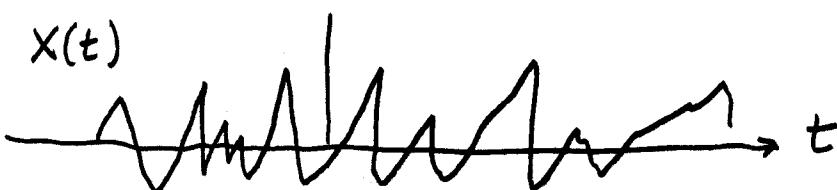


MODULE 2

STOCHASTIC PROCESSES

- The terms "stochastic process" and "random process" are synonyms.
- We use functions to model deterministic signals.
Ex: $x(t) = \cos\omega_0 t$
- We use stochastic processes to model random signals... a.k.a. statistical signals.
- Associated with a random process is an indexing set \mathcal{L} .
 - The indexing set can be uncountable, countable, or finite.
 - Ex: $\mathcal{L} = \mathbb{R}$, $\mathcal{L} = \mathbb{Z}$, $\mathcal{L} = \{-2, -1, 0, 1, 2\}$.
- We usually think of the indexing set as corresponding to time.
- A stochastic process is a collection of RVs, one for each element of \mathcal{L} .
 - Since the indexing set is ordered, the RVs are also ordered (in time).

- Intuitively, the stochastic process $X(t)$ has an RV at every $t \in \mathbb{R}$.
 - Likewise, the stochastic process X_k has an RV at every $k \in \mathbb{Z}$.
 - All of the RV's associated with a stochastic process have the same underlying set of experimental outcomes S .
 - They also have the same "domain", " σ -algebra", or "set of events" \mathcal{G} for their probability measure.
 - However, they may have different probability measures P in general.
- ⇒ For each trial of the experiment, each RV maps the experimental outcome $s \in S$ to a real number.
- Thus, for experimental outcome s , the stochastic process is a function:
- EX: 
- ⇒ The above function is called a "sample function" or "realization" of the process $X(t)$ corresponding to the experimental outcome s .

- Thus, we see that a stochastic process is actually a mapping from $\mathcal{X} \times \mathcal{S}$ into \mathbb{R} .
- A stochastic process $X(t)$ with indexing set $\mathcal{I} = \mathbb{R}$ is called a "continuous-time" stochastic process.
- A stochastic process with indexing set $\mathcal{I} = \mathbb{Z}$ is called a "discrete time" random process.
- To describe a random process completely, it is necessary to specify the joint density or distribution of all of the involved RVs.

NOTE: We have written $X(t)$ to denote a continuous-time process.

This is slightly misleading because the process $X(t)$ is not a function of t in the usual sense.

\Rightarrow However, any single sample function (realization) of the process is a function of t .

- We will sometimes use the alternative notation X_t to denote the process $X(t)$.
 → This gives us better consistency with the discrete-time notation X_k .

Joint Densities for 2 Processes

- Suppose X_t and Y_t are two continuous-time random processes.

- Let $t_1, t_2, \dots, t_m, t'_1, t'_2, \dots, t'_m$ be a set of $2M$ discrete time instants.

- Then we can consider the joint density function $f_{X_{t_1}, X_{t_2}, \dots, X_{t_m}, Y_{t'_1}, Y_{t'_2}, \dots, Y_{t'_m}}(x_{t_1}, \dots, x_{t_m}, y_{t'_1}, \dots, y_{t'_m})$.

→ This is often abbreviated

$$f_{x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_m}(x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_m).$$

- The processes X_t and Y_t are independent if

$$f_{x_1, \dots, x_m, y_1, \dots, y_m}(x_1, \dots, y_m)$$

$$= f_{x_1, \dots, x_m}(x_1, \dots, x_m) f_{y_1, \dots, y_m}(y_1, \dots, y_m)$$

for any choice of $t_1, \dots, t_m, t'_1, \dots, t'_m$.

"Deterministic" Random Process

- This term is an oxymoron!
- The book uses it to describe a process like
 - $X(t) = \theta$, where θ is a $N(m, \sigma^2)$ RV.
 - $X(t) = \theta(2 + \sin\omega t)$, θ an RV.
- For any realization, knowledge of the value of $X(t)$ at any specific t gives us knowledge of the entire realization.

Note: In both cases above, there is really only a single RV involved.

Note: In both cases above, the correlation coefficient between X_{t_1} and X_{t_2} has magnitude 1 $\forall t_1, t_2 \in \mathbb{R}$.

\Rightarrow i.e., this is not a very interesting class of processes.

\Rightarrow The term "deterministic random process"¹⁾ is not standard.

Correlation and Covariance for Discrete-Time Processes

- Let X_k and Y_k be discrete-time processes.
- The "Cross Correlation" function of X_k and Y_k is defined by

$$R_{x,y}(i,j) = E[x_i y_j]$$

- The "cross Covariance" function is

$$\begin{aligned} C_{x,y}(i,j) &= \text{Cov}(x_i, y_j) \\ &= E[(x_i - E[x_i])(y_j - E[y_j])] \\ &= E[x_i y_j] - E[x_i]E[y_j] \end{aligned}$$

Note: $R_{x,y}(i,j) = \text{Cov}(x_i, y_j) + E[x_i]E[y_j]$.

- The "Autocorrelation" of X_k is the cross-correlation of X_k with itself;

$$R_x(i,j) = R_{x,x}(i,j) = E[x_i x_j].$$

Note: $R_x(i,j) = \text{Cov}(x_i, x_j) + E[x_i]E[x_j]$.

NOTE: When X_k is a stochastic process, $\text{Cov}(X_i, X_j)$ is often called the "Autocovariance" of X_k .

Correlation and Covariance for Continuous-Time Processes

- Let X_t and Y_t be continuous-time stochastic processes.
- Let s, t be real variables.
- The "Cross Correlation" function of X_t and Y_t is given by

$$R_{x,y}(s,t) = E[X_s Y_t].$$

- The "Cross Covariance" function is given by

$$\begin{aligned} C_{x,y}(s,t) &= \text{Cov}(X_s, Y_t) \\ &= E[(X_s - E[X_s])(Y_t - E[Y_t])]. \end{aligned}$$

NOTE:

$$R_{x,y}(s,t) = \text{Cov}(X_s, Y_t) + E[X_s]E[Y_t].$$

- The "Autocorrelation" function of X_t is the cross-correlation of X_t with itself:

$$\begin{aligned}
 R_X(s, t) &= R_{X, X}(s, t) \\
 &= E[X_s X_t] \\
 &= \text{Cov}(X_s, X_t) + E[X_s]E[X_t].
 \end{aligned}$$

- When X_t is a stochastic process, $\text{Cov}(X_s, X_t)$ is often called the "Autocovariance" function of the process.

STATIONARY PROCESSES

- A stochastic process is called "strict sense stationary" if the joint density of any number of the involved RV's is invariant under time translation.

EX: $f_{X_{t_1}, X_{t_2}}(x_{t_1}, x_{t_2}) = f_{X_{t_1+\Delta}, X_{t_2+\Delta}}(x_{t_1+\Delta}, x_{t_2+\Delta})$

\Rightarrow For strict sense stationarity, or "SSS", this must be true for all joint densities of all order and $\forall \Delta \in \mathbb{R}$.

$(\forall \Delta \in \mathbb{Z} \text{ if discrete-time}).$

- A stochastic process is called "wide sense stationary" (WSS) if the mean and autocorrelation are invariant under time translation:

continuous time

$$\left\{ \begin{array}{l} E[X_{t_1}] = E[X_{t_1 + \hat{t}}] \\ R_x(t_1, t_2) = R_x(t_1 + \hat{t}, t_2 + \hat{t}) \end{array} \right. \quad \forall t_1, t_2, \hat{t} \in \mathbb{R}$$

discrete time

$$\left\{ \begin{array}{l} E[X_i] = E[X_{i + \hat{k}}] \\ R_x(i, j) = R_x(i + \hat{k}, j + \hat{k}) \end{array} \right. \quad \forall i, j, \hat{k} \in \mathbb{Z}$$

- For WSS processes, the mean is constant and the autocorrelation depends only on the "~~shift amount~~" \hat{t} or k . displacement $|t_2 - t_1|$ or $|j - i|$

→ We write:

$$R_x(t_1, t_2) = R_x(\tau) = R_x(|t_2 - t_1|)$$

$$R_x(i, j) = R_x(k) = R_x(|j - i|)$$

for WSS processes.

Intuition:

- For an SSS process, all joint moments are invariant under time translation.
- For a WSS process, the first and second order moments are invariant under time translation.

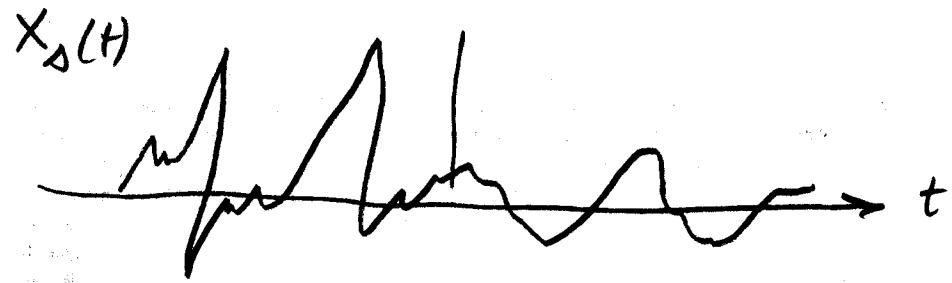
NOTE:

1. SSS implies WSS.
2. The converse is FALSE.

Ergodicity

- The concept of ergodicity is related to but distinct from that of stationarity.
- Let $X(t)$ be a random process with the underlying set of experimental outcomes S .

- Any given trial of the experiment results in a specific outcome $s \in S$, which the process maps to a sample function (realization) $X_s(t)$:



- We can compute the moments of $X_s(t)$:

→ First moment (sample mean):

$$\bar{X}_s = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X_s(t) dt$$

→ Second central moment:

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T [X_s(t) - \bar{X}_s]^2 dt$$

- These moments are usually referred to as "time averages".
- Computing them involves fixing the experimental outcome and averaging over time.

- Alternatively, we could fix time and average over experimental outcomes.
- Let t be fixed at t_0 .

→ First moment (mean of $X(t_0)$):

$$E[X(t_0)] = \int_{-\infty}^{\infty} \theta f_{X(t_0)}(\theta) d\theta$$

→ Second Central Moment:

$$\begin{aligned} E[(X(t_0) - E[X(t_0)])^2] \\ = \int_{-\infty}^{\infty} \{ \theta - E[X(t_0)] \}^2 f_{X(t_0)}(\theta) d\theta \end{aligned}$$

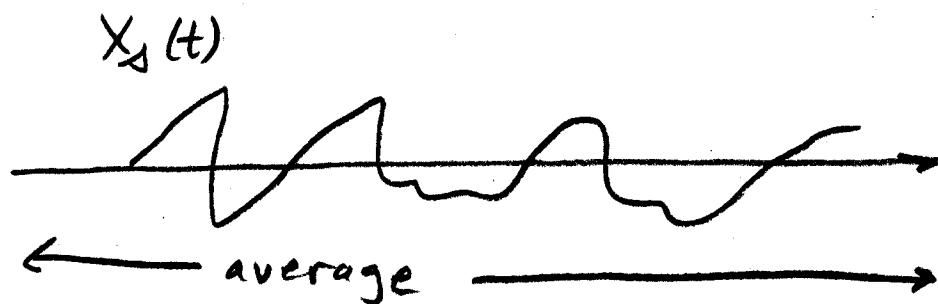
⇒ The above are immediately recognized as the mean and variance of the RV $X(t_0)$.

⇒ In the context of stochastic processes, moments computed by averaging over experimental outcomes for a fixed time are called "ensemble averages".

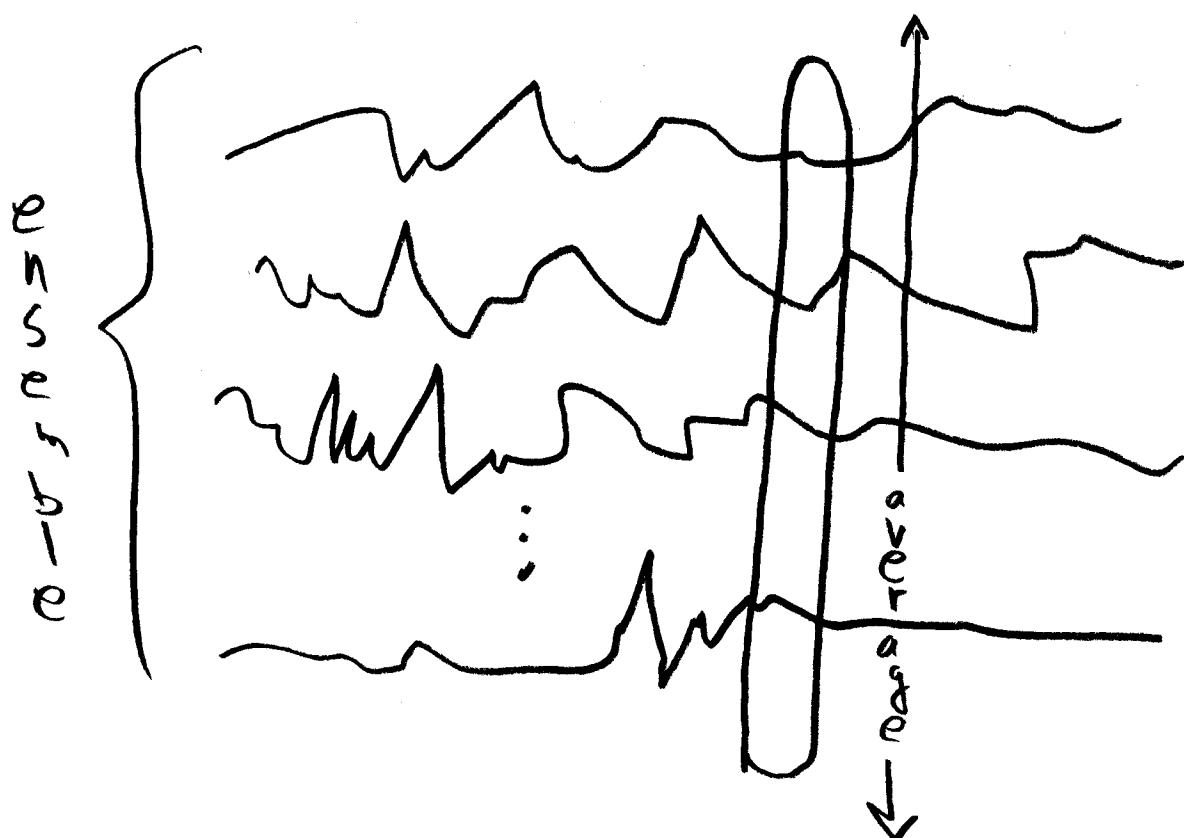
- An "ensemble" is the collection of sample functions generated as the experimental outcome s varies over all of S .

- Thus:

→ Time averaging : fix s :



→ Ensemble averaging : fix t :



- If all time averages are equal to their corresponding ensemble averages, then the process $X(t)$ is called ergodic.
- For an ergodic process, all of the moments -- including autocorrelation and autocovariance -- can be calculated from a single sample function.

EX : For a WSS ergodic process $X(t)$, the autocorrelation function can be computed from a single realization $X_s(t)$ according to

$$R_x(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X_s(t) X_s(t+\tau) dt$$

SUPPLEMENT TO

PAGES 2.10-2.14: Ergodicity.

- Ergodicity is discussed in Section 2.4 of the book.

→ The book is quite informal and vague:

"A random process is said to be "ergodic" if time averaging is equivalent to ensemble averaging!"

- The following is from A. Papoulis, "Probability, Random Variables, and Stochastic Processes," McGraw-Hill, New York, 1984.

- Let $\{s_i\}$ be an indexed set of experimental outcomes, where each $s_i \in S$.

- Let $X(t)$ be a random process and $X_{s_i}(t)$ be the sample function obtained when s_i is the experimental outcome.

- Suppose we wish to find the mean of the process $\eta(t)$.

- Given N sample functions $X_{s_i}(t)$, we can approximate the mean by the ensemble average

$$\eta(t) \approx \eta_N(t) = \frac{1}{N} \sum_i X_{s_i}(t).$$

- In the limit as $N \rightarrow \infty$,

$$\lim_{N \rightarrow \infty} \eta_N(t) = \eta(t).$$

- Given a single sample function $X_{s_i}(t)$, we can also compute the time average (a number)

$$\bar{X} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X_{s_i}(t) dt.$$

- If X is not stationary, so that $\eta(t)$ varies with t , then \bar{X} is clearly a poor estimate of $\eta(t)$. In this case, the process $X(t)$ is not ergodic.
- If the process is stationary, so that $\eta(t) = \eta$ (a number), then it may be that $\bar{X} = \eta$.

→ Ergodicity describes the conditions under which time averages are equal to the corresponding ensemble averages.

DEF: A process $X(t)$ is called "ergodic" if its ensemble averages equal the appropriate time averages.

⇒ This means that, with probability 1, any statistic of $X(t)$ can be determined from a single sample function $X_{s_i}(t)$.

→ Often, we are interested only in a specific set of statistics. This leads to less restricted, specialized types of ergodicity.

DEF: Given a process $X(t)$ with constant mean $E[X(t)] = \eta$ and the time average

$$\bar{\eta}_T = \frac{1}{2T} \int_{-T}^T X(t) dt,$$

the process is called mean-ergodic if

$$\lim_{T \rightarrow \infty} \bar{\eta}_T = \eta$$

with probability 1.

Note: $\bar{\eta}_T$ is an RV. Whether $X(t)$ is mean ergodic or not, we have that

$$E\{\bar{\eta}_T\} = \frac{1}{2T} \int_{-T}^T E\{X(t)\} dt = \eta.$$

→ Thus, $E\{\bar{\eta}_T\} = \eta$ is not a sufficient test for mean ergodicity.

→ The variance $\text{Var}\{\bar{\eta}_T\} = \sigma_T^2$, however, does provide a sufficient test.

→ $X(t)$ is mean ergodic iff

★ ★

$$\lim_{T \rightarrow \infty} \sigma_T^2 = 0.$$

Theorem : A process $X(t)$ with constant mean η is mean ergodic iff its autocovariance $C_X(t_1, t_2)$ satisfies

$$\lim_{T \rightarrow \infty} \frac{1}{4T^2} \int_{-T}^T \int_{-T}^T C_X(t_1, t_2) dt_1 dt_2 = 0.$$

Corollary : A WSS process $X(t)$ with constant mean η is mean ergodic if its autocovariance $C_X(\tau) = R_X(\tau) - \eta^2$ satisfies

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-2T}^{2T} C_X(\tau) \left(1 - \frac{|\tau|}{2T}\right) d\tau = 0.$$

Sufficient Condition: If $X(t)$ is WSS and if

$$\int_{-\infty}^{\infty} |C_X(\tau)| d\tau < \infty,$$

then $X(t)$ is mean ergodic.

Sufficient Condition: If $X(t)$ is WSS, if $C_X(0) < \infty$, and if $\lim_{|\tau| \rightarrow \infty} C_X(\tau) = 0$, then $X(t)$ is mean ergodic.

Distribution Ergodic Processes

- Let $X(t)$ be a SSS process. Then all the RV's that make up $X(t)$ have the same distribution (cdf)

$$F(x) = P\{X(t) \leq x\}.$$

- we wish to determine $F(x)$ from a single sample function $X_s(t)$.
- For each value x , we form the new process

$$y(t) = \begin{cases} 1, & X(t) \leq x \\ 0, & X(t) > x \end{cases}.$$

Note: $P\{y(t) = 1\} = P\{X(t) \leq x\} = F(x)$

$$P\{y(t) = 0\} = P\{X(t) > x\} = 1 - F(x)$$

$$\begin{aligned} \Rightarrow E\{y(t)\} &= P\{y(t) = 1\} \cdot 1 + P\{y(t) = 0\} \cdot 0 \\ &= P\{y(t) = 1\} \\ &= F(x). \end{aligned}$$

- The question is: can we time average a sample function of $y(t)$ to determine $F(x)$?

→ The answer is YES, provided $y(t)$ is mean ergodic. In this case we say that the process $X(t)$ is "distribution ergodic".

- More formally, for a sample function of $X(t)$ let $\tau_1, \tau_2, \dots, \tau_n$ be the lengths of the time intervals where the sample function is $\leq x$.

- Define the time average

$$y_T = \frac{1}{2T} \int_{-T}^T y(t) dt = \frac{\sum_{i=1}^n \tau_i}{2T}$$

DEF: if $\lim_{T \rightarrow \infty} y_T = F(x)$, then we say that $X(t)$ is "distribution ergodic".

Theorem: Let $F(x_1, x_2; \tau) = P\{x(t+\tau) \leq x_1, x(t) \leq x_2\}$.

Then an SSS process $X(t)$ is distribution ergodic iff

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-2T}^{2T} \left(1 - \frac{|\tau|}{2T}\right) [F(x, x; \tau) - F^2(x)] d\tau = 0.$$

Correlation Ergodic Processes

- Let $X(t)$ be WSS.
- Given a single sample function of $X(t)$, we want to find the autocorrelation

$$R_X(\tau) = E\{X(t)X(t+\tau)\}.$$

- For each value τ , we form the new process

$$Z(t) = X(t)X(t+\tau).$$

Note: $E\{Z(t)\} = E\{X(t)X(t+\tau)\} = R_X(\tau).$

- The question is: can we time average a sample function of $Z(t)$ to determine $R_X(\tau)$?

→ The answer is YES, provided $Z(t)$ is mean ergodic.
In this case we say that the process $X(t)$ is "correlation ergodic"!

- More formally, define the time average

$$R_T = \frac{1}{2T} \int_{-T}^T Z(t) dt = \frac{1}{2T} \int_{-T}^T X(t)X(t+\tau) dt.$$

DEF: if $\lim_{T \rightarrow \infty} R_T = R_X(\tau)$, then we say that $X(t)$ is "correlation ergodic"

$$\begin{aligned}\text{Note: } R_z(\lambda) &= E\{z(t)z(t+\lambda)\} \\ &= E\{x(t)x(t+\tau)x(t+\lambda)x(t+\lambda+\tau)\}.\end{aligned}$$

→ THE Autocovariance of $z(t)$ is

$$C_z(\lambda) = R_z(\lambda) - R_x^2(\tau)$$

⇒ $X(t)$ is correlation ergodic iff $C_z(\lambda)$ satisfies the corollary on PAGE 2.14.4;
that is, iff

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{2T} C_z(\lambda) \left(1 - \frac{|\lambda|}{2T}\right) d\lambda = 0.$$

Note: if $X(t)$ is correlation ergodic, then

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X^2(t) dt = E\{X^2(t)\}$$

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \{x(t+\tau) + x(t)\}^2 dt = 2[R_X(0) + R_X(\tau)]$$

Stationary Autocorrelation Function

- Let $X(t)$ be a WSS process with autocorrelation $R_X(\tau)$.

$$1. R_X(0) = E[X_t^2] \geq 0.$$

commutativity
of multiplication

$$\begin{aligned} 2. R_X(-\tau) &= E[X_t X_{t-\tau}] = E[X_{t-\tau} X_t] \\ &= E[X_t X_{t+\tau}] = R_X(\tau) \end{aligned}$$

↑
WSS

$$3. |R_X(\tau)| \leq R_X(0).$$

Follows from Cauchy - Schwartz inequality and the fact that $|\rho_{X_t X_{t+\tau}}| \leq 1$.

Cross-Correlation for WSS Processes

- Let X_t and Y_t be two WSS processes.
- If the cross-correlation function $R_{x,y}(s,t)$ is invariant to time shifts, then it is a function of $|s-t| = \tau$.

- In this case, we write

$$R_{x,y}(\tau) = R_{x,y}(s,t) = R_{x,y}(|s-t|)$$

→ The following properties then hold:

1. $R_{x,y}(0) = R_{y,x}(0)$
2. $R_{x,y}(\tau) = R_{y,x}(-\tau)$
3. $|R_{x,y}(\tau)| \leq [R_x(0) R_y(0)]^{1/2}$

→ The processes X_t and Y_t are called "jointly wide sense stationary".

Power Spectral Density

- For a WSS continuous-time process $X(t)$, the "power spectral density" is the Fourier Transform of the autocorrelation $R_X(\tau)$:

$$S_X(\omega) = \mathcal{F}[R_X(\tau)] = \int_{-\infty}^{\infty} R_X(\tau) e^{-j\omega\tau} d\tau$$

- The book writes $S_X(j\omega)$ instead of $S_X(\omega)$.
- The power spectral density is also known as:

power spectrum
spectral density
PSD

1. $\boxed{- \text{Because } R_X(\tau) \text{ is real and even, } S_X(\omega) \text{ is also real, even, and non-negative.}}$
2. $S_X(\omega)$ is also real, even, and non-negative.

- The autocorrelation can be recovered from the PSD using the inverse Fourier transform

$$R_X(\tau) = \mathcal{F}^{-1}[S_X(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(\omega) e^{j\omega\tau} d\omega$$

- Plugging in $\tau=0$, we obtain

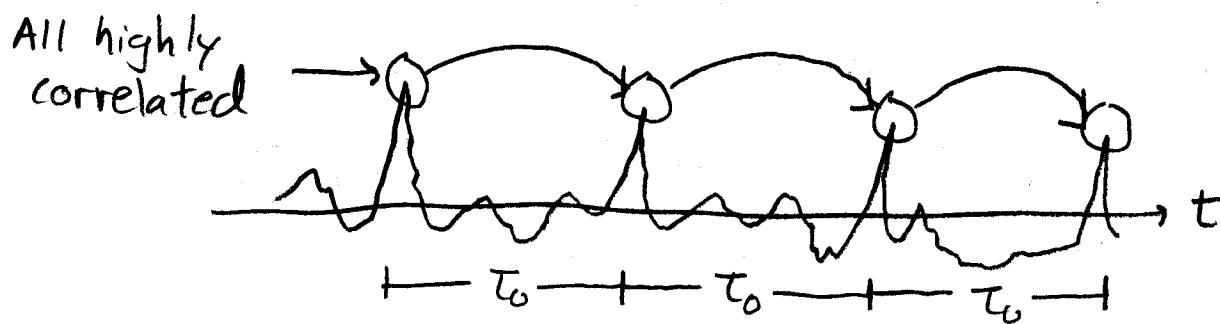
$$R_X(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(\omega) d\omega = E[X^2(t)],$$

the mean power of the process.

- Thus, $S_X(\omega)$ admits an interpretation as power per unit frequency for the process $X(t)$.

* non-negative is shown in section 2.7 of the book.

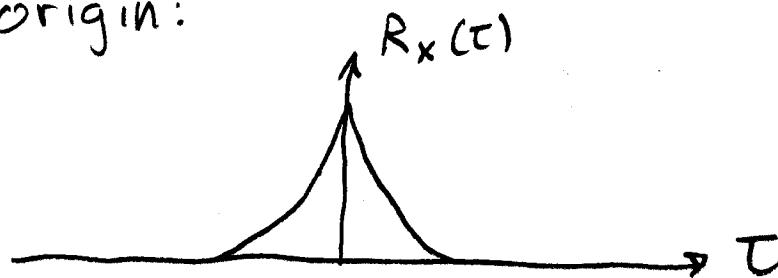
- Suppose there is a strong correlation between the RVs $X(t_0)$ and $X(t_0 + \tau_0)$ for some fixed t_0 and τ_0 .
- Because $X(t)$ is WSS, this implies that there is strong correlation between $X(t)$ and $X(t + \tau_0)$ for any t .
- In particular, there is strong correlation between $X(t + \tau_0)$ and $X((t + \tau_0) + \tau_0) = X(t + 2\tau_0)$.
- Generalizing, we see that the collection of variables $X(t + k\tau_0)$, $k \in \mathbb{Z}$ are all highly correlated with one another.
- This implies that the process $X(t)$ possesses a sort of "pseudo-periodicity" in a statistical sense:



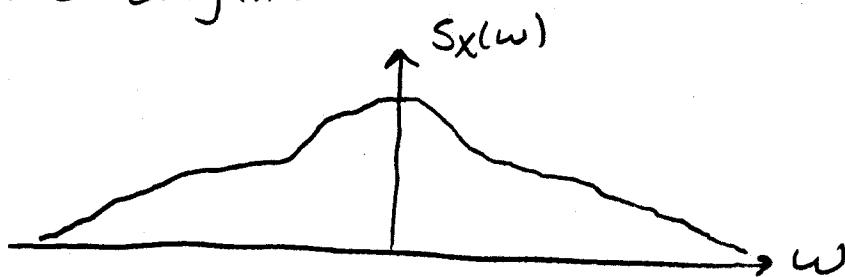
- In this case, the autocorrelation $R_x(\tau)$ will generally exhibit peaks at $R_x(k\tau_0)$, $k \in \mathbb{Z}$.
- The power spectrum $S_x(\omega)$ will generally also have a peak at $\omega_0 = \frac{2\pi}{\tau_0}$.

- If the WSS process $X(t)$ varies rapidly, i.e., there is correlation between the RVs $X(t)$ and $X(t+\tau)$ only for $|\tau|$ small,

→ Then $R_X(\tau)$ falls off rapidly away from the origin:



→ By the reciprocal spreading principle, this implies that $S_X(\omega)$ falls off slowly away from the origin:

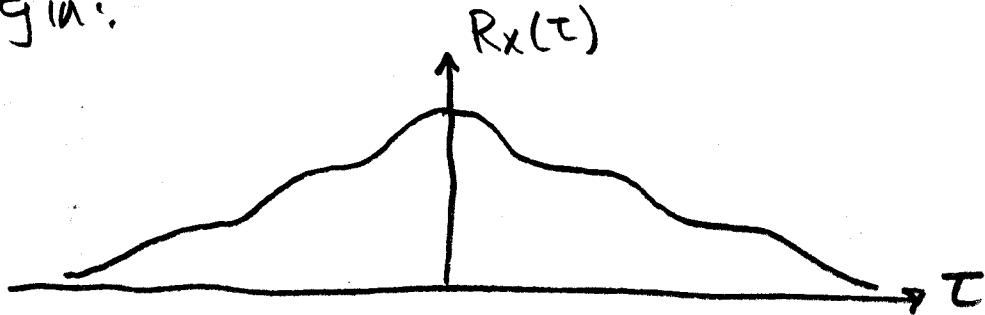


→ Interpretation: the rapidly varying process $X(t)$ has significant high frequency content.

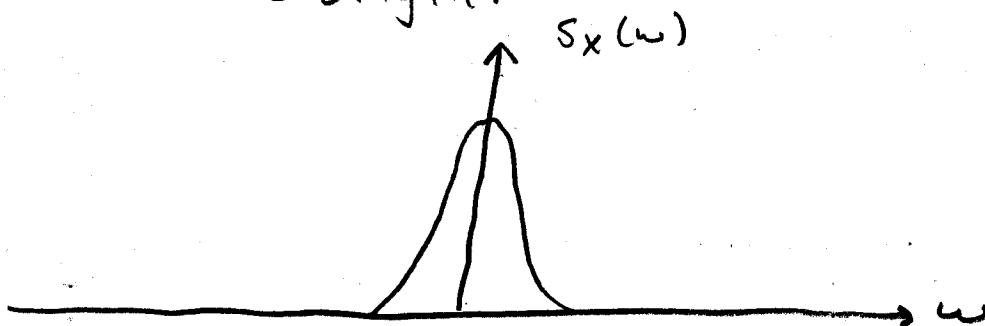
→ We say that this process has a "short correlation length".

- Likewise, if the WSS process $X(t)$ varies slowly, then there is appreciable correlation between $X(t)$ and $X(t+\tau)$ for $|\tau|$ large.

→ Then $R_X(\tau)$ falls off slowly away from the origin:



→ This implies that $S_X(\omega)$ falls off rapidly away from the origin:



→ Intuitively, this means that the process $X(t)$ has substantial low frequency content.

→ We say that this process has a "long correlation length".

- Let $X(t)$ be a WSS process.
- Truncate $X(t)$ to a time interval of length T and denote the truncated process $X_T(t)$.
- Let $X_{s,T}(t)$ be a particular realization (sample function) of $X_T(t)$.
 - Then $X_{s,T}(t)$ is a function in the usual deterministic sense.
 - The periodogram of $X_{s,T}(t)$ is given by
$$\frac{1}{T} |\mathcal{F}\{X_{s,T}(t)\}|^2.$$
- The ensemble average of the periodogram is given by
$$E\left[\frac{1}{T} |\mathcal{F}\{X_T(t)\}|^2\right]$$
- In the book, it is shown that
$$\lim_{T \rightarrow \infty} E\left[\frac{1}{T} |\mathcal{F}\{X_T(t)\}|^2\right] = \mathcal{F}\{R_X(\tau)\} = S_X(\omega).$$
- It is sometimes useful to define the PSD in terms of the Laplace transform:
$$S_X(s) = \mathcal{L}[R_X(\tau)] = \int_{-\infty}^{\infty} R_X(\tau) e^{-s\tau} d\tau$$
- $S_X(\omega)$ and $S_X(s)$ are both called "power spectral density".

- Let X_k be a WSS discrete-time process with autocorrelation $R_x(k)$.
- As in the continuous-time case, the PSD is the Fourier transform of the autocorrelation:

$$S_x(e^{j\omega}) = \sum_{k=-\infty}^{\infty} R_x(k) e^{-jk\omega}$$

$$R_x(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_x(e^{j\omega}) e^{jk\omega} d\omega$$

- The interpretation of $R_x(k)$ is analogous to that of $R_x(\tau)$.
- It is sometimes useful to define the PSD of a discrete-time WSS process in terms of the Z-transform

$$S_x(z) = \sum_{k=-\infty}^{\infty} R_x(k) z^{-k}$$

- $S_x(e^{j\omega})$ and $S_x(z)$ are both referred to as the "power spectral density" of X_k .
- The symmetry of $S_x(z)$ is

$$S_x(z) = S_x(\frac{1}{z})$$

Cross Power Spectrum

- Let $X(t)$ and $Y(t)$ be jointly WSS processes with cross correlation functions $R_{XY}(\tau)$ and $R_{YX}(\tau)$.

NOTE: $R_{YX}(\tau) = R_{XY}(-\tau)$.

- The cross power spectra of $X(t)$ and $Y(t)$ are given by

$$S_{XY}(\omega) = \mathcal{F}[R_{XY}(\tau)] = \int_{-\infty}^{\infty} R_{XY}(\tau) e^{-j\omega\tau} d\tau$$

$$S_{YX}(\omega) = \mathcal{F}[R_{YX}(\tau)] = \int_{-\infty}^{\infty} R_{YX}(\tau) e^{-j\omega\tau} d\tau$$

- The relationship between $S_{XY}(\omega)$ and $S_{YX}(\omega)$ is

$$S_{XY}(\omega) = S_{YX}^*(\omega).$$

- Analogous to the correlation coefficient of two RV's, we define the "coherence function" of the WSS processes $X(t)$ and $Y(t)$ according to

$$\gamma_{XY}^2(\omega) = \frac{|S_{XY}(\omega)|^2}{S_X(\omega) S_Y(\omega)}$$

- The coherence function may be interpreted as a frequency domain correlation coefficient for $X(t)$ and $Y(t)$.

- The magnitude of the coherence function is always less than or equal to 1.
- In the maximum correlation case, we have

$$\gamma_{xx}^2(\omega) = \frac{|S_{xx}(\omega)|^2}{S_x(\omega) S_x(\omega)} = 1$$

- The minimum correlation occurs when $X(t)$ and $Y(t)$ have a zero cross-correlation function. In this case, $\gamma_{xy}^2(\omega) = 0$.

PSD Example: Suppose $X(t)$ and $Y(t)$ are zero mean jointly WSS processes and $Z(t) = X(t) + Y(t)$.

$$\text{Then } S_z(\omega) = S_x(\omega) + S_{xy}(\omega) + S_{yx}(\omega) + S_y(\omega).$$

→ If $X(t)$ and $Y(t)$ have a zero cross-correlation function, then this reduces to

$$S_z(\omega) = S_x(\omega) + S_y(\omega).$$

- For two jointly WSS discrete-time processes X_k and Y_k , the cross spectral density is given by

$$S_{XY}(e^{j\omega}) = \mathcal{F}[R_{XY}(k)]$$

$$= \sum_{k=-\infty}^{\infty} R_{XY}(k) e^{-jk\omega k}$$

or

$$S_{XY}(z) = \mathcal{Z}[R_{XY}(k)]$$

$$= \sum_{k=-\infty}^{\infty} R_{XY}(k) z^{-k}$$

White Noise

- A continuous-time WSS process $X(t)$ is called a "white noise process" if

$$R_X(\tau) = \alpha \delta(\tau), \quad \alpha \text{ constant.}$$

- In this case, the PSD is constant:

$$S_X(\omega) = \alpha.$$

- Light containing all frequencies in equal amounts is white \Rightarrow A process containing all frequencies in equal amounts is a "white noise".

- The RV's that make up a white noise process are mutually uncorrelated for different times; e.g.,

$$R_X(\tau) = \alpha \delta(\tau).$$

→ For this reason, a white noise process is also called an "uncorrelated process".

NOTE: A process that is not white is called "colored noise"

→ light containing different amounts of different frequencies has color.

→ If a white noise is input to a linear filter, then the filter generally causes correlation in the output process.

In this sense, such a filter is often referred to as a "coloring filter" or "coloration filter".

- For a white noise $X(t)$, if the marginal probability density at each time is Gaussian, then $X(t)$ is called a "Gaussian white noise".

- If the WSS discrete-time process X_k is a sequence of zero mean mutually uncorrelated RVs, then X_k is called a (discrete-time) white noise process.
- In this case,

$$R_x(k) = \alpha \delta[k]$$

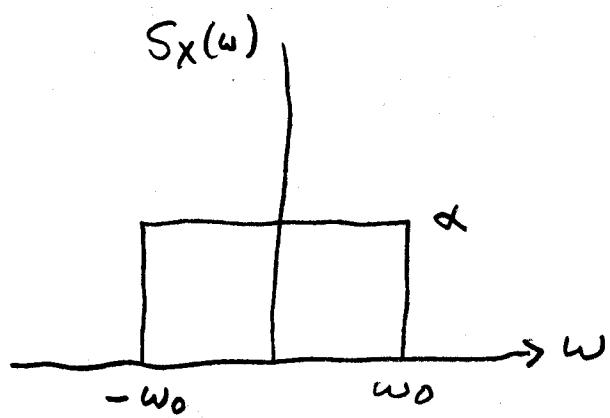
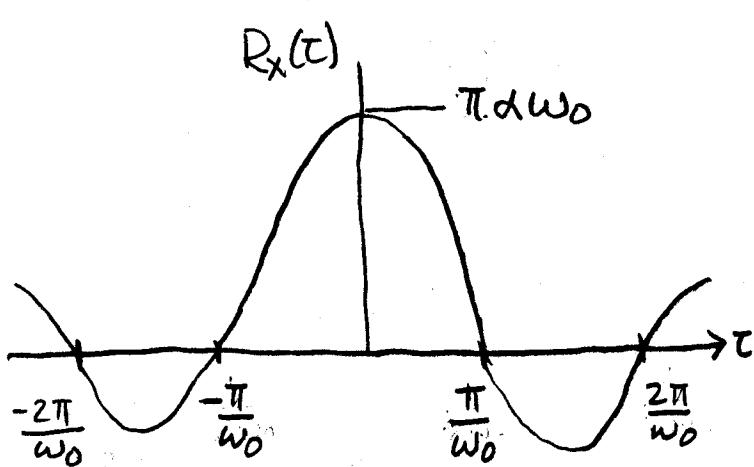
$$S_x(e^{j\omega}) = \alpha$$

Band Limited White Noise

- If $X(t)$ is a WSS process with auto correlation

$$R_x(\tau) = \alpha \omega_0 \frac{\sin \omega_0 \tau}{\pi \omega_0 \tau},$$

then the PSD is constant in the baseband interval $\omega \in [-\omega_0, \omega_0]$, and zero outside this interval:



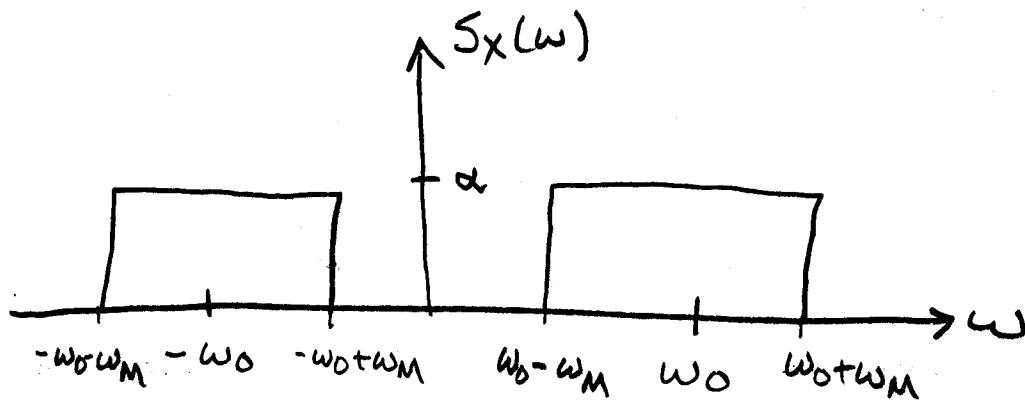
→ In this case, $X(t)$ is called "band limited white noise".

Bandpass White Noise

- If $X(t)$ is a WSS process with autocorrelation

$$R_X(\tau) = 2\alpha \omega_M \frac{\sin \omega_M \tau}{\pi \omega_M \tau} \cos \omega_0 \tau$$

Then the PSD $S_X(\omega)$ is constant in a pair of passbands symmetrically located about the frequency origin:



→ In this case, the process $X(t)$ is called "Bandpass White Noise".

IID Process

- If the RV's comprising the process $X(t)$ or X_k are all mutually independent and all have the same pdf, then the process is called an "independent, identically distributed" process or "IID" process.
- IID implies zero mean; otherwise there would be nonzero correlation between the RV's at different times.
- IID implies that the process is a white noise.
- IID implies strict sense stationarity (SSS).

Gaussian Process

- A random process is called "Gaussian" if any finite set of RV's selected from the process are jointly Gaussian.
- The joint density is then completely specified by the collection of first and second order moments.

Markov Process

- Let X_t be a stochastic process.
- Pick a finite number of ordered time indices from the indexing set $\mathcal{I} = \mathbb{R}$:

$$t_1 < t_2 < \dots < t_{n+1}$$

- Associated with this set of times is a finite collection of RV's $X_{t_1}, X_{t_2}, \dots, X_{t_{n+1}}$.
- The process X_t is called "Markovian" or a "Markov Process" if

$$f_{X_{t_{n+1}} | X_{t_n} \dots X_{t_1}}(x_{t_{n+1}} | x_{t_n}, \dots, x_{t_1}) = f_{X_{t_{n+1}} | X_{t_n}}(x_{t_{n+1}} | x_{t_n}).$$

→ This means that the conditional density of $X_{t_{n+1}}$ conditioned on a series of prior realizations $X_{t_1} = x_{t_1}, \dots, X_{t_n} = x_{t_n}$ is the same as the conditional density of $X_{t_{n+1}}$ conditioned on $X_{t_n} = x_{t_n}$ alone.

→ All of the information about $X_{t_{n+1}}$ that is contained in X_{t_1}, \dots, X_{t_n} is captured by X_{t_n} itself.

- For a particular sample function, let the realizations of the RV's be $X_{t_1} = x_{t_1}, \dots, X_{t_{n+1}} = x_{t_{n+1}}$.

→ We often say that the process is in the "state" x_{t_k} at time t_k . We call the change

from $X_{t_k} = x_{t_k}$ to $X_{t_{k+1}} = x_{t_{k+1}}$ a

"state transition" of the process.

- A Markov process is completely characterized by an initial marginal density

$$f_{X_{t_0}}(x_{t_0})$$

and the set of "transition probabilities"

$$f_{X_{t_{k+1}}|X_{t_k}}(x_{t_{k+1}}|x_{t_k}), \quad t_k < t_{k+1}.$$

- The transition probabilities are conditional pdf's that must satisfy the "Chapman-Kolmogorov" equation

$$f_{X_{t_3}|X_{t_1}}(x_{t_3}|x_{t_1}) = \int_{-\infty}^{\infty} f_{X_{t_3}|X_{t_2}}(x_{t_3}|x_{t_2}) f_{X_{t_2}|X_{t_1}}(x_{t_2}|x_{t_1}) dx_{t_2},$$

$$t_1 < t_2 < t_3.$$

→ This is a sort of transitivity property.

- The definition of a discrete-time Markov process is completely analogous.

Gauss-Markov Process

- A Gauss-Markov process is both Gaussian and Markovian.
- A Gaussian process is Gauss-Markov if and only if the covariances of the involved RVS/ all satisfy the separability condition

$$\text{cov}(X_{t_3}, X_{t_1}) = \frac{\text{cov}(X_{t_3}, X_{t_2}) \text{cov}(X_{t_2}, X_{t_1})}{\sigma_{X_{t_2}}^2}$$

$\forall t_1 < t_2 < t_3.$

- The book considers only WSS Gauss-Markov processes.

→ In this case, the autocorrelation is exponential:

$$R_x(\tau) = \sigma^2 e^{-\beta|\tau|}$$

→ The PSD takes the form

$$S_x(\omega) = \frac{2\sigma^2 \beta}{\omega^2 + \beta^2} \quad S_x(s) = \frac{2\sigma^2 \beta}{-s^2 + \beta^2}$$

where σ and β are parameters that completely specify the process.

- Study the Gauss-Markov and Markov examples in Ex. 2.11 and section 2.11 of the book.
- Study the narrowband Gaussian process example in section 2.12 of the book.
- Study the Brownian-Motion process described in section 2.13 of the book.
- Experimental determination of the autocorrelation and PSD is discussed in section 2.15 of the book.

Stochastic Differentiation

- Let $X(t)$ be a stochastic process and $X_s(t)$, $s \in S$, be a sample function.
 - Clearly, $X_s(t)$ has a derivative in the usual sense:
- $$\dot{X}_s(t) = \frac{d}{dt} X_s(t) = \lim_{\varepsilon \rightarrow 0} \frac{X_s(t+\varepsilon) - X_s(t)}{\varepsilon}$$
- We can also define the derivative $\dot{X}(t)$ of the process $X(t)$ itself. It is the stochastic process $\dot{X}(t)$ satisfying

$$\lim_{\varepsilon \rightarrow 0} E \left\{ \left[\frac{X(t+\varepsilon) - X(t)}{\varepsilon} - \dot{X}(t) \right]^2 \right\} = 0.$$

→ This is known as "mean-square" differentiability.

- Properties of the m.s. derivative:

$$E[\dot{x}(t)] = \frac{d}{dt} E[x(t)]$$

$$R_{x\dot{x}}(t_1, t_2) = \frac{\partial R_{xx}(t_1, t_2)}{\partial t_2}$$

$$R_{\dot{x}\dot{x}}(t_1, t_2) = \frac{\partial R_{xx}(t_1, t_2)}{\partial t_1}$$

$$R_{\ddot{x}\ddot{x}}(t_1, t_2) = \frac{\partial^2 R_{xx}(t_1, t_2)}{\partial t_1 \partial t_2}$$

$$E\left[\frac{d^n}{dt^n} X(t)\right] = \frac{d^n}{dt^n} E[X(t)]$$

$$R_{x^{(n)}x^{(m)}}(t_1, t_2) = \frac{\partial^{n+m} R_{xx}(t_1, t_2)}{\partial t_1^n \partial t_2^m}$$

- For a WSS process, these reduce to

$$R_{x\dot{x}}(\tau) = -\frac{d}{d\tau} R_x(\tau)$$

$$R_{\dot{x}\dot{x}}(\tau) = \frac{d}{d\tau} R_{x\dot{x}}(\tau) = -\frac{d^2}{d\tau^2} R_x(\tau)$$

$$R_{x^{(n)}}(\tau) = (-1)^n \frac{d^{2n}}{d\tau^{2n}} R_x(\tau)$$

$$E[\dot{x}(t)] = \frac{d}{dt} E[x(t)] = 0$$

$$E\left[\frac{d^n}{dt^n} X(t)\right] = \frac{d^n}{dt^n} E[X(t)] = 0$$

- If $X(t)$ and $Y(t)$ are stochastic processes, then

$$R_{X^{(n)} Y^{(m)}}(t_1, t_2) = \frac{d^{n+m}}{\partial t_1^n \partial t_2^m} R_{XY}(t_1, t_2)$$

→ In the case where $X(t)$ and $Y(t)$ are jointly WSS, this reduces to

$$R_{X^{(n)} Y^{(m)}}(\tau) = (-1)^m \frac{d^{n+m}}{d\tau^{n+m}} R_{XY}(\tau).$$

Stochastic Integration

- As with the derivative, sample functions $X_s(t)$ of a process $X(t)$ may be integrated in the usual sense:

$$Z_s = \int_a^b X_s(t) dt.$$

→ Considered over the ensemble $\{X_s(t)\}_{s \in S}$,

Z_s is an ordinary RV.

- We can also consider the mean-square integral of the process $X(t)$ itself. By definition, it is the RV Z satisfying

$$\lim_{\Delta t_i \rightarrow 0} E \left\{ [Z - \sum_i X(t_i) \Delta t_i]^2 \right\} = 0.$$

- A process $X(t)$ is m.s. integrable if

$$\int_a^b \int_a^b |R_X(t_1, t_2)| dt_1 dt_2 < \infty. \text{ PAGE 2.35}$$