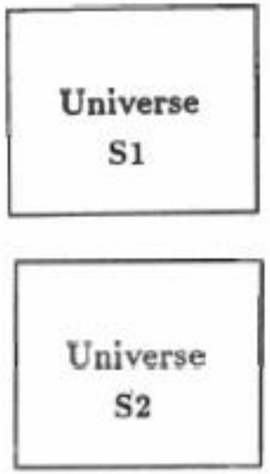
ECE1511
Lecture 2

RANDOM PPOCESSES

* The meaning of *"random"*.



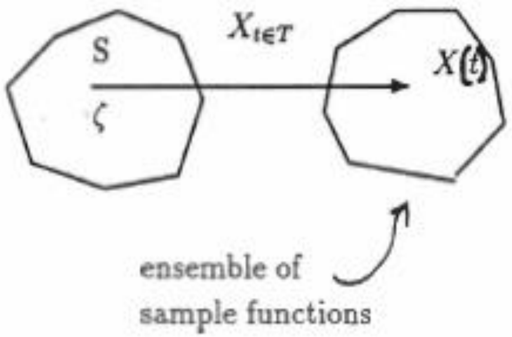
Identical prior to the occurence of an event

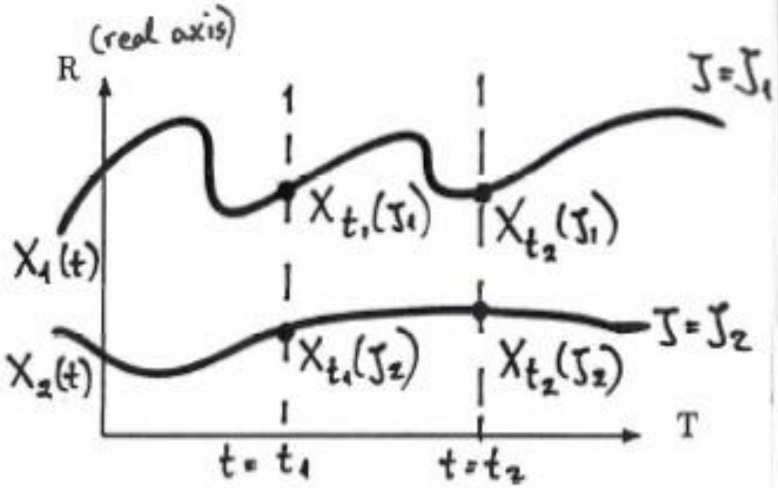
Event occurs in S1 and S2
If S1 and S2 remain identical after the occurence of the event then the event is "deterministic"
If S1 and S2 differ then the event is "random"

In practice, we use the idea of "randomness" (i.e., we use "random" or "probabilistic" models) when we don't know all the factors which might determine the outcome of an experiment. If all the factors are known or assumed known we use deterministic models.

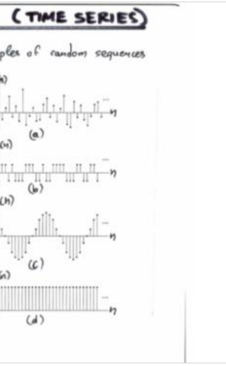
A Random process is simply a mapping from the sample space S to the set of functions defined on T (parameter space)

The set of functions that correspond to elements in S is termed the ensemble of the process. The members of the ensemble represent the possible outcomes which we may observe when we perform the underlying experiment. They are termed the sample functions or realizations





Stochastic Processes may be broadly divided to two classes based on T. If T is countable (i.e., $T = N$) then we say that $\{X_{t\in T}\}$ is a discrete random process. If T is continuous then the $\{X_{t\in T}\}$ is acontinuous stochastic process (i.e., for $T = R$)



RANDOM SEQUENCES (TIME SERIES)

- Parameter space is discrete time.
  $n = 0, \pm 1, \ldots\ldots$

- At time instant $n = n_0$
  $x(n=n_0)$ is a random variable. with probability density function $f_{x(n)}(x)$ and distribution function $F_{x(n)}(x)$

$$F_{x(n)}(x) = P(x(n) < x) \quad , \quad f_{x(n)}(x) = \frac{dF_{x(n)}(x)}{dx}.$$

- Ensemble average or mean

$$m(n) = E\{x(n)\} = \int_{-\infty}^{\infty} x\, f_{x(n)}(x)\, dx.$$

- Ensemble Variance $\sigma^2(n) = E\{|x(n) - m(n)|^2\}$

Examples of random sequences

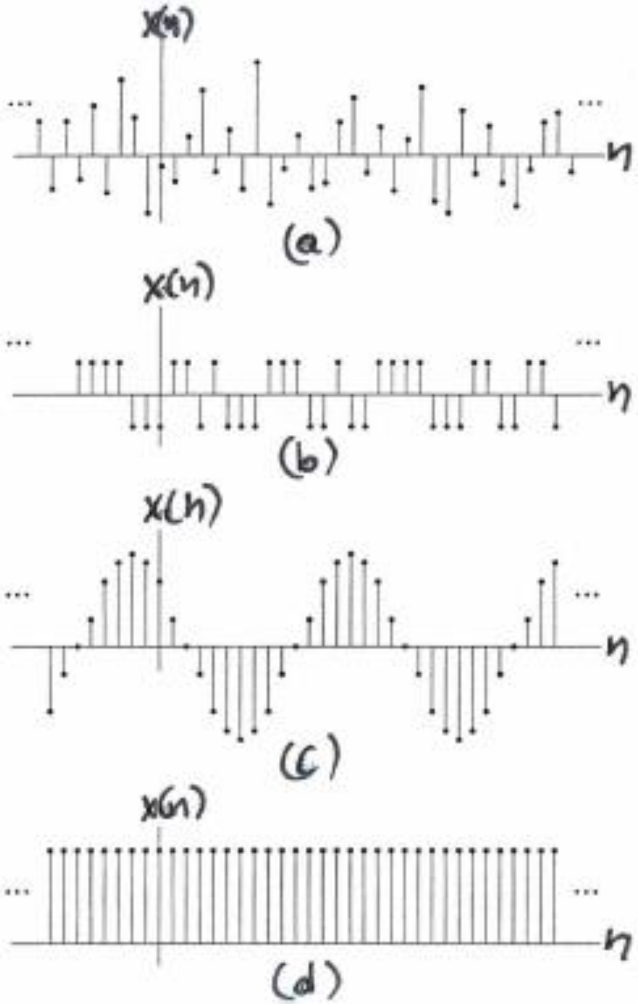
Sample function

Noise type

Binary data

Random Sinusoid

Battery voltage



CHARACTERIZATION OF A RANDOM PROCESS

Full Characterization: K-th order finite order distributions $(K = 1 \rightarrow \infty)$ are the joint distribution functions of the r.vs of the process at any K time instants

$$F_{x(n_1)\dots x(n_k)}(x_1, x_2, \dots x_k) = P\left(X(n_1) \leq x_1, X(n_2) \leq x_2, \dots\dots, X(n_k) < x_k\right)$$

* For all practical purposes a random process is completely characterized by its finite-order distributions

Partial Characterization: In many applications it is sufficient (or by choice) to define of know only a few of the finite-order distributions of the process and thus the corresponding order statistical averages.

* Example: Second order characterization (or second moment analysis) refers to the description of a process in terms of its first and second order distributions and moments (i.e statistical averages)

STATISTICAL AVERAGES - MOMENTS

- mean function : $m(n) = E\{X(n)\} = \int_{-\infty}^{\infty} x\, f_{X(n)}(x)\, dx.$
- Autocorrelation function : $R_X(n_1,n_2) = E\{X^*(n_1)\,X(n_2)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_{n_1}^* x_{n_2}\, f_{X(n_1),X(n_2)}(x_{n_1},x_{n_2})\, dx_1 dx_2$
- Autocovariance function : $R_X(n_1,n_2) - m^*(n_1)\,m(n_2) = E\{(X^*(n_1)-m^*(n_1))(X(n_2)-m(n_2))\}$

- Third order correlation function : $L_X(n_1,n_2,n_3) = E\{X(n_1)X(n_2)X(n_3)\}$

  or $L_X^*(n_1,n_2,n_3) = E\{X(n_1)X^*(n_2)X(n_3)\}$

$\vdots$

Moment Theorem:

moment of order $r = k_1 + \dots + k_n$ $\longrightarrow$ $m_{n_1,n_2,\dots n_n}^{k_1,k_2,\dots k_n} = E\{X^{k_1}(n_1)\cdots X^{k_n}(n_n)\} = (-j)^r \left.\dfrac{\partial^r E\{e^{j\Omega X^T}\}}{\partial \omega_1^{k_1}\cdots\partial\omega_n^{k_n}}\right|_{\omega_i=0}$

cumulant of order $r = k_1 + k_2 + \dots + k_n$ $\longrightarrow$ $c_{n_1\dots n_n}^{k_1\dots k_n} = (-j)^r \left.\dfrac{\partial^r \ln E\{e^{j\Omega X^T}\}}{\partial\omega_1^{k_1}\cdots\partial\omega_n^{k_n}}\right|_{\omega_i=0}$

where $\Omega = [\omega_1\ \omega_2 \dots \omega_n]$ $\quad X = [x(n_1) \dots x(n_n)]$

STATIONARITY

- Strict sense stationarity: All finite-order distributions of the process are time invariant, i.e.

$$\left[F_{x(n_1),\dots\, x(n_k)}(x_1, x_2, \dots\, x_k) = F_{x(n_1+\ell),\dots,\, x(n_k+\ell)}(x_1, x_2, \dots, x_k) \right]$$

for any integer ℓ and any order k.

→ So a process and its shifted version in time are statistically identical
→ statistical description does not depend on n but only on the difference between time instants where the underlying random processes occur. i.e.,

$$E\{x(n)\, x^{k_1}(n+\ell_1) \cdots x^{k_k}(n+\ell_k)\} = L^{k_1+\dots+k_k}(\ell_1, \ell_2, \dots, \ell_k)$$

moment depends only on $\ell_1, \dots, \ell_k$ not on $\underline{\underline{n}}$

→ Dimensionality of statistical averages is reduced.

- K-th order stationarity: All finite order distributions of order $\leq k$ are time invar.

- wide sense stationarity: The mean of the process and its autocovariance are time invariant, that is

$$m_x(n) = \text{constant}, \quad R_x(n_1, n_2) = R_x(n_2 - n_1)$$

ERGODICITY

- Ensemble averages of a process: Based on the statistical averages $E\{\ \}$ of underlying random variables of the process

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

- Time averages: Obtained by time averaging over n one given anyrealization of the process

  $$\left[\langle x(n)\rangle = \lim_{M\to\infty} \frac{1}{2M+1} \sum_{n=-M}^{M} x(n)\right]$$

- Ergodic process (strict sense): A process is ergodic if every time average equals its corresponding ensemble average with probability one.

Ex: ergodic in the mean: $m_n = E\{x(n)\} = \langle x(n)\rangle$

ergodic in autocorrelation: $R_x(k) = E\{x^*(n)\,x(n+k)\} = \langle x^*(n)\,x(n+k)\rangle$

and so on

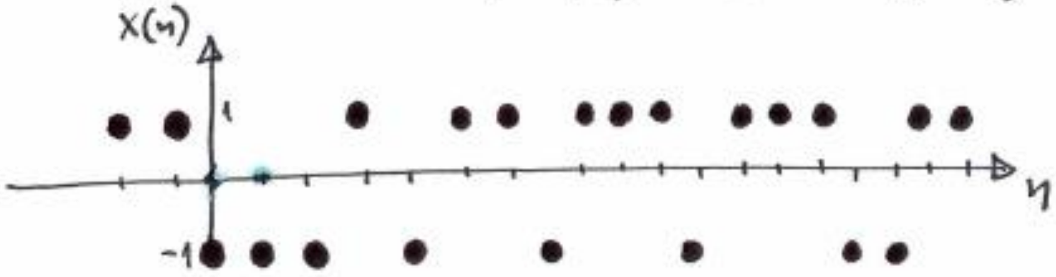
- A process must be stationary in order to be ergodic

* NOTE: Stationarity and ergodicity are convinient assumptions which are approximately met in practice and under certain conditions

THE BERNOULLI PROCESS

- $X(n)$, $n = 0, \pm 1, \pm 2, \ldots$ is a sequence of independent identically distributed (i.i.d) Bernoulli random variables each taking two values 1 and -1 with probabilities p and $1-p$ respectively. $\left\{ \begin{matrix} \text{white binary} \\ \text{noise if } p = \frac{1}{2} \end{matrix} \right\}$

- A typical realization is a random sequence of $\{\pm 1\}$



$$m(n) = E\{X(n)\} = p \cdot 1 - (1-p) \cdot 1 = p - 1 + p = 2p - 1.$$

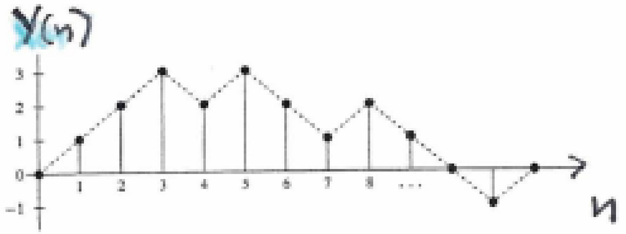
$$R_X(n_1, n_2) = E\{X(n_1) X(n_2)\} = \begin{cases} E\{X(n_1)\} E\{X(n_2)\} & \text{if } n_1 \neq n_2 \\ E\{X^2(n_1)\} & \text{if } n_1 = n_2 \end{cases} = \begin{cases} (2p-1)^2, & n_1 \neq n_2. \\ 1, & n_1 = n_2 \end{cases} = \begin{cases} (2p-1)^2, & n_1 - n_2 \neq 0 \\ 1, & n_1 - n_2 = 0 \end{cases}$$

So the Bernoulli process is Wide Sense Stationary process (WSS)
It is easy to verify that it is also strict sense stationary (SSS) and ergodic.

THE RANDOM WALK PROCESS (or discrete Wiener process)

$$Y(n) = Y_n = \begin{cases} \sum_{i=1}^{n} X_i & n \geq 1 \\ 0 & n = 0 \end{cases}, \qquad X_i = \begin{cases} 1 & prob\, 1/2 \\ -1 & prob\, 1/2 \end{cases}$$

$\{X_i\}$ *i.i.d.*



sample function

Then,

$$E\{Y_n\} = E\{\textstyle\sum_{i=1}^{n} X_i\} = \sum_{i=1}^{n}\{E\{X_i\}\} = 0$$

$$E\{Y_n Y_m\} = E\{\textstyle\sum_{i=1}^{n}\sum_{j=1}^{m} X_i X_j\} = \sum_{i=1}^{n}\sum_{j=1}^{m} E\{X_i X_j\}$$

$$= \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{m} E\{X_i\}\{E\{X_j\}}_{i\neq j} + \underbrace{\sum_{i=1}^{min(n,m)} E\{X_i^2\}}_{i=j}$$

$$= \underbrace{0}_{i\neq j} + \underbrace{min(n,m)}_{i=j} = min(n,m)$$

Concequently, $\eta_Y(n) = 0$, $\underline{R_Y(n,m) = min(n,m) = C_y(n,m)}$ and since we cannot write $min(n,m)$ as a function of $(n-m)$, the random walk process is not WSS as well as SSS However, if we consider the process $\Delta Y_n = Y_{n+m} - Y_n = \sum_{i=n+1}^{n+m} X_i$, then

$$\eta_{\Delta Y}(n) = E\{\Delta Y_n\} = E\{\sum_{i=n+1}^{n+m} X_i\} = E\{Y_{n+m}\} - E\{Y_n\} = 0$$

$$R_{\Delta Y}(n,l) = E\{\Delta Y_n \Delta Y_l\} = E\{Y_{n+m}Y_{l+m}\} + E\{Y_n Y_l\} - E\{Y_{n+m}Y_l\} - E\{Y_n Y_{l+m}\}$$

$$= R_Y(n+m, l+m) - R_Y(n+m, l) - R_Y(n, l+m) + R_Y(n,l)$$

$$= min(n,l) - min(n+m,l) - min(n,l+m) + min(n,l) = \begin{cases} m - |l-n|, & |l-n| < m \\ 0, & |l-n| \geq m \end{cases}$$

Conequently, the process ΔY_n is a W.S.S. process. Actually ΔY_n is a S.S.S process. In other words, the <u>random walk process has stationary increments.</u>

* **martingale process:**
$E\{Y(n)/Y(n-1),\ldots,Y(n_0)\} = Y(n-1)$

* The random walk-process is the starting point for generalizations heading to important continuous random processes such as the Wiener process. which is a continuous Gaussian process with $\begin{cases} m_X(t) = 0 \\ R_X(t_1,t_2) = 6^2 min(t_1,t_2) \end{cases}$

PERIODIC AND ALMOST PERIODIC RANDOM PROCECESS (1)

- **Periodic random process:** There exists an integer P such that.

$$\left[F_{x(n_1)\ldots x(n_L)}(x_1, x_2, \ldots, x_L) = F_{x(n_1+k_1P),\, x(n_2+k_2P)\ldots x(n_L+k_LP)}(x_1, \ldots, x_L) \right]$$

for any $\{n_i\}$, $\{k_i\}$ and L

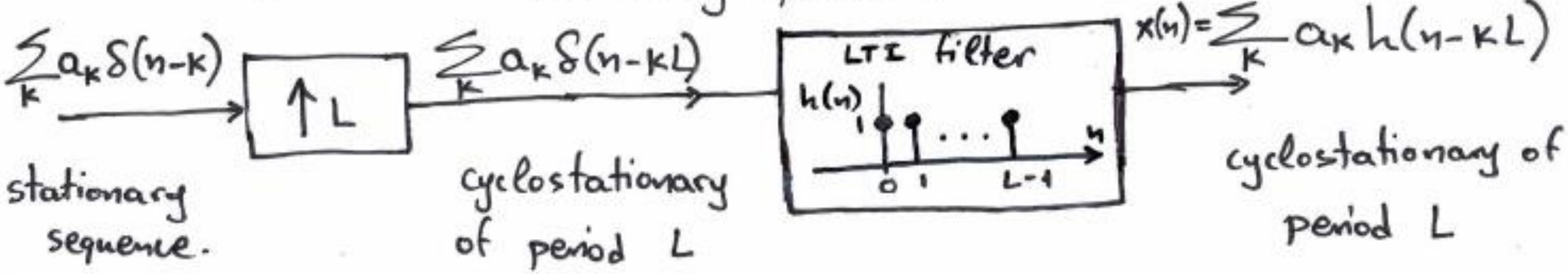
- **Cyclostationary processes:** A periodic process with $k_1 = k_2 = \ldots = k_L = m$
  - → So the statistics do not change when the process is shifted by multiples of P
  - → The mean, autocorrelation and other statistical averages are periodic in n with period P

    i.e. $m_x(n) = m_x(n+mP)$, $R_x(n_1, n_2) = R_x(n_1+mP, n_2+mP)$ .....

  - → Wide Sense cyclostationarity (WSCS): The mean and autocorrelation of the process are periodic with period P.
  - → It is possible to create a stationary process from a cyclostationary by introducing a random time shift independent of the process or by dowsampling the process with the appropriate rate.

PERIODIC AND ALMOST PERIODIC R.PR. (2)

- **Ex.** Let $\{a_n\}$ be a sequence of i.i.d r.v.s taking the values ± 1 with equal probability (that is binary white noise).
  Consider the following operation



  * $x(n)$ is the discrete equivalent of a binary communication signal with rectangular pulse.

- **Ex.** $x(n) = A\cos(\omega_0 n + \varphi)$ where A and φ are independent random variables
  → If $f_0 = \frac{\omega_0}{2\pi} = \frac{K}{P}$: rational then the process is periodic with period P
  → If $\varphi \sim U(0, 2\pi)$ then the process is stationary.

  Otherwise, the process is called "almost periodic".

GAUSSIAN RANDOM PROCESSES (DISCRETE)

* **Gaussian r. process**: Every finite-order distribution of the process is that of a collection of jointly Gaussian random variables (i.e. Given the r.v.s $\underline{X} = [x(n_1)\ x(n_2) \ldots x(n_N)]$, for any N

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{N/2}\,|\underline{C}_x|^{1/2}}\, e^{-\frac{1}{2}(\underline{X}-\underline{m}_x)^T \underline{C}_x^{-1}(\underline{x}-\underline{m}_x)} \qquad \text{(real random vector)}$$

or

$$f_{\underline{X}}(\underline{x}) = \frac{1}{\pi^N\,|\underline{C}_x|}\, e^{(\underline{x}-\underline{m}_x)^T \underline{C}_x^{-1}(\underline{x}-\underline{m}_x)} \qquad \text{(complex random vector)}$$

where, the covariance matrix $\underline{C}_x = E\{|\underline{X}-\underline{m}_x|^H\,|\underline{X}-\underline{m}_x|\}$, $\underline{m}_x = E\{\underline{X}\}$.

* A complex random variable is defined as $x_1(n) + j x_2(n)$ where $x_1(n), x_2(n)$ are r.v.s
* The existence of many Gaussian processes is derived by the Central limit Theorem (CLT)
* A Gaussian random process is completely specified by its mean and autocovariance functions — It is a second order process.
* A Gaussian random process is WSS if and only if it is SSS
* A Gaussian random process remains Gaussian under linear transformations

MARKOV PROCESSES

- A Markov Process is a random process satisfying the property that given the process infinite past, the distribution of $x(n)$ depends only on the previous sample $x(n-1)$, i.e.

$$\left[\, f_{x(n)/x(n-1),x(n-2)\ldots}(x_n/x_{n-1},x_{n-2},\ldots) = f_{x(n)/x(n-1)}(x_n/x_{n-1}) \,\right]$$

  * All independent processes are Markov processes
  * Given the density $f_{x(n)}(x_n)$ of $x(n)$ and the conditional distribution (or density) for $f_{x(n)/x(n-1)}(x_n/x_{n-1})$ we can completely determine the process. since we can easily calculate all finite order distributions (densities) of the process

- Markov chains: A Markov process that takes only discrete values (states)

Then, $$\left[P[X(n)=x_n/\,x(n-1)=x_{n-1}, x(n-2)=x_{n-2}\ \ldots] = P[X(n)=x_n/_{x(n-1)=x_{n-1}}] \right]$$

  * Useful processes for applications in digital signal processing

Ex. The random walk process is a Markov chain with an infinite number of states

SECOND MOMENT ANALYSIS (2nd-order statistics)

* Description of the process is limited to first and second order statistical domains (described by either the first and second order finite distributions or first and second order statistical averages).

* Given an arbitrary process only part of the process story is revealed.

* Natural processing environment for Gaussian processes – The full story about the process is told.

→ mean function: $m_x(n) = E\{x(n)\} \overset{wss}{=} m_x$ constant if wss

→ autocorrelation function: $R_x(n_0,n_1) = E\{x^*(n_0)x(n_1)\} \overset{w.ss}{=} R_x(n_1-n_0)$ time invariant if wss

→ autocovariance: $C_x(n_0,n_1) = E\{(x(n_0)-m_x(n_0))^*(x(n_1)-m_x(n_1))\} \overset{wss}{=} C_x(n_1-n_0)$ time inv. if wss.

Also, $C_x(n_0,n_1) = R_x(n_0,n_1) - m_x^*(n_0)m_x(n_1)$

PROPERTIES OF CORRELATION - COVARIANCE FUNCTIONS

Let $R_x(\ell) = E\{x^*(n)x(n+\ell)\}$, $C_x(\ell) = R_x(\ell) - |m_x|^2$ (w.s.s. assumption)

- Conjugate Symmetry : $\left[R_x(\ell) = R_x^*(-\ell) \quad , \quad C_x(\ell) = C_x^*(-\ell) \right]$

- Positive semidefinite property : $\left[\sum_{n_1=-\infty}^{\infty} \sum_{n_0=-\infty}^{\infty} a^*(n_1) R_x(n_1-n_0) a^*(n_0) \geq 0 \right]$ for any sequence $a(n)$

  $\Downarrow$

  (By replacing R_x with C_x or (n_1-n_0) with (n_1,n_0) the property remains valid.)

- $R_x(0) \geq |R_x(\ell)|$, $C_x(0) = Var(x(n)) \geq |C_x(\ell)|$

  (Proof: use the previous two properties by properly choosing $a(n)$)

- Random process periodic $\longrightarrow$ all moments are periodic with same period.

- Property true for correlation $\underset{\text{unless process is zero mean}}{\overset{\longrightarrow}{\not\longleftarrow}}$ Property true for covariance.

  - $R_x(\ell) = 0 \Rightarrow x(n), x(n+\ell)$ orthogonal
  - $C_x(\ell) = 0 \Rightarrow x(n), x(n+\ell)$ uncorrelated

COVARIANCE MATRIX

Given a process $\{x(n)\}$ and a vector $\underline{x} = [x(n_0), x(n_0+1), \ldots, x(n_0+N-1)]^T$ from the process

Covariance Matrix $(N \times N)$: $\underline{C}_x = E\{(\underline{x}-\underline{m}_x)(\underline{x}-\underline{m}_x)^{*T}\} = \begin{bmatrix} C_x(n_0,n_0) & C_x(n_0,n_0+1) & \cdots & C_x(n_0,n_0+N-1) \\ C_x(n_0+1,n_0) & C_x(n_0+1,n_0+1) & & \\ \vdots & & \ddots & \\ C_x(n_0+N-1,n_0) & & & C_x(n_0+N-1,n_0+N-1) \end{bmatrix}$

Hermitian symmetry but not Toeplitz.

where: $\underline{m}_x = E\{\underline{X}\} = [E\{x(n_0)\} \ldots E\{x(n_0+N-1)\}]^T$

* Correlation matrix is defined similarly as $\underline{R}_x = E\{\underline{X}\underline{X}^H\}$
* Hermitian symmetry: For stationary (wss) processes the correlation and covariance matrices are always Hermitian Symmetric Toeplitz matrices

$$\underline{C}_x = \begin{bmatrix} C_x(0) & C_x(-1) & \cdots & C_x(-N+1) \\ C_x(1) & C_x(0) & & \vdots \\ \vdots & & \ddots & \\ C_x(N-1) & \cdots & & C_x(0) \end{bmatrix}$$

stationarity assumption

* $\underline{R}_x$ and $\underline{C}_x$ are positive semidefinite matrices, i.e. $\underline{a}^H \underline{R}_x \underline{a} \geq 0$ for any vector $\underline{a}$.

CROSS-CORRELATION, CROSS-COVARIANCE

Given two random processes $\{x(n)\}$, $\{y(n)\}$

* $R_{xy}(n_1, n_0) = E\{x(n_1)\, y^*(n_0)\}$ (cross-correlation), $C_{xy}(n_1, n_0) = R_{xy}(n_1, n_0) - m_x(n_1)\, m_y^*(n_0)$ (Cross-Covariance)

Orthogonal processes: $R_{xy}(n_1, n_0) = 0$ Uncorrelates processes: $C_{xy}(n_1, n_0) = 0$

for all (n_1, n_0).

* **Jointly stationary processes**: The marginal and joint finite order distributions of the two processes are all time invariant.

Jointly WSS (or WSS) : $x(n)$, $y(n)$ are WSS and $R_{xy}(n_1, n_0) = R_{xy}(n_1 - n_0)$

→ $R_{xy}(\ell) = R_{yx}^*(-\ell)$

→ The cross-correlation is not positive-semidefinite in general

→ $|R_{xy}(\ell)| \leq \sqrt{R_x(0) R_y(0)}$ and $|R_{xy}(\ell)| \leq \frac{1}{2}\left(R_x(0) + R_y(0)\right)$

B(→ Cross-Correlation matrix $\underline{R}_{xy}$: $\begin{cases} \text{non-symmetric in general.} \\ \text{Toeplitz if jointly stationary} \end{cases}$

ESTIMATION OF COVARIANCE FUNCTION

We are given a segment $X_N(n)$, $n=0,\dots,N-1$ of a stationary discrete random process $X(n)$ and we wish to estimate the mean, autocorrelation and autocovariance functions of the process. Then, assuming that the process is at least second-order ergodic we can estimate these quantities as follows:

$$\left\{\begin{array}{l} \hat{m}_x = \dfrac{1}{N}\displaystyle\sum_{n=0}^{N-1} X(n) \\ \hat{R}_x(\ell) = \dfrac{1}{N}\displaystyle\sum_{k=\max(0,-\ell)}^{\min(N,N-\ell-1)} X^*(k)\,X(k+\ell), \quad |\ell| \leq L \\ \hat{C}_x(\ell) = \hat{R}_x(\ell) - |\hat{m}_x|^2 \end{array}\right.$$

biased estimates

(to obtain un-biased estimates replace division in $\hat{R}_x(\ell)$ by $N-|m|$.)

* The above are biased however assymptotically unbiased estimates ($N \to \infty$)
* Biased estimates are often preferred in practice because they exhibit "better behaviour" in some applications!!

POWER SPECTRAL DENSITY (WSS processes)

By definition: given a WSS process $x(n)$

$$\left[\begin{aligned} S_x(\omega) &= \sum_{k=-\infty}^{\infty} R_x(k)\, e^{-j\omega k} \\ R_x(k) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} S_x(\omega)\, e^{j\omega k}\, d\omega \end{aligned}\right]$$

→ periodic in ω (2π), real and nonnegative.

Wiener-Khinchine Theorem

So the PSD $S_x(\omega)$ is defined as the DTFT of the autocorrelation $R_x(k)$

** Often the mean of the process is removed prior to the estimation of the PSD (or simply power spectrum) to avoid confusion between the spectral components of a non-zero mean process from a process with random mean (dc) component.

** Note that the spectrum (Fourier transform) of $x(n)$ is also random, however the PSD is deterministic.

** It can be shown that under certain conditions $\lim_{N\to\infty} \frac{1}{2N+1} E\left\{\left|\sum_{k=-N}^{N} X(n)\, e^{-j\omega n}\right|^2\right\} = \sum_{k=-\infty}^{\infty} R_x(k)\, e^{-j\omega}$

** For a non-stationary process you may define $S_x(\omega_1,\omega_2) = \sum_{n,\ell=-\infty}^{\infty}\sum R_x(n,\ell)\, e^{-j\omega_1 n} e^{-j\omega_2 \ell}$

or $S_x(n,\omega) = \sum_{\ell=-\infty}^{\infty} R_x(n,\ell)\, e^{-j\omega(\ell-n)}$ which is a time dependent PSD.

Cross-Power Density Spectrum

By definition: given two processes $x(n), y(n)$ jointly stationary.

$$\left[S_{xy}(\omega) = \sum_{\ell=-\infty}^{\infty} R_{xy}(\ell) e^{-j\omega\ell} \right] : \text{periodic in } \omega \ (2\pi), \quad S_{xy}(\omega) = S_{yx}^{*}(\omega)$$

Conjugate symmetric

$$R_{xy}(\ell) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{xy}(\omega) e^{j\omega\ell} d\omega$$

- Normalized cross-spectrum (coherence function)

$$\Gamma_{xy}(\omega) = \frac{S_{xy}(\omega)}{\sqrt{S_x(\omega)}\sqrt{S_y(\omega)}}, \qquad 0 \le |\Gamma_{xy}(\omega)|^2 \le 1$$

* Note: $S_x(\omega)$ carries amplitude (magnitude) information

$S_{xy}(\omega)$ " significant phase information.