INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Koherentní demodulace a její využití v měřicí technice

Teorie, přístroje, aplikace

Učební texty k semináři

Autor:

Prof. Ing. Stanislav Ďaďo, DrSc., ČVUT Fakulta elektrotechnická, Praha

Datum:

18. listopadu 2011

Centrum pro rozvoj výzkumu pokročilých řídicích a senzorických technologií CZ.1.07/2.3.00/09.0031

TENTO STUDIJNÍ MATERIÁL JE SPOLUFINANCOVÁN EVROPSKÝM SOCIÁLNÍM FONDEM A STÁTNÍM ROZPOČTEM ČESKÉ REPUBLIKY

Obsah

1. ODHAD MĚRONOSNÝCH PARAMETRŮ SIGNÁLU

V naprosto převažující většině prací se funkce koherentní demodulace odvozuje od postupů amplitudové demodulace známých z radiotechniky nebo od řízených měřicích usměrňovačů. V tomto příspěvku je použito pohledu na koherentní demodulaci jako zvláštního případu odhadu parametrů (amplitudy) signálu z naměřených dat, jelikož takto lze přirozeným postupem odvodit nejdůležitější vlastnosti koherentní demodulace – možnost měření parametrů signálu hluboce utopených v šumu („denoising“) [1]. Na základě obecných postupů odhadu parametrů signálu popsaných v Přílohách je v první kapitole odvozen vztah pro odhad amplitudy signálu. [11]

1.1. Odhad amplitudy signálu

Odhadováním v měřicí technice obecně rozumíme nalezení odhadu vektoru měronosných parametrů signálu x **(**např. amplituda, frekvence, fáze) z naměřených hodnot, tvořících vektor dat “**y“**. Z několika metod, lišících se kritérii minimalizace chyb odhadu, vybereme postupy nejčastěji užívané v měřicí technice [5].

Odhad amplitudy A_0 amplitudově modulovaného signálu $A_0.s(t)$ přítomného v aditivní směsi $V(t)$ se šumem $N(t)$, tj. platí

$$V(t) = A_0.s(t) + N(t) \tag{1.1}$$

je teoretickým základem postupů koherentního zpracování signálů. Ve vztahu (1.1) je $s(t)$ zobecněný průběh periodické nosné vlny.

Hodnoty (vzorky) průběhů $V(t_i)$, $s(t_i)$ a $N(t_i)$, určené v intervalu pozorování, tvoří sloupcové vektory $\mathbf{V}$, $\mathbf{s}$ a $\mathbf{N}$, které jsou vázány vztahem

$$\mathbf{V} = A_0\mathbf{s} + \mathbf{N} \tag{1.2}$$

Vztah (1.2) mezi naměřenými hodnotami $\mathbf{V}$ a hledanou skalární veličinou amplitudou A_0 je lineární. Proto pro odhad použijeme kritéria LMMSE (P2.14).

Pro úplnost výkladu upravíme tento vztah tak, aby byl analogický s výchozím tvarem uvedeným v příloze P2 (10), tj. aby hledaný odhad byl na levé straně rovnice. Jelikož A_0 je skalární veličina, její přímý výpočet z rovnice (1.2) není možný a pro dodržení pravidel maticového počtu je nutné rovnici upravit násobením zleva jednotkovou maticí $\mathbf{I}$

$$\mathbf{IV} = A_0\mathbf{Is} + \mathbf{IN} \tag{1.3}$$

Vztah (1.3) je ekvivalentní zápisu pomocí diagonálních matic

$$\mathbf{V}_d = A_0 \mathbf{s}_d + \mathbf{N}_d \tag{1.4}$$

kde index d značí diagonální matici.

Amplituda A_0 je pak explicitně určena inverzní funkcí, tj. výrazem

$$A_0 = \mathbf{V}_d \mathbf{s}_d^{-1} - \mathbf{N}_d \mathbf{s}_d^{-1} \tag{1.5}$$

Pak pro odhad amplitudy dle kritéria LMMSE - vztah (P2.14) - platí

$$A_{0LMS} = \boldsymbol{\mu}_{A0} + \mathbf{P}_{A0V}^{\mathrm{T}}.\mathbf{P}_{VV}^{-1}.(\mathbf{V} - \boldsymbol{\mu}_V) \tag{1.6}$$

Pozn.: Jelikož výraz pro odhad (1.6) neobsahuje konstanty $\mathbf{A}$, $\mathbf{b}$, je možné jej použít i bez explicitního vyjádření odhadovaného parametru inverzní lineární funkcí.

Kovarianční matice lze pro tento případ vyjádřit následovně:

$$\mathbf{P}_{A0V} = E\{A_0 \mathbf{V}\} - E\{A_0\} E\{\mathbf{V}\} \tag{1.7}$$

$$E\{A_0 \mathbf{V}\} = E\{A_0 (A_0\, \mathbf{s} + \mathbf{N})\} = E\{A_0^2 \mathbf{s}\} + E\{A_0 \mathbf{N}\} = E\{A_0^2 \mathbf{s}\} \tag{1.8}$$

Zde jsme předpokládali nulovou střední hodnotu šumu $N(t)$.

Dále platí

$$E\{\mathbf{V}\} = E\{A_0 \mathbf{s}\} + E\{\mathbf{N}\} = E\{A_0\} \mathbf{s} \tag{1.9}$$

Po dosazení vztahů (1.8) a (1.9) do (1.6) získáme konečný tvar matice (1.6)

$$\mathbf{P}_{A0V} = E\{A_0^2 \mathbf{s}\} - \mathbf{s}.E\{A_0\}.E\{A_0\} = \mathbf{s}.\left(E\{A_0^2\} - (E\{A_0\})^2\right) = \mathbf{s}.\sigma_{A0}^2 \tag{1.10}$$

kde σ_{A0} je rozptyl hledané amplitudy.

Určeme nyní varianční matici dat s použitím vztahu (1.11)

$$\begin{aligned}\mathbf{P}_{VV} &= \operatorname{var}(A_0 \mathbf{s} + \mathbf{N}) = \operatorname{var} A_0 \mathbf{s} + \operatorname{cov}(A_0 \mathbf{s}, \mathbf{N}) + \operatorname{cov}(\mathbf{N}, A_0 \mathbf{s}) + \operatorname{var} \mathbf{N} = \\ &= \operatorname{var} A_0 \mathbf{s} + \operatorname{var} \mathbf{N}\end{aligned} \tag{1.11}$$

jelikož $\mathbf{s}$ a $\mathbf{N}$ jsou nezávislé procesy.

Použitím vztahů (1.11) a (1.12)

$$\operatorname{var} A_0 \mathbf{s} = \mathbf{s}^{\mathrm{T}} \mathbf{s}.\sigma_{A0}^2 \tag{1.12}$$

lze upravit matici $\mathbf{P}_{VV}$

na tvar

$$\mathbf{P}_{VV} = \sigma_{A0}^2 \mathbf{s}^{\mathrm{T}} \mathbf{s} + \operatorname{var} \mathbf{N} = \sigma_{A0}^2 \mathbf{s}^{\mathrm{T}} \mathbf{s} + \mathbf{P}_{NN} \tag{1.13}$$

Upravme nyní součin matic ze vztahu (1.6) na tvar

$$\mathbf{P}_{VV}^{-1}\,\mathbf{P}_{A0V} = \left(\sigma_{A0}^{2}\mathbf{s}^{\mathrm{T}}\mathbf{s} + \mathbf{P}_{NN}\right)^{-1}\mathbf{s}\,\sigma_{A0}^{2} = \left[\mathbf{P}_{NN}\left(1 + \sigma_{A0}^{2}\mathbf{s}^{\mathrm{T}}\mathbf{P}_{NN}^{-1}\mathbf{s}\right)\right]^{-1}\mathbf{s}.\,\sigma_{A0}^{2} \tag{1.14}$$

Zaveďme kvadratickou formu (skalár)

$$F_S = \mathbf{s}^{\mathrm{T}}\,\mathbf{P}_{NN}^{-1}\,\mathbf{s} \tag{1.15}$$

závislou pouze na průběhu nosné signálu $s(t)$ a úměrnou jeho energii.
S použitím členu F_S lze vztah (1.14) upravit na tvar

$$\mathbf{P}_{VV}^{-1}.\mathbf{P}_{A0V} = \sigma_{A0}^{2}\,\mathbf{P}_{NN}^{-1}\,\mathbf{s}\left(1 + F_S\,\sigma_{A0}^{2}\right)^{-1} \tag{1.16}$$

Za předpokladu nulové střední hodnoty rušení $N(t)$ platí

$$\boldsymbol{\mu}_V = E\{\mathbf{V}\} = E\{\mathbf{s} + \mathbf{N}\} = E\{A_0\mathbf{s}\} = \mathbf{s}\,\mu_{A0} \tag{1.17}$$

Po dosazení do základního vztahu pro odhad (1.6) dostáváme

$$A_{0LMS} = \mu_{A0} + \frac{\sigma_{A0}^{2}}{1 + F_S\,\sigma_{A0}^{2}}\left(\mathbf{V}^{\mathrm{T}}\,\mathbf{P}_{NN}^{-1}\,\mathbf{s} - \mu_{A0}\mathbf{s}^{\mathrm{T}}\,\mathbf{P}_{NN}^{-1}\,\mathbf{s}\right) \tag{1.18}$$

Označme první člen v závorce (kvadratická forma) jako F_V, takže platí

$$F_V = \mathbf{V}^{\mathrm{T}}\,\mathbf{P}_{NN}^{-1}\,\mathbf{s} \tag{1.19}$$

Po úpravách dostáváme

$$A_{0LMS} = \mu_{A0}\left(1 - \frac{\sigma_{A0}^{2}F_S}{1 + \sigma_{A0}^{2}F_S}\right) + \sigma_{A0}^{2}\frac{F_V}{1 + \sigma_{A0}^{2}F_S} = \frac{\mu_{A0} + \sigma_{A0}^{2}F_V}{1 + \sigma_{A0}^{2}F_S} \tag{1.20}$$

Pro třídu signálů splňujících nerovnost

$$\sigma_{A0}^{2}F_S >> 1 \tag{1.21}$$

(jde např. o signály s rovnoměrnou hustotou rozdělení amplitudy) se vztah pro odhad zjednoduší:

$$A_{0LMS} \approx \frac{F_V}{F_S} = \frac{\mathbf{V}^{\mathrm{T}}\,\mathbf{P}_{\mathrm{NN}}^{-1}\mathbf{s}}{\mathbf{s}^{\mathrm{T}}\,\mathbf{P}_{\mathrm{NN}}^{-1}\mathbf{s}} = \frac{\mathbf{V}^{\mathrm{T}}\,\mathbf{P}_{\mathrm{nn}}^{-1}\mathbf{s}}{\mathbf{s}^{\mathrm{T}}\,\mathbf{P}_{\mathrm{nn}}^{-1}\mathbf{s}} \tag{1.22}$$

Zde byla zavedena *normalizovaná* varianční matice šumu

$$\mathbf{P}_{nn} = \frac{1}{\sigma_{\mathrm{N}}^{2}}\mathbf{P}_{NN} \tag{1.23}$$

kde σ_{N}^{2} je střední kvadratická hodnota rušení $N(t)$.

Na základě vztahu v P2 (14a) lze určit *rozptyl chyby odhadu amplitudy*

$$\sigma_{A0LMS}^{2} = \sigma_{A0}^{4}\frac{F_S}{1 + \sigma_{A0}^{2}F_S} \tag{1.24}$$

Za předpokladu (1.21) platí pro rozptyl chyby odhadu

$$\sigma^2_{A0LMS} \approx \sigma^2_{A0} \tag{1.25}$$

tj. odhadem se podstatně nezvětšuje původní rozptyl odhadované veličiny.

1.1.1. *Odhad amplitudy pro případ bílého šumu*

Varianční matice rušení $\mathbf{P}_{nn}$ je v častém případě *bílého šumu* jednotková ($\mathbf{P}_{nn} = \mathbf{I}$), takže pro odhad amplitudy platí

$$A_{0LMS} \approx \frac{\mathbf{V}^{\mathsf{T}}\mathbf{I}\,\mathbf{s}}{\mathbf{s}^{\mathsf{T}}\mathbf{I}\,\mathbf{s}} \tag{1.26}$$

Čitatel vztahu (1.26) odpovídá přibližně diskrétní formě ***vzájemné korelace*** průběhů $V(t)$ a $s(t)$ pro nulové vzájemné zpoždění. Jmenovatel je úměrný energii signálu, takže platí výraz

$$A_{0LMS} = \frac{\sum_{i=1}^{M} V(t_i)s(t_i)}{\sum_{i=1}^{M} s^2(t_i)} = \frac{\frac{1}{M}\sum_{i=1}^{M} V(t_i)s(t_i)}{\frac{1}{M}\sum_{i=1}^{M} s^2(t_i)} \tag{1.27}$$

To ukazuje na souvislost odhadu amplitudy kritériem LMMSE s *korelačními metodami* zmenšování vlivu rušivých signálů.

Výraz pro odhad (1.27) by bylo možné odvodit také použitím alternativního postupu na základě vztahu mezi daty a odhadovanými parametry vyjádřeného rovnicí (1.2).

1.1.2. *Odhad amplitudy pro spojité signály*

Zmenšováním intervalu mezi sousedními vzorky až na limitní nulovou hodnotu budou vzorkované průběhy $V(t)$ a $s(t)$ blízké spojitým a sumace ve vztahu (1.27) mohou být vyjádřeny určitými integrály v intervalu pozorování t_0 až $(t_0 + T_m)$

$$A_{0LMS} = \frac{\frac{1}{T_m}\int_{t_0}^{t_0+T_m} V(t).s(t)dt}{\frac{1}{T_m}\int_{t_0}^{t_0+T_m} s^2(t)dt} \tag{1.28}$$

Tento vztah je základem *analogových koherentních demodulátorů* (model je na Obrázku 2.10)

1.1.3. *Fázová citlivost*

V jednoduchém případě kdy vstupní a referenční signál jsou harmonické o stejné frekvenci je odhadovaná amplituda určena rovnicí

$$A_{0LMS} = V_m \sin(\omega t + \varphi_V) S_m sin(\omega t + \varphi_S) = V_m S_m \, 1/2 \, [cos(\varphi_V - \varphi_S) + \\ + cos(2\omega t + \varphi_V + \varphi_S)] \tag{1.29}$$

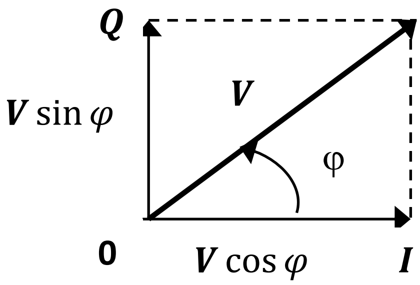
Operací analogickou integraci (výpočtu střední hodnoty), realizovanou dolnofrekvenční propustí, je potlačena složka o frekvenci 2ω α výstupní signál koherentního demodulátoru je úměrný fázovému rozdílu

$$A_{0LMS0} = 1/2 \, V_m S_m cos(\varphi_V - \varphi_S) \tag{1.30}$$

Pro stálé amplitudy signálu a reference je výstupní signál KD úměrný pouze fázovému rozdílu obou průběhů a KD se chová jako známý *fázový detektor*.

Při nulovém fázovém rozdílu je výstupní signál maximální (***soufázová** složka signálu-$\boldsymbol{I}$*). Pro posuv signálu vůči referenci o 90° (***kvadraturní** složka signálu-$\boldsymbol{Q}$*) je naopak výstup fázového detektoru nulový.

Princip fázové detekce lze použít pro určení reálné ($\boldsymbol{I}$) a imaginární ($\boldsymbol{Q}$) složky fázoru harmonického signálu $V(t)$ o amplitudě V.



Obrázek 1.1. Určení složek fázoru V

Dva koherentní demodulátory- fázové detektory- s referenčními signály posunutými o $\pi/2$ jsou základem tzv. *vektor-voltmetrů* (správněji fázor-voltmetrů).

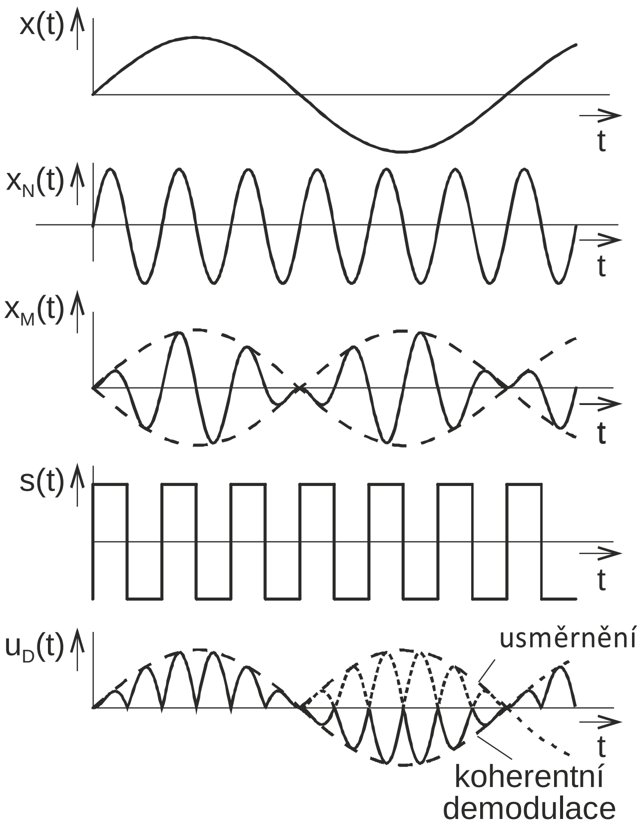
Pro modul fázoru $\boldsymbol{V}$ platí

$$V = \sqrt{I^2 + Q^2} = \sqrt{(Re(V))^2 + (Im(V))^2} = \sqrt{(Vcos\varphi)^2 + (Vsin\varphi)^2} = V \tag{1.31}$$

Hodnotu modulu V lze tedy získat výpočtem i při kolísání fáze.

1.1.4. *Souvislost obecné a koherentní amplitudové demodulace*

Signál $x(t)$ moduluje amplitudu nosné vlny $x(t)$ (Obrázek 1.2), výsledkem je amplitudová modulace s potlačenou nosnou $x_M(t)$. Diodový (neřízený) amplitudový demodulátor dává na výstupu průběh obálky (usměrnění – $u_D(t)$.Koherentní demodulátor (řízený,spínačový) se chová jako obvod násobící střídavě průběh $u_M(t)$ hodnotami +1 a -1 jak udává průběh spínací funkce $s(t)$. Výsledkem je průběh obálky respektující změnu fáze $x_M(t)$ při průchodu $x(t)$ nulou. Při posunutí $s(t)$ o čtvrtinu periody vůči $x_N(t)$ je střední hodnota $u_D(t)$ nulová [10].



Obrázek 1.2. Typické časové průběhy obecně a koherentně demodulovaného signálu

1.1.5. *Souvislost s Fourierovou transformací*

Čitatel vztahu (1.28) pro harmonický průběh referenčního signálu $s(t)$ odpovídá výpočtu harmonické Fourierovy transformace signálu $V(t)$. V modelové interpretaci jde o obvodové řešení *ideálního koherentního demodulátoru* (Obrázek 2.10).

1.1.6. *Souvislost s procesem směšování*

Operace násobení je základem nejen koherentní demodulace, ale také směšování signálu $V(t)$ se signálem místního oscilátoru, v tomto případě referenčního signálu $s(t)$. Na rozdíl od typického směšovače je však vzhledem k rovnosti frekvencí signálu a oscilátoru výstupem nikoliv mezifrekvenční kmitočet $(f_V - f_s)$, ale stejnosměrný signál úměrný amplitudě vstupního signálu.

1.2. Interpretace vztahu pro odhad amplitudy

1.2.1. *Podmínky koherence*

Při opakování pokusu se v závislosti na změně zpoždění signálu $\boldsymbol{\kappa}$ vůči referenčnímu bodu na časové ose hodnota $\boldsymbol{F_V}$ mění i při jinak stálých hodnotách ostatních parametrů. Mohou nastat dva základní případy:

a) $\boldsymbol{\kappa}$ = const, nebo je předem známé. Pak jde o *koherentní* signály, u nichž je zaručeno, že při každém opakování pokusu se získá posloupnost vzorků odebraných ze stejných míst průběhu signálu.

b) $\boldsymbol{\kappa}$ je náhodně proměnné, nebo tvar funkcionálu $s(t)$ předem neznáme. Pak jde o *nekoherentní* signály a určení odhadu dle vztahu (1.22) nepřináší žádné praktické výhody, jelikož vektor $\boldsymbol{s}$ se mění náhodně nebo jej nelze určit.

Je tedy nutné splnit *podmínku* ***koherence*** průběhu řídícího odběr vzorků a nosné vlny $s(t)$. Okamžiky odběru vzorků musí být v přesných časových relacích s nosným průběhem.

V nejčastěji užívaném *ekvidistantním vzorkování* se vyžaduje, aby frekvence odběru vzorků f_{vz} byla celočíselným násobkem nebo podílem základní frekvence f_s nosné vlny $s(t)$ a vzájemné časové posunutí základní harmonické $s(t)$ a vzorkovacího signálu bylo stálé.

Volba f_{vz} je v souladu se vzorkovacím teorémem určena ze *spektra časového průběhu amplitudy* A_0.

Podmínku *apriorní* znalosti průběhu nosné $s(t)$ lze snadno splnit např. u můstkových měření, nebo při zpracování signálů ze senzorů záření přerušovaného rotující clonkou. Základní vlastnosti odhadu dle kritéria LMMSE (např. potlačení rušení) zůstávají však zachovány i tehdy, použijeme-li pro výpočet libovolný *koherentní průběh* se skutečnou nosnou signálu $s(t)$. Tak např. pro harmonický průběh $s(t)$ lze výpočet odhadu realizovat ze vzorků získaných v časových okamžicích odpovídajících např. maximálním hodnotám $s(t)$. Ekvivalentní průběh nosné jsou pak impulsy s jednotkovou amplitudou a šířkou respektující dobu odběru vzorku.

Důležitým důsledkem koherence signálu je skutečnost, že pro odhad jeho parametrů lze použít vztah (1.22), v němž vystupují vzorky dat $\mathbf{V}$ v první mocnině, tj. jde o ***lineární operaci***. V opačném případě by vyšší mocniny $\mathbf{V}$ vedly ke zvýšení úrovně rušivého signálu úměrně k výskytu smíšených členů typu $(A_0 s)^n N^m$, způsobujících výskyt složek signálu o nízké frekvenci blízkých k frekvenci užitečného signálu (obálky). Existence vyšších mocnin přispívá takto mimo jiné k potlačení slabého signálu silným rušením, jak je dobře známo z chování nelineárních detektorů. Proto je možné postupem odhadu parametrů při zachování podmínky koherence *koherentní demodulací* určit i signály pod úrovní šumu.

Naopak, nejsou-li a priori známy parametry nosné (funkcionálu) $\boldsymbol{s}$, takže podmínku koherence nelze splnit, je nutné přítomnost nebo velikost signálu zjišťovat např. z přírůstku energie. tj. použít *nelineární detektory* nebo výpočetní operace. Pak ovšem může nastat potlačení signálu rušením.

1.2.2. *Podmínka koherence pro spojité signály*

Obdobně jako v předchozím případě je přípustné použít místo skutečné nosné $s(t)$ jiný časový průběh $s_{ekv}(t)$, *koherentní* s nosnou. Podmínkou je, aby vzájemná korelace signálů $v(t)$ a $s_{ekv}(t)$ (viz čitatel vztahu (1.28)) byla nenulová. Tuto podmínku splňují průběhy se spektrem obsahujícím alespoň některé spektrální složky shodné se spektrem $s(t)$.

Zajištění koherence je obtížné v situacích kdy zdroj referenčního signálu a vlastní koherentní demodulátor jsou příliš vzdáleny a drátové propojení je nemožné nebo příliš nákladné.

1.2.3. *Vliv statistických parametrů rušení N(t)*

Optimální odhad amplitudy dle kritéria LMMSE vyžaduje znalost kovarianční matice rušení $\mathbf{P}_{nn}$. V případě bílého šumu platí $\mathbf{P}_{nn} = \mathbf{I}$ a výraz pro odhad amplitudy se zjednoduší na střední hodnotu součinu dat $V(t)$ a nosné $s(t)$, tj. diskrétní vzájemné korelace. V praxi se pro jednoduchost výpočet odhadu provádí za předpokladu rušení typu bílý šum, i když výsledek odhadu pak nemusí být optimální.

1.2.4. *Zlepšení poměru signál/šum odhadem amplitudy dle LMMSE*

Poměr signál šum SNR na vstupu (před výpočtem odhadu) je definován vztahem

$$\left(\frac{S}{N}\right)_i = \frac{A_0}{\sigma_N} \tag{1.32}$$

kde σ_N^2 je střední kvadratická hodnota rušení $N(t)$.

Hodnotu užitečné a rušivé složky na výstupu určíme rozborem vztahu (1.22)

$$A_{0LMS} \approx \frac{F_V}{F_s} = \frac{(A_0\mathbf{s}+\mathbf{N})^{\mathrm{T}}\,\mathbf{P}_{nn}^{-1}\mathbf{s}}{F_s} = \frac{A_0\mathbf{s}^{\mathrm{T}}\mathbf{P}_{nn}^{-1}\,\mathbf{s}+\mathbf{N}^{\mathrm{T}}\mathbf{P}_{nn}^{-1}\,\mathbf{s}}{F_s} = A_0 + \frac{\mathbf{N}^{\mathrm{T}}\,\mathbf{P}_{nn}^{-1}\,\mathbf{s}}{F_s} \tag{1.33}$$

Druhý člen ve výrazu (1.33) představuje rušivou složku na výstupu s rozptylem (variancí) určenou podle definice a vztahu (1.20)

$$\sigma_{NO}^2 = \operatorname{var}\left(\frac{\mathbf{N}^{\mathrm{T}}\mathbf{P}_{nn}^{-1}\mathbf{s}}{F_s}\right) = \frac{1}{F_s}(\mathbf{P}_{nn}^{-1}\,\mathbf{s})^{\mathrm{T}}\,\operatorname{var}\mathbf{N}^{\mathrm{T}}\,(\mathbf{P}_{nn}^{-1}\,\mathbf{s})\frac{1}{F_s} \tag{1.34}$$

V souladu s definicí je variance šumu je dána vztahem

$$\operatorname{var}\mathbf{N} = \mathbf{P}_{NN} = \sigma_N^2\mathbf{P}_{nn} \tag{1.35}$$

Po dosazení dostáváme

$$\sigma_{NO}^2 = \frac{1}{F_s^2}(\mathbf{P}_{nn}^{-1}\mathbf{s})^{\mathrm{T}}\sigma_N^2\mathbf{P}_{nn}(\mathbf{P}_{nn}^{-1}\mathbf{s}) = \frac{\sigma_N^2}{F_s} \tag{1.36}$$

Poměr signál šum na výstupu odpovídá vztahu

$$\left(\frac{S}{N}\right)_{\mathrm{O}} = \frac{A_0}{\sigma_{NO}} = \frac{A_0\sqrt{F_s}}{\sigma_N} = \left(\frac{S}{N}\right)_i\sqrt{F_s} \tag{1.37}$$

Zlepšení poměru signál šum G odhadem dle kritéria LMMSE je dáno výrazem

$$G = \frac{\left(\frac{S}{N}\right)_{\mathrm{O}}}{\left(\frac{S}{N}\right)_I} = \sqrt{F_s} = \sqrt{\mathbf{s}^{\mathrm{T}}\,\mathbf{P}_{\mathrm{nn}}^{-1}\mathbf{s}} \tag{1.38}$$

Pro bílý šum je $\mathbf{P}_{nn} = \mathbf{I}$, takže platí

$$G = \sqrt{\mathbf{s}^{\mathrm{T}}\mathbf{I}\,\mathbf{s}} = \left(\sum_{i=1}^{N} s^2(t_i)\right)^{1/2} \tag{1.39}$$

G dosahuje největší hodnoty, jsou-li hodnoty všech vzorků nosného průběhu $s^2(t_i) = 1$. Pak platí

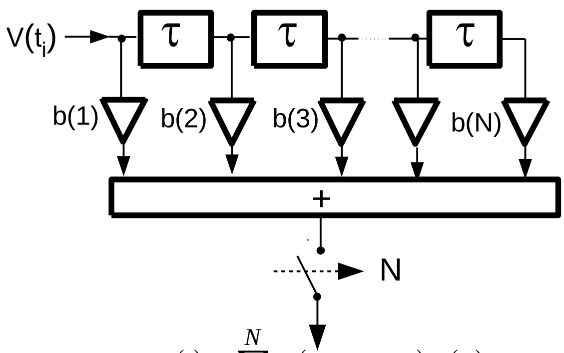
$$G = \sqrt{N} \tag{1.40}$$

Takováto situace odpovídá pravoúhlému průběhu nosné $s(t)$ a je v plném souladu se známým poznatkem o zmenšení rozptylu chyby měření výpočtem průměru N opakovaných měření veličiny za přítomnosti rušení typu bílý šum.

1.3. Odhad LMMSE jako číslicová filtrace [5]

Jelikož jmenovatel výrazu (1.22) je určen pouze průběhem $s(t)$, lze odhad amplitudy signálu dle tohoto vztahu interpretovat jako druh *číslicové filtrace,* realizované transverzálním filtrem typu FIR.

Vstupem filtru jsou vzorky dat $v(t_i)$ a koeficienty filtru $b(n)$ jsou dány prvky sloupcového vektoru, vytvořeného součinem *řádků* $\boldsymbol{p}_n$ $(n = 1, 2, \ldots, N)$ matice $\mathbf{P}_{nn}^{-1}$ a vektoru $\mathbf{s}$, tj. $\mathbf{p}_n \mathbf{s}$.



$$A_0(i) = \sum_{n=1}^{N} V(i - N + n).b(n)$$

Obrázek 1.3. Číslicový filtr pro odhad amplitudy dle kritéria LMMSE

Číslicový filtr na Obrázku 1.3. se od obvyklých filtrů liší vzorkováním výstupu po N vzorcích, jelikož dle vztahu (1.22) je správná hodnota odhadu amplitudy k dispozici až pro $i = N$ vzorcích.

1.3.1. *Přizpůsobený filtr*

Struktura filtru na Obrázku 1.3 odpovídá přizpůsobeným filtrům (*matched filter*), Obrázek 1.4.

Přizpůsobený filtr minimalizuje účinky šumu na měřicí systém a jeho přenos lze odvodit následujícím postupem. Předpokládejme, že vstupní obvod se chová jako lineární filtr s impulsní odezvou $h(t)$ jehož výstupní signál je vzorkován v intervalech T. Původní signál je $s(t)$, avšak po přičtení šumu je na vstupu filtru součet

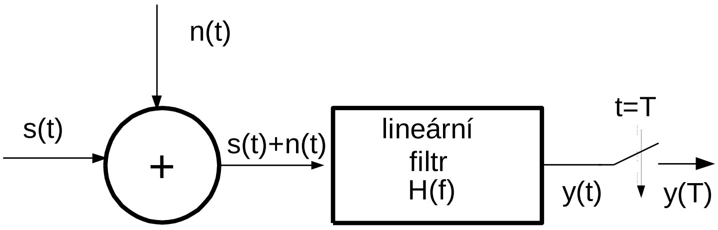
$$y(t) = s(t) + n(t) \quad (1.41)$$

kde $n(t)$ je aditivní šum s nulovou střední hodnotou a výkonovou spektrální hustotou $N_0/2$.

Signál na výstupu filtru je roven

$$y(t) = s_0(t) + n(t) \tag{1.42}$$

kde $s_0(t)$ a $n(t)$ jsou signální a šumová složka.



Obrázek 1.4. K odvození přizpůsobeného filtru

Úkolem je maximalizace okamžité hodnoty výstupního signálu v okamžiku sepnutí spínače (vzorkování) $t = T$. Za předpokladu, že jsou k dispozici informace o signálu $s(t)$ postačující pro splnění podmínky koherence, lze určit T tak, aby poměr kvadrátu špičkové hodnoty signálu k šumu, tj. výraz

$$\eta = \frac{|s_0(T)|^2}{E\{n^2(t)\}} \tag{1.43}$$

byl maximální. Zpětná Fourierova transformace výstupního signálu filtru je určena vztahem

$$s_0(t) = \int_{-\infty}^{\infty} H(f)S(f)e^{j2\pi ft}df \tag{1.44}$$

kde $S(f)$ je Fourierova transformace signálu $s(t)$ a $H(f)$ je přenos filtru.

V okamžiku vzorkování $t = T$ platí pro výstupní signál

$$|s_0(T)|^2 = \left| \int_{-\infty}^{\infty} H(f)S(f)e^{j2\pi fT}df \right|^2 \tag{1.45}$$

Spektrální výkonová hustota šumu na výstupu filtru je rovna

$$S_N(f) = \frac{N_0}{2}|H(f)|^2 \tag{1.46}$$

Pro střední výkon šumové složky pak platí

$$E\left\{n^2(t)\right\}=\frac{N}{2}\int_{-\infty}^{\infty}\left|H(f)\right|^2 df \tag{1.47}$$

Po dosazení do vztahu pro η dostaneme

$$\eta=\frac{\left|\int_{-\infty}^{\infty}H(f)S(f)e^{j2\pi fT}df\right|^2}{\frac{N_0}{2}\int_{-\infty}^{\infty}\left|H(f)\right|^2 df} \tag{1.48}$$

Čitatel zlomku lze upravit na základě *Schwarzovy nerovnosti*, tj. vztahu

$$\left|\int_{-\infty}^{\infty}H(f)S(f)e^{j2\pi fT}df\right|^2 \le \int_{-\infty}^{\infty}\left|H(f)\right|^2 df \int_{-\infty}^{\infty}\left|S(f)\right|^2 df \tag{1.49}$$

S uvážením této nerovnosti vztah pro η může být upraven na tvar

$$\eta \le \frac{2}{N_0}\int_{-\infty}^{\infty}\left|S(f)\right|^2 df \tag{1.50}$$

Hodnota η je maximální, je-li dosaženo rovnosti. Z tohoto důvodu volíme přenos filtru tak, aby platilo

$$H_{opt}(f)=kS^*(f)e^{-j2\pi fT} \tag{1.51}$$

Zde $S^*(f)$ je funkce komplexně sdružená k funkci $S(f)$ a k je libovolná násobící konstanta.

Zpětná Fourierova transformace optimálního přenosu $H_{opt}(f)$ určuje jeho impulsní odezvu

$$h_{opt}(t)=ks(T-t) \tag{1.52}$$

Platí tedy následující závěr: *Impulsní odezva přizpůsobeného filtru je v čase obrácená a zpožděná verze vstupního signálu modifikovaná měřítkem k.*

Číslicový filtr na Obrázku 1.4 odpovídá diskrétní formě přizpůsobeného filtru. Jeho koeficienty jsou v souhlasu se vztahem pro odhad voleny tak, že odpovídají hodnotám vzorků signálu $s(t)$ (přesně pouze pro případ bílého šumu). Odezva na jednotkový impuls $\delta(n)$, která je analogií impulsní odezvy spojitých filtrů, má tvar vstupního signálu. Zpoždění o T je respektováno vzorkovacím obvodem na výstupu.

1.4. Odhad frekvence a fáze harmonického signálu

Předpokládejme sinusový signál o jednotkové amplitudě

$$s(t) = \sin(2\pi t + \varphi) \tag{1.53}$$

vzorkovaný v okamžicích t_i , $i = 1, 2, \ldots, N$.
Vzorkovaná data

$$s(t_i) = \sin(2\pi f t_i + \varphi) + e(t_i) \tag{1.54}$$

jsou zatížena aditivním šumem $e(t_i) = e_i$. Použitím inverzní funkce dostáváme $y_i = \sin^{-1}(s(t_i)) = \arcsin[\sin(2\pi f t_i + \varphi) + e_i]$

Taylorův rozvoj funkce arcsin y v okolí y_0 pro malé přírůstky je dán vztahem

$$\arcsin(y_0 + \Delta) = \arcsin y_0 + \Delta \frac{\mathrm{d}}{\mathrm{d}y}(\arcsin y)\Big|_{y=y_0} = \arcsin y_0 + \frac{\Delta}{\sqrt{1 - y_0^2}} \tag{1.55}$$

Vyberme interval, kde je sinusoida rostoucí (popř. klesající) a stanovme inverzní funkci sin $^{-1}(x)$ jako $\arcsin(x)$, popř. π - $\arcsin(x)$. Potom použitím rovnice (1.55) přibližně platí

$$\arcsin s(t_i) = 2\pi f t_i + \theta + \frac{e_i}{\cos(2\pi f t_i + \theta)} \tag{1.56}$$

kde fázový úhel θ odpovídá φ posunutému o násobek 2π.
Výchozí vztahu pro odhad má tvar (viz Příloha,(18))

$$\mathbf{y} = \mathbf{Mx} + \boldsymbol{\delta} \tag{1.57}$$

kde pro prvky vektoru $\boldsymbol{\delta}$ na základě vztahu (1.54) platí

$$\delta_i = \frac{e_i}{\cos(2\pi f t_i + \theta)} \tag{1.58}$$

Vektor $\boldsymbol{\delta}$ lze zapsat v maticovém tvaru

$$\boldsymbol{\delta} = \begin{vmatrix} 1/\cos(2\pi f t_1 + \theta) & 0 & 0 \\ 0 & . & 0 \\ 0 & 0 & 1/\cos(2\pi f t_n + \theta) \end{vmatrix} . \begin{vmatrix} e_1 \\ . \\ e_n \end{vmatrix} = \mathbf{w}.\mathbf{e} \tag{1.59}$$

Pro kovarianční matici rušení nyní platí

$$\operatorname{var} \boldsymbol{\delta} = \mathbf{w}^T (\operatorname{var} \mathbf{e}) \mathbf{w} = \sigma^2 \mathbf{V} \tag{1.60}$$

Členy diagonální matice $\mathbf{V}$ mají tvar

$$\left\{V_{ij}\right\}=\frac{1}{\cos^2(2\pi f t_{ii}+\theta)}=\frac{1}{W_i} \tag{1.61}$$

Vektor odhadovaných parametrů, tj. kmitočtu f a fáze θ má tvar

$$\mathbf{x}=\begin{vmatrix} \theta \\ 2\pi f \end{vmatrix} \tag{1.62}$$

Matice $\boldsymbol{M}$ v souladu s rovnicí (1.57) má rozměr $n \times 2$ a tvar

$$\mathbf{M}=\begin{vmatrix} 1 & t_1 \\ 1 & t_2 \\ . & . \\ . & . \\ 1 & t_n \end{vmatrix} \tag{1.63}$$

Jelikož matice $\mathbf{V}$ není jednotková, použijeme pro odhad obecného vztahu (P3-(23), v němž se uplatní členy W_i. Jejich účinkem se zmenšuje váha dat y_i z kritické oblasti kolem úhlu $\pi/2$, kdy použitá aproximace funkce $\arcsin(x)$ trpí značnou chybou (přídavným šumem jakoby nabývá $\sin x$ hodnot >1).

Vztah P3-(23) udává odhad vektoru $\mathbf{x}_{LMS}$ ve tvaru

$$\hat{\mathbf{x}}_{LMS}=(\mathbf{M}^T\mathbf{V}^{-1}\mathbf{M})^{-1}\mathbf{M}^T\mathbf{V}^{-1}\mathbf{y} \tag{1.64}$$

Určeme nyní postupně jednotlivé části tohoto vztahu

$$\mathbf{V}^{-1}=\frac{1}{\prod\limits_{i=1}^{N}\frac{1}{W_i}}\begin{vmatrix} \prod\limits_{i\neq 1}\frac{1}{W_i} & 0 & 0 \\ 0 & \prod\limits_{i\neq 2}\frac{1}{W_i} & 0 \\ 0 & 0 & \prod\limits_{i\neq N}\frac{1}{W_i} \end{vmatrix}=\begin{vmatrix} W_1 & 0 & ... & 0 \\ 0 & W_2 & ... & 0 \\ 0 & 0 & ... & W_N \end{vmatrix} \tag{1.65}$$

Pro další část vztahu platí

$$\left(\mathbf{M}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{M}\right)^{-1}=\left(\begin{vmatrix} 1 & 1\,... & 1 \\ t_1 & t_2\,... & t_N \end{vmatrix}\begin{vmatrix} W_1 & 0... & 0 \\ 0 & W_2... & 0 \\ 0 & 0... & W_N \end{vmatrix}\begin{vmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_N \end{vmatrix}\right)^{-1}=\begin{pmatrix} \sum\limits_{1}^{N} t_i W_i & \sum\limits_{1}^{N} t_i^2 W_i \\ \sum\limits_{1}^{N} W_i & \sum\limits_{1}^{N} t_i W_i \end{pmatrix}^{-1} \tag{1.66}$$

Dále platí

$$\left(\mathbf{M}^{\mathrm{T}}\mathbf{V}^{-1}\right)\mathbf{y} = \begin{vmatrix} W_1 & W_2 \dots & W_N \\ t_1 W_1 & t_2 W_2 \dots & t_N W_N \end{vmatrix} \begin{vmatrix} y_1 \\ y_2 \\ . \\ y_N \end{vmatrix} = \begin{vmatrix} \sum_1^N W_i y_i \\ \sum_1^N t_i W_i y_i \end{vmatrix} \tag{1.67}$$

Po dosazení dílčích výrazů do výchozího vztahu dostáváme konečný výraz pro odhad fáze a frekvence

$$\begin{vmatrix} \hat{\theta} \\ 2\pi \hat{f} \end{vmatrix} = \frac{1}{\sum_1^N W_i \sum_1^N t_i^2 W_i - \left(\sum_1^N t_i W_i \right)^2} \begin{vmatrix} \sum_1^N t_i^2 W_i \sum_1^N W_i y_i & -\sum_1^N t_i W_i \sum_1^N t_i W_i y_i \\ \sum_1^N t_i W_i \sum_1^N W_i y_i & +\sum_1^N W_i \sum_1^N t_i W_i y_i \end{vmatrix} \tag{1.68}$$

Pro varianci odhadu platí

$$\operatorname{var} \hat{\mathbf{x}} = \sigma^2 (\mathbf{M}^T \mathbf{V}^{-1} \mathbf{M})^{-1} = \frac{\sigma^2}{\sum_1^N W_i \sum_1^N t_i^2 W_i - \left(\sum_1^N t_i W_i \right)^2} \begin{vmatrix} \sum_1^N t_i^2 W_i & -\sum_1^N t_i W_i \\ -\sum_1^N t_i W_i & \sum_1^N W_i \end{vmatrix} \tag{1.69}$$

Odhadovaná hodnota frekvence je rovna

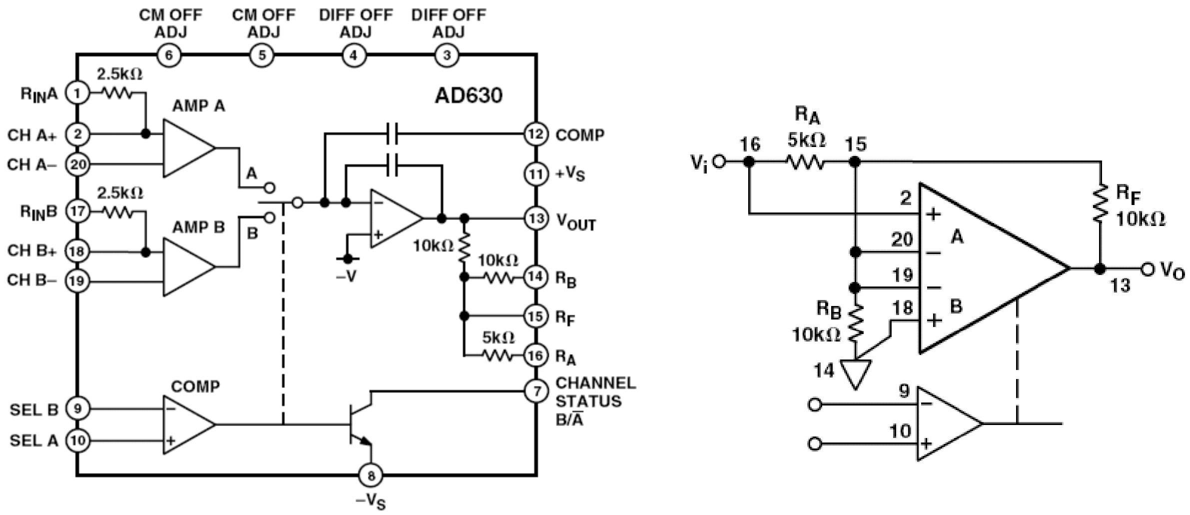
$$2\pi \hat{f} = \frac{\sum_1^N W_i \sum_1^N t_i W_i y_i - \sum_1^N t_i W_i \sum_1^N W_i y_i}{\sum_1^N W_i \sum_1^N t_i^2 W_i - \left(\sum_1^N t_i W_i \right)^2} \tag{1.70}$$

2. Obvody a přístroje využívající principu koherence.

2.1. Analogové koherentní demodulátory

2.1.1. *Obvodová řešení*

Za současného stavu moderní součástkové základny se pro účely koherentní demodulace využívá integrovaných obvodů, z nichž nejznámější jsou obvody AD630 a AD698.

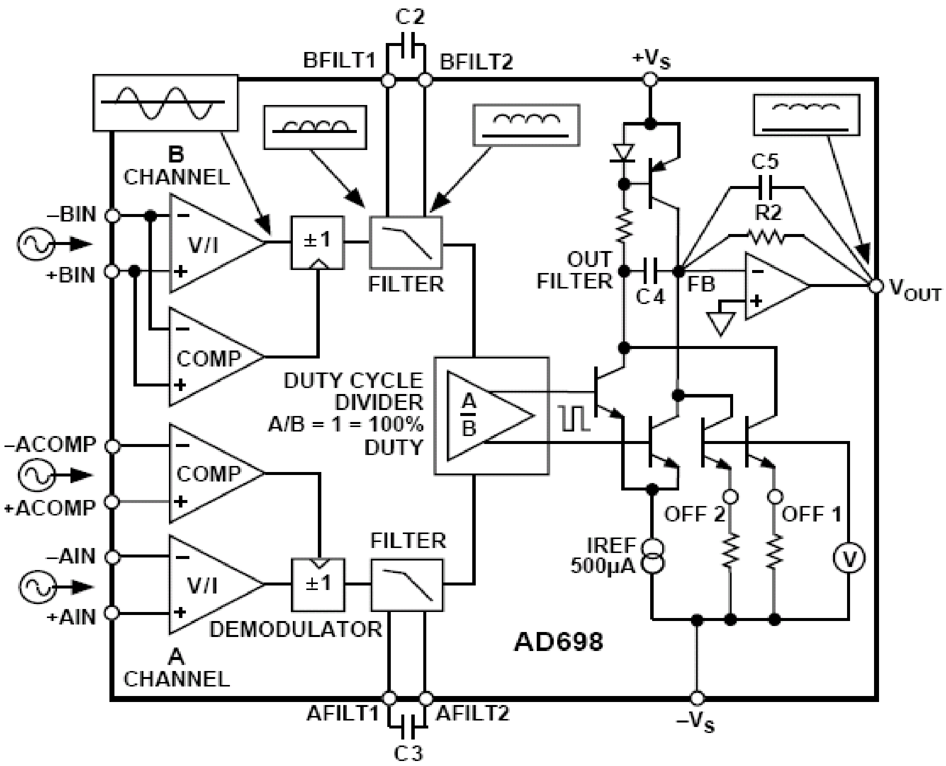


Obrázek 2.1. Vyvážený modulátor – demodulátor AD 630 (Analog Devices)

Kromě nominální aplikace tj. vyvážený modulátor a demodulátor je obvod AD 630 použitelný pro koherentní demodulaci, jednoduchý LIA, fázově citlivý detektor, a násobičku pravoúhlým signálem.

V rytmu referenčního napětí jsou na vstup koncového zesilovače střídavě připojovány výstupy zesilovačů A nebo B. Celý obvod pak působí jako násobička vstupního signálu střídavě hodnotami $-R_F/R_A$ (vybrán B) nebo $1 + R_F/R_B$ (vybrán A). Pro hodnoty vestavěných odporů uvedených na schématu jde o násobení *stálými* hodnotami (oba vestavěné odpory na stejné teplotě !) +2 a -2. Praktické zkušenosti ukazují jisté problémy vyvolané změnou vstupního odporu při přepínání z režimu invertujícího do neinvertujícího. Výstupní odpor zdroje signálu zapojeného na svorku V_i by měl být podstatně menší než odpor R_A.

2.1.2. *AD 698 Obvod pro zpracování signálů z lineárního diferenciálního transformátoru (LVDT)*



Obrázek 2.2.. Integrovaný obvod se dvěma KD vhodný pro poměrová měření

Struktura obvodu je koncipovaná tak, aby použitím metody *poměrového měření* se vyloučila závislost výstupního signálu na kolísání budícího napětí na primárním vinutí transformátoru (vstup B). Poloze jádra úměrné rozdílové napětí na diferenciálním sekundárním vinutí transformátoru je připojeno na vstup A. Koherentní demodulátory, znázorněné jako násobičky hodnotami +/-1 a filtry, dávají na výstupu filtrů stejnosměrné signály úměrné signálům A a B. Výpočet poměru A/B, nezávislého na buzení, obstarává podílový obvod pracující na impulzním principu (mění se činitel plnění impulzního signálu). Pravoúhlé referenční signály pro demodulátory jsou odvozeny z interního oscilátoru (na obrázku nezakreslen) pomocí komparátorů (COMP).

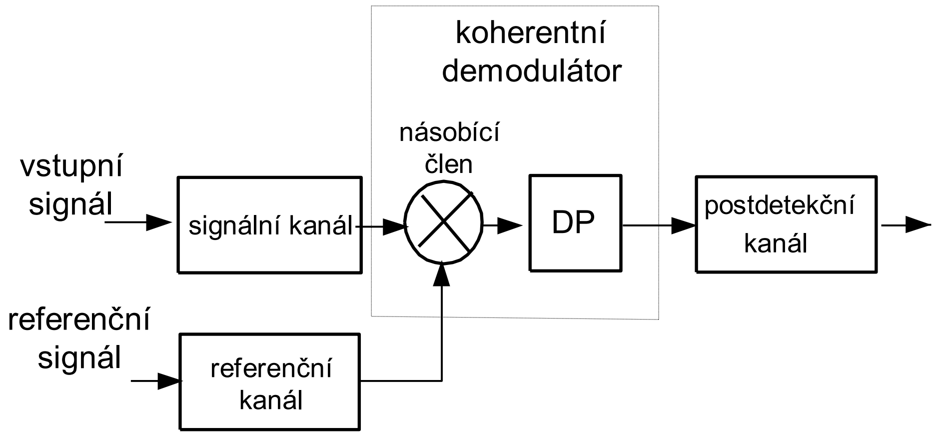
Obvod je vhodný i pro střídavá můstková měření impedancí, signál B je úměrný napájecímu napětí můstku a signál A je odvozen z indikátorové úhlopříčky. Výhoda potlačení závislosti na budícím napětí samozřejmě zůstává.

2.2. Analogové přístroje pro koherentní demodulaci (Lock-in Amplifiers-LIA)

2.1.3. *Historie přístrojů pro koherentní demodulaci*

Jedinečná schopnost koherentní demodulace zpracovat i signály ležící hluboko pod úrovní šumu má zásadní význam pro zpracování signálů v měřicí technice. Tvoří podstatu přístrojů označovaných jako Lock-in Amplifiers (LIA) a patřících do vybavení většiny fyzikálních laboratoří. Za původce prvních typů LIA se považuje tým pracovníků vědecko-výzkumných laboratoří Princeton Applied Research – PAR. První komerčně dostupný LIA z dílny PAR se objevil na trhu v r. 1962. První zmínka o koherentní demodulaci a jejím využití pochází z článku Waltera C. Michelse z r. 1941 (*Review of Scientific Instruments*), v němž se poprvé používá označení LIA. V článku jsou také citované první práce o využití koherentní demodulace publikované již v r. 1934, tj. 30 let před založením PAR. Převážná většina přístrojů patřících do kategorie LIA byla v minulém století realizována analogovou technikou. Pokroky součástkové základny moderní polovodičové elektroniky dovolují v posledních 20-ti letech realizovat vysoce kvalitní LIA téměř výhradně číslicovými metodami.

2.1.4. *Základní části koherentních demodulátorů*



Obrázek 2.3. Základní části přístrojové realizace koherentní demodulace

Přístroje používající principu odhadu parametrů signálů splňujících podmínku koherence jsou složeny ze čtyř základních částí, tj.:

- předdetekčního zpracování signálu (signální kanál)

- vlastního koherentního demodulátoru (varianty násobícího obvodu s integračním článkem DP na výstupu)
- zpracování referenčního signálu (referenční kanál)
- postdetekčního zpracování signálu

2.1.5. *Předdetekční zpracování signálu*

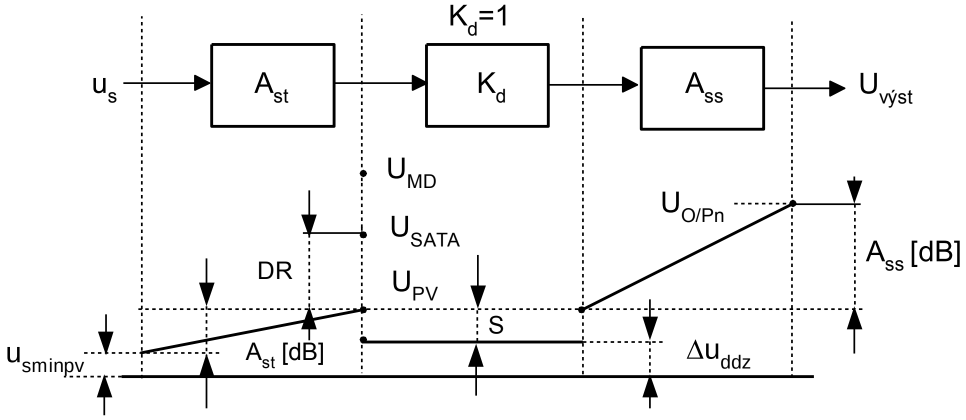
Hlavním úkolem této části je dosáhnout požadovaného poměru užitečného výstupního signálu U_{PV} k rušení (driftu) koherentního demodulátoru a obvodů pro postdetekční zpracování a současně zajistit, aby přístroj mohl bez následků zpracovávat co největší hodnotu rušivého signálu (tzv. *dynamická rezerva*).

Pod dynamickou rezervou (DR) se rozumí největší úroveň rušivého signálu, která vyvolá změnu výstupní úrovně menší než je dovolená chyba. DR se udává v násobcích úrovně užitečného signálu U_{PSV}, nutného pro plnou výchylku.

Vztah dynamické rezervy k základním parametrům obvodové struktury, tj. zisku předdetekčního řetězce A_{st}, přenosu koherentního demodulátoru K_D (zpravidla K_D = 1) a zisku postdetekčního kanálu, je znázorněn na Obrázku 2.4. V obrázku značí:

Δu_{DDZ} drift vlastního demodulátoru a postdetekčního zesilovače,
U_{PV} úroveň napětí na vstupu KD nutná pro plnou výchylku ,
$s = U_{PV}/\Delta u_{DDZ}$ činitel stálosti výstupního napětí,
U_{SATA} hranice linearity převodní charakteristiky zesilovače,
U_{MD} hranice linearity převodní charakteristiky.

Nejznámější výrobci přístrojů na principu koherentních demodulátorů (*lock-in amplifier*), např. EG&G, ITHACO, SR udávají dovolenou úroveň asynchronního signálu U_{ASM}, která po superpozici s užitečným signálem o úrovni nutné pro plnou výchylku vyvolá změnu výstupního údaje o méně než 5 %.



Obrázek 2.4. K výkladu dynamické rezervy

Z úrovňového diagramu je zřejmá důležitost minimalizovat hodnotu Δu_{DDZ}, ačkoliv by se při povrchním pohledu zdálo, že vzhledem ke značným úrovním zesíleného signálu na vstupu koherentního demodulátoru nejsou jeho vlastnosti pro malé signály kritické.

Dynamická rezerva, definovaná jako poměr

$$DR = \frac{U_{SATA}}{U_{PV}} \tag{2.1}$$

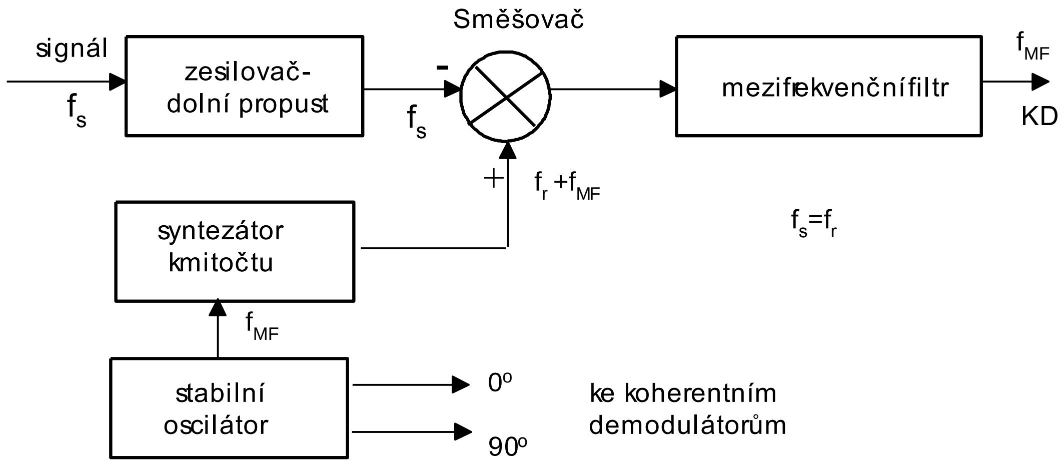
dosahuje u kvalitních přístrojů hodnot kolem 120 dB a je pro některé aplikace základním parametrem. Je tomu tak např. pro jistou skupinu magnetických měření, kdy užitečná informace je nesena druhou harmonickou signálu, jehož základní harmonická, představující v této situaci rušivý signál, má úroveň až o několik řádů vyšší.

Neméně významná je hodnota DR v obecném případě obnovy signálu s úrovní hluboko pod rušením. Proto je prakticky celá koncepce signálního kanálu podmíněna minimalizací účinků rušení na výstupní signál, tj. vedena snahou zvýšit hodnotu DR.

Nejúčinnější jsou přístupy založené na zmenšování úrovně rušivého signálu na výstupu zesilovače filtrací rušení. Jelikož střední kvadratická hodnota rušení je úměrná šumové šíři pásma $B_{š}$, měl by být signální kanál koncipován jako úzkopásmový obvod. To je ovšem v rozporu s požadavkem provozu přístroje v širokém rozsahu kmitočtů, jelikož *úzkopásmová propust vykazuje značné změny fáze i při relativně malých odchylkách kmitočtu signálu* a referenčního průběhu. Přesně reprodukovatelné přelaďování úzkopásmového filtru, nutné při změnách kmitočtu signálu, by bylo při obvyklých metodách ladění velmi obtížně realizovatelné. Východiskem je *heterodynní filtrace*.

2.1.6. *Heterodynní filtrace*

Heterodynní filtrace využívá směšovacího principu. Jak je ukázáno na Obrázku 2.5, je ve směšovači zpracováván zesílený vstupní signál o frekvenci f_S a signál z referenčního kanálu s frekvencí $f_S + f_{MF}$. Výstupní signál směšovače na rozdílové, mezifrekvenční frekvenci f_{MF} je úzkopásmově zesílen mezifrekvenčními obvody. V referenčním kanálu je, jak bude dále vysvětleno, generován smyčkou fázového závěsu signál o frekvenci o f_{MF} vyšší, než je frekvence signálu f_S. Dodržení podmínky koherence vyžaduje, aby platila rovnost $f_S = f_r$, nebo obecně musí být hodnota frekvence f_s předem známá.



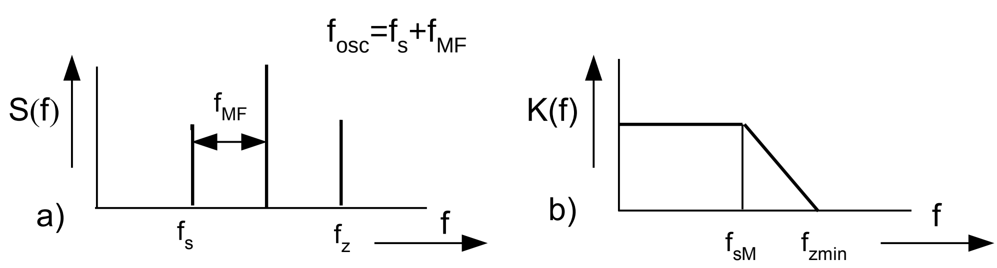
Obrázek 2.5. Signální kanál s heterodynní filtrací

Hlavní výhody principu heterodynní filtrace jsou tyto:

- koherentní demodulátor pracuje na jedné frekvenci f_{MF}, proto je snadné optimalizovat, po případě korigovat, jeho vlastnosti,
- selektivitu lze podstatně zvýšit např. pevně laděnými piezokeramickými filtry s ekvivalentním činitelem jakosti Q řádově 10^4, tj. lze dosáhnout šíře propustného pásma řádově 100 Hz na střední frekvenci 10 MHz. Pozor však na změny fáze při kolísání kmitočtu !,
- na stálé frekvenci f_{MF} lze snadno generovat referenční signály navzájem fázově posunuté o $\pi/2$,
- při správném výběru frekvence f_{MF} se vyšší harmonické signálu neprojeví na výstupu KD.

Nevýhody heterodynní filtrace jsou následující:

- větší složitost,
- možnost vzniku rušivých signálů na *nelinearitách směšovače*, tj. signálů na frekvencích odpovídajících lineární kombinaci frekvencí, se kterými právě obvod směšovače pracuje,
- rušivé působení *zrcadlových frekvencí*.



Obrázek 2.6 K potlačení zrcadlových frekvencí

K potlačení zrcadlových frekvencí slouží dolnofrekvenční propust DP před směšovačem (Obrázek 2.6). Její parametry jsou navrženy tak, aby i v nejméně příznivém případě, kdy je zrcadlová frekvence minimální, byla úroveň signálu na této frekvenci před směšovačem menší než přípustná.

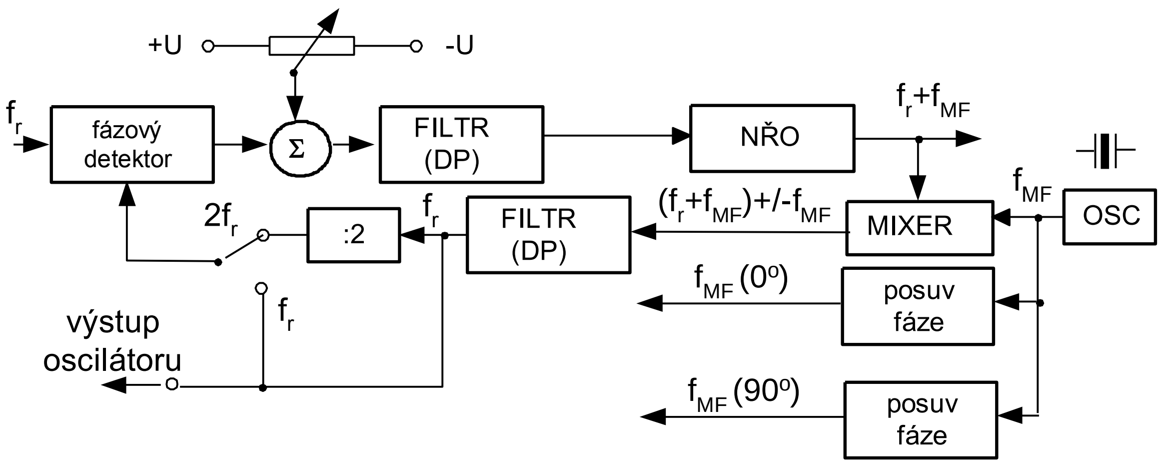
Jelikož platí (Obrázek 2.6 a)

$$f_{z\,min} = f_{s\,min} + 2 f_{MF} \tag{2.2}$$

kde f_{smin} je minimální frekvence signálu, jsou nároky DP tím menší, čím je mezifrekvenční frekvence f_{MF} větší. Tu však nelze příliš zvyšovat, jelikož při vyšších frekvencích by bylo obtížné zajistit správnou funkci vlastního KD (násobičky nebo spínače).

2.1.7. *Referenční kanál*

Typická architektura referenčního kanálu přístrojů pro koherentní demodulaci je na Obrázku 2.7.



Obrázek 2.7 Referenční kanál koherentního demodulátoru

Jeho hlavním úkolem je zajistit splnění *podmínky koherence* funkcionálu signálu $s(t)$ a referenčního průběhu, a to i při změnách frekvence signálu v dovolených mezích. Obecně to vyžaduje apriorní znalost těch parametrů signálu, z nichž lze odvodit jeho frekvenci a fázi. V obecné situaci zcela neznámého signálu je zřejmě tato úloha neřešitelná. Avšak správnou strategií návrhu experimentu je v převážné většině případů možné upravit celý pokus tak, aby referenční průběh $s(t)$ byl znám předem. Jednoduchým příkladem je měření veličin typu záření, kdy místo přímého měření účinků záření na objekt *modulujeme záření* průběhem $s(t)$, např. jeho přerušováním rotující clonkou.

Splnění podmínky koherence je snadné u můstkových měření, kdy napájecí napětí můstku lze přímo odvodit z generátoru v referenčním kanálu koherentního demodulátoru.

Základem je smyčka fázového závěsu (PLL) se směšovačem ve zpětnovazební větvi. Výstupní signál napětím řízeného oscilátoru NŘO o frekvenci $f_r + f_{MF}$ představuje signál "místního oscilátoru" pro směšovač heterodynního filtru v signálním kanálu. Stabilita frekvence výstupního signálu smyčky je odvozena z krystalem řízeného oscilátoru, pracujícího na jediné frekvenci f_{MF}. Vstupní informace o frekvenci signálu $s(t)$ pro fázový detektor PLL, na základě které má být zajištěna koherence, je po případných tvarových a úrovňových úpravách získána z experimentu, nejčastěji ze zdroje můstkového napájení nebo modulátoru zářivého toku. V případech, kdy volba frekvence f_s není z hlediska charakteru pokusu kritická, lze pro účely experimentu použít přímo výstupní napětí za směšovačem referenčního kanálu. Při tomto tzv. *interním* režimu je NŘO předladěn napětím ovládaným z čelního panelu na požadovanou hodnotu f_s a úloha PLL se zredukuje na generaci signálu o frekvenci $f_r + f_{MF}$.

V *externím* režimu lze z podstaty činnosti PLL očekávat, že i při krátkodobých výpadcích informace o frekvenci signálu bude dále generován signál o frekvenci blízké požadované hodnotě.

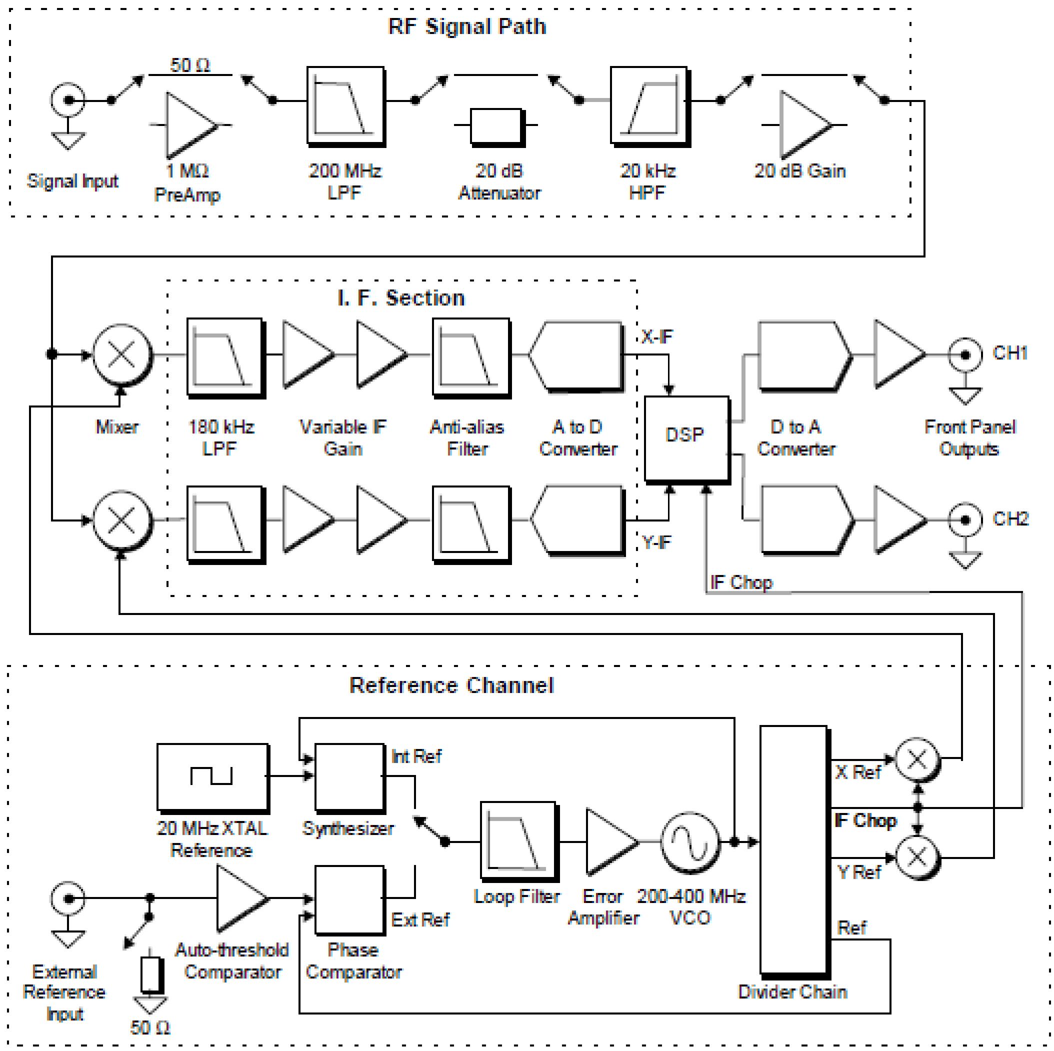
Dále lze díky vlastnostem PLL snadno zařídit, aby frekvence f_r byla celočíselným násobkem (harmonickou) základní frekvence f_s. Toho se dosahuje zařazením děliče frekvence do větve zpětné vazby PLL. Patrně nejznámější je již zmíněná situace v jisté třídě magnetických měření, kdy užitečná informace je nesena druhou harmonickou signálu. Pro tento případ je v uspořádání na Obrázku 2.7 použita dělička frekvence modulo 2.

Jednoduché obvody PLL nevyhoví při větších nárocích na rozsah změn frekvence signálu. Pak je nutné referenční kanál koncipovat jako frekvenční syntezátor. Naopak u jednoduchých aplikací (např. levné jednoúčelové přístroje) se referenční kanál zredukuje na oscilátor a tvarovací obvody.

Vyžaduje-li příslušná aplikace dva ortogonální referenční signály, doporučuje se použít zapojení tzv. *stavových oscilátorů* (metodami analogových počítačů "řeší" rovnici oscilátoru), u nichž jsou snadno dostupné dva signály s fází 0° a 90° nezávislou na změnách frekvence.

2.1.8. *Příklad koncepce moderního heterodynního číslicového LIA (SR844)*

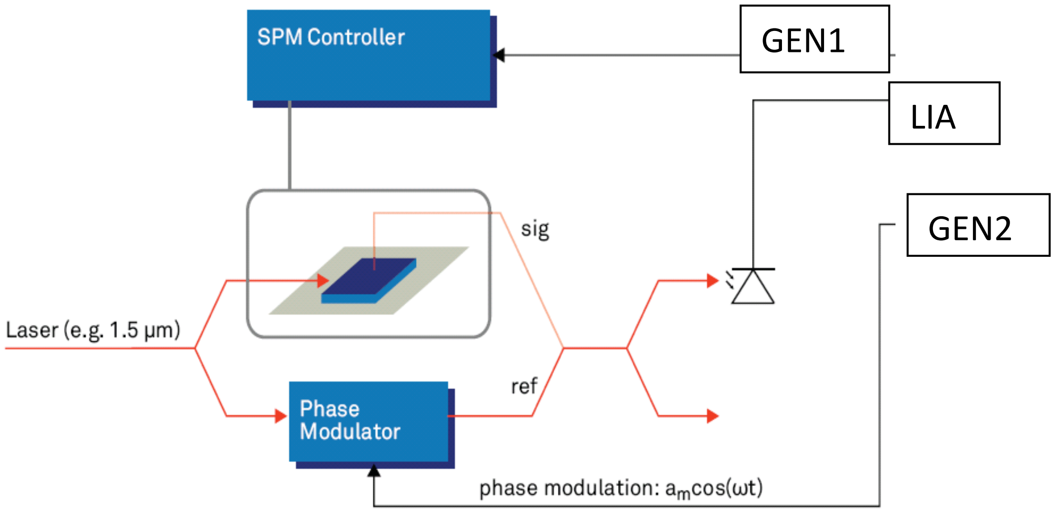
Koncepce odpovídá heterodynnímu demodulátoru podle Obrázku 2.5. Rozdíl spočívá zejména v použití dvou směšovačů s referenčními signály (oscilátory) posunutými navzájem o 90° generovanými v referenčním kanálu. Vstupní signál je ve směšovačích převeden na mezifrekvenční kmitočet 180 kHz. Po číslicově realizované koherentní demodulaci (číslicový součin) se na výstupu DSP získá soufázová a kvadraturní složka vstupního signálu.



Obrázek 2.8 Obvodová struktura moderního heterodynního LIA pro frekvence do 200 MHz

2.1.9. *Pseudo-heterodynní koncepce*

Frekvence vstupního a referenčního signálu jsou stejné, avšak moduluje se fáze referenčního signálu (Obrázek 2.9). Tato koncepce se používá často u optických interferenčních měření. Fáze referenčního svazku je elektrooptickým modulátorem měněna podle vhodného časového průběhu. Výhodou je podstatně vyšší citlivost ve srovnání s jednoduchým interferometrickým uspořádáním a střídavý výstupní signál demodulátoru snadno dále upravitelný střídavými zesilovači.



Obrázek 2.9 Aplikace pseudoheterodynní metody pro laserovou koherentní interferometrii. Zařízení je určeno pro vizualizaci optických vlastností planárních vlnovodných struktur. Polohu objektu řídí SPM kontrolér, fázový modulátor deformuje referenční optické vlákno piezostrikčním elementem.

2.1.10. *Obvody vlastního koherentního demodulátoru*

Ideální koherentní demodulátor by měl realizovat matematické operace vyskytující se ve vztazích pro odhad amplitudy signálu. V praxi se v přístrojovém provedení používají *suboptimální algoritmy* vhodné pro rušení typu bílý šum, kde pak základním modelem je analogová násobička s filtrací výstupního signálu dolnofrekvenční propustí.

I v případě bílého šumu by však respektování principu přizpůsobeného filtru vyžadovalo úpravu tvaru referenčního signálu s ohledem na průběh $s(t)$ v dané aplikaci. V praxi se nejčastěji vyskytují situace s harmonickým průběhem $s(t)$ - např. můstková měření, nebo s obdélníkovým signálem, např. signály ze senzorů vyvolané modulovaným zářením. V obou případech by nejlépe vyhověla analogová násobička s výstupním signálem úměrným součinu vstupního průběhu $V(t)$ a obecné nosné signálu $s(t)$ s harmonickým nebo obdélníkovým průběhem.

2.1.11. *Ekvivalentní amplitudová frekvenční charakteristika KD*

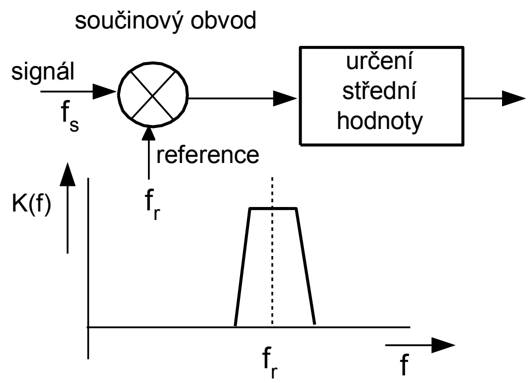
Ekvivalentní amplitudová frekvenční charakteristika (AFCH) koherentního demodulátoru se liší od obvykle užívaného pojetí pro lineární obvody, v němž každé frekvenci vstupního signálu odpovídá výstupní signál na stejné frekvenci, avšak s jinou amplitudou a fází. Jak plyne z principu KD, pod AFCH se zde rozumí vyjádření oblastí frekvencí, které vyvolají příspěvek k výstupnímu signálu za dolnofrekvenční propustí demodulátoru.

Spektrum výstupního signálu násobičky je dáno konvolucí Fourierovy transformace spektra vstupního signálu $v(t)$ a referenčního průběhu $s(t)$, takže platí

$$\boldsymbol{Y}(j\omega) = \boldsymbol{V}(j\omega) * \boldsymbol{S}(j\omega) \tag{2.3}$$

kde * znamená konvoluci.

Je-li $s(t)$ harmonický signál, pak $\boldsymbol{S}(j\omega) = \boldsymbol{S}_0 \delta(j\omega - j\omega_r)$ je diskrétní spektrální čára a ekvivalentní amplitudová frekvenční charakteristika, Obrázek 2.10, má tvar pásmové propusti s průběhem kopírujícím přenos dolní propusti $\boldsymbol{K}_{DP}(j\omega)$ zrcadlově symetricky kolem frekvence referenčního průběhu f_r.



Obrázek 2.10 Amplitudová frekvenční charakteristika ideálního KD

Takovýto průběh je jistě výhodný z hlediska minimalizace šumové šíře pásma.

2.1.12. *Šumová šíře pásma ideálního KD*

Šumová šíře pásma soustavy s přenosem $H(j\omega)$ je definována vztahem

$$\Delta f = \frac{1}{2\pi} \int_0^\infty \left| \frac{H(j\omega)}{H_{max}} \right|^2 d\omega \tag{2.4}$$

v němž H_{max} je maximální hodnota přenosové funkce $H(j\omega)$. Jde o náhradu průběhu $H(j\omega)$ pravoúhlou pásmovou propustí o šíře pásma Δf přenášející stejný výkon šumu.

Obvodová realizace násobičky, vyhovující požadavkům na dynamickou rezervu, je obtížná, jelikož při větších signálech se velmi nepříznivě uplatní její nelinearity. Vyšší mocniny, popisující nelineární převodní charakteristiku reálné násobičky, vyvolávají vznik rušivých signálů zázněji mezi složkami šumu a signálu. Tak dochází k již zmíněnému jevu potlačení signálu šumem. Proto jsou koherentní demodulátory s analogovou násobičkou vhodné pro případy, kdy je primárním cílem regenerace signálu ze šumu a přitom maximální úrovně rušení jsou relativně malé ve srovnání s pásmem linearity obvodů násobičky.

2.1.13. *Spínačové koherentní demodulátory*

Periodicky pracující spínač představuje obvod s časově proměnným přenosem, ovládaným referenčním signálem. Výstupní napětí za spínačem lze popsat jako *součin* průběhu vstupního signálu a v rytmu referenčního signálu proměnné spínací funkce $K(t)$, tj.:

$$y(t) = v(t).K(t) \tag{2.5}$$

a jeho spektrum je dáno konvolucí

$$\boldsymbol{Y}(j\omega) = \boldsymbol{V}(j\omega) * \boldsymbol{K}(j\omega) \tag{2.6}$$

Spektrum $\boldsymbol{K}(j\omega)$ odpovídá pravoúhlému tvaru spínací funkce $K(t)$

$$\boldsymbol{K}(j\omega) = \sum_{k=-\infty}^{k=\infty} \boldsymbol{S}_k \delta(j\omega - jk\omega_0) \tag{2.7}$$

Pak platí

$$\begin{aligned} \boldsymbol{Y}(j\omega) &= \int_{-\infty}^{\infty} \boldsymbol{V}(j\nu) \sum_{k=-\infty}^{k=\infty} \boldsymbol{S}_k \delta(j\omega - jk\omega_0 - j\nu) d\nu = \\ &\sum_{k=-\infty}^{k=\infty} \boldsymbol{S}_k \int_{-\infty}^{\infty} \boldsymbol{V}(j\nu) \delta(j\omega - jk\omega_0 - j\nu) d\nu \end{aligned} \tag{2.8}$$

a spektrum na výstupu demodulátoru (za dolnofrekvenční propustí) je dáno vztahy

$$\boldsymbol{Z}(j\omega) = \boldsymbol{Y}(j\omega)\boldsymbol{K_D}(j\omega) = \sum_{k=-\infty}^{k=\infty} \boldsymbol{S}_k \boldsymbol{K_D}(j\omega) \int_{-\infty}^{\infty} \boldsymbol{V}(j\nu) \delta(j\omega - jk\omega_0 - j\nu) d\nu \tag{2.9}$$

Pro další úpravu využijeme vztahů platných pro δ funkci

$$\delta(-\tau) = \delta(\tau), \quad \int_{-\infty}^{\infty} f(t)\delta(t - t_0) = f(t_0) \tag{2.10}$$

Integrál po úpravě se redukuje na vztah

$$\int_{-\infty}^{\infty} \boldsymbol{V}(j\nu)\delta(j\omega - jk\omega_0 - j\nu) d\nu = \boldsymbol{V}(j\omega - jk\omega_0) \tag{2.11}$$

a po dosazení dostáváme

$$\boldsymbol{Z}(j\omega) = \sum_{k=-\infty}^{k=\infty} \boldsymbol{S}_k \boldsymbol{K_D}(j\omega)\boldsymbol{V}(j\omega - jk\omega_0)\boldsymbol{K_D}(j\omega) \tag{2.12}$$

Po zavedení substituce platí

$$\omega - k\omega_0 = \Omega, \quad \boldsymbol{Z}(j\Omega + jk\omega_0) = \sum_{k=-\infty}^{k=\infty} \boldsymbol{S}_k \boldsymbol{K_D}(j\Omega + jk\omega_0)\boldsymbol{V}(j\Omega) \tag{2.13}$$

K určení příspěvků spektra vstupního signálu předpokládejme, že $\boldsymbol{V}(j\Omega) = V = \text{konst.}$ Pak platí

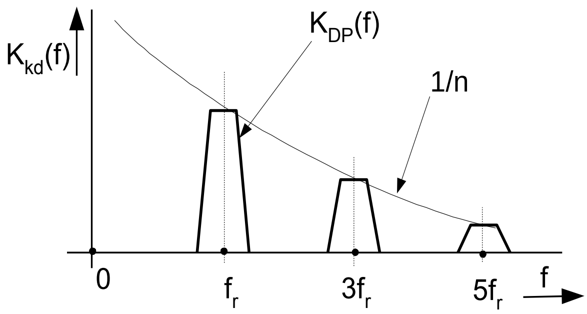
$$\boldsymbol{Z}(j\Omega + jk\omega_0) = V \sum_{k=-\infty}^{k=\infty} \boldsymbol{S}_k \boldsymbol{K_D}(\Omega + jk\omega_0) \tag{2.14}$$

Po záměně sčítacích indexů $k = -l$ dostaneme výsledný vztah pro ekvivalentní přenos koherentního demodulátoru

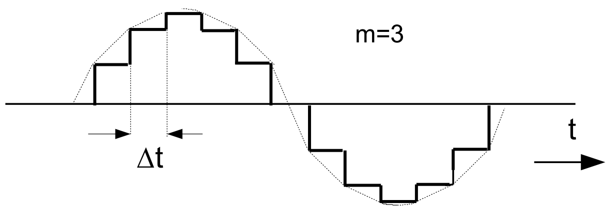
$$\boldsymbol{Z}(j\Omega - jl\omega_0) = V \sum_{l=\infty}^{l=-\infty} \boldsymbol{S}_k \boldsymbol{K_D}(j\Omega - jl\omega_0) \tag{2.15}$$

Výsledná AFCH spínačového KD na Obrázku 2.11 má tvar *hřebenového filtru* složeného z kopií AFCH dolní propusti $K_{DP}(f)$, symetricky rozložených kolem harmonických $K(j\omega)$.

Spínačový KD má vlastnosti *přizpůsobeného filtru* pro signály s pravoúhlým průběhem nosné. To lze snadno ukázat také z AFCH, která má propustná okna kolem všech harmonických $K(t)$ přispívajících k výstupnímu signálu.



Obrázek 2.11 AFCH spínačového KD



Obrázek 2.12 Stupňovitá aproximace sinusovky

Z hlediska konstrukčního lze u spínačových KD dosáhnout největší hodnoty dynamické rezervy, jelikož se u spínačových obvodů principiálně snáze dosahuje stálosti přenosové charakteristiky $K(t)$ a to i při velkých signálech na vstupu spínače.

Pro omezení průniku řídicího signálu na výstup spínače se pro nejvyšší nároky používá galvanicky odděleného zdroje řídicího napětí, téměř výhradně optoelektronické vazby.

Vlastnosti spínačových KD jsou značně závislé na stálosti přepínací funkce $K(t)$ a v nejčastějším případě dvoucestného spínaní na její přísné symetrii. Nesymetrie je způsobena zejména kolísáním doby přepnutí Δt_a mezi oběma krajními hodnotami +1 a -1 průběhu $K(t)$. Nesymetrie způsobuje vznik stejnosměrné složky, která při změnách Δt_a vlivem teploty a především amplitudy signálu působí jako přídavný drift a tím zmenšuje dosažitelnou hodnotu DR. Např. požadujeme-li stálost nuly KD rovnou 0,01%, pak musí být kolísaní Δt_a < 10^{-4} T, takže při T = 0,1 ms nesmí kolísat spínací časy o více než 10 ns. Proto i při poměrně nízkých kmitočtech desítky kHz jsou nároky na rychlost spínačů mimořádné.

2.1.14. *Potlačení vlivu vyšších harmonických spínací funkce*

Spínačový KD je vzhledem k hřebenovému tvaru AFCH nevýhodný pro zpracování harmonických signálů pod úrovní šumu. Potlačení vlivu příspěvků na vyšších harmonických se dosahuje:

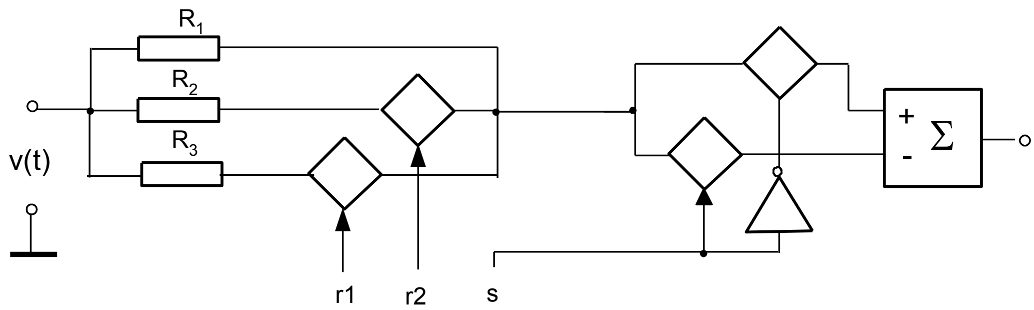
a) **Realizací *násobícího obvodu na principu převodníku Č/A*** s nízkou rozlišovací schopností. Zde se vychází ze skutečnosti, že tvar $K(t)$, který má nyní průběh odpovídající aproximaci sinusové závislosti stupňovitým tvarem, má ve svém spektru harmonické na kmitočtech

$$f_n = (4mn \pm 1) f, \quad n = 1,2,\ldots \quad f = 1/T \,. \tag{2.16}$$

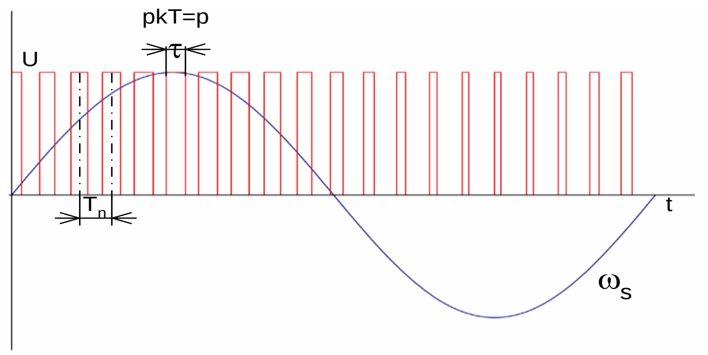
Zde m je počet úrovní aproximačního stupňovitého průběhu za čtvrtinu periody (Obrázek 2.12).

Je zřejmé, že i pro hrubou aproximaci sinusovky podle Obrázku 2.12 (m = 3) se nejbližší propustné pásmo nachází na 11-té harmonické a ne na třetí, jako u klasického spínačového KD. Poměrně jednoduše realizovatelný 3-bitový převodník Č/A (Obrázek 2.13) vyžaduje pouze 3 přesné odpory. Ovládací logické signály r_1, r_2 a s se odvodí zpožděním hodinových signálů v posuvném registru, takže ani řízení jejich fáze vzhledem k signálu není obtížné. Z hlediska dynamických vlastností je výhodné použít proudových spínačů.

b) ***Impulsní šířkovou modulací***. U těchto spínačových KD se časový interval sepnutí spínače šířkově moduluje harmonickým signálem o kmitočtu signálu f_s (Obrázek 2.14).



Obrázek 2.13 Realizace spínačového KD násobícím převodníkem Č/A



Obrázek 2.14 Průběh impulsní šířkové modulace, užívané pro konstrukci spínačových KD

Fourierova řada popisující tento průběh má tvar

$$s(t) = kU + kUp\sin\omega_s t + \sum_{n=-\infty}^{n=\infty} \frac{2}{\pi}\sin\left[n\pi(kU + kpU\sin\omega_s t)\right]\cos n\Omega_n t \qquad (2.17)$$

kde $T_n = 1/\Omega_n$ je perioda nosného kmitočtu, kT_n je základní šíře impulsu při nulové hodnotě modulačního činitele p, proměnného v mezích $0 < p < 1$.

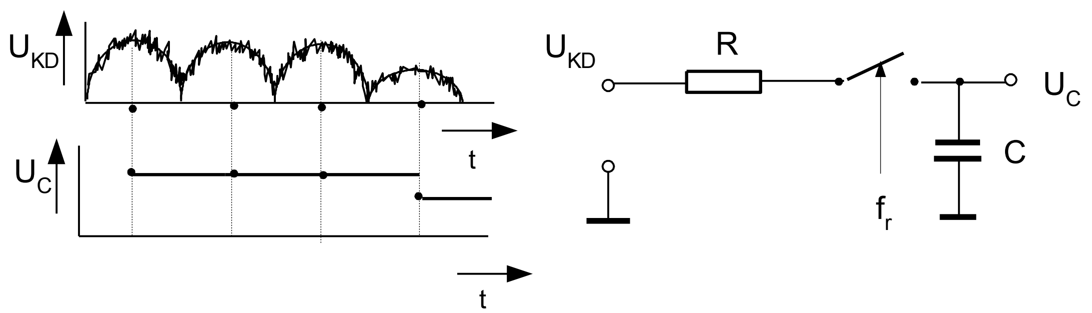
Základní kmitočet impulsní nosné Ω_n je mnohonásobně vyšší než f_s, takže nároky na spínací časy se zvýší. Hlavním důvodem aplikace tohoto složitého postupu je opět zlepšení "selektivity", jelikož nejbližší propustné pásmo KD leží v okolí kmitočtu $\Omega_n - \omega_s$. To lze usoudit ze vztahu pro spektrum impulsní šířkové modulace.

K vytváření signálů složitějších pravoúhlých tvarů lze využít poznatků syntézy signálů na základě *Walshových funkcí* (skládání signálu z elementárních binárních funkcí místo klasické syntézy s harmonickými elementárními funkcemi, jak je užívá Fourierova analýza). Hlavní výhodou zůstává, že pro realizaci vystačíme se spínači, takže nároky na přesnost se přesouvají na snazší oblast dodržování časových intervalů a ne na stálost hodnot analogových prvků.

2.1.15. *Postdetekční obvody*

Jak bylo uvedeno při rozboru dynamické rezervy DR, započítává se drift a jakékoliv nestálosti dolnofrekvenční propusti a případně následujícího stejnosměrného zesilovače jako ekvivalentní část driftu Δu_{DDZ}. Ačkoliv by se na první pohled zdálo, že vzhledem k vyšším úrovním detektorem zpracovávaných signálů hlediska driftu již nejsou závažná, je nutné navrhovat postdetekční obvody i z hlediska minimálního driftu.

Dolnofrekvenční propust má být navržena jako kompromis mezi potlačením střídavých složek po detekci a rychlostí reakce na změnu amplitudy signálu. Tento rozpor požadavků je možné zmírnit *"spínanou" propustí* (Obrázek 2.15), v níž je kapacita C_H nabíjena na hodnotu výstupního signálu za spínačovým KD v přesně definovaném okamžiku maxima usměrněného průběhu.

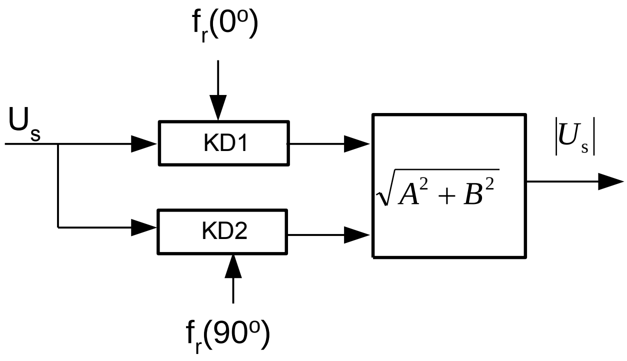


Obrázek 2.15 Použití vzorkovacího obvodu jako dolní propusti pro KD

Hlavní výhodou je, že skoková změna amplitudy vyhodnocovaného signálu je zjistitelná již za dobu půlperiody signálu $T/2$, pokud je časová konstanta RC podstatně kratší než $T/2$. Tento postup je však nevhodný pro silně rušené signály, kdy okamžité hodnoty zjišťované tímto obvodem mohou silně kolísat. Ze systémového hlediska je tento typ dolnofrekvenční filtrace *zvláštním případem koherentní demodula*ce, v níž referenčním signálem je úzký impuls ovládající spínač s opakovacím kmitočtem rovným dvojnásobku f_s a fází rovnou $\pi/2$.

2.1.16. *Vektor (fázor) voltmetr*

Součástí postdetekčních obvodů je často zapojení pro určení modulu fázoru měřeného signálu. Jde v podstatě o známý *vektorvoltmetr* pracující se dvěma KD, jejichž referenční signály jsou navzájem v kvadratuře. Obvodem (Obrázek 2.16) nebo výpočtem se z výstupů A, B obou KD určí modul U_s, případně fáze (jako $\mathrm{arctg}\ B/A$). Takto lze vyhodnotit amplitudu (modul) signálu i při neznámé nebo proměnné vzájemné fázi signálu a reference.



Obrázek 2.16 Určení složek fázoru dvěma KD (princip vektorvoltmetru)

2.2. Koherentní demodulátor na principu koincidence polarity

2.2.1. *Princip činnosti [5]*

Aditivní směs signálu a šumu $V(t)$ se vzorkuje krátkými impulsy s opakovacím kmitočtem (koherentní, referenční průběh). Každá shoda polarity referenčního průběhu a vzorků $v(t_i)$ se registruje v čítači. Je-li na vstupu pouze šumové napětí, pak shoda jeho polarity s polaritou vzorkovacích impulsů nastává při dostatečně dlouhé době pozorování přibližně pro polovinu vzorků. Je-li přítomen signál, bude zaregistrováno více nebo méně shod v závislosti na amplitudě signálu a fázi mezi signálem a vzorkovacími impulsy. Z hlediska realizace je výhodnější místo vzorkování průběhu $v(t)$ na nízké úrovni provést jeho omezení v komparátoru a vzorkovat jeho výstupní napětí. Pak lze pro zjišťování koincidence polarity použít jednoduchého logického součinového hradla. Na jednom vstupu hradla (Obrázek 2.17) působí výstup komparátoru (tj. omezený průběh $v(t)$) l, na druhém vzorkovací impulsy. Impulsy na výstupu hradla, vzniklé při koincidenci, se čítají po dostatečně dlouhou dobu pozorování (akumulace).

Odvoďme nyní závislost počtu shod polarity na amplitudě harmonického signálu

$$u_s(t) = U\sqrt{2}\sin(\omega t + \varphi)$$

který je doprovázen úzkopásmovým šumem $n(t)$ s normálním rozdělením hustoty pravděpodobností a disperzi σ^2_N . Jak známo lze úzkopásmový šum $n(t)$ přibližně považovat za fázor s náhodně proměnnou amplitudou a fází. Po rozkladu do složek platí

$$n(t) = n_s \sin \omega t + n_c \cos \omega t$$

Amplitudy n_s a n_c mají normální hustotu rozdělení, nulové střední hodnoty, jsou na sobě nezávislé a pro jejich sdruženou hustotu platí

$$r(n_s, n_c) = \frac{1}{2\pi\sigma_N^2} e^{-\frac{n_s^2 + n_c^2}{2\sigma_N^2}} \tag{2.18}$$

Směs signálu a šumu tj. součet

$$V(t) = s(t) + n(t)$$

může být upraven na tvar

$$V(t) = U_s sin\omega t + U_c cos\omega t$$

kde

$$U_s = n_s + U\sqrt{2}sin\varphi \text{ a } U_c = n_c + U\sqrt{2}.cos\varphi \tag{2.19}$$

Zápis $V(t)$ ve tvaru dvou složek napovídá, že zjišťováním koincidence polarity se dvěma sledy vzorkovacích impulsů vzájemně posunutých o *1/4* periody signálu můžeme obdržet informaci o U_s a U_c a následně určit amplitudu a fázi signálu. Pravděpodobnost koincidence polarity složky sinové s příslušným sledem vzorkovacích impulsů je dána pravděpodobností, že omezený průběh $U_s \sin \omega t$ je kladný v okamžiku vzorkování, tj. $U_s \geq O$ (obecněji $U_s \geq U_p$, kde U_p je srovnávací napětí komparátoru).
Je-li n_o celkový počet vzorků za dobu pozorování, pak poměrná část počtu soufázových vzorkovacích impulsů, při nichž zjistíme shodu polarity je dána vztahem

$$P(U_c \geq U_p) = \int_{-\infty}^{\infty} \int_{U_p}^{\infty} r(U_s, U_c) dU_c = \frac{n(0^0)}{n_0} \tag{2.20}$$

Hustotu pravděpodobnosti $r(U_S, U_c)$ získáme snadno z $r(n_s, n_s)$ po dosazení za n_s a n_s ze vztahů (C). Obdobně postupujeme při určení poměru $n(90^o)/n_o$. Výsledkem jsou vztahy

$$\frac{n(0^0)}{n_0} = \frac{1}{2}\left[1 - erf\left(\frac{U_p}{\sigma_N\sqrt{2}} - \frac{U\cos\varphi}{\sigma_N}\right)\right]$$
$$\frac{n(90^0)}{n_0} = \frac{1}{2}\left[1 - erf\left(\frac{U_p^{,}}{\sigma_N\sqrt{2}} - \frac{U\sin\varphi}{\sigma_N}\right)\right] \qquad (2.21)$$

Zde je použito chybové funkce $erf(x)$ a zjišťování shod polarity o čtvrt periody posunutých vzorkovacích impulsů $n($90°$/n($0°$)$ v jiném kanálu s odlišnou hodnotou U_p'.
Pro malé hodnoty argumentů $x \ll 1$ lze chybovou funkcí upravit dle vztahu

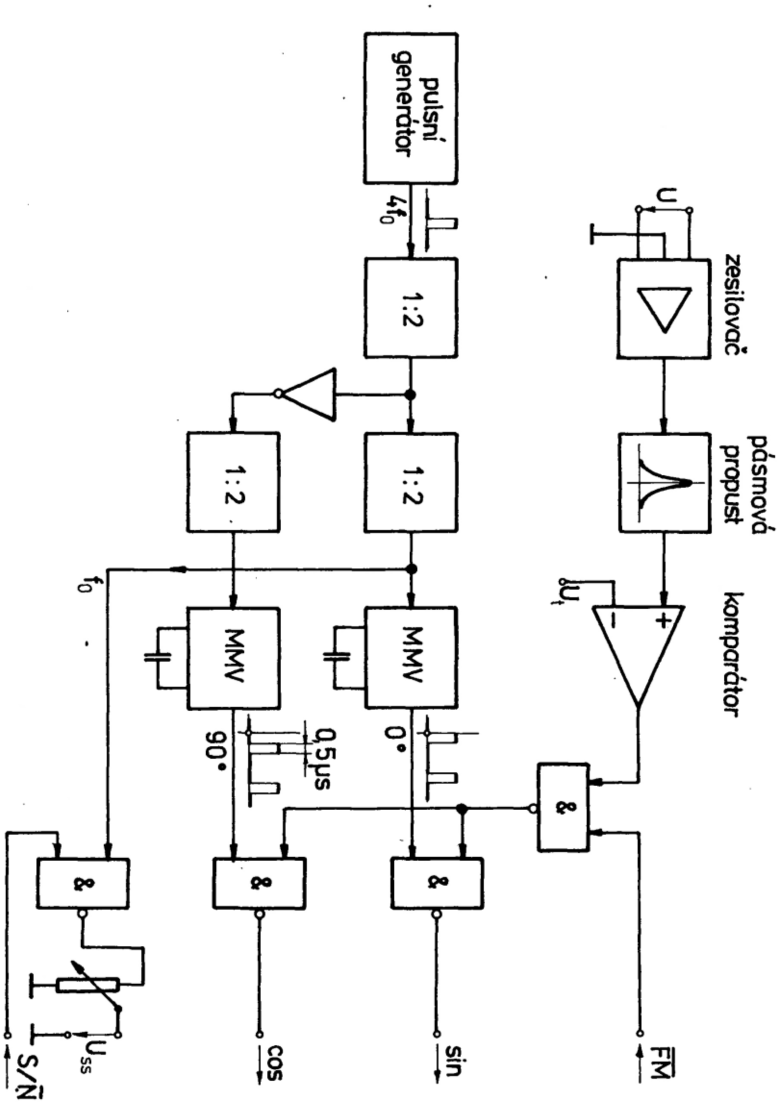
$$erf(x) \approx \frac{2}{\sqrt{\pi}}x \qquad (2.22)$$

Pokud jsou *amplitudy signálu podstatně menší než rozptyl* σ_N^2 , lze vztahy upravit na tvar

$$\frac{n(0^o)}{n_0} \approx \frac{1}{2} - \frac{1}{\sqrt{\pi}\sigma_N}\left(\frac{U_p}{\sqrt{2}} - U\cos\varphi\right)$$
$$\frac{n(90^o)}{n_0} \approx \frac{1}{2} - \frac{1}{\sqrt{\pi}\sigma_N}\left(\frac{U_p^{`}}{\sqrt{2}} - U\sin\varphi\right) \qquad (2.23)$$
$$\frac{U}{\sigma} = \frac{\sqrt{\pi}}{n_o}\sqrt{\left(n(0^o) - n_N(0^o)\right)^2 + \left(n(90^o) - n_N(90^o)\right)^2}$$

Lineární charakter těchto vztahů usnadňuje měření, které probíhá takto:

1. Při odpojeném signálu (U=0) se určí poměry n_N(90°)/n_0 a n_N(90°)/n_0, odpovídající působení samotného šumu.
2. Změří se hodnoty n_N (90°) a n_N (0°) při působení signálu a šumu.
3. Určí se rozdíly n_N (0°)/ n_0 - n_N (90°)/ n_0. Z nich je možné určit složky fázoru
4. Vypočteme amplitudu signálu dle vztahu
$$\frac{U}{\sigma} = \frac{\sqrt{\pi}}{n_o}\sqrt{\left(n(0^o) - n_N(0^o)\right)^2 + \left(n(90^o) - n_N(90^o)\right)^2} \qquad (2.24)$$
Určení amplitudy U je nejčastějším úkolem. Jak patrno z předchozího vztahu poměr U/σ nezávisí na fázi. Tím se vyloučí chyby vlivem kolísání fáze, které se mohou vyskytovat zejména při dlouhotrvajících měřeních. Požaduje se pouze, aby se fáze neměnila během jednoho odměru.



Obrázek 2.17 Příklad skupinového schématu realizace KD na principu koincidence polarity

KD na principu koincidence polarity je vhodný pro případy, kdy extrémně nízká úroveň signálu vyžaduje velké zesílení zesilovače. Pak při vysoké úrovni šumu dochází k omezení (nelinearitám) na výstupu zesilovače signálu. Na rozdíl od obvyklé situace zkreslení výstupního signálu zesilovače není na závadu a koincidenčním koherentním demodulátorem je možné určit amplitudu signálu.
Princip koincidenčního demodulátoru je do jisté míry příbuzný metodám určování spektra z průchodu signálu nulou [7].

2.3. Číslicové koherentní demodulátory.

Vztah pro odhad amplitudy signálu je přímo realizován číslicovým procesorem, zpravidla výpočtem v reálném čase. Na první pohled jsou nároky na převod signálu $v(t)$ do číslicového tvaru nadměrné, zejména pro případ obnovy signálu s úrovní hluboko pod rušením. Tak např. při úloze vyhodnotit amplitudu signálu s přesností odpovídající 7 bitům, "utopeného" o 90 dB pod rušením, se vyžaduje A/Č převodník s rozlišením odpovídajícím 23 bitům. Realizace takového převodníku, zejména pro vyšší kmitočty, by byla velmi náročná. Proto je tento postup užíván pouze pro pomalu proměnné signály nebo v případech, kdy nehrozí přetížení signálního kanálu nadměrným rušivým napětím. Naštěstí lze však obejít nutnost použít převodníků A/Č s vysokým rozlišením využitím principu zvyšování rozlišovací schopnosti A/Č převodu přídavným rozmítacím signálem. Jako přídavný signál náhodného charakteru slouží zesílený vstupní šum procesu. Princip této metody je popsán v následující kapitole.

2.3.1. *Zvýšení rozlišovací schopnosti přídavným signálem*

Ideální A/Č převodník s nekonečnou rozlišovací schopností by měl lineární a nikoliv stupňovitou převodní charakteristiku. Konečná rozlišovací schopnost A/Č převodníku je v podstatě důsledkem jeho nelineární převodní charakteristiky.

Linearizaci převodní charakteristiky pro účely zvýšení rozlišovací schopnosti A/Č převodu lze realizovat na základě *teorému ekvivalentní nelinearity* (TEN).

Teorém je založen na určení podmíněné střední hodnoty směsi signálu s a šumu n, tj. $x = s + n$, na výstupu členu s nelineární převodní charakteristikou, popsanou funkcí $y(x)$

$$E\{y(x) \mid s\} \equiv \int_{-\infty}^{\infty} y(s+n) r[(s+n) \mid s] \,\mathrm{d}n \tag{2.25}$$

kde $r[(s+n)|s]$ je podmíněná hustota pravděpodobnosti součtu $(s + n)$ při dané hodnotě s.

Pro další úpravu použijeme Bayesův vztah

$$r[(s+n) \mid s] = \frac{p[(s+n), s]}{q(s)} \tag{2.26}$$

Pro nezávislé procesy platí

$$p[(s+n), s] = p_1(s+n).q(s) \tag{2.27}$$

a po dosazení

$$r[(s+n) \mid s] = p_1(s+n) \tag{2.28}$$

Hustotu pravděpodobnosti součtu nezávislých veličin $R = P + Q$ lze vypočítat jako konvoluci jejich hustot pravděpodobnosti $p_p(p)$ a $p_q(q)$ podle vztahu

$$p_r(r) = \int p_p(u) p_q(r-u) du \tag{2.29}$$

Ve sledovaném případě se v *konvolučním* vztahu uplatní hustota pravděpodobnosti veličiny s, vyjádřená delta funkcí $\delta(\xi - s)$ (jde o stálou veličinu), a hustota pravděpodobnosti veličiny n, tj. $m(n)$.

Pravou stranu rovnice (2.28) lze tedy vyjádřit konvolučním vztahem

$$p_1(s+n) \; = \; \delta(\xi - s) * m(s+n) \; = \int_{-\infty}^{\infty} \delta(\xi - s) m(s+n-\xi) \mathrm{d}\xi \; = m(n) \tag{2.30}$$

Po dosazení do (2.25) dostaneme

$$E\{y(x) \mid s\} = \int_{-\infty}^{\infty} y(s+n).m(n) \, \mathrm{d}n = R_{ym}(s) = R_{ym}(x)\big|_{x=s} \tag{2.31}$$

kde $R_{ym}(s)$ je vzájemná korelace funkcí $y(x)$ a $m(n)$.

Zde bylo využito podobnosti definičních vztahů korelace a konvoluce

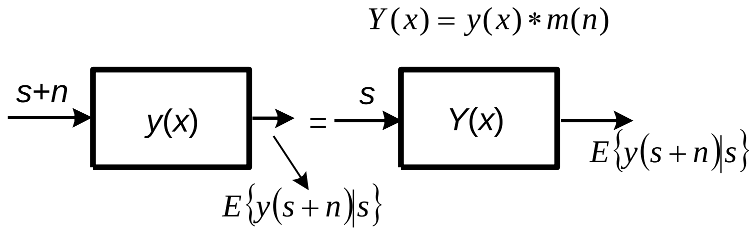
$$E\{y(x) \mid s\} = y(x) * m(-x)\big|_{x=s} = Y(x)\big|_{x=s} \tag{2.32}$$

Z hlediska podmíněné střední hodnoty signálu s se nelineární převodní charakteristika transformuje na ekvivalentní $Y(x)$, danou konvolucí původní převodní charakteristiky $y(x)$ a hustoty pravděpodobnosti přídavného signálu $m(x)$.

2.3.2. *Teorém ekvivalentní nelinearity (TEN)*

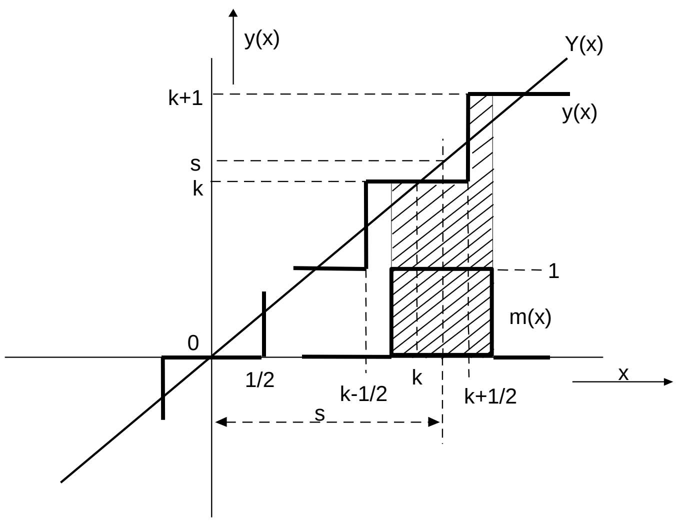
Pro určení podmíněné střední hodnoty součtu náhodných proměnných na výstupu členu s nelineární převodní charakteristikou lze původní převodní charakteristiku $y(x)$ nahradit ekvivalentní charakteristikou $Y(x)$, určenou konvolucí původní charakteristiky $y(x)$ s hustotou pravděpodobnosti přídavného signálu $m(n)$. Vstupním signálem obvodu s ekvivalentní převodní charakteristikou $Y(x)$ je proměnná s (Obrázek 2.18).

Na základě teorému ekvivalentní nelinearity [9] je principiálně možné modifikovat libovolnou převodní charakteristiku přídavným signálem s vhodně zvoleným průběhem hustoty pravděpodobnosti. Při praktických aplikacích je důležité si uvědomit, že náhrada ekvivalentní nelinearitou platí pouze pro *střední hodnotu* signálu. Musí být tedy zajištěno potlačení rušivých účinků přídavného signálu (šumu) na střední hodnotu signálu - nejjednodušeji filtrací, nebo volbou spektra přídavného signálu mimo oblast spektra (obecně časově proměnné) střední hodnoty signálu.



Obrázek 2.18 K teorému ekvivalentní nelinearity

Přídavným signálem lze modifikovat schodovitou převodní charakteristiku A/Č převodníku tak, aby ekvivalentní charakteristika měla lineární průběh, odpovídající ideálně nekonečné rozlišovací schopnosti převodníku. Takovýto postup odpovídá zvýšení rozlišovací schopnosti a označuje se jako *dithering*.
Číslicový výpočet střední hodnoty tj. operace $E\{y(s+n)|s\}$ po detekci-je podstatně dokonalejší než u analogové verze, jelikož se využívá koherentního průměrování. Vzorky výstupního signálu číslicového KD jsou odebírány se stejnou frekvencí jako u převodníku A/Č, takže při vzorkovací frekvenci např. 100 MHz je za 1 s vzroste rozlišovací schopnost 10^8 krát a rozptyl průměru z 10^8 hodnot klesne asi 10^4 krát.



Obrázek 2.19 Princip "*ditheringu*": původní stupňovitá převodní charakteristika *y*(*x*) je upravena přídavným signálem s rovnoměrně rozloženou hustotou pravděpodobnosti *m*(*x*) na ekvivalentní lineární průběh *Y*(*x*)

Linearizace stupňovité převodní charakteristiky vyžaduje, aby přídavný signál $m(n)$ měl v intervalu $\langle -q/2, q/2 \rangle$ rovnoměrnou hustotu rozdělení pravděpodobnosti. Na Obrázku 2.19 je postupem *grafické konvoluce* ukázáno, že pro vstupní signál o hodnotě s střední hodnota plochy součinu překrývajících se částí $m(x)$ a $y(x)$ (tj. konvoluce) je rovna s, jak je ukázáno na následujícím příkladu výpočtu:

$$Y(x)\big|_{x=s} = \left(\frac{3}{4} q.k + \frac{1}{4} q(k+1) \right)\Big|_{q=1} = k + \frac{1}{4} = s \tag{2.33}$$

V příkladu je z důvodů jednoduchosti uvažován jednotkový kvantizační krok ($q = 1$). To dokazuje, že ekvivalentní převodní charakteristika je lineární.

Přídavným signálem o amplitudách větších než q lze také potlačit případné nelinearity, zasahující větší úseky převodní charakteristiky (např. různé typy integrálních nelinearit).

Zvýšení rozlišovací schopnosti přídavným signálem vede na myšlenku používat levné převodníky A/Č s malým počtem bitů (v krajním případě jedním - komparátor) a rozlišovací schopnost zvýšit přídavným signálem. Tato principiálně správná možnost však vyžaduje přídavný signál s velkou amplitudou pokrývající plný rozsah převodníku a tím podstatně rostou nároky na obvody potlačující účinky přídavného signálu na střední hodnotu užitečného signálu.

Konvoluční postup výpočtu $Y(x)$ lze na základě analogie vztahu mezi konvolucí a spektrem při Fourierově transformaci nahradit násobením Fourierových transformací převodní charakteristiky A/Č převodníku a hustoty pravděpodobnosti přídavného signálu. Fourierova transformace hustoty pravděpodobnosti se označuje jako *charakteristická funkce*.

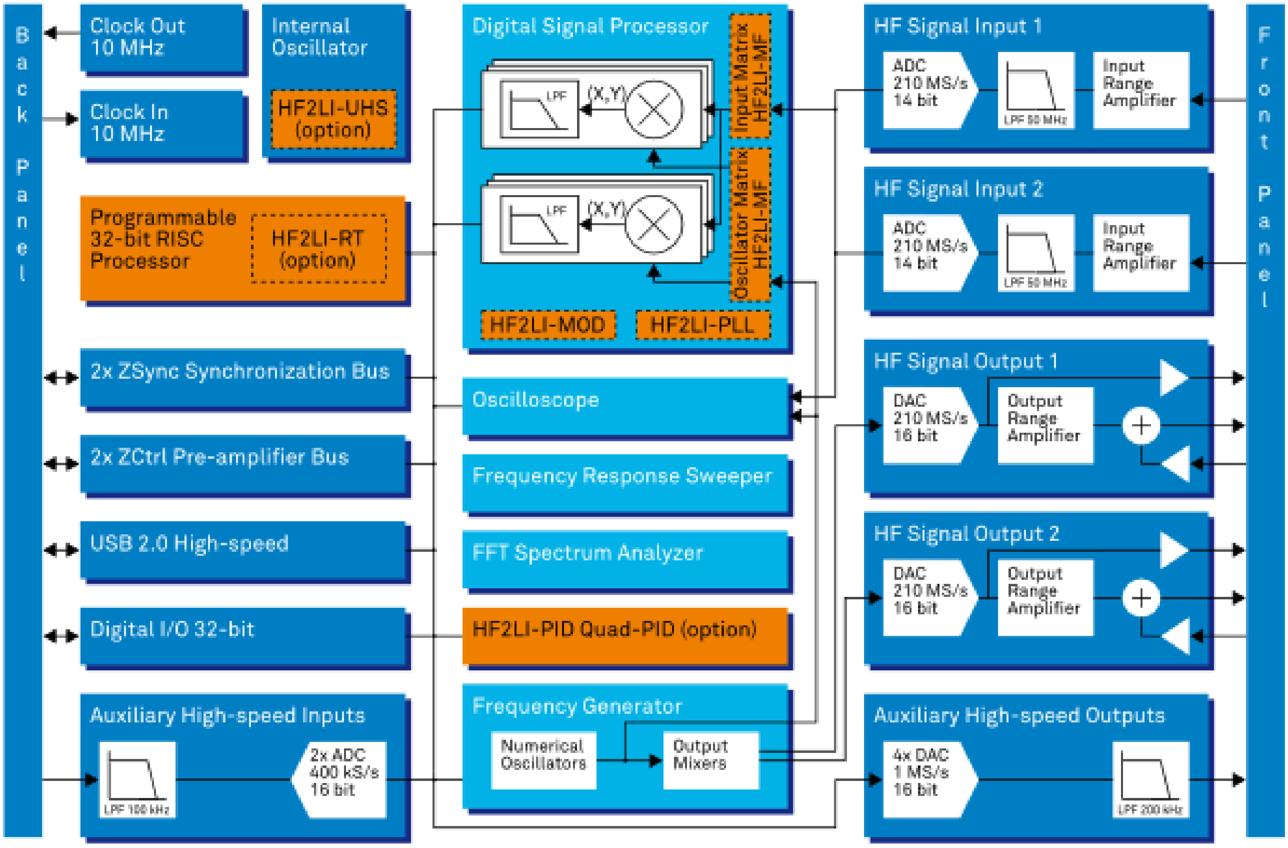
2.3.3. *Výhody číslicové koherentní demodulace*

1. Číslicové násobení (např. 128 bitů) vstupního signálu referenčním signálem přísně harmonického tvaru, získaným přímou číslicovou syntézou (DDS). Referenční signál je počítán s přesností 20bitů a vyšší harmonické jsou potlačeny o 120 dB. Neuplatní se okna přenosu na harmonických referenčního signálu.

2. Amplituda, frekvence a fáze referenčního signálu a tím i průběh AFCH KD je známý s vysokou přesností, tj. zisk KD je stálý a známý

3. Číslicové zpracování signálu po koherentní demodulaci. Číslicově realizovaná dolní propust (integrace-určení střední hodnoty) je téměř ideální. Takto je potlačena nestálost nuly známá u analogových přístrojů.

4. Zatímco obvyklá dynamická rezerva analogových KD je kolem 60 dB, převodník A/Č užívaný v číslicových KD je extrémně lineární a využívá ditheringu, takže dynamická rezerva snadno dosahuje až 130 dB



Obrázek 2.20 Moderní koncepce číslicového koherentního demodulátoru (Zurich Instruments)

Zvláštností jsou velmi rychlé A/D převodníky (210 MS/s, 14 bitů), takže není nutné používat směšování jako u heterodynních demodulátorů analogových.
5. Méně kvalitní externí referenční signál se „*obnovuje*“ v číslicovém oscilátoru.

2.4. Koherentní průměrování signálů

Pod pojmem "průměrování" (averaging) ve smyslu zpracování signálu se rozumí určení aritmetického průměru - střední hodnoty - z jistého úseku signálu. Realizace této operace analogovými metodami se v zásadě děje integračními obvody, přesněji filtry typu dolnofrekvenční propust, jak je známe např. z převodníků parametrů střídavých signálů.

2.4.1. *Teorém fúze informací*

Číslicové postupy zpracování signálu ve formě číslicových filtrů typu FIR jsou téměř ideální pro realizaci operací typu průměrování. Typickým příkladem je transverzální filtr FIR s přenosovou charakteristikou odpovídající dolnofrekvenční propusti.

S problematikou koherentního průměrování souvisí *teorém fúze informací*, určující odhad neznámé veličiny $\mathbf{X}$ (rozměr $p \times 1$) z měřených dat $\mathbf{Y}_i$ (rozměr $p \times 1$) zatížených aditivními a vzájemně nezávislými chybami $\mathbf{N}_i$ Každá chyba je charakterizována kovarianční maticí $\boldsymbol{S}_i$ o rozměru $p \times p$. Optimální odhad $\widehat{\boldsymbol{X}}$ a jeho kovarianční matice $\widehat{\boldsymbol{S}}$ jsou dány následujícími vztahy

$$\widehat{\mathbf{X}} = [S_1^{-1} + S_2^{-1} + \cdots + S_n^{-1}]^{-1}[S_1^{-1}Y_1 + S_2^{-1}Y_2 + \cdots + S_n^{-1}Y_n] \quad (2.34)$$

$$\widehat{\boldsymbol{S}} = [S_1^{-1} + S_2^{-1} + \cdots + S_n^{-1}]^{-1}$$

V nejjednodušším případě jsou Y_i jednotlivé hodnoty neznámé veličiny při opakovaném měření zatížené chybou s variancí $S_i = \sigma_i^2$. Předpokládejme, že variance jednotlivých měření jsou stejné, pak je variance odhadu po n měřeních rovna

$$\hat{S} = [n\sigma^{-2}]^{-1} = \frac{\sigma^2}{n} \quad (2.35)$$

Docházíme k opět k známému poznatku, že *n*-krát opakovaným měřením stejné veličiny, nebo použitím n senzorů (zdrojů informace) pro *jedinou* veličinu klesá variance odhadu úměrně počtu měření, respektive počtu senzorů. Pro potlačení periodických rušivých složek superponovaných na pomalu proměnném signálu jsou zvláště vhodné filtry na principu *klouzavého průměru* s celočíselným počtem vzorků na periodu rušivého signálu.

Zvláštním případem obecné procedury určování průměru signálu je *koherentní průměrování*, které je základem většiny procesů pro obnovu signálu ze šumu. Zde se využívá podmínky koherence, jež je vlastně důsledkem apriorních znalostí o sledovaném jevu.

Na rozdíl od jednoduchého úkolu určení amplitudy signálu postupem koherentní demodulace se často vyžaduje *regenerace tvaru signálu*. Podstatou všech postupů průměrování s cílem obnovy tvaru signálu je sumace vzorků signálu odebíraných při opakovaných měřeních stále ze stejných (nebo předem známých) míst signálu. Pokud je průběh signálu během sčítání stálý, vzroste obsah jednotlivých paměťových míst (jejich počet odpovídá počtu vzorků na jeden průběh signálu) úměrně počtu opakování měření N. V souladu s důsledky teorému o fusi informací se však příspěvky náhodného rušení (variance) uplatní úměrně odmocnině N. Předpokladem splnění podmínky koherence je požadavek opakování zpracovaného signálu při každém stimulu, tj. předpokládá se, že signál je přísně determinovaný. Dále se vychází z podmínky aditivního působení šumové složky $n(t)$ a konečně za předpokladu, že šum při jednotlivých pokusech není navzájem korelován nebo vyvoláván stimulem. Zpravidla je nejhůře splnitelná podmínka deterministického průběhu signálu. Proto při komplikovanějších aplikacích se používá modifikovaného postupu průměrování, založeného na zjišťování vzájemné korelace mezi jednotlivými výsledky průměrování.

Koherentní průměrování lze realizovat:

- sériově
- paralelně

Sériová (jednokanálová) realizace průměrování spočívá v odběru a sumaci vzorků z místa stejně vzdáleného od počátku pozorování při každé realizaci signálu. V další sérii realizací se sčítají vzorky signálu z místa posunutého o dostatečně jemný časový inkrement Δt. Postup se opakuje pro celý průběh signálu.

Na tomto principu jsou založeny tzv. *spínané koherentní integrátory* (stroboskopické nebo synchronní integrátory, *box-car integrátory*), užívané zejména pro obnovu signálu odezvy fyzikální soustavy na opakované stimuly.

Paralelní (vícekanálová) realizace průměrování spočívá v přiřazení paměťových míst k jednotlivým bodům průběhu signálu a opakovaném přičítání vzorků signálu do odpovídajících míst v paměti. Tento postup vede k cíli rychleji za cenu náročnější realizace.

2.6 Přístroje pro koherentní průměrování

2.4.2. *Koherentní integrátory*

Základní princip činnosti tohoto přístroje (Obrázek 2.21), známého také jako synchronní integrátor nebo *box-car integrátor* (BCI), je v podstatě vzorkovací obvod, avšak s parametry, zásadně se lišícími od hodnot vyžadovaných pro ideální realizaci vzorkování. Jde především o splnění podmínky koherence, která zde prakticky znamená, že musí být k dispozici informace o přesném počátku sledovaného periodického signálu [6]. Pak je poměrně snadné odvodit okamžiky vzorkování, tj. řídicí impulsy pro spínač vzorkovacího obvodu.

Další odlišností je hodnota časové konstanty $\tau = RC$, která musí být podstatně větší než doba sepnutí spínače τ_G. Pak je činnost obvodu následující:

Po každém spouštěcím impulsu, odvozeném z generování stimulu soustavy a opakujícím se každých T_p sekund, se po dobu τ_G nabíjí paměťová kapacita C. Výstupní napětí roste exponenciálně, obdobně jako při nabíjení kondenzátoru ze zdroje napětí při sepnutém spínači. Ustálený stav je dosažen - s odchylkou 0,7 % - za dobu t_u = 5 RC. Počet sepnutí spínače N, nutný k dosažení přibližné ustálené hodnoty, je roven

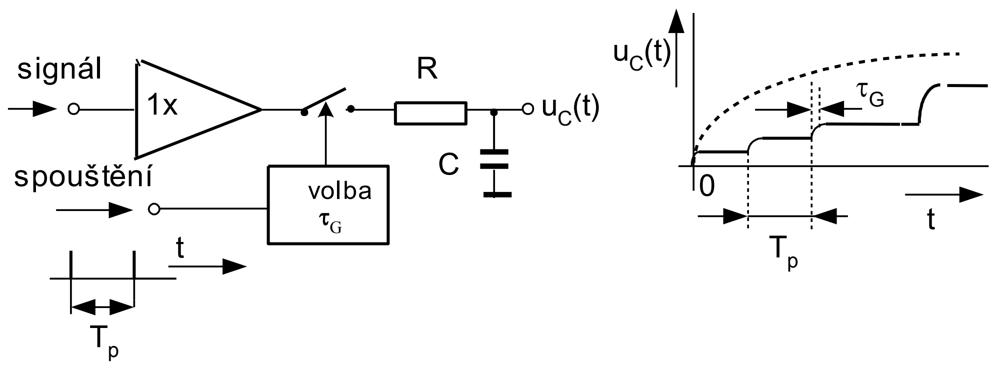
$$N = \frac{t_u}{\tau_G} = 5\frac{RC}{\tau_G} \tag{2.36}$$

kde N je současně počet vzorků signálu přivedených na vstup obvodu.

Jelikož spínání se opakuje s periodou T_p, lze správnou hodnotu vzorku signálu získat na výstupu až za dobu

$$t_u = NT_p = 5RC\frac{T_p}{\tau_G} \tag{2.37}$$

od prvního sepnutí spínače.



Obrázek 2.21 Princip koherentní integrace (box-car integrator)

2.4.3. *Zlepšení poměru signál/šum koherentní integrací*

Zlepšení poměru signálu k šumu dosažené koherentním průměrováním pro rušivý signál se statistickými parametry blízkými bílému šumu je, jak známo, opět úměrné $\sqrt{N}$. Exaktní analýza respektující reálné vlastnosti použitého obvodu dává menší hodnotu zlepšení dle vztahu

$$H = \sqrt{2\frac{RC}{\tau_G}} \tag{2.38}$$

Je-li doba trvání průběhu signálu T_m, pak celkový počet zpracovaných vzorků signálu je dán vztahem

$$N_C = \frac{T_m}{\tau_G} N = 5 \frac{RC}{\tau_G^2} T_m \qquad (2.39)$$

Všechny vzorky signálu jsou odebrány za časový interval

$$T_{sc} = N_C T_p = 5 \frac{RCT_m}{\tau_G^2 f}, \quad f = \frac{1}{T_p} \qquad (2.40)$$

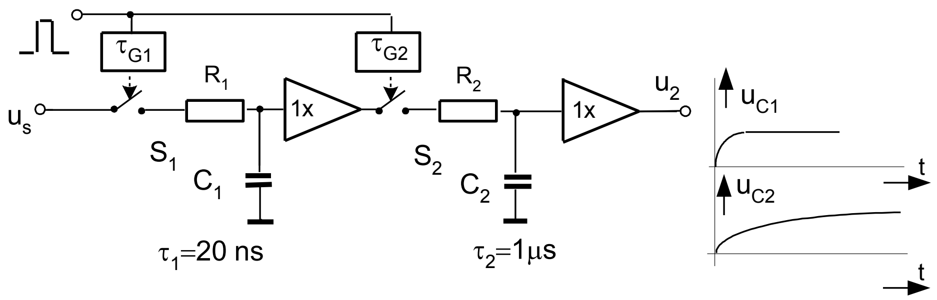
Podmínkou správné rekonstrukce signálu je stálost jeho tvaru nejméně po dobu T_{sc}.

Doba sepnutí spínače τ_G odpovídá svým významem době apertury vzorkovacího obvodu a musí splňovat kritéria obvyklá pro vzorkovací obvody, tj. musí být podstatně kratší než doba periody nejvyšší harmonické složky signálu. Typická hodnota τ_G u přístrojů pro průměrování signálů, vyskytujících se v experimentální fyzice, je řádově několik desítek ps.

2.4.4. *Obvodová koncepce přístrojů pro koherentní průměrování*

Obvody pro vzorkování

Rychle proměnné časové průběhy vyžadují malé hodnoty paměťové kapacity C, aby přírůstky napětí $U.\ \tau_\Gamma/RC$ převýšily šum spínače a zesilovače. Avšak malé hodnoty časové konstanty znamenají malé zlepšení poměru s/š.



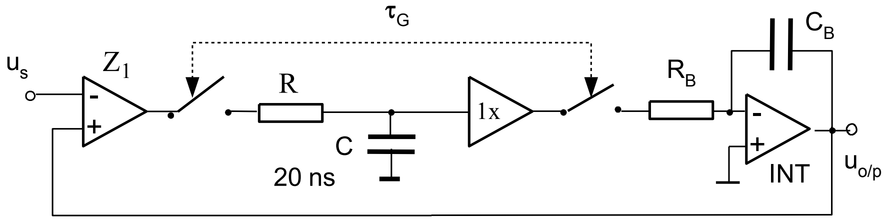
Obrázek 2.22 Dvoustupňové spínání - koherentní integrátor s předřazeným vzorkováním

Řešením je *dvojstupňové spínání* (Obrázek 2.22). První spínač má velmi krátkou spínací dobu, např. 20 ns, další spínač je sepnut po dobu podstatně delší a neměnnou, např. 1 μs. Na malé kapacitě prvního stupně tak vznikne i při velmi krátké době sepnutí značné napětí. Časová konstanta druhého stupně je proměnná a respektuje opakovací periodu signálu T_p.

Zpětnovazební obvody pro vzorkování

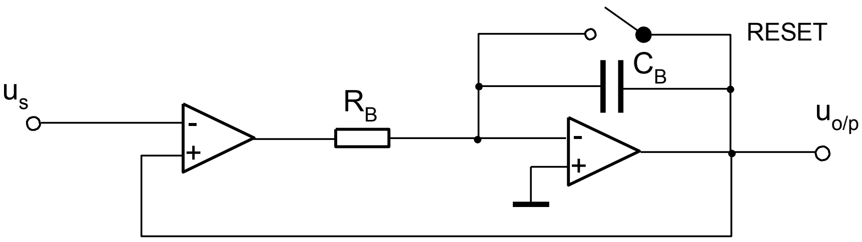
Využívají obdobně jako běžné vzorkovací obvody příznivé účinky záporné zpětné vazby pro potlačení vlivu rušivých parametrů spínače a závislosti na vnitřním odporu zdroje signálu.

Dvoustupňový zpětnovazební vzorkovací obvod na Obrázku 2.22 lze za předpokladu, že platí $RC << R_B\, C_B$, nahradit obvodem dle Obrázku 2.23.



Obrázek 2.23 Zpětnovazební koherentní integrátor

Obrázek 2.24 Náhradní obvod k zapojení na Obrázku 2.24.



Výstupní napětí náhradního obvodu $U_{o/p}$ lze určit z obvodových rovnic (předpokládá se, že oba spínače jsou sepnuty):

$$u_{o/p} = \left(1 + \frac{1}{p\tau}\right) = u_s \frac{1}{p\tau} \Rightarrow \frac{u_{o/p}}{u_s} = \frac{1}{1 + p\tau} \tag{2.41}$$

odpovídá přenosu integračního článku s časovou konstantou $R_B C_B$. Základní princip koherentního průměrování se tedy nemění, jde opět o tzv. *exponenciální* průměrování (tj. průběh výstupního napětí se řídí známou exponenciální závislostí nabíjení kondenzátoru).

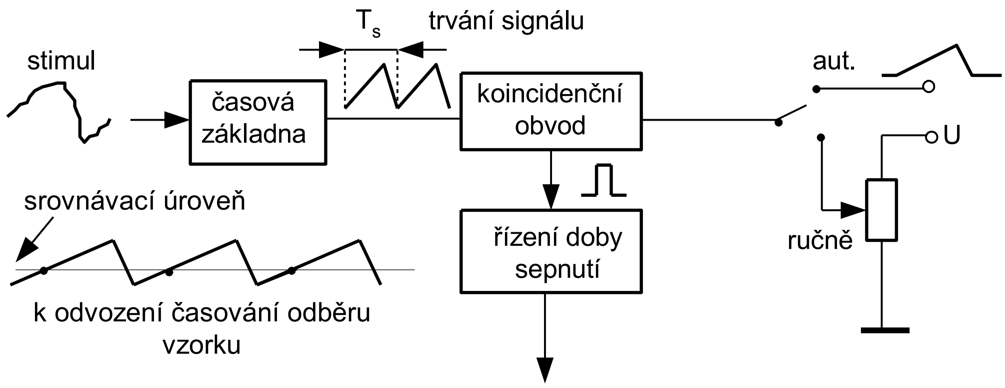
Další alternativou je *lineární* průměrování, kdy je hlavní smyčka zpětné vazby rozpojena a vstupní signál se během doby sepnutí spínače integruje zpětnovazebním integračním obvodem.

Lineární průměrování dává větší hodnotu signálu při stejných hodnotách τ_G a RC než exponenciální. Je výhodné zejména tehdy, je-li počet vzorků signálu, které jsou k dispozici, omezen (např. rychle se měnící okolní podmínky pokusu), nebo když se projevují silné rušivé složky. Nevýhodou je možnost saturace integrátoru po velkém počtu vzorků. Proto je nutné kapacitor integrátoru vybít dříve, než dojde k nasycení. Zachování informace před vybitím se děje po převodu na číslicovou veličinu uložením do paměti.

Exponenciální průměrování probíhá v základní konfiguraci obvodů pro koherentní průměrování a z principu vylučuje jev saturace i při velkém počtu zpracovávaných vzorků.

Časová základna

Postupné odebírání vzorků z průběhu sledovaného signálu se realizuje řízeným zpožděním spínáním vzhledem k počátku sledovaného jevu. Jak je ukázáno na Obrázku 2.25 děje se tak komparací pilovitého průběhu časové základny, spouštěné impulsy odvozenými z počátku experimentu, s podstatně pomalejším pilovitým průběhem rozmítacího napětí. Rychlost růstu rozmítacího napětí je volena tak, aby i při požadavku mnohonásobně opakovaného odběru vzorku signálu ze stejného místa nedošlo ke změnám okamžiku vzorkování větším než dovolená chyba.



Obrázek 2.25 Časová základna koherentního integrátoru

2.4.5. *Číslicové koherentní spínané integrátory*

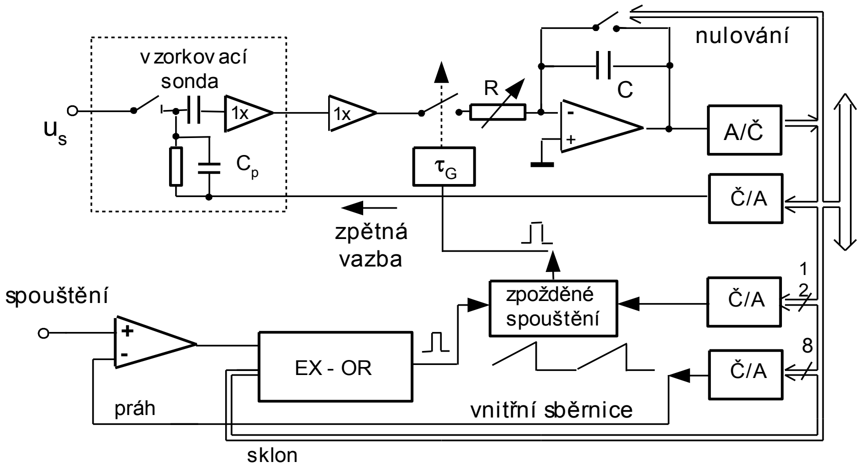
Nedokonalosti analogových koherentních spínaných integrátorů, tj. především mimořádné nároky na linearitu časové základny pro postupné spínaní vzorkujícího spínače při velkém požadovaném zlepšení poměru s/š a změna akumulovaných okamžitých hodnot vybíjením paměťových kapacit, lze podstatně omezit uplatněním číslicových přístupů.

Základem signálního kanálu je i v číslicové verzi spínaný integrátor, avšak s minimální časovou konstantou RC, takže tvoří vlastně vzorkovací obvod. Napětí na výstupu této "vzorkovací jednotky" je převedeno do číslicového tvaru a akumulováno postupně v paměťových místech, přiřazených jednotlivým bodům sledovaného průběhu. Chyby vybíjením paměťové kapacity v období mezi opakováním stimulů jsou tak zásadně potlačeny.

Číslicová koncepce časové základny rovněž přináší podstatné zlepšení. Jak plyne z podstaty automatického postupu odběru vzorků, jsou při déletrvajících signálech a větším počtu opakování stimulů nároky na linearitu časové základny mimořádně přísné. Jinak dochází k deformaci tvaru signálu vlivem nelinearity časové osy. Proto se generuje zpoždění sepnutí spínačů integrátoru číslicovými metodami na bázi čítačů. Pak je poměrně snadné řídit časovou základnu tak, aby sekvence odběru vzorků byla odlišná od běžně užívané "lineární" posloupnosti. Např. pro rychle proměnné signály je výhodné sledovat genezi průběhu, tj. zobrazovat vzorky nejprve z poloviny průběhu, pak z 1/4 a 3/4 celkové doby pozorování, atd.

Příklad přístrojového řešení koherentního integrátoru řízeného počítačem je na Obrázku 2.26. Za zmínku stojí použití převodníků A/Č a Č/A ve zpětné vazbě, odvození spouštění časové základny od náběžné nebo sestupné hrany spouštěcího impulsu (hradlo EX-OR je ve stavu H pouze pro lichý počet vstupů ve stavu H), nastavení prahu spouštění programem, generování zpoždění odběru vzorku komparací pomalého stupňovitého průběhu na výstupu převodníku Č/A s pilovitým průběhem vznikajícím v obvodu pro zpožděné spouštění.

Číslicové koherentní integrátory kladou velké nároky na rozlišovací schopnost převodníků A/Č, jelikož je nutné digitalizovat i signály ležící hluboko pod úrovní šumu.

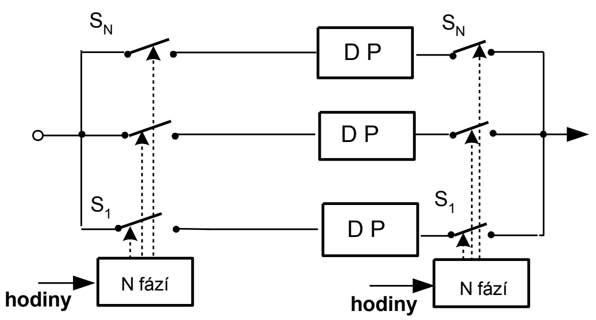


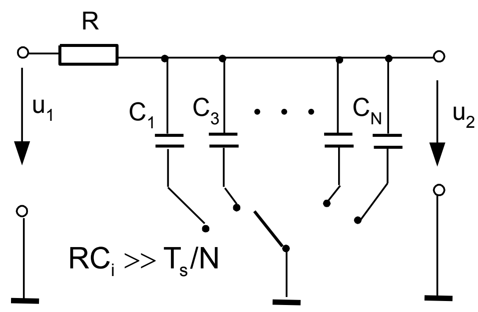
Obrázek 2.26 Koherentní integrátor řízený počítačem

2.5. Koherentní filtrace

Výstupem koherentního demodulátoru je stejnosměrné nebo pomalu proměnné napětí objevující se za dolnofrekvenční propustí. Jak bylo již ukázáno, zvláštností přenosové charakteristiky koherentního demodulátoru je skutečnost, že výstupní napětí je na jiném kmitočtu než vstupní signál, tj. dochází k transpozici spektra signálu.

Pokud je z nějakých důvodů vhodné, aby výstupní signál KD byl na stejném kmitočtu jako vstupní, je možné výstupní napětí KD amplitudově namodulovat na nosnou vlnu s harmonickým průběhem.





Obrázek 2.27 Princip koherentní filtrace Obrázek 2.28 Princip obvodového řešení KF

Jedna z možných realizací tohoto postupu je znázorněna na Obrázku 2.27. Jde v podstatě o skupinu N obvodů pro koherentní integraci (box-car integrátorů), odebírajících postupně cyklicky N vzorků vstupního signálu. Na každém z N kapacitorů (výstupů dolnofrekvenčních propustí) se po jistém počtu cyklů ustálí hodnoty napětí odpovídající úrovním vzorků vstupního periodického signálu v okamžicích $k.T_S/N$, kde T_S je perioda signálu. V jednoduchém případě harmonického signálu napětí na kapacitorech odpovídají aproximaci sinusovky N stupni.

Obdobně jako u koherentních integrátorů musí platit

$$RC_i = \tau_D >> \frac{T_S}{N} \tag{2.42}$$

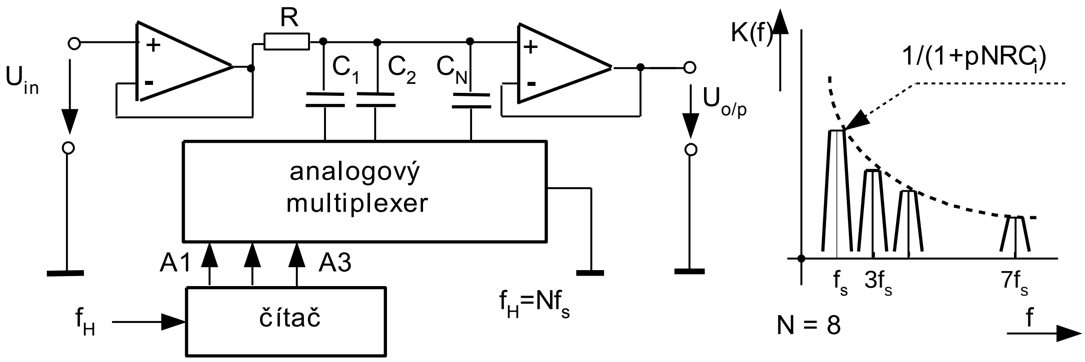
Synchronně pracujícími spínači multiplexeru (Obrázek 2.28) se napětí na kapacitách postupně připojují na výstupní svorku obvodu.

Pro praktickou realizaci stačí jediná sada spínačů pracujících navíc s uzemněným společným kontaktem. Spínače mohou být vytvořeny analogovým multiplexerem, jehož řídicí signály jsou odvozeny z výstupů čítače (Obrázek 2.29) s hodinovým kmitočtem $f_H = N f_S$.

Amplitudová frekvenční charakteristika má průběh hřebenového filtru s pásmy propustnosti, kopírující AFCH dolnofrekvenčních propustí, zrcadlově umístěné kolem násobků základního kmitočtu signálu f_S. Ekvivalentní časová konstanta propusti však odpovídá N-násobku RC .

Z integračního charakteru procedury akumulace vzorků na kapacitorech je patrné, že signály na vstupu o periodě, rovné násobku T_S/N, nepřispívají k výstupnímu signálu, takže na AFCH filtru schází propustná okénka na kmitočtech $N.f_S$ Výstupní střídavý signál lze dále pohodlně zesílit a případně následnou filtrací vyhladit schodovitý průběh.

Koherentní filtrací dochází k zlepšení poměru signál/šum úměrně odmocnině počtu cyklů rotujícího spínače obdobně jako u koherentních integrátorů.



a) b)

Obrázek 2.28 Koherentní filtr: a)obvodová realizace b) AFCH

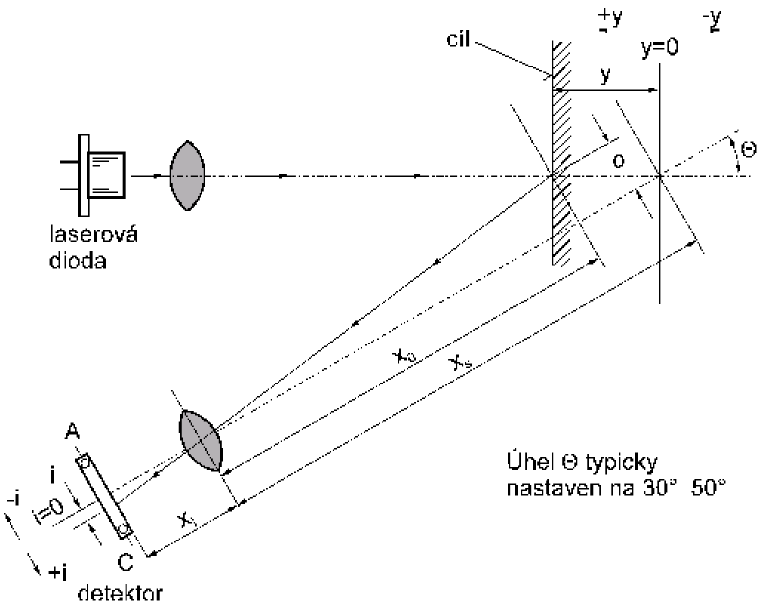
3. Vybrané aplikace koherentních demodulátorů

Kromě obvyklých použití KD pro amplitudovou demodulaci (měření signálu utopeného v šumu, můstková měření, určení složek fázoru) princip koherentní demodulace nachází uplatnění v jiných, méně rozšířených aplikacích. Vybrané typické příklady budou uvedeny v následujících odstavcích.

3.1.1. *Měření využívající senzorů záření*

Z fyzikální podstaty a konstrukce senzorů záření (tepelného, optického, ionizačního) plynou problémy způsobené nedostatečnou dlouhodobou stálostí jejich parametrů. Na vstupu soustavu měřící elektrické parametry senzoru se tyto rušivé efekty projeví jako ekvivalentní šum, v jehož spektru převažují složky na nízkých frekvencích (podobně jako u šumu typu $1/f$).

Obdobné je chování samovolného posuvu nuly-driftu stejnosměrných zesilovačů. Známým řešením tohoto problému je posuv spektra signálu do oblasti vyšších kmitočtů modulací realizovanou co nejblíže k začátku měřicího řetězce. Prakticky to znamená přerušovat tok záření mechanicky („optical chopper“) nebo elektricky (např. impulzním napájením LED příp. laserové diody).



Obrázek 3.1 Triangulační měření vzdálenosti s přerušovaným zářením

Příkladem může být triangulační měření vzdálenosti (Obrázek 3.1), kdy svazek infračerveného záření laserové diody dopadá na povrch objektu s nenulovým koeficientem odrazu. Při změně polohy objektu o vzdálenost x se posune stopa difúzně odraženého paprsku na detektoru PSD o úměrnou vzdálenost i. Střídavý signál z detektoru o modulačním kmitočtu (např. 16 kHz) s amplitudou úměrnou poloze je dále pohodlně zesilován střídavým zesilovačem a zpracován koherentním demodulátorem.

Při praktické realizaci mohou nastat problémy se zajištěním podmínky koherence, pokud zdroj záření je značně vzdálen od měřicího obvodu s detektorem a pro přívod referenčního signálu od zdroje nelze použít běžné propojení vodiči. Možné řešení je bezdrátový přenos vysílačem nebo případně použití signálů GSM.

3.1.2. *Měření šumu*

Amplitudová frekvenční charakteristika koherentního demodulátoru má tvar hřebenového filtru, obecně složeného z řady propustných „oken“, s přenosovou charakteristikou odpovídající přenosu dolnofrekvenční propusti za součinovým členem (Obrázek 2.11). Širokopásmový šum působící na vstupu LIA vyvolá na výstupech demodulátorů napětí úměrné součtu soufázových (I) a kvadraturních(Q) složek signálů z jednotlivých propustí. Předpokládejme, že příspěvky z pásem odpovídajícím propustím na vyšších harmonických, jsou zanedbatelné. Pak je měřená výkonová hustota šumu dána vztahem

$$e_n = \frac{1}{G}\sqrt{\frac{S_I^2 + S_Q^2}{2\Delta f}} \tag{3.1}$$

kde S_I a S_Q jsou úrovně na výstupech LIA pro soufázovou a kvadraturní složku, G je celkový zisk LIA, a

Δf je šumová šíře pásma dolní propusti. Pro dolní propust s časovou konstantou τ a se sklonem 12 dB je

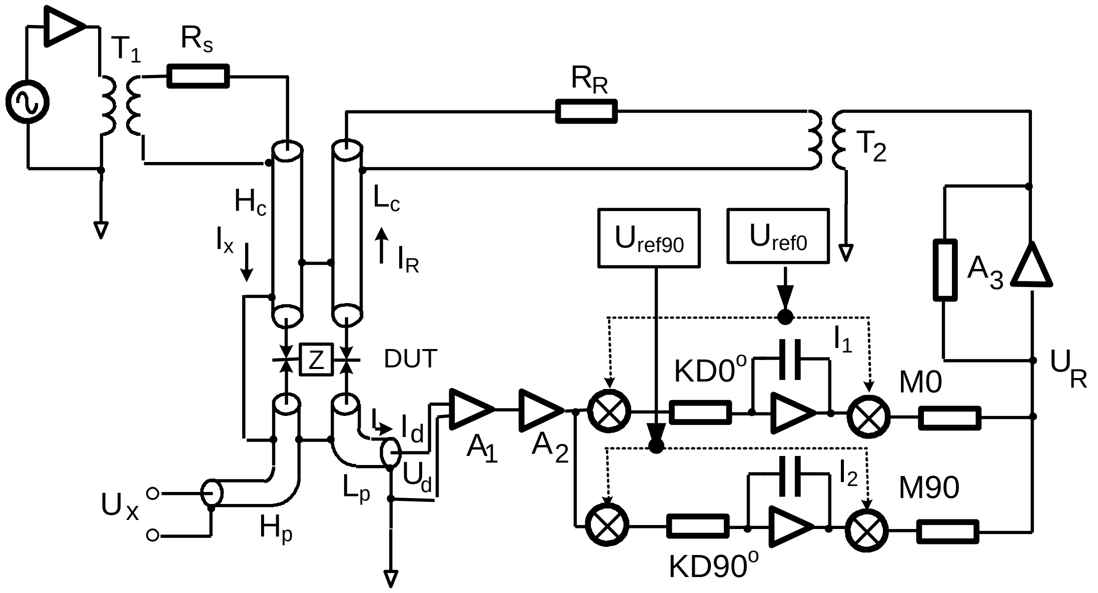
$$\Delta f = \frac{1}{8\tau}$$

Reprodukovatelnost hodnot e_n je závislá na počtu vzorků (odměrů) N. Např. chybu 5% na konfidenční úrovni 95 % je nutné provést N = 800 odměrů.

3.1.3. *Číslicové měřiče impedancí*

Z celé řady možností číslicového měření impedancí metody automaticky vyvažovaných můstků. Pro přesnost těchto metod má základní význam vyloučení vlivu parazitních impedancí uchycení (držáku) měřeného objektu (DUT - *Device Under Test*) na panelu přístroje a příp. přívodů k DUT. Jelikož se zde užívá měření složek fázoru napětí metodou koherentní demodulace, má přesnost a stálost fáze referenčních napětí použitých koherentních demodulátorů zásadní význam.

Zjednodušené zapojení automaticky vyvažovaného můstku ("elektronický můstek") na Obrázku 3.2 zobrazuje podrobněji zejména připojení DUT upravenou čtyřsvorkovou metodou [8]. Proud I_x z můstkového zdroje protéká vnitřním odporem zdroje R_S, středním vodičem koaxiálního kabelu H_C, DUT, L_C, referenčním odporem R_R a vrací se po pláštích H_C , L_C zpět k můstkovému zdroji. Jde o záměrné paralelní uspořádání přívodů tak, aby proudy středními vodiči i plášti měly navzájem opačný směr s cílem dosáhnout minimální hodnoty magnetického pole a tedy i minimální vzájemné indukčnosti mezi přívody k DUT.



Obrázek 3.2 Princip elektronicky vyvažovaného můstku pro měření impedancí

Transformátory T_1 a T_2 slouží k oddělení stejnosměrného napětí připojovaného na DUT např. při měření závislosti kapacity na napětí u polovodičových

součástek. Vstupní napětí A_1 vyvolané proudem I_D je příznakem nerovnováhy elektronického můstku a po úpravách slouží ke generaci vyvažovacího napětí na sekundáru T_2. Výsledkem je dosažení rovnováhy, tj. $\mathbf{I}_D = 0$ a $\mathbf{I}_X = \mathbf{I}_R$. Při rovnováze je také hodnota vzájemné indukčnosti mezi přívody ideálně nulová.
Při harmonickém průběhu napájení rovnováha znamená rovnost fázorů $\mathbf{I}_X$ a $\mathbf{I}_R$, které lze dosáhnout pouze za předpokladu, že lze měnit obě složky fázoru $\mathbf{I}_R$. K tomu slouží dva *koherentní demodulátory* spínačového typu s referenčními napětími vzájemně posunutými o 90°, které provádějí rozklad napětí nerovnováhy na reálnou a imaginární složku. Použití koherentních demodulátorů současně dovoluje *regenerovat* signál k vyvažování i při hodnotách $\mathbf{I}_D$ ležících pod úrovní šumu.
Výstupní napětí demodulátorů (dolní propusti jsou zde realizovány integrátory) je v násobících členech modulováno na kmitočet napájecího napětí můstku. Takto vzniklý fázor slouží k dosažení můstkové rovnováhy, pro kterou platí

$$\mathbf{I}_x = \frac{\mathbf{U}_x}{\mathbf{Z}_x} = \mathbf{I}_R = -\frac{\mathbf{U}_R}{R_R} \Rightarrow \mathbf{Z}_x = -R_R \frac{\mathbf{U}_X}{\mathbf{U}_R} \tag{3.2}$$

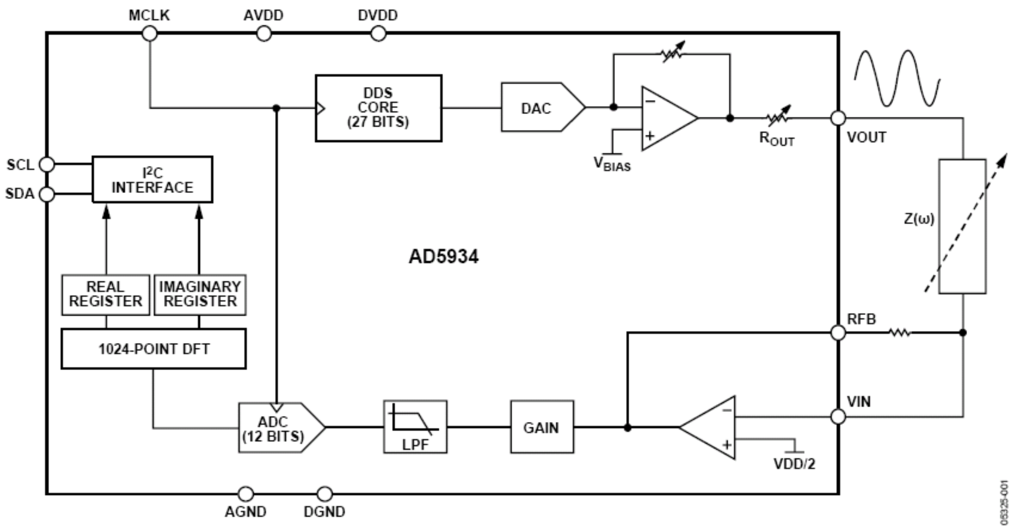
K určení impedance je nutné znát poměr napětí $\mathbf{U}_X/\mathbf{U}_R$ a spolehnout se na přesnost etalonového odporu R_R. Závislost hodnoty $\mathbf{Z}_X$ na *poměru* napětí je výhodná, jelikož potlačuje vliv kolísání napájecího napětí můstku.

Při měření na frekvencích asi nad 1 MHz se nepříznivě uplatňují přenosové vlastnosti přívodních kabelů, zdůrazněné např. nedokonalostí transformátoru T_1, nenulovou hodnotou vnitřního odporu zdroje R_S, dále zbytkovou impedancí kontaktů s DUT a parazitními signály pronikajícími na vstup zesilovače A_1. Při vyváženém můstku jsou živé vodiče kabelů L_C a L_P na nulovém potenciálu a jejich vliv je minimální. Přívody H_C, H_P a vstupní impedance obvodu měřícího $\mathbf{U}_X$ představují zátěž generátoru můstkového napětí a uplatní se pouze při větších hodnotách jeho vnitřního odporu R_S. Přesto je pro větší přesnosti (0,1 %) nutné pracovat s pevnou délkou přívodů, zpravidla lze zvolit 0 nebo 1 m, jinak může dojít k chybám až 30 %. V druhém případě je do cesty výstupu napětí $\mathbf{U}_R$ zařazen úsek kompenzačního koaxiálního kabelu.
Dalším prostředkem potlačení vlivu přívodů je procedura *automatické kalibrace,* probíhající na začátku měření. Hodnoty zbytkových impedancí, naměřené při otevřených a zkratovaných přípojných svorkách, se uloží do paměti a slouží ke korekci naměřené hodnoty výpočtem. Korekční vztahy jsou

poměrně složité a hlavně platné pouze za jistých zjednodušujících předpokladů. Proto se nelze na techniku automatické korekce při vyšších nárocích na přesnost zcela spolehnout a k potlačení vlivu impedancí přívodů použít současně všech dalších adekvátních prostředků.

Demodulátory KD 0°, KD 90° a modulátory M 0°, M 90° elektronického můstku spolehlivě pracují na frekvencích několik MHz. Pokud je nutné měřit impedanci nad tímto základním pásmem, lze směšováním rozšířit pásmo do několika desítek MHz. Pak součtovým směšovačem, vloženým za sumační zesilovač A_3, zvýšíme pracovní frekvenci můstku a změnou frekvence místního oscilátoru směšovače jej můžeme měnit v rozmezí základního rozsahu. Převod frekvence zpět do základního pásma zařídí rozdílový směšovač vložený před blok demodulátorů a modulátorů a pracující s identickým místním oscilátorem. Tento postup je blíže popsán v souvislosti s výkladem činnosti zapojení na Obrázku 2.5.

3.1.4. *Měření impedance integrovaným obvodem AD5934*



Obrázek 3.3 Impedanční konvertor AD5934–blokové schéma

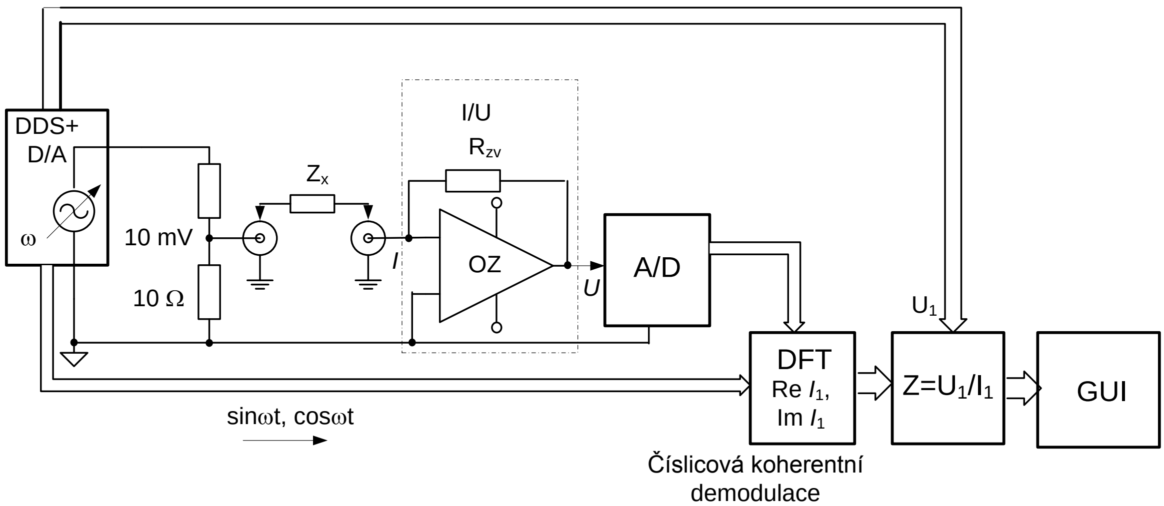
Koncepce obvodu je založena na koherentním zpracování signál, takže dovoluje měření i při nízkých úrovních signálu. Z číslicového generátoru signálu, realizovaného metodou DDS, jsou odvozeny dvě ortogonální harmonické referenční napětí použité pro výpočet první harmonické proudu impedancí algoritmem DFT.

V případě *nelineární* impedance, tj. s hodnotou závislou na přiloženém napětí, je impedance definována nejčastěji jako podíl prvních harmonických napětí

a proudu. Jedna z možností je výpočet první harmonické proudu po digitalizaci signálu číslicovou Fourierovou transformací [2].
Vztahy pro výpočet složek první harmonické Fourierovou transformací, tj.

$$a_1 = \frac{1}{2T}\int_{-T}^{T} i(t)\cos\omega t\,dt,\; b_1 = \frac{1}{2T}\int_{-T}^{T} i(t)\sin\omega t\,dt \tag{3.3}$$

jsou v podstatě ekvivalentní procesu koherentní demodulace, jelikož jde o násobení proudu $i(t)$ referenčními průběhy $\sin\omega t$ a $\cos\omega t$, odvozenými ze zdroje pracovního napětí.
Tento součin je, v souladu s definicí koherentní demodulace, podroben operaci střední hodnota.
Obvod AD 5934 byl použit pro měření impedance buněk ($\mathbf{Z}_x$). Bezpečná úroveň pracovního napětí je získána odporovým děličem (Obrázek 3.4). Vliv parazitních impedancí přívodů je potlačen převodníkem proud/napětí (I/U). Výstupní napětí převodníku je po digitalizaci podrobeno Fourierovy transformaci a její výstupní údaje slouží pro výpočet reálné a imaginární složky, resp. amplitudy a fáze impedance.
Měření impedance je řízeno mikroprocesorem s programovým vybavením odpovídajícím zásadám GUI a připojeným k PC rozhraním USB.



Obrázek 3.4 Implementace IO AD5934 pro měření impedance buněk

Seznam použité literatury

[1] Ďaďo S., Vedral J.: Číslicová měření-metody a přístroje, Vydavatelství ČVUT, Praha 2002

[2] Ďaďo S.:Měření impedance buněk. In: Sborník Nové smery zpracovania signálov, NSZS 2010, Tatranské Zruby květen 2010

[3] Ďaďo S., Bejček L., Platil A.: Měření průtoku a výšky hladiny. Nakladatelství BEN,2006

[4] Ďaďo S, Janošek M,: Single lockin detection of AC magnetic field by fluxgate sensors .Sensors and Actuators. No 4., Vol.152, 2010

[5] Ďaďo S,:Číslicová koherentní demodulace, Doktorská dizertační práce, ČVUT, Praha 1987

[6] Model 162 Boxcar Averager System. EG&G Princeton Applied Research,1987

[7] Steven M.Kay,Sudhaker R.: A Zero Crossing – Based Spestrum Analyzer.IEEE Transaction s on Acoustics, Speech and Signal Processing, Vol.ASSP-14,NO.1.February 1986, pp. 96-104

[8] HP 3588A Spectrum Analyzer, Hewlett Packard Journal, June 1991, pp.46-48

[9] Ďaďo, S.: Dithering - a Special Case of Non-Linearity Correction In: Dithering in Measurement: Theory and Applications. Prague : CTU, Faculty of Electrical Engineering, Department of Measurements, 1998, p. 6-10. ISBN 80-01-01806-7.

[10] Ďaďo, S. - Kreidl, M.: Sensors and Measuring Circuits 1. ed. Prague : CTU, 1996. 315 p. ISBN 80-01-01500-9. (in Czech). ČVUT: Faculty of Electrical Engineering, Department of Measurements 1996

[11] Ďaďo, S.: Zpracování výstupních signálů senzorů koherentními metodami. In: Pragoregula '98. Praha : Masarykova akademie, 1998, p. 23-26. ISBN 80-902131-2-X.

PŘÍLOHY

P1. ODHAD MINIMALIZUJÍCÍ STŘEDNÍ KVADRATICKOU HODNOTU ODCHYLKY (MMSE)

V dalším výkladu jsou vektory uvažovány jako *sloupcové* a bude použito následujícího označení:

$\mathbf{x}$ **-** vektor, tvořený složkami, jejichž hodnoty jsou cílem odhadu, rozměr $[n \times 1]$
$\mathbf{y}$ **-** vektor naměřených dat obsahujících informaci o vektoru $\mathbf{x}$**,** rozměr $[n \times 1]$

$p(\mathbf{x}, \mathbf{y})$ **-**sdružená hustota pravděpodobnosti vektorů $\mathbf{x}$ a $\mathbf{y}$ (sdružené rozdělení)

Z vektoru dat $\mathbf{y}$ určíme vektor odhadu $\hat{\mathbf{x}}_{MS}(\mathbf{y})$ tak, aby střední kvadratická hodnota chyby (Mean Square Error) byla minimální. Odhad dle tohoto kritéria, ve zkratce odhad MMSE (Minimum MSE), tedy minimalizuje výraz

$$J_{MS} = \iint (\mathbf{x} - \hat{\mathbf{x}}_{MS}(\mathbf{y}))\left(\mathbf{x} - \hat{\mathbf{x}}_{MS}(\mathbf{y})\right)^T d\mathbf{x}d\mathbf{y} \tag{1}$$

Upravme tento vztah podle definice s využitím sdruženého rozdělení, tj.

$$J_{MS} = \iint (\mathbf{x} - \hat{\mathbf{x}}_{MS}(\mathbf{y}))\left(\mathbf{x} - \hat{\mathbf{x}}_{MS}(\mathbf{y})\right)^T p(\mathbf{x}, \mathbf{y}) d\mathbf{x}d\mathbf{y} \tag{2}$$

Dosaďme za sdružené rozdělení výraz

$$p(\mathbf{x}, \mathbf{y}) = p(\mathbf{x} \mid \mathbf{y}) p(\mathbf{y}) \tag{3}$$

$$J_{MS} = \iint (\mathbf{x} - \hat{\mathbf{x}}_{MS}(\mathbf{y}))\left(\mathbf{x} - \hat{\mathbf{x}}_{MS}(\mathbf{y})\right)^T p(\mathbf{x} \mid \mathbf{y}) p(\mathbf{y}) d\mathbf{x}d\mathbf{y} \tag{4}$$

Vnitřní integrál a rozdělení $p(\mathbf{y})$ jsou nezáporné, proto minimalizace vnitřního integrálu pro každé $\mathbf{y}$ odpovídá minimalizaci výrazu

$$J'_{MS} = \int (\mathbf{x} - \hat{\mathbf{x}}_{MS}(\mathbf{y}))\left(\mathbf{x} - \hat{\mathbf{x}}_{MS}(\mathbf{y})\right)^T p(\mathbf{x} \mid \mathbf{y}) d\mathbf{x} \tag{5}$$

Upravme vztah (5) na základě definice operátora střední hodnoty $E\{.\}$

$$J'_{MS} = E\left\{\mathbf{x}\mathbf{x}^{\mathrm{T}} | \mathbf{y}\right\} - E\{\mathbf{x} \mid \mathbf{y}\}\, \hat{\mathbf{x}}_{MS}(\mathbf{y})^T - \hat{\mathbf{x}}_{MS}(\mathbf{y}) E\left\{\mathbf{x}^T \mid \mathbf{y}\right\} + \hat{\mathbf{x}}_{MS}(\mathbf{y})\ \hat{\mathbf{x}}_{MS}(\mathbf{y})^T \tag{6}$$

Operátor podmíněné střední hodnoty je označen $E\{\mathbf{z}|\mathbf{p}\}$ (s významem střední hodnoty vektoru náhodné proměnné $\mathbf{z}$ za podmínky stálé hodnoty vektoru $\mathbf{p}$).

Minimum určíme obvyklým postupem hledání extrému

$$\frac{dJ'_{MS}}{d\hat{\mathbf{x}}_{MS}(\mathbf{y})} = -2E\{\mathbf{x} \mid \mathbf{y}\} + 2\hat{\mathbf{x}}_{MS}(\mathbf{y}) = 0 \tag{7}$$

Pro odhad splňující kritérium MMSE tedy platí

$$\hat{\mathbf{x}}_{MS}(\mathbf{y}) = \mathbf{E}\{\mathbf{x}|\mathbf{y}\} \tag{8}$$

Odhad minimalizující střední kvadratickou chybu je určen *podmíněnou střední hodnotou* odhadované veličiny $E\{\mathbf{x} \mid \mathbf{y}\}$ náhodného vektoru $\mathbf{x}$.

Lze ukázat, že pro střední hodnotu odhadu platí

$$E\{\hat{\mathbf{x}}_{MS}(\mathbf{y})\} = E\{\mathbf{x}\} \tag{9}$$

Vztah (9) vyjadřuje, že odhad je *nevychýlený* (centrovaný).

Odhad dle obecného kritéria MMSE se v praxi určuje obtížně, jelikož vyžaduje znalost sdruženého rozdělení $p(\mathbf{x}, \mathbf{y})$ a výpočet podmíněné střední hodnoty $E\{\mathbf{x} \mid \mathbf{y}\}$.

V řadě případů lze však použít (alespoň přibližně) snáze realizovatelného odhadu podle kritéria LMMSE, které bude dále odvozeno.

P.2 Odhady parametrů podle kritéria LMMSE

(<u>L</u>INEAR <u>M</u>INIMUM <u>M</u>EAN <u>S</u>QUARE <u>E</u>RROR)

Odhad dle tohoto kritéria je vhodný pro situace, kdy závislost mezi vektory **x** a vektorem naměřených hodnot **y** je lineární, tj. platí

$$\mathbf{x} = \mathbf{A}\mathbf{y} + \mathbf{b} \tag{10}$$

Přitom matice $\mathbf{A}$ ani vektor $\mathbf{b}$ nejsou závislé na odhadovaných parametrech ani na naměřených hodnotách. Hledaný vektor odhadovaných parametrů určíme opět lineárním vztahem

$$\hat{\mathbf{x}}_{LMS}(\mathbf{y}) = \boldsymbol{\alpha}\mathbf{y} + \boldsymbol{\beta} \tag{11}$$

Matice (vektory) konstant α, β se určí tak, aby bylo splněno kritérium LMMSE.

To znamená, že opět minimalizujeme výraz pro střední kvadratickou hodnotu odchylky odhadu

$$E\{(\mathbf{x} - \hat{\mathbf{x}}_{LMS})(\mathbf{x} - \hat{x}_{LMS})^T\} \tag{12}$$

Tento vztah odpovídá matici, jejíž stopa obsahuje hledané kvadráty rozdílů dat $\mathbf{x}$ a odhadovaného parametru $\hat{\mathbf{x}}_{LMS}$

Minimalizujeme za použití následujících pravidel pro derivaci stopy matice:

$$\frac{d}{d\mathbf{A}}\left\{tr\left[\mathbf{ABA}^T\right]\right\} = 2\mathbf{BA}$$
$$\frac{d}{d\mathbf{A}}\left\{tr\left[\mathbf{CAB}\right]\right\} = \mathbf{C}^T\mathbf{B}^T$$

Po dosazení za $\hat{\mathbf{x}}_{LMS} = \boldsymbol{\alpha}\mathbf{y} + \boldsymbol{\beta}$

do stopy matice (12) dostáváme

$$tr\left\{E\left[\mathbf{x}-\boldsymbol{\alpha}\mathbf{y}-\boldsymbol{\beta}\right]\left[\mathbf{x}-\boldsymbol{\alpha}\mathbf{y}-\boldsymbol{\beta}\right]^T\right\} = tr\left\{E\left[\mathbf{x}\mathbf{x}^T\right] - E\left[\boldsymbol{\alpha}^T\mathbf{x}\mathbf{y}^T\right] - E\left[\mathbf{x}\boldsymbol{\beta}^T\right] - E\left[\boldsymbol{\alpha}\mathbf{y}\mathbf{x}^T\right]\right\} +$$
$$+tr\left\{E\left[\boldsymbol{\alpha}\boldsymbol{\alpha}^T\mathbf{y}\mathbf{y}^T\right] + E\left[\boldsymbol{\alpha}\mathbf{y}\boldsymbol{\beta}^T\right] - E\left[\boldsymbol{\beta}\mathbf{x}^T\right] + \boldsymbol{\alpha}^T\boldsymbol{\beta}E\left[\mathbf{y}^T\right] + E\left[\boldsymbol{\beta}\boldsymbol{\beta}^T\right]\right\} =$$
$$= tr\left\{\mathbf{P}_{xx} - \boldsymbol{\alpha}^T\mathbf{P}_{xy} - E\left[\mathbf{x}\right]\boldsymbol{\beta}^T - \boldsymbol{\alpha}\mathbf{P}_{yx} + \boldsymbol{\alpha}\mathbf{P}_{yy}\boldsymbol{\alpha}^T + \boldsymbol{\alpha}E\left[\mathbf{y}\right]\boldsymbol{\beta}^T - \boldsymbol{\beta}E\left[\mathbf{x}^T\right] + \boldsymbol{\alpha}^T\boldsymbol{\beta}E\left[\mathbf{y}^T\right] + \boldsymbol{\beta}\boldsymbol{\beta}^T\right\} = tr\left\{\mathbf{M}\right\}$$
$$\frac{d}{d\boldsymbol{\beta}}tr\left\{\mathbf{M}\right\} = -E\left[\mathbf{x}\right] + \boldsymbol{\alpha}E\left[\mathbf{y}\right] + \boldsymbol{\beta}$$

Zde byly zavedeny kovarianční matice

$$\mathbf{P_{xx}} = \mathbf{E}[\mathbf{xx^T}], \mathbf{P_{yy}} = \mathbf{E}[\mathbf{xy^T}], \mathbf{P_{yy}} = \mathbf{E}[\mathbf{yy^T}] \qquad (13)$$

Minimální hodnoty $\boldsymbol{\beta}$ a $\boldsymbol{\beta}^T$ určíme ze vztahů

$$\boldsymbol{\alpha}E\left[\mathbf{y}\right] - E\left[\mathbf{x}\right] + \boldsymbol{\beta} = 0 \Rightarrow \boldsymbol{\beta} = E\left[\mathbf{x}\right] - \boldsymbol{\alpha}E\left[\mathbf{y}\right]$$
$$\boldsymbol{\beta}^T = E\left[\mathbf{x}^T\right] - E\left[\mathbf{y}^T\right]\boldsymbol{\alpha}^T$$

Využitím vztahů pro $\boldsymbol{\beta}$ *a* $\boldsymbol{\beta}^T$

$$tr\left\{\mathbf{P}_{xx} - \boldsymbol{\alpha}^T\mathbf{P}_{xy} + \boldsymbol{\alpha}\mathbf{P}_{yy}\boldsymbol{\alpha}^T - \boldsymbol{\alpha}\mathbf{P}_{yx} + \boldsymbol{\beta}^T\left(\boldsymbol{\alpha}E\left[\mathbf{y}\right] - E\left[\mathbf{x}\right]\right) + \boldsymbol{\beta}(E\left[\mathbf{y}^T\right]\boldsymbol{\alpha}^T - E\left[\mathbf{x}^T\right]) + \boldsymbol{\beta}\boldsymbol{\beta}^T\right\} =$$
$$= tr\left\{\mathbf{P}_{xx} - \boldsymbol{\alpha}^T\mathbf{P}_{xy} + \boldsymbol{\alpha}\mathbf{P}_{yy}\boldsymbol{\alpha}^T - \boldsymbol{\alpha}\mathbf{P}_{yx} - \boldsymbol{\beta}\boldsymbol{\beta}^T - \boldsymbol{\beta}\boldsymbol{\beta}^T + \boldsymbol{\beta}\boldsymbol{\beta}^T\right\} = tr\left\{\mathbf{P}_{xx} - \boldsymbol{\alpha}^T\mathbf{P}_{xy} - \boldsymbol{\alpha}\mathbf{P}_{yx}^T + \boldsymbol{\alpha}\mathbf{P}_{yy}\boldsymbol{\alpha}^T - \boldsymbol{\beta}\boldsymbol{\beta}^T\right\} =$$
$$= tr\left\{\mathbf{N}\right\}$$

$$\frac{d}{d\boldsymbol{\alpha}}tr\left\{\mathbf{N}\right\} = -\mathbf{P}_{yx}^T + \mathbf{P}_{yy}^T\boldsymbol{\alpha}$$
$$-\mathbf{P}_{yx}^T + \boldsymbol{\alpha}\mathbf{P}_{yy}^T = 0 \Rightarrow \boldsymbol{\alpha} = \mathbf{P}_{yx}^T(\mathbf{P}_{yy}^{-1})^T$$
$$\boldsymbol{\alpha}^T = \mathbf{P}_{yx}\mathbf{P}_{yy}^{-1}$$
$$\hat{\mathbf{x}}_{LMS} = \boldsymbol{\alpha}\mathbf{y} + \boldsymbol{\beta} = \boldsymbol{\alpha}\mathbf{y} + E\left[\mathbf{x}\right] - \boldsymbol{\alpha}E\left[\mathbf{y}\right] = E\left[\mathbf{x}\right] + \boldsymbol{\alpha}\left(\mathbf{y} - E\left[\mathbf{y}\right]\right) = \boldsymbol{\mu}_x + \mathbf{P}_{yx}^T(\mathbf{P}_{yy}^{-1})^T\left(\mathbf{y} - \boldsymbol{\mu}_y\right) =$$
$$= \boldsymbol{\mu}_x + \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}\left(\mathbf{y} - \boldsymbol{\mu}_y\right)$$

Použijme nové symboly pro střední hodnoty

$$\boldsymbol{\mu}_x = \boldsymbol{E}[\boldsymbol{x}] \text{ a } \boldsymbol{\mu}_y = \boldsymbol{E}[\boldsymbol{y}],$$

a využijme symetrie matice $\mathbf{P}_{yy}$ a rovnosti $\mathbf{P}_{yx}^T = \mathbf{P}_{xy}$.Pro odhadovaný vektor parametrů pak platí

$$\hat{\mathbf{x}}_{LMS} = \boldsymbol{\mu}_x + \mathbf{P}_{xy}^T \mathbf{P}_{yy}^{-1} (\mathbf{y} - \boldsymbol{\mu}_y) \tag{14}$$

Chybu odhadu získáme dosazením do rovnice (12) za vektor $\mathbf{x}_{LMS}$

$$E\left\{ [\mathbf{x} - \hat{\mathbf{x}}_{LMS}] \; [\mathbf{x} - \hat{\mathbf{x}}_{LMS}]^T \right\} = E\left\{ (\mathbf{x} - \boldsymbol{\mu}_x)(\mathbf{x} - \boldsymbol{\mu}_x)^T \right\} -$$
$$-E\left\{ 2(\mathbf{x} - \boldsymbol{\mu}_x)\mathbf{P}_{xy}^T \mathbf{P}_{yy}^{-1} \right\} + E\left\{ (\mathbf{y} - \boldsymbol{\mu}_y)\mathbf{P}_{xy}^T \mathbf{P}_{yy}^{-1} ((\mathbf{y} - \boldsymbol{\mu}_y)\mathbf{P}_{xy}^T \mathbf{P}_{yy}^{-1}))^T \right\}$$

$$= \mathbf{P}_{xx} - 2\mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}\mathbf{P}_{xy}^T + \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}(\mathbf{P}_{xy}\mathbf{P}_{yy}^{-1})^T E\left\{ (\mathbf{y} - \boldsymbol{\mu}_y)(\mathbf{y} - \boldsymbol{\mu}_y)^T \right\} =$$
$$| \, E\left\{ (\mathbf{y} - \boldsymbol{\mu}_y)(\mathbf{y} - \boldsymbol{\mu}_y)^T \right\} = \mathbf{P}_{yy} \, | \tag{14a}$$
$$= \mathbf{P}_{xx} - \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}\mathbf{P}_{xy}^T = \mathbf{P}_{xx} - \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}\mathbf{P}_{yx}$$

Výsledný vztah pro chybu odhadu má pak tvar

$$E\left\{ [\mathbf{x} - \hat{\mathbf{x}}_{LMS}] \, [\mathbf{x} - \hat{\mathbf{x}}_{LMS}]^T \right\} = \mathbf{P}_{xx} - \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}\mathbf{P}_{xy}^T = \mathbf{P}_{xx} - \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}\mathbf{P}_{yx}$$

Chyba odhadu dle kritéria LMMSE je *ortogonální* k hodnotám dat **y,** tj. platí

$$E\left\{ (\mathbf{x} - \hat{\mathbf{x}}_{LMS})\mathbf{y}^{\mathrm{T}} \right\} = 0 \tag{15}$$

Pro střední hodnotu chyby odhadu platí

$$E\left\{ \mathbf{1}.\left((\mathbf{x} - \boldsymbol{\mu}_x) - \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \right)^{\mathrm{T}} \right\} = E\left\{ \mathbf{1}.\mathbf{x}^{\mathrm{T}} \right\} - \boldsymbol{\mu}_x^{\mathrm{T}} - \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}\left(E\left\{ \mathbf{1}.\mathbf{y}^{\mathrm{T}} \right\} - \boldsymbol{\mu}_y^{\mathrm{T}} \right) = 0 \tag{16}$$

Proto platí také

$$E\left\{ \; \mathbf{y}\left(\mathbf{x} - \hat{\mathbf{x}}_{\mathrm{LMS}}(\mathbf{y}) \right)^{\mathrm{T}} \right\} = E\left\{ \; \mathbf{y}\left((\mathbf{x} - \boldsymbol{\mu}_x) - \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \right)^{\mathrm{T}} \right\} = 0 \tag{17}$$

Prakticky to znamená, že korelace chyby odhadu a naměřených dat je nulová. Odhad je *ortogonální* ke všem lineárním funkcím dat

$$g_{lin}(\mathbf{y}) = \mathbf{A}\mathbf{y} + \mathbf{b}$$

P.3. ZJEDNODUŠENÝ TVAR VZTAHU PRO ODHAD LMMSE

V situacích, kdy je na rozdíl od rovnice (10) vektor dat $\mathbf{y}$ explicitně vyjádřen jako lineární funkce vektoru odhadovaných parametrů $\mathbf{x}$, tj. ve tvaru

$$\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{e} \tag{18}$$

v němž $\mathbf{M}$ je matice konstant o rozměru $[n \times p]$, p je počet odhadovaných parametrů, a $\mathbf{e}$ je vektor rušení s kovarianční maticí

$$\mathbf{V} = \mathrm{var}(\mathbf{e}) = E\left\{ \mathbf{e}^{\mathrm{T}}\mathbf{e} \right\} \tag{19}$$

Odhad LMMSE je možné určit z tzv. *normální rovnice*

$$\mathbf{M}^{\mathrm{T}}\,\mathbf{M}\mathbf{x} = \mathbf{M}^{\mathrm{T}}\,\mathbf{y} \tag{20}$$

Za předpokladu že jde o rušení typu bílý šum, platí vztahy

$$E\{\mathbf{e}\} = 0 \qquad \text{a} \qquad \operatorname{var}\mathbf{e} = E\{\mathbf{e}^{\mathrm{T}}\mathbf{e}\} = \mathbf{V} = \sigma^2\mathbf{I} \tag{21}$$

Pro vektor odhadovaných parametrů pak platí

$$\hat{\mathbf{x}}_{\mathrm{LMS}} = (\mathbf{M}^{\mathrm{T}}\mathbf{M})^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{y} \tag{22}$$

V obecném případě, pokud kovarianční matice rušení $\mathbf{V}$ není jednotková, se vztah pro odhad modifikuje na tvar

$$\hat{\mathbf{x}}_{\mathrm{LMS}} = (\mathbf{M}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{M})^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{y} \tag{23}$$

Rozptyl odhadu je obecně dán vztahem

$$\operatorname{var}\hat{\mathbf{x}} = \sigma^2(\mathbf{M}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{M})^{-1} \tag{24}$$

Vyjádření odhadu v tomto tvaru je fyzikálně méně názorné, jelikož v něm nejsou explicitně použity vztahy pro variance a kovariance veličin.

P.4. ÚPRAVA ODHADU LMMSE PRO NELINEÁRNÍ FUNKCE

Odhadu LMMSE platného pro lineární vztah mezi vektorem dat $\mathbf{y}$ a vektorem odhadovaných parametrů $\mathbf{x}$ je s možné s jistou přibližností použít také pro případ obecné funkční závislosti mezi $\mathbf{x}$ a $\mathbf{y}$ vyjádřené zápisem

$$f(\mathbf{x},\mathbf{y}) + \mathbf{e} = 0 \tag{25}$$

Vztah (24) může být aproximován prvním členem Taylorova rozvoje

$$f(\mathbf{x},\mathbf{y}) + \mathbf{e} \cong f(\hat{\mathbf{x}}_0,\mathbf{y}) + (\mathbf{x}-\hat{\mathbf{x}}_0)\frac{df(\mathbf{x},\mathbf{y})}{d\mathbf{x}}\Big|_{\hat{\mathbf{x}}_0} + \mathbf{e} = 0 \tag{26}$$

Zde $\hat{\mathbf{x}}_0$ znamená počáteční hodnotu vektoru odhadovaných parametrů.

Rovnici (25) upravíme na výchozí tvar pro odhadování

$$\mathbf{W}(\mathbf{x}-\hat{\mathbf{x}}_0) + \mathbf{f}_0 \approx 0 \tag{27}$$

kde $\mathbf{W}$ je matice s rozměrem $[n \times p]$, složená z prvních derivací funkce podle odhadovaných parametrů, tj. $f'(\mathbf{x},\mathbf{y})|_{\hat{\mathbf{x}}_0}$ je vektor hodnot první derivace funkce v bodě počátečního odhadu $\hat{\mathbf{x}}_0$. Předpokládejme, že ve vztahu (26) má vektor $\mathbf{f}_0$ význam vektoru dat $\mathbf{y}$. Pak je možné pro odhad v dalším kroku použít vztahu (27), tj.

$$\mathbf{x} - \hat{\mathbf{x}}_0 = -(\mathbf{W}^{\mathrm{T}}\mathbf{W})^{-1}\mathbf{W}^{\mathrm{T}}\mathbf{f}_0 \tag{28}$$

Pro novou hodnotu odhadu platí

$$\mathbf{x}_n = \mathbf{x}_0 - (\mathbf{W}^{\mathrm{T}}\mathbf{W})^{-1}\mathbf{W}^{\mathrm{T}}\mathbf{f}_0 \tag{29}$$

Vypočtenou hodnotu $\mathbf{x}_n$ dosadíme místo $\mathbf{x}_0$ a po několika iteracích dosáhneme podstatně přesnější hodnoty odhadu.

Tento postup je označován jako *Gaussova – Seidelova iterace* a vyznačuje se přibližně kvadratickou konvergencí výsledků iteračních kroků.

Jiná metoda odhadu parametrů v nelineárních soustavách je založena na použití váhových koeficientů korigujících účinky nelinearit. Tohoto postupu se používá např. při odhadu frekvence a fáze harmonických signálů.

Centrum pro rozvoj výzkumu pokročilých řídicích a senzorických technologií
CZ.1.07/2.3.00/09.0031

Ústav automatizace a měřicí techniky
VUT v Brně
Kolejní 2906/4
612 00 Brno
Česká Republika

http://www.crr.vutbr.cz

info@crr.vutbr.cz