Introduction to Random Processes

UDRC Summer School, 27th June 2016

Dr James R. Hopgood

James.Hopgood@ed.ac.uk

Room 2.05

Alexander Graham Bell Building

The King's Buildings

Institute for Digital Communications

School of Engineering

College of Science and Engineering

University of Edinburgh

Blank Page

This slide is intentionally left blank.

Handout 2
Stochastic Processes

Definition of a Stochastic Process

- Natural discrete-time signals can be characterised as random signals, since their values cannot be determined precisely; they are **unpredictable**.

Definition of a Stochastic Process

- Natural discrete-time signals can be characterised as random signals, since their values cannot be determined precisely; they are **unpredictable**.

- Consider an experiment with outcomes $\mathcal{S} = \{\zeta_k,\ k \in \mathbb{Z}^+\}$, each occurring with probability $\Pr(\zeta_k)$. Assign to each $\zeta_k \in \mathcal{S}$ a deterministic sequence $x[n, \zeta_k]\,,\ n \in \mathbb{Z}$.

Definition of a Stochastic Process

- Natural discrete-time signals can be characterised as random signals, since their values cannot be determined precisely; they are **unpredictable**.

- Consider an experiment with outcomes $\mathcal{S} = \{\zeta_k,\ k \in \mathbb{Z}^+\}$, each occurring with probability $\Pr(\zeta_k)$. Assign to each $\zeta_k \in \mathcal{S}$ a deterministic sequence $x[n, \zeta_k]\,,\ n \in \mathbb{Z}$.

- The sample space $\mathcal{S}$, probabilities $\Pr(\zeta_k)$, and the sequences $x[n, \zeta_k]\,,\ n \in \mathbb{Z}$ constitute a **discrete-time stochastic process**, or **random sequence**.

Definition of a Stochastic Process

- Natural discrete-time signals can be characterised as random signals, since their values cannot be determined precisely; they are **unpredictable**.

- Consider an experiment with outcomes $\mathcal{S} = \{\zeta_k,\ k \in \mathbb{Z}^+\}$, each occurring with probability $\Pr(\zeta_k)$. Assign to each $\zeta_k \in \mathcal{S}$ a deterministic sequence $x[n, \zeta_k]\,,\ n \in \mathbb{Z}$.

- The sample space $\mathcal{S}$, probabilities $\Pr(\zeta_k)$, and the sequences $x[n, \zeta_k]\,,\ n \in \mathbb{Z}$ constitute a **discrete-time stochastic process**, or **random sequence**.

- Formally, $x[n, \zeta_k]\,,\ n \in \mathbb{Z}$ is a random sequence or **stochastic process** if, for a fixed value $n_0 \in \mathbb{Z}^+$ of n, $x[n_0, \zeta]\,,\ n \in \mathbb{Z}$ is a random variable.

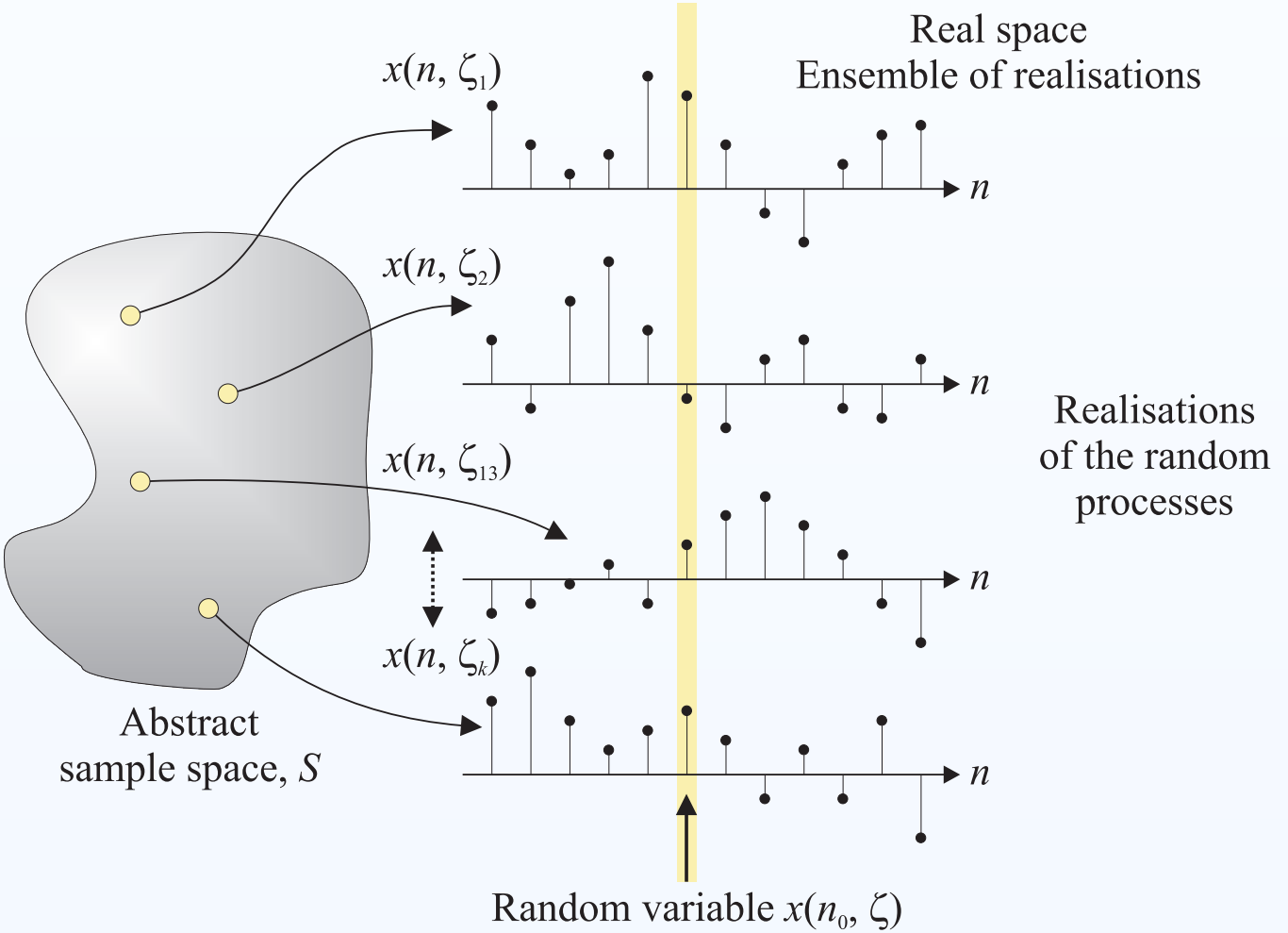
Definition of a Stochastic Process

- Natural discrete-time signals can be characterised as random signals, since their values cannot be determined precisely; they are **unpredictable**.
- Consider an experiment with outcomes $\mathcal{S} = \{\zeta_k,\, k \in \mathbb{Z}^+\}$, each occurring with probability $\Pr(\zeta_k)$. Assign to each $\zeta_k \in \mathcal{S}$ a deterministic sequence $x[n, \zeta_k]\,,\ n \in \mathbb{Z}$.
- The sample space $\mathcal{S}$, probabilities $\Pr(\zeta_k)$, and the sequences $x[n, \zeta_k]\,,\ n \in \mathbb{Z}$ constitute a **discrete-time stochastic process**, or **random sequence**.
- Formally, $x[n, \zeta_k]\,,\ n \in \mathbb{Z}$ is a random sequence or **stochastic process** if, for a fixed value $n_0 \in \mathbb{Z}^+$ of n, $x[n_0, \zeta]\,,\ n \in \mathbb{Z}$ is a random variable.
- Also known as a **time series** in the statistics literature.

Definition of a Stochastic Process

- Natural discrete-time signals can be characterised as random signals, since their values cannot be determined precisely; they are **unpredictable**.
- Consider an experiment with outcomes $\mathcal{S} = \{\zeta_k,\, k \in \mathbb{Z}^+\}$, each occurring with probability $\Pr(\zeta_k)$. Assign to each $\zeta_k \in \mathcal{S}$ a deterministic sequence $x[n, \zeta_k]\,,\ n \in \mathbb{Z}$.
- The sample space $\mathcal{S}$, probabilities $\Pr(\zeta_k)$, and the sequences $x[n, \zeta_k]\,,\ n \in \mathbb{Z}$ constitute a **discrete-time stochastic process**, or **random sequence**.
- Formally, $x[n, \zeta_k]\,,\ n \in \mathbb{Z}$ is a random sequence or **stochastic process** if, for a fixed value $n_0 \in \mathbb{Z}^+$ of n, $x[n_0, \zeta]\,,\ n \in \mathbb{Z}$ is a random variable.
- Also known as a **time series** in the statistics literature.

Interpretation of Sequences



A graphical representation of a random process.

Interpretation of Sequences

The set of all possible sequences $\{x[n, \zeta]\}$ is called an **ensemble**, and each individual sequence $x[n, \zeta_k]$, corresponding to a specific value of $\zeta = \zeta_k$, is called a **realisation** or a **sample sequence** of the ensemble.

Interpretation of Sequences

The set of all possible sequences $\{x[n, \zeta]\}$ is called an **ensemble**, and each individual sequence $x[n, \zeta_k]$, corresponding to a specific value of $\zeta = \zeta_k$, is called a **realisation** or a **sample sequence** of the ensemble.

There are four possible interpretations of $x[n, \zeta]$:

	ζ Fixed	ζ Variable
n Fixed	Number	Random variable
n Variable	Sample sequence	Stochastic process

Interpretation of Sequences

The set of all possible sequences $\{x[n, \zeta]\}$ is called an **ensemble**, and each individual sequence $x[n, \zeta_k]$, corresponding to a specific value of $\zeta = \zeta_k$, is called a **realisation** or a **sample sequence** of the ensemble.

There are four possible interpretations of $x[n, \zeta]$:

	ζ Fixed	ζ Variable
n Fixed	Number	Random variable
n Variable	Sample sequence	Stochastic process

Use simplified notation $x[n] \equiv x[n, \zeta]$ to denote both a stochastic process, and a single realisation. Use the terms **random process** and **stochastic process** interchangeably throughout this course.

Predictable Processes

The unpredictability of a random process is, in general, the combined result of the following two characteristics:

1. The selection of a single realisation is based on the outcome of a random experiment;

2. No functional description is available for *all* realisations of the *ensemble*.

Predictable Processes

The unpredictability of a random process is, in general, the combined result of the following two characteristics:

1. The selection of a single realisation is based on the outcome of a random experiment;

2. No functional description is available for *all* realisations of the *ensemble*.

In some special cases, however, a functional relationship is available. This means that after the occurrence of all samples of a particular realisation up to a particular point, n, all future values can be predicted exactly from the past ones.

Predictable Processes

The unpredictability of a random process is, in general, the combined result of the following two characteristics:

1. The selection of a single realisation is based on the outcome of a random experiment;

2. No functional description is available for *all* realisations of the *ensemble*.

In some special cases, however, a functional relationship is available. This means that after the occurrence of all samples of a particular realisation up to a particular point, n, all future values can be predicted exactly from the past ones.

If this is the case for a random process, then it is called **predictable**, otherwise it is said to be **unpredictable** or a **regular process**.

Predictable Processes

As an example of a predicatble process, consider the signal:

$$x[n, \zeta] = A \sin(\omega n + \phi)$$

where A is a known amplitude, ω is a known normalised angular frequency, and ϕ is a random phase, where $\phi \sim f_\Phi(\phi)$ is its pdf.

Description using pdfs

For fixed $n = n_0$, $x[n_0, \zeta]$ is a random variable. Moreover, the random vector formed from the k random variables $\{x[n_j]\,,\, j \in \{1, \ldots\, k\}\}$ is characterised by the cumulative distribution function (cdf) and pdfs:

$$F_X\left(x_1\ \ldots\ x_k \mid n_1\ \ldots\ n_k\right) = \Pr\left(x[n_1] \leq x_1,\ \ldots,\ x[n_k] \leq x_k\right)$$

$$f_X\left(x_1\ \ldots\ x_k \mid n_1\ \ldots\ n_k\right) = \frac{\partial^k F_X\left(x_1\ \ldots\ x_k \mid n_1\ \ldots\ n_k\right)}{\partial x_1 \cdots \partial x_k}$$

Description using pdfs

For fixed $n = n_0$, $x[n_0, \zeta]$ is a random variable. Moreover, the random vector formed from the k random variables $\{x[n_j]\,,\, j \in \{1, \ldots\, k\}\}$ is characterised by the cdf and pdfs:

$$F_X\left(x_1 \ \ldots\ x_k \mid n_1 \ \ldots\ n_k\right) = \Pr\left(x[n_1] \leq x_1,\ \ldots,\ x[n_k] \leq x_k\right)$$

$$f_X\left(x_1 \ \ldots\ x_k \mid n_1 \ \ldots\ n_k\right) = \frac{\partial^k F_X\left(x_1 \ \ldots\ x_k \mid n_1 \ \ldots\ n_k\right)}{\partial x_1 \cdots \partial x_k}$$

In exactly the same way as with random variables and random vectors, it is:

- difficult to estimate these probability functions without considerable additional information or assumptions;
- possible to frequently characterise stochastic processes usefully with much less information.

Second-order Statistical Description

Mean and Variance Sequence At time n, the **ensemble** mean and variance are given by:

$$\mu_x[n] = \mathbb{E}\left[x[n]\right]$$

$$\sigma_x^2[n] = \mathbb{E}\left[|x[n] - \mu_x[n]\,|^2\right] = \mathbb{E}\left[|x[n]\,|^2\right] - |\mu_x[n]\,|^2$$

Both $\mu_x[n]$ and $\sigma_x^2[n]$ are deterministic sequences.

Second-order Statistical Description

Mean and Variance Sequence At time n, the **ensemble** mean and variance are given by:

$$\mu_x[n] = \mathbb{E}\left[x[n]\right]$$

$$\sigma_x^2[n] = \mathbb{E}\left[|x[n] - \mu_x[n]\,|^2\right] = \mathbb{E}\left[|x[n]\,|^2\right] - |\mu_x[n]\,|^2$$

Both $\mu_x[n]$ and $\sigma_x^2[n]$ are deterministic sequences.

Autocorrelation sequence The second-order statistic $r_{xx}[n_1, n_2]$ provides a measure of the dependence between values of the process at two different times; it can provide information about the time variation of the process:

$$r_{xx}[n_1, n_2] = \mathbb{E}\left[x[n_1]\; x^*[n_2]\right]$$

Second-order Statistical Description

Autocovariance sequence The autocovariance sequence provides a measure of how similar the deviation from the mean of a process is at two different time instances:

$$\begin{aligned}\gamma_{xx}[n_1, n_2] &= \mathbb{E}\left[(x[n_1] - \mu_x[n_1])(x[n_2] - \mu_x[n_2])^*\right] \\ &= r_{xx}[n_1, n_2] - \mu_x[n_1]\ \mu_x^*[n_2]\end{aligned}$$

Second-order Statistical Description

Autocovariance sequence The autocovariance sequence provides a measure of how similar the deviation from the mean of a process is at two different time instances:

$$\begin{aligned}\gamma_{xx}[n_1, n_2] &= \mathbb{E}\left[(x[n_1] - \mu_x[n_1])(x[n_2] - \mu_x[n_2])^*\right] \\ &= r_{xx}[n_1, n_2] - \mu_x[n_1]\ \mu_x^*[n_2]\end{aligned}$$

To show how these deterministic sequences of a stochastic process can be calculated, several examples are considered in detail below.

Example of calculating autocorrelations

Example ([Manolakis:2000, Ex 3.9, page 144]). The harmonic process $x[n]$ is defined by:

$$x[n] = \sum_{k=1}^{M} A_k \cos(\omega_k n + \phi_k), \quad \omega_k \neq 0$$

where M, $\{A_k\}_1^M$ and $\{\omega_k\}_1^M$ are constants, and $\{\phi_k\}_1^M$ are pairwise independent random variables uniformly distributed in the interval $[0, 2\pi]$.

1. Determine the mean of $x(n)$.
2. Show the autocorrelation sequence is given by

$$r_{xx}[\ell] = \frac{1}{2} \sum_{k=1}^{M} |A_k|^2 \cos \omega_k \ell, \quad -\infty < \ell < \infty \qquad \bowtie$$

Example of calculating autocorrelations

Example ([Manolakis:2000, Ex 3.9, page 144]). SOLUTION. 1. The expected value of the process is straightforwardly given by:

$$\mathbb{E}\left[x(n)\right] = \mathbb{E}\left[\sum_{k=1}^{M} A_k \cos(\omega_k n + \phi_k)\right] = \sum_{k=1}^{M} A_k \, \mathbb{E}\left[\cos(\omega_k n + \phi_k)\right]$$

Since a co-sinusoid is zero-mean, then:

$$\mathbb{E}\left[\cos(\omega_k n + \phi_k)\right] = \int_0^{2\pi} \cos(\omega_k n + \phi_k) \times \frac{1}{2\pi} \times d\phi_k = 0$$

Hence, it follows:

$$\mathbb{E}\left[x(n)\right] = 0, \quad \forall n \qquad \square$$

Example of calculating autocorrelations

Example ([Manolakis:2000, Ex 3.9, page 144]). SOLUTION. 1.

$$r_{xx}(n_1, n_2) = \mathbb{E}\left[\sum_{k=1}^{M} A_k \cos(\omega_k n_1 + \phi_k) \sum_{j=1}^{M} A_j^* \cos(\omega_j n_2 + \phi_j)\right]$$

$$= \sum_{k=1}^{M}\sum_{j=1}^{M} A_k A_j^* \mathbb{E}\left[\cos(\omega_k n_1 + \phi_k) \cos(\omega_j n_2 + \phi_j)\right]$$

After some algebra, it can be shown that:

$$\mathbb{E}\left[\cos(\omega_k n_1 + \phi_k) \cos(\omega_j n_2 + \phi_j)\right] = \begin{cases} \frac{1}{2}\cos\omega_k(n_1 - n_2) & k = j \\ 0 & \text{otherwise} \end{cases}$$

where $g(\phi_k) = \cos(\omega_k n_1 + \phi_k)$ and $h(\phi_k) = \cos(\omega_j n_2 + \phi_j)$, and the fact that ϕ_k and ϕ_j are independent implies the expectation function may be factorised.

$$\mathbb{E}\left[\cos(\omega_k n_1 + \phi_k) \cos(\omega_j n_2 + \phi_j)\right] = \frac{1}{2}$$

Types of Stochastic Processes

Independence A stochastic process is independent if, and only if, (iff)

$$f_X\left(x_1, \ldots, x_N \mid n_1, \ldots, n_N\right) = \prod_{k=1}^{N} f_{X_k}\left(x_k \mid n_k\right)$$

$\forall N$, n_k, $k \in \{1, \ldots, N\}$. Here, therefore, $x(n)$ is a sequence of independent random variables.

Types of Stochastic Processes

Independence A stochastic process is independent iff

$$f_X\left(x_1, \ldots, x_N \mid n_1, \ldots, n_N\right) = \prod_{k=1}^{N} f_{X_k}\left(x_k \mid n_k\right)$$

$\forall N,\, n_k,\, k \in \{1, \ldots, N\}$. Here, therefore, $x(n)$ is a sequence of independent random variables.

An independent and identically distributed (i. i. d.) process is one where all the random variables $\{x(n_k,\, \zeta),\, n_k \in \mathbb{Z}\}$ have the same pdf, and $x(n)$ will be called an **i. i. d.** random process.

Types of Stochastic Processes

Independence A stochastic process is independent iff

$$f_X\left(x_1, \ldots, x_N \mid n_1, \ldots, n_N\right) = \prod_{k=1}^{N} f_{X_k}\left(x_k \mid n_k\right)$$

$\forall N,\, n_k,\, k \in \{1, \ldots, N\}$. Here, therefore, $x(n)$ is a sequence of independent random variables.

An i. i. d. process is one where all the random variables $\{x(n_k,\, \zeta),\, n_k \in \mathbb{Z}\}$ have the same pdf, and $x(n)$ will be called an **i. i. d.** random process.

An uncorrelated processes is a sequence of uncorrelated random variables:

$$\gamma_{xx}(n_1, n_2) = \sigma_x^2(n_1)\, \delta(n_1 - n_2)$$

Types of Stochastic Processes

An orthogonal process is a sequence of **orthogonal random variables**, and is given by:

$$r_{xx}(n_1, n_2) = \mathbb{E}\left[|x(n_1)|^2\right] \delta(n_1 - n_2)$$

If a process is zero-mean, then it is both **orthogonal** and **uncorrelated** since $\gamma_{xx}(n_1, n_2) = r_{xx}(n_1, n_2)$.

Types of Stochastic Processes

An orthogonal process is a sequence of **orthogonal random variables**, and is given by:

$$r_{xx}(n_1, n_2) = \mathbb{E}\left[|x(n_1)|^2\right] \delta(n_1 - n_2)$$

If a process is zero-mean, then it is both **orthogonal** and **uncorrelated** since $\gamma_{xx}(n_1, n_2) = r_{xx}(n_1, n_2)$.

A stationary process is a random process where its statistical properties do not vary with time. Processes whose statistical properties **do** change with time are referred to as **nonstationary**.

Stationary Processes

A random process $x(n)$ has been called **stationary** if its statistics determined for $x(n)$ are equal to those for $x(n+k)$, for every k. There are various formal definitions of **stationarity**, along with **quasi-stationary** processes, which are discussed below.

- **Order-N and strict-sense stationarity**
- **Wide-sense stationarity**
- **Wide-sense periodicity** and **cyclo-stationarity**
- Local- or **quasi-stationary** processes

After this, some examples of various stationary processes will be given.

Order-N and strict-sense stationarity

Definition (Stationary of order-N). A stochastic process $x(n)$ is called **stationary of order-N** if:

$$f_X\left(x_1,\ldots,x_N \mid n_1,\ldots,n_N\right) = f_X\left(x_1,\ldots,x_N \mid n_1+k,\ldots,n_N+k\right)$$

$\diamondsuit$

for any value of k. If $x(n)$ is stationary for all orders $N \in \mathbb{Z}^+$, it is said to be **strict-sense stationary (SSS)**.

Order-N and strict-sense stationarity

Definition (Stationary of order-N). A stochastic process $x(n)$ is called **stationary of order-N** if:

$$f_X\left(x_1,\ldots,x_N \mid n_1,\ldots,n_N\right) = f_X\left(x_1,\ldots,x_N \mid n_1+k,\ldots,n_N+k\right)$$

$\diamondsuit$

for any value of k. If $x(n)$ is stationary for all orders $N \in \mathbb{Z}^+$, it is said to be **SSS**.

An independent and identically distributed process is SSS since, in this case, $f_{X_k}\left(x_k \mid n_k\right) = f_X\left(x_k\right)$ is independent of n, and therefore also of $n+k$.

However, SSS is more restrictive than necessary in practical applications, and is a rarely required property.

Wide-sense stationarity

A more relaxed form of stationarity, which is sufficient for practical problems, occurs when a random process is stationary order-2; such a process is **wide-sense stationary (WSS)**.

Wide-sense stationarity

Definition (Wide-sense stationarity). A random signal $x(n)$ is called wide-sense stationary if:

- the mean and variance is constant and independent of n:

$$\mathbb{E}\left[x(n)\right] = \mu_x$$
$$\text{var}\left[x(n)\right] = \sigma_x^2$$

- the autocorrelation depends only on the time difference $l = n_1 - n_2$, called the lag:

$$\begin{aligned} r_{xx}(n_1, n_2) &= r_{xx}^*(n_2, n_1) = \mathbb{E}\left[x(n_1)\, x^*(n_2)\right] \\ &= r_{xx}(l) = r_{xx}(n_1 - n_2) = \mathbb{E}\left[x(n_1)\, x^*(n_1 - l)\right] \\ &= \mathbb{E}\left[x(n_2 + l)\, x^*(n_2)\right] \end{aligned} \qquad \diamondsuit$$

Wide-sense stationarity

- The autocovariance function is given by:

$$\gamma_{xx}(l) = r_{xx}(l) - |\mu_x|^2$$

- Since 2nd-order moments are defined in terms of 2nd-order pdf, then strict-sense stationary are always WSS, but not necessarily *vice-versa*, except if the signal is Gaussian.

- In practice, however, it is very rare to encounter a signal that is stationary in the wide-sense, but not stationary in the strict sense.

Wide-sense cyclo-stationarity

Two classes of **nonstationary process** which, in part, have properties resembling stationary signals are:

1. **A wide-sense periodic (WSP) process** is classified as signals whose mean is periodic, and whose autocorrelation function is periodic in both dimensions:

$$\mu_x(n) = \mu_x(n+N)$$

$$r_{xx}(n_1, n_2) = r_{xx}(n_1 + N, n_2) = r_{xx}(n_1, n_2 + N)$$
$$= r_{xx}(n_1 + N, n_2 + N)$$

for all n, n_1 and n_2. These are quite tight constraints for real signals.

Wide-sense cyclo-stationarity

2. A wide-sense cyclo-stationary process has similar but less restrictive properties than a WSP process, in that the mean is periodic, but the autocorrelation function is now just invariant to a shift by N in both of its arguments:

$$\mu_x(n) = \mu_x(n+N)$$
$$r_{xx}(n_1, n_2) = r_{xx}(n_1+N, n_2+N)$$

for all n, n_1 and n_2. This type of nonstationary process has more practical applications, as the following example will show.

Quasi-stationarity

At the introduction of this lecture course, it was noted that in the analysis of speech signals, the speech waveform is broken up into short segments whose duration is typically 10 to 20 milliseconds.

Quasi-stationarity

At the introduction of this lecture course, it was noted that in the analysis of speech signals, the speech waveform is broken up into short segments whose duration is typically 10 to 20 milliseconds.

This is because speech can be modelled as a **locally stationary** or **quasi-stationary** process.

Quasi-stationarity

At the introduction of this lecture course, it was noted that in the analysis of speech signals, the speech waveform is broken up into short segments whose duration is typically 10 to 20 milliseconds.

This is because speech can be modelled as a **locally stationary** or **quasi-stationary** process.

Such processes possess statistical properties that change *slowly* over short periods of time.

Quasi-stationarity

At the introduction of this lecture course, it was noted that in the analysis of speech signals, the speech waveform is broken up into short segments whose duration is typically 10 to 20 milliseconds.

This is because speech can be modelled as a **locally stationary** or **quasi-stationary** process.

Such processes possess statistical properties that change *slowly* over short periods of time.

They are *globally* nonstationary, but are approximately *locally* stationary, and are modelled as if the statistics *actually are* stationary over a short segment of time.

WSS Properties

The average power of a WSS process $x(n)$ satisfies:

$$r_{xx}(0) = \sigma_x^2 + |\mu_x|^2$$

$$r_{xx}(0) \geq r_{xx}(l), \quad \text{for all } l$$

WSS Properties

The average power of a WSS process $x(n)$ satisfies:

$$r_{xx}(0) = \sigma_x^2 + |\mu_x|^2$$
$$r_{xx}(0) \geq r_{xx}(l), \quad \text{for all } l$$

The expression for power can be broken down as follows:

Average DC Power: $|\mu_x|^2$

Average AC Power: σ_x^2

Total average power: $r_{xx}(0)$

Total average power = Average DC power + Average AC power

WSS Properties

The average power of a WSS process $x(n)$ satisfies:

$$r_{xx}(0) = \sigma_x^2 + |\mu_x|^2$$
$$r_{xx}(0) \geq r_{xx}(l), \quad \text{for all } l$$

The expression for power can be broken down as follows:

Average DC Power: $|\mu_x|^2$

Average AC Power: σ_x^2

Total average power: $r_{xx}(0)$

Total average power $=$ Average DC power $+$ Average AC power

Moreover, it follows that $\gamma_{xx}(0) \geq \gamma_{xx}(l)$.

WSS Properties

It is left as an exercise to show that the autocorrelation sequence $r_{xx}(l)$ is:

- a conjugate symmetric function of the lag l:

$$r_{xx}^*(-l) = r_{xx}(l)$$

- a **nonnegative-definite** or **positive semi-definite** function, such that for any sequence $\alpha(n)$:

$$\sum_{n=1}^{M}\sum_{m=1}^{M} \alpha^*(n)\, r_{xx}(n-m)\alpha(m) \geq 0$$

WSS Properties

It is left as an exercise to show that the autocorrelation sequence $r_{xx}(l)$ is:

- a conjugate symmetric function of the lag l:

$$r_{xx}^*(-l) = r_{xx}(l)$$

- a **nonnegative-definite** or **positive semi-definite** function, such that for any sequence $\alpha(n)$:

$$\sum_{n=1}^{M}\sum_{m=1}^{M} \alpha^*(n)\, r_{xx}(n-m)\alpha(m) \geq 0$$

Note that, more generally, even a correlation function for a nonstationary random process is **positive semi-definite**:

$$\sum_{n=1}^{M}\sum_{m=1}^{M} \alpha^*(n)\, r_{xx}(n,m)\alpha(m) \geq 0 \quad \text{for any sequence } \alpha(n)$$

Estimating statistical properties

- A stochastic process consists of the ensemble, $x(n, \zeta)$, and a probability law, $f_X\left(\{x\} \mid \{n\}\right)$. If this information is available $\forall n$, the statistical properties are easily determined.

Estimating statistical properties

- A stochastic process consists of the ensemble, $x(n, \zeta)$, and a probability law, $f_X\left(\{x\} \mid \{n\}\right)$. If this information is available $\forall n$, the statistical properties are easily determined.

- In practice, only a limited number of realisations of a process is available, and often only one: i.e. $\{x(n, \zeta_k),\ k \in \{1, \ldots, K\}\}$ is known for some K, but $f_X\left(x \mid n\right)$ is unknown.

Estimating statistical properties

- A stochastic process consists of the ensemble, $x(n, \zeta)$, and a probability law, $f_X\left(\{x\} \mid \{n\}\right)$. If this information is available $\forall n$, the statistical properties are easily determined.

- In practice, only a limited number of realisations of a process is available, and often only one: i.e. $\{x(n, \zeta_k),\ k \in \{1, \ldots, K\}\}$ is known for some K, but $f_X\left(x \mid n\right)$ is unknown.

- Is is possible to infer the statistical characteristics of a process from a single realisation? Yes, for the following class of signals:

Estimating statistical properties

- A stochastic process consists of the ensemble, $x(n, \zeta)$, and a probability law, $f_X\left(\{x\} \mid \{n\}\right)$. If this information is available $\forall n$, the statistical properties are easily determined.

- In practice, only a limited number of realisations of a process is available, and often only one: i.e. $\{x(n, \zeta_k),\ k \in \{1, \ldots, K\}\}$ is known for some K, but $f_X\left(x \mid n\right)$ is unknown.

- Is is possible to infer the statistical characteristics of a process from a single realisation? Yes, for the following class of signals:

  - **ergodic** processes;

Estimating statistical properties

- A stochastic process consists of the ensemble, $x(n, \zeta)$, and a probability law, $f_X\left(\{x\} \mid \{n\}\right)$. If this information is available $\forall n$, the statistical properties are easily determined.

- In practice, only a limited number of realisations of a process is available, and often only one: i.e. $\{x(n, \zeta_k),\ k \in \{1, \ldots, K\}\}$ is known for some K, but $f_X\left(x \mid n\right)$ is unknown.

- Is is possible to infer the statistical characteristics of a process from a single realisation? Yes, for the following class of signals:

  - **ergodic** processes;

  - nonstationary processes where additional structure about the autocorrelation function is known (beyond the scope of this course).

Ensemble and Time-Averages

Ensemble averaging, as considered so far in the course, is not frequently used in practice since it is impractical to obtain the number of realisations needed for an accurate estimate.

Ensemble and Time-Averages

Ensemble averaging, as considered so far in the course, is not frequently used in practice since it is impractical to obtain the number of realisations needed for an accurate estimate.

A statistical average that can be obtained from a **single** realisation of a process is a **time-average**, defined by:

$$\langle g(x(n)) \rangle \triangleq \lim_{N \to \infty} \frac{1}{2N+1} \sum_{n=-N}^{N} g(x(n))$$

For every ensemble average, a corresponding time-average can be defined; the above corresponds to: $\mathbb{E}\left[g(x(n))\right]$.

Ensemble and Time-Averages

Ensemble averaging, as considered so far in the course, is not frequently used in practice since it is impractical to obtain the number of realisations needed for an accurate estimate.

A statistical average that can be obtained from a **single** realisation of a process is a **time-average**, defined by:

$$\langle g(x(n)) \rangle \triangleq \lim_{N \to \infty} \frac{1}{2N+1} \sum_{n=-N}^{N} g(x(n))$$

For every ensemble average, a corresponding time-average can be defined; the above corresponds to: $\mathbb{E}\left[g(x(n))\right]$.

Time-averages are random variables since they implicitly depend on the particular realisation, given by ζ. Averages of deterministic signals are fixed numbers or sequences, even though they are given by the same expression.

Ergodicity

A stochastic process, $x(n)$, is **ergodic** if its ensemble averages can be estimated from a single realisation of a process using time averages.

The two most important degrees of ergodicity are:

Mean-Ergodic (or ergodic in the mean) processes have identical expected values and sample-means:

$$\langle x(n) \rangle = \mathbb{E}\left[x(n)\right]$$

Covariance-Ergodic Processes (or ergodic in correlation) have the property that:

$$\langle x(n)\, x^*(n-l) \rangle = \mathbb{E}\left[x(n)\, x^*(n-l)\right]$$

Ergodicity

- It should be intuitiveness obvious that ergodic processes must be stationary and, moreover, that a process which is ergodic both in the mean and correlation is WSS.
- WSS processes are not necessarily ergodic.
- Ergodic is often used to mean both ergodic in the mean and correlation.
- In practice, only finite records of data are available, and therefore an estimate of the time-average will be given by

$$\langle g(x(n)) \rangle = \frac{1}{N} \sum_{n \in \mathcal{N}} g(x(n))$$

where N is the number of data-points available. The performance of this estimator will be discussed later in this course.

Joint Signal Statistics

Cross-correlation and cross-covariance A measure of the dependence between values of two *different* stochastic processes is given by the **cross-correlation** and **cross-covariance** functions:

$$r_{xy}(n_1, n_2) = \mathbb{E}\left[x(n_1)\, y^*(n_2)\right]$$

$$\gamma_{xy}(n_1, n_2) = r_{xy}(n_1, n_2) - \mu_x(n_1)\, \mu_y^*(n_2)$$

Joint Signal Statistics

Cross-correlation and cross-covariance A measure of the dependence between values of two *different* stochastic processes is given by the **cross-correlation** and **cross-covariance** functions:

$$r_{xy}(n_1, n_2) = \mathbb{E}\left[x(n_1)\, y^*(n_2)\right]$$

$$\gamma_{xy}(n_1, n_2) = r_{xy}(n_1, n_2) - \mu_x(n_1)\, \mu_y^*(n_2)$$

Normalised cross-correlation (or cross-covariance) The cross-covariance provides a measure of similarity of the deviation from the respective means of two processes. It makes sense to consider this deviation relative to their **standard deviations**; thus, **normalised cross-correlations**:

$$\rho_{xy}(n_1, n_2) = \frac{\gamma_{xy}(n_1, n_2)}{\sigma_x(n_1)\, \sigma_y(n_2)}$$

Types of Joint Stochastic Processes

Statistically independence of two stochastic processes occurs when, for every n_x and n_y,

$$f_{XY}\left(x, y \mid n_x, n_y\right) = f_X\left(x \mid n_x\right)\, f_Y\left(y \mid n_y\right)$$

Uncorrelated stochastic processes have, for all n_x & $n_y \neq n_x$:

$$\gamma_{xy}(n_x, n_y) = 0$$
$$r_{xy}(n_x, n_y) = \mu_x(n_x)\,\mu_y(n_y)$$

Types of Joint Stochastic Processes

Statistically independence of two stochastic processes occurs when, for every n_x and n_y,

$$f_{XY}\left(x, y \mid n_x, n_y\right) = f_X\left(x \mid n_x\right)\, f_Y\left(y \mid n_y\right)$$

Uncorrelated stochastic processes have, for all n_x & $n_y \neq n_x$:

$$\gamma_{xy}(n_x, n_y) = 0$$
$$r_{xy}(n_x, n_y) = \mu_x(n_x)\,\mu_y(n_y)$$

Joint stochastic processes that are statistically independent are uncorrelated, but not necessarily vice-versa, except for Gaussian processes. Nevertheless, a measure of uncorrelatedness is often used as a measure of independence. More on this later.

Types of Joint Stochastic Processes

Orthogonal joint processes have, for every n_1 and $n_2 \neq n_1$:

$$r_{xy}(n_1, n_2) = 0$$

Joint WSS is a similar to WSS for a single stochastic process, and is useful since it facilitates a spectral description, as discussed later in this course:

$$r_{xy}(l) = r_{xy}(n_1 - n_2) = r^*_{yx}(-l) = \mathbb{E}\left[x(n)\, y^*(n-l)\right]$$
$$\gamma_{xy}(l) = \gamma_{xy}(n_1 - n_2) = \gamma^*_{yx}(-l) = r_{xy}(l) - \mu_x\, \mu^*_y$$

Joint-Ergodicity applies to two ergodic processes, $x(n)$ and $y(n)$, whose ensemble cross-correlation can be estimated from a time-average:

$$\langle x(n)\, y^*(n-l)\rangle = \mathbb{E}\left[x(n)\, y^*(n-l)\right]$$

Correlation Matrices

Let an M-dimensional random vector $\mathbf{X}(n,\zeta) \equiv \mathbf{X}(n)$ be derived from the random process $x(n)$ as follows:

$$\mathbf{X}(n) \triangleq \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-M+1) \end{bmatrix}^T$$

Correlation Matrices

Let an M-dimensional random vector $\mathbf{X}(n,\zeta) \equiv \mathbf{X}(n)$ be derived from the random process $x(n)$ as follows:

$$\mathbf{X}(n) \triangleq \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-M+1) \end{bmatrix}^T$$

Then its mean is given by an M-vector

$$\boldsymbol{\mu}_{\mathbf{X}}(n) \triangleq \begin{bmatrix} \mu_x(n) & \mu_x(n-1) & \cdots & \mu_x(n-M+1) \end{bmatrix}^T$$

Correlation Matrices

Let an M-dimensional random vector $\mathbf{X}(n,\zeta) \equiv \mathbf{X}(n)$ be derived from the random process $x(n)$ as follows:

$$\mathbf{X}(n) \triangleq \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-M+1) \end{bmatrix}^T$$

Then its mean is given by an M-vector

$$\boldsymbol{\mu}_{\mathbf{X}}(n) \triangleq \begin{bmatrix} \mu_x(n) & \mu_x(n-1) & \cdots & \mu_x(n-M+1) \end{bmatrix}^T$$

and the $M \times M$ correlation matrix is given by:

$$\mathbf{R}_{\mathbf{X}}(n) \triangleq \begin{bmatrix} r_{xx}(n,n) & \cdots & r_{xx}(n,n-M+1) \\ \vdots & \ddots & \vdots \\ r_{xx}(n-M+1,n) & \cdots & r_{xx}(n-M+1,n-M+1) \end{bmatrix}$$

Correlation Matrices

For stationary processes, the correlation matrix has an interesting additional structure. Note that:

1. $\mathbf{R_X}(n)$ is a constant matrix $\mathbf{R_X}$;

2. $r_{xx}(n-i, n-j) = r_{xx}(j-i) = r_{xx}(l)$, $l = j - i$;

3. conjugate symmetry gives $r_{xx}(l) = r^*_{xx}(-l)$.

Correlation Matrices

For stationary processes, the correlation matrix has an interesting additional structure. Note that:

1. $\mathbf{R_X}(n)$ is a constant matrix $\mathbf{R_X}$;

2. $r_{xx}(n-i, n-j) = r_{xx}(j-i) = r_{xx}(l)$, $l = j - i$;

3. conjugate symmetry gives $r_{xx}(l) = r^*_{xx}(-l)$.

Hence, the matrix $\mathbf{R}_{xx}$ is given by:

$$\mathbf{R_X} \triangleq \begin{bmatrix} r_{xx}(0) & r_{xx}(1) & r_{xx}(2) & \cdots & r_{xx}(M-1) \\ r^*_{xx}(1) & r_{xx}(0) & r_{xx}(1) & \cdots & r_{xx}(M-2) \\ r^*_{xx}(2) & r^*_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(M-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r^*_{xx}(M-1) & r^*_{xx}(M-2) & r^*_{xx}(M-3) & \cdots & r_{xx}(0) \end{bmatrix}$$

Markov Processes

A powerful model for a stochastic process known as a **Markov model** is introduced; such a process that satisfies this model is known as a **Markov process**.

Quite simply, a Markov process is one in which the probability of any particular value in a sequence is dependent upon the preceding sample values.

Markov Processes

A powerful model for a stochastic process known as a **Markov model** is introduced; such a process that satisfies this model is known as a **Markov process**.

Quite simply, a Markov process is one in which the probability of any particular value in a sequence is dependent upon the preceding sample values.

The simplest kind of dependence arises when the probability of any sample depends only upon the value of the *immediately preceding* sample, and this is known as a **first-order Markov process**.

Markov Processes

A powerful model for a stochastic process known as a **Markov model** is introduced; such a process that satisfies this model is known as a **Markov process**.

Quite simply, a Markov process is one in which the probability of any particular value in a sequence is dependent upon the preceding sample values.

The simplest kind of dependence arises when the probability of any sample depends only upon the value of the *immediately preceding* sample, and this is known as a **first-order Markov process**.

This simple process is a surprisingly good model for a number of practical signal processing, communications and control problems.

Markov Processes

As an example of a Markov process, consider the process generated by the difference equation

$$x(n) = -a\,x(n-1) + w(n)$$

where a is a known constant, and $w(n)$ is a sequence of zero-mean i. i. d. Gaussian random variables with variance σ_W^2 density:

$$f_W\left(w(n)\right) = \frac{1}{\sqrt{2\pi\sigma_W^2}} \exp\left\{-\frac{w(n)^2}{2\sigma_W^2}\right\}$$

Markov Processes

As an example of a Markov process, consider the process generated by the difference equation

$$x(n) = -a\,x(n-1) + w(n)$$

where a is a known constant, and $w(n)$ is a sequence of zero-mean i. i. d. Gaussian random variables with variance σ_W^2 density:

$$f_W\left(w(n)\right) = \frac{1}{\sqrt{2\pi\sigma_W^2}} \exp\left\{-\frac{w(n)^2}{2\sigma_W^2}\right\}$$

The conditional density of $x(n)$ given $x(n-1)$ is also Gaussian,

$$f_X\left(x(n) \mid x(n-1)\right) = \frac{1}{\sqrt{2\pi\sigma_W^2}} \exp\left\{-\frac{(x(n)+ax(n-1))^2}{2\sigma_W^2}\right\}$$

Markov Processes

Definition (Markov Process). A random process is a Pth-order Markov process if the distribution of $x(n)$, given the infinite past, depends only on the previous P samples $\{x(n-1), \ldots, x(n-P)\}$; that is, if:

$$f_X\left(x(n) \mid x(n-1),\, x(n-2),\, \ldots\right) = f_X\left(x(n) \mid x(n-1),\, \ldots,\, x(n-P)\right)$$

$\diamondsuit$

Markov Processes

Definition (Markov Process). A random process is a Pth-order Markov process if the distribution of $x(n)$, given the infinite past, depends only on the previous P samples $\{x(n-1), \ldots, x(n-P)\}$; that is, if:

$$f_X\left(x(n) \mid x(n-1),\, x(n-2),\, \ldots\right) = f_X\left(x(n) \mid x(n-1),\, \ldots,\, x(n-P)\right)$$

$\diamondsuit$

Finally, it is noted that if $x(n)$ takes on a countable (discrete) set of values, a Markov random process is called a **Markov chain.**

Handout 3
Power Spectral Density

Introduction

Frequency- and transform-domain methods are very powerful tools for the analysis of deterministic sequences. It seems natural to extend these techniques to analysis stationary **random processes**.

Introduction

Frequency- and transform-domain methods are very powerful tools for the analysis of deterministic sequences. It seems natural to extend these techniques to analysis stationary **random processes**.

So far in this course, **stationary stochastic processes** have been considered in the time-domain through the use of the **autocorrelation function (ACF)**.

Introduction

Frequency- and transform-domain methods are very powerful tools for the analysis of deterministic sequences. It seems natural to extend these techniques to analysis stationary **random processes**.

So far in this course, **stationary stochastic processes** have been considered in the time-domain through the use of the **ACF**.

Since the ACF for a stationary process is a function of a single-discrete time process, then the question arises as to what the discrete-time Fourier transform (DTFT) corresponds to.

Introduction

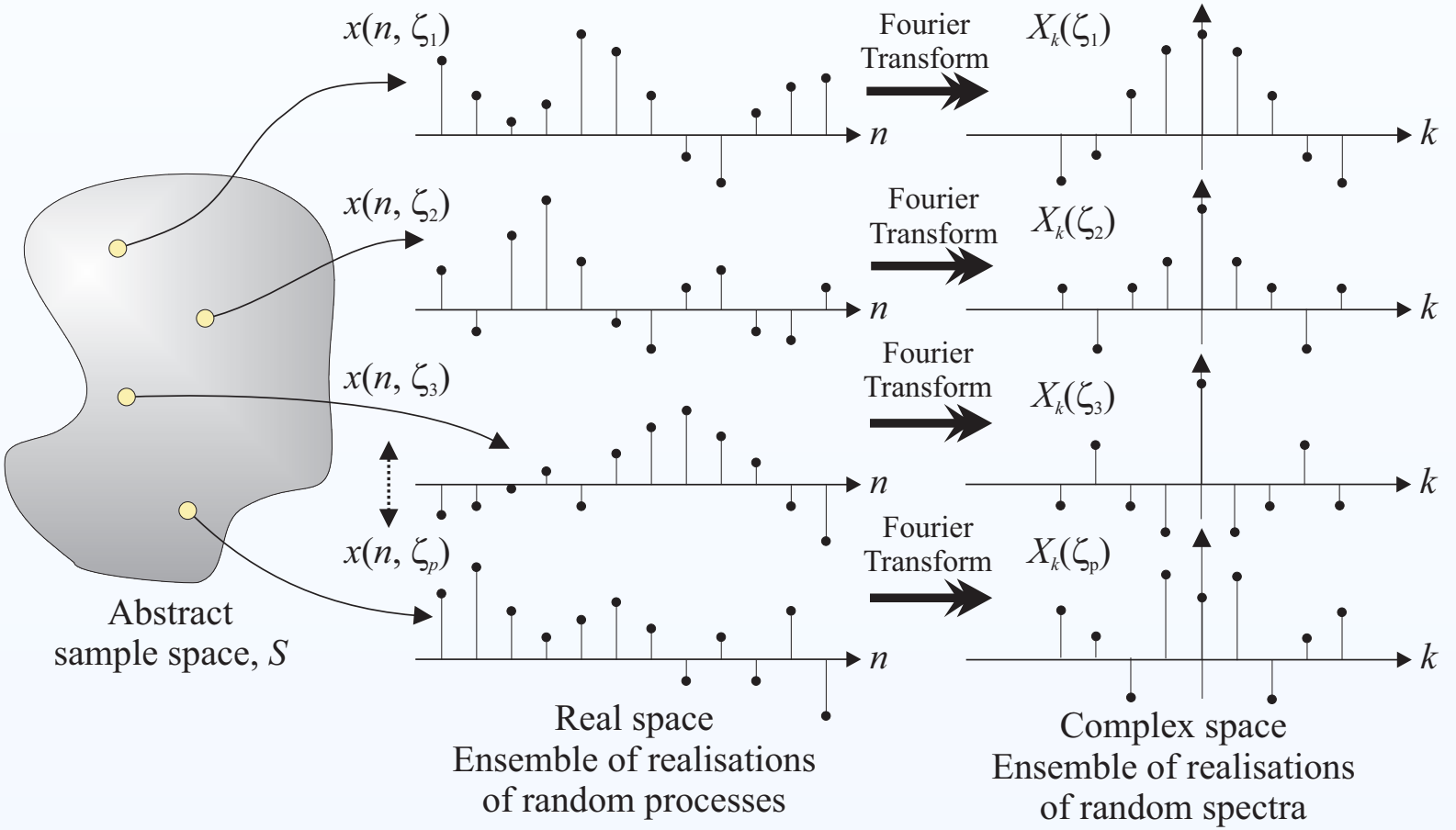
Frequency- and transform-domain methods are very powerful tools for the analysis of deterministic sequences. It seems natural to extend these techniques to analysis stationary **random processes**.

So far in this course, **stationary stochastic processes** have been considered in the time-domain through the use of the **ACF**.

Since the ACF for a stationary process is a function of a single-discrete time process, then the question arises as to what the DTFT corresponds to.

It turns out to be known as the **power spectral density (PSD)** of a stationary random process, and the PSD is an extremely powerful and conceptually appealing tool in statistical signal processing.

Introduction



A graphical respresentation of random spectra.

Introduction

In signal theory for deterministic signals, spectra are used to represent a function as a superposition of exponential functions. For random signals, the notion of a spectrum has two interpretations:

Introduction

In signal theory for deterministic signals, spectra are used to represent a function as a superposition of exponential functions. For random signals, the notion of a spectrum has two interpretations:

Transform of averages The first involves transform of averages (or moments). As will be seen, this will be the Fourier transform of the autocorrelation function.

Introduction

In signal theory for deterministic signals, spectra are used to represent a function as a superposition of exponential functions. For random signals, the notion of a spectrum has two interpretations:

Transform of averages The first involves transform of averages (or moments). As will be seen, this will be the Fourier transform of the autocorrelation function.

Stochastic decomposition The second interpretation represents a stochastic process as a superposition of exponentials, where the coefficients are themselves random variables. Hence, $x(n)$ can be represented as:

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega})\, e^{j\omega n}\, d\omega, \quad n \in \mathbb{R}$$

where $X(e^{j\omega})$ is a random variable for a given value of ω.

The power spectral density

The discrete-time Fourier transform of the autocorrelation function of a stationary stochastic process $x[n, \zeta]$ is known as the **power spectral density (PSD)**, is denoted by $P_{xx}(e^{j\omega})$, and is given by:

$$P_{xx}(e^{j\omega}) = \sum_{\ell \in \mathbb{Z}} r_{xx}[\ell] \; e^{-j\omega\ell}$$

where ω is frequency in radians per sample.

The power spectral density

The discrete-time Fourier transform of the autocorrelation function of a stationary stochastic process $x[n, \zeta]$ is known as the **power spectral density (PSD)**, is denoted by $P_{xx}(e^{j\omega})$, and is given by:

$$P_{xx}(e^{j\omega}) = \sum_{\ell \in \mathbb{Z}} r_{xx}[\ell]\ e^{-j\omega\ell}$$

where ω is frequency in radians per sample.

The autocorrelation function, $r_{xx}[\ell]$, can be recovered from the **PSD** by using the inverse-**DTFT**:

$$r_{xx}[\ell] = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_{xx}(e^{j\omega})\, e^{j\omega\ell}\, d\omega, \quad \ell \in \mathbb{Z}$$

Properties of the power spectral density

- $P_{xx}(e^{j\omega}) : \omega \rightarrow \mathbb{R}^{+}$; in otherwords, the PSD is real valued, and nonnegative definite. i.e.

$$P_{xx}(e^{j\omega}) \geq 0$$

Properties of the power spectral density

- $P_{xx}(e^{j\omega}) : \omega \rightarrow \mathbb{R}^{+}$; in otherwords, the PSD is real valued, and nonnegative definite. i.e.

$$P_{xx}(e^{j\omega}) \geq 0$$

- $P_{xx}(e^{j\omega}) = P_{xx}(e^{j(\omega+2n\pi)})$; in otherwords, the PSD is periodic with period 2π.

Properties of the power spectral density

- $P_{xx}(e^{j\omega}) : \omega \rightarrow \mathbb{R}^{+}$; in otherwords, the PSD is real valued, and nonnegative definite. i.e.

$$P_{xx}(e^{j\omega}) \geq 0$$

- $P_{xx}(e^{j\omega}) = P_{xx}(e^{j(\omega+2n\pi)})$; in otherwords, the PSD is periodic with period 2π.

- If $x[n]$ is real-valued, then:

Properties of the power spectral density

- $P_{xx}(e^{j\omega}) : \omega \rightarrow \mathbb{R}^+$; in otherwords, the PSD is real valued, and nonnegative definite. i.e.

$$P_{xx}(e^{j\omega}) \geq 0$$

- $P_{xx}(e^{j\omega}) = P_{xx}(e^{j(\omega+2n\pi)})$; in otherwords, the PSD is periodic with period 2π.

- If $x[n]$ is real-valued, then:
  - $r_{xx}[\ell]$ is real and even;

Properties of the power spectral density

- $P_{xx}(e^{j\omega}) : \omega \rightarrow \mathbb{R}^+$; in otherwords, the PSD is real valued, and nonnegative definite. i.e.

$$P_{xx}(e^{j\omega}) \geq 0$$

- $P_{xx}(e^{j\omega}) = P_{xx}(e^{j(\omega+2n\pi)})$; in otherwords, the PSD is periodic with period 2π.

- If $x[n]$ is real-valued, then:
  - $r_{xx}[\ell]$ is real and even;
  - $P_{xx}(e^{j\omega}) = P_{xx}(e^{-j\omega})$ is an even function of ω.

Properties of the power spectral density

- $P_{xx}(e^{j\omega}) : \omega \to \mathbb{R}^+$; in otherwords, the PSD is real valued, and nonnegative definite. i.e.

$$P_{xx}(e^{j\omega}) \geq 0$$

- $P_{xx}(e^{j\omega}) = P_{xx}(e^{j(\omega+2n\pi)})$; in otherwords, the PSD is periodic with period 2π.

- If $x[n]$ is real-valued, then:
  - $r_{xx}[\ell]$ is real and even;
  - $P_{xx}(e^{j\omega}) = P_{xx}(e^{-j\omega})$ is an even function of ω.

- The area under $P_{xx}(e^{j\omega})$ is nonnegative and is equal to the average power of $x[n]$. Hence:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} P_{xx}(e^{j\omega})\, d\omega = r_{xx}[0] = \mathbb{E}\left[|x[n]|^2\right] \geq 0$$

General form of the PSD

A process, $x(n)$, and therefore its corresponding autocorrelation function (ACF), $r_{xx}(l)$, can always be decomposed into a zero-mean aperiodic component, $r_{xx}^{(a)}(l)$, and a non-zero-mean periodic component, $r_{xx}^{(p)}(l)$:

$$r_{xx}(l) = r_{xx}^{(a)}(l) + r_{xx}^{(p)}(l)$$

General form of the PSD

A process, $x(n)$, and therefore its corresponding autocorrelation function (ACF), $r_{xx}(l)$, can always be decomposed into a zero-mean aperiodic component, $r_{xx}^{(a)}(l)$, and a non-zero-mean periodic component, $r_{xx}^{(p)}(l)$:

$$r_{xx}(l) = r_{xx}^{(a)}(l) + r_{xx}^{(p)}(l)$$

Theorem (PSD of a non-zero-mean process with periodic component). The most general definition of the PSD for a non-zero-mean stochastic process with a periodic component is

$$P_{xx}(e^{j\omega}) = P_{xx}^{(a)}(e^{j\omega}) + \frac{2\pi}{K} \sum_{k \in \mathcal{K}} P_{xx}^{(p)}(k)\, \delta\left(\omega - \omega_k\right) \qquad \diamondsuit$$

$P_{xx}^{(a)}(e^{j\omega})$ is the DTFT of $r_{xx}^{(a)}(l)$, while $P_{xx}^{(p)}(k)$ are the discrete Fourier transform (DFT) coefficients for $r_{xx}^{(p)}(l)$.

General form of the PSD

Example ([Manolakis:2001, Harmonic Processes, Page 110-111]).
Determine the PSD of the **harmonic process** defined by:

$$x[n] = \sum_{k=1}^{M} A_k \cos(\omega_k n + \phi_k), \quad \omega_k \neq 0$$

⋈

General form of the PSD

Example ([Manolakis:2001, Harmonic Processes, Page 110-111]). Determine the PSD of the **harmonic process** defined by:

$$x[n] = \sum_{k=1}^{M} A_k \cos(\omega_k n + \phi_k), \quad \omega_k \neq 0$$

SOLUTION. $x[n]$ is a stationary process with zero-mean, and autocorrelation function (ACF):

$$r_{xx}[\ell] = \frac{1}{2} \sum_{k=1}^{M} |A_k|^2 \cos \omega_k \ell, \quad -\infty < \ell < \infty$$

□

General form of the PSD

Example ([Manolakis:2001, Harmonic Processes, Page 110-111]). Determine the PSD of the **harmonic process** defined by:

$$x[n] = \sum_{k=1}^{M} A_k \cos(\omega_k n + \phi_k), \quad \omega_k \neq 0$$

SOLUTION. Hence, the ACF can be written as:

$$r_{xx}[\ell] = \sum_{k=-M}^{M} \frac{|A_k|^2}{4} e^{j\omega_k \ell}, \quad -\infty < \ell < \infty \qquad \square$$

where the following are defined: $A_0 = 0$, $A_k = A_{-k}$, and $\omega_{-k} = -\omega_k$.

General form of the PSD

Example ([Manolakis:2001, Harmonic Processes, Page 110-111]). Determine the PSD of the **harmonic process** defined by:

$$x[n] = \sum_{k=1}^{M} A_k \cos(\omega_k n + \phi_k), \quad \omega_k \neq 0$$

SOLUTION. Hence, the ACF can be written as:

$$r_{xx}[\ell] = \sum_{k=-M}^{M} \frac{|A_k|^2}{4} e^{j\omega_k \ell}, \quad -\infty < \ell < \infty$$

where the following are defined: $A_0 = 0$, $A_k = A_{-k}$, and $\omega_{-k} = -\omega_k$.

Hence, it directly follows

$$P_{xx}(e^{j\omega}) = 2\pi \sum_{k=-M}^{M} \frac{|A_k|^2}{4} \delta(\omega - \omega_k) = \frac{\pi}{2} \sum_{k=-M}^{M} |A_k|^2 \delta(\omega - \omega_k)$$

The cross-power spectral density

The cross-power spectral density (CPSD) of two jointly stationary stochastic processes, $x(n)$ and $y(n)$, provides a description of their statistical relations in the frequency domain.

It is defined, naturally, as the DTFT of the cross-correlation, $r_{xy}(\ell) \triangleq \mathbb{E}\left[x(n)\, y^*(n-\ell)\right]$:

$$P_{xy}(e^{j\omega}) = \mathcal{F}\{r_{xy}(\ell)\} = \sum_{l\in\mathbb{Z}} r_{xy}(\ell)\, e^{-j\omega\ell}$$

The cross-power spectral density

The cross-power spectral density (CPSD) of two jointly stationary stochastic processes, $x(n)$ and $y(n)$, provides a description of their statistical relations in the frequency domain.

It is defined, naturally, as the DTFT of the cross-correlation, $r_{xy}(\ell) \triangleq \mathbb{E}\left[x(n)\, y^*(n-\ell)\right]$:

$$P_{xy}(e^{j\omega}) = \mathcal{F}\{r_{xy}(\ell)\} = \sum_{l\in\mathbb{Z}} r_{xy}(\ell)\, e^{-j\omega\ell}$$

The cross-correlation $r_{xy}(l)$ can be recovered by using the inverse-DTFT:

$$r_{xy}(l) = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{xy}(e^{j\omega})\, e^{j\omega l}\, d\omega, \quad l \in \mathbb{R}$$

The cross-power spectral density

The cross-power spectral density (CPSD) of two jointly stationary stochastic processes, $x(n)$ and $y(n)$, provides a description of their statistical relations in the frequency domain.

It is defined, naturally, as the DTFT of the cross-correlation, $r_{xy}(\ell) \triangleq \mathbb{E}\left[x(n)\, y^*(n-\ell)\right]$:

$$P_{xy}(e^{j\omega}) = \mathcal{F}\{r_{xy}(\ell)\} = \sum_{l\in\mathbb{Z}} r_{xy}(\ell)\, e^{-j\omega\ell}$$

The cross-correlation $r_{xy}(l)$ can be recovered by using the inverse-DTFT:

$$r_{xy}(l) = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{xy}(e^{j\omega})\, e^{j\omega l}\, d\omega, \quad l \in \mathbb{R}$$

The cross-spectrum $P_{xy}(e^{j\omega})$ is, in general, a complex function of ω.

The cross-power spectral density

Some properties of the CPSD and related definitions include:

1. $P_{xy}(e^{j\omega})$ is periodic in ω with period 2π.

2. Since $r_{xy}(l) = r^*_{yx}(-l)$, then it follows:

$$P_{xy}(e^{j\omega}) = P^*_{yx}(e^{j\omega})$$

3. If the process $x(n)$ is real, then $r_{xy}(l)$ is real, and:

$$P_{xy}(e^{j\omega}) = P^*_{xy}(e^{-j\omega})$$

4. The **coherence function**, is given by:

$$\Gamma_{xy}(e^{j\omega}) \triangleq \frac{P_{xy}(e^{j\omega})}{\sqrt{P_{xx}(e^{j\omega})}\sqrt{P_{yy}(e^{j\omega})}}$$

Complex Spectral Density Functions

The second moment quantities that described a random process in the transform domain are known as the **complex spectral density** and **complex cross-spectral density** functions.

Complex Spectral Density Functions

The second moment quantities that described a random process in the transform domain are known as the **complex spectral density** and **complex cross-spectral density** functions.

Hence, $r_{xx}(l) \overset{z}{\rightleftharpoons} P_{xx}(z)$ and $r_{xy}(l) \overset{z}{\rightleftharpoons} P_{xy}(z)$, where:

$$P_{xx}(z) = \sum_{l \in \mathbb{Z}} r_{xx}(l)\, z^{-l}$$

$$P_{xy}(z) = \sum_{l \in \mathbb{Z}} r_{xy}(l)\, z^{-l}$$

Complex Spectral Density Functions

The second moment quantities that described a random process in the transform domain are known as the **complex spectral density** and **complex cross-spectral density** functions.

Hence, $r_{xx}(l) \overset{z}{\rightleftharpoons} P_{xx}(z)$ and $r_{xy}(l) \overset{z}{\rightleftharpoons} P_{xy}(z)$, where:

$$P_{xx}(z) = \sum_{l \in \mathbb{Z}} r_{xx}(l)\, z^{-l}$$

$$P_{xy}(z) = \sum_{l \in \mathbb{Z}} r_{xy}(l)\, z^{-l}$$

If the unit circle, defined by $z = e^{j\omega}$ is within the region of convergence of these summations, then:

$$P_{xx}(e^{j\omega}) = P_{xx}(z)\big|_{z=e^{j\omega}}$$

$$P_{xy}(e^{j\omega}) = P_{xy}(z)\big|_{z=e^{j\omega}}$$

Complex Spectral Density Functions

The inverse of the complex spectral and cross-spectral densities are given by the contour integral:

$$r_{xx}(l) = \frac{1}{2\pi j} \oint_C P_{xx}(z)\, z^{l-1}\, dz$$

$$r_{xy}(l) = \frac{1}{2\pi j} \oint_C P_{xy}(z)\, z^{l-1}\, dz$$

Complex Spectral Density Functions

The inverse of the complex spectral and cross-spectral densities are given by the contour integral:

$$r_{xx}(l) = \frac{1}{2\pi j} \oint_C P_{xx}(z)\, z^{l-1}\, dz$$

$$r_{xy}(l) = \frac{1}{2\pi j} \oint_C P_{xy}(z)\, z^{l-1}\, dz$$

Some properties of the complex spectral densities include:

1. Conjugate-symmetry:

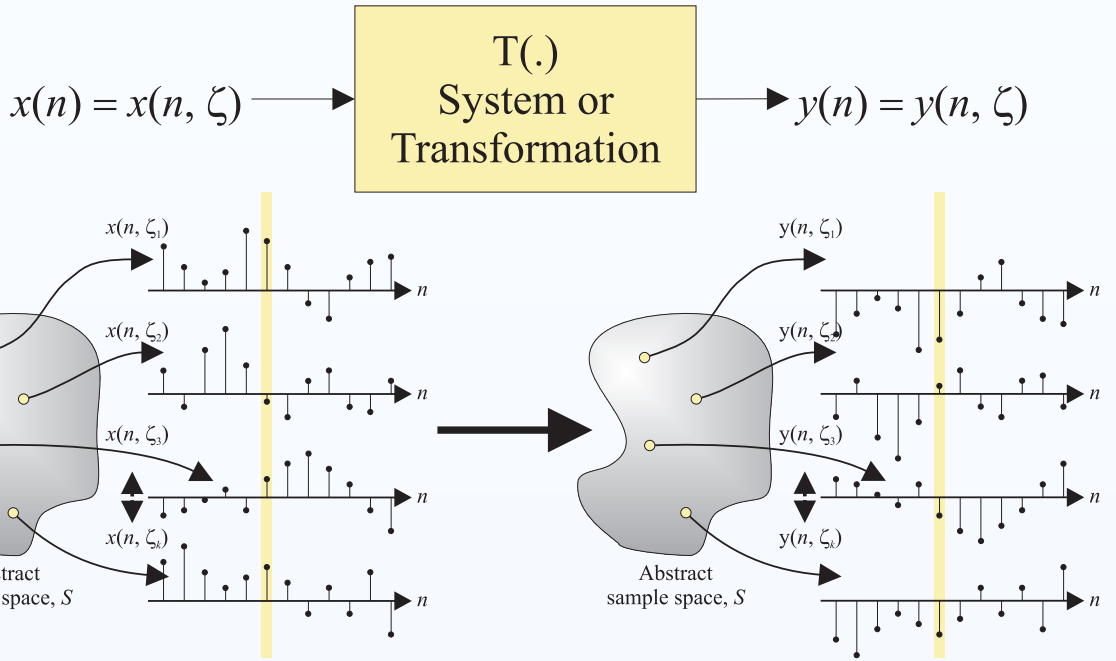
$$P_{xx}(z) = P^*_{xx}(1/z^*) \quad \text{and} \quad P_{xy}(z) = P^*_{xy}(1/z^*)$$

2. For the case when $x(n)$ is real, then:

$$P_{xx}(z) = P_{xx}(z^{-1})$$

Handout 4
Linear Systems Theory

Systems with Stochastic Inputs



A graphical representation of a random process at the output of a system in relation to a random process at the input of the system.

What does it mean to apply a stochastic signal to the input of a system?

Systems with Stochastic Inputs

In principle, the statistics of the output of a system can be expressed in terms of the statistics of the input. However, in general this is a complicated problem except in special cases.

Systems with Stochastic Inputs

In principle, the statistics of the output of a system can be expressed in terms of the statistics of the input. However, in general this is a complicated problem except in special cases.

A special case is that of *linear systems*, and this is considered next.

LTI Systems with Stationary Inputs

Since each sequence (realisation) of a stochastic process is a deterministic signal, there is a well-defined input signal producing a well-defined output signal corresponding to a single realisation of the output stochastic process:

$$y(n,\zeta) = \sum_{k=-\infty}^{\infty} h(k)\, x(n-k,\zeta)$$

LTI Systems with Stationary Inputs

Since each sequence (realisation) of a stochastic process is a deterministic signal, there is a well-defined input signal producing a well-defined output signal corresponding to a single realisation of the output stochastic process:

$$y(n,\zeta) = \sum_{k=-\infty}^{\infty} h(k)\, x(n-k,\zeta)$$

- A complete description of $y(n,\zeta)$ requires the computation of an infinite number of convolutions, corresponding to each value of ζ.

- Thus, a better description would be to consider the statistical properties of $y(n,\zeta)$ in terms of the statistical properties of the input and the characteristics of the system.

LTI Systems with Stationary Inputs

To investigate the statistical input-output properties of a linear system, note the following fundamental theorem:

Theorem (Expectation in Linear Systems). For any linear system,

$$\mathbb{E}\left[L[x(n)]\right] = L[\mathbb{E}\left[x(n)\right]]$$

In other words, the mean $\mu_y(n)$ of the output $y(n)$ equals the response of the system to the mean $\mu_x(n)$ of the input:

$$\mu_y(n) = L[\mu_x(n)] \qquad \diamondsuit$$

Input-output Statistics of a LTI System

If a stationary stochastic process $x[n]$ with mean value μ_x and correlation $r_{xx}[\ell]$ is applied to the input of a LTI system with impulse response $h[n]$ and transfer function $H(e^{j\omega})$, then the:

Output mean value is given by:

$$\mu_y = \mu_x \sum_{k=-\infty}^{\infty} h[k] = \mu_x \, H(e^{j0})$$

Input-output Statistics of a LTI System

If a stationary stochastic process $x[n]$ with mean value μ_x and correlation $r_{xx}[\ell]$ is applied to the input of a LTI system with impulse response $h[n]$ and transfer function $H(e^{j\omega})$, then the:

Output mean value is given by:

$$\mu_y = \mu_x \sum_{k=-\infty}^{\infty} h[k] = \mu_x \, H(e^{j0})$$

Input-output cross-correlation is given by:

$$r_{xy}[\ell] = h^*[-\ell] * r_{xx}[\ell] = \sum_{k=-\infty}^{\infty} h^*[-k] \; r_{xx}[\ell - k]$$

Similarly, it follows that $r_{yx}(l) = h(l) * r_{xx}(l)$.

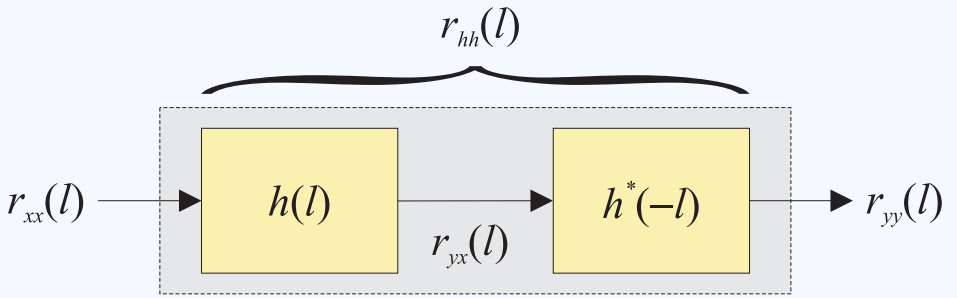
Input-output Statistics of a LTI System

Output autocorrelation is obtained by pre-multiplying the system-output by $y^*(n-l)$ and taking expectations:

$$r_{yy}(l) = \sum_{k=-\infty}^{\infty} h(k)\,\mathbb{E}\left[x(n-k)\,y^*(n-l)\right] = h(l) * r_{xy}(l)$$

Substituting the expression for $r_{xy}(l)$ gives:

$$r_{yy}(l) = h(l) * h^*(-l) * r_{xx}(l) = r_{hh}(l) * r_{xx}(l)$$



An equivalent LTI system for autocorrelation filtration.

Input-output Statistics of a LTI System

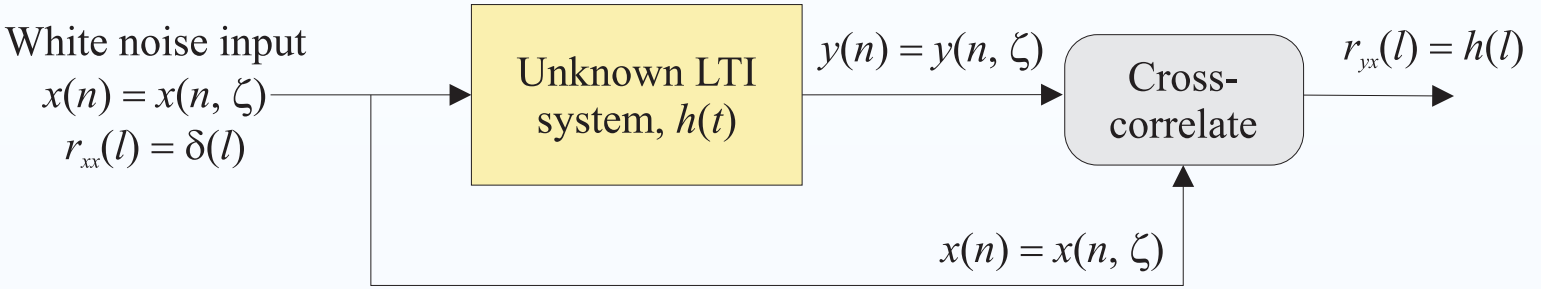
Output-power of the process $y(n)$ is given by $r_{yy}(0) = \mathbb{E}\left[|y(n)|^2\right]$, and therefore since $r_{yy}(l) = r_{hh}(l) * r_{xx}(l)$,

Noting power, P_{yy}, is real, then taking complex-conjugates using $r^*_{xx}(-l) = r_{xx}(l)$:

$$\mathrm{P}_{yy} = \sum_{k=-\infty}^{\infty} r^*_{hh}(k)\, r_{xx}(k) = \sum_{n=-\infty}^{\infty} h^*(n) \sum_{k=-\infty}^{\infty} r_{xx}(n+k)\, h(k)$$

Output pdf In general, it is very difficult to calculate the pdf of the output of a LTI system, except in special cases, namely Gaussian processes.

System identification



System identification by cross-correlation.

The system is excited with a white Gaussian noise (WGN) input with autocorrelation function:

$$r_{xx}(l) = \delta(l)$$

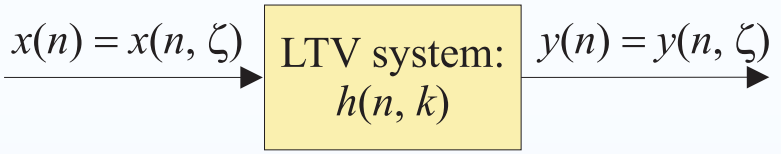
Since the output-input cross-correlation can be written as:

$$r_{yx}(l) = h(l) * r_{xx}(l)$$

then, with $r_{xx}(l) = \delta(l)$, it follows:

$$r_{yx}(l) = h(l) * \delta(l) = h(l)$$

LTV Systems with Nonstationary Inputs



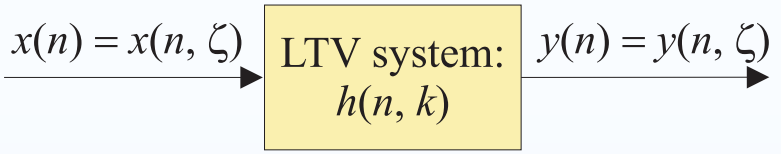
General LTV system with nonstationary input

The input and output are related by the generalised convolution:

$$y(n) = \sum_{k=-\infty}^{\infty} h(n,k)\, x(k)$$

where $h(n,k)$ is the response at time-index n to an impulse occurring at the system input at time-index k.

LTV Systems with Nonstationary Inputs



General LTV system with nonstationary input

The input and output are related by the generalised convolution:

$$y(n) = \sum_{k=-\infty}^{\infty} h(n,k)\, x(k)$$

where $h(n,k)$ is the response at time-index n to an impulse occurring at the system input at time-index k.

The mean, autocorrelation and autocovariance sequences of the output, $y(n)$, as well as the cross-correlation and cross-covariance functions between the input and the output, can be calculated in a similar way as for LTI systems with stationary inputs.

Difference Equation

Consider a LTI system that can be represented by a difference equation:

$$\sum_{p=0}^{P} a_p \, y(n-p) = \sum_{q=0}^{Q} b_q \, x(n-q)$$

where $a_0 \triangleq 1$.

Difference Equation

Consider a LTI system that can be represented by a difference equation:

$$\sum_{p=0}^{P} a_p \, y(n-p) = \sum_{q=0}^{Q} b_q \, x(n-q)$$

where $a_0 \triangleq 1$.

Assuming that both $x(n)$ and $y(n)$ are stationary processes, then taking expectations of both sides gives,

$$\mu_y = \frac{\sum_{q=0}^{Q} b_q}{1 + \sum_{p=1}^{P} a_p} \mu_x$$

Difference Equation

Next, multiplying the system equation throughout by $y^*(m)$ and taking expectations gives:

$$\sum_{p=0}^{P} a_p \, r_{yy}(n-p, m) = \sum_{q=0}^{Q} b_q \, r_{xy}(n-q, m)$$

Difference Equation

Next, multiplying the system equation throughout by $y^*(m)$ and taking expectations gives:

$$\sum_{p=0}^{P} a_p \, r_{yy}(n-p, m) = \sum_{q=0}^{Q} b_q \, r_{xy}(n-q, m)$$

Similarly, instead multiply though by $x^*(m)$ to give:

$$\sum_{p=0}^{P} a_p \, r_{yx}(n-p, m) = \sum_{q=0}^{Q} b_q \, r_{xx}(n-q, m)$$

These two difference equations may be used to solve for $r_{yy}(n_1, n_2)$ and $r_{xy}(n_1, n_2)$. Similar expressions can be obtained for the covariance functions.

Difference Equation

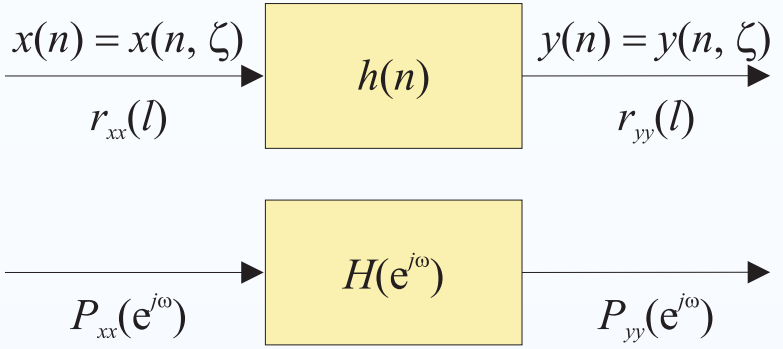
Example ([Manolakis:2000, Example 3.6.2, Page 141]). Let $x(n)$ be a random process generated by the first order difference equation given by:

$$x(n) = \alpha\, x(n-1) + w(n), \quad |\alpha| \leq 1, n \in \mathbb{Z} \qquad \bowtie$$

where $w(n) \sim \mathcal{N}\left(\mu_w, \sigma_w^2\right)$ is an i. i. d. WGN process.

- Demonstrate that the process $x(n)$ is stationary, and determine the mean μ_x.
- Determine the autocovariance and autocorrelation function, $\gamma_{xx}(l)$ and $r_{xx}(l)$.

Frequency-Domain Analysis of LTI systems



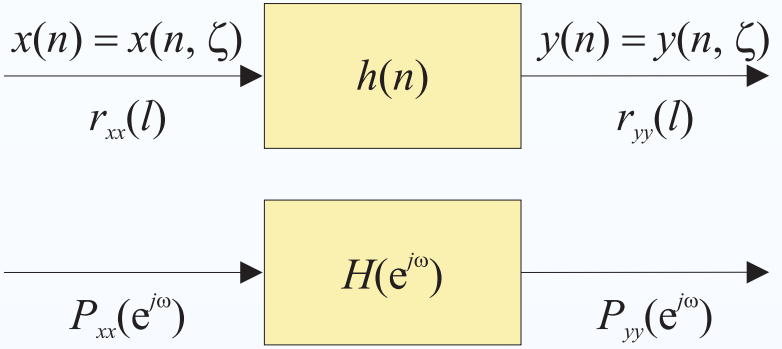
The PSD at the input and output of a LTI system with stationary input.

$$P_{xy}(e^{j\omega}) = H^*(e^{j\omega})\, P_{xx}(e^{j\omega})$$

$$P_{yx}(e^{j\omega}) = H(e^{j\omega})\, P_{xx}(e^{j\omega})$$

$$P_{yy}(e^{j\omega}) = |H(e^{j\omega})|^2\, P_{xx}(e^{j\omega})$$

Frequency-Domain Analysis of LTI systems



The PSD at the input and output of a LTI system with stationary input.

$$P_{xy}(e^{j\omega}) = H^*(e^{j\omega})\, P_{xx}(e^{j\omega})$$

$$P_{yx}(e^{j\omega}) = H(e^{j\omega})\, P_{xx}(e^{j\omega})$$

$$P_{yy}(e^{j\omega}) = |H(e^{j\omega})|^2\, P_{xx}(e^{j\omega})$$

- If the input and output autocorrelations or autospectral densities are known, the magnitude response of a system $|H(e^{j\omega})|$ can be determined, but not the phase response.

Handout 5
Linear Signal Models

Abstract

- This lecture looks at the special class of stationary signals that are obtained by driving a LTI system with white noise. A particular focus is placed on **system function**s that are rational; that is, they can be expressed at the ratio of two polynomials.

Abstract

- This lecture looks at the special class of stationary signals that are obtained by driving a LTI system with white noise. A particular focus is placed on **system function**s that are rational; that is, they can be expressed at the ratio of two polynomials.

- The following models are considered in detail:

Abstract

- This lecture looks at the special class of stationary signals that are obtained by driving a LTI system with white noise. A particular focus is placed on **system function**s that are rational; that is, they can be expressed at the ratio of two polynomials.

- The following models are considered in detail:

  - **All-pole** systems and autoregressive (AR) processes;

Abstract

- This lecture looks at the special class of stationary signals that are obtained by driving a LTI system with white noise. A particular focus is placed on **system function**s that are rational; that is, they can be expressed at the ratio of two polynomials.

- The following models are considered in detail:
  - **All-pole** systems and autoregressive (AR) processes;
  - **All-zero** systems and moving average (MA) processes;

Abstract

- This lecture looks at the special class of stationary signals that are obtained by driving a LTI system with white noise. A particular focus is placed on **system function**s that are rational; that is, they can be expressed at the ratio of two polynomials.

- The following models are considered in detail:
  - **All-pole** systems and autoregressive (AR) processes;
  - **All-zero** systems and moving average (MA) processes;
  - and **pole-zero** systems and autoregressive moving average (ARMA) processes.

Abstract

- This lecture looks at the special class of stationary signals that are obtained by driving a LTI system with white noise. A particular focus is placed on **system function**s that are rational; that is, they can be expressed at the ratio of two polynomials.

- The following models are considered in detail:
  - **All-pole** systems and autoregressive (AR) processes;
  - **All-zero** systems and moving average (MA) processes;
  - and **pole-zero** systems and autoregressive moving average (ARMA) processes.

- **Pole-zero** models are widely used for modelling stationary signals with short memory; the concepts will be extended, in overview at least, to nonstationary processes.

Abstract

- This lecture looks at the special class of stationary signals that are obtained by driving a LTI system with white noise. A particular focus is placed on **system function**s that are rational; that is, they can be expressed at the ratio of two polynomials.

- The following models are considered in detail:
  - **All-pole** systems and autoregressive (AR) processes;
  - **All-zero** systems and moving average (MA) processes;
  - and **pole-zero** systems and autoregressive moving average (ARMA) processes.

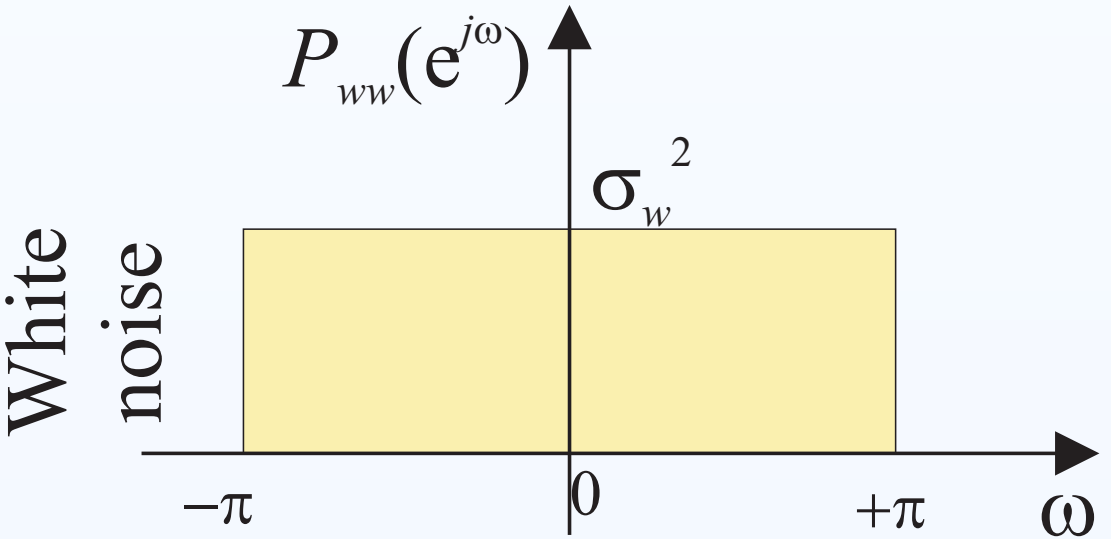
- **Pole-zero** models are widely used for modelling stationary signals with short memory; the concepts will be extended, in overview at least, to nonstationary processes.

The Ubiquitous WGN Sequence

The simplest random signal model is the WSS WGN sequence:

$$w(n) \sim \mathcal{N}\left(0,\, \sigma_w^2\right)$$

The sequence is i. i. d., and $P_{ww}(e^{j\omega}) = \sigma_w^2,\ -\pi < \omega \leq \pi.$



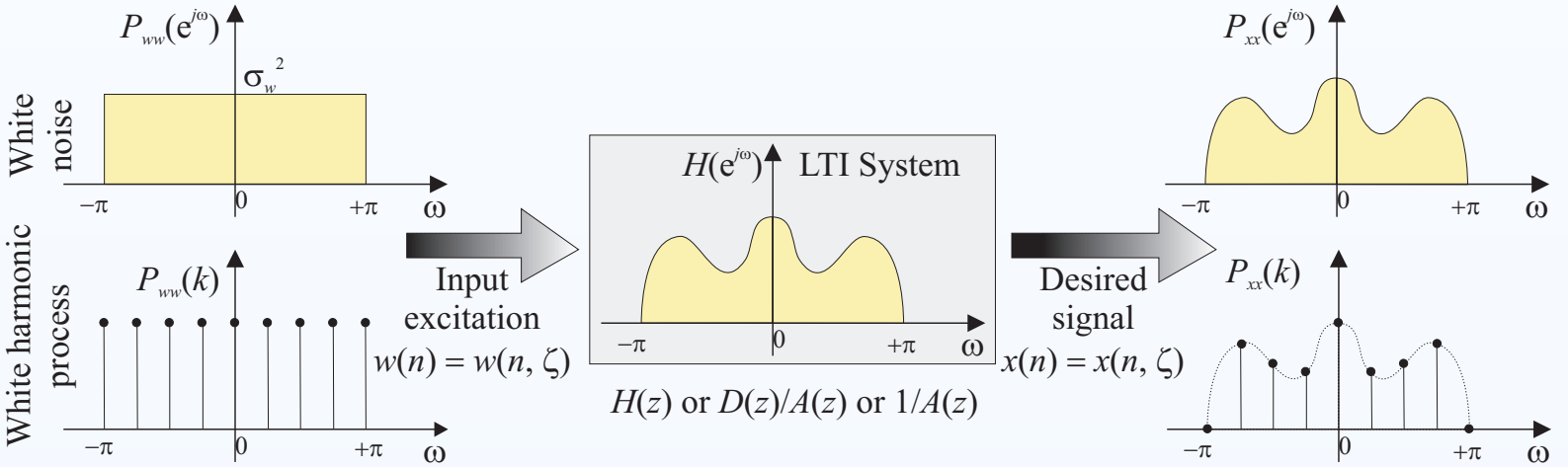
White noise PSD.

Filtration of WGN

By filtering a WGN through a stable LTI system, it is possible to obtain a stochastic signal at the output with almost any arbitrary aperiodic correlation function or continuous PSD.

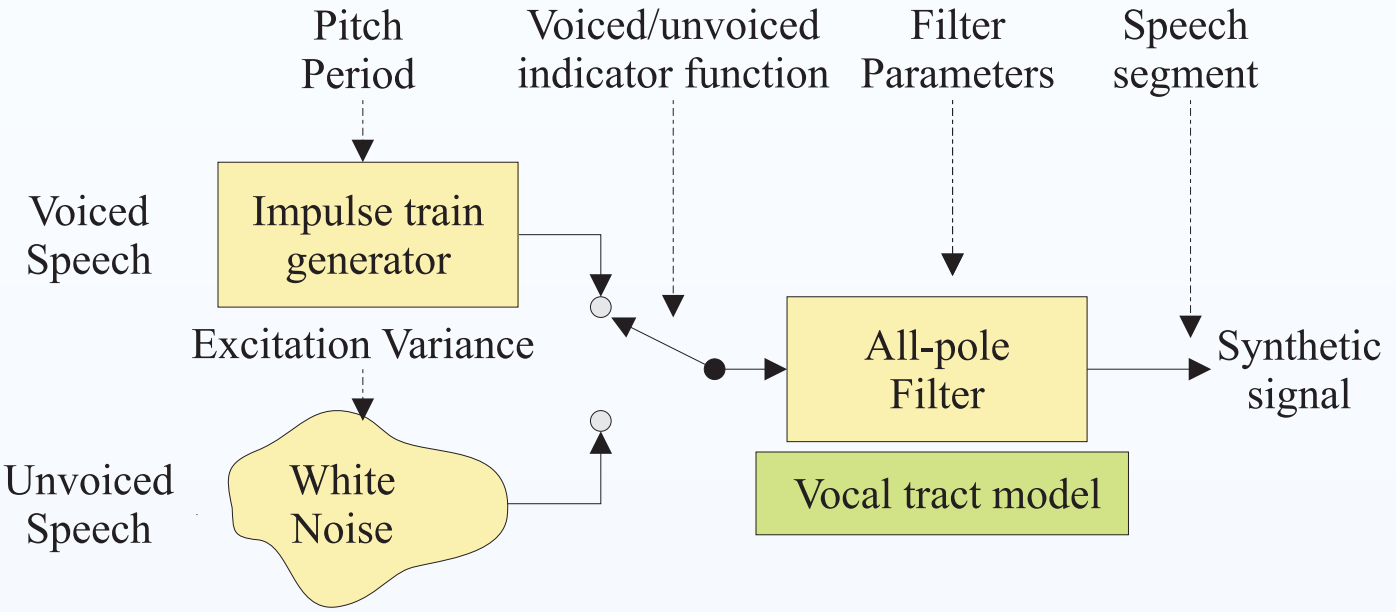
Filtration of WGN

By filtering a WGN through a stable LTI system, it is possible to obtain a stochastic signal at the output with almost any arbitrary aperiodic correlation function or continuous PSD.



Signal models with continuous and discrete (line) power spectrum densities.

Filtration of WGN



The speech synthesis model.

Nonparametric and parametric models

Nonparametric models have no restriction on its form, or the number of parameters characterising the model. For example, specifying a LTI filter by its impulse response is a nonparametric model.

Parametric models, on the other hand, describe a system with a finite number of parameters. For example, if a LTI filter is specified by a finite-order rational **system function**, it is a parametric model.

Nonparametric and parametric models

Nonparametric models have no restriction on its form, or the number of parameters characterising the model. For example, specifying a LTI filter by its impulse response is a nonparametric model.

Parametric models, on the other hand, describe a system with a finite number of parameters. For example, if a LTI filter is specified by a finite-order rational **system function**, it is a parametric model.

Two important analysis tools present themselves for parametric modelling:

1. given the model parameters, analyse the characteristics of that model (in terms of moments etc.);

2. design of a parametric system model to produce a random signal with a specified **autocorrelation** function or PSD. This is **signal modelling**.

Parametric Pole-Zero Signal Models

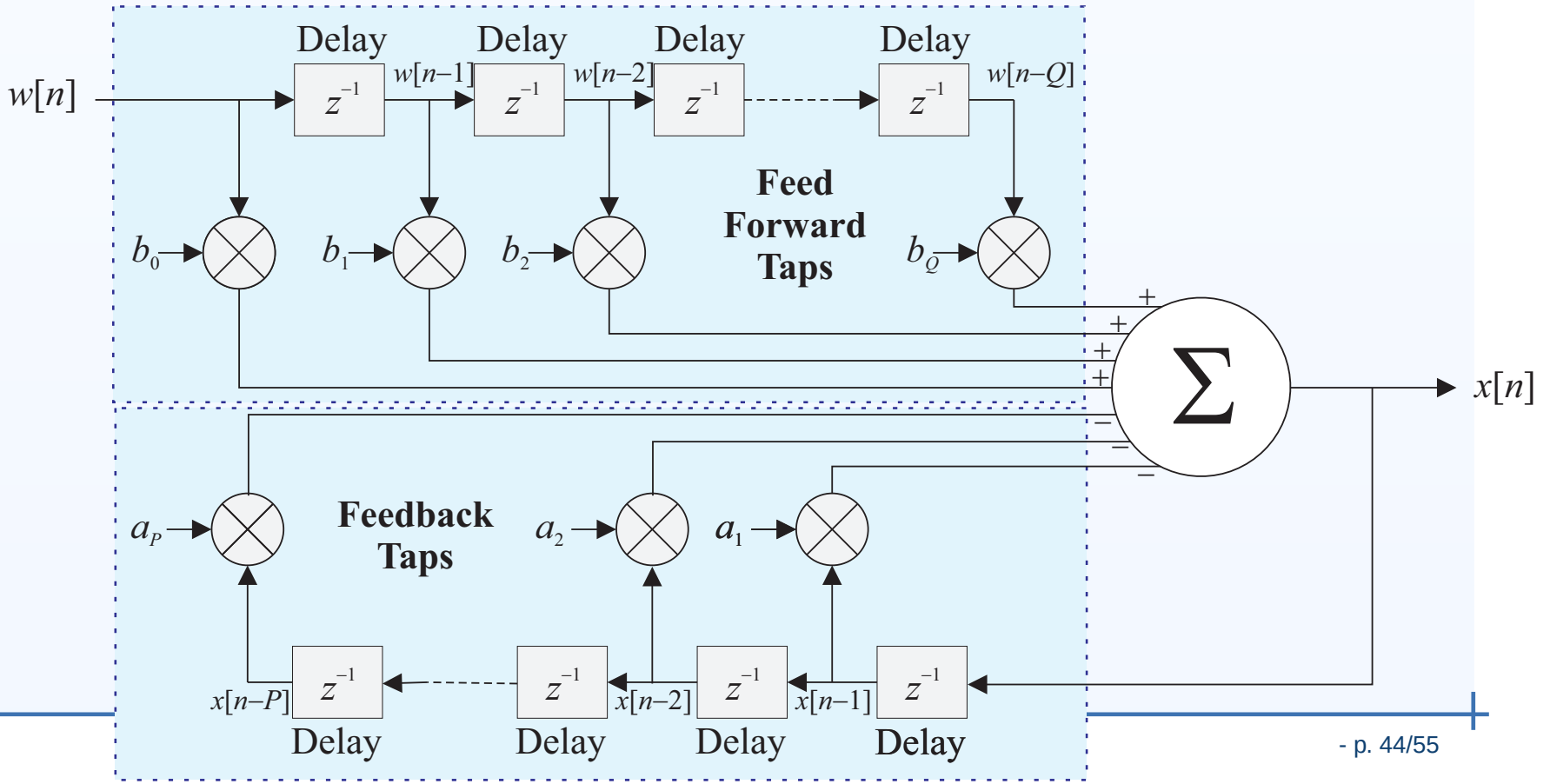
Consider a system described by the following linear constant-coefficient difference equation:

$$x(n) = -\sum_{k=1}^{P} a_k\, x(n-k) + \sum_{k=0}^{Q} d_k\, w(n-k)$$

Parametric Pole-Zero Signal Models

Consider a system described by the following linear constant-coefficient difference equation:

$$x(n) = -\sum_{k=1}^{P} a_k\, x(n-k) + \sum_{k=0}^{Q} d_k\, w(n-k)$$



Types of pole-zero models

All-pole model when $Q = 0$. The input-output difference equation is given by:

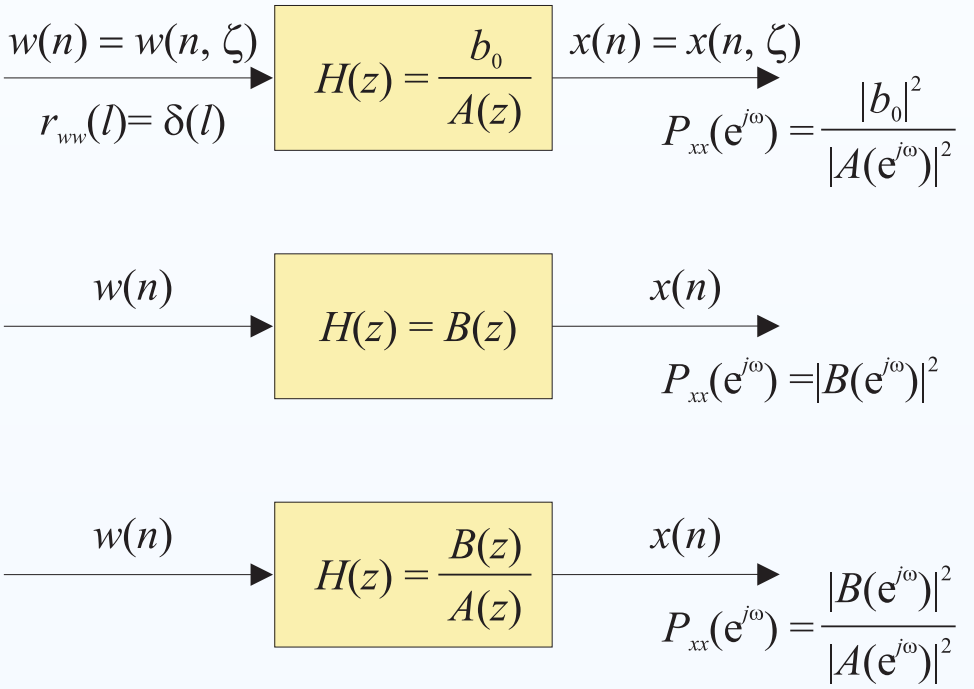
$$x(n) = -\sum_{k=1}^{P} a_k \, x(n-k) + d_0 \, w(n)$$

All-zero model when $P = 0$. The input-output relation is given by:

$$x(n) = \sum_{k=0}^{Q} d_k \, w(n-k)$$

Pole-zero model when $P > 0$ and $Q > 0$.

Types of pole-zero models



Different types of linear model

Types of pole-zero models

If a parametric model is *excited* with WGN, the resulting output signal has second-order moments determined by the parameters of the model.

Types of pole-zero models

If a parametric model is *excited* with WGN, the resulting output signal has second-order moments determined by the parameters of the model.

These **stochastic processes** have special names in the literature, and are known as:

a moving average (MA) process when it is the output of an all-zero model;

an autoregressive (AR) process when it is the output of an all-pole model;

an autoregressive moving average (ARMA) process when it is the output of an pole-zero model;

each subject to a WGN process at the input.

All-pole Models

All-pole models are frequently used in signal processing applications since they are:

- mathematically convenient since model parameters can be estimated by solving a set of linear equations, and
- they widely parsimoniously approximate rational transfer functions, especially resonant systems.

All-pole Models

All-pole models are frequently used in signal processing applications since they are:

- mathematically convenient since model parameters can be estimated by solving a set of linear equations, and
- they widely parsimoniously approximate rational transfer functions, especially resonant systems.

There are various model properties of the all-pole model that are useful; these include:

1. the systems impulse response;
2. the autocorrelation of the impulse response;
3. and minimum-phase conditions.

Frequency Response of an All-Pole Filter

The all-pole model has form:

$$H(z) = \frac{d_0}{A(z)} = \frac{d_0}{1 + \sum_{k=1}^{P} a_k \, z^{-k}} = \frac{d_0}{\prod_{k=1}^{P} (1 - p_k \, z^{-1})}$$

and therefore its frequency response is given by:

$$H(e^{j\omega}) = \frac{d_0}{1 + \sum_{k=1}^{P} a_k \, e^{-jk\omega}} = \frac{d_0}{\prod_{k=1}^{P} (1 - p_k \, e^{-j\omega})}$$

Frequency Response of an All-Pole Filter

The all-pole model has form:

$$H(z) = \frac{d_0}{A(z)} = \frac{d_0}{1 + \sum_{k=1}^{P} a_k \, z^{-k}} = \frac{d_0}{\prod_{k=1}^{P} (1 - p_k \, z^{-1})}$$

and therefore its frequency response is given by:

$$H(e^{j\omega}) = \frac{d_0}{1 + \sum_{k=1}^{P} a_k \, e^{-jk\omega}} = \frac{d_0}{\prod_{k=1}^{P} (1 - p_k \, e^{-j\omega})}$$

When each of the poles are written in the form $p_k = r_k e^{j\omega_k}$, then the frequency response can be written as:

$$H(e^{j\omega}) = \frac{d_0}{\prod_{k=1}^{P} (1 - r_k \, e^{-j(\omega - \omega_k)})}$$

Hence, it can be deduced that resonances occur near the frequencies corresponding to the phase position of the poles.

Frequency Response of an All-Pole Filter

Hence, the PSD of the output of an all-pole filter is given by:

$$P_{xx}(e^{j\omega}) = \sigma_w^2 \left|H(e^{j\omega})\right|^2 = \frac{G^2}{\prod_{k=1}^{P} \left|1 - r_k\, e^{-j(\omega-\omega_k)}\right|^2}$$

where $G = \sigma_w\, d_0$ is the overall gain of the system.

Frequency Response of an All-Pole Filter

Hence, the PSD of the output of an all-pole filter is given by:

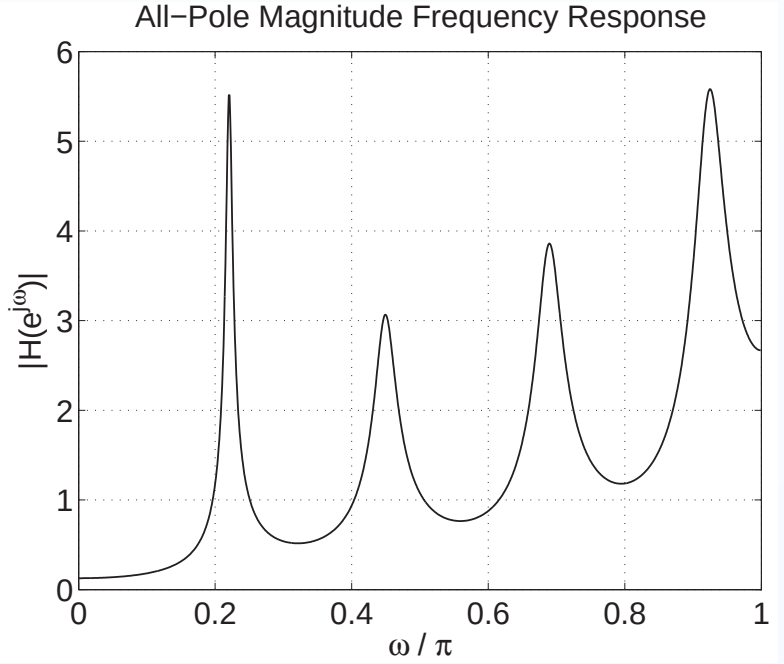
$$P_{xx}(e^{j\omega}) = \sigma_w^2 \left|H(e^{j\omega})\right|^2 = \frac{G^2}{\prod_{k=1}^{P} \left|1 - r_k\, e^{-j(\omega-\omega_k)}\right|^2}$$

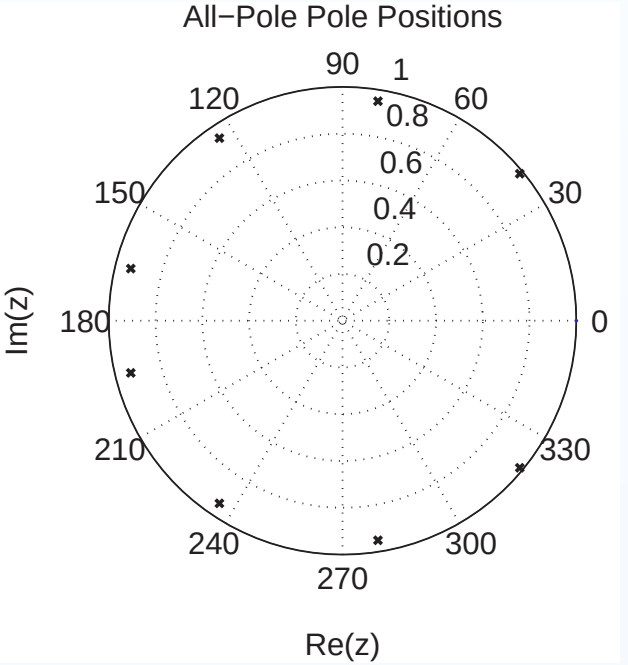
where $G = \sigma_w\, d_0$ is the overall gain of the system.

Consider the all-pole model with poles at positions:

$$\{p_k\} = \{r_k\, e^{j\omega_k}\} \quad \text{where} \quad \begin{cases} \{r_k\} &= \{0.985, 0.951, 0.942, 0.933\} \\ \{\omega_k\} &= 2\pi \times \{270, 550, 844, 1131\}/2450; \end{cases}$$

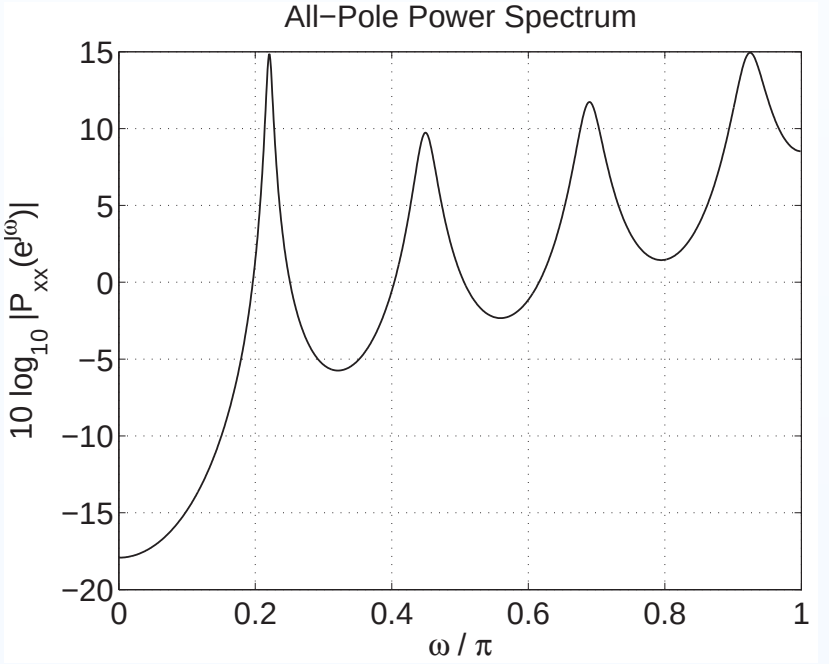
Frequency Response of an All-Pole Filter





The frequency response and position of the poles in an all-pole system.

Frequency Response of an All-Pole Filter



Power spectral response of an all-pole model.

Impulse Response of an All-Pole Filter

The impulse response of the all-pole filter satisfies the equation:

$$h(n) = -\sum_{k=1}^{P} a_k\, h(n-k) + d_0\, \delta(n)$$

If $H(z)$ has its poles inside the unit circle, then $h(n)$ is a causal, stable sequence, and the system is **minimum-phase**.

Impulse Response of an All-Pole Filter

The impulse response of the all-pole filter satisfies the equation:

$$h(n) = -\sum_{k=1}^{P} a_k\, h(n-k) + d_0\, \delta(n)$$

If $H(z)$ has its poles inside the unit circle, then $h(n)$ is a causal, stable sequence, and the system is **minimum-phase**.

Assuming causality, such that $h(n) = 0, \quad n < 0$ then it follows $h(-k) = 0,\ k > 0$, and therefore:

$$h(n) = \begin{cases} 0 & \text{if } n < 0 \\ d_0 & \text{if } n = 0 \\ -\sum_{k=1}^{P} a_k\, h(n-k) & \text{if } n > 0 \end{cases}$$

All-Pole Modelling and Linear Prediction

A **linear predictor** forms an estimate, or *prediction*, $\hat{x}(n)$, of the present value of a stochastic process $x(n)$ from a linear combination of the past P samples; that is:

$$\hat{x}(n) = -\sum_{k=1}^{P} a_k \, x(n-k)$$

All-Pole Modelling and Linear Prediction

A **linear predictor** forms an estimate, or *prediction*, $\hat{x}(n)$, of the present value of a stochastic process $x(n)$ from a linear combination of the past P samples; that is:

$$\hat{x}(n) = -\sum_{k=1}^{P} a_k \, x(n-k)$$

The coefficients $\{a_k\}$ of the linear predictor are determined by attempting to minimise some function of the **prediction error** given by:

$$e(n) = x(n) - \hat{x}(n)$$

Usually the objective function is equivalent to mean-squared error (MSE), given by $E = \sum_n e^2(n)$.

All-Pole Modelling and Linear Prediction

Hence, the prediction error can be written as:

$$e(n) = x(n) - \hat{x}(n) = x(n) + \sum_{k=1}^{P} a_k \, x(n-k)$$

All-Pole Modelling and Linear Prediction

Hence, the prediction error can be written as:

$$e(n) = x(n) - \hat{x}(n) = x(n) + \sum_{k=1}^{P} a_k \, x(n-k)$$

- Thus, the prediction error is equal to the excitation of the all-pole model; $e(n) = w(n)$. Clearly, **finite impulse response (FIR) linear prediction** and **all-pole modelling** are closely related.

- Many of the properties and algorithms developed for either **linear prediction** or **all-pole modelling** can be applied to the other.

All-Pole Modelling and Linear Prediction

Hence, the prediction error can be written as:

$$e(n) = x(n) - \hat{x}(n) = x(n) + \sum_{k=1}^{P} a_k \, x(n-k)$$

- Thus, the prediction error is equal to the excitation of the all-pole model; $e(n) = w(n)$. Clearly, **FIR linear prediction** and **all-pole modelling** are closely related.

- Many of the properties and algorithms developed for either **linear prediction** or **all-pole modelling** can be applied to the other.

- To all intents and purposes, linear prediction, **all-pole modelling**, and **AR processes** (discussed next) are equivalent terms for the same concept.

Autoregressive Processes

While **all-pole models** refer to the properties of a rational system containing only poles, **AR processes** refer to the resulting stochastic process that occurs as the result of **WGN** being applied to the input of an **all-pole filter**.

Autoregressive Processes

While **all-pole models** refer to the properties of a rational system containing only poles, **AR processes** refer to the resulting stochastic process that occurs as the result of **WGN** being applied to the input of an **all-pole filter**.

As such, the same input-output equations for all-pole models still apply.

Autoregressive Processes

While **all-pole models** refer to the properties of a rational system containing only poles, **AR processes** refer to the resulting stochastic process that occurs as the result of **WGN** being applied to the input of an **all-pole filter**.

As such, the same input-output equations for all-pole models still apply.

Thus:

$$x(n) = -\sum_{k=1}^{P} a_k\, x(n-k) + w(n), \quad w(n) \sim \mathcal{N}\left(0,\, \sigma_w^2\right)$$

The autoregressive output, $x(n)$, is a stationary sequence with a mean value of zero, $\mu_x = 0$.

Autoregressive Processes

While **all-pole models** refer to the properties of a rational system containing only poles, **AR processes** refer to the resulting stochastic process that occurs as the result of **WGN** being applied to the input of an **all-pole filter**.

Thus:

$$x(n) = -\sum_{k=1}^{P} a_k\, x(n-k) + w(n), \quad w(n) \sim \mathcal{N}\left(0,\, \sigma_w^2\right)$$

The autoregressive output, $x(n)$, is a stationary sequence with a mean value of zero, $\mu_x = 0$.

The autocorrelation function (ACF) can be calculated in a similar approach to finding the output autocorrelation and cross-correlation for linear systems.

Autoregressive Processes

Multiply the difference through by $x^*(n-l)$ and take expectations to obtain:

$$r_{xx}(l) + \sum_{k=1}^{P} a_k \, r_{xx}(l-k) = r_{wx}(l)$$

Autoregressive Processes

Multiply the difference through by $x^*(n-l)$ and take expectations to obtain:

$$r_{xx}(l) + \sum_{k=1}^{P} a_k \, r_{xx}(l-k) = r_{wx}(l)$$

Observing that $x(n)$ cannot depend on future values of $w(n)$ since the system is causal, then $r_{wx}(l) = \mathbb{E}\left[w(n)\, x^*(n-l)\right]$ is zero if $l > 0$, and σ_w^2 if $l = 0$.

Autoregressive Processes

Multiply the difference through by $x^*(n-l)$ and take expectations to obtain:

$$r_{xx}(l) + \sum_{k=1}^{P} a_k \, r_{xx}(l-k) = r_{wx}(l)$$

Thus, writing Equation **??** for $l = \{0, 1, \ldots, P\}$

matrix-vector form (noting that $r_{xx}(l) = r_{xx}^*(-l)$ and that the parameters $\{a_k\}$ are real) as:

$$\begin{bmatrix} r_{xx}(0) & r_{xx}^*(1) & \cdots & r_{xx}^*(P) \\ r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}^*(P-1) \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}(P) & r_{xx}(P-1) & \cdots & r_{xx}^*(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_P \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

All-Zero models

Whereas **all-pole** models can capture resonant features of a particular PSD, it cannot capture *nulls* in the frequency response. These can only be modelled using a pole-zero or **all-zero** model.

All-Zero models

Whereas **all-pole** models can capture resonant features of a particular PSD, it cannot capture *nulls* in the frequency response. These can only be modelled using a pole-zero or **all-zero** model.

The output of an all-zero model is the weighted average of delayed versions of the input signal. Thus, assume an all-zero model of the form:

$$x(n) = \sum_{k=0}^{Q} d_k \, w(n-k)$$

where Q is the order of the model, and the corresponding system function is given by:

$$H(z) = D(z) = \sum_{k=0}^{Q} d_k \, z^{-k}$$

Frequency Response of an All-Zero Filter

The all-zero model has form:

$$H(z) = D(z) = \sum_{k=0}^{Q} d_k \, z^{-k} = d_0 \prod_{k=1}^{Q} \left(1 - z_k \, z^{-1}\right)$$

Therefore, its frequency response is given by:

$$H(e^{j\omega}) = \sum_{k=0}^{Q} d_k \, e^{-jk\omega} = d_0 \prod_{k=1}^{Q} \left(1 - z_k \, e^{-j\omega}\right)$$

Frequency Response of an All-Zero Filter

The all-zero model has form:

$$H(z) = D(z) = \sum_{k=0}^{Q} d_k \, z^{-k} = d_0 \prod_{k=1}^{Q} \left(1 - z_k \, z^{-1}\right)$$

Therefore, its frequency response is given by:

$$H(e^{j\omega}) = \sum_{k=0}^{Q} d_k \, e^{-jk\omega} = d_0 \prod_{k=1}^{Q} \left(1 - z_k \, e^{-j\omega}\right)$$

When each of the zeros are written in the form $z_k = r_k e^{j\omega_k}$, then the frequency response can be written as:

$$H(e^{j\omega}) = d_0 \prod_{k=1}^{Q} \left(1 - r_k \, e^{-j(\omega - \omega_k)}\right)$$

Hence, it can be deduced that troughs or nulls occur near frequencies corresponding to the phase position of the zeros.

Frequency Response of an All-Zero Filter

Hence, the PSD of the output of an all-zero filter is given by:

$$P_{xx}(e^{j\omega}) = \sigma_w^2 \left|H(e^{j\omega})\right|^2 = G^2 \prod_{k=1}^{Q} \left|1 - r_k\, e^{-j(\omega-\omega_k)}\right|^2$$

where $G = \sigma_w\, d_0$ is the overall gain of the system.

Frequency Response of an All-Zero Filter

Hence, the PSD of the output of an all-zero filter is given by:

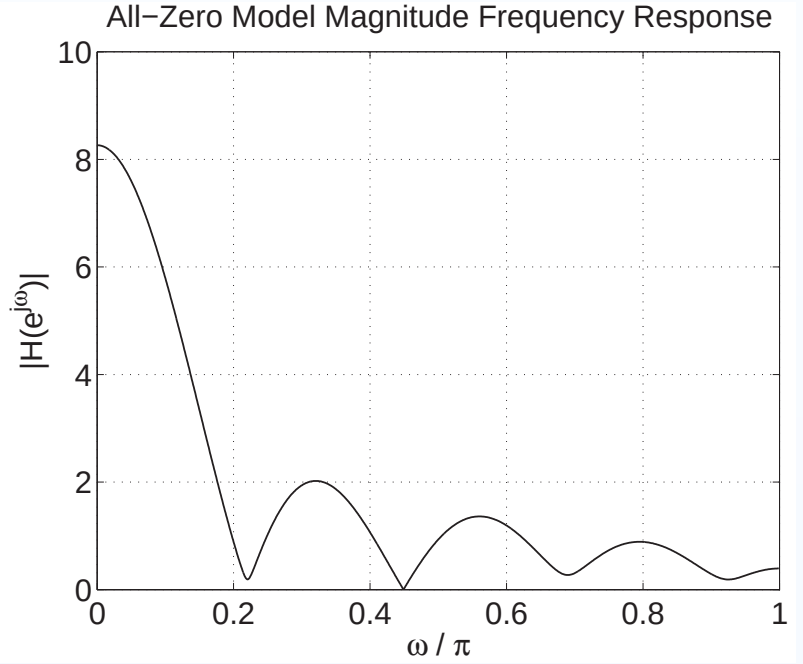
$$P_{xx}(e^{j\omega}) = \sigma_w^2 \left|H(e^{j\omega})\right|^2 = G^2 \prod_{k=1}^{Q} \left|1 - r_k\, e^{-j(\omega-\omega_k)}\right|^2$$

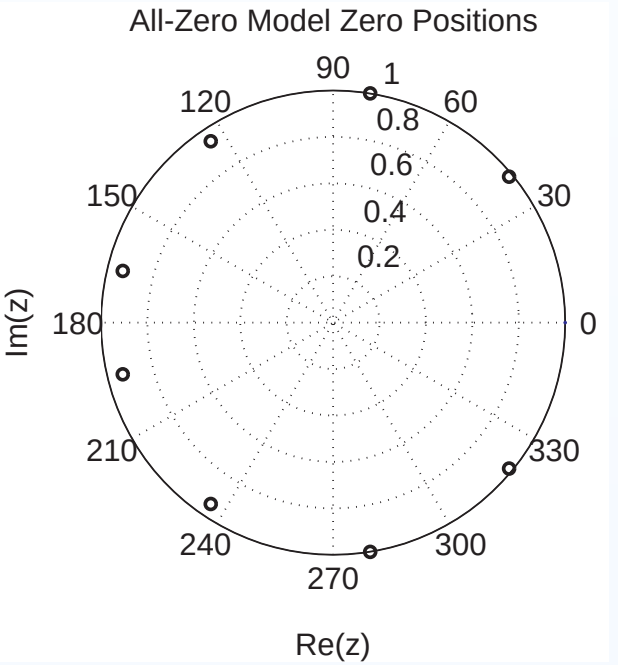
where $G = \sigma_w\, d_0$ is the overall gain of the system.

Consider the all-zero model with zeros at positions:

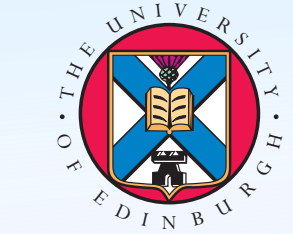
$$\{z_k\} = \{r_k\, e^{j\omega_k}\} \quad \text{where} \quad \begin{cases} \{r_k\} = & \{0.985, 1, 0.942, 0.933\} \\ \{\omega_k\} = & 2\pi \times \{270, 550, 844, 1131\}/2450; \end{cases}$$

Frequency Response of an All-Zero Filter

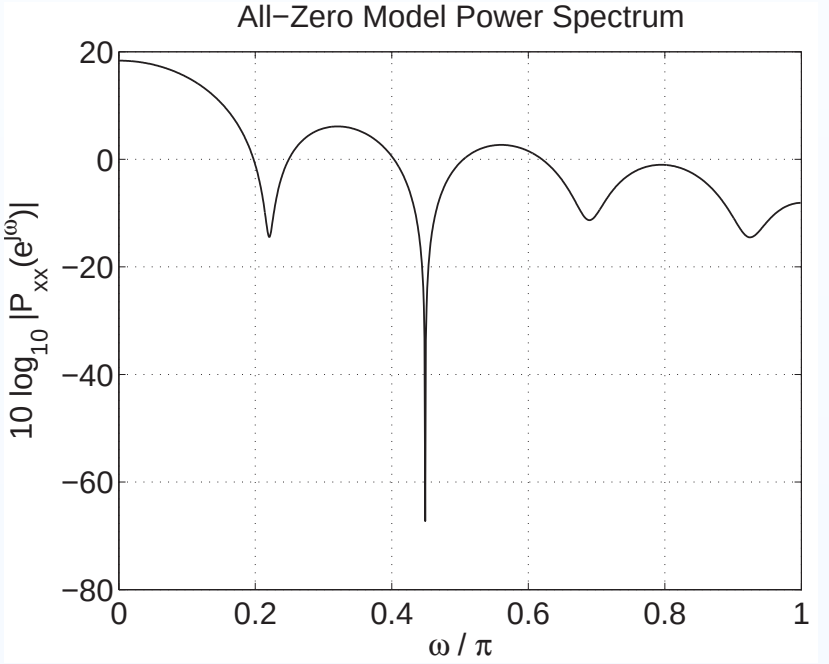




The frequency response and position of the zeros in an all-zero system.



Frequency Response of an All-Zero Filter



Power spectral response of an all-zero model.

Moving-average processes

A **MA process** refers to the stochastic process that is obtained at the output of an **all-zero filter** when a WGN sequence is applied to the input.

Moving-average processes

A **MA process** refers to the stochastic process that is obtained at the output of an **all-zero filter** when a WGN sequence is applied to the input.

Thus, a MA process is an $AZ(Q)$ model with $d_0 = 1$ driven by WGN.

$$x[n] = w[n] + \sum_{k=1}^{Q} d_k \, w[n-k] \,, \quad w[n] \sim \mathcal{N}\left(0, \, \sigma_w^2\right)$$

Moving-average processes

A **MA process** refers to the stochastic process that is obtained at the output of an **all-zero filter** when a WGN sequence is applied to the input.

Thus, a MA process is an $AZ(Q)$ model with $d_0 = 1$ driven by WGN.

$$x[n] = w[n] + \sum_{k=1}^{Q} d_k \, w[n-k], \quad w[n] \sim \mathcal{N}\left(0, \sigma_w^2\right)$$

The output $x(n)$ has zero-mean, and variance of

$$\sigma_x^2 = \sigma_w^2 \left[1 + \sum_{k=1}^{Q} |d_k|^2\right]$$

The autocorrelation function is given by:

$$r_{xx}[\ell] = \sigma_w^2 r_{hh}[\ell] = \sigma_w^2 \sum^{Q-\ell} d_{k+l} \, d_k^*, \qquad \text{for } 0 \leq \ell \leq Q$$

Pole-Zero Models

The output of a causal pole-zero model is given by the recursive input-output relationship:

$$x[n] = -\sum_{k=1}^{P} a_k\, x[n-k] + \sum_{k=0}^{Q} d_k\, w[n-k]$$

The corresponding system function is given by:

$$H(z) = \frac{D(z)}{A(z)} = \frac{\sum_{k=0}^{Q} d_k\, z^{-k}}{1 + \sum_{k=1}^{P} a_k\, z^{-k}}$$

Pole-Zero Frequency Response

The pole-zero model can be written as

$$H(z) = \frac{D(z)}{A(z)} = d_0 \frac{\prod_{k=1}^{Q} \left(1 - z_k\, z^{-1}\right)}{\prod_{k=1}^{P} \left(1 - p_k\, z^{-1}\right)}$$

Therefore, its frequency response is given by:

$$H(e^{j\omega}) = d_0 \frac{\prod_{k=1}^{Q} \left(1 - z_k\, e^{-j\omega}\right)}{\prod_{k=1}^{P} \left(1 - p_k\, e^{-j\omega}\right)}$$

Pole-Zero Frequency Response

The pole-zero model can be written as

$$H(z) = \frac{D(z)}{A(z)} = d_0 \frac{\prod_{k=1}^{Q} \left(1 - z_k\, z^{-1}\right)}{\prod_{k=1}^{P} \left(1 - p_k\, z^{-1}\right)}$$

Therefore, its frequency response is given by:

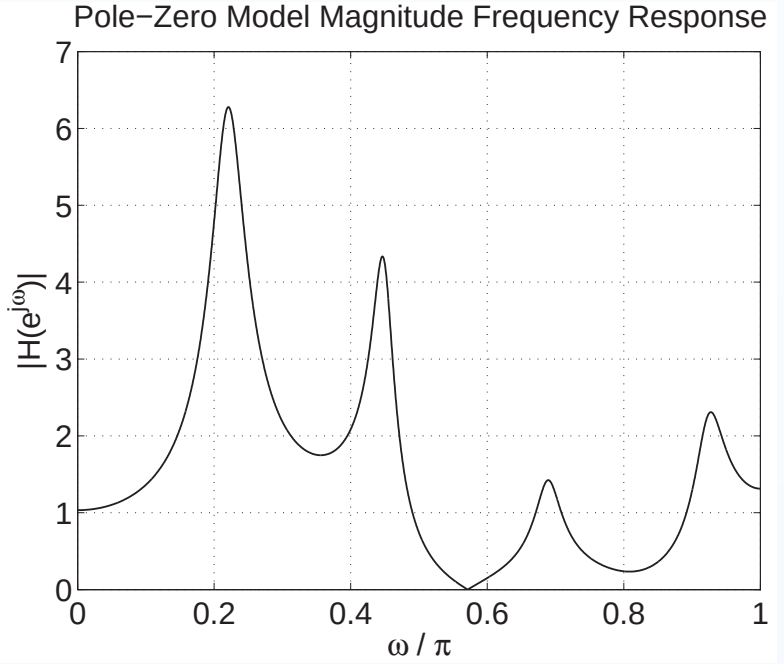
$$H(e^{j\omega}) = d_0 \frac{\prod_{k=1}^{Q} \left(1 - z_k\, e^{-j\omega}\right)}{\prod_{k=1}^{P} \left(1 - p_k\, e^{-j\omega}\right)}$$

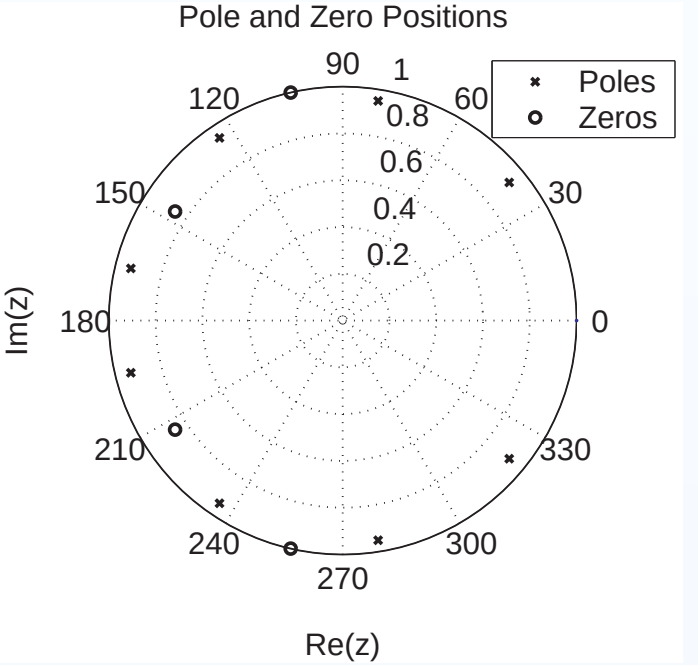
The PSD of the output of a pole-zero filter is given by:

$$P_{xx}(e^{j\omega}) = \sigma_w^2 \left|H(e^{j\omega})\right|^2 = G^2 \frac{\prod_{k=1}^{Q} \left|1 - z_k\, e^{-j\omega}\right|^2}{\prod_{k=1}^{P} \left|1 - p_k\, e^{-j\omega}\right|^2}$$

where $G = \sigma_w\, d_0$ is the overall gain of the system.

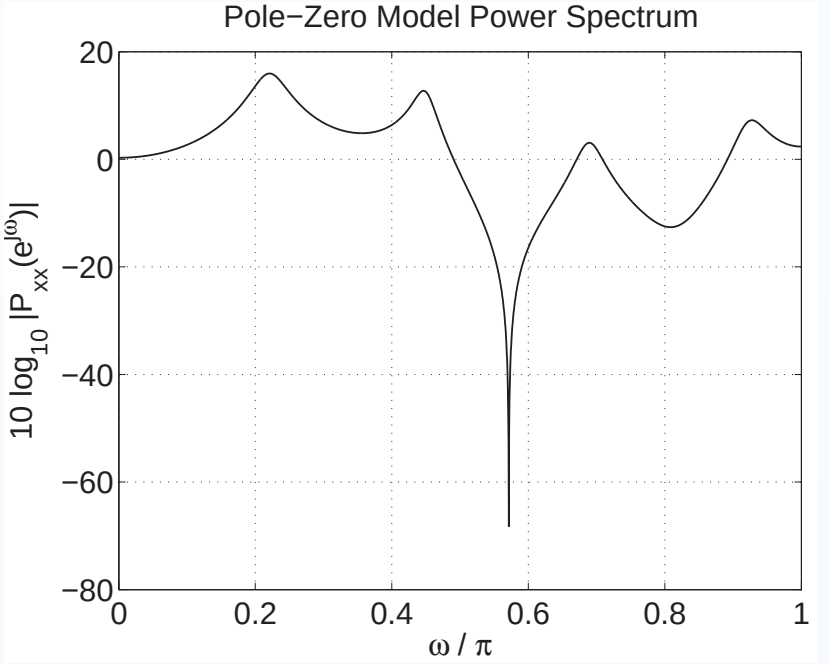
Pole-Zero Frequency Response





The frequency response and position of the poles and zeros in an pole-zero system.

Pole-Zero Frequency Response



Power spectral response of an pole-zero model.