

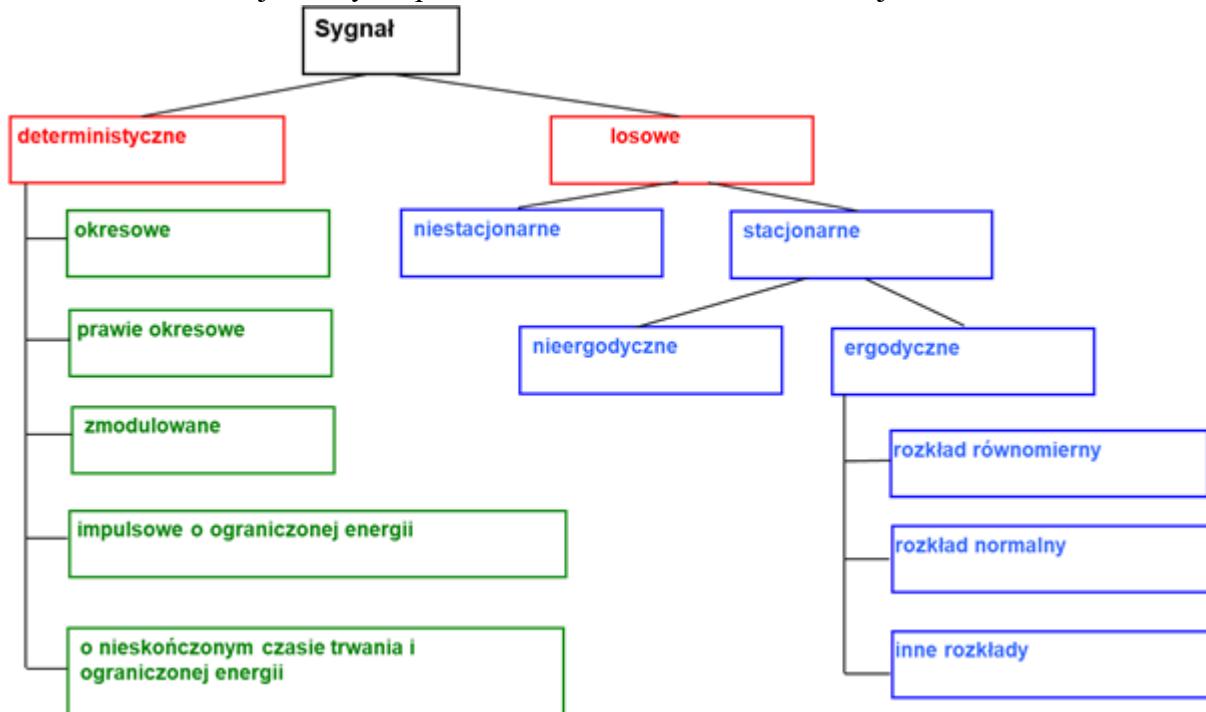
Cyfrowe Przetwarzanie Sygnałów – konspekt z wykładów

Sygnały i ich parametry

1. Sygnały i ich parametry

Pojęcia podstawowe.

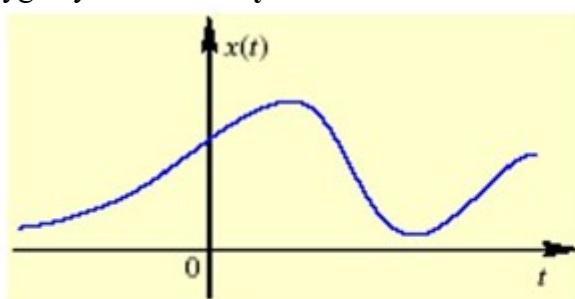
pojęcie **sygnału** jest rozumiane jako proces zmian w czasie pewnej wielkości fizycznej lub stanu obiektu fizycznego w funkcji jakiegoś argumentu (jednego lub wielu). Stosowane są też funkcje wektorowe. Jest nośnikiem informacji. Służy do przenoszenia lub zbierania informacji.



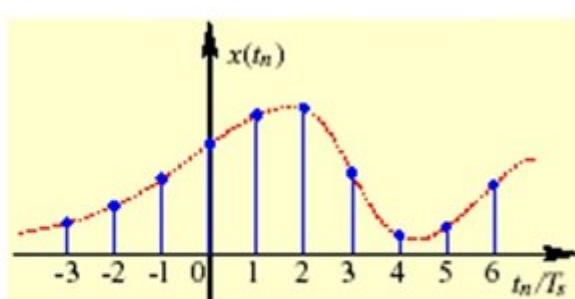
W sygnałach losowych rozróżniamy także sygnały quasi-deterministyczne, dla których losowość jest związana z losowym charakterem tylko niektórych zmiennych.

Ogólnie wszystkie sygnały możemy podzielić na :

- funkcje różnych argumentów, np. **czasu lub położenia** (odległości),
- funkcje różnej liczby argumentów (wieloargumentowe), np. **jedno-, dwu lub wielowymiarowe**,
- funkcje przyjmujące różne wartości, np. **rzeczywiste lub zespolone**,
- sygnały **ciągłe (analogowe)** $x(t)$ oraz **dyskretnie** $x_k(t)$, $x(n)$ i $x_k(n)$,
- sygnały **deterministyczne i losowe**.



Sygnał ciągły w czasie



Sygnał dyskretny w czasie

Sygnały o skończonym czasie trwania = sygnały impulsowe

Ze względu na charakter dziedziny i przeciwdziedziny sygnały dzielimy zatem na:

- ciągłe w czasie i ciągłe w amplitudzie (nazywane także analogowymi),
- ciągłe w czasie i dyskretnie w amplitudzie,
- dyskretnie w czasie i ciągłe w amplitudzie,
- dyskretnie w czasie i dyskretnie w amplitudzie.

Parametry (raczej nie zapamiętujcie wzorów, nie ma po co)

Wartość średnia analogowego impulsowego sygnału deterministycznego $x(t)$ określonego w przedziale $[t_1, t_2]$ jest całka z tego sygnału w przedziale $[t_1, t_2]$ odniesiona do szerokości tego przedziału. Gdy sygnał jest nieskończony średnia jest określana jako wielkość graniczna. Chyba że sygnał jest okresowy, to wtedy liczy się jeden okres (jest równoznaczne z uśrednianiem w nieskończoności).

Wartość skuteczna to pierwiastek mocy.

Energia i moc definiują czy sygnał deterministyczny ma skończoną moc czy skończoną energię (zeruje się to drugie)

Parametr	Sygnal ciągły $x(t)$	Sygnal dyskretny $x(n)$
Wartość średnia sygnału w przedziale	$\bar{x} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x(t) dt$	$\bar{x} = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} x(n)$
Wartość średnia całego sygnału	$\bar{x} = \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \int_{-\tau}^{\tau} x(t) dt$	$\bar{x} = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N x(n)$
Wartość średnia sygnału okresowego	$\bar{x}_T = \frac{1}{T} \int_{t_0}^{t_0+T} x(t) dt, T - okres$	$\bar{x}_N = \frac{1}{N} \sum_{n=n_0}^{n_0+(N-1)} x(n), N - okres$
Średnia bieżąca	$\bar{x}_T = \frac{1}{2T} \int_{t_0-T}^{t_0+T} x(\tau) d\tau, T - stalt$	$\bar{x}_N = \frac{1}{2N+1} \sum_{k=n-N}^{n+N} x(k), N - stalt$
Energia sygnału	$E_x = \int_{-\infty}^{+\infty} x^2(t) dt$	$E_x = \sum_{-\infty}^{+\infty} x^2(n)$
Moc średnia sygnału w przedziale	$P_x(t_1, t_2) = \bar{x^2} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x^2(t) dt$	$P_x(n_1, n_2) = \bar{x^2} = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} x^2(n)$
Moc średnia całego sygnału (wartość średniokwadratowa)	$P_x = \bar{x^2} = \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \int_{-\tau}^{\tau} x^2(t) dt$	$P_x = \bar{x^2} = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N x^2(n)$
Moc średnia sygnału okresowego	$P_x = \bar{x^2} = \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \int_{-\tau}^{\tau} x^2(t) dt$	$P_x^N = \bar{x^2} = \frac{1}{N} \sum_{n=n_0}^{n_0+(N-1)} x^2(n), N - okres$
Wartość skuteczna	$\sqrt{P_x}$	$\sqrt{P_x}$

Wariancja całego sygnału (wokół \bar{x})	$\sigma_x^2 = \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \int_{-\tau}^{\tau} [x(t) - \bar{x}]^2 dt$	$\sigma_x^2 = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N [x(n) - \bar{x}]^2$
Wariancja sygnału w przedziale (wokół \bar{x})	$\sigma_x^2 = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} [x(t) - \bar{x}]^2 dt$	$\sigma_x^2 = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} [x(n) - \bar{x}]^2$
Momenty zwykłe rzędu m, m= 1, 2, 3,...	$\bar{\tau}_x^m = \int_{-\infty}^{+\infty} t^m x(t) dt$	$\bar{k}_x^m = \sum_{n=-\infty}^{\infty} n^m x(n)$
Momenty centralne rzędu m, m= 1, 2, 3,...	$\sigma_x^m = \int_{-\infty}^{+\infty} (t - \bar{\tau}_x^1)^m x(t) dt$	$\sigma_x^m = \sum_{n=-\infty}^{\infty} (n - \bar{k}_x^1)^m x(n)$
Unormowane momenty zwykłe rzędu m	$\bar{\tau}_x^m = \int_{-\infty}^{+\infty} t^m x(t) dt / \int_{-\infty}^{+\infty} x(t) dt$	$\bar{n}_x^m = \sum_{n=-\infty}^{\infty} n^m x(n) / \sum_{n=-\infty}^{\infty} x(n)$
Unormowane momenty centralne rzędu m	$\sigma_x^m = \int_{-\infty}^{+\infty} (t - \bar{\tau}_x^1)^m x(t) dt / \int_{-\infty}^{+\infty} x(t) dt$	$\sigma_x^m = \sum_{n=-\infty}^{\infty} (n - \bar{k}_x^1)^m x(n) / \sum_{n=-\infty}^{\infty} x(n)$
Odcięta średnia ciężkości kwadratu sygnału	$\bar{\tau}_{x^2}^m = \int_{-\infty}^{+\infty} t x^2(t) dt / \int_{-\infty}^{+\infty} x^2(t) dt$	$\bar{n}_{x^2}^1 = \sum_{n=-\infty}^{\infty} n x^2(n) / \sum_{n=-\infty}^{\infty} x^2(n)$
Wariancja kwadratu sygnału wokół	$\sigma_{x^2}^m = \int_{-\infty}^{+\infty} (t - \bar{\tau}_{x^2}^1)^2 x^2(t) dt / \int_{-\infty}^{+\infty} x^2(t) dt$	$\sigma_{x^2}^2 = \sum_{n=-\infty}^{\infty} (n - \bar{n}_{x^2}^1)^2 x^2(n) / \sum_{n=-\infty}^{\infty} x^2(n)$
Szerokość średniokwadratowa	$\sqrt{\sigma_{x^2}^2}$	$\sqrt{\sigma_{x^2}^2}$

Sygnały **ciągłe czasu ciągłego** $x(t)$ są opisane ciągłymi funkcjami czasu, przyjmującymi wartości ze zbioru liczb rzeczywistych.

Sygnały **dyskretne czasu ciągłego** $xk(t)$ są sygnałami ciągłymi w czasie, przyjmującymi wartości dyskretne. (sygnał muzyki z płyty CD)

Sygnały **ciągłe czasu dyskretnego** $x(n)$ powstają w wyniku *dyskretyzacji (spróbkowania)* w czasie sygnałów ciągłych

Sygnały **cyfrowe** $xk(n)$, czyli **sygnały dyskretne czasu dyskretnego**, to sygnały ciągłe czasu dyskretnego $x(n)$, w których dodatkowo dokonano *kwantowania* wartości sygnału

W przypadku sygnałów **deterministycznych** dany jest „przepis” matematyczny na wartość, jaką przyjmie sygnał w każdej chwili czasu

Sygnały stacjonarne dzielą się dodatkowo na sygnały **ergodyczne i nieergodyczne**. Sygnały stacjonarne mają dla każdej chwili czasowej takie same wartości podstawowych parametrów statystycznych (typu średnia, wariancja) w zbiorze ich wielu realizacji. Dla stacjonarnych sygnałów ergodycznych podstawowe parametry statystyczne typu średnia/wariancja dla jednej realizacji są takie same jak po zbiorze wielu realizacji

Sygnały zespolone

Sygnały deterministyczne mogą przyjmować wartości zespolone i mają postać:

$$z(t) = x_r(t) + jx_u(t)$$

gdzie: $j = \sqrt{-1}$

Podobnie jak każda liczbę zespoloną, sygnał zespolony można także zapisać w postaci biegunowej

Najbardziej znanymi zdeterminowanymi sygnałami są:

- **zespolony przebieg wykładniczy:**

$$z(t) = e^{j\omega t} = \cos(\omega t) + j \sin(\omega t)$$

- tzw. **sygnał analityczny:**

$$z(t) = x_r(t) + jx_u(t), \quad x_u(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} x_r(\tau) \frac{1}{t-\tau} d\tau$$

gdzie $x_u(t)$ jest transformatą Hilberta sygnału $x_r(t)$.

Również sygnały zespolone dzielimy na sygnały o ograniczonej energii i ograniczonej mocy

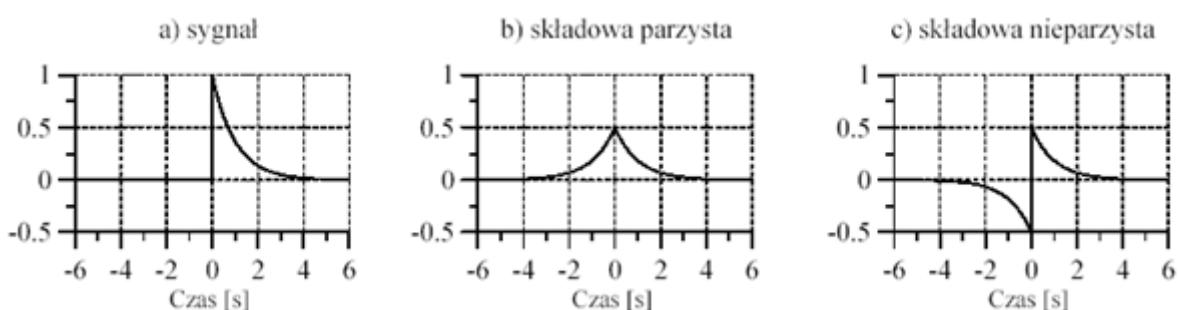
Sygnały deterministyczne można rozłożyć na składowe:

- **rzeczywistą i urojoną:**

$$x_r(t) = \operatorname{Re}(z(t)) = \frac{1}{2} [z(t) + z^*(t)], \quad x_u(t) = \operatorname{Im}(z(t)) = \frac{1}{2} [z(t) - z^*(t)]$$

- **stałą (wartość średnia \bar{x}) i zmienną ($\tilde{x}(t) = x(t) - \bar{x}$),**
- **parzystą i nieparzystą:**

$$x_p(t) = \frac{1}{2} [x(t) + x(-t)], \quad x_n(t) = \frac{1}{2} [x(t) - x(-t)]$$



Sygnalem analitycznym, reprezentującym rzeczywisty sygnał $x(t)$, nazywamy sygnał zespolony którego częścią rzeczywistą jest sygnał $x(t)$, a częścią urojoną – transformata Hilberta $\hat{x}(t)$

Sygnał analityczny stanowi szczególnego rodzaju zespolona reprezentacja sygnału rzeczywistego, często stosowana w zagadnieniach modulacji sygnałów.

Sygnały dyskretne

Sygnały dyskretne występujące w technice otrzymujemy z reguły w wyniku próbkowania sygnałów analogowych.

Splot sygnałów

Splot funkcji ciągłych:

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)d\tau = \int_{-\infty}^{+\infty} h(\tau)x(t-\tau)d\tau$$

Splot opisuje operację filtracji jednego sygnału przez drugi.

Kolejność operacji podczas splotu jest następująca:

- 1) odwróć w czasie drugi z sygnałów
- 2) przesuń go w czasie o czas
- 3) wymnóż pierwszy sygnał ze zmodyfikowanym drugim
- 4) scałkuj wynik mnożenia.

Splot dwóch sygnałów:

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k) = \sum_{k=-\infty}^{\infty} h(k)x(n-k)$$

Obliczanie splotu dyskretnego (filtracji cyfrowej) przebiega w ten sam sposób jak splotu ciągłego.

Równania splotu mają fundamentalne znaczenie w teorii przetwarzania układów analogowych i cyfrowych, ponieważ opisują „mechanizm” *filtracji* sygnałów. Sygnał „filtrujący” jest zazwyczaj *odpowiedzią impulsową układu (filtra)* analogowego lub cyfrowego przez który przechodzi sygnał

Podstawy matematyczne analizy sygnałów deterministycznych

Tego rozdziału nie ma w zagadnieniach od WC, ale warto wiedzieć co nieco z niego.

Sygnały, podobnie jak funkcje, tworzą przestrzenie, czyli zbiory o określonych właściwościach.

Przestrzenią metryczną nazywamy zbiór sygnałów, w którym jest zdefiniowana metryka, czyli funkcjonał przyporządkujący dowolnym dwóm sygnałom x i y nieujemną liczbę rzeczywistą $\rho(x,y) \geq 0$. Liczba ta może być interpretowana jako odległość między tymi sygnałami. Metryka powinna spełniać warunki: tożsamości, symetrii i nierówności trójkąta. Ciąg elementów przestrzeni metrycznej nazywamy ciągiem Cauchy'ego, jeśli odległość pomiędzy wyrazami ciągu maleje do zera wraz ze wzrostem numerów jego elementów. Przestrzeń metryczną nazywamy zupełną jeśli jej każdy ciąg Cauchy'ego jest zbieżny do jakiegoś elementu tej przestrzeni (ma granicę należącą do tej przestrzeni).

Przestrzenią liniową nazywamy zbiór sygnałów z operacjami dodawania sygnałów x oraz mnożenia sygnałów przez liczbę a . W przestrzeni liniowej musi istnieć sygnał zerowy 0 , taki że $0*x=0$ oraz skalar identycznościowy 1 , taki że $1*x=x$.

Przestrzenią liniową unormowaną nazywamy przestrzenie liniową, w której zdefiniowane jest odwzorowanie, przyporządkowujące każdemu sygnałowi x liczبę rzeczywistą nieujemną (normę tego sygnału).

Przestrzenią Banacha nazywamy przestrzeń liniową unormowaną, która równocześnie jest przestrzenią metryczną zupełną, która jest równocześnie przestrzenią metryczną zupełną.

Przestrzenią unitarną nazywamy przestrzeń liniową, w której określony jest iloczyn skalarny $\{x,y\}$ dwóch sygnałów x i y , i która jest unormowana przez normę związaną z tym iloczynem. Dwa sygnały x i y w przestrzeni unitarnej są prostopadłe jeśli ich iloczyn skalarny jest równy zero. Jeśli dodatkowo $\{x,x\} = 1$ i $\{y,y\} = 1$ to są one ortonormalne.

Przestrzenią Hilberta nazywamy przestrzeń unitarną, która także jest przestrzenią metryczną zupełną.

Ciągle reprezentacje ciągłych sygnałów deterministycznych – przekształcenia całkowe

- przekształcenie Laplace'a
- przekształcenie Fouriera (czyli przekształcenie Laplace'a dla $s=j2\pi f=j\omega$)
- przekształcenie Hilberta (sygnał analityczny)
- przekształcenie Hankela, kosinusowe i sinusowe przekształcenie Fouriera
- przekształcenie Mellina

Szereg Fouriera i całkowe przekształcenie Fouriera

Szereg Fouriera - Każdy przebieg okresowy można przedstawić jako sumę nieskończonego szeregu składowych sinusoidalnych i kosinusoidalnych o odpowiednich amplitudach i częstotliwościach.

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(k\omega_0 t) + b_k \sin(k\omega_0 t))$$

$$a_0 = \frac{2}{T} \int_0^T x(t) dt$$

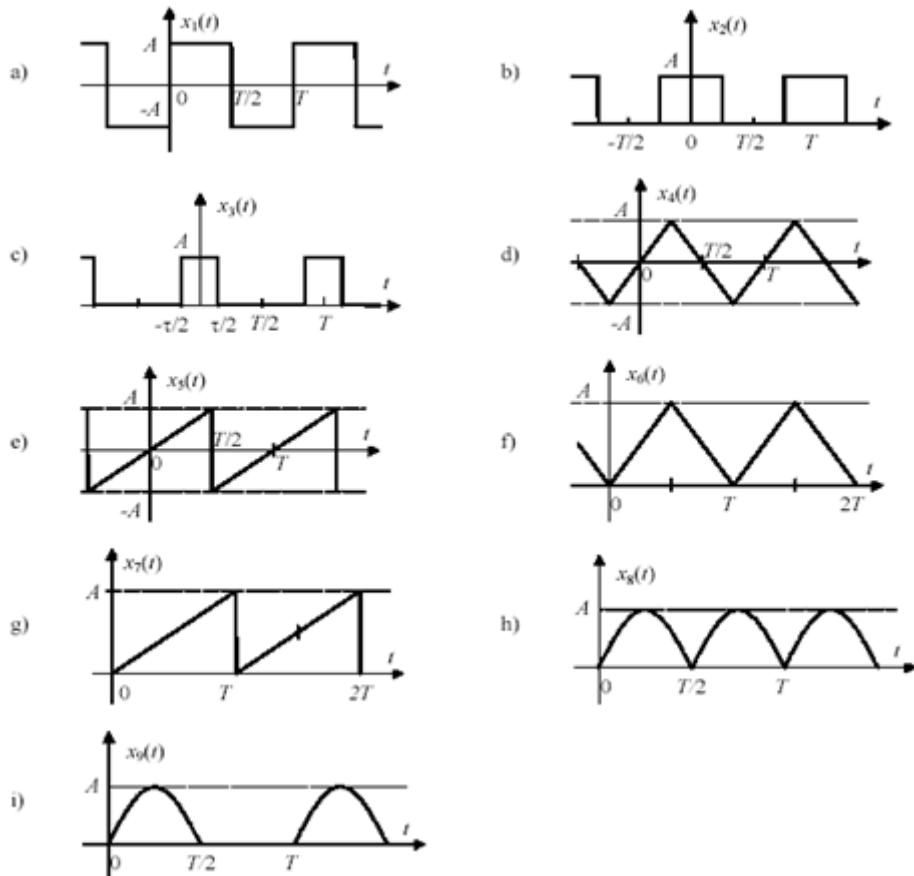
$$a_k = \frac{2}{T} \int_0^T x(t) \cos(k\omega_0 t) dt$$

$$b_k = \frac{2}{T} \int_0^T x(t) \sin(k\omega_0 t) dt$$

Zespolony szereg Fouriera

$$x(t) = \sum_{k=-\infty}^{\infty} \alpha_k \frac{1}{\sqrt{T_0}} e^{jk\frac{2\pi}{T_0}t}$$

Sygnal	Współczynniki szeregu Fouriera
Prostokątny, bipolarny rys. 3.1a	$x(t) = \frac{4A}{\pi} \left(\sin \omega_0 t + \frac{1}{3} \sin 3\omega_0 t + \frac{1}{5} \sin 5\omega_0 t + \dots \right)$
Prostokątny, unipolarny wypełnienie 1/2 rys. 3.1b	$x(t) = \frac{A}{2} + \frac{2A}{\pi} \left(\cos \omega_0 t - \frac{1}{3} \cos 3\omega_0 t + \frac{1}{5} \cos 5\omega_0 t - \dots \right)$
Prostokątny, unipolarny, wypełnienie dowolne rys. 3.1c	$x(t) = \frac{A\tau}{T} + \frac{2A\tau}{T} \sum_{k=1}^{\infty} \frac{\sin(\pi k\tau/T)}{\pi k\tau/T} \cos k\omega_0 t$
Trójkątny, bipolarny 1 rys. 3.1d	$x(t) = \frac{8A}{\pi^2} \left(\sin \omega_0 t - \frac{1}{3^2} \sin 3\omega_0 t + \frac{1}{5^2} \sin 5\omega_0 t - \dots \right)$
Trójkątny, bipolarny 2 rys. 3.1e	$x(t) = \frac{2A}{\pi} \left(\sin \omega_0 t - \frac{1}{2} \sin 2\omega_0 t + \frac{1}{3} \sin 3\omega_0 t - \dots \right)$
Trójkątny, unipolarny 1 rys. 3.1f	$x(t) = \frac{A}{2} - \frac{4A}{\pi^2} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} \cos(2k+1)\omega_0 t$
Trójkątny, unipolarny 2 rys. 3.1g	$x(t) = \frac{A}{2} - \frac{A}{\pi} \sum_{k=1}^{\infty} \frac{\sin k\omega_0 t}{k}$
Sinusoidalny wyprostowany dwupolówkowo rys. 3.1h	$x(t) = \frac{2A}{\pi} - \frac{4A}{\pi} \sum_{k=1}^{\infty} \frac{1}{4k^2 - 1} \cos 2k\omega_0 t$
Sinusoidalny wyprostowany jednopolówkowo rys. 3.1i	$x(t) = \frac{A}{\pi} + \frac{A}{2} \sin \omega_0 t - \frac{2A}{\pi} \sum_{k=1}^{\infty} \frac{1}{4k^2 - 1} \cos 2k\omega_0 t$



Rys. 3.1. Kształt sygnałów, dla których podano współczynniki szeregu Fouriera w tabeli 3-1. a) – bipolarna fala prostokątna, b), c) – unipolarnie fale prostokątne, d), e), f), g) – różne bipolarne i unipolarnie fale trójkątne, h), i) – sygnał sinusoidalny wyprostowany dwupolówkowo i jednopolówkowo

Całkowe przekształcenie Fouriera jest określone parą następujących transformacji:

$$X(j\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt \quad x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\omega)e^{j\omega t} d\omega$$

wiązających sygnał $x(t)$ z jego widmem Fouriera $X(j\omega)$. Pierwsze równanie jest równaniem „analizy” (wyznaczenie współczynników Fouriera), a drugie - równaniem „syntezy” (zsyntezowanie sygnału z jego współczynników Fouriera). Aby sygnał miał transformatę Fouriera $X(j\omega)$ (tzn. aby całka była określona w sensie Cauchy'ego), musi on spełniać tzw. warunki Dirichleta:

a) funkcja $x(t)$ jest bezwzględnie całkowalna

$$\int_{-\infty}^{\infty} |x(t)| dt < \infty$$

b) mieć skończone wartości maksimów i minimów w każdym skończonym przedziale,

c) mieć skońzoną liczbę punktów nieciągłości w każdym skończonym przedziale.

Transformata Fouriera $X(j\omega)$ sygnału $x(t)$ jest nazywana widmem („częstotliwościowym”) tego sygnału, gdyż informuje nas o jego „zawartości” widmowej („częstotliwościowej”). Ponieważ $X(j\omega)$ jest w ogólności liczbą zespoloną, można ją zapisać w układzie biegunkowym za pomocą modułu i fazy (długości promienia i kąta):

$$X(j\omega) = |X(j\omega)| e^{j\angle X(j\omega)}$$

i przedstawić równanie odwrotnego przekształcenia Fouriera jako:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(j\omega)| e^{j(\omega t + \angle X(j\omega))} d\omega$$

Wynika z niego, że podczas operacji syntezy wynikowy sygnał $x(t)$ powstaje w wyniku sumowania nieskończonej liczby sygnałów harmonicznych $e^{j\omega t}$, z których każdy ma swoją amplitudę $|X(j\omega)|$ i przesunięcie fazowe $X(j\omega)$.

Transformacja Fouriera

Prosta i odwrotna transformacja Fouriera są zdefiniowane następującą parą równań:

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt \quad x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} df$$

Sygnały okresowe o okresie T przedstawia się za pomocą szeregu Fouriera:

$$X(kf_0) = \frac{1}{T} \int_0^T x(t) e^{-j2\pi(kf_0)t} dt \quad x(t) = \sum_{k=-\infty}^{\infty} X(kf_0) e^{j2\pi(kf_0)t}$$

Natomiast wzór na szereg Fouriera dla sygnałów dyskretnych $x(n)$ mają następującą postać

$$X(kf_0) = \frac{1}{N} \sum_{n=0}^N x(n\Delta t) e^{-j2\pi(k/N)n} \quad x(n\Delta t) = \sum_{k=0}^{N-1} X(kf_0) e^{j2\pi(k/N)n}$$

gdzie: k jest indeksem częstotliwości, n – indeksem czasu

Funkcja korelacji własnej i wzajemnej

Funkcja korelacji wzajemnej pomiędzy deterministycznymi sygnałami $x(t)$ i $y(t)$ zdeterminowana jest w tej klasie jako:

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)y^*(t-\tau)dt$$

Funkcja zaś korelacji własnej sygnału $x(t)$ dana jest wzorem:

$$R_{xx}(\tau) = R_x(\tau) = \int_{-\infty}^{\infty} x(t)x^*(t-\tau)dt$$

Jak się później przekonamy $R_{xx}(t)$ to iloczyn skalarny dwóch sygnałów w funkcji przesunięcia jednego z nich. W funkcji korelacji drugi sygnał opóźnia się w stosunku do pierwszego o czas t , następnie oba sygnały wymnaża się przez siebie i całkuje iloczyn. W ten sposób dla każdego t otrzymuje się liczbę, mówiącą na ile opóźniony drugi sygnał jest podobny do sygnału pierwszego (skorelowany z nim).

Funkcja korelacji własnej $R_{xx}(\tau)$ jest wykorzystywana do badania „powtarzalności” (okresowości) sygnału, ponieważ przyjmuje ona wartości maksymalne dla wartości przesunięcia τ równego wielokrotności okresu sygnału (wówczas bowiem mnożone są przez siebie wartości dodatnie sygnałów $x(t)$ i $x^*(t-\tau)$ oraz wartości ujemne tych sygnałów, dlatego iloczyn $x(t)x^*(t-\tau)$ ma zawsze wartość dodatnią, a całka z niego wartość maksymalną; w innym przypadku dla jakiegoś t iloczyn $x(t)x^*(t-\tau)$ może być ujemny i wartość całki z tego iloczynu po czasie jest mniejsza). Przykładowo funkcja $R_{xx}(\tau)$ jest stosowana do wyznaczania okresu głosek dźwięcznych sygnału mowy.

Funkcja korelacji wzajemnej $R_{xy}(\tau)$ może być stosowana do detekcji odbić w sygnale odebranym w echografii impulsowej. Wówczas impulsowy sygnał wysłany jest korelowany (przesuwany w czasie i wymnażany) z sygnałem odebranym, w którym występują „kopie” sygnału oryginalnego (nadane), odbite od różnych obiektów. Maksima funkcji korelacji informują nas o obecności i położeniu impulsów odbitych, czas opóźnienia zaś tych impulsów w stosunku do impulsu wysłanego o drodze przebytej przez sygnał, czyli o odległość obiektu od nadajnika.

Funkcje korelacji $R(\cdot)$ i kowariancji $C(\cdot)$ definiują się dla stacjonarnych sygnałów losowych $x(t)$ i $y(t)$ w następujący sposób

Sygnal ciągły

$$R_{xx}(\tau) = E[x(t)x(t-\tau)]$$

Sygnal dyskretny

$$R_{xx}(m) = E[x(n)x(n-m)]$$

Gdzie $E[\cdot]$ oznacza wartość oczekiwana po zbiorze realizacji (sygnałów) procesu ustalonego t . Definicja funkcji kowariancji jest analogiczna do funkcji korelacji, z tą różnicą, że odejmuje się w niej od zmiennych losowych, związanych z konkretną chwilą czasową, ich wartości oczekiwane („średnie”, najbardziej prawdopodobne) Dla sygnałów stacjonarnych powyższych wartości oczekiwane nie zależą od czasu t (indeksu n).

Funkcja gęstości widmowej mocy

Do analizy częstotliwościowej sygnałów stosuje się funkcję gęstości widmowej mocy. Jest ona zdefiniowana jako transformacja Fouriera funkcji autokorelacji:

$$P_{xx}(f) = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-j2\pi f\tau} d\tau$$

$$P_{xx}(f) = \sum_{m=-\infty}^{\infty} R_{xx}(m) e^{-j2\pi(f/f_{pr})m}$$

Para równań ww. „odwrotności” (R zamienione miejscami z P) nosi nazwę równań Wienera-Chinczyna dla sygnałów ciągłych i dyskretnych.

Sygnały losowe

Dla sygnałów losowych (stochastycznych) $x(t)$ nie istnieje „przepis” matematyczny na wartości przyjmowane w kolejnych chwilach czasowych, tylko „przepis” na prawdopodobieństwo przyjęcia przez sygnał określonej wartości.

W teorii sygnałów losowych stosuje się pojęcia procesu (sygnału) losowego, jego konkretnych realizacji czasowych oraz związanych z nimi zmiennych losowych.

Po swoim wystąpieniu konkretne realizacje czasowe procesu losowego są już sygnałami zdeterminowanymi. Zmienna losowa x to zmienna przyjmująca wartości w sposób losowy (stochastyczny), ale zgodnie ze z góry określonym prawdopodobieństwem.

Dystrybuantą $F(x_0)$ zmiennej losowej x jest nazywane prawdopodobieństwo przyjęcia przez nią wartości x mniejszej lub równej od x_0 , tzn.

$$F(x_0) = \Pr[x \leq x_0]$$

Kiedy $x_0 \rightarrow -\infty$, wówczas $F(x_0) \rightarrow 0$, a dla $x_0 \rightarrow \infty$ mamy $F(x_0) \rightarrow 1$. Jeżeli x_0 przybiera wartości ciągłe, to równoważnym do dystrybuanty opisem zmiennej losowej x jest jej funkcja gęstości prawdopodobieństwa, definiowana jako pochodna dystrybuanty:

$$p(x_0) = \frac{\partial F(x_0)}{\partial x_0}$$

Określa ona, jakie jest prawdopodobieństwo przyjęcia przez zmienną losową x wartości x_0 . Ponieważ zmienna losowa musi przyjąć jakąś wartość, prawdopodobieństwo przyjęcia dowolnej wartości jest równe jedności:

$$\int_{-\infty}^{\infty} p(x_0) dx_0 = 1$$

Z definicji funkcji gęstości prawdopodobieństwa wynika, że:

$$F(x_0) = \int_{-\infty}^{x_0} p(\zeta) d\zeta$$

Dla zmiennych losowych przyjmujących wartości dyskretne definiuje się masę prawdopodobieństwa:

$$p(x_k) = \Pr[x = x_k], \quad \sum_k p(x_k) = 1$$

Wartość oczekiwana zmiennej losowej x , przyjmującej wartości w zbiorze ciągłym, jest określona zależnością:

$$\bar{x} = E[x] = \int_{-\infty}^{+\infty} x_0 p(x_0) dx_0$$

i jest to wartość, która jest przyjmowana przez zmienną losową z największym prawdopodobieństwem (najbardziej „prawdopodobna”). Natomiast dla sygnałów o wartościach dyskretnych mamy:

$$\bar{x} = E[x] = \sum_k x_k p(x_k)$$

Z kolei wartości średniokwadratowe i wariancja σ_x^2 zmiennej losowej x zdefiniowane są następująco:

$$E[x^2] = \int_{-\infty}^{+\infty} x_0^2 p(x_0) dx_0, \quad E[x^2] = \sum_k x_k^2 p(x_k)$$

Funkcja korelacji wzajemnej dwóch zmiennych losowych x i y jest zdefiniowana jako:

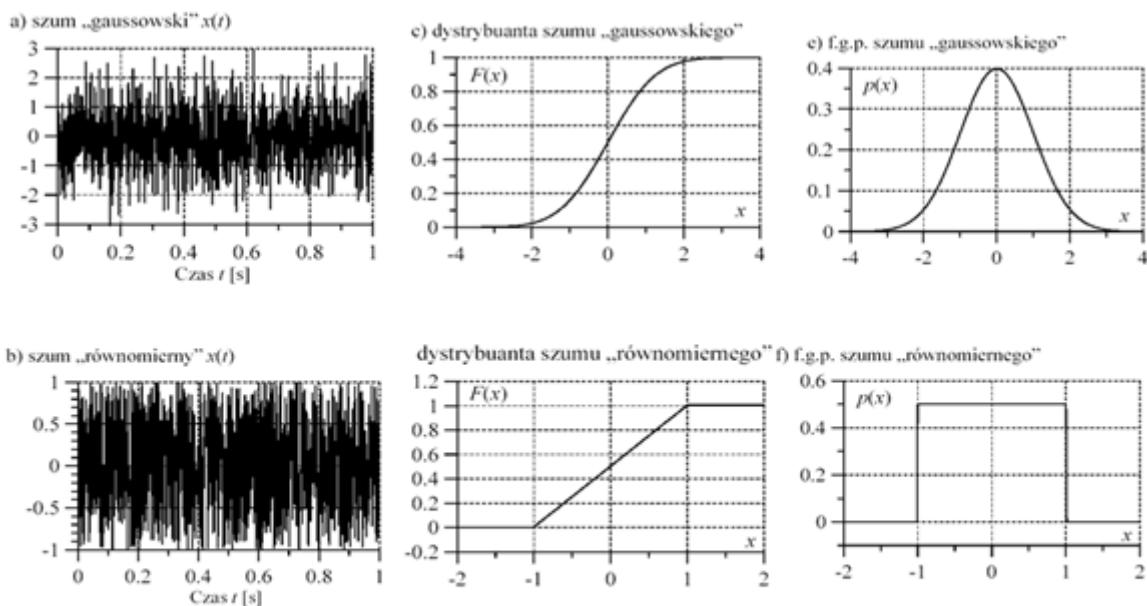
$$R_{xy} = E[xy] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy p(xy) dx dy$$

gdzie $p(xy)$ jest gęstością prawdopodobieństwa przyjęcia przez iloczyn tych zmiennych wartości xy , a $d(xy) = dx dy$ jest różniczką tego iloczynu. Dwie niezależne zmienne losowe nie są ze sobą skorelowane: $E[xy] = 0$. Podsumowując można powiedzieć, że przyjęcie wartości przez zmienną losową jest jak wynik rzutu kostką do gry: prawdopodobieństwo otrzymania poszczególnych pól jest znane i równe $1/6$, ale nie wiadomo jaki będzie wynik obecnego rzutu.

Najpopularniejsze są zmienne losowe o następujących rozkładach funkcji gęstości prawdopodobieństwa:

1) **normalnym**: $p(x)$ jest funkcją Gaussa o określonej wartości średniej x i wariancji σ^2 :

2) **równomiernym**: $p(x)$ jest funkcją „prostokątną” równą (w szczególności $[a=0, b=1]$ lub $[a = -1, b = 1]$): Dla rozkładu normalnego o wartości średniej równej zero i wariancji σ^2 równej 1, prawdopodobieństwo przyjęcia przez zmienną losową wartości z przedziału $[-1, 1]$ jest równe 0,682, wartości z przedziału $[-2, 2]$ - 0,954, wartości zaś z przedziału $[-3, 3]$ - 0,997.



Przykłady fragmentów sygnałów losowych o rozkładzie normalnym (a) i równomiernym (b) oraz ich dystrybuanty (c, d) i funkcje gęstości prawdopodobieństwa (e, f)

Sygnały losowe ciągłe i dyskretnie

Sygnały losowe związane są z tzw. procesami losowymi. Obserwowane sygnały są konkretnymi realizacjami („zapisami”, „próbami”) tych procesów. Aby określić rodzaj procesu losowego X, należy zebrać wiele jego realizacji, czyli sygnałów $x_k(t)$, $k = 1, 2, 3, \dots$. Zbiór tych realizacji określa właściwości generującego je procesu.

Aby poznać te właściwości, należy wyznaczyć i porównać podstawowe parametry statystyczne obserwowanych danych, uśrednionych po zbiorze wszystkich realizacji $\{x_k(t), k = 1, 2, \dots\}$, tzn. dla każdego t . Zakłada się, że zbiór wszystkich wartości przyjmowanych przez proces X w chwili t w poszczególnych realizacjach jest niezależną, dyskretną zmienną losową x_t , mającą określoną dystrybuantę $F_t(x)$ i funkcję gęstości prawdopodobieństwa $p_t(x)$.

Podstawowymi parametrami statystycznymi zmiennej x_t jest średnia wariancja oraz współczynnik korelacji tej zmiennej ze zmienią „oddaloną” od niej o czas τ .

Procesem stacjonarnym w szerszym sensie jest nazywany proces, dla którego wartość średnia i wariancja przyjmowanych wartości po zbiorze jego realizacji oraz wartości funkcji korelacji pomiędzy różnymi chwilami czasowymi nie zależą od t , tzn. dla wszystkich wartości t są takie same:

$$\bar{x}_t = E[x_t] = \text{const}, \quad \sigma_{x_t}^2 = E[(x_t - \bar{x}_t)^2] = \text{const}$$

Procesem stacjonarnym w węższym sensie jest nazywany proces, dla którego wszystkie jednowymiarowe momenty statystyczne wyższych rzędów zmiennych losowych x_t oraz momenty łączne tych zmiennych są takie same, tzn. nie zależą od chwili t .

Procesem ergodycznym jest taki proces stacjonarny, dla którego wartości parametrów statystycznych (wartość średnia, wariancja, funkcja korelacji) po zbiorze realizacji są równe wartościami tych parametrów wyznaczonymi z jego dowolnej realizacji czasowej $x_k(t)$ (czyli po czasie). Dla procesów ergodycznych stacjonarność w szerszym sensie wymaga niezależności od czasu t_0 bieżącej wartości średniej i bieżącej funkcji autokorelacji sygnału $x_k(t)$ (dla dostatecznie dużego T):

$$\bar{x}(t_0) = \frac{1}{T} \int_{t_0}^{t_0+T} x(t) dt$$

Zgodnie z twierdzeniem Wolda każdy dyskretny sygnał (proces) losowy $x(n)$, stacjonarny w szerokim sensie, może być przedstawiony jako suma dwóch składowych: deterministycznej $x_d(n)$ oraz czysto losowej $x_l(n)$:

$$x(n) = x_d(n) + x_l(n)$$

Składowa deterministyczna jest idealnie (bez błędu) przewidywalna na podstawie nieskończonej przeszłości (historii) sygnału, tzn.

$$x_d(n) = - \sum_{k=1}^{\infty} \alpha_k x_d(n-k), \quad \alpha_k - \text{stale predykcji.}$$

Przykładem takiego sygnału może być suma sinusoidy o losowej fazie oraz szumu.

Dla ciągłych sygnałów stacjonarnych, ergodycznych funkcja gęstości widmowej mocy jest równa:

$$P_{xx}(f) = \lim_{T \rightarrow \infty} E \left[\frac{1}{2T} \left| \int_{-T}^T x(t) e^{-j2\pi ft} dt \right|^2 \right] = \lim_{T \rightarrow \infty} E \left[\frac{1}{2T} |X_T(f)|^2 \right] = \lim_{T \rightarrow \infty} E \left[\frac{1}{2T} X_T(f) X_T^*(f) \right]$$

gdzie $E[\cdot]$ oznacza wartość oczekiwana, $X_T(f)$ jest transformatą Fouriera fragmentu sygnału z przedziału czasowego $[-T, T]$ (czyli widmem amplitudy tego fragmentu), funkcja $1/(2T)^*|X_T(f)|^2$ nazywana jest zaś periodogramem i ma taki sam wymiar jak funkcja widmowej gęstości mocy.

Dla dyskretnych sygnałów stacjonarnych mamy:

$$P_{xx}(f) = \lim_{N \rightarrow \infty} E \left[\frac{1}{2N+1} \left| \sum_{n=-N}^N x(n) e^{-j\frac{2\pi nf}{f_p}} \right|^2 \right] = \lim_{N \rightarrow \infty} E \left[\frac{1}{2N+1} |X_N(f)|^2 \right] = \lim_{N \rightarrow \infty} E \left[\frac{1}{2N+1} X_N(f) X_N^*(f) \right]$$

W tym przypadku periodogram jest zdefiniowany jako $1/(2N+1)*|X_N(f)|^2$. Zależność dla dyskretnych sygnałów stacjonarnych stanowi podstawę praktycznych metod szacowania wartości funkcji $P_{xx}(f)$, na przykład metody Welcha.

Szum nazywa się białym, jeśli jego $P_{xx}(f)$ jest stałe i nie zależy od częstotliwości. Nazwa szum wyprowadzona została przez analogię do światła: światło białe to de facto szum elektromagnetyczny mieszaniny wszystkich możliwych barw o całkowicie płaskim widmie w zakresie widzialnym. Jeśli natomiast tak nie jest, to szum jest kolorowy.

Szczególnym rodzajem szumu kolorowego jest idealny szum dolnopasmowy, dla którego funkcja $P_{xx}(f)$ ma kształt prostokątny, tzn. ma wartość stałą, różną od zera dla częstotliwości z przedziału $(-f_{max}, f_{max})$ oraz równą zero poza tym przedziałem.

Innymi przykładami szumu kolorowego jest szum różowy i niebieski oraz czerwony i fioletowy zwany też purpurowym. Dla pierwszego z nich funkcja $P_{xx}(f)$ maleje 3 decybeli na oktawę (10 decybeli na dekadę), dla drugiego zaś - rośnie o 3 decybeli na oktawę (10 decybeli na dekadę). Dla szumu czerwonego funkcja $P_{xx}(f)$ maleje o 6 decybeli na oktawę (20 decybeli na dekadę), dla fioletowego zaś - rośnie o 6 decybeli na oktawę (20 decybeli na dekadę).

Szum o rozkładzie normalnym i wariancji σ^2 ma następujące parametry:

$$R(\tau) = \begin{cases} \sigma^2, & \tau = 0 \\ 0, & \tau \neq 0 \end{cases} \quad R(m) = \begin{cases} \sigma^2, & m = 0 \\ 0, & m \neq 0 \end{cases}$$

Estymatory parametrów i funkcji

Najprostsza jest analiza losowych procesów (sygnałów) ergodycznych, ponieważ przeprowadza się ją na podstawie tylko jednej realizacji zmiennej losowej. Jednak wówczas nie jesteśmy w stanie dokładnie wyznaczyć parametrów procesu losowego, gdyż nie dysponujemy nieskończonym długim zapisem sygnału, czyli nie dysponujemy pełną informacją statystyczną o nim, i z tego powodu musimy te wartości estymować. Funkcja, według której jest przeprowadzane to „szacowanie”, nazywa się estymatorem. Estymator nazywa się nieobciążonym, jeśli jego wartość oczekiwana jest równa rzeczywistej wartości estymowanego parametru (tzn. jest najbardziej prawdopodobne, że właśnie ją otrzymamy). W przeciwnym przypadku estymator jest obciążony, tzn. występuje przesunięcie pomiędzy rzeczywistą wartością parametru a wartością oczekiwana estymatora. Z kolei wariancja estymatora określa jego „wahania” („rozrzut”) wokół wartości oczekiwanej. Estymator nazywa się zgodnym, jeśli jego wariancja dąży do zera przy wzroście liczby danych, służących do jego wyznaczenia. Dobry estymator to estymator zgodny, nieobciążony.

Sygnały ciągłe. Założmy, że dysponujemy jedną realizacją ciągłego sygnału losowego $x(t)$ o skończonej długości T_s (t_0, t_0+T_s). Prawdopodobieństwo przyjęcia przez ten „fragment” sygnału wartości z przedziału $(x, x+D_x)$ jest równe T_x/T_s , gdzie T_x jest sumą odcinków czasu D_{ti} , w których wartości chwilowe realizacji znajdują się w tym przedziale: $T_x = \sum(D_{ti})$. W granicy dla T_s dążącego do nieskończoności otrzymujemy prawdopodobieństwo przyjęcia przez całą realizację wartości z przedziału $(x, x+D_x)$.

Sygnały dyskretne. Obecnie założymy, że mamy N próbek jednej realizacji dyskretnego sygnału losowego $x(n)$, (n_0, n_0+N-1) . Prawdopodobieństwo przyjęcia przez ten fragment sygnału wartości z przedziału $(x, x+D_x)$ jest równe N_x/N , gdzie N_x jest liczbą próbek, których wartości znajdują się w tym przedziale. W granicy dla N dążącego do nieskończoności otrzymujemy prawdopodobieństwo przyjęcia przez całą realizację wartości z przedziału $(x, x+D_x)$.

Estymat „ksztaltu” funkcji gęstości prawdopodobieństwa $p(x)$ ergodycznego, dyskretnego procesu otrzymuje się obliczając tzw. histogram, czyli

- 1) biorąc odpowiednio dużą liczbę N próbek sygnału,
- 2) dzieląc obserwowany zakres zmienności zmiennej x na przedziały o szerokości D_x ,
- 3) zliczając liczbę próbek, występujących w poszczególnych przedziałach. Po unormowaniu histogramu, tzn. po jego podzieleniu przez liczbę wszystkich próbek N oraz szerokość przedziału D_x , otrzymuje się „pełną” estymat funkcji $p(x)$.

Estymata ta jest tym lepsza, im większe jest N oraz mniejsze D_x. Weryfikacja hipotez „kształtu” (rodzaju) rozkładu zmiennej losowej jest dokonywana za pomocą testów zgodności, np. testu chi kwadrat. Zgodnie z twierdzeniem granicznym większość sygnałów występujących w przyrodzie ma rozkład normalny (gaussowski), wynikający z równoczesnego występowania wielu czynników, oddziaływających w różnych kierunkach.

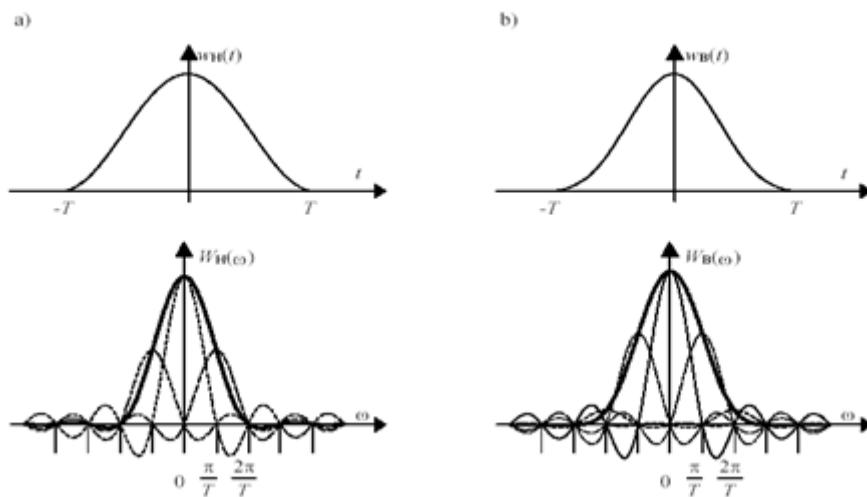
Ponieważ w praktyce nie dysponuje się nieskończoną ilością próbek sygnału, wartości funkcji estymuje się (przybliża) na podstawie dostępnych N danych za pomocą jednej z następujących zależności ($-N+1 \leq k \leq N-1$)

Estymator funkcji widmowej gęstości mocy definiuje się jako wynik transformacji Fouriera estymatora funkcji autokorelacji. Nie jest to estymator zgodny, ponieważ jego wariancja nie dąży do zera przy wzroście długości segmentu danych N. Jedną z metod zmniejszania wariancji estymatora stanowi metoda Blackmana-Tukeya.

Poddaje się w niej transformacji Fouriera nie estymator funkcji autokorelacji, tylko jego iloczyn z wybraną funkcją wagową w(n) (np. oknem Hanninga lub Hamminga),

Inną metodą jest procedura Welch'a, w której wykorzystuje się związek pomiędzy funkcją gęstości widmowej mocy a periodogramem.

Nazwa okna	$w(t)$	$W(\omega)$
Prostokątne	$p_T(t) = \begin{cases} 1 & \text{dla } t \leq T \\ 0 & \text{dla } t > T \end{cases}$	$2 \frac{\sin \omega T}{\omega}$
Trójkątne (Bartletta)	$q_T(t) = \begin{cases} 1 - t /T & \text{dla } t \leq T \\ 0 & \text{dla } t > T \end{cases}$	$T \left[\frac{\sin(\omega T/2)}{\omega T/2} \right]^2$
Hanninga (Hanna)	$[0,5 + 0,5\cos(\pi t/T)] p_T(t)$	$\frac{\pi^2 \sin(\omega T)}{\omega(\pi^2 - T^2 \omega^2)}$
Hamminga	$[0,54 + 0,46\cos(\pi t/T)] p_T(t)$	$\frac{(1,08\pi^2 - 0,16T^2\omega^2)}{\omega(\pi^2 - T^2\omega^2)} \sin(\omega T)$



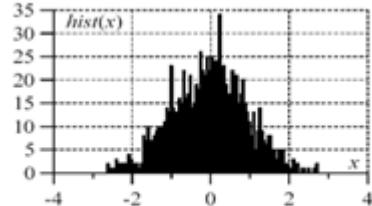
Graficzna ilustracja w dziedzinie częstotliwości zasady konstrukcji okien kosinusoidalnych:

- okno Hanninga $w(t) = [0,5 + 0,5\cos(\pi t/T)] * p_T(t)$,
 - okno Blackmana $w(t) = [0,42 + 0,5\cos(\pi t/T) + 0,08\cos(2\pi t/T)] * p_T(t)$,
- góra – kształt okna, dół – jego widmo częstotliwościowe jako wynik zsumowania kilku przeskalowanych w amplitudzie i przesuniętych w częstotliwości widm PT(omega) okna prostokątnego PT(t)

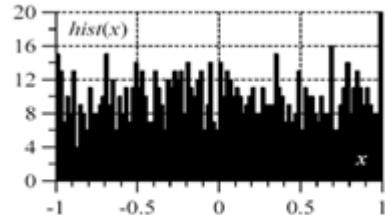
Metoda Welch'a sprowadza się do uśrednienia kilku zmodyfikowanych periodogramów. W metodzie tej ciąg analizowanych N próbek sygnału jest dzielony na L fragmentów o długości M, które zachodzą na siebie lub nie w zależności od przesunięcia D. Następnie wektory danych są wymnażane z wybraną funkcję wagową $w(n)$ (tzw. „oknem”, np. Hamminga) i dla każdego z nich ($i = 0, 1, \dots, L-1$) wyznacza się zmodyfikowany periodogram jako kwadrat modułu transformaty Fouriera nie oryginalnych lecz „zważonych” próbek sygnału. Końcowy estymator funkcji gęstości widmowej mocy jest określony wzorem:

$$\hat{P}_M^{(i)}(f) = \frac{1}{E_w} \left| \sum_{n=0}^{M-1} x^{(i)}(n)w(n)e^{-j2\pi(f/f_{pr})n} \right|^2$$

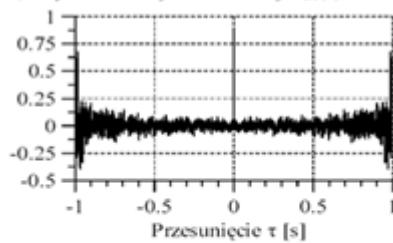
a) histogram szumu „gaussowskiego” $x_1(t)$



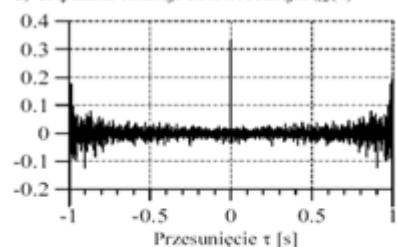
b) histogram szumu „równomiernego” $x_2(t)$



c) estymata funkcji autokorelacji $R_{x1}(\tau)$



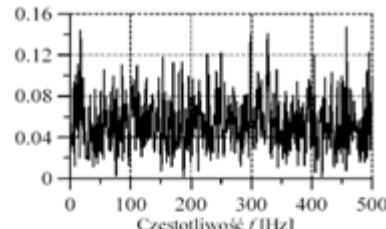
d) estymata funkcji autokorelacji $R_{x2}(\tau)$



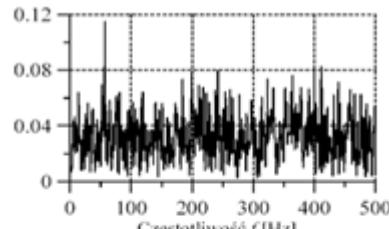
Przykładowe wyniki analizy sygnałów losowych szumu : po lewej - szumu o rozkładzie normalnym, po prawej - szumu o rozkładzie równomiernym,

a) i b) histogram, c) o d) estymata funkcji autokorelacji

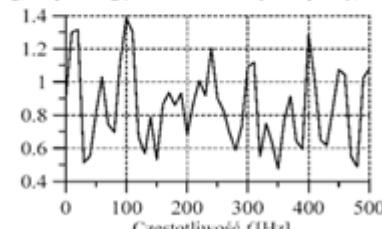
e) widmo Fouriera $2|X_1(f)|$



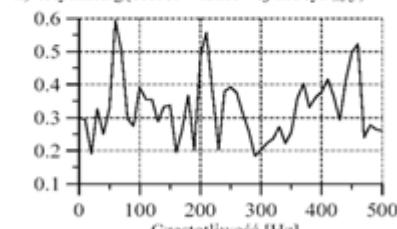
f) widmo Fouriera $2|X_2(f)|$



g) estymata gęstości widmowej mocy $P_{x1}(f)$



h) estymata gęstości widmowej mocy $P_{x2}(f)$



Przykładowe wyniki analizy sygnałów losowych szumu : po lewej - szumu o rozkładzie normalnym, po prawej - szumu o rozkładzie równomiernym,

e) i f) widmo amplitudowe Fouriera, g) i h) estymata funkcji gęstości widmowej mocy

Sygnały cyfrowe

Twierdzenie o próbkowaniu

Załóżmy istnienie sygnału $x(t)$ o paśmie ograniczonym pulsacją ω_m i widmie $X(\omega)$.

Podstawowe twierdzenie o próbkowaniu:

Jeżeli widmo $X(\omega)$ sygnału $x(t)$ spełnia warunek $X(\omega) = 0$ dla $|\omega| \geq \omega_m$, to sygnał ten można odtworzyć z pełną dokładnością na podstawie jego próbek pobieranych z okresem $T_s \leq \frac{\pi}{\omega_m}$. **Przyjmiemy tu założenie, że widmo $X(\omega)$ sygnału $x(t)$ nie zawiera składowych częstotliwościowych dla $|\omega| \geq \omega_m$.**

Twierdzenie o próbkowaniu z częstotliwością Nyquista:

Niech $x(t)$ będzie dowolnym sygnałem, którego widmo spełnia warunek $X(\omega) = 0$ dla $|\omega| \geq \omega_m$. Jeżeli sygnał ten jest próbkowany z okresem $T_s = \frac{\pi}{\omega_m}$ (częstotliwością $f_s = 2 \cdot f_m$), to jego wartości między chwilami próbkowania można odtworzyć na podstawie próbek $x(n \cdot T_s)$ zgodnie ze **wzorem interpolacyjnym Kotelnikowa-Shannona**:

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT_s) \operatorname{Sa} \frac{\omega_m}{\pi} (t - nT_s).$$

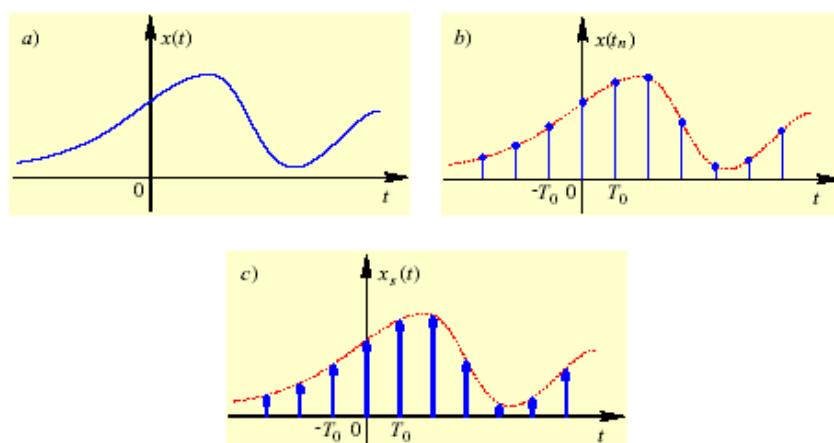
Właściwości twierdzenia o próbkowaniu

1. Właściwość próbkowania

W wyniku mnożenia dowolnego sygnału $x(t)$ przez dystrybucję grzebieniową $\delta_{T_0}(t)$ otrzymujemy ciąg powtarzanych z okresem T_0 impulsów Diraca o wysokościach określonych przez próbki $x(n \cdot T_0)$ sygnału:

$$x(t)\delta_{T_0}(t) = \sum_{n=-\infty}^{\infty} x(nT_0)\delta(t - nT_0)$$

Obwiednią tych impulsów jest sygnał $x(t)$ (otrzymujemy „grzebień”, którego wierzchołki „zębów” układają się w kształt sygnału).



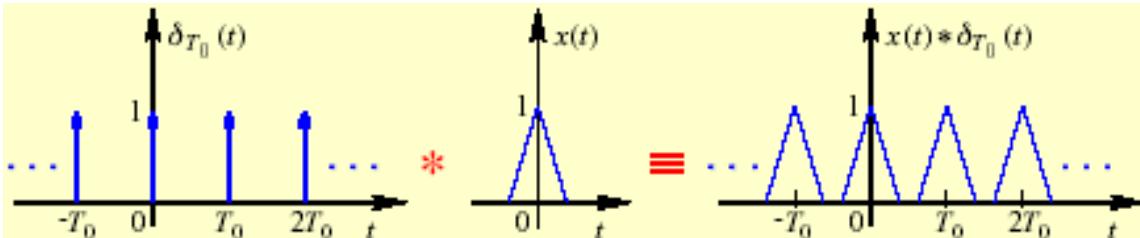
Sygnał dystrybucyjny wyżej zaprezentowany stanowi dystrybucyjną reprezentację spróbkowanego sygnału $x(t)$ (impulsowy sygnał spróbkowany).

2. Właściwość powielenia okresowego

W wyniku splecenia (wymnożenia) dowolnego sygnału $x(t)$ z dystrybucją grzebieniową $\delta_{T_0}(t)$ następuje powielenie okresowe tego sygnału z okresem T_0 :

$$x(t) * \delta_{T_0}(t) = \sum_{n=-\infty}^{\infty} x(t - nT_0)$$

Jeśli $x(t)$ jest sygnałem impulsowym o czasie trwania krótszym niż okres dystrybucji grzebieniowej, sygnał powielony jest ciągiem wiernych kopii sygnału $x(t)$ powtarzonych co przedział T_0 (w efekcie splecenia sygnału z dystrybucją grzebieniową na każdym jej „żębie” pozostawia on swój wierny ślad).



Teoretycznie, równomierne (okresowe) próbkowanie sygnałów analogowych może być przedstawione jako ich wymnożenie z sumą impulsów Diraca:

$$x_\delta(t) = x(t) \sum_{k=-\infty}^{\infty} \delta(t - kT)$$

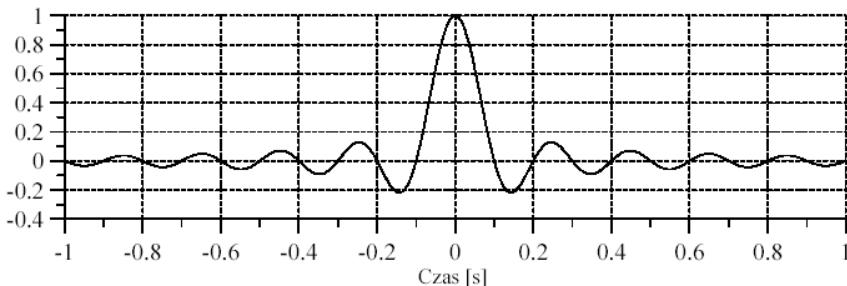
Ponieważ iloczyn w dziedzinie czasu jest równoważny splotowi w dziedzinie częstotliwości - pulsacji (i odwrotnie):

$$\begin{aligned} X_\delta(j\omega) &= \frac{1}{2\pi} \left[X(j\omega) \otimes \omega_p \sum_{k=-\infty}^{\infty} \delta(\omega - k\omega_p) \right] = \frac{1}{2\pi} \left[\omega_p \sum_{k=-\infty}^{\infty} (X(j\omega) \otimes \delta(\omega - k\omega_p)) \right] = \frac{1}{2\pi} \left[\omega_p \sum_{k=-\infty}^{\infty} X(j(\omega - k\omega_p)) \right] \\ f(v) \otimes \delta(v) &= \int_{-\infty}^{+\infty} f(w) \delta(v - w) dw = f(v) \\ f(v) \otimes \delta(v - v_0) &= \int_{-\infty}^{+\infty} f(w) \delta(v - v_0 - w) dw = f(v - v_0) \end{aligned}$$

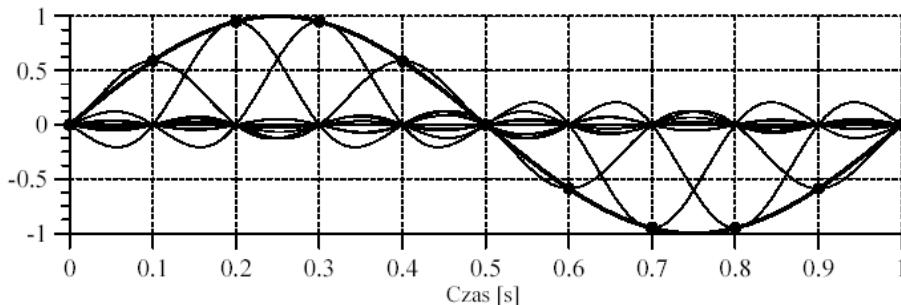
Widmo sygnału po spróbkowaniu jest równe sumie poprzesuwanych widm sygnału oryginalnego. Aby widma te nie „zlały się” i aby było możliwe odtworzenie widma $X(j\omega)$ z widma $X_\delta(j\omega)$, sygnał $x(t)$ powinien mieć widmo ograniczone, tzn. niezerowe tylko dla wybranego przedziału (przedziałów) częstotliwości.

Załóżmy na początek, że $X(j\omega) = 0$ dla $|\omega| > \omega_m$. Jeśli $\omega_m \leq \frac{\omega_p}{2}$, to powyższy warunek jest spełniony i filtrując sygnał $x_\delta(t)$ idealnym filtrem dolnoprzepustowym o pulsacji granicznej (pasma przenoszenia) ω_g , takiej, że $\omega_m \leq \omega_g \leq \omega_p - \omega_m$, jesteśmy w stanie odfiltrować z sygnału $x_\delta(t)$ widma przesunięte $X(j(\omega - k\omega_0))$, $k \neq 0$. Jeśli znamy ω_m sygnału, to najniższą pulsacją próbkowania, nie powodującą utraty informacji „analogowej”, jest $\omega_p = 2\omega_m$. Wówczas idealny filtr ma pulsację graniczną $\omega_g = \omega_m$.

Przykład funkcji rekonstruującej typu sinc, kiedy próbkowanie sygnału ciągłego (analogowego) jest przeprowadzane z częstotliwością $f_{pr} = 10$ herców:



Przykład rekonstrukcji ciągłej sinusoidy (linia pogrubiona) o częstotliwości 1 herca na podstawie jej próbek (punkty „•”), pobranych z częstotliwością 10 herców. Wykorzystano funkcję rekonstruującą typu sinc wyżej pokazaną. Sygnał „zrekonstruowany” jest sumą poprzesuwanych funkcji typu sinc, wziętych z wagami równymi wartościami próbek „•”:



Filtr jest nazywany filtrem rekonstruującym, gdyż jego zastosowanie umożliwia dokładne odtworzenie sygnału analogowego $x(t)$ na podstawie jego próbek $x(kT)$ (właściwy dobór pulsacji ω_g i ω_p dla danego ω_m gwarantuje równość $\hat{x}(t) = x(t)$).

Odtwarzanie przeprowadza się następująco: przesuwa się sygnały:

$$\text{sinc}(\omega_m t) = \frac{\sin(\omega_m t)}{\omega_m t}$$

o czas kT , ($-\infty \leq k \leq \infty$), wymaga się je przez odpowiednie czynniki skalujące w amplitudzie, czyli przez próbki $x(kT)$, oraz sumuje. Jeśli $\omega_p = \frac{2\pi}{T} = 2\omega_m$, to

$$\text{sinc}(\omega_m(t-kT)) = \frac{\sin(\omega_m(t-kT))}{\omega_m(t-kT)} = \frac{\sin\left(\frac{\omega_p}{2}(t-kT)\right)}{\frac{\omega_p}{2}(t-kT)} = \frac{\sin\left(\frac{\pi}{T}(t-kT)\right)}{\frac{\pi}{T}(t-kT)}$$

i dla $t = mT$, $m \neq k$, mamy

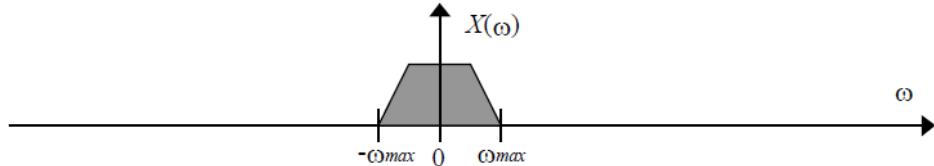
$$\text{sinc}(\omega_m(mT-kT)) = \frac{\sin(\pi(m-k))}{\pi(m-k)} = 0$$

Dla $k = m$ funkcja przyjmuje wartość 1. Wynika stąd, że funkcja rekonstruująca $\text{sinc}()$ przesunięta do punktu czasowego kT i wymnożona przez $x(kT)$ w punkcie tym ma wartość $x(kT)$, a w pozostałych chwilach czasowych próbkowania - zeruje się. Ponieważ jest tak dla wszystkich wartości przesunięć, więc sygnał $\hat{x}(t)$ w chwilach czasowych kT ma dokładnie takie same wartości jak sygnał $x(t)$, czyli $x(kT)$, natomiast „rekonstrukcja” dotyczy przedziałów pomiędzy chwilami próbkowania.

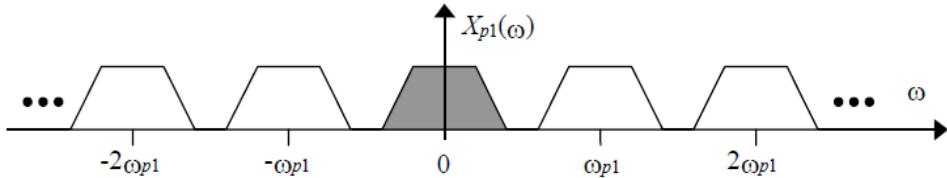
W przypadku kiedy częstotliwość próbkowania jest za mała ($\omega_p < 2\omega_m$) wówczas poprzesuwane „kopie” widma oryginalnego $X(j\omega)$ „zlewają się” (aliasing) i nie jest możliwe odzyskanie $X(j\omega)$ z sumy stosując filtrację.

Konsekwencje częstotliwościowe próbkowania ciągłych sygnałów dolnopasmowych i możliwości ich rekonstrukcji z sygnałów dyskretnych w czasie można zilustrować graficznie:

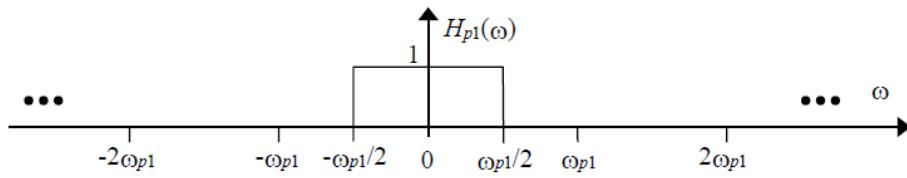
widmo $X(\omega)$ sygnału ciągłego:



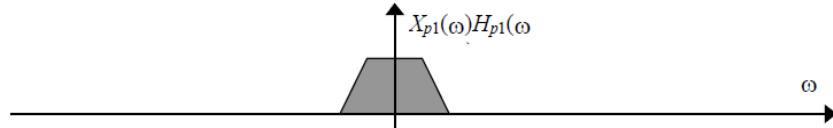
widmo $X_{p1}(\omega)$, sygnału ciągłego po spróbkowaniu z częstotliwością f_{p1} (pulsacją ω_{p1}), taką że $\frac{f_{p1}}{2} > f_{max} \rightarrow f_{p1} > 2 \cdot f_{max}$:



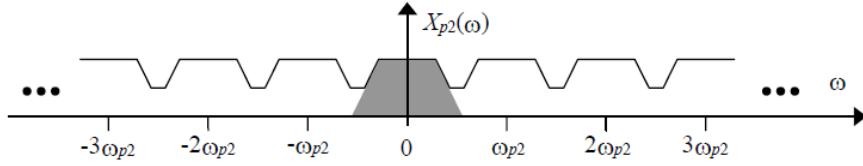
idealna charakterystyka częstotliwościowa $H_{p1}(\omega)$ filtra rekonstruującego o częstotliwości granicznej $f_{g1} = \frac{f_{p1}}{2}$ (pulsacji $\omega_{g1} = \frac{\omega_{p1}}{2}$):



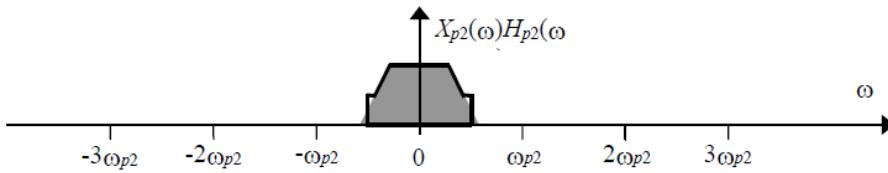
poprawnie „zrekonstruowane” widmo sygnału ciągłego, czyli iloczyn $X_{p1}(\omega) \cdot H_{p1}(\omega)$:



widmo $X_{p2}(\omega)$, sygnału ciągłego po spróbkowaniu z częstotliwością f_{p2} (pulsacją ω_{p2}), taką że $\frac{f_{p2}}{2} < f_{max} \rightarrow f_{p2} < 2 \cdot f_{max}$:



niepoprawnie „zrekonstruowane” widmo sygnału ciągłego, czyli iloczyn $X_{p2}(\omega) \cdot H_{p2}(\omega)$, dla częstotliwości granicznej filtra $H_{p2}(\omega)$ równej $f_{g2} = \frac{f_{p2}}{2}$ (pulsacji $\omega_{g2} = \frac{\omega_{p2}}{2}$):



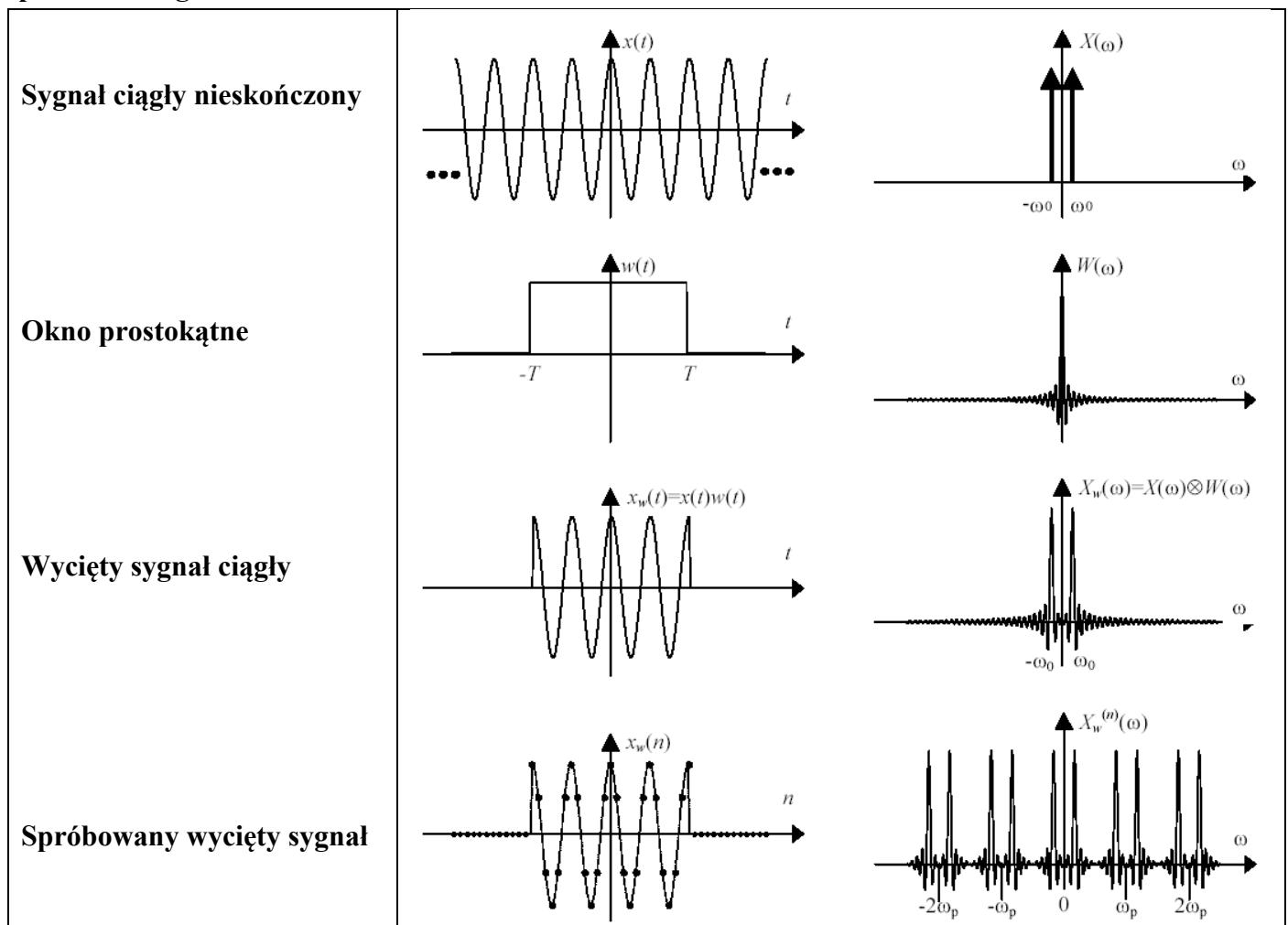
Aliasing

Efekt nakładania się powielonych okresowo kopii widma sygnału w wyniku próbkowania sygnału z częstotliwością mniejszą od częstotliwości Nyquista. Aliasing powoduje niepoprawne zrekonstruowanie widma sygnału ciągłego.

Okna ciągłe $w(t)$ i ich widma $W(\omega)$

Nazwa okna	$w(t)$	$W(\omega)$
Prostokątne	$p_T(t) = \begin{cases} 1 & \text{dla } t \leq T \\ 0 & \text{dla } t > T \end{cases}$	$\frac{2 \sin \omega T}{\omega}$
Trójkątne (Bartletta)	$q_T(t) = \begin{cases} 1 - t /T & \text{dla } t \leq T \\ 0 & \text{dla } t > T \end{cases}$	$T \left[\frac{\sin(\omega T/2)}{\omega T/2} \right]^2$
Hanninga (Hanna)	$[0,5 + 0,5 \cos(\pi t/T)] p_T(t)$	$\frac{\pi^2 \sin(\omega T)}{\omega (\pi^2 - T^2 \omega^2)}$
Hamminga	$[0,54 + 0,46 \cos(\pi t/T)] p_T(t)$	$\frac{(1,08\pi^2 - 0,16T^2\omega^2)}{\omega(\pi^2 - T^2\omega^2)} \sin(\omega T)$

Przykładowe widmo fragmentu sygnału kosinusoidalnego, „wyciętego” przez okno prostokątne i potem spróbkowanego:



Zwyczajowo sygnały, które są tylko spróbkowane w czasie, są nazywane sygnałami dyskretnymi (ang. discrete-time), sygnały zaś dyskretne, które dodatkowo są poddane kwantyzacji przyjmowanych wartości i kodowaniu - sygnałami **cyfrowymi**.

Proces przetwarzania sygnałów ciągłych na dyskretne składa się z operacji:

- **próbkowania w czasie,**

- **kwantowania wartości,**
- **kodowania.**

Operacja próbkowania (dyskretyzacji) sygnału w czasie polega na pobraniu z sygnału ciągłego $x(t)$ jego „próbek” w wybranych, najczęściej równoodległych, chwilach czasowych.

Kwantyzacja to sprowadzenie zbioru wartości, przyjmowanych przez sygnał $x(t)$, najczęściej nieskończonego zbioru liczb rzeczywistych, do jego skończonego podzbioru. W jej wyniku powstaje sygnał $x_q(t)$, czyli sygnał ciągły przyjmujący tylko wartości z określonego zbioru. Kwantyzacja sygnału wynika z konieczności zastosowania przetwornika analogowo-cyfrowego przed wprowadzeniem „danych” do komputera oraz z faktu, że w komputerze wszystkie liczby, a więc także i „próbki” sygnału, są przechowywane w określonej liczbie bitów (8, 16, 24, 32, 48, 64) w jednym z możliwych formatach zapisu. Formaty te zaś mają ograniczoną precyzję i nie umożliwiają przedstawienia dowolnej liczby rzeczywistej.

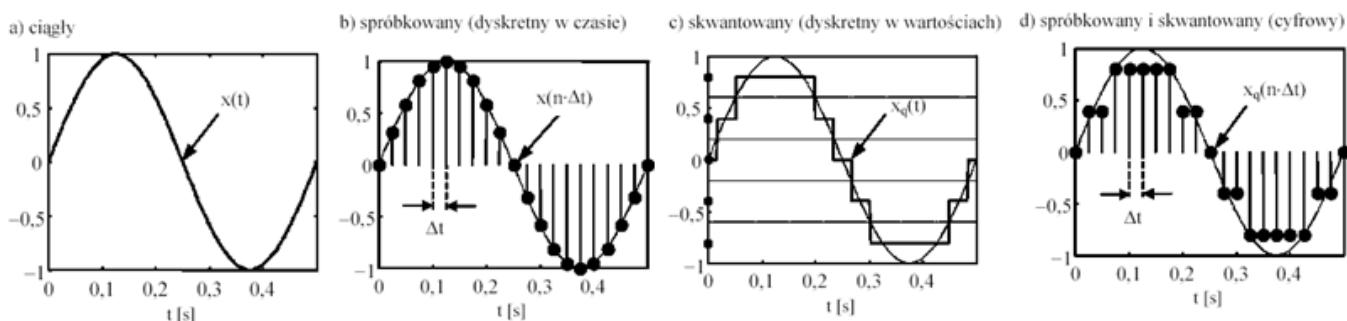
Kodowanie - wybór określonego formatu oraz zapis liczby skwantowanej w tym formacie.

Sygnały ciągłe czasu ciągłego $x(t)$ są opisane ciągłymi funkcjami czasu, przyjmującymi wartości ze zbioru liczb rzeczywistych.

Sygnały dyskretne czasu ciągłego $x_k(t)$ są sygnałami ciągłymi w czasie, przyjmującymi wartości dyskretne. Przykładem takiego sygnału jest sygnał wyjściowy z przetwornika cyfrowo-analogowego C/A, którego wartości są skwantowane. W szczególności sygnał muzyki z płyty CD po przetworniku C/A a przed dolnoprzepustowym filtrem rekonstruującym.

Sygnały ciągłe czasu dyskretnego $x(n)$ powstają w wyniku dyskretyzacji (spróbkowania) w czasie sygnałów ciągłych, tzn. z sygnału ciągłego pobierane są wartości (próbki) tylko w wybranych chwilach czasowych. Próbkowanie może być **równomierne** (równe odstępy czasowe pomiędzy chwilami pobierania próbek) lub **nierównomierne**. W przypadku próbkowania równomiernego odstęp pomiędzy próbami Δt nazywa się okresem próbkowania, a jego odwrotność – częstotliwością próbkowania $f_p = \frac{1}{\Delta t}$. Zapis $x(n)$ oznacza wartość sygnału n-tej chwili czasowej, tzn. $x(n) = x(n\Delta t)$.

Sygnały cyfrowe $x_k(n)$, czyli sygnały dyskretne czasu dyskretnego - sygnały ciągłe czasu dyskretnego $x(n)$, w których dodatkowo dokonano **kwantowania** wartości sygnału, przykładowo zaokrąglono wartości rzeczywiste do najbliższych liczb całkowitych. Sygnały cyfrowe otrzymywane są z tzw. przetworników analogowo-cyfrowych (A/C), w których przeprowadza się równocześnie dyskretyzację czasu i kwantowanie wartości sygnałów analogowych (ciągłych).



Układy analogowe

1. Układy LTI

Układy LTI (Linear Time-Invariant) to układy liniowe niezmienne w czasie.

Liniowy

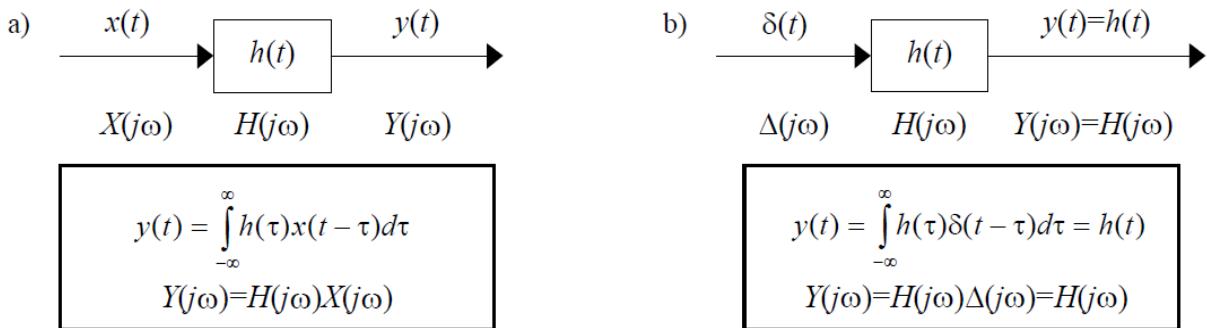
$$x_1(t) \rightarrow y_1(t) \quad x_2(t) \rightarrow y_2(t) \quad ax_1(t) + bx_2(t) \rightarrow ay_1(t) + by_2(t)$$

Czyli odpowiedź układu na sumę wymuszeń jest równa sumie odpowiedzi na poszczególne wymuszenia działające osobno.

Niezmienny w czasie

$$x(t) \rightarrow y(t) \quad x(t - t_0) \rightarrow y(t - t_0)$$

Czyli odpowiedź układu na wymuszenie opóźnione o t_0 jest taka sama jak na oryginalne wymuszenie, tylko opóźniona.



Rys. 5.1. Schemat blokowy i równania liniowego układu analogowego niezmiennego w czasie: $x(t)$ – sygnał wejściowy, $y(t)$ – sygnał wyjściowy, $h(t)$ – odpowiedź impulsowa układu, $\delta(t)$ – impuls Diraca, $X(j\omega)$, $Y(j\omega)$, $H(j\omega)$ i $\Delta(j\omega)$ – transformaty Fouriera sygnałów czasowych, $\Delta(j\omega) = 1$

Układy LTI przetwarzają sygnały wejściowe $x(t)$ na sygnały wyjściowe $y(t)$ z wykorzystaniem swojej *odpowiedzi impulsowej* $h(t)$ - rys. a)

$$x(t) \rightarrow y(t) = \int_{-\infty}^{\infty} h(\tau)x(t - \tau)d\tau$$

Odpowiedzią impulsową $h(t)$ układu nazywamy, jak sama nazwa wskazuje, jego odpowiedź na *impuls* Diraca, czyli sygnał $\delta(t)$ - rys. b)

$$\delta(t) \rightarrow y(t) = \int_{-\infty}^{\infty} h(\tau)\delta(t - \tau)d\tau = h(t)$$

$$\delta(t) \rightarrow h(t)$$

(definicja odpowiedzi impulsowej)

$$\delta(t - \tau) \rightarrow h(t - \tau)$$

(opóźnienie o czas τ , niezmienność w czasie)

$$x(\tau)\delta(t - \tau) \rightarrow x(\tau)h(t - \tau)$$

(skalowanie przez współczynnik $x(\tau)$)

$$\int_{-\infty}^{+\infty} x(\tau)\delta(t - \tau)d\tau \rightarrow \int_{-\infty}^{+\infty} x(\tau)h(t - \tau)d\tau$$

(sumowanie wielu opóźnień, liniowość)

$$x(t) \rightarrow y(t)$$

(ogólnienie)

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau \Leftrightarrow y(t) = x(t) \otimes h(t)$$

Czyli odpowiedź układu LTI jest równa splotowi sygnału wejściowego z funkcją przejścia układu.

Z właściwości splotu przekształcenia Fouriera wynika, że splotowi sygnałów w dziedzinie czasu odpowiada iloczyn ich widm Fouriera w dziedzinie częstotliwości:

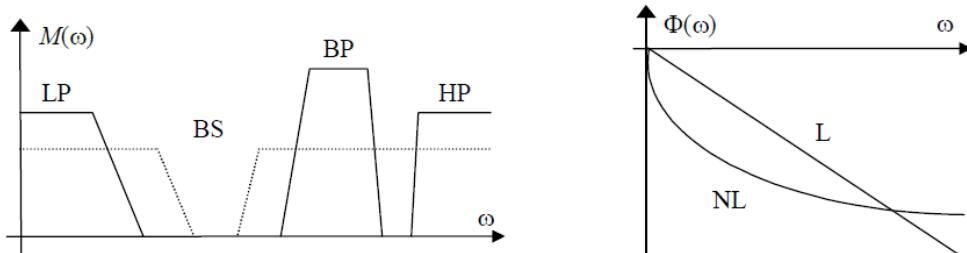
$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau \Leftrightarrow Y(j\omega) = X(j\omega)H(j\omega)$$

Sygnal wyjściowy jest równy sygnalowi wejściowemu, pomnożonemu przez konkretną (jedną) liczbę zespoloną $H(j\omega)$. Liczbę tę można przedstawić we współrzędnych biegunkowych jako promień i kąt, czyli $moduł M(\omega) = |H(j\omega)|$ i fazę $\Phi(\omega) = \angle H(j\omega) = \arctg(\text{Im}(H(j\omega))/\text{Re}(H(j\omega)))$,

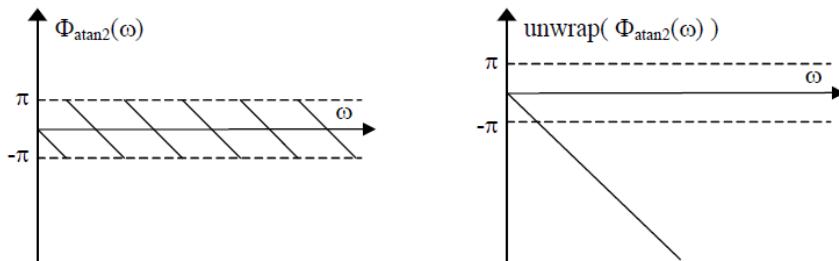
$$H(j\omega) = M(\omega)e^{j\Phi(\omega)}, \quad M(\omega) = |H(j\omega)|, \quad \Phi(\omega) = \angle H(j\omega) = \arctg[\text{Im}(H(j\omega)) / \text{Re}(H(j\omega))]$$

$$y(t) = \left\{ M(\omega)e^{j\Phi(\omega)} \right\} e^{j\omega t} = M(\omega)e^{j(\omega t + \Phi(\omega))}$$

Jak widać $M(\omega)$ wpływa na amplitudę sygnału wyjściowego, $\Phi(\omega)$ przesuwa go zaś w fazie. Z tego powodu zbiór wszystkich wartości $M(\omega)$ dla różnych pulsacji ω nazywa się charakterystką amplitudowo-częstotliwościową układu analogowego, a $\Phi(\omega)$ – jego charakterystyką fazowo-częstotliwościową. Samo $H(j\omega)$ często nazywa się *odpowiedzią* częstotliwościową układu. Na rysunku 5.2a zostały przedstawione różne typy charakterystyk amplitudowo-częstotliwościowych $M(\omega)$: dolnoprzepustowa LP (ang. LowPass), górnoprzepustowa HP (ang. HighPass), pasmowoprzepustowa BP (ang. BandPass) i pasmowozaporowa BS (ang. BandStop). Na rysunku 5.2b pokazano natomiast dwie przykładowe charakterystyki fazowo-częstotliwościowe $\Phi(\omega)$: liniową (L) i nieliniową (NL).



Rys. 5.2. Przykład różnych charakterystyk amplitudowo-częstotliwościowych $M(\omega)$ i fazowo-częstotliwościowych $\Phi(\omega)$ układów analogowych. Oznaczenia: LP – układ dolnoprzepustowy (ang. LowPass), HP – układ górnoprzepustowy (ang. HighPass), BP – układ pasmowoprzepustowy (ang. BandPass), BS – układ pasmowozaporowy (ang. BandStop), L – liniowa charakterystyka fazowa, NL – nieliniowa charakterystyka fazowa



Rys. 5.3. Ilustracja problemu wyznaczania charakterystyki fazowo-częstotliwościowej układu ze wzoru $\Phi(\omega) = \arctg(\text{Im}(H(j\omega)) / \text{Re}(H(j\omega)))$ z dodatkowym uwzględnieniem znaku dzielnej i dzielnika (funkcja atan2(imag(H), real(H)) w języku Matlab)

Jeśli przez układ jest przetwarzany sygnał harmoniczny $x(t) = \exp(j\omega t)$, to otrzymujemy

$$y(t) = M(\omega)e^{j(\omega t + \Phi(\omega))} = M(\omega)e^{j\omega(t-\alpha)}$$

czyli sygnał opóźniony w czasie o α sekund. W przypadku sumy dwóch sygnałów o różnych pulsacjach ω_1 i ω_2 mamy (liniowość układu):

$$y(t) = M(\omega_1)e^{j\omega_1(t-\alpha)} + M(\omega_2)e^{j\omega_2(t-\alpha)}$$

Wynika stąd, że liniowość fazy zapewnia jednakowe opóźnienie czasowe wszystkich składowych sygnału na wyjściu, niezależnie od ich częstotliwości. A to z kolei powoduje, że jeśli tylko $M(\omega_1) = M(\omega_2) = \dots$ itd., to układ nie zmienia kształtu przenoszonego sygnału.

Opóźnienie czasowe sygnału o pulsacji ω na wyjściu układu liniowego o transmitancji $H(j\omega) = M(\omega)e^{j\Phi(\omega)}$ jest określone następującym wzorem:

$$\tau(\omega) = -\frac{d\Phi(\omega)}{d\omega}$$

Funkcja $\tau(\omega)$ jest nazywana funkcją opóźnienia grupowego. Dla $\Phi(\omega) = -\alpha\omega$ mamy $\tau(\omega) = \alpha$.

2. Transmitancja układu analogowego, zera i bieguny

Układy analogowe LTI najczęściej są opisywane za pomocą zależności wiążącej sygnał wyjściowy $y(t)$ z układu, i jego kolejne pochodne, z sygnałem wejściowym $x(t)$, i jego pochodnymi:

$$\begin{aligned} a_N \frac{d^N y(t)}{dt^N} + a_{N-1} \frac{d^{N-1} y(t)}{dt^{N-1}} + \dots + a_1 \frac{dy(t)}{dt} + a_0 y(t) &= b_M \frac{d^M x(t)}{dt^M} + b_{M-1} \frac{d^{M-1} x(t)}{dt^{M-1}} + \dots + b_1 \frac{dx(t)}{dt} + b_0 x(t) \\ \sum_{n=0}^N a_n \frac{d^n y(t)}{dt^n} &= \sum_{m=0}^M b_m \frac{d^m x(t)}{dt^m} \end{aligned}$$

Parametrami układu jest liczba pochodnych wyjścia (rzad równania różniczkowego N), liczba pochodnych wejścia (rzad równania różniczkowego M) oraz wartości współczynników a_n i b_m , $n = 0, 1, \dots, N$ i $m = 0, 1, \dots, M$.

Projektowanie układu sprowadza się do przyjęcia konkretnych wartości dla wszystkich parametrów. Kryterium doboru tych wartości stanowi wymagana (pożądana) charakterystyka częstotliwościowa układu. Korzystając z właściwości pochodnej przekształcenia Fouriera (rozdz. 4.2 Zieliński), dokonuje się transformacji powyższego równania

$$j\omega - z_m = B_m e^{j\theta_m}, \quad j\omega - p_n = A_n e^{j\varphi_n}$$

gdzie

$$B_m = |j\omega - z_m|, \quad A_n = |j\omega - p_n|, \quad \theta_m = \arg(j\omega - z_m), \quad \varphi_n = \arg(j\omega - p_n)$$

Widzimy, że współczynniki a_n i b_m równań różniczkowych są równocześnie współczynnikami wielomianów zmiennej zespolonej ($j\omega$) odpowiedzi częstotliwościowej $H(j\omega)$: współczynniki b_m pochodnych sygnału wejściowego są współczynnikami wielomianu licznika $H(j\omega)$, a współczynniki a_n pochodnych sygnału wyjściowego są współczynnikami wielomianu mianownika. Dobierając wartości a_n i b_m wpływamy na $H(j\omega)$.

Obecnie zajmiemy się **metodologią projektowania** układów analogowych o zadanej odpowiedzi częstotliwościowej $H(j\omega)$. Równanie (5.14) można zapisać jako:

$$H(j\omega) = \frac{Y(j\omega)}{X(j\omega)} = \frac{b_M(j\omega - z_1)(j\omega - z_2)\dots(j\omega - z_M)}{a_N(j\omega - p_1)(j\omega - p_2)\dots(j\omega - p_N)}$$

gdzie z_1, z_2, \dots, z_M oznaczają miejsca zerowe (w ogólności liczby zespolone, parami sprzężone) wielomianu licznika zmiennej $j\omega$, a p_1, p_2, \dots, p_N – miejsca zerowe (liczby zespolone, także parami sprzężone) wielomianu mianownika tej zmiennej. Równanie (5.16) można z kolei przedstawić w postaci bardziej „zwartej” jako:

$$H(j\omega) = \frac{Y(j\omega)}{X(j\omega)} = \frac{b_M \prod_{m=1}^M (j\omega - z_m)}{a_N \prod_{n=1}^N (j\omega - p_n)}$$

gdzie dla konkretnego układu liczby zespolone $z_1, z_2, \dots, z_M, p_1, p_2, \dots, p_N$ są stałe, pulsacja ω jest zaś zmienna. Aby wyzerować odpowiedź częstotliwościową $H(j\omega)$ dla wybranej pulsacji ω , należy spowodować, aby jedno z wyrażeń $(j\omega - z_m)$ było równe zero dla tej pulsacji, czyli aby jedno z zer wielomianu licznika było równe $z_m = j\omega$. Odwrotnie, aby układ wzmacniał sygnał wejściowy dla wybranej pulsacji ω , należy spowodować, aby jedno z wyrażeń $(j\omega - p_n)$ było bliskie zeru dla tej pulsacji, czyli aby jedno z zer wielomianu mianownika p_n było prawie równe $j\omega$.

$$\begin{aligned} (a_N(j\omega)^N + a_{N-1}(j\omega)^{N-1} + \dots + a_1(j\omega)^1 + a_0)Y(j\omega) &= (b_M(j\omega)^M + b_{M-1}(j\omega)^{M-1} + \dots + b_1(j\omega)^1 + b_0)X(j\omega) \\ \sum_{n=0}^N a_n(j\omega)^n Y(j\omega) &= \sum_{m=0}^M b_m(j\omega)^m X(j\omega) \end{aligned} \quad (5.13)$$

skąd:

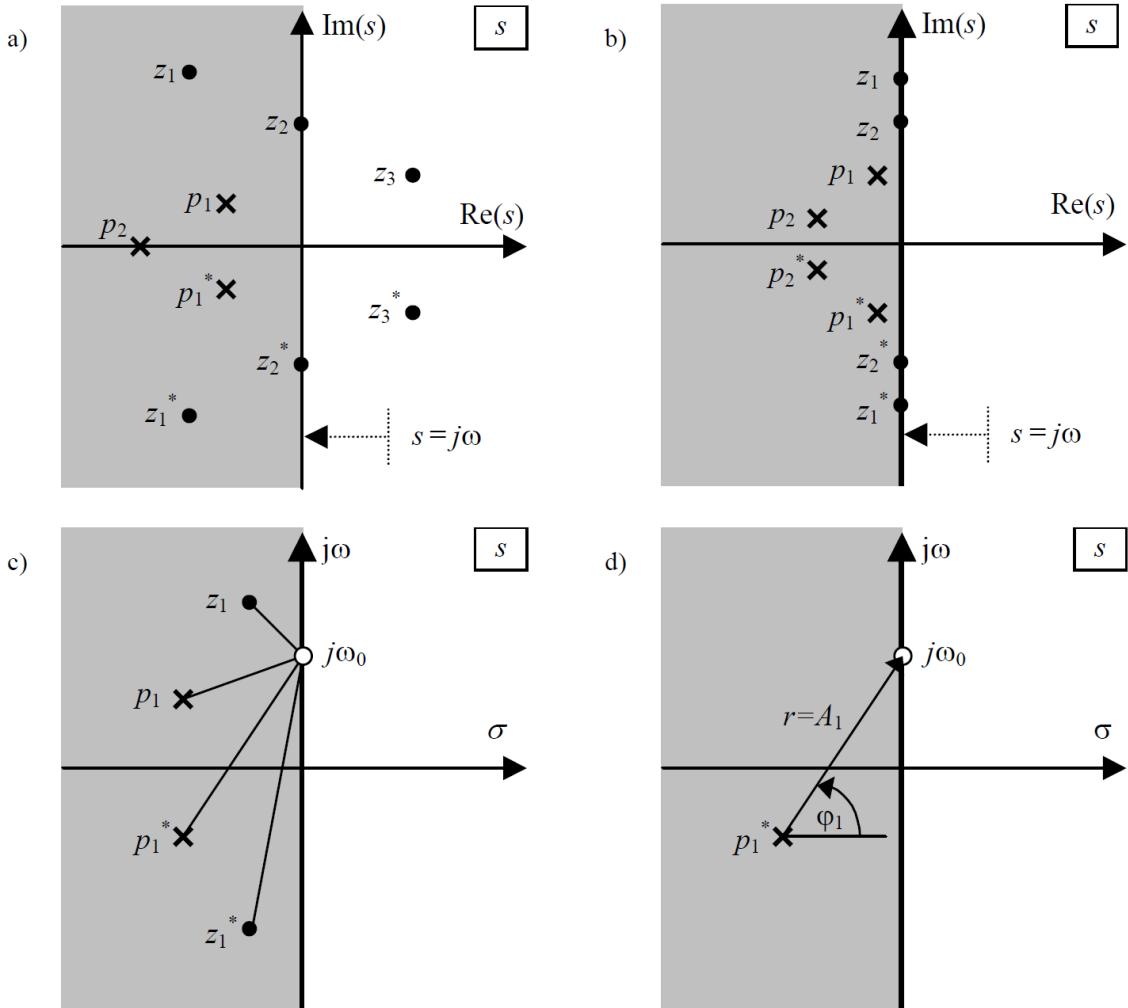
$$H(j\omega) = \frac{Y(j\omega)}{X(j\omega)} = \frac{b_M(j\omega)^M + b_{M-1}(j\omega)^{M-1} + \dots + b_1(j\omega)^1 + b_0}{a_N(j\omega)^N + a_{N-1}(j\omega)^{N-1} + \dots + a_1(j\omega)^1 + a_0} = \frac{\sum_{m=0}^M b_m(j\omega)^m}{\sum_{n=0}^N a_n(j\omega)^n} \quad (5.14)$$

Oczywiście wynikowa odpowiedź impulsowa $h(t)$ układu opisanego równaniem (5.12) jest równa odwrotnemu przekształceniu Fouriera z $H(j\omega)$ (5.14):

$$h(t) = \text{Fourier}^{-1}(H(j\omega)) \quad (5.15)$$

Aby sprawdzić charakterystykę częstotliwościową zaprojektowanego układu należy przesuwać się z częstotliwością wzdłuż osi urojonej i dla każdego $j\omega$ (ω jest zmienne) obliczyć wartości $B_m, A_n, \theta_m, \phi_n, n = 0, 1, \dots, N$ i $m = 0, 1, \dots, M$, czyli liczby zespolone $(j\omega - z_m)$ i $(j\omega - p_n)$ oraz ich moduły (odległości zer licznika i mianownika transmitancji $H(j\omega)$ od punktu $j\omega$) i kąty, a następnie skorzystać z wzorów. $|.|$ oznacza moduł liczby zespolonej, a $\angle(\cdot)$ jej kąt. Wówczas

$$H(j\omega) = M(\omega)e^{j\Phi(\omega)} = \frac{b_M \prod_{m=1}^M B_m e^{j\theta_m}}{a_N \prod_{n=1}^N A_n e^{j\phi_n}} = \frac{B(j\omega)}{A(j\omega)}, \quad M(\omega) = \frac{b_M \prod_{m=1}^M B_m}{a_N \prod_{n=1}^N A_n}, \quad \Phi(\omega) = \sum_{m=1}^M \theta_m - \sum_{n=1}^N \phi_n$$



Rys. 5.4. Ilustracja graficzna metodologii projektowania układów analogowych: a) „dozwolone” położenie zer „●” i biegunów „×” transmitancji układu (bieguny tylko w lewej półpłaszczyźnie), b) przykładowe położenie zer i biegunów dla układu dolnoprzepustowego (zera – „zerują”, bieguny – „podbijają” odpowiednie przedziały odpowiedzi częstotliwościowej układu), c), d) przykładowa konstrukcja wektorów $j\omega_0 - z_k$ i $j\omega_0 - p_k$ (wzór (5.18)), wpływających na kształt $H(j\omega)$ zgodnie z równaniem (5.20). Oznaczenie: $s = \text{Re}(s) + j\text{Im}(s) = \sigma + j\omega$

Zera $\{z_1, z_2, \dots, z_M\}$ i bieguny $\{p_1, p_2, \dots, p_N\}$ transmitancji $H(j\omega)$ są liczbami należącymi do przestrzeni liczb zespolonych s .

Przy założeniu, że współczynniki wielomianów transmitancji $\{a_n, n = 0, 1, \dots, N\}$ i $\{b_m, m = 0, 1, \dots, M\}$ są rzeczywiste, występują one zawsze w parach sprzężonych.

Zera z_k , oznaczone na rysunku przez „●”, mogą leżeć w dowolnym punkcie przestrzeni s , natomiast bieguny p_k , reprezentowane symbolem „X”, ze względu na stabilność układu powinny leżeć tylko w lewej półpłaszczyźnie tej przestrzeni.

Zero $z_k = j\omega_0$ leżące na osi urojonej, zeruje charakterystykę amplitudowo-częstotliwościową układu dla pulsacji $\omega = \omega_0$, natomiast biegun $p_k = \sigma_k + j\omega_k$ powoduje tym większy wzrost wzmocnienia układu dla pulsacji ω_k , im leży bliżej osi urojonej $s = j\omega$, czyli im jest mniejsze σ_k .

Projektowanie $H(j\omega)$ o zadanych właściwościach sprowadza się więc do umieszczenia w odpowiednim miejscu płaszczyzny zespolonej zer i biegunów $H(j\omega)$, czyli, odpowiednio, zer licznika i mianownika.

Z teorii układów analogowych wynika, że aby układ był **stabilny** (odpowiedzią na sygnał wejściowy o ograniczonej amplitudzie jest zawsze sygnał wyjściowy o ograniczonej amplitudzie) rzad wielomianu mianownika $H(j\omega)$ musi być co najmniej taki jak rzad licznika (czyli $N \geq M$) oraz bieguny $H(j\omega)$ (zera mianownika, parami sprzężone liczby zespolone) mogą leżeć tylko w lewej półpłaszczyźnie płaszczyzny liczb zespolonych.

Zera $H(j\omega)$ (zera licznika) mogą natomiast leżeć w dowolnym miejscu tej płaszczyzny. Jeśli jednak i one leżą w lewej półpłaszczyźnie płaszczyzny liczb zespolonych, to układ odwrotny do $H(j\omega)$, czyli $1/H(j\omega)$, jest także zawsze stabilny, jeśli tylko $M = N$. Układ taki nazywa się minimalnofazowym.

3. Przekształcenie Laplace'a, transmitancja Laplace'a

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st} dt$$

$s = \sigma + j\omega$ - zmienna zespolona

Dla $s = j\omega$ przekształcenie Laplace'a przechodzi w jednostronne przekształcenie Fouriera:

$$X(j\omega) = \int_0^{\infty} x(t)e^{-j\omega t} dt$$

Jednostronne i dwustronne przekształcenie Fouriera są takie same dla sygnałów przyczynowych ($x(t) = 0$ dla $t < 0$). Jednostronne przekształcenie Laplace'a ma wszystkie podstawowe właściwości jednostronnego przekształcenia Fouriera, w tym właściwości splotu, iloczynu i pochodnej. Wynika stąd, że:

$$z(t) = \int_{-\infty}^{\infty} x(\tau)y(t-\tau)d\tau \Leftrightarrow Z(s) = X(s)Y(s)$$

$$z(t) = x(t)y(t) \Leftrightarrow Z(s) = \frac{1}{2\pi j} \int_{c-\infty}^{c+\infty} X(r)Y(s-r)dr$$

$$\frac{dx(t)}{dt} \Leftrightarrow sX(s) - x(0+)$$

$$\frac{d^n x(t)}{dt^n} = s^n X(s) - s^{n-1}x(0+) - s^{n-2}x^{(1)}(0+) - \dots - x^{(n-1)}(0+)$$

gdzie $x^{(i)}(0+)$, $i = 1, 2, 3, \dots, n-1$, oznacza i -tą pochodną sygnału $x(\cdot)$ w chwili czasowej 0+.

Wzór na transmitancję Laplace'a układu $H(s)$

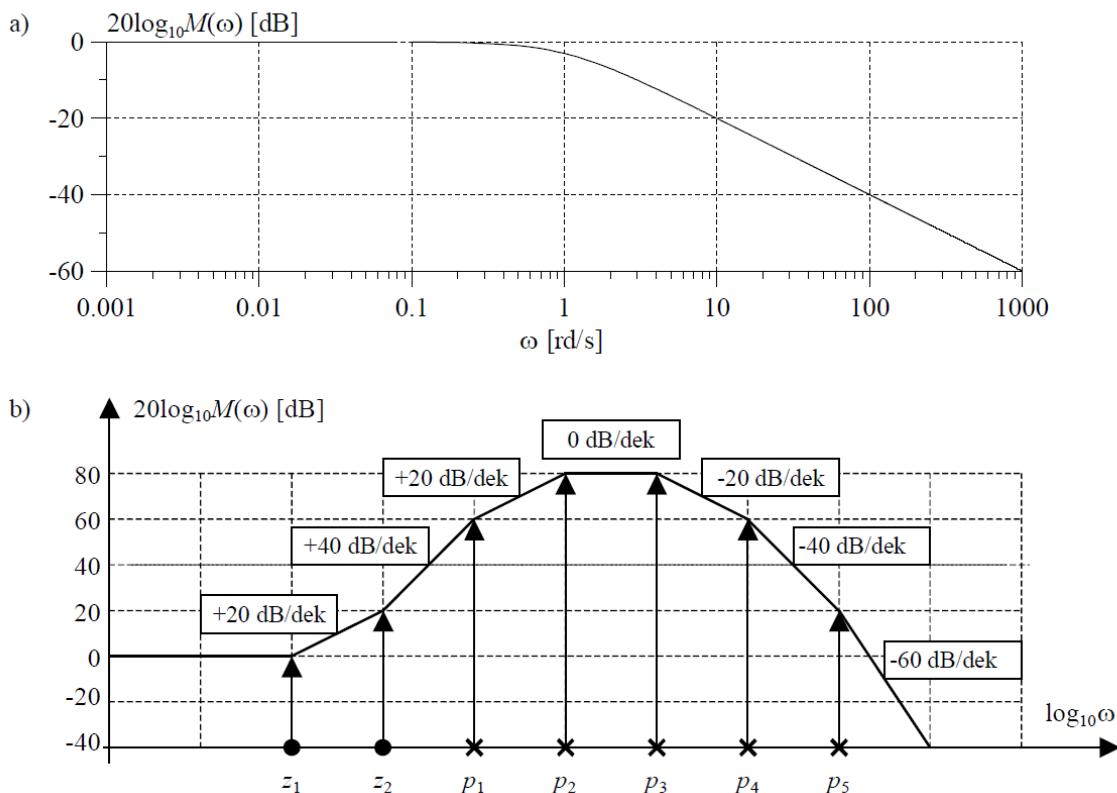
$$H(s) = \frac{Y(s)}{X(s)} = \frac{b_M s^M + b_{M-1} s^{M-1} + \dots + b_1 s^1 + b_0}{a_N s^N + a_{N-1} s^{N-1} + \dots + a_1 s^1 + a_0} = \frac{b_M (s-z_1)(s-z_2)\dots(s-z_M)}{a_N (s-p_1)(s-p_2)\dots(s-p_N)}$$

$$H(s) = \frac{Y(s)}{X(s)} = \frac{b_M \prod_{m=1}^M (s-z_m)}{a_N \prod_{n=1}^N (s-p_n)}$$

Nie należy zapominać, że zmienna $s = \sigma + j\omega$ jest teraz zmienną zespoloną, mającą także część rzeczywistą. Jednak aby uzyskać interpretację częstotliwościową $H(s)$ należy zastosować podstawienie $s = j\omega$, sprowadzające przekształcenie (transformację $H(s)$) Laplace'a w przekształcenie Fouriera (transmitancję $H(j\omega)$).

4. Wykresy Bodego

W zastosowaniach inżynierskich często stosuje się tzw. wykresy Bodego. Nazywa się w ten sposób charakterystyką amplitudowo-częstotliwościową $M(\omega) = |H(j\omega)| = |H(s)|s=j\omega$ układu, w której oś pulsacji $\omega = 2\pi f$ (lub częstotliwości f) jest przedstawiona w skali logarytmicznej, a wzmacnienie układu jest wyrażone w decybelach:

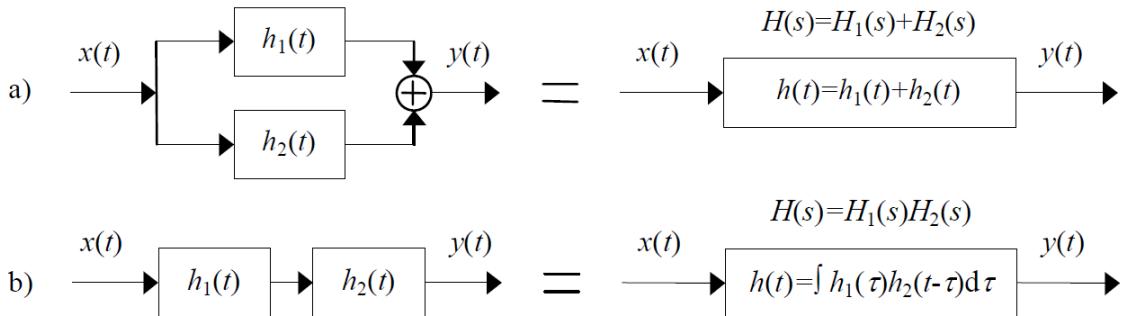


Rys. 5.6. Ilustracja wpływu pulsacji zer z_m i biegunków p_n na nachylenie charakterystyki amplitudowej $M(\omega)$ układów analogowych: a) charakterystyka $M(\omega)$ układu o transmitancji $H(s) = 1/(s+1)$, mającego tylko pojedynczy biegun rzeczywisty $p = -1$, b) orientacyjna, przybliżona linią prostym charakterystyka $M(\omega)$ układu o transmitancji $H(s)$ z wieloma jednokrotnymi zerami i biegunkami zespolonymi, nie występującymi w parach sprzężonych (transmitancja o współczynnikach zespolonych). „●” i „×” oznaczają wartości części urojonych zer i biegunków transmitancji

$$20\log_{10}|H(j\omega)| = 20\log_{10} \frac{b_M \prod_{m=1}^M |j\omega - z_m|}{a_N \prod_{n=1}^N |j\omega - p_n|} = 20\log_{10} \frac{b_M}{a_N} + \sum_{m=1}^M 20\log_{10}|j\omega - z_m| - \sum_{n=1}^N 20\log_{10}|j\omega - p_n|$$

5. Złożone układy analogowe LTI

Prosty układ analogowy LTI ma odpowiedź impulsową $h(t)$, odpowiedź częstotliwościową $H(j\omega)$, transmitancję $H(s)$ oraz realizuje funkcję przetwarzania sygnału, określoną splotem. Z układów prostych można budować jednak układy bardziej złożone, łącząc je szeregowo, jeden za drugim, lub równolegle.



Rys. 5.8. Równoległe (a) i szeregowe (b) połączenie układów LTI

Sygnal wejściowy $x(t)$ wchodzi niezależnie na wejście każdego z nich, a sygnały wyjściowe z obu układów są sumowane. Wówczas:

$$y(t) = y_1(t) + y_2(t) = \int_{-\infty}^{\infty} x(\tau)h_1(t-\tau)d\tau + \int_{-\infty}^{\infty} x(\tau)h_2(t-\tau)d\tau = \int_{-\infty}^{\infty} x(\tau)[h_1(t-\tau) + h_2(t-\tau)]d\tau$$

czyli otrzymujemy układ LTI o odpowiedzi $h(t)$, równej sumie odpowiedzi $h_1(t)$ i $h_2(t)$:

$$h(t) = h_1(t) + h_2(t)$$

Ponieważ całkowe przekształcenia Fouriera i Laplace'a są liniowe, stąd:

$$H(j\omega) = H_1(j\omega) + H_2(j\omega), \quad H(s) = H_1(s) + H_2(s)$$

W konsekwencji transmitancja $H(s)$ otrzymanego układu jest równa sumie transmitancji $H_1(s)$ i $H_2(s)$ układów składowych.

Dla **połączenia szeregowego** (kaskadowego) (rysunek 5.8b) wyjście z pierwszego układu stanowi wejście drugiego układu, czyli jego odpowiedź impulsowa jest równa splotowi odpowiedzi impulsowych obu układów:

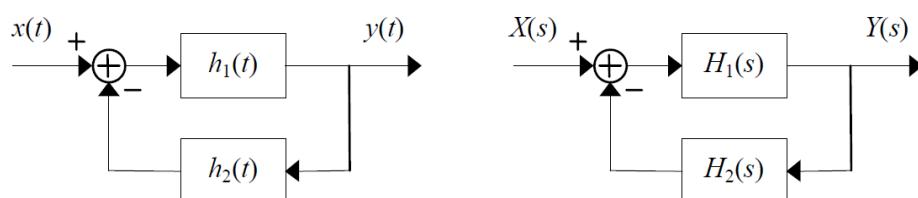
$$\delta(t) \rightarrow [h_1(t)] \rightarrow h_1(t) \rightarrow [h_2(t)] \rightarrow h(t) = \int_{-\infty}^{\infty} h_1(\tau)h_2(t-\tau)d\tau$$

Z właściwości splotu całkowego przekształcenia Fouriera i Laplace'a wynika natomiast, że:

$$H(j\omega) = H_1(j\omega)H_2(j\omega), \quad H(s) = H_1(s)H_2(s)$$

W wyniku połączenia równoległego i szeregowego układów LTI także uzyskujemy układy LTI, czyli liniowe i niezmienne w czasie, ponieważ odpowiedź układu złożonego na sumę dwóch sygnałów jest równa sumie odpowiedzi na każdy z nich z osobna, oraz odpowiedź układu na sygnał przesunięty w czasie jest równa przesuniętej odpowiedzi na pobudzenie oryginalne

Układy ze sprzężeniem zwrotnym



Rys. 5.9. Przykład układu LTI ze sprzężeniem zwrotnym

$$Y(s) = H_1(s)[X(s) - H_2(s)Y(s)]$$

z której wynika, że jego transmitancja jest równa:

$$H(s) = \frac{Y(s)}{X(s)} = \frac{H_1(s)}{1 + H_1(s)H_2(s)}$$

1. Dyskretne przekształcenie Fouriera DFT

W praktyce nie jest możliwe wykonanie sumowania w nieskończonych granicach dlatego robimy to w skończonych granicach i co za tym idzie transformujemy fragment sygnału a nie jego całość.

$$X^{(N)}\left(e^{j\Omega}\right) = \sum_{n=0}^{N-1} x(n)e^{-j\Omega n} \quad x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X^{(N)}\left(e^{j\Omega}\right) e^{j\Omega n} d\Omega$$

Można określić że wyznaczmy transformatę Fouriera iloczynu sygnału i okna prostokątnego.

- Okno wycina fragment sygnału
- Właściwość iloczynu sygnałów ciągłego przekształcenia Fouriera oznacza że w analizie wyznaczamy splot sygnału i okna
- Pulsacja przyjmuje wartości rzeczywiste z przedziału $(-\pi, \pi)$. Mamy więc nieskończoną ich liczbę do wyboru. Z przyczyn praktycznych wybieramy określona ilość, im więcej tym lepiej bo im więcej prążków (próbek) tym lepiej oddany rzeczywisty kształt funkcji "ciąglej". Najczęściej wybieramy tyle prążków ile próbek (N). Upraszczając: przyjmujemy że sygnał jest sygn. Okresowym o okresie N i obliczamy jego N składowych harmonicznych.

W pełni dyskretna realizowana obliczeniowo transformacja Fouriera wyprowadzana jest więc przez dyskretyzację szeregu Fouriera dla sygnałów ciągłych i nazywa się ją **Dyskretną Transformacją Fouriera DFT**:

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}kn} \quad k = 0, 1, 2, \dots, N-1$$

$$x(n) = \sum_{k=0}^{N-1} X(k) e^{j\frac{2\pi}{N}kn} \quad n = 0, 1, 2, \dots, N-1$$

Właściwości DFT:

- Liniowość

$$\sum_{n=0}^{N-1} [ax(n) + by(n)] e^{-j\left(\frac{2\pi}{N}k\right)n} = a \sum_{n=0}^{N-1} x(n) e^{-j\left(\frac{2\pi}{N}k\right)n} + b \sum_{n=0}^{N-1} y(n) e^{-j\left(\frac{2\pi}{N}k\right)n} = aX(k) + bY(k)$$

- Niewrażliwość na przesunięcie w module ($m = n - n_0$):

$$\sum_{n=0}^{N-1} x(n - n_0) e^{-j\left(\frac{2\pi}{N}k\right)n} = \sum_{m=0}^{N-1} x(m) e^{-j\left(\frac{2\pi}{N}k\right)(m+n_0)} = e^{-j\left(\frac{2\pi}{N}k\right)n_0} X(k)$$

- Własność splotu

$$\begin{aligned} z(l) &= \frac{1}{N} \sum_{k=0}^{N-1} Z(k) e^{\frac{j2\pi}{N}kl} = \frac{1}{N} \sum_{k=0}^{N-1} [X(k)Y(k)] e^{\frac{j2\pi}{N}kl} = \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \left[\left(\sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}kn} \right) \left(\sum_{m=0}^{N-1} y(m) e^{-j\frac{2\pi}{N}km} \right) \right] e^{\frac{j2\pi}{N}kl} = \frac{1}{N} \sum_{n=0}^{N-1} x(n) \sum_{m=0}^{N-1} y(m) \left[\sum_{k=0}^{N-1} e^{\frac{j2\pi}{N}k(l-m-n)} \right] \end{aligned}$$

- Własność iloczynu

Iloczyn dwóch sygnałów okresowych

$$z(n) = x(n)y(n) \quad n = 0, 1, 2, \dots, N-1$$

Odpowiada cykliczny splot ich widm

$$Z(k) = \frac{1}{N} \sum_{p=0}^{N-1} X(p)Y(k-p) \quad k = 0, 1, 2, \dots, N-1$$

- **(A)symetria DFT sygnału rzeczywistego.**

$$X\left(\frac{N}{2}+k\right) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi}{N}\left(\frac{N}{2}+k\right)n} = \sum_{n=0}^{N-1} x(n)e^{-j\pi n}e^{-j\frac{2\pi}{N}kn} = \sum_{n=0}^{N-1} [x(n)(-1)^n]e^{-j\frac{2\pi}{N}kn}$$

Analogicznie:

$$X\left(\frac{N}{2}-k\right) = \sum_{n=0}^{N-1} [x(n)(-1)^n]e^{j\frac{2\pi}{N}kn}$$

Stąd mamy:

$$X\left(\frac{N}{2}+k\right) = X^*\left(\frac{N}{2}-k\right)$$

$N/2$ prążek widma leży na osi symetrii, względem niego część rzeczywista jest symetryczna a urojona symetryczna.

Użyteczna informacja o widmie jest w pierwszych $N/2-1$ prążkach.

2. Szybka transformata Fouriera

Bezpośrednia metoda wykonania obliczeń równań dyskretnej transformacji Fouriera jest bardzo wymagająca i zakłada ogromną liczbę operacji matematycznych. Równania te jednak się optymalizuje numerycznie dochodząc do algorytmów tzw. szybkich transformat Fouriera (FFT). Najczęściej stosowane:

- podział w dziedzinie czasu - DIT (Decimation in time)
- podział w dziedzinie częstotliwości - DIF (Decimation in frequency)

2.1. Podział w dziedzinie czasu - DIT (ang. Decimation in Time)

Ideeą, którą się stosuje w algorytmach typu DIT FFT jest podział próbek transformowanego sygnału na te o indeksach parzystych (0, 2, 4, ...) i nieparzystych (1, 3, 5, ...), wykonanie DFT na każdym z tych zbiorów, a następnie odtworzenie widma „całego” sygnału z dwóch widm „częstokowych”.

W algorytmie radix-2 (podział na „dwa” podzbiory) DIT (decymacja w czasie, czyli próbki sygnału) FFT wymaga się, aby transformowany sygnał składał się z $N = 2p$ próbek, a następnie:

- 1) dokonuje się zmiany kolejności próbek, dzieląc je rekurencyjnie na próbki o indeksach parzystych i nieparzystych, aż do uzyskania zbiorów dwuelementowych;
- 2) wykonuje się serię $N/2$ dwupunktowych DFT;
- 3) następnie składa się widma dwuprążkowe w widma czteroprążkowe, czteroprążkowe w ośmioprążkowe itd., aż do momentu odtworzenia widma N -prążkowego, czyli widma całego sygnału.

Przykład:

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{-kn} \quad k = 0, 1, 2, \dots, N-1. \quad W_N = e^{j2\pi/N}$$

Dzielimy na parzyste i nieparzyste próbki

$$X(k) = \sum_{n=0}^{N/2-1} x(2n)W_N^{-k(2n)} + \sum_{n=0}^{N/2-1} x(2n+1)W_N^{-k(2n+1)} \quad k = 0, 1, 2, \dots, N-1$$

Ponieważ:

$$W_N^{-2kn} = \exp\left(\frac{2\pi}{N}(-2kn)\right) = \exp\left(\frac{2\pi}{N/2}(-kn)\right) = W_{N/2}^{-kn}$$

Mamy:

$$X(k) = \left[\sum_{n=0}^{N/2-1} x(2n)W_{N/2}^{-kn} \right] + W_N^{-k} \left[\sum_{n=0}^{N/2-1} x(2n+1)W_{N/2}^{-kn} \right] \quad k = 0, 1, 2, \dots, N-1$$

$$W_{N/2}^{-(k+N/2)n} = W_{N/2}^{-kn}W_{N/2}^{-(N/2)n} = W_{N/2}^{-kn} \exp\left(-\frac{2\pi}{N/2}(N/2)n\right) = W_{N/2}^{-kn} \exp(-2\pi n) = W_{N/2}^{-kn}$$

Dla $k=0,1,2,\dots,N/2-1$ możemy zapisać równanie:

$$X(k) = X_{2n}(k) + W_N^{-k}X_{2n+1}(k), \quad k = 0, 1, 2, \dots, N/2-1$$

gdzie $X_{2n}(k)$ i $X_{2n+1}(k)$ oznaczają $N/2$ -punktowe, nieunormowane (brak dzielenia przez $N/2$) transformaty DFT próbek parzystych ($2n$) i nieparzystych ($2n+1$). Prążki $X(k)$ dla $k = N/2, \dots, N-1$ oblicza się natomiast z zależności:

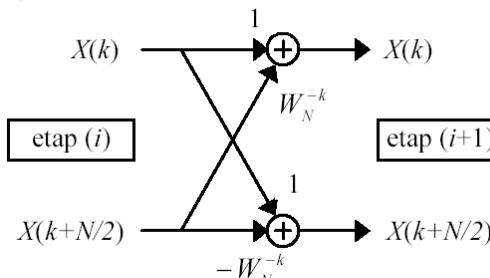
$$X\left(k + \frac{N}{2}\right) = X_{2n}(k) + W_N^{-(k+N/2)}X_{2n+1}(k) \quad k = 0, 1, 2, \dots, N/2-1$$

Co możemy też zapisać jako:

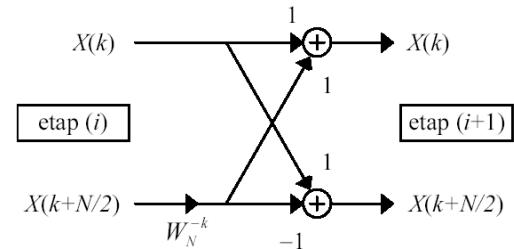
$$X\left(k + \frac{N}{2}\right) = X_{2n}(k) - W_N^{-k}X_{2n+1}(k)$$

$$\begin{bmatrix} X(k) \\ X(k+N/2) \end{bmatrix} = \begin{bmatrix} 1 & W_N^{-k} \\ 1 & -W_N^{-k} \end{bmatrix} \begin{bmatrix} X(k) \\ X(k+N/2) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} X(k) \\ W_N^{-k}X(k+N/2) \end{bmatrix}$$

a)



b)

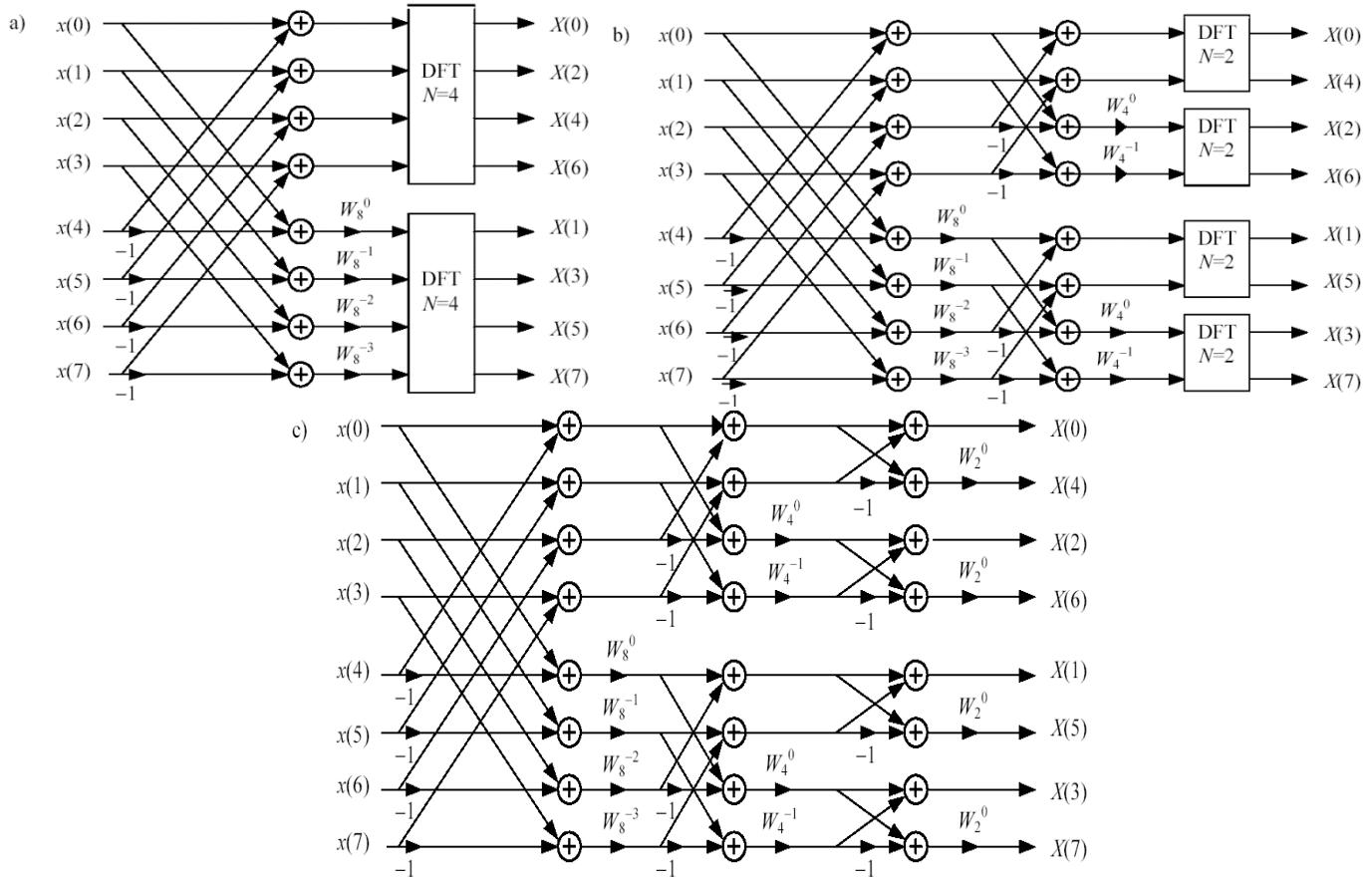


Struktura obliczeniowa podstawowego bloku obliczeniowego algorytmu DIT FFT radix-2, czyli tzw. „motylka”. a) wersja „pełna”, b) wersja „prostsza” o zredukowanej liczbie operacji arytmetycznych, po przesunięciu (wyłączeniu) czynnika .

2.2 Podział w dziedzinie częstotliwości - DIF

W tym przypadku nie dzielimy próbek tylko prążki na parzyste i nieparzyste

$$X(2k) = \sum_{n=0}^{N/2-1} [x(n) + x(n+N/2)] W_{N/2}^{-kn} \quad X(2k+1) = \sum_{n=0}^{N/2-1} \{[x(n) - x(n+N/2)] W_N^{-n}\} W_{N/2}^{-kn}$$



Schemat blokowy algorytmu DIF FFT radix-2:
a) pierwszy, b) drugi, c) końcowy etap dekompozycji

Układy dyskretnie

Układy dyskretne LTI

Niech $x(n)$ oznacza sygnał wejściowy $y(n)$ - sygnał wyjściowy, a $h(n)$ - odpowiedź impulsową układu, czyli jego odpowiedź na pobudzenie dyskretnym impulsem jednostkowym (delta Kroneckera) $\delta(n)$:

$$\delta(n) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$$

Zachodzą dla niego następujące zależności, wiążące sygnał wejściowy i wyjściowy:

$$\delta(n) \rightarrow h(n)$$

$$x(n) \rightarrow y(n)$$

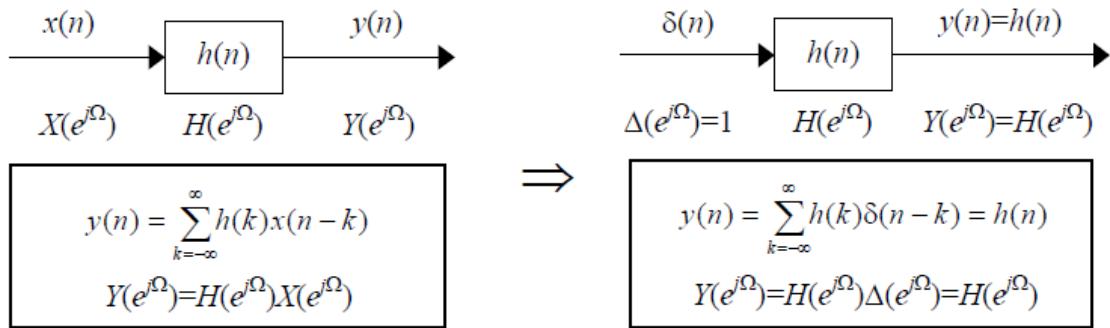
Układ jest niezmienny w czasie, jeśli na opóźnione pobudzenie odpowiada w taki sam sposób, tyleż z opóźnieniem:

Jest natomiast *liniowy*, kiedy na sumę wymuszeń odpowiada sumą odpowiedzi na poszczególne wymuszenia:

$$x(n) \rightarrow y(n) \Rightarrow x(n-n_0) \rightarrow y(n-n_0)$$

Schemat blokowy i równania liniowego układu dyskretnego niezmiennego w czasie: $x(n)$ - sygnał wejściowy, $y(n)$ - sygnał wyjściowy, $h(n)$ - odpowiedź impulsowa układu, $\delta(n)$ - impuls jednostkowy (delta Kroneckera), $X(e^{j\Omega})$, $Y(e^{j\Omega})$, $H(e^{j\Omega})$ transformaty Fouriera dyskretnych sygnałów czasowych. $\Delta(e^{j\Omega})=1$

$$\{x_1(n) \rightarrow y_1(n), \quad x_2(n) \rightarrow y_2(n)\} \Rightarrow \{ax_1(n) + bx_2(n) \rightarrow ay_1(n) + by_2(n)\}$$



Każda próbka $x(n)$ dowolnego sygnału może być przedstawiona jako suma przesuniętych impulsów jednostkowych $\delta(n-k)$ wziętych z „wagami” $x(k)$. Odpowiedzią układu LTI na przesunięty impuls jednostkowy jest przesunięta odpowiedź impulsowa, a odpowiedzią na ważoną sumę przesuniętych impulsów jednostkowych jest ważona suma przesuniętych odpowiedzi impulsowych. Tak więc wyjście $y(n)$ układu dyskretnego LTI jest splotem jego wejścia $x(n)$ z odpowiedzią impulsową tego układu $h(n)$. Po podstawieniu $n-k=m$, można równanie przyjmować postać:

$$y(n) = \sum_{m=-\infty}^{\infty} h(m)x(n-m) = h(n) \otimes x(n)$$

Interpretację częstotliwościową dyskretnych układów LTI podano poniżej.

gdzie $H(e^{j\Omega})$ oznacza zespolone widmo Fouriera dyskretnej odpowiedzi impulsowej $h(n)$. Wynika stąd, że wartość modułu $M(\Omega)$ liczby zespolonej $H(e^{j\Omega})$ będzie decydowała o ewentualnej zmianie amplitudy harmonicznego sygnału wyjściowego, a faza $\Phi(\Omega)$ liczby zespolonej $H(e^{j\Omega})$ - o przesunięciu fazowym (opóźnieniu), wnoszonym przez układ.

$$\begin{aligned} y(n) &= \sum_{k=-\infty}^{\infty} h(k)x(n-k) = \sum_{k=-\infty}^{\infty} h(k)e^{j\Omega(n-k)} = e^{j\Omega n} \sum_{k=-\infty}^{\infty} h(k)e^{-j\Omega k} = x(n)H(e^{j\Omega}) = \\ &= e^{j\Omega n} H(e^{j\Omega}) = e^{j\Omega n} M(\Omega)e^{j\Phi(\Omega)} = M(\Omega)e^{j[\Omega n + \Phi(\Omega)]} \end{aligned}$$

Oznacza to, że projektowanie dyskretnego układu LTI sprowadza się do takiego doboru próbek jego odpowiedzi impulsowej $h(n)$, aby ich transformata Fouriera miała określone właściwości częstotliwościowe („przepuszczanie” i „zerowanie” wybranych częstotliwości).

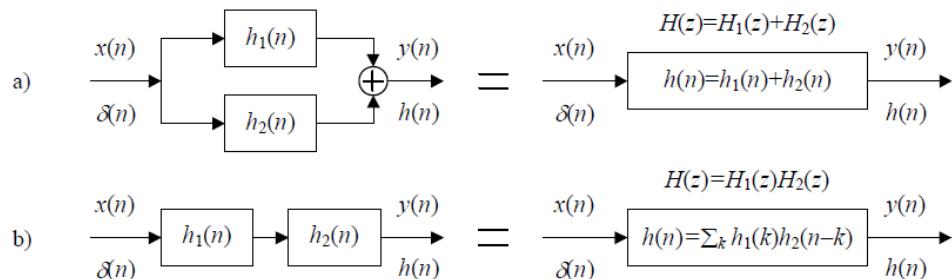
Jeśli układ jest **przyczynowy**, to jego odpowiedź nie wyprzedza pobudzenia. Z definicji dyskretnego impulsu jednostkowego $\delta(n)$ i odpowiedzi impulsowej $h(n)$ wynika więc, że odpowiedź impulsowa $h(n)$ układu LTI przyjmuje wartości równe zero dla $n < 0$. Wówczas ww. zależność upraszcza się do następującej postaci:

Z kolei układ jest **stabilny** w sensie BIBO (ang. *Bounded Input Bounded Output*), kiedy jego odpowiedź na pobudzenie o ograniczonej amplitudzie ($|x(n)| \leq Mx \leq \infty$) ma także zawsze ograniczoną amplitudę ($|y(n)| \leq My \leq \infty$). Wówczas układ się nie wzbuza. Warunkiem koniecznym i wystarczającym stabilności układów LTI jest bezwzględna sumowalność ich odpowiedzi impulsowej:

$$y(n) = \sum_{m=0}^{\infty} h(m)x(n-m) \quad \sum_{n=-\infty}^{\infty} |h(n)| < \infty$$

Równoległe (a) i kaskadowe (b) połączenie dwóch dyskretnych układów LTI:

Uwzględniając w zależności większą liczbę elementów opóźniających, otrzymujemy uogólnioną postać układów LTI ze sprzężeniem zwrotnym, natomiast w wyniku kaskadowego połączenia przyczynowego układu LTI z układem LTI ze sprzężeniem zwrotnym, otrzymujemy uogólnione równanie przyczynowych układów LTI pamiętających zarówno stare „wejścia” jak i „wyjścia”.



$$y(n) + \sum_{k=1}^{\infty} g(k)y(n-k) = \sum_{m=0}^{\infty} h(m)x(n-m)$$

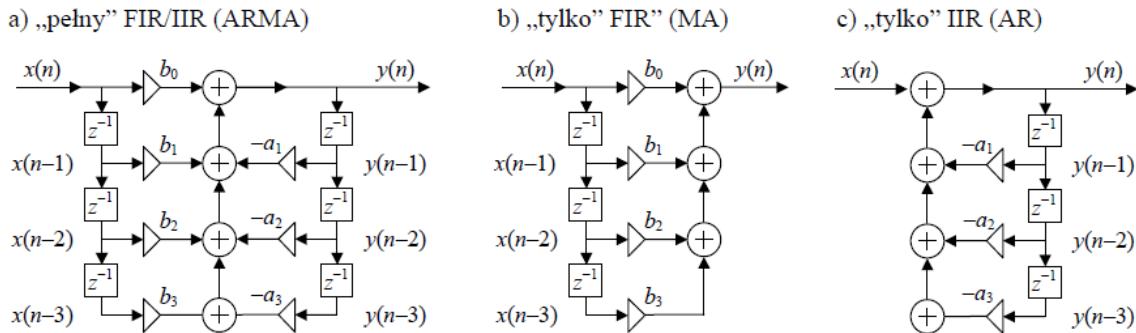
$$y(n) = \sum_{m=0}^{\infty} h(m)x(n-m) - \sum_{k=1}^{\infty} g(k)y(n-k)$$

W praktycznej realizacji „cyfrowej” (mikroprocesorowej) nieskończone sumy zmieniamy na układy LTI o funkcji „przejścia” postaci:

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

$$y(n) = \sum_{m=0}^M b_m x(n-m) - \sum_{k=1}^N a_k y(n-k)$$

Schemat blokowy filtra cyfrowego dla $M = N = 3$: a) pełny, b) linia opóźniająca tylko na wejściu, c) linia opóźniająca tylko na wyjściu. Oznaczenia (angielskie): FIR - *Finite Impulse Response*, IIR - *Infinite Impulse Response*, ARMA - *Autoregressive Moving Average*, MA - *Moving Average*, AR - *Autoregressive*



Transformacja Z

Transformacja Z jest zdefiniowana za pomocą następującego równania:

gdzie „z” jest zmienną zespoloną, pełniącą podobną rolę jak zmienna zespolona „s” w transformacji Laplace'a. Wzór ww. to tzw. szereg Laurenta. Aby istniała transformata $X(z)$ sygnału $x(n)$, szereg ten musi być zbieżny. Dla danego sygnału określa się obszar zmienności wartości „z”, dla którego zbieżność zachodzi:

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n}$$

Dla sygnałów impulsowych o skończonej liczbie niezerowych próbek (skończonym czasie trwania), liczba wyrazów ww. wielomianu jest także skończona i zawsze istnieje obszar zmienności zmiennej „z” zapewniający zbieżność szeregu.

$$0 \leq r_- < |z| < r_+ \leq \infty$$

Przykładowo, wyznaczmy transformaty Z kilku sygnałów, z których każdy ma tylko trzy niezerowe próbki o wartościach 0,25; 0,5 i 0,25; lecz występujące w różnych chwilach czasowych:

$$X_4(z) = x(0) + x(1)z + x(2)z^{-2}, \quad X_4(z) = 0,25 + 0,5z^{-1} + 0,25z^{-2}, \quad z \neq 0$$

$$X_5(z) = x(1)z^{-1} + x(2)z^{-2} + x(3)z^{-3}, \quad X_5(z) = 0,25z^{-1} + 0,5z^{-2} + 0,25z^{-3}, \quad z \neq 0$$

$$X_1(z) = x(-3)z^3 + x(-2)z^2 + x(-1)z, \quad X_1(z) = 0,25z^3 + 0,5z^2 + 0,25z, \quad z \neq \infty$$

$$X_2(z) = x(-2)z^2 + x(-1)z^1 + x(0), \quad X_2(z) = 0,25z^2 + 0,5z + 0,25, \quad z \neq \infty$$

$$X_3(z) = x(-1)z + x(0) + x(1)z^{-1}, \quad X_3(z) = 0,25z + 0,5 + 0,25z^{-1}, \quad z \neq \infty \text{ i } z \neq 0$$

Pierwsze trzy sygnały są nieprzyczynowe, gdyż przyjmują wartości niezerowe dla $n < 0$, natomiast kolejne dwa to sygnały przyczynowe.

Transformata Z sygnałów nieprzyczynowych nie jest określona dla z równego nieskończoności ($z = \infty$), natomiast sygnałów przyczynowych, dla z równego zero ($z = 0$), gdyż wówczas potęga zmiennej z jest równa nieskończoności.

Odwrotna transformacja Z

Zadaniem *odwrotnej transformacji Z* jest wyznaczenie sygnału $x(n)$ na podstawie jego transformaty $X(z)$ i informacji o jej obszarze zbieżności. Jej definicja jest następująca:

gdzie Γ jest „zegaroskrętnym” konturem całkowania, obejmującym początek układu współrzędnych. Zgodnie z twierdzeniem całkowym Cauchego:

$$x(n) = \frac{1}{2\pi j} \oint_{\Gamma} X(z) z^{n-1} dz \quad \frac{1}{2\pi j} \oint_{\Gamma} z^{n-1} dz = \begin{cases} 1, & n=0 \\ 0, & n \neq 0 \end{cases}$$

W najprostszym przypadku, kiedy $X(z)$ jest zwykłym wielomianem zmiennej „ z ”:

$$\frac{1}{2\pi j} \oint_{\Gamma} X(z) z^{n-1} dz = \frac{1}{2\pi j} \oint_{\Gamma} \left[\sum_{k=-\infty}^{\infty} x(k) z^{-k} \right] z^{n-1} dz = \sum_{k=-\infty}^{\infty} \left[x(k) \frac{1}{2\pi j} \oint_{\Gamma} z^{-k+n-1} dz \right] = x(n)$$

Właściwości transformacji Z

$$X(z) = \sum_{n=n_1}^{n_2} b_n z^{-n}$$

Transformata Z jest *liniowa* oraz „*niewrażliwa* w module na przesunięcie sygnału ($m = n-n_0$):

$$\sum_{n=-\infty}^{\infty} [ax_1(n) + bx_2(n)] z^{-n} = a \sum_{n=-\infty}^{\infty} x_1(n) z^{-n} + b \sum_{n=-\infty}^{\infty} x_2(n) z^{-n} = aX_1(z) + bX_2(z)$$

$$\sum_{n=-\infty}^{\infty} x(n-n_0) z^{-n} = \sum_{n=-\infty}^{\infty} x(m) z^{-(m+n_0)} = z^{-n_0} \sum_{n=-\infty}^{\infty} x(m) z^{-m} = z^{-n_0} X(z)$$

Dodatkowo transformata Z ma *właściwość splotu* ($m = n-k$):

$$\sum_{n=-\infty}^{\infty} \left(\sum_{k=-\infty}^{\infty} x(k) y(n-k) \right) z^{-n} = \sum_{m=-\infty}^{\infty} \left(\sum_{k=-\infty}^{\infty} x(k) y(m) \right) z^{-(m+k)} = \left(\sum_{k=-\infty}^{\infty} x(k) z^{-k} \right) \left(\sum_{m=-\infty}^{\infty} y(m) z^{-m} \right) = X(z)Y(z)$$

i *właściwość iloczynu* (korzystamy z definicji odwrotnej transformacji Z ; $\xi = z/v$):

$$\sum_{n=-\infty}^{\infty} [x(n)y(n)] z^{-n} = \sum_{n=-\infty}^{\infty} \left[x(n) \left(\frac{1}{2\pi j} \oint_{\Gamma} Y(\nu) \nu^{n-1} d\nu \right) \right] z^{-n} = \frac{1}{2\pi j} \oint_{\Gamma} \left[\sum_{n=-\infty}^{\infty} x(n)(z/\nu)^{-n} \right] Y(\nu) \nu^{-1} d\nu =$$

$$= \frac{1}{2\pi j} \oint_{\Gamma} X(z/\nu) Y(\nu) \nu^{-1} d\nu$$

Transmitancja układów dyskretnych

Ponieważ dyskretne układy LTI są opisane równaniem splotu dyskretnego to z właściwości splotu transformacji Z wynika, że powyższemu opisowi układu w dziedzinie czasu odpowiada następujące równanie w „dziedzinie” transformaty Z :

$$y(n) = \sum_{m=-\infty}^{\infty} h(m)x(n-m) = \sum_{m=-\infty}^{\infty} x(m)h(n-m)$$

czyli transformata $H(z)$ odpowiedzi impulsowej $h(n)$ (przyczynowej!) stanowi transmitancję tego układu. W przypadku zaś kiedy dyskretny układ LTI jest opisany następującym równaniem:

$$Y(z) = H(z)X(z) \Rightarrow H(z) = \frac{Y(z)}{X(z)} = \sum_{n=-\infty}^{\infty} h(n) z^{-n} = \sum_{n=0}^{\infty} h(n) z^{-n} \quad y(n) = \sum_{m=0}^M b_m x(n-m) - \sum_{k=1}^N a_k y(n-k)$$

to obliczając transformatę Z jego obu stron i korzystając z właściwości liniowości i „niezmienności” na przesunięcie przekształcenia Z , otrzymujemy:

$$Y(z) = \left[\sum_{m=0}^M b_m z^{-m} \right] X(z) - \left[\sum_{k=1}^N a_k z^{-k} \right] Y(z)$$

skąd:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_M z^{-M}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_N z^{-N}} = \frac{(1-z_1 z^{-1})(1-z_2 z^{-1}) \dots (1-z_M z^{-1})}{(1-p_1 z^{-1})(1-p_2 z^{-1}) \dots (1-p_N z^{-1})}$$

gdzie z_k oznaczają zera wielomianu licznika (zera transmitancji), a p_k - zera wielomianu mianownika (bieguny transmitancji). W przypadku kiedy układ nie ma pętli sprzężenia zwrotnego z „wyjścia” ($a_k = 0$), to wówczas transmitancja ma tylko wielomian w liczniku i współczynniki b_m tego wielomianu są równe wartościami próbek odpowiedzi impulsowej $h(n)$. Mnożąc równocześnie licznik i mianownik transmitancji przez z^M i z^N , otrzymujemy ($a_0 = 1$):

czyli transmitancję podobną do transmitancji układów analogowych $H(s)$, w której występują tylko dodatnie potęgi zmiennej zespolonej „ z ”. Układy cyfrowe opisane ww. transmitancjami są stabilne, jeśli ich bieguny p_k , $k = 1, 2, 3, \dots, N$ leżą wewnątrz okręgu jednostkowego na płaszczyźnie zmiennej zespolonej z .

$$H(z) = \frac{z^N b_0 z^M + b_1 z^{M-1} + \dots + b_{M-1} z + b_M}{z^M a_0 z^N + a_1 z^{N-1} + \dots + b_{N-1} z + b_N} = z^{N-M} \frac{b_0 (z - z_1)(z - z_2) \dots (z - z_M)}{a_0 (z - p_1)(z - p_2) \dots (1 - p_N)}$$

Interpretacja częstotliwościowa.

Aby nadać „sens” częstotliwościowy $H(z)$, wystarczy znaleźć związek pomiędzy transformacją Z a transformacją Fouriera dla sygnałów dyskretnych. Ponieważ są one zdefiniowane następująco:

więc widzimy, że transformata $X(z)$ sprowadza się do transformacji Fouriera $X(e^{j\Omega})$ sygnałów dyskretnych dla:

transformacja Z :
$$X(z) = \sum_{n=-\infty}^{\infty} x(n) z^{-n}$$

transformacja Fouriera:
$$X(e^{j\Omega}) = \sum_{n=-\infty}^{+\infty} x(n) e^{-jn\Omega}$$

gdzie Ω jest pulsacją unormowaną względem częstotliwości próbkowania i równą $2\pi F = 2\pi f / f_p$. Wartości z na okręgu jednostkowym są związane wartościami częstotliwości unormowanej $F = f / f_p$, nie istniejącymi poza tym okręgiem.

$$z = e^{j\Omega}$$

Interpretacja częstotliwościowa transmitancji $H(z)$,

Podobnie wynika stąd, że jeśli zastosujemy podstawienie $z = e^{j\Omega}$, to z $H(z)$ widać, jak układ „przetworzy” konkretną pulsację Ω : o ile ją wzmacni (moduł liczby zespolonej $H(e^{j\Omega})$) oraz przesunie w fazie (kąt liczby zespolonej $H(e^{j\Omega})$).

$$\begin{aligned} y(n) &= \sum_{k=-\infty}^{\infty} h(k) x(n-k) = \sum_{k=-\infty}^{\infty} h(k) e^{j\Omega(n-k)} = e^{j\Omega n} \sum_{k=-\infty}^{\infty} h(k) e^{-j\Omega k} = x(n) H(e^{j\Omega}) = x(n) H(z) \Big|_{z=e^{j\Omega}} \\ H(e^{j\Omega}) &= (e^{j\Omega})^{N-M} \frac{b_0 (e^{j\Omega} - z_1)(e^{j\Omega} - z_2) \dots (e^{j\Omega} - z_M)}{a_0 (e^{j\Omega} - p_1)(e^{j\Omega} - p_2) \dots (e^{j\Omega} - p_N)} \\ e^{j\Omega} - z_m &= B_m e^{j\theta_m}, \quad e^{j\Omega} - p_n = A_n e^{j\varphi_n} \end{aligned}$$

gdzie ($|.|$ – moduł liczby zespolonej, $\angle(.)$ – jej kąt):

$$B_m = |e^{j\Omega} - z_m|, \quad A_n = |e^{j\Omega} - p_n|, \quad \theta_m = \angle(e^{j\Omega} - z_m), \quad \varphi_n = \angle(e^{j\Omega} - p_n)$$

zero transmitancji, leżące na okręgu jednostkowym w punkcie $z = e^{j\Omega}$, zeruje transmitancję dla pulsacji Ω , natomiast biegum transmitancji, leżący w punkcie $p = e^{j\Omega}$, $\epsilon \approx 1$, tzn. wewnątrz tego okręgu ale blisko niego, powoduje, że układ wzmacnia pulsację Ω . Wynika to z następujących wzorów:

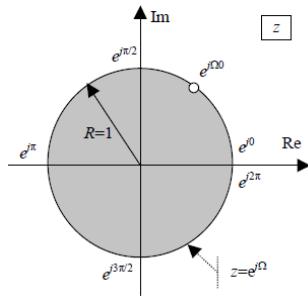
Przykłady projektowania układów dyskretnych metodą „zer i biegunów”

$$\begin{aligned} H(e^{j\Omega}) &= M(\Omega) e^{j\Phi(\Omega)} = (e^{j\Omega})^{N-M} \frac{b_0 \prod_{m=1}^M B_m e^{j\theta_m}}{a_0 \prod_{n=1}^N A_n e^{j\varphi_n}} \\ M(\Omega) &= \frac{b_0 \prod_{m=1}^M B_m}{a_0 \prod_{n=1}^N A_n}, \quad \Phi(\Omega) = \Omega(N - M) + \sum_{m=1}^M \theta_m - \sum_{n=1}^N \varphi_n \end{aligned}$$

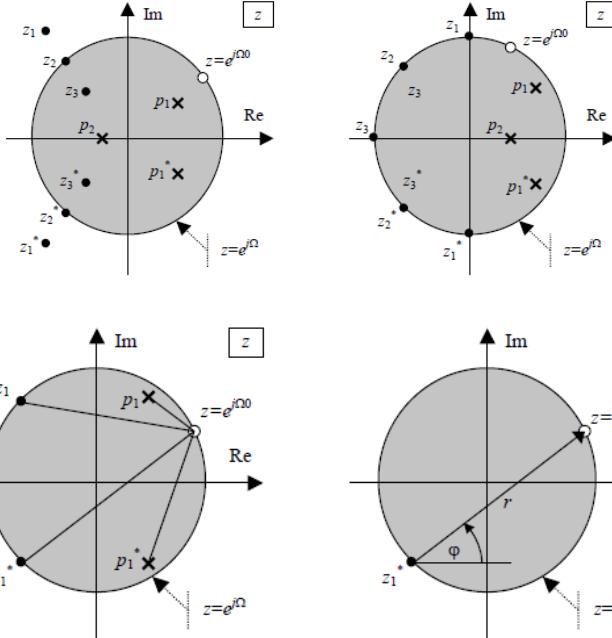
Ilustracja graficzna projektowania transmitancji $H(z)$ metodą zer i biegunów:

a) aby wyznaczyć charakterystykę transmitancji $H(z)$ podstawiamy $z = e^{j\Omega}$, $\Omega = 2\pi f / f_p$, czyli poruszamy się po okręgu o promieniu jednostkowym (zmienności - $[0, 2\pi]$ odpowiada zmienność $f [0, f_m]$);

b) wszystkie zera „ \bullet ” i bieguny „ x ”, jeśli są zespolone, to występują w parach sprzężonych; bieguny muszą leżeć wewnątrz okręgu, a zera nie; zera zazwyczaj leżą na okręgu i zerują charakterystykę filtra dla wybranych pulsacji (częstotliwości), bieguny zaś leżą w pobliżu okręgu i wpływają w ten sposób na wzmacnianie „okolicznych” pulsacji (częstotliwości) przez układ;

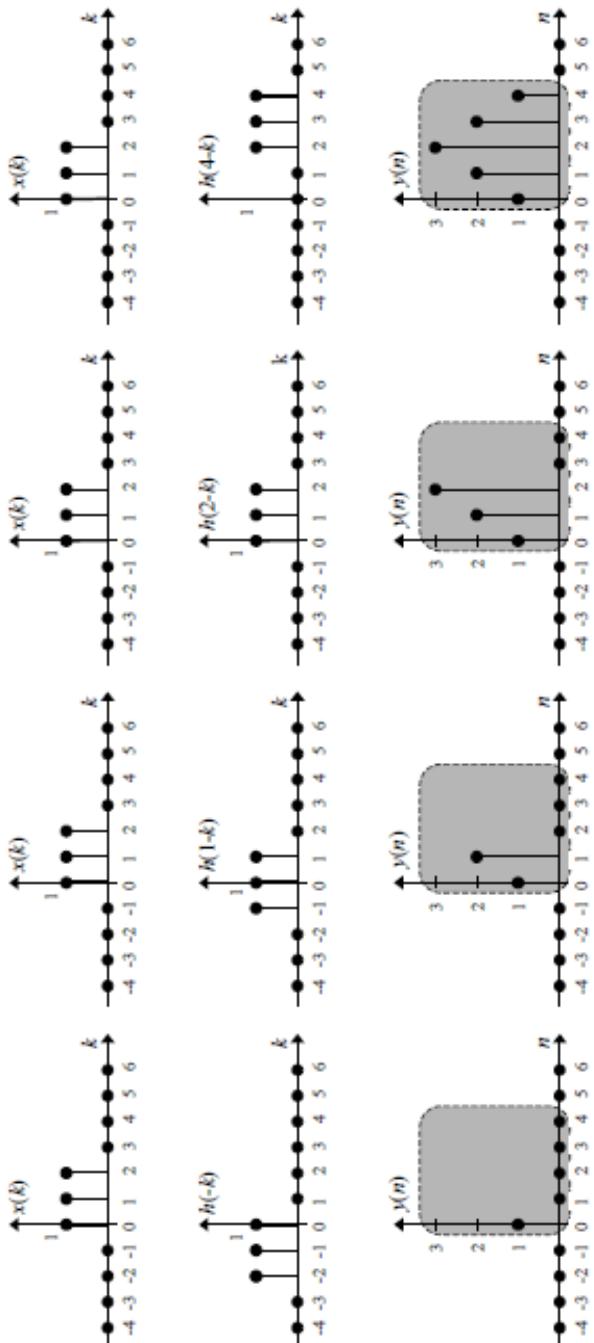


c) ilustracja zasady „konstrukcji” charakterystyki amplitudowo-częstotliwościowej układu (iloraz dwóch iloczynów: iloczynu odległości wszystkich zer od wybranej pulsacji (częstotliwości) i analogicznego iloczynu odległości biegunów) oraz charakterystyki fazowo-częstotliwościowej (suma kątów zer minus suma kątów biegunów)



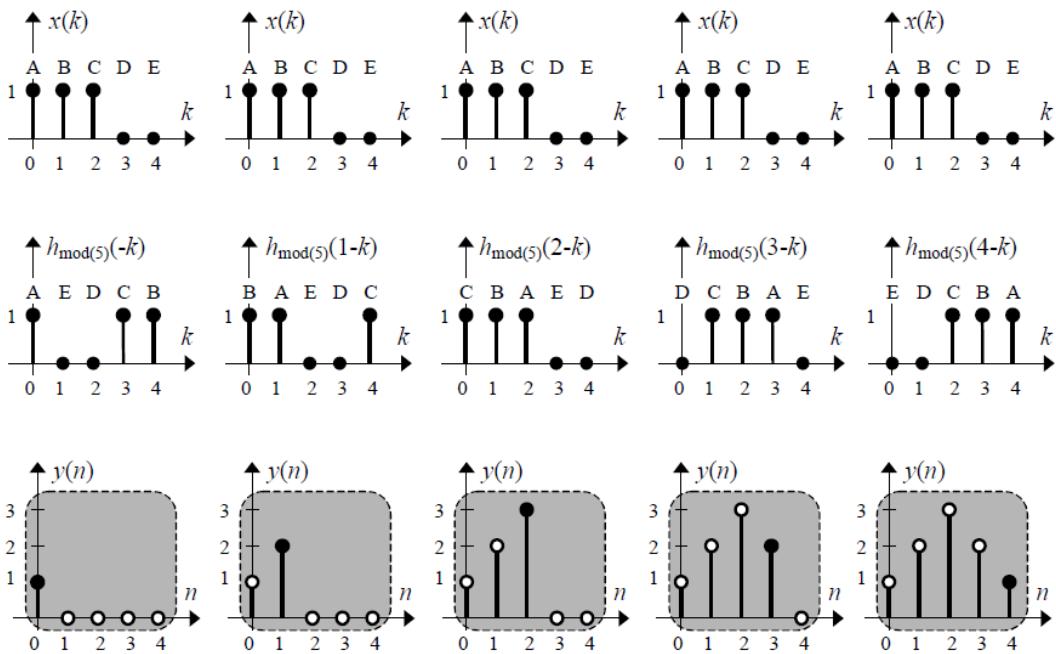
Splot liniowy i kołowy (nie ma w wykładzie, ale jest w zagadnieniach od Doktorka) Splot liniowy

Na rysunku poniżej pokazano kolejne etapy splotu liniowego dwóch prostokątnych sygnałów dyskretnych $x(k)$ i $h(k)$. Oba sygnały są identyczne i mają tylko trzy próbkę niezerowe równe 1 dla $k = 0, 1, 2$. W pierwszym kroku drugi z sygnałów jest odwracany „w czasie”, w wyniku czego otrzymujemy $h(-k)$. W drugim kroku są mnożone parami wszystkie próbki sygnałów $x(k)$ i $h(-k)$ dla tej samej wartości k : $x(k)h(-k)$. W trzecim kroku wszystkie iloczyny są sumowane, a otrzymana w ten sposób liczba jest traktowana jako wartość sygnału wyjściowego z filtra w chwili „zerowej”: $y(0) = \sum k[x(k)h(-k)]$. Następnie sygnał $h(-k)$ jest przesuwany o jedną próbke w prawo i są powtarzane kroki dwa i trzy, ale dla sygnału $h(1-k)$. Otrzymana w ten sposób liczba stanowi wyjście filtra w chwili „pierwszej”: $y(1) = \sum k[x(k) \cdot h(1-k)]$. I cała operacja powtarza się ponownie. Po n_0 przesunięciach sygnału $h(-k)$ w prawo otrzymujemy więc próbke sygnału wyjściowego o numerze n_0 , stanowiącą wynik splotu: $y(n0) = \sum k[x(k)h(n0-k)]$.



Splot kołowy

Splot kołowy dwóch identycznych sygnałów dyskretnych $x(k)$ i $h(k)$ o długości $N = 5$ próbek jest przedstawiony na rysunku poniżej. Dla $k = 0, 1$ i 2 przyjmują one wartość jeden, a dla $k = 3$ i 4 – wartość zero. W splocie kołowym indeksowanie próbek sygnału $h(k)$ jest wykonywane w trybie modulo $h((n0-k)\text{mod}(N))$, dlatego na rysunku sygnał ten został oznaczony jako $x\text{mod}(5)(k)$. W pierwszym kroku sygnał $h\text{mod}(5)(k)$ jest odwracany „w czasie” (**w sposób kołowy!**). W kroku drugim są wymnażane przez siebie pary próbek sygnałów $x(k)$ i $h\text{mod}(5)(-k)$ dla $k = 0, 1, 2, 3, 4$. W kroku trzecim są sumowane wyniki pięciu iloczynów i wynikowa wartość stanowi próbke sygnału wyjściowego o indeksie „0”: $y(0) = \sum k[x(k)h\text{mod}(5)(-k)]$. Następnie sygnał $h\text{mod}(5)(-k)$ jest przesuwany w prawo o jedną próbke (**w sposób kołowy!**). Próbki otrzymanego sygnału $h\text{mod}(5)(1-k)$ są ponownie mnożone z odpowiadającymi im próbками sygnału $x(k)$ o tym samym indeksie k . Zsumowana wartość wszystkich iloczynów stanowi próbke sygnału wyjściowego o indeksie „1”: $y(1) = \sum k[x(k)h\text{mod}(5)(1-k)]$. Postępując w sposób identyczny, po $n0$ przesunięciach w prawo (**kołowych!**) sygnału $h\text{mod}(5)(-k)$ otrzymujemy $n0$ -ową próbke wyniku splotu kołowego: $y(n0) = \sum k[x(k) \cdot h\text{mod}(5)(1-k)]$.



Struktury filtrów cyfrowych

Filtr cyfrowy typu LTI

Klasyczny liniowy filtr cyfrowy typu LTI (**Linear Time-Invariant**) jest zdefiniowany równaniem:

$$y(n) = \sum_{m=0}^M b_m x(n-m) - \sum_{k=1}^N a_k y(n-k)$$

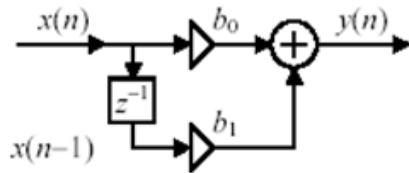
$x(n)$ - sygnał wejściowy, $y(n)$ - sygnał wyjściowy, a_k, b_m – współczynniki wagowe decydujące o właściwościach częstotliwościowych układu i jego stabilności.

Klasyczne struktury filtrów cyfrowych

W klasycznych strukturach filtrów cyfrowych istotną informacją jest blok z^{-1} : jest to układ opóźniający sygnał o jedną próbkę.

Podstawowy filtr nierekursywny

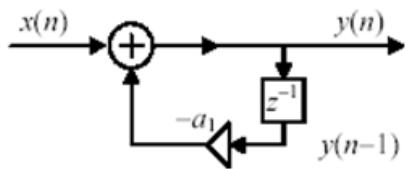
$$y(n) = b_0 x(n) + b_1 x(n-1)$$



Sygnał wyjściowy $y(n)$ jest ważoną sumą dwóch ostatnich próbek wejściowych: obecnej $x(n)$ oraz poprzedniej $x(n-1)$. Wartości wag b_0 i b_1 decydują o charakterze „częstotliwościowym” filtracji, to znaczy w jaki sposób są „przepuszczane” przez układ poszczególne częstotliwości wejściowe.

Podstawowy filtr rekursywny

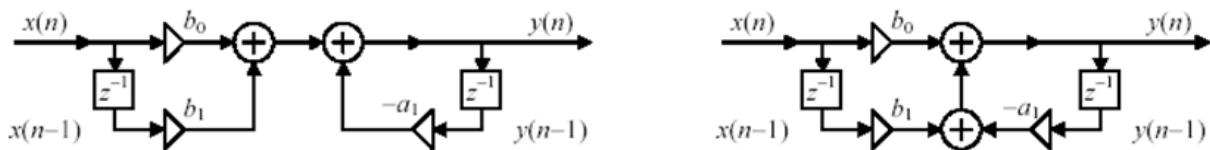
$$y(n) = x(n) - a_1 y(n-1)$$



Filtr rekursywny posiada sprzężenie zwrotne. W tym przypadku każda kolejna próbka wyjściowa $y(n)$ jest ważoną sumą aktualnej próbki wejściowej $x(n)$ oraz poprzedniej próbki wyjściowej $y(n-1)$.

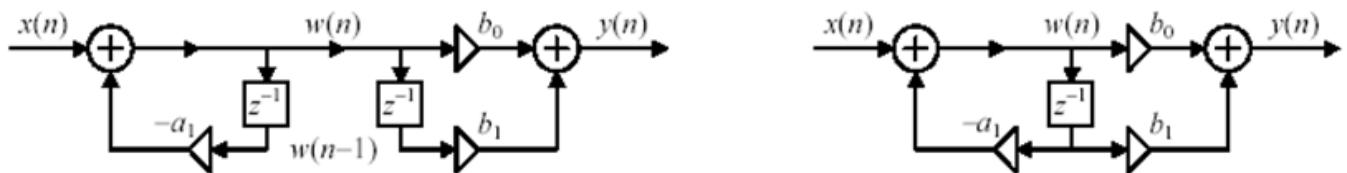
Struktury filtrów połączonych kaskadowo: nierekursywne – rekursywne

kaskadowe połączenie układów z rys. a) i b): $y(n) = b_0 x(n) + b_1 x(n-1) - a_1 y(n-1)$



Struktury filtrów połączonych kaskadowo: rekursywne – nierekursywne

zamiana kolejności układów z rys. a) i b): $w(n) = x(n) - a_1 w(n-1)$, $y(n) = b_0 w(n) + b_1 w(n-1)$

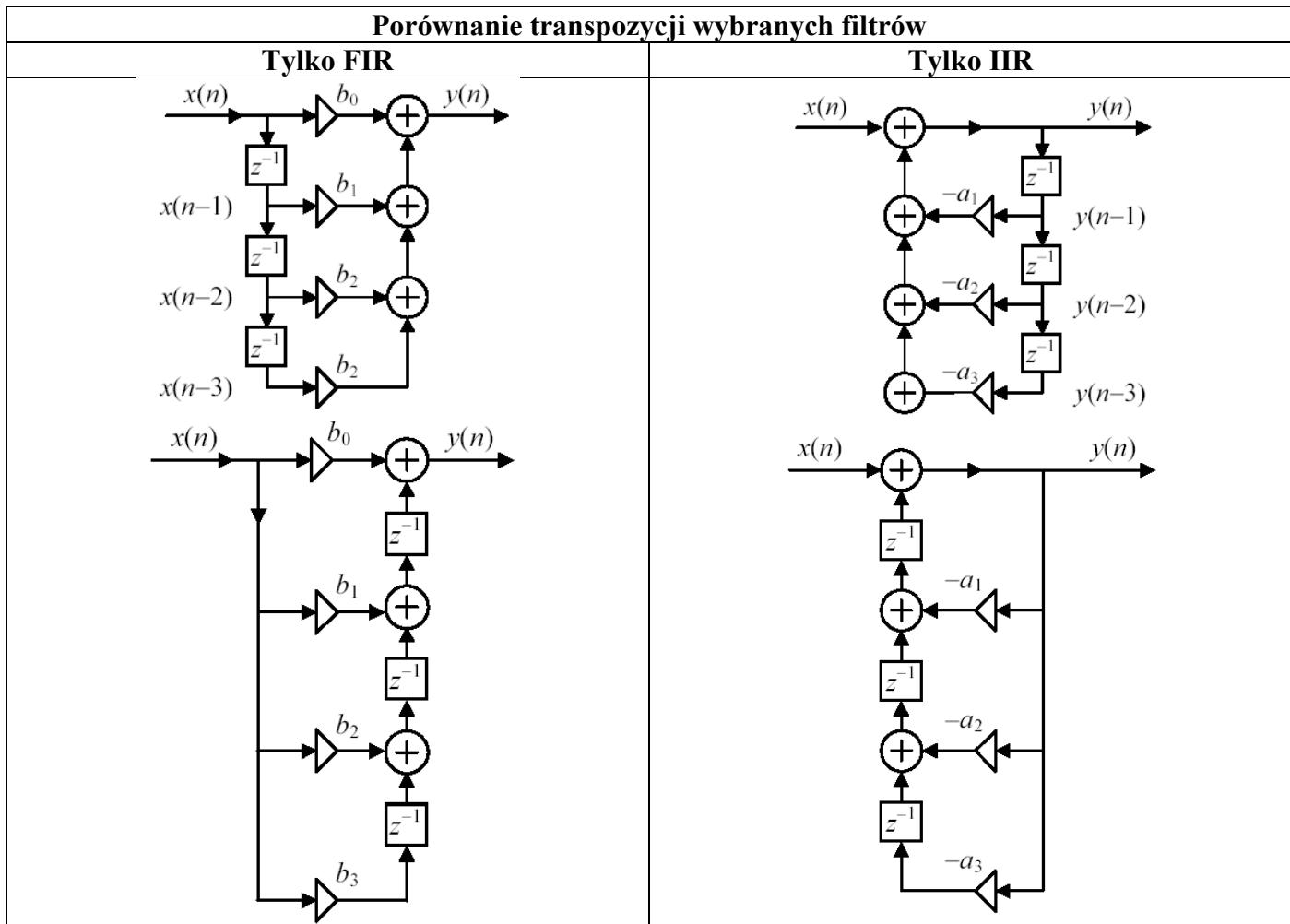
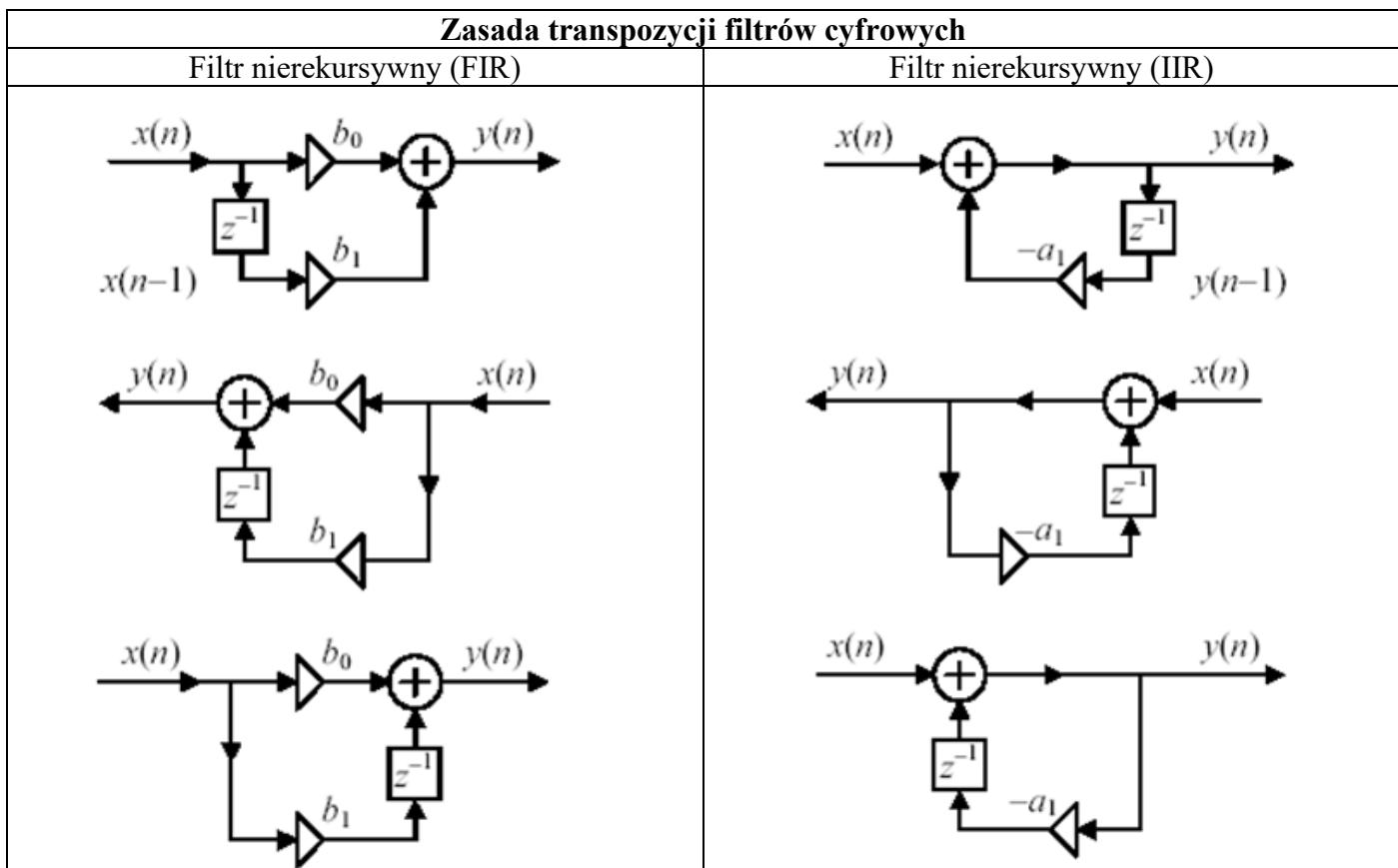


Filtry cyfrowe są przedstawiane za pomocą grafów „przepływowych”, w których informacja przepływa z węzła do węzła i jest pomiędzy tymi węzłami przetwarzana: mnożona przez współczynniki skalujące oraz opóźniana.

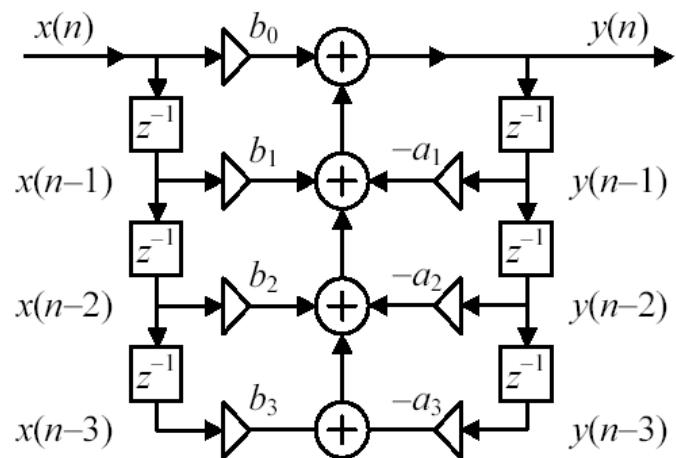
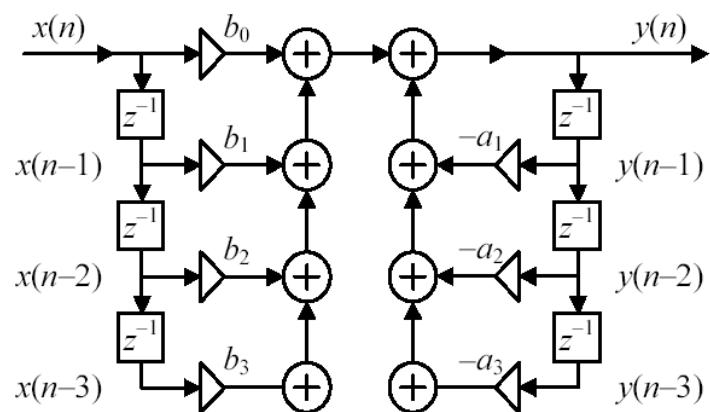
Zasada transpozycji

W teorii grafów liniowych obowiązuje zasada transpozycji, według której, jeśli zmienimy kierunek przepływu danych na przeciwny, to uzyskamy tzw. **układ transponowany**, który realizuje taką samą funkcję jak układ oryginalny.

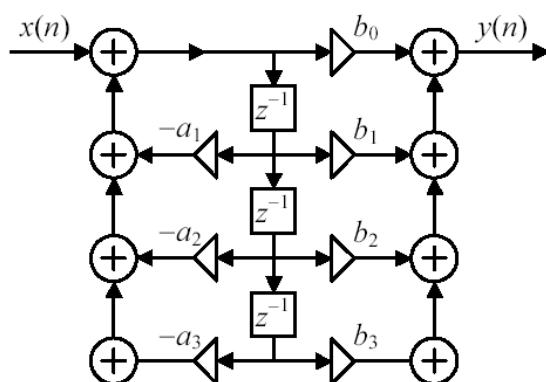
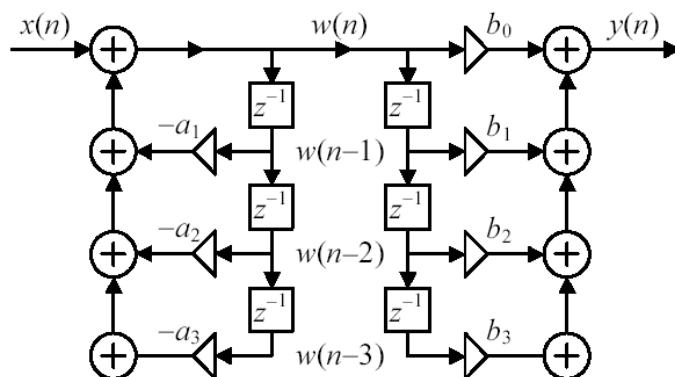
Graficzne przedstawienie zasady transpozycji filtrów cyfrowych (zmiany kierunku przepływu danych na przeciwny) przedstawia poniższa tabela.

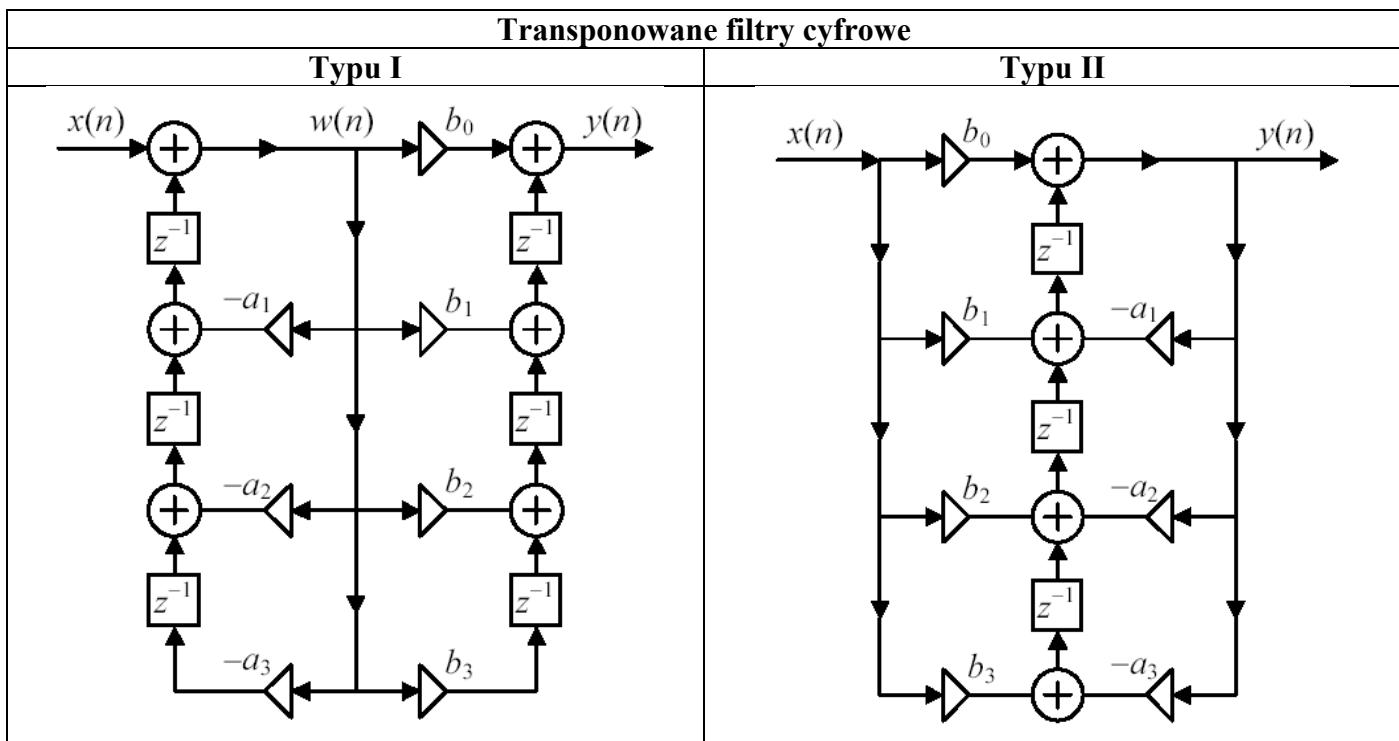


Struktury filtru cyfrowego typu I (kaskadowe połączenie filtrów ‘tylko FIR’ i ‘tylko IIR’)



Struktury filtru cyfrowego typu II (kaskadowe połączenie filtrów ‘tylko IIR’ i ‘tylko FIR’)

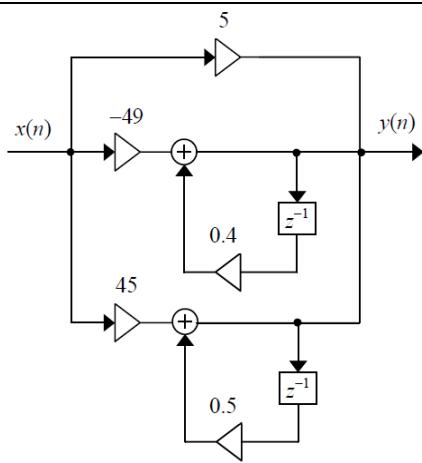




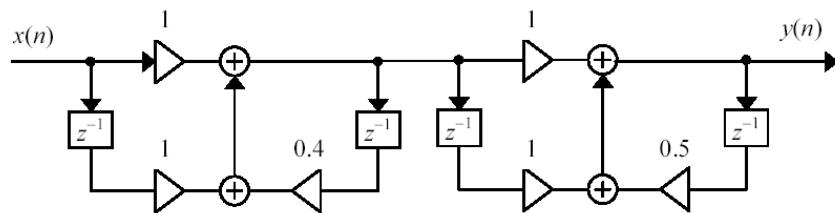
Zamiana klasycznej struktury realizacji filtru

Transmitancja $H(z)$ cyfrowego filtru typu LTI jest **ilorazem dwóch wielomianów zmiennej zespolonej z** . Ponieważ transmitancję tę można zapisać na wiele różnych sposobów, ten sam filtr cyfrowy może mieć bardzo wiele różnych realizacji układowych. **Transmitancję filtra można przedstawić jako sumę lub iloczyn transmitancji prostszych - filtr może być zrealizowany jako kilka filtrów pracujących równolegle lub kaskadowo** (jeden za drugim).

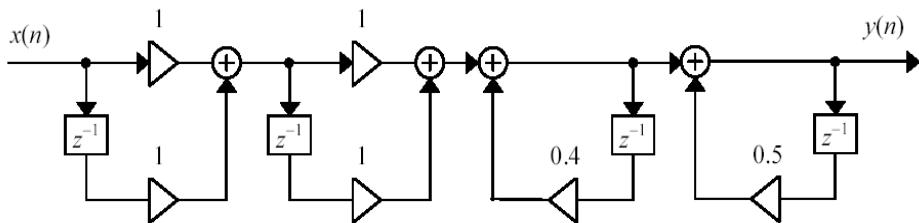
Schematy blokowe układów dyskretnych, realizujących różne zapisy transmitancji $H(z)$	
$H^{(1)}(z) = \frac{1 + 2z^{-1} + z^{-2}}{1 - 0,9z^{-1} + 0,2z^{-2}}$	$H^{(2)}(z) = 5 + \frac{-4 + 6,5z^{-1}}{1 - 0,9z^{-1} + 0,2z^{-2}}$
$H^{(3)}(z) = 5 - \frac{49}{1 - 0,4z^{-1}} - \frac{45}{1 - 0,5z^{-1}}$	



$$H^{(4)}(z) = \left(\frac{1+z^{-1}}{1-0,4z^{-1}} \right) \left(\frac{1+z^{-1}}{1-0,5z^{-1}} \right)$$



$$H^{(5)}(z) = (1+z^{-1})(1+z^{-1}) \left(\frac{1}{1-0,4z^{-1}} \right) \left(\frac{1}{1-0,5z^{-1}} \right)$$



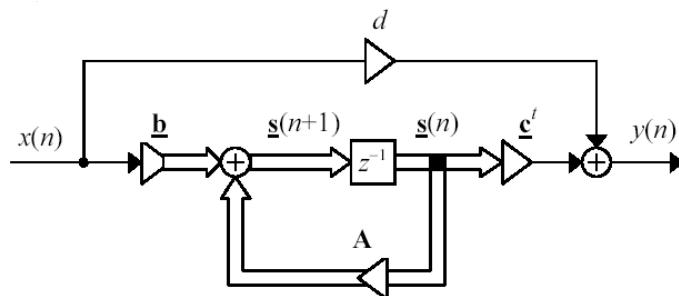
Struktura zmiennych stanu

Filtry cyfrowe mogą być także zapisywane za pomocą równań **zmiennych stanu**. Przykładowo:

Schemat blokowy	Równania zmiennych stanu
	$\begin{cases} s_1(n+1) = -a_1 s_1(n) - a_2 s_2(n) + x(n) \\ s_2(n+1) = s_1(n) \\ y(n) = c_1 s_1(n) + c_2 s_2(n) + d x(n) \end{cases}$

Zmiennymi stanu są sygnały $s_1(n)$ i $s_2(n)$. Zależą one od sygnału wejściowego $x(n)$, same zaś determinują sygnał wyjściowy $y(n)$.

Sygnały wewnętrzne filtra $s_k(n)$ są nazywane zmiennymi stanu, ponieważ podczas jego pracy zmieniają swój stan. Na podstawie rekursywnego filtru cyfrowego drugiego rzędu w postaci macierzowej schematu blokowego:



Sporządzono poniższy układ równań zmiennych stanu:

$$\begin{cases} \underline{s}(n+1) = \mathbf{A}\underline{s}(n) + \underline{b}x(n) \\ y(n) = \underline{c}^T \underline{s}(n) + dx(n) \end{cases}$$

$$\mathbf{A} = \begin{bmatrix} -a_1 & -a_2 \\ 1 & 0 \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \underline{c}^T = [c_1 \quad c_2], \quad \underline{s}(n) = \begin{bmatrix} s_1(n) \\ s_2(n) \end{bmatrix}$$

Dla układu trzeciego rzędu definicje przyjęłyby postać:

$$\mathbf{A} = \begin{bmatrix} -a_1 & -a_2 & -a_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \underline{c}^T = [c_1 \quad c_2 \quad c_3], \quad \underline{s}(n) = \begin{bmatrix} s_1(n) \\ s_2(n) \\ s_3(n) \end{bmatrix}$$

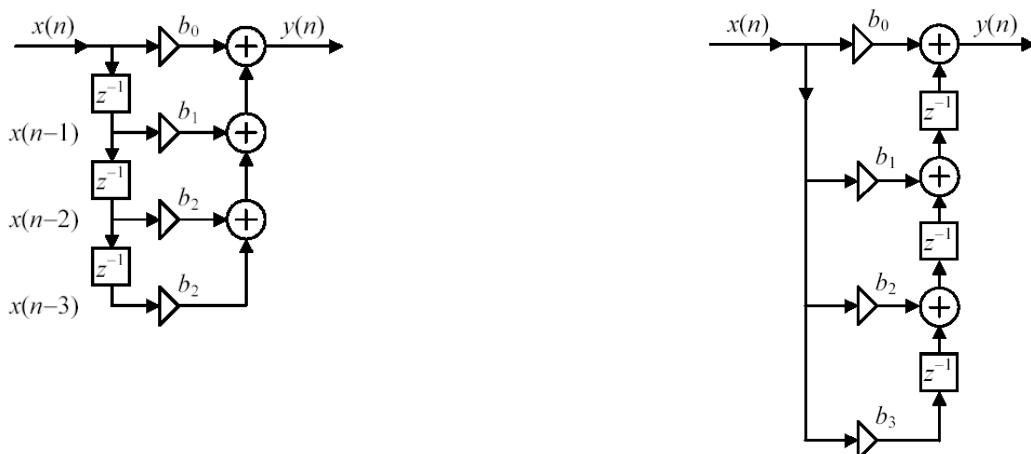
Układ ten ma transmitancję $H(z)$ daną wzorem:

$$H(z) = d + \sum_{n=1}^{\infty} \underline{c}^T \mathbf{A}^{n-1} \underline{b} z^{-n} = d + z^{-1} \underline{c}^T \left[\sum_{k=0}^{\infty} \mathbf{A}^k z^{-k} \right] \underline{b} = d + z^{-1} \underline{c}^T [\mathbf{zI} - \mathbf{A}]^{-1} \underline{b}$$

I - macierz diagonalna jednostkowa (jedynki na przekątnej głównej)

Filtr cyfrowy może występować w wielu różnych postaciach klasycznych, jednak filtr pracuje też w bardziej złożonych układach przetwarzania, na przykład **interpolującą - decymującą**.

W układzie **decymatora cyfrowego** jest korzystniej pod względem obliczeniowym stosować filtr FIR pracujący w układzie interpolatora zaś - jego transpozycji:



Z kolei bardzo ważny zapis filtra cyfrowego w strukturze zmiennych stanu wykorzystuje zmodyfikowany schemat „pełny IIR/FIR” typu II.

Różne struktury pracy filtru wybiera się także ze względu na przyjętą sprzętową arytmetykę obliczeń i jej ograniczenia. Przykładowo w procesorach stałoprzecinkowych (ang. *fixed-point*) należy unikać przepełnień, czyli są istotne występujące w nich maksymalne przejściowe wyniki obliczeń. W procesorach zmiennoprzecinkowych (ang. *floating-point*) należy natomiast dodawać sygnały o zbliżonych wartościach, z powodu ograniczonej liczby bitów mantysy.

Projektowanie rekursywnych filtrów cyfrowych

Istnieją dwie grupy metod: projektowanie bezpośrednie filtrów cyfrowych oraz projektowanie pośrednie, polegające na przekształceniu filtrów analogowych do postaci cyfrowej.

W pierwszym przypadku dobiera się współczynniki wielomianów transmitancji filtru cyfrowego minimalizując błąd średniokwadratowy aproksymacji zadanej charakterystyki częstotliwościowej.

Tak jest na przykład w metodzie **Yule'a-Walkera**. W drugim przypadku natomiast idzie się „na skróty” i wykorzystuje się umiejętność projektowania filtrów analogowych, odpowiednio je przekształcając na filtry cyfrowe. Stosuje się wówczas metody:

- niezmienności odpowiedzi impulsowej,
- dopasowanej transformacji Z,
- transformacji biliniowej.

Wymagania stawiane filtrom cyfrowym

Filtrom cyfrowym stawiane są określone wymagania częstotliwościowe. Zazwyczaj specyfikuje się je dla charakterystyki amplitudowo-częstotliwościowej filtru $H(e^{j\Omega})$, gdzie pulsacja unormowana Ω jest równa $2\pi f/f_{pr}$ (f – częstotliwość, f_{pr} – częstotliwość próbkowania) i zawiera się w przedziale $[-\pi, \pi]$. Definiuje się cztery główne typy filtrów cyfrowych ze względu na zakres (pasmo) przenoszonych częstotliwości:

1) *filtre dolnoprzepustowe LP* (ang. LowPass):

$$|H_{LP}(e^{j\Omega})| = \begin{cases} 1 \pm \delta_{pass} & \text{dla } |\Omega| \leq \Omega_{pass} \\ 0 + \delta_{stop} & \text{dla } |\Omega| \geq \Omega_{stop} \end{cases}$$

3) *filtre pasmowoprzepustowe BP* (ang. BandPass)

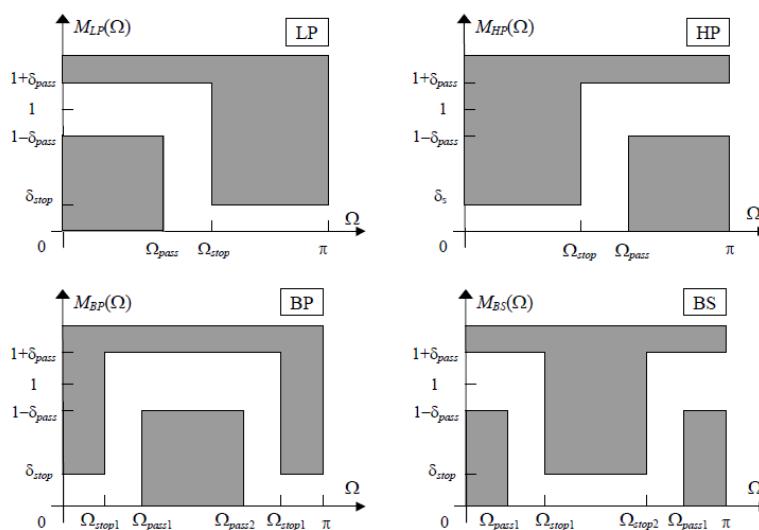
$$|H_{BP}(e^{j\Omega})| = \begin{cases} 1 \pm \delta_{pass} & \text{dla } \Omega_{pass1} \leq |\Omega| \leq \Omega_{pass2} \\ 0 + \delta_{stop} & \text{dla } |\Omega| \leq \Omega_{stop1} \text{ lub } |\Omega| \geq \Omega_{stop2} \end{cases}$$

2) *filtre górnoprzepustowe HP* (ang. HighPass):

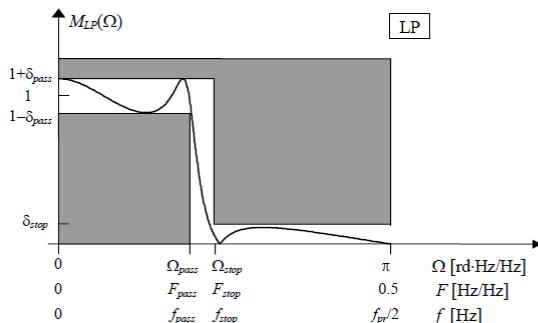
$$|H_{HP}(e^{j\Omega})| = \begin{cases} 0 + \delta_{stop} & \text{dla } |\Omega| \leq \Omega_{stop} \\ 1 \pm \delta_{pass} & \text{dla } |\Omega| \geq \Omega_{pass} \end{cases}$$

4) *filtre pasmowozaporowe BS* (ang. BandStop)

$$|H_{BS}(e^{j\Omega})| = \begin{cases} 0 + \delta_{stop} & \text{dla } \Omega_{stop1} \leq |\Omega| \leq \Omega_{stop2} \\ 1 \pm \delta_{pass} & \text{dla } |\Omega| \leq \Omega_{pass1} \text{ lub } |\Omega| \geq \Omega_{pass2} \end{cases}$$



Rys. 11.2. Orientacyjny kształt charakterystyk amplitudowo-częstotliwościowych $M(\Omega) = |H(e^{j\Omega})|$ (białe „tuneli”) dla czterech podstawowych typów filtrów cyfrowych: dolnoprzepustowego (LP), górnoprzepustowego (HP), pasmowoprzepustowego (BP) i pasmowozaporowego (BS). Ω oznacza pulsację unormowaną równą $\Omega = 2\pi F = 2\pi(f/f_{pr})$, gdzie f oznacza częstotliwość wyrażoną hercach



Rys. 11.1. Orientacyjny kształt charakterystyki amplitudowo-częstotliwościowej $M(\Omega) = |H(e^{j\Omega})|$ (biały „tunel”) dla cyfrowego filtra dolnoprzepustowego (LP). F oznacza częstotliwość unormowaną ($F = f/f_p$), Ω – pulsację unormowaną ($\Omega = 2\pi F$). f_p zaś jest częstotliwością wyrażoną w hercach

Metoda Yule'a-Walkera

Założymy, że realizowany filtr rekursywny IIR ma budowę kaskadową:

$$H(z) = \frac{\sum_{m=0}^M b_m z^{-m}}{1 + \sum_{n=1}^N a_n z^{-n}} = g \prod_{i=1}^L \left(\frac{1 + b_{1i} z^{-1} + b_{2i} z^{-2}}{1 + a_{1i} z^{-1} + a_{2i} z^{-2}} \right)$$

czyli składa się z L sekcji bikwadratowych (transmitancji „elementarnych”, mających wielomiany drugiego rzędu zarówno w liczniku jak i mianowniku). Oznaczmy przez $H_p(e^{j\Omega})$ projektowaną (wymaganą) charakterystykę częstotliwościową filtra cyfrowego i założymy, że jest ona zadana dla K dyskretnych częstotliwości Ω_k , $k = 1, 2, 3, \dots, K$. Zdefiniujmy minimalizowaną średniokwadratową funkcję błędu jako: Nieznane współczynniki filtra $\{g, a_{1i}, a_{2i}, b_{1i}, b_{2i}\}$, $i = 1, 2, \dots, L$ znajduje się wyznaczając pochodne cząstkowe błędu aproksymacji E względem tych współczynników i przyrównując je do zera. Uzyskuje się w ten sposób układ $4L+1$ równań z $4L+1$ niewiadomymi. Jedną z metod rozwiązywania tego układu równań jest

$$E = \sum_{k=1}^K \left| H_p(e^{j\Omega_k}) - H(e^{j\Omega_k}) \right|^2$$

tzw. metoda Flechera-Powella.

Do projektowania filtrów według powyższego scenariusza służy funkcja *yulewalk()* w programie Matlab.

Metoda niezmienności odpowiedzi impulsowej

Założymy, że odpowiedź impulsowa filtra cyfrowego $h_c(n)$ jest spróbowaną odpowiedzią impulsową odpowiedniego filtra analogowego $h_a(t)$:

Mnożenie $h_a(t)$ przez okres próbkowania T ma na celu uniezależnienie wzmacniania filtra cyfrowego od tego okresu. Niech transmitancja filtra analogowego ma postać (5.43):

$$h_c(n) = T h_a(nT) \quad (11.7)$$

i niech p_k , $k = 1, 2, \dots, N$, oznaczają jej biegony (zera wielomianu mianownika transmitancji). Odpowiedź impulsową filtra (11.8) wyznaczamy za pomocą odwrotnej transformaty Laplace'a:

$$H_a(s) = \frac{b_0 + b_1 s + \dots + b_M s^M}{a_0 + a_1 s + \dots + a_N s^N}, \quad N \geq M \quad (11.8)$$

Na podstawie założenia o niezmienności odpowiedzi impulsowej otrzymujemy z powyższego równania zależność na odpowiedź impulsową filtra cyfrowego:

$$h_a(t) = \begin{cases} c_1 e^{p_1 t} + c_2 e^{p_2 t} + \dots + c_N e^{p_N t}, & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (11.9)$$

Transmitancja $H_c(z)$ odpowiedniego filtra cyfrowego jest równa transformacie Z jego odpowiedzi impulsowej $h_c(n)$, czyli z (11.10), (11.11) na podstawie (10.36) mamy:

$$h_c(n) = \begin{cases} h_c^{(1)}(n) + h_c^{(2)}(n) + \dots + h_c^{(N)}(n), & n \geq 0 \\ 0, & n < 0 \end{cases} \quad (11.10)$$

gdzie

$$h_c^{(k)}(n) = \begin{cases} T c_k e^{p_k n T}, & n \geq 0 \\ 0, & n < 0 \end{cases} \quad (11.11)$$

$$X_k(z) = \frac{c_k}{1 - p_k z^{-1}} \Leftrightarrow x(n) = c_k \cdot (p_k)^n \quad (10.36)$$

$$H_c(z) = \sum_{k=1}^N H_c^{(k)}(z), \quad H_c^{(k)}(z) = \frac{T c_k}{1 - e^{p_k T} z^{-1}} \quad (11.12)$$

Bieguny p_k transmitancji filtra analogowego „przechodzą” w bieguny $e^{p_k T}$ filtru cyfrowego. Ponieważ warunkiem stabilności filtru analogowego jest położenie jego biegunów w lewej półpłaszczyźnie liczb zespolonych, stąd

czyli wynikowy filtr cyfrowy jest stabilny i przyczynowy. Charakterystyka częstotliwościowa filtru cyfrowego zaprojektowanego tą metodą jest równa:

$$\operatorname{Re}(p_k) < 0 \Rightarrow |e^{p_k T}| < 1 \quad (11.13)$$

Na podstawie (4.48), (4.49) mamy:

$$H_c(e^{j\Omega}) = \sum_{n=0}^{\infty} h_c(n) e^{-j\Omega n} = \sum_{n=0}^{\infty} T h_a(nT) e^{-j\Omega n} \quad (11.14a)$$

czyli jest ona równa sumie poprzesuwanych charakterystyk filtru analogowego.

$$H_c(e^{j\Omega}) = \sum_{k=-\infty}^{\infty} H_a\left(j\omega - jk \frac{2\pi}{T}\right) \quad (11.14b)$$

PRZYPOMNIENIE :

(4.48), (4.49) – równomierne próbkowanie sygnałów analogowych przedstawione jako ich wymnożenie z sumą impulsów Diraca:

- Jeżeli ktoś chce i ma ochotę nauczyć się tej metody ze zrozumieniem to polecam rozwiązaną przykład

$$\begin{aligned} X_\delta(j\omega) &= \frac{1}{2\pi} \left[X(j\omega) \otimes \omega_p \sum_{k=-\infty}^{\infty} \delta(\omega - k\omega_p) \right] = \frac{1}{2\pi} \left[\omega_p \sum_{k=-\infty}^{\infty} (X(j\omega) \otimes \delta(\omega - k\omega_p)) \right] = \\ &= \frac{1}{2\pi} \left[\omega_p \sum_{k=-\infty}^{\infty} X(j(\omega - k\omega_p)) \right] \end{aligned} \quad (4.49)$$

ze strony 292 z Zielińskiego.

$$x_\delta(t) = x(t) \sum_{k=-\infty}^{\infty} \delta(t - kT) \quad (4.48)$$

Metoda transformacji biliniowej

Załóżmy, że dla każdego filtru analogowego $H_a(s)$ można skonstruować filtr cyfrowy $H_c(z)$, taki że:
czyli zmianom pulsacji analogowej ω w zakresie $(-\infty, +\infty)$ powinny odpowiadać zmiany pulsacji cyfrowej Ω (unormowanej względem częstotliwości próbkowania: $\Omega = 2\pi f/f_{pr}$) w zakresie $(-\pi, +\pi)$. Niech $z = \phi(s)$, wówczas filtr analogowy i odpowiadający mu filtr cyfrowy są związane zależnością:

$$H_c(e^{j\Omega}) = H_a(j\omega) \quad (11.17)$$

Transformacja $z = \phi(s)$ powinna spełniać następujące warunki:

$$H_c(\phi(s)) = H_a(s) \quad (11.19)$$

1. Oś $j\omega$ powinna być przekształcona na okrąg jednostkowy:
 $\phi(j\omega) = e^{j\Omega}$.
2. Powinna istnieć funkcja odwrotna $s = \phi^{-1}(z)$, tak aby można było zaprojektować filtr cyfrowy na podstawie odpowiedniego filtru analogowego:

$$H_c(z) = H_a(\phi^{-1}(z)) \quad (11.20)$$

3. Jeśli układ analogowy opisany transmitancją $H_a(s)$ jest stabilny, to także układ cyfrowy transmitancji $H_c(\phi(s))$ powinien być stabilny. Jeśli więc bieguny transmitancji układu analogowego leżą w lewej

półpłaszczyźnie ($\operatorname{Re}(s) < 0$), to bieguny odpowiadającego mu układu cyfrowego powinny leżeć wewnątrz okręgu jednostkowego, czyli powinien być spełniony warunek $|\phi(s)| < 1$.

- Jeśli „zerowa” częstotliwość ma być zachowana, to powinien być spełniony warunek $\phi(0) = 1$.

W **transformacji biliniowej** przyjmuje się następującą funkcję przekształcającą.

gdzie T jest okresem próbkowania sygnału dyskretnego. Definicja ta spełnia wszystkie powyższe warunki. Transformacja odwrotna do (11.21) ma postać:

$$z = \phi(s) = \frac{(2/T) + s}{(2/T) - s} = \frac{1 + sT/2}{1 - sT/2} \quad (11.21)$$

Podstawiając w (11.22) $s = j\omega$ i $z = e^{j\Omega}$, otrzymujemy:

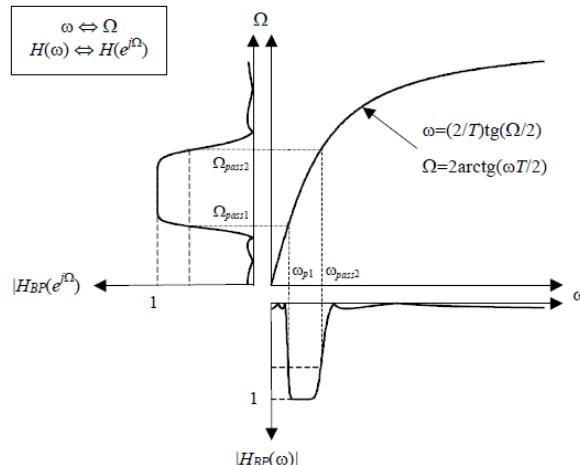
$$s = \phi^{-1}(z) = \frac{2}{T} \frac{z-1}{z+1} \quad (11.22)$$

czyli:

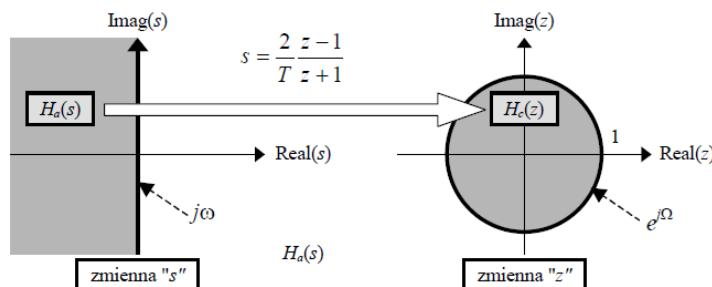
$$j\omega = \frac{2}{T} \frac{e^{j\Omega} - 1}{e^{j\Omega} + 1} = \frac{2}{T} \frac{e^{j\Omega/2} (e^{j\Omega/2} - e^{-j\Omega/2})/2}{e^{j\Omega/2} (e^{j\Omega/2} + e^{-j\Omega/2})/2} = \frac{2}{T} \frac{j \sin(\Omega/2)}{\cos(\Omega/2)}$$

$$\omega = \frac{2}{T} \operatorname{tg}(\Omega/2), \quad \Omega = 2 \operatorname{arctg}(\omega T/2) \quad (11.23a), (11.23b)$$

Kolejność projektowania rekursywnych filtrów cyfrowych metodą transformacji biliniowej jest następująca:



Rys. 11.4. Przykładowa zależność pomiędzy charakterystyką amplitudowo-częstotliwościową pasmowo-przepustowego filtra analogowego $|H_{BP}(\omega)|$ a charakterystyką $|H_{BP}(e^{j\Omega})|$ filtra cyfrowego, jego brata „bliźniaka”, uzyskanego w wyniku transformacji biliniowej ($p1 = \text{pass}$)



Rys. 11.3. Graficzna ilustracja transformacji biliniowej transmitancji $H_a(s)$ filtra analogowego na transmitancję $H_c(z)$ filtra cyfrowego: 1) lewa półpłaszczyzna zmiennej zespolonej „s” transformacji Laplace'a przechodzi we wnętrze okręgu o promieniu jednostkowym na płaszczyźnie zmiennej zespolonej „z” transformacji Z, 2) oś jo przechodzi w okrąg $e^{j\Omega}$, $\Omega = 2\pi f / f_{pr}$

- Podanie zbioru zadanych pulsacji granicznych $\{\Omega_g = 2\pi f_g / f_{pr}\}$ filtru cyfrowego z przedziału $[0, \pi]$ (filtry dolnoprzepustowe, górnoprzepustowe, pasmowoprzepustowe, pasmowozaporowe) oraz wymaganych wartości tłumienia w paśmie przepustowym (A_{pass}) i zaporowym (A_{stop}).
- Wyznaczenie pulsacji ω_g dla filtru analogowego $H_a(s)$, związanego z filtrem cyfrowym $H_c(z)$:
- Zaprojektowanie filtru analogowego, zdefiniowanego przez zbiór parametrów $\{\omega_{g1}, \dots, A_{pass}, A_{stop}\}$.

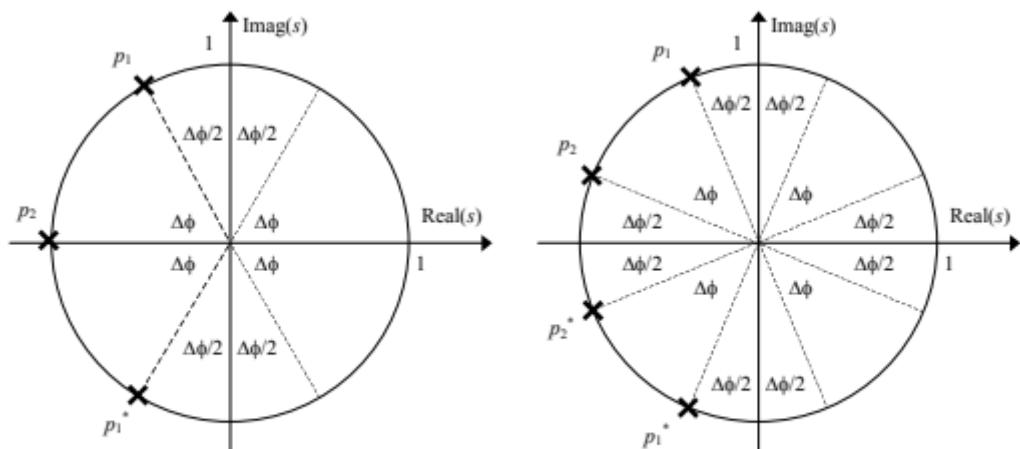
$$\omega_g = \frac{2}{T} \operatorname{tg}\left(\frac{\Omega_g}{2}\right)$$

4. Wyznaczenie współczynników transmitancji filtra cyfrowego na podstawie współczynników transmitancji filtra analogowego:

$$H_c(z) = H_a(s) \Big|_{s=(2/T)(z-1)/(z+1)}$$

Prototypy analogowe:

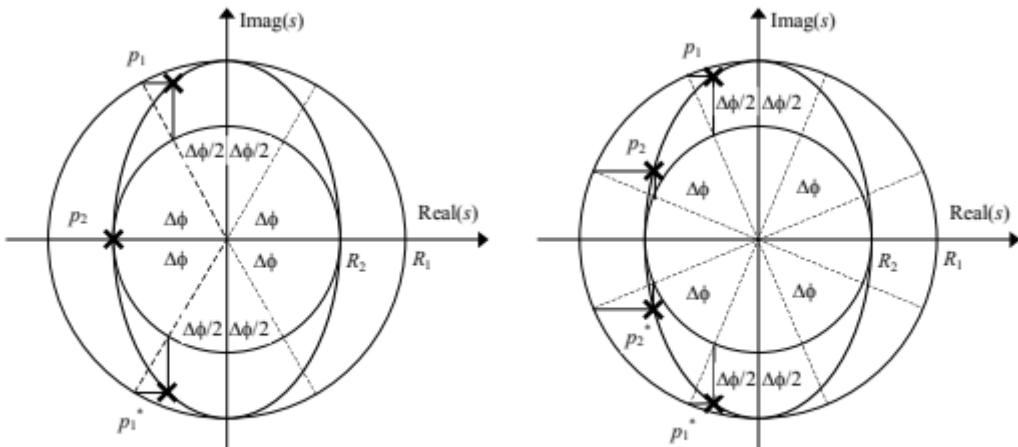
- A. Butterwortha – maksymalnie płaski filtr w pasmie przepustowym o liniowej fazie, Projekt opiera się na rozłożeniu biegunków na okręgu o określonym promieniu i umieszczeniu ich w odległościach $\Delta\phi$ od siebie, zaczynając od kąta $\pi/2 + \Delta\phi/2$ aby być w lewej półpłaszczyźnie i nie zaczynać od osi urojonej. Kąt $\Delta\phi = \pi/N$ gdzie N to rząd filtra określany na podstawie wymogów tłumienia



Rys. 6.5. Położenie biegunków „x” unormowanego ($\omega_0 = 1$ rd/s), dolnoprzepustowego filtra Butterwortha trzeciego i czwartego rzędu ($N = 3$ lub 4)

- B. Czebyszewa I – posiada tężnienia w pasmie przejściowym, konstrukcja opera się na dwóch okręgach o promieniach R_1 i R_2 a następnie określeniu biegunków na wpisanej elipsie za pomocą wzoru:

$$p_k = \operatorname{Re}\{P_k^{(R1)}\} + j \operatorname{Im}\{P_k^{(R2)}\}$$



Rys. 6.8b. Graficzna ilustracja wyznaczania położenia biegunków „x” dolnoprzepustowego filtra Czebyszewa typu I trzeciego i czwartego rzędu ($N = 3$ lub 4)

- C. Czebyszewa II – posiada tężnienia w pasmie zaporowym, zera i biegunki określa się jako odwrotności zer i biegunków otrzymanych dla typu I
D. Eliptyczny – posiada najbardziej nieliniową charakterystykę, zafalowania w obu pasmach ale najbardziej strome pasmo przejściowe.

Wszystkie prototypy analogowe projektuje się na podstawie prototypów LP, następnie za pomocą przekształcenia zmiennej s do s' dokonuje się zmiany typu filtra. Taki filtr następnie przekształca się w cyfrowy.

Projektowanie nierekursywnych filtrów cyfrowych

Filtracja nierekursywna określona jest tylko za pomocą licznika transmitancji, a więc nie występuje sprzężenie zwrotne. Algorytm filtracji nierekursywnej opisany jest splotem sygnału wejściowego i odpowiedzi impulsowej filtra.

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

$x(n)$ – sygnał wejściowy, $h(n)$ – odp. impulsowa filtra, $y(n)$ – wyjście z filtra, b_m – współczynniki (kolejne wartości odp. Impulsowej). Stosuje się sumowanie od 0 ze względu na przyczynowość układu. Sumowanie do nieskończoności ogranicza się do praktycznej implementacji dla pierwszych N próbek odpowiedzi. Odpowiedź impulsowa i sygnały posiadają swoje widmo częstotliwościowe, co pozwala splot w czasie zamienić na iloczyn w częstotliwości. W przypadku sygnałów cyfrowych posługujemy się pulsacją unormowaną względem częstotliwości próbkowania ($-f_{pr}/2$ do $f_{pr}/2$) $\rightarrow (-\pi, \pi)$. Współczynniki dobieramy tak aby widmo częstotliwościowe stało się odpowiednim filtrem (LP, HP, ...), przyjmuje się że pasmo przepustowe powinno wynosić $H(e^{j\Omega_{pass}}) \approx 1$ a nie przepustowe $H(e^{j\Omega_{stop}}) \approx 0$.

$$Y(e^{j\Omega}) = H(e^{j\Omega})X(e^{j\Omega}) \quad \Omega = 2\pi \frac{f}{f_{pr}} \quad H(e^{j\Omega}) = R(\Omega) + jI(\Omega)$$

Cały proces projektowania sprowadza się do określenia modułu i fazy $H(e^{j\Omega})$, jednak moduł nie jest analityczny (nieciągłości), dlatego przyjmuje się funkcję $A(\Omega)$.

$$M(\Omega) = |H(e^{j\Omega})| = \sqrt{R^2(\Omega) + I^2(\Omega)} \quad A(\Omega) = \pm M(\Omega) \quad H(e^{j\Omega}) = A(\Omega)e^{j\Psi(\Omega)}$$

Podstawową zaletą filtrów nierekursywnych jest łatwość uzyskania liniowej charakterystyki fazowej filtra, jednak dla skończonego sumowania liniowość ta jest spełniona w paśmie przepustowym. Stałe c_1 i c_2 są dowolnymi stałymi rzeczywistymi.

$$\Psi(\Omega) = c_1 + c_2\Omega \quad X(e^{j\Omega}) = \sum_{n=-\infty}^{+\infty} x(n)e^{-j\Omega n}$$

Korzystając z transformaty Fouriera w zależności od liczby próbek odpowiedzi impulsowej oraz jej symetryczności względem próbki środkowej (symetria i asymetria) określa się poniższe typy filtrów dla zadanej $h(n)$. Filtry LP, HP, BS, BP, oraz H – Hilberta, D – różniczkujący.

Typ	Właściwości $h(n)$ (a)symetria / długość	Okres	Ω punktów symetrii	Ω punktów asymetrii	Rodzaje filtrów
I	$h(n) = h(N-1-n)$, $N = 2L+1$	2π	$0, \pi$	—	LP, HP, BP, BS
II	$h(n) = h(N-1-n)$, $N = 2L$	4π	$0, 2\pi$	$\pi, 3\pi$	LP, BP
III	$h(n) = -h(N-1-n)$, $N = 2L+1$	2π	—	$0, \pi$	BP, H, D
IV	$h(n) = -h(N-1-n)$, $N = 2L$	4π	$\pi, 3\pi$	$0, 2\pi$	HP, BP, H, D

Metoda próbkowania w dziedzinie częstotliwości

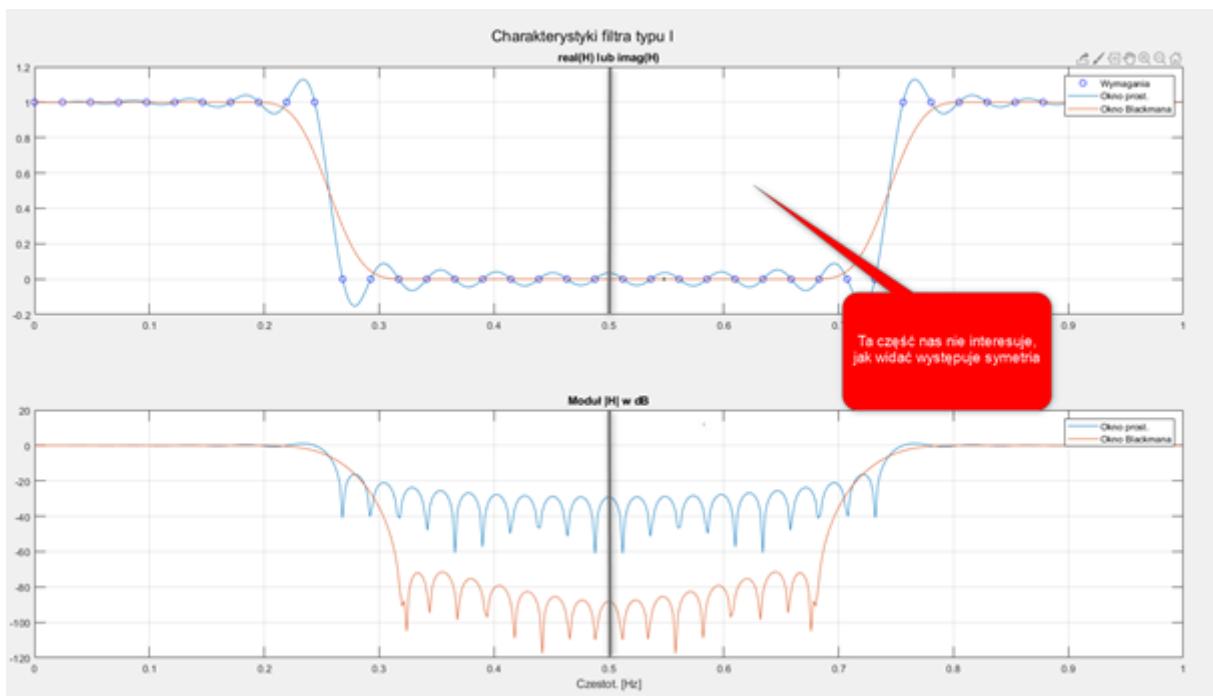
Metoda ta polega na zadawaniu w dziedzinie częstotliwości „próbek” odpowiedzi częstotliwościowej $H(e^{j\Omega})$ dla N pulsacji unormowanych $= k(2\pi/N)$, gdzie k zmienia się od 0 do $N-1$. W matlabie metoda ta kryje się pod poleceniem **FIR2**. Odpowiedź impulsową znajduje się stosując IFFT.

$$h(n) = \sum_{k=0}^{N-1} H(k) e^{\left(\frac{2\pi}{N} k\right)n}, \quad n = 0, 1, 2, \dots, N-1$$

Nie każdy rodzaj filtra można zaprojektować z dowolnego prototypu (I-IV). Przykładowo dla rzeczywistej odpowiedzi $h(n)$ dla parzystych wartości n trzeba umieścić zadane wartości charakterystyki amplitudowo-często. tylko w części rzeczywistej $H(e^{j\Omega})$ (I i II, symetryczne $h(n)$), dla asymetrycznej rzeczywistej $h(n)$ należy umieścić je w części urojonej w nieparzystych n (III i IV, asymetryczne $h(n)$). Odpowiedź częstotliwościowa jest interpolowana pomiędzy prążkami $H(k)$ funkcją (pomijam wyprowadzenie) SINC a więc posiada charakter oscylacyjny pomiędzy prążkami.

$$e^{-j\Omega(N-1)/2} e^{-j\pi k/N} \frac{\sin(\Omega N/2)}{\sin(\Omega/2 - \pi k/N)}$$

Aby zmniejszyć oscylacje należy zdać prążki w paśmie przejściowym filtra lub pomnożyć odpowiedź impulsową przez okno czasowe $w(n)$ co powoduje splot widma filtra i okna czyli redukcje oscylacji.



Metoda optymalizacji średniokwadratowej

W tym podejściu stosuje się optymalizację (aproksymację) średniokwadratową zadanej charakterystyki. Chcemy uzyskać takie wagę filtru $h(n)$ aby jego odpowiedź $H(e^{j\Omega})$ aproksymowała charakterystykę zadaną: $H_p(e^{j\Omega})$ w wybranych L punktach częstości unormowanej.

$$E = \sum_{k=0}^{L-1} \left| H(e^{j\Omega_k}) - H_p(e^{j\Omega_k}) \right|^2 \quad E = \sum_{k=0}^{L-1} \left| A(\Omega_k) - A_p(\Omega_k) \right|^2$$

Dla filtrów typu I-IV można zapisać błąd z wykorzystaniem funkcji $A(\Omega)$. Zakłada się linowe próbkowanie pulsacji, z twierdzenia Parsevala o zachowaniu energii oznacza to, że zminimalizowany jest także błąd aproksymacji $h(n)$. Jeśli wymagania określmy w L punktach a poszukujemy odpowiedzi N -punktowej to „sytuacja” się zmienia. Przykładowo dla $N=2*M+1$:

$$E = \sum_{n=-M}^{M} \left| h(n) - h_p(n) \right|^2 + 2 \sum_{n=M+1}^{(L-1)/2} \left| h_p(n) \right|^2$$

Błąd E osiąga minimum, gdy pierwszy składnik osiąga minimum. W przypadku nierównomiernego próbkowania pulsacji metodologia jest inna, a stosuje się wzory I i IV gdzie pulsacja jest dowolna, określamy $A(\Omega_k)$ dla L-1 pulsacji i otrzymujemy układ równań z M lub M+1 niewiadomych.

I) *symetryczna odpowiedź impulsowa o długości nieparzystej* ($N = 2L+1$, $M = L$):

$$H(e^{j\Omega}) = e^{-j\Omega M} A(\Omega) = e^{-j\Omega M} \left(\sum_{n=0}^{M-1} 2h(n) \cos(\Omega(M-n)) + h(M) \right)$$

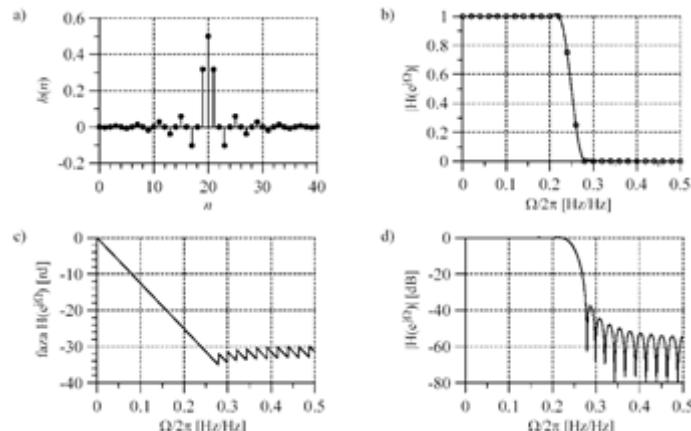
IV) *asymetryczna odpowiedź impulsowa o długości parzystej* ($N = 2L$, $M = L-1/2$):

$$H(e^{j\Omega}) = j e^{-j\Omega M} A(\Omega) = j e^{-j\Omega M} \left(\sum_{n=0}^{N/2-1} 2h(n) \sin(\Omega(M-n)) \right)$$

Podejście to sprowadza się do równania macierzowego na odpowiedź impulsową \mathbf{h} gdzie \mathbf{F} jest macierzą wartości funkcji cosinus (wymiary $L \times (M+1)$), a \mathbf{a} jest L-elementowym wektorem zdanych wartości $A(\Omega)$. Gdy zastosujemy wektor wag otrzymujemy zmodyfikowany wzór. Wagi wpływają na wartość błędu (dodatnio określona macierz diagonalna).

$$\mathbf{h} = (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T \mathbf{a}_p = \text{pinv}(\mathbf{F}) \mathbf{a}_p \quad \mathbf{h} = (\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{W} \mathbf{a}_p$$

Przykład filtra dolnoprzepustowego typu I zaprojektowanego metodą optymalizacji średniokwadratowej bez stosowania wag dla N=41.



Metoda optymalizacji Czebyszewa (alg. Remeza)

Jeden z najczęściej stosowanych algorytmów projektowania filtrów cyfrowych nierekursywnych o liniowej fazie. Wykorzystuje się implementację algorytmu Remeza. W matlabie funkcja **FIRPM**. Polega ona na aproksymacji charakterystyki amplitudowej $A_p(\Omega)$ za pomocą sumy kosinusoid pomnożonych przez wagi c_n . Minimalizuje się błąd $E(\Omega)$, a wagi określa się w wektorze $\mathbf{W}(\Omega)$. Algorytm opiera się na twierdzeniu, że istnieje zbiór $M+2$ pulsacji Ω ze funkcją błędu przyjmuje dla tych pulsacji tylko wartości $+\varepsilon$ i że to są jej ekstremalne wartości w zbiorze pulsacji. Zakładamy ze Ω_m są znane, wówczas otrzymujemy układ $M+2$ równań z $M+2$ niewiadomymi.

$$E(\Omega) = W(\Omega) \left[\sum_{n=0}^M c_n \cos(\Omega n) - A_p(\Omega) \right]$$

$$W(\Omega_m) \left[\sum_{n=0}^M c_n \cos(\Omega_m n) - A_p(\Omega_m) \right] = (-1)^m \varepsilon, \quad m = 1, 2, \dots, M+2$$

$$\begin{bmatrix} 1 & \cos(\Omega_1) & \cdots & \cos(M\Omega_1) & 1/W(\Omega_1) \\ 1 & \cos(\Omega_2) & \cdots & \cos(M\Omega_2) & -1/W(\Omega_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \cos(\Omega_{M+1}) & \cdots & \cos(M\Omega_{M+1}) & (-1)^M/W(\Omega_{M+1}) \\ 1 & \cos(\Omega_{M+2}) & \cdots & \cos(M\Omega_{M+2}) & (-1)^{M+1}/W(\Omega_{M+1}) \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_M \\ \varepsilon \end{bmatrix} = \begin{bmatrix} A_p(\Omega_1) \\ A_p(\Omega_2) \\ \vdots \\ A_p(\Omega_{M+1}) \\ A_p(\Omega_{M+2}) \end{bmatrix}$$

Wyznaczamy z niego $M+1$ współczynników c_k oraz amplitudę oscylacji ε . Jednak pulsacje Ω_m nie są znane, dlatego algorytm ten polega na iteracyjnym poszukiwaniu ich, a następnie rozwiązyaniu układu równan. Kroki algorytmu:

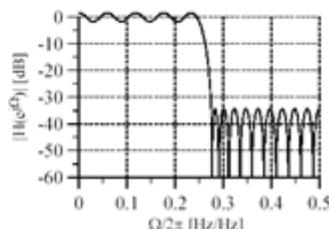
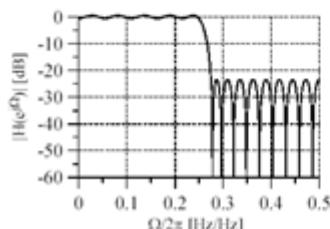
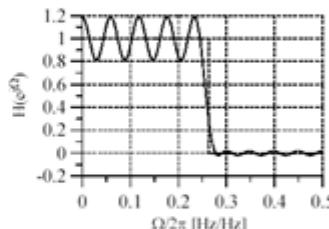
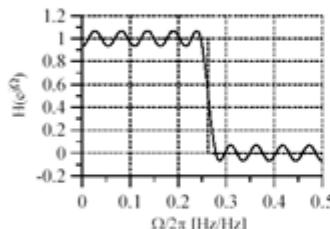
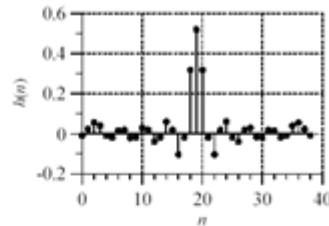
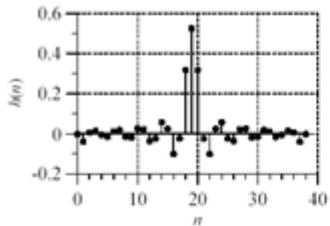
1. Przyjęcie zbioru pulsacji Ω_m ,
2. Rozwiążaniu układu równan, otrzymanie wartości c_n i ε ,
3. Sprawdzenie czy amplituda oscylacji funkcji błędu $E(\Omega)$ dla przedziału $\Omega [0, \pi]$ jest większa niż obliczone ε ; jeśli jest mniejsza to koniec;
4. Jeśli większa to wyznaczenie $M+2$ pulsacji, dla których funkcja błędu E ma ekstrema i przyjęcie ich za nowe Ω_m i powrót do pkt. 2.

Wyznaczenie nowych wartości ekstremalnych E może być wykonane techniką interpolacji Lagrange'a. Przykładowo dla filtry symetrycznego typu I o nieparzystej liczbie próbek używamy wzoru:

$$A(\Omega) = \sum_{n=0}^{M-1} 2h(n) \cos(\Omega(M-n)) + h(M) = \sum_{n=0}^M h_n \cos(\Omega(M-n)) = \sum_{n=0}^M h_{M-n} \cos(\Omega n)$$

Można zauważyć, że próbki o indeksach n i $M-n$ są równe, a więc wagi filtra można uprościć, a po wyznaczeniu współczynników c_n odpowiedź impulsową otrzymamy z zależności $c_n = h_{M-n}$.

$$h_n = \begin{cases} 2h(n), & n = 0, 1, \dots, M-1 \\ h(n), & n = M \end{cases} \quad \left\{ \frac{c_M}{2}, \dots, \frac{c_2}{2}, \frac{c_1}{2}, c_0, \frac{c_1}{2}, \frac{c_2}{2}, \dots, \frac{c_M}{2} \right\}$$



Przykład charakterystyk:

Strona lewa: brak różnicowania pasma przepustowego i zaporowego (wagi: $w_{pass} = 1$, $w_{stop} = 1$), strona prawa - przyznanie większego priorytetu pasmu zaporowemu (wagi: $w_{pass} = 1$, $w_{stop} = 10$). Długość filtrów $N = 39$

Metoda okien

Bardzo prosta metoda dająca dobre efekty, szeroko stosowana. W matlabie funkcja **FIR1**. Składa się z następujących kroków:

1. Wybór rodzaju filtra (LP, ...) i jego pulsacji granicznych czyli określenie $H(e^{j\Omega})$,
2. Analityczne wyznaczenie wzoru na dyskretną odpowiedź impulsową filtra $h(n)$, za pomocą przekształcenia odwrotnego Fouriera $H(e^{j\Omega})$ (podstawowe są wyprowadzone w książce):

$$h(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\Omega}) e^{j\Omega n} d\Omega, \quad -\infty \leq n \leq \infty$$

3. Wymnożenie obliczonego $h(n)$ (nieskończone oscylacje) z wybraną funkcją okna czasowego $w(n)$ o skończonej liczbie M próbek niezerowych

$$h_w(n) = h(n)w(n), \quad -\infty \leq n \leq \infty$$

4. Przesunięcie $h_w(n)$ w prawo o M próbek i pobrania $2*M+1$ próbek:

$$h_w^M(n) = h_w(n-M), \quad 0 \leq n \leq 2M$$

5. Sprawdzenie zgodności uzyskanej charakterystyki częstotliwościowej zaprojektowanego $M_w^{(M)}$ Współczynnik K służy do wyboru długości filtra N o odpowiedniej stromości pasm przejściowych:

$$N > \frac{K}{\Delta f / f_{pr}} = \frac{K}{\Delta_{ml}}$$

gdzie Δ_{ml} – unormowana szerokość ‘listka’ głównego widma okna,

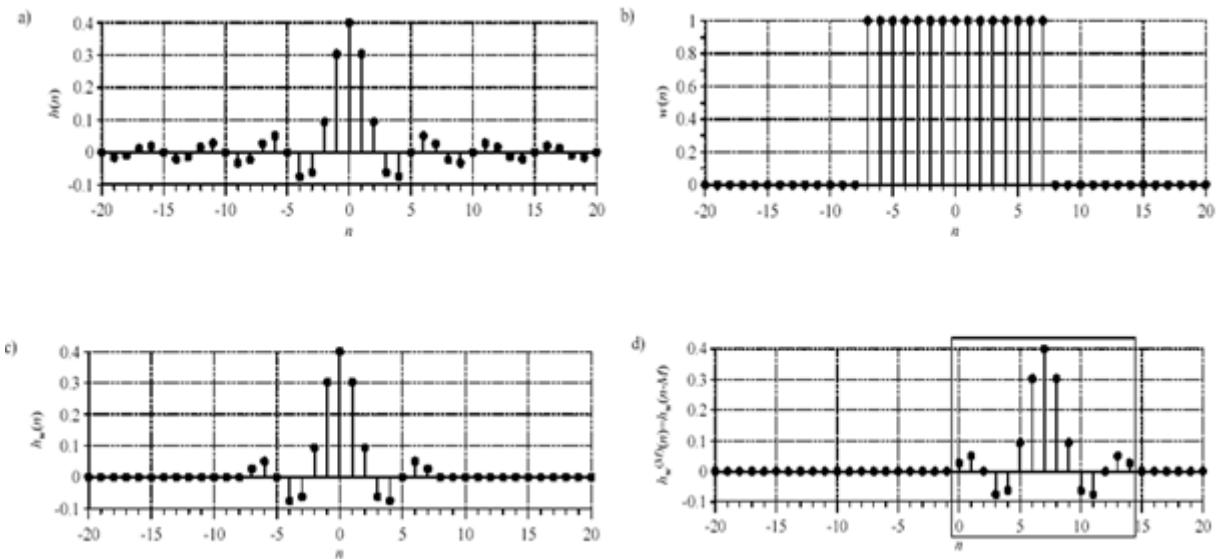
Definicje różnych okien czasowych $w(n)$ o długości $N=2*M+1$, A_{sl} – względny poziom tłumienia listków bocznych do głównego, β – parametr okna Kasiera, $\Delta_{ml} = \Delta f / f_{pr}$ – jak wyżej.

Okno	Definicja $w(n)$, $-M \leq n \leq M$	A_{sl}	$\Delta_{ml} = \Delta f / f_{pr}$ [Hz/Hz]
Prostokątne	1	13,5 dB	2/(2M+1)
Bartletta	$1- n /(M+1)$	27,0 dB	4/(2M+1)
Hanninga	$0,5+0,5\cos(2\pi n/(2M+1))$	31,0 dB	4/(2M+1)
Hamminga	$0,54+0,46\cos(2\pi n/(2M+1))$	41,0 dB	4/(2M+1)
Blackmana	$0,42+0,5\cos(2\pi n/(2M+1))+0,08\cos(4\pi n/(2M+1))$	57,0 dB	6/(2M+1)
Kaisera	$I_0[\beta(1-(n/M)^2)^{0.5}] / I_0(\beta)$, $I_0(\beta)$ – funkcja Bessela	dowolne	dowolne

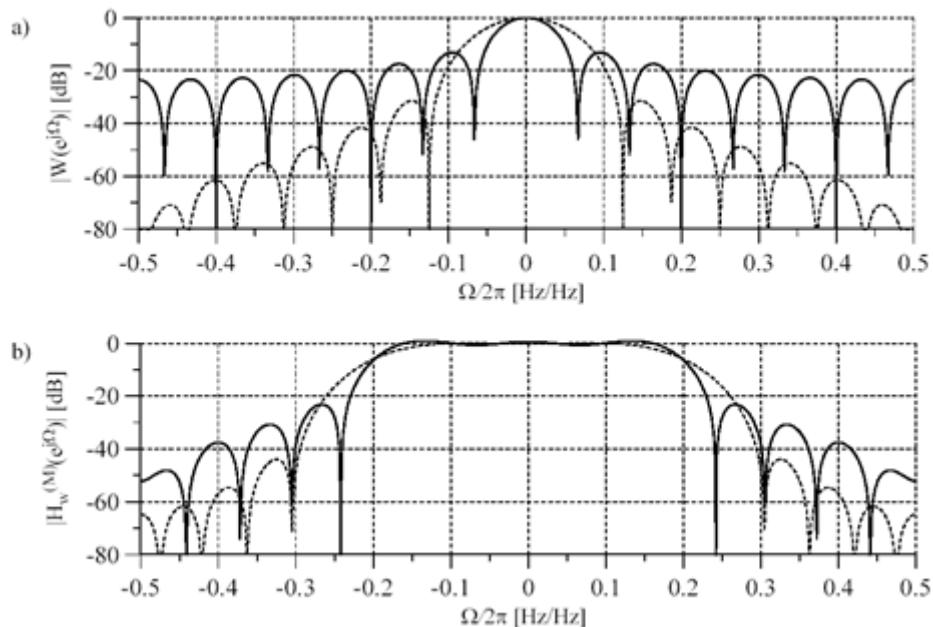
Okno	Tłumienie filtra [*])	Współczynnik K
Prostokątne	21 dB	2,00
Trójkątne (Bartletta)	25 dB	4,00
Hanninga	44 dB	4,00
Hamminga	53 dB	4,00
Blackmana	74 dB	6,00
Kaisera ($\beta = 2,12$)	30 dB	1,54
Kaisera ($\beta = 4,54$)	50 dB	2,93
Kaisera ($\beta = 7,76$)	70 dB	4,32
Kaisera ($\beta = 8,96$)	90 dB	5,71

Metodologia konstrukcji.

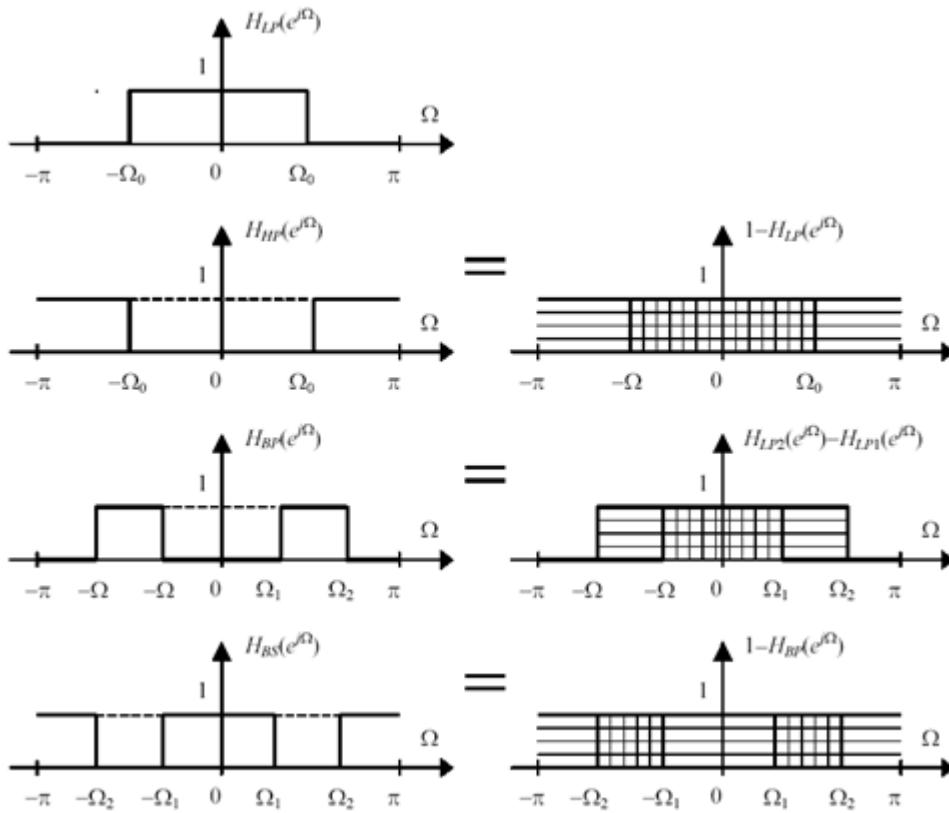
Przykładowa ilustracja graficzna „konstrukcji” odpowiedzi impulsowej $h_w(M)$ filtra LP, $f_0 = 200$ Hz, $f_{pr} = 1000$ Hz) w przypadku zastosowania okna prostokątnego ($M = 7$, $N = 2M+1 = 15$): a) $h(n)$, b) $w(n)$, c) $h_w(n)$, d) $h_w^{(M)}(n)$



Unormowane moduły widma Fouriera ($20\log_{10}|W(e^{j\Omega})/W(0)|$) okna prostokątnego oraz widma okna Hanninga (a) oraz charakterystyki amplitudowo-częstotliwościowe $|H_w^{(M)}(e^{j\Omega})|$



Dla idealnych charakterystyk prostokątnych wzory na odpowiedzi impulsowe są wyznaczone analitycznie. Polega to na znajomości wzoru na filtr LP i konstrukcji pozostałych na podstawie różnic delta Kroneckera i innych filtrów.



Wyprowadzenia pomijam (wiedza zbędna), Wartości i ich zakresy: $0 < F < 1$, $0 < \Omega < 2\pi$, $\Omega_k = 2\pi f_k / f_{pr}$, $k=0,1,2$ (zależy ile pasm granicznych)

Rodzaj filtra	$h(n), -\infty < n < \infty$
Dolnoprzepustowy (LP)	$h_{LP}(n) = \begin{cases} 2F_0 \frac{\sin(\Omega_0 n)}{\Omega_0 n}, & n \neq 0 \\ 2F_0, & n = 0 \end{cases}$
Górnoprzepustowy (HP)	$h_{HP}(n) = \begin{cases} \delta(n) - 2F_0 \frac{\sin(\Omega_0 n)}{\Omega_0 n}, & n \neq 0 \\ 1 - 2F_0, & n = 0 \end{cases}$
Pasmowoprzepustowy (BP)	$h_{BP}(n) = \begin{cases} 2F_2 \frac{\sin(\Omega_2 n)}{\Omega_2 n} - 2F_1 \frac{\sin(\Omega_1 n)}{\Omega_1 n}, & n \neq 0 \\ 2(F_2 - F_1), & n = 0 \end{cases}$
Pasmowozaporowy (BS)	$h_{BS}(n) = \begin{cases} \delta(n) - \left(2F_2 \frac{\sin(\Omega_2 n)}{\Omega_2 n} - 2F_1 \frac{\sin(\Omega_1 n)}{\Omega_1 n} \right), & n \neq 0 \\ 1 - 2(F_2 - F_1), & n = 0 \end{cases}$

Algorytmy filtracji cyfrowej

Jedną z podstawowych i najstarszych metod analizy sygnałów jest analiza częstotliwościowa. Przedstawia się w niej badany sygnał $x(t)$ jako sumę sygnałów elementarnych (bazowych) $g_k(t)$:

$$x(t) = \sum_k a_k g_k(t)$$

gdzie

$$a_k = \int x(t) \gamma_k^*(t) dt \quad (1)$$

i $\gamma_k(t)$ jest sygnałem dualnym do $g_k(t)$, a znak „*” oznacza sprzężenie zespolone.

Jako sygnały bazowe stosuje się funkcje Fouriera, Bessela, Czebyszewa i Haara oraz wielomiany Czebyszewa, Hermite'a, Legendre'a i Laguerre'a.

Każdy z sygnałów bazowych ma inną częstotliwość. Wartości współczynników a_k określają jakie częstotliwości występują w sygnale, a jakie nie. Sygnały bazowe $g_k(t)$ mogą być rzeczywiste lub zespolone. W zależności od pasma częstotliwościowego sygnałów $x(t)$ oraz $g_k(t)$, powyższa suma może być skończona lub nieskończona.

Im kształt funkcji bazowych i ich przesunięcie są lepiej dopasowane do rodzaju rzeczywistych sygnałów składowych obecnych w $x(t)$, tym mniej współczynników a_k ma wartości niezerowe i tym bardziej jest skupione (selektywne) widmo amplitudowe sygnału $S_x(k)$, zdefiniowane jako:

$$S_x(k) = |a_k|^2$$

Funkcje bazowe $g_k(t)$ muszą rozpinać całą przestrzeń sygnałów określonego typu. I można je przedstawić jako sumę sygnałów elementarnych (1).

Ze względu na rodzaj funkcji bazowych $g_k(t)$ rozróżnić można trzy przypadki.

- 1) Jeśli funkcje bazowe są ortonormalne, tzn.

$$\int g_k(t) g_l^*(t) dt = \begin{cases} 1 & \text{dla } k = l \\ 0 & \text{dla } k \neq l \end{cases}$$

$$\gamma_k(t) = g_k(t)$$

i każdy analizowany sygnał $x(t)$ ma unikalne rozwinięcie w tej przestrzeni (reprezentuje go tylko jeden zestaw współczynników $\{a_k\}$).

- 2) Kiedy funkcje bazowe $g_k(t)$ są liniowo niezależne, tzn. żadna z nich nie może być przedstawiona jako liniowa kombinacja pozostałych, wówczas funkcje analizujące $\gamma_k(t)$ są inne niż funkcje syntezujące $g_k(t)$ i wyznacza się je z warunku biortonormalności:

$$\int g_k(t) \gamma_l^*(t) dt = \begin{cases} 1 & \text{dla } k = l \\ 0 & \text{dla } k \neq l \end{cases}$$

Istnieje tylko jeden zbiór $\{\gamma_k(t)\}$. Przyporządkowanie $x(t) \rightarrow \{a_k\}$ jest dalej wzajemnie jednoznaczne.

- 3) Kiedy funkcje bazowe $g_k(t)$ są liniowo zależne, ale rozpinają przestrzeń sygnałów, wówczas istnieje wiele zbiorów funkcji $\gamma_k(t)$ spełniających równanie ww. Zazwyczaj wybiera się ten, w którym funkcje analizujące są najbardziej zbliżone do funkcji syntezujących. Przyporządkowanie $x(t) - \{a_k\}$ nie jest już wzajemnie jednoznaczne i zależy od wyboru zbioru $\{\gamma_k(t)\}$.

Teoria wyboru zbiorów $\{g_k(t)\}$ i projektowania zbiorów $\{\gamma_k(t)\}$ jest nazywana teorią rozpięć (ang. frames).

Funkcje $g_k(t)$ i $\gamma_k(t)$ mogą mieć różnych charakter w zależności od rodzaju analizowanego sygnału.

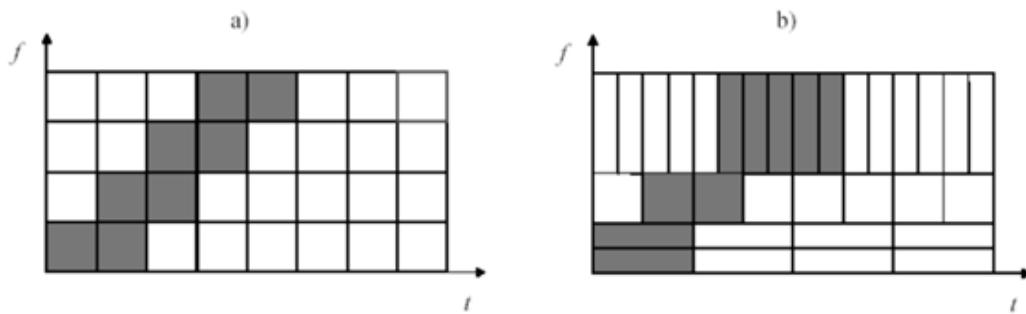
W przypadku sygnałów stacjonarnych (niezmennych w czasie) one także powinny być stacjonarnymi drganiami (oscylacjami) o nieskończonym nośniku, takimi np. jak sygnały (ko)sinusoidalne w bazie fourierowskiej.

Dla sygnałów niestacjonarnych (zmiennych w czasie, impulsowych) funkcje bazowe powinny zaś mieć postać niestacjonarnych oscylacji impulsowych o skończonym nośniku, takich np. jak w transformacie Haara, Gabora lub transformacie falkowej.

W pierwszym przypadku sygnał aproksymowany jest sumą nieskończonych w czasie drgań o różnych częstotliwościach, w przypadku zaś drugim – sumą ograniczonych w czasie przebiegów impulsowych, występujących w różnych chwilach czasowych i mających różne pasmo częstotliwościowe.

Im kształt funkcji bazowych będzie bardziej dopasowany do analizowanego sygnału, tym mniejsza liczba tych funkcji wystarczy do jego aproksymacji.

Dwie podstawowe strategie podziału przestrzeni sygnału i związane z nimi szachownice czasowo-częstotliwościowe:



Podstawowe szachownice dekompozycji czasowo-częstotliwościowej sygnałów: transformacja Gabora - krótkoczasowa transformacja Fouriera (a) oraz transformacja falkowa (b) - Ciemnym kolorem zostały zaznaczone niezerowe współczynniki przykładowej dekompozycji czasowo-częstotliwościowej sygnału z liniowo narastającą częstotliwością.

W przypadku pierwszym komórki szachownicy mają zawsze taką samą wielkość i kształt, w przypadku zaś drugim – taką samą wielkość, lecz różny kształt: dla niskich częstotliwości mają lepszą rozdzielczość częstotliwościową a gorszą czasową, natomiast dla wysokich częstotliwości odwrotnie – gorszą rozdzielczość częstotliwościową a lepszą czasową. Oczywiście sposób dekompozycji powinien być odpowiednio wybrany w zależności od rodzaju sygnału i celu analizy.

Najlepiej żeby struktura szachownicy czasowo-częstotliwościowej była adaptacyjnie dopasowywana do lokalnych cech analizowanego sygnału.

Kryterium jej doboru powinna być minimalizacja niezerowych współczynników dekompozycji sygnału, czyli bardzo dobre dopasowanie kształtu funkcji bazowych do kształtu sygnału oraz chwil ich występowania.

Krótkoczasowa transformacja Fouriera STFT (ang. Short-Time Fourier Transform)

Ciągła, krótkoczasowa transformacja Fouriera STFT (ang. Short-Time Fourier Transform) może być interpretowana jako niezdyskretyzowana w czasie i w częstotliwości transformacja Gabora. Charakteryzuje się ona bardzo dużą redundancją zawartej w niej informacji TF. Podczas analizy i syntezy stosuje się tylko jedno i to samo okno, a nie dwa osobne okna biortogonalne.

Definicja tej transformacji w dziedzinie czasu i częstotliwości jest następująca:

$$STFT_x^T(t, f) = \int_{-\infty}^{+\infty} x(\tau) \gamma^*(\tau - t) e^{-j2\pi f\tau} d\tau \quad (\%1)$$

$$STFT_x^F(t, f) = e^{-j2\pi f t} \int_{-\infty}^{+\infty} X(v) \Gamma^*(v - f) e^{j2\pi v t} dv \quad (\%2)$$

Z powodu braku dyskretyzacji syntezę sygnału przeprowadza się w STFT na podstawie:

$$x(t) = \frac{1}{\gamma(0)} \int_{-\infty}^{+\infty} STFT_x(t, f) e^{j2\pi f t} df$$

Jest to odwrotne przekształcenie Fouriera funkcji STFT(t, f) względem zmiennej f , unormowane przez wartość okna $\gamma(t)$ dla chwili zerowej.

Funkcja $\gamma(t)$ oznacza czasowe okno obserwacji, $\Gamma(f)$ jest zaś jej widmem Fouriera. Równanie to nosi także nazwę metody „przesuwającego się okna” MWM (ang. Moving Window Method) w dziedzinie czasowej lub częstotliwościowej.

W dziedzinie czasowej STFT polega na wykonywaniu prostego przekształcenia Fouriera na kolejnych fragmentach sygnału, „wycinanych” przez przesuwające się okno $\gamma(t)$.

W dziedzinie częstotliwościowej STFT jest natomiast równoważne:

odwrotnemu przekształceniu Fouriera fragmentu widma sygnału $X(v)$, „wyciętemu” przez przesunięte w częstotliwości widmo okna $\Gamma(v - f)$ (jest to filtracja sygnału filtrem pasmowoprzepustowym o częstotliwości środkowej równej f i przesunięciu w częstotliwości sygnału czasowego otrzymanego z 1) do częstotliwości zerowej poprzez jego wymnożenie z $\exp(-j2\pi ft)$.

Metodę STFT w dziedzinie częstotliwościowej udoskonalono. Oblicza się w niej wartości funkcji STFT $x(t, f)$ według (%2), lecz przyporządkowuje się je innym wartościom parametrów t i f , które są związane z częstotliwością chwilową i opóźnieniem grupowym.

STFT można także interpretować jako grzebień równocześnie pracujących filtrów. W interpretacji dolnoprzepustowej dla każdej częstotliwości f sygnał jest najpierw przesuwany w częstotliwości o $-f$ poprzez wymnożenie z $\exp(-j2\pi ft)$, a następnie przepuszczany przez filtr pasmowoprzepustowy o odpowiedzi impulsowej $\gamma^*(-t)$.

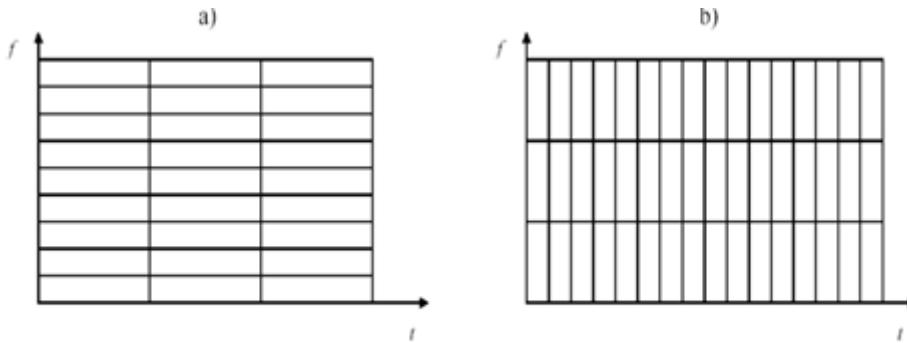
W interpretacji górnoprzepustowej dla każdej częstotliwości f sygnał jest najpierw przepuszczany przez filtr pasmowoprzepustowy o częstotliwości środkowej f i odpowiedzi impulsowej równej $\gamma^*(-t)e^{(j2\pi ft)}$, a następnie przesuwany w częstotliwości do częstotliwości zerowej poprzez wymnożenie z $e^{(-j2\pi ft)}$.

Z krótkoczasową transformacją Fouriera jest związany tzw. spektrogram, który definiuje się jako kwadrat jej modułu:

$$S_x^{SPEC}(t, f) = |STFT_x(t, f)|^2$$

W STFT szerokie okno $\gamma(t)$ powoduje dużą rozdzielcość w osi częstotliwości, a mniejszą w osi czasu (rysunek a poniżej). Wąskie okno daje efekt odwrotny (rysunek b poniżej).

Niemogliwa jest jednoczesna duża rozdzielcość metody w obu osiach.



Przykłady możliwych dekompozycji sygnału w przestrzeni TF dla:
a) długiego okna $q(k)$, b) krótkiego okna $q(k)$

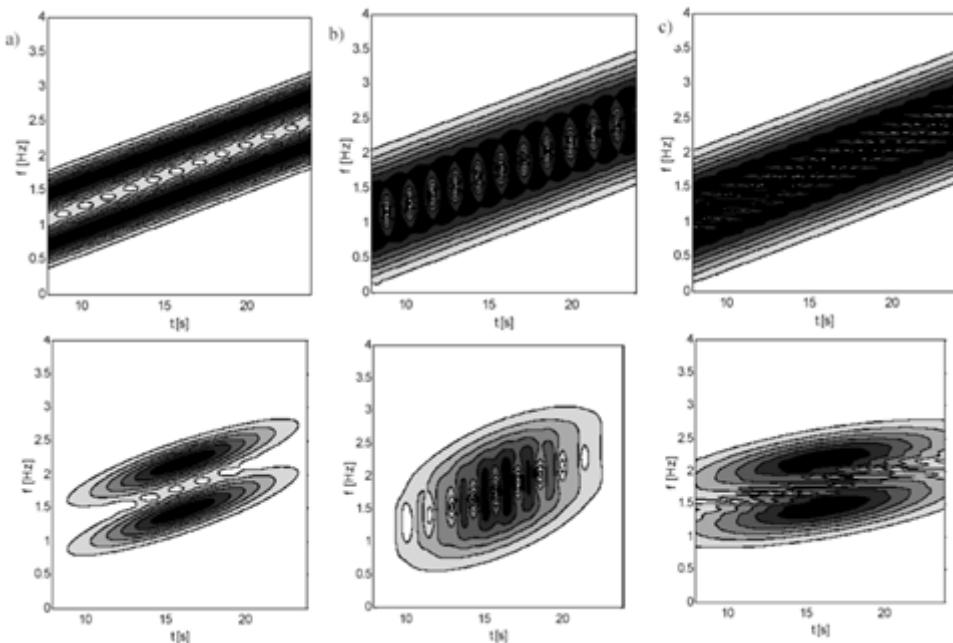
W jaki sposób należy dobierać długość okna analizy $\gamma(t)$?

Niech A oznacza średniokwadratową długość czasu trwania okna, B zaś średniokwadratową szerokość jego widma Fouriera. Podczas analizy sygnałów zmodulowanych liniowo w częstotliwości (LFM) długość okna powinna być taka, aby stosunek szerokości częstotliwościowej B do czasowej A okna był równy stosunkowi przyrostu częstotliwości do czasu, w którym on wystąpił (warunek równości katów nachylenia):

$$(B/A)_{\text{optim1}} \equiv \Delta f / \Delta t$$

Jeśli dodatkowo analizowany sygnał jest zmodulowany w amplitudzie za pomocą okna czasowego, które ma średniokwadratową szerokość czasową A_{okna} i częstotliwościową B_{okna} , to optymalne parametry okna analizy $\gamma(t)$ są określone równaniem:

$$(B/A)_{\text{optim2}} \equiv \Delta f / \Delta t - B_{\text{okna}} / A_{\text{okna}}$$



Przykład wpływu wyboru długości okna analizy $\gamma(k)$ w krótkoczasowej transformacji Fouriera: u góry - suma dwóch sygnałów z liniową modulacją częstotliwości LFM, u dołu - suma tych samych sygnałów LFM po ich dodatkowej modulacji w amplitudzie za pomocą funkcji okna czasowego, a) okno o optymalnej długości, b) okno za krótkie, c) okno za długie.

Transformacja falkowa

Transformacja falkowa jest obecnie jednym z najpopularniejszych i najdynamiczniej rozwijanych się narzędzi analizy częstotliwościowej sygnałów niestacjonarnych.

Koncepcyjnie jest ona równoważna metodzie analizy częstotliwościowej o „stałoprocentowym paśmie”

$\Delta f/f_0 = \text{const}$ (ang. CPB - Constant Percentage Bandwidth), stosowanej w akustyce, lecz inaczej realizowanej.

Zastosowania transformacji falkowej, dotyczące między innymi czasowo-częstotliwościowej identyfikacji stanu obiektów biologicznych i technicznych poprzez analizę generowanych przez nie sygnałów, odszumiania w dziedzinie współczynników transformacji, detekcji punktów skokowej zmiany charakteru sygnałów oraz kodowania i kompresji sygnałów.

Ciągła transformacja falkowa sygnału $x(t)$ jest zdefiniowana w następujący sposób w dziedzinie czasu i częstotliwości:

$$CWT_x^T(t, a) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(\tau) \gamma^* \left(\frac{\tau - t}{a} \right) d\tau$$

$$CWT_x^F(t, a) = \sqrt{|a|} \int_{-\infty}^{+\infty} X(f) \Gamma^*(af) e^{j2\pi f t} df$$

transformacja zaś do niej odwrotna jest określona zależnością:

$$x(t) = \frac{2}{C_\gamma} \int_0^{+\infty} \left[\int_{-\infty}^{+\infty} CWT_x(\tau, a) g \left(\frac{t - \tau}{a} \right) d\tau \right] \frac{da}{a^2}$$

gdzie:

$$C_\gamma = \int_{-\infty}^{+\infty} \frac{|\Gamma(f)|^2}{|f|} df < \infty, \quad \Gamma(0) = 0$$

Funkcja $g(t)$ oznacza falkę syntezującą, $\gamma(t)$ - dualną do niej falkę analizującą, a $G(f)$ i $\Gamma(f)$ są ich widmami Fouriera. Dzielenie i mnożenie przez, ma zapewnić niezmiennosć energii falek i ich widm po przeskalowaniu.

Równanie pierwsze reprezentuje filtrację sygnału analizowanego $x(t)$ przez sygnał analizujący $\gamma(t)$, przeskalowywany w dziedzinie czasu współczynnikiem a („rozciągany” dla $a > 1$ oraz „ścisany” dla $a < 1$).

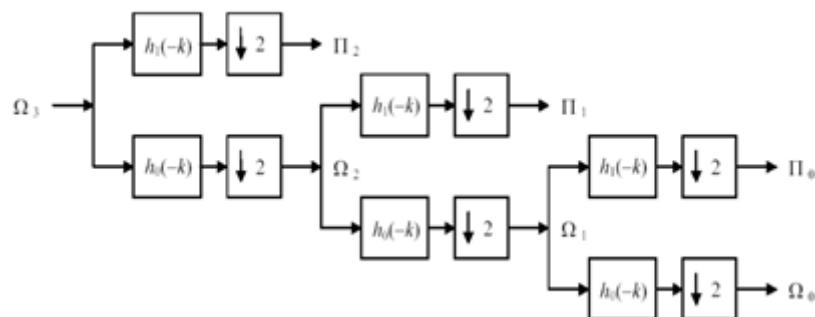
Równoważne mu równanie przedstawia natomiast odwrotne przekształcenie Fouriera iloczynu widma sygnału i przeskalowanego widma falki $\gamma(t)$.

Skalogram, związany z transformacją falkową, jest zdefiniowany jako:

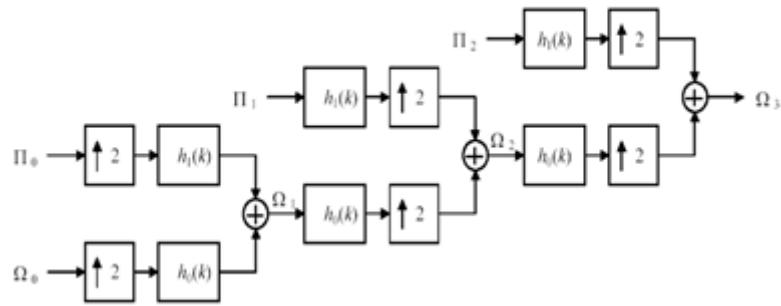
$$S_x^{SCAL}(t, a) = |CWT_x(t, a)|^2$$

Tak jak spektrogram był kwadratem modułu krótkoczasowej transformacji Fouriera STFT, tak skalogram jest kwadratem modułu transformacji falkowej, czyli reprezentacji typu czas-skala.

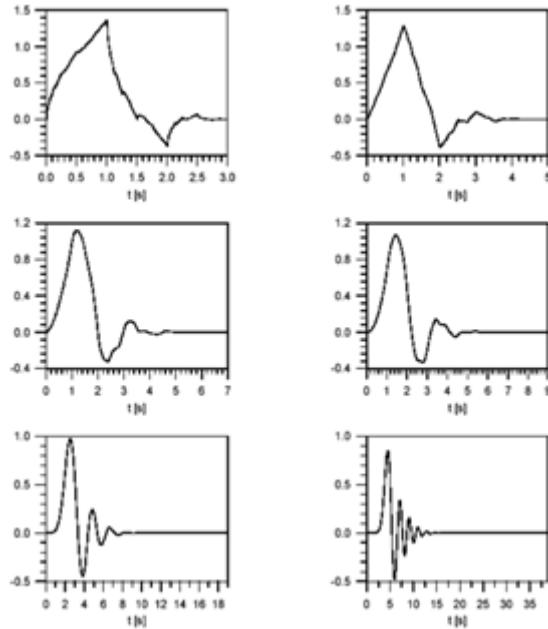
Ponieważ sygnał analizujący $\gamma(t)$, spełniający rolę funkcji bazowej tak określonej dekompozycji, jest sygnałem impulsowym zlokalizowanym w osi czasu wokół $t = 0$ oraz mającym ograniczone, pasmowe widmo częstotliwościowe ($f_0 \pm \Delta f/2$), zapis $S_x^{SCAL}(t, a)$ w dziedzinie czas-skala jest równoważny zapisowi $S_x^{SCAL}(t, (f_0 \pm \Delta f/2)/a)$ w dziedzinie czas-częstotliwość. Oznacza to, że możemy interpretować skalogram jako reprezentację czasowo-częstotliwościową $S_x^{SCAL}(t, f)$.



Trzy poziomy analizy falkowej

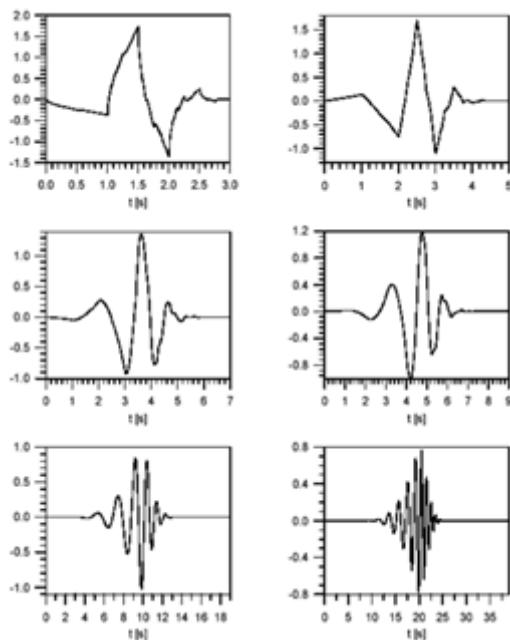


Trzy poziomy syntezy falkowej



Funkcje skalujące Daubechies: φ_{D4} , φ_{D6} , φ_{D8} , φ_{D10} , φ_{D20} , φ_{D40} (kolejno poziomo).

W języku Matlab oznaczane jako φ_{db2} , φ_{db3} , φ_{db4} , φ_{db5} , φ_{db10} , φ_{db20} .



Falki Daubechies: ψ_{d4} , ψ_{d6} , ψ_{d8} , ψ_{d10} , ψ_{d20} , ψ_{d40} (kolejno poziomo).

W języku Matlab oznaczane jako ψ_{db2} , ψ_{db3} , ψ_{db4} , ψ_{db5} , ψ_{db10} , ψ_{db20} .