

Univerzita Palackého v Olomouci  
Přírodovědecká fakulta

DIPLOMOVÁ PRÁCE

2009

Zdeňka Prokšová

Přírodovědecká fakulta Univerzity Palackého v Olomouci  
Katedra matematické analýzy a aplikací matematiky

# Fourierova transformace a její použití na analýzu samohlásek

Zdeňka Prokšová

Diplomová práce



Vedoucí diplomové práce:  
**RNDr. Tomáš Fůrst, Ph.D.**  
Rok odevzdání: 2009

Vypracovala:  
**Zdeňka Prokšová**  
AME, 2. ročník navazující

### **Prohlášení**

Prohlašuji, že jsem diplomovou práci zpracovala samostatně pod vedením RNDr. Tomáše Fürsta, Ph.D. a výhradně s použitím uvedené literatury.

V Olomouci dne 28. listopadu 2009

## **Poděkování**

Velmi ráda bych poděkovala a vyslovila uznání především svému vedoucímu diplomové práce RNDr. Tomáši Füstovi, Ph.D. za jeho trpělivé vedení, množství praktických rad, podněty, připomínky a odbornou pomoc, kterou mi poskytl při vypracování této práce. Také děkuji všem, kteří mi poskytli nahrávky svých samohlásek, za jejich trpělivost a spolupráci při nahrávání.



# Obsah

<b>1</b>	<b>Úvod</b>	<b>5</b>
<b>2</b>	<b>Historie</b>	<b>6</b>
<b>3</b>	<b>Fourierovy řady</b>	<b>7</b>
3.1	Konvergence . . . . .	10
3.1.1	Gibbsův jev . . . . .	15
3.2	Stejněměrná konvergence . . . . .	17
<b>4</b>	<b>Fourierův integrál</b>	<b>20</b>
4.1	Fourierův integrál funkce . . . . .	20
4.2	Věta o reprezentaci funkce Fourierovým integrálem . . . . .	21
4.3	Komplexní tvar Fourierova integrálu a Fourierova transformace . .	22
4.4	Aplikace a využití Fourierovy transformace . . . . .	24
<b>5</b>	<b>Diskrétní fourierova transformace</b>	<b>25</b>
5.1	Samplování . . . . .	25
5.2	Aliasing . . . . .	26
5.3	Přechod od Fourierovy transformace k diskrétní Fourierově transformaci . . . . .	28
5.4	Rychlá Fourierova transformace . . . . .	30
<b>6</b>	<b>Tvorba a šíření zvuku</b>	<b>32</b>
6.1	Vznik zvuku . . . . .	32
6.1.1	Zvukové spektrum, tóny a šumy . . . . .	32
6.1.2	Formanty . . . . .	34
6.2	Vznik hlasu v lidském těle . . . . .	35
6.3	Zjišťování formantů v současnosti . . . . .	40
<b>7</b>	<b>Praktická část</b>	<b>41</b>

7.1	Úvod do praktické části . . . . .	41
7.2	Software a použité m-fily . . . . .	42
7.3	Výsledné hodnoty . . . . .	46
7.4	Komentář výsledků . . . . .	51
7.4.1	Komentář grafů . . . . .	53
<b>8</b>	<b>Komplikace a zajímavosti při zpracování</b>	<b>54</b>
<b>9</b>	<b>Závěr</b>	<b>54</b>
<b>10</b>	<b>CD příloha</b>	<b>58</b>
<b>11</b>	<b>Příloha m-fily</b>	<b>59</b>
<b>12</b>	<b>Obrazová příloha</b>	<b>63</b>

# 1 Úvod

Teorie Fourierovy transformace se v současnosti hojně využívá pro zpracování signálu, přestože pochází z přelomu 18. a 19. století. Užívá se při zaostřování obrazu, odfiltrování šumu apod. Pro nás je důležitá její aplikace v teroii lidské řeči, resp. analýze samohlásek. Jde o diskutované a žádané téma, především v oblasti lékařství, neboť pokud by se podařilo vytvořit program, který by dokázal rozeznat jednotlivé samohlásky z nahrávky na mikrofon, mohl by být využit pro osoby, které se učí znovu mluvit (např. po mozkové obrně a mozkové mrtvici). Sestavený program by těmto osobám ukázal, jak mají změnit tvar mluvidel tak, aby jejich “a” znělo opravdu jako “a”, protože člověk, který se učí znovu mluvit, nerozezná na základě svého sluchu, jestli samohlásku říká správně. Tímto problémem se zabývají odborníci po celém světě. V současnosti jsou vytvořeny softwary jako Praat a Multi Speech, které slouží k nahrávání, analýze a měření charakteristik lidské řeči, ale ne ke korekci mluvící osoby a jejímu vyslovování samohlásek. Sestavení takového programu (spíše softwaru) je samozřejmě nad rámec této práce – chci se věnovat hledání formantů samohlásek pomocí Fourierovy transformace, přičemž bych chtěla zpřesnit oblasti jednotlivých formantů a potvrdit, že teorie formantů je správná a především také vyřešit problém “maskování” formantů za nosnou frekvenci hlásky. Je tedy každá samohláska určena svými formanty? Známe rozmezí frekvence, kde se formanty dané samohlásky nacházejí? Jsou formanty rozdílné pro muže a ženy?

Cílem diplomové práce je pochopit a popsat matematické metody potřebné k analýze digitálních zvukových signálů, zejména teorii Fourierových řad a Fourierovu transformaci a využít tuto teorii na analýzu českých samohlásek, zejména se zřetelem na identifikaci formantu v digitálním záznamu hlasu. Dílčím cílem práce je vytvořit databázi českých vokálů a vytvořit tak praktický základ pro další zpracování tohoto tématu.

Na úvod se seznámíme s trochou historie Fourierových řad a podíváme se na jejich vyjádření a podmínky, za kterých jsou konvergentní. V kapitole o Fourierových řadách je zařazen také Gibbsův jev, protože s konvergencí Fourierových řad velmi úzce souvisí a při aplikaci Fourierových řad i Fourierovy transformace jako takové se s ním můžeme setkat.

Pro odvození Fourierovy transformace je třeba pochopit a popsat Fourierův integrál, kterému se věnujeme v kapitole následující a přejdeme tak plynule k vlastní Fourierově transformaci. Abychom mohli aplikovat Fourierovu transformaci v praxi, je třeba se podívat na podstatu diskrétní Fourierovy transformace a její rychlou verzi. Tím se dostáváme k samotné aplikaci teorie v praxi.

V další kapitole si také objasníme pojmy jako jsou samplování a aliasing, které bychom neměli v souvislosti s digitálním zpracováním zvuku vynechat.

Kapitola 6 je věnována zvuku, a to především jeho vzniku a šíření. Zjistíme zde, jak vlastně vzniká hlas v lidském těle a co rozumíme pod pojmem formanty. V praktické části je uveden konkrétní postup, kterým jsem získala data ke zpracování, a na jehož základě jsem k výsledkům této práce dospěla.

## 2 Historie

Kapitola o historii je s malými změnami převzata od Herrmanna z [6]. Francouzský matematik a inženýr Jean-Baptiste Joseph Fourier (1768–1830) oznámil 21. 12. 1807 v Paříži na zasedání francouzské Akademie, že libovolnou (grafem) zadanou funkci lze reprezentovat trigonometrickou řadou. Toto tvrzení přijali Akademici, mezi nimi i velký analytik Lagrange, se značnou nedůvěrou a dokonce s averzí. Otázkami rozvinutelnosti funkcí do trigonometrických řad se v té době Fourier zabýval v rámci studia šíření tepla. Tato teorie, kterou rozvíjel, byla známa pod názvem Fourier Series Analysis a později zobecněna na Fourier Transform. Akademici chtěli, aby svou teorii upřesnil, a proto šíření tepla bylo vyhlášeno jako soutěžní téma Grand prix de mathématiques na rok 1812. Fourier předložil své výsledky ke konci roku 1811, soutěž vyhrál, avšak posuzovatelé (mezi nimi Laplace a Lagrange) se vyslovili kriticky k přesnosti jeho metod a nedovolili uveřejnit jeho práci v Mémoires de l'Académie des Sciences.



Obrázek 1: Jean-Baptiste Joseph Fourier (1768–1830).

Fourier to pocítoval jako křivdu. Když publikoval svůj spis “Théorie analytique de la chaleur” (Analytická teorie tepla) v r. 1822, zahrnul do něj prakticky

beze změny první část svých tehdejších výsledků. A když se dva roky nato stal sekretářem Akademie, a tím nejvlivnějším vědcem ve Francii, postaral se o to, aby jeho původní práce byla publikována v Mémoires tak, jak byla předložena v r. 1811.

Prioritu si zabezpečil, jeho analýzy však nebyly, a vlastně ani nemohly být, úplně přesné, neboť nebylo vyjasněno, co se rozumí pojmem funkce, součet řady atp. Pro ilustraci uveďme, jak byla v té době chápána funkce: buď jako analytický výraz nebo jako křivka, kterou lze opsat volně vedenou rukou: “*curva quaecunque libero manus ductu descripta*” (Euler). Přesto když Fourier mluvil o možnosti rozvinout do trigonometrické řady “libovolnou” funkci, byl blíže pravdě než jeho kritici, kteří se snažili omezit platnost jeho rozvoje pouze na třídu funkcí analytických (a v důsledku tedy nekonečně diferencovatelných).

Ospravedlněním formalismu, který Fourier zavedl, se zabývalo mnoho matematiků 19. století. Nejdříve bylo třeba vnést pořádek do základních pojmů. Můžeme bez nadsázky říci, že Fourierova práce sice způsobila dočasnou krizi matematiky, ale v zápětí podnítila bouřlivý rozvoj matematických disciplín. Osvětlit podstatu nekonečných rozvoje mohl teprve přesně definovaný pojem limity, který byl rozvinut okolo roku 1825 Gaussem, Cauchym a Abelem. U Dirichleta se při studiu trigonometrických řad objevuje poprvé (1829) moderní pojem funkce jako “přiřazení”. Jako první stanovil podmínky (dodnes zvané Dirichletovy) na funkci, které zaručují konvergenci Fourierovy řady. Jsou to omezenost a existence jen konečného počtu bodů nespojitosti a maxim a minim na intervalu délky periody. Nejjednodušším příkladem funkcí, splňujících Dirichletovy podmínky, jsou funkce po částech spojitě a monotónně na intervalu délky periody.

Při studiu Dirichletovy teorie se setkal Riemann s požadavkem integrovatelnosti funkce. To ho vedlo k vytvoření teorie integrálu – Riemannova integrálu – prvnímu matematicky propracovanému pojmu integrál. A teprve tehdy se staly Fourierovy formule zcela jasnými. Je příznačné, že teorie integrálu byla součástí jeho habilitační práce (z r. 1854, ale publikované až v r. 1868, dva roky po Riemannově smrti) s názvem “Über die Darstellbarkeit einer Funktion durch eine trigonometrische Reihe” (O reprezentovatelnosti funkce trigonometrickou řadou). Riemannovi také vděčíme za zavedení Fourierovy řady. Následující dvě kapitoly čerpají teorii opět od Herrmanna z [6].

### 3 Fourierovy řady

**Definice 1.** Nechtě  $L > 0$ . Funkční řada tvaru

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} \left( a_k \cos \frac{k\pi x}{L} + b_k \sin \frac{k\pi x}{L} \right), \quad (1)$$

kde  $a_k, b_k \in \mathbb{R}$ , se nazývá *trigonometrická řada*. Její částečné součty

$$s_0(x) = \frac{a_0}{2}, \quad s_n(x) = \frac{a_0}{2} + \sum_{k=1}^n \left( a_k \cos \frac{k\pi x}{L} + b_k \sin \frac{k\pi x}{L} \right), \quad n \in \mathbb{N}$$

se nazývají *trigonometrické polynomy* (stupně  $n$ ).

Uvažujme možnost reprezentovat funkci  $f$  ve tvaru součtu trigonometrické řady (1), tj.

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left( a_k \cos \frac{k\pi x}{L} + b_k \sin \frac{k\pi x}{L} \right),$$

kde koeficienty  $a_0, a_1, \dots, b_1, b_2, \dots$  je třeba určit.

**Poznámka 1.** Pro stručnost zavedeme symbol  $P(2L)$ , kterým označíme množinu všech funkcí z  $\mathbb{R}$  do  $\mathbb{R}$ , které jsou:

1. periodické s periodou  $2L$
2. (riemanovsky) integrovatelné na intervalu  $[-L, L]$ .

**Poznámka 2.** V celém textu  $\mathbb{N}_0$  označuje všechna přirozená čísla  $\mathbb{N}$  včetně 0.

Nyní uvedeme vztahy, které budeme pro další postup potřebovat.

Nechť  $L > 0, k, m \in \mathbb{N}$ . Pak  $\cos \frac{k\pi x}{L}, \sin \frac{k\pi x}{L} \in P(2L)$  a platí:

$$\int_{-L}^L dx = 2L; \tag{2}$$

$$\int_{-L}^L \cos \frac{k\pi x}{L} dx = \int_{-L}^L \sin \frac{k\pi x}{L} dx = 0; \tag{3}$$

$$\int_{-L}^L \cos \frac{k\pi x}{L} \sin \frac{k\pi x}{L} dx = 0; \tag{4}$$

$$\int_{-L}^L \cos \frac{k\pi x}{L} \cos \frac{m\pi x}{L} dx = \begin{cases} 0, & \text{je-li } k \neq m, \\ L, & \text{je-li } k = m; \end{cases} \tag{5}$$

$$\int_{-L}^L \sin \frac{k\pi x}{L} \sin \frac{m\pi x}{L} dx = \begin{cases} 0, & \text{je-li } k \neq m, \\ L, & \text{je-li } k = m. \end{cases} \tag{6}$$

Tyto výsledky lze formulovat tak, že trigonometrický systém funkcí

$$\left\{ 1, \cos \frac{k\pi x}{L}, \sin \frac{k\pi x}{L} \right\}_{k \in \mathbb{N}}$$

je ortogonální na intervalu  $[-L, L]$ .

Nyní předpokládáme, že máme danou funkci  $f$ , která je  $2L$ -periodická, a dodáme předpoklad, že je integrovatelná na intervalu délky periody, tj.  $f \in P(2L)$ . V tom případě jsou  $2L$ -periodické a integrovatelné i funkce tvaru

$$f(x) \cos \frac{m\pi x}{L}, \quad f(x) \sin \frac{m\pi x}{L}$$

pro všechna  $m \in \mathbb{N}$ . Koeficienty  $a_k, b_k$  najdeme za předpokladu, že jsou splněny následující podmínky:

- řada konverguje k  $f(x)$  tj. platí

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left( a_k \cos \frac{k\pi x}{L} + b_k \sin \frac{k\pi x}{L} \right), \quad (7)$$

pro všechna  $x \in \mathbb{R}$

- řadu lze po vynásobení  $\cos \frac{m\pi x}{L}$ , kde postupně klademe  $m \in \mathbb{N}_0$ , a  $\sin \frac{m\pi x}{L}$ ,  $m \in \mathbb{N}$ , integrovat přes interval délky periody člen po členu.

Postupným násobením vztahu (7) funkcemi  $\cos \frac{m\pi x}{L}$ ,  $m \in \mathbb{N}_0$  a integrací přes interval  $[-L, L]$  dostáváme

$$\begin{aligned} \int_{-L}^L f(x) \cos \frac{m\pi x}{L} dx &= \frac{a_0}{2} \int_{-L}^L \cos \frac{m\pi x}{L} dx + \\ &+ \sum_{k=1}^{\infty} a_k \int_{-L}^L \cos \frac{k\pi x}{L} \cos \frac{m\pi x}{L} dx + \sum_{k=1}^{\infty} b_k \int_{-L}^L \sin \frac{k\pi x}{L} \cos \frac{m\pi x}{L} dx. \end{aligned}$$

Ze vztahů (4), (5) a (6) plyne, že jediný nenulový člen vpravo je člen v první řadě s  $k = m$ , a tedy

$$\int_{-L}^L f(x) \cos \frac{m\pi x}{L} dx = La_m.$$

Pro  $m = 0$  dostaneme

$$\int_{-L}^L f(x) dx = \frac{a_0}{2} \int_{-L}^L dx = La_0.$$

Zcela analogicky postupným násobením funkcemi  $\sin \frac{m\pi x}{L}$ ,  $m \in \mathbb{N}$  a následnou integrací přes interval  $[-L, L]$ , obdržíme

$$\int_{-L}^L f(x) \sin \frac{m\pi x}{L} dx = Lb_m.$$

Čísla  $a_0, a_1, \dots, b_1, b_2, \dots$  určená těmito vztahy se nazývají *Fourierovy koeficienty funkce  $f$* .

**Definice 2.** Necht  $f \in P(2L)$ . Potom řada (1), v níž

$$a_k = \frac{1}{L} \int_{-L}^L f(x) \cos \frac{k\pi x}{L} dx, \quad k \in \mathbb{N}_0, \quad (8)$$

$$b_k = \frac{1}{L} \int_{-L}^L f(x) \sin \frac{k\pi x}{L} dx, \quad k \in \mathbb{N}, \quad (9)$$

se nazývá *Fourierova (trigonometrická) řada funkce  $f$* .

Skutečnost, že řada (1) je Fourierovou řadou funkce  $f$ , zapisujeme symbolicky

$$f(x) \approx \frac{a_0}{2} + \sum_{k=1}^{\infty} \left( a_k \cos \frac{k\pi x}{L} + b_k \sin \frac{k\pi x}{L} \right).$$

Zajímá nás tedy, kdy je tato řada konvergentní a za jakých podmínek je její součet v  $x$  roven  $f(x)$ .

### 3.1 Konvergence

Konvergence Fourierovy řady je otázkou existence limity posloupnosti částečných součtů

$$s_0(x) = \frac{a_0}{2}, \quad s_n(x) = \sum_{k=1}^n \left( a_k \cos \frac{k\pi x}{L} + b_k \sin \frac{k\pi x}{L} \right), \quad n \in \mathbb{N}, \quad (10)$$

kde  $a_k$  a  $b_k$  jsou definovány předpisy (8) a (9). Předpoklady zaručující konvergenci mohou být kladeny na koeficienty  $a_k$  a  $b_k$  nebo na samotnou funkci  $f$  a právě těmi se budeme zabývat. Fourierovu řadu můžeme utvořit pro libovolnou funkci  $f \in P(2L)$ . Jestliže se hodnoty dvou funkcí liší pouze v konečně mnoha bodech intervalu  $[-L, L]$ , integrály (8) a (9), které definují Fourierovy koeficienty, jsou stejné, to znamená, že obě funkce mají stejnou Fourierovu řadu. Nemůžeme tedy očekávat, že Fourierova řada libovolné funkce (tj. z  $P(2L)$ ) bude konvergovat k  $f(x)$  pro libovolné  $x$ . Ukazuje se, že bohužel ani spojitost nestačí.

Problematika konvergence Fourierových řad je velmi složitá, což ukazuje následující nepatrná část známých výsledků z minulosti. Od A.N. Kolmogorova z r. 1926 (viz [7]) existuje příklad funkce lebesgueovsky integrovatelné, jejíž Fourierova řada diverguje všude. Již v r. 1871 byla známa (Du Bois Reymond) spojitá funkce, jejíž Fourierova řada diverguje na množině, která je hustá v intervalu  $[-L, L]$ . Do r. 1966 nebylo známo, existuje-li pro spojitou funkci neprázdná množina, na níž Fourierova řada konverguje. Kladná odpověď vyplynula teprve z práce L. Carlesona (1966), který dokázal, že každá kvadraticky lebesgueovsky integrovatelná funkce  $f$  má Fourierovu řadu, která konverguje k  $f$  všude až na množinu lebesgueovy míry nula. Základní výsledek o konvergenci vyslovíme ve Větě 1. Následující Lemma 1 a Poznámky 3 a 4 budeme potřebovat pro důkaz této věty.



**Lemma 1.** Je-li periodická funkce integrovatelná na jednom intervalu délky periody, je integrovatelná na každém intervalu délky periody a všechny integrály na intervalech délky periody se rovnají, tj.

$$f \in P(2L) \implies \int_{-L}^L f(x) dx = \int_{\xi-L}^{\xi+L} f(x) dx$$

pro libovolné  $\xi \in \mathbb{R}$ .

Pro další větu budeme potřebovat definovat pojem po částech hladká funkce.

**Definice 3.** Funkce  $f$  se nazývá *po částech hladká na intervalu*  $[a, b]$ , jestliže má (až na konečný počet bodů) derivaci, která je po částech spojitá na  $[a, b]$ .

Přesněji řečeno, funkce  $f$  je po částech hladká, jestliže existuje konečně mnoho bodů

$$a = x_0 < x_1 < \dots < x_q = b$$

takových, že

1.  $f$  má spojitou derivaci na  $(x_{i-1}, x_i), i = 1, \dots, q$ ,
2. existují jednostranné vlastní limity  $f(x_i+), f'(x_i+)$  pro  $i = 0, \dots, q - 1$ , a  $f(x_i-), f'(x_i-)$  pro  $i = 1, \dots, q$ .

**Poznámka 3.**<sup>1</sup> Je-li funkce  $f$  po částech spojitá na intervalu  $[a, b]$ , pak

$$\lim_{\lambda \rightarrow \infty} \int_a^b f(x) \sin \lambda x dx = 0.$$

*Důkaz.* Dokážeme, že  $\int_{\mathbb{R}} f(x) \sin \lambda x$  konverguje k nule pro  $\lambda \rightarrow \infty$ . Toto platí dokonce pro libovolnou  $f \in L^1(\mathbb{R})$ . Platí to, i když sinus nahradíme kosinem.

1. Protože

$$\sin \lambda x = \frac{1}{2i}(e^{i\lambda x} - e^{-i\lambda x})$$

a

$$\cos \lambda x = \frac{1}{2}(e^{i\lambda x} + e^{-i\lambda x})$$

zřejmě stačí dokázat, že

$$\lim_{\lambda \rightarrow \infty} \int_{\mathbb{R}} f(x) e^{i\lambda x} = 0.$$

2. Volme  $f$  charakteristickou funkcí omezeného intervalu  $[a, b]$ . Potom

$$\int_{\mathbb{R}} f(x) e^{i\lambda x} = \int_a^b e^{i\lambda x} = \left[ \frac{e^{i\lambda x}}{i\lambda} \right]_a^b = \frac{1}{i\lambda} (e^{i\lambda b} - e^{i\lambda a}).$$

---

<sup>1</sup>Toto tvrzení je v literatuře označováno jako Riemann–Lebesgueovo lemma.

Tedy

$$\left| \int_{\mathbb{R}} f(x)e^{i\lambda x} \right| \leq \frac{1}{|\lambda|} (|e^{i\lambda b}| + |e^{i\lambda a}|) \leq \frac{2}{|\lambda|},$$

zřejmě konverguje k nule pro  $\lambda \rightarrow \infty$ .

3. Tvrzení tedy platí i pro jakoukoli lineární kombinaci takových charakteristických funkcí, neboli platí pro všechny jednoduché měřitelné nezáporné funkce (j.n.m.) s kompaktním nosičem.

4. Každá měřitelná nezáporná funkce s kompaktním nosičem se dá napsat jako limita j.n.m. s kompaktním nosičem, tedy volme  $\varepsilon > 0$  libovolné, najdeme  $n \in \mathbb{N}$  takové, že  $\|f - s_n\|_1 < \varepsilon$ , kde funkce  $s_n$  jsou j.n.m. s kompaktním nosičem. Potom máme

$$\int_{\mathbb{R}} f(x)e^{i\lambda x} = \int_{\mathbb{R}} (f(x) - s_n(x))e^{i\lambda x} + \int_{\mathbb{R}} s_n(x)e^{i\lambda x}$$

a tedy

$$\left| \int_{\mathbb{R}} f(x)e^{i\lambda x} \right| \leq \int_{\mathbb{R}} |f(x) - s_n(x)| + \left| \int_{\mathbb{R}} s_n(x)e^{i\lambda x} \right|.$$

Proto

$$\lim_{\lambda \rightarrow \infty} \left| \int_{\mathbb{R}} f(x)e^{i\lambda x} \right| \leq \|f - s_n\|_1 + \lim_{\lambda \rightarrow \infty} \left| \int_{\mathbb{R}} s_n(x)e^{i\lambda x} \right|.$$

Tedy

$$\lim_{\lambda \rightarrow \infty} \left| \int_{\mathbb{R}} f(x)e^{i\lambda x} \right| \leq \varepsilon + 0$$

pro  $\varepsilon$  libovolné, neboli

$$\lim_{\lambda \rightarrow \infty} \left| \int_{\mathbb{R}} f(x)e^{i\lambda x} \right| = 0.$$

5. Pokud  $f$  není nezáporná, napíšeme

$$f = f^+ - f^-$$

a dostáváme

$$\lim_{\lambda \rightarrow \infty} \left| \int_{\mathbb{R}} f(x)e^{i\lambda x} \right| = \lim_{\lambda \rightarrow \infty} \left| \int_{\mathbb{R}} f^+(x)e^{i\lambda x} \right| - \lim_{\lambda \rightarrow \infty} \left| \int_{\mathbb{R}} f^-(x)e^{i\lambda x} \right| = 0.$$

6. Nyní volme libovolnou  $f \in L^1(\mathbb{R})$ . Definujme

$$f_n(x) = \text{restrikce } f(x) \text{ na interval } [-n, n].$$

Pro každé  $n$  máme podle bodu 4

$$\lim_{\lambda \rightarrow \infty} \left| \int_{\mathbb{R}} f_n(x)e^{i\lambda x} \right| = 0.$$

Opět napíšeme

$$\left| \int_{\mathbb{R}} f(x) e^{i\lambda x} \right| \leq \int_{\mathbb{R}} |f_n(x) - f(x)| + \left| \int_{\mathbb{R}} f_n(x) e^{i\lambda x} \right|$$

a stejným postupem jako v bodu 4 dostaneme výsledek.  $\square$

**Poznámka 4.** Je-li po částech hladká funkce  $f$  v bodě  $x_i$  spojitá zprava (popř. ji dodefinujeme hodnotou  $f(x_i+)$ ), existuje derivace zprava v bodě  $x_i$  a platí  $f'_+(x_i) = f'(x_i+)$ , tj.

$$\lim_{x \rightarrow x_i+} \frac{f(x) - f(x_i+)}{x - x_i} = \lim_{x \rightarrow x_i+} f'(x).$$

Podobně, je-li  $f$  spojitá zleva v bodě  $x_i$ , popř. ji nahradíme hodnotou  $f(x_i-)$ , existuje derivace zleva v bodě  $x_i$  a platí  $f'_-(x_i) = f'(x_i-)$ . Důkaz plyne z Lagrangeovy věty o střední hodnotě diferenciálního počtu za využití jednostranné spojitosti.

**Věta 1.** Nechť  $f \in P(2L)$  je po částech hladká na  $[-L, L]$ . Potom Fourierova řada funkce  $f$  daná vztahem (1) konverguje v každém bodě  $x \in \mathbb{R}$  k aritmetickému průměru limit zprava a zleva, tj. platí

$$\begin{aligned} & \frac{a_0}{2} + \sum_{k=1}^{\infty} \left( a_k \cos \frac{k\pi x}{L} + b_k \sin \frac{k\pi x}{L} \right) = \\ & = \begin{cases} f(x), & \text{je-li } f \text{ spojitá v } x, \\ \frac{1}{2}[f(x+) + f(x-)], & \text{není-li } f \text{ spojitá v } x, \end{cases} \end{aligned}$$

kde  $a_k, b_k$  jsou definovány vztahy (8) a (9).

*Důkaz.* Východiskem důkazů vět o konvergenci je vyjádření částečných součtů  $s_n$  v integrálním tvaru

$$s_n(x) = \frac{1}{L} \int_{-L}^L f(x-t) D_n \left( \frac{\pi t}{L} \right) dt$$

nebo

$$s_n(x) = \frac{1}{L} \int_0^L [f(x+t) + f(x-t)] D_n \left( \frac{\pi t}{L} \right) dt,$$

kde

$$D_n(\theta) = \frac{1}{2} + \cos \theta + \cos 2\theta + \dots + \cos n\theta.$$

Podoba těchto vztahů je odvozena v tomto důkazu. Využijeme výsledků Poznámky 3 a 4. V integrálním vyjádření částečných součtů se objevuje tzv. Dirichletovo jádro  $D_n$  a začneme odvozením jeho vlastností. Označíme-li

$$D_0(\theta) = \frac{1}{2}, \quad D_n(\theta) = \frac{1}{2} + \cos \theta + \cos 2\theta + \dots + \cos n\theta \quad (= D_{n-1}(\theta) + \cos n\theta),$$

pro každé  $n \in \mathbb{N}$  a  $\theta \in \mathbb{R}$ , platí

$$D_n(\theta) = \begin{cases} n + \frac{1}{2}, & \text{je-li } \theta = 2m\pi, m \in \mathbb{Z}, \\ \frac{\sin[(n+\frac{1}{2})\theta]}{2\sin\frac{\theta}{2}} & \text{je-li } \theta \neq 2m\pi, m \in \mathbb{Z}. \end{cases} \quad (11)$$

Pro každé  $n \in \mathbb{N}_0$  je funkce  $D_n$  spojitá, sudá,  $2\pi$ -periodická a

$$\frac{2}{\pi} \int_0^\pi D_n(\theta) d\theta = 1.$$

Uvedenou trigonometrickou identitu (11) pro  $\theta \neq 2\pi m$  dostaneme  $n$ -násobným použitím vzorce

$$2 \sin \gamma \cos \delta = \sin(\delta + \gamma) - \sin(\delta - \gamma), \quad \gamma, \delta \in \mathbb{R},$$

neboť

$$2 \sin \frac{\theta}{2} \left( \frac{1}{2} + \cos \theta + \cos 2\theta + \dots + \cos n\theta \right) = \sin \left[ \left( n + \frac{1}{2} \right) \theta \right].$$

Uvažujme nyní libovolnou funkci  $f \in P(2L)$ . Dosazením vztahů (8) a (9) do

$$s_n(x) = \frac{a_0}{2} + \sum_{k=1}^n \left( a_k \cos \frac{k\pi x}{L} + b_k \sin \frac{k\pi x}{L} \right),$$

dostáváme pro každé  $n \in \mathbb{N}_0$  při využití vztahu (11)

$$\begin{aligned} s_n(x) &= \frac{1}{L} \int_{-L}^L f(\xi) \left[ \frac{1}{2} + \sum_{k=1}^n \left( \cos \frac{k\pi\xi}{L} \cos \frac{k\pi x}{L} + \sin \frac{k\pi\xi}{L} \sin \frac{k\pi x}{L} \right) \right] d\xi = \\ &= \frac{1}{L} \int_{-L}^L f(\xi) \left[ \frac{1}{2} + \sum_{k=1}^n \cos \frac{k\pi(x-\xi)}{L} \right] d\xi = \frac{1}{L} \int_{-L}^L f(\xi) D_n \left( \frac{\pi(x-\xi)}{L} \right) d\xi. \end{aligned}$$

Použitím  $2L$ -periodicity integrandu lze podle Lemmatu 1 psát

$$s_n(x) = \frac{1}{L} \int_{-L}^L f(x-t) D_n \left( \frac{\pi t}{L} \right) dt,$$

kde nová proměnná  $t = x - \xi$  nebo alternativně (využitím sudosti funkce  $D_n$ )

$$s_n(x) = \frac{1}{L} \int_0^L [f(x+t) + f(x-t)] D_n \left( \frac{\pi t}{L} \right) dt.$$

Nyní, protože platí

$$\frac{2}{L} \int_0^L D_n \left( \frac{\pi t}{L} \right) dt = 1,$$

pak

$$\begin{aligned} s_n(x) - \frac{f(x+) + f(x-)}{2} &= \\ &= \frac{1}{L} \int_0^L [f(x+t) - f(x+) + f(x-t) - f(x-)] D_n \left( \frac{\pi t}{L} \right) dt = \\ &= \frac{1}{\pi} \int_0^L g(t) \sin \left[ \left( n + \frac{1}{2} \right) \frac{\pi t}{L} \right] dt, \end{aligned}$$

kde

$$g(t) = \left[ \frac{f(x+t) - f(x+)}{t} + \frac{f(x-t) - f(x-)}{t} \right] \frac{\frac{\pi t}{2L}}{\sin \frac{\pi t}{2L}}.$$

Užitím výsledku Poznámky 4 je

$$\lim_{t \rightarrow 0} \frac{f(x+t) - f(x+)}{t} = f'(x+), \quad \lim_{t \rightarrow 0} \frac{f(x-t) - f(x-)}{t} = f'(x-),$$

a poněvadž

$$\lim_{t \rightarrow 0} \frac{\frac{\pi t}{2L}}{\sin \frac{\pi t}{2L}} = 1,$$

má funkce  $g$  v 0 vlastní limitu zprava, je tedy po částech spojitá na  $[0, L]$  a podle výsledku Poznámky 3 je tímto věta o konvergenci dokázána.  $\square$

### 3.1.1 Gibbsův jev

“Achillovou patou” aproximace Fourierovou řadou je tzv. Gibbsův jev. Představme si jednoduchou funkci s nespojitostí, např. funkci  $f(x) = x$  pro  $x \in (-\pi, \pi)$ . Nás zajímá, co se bude dít v bodech nespojitosti. V okolí těchto bodů, tj. kde  $f(x+) \neq f(x-)$ , dochází k “překmitu”. Každý částečný součet řady má maximum, které převyšuje přibližně o 9% velikost skoku. Geometricky to znamená, že grafy částečných součtů se blíží při  $n \rightarrow \infty$  a přibližování se bodu nespojitosti k vertikální úsečce nikoli s krajními body  $f(x+)$  a  $f(x-)$ , jak bychom čekali, ale k úsečce podstatně (asi o 9% ) delší. Proč právě 9% ? Podle věty o bodové konvergenci platí: jestliže předpokládáme, že  $f(x)$  je spojitá funkce a spojitě diferencovatelná až na konečný počet bodů, pak pro každé  $a \in (0, 1)$

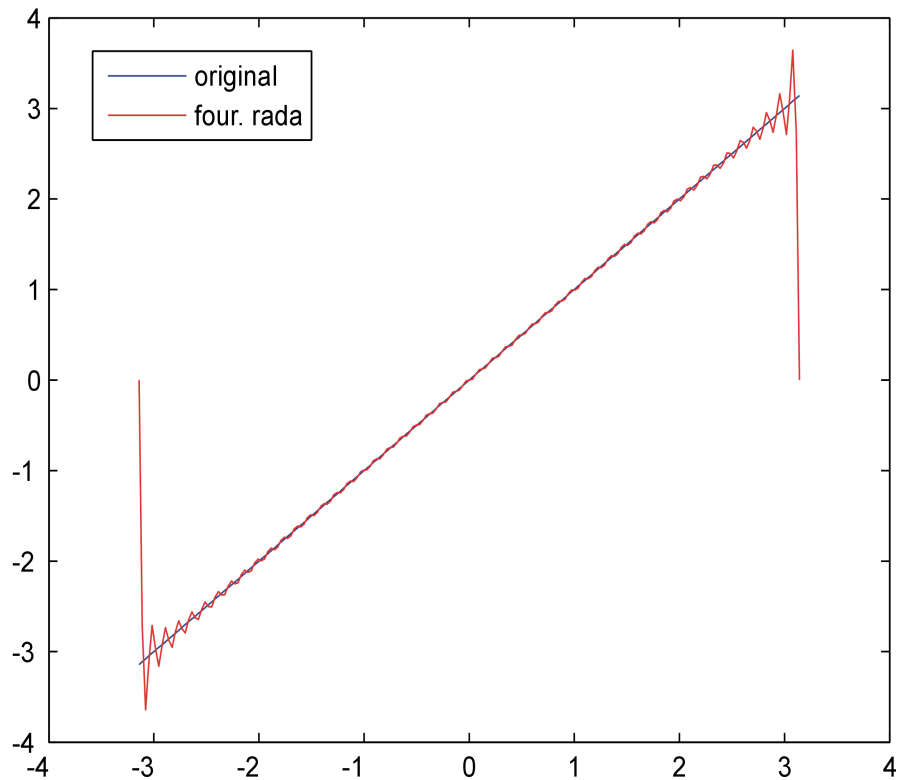
$$S_n(x) \rightarrow \frac{1}{2} \left( \lim_{t \rightarrow a^-} f(x) + \lim_{t \rightarrow a^+} f(x) \right) \quad \text{pro } n \rightarrow \infty$$

a platí, že maximum  $S_n(x)$  je větší než 1 pro  $n \rightarrow \infty$ , přesněji

$$\lim_{n \rightarrow \infty} \max S_n(x) = 1,089490\dots$$

takže přesah nad 1 je téměř 9% .

Na Obrázku 2 můžeme vidět znázornění Gibbsova jevu graficky. Vyjádření jsme získali použitím následujícího m-filu:



Obrázek 2: Grafické zobrazení Gibbsova jevu.

```

h = @(t) t;
N = 44;
tic
koef = zeros(1,2*N+1);
pos = 0;
for k = -N:1:N
    pos = pos+1;
    fun = @(t) h(t).*(exp(-i*k*t));
    c = quadgk(fun, -pi,pi,'MaxIntervalCount',2*N+10);
    koef(pos) = c/(2*pi);
end

xx = -pi:pi/100:pi;
rad = 0*xx;
pos = 0;
for k = -N:1:N
    pos = pos+1;
    c = koef(pos)*exp(i*k*xx);
    rad = rad + c;
end

```

Rychlé změny funkčních hodnot se dějí v rámci velice úzké oblasti a tato oblast se zmenšuje stále více jak roste  $n$ . Částečné součty proto musí v tomto bodě nespojitosti “spěchat” stále více, tím se dostanou výrazně nad danou funkční hodnotu a teprve pak se usadí. Toto chování se stane vždy, když se Fourierova řada musí vypořádat s nespojitostí. Na obou jejích stranách se u částečných součtů objevují oscilace, které jsou na stále užším prostoru, ale jejich velikost se nezmenšuje.

U Fourierových řad má tato nestejněměrná konvergence v okolí bodu nespojitosti zvláštní charakter, který byl poprvé ke konci 18. století podrobněji vyšetřován americkým matematikem a fyzikem J. W. Gibbsem (1839–1903), a proto se v literatuře traduje pod názvem Gibbsův jev.

### 3.2 Stejněměrná konvergence

**Věta 2.** Nechť  $f \in P(2L)$  je po částech hladká na  $[-L, L]$  a spojitá na  $\mathbb{R}$ . Potom řada

$$\left| \frac{a_0}{2} \right| + \sum_{k=1}^{\infty} (|a_k| + |b_k|)$$

je konvergentní a v důsledku toho:

- Fourierova řada (1) funkce  $f$  konverguje absolutně a stejněměrně na  $\mathbb{R}$
- její součet je roven  $f(x)$  pro všechna  $x \in \mathbb{R}$ .

Nyní odvodíme vztahy, které budeme potřebovat pro důkaz Věty 2. Uvažujme funkci  $f$ , která splňuje předpoklady Věty 1 o konvergenci. To znamená, že je  $2L$ -periodická a existuje konečně mnoho bodů  $-L = x_0 < x_1 < \dots < x_q = L$  takových, že  $f$  má spojitou derivaci na  $(x_{i-1}, x_i)$ ,  $i = 1, \dots, q$ , existují jednostranné vlastní limity  $f(x_i+)$ ,  $f'(x_i+)$  pro  $i = 0, \dots, q-1$ , a  $f(x_i-)$ ,  $f'(x_i-)$  pro  $i = 0, \dots, q$ . Pro každé  $k \in \mathbb{N}$  lze tedy psát

$$\begin{aligned} La_k &= \int_{-L}^L f(x) \cos \frac{k\pi x}{L} dx = \int_{-L}^{x_1} f(x) \cos \frac{k\pi x}{L} dx + \int_{x_1}^{x_2} f(x) \cos \frac{k\pi x}{L} dx + \\ &\quad + \dots + \int_{x_{q-1}}^L f(x) \cos \frac{k\pi x}{L} dx. \end{aligned}$$

Odtud integrací per partes dostáváme

$$\begin{aligned} a_k &= \frac{1}{k\pi} \left\{ \left[ f(x) \sin \frac{k\pi x}{L} \right]_{-L}^{x_1} + \left[ f(x) \sin \frac{k\pi x}{L} \right]_{x_1}^{x_2} + \dots + \left[ f(x) \sin \frac{k\pi x}{L} \right]_{x_{q-1}}^L \right\} - \\ &\quad - \frac{1}{k\pi} \left\{ \int_{-L}^{x_1} f'(x) \sin \frac{k\pi x}{L} dx + \int_{x_1}^{x_2} f'(x) \sin \frac{k\pi x}{L} dx + \dots \right\} \end{aligned}$$

$$\cdots + \int_{x_{q-1}}^L f'(x) \sin \frac{k\pi x}{L} dx \Big\},$$

a tedy

$$a_k = \frac{1}{k\pi} \sum_{i=1}^{q-1} [f(x_i-) - f(x_i+)] \sin \frac{k\pi x_i}{L} - \frac{L}{k\pi} b'_k.$$

Podobně

$$b_k = -\frac{1}{k\pi} \sum_{i=1}^{q-1} [f(x_i-) - f(x_i+)] \cos \frac{k\pi x_i}{L} - [f(L-) - f(-L+)] \cos k\pi - \frac{L}{k\pi} a'_k.$$

Zde  $a'_k$  a  $b'_k$  znamenají Fourierovy koeficienty funkce  $f'$ , tj.

$$a'_k = \frac{1}{L} \int_{-L}^L f'(x) \cos \frac{k\pi x}{L} dx, \quad b'_k = \frac{1}{L} \int_{-L}^L f'(x) \sin \frac{k\pi x}{L} dx.$$

Poněvadž funkce  $f'$  je omezená, jsou omezené i posloupnosti  $a'_k$  a  $b'_k$ , a tedy existuje konstanta  $C > 0$  taková, že

$$|a_k| + |b_k| \leq \frac{C}{k}, \quad k \in \mathbb{N}.$$

K tomu, aby platilo

$$a_k = -\frac{L}{k\pi} b'_k, \quad b_k = -\frac{L}{k\pi} a'_k, \quad k \in \mathbb{N}, \quad (12)$$

stačí, aby

$$f(-L+) = f(L-), \quad f(x_i+) = f(x_i-), \quad i = 1, \dots, q-1,$$

tj. aby funkce  $f$  byla spojitá na celém  $\mathbb{R}$ . Tohoto předpokladu využijeme v důkazu Věty 2.

Jestliže doplníme předpoklady Věty 1 o požadavek spojitosti  $f$  na celém  $\mathbb{R}$ , ukážeme ve Větě 2, že řada

$$\left| \frac{a_0}{2} \right| + \sum_{k=1}^{\infty} (|a_k| + |b_k|), \quad (13)$$

je konvergentní a stává se konvergentní majorantou řady

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} \left( a_k \cos \frac{k\pi x}{L} + b_k \sin \frac{k\pi x}{L} \right). \quad (14)$$

Tato skutečnost má důsledky, které jsou uvedeny ve Větě 2.



Tedy ke každému  $\epsilon > 0$  existuje  $n_0 \in \mathbb{N}$  tak, že pro všechna  $n > n_0$  a pro všechna  $x \in \mathbb{R}$  platí

$$\left| f(x) - \left\{ \frac{a_0}{2} + \sum_{k=1}^n \left( a_k \cos \frac{k\pi x}{L} + b_k \sin \frac{k\pi x}{L} \right) \right\} \right| < \epsilon.$$

V následujícím důkazu využijeme Weierstrassovo majorantní kritérium:

Nechť

- $\{f_k(x)\}_{k=0}^{\infty}$  je posloupnost funkcí definovaných na množině  $E \subset \mathbb{R}$
- $|f_k(x)| \leq M_k$  pro všechna  $x \in E, k \in \mathbb{N}_0$
- $\sum_{k=0}^{\infty} M_k$  je konvergentní číselná řada (tzv. konvergentní majoranta řady  $\sum_{k=0}^{\infty} f_k(x)$ ).

Pak řada  $\sum_{k=0}^{\infty} f_k(x)$  je stejnoměrně a absolutně konvergentní na  $E$ .

*Důkaz.* Konvergence řady (14) plyne ze vztahů (12), které jsou důsledkem předpokládané spojitosti funkce  $f$  na celém  $\mathbb{R}$  a uplatněním tzv. *Besselovy nerovnosti* viz [6]

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) \leq \frac{1}{L} \int_{-L}^L f^2 dx. \quad (15)$$

Z elementární nerovnosti  $ab \leq \frac{a^2+b^2}{2}$ , která platí pro libovolná  $a, b \in \mathbb{R}$ , a ze vztahů (12) odvodíme

$$|a_k| = \left| \frac{L}{k\pi} b'_k \right| = \left| \frac{L}{2\pi k^2} + \frac{L}{\pi} b_k'^2 \right|$$

a podobně

$$|b_k| = \left| \frac{L}{k\pi} a'_k \right| = \left| \frac{L}{2\pi k^2} + \frac{L}{\pi} a_k'^2 \right|,$$

tedy dostáváme

$$|a_k| \leq \frac{L}{2\pi} \left( |b'_k|^2 + \frac{1}{k^2} \right), \quad |b_k| \leq \frac{L}{2\pi} \left( |a'_k|^2 + \frac{1}{k^2} \right), \quad k \in \mathbb{N}.$$

Použitím Besselovy nerovnosti (15) pro  $f'$  dostaneme, že řada  $\sum_{k=1}^{\infty} (|a'_k|^2 + |b'_k|^2)$  je konvergentní a poněvadž je známo, že  $\sum_{k=1}^{\infty} \frac{1}{k^2}$  je konvergentní, řada (13) je rovněž konvergentní. Podle Weierstrassova majorantního kritéria je řada (14) absolutně a stejnoměrně konvergentní na  $\mathbb{R}$ .

□

Spojitosť funkce  $f$  (nutná podmínka stejnoměrné konvergence řady (14) ) se tedy stala za předpokladů Věty 1 i podmínkou postačující pro stejnoměrnou i absolutní konvergenci.

**Poznámka 5.** Právě absolutní a stejnoměrná konvergence je charakteristickou vlastností mocninných řad, tj. v intervalu konvergence řada konverguje absolutně a v každém omezeném uzavřeném intervalu, který je částí intervalu konvergence, řada konverguje stejnoměrně.

## 4 Fourierův integrál

V předcházející kapitole jsme se zabývali podmínkami, které umožňují rozvinout funkci  $f$  do Fourierovy trigonometrické řady. Nutnou podmínkou je periodičita funkce  $f$ . Pokud funkce periodická není, lze ji za určitých podmínek vyjádřit pomocí Fourierova integrálu, který si můžeme představit jako limitní případ Fourierovy řady funkce, jejíž perioda roste nade všechny meze. Výsledky jsou základem pro vybudování Fourierovy transformace viz [6].

### 4.1 Fourierův integrál funkce

Funkci  $f$ , která je  $2L$ -periodická, spojitá na  $\mathbb{R}$  a po částech hladká na intervalu  $[-L, L]$ , lze podle Věty 1 napsat ve tvaru

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left( a_k \cos \frac{k\pi x}{L} + b_k \sin \frac{k\pi x}{L} \right),$$

kde  $a_k, b_k$  jsou dány vztahy (8) a (9). Použitím součtového vzorce pro funkci  $\cos$  lze ekvivalentně psát

$$f(x) = \frac{1}{2L} \int_{-L}^L f(\xi) d\xi + \frac{1}{L} \sum_{k=1}^{\infty} f(\xi) \cos \frac{k\pi(x - \xi)}{L} d\xi \quad (16)$$

a tato rovnost platí pro všechna  $x \in \mathbb{R}$ .

Nyní opustíme předpoklad, že  $f$  je periodická funkce. Nechť je  $f$  spojitá na  $\mathbb{R}$  a po částech hladká na intervalu  $[-L, L]$  pro všechna  $L > 0$ . Dá se očekávat, že provedením limitního přechodu pro  $L \rightarrow +\infty$  ve vztahu (16) dostaneme reprezentaci  $f$  pro všechna  $x \in \mathbb{R}$ .

První člen vpravo vztahu (16) konverguje k nule, a to za dodatečného předpokladu existence nevlastního integrálu

$$\int_{-\infty}^{+\infty} |f(x)| dx, \quad (17)$$

tj. absolutní integrovatelnosti funkce  $f$  na  $\mathbb{R}$ . Za tohoto předpokladu totiž pro každé  $L > 0$  platí

$$\left| \frac{1}{2L} \int_{-L}^L f(x) dx \right| \leq \frac{1}{2L} \int_{-L}^L |f(x)| dx \leq \frac{1}{2L} \int_{-\infty}^{+\infty} |f(x)| dx.$$

Ve druhém členu funkce  $f$  (tj. v řadě vztahu (16)) vystupují hodnoty funkce  $\cos[v_k(x - \xi)]$  pro hodnoty  $v_k = \frac{k\pi}{L}$ ,  $k \in \mathbb{N}$ , pro něž je

$$\Delta v_k = v_{k+1} - v_k = \frac{\pi}{L}, \quad \Delta v_k \rightarrow 0$$

pro  $L \rightarrow +\infty$ . Řada tedy strukturou připomíná integrální součet pro funkci

$$g(v) = \frac{1}{\pi} \int_{-L}^L f(\xi) \cos[v(x - \xi)] d\xi.$$

Očekáváme, že pro  $L \rightarrow +\infty$  bude mít pravá strana funkce  $f(x)$  za limitu integrál a dostaneme, pro každé  $x \in \mathbb{R}$ ,

$$f(x) = \frac{1}{\pi} \int_0^{+\infty} \left( \int_{-\infty}^{+\infty} f(\xi) \cos[v(x - \xi)] d\xi \right) dv. \quad (18)$$

Výraz vpravo se nazývá *Fourierův integrál funkce  $f$  (v bodě  $x$ )*.

## 4.2 Věta o reprezentaci funkce Fourierovým integrálem

**Věta 3.** Nechť funkce  $f$  je absolutně integrovatelná na  $\mathbb{R}$  (tj. nevlastní integrál (17) je konvergentní) a po částech hladká na každém omezeném intervalu. Potom Fourierův integrál (18) funkce  $f$  konverguje pro každé  $x \in \mathbb{R}$  a platí

$$\frac{1}{2}[f(x+) + f(x-)] = \frac{1}{\pi} \int_0^{+\infty} \left( \int_{-\infty}^{+\infty} f(\xi) \cos[v(x - \xi)] d\xi \right) dv.$$

*Důkaz.* Z absolutní integrovatelnosti funkce  $f$  (použitím majorantního kritéria) dostáváme, že pro každé  $v, x \in \mathbb{R}$  platí

$$\begin{aligned} \left| \int_{-\infty}^{+\infty} f(\xi) \cos[v(x - \xi)] d\xi \right| &\leq \int_{-\infty}^{+\infty} |f(\xi) \cos[v(x - \xi)]| d\xi \leq \\ &\leq \int_{-\infty}^{+\infty} |f(\xi)| d\xi < +\infty. \end{aligned} \quad (19)$$

Označíme-li, pro každé  $N > 0$ ,

$$I_N(x) = \frac{1}{\pi} \int_0^N \left( \int_{-\infty}^{+\infty} f(\xi) \cos[v(x - \xi)] d\xi \right) dv,$$

je díky vlastnosti (19) možno provést záměnu pořadí integrování (oprávněnost záměny je založena na tzv. stejnoměrné konvergenci vnitřního integrálu vzhledem k  $v$  a ta vyplývá z (19)). Tedy

$$\begin{aligned} I_N(x) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} f(\xi) \left( \int_0^N \cos[v(x - \xi)] dv \right) d\xi = \\ &= \frac{1}{\pi} \int_{-\infty}^{+\infty} f(\xi) \frac{\sin[N(x - \xi)]}{x - \xi} d\xi = \\ &= \frac{1}{\pi} \int_0^{+\infty} [f(x + t) + f(x - t)] \frac{\sin Nt}{t} dt, \end{aligned}$$

kde jsme posunuli meze integrálu, tím místo  $f(\xi)$  píšeme  $f(-\xi) + f(+\xi)$  a opět jsme použili substituci  $t = x - \xi$ .

Z konvergentního nevlastního integrálu

$$\int_{-\infty}^{+\infty} \frac{\sin Nx}{x} dx = \pi, \quad (N \neq 0),$$

který není absolutně konvergentní, plyne, že pro každé  $N > 0$  platí

$$\frac{2}{\pi} \int_0^{+\infty} \frac{\sin Nt}{t} dt = 1.$$

Tedy

$$I_N(x) - \frac{f(x+) + f(x-)}{2} = \frac{1}{\pi} \int_0^{+\infty} [f(x+t) - f(x+) + f(x-t) - f(x-)] \frac{\sin Nt}{t} dt.$$

Integrál napíšeme jako součet dvou integrálů  $I_1$  a  $I_2$  na intervalech  $[0, N_0]$ , resp.  $[N_0, +\infty)$ . V důsledku konvergence integrálu (17) konvergují (s použitím majorantního kritéria) i integrály  $\int_a^{+\infty} \frac{|f(t)|}{t} dt$  pro jakákoli  $a > 0$ . Lze tedy volbou dostatečně velkého  $a$  učinit libovolně malým integrál  $I_2$ . Integrál  $I_1$  na omezeném intervalu  $[0, a]$  pak lze učinit libovolně malým volbou dostatečně velkého  $N$  zcela analogicky, jako v důkazu Věty 1.  $\square$

### 4.3 Komplexní tvar Fourierova integrálu a Fourierova transformace

V této kapitole budeme potřebovat pojem *hlavní hodnota integrálu funkce*. Jestliže pro  $a \in \mathbb{R}$  jsou integrály

$$\int_{-\infty}^a f(x) dx, \quad \int_a^{+\infty} f(x) dx$$

konvergentní, pak říkáme, že zobecněný integrál funkce  $f$  od  $-\infty$  do  $+\infty$  je konvergentní a klademe

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^a f(x) dx + \int_a^{+\infty} f(x) dx. \quad (20)$$

Může se stát, že zobecněný integrál (20) neexistuje, ale existuje (vlastní) limita integrálů v mezích od  $-N$  do  $N$  pro  $N \rightarrow +\infty$ . V takovém případě nazýváme tuto limitu *hlavní hodnotou integrálu funkce  $f$*  od  $-\infty$  do  $+\infty$  a značíme ji V.p. (z francouzského valeur principale). Je tedy

$$V.p. \int_{-\infty}^{+\infty} f(x) dx = \lim_{N \rightarrow +\infty} \int_{-N}^N f(x) dx. \quad (21)$$

Nechť  $f$  je absolutně integrovatelná na  $\mathbb{R}$  a na každém omezeném intervalu je po částech hladká. Poněvadž funkce  $\sin[v(x - \xi)]$  je lichá a funkce  $\cos[v(x - \xi)]$  je sudá (jako funkce  $v$  při pevných  $x, \xi$ ), platí pro každé  $N > 0$

$$\int_{-N}^N \left( \int_{-\infty}^{+\infty} f(\xi) \sin[v(x - \xi)] d\xi \right) dv = 0,$$

$$\int_{-N}^N \left( \int_{-\infty}^{+\infty} f(\xi) \cos[v(x - \xi)] d\xi \right) dv = 2 \int_0^N \left( \int_{-\infty}^{+\infty} f(\xi) \cos[v(x - \xi)] d\xi \right) dv.$$

Použitím Eulerovy formule  $e^{i\alpha} = \cos \alpha + i \sin \alpha$  platné pro všechna  $\alpha \in \mathbb{R}$  a Věty 3 dostaneme

$$\begin{aligned} \frac{1}{2}[f(x+) + f(x-)] &= \frac{1}{\pi} \lim_{N \rightarrow +\infty} \int_0^N \left( \int_{-\infty}^{+\infty} f(\xi) \cos[v(x - \xi)] d\xi \right) dv = \\ &= \frac{1}{2\pi} \lim_{N \rightarrow +\infty} \left( \int_{-N}^N \left( \int_{-\infty}^{+\infty} f(\xi) \cos[v(x - \xi)] d\xi \right) dv \right) + \\ &+ \frac{1}{2\pi} \lim_{N \rightarrow +\infty} \left( i \int_{-N}^N \left( \int_{-\infty}^{+\infty} f(\xi) \sin[v(x - \xi)] d\xi \right) dv \right) = \\ &= \frac{1}{2\pi} \lim_{N \rightarrow +\infty} \int_{-N}^N \left( \int_{-\infty}^{+\infty} f(\xi) e^{iv(x-\xi)} d\xi \right) dv. \end{aligned}$$

Použijeme-li označení V.p. pro hlavní hodnotu integrálu, můžeme psát

$$\frac{1}{2}[f(x+) + f(x-)] = \frac{1}{2\pi} V.p. \int_{-\infty}^{+\infty} \left( \int_{-\infty}^{+\infty} f(\xi) e^{iv(x-\xi)} d\xi \right) dv.$$

V.p. nelze obecně vynechat, neboť nevlastní integrál od  $-\infty$  do  $+\infty$  nemusí existovat. Této úpravě Fourierova integrálu se s výhodou užívá v aplikacích a slouží k zavedení tzv. Fourierovy transformace. Označíme-li

$$\hat{f}(v) = \int_{-\infty}^{+\infty} f(\xi) e^{-iv\xi} d\xi, \quad (22)$$

pak

$$\frac{1}{2}[f(x+) + f(x-)] = \frac{1}{2\pi} V.p. \int_{-\infty}^{+\infty} \hat{f}(v) e^{ivx} dv. \quad (23)$$

Funkce definovaná vztahem (22) se nazývá *Fourierova transformace funkce*  $f$  a vztah (23) představuje tzv. *inverzní Fourierovu formuli*. Zobrazení, které vzoru  $f$  přiřazuje obraz  $\hat{f}$ , se nazývá *Fourierova transformace* a značí se  $F$ , tedy  $Ff = \hat{f}$  nebo podrobněji  $F\{f(x)\}(v) = \hat{f}(v)$ . Fourierův obraz  $Ff$  se nazývá *spektrém signálu* (vzoru)  $f$ . Pro třídu funkcí  $f$ , které jsou absolutně integrovatelné a spojitě na  $\mathbb{R}$  a po částech hladké na každém omezeném intervalu platí podle Věty 3 tzv. Věta o jednoznačnosti: k Fourierovu obrazu je vzor v této třídě jediný (a je určen pomocí inverzní Fourierovy formule).

$F$  je prosté zobrazení a zobrazení, které každému obrazu  $\hat{f}$  přiřazuje funkci definovanou vpravo ve vzorci (23) se nazývá *inverzní Fourierova transformace* a značí  $F^{-1}$ . Platí  $F^{-1}\hat{f} = f$  nebo podrobněji  $F^{-1}\{\hat{f}(v)\}(x) = f(x)$  pro všechna  $x \in \mathbb{R}$ . Poněvadž ovšem definiční obor inverzní Fourierovy transformace, tedy obor hodnot Fourierovy transformace, není snadné charakterizovat, budují se teorie Fourierovy transformace i na jiných třídách a prostorech funkcí a zkoumá se prostota tohoto zobrazení (a tedy existence inverzního zobrazení).

**Poznámka 6.** Z důvodů větší symetrie se někdy Fourierova a inverzní Fourierova transformace definují pomocí vzorců

$$Ff = \hat{f}(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(\xi) e^{-iv\xi} d\xi,$$

$$F^{-1}\hat{f} = \frac{1}{\sqrt{2\pi}} V.p. \int_{-\infty}^{+\infty} \hat{f}(v) e^{ivx} dv.$$

Některé vlastnosti Fourierovy transformace jsou potom vyjádřeny vzorci, které nabývají formálně jednoduššího tvaru.

## 4.4 Aplikace a využití Fourierovy transformace

Fourierova transformace (FT) má opravdu široké využití, pro představu některá uvedeme. Fourierova transformace se používá při řešení diferenciálních rovnic. Podstata spočívá v tom, že FT převádí operaci derivování na operaci násobení obrazu nezávisle proměnnou (viz [11]). Provedeme-li Fourierovu transformaci podle všech nezávislých proměnných, dostaneme pro obraz řešení rovnici, v níž se derivování vůbec nevyskytuje. Když ji vyřešíme, stačí pak najít Fourierův vzor, což bývá to nejobtížnější. Může se ale stát, že řešení rovnice pro obraz nemá žádný vzor. Pak tato metoda nevede k cíli, ale provedeme-li Fourierovu transformaci jen podle části proměnných, dostaneme diferenciální rovnici pro funkci méně proměnných s parametry, která může být snadněji řešitelná, než rovnice původní a pak opět musíme najít vzor.

Pomocí Fourierovy transformace lze provádět spektrální filtraci, tj. modifikaci texturálních struktur či zvýraznění obrazu jako takového a zlepšení jeho kontrastu. Frekvenční filtry odstraňují nízké frekvence, čímž dochází k detekci hran, vysoké frekvence, kde poté dochází ke zvýraznění ploch nebo odstraňují střední frekvence a tím dochází k rozostření malých struktur. V oblasti obrazu je pak FT využita pro úpravu kvality obrazu, jeho zaostření nebo komprese.

Další využití FT spočívá ve frekvenční analýze rotačních strojů, tj. pravidelné sledování strojů, jejich vibrací, vyhodnocování vibračních spekter a určování závad. Na základě vyhodnocení vibračních spekter rotačních strojů lze získat přesný přehled o jeho technickém stavu, predikovat pravděpodobné závady a určit zbývající životnost zařízení.

FT lze použít pro identifikaci a eliminaci periodického šumu obsaženého v družicovém snímku viz [18]. V tomto případě je nejprve provedena Fourierova transformace, posléze aplikována maska (šum se zobrazí ve Fourierově spektru ve tvaru pruhů) a data jsou zpětně transformována do prostorové domény (inverzní Fourierova transformace).

Metody Fourierovy transformace využívají také přístroje pracující na principu interference záření. Tyto přístroje se využívají v oblasti infračervené spektroskopie k identifikaci chemické struktury látek.

## 5 Diskrétní fourierova transformace

Fourierova transformace je účinnou matematickou metodou ke zpracování různých signálů a jejich analyzování. Obraz Fourierovy transformace se nazývá spektrem signálu. Definiční vzorec pro Fourierovu transformaci je integrálem a pro praktickou realizaci není příliš vhodný, protože jeho analytické řešení existuje jen v omezeném počtu případů a je nutno jej řešit numericky. Navíc v případě počítačového zpracování nemáme spojitou funkci, ale jen její hodnoty v diskretních vzorkovacích okamžicích. Z těchto důvodů se definuje diskrétní Fourierova transformace (DFT), jejíž vstupy a výstupy jsou posloupnostmi hodnot tedy diskretizovaný signál.

### 5.1 Samplování

Samplováním (neboli vzorkováním, z anglického slova sample = vzorek) rozumíme proces převodu analogového (neboli spojitého) signálu na digitální (neboli diskrétní). Vstupní signál se digitalizuje pomocí A/D převodníku<sup>2</sup> a periodicky se ukládá po stanovených okamžicích do paměti. Počet sejmutých vzorků za jednu vteřinu je vzorkovací frekvence a spotřeba paměti závisí na této frekvenci přímo úměrně.

---

<sup>2</sup>elektronická součástka určená pro převod signálu z analogového na digitální

Proces vzorkování si můžeme ukázat na následujícím příkladu. Představme si posloupnost velmi úzkých stejně vysokých a širokých impulsů. Řekněme, že impulsů bude 44 100 za sekundu. V okamžiku “vystřelení” impulsu budeme z analogového signálu odebírat vzorky. Výsledným vzorkovaným signálem je tedy řada impulsů a jejich výška je rovna okamžité hodnotě původního analogového signálu. Pokud ze signálu odebereme zmíněných 44 100 vzorků za sekundu, znamená to, že vzorkovací frekvence, označovaná někdy jako SR (sample rate), je 44 100 Hz. Dá se říct, že našim analogovým signálem vlastně modulujeme velikost jednotlivých vzorkovacích impulsů. Proto se také proces vzorkování někdy označuje zkratkou PAM (pulsní amplitudová modulace).

Na první pohled se může zdát, že proces samplování je velice jednoduchý (zapojíme mikrofon, nastavíme vzorkovací frekvenci a nahrajeme určitý zvuk). Faktem ale je, že vytvořit dobrý vzorek zvuku je často poměrně složitý proces, který vyžaduje dobré znalosti z oboru akustiky.

Nyní A/D převodník trochu podrobněji. Tento převodník provádí dvě základní operace. Jednak úrovně kvantování signálu a dále převod úrovně kvantového signálu do digitální podoby. Při úrovněm kvantování zaokrouhlujeme velikosti jednotlivých vzorků signálu na nejbližší hodnotu určitého vhodně zvoleného rasteru. Zaokrouhlujeme proto, abychom mohli vyjádřit vzorky konečným počtem bitů. Bit je základní informační jednotka nabývající buď hodnoty 0 nebo 1. Kvantovaný vzorek může nabývat pouze jedné z  $2^n$  různých hodnot, kde  $n$  je počet bitů příslušného A/D převodníku, což potřebujeme. Kdybychom nezaokrouhlovali (jinými slovy – kdybychom velikost každého vzorku chtěli vyjádřit s nekonečně velkou přesností), potřebovali bychom k tomu nekonečný počet bitů. Po úrovněm kvantování přichází na řadu PCM kódování (zkratka PCM označuje pulsni kódovou modulaci – tedy kódování velikostí impulsů do digitální podoby, reprezentované nulami a jedničkami). Při PCM kódování musíme velikost každého zaokrouhleného vzorku vyjádřit ve dvojkové soustavě. Celý proces je o něco složitější a velmi dobře je popsán Bensonem v [1].

## 5.2 Aliasing

Aliasing je jev, ke kterému může docházet v situacích, kdy se spojitá informace převádí na diskrétní. Slovo aliasing znamenající v češtině falšování (resp. chyba vzorkováním) přesně vystihuje jev, ke kterému dojde při nedodržení Shannonova teorému. Aliasing můžeme popsat jako přítomnost umělých komponent v digitálním signálu, který jsme získali z analogového. Důvod, proč k tomuto jevu dochází je, že samplovací frekvence původního signálu je příliš nízká a tyto frekvence se překrývají. Aby při samplování nedocházelo k aliasingu, musí být vzorkovací frekvence rovna minimálně dvojnásobku nejvyšší frekvence obsažené ve vzorkovaném signálu - tzv. Shannonův teorém (někdy nazývaný Nyquistův, Shannon-



Nyquistův teorém nebo také Shannon-Kotělnikovův teorém):

$$f_m < \frac{f}{2}, \quad (24)$$

kde  $f_m$  je kmitočet signálů přicházejících na vstup a  $f$  je vzorkovací frekvence. Pokud tuto podmínku vzorkovací frekvence nespĺňuje, dochází k překrytí frekvencí vzorkovaného signálu a tedy ke ztrátě informace. Znamou ukázkou aliasingu je například filmový záznam nějakého rychle se otáčejícího předmětu, např. vrtule letadla. Abychom se vyhnuli problémům způsobovaným aliasingem, musí být na vstupu vždy zařazen dolnoproputní analogový filtr (tzv. anti-aliasing filtr), který signály, jejichž frekvence nespĺňují podmínku Shannonova teorému, potlačí. Velmi dobře je aliasing a smysl Shannonova teorému popsán Bensonem v [1]: uvažujme čistě sinusovou vlnu s frekvencí  $\nu$ , např.

$$f(t) = A \cos(2\pi\nu t).$$

Je dána vzorkovací rychlost  $N = \frac{1}{\Delta t}$  vzorků za sekundu a výška funkce v  $M$ -tém vzorku je dána vztahem

$$f(M/N) = A \cos(2\pi\nu M/N)$$

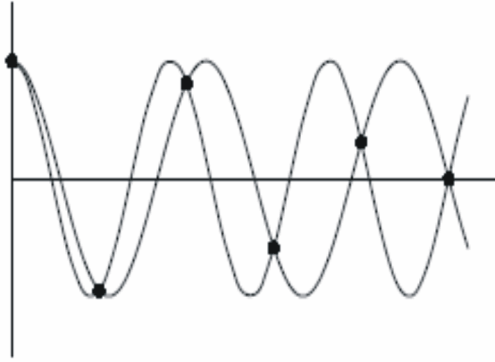
Je-li  $\nu$  větší než  $\frac{N}{2}$ , řekněme  $\frac{N}{2} + \alpha$ , pak

$$\begin{aligned} f(M/N) &= A \cos(2(N/2 + \alpha)M\pi/N) \\ &= A \cos(M\pi + 2\alpha M\pi/N) \\ &= (-1)^M A \cos(2\alpha M\pi/N). \end{aligned}$$

Změna znaménka u  $\alpha$  nemá dopad na výsledek výpočtu, dostaneme tedy stejný výsledek jako u vlny  $\nu = N/2 - \alpha$  namísto  $\nu = N/2 + \alpha$ . Když se na to podíváme z jiného úhlu, vzorkové body v tomto výpočtu jsou přesně ty samé body, kde se grafy funkcí  $A \cos(2(N/2 + \alpha)\pi t)$  a  $A \cos(2(N/2 - \alpha)\pi t)$  kříží, což vidíme na Obrázku 3.

Výsledkem tohoto je, že frekvence, která je větší než polovina vzorkovací frekvence, se odráží skrz polovinu vzorkovací frekvence, zní tedy jako frekvence odpovídající hodnotě menší než polovina vzorkovací frekvence. Tento jev se nazývá *aliasing*. Na Obrázku 3 jsou vzorkové body označeny černými tečkami. Vykreslené dvě frekvence mají mírně vyšší a mírně nižší frekvenci než polovina snímací frekvence. Vzorkové body jsou tedy body, kde se tyto dva grafy na Obrázku 3 kříží.

U vln s přesně polovinou vzorkovací frekvence nastává něco zajímavého – kosinové vlny zůstávají nedotčeny, ale sinové vlny naprosto zmizí. To znamená, že fázová informace je ztracena a informace o amplitudě je zkosena. Závěrem Shannonova teorému je to, že před digitalizací analogového signálu je zásadním krokem nechat ho projít skrz dolní propust, aby se odstranily frekvence nad polovinou vzorkovací frekvence. Jinými slovy, každá frekvence z ní vystoupí v páru se svým odrazem. V případě digitálních kompaktních disků je tou mezní frekvencí polovina z 44,1 KHz, tedy 22,05 KHz. Jelikož je hranice slyšitelnosti u člověka obvykle nižší než 20 KHz, můžeme to považovat za vyhovující bez rezervy.



Obrázek 3: Křížení grafů funkcí (viz [1]).

### 5.3 Přechod od Fourierovy transformace k diskrétní Fourierově transformaci

Výpočet spektra signálu podle vztahu (22) naráží na jeden problém: potřebujeme znát průběh signálu přes nekonečný časový interval. Trvání v praxi se vyskytující signálů je vždy konečné a jejich spektrum tedy, striktně vzato, není definováno. Musíme je tedy nějakým způsobem rozšířit na nekonečný interval a počítat spektrum tohoto náhradního signálu. Přechod od spektra definovaného jako Fourierova transformace spojitěho signálu, které je spíše teoretického významu, k DFT (resp. FFT) konečného počtu vzorků používaný v praxi není zcela přímočarý a je vhodné jej rozložit do několika na sebe navazujících kroků.

Kromě toho, že v praxi se vyskytující signály jsou definovány pouze na omezeném časovém intervalu, jejich spektrum bývá také nenulové pouze na omezeném intervalu. Máme-li signál o spektru omezeném na intervalu  $[-\frac{\pi}{T}, \frac{\pi}{T}]$ , můžeme jej dle Shannonova vzorkovacího teorému (24) beze ztráty informace, a tudíž s možností jeho úplné rekonstrukce, navzorkovat se vzorkovací periodou  $T$ . Necht'  $f_n, n = 0, \dots, N-1$  je původní samplovaný signál v bodech  $nT$ . Náhradou integrálu sumou v definici Fourierovy transformace (22) dostaneme výraz:

$$F(\omega) = \sum_{n=-\infty}^{+\infty} f_n e^{-i\omega n}. \quad (25)$$

Použijeme-li z celého nekonečně dlouhého signálu pouze konečný počet  $N$  vzorků, popř. ostatní vzorky signálu položíme rovny nule, získáme z předchozí transformace (25) výraz:

$$F(\omega) = \sum_{n=0}^{N-1} f_n e^{-i\omega n}. \quad (26)$$

Tento obraz signálu  $f_n$  je tedy, stejně jako v předchozím případě, spojitý a periodický, přesto nese o signálu pro jeho zkrácení méně informací než nesly obrazy

předchozí. To se zpravidla projeví snížením “rozlišení”, tj. jemností detailů podél kmitočtové osy.

Posledním krokem pomyslného přechodu k DFT je navzorkování spojitého obrazu v kmitočtové oblasti (26) v bodech  $\omega_k = \frac{2\pi k}{NT}$ ,  $k = 0, 1, \dots, N-1$ . Toto navzorkování je možné provést při zachování možnosti obnovy původního spojitého průběhu, a to je zajištěno zkrácením vstupního signálu na  $N$  vzorků provedené v předchozím kroku. Výsledkem je *Diskrétní Fourierova transformace (DFT)* definovaná vztahem (značení zjednodušíme, tj.  $F(\omega_k) = F_k$ ):

$$F_k = \sum_{n=0}^{N-1} f_n e^{-2\pi i \frac{kn}{N}}, \quad \text{kde } k = 0, \dots, N-1. \quad (27)$$

tedy zobrazení  $\mathbb{R}^N \rightarrow \mathbb{R}^N$ . Vztah pro inverzní diskrétní Fourierovu transformaci (značíme IDFT) můžeme získat např. vynásobením rovnice (27) činitelem  $e^{ikl\frac{2\pi}{N}}$  a vytvořením sumace podle  $k$ :

$$\sum_{k=0}^{N-1} F_k e^{ikl\frac{2\pi}{N}} = \sum_{n=0}^{N-1} f_n \left[ \sum_{k=0}^{N-1} e^{ik(l-n)\frac{2\pi}{N}} \right]. \quad (28)$$

Sumace v hranatých závorkách je rovna 0 pro  $l \neq n$  a rovna  $N$  pro  $l = n$ , takže dostaneme:

$$f_l = \frac{1}{N} \sum_{k=0}^{N-1} F_k e^{ikl\frac{2\pi}{N}} \quad (l = 0, 1, 2, \dots, N-1). \quad (29)$$

*Důkaz.*

$$\begin{aligned} \frac{1}{N} \sum_{k=0}^{N-1} F_k e^{ikl\frac{2\pi}{N}} &= \frac{1}{N} \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} f_j e^{-ikj\frac{2\pi}{N}} e^{ikl\frac{2\pi}{N}} = \\ &= \frac{1}{N} \sum_{j=0}^{N-1} f_j \sum_{k=0}^{N-1} e^{ik(l-j)\frac{2\pi}{N}} = \frac{1}{N} \sum_{j=0}^{N-1} f_j \delta_N(l-j) = f_l, \end{aligned}$$

kde  $\delta_N(l) = \frac{1}{N} \sum_{k=0}^{N-1} e^{ikl\frac{2\pi}{N}}$ . □

Vztah (29) představuje inverzní DFT, která umožňuje vypočítat diskrétní hodnoty  $f_n$  ze zadaných diskrétních hodnot spektra  $F_k$ , které můžeme např. odečíst ze zadaného spojitého průběhu spektra.

Diskrétní Fourierova transformace je zobrazením, které periodický diskrétní vektor  $f_n$  délky  $N$  převede na periodický diskrétní vektor  $F_k$  délky  $N$ ,  $f_n \rightarrow F_k$  a inverzní diskrétní Fourierova transformace je zobrazením opačným, tj.  $F_k \rightarrow f_n$ .

Diskrétní Fourierova transformace našla velké uplatnění zejména s rozvojem výpočetní techniky. Součástí řady přístrojů jsou jednoúčelové procesory, které realizují tuto transformaci. Její hlavní rozvoj nastal po roce 1965, kdy J.W.Cooley

a J.W.Tukey popsali velmi efektivní algoritmus výpočtu DFT, tzv. rychlou Fourierovu transformaci (Fast Fourier Transform – FFT). Díky tomuto algoritmu se stala diskrétní Fourierova transformace nejrozšířenějším prostředkem pro výpočet Fourierovy transformace. Kromě algoritmu Cooley-Tukey jsou používány další algoritmy, které se liší především vyjádřením FFT a jsou to algoritmy FFT: Bluestein's, Bruun's, Prime-factor, Rader's a Split-radix. Algoritmus FFT je také implementován ve všech nejrozšířenějších matematických programech jako je např. Maple, Matlab, Mathcad, GNU Octave, Mathematica, atd.

## 5.4 Rychlá Fourierova transformace

Kapitola o rychlé Fourierově transformaci (Fast Fourier Transform – FFT) čerpá zejména z literatury [1]. FFT je způsob výpočtu diskrétní Fourierovy transformace přímočarým způsobem, při kterém je třeba méně aritmetických operací než při užití rovnice (27). Vztah pro výpočet diskrétní posloupnosti  $F_k$  ze zadané posloupnosti hodnot  $f_n$  je dán rovnicí (27) a pro inverzní postup vztahem (29). S ohledem na výklad postupu při výpočtech podle těchto rovnic zavedeme substituci:

$$W_N^{nk} = e^{-i\frac{2\pi}{N}nk}. \quad (30)$$

Vztahy pro DFT a IDFT pak budou ve tvaru:

$$DFT = F_k = \sum_0^{N-1} f_n e^{-i\frac{2\pi}{N}nk} = \sum_0^{N-1} f_n W_N^{nk} \quad (31)$$

$$IDFT = f_n = \frac{1}{N} \sum_0^{N-1} F_k e^{i\frac{2\pi}{N}nk} = \frac{1}{N} \sum_0^{N-1} F_k W_N^{-nk}. \quad (32)$$

Výpočet diskrétních hodnot  $F_k$  nebo  $f_n$  podle těchto vztahů je při delší posloupnosti  $N$  zadaných diskrétních hodnot časově velmi náročný. Vyžaduje velké množství aritmetických operací, protože optimálním způsobem pro výpočet polynomů je Hornerovo schéma, které potřebuje  $N - 1$  násobení a sčítání, což pro celou transformaci dává počet operací  $2 \times N \times (N - 1)$  tedy přibližně  $N^2$  operací. Když si představíme, že  $N$  může být např. 0.5 – 1 milion a čas výpočtu stoupá úměrně s  $N^2$ , je výpočet diskrétní Fourierovou transformací nepraktický. Důležitým krokem ve vytváření časové úspory je minimalizace počtu násobení a výpočet pomocí FFT narozdíl od DFT vyžaduje pouze  $N \log_2 N$  operací.

**Příklad 1.** Uvažme situaci, že  $N = 4$ . V tom případě jsou čísla  $e^{\frac{2\pi ik}{N}}$  rovnoměrně rozložena po jednotkové kružnici v komplexní rovině, takže

$$\begin{aligned} e^0 &= 1 & (k = 0) \\ e^{\frac{2\pi i}{4}} &= i & (k = 1) \\ e^{\frac{4\pi i}{4}} &= -1 & (k = 2) \end{aligned}$$

$$e^{\frac{6\pi i}{4}} = -i \quad (k = 3)$$

rovnice se tedy redukuje na

$$F_0 = f_0 + f_1 + f_2 + f_3$$

$$F_1 = f_0 - if_1 - f_2 + if_3$$

$$F_2 = f_0 - f_1 + f_2 - f_3$$

$$F_3 = f_0 + if_1 - f_2 - if_3$$

a

$$f_0 = \frac{1}{4}(F_0 + F_1 + F_2 + F_3)$$

$$f_1 = \frac{1}{4}(F_0 + iF_1 - F_2 - iF_3)$$

$$f_2 = \frac{1}{4}(F_0 - F_1 + F_2 - F_3)$$

$$f_3 = \frac{1}{4}(F_0 - iF_1 - F_2 + iF_3).$$

K vysvětlení fungování rychlé Fourierovy transformace použijeme  $N$  jako sudé číslo. Můžeme tedy rozdělit součet (31) na sudé a liché výrazy:

$$F_k = \sum_{n=0}^{\frac{N}{2}-1} f_{2n} e^{\frac{-2\pi i(2n)k}{N}} + \sum_{n=0}^{\frac{N}{2}-1} f_{2n+1} e^{\frac{-2\pi i(2n+1)k}{N}}.$$

Rozhodující je, že hodnota  $F_{k+\frac{N}{2}}$  je velmi podobná hodnotě  $F_k$ . Zjištěním, že  $e^{-\pi i(2n)k} = 1$  a  $e^{-\pi i(2n+1)k} = (-1)^k$ , obdržíme

$$F_{k+\frac{N}{2}} = \sum_{n=0}^{\frac{N}{2}-1} f_{2n} e^{\frac{-2\pi i(2n)k}{N}} + (-1)^k \sum_{n=0}^{\frac{N}{2}-1} f_{2n+1} e^{\frac{-2\pi i(2n+1)k}{N}}.$$

Můžeme tedy vypočítat hodnoty  $F_{k+\frac{N}{2}}$  za stejný čas s polovičním úsilím, které by bylo jinak potřeba, plus je třeba přičíst další čas na provedení sčítání a odčítání výsledků. Ony dva součty, které počítáme, jsou samy o sobě diskrétními Fourierovými transformacemi (s pravou stranou násobenou  $e^{\frac{-2\pi ik}{N}}$ ) pro  $\frac{N}{2}$  bodů namísto  $N$  bodů, takže je-li  $\frac{N}{2}$  sudé, můžeme zopakovat tentýž postup. Na Příkladu 1 se tímto upravuje výpočet následovně:

$$F_0 = (f_0 + f_2) + (f_1 + f_3)$$

$$F_2 = (f_0 + f_2) - (f_1 + f_3)$$

$$F_1 = (f_0 - f_2) - i(f_1 - f_3)$$

$$F_3 = (f_0 - f_2) + i(f_1 - f_3).$$

## 6 Tvorba a šíření zvuku

Obsah této kapitoly je z velké části převzat z literatury [14]. Začneme tím, jak fyzicky zvuk vzniká.

### 6.1 Vznik zvuku

Zvuk je tvořen vždy, když je porušena rovnováha hustoty (nebo tlaku) plynu, kapaliny nebo pevné látky. Jestliže lokální přerušeni tlaku (hustoty) obsahuje frekvence v rozmezí od 20 do 20 000 Hz (tzv. slyšitelné pásmo), vzniká zvuk. Frekvence pod hranicí 20 Hz se nazývají infrazvuk a frekvence nad 20 000 Hz pak ultrazvuk. Jestliže přerušeni tlaku je pozitivní, vzniká zhušťování ve středu a tím hustota stoupá. Naopak, jestliže je toto přerušeni negativní, vzniká ředění a tedy hustota klesá. Nyní uvedeme tři konkrétní situace s popisem vzniku zvuku.

Představme si válec s pístem uvnitř – pohybem pístu je vzduch střídavě zhušťován (kondenzován) a ředěn. Částice, které jsou v kontaktu s pístem, se snaží následovat jeho kmitavý pohyb, ale protože nejsou s pístem ani s ničím jiným pevně spojeny, vytváří svůj vlastní pohyb nárazem. Hybnost je přenášena na přilehlé částice s časovým zpožděním na bod dále od pístu. Toto “přenášení” přerušeni tlaku tvoří vlastní zvuk.

V případě tleskání je zvuk produkován náhlým přerušeni proudu vzduchu, který je stlačen mezi dlaněmi. Toto dočasné přerušeni proudu vzduchu může být zesíleno vzdálením dlaní od sebe, tím se vytvoří rezonanční komora.

Dalším zajímavým zdrojem zvuku je siréna. Mějme plný kotouč po jehož obvodu jsou pravidelně vzdálené otvory. Proud vzduchu je směřován přímo na otvory po obvodu kola. Periodické rušení proudu vzduchu těmito otvory způsobuje značné přerušeni tlaku a tím velmi intenzivní zvuk. Čím bude rychlejší točení kotouče, tj. pouštění a zastavování proudu vzduchu, tím vyšší bude frekvence takto vzniklého zvuku. Siréna je důležitým zdrojem zvuku pro účast hlasu při mluvení, protože se podobá hlasivkovému zdroji zvuku. V obou případech je vzduch nucený periodicky procházet skrz otvor. Výsledkem je tedy série impulsů.

Zajímavý aspekt je produkce zvuku u lidí a zvířat, protože malá zvířata mohou často vydávat zvuky, které jsou mnohem intenzivnější než ty od zvířat velkého vzrůstu. Je zde určitý kompromis mezi objemem vzduchu, který je hnaný zdrojem, a frekvencí, na které je tento vzduch produkován. “Malé zdroje” mohou být velmi intenzivní, jestliže se jejich proud vzduchu rychle mění.

#### 6.1.1 Zvukové spektrum, tóny a šумы

*Zvukové spektrum* v akustice představuje kombinaci všech frekvencí, ze kterých se většina zvuků (tedy mechanické vlnění) skládá. Většina zvuků, které se v přírodě

vyskytují, nejsou jednoduché tóny o jedné frekvenci, ale mají složitý průběh, který lze analyzovat pomocí přístrojů.

*Tóny* mají pravidelný průběh, který je složen z jedné nebo více frekvencí, na které lze zvuk při analýze jeho spektra rozložit. Tónem v akustice je každý zvuk se stálou frekvencí. Nejnižší frekvence tónu je považována za základní a udává výšku tónu. Ostatní, vyšší frekvence, se označují jako alikvoty<sup>3</sup> a ovlivňují výsledný sluchový dojem nazývaný barva tónu (témbr). Jednotlivé různé tóny se tedy liší frekvencí, ale lidské ucho nevnímá rozdíl těchto frekvencí, nýbrž rozlišuje je na základě jejich podílu. Tón s dvojnásobnou frekvencí zní pro lidské ucho o jednu oktávu výše. Každý tón by mohl být popsán svou frekvencí (vyjádřenou např. v Hz), v praxi by ale bylo takové označení velmi nepohodlné. Frekvence se proto uvádí jen pro základní tón, kterým je tzv. komorní A, které má frekvenci 440 Hz.

**Poznámka 7.** Komorní “A” neboli jednočárkové “A” je základem veškerého moderního ladění. Jako mezinárodní standard bylo toto ladění ustanoveno v Paříži roku 1858, ale jeho používání bylo známo i mnohem dříve viz [23]. Nejprve každá škola či směr používaly jako komorní “A” svoji vlastní normu, která se liší vždy jednotlivou frekvencí tohoto základního tónu. V roce 1958 se mezinárodní komise dohodla na používání komorního “A” o frekvenci 440 Hz. Toto moderní ladění je o něco vyšší než byla ladění používaná v barokní a starší hudbě. Např. při ladění anglických píšťalových varhan v roce 1720 bylo komorní “A” na 380 Hz a při ladění používaném převážně v barokní hudbě na 415 Hz. Norma komorního “A” byla ustanovena hlavně z důvodu, aby bylo zaručeno, že klavíry a jiné nástroje mají stejnou tónovou výšku a díky tomu mohly nástroje společně hrát a nedocházelo tak k nepatřičným zvukovým disonacím či neladění.

Na rozdíl od tónu mají šумы, neboli hluky, nepravidelný průběh kmitů. Mnoho zvuků nejsou čistými tóny a ani čistými šумы, ale mají jak složky tónové (pravidelné kmity o určitých frekvencích), tak i šumové (nepravidelnosti). V lidské řeči jsou samohlásky složenými tóny a souhlásky jsou šумы.

Mezi základní fyzikální vlastnosti zvuku patří:

- intenzita, která je určena velikostí “zvukové energie”, jež projde za sekundu plochou  $1m^2$  kolmou na směr šíření zvuku, intenzity ubývá se čtvercem vzdálenosti od zdroje zvuku
- výška, která je dána u jednoduchého tónu kmitočtem jeho zdroje, u složeného tónu interferencí výšek tónů částkových, udává se v hertzech [Hz]: kmitočet rovný 1 Hz má takový periodický děj, jehož 1 kmit trvá 1 sekundu

---

<sup>3</sup>Alikvotní tón, nebo též vyšší harmonický tón, je tón, který zní společně s tónem základním. Většinou se u každého zvuku vyskytuje množství alikvotních tónů a jejich intenzita je to, co určuje charakteristickou barvu zvuku. Právě díky alikvotním tónům jsme schopni poslechem rozpoznat, o jaký se jedná hudební nástroj.

- barva zvuku, která vzniká u složených zvuků na základě interference výšek (viz Poznámka 8) a intenzit částkových tónů
- kvantita, tj. trvání v čase.

**Poznámka 8.** Interference je charakteristickou vlastností vln a znamená vzájemné ovlivňování či prolínání vln. Při jejich pohybu a prolínání se v určitém bodě vzájemně zesilují, zatímco v jiných bodech vzájemně ruší. Při zpracování nahrávaných samohlásek se projeví tak, že při vstupu dvou tónů na frekvenci přesně 440 Hz, dostaneme na výstupu opět tón na frekvenci 440 Hz, protože vycházíme z goniometrického vzorce

$$\sin x + \sin y = 2 \sin \frac{x+y}{2} \cos \frac{x-y}{2}$$

kde za  $x$  a  $y$  dosazujeme součin:  $2\pi$ , frekvence v Hz a čas v sekundách. Když vezmeme nahrávky dvou komorních A, dostaneme tedy při vstupech  $x = 2\pi 440t$  a  $y = 2\pi 440t$

$$\sin 2\pi 440t + \sin 2\pi 440t = 2 \sin \frac{2\pi 880t}{2} \cos 0 = 2 \sin 2\pi 440t.$$

Ale pokud vezmeme dva tóny o frekvencích s malým rozdílem např. 440 Hz a 460 Hz, ve výsledku dostaneme:

$$\sin 2\pi 460t + \sin 2\pi 440t = 2 \sin \frac{2\pi 900t}{2} \cos \frac{2\pi 20t}{2} = 2 \sin 2\pi 450t \cos 2\pi 10t,$$

což se potom při poslechu tohoto tónu projeví pravidelným “kolísáním” zvuku s frekvencí 10 Hz, tón zní “přerušovaně”. Tímto způsobem se může stát, že součtem tónů s frekvencemi  $f_1$  a  $f_2$  vznikne tón, jehož zdánlivá základní frekvence je výrazně menší než  $f_1$  a  $f_2$ .

Znatelný rozdíl ve výsledné hlásce můžete sami posoudit na vzorcích, které najdete v příloženém CD ve složce Interference. Soubor “stejně” vystihuje situaci první (tj. obě vstupní frekvence jsou stejné a to 440 Hz), naproti tomu hláska “ruzne” vychází z druhé situace, tedy že vstupní frekvence jsou 440 Hz a 460 Hz, rozdíl je slyšitelný.

### 6.1.2 Formanty

Jednotlivé samohlásky se stejnou délkou, výškou a intenzitou se liší barvou tónu (tzv. témbrem), která je dána vlivem spolupodílejících se vedlejších frekvencí – formantů, které vznikají rezonancí dutin (hltanová, nosní a ústní). Rezonance je ovlivňována změnami tvaru těchto dutin během řeči podle polohy mluvidel (artikulátorů). Formanty jsou charakteristické části spektra zvuku – oblasti posílení frekvencí, které identifikují samohlásku pro posluchače. Podle rostoucí frekvence se formanty označují F1, F2 atd., přičemž F1 a F2 a jejich vzájemný poměr mají pro znění hlásky zásadní význam.



První dva formanty se vyskytují zpravidla do hranice 2500 Hz. V Tabulce 1 se můžeme podívat, v jakých oblastech se pohybují tyto formanty pro jednotlivé české samohlásky, v Tabulce 2 jsou formanty pro anglické samohlásky. Konkrétní případy pak uvedeme v praktické části této práce.

Hláška	F1 [Hz]	F2 [Hz]
i	300 - 450	2100 - 2800
e	500 - 700	1600 - 2100
a	800 - 1100	1100 - 1500
o	500 - 700	900 - 1200
u	300 - 500	600 - 1000

Tabulka 1: Hodnoty formantů českých samohlásek (viz [21]).

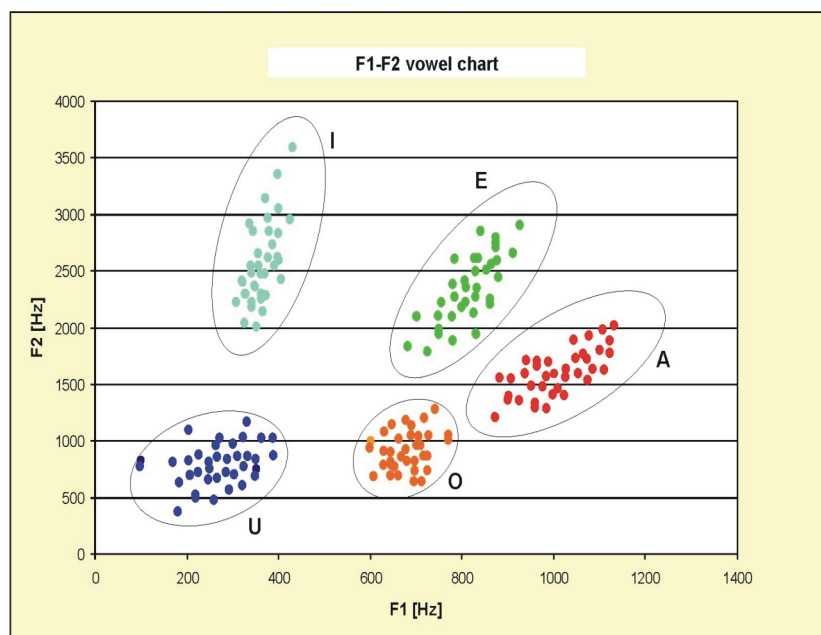
Hláška	F1 [Hz]	F2 [Hz]
i	300 - 450	2000 - 3600
e	650 - 900	1600 - 3000
a	800 - 1200	1100 - 2100
o	600 - 800	500 - 1400
u	100 - 400	400 - 1100

Tabulka 2: Hodnoty formantů anglických samohlásek (viz [21]).

Nutno podotknout, že jako první byly indetifikovány oblasti formantů pro anglické samohlásky. Od českých formantů se mírně liší, což je dáno především odlišnou výslovností samohlásek anglických a českých. Typické pro zobrazení prvního a druhého formantu jsou tzv. “formantové mapy” jak můžeme vidět na Obrázku 4, kde na vodorovné ose jsou frekvence prvního formantu a na svislé ose druhého.

## 6.2 Vznik hlasu v lidském těle

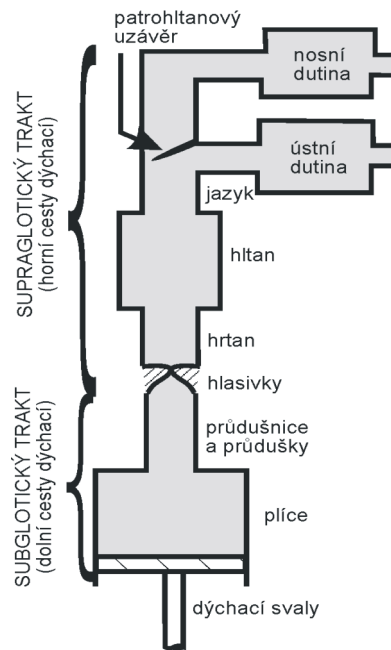
Lidský hlas je poměrně komplikovaný fyzikální jev. Na to, že s ním přicházíme do styku doslova denně, je kupodivu, že (stejně jako mnoho jiných neurofyzilogických procesů) není zcela uspokojivě pochopena ani jeho tvorba ani vnímání. Lidský hlas je periodickým zhušťováním a zředováním vzduchu. Vznik hlasu vysvětlují různé teorie. Nejznámější z nich je myoelastická (tónická, aerodynamická), která vysvětluje vznik hlasu činností hlasivkových svalů a vlivem supraglotálního (viz níže) tlaku vzduchu. Vlastnosti lidského hlasu jsou dány fyziologií hlasového ústrojí. Význam má především délka hlasivek (u žen kolem 1,2 cm, u mužů 1,5 – 2,5 cm). Čím jsou hlasivky kratší, tím rychleji kmitají a běžný mluvený hlas je vyšší.



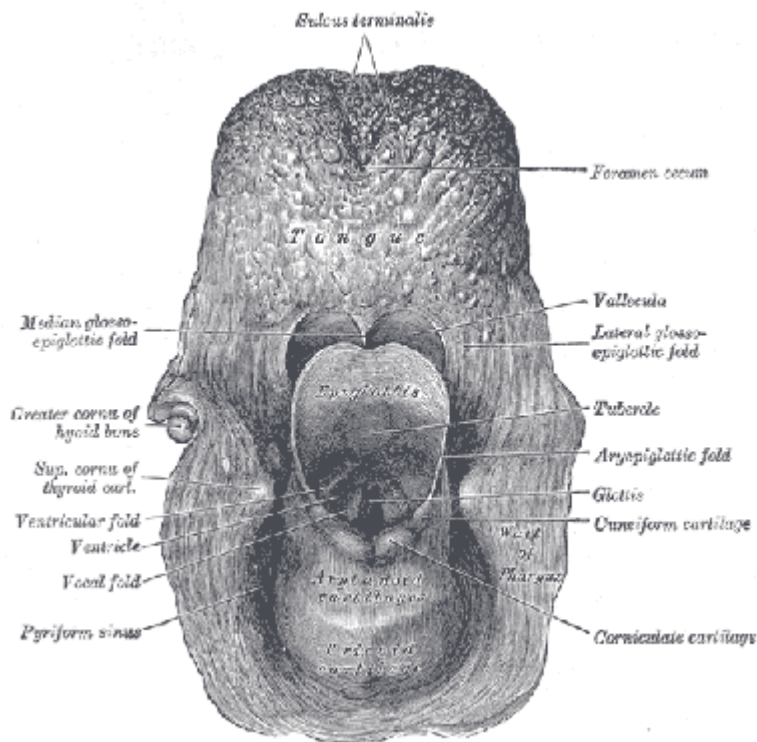
Obrázek 4: Mapa prvního a druhého formantu pro anglické samohlásky.

Hlas se v lidském těle tvoří v dýchacím ústrojí, které anatomicky zahrnuje plíce, průdušky, průdušnici, hrtan (neboli larynx), hltan, jícen a ústní a nosní dutinu. Celý proces je velmi dobře popsán v práci J. Švece viz [13], z které je pro tuto část čerpáno. V souvislosti s vytvářením hlasu hovoříme o dýchacím ústrojí jako o hlasivkovém akustickém systému. Celý proces probíhá ve dvou traktech, a to v subglotickém (pod hlasivkami) a supraglotickém (nad hlasivkami) traktu. Do subglotálního traktu zahrnujeme cestu vzduchu od plic, kde dýchací svaly do nich nátáhnou vzduch a stejně tak jej posílají zpět přes průdušnici a průdušky k hrtanu. V hrtanu jsou umístěny hlasivky, jejichž kmity modulují proud vzduchu, a tím vytvářejí periodické změny vzdušného traktu – vzniká akustický signál. Zde tento trakt končí a začíná trakt supraglotický, do kterého patří hrtan, hltan a dutina ústní a nosní. Nejdůležitějším rezonančním prostorem hlasového akustického systému, jenž se v největší míře podílí na artikulaci, je část právě supraglotického traktu zvaná vokální trakt, kterou je označován prostor mezi hlasivkami a ústním otvorem.

Hlasivky (chordae vocales) bychom mohli popsat jako dva hlasové valy pokryté sliznicí. Mezi hlasivkami se vytváří hlasivková štěrbina (glottis), jejíž dvě třetiny jsou ohraničeny blanitou částí hlasivek. Hlasivky jsou napjaty mezi chrupavkou štítnou, kde se navzájem stále dotýkají, a chrupavkami hlasivkovými. Hlasivkové chrupavky jsou umístěny vzadu a jsou částečně pohyblivé. Jejich kloubní spojení s prstencovou chrupavkou je mimořádně složité. Také způsob jejich svalového a nervového ovládní je velmi komplikovaný a nebyl dosud v úplnosti popsán a vyložen. Při sevření hlasivek pod nimi vzniká přetlak (zhuštění vzduchového sloupce), zatímco nad hlasivkami je podtlak (zředění vzduchu). Vznik přetlaku

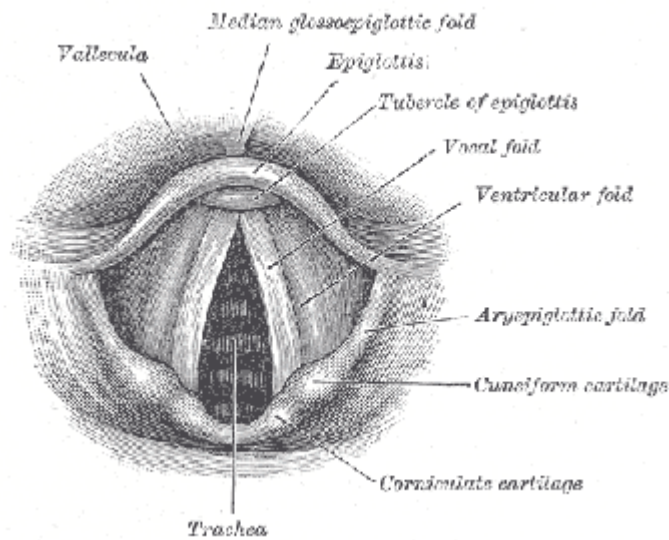


Obrázek 5: Znázornění hlasivkového akustického systému (viz [13]).



Obrázek 6: Hlasivková štěrбина (glottis) zobrazená zepředu viz [20].

je dán tím, že výdechový proud je plynulý a překážka sevřených hlasivek jej na okamžik zbrzdí. Přetlak pod hlasivkami vede k jejich oddálení, množství vzduchu



Obrázek 7: Hlasivková štěrbiná (glottis) zobrazená shora viz [20].



Obrázek 8: Fotografie hlasivkové štěrbiny viz [17].

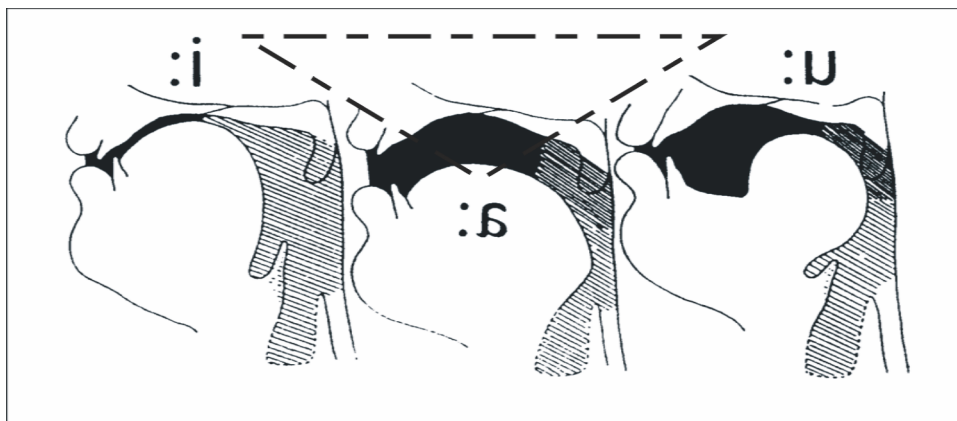
pod hlasivkami unikne, vzniknou podmínky pro nové sblížení atd. Podle nových výzkumů napomáhá sblížení hlasivek i proudění vzduchu kolem nich. Není vyloučeno ani nervové ovlivnění celého procesu.

Hlasivky se chvějí pasivně (jako vlajka ve větru) a aktivně (naší vůlí) můžeme ovlivnit pouze jejich napětí a jiné mechanické vlastnosti. Hlasivky pravidelně přerušují proudící vzduch a vytvářejí tím základní periodický signál, ve kterém je zastoupena pouze základní frekvence (odpovídající výšce tónu) a její násobky (odpovídající vyšším harmonickým frekvencím).

Takový zvuk ještě nemá žádnou barvu a nelze v něm rozeznat žádnou určitou samohlásku. Charakteristické znění, individuální pro jednotlivce, získává až průchodem nadhrtanovými prostory – rezonátory, v nichž se mění některé ze svrchních tónů základního hlasu. Základní periodický signál vytvořený hlasivkami vchází do složitého systému rezonančních dutin, které můžeme různě tvarovat

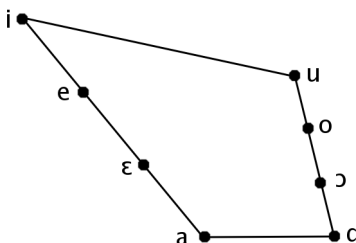
(především pohybem čelistí, jazyka a rtů). Podle toho, jak se nastaví tvar rezonančních dutin, dochází v původním signálu k potlačení či zesílení některých frekvenčních pásem a tato posílená frekvenční pásma se nazývají formanty.

Jazyk funguje při řeči různě. U vokálů se posunuje nahoru a dopředu nebo celá hmota jazyka dozadu, čímž se různě modifikuje rezonanční prostor úst a hrdla. Tento posun je pro jednotlivé samohlásky typický a stal se proto základem pro systemizaci vokálů. Systém vokálů se zobrazuje schématem, jehož východiskem je trojúhelníkový obrazec vytvořený koncem 18. století Ch. F. Hellwagem (podle toho je pojmenován jako “Hellwagův trojúhelník” viz Obrázek 9). Obrazec zachycuje vertikální a horizontální posun jazyka z klidové polohy. Pro potřeby srovnávací fonetiky vytvořil anglický fonetik Daniel Jones v r. 1914 sérii “základních samohlásek” (cardinal vowels) a jejich artikulační a akustické charakteristiky přesně popsal. Základní samohlásky jsou ty, kdy jazyk je v extrémní poloze tj. vpředu, vzadu, nahoře nebo dole. Jonesovo zobrazení je přesnější (než popis Ch. F. Hellwagem) viz Obrázek 10 a naznačuje poněkud i relace výšky různých vokálů. Schémata se stala východiskem dnešních popisů soustav samohlásek různých jazyků.

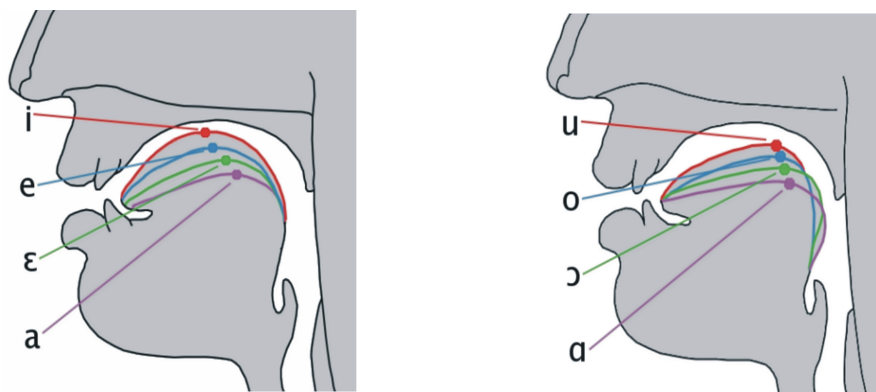


Obrázek 9: Hellwagův trojúhelník.

Na proces formování samohlásek je možné se dívat ze dvou hledisek. Lze sledovat proces formování výsledné akustické vlny nebo proces formování výsledného frekvenčního spektra. Zmíněné dva přístupy odrážejí dvě klasické teorie tvorby



Obrázek 10: Schéma základních samohlásek podle D. Jonese viz [24].



Obrázek 11: Přední a zadní poloha jazyka při základních samohláskách viz [24].

hlasu, které byly formulovány už v minulém století německým fyzikem H. Helmholtzem a německým fyziologem L. Herrmannem. V minulosti probíhaly četné diskuze, která z těchto teorií je správná. Dnes je již zřejmé, že obě teorie jsou v podstatě ekvivalentní a jen vyjadřují různý úhel pohledu na tentýž problém. Původní Herrmannova teorie vysvětluje formování akustické vlny tak, že při každém otevření hlasivek jejich štěrbinou projde puls vzduchu a ten rozezvučí jednotlivé dutiny (rezonátory) vokálního traktu. Ty vydají krátký tlumený tón, který rychle odezní. Tento děj se opakuje každý cyklus znovu a znovu. Charakteristické tóny dutin nazval Herrmann formanty, neboť formují hlásku. Helmholtzova teorie samohlásek vycházela v podstatě z představy dutin vokálního traktu působících jako rezonátory, které zesilují harmonické tóny prvotního hlasového signálu.

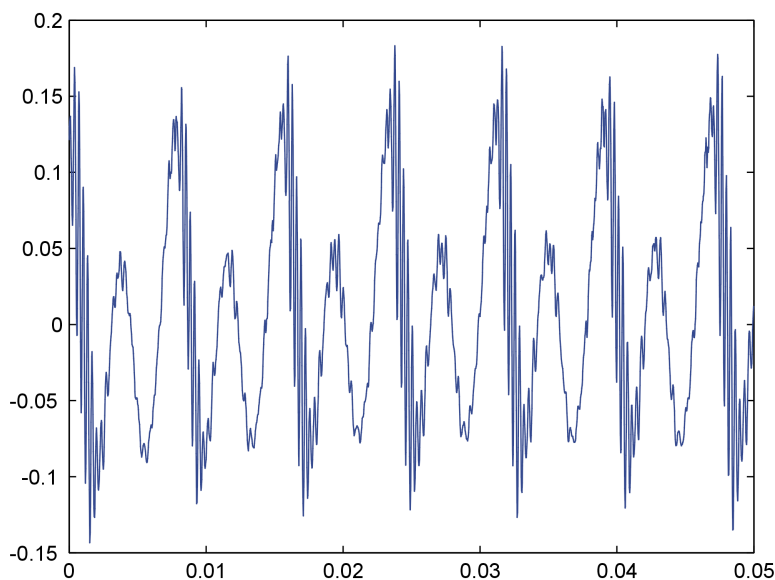
### 6.3 Zjišťování formantů v současnosti

Software, který je v dnešní době využíván pro hledání formantů a analýzu řeči jako takové, vyvinuli Paul Boersma a David Weenink z University of Amsterdam. Jeho název Praat vznikl podle slova “mluvit” v afrikánštině a nizozemštině. Software je zdarma ke stažení, takže je možné jej vyzkoušet. Umožňuje nahrávat celá slova i samotné samohlásky a vykreslit formanty ve formě teček ve spektru dané hlásky. K analýze formantů využívá LPC (Linear Predictive Coding), což je kódovací algoritmus ztrátové komprese pro zajištění spektrální obálky digitalizovaného hlasového signálu v komprimované formě. Je to jeden z nejefektivnějších způsobů analýzy a zpětné konstrukce hlasového signálu. LPC při provádění analýzy vychází z předpokladu, že hlas je generován zdrojem (hlasivkami), kde na hlasivkové štěrbině závisí průběh a intenzita generovaných vokálů a na krku a ústech, které tvoří jakýsi tubus, který pak způsobuje různé rezonance, což dává hlasu jeho charakteristické zabarvení pro každého člověka. Přesné fungování včetně použitých vzorců je mimo rámec této práce, ale naleznete je na [26].

## 7 Praktická část

### 7.1 Úvod do praktické části

Jak již bylo řečeno v úvodu, cílem diplomové práce je vytvořit program, který by rozeznal formanty v samohláskách, a potvrdit, že teorie formantů při rozpoznávání samohlásek je správná. Tedy, že jednotlivé samohlásky jsou určeny oblastmi posílených frekvencí. Především bych chtěla vyřešit problém “maskování” formantů za nosnou frekvenci samohlásky. Abych mohla program aplikovat, pořídila jsem nahrávky samohlásek deseti mužů a deseti žen širokého věkového spektra (18 – 70 let). Nahrávky jsou pořízeny ve formátu “.wav” na mikrofon Genius MIC-01A nahrávacím softwarem Sony Sound Forge 10.0.367 Trial version. Od všech mužů i žen mám nahrávky samohlásek i, e, a, o, u, na frekvenci, na které každý přirozeně mluví. U šesti mužů a pěti žen mám nahrávky samohlásek také intonované tj. každý zpívá ve své poloze danou samohlásku po púltónech přes celou jednočárkovanou oktávu a tím dostávám 12 tónů od každé samohlásky. Jednočárkovaná oktáva je od “c” (261,5 Hz) do “h” a obsahuje komorní A (440Hz). Celá oktáva byla předehrávána na klávesách (keyboardu) typu Casio CTK 530. Celkem mám 72 (tj. 6 x 12) intonovaných nahrávek každé samohlásky od mužů a 60 (tj. 5 x 12) intonovaných nahrávek každé samohlásky od žen. Nahrávky jsou ořezány v sofwaru Sound Forge na délku 1s a uloženy pro další užití opět ve formátu “.wav”. Jednotlivé intonované samohlásky upravené na délku trvání 1s jsem z praktických důvodů pojmenovala i1 – i12, e1 – e12, a1 – a12, o1 – o12, u1 – u12 pro každou osobu zvlášť. Všechny použité nahrávky naleznete na CD příloze.



Obrázek 12: Průběh tlaku vzduchu před ústy při hláске “i”.

Pokud se zaměřím na rozpoznání jednotlivých samohlásek na základě grafického znázornění jejich spektra je potřeba nahranou hlásku upravit. Na Obrázku 12 můžeme vidět záznam průběhu tlaku vzduchu (nahraný při hlásce “i” na mikrofon před ústy), kde žádné posílení frekvence nebo jinou nepravidelnost nepoznáme. Na vodorovné ose je čas v sekundách (zobrazuje se tedy průběh tlaku po dobu dvacetiny vteřiny) a na svislé ose je relativní tlak vzduchu.

Abych mohla jednotlivé samohlásky od sebe rozpoznat, je třeba provést rychlou verzi diskrétní Fourierovy transformace původního signálu.

## 7.2 Software a použité m-fily

Pro nahrávání samohlásek jsem použila software Sony Sound Forge 10.0.367 Trial version a veškeré programy (m-soubory tj. m-fily) byly vytvořeny v softwaru Matlab Version 7.0. První m-file nazvaný “read” (viz níže), kterým jsem upravila původní nahrávky, načte danou samohlásku, smaže z ní vše, co je nad jednu sekundu a funkcí Matlabu “fft” provede rychlou verzi diskrétní Fourierovy transformace. Funkce Matlabu “fft” mi dá diskrétní Fourierovu transformaci vektoru  $x$  (původního signálu) délky  $n$  spočítanou pomocí rychlé Fourierovy transformace

$$X(k) = \sum_{j=1}^n x(j)\omega_n^{(j-1)(k-1)},$$

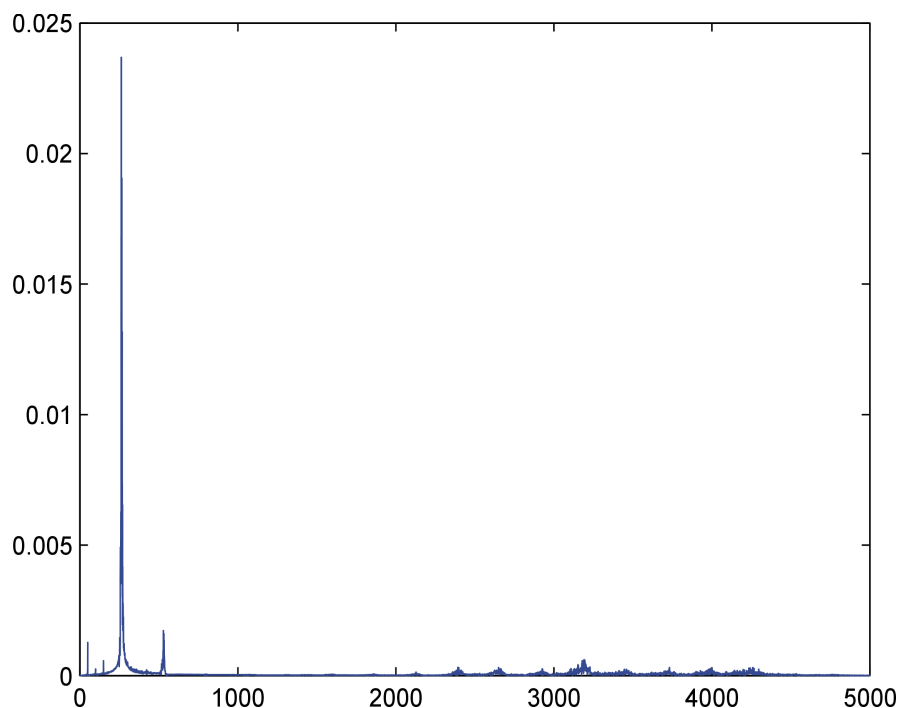
kde  $\omega_n = e^{\frac{-2\pi i}{n}}$  je  $n$ -tý kořen. Přičemž Matlab využívá Cooley-Tukey algoritmus. Dále m-file “read” provede normalizaci spektra a nakonec spektrum vykreslí pomocí funkce Matlabu “plot”. M-file “read”:

```
function [output, sound] = read(x);
origin=wavread(x);
origin = origin(:,1);
m=length(origin);
origin(44100:m)=[];
four=fft(origin);
power=four.*conj(four)/44100;
power(1)=[];
N=5000;
output=sqrt(power);
tot = sum(output);
output = output/tot;
```

Použitím m-filu “read” tedy dostávám spektrum dané samohlásky. Jeho příklad vidíme na spektru samohlásky “i” na Obrázku 13. Tuto konkrétní samohlásku “i” jsem vybrala záměrně, protože má základní frekvenci právě v oblasti formantu. Na ose x je frekvence v Hz a na ose y intenzita.

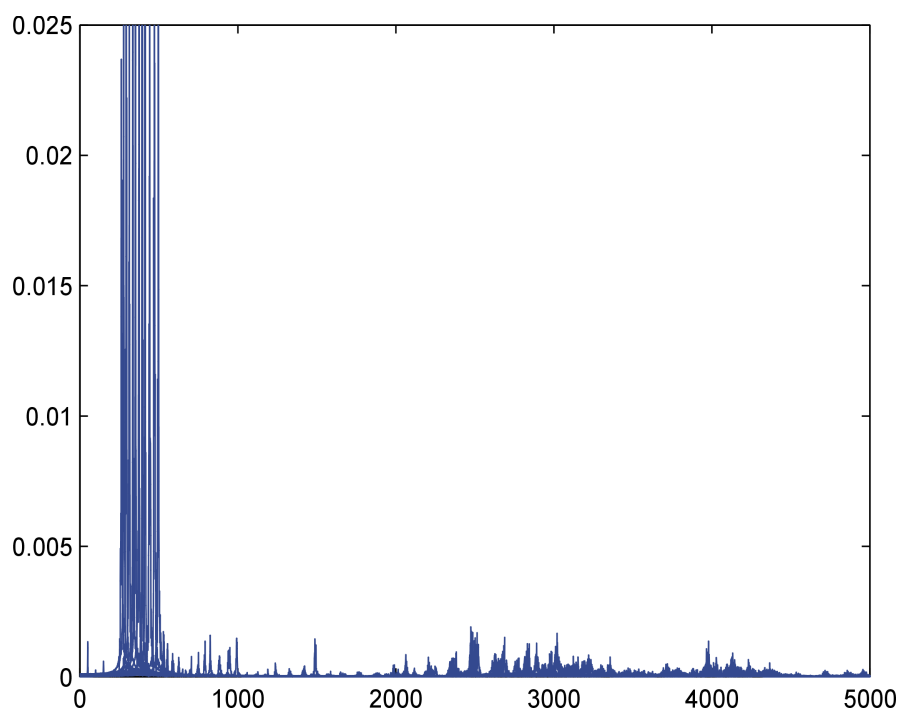


**Poznámka 9.** Důvodem proč analyzuji 12 samohlásek po půltónech je maskování formantů. Může se stát, že samohláska (na dané frekvenci) nemá v oblasti formantu žádnou energii např. pokud řeknu hlásku “i” na frekvenci 600 Hz, není v oblasti prvního formantu (tj. 300–450 Hz) žádná část spektra, která by mohla být posílena. Naopak výrazné spektrální maximum je na 600 Hz, tedy mimo oblast prvního formantu. Oblasti formantů jsou takto “maskovány” nosnou frekvencí a proto k jejich určení potřebujeme součet hlásek na více různých frekvencích. Aby vynikly oblasti posílené frekvence každé samohlásky, nechala jsem opakovaným použitím m-flu “read” vykreslit všechny intonované samohlásky (zvlášť i, e, a, o, u) od jedné osoby do jednoho grafu. Pro ukázkou uvádím graf samohlásky “i” na Obrázku 14, kde jsou již oblasti zesílených frekvencí patrné, ale abych měla přesnější a lépe viditelný výsledek, sestavila jsem m-file, který nám jasněji tyto oblasti určí.



Obrázek 13: Spektrum samohlásky “i” použitím m-flu “read”.

Myšlenka, kterou chci nyní realizovat spočívá ve vytvoření “obecné samohlásky”, kterou dostanu zprůměrováním všech nahraných samohlásek od dané osoby. Takovou obecnou samohlásku použiji jako “obecné pozadí”, na které budu zobrazovat konkrétní samohlásku (i, e, a, o, u) od stejné osoby. Všechny intonované samohlásky od jedné osoby načtu přes m-file “read” do matice “P” tak, že po řádcích jsou uloženy spektra samohlásek  $i_1 - i_{12}$ ,  $e_1 - e_{12}$ ,  $a_1 - a_{12}$ ,  $o_1 - o_{12}$ ,  $u_1 - u_{12}$ . Po sloupcích v této matici sečtu nediferencovaný zvuk a znormuji sumou těch součtů sloupců matice “P”. Tuto obecnou normovanou samohlásku zobrazím jako pozadí v grafu, na které nechám vykreslit znormovanou (opět sumou těch sečtených sloupců matice “P”) samohlásku (zvlášť i, e, a, o, u), která bude tvořena

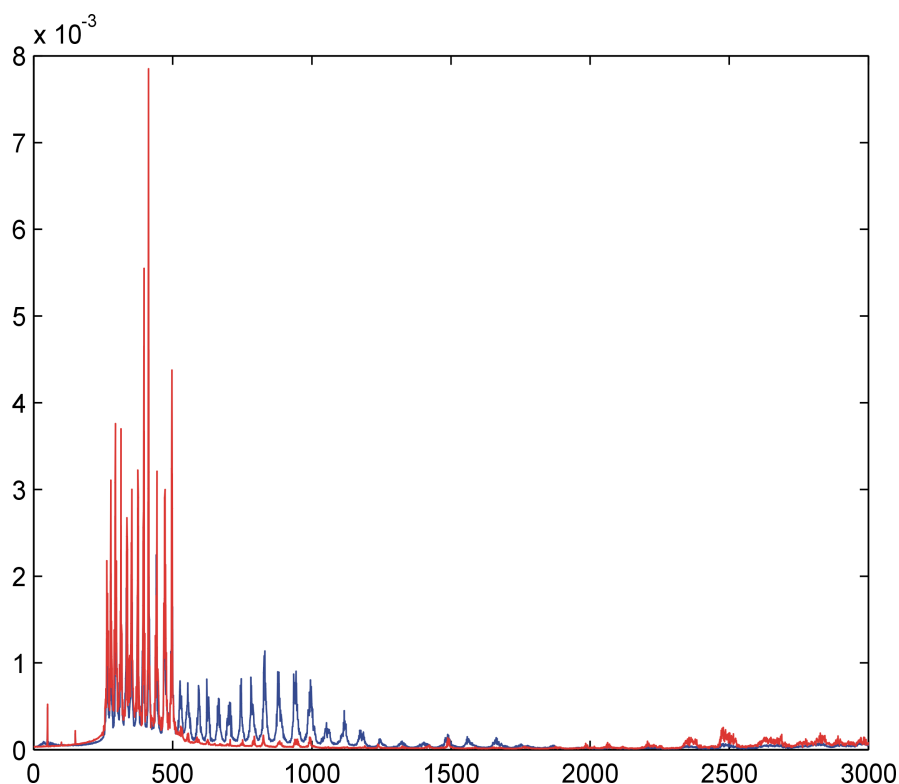


Obrázek 14: Spektrum všech intonovaných “i” použitím m-filu “read”.

všemi intonovanými hláskami od dané osoby. M-file “compare”, který naleznete v CD příloze ve složce “m-fily”, vykreslí celkem 5 grafů, kde v každém bude obecná samohláska jako modré pozadí a přes ní bude červeně vykreslená znormovaná samohláska i, e, a, o, u (zvlášť). Jeden z těchto grafů pro ukázkou uvádím na Obrázku 15. Všechny použité m-fily potom naleznete v Příloze m-fily.

Pro zobrazení oblastí formantů použiji vygenerovaná data obecných samohlásek tj. uni, une, una, uno, unu (pojmenování podle anglického “universal” + konkrétní samohláska), součet “s” a matici “P”. Data načtu a čtyřmi níže uvedenými m-fily vytvořím grafy, kde bude zobrazena obecná samohláska jako pozadí. Původně jsem hledala formanty tam, kde spektrum dané samohlásky (její průměr) přesáhlo spektrum této obecné samohlásky (průměr všech samohlásek). To ale není nejrozumnější, protože v tom obecném pozadí jsou oblasti formantů také posíleny. Proto jsem pozadí proložila exponenciálou a tím vytvořila jakýsi model spektra vokálu těsně nad hlasivkami. Potom budu hledat oblasti, kde daná samohláska přesáhla tuto exponenciálu.

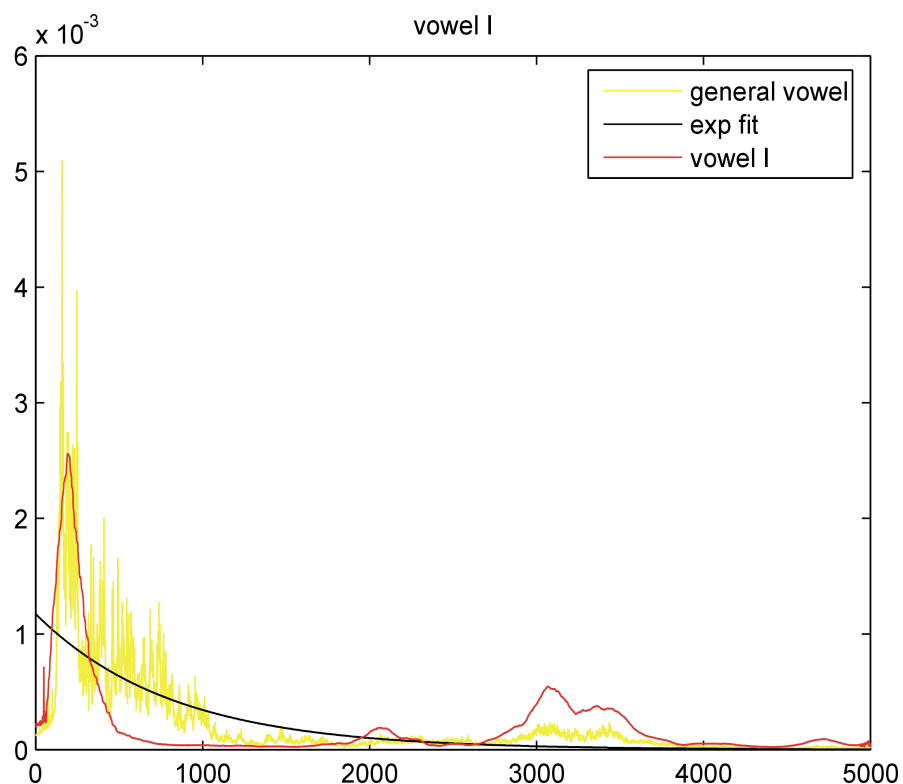
Mám tedy pozadí tvořené obecnou samohláskou, která je proložená exponenciálou, a přes toto pozadí bude vykreslena konkrétní samohláska vyhlazená pomocí filtru viz níže. Filtr je tvořen samostatným m-filem “filter” a vyhlazuje samohlásku pomocí klouzavých průměrů, přičemž šířka okna těchto klouzavých průměrů je v určování oblastí formantů zásadní. Je zde důležité, jak zvolím délku okna, protože pokud by jeho délka byla příliš velká, filtr vyhladí samohlásku tak, že oblasti formantů se začnou “slévat” do jednoho. Na druhou stranu, když



Obrázek 15: Spektrum obecné samohlásky jako modré pozadí s červeným spektrem “i” přes ni.

zvolím délku okna malou, vokál nebude dobře vyhlazen a oblasti formantů se budou objevovat i na frekvencích, kde by pro konkrétní samohlásku neměly být. Samotné grafy vykreslí m-file “findfor” (angl. najdi formant) s využitím pomocných m-filů “expa”, “filter” a “draw”. M-file “expa” vykreslí exponenciálu a “filter” vyhladí spektrum dané samohlásky. M-file “draw” vše zobrazí do jednoho grafu. M-file “findfor” potom načte data, která jsou již vytvořena pomocí m-filu “compare” a proloží obecnou samohlásku exponenciálou. Na konkrétní samohlásku aplikuje zhlazovací filtr a vše vykreslí do jednoho grafu. Výsledek pro samohlásku I můžeme vidět na Obrázku 16. Tam kde spektrum samohlásky přesáhne exponenciálu určí formanty – konkrétně na Obrázku 16 jsou to oblasti 100 – 300 Hz, 2000 – 2100 Hz a 2900 – 3900 Hz. Grafy ostatních samohlásek a všech osob jsou v Obrazové příloze, kde každý graf zobrazuje jednu samohlásku konkrétní osoby, a také na CD příloze ve složce Grafy.

Formanty se tradičně klasifikují jen podle frekvence (lépe řečeno podle rozsahu frekvencí, na kterých se formant vyskytuje). Jak je ale vidět na grafech pro samohlásky “o” a “u” v Obrazové příloze (tj. grafy 7.1 – 7.6, 8.1 – 8.5, 9.1 – 9.6, 10.1 – 10.5), často bývá rozdíl v jejich relativní síle či šířce pásma než v poloze středu jejich pásma. Proto navrhuji formant charakterizovat třemi čísly, která určí střed oblasti formantu, jeho šířku a výšku. Střed označíme  $F_1$  a bude uveden v Hz stejně jako šířka formantu označená  $F_2$ . Jeho výška  $F_3$  měří přesah vrcholu

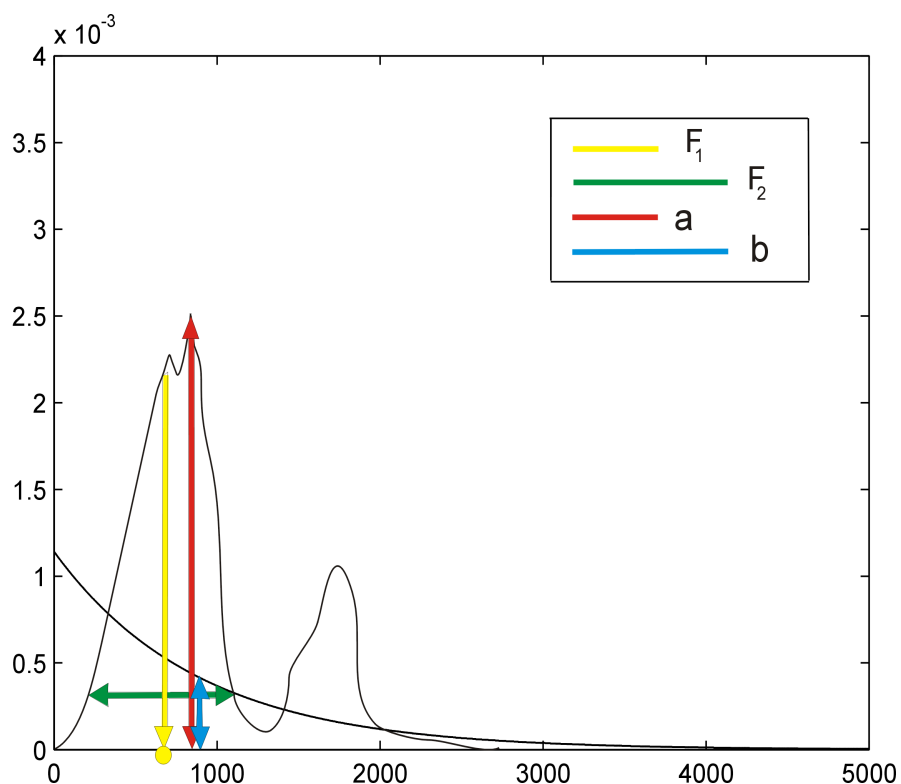


Obrázek 16: Spektrum obecné samohlásky jako žluté pozadí, proložené exponenciálou s červeným vyhlazeným spektrem “i”.

spektra ve formantu nad proloženou exponenciálou a je vyjádřena poměrem  $\frac{a}{b}$ , kde  $a$  je výška spektra v oblasti formantu měřená od osy  $x$  po maximum spektra tohoto formantu a  $b$  je výška exponenciály ve stejném bodě. Grafické znázornění odčítání hodnot z grafu můžete vidět na Obrázku 17. Žlutá šipka směřuje do bodu na ose  $x$ , kde je  $F_1$ . Rozsah, kde měříme  $F_2$ , je zobrazen zelenou šipkou a konečné hodnoty  $a$  a  $b$  jsou znázorněny červenou a modrou šipkou. Tyto charakteristiky jsem postupně odečetla ze všech grafů, kterých je celkem 55 (6 mužů + 5 žen krát 5 samohlásek) a jsou k dispozici pro porovnání v Obrazové příloze.

### 7.3 Výsledné hodnoty

Výsledky, ke kterým jsem dospěla, jsou uvedeny v následujících tabulkách. Jsou rozděleny na muže a ženy, což je vyznačeno v levém horním rohu M/Ž, a podle samohlásek, kterých se týkají – I, E, A, O, U – označeno tamtéž. V prvním sloupci jsou osoby popsány pořadovým číslem (1–6 muži, 1–5 ženy). Následující sloupce již obsahují naměřené hodnoty daných formantů. Jak jsem již zmínila, každý formant je popsán třemi čísly (v dolním indexu) a číslo formantu najdete v horním indexu – tedy značení  $F_1^1$  znamená první charakteristiku prvního formantu,  $F_3^2$  třetí charakteristiku druhého formantu atp. Hodnoty v tabulkách jsem odečítala



Obrázek 17: Grafické znázornění odčítání hodnot z grafu.

přímo z grafu v Matlabu, přičemž první dvě charakteristiky jsou zaokrouhleny na 50 Hz a třetí charakteristika na jedno desetinné místo. Je třeba zdůraznit, že umístění formantu tj.  $F_1$ , je středem oblasti, kde spektrum přesáhlo exponenciálu, nikoli bodem, kde je spektrum nejvyšší. Abych mohla spočítat “průměrné hodnoty” pro každou charakteristiku, bylo by potřeba mít více dat, protože na počet vzorků, které mám k dispozici, není vhodné aplikovat jakoukoli statistickou metodu – výsledky by neměly vypovídající hodnotu. V případech, kde spektrum nepřekročilo danou exponenciálu, jsou místa výsledkových tabulek prázdná. U samohlásky U je tabulka zkrácena jen na dva formanty, protože žádné další se již neprojeví (přestože jsem sledovala spektra až do 5000 Hz).

M I	$F_1^1$	$F_2^1$	$F_3^1$	$F_1^2$	$F_2^2$	$F_3^2$	$F_1^3$	$F_2^3$	$F_3^3$
1	200	350	2,4	2150	350	3,4	2750	450	6,2
2	300	350	2,2	2600	1050	3,4	3650	550	12,2
3	300	300	3	2200	350	2,3	3300	1150	5
4	250	350	2,8	2600	800	3,2	3650	850	6,3
5	200	250	2,7	2100	350	2	3300	1050	20,3
6	200	350	2,4	2550	950	4,6	4550	600	22,6

Tabulka 3: Výsledné hodnoty pro muže a samohlásku I.

Ž I	$F_1^1$	$F_2^1$	$F_3^1$	$F_1^2$	$F_2^2$	$F_3^2$	$F_1^3$	$F_2^3$	$F_3^3$
1	400	350	4,5	2900	1000	1,8	4100	600	2
2	400	350	3,6	2900	600	1,9	4450	600	3,7
3	400	350	3,5	2500	500	2,4	4050	550	3,4
4	400	350	3,5	2900	850	2,1	4150	750	5,1
5	400	350	2,8	2600	500	1,7	3900	550	3,8

Tabulka 4: Výsledné hodnoty pro ženy a samohlásku I.

M E	$F_1^1$	$F_2^1$	$F_3^1$	$F_1^2$	$F_2^2$	$F_3^2$	$F_1^3$	$F_2^3$	$F_3^3$
1	500	300	2,2	1750	350	2,2	2600	600	3,6
2	650	350	2,6	1550	350	2,7	2950	400	4,2
3	600	350	2,7	1600	250	1,7	2850	350	4,4
4	550	450	2,6	-	-	-	-	-	-
5	550	250	2	1600	450	1,8	2300	500	3,1
6	600	350	1,6	1700	450	2,2	2700	500	3

Tabulka 5: Výsledné hodnoty pro muže a samohlásku E.

Ž E	$F_1^1$	$F_2^1$	$F_3^1$	$F_1^2$	$F_2^2$	$F_3^2$	$F_1^3$	$F_2^3$	$F_3^3$
1	700	750	2,5	1600	350	1,6	2900	550	1,6
2	850	450	2,5	-	-	-	-	-	-
3	850	300	2,2	1950	450	3	3250	600	1,9
4	650	750	1,7	1600	300	1,6	2900	450	1,9
5	850	450	1,8	-	-	-	3000	400	2,4

Tabulka 6: Výsledné hodnoty pro ženy a samohlásku E.

M A	$F_1^1$	$F_2^1$	$F_3^1$	$F_1^2$	$F_2^2$	$F_3^2$	$F_1^3$	$F_2^3$	$F_3^3$
1	650	450	2,3	-	-	-	2950	400	2,7
2	650	350	2,7	1100	300	2,1	3100	400	3
3	650	400	2,6	1050	300	1,7	-	-	-
4	600	500	2,5	1150	200	2,8	2600	200	1,7
5	600	250	1,8	950	200	3,2	2600	250	2,8
6	700	350	2	1150	250	2,8	-	-	-

Tabulka 7: Výsledné hodnoty pro muže a samohlásku A.

Ž A	$F_1^1$	$F_2^1$	$F_3^1$	$F_1^2$	$F_2^2$	$F_3^2$	$F_1^3$	$F_2^3$	$F_3^3$
1	950	550	2,2	-	-	-	3300	550	1,9
2	900	550	2,8	1450	200	2	-	-	-
3	900	450	3	1400	350	2,2	3500	650	2
4	900	350	2,9	1450	300	2,3	3050	600	2,6
5	1000	550	3,5	-	-	-	3350	500	2

Tabulka 8: Výsledné hodnoty pro ženy a samohlásku A.

M O	$F_1^1$	$F_2^1$	$F_3^1$	$F_1^2$	$F_2^2$	$F_3^2$	$F_1^3$	$F_2^3$	$F_3^3$
1	350	450	3	-	-	-	-	-	-
2	500	350	4	900	300	2	3500	1000	4,6
3	550	350	3	900	300	2,8	-	-	-
4	400	550	3	950	200	1,9	3350	400	3
5	600	500	2,1	-	-	-	3000	450	5,4
6	450	300	2,2	900	200	2,5	-	-	-

Tabulka 9: Výsledné hodnoty pro muže a samohlásku O.

Ž O	$F_1^1$	$F_2^1$	$F_3^1$	$F_1^2$	$F_2^2$	$F_3^2$	$F_1^3$	$F_2^3$	$F_3^3$
1	450	350	2,2	900	300	2,6	-	-	-
2	400	300	1,9	950	350	3	-	-	-
3	450	400	2,2	800	200	2,3	3250	500	2,3
4	450	400	2	900	300	2,3	-	-	-
5	500	400	2,3	950	300	3	-	-	-

Tabulka 10: Výsledné hodnoty pro ženy a samohlásku O.

M U	$F_1^1$	$F_2^1$	$F_3^1$	$F_1^2$	$F_2^2$	$F_3^2$
1	200	400	2,6	-	-	-
2	300	450	3,7	800	150	1,5
3	350	400	3	850	350	1,8
4	300	450	3,6	-	-	-
5	300	400	3,2	750	200	1,5
6	300	400	3	700	200	1,9

Tabulka 11: Výsledné hodnoty pro muže a samohlásku U.



Ž U	$F_1^1$	$F_2^1$	$F_3^1$	$F_1^2$	$F_2^2$	$F_3^2$
1	400	350	4,3	850	250	1,5
2	450	350	4,2	900	350	3,1
3	350	350	4,7	900	250	1,8
4	400	400	4,1	850	250	1,5
5	450	450	4,1	900	250	2

Tabulka 12: Výsledné hodnoty pro ženy a samohlásku U.

## 7.4 Komentář výsledků

Co tedy mohu říci na základě naměřených hodnot v tabulkách 3 – 12? Nejprve se podívám na rozdílnost mužských a ženských formantů. Je obecně známo, že ženy mají formanty umístěny výš než muži, a to cca o 200 Hz viz [12]. Toto pravidlo mohu potvrdit, neboť když se podíváme na hodnoty  $F_1^i$  výsledkových tabulek a porovnáme je u jedné samohlásky mezi muži a ženami, opravdu jsou ženské formanty posunuty výš (pozor – je potřeba myslet na to, že hodnota  $F_1^i$  nám udává střed daného formantu nikoli rozmezí, kde se formant vyskytuje). U konkrétních samohlásek můžeme pozorovat:

- I: formanty žen jsou o 200 Hz výš než formanty mužů, přičemž se stoupajícím formantem tento rozdíl ještě roste
- E: první formant žen je výš také o 200 Hz, ale druhý formant mají ve stejné oblasti, třetí už potom opět o 200 Hz výš u žen než u mužů
- A: u této samohlásky jsou dokonce všechny ženské formanty o 300 Hz výš než mužské
- O: zde jsou formanty mužů a žen kupodivu stejné
- U: ženské formanty jsou o 100 Hz výš než mužské.

Zajímavá odlišnost mezi mužskými a ženskými formanty je i u druhé charakteristiky  $F_2$  – ženy mají všechny formanty cca o 50 Hz širší než muži, a to u všech samohlásek. Je třeba ale podotknout, že pokud bych měla dostatečné množství vzorků, mohl by tento rozdíl zaniknout (či naopak se posílit). Na druhou stranu není zřejmá odlišnost mezi muži a ženami v třetí charakteristice, tj. v poměru výšky spektra v oblasti formantu a výšky exponenciely (můžeme se na to dívat jako na sílu překročení spektra nad danou exponencielu). Snad jen u samohlásky I, kde rozdíl poměrů u prvního formantu je 1. Jinak u všech samohlásek jsou si hodnoty  $F_3^i$  pro  $i$ -tý formant podobné. To nelze říci o rozdílech mezi samohláskou O a U, ty mají sílu formantů  $F_3^1$  opravdu odlišnou a vzhledem k tomu, že jejich oblasti formantů  $F_1^i$  jsou velmi podobné, je tento rozdíl především u prvního

formantu vcelku zásadní při určení, zda se jedná o formant hlásky O či U. Tento jev je dobře viditelný hlavně v grafech (viz Obrazová příloha).

Když nyní porovnám praktické výsledky s obecně platnými hodnotami formantů v Tabulce 1, je vidět, že je rozumné rozlišovat formanty pro muže a ženy zvlášť už z důvodu posunutí ženských formantů o 200 Hz výš, jak již bylo uvedeno. Pro porovnávání výsledků s hodnotami v Tabulce 1 je třeba z výsledkových tabulek 3 – 12 srovnatelné hodnoty určit, protože charakteristika  $F_1^i$  nám udává střed formantu a  $F_2^i$  jeho šířku narozdíl od hodnot Tabulky 1, kde jsou tyto dvě hodnoty obsaženy v jednom rozsahu Hz. Tedy pro porovnání jsem použila pro dolní hranici formantu (jak je uváděna v Tabulce 1) výpočet  $F_1^i - \frac{F_2^i}{2}$  a pro horní hranici  $F_1^i + \frac{F_2^i}{2}$ . U konkrétních samohlásek mohu o výsledných hodnotách říci:

- I: první a druhý formant odpovídá hodnotám v Tabulce 1 (pokud posuzuji muže i ženy dohromady), třetí formant bych potom mohla definovat v rozmezí 3000 – 4000 Hz pro muže a 3700 – 4500 pro ženy, zde můžeme opět vidět rozdíl v posunutí ženských formantů výš v porovnání s mužskými
- E: u této samohlásky se dle hodnot ve výsledkových tabulkách oblast formantu lehce rozšířila a to na 400 – 1000 Hz pro muže i ženy dohromady, v porovnání s hodnotami v Tabulce 1 a třetí formant se objevuje v rozmezí 2400 – 3000 Hz pro muže a 2650 – 3550 Hz pro ženy (opět se ukazuje rozdíl cca 200 Hz v posunutí formantů ženských nad mužské)
- A: první formant této samohlásky je ve výsledcích hodně posunutý dolů, pro muže se vyskytuje v oblasti 400 – 900 Hz, pro ženy mezi 650 – 1250 Hz, což může být způsobeno malou strmostí exponenciely, druhý formant pro muže je posunutý dolů na 900 – 1200 Hz u žen naopak o 100 Hz nahoru oproti Tabulce 1
- O: spodní hranice prvního formantu je ve srovnání s obecně známými hodnotami o 250 Hz níž pro muže i ženy, horní hranice je stejná, tj. 700 Hz a hodnoty druhého formantu vcelku odpovídají hodnotám Tabulky 1
- U: rozsah prvního formantu ve výsledkových tabulkách vcelku odpovídá hodnotám v Tabulce 1 i když pro některé muže začíná první formant už na 100 Hz, což je o 200 Hz níž než obecně známé hodnoty, rozmezí druhého formantu je stejné jako v Tabulce 1.

**Poznámka 10.** Ve výsledných grafech se projevily také čtvrté formanty, které z důvodu přehlednosti nejsou zahrnuty v tabulkách výsledků, protože se objevily jen zřídka. Ale stojí za povšimnutí u samohlásky A, kde se objevily u čtyř mužů a tří žen. Jejich charakteristiky vyšly celkem přesně: průměr  $F_1^4$  pro muže je 3300 – 3500 Hz,  $F_2^4$  je 550 Hz a poměr  $F_3^4$  vychází 6,1. Pro ženy potom  $F_1^4$  v mezích 4000 – 4200 Hz,  $F_2^4$  je 600 Hz a poměr  $F_3^4$  3,3. Dal by se tedy na základě těchto výsledků charakterizovat čtvrtý formant  $F^4$  pro samohlásku A.

Když mluvím o formantech jako posílených frekvencích každé samohlásky, je zajímavá také myšlenka tzv. antiformantů. Mohly by se takto označit oblasti oslabených frekvencí. Pokud se podíváme na grafy formantů mužů i žen samohlásky O v Obrazové příloze, je zřejmé, že oblast mezi 1500 – 2500 Hz je zcela oslabena. Stejná situace je vidět i na grafech mužů a žen u samohlásky I, zde je tato oblast mezi 1500 a 2000 Hz. Jinými slovy by v těchto oblastech mohl být definován antiformant, ve kterém je frekvence potlačena. Posouzení z tohoto hlediska by mohlo být tedy předmětem dalšího zkoumání a diskuze.

#### 7.4.1 Komentář grafů

Dalším zajímavým výsledkem je pohled na grafické znázornění v Obrazové příloze. Na každé její straně jsou grafy pro jednotlivé osoby jako na Obrázku 16 s tím, že každá strana je pro jednu samohlásku a zvlášť pro muže a ženy. V grafech si můžete povšimnout, že osa y je zobrazena pro každou osobu do jiné výšky resp. jiných hodnot osy y (např.  $4 \cdot 10^{-3}$ ,  $6 \cdot 10^{-3}$ ), ale pro všechny samohlásky jedné osoby jsou tyto hodnoty stejné. Na odčítání charakteristik tato skutečnost nemá vliv.

Když se podívám na grafy samohlásky I, je vidět, co se formantů týče, že je tato samohláska velmi stabilní, tj. přestože ženy zpívaly o oktávu výš než muži, je zřejmé jak graficky tak podle výsledkových tabulek, že formanty jsou ve stejných oblastech. Pro tuto samohlásku je charakteristický výrazný formant v oblasti 300 – 400 Hz a potom další až kolem 2500 Hz, což právě na grafech je dobře vidět. Bylo by také zajímavé zjistit, zda se ve formantech může projevit odlišnost mezi moravským a českým I a v čem vlastně spočívá podstata jejich různé znělosti.

U samohlásky E je dobře vidět, jak se projevuje již zmíněné maskování formantů – když se podíváme na grafy mužů 3.1, 3.2, 3.5 a 3.6, tak v oblasti kolem 250 Hz se zobrazuje lokální maximum (vrchol spektra), které svádí k tomu, považovat jej za formant, ale je to jen posílení základní frekvence, které není formantem. Stejně dobře je tento jev vidět i na grafech 4.2, 4.3, a 4.5 u žen. Řešení vidím v ořezu spektra samohlásky, protože exponenciální regresi počítám z celého spektra 0 – 5000 Hz, což je problém, protože vše, co je ve spektru pod frekvencí nejnižšího tónu je artefakt. Nejnižším tónem rozumíme pro muže 120 Hz a pro ženy 240 Hz. Proto při exponenciální regresi by se měla těmto artefaktům dát nulová váha nebo spektrum ořezat tak, aby hodnoty pod nejnižším tónem neměly na tvar exponenciely vliv. Tím by se vyřešil přesah spektra samohlásky nad exponencielu na začátku grafu, kde nejsou formanty, protože exponenciela by byla v této části grafu strmější. Zajímavé je, že v grafu 3.4 se jiný než první formant neprojevil, což může být zapříčiněno barvou hlasu této osoby a narozdíl od ostatních mužských grafů je zde znatelný čtvrtý formant okolo frekvence 3500 Hz. Grafy 4.1 a 4.4 tj. ženy 1 a 4 jsou téměř stejné a je tomu tak i u ostatních samohlásek – hlasy zde patří sourozencům, je tedy možný vliv genetiky dané barvy hlasu?

Při pohledu na grafy samohlásky A vyniká oblast oslabení v rozmezí 1500 – 2500 Hz a to u mužů i žen (mohli bychom zde zjišťovat existenci antiformantu) a opět se zde projevuje maskování formantů okolo frekvencí 300 – 400 Hz. U této samohlásky je patrné tzv. slévání formantů, které jsme diskutovali při volbě délky okna klouzavých průměrů při vyhlazování samohlásky filtrem. Tento jev by se dal tedy odstranit vhodnější volbou délky tohoto okna. Je zajímavé, že grafy u samohlásky A jsou jakoby dvou typů – průběh spektra na grafech 5.2, 5.3, 5.4, 5.6 a 6.3 je vizuálně podobný narozdíl od grafů 5.1, 5.5, 6.1, 6.2, 6.4 a 6.5.

Samohlásky O a U bych chtěla porovnat společně. Jak jsem již zmínila, jejich oblasti středů formantů  $F_1^i$  jsou téměř stejné, ale právě na jejich grafickém znázornění je patrná odlišnost mezi výškami jejich spektra – nejlépe je to vidět při porovnání grafů O a U u žen. Oblasti posílení jsou stejné, ale u samohlásky U první formant výrazně přesahuje danou exponenciálu narozdíl od O. U mužů tento rozdíl není tak patrný, tam je lépe viditelné mírné posunutí prvního formantu samohlásky U blíž k nule narozdíl od O. U druhého muže na grafu 7.2 je dobře vidět přesah spektra nad exponenciálu v oblasti okolo 3000 – 3500 Hz, což je charakteristické pro konkrétní osobu a je to způsobeno barvou hlasu.

## 8 Komplikace a zajímavosti při zpracování

Nahrávání intonovaných samohlásek bylo trochu obtížné z toho důvodu, že dnes málokdo bez tréninku zpěvu umí intonačně zvládnout a zazpívat celou oktávu po půltónech, přestože každý zvuk byl předehráván na klávesách. Je tedy možné, že výsledky mohly vyjít odlišně, pokud by byly tóny zpívány zcela čistě podle předehrané oktávy. Ale na druhou stranu, čím různější frekvence mám nahrané, tím je větší pravděpodobnost odhalení maskování formantů.

U dvou mužů a jedné ženy se stalo, že poslední nejvyšší hlásku vůbec nevyzpívali - v tom případě jsem místo ní (např. a12) nakopírovala tutéž samohlásku nejvyšší, která byla od dané osoby k dispozici, tj. a11, přepsala jsem ji na a12 a tak s ní i nadále pracovala. Na zpracování formantů by však tato “úprava” neměla mít vliv. Otázkou je, jaký vliv na nahrávky by mělo použití profesionálního mikrofону a techniky, zda by se tím projevil další odlišnosti mezi samohláskami.

Při odčítání hodnot z grafu 9.1 je viditelný druhý formant, který nepřesáhl exponenciálu, ale přesto jsem jej zahrnula do naměřených hodnot.

## 9 Závěr

Základem každé práce s daty je dostatečné množství vzorků, na kterých charakteristiky sledujeme. Bylo by tedy samozřejmě zapotřebí zkoumat formanty na mnohem širším spektru dat, a to v řádu 100. Dalším důležitým předpokladem

výsledků s vypovídající hodnotou je použití dobré statistické metody, která je pro konkrétní účel vhodná. Při zpracování naměřených dat navrhuji, aby při dalším zpracování výsledků k tomuto tématu byla využita vhodná statistická metoda.

Na základě uvedených výsledků mohu potvrdit, že každá samohláska je určena svými formanty a jejich oblasti se podařilo potvrdit. O zpřesnění těchto oblastí lze hovořit jen v souvislosti s rozdělením formantů na mužské a ženské. Výsledky tedy potvrdily, že odlišnost mezi mužskými a ženskými formanty je znatelná a je rozumné charakterizovat je zvlášť pro muže a ženy, neboť jejich rozdílnost je kolem 200 Hz.

Podařilo se vyřešit problematiku maskování formantů a to díky rozsahu použitých frekvencí, na kterých byly samohlásky zpívány. Otázkou stále zůstává, zda jsme schopni na základě zobrazeného grafu spektra konkrétní samohlásky určit, o kterou samohlásku se jedná. Jistě je ale tato metoda na hledání formantů úspěšně použitelná a je možno problematiku tímto směrem dále zkoumat. Také by bylo vhodné při dalším zpracování použít na nahrávání profesionální mikrofon. Tato práce je tedy spíše nastíněním dané problematiky a základem pro její další podrobný výzkum.

Databáze samohlásek je k dispozici na CD příloze. Je možno ji využít pro další zpracování nebo jen pro posouzení dat, ze kterých výsledky vyšly.

Plynulým pokračováním této práce by mohlo být především detailnější zpracování většího množství vzorků a mohla by být základem nebo alespoň nastíněním využití Fourierovy transformace pro sestavení programu, který by na základě nalezených formantů uměl samohlásku určit. Mohli bychom se také zaměřit na rozdíly mezi samohláskami osob, které s hlasem pracují profesionálně (zpěváci, moderátoři apod.), zda jejich samohlásky budou mít stejné charakteristiky na rozdíl od lidí s neškoleným hlasem. V tomto směru je známa teorie pěveckého formantu viz [21], ale mohly by se ukázat i další odlišnosti.

## Reference

- [1] Benson, D.: *Music: a Mathematical Offering, Department of Mathematical Sciences, University of Aberdeen, Scotland, UK, 2007*
- [2] Brigham, E. O.: *Fast Fourier transform and its applications. 2nd ed. Prentice-Hall, Inc., Engelwood Clifs, New Jersey 1987*
- [3] Duběda, T.: *Jazyky a jejich zvuky: univerzálie a typologie ve fonetice a fonologii. Praha, Karolinum, 2005. ISBN: 80-246-1073-6 (brož.)*
- [4] Grepl, M., Fürst, T., Pešák, J.: *The F1-F2 Vowel Chart for Czech Whispered Vowels a, e, i, o, u. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2007, 151(2):353-356*
- [5] Hardy, G. H.: *Fourierovy řady. Praha 1971*
- [6] Herrmann, L.: *Fourierovy řady (Komentované přednášky), Vydavatelství ČVUT, Praha 2002*
- [7] Kolmogorov, A. N.: *Une série de Fourier-Lebesgue divergente partout. 1926, Rend. Acad. Sci. 178, 303-305. Jahrbuch database (Zbl): 52.0269.02*
- [8] Kopáček, J.: *Matematická analýza pro fyziky (IV), Matfyzpress, Univerzita Karlova v Praze, Druhé vydání, Praha 2003*
- [9] Krčmová, M.: *Úvod do fonetiky a fonologie pro bohemisty. Ostravská univerzita v Ostravě, 2007. ISBN: 978-80-7368-405-1 (brož.)*
- [10] Kufner, A.: *Fourierovy řady. Praha 1969*
- [11] Peters, M. T., Williams, J.: *The Fourier Transform in Biomedical Engineering, Alden Bookset, Oxford, England, 1998*
- [12] Sovka, P., Pollák, P.: *Hlasové technologie v podpoře informační společnosti, Souhrnný přehled aktivit řešitelských kolektivů, ČVUT v Praze, 2004*
- [13] Švec, J.G.: *Fyzilogická akustika zpěvního hlasu: Nový pohled na starý problém, Praha, ČVUT & Česká akustická společnost, pp. 219-226 (2000)*
- [14] Titze, I. R.: *Principles of Voice Production, A Paramount Communications Company Englewood Cliffs, 1994*
- [15] Aliasing: <http://herakles.zcu.cz/education/apg-2002-2003/hradek/html/Aliasing.html>, citováno 10/2009
- [16] Aliasing: <http://en.wikipedia.org/wiki/Aliasing>, citováno 9/2009
- [17] Foto hlasivkové štěrby: <http://commons.wikimedia.org/wiki/File:Larynx-normal2a.jpg>, citováno 10/2009

- [18] Fourierova transformace: <http://cs.wikipedia.org/wiki/Fourierova-transformace>, citováno 6/2009
- [19] Fourierovy řady a analýza: <http://en.wikipedia.org/wiki/Fourier-analysis>, citováno 6/2009
- [20] Hlasivková štěrbina: <http://en.wikipedia.org/wiki/Glottis>, citováno 10/2009
- [21] Hodnoty formantů: <http://en.wikipedia.org/wiki/Formant>, citováno 6/2009
- [22] Interference: <http://cs.wikipedia.org/wiki/Interference>, citováno 9/2009
- [23] Komorní A: <http://jlsubs.wordpress.com/2009/01/15/komorni-a-jeho-ladeni/>, citováno 10/2009
- [24] Obrázky k samohláskám podle D. Jonese: <http://en.wikipedia.org/wiki/Cardinal-vowel>, citováno 9/2009
- [25] Praat: <http://www.fon.hum.uva.nl/praat>, citováno 11/2009
- [26] Praat: <http://cs.wikipedia.org/wiki/Linear.predictive.coding>, citováno 11/2009
- [27] Rychlá Fourierova transformace: <http://en.wikipedia.org/wiki/Fast-Fourier-transform>, citováno 6/2009
- [28] Samplování: <http://www2.egr.uh.edu/~glover/applets/Sampling/Sampling.html>, citováno 10/2009

## 10 CD příloha

Na přiloženém CD naleznete celkem 4 složky – Databaze, Grafy, Interference a M-fily. V první z nich jsou data rozdělena do dvou podsložek na muže a ženy. V každé z nich jsou následující 3 podsložky:

- intonovane - obsahuje složky A, E, I, O, U, z nichž v každé jsou nahrávky jen intonovaných samohlásek všech mužů resp. žen
- podle osob - je rozdělena do deseti složek (pro každou osobu jedna), v nichž jsou již jednotlivé samohlásky dané osoby intonované (očíslované od 1 do 12) i mluvené (bez číselného označení)
- vsechny dohromady - obsahuje složky A, E, I, O, U, kde v každé jsou všechny nahrávky (intonované i mluvené) od všech mužů resp. žen.

V další složce s názvem Grafy je celkem 55 souborů, z nichž každý obsahuje graf z Obrazové přílohy. Názvy souborů odpovídají značení v Obrazové příloze. Složka M-fily obsahuje textové soubory, které jsou pojmenovány podle m-filů použitých v praktické části a jejich obsah je popsán tamtéž.



## 11 Příloha m-fily

Compare:

```
total = sum(P);
s = sum(total);
total = total/s;
N = 5000;
figure(1)
plot(total(1:N), 'blue');
title('pozadi')

una = sum(P(1:12,:));
s = sum(una);
una = una/s;
figure(2)
plot(total(1:N), 'blue');
hold on
plot(una(1:N), 'red')
title('modre pozadi a cervene A')

une = sum(P(13:24,:));
s = sum(une);
une = une/s;
figure(3)
plot(total(1:N), 'blue');
hold on
plot(une(1:N), 'red')
title('modre pozadi a cervene E')

uni = sum(P(25:36,:));
s = sum(uni);
uni = uni/s;
figure(4)
plot(total(1:N), 'blue');
hold on
plot(uni(1:N), 'red')
title('modre pozadi a cervene I')

uno = sum(P(37:48,:));
s = sum(uno);
uno = uno/s;
figure(5)
plot(total(1:N), 'blue');
```

```

hold on
plot(uno(1:N),'red')
title('modre pozadi a cervene 0')

```

```

unu = sum(P(49:60,:));
s = sum(unu);
unu = unu/s;
figure(6)
plot(total(1:N),'blue');
hold on
plot(unu(1:N),'red')
title('modre pozadi a cervene U')

```

Draw:

```

function draw(X,total,yy,foba,sound)
figure
plot(X,total,'y');
hold on
plot(X,yy,'k')
plot(X,foba,'r')
word = ['vowel ' sound];
legend('general vowel','exp fit',word)
title(word)

```

Expa:

```

function yhat = expa(b,X)
yhat = b(1)*exp(-b(2)*X)

```

Filter:

```

function fx = filtruj(x,h)

if mod(h,2)~=1
    error('okno neni liche delky')
end

```

```

a = (h+1)/2;
b = (h-1)/2;
N = length(x);

fx = x;
for i = a:1:N-a+1
    bit = x(i-b:i+b);
    p = mean(bit);
    fx(i)=p;
end

```

Findfor:

```

N = 5000;

total = sum(P);
total = total(1:N)';
s = sum(total);
total = total/s;

X = (1:1:N)';
betahat = nlinfit(X,total,@expa,[0;0]);
yy = expa(betahat,X);
s = sum(yy);
yyN = yy/s;

h = 131;

oban = oba(1:N)';
foba = filter(oban,h);
foba = foba/sum(foba);
draw(X,total,yy,foba,'A')

oben = obe(1:N)';
fobe = filter(oben,h);
fobe = fobe/sum(fobe);
draw(X,total,yy,fobe,'E')

obin = obi(1:N)';
fobi = filter(obin,h);
fobi = fobi/sum(fobi);
draw(X,total,yy,fobi,'I')

```

```
obon = obo(1:N)';  
fobo = filter(obon,h);  
fobo = fobo/sum(fobo);  
draw(X,total,yy,fobo,'O')
```

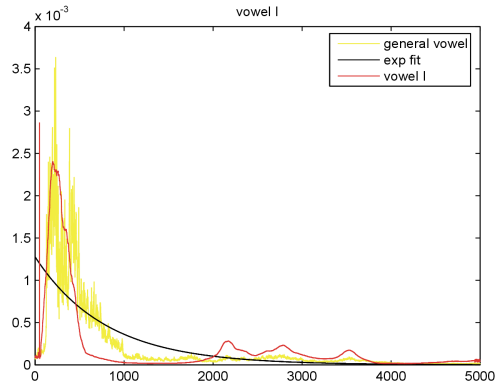
```
obun = obu(1:N)';  
fobu = filter(obun,h);  
fobu = fobu/sum(fobu);  
draw(X,total,yy,fobu,'U')
```

Read:

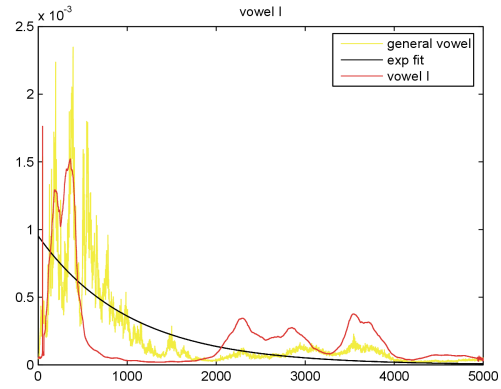
```
function [exit, sound] = read(x);  
original=wavread(x);  
original = original(:,1);  
m=length(original);  
original(44100:m)=[];  
four=fft(original);  
power=four.*conj(four)/44100;  
power(1)=[];  
N=5000;  
exit=sqrt(power);  
totals = sum(exit);  
exit =exit/totals;  
plot(exit(1:N),'red');  
axis([0;5000;0;25*10(-3)]);  
sound = original;
```

## 12 Obrazová příloha

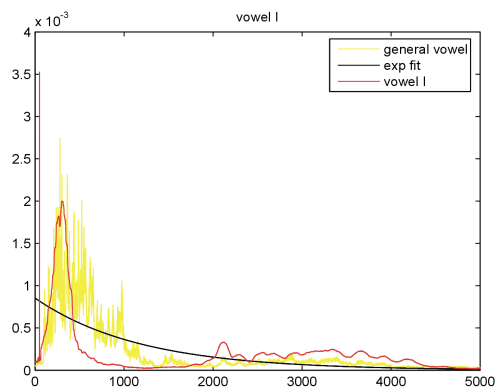
# 1 Grafy formantů mužů samohlásky I



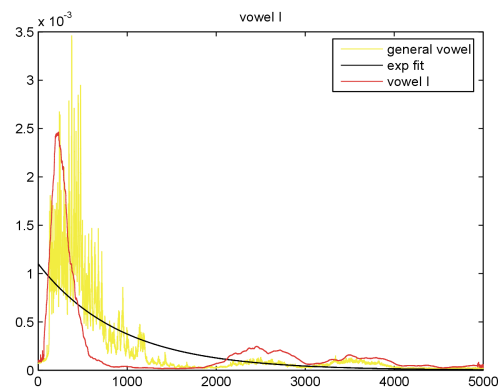
1. 1 graf muže 1



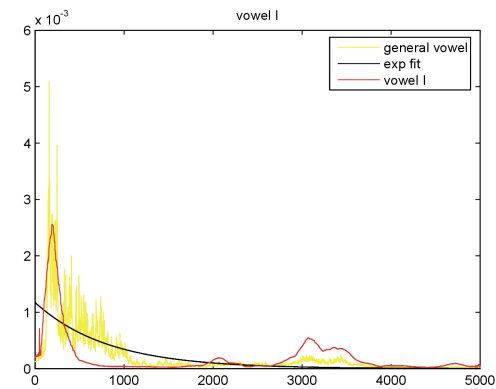
1. 2 graf muže 2



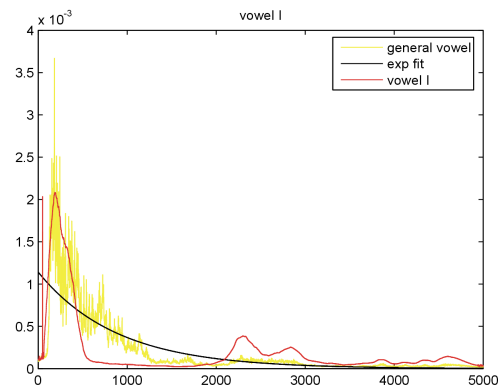
1. 3 graf muže 3



1. 4 graf muže 4

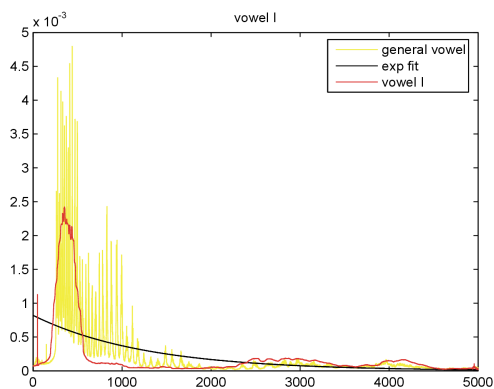


1. 5 graf muže 5

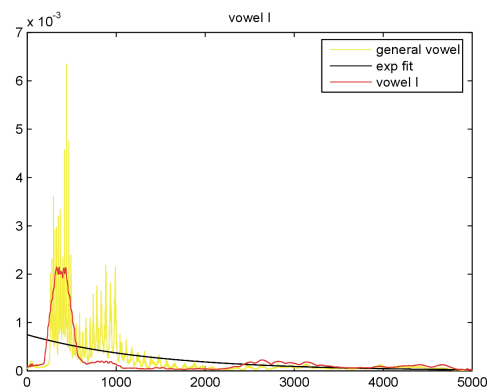


1. 6 graf muže 6

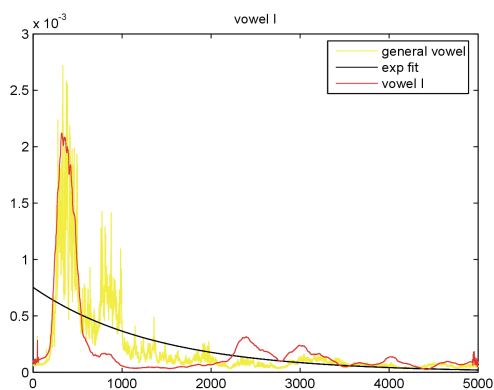
## 2 Grafy formantů žen samohlásky I



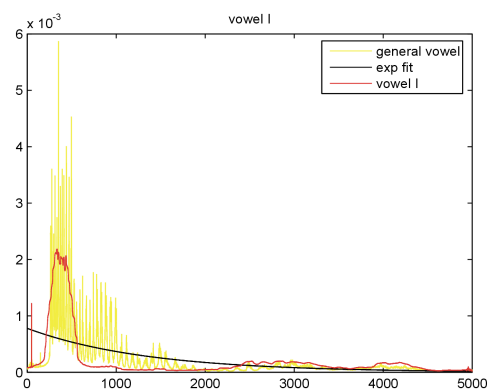
2. 1 graf ženy 1



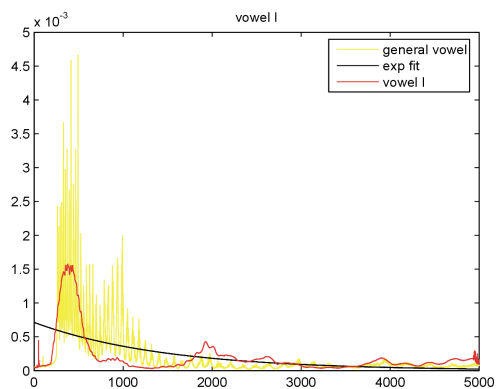
2. 2 graf ženy 2



2. 3 graf ženy 3

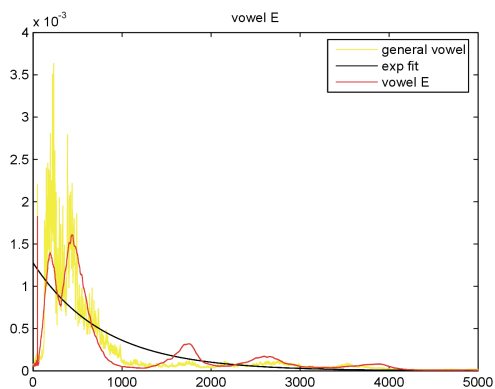


2. 4 graf ženy 4

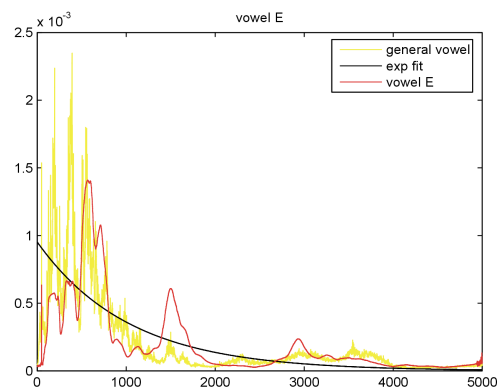


2. 5 graf ženy 5

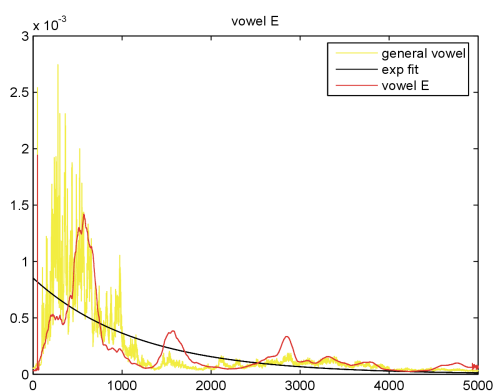
### 3 Grafy formantů mužů samohlásky E



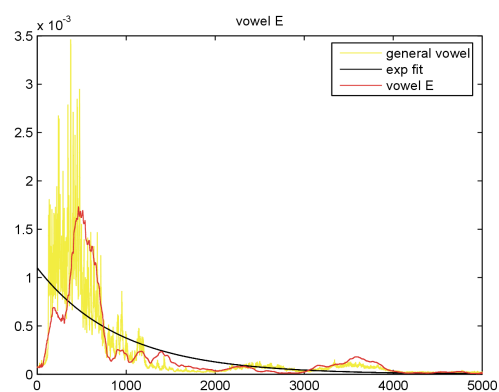
3. 1 graf muže 1



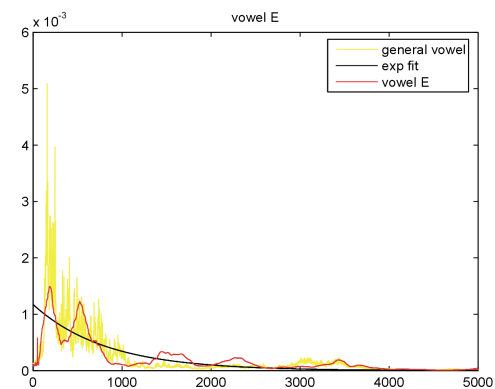
3. 2 graf muže 2



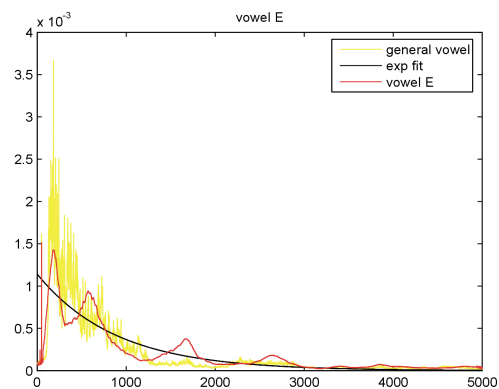
3. 3 graf muže 3



3. 4 graf muže 4



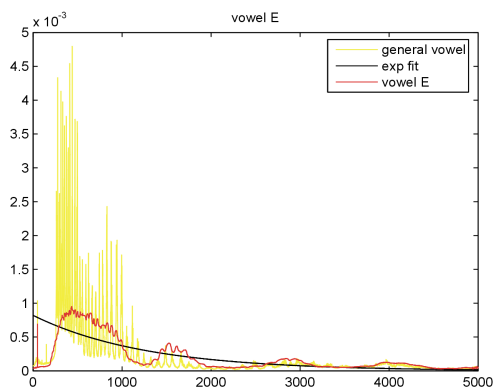
3. 5 graf muže 5



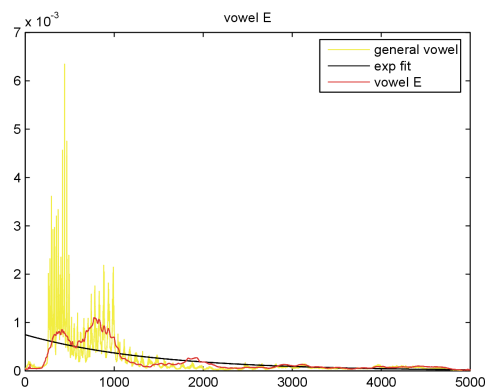
3. 6 graf muže 6



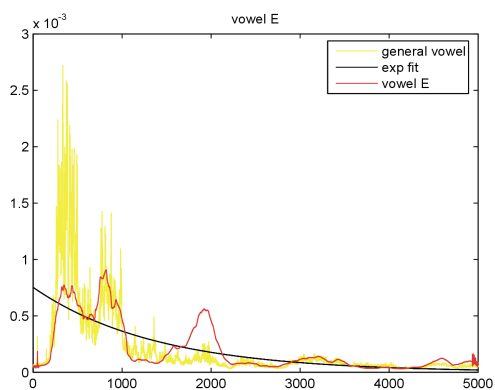
#### 4 Grafy formantů žen samohlásky E



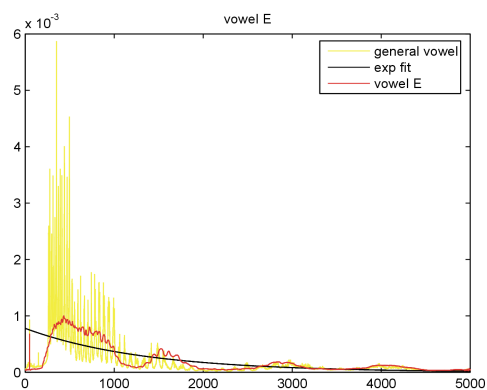
4. 1 graf ženy 1



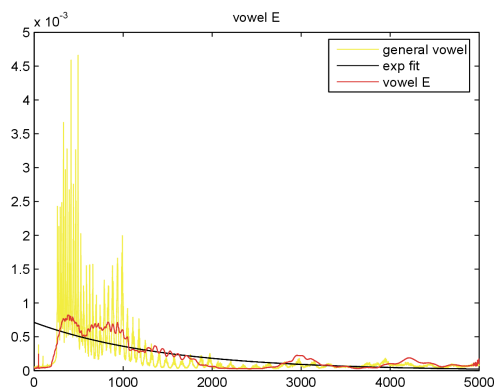
4. 2 graf ženy 2



4. 3 graf ženy 3

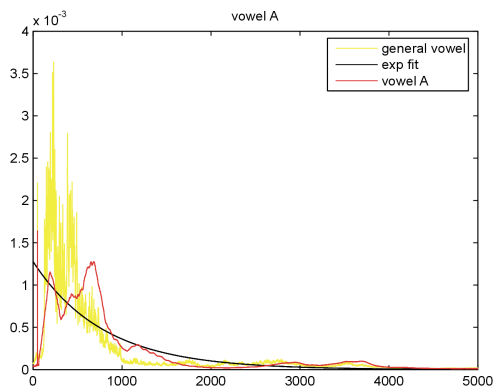


4. 4 graf ženy 4

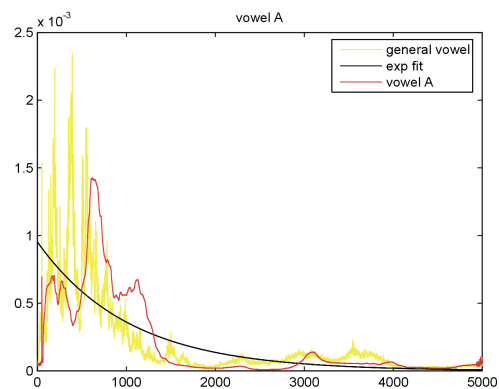


4. 5 graf ženy 5

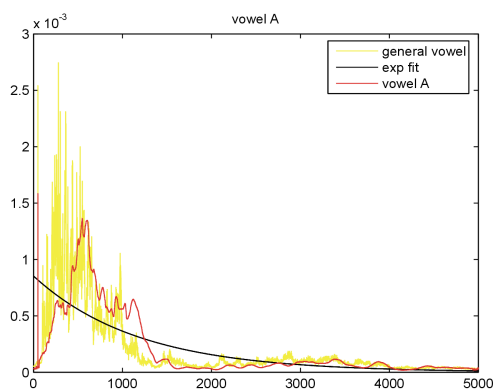
## 5 Grafy formantů mužů samohlásky A



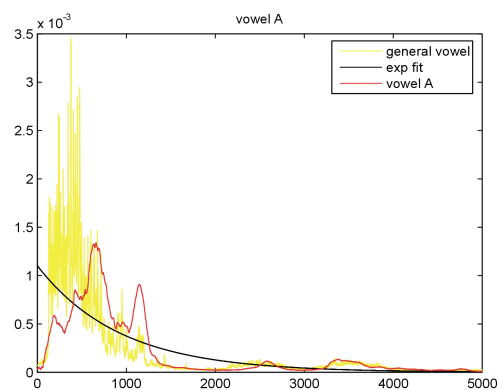
5. 1 graf muže 1



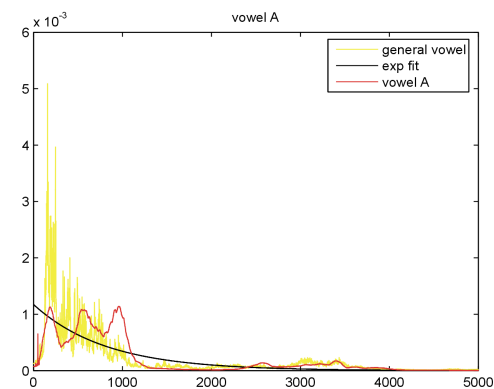
5. 2 graf muže 2



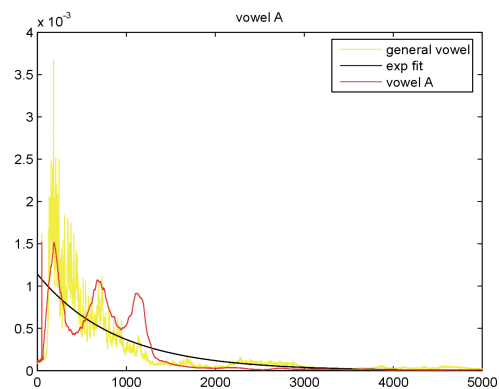
5. 3 graf muže 3



5. 4 graf muže 4

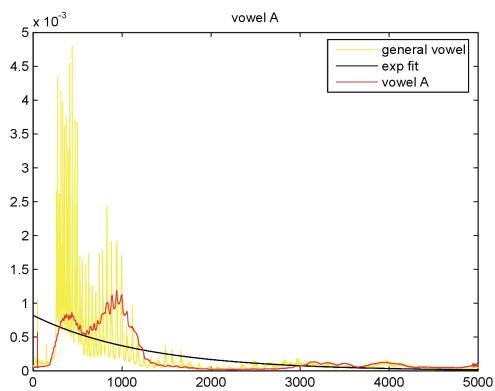


5. 5 graf muže 5

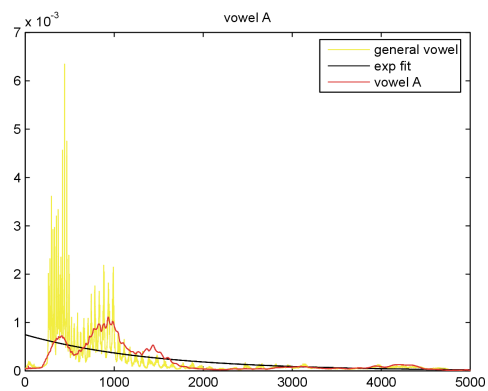


5. 6 graf muže 6

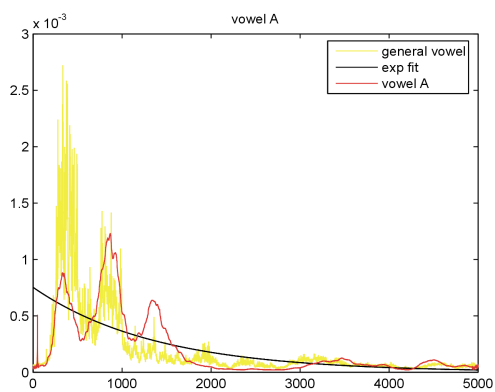
## 6 Grafy formantů žen samohlásky A



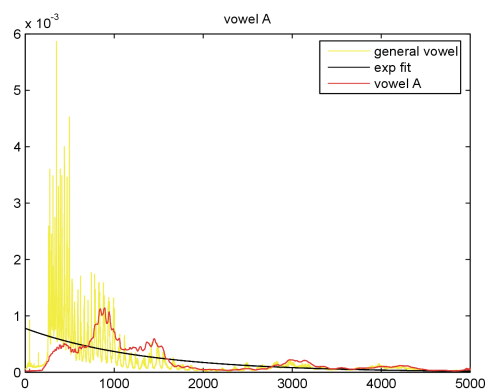
6. 1 graf ženy 1



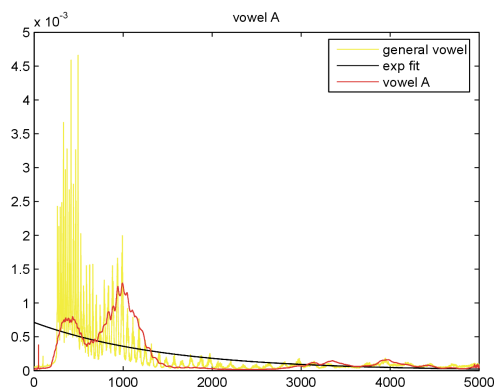
6. 2 graf ženy 2



6. 3 graf ženy 3

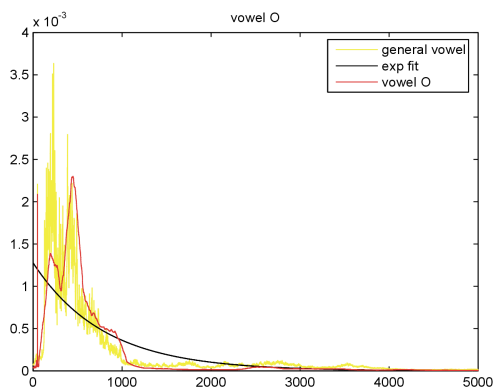


6. 4 graf ženy 4

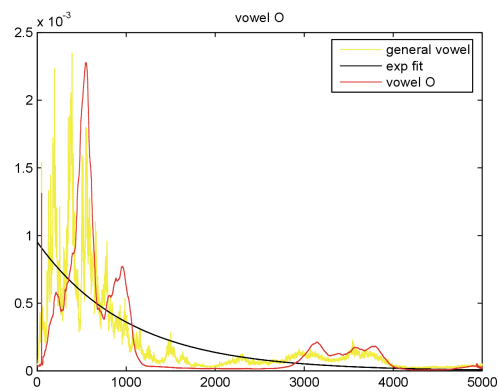


6. 5 graf ženy 5

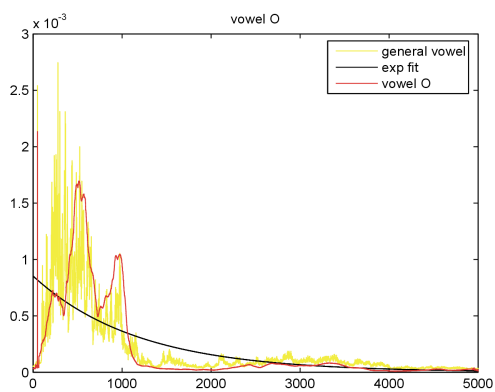
## 7 Grafy formantů mužů samohlásky O



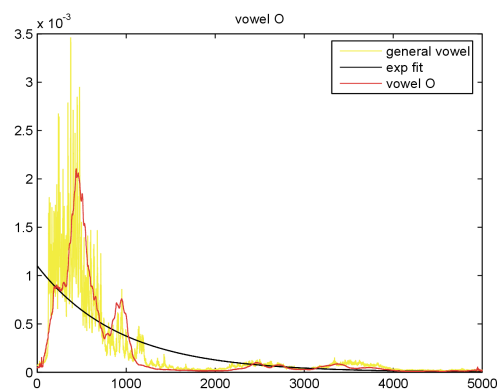
7. 1 graf muže 1



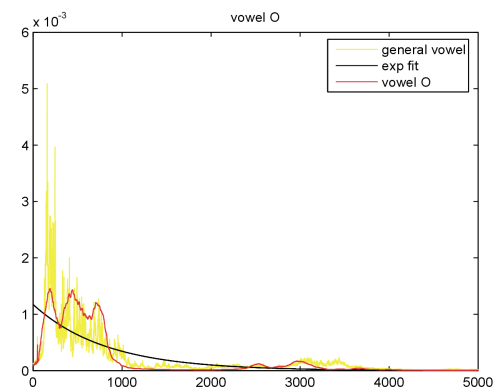
7. 2 graf muže 2



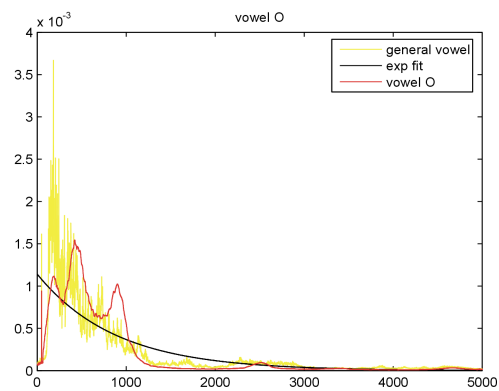
7. 3 graf muže 3



7. 4 graf muže 4

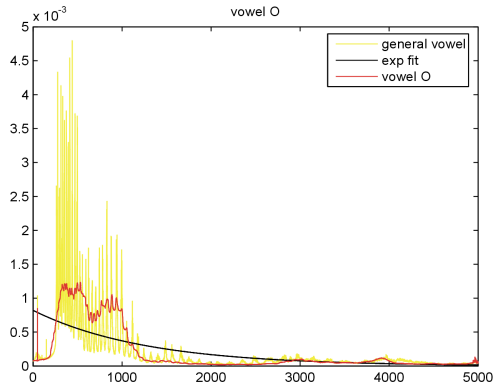


7. 5 graf muže 5

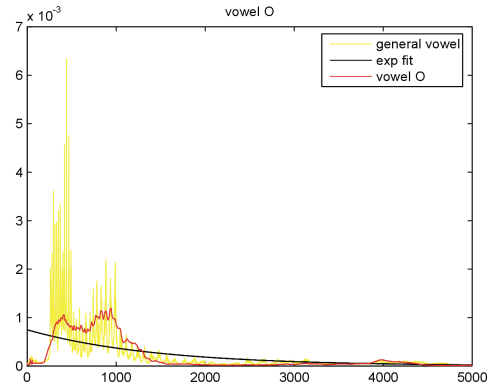


7. 6 graf muže 6

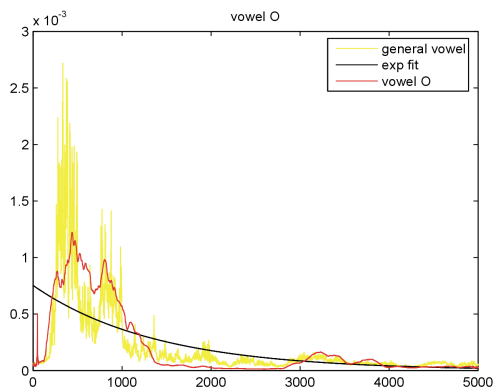
## 8 Grafy formantů žen samohlásky O



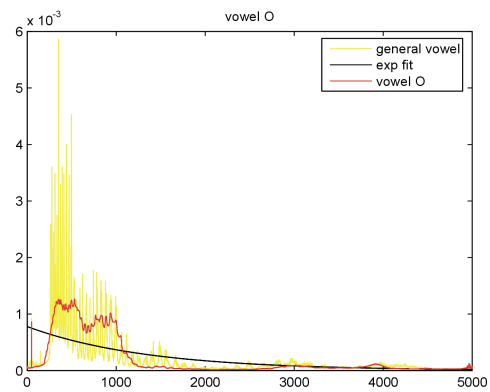
8. 1 graf ženy 1



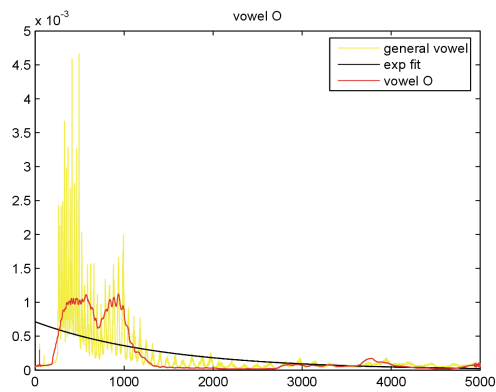
8. 2 graf ženy 2



8. 3 graf ženy 3

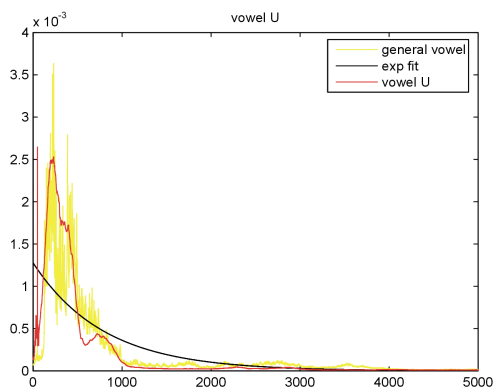


8. 4 graf ženy 4

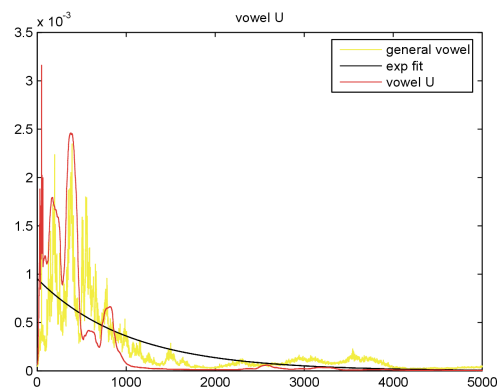


8. 5 graf ženy 5

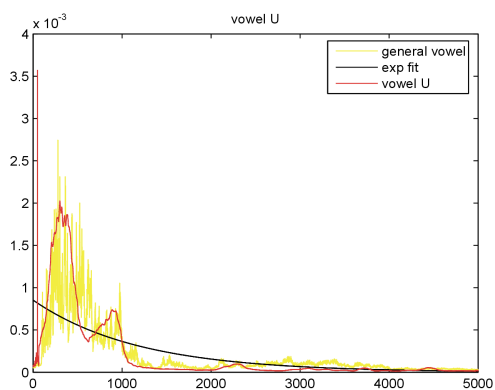
## 9 Grafy formantů mužů samohlásky U



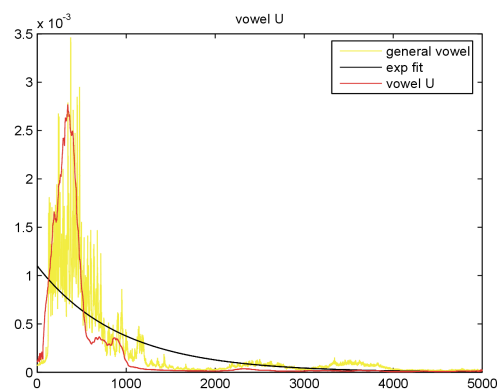
9. 1 graf muže 1



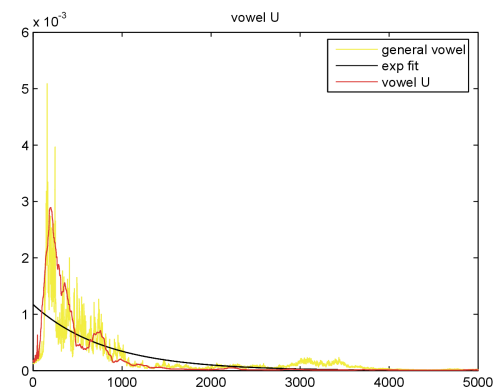
9. 2 graf muže 2



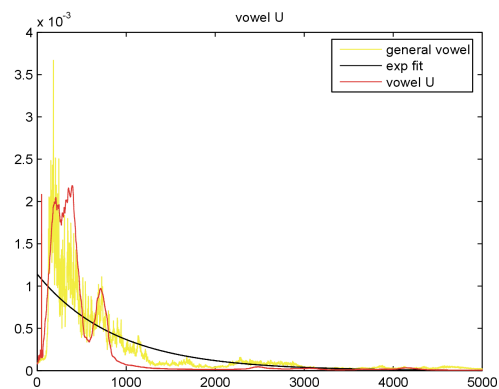
9. 3 graf muže 3



9. 4 graf muže 4

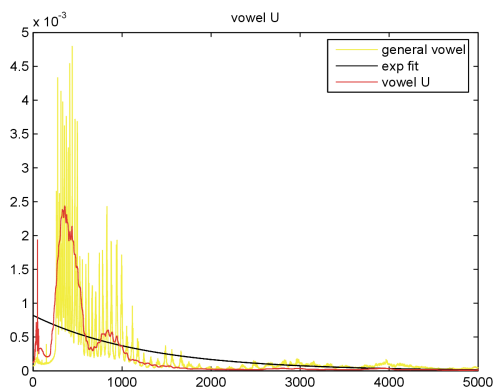


9. 5 graf muže 5

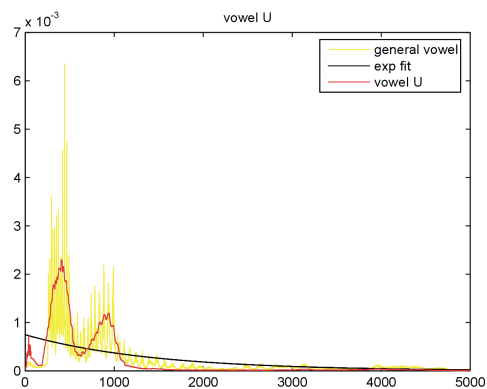


9. 6 graf muže 6

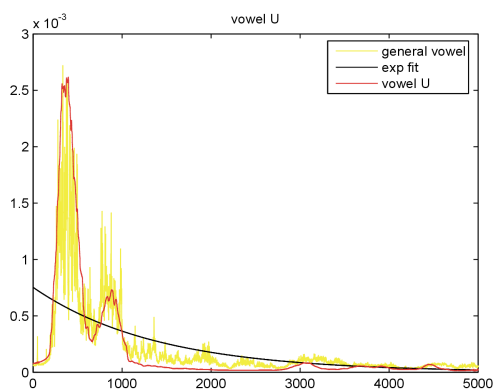
## 10 Grafy formantů žen samohlásky U



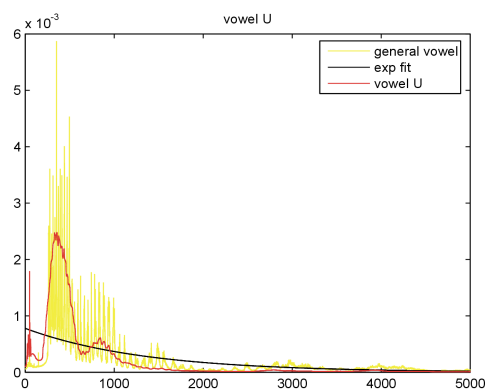
10. 1 graf ženy 1



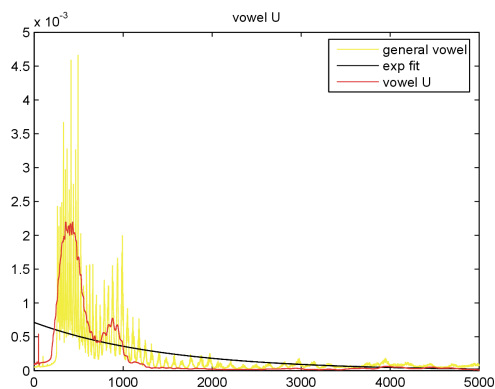
10. 2 graf ženy 2



10. 3 graf ženy 3



10. 4 graf ženy 4



10. 5 graf ženy 5