Chapter 3  Wiener Filters

1. Introduction

- Wiener filters are a class of optimum linear filters which involve linear estimation of a desired signal sequence from another related sequence.

- In the statistical approach to the solution of the linear filtering problem, we assume the availability of certain statistical parameters (e.g. mean and correlation functions) of the useful signal and unwanted additive noise. The problem is to design a linear filter with the noisy data as input and the requirement of minimizing the effect of the noise at the filter output according to some statistical criterion.

- A useful approach to this filter-optimization problem is to minimize the mean-square value of the error signal that is defined as the difference between some desired response and the actual filter output. For stationary inputs, the resulting solution is commonly known as the Weiner filter.

- The Weiner filter is inadequate for dealing with situations in which nonstationarity of the signal and/or noise is intrinsic to the problem. In such situations, the optimum filter has to be assumed a time-varying form. A highly successful solution to this more difficult problem is found in the Kalman filter.
2. Linear Estimation with Mean-Square Error Criterion

- Fig. 3.1 shows the block schematic of a linear discrete-time filter \( W(z) \) for estimating a desired signal \( d(n) \) based on an excitation \( x(n) \)

![Block diagram of a linear discrete-time filter](image)

- We assume that both \( x(n) \) and \( d(n) \) are random processes (discrete-time random signals). The filter output is \( y(n) \) and \( e(n) \) is the estimation error.

- To find the optimum filter parameters, the cost function or performance function must be selected. In choosing a performance function the following points have to be considered:
  
  (1) The performance function must be mathematically tractable.
  (2) The performance function should preferably have a single minimum so that the optimum set of the filter parameters could be selected unambiguously.

- The number of minima points for a performance function is closely related to the filter structure. The recursive (IIR) filters, in general, result in performance function that may have many minima, whereas the non-recursive (FIR) filters are guaranteed to have a single global minimum point if a proper performance function is used.

- In Weiner filter, the performance function is chosen to be
  \[
  \xi = E[|e(n)|^2] 
  \]

This is also called “mean-square error criterion”
3. Wiener Filter - the Transversal, Real-Valued Case

- Fig. 3.2 shows a transversal filter with tap weights $w_0, w_1, \ldots, w_{N-1}$.

Let $\mathbf{W} = [w_0 \ w_1 \ \ldots \ w_{N-1}]^T$

$\mathbf{X}(n) = [x_n \ x_{n-1} \ \ldots \ x_{n-N+1}]^T \ \ldots(1)$

The output is $y(n) = \sum_{i=0}^{N-1} w_i x(n-i)$

$= \mathbf{W}^T \mathbf{X}(n)$

$= \mathbf{X}_i^T(n) \mathbf{W} \ \ldots(2)$

Thus we may write

$e(n) = d(n) - y(n)$

$= d(n) - \mathbf{X}_i^T(n) \mathbf{W} \ \ldots(3)$

The performance function, or cost function, is then given by

$\xi = E[e^2(n)]$

$= E[(d(n) - \mathbf{W}^T \mathbf{X}(n))(d(n) - \mathbf{X}_i^T(n) \mathbf{W})]$ 

$= E[d^2(n)] - \mathbf{W}^T E[\mathbf{X}(n)d(n)] - E[\mathbf{X}_i^T(n)d(n)] \mathbf{W} + \mathbf{W}^T E[\mathbf{X}_i(n)\mathbf{X}_i^T(n)] \mathbf{W} \ \ldots(4)$

Now we define the Nx1 cross-correlation vector

$\mathbf{P} \equiv E[\mathbf{X}(n)d(n)]$

$= [p_0 \ p_1 \ \ldots \ p_{N-1}]^T \ \ldots(5)$
and the \( N \times N \) autocorrelation matrix

\[
\begin{bmatrix}
    r_{00} & r_{01} & \cdots & r_{0,N-1} \\
    r_{10} & r_{11} & \cdots & r_{1,N-1} \\
    \vdots & \vdots & & \vdots \\
    r_{N-1,0} & r_{N-1,1} & \cdots & r_{N-1,N-1}
\end{bmatrix}
\]

...(6)

Also we note that

\[
E[d(n)X^T(n)] = \bar{P}^T
\]

\[
\bar{W}^T \bar{P} = \bar{P}^T \bar{W}
\]

Then we obtain

\[
\xi = E[d^2(n)] - 2\bar{W}^T \bar{P} + \bar{W}^T \bar{R} \bar{W}
\]

...(7)

Equation (7) is a quadratic function of the tap-weight vector \( \bar{W} \) with a single global minimum.

Note that \( R \) has to be a positive definite matrix in order to have a unique minimum point in the \( w \)-space.

- Minimization of performance function

To obtain the set of tap weights that minimize the performance function,

we set \( \frac{\partial \xi}{\partial w_i} = 0 \), for \( i = 0,1,\ldots,N-1 \)

or

\[
\nabla \xi = \bar{0}
\]

...(8)

where \( \nabla \) is the gradient vector defined as the column vector

\[
\nabla = \left[ \frac{\partial}{\partial w_0}, \frac{\partial}{\partial w_1}, \ldots, \frac{\partial}{\partial w_{N-1}} \right]^T
\]

and zero vector \( \bar{0} \) is defined as \( N \)-component vector

\[
\bar{0} = [0 \; 0 \; \ldots \; 0]^T
\]

Equation (7) can be expanded as

\[
\xi = E[d^2(n)] - 2\sum_{l=0}^{N-1} p_l w_l + \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} w_l w_m r_{lm}
\]

...(9)

and

\[
\sum_{l=0}^{N-1} \sum_{m=0}^{N-1} w_l w_m r_{lm}
\]

can be expanded as

\[
\sum_{l=0}^{N-1} \sum_{m=0}^{N-1} w_l w_m r_{lm} = \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} w_l w_m r_{lm} + w_i^2 r_{ii} + w_i \sum_{m \neq i}^{N-1} w_m r_{im} + w_i \sum_{l \neq i}^{N-1} w_l r_{il}
\]

...(10)
Then we obtain
\[
\frac{\partial \xi}{\partial w_i} = -2p_i + \sum_{t=0}^{N-1} W_t (r_t + r_i) \quad \text{...(11)}
\]
for \( i = 0,1,2,\ldots, N - 1 \)

By setting \( \frac{\partial \xi}{\partial w_i} = 0 \), we obtain
\[
\sum_{t=0}^{N-1} W_t (r_t + r_i) = 2p_i \quad \text{...(12)}
\]

Note that
\[ r_i = E[x(n-l)x(n-i)] = \Phi_{x_x}(i-l) \]

The symmetry property of autocorrelation function of real-valued signal,
we have the relation \( r_i = r_d \)

Equation (12) then becomes
\[
\sum_{t=0}^{N-1} r_d w_i = p_i \quad \text{for } i = 0,1,2,\ldots, N - 1
\]

In matrix notation, we then obtain
\[
R \bar{W}_{op} = \bar{P} \quad \text{...(13)}
\]

where \( \bar{W}_{op} \) is the optimum tap-weight vector

Equation (13) is also known as the Wiener-Hopf equation,

which has the solution \( \bar{W}_{op} = R^{-1} \bar{P} \quad \text{...(14)} \)

Assuming that \( R \) has inverse.

The minimum value of \( \xi \) is
\[
\xi_{\min} = E[d^2(n)] - \bar{W}_{op}^T \bar{P}
\]
\[
= E[d^2(n)] - \bar{W}_{op}^T R \bar{W}_{op} \quad \text{...(15)}
\]

Equation (15) can also expressed as
\[
\xi_{\min} = E[d^2(n)] - \bar{P}^T R^{-1} \bar{P} \quad \text{...(16)}
\]
Example 3.1: A Modelling problem

\[ d(n) = 2x(n) + 3x(n-1) + v(n) \]

and \( v(n) \) represents a stationary white process \( \sigma_v^2 = 0.1 \),

\( x(n) \) is also a white process, with unit variance \( \sigma_x^2 = 1 \)

Note:

By this assumption, the actual system function of the plant is \( 2 + 3z^{-1} \)

Design an optimum filter (Weiner filter) to model the plant.

Question:

How to choose the number of taps (N)  

Once N is determined, the procedure to solve the problem is straightforward.

If \( N=2 \) is chosen

then \( R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \)

\( P = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \)

and \( \xi = 13.1 - 4w_0 - 6w_i + w_0^2 + w_i^2 \)

thus \( \begin{bmatrix} w_{0,0} \\ w_{0,1} \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \)

\( \xi_{\text{min}} = 0.1 \)

- The cost function (or performance function) is given by

$$\xi = E[e^2(n)] \quad \ldots (17)$$

$$\frac{\partial \xi}{\partial w_i} = E[2e(n) \frac{\partial e(n)}{\partial w_i}] \quad \ldots (18)$$

for $i = 0, 1, 2, \ldots, N - 1$

where $e(n) = d(n) - y(n)$

Since $d(n)$ is independent of $w_i$, we get

$$\frac{\partial e(n)}{\partial w_i} = -\frac{\partial y(n)}{\partial w_i}$$

$$= -x(n - i) \quad \ldots (19)$$

Then we obtain

$$\frac{\partial \xi}{\partial w_i} = -2E[e(n)x(n - i)] \quad \ldots (20)$$

for $i = 0, 1, 2, \ldots, N - 1$

- When the Wiener filter tap, weights are set to their optimal values, $\frac{\partial \xi}{\partial w_i} = 0$.

Hence, if $e_0(n)$ is the estimation error when $w_i$ are set to their optimal values, then equation (20) becomes

$$E[e_0(n)x(n - i)] = 0 \quad \text{for } i = 0, 1, \ldots, N - 1 \quad \ldots (21)$$

That is, the estimation error is uncorrelated with the filter tap inputs, $x(n-i)$. This is known as “the principles of orthogonality”.

- We can also show that the optimal filter output is also uncorrelated with the estimation error.

That is, $E[e_0(n)y_0(n)] = 0 \quad \ldots (22)$

This result indicates that the optimized Weiner filter output and the estimation error are “orthogonal”.
5. Normalized Performance Function

- If the optimal filter tap weights are expressed by \( w_{0,l}, l=0,1,2,\ldots,N-1 \). The estimation error is then given by
  \[
e_0(n) = d(n) - \sum_{l=0}^{N-1} w_{0,l} x(n-l) \quad \text{...(23)}
  \]
  and then
  \[
d(n) = e_0(n) + y_0(n) \quad \text{...(24)}
  \]
  \[
E[d^2(n)] = E[e_0^2(n)] + E[y_0^2(n)] + 2E[e_0(n)y_0(n)]
  \]

- We may note that
  \[
E[e_0^2(n)] = \bar{\xi}_{\text{min}} \quad \text{...(26)}
  \]
  and we obtain
  \[
\bar{\xi}_{\text{min}} = E[d^2(n)] - E[y_0^2(n)] \quad \text{...(27)}
  \]

- Define \( \rho \) as the normalized performance function
  \[
\rho = \frac{\bar{\xi}}{E[d^2(n)]} \quad \text{...(28)}
  \]

- \( \rho = 1 \) when \( y(n) = 0 \)

- \( \rho \) reaches its minimum value, \( \rho_{\text{min}} \), when the filter tap-weights are chosen to achieve the minimum mean-squared error.

  This gives
  \[
\rho_{\text{min}} = 1 - \frac{E[y_0(n)]}{E[d^2(n)]} \quad \text{...(29)}
  \]

  We have \( 0 < \rho_{\text{min}} < 1 \).
6. Wiener Filter - Complex-Valued Case

- In many practical applications, the random signals are complex-valued. For example, the baseband signal of QPSK & QAM in data transmission systems.

- In the Wiener filter for processing complex-valued random signals, the tap-weights are assumed to be complex variables.

- The estimation error, \( e(n) \), is also complex-valued. We may write
  \[
  \xi = E[|e(n)|^2] = E[e(n)e^*(n)] \quad …(30)
  \]

- The tap-weight \( w_i \) is expressed by
  \[
  w_i = w_{i,R} + jw_{i,I} \quad …(31)
  \]
  The gradient of a function with respect to a complex variable \( w = w_R + jw_I \) is defined as
  \[
  \nabla_w^c \equiv \frac{\partial}{\partial w_R} + j\frac{\partial}{\partial w_I} \quad …(32)
  \]

- The optimum tap-weights of the complex-valued Wiener filter will be obtained from the criterion:
  \[
  \nabla_{w_i} \xi = 0 \quad \text{for} \quad i = 0, 1, 2, ..., N - 1
  \]
  That is,
  \[
  \frac{\partial \xi}{\partial w_{i,R}} = 0 \quad \text{and} \quad \frac{\partial \xi}{\partial w_{i,I}} = 0 \quad …(33)
  \]

- Since \( \xi = E[e(n)e^*(n)] \), we have
  \[
  \nabla_{w_i} \xi = E[e(n)\nabla_{w_i} e^*(n) + e^*(n)\nabla_{w_i} e(n)] \quad …(34)
  \]
  Noting that
  \[
  e(n) = d(n) - \sum_{k=0}^{N-1} w_k x(n - k) \quad …(35)
  \]
  \[
  \nabla_{w_i} e(n) = -x(n - i)\nabla_{w_i} w_i \quad …(36)
  \]
  \[
  \nabla_{w_i} e^*(n) = -x^*(n - i)\nabla_{w_i} w^*_i \quad …(37)
  \]
  Applying the definition (32), we obtain
\[ \nabla_{w_i} w_i = \frac{\partial w_i}{\partial w_{i,R}} + j \frac{\partial w_i}{\partial w_{i,I}} = 1 + j(j) = 0 \]

and \[ \nabla_{w_i}^* w_i^* = \frac{\partial w_i^*}{\partial w_{i,R}} + j \frac{\partial w_i^*}{\partial w_{i,I}} = 1 + j(-j) = 2 \]

Thus, equation (34) becomes
\[ \nabla_{w_i}^* \xi = -2E[e(n)x^*(n-i)] \quad (38) \]

The optimum filter tap-weights are obtained when \( \nabla_{w_i}^* \xi = 0 \). This gives
\[ E[e_0(n)x^*(n-i)] = 0 \quad \text{for } i = 0, 1, 2, ..., N - 1 \quad (39) \]

where \( e_0(n) \) is the optimum estimation error.

- Equation (39) is the “principle of orthogonality” for the case of complex-valued signals in Wiener filter.

- The Wiener-Hopf equation can be derived as follows:

Define \( \bar{x}(n) \equiv [x(n) \quad x(n-1) \quad \cdots \quad x(n-N+1)]^T \quad (40) \)

and \[ \bar{w}(n) \equiv [w_0^* \quad w_i^* \quad \cdots \quad w_{N-1}^*]^T \quad (41) \]

We can also write \( \bar{x}(n) \equiv [x^*(n) \quad x^*(n-1) \quad \cdots \quad x^*(n-N+1)]^H \quad (42) \)

and \[ \bar{w}(n) \equiv [w_0 \quad w_i \quad \cdots \quad w_{N-1}]^H \quad (43) \]

where \( H \) denotes complex-conjugate transpose or Hermitian.

Noting that
\[ e(n) = d(n) - \bar{w}_0^H x(n) \quad (44) \]

and \[ E[e_0(n)\bar{x}(n)] = 0 \quad (45) \quad \text{from equation (39)} \]

We have
\[ E[\bar{x}(n)(d^*(n) - \bar{x}^H(n)\bar{w}_0)] = \bar{0} \quad (46) \]

and then
\[ R\bar{w}_0 = \bar{p} \quad (47) \]
where \( R = E[\bar{x}(n)\bar{x}^H(n)] \)

and \( \bar{p} = E[\bar{x}(n)d^*(n)] \)

- Equation (47) is the Wiener-Hopf equation for the case of complex-valued signals.

- The minimum performance function is then expressed as

\[
\xi_{\text{min}} = E[d^2(n)] - \bar{w}_0^H R \bar{w}_0 \quad \text{(48)}
\]

- Remarks:
  In the derivation of the above Wiener filter we have made assumption that it is causal and finite impulse response, for both real-valued and complex-valued signals.
7. Unconstrained Weiner Filters

- The block diagram of a Wiener filter is shown in Fig. 3.1

![Block Diagram](image)

Fig. 3.1

We assume that the filter \( W(z) \) may be non-causal and/or IIR.

- We consider only the case the signals and the system parameters are real-valued.
Moreover, we assume that the complex variable \( z \) remains on the unit circle, i.e. \( |z|=1 \), thus \( z^*=z^{-1} \), \( |z|=1 \).

7.1 Performance Function

- The performance function is defined as

\[
\xi = E[|e(n)|^2]
\]

Where \( e(n)=d(n)-y(n) \)

In terms of autocorrelation and cross-correlation functions, we have

\[
\xi = E[d^2(n)] + E[y^2(n)] - 2E[y(n)d(n)]
= \phi_{dd}(0) + \phi_{yy}(0) - 2\phi_{yd}(0) \quad \ldots(49)
\]

Using the inverse Z-transform relations, we have

\[
\xi = \phi_{dd}(0) + \frac{1}{2\pi j} \int \Phi_{yy}(z) \frac{dz}{z} - 2 \frac{1}{2\pi j} \int \Phi_{yd}(z) \frac{dz}{z} \quad \ldots(50)
\]

Since \( Y(z)=X(z)W(z) \), \( \Phi_{yd}(z)=W(z)\Phi_{yd}(z) \)

If \( z \) is selected to be on the unit circle in the \( z \)-plane, then

\[
\Phi_{yy}(z) = |W(z)|^2 \Phi_{xx}(z),
\]

\[
|W(z)|^2 = W(z)W^*(z) \quad \text{and} \quad W^*(z) = W(z^{-1})
\]
Using these relations in equation (50), we obtain

\[
\xi = \phi_{dd}(0) + \frac{1}{2\pi j} \oint c |W(z)|^2 \Phi_{xx}(z) \frac{dz}{z} - 2 \frac{1}{2\pi j} \oint c W(z) \Phi_{sd}(z) \frac{dz}{z}
\]

\[
= \phi_{dd}(0) + \frac{1}{2\pi j} \oint c [W^*(z)\Phi_{xx}(z) - 2\Phi_{sd}(z)]W(z) \frac{dz}{z} \quad \text{...(51)}
\]

Where the contour of integration, \(c\), is the unit circle.

- The performance function given in equation (51) covers IIR and FIR. This is a most general form of the Wiener filter performance function.

**Example 3.2**

From general expression of \(\xi\), eq.(51) to demonstrate the special case of real-valued FIR system.

Consider the case where the Wiener filter is an N-tap FIR filter with system function \(W(z) = \sum_{l=0}^{N-1} w_l z^{-l} \) ... (52)

Using eq.(52) in the expression of \(\xi\), given by eq.(51), then

\[
\xi = \phi_{dd}(0) + \frac{1}{2\pi j} \oint c \left( \sum_{l=0}^{N-1} w_l z^{-l} \right) \left( \sum_{m=0}^{N-1} w_m z^{-m} \right) \Phi_{xx}(z) \frac{dz}{z} - 2 \frac{1}{2\pi j} \oint c \left( \sum_{l=0}^{N-1} w_l z^{-l} \right) \Phi_{sd}(z) \frac{dz}{z}
\]

\[
= \phi_{dd}(0) + \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} w_l w_m \Phi_{xx}(m-l) - 2 \sum_{l=0}^{N-1} w_l \Phi_{sd}(-l)
\]

Using the inverse z-transform relation, this gives

\[
\xi = \phi_{dd}(0) + \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} w_l w_m \Phi_{xx}(m-l) - 2 \sum_{l=0}^{N-1} w_l \Phi_{sd}(-l)
\]

Now using the notations

\[
\phi_{dd}(0) = E[d^2(n)]
\]

\[
\phi_{sd}(-l) = p_l
\]

\[
\phi_{xx}(m-l) = \phi_{xx}(l-m) = r_{lm}
\]

We obtain the expression

\[
\xi = E[d^2(n)] - 2 \sum_{l=0}^{N-1} p_l w_l + \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} w_l w_m r_{lm}
\]

Note that

\[
\bar{p} = E[\tilde{x}(n)d(n)]
\]

Thus \(p_l = E[x(n-l)d(n)]\)
By definition, \( \phi_{xy}(k) = E[x(n)y^*(n-k)] \)
\[ \therefore \phi_{xu}(-l) = E[x(n)d(n+l)] = E[x(n-l)d(n)] \]

**Example 3.3**
Modelling with an IIR filter

The plant is modeled by an IIR filter with
\[ W(z) = \frac{1 - w_0 z^{-1}}{1 - w_1 z^{-1}}. \]
We assume that all involved signals and system parameters are real-valued.

\( x(n) \) : input sequence, white process with zero-mean, \( \sigma_x^2 = 1 \).
\( v(n) \) : additive noise.
\( x(n) \& v(n) \) are uncorrelated.

Thus, \( \Phi_{xx}(z) = 1 \) and \( \Phi_{vx}(z) = 0 \)

\( G(z) \) : system function of the unknown plant.

\[ \Phi_{xu}(z) = G(z^{-1})\Phi_{xx}(z) \]

Using equation(51) to obtain the performance function \( \xi \):
\[ \xi = \phi_{dd}(0) + \frac{w_1 - w_0}{w_1} \frac{1 - w_0 w_1}{1 - w_1^2} + \frac{w_0}{w_1} - 2[\frac{w_1 - w_0}{w_1} G(w_1) + \frac{w_0}{w_1} G(\infty)] \]

- We may find that there can be many local minima, and searching for the global minimum of \( \xi \) may not be a trivial task.
7.2 Optimum transfer function

- From the principle of orthogonality, we have
  \[ E[e_0(n)x(n-i)] = 0 \quad \text{for} \quad i = 0,1,2,\ldots,N-1 \quad \text{...(52)} \]
  where \( e_0(n) = d(n) - \sum_{l=\infty}^{\infty} w_{0l} x(n-l) \quad \text{...(53)} \)

Here we assume that all involved signals are real-valued.

- Combining eq.(52)&eq.(53), we obtain
  \[ \sum_{l=\infty}^{\infty} w_{0l} E[x(n-l)x(n-i)] = E[d(n)x(n-i)] \quad \text{...(54)} \]
  with \( E[x(n-l)x(n-i)] = \phi_{xx}(i-l) \)

  and \( E[d(n)x(n-i)] = \phi_{dx}(i) \)

and taking Z-transform on both sides of equation(54), we get
  \[ \Phi_{as}(z)W_0(z) = \Phi_{ds}(z) \quad \text{...(55)} \]
  This is referred to as the “Wiener-Hopf” equation of the unconstrained Wiener filtering.

- The optimum unconstrained Wiener filter is given by
  \[ W_0(z) = \frac{\Phi_{ds}(z)}{\Phi_{as}(z)} \quad \text{...(56)} \]
  and
  \[ W_0(e^{jw}) = \frac{\Phi_{ds}(e^{jw})}{\Phi_{as}(e^{jw})} \quad \text{...(57)} \]
  This is the frequency response of the Wiener filter

\( \Phi_{ds}(e^{jw}) \) : cross-power spectral density of \( d(n) \) and \( x(n) \)

\( \Phi_{as}(e^{jw}) \) : power spectral density of \( x(n) \)
8. Application of Wiener Filters  □ : Modelling

- Consider the modeling problem depicted in Fig. 3.8

□ (n), □ 0(n), and □ i(n) are assumed to be stationary, zero-mean and uncorrelated with one another.

- The input to Wiener filter is given by
  \[ x(n) = \mu(n) + \nu_i(n) \]  
  ...(58)

and the desired output is given by
  \[ d(n) = g_n \ast \mu(n) + \nu_0(n) \]  
  ...(59)

where \( g_n \) is the impulse response sample of the plant.

- The optimum unconstrained Wiener filter transfer function
  \[ W_0(z) = \frac{\Phi_{dx}(z)}{\Phi_{xx}(z)} \]  
  ...(60)

Note that
  \[ \phi_{xx}(k) = E[x(n)x(n-k)] \]
  \[ = E[(\mu(n) + \nu_i(n))(\mu(n-k) + \nu_i(n-k))] \]
  \[ = E[\mu(n)\mu(n-k)] + E[\mu(n)\nu_i(n-k)] + E[\nu_i(n)\mu(n-k)] + E[\nu_i(n)\nu_i(n-k)] \]
  \[ = \phi_{\mu\mu}(k) + \phi_{w\nu_i}(k) \]  
  ...(61)

Taking Z-transform on both sides of eq.(61), we get
  \[ \Phi_{xx}(z) = \Phi_{\mu\mu}(z) + \Phi_{w\nu_i}(z) \]  
  ...(62)
To calculate \( W_0(z) = \frac{\Phi_{xx}(z)}{\Phi_{xx}(z)} \), we must first find the expression for \( \Phi_{xx}(z) \).

We can show that

\[ \Phi_{xx}(z) = \Phi_{x\mu}(z) \]

where \( d'(n) \) is the plant output when the additive noise \( \nu_0(n) \) is excluded from that.

Moreover, we have

\[ \Phi_{d\mu}(z) = G(z)\Phi_{\mu\mu}(z) \]

Thus \( \Phi_{xx}(z) = G(z)\Phi_{\mu\mu}(z) \)

And we obtain

\[ W_0(z) = \frac{\Phi_{\mu\mu}(z)}{\Phi_{\mu\mu}(z) + \Phi_{\nu\nu}(z)} G(z) \quad \text{(63)} \]

We note that \( W_0(z) \) is equal to \( G(z) \) only when \( \Phi_{\nu\nu}(z) \) is equal to zero.

That is, when \( \nu_0(n) \) is zero for all values of \( n \).

The noise sequence \( \nu_0(n) \) may be thought of as introduced by a transducer that is used to get samples of the plant input.

Replacing \( z \) by \( e^{jw} \) in equation(63), we obtain

\[ W_0(e^{jw}) = \frac{\Phi_{\mu\mu}(e^{jw})}{\Phi_{\mu\mu}(e^{jw}) + \Phi_{\nu\nu}(e^{jw})} G(e^{jw}) \quad \text{(64)} \]

Define \( K(e^{jw}) \equiv \frac{\Phi_{\mu\mu}(e^{jw})}{\Phi_{\mu\mu}(e^{jw}) + \Phi_{\nu\nu}(e^{jw})} \quad \text{(65)} \)

We obtain

\[ W_0(e^{jw})=K(e^{jw})G(e^{jw}) \]
With some mathematic manipulation, we can find the minimum mean-square error, \( \xi_{\text{min}} \), expressed by

\[
\xi_{\text{min}} = \phi_{u_0u_0}(0) + \frac{1}{2\pi} \int_{-\pi}^{\pi} (1 - G(e^{j\nu}))G(e^{j\nu}) \, dw \quad \cdots (66)
\]

- The best performance that one can expect from the unconstrained Wiener filter is

\[
\xi_{\text{min}} = \phi_{u,u_0}(0)
\]

and this happens when \( \nu_\nu(n) = 0 \).

- The Wiener filter attempts to estimate that part of the target signal \( d(n) \) that is correlated with its own input \( x(n) \) and leaves the remaining part of \( d(n) \) (i.e. \( \nu_\nu(n) \)) unaffected. This is known as “the principles of correlation cancelling “.
9. Application of Wiener Filters — Inverse Modelling

- Fig. 3.9 depicts a channel equalization scenario.

![Diagram of channel equalization](image)

- When the additive noise at the channel output is absent, the equalizer has the following trivial solution:

$$W_0(z) = \frac{1}{H(z)} \quad \ldots(67)$$

This implies that $y(n) = s(n)$ and thus $e(n) = 0$ for all $n$.

Demonstration: Fig. 3.10

- When the channel noise, $\omega(n)$, is non-zero, the solution provided by equation(67) may not be optimal.

$$x(n) = h_n * s(n) + \omega(n) \quad \ldots(68)$$

and

$$d(n) = s(n) \quad \ldots(69)$$

where $h_n$ is the impulse response of the channel, $H(z)$.

From equation(68), we obtain

$$\Phi_{xx}(z) = \Phi_{xx}(z)|H(z)|^2 + \Phi_{\omega\omega}(z) \quad \ldots(70)$$

Also,

$$\Phi_{dx}(z) = \Phi_{sx}(z) = H(z^{-1})\Phi_{sx}(z) \quad \ldots(71)$$

With $|z| = 1$, we may also write

$$\Phi_{dx}(z) = H^*(z)\Phi_{sx}(z) \quad \ldots(72)$$
and then
\[ W_0(z) = \frac{H^*(z)\Phi_{ss}(z)}{\Phi_{ss}(z)|H(z)|^2 + \Phi_{uu}(z)} \quad \text{(73)} \]

This is the general solution to the equalization problem when there is no constraint on the equalizer length and, also, it may be let to be non-causal.

- Equation (73) can be rewritten as
  \[ W_0(z) = \frac{1}{1 + \frac{\Phi_{uu}(z)}{\Phi_{ss}(z)|H(z)|^2}} \cdot \frac{1}{H(z)} \quad \text{(74)} \]

- Let \( z = e^{jw} \) and define the parameter
  \[ \rho(e^{jw}) \equiv \frac{\Phi_{ss}(e^{jw})|H(e^{jw})|^2}{\Phi_{uu}(e^{jw})} \quad \text{(75)} \]
  where \( \Phi_{ss}(e^{jw})|H(e^{jw})|^2 \) and \( \Phi_{uu}(e^{jw}) \) are the signal power spectral density and the noise power spectral density, respectively, at the channel output.
  
  We obtain
  \[ W_0(e^{jw}) = \frac{\rho(e^{jw})^{-1}}{1 + \rho(e^{jw})^{-1}} \cdot \frac{1}{H(e^{jw})} \quad \text{(76)} \]
  
  Note that \( \rho(e^{jw}) \) is a non-negative quantity, since it is the signal-to-noise power spectral density ratio at the equalizer input.
  
  Also, \( 0 \leq \frac{\rho(e^{jw})^{-1}}{1 + \rho(e^{jw})^{-1}} \leq 1 \quad \text{(77)} \)

- Cancellation of ISI and noise enhancement

Consider the optimized equalizer with frequency response given by
\[ W_0(e^{jw}) = \frac{\rho(e^{jw})^{-1}}{1 + \rho(e^{jw})^{-1}} \cdot \frac{1}{H(e^{jw})} \]

In the frequency regions where the noise is almost absent, the value of \( \rho(e^{jw}) \) is very large and hence
\[ W_0(e^{jw}) \approx \frac{1}{H(e^{jw})} \]
The ISI will be eliminated without any significant enhancement of noise. On the other hand, in the frequency regions where the noise level is high, the value of $\mathcal{F}(e^{jw})$ is not large and hence the equalizer does not approximate the channel inverse well. This is of course, to prevent noise enhancement.
10. Noise Cancellation

- Fig. 3.12 shows a typical noise canceller block diagram

\[ s(n) \]: signal source

\[ ?(n) \]: noise source

\( s(n) \) & \( ?(n) \) are uncorrelated

\( H(z) \) & \( W(z) \) are two system functions used to mixed the two input signals to form \( d(n) \) & \( x(n) \).

\( d(n) \): primary input

\( x(n) \): Reference input

\[
\begin{align*}
  x(n) &= ?(n) + h_n * s(n) \quad \text{...(78)} \\
  d(n) &= s(n) + g_n * ?(n) \quad \text{...(79)} 
\end{align*}
\]

- Since \( s(n) \) & \( ?(n) \) are uncorrelated with each other, we obtain

\[
\Phi_{ds}(z) = \Phi_{ds}^s(z) + \Phi_{ds}^\nu(z) \quad \text{...(81)}
\]

where \( \Phi_{ds}(z) \) when \( ?(n)=0 \) for all \( n \),

and \( \Phi_{ds}^\nu(z) \) is \( \Phi_{ds}(z) \) when \( s(n)=0 \) for all \( n \).

This is because \( s(n) \) and \( ?(n) \) are uncorrelated with each other, their contribution in \( \Phi_{ds}(z) \) can be considered separately.

Thus, we obtain

\[
\Phi_{ds}^s(z) = H^*(z)\Phi_{ss}(z) \quad \text{...(82)}
\]
and $\Phi_{\nu\nu}(z) = G(z)\Phi_{\nu\nu}(z)$ \hspace{1cm} (83)

Recall that $|z| = 1$. Finally, we get

$$\Phi_{\nu\nu}(z) = H'(z)\Phi_{ss}(z) + G(z)\Phi_{\nu\nu}(z) \hspace{1cm} (84)$$

and $W_{\nu}(z) = \frac{H'(z)\Phi_{ss}(z) + G(z)\Phi_{\nu\nu}(z)}{\Phi_{\nu\nu}(z) + \Phi_{ss}(z)|H(z)|^2} \hspace{1cm} (85)$

- Equation (85) for noise cancellation may be thought of as a generalization of the results we obtained for the modeling and inverse modeling scenarios, given by equation (63) and (73).

- To minimize the mean-square value of the output error, we must strike a balance between noise cancellation and signal cancellation at the output of the noise canceller.

Cancellation of the noise $?(n)$ occurs when the Wiener filter $W(z)$ is chosen to be close to $G(z)$, and cancellation of the signal $s(n)$ occurs when the Wiener filter $W(z)$ is close to the inverse of $H(z)$. In the sense, we may note that the noise canceller treats $s(n)$ and $?(n)$ without making any distinction between them.

- Define

$$\rho_{\text{pri}}(e^{jw}) : \text{signal-to-noise PSD at primary input}$$

$$\rho_{\text{ref}}(e^{jw}) : \text{signal-to-noise PSD at reference input}$$

$$\rho_{\text{out}}(e^{jw}) : \text{signal-to-noise PSD at output}$$

$$\rho_{\text{pri}}(e^{jw}) = \frac{\Phi_{ss}(e^{jw})}{|G(e^{jw})|^2\Phi_{\nu\nu}(e^{jw})} \hspace{1cm} (86)$$

$$\rho_{\text{ref}}(e^{jw}) = \frac{|H(e^{jw})|^2\Phi_{ss}(e^{jw})}{\Phi_{\nu\nu}(e^{jw})} \hspace{1cm} (87)$$

To calculate $\rho_{\text{out}}(e^{jw})$, we note that $s(n)$ reaches the canceller output
through two routes: one direct and one through the cascade of $H(z)$ and $W(z)$:

$$\Phi_{ee}(e^{jw}) = |1 - H(e^{jw})W(e^{jw})|^2 \Phi_{ss}(e^{jw}) \quad \text{(88)}$$

Similarly, $\Phi(n)$ reaches the output through the routes $G(z)$ and $W(z)$:

$$\Phi_{ou}(e^{jw}) = |G(e^{jw}) - W(e^{jw})|^2 \Phi_{wu}(e^{jw}) \quad \text{(89)}$$

Replacing $W(e^{jw})$ by $W_0(e^{jw})$, we obtain

$$\Phi_{oo}(e^{jw}) = \frac{|1 - G(e^{jw})H(e^{jw})|^2 \Phi_{ww}(e^{jw})}{\Phi_{wu}(e^{jw}) + |H(e^{jw})|^2 \Phi_{ss}(e^{jw})}$$

and

$$\Phi_{oe}(e^{jw}) = \frac{|H(e^{jw})|^2 |1 - G(e^{jw})H(e^{jw})|^2 \Phi_{oe}(e^{jw})}{\Phi_{wu}(e^{jw}) + |H(e^{jw})|^2 \Phi_{ss}(e^{jw})} \Phi_{wu}(e^{jw})$$

Hence, $\rho_{out}(e^{jw})$ can now be obtained as

$$\rho_{out}(e^{jw}) = \frac{\Phi_{oe}(e^{jw})}{\Phi_{ee}(e^{jw})} = \frac{\Phi_{wu}(e^{jw})}{|H(e^{jw})|^2 \Phi_{ss}(e^{jw})} = \frac{1}{\rho_{ref}(e^{jw})} \quad \text{(90)}$$

This is known as “power inversion”. (Widrow et al, 1975)

** This result shows that the noise canceller works better when $\rho_{ref}(e^{jw})$ is low.
Example 3.6

Two Omni-directional Antenna in a Receiver

- $s(n)$ arrives the receiver (antennas) in the direction perpendicular to the line connecting A & B. $\tau(n)$ arrives at an angle $\theta_0$ with respect to the direction of $s(n)$.

- $a(n)$ and $\beta(n)$ are narrowband baseband signals. Thus $s(n)$ & $\tau(n)$ can be treated as narrowband signals concentrated around $w = w_0$.

- since $s(n)$ & $\tau(n)$ may be treated as single tones, and thus a filter with two degrees of freedom is sufficient for optimal filtering.

- Wiener filter coefficients can be determined as follows. $s(n)$ arrives at the same time at both omni-antennas. $\tau(n)$ arrives at B first and arrives at A with a delay
  \[ \delta_0 = \frac{l \sin \theta_0}{c} \]
  the delay is normalized by the time step $T$ which corresponds to one increment of time index $n$. This gives
  \[ \delta_0 = \frac{l \sin \theta_0}{cT} \]
Thus
\[ d(n) = \alpha(n) \cos nw_0 + \beta(n) \cos [(n - \delta)w_0] \]
\[ s(n) = \alpha(n) \cos nw_0 + \beta(n) \cos nw_0 \]
\[ \tilde{x}(n) = \alpha(n) \sin nw_0 + \beta(n) \sin nw_0 \]
\[ R = \begin{bmatrix}
E[x^2(n)] & E[x(n)\tilde{x}(n)] \\
E[x(n)\tilde{x}(n)] & E[\tilde{x}^2(n)]
\end{bmatrix} \]

With some mathematical manipulations, we obtain
\[ R = \sigma_a^2 + \sigma_\beta^2 \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix} \]

where \( \sigma_a^2 \) and \( \sigma_\beta^2 \) are variance of \( a(n) \) and \( \beta(n) \), respectively.

Also, \( \bar{P} = \begin{bmatrix}
E[d(n)x(n)] \\
E[d(n)\tilde{x}(n)]
\end{bmatrix} = \frac{1}{2} \begin{bmatrix}
\sigma_a^2 + \sigma_\beta^2 \cos \delta w_0 \\
\sigma_\beta^2 \sin \delta w_0
\end{bmatrix} \]

The Wiener-Hopf equation \( RT_0 = \bar{P} \)

gives \( \bar{W}_0 = \begin{bmatrix}
(\sigma_a^2 + \sigma_\beta^2 \cos \delta w_0) \\
\sigma_\beta^2 \sin \delta w_0
\end{bmatrix} \times \frac{1}{\sigma_a^2 + \sigma_\beta^2} \]

The signal-to-noise ratio at the output is equal to \( \sigma_\beta^2 / \sigma_a^2 \)

while the signal-to-noise ratio at the reference input is equal to \( \sigma_a^2 / \sigma_\beta^2 \).