# SIGNÁLY A LINEÁRNÍ SYSTÉMY

**prof. Ing. Jiří Holčík, CSc.**

holcik@iba.muni.cz, Kamenice 3, 4. patro, dv.č.424

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

# XIV.
# ANALÝZA ČASOVÝCH ŘAD

# ČASOVÁ ŘADA

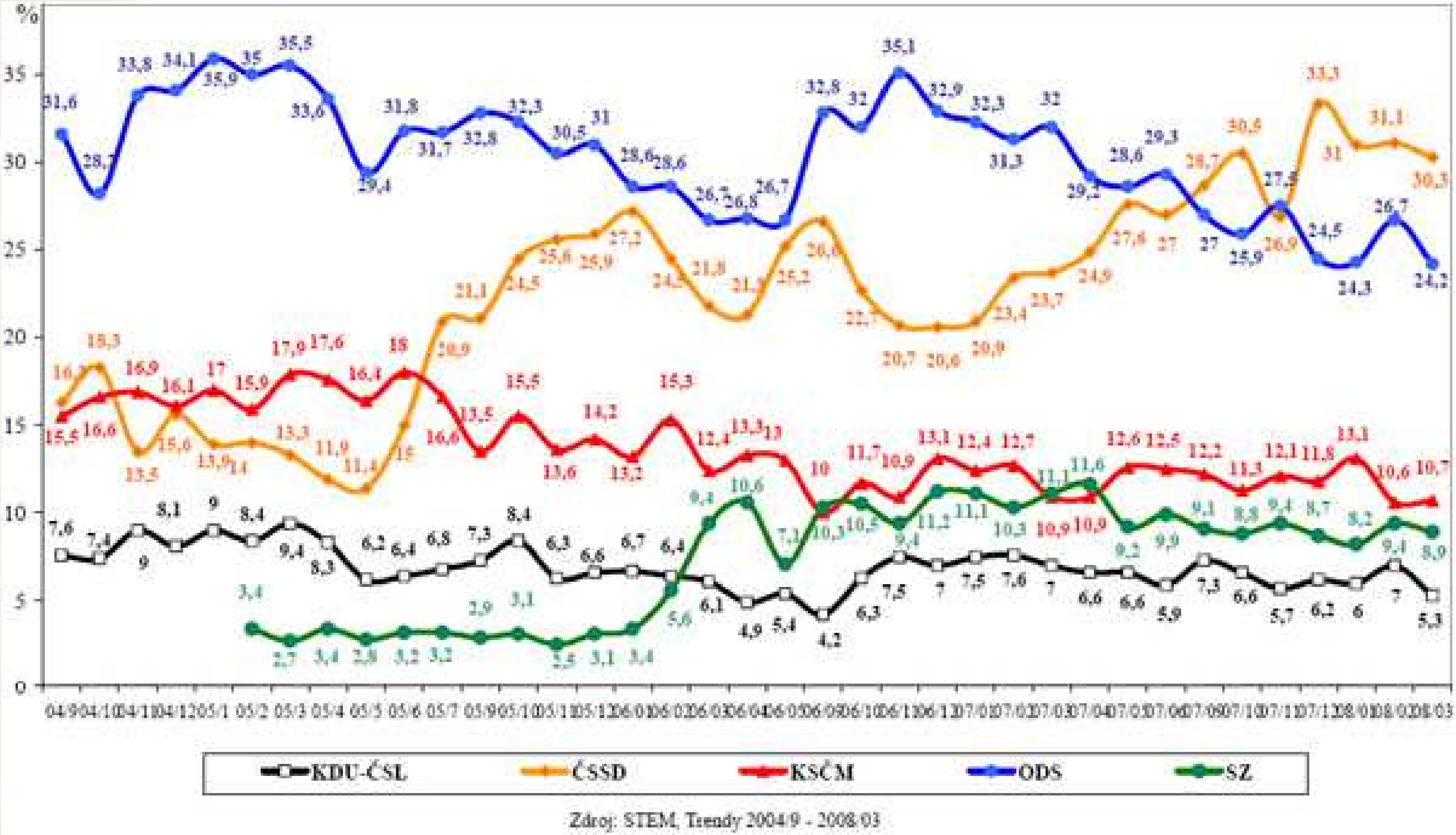
**Časová řada** jsou věcně a prostorově srovnatelná pozorování (dat), která jsou jednoznačně uspořádána z hlediska času ve směru minulost – přítomnost.

# ČASOVÁ ŘADA

**Časová řada** jsou věcně a prostorově srovnatelná pozorování (dat), která jsou jednoznačně uspořádána z hlediska času ve směru minulost – přítomnost.

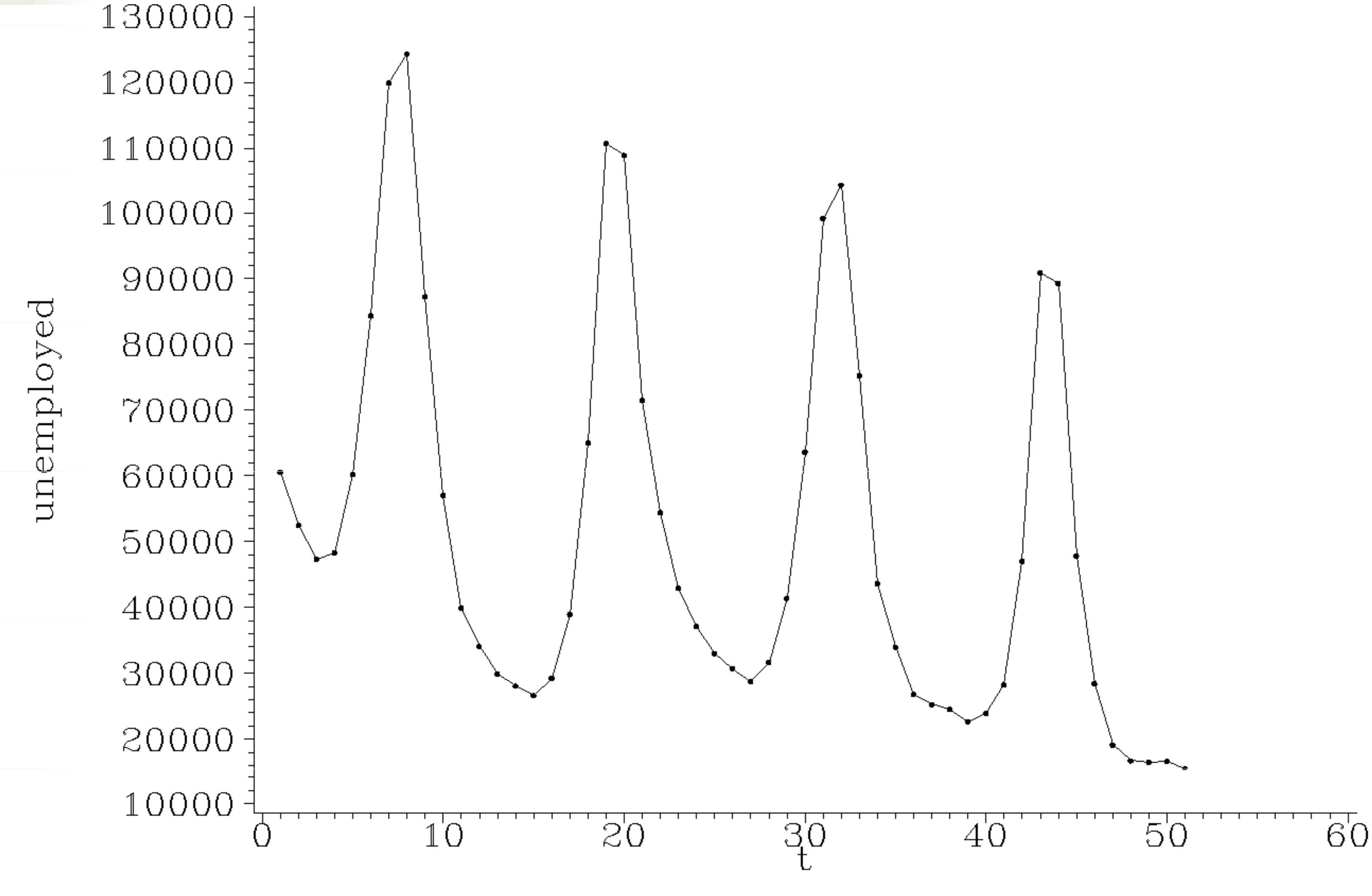
Časová řada ≡ časově závislá posloupnost hodnot sledované veličiny v diskrétních časových okamžicích ≡ diskrétní signál

# PŘÍKLADY



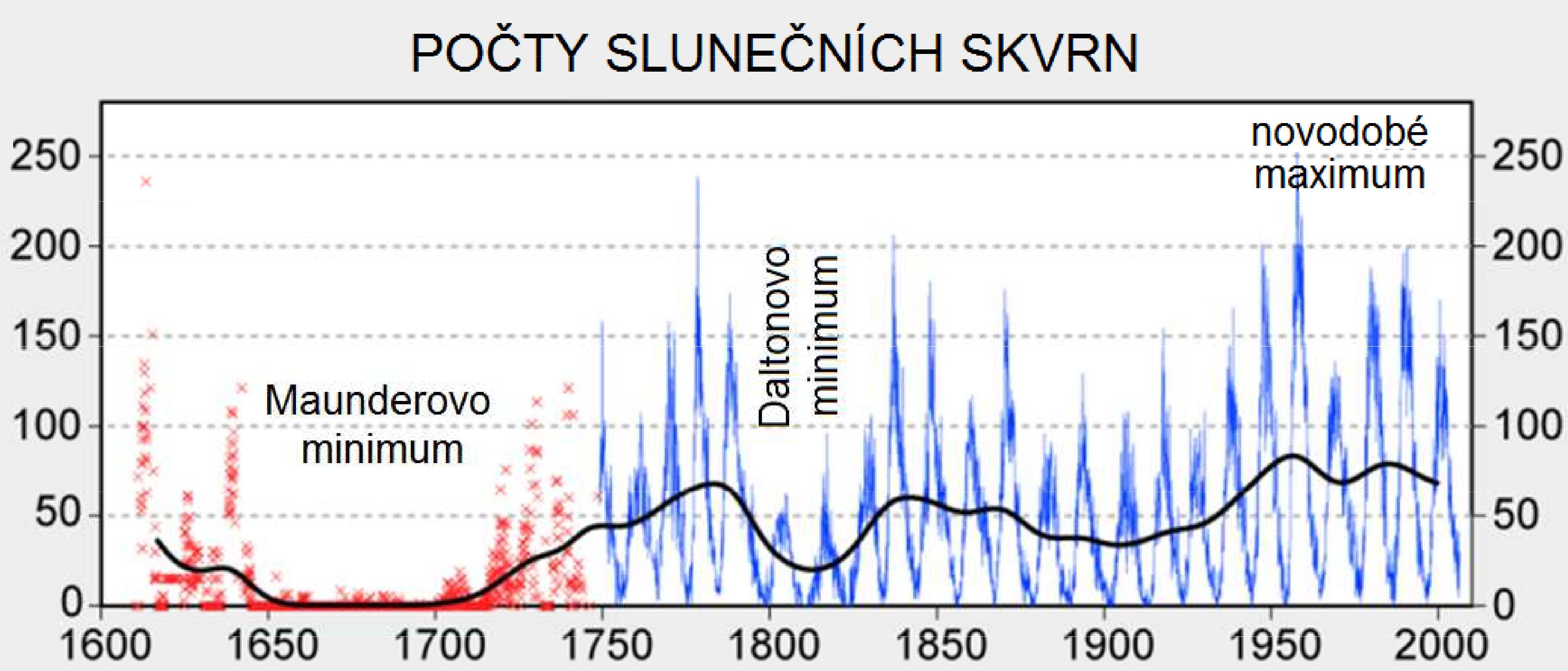

**Preference politických stran v ČR v období od 8/2004 do 3/2008**

# PŘÍKLADY




počty nezaměstnaných v Německu – 07.1975 – 09.1979 (po měsících)

# PŘÍKLADY



POČTY SLUNEČNÍCH SKVRN
250
200
150
100
50
0
1600
1650
1700
1750
1800
1850
1900
1950
2000
Maunderovo minimum
Daltonovo minimum
novodobé maximum

# ČASOVÉ ŘADY – CO S NIMI?

- ☑ **stručný popis jejích vlastností** (pomocí několika některých souhrnných parametrů (statistik)) ⇓

  k popisu spíše funkce než jednoduchá hodnota, např. klouzavý průměr než střední hodnota;
  složky řady – trend, sezónní změny, pomalé a rychlé změny, nepravidelné oscilace – **frekvenční analýza**

- ☑ **modelování průběhu**
  - ➔ pochopení procesů způsobujících vznik dat;
  - ➔ pragmatický nástroj pro splnění výše uvedených cílů

    např. pomocí lineárních systémů – autoregresivní (AR), integrační (I), s klouzavým průměrem (moving average – MA)

# ČASOVÉ ŘADY – CO S NIMI?

- ☑ **predikce budoucích hodnot** – velká část analytických metod pro časové řady; (**Predikce** (z lat. *prae*-, před, a *dicere*, říkat) znamená **předpověď** či prognózu, tvrzení o tom, co se stane nebo nestane v budoucnosti. Na rozdíl od věštění nebo hádání se slovo predikce obvykle užívá pro odhady, opřené o vědeckou hypotézu nebo teorii. srvn. forecasting

- ☑ **monitorování průběhu a detekce významných změn** - např. sledování funkce ledvin po transplantaci;

# ČASOVÁ ŘADA

## JAKÉ FENOMÉNY ROZPOZNÁVÁME V PRŮBĚHU ČASOVÉ ŘADY?

- ☑ trend
- ☑ oscilace

# ČASOVÁ ŘADA

## JAKÉ FENOMÉNY ROZPOZNÁVÁME V PRŮBĚHU ČASOVÉ ŘADY?

- ☑ **trend**

  celkový, obecný sklon, směřování, vývojová tendence (ABZ.cz)

  … {Y(t)} je (nenáhodná, deterministická?) funkce $\mu(t) = E[Y(t)]$, kde E[.] označuje očekávanou, resp. střední hodnotu;

# ČASOVÁ ŘADA

## JAKÉ FENOMÉNY ROZPOZNÁVÁME V PRŮBĚHU ČASOVÉ ŘADY?

- ☑ **trend**
- ☑ **oscilace** –
  - ➔ kolik oscilačních složek obsahuje a jaké mají kmitočty?

# ČASOVÁ ŘADA

## JAKÉ FENOMÉNY ROZPOZNÁVÁME V PRŮBĚHU ČASOVÉ ŘADY?

- ☑ **trend**
- ☑ **oscilace** –
  - ➔ kolik oscilačních složek obsahuje a jaké mají kmitočty?
  - ➔ jak často budeme odebírat vzorky?

# ADITIVNÍ MODEL ČASOVÉ ŘADY

předpokládá, že hodnoty dané časové řady jsou realizacemi náhodné proměnné $Y_n$, která se skládá ze čtyř dílčích složek

$$Y_n = T_n + Z_n + S_n + R_n,$$

kde

$T_n$ je (monotónní) funkce času (n, nT), kterou nazýváme **trend**;

$Z_n$ reprezentuje nenáhodné (deterministické) dlouhodobé cyklické vlivy (ekonomické jevy, roční či déledobější periodicita – všechno je relativní,…);

složky $T_n$ a $Z_n$ se mnohdy slučují k vyjádření dlouhodobého chování časové řady

$$G_n = T_n + Z_n$$

# ADITIVNÍ MODEL ČASOVÉ ŘADY

předpokládá, že hodnoty dané časové řady jsou realizacemi náhodné proměnné $Y_n$, která se skládá ze čtyř dílčích složek

$$Y_n = T_n + Z_n + S_n + R_n,$$

kde

$T_n$ je (monotónní) funkce času (n, nT), kterou nazýváme **trend**;

$Z_n$ reprezentuje nenáhodné (deterministické) dlouhodobé cyklické vlivy (déledobější nesezónní periodicita – všechno je relativní,…ekonomické jevy, doby ledové);

$S_n$ popisuje nenáhodné (deterministické) krátkodobé cyklické (sezónní) vlivy;

$R_n$ je náhodná veličina, která zahrnuje všechny náhodné odchylky od ideálního deterministického modelu

# ZÁKLADNÍ KONCEPT

**REÁLNÝ OBJEKT** → **PŘÍČINNÝ DETERMINISTICKÝ VZTAH** → **HODNOTÍCÍ VÝROK**

# CÍL VŠECH MOŽNÝCH ANALÝZ

ODHALIT TEN PŘÍČINNÝ DETERMINISTICKÝ VZTAH NAVZDORY VŠEMU TOMU, CO NÁM TO ODHALENÍ KAZÍ

# ADITIVNÍ MODEL ČASOVÉ ŘADY

předpokládá, že hodnoty dané časové řady jsou realizacemi náhodné proměnné $Y_n$, která se skládá ze čtyř dílčích složek

$$Y_n = T_n + Z_n + S_n + R_n,$$

V dalším budeme předpokládat, že střední hodnota chybové proměnné $R_n$ bude nulová, tj. že kladné i záporné odchylky od deterministického modelu jsou v průměru v rovnováze. (Toho lze vždycky dosáhnout šikovnou volbou jedné či více složek $T_n$, $Z_n$, nebo $S_n$.)

# ADITIVNÍ MODEL ČASOVÉ ŘADY

předpokládá, že hodnoty dané časové řady jsou realizacemi náhodné proměnné $Y_n$, která se skládá ze čtyř dílčích složek

$$Y_n = T_n + Z_n + S_n + R_n,$$

Cílem analýzy časových řad je identifikovat a interpretovat jednotlivé složky časové řady.

# MODELY TRENDU

- funkční
  - z nějakého hlediska (kritéria) optimální proložení experimentálních hodnot danou funkcí času – tehdy když už známe, jakou funkční závislostí se sledovaná veličina řídí, tj. už známe řešení diferenční rovnice popisující dynamiku daného procesu
- dynamické
  - vytvoření modelu trendu lineární dynamickou soustavou, tj, známe jak popsat dynamiku procesu, ale neznáme řešení definiční diferenciální rovnice

# NELINEÁRNÍ FUNKČNÍ MODELY TRENDU

předpokládejme pro zjednodušení aditivní model pouze

$$Y_n = T_n + R_n \text{ a } E(R_n) = 0,$$

pak $E(Y_n) = T_n = f(n)$

Obecně je funkce f(n), jejíž funkční předpis známe, závislá na určitých parametrech, tj. $f(n; \beta_1,\ldots, \beta_p)$. Parametry funkce odhadneme z množiny realizací $y_n$ náhodné proměnné $Y_n$.

# NELINEÁRNÍ FUNKČNÍ MODELY TRENDU

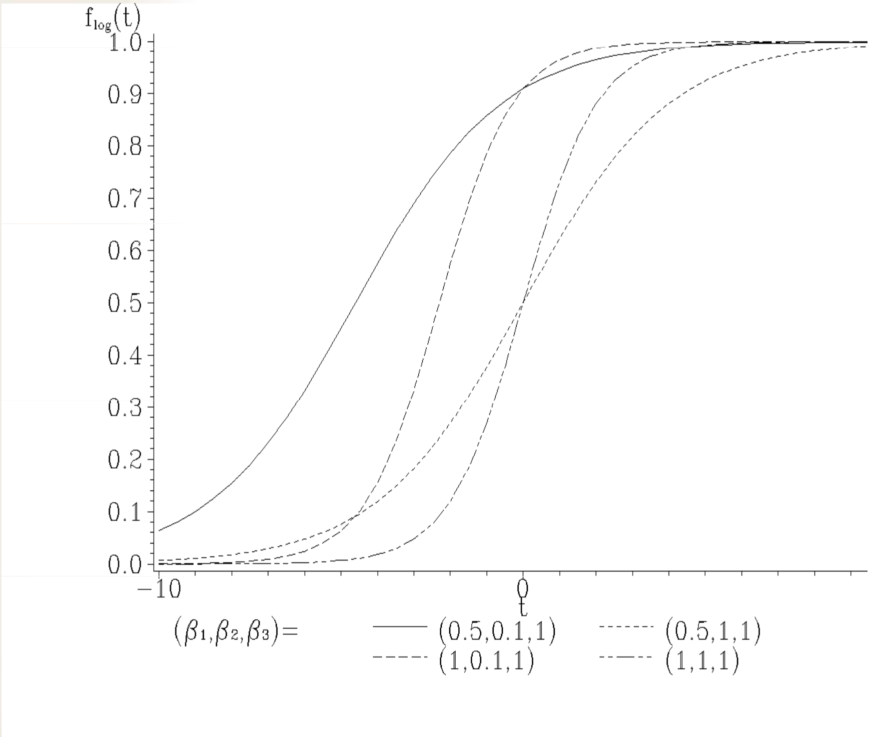
Standardním postupem, jak odhad lze spočítat, je odhad pomocí nejmenších čtverců odchylky

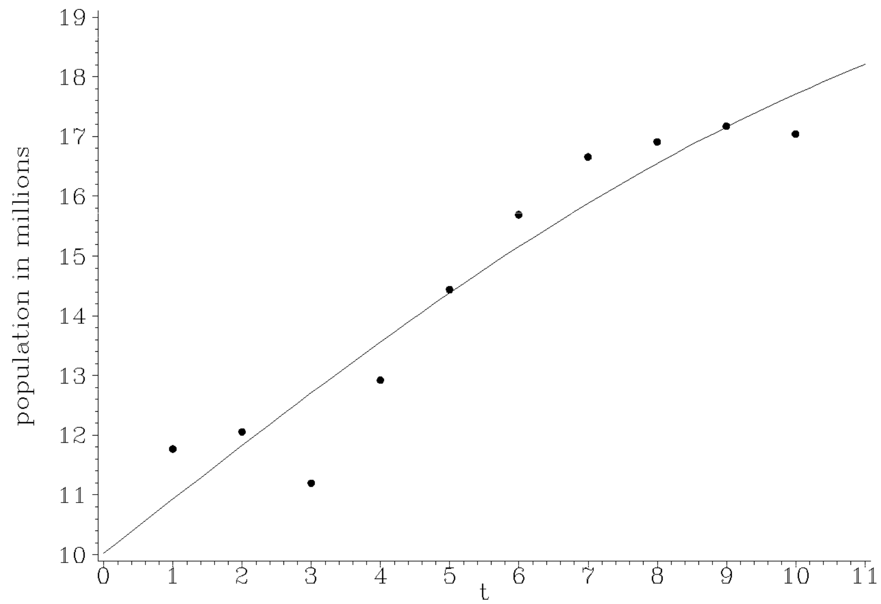
$$\sum_n \left(y_n - f(n,\hat{\beta}_1,\dots,\hat{\beta}_p)\right)^2 = \min_{\beta_1,\dots,\beta_p} \sum_n \left(y_n - f(n,\beta_1,\dots,\beta_p)\right)^2$$

Hodnoty $\hat{y}_n = f(n,\hat{\beta}_1,\dots,\hat{\beta}_p)$ pak slouží jako model chování mimo sledovaný interval (dopředná, zpětná predikce). Hodnoty rozdílů $y_n - \hat{y}_n$ nazýváme rezidua. Obsahují informaci o kvalitě datového modelu.

# LOGISTICKÁ FUNKCE

$$f_{log}(n) = f_{log}(n,\beta_1,\beta_2,\beta_3) = \frac{\beta_3}{1+\beta_2 \exp(-\beta_1 n)}, \quad n,\beta_1,\beta_2,\beta_3 \in R$$



# MITSCHERLICHOVA FUNKCE

$$f_M(n) = f_M(n, \beta_1, \beta_2, \beta_3) = \beta_1 + \beta_2 \exp(\beta_3 n), \quad n \geq 0; \beta_1, \beta_2, \beta_3 \in R$$

- typické použití pro dlouhodobý růst (pokles) v systému;
- počáteční hodnota je $f_M(0) = \beta_1 + \beta_2$;
- když je $\beta_3 < 0$, pak $\lim_{t\to\infty} f_M(n) = \beta_1$; $\beta_1$ je saturační hodnota stavu systému

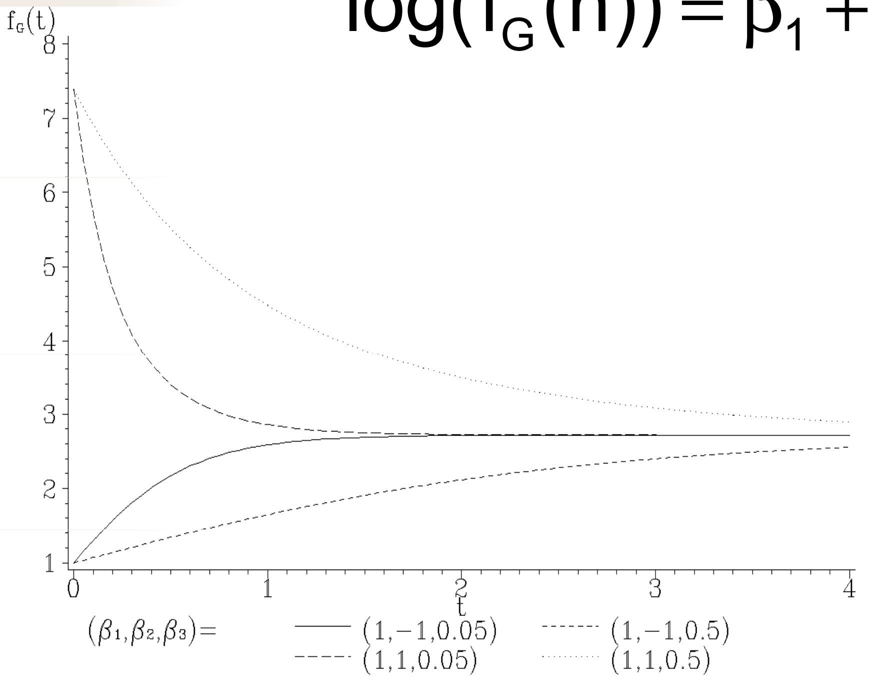
# GOMPERTZOVA FUNKCE

$$f_G(n) = f_G(n, \beta_1, \beta_2, \beta_3) = \exp(\beta_1 + \beta_2 \beta_3 n), \quad \beta_1, \beta_2 \in R; \beta_3 \in (0,1)$$

samozřejmě platí

$$\log(f_G(n)) = \beta_1 + \beta_2 \beta_3{}^n = \beta_1 + \beta_2 \exp(\log(\beta_3).n),$$

a tak $\log(f_G)$ je Mitscherlichovou funkcí s parametry $\beta_1$, $\beta_2$ a $\log(\beta_3)$



# ALOMETRICKÁ FUNKCE

$$f_a(n) = f_a(n, \beta_1, \beta_2) = \beta_2 n^{\beta_1}, \quad n \geq 0, \beta_1 \in R, \beta_2 > 0$$

Protože

$$\log(f_a(n)) = \log(\beta_2) + \beta_1 \log(n), \quad n \geq 0,$$

můžeme použít lineární regresní model pro logaritmická data $\log(y_n)$

$$\log(y_n) = \log(\beta_2) + \beta_1 \log(n) + \varepsilon_n, \; n \geq 1,$$

kde $\varepsilon_n$ reprezentuje chybovou proměnnou

# ALOMETRICKÁ FUNKCE

Alometrie = studium proměnlivých proporcí rozměrů organismu, spojených se změnou jeho velikosti, a to buď v rámci individuálního růstu (a. ontogenetická), nebo ve srovnání s příbuznými organismy různých velikostí (a. fylogenetická)

Alometrická rovnice: nechť **$y = mx^a$**, kde **x** a **y** jsou změřené délky charakteristických částí těla. Tvar rovnice naznačuje nelineární závislost růstu. Předpokládáme **y = 0** pro **x = 0**, je-li **x** celková délka.

Z toho plyne ***ln*(y) = a *ln*(x)**, kde **a = alometrický koeficient** který odráží změny poměru délek **x** k délkám **y**.

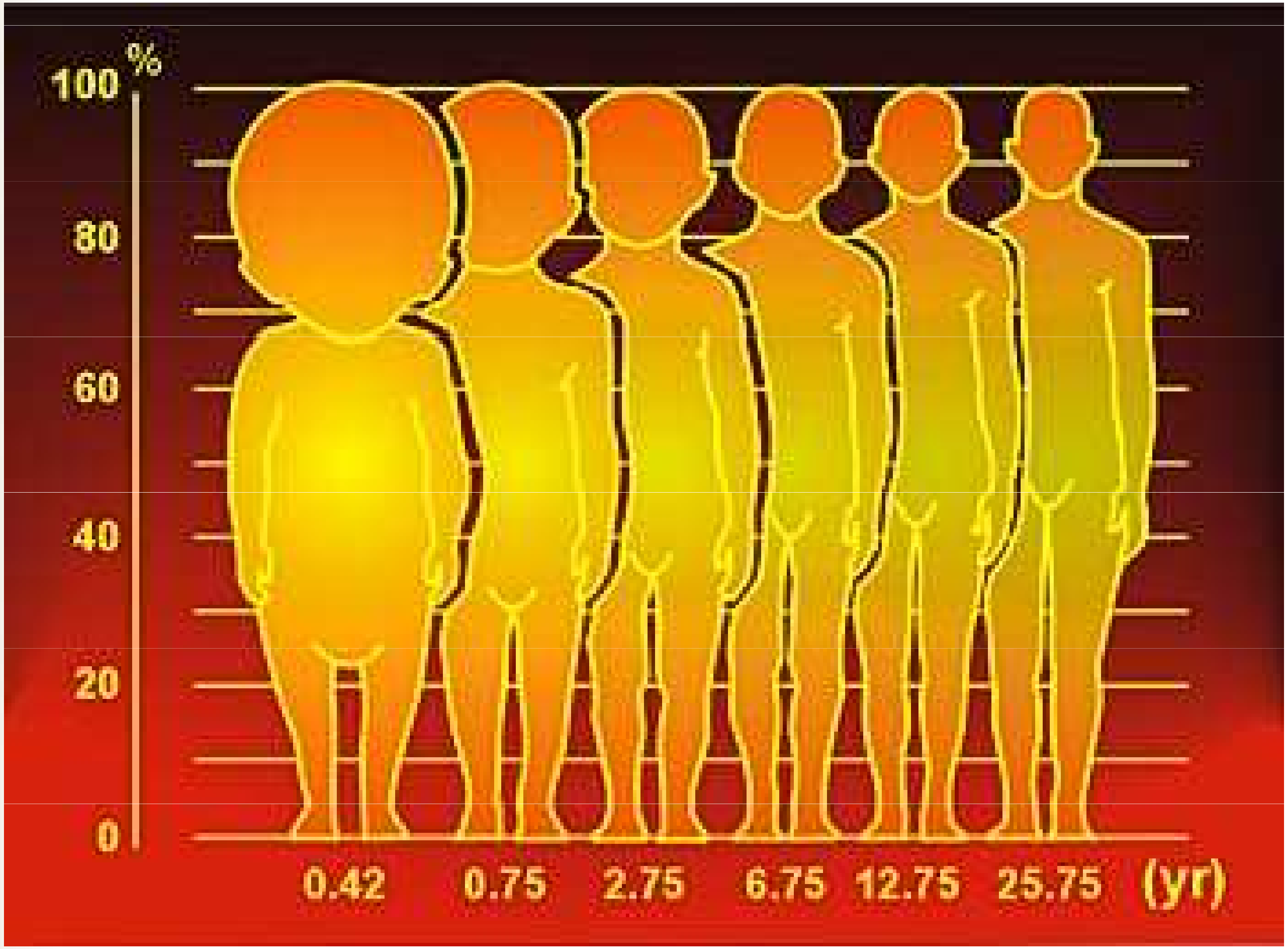
Je-li **a = 1**, potom je **y** úměrné **x** a nedochází k žádné alometrii, **y** je isometrické vzhledem k **x**.

Je-li **a > 1**, roste **y** rychleji než **x** (pozitivní alometrie).

Je-li **a < 1**, roste **y** pomaleji než **x** (negativní alometrie).

Například u člověka roste hlava pomaleji než torzo.

# ALOMETRICKÁ FUNKCE



# LINEÁRNÍ FILTRACE ČASOVÝCH ŘAD

Předpokládejme

$$Y_n = T_n + S_n + R_n,$$

Cílem nechť je stanovení odhadů trendu (driftu) a sezónní složky (složek) $\widehat{T}_n$ a $\widehat{S}_n$ , které posléze odstraníme z časové řady $y_n - \widehat{S}_n$ , resp. $y_n - \widehat{T}_n$ , čímž řadu rozdělíme na „hladké" deterministické složky a nepravidelnou fluktuační složku s nulovou střední hodnotou.

# LINEÁRNÍ FILTRACE ČASOVÝCH ŘAD

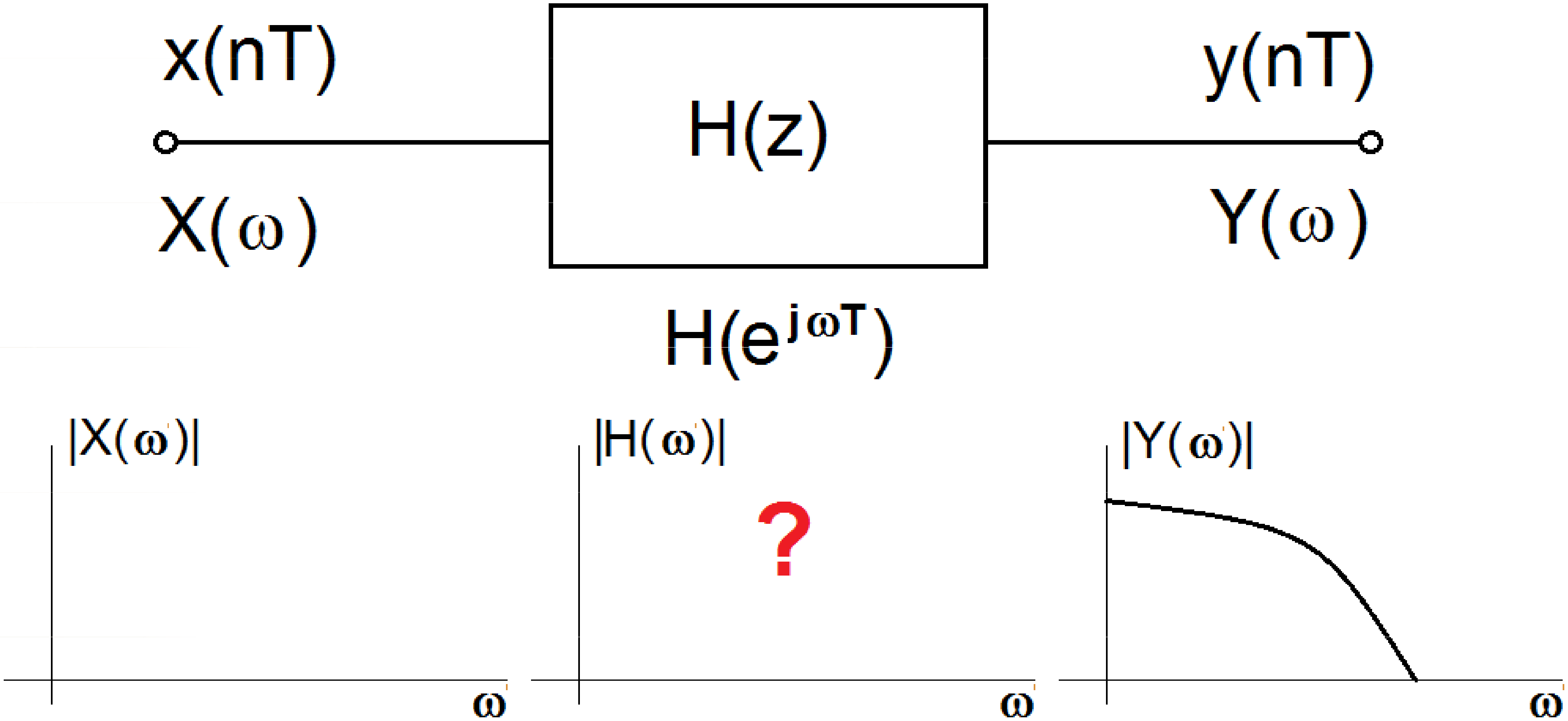
Předpokládejme

$$Y_n = T_n + S_n + R_n,$$

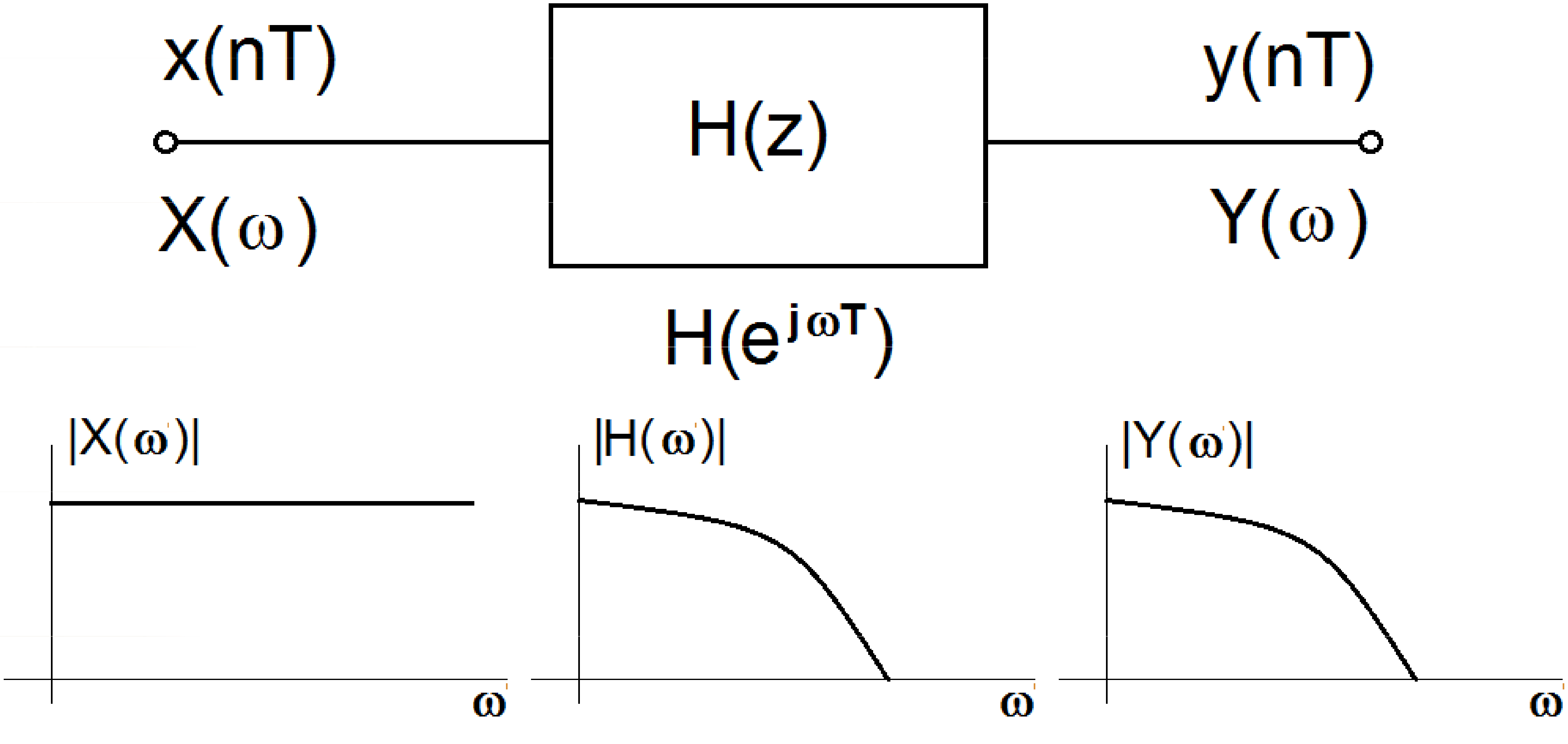
Pokud vymyslíme filtr, který dokáže odhadnout (modelovat) průběh trendu, pak vymyšlený filtr představuje matematický model procesu, který je zdrojem trendové posloupnosti.

Pokud vymyslíme filtr, který dokáže odhadnout (modelovat) průběh sezónních oscilací, pak vymyšlený filtr představuje matematický model procesu, který je zdrojem sezónní posloupnosti.

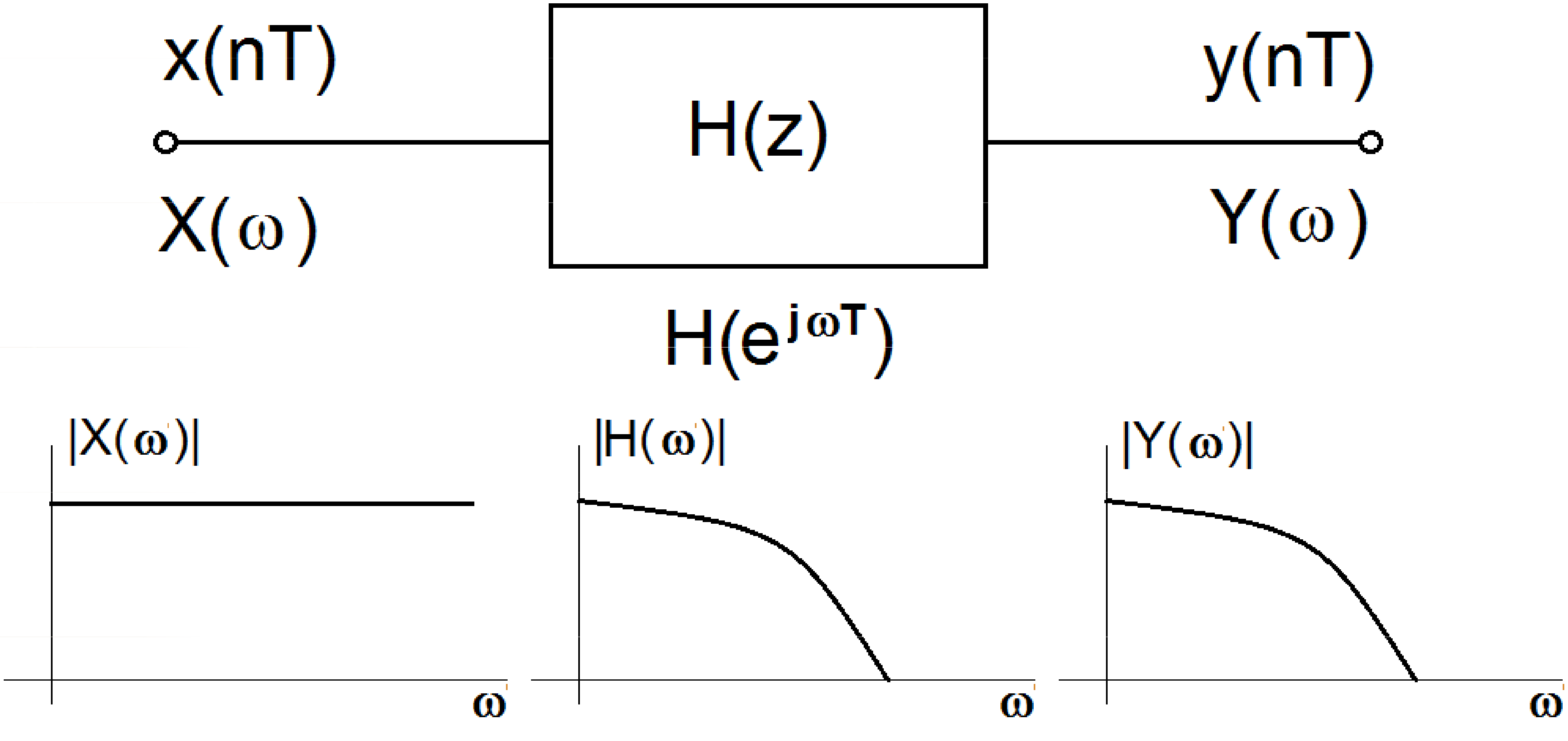
# JAK VYMYSLET TEN SPRÁVNÝ MODEL?



# JAK VYMYSLET TEN SPRÁVNÝ MODEL?



# JAK VYMYSLET TEN SPRÁVNÝ MODEL?




Jaké známe signály s konstantním spektrem?

# JAK VYMYSLET TEN SPRÁVNÝ MODEL?

Jaké známe signály s konstantním spektrem?

# JAK VYMYSLET TEN SPRÁVNÝ MODEL?

Jaké známe signály s konstantním spektrem?

- ☑ deterministický - jednotkový impuls;
- ☑ náhodný - bílý šum;

Chceme zkonstruovat takový filtr, které má, zjednodušeně, impulsní charakteristiku odpovídající modelovanému průběhu.

# ODHAD TRENDU (DRIFTU)

pomalu se měnící složky

$\Downarrow$

**DOLNÍ PROPUST**

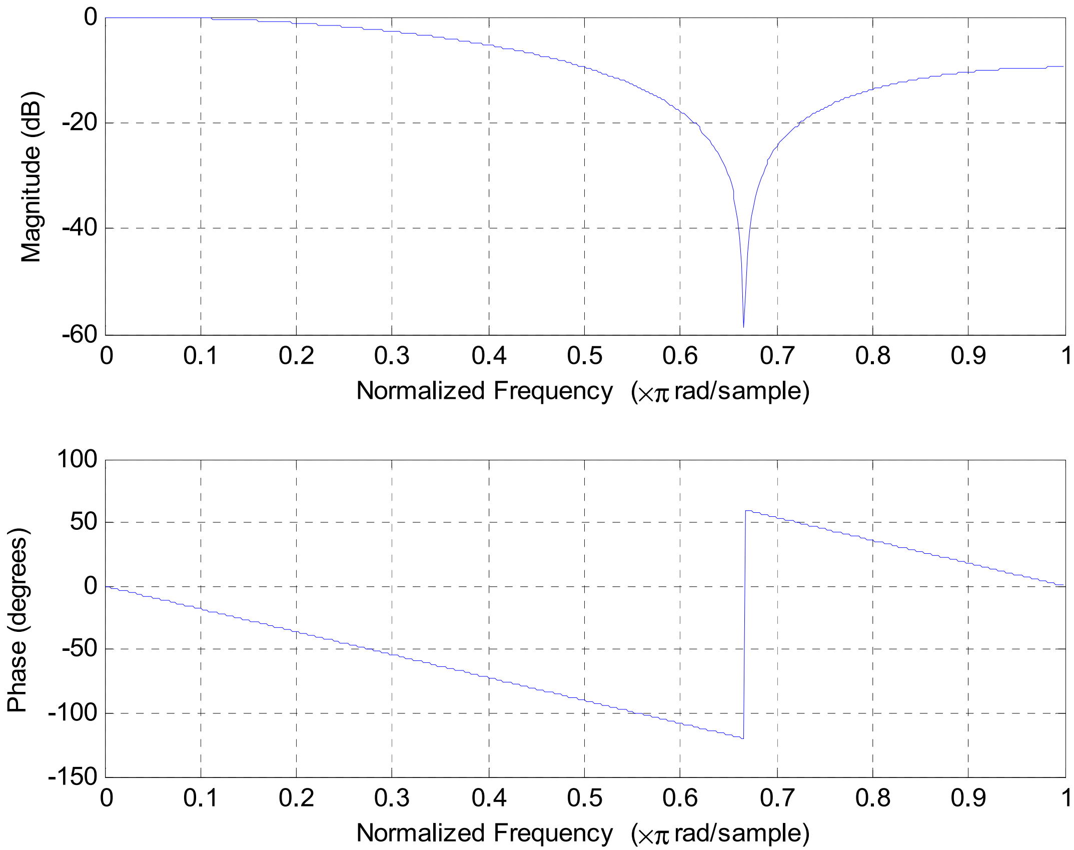
(snadněji se realizuje pomocí MA filtrů)

**Woldův dekompoziční teorém:**

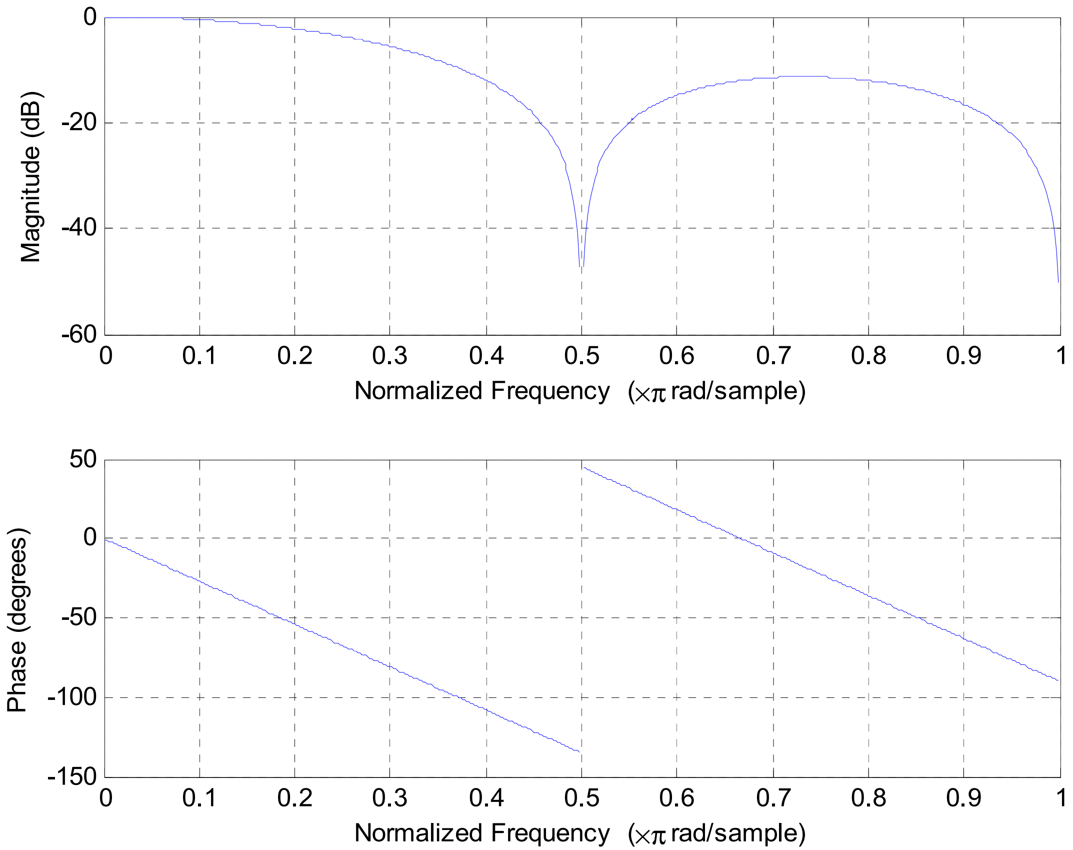
- ➔ jakýkoliv ARMA nebo MA proces může být jednoznačně reprezentován AR modelem max. ∞ řádu;
- ➔ jakýkoliv ARMA nebo AR proces lze reprezentovat MA modelem max. ∞ řádu;

$\Downarrow$

je nám jedno, co použijeme za model, jen by měl mít co nejméně parametrů, které se snadno počítají

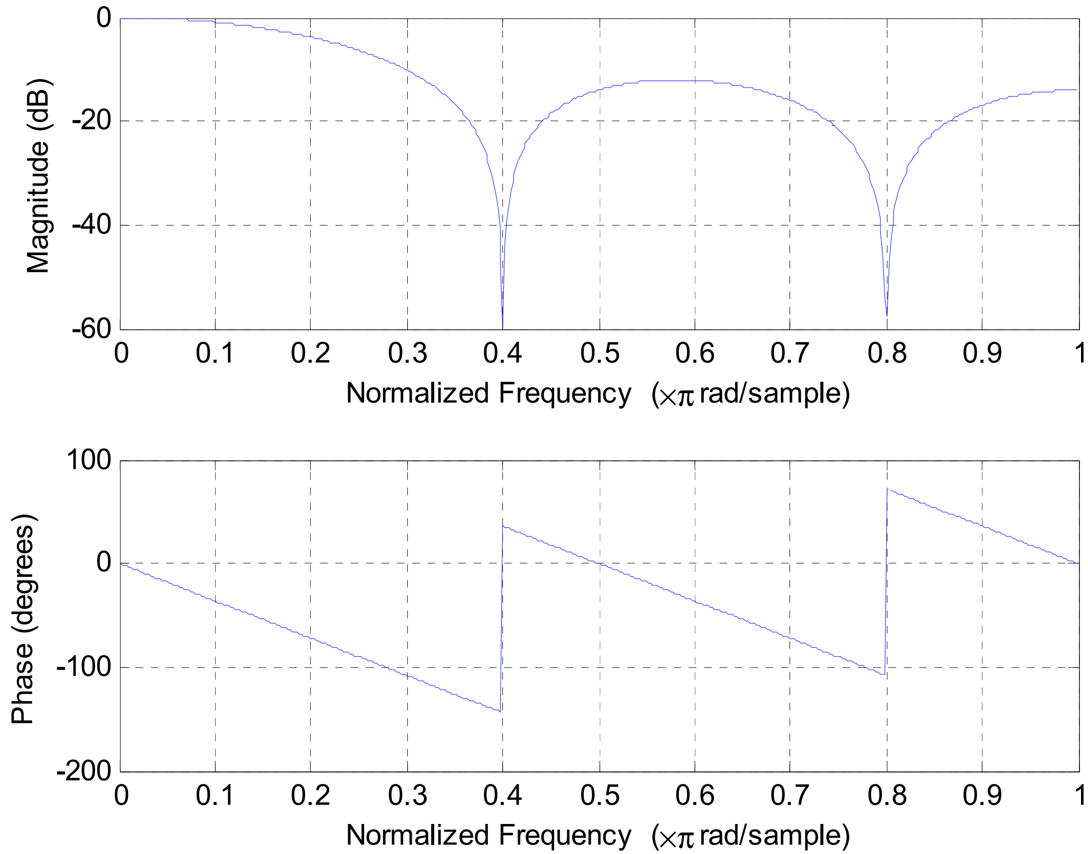

0
-20
-40
-60
Magnitude (dB)
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Normalized Frequency (×π rad/sample)
100
50
0
-50
-100
-150
Phase (degrees)
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Normalized Frequency (×π rad/sample)

b=(1,1,1) a=3

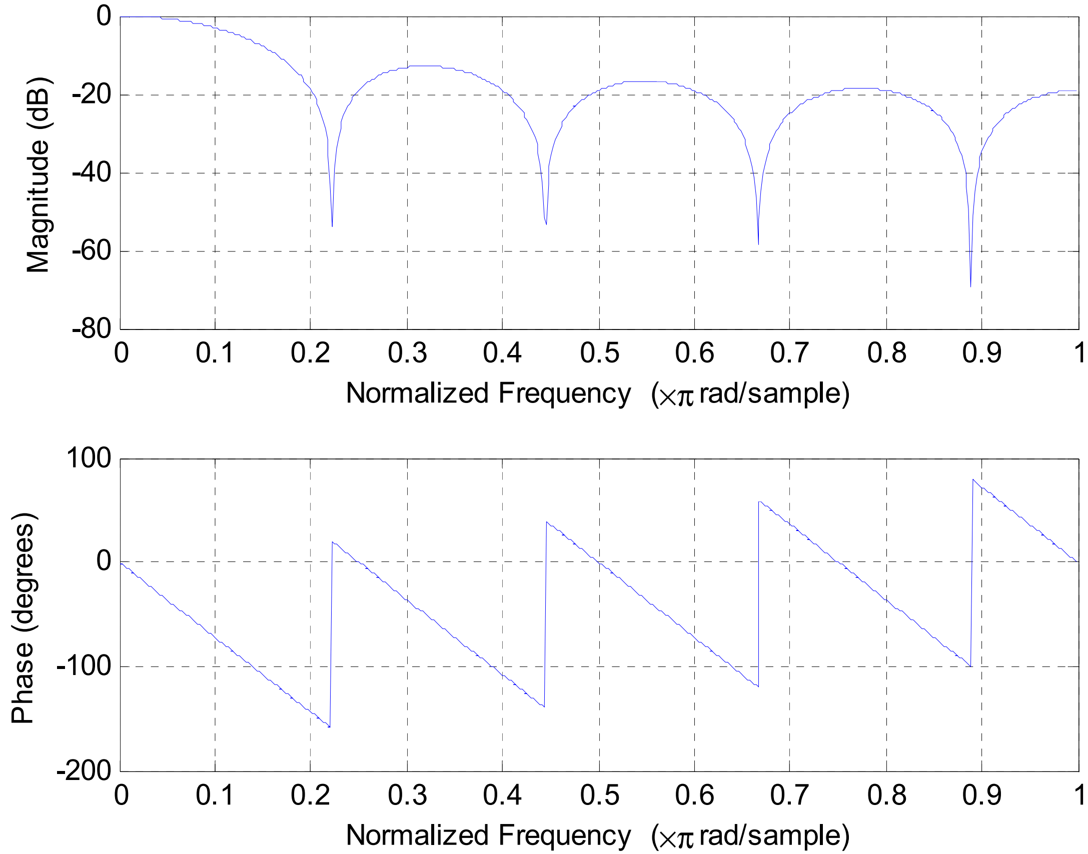


0
-20
-40
-60
Magnitude (dB)
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Normalized Frequency (×π rad/sample)
50
0
-50
-100
-150
Phase (degrees)
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Normalized Frequency (×π rad/sample)

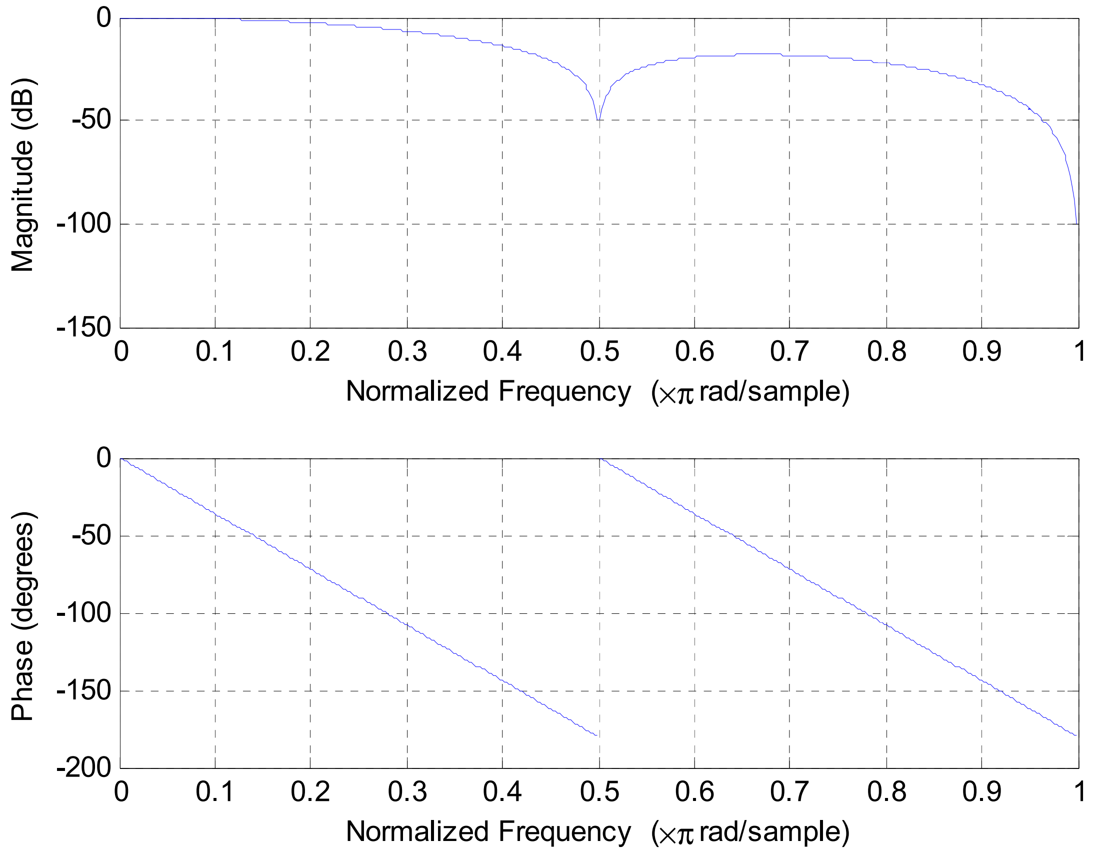

b=(1,1,1,1) a=4


0
-20
-40
-60
Magnitude (dB)
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Normalized Frequency (×π rad/sample)
100
0
-100
-200
Phase (degrees)
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Normalized Frequency (×π rad/sample)

b=(1,1,1,1,1) a=5

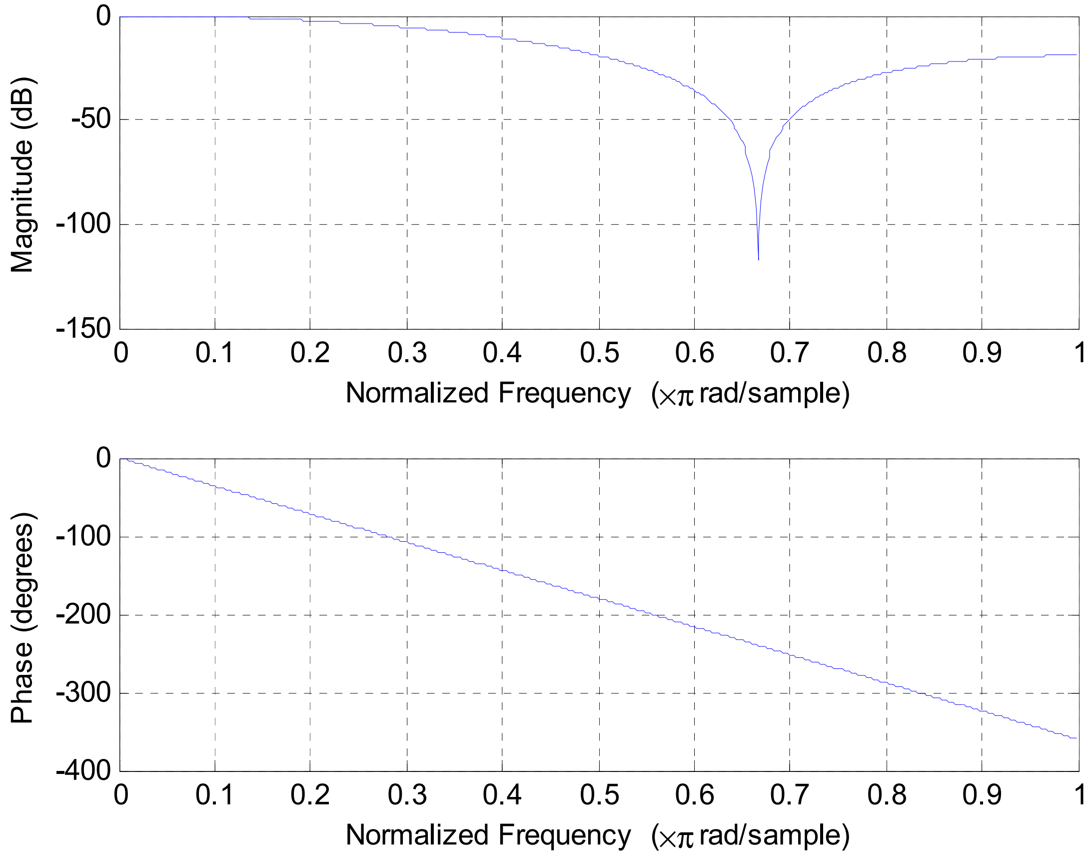



b=(1,1,1,1,1,1,1,1,1) a=9



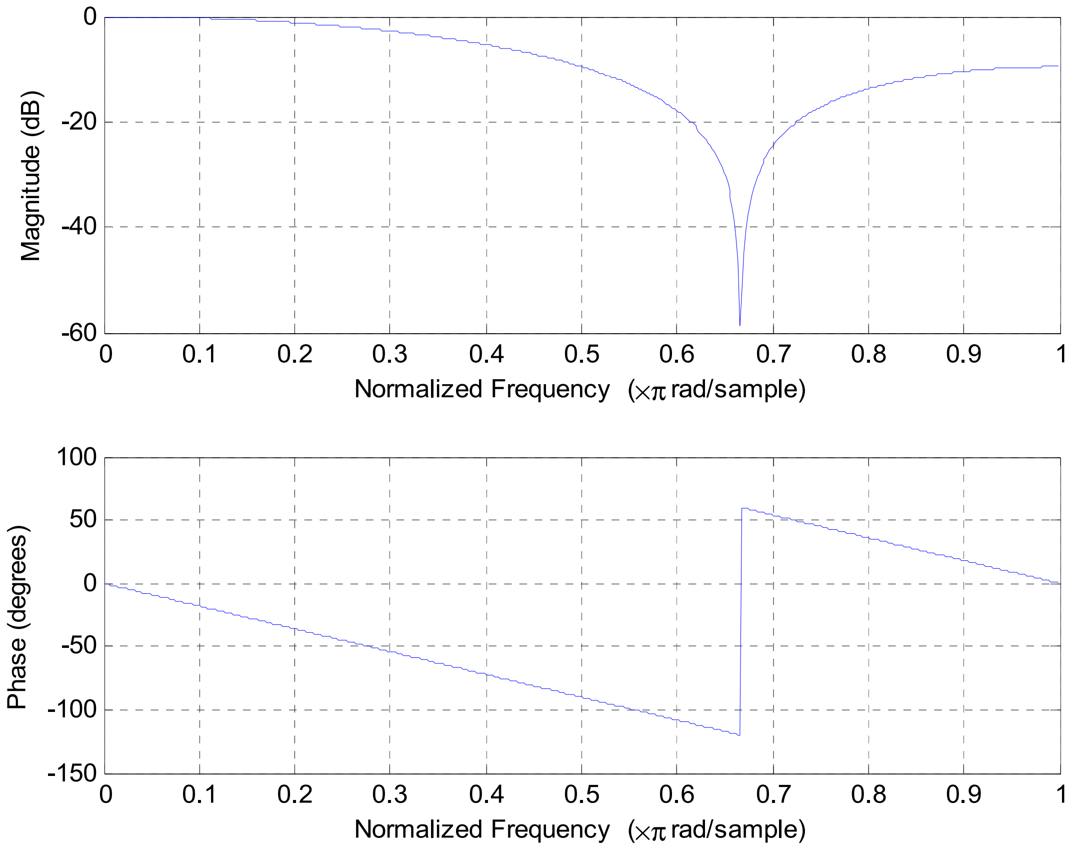
Magnitude (dB)
0
-50
-100
-150
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Normalized Frequency (×π rad/sample)
Phase (degrees)
0
-50
-100
-150
-200
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Normalized Frequency (×π rad/sample)


b=(1,2,2,2,1) a=8



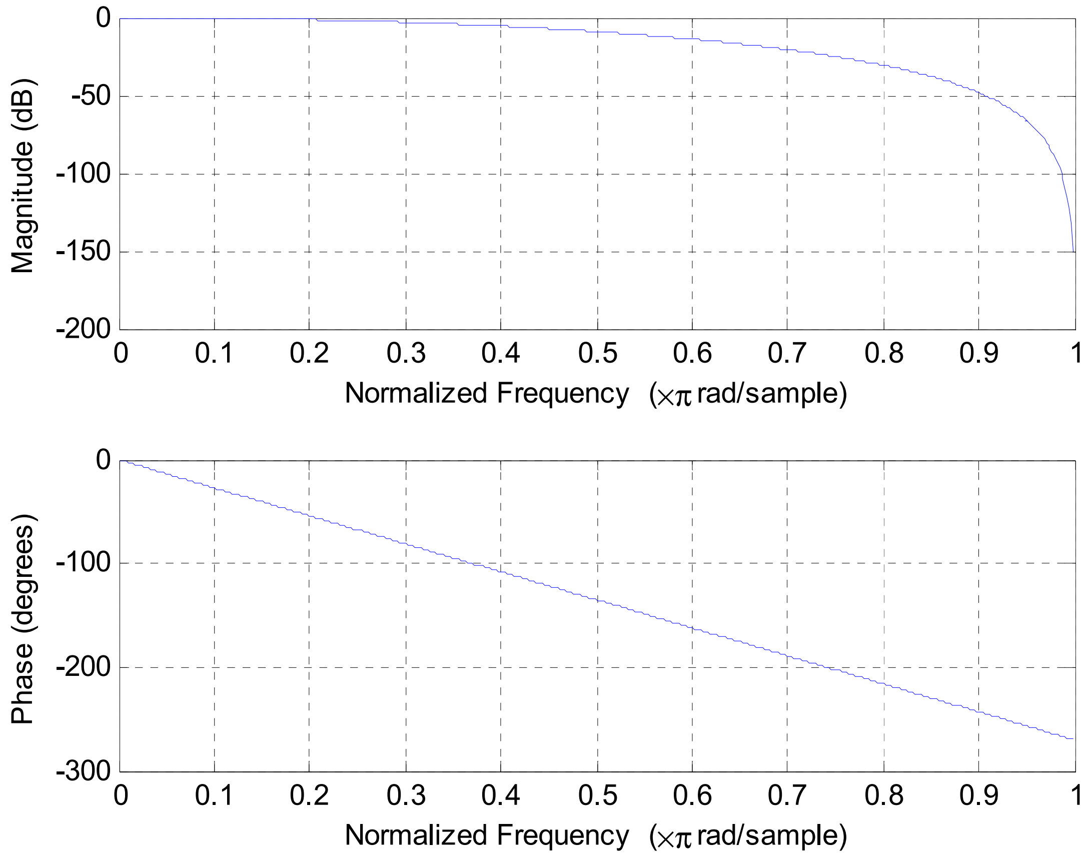

b=(1,2,3,2,1) a=9

$H(z) = (1+z^{-1}+z^{-2}).(1+z^{-1}+z^{-2})$



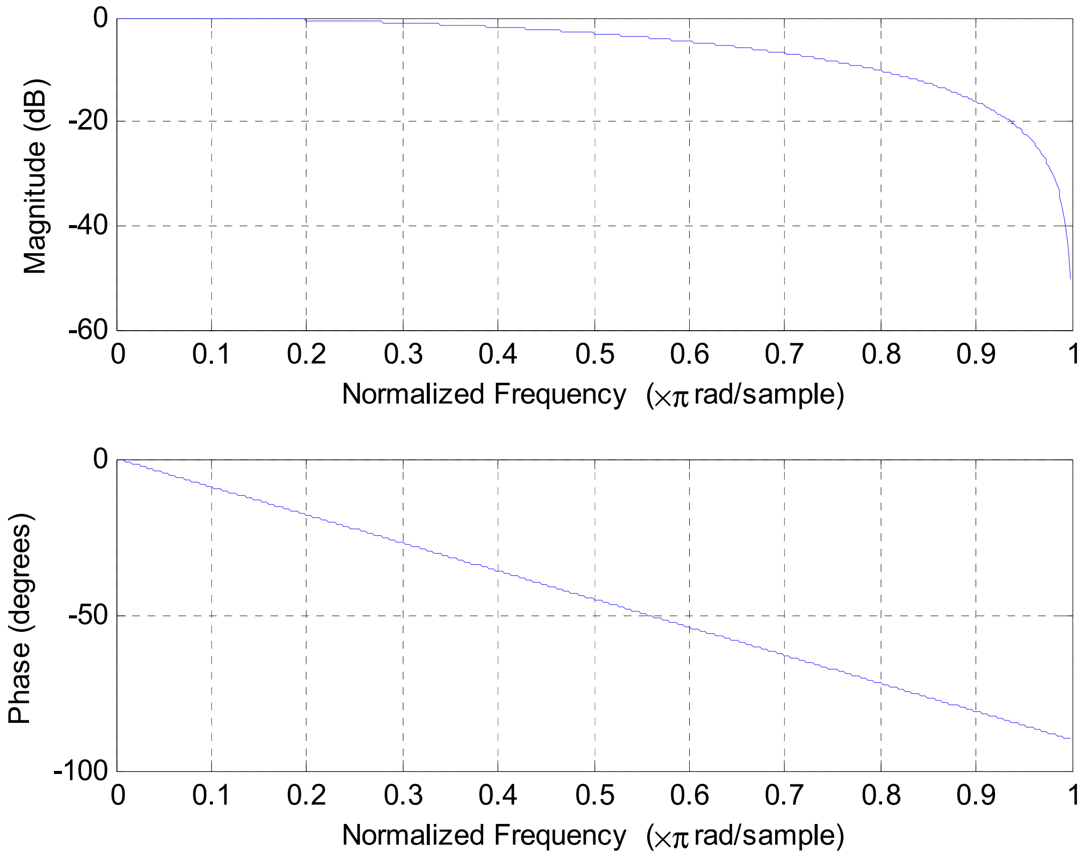

b=(1,1,1) a=3

$H(z)= (1+z^{-1}+z^{-2})$




b=(1,3,3,1) a=8

$H(z)= (1+z^{-1})^3$


0
-20
-40
-60
Magnitude (dB)
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Normalized Frequency (×π rad/sample)
0
-50
-100
Phase (degrees)
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Normalized Frequency (×π rad/sample)

$b=(1,1)$ $a=2$

$H(z)= (1+z^{-1})$

# ODHAD TRENDU (DRIFTU)

pomalu se měnící složky

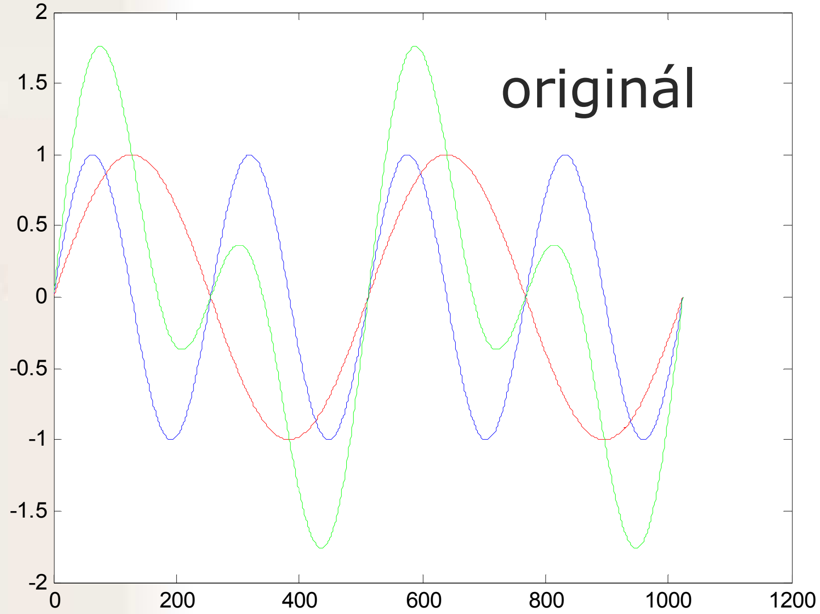
$\Downarrow$

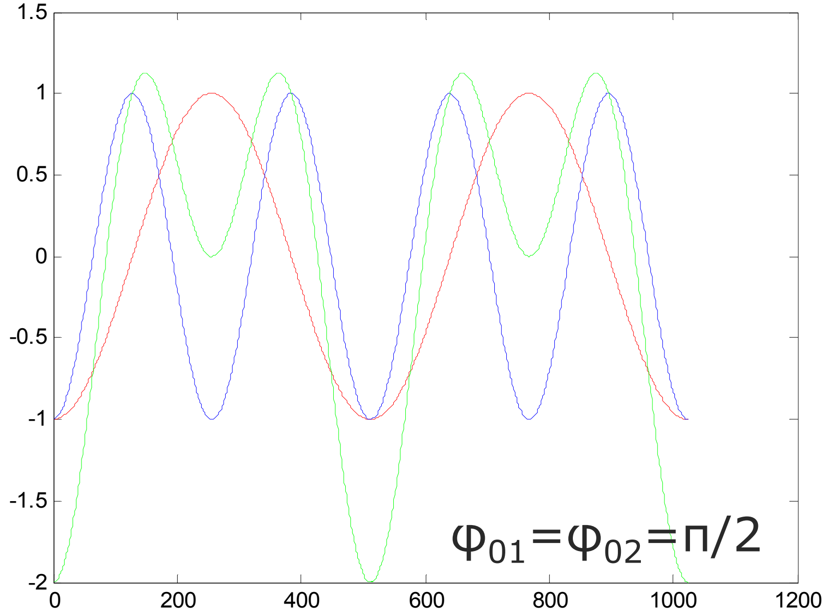
**DOLNÍ PROPUST**

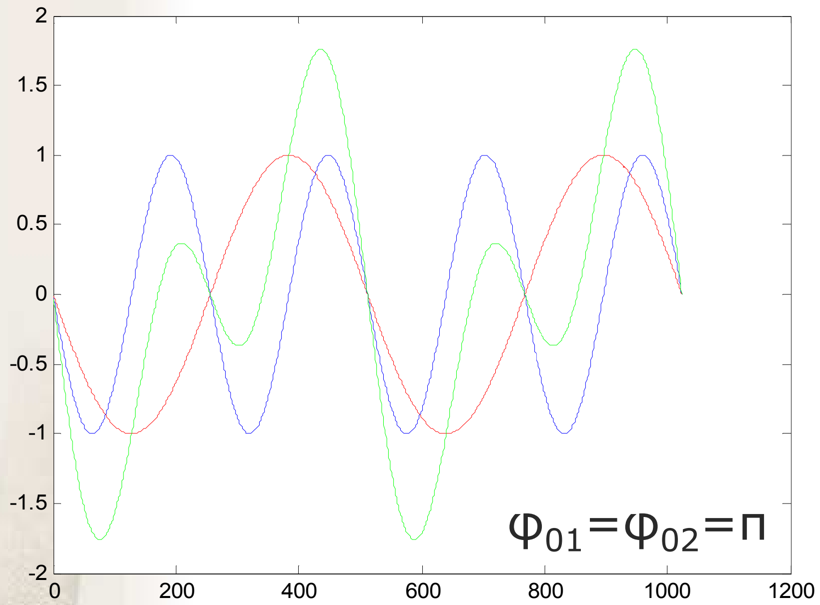
(snadněji se realizuje pomocí MA filtrů)

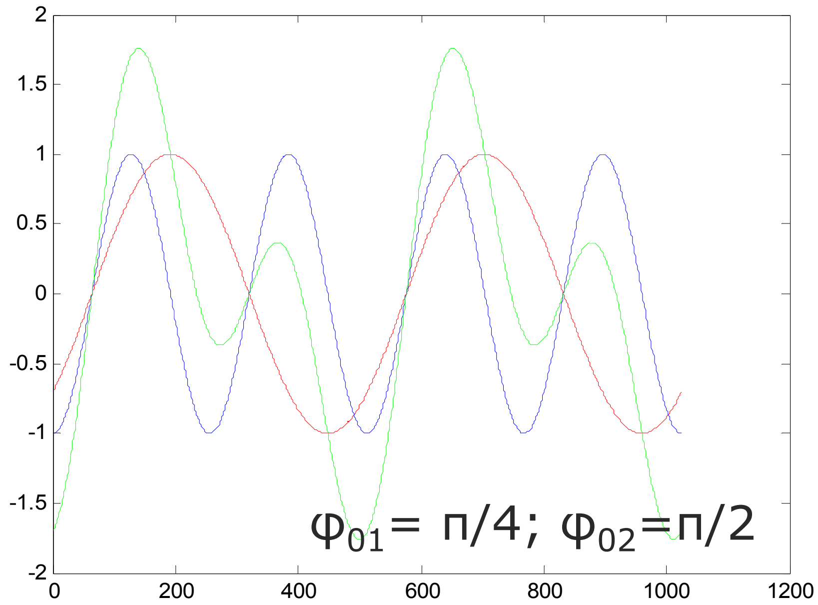
1. aby bylo zesílení filtru jednotkové, je třeba, aby $\sum_{i=1}^{m} |a_i| = 1$;
2. je třeba počítat průměr tak, aby nedocházelo ke zpoždění výstupních hodnot – jak toho dosáhnout?

# HRÁTKY S POČÁTEČNÍ FÁZÍ



# HRÁTKY S POČÁTEČNÍ FÁZÍ

- ☑ ke zkreslení nedošlo, pokud fázová charakteristika byla lineární;
- ☑ lineární je tehdy, pokud je impulsní charakteristika symetrická

# SKUPINOVÉ ZPOŽDĚNÍ

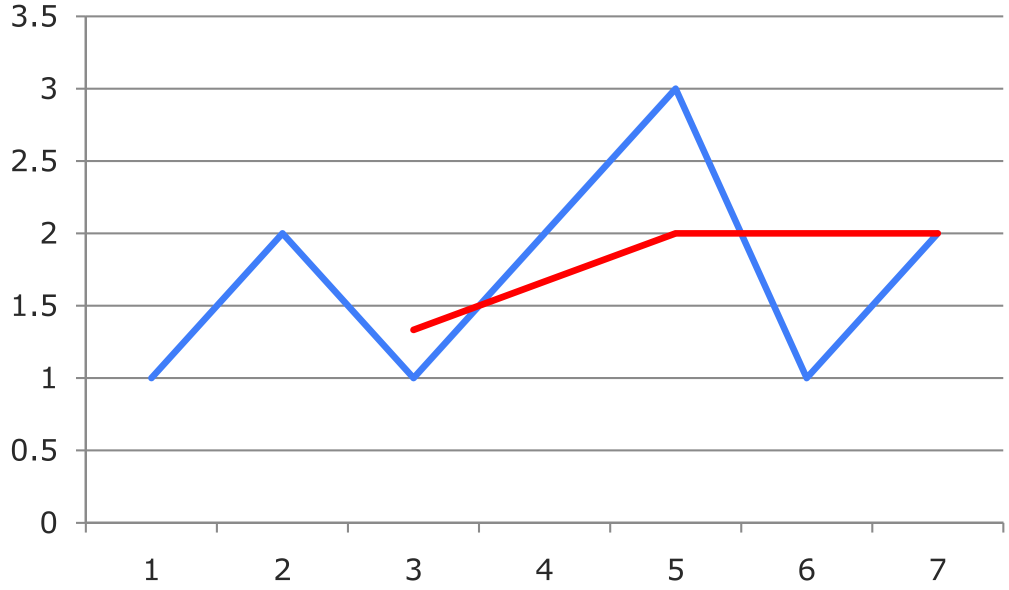
Záporná derivace fázové charakteristiky se nazývá skupinové zpoždění

$$\tau_g(\omega) = -\frac{d\varphi(\omega)}{d\omega}$$

Konstantní skupinové zpoždění znamená, že se všechny složky signálu kmitočtovém pásmu dostanou na výstup se stejným zpožděním.

Filtry řádu m s lineární fází mají skupinové zpoždění m/2 a filtrovaný signál je zpožděn o m/2 kroků.

# PŘECHODOVÝ DĚJ

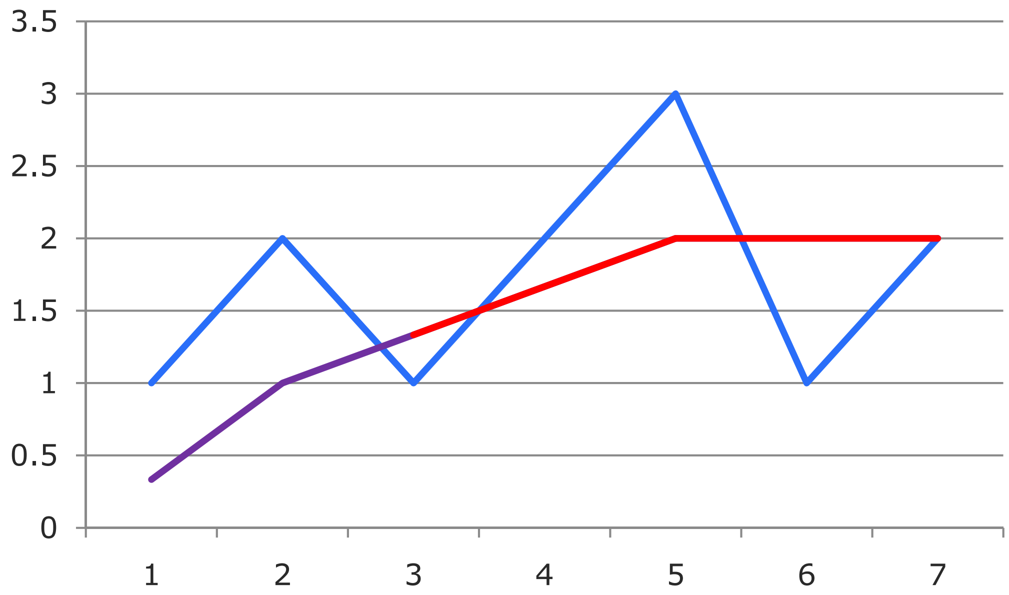



$x(nT) = \{1,2,1,2,3,1,2\}$

$g(nT) = \{1/3,1/3,1/3\}$

$y(nT) = [x(nT)+x(nT-T)+x(nT-2T)]/3$

# PŘECHODOVÝ DĚJ



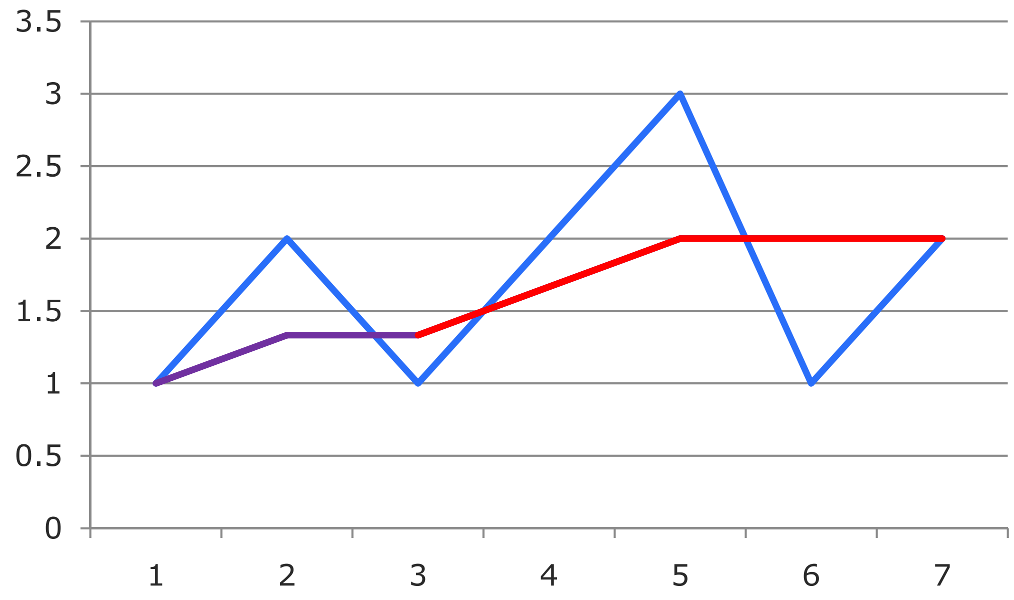

$x(nT) = \{1,2,1,2,3,1,2\}$

$g(nT) = \{1/3,1/3,1/3\}$

nulové počáteční podmínky: $x(0)=0$; $x(-T)=0$

$y(nT) = [x(nT)+x(nT-T)+x(nT-2T)]/3$

# PŘECHODOVÝ DĚJ




$x(nT) = \{1,2,1,2,3,1,2\}$

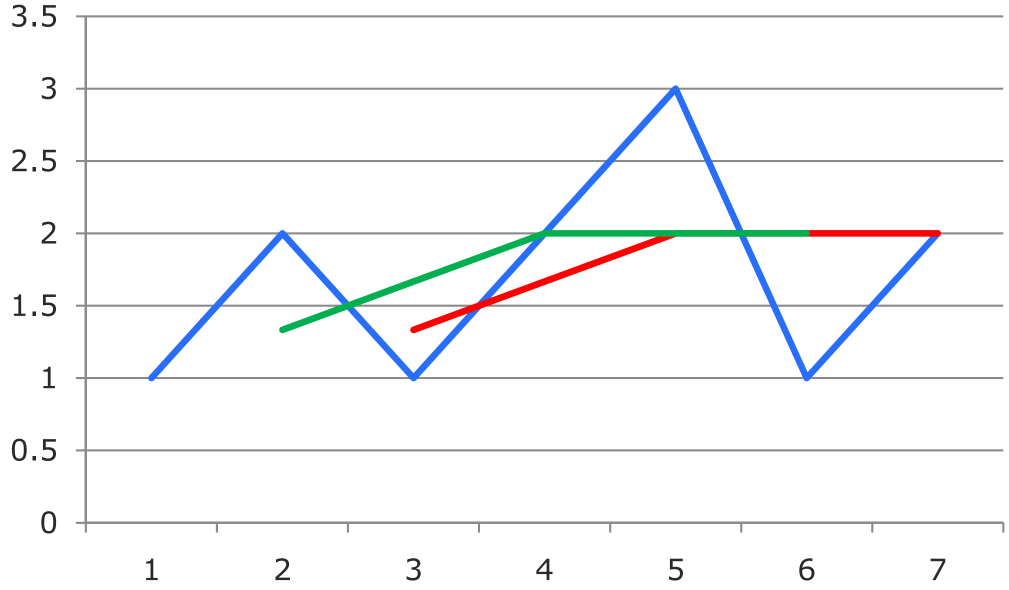
$g(nT) = \{1/3,1/3,1/3\}$

počáteční podmínky určeny první hodnotou vstupu:

$x(0)=1;\ x(-T)=1$

$y(nT) = [x(nT)+x(nT-T)+x(nT-2T)]/3$

# PŘECHODOVÝ DĚJ



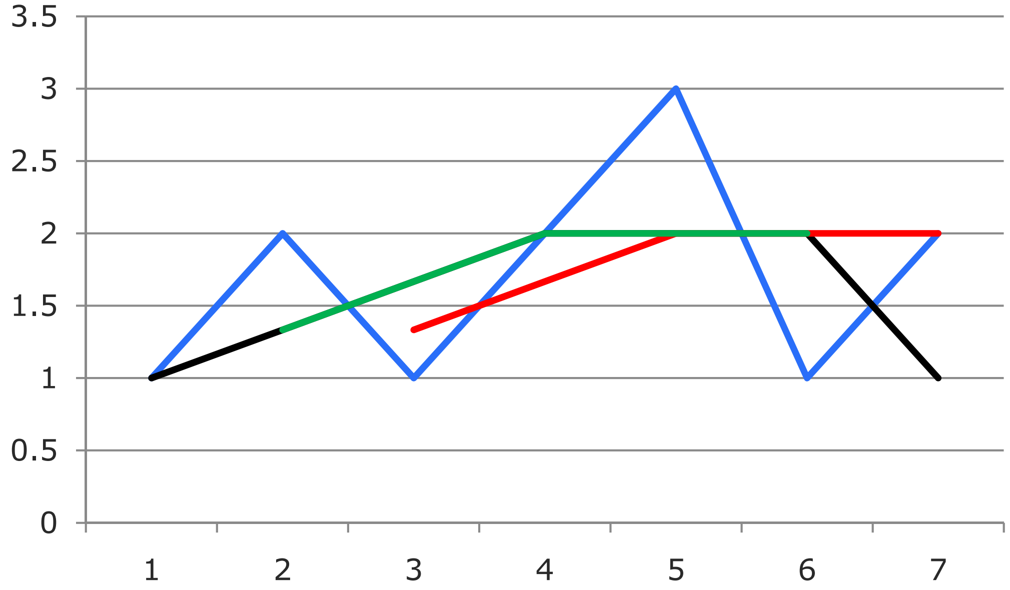

$x(nT) = \{1,2,1,2,3,1,2\}$

$g(nT) = \{1/3,1/3,1/3\}$

výpočet vůči střední hodnotě impulsní odezvy:

$y(nT) = [x(nT+T)+x(nT)+x(nT-T)]/3$

# PŘECHODOVÝ DĚJ




$x(nT) = \{1,2,1,2,3,1,2\}$

$g(nT) = \{1/3,1/3,1/3\}$

výpočet vůči střední hodnotě impulsní odezvy s nulovými počátečními a koncovými podmínkami

$y(nT) = [x(nT+T)+x(nT)+x(nT-T)]/3$

# ODHAD TRENDU (DRIFTU)

pomalu se měnící složky

$\Downarrow$

**DOLNÍ PROPUST**

(snadněji se realizuje pomocí MA filtrů)

1. aby bylo zesílení filtru jednotkové, je třeba, aby $\sum_{i=1}^{m} |a_i| = 1$;
2. je třeba počítat průměr tak, aby nedocházelo ke zpoždění výstupních hodnot jak toho dosáhnout?

# ODHAD TRENDU (DRIFTU)

## DOLNÍ PROPUST

(snadněji se realizuje pomocí MA filtrů)

2. je třeba počítat průměr tak, aby nedocházelo ke zpoždění výstupních hodnot
   - počítat podle standardní diferenční rovnice o výsledek posunout o hodnotu skupinového zpoždění;
   - počítat vůči středu impulsní odezvy.

## DOLNÍ PROPUST

(snadněji se realizuje pomocí MA filtrů)

3. je třeba počítat s přechodným dějem o délce impulsní odezvy na začátku výstupu nebo o polovině délky impulsní odezvy na začátku a na konci

# ODHAD TRENDU (DRIFTU)

## DOLNÍ PROPUST

(snadněji se realizuje pomocí MA filtrů)

Jak odhadnout mezní frekvenci DP?

praktické pravidlo:

určíme oscilační složku o nejnižší frekvenci a mezní frekvenci DP pro odhad driftu stanovíme jako nejvyšší frekvenci propustného pásma filtru, který neovlivní oscilační složku

Příprava nových učebních materiálů
pro obor Matematická biologie

je podporována projektem ESF
č. CZ.1.07/2.2.00/07.0318

## „VÍCEOBOROVÁ INOVACE STUDIA MATEMATICKÉ BIOLOGIE“

EVROPSKÁ UNIE

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

OP Vzdělávání pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ