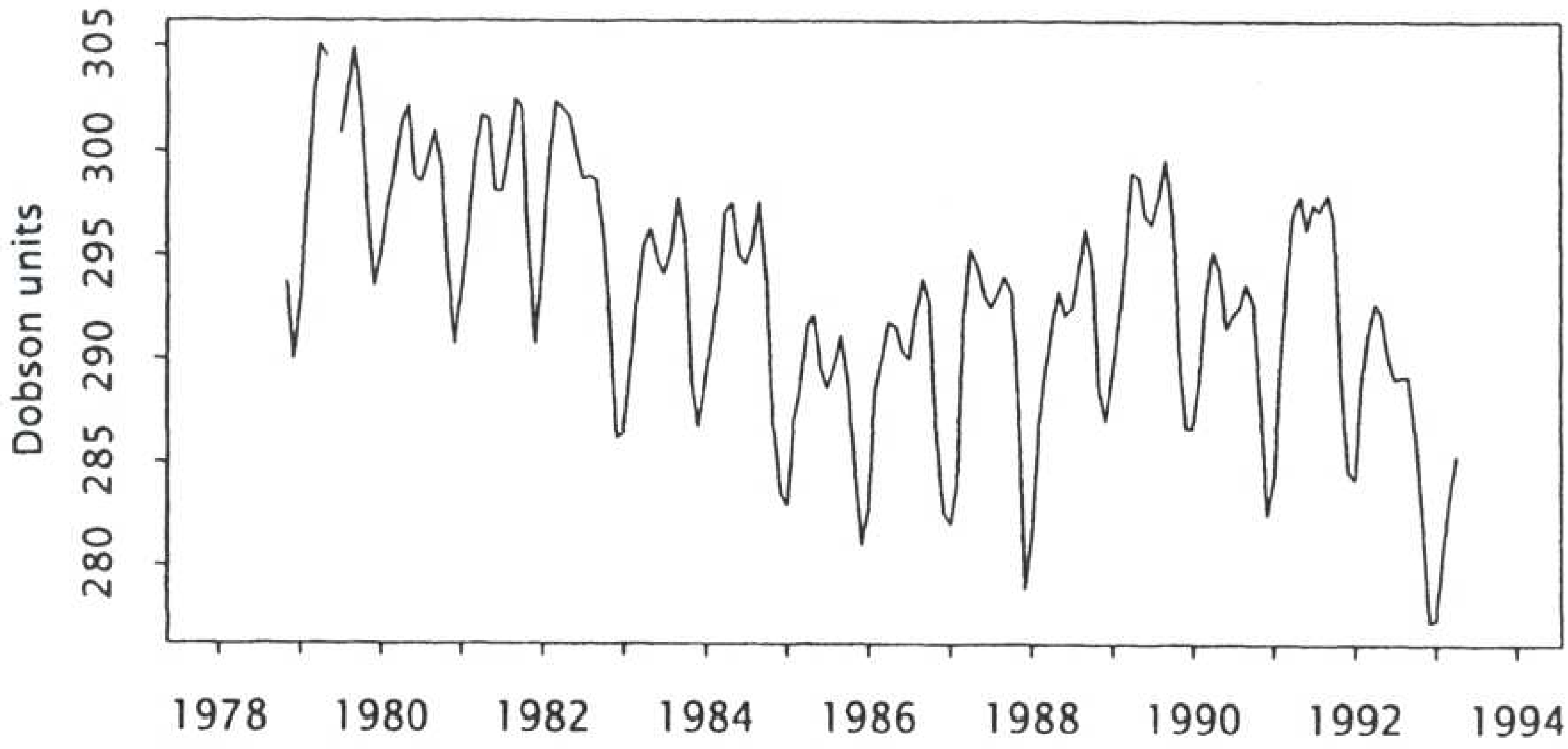
SPEKTRÁLNÍ ANALÝZA ČASOVÝCH ŘAD

prof. Ing. Jiří Holčík, CSc.

holcik@iba.muni.cz

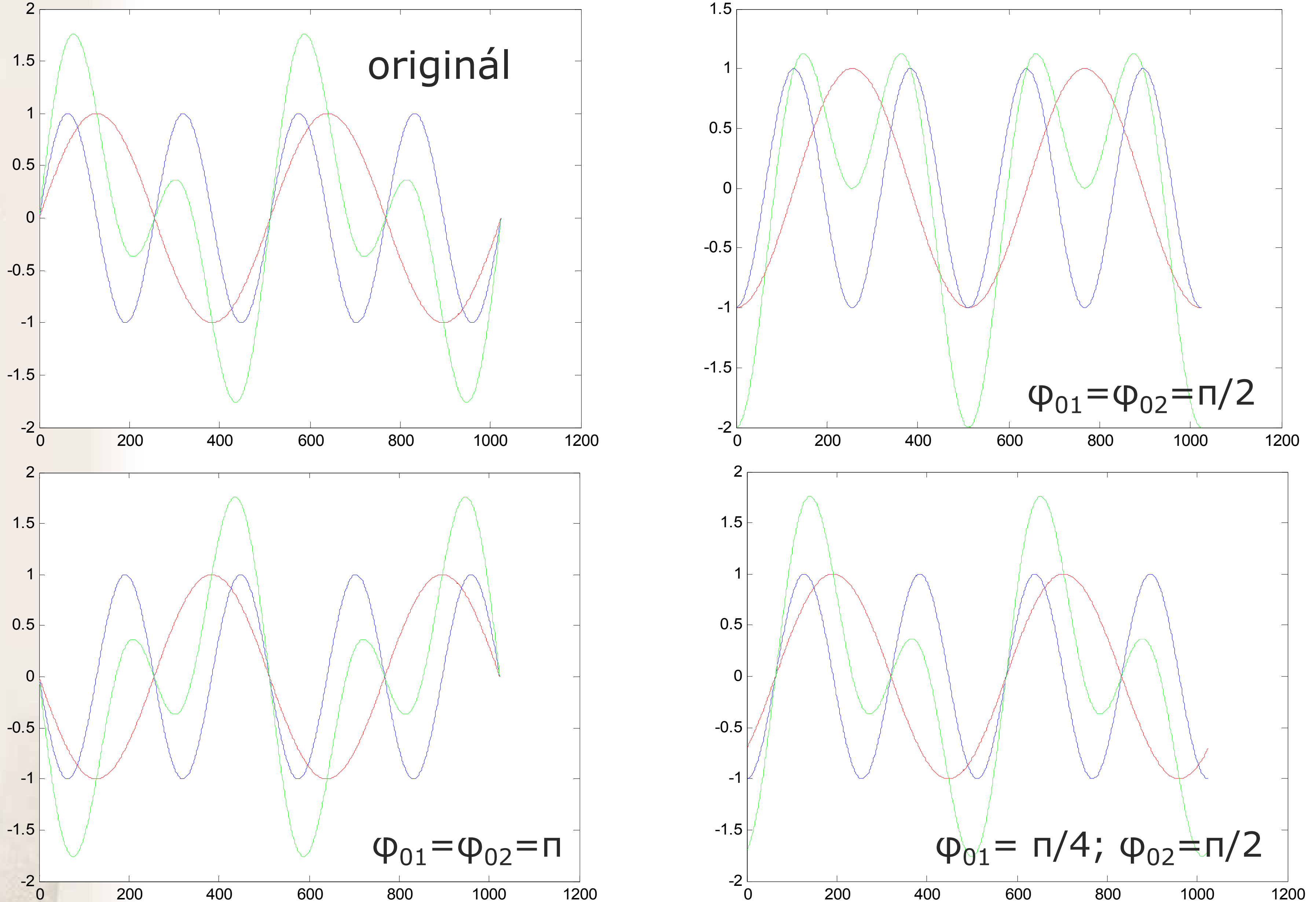
IV. ROZKLAD POMOCÍ OPTIMÁLNÍ APROXIMACE

ROZKLAD POMOCÍ OPTIMÁLNÍ APROXIMACE

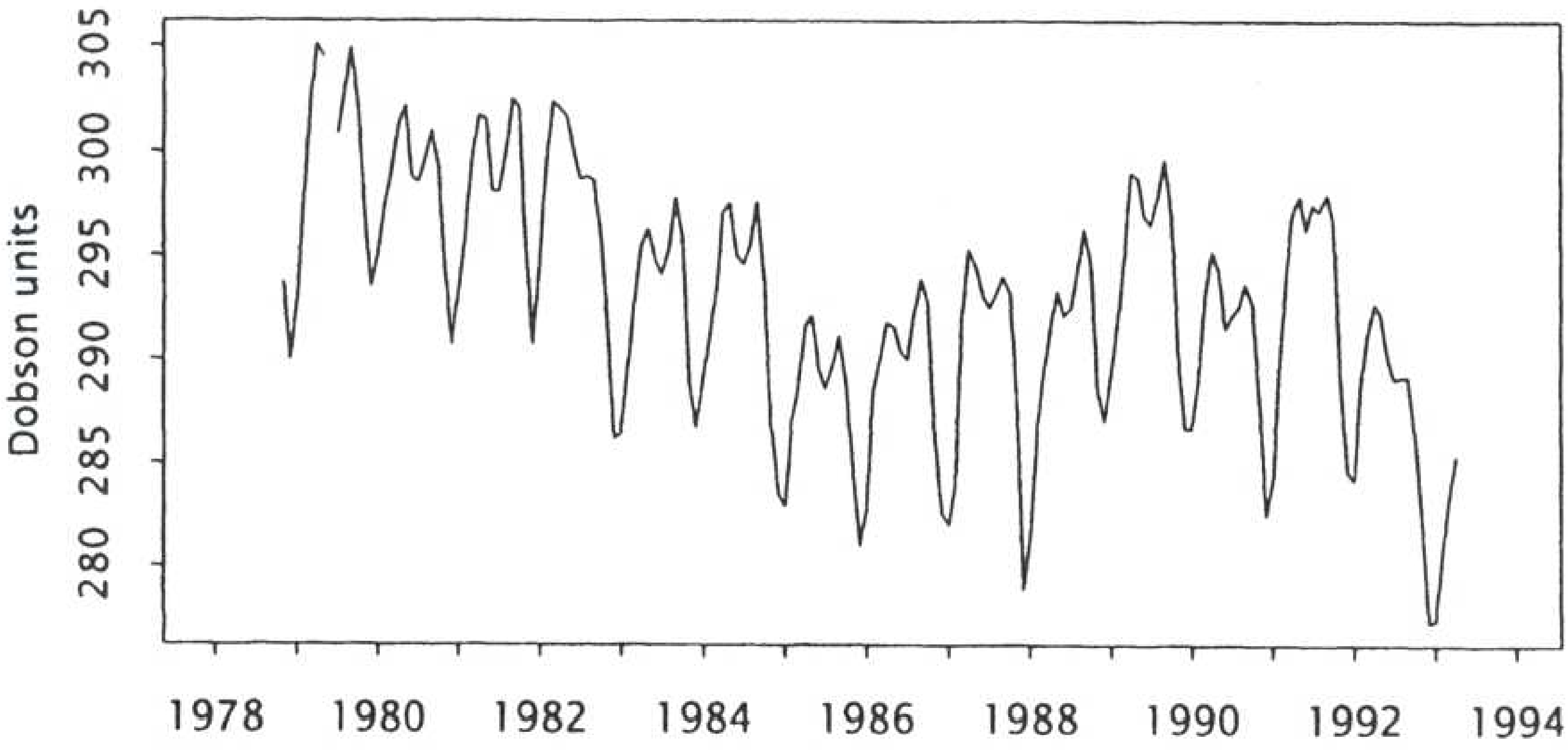


měsíční průměry celkové koncentrace ozónu, 65°S – 65°N

HRÁTKY S POČÁTEČNÍ FÁZÍ



ROZKLAD POMOCÍ OPTIMÁLNÍ APROXIMACE



měsíční průměry celkové koncentrace ozónu, 65°S – 65°N

PRINCIP

dominantní část časové řady lze vyjádřit ve tvaru

$$s(n) = A_0 + A.\cos(2\pi fn + \varphi_0),$$

kde f=1/12 cyklů/měsíc = 1 cyklus/rok
(T_{vz} je 1 měsíc)

neznámé jsou $\mathbf{A_0}$, $\mathbf{A}$ a $\boldsymbol{\varphi_0}$

experimentální data x(n) lze vyjádřit

$$x(n) = s(n) + e(n),$$

e(n) je reziduum n-tého vzorku

(model je dobrý, pokud jsou rezidua malá)

PRINCIP

metoda nejmenších čtverců

$$S(A_0, A, \varphi_0) = \sum_{n=0}^{N-1} \left(x(n) - A_0 - A.\cos(2\pi f n + \varphi_0)\right)^2$$

nejlépe když je splněn požadavek na linearitu vzhledem k parametrům modelu

PRINCIP

metoda nejmenších čtverců

$$S(A_0, A, \varphi_0) = \sum_{n=0}^{N-1} \left(x(n) - A_0 - A.\cos(2\pi fn + \varphi_0)\right)^2$$

nejlépe když je splněn požadavek na linearitu vzhledem k parametrům modelu

$$s(n) = A_0 + C_c.\cos(2\pi fn) + C_s.\sin(2\pi fn),$$

kde $C_c = A.\cos(\varphi_0)$ a $C_s = -A.\sin(\varphi_0)$

nebo také

$$A = \sqrt{C_c^2 + C_s^2} \quad \text{a} \quad \mathrm{tg}\varphi_0 = -\frac{C_s}{C_c} \; *)$$

PRINCIP

$$
{}^{*})\qquad \varphi_0 = \begin{cases} \operatorname{arctg}(-C_s / C_c) & C_c > 0 \\ \operatorname{arctg}(-C_s / C_c) - \pi & C_c < 0, C_s > 0 \\ \operatorname{arctg}(-C_s / C_c) + \pi & C_c < 0, C_s \leq 0 \\ -\pi / 2 & C_c = 0, C_s > 0 \\ \pi / 2 & C_c = 0, C_s < 0 \\ \text{cokoliv} & C_c = 0, C_s = 0 \end{cases}
$$

PRINCIP

rovnice pro určení A_0, C_c a C_s metodou nejmenších čtverců:

PRINCIP

rovnice pro určení A_0, C_c a C_s metodou nejmenších čtverců:

$$\frac{\partial S}{\partial A_0} = \sum \left[x(n) - A_0 - C_c.\cos(2\pi fn) - C_s.\sin(2\pi fn)\right]$$

$$\frac{\partial S}{\partial C_c} = \sum \cos(2\pi fn).\left[x(n) - A_0 - C_c.\cos(2\pi fn) - C_s.\sin(2\pi fn)\right] = 0$$

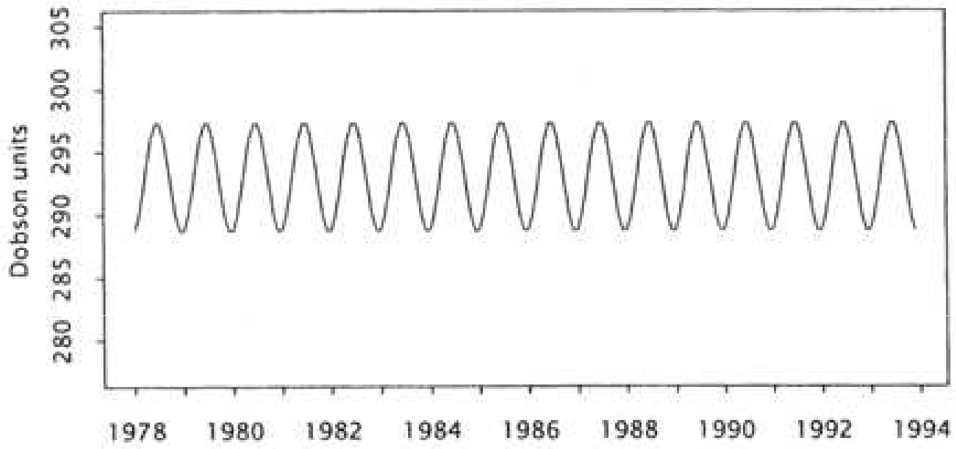
a

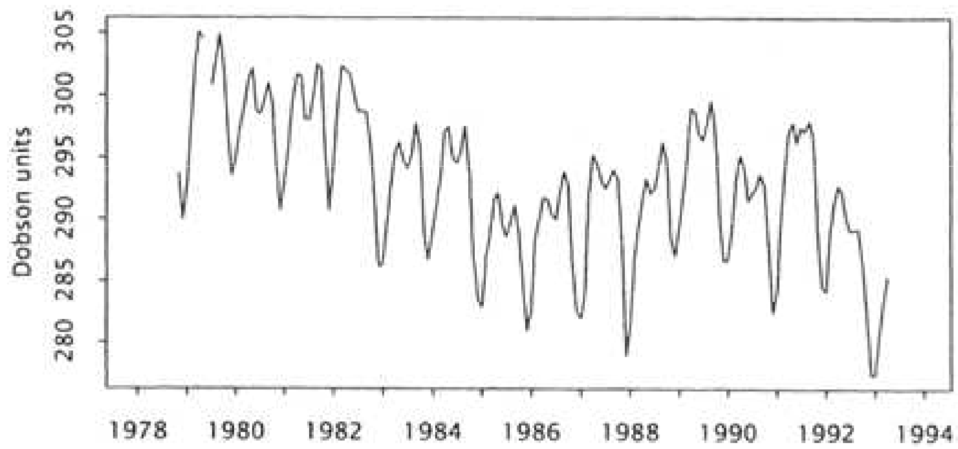
$$\frac{\partial S}{\partial C_s} = \sum \sin(2\pi fn).\left[x(n) - A_0 - C_c.\cos(2\pi fn) - C_s.\sin(2\pi fn)\right] = 0$$

ŘEŠENÍ PRO EXPERIMENTÁLNÍ DATA

$$\hat{x}(n) = 293{,}017 - 4{,}274.\cos 2\pi fn + 1{,}16.\sin 2\pi fn =$$

$$= 293{,}017 + 4{,}429.\cos(2\pi fn - 2{,}878)$$

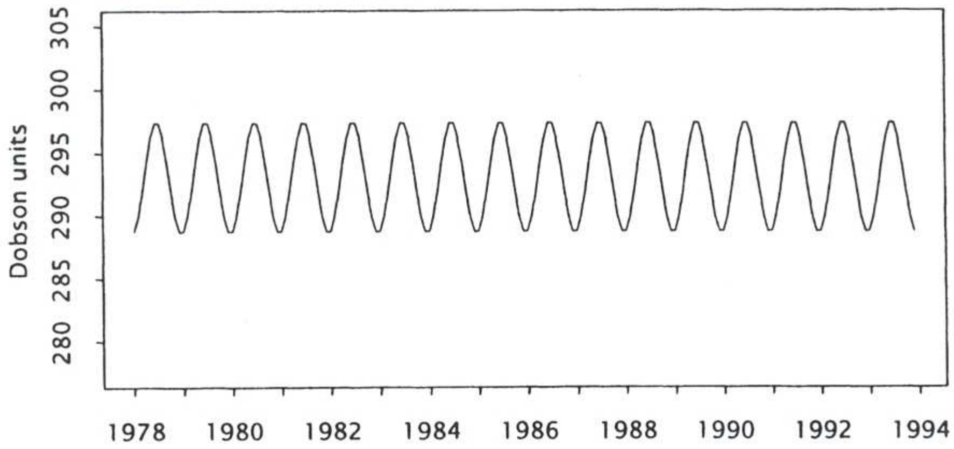


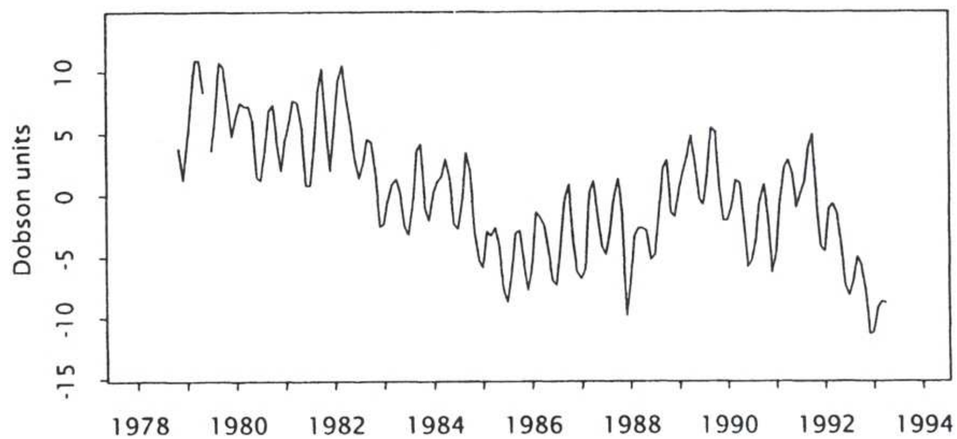


ŘEŠENÍ PRO EXPERIMENTÁLNÍ DATA

$$\hat{x}(n) = 293{,}017 - 4{,}274.\cos 2\pi fn + 1{,}16.\sin 2\pi fn =$$

$$= 293{,}017 + 4{,}429.\cos(2\pi fn - 2{,}878)$$





VÍCE HARMONICKÝCH SLOŽEK

$$x(n) = A_0 + C_{c1}.\cos(2\pi fn) + C_{s1}.\sin(2\pi fn) + \\ + C_{c2}.\cos(2.2\pi fn) + C_{s2}.\sin(2.2\pi fn) + e(n)$$

$$\frac{\partial S}{\partial A_0} = -2\sum_{n=0}^{N-1}(x(n) - A_0 - C_{c1}\cos 2\pi f_1 n - C_{s1}\sin 2\pi f_1 n - \\ - C_{c2}\cos 2\pi f_2 n - C_{s2}\sin 2\pi f_2 n)$$

$$\frac{\partial S}{\partial C_{cj}} = -2\sum_{n=0}^{N-1}\cos 2\pi f_j n(x(n) - A_0 - C_{c1}\cos 2\pi f_1 n - C_{s1}\sin 2\pi f_1 n - \\ - C_{c2}\cos 2\pi f_2 n - C_{s2}\sin 2\pi f_2 n),\ j = 1,2$$

$$\frac{\partial S}{\partial C_{sj}} = -2\sum_{n=0}^{N-1}\sin 2\pi f_j n(x(n) - A_0 - C_{c1}\cos 2\pi f_1 n - C_{s1}\sin 2\pi f_1 n - \\ - C_{c2}\cos 2\pi f_2 n - C_{s2}\sin 2\pi f_2 n),\ j = 1,2$$

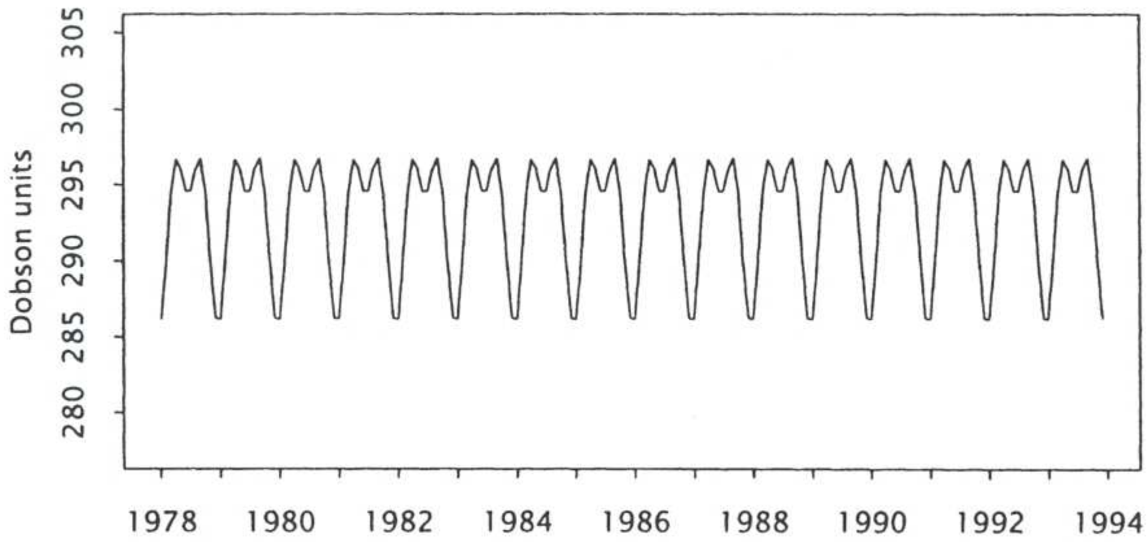
DVĚ HARMONICKÉ SLOŽKY - ŘEŠENÍ

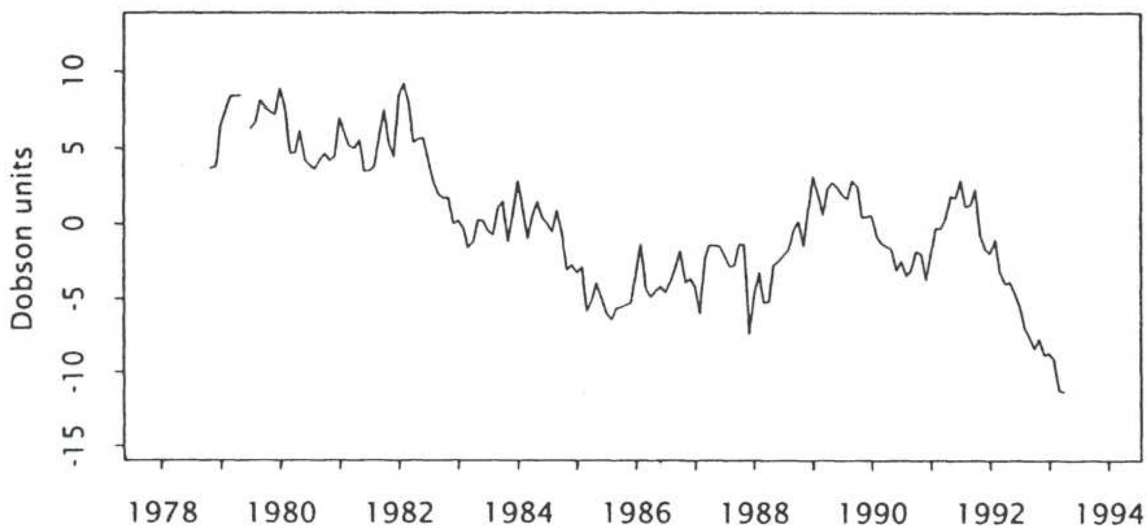
$$\hat{x}(n) = 293 - 4{,}205.\cos 2\pi fn + 1{,}075.\sin 2\pi fn -$$
$$- 2{,}614.\cos 4\pi fn + 1{,}477.\sin 4\pi fn =$$
$$= 293 + 4{,}34.\cos(2\pi fn - 2{,}89) +$$
$$+ 3{,}003.\cos(4\pi fn - 2{,}626)$$

$$\hat{x}(n) = 293{,}017 - 4{,}274.\cos 2\pi fn + 1{,}16.\sin 2\pi fn =$$
$$= 293{,}017 + 4{,}429.\cos(2\pi fn - 2{,}878)$$

?

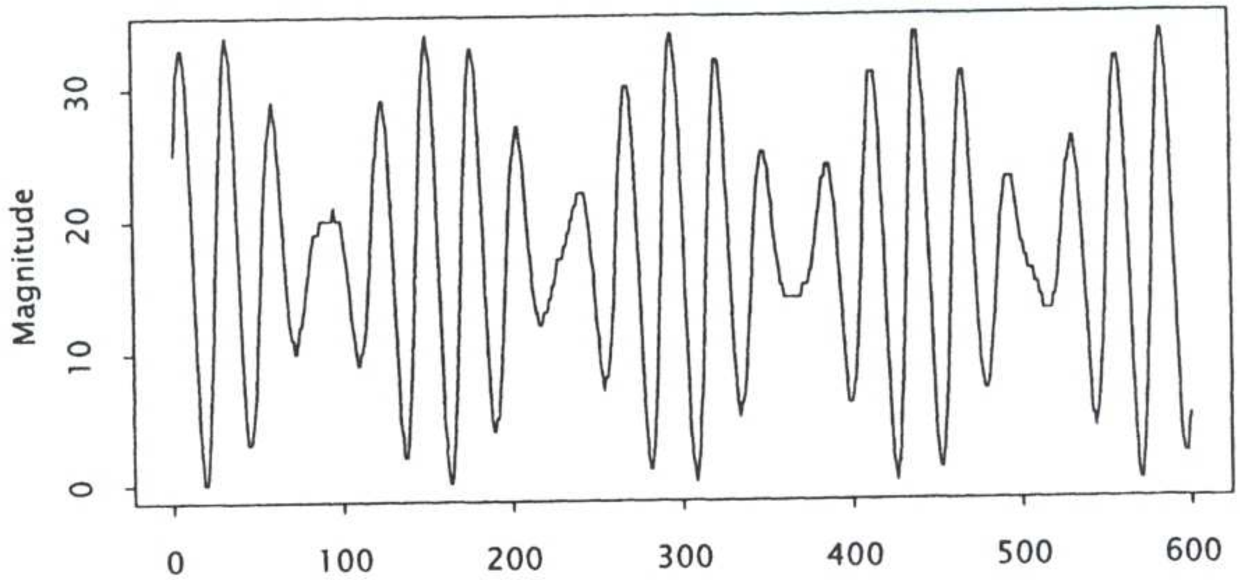
DVĚ HARMONICKÉ SLOŽKY - ŘEŠENÍ





ROZKLAD POMOCÍ OPTIMÁLNÍ APROXIMACE

CO KDYŽ NEZNÁME FREKVENCE?



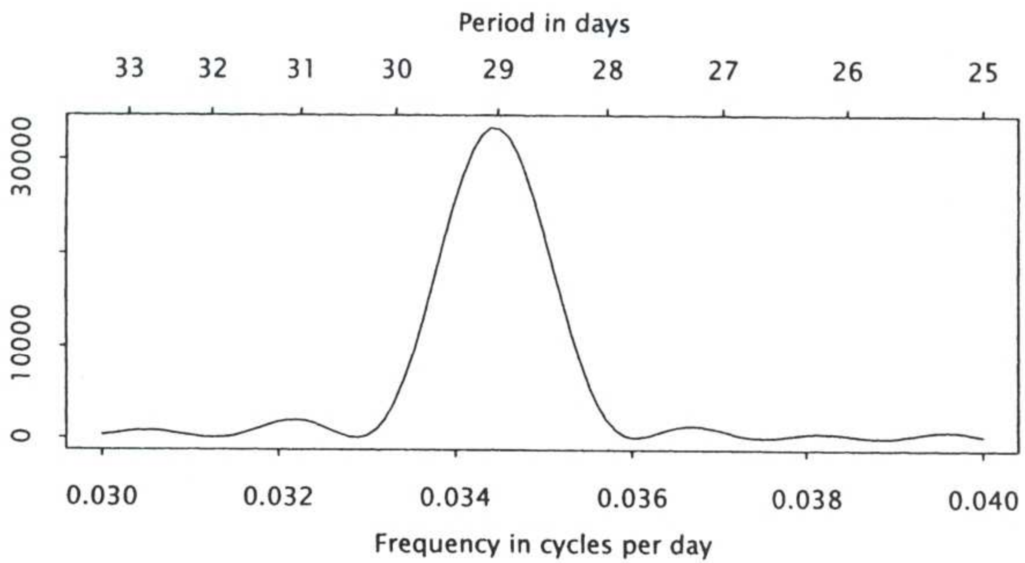
půlnoční magnituda proměnné hvězdy v 600 následných nocích (podle Whittaker,E.T., Robinson G.: The Calculus of Observations, London, Blackie & Son 1944, str.349)

21 vrcholů během 600 dní ⇒ perioda 600/21 ≈ 28,6 dní ⇒
⇒ frekvence f = 1/28,6 ≈ 0,035 cyklů/den

http://www.york.ac.uk/depts/maths/data/ts/, No.26

CO KDYŽ NEZNÁME FREKVENCE?

to je ale jenom odhad



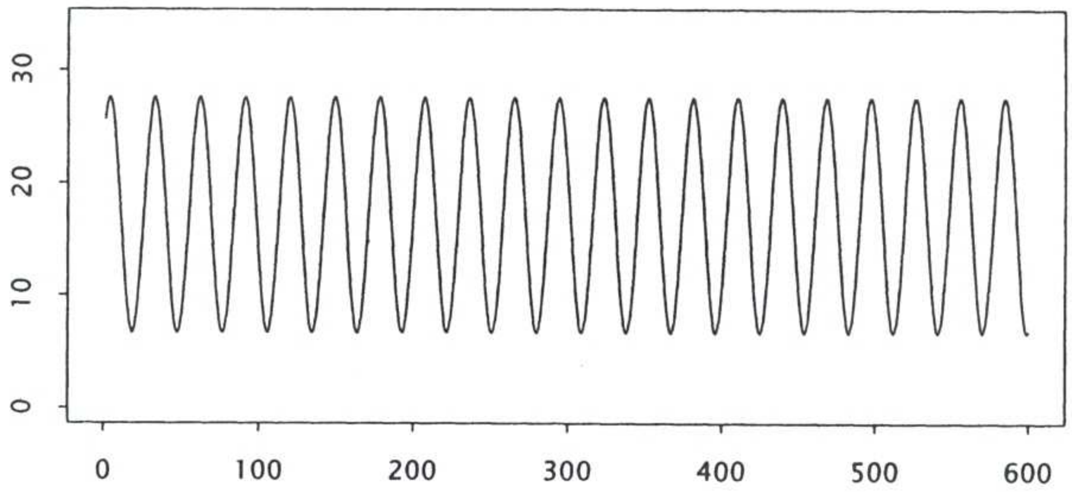
vedlejší laloky v grafu nesignalizují přítomnost dalších periodických komponent

součet čtverců pro data z proměnné hvězdy pro f ∈⟨0,03; 0,04⟩ cyklů/den
maximum pro $\hat{f} = 0{,}034442 = 1/29{,}0343$ cyklů/den

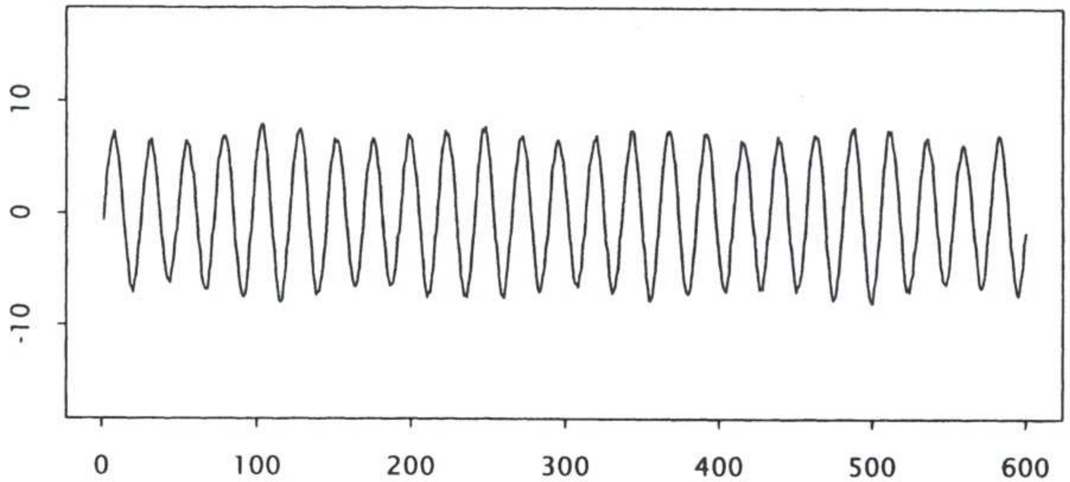
$$\hat{x}(n) = 17{,}084 + 7{,}024\cos 2\pi\hat{f}n + 7{,}82\sin 2\pi\hat{f}n$$
$$= 17{,}084 + 10{,}511\cos 2\pi(\hat{f}n - 0{,}842)$$

ROZKLAD POMOCÍ OPTIMÁLNÍ APROXIMACE

CO KDYŽ NEZNÁME FREKVENCE?



určená kosinusovka pro periodu 29,034 dní a její zbytková funkce



25 vrcholů během 600 dní ⇒ perioda 600/25 = 24 dní ⇒ ⇒ frekvence f = 1/24 ≈ ≈ 0,042 cyklů/den

ROZKLAD POMOCÍ OPTIMÁLNÍ APROXIMACE

CO KDYŽ NEZNÁME FREKVENCE?

analýzou originální časové řady lze ekvivalentně

(založeno na modelu

$$x(n)=\mu+A_2\cos 2\pi f_2 n+B_2\sin 2\pi f_2 n+\varepsilon(n)\quad)$$

určit frekvenci blízkou 0,042 cyklů/den

$$\hat{f} = 0{,}041724 = 1/23{,}967$$

$$\hat{x}(n) = 17{,}115 - 2{,}847\cos 2\pi\hat{f}n + 7{,}234\sin 2\pi\hat{f}n =$$

$$= 17{,}115 + 7{,}774\cos 2\pi(\hat{f}n - 1{,}948)$$

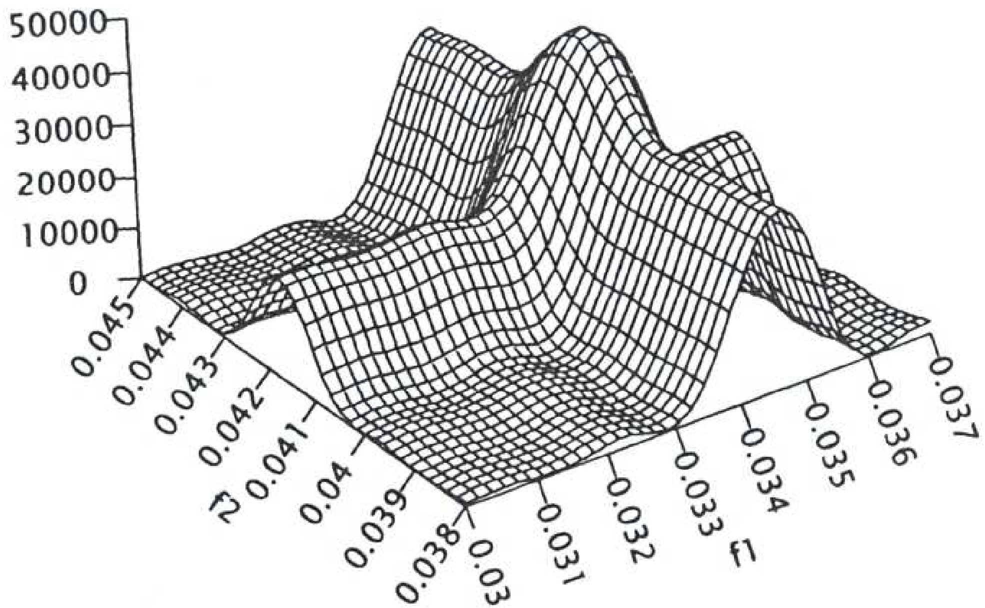
CO KDYŽ NEZNÁME FREKVENCE?

obecně model pro dvě frekvenční složky je

$$x(n) = \mu + A_1\cos 2\pi f_1 n + B_1 \sin 2\pi f_1 n + A_2 \cos 2\pi f_2 n + B_2 \sin 2\pi f_2 n + \varepsilon(n)$$

kriteriální funkce pro minimalizaci

$$S(\mu, A_1, A_2, B_1, B_2, f_1, f_2) = \sum_{n=0}^{N-1} \left(x(n) - \mu - A_1 \cos 2\pi f_1 n - B_1 \sin 2\pi f_1 n - A_2 \cos 2\pi f_2 n - B_2 \sin 2\pi f_2 n\right)^2$$



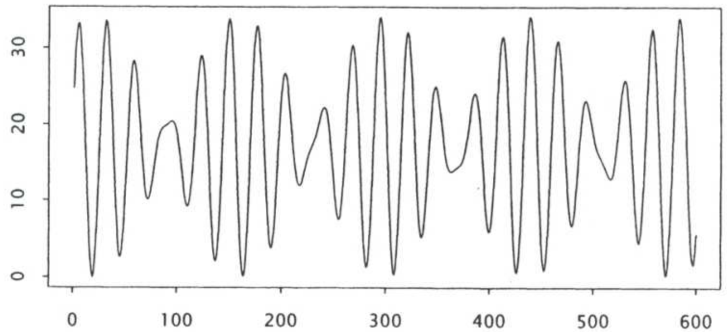
CO KDYŽ NEZNÁME FREKVENCE?

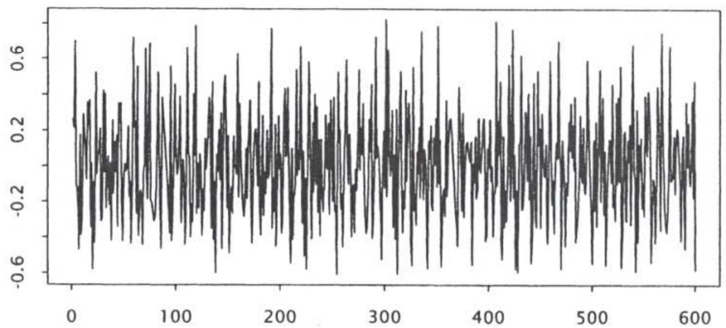
$$\hat{f}_1 = 0{,}034482 \approx 1/29{,}0003$$

$$\hat{f}_2 = 0{,}041666 \approx 1/24{,}0001$$

$$\hat{x}(n) = 17{,}086 + 6{,}074\cos 2\pi\hat{f}_1 n + 7{,}983\sin 2\pi\hat{f}_1 n - 1{,}833\cos 2\pi\hat{f}_2 n + 6{,}843\sin 2\pi\hat{f}_2 n =$$

$$= 17{,}086 + 10{,}031\cos 2\pi(\hat{f}_1 n - 0{,}917) + 7{,}085\cos 2\pi(\hat{f}_2 n - 1{,}835)$$





ÚŽASNÁ DATABÁZE ČASOVÝCH ŘAD

http://www.york.ac.uk/depts/maths/data/ts/welcome.htm

PODMÍNĚNOST SOUSTAVY LINEÁRNÍCH ROVNIC

$$2x + 6y = 8$$

$$2x + 6{,}00001y = 8{,}00001$$

Řešení: ?

PODMÍNĚNOST SOUSTAVY LINEÁRNÍCH ROVNIC

$$2x + 6y = 8$$

$$2x + 6{,}00001y = 8{,}00001$$

Řešení: $x = 1;\ \ y = 1$

PODMÍNĚNOST SOUSTAVY LINEÁRNÍCH ROVNIC

$2x + 6y = 8$

$2x + 6{,}00001 = 8{,}00001$

Řešení: $x = 1;\ \ y = 1$

$2x + 6y = 8$

$2x + 5{,}99999y = 8{,}00002$

Řešení: ?

PODMÍNĚNOST SOUSTAVY LINEÁRNÍCH ROVNIC

$$2x + 6y = 8$$

$$2x + 6,00001 = 8,00001$$

Řešení: $x = 1; \quad y = 1$

$$2x + 6y = 8$$

$$2x + 5,99999y = 8,00002$$

Řešení: $x = 10; \quad y = -2$

PODMÍNĚNOST SOUSTAVY LINEÁRNÍCH ROVNIC

špatně podmíněná soustava:

- normalizace matice soustavy tak, že všechny prvky rozšířené matice soustavy jsou menší (ale ne moc) než 1
- špatně podmíněná soustava lineárních rovnic je taková, která po této normalizaci má inverzní matici soustavy A^{-1} s některými prvky o „velké“ absolutní hodnotě.

PODMÍNĚNOST SOUSTAVY LINEÁRNÍCH ROVNIC

$$2x + 6y = 8$$

$$2x + 6{,}00001y = 8{,}00001$$

$$\mathbf{A} = \begin{bmatrix} 0{,}2 & 0{,}6 \\ 0{,}2 & 0{,}600001 \end{bmatrix}$$

PODMÍNĚNOST SOUSTAVY LINEÁRNÍCH ROVNIC

$$2x + 6y = 8$$

$$2x + 6{,}00001y = 8{,}00001$$

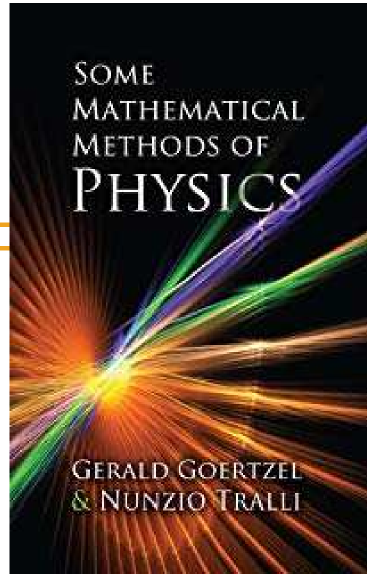
$$\mathbf{A} = \begin{bmatrix} 0{,}2 & 0{,}6 \\ 0{,}2 & 0{,}600001 \end{bmatrix}$$

$$a_{ij}^{-1} = \frac{(-1)^{i+j}|\mathbf{A}_{ij}|}{|\mathbf{A}|}$$

$$\mathbf{A}^{-1} = \begin{bmatrix} 3000005 & -1000000 \\ -3000000 & 1000000 \end{bmatrix}$$

V. GOERTZELŮV ALGORITMUS

GERALD GOERTZEL



- * 18. srpna 1919,
- ✝ 17. července 2002, White Plains, N.Y., U.S.A.
- ☑ americký teoretický fyzik, jaderný inženýr, informatik;

Vzdělání:

- ☑ BSc. (strojní inženýrství) & MSc. (fyzika) , Stevens Inst. Technology, Hoboken, New Jersey, U.S.A.
- ☑ Ph.D. (teoretická fyzika) New York Univ.

Zaměstnání:

- ☑ Manhattan projekt (Nuclear Development Corporation of America)
- ☑ zakladatel Sage Instruments, Kalifornie (testovací přístroje a přístroje pro lékařský výzkum)
- ☑ IBM, Advanced Systems Development Division (28 let) – systémy automatického řízení, komprese dat, digitální tiskárny

Výsledky:

- ☑ Goertzelův algoritmus, 13 U.S.patentů (komprese dat, zpracování 2D dat, dětský inkubátor, …),…

GOERTZELŮV ALGORITMUS (GA)

- efektivní algoritmus pro výpočet jedné hodnoty (vzorku) diskrétní Fourierovy transformace;
- (pro celé spektrum je GA složitější než FFT, ale pro výpočet omezeného počtu vzorků je efektivnější)
- podobně jako definiční vztah DFT počítá GA parametry jedné určité frekvenční složky analyzované časové řady (diskrétního signálu);
- na rozdíl od DFT pro posloupnost reálných čísel používá pouze reálnou aritmetiku

GOERTZELŮV ALGORITMUS (GA)

ALGORITMUS

(protože je podstatou číslicovým filtrem, nazývá se také často Goerzelovým filtrem)

má dva sériově zapojené stupně:

(1): $s(n) = x(n) + 2\cos(2\pi f)s(n-1) - s(n-2)$

$x(n) = 0$ pro $n<0$; $s(-2) = s(-1) = 0$

(2): $y(n) = s(n) - e^{-2\pi jf}.s(n-1)$

hodnota f určuje hodnotu normalizované (počet cyklů na vzorek) frekvence analyzované harmonické složky

!!! výpočet v reálném čase !!!

GOERTZELŮV ALGORITMUS (GA)

ALGORITMUS

přenosové funkce obou dílčích stupňů:

$$\frac{S(z)}{X(z)} = \frac{1}{1-2\cos(2\pi f)z^{-1}+z^{-2}} = \frac{1}{(1-e^{+j2\pi f}z^{-1})(1-e^{-j2\pi f}z^{-1})}$$

póly tohoto filtru leží na $e^{+j2\pi f}$ a $e^{-j2\pi f}$, tj. na jednotkové kružnici na frekvenci odpovídající f, tedy je na mezi stability a jeho stabilita může být tak závislá na numerických chybách, zejména při výpočtech s dlouhou vstupní posloupností a s aritmetikou s nižší přesností

$$\frac{Y(z)}{S(z)} = 1-e^{-j2\pi f}z^{-1}$$

GOERTZELŮV ALGORITMUS (GA)

ALGORITMUS

celková přenosová funkce pak je

$$\frac{S(z)}{X(z)}\cdot\frac{Y(z)}{S(z)}=\frac{Y(z)}{X(z)}=\frac{1-e^{-j2\pi f}z^{-1}}{(1-e^{+j2\pi f}z^{-1})(1-e^{-j2\pi f}z^{-1})}=\frac{1}{(1-e^{+j2\pi f}z^{-1})}$$

po zpětném vyjádření v časové doméně je (po rozvinutí rekurzivního vztahu pro kauzální posloupnost)

$$y(n)=x(n)+e^{+j2\pi f}y(n-1)=\sum_{i=0}^{n}x(i)e^{+j2\pi f(n-i)}=$$

$$=e^{+j2\pi fn}\sum_{i=0}^{n}x(i)e^{-j2\pi fi}\qquad(@)$$

GOERTZELŮV ALGORIMUS - VÝPOČET

PŘEDPOKLADY

- ☑ filtrace končí posledním N-tým vzorkem;
- ☑ hodnoty frekvence, pro které se realizuje výpočet (DFT) jsou dány vztahem

$$f = \frac{k}{N(\cdot T_{vz})},$$

kde k je celé číslo $k \in \{0, 1, \ldots, N-1\}$.

Po dosazení do (@) je $$y(N) = e^{+j2\pi fN} \sum_{i=0}^{N} x(i) e^{-j2\pi ik/N} =$$

$$= \sum_{i=0}^{N} x(i) e^{-j2\pi ik/N}, \text{ protože } e^{+j2\pi fN} = 1.$$

což je vztah lišící se od definice DFT pouze horní mezí součtu.

GOERTZELŮV ALGORIMUS - VÝPOČET

Pokud vložíme x(N)=0, lze beztrestně snížit horní mez na N-1.

Z definiční rekurzivní diferenční rovnice (1) je při nulovém posledním vzorku x(N)=0

$$s(N) = 2\cos(2\pi f)s(N-1) - s(N-2). \quad (*)$$

Z toho je výsledný algoritmus:

- ☑ ukončit výpočet podle vztahu (1) pro x(N-1);
- ☑ pro výpočet s(N) z hodnot s(N-1) a s(N-2) použít vztah (*);
- ☑ výslednou hodnotu y(N) z s(N) a s(N-1) určit ze vztahu (2).

GOERTZELŮV ALGORIMUS - VÝPOČET

Poslední dva kroky algoritmu lze zjednodušit

$$y(N) = s(N) - e^{-2\pi jk/N}.s(N-1) =$$

$$= (2\cos(2\pi f)s(N-1) - s(N-2)) - e^{-2\pi jk/N}.s(N-1) =$$

$$= e^{2\pi jk/N}.s(N-1) - s(N-2)$$

GOERTZELŮV ALGORIMUS - APLIKACE

VZOREK VÝKONOVÉHO SPEKTRA

$$X(k).X^*(k) = y(N).y^*(N) =$$

$$= \left(e^{2\pi jk/N}.s(N-1) - s(N-2)\right).\left(e^{-2\pi jk/N}.s(N-1) - s(N-2)\right) =$$

$$= \left(e^{2\pi jk/N}.e^{-2\pi jk/N}.s^2(N-1) - e^{2\pi jk/N}.s(N-1).s(N-2) - e^{-2\pi jk/N}.s(N-1).s(N-2) + s^2(N-2)\right) =$$

$$= s^2(N-1) - 2\cos(2\pi k/N).s(N-1).s(N-2) + s^2(N-2)$$

GOERTZELŮV ALGORIMUS - APLIKACE

VZOREK DFT POMOCÍ REÁLNÉ ARITMETIKY

první část algoritmu pracuje pouze s reálnými čísly (pokud jsou hodnoty vstupní posloupnosti reálné), komplexní výpočet reprezentuje pouze poslední krok algoritmu

$$y(N) = e^{2\pi jk/N}.s(N-1) - s(N-2)$$

což můžeme rozepsat podle Eulerova vztahu

$$\begin{aligned} y(N) &= \left(\cos(2\pi k/N) + j\sin(2\pi k/N)\right).s(N-1) - s(N-2) = \\ &= \left(\cos(2\pi k/N).s(N-1) - s(N-2)\right) + j\sin(2\pi k/N).s(N-1) \end{aligned}$$

potom lze snadno určit i modul a fázi komplexního výsledku

je-li vstup komplexní, lze jej rozložit na reálnou a imaginární část a GA zpracuje obě části separátně