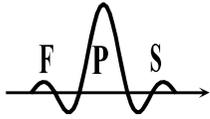




Sumário

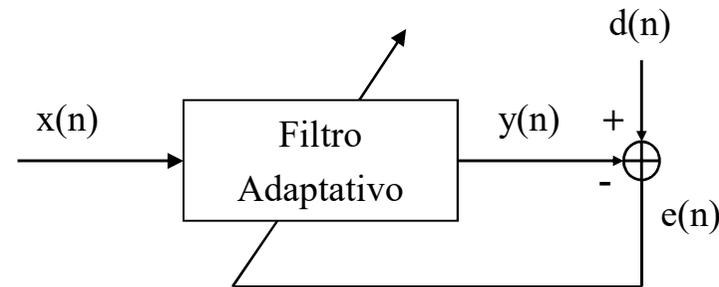
- *Filtragem adaptativa numa perspectiva prática*
 - *conceito e exemplos*
 - *algoritmos de gradiente*
 - *filtragem óptima de Wiener*
 - *o método do gradiente mais negativo*
 - *o algoritmo LMS*
 - *caso particular: eliminação de perturbação sinusoidal*



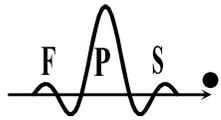
Filtros adaptativos

- Conceito

- um filtro adaptativo é um filtro cujos coeficientes são ajustados, de forma adaptativa, em função de objectivos ou condições variáveis no tempo e traduzidos num sinal de erro



- a aplicação típica consiste em suprimir ou modelizar certas componentes (possivelmente indesejáveis), representadas por $x(n)$ e projectadas num sinal $d(n)$, de acordo com algum critério estatístico incidindo sobre o sinal de erro $e(n)$
- o critério típico e prático para a adaptação dos coeficientes do filtro e optimização do seu desempenho é a minimização do valor médio quadrático do sinal de erro

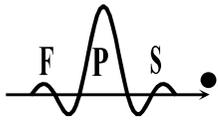


Conceito (*cont.*)

- para entradas estacionárias e usando o critério de minimização de erro médio quadrático, o filtro óptimo é único e designa-se por filtro de *Wiener*; para entradas não estacionárias, os filtros de *Kalman* são mais adequados e eficientes (mas também implicam maior complexidade)
- os filtros adaptativos podem ser do tipo FIR ou IIR, porém são quase sempre preferidas estruturas FIR já que, como se sabe, estas são intrinsecamente estáveis

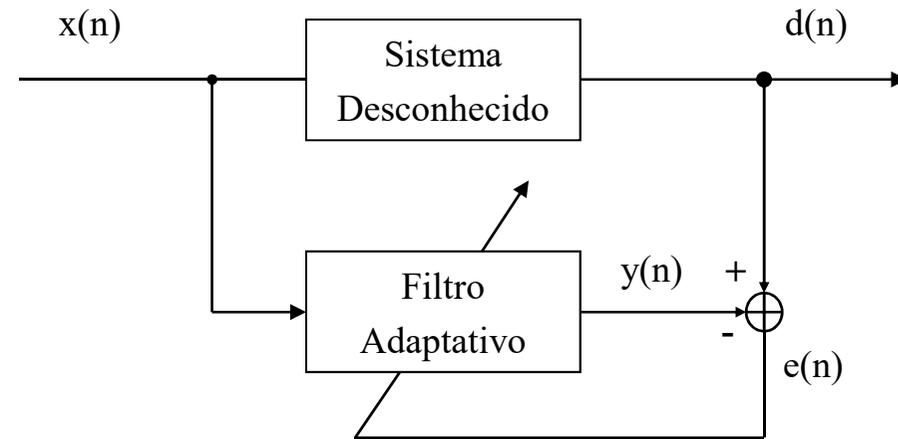
- Exemplos

- os filtros adaptativos são usados em diversas aplicações e contextos, tudo dependendo da forma como a resposta desejada do sistema adaptativo é caracterizada e extraída. Há essencialmente quatro tipos de aplicações: identificação, modelização inversa, predição, cancelamento de interferência.



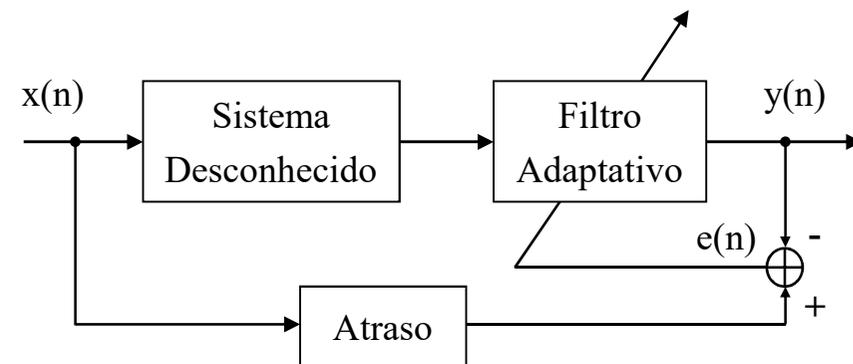
Identificação

- nestas aplicações, o filtro adaptativo é usado para fornecer um modelo linear que representa a melhor aproximação, de acordo com algum critério, a um sistema desconhecido



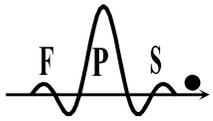
- **Modelização inversa (igualização)**

- nestas aplicações o filtro adaptativo tem por objectivo fornecer um modelo inverso que representa a melhor aproximação, de acordo com algum critério, a um sistema desconhecido



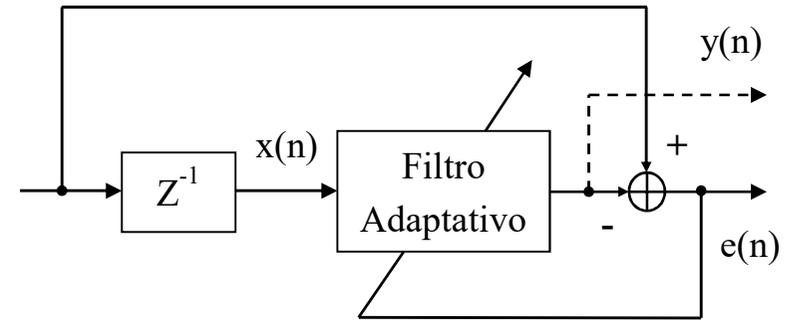
Exemplos de aplicações:

- descorrelação preditiva
- igualização adaptativa



Predição

- nestas aplicações, o filtro adaptativo tem por objectivo fornecer a melhor predição para a evolução de um sinal com base no conhecimento do seu passado

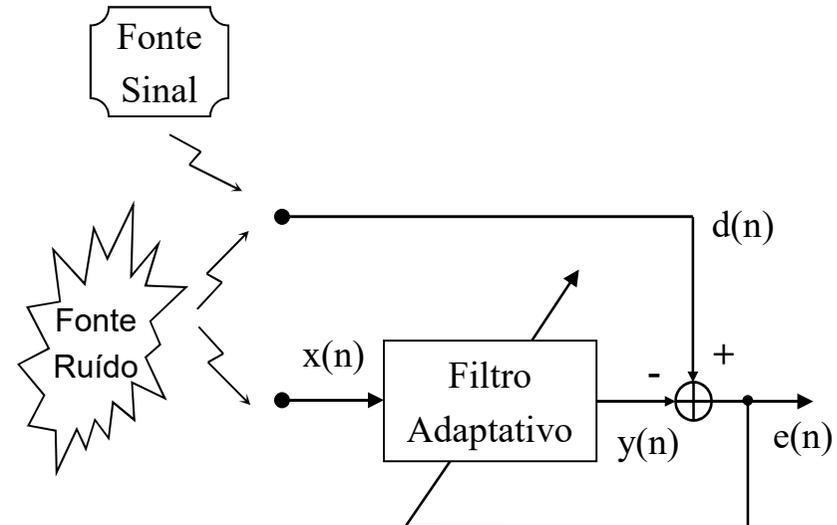


Exemplos de aplicações:

- codificação linear preditiva (LPC)
- ADPCM
- análise espectral autoregressiva (AR)

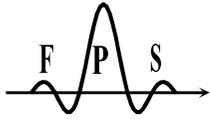
Cancelamento de interferência

- nestas aplicações o filtro adaptativo tem por objectivo cancelar uma interferência desconhecida e contida em $d(n)$ mas reflectida também no sinal de referência $x(n)$. O objectivo é simplesmente subtraír o “ruído” do sinal principal, de forma adaptativa, de modo a melhorar, à saída, a relação sinal-ruído.



Exemplos de aplicações:

- cancelamento adaptativo de ruído
- cancelamento de eco



Dado que o primeiro e último casos servem de base ao trabalho de laboratório, retiram-se a seguir algumas conclusões pertinentes a partir da análise aos sinais e operações implicados. Considerando que o sinal $d(n)=s(n)+x'(n)$ contém uma parcela de informação útil a preservar $s(n)$ e também uma componente de ruído a eliminar $x'(n)$ e que este é também projectado em $x(n)$, tomam-se, como hipóteses realistas, as seguintes:

$$E[s(n)x'(n-k)] = 0, \quad \forall k$$

$$E[s(n)x(n-k)] = 0, \quad \forall k$$

$$E[x(n)x'(n-k)] = p(+k)$$

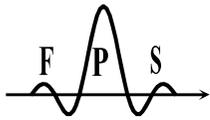
ou seja, $s(n)$ é descorrelacionado com qualquer uma das duas representações do ruído e estas, por sua vez, exibem uma correlação cruzada dada por $p(k)$ para uma distância k .

A saída do filtro adaptativo $y(n)$ e o sinal de erro $e(n)$ são dados por:

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

$$e(n) = d(n) - y(n) = s(n) + x'(n) - y(n)$$

o que coloca em evidência que o sinal $s(n)$ faz parte do sinal de erro. Porém, dadas as hipóteses anteriores, minimizar o erro médio quadrático de $e(n)$ equivale a minimizar o erro médio quadrático do ruído indesejado dado por $e(n)-s(n)=x'(n)-y(n)$, o que significa que $s(n)$ não é afectado pelo processo de filtragem adaptativa.

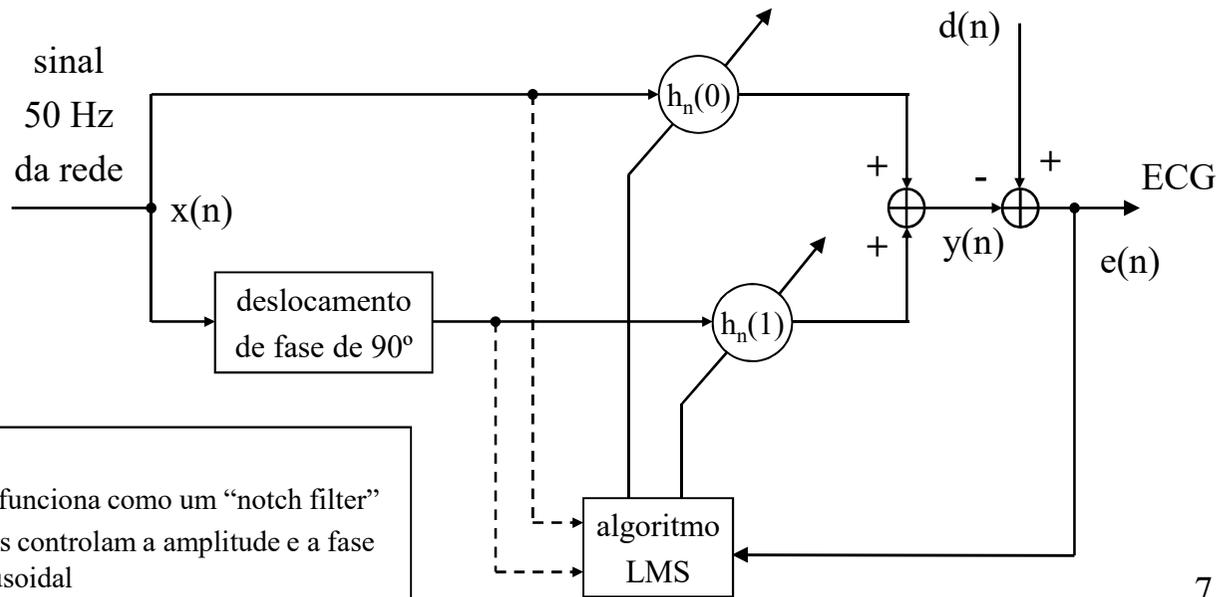


Como casos limite de operação do filtro adaptativo temos duas situações, a primeira em que $y(n)=x'(n)$ e a segunda em que $E[d(n)x(n-k)] = 0, \forall k$.

No primeiro caso a operação de filtragem é efectiva, cancelando todo o ruído. No segundo caso, o sinal de referência é completamente descorrelacionado com o sinal primário, fazendo com que o filtro adaptativo se “auto-desligue”, produzindo $y(n)=0$, o que não afecta o sinal principal.

– EXEMPLO 1

- Cancelamento da interferência sinusoidal de rede (50 Hz) num electrocardiograma usando somente dois coeficientes



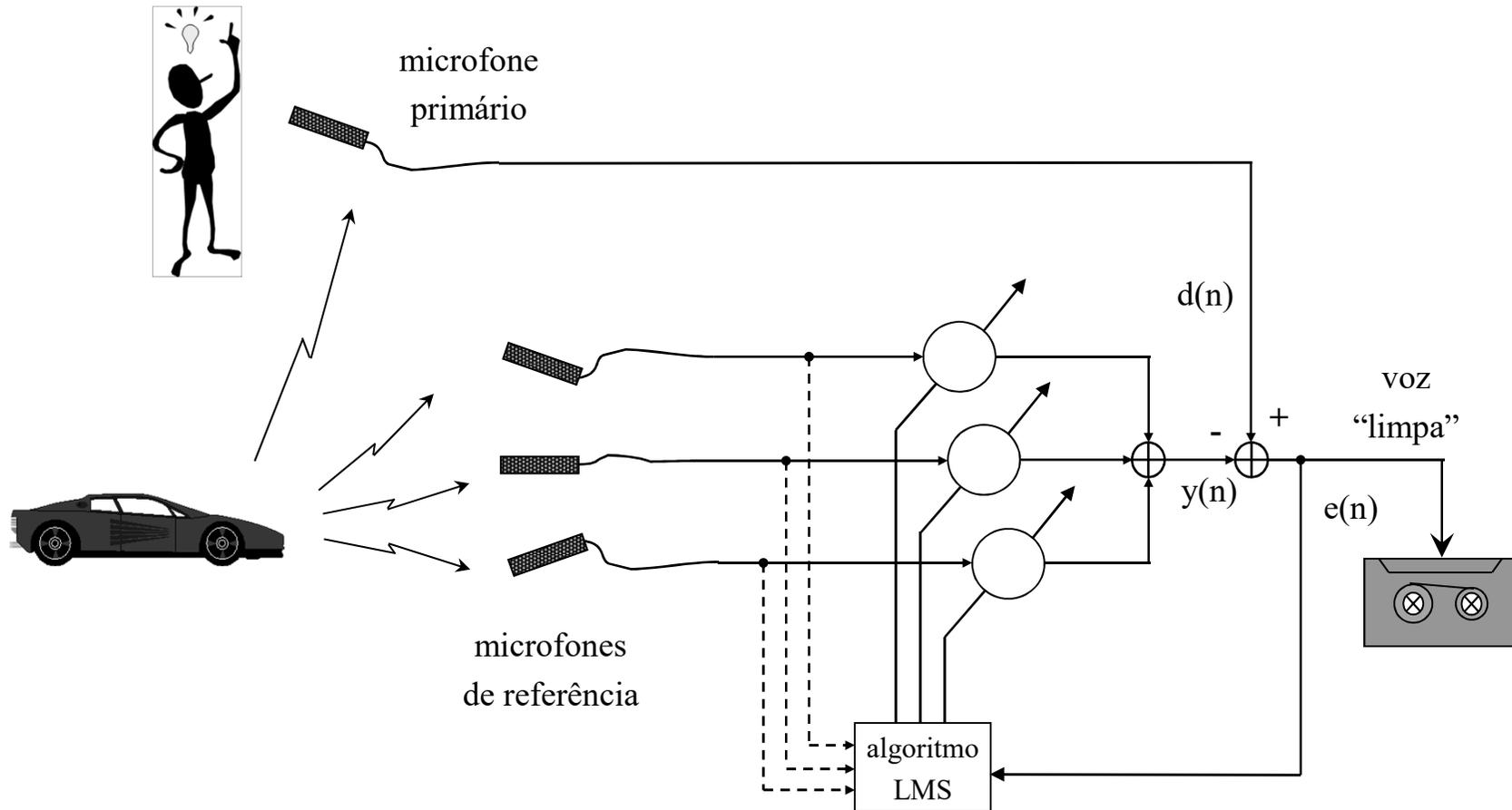
NOTAS:

- o filtro adaptativo funciona como um “notch filter”
- os dois coeficientes controlam a amplitude e a fase da interferência sinusoidal



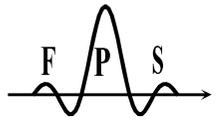
– EXEMPLO 2

- Redução / cancelamento de ruído acústico em sinais de voz (é possível melhorar a a relação sinal/ruído em mais do que 10 dB)



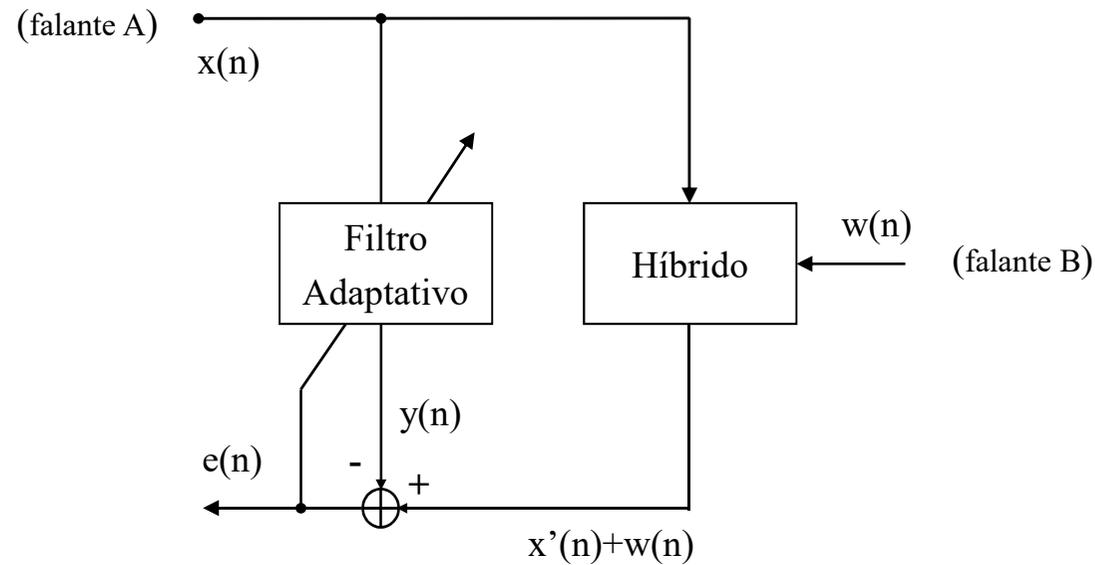
NOTAS:

- o algoritmo de adaptação pode ser o “LMS” ou outro alternativo
- os microfones de referência devem situar-se suficientemente longe do falante de modo a captarem quase exclusivamente só o ruído



– EXEMPLO 3

- Cancelamento de eco em circuitos telefónicos

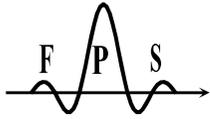


NOTAS:

- o objectivo do filtro adaptativo é sintetizar uma réplica do eco, perto do seu ponto de geração, e subtraí-la do sinal de retorno
- idealmente, o filtro adaptativo aproxima a função de transferência de retorno do eco

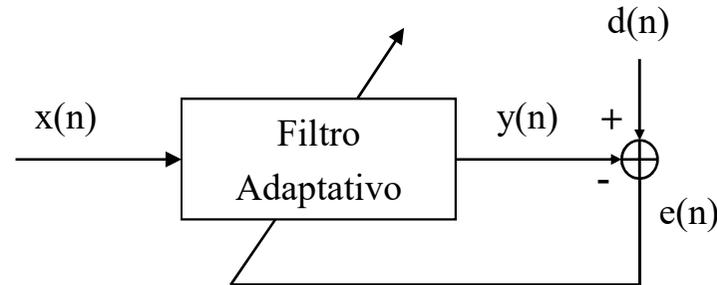
Nota introdutória sobre o algoritmo LMS, o que é afinal ?

trata-se de um algoritmo de gradiente estocástico, no sentido em que itera o peso de cada coeficiente de um filtro transversal (FIR) na direcção do gradiente negativo da função amplitude quadrada de um sinal de erro, em ordem ao peso desse coeficiente; detalhes seguem nos 'slides' seguintes ...



Algoritmos de Gradiente

- Filtragem de *Wiener*
 - dado o sistema adaptativo já ilustrado anteriormente:



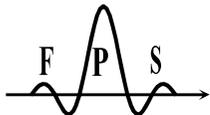
será:
$$e(n) = d(n) - y(n) = d(n) - \sum_{k=0}^{N-1} h_n(k)x(n-k)$$

admitindo as hipóteses anteriormente consideradas e admitindo vectores complexos, será para a função de erro a minimizar:

$$J = E\{e(n)e^*(n)\} = E\left\{ \left[d(n) - \sum_{k=0}^{N-1} h_n(k)x(n-k) \right] \left[d^*(n) - \sum_{\ell=0}^{N-1} h_n^*(\ell)x^*(n-\ell) \right] \right\}$$

$$J = E\left\{ \underbrace{d(n)^2}_{\sigma_d^2} \right\} - \sum_{\ell=0}^{N-1} h_n^*(\ell) \underbrace{E\{d(n)x^*(n-\ell)\}}_{p(+\ell)} - \sum_{k=0}^{N-1} h_n(k) \underbrace{E\{d^*(n)x(n-k)\}}_{p^*(+k)} + \sum_{k=0}^{N-1} \sum_{\ell=0}^{N-1} h_n(k) h_n^*(\ell) \underbrace{E\{x(n-k)x^*(n-\ell)\}}_{r(\ell-k)}$$

sendo σ^2 a variância de $d(n)$ cuja média se supõe nula, $p(\ell)$ e $p^*(k)$ representam a correlação cruzada entre o sinal $d(n)$ e a entrada $x(n)$ e $r(\ell-k)$ representa a correlação da entrada para uma distância $\ell-k$.



Abreviadamente será:
$$J = \sigma_d^2 - \sum_{\ell=0}^{N-1} h_n^*(\ell)p(+\ell) - \sum_{k=0}^{N-1} h_n(k)p^*(+k) + \sum_{k=0}^{N-1} \sum_{\ell=0}^{N-1} h_n(k)h_n^*(\ell)r(\ell - k)$$

o que revela que quando os sinais $x(n)$ e $d(n)$ são conjuntamente estacionários, a função do erro quadrático médio é uma função de 2ª ordem dos coeficientes do filtro FIR. Mostra-se que [Simon Haykin, “Adaptive Filter Theory”, Prentice-Hall, 1991] que se este filtro possuir N coeficientes, a superfície de erro é uma superfície hiperparabolóide de $N+1$ dimensões com um mínimo global e sem mínimos locais.

O mínimo global ocorre em J_{MIN} onde se verifica que o vector gradiente é identicamente nulo, ou seja:
$$\nabla_k(J) = 0, \quad k = 0, 1, \dots, N-1$$

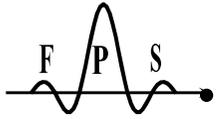
Assumindo coeficientes $h(k)$ reais, o vector gradiente,
$$\nabla_k(J) = \frac{\partial J}{\partial h(k)}$$
 concretizado na expressão J produz:

$$\nabla_k(J) = -2p(+k) + \sum_{\ell=0}^{N-1} h_n(\ell)r(\ell - k) + \sum_{s=0}^{N-1} h_n(s)r(k - s) = -2p(+k) + 2\sum_{\ell=0}^{N-1} h_n(\ell)r(\ell - k)$$

pelo que se conclui que os coeficientes óptimos são os que verificam:

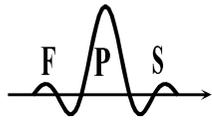
$$\sum_{\ell=0}^{N-1} h_n(\ell)r(\ell - k) = p(+k)$$

a solução deste sistema de equações é conhecida como a solução de Wiener-Hopf que fornece o filtro óptimo (filtro de Wiener).



O método do gradiente

- Em vez de usarmos o procedimento determinístico anterior que é aplicável a sinais estacionários para os quais se conhecem previamente as estatísticas associadas, podemos, não conhecendo estas, optar por um método iterativo de aproximação ao filtro óptimo de Wiener.
- O método iterativo de adaptação mais conhecido que converge para a solução óptima (filtro de Wiener), é o método de gradiente mais negativo (“steepest gradient” = *direction of the negative of the gradient vector*) que referiremos simplesmente pelo método do gradiente e que consiste no seguinte procedimento:
 1. Parte-se de um valor inicial para os coeficientes do filtro, tipicamente o vector nulo,
 2. Calcula-se o vector gradiente, determinado em ordem aos coeficientes do filtro no instante n ,
 3. Estimam-se os coeficientes do filtro para a próxima iteração modificando a estimativa actual no sentido oposto ao do vector gradiente,
 4. Retorna-se ao ponto 2 e repete-se o processo.



- De acordo com este procedimento, será então:
$$h_{n+1}(k) = h_n(k) + \frac{1}{2} \mu [-\nabla_k [J(n)]]$$
 em que μ é uma constante real e positiva que traduz o *passo de adaptação* do algoritmo.

Como vimos anteriormente,
$$\nabla_k [J(n)] = -2p(+k) + 2 \sum_{\ell=0}^{N-1} h_n(\ell) r(\ell - k)$$
 porém, como não conhecemos as estatísticas, podemos optar por uma expressão alternativa usando a definição de $p(k)$ e de $r(\ell - k)$ e admitindo dados e coeficientes reais:

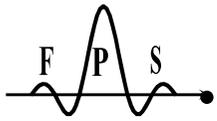
$$\nabla_k [J(n)] = -2E\{d(n)x(n-k)\} + 2 \sum_{\ell=0}^{N-1} h_n(\ell) E\{x(n-k)x(n-\ell)\}$$

$$\nabla_k [J(n)] = -2E\left\{x(n-k) \left[d(n) - \sum_{\ell=0}^{N-1} h_n(\ell) x(n-\ell) \right]\right\} = -2E\{x(n-k)e(n)\}$$

Este resultado traduz o princípio da ortogonalidade: para a solução óptima, $x(n-k)$ e $e(n)$ são ortogonais. Como corolário desta afirmação, resulta que para a solução óptima, $y(n)$ e $e(n)$ são ortogonais.

Será então:
$$h_{n+1}(k) = h_n(k) + \mu E\{x(n-k)e(n)\}$$

Esta é a solução do método iterativo de gradiente mais negativo que evolui suavemente para a solução óptima (*i.e.* para o filtro de Wiener).



O algoritmo LMS

- O algoritmo anterior presume a obtenção de estatísticas durante o procedimento iterativo e oferece a garantia de convergência suave para o filtro de Wiener. Podemos contudo ter, a partir do algoritmo anterior, uma outra aproximação à solução óptima baseada no gradiente “instantâneo” (dito *estocástico*), em que se elimina o operador de média:

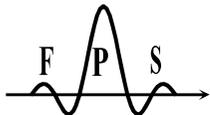
$$h_{n+1}(k) = h_n(k) + \mu x(n-k)e(n)$$

Dada a sua simplicidade, esta solução é muito popular e constitui o conhecido algoritmo de filtragem adaptativa LMS (de *Least Mean Squares*). Porém, o custo da simplicidade traduz-se em ruído de gradiente no cálculo recursivo de cada coeficiente do filtro adaptativo. Este ruído de gradiente pode conduzir (assumindo convergência):

1. Ou a uma solução aleatória do filtro adaptativo em torno do filtro óptimo de Wiener,
2. Ou a um deslocamento fixo do filtro adaptativo relativamente ao filtro óptimo (desalinhamento), comprometendo portanto o seu desempenho.

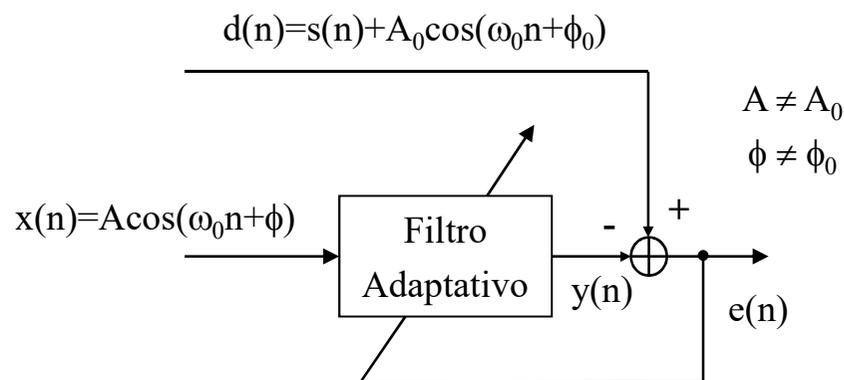
Estes aspectos são controlados através do factor de adaptação μ : um factor mais pequeno dá lugar a um menor ruído de gradiente o que minimiza o problema de desalinhamento, mas implica também, em contra-partida, uma convergência mais lenta para a solução óptima.

Para saber mais: há outros algoritmos de filtragem adaptativa que possuem melhor desempenho do que o LMS, e.g., o algoritmo *Recursive Least Squares* (RLS).



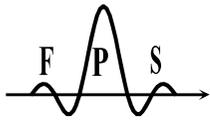
Caso de Eliminação de Perturbação sinusoidal

- O sistema seguinte tem por objectivo suprimir a interferência sinusoidal que corrompe a informação principal $s(n)$, contida no sinal $d(n)$, mesmo se a sua frequência ω_0 variar lentamente :



Supõe-se que $s(n)$ e $A_0 \cos(\omega_0 n + \phi_0)$ são descorrelacionados. Supõe-se ainda que $x(n)$ é um sinal fortemente correlacionado com a interferência sinusoidal que se pretende remover.

Os coeficientes do filtro FIR são adaptados de acordo com o algoritmo LMS. O algoritmo de adaptação usa a referência $x(n)$ para produzir à saída do filtro uma estimativa da sinusóide interferente que depois de subtraída a $d(n)$ deverá diminuir o efeito perturbador desta.



– Pelo algoritmo LMS, a adaptação dos coeficientes do filtro FIR processa-se de acordo com um cálculo iterativo usando as seguintes expressões:

- 1.
$$y(n) = \sum_{k=0}^{N-1} h_n(k)x(n-k)$$
- 2.
$$e(n) = d(n) - y(n)$$
- 3.
$$h_{n+1}(k) = h_n(k) + \mu x(n-k)e(n)$$

Sendo N o comprimento do filtro e μ o parâmetro de adaptação. Prova-se [Simon Haykin, “Adaptive Filter Theory”, Prentice-Hall, 1991] que o sistema ilustrado implementa o equivalente a um “*notch filter*” de 2ª ordem, colocando os zeros em $e^{\pm j\omega_0}$ e colocando, para μ pequeno, os polos aproximadamente em $(1 - \mu NA^2/4)e^{\pm j\omega_0}$. A qualidade do “*notch*” depende da proximidade entre os seus zeros e polos, convindo que μ seja pequeno para que a largura B (a 3 dB) do “*notch*” seja também pequena: $B \approx \mu NA^2/2$ rad. Estas aproximações são válidas se para a frequência ω_0 se verificar:

$$\frac{\sin(N\omega_0)}{N \sin(\omega_0)} \approx 0$$

← O que é que esta relação implica ? →

Uma regra prática para garantir convergência no processo de adaptação é considerar: $0 \leq \mu \leq \frac{1}{N \cdot P_x}$ sendo P_x a potência do sinal de

referência, estimada por: $P_x = \frac{1}{M} \sum_{m=0}^{M-1} x^2(m)$ para um dado inteiro M .