

UNIVERZITA PALACKÉHO V OLMOUCI  
PŘÍRODOVĚDECKÁ FAKULTA  
KATEDRA MATEMATICKÉ ANALÝZY A APLIKACÍ MATEMATIKY

## BAKALÁŘSKÁ PRÁCE

Lombův periodogram a jeho použití na  
frekvenční analýzu EKG signálu



Vedoucí bakalářské práce:  
**RNDr. Jitka Machalová, Ph.D.**  
Rok odevzdání: 2013

Vypracovala:  
**Markéta Mraznicová**  
MATAP, III. ročník

### **Prohlášení**

Prohlašuji, že jsem bakalářskou práci zpracovala samostatně pod vedením RNDr. Jitky Machalové, Ph.D. a všechny použité zdroje jsem uvedla v seznamu literatury v závěru práce.

V Olomouci dne 22. 4. 2013

## **Poděkování**

Děkuji RNDr. Tomáši Fürstovi, Ph.D. za návrh zajímavého tématu práce a za výraznou pomoc s programovou stránkou práce. Dále bych chtěla poděkovat RNDr. Jitce Machalové, Ph.D. za převzetí vedení mé práce, vstřícný přístup a cenné rady, které mi značně pomohly při zpracování práce.

# Obsah

Úvod	4
<b>1 Potřebné definice</b>	<b>6</b>
<b>2 Nalezení významných frekvencí ve stejnoměrně vzorkovaném signálu</b>	<b>9</b>
2.1 Definice periodogramu	9
2.2 Test R. A. Fishera	14
<b>3 Nalezení významných frekvencí v nestejně vzorkovaném signálu</b>	<b>18</b>
3.1 Vaníčkova metoda	18
3.2 Lombův periodogram	19
3.2.1 Metoda nejmenších čtverců	19
3.2.2 Vztahy pro spektrum získané metodou nejmenších čtverců	22
3.2.3 Spektrum vzhledem k sinusovému signálu	26
<b>4 Srovnání frekvenčních charakteristik reálných EKG signálů</b>	<b>30</b>
4.1 Popis datového souboru	31
4.1.1 Stanovení RR intervalů	32
4.2 Frekvenční charakteristiky srdečního rytmu	35
4.2.1 Aplikace Lombova periodogramu na EKG signálech	37
<b>Závěr</b>	<b>41</b>
<b>Přílohy</b>	<b>42</b>
<b>Literatura</b>	<b>58</b>

# Úvod

Pojem perioda pochází z řeckého slova „períodos“, což v překladu znamená „chození dokola“. Periodicita je tedy jev opakující se v pravidelných intervalech, respektive po uplynutí určitého časového období — periody. Periodicita je běžný jev vyskytující se téměř všude okolo nás. Periodicita se vyskytuje i v běžných jevech, které bereme jako samozřejmost, např. střídání dne a noci, či pravidelné střídání ročních období. V životě každého z nás je periodicita velmi důležitá pro správnou funkci lidského organismu, jako je například chůze (střídání pohybu obou končetin), krevní oběh či činnost srdce.

Stahy srdečního svalu jsou nejpravidelnějším jevem v lidském těle. Elektrickou aktivitu srdce snímá přístroj zvaný elektrokardiograf. Následně pořízený záznam se nazývá elektrokardiogram, zkráceně EKG. Na záznam EKG můžeme z matematického hlediska nahlížet jako na časovou řadu, u které lze pozorovat jistou míru periodičnosti. Hlavním z cílů této bakalářské práce je pomocí statistických nástrojů popsaných v následujících kapitolách zjistit a otestovat významnost period, tedy i frekvencí, reálných EKG záznamů.

Vhodná volba metody určená k analýze časových řad je založena na mnoha faktorech, jako je např. účel analýzy, typ časové řady a spousta dalších. Základní postupy analýzy časových řad lze rozdělit na dvě velké kategorie. Jako první se můžeme setkat s tzv. analýzou časových řad v časové doméně (např. dekompozice čas. řad). Jiný přístup, ve kterém považujeme zkoumanou řadu za směs sinusových a kosinusových křivek o různých amplitudách a frekvencích, se označuje jako *analýza časových řad ve spektrální doméně*, nebo také *spektrální analýza časových řad*, viz [3]. Touto kategorií se budeme, z matematického pohledu, podrobně zabývat v této práci a uvedeme podrobné odvození metod.

V současnosti se využívá dvou základních metod pro popis a zpracování periodických časových řad. První z nich je Fourierova transformace a druhou metodu poskytuje periodogram, jehož pochopení a popis je další z cílů naší práce. Periodogram byl jedním z prvních statistických nástrojů, pomocí kterých se studují

periodické vlastnosti časových řad. Před vývojem periodogramu byly veškeré takové analýzy příliš zdlouhavé a velmi obtížně realizovatelné.

Periodogram lze správným způsobem aplikovat pouze na stejnoměrně vzorkované signály, což na druhé straně může být i základní nevýhodou. Výrazný pokrok ve spektrální analýze nestejnoměrně rozložených dat učinil v roce 1971 vědec pocházející z Československa, Petr Vaníček. Na Vaníčkovu metodu navázal astronom Nicholas R. Lomb v roce 1976, který poukázal na blízké spojení s problematikou použití periodogramu, což bude blíže prozkoumáno v dalším textu.

Spektrální analýzu budeme využívat k vyšetření variability tepové frekvence, tedy k nalezení významných frekvencí EKG signálů.

# 1. Potřebné definice

V této kapitole uvádíme potřebné teoretické podklady nutné pro zřetelnější pochopení dalšího textu. Předpokládáme, že budoucí čtenář má základní znalosti a ovládá základní pojmy z oblasti pravděpodobnosti a statistiky, jako je např.  $\sigma$ -algebra, náhodný jev apod. V textu čerpáme z [1], [2], [7] a [15]. Pro přehlednost nebudeme uvádět přímé citace.

**Definice 1.** Nechť je dána množina  $\Omega$ ,  $\Omega \neq \emptyset$  a na ní  $\sigma$ -algebra  $\Theta$  náhodných jevů. *Pravděpodobností* nazveme každou funkci  $P : \Theta \rightarrow \mathbb{R}$  splňující následující podmínky:

- a)  $P(\Omega) = 1$
- b)  $P(A) \geq 0, \forall A \in \Theta$
- c) pro libovolnou posloupnost neslučitelných náhodných jevů  $A_n \in \Theta$ ,  $n = 1, 2, \dots$ , platí

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

**Definice 2.** Uspořádanou trojici  $(\Omega, \Theta, P)$  nazýváme *pravděpodobnostní prostor*.

**Definice 3.** Nechť je dán pravděpodobnostní prostor  $(\Omega, \Theta, P)$ . Funkci  $X : \Omega \rightarrow \mathbb{R}$  nazýváme *náhodnou veličinou*, jestliže pro tuto funkci a pro každé  $x \in \mathbb{R}$  platí

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \Theta,$$

tj. jestliže funkce  $X$  je  $\Omega$ -měřitelná.

**Definice 4.** *Rozdělení pravděpodobnosti náhodné veličiny*  $X$  je množinová funkce  $P_X(B) : \mathcal{B} \rightarrow \mathbb{R}$  definovaná vztahem

$$P_X(B) = P(X \in B), \quad B \in \mathcal{B},$$

kde  $\mathcal{B}$  je systém borelovských podmnožin  $X$ .

**Definice 5.** Necht  $X$  je náhodná veličina definovaná na pravděpodobnostním prostoru  $(\Omega, \Theta, P)$ . Funkce  $F_X(x)$  definovaná na  $\mathbb{R}$  předpisem

$$F_X(x) = P(X \leq x), \quad x \in \mathbb{R},$$

se nazývá *distribuční funkce náhodné veličiny  $X$* .

**Definice 6.** Distribuční funkce  $F_X(x)$  náhodné veličiny  $X$  se nazývá *absolutně spojitá*, existuje-li nezáporná, borelovsky měřitelná funkce  $f_X(x) : \mathbb{R} \rightarrow \mathbb{R}$  taková, že

$$F_X(x) = \int_{-\infty}^x f_X(t) dt, \quad \forall x \in \mathbb{R}.$$

Funkce  $f_X$  se nazývá *hustota* (rozdělení pravděpodobnosti) náhodné veličiny  $X$ .

**Definice 7.** *Normální rozdělení (Gaussovo)* s parametry  $\mu, \sigma$  má náhodná veličina  $X$ , která má hustotu

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \forall x \in \mathbb{R}.$$

**Definice 8.** Necht  $X$  je náhodná veličina definovaná na pravděpodobnostním prostoru  $(\Omega, \Theta, P)$ . *Střední hodnota  $E(X)$  náhodné veličiny  $X$*  je definována jako integrál z  $X$  podle míry  $P$ , tj.

$$\mu = E(X) = \int_{\Omega} X(\omega) dP(\omega).$$

Výraz

$$E(X - E(X))^r$$

nazýváme  *$r$ -tý centrální moment*.

**Definice 9.** Druhý centrální moment náhodné veličiny  $X$  nazýváme *rozptyl náhodné veličiny  $X$* , značíme  $\sigma^2$  resp.  $var(X)$ , a platí

$$\sigma^2 = var(X) = E(X - E(X))^2.$$



**Definice 10.** Střední hodnotu součinu odchylek dvou náhodných veličin  $X$  a  $Y$  od jejich středních hodnot nazýváme *kovariance* a značíme

$$\text{cov}(X, Y) = E([X - E(X)][Y - E(Y)]) = E(XY) - E(X)E(Y).$$

**Definice 11.** Posloupnost  $n$  nezávislých náhodných veličin  $X_1, \dots, X_n$  se stejným rozdělením pravděpodobnosti  $F$  nazýváme *náhodný výběr* o rozsahu  $n$  z rozdělení  $F$ .

**Věta 1.** Nechť  $(X_1, \dots, X_n)^T$  je náhodný výběr příslušný náhodné veličině  $X$ , která má konečnou střední hodnotu  $E(X) = \mu$  a konečný rozptyl  $\text{var}(X) = \sigma^2$ . Pak platí, že rozptyl výběrového průměru  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  je roven

$$\text{var}(\bar{X}_n) = \frac{\sigma^2}{n}.$$

**Důkaz:** viz [7], str. 136-137.

**Definice 12.** Časovou řadou nazveme posloupnost věcně a prostorově srovnatelných pozorování (dat), která jsou jednoznačně uspořádaná z časového hlediska od minulosti do přítomnosti.

## 2. Nalezení významných frekvencí ve stejnoměrně vzorkovaném signálu

EKG signál můžeme chápat jako časovou řadu vykazující jistou dávku periodicity. To znamená, že na takovou řadu můžeme pohlížet jako na soubor periodických složek o různých frekvencích.

Pro nalezení významných periodických složek, tzn. nalezení významných frekvencí v analyzované řadě, použijeme analýzu pomocí statistického nástroje zvaného *periodogram*. Termín „periodogram“ byl poprvé uveden v roce 1897, kdy Arthur Schuster představil postup, který může poskytnout informace o periodických složkách časových řad, což se můžeme dočíst v [18]. Tato metoda umožňuje pozorovat vývoj spektra v závislosti na čase a využívá se především k výpočtu odhadu složek spektrálního výkonu.

Princip této metody spočívá v převedení jednotlivých členů časové řady na tvary, ve kterých vystupují goniometrické funkce sinus a kosinus. Pokud řada obsahuje periodický člen o určité periodě, nalezneme v tomto bodě globální extrém, ale také mnoho jiných vrcholů, jež jsou odezvami tohoto extrému. V literatuře je tento jev nazýván jako tzv. aliasing.

K otestování, zda časová řada významnou periodu obsahuje, slouží tzv. test R. A. Fishera, který bude popsán v druhé části této kapitoly.

### 2.1. Definice periodogramu

Před tím než přistoupíme k definici periodogramu, je nutné uvést několik základních pojmů a vztahů. Periodický děj či periodická posloupnost hodnot je charakterizována periodou  $T$ , která vyjadřuje nejkratší možný čas, po kterém se děj či hodnoty posloupnosti opakují. Frekvence  $f$ , jejíž jednotkou je hertz ( $\text{Hz} = \text{s}^{-1}$ ), určuje počet takových opakování za jednotku času, je tedy dána vztahem

$$f = \frac{1}{T}.$$

Jak bylo řečeno v úvodu této kapitoly, na EKG signál lze pohlížet jako na časovou řadu. V souvislosti s časovými řadami budeme místo frekvence  $f$  užívat tzv. úhlovou frekvenci  $\omega$ , jež je definovaná jako změna fáze za jednotku času, a tedy jednotkou je  $rad \cdot s^{-1}$ . Vzájemný vztah předchozích tří veličin je

$$\omega = \frac{2\pi}{T} = 2\pi f.$$

Pokud budeme pracovat s posloupností o délce  $N$ , která bude menší ve srovnání s délkou skutečné periody dané vztahem

$$T = \frac{2\pi}{\omega},$$

pak bude tato perioda připomínat spíše trend. Největší možná zjistitelná délka periody má hodnotu počtu pozorování, tj.

$$T_{max} = N,$$

které odpovídá minimální frekvence

$$\omega_{min} = \frac{2\pi}{N}.$$

Naopak nejmenší možná zjistitelná perioda má délku

$$T = 2,$$

a odpovídá jí tzv. *Nyquistova frekvence* viz [1], str. 77.

$$\omega = \pi.$$

Tato frekvence je největší možná, která umožňuje z pozorované funkce usuzovat o charakteru periodického chování, jak bylo uvedeno v [3], str. 206.

Periodogram lze definovat, podle [1], str. 75, následovně:

**Definice 13.** Nechť je dána konečná posloupnost náhodných veličin  $X_1, \dots, X_N$ . *Periodogramem* posloupnosti  $X_1, \dots, X_N$  nazýváme funkci

$$I(\omega) = \frac{1}{2\pi N} \left| \sum_{t=1}^N X_t e^{-it\omega} \right|^2, \quad (1)$$

kde  $-\pi \leq \omega \leq \pi$ .

Periodogram lze definovat také pro spojitý případ (viz [1], str. 75), ale jelikož takový periodogram nebudeme využívat, omezíme se v dalším textu pouze na diskrétní případ periodogramu, tj. vztah (1). Podívejme se na jiné vyjádření periodogramu, které nám umožní pracovat se členy časové řady vyjádřenými pomocí goniometrických funkcí. Nyní představíme lemma převzaté z [1], str. 77-78, kde je uveden také nástin důkazu. Pro lepší pochopení postupu v důkazu uvádíme podrobnější popis.

**Lemma 1.** Necht

$$A(\omega) = \sqrt{\frac{2}{N}} \sum_{t=1}^N X_t \cos \omega t, \quad B(\omega) = \sqrt{\frac{2}{N}} \sum_{t=1}^N X_t \sin \omega t. \quad (2)$$

Poté platí

$$I(\omega) = \frac{1}{4\pi} [A^2(\omega) + B^2(\omega)]. \quad (3)$$

**Důkaz:** K důkazu lemmatu využijeme vztahu vyjadřující komplexní exponenciální funkci

$$e^{\alpha i} = \cos \alpha + i \sin \alpha.$$

Pokud za  $e^{-it\omega}$  dosadíme do definice periodogramu (1), dostaneme

$$\begin{aligned} I(\omega) &= \frac{1}{2\pi N} \left| \sum_{t=1}^N X_t e^{-\omega t i} \right|^2 = \frac{1}{2\pi N} \left| \sum_{t=1}^N X_t (\cos(-\omega t) + i \sin(-\omega t)) \right|^2 = \\ &= \frac{1}{2\pi N} \left| \sum_{t=1}^N X_t (\cos \omega t - i \sin \omega t) \right|^2 = \frac{1}{2\pi N} \left| \sum_{t=1}^N X_t \cos \omega t - i \sum_{t=1}^N X_t \sin \omega t \right|^2. \end{aligned}$$

Dále pomocí vzorce pro velikost (absolutní hodnotu) komplexního čísla

$$|z| = \sqrt{x^2 + y^2} \quad \forall z \in \mathbb{C}, \text{ kde } z = x + iy$$

můžeme psát

$$\begin{aligned}
I(\omega) &= \frac{1}{2\pi N} \left| \sum_{t=1}^N X_t \cos \omega t - i \sum_{t=1}^N X_t \sin \omega t \right|^2 = \\
&= \frac{1}{2\pi N} \left\{ \left( \sum_{t=1}^N X_t \cos \omega t \right)^2 + \left( \sum_{t=1}^N X_t \sin \omega t \right)^2 \right\} = \\
&= \frac{1}{4\pi} [A^2(\omega) + B^2(\omega)].
\end{aligned}$$

□

Nyní ukážeme, že periodogram je opravdu vhodným a dobrým ukazatelem periodicit. Předpokládejme, že konečná posloupnost náhodných veličin  $\{X_t\}$  splňuje

$$X_t = \sum_{k=1}^p a_k e^{it\omega_k} + Y_t, \quad (4)$$

kde  $t = 1, 2, \dots, N$ ,  $a_k \in \mathbb{R}$ ,  $a_k \neq 0$  pro každé  $k = 1, \dots, p$ ,  $\omega_k \in (-\pi, \pi)$  takové, že  $\omega_i \neq \omega_j$  pro  $i \neq j$ ,  $\{Y_t\}$  je posloupnost vzájemně nekorelovaných náhodných veličin, tj.  $\text{cov}(Y_i, Y_j) = 0 \forall i \neq j$ ,  $i, j = 1, \dots, p$ , jež mají nulovou střední hodnotu a stejný kladný rozptyl  $\sigma^2 > 0$ .

Nyní vyjádříme periodogram posloupnosti  $\{X_t\}_{t=1}^N$  dané vztahem (4) a dostaneme

$$\begin{aligned}
I(\omega) &= \frac{1}{2\pi N} \left| \sum_{t=1}^N X_t e^{-it\omega} \right|^2 = \frac{1}{2\pi N} \left| \sum_{t=1}^N \left( \sum_{k=1}^p a_k e^{it\omega_k} + Y_t \right) e^{-it\omega} \right|^2 = \\
&= \frac{1}{2\pi} \left| \left( \frac{1}{\sqrt{N}} \sum_{k=1}^p a_k \sum_{t=1}^N e^{ti(\omega_k - \omega)} \right) + \left( \frac{1}{\sqrt{N}} \sum_{t=1}^N Y_t e^{-it\omega} \right) \right|^2.
\end{aligned} \quad (5)$$

Pro snadnější práci s úpravou výrazu (5), označme

$$A = \frac{1}{\sqrt{N}} \sum_{k=1}^p a_k \sum_{t=1}^N e^{ti(\omega_k - \omega)}, \quad (6)$$

$$B = \frac{1}{\sqrt{N}} \sum_{t=1}^N Y_t e^{-it\omega}. \quad (7)$$

Tedy výraz (5) lze psát ve tvaru

$$I(\omega) = \frac{1}{2\pi} \left| A + B \right|^2. \quad (8)$$

Je-li  $\omega \neq \omega_k$  pro každé  $k = 1, \dots, p$ , lze využít pro úpravu výrazu (6) vzorce pro součet konečné geometrické řady, který bude v našem případě ve tvaru

$$\sum_{t=1}^N e^{ti(\omega_k - \omega)} = e^{i(\omega_k - \omega)} \frac{e^{Ni(\omega_k - \omega)} - 1}{e^{i(\omega_k - \omega)} - 1}.$$

A tedy výraz  $A$ , vztah (6), můžeme vyjádřit jako

$$A = \frac{1}{\sqrt{N}} \sum_{k=1}^p a_k e^{i(\omega_k - \omega)} \frac{e^{Ni(\omega_k - \omega)} - 1}{e^{i(\omega_k - \omega)} - 1}.$$

Dále lze dokázat, že tento člen konverguje k 0 pro  $N \rightarrow \infty$ , viz [1], str. 76, což znamená, že

$$\lim_{N \rightarrow \infty} \frac{1}{\sqrt{N}} \sum_{k=1}^p a_k e^{i(\omega_k - \omega)} \frac{e^{Ni(\omega_k - \omega)} - 1}{e^{i(\omega_k - \omega)} - 1} = 0.$$

Další možná situace může nastat, jestliže  $\exists j \in \{1, 2, \dots, p\}$  takové, že  $\omega = \omega_j$ .

Pak platí

$$\begin{aligned} A &= \frac{1}{\sqrt{N}} N a_j + \frac{1}{\sqrt{N}} \sum_{k=1, k \neq j}^p a_k e^{i(\omega_k - \omega_j)} \frac{e^{Ni(\omega_k - \omega_j)} - 1}{e^{i(\omega_k - \omega_j)} - 1} = \\ &= \sqrt{N} a_j + \frac{1}{\sqrt{N}} \sum_{k=1, k \neq j}^p a_k e^{i(\omega_k - \omega_j)} \frac{e^{Ni(\omega_k - \omega_j)} - 1}{e^{i(\omega_k - \omega_j)} - 1}. \end{aligned}$$

Absolutní hodnota tohoto členu roste nade všechny meze pro  $N \rightarrow \infty$ , viz [1], str. 76. Nyní se blíže podíváme na výraz  $B$ , vztah (7). Jedná se o náhodnou veličinu s nulovou střední hodnotou z důvodu uvedeného v předpokladech této části kapitoly, tj.  $\{Y_t\}$  je posloupnost vzájemně nekorelovaných náhodných veličin s nulovou střední hodnotou a stejným kladným rozptylem  $\sigma^2$ . Podíváme se, jak

vypadá rozptyl této náhodné veličiny, tedy

$$\begin{aligned}
\text{var}\left(\frac{1}{\sqrt{N}} \sum_{t=1}^N Y_t e^{-it\omega}\right) &= E\left|\frac{1}{\sqrt{N}} \sum_{t=1}^N Y_t e^{-it\omega}\right|^2 = E\left(\frac{1}{N} \sum_{t=1}^N Y_t e^{-it\omega} \sum_{r=1}^N Y_r e^{-ir\omega}\right) = \\
&= \frac{1}{N} E\left(\sum_{t=1}^N Y_t e^{-it\omega} \sum_{r=1}^N Y_r e^{-ir\omega}\right) = \\
&= \frac{1}{N} \sum_{t=1}^N \sum_{r=1}^N e^{-i\omega(t-r)} E(Y_t Y_r) = \frac{1}{N} \sum_{t=1}^N E(Y_t^2) = \\
&= \frac{1}{N} N \sigma^2 = \sigma^2.
\end{aligned}$$

Jak vidíme, rozptyl uvedené náhodné veličiny nezávisí na  $N$ . Jestliže tedy budeme uvažovat model (5), lze konstatovat, že periodogram dosáhne významně velkých hodnot v bodech  $\omega_k$ , pro  $k = 1, \dots, p$ . Tato významná lokální maxima identifikují periodické chování uvažované situace. Mimo tyto body budou hodnoty poměrně malé, kolísající okolo hodnoty  $\frac{\sigma^2}{2\pi}$ , která byla odvozena podle věty 1. Jako poslední krok je nutné otestovat, které hodnoty periodogramu můžeme s jistotou považovat za významně velké v porovnání s ostatními. K takovému ověření použijeme test R. A. Fishera.

## 2.2. Test R. A. Fishera

V předchozí podkapitole jsme zjistili, že periodogram dosahuje výrazně velkých hodnot v bodech odpovídajících frekvencím  $\omega_1, \dots, \omega_p$ . Nyní je ovšem nutné zvolit vhodný test pro zjištění, které z těch hodnot jsou statisticky významné. Pro tyto účely slouží test R. A. Fishera, jež byl navržen roku 1929. Postup tohoto testu bude popsán v této kapitole, ve které čerpáme z [1] a [12].

Nechť je dána posloupnost náhodných veličin  $X_1, \dots, X_N$ . Pro zjednodušení a s ohledem na praktickou stránku předpokládejme, že  $N \geq 3$  a  $N$  je liché číslo, tedy

$$N = 2m + 1,$$

kde  $m \in \mathbb{N}$ . Pro kompletnost doplníme, že při sudém počtu náhodných veličin vždy první z nich vypouštíme, jelikož tato náhodná veličina je pro nás nejvzdálenější od současnosti.

Testujeme hypotézu, že uvažovaná posloupnost náhodných veličin je nezávislá a má normální rozdělení s nulovou střední hodnotou a rozptylem  $\sigma^2$ , tedy

$$H_0 : X_t \sim N(0, \sigma^2),$$

což v našem případě znamená, že periodická složka není v posloupnosti obsažena. Alternativní hypotézou je, že existuje alespoň jedna významná perioda.

S využitím vztahu (1) vypočteme hodnoty periodogramu  $I(\omega)$  v bodech

$$\omega_r = \frac{2\pi r}{N}, \quad r = 1, 2, \dots, m.$$

Za platnosti nulové hypotézy nebude žádná z hodnot výrazně převyšovat ostatní hodnoty periodogramu.

Srovnáme vypočítané hodnoty periodogramu  $I(\omega_1), \dots, I(\omega_m)$  sestupně dle velikosti, přičemž

$$V_1 = \max_{1 \leq r \leq m} I(\omega_r)$$

označíme největší z těchto hodnot a podobně  $V_2, \dots, V_{m-1}$  až

$$V_m = \min_{1 \leq r \leq m} I(\omega_r)$$

nejmenší hodnotu. Označme  $W$  jako tzv. *Fisherovu testovou statistiku* definovanou vztahem

$$W^{(1)} = \frac{V_1}{\sum_{r=1}^m V_r},$$

kde horní index (1) u testové statistiky  $W$  značí, že testujeme významnost hodnoty  $V_1$ . Pro tuto statistiku platí

$$W^{(1)} \in (0, 1).$$

Jestliže si budou veličiny  $V_1, \dots, V_m$  přibližně rovny, bude hodnota Fisherovy testové statistiky blízká číslu  $\frac{1}{m}$ . Jestliže ale budou hodnoty veličiny  $V_1$  vyšší



v porovnání s ostatními veličinami  $V_2, \dots, V_m$ , bude hodnota testové statistiky blízká jedné. Je tedy zřejmé, že kritický obor naší hypotézy budou tvořit velké hodnoty testové statistiky (hodnoty blízké jedné). Hustota  $p(w)$  náhodné veličiny  $W$  je podle literatury [1], str. 82, rovna

$$p(w) = m(m-1) \sum_{j=1}^{\lfloor \frac{1}{w} \rfloor} (-1)^{j-1} \binom{m-1}{j-1} (1-jw)^{m-2}, \quad (9)$$

kde  $\lfloor \frac{1}{w} \rfloor$  označuje dolní celou část čísla  $\frac{1}{w}$ . Jak můžeme vidět, hustota  $p(w) = 0$  pro  $w \geq 1$ , což lze ověřit pouhým dosazením hodnot, podrobněji viz [12], str. 191. Nulovou hypotézu zamítáme ve prospěch její alternativy, pokud  $W > x$ , kde  $x$  je kritická hodnota Fisherova testu příslušná zvolené hladině významnosti  $\alpha$ . Tedy nás zajímá  $P(W > x)$ , a proto integrujeme vztah (9) v mezích od  $x$  do 1, viz definice 5 a 6. Lze tedy psát

$$\begin{aligned} P(W > x) &= m(m-1) \int_x^1 \sum_{j=1}^{\lfloor \frac{1}{x} \rfloor} (-1)^{j-1} \binom{m-1}{j-1} (1-jw)^{m-2} dw = \\ &= m(m-1) \left[ \sum_{j=1}^{\lfloor \frac{1}{x} \rfloor} (-1)^{j-1} \frac{(m-1)!}{(j-1)!(m-j)!} \frac{1}{(-j)} \frac{(1-jw)^{m-1}}{m-1} \right]_x^1 = \\ &= \left[ \sum_{j=1}^{\lfloor \frac{1}{x} \rfloor} (-1)^j \binom{m}{j} (1-jw)^{m-1} \right]_x^1 = \sum_{j=1}^{\lfloor \frac{1}{x} \rfloor} (-1)^{j+1} \binom{m}{j} (1-jx)^{m-1} \end{aligned} \quad (10)$$

pro  $0 < x < 1$ . Vztah (10) lze explicitně přepsat následovně

$$P(W > x) = m(1-x)^{m-1} - \binom{m}{2} (1-2x)^{m-1} + \binom{m}{3} (1-3x)^{m-1} - \dots, \quad (11)$$

kde sčítáme, dokud platí  $(1-jx) > 0$ .  $\alpha$ -kritická hodnota Fisherova testu je takové číslo  $x$ , pro které platí

$$P(W^{(1)} > x) = \alpha.$$

Nulovou hypotézu  $H_0$  zamítáme na hladině významnosti  $\alpha$ , překročí-li hodnoty testové statistiky  $\alpha$ -kritickou hodnotu, tedy je-li

$$x < W^{(1)},$$

což znamená, že periodická složka o frekvenci  $\omega_1$  je významná. Tímto způsobem jsme otestovali největší hodnotu periodogramu  $V_1$ . Periodogram může obsahovat i více významných hodnot. Z tohoto důvodu, po zamítnutí nulové hypotézy pro  $V_1$ , testujeme významnost hodnoty  $V_2$  tak, že hodnotu  $V_1$  jednoduše vynecháme. Tedy  $V_2$  nyní bereme jako maximální hodnotu periodogramu. Fisherova testová statistika bude nyní tvaru

$$W^{(2)} = \frac{V_2}{\sum_{r=2}^m V_r}.$$

Opakujeme stejný postup, jako pro  $V_1$ . Pokud nulovou hypotézu opět zamítneme, pokračujeme stejným způsobem s další velkou hodnotou periodogramu, dokud nulovou hypotézu nelze zamítnout.

V mnoha případech dáváme přednost vypočítání hodnoty  $P(W > x)$  před pouhým porovnáním s  $\alpha$ -kritickou hodnotou. Výpočet podle vztahu (11) může pro  $m > 50$  způsobit jisté komplikace. R. A. Fisher dokázal, že první člen na pravé straně vztahu (11) je velmi dobrou aproximací pravděpodobnosti  $P(W > x)$ , tedy

$$P(W > x) \cong m(1 - x)^{m-1}.$$

Tuto aproximaci můžeme považovat za vhodnou pouze pro  $m \geq 50$ . V tabulce 1 (viz Přílohy) jsou uvedeny kritické hodnoty testu R. A. Fishera pro  $\alpha = 0,05$ , kde můžeme pozorovat velmi dobré aproximace této kritické hodnoty.

### 3. Nalezení významných frekvencí v nestejnoměrně vzorkovaném signálu

V mnoha oblastech vědy, například v astronomii či medicíně, je často nutné zkoumat data uspořádaná v nestejnoměrně rozložených intervalech. V podobně uspořádaném komplexu není perioda jednoznačně definována. Každý skutečný vrchol vyvolává řadu dalších vrcholů nejrůznějších výšek rozmístěných v celé oblasti frekvenčního spektra. Taková data můžeme z matematického pohledu chápat jako časovou řadu. Na spektrum této časové řady můžeme pohlížet jako na zobrazení o určitých frekvencích zachycující velikost intenzity každé frekvence, která je v řadě zastoupena, viz [3]. V důsledku toho může být stanovena nejvýše jedna konkrétní perioda pro výpočet spektra, aby se zabránilo jistému zmatku v souboru nevýznamných vrcholů příslušných tomu jedinému skutečnému vrcholu. Další periody mohou být nalezeny postupným výpočtem z periody předešlé, viz [8]. *Spektrum* časové řady je pojem, který se používá jako zjednodušení tzv. *spektrální hustoty*.

Nejběžněji používanou metodou pro výpočet spektra nerovnoměrně rozmístěných dat je tzv. Lombův periodogram, známý také jako tzv. spektrální analýza nejmenších čtverců (z anglického Least-squares spectral analysis). Podle spektra si následně můžeme udělat představu o tom, jak intenzivně se v dané časové řadě vyskytují periodické složky, viz [3].

Tato metoda je nejčastěji používána ve spojitosti se zpracováním biologických signálů, jak bude ukázáno v praktické části této bakalářské práce, viz kapitola 4.

#### 3.1. Vaníčková metoda

Lombův periodogram navazuje a zjednodušuje tzv. Vaníčkovu metodu pojmenovanou po významném česko-kanadském geodetovi Petru Vaníčkově.

Petr Vaníček se narodil v Sušici roku 1935. Těsně před sovětskou invazí v roce 1968 opustil spolu se svou rodinou Československo a odjel do Anglie, kde začal pracovat na univerzitě v Liverpoolu. Do Kanady emigroval roku 1969. Podstat-

ným úspěchem bylo, po více než třiceti letech pedagogické profese na univerzitách v Torontu a New Brunswicku, získání zaslouženého titulu „professor emeritus“ v roce 2002. Dodnes žije ve městě Fredericton spolu se svou rodinou, viz [20].

Roku 1971 Petr Vaníček publikoval článek, v němž popsal novou metodu spektrální analýzy. Jedním z hlavních požadavků této metody je schopnost odstranit systematický šum z časové řady za podmínky minimálního narušení spektra zbývající řady, jak bylo uvedeno v [4].

Vaníčková metoda pracuje se souborem dat diskrétního typu, jenž je aproximován váženým součtem sinusoid o postupně stanovených frekvencích.

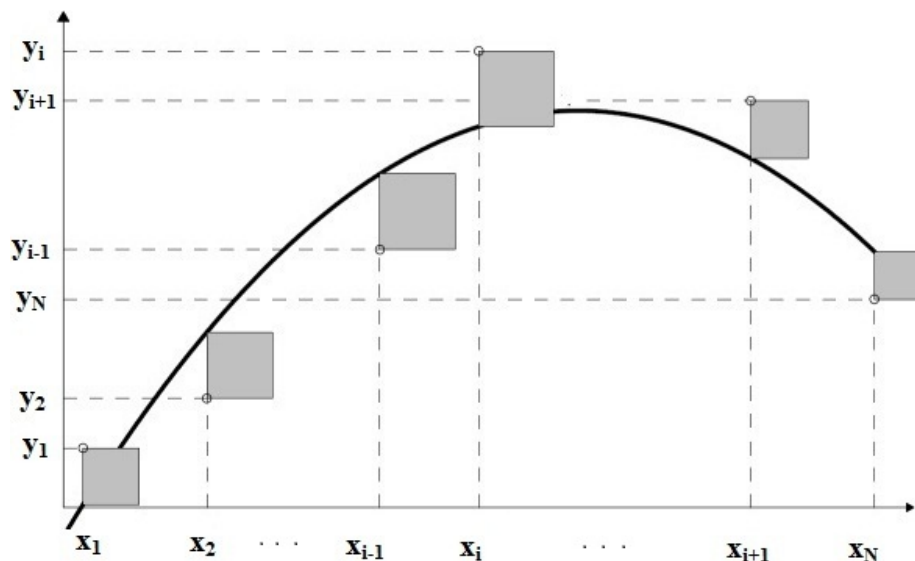
## 3.2. Lombův periodogram

Lombův periodogram, neboli spektrální analýza nejmenších čtverců, poskytuje způsob odhadu frekvenčního spektra. Metoda je významným vylepšením Fourierovy analýzy, nejpoužívanější spektrální metody v oblasti vědy. Jistou nevýhodou Fourierovy analýzy je například zvýšený výskyt šumu při dlouhých mezerách v záznamu signálu. Tento problém výrazně redukuje použití Lombova periodogramu.

V celé této kapitole čerpáme z článků [8], [13] a [17]. Tyto články poskytují stručný přehled problematiky Lombova periodogramu a použitých vztahů této metody. Proto budeme uvádět i postup s bližším vysvětlením, což je jedním z přínosů naší bakalářské práce, a povede tak k snadnějšímu pochopení uvedené problematiky budoucím čtenářům.

### 3.2.1. Metoda nejmenších čtverců

Metoda nejmenších čtverců je matematická metoda, používaná v numerice a statistice, která se využívá při zpracování dat získaných například měřeními. Hlavním smyslem metody je aproximace (prokládání) naměřených datových hodnot funkcí, jež nemusí přímo procházet zadanými body. V této podkapitole čerpáme z literatury [9].



Obr. 3.1: Grafické znázornění metody nejmenších čtverců

Nechť jsou zadané body  $x_i$ ,  $i = 1, 2, \dots, N$  a naměřené hodnoty v těchto bodech  $y_i = f(x_i)$ ,  $i = 1, 2, \dots, N$ , kde  $f$  je neznámá funkce. Uvažujme funkce  $\varphi_1(x), \dots, \varphi_m(x)$ . Princip metody nejmenších čtverců spočívá v nalezení funkce tvaru

$$P_m(x) = \sum_{j=1}^m c_j \varphi_j(x)$$

tak, aby výraz

$$\rho^2(f, P_m) = \rho^2(c_1, \dots, c_m) = \sum_{i=1}^N w_i [f(x_i) - P_m(x_i)]^2 \quad (12)$$

byl minimální. Hledání funkce  $P_m(x)$  je ekvivalentní s hledáním neznámých konstant  $c_1, \dots, c_m$ . Hodnoty  $w_i$ ,  $i = 1, \dots, N$  nazýváme váhy. Váhy jsou kladné konstanty, jež klademe většinou rovny 1. U bodů, které jsou naměřeny s vyšší přesností než ostatní, lze klást hodnoty vah větší než jedna. Tyto konstanty si uživatel volí dle svého uvážení. Pro naše případy, analýzu EKG signálu v další kapitole, budeme váhu  $w_i$  pokládat rovnu 1 pro každé  $i = 1, 2, \dots, N$ .

**Definice 14.** Funkce  $P_m(x) = \sum_{j=1}^m c_j \varphi_j(x)$ , pro kterou výraz  $\rho^2(f, P_m)$  nabývá své minimální hodnoty se nazývá *nejlepší aproximací funkce  $f(x)$  ve smyslu metody nejmenších čtverců*.

Případy, kdy  $m \geq N$  se řeší pomocí interpolace, tudíž předpokládejme  $m \ll N$ . Budeme hledat konstanty  $c_1, \dots, c_m$ , pro které by výraz (12) nabýval své nejmenší hodnoty. Hledáme tedy extrémy funkce  $\rho^2(c_1, \dots, c_m)$ . Jelikož musíme nalézt stacionární body funkce  $\rho^2$ , položíme

$$\frac{\partial \rho^2}{\partial c_k} = 0 \quad \forall k = 1, 2, \dots, m.$$

Tedy dostáváme

$$\begin{aligned} \frac{\partial \rho^2}{\partial c_1} &= -2 \sum_{i=1}^N w_i [f(x_i) - \sum_{j=1}^m c_j \varphi_j(x_i)] \varphi_1(x_i) = 0 \\ &\vdots \\ \frac{\partial \rho^2}{\partial c_m} &= -2 \sum_{i=1}^N w_i [f(x_i) - \sum_{j=1}^m c_j \varphi_j(x_i)] \varphi_m(x_i) = 0. \end{aligned}$$

Uvažujme nyní  $k$ -tou rovnicí a upravíme ji následujícím způsobem.

$$\begin{aligned} \sum_{i=1}^N w_i [f(x_i) - \sum_{j=1}^m c_j \varphi_j(x_i)] \varphi_k(x_i) &= 0 \\ \sum_{i=1}^N w_i f(x_i) \varphi_k(x_i) - \sum_{i=1}^N w_i \sum_{j=1}^m c_j \varphi_j(x_i) \varphi_k(x_i) &= 0. \end{aligned}$$

Odtud dostáváme

$$\sum_{j=1}^m c_j \sum_{i=1}^N w_i \varphi_j(x_i) \varphi_k(x_i) = \sum_{i=1}^N w_i f(x_i) \varphi_k(x_i),$$

což lze psát

$$\sum_{j=1}^m c_j (\varphi_j, \varphi_k) = (f, \varphi_k),$$

kde  $(f, g)$  představuje *diskrétní skalární součin* definovaný vztahem

$$(f, g) = \sum_{i=1}^N w_i f(x_i) g(x_i).$$

Pokud tímto způsobem upravíme všechny rovnice, dostáváme tzv. *normální soustavu*, jejíž determinant se nazývá *Grammův determinant*. V našem případě má normální soustava tvar

$$c_1(\varphi_1, \varphi_1) + c_2(\varphi_2, \varphi_1) + \dots + c_m(\varphi_m, \varphi_1) = (f, \varphi_1)$$

$$c_1(\varphi_1, \varphi_2) + c_2(\varphi_2, \varphi_2) + \dots + c_m(\varphi_m, \varphi_2) = (f, \varphi_2)$$

⋮

$$c_1(\varphi_1, \varphi_m) + c_2(\varphi_2, \varphi_m) + \dots + c_m(\varphi_m, \varphi_m) = (f, \varphi_m).$$

**Věta 2.** Necht' funkce  $\varphi_j(x)$ ,  $j = 1, 2, \dots, m$  jsou lineárně nezávislé na množině bodů  $\{x_i\}_{i=1}^N$ . Pak normální soustava má jediné řešení  $c_1^*, \dots, c_m^*$  a funkce  $P_m(x) = \sum_{j=1}^m c_j^* \varphi_j(x)$  minimalizuje výraz  $\rho^2(c_1, \dots, c_m)$ , tj. jedná se o nejlepší aproximaci funkce  $f(x)$  na množině bodů  $\{x_i\}_{i=1}^N$  ve smyslu metody nejmenších čtverců.

**Důkaz:** viz [9].

### 3.2.2. Vztahy pro spektrum získané metodou nejmenších čtverců

Necht' jsou dány časové body  $t_i$ . Uvažujme soubor  $N$  nezávislých pozorování  $y_i = f(t_i)$ ,  $i = 1, 2, \dots, N$ , s nulovou střední hodnotou. Hledáme funkci  $P(t)$  ve tvaru

$$P(t) = a \cos(2\pi ft) + b \sin(2\pi ft),$$

kde  $a$  a  $b$  jsou neznámé konstanty a  $f$  je daná frekvence. Ve smyslu metody nejmenších čtverců chceme minimalizovat výraz

$$\rho^2(a, b) = \sum_{i=1}^N (a \cos(2\pi ft_i) + b \sin(2\pi ft_i) - y_i)^2 \quad (13)$$

(viz podkapitola 3.2.1). Abychom dostali soustavu normálních rovnic, zderivujeme výraz (13) podle jednotlivých proměnných, tedy

$$\frac{\partial \rho^2(a, b)}{\partial a} = 2 \sum_{i=1}^N (a \cos(2\pi f t_i) + b \sin(2\pi f t_i) - y_i) \cos(2\pi f t_i)$$

$$\frac{\partial \rho^2(a, b)}{\partial b} = 2 \sum_{i=1}^N (a \cos(2\pi f t_i) + b \sin(2\pi f t_i) - y_i) \sin(2\pi f t_i).$$

Derivace položíme rovny nule, upravíme a vyjádříme maticovou formou. Normální soustava je tedy ve tvaru

$$\begin{pmatrix} CC & CS \\ CS & SS \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} YC \\ YS \end{pmatrix},$$

kde kvůli přehlednosti a zjednodušení následného upravování výrazů definujeme následující značení:

$$CC = \sum_{i=1}^N \cos^2(2\pi f t_i), \quad SS = \sum_{i=1}^N \sin^2(2\pi f t_i),$$

$$CS = \sum_{i=1}^N \cos(2\pi f t_i) \sin(2\pi f t_i),$$

$$YC = \sum_{i=1}^N y_i \cos(2\pi f t_i), \quad YS = \sum_{i=1}^N y_i \sin(2\pi f t_i).$$

Nyní dokážeme, že matice normální soustavy je regulární. Využijeme tvrzení, že matice je regulární právě tehdy, když je determinant této matice různý od nuly. Označme  $D$  determinant matice normální soustavy, tedy platí

$$D = CC \cdot SS - (CS)^2$$

$$D = \sum_{i=1}^N \cos^2(2\pi f t_i) \sum_{i=1}^N \sin^2(2\pi f t_i) - \left( \sum_{i=1}^N \cos(2\pi f t_i) \sin(2\pi f t_i) \right)^2.$$



S použitím Cauchy-Schwarzovy nerovnosti je

$$\sum_{i=1}^N \cos^2(2\pi ft_i) \sum_{i=1}^N \sin^2(2\pi ft_i) \geq \left( \sum_{i=1}^N \cos(2\pi ft_i) \sin(2\pi ft_i) \right)^2,$$

tedy  $D \neq 0$ .

Jelikož je matice normální soustavy regulární a je určena jednoznačně, existuje právě jedno řešení normální soustavy. Z uvedených vlastností plyne existence inverzní matice normální soustavy. Následující postup povede ke tvaru periodogramu, popsaného v předchozí kapitole. Definujme nyní *redukcí v součtu čtverců* vztahem

$$\begin{aligned} \Delta R(f) &:= (YC \ YS) \begin{pmatrix} a \\ b \end{pmatrix} = (YC \ YS) \begin{pmatrix} CC & CS \\ CS & SS \end{pmatrix}^{-1} \begin{pmatrix} YC \\ YS \end{pmatrix} = \\ &= (YC \ YS) \begin{pmatrix} \frac{SS}{D} & -\frac{CS}{D} \\ -\frac{CS}{D} & \frac{CC}{D} \end{pmatrix} \begin{pmatrix} YC \\ YS \end{pmatrix}. \end{aligned} \quad (14)$$

Pro numerickou práci je jednodušší vyjádření vztahu (14) ve tvaru  $A^2 + B^2$ , což usnadní statistický popis spektra nejmenších čtverců v následující podkapitole. Pro tento přepis využijeme vyjádření dat ve tvaru

$$P(t) = a \cos(2\pi f(t - \tau)) + b \sin(2\pi f(t - \tau)), \quad (15)$$

místo

$$P(t) = a \cos(2\pi ft) + b \sin(2\pi ft),$$

kde  $\tau$  je časové zpoždění definované vztahem

$$\tan(4\pi f\tau) = \frac{\sum_{i=1}^N \sin(4\pi ft_i)}{\sum_{i=1}^N \cos(4\pi ft_i)},$$

viz [17].  $\tau$  volíme tak, aby platilo  $CS = 0$ . Dosazením do vztahu (14) dostáváme

$$\Delta R(f) = (YC \ YS) \begin{pmatrix} \frac{1}{CC} & 0 \\ 0 & \frac{1}{SS} \end{pmatrix} \begin{pmatrix} YC \\ YS \end{pmatrix},$$

protože platí

$$\frac{SS}{D} = \frac{SS}{CC \cdot SS} = \frac{1}{CC}$$

$$\frac{CC}{D} = \frac{CC}{CC \cdot SS} = \frac{1}{SS},$$

kde nyní, například

$$CC = \sum_{i=1}^N \cos^2(2\pi f(t_i - \tau)).$$

A tedy konečně můžeme psát

$$\Delta R(f) = \left( \frac{1}{\sqrt{CC}} YC \right)^2 + \left( \frac{1}{\sqrt{SS}} YS \right)^2. \quad (16)$$

Vzhledem k dalším numerickým úpravám vztahu (16) v následující sekci, označme jednotlivé jeho členy následovně:

$$C(f) = \frac{1}{\sqrt{CC}} YC, \quad S(f) = \frac{1}{\sqrt{SS}} YS. \quad (17)$$

Podíváme-li se na vztah (16), můžeme pozorovat jistou podobnost s periodogramem uvedeným v lemmatu 1. Ve skutečnosti vztah (3) pro periodogram je aproximace vztahu (16). Použijeme-li dvou předpokladů: pro všechny hodnoty  $\tau$  je  $CS = 0$  a pokud navíc  $CC = SS = \frac{n}{2}$ , můžeme vztah (16) převést na tvar

$$\Delta R(f) = \left( \sqrt{\frac{2}{n}} YC \right)^2 + \left( \sqrt{\frac{2}{n}} YS \right)^2.$$

Tato formule je používaná v analýze periodogramu, jak jsme mohli vidět v předchozím textu, viz lemma 1.

Položme

$$R = \sum_{i=1}^N y_i^2,$$

pak definujme *normovanou spektrální funkci* vztahem

$$p(f) = \frac{\Delta R(f)}{R} = \frac{C^2(f) + S^2(f)}{R},$$

kde po dosazení všech pomocných výrazů můžeme psát

$$p(f) = \frac{1}{\sum_{i=1}^N y_i^2} \left( \frac{[\sum_{i=1}^N y_i \cos(2\pi f(t_i - \tau))]^2}{\sum_{i=1}^N \cos^2(2\pi f(t_i - \tau))} + \frac{[\sum_{i=1}^N y_i \sin(2\pi f(t_i - \tau))]^2}{\sum_{i=1}^N \sin^2(2\pi f(t_i - \tau))} \right) \quad (18)$$

Posledním důležitým bodem spektrální analýzy je test významnosti vrcholů ve spektru  $p(f)$ . Z [17] plyne, že důležitým aspektem je přítomnost Gaussova náhodného šumu v naměřených hodnotách. Jestliže se data skládají z tohoto šumu, má spektrální funkce (18) exponenciální rozdělení pravděpodobnosti. Jinými slovy, pravděpodobnost, že  $p(f)$  bude v rozmezí mezi nějakou kladnou hodnotou  $z$  a  $z + dz$  je  $e^{-z}dz$ , viz [13]. Abychom nejvýznamnější hodnotu spektra mohli určit, vypočítáme nejprve tzv. *falešný poplach pravděpodobnosti* (dále jen *FAP* z anglického false alarm probability) daný vztahem

$$FAP = 1 - (1 - e^{-z})^M,$$

kde  $M$  je počet nezávislých frekvencí spektra. Závěr je takový, že pokud jsou hodnoty *FAP* malé, indikuje tento fakt vysokou významnost hodnoty frekvence spektra.

### 3.2.3. Spektrum vzhledem k sinusovému signálu

Sinusový signál o frekvenci  $f_1$  může být reprezentován vztahem

$$y_i = a \cos(2\pi f_1(t_i - \tau_1)) + b \sin(2\pi f_1(t_i - \tau_1)),$$

který koresponduje s funkcí  $P(t)$  danou vztahem (15) pro  $f_1$  a  $\tau_1$ , viz podkapitola 3.2.2. Jestliže  $y_i$  jsou naše pozorování, můžeme pro kteroukoli frekvenci  $f_2$  dosadit tento vztah do vztahu pro  $C(f)$ , tedy dostaneme

$$C(f_2) = \frac{1}{\sqrt{C_2 C_2}} Y C_2,$$

kde zavedeným značením rozumíme

$$C_2 C_2 = \sum_{i=1}^N \cos^2(2\pi f_2(t_i - \tau_2)),$$

$$YC_2 = \sum_{i=1}^N y_i \cos(2\pi f_2(t_i - \tau_2)).$$

$YC_2$  dále upravíme dosazením vztahu pro  $y_i$  (viz výše)

$$\begin{aligned} YC_2 &= \sum_{i=1}^N [a \cos(2\pi f_1(t_i - \tau_1)) + b \sin(2\pi f_1(t_i - \tau_1))] \cos(2\pi f_2(t_i - \tau_2)) = \\ &= aC_1C_2 + bS_1C_2. \end{aligned}$$

Tedy po dosazení do vztahu pro  $C(f_2)$  dostáváme

$$C(f_2) = \frac{1}{\sqrt{C_2C_2}} (aC_1C_2 + bS_1C_2),$$

kde

$$\begin{aligned} C_1C_2 &= \sum_{i=1}^n \cos(2\pi f_1(t_i - \tau_1)) \cos(2\pi f_2(t_i - \tau_2)), \\ S_1C_2 &= \sum_{i=1}^n \sin(2\pi f_1(t_i - \tau_1)) \cos(2\pi f_2(t_i - \tau_2)). \end{aligned}$$

Podobně lze psát

$$S(f_2) = \frac{1}{\sqrt{S_2S_2}} (aC_1S_2 + bS_1S_2).$$

Pokles v součtu čtverců pro frekvenci  $f_2$  dostaneme užitím vztahu (16).

$$\begin{aligned} \Delta R(f_2) &= C^2(f_2) + S^2(f_2) = \frac{1}{C_2C_2} (aC_1C_2 + bS_1C_2)^2 + \frac{1}{S_2S_2} (aC_1S_2 + bS_1S_2)^2 = \\ &= a^2 \left( \frac{(C_1C_2)^2}{C_2C_2} + \frac{(C_1S_2)^2}{S_2S_2} \right) + b^2 \left( \frac{(S_1C_2)^2}{C_2C_2} + \frac{(S_1S_2)^2}{S_2S_2} \right) + \\ &\quad + 2ab \left( \frac{C_1C_2 \cdot S_1C_2}{C_2C_2} + \frac{C_1S_2 \cdot S_1S_2}{S_2S_2} \right). \end{aligned}$$

V dalším postupu provedeme jisté úpravy, které v důsledku povedou k formuli pro normované spektrum. Rozšíříme postupně jednotlivé závorky ve vztahu  $\Delta R(f_2)$  členy

$$\frac{C_1C_1}{C_1C_1}, \quad \frac{S_1S_1}{S_1S_1}, \quad \frac{\sqrt{C_1C_1 \cdot S_1S_1}}{\sqrt{C_1C_1 \cdot S_1S_1}}$$

a dostaneme

$$\begin{aligned}
\Delta R(f_2) &= a^2 \frac{C_1 C_1}{C_1 C_1} \left( \frac{(C_1 C_2)^2}{C_2 C_2} + \frac{(C_1 S_2)^2}{S_2 S_2} \right) + b^2 \frac{S_1 S_1}{S_1 S_1} \left( \frac{(S_1 C_2)^2}{C_2 C_2} + \frac{(S_1 S_2)^2}{S_2 S_2} \right) + \\
&\quad + 2ab \frac{\sqrt{C_1 C_1 \cdot S_1 S_1}}{\sqrt{C_1 C_1 \cdot S_1 S_1}} \left( \frac{C_1 C_2 \cdot S_1 C_2}{C_2 C_2} + \frac{C_1 S_2 \cdot S_1 S_2}{S_2 S_2} \right) = \\
&= a^2 C_1 C_1 \left( \frac{(C_1 C_2)^2}{C_1 C_1 \cdot C_2 C_2} + \frac{(C_1 S_2)^2}{C_1 C_1 \cdot S_2 S_2} \right) + \\
&\quad + b^2 S_1 S_1 \left( \frac{(S_1 C_2)^2}{S_1 S_1 \cdot C_2 C_2} + \frac{(S_1 S_2)^2}{S_1 S_1 \cdot S_2 S_2} \right) + \\
&\quad + 2ab \sqrt{C_1 C_1 \cdot S_1 S_1} \left( \frac{C_1 C_2 \cdot S_1 C_2}{\sqrt{C_1 C_1 \cdot (C_2 C_2)^2 \cdot S_1 S_1}} + \frac{C_1 S_2 \cdot S_1 S_2}{\sqrt{C_1 C_1 \cdot (S_2 S_2)^2 \cdot S_1 S_1}} \right).
\end{aligned}$$

Definujme nyní následovně koeficient

$$\rho_{C_1 C_2} = \frac{C_1 C_2}{\sqrt{C_1 C_1 \cdot C_2 C_2}},$$

a stejným způsobem definujeme i ostatní koeficienty  $\rho_{C_1 S_2}$ ,  $\rho_{S_1 S_2}$ , atd. Dále označme koeficienty  $\alpha$  a  $\beta$  následovně

$$\alpha = a\sqrt{C_1 C_1}, \quad \beta = b\sqrt{S_1 S_1}.$$

S použitým označením dostáváme

$$\Delta R(f_2) = \alpha^2 (\rho_{C_1 C_2}^2 + \rho_{C_1 S_2}^2) + \beta^2 (\rho_{S_1 C_2}^2 + \rho_{S_1 S_2}^2) + 2\alpha\beta (\rho_{C_1 C_2} \rho_{S_1 C_2} + \rho_{C_1 S_2} \rho_{S_1 S_2}).$$

Jestliže  $\tau_1$  volíme tak, aby  $C_1 S_1 = 0$ , pak celkový součet čtverců je dán vztahem

$$R = \sum_{i=1}^N y_i^2 = (\alpha + \beta)^2 = a^2 C_1 C_1 + b^2 S_1 S_1 + 2ab C_1 S_1 = \alpha^2 + \beta^2.$$

Užitím vztahu (18) dostáváme formuli pro *normované spektrum*, tj.

$$\begin{aligned}
p(f_2) = \frac{\Delta R(f_2)}{R} &= \frac{\alpha^2}{\alpha^2 + \beta^2} (\rho_{C_1 C_2}^2 + \rho_{C_1 S_2}^2) + \frac{\beta^2}{\alpha^2 + \beta^2} (\rho_{S_1 C_2}^2 + \rho_{S_1 S_2}^2) + \\
&\quad + \frac{2\alpha\beta}{\alpha^2 + \beta^2} (\rho_{C_1 C_2} \rho_{S_1 C_2} + \rho_{C_1 S_2} \rho_{S_1 S_2}). \tag{19}
\end{aligned}$$

Hodnoty tohoto spektra se budou mírně lišit v závislosti na změně  $\alpha$  a  $\beta$ , což koresponduje se změnou fáze signálu. Abychom mohli stanovit střední hodnotu spektra, položíme

$$\sin^2 \omega = \frac{\alpha^2}{\alpha^2 + \beta^2}, \quad \cos^2 \omega = \frac{\beta^2}{\alpha^2 + \beta^2}.$$

Dosazením do vztahu (19) máme

$$\begin{aligned} p(f_2) = & (\rho_{C_1 C_2}^2 + \rho_{C_1 S_2}^2) \sin^2 \omega + (\rho_{S_1 C_2}^2 + \rho_{S_1 S_2}^2) \cos^2 \omega + \\ & + (\rho_{C_1 C_2} \rho_{S_1 C_2} + \rho_{C_1 S_2} \rho_{S_1 S_2}) \sin 2\omega. \end{aligned} \quad (20)$$

Vztah (20) souhrně popisuje spektrum datového souboru vzhledem k sinusovému signálu. Pro stanovení kvalitního zobrazení tvaru spektra sinusového signálu je nutné zjednodušení vztahů pomocí jistých aproximací. Užitím předpokladů, stejně jako v předchozí podkapitole, tj. že pro všechny hodnoty  $\tau$  je  $CS = 0$  a  $CC = SS = \frac{n}{2}$ , můžeme dosáhnout přesnější aproximace. Tak lze pozorovat přiblížení k periodogramu, jenž byl popsán v podkapitole 3.2.2.

## 4. Srovnání frekvenčních charakteristik reálných EKG signálů

Jedním z cílů naší bakalářské práce je aplikovat postup Lombova periodogramu, popsaného v předchozí kapitole, k získání frekvenčních charakteristik na reálných EKG signálech pacientů, kteří prodělali tzv. synkopu.

*Synkopa* je náhlá, krátkodobá ztráta vědomí způsobená nedostatkem okysličené krve v mozku (tzv. hypoxií). Hypoxie bývá důsledkem poklesu krevního tlaku. Synkopa může nastat zcela bez varování i u zcela zdravých lidí následkem citového vypětí, jako je silné vzrušení či strach, nebo při dlouhém stání, viz [22].

V současné době se vědci domnívají, že EKG záznam pacienta se synkopou v sobě může zahrnovat jistou informaci, jež by mohla dát lékařům impulz, že by u tohoto pacienta v budoucnu mohly nastat jiné kardiologické potíže, a je tedy nutné aplikovat jistá medicínská opatření.

Lombův periodogram je jednou z metod odhadu tzv. *výkonové spektrální hustoty* (vyjadřuje výkon obsažený v úzkém intervalu spektra, dále jen PSD z anglického power spectral density) v nepravidelně vzorkovaném signálu, jako je např. časová řada okamžité srdeční frekvence, která je přirozeně rozložena do nepravidelných intervalů. Standardní odhady PSD vyžadují převzorkování na stejnoměrně vzorkované intervaly. V tomto ohledu poskytuje Lombův periodogram výrazné vylepšení, protože umožňuje získat tento odhad přímo z nepravidelně rozloženého signálu a vyhnout se tak problémům spojeným se složitým převzorkováním signálu, viz [10].

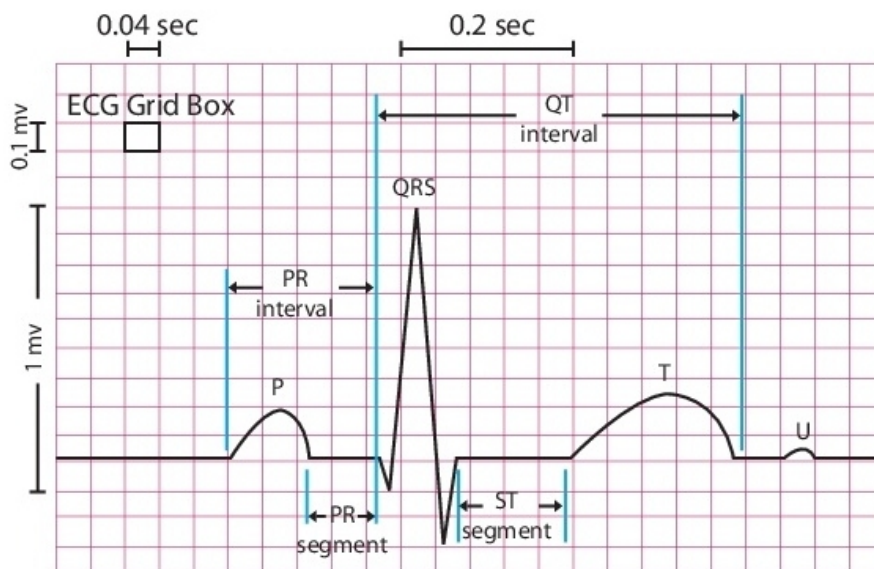
Programovou stránku této bakalářské práce tvoříme v matematickém programu Matlab a využíváme návody a některé programové části z [23]. Jelikož je programový materiál velice rozsáhlý, nebudeme se zabývat jeho vysvětlením, ale spíše přímou aplikací a popisem na datovém souboru. Veškeré kódy s příslušnými komentáři jsou přiloženy na CD, viz Přílohy.

## 4.1. Popis datového souboru

Pro účely této bakalářské práce máme k dispozici 22 reálných záznamů EKG signálů zcela anonymních pacientů, kteří prodělali synkopu. Tyto EKG záznamy byly pořízeny v rozmezí 24 hodin přístrojem zvaným EKG holter. Tato vyšetřovací metoda umožňuje sledování srdeční aktivity pacienta po celý den a noc, a to i během běžných denních činností. Tím je možné zachytit abnormality, které by se při ambulantním EKG vyšetření nemusely projevit. Výsledkem tohoto vyšetření je tedy 24 hodinový záznam EKG křivky, což je podrobně popsáno v [19]. Vzhledem k velikosti datového souboru a citlivosti údajů pacientů nebudou data, ze kterých čerpáme, přiložena na CD.

V dalším textu budeme pro názornost uvádět veškeré aplikace naší programové práce na EKG záznamu prvního pacienta. Aplikace na ostatních pacientech jsou uvedeny v Příloze.

Nyní se podrobněji zaměříme na popis EKG křivky, jež je zobrazena na Obr. 4.1. Pro naše účely se omezíme pouze na stručný přehled prvků v EKG záznamu, protože podrobné pochopení struktury a funkce EKG přesahuje rámec této bakalářské práce.



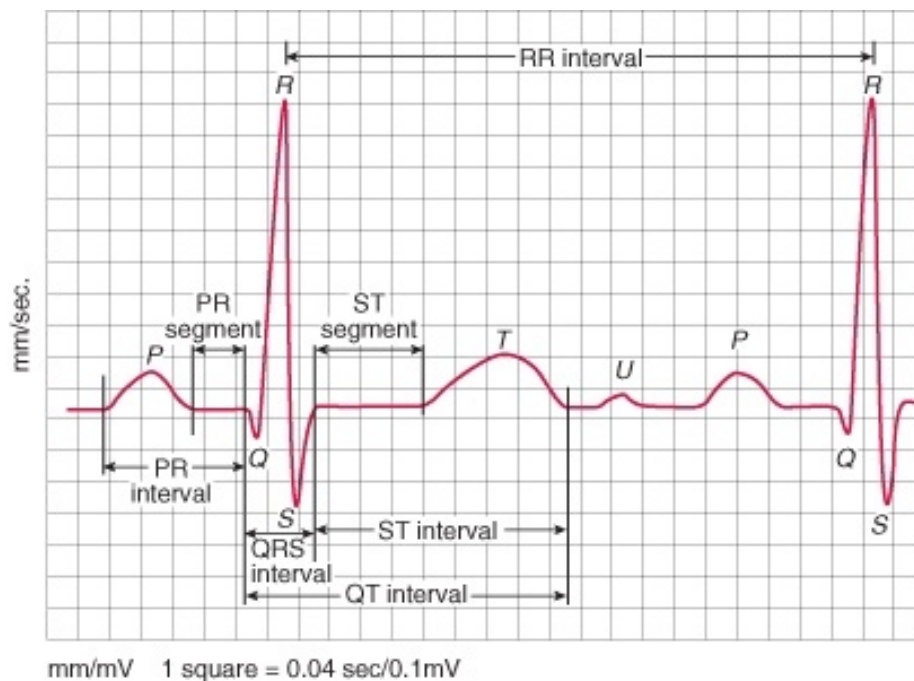
Obr. 4.1: Podrobné znázornění EKG křivky



Celé EKG, tedy jednotlivé kmity a vlny, vychází z depolarizace a repolarizace srdečního svalu. Základní křivka EKG se skládá z vlny P, T a U, komplexu kmitů Q, R a S, a intervalů PR (někdy označovaný PQ), QT, segmentů PR a ST. Význam vlny P spočívá v depolarizaci síní. Podstatou QRS komplexu je depolarizace komor. Repolarizaci komor vyjadřuje vlna T, viz [24].

#### 4.1.1. Stanovení RR intervalů

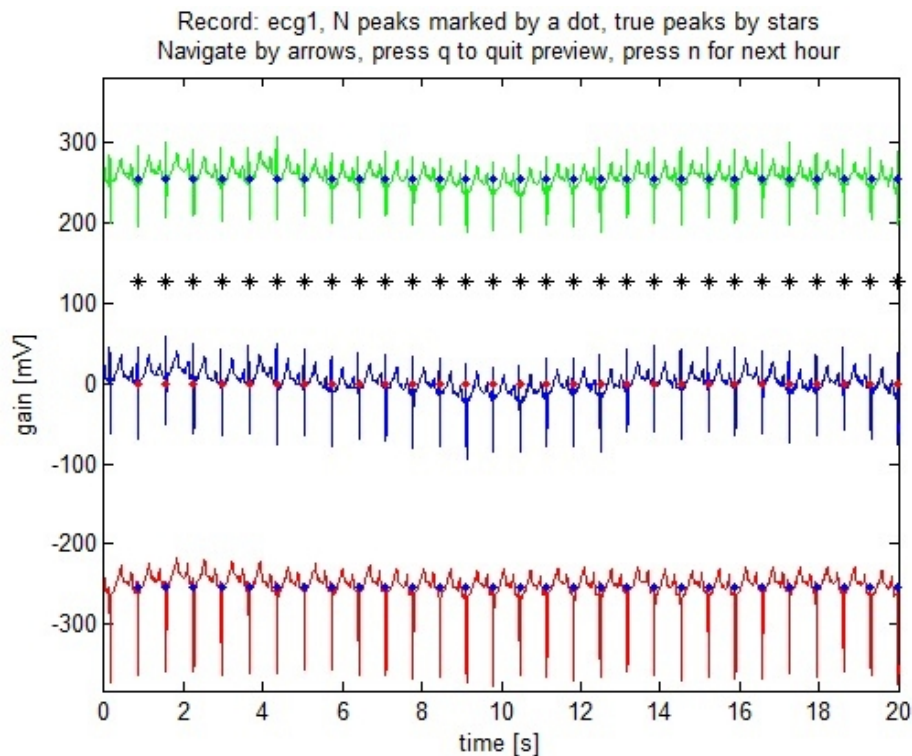
Prvním a velmi důležitým krokem pro stanovení frekvenčních charakteristik EKG signálů je vymezení tzv. RR intervalů, popř. NN intervalů. NN intervaly (z anglického normal to normal) jsou časové intervaly mezi po sobě jdoucími srdečními stahy od začátku QRS komplexu do začátku dalšího QRS komplexu, RR intervaly (z anglického R to R) jsou časové intervaly od kmitu R do dalšího kmitu R, jak můžeme podrobně vidět na Obr. 4.2.



Obr. 4.2: Grafické znázornění RR intervalu v EKG signálu

Jelikož EKG snímá elektrickou aktivitu srdce, jednotkou signálu, se kterým budeme pracovat, jsou  $V$  (volty). V našem případě to budou  $mV$  (milivolty, platí

$1mV = 10^{-3}V$ ). Vzorkovací frekvence EKG je  $200\text{ Hz}$ . EKG signál přijímáme ve třech kanálech, z nichž každý kanál obsahuje své R kmity, jež jsou na Obr. 4.3 znázorněny formou teček. Obr. 4.3 je výstupem naší programové části v Matlabu a je ukázkou, jak vypadá vstupní signál EKG prvního pacienta během prvních 20 s měření. Tento záznam samozřejmě pokračuje po dobu celých 24 hodin, což je možné si prohlédnout na přiloženém CD.



Obr. 4.3: Grafické znázornění přijímaného EKG signálu prvního pacienta na 3 kanálech

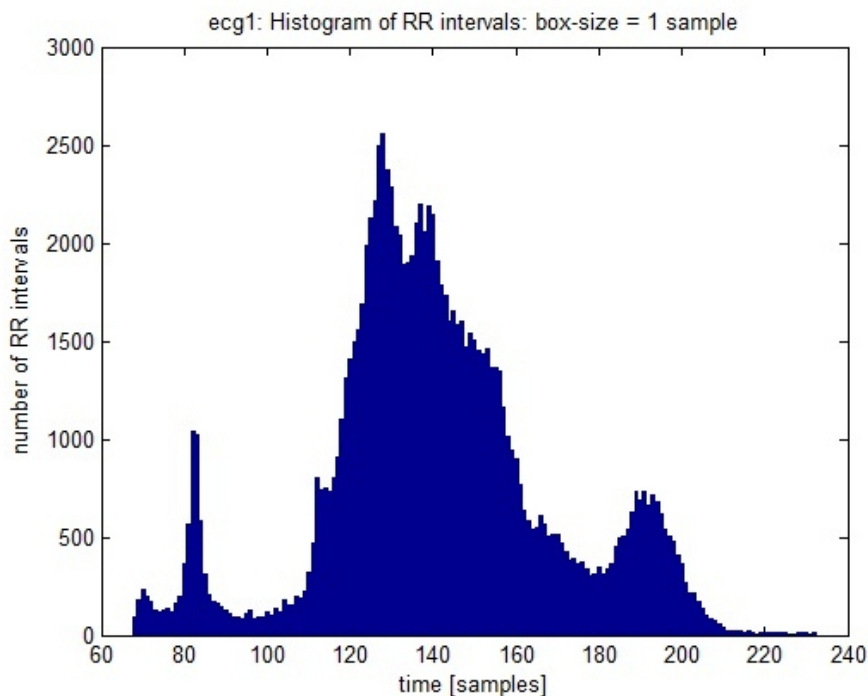
Na základě hodnot R kmitů v čase jsme obdrželi systém RR intervalů pro každý kanál. Tyto systémy je nyní nutné sjednotit do jednoho. Během záznamu může docházet k různým chybám měření (chyba přístroje, lidský faktor - pohybové artefakty, atd.), jež se v záznamu projeví například výpadkem signálu na jednom z kanálů. Abychom tedy mohli sjednocený R kmit považovat za „pravdivý“, požadujeme, aby se v daném čase objevil alespoň na dvou kanálech. Tyto pravdivé kmity (v matlabovském programu označovány jako *true peaks*) jsou na Obr. 4.3

znázorněny formou hvězdiček. Tímto postupem jsme obdrželi posloupnost RR intervalů pro každý EKG signál.

Dalším důležitým krokem je odfiltrování příliš dlouhých a naopak příliš krátkých RR intervalů, které znázorňují chyby při měření EKG. Přehled smazaných dlouhých, krátkých a celkový počet výsledných intervalů je uveden v Tabulce 2b (viz Přílohy). Abychom měli kompletní přehled pro EKG prvního pacienta, uvádíme tento přehled již zde.

EKG	Dlouhé intervaly	Krátké intervaly	Počet RR intervalů
ekg 1	2 621	12 629	110 395

Tabulka 2a: Přehled smazaných a výsledných RR intervalů EKG prvního pacienta



Obr. 4.4: Histogram výsledných RR intervalů EKG prvního pacienta

Někdy je výhodné mít k dispozici jistý grafický přehled RR intervalů, na-

příklad formou histogramu. Touto výhodou může být například rychlý pohled na počet stanovených intervalů při jejich bližší analýze. Generovaný histogram z matlabovského kódu můžeme vidět na Obr. 4.4.

## 4.2. Frekvenční charakteristiky srdečního rytmu

Nyní přejdeme ke stanovení frekvenčních charakteristik našich EKG signálů. Variabilita srdečního rytmu (z anglického heart rate variability - HRV) je určena ke sledování odchylek v intervalech mezi jednotlivými stahy srdce, viz [27]. Může souviset s mnoha onemocněními. Změna této variability může poukazovat na onemocnění, aniž by pacient zaznamenal jakékoli příznaky. Variabilitu srdečního rytmu hodnotíme dvěma způsoby, viz [5],

1. *časovou metodou analýzy posloupnosti RR intervalů* (tzv. time domain analysis),
2. *spektrální analýzou řady RR intervalů* (tzv. frequency domain analysis).

### Časová analýza

Tento druh analýzy EKG signálu je podstatně jednodušší oproti spektrální metodě. Základním smyslem je identifikace časové polohy QRS komplexů, z nichž se následně určí NN intervaly (viz 4.1.1.). Dle [16] lze časovou analýzu rozdělit do dvou podkategorií, statistickou a geometrickou. *Geometrická metoda* využívá grafické zpracování NN, resp. RR intervalů (např. histogram, viz Obr. 4.4). *Statistická metoda* užívá frekvenční charakteristiky založené na statistických aparátech, mezi základní patří:

- AVNN (ms): průměr všech NN intervalů.  
SDNN (ms): směrodatná odchylka všech NN intervalů.  
SDANN (ms): směrodatná odchylka průměru NN intervalů ve všech 5-minutových segmentech 24-hodinového záznamu.  
ASDNN (ms): průměr směrodatných odchylek všech NN intervalů.  
RMSSD (ms): odmocnina průměru druhých mocnin rozdílů mezi po sobě jdoucími NN intervaly.  
PNN50 (%): procentuální zastoupení rozdílů mezi po sobě jdoucími NN intervaly, které se liší o více než 50 ms.

Vypočítané hodnoty těchto charakteristik na našich EKG signálech jsou uvedeny v tabulce 3b (viz Přílohy). Opět pro přehlednost uvádíme v tabulce 3a charakteristiky EKG signálu prvního pacienta. Jelikož se nejedná o hlavní programovou část naší práce, využili jsme pro vypočítání těchto charakteristik funkce z [23].

EKG	AVNN	SDNN	SDANN	ASDNN	RMSSD	PNN50
ekg 1	706,886	138,508	114,169	68,295	101,300	0,111

Tabulka 3a: Charakteristiky získané statistickou metodou EKG prvního pacienta

### Spektrální analýza

Nyní se zaměříme na spektrální analýzu řady RR intervalů. Tyto časové hodnoty jsou transformovány do frekvenční oblasti. Odhad spektrální hustoty nám poslouží k získání informace o rozložení spektrálního výkonu v závislosti na frekvenci. Frekvenční spektrum srdečního rytmu lze zaznamenat v několika oblastech:

- TPF ( $ms^2$ ): oblast celkové frekvence ( $0 - 0,4 Hz$ ).  
ULF ( $ms^2$ ): oblast „ultra nízké“ frekvence ( $0 - 0,0033 Hz$ ).  
VLF ( $ms^2$ ): oblast „velmi nízké“ frekvence ( $0,0033 - 0,04 Hz$ ).  
LF ( $ms^2$ ): oblast nízké frekvence ( $0,04 - 0,15 Hz$ ).  
HF ( $ms^2$ ): oblast vysoké frekvence ( $0,15 - 0,4 Hz$ ).  
LF/HF: poměr LF a HF.

Každá z těchto oblastí koresponduje s jistými fyziologickými procesy v lidském organismu, jako je např. termoregulace či dýchání. V některých studiích se

jako jednotky preferují tzv. normalizované jednotky, jak je uvedeno v [5]. Tyto jednotky představují poměr jednotlivých komponent k celkové hodnotě spektra po odečtení VLF, tedy platí

$$LF_{norm} = \frac{LF(ms^2) \cdot 100}{TPF(ms^2) - VLF(ms^2)}, \quad HF_{norm} = \frac{HF(ms^2) \cdot 100}{TPF(ms^2) - VLF(ms^2)}.$$

Tyto charakteristiky určíme později, jelikož k jejich výpočtu potřebujeme znát spektrální výkon EKG signálů, který vypočítáme pomocí Lombova periodogramu.

#### 4.2.1. Aplikace Lombova periodogramu na EKG signálech

V této podkapitole se zaměříme na popis aplikace Lombova periodogramu, popsaného v kapitole 3.2., na EKG signálech našich pacientů. Tuto metodu využíváme ke stanovení významných frekvencí v oblasti spektra EKG signálů.

Konkrétní výpočet Lombova periodogramu provádíme programem *lomb.m*, který byl implementován na základě popsaného postupu v kapitole 3.2. Tento program jsme převzali z oficiálních internetových stránek matematického programu Matlab, viz [28].

Před vlastním výpočtem je nutné převést hodnoty RR intervalů na *ms* (milisekundy), čímž docílíme toho, že spektrální výkon bude vycházet v jednotkách  $ms^2$ . Jako vstupní hodnoty tohoto programu zadáváme vektor RR intervalů. Jelikož se velikost tohoto vektoru pohybuje v řádech stovek tisíců hodnot, což působí jisté výpočetní problémy s pamětí počítače, budeme počítat s vektorem menší velikosti. Nyní je potřeba určit počet bodů – frekvencí, které budeme pomocí Lombova periodogramu zkoumat, jestli se v signálu vyskytují, a vytvořit vektor těchto frekvencí. Pro naše účely jsme vygenerovali vektor frekvencí v rozsahu 0–1500 *Hz* s krokem 0,48 *Hz*, tedy celkem 3126 hodnot. Na stejný počet hodnot se omezíme také ve vektoru RR intervalů.

Výstupní parametry programu *lomb.m* jsou tři. Parametr *prob* je vektor hodnot *FAP*, *f* udává frekvenci [*Hz*] a parametr *P* obsahuje hodnoty spektrálního výkonu [ $ms^2$ ]. Program porovnává hodnoty vektoru *FAP* a vybere index *jmax* s nejmenší hodnotou *FAP*, dále najde hodnotu s indexem *jmax* ve vektoru frek-

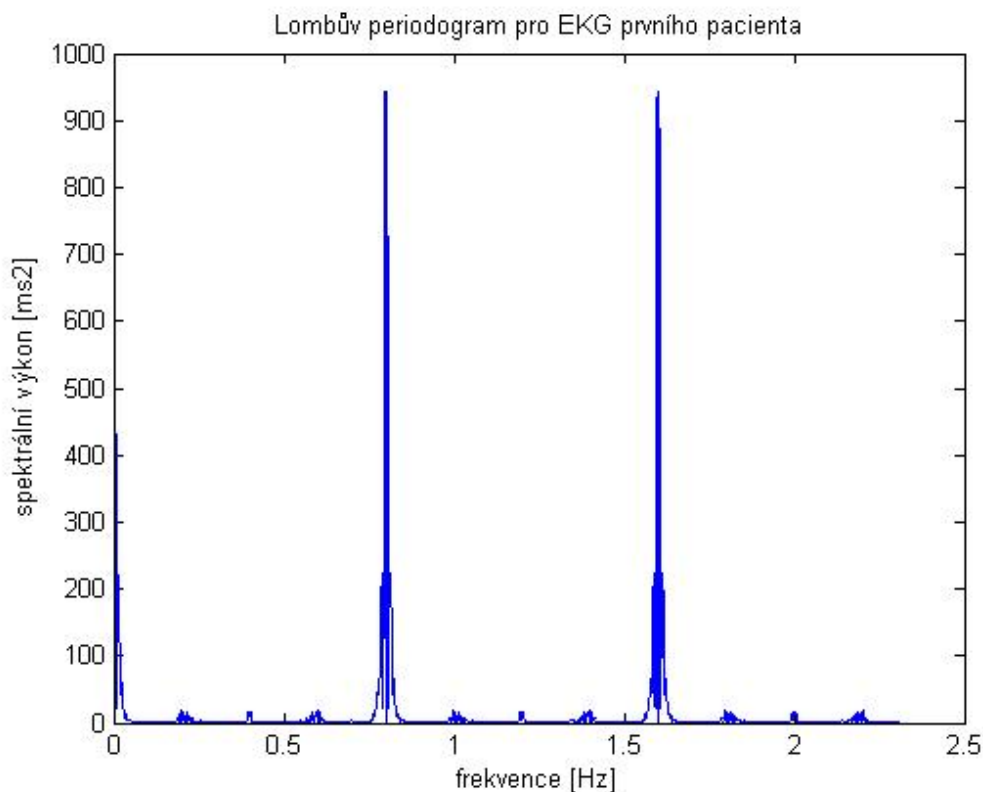
vencí, což je jeden z požadovaných výsledků, jedná se totiž o nejméně významnější frekvenci. Stejným způsobem vybereme hodnotu  $P_{max}$  z vektoru  $P$ , což je největší z hodnot spektrálního výkonu.

Výstup naší práce, tedy vypočítané hodnoty nejméně významnějších frekvencí  $f$ , period  $T$ ,  $FAP$ ,  $P_{max}$  a  $j_{max}$  pro EKG prvního pacienta je uvedeno v tabulce 4a. Grafické znázornění spektrálního výkonu v závislosti na frekvenci EKG prvního pacienta lze pozorovat na Obr. 4.4.

Ostatní výsledky jsou uvedeny v Příloze.

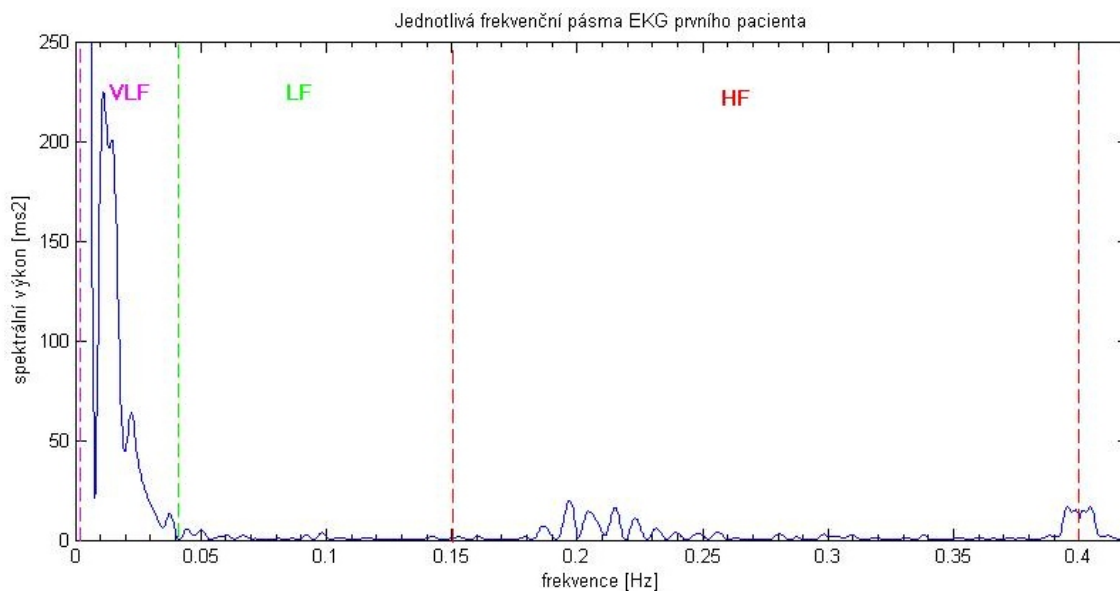
EKG	Nejméně významnější frekvence $f$	Nejméně významnější periody $T$	$FAP$	$P_{max}$	$j_{max}$
ekg 1	1,602	0,624	0	942,284	4341

Tabulka 4a: Nejméně významnější frekvence, periody,  $FAP$ ,  $P_{max}$  a  $j_{max}$  EKG signálu prvního pacienta



Obr. 4.4: Spektrální výkon EKG prvního pacienta

Pokud se na spektrální výkon podíváme podrobněji ve frekvenčním pásmu 0–0,4 Hz, můžeme identifikovat jednotlivé složky v oblastech frekvenčního spektra srdečního rytmu, viz Obr. 4.5. Integrací těchto složek v příslušných oblastech získáme spektrální charakteristiky. Vypočítané hodnoty spektrálních charakteristik našich EKG signálů jsou uvedeny v tabulce 5b (viz Přílohy). Opět pro přehlednost uvádíme v tabulce 5a charakteristiky EKG signálu prvního pacienta.



Obr. 4.5: Jednotlivá frekvenční pásma srdečního rytmu prvního pacienta

EKG	TPF	ULF	VLF	LF	HF	LF/HF
ekg1	17 514,700	12 087,100	1 678,660	1 436,290	2 312,710	0,621

Tabulka 5a: Spektrální analýza EKG prvního pacienta

Aby tyto hodnoty korespondovaly s hodnotami ve výše uvedeném grafu, je nutné je převést na normalizované jednotky, tedy

$$LF_{norm} = \frac{LF(ms^2) \cdot 100}{TPF(ms^2) - VLF(ms^2)} = \frac{1436,290 \cdot 100}{17514,700 - 1678,660} = 9,069,$$

$$HF_{norm} = \frac{HF(ms^2) \cdot 100}{TPF(ms^2) - VLF(ms^2)} = \frac{2312,710 \cdot 100}{17514,700 - 1678,660} = 14,604.$$



Abychom měli představu o hodnotách těchto frekvenčních charakteristik EKG signálů, uvedeme několik srovnání, viz [26]. Pro dospělou zdravou osobu se typické hodnoty LF a HF pohybují kolem stovek až tisíc  $ms^2$ . Hodnoty ULF a VLF jsou obecně vyšší, pohybují se v hodnotách tisíců  $ms^2$ . Důležitou charakteristikou je poměr LF a HF, který by se u zdravého dospělého člověka měl pohybovat v rozmezí 1,5 – 4,5.

## Závěr

Cílem naší práce bylo popsat metody zpracování periodických signálů a metody, pomocí kterých se testuje významnost frekvencí na ekvidistantně i neekvidistantně vzorkovaném signálu.

Takovou metodu pro stejnoměrně vzorkované signály představuje periodogram a k testování významnosti frekvencí jsme uvedli test R. A. Fishera. Dále jsme popsali metodu Lombova periodogramu, používanou ve spektrální analýze nestejně rozložených signálů, která poskytuje značné vylepšení oproti metodě Fourierovy transformace, protože nevyžaduje převzorkování na stejnoměrně rozložený signál. Tato metoda byla podrobně matematicky prozkoumána a bylo uvedeno velké množství odvození jednotlivých vztahů, jež povede ke snadnějšímu pochopení uvedené problematiky budoucím čtenářům. Tuto metodu jsme aplikovali na reálné EKG signály pro získání významných frekvencí a frekvenčních charakteristik (např. podíl LF a HF).

Nejvýznamnější frekvence a procentuální podíl LF a HF jsou jedny z nejpožívanějších prostředků k hodnocení variability tepové frekvence ve frekvenční oblasti. Spektrální analýza je běžně užívána v klinické praxi k stanovení či posouzení stavu pacienta (např. po infarktu myokardu k posouzení dalších možných rizik). Medicínské posouzení těchto hodnot je velmi složité z hlediska znalosti anatomie lidského těla, což vede k dalšímu možnému pokračování výzkumu a rozšíření získaných výsledků. Podle [26] lze usoudit, že pouze pět pacientů z našeho testovaného souboru dvaceti dvou pacientů je zdravých, což vyplývá z podmínky pro poměr LF a HF (viz tabulka 5b).

## Přílohy

V této kapitole uvádíme tabulky a grafy, které byly prezentovány v průběhu celé práce. První tabulka uvádí hodnoty Fisherova testu, která je převzata z [1]. Další tabulky jsou již výstupem našich matematických výpočtů z programu Matlab. Všechny hodnoty jsou uvedeny se třemi desetinnými místy. Jedinou výjimku tvoří hodnota  $FAP$  v tabulce 4b, kde vypočítané hodnoty uvádíme s vyšší přesností. Přesnost těchto hodnot byla nastavena použitím programu *lomb.m*.

m	5%-ní kritická hodnota	Aproximace kritické hodnoty
5	0,68377	0,68377
10	0,44495	0,44495
15	0,33462	0,33463
20	0,27040	0,27046
25	0,22805	0,22813
30	0,19784	0,19794
35	0,17513	0,17525
40	0,15738	0,15752
45	0,14310	0,14324
50	0,13135	0,13149

Tabulka 1: Kritické hodnoty Fisherova testu pro  $\alpha = 0,05$  a její aproximace

EKG	Dlouhé intervaly	Krátké intervaly	Počet RR intervalů
ekg 1	2 621	12 629	110 395
ekg 2	380	4 736	91 849
ekg 3	11 477	807	83 513
ekg 4	263	832	124 643
ekg 5	1 948	5 335	97 667
ekg 6	881	4 151	105 249
ekg 7	2 946	5 941	102 675
ekg 8	547	973	100 333
ekg 9	716	667	106 143
ekg 10	6 735	7 120	59 862
ekg 11	634	9 981	83 421
ekg 12	1 302	15 040	61 336
ekg 13	378	3 216	82 344
ekg 14	1 657	6 978	120 265
ekg 15	1 217	1 618	97 639
ekg 16	200	16 608	91 976
ekg 17	283	808	62 760
ekg 18	414	1 292	93 908
ekg 19	892	36 309	76 557
ekg 20	236	4 382	72 933
ekg 21	3 974	5 845	125 859
ekg 22	204	77	112 887

Tabulka 2b: Přehled smazaných a výsledných RR intervalů EKG signálů

EKG	AVNN	SDNN	SDANN	ASDNN	RMSSD	PNN50
ekg 1	706,886	138,508	114,169	68,295	101,300	0,111
ekg 2	815,094	144,309	120,393	74,695	57,487	0,125
ekg 3	817,588	207,104	138,510	149,326	215,555	0,409
ekg 4	687,340	72,219	65,339	25,669	28,506	0,010
ekg 5	718,008	137,018	122,297	43,803	78,682	0,079
ekg 6	797,097	166,692	150,178	79,253	69,900	0,144
ekg 7	771,432	136,672	109,291	70,910	95,756	0,150
ekg 8	736,143	151,159	142,124	49,354	53,162	0,051
ekg 9	801,815	231,933	208,318	95,049	64,619	0,238
ekg 10	694,564	177,029	81,503	129,816	220,479	0,371
ekg 11	891,403	211,457	180,058	127,968	100,538	0,374
ekg 12	1 131,430	196,374	131,766	115,981	216,804	0,457
ekg 13	996,958	106,934	80,775	57,914	85,055	0,117
ekg 14	669,383	189,242	169,778	73,839	76,935	0,212
ekg 15	759,485	185,564	160,725	93,071	136,294	0,203
ekg 16	866,628	151,432	134,549	65,011	35,327	0,078
ekg 17	770,049	153,073	136,760	63,345	46,490	0,100
ekg 18	767,330	158,124	150,070	54,925	34,680	0,077
ekg 19	953,960	207,471	157,660	111,025	202,378	0,335
ekg 20	992,613	165,131	141,306	81,196	77,919	0,307
ekg 21	628,387	89,967	72,407	48,385	50,092	0,056
ekg 22	760,310	79,211	067,262	33,775	47,770	0,029

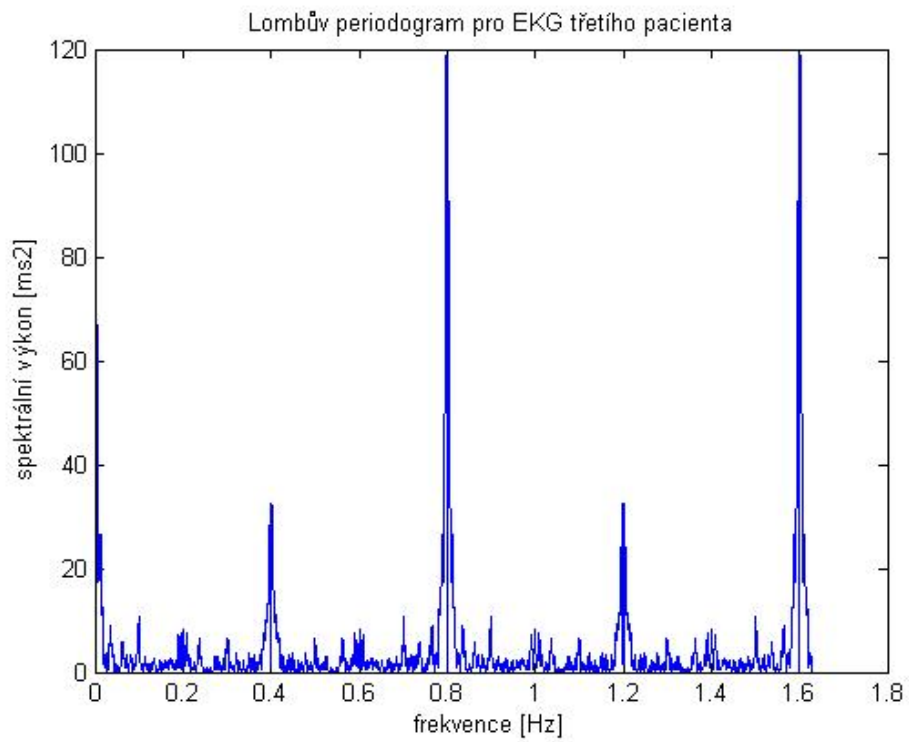
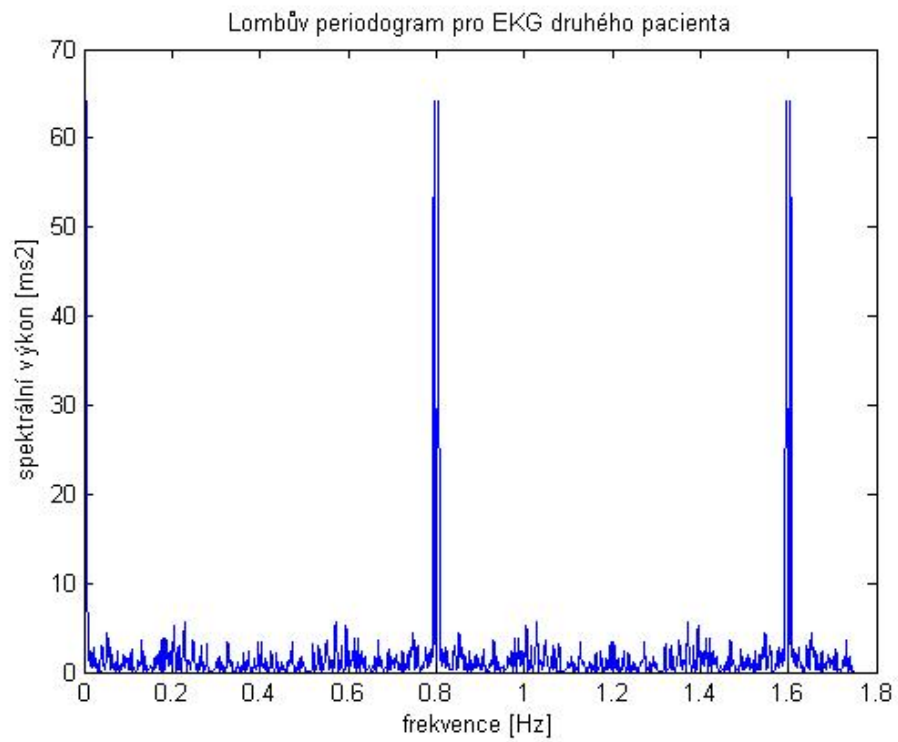
Tabulka 3b: Charakteristiky získané statistickou metódou

EKG	Nejvýznamější frekvence $f$	Nejvýznamější periody $T$	$FAP$	$P_{max}$	$j_{max}$
ekg 1	1,602	0,624	0	942,284	4341
ekg 2	1,605	0,623	3,883e-25	64,256	5744
ekg 3	0,798	1,252	8,756e-049	118,705	3066
ekg 4	8,791	0,114	5,702e-265	616,492	6066
ekg 5	0,007	142,174	1,482e-096	228,703	23
ekg 6	1,598	0,626	3,246e-147	345,350	5769
ekg 7	0,798	1,253	7,702e-100	236,265	2906
ekg 8	0,795	1,258	2,959e-036	89,856	2702
ekg 9	0,802	1,247	1,719e-044	108,820	2904
ekg 10	0,797	1,254	4,876e-134	315,010	3712
ekg 11	0,799	1,252	7,697e-297	689,874	2836
ekg 12	0,006	181,379	2,491e-067	161,408	29
ekg 13	0,796	1,257	1,449e-110	260,961	3660
ekg 14	1,599	0,625	2,454e-126	297,276	4773
ekg 15	1,596	0,627	0	838,865	5666
ekg 16	0,798	1,253	3,143e-116	242,002	2467
ekg 17	0,801	1,248	5,892e-235	547,381	2937
ekg 18	1,602	0,624	3,824e-048	117,230	5471
ekg 19	0,003	397,500	5,531e-191	446,130	10
ekg 20	0,802	1,247	1,548e-038	95,109	3706
ekg 21	1,594	0,627	6,879e-270	627,817	4718
ekg 22	1,594	0,627	2,804e-063	152,079	5261

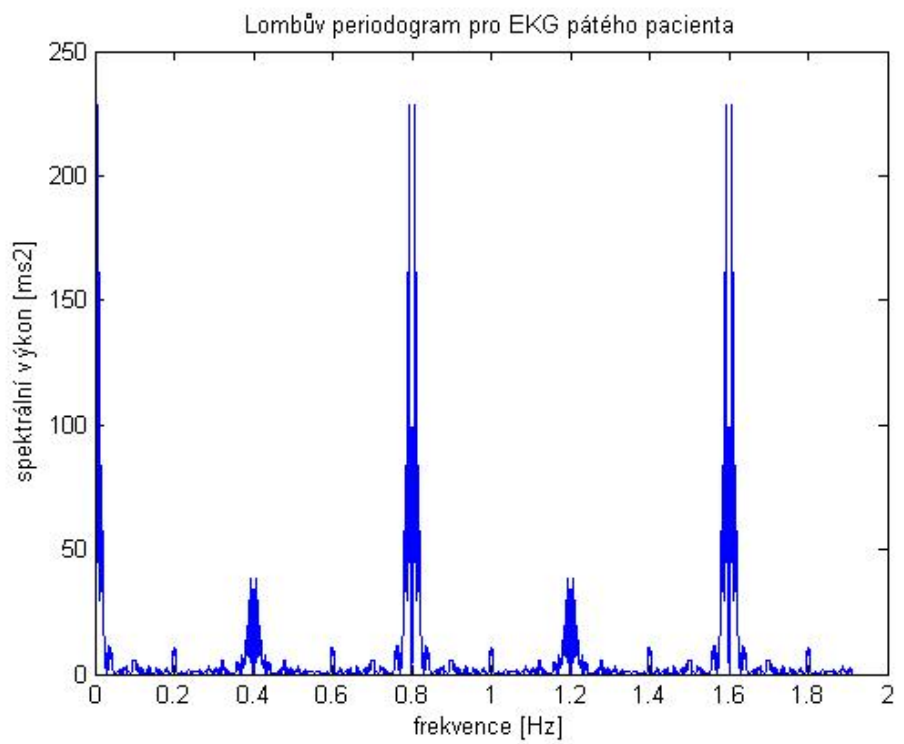
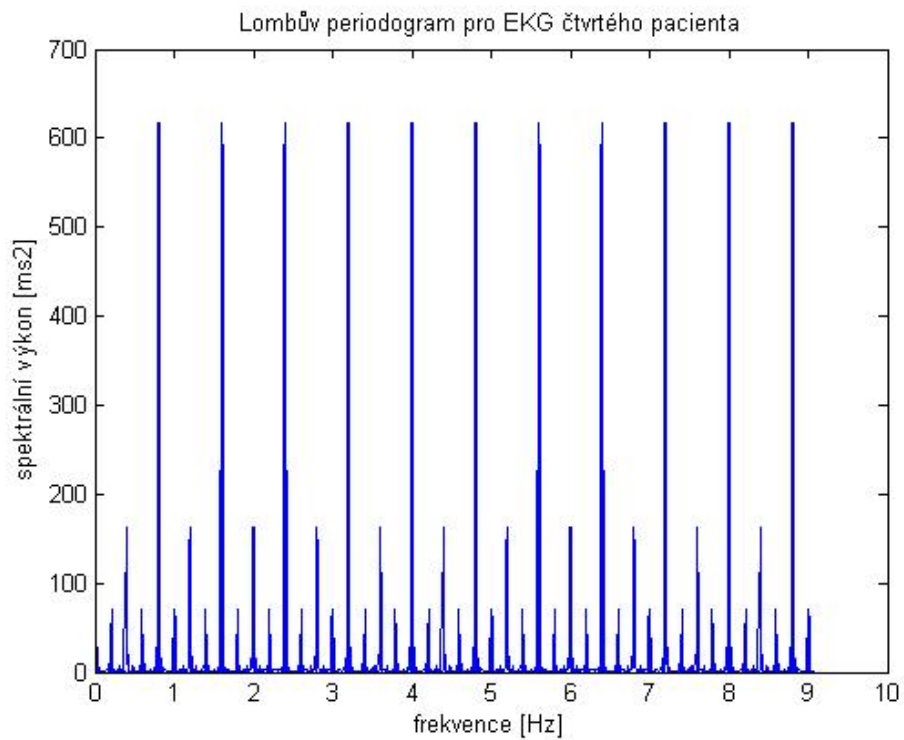
Tabulka 4b: Nejvýznamější frekvence, periody,  $FAP$ ,  $P_{max}$  a  $j_{max}$  EKG signálů

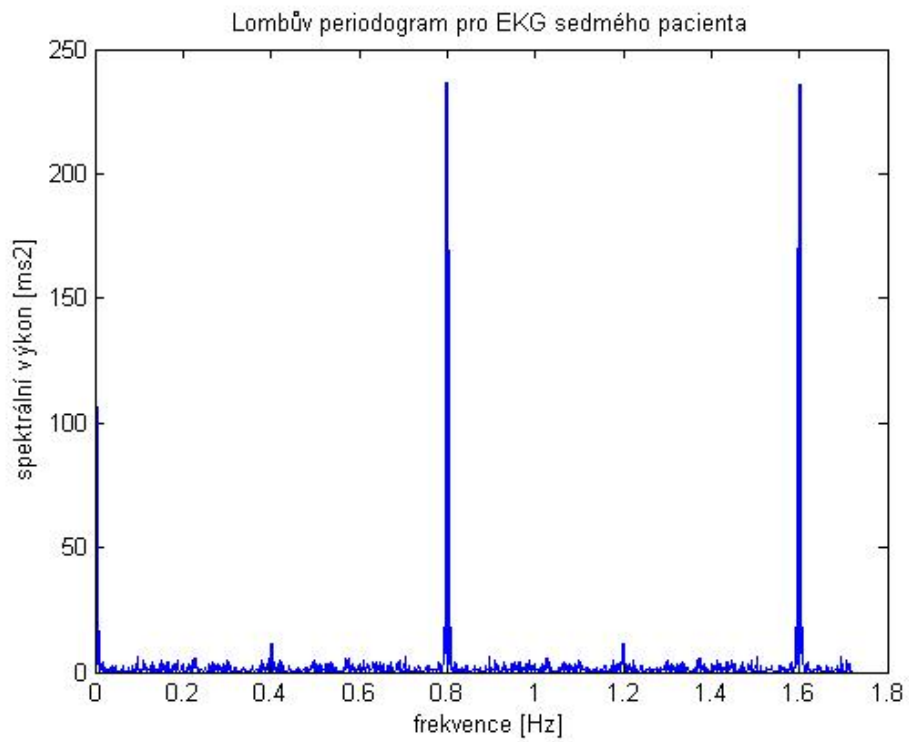
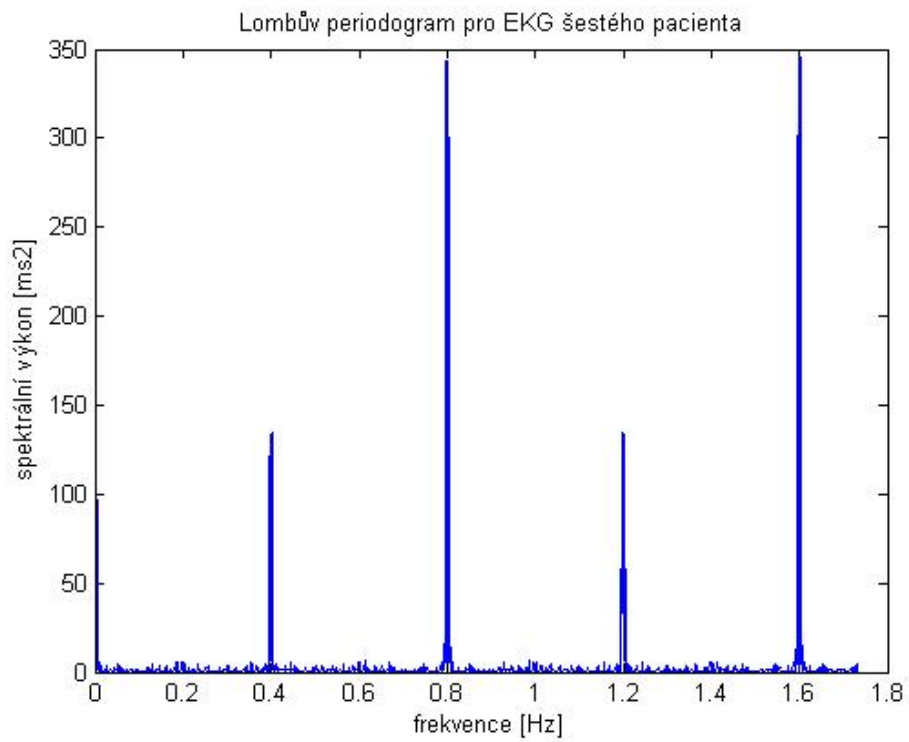
EKG	TPF	ULF	VLF	LF	HF	LF/HF
ekg1	17 514,700	12 087,100	1 678,660	1 436,290	2 312,710	0,621
ekg2	21 665,000	16 051,300	2 773,730	1 930,180	909,860	2,121
ekg3	38 192,700	22 615,800	2 254,260	3 521,790	9 800,860	0,359
ekg4	5 210,900	4 444,200	420,290	170,270	176,140	0,967
ekg5	19 375,200	16 398,400	828,930	704,120	1 443,720	0488
ekg6	27 497,000	21 970,700	3 097,260	1 199,660	1 229,380	0,976
ekg7	17 991,500	12 799,000	1 447,920	1 546,890	2 197,660	0,704
ekg8	23 484,700	20 751,800	1 284,980	684,140	763,840	0,896
ekg9	57 237,100	49 390,700	5 106,790	1 654,000	1 085,690	1,524
ekg10	27 474,100	7 121,500	2 621,240	5 308,160	12 423,200	0,427
ekg11	47 334,600	35 114,400	5 882,550	3 114,960	3 222,660	0,967
ekg12	38 318,600	19 275,700	3 242,110	5 053,170	10 747,600	0,470
ekg13	11 914,900	7 637,800	1 357,530	908,050	2 011,530	0,451
ekg14	37 903,400	32 314,800	2 067,280	1 727,250	1 794,160	0,963
ekg15	32 043,800	25 106,700	1 866,670	1 445,890	3 624,520	0,399
ekg16	24 568,200	20 275,300	2 663,220	1 252,520	377,200	3,321
ekg17	26 805,800	23 349,900	1 786,820	1 046,160	623,030	1,679
ekg18	26 567,300	23 656,000	1 970,290	650,340	290,720	2,237
ekg19	42 055,700	27 392,700	2 852,820	3 343,230	8 466,920	0,395
ekg20	28 570,700	21 412,500	3 826,620	1 531,030	1 800,490	0,850
ekg21	7 204,200	5 413,100	826,260	528,540	436,260	1,212
ekg22	6 363,800	5 478,600	336,020	174,840	374,330	0,467

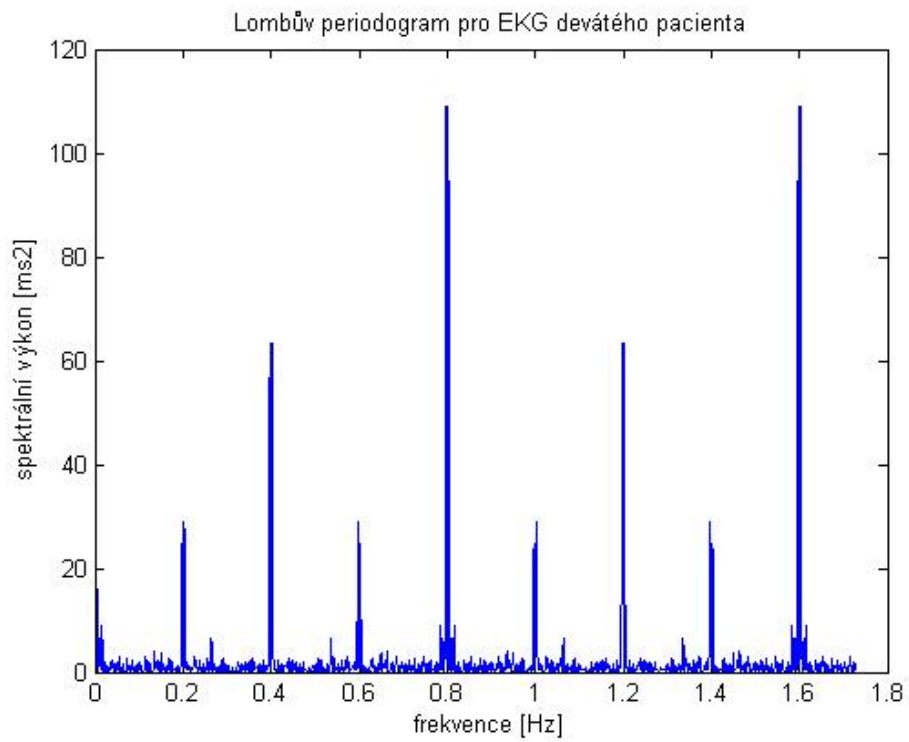
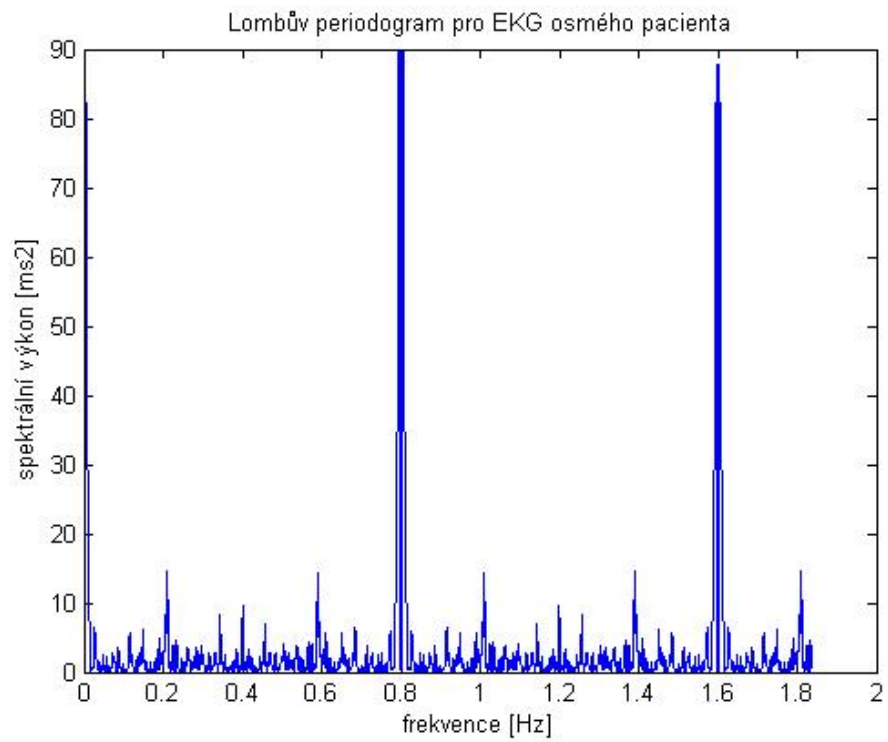
Tabulka 5b: Spektrální analýza

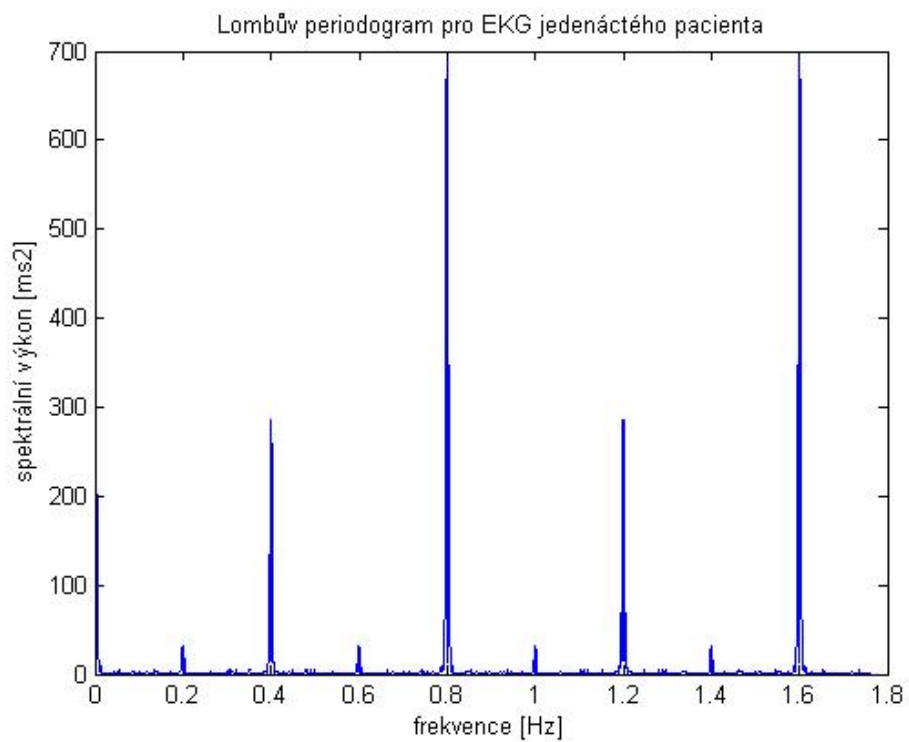
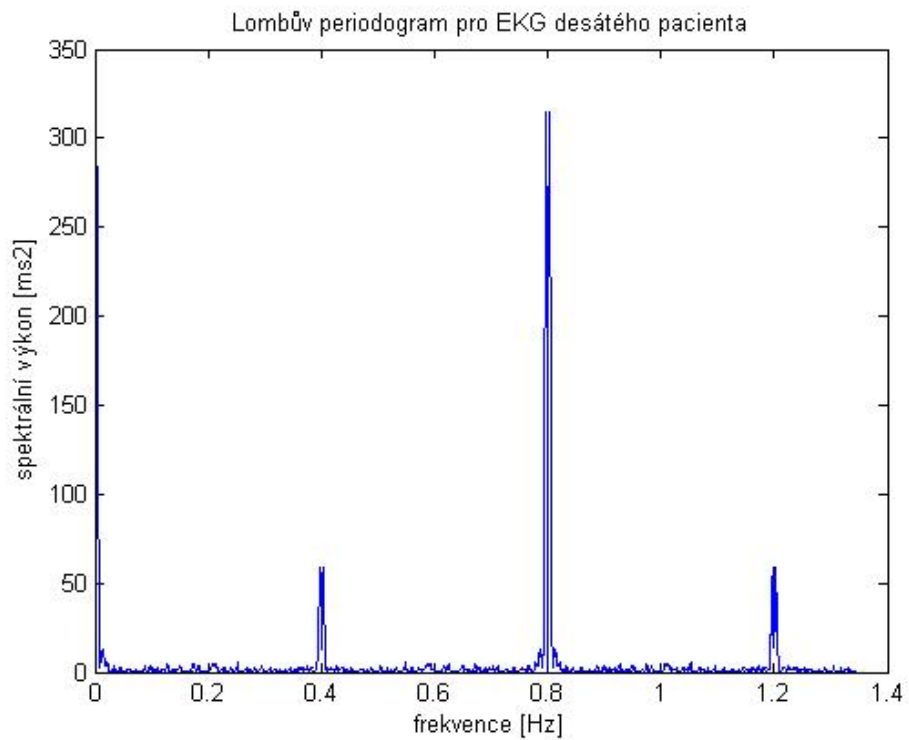


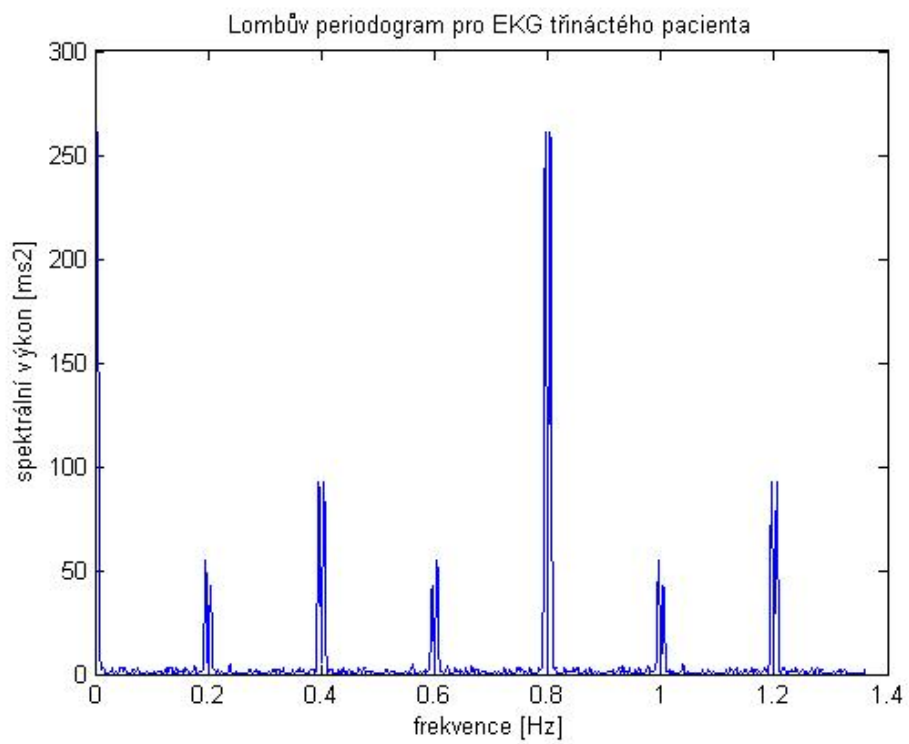
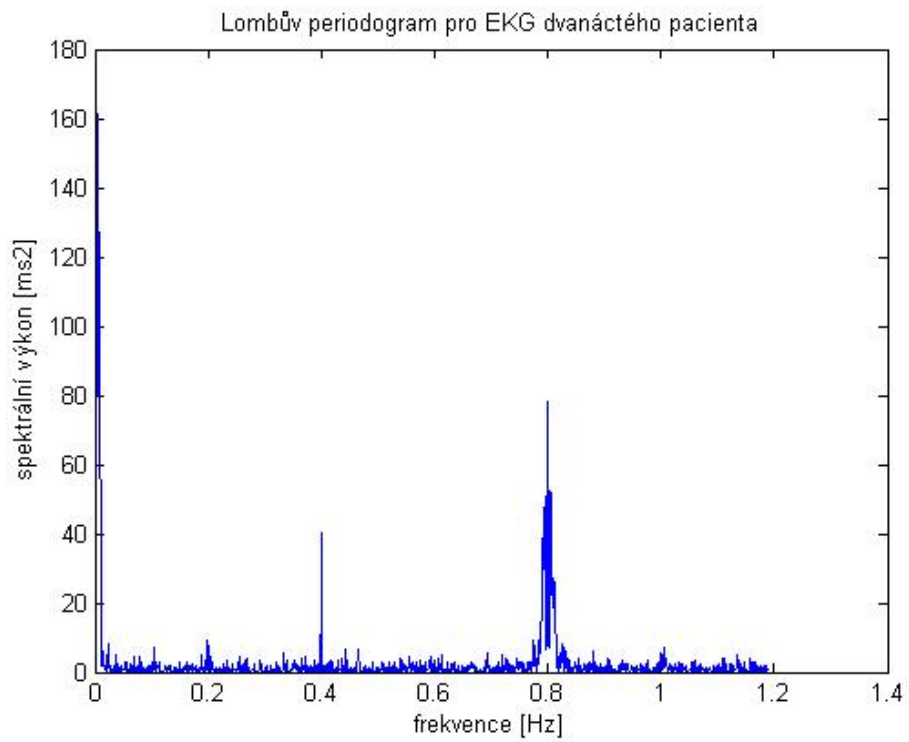


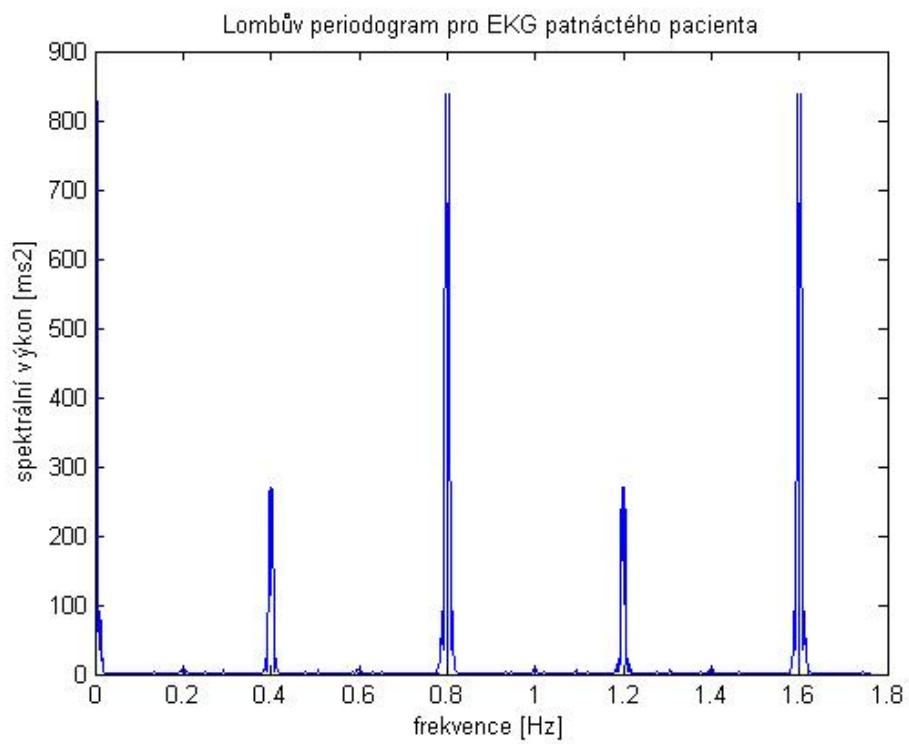
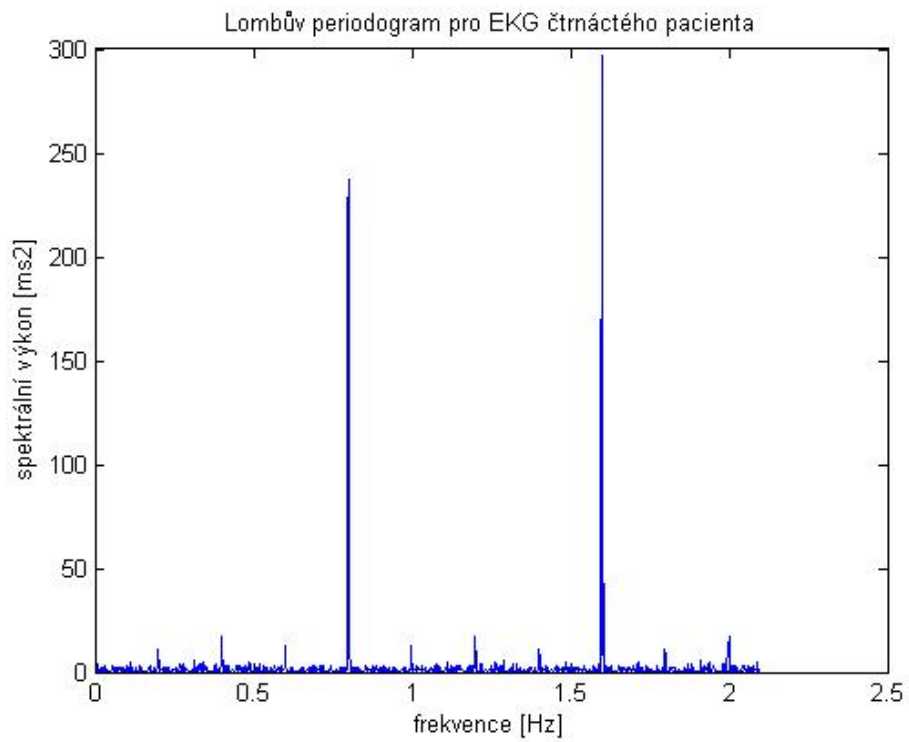


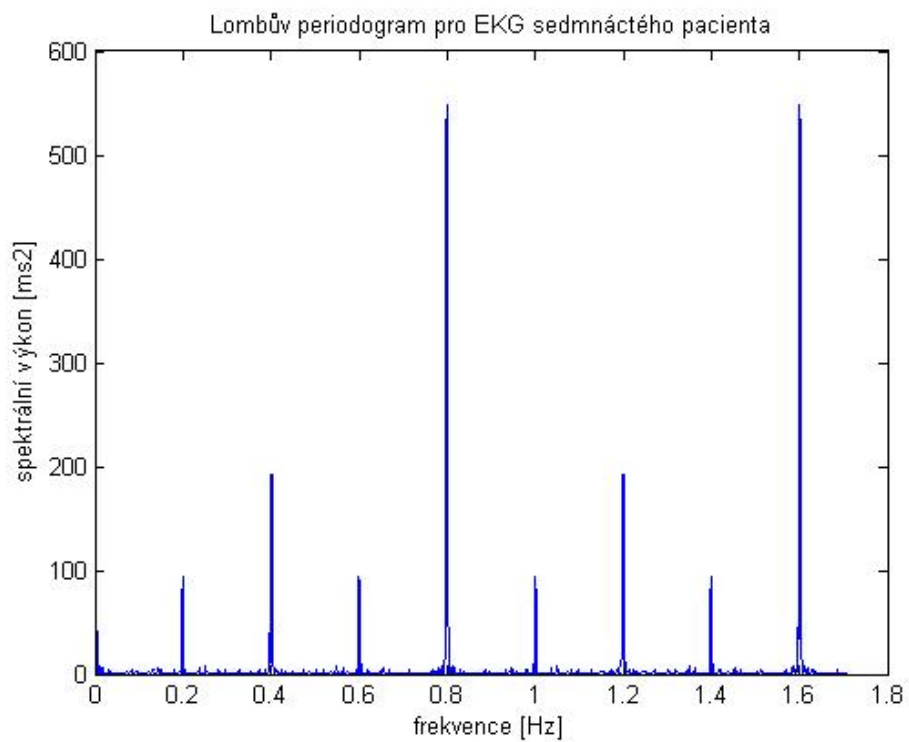
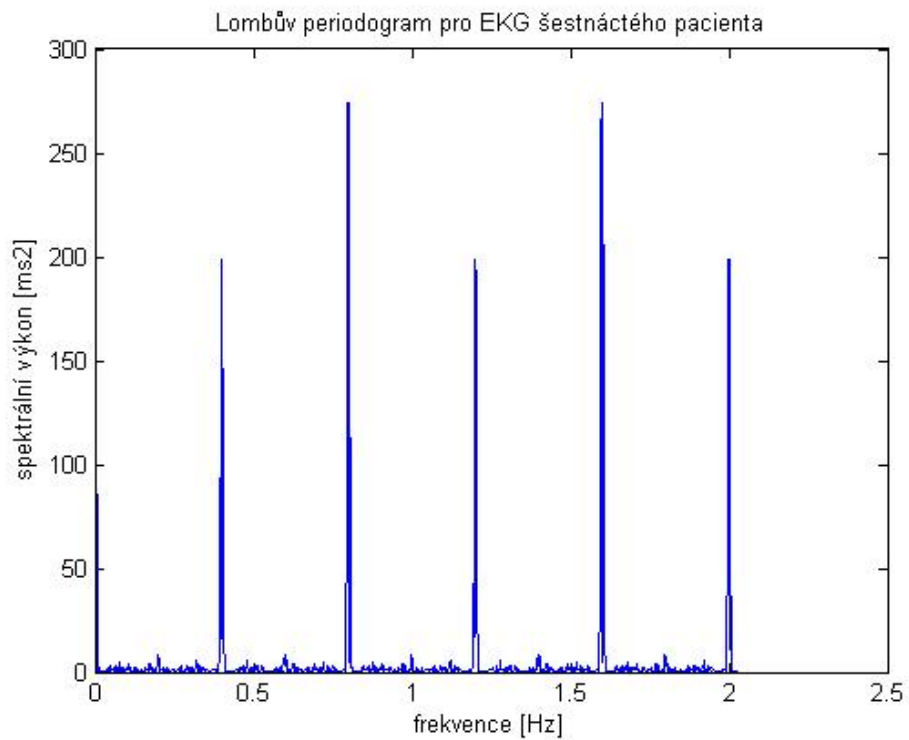


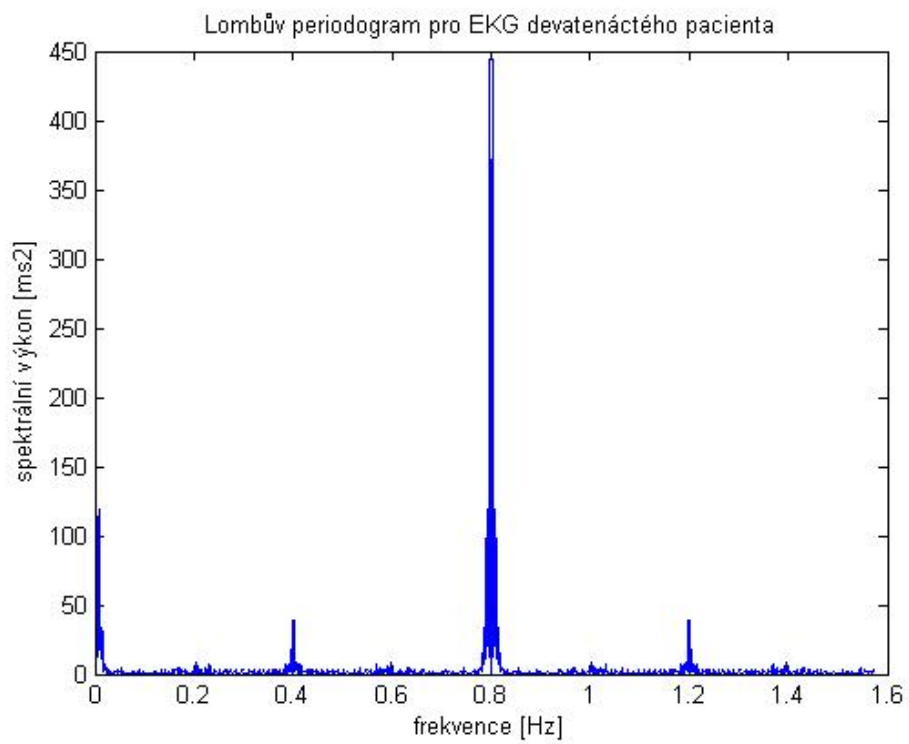
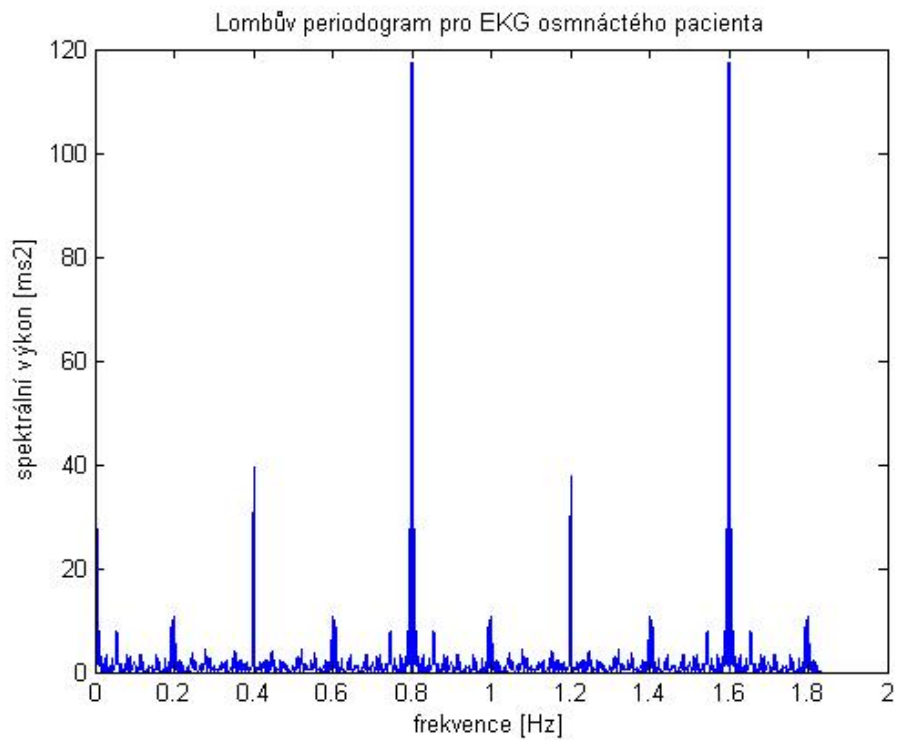




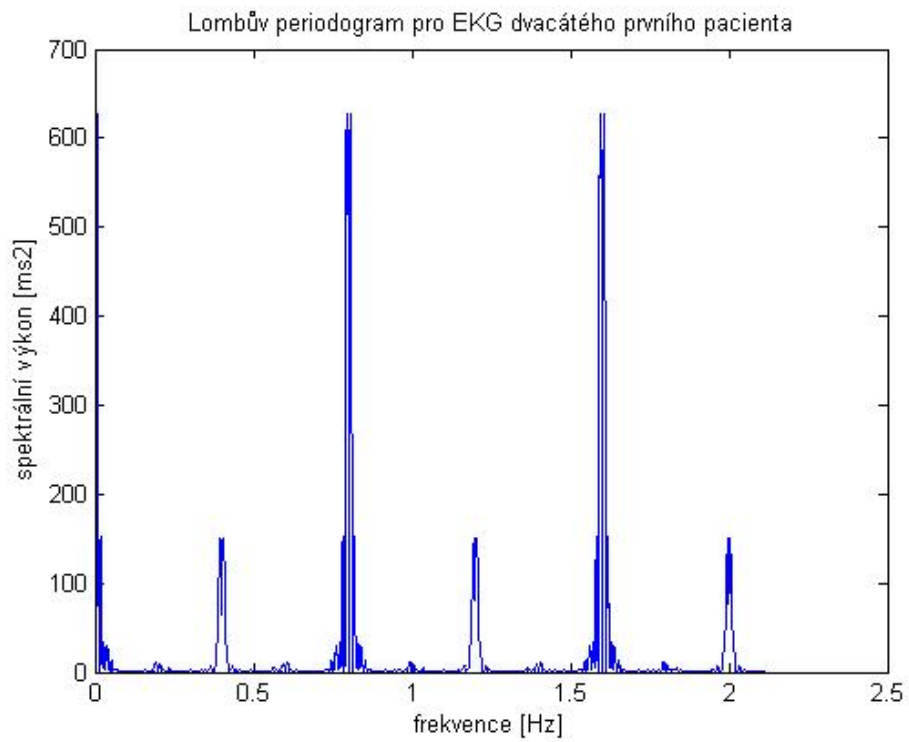
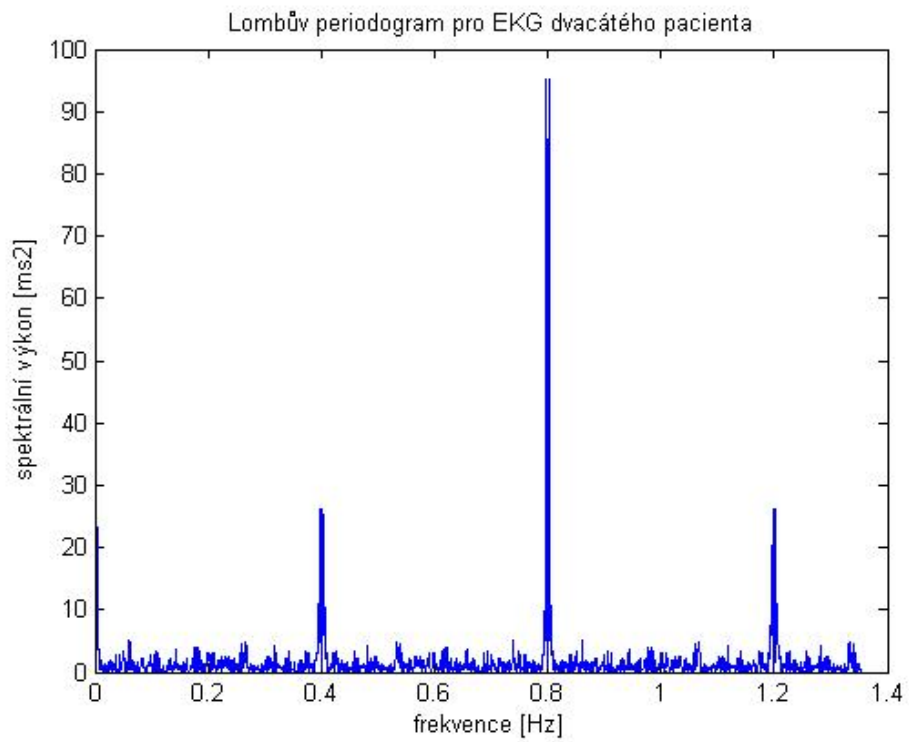


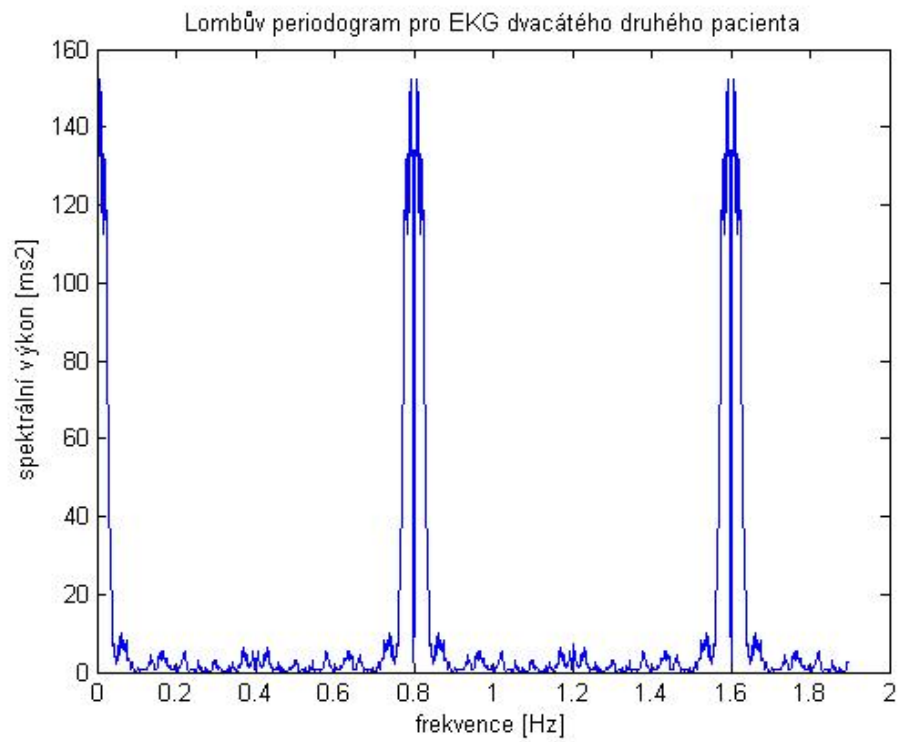












## Literatura

- [1] Anděl, J., Statistická analýza časových řad, 1. vydání. Praha: STNL, 1976.
- [2] Anděl, J., Základy matematické statistiky, 3. vydání. Praha: Matfyzpress, 2011.
- [3] Cipra, T., Analýza časových řad s aplikacemi v ekonomii. Praha: STNL, 1986.
- [4] Craymer, M. R., The least squares spectrum, its inverse transform and autocorrelation function: theory and some applications in geodesy. Ph.D. dissertation, University of Toronto, Canada, 1998.
- [5] Heinc, P., Vyšetřování srdeční stability, Kardiologické revue 2006, 8(4): p. 156-165.
- [6] Kropáč, J., Statistika B. Jednorozměrné a dvourozměrné datové soubory, regresní analýza, časové řady, 2. doplněné vydání. Brno: Fakulta podnikatelská, VUT v Brně, 2009.
- [7] Kunderová, P., Základy pravděpodobnosti a matematické statistiky. Olomouc: Vydavatelství UP, 2004.
- [8] Lomb, N. R., Least-squares frequency analysis of unequally spaced data, Astrophysics and Space Science 1976, 39: p. 447-462.
- [9] Machalová, J., Numerické metody (učební text), 2011.
- [10] Moody, G. B., Spectral Analysis of Heart Rate Without Resampling, Harvard-M.I.T. Division of Health Sciences and Technology, Cambridge, MA, USA. Computers in Cardiology 1993, 20: p. 715-718.
- [11] P. T. Ahamed Sezd, V. I. Thajudin Ahamed, Jeevamma Jacob, Paul Joseph K, Time and Frequency Domain Analysis of Heart Rate Variability and their Correlations in Diabetes Mellitus, International Journal of Biological and Life Sciences 2008, 4:1: p. 24-27.
- [12] Poměnková, J., Maršálek, R., Industrial production periodicity testing using R. A. Fisher test. Acta univ. agric. et silvic. Mendel. Brun., 2010, LVIII, No. 3: p. 189-196.
- [13] Press, W. H. Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., Numerical recipes in C: the art of scientific computing, Cambridge Univ. Press, 1992: p. 575-584.
- [14] Taylor J., Hamilton S., Some tests of the Vaníček Method of spectral analysis, Astrophysics and Space Science 1972, 17: p. 357-367.

- [15] Tomáš J., Zapletal J., Úvod do analýzy časových řad. Fakulta elektrotechniky a komunikačních technologií, VUT v Brně, 2003.
- [16] Vyskočilová, M., Analýza variability srdečního rytmu ve frekvenční oblasti. Bakalářská práce. Brno: FEKT VUT v Brně, 2011. 45 s, 1 příl. Vedoucí bakalářské práce doc. Ing. Jiří Kozumplík CSc.
- [17] Zechmeister, M., Kürster, M., The generalised Lomb-Scargle periodogram, *Astronomy and Astrophysics*, 2009, 496.2: p. 577-584.
- [18] Wikipedia.org: Periodogram [online 12. 8. 2012], dostupné z <http://en.wikipedia.org/wiki/Periodogram>
- [19] Kardiologická ambulance: EKG Holter [online 15. 3. 2013], dostupné z [http://www.kardioprachatice.cz/pages/ekg\\_holter/custom\\_page.htm;jsessionid=83D996B64F0B4BB1DD6A8AC30C3C7339?execution=e1s1](http://www.kardioprachatice.cz/pages/ekg_holter/custom_page.htm;jsessionid=83D996B64F0B4BB1DD6A8AC30C3C7339?execution=e1s1)
- [20] Czechoslovak Society of Arts & Sciences (SVU): Petr Vaníček [online 3. 10. 2012], dostupné z <http://www.svu2000.org/biosketches/petr-vanicek/>
- [21] Wikipedia.org: Least-Squares spectral analysis [online 12. 8. 2012], dostupné z [http://en.wikipedia.org/wiki/Least-squares\\_spectral\\_analysis](http://en.wikipedia.org/wiki/Least-squares_spectral_analysis)
- [22] Wikipedia.org: Synkopa [online 3. 3. 2013], dostupné z [http://cs.wikipedia.org/wiki/Synkopa\\_\(nemoc\)](http://cs.wikipedia.org/wiki/Synkopa_(nemoc))
- [23] Physionet.org: Lomb periodogram [online 15. 8. 2012], dostupné z <http://www.physionet.org/tutorials/hrv/>
- [24] Základy EKG [online 12. 1. 2013], dostupné z <http://ekg.kvalitne.cz/start.htm>
- [25] Kapitoly sportovní medicíny: zátěžové testy: variabilita srdeční frekvence [online 10. 3. 2013], dostupné z <http://is.muni.cz/do/fsps/e-learning/kapitolysportmed/pages/18-10-zatezove-testy.html>
- [26] Physionet.org: George B. Moody: Frequency Domain Measures: The Fourier Transform, the Lomb Periodogram, and Other Methods [online 13. 3. 2013], dostupné z <http://www.physionet.org/events/hrv-2006/moody-1.html>
- [27] M centrum: Analýza variability srdeční frekvence (HRV) [online 31. 3. 2013], dostupné z <http://www.mcentrum.eu/nasa-dagnostika/>
- [28] Matlab Central: Lomb [online 14. 11. 2012], dostupné z <http://www.mathworks.com/matlabcentral/fileexchange/20004-lomb-lomb-scargle-periodogram>