Statistical Signal Processing

Lecture 2:
Stochastic Processes (or Random Signals)

Carlas Smith & Peyman Mohajerin Esfahani

Delft Center for Systems and Control

Delft University of Technology

Part I:

1. **What is a Random Process (RP)**
2. Characterization of multiple (2) RPs
3. Gaussian Processes
4. Stationairy RPs
5. Property of Wide Sense Stationary (WSS) RPs
6. Autocorrelation matrix of a WSS RP.

Definition Random Process (part 1)

Definition (part 1): A Random Process (RP) $x(n)$ is an indexed sequence of random variables,

$$\cdots, x(-2), x(-1), x(0), x(1), x(2), \cdots$$

with each $x(n)$ a **random variable** (see Lecture 2). This RV is characterized by a PDF or pdf. This is a mapping from its sample space Ω_n ($\subset$ real (complex) numbers) $\rightarrow \mathbb{R}^+$,

$$F_{x(n)}(\alpha) = Pr(x(n) \leq \alpha) \quad f_{x(n)}(\alpha) = \frac{dF_{x(n)}(\alpha)}{d\alpha}$$

Definition of Realization: A realization of a RP is the indexed sequence $x(n)$ with a single sample for each time instance n,

$\{x_i(n)\}$(denoted in short as $\{x(n)\}$)   `Brownian_bup`

Examples of Random Processes

Example 1: Let A be the outcome of a roll of a die, then for a given ω_0,

$$x(n) = Acos(n\omega_0)$$

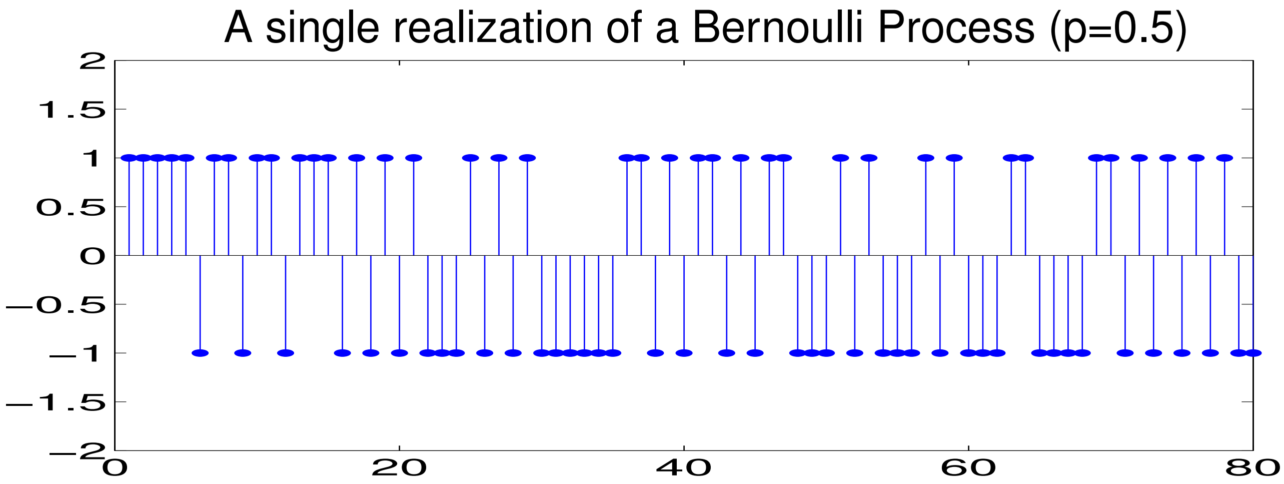
with $A \in \{i\}$ for $i = 1, 2, \cdots, 6$ and $Pr(A = i) = \frac{1}{6}$. Therefore, $x(n)$ is an ensemble (“outcomes”) of 6 different and equally probable discrete-time signals (6 realizations).

Examples of Random Processes

Example 2: Let $x(n)$ be a binary Random Process (or Bernoulli process):

$$x(n) = \sum_{k=-\infty}^{\infty} a_k \Delta(n-k) \quad a_k \text{ a Bernoulli RV } (p = \frac{1}{2})$$

with a_k and a_j uncorrelated $\forall k, j$.



Definition Random Process (part 2)

Definition (part 2): A Random Process (RP) $x(n)$ is an indexed sequence of random variables,

$$\{x(n)\}_{n=-\infty}^{\infty}$$

with each $x(n_i)$ a **random variable** (see Lecture 2). This RV is completely characterized by the joint PDF (pdf),

$$F_{x(n_1),\cdots,x(n_k)}(\alpha_1,\cdots,\alpha_k) = Pr(x(n_1) \leq \alpha_1, \cdots x(n_k) \leq \alpha_k)$$

for any collection of $n_1, \cdots, n_k$.

Characterizatioin of RPs via Ensemble Averages

Since $x(n)$ is an RV, the following ensemble averages may be defined.

Mean: $m_x(n) = E[x(n)]$

Variance: $\sigma_x^2(n) = E[|x(n) - m_x(n)|^2]$

Autocovariance: $c_x(k,\ell) = E[(x(k) - m_x(k))(x(\ell) - m_x(\ell))^*]$

Autocorrelation: $r_x(k,\ell) = E[x(k)x^*(\ell)]$

Standing Assumption: The RVs considered in the rest of the course have mean zero.

Examples of Single Random Processes

Example 3: Let the RP be defined as:

$$x(n) = Asin(n\omega_0 + \phi) = \frac{A\left(e^{j(n\omega_0+\phi)} - e^{-j(n\omega_0+\phi)}\right)}{2j} \quad A, \omega_0 \text{ fixed}$$

and ϕ an RV with pdf, $f_\phi(\alpha) = \begin{cases} \frac{1}{2\pi} & -\pi \leq \alpha < \pi \\ 0 & \text{otherwise} \end{cases}$

Mean:

Autocorrelation

Examples of Single Random Processes

Example 3: Let the RP be defined as:

$$x(n) = Asin(n\omega_0 + \phi) = \frac{A\left(e^{j(n\omega_0+\phi)} - e^{-j(n\omega_0+\phi)}\right)}{2j} \quad A, \omega_0 \text{ fixed}$$

and ϕ an RV with pdf, $f_\phi(\alpha) = \begin{cases} \frac{1}{2\pi} & -\pi \leq \alpha < \pi \\ 0 & \text{otherwise} \end{cases}$

Mean: $m_x(n) = \int_{-\infty}^{\infty} Asin(n\omega_0 + \alpha) f_\phi(\alpha) d\alpha = \frac{A}{2\pi} \int_{-\pi}^{\pi} sin(n\omega_0 + \alpha) d\alpha = ($

Autocorrelation

Examples of Single Random Processes

Example 3: Let the RP be defined as:

$$x(n) = Asin(n\omega_0 + \phi) = \frac{A\Big(e^{j(n\omega_0+\phi)} - e^{-j(n\omega_0+\phi)}\Big)}{2j} \qquad A, \omega_0 \ \text{fixed}$$

and ϕ an RV with pdf, $f_\phi(\alpha) = \begin{cases} \frac{1}{2\pi} & -\pi \leq \alpha < \pi \\ 0 & \text{otherwise} \end{cases}$

Mean: $m_x(n) = \int_{-\infty}^{\infty} Asin(n\omega_0 + \alpha) f_\phi(\alpha) d\alpha = \frac{A}{2\pi} \int_{-\pi}^{\pi} sin(n\omega_0 + \alpha) d\alpha =$

Autocorrelation

$$\begin{aligned} r_x(k,\ell) &= E[x(k)x^*(\ell)] = E[A^2 \frac{\Big(e^{j(k\omega_0+\phi)} - e^{-j(k\omega_0+\phi)}\Big)}{2j} \frac{\Big(e^{-j(\ell\omega_0+\phi)} - e^{j(\ell\omega_0+\phi)}\Big)}{-2j}] \\ &= \frac{A^2}{2} E[cos(k-\ell)\omega_0] - \frac{A^2}{2} E[cos\Big((k+\ell)\omega_0 + 2\phi\Big)] \\ &= \frac{A^2}{2} E[cos(k-\ell)\omega_0] \end{aligned}$$

Examples of Single Random Processes

Example 3: Let the RP be defined as:

$$x(n) = Asin(n\omega_0 + \phi) = \frac{A\Big(e^{j(n\omega_0+\phi)} - e^{-j(n\omega_0+\phi)}\Big)}{2j} \quad A, \omega_0 \text{ fixed}$$

and ϕ an RV with pdf, $f_\phi(\alpha) = \begin{cases} \frac{1}{2\pi} & -\pi \leq \alpha < \pi \\ 0 & \text{otherwise} \end{cases}$

Mean: $m_x(n) = \int_{-\infty}^{\infty} Asin(n\omega_0 + \alpha) f_\phi(\alpha) d\alpha = \frac{A}{2\pi} \int_{-\pi}^{\pi} sin(n\omega_0 + \alpha) d\alpha = 0$

Autocorrelation

$$\begin{aligned} r_x(k,\ell) &= E[x(k)x^*(\ell)] = E[A^2 \frac{\Big(e^{j(k\omega_0+\phi)} - e^{-j(k\omega_0+\phi)}\Big)}{2j} \frac{\Big(e^{-j(\ell\omega_0+\phi)} - e^{j(\ell\omega_0+\phi)}\Big)}{-2j}] \\ &= \frac{A^2}{2} E[cos(k-\ell)\omega_0] - \frac{A^2}{2} E[cos\Big((k+\ell)\omega_0 + 2\phi\Big)] \\ &= \frac{A^2}{2} cos(k-\ell)\omega_0 \end{aligned}$$

Part I:

1. What is a Random Process (RP)
2. **Characterization of multiple (2) RPs**
3. Gaussian Processes
4. Stationairy RPs
5. Property of Wide Sense Stationary (WSS) RPs
6. Autocorrelation matrix of a WSS RP.

Two Random Processes

Let $x(n)$ and $y(n)$ be RPs with means resp. $m_x(n), m_y(n)$, then the cross-covariance is defined as,

$$c_{xy}(k,\ell) = E[(x(k) - m_x(k))(y(\ell) - m_y(\ell))^*]$$

and the cross-correlation as,

$$r_{xy}(k,\ell) = E[x(k)y^*(\ell)]$$

Corrolary: $c_{xy}(k,\ell) = r_{xy}(k,\ell) - m_x(k)m_y^*(\ell)$.

Uncorrelated RPs $x(n)$ and $y(n)$: $c_{xy}(k,\ell) = 0 \ \ \forall k,\ell$

Orthogonal RPs $x(n)$ and $y(n)$: $r_{xy}(k,\ell) = 0 \ \ \forall k,\ell$

Examples of 2 “correlated” Random Processes

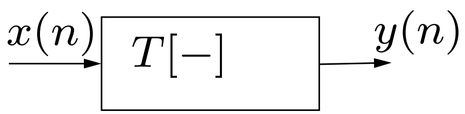
Example 4: Let $x(n)$ be an RP and $T[-]$ be an LSI filter characterized by its impulse response $\{h(n)\}_{n=-\infty}^{\infty}$,

$$x(n) \longrightarrow \boxed{T[-]} \longrightarrow y(n) \qquad\qquad y(n) = \sum_{m=-\infty}^{\infty} h(m)x(n-m)$$

Then, $r_{xy}(k, \ell)$ is given as:

Examples of 2 “correlated” Random Processes

Example 4: Let $x(n)$ be an RP and $T[-]$ be an LSI filter characterized by its impulse response $\{h(n)\}_{n=-\infty}^{\infty}$,

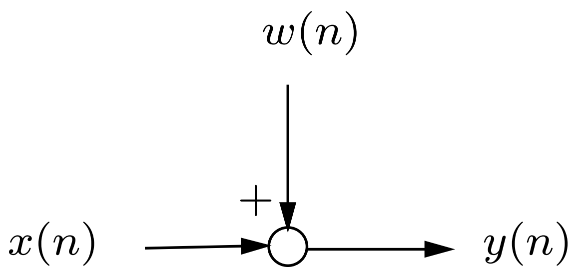


$$y(n) = \sum_{m=-\infty}^{\infty} h(m)x(n-m)$$

Then, $r_{xy}(k,\ell)$ is given as:

$$\begin{aligned}
r_{xy}(k,\ell) &= E[x(k)y^*(\ell)] = E\Big[x(k)\sum_{m=-\infty}^{\infty} h^*(m)x^*(\ell-m)\Big] \\
&= \sum_{m=-\infty}^{\infty} h^*(m)E\Big[x(k)x^*(\ell-m)\Big] \\
&= \sum_{m=-\infty}^{\infty} h^*(m)r_x(k,\ell-m)
\end{aligned}$$

Example 5:

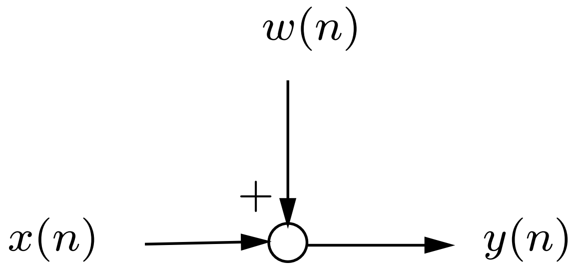


When the signal $x(n)$ and the noise $w(n)$ are *zero mean* and *uncorrelated*:

$$r_y(k,\ell) = r_x(k,\ell) + r_w(k,\ell)$$

And of 2 “uncorrelated” Random Processes

Example 5:



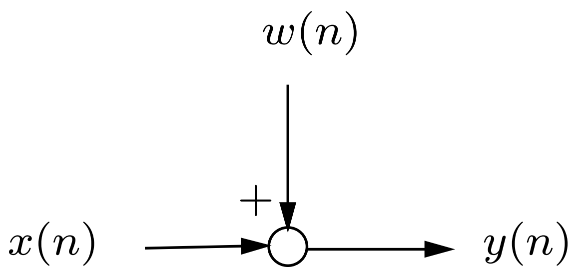
When the signal $x(n)$ and the noise $w(n)$ are *zero mean* and *uncorrelated*:

$$r_y(k,\ell) = r_x(k,\ell) + r_w(k,\ell)$$

$$\begin{aligned} r_y(k,\ell) &= E[y(k)y^*(\ell)] \\ &= E[\Big(x(k)+w(k)\Big)\Big(x^*(\ell)+w^*(\ell)\Big)] \end{aligned}$$

And of 2 “uncorrelated” Random Processes

Example 5:



When the signal $x(n)$ and the noise $w(n)$ are *zero mean* and *uncorrelated*:

$$r_y(k,\ell) = r_x(k,\ell) + r_w(k,\ell)$$

$$\begin{aligned} r_y(k,\ell) &= E[y(k)y^*(\ell)] \\ &= E\Big[\Big(x(k) + w(k)\Big)\Big(x^*(\ell) + w^*(\ell)\Big)\Big] \end{aligned}$$

$$\begin{aligned} r_y(k,\ell) &= E[x(k)x^*(\ell)] + E[x(k)w^*(\ell)] + E[w(k)x^*(\ell)] + E[w(k)w^*(\ell)] \\ &= r_x(k,\ell) + r_w(k,\ell) \end{aligned}$$

Part I:

1. What is a Random Process (RP)
2. Characterization of multiple (2) RPs
3. **Gaussian Processes**
4. Stationairy RPs
5. Property of Wide Sense Stationary (WSS) RPs
6. Autocorrelation matrix of a WSS RP.

Gaussian Processes (illustration for $k = 3$)

Definition: An RP $x(n) \in \mathbb{R}$ is a Gaussian Process, if every finite collection of real samples $x(n_i)$ for $i = 1 : k$ are jointly Gaussian with pdf:

$$f_x(\mathbf{x}) = \frac{1}{(2\pi)^{k/2}|\det(C_x)|^{1/2}} exp\Big[-\frac{1}{2}(\mathbf{x}-\mathbf{m_x})^T C_x^{-1}(\mathbf{x}-\mathbf{m_x})\Big]$$

with $\mathbf{x} = \begin{bmatrix} x(n_1) & \cdots & x(n_k) \end{bmatrix}^T$ and $\mathbf{m_x} = \begin{bmatrix} m_x(n_1) & \cdots & m_x(n_k) \end{bmatrix}^T$ and C_x is a symmetric positive definite matrix with entries (for $k = 3$ and $m_x(n_i) \equiv 0 \ \forall i$):

$$C_x = \begin{bmatrix} E[x(n_1)^2] & E[x(n_1)x(n_2)] & E[x(n_1)x(n_3)] \\ E[x(n_2)x(n_1)] & E[x(n_2)^2] & E[x(n_2)x(n_3)] \\ E[x(n_3)x(n_1)] & E[x(n_3)x(n_2)] & E[x(n_3)^2] \end{bmatrix}$$

Part I:

1. What is a Random Process (RP)
2. Characterization of multiple (2) RPs
3. Gaussian Processes
4. **Stationairy RPs**
5. Property of Wide Sense Stationary (WSS) RPs
6. Autocorrelation matrix of a WSS RP.

Stationarity

Notions of “time-invariance” imposed on the statistical quantities that are used to characterize the RP (PDF, pdf or ensemble averages).

First-order Stationarity

Definition: A RP $x(n)$ is first-order stationairy iff,

$$f_{x(n)}(\alpha) = f_{x(n+k)}(\alpha) \quad \forall k$$

Corollary:

The mean of a first-order stationary RP $x(n)$ is constant, i.e., $m_x(n) = m_x$.

First-order Stationarity

Definition: A RP $x(n)$ is first-order stationairy iff,

$$f_{x(n)}(\alpha) = f_{x(n+k)}(\alpha) \quad \forall k$$

Corollary:

$$\boxed{\text{The mean of a first-order stationary RP } x(n) \text{ is constant, i.e., } m_x(n) = m_x.}$$

$$\begin{aligned} m_x(n+k) &= \int_{-\infty}^{\infty} \alpha f_{x(n+k)}(\alpha) d\alpha \\ &= \int_{-\infty}^{\infty} \alpha f_{x(n)}(\alpha) d\alpha \\ &= m_x(n) \end{aligned}$$

Second-order Stationarity

Definition: An RP $x(n)$ is second-order stationairy iff,

$$f_{x(n_1),x(n_2)}(\alpha_1, \alpha_2) = f_{x(n_1+k),x(n_2+k)}(\alpha_1, \alpha_2) \quad \forall k$$

Second-order Stationarity

Definition: An RP $x(n)$ is second-order stationairy iff,

$$f_{x(n_1),x(n_2)}(\alpha_1, \alpha_2) = f_{x(n_1+k),x(n_2+k)}(\alpha_1, \alpha_2) \quad \forall k$$

$$\begin{aligned} r_x(n+k, m+k) &= \int_{-\infty}^{\infty} \alpha_1 \alpha_2 f_{x(n+k),x(m+k)}(\alpha_1, \alpha_2) d\alpha_1 d\alpha_2 \\ &= \int_{-\infty}^{\infty} \alpha_1 \alpha_2 f_{x(n),x(m)}(\alpha_1, \alpha_2) d\alpha_1 d\alpha_2 \\ &= r_x(n, m) \end{aligned}$$

Second-order Stationarity

Definition: An RP $x(n)$ is second-order stationairy iff,

$$f_{x(n_1),x(n_2)}(\alpha_1,\alpha_2) = f_{x(n_1+k),x(n_2+k)}(\alpha_1,\alpha_2) \quad \forall k$$

$$\begin{aligned} r_x(n+k,m+k) &= \int_{-\infty}^{\infty} \alpha_1\alpha_2 f_{x(n+k),x(m+k)}(\alpha_1,\alpha_2)d\alpha_1 d\alpha_2 \\ &= \int_{-\infty}^{\infty} \alpha_1\alpha_2 f_{x(n),x(m)}(\alpha_1,\alpha_2)d\alpha_1 d\alpha_2 \\ &= r_x(n,m) \end{aligned}$$

Corollary:

The auto correlation function of a second-order stationary RP $x(n)$ satisfies,

$$r_x(n,m) = r_x(n-m,0) = r_x(n-m)$$

Part I:

1. What is a Random Process (RP)
2. Characterization of multiple (2) RPs
3. Gaussian Processes
4. Stationairy RPs
5. **Property of Wide Sense Stationary (WSS) RPs**
6. Autocorrelation matrix of a WSS RP.

Wide-sense Stationarity

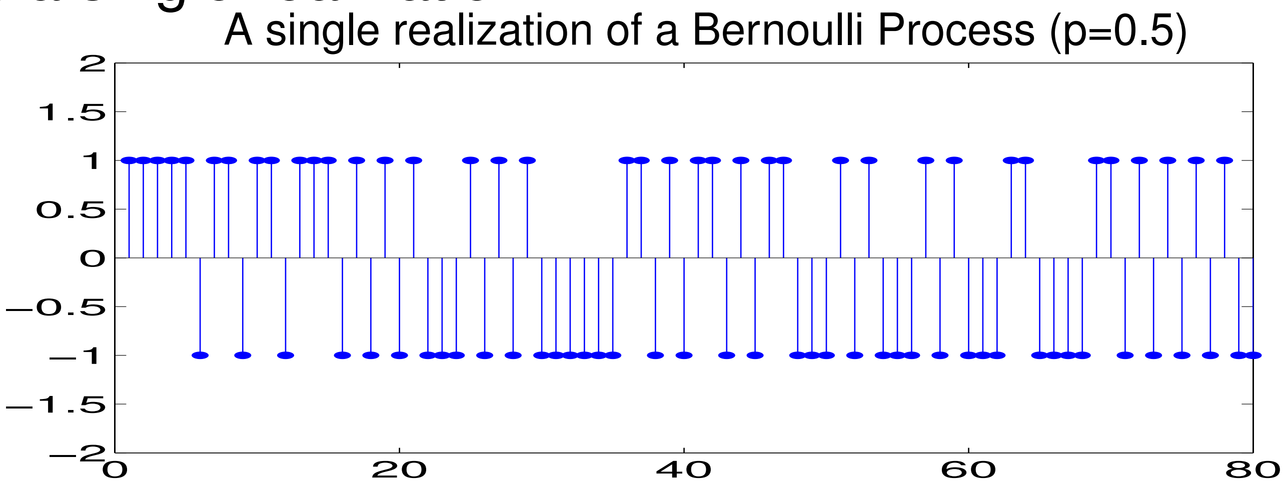
Definition [WSS]: A discrete-time random process $x(n)$ is WSS if

1. $m_x(k) = m_x < \infty$
2. $r_x(k,\ell) = r_x(k-\ell)\ \forall k,\ell$
3. $c_x(0) < \infty$ (the variance is finite).

WSS is a weaker notion of stationarity: For example, $x(n)$ is first order stationary $\Rightarrow m_x(n) = m_x$ (but the reverse is not necessarily true)!

Examples of Random Processes

Example 6: A Bernoulli process $x(n)$ is WSS. The following figure displays a single realization.



Then,

1. $E[x(k)] = 1\frac{1}{2} - 1\frac{1}{2} = 0 \Rightarrow$ mean is constant.
2. $E[x(k)x^*(\ell)] = 0$ for $k \neq \ell$ as the samples are drawn independently.
3. $E[x(k)^2] = 1\frac{1}{2} + 1\frac{1}{2} = 1 < \infty$.

Properties of a (zero-mean) WSS RP

Let $x(n)$ be a WSS RP and recall $r_x(k) = E[x(n)x^*(n-k)]$ then,

Property 1: If $x(n) \in \mathbb{C}$ then $r_x(k) = r_x^*(-k)$ (conjugate symmetric).
If $x(n) \in \mathbb{R}$ then $r_x(k) = r_x(-k)$ (symmetric).

Property 2: $r_x(0) = E[|x(n)|^2] \geq 0$.

Property 3: $r_x(0) \geq |r_x(k)| \quad \forall k$

Property 4: If $\exists k_0 : r_x(k_0) = r_x(0) \Rightarrow r_x(k)$ is periodic with period k_0 and further

$$E[|x(n) - x(n-k_0)|^2] = 0$$

$x(n)$ is said to be *mean-square periodic*.

Nice exercises

WSS joint RPs

For 2 RPs $x(n)$ and $y(n)$ that have mean zero, the crosscorrelation function is defined as:

$$r_{xy}(k,\ell) = E[x(k)y^*(\ell)]$$

Definition: [WSS 2 random processes] Two random processes $\{x(n)\}$ and $\{y(n)\}$ are jointly WSS if

1. $\{x(n)\}$ WSS
2. $\{y(n)\}$ WSS
3. $r_{xy}(k,\ell) = r_{xy}(k-\ell,0) := r_{xy}(k-\ell)$

Properties of two WSS RP

Consider two WSS RPs $x(n)$ and $y(n)$, then their *cross-correlation function* $r_{xy}(k) = E[x(n)y(n-k)^*]$ satisfies,

Property 1: If $x(n), y(n) \in \mathbb{C}$ then $r_{yx}(k) = r_{xy}^*(-k)$ (change of order index arguments!)

If $x(n), y(n) \in \mathbb{R}$ then $r_{yx}(k) = r_{xy}(-k)$(not symmetric)

Property 2: $|r_{xy}(k)| \leq \sqrt{r_x(0) r_y(0)}$

Property 3: $|Re\Big(r_{xy}(k)\Big)| \leq \frac{1}{2}[r_x(0) + r_y(0)]$

Summary Definition RPs

A (jointly) WSS RP $x(n) \in \mathbb{C}$ (and a RP $y(n) \in \mathbb{C}$) is (are) fully characterized by

- by its mean $E[x(n)](= 0)$ and same for $E[y(n)]$.

Summary Definition RPs

A (jointly) WSS RP $x(n) \in \mathbb{C}$ (and a RP $y(n) \in \mathbb{C}$) is (are) fully characterized by

- by its mean $E[x(n)](= 0)$ and same for $E[y(n)]$.
- and its auto-correlation functions $r_x(k), r_y(k)$ and cross-correlation function $r_{xy}(k) = E[x(n)y(n-k)^*]$.

Exercise

Consider the two perturbed sinusoids:

$$\begin{aligned} x(n) &= \sin(\omega n) + v_1(n) \\ y(n) &= \sin(\omega n + \varphi) + v_2(n) \end{aligned}$$

with $v_1(n), v_2(n)$ zero-mean and uncorrelated, ω given and φ unknown (but assumed to be constant), then use the cross-correlation function $r_{yx}(n, n-k)$ to retrieve the unknown phase φ. Show first whether or not $x(n)$ and $y(n)$ are jointly WSS?

Part I:

1. What is a Random Process (RP)
2. Characterization of multiple (2) RPs
3. Gaussian Processes
4. Stationairy RPs
5. Property of Wide Sense Stationary (WSS) RPs
6. **Autocorrelation matrix of a WSS RP.**

The Autocorrelation Matrix of a WSS RP

Let $x(n)$ be a WSS RP and recall $r_x(k) = E[x(n)x^*(n-k)]$.

An important second-order statistical characterization of a RP $x(n)$ is the covariance matrix R_x. For the 3×3 case:

$$
\begin{aligned}
E[\mathbf{x}\mathbf{x}^H] &= E[\begin{bmatrix} x(n) \\ x(n+1) \\ x(n+2) \end{bmatrix} \begin{bmatrix} x^*(n) & x^*(n+1) & x^*(n+2) \end{bmatrix}] \\
&= \begin{bmatrix} r_x(0) & r_x^*(1) & r_x^*(2) \\ r_x(1) & r_x(0) & r_x^*(1) \\ r_x(2) & r_x(1) & r_x(0) \end{bmatrix} \\
&= R_x
\end{aligned}
$$

Exercise: The Autocorrelation Matrix of a WSS RP

Let $x(n)$ be a WSS RP and recall $r_x(k) = E[x(n)x^*(n-k)]$.

Then show that,

$$
\begin{aligned}
E[\begin{bmatrix} x^*(n) \\ x^*(n-1) \\ x^*(n-2) \end{bmatrix} \begin{bmatrix} x(n) & x(n-1) & x(n-2) \end{bmatrix}] &= \begin{bmatrix} r_x(0) & r_x^*(1) & r_x^*(2) \\ r_x(1) & r_x(0) & r_x^*(1) \\ r_x(2) & r_x(1) & r_x(0) \end{bmatrix} \\
&= R_x
\end{aligned}
$$

Properties Autocorrelation matrix of a WSS RP

Property 1: R_x is Hermitian Toeplitz.

However the reverse is not true!

Property 2: $R_x > 0$ and therefore $\lambda_k(R_x)$ are real and non-negative.

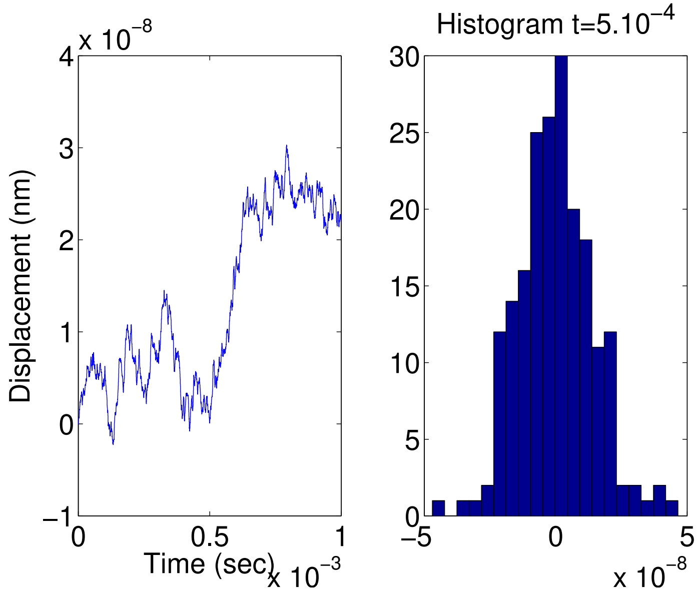
Part II:

1. **Ergodicity in the mean**
2. White noise RP
3. Representing RPs in the Frequency Domain: Signal Spectra
   - Power Spectra
   - Cross Spectra

Obtaining Characteristics about RPs from Experiments?

Statistical Characterizations of an RP (like pdf, mean, Autocorrelation, etc.) are ensemble averages.

Example 1: A single realization of a free Brownian particle and the Histogram at time $500\mu s$.



In practice, we usually have only access to a single realization.

Question: What are the conditions on the statistical characteristic(s) of the RPs to replace the ensemble average by a time-average?

`Brownian_bup, ergodic_min.m`

Ergodicity in the mean

Question: Let $x(n)$ be a RP, when does the following time-average (sample mean),

$$\hat{m}_x(N) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)$$

become equal to the mean (ensemble average) m_x.

Ergodicity in the mean

Question: Let $x(n)$ be a RP, when does the following time-average (sample mean),

$$\hat{m}_x(N) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)$$

become equal to the mean (ensemble average) m_x.

Definition Ergodic in the mean When the sample mean of a WSS RP $x(n)$ converges in the mean-square sense, i.e.

$$\lim_{N\to\infty} E[|\hat{m}_x(N) - m_x|^2] = 0 \quad \text{or} \quad \Pr\left[\lim_{N\to\infty} \hat{m}_x(N) = m_x\right] = 1$$

then the RP is ergodic in the mean

Theorem: Ergodicity in the mean

A WSS RP $x(n)$ is ergodic in the mean iff,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{\ell=-N+1}^{N-1} (1 - \frac{|\ell|}{N}) c_x(\ell) = 0$$

Sufficiency conditions

1. $c_x(0) < \infty$
2. $\lim_{\ell\to\infty} c_x(\ell) = 0$

Proof Theorem Ergodicity in the mean

$$
\begin{aligned}
\mathrm{Var}(\hat{m}_x(N)) &= E[|\hat{m}_x(N) - m_x|^2] \\
&= E\left[\left|\frac{1}{N}\sum_{n=0}^{N-1}(x(n) - m_x)\right|^2\right]
\end{aligned}
$$

$$=$$

$$=$$

$$=$$

$$=$$

$$=$$

Proof Theorem Ergodicity in the mean

$$
\begin{aligned}
\mathrm{Var}(\hat{m}_x(N)) &= E[|\hat{m}_x(N) - m_x|^2] \\
&= E\left[\left| \tfrac{1}{N} \textstyle\sum_{n=0}^{N-1} (x(n) - m_x) \right|^2\right]
\end{aligned}
$$

$$
= \frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} E[(x(m) - m_x)(x(n) - m_x)^*]
$$

$$=$$

$$=$$

$$=$$

$$=$$

Proof Theorem Ergodicity in the mean

$$
\begin{aligned}
\mathrm{Var}(\hat{m}_x(N)) &= E[|\hat{m}_x(N) - m_x|^2] \\
&= E\Big[\Big| \tfrac{1}{N} \textstyle\sum_{n=0}^{N-1} (x(n) - m_x)\Big|^2\Big]
\end{aligned}
$$

$$
= \frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} E[(x(m) - m_x)(x(n) - m_x)^*]
$$

$$
= \frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} c_x(m-n)
$$

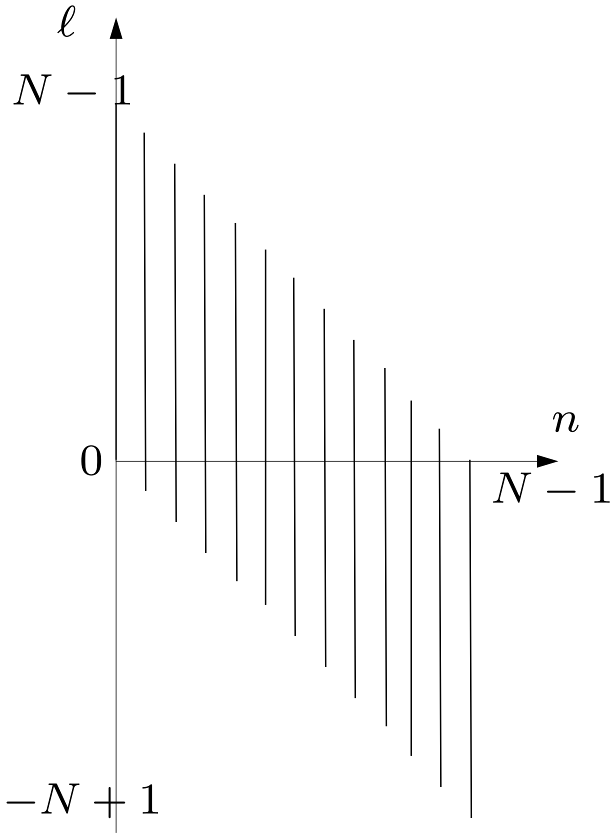
$=$

$=$

$=$

Proof Theorem Ergodicity in the mean

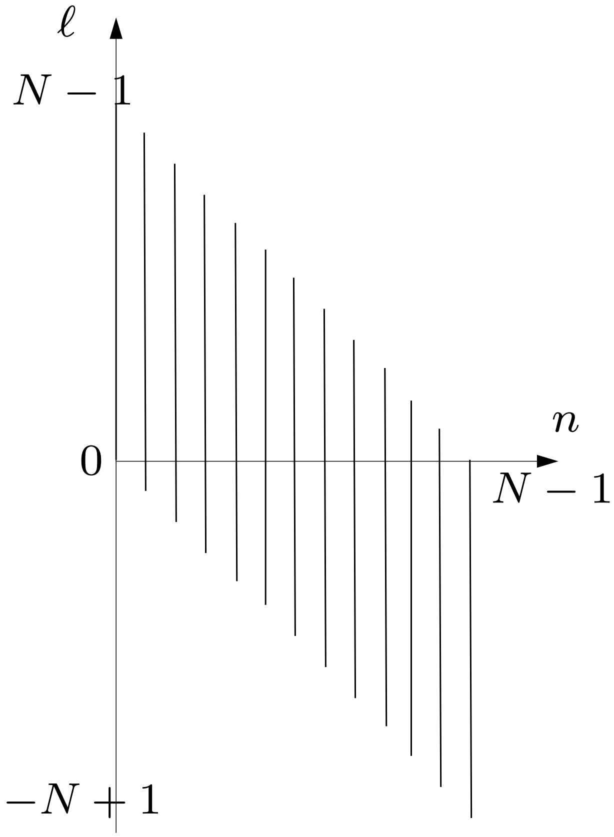
$$
\begin{aligned}
\mathrm{Var}(\hat{m}_x(N)) &= E[|\hat{m}_x(N) - m_x|^2] \\
&= E\Big[\Big|\tfrac{1}{N}\textstyle\sum_{n=0}^{N-1}(x(n) - m_x)\Big|^2\Big] \\
&= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} E[(x(m) - m_x)(x(n) - m_x)^*] \\
&= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} c_x(m-n) \\
&= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{\ell=-n}^{N-1-n} c_x(\ell) \\
&= \\
&=
\end{aligned}
$$



TU Delft

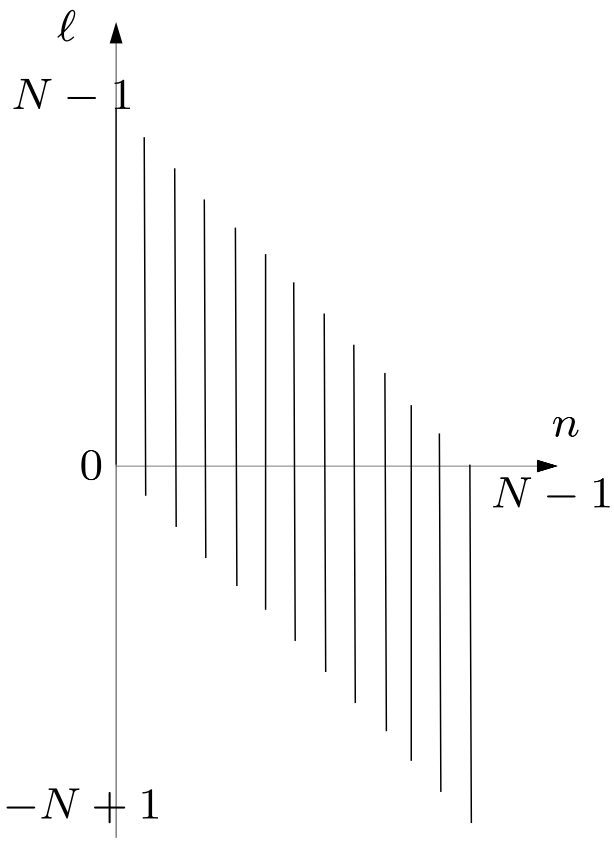
Proof Theorem Ergodicity in the mean

$$
\begin{aligned}
\mathrm{Var}(\hat{m}_x(N)) &= E[|\hat{m}_x(N) - m_x|^2] \\
&= E\Big[\Big|\tfrac{1}{N}\textstyle\sum_{n=0}^{N-1}(x(n) - m_x)\Big|^2\Big] \\
&= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} E[(x(m) - m_x)(x(n) - m_x)^*] \\
&= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} c_x(m-n) \\
&= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{\ell=-n}^{N-1-n} c_x(\ell) \\
&= \frac{1}{N^2}\sum_{\ell=-N+1}^{N-1} (N - |\ell|)c_x(\ell) \\
&=
\end{aligned}
$$



Proof Theorem Ergodicity in the mean

$$
\begin{aligned}
\mathrm{Var}(\hat{m}_x(N)) &= E[|\hat{m}_x(N) - m_x|^2] \\
&= E\left[\left|\tfrac{1}{N}\textstyle\sum_{n=0}^{N-1}(x(n) - m_x)\right|^2\right] \\
&= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} E[(x(m) - m_x)(x(n) - m_x)^*] \\
&= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} c_x(m-n) \\
&= \frac{1}{N^2}\sum_{n=0}^{N-1}\sum_{\ell=-n}^{N-1-n} c_x(\ell) \\
&= \frac{1}{N^2}\sum_{\ell=-N+1}^{N-1} (N - |\ell|)c_x(\ell) \\
&= \frac{1}{N}\sum_{\ell=-N+1}^{N-1} \left(1 - \frac{|\ell|}{N}\right)c_x(\ell)
\end{aligned}
$$



Ergodicity in other Ensemble Averages

Let $\{x_i(n)\}_{n=-k}^{N-1}$ be a single realization of the RP $x(n)$, then the time-average estimate (sample mean) of Autocorrelation function is:

$$\hat{r}_x(k,N) = \frac{1}{N}\sum_{n=0}^{N-1} x_i(n)x_i^*(n-k) \quad \Big(r_x(k) = E[x(n)x^*(n-k)]\Big)$$

Ergodicity in other Ensemble Averages

Let $\{x_i(n)\}_{n=-k}^{N-1}$ be a single realization of the RP $x(n)$, then the time-average estimate (sample mean) of Autocorrelation function is:

$$\hat{r}_x(k,N) = \frac{1}{N}\sum_{n=0}^{N-1} x_i(n)x_i^*(n-k) \quad \Big(r_x(k) = E[x(n)x^*(n-k)]\Big)$$

Defintion of Autocorrelation Ergodic: A WSS RP is Autocorrelation Ergodic if the Autocorrelation sample average $\hat{r}_x(k,N) = \frac{1}{N}\sum_{n=0}^{N-1} x_i(n)x_i^*(n-k)$ converges in the mean-square sense, given as,

$$\lim_{N\to\infty} E[|\hat{r}_x(k,N) - r_x(k)|^2] = 0 \quad \text{or} \quad \Pr\left[\lim_{N\to\infty} \hat{r}_x(k,N) = r_x(k)\right] = 1$$

then we can write, $\lim_{N\to\infty} \hat{r}_x(k,N) = r_x(k)$

Part II:

1. Ergodicity in the mean
2. **White noise RP**
3. Representing RPs in the Frequency Domain: Signal Spectra
   - Power Spectra
   - Cross Spectra

Zero-mean White Noise (ZMWN) - `demo_zmwn.m`

The standardized source in generating RPs (by filtering RPs next lecture) is ZMWN.

Definition: A WSS RP $v(n)$ with $v(n) \in \mathbb{R}$ or $\in \mathbb{C}$ is ZMWN if it has mean equal to zero and its autocovariance function equals,

$$c_v(k) = \sigma_v^2 \Delta(k)$$

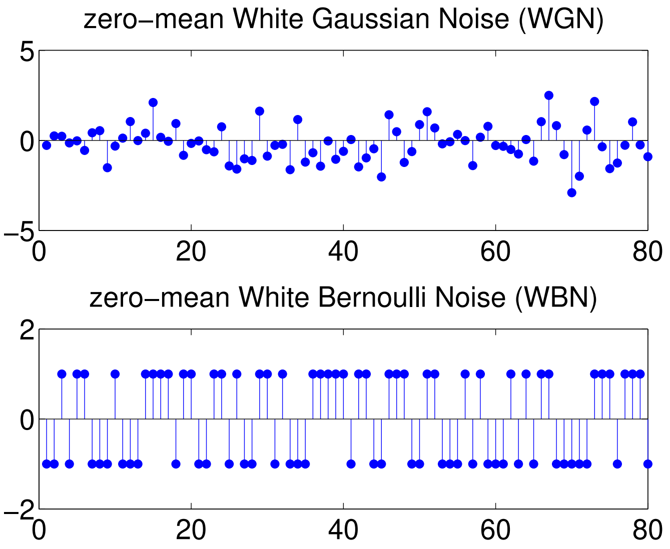
Zero-mean White Noise (ZMWN) - `demo_zmwn.m`

The standardized source in generating RPs (by filtering RPs next lecture) is ZMWN.

Definition: A WSS RP $v(n)$ with $v(n) \in \mathbb{R}$ or $\in \mathbb{C}$ is ZMWN if it has mean equal to zero and its autocovariance function equals,

$$c_v(k) = \sigma_v^2 \Delta(k)$$

Example 2:



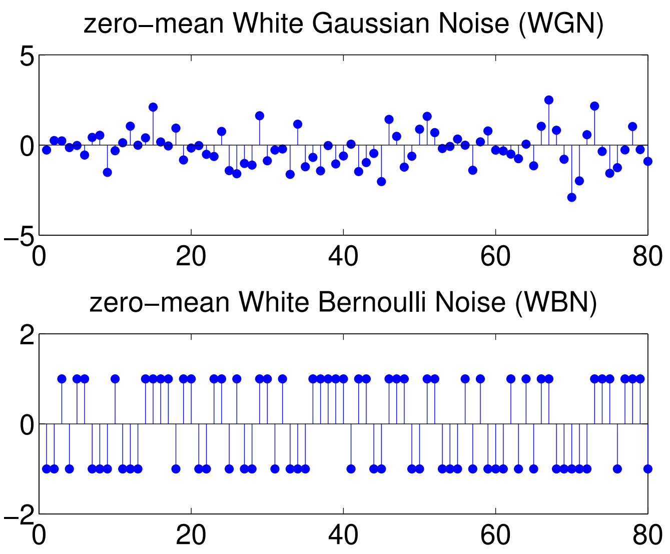
Zero-mean White Noise (ZMWN) - `demo_zmwn.m`

The standardized source in generating RPs (by filtering RPs next lecture) is ZMWN.

Definition: A WSS RP $v(n)$ with $v(n) \in \mathbb{R}$ or $\in \mathbb{C}$ is ZMWN if it has mean equal to zero and its autocovariance function equals,

$$c_v(k) = \sigma_v^2 \Delta(k)$$

Example 2:



Remark: If $v(n) \in \mathbb{C}$ and $v(n)$ is ZMWN, then

$$v(n) = v_r(n) + jv_i(n) \Rightarrow \mathrm{Var}(v(n)) = \mathrm{Var}(v_r(n)) + \mathrm{Var}(v_i(n))$$

since $v_r(n)$ and $v_j(n)$ are uncorrelated (and have mean zero).

Part II:

1. Ergodicity in the mean
2. White noise RP
3. **Representing RPs in the Frequency Domain: Signal Spectra**
   - Power Spectra
   - Cross Spectra

Review Fourier transform

For absolute integrable function (series) $\int_{-\infty}^{\infty} |x(t)|dt < \infty$ $(\sum_{n=-\infty}^{\infty} |x(n)| < \infty)$:

Transform

$$X_c(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$$

$$X_d(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(nT_s)e^{-j\omega nT_s}$$

Inverse

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X_c(\omega)e^{j\omega t}d\omega$$

$$x(nT_s) = \frac{T_s}{2\pi}\int_{-\frac{\pi}{T_s}}^{\frac{\pi}{T_s}} X_d(e^{j\omega})e^{j\omega kT_s}d\omega$$

Special Case $T_s = 1$(sampling time):

$$X_d(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n}$$

$$x(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X_d(e^{j\omega})e^{j\omega n}d\omega$$

Properties DTFT

Property 1: If $x(n) \in \mathbb{R}$ then $X(e^{j\omega}) = X^*(e^{-j\omega})$ (i.e. $X(e^{j\omega})$ is conjugate symmetric.)

Property 2: $X(e^{j\omega})$ is periodic with a period of 2π.

Property 3: *Parseval's theorem:*

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

Property 4: *convolution:* $y(n) = \sum_{m=-\infty}^{\infty} h(m)x(n-m)$ then,

$$\begin{aligned}
Y(e^{j\omega}) &= \sum_{n=-\infty}^{\infty} y(n)e^{-j\omega n} = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} h(m)x(n-m)e^{-j\omega n} \\
&= \sum_{m=-\infty}^{\infty} h(m) \sum_{n=-\infty}^{\infty} x(n-m)e^{-j\omega n} \overset{Delay}{=} \sum_{m=-\infty}^{\infty} h(m)e^{-j\omega m}X(e^{j\omega}) \\
&= H(e^{j\omega})X(e^{j\omega})
\end{aligned}$$

The Power Spectrum

Definition: The power spectrum (or power spectral density) of a WSS RP $x(n)$ is the DTFT of its Autocorrelation function:

$$P_x(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_x(k) e^{-j\omega k}$$

The Power Spectrum

Definition: The power spectrum (or power spectral density) of a WSS RP $x(n)$ is the DTFT of its Autocorrelation function:

$$P_x(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_x(k) e^{-j\omega k}$$

and the inverse DTFT yields:

$$r_x(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_x(e^{j\omega}) e^{j\omega k} d\omega$$

The Power Spectrum

Definition: The power spectrum (or power spectral density) of a WSS RP $x(n)$ is the DTFT of its Autocorrelation function:

$$P_x(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_x(k) e^{-j\omega k}$$

and the inverse DTFT yields:

$$r_x(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_x(e^{j\omega}) e^{j\omega k} d\omega$$

Using the z-transform we also have:

$$P_x(z) = \sum_{k=-\infty}^{\infty} r_x(k) z^{-k}$$

This will also be referred to as the power spectrum of $x(n)$.

Example 3: Consider the RP $x(n)$ to be generated as:

$$x(n) = ax(n-1) + v(n) \quad a \in]-1, 1[$$

with $v(n)$ ZMWN(σ_v^2). It may be assume that $x(n)$ is WSS (see next Lecture). Then its autocorrelation function is given as:

$$r_x(k) = \frac{\sigma_v^2}{1-a^2} a^{|k|}$$

Therefore its Power spectrum is,

$$P_x(e^{j\omega}) = \frac{\sigma_v^2}{1 + a^2 - 2acos(\omega)}$$

`Powspec.m`

Properties of the Power Spectrum $P_x(e^{j\omega})$

Property 1: *real and (conjugate) Symmetry*.

$$P_x(e^{j\omega}) = P_x^*(e^{j\omega}) \quad \text{i.e.} \quad P_x(e^{j\omega}) \in \mathbb{R} \qquad P_x(z) = P_x^*(1/z^*)$$

If $x(n) \in \mathbb{R}$ then $P_x(e^{j\omega}) = P_x(e^{-j\omega})$ is an even function in ω

Property 2: *Positivity*.

$$P_x(e^{j\omega}) \geq 0 \quad \text{Proof later.}$$

Property 3: *Total Power*.

$$E[|x(n)|^2] = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_x(e^{j\omega}) d\omega$$

Property 4: If $x(n)$ contains periodic components, then $P_x(e^{j\omega})$ contains unit sample functions. `PeriodBias.m`

A sample average approximation of the The Power Spectrum

Question: What would be a good approximation of the Power Spectrum when having only a finite data set?

Let us inspect the definition more closely.

$$
\begin{aligned}
P_x(e^{j\omega}) &= \sum_{k=-\infty}^{\infty} r_x(k)e^{-j\omega k} = \sum_{k=-\infty}^{\infty} E[x(n)x^*(n-k)]e^{-j\omega k} \\
= \\
&\approx \\
&\approx \\
&=
\end{aligned}
$$

A sample average approximation of the The Power Spectrum

Question: What would be a good approximation of the Power Spectrum when having only a finite data set?

Let us inspect the definition more closely.

$$
\begin{aligned}
P_x(e^{j\omega}) &= \sum_{k=-\infty}^{\infty} r_x(k)e^{-j\omega k} = \sum_{k=-\infty}^{\infty} E[x(n)x^*(n-k)]e^{-j\omega k} \\
&= E[x(n)\sum_{k=-\infty}^{\infty} x^*(n+k)e^{j\omega k}] = E[x(n)\Big[\sum_{k=-\infty}^{\infty} x(n+k)e^{-j\omega k}\Big]^*] = E[x(n)e^{-j\omega n}X^*(e^{j\omega})] \\
&\approx \\
&\approx \\
&=
\end{aligned}
$$

A sample average approximation of the The Power Spectrum

Question: What would be a good approximation of the Power Spectrum when having only a finite data set?

Let us inspect the definition more closely.

$$
\begin{aligned}
P_x(e^{j\omega}) &= \sum_{k=-\infty}^{\infty} r_x(k)e^{-j\omega k} = \sum_{k=-\infty}^{\infty} E[x(n)x^*(n-k)]e^{-j\omega k} \\
&= E[x(n) \sum_{k=-\infty}^{\infty} x^*(n+k)e^{j\omega k}] = E[x(n)\Big[\sum_{k=-\infty}^{\infty} x(n+k)e^{-j\omega k}\Big]^*] = E[x(n)e^{-j\omega n}X^*(e^{j\omega})] \\
&\approx \frac{1}{2N+1}\sum_{n=-N}^{N} x(n)e^{-j\omega n}X^*(e^{j\omega}) \\
&\approx \\
&=
\end{aligned}
$$

A sample average approximation of the The Power Spectrum

Question: What would be a good approximation of the Power Spectrum when having only a finite data set?
Let us inspect the definition more closely.

$$
\begin{aligned}
P_x(e^{j\omega}) &= \sum_{k=-\infty}^{\infty} r_x(k)e^{-j\omega k} = \sum_{k=-\infty}^{\infty} E[x(n)x^*(n-k)]e^{-j\omega k} \\
&= E[x(n)\sum_{k=-\infty}^{\infty} x^*(n+k)e^{j\omega k}] = E[x(n)\Big[\sum_{k=-\infty}^{\infty} x(n+k)e^{-j\omega k}\Big]^*] = E[x(n)e^{-j\omega n}X^*(e^{j\omega})] \\
&\approx \frac{1}{2N+1}\sum_{n=-N}^{N} x(n)e^{-j\omega n}X^*(e^{j\omega}) \\
&\approx \frac{1}{2N+1}\sum_{n=-N}^{N} x(n)e^{-j\omega n}\sum_{m=-N}^{N} x^*(m)e^{j\omega m} \\
&=
\end{aligned}
$$

A sample average approximation of the The Power Spectrum

Question: What would be a good approximation of the Power Spectrum when having only a finite data set?
Let us inspect the definition more closely.

$$\begin{aligned}
P_x(e^{j\omega}) &= \sum_{k=-\infty}^{\infty} r_x(k)e^{-j\omega k} = \sum_{k=-\infty}^{\infty} E[x(n)x^*(n-k)]e^{-j\omega k} \\
&= E[x(n)\sum_{k=-\infty}^{\infty} x^*(n+k)e^{j\omega k}] = E[x(n)\Big[\sum_{k=-\infty}^{\infty} x(n+k)e^{-j\omega k}\Big]^*] = E[x(n)e^{-j\omega n}X^*(e^{j\omega})] \\
&\approx \frac{1}{2N+1}\sum_{n=-N}^{N} x(n)e^{-j\omega n}X^*(e^{j\omega}) \\
&\approx \frac{1}{2N+1}\sum_{n=-N}^{N} x(n)e^{-j\omega n}\sum_{m=-N}^{N} x^*(m)e^{j\omega m} \\
&= \frac{1}{2N+1}X_N(e^{j\omega})X_N^*(e^{j\omega}) = \frac{1}{2N+1}|X_N(e^{j\omega})|^2
\end{aligned}$$

In what sense is this an approximator?

Theorem: When

$$\sum_{k=-\infty}^{\infty} |k| r_x(k) < \infty$$

then, the approximation $P_N(e^{j\omega}) = \frac{1}{2N+1}|X_N(e^{j\omega})|^2$ is related to the power spectrum $P_x(e^{j\omega})$ in the following asymptotic manner:

$$P_x(e^{j\omega}) = \lim_{N\to\infty} E[P_N(e^{j\omega})]$$

[$P_N(e^{j\omega})$ is called the Periodogram.]

Example 3 (Ct'd): Consider the RP $x(n)$ to be generated as:

$$x(n) = ax(n-1) + v(n) \quad a \in]-1, 1[$$

with $v(n)$ ZMWN(σ_v^2). It may be assumed that $x(n)$ is WSS (see next Lecture). The theoretical Power Spectrum was:

$$P_x(e^{j\omega}) = \frac{\sigma_v^2}{1 + a^2 - 2acos(\omega)}$$

Example Power Spectrum Calculation

Example 3 (Ct'd): Consider the RP $x(n)$ to be generated as:

$$x(n) = ax(n-1) + v(n) \quad a \in]-1, 1[$$

with $v(n)$ ZMWN(σ_v^2). It may be assumed that $x(n)$ is WSS (see next Lecture). The theoretical Power Spectrum was:

$$P_x(e^{j\omega}) = \frac{\sigma_v^2}{1 + a^2 - 2acos(\omega)}$$

Its approximation via $P_N(e^{j\omega})$ is analysed in `PowspecPer.m`.

Example Power Spectrum Calculation

Example 3 (Ct'd): Consider the RP $x(n)$ to be generated as:

$$x(n) = ax(n-1) + v(n) \quad a \in]-1, 1[$$

with $v(n)$ ZMWN(σ_v^2). It may be assumed that $x(n)$ is WSS (see next Lecture). The theoretical Power Spectrum was:

$$P_x(e^{j\omega}) = \frac{\sigma_v^2}{1 + a^2 - 2acos(\omega)}$$

Its approximation via $P_N(e^{j\omega})$ is analysed in `PowspecPer.m`.

How to improve this approximation via averaging is analysed in `PowspecAveraging.m`?

Cross-power spectra

Definition

$$\begin{aligned} P_{xy}(e^{j\omega}) &= \sum_{k=-\infty}^{\infty} r_{xy}(k)e^{-j\omega k} \\ r_{xy}(k) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{xy}(e^{j\omega})e^{j\omega k}d\omega \end{aligned}$$

Then, if $x(n)$ and $y(n)$ are real,

- $P_{xy}(\omega)$ is complex valued ($\in \mathbb{C}$)
- $P_{xy}(\omega) = P_{xy}^{*}(-\omega)$ *Re* part is even; *Im* part is oneven
- $P_{xy}(\omega) = P_{yx}^{*}(\omega)$

When $x(n)$ and $y(n)$ are complex, this extends to,

$$P_{xy}(z) = P_{yx}^{*}(1/z^{*})$$

Next steps forward to improve your chances to succeed ...

Instruction session for explanation of the abstract notions and getting hands-on-exerpience!

Preparation:

Study Chapter 5 (5.1 - 5.3)

Next Instruction/lecture see Course Overview