8.1 Basic Concepts

A random process is a collection of time functions and an associated probability description.

When a continuous or discrete or mixed process in time/space can be describe mathematically as a function containing one or more random variables.

- A sinusoidal waveform with a random amplitude.
- A sinusoidal waveform with a random phase.
- A sequence of digital symbols, each taking on a random value for a defined time period (e.g. amplitude, phase, frequency).
- A random walk (2-D or 3-D movement of a particle)

The entire collection of possible time functions is an ensemble, designated as \( \{x(t)\} \), where one particular member of the ensemble, designated as \( x(t) \), is a sample function of the ensemble. In general only one sample function of a random process can be observed!

Think of:

\[
X(t) = A \cdot \sin(w \cdot t + \theta), \quad 0 \leq \theta \leq 2 \cdot \pi
\]

where \( A \) and \( w \) are known constants.

Note that once a sample has been observed …

\[
x(t_1) = A \cdot \sin(w \cdot t_1 + \theta)
\]

the function is known for all time, \( t \).

Note that, \( x(t_2) \) is a second time sample of the same random process and does not provide any “new information” about the value of the random variable.

\[
x(t_2) = A \cdot \sin(w \cdot t_2 + \theta)
\]

There are many similar ensembles in engineering, where the sample function, once known, provides a continuing solution. In many cases, an entire system design approach is based on either assuming that randomness remains or is removed once actual measurements are taken!

For example, in communications there is a significant difference between coherent (phase and frequency) demodulation and non-coherent (i.e. unknown starting phase) demodulation.

On the other hand, another measurement in a different environment might measure

\[
x_2(t_1) = A_2 \cdot \sin(w \cdot t_1 + \theta_2)
\]

In this “space” the random variables could take on other values within the defined ranges. Thus an entire “ensemble” of possibilities may exist based on the random variables defined in the random process.
For example, assume that there is a known AM signal transmitted:

\[ s(t) = (1 + b \cdot A(t)) \cdot \sin(w \cdot t) \]

at an undetermined distance the signal is received as

\[ y(t) = (1 + b \cdot A(t)) \cdot \sin(w \cdot t + \theta), \quad 0 \leq \theta \leq 2 \cdot \pi \]

The received signal is mixed and low pass filtered …

\[
\begin{align*}
    x(t) &= h(t) \otimes [y(t) \cdot \cos(w \cdot t)] = h(t) \otimes [(1 + b \cdot A(t)) \cdot \sin(w \cdot t + \theta) \cdot \cos(w \cdot t)], 0 \leq \theta \leq 2 \cdot \pi \\
    x(t) &= h(t) \otimes [y(t) \cdot \cos(w \cdot t)] = h(t) \otimes [(1 + b \cdot A(t)) \cdot 0.5 \cdot \sin(2 \cdot w \cdot t + \theta) + \sin(\theta)], 0 \leq \theta \leq 2 \cdot \pi 
\end{align*}
\]

If the filter removes the 2wt term, we have

\[ x(t) = h(t) \otimes [y(t) \cdot \cos(w \cdot t)] = \left( \frac{1 + b \cdot A(t)}{2} \right) \cdot \sin(\theta), \quad 0 \leq \theta \leq 2 \cdot \pi \]

Notice that based on the value of the random variable, the output can change significantly! From producing no output signal, \((\theta = 0, \pi)\), to having the output be positive or negative \((\theta = 0 \text{ to } \pi \text{ or } \pi \text{ to } 2 \pi)\). P.S. This is not how you perform non-coherent AM demodulation.

To perform coherent AM demodulation, all I need to do is measured the value of the random variable and use it to insure that the output is a maximum (i.e. mix with \(\cos(w \cdot t + \theta_m)\), where. \(\theta_m = \theta(t_i)\))

Note: the phase is a function of frequency, time, and distance from the transmitter.
From our textbook

Random Stochastic Sequence

Definition 8.1-1. Let \( (\Omega, \mathcal{F}, P) \) be a probability space. Let \( \zeta \in \Omega \). Let \( X[n, \Omega] \) be a mapping of the sample space \( \Omega \) into a space of complex-valued sequences on some index set \( Z \). If, for each fixed integer \( n \in Z \), \( X[n, \zeta] \) is a random variable, then \( X[n, \Omega] \) is a random (stochastic) sequence. The index set \( Z \) is all integers, \( -\infty < n < +\infty \), padded with zeros if necessary.

Example sets of random sequences.

Figure 8.1-1 Illustration of the concept of random sequence \( X(n, \zeta) \), where the \( \zeta \) domain (i.e., the sample space \( \Omega \)) consists of just ten values. (Samples connected only for plot.)

The sequences can be thought of as “realizations” of the random sequence or sample sequences.

The absolute sequence is the realization of individual random variables in time.

- One the realization exists; it becomes statistical data related to one instantiation of the Random Sequence.
- Prior to collecting a realization, the Random Sequence can be defined probabilistically.
Example 8.1-1&2

Separable random sequences may be constructed by combining a deterministic sequence with one or more random variables.

The classic example already shown is a sinusoid with random amplitude and phase:

\[ X[n, \zeta] = A(\zeta) \cdot \sin \left( 2\pi \cdot \frac{1}{20} \cdot n + \Theta(\zeta) \right) \]

Where the amplitude and phase are R.V. defined based on the probability space selected.

Example 8.1-3

A random sequence with finite support is limited in time or samples.

\[ X[n, \zeta] = \begin{cases} X_n(\zeta), & 1 \leq n \leq N \\ 0, & \text{else} \end{cases} \]

Definition 8.1-2 Independent Random Sequence

An independent random sequence is one whose random variables at any time are jointly independent for all positive integers.

\[ E(X_i(\zeta) \cdot X_j(\zeta)) = \begin{cases} \sigma^2 + \mu^2, & \text{for } i = j \\ \mu^2, & \text{for } i \neq j \end{cases} \]
Figure 8.1-3  Example of a sample sequence of a random sequence.  (Samples connected only for plot.)

Figure 8.1-4  Close-up view of portion of sample sequence.
Infinite-length Bernoulli Trials

Defining an infinite space consisting of all possible sequential outcomes of the R.V over a defined interval. It also defines the probability of the infinite space as the product of the probability that generated one such instantiation.

$$\Omega_{\infty} = \bigotimes_{n=1}^{\infty} \Omega_n$$

the infinite cross product that the points in the infinite space consist of all the infinite-length sequences of events.

Example 8.1-5: Building correlated noise

For a Bernoulli random sequence W(m) just described, We form the “filtered” output

$$X[n] = \sum_{m=1}^{n} a^{-m} \cdot W(m), \quad \text{for } n \geq 1$$

For the Bernoulli random sequence, we know that $$P[W(m) = 1] = p$$ and $$P[W(m) = 0] = 1 - p$$.

We can define the mean of the output as

$$E[X[n]] = E\left[\sum_{m=1}^{n} a^{-m} \cdot W(m)\right] = \sum_{m=1}^{n} a^{-m} \cdot E[W(m)]$$

$$E[X[n]] = \sum_{m=1}^{n} a^{-m} \cdot p = a^n \cdot p \cdot \sum_{m=1}^{n} a^{-m} = a^n \cdot p \cdot a^{-1} \cdot \sum_{m=0}^{n-1} a^{-m}$$

$$E[X[n]] = a^n \cdot p \cdot a^{-1} \cdot \frac{1 - a^{-n}}{1 - a^{-1}} = p \cdot \left(\frac{a^n - 1}{a - 1}\right) = p \cdot \left(\frac{1 - a^n}{a - 1}\right)$$

The new sequence created involved the sum of random variables and as such is a random sequence; however, the individual elements are no longer IID.

We compute the cross-product of the first two terms:

$$E[X[2] \cdot X[1]] = E[(a \cdot W(1) + W(2)) \cdot W(1))$$

$$E[X[2] \cdot X[1]] = a \cdot E[W(1)^2] + E[W(2) \cdot W(1)]$$

$$E[X[2] \cdot X[1]] = a \cdot p + p \cdot p$$

Note that

\[ E[X[1]] = p \cdot \left( \frac{1-a^1}{1-a} \right) = p \quad \text{and} \quad E[X[2]] = p \cdot \left( \frac{1-a^2}{1-a} \right) = p \cdot (1+a) \]

Therefore,

\[ E[X[2] \cdot X[1]] = a \cdot p + p \cdot p = p^2 \cdot (1+a) = E[X[1]] \cdot E[X[2]] \]

The outputs are correlated (as you might expect).

The variance of the new sequence can be computed

\[ Var[X[n]] = Var \left[ \sum_{m=1}^{n} a^{n-m} \cdot W(m) \right] = \sum_{m=1}^{n} a^{2(n-m)} \cdot Var[W(m)] \]

\[ Var[X[n]] = \left( \frac{1-a^{2n}}{1-a^2} \right) \cdot p \cdot q = \left( \frac{1-a^{2n}}{1-a^2} \right) \cdot p \cdot (1-p) \]

The random outputs can be generated recursively as

\[ X[n] = \sum_{m=1}^{n} a^{n-m} \cdot W(m), \quad \text{for} \ n \geq 1 \]

\[ X[n] = a^0 \cdot W(n) + a \cdot \sum_{m=1}^{n-1} a^{n-1-m} \cdot W(m), \quad \text{for} \ n \geq 1 \]

\[ X[n] = W(n) + a \cdot X[n-1], \quad \text{for} \ n \geq 1 \]

Then

![Feedback Filter Diagram](image)

**Figure 8.1-5** A feedback filter that generates correlated noise X[n] from an uncorrelated sequence W[n].


B.J. Bazuin, Fall 2016 8 of 54 ECE 3800
Matlab simulation of the autoregressive sequence

Data Generation
numsamples = 400;
p = 0.50; \% X(0) prob (1-p) x(1) prob p
q = 1-p;

x = zeros(numsamples,1);

u = rand(numsamples,1);
w = double(u>=q);

As a Bernoulli R.V. p and q are defined as well as defining P[0]=q and P[1]=p.

The Auto-Regressive (AR) Sequence is generated as
\% The autoregressive function
alpha = 0.95;

x(1) = w(1);
for ii=2:numsamples
    x(ii) = alpha\*x(ii-1)+w(ii);
end

\% The IIR filter implementation
b = 1.0;
a = \{1.0 -alpha\};

xx = filter(b,a,w);

It can be done literally based on the regressive equation or as an IIR filter …
Knowing something about an IIR filter, the maximum output for a unity gain input is expected to be ….

\[ X[n \to \infty] = \sum_{m=1}^{n} a^{n-m} \cdot 1 = \frac{1}{1-a} = \frac{1}{1-0.95} = 20 \]

With mean value of the input of 0.5, I would expect results to be less than 20 and possible randomly moving about 10 …

Note that the R.V. statistics are generated from the data for comparison purposes …

\[ E[W[n]] = p \]
\[ Var[W[n]] = p \cdot (1 - p) \]

\[ E[X[n]] = p \cdot \left( \frac{1-a^n}{1-a} \right) \]
\[ Var[X[n]] = \left( \frac{1-a^{2n}}{1-a^2} \right) \cdot p \cdot (1 - p) \]
Continuity of Probability Measure  

Relating to Countable Additively...

\[ P \left[ \bigcup_{n=1}^{N} A_n \right] = \sum_{n=1}^{N} P[A_n] \]

for a mutually exclusive infinite collection of events.

**For an increasing sequence of events** where each \( B_n \) includes all previous \( B_j \) for \( j<n \) or

\[ B_n \subset B_{n-1} \subset \ldots \subset B_1 \]

We can define

\[ B_\infty = \bigcup_{n=1}^{N} B_n \]

and define proper disjoint sets such that

\[ A_1 = B_1 \quad \text{and} \quad A_n = B_n \cdot B_{n-1}^C \]

Then

\[ P[B_n] = P \left[ \bigcup_{n=1}^{N} B_n \right] = P \left[ \bigcup_{n=1}^{N} A_n \right] = \sum_{n=1}^{N} P[A_n] \]

This is a total probability statement and a statement of subdividing the inclusive area \( B \) into disjoint subsets.

**Inverse Corollary**

**For decreasing sequence of events** where each \( B_n \) is a subset of all previous \( B_j \) for \( j<n \) or

\[ B_n \supset B_{n-1} \supset \ldots \supset B_1 \]

\[ B_\infty = \bigcap_{n=1}^{N} B_n \]
Example 8.1-8 CDF continuity from the right at step discontinuities …

Using the decreasing sequence of events concept …

\[ B_n \supseteq B_{n-1} \supseteq \ldots \supseteq B_1 \]

For a CDF

\[ \lim_{n \to \infty} F_X \left( x + \frac{1}{n} \right) = F_X(x) \]

But for every successive value of n, we can define

\[ B_n = \left\{ \zeta : X(\zeta) \leq x + \frac{1}{n} \right\} \]

With

\[ P[B_n] = P\left[ \bigcap_{n=1}^{N} B_n \right] \]

Leading to

\[ \lim_{n \to \infty} F_X \left( x + \frac{1}{n} \right) = \lim_{n \to \infty} P[B_n] = P[B_\infty] = F_X(x) \]
**Statistical Specification of a Random Sequence**

In general we are looking developing properties for developing random processes where:

1. The statistical specification for the random sequence matches the probabilistic (or axiomatic) specification for the random variables used to generate the sequence.

2. We will be interested in stationary sequences where the statistics do not change in time. We will be defining a wide-sense stationary random process definition where only the mean and the variance need to be constant in time.

A random sequence $X[n]$ is said to be statistically specified by knowing its $N$th-order CDFs for all integers $N \geq 1$, and for all times, $n$, $n+1$, \ldots, $n+N-1$. That states that we know ...

$$F_X(x_n, x_{n+1}, \ldots, x_{n+N-1}; n, n+1, \ldots, n + N - 1) = P[X(n) \leq x_n, X(n+1) \leq x_{n+1}, \ldots, X(n+N-1) \leq x_{n+N-1}]$$

If we specify all these infinite-order joint distributions at all finite times, using continuity of the probability measures, we can calculate the probabilities of events involving an infinite number of random variables via limiting operations involving the finite order CDFs.

Consistency can be guaranteed by construction \ldots constructing models of stochastic sequences and processes.

Moments play an important role and, for Ergodic Sequences, they can be estimated from a single sample sequence of the infinite number that may be possible.

Therefore,

$$\mu_X[n] = E[X[n]] = \int_{-\infty}^{\infty} x \cdot f_X(x; n) \cdot dx$$

$$= \int_{-\infty}^{\infty} x_n \cdot f_X(x_n) \cdot dx_n$$

and for a discrete valued random sequences

$$\mu_X[n] = E[X[n]] = \sum_{k=-\infty}^{\infty} x_k \cdot P(X[n] = x_k)$$

The Autocorrelation Function

The expected value of a random sequence evaluated at offset times can be determined.

\[ R_{kk}[k, l] = E[X[k] \cdot X[l]] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_k \cdot x_l \cdot f_x(x_k, x_l) \cdot dx_k \cdot dx_l \]

For sequences of finite average power \( E[|X[k]|^2] < \infty \), then the correlation function will exist.

We can also describe the “centered” autocorrelation sequence as the autocovariance.

\[ K_{kk}[k, l] = E[(X[k] - \mu_X[k]) \cdot (X[l] - \mu_X[l])^*] \]

Note that

\[ K_{kk}[k, l] = E[(X[k] - \mu_X[k]) \cdot (X[l] - \mu_X[l])^*] \]

\[ = E[X[k] \cdot X[l]^*] - X[k] \cdot \mu_X[l]^* - \mu_X[k] \cdot X[l]^* + \mu_X[k] \cdot \mu_X[l]^* \]

\[ = E[X[k] \cdot X[l]^*] - \mu_X[k] \cdot \mu_X[l]^* \]

\[ K_{kk}[k, l] = R_{kk}[k, l] - \mu_X[k] \cdot \mu_X[l]^* \]

**Basic properties of the function:**

Hermitian symmetry \( R_{kk}[k, l] = R_{kk}[l, k]^* \)

Hermitian symmetry \( K_{kk}[k, l] = K_{kk}[l, k]^* \)

Deriving other functions

\[ R_{kk}[k, k] = E[|X[k]|^2] \]

\[ K_{kk}[k, k] = \sigma_X^2 \]
Example 8.1-1 & 10 functions consisting of R.V and deterministic sequences

Let

\[ X[n, \zeta] = X(\zeta) \cdot f(n) \]

where \( X \) is a random variable and \( f \) is a deterministic function in sample time \( n \).

Note then,

\[ E[X[n, \zeta]] = E[X(\zeta)] \cdot f(n) = \mu_X \cdot f(n) \]

The autocorrelation function becomes

\[ R_{kk}[k,l] = E[X[k, \zeta] \cdot X[l, \zeta]^*] = \int_{-\infty}^{\infty} [X(\zeta) \cdot f(k)] \cdot [X(\zeta) \cdot f(l)]^* \cdot f_X(x) \cdot dx \]

If \( X \) is a real R.V.

\[ R_{kk}[k,l] = f(k) \cdot f(l)^* \cdot (\sigma_X^2 + \mu_X^2) \]

Similarly

\[ K_{xx}[k,l] = f(k) \cdot f(l)^* \cdot E[(X(\zeta) - \mu_X) \cdot (X(\zeta) - \mu_X)^*] \]

Example 8.1-11 Waiting times in a line

consider the random sequence of IID “exponential random variable” waiting times in a line. Assume that each of the waiting times per individual \( t(k) \) is based on the exponential.

\[ f_r(t; n) = f_r(t) = \begin{cases} 0, & t < 0 \\ \lambda \cdot \exp(-\lambda \cdot t), & t \geq 0 \end{cases} \]

The waiting time is then described as

\[ T(n) = \sum_{k=1}^{n} \tau[k] \]

where 

\[ T(1) = \tau[1], \quad T(2) = \tau[1] + \tau[2], \ldots, T(n) = \tau[1] + \tau[2] + \ldots + \tau[n] \]
This calls for a summation of random variables, where the new pdf for each new sum is the convolution of the exponential pdf with the previous pdf or

\[ f_t(t;2) = f_t(t) * f_t(t) \]

\[ f_t(t;3) = f_t(t;2) * f_t(t) = [f_t(t)*f_t(t)]*f_t(t) \]

Exam #1 derived the first convolution

\[
f_t(t;2) = \int_0^t \lambda \cdot \exp(-\lambda \cdot \tau) \cdot \lambda \cdot \exp(-\lambda \cdot (t - \tau)) \cdot d\tau
\]

\[
f_t(t;2) = \lambda^2 \cdot \exp(-\lambda \cdot t) \cdot \int_0^t 1 \cdot d\tau = \lambda^2 \cdot t \cdot \exp(-\lambda \cdot t)
\]

Repeating

\[
f_t(t;3) = \int_0^t \lambda^2 \cdot \tau \cdot \exp(-\lambda \cdot \tau) \cdot \lambda \cdot \exp(-\lambda \cdot (t - \tau)) \cdot d\tau
\]

\[
f_t(t;3) = \lambda^3 \cdot \exp(-\lambda \cdot t) \cdot \int_0^t \tau \cdot d\tau = \lambda^3 \cdot \frac{\tau^2}{2} \cdot \exp(-\lambda \cdot t)
\]

If you see the pattern …. we can jump to the nth summation where

\[
f_t(t;n) = \lambda^n \cdot \frac{\tau^{n-1}}{(n-1)!} \cdot \exp(-\lambda \cdot t) = \frac{(\lambda \cdot \tau)^{n-1}}{(n-1)!} \cdot \lambda \cdot \exp(-\lambda \cdot t)
\]

This is called the Erlang probability density function …

It is used to determine waiting times in lines and software queues … how long until your internet request can be processed!

The mean and variance of the Erlang pdf is

\[
\mu_T = n \cdot \mu_t = \frac{n}{\lambda}
\]

\[
\sigma_T^2 = n \cdot Var[\tau] = \frac{n}{\lambda^2}
\]
Gaussian Random Sequences

There are a few developments to consider:
- The sequence of Gaussian R.V. \( W[n] \), \( \mu_W = 0 \)

A random sequence based on the Gaussian.
Assume iid Gaussian R.V. with zero mean and a variance
For \( E[W[n]] = \mu_W = 0 \) and \( E[W[n]^2] = \sigma_W^2 \)

What about the autocorrelation
\[
R_{kk}[k,l] = E[W[k] \cdot W[l]^*] = \begin{cases} 
E[W[k]^2] = \sigma_W^2 + \mu_W^2 = \sigma_W^2, & k = l \\
\mu_W^2 = 0 & k \neq l 
\end{cases}
\]
or it can be written as
\[ R_{kk}[k,l] = \sigma_W^2 \cdot \delta(k-l) \]

A random sequence based on the sum of two Gaussians.
Assume iid Gaussian R.V. with zero mean and a variance
For \( X[n] = W[n] + W[n-1] \)

Then,
\[
E[X[n]] = E[W[n] + W[n-1]] = \mu_W + \mu_W = 2 \cdot \mu_W = 0
\]
\[
= \sigma_W^2 + 2 \cdot \mu_W^2 + \sigma_W^2 \\
= 2 \cdot \sigma_W^2
\]

also
\[
R_{kk}[k,l] = E[(W[k] + W[k-1]) \cdot (W[l] + W[l-1])^*] \\
But then \( R_{kk}[k,l] = \sigma_W^2 \cdot \delta(k-l) + \sigma_W^2 \cdot \delta(k-l+1) + \sigma_W^2 \cdot \delta(k-1-l) + \sigma_W^2 \cdot \delta(k-1+l-1) \)
\]
\[
R_{kk}[k,l] = 2 \cdot \sigma_W^2 \cdot \delta(k-l) + \sigma_W^2 \cdot \delta(k-l+1) + \sigma_W^2 \cdot \delta(k-l-1)
\]
Note that this is a preferred structure for an autocorrelation function …

\[ R_{kk}[k,l] = R_{kk}[k-l] = 2 \cdot \sigma_w^2 \cdot \delta(k-l) + \sigma_w^2 \cdot \delta(k-l+1) + \sigma_w^2 \cdot \delta(k-l-1) \]

which can be written as

\[ R_{kk}[k-l] = R_{kk}[n] = 2 \cdot \sigma_w^2 \cdot \delta(n) + \sigma_w^2 \cdot \delta(n+1) + \sigma_w^2 \cdot \delta(n-1) \]

The result shows that the time samples are not important individually, only the distance between the time samples!

**Example 8.1-13 Random Walk**

Continuing with infinite length Bernoulli trials, we will define taking steps either forward or backward and see if there is any progress in time.

Therefore let

\[ W[k] = \begin{cases} +s, & \zeta = \text{Heads} \\ -s, & \zeta = \text{Tails} \end{cases} \]

\[ X(n) = \sum_{k=1}^{n} W[k] \]

The current position is based on the sum of time history of heads and tails. Let,

\[ X(n) = r \cdot s = k \cdot s - (n-k) \cdot s = (2 \cdot k - n) \cdot s \]

Where there would be k heads and n-k tails. Based on the current position, we can “solve for k” in the Bernoulli equation as

\[ k = \frac{r+n}{2} \]

But k is based on the Binomial R.V.

\[ p_k = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}, \quad \text{for } k = 0, 1, \ldots, n \]

\[ P\{X(n) = r \cdot s\} = P\left[ \frac{n+r}{2} \right] = \begin{cases} \binom{n}{n+r/2} \cdot p^{n+r/2} \cdot (1-p)^{n-r/2}, & n+r/2 \text{ an integer with } r \leq n \\ 0, & \text{else} \end{cases} \]
We can find the mean and variance
\[ E[X(n)] = \sum_{k=1}^{n} E[W[k]] = \sum_{k=1}^{n} (p \cdot s + q \cdot (-s)) = n \cdot (2 \cdot p - 1) \cdot s \]

Assuming equal probability
\[
P\{X(n) = r \cdot s\} = P\left[\frac{n+r}{2}\right] = \begin{cases} \binom{n}{\frac{n+r}{2}} \cdot 2^{-2}, & n+r/2 \text{ an integer with } r \leq n \\ 0, & \text{else} \end{cases}
\]

\[ E[X(n)] = n \cdot (2 \cdot p - 1) \cdot s = 0 \]

For W[n] IID
\[
E[X(n)^2] = E\left[\sum_{k=1}^{n} W[k] \cdot \sum_{j=1}^{n} W[j]\right] = E\left[\sum_{k=1}^{n} W[k]^2 + \sum_{k=1}^{n} \sum_{j=k+1}^{n} W[k] \cdot W[j]\right]
\]

\[ E[X(n)^2] = \sum_{k=1}^{n} E[W[k]^2] + \sum_{k=1}^{n} \sum_{j=k+1}^{n} E[W[k] \cdot W[j]] \]

\[
E[X(n)^2] = \sum_{k=1}^{n} [p \cdot s^2 + q \cdot (-s)^2] + (n^2 - n) \cdot \mu_w^2
\]

\[ E[X(n)^2] = n \cdot (p + q) \cdot s^2 = n \cdot s^2 \]

If we define a “normalized” random variable
\[ \tilde{X}(n) = \frac{1}{\sqrt{n}} \cdot X(n) \]

The normalized random variable based on the Central Limit Theorem would be equivalent to a Gaussian Normal with standard deviation s!

A probability bound on the range of X(n) could then be defined as …
\[ P[a < \tilde{X}(n) < b] = \Phi\left(\frac{b}{s}\right) - \Phi\left(\frac{a}{s}\right) \]

and
\[ P[a < \tilde{X}(n) < b] = P[\sqrt{n} \cdot a < X(n) < \sqrt{n} \cdot b] = \Phi\left(\frac{b}{s}\right) - \Phi\left(\frac{a}{s}\right) \]
Unique properties of Erlang and Random Walk

The Erlang waiting time and Random Walk have a property called “independent-increments”, where the increments in value could be considered jointly independent. Unfortunately, the mean, the variance or both varied based on the time increments.

*For many applications, we prefer that the probability is non-time or sample varying ... that the moments are stationary.*

Non-stationary means and variances are considerably harder to deal with in signal processing …

This is indicative of classes of sequences and processes … those that are stationary and those the vary as time or the number of samples increases.
Stationary vs. Nonstationary Random Processes

The probability density functions for random variables in time have been discussed, but what is the dependence of the density function on the value of time, t or n, when it is taken?

If all marginal and joint density functions of a process do not depend upon the choice of the time origin, the process is said to be stationary (that is it doesn’t change with time). All the mean values and moments are constants and not functions of time!

For nonstationary processes, the probability density functions change based on the time origin or in time. For these processes, the mean values and moments are functions of time.

In general, we always attempt to deal with stationary processes … or approximate stationary by assuming that the process probability distribution, means and moments do not change significantly during the period of interest.

Examples:
- Resistor values (noise varies based on the local temperature)
- Wind velocity (varies significantly from day to day)
- Humidity (though it can change rapidly during showers)

The requirement that all marginal and joint density functions be independent of the choice of time origin is frequently more stringent (tighter) than is necessary for system analysis.

A more relaxed requirement is called stationary in the wide sense: where the mean value of any random variable is independent of the choice of time, t, and that the correlation of two random variables depends only upon the time difference between them. That is

\[ E[X(t)] = \bar{X} = \mu_X \quad \text{and} \]

\[ E[X(t_1) \cdot X(t_2)] = E[X(0) \cdot X(t_2 - t_1)] = \bar{X}(0) \cdot \bar{X}(\tau) = R_{XX}(\tau) \quad \text{for} \quad \tau = t_2 - t_1 \]

You will typically deal with **Wide-Sense Stationary Signals**.

See Definition 8.1-5, p. 464, for the text definition of stationary. The definition of wide-sense stationary (WSS) random sequence also appears on p. 465, where the mean and the covariance of a random sequence are not based on time or sample number.
Stationary Systems Properties

Mean Value

\[ \mu_X(n) = E[X(n)] = \int_{-\infty}^{\infty} x \cdot f_X(x;n) \cdot dx = \int_{-\infty}^{\infty} x \cdot f_X(x;0) \cdot dx = \mu_X(0) \]

The mean value is not dependent upon the sample in time.

Autocorrelation

\[ R_{KK}[k,l] = E[X(k) \cdot X(l)^*] = \int_{-\infty}^{\infty} x_k \cdot x_l^* \cdot f_X(x_k,x_l;k,l) \cdot dx_k \cdot dx_l \]

\[ R_{KK}[k,l] = \int_{-\infty}^{\infty} x_{k+n} \cdot x_{l+n}^* \cdot f_X(x_{k+n},x_{l+n};k+n,l+n+1) \cdot dx_{k+n} \cdot dx_{l+n+1} \]

\[ R_{KK}[k,l] = E[X(k+n) \cdot X(l+n)^*] = R_{KK}[k+n,l+n] \]

And in particular

\[ R_{KK}[k,l] = R_{KK}[k-l,0] = R_{KK}[k-l] \Rightarrow R_{KK}[k,0] = R_{KK}[k] \]

Autocovariance

\[ K_{KK}[k,l] = E[(X(k) - \mu_X(k)) \cdot (X(l) - \mu_X(l))^*] = E[(X(k) - \mu_X(k)) \cdot (X(l) - \mu_X(l))^*] \]

\[ = E[(X(k+n) - \mu_X(k)) \cdot (X(l+n) - \mu_X(l))^*] = K_{KK}[k+n,l+n] \]

And in particular

\[ K_{KK}[k,l] = K_{KK}[k-l,0] = K_{KK}[k-l] \Rightarrow K_{KK}[k,0] = K_{KK}[k] \]

The autocorrelation and autocovariance are functions of the time difference and not the absolute time.

Of course as an example the text starts with something confusing and difficult …
Example 8.1-15 A sequence with memory ....

A two random state function.

\[
X[k] = a \quad \text{and} \quad \sim X[k] = b
\]

Let

\[
X[n] = \begin{cases} 
X[n-1] & \text{with } P = p \\
\sim X[n-1] & \text{with } P = q = 1 - p
\end{cases}
\]

Mean value

\[
E[X[1]] = E[p \cdot X[0] + q \cdot \sim X[0]] = p \cdot a + q \cdot b
\]

\[
E[\sim X[1]] = E[p \cdot \sim X[0] + q \cdot X[0]] = p \cdot b + q \cdot a
\]

\[
E[X[2]] = E[p \cdot X[1] + q \cdot \sim X[1]] = p^2 \cdot a + 2 \cdot p \cdot q \cdot b + q^2 \cdot a
\]

\[
E[\sim X[2]] = E[p \cdot \sim X[1] + q \cdot X[1]] = p^2 \cdot b + 2 \cdot p \cdot q \cdot a + q^2 \cdot b
\]

\[
E[X[3]] = E[p \cdot X[2] + q \cdot \sim X[2]] = p^3 \cdot a + 2 \cdot p^2 \cdot q \cdot b + p \cdot q^2 \cdot a + q \cdot p^2 \cdot b + 2 \cdot p \cdot q^2 \cdot a + q^3 \cdot b
\]

\[
E[X[3]] = E[p \cdot X[2] + q \cdot \sim X[2]] = (p^3 + 2 \cdot p^2 \cdot q^2 + p^2 \cdot q^2) \cdot a + (q^3 + 2 \cdot p^2 \cdot q + q^2 \cdot p^2) \cdot b
\]

\[
E[X[3]] = E[p \cdot X[2] + q \cdot \sim X[2]] = p \cdot (p + q)^2 \cdot a + q \cdot (p + q)^2 \cdot b = p \cdot a + q \cdot b = E[X[1]]
\]

\[
E[\sim X[3]] = E[p \cdot \sim X[2] + q \cdot X[2]] = p \cdot b + q \cdot a = E[\sim X[1]]
\]

The expected value “oscillates” from even to odd and back again ...
There is a binomial growth of probabilities based on the initial state chosen. When N is odd, all products of two transitions or an integer number time two transitions will be a. All others will be b.

As an interesting note … when \( p=\frac{1}{2} \) and N is odd, \( E[X[1]] = E[X[odd]] = p \cdot b + q \cdot a = \frac{1}{2} \cdot b + \frac{1}{2} \cdot a \)

Furthermore

\[
E[X[2]] = E[X[even]] = \frac{1}{4} \cdot a + \frac{1}{4} \cdot b + \frac{1}{4} \cdot a = \frac{1}{2} \cdot a + \frac{1}{2} \cdot b
\]

The number of possible “even” transitions in state due to p and q are equal to the number of “odd” transitions in state due to p and q probabilities. But we know this from the Binomial pmf!

\[
\binom{N}{k} : \binom{3}{k} \Rightarrow 1,3,3,1 \quad \binom{5}{k} \Rightarrow 1,5,10,10,5,1
\]

Example 8.1-16 A sequence with memory …. Autocorrelation

\[
R_{kk}[k,n+k] = E[X(k) \cdot X(n+k)]
\]

\[
R_{kk}[k,n+k] = \begin{cases} 
  b \cdot b \cdot P(b,b;k,n+k) + a \cdot a \cdot P(a,a;k,n+k) + \\
  a \cdot b \cdot P(a,b;k,n+k) + b \cdot a \cdot P(b,a;k,n+k)
\end{cases}
\]

Note if we set \( a=0 \) this greatly simplifies ...

\[
R_{kk}[k,n+k] = b \cdot b \cdot P(b,b;k,n+k)
\]

As a sequence of events, for the second event following the first, we can deal with conditional probability … given \( X[k]=b \) then

\[
R_{kk}[k,n+k] = b \cdot b \cdot P(b,b;k,n+k) = b \cdot b \cdot P(b,k) \cdot P(b;k+n|b;k)
\]

or

\[
R_{kk}[k,n+k] = b \cdot b \cdot P(X[k]=b) \cdot P(X[k+n]=b | X[k]=b)
\]

Based on the mean value discussion

\[
P(X[k]=b) = 0.5
\]

\[
R_{kk}[k,n+k] = \frac{b^2}{2} \cdot P(X[k+n]=b | X[k]=b)
\]
But now, the conditional probability is based on an even number of state transitions in k steps. This is a summation of the Binomial probability (with an adjustment for “change” and “no change”).

\[ p_l = \binom{n}{l} \cdot (1 - p)^l \cdot p^{n-l} \]

\[ P(X[k+n] = b \mid X[k] = b) = \sum_{l \text{ even}}^{n} \binom{n}{l} \cdot (1 - p)^l \cdot p^{n-l} \]

Using the textbook trick

Note, I haven’t seen this result before

\[ \sum_{l=0}^{n} \binom{n}{l} \cdot (p-1)^l \cdot p^{n-l} = (2p-1)^n \]

\[ P(X[k+n] = b \mid X[k] = b) = \sum_{l \text{ even}}^{n} \binom{n}{l} \cdot (1 - p)^l \cdot p^{n-l} = \frac{1}{2} \cdot [(2p-1)^n + 1] \]

And the result becomes

\[ R_{kk}[n] = \frac{b^2}{2} \cdot \frac{1}{2} \cdot [(2p-1)^n + 1] \]

\[ R_{kk}[n] = \frac{b^2}{4} \cdot (2p-1)^n + 1 \]

Note that the autocovariance is

\[ K_{kk}[n] = R_{kk}[n] - \mu_x^2 = \frac{b^2}{4} \cdot [(2p-1)^n + 1] - \frac{b^2}{4} \]

\[ K_{kk}[n] = \frac{b^2}{4} \cdot (2p-1)^n \]

We can tackle the more complex problem of a not equal to zero now … we needed Ae and Ao.

\[ R_{kk}[k, n + k] = \begin{cases} b \cdot b \cdot P(b; k) \cdot P(b; k + n \mid b; k) + a \cdot a \cdot P(a; k) \cdot P(a; k + n \mid a; k) + \\ a \cdot b \cdot P(a; k) \cdot P(b; k + n \mid a; k) + b \cdot a \cdot P(b; k) \cdot P(a; k + n \mid b; k) \end{cases} \]

\[ R_{kk}[k, n + k] = \begin{cases} b \cdot b \cdot P(b; k) \cdot A_e + a \cdot a \cdot P(a; k) \cdot A_e + \\ a \cdot b \cdot P(a; k) \cdot A_o + b \cdot a \cdot P(b; k) \cdot A_o \end{cases} \]
For

\[ A_E = \frac{1}{2} \left[ 1 + (2p - 1)^n \right] \]

\[ A_O = \frac{1}{2} \left[ 1 - (2p - 1)^n \right] \]

\[ R_{kk}[k, n + k] = \left\{ \left( \frac{b^2}{4} + \frac{a^2}{2} \right) \cdot A_E + a \cdot b \cdot A_O \right\} \]

\[ R_{kk}[k, n + k] = \left\{ \left( \frac{b^2}{4} + \frac{a^2}{2} \right) \cdot [1 + (2p - 1)^n] + \frac{a \cdot b}{2} \cdot [1 - (2p - 1)^n] \right\} \]

\[ R_{kk}[n] = \frac{(a + b)^2}{4} + \frac{(a - b)^2}{4} \cdot (2p - 1)^n \]

Note: the simulation has “flipped p and q so that

\[ R_{kk}[n] = \frac{(a + b)^2}{4} + \frac{(a - b)^2}{4} \cdot (2q - 1)^n = \frac{(a + b)^2}{4} + \frac{(a - b)^2}{4} \cdot (1 - 2p)^n \]

p=0.75, a=-1, b=1

p=0.25, a=-1, b=1


B.J. Bazuin, Fall 2016 26 of 54 ECE 3800
Note: the following examples are going to use continuous time instead of discrete time. These examples could be made discrete, but the concept remains the same. In the textbook, the continuous time derivation is in Chapter 9.

Example: \( x(t) = A \cdot \sin(2\pi \cdot f \cdot t + \theta) \) for \( \theta \) a uniformly distributed random variable \( \theta \in [0, 2\pi] \)

\[
R_{xx}(t_1, t_2) = E[X(t_1) \cdot X(t_2)] = E[A \cdot \sin(2\pi \cdot f \cdot t_1 + \theta) \cdot A \cdot \sin(2\pi \cdot f \cdot t_2 + \theta)]
\]

\[
R_{xx}(t_1, t_2) = E[X(t_1) \cdot X(t_2)] = A^2 \cdot E\left[\frac{1}{2} \cdot \{\cos(2\pi \cdot f \cdot (t_1 - t_2)) - \cos(2\pi \cdot f \cdot (t_1 + t_2) + 2 \cdot \theta)\}\right]
\]

\[
R_{xx}(t_1, t_2) = E[X(t_1) \cdot X(t_2)] = \frac{A^2}{2} \cdot \cos(2\pi \cdot f \cdot (t_1 - t_2)) - \frac{A^2}{2} \cdot E[\cos(2\pi \cdot f \cdot (t_1 + t_2) + 2 \cdot \theta)]
\]

Of note is that the phase need only be uniformly distributed over 0 to \( \pi \) in the previous step!

\[
R_{xx}(t_1, t_2) = E[X(t_1) \cdot X(t_2)] = \frac{A^2}{2} \cdot \cos(2\pi \cdot f \cdot (t_1 - t_2))
\]

for \( \tau = t_1 - t_2 \)

\[
R_{xx}(\tau) = \frac{A^2}{2} \cdot \cos(2\pi \cdot f \cdot \tau)
\]

but

\[
R_{xx}(\tau) = R_{xx}(-\tau) = \frac{A^2}{2} \cdot \cos(2\pi \cdot f \cdot (-\tau))
\]

Assuming a uniformly distributed random phase “simplifies the problem” …

Also of note, if the amplitude is an independent random variable, then

\[
R_{xx}(\tau) = \frac{E[A^2]}{2} \cdot \cos(2 \cdot \pi \cdot f \cdot \tau)
\]
**Example:** \( x(t) = B \cdot \text{rect}\left(\frac{t-t_0}{T}\right) \) for \( B = \pm A \) with probability \( p \) and \( (1-p) \) and \( t_0 \) a uniformly distributed random variable \( t_0 \in \left[ -\frac{T}{2}, \frac{T}{2} \right] \). Assume \( B \) and \( t_0 \) are independent.

\[
R_{xx}(t_1, t_2) = E[X(t_1) \cdot X(t_2)] = E\left[ B \cdot \text{rect}\left(\frac{t_1-t_0}{T}\right) \cdot \text{rect}\left(\frac{t_2-t_0}{T}\right) \right]
\]

\[
R_{xx}(t_1, t_2) = E[X(t_1) \cdot X(t_2)] = E\left[ B^2 \cdot \text{rect}\left(\frac{t_1-t_0}{T}\right) \cdot \text{rect}\left(\frac{t_2-t_0}{T}\right) \right]
\]

As the RV are independent

\[
R_{xx}(t_1, t_2) = E[X(t_1) \cdot X(t_2)] = E[B^2] \cdot E\left[ \text{rect}\left(\frac{t_1-t_0}{T}\right) \cdot \text{rect}\left(\frac{t_2-t_0}{T}\right) \right]
\]

\[
R_{xx}(t_1, t_2) = (A^2 \cdot p + (-A)^2 \cdot (1-p)) \cdot E\left[ \text{rect}\left(\frac{t_1-t_0}{T}\right) \cdot \text{rect}\left(\frac{t_2-t_0}{T}\right) \right]
\]

\[
R_{xx}(t_1, t_2) = A^2 \cdot \int_{-\frac{T}{2}}^{\frac{T}{2}} \text{rect}\left(\frac{t_1-t_0}{T}\right) \cdot \text{rect}\left(\frac{t_2-t_0}{T}\right) \cdot \frac{1}{T} \cdot dt_0
\]

For \( t_1 = 0 \) and \( t_2 = \tau \)

\[
R_{xx}(0, \tau) = A^2 \cdot \int_{-\frac{T}{2}}^{\frac{T}{2}} 1 \cdot \text{rect}\left(\frac{\tau-t_0}{T}\right) \cdot \frac{1}{T} \cdot dt_0
\]

The integral can be recognized as being a triangle, extending from \(-T\) to \(T\) and zero everywhere else.

\[
R_{xx}(\tau) = A^2 \cdot \text{tri}\left(\frac{\tau}{T}\right)
\]

\[
R_{xx}(\tau) = \begin{cases} 
0, & \tau \leq -T \\
A^2 \cdot \frac{1}{T} \cdot (T + \tau), & -T < \tau \leq 0 \\
A^2 \cdot \frac{1}{T} \cdot (T - \tau), & 0 < \tau \leq T \\
0, & T < \tau 
\end{cases}
\]
Properties of Autocorrelation Functions (with WSS property)

1) \[ R_{XX}(0) = E[XX] = \bar{X}^2 \quad \text{or} \quad R_{XX}(0) = \langle x(t)^2 \rangle \]

the mean squared value of the random process can be obtained by observing the zeroth lag of the autocorrelation function.

2) \[ R_{XX}(\tau) = R_{XX}(-\tau) \]

The autocorrelation function is an even function in time. Only positive (or negative) needs to be computed for an ergodic, WSS random process. (Conjugate symmetry if complex)

3) \[ |R_{XX}(\tau)| \leq R_{XX}(0) \]

The autocorrelation function is a maximum at 0. For periodic functions, other values may equal the zeroth lag, but never be larger.

4) If \( X \) has a DC component, the \( R_{XX} \) has a constant factor.

\[
X(t) = \bar{X} + N(t)
\]

\[
R_{XX}(\tau) = \bar{X}^2 + R_{NN}(\tau)
\]

Note that the mean value can be computed from the autocorrelation function constants!

5) If \( X \) has a periodic component, then \( R_{XX} \) will also have a periodic component of the same period.

Think of:

\[
X(t) = A \cdot \cos(w \cdot t + \theta), \quad 0 \leq \theta \leq 2 \cdot \pi
\]

where \( A \) and \( w \) are known constants and theta is a uniform random variable.

\[
R_{XX}(\tau) = E[X(t)X(t+\tau)] = \frac{A^2}{2} \cdot \cos(w \cdot \tau)
\]

5b) For signals that are the sum of independent random variable, the autocorrelation is the sum of the individual autocorrelation functions.

\[
W(t) = X(t) + Y(t)
\]

\[
R_{WW}(\tau) = R_{XX}(\tau) + R_{YY}(\tau) + 2 \cdot \mu_X \cdot \mu_Y
\]


B.J. Bazuin, Fall 2016

30 of 54

ECE 3800
For non-zero mean functions, (let w, x, y be zero mean and W, X, Y have a mean)

\[ R_{WW}(\tau) = R_{XX}(\tau) + R_{YY}(\tau) + 2 \cdot \mu_X \cdot \mu_Y \]

\[ R_{WW}(\tau) = R_{ww}(\tau) + \mu_W^2 = R_{xx}(\tau) + \mu_X^2 + R_{yy}(\tau) + \mu_Y^2 + 2 \cdot \mu_X \cdot \mu_Y \]

\[ R_{WW}(\tau) = R_{ww}(\tau) + \mu_W^2 = R_{xx}(\tau) + R_{yy}(\tau) + 2 \cdot \mu_X \cdot \mu_Y + \mu_Y^2 \]

Then we have

\[ R_{ww}(\tau) = R_{xx}(\tau) + R_{yy}(\tau) \]

6) If X is ergodic and zero mean and has no periodic component, then

\[ \lim_{|\tau| \to \infty} R_{XX}(\tau) = 0 \]

No good proof provided.

The logical argument is that eventually new samples of a non-deterministic random process will become uncorrelated. Once the samples are uncorrelated, the autocorrelation goes to zero.

7) Autocorrelation functions cannot have an arbitrary shape. One way of specifying shapes permissible is in terms of the Fourier transform of the autocorrelation function. That is, if

\[ \mathcal{Z}[R_{XX}(\tau)] = \int_{-\infty}^{\infty} R_{XX}(\tau) \cdot \exp(-jwt) \cdot dt \]

then the restriction states that

\[ \mathcal{Z}[R_{xx}(\tau)] \geq 0 \quad \text{for all } w \]
Additional concepts that are useful … dealing with constant:

\[ X(t) = a \cdot N(t) \]

\[ R_{XX}(\tau) = E[X(t)X(t + \tau)] = E[(a \cdot N(t)) \cdot (a \cdot N(t + \tau))] \]

\[ R_{XX}(\tau) = a^2 \cdot E[N(t) \cdot N(t + \tau)] = a^2 \cdot R_{NN}(\tau) \]

\[ X(t) = a + Y(t) \]

\[ R_{XX}(\tau) = E[(a + Y(t)) \cdot (a + Y(t + \tau))] \]

\[ R_{XX}(\tau) = E[a^2 + a \cdot Y(t) + a \cdot Y(t + \tau) + Y(t) \cdot Y(t + \tau)] \]

\[ R_{XX}(\tau) = a^2 + a \cdot E[Y(t)] + a \cdot E[Y(t + \tau)] + E[Y(t) \cdot Y(t + \tau)] \]

\[ R_{XX}(\tau) = a^2 + 2 \cdot a \cdot E[Y(t)] + R_{YY}(\tau) \]
8.2 Basic Principles of Discrete-Time Linear Systems

We get to do convolutions some more … in the discrete time domain!

Note: if you are in ECE 3710, this should be normal; otherwise, ECE 3100 probably talked about linear systems being a convolution.

For a “causal” discrete finite impulse response linear system we will have ….

\[ y(n) = \sum_{k=0}^{\infty} h(k) \cdot x(n-k) = \sum_{m=-\infty}^{n} h(n-m) \cdot x(m) \]

For a “non-causal” discrete linear system we will have ….

\[ y(n) = \sum_{k=-\infty}^{\infty} h(k) \cdot x(n-k) = \sum_{m=-\infty}^{\infty} h(n-m) \cdot x(m) \]

For a linear system, superposition applies

\[ y(n) = \sum_{k=0}^{\infty} h(k) \left[ a_1 \cdot x_1(n-k) + a_2 \cdot x_2(n-k) \right] \]

\[ = \sum_{k=0}^{\infty} h(k) \cdot a_1 \cdot x_1(n-k) + \sum_{k=0}^{\infty} h(k) \cdot a_2 \cdot x_2(n-k) \]

\[ = a_1 \cdot \sum_{k=0}^{\infty} h(k) \cdot x_1(n-k) + a_2 \cdot \sum_{k=0}^{\infty} h(k) \cdot x_2(n-k) \]

\[ = y_1(n) + y_2(n) \]

For a filter with poles and zeros …. the filter may be autoregressive as well!

\[ y(n) = \sum_{k=0}^{\infty} b(k) \cdot x(n-k) + \sum_{k=1}^{\infty} a(k) \cdot y(n-k) \]

This is called a linear constant coefficient difference equation (LCCDE) in the text.
**Solving difference equations**

Example 8.2-1

\[ y(n) = 1.7 \cdot y(n-1) - 0.72 \cdot y(n-2) + 1.0 \cdot u(n) \]

Need assumptions: \( y(-1)=0 \) and \( y(-2)=0 \)

First solve the recursive part of the equation (homogeneous equation) then solve for the input and combine the two solutions (superposition works).

\[ y_h(n) = 1.7 \cdot y_h(n-1) - 0.72 \cdot y_h(n-2) \]

A prototype solution is:

\[ y_h(n) = A \cdot r^n \]

Plugging it in …

\[ y_h(n) - 1.7 y_h(n-1) + 0.72 y_h(n-2) = 0 \]

\[ A \cdot r^n - 1.7 \cdot A \cdot r^{n-1} + 0.72 \cdot A \cdot r^{n-2} = 0 \]

\[ r^2 - 1.7 \cdot r + 0.72 = 0 \]

Solve for the roots

\[ r_{1,2} = \frac{b}{2 \cdot a} \pm \frac{\sqrt{b^2 - 4 \cdot a \cdot c}}{2 \cdot a} \]

\[ r_{1,2} = -\frac{1.7}{2} \pm \frac{\sqrt{(1.7)^2 - 4 \cdot 0.72}}{2} \]

\[ r_{1,2} = 0.85 \pm \frac{2.89 - 2.88}{2} = 0.85 \pm 0.05 = 0.8, 0.9 \]

Therefore, since the roots are less than one, the equation will be stable and …

\[ y_h(n) = A_1 \cdot 0.8^n + A_2 \cdot 0.9^n \]

Looking at the particular solution …

The input is a unit step function, \( x(n) = 1 \) for \( n \geq 0 \). There should be a steady state solution as \( n \) goes to infinity. If so, \( y(n)=y(n-1)-y(n-2) \) … for very large \( n \)

\[ y(\infty) - 1.7 \cdot y(\infty) + 0.72 \cdot y(\infty) = 1 \]

Therefore,

\[ y(\infty) \cdot (1 - 1.7 + 0.72) = 1 \]

\[ y(\infty) = \frac{1}{0.02} = 50 \]
Then,
\[ y(n) = y_h(n) + y(\infty) = A_1 \cdot 0.8^n + A_2 \cdot 0.9^n + 50 \]

Now for the specifics, determine the result for \( y(1) \) and \( y(2) \) and then solve for the two remaining unknowns in A. Given \( y(-1) = 0 \) and \( y(-2) = 0 \) …

\[ y(0) = 1.7 \cdot y(-1) - 0.72 \cdot y(-2) + 1.0 = 1.0 \]
\[ y(1) = 1.7 \cdot y(0) - 0.72 \cdot y(-1) + 1.0 = 1.7 + 1.0 = 2.7 \]

Using the prototype solution

\[ y(0) = A_1 \cdot 0.8^0 + A_2 \cdot 0.9^0 + 50 = 1 \]
\[ y(1) = A_1 \cdot 0.8^1 + A_2 \cdot 0.9^1 + 50 = 2.7 \]

You like doing a matrix solution?

\[
\begin{bmatrix}
1 & 1 \\
0.8 & 0.9
\end{bmatrix}
\begin{bmatrix}
A_1 \\
A_2
\end{bmatrix}
= 
\begin{bmatrix}
-49.0 \\
-47.8
\end{bmatrix}
\]

\[
\begin{bmatrix}
A_1 \\
A_2
\end{bmatrix}
= 
\begin{bmatrix}
0.9 & -1 \\
-0.8 & 1 \\
0.9 & -0.8
\end{bmatrix}
\begin{bmatrix}
-49.0 \\
-47.8
\end{bmatrix}
= 
\begin{bmatrix}
32.0 \\
-81.0
\end{bmatrix}
\]

and

\[ y(n) = 32 \cdot 0.8^n - 81 \cdot 0.9^n + 50, \quad \text{for } 0 \leq n \]

**Linear time invariant and linear shift invariant**

\[ y(n+k) = L[x(n+k)], \quad \text{for all } n \]

A time offset in the input will not change the response at the output! This is key to the convolution theory!

**System Impulse response**

The response to a unit impulse is the impulse response

\[ y(n) = L[\delta(n)] = h(n) \]
The discrete Fourier Transform exists (if \( x(n) \) is bounded).
Dr. Bazuin may also call this the discrete-time, continuous-frequency Fourier Transform (something from ECE 4550)

\[
X(w) = \mathcal{F}\{x(n)\} = \sum_{n=-\infty}^{\infty} x(n) \cdot \exp(-j \cdot w \cdot n), \quad \text{for} \quad -\pi \leq w \leq \pi
\]

and \( X(w) \) periodic in \( 2\pi \) outside the range (it repeats in the frequency domain)

The inverse transform is defined as

\[
x(n) = \mathcal{F}^{-1}\{X(w)\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(w) \cdot \exp(j \cdot w \cdot n) \cdot dw
\]

**Theorem 8.2-1:** Convolution in the time domain is equivalent to multiplication in the frequency domain

\[
Y(w) = \sum_{n=-\infty}^{\infty} y(n) \cdot \exp(-j \cdot w \cdot n) = \sum_{n=-\infty}^{\infty} [x(n) \ast h(n)] \cdot \exp(-j \cdot w \cdot n)
\]

\[
Y(w) = \sum_{n=-\infty}^{\infty} \left[ \sum_{m=-\infty}^{\infty} h(n-m) \cdot x(m) \right] \cdot \exp(-j \cdot w \cdot n)
\]

Switch the order of summation and maintain terms …

\[
Y(w) = \sum_{m=-\infty}^{\infty} \left[ \sum_{n=-\infty}^{\infty} h(n-m) \cdot \exp(-j \cdot w \cdot n) \right] \cdot x(m)
\]

\[
Y(w) = \sum_{m=-\infty}^{\infty} \left[ \sum_{n=-\infty}^{\infty} h(n-m) \cdot \exp(-j \cdot w \cdot (n-m)) \right] \cdot x(m) \cdot \exp(-j \cdot w \cdot m)
\]

The inside summation is a discrete FT

\[
Y(w) = \sum_{m=-\infty}^{\infty} H(w) \cdot x(m) \cdot \exp(-j \cdot w \cdot m) = H(w) \cdot \sum_{m=-\infty}^{\infty} x(m) \cdot \exp(-j \cdot w \cdot m)
\]

\[
Y(w) = H(w) \cdot X(w)
\]

**The z-Transform exists (with the appropriate region of convergence (ROC)).**

\[
X(z) = \sum_{n=-\infty}^{\infty} x(n) \cdot z^{-n}
\]
8.3 Random Sequences and Linear Systems

For a “non-causal” discrete linear system we will have …. (FIR filter)

\[
E[y(n)] = E\left[\sum_{k=-\infty}^{\infty} h(k) \cdot x(n-k)\right] = E\left[\sum_{m=-\infty}^{\infty} h(n-m) \cdot x(m)\right]
\]

\[
E[y(n)] = \sum_{k=-\infty}^{\infty} h(k) \cdot E[x(n-k)] = \sum_{m=-\infty}^{\infty} h(n-m) \cdot E[x(m)]
\]

If WSS

\[
E[y(n)] = \mu_Y = \sum_{k=-\infty}^{\infty} h(k) \cdot \mu_X = \sum_{m=-\infty}^{\infty} h(n-m) \cdot \mu_X
\]

\[
E[y(n)] = \mu_Y = \mu_X \cdot \sum_{k=-\infty}^{\infty} h(k)
\]

The mean times the coherent gain of the filter.

For a filter with poles and zeros …. the filter may be autoregressive as well!

\[
E[y(n)] = E\left[\sum_{k=0}^{\infty} b(k) \cdot x(n-k)\right] + E\left[\sum_{k=1}^{\infty} a(k) \cdot y(n-k)\right]
\]

\[
E[y(n)] = \sum_{k=0}^{\infty} b(k) \cdot E[x(n-k)] + \sum_{k=1}^{\infty} a(k) \cdot E[y(n-k)]
\]

If WSS

\[
E[y(n)] = \mu_Y = \sum_{k=0}^{\infty} b(k) \cdot \mu_X + \sum_{k=1}^{\infty} a(k) \cdot \mu_Y
\]

\[
\mu_Y \cdot \left(1 - \sum_{k=1}^{\infty} a(k)\right) = \mu_X \cdot \sum_{k=0}^{\infty} b(k)
\]

\[
\mu_Y = \mu_X \cdot \frac{\sum_{k=0}^{\infty} b(k)}{1 - \sum_{k=1}^{\infty} a(k)}
\]
If you are familiar with the z-transform

\[ X(z) = \sum_{n=-\infty}^{\infty} x(n) \cdot z^{-n} \]

\[ H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^{\infty} b(k) \cdot z^{-k}}{1 - \sum_{k=1}^{\infty} a(k) \cdot z^{-k}} \]

Therefore

\[ \mu_Y = \mu_X \cdot \frac{\sum_{k=0}^{\infty} b(k)}{1 - \sum_{k=1}^{\infty} a(k)} = \mu_X \cdot H(z = 1) \]

If in the Fourier Domain

\[ \mu_Y = \mu_X \cdot \frac{\sum_{k=0}^{\infty} b(k)}{1 - \sum_{k=1}^{\infty} a(k)} = \mu_X \cdot H(w = 0) \]

For an analog system …

For he Laplace transform, the “final value theorem” or steady state at infinite time is

\[ x(\infty) = \lim_{s \to 0} (s \cdot X(s)) \]

But for a unit step response input to a transfer function,

\[ u(t) \Rightarrow \frac{1}{s} \]

Therefore, for a unit step “DC level” input, the steady state output is

\[ y(\infty) = \lim_{t \to \infty} [h(t) \ast u(t)] = \lim_{s \to 0} \left( s \cdot H(s) \cdot \frac{1}{s} \right) = H(0) \]

For

\[ H(s) = \frac{A(s)}{B(s)} \]

It is just the sum of the numerator coefficients divided by the sum of the denominator coefficients!
Auto- and Cross-Correlation

For a “causal” discrete finite impulse response linear system we will have … (impulse response based)

\[ y(n) = \sum_{k=0}^{\infty} h(k) \cdot x(n-k) = \sum_{m=-\infty}^{n} h(n-m) \cdot x(m) \]

And performing a cross-correlation (assuming real R.V. and processing)

\[
E[x(n_1) \cdot y(n_2)] = E \left[ x(n_1) \cdot \sum_{k=0}^{\infty} h(k) \cdot x(n_2-k) \right]
\]

\[
E[x(n_1) \cdot y(n_2)] = E \left[ \sum_{k=0}^{\infty} h(k) \cdot x(n_1) \cdot x(n_2-k) \right]
\]

\[
E[x(n_1) \cdot y(n_2)] = \sum_{k=0}^{\infty} h(k) \cdot E[x(n_1) \cdot x(n_2-k)]
\]

\[
E[x(n_1) \cdot y(n_2)] = \sum_{k=0}^{\infty} h(k) \cdot R_{XX}(n_1, n_2-k)
\]

For x(n) WSS

\[
E[x(n) \cdot y(n+m)] = R_{XY}(m) = \sum_{k=0}^{\infty} h(k) \cdot R_{XX}(n+m-k-n)
\]

\[
E[x(n) \cdot y(n+m)] = R_{XY}(m) = \sum_{k=0}^{\infty} h(k) \cdot R_{XX}(m-k)
\]

\[
E[x(n) \cdot y(n+m)] = R_{XY}(m) = h(m) \cdot R_{XX}(m)
\]

What about the other way … YX instead of XY

\[
E[y(n_1) \cdot x(n_2)] = E \left[ \sum_{k=0}^{\infty} h(k) \cdot x(n_1-k) \cdot x(n_2) \right]
\]

\[
E[y(n_1) \cdot x(n_2)] = \sum_{k=0}^{\infty} h(k) \cdot E[x(n_1-k) \cdot x(n_2)]
\]

\[
E[y(n_1) \cdot x(n_2)] = \sum_{k=0}^{\infty} h(k) \cdot R_{XX}(n_1-k, n_2)
\]

For x(n) WSS … see the next page
For x(n) WSS
\[ E[y(n) \cdot x(n + m)] = R_{yx}(m) = \sum_{k=0}^{\infty} h(k) \cdot R_{XX}(n + m - (n - k)) \]
\[ E[y(n) \cdot x(n + m)] = R_{yx}(m) = \sum_{k=0}^{\infty} h(k) \cdot R_{XX}(m + k) \]

Perform a change of variable for k to “-l” (assuming h(t) is real, see text for complex)
\[ E[y(n) \cdot x(n + m)] = R_{yx}(m) = \sum_{l=0}^{\infty} h(-l) \cdot R_{XX}(m - l) \]

Therefore
\[ E[y(n) \cdot x(n + m)] = R_{yx}(m) = h(-m) \ast R_{XX}(m) \]

**What about the auto-correlation of y(n)?**

And performing an auto-correlation (assuming real R.V. and processing)
\[ E[y(n_1) \cdot y(n_2)] = E \left[ \sum_{k_1=0}^{\infty} h(k_1) \cdot x(n_1 - k_1) \cdot \sum_{k_2=0}^{\infty} h(k_2) \cdot x(n_2 - k_2) \right] \]
\[ E[y(n_1) \cdot y(n_2)] = E \left[ \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} h(k_1) \cdot h(k_2) \cdot x(n_1 - k_1) \cdot x(n_2 - k_2) \right] \]
\[ E[y(n_1) \cdot y(n_2)] = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} h(k_1) \cdot h(k_2) \cdot R_{XX}(n_1 - k_1, n_2 - k_2) \]

For x(n) WSS
\[ E[y(n) \cdot y(n + m)] = R_{yy}(m) = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} h(k_1) \cdot h(k_2) \cdot R_{XX}(n + m - k_2 - (n - k_1)) \]
\[ E[y(n) \cdot y(n + m)] = R_{yy}(m) = \sum_{k_1=0}^{\infty} h(k_1) \cdot \sum_{k_2=0}^{\infty} h(k_2) \cdot R_{XX}(m + k_1 - k_2) \]

The output autocorrelation can also be defined in terms of the cross-correlation as
\[ E[y(n) \cdot y(n + m)] = R_{yy}(m) = \sum_{k_1=0}^{\infty} h(k_1) \cdot R_{yx}(m + k_1) \]
\[ E[y(n) \cdot y(n + m)] = R_{yy}(m) = h(-m) \ast R_{XX}(m) = R_{yx}(m) \ast h(-m) \]

The cross-correlation can be used to determine the output auto-correlation!
Continue in this concept, the cross correlation is also a convolution. Therefore,

\[ E[y(n) \cdot y(n + m)] = R_{yy}(m) = R_{xy}(m) \ast h(-m) \]

\[ E[y(n) \cdot y(n + m)] = R_{yy}(m) = R_{xx}(m) \ast h(m) \ast h(-m) \]

If \( h(n) \) is complex, the term in \( h(-n) \) must be a conjugate.

Summary: For \( x(n) \) WSS and a real filter

\[ E[x(n) \cdot y(n + m)] = R_{xy}(m) = h(m) \ast R_{xx}(m) \]

\[ E[y(n) \cdot x(n + m)] = R_{yx}(m) = h(-m) \ast R_{xx}(m) \]

\[ E[y(n) \cdot y(n + m)] = R_{yy}(m) = R_{xx}(m) \ast h(m) \ast h(-m) \]

**The Mean Square Value at a System Output**

\[ E[y(n)^2] = R_{yy}(0) = \sum_{k_1=0}^{\infty} h(k_1) \cdot \sum_{k_2=0}^{\infty} R_{xx}(0 + k_1 - k_2) \]

\[ E[y(n)^2] = R_{yy}(0) = \sum_{k_1=0}^{\infty} h(k_1) \cdot \sum_{k_2=0}^{\infty} h(k_2) \cdot R_{xx}(k_1 - k_2) \]

\[ E[y(n)^2] = R_{yy}(0) = \sum_{k=0}^{\infty} h(k) \cdot [h(k) \ast R_{xx}(k)] \]

or

\[ E[y(n)^2] = R_{yy}(0) = \sum_{k=0}^{\infty} h(k) \cdot R_{xy}(k) \]
Example: White Noise Inputs to a causal filter

Let

\[ R_{xx}(n) = \frac{N_0}{2} \cdot \delta(n) \]

\[ E[y(n)^2] = R_{yy}(0) = \sum_{k_1=0}^{\infty} h(k_1) \cdot \sum_{k_2=0}^{\infty} h(k_2) \cdot R_{xx}(k_1 - k_2) \]

\[ E[y(n)^2] = R_{yy}(0) = \sum_{k_1=0}^{\infty} h(k_1) \cdot \sum_{k_2=0}^{\infty} h(k_2) \cdot \frac{N_0}{2} \cdot \delta(k_1 - k_2) \]

\[ E[y(n)^2] = R_{yy}(0) = \frac{N_0}{2} \cdot \sum_{k_1=0}^{\infty} h(k_1) \cdot h(k_1) \]

\[ E[y(n)^2] = R_{yy}(0) = \frac{N_0}{2} \cdot \sum_{k=0}^{\infty} h(k)^2 \]

For a white noise process, the mean squared (or 2\textsuperscript{nd} moment) is proportional to the filter power.

Typically, there are similar derivations for sampled systems and continuous systems.
The power spectral density output of linear systems

The discrete Power Spectral Density is defined as:

$$S_{XX}(w) = \sum_{n=-\infty}^{\infty} R_{XX}(n) \cdot \exp(-j \cdot w \cdot n)$$

The inverse transform is defined as

$$R_{XX}(n) = \mathcal{F}^{-1}[S_{XX}(w)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{XX}(w) \cdot \exp(j \cdot w \cdot n) \cdot dw$$

Properties:

1. $S_{xx}(w)$ is purely real as $R_{xx}(n)$ is conjugate symmetric
2. If $X(n)$ is a real-valued WSS process, then $S_{xx}(w)$ is an even function, as $R_{xx}(n)$ is real and even.
3. $S_{xx}(w) \geq 0$ for all $w$.
4. $R_{xx}(m)=0$ for all $m>N$ for some finite integer. This is the condition for the Fourier transform to exist ... finite energy.

Cross-Spectral Density

Since we have already shown the convolution formula, we can progress to the cross-spectral density functions

$$S_{XY}(w) = \sum_{n=-\infty}^{\infty} R_{XY}(n) \cdot \exp(-j \cdot w \cdot n)$$

$$S_{YX}(w) = \sum_{n=-\infty}^{\infty} [h(n) * R_{XX}(n)] \cdot \exp(-j \cdot w \cdot n)$$

Then

$$S_{XY}(w) = S_{XX}(w) \cdot H(w)$$
And for the other cross spectral density

\[ S_{YX}(w) = \sum_{n=-\infty}^{\infty} R_{YX}(n) \cdot \exp(-j \cdot w \cdot n) \]

\[ S_{XY}(w) = \sum_{n=-\infty}^{\infty} [h(-m) \ast R_{XX}(m)] \cdot \exp(-j \cdot w \cdot n) \]

Then for a real filter

\[ S_{YX}(w) = S_{XX}(w) \cdot H(-w) = S_{XX}(w) \cdot H(w)^* \]

The output power spectral density becomes

\[ S_{YY}(w) = \sum_{n=-\infty}^{\infty} R_{YY}(n) \cdot \exp(-j \cdot w \cdot n) \]

\[ S_{YY}(w) = \sum_{n=-\infty}^{\infty} \left[R_{XX}(m) \ast h(m) \ast h(-m)^*\right] \cdot \exp(-j \cdot w \cdot n) \]

For all systems

\[ S_{YY}(w) = S_{XX}(w) \cdot H(w) \cdot H(w)^* = S_{XX}(w) \cdot |H(w)|^2 \]
Table 8.4-1 Input/Output Relations for WSS Sequences and Linear Systems

<table>
<thead>
<tr>
<th>Random Sequence:</th>
<th>Output Mean:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y[n] = h[n] \ast X[n]$</td>
<td>$\mu_Y = H(0)\mu_X$</td>
</tr>
<tr>
<td>Crosscorrelations:</td>
<td>Cross–Power Spectral Densities:</td>
</tr>
<tr>
<td>$R_{XY}[m] = R_{XX}[m] \ast h^*[−m]$</td>
<td>$S_{XY}(\omega) = S_{XX}(\omega)H^*(\omega)$</td>
</tr>
<tr>
<td>$R_{YX}[m] = h[m] \ast R_{XX}[m]$</td>
<td>$S_{YX}(\omega) = H(\omega)S_{XX}(\omega)$</td>
</tr>
<tr>
<td>$R_{YY}[m] = R_{YX}[m] \ast h^*[−m]$</td>
<td>$S_{YY}(\omega) = S_{YX}(\omega)H^*(\omega)$</td>
</tr>
<tr>
<td>Autocorrelation:</td>
<td>Power Spectral Density:</td>
</tr>
<tr>
<td>$R_{YY}[m] = h[m] \ast h^*[−m] \ast R_{XX}[m]$</td>
<td>$S_{YY}(\omega) =</td>
</tr>
<tr>
<td></td>
<td>$\quad = G(\omega)S_{XX}(\omega)$</td>
</tr>
</tbody>
</table>

Output Power and Variance:

$$E\{[Y[n]]^2\} = R_{YY}[0] = \frac{1}{2\pi} \int_{-\pi}^{+\pi} |H(\omega)|^2 S_{XX}(\omega) d\omega$$

$$\sigma_Y^2 = R_{YY}[0] - |\mu_Y|^2$$

Note: The author and Dr. Bazuin differ in the ordering of the cross-spectral densities.

There is a philosophy to the notations …

Do you define it as

$$E\{y(n) \cdot y(n + m)^*\} = R_{yy}(m)$$

$$E\{y(n_1) \cdot y(n_2)^*\} = R_{yy}(n_2 - n_1)$$

or as

$$E\{y(n + m) \cdot y(n)^*\} = R_{yy}(m) \quad \text{where} \quad E\{y(n) \cdot y(n + m)^*\} = R_{yy}(-m)$$

$$E\{y(n_1) \cdot y(n_2)^*\} = R_{yy}(n_1 - n_2) = R_{yy}(-n_2 + n_1)$$

Dr. Bazuin prefers the former while the authors prefer the later!

For real R.V. or functions, it really doesn’t matter for auto-correlation

$$R_{yy}(m) = R_{yy}(-m)$$

But, for cross-correlation functions you must be consistent using one way or the other and not mixing them! Although again for real functions

$$R_{xy}(m) = R_{yx}(-m)$$


B.J. Bazuin, Fall 2016 45 of 54 ECE 3800
Examples using the Autocorrelation and PSD

Example 8.4-2 Edge Detector (1st order difference or derivative)

\[ Y(n) = X(n) - X(n-1) \]

The filter that generates this can be described as

\[ h(n) = \delta(n) - \delta(n-1) \]

If the output autocorrelation can be described as

\[ E[y(n) \cdot y(n+m)] = R_{yy}(m) = R_{xx}(m) * h(m) * h(-m) \]

We can form

\[ E[y(n) \cdot y(n+m)] = R_{yy}(m) = R_{xx}(m) * g(m) \]

where

\[ g(m) = h(m) * h(-m) \]

For this problem we can compute

\[ g(m) = \sum_{k=-\infty}^{\infty} h(-k) \cdot h(m-k) \]

\[ g(m) = \sum_{k=-\infty}^{\infty} [\delta(-k) - \delta(-k-1)] \cdot [\delta(m-k) - \delta(m-k-1)] \]

There will be four terms to consider:

\[ g(m) = \sum_{k=-\infty}^{\infty} [\delta(-k) \cdot \delta(m-k) - \delta(-k) \cdot \delta(m-k-1) - \delta(-k-1) \cdot \delta(m-k) + \delta(-k-1) \cdot \delta(m-k-1)] \]

Based on arbitrary time differences … these reduce to

\[ g(m) = \sum_{k=-\infty}^{\infty} [\delta(m) - \delta(m-1) - \delta(m+1) + \delta(m)] = 2 \cdot \delta(m) - \delta(m-1) - \delta(m+1) \]

Then

\[ R_{yy}(m) = \sum_{k=-\infty}^{\infty} R_{xx}(m-k) \cdot g(k) \]

\[ R_{yy}(m) = \sum_{k=1}^{11} R_{xx}(m-k) \cdot [2 \cdot \delta(k) - \delta(m-1) - \delta(k+1)] \]

\[ R_{yy}(m) = 2 \cdot R_{xx}(m) - R_{xx}(m-1) - R_{xx}(m+1) \]

If the input autocorrelation is defined as

\[ R_{xx}(m) = \alpha^{|m|} \]

Figure 8.3-2 Input correlation function for edge detector with \( \alpha = 0.7 \).

Then

\[ R_{yy}(m) = 2 \cdot \alpha^{|m|} - \alpha^{|m-1|} - \alpha^{|m+1|} \]

What about the power spectral density?

The first term is easy, but the others are tougher … ie is easier to calculate as the product of the input PSD and the filter magnitude squared.
Power spectral density approach

\[
S_{yy}(w) = S_{xx}(w) \cdot H(w) \cdot H(w)^* = S_{xx}(w) \cdot |H(w)|^2
\]

From the filter

\[
H(w) = \sum_{n=-\infty}^{\infty} h(n) \cdot \exp(-j \cdot w \cdot n) = \sum_{n=-\infty}^{\infty} \left[\delta(n) - \delta(n - 1)\right] \cdot \exp(-j \cdot w \cdot n)
\]

\[
H(w) = 1 - \exp(-j \cdot w)
\]

\[
|H(w)|^2 = H(w) \cdot H(w)^* = \left[1 - \exp(-j \cdot w)\right] \cdot \left[1 - \exp(j \cdot w)\right]
\]

\[
|H(w)|^2 = 1 - \exp(-j \cdot w) - \exp(j \cdot w) + 1 = 2 - 2 \cdot \cos(w)
\]

From the input (the easy term from before)

\[
S_{xx}(w) = \sum_{n=-\infty}^{\infty} R_{xx}(n) \cdot \exp(-j \cdot w \cdot n)
\]

\[
S_{xx}(w) = \sum_{n=-\infty}^{\infty} \alpha^n \cdot \exp(-j \cdot w \cdot n)
\]

\[
S_{xx}(w) = \sum_{n=0}^{\infty} \alpha^n \cdot \exp(-j \cdot w \cdot n) + \sum_{n=-\infty}^{-1} \alpha^n \cdot \exp(-j \cdot w \cdot n)
\]

\[
S_{xx}(w) = -1 + \sum_{n=0}^{\infty} (\alpha \cdot \exp(-j \cdot w))^n + \sum_{n=0}^{\infty} (\alpha \cdot \exp(j \cdot w))^n
\]

\[
S_{xx}(w) = -1 + \frac{1}{1 - \alpha \cdot \exp(-j \cdot w)} + \frac{1}{1 - \alpha \cdot \exp(j \cdot w)}
\]

\[
S_{xx}(w) = -1 + \frac{1 - \alpha \cdot \exp(j \cdot w) + 1 - \alpha \cdot \exp(-j \cdot w)}{1 - \alpha \cdot \exp(-j \cdot w) \cdot - \alpha \cdot \exp(j \cdot w) + \alpha^2}
\]

\[
S_{xx}(w) = -1 + \frac{2 - 2 \cdot \alpha \cos(w)}{1 - 2 \cdot \alpha \cdot \cos(w) + \alpha^2} = \frac{2 - 2 \cdot \alpha \cos(w) - 1 + 2 \cdot \alpha \cdot \cos(w) - \alpha^2}{1 - 2 \cdot \alpha \cdot \cos(w) + \alpha^2}
\]

\[
S_{xx}(w) = \frac{1 - \alpha^2}{1 - 2 \cdot \alpha \cdot \cos(w) + \alpha^2}
\]
And finally

\[ S_{yy}(w) = S_{xx}(w) \cdot |H(w)|^2 = S_{xx}(w) \cdot |H(w)|^2 \]

\[ S_{yy}(w) = \left( \frac{1-\alpha^2}{1-2 \cdot \alpha \cdot \cos(w) + \alpha^2} \right) \cdot (2 - 2 \cdot \cos(w)) \]

\[ S_{yy}(w) = 2 \cdot \left( \frac{(1-\alpha^2) \cdot (1-\cos(w))}{1-2 \cdot \alpha \cdot \cos(w) + \alpha^2} \right) \]

So is this better or worse than doing a direct Fourier transform?

Frequency response notes:

\[ S_{xx}(0) = \frac{1-\alpha^2}{1-2 \cdot \alpha \cdot \cos(w) + \alpha^2} = \frac{1-\alpha^2}{1-2 \cdot \alpha + \alpha^2} = \frac{(1-\alpha) \cdot (1+\alpha)}{(1-\alpha)^2} = \frac{(1+\alpha)}{(1-\alpha)} \]

\[ |H(0)|^2 = 2 - 2 \cdot \cos(w) = 0 \]

\[ S_{yy}(0) = 2 \cdot \left( \frac{(1-\alpha^2) \cdot (1-\cos(w))}{1-2 \cdot \alpha \cdot \cos(w) + \alpha^2} \right) = 2 \cdot \left( \frac{(1+\alpha) \cdot (1-1)}{1-\alpha} \right) = 0 \]

and

\[ S_{xx}(\pi) = \frac{1-\alpha^2}{1-2 \cdot \alpha \cdot \cos(w) + \alpha^2} = \frac{1-\alpha^2}{1+2 \cdot \alpha + \alpha^2} = \frac{(1-\alpha) \cdot (1+\alpha)}{(1+\alpha)^2} = \frac{(1-\alpha)}{(1+\alpha)} \]

\[ |H(\pi)|^2 = 2 - 2 \cdot \cos(w) = 4 \]

\[ S_{yy}(\pi) = 2 \cdot \left( \frac{(1-\alpha^2) \cdot (1-\cos(w))}{1-2 \cdot \alpha \cdot \cos(w) + \alpha^2} \right) = 2 \cdot \left( \frac{(1-\alpha) \cdot (1+1)}{1+\alpha} \right) = 4 \cdot \frac{1-\alpha}{1+\alpha} \]

see Example_SW_8_4_2.m
Synthesis of Random Sequences and Discrete-Time Simulations

We can generate a transfer function to provide a random sequence with a specified PSD or correlation function. Starting with a digital filter

\[ y(n) = \sum_{k=0}^{\infty} b(k) \cdot x(n-k) + \sum_{k=1}^{\infty} a(k) \cdot y(n-k) \]

The Fourier transform is

\[ H(w) = \frac{Y(w)}{X(w)} = \frac{A(w)}{B(w)} \]

where

\[ B(w) = \sum_{n=0}^{\infty} b(n) \cdot \exp(-j \cdot w \cdot n) \quad \text{and} \quad A(w) = 1 - \sum_{n=1}^{\infty} a(n) \cdot \exp(-j \cdot w \cdot n) \]

The signal input to the filter is white noise, producing a constant magnitude frequency response. Therefore,

\[ S_{yy}(w) = \frac{N_0}{2} \cdot H(w) \cdot H^*(w) = \frac{N_0}{2} \cdot |H(w)|^2 \]

For real causal coefficients, this is also equivalent to

\[ S_{yy}(w) = \frac{N_0}{2} \cdot H(w) \cdot H(-w) \]

Or using z-transform notation for \( z = \exp(-j \cdot w) \) then \( z^{-1} = \exp(j \cdot w) \) this can be written as

\[ S_{yy}(z) = \frac{N_0}{2} \cdot H(z) \cdot H(z^{-1}) \]

In the z-domain, the unit circle is a key component where this implies that there is a mirror image about the unit circle for poles and zeros in the z-domain. As an added point of interest, minimum phase, stable filters will have all their poles and zeros inside the unit circle. The mirror image elements form the “inverse filter”.
Example 8.4-5 Filter generation

If a desired psd can be stated as

\[ S_{xx}(w) = \frac{\sigma_n^2}{1 - 2 \cdot \rho \cdot \cos(w) + \rho^2} \]

For \( 2 \cdot \cos(w) = \exp(j \cdot w) + \exp(-j \cdot w) \) and \( 1 = \exp(j \cdot w) \cdot \exp(-j \cdot w) \)

The equivalent z-transform representation is

\[ 2 \cdot \cos(w) = z^{-1} + z \quad \text{and} \quad 1 = z^{-1} \cdot z \]

Then

\[ S_{xx}(z) = \frac{\sigma_n^2}{1 - \rho \cdot (z^{-1} + z) + (\rho \cdot z^{-1}) \cdot (\rho \cdot z)} \]

\[ S_{xx}(z) = \sigma_n^2 \cdot \frac{1}{(1 - \rho \cdot z^{-1}) \cdot (1 - \rho \cdot z)} = \sigma_n^2 \cdot \frac{1}{1 - \rho \cdot z^{-1}} \cdot \frac{1}{1 - \rho \cdot z} = \sigma_n^2 \cdot H(z^{-1}) \cdot H(z) \]

The desired filter is then

\[ H(z) = \frac{1}{1 - \rho \cdot z} \]

The inverse transform becomes

\[ h(n) = \rho^n \cdot u(n) \]

or

\[ y(n) = x(n) + \rho \cdot y(n-1) \]

Note: this should look similar to Example 8.4-6 . . .

If you are interested in

Decimation 496
Interpolation 497

the basis for this will be shown in ECE 4550 and you could see me for a multirate signal processing (graduate) course . . .
Discrete Implementation – Notes on the Discrete Fourier Transform

The discrete Fourier Transform

\[ X_D(k) = \sum_{n=0}^{N-1} W_N^{nk} \cdot x(n) \]

\[ \hat{x}(k) = \frac{1}{N} \sum_{n=0}^{N-1} W_N^{-nk} \cdot X_D(k) \]

where \( W_N^{nk} = \exp\left(-j\frac{2\pi nk}{N}\right) \)

Equivalence of the continuous and discrete transforms …

1. Discretization of the input \( x(n) \equiv x(n \cdot \Delta t) \) for \( \Delta t = \frac{1}{f_{\text{sample}}} \)

2. Spectral bins “actual center freq” \( X_C\left(\frac{k}{N \cdot \Delta t}\right) = X_C\left(k \cdot \Delta f\right) = X_D(k) \)

where \( k=0, +/-1, \ldots +/- N/2, \) and \( \Delta f = \frac{f_{\text{sample}}}{N} \)

DFTs get it right at the frequency sample points … based on the applied or actual window used.

3. Spectral bin concept

\[ \frac{f_{\text{sample}}}{\Delta f} \int_{-f_{\text{sample}}/2N}^{f_{\text{sample}}/2N} X_C\left(\frac{k}{N \cdot \Delta t}\right) \cdot df = X_D(k) \]

This is a concept; reality requires that the discrete point consist of the complete “windowed” signal response. (see the Matlab windows). This can be thought of as a piecewise linear estimate of the bin energy.
DFT Symmetry for real numbers

\[ X(k) = \sum_{n=0}^{N-1} W_N^{nk} \cdot x(n) \]

\[ X(N - k) = \sum_{n=0}^{N-1} W_N^{n(N-k)} \cdot x(n) = \sum_{n=0}^{N-1} W_N^{nk} \cdot W_N^{-nk} \cdot x(n) \]

\[ X(N - k) = X^*(k) \]

\[ Y(k) = i \sum_{n=0}^{N-1} W_N^{nk} \cdot y(n) \]

\[ Y(N - k) = i \sum_{n=0}^{N-1} W_N^{n(N-k)} \cdot y(n) = i \sum_{n=0}^{N-1} W_N^{nk} \cdot W_N^{-nk} \cdot y(n) \]

\[ Y(N - k) = i \sum_{n=0}^{N-1} W_N^{-nk} \cdot y(n) \]

\[ Y^*(N - k) = -i \sum_{n=0}^{N-1} W_N^{nk} \cdot y(n) \]

\[ Y(N - k) = -Y^*(k) \]

Notice the spectral symmetries involved. Purely real inputs are conjugate symmetric about \( N \), while imaginary inputs are conjugate anti-symmetric about \( N \).

For real input, bins 0 and \( N/2 \) are special, while

\[ X(N - 1) = X^*(1) \]

\[ X(N - 2) = X^*(2) \]

\[ \vdots \]

\[ X\left(\frac{N}{2} + 1\right) = X^*\left(\frac{N}{2} - 1\right) \]

Note that \( k=0 \) to \( N/2-1 \) represent the positive spectrum, while \( N/2+1 \) to \( N-1 \) represent the negative spectrum.
DFT Symmetry for even symmetric sequences

\[ x(n) = x(2 \cdot N - 2 - n), \quad n = 0 : N - 1 \]

\[ x(0) = x(2 \cdot N - 2) \]
\[ x(1) = x(2 \cdot N - 3) \]
\[ \vdots \]
\[ x(N - 2) = x(N) \]
\[ x(N - 1) = x(N - 1) \]

Note that \( x(N-1) \) is singular … it is symmetric to itself

\[ X(k) = \sum_{n=0}^{2N-2} W_{2N-2}^{nk} \cdot x(n) \]

\[ X(k) = \sum_{n=0}^{N-2} W_{2N-2}^{nk} \cdot x(n) + W_{2N-2}^{(N-1)k} \cdot x(N - 1) + \sum_{n=N}^{2N-2} W_{2N-2}^{nk} \cdot x(n) \]

\[ X(k) = \sum_{n=0}^{N-2} W_{2N-2}^{nk} \cdot x(n) + W_{2N-2}^{k} \cdot x(N - 1) + \sum_{n=0}^{N-2} W_{2N-2}^{(2N-2-n)k} \cdot x(2N - 2 - n) \]

\[ X(k) = \sum_{n=0}^{N-2} W_{2N-2}^{nk} \cdot x(n) + W_{2N-2}^{k} \cdot x(N - 1) + \sum_{n=0}^{N-2} W_{2N-2}^{(2N-2-n)k} \cdot W_{2N-2}^{-nk} \cdot x(2N - 2 - n) \]

\[ X(k) = \sum_{n=0}^{N-2} W_{2N-2}^{nk} \cdot x(n) + W_{2N-2}^{k} \cdot x(N - 1) + \sum_{n=0}^{N-2} W_{2N-2}^{-nk} \cdot x(n) \]

\[ X(k) = \sum_{n=0}^{N-2} \left( W_{2N-2}^{nk} + W_{2N-2}^{-nk} \right) \cdot x(n) + W_{2N-2}^{k} \cdot x(N - 1) \]

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

Therefore, \( X(k) \) is purely real … no phase components as we would require.

Since \( x(n) \) is real, we must also have … “conjugate symmetry” but there is no imaginary part!

\[ X(k) = X(N - 2 - k) \]

Therefore, only \( k=0 \) to \( N-1 \) must be computed!