Part 1 Random Processes for Communications

System Models

- A good mathematical model for a system is the basis of its analysis.
- Two models are often considered:
  - Deterministic model
    - No uncertainty about its time-dependent behavior at any instance of time
  - Random or stochastic model
    - Uncertain about its time-dependent behavior at any instance of time
    - but certain on the *statistical* behavior at any instance of time

Examples of Stochastic Models

- Channel noise and interference
- Source of information, such as voice

Notion of Relative Frequency

- How to determine the probability of "head appearance" for a coin?
- *Answer:* Relative frequency.

  Specifically, by carrying out n coin-tossing experiments, the *relative frequency* of head appearance is equal to $N_n(A)/n$, where $N_n(A)$ is the number of head appearance in these n random experiments.

Notion of Relative Frequency

- Is *relative frequency* close to the *true probability* (of head appearance)?
  - It could occur that **4-out-of-10** tossing results are “head” for a **fair** coin!
- Can one guarantee that the true “head appearance probability” remains *unchanged* (i.e., *time-invariant*) in each experiment performed at different time instance?

Notion of Relative Frequency

- Similarly, the previous question can be extended to "In a communication system, can we estimate the noise by *repetitive measurements* at *consecutive but different* time instance?"

- Some *assumptions* on the statistical models are necessary!

Conditional Probability

- Definition of conditional probability

$$P(B \mid A) \quad \left(\approx \frac{N_n(A \cap B)}{N_n(A)} \right) = \frac{P(A \cap B)}{P(A)}$$

- Independence of events $P(B \mid A) = P(B)$
  - A knowledge of occurrence of event A tells us no more about the probability of occurrence of event B than we knew without this knowledge.
  - Hence, they are *statistically* independent.

Random Variable

- A non-negative function $f_X(x)$ satisfies

$$F_X(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f_X(t)\, dt$$

is called the *probability density function* (pdf) of random variable X.

- If the pdf of X exists, then

$$f_X(x) = \frac{\partial F_X(x)}{\partial x}$$

Random Variable

- It is not necessarily true that
  - If

$$f_X(x) = \frac{\partial F_X(x)}{\partial x},$$

    then the pdf of X exists and equals $f_X(x)$.

Random Vector

- If its joint density $f_{X,Y}(x,y)$ exists, then

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

where $F_{X,Y}(x,y) = \Pr[X \leq x \text{ and } Y \leq y]$

- The conditional density of Y given that $[X = x]$ is

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

provided that $f_X(x) \neq 0$.

Random Process

- Random process: An extension of *multi-dimensional* random vectors
  - Representation of two-dimensional random vector
    - $(X,Y) = (X(1), X(2)) = \{X(j), j \in I\}$, where the index set I equals $\{1, 2\}$.
  - Representation of m-dimensional random vector
    - $\{X(j), j \in I\}$, where the index set I equals $\{1, 2, \ldots, m\}$.

Random Process

- How about $\{X(t), t \in \Re\}$?
  - It is no longer a random vector since the index set is continuous!
  - This is a suitable model for, e.g., a noise because a noise often exists continuously in time.

Stationarity

- The statistical property of a random process encountered in real world is often *independent* of the time at which the observation (or experiment) is initiated.
- Mathematically, this can be formulated as that for any t_1, t_2, …, t_k and τ:

$$F_{X(t_1+\tau),X(t_2+\tau),\dots,X(t_k+\tau)}(x_1,x_2,\dots,x_k)$$
$$=F_{X(t_1),X(t_2),\dots,X(t_k)}(x_1,x_2,\dots,x_k)$$

Stationarity

- Why introducing "stationarity?"
  - With stationarity, we can be certain that the observations made at different instances of time have the same distributions!
  - For example, $X(0), X(T), X(2T), X(3T), \ldots$.
  - Suppose that $\Pr[X(0) = 0] = \Pr[X(0)=1] = 1/2$. Can we guarantee that the *relative frequency* of "1's appearance" for experiments performed at **several** different instances of time approach ½ by stationarity? No, we need an additional assumption!

Mean Function

- The mean of a random process $X(t)$ at time t is equal to:

$$\mu_X(t) = E[X(t)] = \int_{-\infty}^{\infty} x \cdot f_{X(t)}(x)dx$$

where $f_{X(t)}(\cdot)$ is the pdf of $X(t)$ at time t.

- If $X(t)$ is *stationary*, $\mu_X(t)$ is a constant for all t.

Autocorrelation

- The autocorrelation function of a (possibly complex) random process $X(t)$ is given by:

$$
\begin{aligned}
R_X(t_1, t_2) &= E[X(t_1)X^*(t_2)] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2^* f_{X(t_1),X(t_2)}(x_1, x_2)\, dx_1 dx_2
\end{aligned}
$$

- If $X(t)$ is *stationary*, the autocorrelation function $R_X(t_1, t_2)$ is equal to $R_X(t_1 - t_2, 0)$.

Autocorrelation

$$\begin{aligned}
R_X(t_1,t_2) &= E[X(t_1)X^*(t_2)] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2^* f_{X(t_1),X(t_2)}(x_1,x_2)\,dx_1 dx_2 \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2^* f_{X(t_1-t_2),X(0)}(x_1,x_2)\,dx_1 dx_2 \\
&= E[X(t_1-t_2)X^*(0)] \\
&= R_X(t_1-t_2,0) \\
&= \underline{R_X(t_1-t_2)}
\end{aligned}$$

A short-hand for autocorrelation function of a stationary process

Autocorrelation

- Conceptually,
  - Autocorrelation function = “power correlation” between two time instances t_1 and t_2.
  - “Variance” is the degree of variation to the standard value (i.e., mean).

Autocovariance

$$\begin{aligned}
C_X(t_1,t_2) &= E[(X(t_1)-\mu_X(t_1))(X(t_2)-\mu_X(t_2))^*] \\
&= E[X(t_1)X^*(t_2) - X(t_1)\mu_X^*(t_2) \\
&\quad -\mu_X(t_1)X^*(t_2) + \mu_X(t_1)\mu_X^*(t_2)] \\
&= E[X(t_1)X^*(t_2)] - E[X(t_1)]\mu_X^*(t_2) \\
&\quad -\mu_X(t_1)E[X^*(t_2)] + \mu_X(t_1)\mu_X^*(t_2) \\
&= R_X(t_1,t_2) - \mu_X(t_1)\mu_X^*(t_2) \\
&\quad -\mu_X(t_1)\mu_X^*(t_2) + \mu_X(t_1)\mu_X^*(t_2) \\
&= R_X(t_1,t_2) - \mu_X(t_1)\mu_X^*(t_2)
\end{aligned}$$

Autocovariance

- If $X(t)$ is *stationary*, $C_X(t_1, t_2)$ becomes

$$
\begin{aligned}
C_X(t_1, t_2) &= R_X(t_1, t_2) - \mu_X(t_1)\mu_X^*(t_2) \\
&= R_X(t_1 - t_2, 0) - |\mu_X|^2 \\
&= C_X(t_1 - t_2, 0) \\
&= \underline{C_X(t_1 - t_2)}
\end{aligned}
$$

Wide-Sense Stationary (WSS)

- Since in most cases of practical interest, only the first two moments (i.e., $\mu_X(t)$ and $C_X(t_1, t_2)$) are concerned, an alternative definition of stationarity is introduced.
- Definition (**Wide-Sense Stationarity**) A random process $X(t)$ is WSS if

$$\begin{cases} \mu_X(t) = \text{constant}; \\ C_X(t_1, t_2) = C_X(t_1 - t_2) \end{cases} \quad \text{or} \quad \begin{cases} \mu_X(t) = \text{constant}; \\ R_X(t_1, t_2) = R_X(t_1 - t_2). \end{cases}$$

Wide-Sense Stationary (WSS)

- Alternative names for WSS
  - weakly stationary
  - stationary in the weak sense
  - second-order stationary

- If the first two moments of a random process exist (i.e., are finite), then *strictly stationary* implies *weakly stationary* (but not vice versa).

Cyclostationarity

- Definition (**Cyclostationarity**) A random process $X(t)$ is cyclostationary if there exists a constant T such that

$$\begin{cases} \mu_X(t+T) = \mu_X(t); \\ C_X(t_1+T, t_2+T) = C_X(t_1, t_2). \end{cases}$$

Properties of Autocorrelation Function for WSS Random Process

1. **Mean Square Value:** $R_X(0) = E[|X(t)|^2]$
2. **Conjugate Symmetry:**

$$R_X(\tau) = R_X^*(-\tau)$$

- Recall that autocorrelation function = "power correlation" between two time instances t_1 and t_2.
- For a WSS process, this "power correlation" only depends on time difference.
- Hence, we only need to deal with $R_X(\tau)$ here.

Properties of Autocorrelation Function for WSS Random Process

3. Real Part Peaks at zero: $|\mathrm{Re}\{R_X(\tau)\}| \leqq R_X(0)$

Proof:

$$\begin{aligned}
0 &\leq E\left[|X(t+\tau) \pm X(t)|^2\right] \\
&= E[|X(t+\tau)|^2] + E[|X(t)|^2] \pm E[X(t+\tau)X^*(t)] \pm E[X(t)X^*(t+\tau)] \\
&= R_X(0) + R_X(0) \pm R_X(\tau) \pm R_X(-\tau) \quad (R_X(-\tau) = R_X^*(\tau)) \\
&= 2R_X(0) \pm 2\mathrm{Re}\{R_X(\tau)\}
\end{aligned}$$

Hence, $-R_X(0) \leq \mathrm{Re}\{R_X(\tau)\} \leq R_X(0)$

with equality holding when

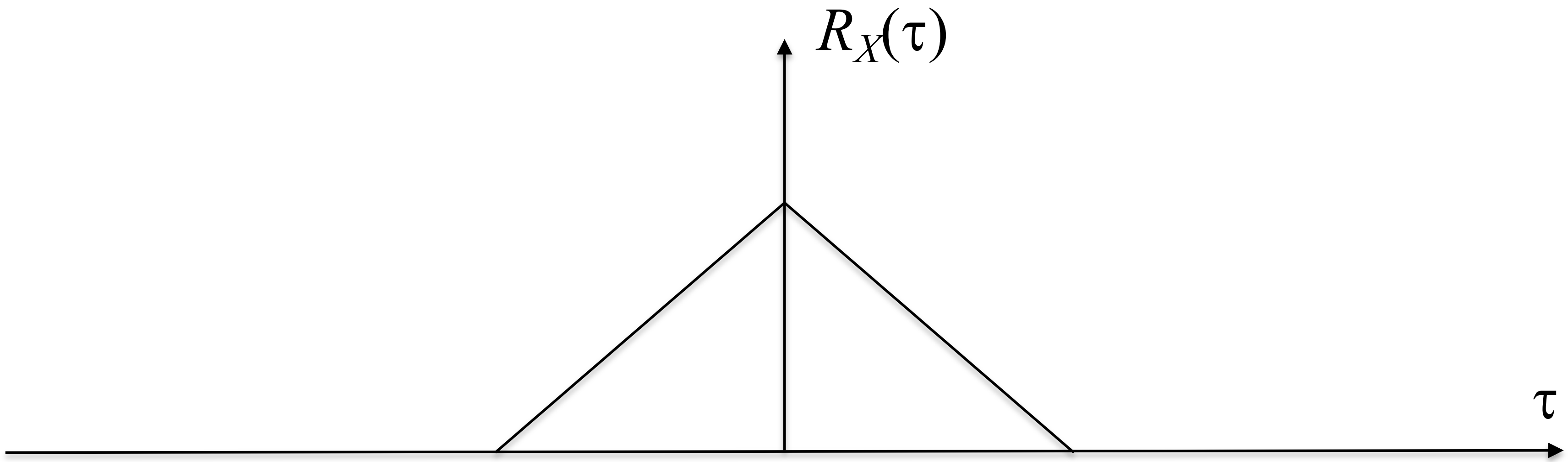
$$\Pr[X(t+\tau) = X(t)] = \Pr[X(\tau) = X(0)] = 1$$
$$\text{or } \Pr[X(t+\tau) = -X(t)] = \Pr[X(\tau) = -X(0)] = 1$$

Properties of Autocorrelation Function for WSS Random Process

- Operational meaning of autocorrelation function:
  - The "power" correlation of a random process at τ seconds apart.
  - The smaller $R_X(\tau)$ is, the less the correlation between $X(t)$ and $X(t+\tau)$.
    - Here, we assume $X(t)$ is a real-valued random process.

Properties of Autocorrelation Function for WSS Random Process

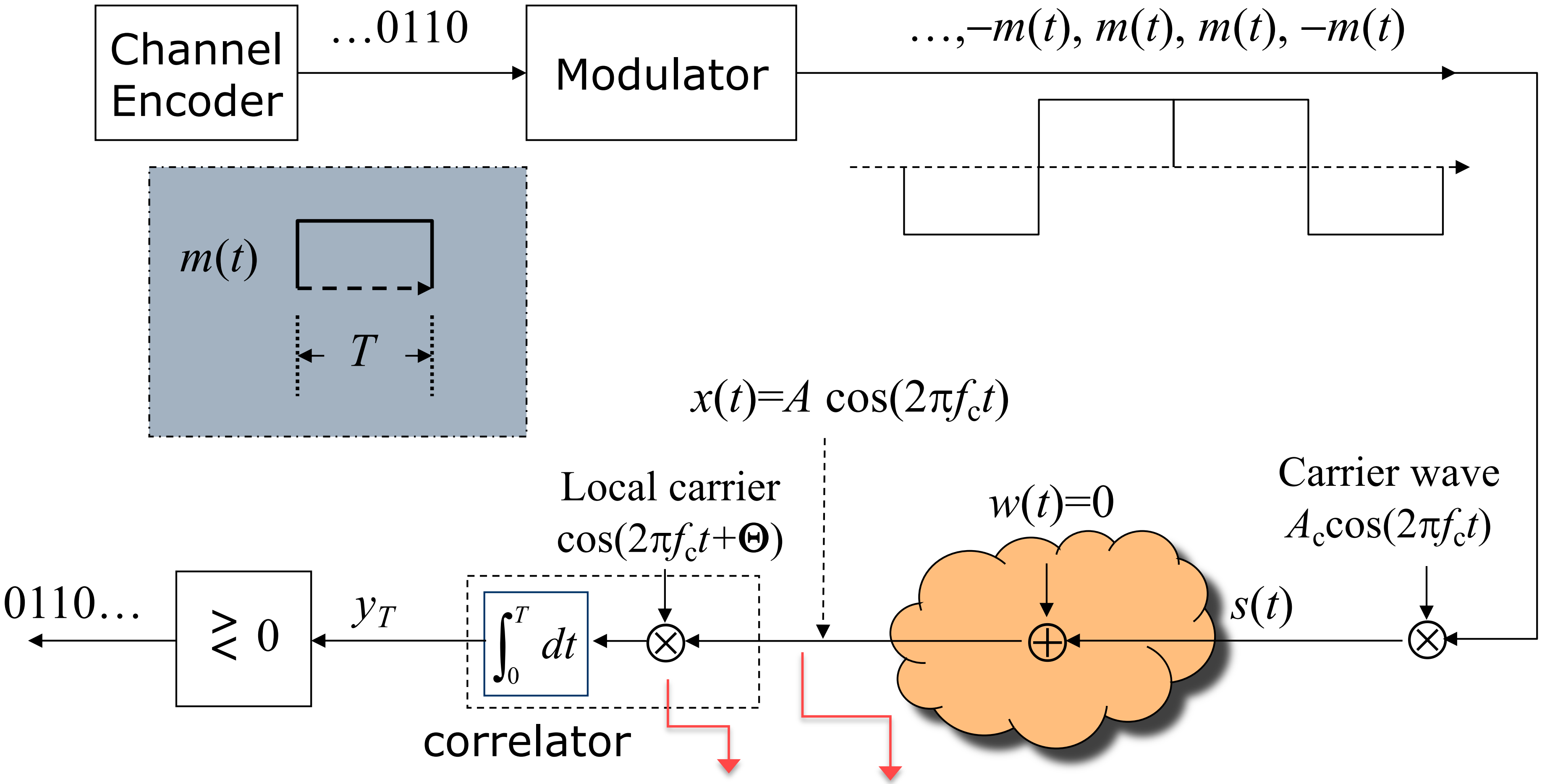
- If $R_X(\tau)$ decreases faster, the correlation decreases faster.



Example: Signal with Random Phase

- Let $X(t) = A\cos(2\pi f_c t + \Theta)$, where Θ is uniformly distributed over $[-\pi, \pi)$.
  - *Application*: A local carrier at the receiver side may have a random "phase difference" with respect to the phase of the carrier at the transmitter side.

Example: Signal with Random Phase



An equivalent view: | Local carrier $\cos(2\pi f_c t)$ | $X(t)=A\cos(2\pi f_c t+\Theta)$

Example: Signal with Random Phase

Then

$$
\begin{aligned}
\mu_X(t) &= E[A\cos(2\pi f_c t + \Theta)] \\
&= \int_{-\pi}^{\pi} A\cos(2\pi f_c t + \theta)\frac{1}{2\pi}\,d\theta \\
&= \frac{A}{2\pi}\int_{-\pi}^{\pi} \cos(\theta + 2\pi f_c t)\,d\theta \\
&= \frac{A}{2\pi}\left(\sin(\theta + 2\pi f_c t)\right)\Big|_{-\pi}^{\pi} \\
&= \frac{A}{2\pi}\left(\sin(\pi + 2\pi f_c t) - \sin(-\pi + 2\pi f_c t)\right) \\
&= 0.
\end{aligned}
$$

Example: Signal with Random Phase

$$R_X(t_1,t_2) = E[A\cos(2\pi f_c t_1 + \Theta)\cdot A\cos(2\pi f_c t_2 + \Theta)]$$

$$= A^2 \int_{-\pi}^{\pi} \cos(\theta + 2\pi f_c t_1)\cos(\theta + 2\pi f_c t_2)\frac{1}{2\pi}d\theta$$

$$= \frac{A^2}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\big(\cos[(\theta+2\pi f_c t_1)+(\theta+2\pi f_c t_2)]$$

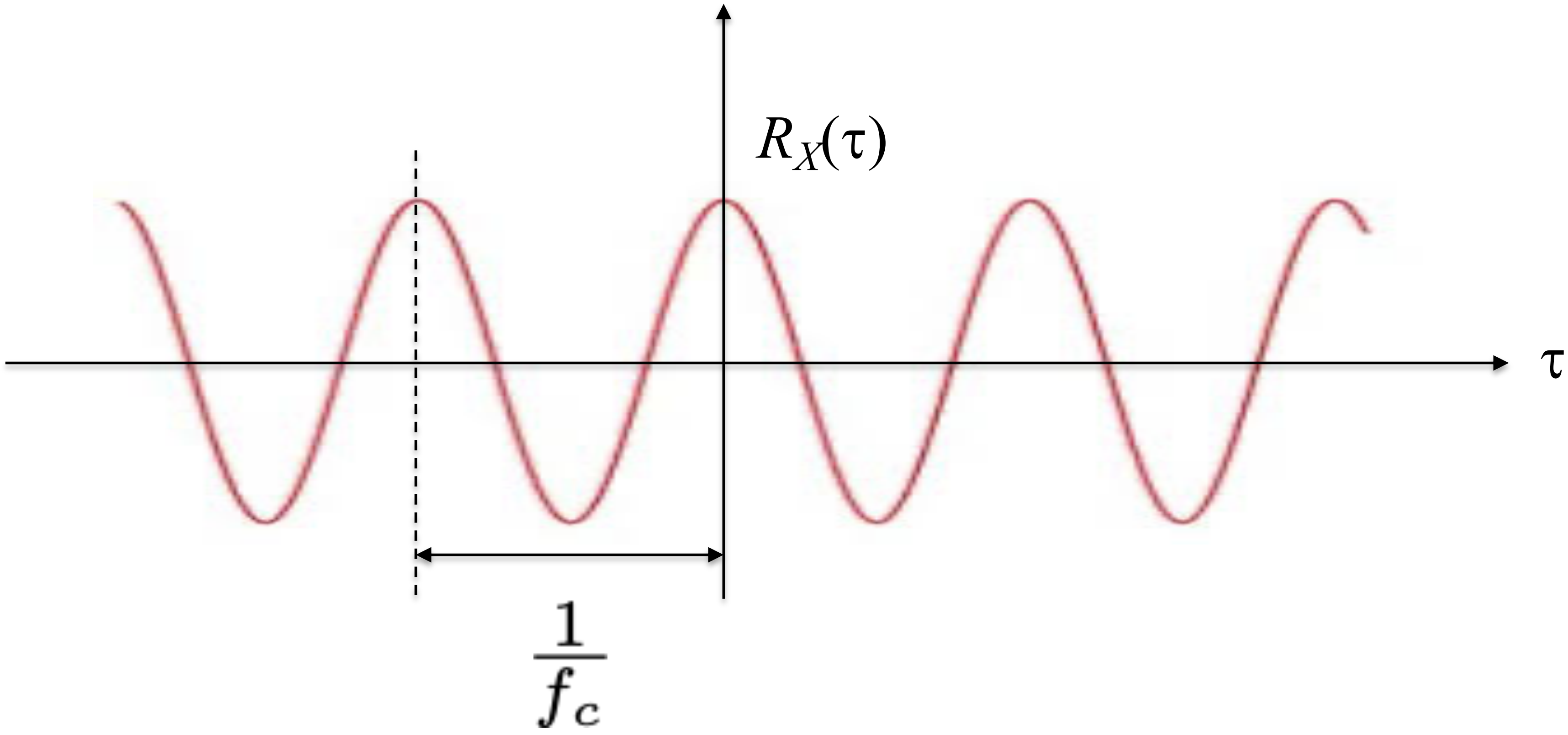
$$+\cos[(\theta+2\pi f_c t_1)-(\theta+2\pi f_c t_2)]\big)d\theta$$

$$= \frac{A^2}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\Big(\underbrace{\cancel{\cos(2\theta+2\pi f_c(t_1+t_2))}}_{=0}+\cos(2\pi f_c(t_1-t_2))\Big)d\theta$$

$$= \frac{A^2}{2}\cos(2\pi f_c(t_1-t_2)).$$

Hence, $X(t)$ is WSS.

Example: Signal with Random Phase



Example: Signal with Random Delay

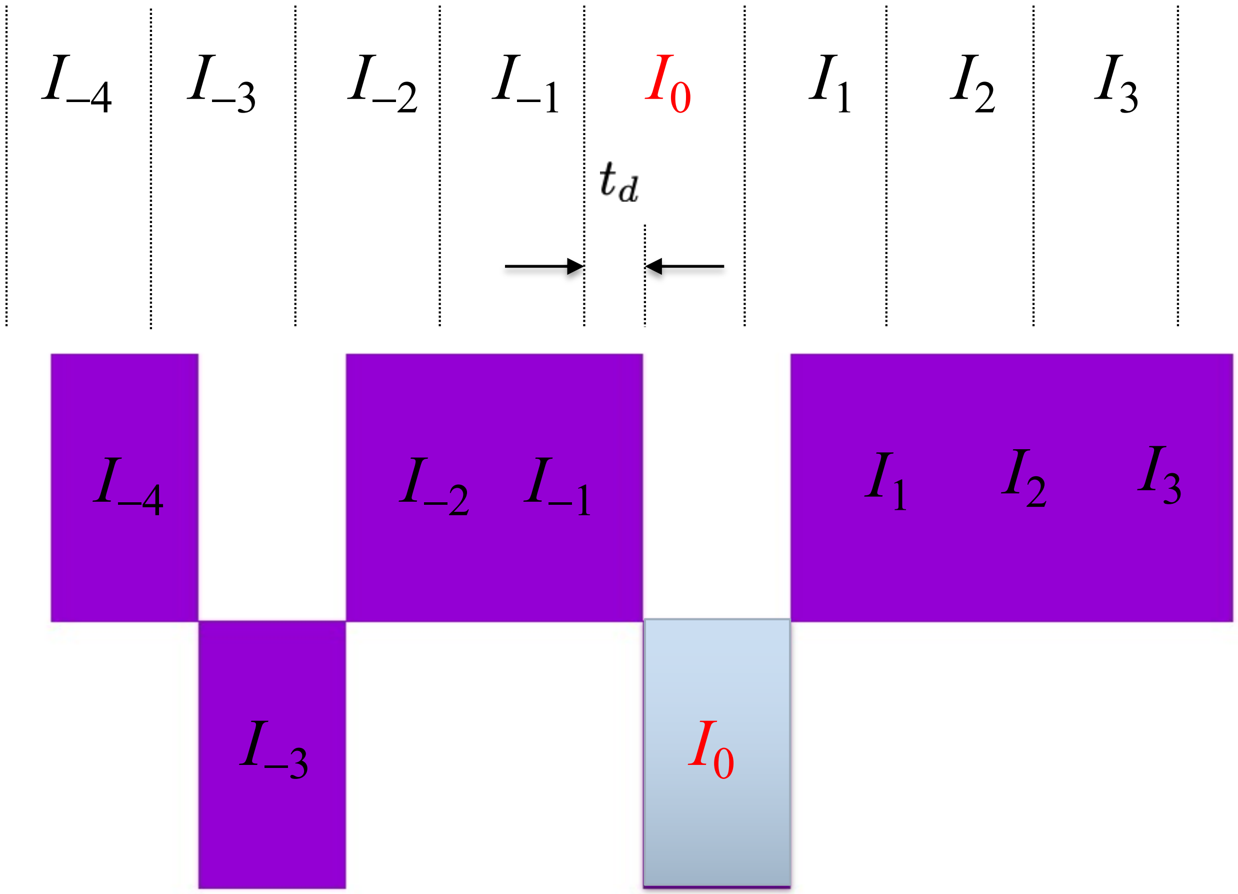
- Let

$$X(t)=\sum_{n=-\infty}^{\infty} A\cdot I_n \cdot p(t-nT-t_d)$$

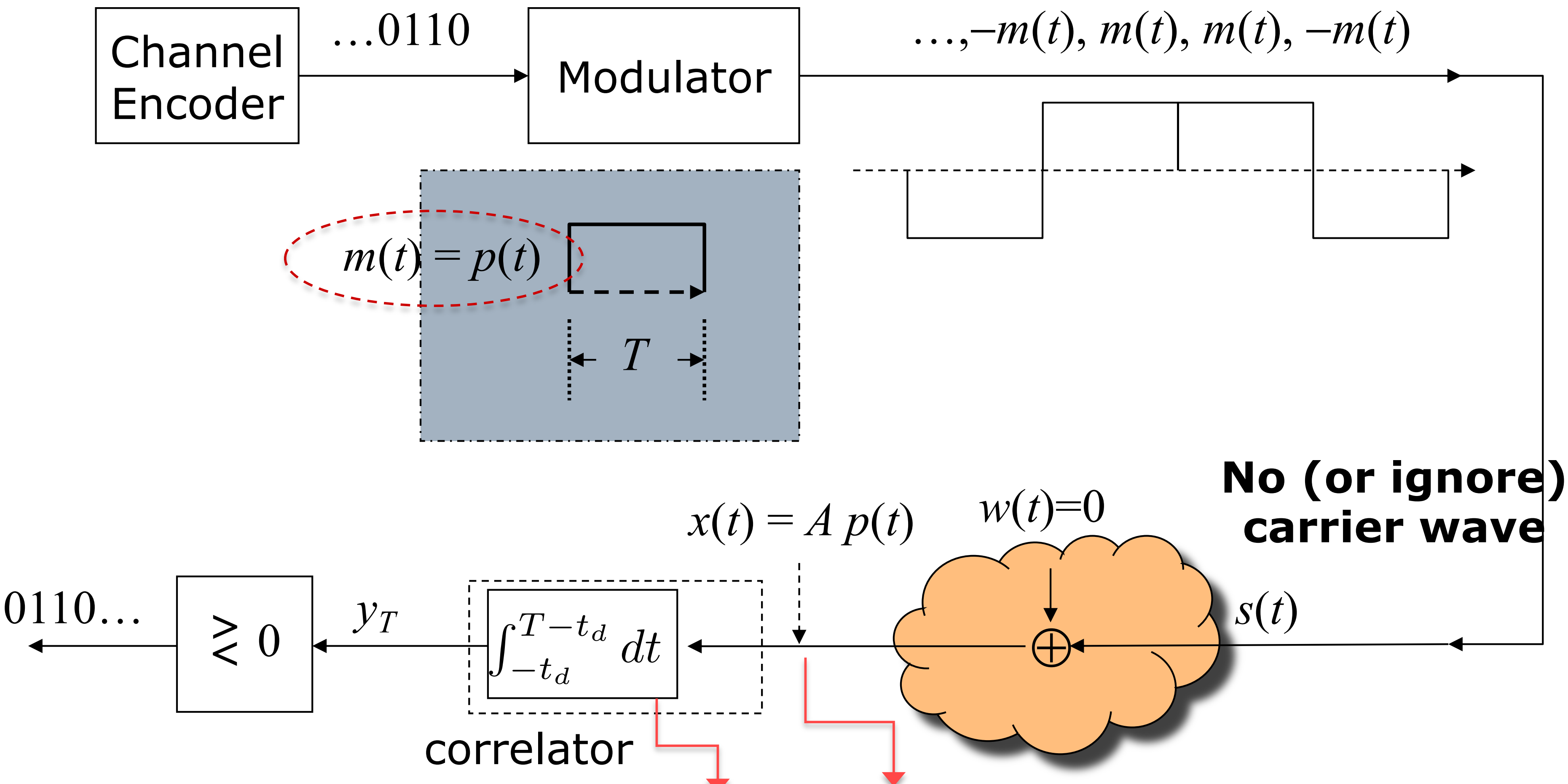
where …, I_{-2}, I_{-1}, I_0, I_1, I_2, … are independent, and each I_j is either -1 or $+1$ with equal probability, and

$$p(t)=\begin{cases}1, & 0\le t<T\\ 0, & \text{otherwise}\end{cases}$$

Example: Signal with Random Delay



Example: Signal with Random Delay



An equivalent view: $\int_0^T dt$ $X(t) = A\,p(t-t_d)$

Example: Signal with Random Delay

- By assuming that t_d is uniformly distributed over $[0, T)$, we obtain:

$$
\begin{aligned}
\mu_X(t) &= E\left[\sum_{n=-\infty}^{\infty} A \cdot I_n \cdot p(t - nT - t_d)\right] \\
&= \sum_{n=-\infty}^{\infty} A \cdot E[I_n] \cdot E[p(t - nT - t_d)] \\
&= \sum_{n=-\infty}^{\infty} A \cdot 0 \cdot E[p(t - nT - t_d)] \\
&= 0
\end{aligned}
$$

Example: Signal with Random Delay

- A useful probabilistic rule: $\boxed{E[X] = E[E[X|Y]]}$

  So, we have:

$$E[X(t_1)X(t_2)] = E\big[E\big[X(t_1)X(t_2)\big|t_d\big]\big]$$

Note: $\begin{cases} E[X|Y] = \int_{\mathcal{X}} x\, f_{X|Y}(x|y)dx = g(y) \\ E[E[X|Y]] = \int_{\mathcal{Y}} g(y)\, f_Y(y)dy \end{cases}$

$$E\left[X(t_1)X(t_2)\middle|t_d\right]$$

$$= E\left[\left(\sum_{n=-\infty}^{\infty} A\cdot I_n \cdot p(t_1 - nT - t_d)\right)\left(\sum_{m=-\infty}^{\infty} A\cdot I_m \cdot p(t_2 - mT - t_d)\right)\middle| t_d\right]$$

$$= A^2 \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} E[I_n I_m \mid t_d] E[p(t_1 - nT - t_d)p(t_2 - mT - t_d) \mid t_d]$$

$$= A^2 \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} E[I_n I_m] p(t_1 - nT - t_d)p(t_2 - mT - t_d)$$

$$= A^2 \sum_{n=-\infty}^{\infty} E[I_n^2] p(t_1 - nT - t_d)p(t_2 - nT - t_d)$$

$$= A^2 \sum_{n=-\infty}^{\infty} p(t_1 - nT - t_d)p(t_2 - nT - t_d)$$

Since $E[I_n I_m] = E[I_n]E[I_m] = 0$ for $n \neq m$.

Among $-\infty < n < \infty$, there is at most one n that can make

$$p(t_1 - nT - t_d)p(t_2 - nT - t_d) = 1.$$

Without loss of generality, we let $t_1 = mT + \xi$ and $t_2 = t_1 - \tau$, where m is an integer and $0 \le \xi < T$.

$$\begin{aligned}
& p(t_1 - nT - t_d) = p(t_2 - nT - t_d) = 1 \\
\Leftrightarrow\quad & 0 \le t_1 - nT - t_d < T \text{ and } 0 \le t_2 - nT - t_d < T \\
\Leftrightarrow\quad & \frac{t_1 - t_d}{T} - 1 < n \le \frac{t_1 - t_d}{T} \text{ and } \frac{t_2 - t_d}{T} - 1 < n \le \frac{t_2 - t_d}{T} \\
\Leftrightarrow\quad & \left\lfloor \frac{t_1 - t_d}{T} \right\rfloor = \left\lfloor \frac{t_2 - t_d}{T} \right\rfloor \\
\Leftrightarrow\quad & \left\lfloor \frac{mT + \xi - t_d}{T} \right\rfloor = \left\lfloor \frac{mT + \xi - \tau - t_d}{T} \right\rfloor
\end{aligned}$$

$$\iff \quad m + \left\lfloor \frac{\xi - t_d}{T} \right\rfloor = m + \left\lfloor \frac{\xi - \tau - t_d}{T} \right\rfloor$$

$$\iff \quad \left\lfloor \frac{\xi - t_d}{T} \right\rfloor = \left\lfloor \frac{\xi - \tau - t_d}{T} \right\rfloor$$

$$\iff \quad \left\lfloor \frac{\xi - t_d}{T} \right\rfloor = \left\lfloor \frac{\xi - \tau - t_d}{T} \right\rfloor = -1 \text{ or } 0$$

(Note that $\left\lfloor \frac{\xi - t_d}{T} \right\rfloor$ can only be either -1 or 0 since $0 \leq \xi, t_d < T$.)

$$\iff \quad (\xi < t_d \leq \xi + T \text{ and } \xi - \tau < t_d \leq \xi - \tau + T)$$
$$\text{or } (\xi - T < t_d \leq \xi \text{ and } \xi - \tau - T < t_d \leq \xi - \tau)$$

$$\iff \quad (\xi < t_d \leq \xi - \tau + T) \quad \text{or} \quad (\xi - T < t_d \leq \xi - \tau) \quad \text{for} \quad 0 \leq \tau < T$$

$$\iff \quad \begin{cases} (\xi < t_d < T) \quad \text{or} \quad (0 \leq t_d \leq \xi - \tau), & 0 \leq \tau \leq \xi; \\ \xi < t_d \leq \xi - \tau + T, & \xi < \tau < T, \end{cases}$$

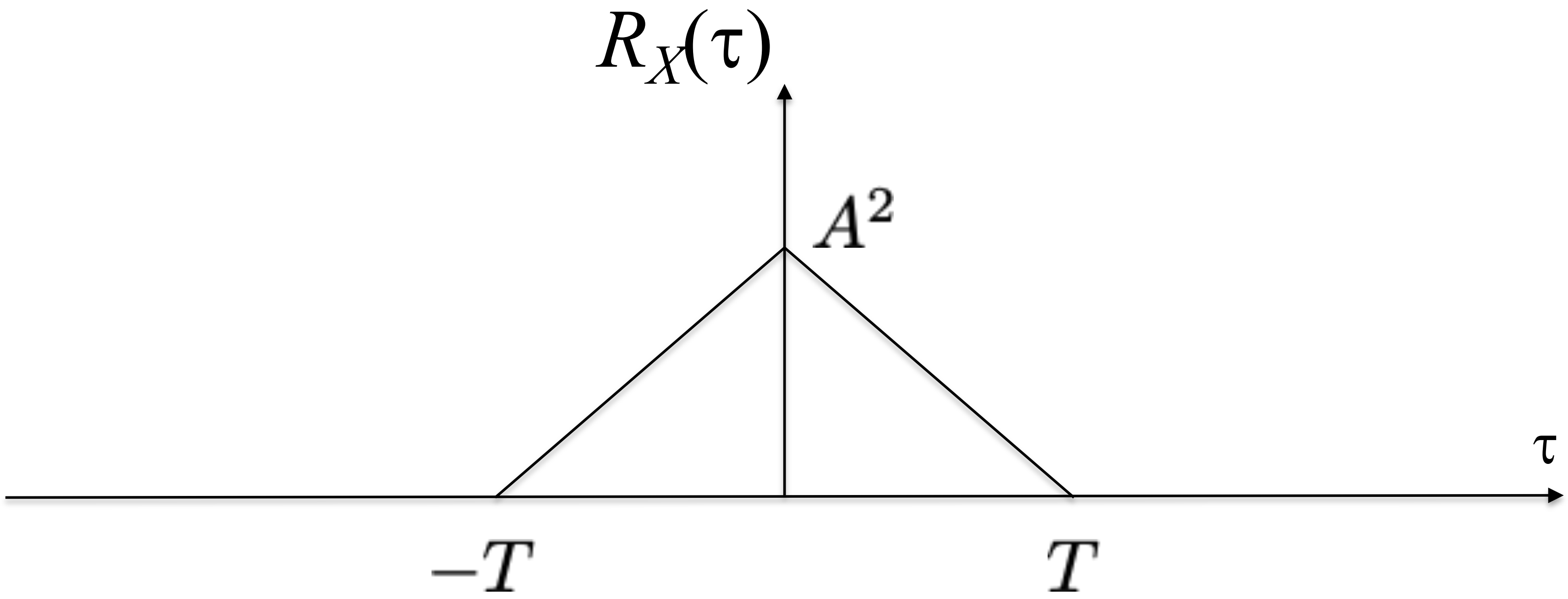
where in the last step, we use again $0 \leq t_d < T$.

From Slide 1-24, $R_X(\tau) = R_X^*(-\tau)$; hence, it suffices to consider $0 \leq \tau < T$.

As a result,

$$
\begin{aligned}
E[X(t_1)X(t_2)] &= E\big[\,E[X(t_1)X(t_2)]\,|t_d]\,\big] \\
&= \begin{cases} \displaystyle\int_\xi^T A^2\frac{1}{T}dt_d + \int_0^{\xi-\tau} A^2\frac{1}{T}dt_d, & 0 \le \tau \le \xi; \\ \displaystyle\int_\xi^{\xi-\tau+T} A^2\frac{1}{T}dt_d, & \xi < \tau < T, \end{cases} \\
&= A^2\left(\frac{T-\tau}{T}\right), 0 \le \tau < T \\
\Bigg(\quad &= \begin{cases} A^2\left(1-\frac{|\tau|}{T}\right), & |\tau| < T; \\ 0, & \text{otherwise} \end{cases} \Bigg)
\end{aligned}
$$

Example: Signal with Random Delay



Cross-Correlation

- The cross-correlation between two processes $X(t)$ and $Y(t)$ is:

$$R_{X,Y}(t,u) = E[X(t)Y^*(u)]$$

- Sometimes, their correlation matrix is given instead for convenience:

$$\mathbf{R}_{X,Y}(t,u) = \begin{bmatrix} R_X(t,u) & R_{X,Y}(t,u) \\ R_{Y,X}(t,u) & R_Y(t,u) \end{bmatrix}$$

Cross-Correlation

- If $X(t)$ and $Y(t)$ are jointly WSS, then

$$\mathbf{R}_{X,Y}(t,u) = \mathbf{R}_{X,Y}(t-u)$$

$$= \begin{bmatrix} R_X(t-u) & R_{X,Y}(t-u) \\ R_{Y,X}(t-u) & R_Y(t-u) \end{bmatrix}$$

Example: Quadrature-Modulated Random Delay Processes

- Consider a pair of quadrature decomposition of $X(t)$ as:

$$\begin{cases} X_I(t) = X(t)\cos\left(2\pi f_c t + \Theta\right) \\ X_Q(t) = X(t)\sin(2\pi f_c t + \Theta) \end{cases}$$

where Θ is independent of $X(t)$ and is uniformly distributed over $[0, 2\pi)$, and

$$X(t) = \sum_{n=-\infty}^{\infty} A \cdot I_n \cdot p(t - nT - t_d).$$

Example: Quadrature-Modulated Random Delay Processes

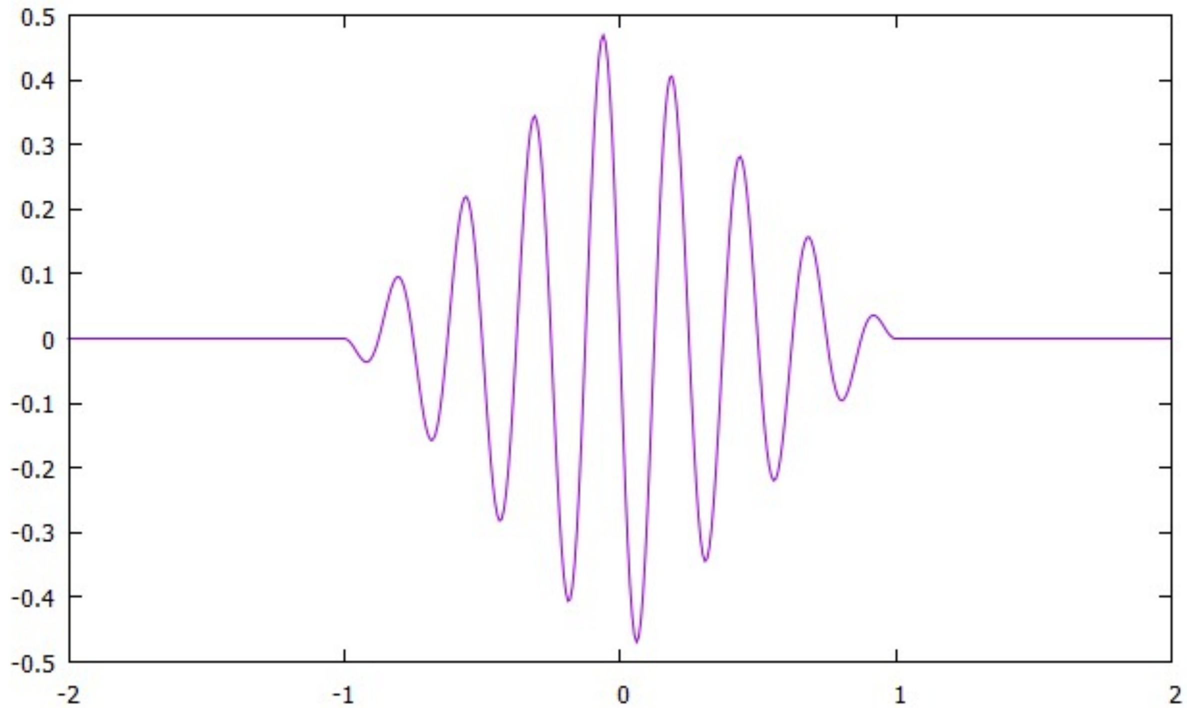
$$
\begin{aligned}
R_{X_I,X_Q}(t,u) &= E[X_I(t)X_Q(u)] \\
&= E[X(t)\cos(2\pi f_c t+\Theta)\cdot X(u)\sin(2\pi f_c u+\Theta)] \\
&= E[X(t)X(u)]E[\sin(2\pi f_c u+\Theta)\cos(2\pi f_c t+\Theta)] \\
&= R_X(t,u)E\left[\frac{\overset{=0}{\cancel{\sin(2\pi f_c(t+u)+2\Theta)}}+\sin(2\pi f_c(u-t))}{2}\right] \\
&= -\frac{1}{2}\sin(2\pi f_c(t-u))R_X(t,u)
\end{aligned}
$$

Example: Quadrature-Modulated Random Delay Processes

$$R_X(t,u) = A^2\left(1-\frac{|t-u|}{T}\right)\mathbf{1}\{|t-u|<T\}$$

Set $A = T = 1$ and $f_c = 4$.

$R_{X_I,X_Q}(\tau)$



Example: Quadrature-Modulated Random Delay Processes

- Notably, if $t = u$, i.e., two quadrature components are synchronized, then

$$R_{X_I,X_Q}(t,t) = 0$$

which indicates that simultaneous observations of the quadrature-modulated processes are "orthogonal" to each other!

(See Slide 1-59 for a formal definition of orthogonality.)

Ergodicity

- For a random-process-modeled noise (or random-process-modeled source) $X(t)$, how can we know its mean and variance?
  - Answer: Relative frequency.
  - How can we get the relative frequency?
    - By measuring $X(t_1)$, $X(t_2)$, …, $X(t_n)$, and calculating their *average*, it is expected that this *time average* will be close to its *mean*.
- Question: Will this *time average* be close to its *mean*, if $X(t)$ is stationary ?
  - Even if for a stationary process, the mean function $\mu_X(t)$ is independent of time t, the answer is negative!

Ergodicity

- An additional *ergodicity* assumption is necessary for *time average* converging to the *ensemble average* μ_X.

Time Average versus Ensemble Average

- Example
  - $X(t)$ is stationary.
  - For any t, $X(t)$ is uniformly distributed over $\{1, 2, 3, 4, 5, 6\}$.
  - Then, *ensemble average* is equal to:

$$1\cdot\frac{1}{6}+2\cdot\frac{1}{6}+3\cdot\frac{1}{6}+4\cdot\frac{1}{6}+5\cdot\frac{1}{6}+6\cdot\frac{1}{6}=3.5$$

Time Average versus Ensemble Average

- We make a series of observations at time 0, T, $2T$, …, $10T$ to obtain 1, 2, 3, 4, 3, 2, 5, 6, 4, 1. (These observations are deterministic!)
- Then, the *time average* is equal to:

$$\frac{1+2+3+4+3+2+5+6+4+1}{10} = 3.1$$

Ergodicity

- **Definition.** A stationary process $X(t)$ is *ergodic in the mean* if

$$1.\quad \Pr\Big[\lim_{T\to\infty}\underbrace{\mu_X(T)}_{\text{Time average}} = \underbrace{\mu_X}_{\text{Ensemble average}}\Big] = 1, \text{ and}$$

$$2.\quad \lim_{T\to\infty}\operatorname{Var}[\mu_X(T)] = 0$$

where

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

Ergodicity

- **Definition**. A stationary process $X(t)$ is *ergodic in the autocorrelation function* if

$$1.\ \Pr\Big[\lim_{T\to\infty} \underbrace{R_X(\tau;T)}_{\text{Time average}} = \underbrace{R_X(\tau)}_{\text{Ensemble average}}\Big] = 1, \text{ and}$$

$$2.\ \lim_{T\to\infty} \mathrm{Var}\big[R_X(\tau;T)\big] = 0$$

where

$$R_X(\tau;T) = \frac{1}{2T}\int_{-T}^{T} X(t+\tau)X^*(t)\mathrm{d}t$$

Ergodicity

- Experiments (or observations) on the same process can only be performed at different time.
- "Stationarity" only guarantees that the observations made at different time come from the same distribution.
  - Example. Rolling two different fair dices will get two results but the two results have the same distribution.

Statistical Average of Random Variables

- Alternative names of *ensemble average*
  - *Mean*
  - *Expected value, or expectation value*
  - *Sample average*
- How about the sample average of a function $g(\ )$ of a random variable X?

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

Statistical Average of Random Variables

- The *n*th *moment* of random variable *X*

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

  - The *2nd moment* is also named *mean-square value.*
- The *n*th *central moment* of random variable *X*

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

  - The *2nd central moment* is also named *variance.*
  - *Square root of the 2nd central moment* is also named *standard deviation.*

Joint Moments

- The joint moment of X and Y is given by:

$$E[X^i(Y^*)^j] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^i(y^*)^j f_{X,Y}(x,y)dxdy$$

  - When $i = j = 1$, the joint moment is specifically named *correlation*.
  - The correlation of centered random variables is specifically named *covariance*.

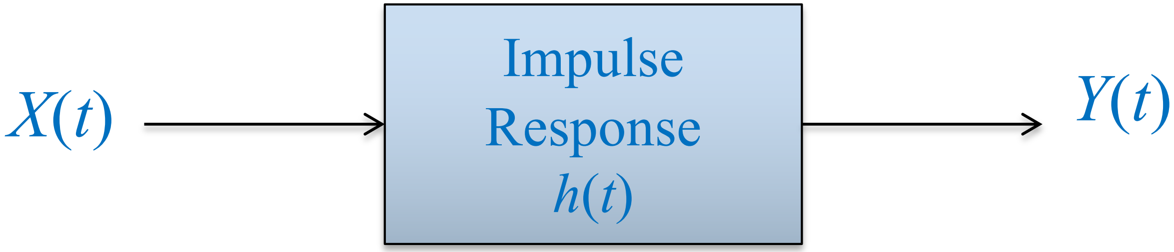
$$\text{Cov}[X,Y] = E[(X-\mu_X)(Y-\mu_Y)^*] = E[XY^*] - \mu_X\mu_Y^*$$

Joint Moments

- Two random variables, X and Y, are *uncorrelated* if $\mathrm{Cov}[X, Y] = 0$.
- Two random variables, X and Y, are *orthogonal* if $E[XY^*] = 0$.
- The covariance, normalized by two standard deviations, is named ***correlation coefficient*** of X and Y.

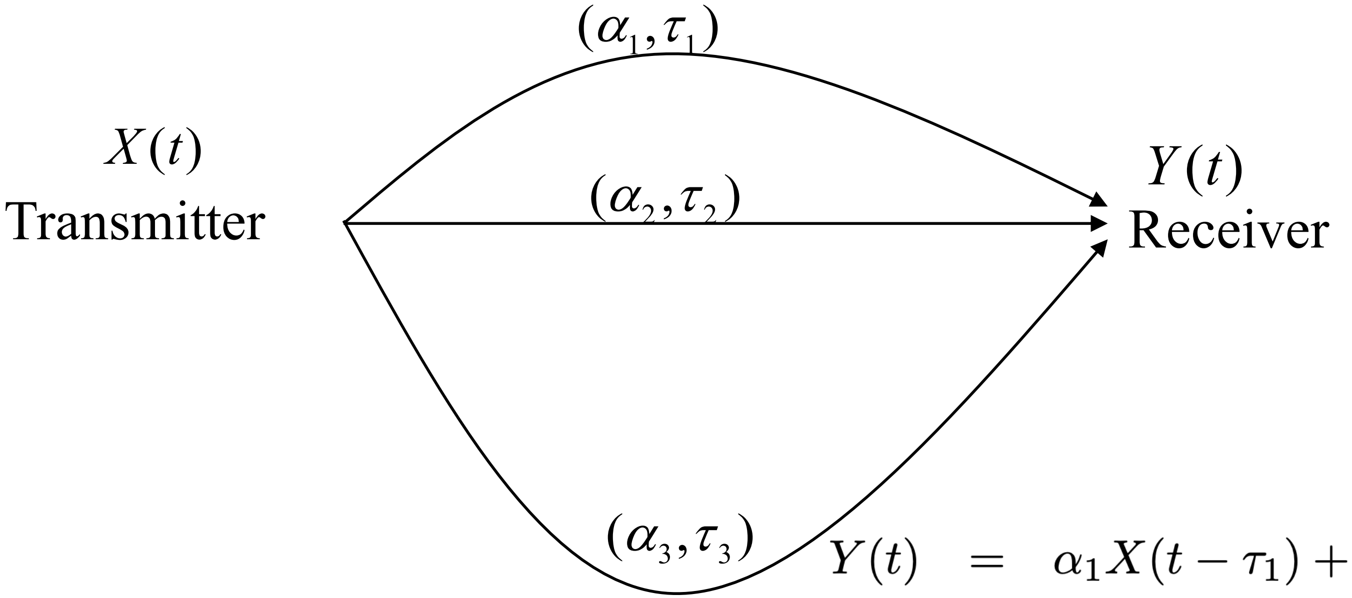
$$\rho = \frac{\mathrm{Cov}[X, Y]}{\sigma_X \sigma_Y}$$

Stable Linear Time-Invariant (LTI) System

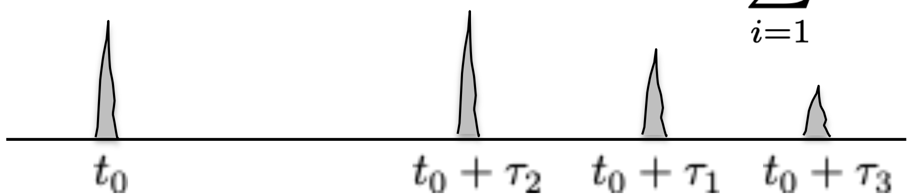


- Linear
  - $Y(t)$ is a linear function of $X(t)$.
  - Specifically, $Y(t)$ is a weighted sum of $X(t)$.
- Time-invariant
  - The weights are time-independent.
- Stable
  - *Dirichlet's condition* (defined later) and $\int_{-\infty}^{\infty} |h(\tau)|^2 \, d\tau < \infty$
  - "Stability" implies that if the input is an energy function (i.e., finite energy), the output is an energy function.

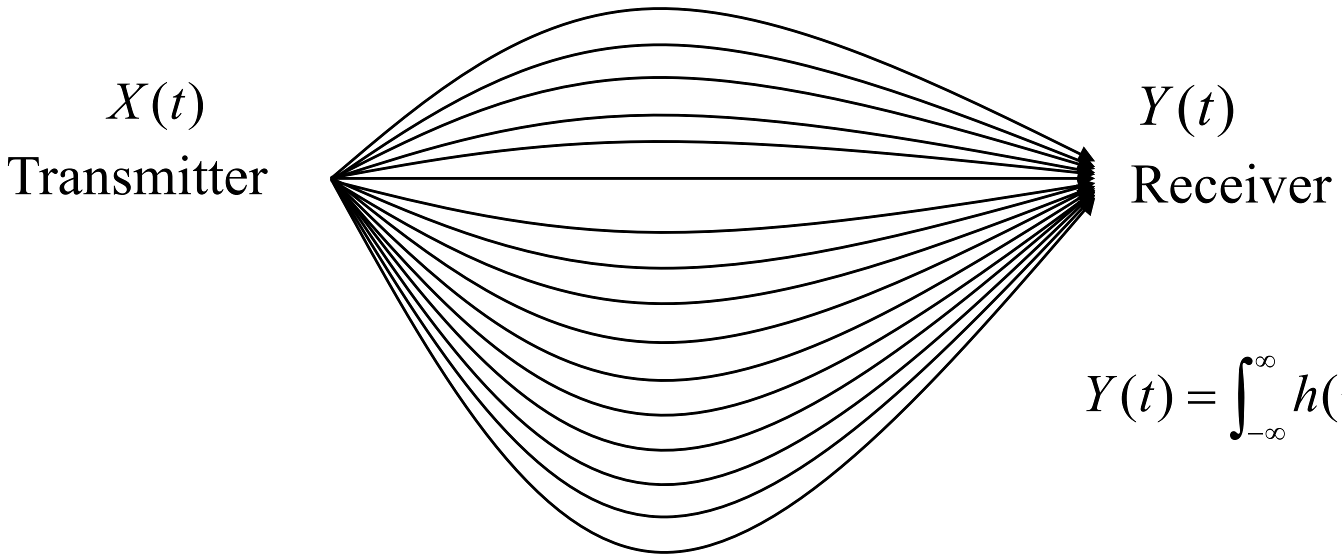
Example of LTI Filter: Mobile Radio Channel



$$
\begin{aligned}
Y(t) &= \alpha_1 X(t-\tau_1) + \alpha_2 X(t-\tau_2) + \alpha_3 X(t-\tau_3) \\
&= \sum_{i=1}^{3} h(\tau_i) s(t-\tau_i), \text{ where } h(\tau_i) = \alpha_i.
\end{aligned}
$$



Example of LTI Filter: Mobile Radio Channel



$$Y(t) = \int_{-\infty}^{\infty} h(\tau) X(t-\tau) d\tau$$

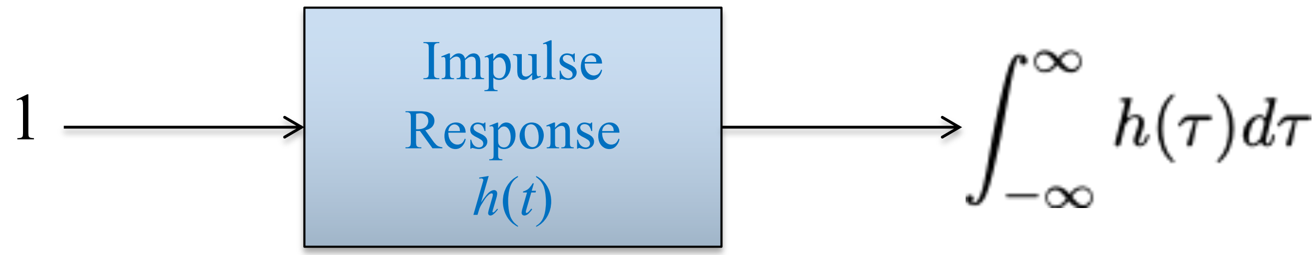
Stable Linear Time-Invariant (LTI) System

- What are the mean and autocorrelation functions of the LTI filter output $Y(t)$?
  - Suppose $X(t)$ is stationary and has finite mean.
  - Suppose $\int_{-\infty}^{\infty} |h(\tau)|\, d\tau < \infty$
  - Then

$$\mu_Y(t) = E[Y(t)] = E\left[\int_{-\infty}^{\infty} h(\tau) X(t-\tau)\, d\tau\right]$$

$$= \int_{-\infty}^{\infty} h(\tau) E[X(t-\tau)]\, d\tau = \mu_X \int_{-\infty}^{\infty} h(\tau)\, d\tau$$

Zero-Frequency (ZF) or Direct Current (DC) Response



- The mean of the *LTI filter output process* is equal to the mean of the *stationary filter input* multiplied by the *DC response* of the system.

$$\mu_Y = \mu_X \int_{-\infty}^{\infty} h(\tau)d\tau$$

Autocorrelation Relation of LTI system

$$\begin{aligned}
R_Y(t,u) &= E[Y(t)Y^*(u)] \\
&= E\left[\left(\int_{-\infty}^{\infty} h(\tau_1)X(t-\tau_1)d\tau_1\right)\left(\int_{-\infty}^{\infty} h(\tau_2)X(u-\tau_2)d\tau_2\right)^*\right] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h^*(\tau_2)E[X(t-\tau_1)X^*(u-\tau_2)]d\tau_2 d\tau_1 \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h^*(\tau_2)R_X(t-\tau_1, u-\tau_2)d\tau_2 d\tau_1
\end{aligned}$$

$\Rightarrow$ If $X(t)$ WSS,

then $R_Y(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h^*(\tau_2)R_X(\tau-\tau_1+\tau_2)d\tau_2 d\tau_1.$

Important Fact: WSS Input Induces WSS Output

- From the above derivations, we conclude:
  - For a stable LTI filter, a WSS input guarantees to induce a WSS output.
  - In general (not necessarily WSS),

$$\mu_Y(t) = \int_{-\infty}^{\infty} h(\tau)\mu_X(t-\tau)d\tau$$

$$R_Y(t,u) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h^*(\tau_2)R_X(t-\tau_1,u-\tau_2)d\tau_2 d\tau_1$$

  - As the above two quantities also relate in the "convolution" form, a spectrum analysis is perhaps better in characterizing their relationship.

Summary

- Random variable, random vector and random process
- Autocorrelation and crosscorrelation
- Definition of WSS
- Why ergodicity?
  - Time average as a good “estimate” of ensemble average