

IIR filtry

Vlastnosti IIR filtrů:

Výhody:

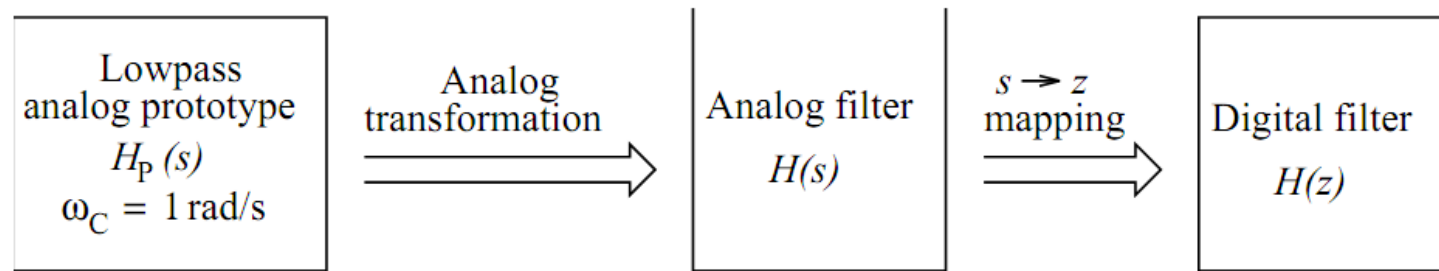
- jsou výrazně nižšího řádu než Fir filtry se stejnými vlastnostmi a z toho vyplývá že mají:
 - nižší výpočetní složitost v porovnání s Fir filtrem
 - kratší zpoždění mezi vstupem a výstupem

Nevýhody:

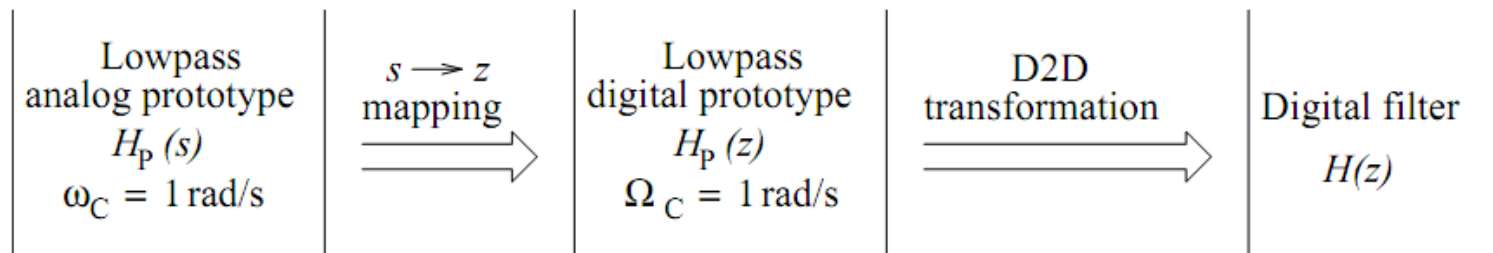
- není u nich zaručena stabilita
- fáze není lineární, a doladění fázové charakteristiky je obtížné
- obecný návrh je relativně složitý v porovnání s FIR
- jsou citlivé na numerickou přesnost výpočtů

Základní metody návrhu IIR filtrů

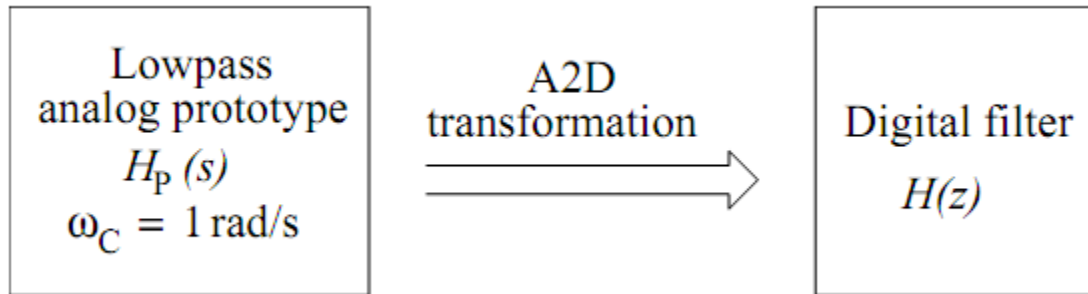
1. Návrh analogové dolní propusti $H(s)$ a konverze na číslicovou $H(z)$
Nelze použít pro horní a pásmovou propust.



2. Návrh analogové dolní propusti, konverze na číslicovou dolní propust a spektrální transformace (A2D) pro konverzi na libovolný typ filtru.



3. Úprava frekvenční charakteristiky a použití bilineární transformace ke konverzi analogového prototypu na požadovaný číslicový filtr



Návrh IIR filtrů

3 základní kroky návrhu:

1. stanovení parametrů frekvenční charakteristiky filtru a návrh ekvivalentního analogového filtru
2. transformace přenosové funkce analogového filtru $H(p)$ (nebo $H(s)$) do diskrétní oblasti $H(z)$
3. HW nebo SW realizace filtru

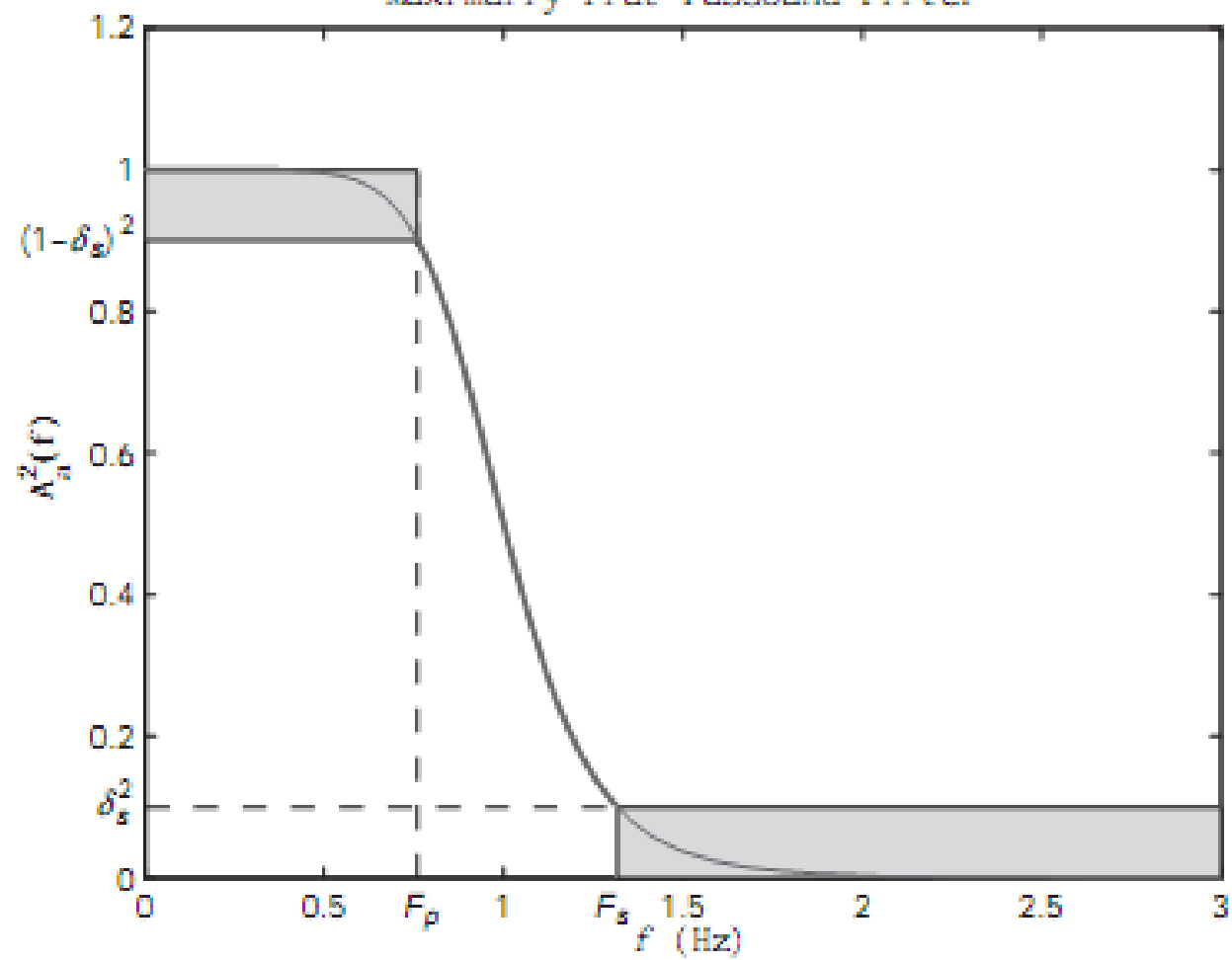
Ad 1) Určíme parametry filtru (toleranční schéma) a aproximujeme navrženou amplitudovou frekvenční charakteristiku. Jednotlivé typy aproximací:

- Butterworthova aproximace
- Čebyševova aproximace 1
- Čebyševova aproximace 2 (inverzní Čebyšev)
- Eliptická aproximace

Butterworthova aproximace

- požadovaná frekvenční odezva se v propustném i nepropustném pásmu aproximuje Taylorovou řadou
- aproximovaná frekvenční charakteristika je maximálně plochá a monotónní bez zákmitů od hodnoty $|H(j\Omega)|=1$ pro $\Omega=0$ k hodnotě $|H(j\Omega)|=0$ pro $\Omega=\infty$
- návrh tohoto typu je jednoduchý
- nevýhoda: široké přechodové pásmo mezi propustným a nepropustným pásmem, které lze zúžit za cenu vyššího řádu (a tím pádem složitější realizace)

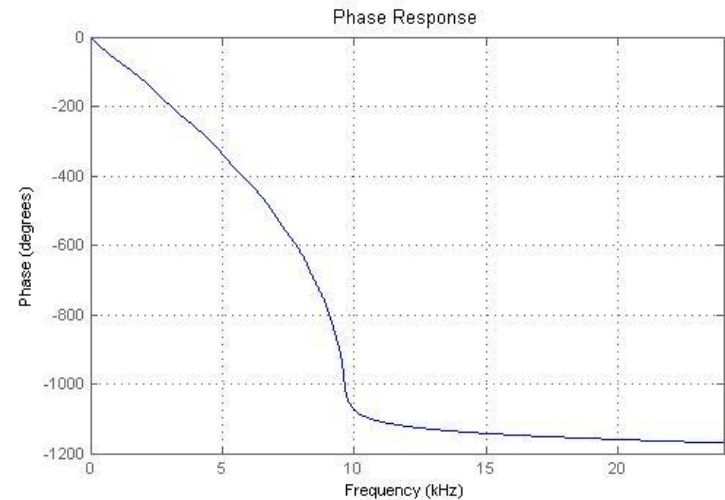
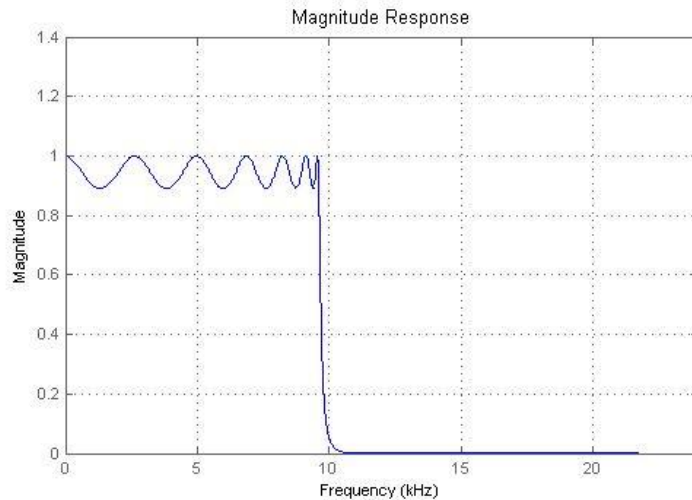
Maximally Flat Passband Filter



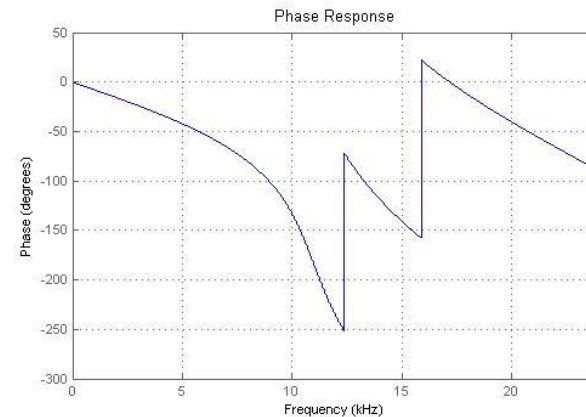
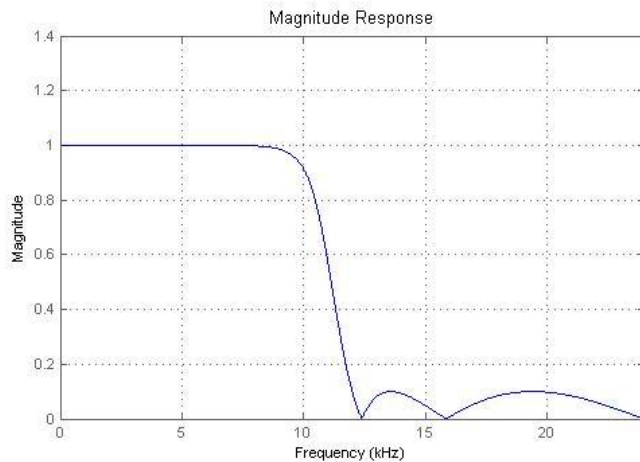
Čebyševova aproximace

- Používají se 2 druhy aproximací v závislosti na tom v které oblasti připustíme zvlnění

Čebyšev I – v propustném pásmu se aproximuje Čebyševovým polynomem → zvlněný (izoextremální) průběh, nepropustná část je hladká (aproximace Taylorovým polynomem)

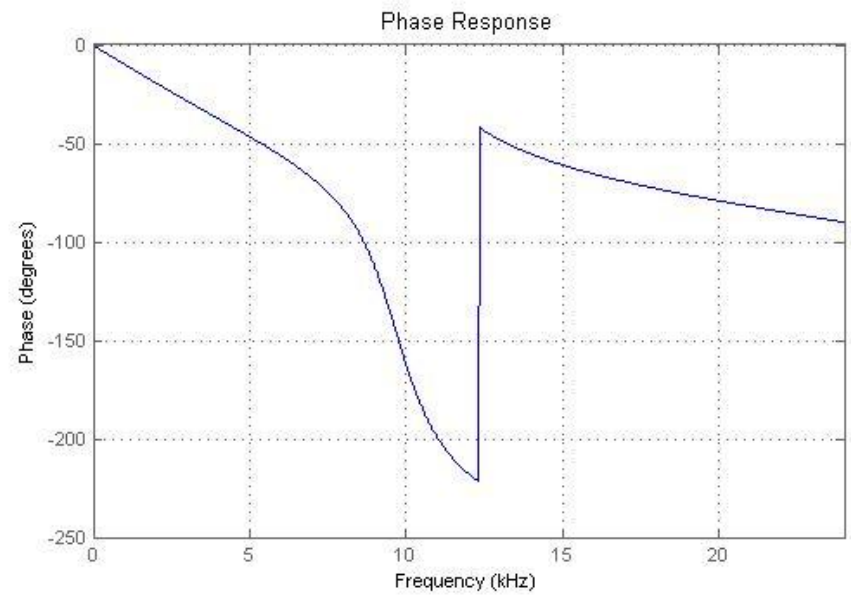
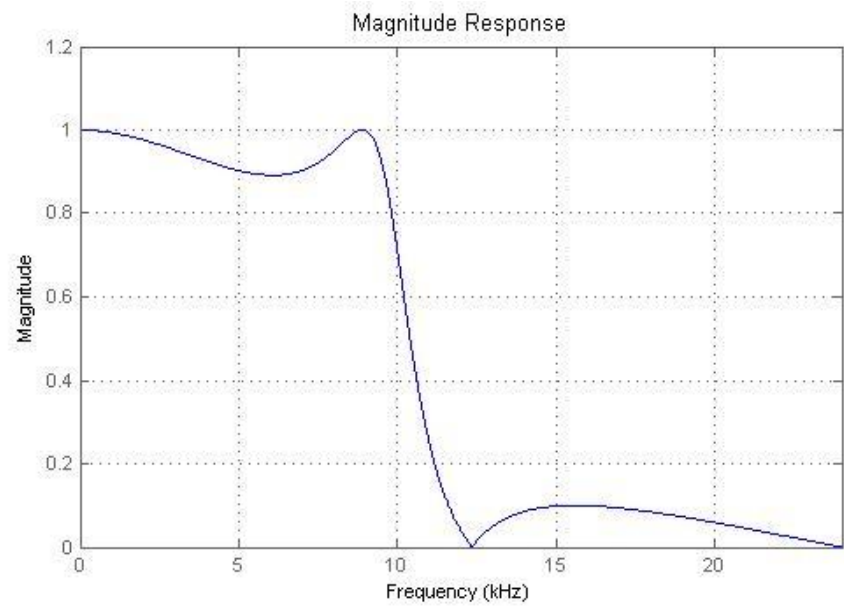


- Čebyšev II (inverzní Čebyšev) – aproximace v propustném pásmu Taylorovým polynomem (hladký průběh), v nepropustném pásmu Čebyševovým polynomem (zvlněný průběh)



Elíptická aproximace (Cauerův filtr)

- Propustná i nepropustná část je aproximována Čebyševovým polynomem (zvlněný průběh)
- Pro zadané požadavky lze tímto způsobem navrhnout filtr nejnižšího řádu, ruční návrh je z uvedených typů filtrů nejsložitější
- Obvykle se zadávají 3 z následujících 4 parametrů, čtvrtý parametr vychází minimální
 - maximální zvlnění v propustném pásmu δ_1
 - šířka přenosového pásma
 - maximální zvlnění v nepropustném pásmu δ_2
 - řád filtru N



Obecný postup při návrhu IIR filtru

Vstup : toleranční charakteristika tj. mezní frekvence f_p , f_s , pro propustné a nepropustné pásmo, útlum v propustné a nepropustné části , vzorkovací frekvence S .

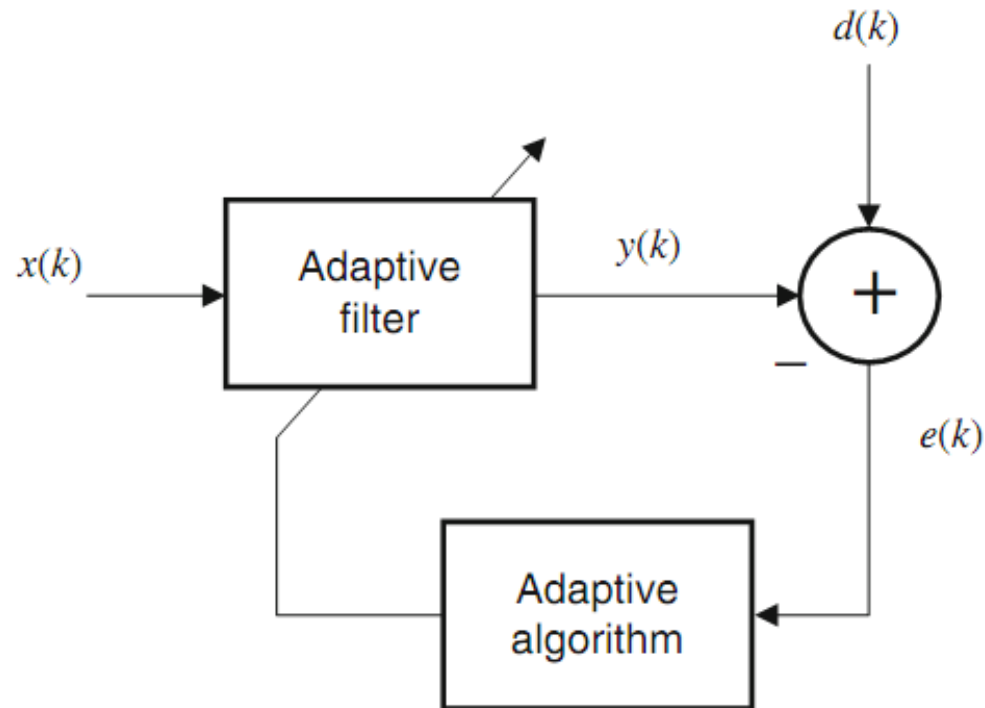
Postup:

1. Normalizace mezních frekvencí a převedení na F_p , F_s . To umožňuje použití vzorkovacího intervalu $t_s = 1$.
2. Použití normalizovaných frekvencí a útlumů pro návrh analogové dolní propusti $H_p(s)$ s mezní úhlovou frekvencí $\omega_c = 1\text{rad/s}$
3. Konverze z $H_p(s)$ na číslicovou dolní propust $H_p(z)$ (s $\Omega_d = 1$) zvolenou mapovací funkcí (s hodnotou $t_s = 1$).
4. Konverze $H_p(z)$ na $H(z)$ spektrální transformací (s $\Omega_d = 1$).
5. Pro bilineární transformaci použít A2D transformaci pro konverzi z $H_p(s)$ na $H(z)$.

Adaptivní filtrace

Adaptivní filtry jsou nelineární systémy, které přizpůsobují své chování parametrům vstupního signálu.

Obecná struktura adaptivního filtru:

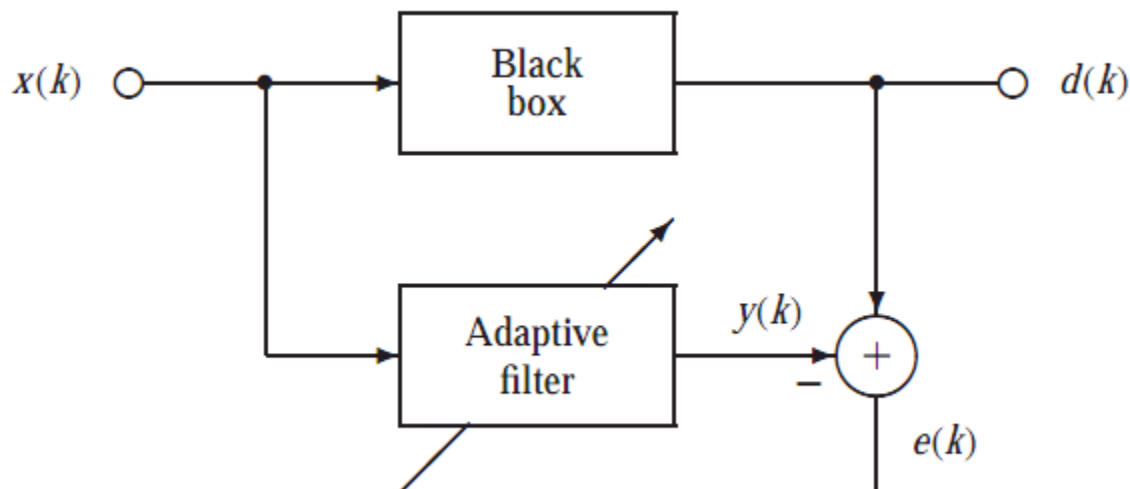


Kompletní specifikace adaptivního systému (filtru) se skládá ze tří kroků:

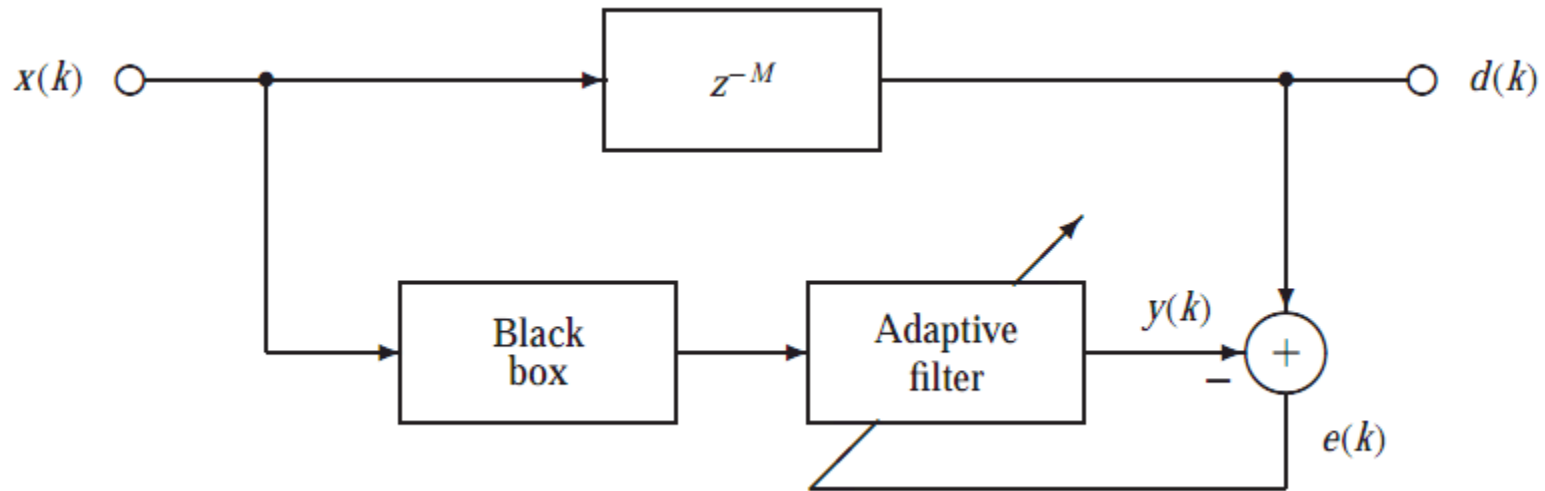
1. Volba typu aplikace (volba signálů získaných z prostředí tj. vstupního a požadovaného signálu)
2. Volba struktury adaptivního filtru
 - a) Adaptivní FIR – nejčastěji se používá tzv. transversální struktura
 - b) Adaptivní IIR – nejčastěji kanonická přímá forma realizace
3. Volba adaptivního algoritmu pro nastavení parametrů filtru
 - Definice funkce F která je funkcí vstupu, požadovaného výstupu
 - Definice minimalizačního algoritmu pro Funkci F

Typické aplikace adaptivního filtru:

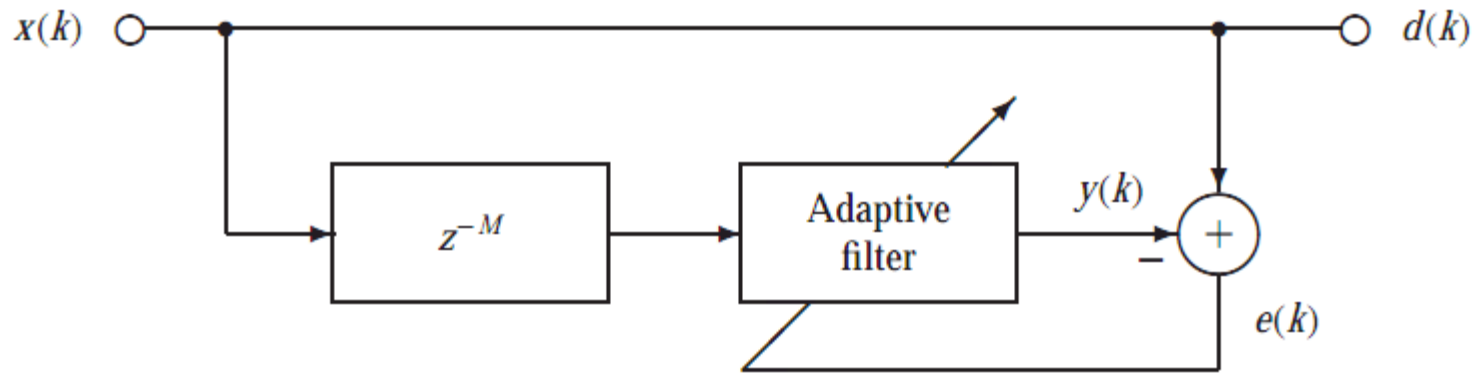
1. Identifikace systému - máme neznámý systém (černou skříňku) a snažíme se určit jeho charakteristiku (impulzní odezvu)



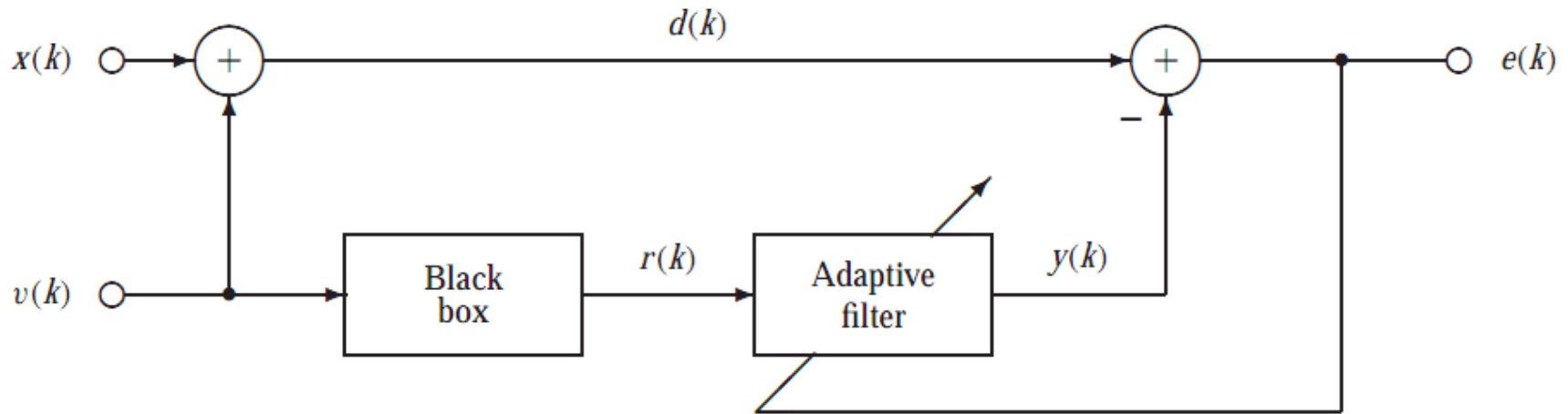
2. Ekvalizace kanálu – použití v komunikacích (obvykle při spojení dvou modemů). Filtr eliminuje charakteristiku komunikační linky, která může zatěžovat přenášený signál šumem. Černá skříňka (black box) reprezentuje neznámý komunikační kanál. Zpožďovací linka kompenzuje zpoždění sériového spojení filtru a komunikační linky.



3. Predikce signálu - používá se při kompresi řečového signálu před uložením nebo přenosem přes komunikační kanál.



4. Potlačení šumu - používá se k potlačení okolního rušení a zvýraznění užitečného signálu (např. telefonování za jízdy v automobilu).



$x(k)$ – řečový signál

$v(k)$ – okolní rušení (snímané dalším vhodně umístěným mikrofonem)

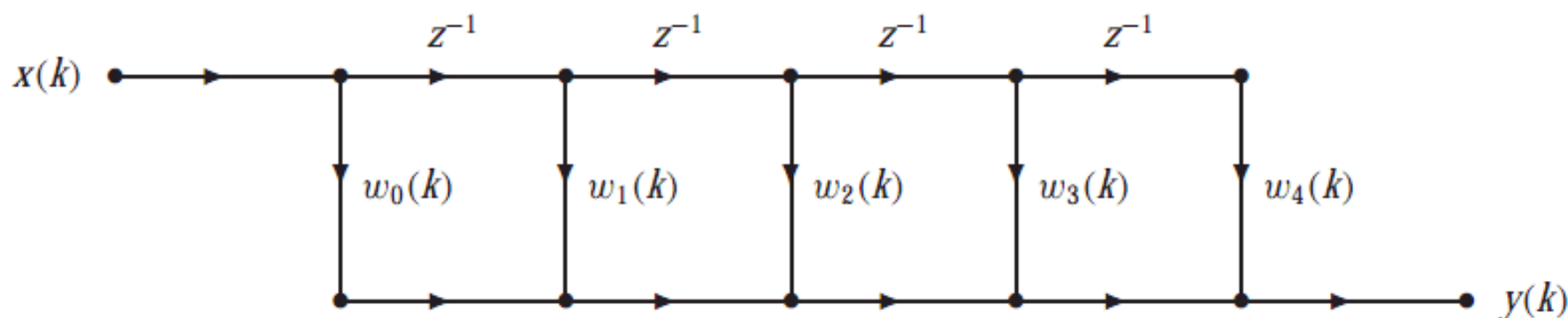
$e(k)$ – vyčištěný signál

Jako adaptivní filtr se nejčastěji používá FIR transversální filtr jehož diferenční rovnice má tvar

$$y(k) = \sum_{i=0}^m w_i(k) x(k-i)$$

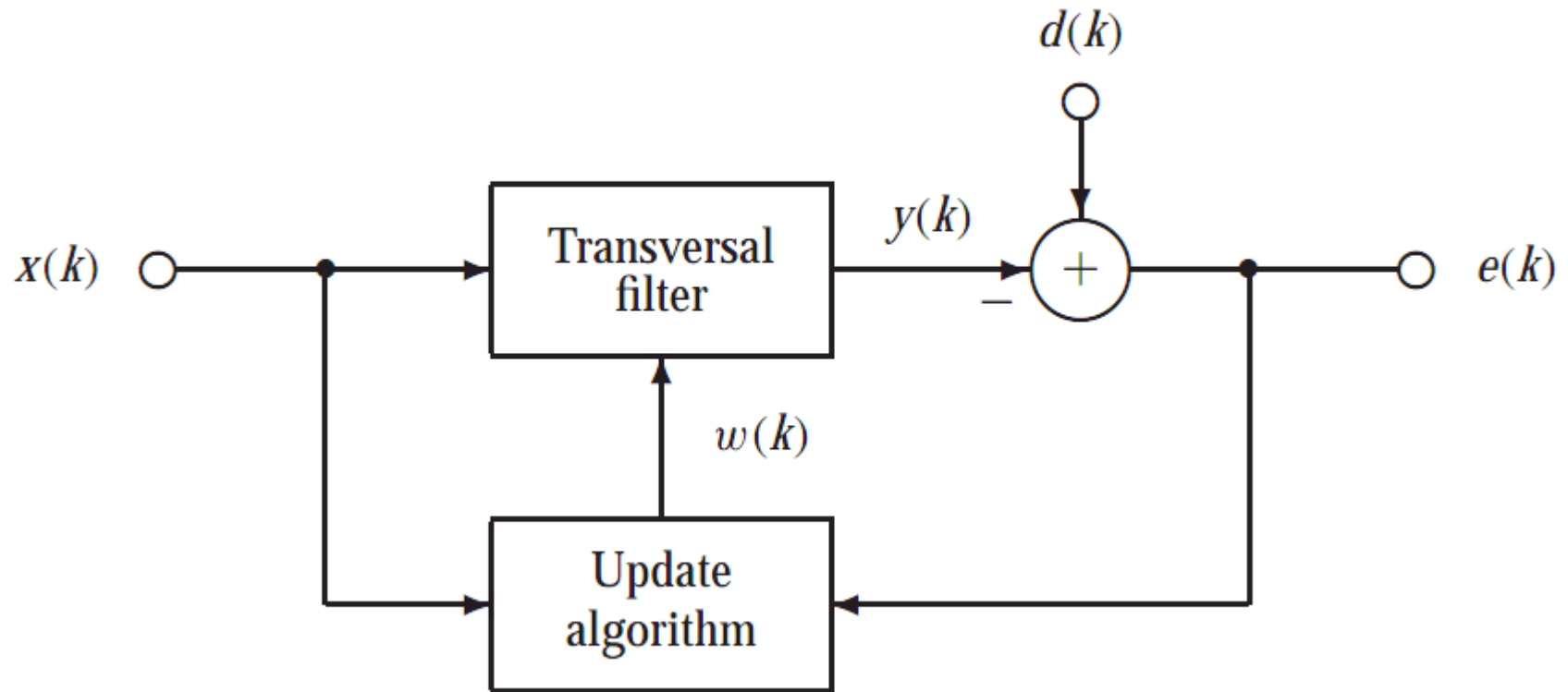
$$u(k) \triangleq [x(k), x(k-1), \dots, x(k-m)]^T$$

$$w(k) \triangleq [w_0(k), w_1(k), \dots, w_m(k)]^T$$



$$y(k) = w^T(k) u(k), \quad k \geq 0$$

Detailní struktura adaptivního filtru



Adaptační algoritmus:

- LMS (Least Mean Square)
- modifikovaný LMS
- RLS (Recursive Least-Square)

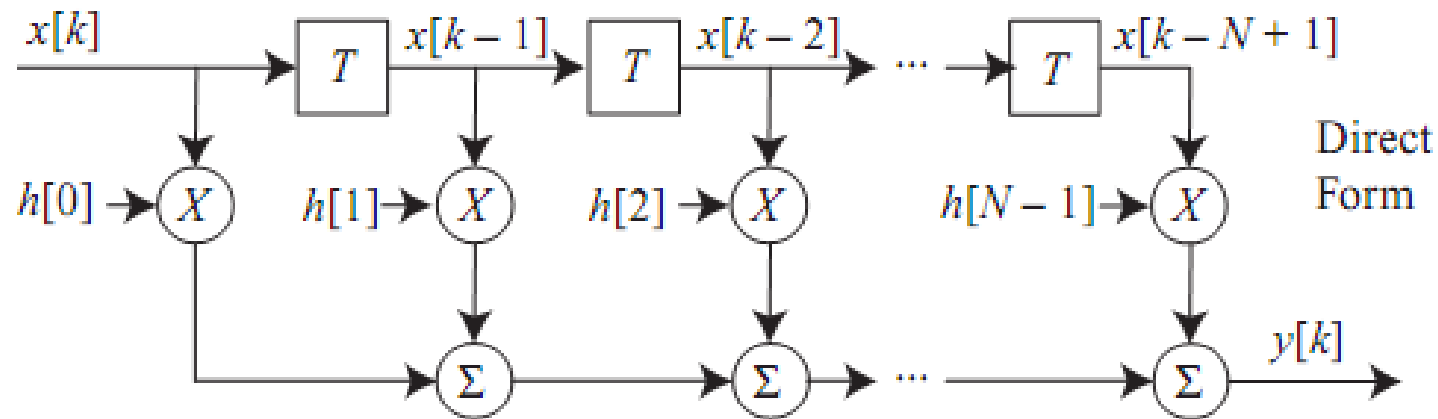
9.34 (Decoding a Mystery Message) During transmission, a message signal gets contaminated by a low-frequency signal and high-frequency noise. The message can be decoded only by displaying it in the time domain. The contaminated signal $x[n]$ is provided on the author’s website as `mystery1.mat`. Load this signal into MATLAB (use the command `load mystery1`). In an effort to decode the message, try the following methods and determine what the decoded message says.

- (a) Display the contaminated signal. Can you “read” the message? Display the DFT of the signal to identify the range of the message spectrum.
- (b) Use the bilinear transformation to design a second-order IIR bandpass filter capable of extracting the message spectrum. Filter the contaminated signal and display the filtered signal to decode the message. Use both the `filter` (filtering) and `filfilt` (zero-phase filtering) commands.
- (c) As an alternative method, first zero out the DFT component corresponding to the low-frequency contamination and obtain the IDFT $y[n]$. Next, design a lowpass IIR filter (using impulse invariance) to reject the high-frequency noise. Filter the signal $y[n]$ and display the filtered signal to decode the message. Use both the `filter` and `filtfilt` commands.
- (d) Which of the two methods allows better visual detection of the message? Which of the two filtering routines (in each method) allows better visual detection of the message?

Realizační struktury filtrů

FIR filtry:

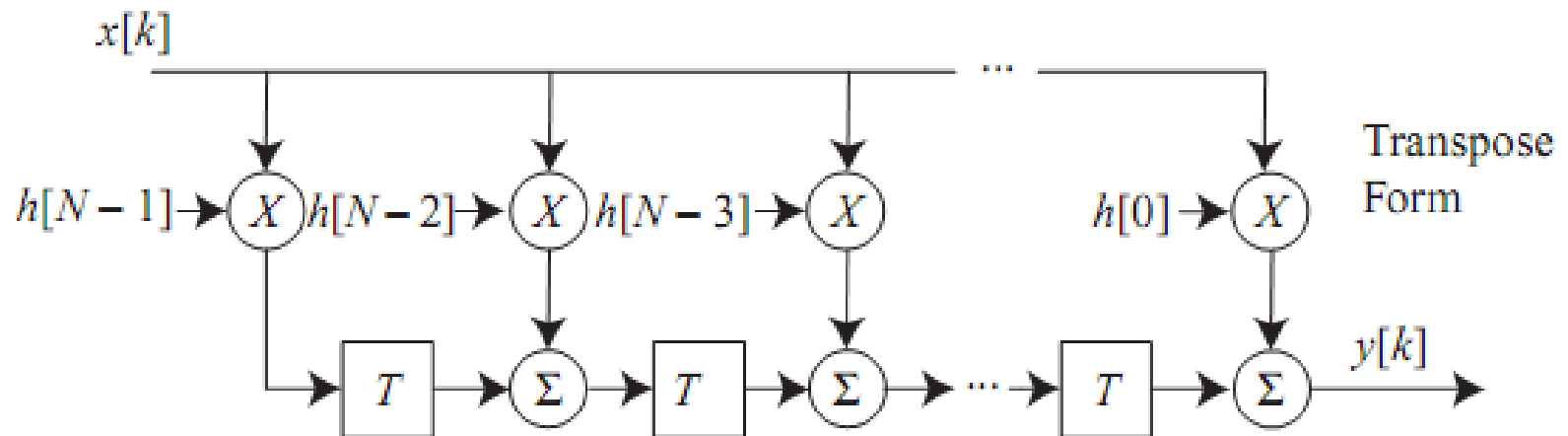
1. Přímá forma FIR (Direct-form FIR, transversal structure)



$$y(n) = \sum_{m=0}^{N-1} h(m)x(n-m)$$

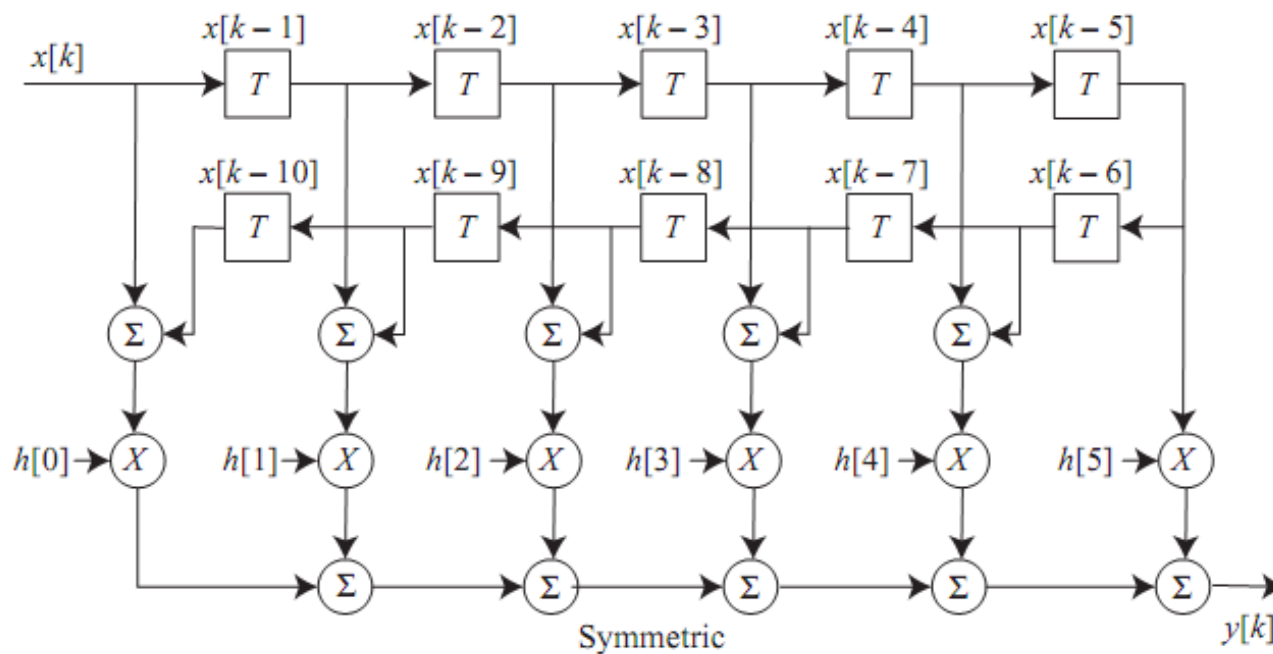
- Struktura filtru vychází přímo z výstupní rovnice filtru
- pro filtr N-tého řádu je potřeba:
 - N-1 posuvných registrů
 - N-1 paměťových buněk pro uchování vstupních vzorků
 - N buněk pro uchování koeficientů
 - N násobiček
 - N – 1 sčítaček
- snadná realizace v DSP procesorech, je to jedna z nejčastěji používaných struktur pro FIR filtry (používá se i v adaptivních FIR filtrech)

Transpose architektura - je variací na přímou formu, má podobné vlastnosti



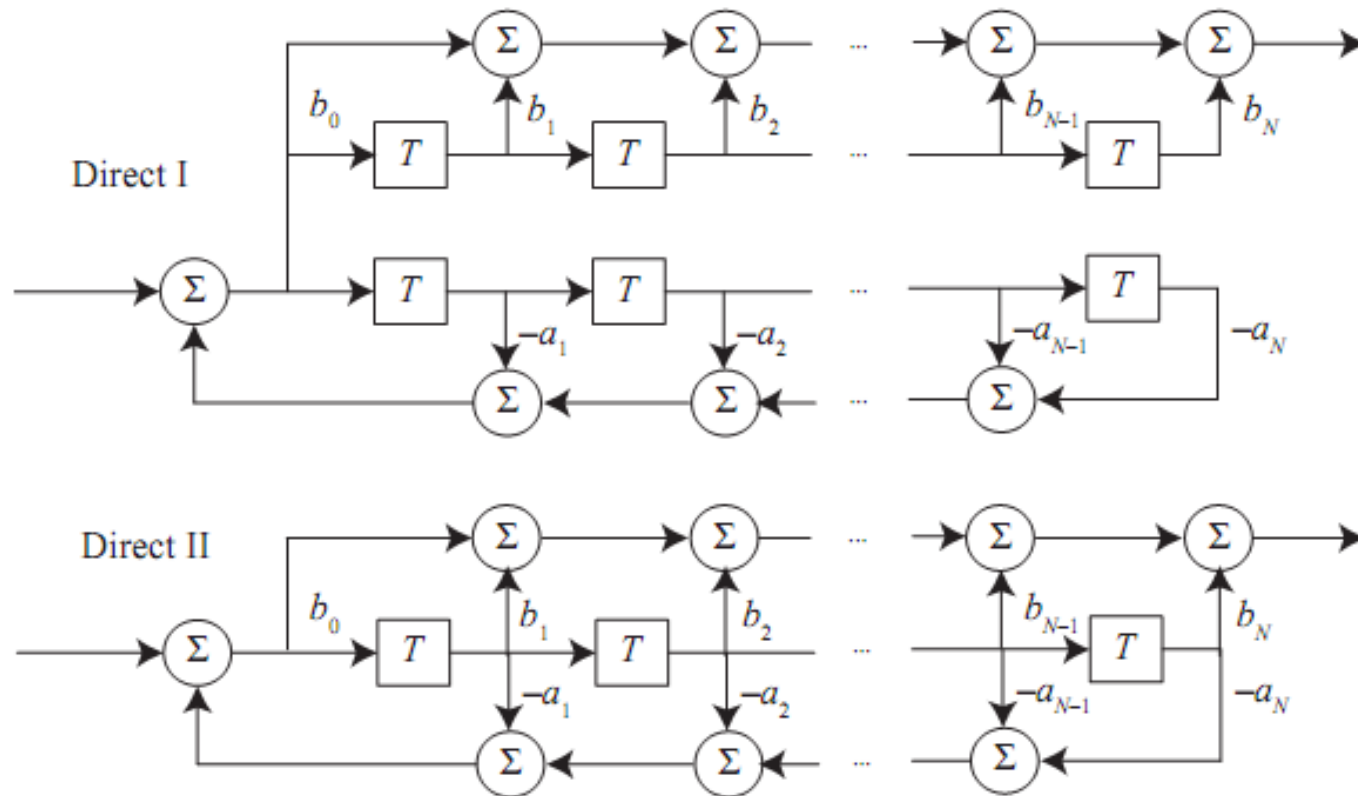
Symetrická FIR architektura

- používá se k realizaci FIR filtrů s lineární fází u kterých jsou symetrické koeficienty impulzní odezvy \rightarrow je potřeba poloviční počet násobiček \rightarrow rychlejší zpracování .



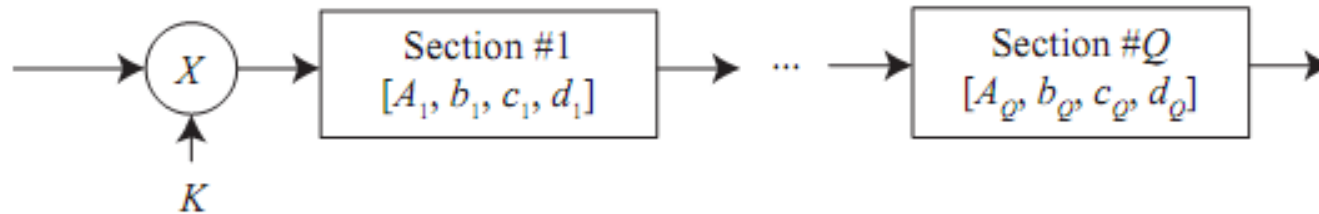
- **struktura není vhodná pro zpracování na moderních DSP procesorech**
- ty jsou optimalizovány pro provádění MAC operací (multiply-accumulate) a ne operací typu add-multiply-add , které jsou použity v symetrické struktuře,

- IIR filtry
- Přímá forma I, přímá forma 2



- Přímá forma 2 je realizačně jednodušší, pro filtr řádu N je potřeba pouze N posuvných registrů (pro přímou formu I je to $2 \cdot N$)

- kaskádní architektura



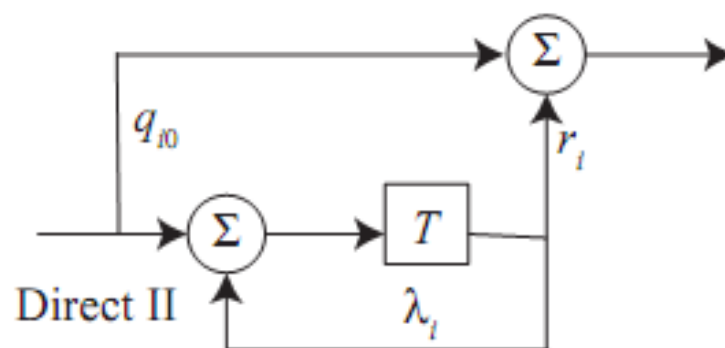
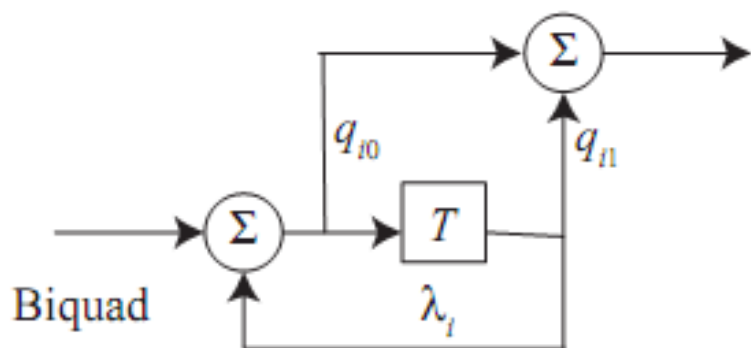
$$H(z) = \frac{N(z)}{D(z)} = G \prod_{i=1}^Q H_i(z); \sum_{i=1}^Q \text{order}(H_i(z)) = N.$$

- Kaskádní architektura implementuje výsledný filtr nejčastěji jako kaskádu filtrů 2. řádu (popř. filtru 1. řádu) popsaného stavovým modelem $[A_i, b_i, c_i, d_i]$

- Filtry 1. řádu

$$H_i(z) = \frac{q_{i0} + q_{i1}z^{-1}}{1 - p_i z^{-1}} \text{ biquadartic (biquad),}$$

$$H_i(z) = q_{i0} + \frac{r_i z^{-1}}{1 - p_i z^{-1}}; r_i = q_{i1} + p_i q_{i0} \text{ direct II.}$$

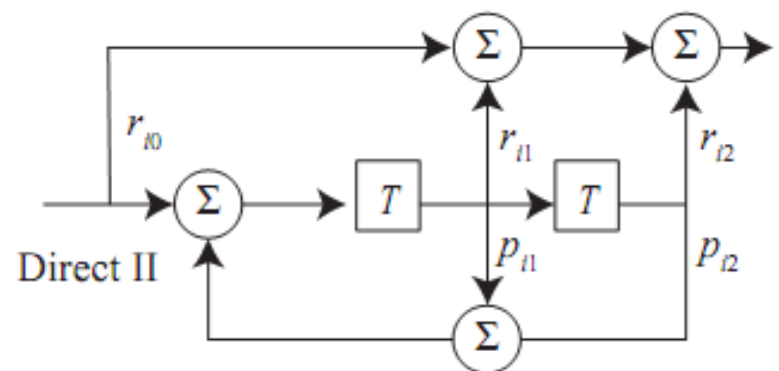
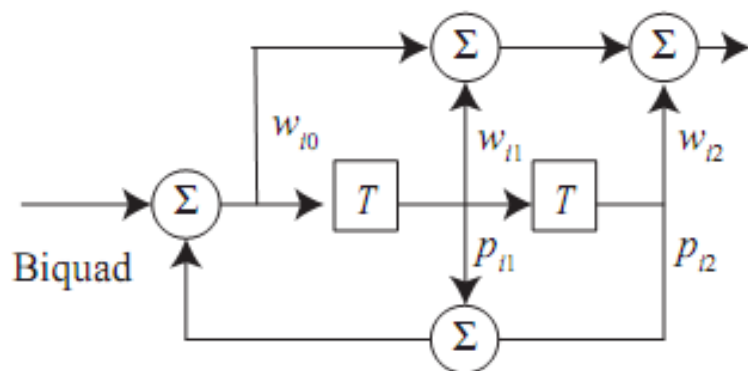


- Filtry 2. Řádu

$$H_i(z) = \frac{w_{i0} + w_{i1}z^{-1} + w_{i2}z^{-2}}{1 - v_{i1}z^{-1} - v_{i2}z^{-2}}; (z - p_i)(z - p_i^*) = z^2 - v_{i1}z - v_{i2}; \text{biquadartie (biquad)}$$

$$H_i(z) = d_{i0} + \frac{r_{i1}z^{-1} + r_{i2}z^{-2}}{1 - v_{i1}z^{-1} - v_{i2}z^{-2}}; (z - p_i)(z - p_i^*) = z^2 - v_{i1}z - v_{i2}; \text{direct II.}$$

$d_{i0} = w_{i0}; r_{i1} = w_{i1} + v_{i1}w_{i0}; r_{i2} = w_{i2} + v_{i2}w_{i0}$



$$H_i(z) = \frac{B_i(z)}{A_i(z)} = \frac{b_{0i} + b_{1i}z^{-1} + b_{2i}z^{-n}}{1 + a_{1i}z^{-1} + \dots + a_{2i}z^{-m}}; i \in [1, L],$$

$$\text{SOS} = \begin{bmatrix} b_{01} & b_{11} & b_{21} & 1 & a_{11} & a_{21} \\ b_{02} & b_{12} & b_{22} & 1 & a_{12} & a_{22} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ b_{0L} & b_{1L} & b_{2L} & 1 & a_{1L} & a_{2L} \end{bmatrix}.$$

Výpočet sos (second-order- section) matice:

1. Výpočet nul a pólů
2. Seskupení nul a pólů do komplexně sdružených dvojic na základě blízkosti `[SOS, G]=TF2SOS(N, D)`
3. Implementace subfiltrů 2. řádu.

Příklad:

$$H(z) = \frac{N(z)}{D(z)} = \frac{1 - 0.5z^{-1} - 0.315z^{-2} - 0.185z^{-3}}{1 - 0.5z^{-1} + 0.5z^{-2} - 0.25z^{-3}}.$$

`N=[1 -.5 -.315 -.0185]; D=[1 -.5 .5 -.25];`

`[SOS,G]=TF2SOS(N,D)`

`SOS = 1.0000 -0.8813 0.0000 1.0000 -0.5000 0.0000`

`1.0000 0.3813 0.0210 1.0000 0.0000 0.5000`

`G = 1`

$$H(z) = GH_1(z)H_2(z) = \frac{1 - 0.8813z^{-1}}{1 - 0.5z^{-1}} \frac{1 + 0.3813z^{-1} + 0.021z^{-2}}{1 + 0.5z^{-2}}$$

