnejšej (nezátvorkovanej) úrovni sa vyskytuje práve jedna suma pre každý komponent transformácie. Perióda exponenciálneho násobiteľa je v  $n = 8$ . Tým sa počíta  $1 \times 8 = 8$  súm a druhú polovicu možno opäť získať zmenou znamienok v prvej.

V každom stupni výpočtu je potrebných 8 súčtov. V zmysle výpočtovej zložitosti to predstavuje  $\log_2 N$  úrovní a pre každú z nich  $N$  súm. Dohromady dáva tento postup časovú zložitost'  $O(N \log_2 N)$ . Rovnaký počet súčtov pre všetky úrovne umožňuje spracovať tie isté dáta bez potreby ich rozdelenia na vstupné a výstupné. Uvedený príklad ilustroval podstatu FFT. Ďalej sa bude pojednávať o inštrukciách krok po kroku.

### 4.3 Algoritmus

Na základe predchádzajúcej analýzy vlastností DFT možno jednoducho zapísať algoritmus FFT nasledovne [10]:

1. Príprava vstupných dát - preusporiadanie vzoriek do vhodného poradia.
2. Pre každú sumačnú úroveň (zátvorkovania):
  - (a) Pre každý exponenciálny činiteľ (po polperiódach):
    - i. Výpočet činiteľa.
    - ii. Pre každý súčet výrazov, v ktorom sa objavuje daný činiteľ:
      - A. Výpočet súčinu činiteľa a druhého člena súčtu.
      - B. Výpočet súčtu.
      - C. Výpočet rozdielu.

Prvým krokom je preusporiadať vstupné dáta do ich prirodzeného poradia:

$$(f_0, f_1, f_2, f_3, f_4, f_5, f_6, f_7) \rightarrow (f_0, f_4, f_2, f_6, f_1, f_5, f_3, f_7) \quad (4.9)$$

Toto nové poradie bolo získané ako výsledok postupného triedenia členov na párne a nepárne. V binárnej reprezentácii ide o obrátenie (reverzia, zrkadlenie) bitov podľa významnosti ako naznačuje nasledovná tabuľka:

0	000	000	000	0
1	001	010	100	4
2	010	100	010	2
3	011	110	110	6
4	100	001	001	1
5	101	011	101	5
6	110	101	011	3
7	111	111	111	7

Tabuľka 4.1: Preusporiadanie dát.

Táto zmena poradia vedie k „zozrkadlenej“ aritmetike. Dvojnásobným preusporiadaním sa takýmto spôsobom dospeje znovu k pôvodnej postupnosti.

Sumačné úrovne sú určené zátvorkovaním. Vnárание vedie k iteráciám na pároch, štvoricach, osemčlenných skupinách atď.

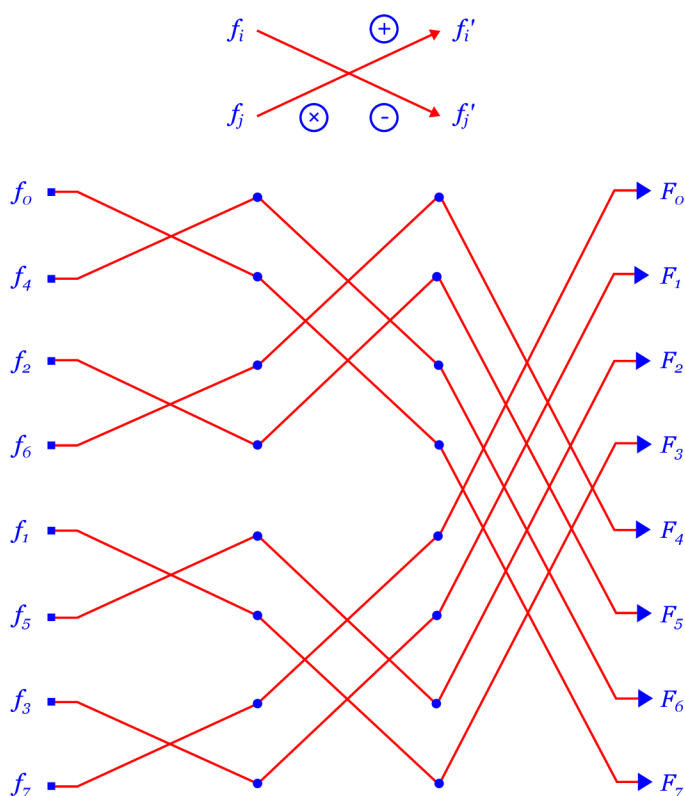
Ďalej, pri iteráciách po polperióde, druhú polperiódu od sumy odčítavame namiesto pričítavania. Pre najhlbšiu úroveň zátvorkovania perióda nastáva pri  $n = 2$  a polperióda pri  $n = 1$ , čo znamená, že cyklus bude vykonaný iba raz. V druhej úrovni je to pri  $n = 4$  a  $n = 2$ , cyklus sa tak spustí dvakrát. Vo všeobecnosti sa počet cyklov vždy zdvojnásobí, vyvíja sa teda geometricky 1, 2, 4, 8, ....

Na výpočet komplexného činiteľa sa z dôvodu obmedzenia počtu volaní trigonometrických funkcií použije rekurentný vzťah:

$$e^{i(\phi+\delta)} = e^{i\phi}e^{i\delta} = e^{i\phi} + e^{i\phi} \left( -2 \sin^2 \frac{\delta}{2} + i \sin \delta \right) \quad (4.10)$$

Výraz je rozpísaný v uvedenom tvare kôli potrebe zachovania významu pre malé  $\delta$ . V tomto prípade  $\cos \delta \approx 1$  a  $\sin \delta \approx \delta$ , z čoho vyplýva, že  $\cos \delta$  bude v pamäti uložený ako .999999(9), ale pre  $\sin \delta$  to bude omnoho užitočnejšia informácia. Vzťah 4.10 predstavuje spôsob ako eliminovať  $\cos \delta$  z výpočtu. Pri transformáciách s vysokým počtom vzoriek  $N$  by bola jeho hodnota veľmi nízka.

Ďalší výpočet FFT je vykonávaný pomocou zloženej operácie označovanej ako *motýľ*. Všeobecnú schému motýľa ilustruje horná časť obrázku 4.1. Spodný člen naľavo je násobený komplexným činiteľom a potom je jeho súčet s vrchým ľavým uložený do pravého horného a ich rozdiel do dolného poľa [27].



Obrázok 4.1: Motýľ: Všeobecná schéma (hore), schéma pre  $N = 8$  (dole).

## 4.4 Inverzná transformácia

Vzorec pre inverznú diskretnú Fourierovu transformáciu je:

$$f_k = \frac{1}{N} \sum_{n=0}^{N-1} F_n e^{i \frac{2\pi}{N} kn} \quad (4.11)$$

Rozdiel medzi doprednou a spätnou formou je iba v jednom znamienku a normalizačnom faktore, ktorý nie je vždy nevyhnutný. Niekedy netreba vedieť absolútne hodnoty transformácie, ale iba ich pomer. Dopredný postup je teda za potreby k výpočtu inverznej transformácie modifikovať iba minimálne.

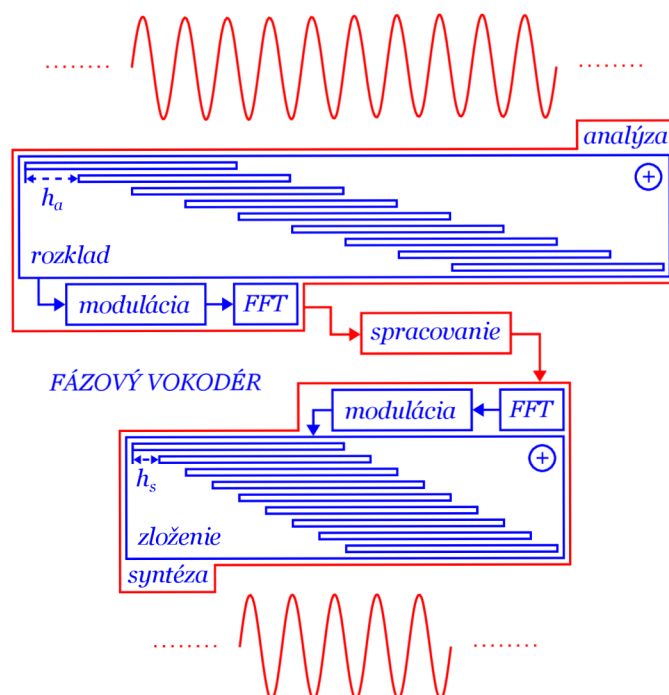
## Kapitola 5

# Fázový vokodér

V tejto kapitole sa pozornosť venuje fázovému vokodéru, konkrétnejšie popisuje ako pracuje.

### 5.1 Prehľad

Jednou z populárnych aplikácií časovo-frekvenčnej dekompozície je fázový vokodér. Jedná sa o analyzačno-resyntetizačný systém založený na STFT [13]. V užšom slova zmysle predstavuje techniku spracovania zvukových signálov, ktorá umožňuje zmenu trvania zvukovej nahrávky pri zachovaní výšky tónu a opačne.



Obrázok 5.1: Fungovanie fázového vokodéru.

Fázový vokodér vynašli J. L. Flanagan a R. M. Golden z Bell Telephone Laboratories. Prvá implementácia sa historicky datuje od roku 1966. Ide tak o časom overený, stále živý

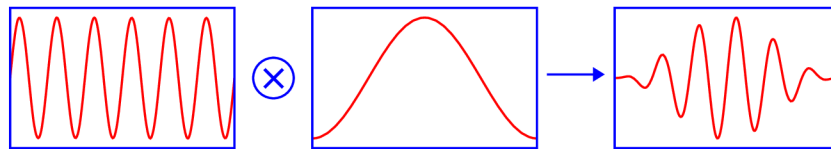
koncept s pomerne dlhou tradíciou. Pôvodne bol fázový vokodér určený ku kódovaniu reči s cieľom znížiť šírku pásma. Vďaka urýchleniu využitím FFT sa neskôr stal významný v oblasti počítačovej hudby [31].

## 5.2 Princíp

Fungovanie fázového vokodéru možno pri vyššom stupni abstrakcie rozdeliť do troch etáp. Sú nimi analýza, spracovanie a syntéza. Prvá a tretia fáza zodpovedá v podstate STFT, respektívne jej inverzii. Bloková schéma je zakreslená na obrázku 5.1.

## 5.3 Analýza

V štádiu analýzy sa prichádzajúci signál postupne segmentuje do rovnako vzdialených, čiastočne prekrývajúcich sa rámcov jednotnej dĺžky. Tie sa následne modulujú oknovou funkciou. Súčin s obálkou (obrázok 5.2) sa vykonáva z dôvodu zlepšenia frekvenčnej odozvy. Takáto „porcia“ signálu sa potom transformuje do frekvenčnej oblasti DFT, najčastejšie algoritmom FFT [32].



Obrázok 5.2: Modulácia - skalárny súčin rámca s oknom.

## 5.4 Spracovanie

DFT segmentu dĺžky  $N$  popisuje spektrálne komponenty začínajúc v jednosmernej zložke 0 až po Nyquistovu hranicu  $\frac{f_s}{2}$  s krokom  $\frac{f_s}{N}$ , kde  $f_s$  predstavuje vzorkovaciu frekvenciu signálu. Taktiež zachytáva záporné frekvencie od  $-\frac{f_s}{N}$  po  $\frac{f_s}{N} - \frac{f_s}{2}$ . Spektrum reálneho signálu je však symetrické a preto sa časť informácie stáva nadbytočnou. Energie frekvencií, ktoré presne nezapadajú do stupnice DFT sú distribuované do okolia. Z postupnosti komplexných čísel sa fáza a amplitúda získa ako argument a absolútna hodnota. Poskytovaná informácia o fáze je z dôvodu periodicity komplexnej exponenciály z intervalu  $(-\pi, \pi)$ .

Nech v ďalšom výklade  $a$ ,  $s$ ,  $k$ ,  $i$  značia analytický, syntetický aspekt, index frekvenčnej komponenty a poradie rámca. Pri modifikácii vzdialenosti dvoch po sebe idúcich segmentov sú podstatné fázové rozdiely jednotlivých komponent [16]. Možno ich tiež nazvať fázovými posunmi  $\Delta\phi[k]_i = \phi[k]_i - \phi[k]_{i-1}$ . Bez cyklickej aritmetiky by bolo možné vypočítať analytickú frekvenciu jednoducho:

$$\omega_a[k]_i = \frac{\Delta\phi_a[k]_i}{\Delta t_a} \quad (5.1)$$

kde  $\Delta t_a = \frac{h_a}{f_s}$  a ide o čas prehrávania analytickej vzdialenosti. Problém sa však stáva zložitejší odkedy je fáza modulárna. V tomto prípade sa najkôr vypočíta predpokladaná

frekvenčná odchýlka:

$$A = \frac{\Delta\phi_a[k]_i}{\Delta t_a} - \omega[k] \quad (5.2)$$

kde symbol  $\omega[k]$  označuje spektrum. Následne je zbalená a namapovaná do základného intervalu:

$$B = \text{mod}(A + \pi, 2\pi) - \pi \quad (5.3)$$

Jej pripočítaním k frekvencii DFT sa získa už skutočná hodnota:

$$\omega_a[k]_i = B + \omega[k] \quad (5.4)$$

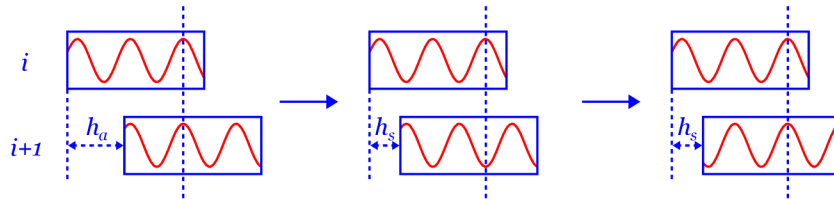
Nová syntetická fáza vznikne rekurentne:

$$\phi_s[k]_i = \phi_s[k]_{i-1} + \omega_a[k]_i \times \Delta t_s \quad (5.5)$$

kde  $\Delta t_s$  analogicky vyjadruje syntetickú rýchlosť. Pre konečné hodnoty magnitúdy a fázy spektra platí:

$$|X_s[k]_i| = |X_a[k]_i| \quad \angle X_s[k]_i = \phi_s[k]_i \quad (5.6)$$

Vizuálny priebeh zachovania koherencie je znázornený na nasledujúcom obrázku.



Obrázok 5.3: Vyladenie fázy.

## 5.5 Syntéza

Prispôbením fázy novej segmentačnej vzdialenosti je nutné previesť rámec naspäť do časovej domény. K tomuto účelu posluží inverzná DFT. Jej výsledok je opäť vynásobený obálkou a normalizačným faktorom. Jednotlivé prekrývajúce sa segmenty je za potreby ešte spočítať. Týmto krokom výpočtu je už generovaný samotný výstupný signál [12].

Ak je v záujme použitia aj hudobná transpozícia (posun tónu), zvuk sa dodatočne prevzorkuje pomocou lineárnej alebo inej interpolácie. Dĺžka sa však zmení. Preto pri znížení tónu o oktávu treba najskôr zkrátiť signál na polovicu a následne jeho trvanie zdvojnásobiť interpoláciou. Oba typy efektov však možno použiť súčasne.

# Kapitola 6

## Implementácia

V rámci tohto projektu bol napísaný zdrojový text jednoduchej aplikácie fázového vokodéru. Nasledujúca kapitola sa zaoberá podrobnejším popisom implementácie.

### 6.1 Návrh architektúry programu

Keďže fázový vokodér predstavuje v ponímaní spracovania signálov systém, možno preň špecifikovať vstupy a výstupy. Na rozdiel od klasického prístupu je však daný program rozšírený o vykresľovanie spektrogramu a prevzorkovanie. Vstupom tak zostáva jeden audio signál. Výstupmi sú buď obraz alebo zvuk.

Obrázok 6.1 ilustruje architektúru sústavy vykonávajúcu sa v cykle, pokiaľ je na vstupe prítomný signál. Vyrovnávacia pamäť, cez ktorú plynie, vždy zachytí dva susedné segmenty  $x$  a  $y$ . Tie sa rovnako normalizujú a vynásobia oknom. Oba potom putujú do bloku FFT. Tu prvý segment predstavuje reálnu a druhý imaginárnu časť, pretože FFT pracuje v obore komplexných čísel. Spracovať rámce po jednom by bolo tiež možné tak, že imaginárna zložka by sa nastavila na nulu. Riešenie pomocou pseudo-reálneho signálu je jednoduchšie ale neefektívne.

Po transformácii je nutné oddeliť spektrá  $X$ ,  $Y$ , ktoré sú vo výsledku zakódované, týmto spôsobom [23]:

$$\begin{aligned} r(X[f]) &= r\left(\frac{Z[f] + Z[-f]}{2}\right) & i(X[f]) &= i\left(\frac{Z[f] - Z[-f]}{2}\right) \\ r(Y[f]) &= i\left(\frac{Z[f] + Z[-f]}{2}\right) & i(Y[f]) &= r\left(\frac{Z[-f] - Z[f]}{2}\right) \end{aligned} \quad (6.1)$$

kde  $Z$  je FFT zmiešaných segmentov,  $f$  predstavuje kladné,  $-f$  záporne frekvencie,  $r$  reálnu a  $i$  imaginárnu časť hodnoty. Vypočítať dva reálne signály naraz umožňuje linearita operácie.

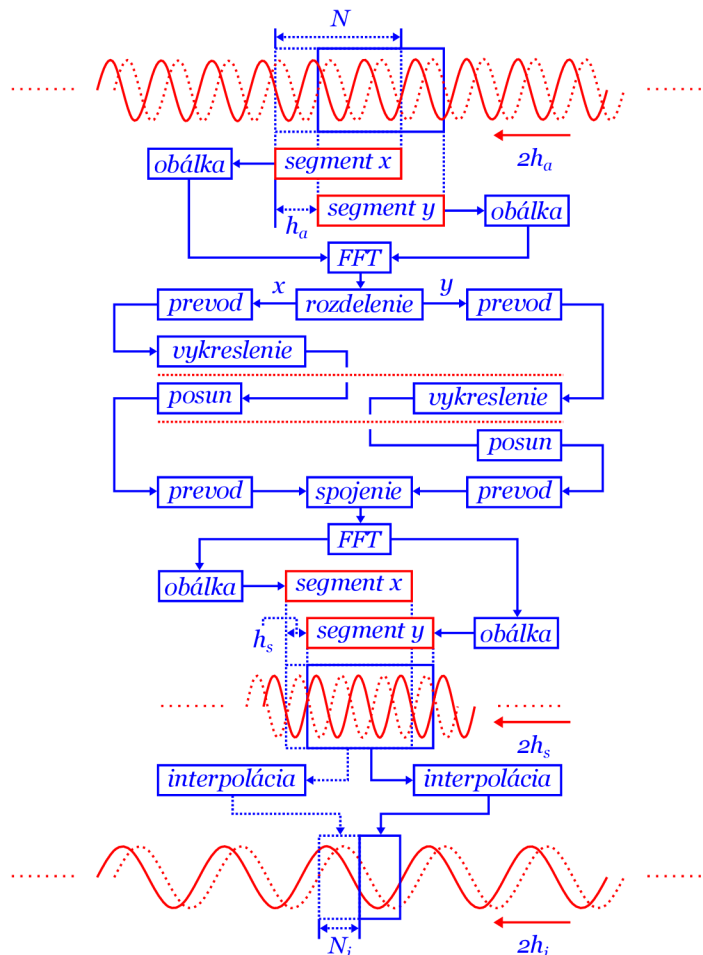
Ďalším krokom je konverzia komplexných postupností do magnitudovo-fázovej reprezentácie. Na poradí segmentov pri vykonávaní nezáleží, prevody teda môžu prebiehať paralelne.

Fázu vykreslenia je už nutné uskutočniť synchronne. Najskôr so spektrom  $X$ , potom  $Y$ . Graf je trojrozmerný s lineárnou časovou a frekvenčnou osou. Amplitúda rastie logaritmicky.

Po tomto úkone sa zosúladiť fázy v určenom poradí. Upravené segmenty sa späťne prevedú do komplexnej podoby a analogicky k rozkladu spoja do „zlepeného“ tvaru. Inverzná FFT nasledovne syntetizuje dva segmenty, ktoré sa normalizujú a modulujú tou istou obálkou ako pri čítaní.



Segmenty sú potom pripočítané do akumuláčného posuvného poľa. Časti, ktoré pri vytesňovaní novými rámcami odchádzajú sú napokon prevzorkované prostredníctvom lineárnej interpolácie a sekvenčne zapísané na výstup.



Obrázok 6.1: Jadro slučky.

## 6.2 Graf hustoty energie

Časová a frekvenčná os grafického výstupu sú škálované lineárne. Čas je zaznamenaný v analytickom skoku, ktorý sa v extrémnom prípade dá nastaviť na jednu vzorku. Frekvencie sú zachytené všetky práve raz, presne podľa DFT. Veľkosť energie daného bodu je vykreslená v logaritmickú mierke z dôvodu zvýraznenia nižších hodnôt. Odtieň pixelu tak zodpovedá:

$$P = \log_a |x[k]|^2 \quad (6.2)$$

kde  $a$  je základ logaritmu a  $|x[k]|^2$  je druhá mocnina absolútnej hodnoty frekvenčnej komponenty  $[k]$ , čiže energia. Výraz sa dá ďalej zjednodušiť na:

$$P = 2 \frac{\ln |x[k]|}{\ln a} \quad (6.3)$$

Keďže paleta farieb je kvantovaná na 256 úrovní s maximálnym stupňom 255 a najvyššia možná energia jednej spektrálnej zložky zas  $2^{15}$  (FFT 16-bitového bipolárneho signálu normalizovaná dĺžkou), platí tento vzťah:

$$\log_a (2^{15})^2 = 255 \quad (6.4)$$

ktorý možno ďalej previesť na:

$$2 \times 15 \log_a 2 = 255 \quad (6.5)$$

čo sa dá ešte vykrátiť a získať tak:

$$\frac{\ln 2}{\ln a} = \frac{17}{2} \quad (6.6)$$

Požadovaný výraz pre menovateľ je:

$$\ln a = \frac{2 \ln 2}{17} \quad (6.7)$$

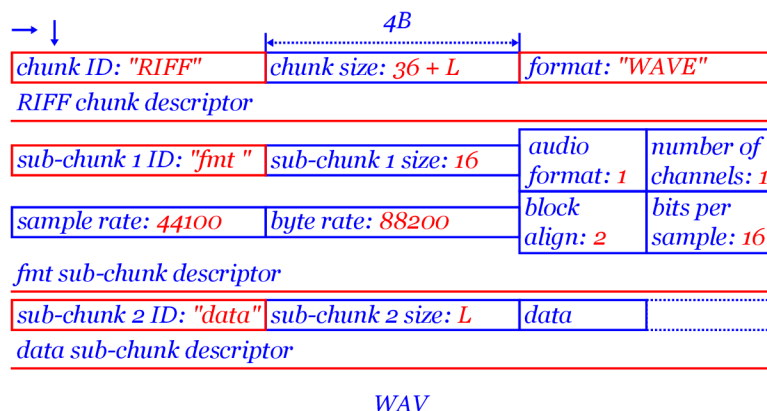
a výsledná hodnota pixelu:

$$P = \frac{17}{\ln 2} \ln |x[k]| \quad (6.8)$$

Všetky obrázky grafu hustoty energie v tejto práci boli invertované na negatív aby sa ušetrila farba pri tlači a zmenšené do požadovaných rozmerov.

### 6.3 Prostriedky

Zdrojový kód aplikácie je naprogramovaný v čistej forme jazyka C (C99). Nevyužíva žiadne neštandardné externé knižnice. Je maximálne prenositeľný, pretože program možno zostaviť sadou kompilátorov GCC. Vyvíjaný bol pod operačným systémom Windows 7 v integrovanom prostredí Code::Blocks, ktoré využíva MinGW. Spúšťa sa v konzole a s okolím komunikuje prostredníctvom súborov podľa WAV a BMP špecifikácie. Monofonický zvuk a monochromatický (odtiene sivej) obraz musia spĺňať obmedzenia na obrázkoch 6.2 a 6.3.

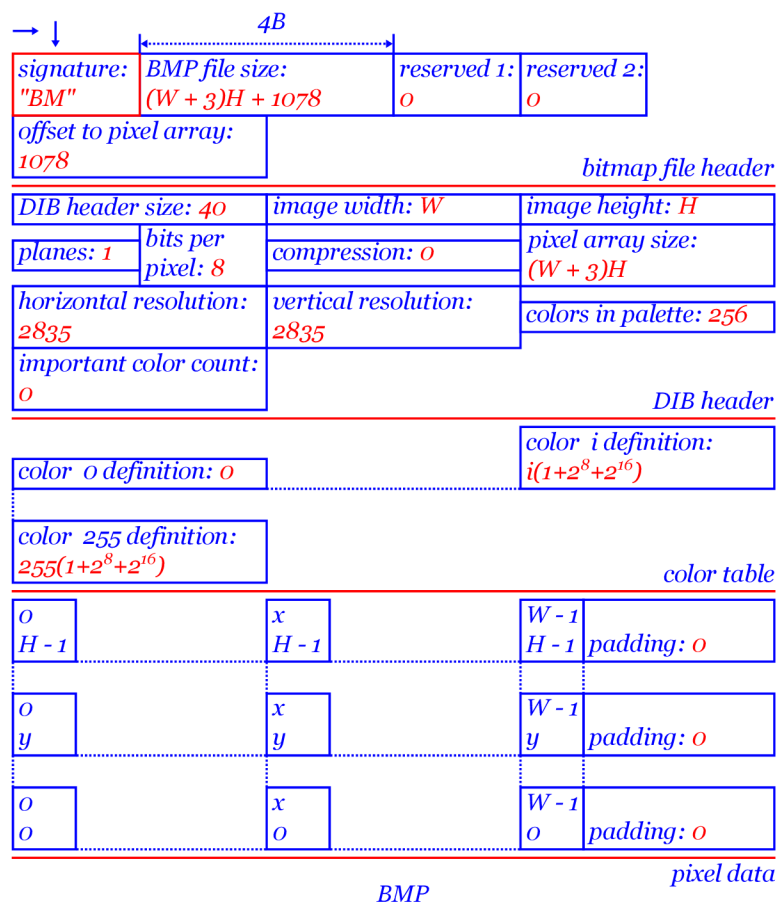


Obrázok 6.2: Špecifikácia WAV súboru.

V prípade WAV sa jedná o nekomprimovaný, jednokanálový, pulzne-kódovo modulovaný zvuk so 16-bitovou lineárnou kvantizáciou a vzorkovaním 44.1 kHz. Ako možno vidieť

na diagrame, súbor sa skladá z hlavičiek RIFF, fmt a data, ktoré sa do seba zanárajú a deklarujú typ. Ďalej nasledujú samotné údaje o bodov signálu.

Čo sa týka BMP, ten je zas nekomprimovaný s 256 úrovňami intenzity pixelu. Obdobne je rozdelený na bitmapovú, DIB hlavičku s definíciou farebnej palety a dátami.



Obrázok 6.3: Špecifikácia BMP súboru.

## 6.4 Zložitosť a rýchlosť

Časová zložitosť výpočtu je lineárna  $O(N)$ , čiže priamo úmerne závisí od dĺžky vstupu. Pamäťová je konštantná  $O(1)$  aj vďaka tomu, že spektrogram je generovaný po riadkoch. V prípade obrátenia o 90 stupňov v smere hodinových ručičiek (tak ako býva bežne orientovaný), by musel byť najskôr načítaný celý do pamäte, pretože BMP ukladá dáta po riadkoch.

# Kapitola 7

## Výsledky

Účelom tejto kapitoly je prezentácia dosiahnutých výsledkov od grafického výstupu analýzy cez preukázanie princípu neurčitosti a testovanie rýchlosti aplikácie po rozbor spracovania signálu fázovým vokodérom.

### 7.1 Spektrogramy

Ako už bolo spomenuté, súčasťou aplikácie je vizualizácia časovo-frekvenčnej analýzy pomocou STFT. Obrázok 7.1 zachytáva grafy piatich rôznorodých signálov. Všetky sú dostatočne krátke a vygenerované pri dĺžke okna 1024 so 75% faktorom prekrytia.

Prvým (a) je 8-sekundová zvučka elektronickej hudby, konkrétne inštrumentálnej skladby Axel F z filmu Policajt v Beverly Hills. V grafickej reprezentácii sa dajú jasne rozpoznať údery bicích nástrojov, ktorých charakteristiky zaberajú skoro celú šírku frekvenčného pásma a sú blízke bielemu šumu.

Druhým signálom (b) je orchestrálna fráza z kompozície The Legend of Zelda Main Theme Medley s dĺžkou 8 sekúnd. Tu nie sú prítomné žiadne perkusie, no zreteľne sa javí vývoj melódie vo vyšších frekvenciách.

Ďalšou ukážkou (c) je 2 sekundy trvajúci rečový signál „Alfréde, jedem na nákup“ z filmu Batman.

Štvrtým a piatym príkladom sú 2-sekundové testovacie signály podobné cvrlikaniu, takzvaný lineárny (d) a exponenciálny chirp (e). V oboch prípadoch ide o sínusovú vlnu, ktorej perióda sa časom spojitely skraca. Oba štartujú od jednosmernej zložky a končia v Nyquistovej frekvencii.

Lineárny chirp bol vytvorený na základe nasledovnej definície:

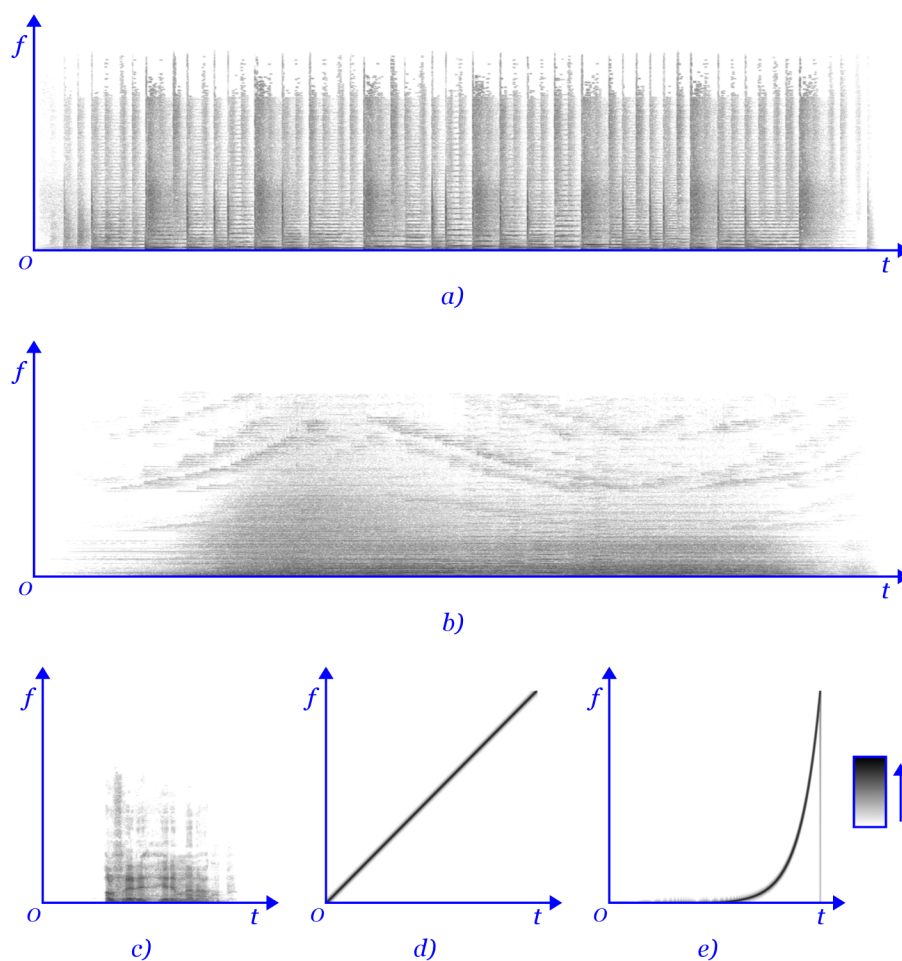
$$x(t) = \sin \left[ 2\pi \left( f_0 + \frac{k}{2} t \right) t \right] \quad (7.1)$$

kde  $k$  je tempo nárastu a  $f_0$  počiatočná frekvencia.

Exponenciálne urýchľovanie vlnenia sa vykonáva zas týmto štýlom:

$$x(t) = \sin \left[ 2\pi f_0 \frac{k^t - 1}{\ln(k)} \right] \quad (7.2)$$

V prvom prípade sa  $f_0$  rovná 0, v druhom 1. Pre oba signály je  $k$  polovica zo vzorkovacej frekvencie, čiže 22050.



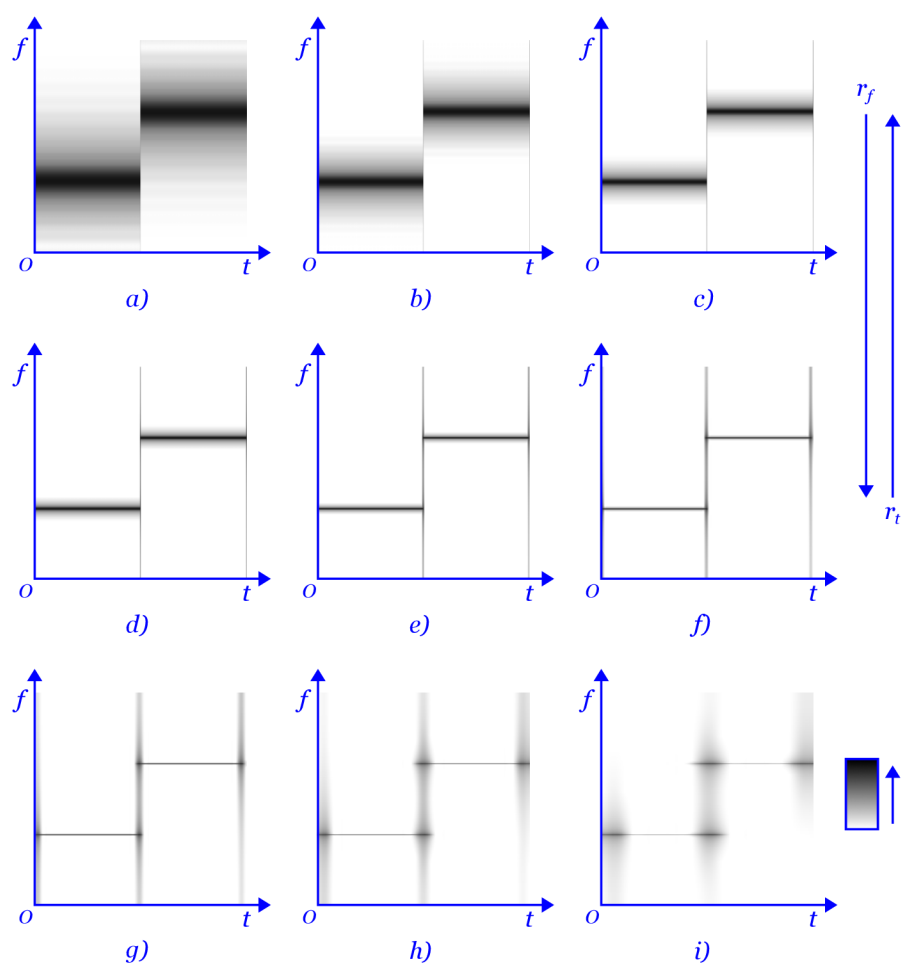
Obrázok 7.1: Časovo-frekvenčná analýza rôznorodých signálov: a) elektro b) symfonická hudba c) reč (mužský hlas) d) lineárne a d) exponenciálne frekvenčne modulovaný sínus

## 7.2 Princíp neurčitosti

Parametrizácia operácie STFT umožňuje demonštráciu Heisenbergovho princípu neurčitosti v praxi (obrázok 7.2). Pre názornosť je skúšobný signál veľmi jednoduchý. Ide o postupnosť dvoch sínusov s frekvenciami šesty a tretiny Nyquistovej. Od spektrogramu (a) po (i) sa dĺžka rámca vždy zdvojnásobí, tak že sa začína s  $2^6$  a končí pri  $2^{14}$ . Faktor prekrytia je zakaždým 75%.

Je celkom očividné, že sa pri zväčšovaní dĺžky segmentu zlepšuje spektrálne rozlíšenie  $r_f$  (horizontálne pásy sa zužujú). Rozhrania (vetikálne pruhy, v dôsledku skoku) sa naopak rozširujú, čo znamená znižovanie časového rozlíšenia  $r_t$ . S narastajúcou dĺžkou možno tiež pozorovať aj to ako sa začínajú za sebou idúce vlny prekrývať.

Presnosť časovej a frekvenčnej lokalizácie sa tak vzájomne vylučuje. Jedna je na úkor druhej. Tento problém čiastočne riešia metódy s viacnásobným rozlíšením (CWT), ktoré vychádzajú z predpokladu, že nízke frekvencie vyžadujú lepšie spektrálne a vysoké, časové určenie.



Obrázok 7.2: Experimentálny dôkaz Heisenbergovho princípu.

### 7.3 Porovnanie rýchlostí módov

Fázový vokóder je nastaviteľný tak, aby vykonávanie jednotlivých operácií mohlo bežať vo viacerých režimoch. Doba výpočtu je závislá najmä od dĺžky okna  $\mathbf{N}$  a vstupného faktoru prekrytia  $\mathbf{O}$ .

Pre lepšiu predstavu o rozdieloch v rozličných módoch bola preskúšaná hudobná transpozícia (tonálny posun) o oktávu nižšie. Trvanie spracovania referenčného 2-sekundového súboru na CPU AMD Turion™64 X2 Dual-Core Mobile Technology TL-58 1.9 GHz je zaznamenané v tabuľke 7.1. Červenou farbou sú vyznačené hodnoty, pre ktoré by daný procesor nestíhal uskutočňovať analýzu a resyntézu v reálnom čase. Z nameraných hodnôt možno usúdiť, že predlžovaním rámca ako aj zvyšovaním faktoru prekrytia rastie časová náročnosť.

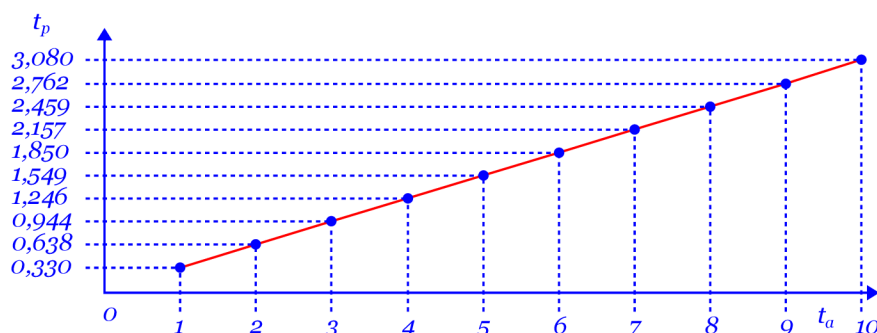
K tomu, aby bolo možné prepracovať program do real-time verzie (nemá zmysel pri zmene trvania), je nutné, aby spĺňal podmienky pre lineárnu časovú a konštantnú pamäťovú zložitosť. Ak by výpočet prebiehal rýchlejšie ako priamo úmerne k času, dáta, ktoré čakajú na zápis by sa hromadili a pamäť pretiekla. S vyššou zložitou fungovať už tiež nemôže, lebo by rovnomerný tok údajov nezvládol.

Preukázanie o linearite rýchlosti poskytuje obrázok 7.3, kde  $t_a$  je čas prehrávania zvuku

O	N	$2^6$	$2^7$	$2^8$	$2^9$	$2^{10}$	$2^{11}$	$2^{12}$	$2^{13}$	$2^{14}$
$8/16 = 1/2$		0,355	0,358	0,359	0,368	0,369	0,417	0,430	0,506	0,544
$9/16$		0,398	0,406	0,415	0,422	0,466	0,471	0,489	0,545	0,575
$10/16 = 5/8$		0,462	0,479	0,494	0,503	0,522	0,542	0,546	0,641	0,666
$11/16$		0,530	0,542	0,547	0,562	0,592	0,603	0,623	0,702	0,878
$12/16 = 3/4$		0,676	0,703	0,711	0,734	0,755	0,771	0,798	0,952	0,972
$13/16$		0,906	0,937	0,948	0,953	0,956	0,988	1,021	1,203	1,369
$14/16 = 7/8$		1,315	1,319	1,359	1,375	1,382	1,486	1,498	1,755	1,966
$15/16$		2,578	2,609	2,641	2,656	2,664	2,727	2,882	3,423	3,896

Tabuľka 7.1: Čas výpočtu v rôznych režimoch (jednotka - sekunda).

( $t_{\text{audio}}$ ) a  $t_p$  zas trvanie spracovania ( $t_{\text{processing}}$ ) v sekundách. Testovaná funkcionality bola opäť transpozícia na polovičnú frekvenciu so 75% prekrytím a oknom dlhým 1024 vzorkov.



Obrázok 7.3: Lineárna výpočtová zložitosť (závislosť doby spracovania od trvania nahrávky, meraná v sekundách).

## 7.4 Zmena trvania a tónu nahrávky

Nech už ide o skrátenie, či predĺženie trvania, zvýšenie alebo zníženie tónu zvukového signálu, všetky tieto úkony prevádza fázový vokodér deštruktívne. Znamená to, že výstup je čiastočne zdeformovaný, spracovanie zanecháva typické artefakty a inverznou operáciou už nemožno dosiahnuť dokonalý pôvodný zvuk. Ak by to naopak išlo, jednalo by sa o silný nástroj bezstratovej kompresie (v prípade skracovania).

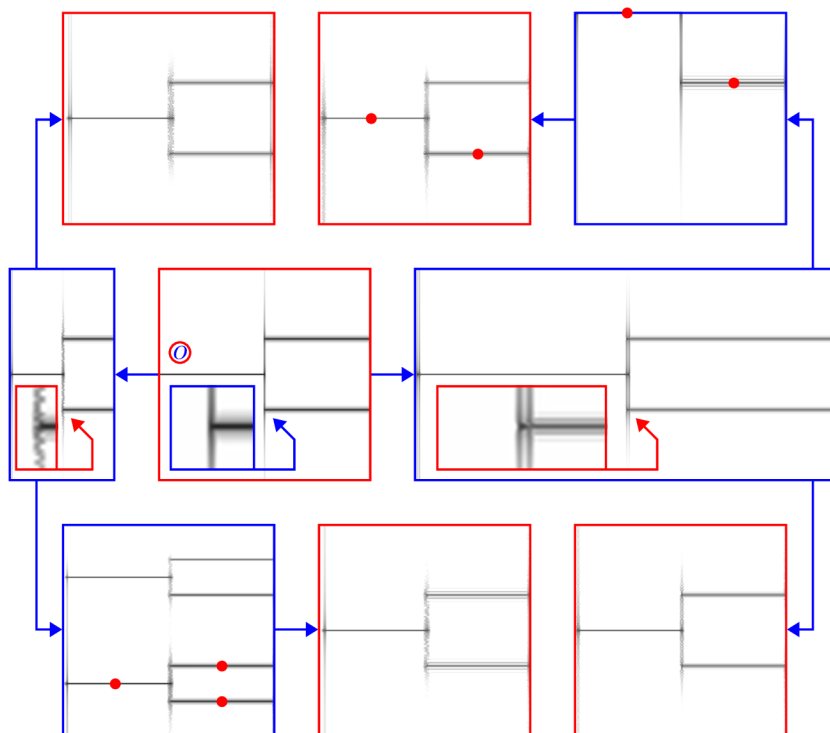
Na obrázku 7.4 sú zaskreslené štyri reťazce od zdroja (označeného 0) po jeho rekonštrukcie cez všetky manipulácie. Referenčný 2-sekundový signál predstavuje prechod od sínusovej vlny s polovicou do tretiny a šestiny Nyquistovej frekvencie. Štvornásobné priblíženie 50%-ne stlačeného a 200%-ne roztiahnutého signálu ilustruje vplyv fázového vokodéru na náhle frekvenčné zmeny.

Rozhranie predĺženej verzie je akoby zdvojené a hrany skrátenej nie sú zas hladké. Vrátením späť do pôvodnej dĺžky sa chyby kombinujú.

Reharmonizáciou o oktávu vyššie zanikajú frekvencie presahujúce pásmo DFT. Pri

znížení sa zas zrkadlovo cez stred duplikujú. Pravdepodobne to spôsobuje lineárna interpolácia, ktorá sa zároveň správa ako anti-aliasingový filter. Prevzorkovanie by však mohlo byť uskutočnené aj ďalšími typmi interpolácie (polynomiálna, spline).

Priemerný výkon výstupného signálu je z dôvodu sekundárnej modulácie oknom vždy nižší ako pri pôvodnej nahrávke. Veľkosť úhrnej energie sa však mení. Rozťahovaním a posunom smerom nahor rastie, skracovaním a podlaďovaním klesá.



Obrázok 7.4: Čiastočná deštruktívnosť systému na výstupy a spätná rekonštrukcia.



# Kapitola 8

## Záver

Tento diplomový projekt bol vypracovaný v dvoch etapách zodpovedajúcich zimmému a letnému semestru školského roku.

Prvá polovica bola venovaná štúdiu teórie s následným zdokumentovaním nadobudnutých vedomostí. Pozornosť sa sústredila predovšetkým na základy CWT, ktorá sa neskôr, v praktickej fáze, stala problematickou. V rámci teórie bola spomenutá a stručnejšie popísaná aj STFT v porovnaní kritických vlastností s uvádzanou metódou.

V druhom štádiu sa už riešila praktická stránka. Hlavným zámerom bolo porovnať výkonnosť a efektivitu implementácií metód CWT a STFT v jednej aplikácii časovo-frekvenčnej analýzy. Úspešne sa však podarilo naprogramovať zmenu trvania dĺžky zvukovej nahrávky zmysluplne iba jedným spôsobom a to STFT. Pri snahe o realizáciu CWT sa narazilo na viacero prekážok. Výpočet podľa pôvodnej formulácie konvolúciou sa preukázal ako extrémne časovo zložitý a neprichádzal už viac do úvahy. V prípade alternatívnej definície (operácie vo frekvenčnej doméne) by zas CWT pravdepodobne zodpovedala viacnásobnej STFT so spojenou zmenou dĺžky okna a pásmovým filtrovaním, čo už nepredstavuje zrovna triviálny problém. Experimentálna implementácia CWT tak bola z projektu vylúčená a úsilie sa po neúspechu preorientovalo na osvedčenú techniku fázového vokodéru.

Zdá sa, že program, ktorý bol v rámci projektu vyvinutý, pracuje správne a efektívne. Samozrejme, poskytuje ešte veľký priestor pre odstránenie nedostatkov, rozšírenie a optimalizáciu. Prvým významným vylepšením by mohla byť normalizácia energie signálu, pretože pri modulácii spracovaného rámca oknovou funkciou jej časť zaniká. Ďalej sa dá prepísať zdrojový kód na aplikáciu reálneho času, tak aby bol vykonávaný na jednotke DSP. Pre spríjemnenie používania by bolo zas vhodné dodať grafické užívateľské rozhranie a umožniť manipuláciu s viacerými formátmi.

Aj keď fázový vokodér nepatrí práve k najnovším výmyslom, snáď by mohol byť jednoduchý program aspoň drobným prínosom pre tých, ktorí nemajú zakúpený komerčný softvér, alebo ho ocenia používatelia uprednostňujúci minimalistické aplikácie k danému účelu pred komplexnými systémami.

Časovo-frekvenčná analýza ako téma bola pre autora nová, ale zato veľmi zaujímavá, zábavná a náučná. Za prípadné chyby sa čitateľovi ospravedlňuje.

# Literatúra

- [1] Addison, P. S.: *The Illustrated Wavelet Transform Handbook*. IOP Publishing, 2002, ISBN 0-7503-0692-0.
- [2] Akansu, A.; Haddad, R.: *Multiresolution Signal Decomposition: Transforms, Subbands, Wavelets*. Academic Press, 1992, ISBN 0-12-047140-X.
- [3] Artin, M.: *Algebra*. Prentice Hall, 1991, ISBN 978-0-89871-510-1.
- [4] Bařina, D.: Vlnková transformace: Pojednání k tématu disertační práce. 2011.
- [5] Blatter, C.: *Wavelets: A Primer*. A K Peters, 1998, ISBN 1-56881-095-4.
- [6] Boashash, B.: *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*. Elsevier Science, 2003, ISBN 0-08-044335-4.
- [7] Bourbaki, N.: *Topological Vector Spaces*. Springer-Verlag, 1987, ISBN 978-3540136279.
- [8] Bourke, P.: DFT, FFT [online]. <http://paulbourke.net/miscellaneous/dft/>, 1993 [cit. 2012-01-17].
- [9] Brown, W.: *Matrices and Vector Spaces*. CRC Press, 1991, ISBN 978-0-8247-8419-5.
- [10] Chernenko, S.: Fast Fourier transform [online]. <http://www.librow.com/articles/article-10>, 2007 - 2011 [cit. 2012-01-17].
- [11] Daubechies, I.: *Ten Lectures on Wavelets*. SIAM, 1992, ISBN 0-89871-274-24.
- [12] Dolson, M.: The Phase Vocoder: A Tutorial [online]. <http://www.panix.com/~jens/pvoc-dolson.par>, 2000-11-17 [cit. 2012-02-17].
- [13] Flanagan, J. L.; Golden, R. M.: Bell System Technical Journal 45. 1966.
- [14] Gençay, R.; Selçuk, F.; Whitcher, B.: *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Academic Press, 2001, ISBN 0-12-279670-5.
- [15] Goswami, J. C.; Chan, A. K.: *Fundamentals of Wavelets: Theory, Algorithms, and Applications*. Wiley, 2011, ISBN 978-0-470-48413-5.
- [16] Grondin, F.; Vakili, A.; Demers, L.: Guitar Pitch Shifter [online]. <http://www.guitarpitchshifter.com/algorithm.html>, 2009 [cit. 2012-2-15].

- [17] Hubbard, B. B.: *The World According to Wavelets*. A K Peters, 1996, ISBN 1-56881-047-4.
- [18] Jan, J.: *Číslíková filtrace, analýza a restaurace signálů*. VUTIUM, 2002, ISBN 80-214-1558-4.
- [19] Kaiser, G.: *A Friendly Guide to Wavelets*. Birkhäuser, 2010, ISBN 978-0-8176-8110-4.
- [20] Lathi, B. P.: *Signal Processing and Linear Systems*. Oxford University Press, 2000, ISBN 978-0195219173.
- [21] Mallat, S.: *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, 2008, ISBN 978-0-12-374370-1.
- [22] May, E.: *Auto-Tune vs. the Vocoder: Differences, Similarities, and Artists Who Use Them*. Webster's Digital Services, 2011, ISBN 978-1241724283.
- [23] McGee, K. J.: *An Introduction to Signal Processing and Fast Fourier Transform*. 2008.
- [24] Miles, J. W.: *Integral Transforms in Applied Mathematics*. Cambridge University Press, 2008, ISBN 978-0521090681.
- [25] Mohlenkamp, M. J.; Pereyra, M. C.: *Wavelets, Their Friends, and What They Can Do for You*. European Mathematical Society, 2008, ISBN 978-3-03719-018-0.
- [26] Ondráček, O.: *Signály a systavy*. Vydavatelstvo STU, 2008, ISBN 978-80-227-2956-7.
- [27] Oppenheim, A. V.; Schaffer, R. W.: *Digital Signal Processing*. Prentice Hall, 1975, ISBN 978-0132146357.
- [28] Percival, D. B.; Walden, A. T.: *Wavelet Methods for Time Series Analysis*. Cambridge University Press, 2006, ISBN 0-5216-8508-7.
- [29] Polikar, R.: The Engineer's Ultimate Guide to Wavelet Analysis: The Wavelet Tutorial [online]. <http://users.rowan.edu/~polikar/wavelets/wttutorial.html>, 2001-01-12 [cit. 2011-12-12].
- [30] Radunović, D. P.: *Wavelets: From Math to Practice*. Springer, 2009, ISBN 978-3-642-00613-5.
- [31] Roads, C.: *The Computer Music Tutorial*. The MIT Press, 1996, ISBN 0-262-68082-3.
- [32] Roads, C.; Pope, S. T.; Piccialli, A.; aj.: *Musical Signal Processing*. Routledge, 1997, ISBN 90-265-1482-4.
- [33] Ruch, D. K.; Fleet, P. J. V.: *Wavelet Theory: An Elementary Approach with Applications*. Wiley, 2009, ISBN 978-0-470-38840-2.
- [34] Schaumann, R.: *Design of Analog Filters*. Oxford University Press, 2001, ISBN 978-0195118773.
- [35] Vargic, R.: *Wavelety a banky filtrův*. Vydavatelstvo STU, 2004, ISBN 80-227-2093-3.

- [36] Vetterli, M.; Kovacevic, J.: *Wavelets and Subband Coding*. Prentice Hall, 1995, ISBN 0-13-097080-8.
- [37] Walker, J. S.: *Fast Fourier Transforms (Studies in Advanced Mathematics)*. CRC Press, 1991, ISBN 978-0849371547.
- [38] Weidmann, J.: *Linear Operators in Hilbert Spaces*. Springer-Verlag, 1980, ISBN 978-0-387-90427-6.
- [39] Wickerhauser, M. V.: *Adapted Wavelet Analysis: From Theory to Software*. A K Peters Ltd, 1994, ISBN 1-56881-041-5.
- [40] Young, I. T.; van Vliet, L. J.; van Ginkel, M.: Recursive Gabor Filtering. 2000 [cit. 2012-4-16].
- [41] Young, N.: *An Introduction to Hilbert Space*. Cambridge University Press, 1988, ISBN 0-521-33071-8.
- [42] Zimmermann, J.: *Spektrografická a škálografická analýza akustického rečového signálu*. Náuka, 2002, ISBN 80-89038-22-0.

# Príloha A: Obsah CD

Súbory nachádzajúce sa na priloženom CD sú utriedené v tejto štruktúre:

- / : sprievodca `readme.txt`
- /con : príspevok do zborníku EEICT - `pvoc.pdf`
  - /tex : zdrojové texty v  $\text{\LaTeX}$
- /doc : dokument - `diplo.pdf`
  - /tex :  $\text{\LaTeX}$
- /prg : program
- /bin : spustiteľné súbory
  - /linux : linux, binárny kód `pvc` a testovací skript `go.sh`
  - /windows, strojový kód `pvc.exe` a skript `go.bat`
- /dat : dáta
  - /in : vstupné
  - /out : výstupné
- /src : zdrojový kód v C a `Makefile`
- /sld : prezentácia - `pre.pdf`
  - /tex :  $\text{\LaTeX}$

# Príloha B: Manuál

Obsluha programu `pvc` je jednoduchá. Spúšťa sa v konzole. Používateľ volí jeden z nasledovných režimov zadaním reťazca argumentov:

1. `h`
2. `s [i] [o] [m] [w] [f]`
3. `p [i] [o] [m] [w] [f] [t] [s]`

kde módmí sú:

- `h` - nápoveda
- `s` - spektrogram
- `p` - fázový vokodér

a premennými:

- `[i]` - cesta pre vstupný súbor
- `[o]` - cesta pre výstupný súbor
- `[m]` - exponent: dĺžka rámca  $n = 2^m$ ,  $m \in \{6, 7..14\}$
- `[w]` - typ okna: 1 (trojuholníkové), 2 (Hannovo), 3 (Hammingovo), 4 (Gaussovo)
- `[f]` - faktor prerokrytia:  $f \in (0.5, 1)$
- `[t]` - faktor zmeny trvania:  $t > 0$
- `[s]` - faktor posunutia tónu:  $s > 0$

! faktory musia spĺňať kritické obmedzenie:  $(1 - f) \times t \times s \leq 0.5$

! kompatibilné sú iba typy súborov podľa špecifikácií na obrázkoch 6.2 a 6.3.

Príklady:

```
./pvc s a.wav i.bmp 12 1 0.875
```

```
./pvc p x.wav y.wav 10 3 0.75 1.25 1.5
```