TABLE OF CONTENTS

Syllabus

Course Catalog

3 Credit hours (3 h lectures). DSP fundamentals, such as the z-transform, DFT, FFT, IIR, and FIR filters, and so on. Optimum filtering. Various physical layer issues in communications are addressed, including channel estimation and adaptive equalization.

Textbook

Several books and journal articles.

References

Books

1. Alan V. Oppenheim, Ronald W. Schafer and John R. Buck, *Discrete-Time Signal Processing*, Prentice-Hall, 1999
2. Giorgio M. Vitetta, Desmond P. Taylor, Giulio Colavolpe, Fabrizio Pancaldi and Philippa A. Martin, *Wireless Communications Algorithmic Techniques,* Wiley, 2013
3. Frank A. Dietrich, *Robust Signal Processing for Wireless Communications,* Springer, 2008

Instructor

Instructor: Dr. Mohammad M. Banat

Email Address: banat@just.edu.jo

Prerequisites

Background in linear algebra, signal analysis, random processes, communication systems and digital signal processing.

Topics Covered

Week	Topics
1-6	DSP Background
7-9	Optimum Filtering
10-12	Channel Estimation
13-16	Channel Equalization

Evaluation

Assessment Tool	**Expected Due Date**	**Weight**
Computer Assignments		10%
Mid-Term Exam		20%
Term Project Report		10%
Presentations		10%
Final Exam	According to the university final examination schedule	50%

0-Course Catalog

I. DSP BACKGROUND

I.1. What Is Digital Signal Processing?

A signal is a formal description of a phenomenon evolving over time or space. Signal processing means any operation which modifies, analyzes or otherwise manipulates the information contained in a signal. Digital signal processing is a flavor of signal processing in which everything including time is described in terms of integer numbers.

I.2. Discrete-Time Signals (Sequences)

Discrete-time signals are represented mathematically as sequences of numbers. A discrete-time signal is denoted as $x[n]$, where the variable n denotes discrete time. In a practical setting, such sequences can arise from periodic sampling of an analog signal $x_a(t)$. In this case,

$$x[n] = x_a(nT) \tag{I.1}$$

The quantity T is called the sampling period, and its reciprocal $f_s = 1/T$ is the called sampling frequency. According to the sampling theorem, the original analog signal can be reconstructed as accurately as desired from a corresponding sequence of samples if the samples are taken frequently enough. Sampling rate should be equal to or higher than the Nyquist rate. For baseband signals, the Nyquist rate is equal to twice the highest frequency component of the signal.

Although sequences do not always arise from sampling analog waveforms, it is convenient to refer to $x[n]$ as the nth sample of the sequence. Discrete-time signals (i.e., sequences) are often depicted graphically as in Figure I.1 below. Although the abscissa is drawn as a continuous line, it is important to recognize that $x[n]$ is defined only for integer values of n. It is not correct to think of $x[n]$ as being zero when n is not an integer; $x[n]$ is simply undefined for non-integer values of n.

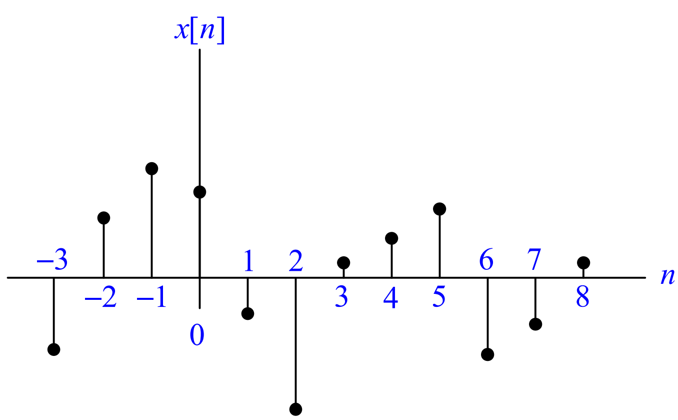


Figure I.1: Discrete-Time Signal

I.3. Basic Sequences and Sequence Operations

In the analysis of discrete-time signal-processing systems, sequences are manipulated in several basic ways. The product and sum of two sequences are defined as the sample-by-sample product and sum, respectively. Multiplication of a sequence by a number a is defined as multiplication of the value of each sample by a. A sequence $y[n]$ is said to be a delayed or shifted version of a sequence $x[n]$ if

$$y[n] = x[n-k] \tag{I.2}$$

where k is an integer.

The unit sample sequence is defined as the sequence

$$\delta[n] = \begin{cases} 1, & n = 0 \\ 0, & \text{otherwise} \end{cases} \tag{I.3}$$

For convenience, the unit sample sequence is often referred to as a discrete-time impulse or simply as an impulse. It is important to note that a discrete-time impulse does not suffer from the mathematical complications of the continuous-time impulse; its definition is simple and precise. Compare (I.3) to following definition of $\delta(t)$

$$\delta(t) = \begin{cases} \text{undefined}, & t = 0 \\ 0, & \text{otherwise} \end{cases} \tag{I.4}$$

An arbitrary sequence can be represented as a sum of scaled, delayed impulses. For example, the signal in Figure I.2 can be expressed as follows:

$$x[n] = a_{-3}\delta[n+3] + a_1\delta[n-1] + a_2\delta[n-2] + a_7\delta[n-7] \tag{I.5}$$

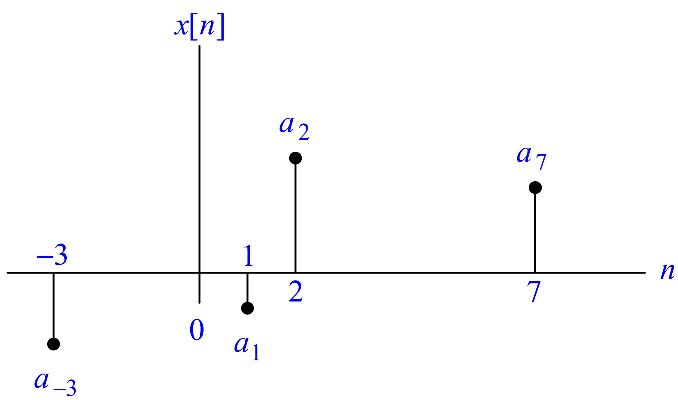


Figure I.2: Discrete-Time Signal

Any sequence can be expressed as

$$x[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k] \tag{I.6}$$

The unit step sequence is given by

$$u[n] = \begin{cases} 1, & n \ge 0 \\ 0, & n < 0 \end{cases} \tag{I.7}$$

The unit step is related to the unit impulse by

$$\begin{aligned} u[n] &= \sum_{k=0}^{\infty} \delta[n-k] \\ &= \sum_{k=-\infty}^{n} \delta[k] \end{aligned} \tag{I.8}$$

Conversely, the unit impulse sequence can be expressed as the first backward difference of the unit step sequence, i.e.,

$$\delta[n] = u[n] - u[n-1] \tag{I.9}$$

The general form of an exponential sequence is

$$x[n] = A\alpha^n \tag{I.10}$$

where A and α are constants that are generally complex. If A and α are real numbers, then the sequence is real-valued.

A sinusoidal sequence has the general form

$$x[n] = A\cos(\omega_0 n + \phi) \tag{I.11}$$

where A ($|A| = \text{amplitude}$), ω_0 (radian frequency), and ϕ (phase) are all real.

A complex exponential sequence has the form

$$x[n] = A\alpha^n \tag{I.12}$$

where A and α are generally complex. Let

$$\begin{aligned} A &= |A| e^{j\phi} \\ \alpha &= |\alpha| e^{j\omega_0} \end{aligned} \tag{I.13}$$

Then,

$$\begin{aligned} x[n] &= |A||\alpha|^n e^{j(\omega_0 n + \phi)} \\ &= |A||\alpha|^n \left[\cos(\omega_0 n + \phi) + j\sin(\omega_0 n + \phi)\right] \end{aligned} \tag{I.14}$$

The fact that n is always an integer leads to some important differences between the properties of discrete-time and continuous-time complex exponential sequences and sinusoidal sequences. A major difference between continuous-time and discrete-time complex sinusoids is seen when we consider a frequency $\omega_0 + 2\pi r$. Let r be an integer, and consider the sequence

$$x[n] = Ae^{j(\omega_0 + 2\pi r)n} = Ae^{j\omega_0 n} \tag{I.15}$$

This shows that complex exponential sequences with frequencies $\omega_0 + 2\pi r$ are indistinguishable from one another. An identical statement holds for sinusoidal sequences. We conclude that, when discussing complex exponential signals or real sinusoidal signals, we need only consider frequencies in an interval of length 2π, such as $-\pi \le \omega_0 < \pi$ or $0 \le \omega_0 < 2\pi$.

Another important difference between continuous-time and discrete-time complex exponentials and sinusoids concerns their periodicity. In the continuous-time case, a sinusoidal signal and a complex exponential signal are both periodic, with the period equal to 2π divided by the (angular) frequency. In the discrete-time case, a periodic sequence is a sequence for which

$$x[n+N] = x[n] \tag{I.16}$$

where the period is N, which is necessarily an integer. If this condition for periodicity is tested for the discrete-time sinusoid, then

$$A\cos(\omega_0 n + \omega_0 N + \phi) = A\cos(\omega_0 n + \phi) \tag{I.17}$$

which requires that

$$\omega_0 N = 2\pi k \tag{I.18}$$

where k is an integer. Consequently, complex exponential and sinusoidal sequences are not necessarily periodic in n with period $2\pi/\omega_0$ and, depending on the value of ω_0, may not be periodic at all.

Actually, if the sequence is periodic, then the period is given by the following expression, where k is the smallest integer that makes the calculated value of N an integer:

$$N = \frac{2\pi}{\omega_0} k \tag{I.19}$$

Example I-1

Consider the sequence $x_1[n] = \cos(\pi n/4)$. This sequence has a period $N = 8$. To show this, note that $x_1[n+8] = \cos(\pi(n+8)/4) = \cos(\pi n/4 + 2\pi) = \cos(\pi n/4) = x_1[n]$. Satisfying the definition of a discrete-time periodic sequence. Contrary to our intuition from continuous-time sinusoids, increasing the frequency of a discrete-time sinusoid does not necessarily decrease the period of the sequence. Consider the discrete-time sinusoid $x_2[n] = \cos(3\pi n/8)$, which has a higher frequency than the sequence $x_1[n]$. However, $x_2[n]$ is not periodic with period 8, since $x_2[n+8] \ne x_2[8]$. Using an argument analogous to the one for $x_1[n]$, we can show that $x_2[n]$ has a period $N = 16$. Thus, increasing the frequency from $\omega_0 = 2\pi/8$ to $\omega_0 = 3\pi/8$ also increases the period of the sequence. This occurs because discrete-time sequence are defined only for integer indices n.

The integer restriction on n may cause some sinusoidal sequence not to be periodic at all. For example, there is no integer N such that the sequence $x_3[n]=\cos(n)$ satisfies the condition $x_3[n+N]=x_3[n]$ for all n.

These and other properties of discrete-time sinusoids that run differently to their continuous-time counterparts are caused by the limitation of the time index n to integers for discrete-time signals and systems.

When we combine the condition of (I.18) with our previous observation that ω_0 and $\omega_0+2\pi r$ are indistinguishable frequencies, it becomes clear that there are N distinguishable frequencies for which the corresponding sequences are periodic with period N. To see this, let's choose $r=1$ and check the periodicity of sequences with angular frequencies $\omega_0+l\frac{2\pi}{N}$, for $l=0,\cdots,N-1$. Note that the number of choices for the angular frequency is N. According to (I.18), the period will be equal to N if

$$\begin{aligned}\left(\omega_0+l\frac{2\pi}{N}\right)N=2\pi k &\Rightarrow\\ \omega_0 N+2\pi l=2\pi k &\Rightarrow\\ \omega_0 N=2\pi(k-l)&\end{aligned} \tag{I.20}$$

Since $k-l$ is an integer, then all sequences with angular frequencies $\omega_0+l\frac{2\pi}{N}$ are periodic with period N.

Example I-2

Let

$$x[n]=\cos\left(\frac{3\pi}{5}n\right)$$

Obviously,

$$\omega_0=\frac{3\pi}{5}$$

To satisfy (I.19) for the smallest integer k, we should have

$$\begin{aligned}N&=\frac{2\pi}{3\pi/5}k\\&=\frac{10}{3}k\end{aligned}$$

Note that the smallest possible integer k is 3. Therefore,

$$N=10$$

Sequences with any of the following 10 frequencies have the same period $N = 10$:

$$\omega_0 = \frac{3\pi}{5}$$
$$\omega_1 = \frac{3\pi}{5} + \frac{2\pi}{10}$$
$$\vdots$$
$$\omega_9 = \frac{3\pi}{5} + 9 \times \frac{2\pi}{10}$$

Exercise I-1

Show that the period in Example I-2 is equal to 10.

Show that the 10 sequences in Example I-2 have frequencies $0, 0.2\pi, \cdots, 1.8\pi$ and that they all have the same period $N = 10$.

Generally, the set of signals with frequencies $\omega_k = 2\pi k/N$, $k = 0, 1, \cdots, N-1$ all have the same period N. These properties of complex exponential and sinusoidal sequences are basic to both the theory and the design of computational algorithms for discrete-time Fourier analysis.

The interpretation of high and low frequencies is somewhat different for continuous-time and discrete-time sinusoidal and complex exponential signals. For a continuous-time sinusoidal signal $x(t) = A\cos(\Omega_0 t + \phi)$, as Ω_0 increases, $x(t)$ oscillates more and more rapidly. For the discrete-time sinusoidal signal $x[n] = A\cos(\omega_0 n + \phi)$, as ω_0 increases from $\omega_0 = 0$ toward $\omega_0 = \pi$, $x[n]$ oscillates more and more rapidly. However, as ω_0 increases from $\omega_0 = \pi$ to $\omega_0 = 2\pi$, the oscillations become slower. Values of ω_0 in the vicinity of $\omega_0 = 2\pi k$ for any integer value of k are typically referred to as low frequencies (relatively slow oscillations), while values of ω_0 in the vicinity of $\omega_0 = \pi + 2\pi k$ for any integer value of k are typically referred to as high frequencies (relatively rapid oscillations).

I.4. Discrete-Time Systems

A discrete-time system is a transformation or an operator that maps an input sequence $x[n]$ into an output sequence $y[n]$, i.e.,

$$y[n] = T\left[x[n]\right] \tag{I.21}$$

Example I-3

$$
\begin{aligned}
y[n] &= \frac{1}{M_1+M_2+1}\sum_{k=-M_1}^{k=M_2} x[n-k] \\
&= \frac{1}{M_1+M_2+1}\left(x[n+M_1]+x[n+M_1-1]+\cdots x[n]+x[n-1]+\cdots+x[n-M_2]\right)
\end{aligned}
$$

I.4.A. MEMORYLESS SYSTEMS

A system is referred to as memoryless if the output $y[n]$ at every value of n depends only on the input $x[n]$ at the same value of n.

I.4.B. LINEAR SYSTEMS

The class of linear systems is defined by the principle of superposition. If $y_1[n]$ and $y_2[n]$ are the responses of a system when $x_1[n]$ and $x_2[n]$ are the respective inputs, then the system is linear if and only if

$$
\begin{aligned}
T\left[x_1[n]+x_2[n]\right] &= T\left[x_1[n]\right]+T\left[x_2[n]\right] \\
&= y_1[n]+y_2[n]
\end{aligned}
\tag{I.22}
$$

and

$$
\begin{aligned}
T\left[ax[n]\right] &= aT\left[x[n]\right] \\
&= ay[n]
\end{aligned}
\tag{I.23}
$$

where a is an arbitrary constant. The first property is called the additivity property, and the second is called the homogeneity or scaling property. These two properties can be combined into the principle of superposition, stated as

$$T\left[a_1x_1[n]+a_2x_2[n]\right] = a_1T\left[x_1[n]\right]+a_2T\left[x_2[n]\right] \tag{I.24}$$

This equation can be generalized to the superposition of more than two inputs. Specifically, if

$$x[n] = \sum_{k=1}^{K} a_k x_k[n] \tag{I.25}$$

then the output of a linear system will be

$$y[n] = \sum_{k=1}^{K} a_k y_k[n] \tag{I.26}$$

where $y_k[n]$ is the system response to the input $x_k[n]$.

Example I-4

The system defined by the input-output equation

$$y[n] = \sum_{k=-\infty}^{n} x[k]$$

is called the accumulator system; since the output at time n is just the sum of the present and all previous input samples.

The accumulator system is a linear system. In order to prove this, we must show that it satisfies the superposition principle for all inputs, not just any specific set of inputs. We begin by defining two arbitrary inputs $x_1[n]$ and $x_2[n]$ and their corresponding outputs

$$y_1[n] = \sum_{k=-\infty}^{n} x_1[k]$$

$$y_2[n] = \sum_{k=-\infty}^{n} x_2[k]$$

When the input is $x_3[n] = a_1 x_1[n] + a_2 x_2[n]$, the superposition principle requires the output $y_3[n] = a_1 y_1[n] + a_2 y_2[n]$ for all possible choices of a_1 and a_2. We can show this by starting from

$$\begin{aligned} y_3[n] &= \sum_{k=-\infty}^{n} x_3[k] \\ &= \sum_{k=-\infty}^{n} \left(a_1 x_1[k] + a_2 x_2[k]\right) \\ &= a_1 \sum_{k=-\infty}^{n} x_1[k] + a_2 \sum_{k=-\infty}^{n} x_2[k] \\ &= a_1 y_1[n] + a_2 y_2[n] \end{aligned}$$

Thus, the accumulator system satisfies the superposition principle for all inputs, and is therefore a linear system.

I.4.C. Time-Invariant Systems

A time-invariant system (often referred to equivalently as a shift-invariant system) is a system for which a time shift or delay of the input sequence causes a corresponding shift in the output sequence. Specifically, suppose that a system transforms the input sequence with values $x[n]$ into the output sequence with values $y[n]$. Then the system is said to be time invariant if, for all n_0, the input sequence with values $x_1[n] = x[n-n_0]$ produces the output sequence with values $y_1[n] = y[n-n_0]$.

Example I-5

Consider the accumulator system in Example I-4. We define $x_1[n] = x[n-n_0]$. To show time invariavce, we solve for both $y[n-n_0]$ and $y_1[n]$ and compare them to see whether they are equal. First,

$$y[n-n_0] = \sum_{k=-\infty}^{n-n_0} x[k]$$

Next, we find

$$y_1[n] = \sum_{k=-\infty}^{n} x_1[k]$$
$$= \sum_{k=-\infty}^{n} x[k-n_0]$$

Substituting the change of variables $k_1 = k - n_0$ into the summation gives

$$y_1[n] = \sum_{k_1=-\infty}^{n-n_0} x[k_1]$$
$$= y[n-n_0]$$

Thus, the accumulator is a time-invariant system.

Example I-6

The system defined by the relation

$$y[n] = x[Mn], \quad -\infty < n < \infty$$

with M a positive integer, is called a compressor (or downsampler). Specifically, it discards $M-1$ samples out of M; i.e., it creates the output sequence by selecting every M th sample. This system is not time invariant. We can show that it is not by considering the response $y_1[n]$ to the input $x_1[n] = x[n-n_0]$. In order for the system to be time invariant, the output of the system when the input is $x_1[n]$ must be equal to $y[n-n_0]$. The output $y_1[n]$ that results from the input $x_1[n]$ can be directly computed from the definition of $y[n]$ above to be

$$y_1[n] = x_1[Mn]$$
$$= x[Mn-n_0]$$

Delaying the output $y[n]$ by n_0 samples yields

$$y[n-n_0] = x[M(n-n_0)]$$

Comparing these two outputs, we see that $y[n-n_0]$ is not equal to $y_1[n]$ for all M and n_0, and therefore, the system is not time invariant.

It is also possible to prove that a system is not time invariant by finding a single counterexample that violates the time-invariance property. For instance, a counterexample for the compressor is the case when $M=2$, $x[n]=\delta[n]$, and $x_1[n]=\delta[n-1]$. For these choices, $y[n]=\delta[n]$, but $y_1[n]=0$; thus, it is clear that $y_1[n]\neq y[n-1]$ for this system.

Exercise I-2

Consider that the upsampler system

$$y[n]=\begin{cases} x\left[\dfrac{n}{M}\right], & \dfrac{n}{M}\in\mathbb{N} \\ 0, & \text{otherwise} \end{cases}, \quad -\infty<n<\infty$$

Show that this system is not time-invariant.

I.4.D. Causal Systems

A system is causal if, for every choice of n_0, the output sequence value at the index $n=n_0$ depends only on the input sequence values for $n\le n_0$. This implies that if $x_1[n]=x_2[n]$ for all $n\le n_0$, then $y_1[n]=y_2[n]$ for $n\le n_0$. That is, the system is non-anticipative.

Example I-7

Consider the forward difference system defined by the relationship

$$y[n]=x[n+1]-x[n]$$

This system is not causal, since the current value of the output depends on a future value of the input. The violation of causality can be demonstrated by considering the two inputs $x_1[n]=\delta[n-1]$ and $x_2[n]=0$ and their corresponding outputs $y_1[n]=\delta[n]-\delta[n-1]$ and $y_2[n]=0$. Note that $x_1[n]=x_2[n]$ for $n\le 0$, so the definition of causality requires that $y_1[n]=y_2[n]$ for $n\le 0$, which is clearly not the case for $n=0$. Thus, by this counterexample, we have shown that the system is not causal.

Example I-8

The backward difference system, defined as

$$y[n]=x[n]-x[n-1]$$

has an output that depends only on the present and past values of the input. Because there is no way for the output at a specific time $y[n_0]$ to incorporate values of the input for $n>n_0$, the system is causal.

I.4.E. Stable Systems

A system is stable in the bounded-input, bounded-output (BIBO) sense if and only if every bounded input sequence produces a bounded output sequence. The input $x[n]$ is bounded if there exists a fixed positive finite value B_x such that

$$|x[n]| \le B_x, \forall n \tag{I.27}$$

Stability in the BIBO sense requires that, for every bounded input, there exist a fixed positive finite value B_y such that

$$|y[n]| \le B_y, \forall n \tag{I.28}$$

It is important to emphasize that the properties we have defined in this section are properties of systems, not of the inputs to a system. That is, we may be able to find inputs for which the properties hold, but the existence of the property for some inputs does not mean that the system has the property. For the system to have the property, it must hold for all inputs.

For example, an unstable system may have some bounded inputs for which the output is bounded, but for the system to have the property of stability, it must be true that for all bounded inputs, the output is bounded. If we can find just one input for which the system property does not hold, then we have shown that the system does not have that property.

I.5. Linear Time-Invariant (LTI) Systems

A particularly important class of systems consists of those systems that are both linear and time invariant. These two properties in combination lead to especially convenient representations for such systems. Most importantly, this class of systems has significant signal-processing applications.

If the linearity property is combined with the representation of a general sequence as a linear combination of delayed impulses as in (I.6), it follows that a linear system can be completely characterized by its impulse response.

Specifically, let $h_k[n]$ be the response of the system to $\delta[n-k]$. Then,

$$y[n] = T\left[\sum_{k=-\infty}^{\infty} x[k]\delta[n-k]\right] \tag{I.29}$$

From the principle of superposition,

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} x[k]T\left[\delta[n-k]\right] \\ &= \sum_{k=-\infty}^{\infty} x[k]h_k[n] \end{aligned} \tag{I.30}$$

Therefore, the system response to any input can be expressed in terms of the responses of the system to the sequences $\delta[n-k]$. If only linearity is imposed, $h_k[n]$ will depend on both n and k, in which case the computational usefulness of (I.30) is limited. We obtain a more useful result if we impose the additional constraint of time invariance.

The property of time invariance implies that if $h[n]$ is the response to $\delta[n]$, then the response to $\delta[n-k]$ is $h[n-k]$. With this additional constraint, (I.30) becomes

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} x[k]h[n-k] \\ &= \sum_{k=-\infty}^{\infty} h[k]x[n-k] \end{aligned} \tag{I.31}$$

Equation (I.31) is commonly called the convolution sum. If $y[n]$ is a sequence whose values are related to the values of two sequences $h[n]$ and $x[n]$ as in (I.31), we say that $y[n]$ is the convolution of $x[n]$ with $h[n]$ and represent this by the notation

$$y[n] = x[n] * h[n] \tag{I.32}$$

Equation (I.31) expresses each sample of the output sequence in terms all of the samples of the input and impulse response sequences.

Although the convolution-sum expression is analogous to the convolution integral of continuous-time linear system theory, the convolution sum should not be thought of as an approximation to the convolution integral. The convolution integral plays mainly a theoretical role in continuous-time linear system theory; we will see that the convolution sum, in addition to its theoretical importance, often serves as an explicit realization of a discrete-time linear system.

I.6. Properties of LTI Systems

Since all linear time-invariant systems are described by the convolution sum, the properties of this class of systems are defined by the properties of discrete-time convolution.

Convolution Operation is Commutative

$$x[n] * h[n] = h[n] * x[n] \tag{I.33}$$

Convolution Operation Distributes over Addition

$$x[n] * \left[h_1[n] + h_2[n]\right] = x[n] * h_1[n] + x[n] * h_2[n] \tag{I.34}$$

Cascade Connection

In a cascade connection of systems, the output of the first system is the input to the second, the output of the second is the input to the third, etc. The output of the last system is the overall output. Two linear time-invariant systems in cascade correspond to a linear time-invariant system with an impulse response that is the convolution of the impulse responses of the two systems.

$$h[n] = h_1[n] * h_2[n] \tag{I.35}$$

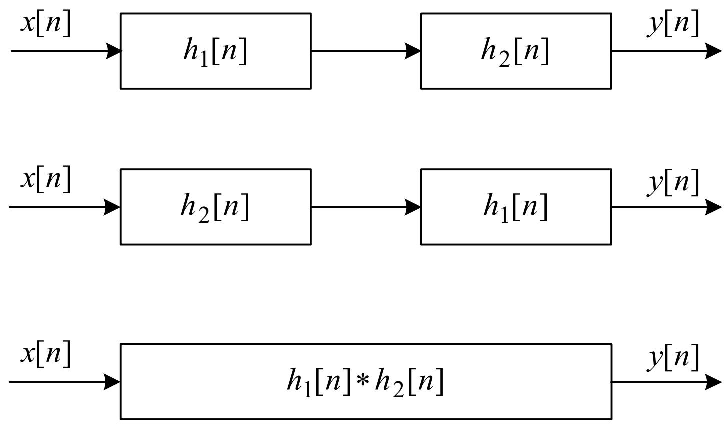


Figure I.3: Cascade Connection

Parallel Connection

In a parallel connection, the systems have the same input, and their outputs are summed to produce an overall output. It follows from the distributive property of convolution that the connection of two linear time-invariant systems in parallel is equivalent to a single system whose impulse response is the sum of the individual impulse responses; i.e.,

$$h[n] = h_1[n] + h_2[n] \tag{I.36}$$

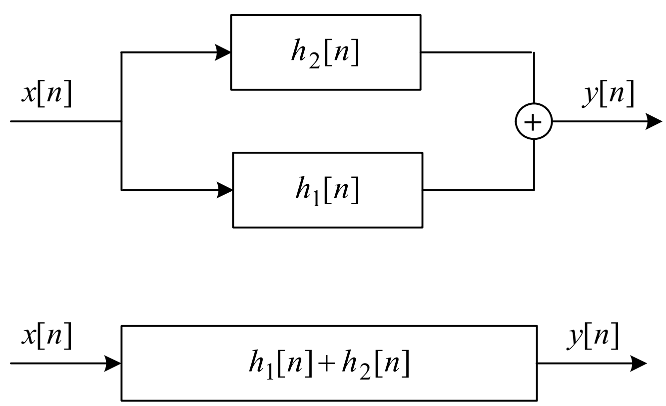


Figure I.4: Parallel Connection

Stability and Causality

The constraints of linearity and time invariance define a class of systems with very special properties. Stability and causality represent additional properties, and it is essential to know whether a linear time-invariant system is stable and whether it is causal. Recall that a stable system is a system for which every bounded input produces a bounded output. Linear time-invariant systems are stable if and only if the impulse response is absolutely summable, i.e., if

$$S = \sum_{k=-\infty}^{\infty} \left| h[k] \right| < \infty \tag{I.37}$$

The class of causal systems was defined earlier to represent systems for which the output $y[n_0]$ depends only on past samples of the input $x[n]$, i.e., for $n \le n_0$. It follows from (I.31) that this definition implies the following condition for causality of linear time-invariant systems.

$$h[n] = 0, \forall n < 0 \tag{I.38}$$

For this reason, it is sometimes convenient to refer to a sequence that is zero for $n < 0$ as a causal sequence, meaning that it could be the impulse response of a causal system.

Although the impulse response of nonlinear or time-varying systems can be found, it is generally of limited interest, since the convolution-sum formula and (I.37) and (I.38), expressing stability and causality, do not apply to such systems.

Ideal Delay

$$y[n] = x[n - n_d] \tag{I.39}$$

$$h[n] = \delta[n - n_d] \tag{I.40}$$

$$S = 1 \tag{I.41}$$

System is stable. System is causal as long as $n_d \ge 0$.

Moving Average (MA)

$$y[n] = \frac{1}{M_1 + M_2 + 1} \sum_{k=-M_1}^{M_2} x[n-k] \tag{I.42}$$

$$h[n] = \frac{1}{M_1 + M_2 + 1} \sum_{k=-M_1}^{M_2} \delta[n-k] \tag{I.43}$$

$$h[n] = \begin{cases} \dfrac{1}{M_1 + M_2 + 1}, & -M_1 \le n \le M_2 \\ 0, & \text{otherwise} \end{cases} \tag{I.44}$$

$$S = 1 \tag{I.45}$$

System is stable. System is causal as long as $-M_1 \ge 0$ and $M_2 \ge 0$.

Exercise I-3

Write and execute Matlab code to simulate a moving average system with different values of M_1 and M_2.

- Feed the systems with different random and deterministic input signals.
- Notice the smoothing effect of the MA system.

Accumulator

$$y[n] = \sum_{k=-\infty}^{n} x[k] \tag{I.46}$$

$$\begin{aligned} h[n] &= \sum_{k=-\infty}^{n} \delta[k] \\ &= u[n] \end{aligned} \tag{I.47}$$

$$S = \infty \tag{I.48}$$

System is unstable. System is causal.

Forward Difference

$$y[n] = x[n+1] - x[n] \tag{I.49}$$

$$h[n] = \delta[n+1] - \delta[n] \tag{I.50}$$

$$S = 2 \tag{I.51}$$

System is stable. System is noncausal.

Backward Difference

$$y[n] = x[n] - x[n-1] \tag{I.52}$$

$$h[n] = \delta[n] - \delta[n-1] \tag{I.53}$$

$$S = 2 \tag{I.54}$$

System is stable. System is causal.

Exponential Impulse Response

$$h[n] = a^n u[n] \tag{I.55}$$

$$\begin{aligned} S &= \sum_{k=-\infty}^{\infty} \left|h[k]\right| \\ &= \sum_{k=-\infty}^{\infty} \left|a^k u[k]\right| \\ &= \sum_{k=0}^{\infty} \left|a^k\right| = \sum_{k=0}^{\infty} |a|^k = \frac{1}{1-|a|}, \text{ if } |a| < 1 \end{aligned} \tag{I.56}$$

System is stable only if $|a| < 1$. System is causal.

For the ideal delay, moving-average, forward difference, and backward difference examples, it is clear that $S<\infty$, since the impulse response has only a finite number of nonzero samples. Such systems are called finite-duration impulse response (FIR) systems. Clearly, FIR systems will always be stable, as long as **each of the impulse response values is finite in magnitude**.

Exercise I-4

Consider the generalized form of an MA system

$$y[n]=\frac{1}{M_1+M_2+1}\sum_{k=-M_1}^{M_2}a_k x[n-k]$$

where $a_k\big|_{k=-M_1}^{M_2}$ are constants. This is actually an FIR system. Specify the parameters of the ideal delay, forward difference, and backward difference systems (M_1, M_2 and $a_k\big|_{k=-M_2}^{M_1}$) such that these systems are seen as special cases of this FIR system.

The accumulator, however, is unstable because $S=\infty$. The impulse response of the accumulator is infinite in duration. This is an example of the class of systems referred to as infinite-duration impulse response (IIR) systems. Not all IIR systems are unstable. An example of an IIR system that is stable is a system whose impulse response is

$$h[n]=a^n u[n], |a|<1 \tag{I.57}$$

In this case,

$$\begin{aligned} S &= \sum_{k=0}^{\infty}|a|^n \\ &= \frac{1}{1-|a|}<\infty \end{aligned} \tag{I.58}$$

Note that if $|a|\geq 1$, the series in (I.58) diverges and the system is unstable.

Noncausal to Causal Conversion

Consider the system in Figure I.5 below which consists of a forward difference system cascaded with an ideal delay of one sample.

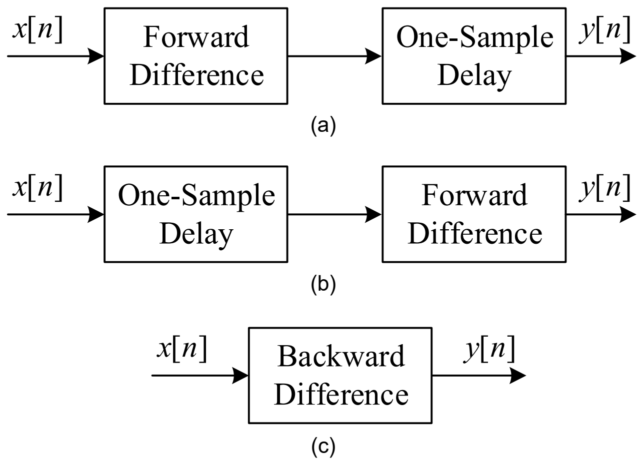


Figure I.5: Noncausal to Causal Conversion

According to the commutative property of convolution, the order in which systems are cascaded does not matter, as long as they are linear and time invariant. Therefore, we obtain the same result when we compute the forward difference of a sequence and delay the result (part a) as when we delay the sequence first and then compute the forward difference (part b).

The overall impulse response of the cascade system is

$$\begin{aligned} h[n] &= \left[\delta[n+1]-\delta[n]\right]*\delta[n-1] \\ &= \delta[n]-\delta[n-1] \end{aligned} \tag{I.59}$$

Thus, $h(n)$ is identical to the impulse response of the backward difference system; that is, the cascaded systems of part (a) and part (b) can be replaced by a backward difference system, as shown in part (c). Note that the noncausal forward difference systems in parts a and b have been converted to causal systems by cascading them with a delay. In general, any noncausal FIR system can be made causal by cascading it with a sufficiently long delay.

Inverse System

Consider the cascade of systems in Figure I.6 below.

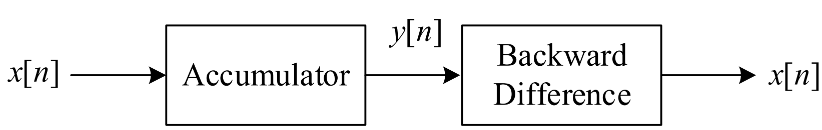


Figure I.6: Inverse System Example

The impulse response of the cascade system is

$$\begin{aligned} h[n] &= u[n]*\left[\delta[n]-\delta[n-1]\right] \\ &= u[n]-u[n-1] \\ &= \delta[n] \end{aligned} \tag{I.60}$$

That is, the cascade combination of an accumulator followed by a backward difference (or vice versa) yields a system whose overall impulse response is the impulse. Thus, the output of the cascade combination will always be equal to the input. In this case, the backward difference system compensates exactly for (or inverts) the effect of the accumulator; that is, the backward difference system is the inverse system for the accumulator. From the commutative property of convolution, the accumulator is likewise the inverse system for the backward difference system.

I.7. Linear Constant-Coefficient Difference Equations

I.7.A. General Equation

An important subclass of linear time-invariant systems consists of those systems for which the input $x[n]$ and the output $y[n]$ satisfy a linear constant-coefficient difference equation of the form

$$\sum_{k=0}^{K} a_k y[n-k] = \sum_{m=0}^{M} b_m x[n-m] \tag{I.61}$$

Example I-9

For the accumulator,

$$\begin{aligned} y[n] &= y[n-1] + x[n] \\ y[n] - y[n-1] &= x[n] \end{aligned} \tag{I.62}$$

Note that

$$K = 1,\ M = 0,\ a_0 = 1,\ a_1 = -1,\ b_0 = 1$$

Exercise I-5

Determine the difference equation representation of the moving average system. Note that there can be more than one answer.

Sketch a block diagram of the moving average system.

A linear constant coefficient difference equation for discrete-time systems does not provide a unique specification of the output for a given input. Specifically, suppose that, for a given input $x_p[n]$, we have determined by some means one output sequence $y_p[n]$, so that an equation of the form of (I.61) is satisfied. Let $y_h[n]$ be the solution to (I.61) when $x[n] = 0$, i.e.,

$$\sum_{k=0}^{K} a_k y_h[n-k] = 0 \tag{I.63}$$

Equation (I.63) is referred to as the homogeneous equation and $y_h[n]$ the homogeneous solution. Clearly enough, the output $y[n] = y_p[n] + y_h[n]$ and the input $x[n] = x_p[n]$ satisfy (I.61), meaning that $y[n] = y_p[n] + y_h[n]$ can be the system response to $x_p[n]$.

The sequence $y_h[n]$ is in fact a member of a family of solutions of the form

$$y_h[n] = \sum_{l=1}^{K} A_l z_l^n \tag{I.64}$$

Substituting (I.64) into (I.63) shows that the complex numbers z_l must be roots of the polynomial

$$\sum_{k=0}^{K} a_k z^{-k} = 0 \tag{I.65}$$

Equation (I.64) assumes that all K roots of the polynomial in (I.65) are distinct. The form of terms associated with multiple roots is slightly different, but there are always K undetermined coefficients. Since $y_h[n]$ has K undetermined coefficients, a set of K auxiliary conditions is required for the unique specification of $y[n]$ for a given $x[n]$.

I.7.B. Recursive Computation of Difference Equations

Consider the difference equation

$$y[n] - ay[n-1] = x[n] \tag{I.66}$$

Let the input sequence be $x[n] = G\delta[n]$, where G is an arbitrary constant. Let $y[-1] = c$, where c is an arbitrary constant. Beginning with this value, the output for $n > -1$ can be computed recursively be first rewriting (I.66) in the form

$$y[n] = ay[n-1] + x[n] \tag{I.67}$$

Then,

$$y[0] = ac + G$$

$$y[1] = a^2 c + aG$$

$$\vdots$$

$$y[n] = a^{n+1} c + a^n G$$

To determine the output for $n < 0$, we express the difference equation in the form

$$\begin{aligned} y[n-1] &= a^{-1}\left[y[n] - x[n]\right] \\ y[n] &= a^{-1}\left[y[n+1] - x[n+1]\right] \end{aligned} \tag{I.68}$$

Using the auxiliary condition $y[-1] = c$, we can compute $y[n]$ for $n < -1$ as follows:

$$y[-2] = a^{-1} c$$

$$y[-3] = a^{-2} c$$

$$\vdots$$

$$y[n] = a^{n+1}c$$

Combining the results for $n \geq 0$ and $n < 0$, we obtain

$$y[n] = a^{n+1}c + a^{n}Gu[n], \forall n \tag{I.69}$$

Several important points are illustrated by the above procedure. First, note that we implemented the system by recursively computing the output in both the positive and the negative directions, beginning with $n = -1$. Clearly, this procedure is noncausal. Also, note that when $G = 0$ (zero input), $y[n] = a^{n+1}c$. A linear system requires that the output be zero for all time when the input is zero for all time. Consequently, this system is not linear. Furthermore, if the input were shifted by n_0 samples, i.e., $x_1[n] = G\delta[n - n_0]$, the output would be

$$\begin{aligned} y_1[n] &= a^{n+1}c + Ga^{n-n_0}u[n-n_0] \\ &\neq y[n-n_0] \end{aligned} \tag{I.70}$$

The system is therefore not time invariant.

Our principal interest is in systems that are linear and time invariant, in which case the auxiliary conditions must be consistent with these additional requirements. When we discuss the solution of difference equations using the z-transform, we implicitly incorporate conditions of linearity and time invariance. As we will see in that discussion, even with the additional constraints of linearity and time invariance, the solution to the difference equation, and therefore the system, is not uniquely specified. In particular, there are in general, both causal and noncausal linear time-invariant systems consistent with a given difference equation.

If a system is characterized by a linear constant-coefficient difference equation and is further specified to be linear, time invariant, and causal, the solution is unique. In this case, the auxiliary conditions are often stated as initial-rest conditions.

Exercise I-6

Assuming initial-rest conditions, solve the difference equation

$$y[n] - ay[n-1] = G\delta[n]$$

Study linearity, causality and time invariance of the system.

I.8. Frequency Domain Representation of Discrete-time Signals and Systems

As with continuous-time signals, discrete-time signals may be represented in a number of different ways. For example, sinusoidal and complex exponential sequences play a particularly important role in representing discrete-time signals. This is because complex exponential sequences are eigenfunctions of linear time-invariant systems, and the response to a sinusoidal input is sinusoidal with the same frequency as the input and with amplitude and phase that are determined by the system.

Consider an input sequence $x[n]=e^{j\omega n}$ for $-\infty<n<\infty$. The corresponding output of a linear time-invariant system with impulse response $h[n]$ is

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} h[k]e^{j\omega(n-k)} \\ &= e^{j\omega n}\sum_{k=-\infty}^{\infty} h[k]e^{-j\omega k} \end{aligned} \tag{I.71}$$

Let's define

$$H(e^{j\omega})=\sum_{k=-\infty}^{\infty} h[k]e^{-j\omega k} \tag{I.72}$$

Note that $H(e^{j\omega})$ depends only on the system impulse response and the input frequency. $H(e^{j\omega})$ is called the frequency response of the system. Note that the output can be written in the form

$$y[n]=H(e^{j\omega})e^{j\omega n} \tag{I.73}$$

Consequently, $e^{j\omega n}$ is an eigenfunction of the system, and the associated eigenvalue is $H(e^{j\omega})$. In general, $H(e^{j\omega})$ is complex and can be expressed in terms of its real and imaginary parts as

$$H(e^{j\omega})=H_R(e^{j\omega})+jH_I(e^{j\omega}) \tag{I.74}$$

$H(e^{j\omega})$ can also be written in terms of its magnitude and phase as

$$H(e^{j\omega})=\left|H(e^{j\omega})\right|e^{j\angle H(e^{j\omega})} \tag{I.75}$$

Exercise I-7

- Determine and sketch the magnitude and phase of the frequency response of the ideal delay system.
- Determine and sketch the magnitude and phase of the frequency response of the backward difference system.

A broad class of signals can be represented as a linear combination of complex exponentials in the form

$$x[n]=\sum_{k}\alpha_k e^{j\omega_k n} \tag{I.76}$$

From the principle of superposition, the corresponding output of a linear time-invariant system is

$$y[n]=\sum_{k}\alpha_k H(e^{j\omega_k})e^{j\omega_k n} \tag{I.77}$$

Exercise I-8

1- Determine the response of an LTI system to a sinusoidal input of the form $x[n] = A_1 \cos(\omega_1 n + \phi_1) + A_2 \cos(\omega_2 n + \phi_2)$.
2- What is the output when the system is an ideal delay?
3- What is the output when $h[n] = \alpha^n u[n]$, where α is generally complex with $|\alpha| < 1$?

The concept of the frequency response of linear time-invariant systems is essentially the same for continuous-time and discrete-time systems. However, an important distinction arises because the frequency response of discrete-time linear time-invariant systems is always a periodic function of the frequency variable with period 2π. The periodicity of $H(e^{j\omega})$ can be seen from (I.72). It, therefore, follows that we need to specify $H(e^{j\omega})$ only over an interval of length 2π, e.g., $0 \le \omega < 2\pi$ or $-\pi \le \omega < \pi$. It is generally convenient to specify $H(e^{j\omega})$ over the interval $-\pi \le \omega < \pi$. With respect to this interval, the "low frequencies" are frequencies close to zero, while the "high frequencies" are frequencies close to $\pm\pi$.

Assignment I.1

Frequency Response of Ideal Frequency Selective Filters: e.g., LPF, HPF, …

I.9. Representation of Sequences by Fourier Transforms

Many sequences can be represented by a Fourier integral of the form

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} d\omega \tag{I.78}$$

where

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n} \tag{I.79}$$

Equations (I.78) and (I.79) form a Fourier representation for the sequence. Equation (I.78), the inverse Fourier transform, is a synthesis formula, while (I.79) is an analysis formula. In general, the Fourier transform is a complex-valued function of ω. As with the frequency response, we may either express $X(e^{j\omega})$ in rectangular form as

$$X(e^{j\omega}) = X_R(e^{j\omega}) + jX_I(e^{j\omega}) \tag{I.80}$$

or in polar form as

$$X(e^{j\omega}) = \left| X(e^{j\omega}) \right| e^{j\angle X(e^{j\omega})} \tag{I.81}$$

The Fourier transform is sometimes referred to as the Fourier spectrum or, simply, the spectrum. Also, the terminology magnitude spectrum or amplitude spectrum is sometimes used to refer to $\left|X(e^{j\omega})\right|$, and the angle or phase $\angle X(e^{j\omega})$ is sometimes called the phase spectrum.

The frequency response of a linear time-invariant system is simply the Fourier transform of the impulse response and, therefore, the impulse response can be obtained from the frequency response by applying the inverse Fourier transform integral; i.e.,

$$h[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} H(e^{j\omega})e^{j\omega n}d\omega \tag{I.82}$$

Equation (I.79) is of the form of a Fourier series for the continuous-variable periodic function $X(e^{j\omega})$, and (I.78), which expresses the sequence values $x[n]$ in terms of the periodic function $X(e^{j\omega})$, is of the form of the integral that would be used to obtain the coefficients in the Fourier series.

Exercise I-9

Show that equations (I.78) and (I.79) are indeed inverses of each other.

Show that for $x[n]$ to have a Fourier transform $X(e^{j\omega})$, $x(n)$ must be absolutely summable, i.e.,

$$\sum_{n=-\infty}^{\infty} \left|x[n]\right| < \infty$$

Since a stable sequence is, by definition, absolutely summable, all stable sequences have Fourier transforms. It also follows, then, that any stable system will have a finite continuous periodic frequency response.

Absolute summability is a sufficient condition for the existence of a Fourier transform representation. Clearly, any finite length sequence with finite sequence values is absolutely summable and thus will have a Fourier transform representation. In the context of linear time-invariant systems, any FIR system with finite impulse response values will be stable and therefore will have a finite continuous periodic frequency response. When a sequence has infinite length, we must be concerned about convergence of the infinite sum.

Absolute summability is a sufficient condition for the existence of a Fourier transform representation. Some sequences are not absolutely summable, but are square summable, i.e.,

$$\sum_{n=-\infty}^{\infty} \left|x[n]\right|^2 < \infty \tag{I.83}$$

Such sequences can be represented by a Fourier transform if we are willing to relax the condition of uniform convergence of the infinite sum defining $X(e^{j\omega})$. Specifically, in this case we have mean-square convergence; that is, with

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \tag{I.84}$$

and

$$X_M(e^{j\omega}) = \sum_{n=-M}^{M} x[n]e^{-j\omega n} \tag{I.85}$$

we have

$$\lim_{M\to\infty} \int_{-\pi}^{\pi} \left| X(e^{j\omega}) - X_M(e^{j\omega}) \right|^2 d\omega = 0 \tag{I.86}$$

Exercise I-10

Determine the Fourier transforms of:

1. $x[n] = 1, -\infty < n < \infty$
2. $x[n] = e^{j\omega_0 n}, -\infty < n < \infty$
3. $x[n] = u[n]$

I.10. Symmetry Properties of the Fourier Transform

A conjugate-symmetric sequence $x_e[n]$ is defined as a sequence for which

$$x_e[n] = x_e^*[-n] \tag{I.87}$$

A conjugate-anti-symmetric sequence $x_o[n]$ is defined as a sequence for which

$$x_o[n] = -x_o^*[-n] \tag{I.88}$$

Any sequence $x[n]$ can be expressed as a sum of a conjugate-symmetric and conjugate-anti-symmetric sequence. Specifically,

$$x[n] = x_e[n] + x_o[n] \tag{I.89}$$

where

$$x_e[n] = \frac{1}{2}\left[x[n] + x^*[-n] \right] \tag{I.90}$$

$$x_o[n] = \frac{1}{2}\left[x[n] - x^*[-n] \right] \tag{I.91}$$

A real sequence that is conjugate symmetric is called an even sequence, and a real sequence that is conjugate anti-symmetric is called an odd sequence.

A Fourier transform $X(e^{j\omega})$ can be decomposed into a sum of conjugate-symmetric and conjugate-anti-symmetric functions as

$$X(e^{j\omega}) = X_e(e^{j\omega}) + X_o(e^{j\omega}) \tag{I.92}$$

where

$$X_e(e^{j\omega}) = \frac{1}{2}\left[X(e^{j\omega}) + X^*(e^{-j\omega})\right] \tag{I.93}$$

$$X_o(e^{j\omega}) = \frac{1}{2}\left[X(e^{j\omega}) - X^*(e^{-j\omega})\right] \tag{I.94}$$

Note that

$$X_e(e^{j\omega}) = X_e^*(e^{-j\omega}) \tag{I.95}$$

$$X_o(e^{j\omega}) = -X_o^*(e^{-j\omega}) \tag{I.96}$$

Table I.1: Symmetry Properties of the Fourier Transform

Sequence	Fourier Transform
$x[n]$	$X(e^{j\omega})$
$x^*[n]$	$X^*(e^{-j\omega})$
$x^*[-n]$	$X^*(e^{j\omega})$
$\mathrm{Re}\{x[n]\}$	$X_e(e^{j\omega})$
$j\,\mathrm{Im}\{x[n]\}$	$X_o(e^{j\omega})$
$x_e[n]$	$X_R(e^{j\omega}) = \mathrm{Re}\left\{X(e^{j\omega})\right\}$
$x_o[n]$	$jX_I(e^{j\omega}) = j\,\mathrm{Im}\left\{X(e^{j\omega})\right\}$

Symmetry Properties of Fourier Transforms of Real Sequences

In all the equations below, $x[n]$ is assumed to be real-valued.

$$X_R(e^{j\omega}) = X_R(e^{-j\omega}) \tag{I.97}$$

$$X_I(e^{j\omega}) = -X_I(e^{-j\omega}) \tag{I.98}$$

$$\left|X(e^{j\omega})\right| = \left|X(e^{-j\omega})\right| \tag{I.99}$$

$$\sphericalangle X(e^{j\omega}) = -\sphericalangle X(e^{-j\omega}) \tag{I.100}$$

$$\mathcal{F}\{x_e[n]\} = X_R(e^{j\omega}) \tag{I.101}$$

$$\mathcal{F}\{x_o[n]\} = jX_I(e^{j\omega}) \tag{I.102}$$

I.11. Fourier Transform Theorems

Linearity of the Fourier Transform

If

$$x_1[n] \overset{\mathcal{F}}{\rightleftharpoons} X_1(e^{j\omega})$$
$$x_2[n] \overset{\mathcal{F}}{\rightleftharpoons} X_2(e^{j\omega})$$

then,

$$a_1x_1[n] + a_2x_2[n] \overset{\mathcal{F}}{\rightleftharpoons} a_1X_1(e^{j\omega}) + a_2X_2(e^{j\omega}) \tag{I.103}$$

Time Shifting and Frequency Shifting

$$x[n-n_d] \overset{\mathcal{F}}{\rightleftharpoons} e^{-j\omega n_d} X(e^{j\omega}) \tag{I.104}$$

$$e^{j\omega_0 n}x[n] \overset{\mathcal{F}}{\rightleftharpoons} X(e^{j(\omega-\omega_0)}) \tag{I.105}$$

Time Reversal

$$x[-n] \overset{\mathcal{F}}{\rightleftharpoons} X(e^{-j\omega}) \tag{I.106}$$

Parseval's Theorem

$$\sum_{n=-\infty}^{\infty} \left|x[n]\right|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left|X(e^{j\omega})\right|^2 d\omega \tag{I.107}$$

The function $\left|X(e^{j\omega})\right|^2$ is called the energy density spectrum, since it determines how the energy is distributed in the frequency domain. Necessarily, the energy density spectrum is defined only for finite-energy signals.

Convolution Theorem

$$x[n] * h[n] \overset{\mathcal{F}}{\rightleftharpoons} X(e^{j\omega})H(e^{j\omega}) \tag{I.108}$$

$$
\begin{aligned}
x[n]g[n] &\xleftrightarrow{\mathcal{F}} \frac{1}{2\pi} X(e^{j\omega}) * G(e^{j\omega}) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\lambda}) G(e^{j(\omega-\lambda)}) d\lambda
\end{aligned}
\tag{I.109}
$$

The operation $x[n]g[n]$ is known as modulation or windowing. Equation (I.109) is a periodic convolution, i.e., a convolution of two periodic functions with the limits of integration extending over only one period.

In contrast to the continuous-time case, where this duality is complete, in the discrete-time case fundamental differences arise because the Fourier transform is a sum while the inverse transform is an integral with a periodic integrand. Although for continuous time we can state that convolution in the time domain is represented by multiplication in the frequency domain and vice versa, in discrete time this statement must be modified somewhat. Specifically, discrete-time convolution of sequences (the convolution sum) is equivalent to multiplication of corresponding periodic Fourier transforms, and multiplication of sequences is equivalent to periodic convolution of corresponding Fourier transforms.

Exercise I-11

Determine

1. the DTFT of $x[n] = a^{2n} u[n-3]$
2. the IDTFT of $X(e^{j\omega}) = \dfrac{c}{\left(a - e^{-j\omega}\right)\left(b - e^{-j\omega}\right)}$
3. the impulse response when the frequency response is $H(e^{j\omega}) = \begin{cases} e^{-j\omega n_0}, & \omega_c < |\omega| < \pi \\ 0, & |\omega| < \omega_c \end{cases}$
4. the impulse response when $y[n] - \frac{1}{2} y[n-2] = x[n] + \frac{1}{3} x\left[n-3\right]$

Table I.2: Properties of the Fourier Transform

Sequence	Fourier Transform
$x[n]$	$X(e^{j\omega})$
$y[n]$	$Y(e^{j\omega})$
$ax[n] + by[n]$	$aX(e^{j\omega}) + bY(e^{j\omega})$
$x[n-n_d]$	$e^{-j\omega n_d} X(e^{j\omega})$
$e^{j\omega_0 n} x[n]$	$X(e^{j(\omega-\omega_0)})$
$x[-n]$	$X(e^{-j\omega})$

$nx[n]$	$j\frac{d}{d\omega}X(e^{j\omega})$
$x[n]*y[n]$	$X(e^{j\omega})Y(e^{j\omega})$
$x[n]y[n]$	$\frac{1}{2\pi}\int_{-\pi}^{\pi}X(e^{j\theta})Y(e^{j(\omega-\theta)})d\theta$

Parseval's Theorem

$$\sum_{n=-\infty}^{\infty}|x[n]|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|X(e^{j\omega})\right|^2 d\omega \tag{I.110}$$

$$\sum_{n=-\infty}^{\infty}x[n]y^*[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi}X(e^{j\omega})Y^*(e^{j\omega})d\omega \tag{I.111}$$

Table I.3: Fourier Transform Pairs

Sequence	Fourier Transform
$\delta[n]$	1
$\delta[n-n_0]$	$e^{-j\omega n_0}$
1	$2\pi\sum_{k=-\infty}^{\infty}\delta(\omega+2\pi k)$
$a^n u[n], \|a\|<1$	$\frac{1}{1-ae^{-j\omega}}$
$u[n]$	$\frac{1}{1-e^{-j\omega}}+\pi\sum_{k=-\infty}^{\infty}\delta(\omega+2\pi k)$
$(n+1)a^n u[n], \|a\|<1$	$\frac{1}{(1-ae^{-j\omega})^2}$
$\frac{r^n \sin\omega_p(n+1)}{\sin\omega_p}u[n], \|r\|<1$	$\frac{1}{1-2r\cos\omega_p e^{-j\omega}+r^2e^{-j2\omega}}$
$\frac{\sin\omega_c n}{\pi n}$	$\begin{cases}1, & \|\omega\|<\omega_c \\ 0, & \omega_c<\|\omega\|\le\pi\end{cases}$

$\begin{cases} 1, & 0 \le n \le M \\ 0, & \text{otherwise} \end{cases}$	$\frac{\sin\left[\omega(M+1)/2\right]}{\sin(\omega/2)} e^{-j\omega M/2}$
$e^{j\omega_0 n}$	$2\pi \sum_{k=-\infty}^{\infty} \delta(\omega - \omega_0 + 2\pi k)$
$\cos(\omega_0 n + \phi)$	$\pi \sum_{k=-\infty}^{\infty} \left[e^{j\phi} \delta(\omega - \omega_0 + 2\pi k) + e^{-j\phi} \delta(\omega + \omega_0 + 2\pi k) \right]$

I.12. Discrete-Time Random Signals

In many situations, the processes that generate signals are complex enough to make precise description of a signal extremely difficult or undesirable, if not impossible. In such cases, modeling the signal as a stochastic process is analytically useful. The term stochastic process, or random process, is used to describe the time evolution of a statistical phenomenon according to probabilistic laws. The time evolution of the phenomenon means that the stochastic process is a function of time, defined on some observation interval. The statistical nature of the phenomenon means that, before conducting an experiment, it is not possible to define exactly the way it evolves in time. Examples of a stochastic process include speech signals, television signals, radar signals, digital computer data, the output of a communication channel, seismological data, and noise.

The form of a stochastic process that is of interest to us is one that is defined at discrete and uniformly spaced instants of time. Such a restriction may arise naturally in practice, as in the case of digital computer data. Alternatively, the stochastic process may be defined originally for a continuous range of values of time; however, before processing, it is sampled uniformly in time, with the sampling rate chosen to be greater than twice the highest frequency component of the process.

A stochastic signal is considered to be a member of an infinite ensemble of discrete-time signals that is characterized by a set of probability density functions. More specifically, for a particular signal at a particular time, the amplitude of the signal sample at that time is assumed to have been determined by an underlying scheme of probabilities. For convenience of notation, we normalize time with respect to the sampling period. Each individual sample $x[n]$ of a particular signal is assumed to be an outcome of some underlying random variable x_n. The entire signal is represented by a collection of such random variables, one for each sample time, $-\infty < n < \infty$. This collection of random variables is called a random process. To completely describe the random process, we need to specify the individual and joint probability distributions of all the random variables.

An individual random variable x_n is described by the cumulative distribution function (CDF)

$$F_{x_n}(\xi_n, n) = \Pr\left\{x_n \le \xi_n\right\} \tag{I.112}$$

If x_n takes on a continuous range of values, it is equivalently specified by the probability density function

$$f_{x_n}(x,n) = \frac{\partial F_{x_n}(x,n)}{\partial x} \tag{I.113}$$

The interdependence of two random variables x_n and x_m of a random process is described by the joint CDF

$$F_{x_n,x_m}(\xi_n, n, \xi_m, m) = \Pr\left\{x_n \le \xi_n, x_m \le \xi_m\right\} \tag{I.114}$$

The two random variables have the joint probability density function

$$f_{x_n,x_m}(x,n,y,m) = \frac{\partial F_{x_n,x_m}(x,n,y,m)}{\partial x \partial y} \tag{I.115}$$

Two random variables are statistically independent if knowledge of the value of one does not affect the probability density of the other

$$F_{x_n,x_m}(\xi_n, n, \xi_m, m) = F_{x_n}(\xi_n, n) F_{x_m}(\xi_m, m) \tag{I.116}$$

A complete characterization of a random process requires the specification of all possible joint distributions. As we have indicated, these distributions may be functions of the time indices. In the case where all the distributions are independent of a shift of time origin, the random process is said to be stationary. For example, the second-order distribution of a stationary process satisfies

$$F_{x_{n+k},x_{m+k}}(\xi_{n+k}, n, \xi_{m+k}, m) = F_{x_n,x_m}(\xi_n, n, \xi_m, m) \tag{I.117}$$

In many applications of discrete-time signal processing, random processes serve as models for signals in the sense that a particular signal can be considered a sample sequence of a random process. Although the details of such signals are unpredictable – making a deterministic approach to signal representation inappropriate – certain average properties of the ensemble can be determined, given the probability law of the process. These average properties often serve as a useful, although incomplete, characterization of such signals.

I.12.A. Statistical Averages

The average, or mean, of a random process is defined as

$$\begin{aligned} \mu_{x_n} &= \mathrm{E}\left[x_n\right] \\ &= \int_{-\infty}^{\infty} x f_{x_n}(x) dx \end{aligned} \tag{I.118}$$

Stochastic signals are not absolutely summable nor square summable and, consequently, do not directly have Fourier transforms. Many (but not all) of the properties of such signals can be summarized in terms of averages such as the autocorrelation or auto covariance sequence, for which the Fourier transform often exists. The Fourier transform of the autocovariance sequence has a useful interpretation in terms of the frequency distribution of the power in the signal. The use of the autocovariance sequence and its transform has another important advantage: The effect of processing stochastic signals with a discrete-time linear system can be conveniently described in terms of the effect of the system on the autocovariance sequence.

Read
Appendix A – Oppenheim

We focus on wide-sense stationary random signals and their representations in the context of processing with linear time-invariant systems.

Consider a stable linear time-invariant system with real impulse response $h[n]$. Let $x[n]$ be a real-valued sequence that is a sample sequence of a wide-sense stationary discrete-time random process. Then the output of the linear system is also a sample function of a random process related to the input process by the linear transformation

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} h[n-k]x[k] \\ &= \sum_{k=-\infty}^{\infty} h[k]x[n-k] \end{aligned} \tag{I.119}$$

The input signal may be characterized by its mean m_x and its autocorrelation function $\varphi_{xx}[m]$, or we may also have additional information about first- or even second-order probability distributions. In characterizing the output random process $y[n]$ we desire similar information. For many applications, it is sufficient to characterize both the input and output in terms of simple averages, such as the mean, variance, and autocorrelation.

The mean of the output process is

$$\begin{aligned} m_y[n] &= \mathrm{E}\left[y[n]\right] \\ &= \sum_{k=-\infty}^{\infty} h[k]\mathrm{E}\left[x[n-k]\right] \\ &= \sum_{k=-\infty}^{\infty} h[k]m_x[n-k] \\ &= h[n] * m_x[n] \end{aligned} \tag{I.120}$$

Since the input is stationary,

$$m_x[n-k] = m_x \tag{I.121}$$

Consequently,

$$\begin{aligned} m_y[n] &= m_x \sum_{k=-\infty}^{\infty} h[k] \\ &= m_x H(e^{j0}) \end{aligned} \tag{I.122}$$

The mean of the output is constant. The autocorrelation function of the output process (assuming a real input) is

$$\begin{aligned}\varphi_{yy}[n,n+m] &= \mathrm{E}\left[y[n]y[n+m]\right] \\ &= \mathrm{E}\left[\sum_{k=-\infty}^{\infty}\sum_{r=-\infty}^{\infty} h[k]h[r]x[n-k]x[n+m-r]\right] \\ &= \sum_{k=-\infty}^{\infty} h[k]\sum_{r=-\infty}^{\infty} h[r]\,\mathrm{E}\,x[n-k]x[n+m-r]\end{aligned} \qquad \text{(I.123)}$$

Since $x[n]$ is assumed to be stationary, $\mathrm{E}\left[x[n-k]x[n+m-r]\right]$ depends only on the time difference $m+k-r$. Therefore,

$$\begin{aligned}\varphi_{yy}[n,n+m] &= \sum_{k=-\infty}^{\infty} h[k]\sum_{r=-\infty}^{\infty} h[r]\varphi_{xx}[m+k-r] \\ &= \varphi_{yy}[m]\end{aligned} \qquad \text{(I.124)}$$

That is, the output autocorrelation sequence also depends only on the time difference m. Thus, for a linear time-invariant system having a wide-sense stationary input, the output is also wide-sense stationary.

By making the substitution $l = r-k$,

$$\begin{aligned}\varphi_{yy}[m] &= \sum_{l=-\infty}^{\infty} \varphi_{xx}[m-l]\sum_{k=-\infty}^{\infty} h[k]h[l+k] \\ &= \sum_{l=-\infty}^{\infty} \varphi_{xx}[m-l]c_{hh}[l] \\ &= \varphi_{xx}[m]*c_{hh}[m]\end{aligned} \qquad \text{(I.125)}$$

where

$$c_{hh}[l] = \sum_{k=-\infty}^{\infty} h[k]h[l+k] \qquad \text{(I.126)}$$

The sequence $c_{hh}[l]$ is called a deterministic (or time) correlation sequence or, simply, the correlation sequence of $h[n]$. It should be emphasized that $c_{hh}[l]$ is the time correlation of an aperiodic, finite-energy-sequence and should not be confused with the autocorrelation of an infinite-energy random sequence. It can also be seen that

$$c_{hh}[l] = h[l]*h[-l] \qquad \text{(I.127)}$$

The last result allows (I.125) to be rewritten as

$$\varphi_{yy}[m] = \varphi_{xx}[m]*h[m]*h[-m] \qquad \text{(I.128)}$$

Assume, for convenience, that $m_x = 0$; i.e., the autocorrelation and autocovariance sequences are identical. Then,

$$\Phi_{yy}(e^{j\omega}) = C_{hh}(e^{j\omega})\Phi_{xx}(e^{j\omega}) \tag{I.129}$$

From (I.127), we have

$$\begin{aligned} C_{hh}(e^{j\omega}) &= H(e^{j\omega})H^*(e^{j\omega}) \\ &= \left|H(e^{j\omega})\right|^2 \end{aligned} \tag{I.130}$$

Therefore,

$$\Phi_{yy}(e^{j\omega}) = \Phi_{xx}(e^{j\omega})\left|H(e^{j\omega})\right|^2 \tag{I.131}$$

Note that the total average output power is equal to

$$\begin{aligned} \mathrm{E}\left[y^2[n]\right] &= \varphi_{yy}[0] \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{yy}(e^{j\omega})d\omega \end{aligned} \tag{I.132}$$

Substituting (I.131) into (I.132),

$$\mathrm{E}\left[y^2[n]\right] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{xx}(e^{j\omega})\left|H(e^{j\omega})\right|^2 d\omega \tag{I.133}$$

Exercise I-12

A zero-mean white noise sequence $x[n]$ with an autocorrelation function $\varphi_{xx}[m] = \sigma_x^2\delta[m]$ is the input of an ideal BPF that passes frequencies $\omega_a \le |\omega| \le \omega_b$. Determine

1. the average noise power at the filter input
2. the output noise power spectral density
3. the average output noise power
4. the output noise autocorrelation function

Exercise I-13

Repeat the exercise above if the BPF is replaced by a filter whose frequency response is $H(e^{j\omega}) = \dfrac{1}{1 - ae^{-j\omega}}$. Furthermore, sketch the magnitude and phase responses of the filter.

The cross-correlation between the input and output of a linear time-invariant system is given by

$$\begin{aligned}\varphi_{xy}[m] &= \mathrm{E}\left[x[n]y[n+m]\right] \\ &= \mathrm{E}\left[x[n]\sum_{k=-\infty}^{\infty} h[k]x[n+m-k]\right] \\ &= \sum_{k=-\infty}^{\infty} h[k]\varphi_{xx}[m-k] \\ &= \varphi_{xx}[m]*h[m]\end{aligned} \tag{I.134}$$

Using the DTFT,

$$\Phi_{xy}(e^{j\omega}) = H(e^{j\omega})\Phi_{xx}(e^{j\omega}) \tag{I.135}$$

This result has a useful application when the input is white noise, i.e., when $\varphi_{xx}[m] = \sigma_x^2\delta[m]$. In this case

$$\varphi_{xy}[m] = \sigma_x^2 h[m] \tag{I.136}$$

That is, for a zero-mean white-noise input, the cross-correlation between input and output of a linear system is proportional to the impulse response of the system. Similarly, the power spectrum of a white-noise input is

$$\Phi_{xy}(e^{j\omega}) = \sigma_x^2 H(e^{j\omega}) \tag{I.137}$$

In other words, the cross power spectrum is in this case proportional to the frequency response of the system.

***

II. The z-Transform

II.1. Definitions and Region of Convergence

The z-transform of a sequence $x(n)$ is defined as

$$X(z)=\sum_{n=-\infty}^{\infty}x[n]z^{-n} \tag{II.1}$$

This equation is, in general, an infinite sum or infinite power series, with z being a complex variable. The correspondence between a sequence and its z-transform is indicated by the notation

$$x[n]\xleftrightarrow{\mathcal{Z}}X(z) \tag{II.2}$$

The z-transform, as we have defined it in (II.1), is often referred to as the two-sided or bilateral z-transform, in contrast to the one-sided or unilateral z-transform, which is defined as

$$X(z)=\sum_{n=0}^{\infty}x[n]z^{-n} \tag{II.3}$$

If we replace the complex variable z in (II.1) with the complex variable $e^{j\omega}$, then the z-transform reduces to the Fourier transform. This is one motivation for the notation $X(e^{j\omega})$ for the Fourier transform; when it exists, the Fourier transform is simply $X(z)$ with $z=e^{j\omega}$. This corresponds to restricting z to have unity magnitude; i.e., for $|z|=1$, the z-transform corresponds to the Fourier transform.

We can express the complex variable z in polar form as

$$z=re^{j\omega} \tag{II.4}$$

Therefore,

$$\begin{aligned}X(re^{j\omega})&=\sum_{n=-\infty}^{\infty}x[n]\left(re^{j\omega}\right)^{-n}\\&=\sum_{n=-\infty}^{\infty}\left(x[n]r^{-n}\right)e^{-j\omega n}\end{aligned} \tag{II.5}$$

In the z-plane, the contour corresponding to $|z|=1$ is a circle of unit radius:

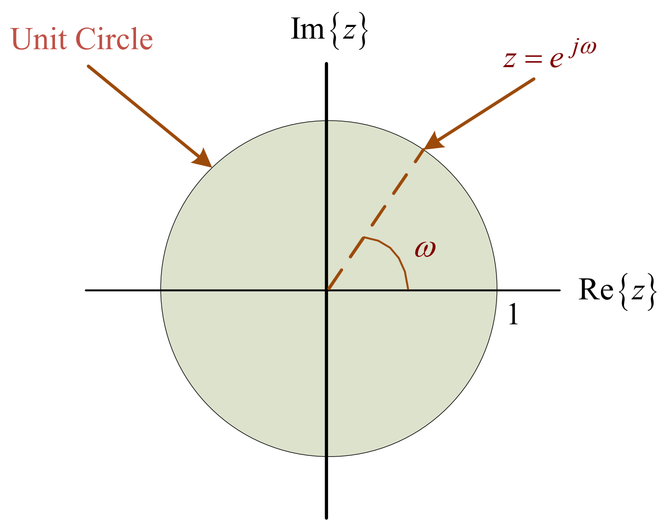


Figure II.1: Unit circle in z-plane

The z-transform evaluated on the unit circle corresponds to the Fourier transform. Note that ω is the angle between the vector to a point z on the unit circle and the real axis of the complex z-plane.

Interpreting the Fourier transform as the z-transform on the unit circle in the z-plane corresponds conceptually to wrapping the linear frequency axis around the unit circle with $\omega = 0$ at $z = 1$ and $\omega = \pi$ at $z = -1$. With this interpretation, the inherent periodicity in frequency of the Fourier transform is captured naturally, since a change of angle of 2π radians in the z-plane corresponds to traversing the unit circle once and returning to exactly the same point.

The z-transform does not converge for all sequences or for all values of z. For any given sequence, the set of values of z for which the z-transform converges is called the region of convergence, which we abbreviate ROC. For convergence of the z-transform,

$$\sum_{n=-\infty}^{\infty} \left| x(n) r^{-n} \right| < \infty \tag{II.6}$$

Because of the multiplication of the sequence by the real exponential r^{-n}, it is possible for the z-transform to converge even if the Fourier transform does not.

For example, the sequence $x(n) = u(n)$ is not absolutely summable, and therefore, the Fourier transform does not converge absolutely. However, $r^{-n}u(n)$ is absolutely summable if $r > 1$. This means that the z-transform for the unit step exists with a region of convergence $|z| > 1$.

Convergence of z-transform depends only on $|z|$, since $|X(z)| < \infty$ if

$$\sum_{n=-\infty}^{\infty} |x(n)| |z|^{-n} < \infty \tag{II.7}$$

i. e., the region of convergence consists of all values of z such that the inequality in (II.7) holds. Thus, if some value $z = z_1$ is in the ROC, then all values of z on the circle defined by $|z| = |z_1|$ will also be in the ROC. As one consequence of this, the region of convergence will consist of a ring in the z-plane centered about the origin. Its outer boundary will be a circle (or the ROC may extend outward to infinity), and its inner boundary will be a circle (or it may extend inward to include the origin).

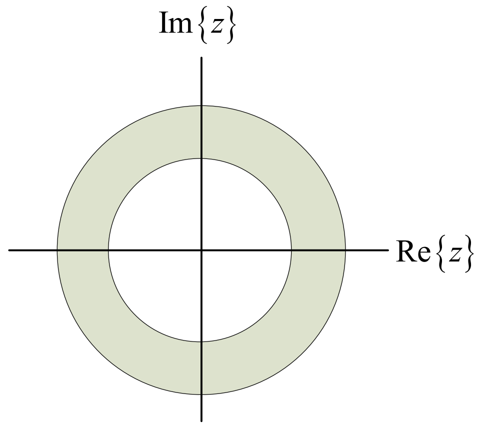


Figure II.2: Ring region of convergence

If the ROC includes the unit circle, this implies convergence of the z-transform for $|z| = 1$, or equivalently, the Fourier transform of the sequence converges. Conversely, if the ROC does not include the unit circle, the Fourier transform does not converge absolutely.

Among the most important and useful z-transforms are those for which $X(z)$ is a rational function inside the region of convergence, i.e.,

$$X(z) = \frac{P(z)}{Q(z)} \tag{II.8}$$

where $P(z)$ and $Q(z)$ are polynomials in z. The values of z for which $X(z) = 0$ are called the zeros of $X(z)$, and the values of z for which $X(z)$ is infinite are called the poles of $X(z)$. The poles of $X(z)$ for finite values of z are the roots of the denominator polynomial. In addition, poles may occur at $z = 0$ or $z = \infty$. For rational z-transforms, a number of important relationships exist between the locations of poles of $X(z)$ and the region of convergence of the z-transform.

Example II-1

Let $x(n) = a^n u(n)$. This is a right-sided sequence.

$$X(z) = \sum_{n=0}^{\infty} a^n z^{-n}$$
$$= \sum_{n=0}^{\infty} \left(az^{-1}\right)^n$$

For convergence, we must have

$$\left|az^{-1}\right| < 1$$

which means that

$$|z| > |a|$$

When the sequence is convergent,

$$X(z) = \frac{1}{1 - az^{-1}}$$

The DTFT is evaluated when $|z| = 1$. Therefore, the DTFT sum converges when $|a| < 1$. Note that

$$X(z) = \frac{z}{z - a}$$

Therefore, there is a pole at $z = a$ and a zero at $z = 0$.

Exercise II-1

Repeat the example above for the left-sided sequence $x(n) = -a^n u(-n-1)$.

In the above, the algebraic expressions for $X(z)$ and the corresponding pole-zero plots are identical for the right-sided and left-sided sequences. The only difference is in the region of convergence. This emphasizes the need for specifying both the algebraic expression and the region of convergence for the z-transform of a given sequence. Also, in both cases, the sequences were exponentials and the resulting z-transforms were rational. In fact, $X(z)$ will be rational whenever $x(n)$ is a linear combination of real or complex exponentials.

Exercise II-2

Determine the z-transform and its ROC of the sequence $x(n) = \left(\frac{1}{2}\right)^n u(n) + \left(\frac{-1}{3}\right)^n u(n)$.

Exercise II-3

Determine the z-transform and its ROC of the sequence $x(n) = \delta(n) + \delta(n-3)$.

Exercise II-4

Determine the poles of

- $X(z)=\dfrac{1}{a^N - z^N}$, where $N > 1$ is an integer.
- $X(z)=1+\dfrac{1}{2}z^{-1}-\dfrac{1}{3}z^{-2}$.

II.1.A. Inspection Method

The inspection method consists simply of becoming familiar with, or recognizing "by inspection", certain transform pairs. It is known that

$$a^n u(n) \overset{z}{\longleftrightarrow} \frac{1}{1-az^{-1}}, \quad |z|>|a| \tag{II.9}$$

If we need to find the inverse z-transform of

$$X(z)=\frac{1}{1-\frac{1}{2}z^{-1}}, \quad |z|>\frac{1}{2} \tag{II.10}$$

we would recognize "by inspection" the associated sequence as

$$x(n)=\left(\frac{1}{2}\right)^n u(n) \tag{II.11}$$

II.1.B. Partial Fractions Expansion

When $X(z)$ is a rational function, we can obtain a partial fraction expansion and easily identify the sequences corresponding to the individual terms. Let

$$X(z)=\frac{\sum_{k=0}^{M} b_k z^{-k}}{\sum_{k=0}^{N} a_k z^{-k}} \tag{II.12}$$

An equivalent expression is

$$X(z)=\frac{z^N \sum_{k=0}^{M} b_k z^{M-k}}{z^M \sum_{k=0}^{N} a_k z^{N-k}} \tag{II.13}$$

Equation (II.13) explicitly shows that for such functions, there will be M zeros and N poles at nonzero locations in the z-plane. In addition, there will be either $M-N$ poles at $z=0$ if $M>N$ or $N-M$ zeros at $z=0$ if $N>M$. In other words, z-transforms of the form of (II.12) always have the same number of poles and zeros in the finite z-plane.

It is most convenient to note that $X(z)$ could be expressed in the form

$$X(z)=\frac{b_0}{a_0}\frac{\prod_{k=1}^{M}\left(1-c_k z^{-1}\right)}{\prod_{k=1}^{N}\left(1-d_k z^{-1}\right)} \tag{II.14}$$

where the c_k's are the nonzero zeros of $X(z)$ and the d_k's are the nonzero poles of $X(z)$. If $M<N$ and the poles are all first order, then $X(z)$ can be expressed as

$$X(z)=\sum_{k=1}^{N}\frac{A_k}{1-d_k z^{-1}} \tag{II.15}$$

Multiplying both sides of (II.15) by $\left(1-d_l z^{-1}\right)$, and evaluating for $z=d_l$ shows that the coefficient A_l, can be found from

$$A_l=\left(1-d_l z^{-1}\right)X(z)\Big|_{z=d_l} \tag{II.16}$$

Example II-2

Let

$$X(z)=\frac{1}{\left(1-\frac{1}{4}z^{-1}\right)\left(1-\frac{1}{2}z^{-1}\right)},\quad |z|>\frac{1}{2}$$

$$X(z)=\frac{A_1}{1-\frac{1}{4}z^{-1}}+\frac{A_2}{1-\frac{1}{2}z^{-1}}$$

$$A_1=\left(1-\frac{1}{4}z^{-1}\right)X(z)\Big|_{z=\frac{1}{4}}=-1$$

$$A_2=\left(1-\frac{1}{2}z^{-1}\right)X(z)\Big|_{z=\frac{1}{2}}=2$$

$$X(z)=\frac{-1}{1-\frac{1}{4}z^{-1}}+\frac{2}{1-\frac{1}{2}z^{-1}}$$

Since $x(n)$ is right-sided,

$$x(n) = -\left(\frac{1}{4}\right)^n u(n) + 2\left(\frac{1}{2}\right)^n u(n)$$
$$= \left(2\left(\frac{1}{2}\right)^n - \left(\frac{1}{4}\right)^n\right) u(n)$$

If $M \geq N$, a polynomial must be added to the right-hand side of (II.15), the order of which is $M - N$. Thus, for $M \geq N$, the complete partial fraction expansion would have the form

$$X(z) = \sum_{k=0}^{M-N} B_k z^{-k} + \sum_{k=1}^{N} \frac{A_k}{1 - d_k z^{-1}} \tag{II.17}$$

Exercise II-5

Determine the inverse z-transform of $X(z) = \frac{1 + 2z^{-1} + z^{-2}}{1 - 2z^{-1} + z^{-2}}$, $|z| > 1$. Note that $X(z)$ has repeated poles.

Exercise II-6

Use power series expansion to determine the inverse z-transform of $X(z) = z^2\left(1 + \frac{1}{2}z^{-1}\right)\left(1 - z^{-2}\right)$.

Exercise II-7

Use long division to determine the inverse z-transform of $X(z) = \frac{1}{1 - az^{-1}}$.

II.2. z-Transform Properties

$$x[n] \overset{z}{\longleftrightarrow} X(z), \quad \text{ROC} = R_x$$
$$x_1[n] \overset{z}{\longleftrightarrow} X_1(z), \quad \text{ROC} = R_{x_1}$$
$$x_2[n] \overset{z}{\longleftrightarrow} X_2(z), \quad \text{ROC} = R_{x_2}$$

Linearity

$$ax_1[n] + bx_2[n] \overset{z}{\longleftrightarrow} aX_1(z) + bX_2(z), \quad \text{ROC} \supseteq R_{x_1} \cap R_{x_2} \tag{II.18}$$

Exercise II-8

Determine the ROC of the z-transform of $x[n] = a^n u[n] - a^n u[n - n_0]$.

Time Shifting

$$x[n-n_0] \stackrel{z}{\longleftrightarrow} z^{-n_0} X(z) \qquad \text{(II.19)}$$

As in the case of linearity, the ROC can be changed, since the factor z^{-n_0} can alter the number of poles at $z=0$ or $z=\infty$.

Exercise II-9

Determine the inverse z-transform of $X(z)=\dfrac{1}{z-\frac{1}{4}}, \quad |z|>\dfrac{1}{4}$.

Multiplication by an Exponential Sequence

$$\alpha^n x[n] \stackrel{z}{\longleftrightarrow} X\left(\frac{z}{\alpha}\right), \quad \text{ROC}=|\alpha| R_x \qquad \text{(II.20)}$$

Exercise II-10

Determine the z-transform of $x[n]=r^n \cos(\omega_0 n) u[n]$.

Differentiation

$$n x[n] \stackrel{z}{\longleftrightarrow} -z\frac{d}{dz}X(z), \quad \text{ROC}=R_x \qquad \text{(II.21)}$$

Exercise II-11

Determine the inverse z-transform of $X(z)=\log\left(1+az^{-1}\right), \quad |z|>|a|$.

Conjugation

$$x^*[n] \stackrel{z}{\longleftrightarrow} X^*(z), \quad \text{ROC}=R_x \qquad \text{(II.22)}$$

Time Reversal

$$x^*[-n] \stackrel{z}{\longleftrightarrow} X^*\left(\frac{1}{z^*}\right), \quad \text{ROC}=\frac{1}{R_x} \qquad \text{(II.23)}$$

$$x[-n] \stackrel{z}{\longleftrightarrow} X\left(\frac{1}{z}\right), \quad \text{ROC}=\frac{1}{R_x} \qquad \text{(II.24)}$$

Convolution

$$x_1[n]*x_2[n] \stackrel{z}{\longleftrightarrow} X_1(z)X_2(z), \quad \text{ROC}=\supseteq R_{x_1}\cap R_{x_2} \qquad \text{(II.25)}$$

Exercise II-12

Use the z-transform and the inverse z-transform to determine the output of an LTI system when $h[n] = a^n u[n]$ and $x[n] = u[n]$.

Initial-Value Theorem

If $x[n]$ is zero for $n < 0$ (i.e., if $x[n]$ is causal), then

$$x[0] = \lim_{z \to \infty} X(z) \qquad \text{(II.26)}$$

***

III. Discrete Fourier Transform

For finite-duration sequences, it is possible to develop a Fourier representation, referred to as the discrete Fourier transform (DFT). The DFT is itself a sequence rather than a function of a continuous variable, and it corresponds to samples, equally spaced in frequency, of the Fourier transform of the signal.

In addition to its theoretical importance as a Fourier representation of sequences, the DFT plays a central role in the implementation of a variety of digital signal-processing algorithms. This is because efficient algorithms exist for the computation of the DFT.

III.1. Discrete Fourier Series (DFS)

Consider a sequence $\tilde{x}[n]$ that is periodic with period N. As with continuous-time periodic signals, such a sequence can be represented by a Fourier series corresponding to a sum of harmonically related complex exponential sequences, i.e., complex exponentials with frequencies that are integer multiples of the fundamental frequency $2\pi/N$ that is associated with the periodic sequence $\tilde{x}[n]$.

These periodic complex exponentials are of the form

$$\begin{aligned} e_k[n] &= e^{j\frac{2\pi}{N}kn} \\ &= e_k[n+rN] \end{aligned} \tag{III.1}$$

where k is an integer, and the Fourier series representation then has the form

$$\tilde{x}[n] = \frac{1}{N}\sum_k \tilde{X}[k]e^{j\frac{2\pi}{N}kn} \tag{III.2}$$

The Fourier series representation of a continuous-time periodic signal generally requires infinitely many harmonically related complex exponentials, whereas the Fourier series for any discrete-time signal with period N requires only N harmonically related complex exponentials. To see this, note that the harmonically related complex exponentials $e_k[n]$ are identical for values of k separated by integer multiples of N, i.e.,

$$e_{k+lN}[n] = e_k[n] \tag{III.3}$$

Consequently, the set of N periodic complex exponentials $\{e_k[n]\}_{k=0}^{N-1}$ defines all the distinct periodic complex exponentials with frequencies that are integer multiples of $2\pi/N$. Thus, the Fourier series representation of a periodic sequence $\tilde{x}[n]$ need contain only N of these complex exponentials, and hence, it has the form

$$\tilde{x}[n] = \frac{1}{N}\sum_{k=0}^{N-1} \tilde{X}[k]e^{j\frac{2\pi}{N}kn} \tag{III.4}$$

The Fourier series coefficients $\tilde{X}(k)$ are obtained from $\tilde{x}(n)$ by the relation

$$\tilde{X}(k)=\sum_{n=0}^{N-1}\tilde{x}(n)e^{-j\frac{2\pi}{N}kn} \tag{III.5}$$

Note that the sequence $\tilde{X}(k)$ is periodic with period N, i.e.,

$$\tilde{X}(k+N)=\tilde{X}(k) \tag{III.6}$$

The Fourier series coefficients can be interpreted to be a sequence of finite length, for $k=0,1,\cdots,N-1$, and zero otherwise, or as a periodic sequence defined for all k. An advantage of interpreting the Fourier series coefficients $\tilde{X}(k)$ as a periodic sequence is that there is then a duality between the time and frequency domains for the Fourier series representation of periodic sequences.

Let

$$W_N=e^{-j\frac{2\pi}{N}} \tag{III.7}$$

With this notation, the DFS analysis-synthesis pair is expressed as follows:

$$\tilde{X}(k)=\sum_{n=0}^{N-1}\tilde{x}(n)W_N^{kn} \tag{III.8}$$

$$\tilde{x}(n)=\frac{1}{N}\sum_{k=0}^{N-1}\tilde{X}(k)W_N^{-kn} \tag{III.9}$$

In both of these equations, $\tilde{X}(k)$ and $\tilde{x}(n)$ are periodic sequences. We will sometimes find it convenient to use the notation

$$\tilde{x}(n)\xleftrightarrow{\mathcal{DFS}}\tilde{X}(k) \tag{III.10}$$

Exercise III-1

Determine the DFS of $\tilde{x}(n)=\begin{cases}1, & n=rN,\ r \text{ is any integer}\\ 0, & \text{otherwise}\end{cases}$.

III.2. Properties of The Discrete Fourier Series

Linearity

$$\tilde{x}_1(n)\xleftrightarrow{\mathcal{DFS}}\tilde{X}_1(k) \tag{III.11}$$

$$\tilde{x}_2(n)\xleftrightarrow{\mathcal{DFS}}\tilde{X}_2(k) \tag{III.12}$$

$$a\tilde{x}_1(n)+b\tilde{x}_2(n)\xleftrightarrow{\mathcal{DFS}}a\tilde{X}_1(k)+b\tilde{X}_2(k) \tag{III.13}$$

Shifting

$$\tilde{x}(n-m)\xleftrightarrow{\mathcal{DFS}}W_N^{km}\tilde{X}(k) \tag{III.14}$$

$$W_N^{-nl}\tilde{x}(n)\xleftrightarrow{\mathcal{DFS}}\tilde{X}(k-l) \tag{III.15}$$

Duality

$$\tilde{X}(n)\xleftrightarrow{\mathcal{DFS}}N\tilde{x}(-k) \tag{III.16}$$

Symmetry Properties

$$\tilde{x}^*(n)\xleftrightarrow{\mathcal{DFS}}\tilde{X}^*(-k) \tag{III.17}$$

$$\tilde{x}^*(-n)\xleftrightarrow{\mathcal{DFS}}\tilde{X}^*(k) \tag{III.18}$$

$$\mathrm{Re}\left\{\tilde{x}(n)\right\}\xleftrightarrow{\mathcal{DFS}}\tilde{X}_e(k)=\frac{1}{2}\left[\tilde{X}(k)+\tilde{X}^*(-k)\right] \tag{III.19}$$

$$j\,\mathrm{Im}\left\{\tilde{x}(n)\right\}\xleftrightarrow{\mathcal{DFS}}\tilde{X}_o(k)=\frac{1}{2}\left[\tilde{X}(k)-\tilde{X}^*(-k)\right] \tag{III.20}$$

$$\tilde{x}_e(n)=\frac{1}{2}\left[\tilde{x}(n)+\tilde{x}^*(-n)\right]\xleftrightarrow{\mathcal{DFS}}\mathrm{Re}\left\{\tilde{X}(k)\right\} \tag{III.21}$$

$$\tilde{x}_o(n)=\frac{1}{2}\left[\tilde{x}(n)-\tilde{x}^*(-n)\right]\xleftrightarrow{\mathcal{DFS}}j\,\mathrm{Im}\left\{\tilde{X}(k)\right\} \tag{III.22}$$

When $\tilde{x}(n)$ is real:

$$\tilde{X}(k)=\tilde{X}^*(-k) \tag{III.23}$$

$$\tilde{x}_e(n)=\frac{1}{2}\left[\tilde{x}(n)+\tilde{x}(-n)\right]\xleftrightarrow{\mathcal{DFS}}\mathrm{Re}\left\{\tilde{X}(k)\right\} \tag{III.24}$$

$$\tilde{x}_o(n)=\frac{1}{2}\left[\tilde{x}(n)-\tilde{x}(-n)\right]\xleftrightarrow{\mathcal{DFS}}j\,\mathrm{Im}\left\{\tilde{X}(k)\right\} \tag{III.25}$$

Periodic Convolution

If

$$\tilde{X}_3(k)=\tilde{X}_1(k)\tilde{X}_2(k) \tag{III.26}$$

then the periodic sequence $\tilde{x}_3(n)$ with Fourier series coefficients $\tilde{X}_3(k)$ is

$$\tilde{x}_3(n) = \sum_{m=0}^{N-1} \tilde{x}_1(m)\tilde{x}_2(n-m) \tag{III.27}$$

The sequences in (III.27) are all periodic with period N, and the summation is over only one period. A convolution in the form of (III.27) is referred to as a periodic convolution. Just as with aperiodic convolution, periodic convolution is commutative; i.e.,

$$\tilde{x}_3(n) = \sum_{m=0}^{N-1} \tilde{x}_2(m)\tilde{x}_1(n-m) \tag{III.28}$$

If

$$\tilde{x}_3(n) = \tilde{x}_1(n)\tilde{x}_2(n) \tag{III.29}$$

then,

$$\tilde{X}_3(k) = \frac{1}{N}\sum_{l=0}^{N-1} \tilde{X}_1(l)\tilde{X}_2(k-l) \tag{III.30}$$

III.3. Fourier Transform of Periodic Signals

As discussed earlier, uniform convergence of the Fourier transform of a sequence requires that the sequence be absolutely summable, and mean-square convergence requires that the sequence be square summable. Periodic sequences satisfy neither condition, because they do not approach zero as time approaches $\pm\infty$. However, sequences that can be expressed as a sum of complex exponentials can be considered to have a Fourier transform representation in the form of a train of impulses. Similarly, it is often useful to incorporate the discrete Fourier series representation of periodic signals within the framework of the Fourier transform. This can be done by interpreting the Fourier transform of a periodic signal to be an impulse train in the frequency domain with the impulse values proportional to the DFS coefficients for the sequence. Specifically, if $\tilde{x}(n)$ is periodic with period N and the corresponding discrete Fourier series coefficients are $\tilde{X}(k)$, then the Fourier transform of $\tilde{x}(n)$ is defined to be the impulse train

$$\tilde{X}(e^{j\omega}) = \frac{2\pi}{N}\sum_{k=-\infty}^{\infty} \tilde{X}(k)\delta\left(\omega - \frac{2\pi k}{N}\right) \tag{III.31}$$

Exercise III-2

Show that $\tilde{X}(e^{j\omega})$ as defined above is periodic with period 2π.

Exercise III-3

1. Use the inverse DTFT expression to show that $\tilde{X}(e^{j\omega})$ is a Fourier transform of $\tilde{x}(n)$.
2. Determine the DTFT of $\tilde{p}(n) = \sum_{m=-\infty}^{\infty} \delta(n-mN)$.

Consider a finite-length signal $x(n)$ such that $x(n)=0$ except in the interval $0 \le n < N$, and consider the convolution of $x(n)$ with the periodic impulse train $\tilde{p}(n)$:

$$\begin{aligned}\tilde{x}(n) &= x(n) * \tilde{p}(n) \\ &= x(n) * \sum_{m=-\infty}^{\infty} \delta(n-mN) \\ &= \sum_{m=-\infty}^{\infty} x(n-mN)\end{aligned} \tag{III.32}$$

Note that $\tilde{x}(n)$ is composed of a set of periodically repeated copies of the finite-length sequence $x(n)$. Therefore,

$$\begin{aligned}\tilde{X}(e^{j\omega}) &= X(e^{j\omega})\tilde{P}(e^{j\omega}) \\ &= X(e^{j\omega})\frac{2\pi}{N}\sum_{k=-\infty}^{\infty}\delta\left(\omega-\frac{2\pi}{N}k\right) \\ &= \frac{2\pi}{N}\sum_{k=-\infty}^{\infty} X\left(e^{j\frac{2\pi}{N}k}\right)\delta\left(\omega-\frac{2\pi}{N}k\right)\end{aligned} \tag{III.33}$$

Comparing (III.33) with (III.31) yields

$$\begin{aligned}\tilde{X}(k) &= X\left(e^{j\frac{2\pi}{N}k}\right) \\ &= X(e^{j\omega})\Big|_{\omega=\frac{2\pi}{N}k}\end{aligned} \tag{III.34}$$

In other words, the periodic sequence $\tilde{X}(k)$ of DFS coefficients has an interpretation as equally spaced samples of the Fourier transform of the finite-length sequence obtained by extracting one period of $\tilde{x}(n)$; i.e.,

$$x(n) = \begin{cases}\tilde{x}(n), & 0 \le n < N \\ 0, & \text{otherwise}\end{cases} \tag{III.35}$$

III.4. Sampling the Fourier Transform

Consider an aperiodic sequence $x(n)$ with Fourier transform $X(e^{j\omega})$, and assume that a sequence $\tilde{X}(k)$ is obtained by sampling $X(e^{j\omega})$ at frequencies $\omega_k = 2\pi k/N$, i.e.,

$$\tilde{X}(k) = X(e^{j\omega})\Big|_{\omega=\frac{2\pi}{N}k} \tag{III.36}$$

We make no assumptions about $x(n)$ other than that its Fourier transform exists, i.e.,

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} \tag{III.37}$$

Therefore,

$$\tilde{X}(k) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\frac{2\pi}{N}kn} \tag{III.38}$$

Since the Fourier transform is periodic in ω with period 2π, the resulting sequence is periodic in k with period N. Also, since the Fourier transform is equal to the z-transform evaluated on the unit circle, it follows that $\tilde{X}(k)$ can also be obtained by sampling $X(z)$ at N equally spaced points on the unit circle. Thus,

$$\tilde{X}(k) = X(z)\Big|_{z=e^{j\frac{2\pi}{N}k}} \tag{III.39}$$

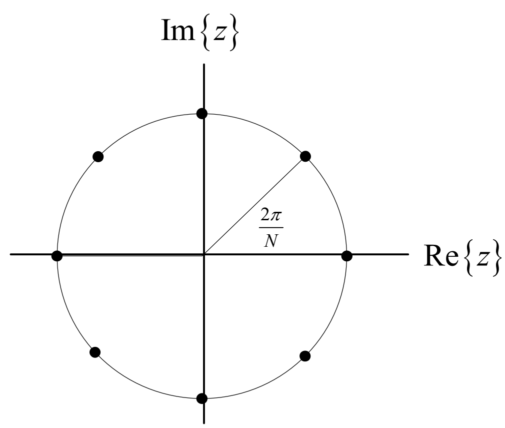


Figure III.1

The above figure makes it clear that the sequence of samples is periodic, since the N points are equally spaced starting with zero angle. Therefore, the same sequence repeats as k varies outside the range $0 \le k < N$.

Note that the sequence of samples $\tilde{X}(k)$, being periodic with period N, could be the sequence of discrete Fourier series coefficients of a sequence $\tilde{x}(n)$, given by

$$\tilde{x}(n) = \frac{1}{N}\sum_{k=0}^{N-1} \tilde{X}(k)W_N^{-kn} \tag{III.40}$$

Substituting (III.37) into (III.36) and then substituting the resulting expression for $\tilde{X}(k)$ into (III.40) gives

$$\begin{aligned}\tilde{x}(n) &= \sum_{m=-\infty}^{\infty} x(m)\tilde{p}(n-m) \\ &= \sum_{r=-\infty}^{\infty} x(n-rN)\end{aligned} \tag{III.41}$$

The periodic sequence $\tilde{x}(n)$, corresponding to $\tilde{X}(k)$ obtained by sampling $X(e^{j\omega})$, is formed from $x(n)$ by adding together an infinite number of shifted replicas of $x(n)$. The shifts are all the positive and negative integer multiples of N, the period of the sequence $\tilde{X}(k)$.

III.5. The Discrete Fourier Transform

To each finite-length sequence of length N. we can always associate a periodic sequence

$$\tilde{x}(n) = \sum_{r=-\infty}^{\infty} x(n-rN) \tag{III.42}$$

The finite-length sequence $x(n)$ can be recovered from $\tilde{x}(n)$ through

$$x(n) = \begin{cases} \tilde{x}(n), & 0 \le n < N \\ 0, & \text{otherwise} \end{cases} \tag{III.43}$$

The DFS coefficients of $\tilde{x}(n)$ are samples (spaced in frequency by $2\pi/N$) of the Fourier transform of $x(n)$. Since $x(n)$ is assumed to have finite length N, there is no overlap between the terms $x(n-rN)$ for different values of r. Thus,

$$\begin{aligned}\tilde{x}(n) &= x\left(n \bmod N\right) \\ &= x\left((n)_N\right)\end{aligned} \tag{III.44}$$

Note that (III.44) is equivalent to (III.42) only when $x(n)$ has length less than or equal to N. The finite-duration sequence $x(n)$ is obtained from $\tilde{x}(n)$ by extracting one period, as in (III.43).

The sequence of discrete Fourier series coefficients $\tilde{X}(k)$ of the periodic sequence $\tilde{x}(n)$ is itself a periodic sequence with period N. To maintain a duality between the time and frequency domains, we will choose the Fourier coefficients that we associate with a finite-duration sequence to be a finite-duration sequence corresponding to one period of $\tilde{X}(k)$. This finite-duration sequence $X(k)$ will be referred to as the discrete Fourier transform (DFT). Thus, the DFT $X(k)$ is related to the DFS coefficients $\tilde{X}(k)$ by

$$X(k) = \begin{cases} \tilde{X}(k), & 0 \le k < N \\ 0, & \text{otherwise} \end{cases} \tag{III.45}$$

and

$$\begin{aligned}\tilde{X}(k) &= X\left(k \bmod N\right) \\ &= X\left((k)_N\right)\end{aligned} \tag{III.46}$$

Therefore,

$$X(k) = \begin{cases} \sum_{n=0}^{N-1} x(n) W_N^{kn}, & 0 \le k < N \\ 0, & \text{otherwise} \end{cases} \tag{III.47}$$

$$x(n) = \begin{cases} \sum_{k=0}^{N-1} X(k) W_N^{-kn}, & 0 \le n < N \\ 0, & \text{otherwise} \end{cases} \tag{III.48}$$

The relationship between $x(n)$ and $X(k)$ will sometimes be denoted as

$$x(n) \xleftrightarrow{\mathcal{DFT}} X(k) \tag{III.49}$$

Exercise III-4

1. Determine the DFT of a rectangular pulse with duration N.
2. Determine the DFT of $x(n) = \begin{cases} 1, & 0 \le n < \frac{N}{2} \\ 0, & \frac{N}{2} \le n < N \end{cases}$.

III.6. Properties of the DFT

Linearity

If

$$\begin{aligned} x_1(n) &\xleftrightarrow{\mathcal{DFT}} X_1(k) \\ x_2(n) &\xleftrightarrow{\mathcal{DFT}} X_2(k) \end{aligned} \tag{III.50}$$

then

$$x_3(n) = a_1 x_1(n) + a_2 x_2(n) \xleftrightarrow{\mathcal{DFT}} a_1 X_1(k) + a_2 X_2(k) = X_3(k) \tag{III.51}$$

Clearly, if $x_1(n)$ has length N_1 and $x_2(n)$ has length N_2, then the maximum length of $x_3(n)$ will be $N_3 = \max\left(N_1, N_2\right)$. Thus, in order for (III.51) to be meaningful, both DFTs must be computed with the same length $N \ge N_3$. If, for example, $N_1 < N_2$, then $X_1(k)$ is the DFT of the sequence $x_1(n)$ augmented by $N_2 - N_1$ zeros. That is, the N_2-point DFT of $x_1(n)$ is

$$X_1(k) = \sum_{n=0}^{N_1-1} x_1(n) W_{N_2}^{kn}, \quad 0 \le k < N_2 \tag{III.52}$$

and the N_2-point DFT of $x_2(n)$ is

$$X_2(k) = \sum_{n=0}^{N_2-1} x_2(n) W_{N_2}^{kn}, \quad 0 \le k < N_2 \tag{III.53}$$

Circular Shift of a Sequence

If

$$x_1(n) = \begin{cases} x\big((n-m)_N\big), & 0 \le n < N \\ 0, & \text{otherwise} \end{cases} \tag{III.54}$$

then

$$X_1(k) = e^{-j\frac{2\pi}{N}km} X(k) \tag{III.55}$$

Duality

$$X(n) \xleftrightarrow{\mathcal{DFT}} X\big((-k)_N\big), \quad 0 \le k < N \tag{III.56}$$

Symmetry Properties

$$x^*(n) \xleftrightarrow{\mathcal{DFT}} X^*\big((-k)_N\big), \quad 0 \le k < N \tag{III.57}$$

$$x^*\big((-n)_N\big) \xleftrightarrow{\mathcal{DFT}} X^*(k), \quad 0 \le k < N \tag{III.58}$$

The even and odd parts of $x(n)$ are given by

$$x_{ep}(n) = \frac{1}{2}\Big(x\big((n)_N\big) + x^*\big((-n)_N\big)\Big), \quad 0 \le n < N \tag{III.59}$$

$$x_{op}(n) = \frac{1}{2}\Big(x\big((n)_N\big) - x^*\big((-n)_N\big)\Big), \quad 0 \le n < N \tag{III.60}$$

Note that for $0 \le n < N$,

$$\begin{aligned} (n)_N &= n \\ (-n)_N &= N - n \end{aligned} \tag{III.61}$$

Therefore,

$$x_{ep}(n) = \frac{1}{2}\Big(x(n) + x^*(N-n)\Big), \quad 0 \le n < N \tag{III.62}$$

$$x_{op}(n) = \frac{1}{2}\left(x(n) - x^*(N-n)\right), \quad 0 \le n < N \tag{III.63}$$

$$x_{ep}(0) = \mathrm{Re}\{x(0)\} \tag{III.64}$$

$$x_{op}(0) = j\,\mathrm{Im}\{x(0)\} \tag{III.65}$$

Exercise III-5

Show that for $0 \le n < N$,

$$x_{ep}(n) = x_e(n) + x_e(n-N)$$
$$x_{op}(n) = x_o(n) + x_o(n-N)$$

The sequences $x_{ep}(n)$ and $x_{op}(n)$ will be referred to as the periodic conjugate-symmetric and periodic conjugate-antisymmetric components, respectively, of $x(n)$. When $x_{ep}(n)$ and $x_{op}(n)$ are real, they will be referred to as the periodic even and periodic odd components, respectively.

Exercise III-6

Show that $x_{ep}(n)$ and $x_{op}(n)$ are not periodic sequences. Show that, instead, they are finite-length sequences that are equal to one period of the periodic sequences $\tilde{x}_e(n)$ and $\tilde{x}_o(n)$, respectively.

Note that

$$x(n) = x_{ep}(n) + x_{op}(n) \tag{III.66}$$

The symmetry properties of the DFT now follow in a straightforward way:

$$\mathrm{Re}\{x(n)\} \xleftrightarrow{\mathcal{DFT}} X_{ep}(k) \tag{III.67}$$

$$j\,\mathrm{Im}\{x(n)\} \xleftrightarrow{\mathcal{DFT}} X_{op}(k) \tag{III.68}$$

$$x_{ep}(n) \xleftrightarrow{\mathcal{DFT}} \mathrm{Re}\{X(k)\} \tag{III.69}$$

$$x_{op}(n) \xleftrightarrow{\mathcal{DFT}} j\,\mathrm{Im}\{X(k)\} \tag{III.70}$$

Circular Convolution

Consider two finite-duration sequences $x_1(n)$ and $x_2(n)$, both of length N, with DFTs $X_1(k)$ and $X_2(k)$, respectively. Let the sequence $x_3(n)$ have the DFT $X_3(k) = X_1(k)X_2(k)$. Note that $X_3(k)$ is the first period of $\tilde{X}_3(k) = \tilde{X}_1(k)\tilde{X}_2(k)$. According to (III.27),

$$\tilde{x}_3(n) = \sum_{m=0}^{N-1} \tilde{x}_1(m)\tilde{x}_2(n-m) \quad \text{(III.71)}$$

The sequence $x_3(n)$ corresponds to one period of $\tilde{x}_3(n)$, and is given by

$$x_3(n) = \sum_{m=0}^{N-1} \tilde{x}_1(m)\tilde{x}_2(n-m), \quad 0 \le n < N \quad \text{(III.72)}$$

Equivalently,

$$x_3(n) = \sum_{m=0}^{N-1} x_1\left((m)_N\right) x_2\left((n-m)_N\right), \quad 0 \le n < N \quad \text{(III.73)}$$

Since $(m)_N = m$ for $0 \le m < N$, (III.73) can be written as

$$x_3(n) = \sum_{m=0}^{N-1} x_1(m) x_2\left((n-m)_N\right), \quad 0 \le n < N \quad \text{(III.74)}$$

This is known as circular convolution, and is usually denoted as

$$\begin{aligned} x_3(n) &= x_1(n) *_{(N)} x_2(n) \\ &= x_2(n) *_{(N)} x_1(n) \end{aligned} \quad \text{(III.75)}$$

Exercise III-7

Let $x_1(n)$ and $x_2(n)$ be two unit rectangular sequences of duration L. Determine $x_3(n) = x_1(n) *_{(N)} x_2(n)$ when

1. $N = L$
2. $N = 2L$

In view of the duality of the DFT relations, it is not surprising that the DFT of a product of two N point sequences is the circular convolution of their respective discrete Fourier transforms. Specifically, if $x_3(n) = x_1(n)x_2(n)$, then

$$X_3(k) = \frac{1}{N} \sum_{l=0}^{N-1} X_1(l) X_2\left((k-l)_N\right) \quad \text{(III.76)}$$

or

$$x_1(n)x_2(n) \xleftrightarrow{\mathcal{DFT}} \frac{1}{N} X_1(k) *_{(N)} X_2(k) \quad \text{(III.77)}$$

Assignment III.1

Linear Convolution Using the Discrete Fourier Transform.

***

IV. LINEAR OPTIMUM FILTERING

IV.1. Introduction

Consider the following block diagram

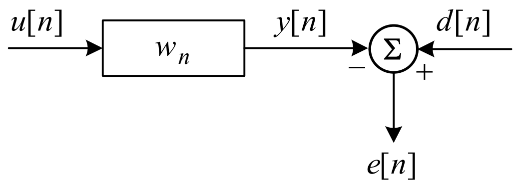


Figure IV.1

This discrete-time filter output is used to provide an estimate of some desired response. The filter input and the desired response represent single realizations of respective stochastic processes. The estimation is accompanied by an error with statistical characteristics of its own. In particular, the estimation error, denoted by $e[n]$, is defined as the difference between the desired response $d[n]$ and the filter output $y[n]$.

The requirement is to make the estimation error $e[n]$ "as small as possible" in some statistical sense. Two restrictions are placed on the filter:

- The filter is linear, which makes the mathematical analysis easy to handle.
- The filter operates in discrete time, which makes it possible for the filter to be implemented using digital hardware/software.

The final details of the filter specifications, however, depend on two other choices that have to be made:

- Whether the impulse response of the filter has finite or infinite duration.
- The type of statistical criterion used for the optimization.

The choice of a finite-duration impulse response (FIR) or an infinite-duration impulse response (IIR) for the filter is dictated by practical considerations. The choice of a statistical criterion for optimizing the filter design is influenced by mathematical tractability.

We will confine our attention to the use of FIR filters. An FIR filter is inherently stable, because its structure involves the use of forward paths only. Indeed, it is this form of signal transmission through the filter that limits its impulse response to a finite duration.

On the other hand, an IIR filter involves both feedforward and feedback paths. The presence of feedback means that portions of the filter output and possibly other internal variables in the filter are fed back to the input. Consequently, unless it is properly designed, feedback in the filter can indeed make it unstable. This kind of operation is clearly unacceptable when the requirement is that of filtering for which stability is a "must".

By itself, the stability problem in IIR filters is manageable in both theoretical and practical terms. However, when the filter is required to be adaptive, bringing with it stability problems of its own, the inclusion of adaptivity combined with feedback that is inherently present in an IlR filter creates a problem that is much more difficult to handle. It is for this reason that we find that in the majority of applications requiring the use of adaptivity, the use of an FIR filter is preferred over an IIR filler even though the latter is less demanding in computational requirements.

Turning next to the issue of what criterion to choose for statistical optimization, there are indeed several criteria that suggest themselves. Specifically, we may consider optimizing the filter design by minimizing a cost function, or index of performance, selected from the following short list of possibilities:

- Mean-square value of the estimation error
- Expectation of the absolute value of the estimation error
- Expectation of third or higher powers of the absolute value of the estimation error

The first option has a clear advantage over the other two, because it leads to tractable mathematics. In particular, the choice of the mean-square error criterion results in a second-order dependence for the cost function on the unknown coefficients in the impulse response of the filter. Moreover, the cost function has a distinct minimum that uniquely defines the optimum statistical design of the filter.

We may now summarize the essence of the filtering problem by making the following statement:

Design a linear discrete-time filter, with an impulse response w_n, that uses an input sequence $u[n]$ to produce an output sequence $y[n]$ that is meant as an estimate of a desired response $d[n]$, such that the mean-square value of the estimation error $e[n]$, defined as the difference between the desired response $d[n]$ and the actual response $y[n]$, is minimized.

We may develop the mathematical solution to this statistical optimization problem by following two entirely different approaches that are complementary. One approach leads to the development of an important theorem commonly known as the principle of orthogonality. The other approach highlights the error performance surface that describes the second-order dependence of the cost function on the filter coefficients. We will proceed by deriving the principle of orthogonality first, because the derivation is relatively simple and because the principle of orthogonality is highly insightful.

IV.2. Principle of Orthogonality

Consider a discrete-time linear time-invariant filter whose input is denoted by $u[n]$. Let w_n be the filter impulse response sequence. Let both $u[n]$ and w_n have complex values and infinite duration. The filter output is defined by the linear convolution sum:

$$y[n] = \sum_{k=0}^{\infty} w_k^* u[n-k] \tag{IV.1}$$

Note that in complex terminology, the term $w_k^* u[n-k]$ represents the scalar version of an inner product of the filter coefficient w_k and the filter input $u[n-k]$. The purpose of the filter is to produce an estimate of the desired response $d[n]$. We assume that the filter input and the desired response are single realizations of jointly wide-sense stationary stochastic processes, both with zero mean. Accordingly, the estimation of $d[n]$ is accompanied by an error defined by the difference

$$e[n] = d[n] - y[n] \tag{IV.2}$$

The estimation error $e[n]$ is the sample value of a random variable. To optimize the filter design, we choose to minimize the mean-square value of the estimation error $e[n]$. We may thus define the cost function as the mean-squared error

$$\begin{aligned} J &= \mathrm{E}\left[e[n]e^*[n]\right] \\ &= \mathrm{E}\left[\left|e[n]\right|^2\right] \end{aligned} \tag{IV.3}$$

The problem is therefore to determine the operating conditions for which J attains its minimum value. For complex input data, the filter coefficients are in general complex, too. Let the k th filter coefficient w_k be denoted in terms of its real and imaginary parts as follows:

$$w_k = a_k + jb_k \tag{IV.4}$$

Let us define the vector gradient operator $\underline{\nabla}$, the k th element of which is

$$\nabla_k = \frac{\partial}{\partial a_k} + j\frac{\partial}{\partial b_k} \tag{IV.5}$$

Applying the operator ∇_k to the cost function J gives

$$\nabla_k J = \frac{\partial J}{\partial a_k} + j\frac{\partial J}{\partial b_k} \tag{IV.6}$$

Note that for the definition of the complex gradient given in (IV.6) to be valid, it is essential that J be real. For the cost function J to attain its minimum value, all elements of the gradient vector $\underline{\nabla}J$ must be simultaneously equal to zero, as shown by

$$\nabla_k J = 0, \quad k = 0, 1, \cdots \tag{IV.7}$$

Under this set of conditions, the filter is said to be optimum in the mean-squared-error sense. According to (IV.3), the cost function is a scalar that is independent of time n. Hence, substituting the first line of (IV.3) in (IV.6), and using (IV.2) and (IV.4), we get the following partial derivatives:

$$\begin{aligned}\frac{\partial e[n]}{\partial a_k} &= -u[n-k] \\ \frac{\partial e[n]}{\partial b_k} &= ju[n-k] \\ \frac{\partial e^*[n]}{\partial a_k} &= -u^*[n-k] \\ \frac{\partial e^*[n]}{\partial b_k} &= -ju^*[n-k]\end{aligned} \qquad \text{(IV.8)}$$

Thus, substituting these partial derivatives in (IV.7) and then canceling common terms we finally get the result

$$\nabla_k J = -2\,\mathrm{E}\left[u[n-k]e^*[n]\right] \qquad \text{(IV.9)}$$

Let $e_o[n]$ denote the special value of the estimation error that results when the filter operates in its optimum condition. In other words,

$$\mathrm{E}\left[u[n-k]e_o^*[n]\right] = 0, \quad k = 0,1,\cdots \qquad \text{(IV.10)}$$

The last result can be stated as follows: The necessary and sufficient condition for the cost function J to attain its minimum value is that the corresponding value of the estimation error $e_o[n]$ is orthogonal to each input sample that enters into the estimation of the desired response at time n. Indeed, this statement constitutes the principle of orthogonality; it represents one of the most elegant theorems in the subject of linear optimum filtering. It also provides the mathematical basis of a procedure for testing that the linear filter is operating in its optimum condition.

Consider the following linear system, in which $\underline{x}$ is an $n\times1$ vector and $\underline{b}$ is an $m\times1$ vector

$$A\underline{x} = \underline{b} \qquad \text{(IV.11)}$$

where the $m\times n$ matrix A can be written in terms of its $m\times1$ column vectors in the form

$$A = \left[\underline{a}_1 \quad \underline{a}_2 \quad \cdots \quad \underline{a}_n\right] \qquad \text{(IV.12)}$$

It is well-known that if $\underline{b}$ is not in the column space of A then there is no solution to the system in (IV.11). Since the column space is composed of all linear combinations of $\{\underline{a}_i\}_{i=1}^n$, then $\underline{b}$ can be replaced by its projection onto the column space, and in that case there will be a solution. Projection is illustrated by the simple example in Figure IV.2.

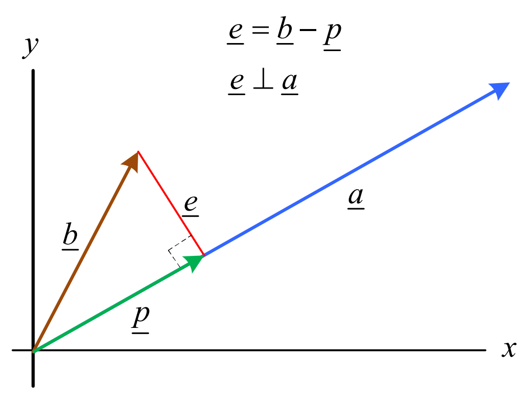


Figure IV.2: Projection and orthogonality

Corollary to the Principle of Orthogonality

Note that

$$\begin{aligned} \mathrm{E}\left[y[n]e^*[n] \right] &= \mathrm{E}\left[\sum_{k=0}^{\infty} w_k^* u[n-k] e^*[n] \right] \\ &= \sum_{k=0}^{\infty} w_k^* \, \mathrm{E}\left[u[n-k] e^*[n] \right] \end{aligned} \quad \text{(IV.13)}$$

Let $y_o[n]$ denote the output produced by the filter optimized in the mean-squared-error sense, with $e_o[n]$ denoting the corresponding estimation error. Using the principle of orthogonality, we get the desired result

$$\mathrm{E}\left[y_o[n] e_o^*[n] \right] = 0 \quad \text{(IV.14)}$$

We may thus state the corollary to the principle of orthogonality as follows: When the filter operates in its optimum condition, the estimate of the desired response defined by the filter output $y_o[n]$ and the corresponding estimation error $e_o[n]$ are orthogonal to each other.

IV.3. Minimum Mean-Squared Error

When the linear discrete-time filter operates in its optimum condition, (IV.2) takes on the following special form

$$e_o[n] = d[n] - y_o[n] \quad \text{(IV.15)}$$

Rearranging terms we get

$$d[n] = y_o[n] + e_o[n] \quad \text{(IV.16)}$$

Let $J_{\min}$ denote the minimum mean-squared error, defined by

$$J_{\min} = \mathrm{E}\left[\left|e_o[n]\right|^2\right] \tag{IV.17}$$

Hence, evaluating the mean-square values of both sides of (IV.16), and applying to it the corollary to the principle of orthogonality described by (IV.14), we get

$$\sigma_d^2 = \sigma_{\hat{d}}^2 + J_{\min} \tag{IV.18}$$

Solving (IV.18) for the minimum mean-squared error, we get

$$J_{\min} = \sigma_d^2 - \sigma_{\hat{d}}^2 \tag{IV.19}$$

This relation shows that for the optimum filter, the minimum mean-squared error equals the difference between the variance of the desired response and the variance of the estimate that the filter produces at its output.

It is convenient to normalize the expression in (IV.19) in such a way that the minimum value of the mean-squared error always lies between zero and one. We may do this by dividing both sides of (IV.19) by σ_d^2, obtaining

$$\frac{J_{\min}}{\sigma_d^2} = 1 - \frac{\sigma_{\hat{d}}^2}{\sigma_d^2} \tag{IV.20}$$

Let

$$\varepsilon = \frac{J_{\min}}{\sigma_d^2} \tag{IV.21}$$

The quantity ε is called the normalized mean-squared error, in terms of which we may rewrite (IV.20) in the form

$$\varepsilon = 1 - \frac{\sigma_{\hat{d}}^2}{\sigma_d^2} \tag{IV.22}$$

Exercise IV-1

Show that $0 \le \varepsilon \le 1$.

IV.4. Wiener-Hopf Equations

The principle of orthogonality specifies the necessary and sufficient condition for the optimum operation of the filter. We may reformulate the necessary and sufficient condition for optimality by substituting (IV.1) and (IV.2) in the orthogonality principle equation (IV.10). In particular, we may write

$$\mathrm{E}\left[u[n-k]\left(d^*[n] - \sum_{i=0}^{\infty} w_{oi} u^*[n-i]\right)\right] = 0 \tag{IV.23}$$

where w_{oi} is the i th coefficient in the impulse response of the optimum filter. Expanding this equation and rearranging terms, we get

$$\sum_{i=0}^{\infty} w_{oi} \, \mathrm{E}\left[u[n-k]u^*[n-i]\right] = \mathrm{E}\left[u[n-k]d^*[n]\right] \tag{IV.24}$$

The two expectations in (IV.24) may be interpreted as follows:

- The expectation $\mathrm{E}\left[u[n-k]u^*[n-i]\right]$ is equal to the autocorrelation function of the filter input for a lag of $i-k$. We may thus express this expectation as

$$r[i-k] = \mathrm{E}\left[u[n-k]u^*[n-i]\right] \tag{IV.25}$$

- The expectation $\mathrm{E}\left[u[n-k]d^*[n]\right]$ is equal to the cross-correlation between the filter input $u[n-k]$ and the desired response $d[n]$ for a lag of $-k$. We may thus express this second expectation as

$$p[-k] = \mathrm{E}\left[u[n-k]d^*[n]\right] \tag{IV.26}$$

Accordingly,

$$\sum_{i=0}^{\infty} w_{oi} r[i-k] = p[-k] \tag{IV.27}$$

The system of equations (IV.27) defines the optimum filter coefficients, in the most general setting, in terms of two correlation functions: the autocorrelation function of the filter input, and the cross-correlation between the filter input and the desired response. These equations are called the Wiener-Hopf equations.

It should also be noted that the defining equation for a linear optimum filter was formulated originally by Wiener and Hopf (1931) for the case of a continuous-time filter, whereas, of course the system of (IV.27) is formulated for a discrete-time filter.

IV.5. Solution of the Wiener-Hopf Equations for Linear Transversal Filters

The solution of the set of Wiener-Hopf equations is greatly simplified for the special case when a linear transversal filter, or FIR filter, is used to perform the estimation of desired response $d(n)$.

The impulse response of the transversal filter is defined by the finite set of tap weights $w_0, w_1, \cdots, w_{M-1}$. Accordingly, the Wiener-Hopf equations reduce to a system of M simultaneous equations, as shown by

$$\sum_{i=0}^{M-1} w_{oi} r(i-k) = p(-k), \quad k = 0,1,\cdots,M-1 \tag{IV.28}$$

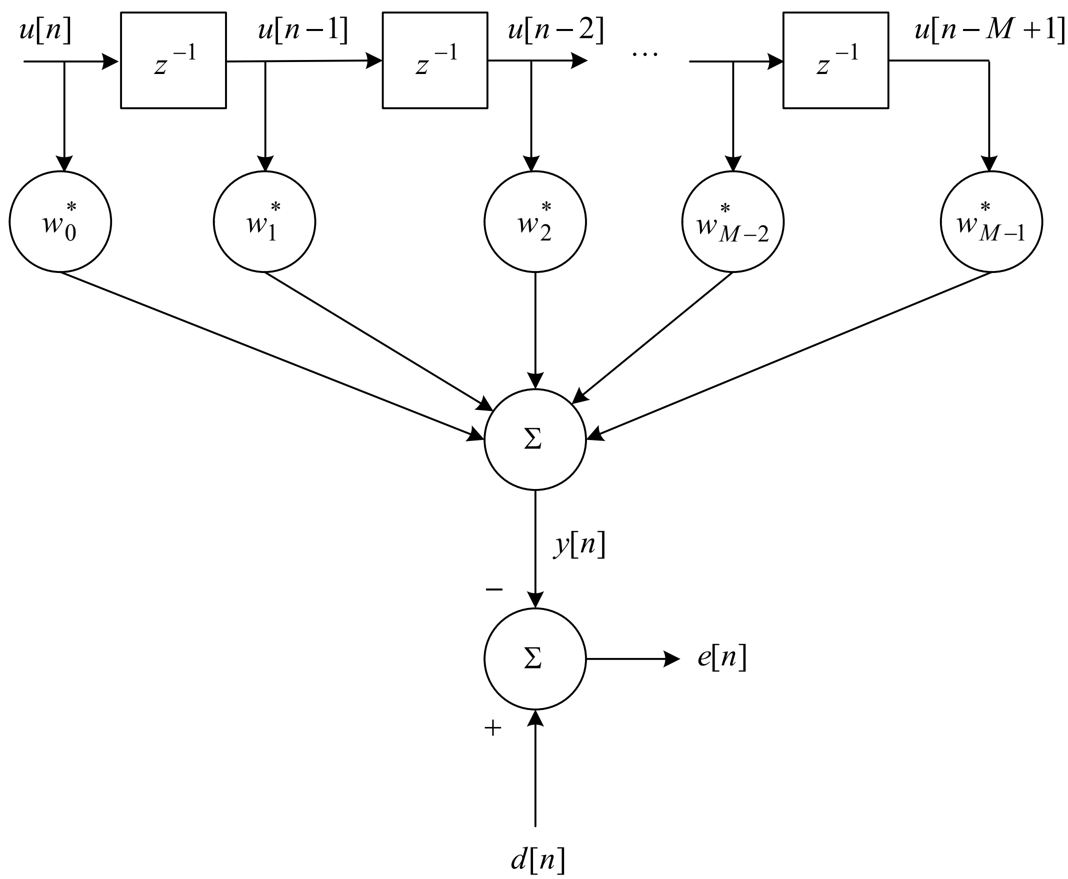


Figure IV.3

Let R denote the $M \times M$ correlation matrix of the tap inputs $u(n), u(n-1), \cdots, u(n-M+1)$ in the transversal filter, i.e.,

$$R = \mathrm{E}\left[\underline{u}(n)\underline{u}^{H}(n)\right] \tag{IV.29}$$

where $\underline{u}(n)$ is the $M \times 1$ tap-input vector, defined as

$$\underline{u}(n) = \left[u(n) \quad u(n-1) \quad \cdots \quad u(n-M+1)\right]^{T} \tag{IV.30}$$

Exercise IV-2

Explain why the correlation matrix is not necessarily a rank-one matrix.

In expanded form, we have

$$R=\begin{bmatrix} r(0) & r(1) & \cdots & r(M-1) \\ r(-1) & r(0) & \cdots & r(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ r(-M+1) & r(-M+2) & \cdots & r(0) \end{bmatrix} \tag{IV.31}$$

Correspondingly, let $\underline{p}$ denote the $M\times 1$ cross-correlation vector between the tap inputs of the filter and the desired response $d(n)$:

$$\underline{p}=\left[p(0) \quad p(-1) \quad \cdots \quad p(1-M)\right]^T \tag{IV.32}$$

Note that the lags used in the definition of $\underline{p}$ are either zero or else negative. We may thus rewrite the Wiener-Hopf equations in the compact matrix form:

$$R\underline{w}_o=\underline{p} \tag{IV.33}$$

where

$$\underline{w}_o=\left[w_{o,0} \quad w_{o,1} \quad \cdots \quad w_{o,M-1}\right]^T \tag{IV.34}$$

To solve the Wiener-Hopf equations for $\underline{w}_o$ we assume that the correlation matrix R is nonsingular, yielding

$$\underline{w}_o=R^{-1}\underline{p} \tag{IV.35}$$

The computation of the optimum tap-weight vector $\underline{w}_o$ requires knowledge of two quantities: (1) the correlation matrix R of the tap-input vector $\underline{u}(n)$ and (2) the cross-correlation vector $\underline{p}$ between the tap-input vector $\underline{u}(n)$ and the desired response $d(n)$.

IV.6. Properties of the Correlation Matrix

The correlation matrix of a stationary discrete-time stochastic process has the following properties:

- The correlation matrix of a stationary discrete-time stochastic process is Hermitian.

$$R^H=R \tag{IV.36}$$

Another way of stating the Hermitian property is

$$r(-k)=r^*(k) \tag{IV.37}$$

Therefore, the correlation matrix can be rewritten in the form

$$R = \begin{bmatrix} r(0) & r(1) & \cdots & r(M-1) \\ r^*(1) & r(0) & \cdots & r(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ r^*(M-1) & r^*(M-2) & \cdots & r(0) \end{bmatrix} \tag{IV.38}$$

For the special case of real-valued data, the autocorrelation function $r(k)$ is real for all k, and the correlation matrix is symmetric.

- The correlation matrix of a stationary discrete· time stochastic process is Toeplitz.

We say that a square matrix is Toeplitz if all the elements on its main diagonal are equal, and if the elements on any other diagonal parallel to the main diagonal are also equal. It is important to recognize, however, that the Toeplitz property of the correlation matrix is a direct consequence of the assumption that the discrete-time stochastic process represented by the observation vector $\underline{u}(n)$ is wide-sense stationary.

Indeed, we may state that if the discrete-time stochastic process is wide-sense stationary, then its correlation matrix must be Toeplitz; and, conversely, if the correlation matrix is Toeplitz, then the discrete-time stochastic process must be wide-sense stationary.

- The correlation matrix of a discrete-time stochastic process is always nonnegative definite and almost always positive definite.

Let $\underline{x}$ be an arbitrary (nonzero) $M \times 1$ complex-valued vector. Define the scalar random variable y as the inner product of $\underline{x}$ and the observation vector $\underline{u}(n)$, as shown by

$$y = \underline{x}^H \underline{u}(n) \tag{IV.39}$$

Taking the Hermitian transpose of both sides and recognizing that y is a scalar, we get

$$y^* = \underline{u}^H(n)\underline{x} \tag{IV.40}$$

The mean-square value of the random variable y equals

$$\begin{aligned} \mathrm{E}\left[|y|^2\right] &= \mathrm{E}\left[yy^*\right] \\ &= \mathrm{E}\left[\underline{x}^H \underline{u}(n)\underline{u}^H(n)\underline{x}\right] \\ &= \underline{x}^H \,\mathrm{E}\left[\underline{u}(n)\underline{u}^H(n)\right]\underline{x} \\ &= \underline{x}^H R\underline{x} \end{aligned} \tag{IV.41}$$

The expression $\underline{x}^H R\underline{x}$ is called a Hermitian form. Since

$$\mathrm{E}\left[|y|^2\right] \geq 0 \tag{IV.42}$$

it follows that

$$\underline{x}^H R \underline{x} \geq 0 \tag{IV.43}$$

A Hermitian form that satisfies this condition for every nonzero $\underline{x}$ is said to be nonnegative definite or positive semidefinite. Accordingly, we may state that the correlation matrix of a wide-sense stationary process is always nonnegative definite.

If the Hermitian form $\underline{x}^H R \underline{x}$ satisfies the condition

$$\underline{x}^H R \underline{x} > 0 \tag{IV.44}$$

for every nonzero $\underline{x}$, we say that the correlation matrix R is positive definite. This condition is satisfied for a wide-sense stationary process unless there are linear dependencies between the random variables that constitute the M elements of the observation vector. Such a situation arises essentially only when the process $\underline{u}(n)$ consists of the sum of K sinusoids with $K \leq M$. In practice, we find that this idealized situation is so rare in occurrence that the correlation matrix is almost always positive definite.

The positive definiteness of a correlation matrix implies that its determinant is greater than zero. This implies that the correlation matrix is nonsingular. We say that a matrix is nonsingular if its inverse exists; otherwise, it is singular. Accordingly, we may state that a correlation matrix is almost always nonsingular.

- When the elements that constitute the observation vector of a stationary discrete-time stochastic process are rearranged backward, the effect is equivalent to the transposition of the correlation matrix of the process.

Let $\underline{u}^B(n)$ denote the $M \times 1$ vector obtained by rearranging the elements that constitute the observation vector $\underline{u}(n)$ backward. We illustrate this operation by writing

$$\underline{u}^B(n) = \begin{bmatrix} u(n-M+1) & u(n-M+2) & \cdots & u(n) \end{bmatrix}^T \tag{IV.45}$$

Then

$$\begin{aligned} R^B &= \mathrm{E}\left[\underline{u}^B(n)\underline{u}^{BH}(n)\right] \\ &= R^T \end{aligned} \tag{IV.46}$$

- The correlation matrices R_M and R_{M+1} of a stationary discrete-time stochastic process, pertaining to M and $M+1$ observations of the process, respectively are related by

$$\begin{aligned} R_{M+1} &= \begin{bmatrix} r(0) & \underline{r}^H \\ \underline{r} & R_M \end{bmatrix} \\ &= \begin{bmatrix} R_M & \underline{r}^{B*} \\ \underline{r}^{BT} & r(0) \end{bmatrix} \end{aligned} \tag{IV.47}$$

IV.7. Eigenanalysis of the Correlation Matrix

Let the Hermitian matrix R denote the $M \times M$ correlation matrix of a wide-sense stationary discrete-time stochastic process represented by the $M \times 1$ observation vector $\underline{u}(n)$. In general, this matrix may contain complex elements. We wish to find an $M \times 1$ vector $\underline{q}$ that satisfies the condition for some constant λ:

$$R\underline{q} = \lambda \underline{q} \tag{IV.48}$$

This condition states that the vector $\underline{q}$ is linearly transformed to the vector $\lambda\underline{q}$ by the Hermitian matrix R. Since λ is a constant, the vector $\underline{q}$ therefore has special significance in that it is left invariant in direction (in the M-dimensional space) by a linear transformation. For a typical $M \times M$ matrix there will be M such vectors. To show this, we first rewrite (IV.48) in the form

$$(R - \lambda I)\underline{q} = 0 \tag{IV.49}$$

Equation (IV.49) will have a solution only when the matrix $R - \lambda I$ is singular. This happens when

$$\det(R - \lambda I) = 0 \tag{IV.50}$$

This determinant can be expanded in the form of a polynomial in λ of degree M. We thus find that, in general, (IV.50) has M distinct roots. Correspondingly, (IV.50) has M solutions in the form of vector $\underline{q}$.

Equation (IV.50) is called the characteristic equation of the matrix R. Let $\lambda_1, \lambda_2, \cdots, \lambda_M$ denote the M roots of this equation. These roots are called the eigenvalues of the matrix R. Although R has M eigenvalues, they need not be distinct. When the characteristic equation has multiple roots, the matrix R is said to have degenerate eigenvalues.

Let λ_i denote the i th eigenvalue of the matrix R. Also. let $\underline{q}_i$ be a nonzero vector such that

$$R\underline{q}_i = \lambda_i \underline{q}_i \tag{IV.51}$$

The vector $\underline{q}_i$ is called the eigenvector associated with λ_i. An eigenvector can correspond to only one eigenvalue. However, an eigenvalue may have many eigenvectors. For example, if $\underline{q}_i$ is an eigenvector associated with eigenvalue λ_i, then so is $a\underline{q}_i$ for any scalar $a \neq 0$.

IV.8. Properties of Eigenvalues and Eigenvectors

- If $\lambda_1, \lambda_2, \cdots, \lambda_M$ denote the eigenvalues of the correlation matrix R, then the eigenvalues of the matrix R^k equal $\lambda_1^k, \lambda_2^k, \cdots, \lambda_M^k$ for any integer $k > 0$.
- Let $\underline{q}_1, \underline{q}_2, \cdots, \underline{q}_M$ be the eigenvectors corresponding to the distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_M$ of the $M \times M$ correlation matrix R, respectively. Then the eigenvectors $\underline{q}_1, \underline{q}_2, \cdots, \underline{q}_M$ are linearly independent.

Vandermonde Matrix

A matrix in the following form is called a Vandermonde matrix:

$$S = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{M-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{M-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_M & \lambda_M^2 & \cdots & \lambda_M^{M-1} \end{bmatrix} \tag{IV.52}$$

When $\lambda_1, \lambda_2, \cdots, \lambda_M$ are distinct, the Vandermonde matrix S is nonsingular.

- Let $\lambda_1, \lambda_2, \cdots, \lambda_M$ be the eigenvalues of the $M \times M$ correlation matrix R. Then all these eigenvalues are real and nonnegative.
- Let $\underline{q}_1, \underline{q}_2, \cdots, \underline{q}_M$ be the eigenvectors corresponding to the distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_M$ of the $M \times M$ correlation matrix R, respectively. Then the eigenvectors $\underline{q}_1, \underline{q}_2, \cdots, \underline{q}_M$ are orthogonal to each other:

$$\underline{q}_i^H \underline{q}_j = 0, \quad \forall i \neq j \tag{IV.53}$$

- Let $\underline{q}_1, \underline{q}_2, \cdots, \underline{q}_M$ be the eigenvectors corresponding to the distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_M$ of the $M \times M$ correlation matrix R, respectively. Define the $M \times M$ matrix

$$Q = \begin{bmatrix} \underline{q}_1 & \underline{q}_2 & \cdots & \underline{q}_M \end{bmatrix} \tag{IV.54}$$

where $\left\{\underline{q}_i\right\}_{i=1}^{M}$ constitutes a set of orthonormal vectors. i.e.,

$$\underline{q}_i^H \underline{q}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{IV.55}$$

Matrix Q is called a unitary matrix (orthogonal matrix for real entries). Matrix Q has the property

$$Q^{-1} = Q^H \tag{IV.56}$$

Matrix Q is nonsingular.

Define

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_M \end{bmatrix} \tag{IV.57}$$

Then,

$$\Lambda = Q^H R Q \tag{IV.58}$$

Using (IV.56) in (IV.58),

$$RQ = Q\Lambda \tag{IV.59}$$

We have proved that the correlation matrix R may be diagonalized by a unitary similarity transformation.

By postmultiplying both sides of (IV.59) by Q^{-1} and then using (IV.56), we may also write

$$\begin{aligned} R &= Q\Lambda Q^H \\ &= \sum_{i=1}^{M} \lambda_i \underline{q}_i \underline{q}_i^H \end{aligned} \tag{IV.60}$$

Let the projection P_i denote the outer product $\underline{q}_i \underline{q}_i^H$. Then, it is a straightforward matter to show that

$$\begin{aligned} P_i &= P_i^2 \\ &= P_i^H \end{aligned} \tag{IV.61}$$

Note that P_i is a rank-one matrix. The above shows that the correlation matrix of a wide-sense stationary process equals the linear combination of rank-one projections, with each projection being weighted by the respective eigenvalue. This result is known as Mercer's theorem. It is also referred to as the spectral theorem.

- Let $\lambda_1, \lambda_2, \cdots, \lambda_M$ be the eigenvalues of the $M \times M$ correlation matrix R. Then the sum of these eigenvalues equals the trace of matrix R. The trace of a square matrix is defined as the sum of the diagonal elements of the matrix.
- The correlation matrix R is ill conditioned if the ratio of the largest eigenvalue to the smallest eigenvalue of R is large. When R is ill-conditioned, the solution of $R\underline{w}_o = \underline{p}$ is very sensitive to changes in $\underline{p}$.
- The eigenvalues of the correlation matrix of a discrete-time stochastic process are bounded by the minimum and maximum values of the power spectral density of the process.

Let $\{\lambda_i\}_{i=1}^{M}$ and $\{\underline{q}_i\}_{i=1}^{M}$ denote the eigenvalues of the $M \times M$ correlation matrix R of a discrete-time stochastic process $u(n)$ and their associated eigenvectors, respectively. Note that

$$\underline{q}_i^H R \underline{q}_i = \lambda_i \underline{q}_i^H \underline{q}_i \tag{IV.62}$$

Since $\underline{q}_i^H \underline{q}_i > 0$, we can determine λ_i as

$$\lambda_i = \frac{\underline{q}_i^H R \underline{q}_i}{\underline{q}_i^H \underline{q}_i} \tag{IV.63}$$

The Hermitian form in the numerator may be expressed in its expanded form as follows

$$\underline{q}_i^H R \underline{q}_i = \sum_{k=1}^{M} \sum_{l=1}^{M} q_{ik}^* r(l-k) q_{il} \tag{IV.64}$$

Expressing the autocorrelation function as the inverse Fourier transform of the power spectral density yields

$$r(l-k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_u(e^{j\omega}) e^{j\omega(l-k)} d\omega \tag{IV.65}$$

Substituting (IV.65) into (IV.64),

$$\begin{aligned} \underline{q}_i^H R \underline{q}_i &= \frac{1}{2\pi} \sum_{k=1}^{M} \sum_{l=1}^{M} q_{ik}^* q_{il} \int_{-\pi}^{\pi} S_u(e^{j\omega}) e^{j\omega(l-k)} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} S_u(e^{j\omega}) \left\{ \sum_{k=1}^{M} q_{ik}^* e^{-j\omega k} \sum_{l=1}^{M} q_{il} e^{j\omega l} \right\} d\omega \end{aligned} \tag{IV.66}$$

Let's denote the DTFT of $\{q_{il}\}_{l=1}^{M}$ as

$$Q_i(e^{j\omega}) = \sum_{l=1}^{M} q_{il} e^{j\omega l} \tag{IV.67}$$

Then,

$$\underline{q}_i^H R \underline{q}_i = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_u(e^{j\omega}) \left| Q_i(e^{j\omega}) \right|^2 d\omega \tag{IV.68}$$

Similarly, it can be shown that

$$\underline{q}_i^H \underline{q}_i = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| Q_i(e^{j\omega}) \right|^2 d\omega \tag{IV.69}$$

Accordingly,

$$\lambda_i = \frac{\int_{-\pi}^{\pi} S_u(e^{j\omega}) \left| Q_i(e^{j\omega}) \right|^2 d\omega}{\int_{-\pi}^{\pi} \left| Q_i(e^{j\omega}) \right|^2 d\omega} \tag{IV.70}$$

Let $S_{\min}$ and $S_{\max}$ denote the absolute minimum and maximum values of the power spectral density $S_u(e^{j\omega})$, respectively. Then it follows that

$$\int_{-\pi}^{\pi} S_u(e^{j\omega}) \left| Q_i(e^{j\omega}) \right|^2 d\omega \geq S_{\min} \int_{-\pi}^{\pi} \left| Q_i(e^{j\omega}) \right|^2 d\omega \tag{IV.71}$$

and

$$\int_{-\pi}^{\pi} S_u(e^{j\omega}) \left| Q_i(e^{j\omega}) \right|^2 d\omega \leq S_{\max} \int_{-\pi}^{\pi} \left| Q_i(e^{j\omega}) \right|^2 d\omega \tag{IV.72}$$

Equations (IV.71) and (IV.72) can be rewritten as follows:

$$\lambda_i \geq S_{\min} \tag{IV.73}$$

and

$$\lambda_i \leq S_{\max} \tag{IV.74}$$

Combining (IV.73) and (IV.74) we get

$$S_{\min} \leq \lambda_i \leq S_{\max} \tag{IV.75}$$

Correspondingly, the eigenvalue spread $\mathcal{X}(R)$ is bounded as

$$\begin{aligned} \mathcal{X}(R) &= \frac{\lambda_{\max}}{\lambda_{\min}} \\ &\leq \frac{S_{\max}}{S_{\min}} \end{aligned} \tag{IV.76}$$

IV.8.A. KARHUNEN-LOÉVE EXPANSION:

$$\underline{u}(n) = \sum_{i=1}^{M} c_i(n) \underline{q}_i \tag{IV.77}$$

The coefficients of the expansion are zero-mean, uncorrelated random variables defined by the inner product

$$c_i(n) = \underline{q}_i^H \underline{u}(n) \tag{IV.78}$$

The coefficients of the expansion are random variables characterized as follows:

$$\mathrm{E}\left[c_i(n)\right] = 0 \tag{IV.79}$$

and

$$\mathrm{E}\left[c_i(n)c_j^*(n)\right] = \begin{cases} \lambda_i, & i = j \\ 0, & i \neq j \end{cases} \tag{IV.80}$$

For a physical interpretation of the Karhunen- Loéve expansion, we may view the eigenvectors as the coordinates of an M -dimensional space, and thus represent the random vector $\underline{u}(n)$ by the set of its projections $\{c_i(n)\}_{i=1}^{M}$ onto these axes, respectively.

Moreover, we deduce from (IV.77) that

$$\begin{aligned} \sum_{i=1}^{M} \left|c_i(n)\right|^2 &= \underline{u}^H(n)\underline{u}(n) \\ &= \left\|\underline{u}(n)\right\|^2 \end{aligned} \tag{IV.81}$$

That is to say, the coefficient $c_i(n)$ has an energy equal to that of the observation vector $\underline{u}(n)$ measured along the i^{th} coordinate. Naturally, this energy is a random variable whose mean value equals the i^{th} eigenvalue, as shown by

$$\mathrm{E}\left[\left|c_i(n)\right|^2\right] = \lambda_i \tag{IV.82}$$

IV.9. Error Performance Surface

The Wiener-Hopf equations are traceable to the principle of orthogonality. We may also derive the Wiener-Hopf equations by examining the dependence of the cost function J on the tap weights of the transversal filter. First, we write the estimation error $e(n)$ as follows:

$$e(n) = d(n) - \sum_{k=0}^{M-1} w_k^* u(n-k) \tag{IV.83}$$

Accordingly, we may define the cost function for the transversal filter structure as

$$J = \mathrm{E}\left[e(n)e^*(n)\right] \tag{IV.84}$$

Therefore,

$$J = \mathrm{E}\left[\left|d(n)\right|^2\right] - \sum_{k=0}^{M-1} w_k^* \,\mathrm{E}\left[u(n-k)d^*(n)\right] - \sum_{k=0}^{M-1} w_k \,\mathrm{E}\left[u^*(n-k)d(n)\right] + \sum_{k=0}^{M-1}\sum_{i=0}^{M-1} w_k^* w_i \,\mathrm{E}\left[u(n-k)u^*(n-i)\right] \quad \text{(IV.85)}$$

Note that

$$\mathrm{E}\left[\left|d(n)\right|^2\right] = \sigma_d^2 \quad \text{(IV.86)}$$

where σ_d^2 is the variance of the desired response, assumed to be of zero mean.

$$\mathrm{E}\left[u(n-k)d^*(n)\right] = p(-k) \quad \text{(IV.87)}$$

$$\mathrm{E}\left[u^*(n-k)d(n)\right] = p^*(-k) \quad \text{(IV.88)}$$

$$\mathrm{E}\left[u(n-k)u^*(n-i)\right] = r(i-k) \quad \text{(IV.89)}$$

Using the last few equations in (IV.85) yields

$$J = \sigma_d^2 - \sum_{k=0}^{M-1} w_k^* p(-k) + \sum_{k=0}^{M-1} w_k^* p^*(-k) + \sum_{k=0}^{M-1}\sum_{i=0}^{M-1} w_k^* w_i r(i-k) \quad \text{(IV.90)}$$

Equation (IV.90) states that for the case when the tap inputs of the transversal filter and the desired response are jointly stationary, the cost function, or mean-squared error, J is precisely a second-order function of the tap weights in the filter. Consequently, we may visualize the dependence of the cost function on the tap weights as a bowl-shaped $M+1$-dimensional surface with M degrees of freedom represented by the tap weights of the filter. This surface is characterized by a unique minimum. We refer to the surface so described as the error-performance surface of the transversal filter.

At the bottom or minimum point of the error-performance surface, the cost function attains its minimum value denoted by $J_{\min}$. At this point, the gradient vector $\underline{\nabla}J$ is identically zero. In other words,

$$\nabla_k J = 0, \quad k = 0,1,\cdots,M-1 \quad \text{(IV.91)}$$

Let

$$w_k = a_k + jb_k \quad \text{(IV.92)}$$

Applying (IV.91) to (IV.90), and replacing w_i by $w_{o,i}$ we get

$$\begin{aligned}\nabla_k J &= \frac{\partial J}{\partial a_k} + j\frac{\partial J}{\partial b_k} \\ &= -2p(-k) + 2\sum_{i=0}^{M-1} w_{o,i} r(i-k) \\ &= 0\end{aligned} \qquad \text{(IV.93)}$$

This simplifies to

$$\sum_{i=0}^{M-1} w_{o,i} r(i-k) = p(-k), \quad k = 0, 1, \cdots, M-1 \qquad \text{(IV.94)}$$

This system of equations is identical to the Wiener-Hopf equations derived earlier.

IV.10. Minimum Mean-Squared Error

Let $\hat{d}\left(n|\mathcal{U}_n\right)$ denote the estimate of the desired response $d(n)$, produced at the output of the transversal filter, that is optimized in the mean-squared-error sense, given the tap inputs $u(n), u(n-1), \cdots, u(n-M+1)$ that span the space $\mathcal{U}_n$. Then,

$$\begin{aligned}\hat{d}\left(n|\mathcal{U}_n\right) &= \sum_{k=0}^{M-1} w_{o,k}^* u(n-k) \\ &= \underline{w}_o^H \underline{u}(n)\end{aligned} \qquad \text{(IV.95)}$$

Note that $\underline{w}_o^H \underline{u}(n)$ denotes an inner product of the optimum tap-weight vector $\underline{w}_o$ and the tap-input vector $\underline{u}(n)$. We assume that $\underline{u}(n)$ has zero mean, making the estimate $\hat{d}\left(n|\mathcal{U}_n\right)$ have zero mean too. Hence,

$$\begin{aligned}\sigma_{\hat{d}}^2 &= \mathrm{E}\left[\underline{w}_o^H \underline{u}(n)\underline{u}^H(n)\underline{w}_o\right] \\ &= \underline{w}_o^H \mathrm{E}\left[\underline{u}(n)\underline{u}^H(n)\right]\underline{w}_o \\ &= \underline{w}_o^H R \underline{w}_o\end{aligned} \qquad \text{(IV.96)}$$

where R is the correlation matrix of the tap-weight vector $\underline{u}(n)$. Note that we can write

$$\begin{aligned}\sigma_{\hat{d}}^2 &= \underline{w}_o^H \underline{p} \\ &= \underline{p}^H \underline{w}_o\end{aligned} \qquad \text{(IV.97)}$$

Now, since

$$J_{\min} = \sigma_d^2 - \sigma_{\hat{d}}^2 \qquad \text{(IV.98)}$$

We get

$$\begin{aligned} J_{\min} &= \sigma_d^2 - \underline{p}^H \underline{w}_o \\ &= \sigma_d^2 - \underline{p}^H R^{-1} \underline{p} \end{aligned} \tag{IV.99}$$

IV.11. Canonical Form of the Error Performance Surface

Equation (IV.85) defines the expanded form of the mean-squared error produced by the transversal filter. We may rewrite this equation in matrix form, by using the definitions for the correlation matrix R and the cross-correlation vector $\underline{p}$:

$$J(\underline{w}) = \sigma_d^2 - \underline{w}^H \underline{p} - \underline{p}^H \underline{w} + \underline{w}^H R \underline{w} \tag{IV.100}$$

The correlation matrix R is almost always positive definite, so that the inverse matrix R^{-1} exists. Accordingly, expressing $J(\underline{w})$ as a "perfect square" in $\underline{w}$, we may rewrite (IV.100) in the form

$$J(\underline{w}) = \sigma_d^2 - \underline{p}^H R^{-1} \underline{p} + \left(\underline{w} - R^{-1}\underline{p}\right)^H R\left(\underline{w} - R^{-1}\underline{p}\right) \tag{IV.101}$$

We now immediately see that

$$\min_{\underline{w}} J(\underline{w}) = \sigma_d^2 - \underline{p}^H R^{-1} \underline{p} \tag{IV.102}$$

This happens when

$$\underline{w}_o = R^{-1}\underline{p} \tag{IV.103}$$

Note that we can write

$$J(\underline{w}) = J_{\min} + (\underline{w} - \underline{w}_o)^H R(\underline{w} - \underline{w}_o) \tag{IV.104}$$

This equation shows explicitly the unique optimality of the minimizing tap-weight vector $\underline{w}_o$.

Although the quadratic form on the right-hand side of (IV.104) is quite informative, nevertheless, it is desirable to change the basis on which it is defined so that the representation of the error-performance surface is considerably simplified. To do this, we recall that

$$R = Q\Lambda Q^H \tag{IV.105}$$

Hence, substituting (IV.105) into (IV.104), we get

$$J = J_{\min} + \left(\underline{w} - \underline{w}_o\right)^H Q\Lambda Q^H \left(\underline{w} - \underline{w}_o\right) \tag{IV.106}$$

Define a transformed version of the difference between the tap-weight vector $\underline{w}$ and the optimum solution $\underline{w}_o$ as

$$\underline{v} = Q^H \left(\underline{w} - \underline{w}_o\right) \tag{IV.107}$$

Then we may put the quadratic form of (IV.106) into its canonical form defined by

$$J = J_{\min} + \underline{v}^H \Lambda \underline{v} \tag{IV.108}$$

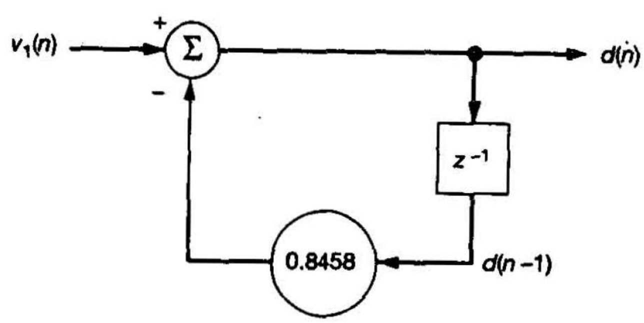
This new formulation of the mean-squared error contains no cross-product terms, as shown by

$$\begin{aligned} J &= J_{\min} + \sum_{k=1}^{M} \lambda_k v_k v_k^* \\ &= J_{\min} + \sum_{k=1}^{M} \lambda_k \left| v_k \right|^2 \end{aligned} \tag{IV.109}$$

The feature that makes the canonical form of (IV.108) a rather useful representation of the error-performance surface is the fact that the components of the transformed coefficient vector $\underline{v}$ constitute the principal axes of the error-performance surface.

IV.12. Numerical Example

Consider the example depicted below.

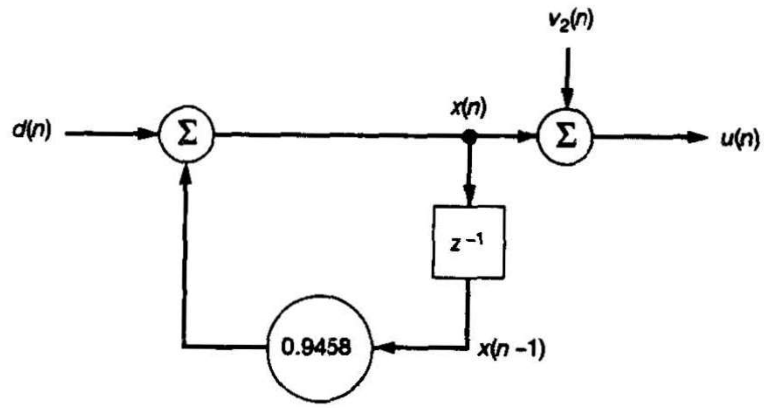


The desired response $d(n)$ is modeled as an autoregressive (AR) process of order 1; that is, it may be produced by applying a white-noise process $v_1(n)$ of zero mean and variance $\sigma_1^2 = 0.27$ to the input of an all-pole filter of order 1, whose transfer function equals

$$H_1(z) = \frac{1}{1 + 0.8458 z^{-1}} \tag{IV.110}$$

The process $d(n)$ is applied to a communication channel modeled by the first order all-pole transfer function

$$H_2(z) = \frac{1}{1 - 0.9458 z^{-1}} \tag{IV.111}$$



(b)

The channel output $x(n)$ is corrupted by an additive white-noise process $v_2(n)$ of zero mean and variance $\sigma_2^2 = 0.1$ so a sample of the received signal $u(n)$ equals

$$u(n) = x(n) + v_2(n) \tag{IV.112}$$

The white-noise processes $v_1(n)$ and $v_2(n)$ are uncorrelated. It is also assumed that $d(n)$ and $u(n)$, and therefore $v_1(n)$ and $v_2(n)$ are all real valued.

The requirement is to specify a Wiener filter consisting of a transversal filter with two taps, which operates on the received signal $u(n)$ so as to produce an estimate of the desired response $d(n)$ that is optimum in the mean-square sense.

We begin the analysis by considering the difference equations that characterize the various processes described above. The generation of the desired response $d(n)$ is governed by the first-order difference equation

$$d(n) + a_1 d(n-1) = v_1(n) \tag{IV.113}$$

where $a_1 = 0.8458$. The variance of the process $d(n)$ equals

$$\begin{aligned}\sigma_d^2 &= \frac{\sigma_1^2}{1 - a_1^2} \\ &= 0.9486\end{aligned} \tag{IV.114}$$

Exercise IV-3

Prove (IV.114).

The channel output $x(n)$ is related to the channel input $d(n)$ by the first-order difference equation

$$x(n) + b_1 x(n-1) = d(n) \tag{IV.115}$$

where $b_1 = -0.9458$. Note that the channel output $x(n)$ may be generated by applying the white-noise process $v_1(n)$ to a second-order all-pole filter whose transfer function equals

$$\begin{aligned} H(z) &= H_1(z)H_2(z) \\ &= \frac{1}{\left(1+a_1 z^{-1}\right)\left(1+a_2 z^{-1}\right)} \end{aligned} \qquad \text{(IV.116)}$$

Accordingly, $x(n)$ is a second-order AR process described by the difference equation

$$x(x) + \gamma_1 x(n-1) + \gamma_2 x(n-2) = v_1(n) \qquad \text{(IV.117)}$$

where $\gamma_1 = -0.1$ and $\gamma_2 = -0.8$. Note that both AR processes $d(n)$ and are $x(n)$ wide-sense stationary.

Introducing $\gamma_0 = 1$ and $M = 2$, equation (IV.117) can be rewritten in the form

$$\sum_{k=0}^{M} \gamma_k x(n-k) = v_1(n) \qquad \text{(IV.118)}$$

Multiplying both sides of (IV.118) by $x(n-l)$ and applying the expectation operation, we get

$$\mathrm{E}\left[\sum_{k=0}^{M} \gamma_k x(n-k)x(n-l)\right] = \mathrm{E}\left[v_1(n)x(n-l)\right] \qquad \text{(IV.119)}$$

Note that

$$\mathrm{E}\left[x(n-k)x(n-l)\right] = r_x(l-k) \qquad \text{(IV.120)}$$

Note also that $x(n-l)$ only involves samples of the white-noise process at the filter input up to time $n-l$. This mean that

$$\mathrm{E}\left[v(n)x(n-l)\right] = 0, \quad l > 0 \qquad \text{(IV.121)}$$

Therefore, (IV.119) simplifies to

$$\sum_{k=0}^{M} \gamma_k r_x(l-k) = 0, \quad l > 0 \qquad \text{(IV.122)}$$

The general solution of (IV.122) can be expressed in the form

$$r_x(m) = \sum_{k=1}^{M} C_k p_k^m \qquad \text{(IV.123)}$$

where $\{C_k\}_{k=1}^{M}$ are constants, and $\{p_k\}_{k=1}^{M}$ are the poles of $H(z)$.

Since the processes $x(n)$ and $v_2(n)$ are uncorrelated, it follows that the correlation matrix R equals the correlation matrix of $x(n)$ plus the correlation matrix of $v_2(n)$. That is,

$$R = R_x + R_2 \tag{IV.124}$$

For the correlation matrix R_x we write [since the process $x(n)$ is real valued]:

$$R_x = \begin{bmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{bmatrix} \tag{IV.125}$$

The values of γ_1 and γ_2 allow $x(n)$ to be asymptotically stationary. Hence, the autocorrelation sequence of $x(n)$ satisfies the difference equation

$$r_x(l) + \gamma_1 r_x(l-1) + \gamma_2 r_x(l-2) = 0, \quad l > 0 \tag{IV.126}$$

Note that

$$r_x(0) = \sigma_x^2 \tag{IV.127}$$

Solving (IV.126) for $l = 1$ and $l = 2$ yields

$$r_x(1) = \frac{-\gamma_1}{1+\gamma_2}\sigma_x^2 \tag{IV.128}$$

$$r_x(2) = \left(-\gamma_2 + \frac{\gamma_1^2}{1+\gamma_2}\right)\sigma_x^2 \tag{IV.129}$$

Note that letting $l = 0$ in $\mathrm{E}\left[v_1(n)x(n-l)\right]$, yields

$$\begin{aligned} \mathrm{E}\left[v_1(n)x(n)\right] &= \mathrm{E}\left[v_1^2(n)\right] \\ &= \sigma_1^2 \end{aligned} \tag{IV.130}$$

Using this result in (IV.119) for $l = 0$ produces

$$\begin{aligned} \sigma_1^2 &= \sum_{k=0}^{M} \gamma_k r_x(k) \\ &= r_x(0) + \gamma_1 r_x(1) + \gamma_2 r_x(2) \end{aligned} \tag{IV.131}$$

Using (IV.128) and (IV.129) in (IV.131) and solving for σ_x^2 yields

$$\sigma_x^2 = \frac{1+\gamma_2}{1-\gamma_2}\frac{1}{\left(1+\gamma_2\right)^2 - \gamma_1^2}\sigma_1^2 \tag{IV.132}$$

Substituting for σ_1^2, γ_1 and γ_2 in (IV.132) results in

$$\sigma_x^2 = 1 \tag{IV.133}$$

And hence, we get from (IV.127) and (IV.128)

$$r_x(0) = 1 \tag{IV.134}$$

$$r_x(1) = 0.5 \tag{IV.135}$$

Which make R_x equal to

$$R_x = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \tag{IV.136}$$

Note that since $v_2(n)$ is a white-noise process of zero mean and variance $\sigma_2^2 = 0.1$, the 2×2 correlation matrix R_2 equals

$$R_2 = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix} \tag{IV.137}$$

Therefore,

$$R = \begin{bmatrix} 1.1 & 0.5 \\ 0.5 & 1.1 \end{bmatrix} \tag{IV.138}$$

Recall that the vector $\underline{p}$ is given by

$$\underline{p} = \begin{bmatrix} p(0) \\ p(-1) \end{bmatrix} \tag{IV.139}$$

For real processes, $p(k)$ can be found from

$$\begin{aligned} p(k) &= \mathrm{E}\left[u(n-k)d(k)\right] \\ &= p(-k) \end{aligned} \tag{IV.140}$$

Substituting (IV.113) and (IV.115) into (IV.140) yields

$$p(k) = r_x(k) + b_1 r_x(k-1) \tag{IV.141}$$

From the last result we get

$$\underline{p} = \begin{bmatrix} 0.5272 \\ -0.4458 \end{bmatrix} \tag{IV.142}$$

The dependence of the mean-squared error on the tap-weight vector $\underline{w} = \begin{bmatrix} w_0 & w_1 \end{bmatrix}^T$ is defined by (IV.100). Substituting all the needed quantities in (IV.100),

$$J(w_0, w_1) = 0.9486 - 1.0544w_0 + 0.8916w_1 + 1.1\left(w_0^2 + w_1^2\right) \tag{IV.143}$$

The inverse of the autocorrelation matrix is calculated as

$$R^{-1} = \begin{bmatrix} 1.1456 & -0.5208 \\ -0.5208 & 1.1456 \end{bmatrix} \tag{IV.144}$$

And hence, the Wiener filter coefficient vector is given by

$$\begin{aligned} \underline{w}_0 &= R^{-1}\underline{p} \\ &= \begin{bmatrix} 0.8360 \\ -0.7853 \end{bmatrix} \end{aligned} \tag{IV.145}$$

The MMSE can be found from (IV.102) to be equal to

$$J_{\min} = 0.1579 \tag{IV.146}$$

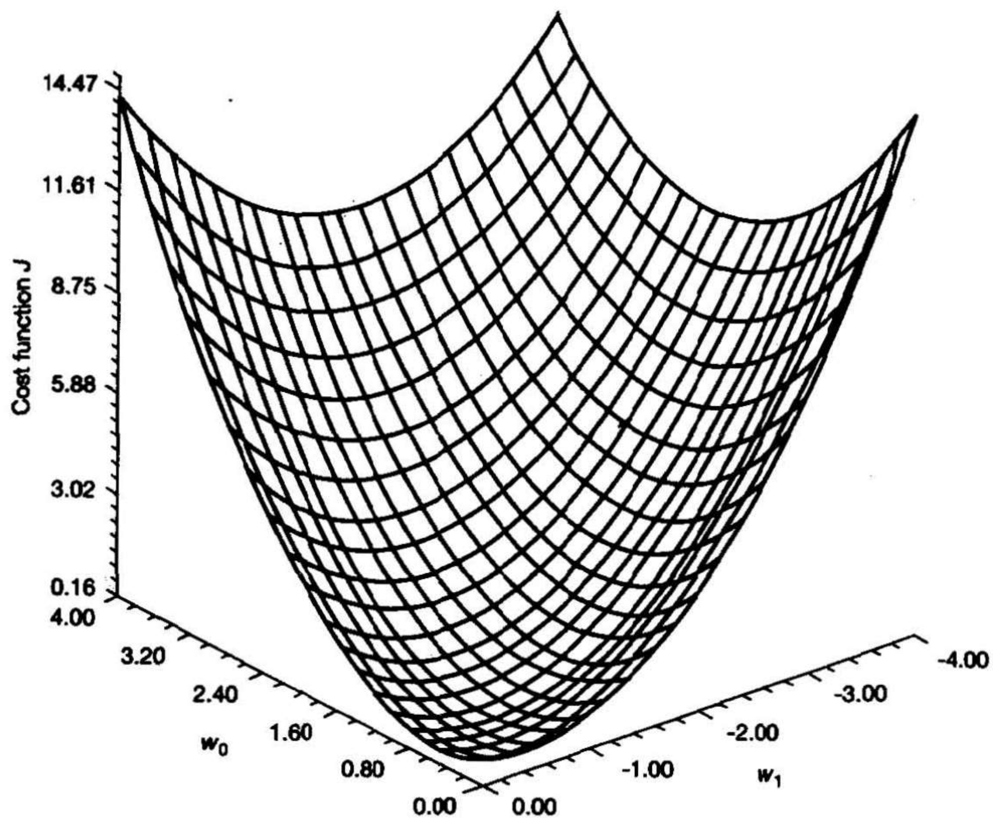


Figure IV.4

***