1 Review of Random Variables

1.1 Probability Space

- $\mathcal{U}$

  - $\mathcal{U}$ is called certain space, certain event, universe space, or space.
  - $\mathcal{U}$ is a set of all experiment outcomes.
  - Each element in $\mathcal{U}$ is called an outcome.
  - Each subset in $\mathcal{U}$ is called an event, which is a set of outcomes.

  Ex: Rolling a dice. $\mathcal{U} = \{1, 2, 3, 4, 5, 6\}$. $\{1, 3, 5\}$ is an event.

  To each event, say A, assign a number (or weight) $P(A)$ such that the following conditions (commonly referred as the *Three Axioms of Probability*) are met simultaneously:

$$\begin{cases} (a)\ P(A) \geq 0 \\ (b)\ P(\mathcal{U}) = 1 \\ (c)\ \text{If } A \cap B = \Phi \text{ with } \Phi \text{ an empty set, then } P(A \cup B) = P(A) + P(B). \end{cases}$$

  We call this number the probability of the event A.

  Note: A and B are mutually exclusive or disjoint iff (i.e., if and only if) $A \cap B = \Phi$.

- Definition of a Probability Space : $(\mathcal{U}, \mathcal{F}, P)$

  - $\mathcal{U}$ is the set of all experiment outcomes.
  - $\mathcal{F}$ is the Borel field of $\mathcal{U}$, i.e., the class of all subsets (events) of $\mathcal{U}$.
  - P is the probability of events in $\mathcal{F}$.

  Note: $\mathcal{F}$ is a field defined on $\mathcal{U}$ iff

  1. $\mathcal{U} \in \mathcal{F}$, $\Phi \in \mathcal{F}$.
  2. If $A \in \mathcal{F}$, then $\overline{A} \in \mathcal{F}$, with $\overline{A}$ being the complement of A.
  3. If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$.

- Reading Assignment: Papoulis and Pillai, 4th ed., Chaps. 1–3.

- Recommended Self-Exercise: (Papoulis and Pillai, 4th ed.)
  Ch 2– 3, 5, 10, 14, 19, 22, 27.
  Ch 3– 2, 4, 5, 9.

1.2 Random Variables

- Defn: A random variable (rv) $X(\mu)$ is a mapping from a probability space $(\mathcal{U}, \mathcal{F}, P)$ to a set of real or complex numbers.

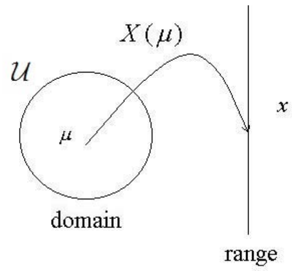


Figure 1:

- Alternative Definition (Papoulis and Pillai)
  An rv $X(\mu)$ is a process of assigning a number $X(\mu)$ to every experiment outcome μ.
  The following two conditions should be satisfied.

  1. The set $\{X(\mu) \leq x\}$ is an event for every x.
  2. $\Pr\{X(\mu) = \infty\} = \Pr\{X(\mu) = -\infty\} = 0$.

- Defn: The probability of the event $\{X(\mu) \leq x\}$ is called the (cumulative) distribution function (cdf) of the rv $X(\mu)$, denoted by

$$F_X(x) \triangleq \Pr\{X(\mu) \leq x\}$$

  for every x.

- Properties of Distribution Function:

  1. $F_X(+\infty) = 1$ and $F_X(-\infty) = 0$.

2. $F_X(x)$ is a nondecreasing function of x, i.e., if $x_1 < x_2$, then $F_X(x_1) \leq F_X(x_2)$.
3. If $F_X(x_0) = 0$, then $F_X(x) = 0$ for every $x \leq x_0$.
4. $\Pr\{X(\mu) > x\} = 1 - \Pr\{X(\mu) \leq x\} = 1 - F_X(x)$.
   Note: $\mathcal{U} = \{X(\mu) > x\} \cup \{X(\mu) \leq x\}$.
5. $F_X(x)$ is continuous from the right, i.e., $\lim_{\epsilon \to 0, \epsilon > 0} F_X(x+\epsilon) = F_X(x)$.
6. $\Pr\{x_1 < X(\mu) \leq x_2\} = F_X(x_2) - F_X(x_1)$ for $x_1 < x_2$.
7. $\Pr\{X(\mu) = x\} = F_X(x) - \lim_{\epsilon \to 0, \epsilon > 0} F_X(x - \epsilon)$.
8. $\Pr\{x_1 \leq X(\mu) \leq x_2\} = \Pr\{X(\mu) \leq x_2\} - \Pr\{X(\mu) < x_1\}$
   $= F_X(x_2) - \lim_{\epsilon \to 0, \epsilon > 0} F_X(x_1 - \epsilon)$.

Notes:

1. $X(\mu)$ is called of continuous type iff $F_X(x)$ is continuous. In the case, $F_X(x)$ is continuous from the left as well and $\Pr\{X(\mu) = x\} = 0$ for all x.
2. $X(\mu)$ is called of discrete type iff $F_X(x)$ is a staircase function, as

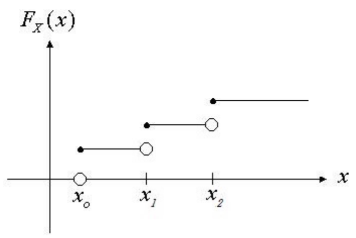


Figure 2:

3. $X(\mu)$ is called of mixed type iff $F_X(x)$ is discontinuous but not a staircase.

1.3 Probability Density Function (For Continuous RV's)

- The derivative $f_X(x) = \frac{dF_X(x)}{dx}$ of continuous and first-order-differentiable $F_X(x)$ is called the (probability) density function (pdf) of $X(\mu)$.

  Note:
  $$F_X(x) = \int_{-\infty}^{x} f_X(y)dy$$
  from Existence theorem.

- Properties of Density Function:

  (1) $f_X(x) \geq 0$.

  (2) $\int_{-\infty}^{+\infty} f_X(x)dx = 1$.

  (3) $\Pr\{x_1 < X(\mu) \leq x_2\} = \int_{x_1}^{x_2} f_X(x)dx$.

  (4) $f_X(x) = \lim\limits_{\Delta x \to 0, \Delta x > 0} \frac{\Pr\{x < X(\mu) \leq x + \Delta x\}}{\Delta x}$.

1.4 Some Special and Useful RV's

- Normal or Gaussian RV:

  An rv $X(\mu)$ is called normal or Gaussian if
  $$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\eta)^2}{2\sigma^2}}$$
  $$F_X(x) = 1 - Q(\frac{x-\eta}{\sigma})$$
  where $Q(x) \triangleq \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}} dy$ is called Q function (which is a well-tabulated function with tight polynomial approximation, upper and lower bounds), η is the mean of $X(\mu)$, and σ is the standard deviation of $X(\mu)$ with $\sigma > 0$ by convention.

- Uniform RV:

  An rv $X(\mu)$ is called uniform between x_1 and x_2 (with $x_1 < x_2$) iff
  $$f_X(x) = \begin{cases} \frac{1}{x_2 - x_1}, & x_1 < x < x_2 \\ 0, & \text{otherwise} \end{cases}.$$

- Binomial RV:
  An rv $X(\mu)$ has a binomial distribution of order n iff it takes value in $\{0, 1, 2, ..., n\}$ and

$$\Pr\{X(\mu) = k\} = \binom{n}{k} p^k (1-p)^{n-k}, \ k = 0, 1, ..., n$$

  with $0 \le p \le 1$.

- Poisson RV:
  An rv $X(\mu)$ is Poisson-distributed with parameter a iff it takes values $0, 1, 2, ...$ with

$$\Pr\{X(\mu) = k\} = e^{-a}\frac{a^k}{k!}, \ k = 0, 1, ...$$

  where $a > 0$.

- Gamma RV:
  An rv $X(\mu)$ is called Gamma-distributed iff, for $b \ge 0$ and $c > 0$,

$$f_X(x) = \begin{cases} \frac{c^{b+1}}{\Gamma(b+1)} x^b e^{-cx}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

  where $\Gamma(b) = \int_0^\infty y^{b-1} e^{-y} dy$, the Gamma function. If b is an integer, $\Gamma(b+1) = b!$ and

$$f_X(x) = \begin{cases} \frac{c^{b+1}}{b!} x^b e^{-cx}, & x \ge 0 \\ 0, & x > 0 \end{cases}$$

  is called the Erlang density. Note further that, when $b = 0$

$$f_X(x) = \begin{cases} ce^{-cx}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

  is an exponential density.

1.5 Conditional Distribution

- $F_X(x|A)$ is called the conditional distribution of an rv $X(\mu)$, assuming the event A, and is defined by

$$F_X(x|A) = \Pr\{X(\mu) \le x | A\} = \frac{\Pr\{X(\mu) \le x, A\}}{P(A)}$$

provided with $P(A) \neq 0$, where $\{X(\mu) \leq x, A\}$ represents the intersection of the events $\{X(\mu) \leq x\}$ and A.

Note: $F_X(x|A)$ has the same properties as $F_X(x)$.

- $f_X(x|A)$ is called the conditional density function of an rv $X(\mu)$, assuming the event A. If $X(\mu)$ is continuous and has first-order-differentiable $F_X(x)$, $f_X(x|A)$ is defined by

$$f_X(x|A) = \frac{dF_X(x|A)}{dx}$$

Note: $f_X(x|A)$ has the same properties as $f_X(x)$.

1.6 Some Results on Conditional Distributions

1. Let $A_1, A_2, \ldots, A_n$ be a partition of $\mathcal{U}$ (i.e., $A_1 \cup A_2 \cup \ldots \cup A_n = \mathcal{U}$ and each single outcome belongs to only one event, e.g., for $n = 3$,

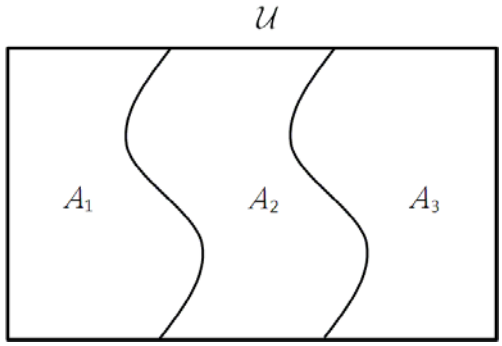


Figure 3:

) with $P(A_k) \neq 0$ for all k. Also, let $B = \{X(\mu) \leq x\}$. Then,

$$\Pr\{X(\mu) \leq x\} = P(B) = \sum_{k=1}^{n} \Pr\{X(\mu) \leq x, A_k\}$$

$$\Rightarrow \quad \Pr\{X(\mu) \leq x\} = \sum_{k=1}^{n} \Pr\{X(\mu) \leq x | A_k\} P(A_k).$$

Thus,

$$\begin{aligned} F_X(x) &= \sum_{k=1}^{n} F_X(x|A_k)P(A_k) \\ f_X(x) &= \sum_{k=1}^{n} f_X(x|A_k)P(A_k). \end{aligned}$$

2.

$$\begin{aligned} P(A|X(\mu) \le x) &= \frac{\Pr\{A, X(\mu) \le x\}}{\Pr\{X(\mu) \le x\}} = \frac{\Pr\{X(\mu) \le x|A\}P(A)}{\Pr\{X(\mu) \le x\}} \\ &= \frac{F_X(x|A)}{F_X(x)}P(A). \end{aligned}$$

$$\begin{aligned} P(A|x_1 < X(\mu) \le x_2) &= \frac{\Pr\{x_1 < X(\mu) \le x_2|A\}P(A)}{\Pr\{x_1 < X(\mu) \le x_2\}} \\ &= \frac{F_X(x_2|A) - F_X(x_1|A)}{F_X(x_2) - F_X(x_1)}P(A). \end{aligned}$$

3.

$$P(A|X(\mu) = x) = \frac{f_X(x|A)}{f_X(x)}P(A).$$

4. Total Probability Theorem:

$$P(A) = \int_{-\infty}^{\infty} P(A|X(\mu) = x)f_X(x)dx$$

for continuous $X(\mu)$, and

$$P(A) = \sum_{x_i} P(A|X(\mu) = x_i)\Pr\{X(\mu) = x_i\}$$

for discrete $X(\mu)$.

5. Bayes' Theorem:

$$f_X(x|A) = \frac{P(A|X(\mu) = x)}{P(A)}f_X(x).$$

1.7 Mean, Variance, and Moments of an RV

- We give definitions for continuous rv's here.
- The mean, or expected value, of an rv $X(\mu)$ is defined by

$$E\{X(\mu)\} \equiv \eta_X \triangleq \int_{-\infty}^{\infty} x f_X(x) dx.$$

- The variance of an rv $X(\mu)$ is defined by

$$\begin{aligned} Var\{X(\mu)\} &\equiv \sigma_X^2 \triangleq \int_{-\infty}^{\infty} (x - E\{X(\mu)\})^2 f_X(x) dx \\ &= E\{(X(\mu) - E\{X(\mu)\})^2\} \\ &= E\{X(\mu)^2\} - E^2\{X(\mu)\}. \end{aligned}$$

  Ex: For a Poisson-distributed rv with pdf $\Pr\{X(\mu) = k\} = e^{-a}\frac{a^k}{k!}$, $k = 0, 1, 2, ...$, its mean and variance are $E\{X(\mu)\} = Var\{X(\mu)\} = a$.

- By convention, the positive σ_X ($\sigma_X \geq 0$) is called the standard deviation of the rv $X(\mu)$.
- Likewise, the conditional mean and the conditional variance are defined respectively by

$$\begin{aligned} E\{X(\mu)|A\} &= \int_{-\infty}^{\infty} x f_X(x|A) dx \\ Var\{X(\mu)|A\} &= \int_{-\infty}^{\infty} (x - E\{X(\mu)|A\})^2 f_X(x|A) dx. \end{aligned}$$

- Moments of an RV $X(\mu)$:

$$m_n \triangleq E\{X(\mu)^n\} = \int_{-\infty}^{\infty} x^n f_X(x) dx, \; n = 1, 2, ...$$

- Central Moments of an RV $X(\mu)$:

$$u_n \triangleq E\{(X(\mu) - \eta_X)^n\} = \int_{-\infty}^{\infty} (x - \eta_X)^n f_X(x) dx, \; n = 1, 2, ...$$

- Absolute Moments of an RV $X(\mu)$: $E\{|X(\mu)|^n\}, E\{|X(\mu) - \eta_X|^n\}$.

- Generalized Moments of an RV $X(\mu)$: $E\{(X(\mu)-a)^n\}, E\{|X(\mu)-a|^n\}$.

- Ex: For a Gaussian rv with mean zero and variance σ_X^2,

$$E\{X(\mu)^n\} = \begin{cases} 0, & n = 2k+1 \\ (n-1)!!\sigma_X^n, & n = 2k \end{cases}$$
$$E\{|X(\mu)|^n\} = \begin{cases} 2^k k! \sigma_X^{2k+1}\sqrt{\frac{2}{\pi}}, & n = 2k+1 \\ (n-1)!!\sigma_X^n, & n = 2k \end{cases}$$

where we have defined

$$n!! = \begin{cases} n(n-2)(n-4)...2, & \text{if } n \text{ is even} \\ n(n-2)(n-4)...1, & \text{if } n \text{ is odd} \end{cases}.$$

- Two Notes:

1. For arbitrary a and integer n, and for $\epsilon > 0$,

$$\Pr\{|X(\mu) - a| \geq \epsilon\} \leq \frac{E\{|X(\mu) - a|^n\}}{\epsilon^n}.$$

This result is known as the inequality of Bienaymé. When $a = \eta_X$ and $n = 2$, it is called Tchebycheff's inequality.

2. Also, $m_1 = E\{X(\mu)\} = \eta_X$ and $u_2 = E\{(X(\mu) - \eta_X)^2\} = Var\{X(\mu)\}$.

1.8 Characteristic Function and Moment Generating Function

- The characteristic function of $X(\mu)$ is defined by

$$\Phi_X(\omega) \triangleq E\{e^{j\omega X(\mu)}\} = \int_{-\infty}^{\infty} f_X(x)e^{j\omega x}dx$$

with ω being real and $j = \sqrt{-1}$.

Notes:

1. $f_X(x) \xleftrightarrow{\mathcal{F}} \Phi_X(-\omega)$, i.e., $\Phi_X(-\omega)$ is the Fourier transform of $f_X(x)$.
2. $|\Phi_X(\omega)| \leq \Phi_X(0)$ because

$$\begin{aligned} |\Phi_X(\omega)| &= |\int_{-\infty}^{\infty} e^{j\omega x} f_X(x)dx| \\ &\leq \int_{-\infty}^{\infty} |e^{j\omega x} f_X(x)|dx = \int_{-\infty}^{\infty} f_X(x)dx \\ &= 1 = \Phi_X(0). \end{aligned}$$

- The moment generating function of $X(\mu)$ is defined by

$$\Phi_X(s) \triangleq E\{e^{sX(\mu)}\} = \int_{-\infty}^{\infty} f_X(x)e^{sx}dx$$

with s a complex number. $\Phi_X(\omega)$ is obtained from $\Phi_X(s)$ by setting $s = j\omega$. Note that $\Phi_X(-s)$ is the Laplace transform of $f_X(x)$.

- Ex: For a Gaussian rv with mean η_X and variance σ_X^2, $\Phi_X(\omega) = \exp\{j\eta_X\omega - \frac{1}{2}\sigma_X^2\omega^2\}$ and $\Phi_X(s) = \exp\{\eta_X s + \frac{1}{2}\sigma_X^2 s^2\}$.

- The moments m_n's of the rv $X(\mu)$ are related to its characteristic function $\Phi_X(\omega)$ by $m_n = (-j)^n \frac{d^n\Phi_X(\omega)}{d\omega^n}|_{\omega=0}$, and its moment generating function $\Phi_X(s)$ by $m_n = \frac{d^n\Phi_X(s)}{ds^n}|_{s=0}$.

- A random variable is completely statistically characterized by any one among its pdf, cdf, cf, mgf, and all moments.

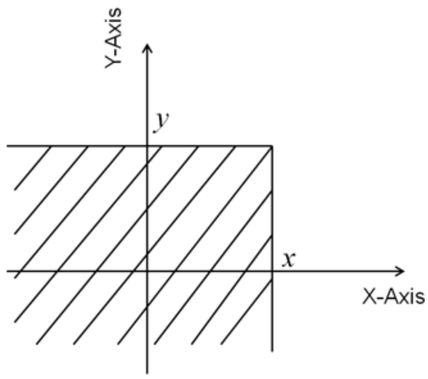
- Reading Assignment: Papoulis and Pillai, 4th ed., Chaps. 4-5

- Recommended Self-Exercise: (Papoulis and Pillai, 4th ed.)
  Ch4 – 1, 2, 7, 9, 16, 18.
  Ch5 – 6, 8, 13, 22, 27, 42, 43, 46.

1.9 Joint Statistics

- The joint distribution $F_{X,Y}(x,y)$ of two rv's $X(\mu)$ and $Y(\mu)$ is the probability of the event $\{X(\mu) \leq x, Y(\mu) \leq y\}$, i.e.,

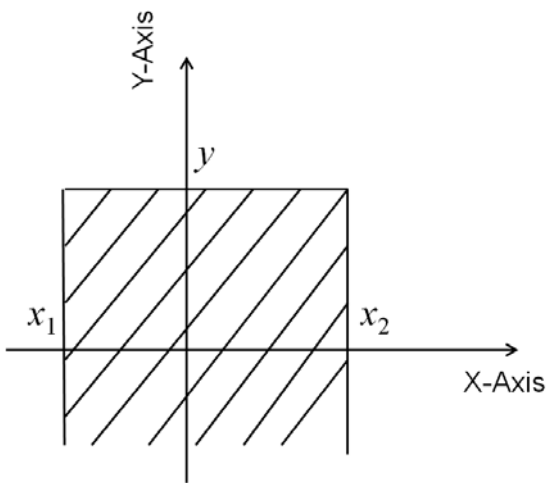
$$F_{X,Y}(x,y) = \Pr\{X(\mu) \leq x, Y(\mu) \leq y\} \ ,\forall x, \forall y.$$



- Properties:

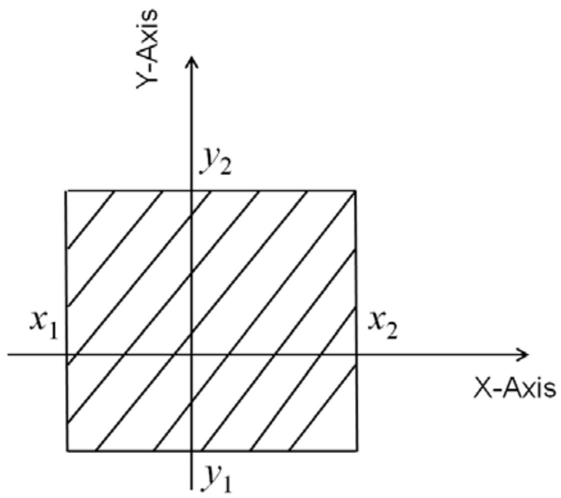
  1. $F_{X,Y}(-\infty, y) = 0$, $F_{X,Y}(x, -\infty) = 0$, and $F_{X,Y}(\infty, \infty) = 1$.
  2.

$$\begin{aligned}
& \Pr\{x_1 < X(\mu) \leq x_2, Y(\mu) \leq y\} \\
= \; & \Pr\{X(\mu) \leq x_2, Y(\mu) \leq y\} - \Pr\{X(\mu) \leq x_1, Y(\mu) \leq y\} \\
= \; & F_{X,Y}(x_2, y) - F_{X,Y}(x_1, y). \\
& \Pr\{X(\mu) \leq x, y_1 < Y(\mu) \leq y_2\} \\
= \; & F_{X,Y}(x, y_2) - F_{X,Y}(x, y_1).
\end{aligned}$$



3.

$$\begin{aligned} & \Pr\{x_1 < X(\mu) \le x_2, y_1 < Y(\mu) \le y_2\} \\ = \; & F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_2) - F_{X,Y}(x_2, y_1) + F_{X,Y}(x_1, y_1). \end{aligned}$$



- The joint density of two continuous rv's $X(\mu)$ and $Y(\mu)$ with continuous and first-order-differentiable $F_{X,Y}(x, y)$ is defined by

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

It follows that

1. $F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(\alpha, \beta) d\alpha d\beta$ from the existence theorem.
2. $\Pr\{(X(\mu), Y(\mu)) \in D\} = \int\int_D f_{X,Y}(x, y) dx dy$ with D a two-dimensional region.

- Note: In our study of several rv's, the statistic of each rv is called marginal. For example,

the marginal distribution of $X(\mu)$ is $F_X(x) = F_{X,Y}(x, \infty)$, and

the marginal density of $X(\mu)$ is $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$.

- Ex: Two rv's $X(\mu)$ and $Y(\mu)$ are called jointly normal (or, jointly Gaussian) if their joint density is

$$\begin{aligned} f_{X,Y}(x,y) &= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\gamma_{XY}^2}}\exp\{\frac{-1}{2(1-\gamma_{XY}^2)}[\frac{(x-\eta_X)^2}{\sigma_X^2} \\ &\quad -2\gamma_{XY}\frac{(x-\eta_X)(y-\eta_Y)}{\sigma_X\sigma_Y}+\frac{(y-\eta_Y)^2}{\sigma_Y^2}]\} \end{aligned}$$

for $|\gamma_{XY}|<1$, where $\eta_X = E\{X(\mu)\}$, $\eta_Y = E\{Y(\mu)\}$, $\sigma_X^2 = Var\{X(\mu)\}$, $\sigma_Y^2 = Var\{Y(\mu)\}$, and $\gamma_{XY} \triangleq \frac{E\{X(\mu)Y(\mu)\}-\eta_X\eta_Y}{\sigma_X\sigma_Y}$ with $E\{X(\mu)Y(\mu)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)dxdy$.

- $X(\mu)$ and $Y(\mu)$ are called (statistically) (mutually) independent, iff the events $\{X(\mu)\in A\}$ and $\{Y(\mu)\in B\}$ are independent for all A and all B (noting here that A and B represent two sets of numbers, not events), i.e., iff

$$\Pr\{X(\mu)\in A, Y(\mu)\in B\} = \Pr\{X(\mu)\in A\}\Pr\{Y(\mu)\in B\}$$

for all A and all B. It can be shown that

$$\begin{aligned} & \Pr\{X(\mu)\in A, Y(\mu)\in B\} = \Pr\{X(\mu)\in A\}\Pr\{Y(\mu)\in B\}, \forall A, B \\ \Longleftrightarrow\ & F_{X,Y}(x,y) = F_X(x)F_Y(y), \forall x,y \\ \Longleftrightarrow\ & f_{X,Y}(x,y) = f_X(x)f_Y(y), \forall x,y. \end{aligned}$$

- If $X(\mu)$ and $Y(\mu)$ are independent, then $Z(\mu) = g(X(\mu))$ and $W(\mu) = h(Y(\mu))$ are also independent.

  Proof: Define A_z and B_w as $A_z = \{x|g(x)\le z\}$ and $B_w = \{y|h(y)\le w\}$. Then, the following equivalences hold

$$\begin{aligned} \{X(\mu)\in\ & A_z\} = \{g(X(\mu))\le z\} = \{Z(\mu)\le z\} \\ \{Y(\mu)\in\ & B_w\} = \{h(Y(\mu))\le w\} = \{W(\mu)\le w\}. \end{aligned}$$

  Therefore, if $\{X(\mu)\in A_z\}$ and $\{Y(\mu)\in B_w\}$ are independent for all z and all w, then $\{Z(\mu)\le z\}$ and $\{W(\mu)\le w\}$ are independent for all z and all w. Q.E.D.

- The covariance C_{XY} of rv's $X(\mu)$ and $Y(\mu)$ is defined by

$$\begin{aligned} C_{XY} &\triangleq E\{(X(\mu)-\eta_X)(Y(\mu)-\eta_Y)\} \\ &= E\{X(\mu)Y(\mu)\} - \eta_X\eta_Y. \end{aligned}$$

The correlation coefficient γ_{XY} of rv's $X(\mu)$ and $Y(\mu)$ is defined by $\gamma_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y}$.

Note: $|\gamma_{XY}| \leq 1$ and $|C_{XY}| \leq \sigma_X \sigma_Y$. (Prove them by Schwartz's inequality.)

- Two rv's $X(\mu)$ and $Y(\mu)$ are called uncorrelated if $C_{XY} = 0$.

  Two rv's $X(\mu)$ and $Y(\mu)$ are called orthogonal if $E\{X(\mu)Y(\mu)\} = 0$.

  (Note: $E\{X(\mu)Y(\mu)\}$ is also called the correlation of $X(\mu)$ and $Y(\mu)$.)

- Theorem : If $X(\mu)$ and $Y(\mu)$ are independent, then they are uncorrelated.

  Proof : You show it.

  In general, *uncorrelated* does NOT imply *independent*.

  However, for jointly Gaussian rv's, *uncorrelated* is equivalent to *independent*, i.e.,

  if $X(\mu)$ and $Y(\mu)$ are jointly Gaussian, *uncorrelated* $\Leftrightarrow$ *independent*.

  That is, for $X(\mu)$ and $Y(\mu)$,

$$\begin{aligned} &uncorrelated \underset{in\ general}{\Longleftarrow} independent \\ &uncorrelated \underset{in\ general}{\not\Rightarrow} independent \\ &uncorrelated \underset{jointly\ Gaussian}{\Longrightarrow} independent. \end{aligned}$$

- A Naive Question: If $X(\mu)$ and $Y(\mu)$ are uncorrelated, does it mean that $h(X(\mu))$ and $g(Y(\mu))$ are also uncorrelated?

  Answer: Not in general.

  Counter-Example: Let $X(\mu)$ be a real-valued rv with a symmetric pdf $f_X(x) = f_X(-x)$ for all x. Also, let $Y(\mu)$=$X^2(\mu)$. Then, $E\{X(\mu)Y(\mu)\} = E\{X^3(\mu)\} = \int_{-\infty}^{\infty} x^3 f_X(x)dx = 0$ and thus $X(\mu)$ and $Y(\mu)$ are uncorrelated. However, if $Z(\mu) \triangleq X^2(\mu) = h(X(\mu))$ and $W(\mu) \triangleq Y(\mu) = X^2(\mu) = g(Y(\mu))$, then $E\{h(X(\mu))g(Y(\mu))\} = E\{X^4(\mu)\} = \int_{-\infty}^{\infty} x^4 f_X(x)dx$ $\neq E\{h(X(\mu))\}E\{g(Y(\mu))\}$ in general.

- More on Event Independence:

N events $A_1, A_2, ..., A_N$ are called (mutually) independent iff, for any $2 \leq k \leq N$,

$$P(B_1 \cap B_2 \cap \cdots \cap B_k) = \prod_{i=1}^{k} P(B_i) \qquad \circledast$$

where $B_i \in \{A_1, A_2, ..., A_N\}$ and $B_i \neq B_j$ if $i \neq j$, holds true.

Ex *: $\mathcal{U} = \{1, 2, 3, 4\}$ and $P(\{i\}) = \frac{1}{4}$ for $i = 1, 2, 3, 4$. Define $A = \{1, 2\}$, $B = \{2, 3\}$, and $C = \{3, 1\}$. Now,

$$\begin{aligned} P(A) &= P(B) = P(C) = \frac{1}{2} \\ P(A \cap B) &= \frac{1}{4} = P(B \cap C) = P(A \cap C). \end{aligned}$$

Thus, A, B, C are independent in pairs. But,

$$P(A \cap B \cap C) = P(\Phi) = 0 \neq \frac{1}{8} = P(A)P(B)P(C).$$

This implies that A, B, C are not (mutually) independent.

This example tells us that *independence in pairs* does NOT imply *mutual independence.*

Ex: (Papoulis and Pillai) Show that "$2^N - (N + 1)$" equations are needed to establish the independence of N events.

Ans: For any k event, $\circledast$ needs to be satisfied; in this case, we have $\binom{N}{k}$ equations to ensure the independence in groups of k events. Now, $N \geq k \geq 2$ implies a totality of

$$\begin{aligned} \binom{N}{2} + \binom{N}{3} + \cdots + \binom{N}{N} &= \sum_{k=0}^{N} \binom{N}{k} - \binom{N}{0} - \binom{N}{1} \\ &= 2^N - 1 - N \qquad \text{equations.} \end{aligned}$$

- $X_1(\mu), X_2(\mu), ..., X_N(\mu)$ are called (mutually) independent,

  iff any group of m rv's with $m \in \{2, 3, ..., N\}$ are independent, i.e., for m distinct numbers $i_1, i_2, ..., i_m$ in $\{1, 2, ..., N\}$, $\{X_{i_1}(\mu) \in A_{i_1}\}, \{X_{i_2}(\mu) \in A_{i_2}\}, ..., \{X_{i_m}(\mu) \in A_{i_m}\}$ are independent for all $A_{i_1}, A_{i_2}, ..., A_{i_m}$.

  iff the events $\{X_1(\mu) \leq x_1\}, ..., \{X_N(\mu) \leq x_N\}$ are independent, $\forall x_i' s$.

iff $f_{X_1,X_2,\ldots,X_N}(x_1,x_2,\ldots,x_N)=\prod_{i=1}^{N} f_{X_i}(x_i)$, $\forall x_i's$.

iff $F_{X_1,X_2,\ldots,X_N}(x_1,x_2,\ldots,x_N)=\prod_{i=1}^{N} F_{X_i}(x_i)$, $\forall x_i's$.

iff $\Phi_{X_1,X_2,\ldots,X_N}(\omega_1,\omega_2,\ldots,\omega_N)=\prod_{i=1}^{N} \Phi_{X_i}(\omega_i)$, $\forall \omega_i's$.

Here, we have defined the joint CF

$$\Phi_{X_1,X_2,\ldots,X_N}(\omega_1,\omega_2,\ldots,\omega_N) \triangleq E\{\exp\{j\sum_{i=1}^{N}\omega_i X_i(\mu)\}\}.$$

For N rv's, one "only" needs to show that

$$\begin{aligned} F_{X_1,\ldots,X_N}(x_1,\ldots,x_N) &= \Pr\{X_1(\mu)\le x_1,\ldots,X_N(\mu)\le x_N\} \\ &= \Pr\{X_1(\mu)\le x_1\}\cdots\Pr\{X_N(\mu)\le x_N\} \\ &= F_{X_1}(x_1)\cdots F_{X_N}(x_N) \\ \text{or} \qquad f_{X_1,\ldots,X_N}(x_1,\ldots,x_N) &= \prod_{i=1}^{N} f_{X_i}(x_i). \end{aligned}$$

This is because

$$\begin{aligned} f_{X_1,\ldots,X_{N-1}}(x_1,\ldots,x_{N-1}) &= \int_{-\infty}^{\infty} f_{X_1,\ldots,X_N}(x_1,\ldots,x_N)dx_N \\ &= \int_{-\infty}^{\infty}\prod_{i=1}^{N-1} f_{X_i}(x_i) f_{X_N}(x_N)dx_N \\ &= \prod_{i=1}^{N-1} f_{X_i}(x_i) \end{aligned}$$

and repeating this procedure recursively can yield all equations required by $\circledast$.

- If $X_1(\mu), X_2(\mu), \ldots, X_N(\mu)$ are (mutually) independent, then
  1. $\Phi_{X_1+X_2+\cdots+X_N}(\omega)=\prod_{i=1}^{N}\Phi_{X_i}(\omega)$.
  2. $f_{X_1+X_2+\cdots+X_N}(z)=f_{X_1}(z)*f_{X_2}(z)*\cdots*f_{X_N}(z)$ where $*$ is the convolution operator, i.e.,

$$f(z)*g(z)=\int_{-\infty}^{\infty} f(x)g(z-x)dx=\int_{-\infty}^{\infty} g(x)f(z-x)dx.$$

- $X_1(\mu), X_2(\mu), ..., X_N(\mu)$ are called independent and identically distributed (iid), iff they are mutually independent and have the same marginal density (denoted by $f_X(x)$ with the corresponding characteristic function $\Phi_X(\omega)$).

  In the case,

$$
\begin{aligned}
f_{X_1,\ldots,X_N}(x_1,\ldots,x_N) &= \prod_{i=1}^{N} f_{X_i}(x_i) = \prod_{i=1}^{N} f_X(x_i) \\
\Phi_{X_1+X_2+\cdots+X_N}(\omega) &= \prod_{i=1}^{N} \Phi_{X_i}(\omega) = [\Phi_X(\omega)]^N.
\end{aligned}
$$

- Central Limit Theorem

  If $X_n(\mu)$, $n = 1, 2, 3, ...$, is a sequence of independent rv's with arbitrary pdf's and $\sigma^2_{X_i} < \infty$, $\forall i$, then the random sequence of sums $Y_1(\mu), Y_2(\mu), ...$, where

$$
Y_n(\mu) \triangleq \frac{S_n(\mu) - E\{S_n(\mu)\}}{\sqrt{Var\{S_n(\mu)\}}}
$$

  with $S_n(\mu) \triangleq \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i(\mu)$, converges in distribution to a Gaussian rv with mean zero and variance one.

  Special Case: When $X_n(\mu)'s$ are iid, we can restate and prove the theorem as follows.

  Let $\eta = E\{X_n(\mu)\}$ and $\sigma^2 = Var\{X_n(\mu)\} < \infty, \forall n.$ First, let us define the new rv's

$$
U_n(\mu) = \frac{X_n(\mu) - \eta}{\sigma}, \forall n
$$

  It follows that $U_n(\mu)$ has zero mean and unit variance. Now, if we define

$$
Y_n(\mu) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} U_i(\mu)
$$

  then $Y_1(\mu), Y_2(\mu), ...$ should tend to a Gaussian rv with zero mean and unit variance, in the distribution sense, according to Central Limit Theorem.

Proof: Let us check the characteristic function of $Y_n(\mu)$,

$$\begin{aligned}
\Phi_{Y_n}(\omega) &= E\{e^{j\omega Y_n(\mu)}\} \\
&= E\{\exp\{j\omega\frac{1}{\sqrt{n}}\sum_{i=1}^{n}U_i(\mu)\}\} \\
&= E\{\prod_{i=1}^{n}\exp\{j\frac{\omega}{\sqrt{n}}U_i(\mu)\}\} \\
&= \prod_{i=1}^{n}E\{\exp\{j\frac{\omega}{\sqrt{n}}U_i(\mu)\}\} \ (\because \text{ independent}) \\
&= [\Phi_U(\frac{\omega}{\sqrt{n}})]^n. \qquad \circledast \ (\because \text{ identically distributed})
\end{aligned}$$

Expanding $\Phi_U(\frac{\omega}{\sqrt{n}})$ in a Taylor series, we have

$$\begin{aligned}
\Phi_U(\frac{\omega}{\sqrt{n}}) &= E\{e^{j\frac{\omega}{\sqrt{n}}U(\mu)}\} \qquad (e^x = \sum_{k=0}^{\infty}\frac{x^k}{k!}) \\
&= 1 + j\frac{\omega}{\sqrt{n}}E\{U(\mu)\} - \frac{\omega^2}{2n}E\{U^2(\mu)\} \\
&\quad + \frac{(j\omega)^3}{6(\sqrt{n})^3}E\{U^3(\mu)\} - \cdots \\
&= 1 - \frac{\omega^2}{2n} + \frac{1}{n}R(\omega, n)
\end{aligned}$$

with $\frac{R(\omega,n)}{n}$ denoting the remainder. We first note that $\lim_{n\to\infty} R(\omega, n) = 0$. Next, substituting the above series into $\circledast$, we obtain $\Phi_{Y_n}(\omega)$ in the form of

$$\Phi_{Y_n}(\omega) = [1 - \frac{\omega^2}{2n} + \frac{1}{n}R(\omega, n)]^n.$$

Taking the natural logarithm,

$$\ln \Phi_{Y_n}(\omega) = n \ln[1 - \frac{\omega^2}{2n} + \frac{1}{n}R(\omega, n)].$$

For small $|x|$, i.e., $|x| \ll 1$,

$$\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$$

(noting that $\ln(1+x) = \sum_{n=1}^{\infty} \frac{1}{n}(-1)^{n+1} x^n$ for $|x| < 1$), which expands $\ln \Phi_{Y_n}(\omega)$ into

$$\begin{aligned}
\ln \Phi_{Y_n}(\omega) &= n[-\frac{\omega^2}{2n} + \frac{1}{n}R(\omega, n) \\
&\quad -\frac{1}{2}(-\frac{\omega^2}{2n} + \frac{1}{n}R(\omega, n))^2 + \cdots] \\
&= -\frac{\omega^2}{2} + \underset{\substack{\downarrow \\ 0}}{R(\omega, n)} - \frac{1}{n}\underset{\substack{\downarrow \\ 0}}{[\cdots]}.
\end{aligned}$$

When $n \to \infty$, we can find that

$$\lim_{n\to\infty} \Phi_{Y_n}(\omega) = \exp\{-\frac{\omega^2}{2}\}$$

which is the characteristic function of a Gaussian rv with zero mean and unit variance. Q.E.D.

Note: There are several versions of Central Limit Theorem. For example, there exist versions for correlated sequences. For more exposure, read the reference: P. Billingsley, "Probability and Measure," 2^{nd} edition, Wiley, 1986 (Section 27).

1.10 Function of RV's

- (Material from J. G. Proakis, "Digital Communications," 3rd edition, MaGraw-Hill, 1995, Sec 2.1.2.)
  Question: Given an rv $X(\mu)$ as well as its pdf $f_X(x)$, how can we determine the pdf of $Y(\mu)$ defined by $Y(\mu) = g(X(\mu))$, where $g(x)$ is a given function of x?

  Note: The mapping $g(x)$ may not be a one-to-one mapping.

  Ex 1: $Y(\mu) = aX(\mu) + b$, where a and b are constants, and $a > 0$.
  First, look at

$$\begin{aligned}
F_Y(y) &= \Pr\{Y(\mu) \le y\} \\
&= \Pr\{aX(\mu) + b \le y\} \\
&= \Pr\{X(\mu) \le \frac{1}{a}(y-b)\} \\
&= F_X(\frac{1}{a}(y-b)) \\
\Rightarrow f_Y(y) &= \frac{1}{a} f_X(\frac{1}{a}(y-b)).
\end{aligned}$$

This example illustrates that the answer to Question is straightforward if the mapping $g(x)$ is one-to-one.

Ex 2: $Y(\mu) = aX^2(\mu) + b$, where a and b are constants, and $a > 0$.
Again,

$$\begin{aligned} F_Y(y) &= \Pr\{Y(\mu) \le y\} \\ &= \Pr\{aX^2(\mu) + b \le y\} \\ &= \Pr\{X^2(\mu) \le \frac{1}{a}(y-b)\} \\ &= \begin{cases} 0, & y \le b \\ F_X(\sqrt{\frac{1}{a}(y-b)}) - F_X(-\sqrt{\frac{1}{a}(y-b)}), & y > b \end{cases}. \end{aligned}$$

For $y \le b$, $f_Y(y) = 0$.
For $y > b$,

$$f_Y(y) = \frac{f_X(\sqrt{\frac{1}{a}(y-b)})}{2a\sqrt{\frac{1}{a}(y-b)}} + \frac{f_X(-\sqrt{\frac{1}{a}(y-b)})}{2a\sqrt{\frac{1}{a}(y-b)}}.$$

- In general, suppose that $x_1, x_2, \ldots, x_n$ are the real roots of the equation $g(x) = y$. Then, the pdf of rv $Y(\mu) = g(X(\mu))$ may be expressed as

$$f_Y(y) = \sum_{i=1}^{n} \frac{f_X(x_i)}{|g'(x_i)|}$$

where the roots x_i, $i = 1, 2, \ldots, n$, are functions of y and g is first-order differentiable with $g^{'}$ being its first-order derivative (Jacobian Formula).

- Function of Multidimensional RV's:

Let $X_i(\mu), i = 1, 2, \ldots, n$, be rv's with the given joint pdf $f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n)$.
Let $Y_i(\mu), i = 1, 2, \ldots, n$, be another set of rv's defined by

$$Y_i(\mu) = g_i(X_1(\mu), X_2(\mu), \ldots, X_n(\mu)), i = 1, 2, \ldots, n.$$

The mapping between $(X_1(\mu), X_2(\mu), \ldots, X_n(\mu))$ and $(Y_1(\mu), Y_2(\mu), \ldots, Y_n(\mu))$ is deterministic and one-to-one.
Assume that g_i's are single-valued functions with continuous first partial derivatives and inverse functions. Let g_i^{-1} denote the inverse of g_i.
Then, the joint pdf of $Y_i(\mu)$'s is

$$f_{X_1,X_2,\ldots,X_n}(g_1^{-1}(y_1, y_2, \ldots, y_n), g_2^{-1}(y_1, y_2, \ldots, y_n), \ldots, g_n^{-1}(y_1, y_2, \ldots, y_n)) \cdot |J| \quad (+)$$

where J is the Jacobian, defined by

$$J = \begin{vmatrix} \frac{\partial g_1^{-1}}{\partial y_1} & \frac{\partial g_2^{-1}}{\partial y_1} & \cdots & \frac{\partial g_n^{-1}}{\partial y_1} \\ \frac{\partial g_1^{-1}}{\partial y_2} & \frac{\partial g_2^{-1}}{\partial y_2} & \cdots & \frac{\partial g_n^{-1}}{\partial y_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1^{-1}}{\partial y_n} & \frac{\partial g_2^{-1}}{\partial y_n} & \cdots & \frac{\partial g_n^{-1}}{\partial y_n} \end{vmatrix}.$$

Note: The Jacobian formula (+) can be generalized to the many-to-one mapping from $(X_1(\mu), X_2(\mu), \ldots, X_n(\mu))$ to $(Y_1(\mu), Y_2(\mu), \ldots, Y_n(\mu))$ with $f_{X_1,X_2,\ldots,X_n} \cdot |J|$ summed over all possible roots of the mapping.

1.11 Some Useful Distributions

- Chi-Square Distributions

Let

$$Y(\mu) = \sum_{i=1}^{n} X_i^2(\mu)$$

where n is **even** and where $X_i(\mu), i = 1, 2, \ldots, n$, are mutually statistically independent Gaussian rv's with means η_i, $i = 1, 2, \ldots, n$ and identical variance equal to σ^2. Then, $Y(\mu)$ is a noncentral chi-square rv with n degrees of freedom. Its pdf is

$$f_Y(y) = \frac{1}{2\sigma^2}(\frac{y}{s^2})^{\frac{n-2}{4}} e^{-\frac{y+s^2}{2\sigma^2}} I_{\frac{n}{2}-1}(\frac{s}{\sigma^2}\sqrt{y}), \ y \geq 0$$

and $f_Y(y) = 0$ for $y < 0$, where $s^2 = \sum_{i=1}^{n} \eta_i^2$ is called the noncentrality parameter of the distribution with $s > 0$. In the above formulation, $I_\alpha(x)$ is the α^{th}-order modified Bessel function of the first kind, which can be represented by the infinite series

$$I_\alpha(x) = \sum_{k=0}^{\infty} \frac{(\frac{x}{2})^{\alpha+2k}}{k!\Gamma(\alpha+k+1)}, \ x \geq 0$$

with α an integer. The cdf of $Y(\mu)$ is

$$F_Y(y) = 1 - Q_{\frac{n}{2}}(\frac{s}{\sigma}, \frac{\sqrt{y}}{\sigma}), \ y \geq 0$$

where $Q_m(a,b)$ is called the Marcum Q-function, defined by

$$\begin{aligned} Q_m(a,b) &= \int_b^\infty x(\frac{x}{a})^{m-1} e^{-\frac{x^2+a^2}{2}} I_{m-1}(ax)dx \\ &= e^{-\frac{1}{2}(a-b)^2} \sum_{i=1-m}^{\infty} \left(\frac{a}{b}\right)^i e^{-ab} I_i(ab), \text{ for } a < b \\ &= 1 - e^{-\frac{1}{2}(a-b)^2} \sum_{i=m}^{\infty} \left(\frac{b}{a}\right)^i e^{-ab} I_i(ab), \text{ for } a \geq b. \end{aligned}$$

See 'D. A. Shnidmen,"The calculation of the probability of detection and the generalized Marcum Q-function," *IEEE Transactions on Information Theory*, vol. 35, pp. 389-400, March 1989.' for efficient computation of $Q_m(a,b)$.
The characteristic function of $Y(\mu)$ is

$$\Phi_Y(\omega) = \frac{1}{(1-j2\sigma^2\omega)^{\frac{n}{2}}} \exp\{\frac{j\omega s^2}{1-j2\omega\sigma^2}\}.$$

The first two moments and variance of $Y(\mu)$ are

$$\begin{aligned} E\{Y(\mu)\} &= n\sigma^2 + s^2 \\ E\{Y^2(\mu)\} &= 2n\sigma^4 + 4\sigma^2 s^2 + (n\sigma^2 + s^2)^2 \\ \sigma_Y^2 &= 2n\sigma^4 + 4\sigma^2 s^2. \end{aligned}$$

When $\eta_i = 0, i = 1, 2, \ldots, n$, $Y(\mu)$ is called (central) chi-square (or gamma) rv with n degrees of freedom. Its pdf, cdf, characteristic function, first two moments and variance are

$$f_Y(y) = \begin{cases} \frac{1}{\sigma^n 2^{\frac{n}{2}} \Gamma(\frac{n}{2})} y^{\frac{n}{2}-1} e^{-\frac{y}{2\sigma^2}}, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

$$F_Y(y) = \begin{cases} 1 - e^{-\frac{y}{2\sigma^2}} \sum_{k=0}^{\frac{n}{2}-1} \frac{1}{k!}(\frac{y}{2\sigma^2})^k, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

$$\Phi_Y(\omega) = \frac{1}{(1-j2\omega\sigma^2)^{\frac{n}{2}}}$$

$$E\{Y(\mu)\} = n\sigma^2$$

$$E\{Y^2(\mu)\} = 2n\sigma^4 + n^2\sigma^4$$

$$Var\{Y(\mu)\} = 2n\sigma^4.$$

- Rayleigh Distribution
The rv $R(\mu)$ is Rayleigh distributed if $R(\mu) = \sqrt{Y(\mu)}$ where $Y(\mu)$ is a chi-square rv with two degrees of freedom. It has the following statistical characteristics:

$$\begin{aligned}
f_R(r) &= \begin{cases} \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, & r \geq 0 \\ 0, & r < 0 \end{cases} \\
F_R(r) &= \begin{cases} 1 - e^{-\frac{r^2}{2\sigma^2}}, & r \geq 0 \\ 0, & r < 0 \end{cases} \\
\Phi_R(\omega) &= \int_0^\infty f_R(r) e^{j\omega r} dr \\
E\{R^k(\mu)\} &= (2\sigma^2)^{\frac{k}{2}} \Gamma(1 + \frac{k}{2}) \\
\sigma_R^2 &= (2 - \frac{\pi}{2})\sigma^2.
\end{aligned}$$

- Rician Distribution
The rv $R(\mu)$ is Rician-distributed if $R(\mu) = \sqrt{Y(\mu)}$ where $Y(\mu)$ is a noncentral chi-square rv with two degrees of freedom and noncentrality parameter s^2. Then, $R(\mu)$ is characterized by

$$f_R(r) = \begin{cases} \frac{r}{\sigma^2} e^{-\frac{1}{2\sigma^2}(s^2+r^2)} I_o(\frac{rs}{\sigma^2}), & r \geq 0 \\ 0, & r < 0 \end{cases}$$

$$F_R(r) = \begin{cases} 1 - Q(\frac{s}{\sigma}, \frac{r}{\sigma}), & r \geq 0 \\ 0, & r < 0 \end{cases}$$

where $Q(a,b) = Q_1(a,b)$ is a generalized Q-function (i.e., a Marcum Q-function with order one), and

$$E\{R^k(\mu)\} = (2\sigma^2)^{\frac{k}{2}} e^{-\frac{s^2}{2\sigma^2}} \Gamma(\frac{k}{2} + 1) \cdot {}_1F_1(\frac{k}{2} + 1, 1; \frac{s^2}{2\sigma^2})$$

for $k \geq 0$, where ${}_1F_1(\alpha, \beta; x)$ is the confluent hypergeometric function defined as

$${}_1F_1(\alpha, \beta; x) = \sum_{k=0}^{\infty} \frac{\Gamma(\alpha + k)\Gamma(\beta)x^k}{\Gamma(\alpha)\Gamma(\beta + k)k!}, \ \beta \neq 0, 1, 2, \ldots$$

- Multivariate Gaussian Distribution

Defn: $X_1(\mu), X_2(\mu), ..., X_N(\mu)$ are called jointly Gaussian if every of their linear combinations is marginally Gaussian, i.e., $\sum_{n=1}^N a_n X_n(\mu)$ is marginally Gaussian for all a_n's which are not zeros simultaneously.

Assume that $X_i(\mu)$, $i = 1, 2, \ldots, N$, are jointly Gaussian rv's with mean η_i, $i = 1, 2, \ldots, N$, variance σ_i^2, $i = 1, 2, \ldots, N$, and covariance $C_{ij} \triangleq E\{(X_i(\mu) - \eta_i)(X_j(\mu) - \eta_j)\}$, $i, j = 1, 2, \ldots, N$. (Clearly, $C_{ii} = \sigma_i^2$) Now, let M denote $N \times N$ covariance matrix, i.e.,

$$M = \left[\ C_{ij}\ \right].$$

Let $\underline{X}(\mu)$ denote $N \times 1$ column vector of rv's, i.e.,

$$\underline{X}(\mu) = \begin{bmatrix} X_1(\mu) \\ X_2(\mu) \\ \vdots \\ X_N(\mu) \end{bmatrix}.$$

Let $\underline{\eta}_X$ denote $N \times 1$ column vector of means $\eta_i, i = 1, 2, \ldots, N$, i.e.,

$$\underline{\eta}_X = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_N \end{bmatrix}.$$

The joint pdf of $X_i(\mu)$, $i = 1, 2, \ldots, N$, is defined as

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{det(M)}} \cdot \exp\{-\frac{1}{2}(\underline{x} - \underline{\eta}_X)^T M^{-1}(\underline{x} - \underline{\eta}_X)\}$$

with $\underline{x} = [x_1, x_2, ..., x_N]^T$, where the exponent is a quadratic form of $\underline{x} - \underline{\eta}_X$, $det(M)$ and M^{-1} are respectively the determinant and the inverse of M, and $\underline{a}^T$ is the transpose of $\underline{a}$. Its characteristic function is

$$\Phi_{\underline{X}}(\underline{\omega}) \triangleq E\{\exp\{j\underline{\omega}^T \underline{X}(\mu)\}\} = \exp\{j\underline{\eta}_X^T \underline{\omega} - \frac{1}{2}\underline{\omega}^T M \underline{\omega}\}$$

with $\underline{\omega} = [\omega_1, \omega_2, ..., \omega_N]^T$.

Notes:

1. The first two moments are sufficient to fully characterize the Gaussian random vector.

2. Recall: For n rv's, independence in pairs does not necessarily imply mutual independence.
3. For jointly Gaussian rv's, independence in pairs does imply mutual independence. Now, N jointly Gaussian rv's have the joint characteristic function of

$$\Phi_{\underline{X}}(\underline{\omega}) = \exp\{j\sum_{i=1}^{N}\eta_i\omega_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{N}\omega_i\omega_k C_{ik}\}.$$

Independence in pairs implies $C_{ik} = 0, i \neq k$. In this case,

$$\begin{aligned} \Phi_{\underline{X}}(\underline{\omega}) &= \exp\{j\sum_{i=1}^{N}\eta_i\omega_i - \frac{1}{2}\sum_{i=1}^{N}\omega_i^2\sigma_i^2\} \\ &= \prod_{i=1}^{N}\exp\{j\eta_i\omega_i - \frac{1}{2}\omega_i^2\sigma_i^2\} \\ &= \prod_{i=1}^{N}\Phi_{X_i}(\omega_i) \\ \Rightarrow f_{\underline{X}}(\underline{x}) &= \prod_{i=1}^{N}f_{X_i}(x_i). \end{aligned}$$

Thus, mutual independence is ensured.
In summary, for N rv's,

$$\text{independence in pairs} \quad \underset{\text{in general}}{\Longleftarrow} \quad \underset{\text{in general}}{\not\Rightarrow} \quad \underset{\text{for jointly Gaussian rv's}}{\Longrightarrow} \quad \text{mutual independence.}$$

- Reading Assignment: Papoulis and Pillai, 4th ed., Chap. 6.
- Recommended Self-Exercise: (Papoulis and Pillai, 4th ed.)

  Chap 6: 1, 6, 10, 13, 18, 26, 40.

1.12 Convergence Concept

- (Materials from Papoulis and Pillai, Chap 7)
  We introduce various convergence modes involving sequences of random variables.

- Defn: A random sequence is a sequence of rv's $X_1(\mu), X_2(\mu), \ldots$

  For a given outcome μ_0, $X_1(\mu_0), X_2(\mu_0), \ldots$ is a deterministic sequence which may or may not converge.

- Convergence Everywhere (e)

  A random sequence is said to converge everywhere to $X(\mu)$ iff, given $\epsilon > 0$, we can find a number n, such that

$$|X_k(\mu) - X(\mu)| < \epsilon \qquad\qquad \forall\mu$$

  for every $k > n$. We use $\lim_{k\to\infty} X_k(\mu) = X(\mu)$ to denote this mode.

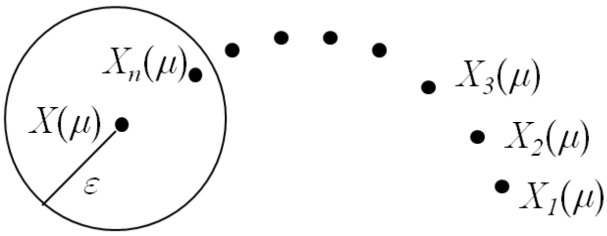


Figure 4:

- Convergence Almost Everywhere (a.e)

  Iff the set $\left\{\mu | \lim_{n\to\infty} X_n(\mu) = X(\mu)\right\}$ exists and has probability one, then we say that the sequence converges almost everywhere to $X(\mu)$ (or with probability one). We use the form

$$\Pr\left\{\lim_{n\to\infty} X_n(\mu) = X(\mu)\right\} = 1$$

  to denote it .

  Strong Law of Large Number is defined under the convergence with probability one.

- Convergence in the Mean-Square (MS) Sense

  The sequence $X_1(\mu), X_2(\mu), ...$ converges to rv $X(\mu)$ in the mean-square sense iff

$$\lim_{n\to\infty} E\left\{|X_n(\mu) - X(\mu)|^2\right\} = 0.$$

  We denote this mode by

$$\underset{n\to\infty}{l.i.m.} X_n(\mu) = X(\mu)$$

where $l.i.m.$ means "limit in the mean square sense".

Karhunen-Loéve Expansion is defined under the convergence in the mean-square sense.

- Convergence in Probability (p)

  Iff

  $$\lim_{n\to\infty} \Pr\{|X_n(\mu) - X(\mu)| > \epsilon\} = 0$$

  for any $\epsilon > 0$, then the random sequence $X_1(\mu), X_2(\mu), ...$ is said to converge to $X(\mu)$ in probability.

  Weak Law of Large Number is defined under the convergence in probability.

- Convergence in Distribution (d)

  Iff

  $$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

  for every point x of continuity of $F_X(x)$, then the random sequence $X_1(\mu), X_2(\mu), ...$ is said to converge in distribution to $X(\mu)$.

  Central Limit Theorem is defined under the convergence in distribution.

- Note: The relationship among various convergence modes is depicted by

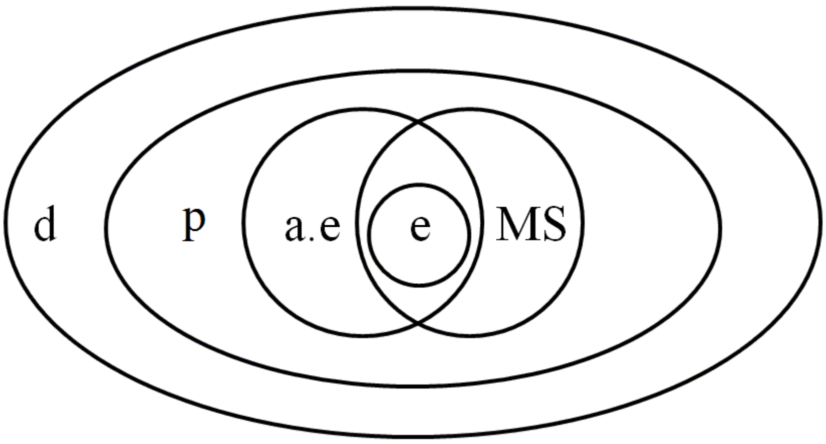


Figure 5:

i.e., e $\Rightarrow$ a.e $\Rightarrow$ p $\Rightarrow$ d

e $\Rightarrow$ MS $\Rightarrow$ p $\Rightarrow$ d

- Reading Assignment: Papoulis and Pillai, 4th ed., Chap. 7.
- Recommended Self-Exercise: Papoulis and Pillai, 4th ed., Chap 7: 5(a), 7, 13, 16, 17.
- Similar concept exists for the equivalence of random variables:

  1. $X(\mu)$ and $Y(\mu)$ are equal everywhere iff $X(\mu) = Y(\mu)$ for every outcome μ.
  2. $X(\mu)$ and $Y(\mu)$ are equal almost everywhere (or with probability one) iff $\Pr\{X(\mu) = Y(\mu)\} = 1$.
  3. $X(\mu)$ and $Y(\mu)$ are equal in the mean square sense iff $E\{|X(\mu) - Y(\mu)|^2\} = 0$.
  4. $X(\mu)$ and $Y(\mu)$ are equal in probability iff $\Pr\{|X(\mu) - Y(\mu)| > \epsilon\} = 0$ for any $\epsilon > 0$.
  5. $X(\mu)$ and $Y(\mu)$ are equal in distribution iff $F_X(x) = F_Y(x)$ for every point x of continuity of $F_X(x)$ and $F_Y(x)$.

2 Introduction to Random Process

(Material from Papoulis and Pillai, Chap. 10)

- Recall: An rv $X(\mu)$ is a mapping, or rule, for assigning to every experiment outcome μ a number $X(\mu)$.

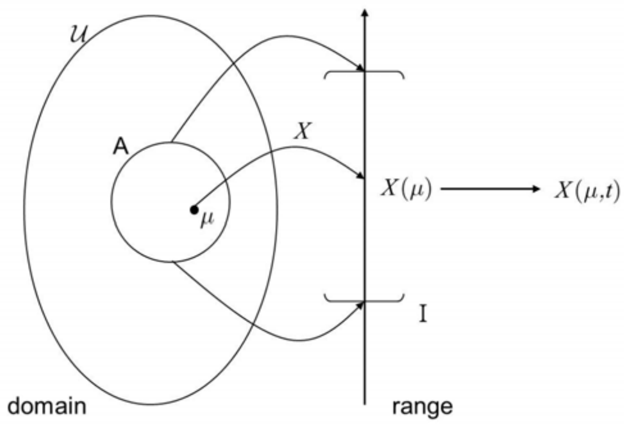


Figure 6:

- Defn: A random process (rp) $X(\mu, t)$ is a mapping from a probability space $(\mathcal{U}, \mathcal{F}, P)$ to a set of functions indexed by t. Therefore, an rp $X(\mu, t)$ is a family of functions, indexed by a parameter t (may or may not be time), depending on the experimental outcome μ, or equivalently, a function of μ and t. The domain of μ is the certain event $\mathcal{U}$ and the domain of t is a set of real numbers $\mathcal{L}$. We treat t as a timing parameter here. If $\mathcal{L}$ is the real axis, $X(\mu, t)$ is called a continuous-time random process. If $\mathcal{L}$ is the set of integers, then $X(\mu, t)$ is called a discrete-time random process. In this part, we consider only the continuous-time random processes.

2.1 Interpretations of an RP $X(\mu, t)$

1. $X(\mu, t)$ represents a family of functions of t and μ.
2. For a fixed outcomes μ_0, $X(\mu_0, t)$ is a sample time function.
3. For a fixed time t_0, $X(\mu, t_0)$ is an rv.
4. For fixed t_0 and μ_0, $X(\mu_0, t_0)$ is a number.

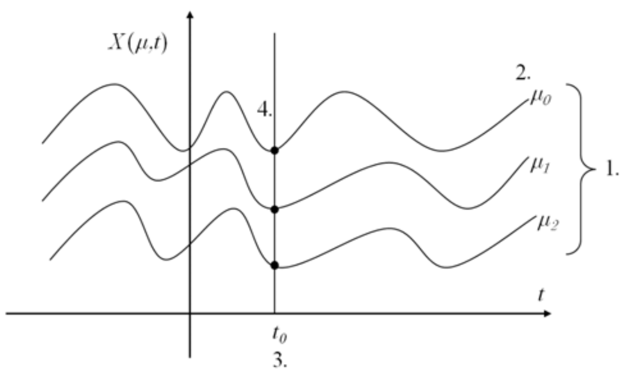


Figure 7:

- Ex: $X(\mu, t) = r(\mu)\cos(\omega t + \phi(\mu))$ is a random process (waveform) where the amplitude $r(\mu)$ and phase $\phi(\mu)$ are rv's. In this case, $X(\mu, t)$ consists of a family of pure cosinusoidal waves. For an outcome μ_0, its sample function is $X(\mu_0, t) = r(\mu_0)\cos(\omega t + \phi(\mu_0))$.

2.2 Statistics of Continuous-Time Random Processes

- First-Order Statistics

  For a specific t, $X(\mu, t)$ is an rv with "first-order distribution"

  $$F_X(x;t) \triangleq \Pr\{X(\mu,t) \le x\}$$

  and "first-order density"

  $$f_X(x;t) \triangleq \frac{\partial F_X(x;t)}{\partial x}$$

  provided that $F_X(x;t)$ is continuous and first-order differentiable in x, i.e., $X(\mu, t)$ is a continuous rv for any t.

- Second-Order Statistics

  For any two $t_1$and t_2, the second-order distribution of $X(\mu, t)$ is the joint distribution of rv's $X(\mu, t_1)$ and $X(\mu, t_2)$, i.e.,

  $$F_X(x_1, x_2; t_1, t_2) \triangleq \Pr\{X(\mu, t_1) \le x_1, X(\mu, t_2) \le x_2\}.$$

  The second-order density of $X(\mu, t)$ is thus defined by

  $$f_X(x_1, x_2; t_1, t_2) \triangleq \frac{\partial^2}{\partial x_1 \partial x_2} F_X(x_1, x_2; t_1, t_2)$$

  provided that $F_X(x_1, x_2; t_1, t_2)$ is continuous and first-order differentiable in x_1 and x_2, i.e., $X(\mu, t_1)$ and $X(\mu, t_2)$ are continuous rv's for any t_1 and t_2.

- Notes:

  1. The first-order distribution is related to the second-order distribution by

     $$F_X(x_1; t_1) = F_X(x_1, \infty; t_1, t_2).$$

  2. The first-order density is related to the second-order density by

     $$\begin{aligned} f_X(x_1; t_1) &= \frac{\partial}{\partial x_1} F_X(x_1, \infty; t_1, t_2) \\ &= \frac{\partial}{\partial x_1} \int_{-\infty}^{\infty} \int_{-\infty}^{x_1} f_X(x_1', x_2'; t_1, t_2) dx_1' dx_2' \\ &= \int_{-\infty}^{\infty} f_X(x_1, x_2'; t_1, t_2) dx_2'. \end{aligned}$$

     (A property as defined in Joint Statistics of Two RV's.)

- N-th-Order Statistics

  The N-th-order distribution of $X(\mu, t)$ is the joint distribution $F_X(x_1, x_2, ..., x_N; t_1, t_2, ..., t_N)$ of the rv's $X(\mu, t_1), X(\mu, t_2), ..., X(\mu, t_N)$, and its density can be defined accordingly. To determine the statistical property of an rp $X(\mu, t)$, knowledge of the function $F_X(x_1, x_2, ..., x_N; t_1, t_2, ..., t_N)$ is required for every x_i's , t_i's and N.

  For many applications in engineering, only the first two order statistics are used to express the underlined process.

- In what follows, we consider continuous-time continuous-valued random processes, unless mentioned otherwise.

2.3 First-Order and Second-Order Properties

- Mean Function: The mean function $\eta_X(t)$ of $X(\mu, t)$ is the expected value of the rv $X(\mu, t)$ for a fixed t, i.e.,

$$\eta_X(t) \triangleq E\{X(\mu, t)\} = \int_{-\infty}^{\infty} x f_X(x; t) dx.$$

- Autocorrelation Function: The autocorrelation $R_X(t_1, t_2)$ of real-valued $X(\mu, t)$ is the expected value of the product $X(\mu, t_1)X(\mu, t_2)$, i.e.,

$$\begin{aligned} R_X(t_1, t_2) &\triangleq E\{X(\mu, t_1)X(\mu, t_2)\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_X(x_1, x_2; t_1, t_2) dx_1 dx_2. \end{aligned}$$

  The value $R_X(t, t)$ is called the average power of real-valued $X(\mu, t)$, which equals to

$$R_X(t, t) = E\{X^2(\mu, t)\}$$

  i.e., the second moment of the real-valued rv $X(\mu, t)$ for a fixed t.

- Autocovariance Function: The autocovariance $C_X(t_1, t_2)$ of real-valued $X(\mu, t)$ is the covariance of rv's $X(\mu, t_1)$ and $X(\mu, t_2)$, i.e.,

$$\begin{aligned} C_X(t_1, t_2) &\triangleq E\{(X(\mu, t_1) - \eta_X(t_1))(X(\mu, t_2) - \eta_X(t_2))\} \\ &= R_X(t_1, t_2) - \eta_X(t_1)\eta_X(t_2). \end{aligned}$$

  The value $C_X(t, t) = Var\{X(\mu, t)\}$ is the variance of real-valued rv $X(\mu, t)$ for a fixed t.

- Ex: $X(\mu,t) = r(\mu)\cos(\omega t + \phi(\mu))$ is an rp where $r(\mu)$ and $\phi(\mu)$ are independent, and where $\phi(\mu)$ is uniform in $[-\pi,\pi)$. Then, the mean function $\eta_X(t)$ is

$$\begin{aligned}\eta_X(t) &= E\{X(\mu,t)\} \\ &= E\{r(\mu)\cos(\omega t + \phi(\mu))\} \\ &= E\{r(\mu)\}E\{\cos(\omega t + \phi(\mu))\} \qquad (\because \text{ independent}) \\ &= 0\end{aligned}$$

since

$$E\{\cos(\omega t + \phi(\mu))\} = \int_{-\pi}^{\pi} \frac{1}{2\pi}\cos(\omega t + \phi)d\phi \ = 0.$$

The autocorrelation $R_X(t_1,t_2)$ is

$$\begin{aligned}R_X(t_1,t_2) &= E\{X(\mu,t_1)X(\mu,t_2)\} \\ &= E\{r^2(\mu)\cos(\omega t_1 + \phi(\mu))\cos(\omega t_2 + \phi(\mu))\} \\ &= E\{r^2(\mu)\}E\{\cos(\omega t_1 + \phi(\mu))\cos(\omega t_2 + \phi(\mu))\} \ (\because \text{ independent}).\end{aligned}$$

Now ,

$$\begin{aligned}& E\{\cos(\omega t_1 + \phi(\mu))\cos(\omega t_2 + \phi(\mu))\} \\ = \ & \frac{1}{2}E\{\cos(\omega(t_1+t_2) + 2\phi(\mu)) + \cos(\omega(t_1 - t_2))\} \\ = \ & \frac{1}{2}\cos(\omega(t_1 - t_2)).\end{aligned}$$

Thus ,

$$R_X(t_1,t_2) = \frac{1}{2}E\{r^2(\mu)\}\cos(\omega(t_1-t_2)).$$

Also, the autocovariance is $C_X(t_1,t_2) = R_X(t_1,t_2)$. The average power of $X(\mu,t)$ is

$$E\{X^2(\mu,t)\} = R_X(t,t) = \frac{1}{2}E\{r^2(\mu)\}.$$

2.4 Complex-Valued Random Processes

- The joint statistic of two real-valued rp's $X(\mu,t)$ and $Y(\mu,t)$ is determined in terms of the joint distribution of the rv's $X(\mu,t_1), X(\mu,t_2), ...,$ $X(\mu,t_n)$ and $Y(\mu,t'_1), Y(\mu,t'_2), ..., Y(\mu,t'_m)$ for all t_i's, t'_j's, n and m. The complex-valued rp

$$Z(\mu,t) = X(\mu,t) + jY(\mu,t)$$

is then specified by this joint statistic.

- The autocorrelation of a complex-valued rp $Z(\mu, t)$ is by definition

$$R_Z(t_1, t_2) = E\{Z(\mu, t_1)Z^*(\mu, t_2)\}$$

where the conjugate operator is associated with the second variable in $R_Z(t_1, t_2)$.

Therefore,

$$\begin{aligned} R_Z(t_2, t_1) &= E\{Z(\mu, t_2)Z^*(\mu, t_1)\} \\ &= E^*\{Z(\mu, t_1)Z^*(\mu, t_2)\} \\ &= R_Z^*(t_1, t_2) \end{aligned}$$

(i.e., Hermitian symmetry is guaranteed).

Furthermore, we notice that

$$R_Z(t, t) = E\{|Z(\mu, t)|^2\} \equiv \text{average power of } Z(\mu, t) \geq 0.$$

- Properties :

1. The autocorrelation of an rp $Z(\mu, t)$ is nonnegative definite, i.e., for any a_i's,

$$\begin{aligned} \sum_i \sum_k a_i a_k^* R_Z(t_i, t_k) &= \sum_i \sum_k E\{a_i a_k^* Z(\mu, t_i)Z^*(\mu, t_k)\} \\ &= E\{\left|\sum_i a_i Z(\mu, t_i)\right|^2\} \geq 0. \end{aligned}$$

2. The autocovariance of an rp $Z(\mu, t)$ is nonnegative definite (which can be similarly proved).
3. Given a nonnegative definite function $R_Z(t_1, t_2)$, one can find an rp $Z(\mu, t)$ with autocorrelation $R_Z(t_1, t_2)$.

- The autocovariance of a complex-valued rp $Z(\mu, t)$ is by definition

$$\begin{aligned} C_Z(t_1, t_2) &\triangleq E\{(Z(\mu, t_1) - E\{Z(\mu, t_1)\})(Z(\mu, t_2) - E\{Z(\mu, t_2)\})^*\} \\ &= R_Z(t_1, t_2) - \eta_Z(t_1)\eta_Z^*(t_2) \end{aligned}$$

with the mean $\eta_Z(t) = E\{Z(\mu, t)\}$. The correlation coefficient of $Z(\mu, t)$ is defined by

$$r_Z(t_1, t_2) = \frac{C_Z(t_1, t_2)}{\sqrt{C_Z(t_1, t_1)C_Z(t_2, t_2)}}.$$

Note: $C_Z(t,t) = E\{|Z(\mu,t) - E\{Z(\mu,t)\}|^2\} \geq 0$.

Since $C_Z(t_1,t_2)$ is the autocorrelation of a new rp $Z(\mu,t) - E\{Z(\mu,t)\}$, it is nonnegative definite. Also, since $r_Z(t_1,t_2)$ is the autocorrelation of a new rp

$$\frac{Z(\mu,t) - E\{Z(\mu,t)\}}{\sqrt{C_Z(t,t)}}$$

it is also nonnegative definite. Note also that $|r_Z(t_1,t_2)| \leq 1$ and $r_Z(t,t) = 1$.

- The cross-correlation of two complex-valued rp's $X(\mu,t)$ and $Y(\mu,t)$ is defined by

$$R_{X,Y}(t_1,t_2) = E\{X(\mu,t_1)Y^*(\mu,t_2)\} = R^*_{Y,X}(t_2,t_1).$$

The cross-covariance of two complex-valued rp's $X(\mu,t)$ and $Y(\mu,t)$ is

$$\begin{aligned} C_{X,Y}(t_1,t_2) &\triangleq E\{(X(\mu,t_1) - \eta_X(t_1))(Y(\mu,t_2) - \eta_Y(t_2))^*\} \\ &= R_{X,Y}(t_1,t_2) - \eta_X(t_1)\eta^*_Y(t_2). \end{aligned}$$

- $X(\mu,t)$ and $Y(\mu,t)$ are called mutually orthogonal iff $R_{X,Y}(t_1,t_2) = 0$ for all t_1 and t_2. They are called uncorrelated iff $C_{X,Y}(t_1,t_2) = 0$ for all t_1 and t_2. We call a process $Z(\mu,t)$ as white noise iff $C_Z(t_1,t_2) = \alpha\delta(t_1-t_2)$ where α is a positive constant and $\delta(t)$ is Dirac delta function.

- $X(\mu,t)$ and $Y(\mu,t)$ are called independent iff $\{X(\mu,t_1), X(\mu,t_2), ..., X(\mu,t_n)\}$ and $\{Y(\mu,t'_1), Y(\mu,t'_2), ..., Y(\mu,t'_m)\}$ are two independent sets of random variables for all t_i's, t'_k's, n and m.

- $X(\mu,t)$ is called independent iff $X(\mu,t_1), X(\mu,t_2), ..., X(\mu,t_n)$ are independent for all t_i's and n.

- $X(\mu,t)$ is called normal or Gaussian iff the rv's $X(\mu,t_1), X(\mu,t_2), ...,$ $X(\mu,t_n)$ are jointly normal or jointly Gaussian for all t_i's and n.

- Iff $X(\mu,t)$ is both Gaussian and white, we call this process a white Gaussian noise (WGN).

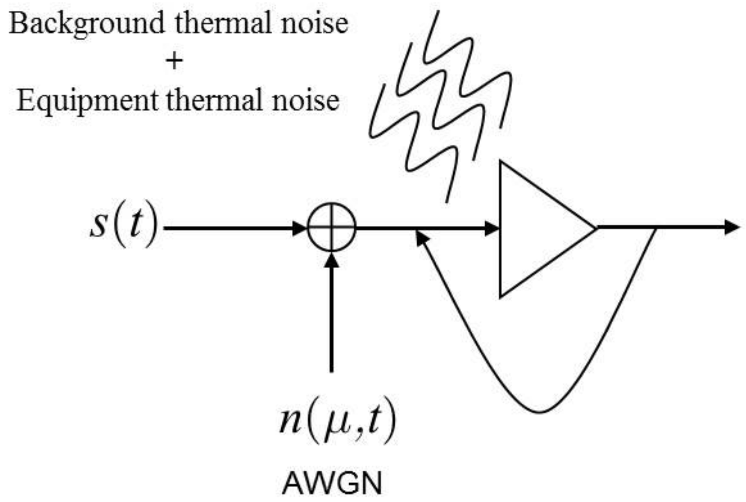


Figure 8:

- The statistic of a real-valued Gaussian rp $X(\mu, t)$ is completely determined by its mean function $\eta_X(t)$ and autocovariance $C_X(t_1, t_2)$. This is because the n^{th}-order characteristic function of $X(\mu, t)$ is defined and given by

$$E\{\exp\{j\sum_{i=1}^{n}\omega_i X(\mu, t_i)\}\} = \exp\{j\sum_{i=1}^{n}\eta_X(t_i)\omega_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{n}\omega_i\omega_k C_X(t_i, t_k)\}$$

which will uniquely determine its n-th order density by an n-dimensional inverse Fourier transform.

- Notice that all the above discussion in this section can also be applied to the real-valued rp's.

- Common Properties of $R_X(t_1, t_2)$, $C_X(t_1, t_2)$ (i.e., Autocorrelation of $X(\mu, t) - \eta_X(t)$) and $r_X(t_1, t_2)$ (i.e., Autocorrelation of $\frac{X(\mu, t) - \eta_X(t)}{\sqrt{C_X(t, t)}}$)

  1. $R_X(t_1, t_2)$ must be Hermitian symmetric, i.e., $R_X(t_1, t_2) = R_X^*(t_2, t_1)$.
  2. $R_X(t_1, t_2)$ must be nonnegative definite, i.e., for any a_i's and t_i's,

$$\sum_i \sum_k a_i a_k^* R_X(t_i, t_k) \geq 0.$$

3. $R_X(t_1, t_2)$ satisfies the Schwartz's inequality, i.e.,

$$|R_X(t_1, t_2)| \le \sqrt{R_X(t_1, t_1)}\sqrt{R_X(t_2, t_2)}.$$

Proof: The proof is given only for real-valued rp's. Now, we have to show that

$$R_X^2(t_1, t_2) \le R_X(t_1, t_1)R_X(t_2, t_2).$$

Consider that $E\{[X(\mu, t_1)+aX(\mu, t_2)]^2\} \ge 0$ for any $a \in R$. Thus, for any $a \in R$,

$$\begin{aligned} & E\{X^2(\mu, t_1)\} + 2E\{X(\mu, t_1)X(\mu, t_2)\}a + E\{X^2(\mu, t_2)\}a^2 \\ \Rightarrow\ & R_X(t_1, t_1) + 2R_X(t_1, t_2)a + R_X(t_2, t_2)a^2 \ge 0 \\ \Rightarrow\ & \text{Discriminant} = 4R_X^2(t_1, t_2) - 4R_X(t_1, t_1)R_X(t_2, t_2) \le 0 \\ \Rightarrow\ & R_X^2(t_1, t_2) \le R_X(t_1, t_1)R_X(t_2, t_2). \quad Q.E.D. \end{aligned}$$

You try to prove it for complex-valued rp's.

- For a complex-valued rp $Z(\mu, t)$, $E\{Z(\mu, t_1)Z^*(\mu, t_2)\}$ is nonnegative definite, but $E\{Z(\mu, t_1)Z(\mu, t_2)\}$ is not. This is why $E\{Z(\mu, t_1)Z(\mu, t_2)\}$ is not used to define the autocorrelation of $Z(\mu, t)$.

2.5 Stationary Processes

- Strict-Sense Stationary (SSS) Processes
  An rp $X(\mu, t)$ is called SSS iff rp's $X(\mu, t)$ and $X(\mu, t + c)$ have the same statistic for any c.
  Two rp's $X(\mu, t)$ and $Y(\mu, t)$ are called jointly SSS iff the joint statistic of rp's $X(\mu, t)$ and $Y(\mu, t)$ is the same as the joint statistic of $X(\mu, t+c)$ and $Y(\mu, t + c)$ for any c.

  A complex-valued rp $Z(\mu, t) = X(\mu, t) + jY(\mu, t)$ is SSS iff $X(\mu, t)$ and $Y(\mu, t)$ are jointly SSS.

- Properties

  1. The n^{th}-order density of a real-valued SSS rp $X(\mu, t)$ satisfies

  $$\begin{aligned} & f_X(x_1, x_2, \ldots, x_n; t_1, t_2, \ldots, t_n) \\ = \ & f_X(x_1, x_2, \ldots, x_n; t_1 + c, t_2 + c, \ldots, t_n + c) \end{aligned}$$

  for any real c. Therefore, this $X(\mu, t)$ has the following properties.

2. For a real-valued SSS rp $X(\mu, t)$,

$$\begin{aligned} f_X(x;t) &= f_X(x;t+c) \text{ for any } c \\ &= f_X(x) \end{aligned}$$

which is independent of t. Thus, the 1^{st}-order statistic is independent of time. Now,

$$\begin{aligned} \eta_X(t) &= \int_{-\infty}^{\infty} x f_X(x;t)dx \\ &= \int_{-\infty}^{\infty} x f_X(x)dx \end{aligned}$$

which is independent of t. Thus, the mean function is independent of time.

3. For a real-valued SSS rp $X(\mu, t)$,

$$\begin{aligned} f_X(x_1, x_2; t_1, t_2) &= f_X(x_1, x_2; t_1 + c, t_2 + c) \text{ for any } c \\ &= f_X(x_1, x_2; \tau) \end{aligned}$$

which depends only on τ, with $\tau = t_1 - t_2$. Thus, the 2^{nd}-order statistic is only dependent on the time difference but independent of any specific times.

The following discussion applies to both real-valued and complex-valued rp's.

- Wide-Sense Stationary (WSS) Processes

  $X(\mu, t)$ is WSS iff

  (a) $\eta_X(t) = \text{constant}$,

  (b) $E\{|X(\mu, t)|^2\} < \infty$,

  (c) $E\{X(\mu, t+\tau)X^*(\mu, t)\} = R_X(\tau)$ (i.e., a function of time difference only).

- Notice that SSS implies WSS, but the converse does not hold in general. However, for a Gaussian random process, WSS implies SSS. This is because the n^{th}-order characteristic function of a real-valued Gaussian WSS random process has a form of

$$\exp\{j\eta_X \sum_{i=1}^{n} \omega_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} C_X(t_i - t_k)\omega_i\omega_k\}$$

which is invariant to a shift in time. The same argument can be applied to complex-valued Gaussian WSS random processes.

Ex A: If a real-valued random process $X(\mu, t)$ is WSS and $I(\mu, T) = \int_{-T}^{T} X(\mu, t)dt$, then

$$Var\{I(\mu, T)\} = \int_{-2T}^{2T} (2T - |\tau|)C_X(\tau)d\tau.$$

Proof: By definition,

$$\begin{aligned}
Var\{I(\mu, T)\} &= Var\{\int_{-T}^{T} X(\mu, t)dt\} \\
&= E\{(\int_{-T}^{T} X(\mu, t)dt)^2\} - (\int_{-T}^{T} \eta_X dt)^2 \\
&= \int_{-T}^{T}\int_{-T}^{T} E\{X(\mu, t_1)X(\mu, t_2)\}dt_1dt_2 - \int_{-T}^{T}\int_{-T}^{T} \eta_X^2 dt_1dt_2 \\
&= \int_{-T}^{T}\int_{-T}^{T} [E\{X(\mu, t_1)X(\mu, t_2)\} - \eta_X^2]dt_1dt_2 \\
&= \int_{-T}^{T}\int_{-T}^{T} C_X(t_2 - t_1)dt_1dt_2.
\end{aligned}$$

Substituting the variable $\tau = t_2 - t_1$ into the above integral, i.e., $t_2 = \tau + t_1$.

$$\begin{aligned}
&\Rightarrow \quad -T \leq \tau + t_1 \leq T \Rightarrow -T - \tau \leq t_1 \leq T - \tau \\
&\Rightarrow \quad \max\{-T, -T - \tau\} \leq t_1 \leq \min\{T, T - \tau\}
\end{aligned}$$

we have

$$Var\{I(\mu, T)\} = \int_{-2T}^{2T}\int_{\max\{-T, -T-\tau\}}^{\min\{T, T-\tau\}} C_X(\tau)dt_1d\tau.$$

Because

$$\begin{aligned}
\int_{\max\{-T, -T-\tau\}}^{\min\{T, T-\tau\}} dt_1 &= \begin{cases} \int_{-T}^{T-\tau} dt_1 = 2T - \tau, & \text{if } \tau \geq 0 \\ \int_{-T-\tau}^{T} dt_1 = 2T + \tau, & \text{if } \tau < 0 \end{cases} \\
&= 2T - |\tau|
\end{aligned}$$

we end up with

$$Var\{I(\mu,T)\} = \int_{-2T}^{2T} C_X(\tau)(2T - |\tau|)d\tau. \quad Q.E.D.$$

- A real-valued random process $X(\mu, t)$ is called the N^{th}-order stationary iff

$$\begin{aligned} & f_X(x_1, x_2, ..., x_n; t_1, t_2, ..., t_n) \\ = \; & f_X(x_1, x_2, ..., x_n; t_1 + c, t_2 + c, ..., t_n + c) \end{aligned}$$

for any x_i's, t_i's, and for any c, only holds for $n \leq N$.

- Reading Assignment: Papoulis and Pillai, Section 10.1.

- In practice, we can estimate $\eta_X(t)$ of a random process $X(\mu, t)$ by the following.

  1. For a fixed t, if we can observe a large number of trial samples, say $X(\mu^1, t), X(\mu^2, t), ...,$ then we can use their *average* as the estimate, i.e.,

  $$\eta_X(t) \approx \frac{1}{n}\sum_{i=1}^{n} X(\mu^i, t)$$

     where μ^i represents the i-th trial result of the underlined random experiment. For an extremely large set of samples, this provides an adequate estimate. Note that a true *ensemble average* (or *statistical average*) is given by

  $$\eta_X(t) = \sum_{\text{all outcomes } \mu_i} X(\mu_i, t)P(\{\mu_i\}).$$

  2. If we can only obtain one sample function $X(\mu_0, t)$ for a fixed μ_0, then we "may" use its *time average*

  $$\overline{X} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} X(\mu_0, t)dt$$

     as the estimate.

     Note: If $\eta_X(t)$ is not a constant, then the time average is not adequate as such an estimate.

- Defn: A random process $X(\mu, t)$ is called ergodic (in a certain sense) iff its ensemble average equals appropriate time average.

  This means that a certain statistic of $X(\mu, t)$ can be determined from a sample function, in a certain probabilistic sense.

- Consider real-valued random processes below.

- Defn: Mean-Ergodic

  An rp $X(\mu, t)$ is said to be mean-ergodic iff

  1. $E\{X(\mu, t)\} = \eta_X$, a constant.
  2. In the mean square sense,

  $$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X(\mu, t)dt = \eta_X \qquad \circledast$$

  (i.e., $\lim\limits_{T\to\infty} E\{|\frac{1}{2T}\int_{-T}^{T} X(\mu, t)dt - \eta_X|^2\} = 0$).

  Since $\frac{1}{2T}\int_{-T}^{T} X(\mu, t)dt$ is an rv, $\circledast$ can only be true iff

  $$\lim_{T\to\infty} Var\{\frac{1}{2T} \int_{-T}^{T} X(\mu, t)dt\} = 0. \qquad \oplus$$

  Therefore, to test whether $X(\mu, t)$ is mean-ergodic or not, it suffices to test $E\{X(\mu, t)\} = \eta_X$ and $\oplus$.

- Theorem: An rp $X(\mu, t)$ with $E\{X(\mu, t)\} = \eta_X$ is mean-ergodic iff $\lim\limits_{T\to\infty} \frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T} C_X(t_1, t_2)\ dt_1 dt_2 = 0$.

  Proof: Obvious since

  $$Var\{\frac{1}{2T} \int_{-T}^{T} X(\mu, t)dt\} \underset{\substack{\uparrow \\ \text{see Ex A}}}{=} \frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T} C_X(t_1, t_2)dt_1 dt_2.$$

  Q.E.D.

- Corollary: A WSS process $X(\mu, t)$ is mean-ergodic iff its autocovariance $C_X(\tau) = R_X(\tau) - \eta_X^2$ satisfies

  $$\lim_{T\to\infty} \frac{1}{2T} \int_{-2T}^{2T} C_X(\tau)(1 - \frac{|\tau|}{2T})d\tau = 0. \qquad \circledcirc$$

  Proof: Check Ex A and apply the result to the above Theorem.

- There are two sufficient conditions for mean-ergodicity:

  1. If $X(\mu, t)$ is WSS and $\int_{-\infty}^{\infty} |C_X(\tau)| d\tau < \infty$, then $X(\mu, t)$ is mean-ergodic.

     Proof: From the above corollary,

     $$\left| \frac{1}{2T} \int_{-2T}^{2T} C_X(\tau)(1 - \frac{|\tau|}{2T}) d\tau \right| < \frac{1}{2T} \int_{-2T}^{2T} |C_X(\tau)| d\tau.$$

     Q.E.D.

  2. If $X(\mu, t)$ is WSS, $C_X(0) < \infty$ and $\lim_{|\tau| \to \infty} C_X(\tau) = 0$, then $X(\mu, t)$ is mean-ergodic.

     Proof: Since $\lim_{|\tau| \to \infty} C_X(\tau) = 0$, we can find a constant a for any given $\epsilon > 0$, such that $|C_X(\tau)| < \epsilon$ for every $|\tau| > a$. Therefore, for $2T > a$,

     $$\begin{aligned} & \left| \int_{-2T}^{2T} C_X(\tau)(1 - \frac{|\tau|}{2T}) d\tau \right| \\ \leq & \int_{-a}^{a} |C_X(\tau)(1 - \frac{|\tau|}{2T})| d\tau + \int_{a < |\tau| < 2T} |C_X(\tau)(1 - \frac{|\tau|}{2T})| d\tau \\ \leq & \int_{-a}^{a} \underbrace{|C_X(\tau)|}_{\substack{\leq C_X(0) \\ \text{by Schwartz's inequality}}} d\tau + \int_{a < |\tau| < 2T} \underbrace{|C_X(\tau)|}_{\leq \epsilon} d\tau \\ \leq & 2aC_X(0) + 2\epsilon(2T - a). \end{aligned}$$

     This gives

     $$\begin{aligned} & \lim_{T \to \infty} \frac{1}{2T} \left| \int_{-2T}^{2T} C_X(\tau)(1 - \frac{|\tau|}{2T}) d\tau \right| \\ \leq & \lim_{T \to \infty} \frac{1}{2T} (2a \underbrace{C_X(0)}_{< \infty} + 2\epsilon(2T - a)) \\ = & 0 + 2\epsilon \end{aligned}$$

     for every $\epsilon > 0$. Thus, $\circledcirc$ holds. *Q.E.D.*

- Defn: Correlation-Ergodic

A WSS rp $X(\mu,t)$ is said to be correlation-ergodic iff the process $X(\mu,t+\lambda)X(\mu,t)$ is WSS and mean-ergodic for any λ, i.e., iff

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}X(\mu,t+\lambda)X(\mu,t)dt = R_X(\lambda)$$

for any λ. Form the Corollary introduced in the Mean-Ergodic case, we conclude that $X(\mu,t)$ is correlation-ergodic given that $Z(\mu,t) \triangleq X(\mu,t+\lambda)X(\mu,t)$ is WSS for any λ iff

$$C_Z(\tau) = R_Z(\tau) - R_X^2(\lambda)$$

satisfies

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-2T}^{2T}C_Z(\tau)(1-\frac{|\tau|}{2T})d\tau = 0.$$

2.6 Deterministic System with Stochastic Input/Output

- Consider $Y(\mu,t) = \mathcal{T}[X(\mu,t)]$ with $\mathcal{T}$ being the system operator.
  If $\mathcal{T}$ operates only on t with μ fixed, then the system is called deterministic.
  If $\mathcal{T}$ operates on both t and μ, it is called stochastic.

- For a deterministic system $\mathcal{T}$,
  $X(\mu_1,t) = X(\mu_2,t)$ for every $t \Rightarrow Y(\mu_1,t) = Y(\mu_2,t)$.
  However, for a stochastic system $\mathcal{T}$,
  $X(\mu_1,t) = X(\mu_2,t)$ for every $t \not\Rightarrow Y(\mu_1,t) = Y(\mu_2,t)$.

Let us consider the deterministic system in the following.

- Memoryless System

  A system is memoryless iff $Y(\mu,t) = g(X(\mu,t))$ where $g(x)$ is a function of x.

  At any specific time t_1, $Y(\mu,t_1)$ depends "only" on $X(\mu,t_1)$, and not on any other $X(\mu,t)$.

  Therefore, for a memoryless system with real input and output, the n^{th}-order density $f_Y(y_1,y_2,...,y_n;t_1,t_2,...,t_n)$ can be obtained by

$$f_Y(y_1,y_2,...,y_n;t_1,t_2,...,t_n) = \sum_{\substack{all\\roots}}\frac{f_X(x_1,x_2,...,x_n;t_1,t_2,...,t_n)}{|J|}$$

where J is the Jacobian

$$J = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \cdot & \cdot & \cdot & \frac{\partial g_1}{\partial x_n} \\ \cdot & \cdot & & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \frac{\partial g_n}{\partial x_1} & \cdot & \cdot & \cdot & \frac{\partial g_n}{\partial x_n} \end{vmatrix}$$

and for $i = 1, 2, ..., n$,

$$\begin{aligned} Y(\mu, t_i) &= g(X(\mu, t_i)) \\ &= g_i(X(\mu, t_1), X(\mu, t_2), ..., X(\mu, t_n)). \end{aligned}$$

Notes:

1. If $X(\mu, t)$ is SSS, so is $Y(\mu, t)$.
2. If $X(\mu, t)$ is stationary of order N, so is $Y(\mu, t)$.
3. If $X(\mu, t)$ is WSS, then $Y(\mu, t)$ may NOT be stationary in any sense.

Example: A square-law device with WSS input $X(\mu, t)$.

1. Consider $Y(\mu, t) = X^2(\mu, t)$, i.e., $g(x) = x^2$ a memoryless system. Now,

$$\begin{aligned} E\{Y(\mu, t)\} &= E\{X^2(\mu, t)\} = R_X(0), \text{ a constant} \\ E\{Y(\mu, t_1)Y(\mu, t_2) &= E\{X^2(\mu, t_1)X^2(\mu, t_2)\} \\ &\neq \text{a function of } t_1 - t_2 \text{ in general.} \end{aligned}$$

2. For any fixed t, the first-order density of $Y(\mu, t)$ can be derived by Jacobian approach, as given by

$$f_Y(y; t) = \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}; t) + f_X(-\sqrt{y}; t)], \ y \geqslant 0.$$

Similarly, the second-order density of $Y(\mu, t)$ is derived as

$$f_Y(y_1, y_2; t_1, t_2) = \frac{1}{4\sqrt{y_1 y_2}} \sum_{a_1, a_2 = \pm 1} f_X(a_1\sqrt{y_1}, a_2\sqrt{y_2}; t_1, t_2)$$

for $y_1, y_2 \geqslant 0$ and it is zero otherwise.

If $X(\mu,t)$ is normal and stationary with mean zero and autocorrelation $R_X(\tau)$, then

$$f_Y(y;t) = f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi R_X(0) y}} e^{-\frac{y}{2R_X(0)}}, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

and

$$\begin{aligned} R_Y(t_1,t_2) &= E\{X^2(\mu,t_1)X^2(\mu,t_2)\} \\ &= E\{X^2(\mu,t_1)\}E\{X^2(\mu,t_2)\} + 2E^2\{X(\mu,t_1)X(\mu,t_2)\} \\ &= R_X^2(0) + 2R_X^2(t_1-t_2) \\ &= R_Y(t_1-t_2). \end{aligned}$$

Note here that for a normal input $X(\mu,t)$,

$$SSS \Leftrightarrow WSS$$

which implies that the output $Y(\mu,t)$ is SSS, and thus WSS.

- Lemma: For a jointly normal $X_1(\mu)$, $X_2(\mu)$, $X_3(\mu)$, and $X_4(\mu)$ which have zero mean,

$$\begin{aligned} E\{X_1(\mu)X_2(\mu)X_3(\mu)X_4(\mu)\} &= E\{X_1(\mu)X_2(\mu)\}E\{X_3(\mu)X_4(\mu)\} \\ &\quad + E\{X_1(\mu)X_3(\mu)\}E\{X_2(\mu)X_4(\mu)\} \\ &\quad + E\{X_1(\mu)X_4(\mu)\}E\{X_2(\mu)X_3(\mu)\}. \end{aligned}$$

Proof: To be done later.

- Linear System

$Y(\mu,t) = \mathcal{T}[X(\mu,t)]$ is a linear system iff, for any n rp's $X_1(\mu,t), X_2(\mu,t), ..., X_n(\mu,t)$, and any n constants $a_1, a_2, ..., a_n$,

$$\mathcal{T}[\sum_{i=1}^{n} a_i X_i(\mu,t)] = \sum_{i=1}^{n} a_i \mathcal{T}[X_i(\mu,t)].$$

(This is called the principle of superposition).

For continuous I/O, a linear system satisfying the principle of superposition can be represented by

$$Y(\mu,t) = \int_{-\infty}^{\infty} X(\mu,\alpha) h(t,\alpha) d\alpha.$$

- Time-Invariant System

  A system is called time-invariant iff $Y(\mu, t+c) = \mathcal{T}[X(\mu, t+c)]$ for any t and c.

- Linear and Time-Invariant System

  A system is called linear and time-invariant (LTI) iff $Y(\mu, t) = \mathcal{T}[X(\mu, t)]$ is both linear and time-invariant.

  For continuous I/O, an LTI system can be represented by

  $$Y(\mu, t) = \int_{-\infty}^{\infty} X(\mu, \alpha) h(t-\alpha) d\alpha.$$

  Proof: Since an LTI system is linear,

  $$\begin{aligned} Y(\mu, t+c) &= \int_{-\infty}^{\infty} X(\mu, \alpha) h(t+c, \alpha) d\alpha \text{ for any } t \text{ and } c \\ &= \int_{-\infty}^{\infty} X(\mu, \alpha'+c) h(t+c, \alpha'+c) d\alpha' \quad (\alpha' = \alpha - c) \end{aligned}$$

  which is a system $\mathcal{T}$ with input $X(\mu, t+c)$ and output $Y(\mu, t+c)$. Now, because the system is also time-invariant,

  $$\begin{aligned} Y(\mu, t) &= \int_{-\infty}^{\infty} X(\mu, \alpha') h(t+c, \alpha'+c) d\alpha' \qquad \text{for any } t \text{ and } c \\ & \quad \text{(there are infinitely many equations satisfying the} \\ & \quad \text{above equality for a fixed } t \text{ value with } c \text{ varied)} \\ &= \int_{-\infty}^{\infty} X(\mu, \alpha') h((t+c) - (\alpha'+c)) d\alpha' \\ &= \int_{-\infty}^{\infty} X(\mu, \alpha') h(t-\alpha') d\alpha' \end{aligned}$$

  for any t. Q.E.D.

  Notes:

  1. The input/output relation can be represented simply by a convolution operator $*$, i.e.,

     $$Y(\mu, t) = X(\mu, t) * h(t)$$

     where the convolution operator is defined by

     $$\begin{aligned} x(t) * y(t) &= \int_{-\infty}^{\infty} x(t-\alpha) y(\alpha) d\alpha \\ &= \int_{-\infty}^{\infty} x(\alpha) y(t-\alpha) d\alpha. \end{aligned}$$

By this definition, $*$ is commutative (i.e., $x(t)*y(t) = y(t)*x(t)$), associative (i.e., $x(t)*(y(t)*z(t)) = (x(t)*y(t))*z(t)$), and distributive over addition/subtraction (i.e., $x(t)*(y(t) \pm z(t)) = (x(t)*y(t)) \pm (x(t)*z(t))$).

2. $h(t)$ is called the impulse response of the LTI system because

$$h(t) = \int_{-\infty}^{\infty} \delta(\alpha')h(t-\alpha')d\alpha' = \delta(t)*h(t).$$

- On Dirac Delta Function $\delta(x)$:

  (Materials come from D.S. Jones, "The Theory of Generalized Functions," Cambridge University Press, 1980)

  $\delta(x)$ is a generalized function which has the property

  $$\int_{-\infty}^{\infty} \delta(x)f(x)dx = f(0)$$

  for any suitably continuous function $f(x)$ which is continuous at $x = 0$. Note that no function in the ordinary sense has the above property. However, one can imagine a sequence of functions which have progressively taller and thinner peaks at $x = 0$, with the area under the curve remaining equal to one, while the value of the function tends to 0 at every point except $x = 0$ where it tends to infinity. For example, the limit of a sequence $g_1(x), g_2(x), ...$ with $g_n(x) = n$ for $|x| < \frac{1}{2n}$ and $g_n(x) = 0$ otherwise, can be used to imagine $\delta(x)$. The following is a list of features on $\delta(x)$:

  (1) $\delta(x) = 0$ for all $x \neq 0$, but $\delta(x)$ does not really exist at $x = 0$.

  (2) $\int_{-\varepsilon}^{\varepsilon} \delta(x)dx = 1$ if $\varepsilon > 0$.

  (3) $\int_{-\infty}^{\infty} \delta(x-x_0)f(x)dx = f(x_0)$ for any suitably continuous function $f(x)$ which is continuous at $x = x_0$.

  (4) $\int_{-\infty}^{\varepsilon} \delta(x)dx = u(\varepsilon)$ if $\varepsilon \neq 0$ where $u(\varepsilon)$ is a unit step function defined by $u(\varepsilon) = 1$ if $\varepsilon > 0$ and $u(\varepsilon) = 0$ if $\varepsilon < 0$, but $u(\varepsilon)$ does not really exist if $\varepsilon = 0$. Note that $u(\varepsilon)$ is not an ordinary function and $u(0) = 1$ is used to be defined in an ordinary sense.

  (5) $\delta(x) = \frac{du(x)}{dx}$ is a convenient notation commonly adopted. Note that the derivative of $u(x)$ does not really exist for $x = 0$.

  (6) $f(x)\delta(x-x_0) = f(x_0)\delta(x-x_0)$ can be conveniently represented for any function $f(x)$ which is continuous at $x = x_0$.

  (7) $f(x)*\delta(x-x_0) = f(x-x_0)$ for any suitably continuous function $f(x)$.

- Let us now consider the LTI system below, where $h(t)$ denotes the impulse response for an LTI system.

- Fundamental Theorem: For any linear system, $E\{\mathcal{T}[X(\mu,t)]\} = \mathcal{T}[E\{X(\mu,t)\}]$. For an LTI system, this implies

$$E\{Y(\mu,t)\} = E\{\int_{-\infty}^{\infty} X(\mu,\alpha)h(t-\alpha)d\alpha\}$$

$$\begin{aligned}\Rightarrow \eta_Y(t) &= \int_{-\infty}^{\infty} \eta_X(\alpha)h(t-\alpha)d\alpha \\ &= \eta_X(t) * h(t).\end{aligned}$$

The relation $\eta_Y(t) = \eta_X(t) * h(t)$ indicates

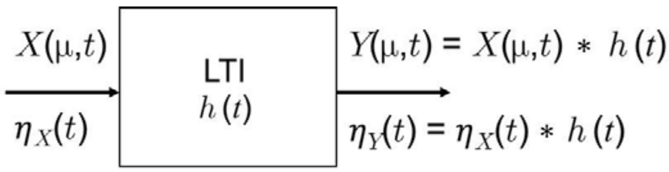


Figure 9:

- Properties: (Show them yourself)

  (1) If $Y(\mu,t) = \mathcal{T}[X(\mu,t)]$ is a time-invariant system and $X(\mu,t)$ is SSS, then so is $Y(\mu,t)$ SSS.

  (2) If $Y(\mu,t) = \mathcal{T}[X(\mu,t)]$ is an LTI system and $X(\mu,t)$ is WSS, then $Y(\mu,t)$ is also WSS.

- Furthermore, for an LTI system with real-valued $X(\mu,t)$ as input, we can determine the output autocorrelation $R_Y(t_1,t_2)$ as follows

$$\begin{aligned}R_Y(t_1,t_2) &= E\{Y(\mu,t_1)Y(\mu,t_2)\} \\ &= E\{\int_{-\infty}^{\infty} X(\mu,t_1-\alpha)h(\alpha)d\alpha \int_{-\infty}^{\infty} X(\mu,t_2-\beta)h(\beta)d\beta\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_X(t_1-\alpha,t_2-\beta)h(\alpha)h(\beta)d\alpha d\beta \\ &= \int_{-\infty}^{\infty} [R_X(t_1,t_2-\beta) * h(t_1)]h(\beta)d\beta \\ &= R_X(t_1,t_2) * h(t_1) * h(t_2).\end{aligned}$$

However, conceptually and operationally, it is preferable to find $R_Y(t_1,t_2)$ by a two-stage approach as

$$\begin{aligned}
R_Y(t_1,t_2) &= E\{Y(\mu,t_1)Y(\mu,t_2)\} \\
&= R_{YY}(t_1,t_2) \\
&= E\{Y(\mu,t_1)\int_{-\infty}^{\infty} X(\mu,t_2-\beta)h(\beta)d\beta\} \\
&= \int_{-\infty}^{\infty} R_{YX}(t_1,t_2-\beta)h(\beta)d\beta \\
&= R_{YX}(t_1,t_2) * h(t_2).
\end{aligned}$$

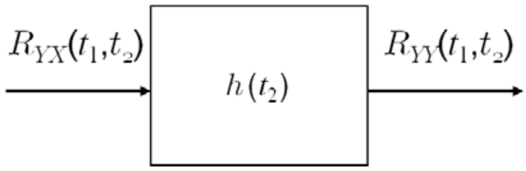


Figure 10:

Also,

$$\begin{aligned}
R_{YX}(t_1,t_2) &= E\{Y(\mu,t_1)X(\mu,t_2)\} \\
&= E\{\int_{-\infty}^{\infty} X(\mu,t_1-\alpha)h(\alpha)d\alpha X(\mu,t_2)\} \\
&= \int_{-\infty}^{\infty} R_{XX}(t_1-\alpha,t_2)h(\alpha)d\alpha \\
&= R_{XX}(t_1,t_2) * h(t_1).
\end{aligned}$$

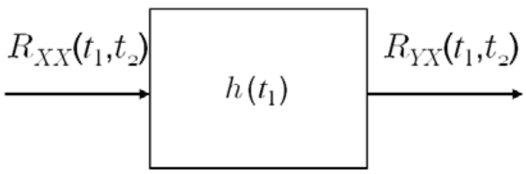


Figure 11:

That is, $R_Y(t_1,t_2) \equiv R_{YY}(t_1,t_2)$ can be determined by $R_X(t_1,t_2) \equiv R_{XX}(t_1,t_2)$ and $h(t)$ through:

1. $R_{YX}(t_1,t_2) = R_{XX}(t_1,t_2) * h(t_1)$.
2. $R_{YY}(t_1,t_2) = R_{YX}(t_1,t_2) * h(t_2)$.

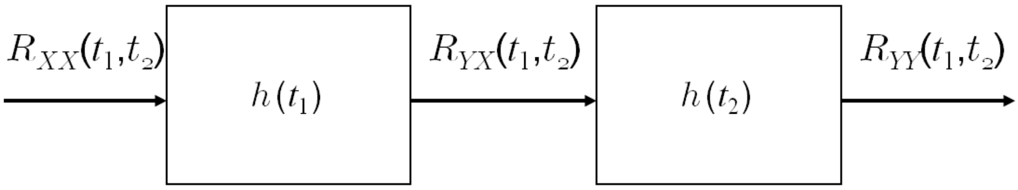


Figure 12:

Similarly, we can obtain $C_Y(t_1,t_2) \equiv C_{YY}(t_1,t_2)$ from $C_X(t_1,t_2) \equiv C_{XX}(t_1,t_2)$ by the same approach as

1. $C_{XY}(t_1,t_2) = C_{XX}(t_1,t_2) * h(t_2)$.
2. $C_{YY}(t_1,t_2) = C_{XY}(t_1,t_2) * h(t_1)$.

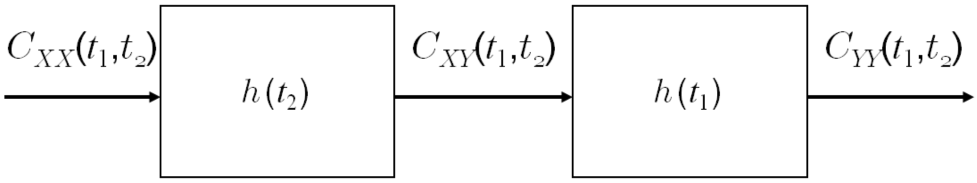


Figure 13:

Notes:

1. In the previous presentation, we use $R_{XX}(t_1,t_2) \equiv R_X(t_1,t_2)$, $R_{YY}(t_1,t_2) \equiv R_Y(t_1,t_2)$, $C_{XX}(t_1,t_2) \equiv C_X(t_1,t_2)$, $C_{YY}(t_1,t_2) \equiv C_Y(t_1,t_2)$ to help our discussion.
2. The above results hold for any real-valued rp and any LTI system.

- The preceding results can be extended to complex-valued rp's and to systems with complex impulse response $h(t)$, which states as

$$\begin{aligned} R_{YX}(t_1,t_2) &= R_{XX}(t_1,t_2) * h(t_1) \\ R_{YY}(t_1,t_2) &= R_{YX}(t_1,t_2) * h^*(t_2). \end{aligned}$$

- Self-exercise: Express the two-stage approach for LTI systems with WSS I/O.
- Ex: $X(\mu,t)$ is a complex-valued noise with zero mean and $R_{XX}(t_1,t_2) = q(t_1)\delta(t_2-t_1)$. Let $Y(\mu,t) = \mathcal{T}[X(\mu,t)]$ be an LTI system. Then,

$$\begin{aligned} E\{|Y(\mu,t)|^2\} &= q(t) * |h(t)|^2 \\ &= \int_{-\infty}^{\infty} q(t-\alpha)|h(\alpha)|^2 d\alpha. \end{aligned}$$

Proof: First, we derive

$$\begin{aligned} R_{XY}(t_1,t_2) &= R_{XX}(t_1,t_2) * h^*(t_2) \\ &= q(t_1)\delta(t_2-t_1) * h^*(t_2) \\ &= q(t_1)h^*(t_2-t_1). \end{aligned}$$

Then,

$$\begin{aligned} R_{YY}(t_1,t_2) &= R_{XY}(t_1,t_2) * h(t_1) \\ &= q(t_1)h^*(t_2-t_1) * h(t_1) \\ &= \int_{-\infty}^{\infty} q(\alpha)h^*(t_2-\alpha)h(t_1-\alpha)d\alpha. \end{aligned}$$

Now,

$$\begin{aligned} E\{|Y(\mu,t)|^2\} &= R_Y(t,t) \\ &= R_{YY}(t,t) \\ &= \int_{-\infty}^{\infty} q(\alpha)|h(t-\alpha)|^2 d\alpha \\ &= q(t) * |h(t)|^2. \qquad \text{QED} \end{aligned}$$

- Ex: Matched Filter

Let us define the real-valued rp $r(\mu,t) = As(t) + n(\mu,t)$ where

1. $s(t) = 0$ for $t \notin [0,T]$,
2. A is a positive constant,
3. $n(\mu,t)$ is a stationary additive white Gaussian noise (AWGN) with $\eta_n(t) = 0$ and $R_n(\tau) = c\delta(\tau)$ with a positive constant c.

Consider an LTI system with $r(\mu, t)$ as the input and the impulse response

$$h(t) = \begin{cases} s(T-t), & 0 \le t \le T \\ 0, & \text{otherwise} \end{cases}.$$

Such a system is called the matched filter (matched to the desired signal $s(t)$).

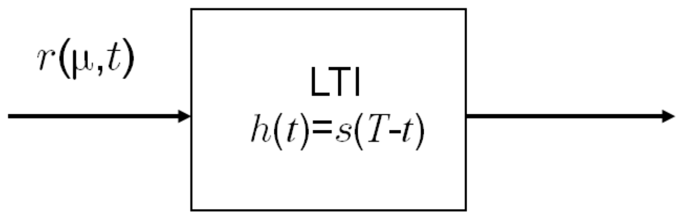


Figure 14:

The output $Y(\mu, t) = \int_{-\infty}^{\infty} r(\mu, t-\alpha) h(\alpha) d\alpha$ possesses the following statistical characteristics:

1. It is a Gaussian process since $\mathcal{T}$ is a linear operator. (This will be shown later.)
2. First, the output mean function is

$$\begin{aligned} \eta_Y(t) &= \eta_r(t) * h(t) \\ &= As(t) * h(t) \\ &= A \int_{-\infty}^{\infty} s(t-\alpha) h(\alpha) d\alpha \\ &= A \int_{-\infty}^{\infty} s(t-\alpha) s(T-\alpha) d\alpha. \end{aligned}$$

   The output mean at time $t = T$ is $\eta_Y(T) = A \int_0^T s^2(T-\alpha) d\alpha = A \int_0^T s^2(\alpha) d\alpha$.

3. Second, the output autocorrelation is

$$
\begin{aligned}
R_Y(t_1,t_2) &= E\{Y(\mu,t_1)Y(\mu,t_2)\} \\
&= E\{\int_{-\infty}^{\infty} r(\mu,t_1-\alpha)h(\alpha)d\alpha \int_{-\infty}^{\infty} r(\mu,t_2-\beta)h(\beta)d\beta\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \underbrace{E\{r(\mu,t_1-\alpha)r(\mu,t_2-\beta)\}}_{E\{(As(t_1-\alpha)+n(\mu,t_1-\alpha))(As(t_2-\beta)+n(\mu,t_2-\beta))\}} h(\alpha)h(\beta)d\alpha d\beta \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [A^2 s(t_1-\alpha)s(t_2-\beta) \\
&\quad + \underbrace{R_n(t_1-\alpha,t_2-\beta)}_{c\delta(t_1-t_2-\alpha+\beta)}]h(\alpha)h(\beta)d\alpha d\beta \\
&= A^2[\int_{-\infty}^{\infty} s(t_1-\alpha)h(\alpha)d\alpha][\int_{-\infty}^{\infty} s(t_2-\beta)h(\beta)d\beta] \\
&\quad + c\int_{-\infty}^{\infty} h(\alpha)h(\alpha-t_1+t_2)d\alpha
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
R_Y(T,T) &= A^2[\int_{-\infty}^{\infty} s^2(T-\alpha)d\alpha]^2 + c\int_{-\infty}^{\infty} h^2(\alpha)d\alpha \\
&= \eta_Y^2(T) + \frac{c}{A}\eta_Y(T).
\end{aligned}
$$

The output variance at time $t = T$ is $Var\{Y(\mu,T)\} = R_Y(T,T) - \eta_Y^2(T) = \frac{c}{A}\eta_Y(T)$.

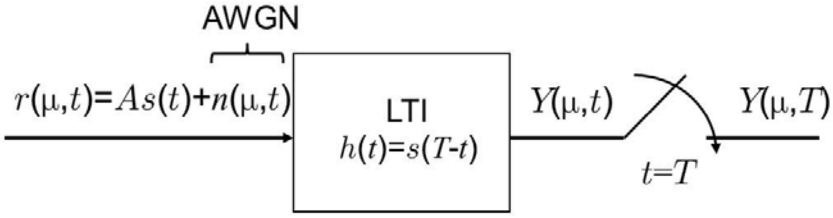


Figure 15:

Therefore, if we sample the system output at $t = T$, we can obtain the output signal-to-noise power ratio (SNR) as

–

$$
\begin{aligned}
SNR_{out} &\triangleq \frac{\eta_Y^2(T)}{Var\{Y(\mu,T)\}} \\
&= \frac{A}{c}\eta_Y(T) \\
&= \frac{A^2}{c}\int_0^T s^2(T-\alpha)d\alpha \\
&= \frac{A^2}{c}\int_0^T s^2(\alpha)d\alpha \\
&= \frac{A^2}{c}E_s
\end{aligned}
$$

where $E_s \triangleq \int_0^T s^2(\alpha)d\alpha$ is the signal energy of $s(t)$ over the time interval $[0,T]$.

Note: If we define $SNR(t) \triangleq \frac{\eta_Y^2(t)}{Var\{Y(\mu,t)\}}$, then it can be proven that $SNR(T) = \max_t SNR(t)$. That is, the matched filter yields the maximum SNR in the AWGN channel.

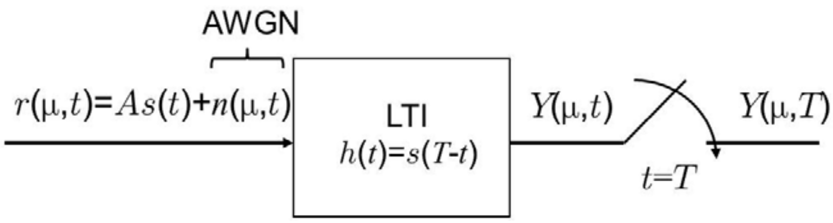


Figure 16:

2.7 Correlations and Spectra

- Defn: Two rp's $X(\mu,t)$ and $Y(\mu,t)$ are jointly WSS iff each is WSS and their cross correlations depend on time difference, i.e., $R_{XY}(t_1,t_2) = R_{XY}(t_1-t_2)$ and $R_{YX}(t_1,t_2) = R_{YX}(t_1-t_2)$.

  Note that a complex-valued rp $Z(\mu,t) = X(\mu,t) + jY(\mu,t)$ is WSS if real-valued $X(\mu,t)$ and $Y(\mu,t)$ are jointly WSS.

- Properties of Correlations for WSS and Jointly WSS RP's

  1. $R_X(-\tau) = R_X^*(\tau)$ (i.e., Hermitian symmetric)

2. $R_{XY}(-\tau) = R^*_{YX}(\tau)$.
3. If $X(\mu, t)$ is real, $R_X(-\tau) = R_X(\tau)$, an even function.
4. If $X(\mu, t)$ and $Y(\mu, t)$ are jointly WSS, then $Z(\mu, t) = aX(\mu, t) + bY(\mu, t)$ is WSS, and

$$R_Z(\tau) = |a|^2 R_X(\tau) + |b|^2 R_Y(\tau) + ab^* R_{XY}(\tau) + a^* b R_{YX}(\tau).$$

5. If $X(\mu, t)$ and $Y(\mu, t)$ are real and jointly WSS, then $Z(\mu, t) = X(\mu, t) + jY(\mu, t)$ is WSS, and

$$R_Z(\tau) = [R_X(\tau) + R_Y(\tau)] + j[R_{YX}(\tau) - R_{XY}(\tau)].$$

6. In general, $\mathrm{Re}\{Z(\mu, t)\}$ of a complex-valued WSS $Z(\mu, t)$ is NOT WSS. But, if $Z^*(\mu, t)$ and $Z(\mu, t)$ are jointly WSS, then $\mathrm{Re}\{Z(\mu, t)\}$ is WSS.
7. $R_X(\tau)$ is nonnegative definite, i.e., for any a_i's, τ_i's, and n,

$$\sum_{i=1}^{n}\sum_{k=1}^{n} a_i a_k^* R_X(\tau_i - \tau_k) \geq 0.$$

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

- We define the power spectrum $S_X(\omega)$ of a WSS rp $X(\mu, t)$ by

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

(i.e., the Fourier transform of $R_X(\tau)$). Therefore, $R_X(\tau)$ can be expressed as

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

(i.e., the inverse Fourier transform of $S_X(\omega)$).

Since $R_X(-\tau) = R^*_X(\tau)$ (Hermitian symmetric),

$$\begin{aligned} S^*_X(\omega) &= \int_{-\infty}^{\infty} R^*_X(\tau) e^{j\omega\tau} d\tau = \int_{-\infty}^{\infty} R_X(-\tau) e^{j\omega\tau} d\tau \\ &\stackrel{\tau' = -\tau}{=} \int_{-\infty}^{\infty} R_X(\tau') e^{-j\omega\tau'} d\tau' = S_X(\omega). \end{aligned}$$

That is, $S_X(\omega)$ is real-valued. Also, if $S_X(\omega)$ is real-valued, then

$$R_X^*(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_X^*(\omega)e^{-j\omega\tau}d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_X(\omega)e^{-j\omega\tau}d\omega = R_X(-\tau).$$

This shows that

$$R_X(\tau) \text{ is Hermitian symmetric} \Longleftrightarrow S_X(\omega) \text{ is real.}$$

In fact, no matter if $X(\mu, t)$ is real- or complex-valued, $S_X(\omega)$ is real.

- Furthermore, if $X(\mu, t)$ is real-valued, then $R_X(\tau) = R_X(-\tau)$ is also even and real; henceforth, $S_X(\omega)$ is also real and even. In this case, $R_X(\tau)$ and $S_X(\omega)$ form a cosine transform pair.

$$\begin{aligned} S_X(\omega) &= \int_{-\infty}^{\infty} R_X(\tau)\cos(\omega\tau)d\tau \\ R_X(\tau) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_X(\omega)\cos(\omega\tau)d\omega. \end{aligned}$$

- Next, we define the cross-power spectrum $S_{XY}(w)$ of two jointly WSS rp's $X(\mu, t)$ and $Y(\mu, t)$ by

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

whose inverse yields

$$R_{XY}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XY}(\omega)e^{j\omega\tau}d\tau.$$

Notes:

1. $S_{XY}(\omega) = S_{YX}^*(\omega)$ since $R_{XY}(\tau) = R_{YX}^*(-\tau)$.
2. $S_{XY}(\omega)$ is in general complex-valued even when $X(\mu, t)$ and $Y(\mu, t)$ are real-valued.

- Ex: Let $X(\mu, t) = a\cos(\omega(\mu)t + \phi(\mu))$ where $\omega(\mu)$ is an rv with pdf $f_\omega(\omega)$, and $\phi(\mu)$ is an rv uniformly distributed in $[-\pi, \pi)$ and independent of $\omega(\mu)$. Then,

$$\begin{aligned} R_X(t_1, t_2) &= E\{X(\mu, t_1)X(\mu, t_2)\} \\ &= \frac{a^2}{2}\{E\{\cos(\omega(\mu)(t_1+t_2) + 2\phi(\mu))\} \\ &\quad +E\{\cos(\omega(\mu)(t_1 - t_2))\}\} \\ &= \frac{a^2}{2}E\{\cos(\omega(\mu)(t_1 - t_2))\} \end{aligned}$$

since

$$\begin{aligned} & E\{\cos(\omega(\mu)(t_1+t_2)+2\phi(\mu))\} \\ = \ & E\{E\{\cos(\omega(\mu)(t_1+t_2)+2\phi(\mu))|\omega(\mu)\}\} \\ = \ & E\{\int_{-\pi}^{\pi}\frac{1}{2\pi}\cos(\omega(\mu)(t_1+t_2)+2\phi)d\phi\} \\ = \ & 0. \end{aligned}$$

(Note that $E\{g(A(\mu),B(\mu))\} = E\{E\{g(A(\mu),B(\mu))|B(\mu)\}\}$ for any well-defined function $g(\cdot,\cdot)$.) Thus,

$$\begin{aligned} R_X(t_1,t_2) &= R_X(\tau) \\ &= \frac{a^2}{2}\int_{-\infty}^{\infty}\cos(\omega\tau)f_\omega(\omega)d\omega. \end{aligned}$$

By expanding $\cos(\omega\tau) = \frac{1}{2}(e^{j\omega\tau}+e^{-j\omega\tau})$,

$$\begin{aligned} R_X(\tau) &= \frac{a^2}{4}\int_{-\infty}^{\infty}[f_\omega(\omega)+f_\omega(-\omega)]e^{j\omega\tau}d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}[\frac{a^2\pi}{2}(f_\omega(\omega)+f_\omega(-\omega))]e^{j\omega\tau}d\omega \\ \Rightarrow S_X(\omega) &= \frac{a^2\pi}{2}(f_\omega(\omega)+f_\omega(-\omega)). \end{aligned}$$

- Theorem :
  For a WSS rp $X(\mu,t)$, $S_X(\omega)\geq 0 \Longleftrightarrow R_X(\tau)$ is nonnegative definite.
  Proof: " $\Rightarrow$"

$$\begin{aligned} & \int_{-\infty}^{\infty}S_X(\omega)|\sum_i a_i e^{j\omega\tau_i}|^2\frac{d\omega}{2\pi}\geq 0 \\ \Rightarrow \ & \sum_i\sum_k a_i a_k^* \underbrace{\int_{-\infty}^{\infty}S_X(\omega)e^{j\omega(\tau_i-\tau_k)}\frac{d\omega}{2\pi}}_{\substack{\| \\ R_X(\tau_i-\tau_k)}} \geq 0 \\ \Rightarrow \ & \sum_i\sum_k a_i a_k^* R_X(\tau_i-\tau_k)\geq 0 \text{ for all } a_k\text{'s and } \tau_k\text{'s} \\ \Rightarrow \ & R_X(\tau) \text{ is nonnegative definite.} \end{aligned}$$

" $\Longleftarrow$"

$$\sum_{i=1}^{n}\sum_{k=1}^{n}a_i a_k^* R_X(\tau_i-\tau_k)\geq 0,\ \forall a_k\text{'s },\ \forall\tau_k\text{'s},\ \forall n$$

Now, we have for all a_k's, τ_k's, and n,

$$\sum_{i=1}^{n}\sum_{k=1}^{n} a_i a_k^* \int_{-\infty}^{\infty} S_X(\omega) e^{j\omega(\tau_i-\tau_k)} \frac{d\omega}{2\pi} \geq 0$$
$$\Rightarrow \int_{-\infty}^{\infty} S_X(\omega) |\sum_{k=1}^{n} a_k e^{j\omega\tau_k}|^2 \frac{d\omega}{2\pi} \geq 0.$$

By letting for some i

$$a_k = \begin{cases} \sqrt{2\pi}\exp\{-j\omega_0\tau_i\}\sqrt{\delta(\omega-\omega_0)}, & k = i \\ 0, & k \neq i \end{cases}$$

we can show that $S_X(\omega_0) \geq 0$ for all ω_0. Q.E.D.

- Thus, for a WSS rp $X(\mu, t)$, $S_X(\omega) \geq 0$ for all ω, no matter if $X(\mu, t)$ is real- or complex-valued. The average power of $X(\mu, t)$ is

$$E\{X^2(\mu, t)\} = R_X(0) = \int_{-\infty}^{\infty} S_X(f) df$$

with $\omega = 2\pi f$, where $S_X(f)$ denotes a spectrum in f (in Hz). Thus, the area of $S_X(f)$ is the average power of rp $X(\mu, t)$.

- For deterministic signals, we define the power spectrum as the limit of the integral

$$S_T(\omega) = \frac{1}{2T} |\int_{-T}^{T} X(t) e^{-j\omega t} dt|^2.$$

If we use a WSS rp $X(\mu, t)$ instead of $X(t)$ in $S_T(\omega)$, then it can be proven that

$$\lim_{T\to\infty} E\{S_T(\omega)\} = S_X(\omega).$$

Note that $S_T(\omega)$ is not used as the definition of the power spectrum of $X(\mu, t)$. Since $S_T(\omega)$ for any T is an rv with nonzero variance, $S_T(\omega)$ can not provide a stationary estimate of $S_X(\omega)$ for $T < \infty$.

2.8 Deterministic LTI System With WSS RP as Input

- Consider the LTI system with $Y(\mu, t) = X(\mu, t) * h(t)$.

  $h(t)$ is the impulse response of the LTI system.

$H(\omega) = \mathcal{F}\{h(t)\}$ is the system function of the LTI system.

We call $|H(\omega)|^2$ the energy spectrum of the LTI system, which is the Fourier transform of a function $\rho(t)$ defined by

$$\rho(t) = h(t) * h^*(-t) = \int_{-\infty}^{\infty} h(t-\alpha)h^*(-\alpha)d\alpha.$$

This function $\rho(t)$ is called the deterministic autocorrelation of $h(t)$.

- Now, apply a WSS rp $X(\mu, t)$ to the system. It then follows that the system output $Y(\mu, t)$ is also WSS. By simple mathematical treatment, we can derive the following properties: (Note: $Y(\mu, t) = \int_{-\infty}^{\infty} X(\mu, t-\alpha)h(\alpha)d\alpha$.)

  1. $$\eta_Y = E\{\int_{-\infty}^{\infty} X(\mu, t-\alpha)h(\alpha)d\alpha\} = \int_{-\infty}^{\infty} \eta_X h(\alpha)d\alpha = \eta_X H(0).$$

  2. $$\begin{aligned} R_{XY}(\tau) &= R_{XX}(\tau) * h^*(-\tau) \\ R_{YY}(\tau) &= R_{XY}(\tau) * h(\tau) \end{aligned}$$

     (a two-stage procedure to obtain $R_{YY}(\tau)$).

  3. $$\begin{aligned} S_{XY}(\omega) &= S_{XX}(\omega)H^*(\omega) \\ S_{YY}(\omega) &= S_{XY}(\omega)H(\omega). \end{aligned}$$

  4. $$\begin{aligned} R_{YY}(\tau) &= R_{XY}(\tau) * h(\tau) \\ &= R_{XX}(\tau) * \underbrace{h^*(-\tau) * h(\tau)}_{\rho(\tau)} \\ &= R_{XX}(\tau) * \rho(\tau). \end{aligned}$$

  5. $$\begin{aligned} S_{YY}(\omega) &= S_{XY}(\omega)H(\omega) \\ &= S_{XX}(\omega)H^*(\omega)H(\omega) \\ &= S_{XX}(\omega)|H(\omega)|^2. \end{aligned}$$

  The above properties can be depicted by the following picture.

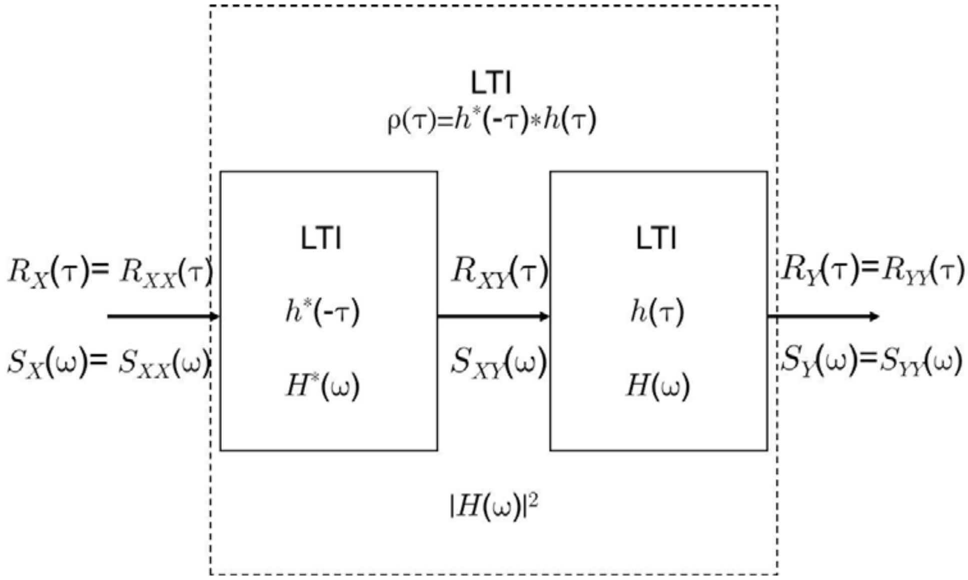


Figure 17:

6.

$$
\begin{aligned}
E\{|Y(\mu,t)|^2\} &= R_{YY}(0) \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{YY}(\omega)d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XX}(\omega)|H(\omega)|^2 d\omega \\
&= \int_{-\infty}^{\infty} R_{XX}^*(\tau)\rho(\tau)d\tau
\end{aligned}
$$

where the last equality is obtained from Parseval's relation or from

$$R_{YY}(0) = \int_{-\infty}^{\infty} R_{XX}(0-\tau)\rho(\tau)d\tau$$

and $R_{XX}^*(-\tau) = R_{XX}(\tau)$.
Note: For any two Fourier transform pairs $H(\omega) = \mathcal{F}\{h(t)\}$ and $G(\omega) = \mathcal{F}\{g(t)\}$, the Parseval's relation states that

$$\int_{-\infty}^{\infty} H(f)G^*(f)df = \int_{-\infty}^{\infty} h(t)g^*(t)dt.$$

- Ex: Consider the moving average of $X(\mu, t)$ in the interval $(t-T, t+T)$,

$$\begin{aligned} Y(\mu,t) &= \frac{1}{2T}\int_{-T+t}^{T+t} X(\mu,\alpha)d\alpha \\ &= X(\mu,t) * h(t) \\ &= \int_{-\infty}^{\infty} X(\mu,\alpha)h(t-\alpha)d\alpha. \end{aligned}$$

Clearly, the system is LTI with

$$h(t) = \begin{cases} \frac{1}{2T}, & -T < t < T \\ 0, & \text{otherwise} \end{cases}$$

$$H(\omega) = \mathcal{F}\{h(t)\} = \frac{1}{2T}\int_{-T}^{T} e^{-j\omega t}dt = \frac{\sin(\omega T)}{\omega T}$$
$$\Rightarrow \quad S_{YY}(\omega) = S_{XX}(\omega)|H(\omega)|^2 = S_{XX}(\omega)\frac{\sin^2(\omega T)}{(\omega T)^2}.$$

Since $\rho(t) = h(t) * h^*(-t) = h(t) * h(t)$,

$$\begin{aligned} R_{YY}(\tau) &= R_{XX}(\tau) * \rho(\tau) \\ &= \frac{1}{2T}\int_{-2T}^{2T} R_{XX}(\tau-\alpha)(1-\frac{|\alpha|}{2T})d\alpha. \end{aligned}$$

- Ex: If the input $X(\mu, t)$ to the LTI system is a real-valued white noise with zero mean and $R_{XX}(\tau) = q\delta(\tau)$ (with $q > 0$), then

$$\begin{aligned} R_{XY}(\tau) &= R_{XX}(\tau) * h^*(-\tau) \\ &= q\delta(\tau) * h^*(-\tau) \\ &= qh^*(-\tau) \end{aligned}$$

$$\begin{aligned} R_{YY}(\tau) &= R_{XY}(\tau) * h(\tau) \\ &= qh^*(-\tau) * h(\tau) \\ &= q\rho(\tau) \end{aligned}$$

$$S_{YY}(\omega) = q|H(\omega)|^2.$$

- Reading Assignment: Papoulis and Pillai, 4th ed., Sections 9-1, 9-2, 9-3

- Recommended Self-Exercise: Papoulis and Pillai, 4th ed.
  Chap 9: 2, 8, 13, 15, 20, 35

3 Real-Valued Gaussian Random Vectors and Real-Valued Gaussian Random Processes

- (Materials come from Larson and Shubert.)

3.1 Real-Valued Gaussian Random Vectors

- Let $\underline{Y}(\mu) = (Y_1(\mu), Y_2(\mu), ..., Y_n(\mu))^T$ be a Gaussian random vector. Its mean vector is denoted by

$$\underline{m}_Y \triangleq E\{\underline{Y}(\mu)\} = (m_1, m_2, ..., m_n)^T$$

and its covariance matrix by

$$\Lambda_Y \triangleq E\{(\underline{Y}(\mu) - \underline{m}_Y)(\underline{Y}(\mu) - \underline{m}_Y)^T\} = [\lambda_{ik}]$$

with $\lambda_{ik} = E\{(Y_i(\mu) - m_i)(Y_k(\mu) - m_k)\}$. First, it is straightforward that

$$\Lambda_Y = E\{\underline{Y}(\mu)\underline{Y}(\mu)^T\} - \underline{m}_Y \underline{m}_Y^T.$$

- Defn: The n-dimensional Gaussian density of $\underline{Y}(\mu)$ is given by

$$f_{\underline{Y}}(\underline{y}) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\Lambda_Y|}} \exp\{-\frac{1}{2}(\underline{y} - \underline{m}_Y)^T \Lambda_Y^{-1}(\underline{y} - \underline{m}_Y)\}$$

assuming nonsingular Λ_Y. Notationally, let us denote

$$f_{\underline{Y}}(\underline{y}) = G(\underline{y}; \underline{m}_Y, \Lambda_Y).$$

Notes:

1. The exponent is a real value and is also a quadratic form in the variables $y_i - m_i$, $\forall i$, i.e., it has a form of

$$\sum_{i=1}^{n}\sum_{k=1}^{n}(y_i - m_i)(y_k - m_k)a_{ik}.$$

2. The n-dimensional Gaussian density is completely specified by its mean vector and covariance matrix. In the other words, knowledge of the first and second moments is sufficient to completely describe the n-dimensional Gaussian density.

3. The density can be expressed explicitly if Λ_Y is strictly positive definite, since then Λ_Y^{-1} exists. If not, the density is concentrated on a lower dimensional hyperplane.

- Theorem 1: The characteristic function of an n-dimensional Gaussian random vector is

$$C_{\underline{Y}}(\underline{\omega};\underline{m}_Y,\Lambda_Y) \triangleq E\{\exp\{j\underline{\omega}^T\underline{Y}(\mu)\}\} = \exp\{j\underline{\omega}^T\underline{m}_Y - \frac{1}{2}\underline{\omega}^T\Lambda_Y\,\underline{\omega}\}$$

Proof:

$$\begin{aligned}
& C_{\underline{Y}}(\underline{\omega};\underline{m}_Y,\Lambda_Y) \\
= \quad & E\{\exp\{j\underline{\omega}^T\underline{Y}(\mu)\}\} \\
= \quad & \underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{n-fold} e^{j\underline{\omega}^T\underline{y}}\frac{1}{\sqrt{(2\pi)^n}\sqrt{|\Lambda_Y|}}\exp\{-\frac{1}{2}(\underline{y}-\underline{m}_Y)^T\Lambda_Y^{-1}(\underline{y}-\underline{m}_Y)\}d\underline{y}
\end{aligned}$$

$$\underline{z}=\underline{y}-\underline{m}_Y$$
$$\downarrow$$
$$= \quad e^{j\underline{\omega}^T\underline{m}_Y}\underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{n-fold}\frac{1}{\sqrt{(2\pi)^n}\sqrt{|\Lambda_Y|}}\exp\{j\underline{\omega}^T\underline{z}-\frac{1}{2}\underline{z}^T\Lambda_Y^{-1}\underline{z}\}d\underline{z}.$$

From factorization theorem,

$$\Lambda_Y = EDE^T \Rightarrow \Lambda_Y^{-1} = ED^{-1}E^T$$

where $E \equiv$ matrix of eigenvectors ($E=[\underline{e}_1,\underline{e}_2,...,\underline{e}_n]$ with $\Lambda_Y\underline{e}_i=\lambda_i\underline{e}_i$, $i=1,2,...,n$, and $EE^T=E^TE=I$) and $D\equiv$ diagonal matrix of eigenvalues $\lambda_1,\lambda_2,...,\lambda_n$, i.e.,

$$D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

Since Λ_Y is nonsingular and positive definite (i.e., $|\Lambda_Y|\neq 0$ and $\underline{g}_i^T\Lambda_Y\underline{g}_i>0$ for all $\underline{g}_i$'s), $\lambda_i>0$, $\forall i=1,2,...,n$.
Now, letting $\underline{x}=E^T\underline{z}\Rightarrow\underline{z}=E\underline{x}$, we have

$$C_Y(\underline{\omega};\underline{m}_Y,\Lambda_Y) = e^{j\underline{\omega}^T\underline{m}_Y}\underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{n-fold}\frac{\exp\{j\underline{\omega}^TE\underline{x}-\frac{1}{2}(E\underline{x})^T\Lambda_Y^{-1}(E\underline{x})\}}{(\sqrt{2\pi})^n\sqrt{|\Lambda_Y|}}d\underline{x}.$$

Next, using the fact that $|\Lambda_Y| = \prod_{i=1}^n \lambda_i$ and simplifying the exponent, we obtain

$$C_Y(\underline{\omega}; \underline{m}_Y, \Lambda_Y) = e^{j\underline{\omega}^T \underline{m}_Y} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n-fold} \frac{\exp\{j\overbrace{\underline{\omega}^T E}^{g^T}\underline{x} - \frac{1}{2}\underline{x}^T D^{-1}\underline{x}\}}{(\sqrt{2\pi})^n \sqrt{\prod_{i=1}^n \lambda_i}} d\underline{x}.$$

Let $\underline{g} = E^T \underline{\omega}$. Then, the above becomes

$$C_Y(\underline{\omega}; \underline{m}_Y, \Lambda_Y) = e^{j\underline{\omega}^T \underline{m}_Y} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n-fold} \frac{\exp\{j\sum_{i=1}^n g_i x_i - \frac{1}{2}\sum_{i=1}^n \frac{x_i^2}{\lambda_i}\}}{(\sqrt{2\pi})^n \prod_{i=1}^n \sqrt{\lambda_i}} d\underline{x}$$

where $\underline{x} = [x_1, x_2, ..., x_n]^T$ and $\underline{g} = [g_1, g_2, ..., g_n]^T$. Here, we observe that the multiple integral can be expressed as a product of independent terms. Therefore,

$$C_Y(\underline{\omega}; \underline{m}_Y, \Lambda_Y) = e^{j\underline{\omega}^T \underline{m}_Y} \prod_{i=1}^n \underbrace{\int_{-\infty}^{\infty} \frac{\exp\{jg_i x_i - \frac{1}{2}\frac{x_i^2}{\lambda_i}\}}{\sqrt{2\pi\lambda_i}} dx_i}_{=\square}.$$

Since

$$\begin{aligned} \square &= e^{-\frac{1}{2}g_i^2\lambda_i} \int_{-\infty}^{\infty} \underbrace{\frac{\exp\{-\frac{1}{2\lambda_i}(x_i - jg_i\lambda_i)^2\}}{\sqrt{2\pi\lambda_i}}}_{\text{Gaussian density}} dx_i \\ &= e^{-\frac{1}{2}g_i^2\lambda_i} \end{aligned}$$

we further have

$$\begin{aligned} C_Y(\underline{\omega}; \underline{m}_Y, \Lambda_Y) &= e^{j\underline{\omega}^T \underline{m}_Y} \exp\{-\frac{1}{2}\overbrace{\sum_{i=1}^n g_i^2 \lambda_i}^{\underline{g}^T D\, \underline{g}}\} \\ &= \exp\{j\underline{\omega}^T \underline{m}_Y - \frac{1}{2}\underline{g}^T D \underline{g}\} \\ &= \exp\{j\underline{\omega}^T \underline{m}_Y - \frac{1}{2}(E^T\underline{\omega})^T D\ (E^T\underline{\omega})\} \\ &= \exp\{j\underline{\omega}^T \underline{m}_Y - \frac{1}{2}\underline{\omega}^T \Lambda_Y \underline{\omega}\}. \qquad \textit{Q.E.D.} \end{aligned}$$

- Let $\underline{Z}(\mu)$ represent an n-dimensional Gaussian random vector. A linear transformation of $\underline{Z}(\mu)$ into $\underline{Y}(\mu)$ can be expressed by a matrix equation
$$\underline{Y}(\mu) = A\underline{Z}(\mu)$$
where $\underline{Z}(\mu)$ is $n \times 1$, $\underline{Y}(\mu)$ is $m \times 1$, and A is $m \times n$.

- Theorem 2: Linear transformation on Gaussian random vectors produces Gaussian random vectors.

  Proof: We will show that the characteristic function $C_Y(\underline{\omega}; \underline{m}_Y, \Lambda_Y)$ is that of a Gaussian random vector. Now,
$$\begin{aligned} C_Y(\underline{\omega}; \underline{m}_Y, \Lambda_Y) &= E\{\exp\{j\underline{\omega}^T \underline{Y}(\mu)\}\} \\ &= E\{\exp\{j\underline{\omega}^T A\underline{Z}(\mu)\}\} \\ &= E\{\exp\{j(A^T\underline{\omega})^T \underline{Z}(\mu)\}\} \\ &= C_Z(A^T\underline{\omega}; \underline{m}_Z, \Lambda_Z). \end{aligned}$$
  From Theorem 1,
$$C_Z(\underline{v}; \underline{m}_Z, \Lambda_Z) = \exp\{j\underline{v}^T \underline{m}_Z - \frac{1}{2}\underline{v}^T \Lambda_Z \underline{v}\}.$$
  Substituting $A^T\underline{\omega}$ for $\underline{v}$, we have
$$C_Z(A^T\underline{\omega}; \underline{m}_Z, \Lambda_Z) = \exp\{j\underline{\omega}^T (A\underline{m}_Z) - \frac{1}{2}\underline{\omega}^T (A\Lambda_Z A^T)\underline{\omega}\}$$
  which is the form for the characteristic function of a Gaussian random vector. Therefore, $\underline{Y}(\mu)$ is a Gaussian random vector with mean vector $\underline{m}_Y = A\underline{m}_Z$ and covariance matrix $\Lambda_Y = A\Lambda_Z A^T$. *Q.E.D.*

- Theorem 3: If two jointly Gaussian random vectors $\underline{X}(\mu)$ and $\underline{Y}(\mu)$ are uncorrelated, they are statistically independent.
  Proof: Let $\underline{Z}(\mu) = \begin{bmatrix} \underline{X}(\mu) \\ \underline{Y}(\mu) \end{bmatrix}$ where

  1. $\underline{X}(\mu)$ is an $n \times 1$ Gaussian random vector with mean $\underline{m}_X$ and covariance Λ_X;
  2. $\underline{Y}(\mu)$ is an $m \times 1$ Gaussian random vector with mean $\underline{m}_Y$ and covariance Λ_Y.

Denote the cross-covariance matrix of $\underline{X}(\mu)$ and $\underline{Y}(\mu)$ as

$$C_{XY} = E\{(\underline{X}(\mu) - \underline{m}_X)(\underline{Y}(\mu) - \underline{m}_Y)^T\}.$$

Then, we have

$$C_{XY}^T = C_{YX}$$

$$\begin{aligned}
\Lambda_Z &= E\left\{ \begin{bmatrix} \underline{X}(\mu) - \underline{m}_X \\ \underline{Y}(\mu) - \underline{m}_Y \end{bmatrix} \begin{bmatrix} \underline{X}(\mu) - \underline{m}_X \\ \underline{Y}(\mu) - \underline{m}_Y \end{bmatrix}^T \right\} \\
&= \begin{bmatrix} \Lambda_X & C_{XY} \\ C_{YX} & \Lambda_Y \end{bmatrix}_{(m+n)\times(m+n)}.
\end{aligned}$$

Also,

$$\begin{aligned}
f_{\underline{Z}}(\underline{x}, \underline{y}) &= f_{\underline{Z}}(\underline{z}) = G\left(\begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix}; \begin{bmatrix} \underline{m}_X \\ \underline{m}_Y \end{bmatrix}, \begin{bmatrix} \Lambda_X & C_{XY} \\ C_{YX} & \Lambda_Y \end{bmatrix} \right) \\
&= \frac{1}{(2\pi)^{\frac{m+n}{2}} \sqrt{\begin{vmatrix} \Lambda_X & C_{XY} \\ C_{YX} & \Lambda_Y \end{vmatrix}}} \\
&\quad \cdot \exp\left\{ -\frac{1}{2} \begin{bmatrix} \underline{x} - \underline{m}_X \\ \underline{y} - \underline{m}_Y \end{bmatrix}^T \begin{bmatrix} \Lambda_X & C_{XY} \\ C_{YX} & \Lambda_Y \end{bmatrix}^{-1} \begin{bmatrix} \underline{x} - \underline{m}_X \\ \underline{y} - \underline{m}_Y \end{bmatrix} \right\}.
\end{aligned}$$

If $\underline{X}(\mu)$ and $\underline{Y}(\mu)$ are uncorrelated, $C_{XY} = 0$,

$$\begin{aligned}
\begin{vmatrix} \Lambda_X & 0 \\ 0 & \Lambda_Y \end{vmatrix} &= |\Lambda_X|\,|\Lambda_Y| \\
\begin{bmatrix} \Lambda_X & 0 \\ 0 & \Lambda_Y \end{bmatrix}^{-1} &= \begin{bmatrix} \Lambda_X^{-1} & 0 \\ 0 & \Lambda_Y^{-1} \end{bmatrix}.
\end{aligned}$$

In the case,

$$\begin{aligned}
f_{\underline{X},\underline{Y}}(\underline{x}, \underline{y}) &= \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{|\Lambda_X|}} \exp\{-\frac{1}{2}(\underline{x} - \underline{m}_X)^T \Lambda_X^{-1} (\underline{x} - \underline{m}_X)\} \\
&\quad \times \frac{1}{(2\pi)^{\frac{m}{2}} \sqrt{|\Lambda_Y|}} \exp\{-\frac{1}{2}(\underline{y} - \underline{m}_Y)^T \Lambda_Y^{-1} (\underline{y} - \underline{m}_Y)\} \\
&= f_{\underline{X}}(\underline{x}) f_{\underline{Y}}(\underline{y})
\end{aligned}$$

which indicates that $\underline{X}(\mu)$ and $\underline{Y}(\mu)$ are independent. *Q.E.D.*

- Theorem 4:

  If $\underline{X}(\mu)$ and $\underline{Y}(\mu)$ are two jointly Gaussian random vectors, n and m dimensional, respectively, with mean vectors $\underline{m}_X$ and $\underline{m}_Y$, covariance matrices Λ_X and Λ_Y, and cross-covariance matrix C_{XY}, then the conditional density of the random vector $\underline{X}(\mu)$, given $\underline{Y}(\mu)$, is also Gaussian with conditional mean

$$E\{\underline{X}(\mu)|\underline{Y}(\mu)\} = \underline{m}_X + C_{XY}\Lambda_Y^{-1}(\underline{Y}(\mu) - \underline{m}_Y)$$

  and conditional covariance matrix $\Lambda_{X|Y} \triangleq E\{(\underline{X}(\mu) - \underline{m}_X)(\underline{X}(\mu) - \underline{m}_X)^T|\underline{Y}(\mu)\} = \Lambda_X - C_{XY}\Lambda_Y^{-1}C_{YX}$.

  Proof:

  (1) First, $f_{\underline{X}|\underline{Y}}(\underline{x}|\underline{y}) = \frac{f_{\underline{X},\underline{Y}}(\underline{x},\underline{y})}{f_{\underline{Y}}(\underline{y})}$ where

$$\begin{aligned} f_{\underline{X},\underline{Y}}(\underline{x},\underline{y}) &= (2\pi)^{-\frac{1}{2}(m+n)}\left(\begin{vmatrix} \Lambda_X & C_{XY} \\ C_{YX} & \Lambda_Y \end{vmatrix}\right)^{-\frac{1}{2}} \\ &\quad \cdot \exp\left\{-\frac{1}{2}\begin{bmatrix} \underline{x}-\underline{m}_X \\ \underline{y}-\underline{m}_Y \end{bmatrix}^T \begin{bmatrix} \Lambda_X & C_{XY} \\ C_{YX} & \Lambda_Y \end{bmatrix}^{-1} \begin{bmatrix} \underline{x}-\underline{m}_X \\ \underline{y}-\underline{m}_Y \end{bmatrix}\right\} \\ f_{\underline{Y}}(\underline{y}) &= (2\pi)^{-\frac{m}{2}}\left|\Lambda_Y\right|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\underline{y}-\underline{m}_Y)^T\Lambda_Y^{-1}(\underline{y}-\underline{m}_Y)\right\}. \end{aligned}$$

$$\begin{aligned} \Rightarrow f_{\underline{X}|\underline{Y}}(\underline{x}|\underline{y}) &= (2\pi)^{-\frac{n}{2}}\left(\left|\Lambda_Y^{-1}\right|\begin{vmatrix} \Lambda_X & C_{XY} \\ C_{YX} & \Lambda_Y \end{vmatrix}\right)^{-\frac{1}{2}} \\ &\quad \exp\left\{-\frac{1}{2}\begin{bmatrix} \underline{x}-\underline{m}_X \\ \underline{y}-\underline{m}_Y \end{bmatrix}^T \begin{bmatrix} \Lambda_X & C_{XY} \\ C_{YX} & \Lambda_Y \end{bmatrix}^{-1} \begin{bmatrix} \underline{x}-\underline{m}_X \\ \underline{y}-\underline{m}_Y \end{bmatrix}\right. \\ &\quad \left.+\frac{1}{2}(\underline{y}-\underline{m}_Y)^T\Lambda_Y^{-1}(\underline{y}-\underline{m}_Y)\right\}. \end{aligned}$$

  (2) To proceed, we need

$$\begin{aligned} \begin{vmatrix} \Lambda_X & C_{XY} \\ C_{YX} & \Lambda_Y \end{vmatrix} &\underset{\substack{\uparrow \\ \text{elementary} \\ \text{row operation}}}{=} \begin{vmatrix} \Lambda_X - C_{XY}\Lambda_Y^{-1}C_{YX} & \overbrace{C_{XY} - C_{XY}\Lambda_Y^{-1}\Lambda_Y}^{=0} \\ C_{YX} & \Lambda_Y \end{vmatrix} \\ &= \left|\Lambda_X - C_{XY}\Lambda_Y^{-1}C_{YX}\right|\left|\Lambda_Y\right| \end{aligned}$$

and

$$\begin{bmatrix} \Lambda_X & C_{XY} \\ C_{YX} & \Lambda_Y \end{bmatrix}^{-1} = \begin{bmatrix} \alpha & \beta \\ \gamma & \eta \end{bmatrix}$$

where $\alpha = \left[\Lambda_X - C_{XY}\Lambda_Y^{-1}C_{YX}\right]^{-1}$, $\beta = -\alpha C_{XY}\Lambda_Y^{-1}$, $\gamma = -\Lambda_Y^{-1}C_{YX}\alpha$ and $\eta = \Lambda_Y^{-1} - \Lambda_Y^{-1}C_{YX}\beta$.

Self-Exercise: Show this by multiplying $\begin{bmatrix} \alpha & \beta \\ \gamma & \eta \end{bmatrix}$ by $\begin{bmatrix} \Lambda_X & C_{XY} \\ C_{YX} & \Lambda_Y \end{bmatrix}$, and $\begin{bmatrix} \Lambda_X & C_{XY} \\ C_{YX} & \Lambda_Y \end{bmatrix}$ by $\begin{bmatrix} \alpha & \beta \\ \gamma & \eta \end{bmatrix}$.

(3) Substituting (2) into (1), we have

$$\begin{aligned} f_{\underline{X}|\underline{Y}}(\underline{x}|\underline{y}) \;=\; & (2\pi)^{-n/2}\left|\Lambda_X - C_{XY}\Lambda_Y^{-1}C_{YX}\right|^{-1/2} \\ & \cdot\exp\left\{\frac{-1}{2}(\underline{x}-\underline{m}_X)^T\left[\Lambda_X - C_{XY}\Lambda_Y^{-1}C_{YX}\right]^{-1}(\underline{x}-\underline{m}_X)\right. \\ & -(\underline{x}-\underline{m}_X)^T\left[\Lambda_X - C_{XY}\Lambda_Y^{-1}C_{YX}\right]^{-1}C_{XY}\Lambda_Y^{-1}(\underline{y}-\underline{m}_Y) \\ & -(\underline{y}-\underline{m}_Y)^T\Lambda_Y^{-1}C_{YX}\left[\Lambda_X - C_{XY}\Lambda_Y^{-1}C_{YX}\right]^{-1}(\underline{x}-\underline{m}_X) \\ & +(\underline{y}-\underline{m}_Y)^T(\Lambda_Y^{-1} + \Lambda_Y^{-1}C_{YX}\left[\Lambda_X - C_{XY}\Lambda_Y^{-1}C_{YX}\right]^{-1} \\ & \left.\cdot C_{XY}\Lambda_Y^{-1})(\underline{y}-\underline{m}_Y) - (\underline{y}-\underline{m}_Y)^T\Lambda_Y^{-1}(\underline{y}-\underline{m}_Y)\right\}. \qquad \circledast \end{aligned}$$

(4) Self-Exercise: Show that $\circledast$ can be simplified to

$$\begin{aligned} f_{\underline{X}|\underline{Y}}(\underline{x}|\underline{y}) = \; & (2\pi)^{-\frac{n}{2}}\left|\Lambda_X - C_{XY}\Lambda_Y^{-1}C_{YX}\right|^{-\frac{1}{2}} \\ & \cdot\exp\{-\tfrac{1}{2}(\underline{x} - [\underline{m}_X + C_{XY}\Lambda_Y^{-1}(\underline{y}-\underline{m}_Y)])^T \\ & \cdot\left[\Lambda_X - C_{XY}\Lambda_Y^{-1}C_{YX}\right]^{-1} \\ & \cdot(\underline{x} - [\underline{m}_X + C_{XY}\Lambda_Y^{-1}(\underline{y}-\underline{m}_Y)])\}. \end{aligned}$$

(5) This means that, given $\underline{Y}(\mu)$, $\underline{X}(\mu)$ is conditionally Gaussian with mean $\underline{m}_X + C_{XY}\Lambda_Y^{-1}(\underline{y}-\underline{m}_Y)$ and covariance $\Lambda_X - C_{XY}\Lambda_Y^{-1}C_{YX}$. Q.E.D.

3.2 Real-Valued Gaussian Random Processes

- Defn: A real-valued rp $X(\mu,t)$ is said to be a Gaussian rp iff, for every set of time instants $\{t_i\}$, the real-valued random variables $\{X(\mu,t_i)\}$ have a Gaussian joint probability density. The density of the real-valued Gaussian random vector $\underline{X}(\mu) = [X(\mu,t_1), X(\mu,t_2), ..., X(\mu,t_n)]^T$

may be determined from its mean vector $\underline{m}_X$ and its covariance matrix Λ_X, i.e.,

$$\begin{aligned}
\underline{m}_X &= [E\{X(\mu,t_1)\}, E\{X(\mu,t_2)\}, ..., E\{X(\mu,t_n)\}]^T \\
&= [\eta_X(t_1), \eta_X(t_2), ..., \eta_X(t_n)]^T \\
\Lambda_X &= [\lambda_{ik}]
\end{aligned}$$

with $\lambda_{ik} \triangleq E\{[X(\mu,t_i)-E\{X(\mu,t_i)\}][X(\mu,t_k)-E\{X(\mu,t_k)\}]\} = R_X(t_i,t_k)$ $-\eta_X(t_i)\eta_X(t_k)$.

Note that a Gaussian rp is completely statistically characterized by its mean function and autocorrelation function.

- If this rp is WSS, then $\underline{m}_X = [\eta_X, \eta_X, ..., \eta_X]^T$ with $\eta_X = E\{X(\mu,t)\}$ and $\lambda_{ik} = R_X(t_i - t_k) - \eta_X^2$. In the case, the n-dimensional pdf is only a function of time differences $\{t_i - t_k\}$. Thus, a WSS Gaussian rp is SSS.

3.3 Linear Transformation of a Real-Valued Gaussian Random Process

- Consider the linear system

$$Y(u,t) = \int_{a(t)}^{b(t)} X(u,\tau)h(t,\tau)d\tau$$

where the system response function is not necessarily time invariant, and $X(u,t)$ is a Gaussian random process but not assumed to be stationary.

Our goal here is to show that $Y(u,t)$ is also a Gaussian random process. Without loss of generality, let us assume zero-mean $X(u,t)$ first and proceed as follows:

(1) We begin the analysis by showing that $Y(u,t)$ is a Gaussian random variable at any specific time t. Let us look at the characteristic function

of $Y(u,t)$ for a fixed t.

$$\begin{aligned}
C_Y(\omega;t) &= E\{e^{j\omega Y(u,t)}\} \\
&= \sum_{m=0}^{\infty}\frac{(j\omega)^m}{m!}E\{Y^m(\mu,t)\} \\
&= \sum_{m=0}^{\infty}\frac{(j\omega)^m}{m!}E\left\{\left[\int_{a(t)}^{b(t)}X(u,\tau)h(t,\tau)d\tau\right]^m\right\} \\
&= \sum_{m=0}^{\infty}\frac{(j\omega)^m}{m!}\underbrace{\int_{a(t)}^{b(t)}\cdots\int_{a(t)}^{b(t)}}_{m-fold}\left(\prod_{i=1}^{m}h(t,\tau_i)\right) \\
&\quad \cdot E\left\{\prod_{l=1}^{m}X(\mu,\tau_l)\right\}d\tau_1 d\tau_2 ... d\tau_m. \quad (\#)
\end{aligned}$$

- Lemma: For zero-mean jointly Gaussian random variables $X_1(\mu), X_2(\mu), ..., X_n(\mu)$,

$$E\{\prod_{i=1}^{n}X_i(\mu)\} = \begin{cases} 0, & n \text{ is odd} \\ \displaystyle\sum_{\substack{\text{all combinations of } \frac{n}{2} \text{ pairs} \\ i_1 \neq i_2 \neq ... \neq i_n \text{ and } i_k \in \{1,2,...,n\}}}^{\frac{n!}{(\frac{n}{2})!2^{n/2}}} \left[\prod_{k=1}^{\frac{n}{2}} E\{X_{i_{2k}}(\mu)X_{i_{2k+1}}(\mu)\}\right], & n \text{ is even} \end{cases}.$$

Ex: For jointly Gaussian zero-mean rv's $X_1(\mu), X_2(\mu), X_3(\mu), X_4(\mu)$,

$$\begin{aligned}
E\{\prod_{i=1}^{4}X_i(\mu)\} &= E\{X_1(\mu)X_2(\mu)\}E\{X_3(\mu)X_4(\mu)\} \\
&\quad +E\{X_1(\mu)X_3(\mu)\}E\{X_2(\mu)X_4(\mu)\} \\
&\quad +E\{X_1(\mu)X_4(\mu)\}E\{X_2(\mu)X_3(\mu)\} \\
E\{\prod_{i=1}^{3}X_i(\mu)\} &= 0.
\end{aligned}$$

Proof: For n-dimensional Gaussian random vector $\underline{X}(\mu) = [X_1(\mu), X_2(\mu), ..., X_n(\mu)]^T$ with zero mean,

$$C_X(\underline{\omega};\underline{0},\Lambda_X) \triangleq E\{\exp\{j\underline{\omega}^T\underline{X}(\mu)\}\} = \exp\{\frac{-1}{2}\underline{\omega}^T\Lambda_X\underline{\omega}\}$$

with $\underline{\omega} = [\omega_1, \omega_2, ..., \omega_n]^T$. Now,

$$\begin{aligned} E\{\exp\{j\underline{\omega}^T \underline{X}(\mu)\}\} &= \sum_{l=0}^{\infty} \frac{1}{l!} E\{(j\underline{\omega}^T \underline{X}(\mu))^l\} \\ \exp\{\frac{-1}{2}\underline{\omega}^T \Lambda_X \underline{\omega}\} &= \sum_{l=0}^{\infty} \frac{1}{l!} (\frac{-1}{2}\underline{\omega}^T \Lambda_X \underline{\omega})^l. \end{aligned}$$

Next, letting $\Lambda_X = [C_{ik}] = [E\{X_i(\mu)X_k(\mu)\}]$, we have

$$\begin{aligned} E\{\exp\{j\underline{\omega}^T \underline{X}(\mu)\}\} &= ... + \frac{j^n}{n!} E\{(\omega_1 X_1(\mu) + ... + \omega_n X_n(\mu))^n\} + ... \\ \exp\{\frac{-1}{2}\underline{\omega}^T \Lambda_X \underline{\omega}\} &= ... + \frac{1}{(\frac{n}{2})!} (\frac{-1}{2})^{\frac{n}{2}} (\sum_{i=1}^{n} \sum_{k=1}^{n} \omega_i \omega_k C_{ik})^{\frac{n}{2}} + ... \end{aligned}$$

For n odd, since there is no terms $\prod_{i=1}^{n} \omega_i$ in $\exp\{\frac{-1}{2}\underline{\omega}^T \Lambda_X \underline{\omega}\}$, $E\{\prod_{i=1}^{n} X_i(\mu)\}$ in $E\{\exp\{j\underline{\omega}^T \underline{X}(\mu)\}\}$ must be zero.
For n even, let us find the coefficient of $\prod_{i=1}^{n} \omega_i$ as follows:

$$E\{\exp\{j\underline{\omega}^T \underline{X}(\mu)\}\} = ... + \frac{j^n}{n!} E\{\prod_{i=1}^{n} X_i(\mu)\}(n!) \prod_{i=1}^{n} \omega_i + ... \qquad (A)$$

$$\exp\{\frac{-1}{2}\underline{\omega}^T \Lambda_X \underline{\omega}\} = ...$$

$$\underbrace{+\frac{1}{(\frac{n}{2})!} (\frac{-1}{2})^{\frac{n}{2}} (\frac{n}{2})! 2^{\frac{n}{2}} (\sum_{\substack{\textit{all combinations of } \frac{n}{2} \textit{ pairs} \\ i_1 \neq i_2 \neq ... \neq i_n \textit{ and } i_k \in \{1,2,...,n\}}}^{\frac{n!}{(\frac{n}{2})! 2^{n/2}}} (\prod_{k=1}^{\frac{n}{2}} C_{i_{2k} i_{2k+1}})}_{j^n \sum_{\substack{\textit{all combinations of } \frac{n}{2} \textit{ pairs} \\ i_1 \neq i_2 \neq ... \neq i_n \textit{ and } i_k \in \{1,2,...,n\}}}^{\frac{n!}{(\frac{n}{2})! 2^{n/2}}} (\prod_{k=1}^{\frac{n}{2}} C_{i_{2k} i_{2k+1}})} \prod_{i=1}^{n} \omega_i + ... (B)$$

Equating the coefficients of $\prod_{i=1}^{n} \omega_i$ in (A) and (B),

$$E\{\prod_{i=1}^{n} X_i(\mu)\} = \sum_{\substack{\textit{all combinations of } \frac{n}{2} \textit{ pairs} \\ i_1 \neq i_2 \neq ... \neq i_n \textit{ and } i_k \in \{1,2,...,n\}}}^{\frac{n!}{(\frac{n}{2})! 2^{n/2}}} \left[\prod_{k=1}^{\frac{n}{2}} C_{i_{2k} i_{2k+1}} \right]. \qquad \text{Q.E.D.}$$

- With this Lemma, the summation in (#) can be taken outside the integral, and we have

$$C_Y(\omega;t) \underset{\substack{\uparrow \\ m=2n}}{=} \sum_{n=0}^{\infty} \frac{(-1)^n \omega^{2n}}{(2n)!}$$

$$\cdot \left\{ \sum_{\substack{\text{all combinations of } n \text{ pairs} \\ i_1 \neq i_2 \neq \ldots \neq i_n \text{ and } i_k \in \{1,2,\ldots,2n\}}}^{\frac{(2n)!}{n!2^n}} \left[\int_{a(t)}^{b(t)} \int_{a(t)}^{b(t)} h(t,\tau_{i_1}) h(t,\tau_{i_2}) R_X(\tau_{i_1},\tau_{i_2}) d\tau_{i_1} d\tau_{i_2} \right]^n \right\}.$$

Note that for any set of pairing, the above double integral gives the same result. Thus,

$$\begin{aligned}
C_Y(\omega;t) &= \sum_{n=0}^{\infty} \frac{(-1)^n (\omega^2)^n}{(2n)!} \frac{(2n)!}{2^n n!} \\
&\quad \cdot \left[\int_{a(t)}^{b(t)} \int_{a(t)}^{b(t)} h(t,\tau_1) h(t,\tau_2) R_X(\tau_1,\tau_2) d\tau_1 d\tau_2 \right]^n \\
&= \exp\{ \frac{-\omega^2}{2} \int_{a(t)}^{b(t)} \int_{a(t)}^{b(t)} h(t,\tau_1) h(t,\tau_2) R_X(\tau_1,\tau_2) d\tau_1 d\tau_2 \} \\
&= \exp\{ \frac{-\omega^2}{2} R_Y(t,t) \}
\end{aligned}$$

where $R_Y(t,t) = \int_{a(t)}^{b(t)} \int_{a(t)}^{b(t)} h(t,\tau_1) h(t,\tau_2) R_X(\tau_1,\tau_2) d\tau_1 d\tau_2$ is the auto-correlation of $Y(\mu,t)$ at time t.

- The above characteristic function is that of a Gaussian random variable and represents the density

$$f_Y(y;t) = \frac{1}{\sqrt{2\pi R_Y(t,t)}} \exp\{ -\frac{y^2}{2R_Y(t,t)} \}.$$

This proves that a linear transform of a zero-mean Gaussian random process, at any time t, yields a Gaussian random variable.

(2) Define $W(\mu,t) = X(\mu,t) + \eta_W(t)$ and

$$\begin{aligned}
Z(\mu,t) &= \int_{a(t)}^{b(t)} h(t,\tau) W(\mu,\tau) d\tau \\
&= Y(\mu,t) + \eta_Z(t)
\end{aligned}$$

where $\eta_Z(t) = \int_{a(t)}^{b(t)} h(t,\tau)\eta_W(\tau)d\tau$. Thus, the first-order density of $Z(\mu, t)$ is

$$f_Z(z;t) = \frac{1}{\sqrt{2\pi R_Y(t,t)}} \exp\{-\frac{(z-\eta_Z(t))^2}{2R_Y(t,t)}\}.$$

From (1) and (2), we show that a linear transform of a Gaussian process yields a Gaussian random variable, when evaluated at any specific time.

(3) To demonstrate that $Y(\mu, t)$, and hence $Z(\mu, t)$, are actually Gaussian random processes, we have to show furthermore that an arbitrary n-dimensional vector of samples of $Y(\mu, t)$ has an n-dimensional Gaussian joint density. The proof of this fact can be derived along the same line as the preceding proof (1)+(2).

- In summary,
  a linear transformation of a Gaussian random process is a Gaussian process.

4 Karhunen-Loéve Representation

- (A Theory of Random Signal Representation)

  (Materials come from Papoulis and Pillai, Chap. 11)

- In this part, we introduce a fundamental theory of random signal representation, known as Karhunen-Loéve representation.

- For a strictly bandlimited random process, i.e., one whose spectral density is nonzero only within some frequency interval $[-F, F]$. Shannon sampling technique states that a random process $X(\mu, t)$ can be represented by

$$X(\mu, t) = \underset{M\to\infty}{l.i.m.} \sum_{k=-M}^{M} X(\mu, t_k) \frac{\sin(2\pi F(t - t_k))}{2\pi F(t - t_k)}$$

  where $t_k = k\Delta t$ with $2F\Delta t = 1$, and the convergence is in a mean square sense, i.e.,

$$\lim_{M\to\infty} E\left\{ \left| X(\mu, t) - \sum_{k=-M}^{M} X(\mu, t_k) \frac{\sin(2\pi F(t - t_k))}{2\pi F(t - t_k)} \right|^2 \right\} = 0.$$

  There exist two inherent disadvantages with this sampling technique:

  1. Strictly bandlimited processes are necessarily not time-limited. But, in practice, we deal often with processes limited in some time interval, which are not bandlimited.
  2. In general, $\{X(\mu, t_k)\}$ are not mutually independent, nor uncorrelated, random variables. (An exception is white Gaussian noise.) This will cause inconvenience when the number of samples is increased and they get closer together.

  Both disadvantages are alleviated if the Karhunen-Loéve expansion is used instead to represent a random process. In the following, let us consider a random process $X(\mu, t)$ defined in $t \in [0, T]$.

- Recall that every autocovariance or autocorrelation function of a complex-valued random process is both Hermitian symmetric, i.e., $C_X(t_1, t_2) = C_X^*(t_2, t_1)$, $\forall t_1, t_2$, and nonnegative definite, i.e.,

$$\int_0^T \int_0^T g^*(t_1) C_X(t_1, t_2) g(t_2) dt_1 dt_2 \geq 0$$

for all $g(t)$'s in $\mathcal{L}_2[0,T]$ (Lebesque-measurable or square-integrable, i.e., $\int_0^T |g(t)|^2 dt < \infty$).

- If $R_X(t,s)$ is an autocorrelation function of complex-valued $X(\mu,t)$, then the total energy in the process in the time interval $[0,T]$ is

$$\begin{aligned}
\int_0^T E\{|X(\mu,t)|^2\}dt &= \int_0^T R_X(t,t)dt \\
&= \sum_{k=1}^{\infty} \int_0^T \lambda_k |\phi_k(t)|^2 dt \qquad \circledast \\
&= \sum_{k=1}^{\infty} \lambda_k \qquad \oplus
\end{aligned}$$

where $\{\phi_k(t)\}$ are the eigenfunctions and $\{\lambda_k\}$ the corresponding eigenvalue of $R_X(t,s)$, i.e.,

$$\begin{aligned}
\int_0^T R_X(t,s)\phi_k(s)ds &= \lambda_k \phi_k(t), \qquad 0 \le t \le T \\
\int_0^T \phi_i(t)\phi_k^*(t)dt &= \delta_{ik}.
\end{aligned}$$

Notes:

1. $\circledast$ results from the Mercer's theorem:
   Consider the linear transformation generated by the continuous Hermitian symmetric Kernel $A(t,s)$. If $A(t,s)$ is nonnegative definite, or equivalently, if all the eigenvalues of $A(t,s)$ are nonnegative, then
   $$A(t,s) = \sum_{i=1}^{\infty} \lambda_i \phi_i(t)\phi_i^*(s)$$
   where λ_i's are the eigenvalues and $\phi_i(t)$'s the (normalized) eigenfunctions of A, and where the convergence is uniform.

   Reference: E. Wong and B. Hajek, "Stochastic Processes in Engineering Systems," Springer-Verlag, 1985.

2. It is shown by $\oplus$ that the total energy in $[0,T]$ is equal to the sum of the eigenvalues of its autocorrelation function. Each eigenvalue may be thought of as the energy associated with the corresponding component of $X(\mu,t)$ in the $\phi_k(t)$ direction during the interval

$[0, T]$. Therefore, if a process $X(\mu, t)$ has a finite energy in $[0, T]$, this means $\sum_{i=1}^{\infty} \lambda_i < \infty$.

3. The process needs not to be stationary.

- Defn: A random process $X(\mu, t)$ is said to be continuous in the mean square sense iff
$$\lim_{\Delta t \to 0} E\{|X(\mu, t) - X(\mu, t + \Delta t)|^2\} = 0, \qquad \forall t.$$
Note that a random process can be continuous in the other senses, including everywhere, almost everywhere (i.e., with probability one), in probability, and in distribution. We are used to saying that a random process is continuous without specifying any probabilistic sense, but usually with an implication of continuity in distribution, i.e., its n-th order distribution is continuous in all time arguments for all positive integers n.

- Lemma: The autocorrelation $R_X(t, s)$ of a random process $X(\mu, t)$ is continuous in s and t for $0 < s < T$ and $0 < t < T$ iff $X(\mu, t)$ is continuous in the mean square sense.
Proof: Consider complex-valued $X(\mu, t)$'s.
" $\Rightarrow$ "
$$\begin{aligned} E\{|X(\mu, t) - X(\mu, t + \Delta t)|^2\} &= (R_X(t, t) - R_X(t, t + \Delta t)) \\ &\quad + (R_X(t + \Delta t, t + \Delta t) - R_X(t + \Delta t, t)). \end{aligned}$$
If $R_X(t, s)$ is continuous in t and s, then
$$R_X(t, t) - R_X(t, t + \Delta t) \stackrel{\Delta t \to 0}{\longrightarrow} 0$$
$$R_X(t + \Delta t, t + \Delta t) - R_X(t + \Delta t, t) \stackrel{\Delta t \to 0}{\longrightarrow} 0.$$
Hence, $X(\mu, t)$ is continuous in the mean square sense.
" $\Leftarrow$ "
$$\begin{aligned} & |E\{[X(\mu, t + \Delta t) - X(\mu, t)]X^*(\mu, s)\}| \\ = \ & |R_X(t + \Delta t, s) - R_X(t, s)| \\ \leq \ & \sqrt{E\{|X(\mu, s)|^2\}} \underbrace{\sqrt{E\{|X(\mu, t + \Delta t) - X(\mu, t)|^2\}}}_{\longrightarrow 0 \ as \ \Delta t \to 0} \\ \to \ & 0 \text{ as } \Delta t \to 0 \end{aligned}$$
by Schwartz's inequality, i.e., $|E\{F(u)G(u)\}|^2 \leq E\{|F(u)|^2\}E\{|G(u)|^2\}$. Therefore, $R_X(t, s)$ is continuous in t. Similarly, we can also prove that $R_X(t, s)$ is continuous in s. Q.E.D.

4.1 Theorem (The Karhunen-Loéve Expansion)

- Without loss of generality, let us assume zero-mean random processes. If not, then we deal with autocovariance instead of autocorrelation.

- Any complex-valued random process, which is continuous in the mean square sense and whose autocorrelation satisfies

$$\int_0^T \int_0^T |R_X(t,s)|^2 dtds < \infty \qquad (\&)$$

has the representation

$$X(\mu,t) = \underset{N\to\infty}{l.i.m.} \sum_{k=1}^{N} a_k(\mu)\phi_k(t)$$

for all $t \in [0,T]$, where the random coefficients $\{a_k(\mu)\}$ are given by

$$a_k(\mu) = \int_0^T X(\mu,t)\phi_k^*(t)dt$$

and where $\phi_k(t)$'s are the eigenfunctions of $R_X(t,s)$ with the associated eigenvalues λ_k's, i.e.,

$$\int_0^T R_X(t,s)\phi_k(s)ds = \lambda_k\phi_k(t) \qquad t \in [0,T]$$

$$\int_0^T \phi_k(t)\phi_i^*(t)dt = \delta_{ki}.$$

These $a_k(\mu)$'s are uncorrelated, i.e., $E\{a_k(\mu)a_i^*(\mu)\} = E\{a_k(\mu)\}E\{a_i^*(\mu)\}$ if $k \neq i$, and $E\{|a_k(\mu)|^2\} = \lambda_k$.

Note: If $X(\mu,t)$ is Gaussian, $a_k(\mu)$'s are independent.

Proof:
(1) First, we notice that

$$\begin{aligned} E\{a_k(\mu)a_l^*(\mu)\} &= E\{\int_0^T X(\mu,t)\phi_k^*(t)dt \int_0^T X^*(\mu,s)\phi_l(s)ds\} \\ &= \int_0^T \int_0^T R_X(t,s)\phi_k^*(t)\phi_l(s)dtds. \end{aligned}$$

From Mercer's theorem

$$R_X(t,s) = \sum_{i=1}^{\infty} \lambda_i \phi_i(t)\phi_i^*(s)$$

we have

$$\begin{aligned} E\{a_k(\mu)a_l^*(\mu)\} &= \sum_{i=1}^{\infty} \lambda_i \overbrace{\int_0^T \underbrace{\int_0^T \phi_i(t)\phi_k^*(t)dt}_{\delta_{ik}} \cdot \phi_i^*(s)\phi_l(s)ds}^{\delta_{il}} \\ &= \sum_{i=1}^{\infty} \lambda_i \delta_{ik}\delta_{il} = \lambda_k \delta_{kl}. \end{aligned}$$

Hence, $a_k(\mu)$'s are mutually uncorrelated and have $E\{|a_k(\mu)|^2\} = \lambda_k$.
(2) Now, to show

$$X(\mu,t) = \underset{N\to\infty}{l.i.m.} \sum_{k=1}^{N} a_k(\mu)\phi_k(t), \qquad t \in [0,T]$$

we have to prove

$$\lim_{N\to\infty} E\{|X(\mu,t) - \sum_{k=1}^{N} a_k(\mu)\phi_k(t)|^2\} = 0, \qquad t \in [0,T].$$

Note: From Theorem 4.1 in the book "Stochastic Processes" by J. Doob, if $Y_1(\mu), Y_2(\mu), ...$ is a sequence of uncorrelated random variables with $E\{|Y_n(\mu)|^2\} = \sigma_n^2$, then $\sum_{n=1}^{\infty} Y_n(\mu)$ converges in the mean square sense iff $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$.
From this theorem, we note that the sequence

$$Y_N(\mu,t) \triangleq \sum_{k=1}^{N} a_k(\mu)\phi_k(t)$$

for a fixed t, approaches a limit in the mean square sense iff the sequences

$$E\{\sum_{k=1}^{N} a_k(\mu)\phi_k(t)\} = 0$$

and

$$Var\{\sum_{k=1}^{N} a_k(\mu)\phi_k(t)\} = \sum_{k=1}^{N} \lambda_k|\phi_k(t)|^2 \quad \underset{\substack{\text{Mercer's}\\ \text{theorem}}}{\overset{N\to\infty}{\longrightarrow}} \quad R_X(t,t) = E\{|X(\mu,t)|^2\}$$

have limits. Since $R_X(t,s)$ satisfies the hypothesis of Mercer's theorem,

$$R_X(t,s) = \sum_{i=1}^{\infty} \lambda_i \phi_i(t)\phi_i^*(s)$$

and thus $R_X(t,t) = E\{|X(\mu,t)|^2\} = \sum\limits_{k=1}^{\infty} \lambda_k|\phi_k(t)|^2$. Because $R_X(t,s)$ satisfies (&), $Var\{\sum\limits_{k=1}^{N} a_k(\mu)\phi_k(t)\}$ has a limit given by finite-valued $E\{|X(\mu,t)|^2\}$. Therefore, the limit $\sum_{k=1}^{\infty} a_k(\mu)\phi_k(t)$ exists in the mean square sense. Now, we can have

$$\begin{aligned}
&\lim_{N\to\infty} E\{|X(\mu,t) - \sum_{k=1}^{N} a_k(\mu)\phi_k(t)|^2\} \\
= &\lim_{N\to\infty} \{R_X(t,t) - \sum_{k=1}^{N} E\{X(\mu,t)a_k^*(\mu)\}\phi_k^*(t) \\
&- \sum_{k=1}^{N} E\{X^*(\mu,t)a_k(\mu)\}\phi_k(t) + \sum_{k=1}^{N} \lambda_k|\phi_k(t)|^2\}.
\end{aligned}$$

Because

$$\begin{aligned}
\sum_{k=1}^{N} E\{X(\mu,t)a_k^*(\mu)\}\phi_k^*(t) &= \sum_{k=1}^{N} E\{X(\mu,t)\int_0^T X^*(\mu,\alpha)\phi_k(\alpha)d\alpha\}\phi_k^*(t) \\
&= \sum_{k=1}^{N} \underbrace{\int_0^T R_X(t,\alpha)\phi_k(\alpha)d\alpha}_{=\lambda_k\phi_k(t)}\phi_k^*(t) \\
&= \sum_{k=1}^{N} \lambda_k|\phi_k(t)|^2 \\
&= \sum_{k=1}^{N} E\{X^*(\mu,t)a_k(\mu)\}\phi_k(t)
\end{aligned}$$

we further have

$$
\begin{aligned}
\lim_{N\to\infty} E\{|X(\mu,t) - \sum_{k=1}^{N} a_k(\mu)\phi_k(t)|^2\} &= \lim_{N\to\infty} \{R_X(t,t) - \sum_{k=1}^{N} \lambda_k|\phi_k(t)|^2\} \\
&= R_X(t,t) - \underbrace{\lim_{N\to\infty} \sum_{k=1}^{N} \lambda_k|\phi_k(t)|^2}_{R_X(t,t)} \\
&= 0.
\end{aligned}
$$

Q.E.D.

- Two important facts:

  1. It is only necessary that $X(\mu,t)$ is continuous in the mean square sense and $R_X(t,s)$ be square-integrable in order for Karhunen-Loéve expansion to exist.
  2. The expansion is impossible if $R_X(t,s)$ is not square-integrable.

4.2 Relationship Between Fourier Series Expansion for Random Processes and KL Expansion

- Defn: An rp $X(\mu,t)$ is called MS periodic with period T iff $E\{|X(\mu,t+T) - X(\mu,t)|^2\} = 0$ for all t.

  Thus, a real-valued WSS random process $X(\mu,t)$ is MS periodic if its $R_X(\tau)$ is periodic with period T. Now, if we expand $R_X(\tau)$ into Fourier series, we have

  $$R_X(\tau) = \sum_{n=-\infty}^{\infty} r_n e^{j2\pi\frac{n}{T}\tau}$$

  with

  $$r_n = \frac{1}{T}\int_0^T R_X(\tau)e^{-j2\pi\frac{n}{T}\tau}d\tau.$$

- Theorem: Given a real-valued zero-mean WSS MS periodic process $X(\mu,t)$ with period T, we can represent it by the series

  $$X(\mu,t) = \underset{N\to\infty}{l.i.m.} \sum_{n=-N}^{N} a_n(\mu)e^{j2\pi\frac{n}{T}t}$$

where

$$a_n(\mu) = \frac{1}{T}\int_0^T X(\mu,t)e^{-j2\pi\frac{n}{T}t}dt.$$

Furthermore, $a_n(\mu)$'s are uncorrelated with

$$\begin{aligned} E\{a_n(\mu)\} &= 0 \qquad \text{(E1)} \\ E\{a_n(\mu)a_m^*(\mu)\} &= \begin{cases} r_n, n=m \\ 0,\ n\neq m \end{cases}. \qquad \text{(E2)} \end{aligned}$$

Note that $r_n \geq 0$ for all n.

Proof:

1. (E1) is straightforward. Let us prove (E2) below:

$$\begin{aligned} E\{a_n(\mu)X^*(\mu,\alpha)\} &= E\{\frac{1}{T}\int_0^T X(\mu,t)e^{-j2\pi\frac{n}{T}t}dt\cdot X^*(\mu,\alpha)\} \\ &= \frac{1}{T}\int_0^T R_X(t-\alpha)e^{-j2\pi\frac{n}{T}t}dt \\ x=t-\alpha \Rightarrow \quad &= \frac{1}{T}\int_{-\alpha}^{T-\alpha} \underbrace{R_X(x)e^{-j2\pi\frac{n}{T}x}}_{\text{periodic with period T}}\ dx\cdot e^{-j2\pi\frac{n}{T}\alpha} \\ &= \underbrace{\frac{1}{T}\int_0^T R_X(x)e^{-j2\pi\frac{n}{T}x}dx}_{r_n}\cdot\ e^{-j2\pi\frac{n}{T}\alpha} \\ &= r_n e^{-j2\pi\frac{n}{T}\alpha} \qquad (\times) \end{aligned}$$

$$\begin{aligned} E\{a_n(\mu)a_m^*(\mu)\} &= E\{\frac{1}{T}\int_0^T a_n(\mu)X^*(\mu,t)e^{j2\pi\frac{m}{T}t}dt\} \\ (\times)\Rightarrow \quad &= \frac{1}{T}\int_0^T r_n e^{-j2\pi\frac{n}{T}t}e^{j2\pi\frac{m}{T}t}dt \\ &= \begin{cases} r_n, n=m \\ 0,\ n\neq m \end{cases}. \end{aligned}$$

2. To prove

$$X(\mu,t) = \underset{N\to\infty}{l.i.m.}\sum_{n=-N}^{N} a_n(\mu)e^{j2\pi\frac{n}{T}t}$$

   we need to show

$$E\{|X(\mu,t) - \sum_{n=-\infty}^{\infty} a_n(\mu)e^{j2\pi\frac{n}{T}t}|^2\} = 0.$$

First, we find that

$$
\begin{aligned}
E\{|\sum_{n=-\infty}^{\infty} a_n(\mu)e^{j2\pi\frac{n}{T}t}|^2\} &= \sum_{n=-\infty}^{\infty} E\{|a_n(\mu)|^2\} \\
&(\because \quad a_n(\mu)\text{'s are uncorrelated with mean zero.}) \\
&= \sum_{n=-\infty}^{\infty} r_n \\
&= R_X(0) \\
&= E\{|X(\mu,t)|^2\}.
\end{aligned}
$$

Next,

$$
\begin{aligned}
E\{\sum_{n=-\infty}^{\infty} a_n(\mu)e^{j2\pi\frac{n}{T}t}X^*(\mu,t)\} &= \sum_{n=-\infty}^{\infty} E\{a_n(\mu)X^*(\mu,t)\}e^{j2\pi\frac{n}{T}t} \\
(\times) \Rightarrow \quad &= \sum_{n=-\infty}^{\infty} r_n \\
&= E\{\sum_{n=-\infty}^{\infty} [a_n(\mu)e^{j2\pi\frac{n}{T}t}]^* X(\mu,t)\}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
& E\{|X(\mu,t) - \sum_{n=-\infty}^{\infty} a_n(\mu)e^{-j2\pi\frac{n}{T}t}|^2\} \\
= \; & E\{|X(\mu,t)|^2\} + E\{|\sum_{n=-\infty}^{\infty} a_n(\mu)e^{-j2\pi\frac{n}{T}t}|^2\} \\
& -E\{X(\mu,t)\cdot \sum_{n=-\infty}^{\infty} a_n^*(\mu)e^{j2\pi\frac{n}{T}t}\} - E\{X^*(\mu,t)\cdot \sum_{n=-\infty}^{\infty} a_n(\mu)e^{-j2\pi\frac{n}{T}t}\} \\
= \; & 0 \qquad Q.E.D.
\end{aligned}
$$

- Note: The above Fourier series expansion for WSS MS periodic processes is a special case of the KL expansion which holds good for both stationary and nonstationary processes!!

4.3 Examples for KL Expansion

- Ex 1:
  Let $X(\mu,t)$ be a real-valued rp with zero mean and

$$
R_X(t_1,t_2) = \alpha \min(t_1,t_2) = \begin{cases} \alpha t_2, t_1 > t_2 \\ \alpha t_1, t_2 \geq t_1 \end{cases}
$$

with $\alpha > 0$. First, let us find the corresponding $\{\phi_n(t)\}$ and $\{\lambda_n\}$ by

$$\int_0^T R_X(t_1, t_2)\phi_n(t_2)dt_2 = \lambda_n \phi_n(t_1) \qquad 0 \le t_1 \le T$$

$$\Rightarrow \alpha \int_0^{t_1} t_2 \phi_n(t_2)dt_2 + \alpha t_1 \int_{t_1}^T \phi_n(t_2)dt_2 = \lambda_n \phi_n(t_1) \qquad 0 \le t_1 \le T. \qquad (*)$$

Differentiating (*) with respect to t_1, we get

$$\alpha t_1 \phi_n(t_1) + \alpha \int_{t_1}^T \phi_n(t_2)dt_2 - \alpha t_1 \phi_n(t_1) = \lambda_n \phi_n^{'}(t_1) \qquad 0 \le t_1 \le T$$

$$\Rightarrow \alpha \int_{t_1}^T \phi_n(t_2)dt_2 = \lambda_n \phi_n^{'}(t_1) \qquad 0 \le t_1 \le T \qquad (+)$$

and differentiating (+) with respect to t_1, we further have

$$-\alpha \phi_n(t_1) = \lambda_n \phi_n^{''}(t_1) \qquad 0 \le t_1 \le T. \qquad (\text{—})$$

Now choosing $\phi_n(t)$'s such that $\phi_n(0) = 0$ and $\phi_n^{'}(T) = 0$, we can solve (—) and obtain

$$\phi_n(t) = \sqrt{\frac{2}{T}} \sin(\omega_n t)$$

where $\omega_n = \sqrt{\frac{\alpha}{\lambda_n}} = \frac{2n+1}{2T}\pi$ with $\lambda_n = \frac{4T^2\alpha}{\pi^2(2n+1)^2}$. Thus, $X(\mu, t)$ can be expanded as

$$X(\mu, t) = \sum_{n=1}^{\infty} a_n(\mu) \cdot \sqrt{\frac{2}{T}} \sin(\omega_n t)$$

in the mean square sense and

$$a_n(\mu) = \sqrt{\frac{2}{T}} \int_0^T X(\mu, t) \sin(\omega_n t)dt$$

where $a_n(\mu)$'s are uncorrelated with $E\{a_n^2(\mu)\} = \lambda_n$.

- Ex 2: A Binary Hypothesis Testing Problem

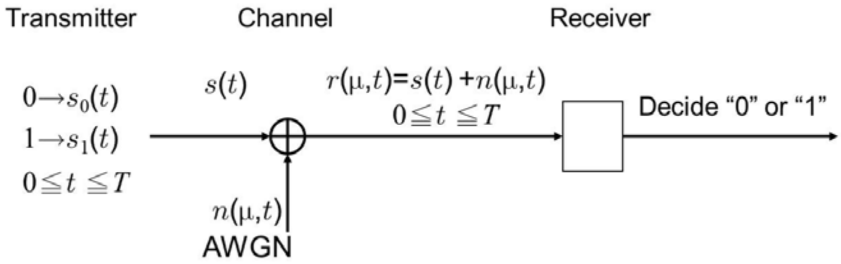


Figure 18:

Consider the binary real-valued waveform source which, over the time interval $[0, T]$, emits $s(t) = s_0(t)$ when bit 0 is generated and $s(t) = s_1(t)$ when bit 1 is generated. The signal transmission is disturbed by the (real-valued) stationary additive white Gaussian noise $n(\mu, t)$, which has zero mean and autocorrelation function $R_n(\tau) = \frac{N_0}{2}\delta(\tau)$ with $\frac{N_0}{2}$ being the two-sided power spectral density (PSD) level. The receiver observes the received waveform $r(\mu, t) = s(t) + n(\mu, t)$ over $t \in [0, T]$ and intends to determine whether bit 0 or 1 is currently transmitted. To facilitate the receiver design, it is advantageous to process samples instead of waveforms, and thus a transform from waveform to samples is required at the receiver frontend. KL expansion is one approach to transform $r(\mu, t)$ into samples without loss of any information in the mean square sense.
First, expand $n(\mu, t)$. Because $R_n(\tau) = \frac{N_0}{2}\delta(\tau)$,

$$\begin{aligned}\lambda_n \phi_n(\tau) &= \int_0^T R_n(\tau - s)\phi_n(s)ds = \int_0^T \frac{N_0}{2}\delta(\tau - s)\phi_n(s)ds \\ &= \frac{N_0}{2}\phi_n(\tau) \qquad \text{for } \tau \in [0, T] \\ \Longrightarrow \lambda_n &= \frac{N_0}{2} \qquad \forall n.\end{aligned}$$

This means that any orthonormal function set $\{\phi_n(\tau)\}$ (i.e., $\int_0^T \phi_n(\tau) \cdot\phi_k(\tau)d\tau = \delta_{nk}$) can be used to expand $r(\mu, t)$, with the same eigenvalue $\lambda_n = \frac{N_0}{2}$ for all n.

If $s_0(t) = -s_1(t)$ (antipodal signaling), then we can choose $\phi_1(t) =$

$s_0(t)$, and represent

$$r(\mu,t)=\begin{cases} s_0(t)+a_1(\mu)\underbrace{\phi_1(t)}_{s_0(t)}+\begin{array}{l}\text{terms independent of } a_1(\mu)\\ \text{and irrelevant to } s_0(t)\end{array} & \text{, if 0 was sent}\\ -s_0(t)+a_1(\mu)\underbrace{\phi_1(t)}_{s_0(t)}+\begin{array}{l}\text{terms independent of } a_1(\mu)\\ \text{and irrelevant to } s_0(t)\end{array} & \text{, if 1 was sent}\end{cases}$$

where $a_1(\mu)=\int_0^T n(\mu,t)\underbrace{\phi_1(t)}_{s_0(t)}dt$. In this case, the following preprocessor can extract sufficient statistic from $r(\mu,t)$:

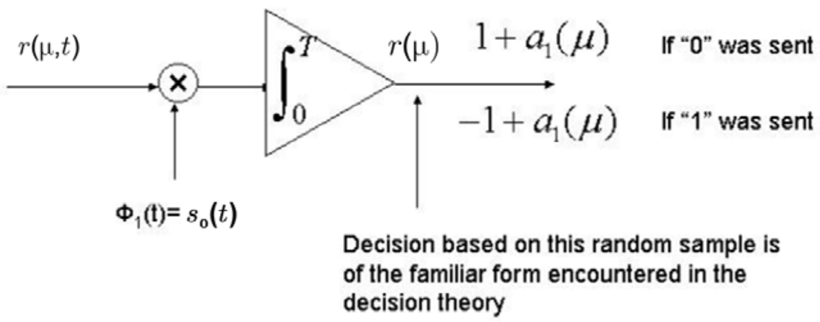


Figure 19:

5 Narrowband Processes and Bandpass Systems

- (Materials come from Davenport and Root)

5.1 Motivation for this topic:

- Most of the signals and systems of interest in the communication world are narrowband and bandpass in nature.

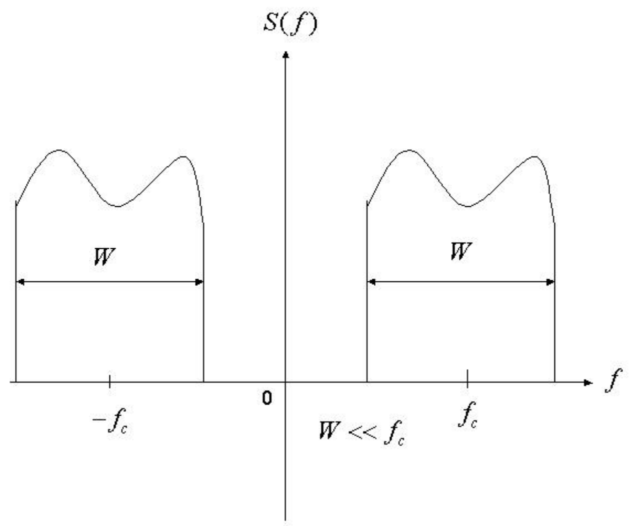


Figure 20:

5.2 Hilbert Transform

- Motivation for using Hilbert transform

  Let $g(t)$ be a real function. Then, its Fourier transform $G(f) = \mathcal{F}\{g(t)\}$ has the Hermitian symmetric property $G(-f) = G^*(f)$. This means that $g(t)$ is totally specified by $G(f)$ for $f \geq 0$. Now,

$$
\begin{aligned}
g(t) &= \int_{-\infty}^{\infty} G(f) e^{j2\pi ft} df \\
&= \int_{0}^{\infty} G(f) e^{j2\pi ft} df + \int_{-\infty}^{0} \underbrace{G(f)}_{G^*(-f)} e^{j2\pi ft} df \\
&= \int_{0}^{\infty} (G(f) e^{j2\pi ft} + G^*(f) e^{-j2\pi ft}) df \\
&= \mathrm{Re}\{\int_{0}^{\infty} 2G(f) e^{j2\pi ft} df\}.
\end{aligned}
$$

Thus, if we define

$$
G_+(f) = \begin{cases} 2G(f), & f \geq 0 \\ 0, & f < 0 \end{cases} = 2U(f)G(f)
$$

where $U(f) = \{\begin{smallmatrix} 1, & f \geq 0 \\ 0, & f < 0 \end{smallmatrix}$ is the unit step function, we have

$$
g(t) = \mathrm{Re}\{\int_{-\infty}^{\infty} G_+(f) e^{j2\pi ft} df\}.
$$

- Defn: $g_+(t) \triangleq \mathcal{F}^{-1}\{G_+(f)\} = \int_{-\infty}^{\infty} G_+(f) e^{j2\pi ft} df$ is called the preenvelope, or the analytical signal, of $g(t)$.

  Note that $g(t) = \mathrm{Re}\{g_+(t)\}$.
  Now, the question arises: if $g(t) = \mathrm{Re}\{g_+(t)\}$, then what is $\mathrm{Im}\{g_+(t)\}$? Let us call $\mathrm{Im}\{g_+(t)\} = \widehat{g}(t)$ where $\widehat{g}(t)$ is the Hilbert transform of $g(t)$. Clearly,

$$
\begin{aligned}
j\widehat{g}(t) &= g_+(t) - g(t) \\
\Rightarrow j\widehat{G}(f) &= G_+(f) - G(f) \\
\Rightarrow \widehat{G}(f) &= \mathcal{F}\{\widehat{g}(t)\} = (-j)(G_+(f) - G(f)) \\
&= -j(2U(f) - 1)G(f) = -jsgn(f)G(f)
\end{aligned}
$$

with $sgn(f) = \{\begin{smallmatrix} +1, & f \geq 0 \\ -1, & f < 0 \end{smallmatrix}$. So,

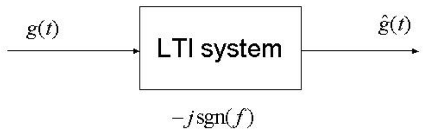


Figure 21:

Let us denote the Hilbert transform by $\widehat{g}(t) = H.T.\{g(t)\}$.

Notes:

1. $H.T.$ takes us from time domain to time domain.
2. Since $\mathcal{F}\left\{\frac{1}{\pi t}\right\} = -jsgn(f)$, we can create a $\widehat{g}(t)$ by passing $g(t)$ through an (ideal) LTI system with impulse response $h(t) = \frac{1}{\pi t}$.
3. For a real signal $g(t)$, its $g_+(t)$ is sufficient to carry all the information with it, but only with half bandwidth.

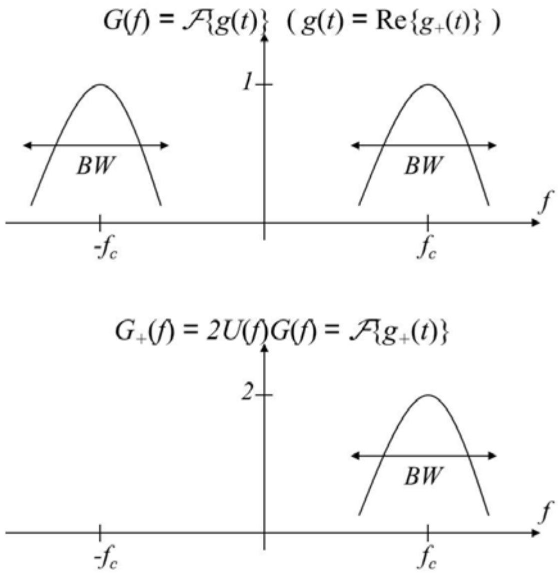


Figure 22:

- Properties of $H.T.$

(1) Now,

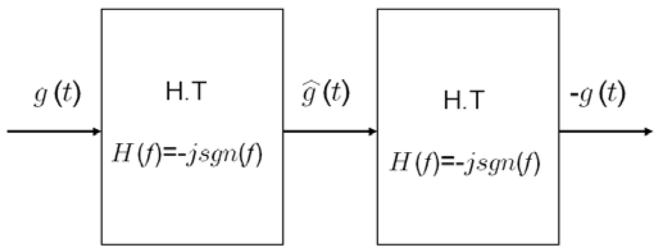


Figure 23:

since $X(f) = G(f)H^2(f) = G(f)(-jsgn(f))^2 = -G(f)$.

(2) $|H(f)| = |-jsgn(f)| = 1$, all pass.

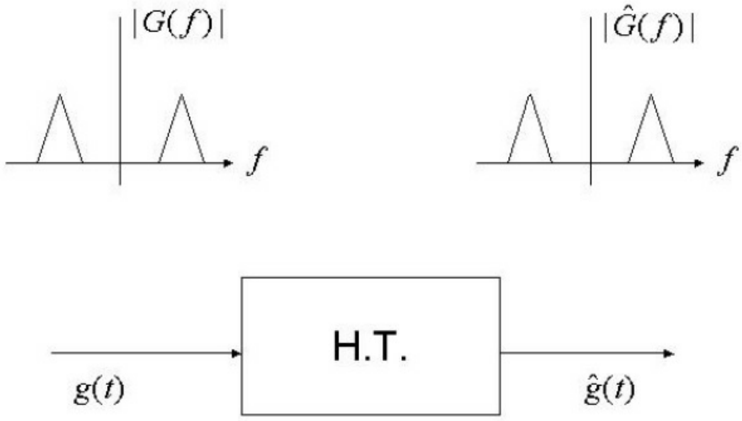


Figure 24:

(3) $\angle H(f) = \begin{cases} -90^\circ, & f \geq 0 \\ +90^\circ, & f < 0 \end{cases}$.
(2), (3)$\Rightarrow$ $H.T.$ is a phase shift system.

(4) Bedrosian's Theorem (1963):
Let $x_L(t)$ and $x_H(t)$ be Fourier transformable real-valued functions with the property that their transforms $X_L(f)$ and $X_H(f)$ are such: For some $\omega > 0$, $X_L(f) = 0$ for $|f| > \omega$ and $X_H(f) = 0$ for $|f| < \omega$. Thus,

$$H.T.[x_L(t)x_H(t)] = x_L(t)H.T.[x_H(t)].$$

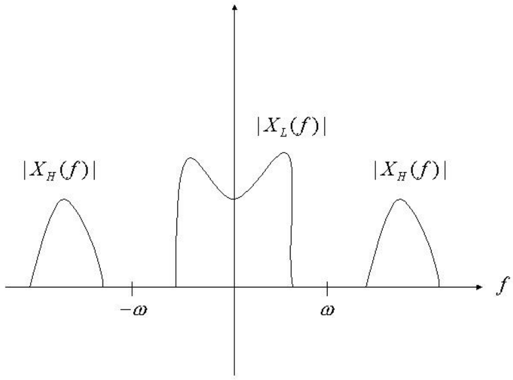


Figure 25:

Example: Let $g(t)$ be a lowpass pulse shape with bandwidth much smaller than f_c. We set $x_L(t) = g(t)$ and $x_H(t) = \cos(2\pi f_c t + \theta)$, and according to the Theorem,

$$H.T.[g(t)\cos(2\pi f_c t + \theta)] = g(t)H.T.[\cos(2\pi f_c t+\theta)] = g(t)\sin(2\pi f_c t+\theta).$$

Proof:

$$\mathcal{F}\{H.T.[x_L(t)x_H(t)]\} \triangleq \widehat{S}(f) = (-jsgn(f))S(f)$$

with

$$\begin{aligned} S(f) &= \mathcal{F}\{x_L(t)x_H(t)\} \\ &= \int_{-\infty}^{\infty} X_H(f-\lambda)X_L(\lambda)d\lambda \\ &= X_H(f) * X_L(f). \end{aligned}$$

So,

$$
\begin{aligned}
\widehat{s}(t) &= \int_{-\infty}^{\infty} \widehat{S}(f) e^{j2\pi f t} df \\
&= \int_{-\infty}^{\infty} (-jsgn(f)) S(f) e^{j2\pi f t} df \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (-jsgn(f)) X_H(f-\lambda) X_L(\lambda) e^{j2\pi f t} df d\lambda \\
(\tau = f - \lambda \Rightarrow) \quad &= (-j) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} sgn(\tau+\lambda)) X_H(\tau) X_L(\lambda) e^{j2\pi(\tau+\lambda)t} d\tau d\lambda.
\end{aligned}
$$

Since $X_H(\tau)X_L(\lambda)$ is nonzero for $|\tau| > \omega$ and $|\lambda| < \omega$, therefore $sgn(\tau+\lambda) = sgn(\tau)$. Thus , we can rewrite

$$
\begin{aligned}
\widehat{s}(t) = &\quad \underbrace{\int_{-\infty}^{\infty} X_L(\lambda) e^{j2\pi\lambda t} d\lambda}_{x_L(t)} \cdot \quad \underbrace{\int_{-\infty}^{\infty} (-jsgn(\tau)) X_H(\tau) e^{j2\pi\tau t} d\tau}_{\widehat{x}_H(t)} \\
= &\quad x_L(t) \cdot H.T.[x_H(t)]. \qquad Q.E.D.
\end{aligned}
$$

(5) A real signal $g(t)$ and its Hilbert transform $\widehat{g}(t)$ have the same spectral density. (This holds for both power and energy densities.)
Proof: For deterministic signals,

$$
|\widehat{G}(f)|^2 = |G(f)H(f)|^2 = |G(f)(-jsgn(f))|^2 = |G(f)|^2.
$$

For random WSS signals,

$$
\begin{aligned}
S_{\widehat{g}}(f) &= S_g(f)|H(f)|^2 = S_g(f)|-jsgn(f)|^2 = S_g(f). \\
& \qquad\qquad\qquad\qquad\qquad\qquad Q.E.D.
\end{aligned}
$$

Corollary: They have the same autocorrelation function if they are random and WSS.

(6) $g(t)$ and $\widehat{g}(t)$ are orthogonal, i.e., $\int_{-\infty}^{\infty} g(t)\widehat{g}(t)dt = 0$.
Proof:

$$
\begin{aligned}
\int_{-\infty}^{\infty} g(t)\widehat{g}(t)dt &= \int_{-\infty}^{\infty} g^*(t)\widehat{g}(t)dt \\
\text{(Parseval's relation)} &= \int_{-\infty}^{\infty} G^*(f) \underbrace{\widehat{G}(f)}_{G(f)(-jsgn(f))} df \\
&= \int_{-\infty}^{\infty} (-jsgn(f))|G(f)|^2 df \\
&= j\int_{-\infty}^{0} \underbrace{|G(f)|^2}_{even} df - j\int_{0}^{\infty} |G(f)|^2 df \\
&= j\int_{0}^{\infty} |G(f)|^2 df - j\int_{0}^{\infty} |G(f)|^2 df \\
&= 0 \qquad Q.E.D.
\end{aligned}
$$

- Example:
  Let $g(t) = \cos(2\pi f_c t), \quad f_c > 0$.
  $\Rightarrow G(f) = \frac{1}{2}[\delta(f + f_c) + \delta(f - f_c)]$
  $\Rightarrow \widehat{G}(f) = G(f)(-jsgn(f)) = \frac{1}{2}[j\delta(f + f_c) - j\delta(f - f_c)]$
  $\Rightarrow \widehat{g}(t) = \sin(2\pi f_c t)$.
  Also, $H.T.[\sin(2\pi f_c t)] = -\cos(2\pi f_c t)$.

5.3 Bandpass Signals and Bandpass Linear Systems

- A real signal $g(t)$ is said to be bandpass if $G(f)$ is virtually nonzero within some band, centered on or around some high (with respect to the bandwidth of $G(f)$) frequency f_c.

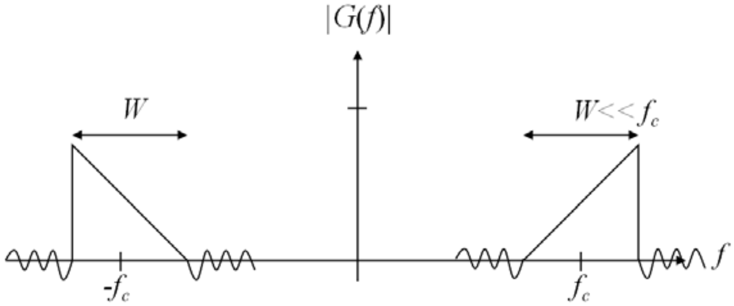


Figure 26:

- Defn: The complex envelope $\widetilde{g}(t)$ is defined as $\widetilde{g}(t) \triangleq g_+(t)e^{-j2\pi f_c t}$. This gives $g_+(t) = \widetilde{g}(t)e^{j2\pi f_c t}$ and $\widetilde{G}(f) = G_+(f + f_c)$.

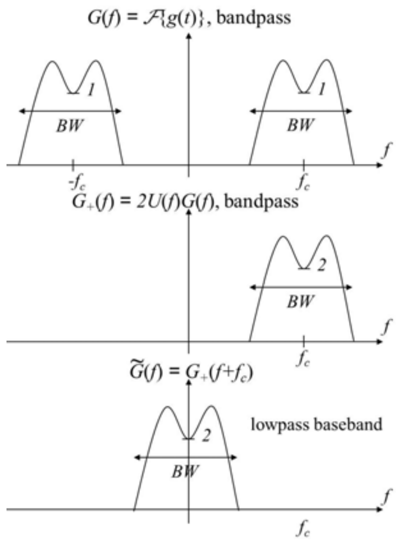


Figure 27:

- Facts: If $g(t)$ is bandpass, so are $\widehat{g}(t)$ and $g_+(t)$. However, $\widetilde{g}(t)$ is lowpass.

- Quadrature Representation of a Bandpass Signal:
First, recall that $g(t) = \mathrm{Re}\{g_+(t)\} = \mathrm{Re}\{\widetilde{g}(t)e^{j2\pi f_c t}\}$.
Now, suppose that we can write $\widetilde{g}(t)$ as $\widetilde{g}(t) = g_c(t) + jg_s(t)$ where $g_c(t)$ and $g_s(t)$ are both real functions of time. Then, we can express $g(t)$ by

$$\begin{aligned} g(t) &= \mathrm{Re}\{(g_c(t) + jg_s(t))e^{j2\pi f_c t}\} \\ \Rightarrow \underbrace{g(t)}_{bandpass} &= \underbrace{g_c(t)}_{lowpass} \cos(2\pi f_c t) - \underbrace{g_s(t)}_{lowpass} \sin(2\pi f_c t) \end{aligned}$$

where $g_c(t)$ and $g_s(t)$ are called the inphase and quadrature-phase components, respectively, of $g(t)$.

- Lemma:

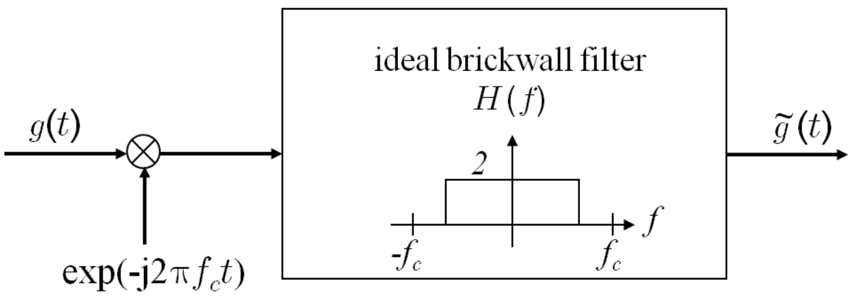


Figure 28:

Proof: Since $g(t) = \mathrm{Re}\{\widetilde{g}(t)e^{j2\pi f_c t}\}$,

$$\begin{aligned} g(t)e^{-j2\pi f_c t} &= \mathrm{Re}\{\widetilde{g}(t)e^{j2\pi f_c t}\}e^{-j2\pi f_c t} \\ &= \frac{1}{2}[\widetilde{g}(t)e^{j2\pi f_c t} + \widetilde{g}^*(t)e^{-j2\pi f_c t}]e^{-j2\pi f_c t} \\ &= \frac{1}{2}\widetilde{g}(t) + \underbrace{\frac{1}{2}\widetilde{g}^*(t)e^{-j2\pi(2f_c)t}}_{\text{filtered out by the filter}} \end{aligned}$$

$\Rightarrow$ The filter output $= \widetilde{g}(t)$. *Q.E.D.*

- Note: The actual implementation structure looks like the following:

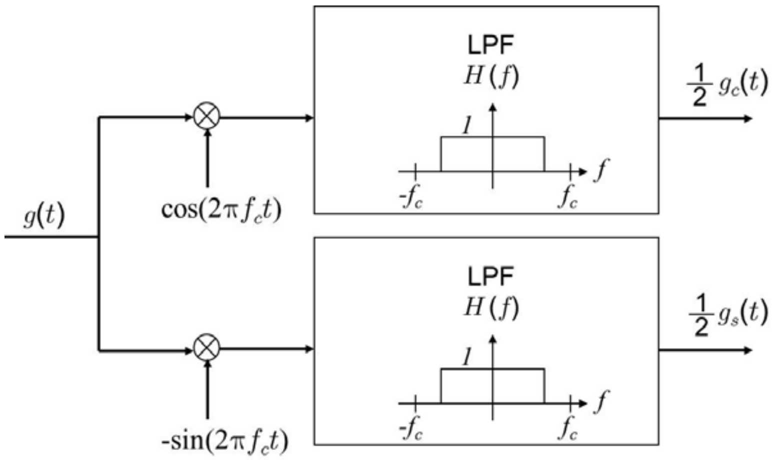


Figure 29:

- Instead of quadrature representation, we could have a polar description of a bandpass signal in terms of amplitude and phase functions. Let

$$\widetilde{g}(t) = v(t)e^{j\phi(t)}$$

where $v(t)$ and $\phi(t)$ are real functions of time. Then,

$$\begin{aligned} g(t) &= \mathrm{Re}\{\widetilde{g}(t)e^{j2\pi f_c t}\} \\ &= \mathrm{Re}\{v(t)e^{j(2\pi f_c t+\phi(t))}\} \\ &= v(t)\cos(2\pi f_c t + \phi(t)) \end{aligned}$$

where

$$\begin{aligned} v(t) &\equiv \text{amplitude, or (nature) envelope, of } g(t) \\ &= |\widetilde{g}(t)| = |g_+(t)| \end{aligned}$$

and

$$\begin{aligned} \phi(t) &\equiv \text{phase of } g(t) \\ &= \angle\widetilde{g}(t). \end{aligned}$$

- Application: Bandpass Signals Through Bandpass Linear Systems

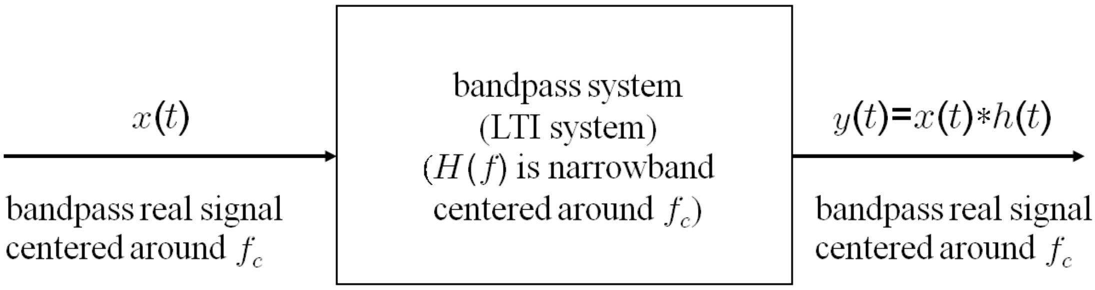


Figure 30:

Theorem: $\widetilde{y}(t) = \widetilde{x}(t) * \widetilde{h}_L(t)$ where $\widetilde{h}_L(t)$ is the lowpass equivalent impulse response of the bandpass filter given by

$$\begin{aligned}\widetilde{h}_L(t) &= \frac{1}{2}\widetilde{h}(t)\\ &= \mathcal{F}^{-1}\{H(f+f_c)U(f+f_c)\}.\end{aligned}$$

Proof: Considering that all functions are bandpass and real,

$$\begin{aligned}y(t) &= x(t) * h(t)\\ &= \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau. \qquad \oplus\end{aligned}$$

Denote the pre-envelope associated with $\widetilde{h}_L(t)$ by

$$\begin{aligned}h_{L,+}(t) &= \widetilde{h}_L(t)e^{j2\pi f_c t}\\ &= \frac{1}{2}\widetilde{h}(t)e^{j2\pi f_c t}.\end{aligned}$$

Then,

$$\begin{aligned}h(t) &= \mathrm{Re}\{\widetilde{h}(t)e^{j2\pi f_c t}\}\\ &= 2\,\mathrm{Re}\{h_{L,+}(t)\}.\end{aligned}$$

Also, let $x_+(t)$ be the pre-envelope of $x(t)$. Then, $x(t) = \mathrm{Re}\{x_+(t)\}$. Now, substituting $x_+(t)$ and $h_{L,+}(t)$ into $\oplus$, we have

$$y(t) = 2\int_{-\infty}^{\infty} \mathrm{Re}\{h_{L,+}(\tau)\}\,\mathrm{Re}\{x_+(t-\tau)\}d\tau.$$

Now that $\mathrm{Re}\{z\}\,\mathrm{Re}\{y\} = \frac{1}{2}\,\mathrm{Re}\{zy\} + \frac{1}{2}\,\mathrm{Re}\{zy^*\}$,

$$y(t) = \mathrm{Re}\{I_1\} + \mathrm{Re}\{I_2\}$$

where

$$\begin{aligned} I_1 &= \int_{-\infty}^{\infty} h_{L,+}(\tau)x_+(t-\tau)d\tau \\ I_2 &= \int_{-\infty}^{\infty} h_{L,+}(\tau)x_+^*(t-\tau)d\tau. \end{aligned}$$

First, we show that $I_2 = 0$: Now, if we let $g_+(\tau) = x_+(t-\tau)$, then

$$\begin{aligned} I_2 &= \int_{-\infty}^{\infty} h_{L,+}(\tau)g_+^*(\tau)d\tau \\ &= \int_{-\infty}^{\infty} H_{L,+}(f)G_+^*(f)df \end{aligned}$$

where $H_{L,+}(f) = \mathcal{F}\{h_{L,+}(\tau)\}$ and $G_+(f) = \mathcal{F}\{g_+(\tau)\}$. But,

$$\begin{aligned} G_+(f) &= \int_{-\infty}^{\infty} g_+(\tau)e^{-j2\pi f\tau}d\tau \\ &= \int_{-\infty}^{\infty} x_+(t-\tau)e^{-j2\pi f\tau}d\tau \\ (t' = t-\tau) &= \int_{-\infty}^{\infty} x_+(t')e^{j2\pi ft'}dt'e^{-j2\pi ft} \\ &= X_+(-f)e^{-j2\pi ft} \end{aligned}$$

where $X_+(f) = \mathcal{F}\{x_+(\tau)\}$. Thus $I_2 = \int_{-\infty}^{\infty} H_{L,+}(f)X_+^*(-f)e^{j2\pi ft}df = 0$ since $X_+(-f) = 0$ for $f > 0$ and $H_{L,+}(f) = 0$ for $f < 0$.

Next, return to I_1:

$$\begin{aligned} y(t) &= \mathrm{Re}\{I_1\} \\ &= \mathrm{Re}\{\int_{-\infty}^{\infty} h_{L,+}(\tau)x_+(t-\tau)d\tau\} \\ &= \mathrm{Re}\{\int_{-\infty}^{\infty} \widetilde{h}_L(\tau)e^{j2\pi f_c\tau}\widetilde{x}(t-\tau)e^{j2\pi f_c(t-\tau)}d\tau\} \\ &= \mathrm{Re}\{e^{j2\pi f_c t}\underbrace{\int_{-\infty}^{\infty} \widetilde{h}_L(\tau)\widetilde{x}(t-\tau)d\tau}_{\widetilde{x}(t)*\widetilde{h}_L(t)}\}. \end{aligned}$$

Thus, $\widetilde{y}(t) = \widetilde{x}(t) * \widetilde{h}_L(t)$.   *Q.E.D.*

In terms of quadrature representation,

$$\begin{aligned} & \widetilde{y}(t) = \widetilde{x}(t) * \widetilde{h}_L(t) \\ \Rightarrow\ & [y_c(t) + jy_s(t)] = [x_c(t) + jx_s(t)] * [h_{L,c}(t) + jh_{L,s}(t)] \\ \Rightarrow\ & \begin{cases} y_c(t) = x_c(t) * h_{L,c}(t) - x_s(t) * h_{L,s}(t) \\ y_s(t) = x_c(t) * h_{L,s}(t) + x_s(t) * h_{L,c}(t) \end{cases}. \end{aligned}$$

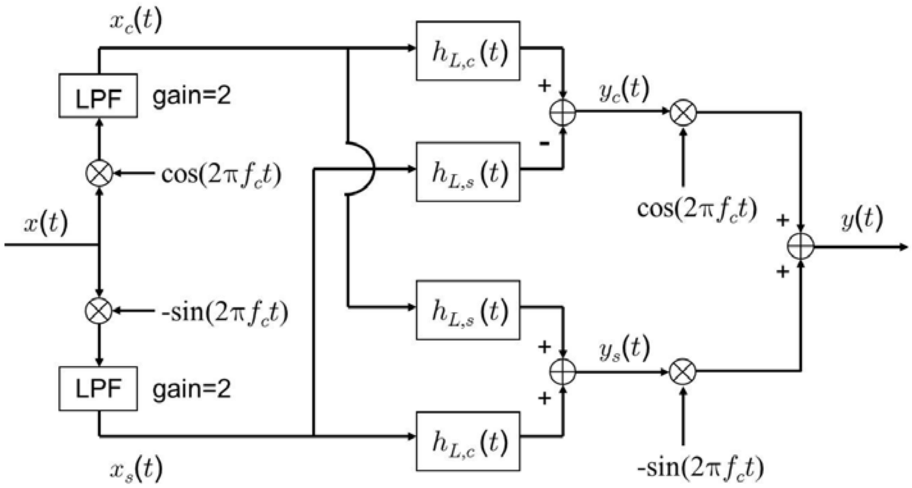


Figure 31:

5.4 Real-Valued Narrowband Noise Processes

- (We emphasize on Gaussian processes here.)
- If $n(\mu,t)$ is the output of passing real-valued WSS white noise $w(\mu,t)$ with PSD one and mean zero through an LTI system with real-valued impulse response $h(t)$, then its PSD $S_n(f)$ is real and even since

$$S_n(f) = 1\cdot |H(f)|^2 = |H^*(f)|^2 \underset{h(t)\text{ is real}}{=} |H(-f)|^2 = S_n(-f).$$

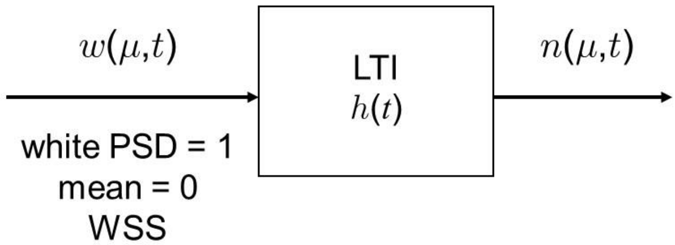


Figure 32:

- A Quadrature Representation of Narrowband $n(\mu,t)$:
  Recall that

$$\begin{aligned} n_+(\mu,t) &= n(\mu,t) + j\widehat{n}(\mu,t) \\ &= \widetilde{n}(\mu,t)e^{j2\pi f_c t} \end{aligned}$$

where f_c is arbitrary and positive. Note that $\widehat{n}(\mu,t) = n(\mu,t) * h(t)$ with $h(t) = \frac{1}{\pi t}$. If we define $\widetilde{n}(\mu,t) = n_c(\mu,t) + jn_s(\mu,t)$, then

$$\left.\begin{aligned} n_c(\mu,t) &= n(\mu,t)\cos(2\pi f_c t) + \widehat{n}(\mu,t)\sin(2\pi f_c t) \\ n_s(\mu,t) &= \widehat{n}(\mu,t)\cos(2\pi f_c t) - n(\mu,t)\sin(2\pi f_c t) \end{aligned}\right\} (A_1)$$

$$n(\mu,t) = n_c(\mu,t)\cos(2\pi f_c t) - n_s(\mu,t)\sin(2\pi f_c t). \qquad (A_2)$$

(A_2) is called the quadrature representation of $n(\mu,t)$.

- Statistics of Zero-Mean WSS $n(\mu,t)$ and Associated Processes:

1. $E\{n(\mu,t)\} = 0 \Rightarrow$

   $E\{\widehat{n}(\mu,t)\} = E\{n_c(\mu,t)\} = E\{n_s(\mu,t)\} = E\{n_+(\mu,t)\} = E\{\widetilde{n}(\mu,t)\} = 0.$

2. $S_{\widehat{n}}(f) = S_n(f)$ (as shown before)
   $\Rightarrow R_{\widehat{n}}(\tau) = R_n(\tau)$
   Since $S_n(f)$ is real and even, so are $R_{\widehat{n}}(\tau)$ and $R_n(\tau)$.

3. $R_{n\widehat{n}}(\tau) = -\widehat{R}_n(\tau)$
   $R_{\widehat{n}n}(\tau) = \widehat{R}_n(\tau) = -R_{n\widehat{n}}(\tau) = -R_{\widehat{n}n}(-\tau).$
   Note: $R_{\widehat{n}n}(\tau)$ is odd and $R_{\widehat{n}n}(0) = 0$. Also, $\widehat{R}_n(0) = 0$.
   Proof: Because

$$\widehat{n}(\mu,t) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{n(\mu,t)}{t-\lambda}d\lambda = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{n(\mu,t-\lambda)}{\lambda}d\lambda$$

   we have

$$\begin{aligned}
R_{n\widehat{n}}(\tau) &= E\{n(\mu,t)\widehat{n}(\mu,t-\tau)\} \\
&= E\{\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{n(\mu,t)n(\mu,t-\tau-\lambda)}{\lambda}d\lambda\} \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{R_n(\lambda+\tau)}{\lambda}d\lambda \\
(\lambda' = \lambda+\tau) &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{R_n(\lambda')}{\lambda'-\tau}d\lambda' \\
&= -\int_{-\infty}^{\infty}\frac{R_n(\lambda')}{\pi(\tau-\lambda')}d\lambda' \\
&= -\widehat{R}_n(\tau).
\end{aligned}$$

   Similarly, $R_{\widehat{n}n}(\tau) = \widehat{R}_n(\tau)$.

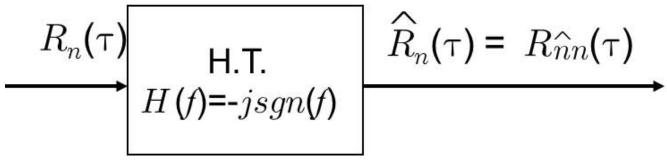


Figure 33:

Next, we show that

$$
\begin{aligned}
R_{\widehat{n}n}(\tau) &= E\{\widehat{n}(\mu, t+\tau)n(\mu, t)\} \\
&= E\{\int_{-\infty}^{\infty} \frac{1}{\pi x} n(\mu, t+\tau-x)dx \cdot n(\mu, t)\} \\
&= \int_{-\infty}^{\infty} \frac{1}{\pi x} R_n(\tau - x)dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\pi x} R_n(x - \tau)dx \\
&\quad \text{(since } R_n(\tau) \text{ is even)} \\
&= E\{\int_{-\infty}^{\infty} \frac{1}{\pi x} n(\mu, t+x-\tau)dx \cdot n(\mu, t)\} \\
&= -E\{\int_{-\infty}^{\infty} \frac{1}{\pi y} n(\mu, t-\tau-y)dy \cdot n(\mu, t)\} \\
\text{(letting } y &= -x) \\
&= -E\{\widehat{n}(\mu, t-\tau)n(\mu, t)\} \\
&= -R_{\widehat{n}n}(-\tau).
\end{aligned}
$$

Thus, $R_{\widehat{n}n}(\tau)$ is odd, and so is $\widehat{R}_n(\tau)$. Q.E.D.

4. If $n(\mu, t)$ is WSS, then $n_c(\mu, t)$ and $n_s(\mu, t)$ are also jointly WSS with

$$
\begin{aligned}
R_{n_c}(\tau) &= R_{n_s}(\tau) = R_n(\tau)\cos(2\pi f_c \tau) + \widehat{R}_n(\tau)\sin(2\pi f_c \tau) \\
R_{n_c n_s}(\tau) &= -R_{n_s n_c}(\tau) = R_n(\tau)\sin(2\pi f_c \tau) - \widehat{R}_n(\tau)\cos(2\pi f_c \tau).
\end{aligned}
$$

(You prove it.)

5. Assume that $n(\mu, t)$ has zero mean and is bandlimited to $|f + f_c| \leq \frac{W}{2}$ and $|f - f_c| \leq \frac{W}{2}$. Then,

$$
S_{n_c}(f) = S_{n_s}(f) = \begin{cases} S_n(f - f_c) + S_n(f + f_c), & |f| \leq \frac{W}{2} \\ 0, & \text{otherwise} \end{cases}.
$$

(You prove it.)

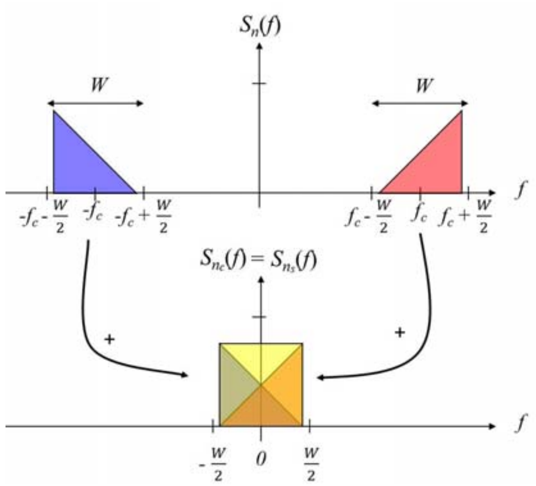


Figure 34:

Therefore, $Var\{n_c(\mu, t)\} = Var\{n_s(\mu, t)\} = Var\{n(\mu, t)\}$.
Proof: Now,

$$\begin{aligned}
Var\{n(\mu, t)\} &= \int_{-\infty}^{\infty} S_n(f)df \\
&= \int_{|f-f_c|\leq\frac{W}{2}} S_n(f)df + \int_{|f+f_c|\leq\frac{W}{2}} S_n(f)df \\
&= \int_{|f|\leq\frac{W}{2}} [S_n(f+f_c) + S_n(f-f_c)]df \\
&= Var\{n_c(\mu, t)\} \\
&= Var\{n_s(\mu, t)\}. \qquad \text{Q.E.D.}
\end{aligned}$$

6.

$$S_{n_c n_s}(f) = -S_{n_s n_c}(f) = \begin{cases} j[S_n(f+f_c) - S_n(f-f_c)], & |f| \leq \frac{W}{2} \\ 0, & \text{otherwise} \end{cases}.$$

(You prove it.)

* Corollary: If $S_n(f)$ is also locally symmetric about f_c, then the processes $n_c(\mu, t)$ and $n_s(\mu, t)$ are orthogonal for all τ, i.e., $E\{n_c(\mu, t+\tau)n_s(\mu, t)\} = R_{n_c n_s}(\tau) = 0$, $\forall t, \tau$.
Proof: Since $S_{n_c n_s}(f) = 0$, $R_{n_c n_s}(\tau) = 0$. Q.E.D.

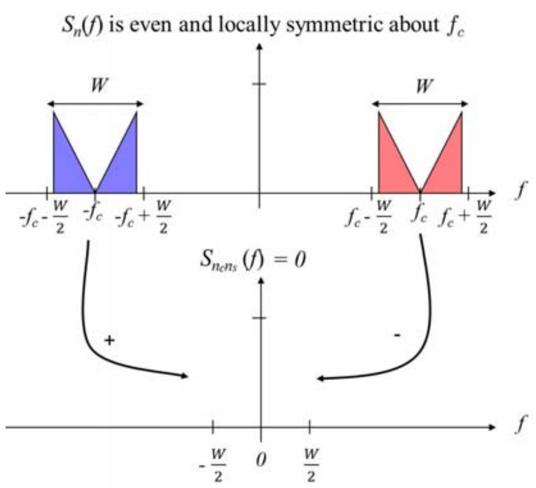


Figure 35:

- Note: If WSS bandpass real-valued $n(\mu, t)$ is also Gaussian with zero mean, then the following properties hold.

  a. $n(\mu, t)$ and $\widehat{n}(\mu, t)$ are jointly Gaussian random processes with zero mean. Because $R_{\widehat{n}n}(0) = 0$, $n(\mu, t)$ and $\widehat{n}(\mu, t)$ are independent Gaussian random variables with zero mean for a fixed t.

  b. $n_c(\mu, t)$ and $n_s(\mu, t)$ are jointly Gaussian random processes with zero mean. Because $R_{n_cn_s}(0) = -\widehat{R}_n(0) = 0$, $n_c(\mu, t)$ and $n_s(\mu, t)$ are independent Gaussian random variables with zero mean for a fixed t.

     Also, if the local symmetry of $S_n(f)$ also holds, then $n_c(\mu, t)$ and $n_s(\mu, t)$ are independent Gaussian random processes.

5.5 More on Statistics of Polar and Quadrature Representations of $n(\mu, t)$

- Polar Representation of $n(\mu, t)$:

  Express $\widetilde{n}(\mu, t)$ in its polar and quadrature representations

$$\begin{aligned} \widetilde{n}(\mu, t) &= v(\mu, t)e^{j\phi(\mu,t)} \\ &= n_c(\mu, t) + jn_s(\mu, t) \end{aligned}$$

where $v(\mu,t) = |\widetilde{n}(\mu,t)|$ and $\phi(\mu,t) = \angle\widetilde{n}(\mu,t)$. This shows that

$$\begin{aligned} n_c(\mu,t) &= v(\mu,t)\cos(\phi(\mu,t)) \\ n_s(\mu,t) &= v(\mu,t)\sin(\phi(\mu,t)). \end{aligned}$$

Assuming that the joint statistic of $n_c(\mu,t)$ and $n_s(\mu,t)$ is known, we discuss below the joint statistic of $v(\mu,t)$ and $\phi(\mu,t)$:

CASE I: **Zero-Mean Noise**

Let $n(\mu,t)$ be zero-mean WSS Gaussian noise with PSD $S_n(f)$.

- Lemma:

$$\begin{aligned} C_{\widetilde{n}\widetilde{n}^*}(\tau) &\triangleq E\{\widetilde{n}(\mu,t+\tau)\widetilde{n}^*(\mu,t)\} \\ &= R_{\widetilde{n}}(\tau) \\ &= R_{n_+}(\tau)e^{-j2\pi f_c\tau} \end{aligned}$$

and

$$\begin{aligned} R_{\widetilde{n}\widetilde{n}}(\tau,t) &= E\{\widetilde{n}(\mu,\tau)\widetilde{n}(\mu,t)\} \\ &= 0 \qquad \text{for all } t \text{ and } \tau. \end{aligned}$$

Proof: First note that

$$\begin{aligned} R_{\widetilde{n}}(\tau) &\triangleq E\{\widetilde{n}(\mu,t+\tau)\widetilde{n}^*(\mu,t)\} \\ &= E\{n_+(\mu,t+\tau)e^{-j2\pi f_c(t+\tau)}n_+^*(\mu,t)e^{j2\pi f_c t}\} \\ &= R_{n_+}(\tau)e^{-j2\pi f_c\tau} \end{aligned}$$

and also that

$$\begin{aligned} R_{\widetilde{n}\widetilde{n}}(\tau,t) &= E\{\widetilde{n}(\mu,\tau)\widetilde{n}(\mu,t)\} \\ &= E\{n_+(\mu,\tau)e^{-j2\pi f_c\tau}n_+(\mu,t)e^{-j2\pi f_c t}\} \\ &= R_{n_+n_+}(\tau,t)e^{-j2\pi f_c(\tau+t)}. \end{aligned}$$

But,

$$\begin{aligned} R_{n_+n_+}(\tau,t) &= E\{n_+(\mu,\tau)n_+(\mu,t)\} \\ &= E\{\int_{-\infty}^{\infty} h(a)n(\mu,\tau-a)da \int_{-\infty}^{\infty} h(b)n(\mu,t-b)db\} \end{aligned}$$

where $h(t) = \mathcal{F}^{-1}\{2U(f)\}$. Thus,

$$\begin{aligned}
R_{n_+n_+}(\tau,t) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(a)h(b)R_n(\tau-a-t+b)dadb \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(a)h(b)\int_{-\infty}^{\infty} S_n(f)e^{j2\pi f(\tau-a-t+b)}df dadb \\
&= \int_{-\infty}^{\infty} S_n(f)e^{j2\pi f(\tau-t)}[\int_{-\infty}^{\infty} h(a)e^{-j2\pi fa}da][\int_{-\infty}^{\infty} h(b)e^{j2\pi fb}db]df \\
&= \int_{-\infty}^{\infty} S_n(f)e^{j2\pi f(\tau-t)}\underbrace{(2U(f))(2U(-f))}_{0}df \\
&= 0.
\end{aligned}$$

This implies that $R_{\widetilde{n}\widetilde{n}}(\tau,t) = 0$, $\forall \tau, t$. Q.E.D.

- For a specific time t, $\widetilde{n}(\mu,t)$ is a complex-valued Gaussian random variable and has the first-order pdf

$$\begin{aligned}
f_{\widetilde{n}}(\widetilde{n};t) &= f_{n_c,n_s}(n_c,n_s) \\
&= \frac{1}{\pi s}e^{-\frac{|\widetilde{n}|^2}{s}}
\end{aligned}$$

where $|\widetilde{n}|^2 = n_c^2 + n_s^2$,

$$\begin{aligned}
s &= R_{\widetilde{n}\widetilde{n}^*}(0) = E\{|\widetilde{n}(\mu,t)|^2\} \\
&= E\{|n_+(\mu,t)|^2\} \\
&= 2E\{|n(\mu,t)|^2\} \\
&\triangleq 2\sigma_n^2
\end{aligned}$$

(which is derived since (1) $n_+(\mu,t) = n(\mu,t) + j\widehat{n}(\mu,t)$, (2) $E\{|n_+(\mu,t)|^2\} = E\{|n(\mu,t)|^2\} + E\{|\widehat{n}(\mu,t)|^2\} = 2E\{|n(\mu,t)|^2\}$) and

$$\sigma_n^2 = \overbrace{E\{|n(\mu,t)|^2\}}^{R_n(0)} = Var\{n(\mu,t)\} = \int_{-\infty}^{\infty} S_n(f)df$$

is the total noise variance or power.) Therefore,

$$\begin{aligned}
f_{n_c,n_s}(n_c,n_s) &= \frac{1}{2\pi\sigma_n^2}e^{-\frac{n_c^2+n_s^2}{2\sigma_n^2}} \\
&= f_{n_c}(n_c)f_{n_s}(n_s)
\end{aligned}$$

i.e., $n_c(\mu,t)$ and $n_s(\mu,t)$ are independent and identically distributed Gaussian random variables with mean zero and variance σ_n^2 for a fixed

t.
Now, using the transformation

$$\begin{aligned} n_c &= v\cos\phi \\ n_s &= v\sin\phi \end{aligned}$$

we can derive by Jacobian approach

$$f_{v,\phi}(v,\phi) = f_{n_c,n_s}(n_c,n_s)|J|$$

with

$$J = \begin{vmatrix} \frac{\partial n_c}{\partial v} & \frac{\partial n_c}{\partial \phi} \\ \frac{\partial n_s}{\partial v} & \frac{\partial n_s}{\partial \phi} \end{vmatrix} = v.$$

Thus,

$$\begin{aligned} f_{v,\phi}(v,\phi) &= \frac{v}{2\pi\sigma_n^2} e^{-\frac{v^2}{2\sigma_n^2}} U(v) \qquad 0 \le \phi < 2\pi \\ &= \frac{1}{2\pi} \cdot \frac{v}{\sigma_n^2} e^{-\frac{v^2}{2\sigma_n^2}} U(v) \\ &= f_\phi(\phi) \cdot f_v(v). \end{aligned}$$

For fixed t, $\phi(\mu,t)$ is uniform in $[0,2\pi)$, $v(\mu,t)$ is Rayleigh, and they are independent.
Note: If $n(\mu,t)$ is narrowband so that we can write

$$S_n(f) = S_L(f-f_c) + S_L(-f-f_c)$$

where $S_L(f) = 0$ for $|f| \ge B$. Then, $\sigma_n^2 = 2\int_{|f|\le B} S_L(f)df$ and

$$\begin{aligned} S_{n_+}(f) &= \begin{cases} 4S_L(f-f_c), & f > 0 \\ 0, & f < 0 \end{cases} \\ S_{\tilde{n}}(f) &= 4S_L(f) = S_{n_+}(f+f_c). \end{aligned}$$

Proof: Now,

$$\begin{aligned} R_{n_+}(\tau) &\triangleq E\{n_+(\mu,t+\tau)n_+^*(\mu,t)\} \\ &= E\{[n(\mu,t+\tau) + j\widehat{n}(\mu,t+\tau)][n(\mu,t) - j\widehat{n}(\mu,t)]\} \\ &= R_n(\tau) + R_{\widehat{n}}(\tau) - jR_{n\widehat{n}}(\tau) + jR_{\widehat{n}n}(\tau). \end{aligned}$$

Since $R_{\widehat{n}n}(\tau) = \widehat{R}_n(\tau) = -R_{n\widehat{n}}(\tau)$ and $R_n(\tau) = R_{\widehat{n}}(\tau)$,

$$R_{n_+}(\tau) = 2(R_n(\tau) + j\widehat{R}_n(\tau))$$

in which $R_n(\tau) + j\widehat{R}_n(\tau)$ represents the pre-envelope of $R_n(\tau)$, i.e., $\mathcal{F}^{-1}\{2U(f)S_n(f)\}$. Thus, $S_{n_+}(f) = 4U(f)S_n(f)$. Next, because $R_{\widetilde{n}}(\tau) = R_{n_+}(\tau)e^{-j2\pi f_c\tau}$, we have

$$S_{\widetilde{n}}(f) = S_{n_+}(f + f_c).$$

Q.E.D.

- Joint Statistics of $v(\mu, t_1) = v_1(\mu)$, $v(\mu, t_2) = v_2(\mu)$, $\phi(\mu, t_1) = \phi_1(\mu)$, and $\phi(\mu, t_2) = \phi_2(\mu)$:
  Let $\widetilde{n}_1(\mu) = \widetilde{n}(\mu, t_1)$ and $\widetilde{n}_2(\mu) = \widetilde{n}(\mu, t_2)$. Note that $\widetilde{n}_1(\mu)$ and $\widetilde{n}_2(\mu)$ are jointly Gaussian random variables with

$$f_{\widetilde{n}_1,\widetilde{n}_2}(\widetilde{n}_1, \widetilde{n}_2) = \frac{1}{\pi^2|C_{\underline{z}\underline{z}^*}|}\exp\{-\begin{bmatrix}\widetilde{n}_1\\ \widetilde{n}_2\end{bmatrix}^{*T} C_{\underline{z}\underline{z}^*}^{-1}\begin{bmatrix}\widetilde{n}_1\\ \widetilde{n}_2\end{bmatrix}\}$$

where $\underline{z} = \begin{bmatrix}\widetilde{n}_1\\ \widetilde{n}_2\end{bmatrix}$ and $C_{\underline{z}\underline{z}^*} = E\{\begin{bmatrix}\widetilde{n}_1\\ \widetilde{n}_2\end{bmatrix}\begin{bmatrix}\widetilde{n}_1 & \widetilde{n}_2\end{bmatrix}^*\} = \begin{bmatrix} s & r\\ r^* & s\end{bmatrix}$ with

$$\begin{aligned}
s &= E\{|\widetilde{n}(\mu, t)|^2\} = 2\sigma_n^2\\
r &= E\{\widetilde{n}_1(\mu)\widetilde{n}_2^*(\mu)\}\\
&= R_{\widetilde{n}}(t_1 - t_2)\\
&= 4\int_{-\infty}^{\infty} S_L(f)e^{j2\pi f(t_1-t_2)}df\\
&= 2(R_{n_c}(t_1 - t_2) - jR_{n_cn_s}(t_1 - t_2)).
\end{aligned}$$

Note that

$$\begin{aligned}
R_{\widetilde{n}}(\tau) &= E\{\widetilde{n}(\mu, t+\tau)\widetilde{n}^*(\mu, t)\}\\
&= E\{[n_c(\mu, t+\tau) + jn_s(\mu, t+\tau)][n_c(\mu, t) - jn_s(\mu, t)]\}\\
&= R_{n_c}(\tau) + R_{n_s}(\tau) - jR_{n_cn_s}(\tau) + jR_{n_sn_c}(\tau)\\
&= 2(R_{n_c}(\tau) - jR_{n_cn_s}(\tau)).
\end{aligned}$$

Let $v_i = |\widetilde{n}_i|$, $i = 1, 2$, and $\phi_i = \angle\widetilde{n}_i, i = 1, 2$. Then, $|J| = |J_1||J_2| = v_1v_2$ and

$$f_{v_1,v_2,\phi_1,\phi_2}(v_1, v_2, \phi_1, \phi_2) = v_1v_2 f_{\widetilde{n}_1,\widetilde{n}_2}(v_1e^{j\phi_1}, v_2e^{j\phi_2})$$

for $v_1, v_2 \geq 0$ and $0 \leq \phi_1, \phi_2 < 2\pi$. Note that

$$C_{\underline{z}\underline{z}^*}^{-1} = \frac{1}{s^2 - |r|^2}\begin{bmatrix} s & -r\\ -r^* & s\end{bmatrix} \text{ and } |C_{\underline{z}\underline{z}^*}| = s^2 - |r|^2.$$

Thus,

$$
\begin{aligned}
& f_{v_1,v_2,\phi_1,\phi_2}(v_1,v_2,\phi_1,\phi_2) \\
= \;& v_1 v_2 \frac{\exp\left\{-\begin{bmatrix} v_1 e^{-j\phi_1} \\ v_2 e^{-j\phi_2} \end{bmatrix}^T C_{\underline{zz}^*}^{-1} \begin{bmatrix} v_1 e^{j\phi_1} \\ v_2 e^{j\phi_2} \end{bmatrix}\right\}}{\pi^2(s^2-|r|^2)} \\
= \;& \frac{v_1 v_2}{\pi^2(s^2-|r|^2)} \exp\left\{-\frac{s(v_1^2+v_2^2)-2v_1v_2|r|\cos(\phi_2-\phi_1+\theta_r)}{(s^2-|r|^2)}\right\}
\end{aligned}
$$

where $\theta_r = \angle r$, $v_1, v_2 \geq 0$, and $0 \leq \phi_1, \phi_2 < 2\pi$.

First, look at $f_{v_1,v_2}(v_1,v_2)$:

$$f_{v_1,v_2}(v_1,v_2) = \int_0^{2\pi}\int_0^{2\pi} f_{v_1,v_2,\phi_1,\phi_2}(v_1,v_2,\phi_1,\phi_2)d\phi_1 d\phi_2.$$

Now, letting $\phi_1^{'} = \phi_1 - \phi_2$ and $\phi_2^{'} = \phi_2$, we have

$$f_{v_1,v_2}(v_1,v_2) = \frac{4v_1v_2}{s^2-|r|^2}\exp\left\{-\frac{s(v_1^2+v_2^2)}{s^2-|r|^2}\right\} I_0\left(\frac{2v_1v_2|r|}{s^2-|r|^2}\right) \qquad v_1, v_2 \geq 0$$

where

$$
\begin{aligned}
I_0(z) &= \frac{1}{2\pi}\int_0^{2\pi} e^{z\cos(\theta+\phi)}d\theta \\
&= \frac{1}{\pi}\int_0^{\pi} e^{z\cos(\theta+\phi)}d\theta \qquad \text{for any } \phi
\end{aligned}
$$

is the zeroth-order modified Bessel-function of the first kind.

Next, look at $f_{\phi_1,\phi_2}(\phi_1,\phi_2)$:

$$f_{\phi_1,\phi_2}(\phi_1,\phi_2) = \int_0^{\infty}\int_0^{\infty} f_{v_1,v_2,\phi_1,\phi_2}(v_1,v_2,\phi_1,\phi_2)dv_1 dv_2.$$

Now, make the transformation

$$
\begin{aligned}
\left\{\begin{array}{l} ye^{2z} = \frac{2s}{s^2-|r|^2}v_1^2 \\ ye^{-2z} = \frac{2s}{s^2-|r|^2}v_2^2 \end{array}\right. \quad &\Rightarrow \quad \left\{\begin{array}{l} y = \frac{2s}{s^2-|r|^2}v_1v_2 \\ z = \frac{1}{2}(\ln v_1 - \ln v_2) \end{array}\right. \\
&\Rightarrow \quad \left\{\begin{array}{l} v_1 = y^{1/2}e^{z}a^{1/2} \\ v_2 = y^{1/2}e^{-z}a^{1/2} \end{array}\right.
\end{aligned}
$$

with $a = (s^2 - |r|^2)/(2s)$. It is obvious that $y \geq 0$, z is real, and the Jacobian for the transformation is given by the absolute of

$$\frac{\partial(v_1, v_2)}{\partial(y, z)} = \det \left\{ a^{\frac{1}{2}} \begin{bmatrix} \frac{1}{2} y^{-1/2} e^{z} & y^{1/2} e^{z} \\ \frac{1}{2} y^{-1/2} e^{-z} & -y^{1/2} e^{-z} \end{bmatrix} \right\} = -a$$

($dv_1 dv_2 = a dy dz$). Thus,

$$\begin{aligned} f_{\phi_1,\phi_2}(\phi_1, \phi_2) &= \frac{1 - \frac{|r|^2}{s^2}}{(2\pi)^2} \int_0^\infty y e^{\beta y} \int_{-\infty}^{\infty} e^{-y\cosh(2z)} dz dy \\ &= \frac{1 - \frac{|r|^2}{s^2}}{(2\pi)^2} \int_0^\infty y e^{\beta y} \cdot 2 \int_0^{\infty} e^{-y\cosh(2z)} dz dy \quad (*) \end{aligned}$$

where $\beta \triangleq \frac{|r|}{s}\cos(\phi_2 - \phi_1 + \theta_r)$ and $\cosh x = \frac{1}{2}(e^x + e^{-x})$. Note that $(*)$ results from the fact that $\cosh x$ is even. Since

$$K_\nu(z) = \int_0^\infty e^{-z\cosh(t)} \cosh(\nu t) dt$$

is the νth-order modified Bessel function of the second kind, if we set $\nu = 0$ and $t = 2z$ in $(*)$, we have

$$f_{\phi_1,\phi_2}(\phi_1, \phi_2) = \frac{1 - \frac{|r|^2}{s^2}}{(2\pi)^2} \int_0^\infty y e^{\beta y} K_0(y) dy.$$

Furthermore, since

$$\int_0^\infty e^{-at} K_0(t) dt = \frac{\cos^{-1} a}{\sqrt{1 - a^2}},$$

differentiating this integral with respect to a yields

$$\int_0^\infty t e^{-at} K_0(t) dt = \frac{1}{1 - a^2} - \frac{a \cos^{-1} a}{(1 - a^2)^{3/2}}.$$

Now, letting $\beta = -a$, we conclude that

$$f_{\phi_1,\phi_2}(\phi_1, \phi_2) = \frac{1 - \frac{|r|^2}{s^2}}{(2\pi)^2} \left[\frac{1}{1 - \beta^2} + \frac{\beta(\pi - \cos^{-1}\beta)}{(1 - \beta^2)^{\frac{3}{2}}} \right] \qquad 0 \leq \phi_1, \phi_2 < 2\pi$$

where

$$\beta = \frac{|r|}{s}\cos(\phi_1 - \phi_2 + \theta_r).$$

Note: Since $r = |r|e^{j\theta_r} = 2[R_{n_c}(\tau) - jR_{n_c n_s}(\tau)]$,

$$\begin{aligned} |r|\cos\theta_r &= 2R_{n_c}(\tau) \\ |r|\sin\theta_r &= -2R_{n_c n_s}(\tau). \end{aligned}$$

Also, because $s = 2\sigma_n^2 = 2R_n(0)$, we have

$$\beta = \frac{R_{n_c(\tau)}}{R_n(0)}\cos(\phi_2 - \phi_1) + \frac{R_{n_c n_s}(\tau)}{R_n(0)}\sin(\phi_2 - \phi_1).$$

CASE II: **Sinusoidal Signal Plus Noise at a Fixed Time Instant**

Let

$$x(\mu, t) = A\cos(\omega_c t + \psi(\mu)) + n(\mu, t)$$

where A is a deterministic amplitude, $\psi(\mu)$ is a uniform phase in $[0, 2\pi)$, and $n(\mu, t)$ is a bandpass Gaussian noise, as defined in CASE I and independent of $\psi(\mu)$. Note that the average signal power for $A\cos(\omega_c t + \psi(\mu))$ is given by $\frac{A^2}{2}$.

We want to derive the statistics for $v(\mu, t)$ and $\phi(\mu, t)$ for which $x(\mu, t)$ has been defined by

$$x(\mu, t) = \mathrm{Re}\{v(\mu, t)e^{j(\omega_c t + \phi(\mu, t))}\}.$$

If we use a quadrature representation of $n(\mu, t)$

$$n(\mu, t) = n_c(\mu, t)\cos(\omega_c t) - n_s(\mu, t)\sin(\omega_c t),$$

then

$$x(\mu, t) = x_c(\mu, t)\cos(\omega_c t) - x_s(\mu, t)\sin(\omega_c t)$$

with

$$\begin{aligned} x_c(\mu, t) &= A\cos\psi(\mu) + n_c(\mu, t) \\ x_s(\mu, t) &= A\sin\psi(\mu) + n_s(\mu, t). \end{aligned}$$

Thus,

$$\begin{aligned} v(\mu, t) &= \sqrt{x_c^2(\mu, t) + x_s^2(\mu, t)} \\ \phi(\mu, t) &= \tan^{-1}\frac{x_s(\mu, t)}{x_c(\mu, t)}. \end{aligned}$$

Now, let us fix the time point $t = t_0$, so that

$$\begin{aligned} x_c(\mu, t_0) &= x_c(\mu),\ x_s(\mu, t_0) = x_s(\mu) \\ v(\mu, t_0) &= v(\mu),\ \phi(\mu, t_0) = \phi(\mu). \end{aligned}$$

Recall that $n_c(\mu, t_0)$ and $n_s(\mu, t_0)$ are independent because (1)

$$\begin{aligned} R_{n_c n_s}(0) &= -R_{n_s n_c}(0) = R_n(0)\sin(0) - \widehat{R}_n(0)\cos(0) \\ &= -\widehat{R}_n(0) = -R_{\widehat{n}n}(0) = 0 \end{aligned}$$

(note: $R_{\widehat{n}n}(\tau)$ is odd), (2) $E\{n_c(\mu, t_0)\} = E\{n_s(\mu, t_0)\} = 0$, and (3) $n_c(\mu, t_0)$ and $n_s(\mu, t_0)$ are jointly Gaussian. Thus, for $t = t_0$, the random variables $x_c(\mu)$ and $x_s(\mu)$ are conditionally independent when $\psi(\mu)$ is given, and

$$\begin{aligned} f_{x_c, x_s, \psi}(x_c, x_s, \psi) &= f_{x_c, x_s|\psi}(x_c, x_s|\psi) f_\psi(\psi) \\ &= (\frac{1}{2\pi}) \cdot (\frac{1}{2\pi\sigma_n^2}) \exp\left\{ -\frac{(x_c - A\cos\psi)^2}{2\sigma_n^2} - \frac{(x_s - A\sin\psi)^2}{2\sigma_n^2} \right\} \\ &= \frac{1}{4\pi^2\sigma_n^2} \exp\left\{ -\frac{1}{2\sigma_n^2} \left[x_c^2 + x_s^2 + A^2 - 2A(x_c\cos\psi + x_s\sin\psi) \right] \right\}. \end{aligned}$$

Changing from $\{x_c, x_s, \psi\}$ to $\{v, \phi, \psi\}$, we have ($f_{x_c, x_s, \psi}(x_c, x_s, \psi) v = f_{v,\phi,\psi}(v, \phi, \psi)$)

$$f_{v,\phi,\psi}(v, \phi, \psi) = \frac{v}{4\pi^2\sigma_n^2} \exp\left\{ -\frac{A^2 + v^2 - 2Av\cos(\phi - \psi)}{2\sigma_n^2} \right\} \qquad v \ge 0 \text{ and } 0 \le \phi, \psi < 2\pi.$$

Now, integrating over ϕ and ψ, we have

$$\begin{aligned} f_v(v) &= \int_0^{2\pi} \int_0^{2\pi} f_{v,\phi,\psi}(v, \phi, \psi) d\phi d\psi \\ (\theta \triangleq \phi - \psi) &= \frac{v}{\sigma_n^2} \exp\left\{ -\frac{v^2 + A^2}{2\sigma_n^2} \right\} \cdot (\frac{1}{2\pi}) \int_0^{2\pi} d\psi \\ &\quad \cdot \underbrace{\int_{-\psi}^{2\pi - \psi} \frac{1}{2\pi} \exp\left\{ \frac{Av}{\sigma_n^2} \cos\theta \right\} d\theta.}_{I_0(\frac{Av}{\sigma_n^2})} \end{aligned}$$

So,

$$f_v(v) = \frac{v}{\sigma_n^2} \exp\left\{ -\frac{v^2 + A^2}{2\sigma_n^2} \right\} I_0(\frac{Av}{\sigma_n^2}) \qquad v \ge 0$$

which is the Rician density.

Next, integrating over ψ only, we have

$$\begin{aligned} f_{v,\phi}(v,\phi) &= \int_0^{2\pi} f_{v,\phi,\psi}(v,\phi,\psi)d\psi \\ &= \underbrace{\frac{1}{2\pi}}_{f_\phi(\phi)} f_v(v) \qquad 0 \leq \phi < 2\pi \text{ and } v \geq 0. \end{aligned}$$

Thus, $v(\mu,t_0)$ and $\phi(\mu,t_0)$ are independent random variables with $\phi(\mu,t_0)$ uniform in $[0,2\pi)$ and $v(\mu,t_0)$ Rician distributed.
Notes:

1. As $A \to 0$ (i.e., average signal power $\to 0$), $f_v(v) \to$ *Rayleigh pdf*.
2. The above form can be expressed as a function of the signal-to-noise power ratio (SNR)

$$SNR = \frac{A^2/2}{\sigma_n^2}.$$

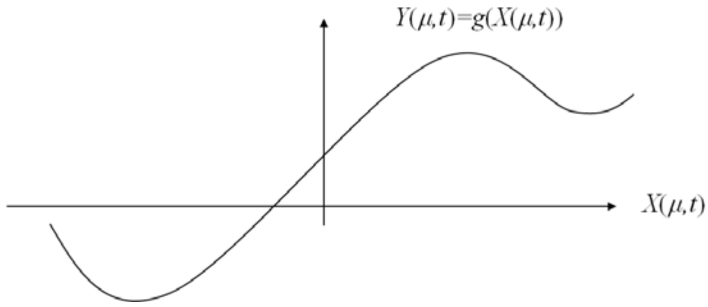
3. It can be shown that

$$\begin{aligned} f_{\phi,\psi}(\phi,\psi) &= \frac{1}{4\pi^2}\exp\left\{-\frac{A^2}{2\sigma_n^2}\right\} + \frac{A\cos(\phi-\psi)}{(2\pi)^{3/2}\sigma_n^2}\exp\left\{-\frac{A^2\sin^2(\phi-\psi)}{2\sigma_n^2}\right\} \\ &\quad \cdot \left[1 - Q\left(\frac{A\cos(\phi-\psi)}{\sigma_n}\right)\right] \end{aligned}$$

with $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy$.

5.6 Memoryless Nonlinear System

- Recall that a system $Y(\mu,t) = T[X(\mu,t)]$ is called memoryless iff the output $Y(\mu,t)$ is a function of the input $X(\mu,t)$ only for the same time instant. For example, $Y(\mu,t) = X(\mu,t-1)$ has memory and $Y(\mu,t) = X^2(\mu,t)$ is memoryless. Note also that $Y(\mu,t) = X^2(\mu,t)$ is nonlinear. We are here interested in memoryless nonlinear systems whose input and output are both real-valued and can be characterized by $Y(\mu,t) = g(X(\mu,t))$ where $g(x)$ is a function of x.



- For memoryless nonlinear systems,

  1. if $X(\mu, t)$ is SSS, so is $Y(\mu, t)$;
  2. if $X(\mu, t)$ is stationary of order N, so is $Y(\mu, t)$;
  3. if $X(\mu, t)$ is WSS, $Y(\mu, t)$ may not be stationary in any sense.

  Therefore, the second-moment description of $X(\mu, t)$ is not sufficient for the second-moment description of $Y(\mu, t)$ in memoryless nonlinear systems.

- Examples of Nonlinearity:

  1. Full-Wave Square Law: $g(x) = ax^2$

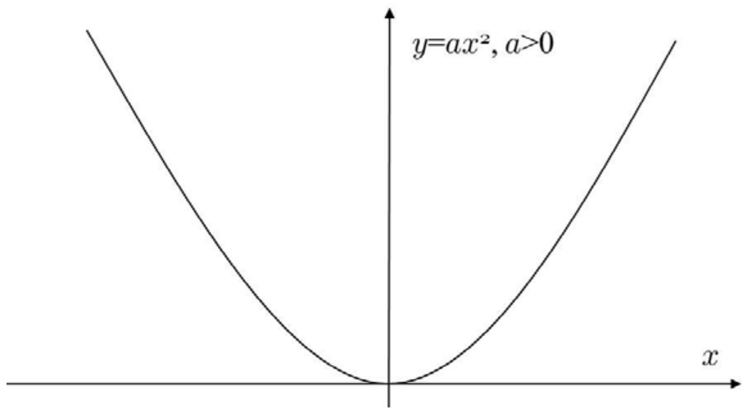


Figure 36:

which is memoryless.

2. Half-Wave Linear Law: $g(x) = ax \cdot u(x)$ with $u(x)$ a unit step function

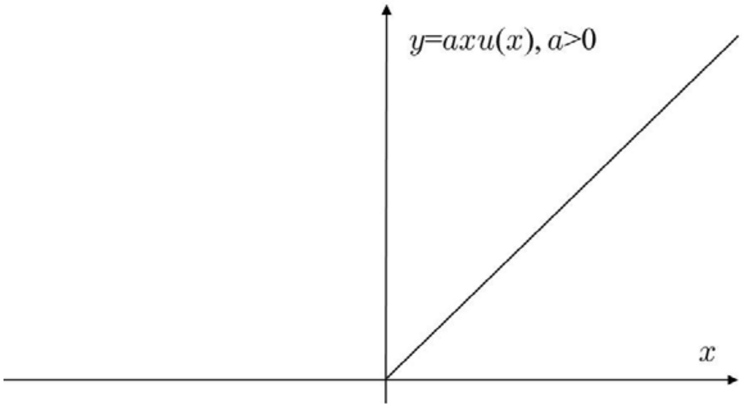


Figure 37:

which is memoryless.

3. Hysteresis Law:

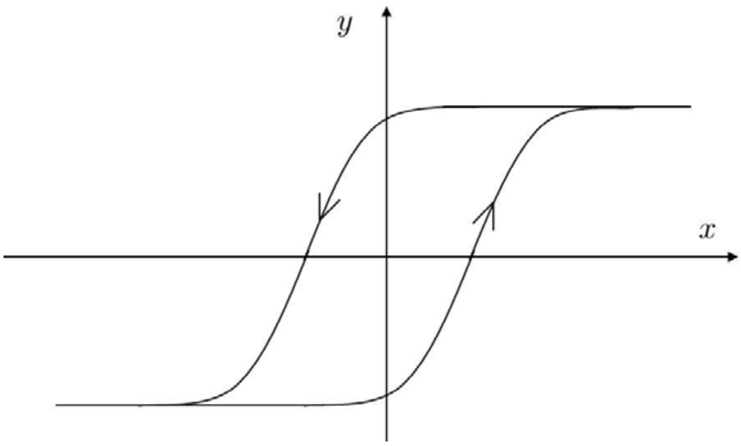


Figure 38:

which is not memoryless.

4. Hard Limiter:

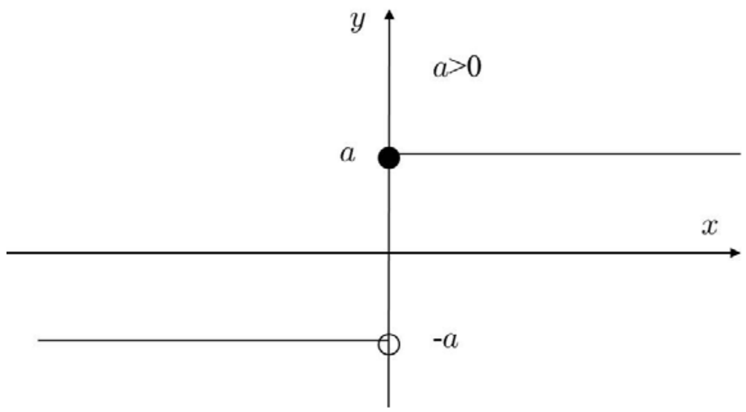
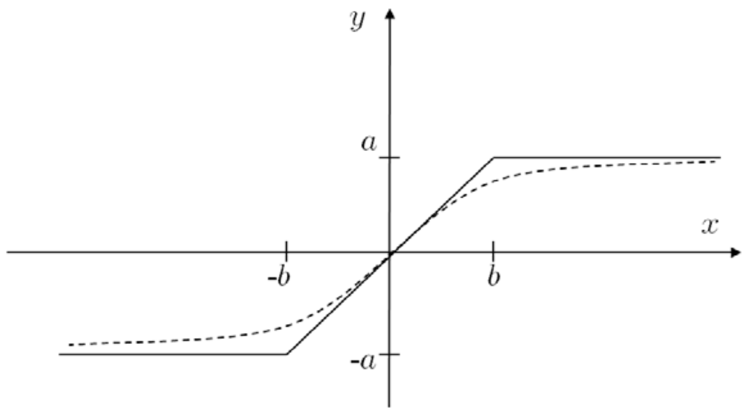


Figure 39:

which is memoryless.

5. Soft Limiter:

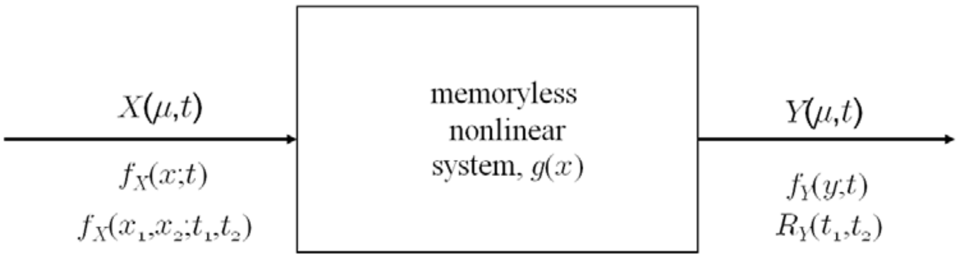


which is memoryless.

- Most of the nonlinear analytical methods concentrate on the second-order statistical description of input and output processes, namely autocorrelations and power spectrums. One famous approach is the direct method which deals with probability density functions and is good for use if $X(\mu, t)$ is Gaussian.

5.6.1 Direct Method

- Consider the memoryless nonlinear system $Y(\mu,t)=g(X(\mu,t))$ where the first-order and second-order densities of input process $X(\mu,t)$, namely $f_X(x;t)$ and $f_X(x_1,x_2;t_1,t_2)$, are given.



Now, the following statistics of output $Y(\mu,t)$ can be obtained

$$\begin{aligned}
f_Y(y;t) &= \sum_{\substack{\text{all roots}\\ \text{of } y=g(x_i)}} f_X(x_i;t)|J(x_i)| \\
E\{Y^n(\mu,t)\} &= \int_{-\infty}^{\infty} g^n(x) f_X(x;t)dx \\
R_Y(t_1,t_2) &= E\{g(X(\mu,t_1))g(X(\mu,t_2))\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1)g(x_2)f_X(x_1,x_2;t_1,t_2)dx_1dx_2.
\end{aligned}$$

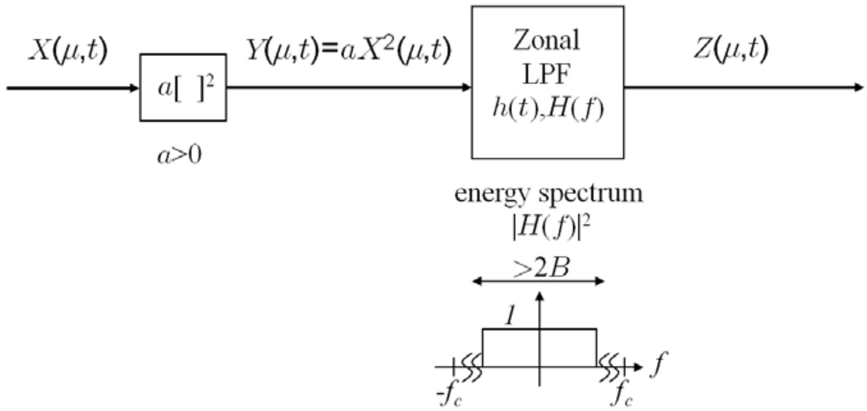
- Consider some examples below.

1. Full-Wave Square Law Device: $Y(\mu,t)=aX^2(\mu,t)$ with $a>0$.

$$f_Y(y;t)=\frac{1}{2\sqrt{ay}}[f_X(\sqrt{y/a};t)+f_X(-\sqrt{y/a};t)], \qquad y\geq 0$$

and $f_Y(y;t)=0$ for $y<0$.

2. Square-Law Detector:



The zonal LPF allows a spectral band to pass undistorted and everything else is filtered perfectly.

CASE 1: Assume further that $X(\mu,t)$ is a narrowband WSS Gaussian noise process $n(\mu,t)$ with zero mean and autocorrelation $R_X(\tau)$. Thus, $X(\mu,t)$ can be expressed in the polar form

$$X(\mu,t) = n(\mu,t) = v(\mu,t)\cos(2\pi f_c t + \theta(\mu,t)).$$

For fixed t, $v(\mu,t)$ and $\theta(\mu,t)$ are independent random variables where $v(\mu,t)$ is Rayleigh distributed with

$$f_v(v;t) = \frac{v}{R_n(0)}\exp\{-\frac{v^2}{2R_n(0)}\} \qquad v \geq 0$$

and $f_v(v;t) = 0$ for $v < 0$, and $\theta(\mu,t)$ is uniform in $[0,2\pi)$. Here, $\sigma_n^2 = R_n(0) = E\{n^2(\mu,t)\}$ and $E\{v^2(\mu,t)\} = 2R_n(0)$. When the bandwidth of the zonal LPF is much smaller than f_c, we have

$$\begin{aligned} Y(\mu,t) &= av^2(\mu,t)\cos^2(2\pi f_c t + \theta(\mu,t)) \\ &= \frac{a}{2}[v^2(\mu,t) + v^2(\mu,t)\cos(4\pi f_c t + 2\theta(\mu,t))] \\ \Rightarrow Z(\mu,t) &= \frac{a}{2}v^2(\mu,t). \end{aligned}$$

Note that the zonal LPF filters out all information about frequency and phase.

Consider further that $S_n(f) = A$ for $|f - f_c| < B/2$ and $|f + f_c| < B/2$ and $S_n(f) = 0$ otherwise, with $A, B > 0$. Zonal bandwidth is assumed larger than $2B$.

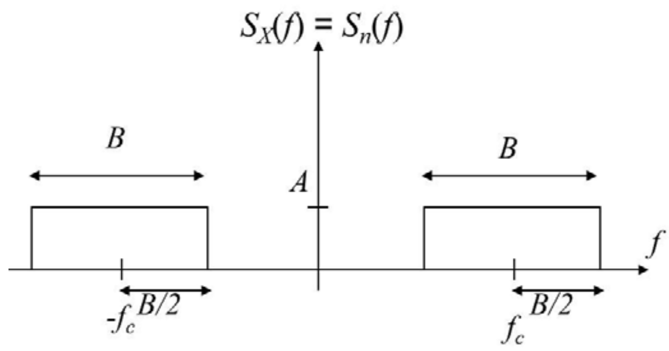


Figure 40:

In the case, $R_X(0) = R_n(0) = 2AB$ and the following can be obtained:

(a) $E\{Y^n(\mu,t)\} = a^n E\{X^{2n}(\mu,t)\} = a^n \frac{(2n)!}{2^n n!}(R_n(0))^n$.

(b) $Var\{Y(\mu,t)\} = E\{Y^2(\mu,t)\} - E^2\{Y(\mu,t)\} = 3a^2 \cdot 4A^2B^2$ $-(a \cdot 2AB)^2 = 8a^2A^2B^2$.

(c) $E\{Z(\mu,t)\} = \frac{a}{2}E\{v^2(\mu,t)\} = \frac{a}{2} \cdot 2R_n(0) = 2aAB$.

$S_Z(f)$ has the triangular shape:

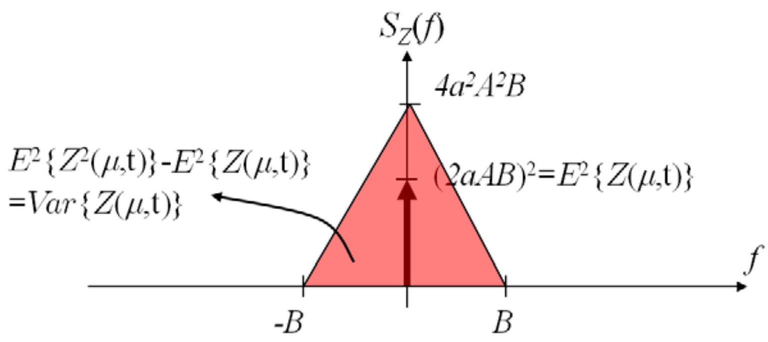


Figure 41:

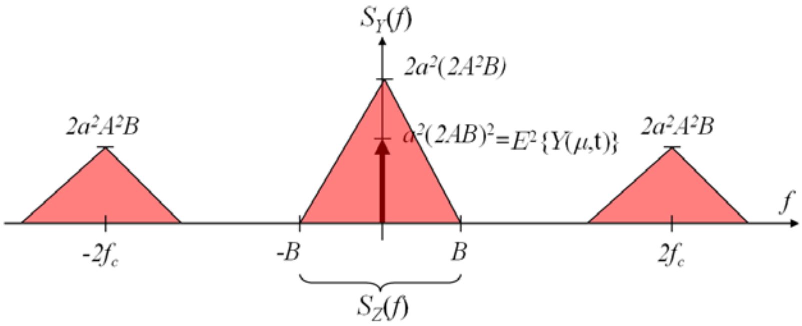
This is derived as follows: First, obtain $R_Y(t_1,t_2)$

$$\begin{aligned} R_Y(t_1,t_2) &= a^2E\{X^2(\mu,t_1)X^2(\mu,t_2)\} = a^2[R_n^2(0) + 2R_n^2(t_1-t_2)] \\ \Rightarrow R_Y(\tau) &= a^2[4A^2B^2 + 2R_n^2(\tau)]. \end{aligned}$$

Thus, the power spectrum $S_Y(f)$ is given by the Fourier transform of $R_Y(\tau)$ as

$$S_Y(f) = 4a^2A^2B^2\delta(f) + 2a^2S_n(f) * S_n(f)$$

which can be obtained as



(You show it!).
Also, $Var\{Z(\mu,t)\} = \frac{1}{2}Var\{Y(\mu,t)\} = 4a^2A^2B^2$.
CASE 2: Assume that $X(\mu,t) = s(\mu,t) + n(\mu,t)$ where the signal $s(\mu,t)$ and the noise $n(\mu,t)$ are independent and jointly WSS processes with zero mean, and $n(\mu,t)$ is as given in CASE 1. Further, we assume that $s^2(\mu,t)$ is WSS. First, $n^2(\mu,t)$ is WSS because $E\{n^2(\mu,t)\} = R_n(0)$ and $R_{n^2}(\tau) = R_n^2(0) + 2R_n^2(\tau)$. Now,

$$\begin{aligned} Y(\mu,t) &= aX^2(\mu,t) \\ &= Y_{S\times S}(\mu,t) + Y_{S\times N}(\mu,t) + Y_{N\times N}(\mu,t) \end{aligned}$$

where $Y_{S\times S}(\mu,t) \triangleq as^2(\mu,t)$, $Y_{S\times N}(\mu,t) \triangleq 2as(\mu,t)n(\mu,t)$, and $Y_{N\times N}(\mu,t) \triangleq an^2(\mu,t)$ represent the signal-times-signal, signal-times-noise and noise-times-noise terms, respectively. $Y(\mu,t)$ has the following statistics:

(a) $E\{Y(\mu,t)\} = a[\sigma_s^2 + \sigma_n^2]$ where $\sigma_s^2 = R_s(0)$ and $\sigma_n^2 = R_n(0)$.
(b) $E\{Y^2(\mu,t)\} = a^2[E\{s^4(\mu,t)\}+6E\{s^2(\mu,t)\}E\{n^2(\mu,t)\}+E\{n^4(\mu,t)\}]$.
(c) $R_Y(t_1,t_2)$ is given by

$$\begin{aligned} R_Y(t_1,t_2) &= a^2E\{[s(\mu,t_1)+n(\mu,t_1)]^2[s(\mu,t_2)+n(\mu,t_2)]^2\} \\ &= a^2[R_{s^2}(t_1-t_2) + 4R_s(t_1-t_2)R_n(t_1-t_2) + 2R_n(0)R_s(0) \\ &\quad +R_{n^2}(t_1-t_2)] \\ \Rightarrow R_Y(\tau) &= R_{S\times S}(\tau) + R_{S\times N}(\tau) + R_{N\times N}(\tau) \end{aligned}$$

where $R_{S\times S}(\tau) = a^2R_{s^2}(\tau)$, $R_{S\times N}(\tau) = 4a^2R_s(\tau)R_n(\tau) + 2a^2\sigma_n^2\sigma_s^2$ and $R_{N\times N}(\tau) = a^2R_{n^2}(\tau)$ represent the autocorrelations corresponding to signal-times-signal, signal-times-noise and noise-times-noise terms, respectively. This yields the power spectrum

$$S_Y(f) = S_{S\times S}(f) + S_{S\times N}(f) + S_{N\times N}(f)$$

where $S_{S\times S}(f) = \mathcal{F}\{R_{S\times S}(\tau)\}$, $S_{S\times N}(f) = \mathcal{F}\{R_{S\times N}(\tau)\}$ and $S_{N\times N}(f) = \mathcal{F}\{R_{N\times N}(\tau)\}$ represent the power spectrums corresponding to signal-times-signal, signal-times-noise and noise-times-noise terms, respectively.

Consider further that $S_n(f) = A$ for $|f - f_c| < B/2$ and $|f + f_c| < B/2$ and $S_n(f) = 0$ otherwise, with $A, B > 0$, and that

$$s(\mu, t) = \sqrt{2S}\cos(2\pi f_c t + \theta(\mu))$$

where S is the signal power and the phase $\theta(\mu)$ is uniform in $[0, 2\pi)$. Note that $\sigma_s^2 = S$ and $\sigma_n^2 = 2AB$. Specifically, we can obtain the following:

(a) $R_{N\times N}(\tau) = a^2 R_{n^2}(\tau) = a^2\sigma_n^4 + 2a^2 R_n^2(\tau)$. Taking Fourier transform gives

$$S_{N\times N}(f) = a^2\sigma_n^4\delta(f) + 2a^2 S_n(f) * S_n(f).$$

(b) Because $R_s(\tau) = S\cos(2\pi f_c\tau)$, $S_s(f) = \frac{S}{2}[\delta(f-f_c)+\delta(f+f_c)]$. This gives

$$\begin{aligned} R_{S\times N}(\tau) &= 2a^2 S\sigma_n^2 + 4a^2 S\cos(2\pi f_c\tau)R_n(\tau) \\ S_{S\times N}(f) &= 2a^2\sigma_n^2 S\delta(f) + 2a^2 S[S_n(f - f_c) + S_n(f + f_c)]. \end{aligned}$$

(c) Because

$$\begin{aligned} R_{S\times S}(\tau) &= a^2 E\{2S\cos^2(2\pi f_c t + \theta(\mu))\cdot 2S\cos^2(2\pi f_c(t-\tau) + \theta(\mu))\} \\ &= a^2 S^2[1 + \frac{1}{2}\cos(4\pi f_c\tau)] \end{aligned}$$

we have

$$S_{S\times S}(f) = a^2 S^2\delta(f) + \frac{1}{4}a^2 S^2[\delta(f - 2f_c) + \delta(f + 2f_c)].$$

Putting together the above results, we have

$$\begin{aligned} R_Y(\tau) &= a^2 S^2[1 + \frac{1}{2}\cos(4\pi f_c\tau)] \\ &\quad + 2a^2[SR_n(0) + 2SR_n(\tau)\cos(2\pi f_c\tau)] \\ &\quad + a^2[R_n^2(0) + 2R_n^2(\tau)] \end{aligned}$$

$$\begin{aligned} S_Y(f) &= a^2\{S^2\delta(f) + \frac{1}{4}S^2[\delta(f - 2f_c) + \delta(f + 2f_c)]\} \\ &\quad + 2a^2\{\sigma_n^2 S\delta(f) + S[S_n(f - f_c) + S_n(f + f_c)]\} \\ &\quad + a^2\{\sigma_n^4\delta(f) + 2S_n(f) * S_n(f)\}. \end{aligned}$$

Graphically, we can depict $S_Y(f)$ as follows

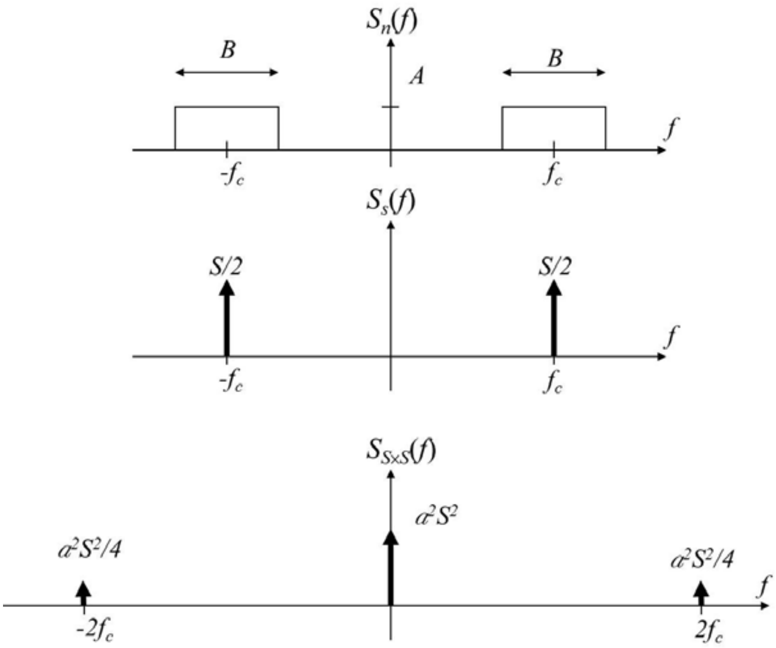
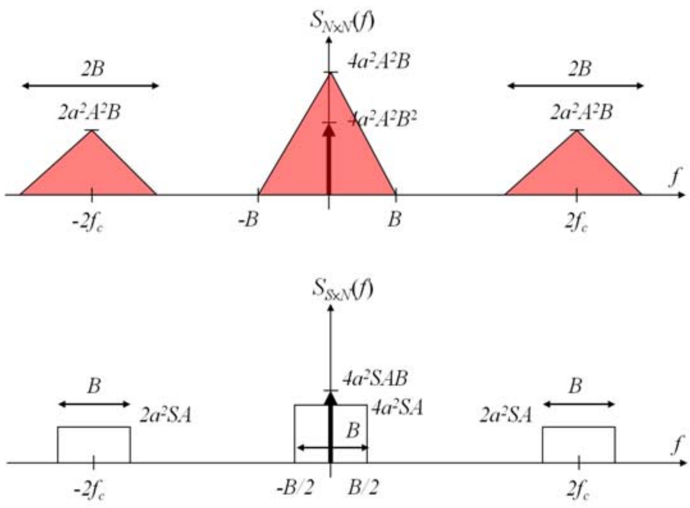
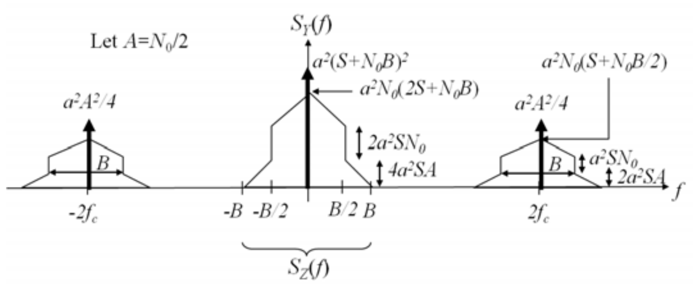
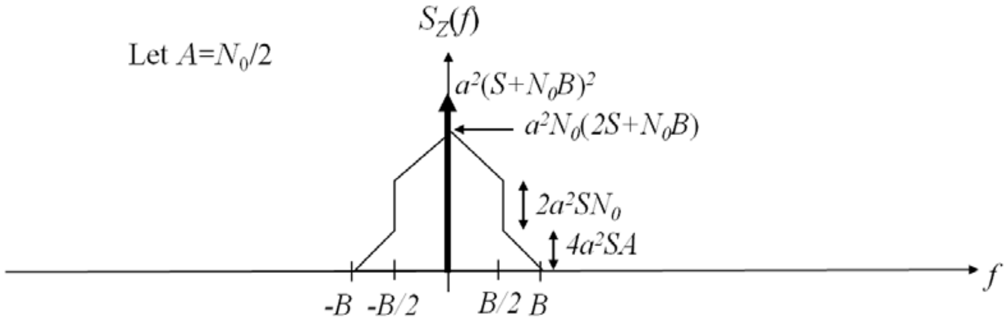


Figure 42:







Applications: Squaring loop for suppressed carrier tracking; energy detection (radiometer) for spectrum sensing.

6 Poisson Processes

- Materials stem from A. Leon-Garcia, Probability and Random Processes for Electrical Engineering, and from H. Larson and B. Shubert, Probabilistic Models in Engineering Sciences.

6.1 Definition

- $N(\mu, t)$, $t \geq 0$, is called a counting process iff all sample functions $N(\mu_0, t)$'s are right-continuous, nondecreasing, and increase only by unit jumps. By default, we set $N(\mu, 0) = 0$.

- A counting process has independent increment iff the number of arrivals in disjoint intervals are independent. The number of arrivals in $(s, t]$ is $N(\mu, t) - N(\mu, s)$. Thus, if we let $s_1 \leq t_1 \leq s_2 \leq t_2 \leq \cdots \leq s_n \leq t_n$, then $\{N(\mu, t_j) - N(\mu, s_j)\}_{j=1}^{n}$ are independent random variables for any choice of n and $\{s_i, t_i\}_{i=1}^{n}$.

- A counting process has stationary increment iff the number of arrivals in a time interval depends only on the length of the time interval. In this case,

  $$\Pr\{N(\mu, t+s) - N(\mu, t) = k\} = \Pr\{N(\mu, s) - N(\mu, 0) = k\} \quad \forall t, s \geq 0, \forall k \in \mathcal{N}$$

  with $\mathcal{N}$ being the set of nonnegative integers.

- Defn: A counting process $\{N(\mu, t), t \geq 0\}$ is called a Poisson process provided that

  1. the process has stationary and independent increments, and
  2. $\Pr\{N(\mu, h) \geq 2\} = o(h)$.

  Notes:

  1. $f(h) = o(h)$ if $\lim_{h \to 0} \frac{f(h)}{h} = 0$.
     This means (1) $\lim_{h \to 0} f(h) = 0$ and (2) $|f(h)|$ is small when compared to a positive h as h approaches to zero, i.e., $|f(h)| \ll h$ as $h \to 0$.
  2. If $f(h)$ is differentiable, then $f(h) = o(h)$ is equivalent to $f(h) \to 0$ and $f'(h) \to 0$ as $h \to 0$.
  3. $o(h) + o(h) = o(h)$, $o(h) - o(h) = o(h)$ and $-o(h) = o(h)$.

- Defn: $p_n(t) \triangleq \Pr\{N(\mu,t)=n\}$.
  Then, for $n>0$,

$$
\begin{aligned}
p_n(t+h) &= \Pr\{N(\mu,t+h)=n\} \quad t,h \geq 0 \\
&= \sum_{k=0}^{n} \Pr\{N(\mu,t+h)=n, N(\mu,t)=k\} \\
&= \sum_{k=0}^{n} \Pr\{N(\mu,t+h)-N(\mu,t)=n-k, N(\mu,t)=k\} \\
&= \sum_{k=0}^{n} \Pr\{N(\mu,t+h)-N(\mu,t)=n-k, N(\mu,t)-N(\mu,0)=k\} \\
&= \sum_{k=0}^{n} \Pr\{N(\mu,t+h)-N(\mu,t)=n-k\} \Pr\{N(\mu,t)-N(\mu,0)=k\} \\
&\quad \text{(independent increments)} \\
&= \sum_{k=0}^{n} \Pr\{N(\mu,h)=n-k\} \Pr\{N(\mu,t)=k\} \\
&\quad \text{(stationary increments)} \\
&= \sum_{k=0}^{n} p_k(t)p_{n-k}(h) \\
&= p_0(h)p_n(t) + p_1(h)p_{n-1}(t) + \sum_{k=2}^{n} p_k(h)p_{n-k}(t). \qquad \oplus
\end{aligned}
$$

  Now let us look at $p_0(h)$:

$$
\begin{aligned}
p_0(h) &= \Pr\{N(\mu,h)=0\} \qquad h \geq 0 \\
p_0(h+s) &= \Pr\{N(\mu,h+s)=0\} \qquad h,s \geq 0 \\
&= \Pr\{N(\mu,h)=0, N(\mu,h+s)-N(\mu,h)=0\} \\
\Rightarrow p_0(h+s) &= p_0(h)p_0(s).
\end{aligned}
$$

- Lemma: With the initial condition

$$p_0(0) = \Pr\{N(\mu,0)=0\} = 1$$

  the only solution to

$$p_0(h+s) = p_0(h)p_0(s) \text{ for } h,s \geq 0 \qquad \circledast$$

  is $p_0(t)=e^{-\lambda t}$, $\lambda \geq 0$.

Note: This λ is called the rate of the Poisson process.

Proof: (a) For arbitrary positive integers n and m, $\circledast$ gives that

$$\begin{aligned}
p_0(\frac{2}{n}) &= p_0(\frac{1}{n}+\frac{1}{n}) = p_0(\frac{1}{n})p_0(\frac{1}{n}) = [p_0(\frac{1}{n})]^2 \\
p_0(\frac{3}{n}) &= p_0(\frac{2}{n}+\frac{1}{n}) = p_0(\frac{2}{n})p_0(\frac{1}{n}) = [p_0(\frac{1}{n})]^3 \\
&\vdots \\
p_0(\frac{m}{n}) &= [p_0(\frac{1}{n})]^m.
\end{aligned}$$

Because

$$p_0(1) = p_0(\frac{n}{n}) = [p_0(\frac{1}{n})]^n$$

we further have that

$$p_0(\frac{1}{n}) = [p_0(1)]^{1/n}. \qquad (+)$$

Thus,

$$p_0(\frac{m}{n}) = [p_0(1)]^{m/n}. \qquad (-)$$

Let us show that $p_0(1) > 0$: Assume $p_0(1) = \Pr\{N(\mu,1) = 0\} = 0$. Then, by $(+)$, $p_0(\frac{1}{n}) = 0$ for all positive integers n. This, in conjunction with the property that $p_0(h) = \Pr\{N(\mu,h) = 0\}$ is continuous from the right, implies that

$$p_0(0) = p_0(\lim_{n\to\infty}\frac{1}{n}) = \lim_{n\to\infty} p_0(\frac{1}{n}) = 0$$

which is a contradiction to the given fact that $p_0(0) = \Pr\{N(\mu,0) = 0\} = 1$. Thus, $p_0(1) > 0$. In the case, we can define $\lambda = -\ln[p_0(1)]$ with $\lambda \geq 0$. This gives

$$p_0(1) = \exp\{-\lambda\}$$

and from $(-)$

$$p_0(\frac{m}{n}) = \exp\{-\lambda\frac{m}{n}\}.$$

Therefore, for any positive rational t,

$$p_0(t) = \exp\{-\lambda t\} \qquad (0)$$

with $\lambda \geq 0$.

(b) Let us prove the same relation as (0) for a positive irrational number t. First, recall from calculus that for each positive integer n, there exists a rational number t_n in $(t, t+\frac{1}{n})$. Since $t < t_n < t+\frac{1}{n}$, $\lim_{n\to\infty} t_n$ exists and is t. Next, because $p_0(h) = \Pr\{N(\mu, h) = 0\}$ is continuous from the right,

$$p_0(t) = p_0(\lim_{n\to\infty} t_n) = \lim_{n\to\infty} p_0(t_n).$$

Because t_n is a positive rational number, (0) implies that

$$p_0(t) = \lim_{n\to\infty} \exp\{-\lambda t_n\} = \exp\{-\lambda t\}$$

because $\exp\{-\lambda t\}$ is continuous in t. Thus, for any positive irrational t, (0) also holds for $\lambda \geq 0$.

Q.E.D.

- Some Notes:

(1) $\circledast$ is also called the memoryless property. This is because

$$\begin{aligned}
& p_0(h+s) = p_0(h)p_0(s) \\
\Leftrightarrow\ & \frac{\Pr\{N(\mu, h+s) = 0\}}{\Pr\{N(\mu, h) = 0\}} = \Pr\{N(\mu, s) = 0\} \\
\Leftrightarrow\ & \frac{\Pr\{N(\mu, h+s) = 0, N(\mu, h) = 0\}}{\Pr\{N(\mu, h) = 0\}} = \Pr\{N(\mu, s) = 0\} \\
\Leftrightarrow\ & \Pr\{N(\mu, h+s) = 0 | N(\mu, h) = 0\} = \Pr\{N(\mu, s) = 0\}
\end{aligned}$$

means that the count $N(\mu, t)$ will continue to be zero for another s seconds does not depend on how long it has been zero.

(2) In general, the memoryless property for a continuous random variable $X(\mu)$ is defined by

$$\Pr\{X(\mu) > h+s | X(\mu) > h\} = \Pr\{X(\mu) > s\} \text{ for } h, s \geq 0. \qquad (++)$$

Note that $\Pr\{X(\mu) > s\}$ is the complementary distribution function for $X(\mu)$.

(3) If we define $T(\mu) = \inf\{t; N(\mu, t) = 1\}$, i.e., the time length that the process remains zero from the beginning (in fact, $T(\mu)$ is the first arrival time), then

$$\Pr\{N(\mu, h+s) = 0 | N(\mu, h) = 0\} = \Pr\{N(\mu, s) = 0\} \text{ for } h, s \geq 0$$

can be equivalently expressed as

$$\Pr\{T(\mu) > h+s | T(\mu) > h\} = \Pr\{T(\mu) > s\} \text{ for } h, s \geq 0.$$

(Here, $\inf\{t; N(\mu,t)=1\}$ denotes the smallest value of t satisfying the condition $N(\mu,t)=1$.) Because $p_0(t) = \Pr\{N(\mu,t)=0\} = \exp\{-\lambda t\}$ for $\lambda > 0$ is the complementary distribution function for an exponential random variable with parameter λ. Thus, the lemma tells us that the exponential random variable is the only continuous random variable that have the memoryless property (++).

(4) If a continuous random variable $X(\mu)$ is memoryless, then $\Pr\{X(\mu) \leq 0\} = 0$. To see this, note from (++) that $\Pr\{X(\mu) \leq h+s | X(\mu) > h\} = \Pr\{X(\mu) \leq s\}$. Letting $h=s=0$, we have $\Pr\{X(\mu) \leq 0 | X(\mu) > 0\} = \Pr\{X(\mu) \leq 0\}$. Because $\Pr\{X(\mu) \leq 0 | X(\mu) > 0\} = 0$, we obtain $\Pr\{X(\mu) \leq 0\} = 0$. This shows that the memoryless property can not be defined for random variables possessing negative values with positive probability.

- Next,

$$\begin{aligned}
\lim_{h \to 0} \frac{1}{h} p_1(h) &= \lim_{h \to 0} \frac{1}{h}[1 - p_0(h) - \underbrace{\sum_{k \geq 2} p_k(h)}_{\Pr\{N(\mu,h) \geq 2\} = o(h)}] \\
&= \lim_{h \to 0} \frac{1}{h}[1 - e^{-\lambda h} - o(h)] \\
\text{(L}'\text{Hopital's rule)} &= \lambda - \lim_{h \to 0} \frac{o(h)}{h} \\
&= \lambda.
\end{aligned}$$

Another way to state this is

$$p_1(h) = \Pr\{N(\mu,h)=1\} = \lambda h + o(h)$$

since

$$\lim_{h \to 0} \frac{1}{h}(\Pr\{N(\mu,h)=1\} - \lambda h) = 0.$$

Similarly, we can find that

$$p_0(h) = 1 - \lambda h + o(h).$$

Notes:

1. $p_0(h) = 1 - p_1(h) - \sum_{k\geq 2} p_k(h) = 1 - (\lambda h + o(h)) - o(h) = 1 - \lambda h + o(h)$.
2. $p_0(h) = e^{-\lambda h} = 1 - \lambda h + \underbrace{\sum_{k\geq 2} \frac{(-\lambda h)^k}{k!}}_{o(h)} = 1 - \lambda h + o(h)$.

- Plugging $p_0(h) = 1 - \lambda h + o(h)$ and $p_1(h) = \lambda h + o(h)$ into $\oplus$, we have for $n > 0$ that

$$p_n(t+h) = (1 - \lambda h + o(h))p_n(t) + (\lambda h + o(h))p_{n-1}(t) + \sum_{k=2}^{n} o(h)p_{n-k}(t)$$

$$\begin{aligned}
\Rightarrow p_n(t+h) - p_n(t) &= -\lambda h p_n(t) + \lambda h p_{n-1}(t) + o(h)\underbrace{\sum_{k=0}^{n} p_{n-k}(t)}_{\leq 1} \\
&= -\lambda h p_n(t) + \lambda h p_{n-1}(t) + o(h) \\
\Rightarrow \tfrac{1}{h}[p_n(t+h) - p_n(t)] &= -\lambda p_n(t) + \lambda p_{n-1}(t) + \tfrac{o(h)}{h}.
\end{aligned}$$

As $h \to 0$,
$$p_n'(t) = -\lambda p_n(t) + \lambda p_{n-1}(t), \ n > 0.$$
Similarly,
$$p_n'(t) = -\lambda p_n(t), \ n = 0.$$
This set of equations is an example of the Kolmogorov forward equations. The initial conditions for these differential equations are

$$\begin{aligned}
p_n(0) &= 0, \ \text{if } n > 0 \\
p_0(0) &= 1.
\end{aligned}$$

- Theorem: For a Poisson Process $\{N(\mu, t); t \geq 0\}$ with rate λ,

$$\Pr\{N(\mu, t) = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}.$$

Proof:

1. $p_0'(t) = -\lambda p_0(t)$ with $p(0) = 1 \Rightarrow p_0(t) = e^{-\lambda t} = e^{-\lambda t}\frac{(\lambda t)^0}{0!}$.
2. Assume that the distribution holds true for $n = 0, 1, ..., k$. Then, for $n = k+1$,
$$p_{k+1}'(t) = -\lambda p_{k+1}(t) + \lambda p_k(t) \qquad \text{with} \qquad p_{k+1}(0) = 0$$
$$\Rightarrow p_{k+1}'(t) + \lambda p_{k+1}(t) = \lambda p_k(t) = \lambda\left(e^{-\lambda t}\frac{(\lambda t)^k}{k!}\right)$$

$\Rightarrow \underbrace{(e^{\lambda t})p_{k+1}^{'}(t) + (\lambda e^{\lambda t})p_{k+1}(t)} = \lambda \frac{(\lambda t)^k}{k!}$

$\Rightarrow \qquad \frac{d}{dt}(e^{\lambda t}p_{k+1}(t)) = \lambda \frac{(\lambda t)^k}{k!}$

$\Rightarrow e^{\lambda t}p_{k+1}(t) - e^{\lambda \cdot 0}p_{k+1}(0) = \frac{(\lambda t)^{k+1}}{(k+1)!} - \frac{(\lambda \cdot 0)^{k+1}}{(k+1)!}$

$\Rightarrow p_{k+1}(t) = e^{-\lambda t}\frac{(\lambda t)^{k+!}}{(k+1)!}.$

This proves the theorem by induction. Q.E.D.

Note: Another proof by moment generating function is also available, see Larson and Shubert for details.

- Corollary: A counting process $\{N(\mu, t); t \geq 0\}$ is a Poisson process iff

  1. the process has stationary and independent increments,
  2. $\Pr\{N(\mu, t) = k\} = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$ for some $\lambda > 0$.

- Defn: $T_1(\mu) = \inf\{t; N(\mu, t) = 1\}$.

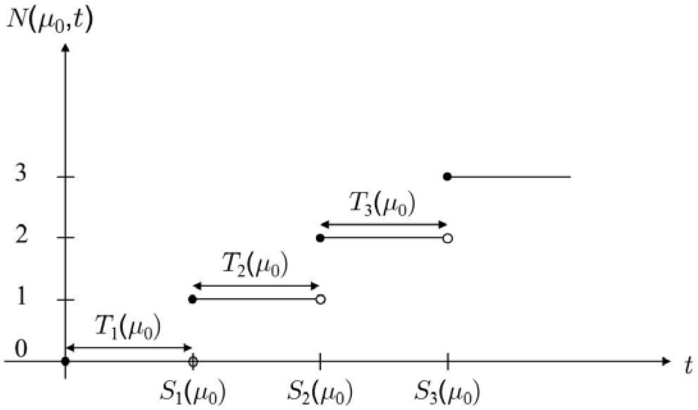


Figure 43:

Then, $\Pr\{T_1(\mu) > t\} = \Pr\{N(\mu, t) = 0\} = e^{-\lambda t}$
which says that $T_1(\mu)$ is an exponential random variable.
Now, define $S_n(\mu) = \inf\{t; N(\mu, t) = n\}$ as the time of the n^{th} arrival.
We notice that $S_1(\mu) = T_1(\mu)$.
Also, define the interarrival time $T_n(\mu) = S_n(\mu) - S_{n-1}(\mu)$ with $S_0(\mu) = 0$ by default.
Thus, $S_n(\mu) = \sum_{k=1}^{n} T_k(\mu)$.

Note: $N(\mu, t) \geq n$ iff $S_n(\mu) \leq t$.

- Proposition: $\{T_i(\mu)\}_{i=1}^{\infty}$ are independent and identically distributed random variables which are exponentially distributed.

  Note: This is true since the Poisson process has independent and stationary increments and since $T_1(\mu)$ is an exponential random variable.

- Proposition: $S_n(\mu)$ is a gamma random variable.

  Proof: Now, $S_1(\mu) = T_1(\mu)$ is an exponential random variable, and thus a gamma random variable. Next,

$$\begin{aligned}
F_{S_n}(t) &\triangleq \Pr\{S_n(\mu) \leq t\} \\
&= \Pr\{N(\mu, t) \geq n\} \\
&= \sum_{k=n}^{\infty} \Pr\{N(\mu, t) = k\} \\
&= \sum_{k=n}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t}.
\end{aligned}$$

  Differentiating both sides for $n \in \{1, 2, ...\}$,

$$\begin{aligned}
f_{S_n}(t) &= \frac{d}{dt} F_{S_n}(t) \\
&= \sum_{k=n}^{\infty} (e^{-\lambda t} \frac{(\lambda t)^{(k-1)}}{(k-1)!} \lambda - \lambda e^{-\lambda t} \frac{(\lambda t)^k}{k!}) \\
&= \lambda e^{-\lambda t} \frac{(\lambda t)^{(n-1)}}{(n-1)!}.
\end{aligned}$$

  Thus, $S_n(\mu)$ is a gamma random variable. Q.E.D.

  Notes:

  (1) $E\{N(\mu, t)\} = \lambda t$, $Var\{N(\mu, t)\} = \lambda t$.
  (2) $E\{T_k(\mu)\} = \frac{1}{\lambda}$, $Var\{T_k(\mu)\} = \frac{1}{\lambda^2}$.
  (3) $E\{S_k(\mu)\} = \frac{k}{\lambda}$, $Var\{S_k(\mu)\} = \frac{k}{\lambda^2}$.
  (4) Distributions:
  counts, $N(\mu, t) \rightarrow$ Poisson process
  arrival times, $S_k(\mu) \rightarrow$ gamma random variables
  interarrival times, $T_k(\mu) \rightarrow$ exponential random variables

- Theorem: If

  (1) $\{N_I(\mu,t); t \geq 0\}$ is a Poisson process with rate λ_I,

  (2) $\{N_{II}(\mu,t); t \geq 0\}$ is a Poisson process with rate λ_{II}, and

  (3) $N_I(\mu,t)$ and $N_{II}(\mu,t)$ are independent,

  then $\{N(\mu,t) = N_I(\mu,t) + N_{II}(\mu,t); t \geq 0\}$ is a Poisson process with rate $\lambda = \lambda_I + \lambda_{II}$.

  Proof: First, $N(\mu,t)$ has stationary increment because

$$
\begin{aligned}
& \Pr\{N(\mu,t+s) - N(\mu,t) = K\} \\
= & \Pr\{[N_I(\mu,t+s) + N_{II}(\mu,t+s)] - [N_I(\mu,t) - N_{II}(\mu,t)] = K\} \\
= & \sum_{k=0}^{K} \Pr\{N_I(\mu,t+s) - N_I(\mu,t) = k, N_{II}(\mu,t+s) - N_{II}(\mu,t) = K-k\} \\
= & \sum_{k=0}^{K} \Pr\{N_I(\mu,t+s) - N_I(\mu,t) = k\} \Pr\{N_{II}(\mu,t+s) - N_{II}(\mu,t) = K-k\} \\
= & \sum_{k=0}^{K} \Pr\{N_I(\mu,s) - N_I(\mu,0) = k\} \Pr\{N_{II}(\mu,s) - N_{II}(\mu,0) = K-k\} \\
= & \sum_{k=0}^{K} \Pr\{N_I(\mu,s) - N_I(\mu,0) = k, N_{II}(\mu,s) - N_{II}(\mu,0) = K-k\} \\
= & \Pr\{[N_I(\mu,s) + N_{II}(\mu,s)] - [N_I(\mu,0) - N_{II}(\mu,0)] = K\} \\
= & \Pr\{N(\mu,s) - N(\mu,0) = k\} \quad \forall t,s \geq 0, \forall K \in \mathcal{N}.
\end{aligned}
$$

  Second, because $N_I(\mu,t)$ and $N_{II}(\mu,t)$ are independent and $\{N_I(\mu,t_j) - N_I(\mu,s_j)\}_{j=0}^{n}$ and $\{N_{II}(\mu,t_j) - N_{II}(\mu,s_j)\}_{j=0}^{n}$ are independent random variables, $\{N(\mu,t_j) - N(\mu,s_j)\}_{j=0}^{n}$ are also independent increments and thus $N(\mu,t)$ has independent increment, where we let $s_1 \leq t_1 \leq s_2 \leq$

$t_2 \leq ... \leq s_n \leq t_n$. Third, it is shown that

$$\begin{aligned}
\Pr\{N(\mu,t)=K\} &= \Pr\{N_I(\mu,t)+N_{II}(\mu,t)=K\} \\
&= \sum_{k=0}^{K} \Pr\{N_I(\mu,t)=k\}\Pr\{N_{II}(\mu,t)=K-k\} \\
&= \sum_{k=0}^{K} [e^{-\lambda_I}\frac{\lambda_I^k}{k!}][e^{-\lambda_{II}}\frac{\lambda_{II}^{K-k}}{(K-k)!}] \\
&= \frac{e^{-\lambda_I+\lambda_{II}}}{K!}\sum_{k=0}^{K}\frac{K!}{k!(K-k)!}\lambda_I^k\lambda_{II}^{K-k} \\
&= e^{-\lambda_I+\lambda_{II}}\frac{(\lambda_I+\lambda_{II})^K}{K!}.
\end{aligned}$$

Because $N(\mu,t)$ has stationary and independent increments and $\Pr\{N(\mu,t) = K\} = e^{-\lambda_I+\lambda_{II}}\frac{(\lambda_I+\lambda_{II})^K}{K!}$, $N(\mu,t)$ is a Poisson process with rate $\lambda_I+\lambda_{II}$. Q.E.D.

- Example A on Poisson Process:
  If cars enter a toll booth according to a Poisson process with a rate of ten cars per minute, and if each car heads west independently of the other, with probability one third.

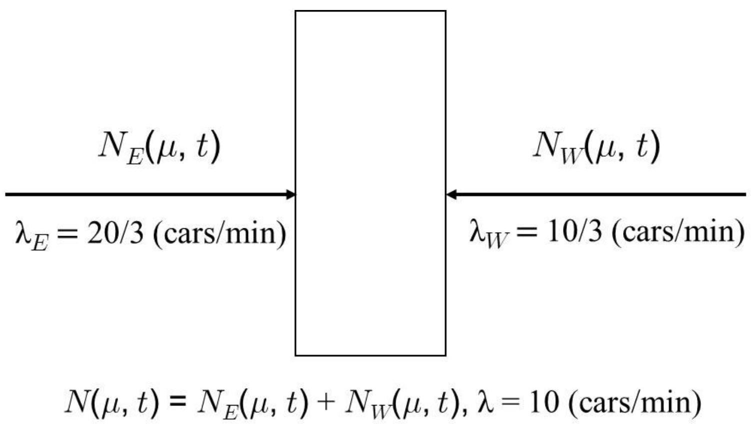


Figure 44:

(a) Find the probability that no one heads west for two minutes.

The number of cars westbound is a Poisson process $N_W(\mu, t)$ with rate 10/3.

$$\Pr\{N_W(\mu,2)=0\} = e^{-\frac{10}{3}*2} = e^{-\frac{20}{3}}.$$

(b) Find the probability that two cars head west if five cars go through at a given minute, i.e.,

$$\Pr\{N_W(\mu,1)=2|N(\mu,1)=5\}.$$

Now,

$$\begin{aligned}
& \Pr\{N_W(\mu,t)=k|N(\mu,t)=n\},\ k\leq n \\
= & \Pr\{N_W(\mu,t)=k, N(\mu,t)=n\}/\Pr\{N(\mu,t)=n\} \\
= & \Pr\{N_W(\mu,t)=k, N_E(\mu,t)=n-k\}/\Pr\{N(\mu,t)=n\} \\
= & \Pr\{N_W(\mu,t)=k\}\Pr\{N_E(\mu,t)=n-k\}/\Pr\{N(\mu,t)=n\} \\
= & (e^{-\lambda_W t}\frac{(\lambda_W t)^k}{k!})(e^{-\lambda_E t}\frac{(\lambda_E t)^{(n-k)}}{(n-k)!})/(e^{-\lambda t}\frac{(\lambda t)^n}{n!}) \\
= & \tbinom{n}{k}\frac{\lambda_W^k\lambda_E^{n-k}}{\lambda^n} \\
= & \tbinom{n}{k}(\frac{\lambda_W}{\lambda})^k(\frac{\lambda_E}{\lambda})^{n-k} \\
\Rightarrow & \Pr\{N_W(\mu,1)=2|N(\mu,1)=5\} = \tbinom{5}{2}(\frac{1}{3})^2(\frac{2}{3})^3.
\end{aligned}$$

(c) If only one car headed west in the first minute, what is the probability that it arrived in the first half minute?
This is equivalent to finding $\Pr\{T_1(\mu)\leq\frac{1}{2}|N_W(\mu,1)=1\}$.

Now,

$$\begin{aligned}
& \Pr\{T_1(\mu) \leq s | N_W(\mu, t) = 1\} \qquad s \leq t \\
= \;& \Pr\{N_W(\mu, s) \geq 1 | N_W(\mu, t) = 1\} \\
= \;& \Pr\{N_W(\mu, s) \geq 1, N_W(\mu, t) = 1\} / \Pr\{N_W(\mu, t) = 1\} \\
= \;& \Pr\{N_W(\mu, s) = 1, N_W(\mu, t) = 1\} / \Pr\{N_W(\mu, t) = 1\} \\
= \;& \frac{\Pr\{N_W(\mu, s) - N_W(\mu, 0) = 1, N_W(\mu, t) - N_W(\mu, s) = 0\}}{\Pr\{N_W(\mu, t) = 1\}} \\
= \;& \frac{\Pr\{N_W(\mu, s) - N_W(\mu, 0) = 1\} \Pr\{N_W(\mu, t) - N_W(\mu, s) = 0\}}{\Pr\{N_W(\mu, t) = 1\}} \\
= \;& \frac{\Pr\{N_W(\mu, s) = 1\} \Pr\{N_W(\mu, t - s) = 0\}}{\Pr\{N_W(\mu, t) = 1\}} \\
= \;& (e^{-\lambda_W s} \frac{(\lambda_W s)^1}{1!})(e^{-\lambda_W (t-s)} \frac{(\lambda_W (t - s))^0}{0!}) / (e^{-\lambda_W t} \frac{(\lambda_W t)^1}{1!}) \\
= \;& \frac{s}{t} \qquad \text{(uniform in } [0, t]). \\
\Rightarrow \;& \Pr\{T_1(\mu) \leq \frac{1}{2} | N_W(\mu, 1) = 1\} = \frac{1}{2}.
\end{aligned}$$

- Example B on Poisson Process:

  Suppose that a network router handles data packets that arrive according to a Poisson process $N(\mu, t)$ with rate λ packets per minute.

  (a) What is the probability that no packet arrives for the first 15 and last 15 seconds of a minute?

  The probability is given by

$$\begin{aligned} & \Pr\{N(\mu, \frac{1}{4}) - N(\mu, 0) = 0, N(\mu, 1) - N(\mu, \frac{3}{4}) = 0\} \\ = & \Pr\{N(\mu, \frac{1}{4}) - N(\mu, 0) = 0\} \Pr\{N(\mu, 1) - N(\mu, \frac{3}{4}) = 0\} \\ = & \exp\{-\frac{\lambda}{2}\}. \end{aligned}$$

  (Poisson process has independent and stationary increments.)

  (b) Find the probability that there are j arrivals in the first $t - d$ minutes when there are k arrivals in a particular period of t minutes, with $j \in \{0, 1, ..., k\}$, k a nonnegative integer, and $0 < d < t$.

  The probability is given by

$$\begin{aligned} & \Pr\{N(\mu, t-d) = j | N(\mu, t) = k\} \\ = & \frac{\Pr\{N(\mu, t-d) = j, N(\mu, t) = k\}}{\Pr\{N(\mu, t) = k\}} \\ = & \frac{\Pr\{N(\mu, t-d) = j\} \Pr\{N(\mu, t) - N(\mu, t-d) = k - j\}}{\Pr\{N(\mu, t) = k\}} \\ = & \frac{\frac{\lambda^j (t-d)^j}{j!} e^{-\lambda(t-d)} \frac{\lambda^{k-j} d^{k-j}}{(k-j)!} e^{-\lambda d}}{\frac{\lambda^k t^k}{k!} e^{-\lambda t}} \\ = & \binom{k}{j} (1 - \frac{d}{t})^j (\frac{d}{t})^{k-j} \end{aligned}$$

  which is binomially distributed.

  (c) Let $I(\mu)$ and $J(\mu)$ be the numbers of data packets arriving in the first minute and the last minute in a specific hour. Find the probability that $I(\mu) + J(\mu) = 1$.

First, find the probability

$$\begin{aligned}
\Pr\{I(\mu)+J(\mu)=1\} &= \Pr\{I(\mu)=0, J(\mu)=1\}+\Pr\{I(\mu)=1, J(\mu)=0\} \\
&= \Pr\{N(\mu,1)-N(\mu,0)=0, N(\mu,60)-N(\mu,59)=1\}+ \\
&\quad \Pr\{N(\mu,1)-N(\mu,0)=1, N(\mu,60)-N(\mu,59)=0\} \\
&= 2\Pr\{N(\mu,1)=1\}\Pr\{N(\mu,1)=0\} \\
&\quad (N(\mu,t) \text{ has independent and stationary increments} \\
&\quad \text{with } N(\mu,0)=0.) \\
&= 2\lambda t e^{-2\lambda t}.
\end{aligned}$$

6.2 An Example Process Derived From Poisson Processes–Random Telegraph Signal

- Let $X(\mu,t)$ be a random process assuming the value ± 1, and with

  (a) $\Pr\{X(\mu,0)=+1\}=\Pr\{X(\mu,0)=-1\}=\frac{1}{2}$,

  (b) $X(\mu,t)$ changes polarity with each occurrence of an event in a Poisson process $N(\mu,t)$ of rate λ.

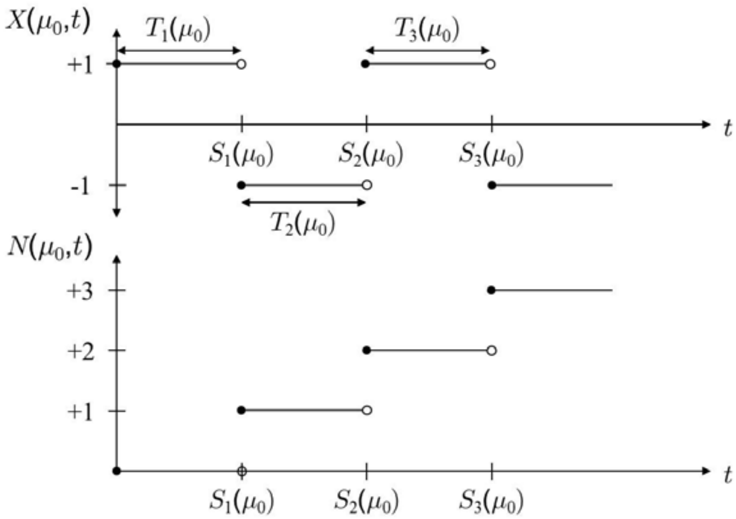


Figure 45:

The number of polarity transitions $N(\mu,t)$ in the interval $[0,t]$ is a Poisson process.

Now,

(i)

$$\begin{aligned}
& \Pr\{X(\mu,t)=+1|X(\mu,0)=+1\} \\
= \ & \Pr\{N(\mu,t)=even\} \\
= \ & \sum_{j=0}^{\infty}\frac{(\lambda t)^{2j}}{(2j)!}e^{-\lambda t} \\
= \ & e^{-\lambda t}\frac{1}{2}(\sum_{n=0}^{\infty}\frac{(\lambda t)^n}{(n)!}+\sum_{n=0}^{\infty}\frac{(-\lambda t)^n}{(n)!}) \\
= \ & e^{-\lambda t}\frac{1}{2}(e^{\lambda t}+e^{-\lambda t}) \\
= \ & \frac{1}{2}(1+e^{-2\lambda t}) \\
= \ & \Pr\{X(\mu,t)=-1|X(\mu,0)=-1\}.
\end{aligned}$$

(ii)

$$\begin{aligned}
& \Pr\{X(\mu,t)=+1|X(\mu,0)=-1\} \\
= \ & \Pr\{X(\mu,t)=-1|X(\mu,0)=+1\} \\
= \ & \Pr\{N(\mu,t)=odd\} \\
= \ & 1-\Pr\{N(\mu,t)=even\} \\
= \ & \frac{1}{2}(1-e^{-2\lambda t}).
\end{aligned}$$

(iii)

$$\begin{aligned}
& \Pr\{X(\mu,t)=+1\} \\
= \ & \Pr\{X(\mu,t)=+1|X(\mu,0)=+1\}\Pr\{X(\mu,0)=+1\}+ \\
& \Pr\{X(\mu,t)=+1|X(\mu,0)=-1\}\Pr\{X(\mu,0)=-1\} \\
= \ & \frac{1}{2}[\frac{1}{2}(1+e^{-2\lambda t})+\frac{1}{2}(1-e^{-2\lambda t})] \\
= \ & \frac{1}{2} \\
= \ & \Pr\{X(\mu,t)=-1\}.
\end{aligned}$$

Therefore, the random telegraph signal is equally likely to be ± 1 at any time $t>0$.

(iv)

$$\begin{aligned}
E\{X(\mu,t)\} &= (+1)(\frac{1}{2}) + (-1)(\frac{1}{2}) = 0 \\
Var\{X(\mu,t)\} &= E\{X^2(\mu,t)\} = 1.
\end{aligned}$$

(v) Assuming $t_1 > t_2$ without loss of generality,

$$\begin{aligned}
R_X(t_1,t_2) &= C_X(t_1,t_2) \\
&= E\{X(\mu,t_1)X(\mu,t_2)\} \\
&= (+1)\Pr\{X(\mu,t_1) = X(\mu,t_2)\} + (-1)\Pr\{X(\mu,t_1) \neq X(\mu,t_2)\} \\
&= 2\Pr\{X(\mu,t_1) = X(\mu,t_2)\} - 1.
\end{aligned}$$

Now, since

$$\begin{aligned}
& \Pr\{X(\mu,t_1) = X(\mu,t_2)\} \\
= & \sum_{\alpha=\pm1} \Pr\{X(\mu,t_1) = \alpha | X(\mu,t_2) = \alpha\} \Pr\{X(\mu,t_2) = \alpha\} \\
= & \frac{1}{2} \sum_{\alpha=\pm1} \Pr\{X(\mu,t_1) = \alpha | X(\mu,t_2) = \alpha\} \\
= & \frac{1}{2} \sum_{\alpha=\pm1} \Pr\{N(\mu,t_1 - t_2) = even\} \\
= & \frac{1}{2}(2)(\frac{1}{2}(1 + e^{-2\lambda(t_1-t_2)})) \\
= & \frac{1}{2}(1 + e^{-2\lambda(t_1-t_2)})
\end{aligned}$$

for $t_1 > t_2$, we obtain

$$R_X(t_1,t_2) = e^{-2\lambda|t_1-t_2|}$$

for all t_1, t_2. Therefore, the time samples of $X(\mu,t)$ is less and less correlated as $|t_1 - t_2|$ increases.

- Notes:

  (a) The random telegraph signal $X(\mu,t)$ is a stationary random process when $\Pr\{X(\mu,0) = \pm1\} = \frac{1}{2}$.

  (b) The random telegraph signal settles into stationary behavior as $t \to \infty$ even when $\Pr\{X(\mu,0) = \pm1\} \neq \frac{1}{2}$.

Proof for (a) and (b): For stationarity, we need to show the equality

$$\Pr\{X(\mu,t_1)=a_1,...,X(\mu,t_k)=a_k\}=\Pr\{X(\mu,t_1+\tau)=a_1,...,X(\mu,t_k+\tau)=a_k\}$$

for any k, any τ, any $t_1<t_2<\cdots<t_k$, and any $a_j=\pm 1$. Now,

$$\begin{aligned}
& \Pr\{X(\mu,t_1)=a_1,...,X(\mu,t_k)=a_k\} \\
= \ & \Pr\{X(\mu,t_1)=a_1\}\Pr\{X(\mu,t_2)=a_2|X(\mu,t_1)=a_1\} \\
& \Pr\{X(\mu,t_3)=a_3|X(\mu,t_1)=a_1,X(\mu,t_2)=a_2\}... \\
& \Pr\{X(\mu,t_k)=a_k|X(\mu,t_1)=a_1,...,X(\mu,t_{k-1})=a_{k-1}\} \\
= \ & \Pr\{X(\mu,t_1)=a_1\}\Pr\{X(\mu,t_2)=a_2|X(\mu,t_1)=a_1\} \\
& \Pr\{X(\mu,t_3)=a_3|X(\mu,t_2)=a_2\}\cdots \\
& \Pr\{X(\mu,t_k)=a_k|X(\mu,t_{k-1})=a_{k-1}\} \qquad \ominus
\end{aligned}$$

since the underlined Poisson process $N(\mu,t)$ has independent increments.

Similarly,

$$\begin{aligned}
& \Pr\{X(\mu,t_1+\tau)=a_1,...,X(\mu,t_k+\tau)=a_k\} \\
= \ & \Pr\{X(\mu,t_1+\tau)=a_1\}\Pr\{X(\mu,t_2+\tau)=a_2|X(\mu,t_1+\tau)=a_1\} \\
& \Pr\{X(\mu,t_3+\tau)=a_3|X(\mu,t_2+\tau)=a_2\}... \\
& \Pr\{X(\mu,t_k+\tau)=a_k|X(\mu,t_{k-1}+\tau)=a_{k-1}\}. \qquad \oplus
\end{aligned}$$

Since (1)

$$\begin{aligned}
& \Pr\{X(\mu,t_{j+1})=a_{j+1}|X(\mu,t_j)=a_j\} \\
= \ & \Pr\{X(\mu,t_{j+1}+\tau)=a_{j+1}|X(\mu,t_j+\tau)=a_j\} \\
= \ & \begin{cases} \frac{1}{2}(1+e^{-2\lambda(t_{j+1}-t_j)}), & \text{if} \quad a_j=a_{j+!} \\ \frac{1}{2}(1-e^{-2\lambda(t_{j+1}-t_j)}), & \text{if} \quad a_j\neq a_{j+!} \end{cases}
\end{aligned}$$

and (2)

$$\Pr\{X(\mu,t_1)=a_1\}=\Pr\{X(\mu,t_1+\tau)=a_1\},$$

we obtain $\oplus=\ominus$ if $\Pr\{X(\mu,0)=\pm 1\}=\frac{1}{2}$ (from (iii)). Thus, (a) is proved.

Next, look at the conditional probabilities

$$\Pr\{X(\mu,t)=a|X(\mu,0)=+1\}=\begin{cases} \frac{1}{2}(1+e^{-2\lambda t}), & \text{if } a=1 \\ \frac{1}{2}(1-e^{-2\lambda t}), & \text{if } a=-1 \end{cases}$$

$$\Pr\{X(\mu,t)=a|X(\mu,0)=-1\} = \begin{cases} \frac{1}{2}(1+e^{-2\lambda t}), & \text{if } a=-1 \\ \frac{1}{2}(1-e^{-2\lambda t}), & \text{if } a=1 \end{cases}.$$

As $t\to\infty$, $\Pr\{X(\mu,t)=a|X(\mu,0)=\pm 1\}\to\frac{1}{2}$ for $a=\pm 1$, and thus

$$\Pr\{X(\mu,t)=a\} = \sum_{b=\pm 1}\Pr\{X(\mu,t)=a|X(\mu,0)=b\}\Pr\{X(\mu,0)=b\}$$
$$\to \frac{1}{2}\sum_{b=\pm 1}\Pr\{X(\mu,0)=b\} = \frac{1}{2}.$$

Therefore, (b) is proved.

6.3 Another Example Process Derived From Poisson Processes–Shot Noise Process

- Let $N(\mu,t)$ be a Poisson process and $S_i(\mu)$ be the arrival time of the ith event. Then,

  $$N(\mu,t)=\sum_{i=1}^{\infty}u(t-S_i(\mu))$$

  with $N(\mu,0)=0$ and $u(t)$ a unit step function, i.e., $u(t)=1$ if $t\geq 0$ and $u(t)=0$ otherwise.

  Now, defining the random process

  $$Z(\mu,t)=\sum_{i=1}^{\infty}\delta(t-S_i(\mu))$$

  with $\delta(t)$ being the Dirac delta function, i.e., $u(t)=\int_{-\infty}^{t}\delta(x)dx$. Then, $N(\mu,t)$ can be regarded as the result of integrating $Z(\mu,t)$, which is a train of Dirac delta functions. That is,

  $$N(\mu,t)=\int_0^t Z(\mu,x)dx.$$

  Now, feeding $Z(\mu,t)$ into an LTI system with impulse response $h(t)$, we can express the output process by

  $$X(\mu,t)=\sum_{i=1}^{\infty}h(t-S_i(\mu)).$$

  Such a process $X(\mu,t)$ is called a shot noise process.

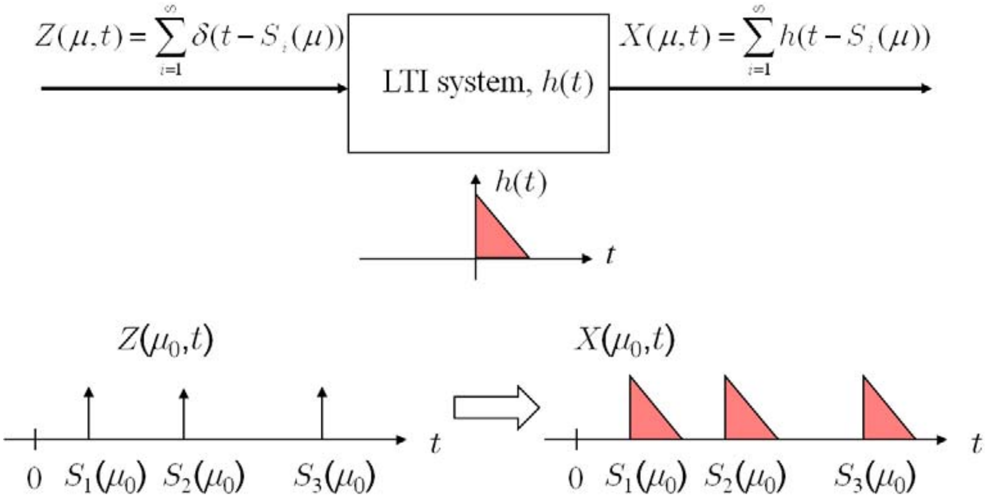


Figure 46:

- We are now interested in finding $E\{X(\mu,t)\}$:

$$\begin{aligned}
E\{X(\mu,t)\} &= E\{E\{X(\mu,t)|N(\mu,t)\}\} \\
&= E\{E\{\sum_{i=1}^{N(\mu,t)} h(t-S_i(\mu))|N(\mu,t)\}\} \\
&= E\{\sum_{i=1}^{N(\mu,t)} E\{h(t-S_i(\mu))|N(\mu,t)\}\}.
\end{aligned}$$

Now, let us look at $E\{h(t-S_i(\mu))|N(\mu,t)\}$: First, given that only one arrival occurred in $[0,t]$, the arrival time of the single customer $S_1(\mu)$ has the distribution function

$$\begin{aligned}
\Pr\{S_1(\mu)\leq x|N(\mu,t)=1\} &= \Pr\{N(\mu,x)=1|N(\mu,t)=1\} \\
&= \frac{\Pr\{N(\mu,x)=1,N(\mu,t)=1\}}{\Pr\{N(\mu,t)=1\}} \\
&= \frac{\Pr\{N(\mu,x)=1\}\Pr\{N(\mu,t-x)=0\}}{\Pr\{N(\mu,t)=1\}} \\
&= \frac{\lambda x\exp\{-\lambda x\}\cdot\exp\{-\lambda(t-x)\}}{\lambda t\exp\{-\lambda t\}} \\
&= \frac{x}{t}
\end{aligned}$$

for $0 \le x \le t$. Note that $S_1(\mu)$ is uniformly distributed in $[0, t]$, given the event $N(\mu, t) = 1$. Next, given that there are k arrivals in $[0, t]$, the arrival time for the j-th arrival $S_j(\mu)$ with $j \in \{1, 2, ..., k\}$ has the distribution function

$$\begin{aligned}
\Pr\{S_j(\mu) \le x | N(\mu, t) = k\} &= \Pr\{N(\mu, x) \ge j | N(\mu, t) = k\} \\
&= \frac{\Pr\{N(\mu, x) \ge j, N(\mu, t) = k\}}{\Pr\{N(\mu, t) = k\}} \\
&= \sum_{l=j}^{k} \frac{\Pr\{N(\mu, x) = l, N(\mu, t) = k\}}{\Pr\{N(\mu, t) = k\}} \\
&= \sum_{l=j}^{k} \frac{\Pr\{N(\mu, x) = l\} \Pr\{N(\mu, t - x) = k - l\}}{\Pr\{N(\mu, t) = k\}} \\
&= \sum_{l=j}^{k} \frac{\frac{(\lambda x)^l}{l!} \exp\{-\lambda x\} \cdot \frac{(\lambda(t-x))^{k-l}}{(k-l)!} \exp\{-\lambda(t - x)\}}{\frac{(\lambda t)^k}{k!} \exp\{-\lambda t\}} \\
&= \sum_{l=j}^{k} \binom{k}{l} (\frac{x}{t})^l (1 - \frac{x}{t})^{k-l}
\end{aligned}$$

for $0 \le x \le t$. This indicates that every individual arrival time is uniformly distributed in $[0, t]$ and independent of the other arrival times, given the event $N(\mu, t) = k$. Thus,

$$\begin{aligned}
E\{h(t - S_i(\mu)) | N(\mu, t)\} &= \int_0^t h(t - s) \cdot \frac{1}{t} ds \\
&= \frac{1}{t} \int_0^t h(s) ds.
\end{aligned}$$

This yields

$$
\begin{aligned}
E\{X(\mu,t)\} &= E\{\sum_{i=1}^{N(\mu,t)} E\{h(t-S_i(\mu))|N(\mu,t)\}\} \\
&= E\{\sum_{i=1}^{N(\mu,t)} \frac{1}{t}\int_0^t h(s)ds\} \\
&= E\{N(\mu,t)\cdot\frac{1}{t}\int_0^t h(s)ds\} \\
&= E\{N(\mu,t)\}\cdot\frac{1}{t}\int_0^t h(s)ds \\
&= \lambda\int_0^t h(s)ds.
\end{aligned}
$$

As $t \to \infty$, $E\{X(\mu,t)\}$ approaches to a constant if $\int_0^\infty h(s)ds$ is finite.

6.4 Renewal Processes

- Defn: A renewal process $\{N(\mu,t); t \geq 0\}$ is a counting process where the interarrival times $\{T_k(\mu); k \geq 1\}$ are iid random variables.

  Notes:

  (1) Poisson process is a renewal process with interarrival times being iid exponentials.

  (2) There are other types of renewal processes. For example, counting processes in independent Bernoulli trials are renewal processes.

- Recall:

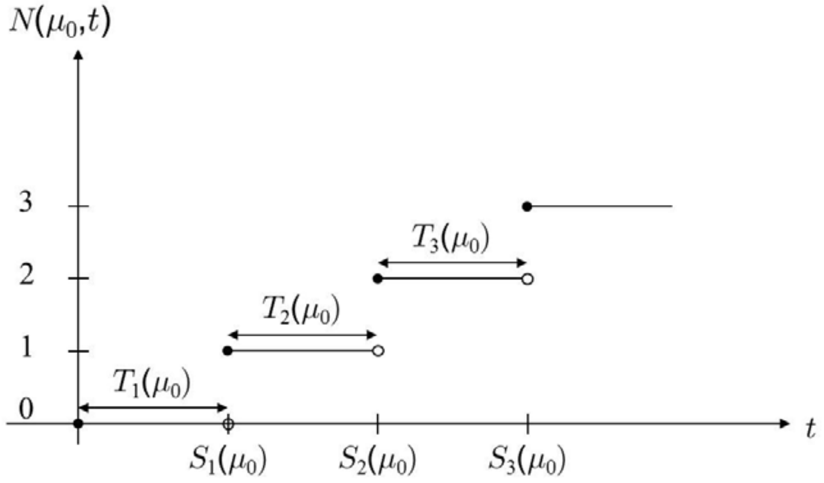


Figure 47:

$S_n(\mu) = \sum_{k=1}^{n} T_k(\mu)$ is called the n-th arrival time.

$F(x) \triangleq \Pr\{T_k(\mu) \leq x\}$ is the interarrival distribution where $F(0) \triangleq \Pr\{T_k(\mu) \leq 0\} = 0$ is assumed.

Note also that $\{N(\mu,t) \geq n\} = \{S_n(\mu) \leq t\}$.

The distribution of $S_2(\mu)$ is

$$\Pr\{S_2(\mu) \leq x\} = \Pr\{T_1(\mu) + T_2(\mu) \leq x\} = (F * F)(x)$$

where the operator $*$ is specified only in this subsection as

$$\begin{aligned} (F * G)(x) &= \int_{-\infty}^{\infty} F(x-\tau) g(\tau) d\tau \\ &= \int_{-\infty}^{\infty} F(x-\tau) dG(\tau) \end{aligned}$$

with $g(x) = \frac{dG(x)}{dx}$. If we set $F_1 = F$, $F_2 = F * F$, and $F_n = F_{n-1} * F$, then $F_n(x)$ is the distribution of $S_n(\mu)$, i.e.,

$$\begin{aligned} \Pr\{S_n(\mu) \leq x\} &= F_n(x) = \Pr\{N(\mu,x) \geq n\} \\ \Rightarrow \Pr\{N(\mu,x) = n\} &= \Pr\{N(\mu,x) \geq n\} - \Pr\{N(\mu,x) \geq n+1\} = F_n(x) - F_{n+1}(x). \end{aligned}$$

Note: $T_k(\mu)$'s are iid with identical marginal density $f(x) = \frac{d}{dx}F(x)$. Thus, the density of $S_n(\mu)$ is given by convolving n $f(x)$'s and results in the distribution

$$\begin{aligned}
F_n(x) &= \int_{-\infty}^{x}\int_{-\infty}^{\infty}\frac{d}{dy}F_{n-1}(y)|_{y=z-\tau}\cdot f(\tau)d\tau dz \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{x}\frac{d}{dy}F_{n-1}(y)|_{y=z-\tau}dz\cdot f(\tau)d\tau \\
&= \int_{-\infty}^{\infty}F_{n-1}(x-\tau)f(\tau)d\tau \\
&= F_{n-1} * F(x).
\end{aligned}$$

The mean of $N(\mu, t)$ can be thus obtained as

$$\begin{aligned}
m(x) &\triangleq E\{N(\mu, x)\} \\
&= \sum_{n=1}^{\infty} n\Pr\{N(\mu, x) = n\} \\
&= \sum_{n=1}^{\infty} n\Pr\{N(\mu, x) \geq n\} - \sum_{n=1}^{\infty} n\Pr\{N(\mu, x) \geq n+1\} \\
&= \sum_{n=1}^{\infty} n\Pr\{N(\mu, x) \geq n\} - \sum_{l=2}^{\infty}(l-1)\Pr\{N(\mu, x) \geq l\} \\
&= \sum_{n=1}^{\infty}\Pr\{N(\mu, x) \geq n\} \\
\Rightarrow m(x) &= \sum_{n=1}^{\infty}F_n(x).
\end{aligned}$$

- Proposition: $m(t) = F(t) + (m * F)(t)$ if $F(0) = 0$.

  Proof:

$$\begin{aligned}
E\{N(\mu, t)|T_1(\mu) = x\} &= \begin{cases} 0, & \text{if } x > t \\ 1 + E\{N(\mu, t-x)\}, & \text{if } x \leq t \end{cases} \\
\Rightarrow E\{N(\mu, t)|T_1(\mu)\} &= (1 + m(t - T_1(\mu)))\cdot 1_{T_1(\mu)\leq t}
\end{aligned}$$

where $1_{condition} = 1$ if condition is met and it is zero otherwise.

Now,

$$\begin{aligned}
m(t) &= E\{N(\mu,t)\} \\
&= E\{E\{N(\mu,t)|T_1(\mu)\}\} \\
&= E\{(1+m(t-T_1(\mu)))\cdot 1_{T_1(\mu)\le t}\} \\
&= \int_0^\infty (1+m(t-x))\cdot 1_{x\le t}\; dF(x) \\
&= \int_0^t (1+m(t-x))\; dF(x) \\
&= F(t) + \underbrace{\int_0^t m(t-x)dF(x)}_{(m*F)(t)} \\
&= F(t) + (m*F)(t).
\end{aligned}$$

Q.E.D.

Note: $m(t) = F(t) + (m*F)(t)$ is called the renewal equation.

7 Markov Processes and Markov Chains

- Materials from A. Leon-Garcia, Probability and Random Processes for Electrical Engineering, and from H. Larson and B. Shubert, Probabilistic Models in Engineering Sciences.

7.1 Markov Processes

- Defn: A process $X(\mu, t)$ is Markov iff for any sequence of times $t_1 < t_2 < \cdots < t_k < t_{k+1}$,

$$\begin{aligned} & f_{X(\mu,t_{k+1})|X(\mu,t_1),X(\mu,t_2),\ldots,X(\mu,t_k)}(x_{k+1}|x_1, x_2, \ldots, x_k) \\ = \ & f_{\underset{\underset{future}{\uparrow}}{t_{k+1}} | \underbrace{t_1, t_2, \ldots,}_{\underset{past}{\uparrow}} \underset{\underset{present}{\uparrow}}{t_k}}(x_{k+1}|x_1, x_2, \ldots, x_k) \\ = \ & f_{t_{k+1}|t_k}(x_{k+1}|x_k) \end{aligned}$$

for all t_i's and all x_i's, i.e., a random process whose past has no influence on the future if its present is specified.

For all densities,

$$\begin{aligned} & f_{t_0,t_1,\ldots,t_k}(x_0, x_1, \ldots, x_k) \\ = \ & f_{t_0}(x_0) f_{t_1|t_0}(x_1|x_0) f_{t_2|t_0,t_1}(x_2|x_0, x_1) \cdots \\ & f_{t_k|t_0,t_1,\ldots,t_{k-1}}(x_k|x_0, x_1, \ldots, x_{k-1}). \end{aligned}$$

Furthermore, if the process is Markovian, we can simplify it to

$$\begin{aligned} & f_{t_0,t_1,\ldots,t_k}(x_0, x_1, \ldots, x_k) \\ = \ & f_{t_0}(x_0) f_{t_1|t_0}(x_1|x_0) f_{t_2|t_1}(x_2|x_1) f_{t_3|t_2}(x_3|x_2) \cdots \\ & f_{t_k|t_{k-1}}(x_k|x_{k-1}) \\ = \ & f_{t_0}(x_0) \prod_{n=1}^{k} f_{t_n|t_{n-1}}(x_n|x_{n-1}) \end{aligned}$$

where $f_{t|t'}(x|x'), t' < t$, is called the transition probability density of the process.

- Theorem: Transition probability density for a Markov process must satisfy the Chapmom-Kolmogorov equation

$$f_{t_3|t_1}(x_3|x_1) = \int_{-\infty}^{\infty} f_{t_3|t_2}(x_3|x_2) f_{t_2|t_1}(x_2|x_1) dx_2$$

where $t_1 < t_2 < t_3$, for all t_1, t_2, t_3, x_1, x_3.

Note: The equation gives a necessary condition for Markov processes.

Proof:

$$
\begin{aligned}
& f_{t_2,t_3|t_1}(x_2, x_3|x_1) \\
= \; & f_{t_3|t_1,t_2}(x_3|x_1, x_2) f_{t_2|t_1}(x_2|x_1) \\
= \; & f_{t_3|t_2}(x_3|x_2) f_{t_2|t_1}(x_2|x_1)
\end{aligned}
$$

So,

$$
f_{t_3|t_1}(x_3|x_1) = \int_{-\infty}^{\infty} f_{t_3|t_2}(x_3|x_2) f_{t_2|t_1}(x_2|x_1) dx_2.
$$

Q.E.D.

- Theorem: Every process with independent transitions (increments or decrements) is Markov.

  Note: This gives a sufficient condition for Markov processes.

  Proof:

$$
\begin{aligned}
& f_{t_k|t_1,t_2,\ldots,t_{k-1}}(x_k|x_1,x_2,\ldots,x_{k-1}) \\
= \; & f_{X(\mu,t_k)|X(\mu,t_1),X(\mu,t_2),\ldots,X(\mu,t_{k-1})}(x_k|x_1,x_2,\ldots,x_{k-1}) \\
= \; & f_{Z_k(\mu)|Z_{k-1}(\mu),\ldots,Z_2(\mu),X(\mu,t_{k-1})}(z_k = x_k - x_{k-1}| \\
& z_{k-1} = x_{k-1} - x_{k-2}, \ldots, z_2 = x_2 - x_1, x_{k-1}) \\
= \; & f_{Z_k(\mu)|X(\mu,t_{k-1})}(z_k = x_k - x_{k-1}|x_{k-1}) \\
= \; & f_{X(\mu,t_k)|X(\mu,t_{k-1})}(x_k|x_{k-1}) \\
= \; & f_{t_k|t_{k-1}}(x_k|x_{k-1})
\end{aligned}
$$

  which is Markovian. Q.E.D.

  Notes: (1) The converse is not true, i.e., not every Markov process has independent transitions. (2) Poisson processes are Markovian since they have independent increments.

- Properties of Markov Process

  1. $E\{X(\mu,t_n)|X(\mu,t_1), X(\mu,t_2), \ldots, X(\mu,t_{n-1})\} = E\{X(\mu,t_n)|X(\mu,t_{n-1})\}$ if $t_1 < t_2 < \ldots < t_n$.
  2. A Markov Process is also Markov if time is reversed, i.e.,

$$
f_{t_k|t_{k+1},t_{k+2},\ldots,t_{k+n}}(x_k|x_{k+1}, x_{k+2}, \ldots, x_{k+n}) = f_{t_k|t_{k+1}}(x_k|x_{k+1})
$$

     if $t_k < t_{k+1} < \ldots < t_{k+n}$.

Proof:

$$\begin{aligned}
& f_{t_k|t_{k+1},t_{k+2},\ldots,t_{k+n}}(x_k|x_{k+1},x_{k+2},\ldots,x_{k+n}) \\
= & \frac{f_{t_k,t_{k+1},t_{k+2},\ldots,t_{k+n}}(x_k,x_{k+1},x_{k+2},\ldots,x_{k+n})}{f_{t_{k+1},t_{k+2},\ldots,t_{k+n}}(x_{k+1},x_{k+2},\ldots,x_{k+n})} \\
= & \frac{f_{t_{k+n}|t_{k+n-1}}(x_{k+n}|x_{k+n-1})f_{t_{k+n-1}|t_{k+n-2}}(x_{k+n-1}|x_{k+n-2})\cdots}{f_{t_{k+n}|t_{k+n-1}}(x_{k+n}|x_{k+n-1})f_{t_{k+n-1}|t_{k+n-2}}(x_{k+n-1}|x_{k+n-2})\cdots} \\
& \frac{f_{t_{k+2}|t_{k+1}}(x_{k+2}|x_{k+1})f_{t_{k+1}|t_k}(x_{k+1}|x_k)f_{t_k}(x_k)}{f_{t_{k+2}|t_{k+1}}(x_{k+2}|x_{k+1})f_{t_{k+1}}(x_{k+1})} \\
= & \frac{f_{t_{k+1}|t_k}(x_{k+1}|x_k)f_{t_k}(x_k)}{f_{t_{k+1}}(x_{k+1})} \\
= & f_{t_k|t_{k+1}}(x_k|x_{k+1}).
\end{aligned}$$

Q.E.D.

Note:

$$f_{t_k,t_{k+1},t_{k+2},\ldots,t_{k+n}}(x_k,x_{k+1},x_{k+2},\ldots,x_{k+n}) = f_{t_{k+n}}(x_{k+n})\cdot\prod_{i=1}^{n} f_{t_{k+i-1}|t_{k+i}}(x_{k+i-1}|x_{k+i}).$$

3. $f_{t_3,t_1|t_2}(x_3,x_1|x_2) = f_{t_3|t_2}(x_3|x_2)f_{t_1|t_2}(x_1|x_2)$ if $\underset{\substack{\uparrow \\ \text{past}}}{t_1} < \underset{\substack{\uparrow \\ \text{present}}}{t_2} < \underset{\substack{\uparrow \\ \text{future}}}{t_3}$,
   i.e., if the present is specified, the past is independent of the future.
   Proof:

$$\begin{aligned}
& f_{t_3,t_1|t_2}(x_3,x_1|x_2) \\
= & \frac{f_{t_3|t_2}(x_3|x_2)f_{t_2|t_1}(x_2|x_1)f_{t_1}(x_1)}{f_{t_2}(x_2)} \\
= & f_{t_3|t_2}(x_3|x_2)f_{t_1|t_2}(x_1|x_2).
\end{aligned}$$

   Q.E.D.

4. $f_{t_2|t_3,t_1}(x_2|x_3,x_1) = \frac{f_{t_1|t_2}(x_1|x_2)}{f_{t_1|t_3}(x_1|x_3)} f_{t_2|t_3}(x_2|x_3)$.
   Proof: You show it!

- Note that the n-th order statistic of a Markov process is completely specified by
  $f_{t_n|t_{n-1}}(x_n|x_{n-1}),\ldots,f_{t_2|t_1}(x_2|x_1), f_{t_1}(x_1)$ or $f_{t_1|t_2}(x_1|x_2), f_{t_2|t_3}(x_2|x_3),\ldots,$
  $f_{t_{n-1}|t_n}(x_n|x_{n-1}), f_{t_n}(x_n)$.
  If a Markov process is stationary, then $f_{t_1+\tau|t_1}(x_1|x_2) = f_{t_2+\tau|t_2}(x_1|x_2)$, $\forall t_1,t_2,\tau$ and $f_{t_n}(x) = f_{t_m}(x)$, $\forall t_m,t_n$. Therefore, if $t_1,t_2,\ldots,t_n$ are equally spaced, then a stationary Markov process can be completely specified by $f_{t_2|t_1}(x_2|x_1)$ and $f_{t_1}(x_1)$.

- Defn: A Markov Process is called homogeneous iff $f_{t_1+\tau|t_1}(x_1|x_2) = f_{t_2+\tau|t_2}(x_1|x_2)$, $\forall t_1, t_2, \tau$, but $f_{t_n}(x)$ may depend on t_n.

  In general, a homogeneous process is not stationary, although in many cases a homogeneous process tends to be stationary as $t_n \to \infty$. For example, the random telegraph signal $X(\mu, t)$ with $\Pr\{X(\mu, 0) = \pm 1\} \neq \frac{1}{2}$ is homogeneous and tends to be stationary as $t_n \to \infty$.

- Examples of Markov Processes

  1. A Poisson process is Markov since it has independent and stationary increments.
  2. A process $w(\mu, t)$, $t \geq 0$, with independent transitions such that $w(\mu, 0) = 0$ and in each $0 \leq t' \leq t$, the increment $w(\mu, t) - w(\mu, t')$ has a Gaussian pdf with zero mean, i.e.,

     $$E\{w(\mu, t) - w(\mu, t')\} = 0$$

     and variance

     $$Var\{w(\mu, t) - w(\mu, t')\} = t - t' \geq 0$$

     is called (standard) Wienner process. It is a Markov Process. Wienner process is nonstationary since its variance goes up with t. In fact, we can define $w(\mu, t)$ by

     $$w(\mu, t) = \int_0^t n(\mu, t')dt'$$

     where $n(\mu, t)$ is a zero-mean, unit-variance white Gaussian noise. Note: $w(\mu, t)$ can not be differentiated since no smooth sample path exists in the process. But, its difference is well-defined.

7.2 Discrete-Time Markov Chains

Reference: Alberto Leon-Garcia, "Probability and Random Processes for Electrical Engineering," Chap. 8.

- Defn: An integer-valued Markov process is called a Markov chain.

  Ex: Poisson process is a Markov chain.

- Now, consider the discrete-time integer-valued Markov random process $X_n(\mu)$, $n = 0, 1, 2, ...$ and denote its state probability by

$$p_j(n) \stackrel{\triangle}{=} \Pr\{X_n(\mu) = j\},\ j = 0, 1, 2, \ldots$$

Let us assume that $X_n(\mu)$ is homogeneous and denote its one-step state transition probability by

$$p_{ij} \stackrel{\triangle}{=} \Pr\{X_{n+1}(\mu) = j | X_n(\mu) = i\},\ \forall n.$$

Then, the $(n+1)$-th order state probability for $X_0(\mu), X_1(\mu), \ldots, X_n(\mu)$ is

$$\begin{aligned} & \Pr\{X_0(\mu) = i_0, X_1(\mu) = i_1, \ldots, X_n(\mu) = i_n\} \\ = \ & p_{i_0}(0) p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}. \end{aligned}$$

It can be shown that $X_n(\mu)$ is completely statistically specified by the initial state probabilities $\{p_i(0)\}$ and the matrix of one-step state transition probabilities $\mathbb{P}$

$$\mathbb{P} = \begin{bmatrix} p_{00} & p_{01} & p_{02} & p_{03} & \cdots \\ p_{10} & p_{11} & p_{12} & p_{13} & \cdots \\ p_{20} & p_{21} & p_{22} & p_{23} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Note that $\sum\limits_j \Pr\{X_{n+1}(\mu) = j | X_n(\mu) = i\} = \sum\limits_j p_{ij} = 1$, i.e., each row in $\mathbb{P}$ must add to one.

- Ex: Homogeneous Markov Model for Packet Speech

  $1 - \alpha$: the probability of silence in the $(n+1)$-th packet, given that the n-th packet contains silence
  $1 - \beta$: the probability of speech activity in the $(n+1)$-th packet, given that the n-th packet contains speech activity

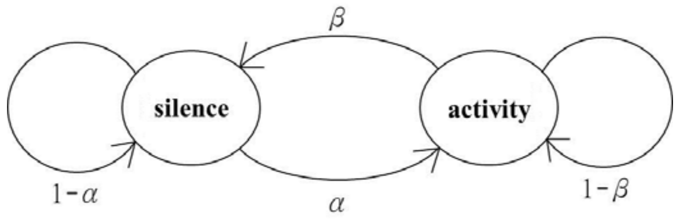


Figure 48:

$$\mathbb{P} = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}.$$

- Consider the discrete-time homogeneous Markov chain $X_n(\mu)$, $n = 0, 1, 2, \ldots$. Define the n-step state transition probability by

$$p_{ij}(n) \stackrel{\triangle}{=} \Pr\{X_{n+k}(\mu) = j | X_k(\mu) = i\} = \Pr\{X_n(\mu) = j | X_0(\mu) = i\}$$

because $X_n(\mu)$ is homogeneous. Now,

$$\begin{aligned} & \Pr\{X_2(\mu) = j, X_1(\mu) = k | X_0(\mu) = i\} \\ = & \Pr\{X_2(\mu) = j | X_1(\mu) = k\} \Pr\{X_1(\mu) = k | X_0(\mu) = i\} \\ = & p_{kj} p_{ik} \end{aligned}$$

$$\begin{aligned} \Rightarrow & \sum_k \Pr\{X_2(\mu) = j, X_1(\mu) = k | X_0(\mu) = i\} = \sum_k p_{ik} p_{kj} \\ \Rightarrow & p_{ij}(2) = \sum_k p_{ik}(1) p_{kj}(1) = \sum_k p_{ik} p_{kj} \qquad \circledast \end{aligned}$$

which is the Chapmom-Kolmogorov equation for discrete-time homogeneous Markov chains. Therefore, if we define the matrix of n-step state transition probabilities

$$\mathbb{P}(n) = [p_{ij}(n)]$$

then $\mathbb{P}(2) \underset{\uparrow \atop \circledast}{=} \mathbb{P}\mathbb{P} = \mathbb{P}^2$.

Following the same line of analysis as in $\circledast$, we can show that

$$\begin{aligned} \mathbb{P}(n) &= \mathbb{P}(n-1)\mathbb{P}(1) \\ &= \mathbb{P}(n-2)\mathbb{P}(1)\mathbb{P}(1) \\ &= \underbrace{\mathbb{P}(1) \cdots \mathbb{P}(1)}_{n} = \mathbb{P}^n. \end{aligned}$$

That is, the n-step state transition probability matrix is the n-th power of the one-step state transition probability matrix.

Define the state probability

$$p_j(n) \triangleq \Pr\{X_n(\mu) = j\}$$

which can be related to $p_j(n-1)$ by

$$\begin{aligned} p_j(n) &= \sum_i \Pr\{X_n(\mu) = j | X_{n-1}(\mu) = i\} \Pr\{X_{n-1}(\mu) = i\} \\ \Rightarrow p_j(n) &= \sum_i p_{ij} p_i(n-1). \end{aligned}$$

Now, if we define the vector of state probabilities at time n by the row vector

$$R(n) \triangleq [p_0(n), p_1(n), p_2(n), ...]$$

then $R(n)$ can be expressed as

$$\begin{aligned} R(n) &= R(n-1)\mathbb{P} \\ &= R(n-2)\mathbb{P}^2 \\ &= ... \\ &= R(0)\mathbb{P}^n. \end{aligned}$$

In other words, the vector of state probabilities at time n is obtained by multiplying the vector of initial state probabilities by $\mathbb{P}^n$.

- Example A: Let $X_1(\mu), X_2(\mu), ...$ be a random sequence where $X_n(\mu)$'s are independent and identically distributed and take the values $+1$ and -1 with $\Pr\{X_n(\mu) = 1\} = \Pr\{X_n(\mu) = -1\} = 1/2$. Define another sequence $Y_1(\mu), Y_2(\mu), ...$ by $Y_m(\mu) = \sum_{n=1}^{m} X_n(\mu)$. Show that the sequence $Y_1(\mu), Y_2(\mu), ...$ is a Markov process. Also. derive its state probability $p_\beta(m) = \Pr\{Y_m(\mu) = \beta\}$ and one-step state transition probability $p_{\alpha\beta}(m) = \Pr\{Y_m(\mu) = \beta | Y_{m-1}(\mu) = \alpha\}$.

  Now,

$$\begin{aligned} & \Pr\{Y_m(\mu) = y_m | Y_{m-1}(\mu) = y_{m-1}, ..., Y_1(\mu) = y_1\} \\ = \; & \Pr\{Y_m(\mu) = y_{m-1} + X_m(\mu) | Y_{m-1}(\mu) = y_{m-1}\} \\ & (X_n(\mu)\text{'s are independent}). \end{aligned}$$

Thus, the sequence is Markovian. Because $X_n(\mu)$'s are independent and identically distributed and take the values $+1$ and -1 with $\Pr\{X_n(\mu) = 1\} = \Pr\{X_n(\mu) = -1\} = 1/2$, $Y_m(\mu) = \sum_{n=1}^{m} X_n(\mu)$ has the moment generating function

$$\begin{aligned}
\Phi_m(s) &= \exp\{sY_m(\mu)\} \\
&= \exp\{s \sum_{n=1}^{m} X_n(\mu)\} \\
&= \prod_{n=1}^{m} \exp\{sX_n(\mu)\} \\
&\quad (X_n(\mu)\text{'s are independent}) \\
&= (\frac{1}{2})^m (\exp\{s\} + \exp\{-s\})^m \\
&\quad (X_n(\mu)\text{'s are identically distributed}) \\
&= \sum_{n=0}^{m} (\frac{1}{2})^m \binom{m}{n} \exp\{(m-2n)s\} \\
&= \sum_{n=0}^{m} \Pr\{Y_m(\mu) = m-2n\} \exp\{(m-2n)s\}.
\end{aligned}$$

Thus, the state probability $p_\beta(m)$ is given by

$$\begin{aligned}
p_\beta(m) &= \Pr\{Y_m(\mu) = \beta\} \\
&= \begin{cases} (\frac{1}{2})^m \binom{m}{n}, & \text{if } \beta = m-2n \text{ and } n \in \{0, 1, ..., m\} \\ 0, & \text{otherwise} \end{cases} .
\end{aligned}$$

where β takes value in the set $\{-m, -m+2, ..., 0\}$. Also, the one-step state transition probability $p_{\alpha\beta}(m) = \Pr\{Y_m(\mu) = \beta | Y_{m-1}(\mu) = \alpha\}$ is given by

$$\begin{aligned}
p_{\alpha\beta}(m) &= \Pr\{Y_m(\mu) = \beta | Y_{m-1}(\mu) = \alpha\} \\
&= \Pr\{X_m(\mu) = \beta - \alpha\} \\
&= \begin{cases} \frac{1}{2}, & \begin{array}{l} \text{if } \beta \in \{\alpha-1, \alpha+1\} \\ \text{and } \alpha = m-1-2n \\ \text{with } n \in \{0, 1, ..., m-1\} \end{array} \\ 0, & \text{otherwise} \end{cases} .
\end{aligned}$$

- Example B: Homogeneous Markov Model for Packet Speech With $\alpha, \beta > 0$

$$\mathbb{P} = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix} = \frac{1}{\alpha+\beta} \begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} + \frac{1-\alpha-\beta}{\alpha+\beta} \begin{bmatrix} \alpha & -\alpha \\ -\beta & \beta \end{bmatrix}$$

$\Rightarrow$

$$\begin{aligned}
\mathbb{P}^2 &= \begin{bmatrix} (1-\alpha)^2+\alpha\beta & \alpha(1-\alpha)+\alpha(1-\beta) \\ \beta(1-\alpha)+\beta(1-\beta) & \alpha\beta+(1-\beta) \end{bmatrix} \\
&= \frac{1}{\alpha+\beta}\begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} + \frac{(1-\alpha-\beta)^2}{\alpha+\beta}\begin{bmatrix} \alpha & -\alpha \\ -\beta & \beta \end{bmatrix}
\end{aligned}$$

$\Rightarrow$

$$R(2) = [p_0(2), p_1(2)] = [p_0(0), p_1(0)]\mathbb{P}^2 \neq R(0) \text{ in general.}$$

It can be shown that

$$\mathbb{P}^n = \frac{1}{\alpha+\beta}\begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} + \frac{(1-\alpha-\beta)^n}{\alpha+\beta}\begin{bmatrix} \alpha & -\alpha \\ -\beta & \beta \end{bmatrix}$$

$$\Rightarrow \lim_{n\to\infty}\mathbb{P}^n = \frac{1}{\alpha+\beta}\begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} \qquad \text{since } 0 < \alpha+\beta < 2.$$

Thus,

$$\begin{aligned}
\lim_{n\to\infty} R(n) &= R(0)\lim_{n\to\infty}\mathbb{P}^n \\
&= R(0)\begin{bmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\beta}{\alpha+\beta}, & \frac{\alpha}{\alpha+\beta} \end{bmatrix}.
\end{aligned}$$

We call such a system steady-state, i.e., a steady-state discrete-time homogeneous Markov chain has

$$\lim_{n\to\infty} p_j(n) = \pi_j \qquad \text{a constant for fixed j.}$$

Here, π_j is called the steady-state probability of state j. Defining $\pi = [\pi_0, \pi_1, \ldots]$, we can define

$$\lim_{n\to\infty} R(n) = \pi.$$

Notes:

1. If a discrete-time homogeneous Markov chain has steady state, then π can be obtained by

$$\begin{aligned}
& \lim_{n\to\infty} R(n+1) = \pi = \lim_{n\to\infty} R(n)\mathbb{P} = \pi\mathbb{P} \\
\Rightarrow\ & \pi = \pi\mathbb{P}.
\end{aligned}$$

2. A discrete-time stationary Markov chain has steady state.
3. Not every discrete-time Markov chain has steady state.

7.3 Continuous-Time Homogeneous Markov Chains

- Reference: Alberto Leon-Garcia, "Probability and Random Processes for Electrical Engineering."
- Consider the continuous-time homogeneous Markov chain $X(\mu, t)$, $t \geq 0$.
- Defns:

  1. State transition probability is denoted by

  $$p_{ij}(t) \stackrel{\triangle}{=} \Pr\{X(\mu, t+s) = j | X(\mu, s) = i\} \quad t \geq 0, \ \forall s.$$

  2. State probability is denoted by

  $$p_i(t) = \Pr\{X(\mu, t) = i\} \quad \forall t.$$

  3. The matrix of state transition probabilities in an interval of length t is denoted by

  $$\mathbb{P}(t) = [p_{ij}(t)] \quad \forall t.$$

- Examples:

  (a) Poisson Process

  $$\begin{aligned} p_{ij}(t) &= \Pr\{j - i \text{ arrivals in } t \text{ seconds}\} \\ &= p_{0\ j-i}(t) \\ &\quad \text{(stationary and independent increments,} \\ &\quad \text{and thus iid interarrival times)} \\ &= \Pr\{X(\mu, t) = j - i | X(\mu, 0) = 0\} \\ &= \Pr\{X(\mu, t) = j - i\} \\ &= e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!} \quad j \geq i. \end{aligned}$$

  $$\Rightarrow \mathbb{P}(t) = \begin{bmatrix} e^{-\lambda t} & (\lambda t) e^{-\lambda t} & \frac{(\lambda t)^2}{2!} e^{-\lambda t} & \frac{(\lambda t)^3}{3!} e^{-\lambda t} & \cdots \\ 0 & e^{-\lambda t} & (\lambda t) e^{-\lambda t} & \frac{(\lambda t)^2}{2!} e^{-\lambda t} & \cdots \\ 0 & 0 & e^{-\lambda t} & (\lambda t) e^{-\lambda t} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

(b) Random Telegraph Signal

$$
\begin{aligned}
p_{11}(t) &= p_{-1-1}(t) = \frac{1}{2}[1 + e^{-2\lambda t}] \\
p_{1-1}(t) &= p_{-11}(t) = \frac{1}{2}[1 - e^{-2\lambda t}]
\end{aligned}
$$

$$
\Rightarrow \mathbb{P}(t) = \begin{array}{c} \\ +1 \\ -1 \end{array} \begin{array}{c} \begin{array}{cc} +1 & -1 \end{array} \\ \left[\begin{array}{cc} \frac{1}{2}[1 + e^{-2\lambda t}] & \frac{1}{2}[1 - e^{-2\lambda t}] \\ \frac{1}{2}[1 - e^{-2\lambda t}] & \frac{1}{2}[1 + e^{-2\lambda t}] \end{array} \right] \end{array}.
$$

- Theorem: A continuous-time homogeneous Markov chain $X(\mu, t)$ remains at a given state for an exponentially distributed random time.

  Proof: Let $T_i(\mu)$ be the time, or state occupancy time, spent in state i. Then,

$$
\begin{array}{ll}
 & \Pr\{T_i(\mu) > t + s | T_i(\mu) > s\} \quad t, s \geq 0 \\
\text{(homogeneous)} & = \Pr\{T_i(\mu) > t + s | X(\mu, s') = i, 0 \leq s' \leq s\} \\
 & = \Pr\{X(\mu, t') = i, 0 \leq t' \leq t + s | X(\mu, s') = i, 0 \leq s' \leq s\} \\
\text{(Markov)} & = \Pr\{X(\mu, t') = i, s \leq t' \leq t + s | X(\mu, s) = i\} \\
\text{(homogeneous)} & = \Pr\{X(\mu, t'') = i, 0 \leq t'' \leq t | X(\mu, 0) = i\} \\
 & = \Pr\{T_i(\mu) > t\}
\end{array}
$$

  and thus

$$
\Pr\{T_i(\mu) > t + s | T_i(\mu) > s\} = \Pr\{T_i(\mu) > t\} \quad t, s \geq 0.
$$

  Only the exponential random variables satisfy this memoryless property. It follows from previous discussion that

$$
\Pr\{T_i(\mu) > t\} = e^{-\nu_i t}
$$

  where $(\nu_i)^{-1} = E\{T_i(\mu)\}$ is called the mean state occupancy time for state i. Q.E.D.

- Therefore, a continuous-time homogeneous Markov Chain can be viewed as follows:

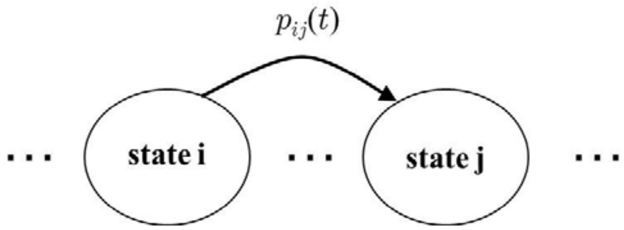


Figure 49:

or in an alternative way

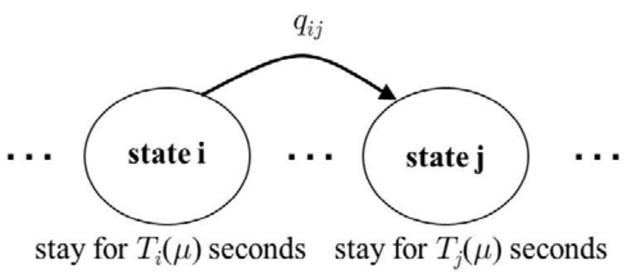


Figure 50:

That is:

1. Each time, a state, say i, is entered, an exponentially distributed state occupancy time $T_i(\mu)$ is selected.
2. When the time is up, the next state j is entered according to a discrete-time homogeneous Markov Chain with transition probability q_{ij}.
3. Then, the new state occupancy time, is selected according to $T_j(\mu)$, and so on.

Therefore,

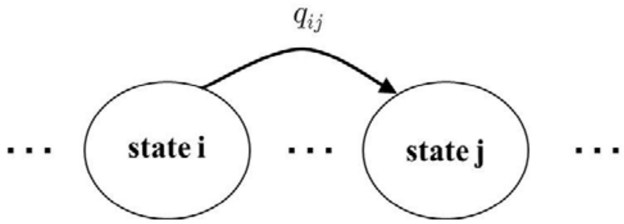


Figure 51:

is a discrete-time Markov Chain with transition probabilities q_{ij}'s. This Markov Chain is called an embedded Markov chain.

Note: $T_i(\mu)$ may be differently distributed from $T_j(\mu)$.

- Ex: Random Telegraph Signal

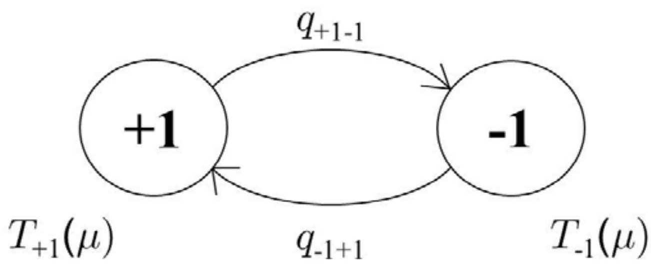


Figure 52:

with $q_{+1,-1} = q_{-1,+1} = 1$ and $q_{+1,+1} = q_{-1,-1} = 0$.

- Now, for small $\delta > 0$,

$$
\begin{aligned}
& \Pr\{T_i(\mu) > \delta\} = e^{-\nu_i \delta} \\
= \; & 1 - \nu_i \delta + \frac{(\nu_i \delta)^2}{2!} - \frac{(\nu_i \delta)^3}{3!} + \ldots \\
= \; & 1 - \nu_i \delta + o(\delta)
\end{aligned}
$$

where we say that the remainder term

$$R(\delta) = \frac{(\nu_i \delta)^2}{2!} - \frac{(\nu_i \delta)^3}{3!} + \cdots$$

is $o(\delta)$ since $\lim_{\delta\to 0}\frac{R(\delta)}{\delta}=0$, i.e., $R(\delta)$ is negligible relative to δ as $\delta\to 0$. Since $\Pr\{T_i(\mu)>\delta\}$ is the probability that the process $X(\mu,t)$ remains in state i during (and at the end of) the interval of length δ,

$$\begin{aligned} p_{ii}(\delta) &= \Pr\{T_i(\mu)>\delta\} \\ &= 1-\nu_i\delta+o(\delta) \end{aligned}$$

where $o(\delta)$ can be regarded as the probability of more than one transition from i to i in time δ. Thus, $1-p_{ii}(\delta)$ is the probability that the process $X(\mu,t)$ leaves state i during (and at the end of) the interval of length δ,

$$1-p_{ii}(\delta)=\nu_i\delta+o(\delta)\,.$$

This ν_i is called the rate (of probability flow) at which the process $X(\mu,t)$ leaves state i.

Further, since we use q_{ij} to denote the probability that, given $X(\mu,t)$ leaves state i, it will enter state j,

$$\begin{aligned} p_{ij}(\delta) &= (1-p_{ii}(\delta))\,q_{ij} \\ &= \nu_i q_{ij}\delta+o(\delta) \\ &= \gamma_{ij}\delta+o(\delta) \end{aligned}$$

for $i\neq j$. Here, we call $\gamma_{ij}\triangleq\nu_i q_{ij}$ for $i\neq j$ the rate (of probability flow) at which the process $X(\mu,t)$ enters state j from state i.
Note: We define $\gamma_{ii}=-\nu_i$ and $q_{ii}=0$ by default.
Next, for $\delta>0$, the state probability is

$$\begin{aligned} p_j(t+\delta) &= \Pr\{X(\mu,t+\delta)=j\} \\ &= \sum_i \Pr\{X(\mu,t+\delta)=j|X(\mu,t)=i\}\Pr\{X(\mu,t)=i\} \\ &= \sum_i p_{ij}(\delta)p_i(t). \end{aligned}$$

Subtracting $p_j(t)$ from both sides gives

$$p_j(t+\delta)-p_j(t)=\sum_{i\neq j}p_{ij}(\delta)p_i(t)+(p_{jj}(\delta)-1)p_j(t).$$

Dividing both sides by δ and then letting $\delta\to 0$, we have

$$p_j'(t) = \sum_{i\neq j}\underbrace{\lim_{\delta\to 0}\frac{p_{ij}(\delta)}{\delta}}_{=\lim_{\delta\to 0}\frac{\gamma_{ij}\delta+o(\delta)}{\delta}=\gamma_{ij}}p_i(t)+\underbrace{\lim_{\delta\to 0}\frac{p_{jj}(\delta)-1}{\delta}}_{=-\nu_j=\gamma_{jj}}p_j(t)$$

$$\Rightarrow p_j'(t) = \sum_i \gamma_{ij} p_i(t) \qquad \forall j \qquad \circledast$$

which is a form of Chapmom-Kolmogorov equations for continuous-time homogeneous Markov chains.
Note that given initial conditions $\{p_j(0)\}$ and $\{\gamma_{ij}\}$, the state probabilities $p_j(t)$'s can be found by solving $\circledast$.

- Ex: Simple Two-State Queuing System
  State 0: The system is idle and wait for a customer to arrive. The idle time $T_0(\mu)$ is an exponential random variable with mean $\frac{1}{\alpha}$.
  State 1: The system is busy serving a customer. The time in the busy state $T_1(\mu)$ is an exponential random variable with mean $\frac{1}{\beta}$.

  The transition state diagram is

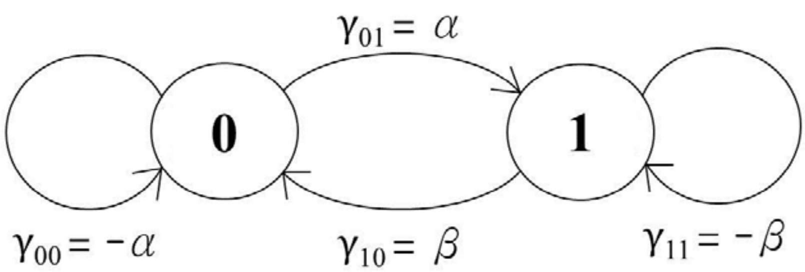


Figure 53:

From $\circledast$,

$$\begin{cases} p_0'(t) = -\alpha p_0(t) + \beta p_1(t) \\ p_1'(t) = \alpha p_0(t) - \beta p_1(t) \end{cases}.$$

Since $p_0(t) + p_1(t) = 1$,

$$\begin{cases} p_0'(t) = -\alpha p_0(t) + \beta(1 - p_0(t)) \\ p_1'(t) = \alpha p_0(t) - \beta p_1(t) \end{cases}.$$

Give $p_0(0) = p_0$, the first equation has the solution

$$p_0(t) = \frac{\beta}{\alpha+\beta} + (p_0 - \frac{\beta}{\alpha+\beta})e^{-(\alpha+\beta)t}.$$

This, in conjunction with $p_1(t) = 1 - p_0(t)$, gives

$$p_1(t) = \frac{\alpha}{\alpha+\beta} + (1 - p_0 - \frac{\alpha}{\alpha+\beta})e^{-(\alpha+\beta)t}.$$

Note that as $t \to \infty$,

$$p_0(t) \to \frac{\beta}{\alpha+\beta} \text{ and } p_1(t) \to \frac{\alpha}{\alpha+\beta} \qquad \oplus$$

i.e., the state probabilities approach constant values and are independent of the initial state probabilities.

- Another way to view the steady-state state probabilities:

  In the steady state, the system maintains in equilibrium; therefore, $p_j'(t) = 0$ as $t \to \infty$ and, thus, $\circledast$ becomes

  $$\sum_i \gamma_{ij} p_i(t) = 0. \qquad \ominus$$

  Thus, in the above example on a "Simple Two-State Queuing System", we have from $\ominus$ that

  $$\begin{cases} p_0(t)\gamma_{00} + p_1(t)\gamma_{10} = 0 \\ p_0(t)\gamma_{01} + p_1(t)\gamma_{11} = 0 \end{cases}.$$

  You can obtain the same steady-state state probability result as $\oplus$ by using $p_0(t) + p_1(t) = 1$.

- Steady-State State Probabilities

  As $t \to \infty$, $p_j(t) \to p_j$ (a constant) and $p_j'(t) \to 0$ for a system that settles into steady state or equilibrium. Thus, $\circledast$ becomes

  $$\begin{aligned} \sum_i \gamma_{ij} p_i &= 0 \qquad \forall j \\ \Rightarrow \nu_j p_j &= \sum_{i \neq j} \gamma_{ij} p_i \qquad \forall j. \ \circledast\circledast \end{aligned}$$

  Note that $1 = \sum_i p_{ji}(t)$. Thus, for small δ and t, both approaching zero,

  $$\begin{aligned} 1 &= \sum_{i \neq j} p_{ji}(t+\delta) + p_{jj}(t+\delta) \\ &= \sum_{i \neq j} (\gamma_{ji}(t+\delta) + o(t+\delta)) + (1 + \gamma_{jj}(t+\delta) + o(t+\delta)). \end{aligned}$$

Further, for small t,

$$\begin{aligned} 1 &= \sum_{i \neq j} p_{ji}(t) + p_{jj}(t) \\ &= \sum_{i \neq j} (\gamma_{ji} t + o(t)) + (1 + \gamma_{jj} t + o(t)). \end{aligned}$$

These yield

$$\lim_{\delta \to 0} \sum_i \frac{p_{ji}(t+\delta) - p_{ji}(t)}{\delta} = \sum_i p'_{ji}(t) = 0 = \sum_i \gamma_{ji}$$

(since $\lim\limits_{\delta \to 0} \frac{o(t+\delta)-o(t)}{\delta} = 0$ for small t) which results in

$$\begin{aligned} \sum_i \gamma_{ji} &= 0 \quad \forall j \\ \Rightarrow \quad -\gamma_{jj} &= \nu_j = \sum_{i \neq j} \gamma_{ji}. \end{aligned}$$

Therefore, we have from ⊛⊛ that

$$p_j \sum_{i \neq j} \gamma_{ji} = \sum_{i \neq j} \gamma_{ij} p_i \ \forall j.$$

This equation is called global balance equation. The equation states that at equilibrium (or in steady state) the weighted rate of probability flow out of state j, namely $\nu_j p_j = p_j \sum\limits_{i \neq j} \gamma_{ji}$, is equal to the accumulated weighted rate of probability flow into state j, namely $\sum\limits_{i \neq j} \gamma_{ij} p_i$.

The transition state diagram is

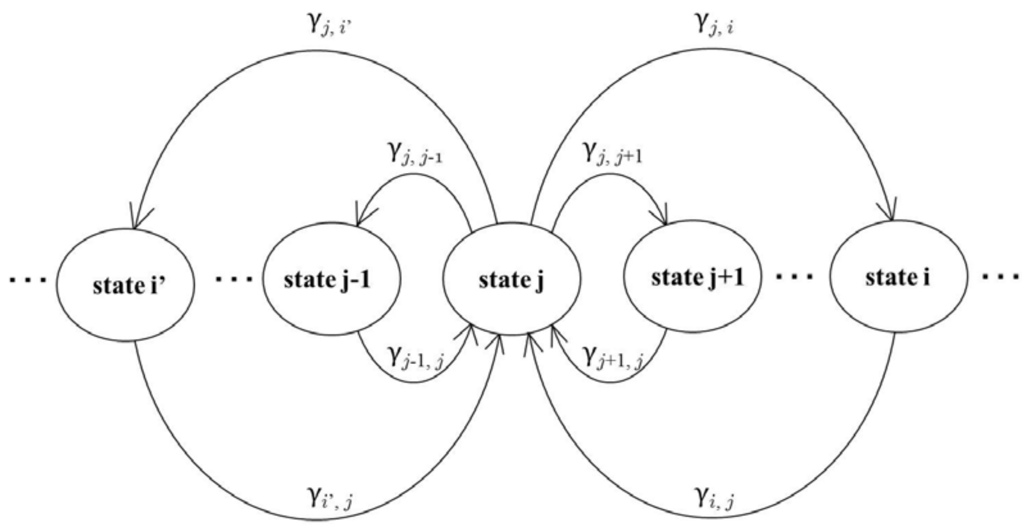


Figure 54:

Solving these equations will yield the steady-state state probabilities of the continuous-time homogeneous Markov chain, if it exists.

Now, if $p_i(0) = p_i$ for all i, then $p_i(t) = p_i$, $\forall t$, i.e., the resulting process is stationary.

- Example: $M/M/1$ Single-Server Queuing System

  Consider a queuing system in which customers are served one at a time in order of arrival. The times between customer arrivals are iid and exponentially distributed with rate λ. The times required to serve customers are iid and exponentially distributed with rate μ. Service times and interarrival times are independent. Such a queuing system is called the $M/M/1$ queue.

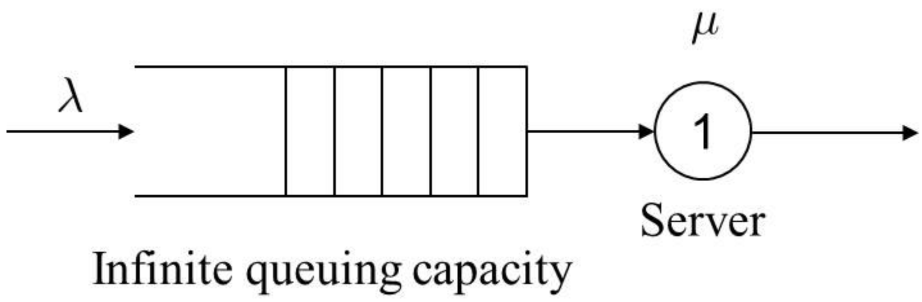
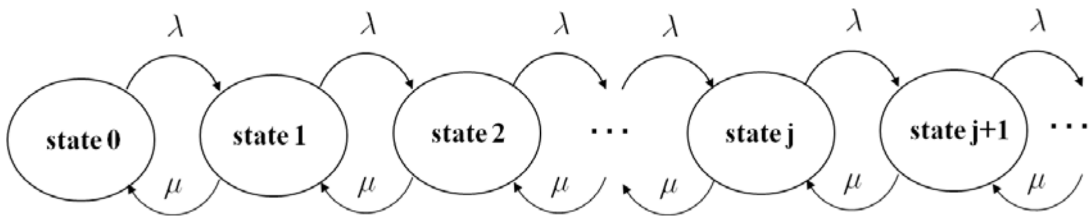


Figure 55:

Let us check its steady-state state probabilities:

Define $N(\mu, t) \equiv$ number of customers in the system at time t.

$N(\mu, t)$ has independent increments and decrements. Thus, $N(\mu, t)$ is a continuous-time Markov chain. Moreover, since service times are iid and interarrival times are iid, $N(\mu, t)$ is homogeneous.
Now, define state i as the state of i customers in the system. It has the transition state diagram given by



By global balance equations

$$
\begin{aligned}
\lambda p_0 &= \mu p_1 \\
(\lambda + \mu) p_j &= \lambda p_{j-1} + \mu p_{j+1} \qquad j = 1, 2, ... \\
\Rightarrow \lambda p_j - \mu p_{j+1} &= \lambda p_{j-1} - \mu p_j \\
&= ... \\
&= \lambda p_0 - \mu p_1 \\
&= 0 \\
\Rightarrow p_{j+1} &= \rho p_j \text{ with } \rho \triangleq \frac{\lambda}{\mu} \qquad j = 0, 1, ... \\
\Rightarrow p_{j+1} &= \rho^{j+1} p_0 \qquad j = 0, 1, ...
\end{aligned}
$$

Since $\sum_{j=0}^{\infty} p_j = 1$,

$$\begin{aligned} & p_0 \sum_{j=0}^{\infty} \rho^j = 1 \\ \Rightarrow \quad & p_0 \frac{1}{1-\rho} = 1 \text{ for } \rho < 1 \\ \Rightarrow \quad & p_0 = 1 - \rho \text{ and } p_j = \rho^j (1-\rho) \text{ for } \rho < 1. \end{aligned}$$

For $\rho \geq 1$, i.e., $\lambda \geq \mu$, the system can not settle into steady state. Thus, the customer arrival rate should be smaller than the service rate.

- Example: Birth and Death Process

  Define state i as the state that there are i persons alive, λ_i as the birth rate when there are i persons alive, assuming the times between births are exponentially distributed and mutually independent, and μ_i as the death rate when there are i persons alive, assuming the times between deaths are exponentially distributed and mutually independent. Assume further that death and birth processes are independent. This is a continuous-time Markov chain since $N(\mu, t)$ ($\equiv$ number of persons alive) has independent increments and decrements. Moreover, $N(\mu, t)$ is homogeneous because interbirth times and interdeath times are respectively iid. The transition state diagram is

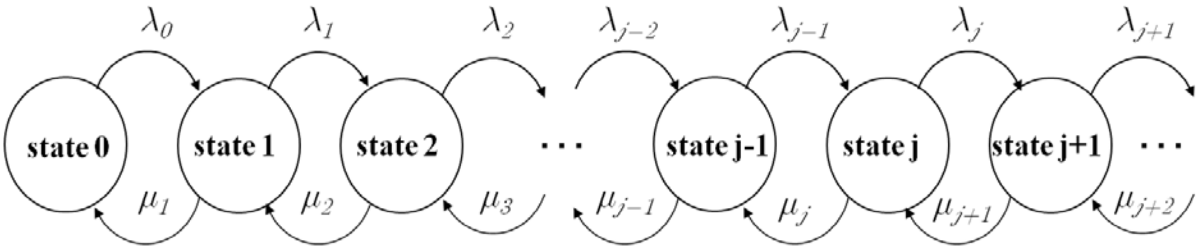


Figure 56:

By global balance equations

$$\begin{aligned}
\lambda_0 p_0 &= \mu_1 p_1 \\
(\lambda_j + \mu_j) p_j &= \lambda_{j-1} p_{j-1} + \mu_{j+1} p_{j+1} \qquad j = 1, 2, ... \\
\Rightarrow \lambda_j p_j - \mu_{j+1} p_{j+1} &= \lambda_{j-1} p_{j-1} - \mu_j p_j \\
&= ... \\
&= \lambda_0 p_0 - \mu_1 p_1 \\
&= 0
\end{aligned}$$

$$\Rightarrow \quad p_j = \frac{\lambda_{j-1}}{\mu_j} p_{j-1} \triangleq \rho_j p_{j-1} \text{ with } \rho_j \triangleq \frac{\lambda_{j-1}}{\mu_j} \qquad j = 1, 2, ...$$

$$\Rightarrow \quad p_j = \rho_j \rho_{j-1} ... \rho_1 p_0 = (\prod_{k=1}^{j} \rho_k) p_0 \qquad j = 1, 2, ...$$

Since $\sum\limits_{j=0}^{\infty} p_j = 1$,

$$p_0 (\sum_{j=0}^{\infty} \prod_{k=1}^{j} \rho_k) = 1$$

$$\Rightarrow \quad p_0 = (\sum_{j=0}^{\infty} \prod_{k=1}^{j} \rho_k)^{-1}$$

$$p_j = \frac{\prod_{k=1}^{j} \rho_k}{\sum\limits_{l=0}^{\infty} \prod_{m=1}^{l} \rho_m} \qquad j = 1, 2, ...$$

where by default $\prod_{m=1}^{0} \rho_m = 1$, if $\sum\limits_{l=0}^{\infty} \prod_{m=1}^{l} \rho_m > 1$. If $\sum\limits_{l=0}^{\infty} \prod_{m=1}^{l} \rho_m \leq 1$, the steady state dose not exist.

7.4 Long-Term Behavior of Markov Chains

- Reference: Alberto Leon-Garcia, "Probability and Random Processes for Electrical Engineering."
- Here, we consider Markov processes with independent transitions (increments or decrements) throughout and want to show under what condition a Markov chain has steady state.
- Let us consider discrete-time homogeneous Markov chains first:

- Defn: State j is said to be accessible from state i iff, for some $n \geq 0$, $p_{ij}(n) > 0$.

  We denote this case by $i \to j$. In the case, there is a sequence of transitions from i to j that has nonzero probability.

- Defn: States i and j communicate iff they are accessible to each other.

  We denote the case by $i \leftrightarrow j$.

  Notes:

  1. $i \leftrightarrow i$ since $p_{ii}(0) = 1$.
  2. If $i \leftrightarrow j$ and $j \leftrightarrow k$, then $i \leftrightarrow k$. This is because $p_{ij}(n) \neq 0$ and $p_{jk}(m) \neq 0$ for some n and m; thus at least $p_{ik}(n+m) \neq 0$.

- Defn: We say that two states i and j belong to the same class iff $i \leftrightarrow j$.

  This definition implies that two different classes of states must not have any state in common.

  Ex: Two classes.

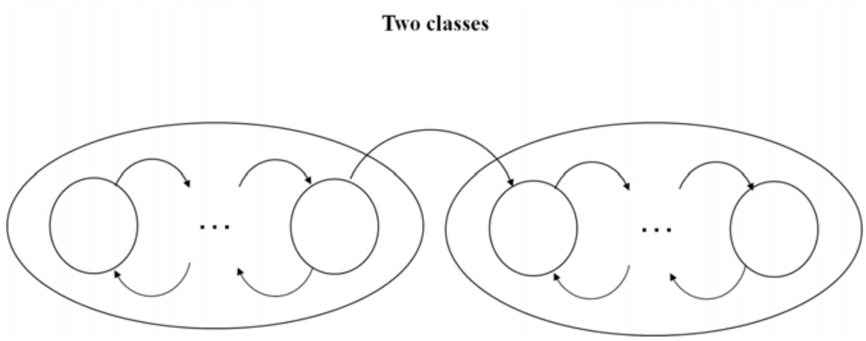


Figure 57:

The states of a Markov chain consist of one or more disjoint communication classes.

- Defn: A Markov chain that consists of a single class is called to be irreducible.

Ex: Three classes.

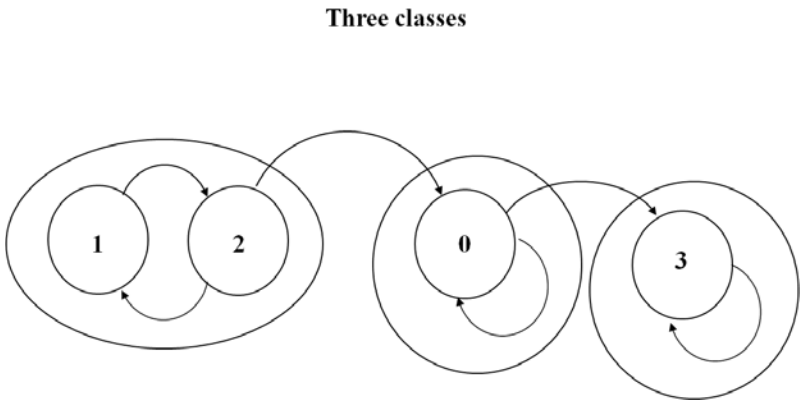


Figure 58:

Ex: One class. Irreducible and periodic Markov chain.

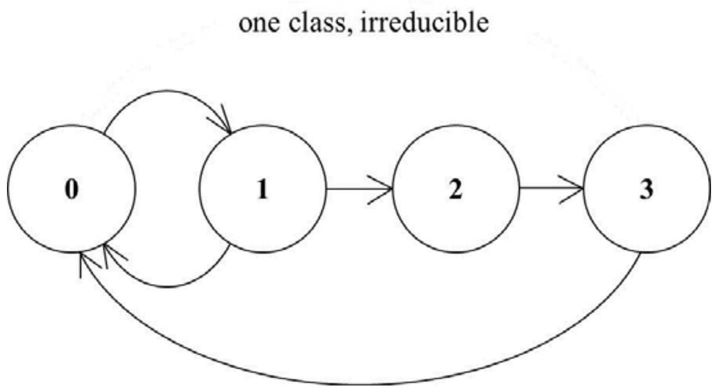


Figure 59:

Ex: Infinite classes.

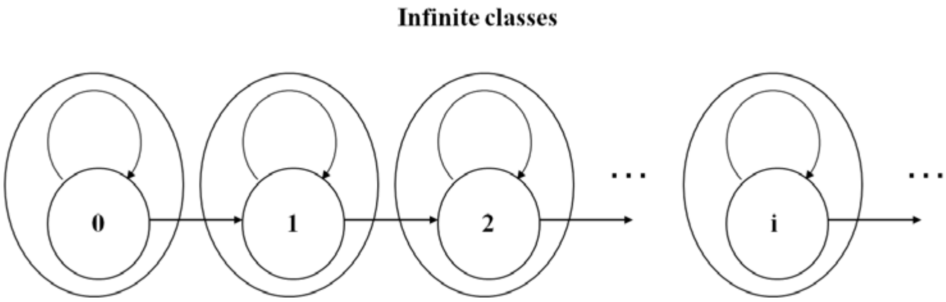


Figure 60:

Ex: One Class. Irreducible chain.

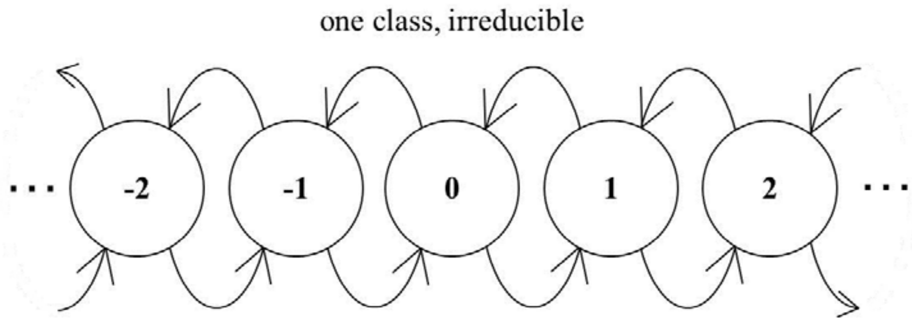


Figure 61:

- Defns:

    1. State i is said to be recurrent iff a Markov chains is started in state i, and the process returns to the state with probability one. That is,

    $$f_i \triangleq \Pr\{\underbrace{\text{ever returning to state } i}_{X_m(\mu)=i \text{ for some } m\neq 0} \mid X_0(\mu) = i\} = 1.$$

    2. State i is said to be transient iff $f_i < 1$.

- Note: Another way to define recurrence and transience is as follows.

    1. State i is recurrent iff "if we start the Markov chain in a recurrent state i, then the state reoccurs an infinite number of times".

2. State i is transient iff "if we start the Markov chain in a transient state i, then the state dose not reoccur after some finite number of returns".

- Now, defining the indicator function $1_{X_n(\mu)=i} = \begin{cases} 1, & \text{if } X_n(\mu) = i \\ 0, & \text{if } X_n(\mu) \neq i \end{cases}$, we can express the expected number of returns to state i as

$$\begin{aligned} & E\{\sum_{n=1}^{\infty} 1_{X_n(\mu)=i} \mid X_0(\mu) = i\} \\ = & \sum_{n=1}^{\infty} E\{1_{X_n(\mu)=i} \mid X_0(\mu) = i\} \\ = & \sum_{n=1}^{\infty} 1 \cdot \Pr\{X_n(\mu) = i \mid X_0(\mu) = i\} + 0 \cdot \Pr\{X_n(\mu) \neq i \mid X_0(\mu) = i\} \\ = & \sum_{n=1}^{\infty} p_{ii}(n). \end{aligned}$$

Note that a state i is recurrent iff it reoccurs an infinite number of times.

Lemma: State i is recurrent iff $\sum_{n=1}^{\infty} p_{ii}(n) = \infty$.

Lemma: State i is transient iff $\sum_{n=1}^{\infty} p_{ii}(n) < \infty$.

Ex:

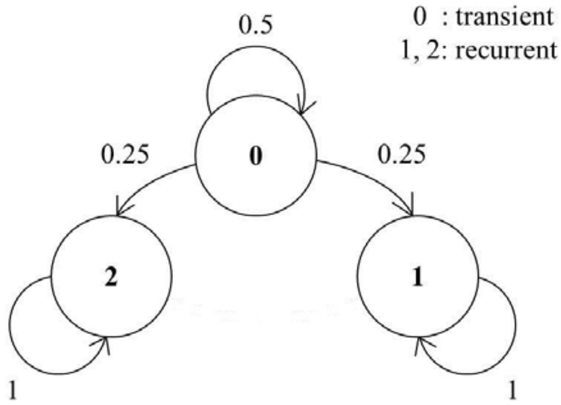


Figure 62:

$$
\begin{aligned}
\sum_{n=1}^{\infty} p_{00}(n) &= \sum_{n=1}^{\infty} (\frac{1}{2})^n = \frac{0.5}{1-0.5} = 1 < \infty \\
\sum_{n=1}^{\infty} p_{11}(n) &= \sum_{n=1}^{\infty} p_{22}(n) = \infty.
\end{aligned}
$$

State 0 is transient but states 1 and 2 are recurrent.

- Properties:

  1. If state i is recurrent and $i \leftrightarrow j$, then state j is also recurrent.
  2. If state i is transient and $i \leftrightarrow j$, then state j is transient.
     Note: Properties 1 and 2 say that both recurrence and transience are class properties.
  3. If a Markov chain is irreducible, then either all its states are transient or all its states are recurrent. If this chain has a finite number of states, then all these states are recurrent.

- Defn: State i is said to have period d iff $p_{ii}(n) = 0$ for all $n \neq d \cdot k$ with k a positive integer.

  Property: All the states in the same class have the same period.

- Defn: An irreducible Markov chain is said to be aperiodic if the states in its single class have period one.

  Ex:

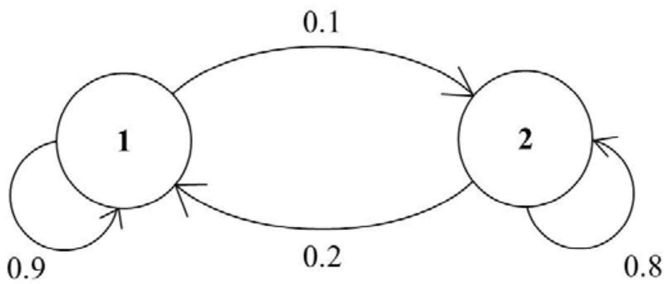


Figure 63:

States 1 and 2 have period one since $p_{ii}(n) > 0$ for $n = 1, 2, 3, ...$

Ex:

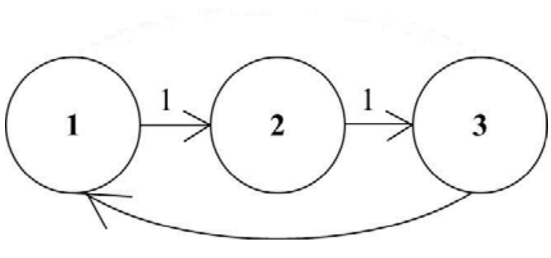


Figure 64:

States 1, 2 and 3 have period 3.

Ex:

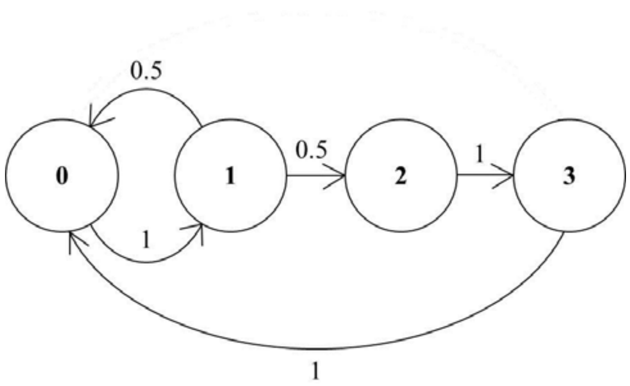


Figure 65:

States 0 and 1 can reoccur at times 2,4,6,8,...

States 2 and 3 can reoccur at times 4,6,8,10,...

Therefore, the Markov chain has period 2.

- Recall: Strong Law of Large Number

  Let $X_1(\mu), X_2(\mu), \dots$ be a sequence of iid random variables with finite common mean $E\{X(\mu)\} = \eta$ and finite common variance. Then, the corresponding sample mean sequence $Y_1(\mu), Y_2(\mu), \dots$ converges to η with probability one (or, almost everywhere), i.e.,

$$\Pr\{\lim_{n\to\infty} Y_n(\mu) = \eta\} = 1$$

  where $Y_n(\mu)$ is the sample mean defined by $Y_n(\mu) = \frac{1}{n}\sum_{k=1}^{n} X_k(\mu)$.

  Notes:

  1. For every outcome μ_0, the sample mean sequence $Y_1(\mu_0), Y_2(\mu_0), \dots$, which is a time average sequence of $X_n(\mu_0)$'s, i.e., $\frac{1}{1}\sum_{k=1}^{1} X_k(\mu_0)$, $\frac{1}{2}\sum_{k=1}^{2} X_k(\mu_0), \dots$, converges to η (ensemble average or statistical average, $E\{X(\mu)\}$) with probability one.
  2. Because convergence with probability one implies convergence in probability, the sample mean sequence $Y_1(\mu), Y_2(\mu), \dots$ converges to η in probability, i.e., for any $\epsilon > 0$,

$$\lim_{n\to\infty} \Pr\{|Y_n(\mu) - \eta| > \epsilon\} = 0.$$

     This is called the weak law of large number.

- Let us now look at the limiting probabilities for an irreducible, recurrent, homogeneous discrete-time Markov chain:

  Define $T_i(k)$ as the random time that elapses between $(k-1)$-th and k-th returns to state i, if we start a Markov chain in a recurrent state i.

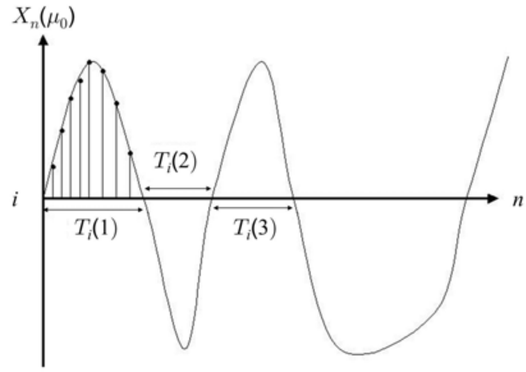


Figure 66:

For a recurrent irreducible homogeneous Markov chain, $T_i(k)$, $k = 1, 2, 3, ...$ are iid since each return time is independent of previous return times. Their common mean, denoted by $E\{T_i\}$, is called mean recurrence time. $\pi_i \triangleq \frac{1}{E\{T_i\}}$ is called the long-term proportion of time spent in state i. This can be shown by the law of large number as follows: Let $Z_n(\mu)$ be the indicator random variable defined by

$$Z_n(\mu) = \begin{cases} 1, & \text{if state } i \text{ occurs at time } n \\ 0, & \text{otherwise} \end{cases}.$$

Note that $Z_0(\mu) = 1$ by default. For $n = 1, 2, ...$, $Z_n(\mu)$'s are iid with $E\{Z_n(\mu)\} = E\{Z(\mu)\} = \frac{1}{E\{T_i\}}$ (See * below!). Then, the number of times that the Markov chain returns to state i up to time n is given by

$$Y_n(\mu) = \sum_{k=1}^{n} Z_k(\mu).$$

The proportion of time spent in state i up to time n is

$$\frac{Y_n(\mu)}{n}$$

which converges to $E\{Z(\mu)\}$ with probability one, according to the strong law of large number. Thus, $\pi_i \triangleq \frac{1}{E\{T_i\}}$.

* $Z_n(\mu)$'s are iid for a recurrent irreducible homogeneous Markov chain. For every experiment outcome μ_0, the time average $(\frac{1}{n}\sum_{k=1}^{n} T_i(k))^{-1}$ converges to p_i and $\frac{1}{n}\sum_{k=1}^{n} T_i(k)$ converges to $E\{T_i\}$, both with probability one. Thus, $E\{Z(\mu)\} = p_i = \frac{1}{E\{T_i\}}$.

- Defns:

  1. If $E\{T_i\} < \infty$, or equivalently $\pi_i > 0$, we say that state i is positive recurrent.
  2. If $E\{T_i\} = \infty$, or equivalently $\pi_i = 0$, state i is called null recurrent.

  One can show that both positive recurrence and null recurrence are class properties.

- Defn: We say that a discrete-time homogenous Markov chain *ergodic* iff it is irreducible, aperiodic, and positive recurrent.

- Since π_i is the proportion of time spent in state i, then $\pi_i p_{ij}$ is the proportion of time in which state j follows state i (from the law of large number). Therefore, we have the following identities:

$$\begin{aligned} \pi_j &= \sum_i \pi_i p_{ij} \qquad \forall j \qquad \circledast \\ \sum_i \pi_i &= 1. \end{aligned}$$

Ex:

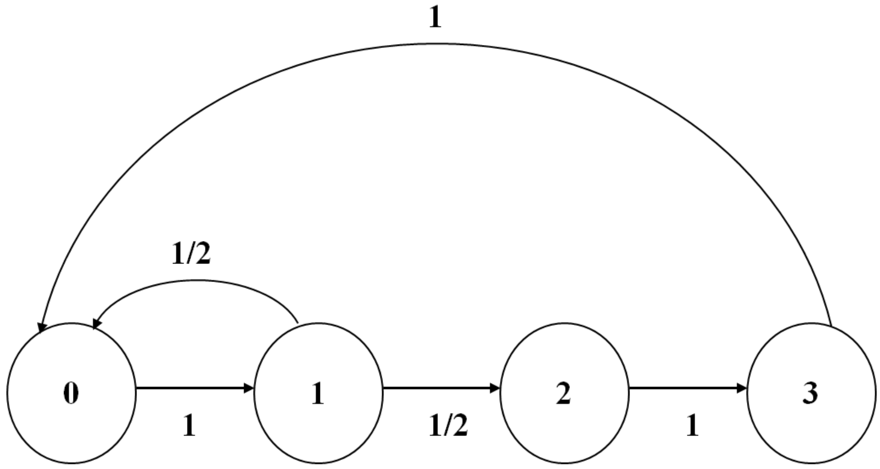


Figure 67:

$$\left.\begin{array}{l}\pi_0 = \pi_3 + \frac{1}{2}\pi_1 \\ \pi_1 = \pi_0 \\ \pi_2 = \frac{1}{2}\pi_1 \\ \pi_3 = \pi_2 \\ \pi_0 + \pi_1 + \pi_2 + \pi_3 = 1\end{array}\right\} \Rightarrow \left\{\begin{array}{l}\pi_1 = \pi_0 = \frac{1}{3} \\ \pi_3 = \pi_2 = \frac{1}{6}\end{array}\right. .$$

Recall in the previous discussion that, for discrete-time homogeneous Markov chains that exhibit steady-state behavior, the n-step state transition probability matrix approaches a fixed matrix as $n \to \infty$.
Let us show under what condition this occurs.

- Theorem: For an irreducible, aperiodic, positive recurrent, and homogeneous discrete-time Markov chains, i.e., for an ergodic discrete-time Markov chain,

$$\lim_{n\to\infty} p_{ij}(n) = \pi_j \qquad \forall j$$

where π_j is the unique solution of $\circledast$ and is also called the stationary probability for state j.

This theorem states that for such Markov chains,

"steady-state state probability (i.e., ensemble average) equals to stationary probability defined from time average" which is the reason we call it ergodic.

- Theorem: For an irreducible, periodic, positive recurrent, and homogeneous discrete-time Markov-chain with period d,

$$\lim_{n\to\infty} p_{jj}(nd) = d\pi_j \qquad \forall j$$

where π_j is the unique nonnegative solution to $\circledast$.

Notes:

1. The probability of occurrence for such a periodic Markov chain is d times the proportion of time spent in state j at the allowable times and zero elsewhere.
2. Such Markov chains have no steady state.

- Now, return to continuous-time Markov chains:

- Theorem: For a continuous-time Markov chain whose embedded Markov chain $X_n(\mu)$ is homogeneous, irreducible and positive recurrent,

$$p_i = \frac{\pi_i/\nu_i}{\sum_j \pi_j/\nu_j}$$

where $\frac{1}{\nu_i}$ is the mean state occupancy time in state i, π_i is the stationary probability of state i for the embedded Markov chain $X_n(\mu)$, obtained form $\circledast$, and p_i's are the unique solution to the global balance equation.

Proof: Let $N_i(n)$ be the number of times state i occurs in the first n transitions of the embedded Markov chain $X_n(\mu)$ and let $T_i(j)$ denote the occupancy time the jth time state i occurs. For a given state i, $T_i(j)$'s are iid because the embedded Markov chain is homogeneous. Then, the proportion of time spent in state i after the first n transitions is (for a given experiment outcome μ_0)

$$\frac{\text{time spent in state } i}{\text{time spent in all states}} = \frac{\sum\limits_{j=1}^{N_i(n)} T_i(j)}{\sum\limits_k \sum\limits_{j=1}^{N_k(n)} T_k(j)} = \frac{\frac{N_i(n)}{n} \cdot \frac{1}{N_i(n)} \sum\limits_{j=1}^{N_i(n)} T_i(j)}{\sum\limits_k \frac{N_k(n)}{n} \cdot \frac{1}{N_k(n)} \sum\limits_{j=1}^{N_k(n)} T_k(j)}.$$

As $n \to \infty$,

$$\frac{N_i(n)}{n} \to \pi_i$$

where π_i is the stationary probability of state i of the embedded Markov-chain.
In addition, by the law of large number,

$$\frac{1}{N_i(n)} \sum_{j=1}^{N_i(n)} T_i(j) \rightarrow E\{T_i(\mu)\} = 1/\nu_i$$

with probability one since $\lim_{n\to\infty} N_i(n) = \infty$ for every state i (because of the recurrence and irreducibility associated with the embedded Markov chain).
Therefore, as $n \rightarrow \infty$,

$$\frac{\text{time spent in state } i}{\text{time spent in all states}} = \frac{\pi_i/\nu_i}{\sum_k \pi_k/\nu_k} \triangleq \widetilde{p}_i.$$

Note that π_i's are the unique solution to

$$\begin{aligned} & \pi_j = \sum_i \pi_i q_{ij} \qquad \forall j \\ \Rightarrow \quad & \nu_j \widetilde{p}_j \sum_k \pi_k/\nu_k = \sum_i \nu_i \widetilde{p}_i (\sum_k \pi_k/\nu_k) q_{ij} \\ \Rightarrow \quad & \nu_j \widetilde{p}_j = \sum_i \nu_i \widetilde{p}_i q_{ij} \end{aligned}$$

(due to the positive recurrence). Since $q_{ij} = \gamma_{ij}/\nu_i$, $i \neq j$ and $q_{ii} = 0$ by definition of γ_{ij}, we have

$$\nu_j \widetilde{p}_j = \sum_{i \neq j} \widetilde{p}_i \gamma_{ij} \qquad \forall j$$

which is the global balance equation of a continuous-time homogeneous Markov chain that settles into steady state. Thus, $\widetilde{p}_j$'s are the unique solution to the global balance equation, i.e.,

$$\widetilde{p}_i = p_i = \frac{\pi_i/\nu_j}{\sum_k \pi_k/\nu_k}.$$

Q.E.D.

7.5 Time-Reversed Markov Chains

- Reference: Alberto Leon-Garcia, "Probability and Random Processes for Electrical Engineering."
- Consider stationary ergodic discrete-time Markov chains.
- We have shown previously that the time-reversed process of a Markov process is still a Markov process, which has the one-step state transition probability

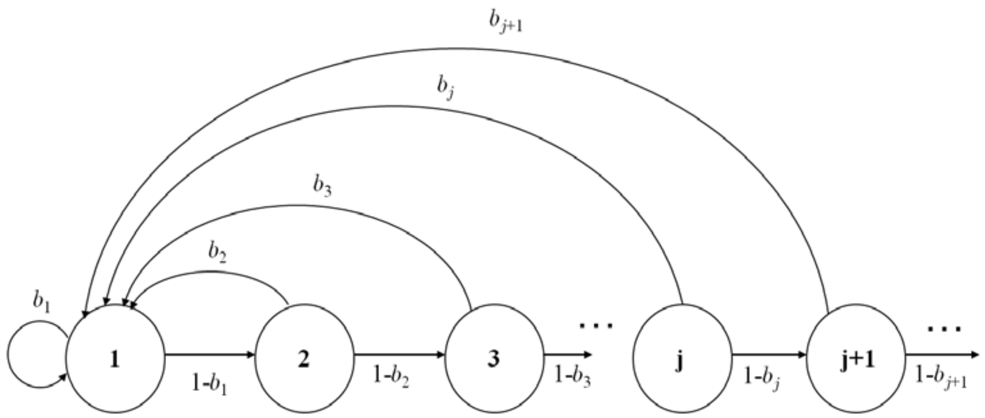
$$\begin{aligned}
q_{ij} &\triangleq \Pr\{X_{n-1}(\mu) = j | X_n(\mu) = i\} \\
&= \Pr\{X_{n-1}(\mu) = j | X_n(\mu) = i, X_{n+1}(\mu) = i_1, ..., X_{n+k}(\mu) = i_k\} \\
&= \frac{\Pr\{X_{n-1}(\mu) = j, X_n(\mu) = i, X_{n+1}(\mu) = i_1, ..., X_{n+k}(\mu) = i_k\}}{\Pr\{X_n(\mu) = i, X_{n+1}(\mu) = i_1, ..., X_{n+k}(\mu) = i_k\}} \\
&= \frac{\pi_j p_{ji} p_{ii_1} \cdots p_{i_{k-1}i_k}}{\pi_i p_{ii_1} \cdots p_{i_{k-1}i_k}} \quad \text{(from forward Markov property)} \\
&= \frac{\pi_j p_{ji}}{\pi_i}
\end{aligned}$$

which gives

$$q_{ij} = \frac{\pi_j p_{ji}}{\pi_i}. \qquad (*)$$

Since $X_n(\mu)$ is irreducible and aperiodic (due to ergodicity), its stationary state probability π_j represents the proportion of time that the state is state j, which does not depend on whether the Markov chain goes forward or backward. Thus, the forward and backward (reverse) Markov chains must have the same stationary state probabilities.

Ex: Consider the Markov chain



which is characterized by

$$
\begin{aligned}
p_{i,i+1} &= 1 - b_i, \qquad i = 1, 2, ... \\
p_{i,1} &= b_i, \qquad i = 1, 2, ... \\
p_{i,j} &= 0, \qquad \text{otherwise}
\end{aligned}
$$

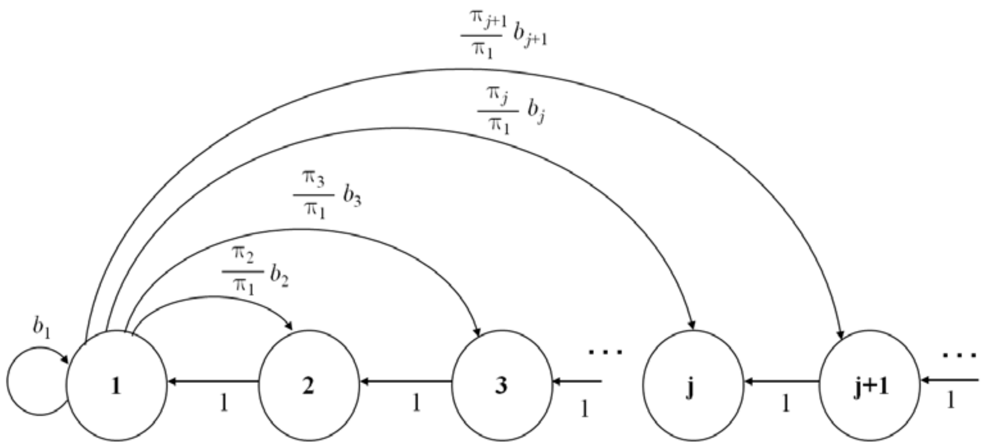
Now, from $(*)$,

$$
\begin{aligned}
q_{i,i-1} &= \frac{\pi_{i-1} p_{i-1,i}}{\pi_i} = \frac{\pi_{i-1}}{\pi_i}(1 - b_{i-1}) = 1, \qquad i = 2, 3, ... \text{ (note @)} \\
q_{1,i} &= \frac{\pi_i p_{i,1}}{\pi_1} = \frac{\pi_i}{\pi_1} b_i, \qquad i = 1, 2, ... \\
q_{i,j} &= 0, \qquad \text{otherwise}
\end{aligned}
$$

Note @: For $i = 2, 3, ...$

$$
\begin{aligned}
& \pi_i = \sum_j \pi_j p_{j,i} = \pi_{i-1} p_{i-1,i} \\
\Rightarrow \quad & \pi_i = \pi_{i-1}(1 - b_{i-1}) \\
\Rightarrow \quad & \frac{\pi_{i-1}}{\pi_i}(1 - b_{i-1}) = q_{i,i-1} = 1.
\end{aligned}
$$

So, the reverse Markov chain is given by



- A stationary ergodic discrete-time Markov chain is said to be reversible iff the one-step state transition probability matrix of the forward and reverse processes are the same, i.e., iff

$$
q_{ij} = p_{ij} \qquad \forall i, j.
$$

Since $\pi_i q_{ij} = \pi_j p_{ji}$ from $(*)$, we say that a stationary ergodic Markov chain is reversible iff

$$\pi_i p_{ij} = \pi_j p_{ji} \qquad \forall i, j$$

i.e., the proportion of transitions from i to j is equal to the proportion of transitions from j to i.

Ex: Consider the state diagram for a birth-and-death process:

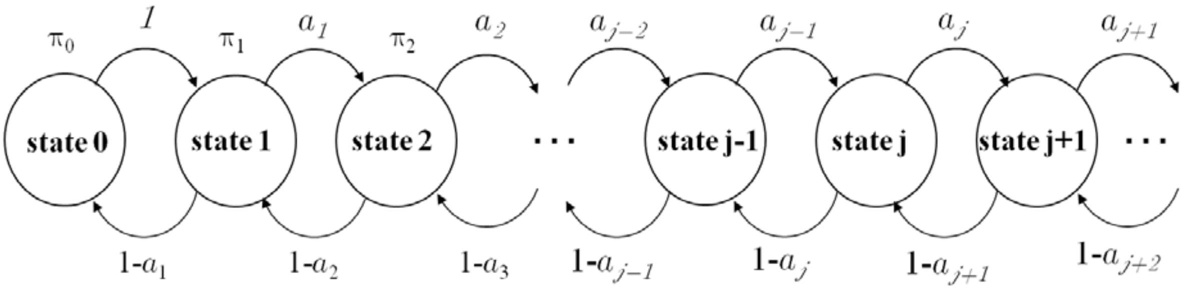


Figure 68:

Suppose that the process is a reversible Markov process. Thus,

$$\pi_i a_i = \pi_{i+1}(1 - a_{i+1}) \Rightarrow \pi_{i+1} = \pi_0 \prod_{j=0}^{i} \frac{a_j}{1 - a_{j+1}}.$$

Because $\sum_{i=0}^{\infty} \pi_i = 1$,

$$\pi_0 = [\sum_{i=0}^{\infty} \prod_{j=0}^{i} \frac{a_j}{1 - a_{j+1}}]^{-1}$$

and thus

$$\pi_{i+1} = [\sum_{l=0}^{\infty} \prod_{j=0}^{l} \frac{a_j}{1 - a_{j+1}}]^{-1} \prod_{k=0}^{i} \frac{a_k}{1 - a_{k+1}}$$

provided that $\sum_{i=0}^{\infty} \prod_{j=0}^{i} \frac{a_j}{1-a_{j+1}} > 1$. When $\sum_{i=0}^{\infty} \prod_{j=0}^{i} \frac{a_j}{1-a_{j+1}} \leq 1$, the Markov chain does not settle into stationarity.

- Now, reverse a stationary continuous-time Markov chain $X(\mu, t)$ which has an irreducible and positive recurrent embedded Markov chain. Then,

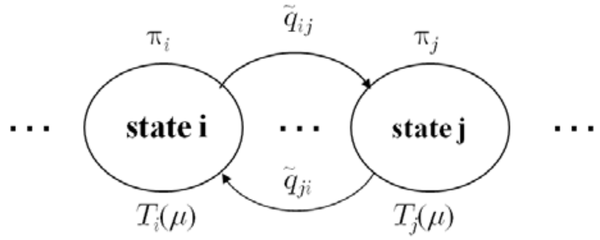
$$\begin{aligned}
& \Pr\{X(\mu, t^{'}) = i \text{ for } t - s < t^{'} < t | X(\mu, t) = i\} \\
= & \frac{\Pr\{X(\mu, t - s) = i, T_i(\mu) > s\}}{\Pr\{X(\mu, t) = i\}} \\
= & \frac{\Pr\{X(\mu, t - s) = i\} \Pr\{T_i(\mu) > s\}}{\Pr\{X(\mu, t) = i\}}
\end{aligned}$$

because state occupancy time $T_i(\mu)$ is independent of the state event $\{X(\mu, t - s) = i\}$. Because the process is stationary, $\Pr\{X(\mu, t - s) = i\} = \Pr\{X(\mu, t) = i\}$ and

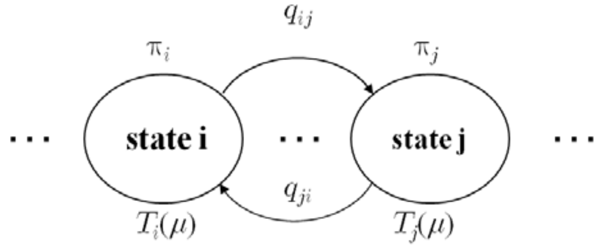
$$\Pr\{X(\mu, t^{'}) = i \text{ for } t - s < t^{'} < t | X(\mu, t) = i\} = \Pr\{T_i(\mu) > s\} = \exp\{-\nu_i s\}.$$

Thus, the reverse process also has the same exponentially distributed state occupancy times. Since the reverse process is also continuous-time Markov, it can be described by the embedded Markov chain

Forward Process



Reverse Process



where from $(*)$

$$\begin{aligned}
q_{ij} &= \frac{\pi_j}{\pi_i} \widetilde{q}_{ji} \\
&= \frac{\pi_j}{\pi_i} \cdot \frac{\gamma_{ji}}{\nu_j}
\end{aligned}$$

with $\widetilde{q}_{ji}$ defined as the state transition probability for the embedded Markov chain associated with the forward process with $\widetilde{q}_{ji} \triangleq \frac{\gamma_{ji}}{\nu_j}$. Also, we can define $q_{ij} = \frac{\gamma'_{ij}}{\nu_i}$ where γ'_{ij} represents the transition rate from state i to state j for the reverse process and is defined by

$$\begin{aligned}\gamma'_{ij} &= q_{ij}\nu_i \\ &= \frac{\pi_j}{\pi_i} \cdot \frac{\nu_i}{\nu_j} \cdot \gamma_{ji} \\ &= (c\frac{\pi_j}{\nu_j}) \cdot (\frac{\nu_i}{c\pi_i}) \cdot \gamma_{ji} \\ &= \frac{p_j}{p_i}\gamma_{ji}\end{aligned}$$

where $c \triangleq (\sum_k \pi_k/\nu_k)^{-1}$ and p_i denotes the steady-state state probability for state i given by

$$p_i = c\frac{\pi_i}{\nu_i}.$$

Thus, the steady-state state probabilities of $X(\mu, t)$, $\{p_i\}$, satisfy

$$p_i\gamma'_{ij} = p_j\gamma_{ji} \qquad (+)$$

where the LHS is for the reverse process and the RHS for the forward process.

- Since both forward and reverse processes have the same distribution for state occupancy time, $X(\mu, t)$ is reversible iff its embedded Markov chain is reversible, i.e., $\pi_j\widetilde{q}_{ji} = \pi_i\widetilde{q}_{ij}$. Since $\pi_j = \nu_j p_j/c$, $X(\mu, t)$ is reversible

$$\begin{aligned} &\text{iff} && \nu_j p_j \widetilde{q}_{ji} = \nu_i p_i \widetilde{q}_{ij} \\ \Rightarrow\ &\text{iff} && p_j\gamma_{ji} = p_i\gamma_{ij} \\ (+)\ \Rightarrow\ &\text{iff} && \gamma'_{ij} = \gamma_{ij}. \qquad (\&)\end{aligned}$$

This (&), together with the relation $q_{ij} = \pi_j\widetilde{q}_{ji}/\pi_i$ and the statistical property of $T_i(\mu)$, can statistically describe the reverse process directly from the reversible forward process $X(\mu, t)$ (characterized by ν_i's and γ_{ij}'s).

8 Queuing Systems

- Reference: Alberto Leon-Garcia, "Probability and Random Processes for Electrical Engineering."

8.1 Elements of A Queuing System

- Consider a first-come first-serve system where

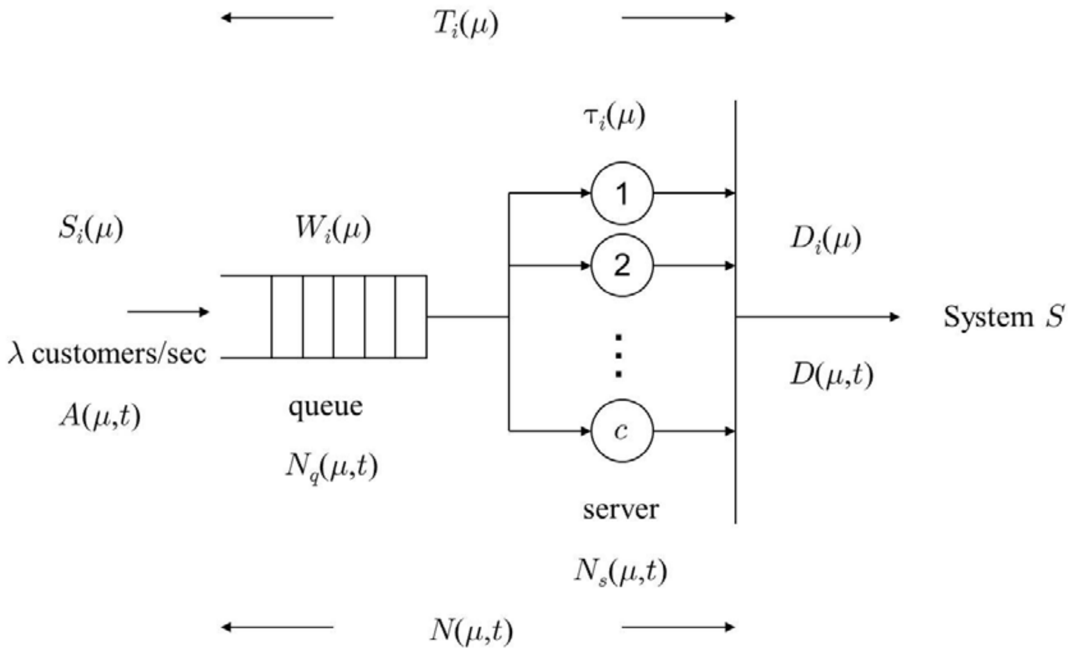


Figure 69:

- $N_q(\mu, t)$ is a random process denoting the number of customers in queue at time t,

- $N_s(\mu, t)$ is a random process denoting the number of customers in service at time t,

- $N(\mu, t) = N_s(\mu, t) + N_q(\mu, t)$ is a random process denoting the number of customers in the system at time t,

- $\lambda \equiv$ the average customer arrival rate,
- $S_i(\mu) \equiv$ the arrival time of the i-th customer,
- $W_i(\mu) \equiv$ the waiting time that the i-th customer stays in queue,
- $\tau_i(\mu) \equiv$ the waiting time that the i-th customer stays in the server,
- $T_i(\mu) = W_i(\mu) + \tau_i(\mu) \equiv$ the total delay of the i-th customer in the system,
- $D_i(\mu) \equiv$ the departure time of the i-th customer.

- If the queue is finite, the customer that arrives and finds no waiting space in the queue will be rejected. Such customers are called "blocked" with a rate of λ_b.
- Let us denote the system by notation $a/b/m/K$ where

    - $a \equiv$ the type of arrival process,
    - $b \equiv$ the type of service time distribution,
    - $m \equiv$ the number of servers,
    - $K \equiv$ the maximum number of customers allowed in the system at any time.

Notation:

- $a = M$ stands for a Poisson arrival process with iid exponential interarrival times and independent of system states,
- $b = M$ stands for iid exponential service times which are independent of system states. When the server array is occupied by at least one customer, this service distribution implies a Poisson departure process with iid exponential interdeparture times.
- $b = G$ stands for iid service times with some general distribution.

8.2 Little's Formula

- Consider mean-ergodic systems with iid total delays.
- Theorem: For systems that reach steady state, the average number of customers in a system is equal to the product of the average customer arrival rate and the average time spent in the system per customer, i.e.,

$$E\{N(\mu, \infty)\} = \lambda E\{T(\mu)\}.$$

This is called the Little's formula.

Notes:

1. Here, λ is the actual arrival rate to the system.
2. Mean-Ergodicity means that all random processes involved in the system are mean-ergodic. Thus, statistical averages involved in the system can be measured by the corresponding time averages of a single sample process (in the mean-square sense or with probability one).
3. Because $T_i(\mu)$'s are iid, $T(\mu)$ represent a random time having the same distribution as $T_i(\mu)$'s.

Proof: Let

- $A(\mu, t) \equiv$ the number of customer arrivals up to time t in the system S,
- $D(\mu, t) \equiv$ the number of customer departures up to time t in the system S.

Thus, $N(\mu, t) = A(\mu, t) - D(\mu, t)$ and $D_i(\mu) = S_i(\mu) + T_i(\mu)$.

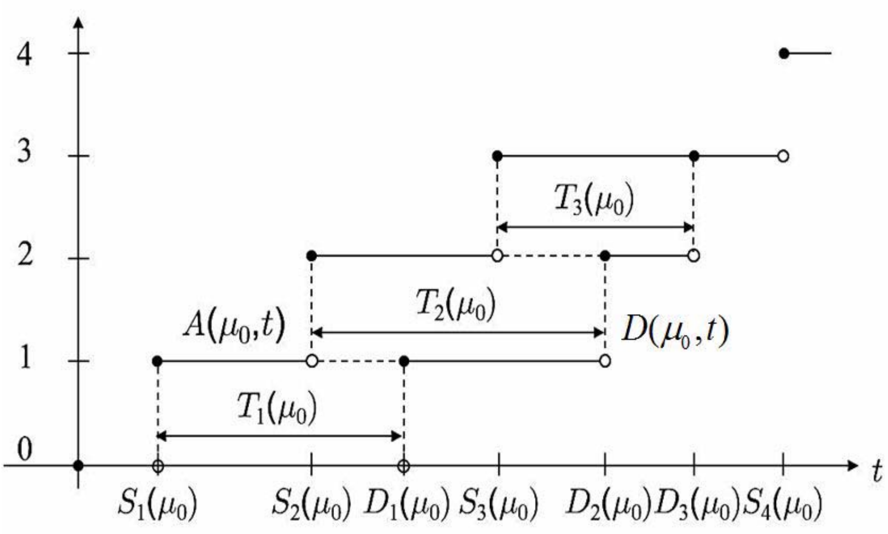


Figure 70:

Now, for $\mu = \mu_0$, define the time average of the number of customers in the system in $[0, t]$ as

$$< N >_t = \frac{1}{t} \int_0^t N(\mu_0, t') dt'.$$

Consider a t_0 such that $N(\mu_0, t_0) = 0$. Then,

$$< N >_{t_0} = \frac{1}{t_0} \int_0^{t_0} N(\mu_0, t') dt' = \frac{1}{t_0} \sum_{i=1}^{A(\mu_0, t_0)} T_i(\mu_0).$$

Also, define the average arrival rate up to time t as

$$\begin{aligned} < \lambda >_t &= \frac{A(\mu_0, t)}{t} \\ \Rightarrow < N >_{t_0} &= \frac{A(\mu_0, t_0)}{t_0} \frac{1}{A(\mu_0, t_0)} \sum_{i=1}^{A(\mu_0, t_0)} T_i(\mu_0) \\ &= < \lambda >_{t_0} \frac{1}{A(\mu_0, t_0)} \sum_{i=1}^{A(\mu_0, t_0)} T_i(\mu_0). \end{aligned}$$

Define $< T >_t$ as the average of the time spent in the system by the first $A(\mu_0, t)$ customers in the interval $[0, t]$. This implies that

$$
\begin{aligned}
< T >_t &= \frac{1}{A(\mu_0, t)} \sum_{i=1}^{A(\mu_0,t)} T_i(\mu_0) \\
\Rightarrow < N >_{t_0} &= < \lambda >_{t_0} \cdot < T >_{t_0} .
\end{aligned}
$$

If the system is mean ergodic and has iid total delays, then

$$
\begin{aligned}
< \; & N >_{t_0} \overset{t_0 \to \infty}{\rightarrow} E\{N(\mu, \infty)\} \\
< \; & \lambda >_{t_0} \overset{t_0 \to \infty}{\rightarrow} \lambda \\
< \; & T >_{t_0} \overset{t_0 \to \infty}{\rightarrow} E\{T(\mu)\}.
\end{aligned}
$$

Q.E.D.

Notes:

1. When $E\{N(\mu, t)\} < \infty$, there are many t_0's such that $N(\mu_0, t_0) = 0$. Thus, the limitation of t to t_0 can be relaxed as $t \to \infty$.
2. It can be shown that Little's formula also holds for systems with "other than" the first-come first-serve service disciplines.

- Ex: Consider the waiting queue with iid $W_i(\mu)$'s.

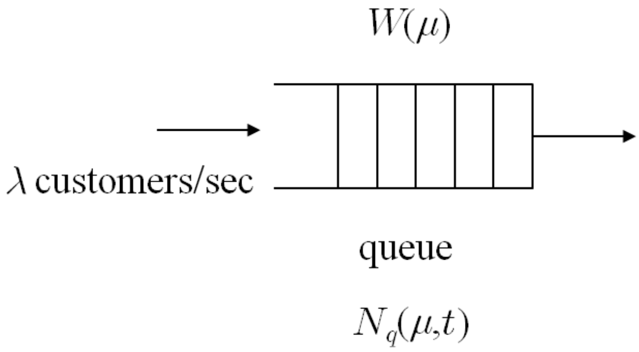


Figure 71:

This can be regarded as a queuing system. Thus, $E\{N_q(\mu, \infty)\} = \lambda E\{W(\mu)\}$.

- Ex: Consider the server with iid $\tau_i(\mu)$'s.

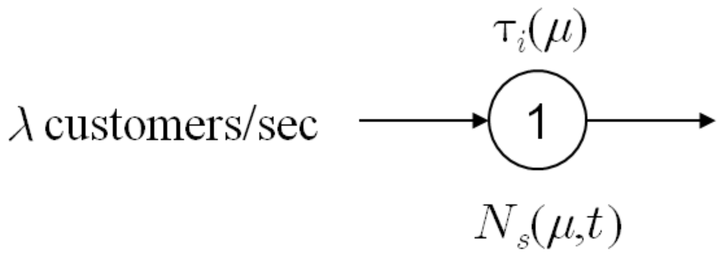


Figure 72:

This is a queuing system. Thus, $E\{N_s(\mu,\infty)\} = \lambda E\{\tau(\mu)\}$.

- In what follows, we consider mean-ergodic systems with iid total delays so that Little's formula holds.

8.3 The $M/M/1$ Queue

- Consider the system with

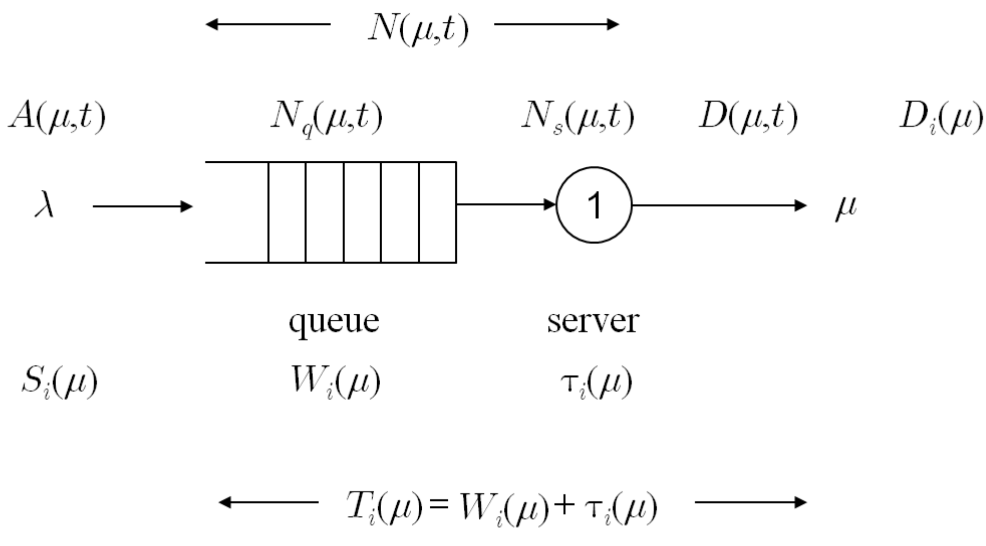


Figure 73:

  - one first-come-first-serve server

- infinite-space queue.
- Poisson arrival with rate λ and iid exponential interarrival times.
- iid exponential service times with rate μ.
- independent interarrival and service times.

* Our goal is to find

  1. the steady-state state probability of $N(\mu, t)$,
  2. $E\{N(\mu, \infty)\}$,
  3. the pdf of $T(\mu)$,
  4. $E\{T(\mu)\}$,
  5. $E\{N_q(\mu, \infty)\}$

  when the system settles into steady state.

(1) Since the process $N(\mu, t)$ has independent increments and decrements, it is a continuous-time Markov chain. Because of iid interarrival times and iid service times, the Markov chain is homogeneous.

Now, consider a small $\delta > 0$.

(a) $\Pr\{A(\mu, \delta) = 1\} = \lambda\delta + o(\delta)$ $\quad$ ($A(\mu, t)$ is Poisson).

(b) $\Pr\{A(\mu, \delta) = k\} = o(\delta)$ for $k = 2, 3, \ldots$ $\quad$ ($A(\mu, t)$ is Poisson).

(c) $\Pr\{\tau(\mu) \le \delta\} = 1 - e^{-\mu\delta} = \mu\delta + o(\delta)$.

(d)

$$\begin{aligned} & \Pr\{A(\mu, \delta) = 0, \tau(\mu) \le \delta\} \\ = & \Pr\{A(\mu, \delta) = 0\} \Pr\{\tau(\mu) \le \delta\} \\ = & (1 - \lambda\delta + o(\delta))(\mu\delta + o(\delta)) \\ = & \mu\delta + o(\delta). \end{aligned}$$

(e) $\Pr\{A(\mu, \delta) = k, \tau(\mu) \le \delta\} = o(\delta)$ for $k = 1, 2, \ldots$

(f) $\Pr\{A(\mu, \delta) = 1, \tau(\mu) > \delta\} = \lambda\delta + o(\delta)$.

(g) Note that $N(\mu, t) = A(\mu, t) - D(\mu, t)$. We have for a small $\delta > 0$ that

$$\begin{aligned} p_{ij}(\delta) &= \Pr\{N(\mu, t + \delta) = j | N(\mu, t) = i\} \\ &= \begin{cases} \Pr\{A(\mu, \delta) = 1\} \overset{(a),(e),(f)}{=} \lambda\delta + o(\delta), & \text{if } j = i + 1 \\ \Pr\{\tau(\mu) \le \delta\} \overset{(d),(e)}{=} \mu\delta + o(\delta), & \text{if } j = i - 1 \\ o(\delta), & \text{if } |i - j| > 1 \end{cases} \end{aligned}$$

and thus

$$\gamma_{i,i+1} = \lambda, \qquad \gamma_{i,i-1} = \mu, \qquad \gamma_{i,j} = 0 \text{ for } |i-j| > 1.$$

The transition state diagram for $N(\mu, t)$ is (state $\equiv$ number of customers in the system)

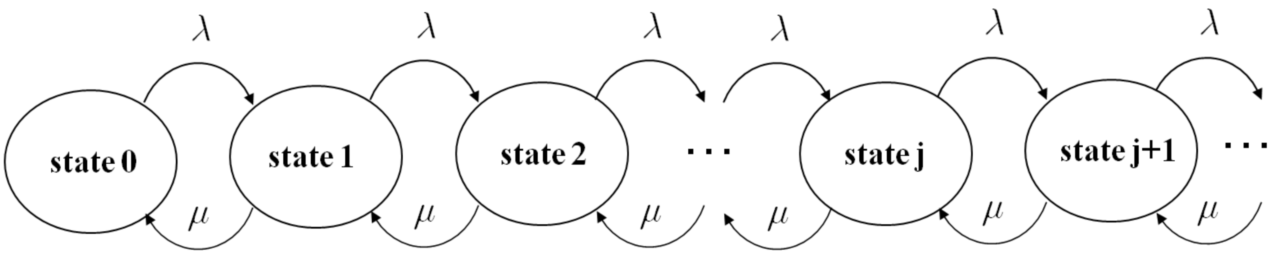


Figure 74:

From global balance equations for steady-state state probabilities

$$\begin{aligned} \lambda p_0 &= \mu p_1 \\ \lambda p_{j-1} + \mu p_{j+1} &= (\lambda + \mu) p_j \qquad j = 1, 2, ... \end{aligned}$$

we can obtain $p_j = \Pr\{N(\mu, \infty) = j\} = (1-\rho)\rho^j$ with $\rho = \frac{\lambda}{\mu}$ if $\rho < 1$. Note that $\rho < 1$ is required to ensure the steady state.

When the system settles into steady state, consider the following:

(2)

$$\begin{aligned} \lim_{t \to \infty} E\{N(\mu, t)\} &= \sum_{j=1}^{\infty} j p_j \\ &= (1-\rho) \sum_{j=1}^{\infty} j \rho^j \\ &= (1-\rho) \rho \frac{d}{d\rho} \sum_{j=1}^{\infty} \rho^j \\ &= \frac{\rho}{1-\rho}. \end{aligned}$$

(4) From Little's formula,

$$E\{T(\mu)\} = \frac{E\{N(\mu,\infty)\}}{\lambda} = \frac{\frac{\rho}{1-\rho}}{\lambda} = \frac{1}{\mu-\lambda}.$$

(5) Since $T(\mu) = W(\mu) + \tau(\mu)$,

$$E\{W(\mu)\} = E\{T(\mu)\} - E\{\tau(\mu)\} = \frac{1}{\mu-\lambda} - \frac{1}{\mu} = \frac{\lambda}{\mu(\mu-\lambda)}.$$

From Little's formula,

$$E\{N_q(\mu,\infty)\} = \lambda E\{W(\mu)\} = \frac{\lambda^2}{\mu(\mu-\lambda)}.$$

(3) Let us find the pdf of $T(\mu)$ below:

Define the process $N_a(\mu,t)$ as the number of customers found in the system by an arriving customer.

$$\begin{aligned}
\Pr\{N_a(\mu,t)=k\} &= \lim_{\delta\to 0}\Pr\{N(\mu,t)=k|A(\mu,t+\delta)-A(\mu,t)=1\}\\
&= \Pr\{\text{A customer that arrives at time } t \text{ finds } k \text{ customers in the system}\}\\
&= \lim_{\delta\to 0}\frac{\Pr\{A(\mu,t+\delta)-A(\mu,t)=1, N(\mu,t)=k\}}{\Pr\{A(\mu,t+\delta)-A(\mu,t)=1\}}\\
&= \lim_{\delta\to 0}\frac{\Pr\{A(\mu,t+\delta)-A(\mu,t)=1|N(\mu,t)=k\}}{\Pr\{A(\mu,t+\delta)-A(\mu,t)=1\}}\Pr\{N(\mu,t)=k\}.
\end{aligned}$$

Since $A(\mu,t)$ is Poisson and independent of systems states,

$$\Pr\{A(\mu,t+\delta)-A(\mu,t)=1|N(\mu,t)=k\} = \Pr\{A(\mu,t+\delta)-A(\mu,t)=1\}$$

(the event that there is one arrival in $(t,t+\delta]$ is independent of the event $\{N(\mu,t)=k\}$); thus,

$$\Pr\{N_a(\mu,t)=k\} = \Pr\{N(\mu,t)=k\}.$$

That is, the number of customers found in the system by an arriving customer is the same as the number of customers in the system at that time. This gives

$$\lim_{t\to\infty}\Pr\{N_a(\mu,t)=k\} = p_k = (1-\rho)\rho^k$$

Note that it takes $k+1$ iid service times for this new customer to leave the system, if $N_a(\mu, t) = k$ (noting that exponential service time is memoryless). Thus, as $t \to \infty$,

$$\begin{aligned} f_T(x|N_a(\mu,\infty)=k) &= \text{pdf of sum of } k+1 \text{ iid exponential random} \\ &\quad\ \text{variables with parameter } \mu \\ &= \frac{(\mu x)^k}{k!}\mu e^{-\mu x} \end{aligned}$$

for $x \geq 0$, which is a gamma pdf. This implies that

$$\begin{aligned} f_T(x) &= \sum_{k=0}^{\infty} f_T(x|N_a(\mu,\infty)=k)\,\Pr\{N_a(\mu,\infty)=k\} \\ &= \sum_{k=0}^{\infty} \frac{(\mu x)^k}{k!}\mu e^{-\mu x}(1-\rho)\rho^k \\ &= \mu(1-\rho)e^{-\mu x}\sum_{k=0}^{\infty}\frac{(\mu\rho x)^k}{k!} \\ &= \mu(1-\rho)e^{-\mu x}e^{\mu\rho x} \\ &= (\mu-\lambda)e^{-(\mu-\lambda)x} \end{aligned}$$

for $x \geq 0$. Similarly, for $x \geq 0$,

$$\begin{aligned} f_W(x|N_a(\mu,\infty)=k) &= \textit{pdf} \text{ of sum of } k \text{ iid exponential random} \\ &\qquad\quad \text{variables with parameter } \mu \\ &= \begin{cases} \frac{(\mu x)^{k-1}}{(k-1)!}\mu e^{-\mu x}, & k = 1, 2, ... \\ \delta(x), & k = 0 \end{cases} \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} f_W(x) &= \sum_{k=1}^{\infty}\frac{(\mu x)^{k-1}}{(k-1)!}\mu e^{-\mu x}(1-\rho)\rho^k + \delta(x)(1-\rho) \\ &= \mu(1-\rho)\rho e^{-\mu x}\sum_{k=1}^{\infty}\frac{1}{(k-1)!}(\mu\rho x)^{k-1} + \delta(x)(1-\rho) \\ &= (\mu-\lambda)\rho e^{-(\mu-\lambda)x} + \delta(x)(1-\rho). \end{aligned}$$

8.4 The $M/M/1$ System with Finite Capacity ($M/M/1/K$ System)

- The customer is turned away if the queue is full at its arrival.

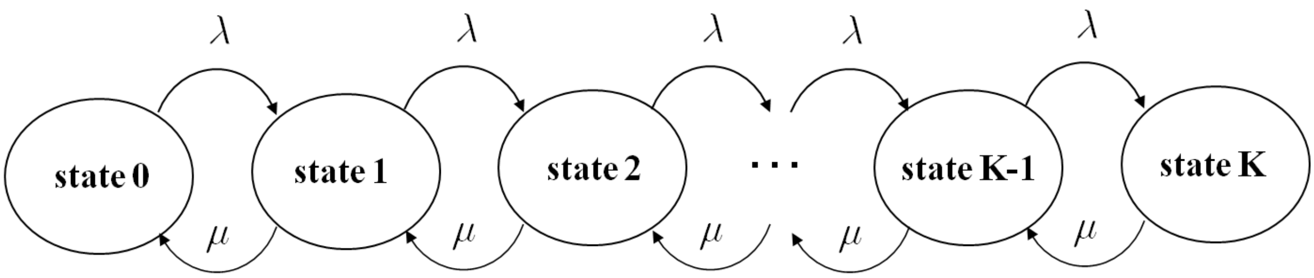


Figure 75:

- The global balance equations are

$$\begin{aligned} \lambda p_0 &= \mu p_1 \\ \lambda p_{j-1} + \mu p_{j+1} &= (\lambda + \mu) p_j, \; j = 1, 2, ..., K-1 \\ \lambda p_{K-1} &= \mu p_K \end{aligned}$$

which have the solutions

$$p_j = \begin{cases} \frac{(1-\rho)\rho^j}{1-\rho^{K+1}}, & j = 0, 1, 2, ..., K \text{ and } \rho \neq 1 \\ \frac{1}{K+1}, & j = 0, 1, 2, ..., K \text{ and } \rho = 1 \end{cases}$$

with $\rho = \frac{\lambda}{\mu}$. Note that ρ can be of any positive value.

- $$E\{N(\mu, \infty)\} = \sum_{j=0}^{K} j p_j = \begin{cases} \frac{\rho}{1-\rho} - \frac{(K+1)^{K+1}}{1-\rho^{K+1}}, & \rho \neq 1 \\ \frac{K}{2}, & \rho = 1 \end{cases}.$$

- When the system is full, the arrival rate "into" the system is zero. The rate at which a customer is blocked is

$$\lambda_b = \underbrace{\lambda}_{\text{arrival rate}} \cdot \underbrace{p_K}_{\substack{\text{the proportion} \\ \text{of time when the system} \\ \text{turns away customers}}}.$$

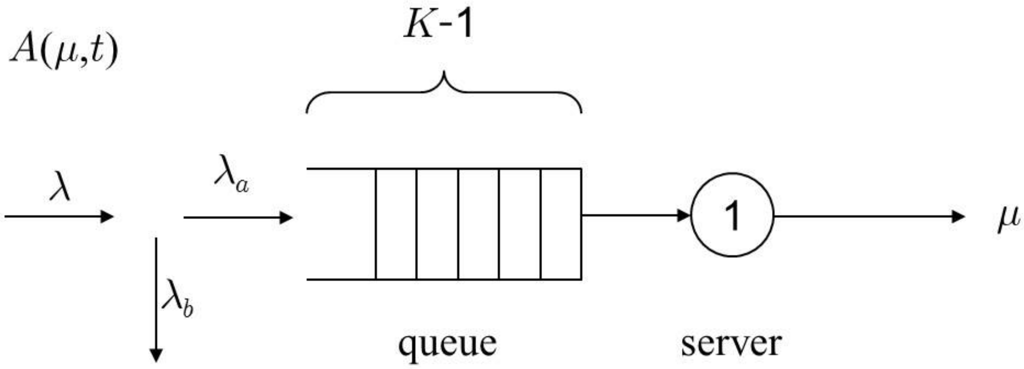


Figure 76:

Thus, the actual arrival rate into the system is

$$\lambda_a = \lambda - \lambda_b = \lambda(1 - p_K)$$

$$\Rightarrow \quad E\{T(\mu)\} = \frac{E\{N(\mu,\infty)\}}{\lambda_a} = \frac{\sum_{j=0}^{K} j p_j}{\lambda(1 - p_K)}.$$

8.5 The $M/M/c$ Queue

- Consider the $M/M/c$ queue with

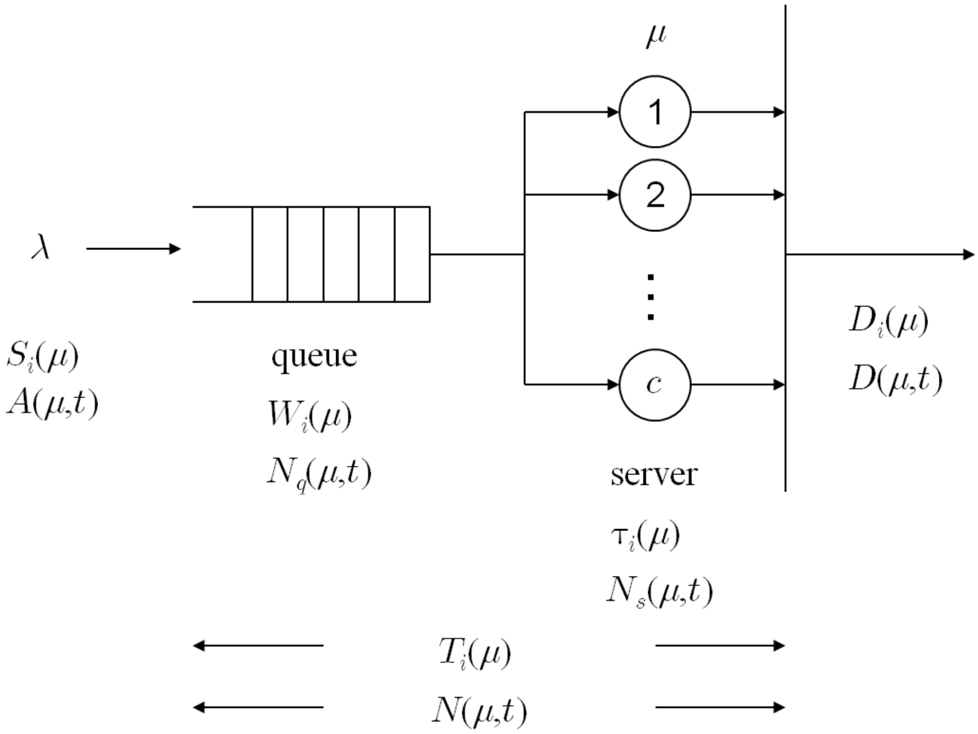


figure A1

  - c first-come-first-serve servers.
  - an infinite-space queue.
  - Poisson arrival with rate λ and iid exponential interarrival times.
  - c servers working independently with $c \geq 2$ and each server serving customers with iid exponential service times at rate μ.
  - independent interarrival and service times.

- Our goal is to find that

  1. the steady-state state probability of $N(\mu, t)$,
  2. $E\{N_q(\mu, \infty)\}$,
  3. $E\{W(\mu)\}$,

4. $E\{T(\mu)\}$,
5. $E\{N(\mu,\infty)\}$,
6. the marginal cdf's of $W(\mu)$ and $T(\mu)$

when the system settles into steady state.

(1) Since the process $N(\mu,t)$ has independent increments and decrements, it is a continuous-time Markov chain. Because of iid interarrival times and iid service times, the Markov chain is homogeneous.

Now, consider a small $\delta > 0$.

(a) $\Pr\{A(\mu,\delta)=1\} = \lambda\delta + o(\delta)$ $\quad$ ($A(\mu,t)$ is Poisson)

(b) $\Pr\{A(\mu,\delta)=k\} = o(\delta)$ for $k = 2,3,...$

(c) Let $X(\mu) = D_i(\mu) - D_{i-1}(\mu)$ be the time until next departure when customer $i-1$ leaves. Now, suppose that k servers are busy right after customer $i-1$ leaves, $k \in \{1,2,...,c\}$. Also, denote $\tau_l^*(\mu)$ as the service time that server l will take to serve its customer, $l \in \{1,2,...,k\}$. Then, $\tau_l^*(\mu)$'s are iid exponentially distributed with rate μ (noting exponential service times are memoryless) and $X(\mu) = \min\{\tau_1^*(\mu), \tau_2^*(\mu), ..., \tau_k^*(\mu)\}$. The complementary distribution of $X(\mu)$ is given by

$$\begin{aligned}
\Pr\{X(\mu) > \delta\} &= \Pr\{\min\{\tau_1^*(\mu), \tau_2^*(\mu), ..., \tau_k^*(\mu)\} > \delta\} \\
&= \Pr\{\tau_1^*(\mu) > \delta, \tau_2^*(\mu) > \delta, ..., \tau_k^*(\mu) > \delta\} \\
&= \Pr\{\tau_1^*(\mu) > \delta\} \Pr\{\tau_2^*(\mu) > \delta\} \cdots \Pr\{\tau_k^*(\mu) > \delta\} \\
&= e^{-\mu\delta} e^{-\mu\delta} \cdots e^{-\mu\delta} \\
&= e^{-k\mu\delta}.
\end{aligned}$$

Thus, $\Pr\{X(\mu) \le \delta\} = k\mu\delta + o(\delta)$ when k servers are busy. This shows that customers depart at rate $k\mu$ when k servers are busy. This indicates that when k servers are busy each service completion is exponentially distributed with parameter $k\mu$.

(d) When k servers are busy

$$\begin{aligned}
\Pr\{A(\mu,\delta)=0, X(\mu) \le \delta\} &= \Pr\{A(\mu,\delta)=0\} \Pr\{X(\mu) \le \delta\} \\
&= (1 - \lambda\delta + o(\delta))(k\mu\delta + o(\delta)) \\
&= k\mu\delta + o(\delta).
\end{aligned}$$

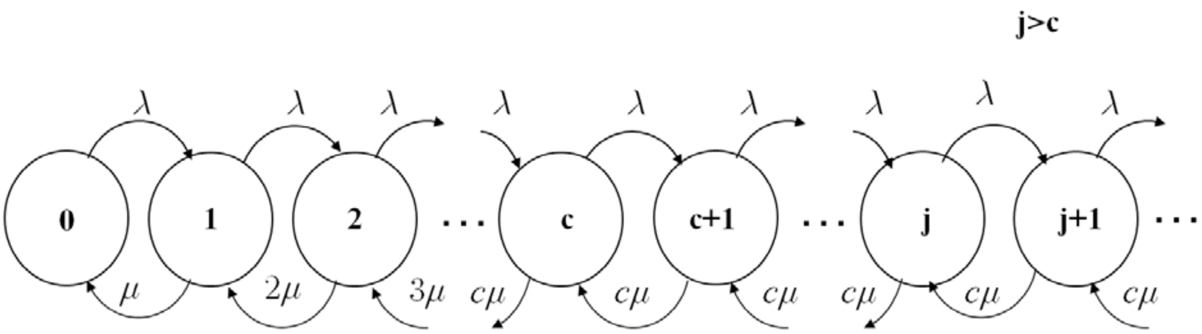
(e) $N(\mu, t) = A(\mu, t) - D(\mu, t)$. When k servers are busy, we have

$$\begin{aligned} p_{ij}(\delta) &= \Pr\{N(\mu, t+\delta) = j | N(\mu, t) = i\} \\ &= \begin{cases} \Pr\{A(\mu,\delta) = 1\} \overset{(a)}{=} \lambda\delta + o(\delta), & \text{if } j = i+1 \\ \Pr\{X(\mu) \le \delta\} \overset{(d)}{=} k\mu\delta + o(\delta), & \text{if } j = i-1 \\ o(\delta)\ , & \text{if } |i-j| > 1 \end{cases} . \end{aligned}$$

and thus

$$\gamma_{i,i+1} = \lambda, \qquad \gamma_{i,i-1} = k\mu, \qquad \gamma_{i,j} = 0 \text{ for } |i-j| > 1.$$

The transition state diagram for $N(\mu, t)$ is (state $\equiv$ number of customers in the system)



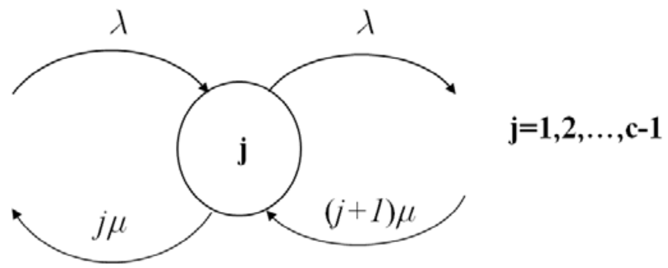


Figure 77:

(a) From global balance equations for steady-state state probabilities,

$$\begin{aligned} \lambda p_0 &= \mu p_1 \\ \lambda p_{j-1} + (j+1)\mu p_{j+1} &= (\lambda + j\mu)p_j \qquad j = 1, 2, ..., c-1 \\ \lambda p_{j-1} + c\mu p_{j+1} &= (\lambda + c\mu)p_j \qquad j = c, c+1, ... \end{aligned}$$

Now, for $j = 1, 2, ..., c-1$,

$$\begin{aligned}(j+1)\mu p_{j+1} - \lambda p_j &= j\mu p_j - \lambda p_{j-1} \\ &= ... \\ &= \mu p_1 - \lambda p_0 \\ &= 0 \\ \Rightarrow p_j &= \frac{\lambda}{j\mu} p_{j-1} \\ &= ... \\ &= \frac{\alpha^j}{j!} p_0\end{aligned}$$

with $\alpha = \lambda/\mu$. Further, for $j = c, c+1, ...,$

$$\begin{aligned}c\mu p_{j+1} - \lambda p_j &= c\mu p_j - \lambda p_{j-1} \\ &= ... \\ &= c\mu p_c - \lambda p_{c-1} \\ &= (c-1)\mu p_{c-1} - \lambda p_{c-2} \\ &= ... \\ &= \mu p_1 - \lambda p_0 \\ &= 0 \\ \Rightarrow p_j &= \frac{\lambda}{c\mu} p_{j-1} \\ &= ... \\ &= \rho^{j-c} p_c \\ &= \frac{\rho^{j-c}\alpha^c}{c!} p_0\end{aligned}$$

with $\rho = \lambda/(c\mu)$. Because $\sum_{j=0}^{\infty} p_j = 1$, we further have

$$p_0 \left\{ \sum_{j=0}^{c-1} \frac{\alpha^j}{j!} + \frac{\alpha^c}{c!} \sum_{j=c}^{\infty} \rho^{j-c} \right\} = 1.$$

The system has steady state if the term inside $\{\}$ is finite, i.e., if $\sum_{j=c}^{\infty} \rho^{j-c}$ converges, and greater than one. Thus, when $\rho < 1$, or equivalently $\lambda < c\mu$ (the average arrival rate is less than the total service rate at which c servers can process customers), and the term inside $\{\}$ is greater than one, the system has steady state and in the case

$$
\begin{aligned}
p_j &= \Pr\{N(\mu,\infty)=j\} \\
&= \begin{cases} \frac{\alpha^j}{j!}\left\{\sum_{l=0}^{c-1}\frac{\alpha^l}{l!}+\frac{\alpha^c}{c!}\sum_{l=c}^{\infty}\rho^{l-c}\right\}^{-1}, & j=1,2,...,c-1 \\ \frac{\rho^{j-c}\alpha^c}{c!}\left\{\sum_{l=0}^{c-1}\frac{\alpha^l}{l!}+\frac{\alpha^c}{c!}\sum_{l=c}^{\infty}\rho^{l-c}\right\}^{-1}, & j=c,c+1,... \end{cases}.
\end{aligned}
$$

Otherwise, the system is not stable.

Suppose that the system settles into steady state and consider the following:

(2) Now, consider the probability that an arriving customer finds all servers busy and has to wait in queue

$$
\begin{aligned}
\Pr\{W(\mu)>0\} &= \Pr\{N(\mu,\infty)\geq c\} \\
&= \sum_{j=c}^{\infty} p_j \\
&= \sum_{j=c}^{\infty} \rho^{j-c} p_c \\
&= \frac{p_c}{1-\rho}
\end{aligned}
$$

for $\rho < 1$. This probability is called the Erlang C formula and denoted by

$$
C(c,\rho,\alpha) = \frac{p_c}{1-\rho} = \Pr\{W(\mu)>0\}.
$$

Now, the mean number of customers waiting in the queue is

$$
\begin{aligned}
E\{N_q(\mu,\infty)\} &= \sum_{j=c}^{\infty}(j-c)p_j \\
&= \sum_{j=c}^{\infty}(j-c)\rho^{j-c}p_c \\
&= \sum_{k=0}^{\infty} k\rho^k p_c \\
&= \frac{\rho}{(1-\rho)^2}p_c \\
&= \frac{\rho}{1-\rho}C(c,\rho,\alpha).
\end{aligned}
$$

(3) The mean waiting time is found from Little's formula as

$$E\{W(\mu)\} = \frac{E\{N_q(\mu,\infty)\}}{\lambda} = \frac{1/\mu}{c(1-\rho)}C(c,\rho,\alpha).$$

(4) The mean total delay in the system is

$$E\{T(\mu)\} = E\{W(\mu)\} + E\{\tau(\mu)\} = \frac{1/\mu}{c(1-\rho)}C(c,\rho,\alpha) + \frac{1}{\mu}.$$

(5) The mean number of customers in the system is found from Little's formula as

$$\begin{aligned} E\{N(\mu,\infty)\} &= \lambda E\{T(\mu)\} \\ &= \lambda(E\{W(\mu)\} + E\{\tau(\mu)\}) \\ &= E\{N_q(\mu,\infty)\} + \alpha \\ &= \frac{\rho}{1-\rho}C(c,\rho,\alpha) + \alpha. \end{aligned}$$

(6) First, consider the conditional probability that there are $j-c$ customers (with $j \geq c$) in queue given that all servers are busy, i.e., $N(\mu,\infty) \geq c$:

$$\begin{aligned} \Pr\{N(\mu,\infty) = j | N(\mu,\infty) \geq c\} &= \frac{\Pr\{N(\mu,\infty) = j, N(\mu,\infty) \geq c\}}{\Pr\{N(\mu,\infty) \geq c\}} \\ &= \frac{\Pr\{N(\mu,\infty) = j\}}{\Pr\{N(\mu,\infty) \geq c\}} \\ &= \frac{p_j}{\sum_{k=c}^{\infty} p_k} \\ &= \frac{\rho^{j-c}p_c}{p_c/(1-\rho)} \\ &= (1-\rho)\rho^{j-c} \end{aligned}$$

which is a geometric distribution.

Next, suppose that a customer arrives when there are k customers in queue. There must be $k+1$ service completions before this customer enters service. Because each service completion is exponentially distributed with parameter $c\mu$ when all c servers are busy, the waiting time for the arriving customer is the sum of $k+1$ iid exponential random variables with parameter $c\mu$ and thus gamma distributed with

$$f_W(x|N(\mu,\infty) = c+k) = \frac{(c\mu x)^k}{k!}c\mu\exp\{-c\mu x\}.$$

Therefore, the conditional cdf of $W(\mu)$ when all c servers are busy, i.e., $N(\mu, \infty) \geq c$, is given by

$$
\begin{aligned}
& F_W(x|N(\mu,\infty) \geq c) \\
= & \sum_{k=0}^{\infty} F_W(x|N(\mu,\infty) = c+k) \Pr\{N(\mu,\infty) = c+k|N(\mu,\infty) \geq c\} \\
= & \sum_{k=0}^{\infty} \int_0^x \frac{(c\mu y)^k}{k!} c\mu \exp\{-c\mu y\} dy \cdot (1-\rho)\rho^k \\
= & (1-\rho) \int_0^x \sum_{k=0}^{\infty} \frac{(c\mu\rho y)^k}{k!} c\mu \exp\{-c\mu y\} dy \\
= & (1-\rho)c\mu \int_0^x \exp\{-c\mu(1-\rho)y\} dy \\
= & 1 - \exp\{-c\mu(1-\rho)x\}
\end{aligned}
$$

for $x \geq 0$. Thus, the cdf of $W(\mu)$ is

$$
\begin{aligned}
\Pr\{W(\mu) \leq x\} &= \Pr\{W(\mu) = 0\} + F_W(x|W(\mu) > 0) \Pr\{W(\mu) > 0\} \\
&= (1 - \Pr\{W(\mu) > 0\}) + F_W(x|N(\mu,\infty) \geq c) \Pr\{W(\mu) > 0\} \\
\text{(from (2))} &= (1 - C(c,\rho,\alpha)) + (1 - \exp\{-c\mu(1-\rho)x\}) C(c,\rho,\alpha) \\
&= 1 - C(c,\rho,\alpha) \exp\{-c\mu(1-\rho)x\}
\end{aligned}
$$

for $x \geq 0$. Because $T(\mu) = W(\mu) + \tau(\mu)$ where $W(\mu)$ and $\tau(\mu)$ are independent and $\tau(\mu)$ is exponentially distributed with parameter μ, the cdf of $T(\mu)$ can be derived as follows: For $x \geq 0$,

$$
\begin{aligned}
\Pr\{T(\mu) \leq x\} &= \Pr\{W(\mu) + \tau(\mu) \leq x\} \\
&= \int_0^x \Pr\{W(\mu) \leq x - \tau\} dF_\tau(\tau) \\
&= \int_0^x \Pr\{W(\mu) \leq x - \tau\} (\mu e^{-\mu\tau}) d\tau \\
&= \int_0^x [1 - C(c,\rho,\alpha) e^{-c\mu(1-\rho)(x-\tau)}] (\mu e^{-\mu\tau}) d\tau \\
&= (1 - e^{-\mu x}) - C(c,\rho,\alpha) e^{-c\mu(1-\rho)x} \int_0^x \exp\{\mu(c - c\rho - 1)\tau\} d\tau \\
&= (1 - e^{-\mu x}) - C(c,\rho,\alpha) e^{-c\mu(1-\rho)x} \\
&\quad \cdot [\mu(c - c\rho - 1)]^{-1} [e^{\mu(c-c\rho-1)x} - 1] \\
&= (1 - e^{-\mu x}) - \frac{C(c,\rho,\alpha)}{\mu(c - c\rho - 1)} [e^{-\mu x} - e^{-c\mu(1-\rho)x}].
\end{aligned}
$$

8.6 The $M/M/c/c$ System Without Queue

- Consider the $M/M/c/c$ queuing system that has c servers but no waiting room.

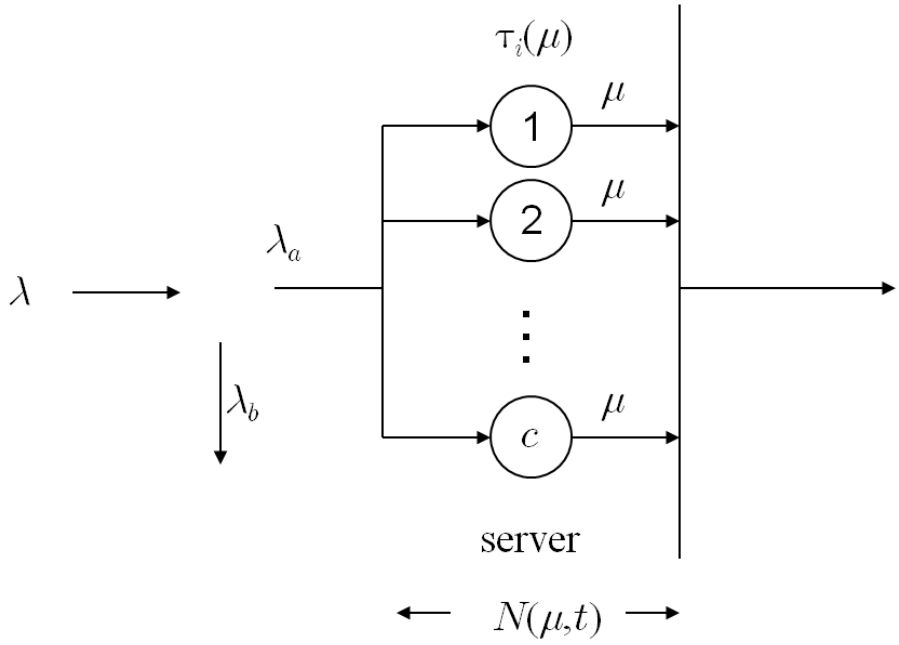
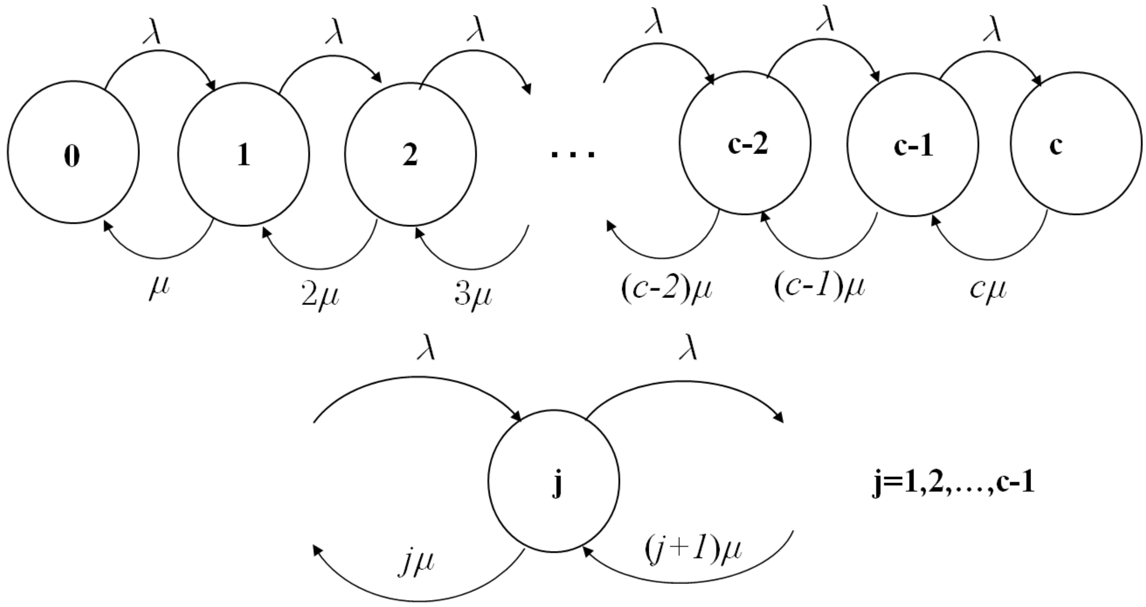


Figure 78:

The customer is turned away if all servers are occupied at his arrival. The transition state diagram is shown as



(1) The global balance equations are

$$
\begin{aligned}
\lambda p_0 &= \mu p_1 \\
\lambda p_{j-1} + (j+1)\mu p_{j+1} &= (\lambda + j\mu) p_j \qquad j = 1, 2, ..., c-1 \\
\lambda p_{c-1} &= c\mu p_c
\end{aligned}
$$

which have the solutions

$$p_j = \frac{\alpha^j}{j!} \left\{ \sum_{l=0}^{c} \frac{\alpha^l}{l!} \right\}^{-1} \qquad j = 0, 1, ..., c$$

with $\alpha = \frac{\lambda}{\mu}$.

(2) The Erlang B formula is defined as the probability that all servers are busy under steady state,

$$B(c, \alpha) = \Pr\{N(\mu, \infty) = c\} = p_c = \frac{\alpha^c}{c!} \left\{ \sum_{l=0}^{c} \frac{\alpha^l}{l!} \right\}^{-1}.$$

(3) When the system is full, the arrival rate "into" the system is zero. The rate at which a customer is blocked is $\lambda_b = \lambda p_c$. Thus, the

actual arrival rate into the system is

$$\begin{aligned}\lambda_a &= \lambda - \lambda_b \\ &= \lambda(1 - p_c) \\ \Rightarrow E\{N(\mu, \infty)\} &= \lambda_a E\{\tau(\mu)\} \\ &= \frac{\lambda_a}{\mu} \\ &= \frac{\lambda}{\mu}(1 - B(c, \alpha)).\end{aligned}$$

8.7 The $M/M/\infty$ System

- Consider the $M/M/\infty$ queuing system that has an infinite number of servers. In the case, no queue is necessary because every arriving customer is served right on its arrival. The transition state diagram is given by

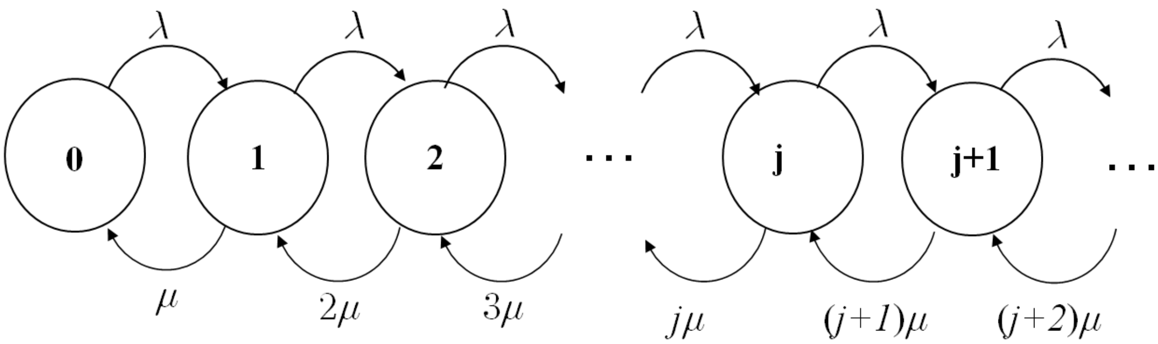


Figure 79:

The global balance equations are

$$\begin{aligned}\lambda p_0 &= \mu p_1 \\ \lambda p_{j-1} + (j+1)\mu p_{j+1} &= (\lambda + j\mu)p_j \qquad j = 1, 2, ...\end{aligned}$$

which have the solution

$$p_j = \frac{\alpha^j}{j!}e^{-\alpha} \qquad j = 0, 1, ...$$

with $\alpha = \frac{\lambda}{\mu}$. Therefore, when the system settles into steady state, the number of customers in the system is a Poisson random variable with mean

$$E\{N(\mu, \infty)\} = \sum_{j=0}^{\infty} j p_j = \alpha.$$

9 Random Walk Process and Brownian Motion Process

- Reference: Alberto Leon-Garcia, "Probability and Random Processes for Electrical Engineering."

9.1 Random Walk Process

- Consider two random sequences $D_1(\mu), D_2(\mu), ...$ and $I_1(\mu), I_2(\mu), ...$ where (1) $I_n(\mu)$'s are iid Bernoulli random variables with $\Pr\{I_n(\mu) = 1\} = p$ and $\Pr\{I_n(\mu) = 0\} = 1 - p$, and (2) $D_n(\mu)$'s are iid ± 1-valued random variables derived from $I_n(\mu)$'s as $D_n(\mu) = 1$ if $I_n(\mu) = 1$ and $D_n(\mu) = -1$ if $I_n(\mu) = 0$. Note that $D_n(\mu) = 2I_n(\mu) - 1$. The means and variances of $I_n(\mu)$ and $D_n(\mu)$ are given by

$$\begin{aligned} E\{I_n(\mu)\} &= p \\ Var\{I_n(\mu)\} &= E\{I_n^2(\mu)\} - p^2 = p(1-p) \\ E\{D_n(\mu)\} &= 2E\{I_n(\mu)\} - 1 = 2p - 1 \\ Var\{D_n(\mu)\} &= 4Var\{I_n(\mu)\} = 4p(1-p). \end{aligned}$$

- Now, we further define the sum sequence $S_1(\mu), S_2(\mu), ...$ by

$$S_n(\mu) = \sum_{k=1}^{n} D_k(\mu).$$

Such a sequence is called a random walk process. Suppose that we have sample sequences corresponding to a particular outcome μ_0, $D_1(\mu_0), D_2(\mu_0), ...$ and $S_1(\mu_0), S_2(\mu_0), ...$ These sample sequences can realize the motion of a particle along one-dimensional quantum space. For example, $D_n(\mu_0) = 1$ means that the particle moves one step to the right and $D_n(\mu_0) = -1$ means that the particle moves one step to the left. If the particle is located at the origin of the space initially at time $k = 0$, then $S_n(\mu_0)$ represents the position of the particle on the space at time n.

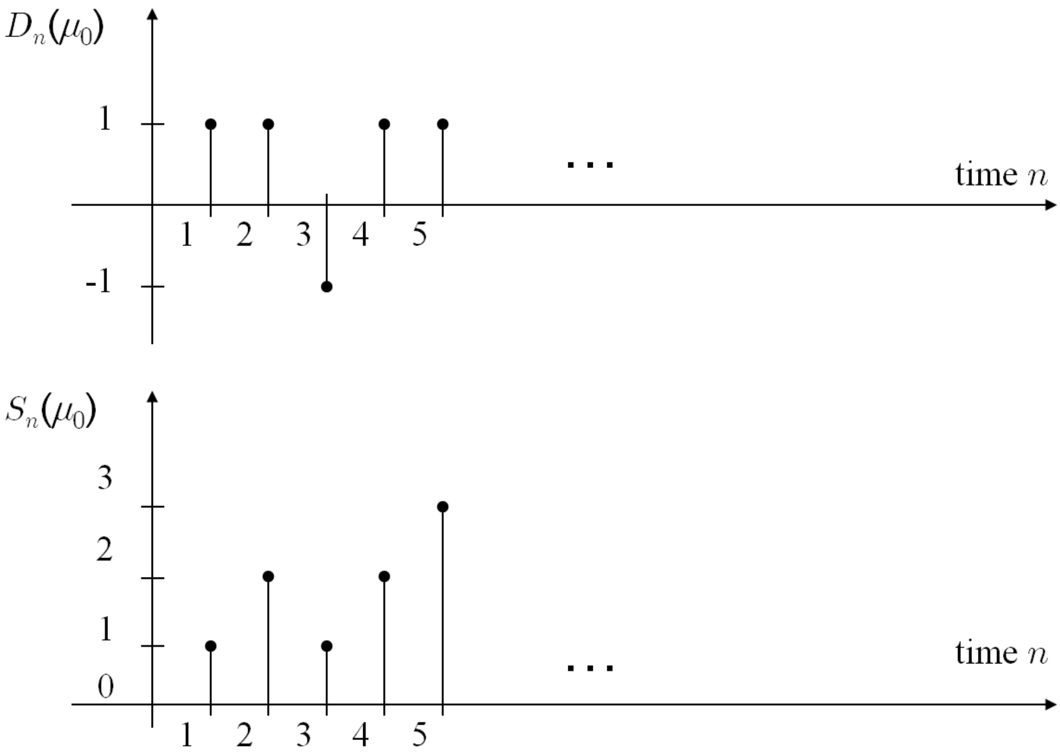


Figure 80:

- We have the following statistical property:
  - The mean and the variance of $S_n(\mu)$ are given by

$$\begin{aligned} E\{S_n(\mu)\} &= \sum_{k=1}^{n} E\{D_k(\mu)\} = n(2p-1) \\ Var\{S_n(\mu)\} &= \sum_{k=1}^{n} Var\{D_k(\mu)\} = 4np(1-p). \end{aligned}$$

  - The pdf of $S_n(\mu)$ is found as follows: Now, the characteristic function of $S_n(\mu)$ is

$$\begin{aligned}
\Phi_{S_n}(\omega) &= E\{\exp\{j\omega S_n(\mu)\}\} \\
&= E\{\exp\{j\omega \sum_{k=1}^{n} D_k(\mu)\}\} \\
&= \prod_{k=1}^{n} E\{\exp\{j\omega D_k(\mu)\}\} \\
&= E^n\{\exp\{j\omega D_k(\mu)\}\} \\
&= (p\exp\{j\omega\} + (1-p)\exp\{-j\omega\})^n \\
&= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \exp\{j\omega(2k-n)\}.
\end{aligned}$$

Thus, the pdf of $S_n(\mu)$ is given by

$$\Pr\{S_n(\mu) = 2k - n\} = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for } k \in \{0, 1, ..., n\}.$$

Note that $S_n(\mu)$'s are correlated and not identically distributed.

– Because $S_n(\mu) = \sum_{k=1}^{n} D_k(\mu)$ has iid unit increments/decrements $D_k(\mu)$'s, the random walk process has independent and stationary increments/decrements for all integer-boundary discrete-time intervals. Thus, the random walk process is a discrete-time homogeneous Markov chain.

9.2 Brownian Motion Process

- Consider the continuous-time process $X_\delta(\mu, t)$, $t \geq 0$, defined by

$$X_\delta(\mu, t) = hS_{\lfloor t/\delta \rfloor}(\mu)$$

where δ is a time unit with $\delta > 0$, h is a magnitude unit with $h > 0$, $\lfloor x \rfloor$ means the integer part of x, and $S_1(\mu), S_2(\mu), ...$ is a random walk process with $p = 1/2$. By default, we let $S_0(\mu) = 0$. This $X_\delta(\mu, t)$ is also a random walk process with rescaled time and magnitude units. Note that the mean and the variance of $X_\delta(\mu, t)$ for a fixed t are given by

$$\begin{aligned}
E\{X_\delta(\mu, t)\} &= hE\{S_{\lfloor t/\delta \rfloor}(\mu)\} = 0 \\
Var\{X_\delta(\mu, t)\} &= h^2 Var\{S_{\lfloor t/\delta \rfloor}(\mu)\} = h^2 \lfloor t/\delta \rfloor .
\end{aligned}$$

- Suppose that we simultaneously shrink δ and h with $h = \sqrt{\alpha\delta}$ and α a positive factor, and then consider the process

$$X(\mu, t) = \lim_{\delta \to 0} X_\delta(\mu, t).$$

For a fixed time t, the mean and the variance of $X(\mu, t)$ are given by

$$\begin{aligned} E\{X(\mu,t)\} &= \lim_{\delta\to 0} E\{X_\delta(\mu,t)\} = 0 \\ Var\{X(\mu,t)\} &= \lim_{\delta\to 0} Var\{X_\delta(\mu,t)\} \\ &= \lim_{\delta\to 0} h^2 \lfloor t/\delta \rfloor = \lim_{\delta\to 0} \alpha\delta \lfloor t/\delta \rfloor = \alpha t. \end{aligned}$$

This new process $X(\mu, t)$, $t \geq 0$, which begins at the origin with $X(\mu, 0) = 0$, has zero mean for all time but has a variance that increases linearly with time.

The process $X(\mu, t)$, $t \geq 0$, is called the Brownian motion process because it is used to model the motion of particles suspended in a fluid that move under the rapid and random impact of neighboring particles.

- Because $S_n(\mu) = \sum_{k=1}^{n} D_k(\mu)$ with $D_n(\mu)$'s being iid ± 1-valued random variables with $\Pr\{D_n(\mu) = \pm 1\} = 1/2$, $X(\mu, t)$ is an infinite sum of iid antipodal-valued and equally likely random variables. According to Central Limit Theorem, the first-order density of $X(\mu, t)$ is the density of a Gaussian random variable with mean zero and variance αt. That is,

$$f_t(x) = \frac{1}{\sqrt{2\pi\alpha t}} \exp\{-\frac{x^2}{2\alpha t}\}.$$

Because $X(\mu, t)$ is a limiting process of the random walk process $X_\delta(\mu, t)$, the Brownian motion process $X(\mu, t)$, $t \geq 0$, also inherits the property of independent and stationary increments/decrements from the random walk process. Therefore, the Brownian motion process is also a continuous-time homogeneous Markov process.

- Now, for $0 < t_1 < t_2 < ... < t_n$, the n-th order density of $X(\mu, t)$ is given by

$$\begin{aligned}
f_{t_1,t_2,\ldots,t_n}(x_1,x_2,\ldots,x_n) &= f_{t_1}(x_1)\prod_{k=2}^{n}\frac{f_{t_1,t_2,\ldots,t_k}(x_1,x_2,\ldots,x_k)}{f_{t_1,t_2,\ldots,t_{k-1}}(x_1,x_2,\ldots,x_{k-1})} \\
&\quad \text{(chain rule)} \\
&= f_{t_1}(x_1)\prod_{k=2}^{n}\frac{f_{t_{k-1},t_k}(x_{k-1},x_k)}{f_{t_{k-1}}(x_{k-1})} \\
&\quad \text{(Markovian property)} \\
&= f_{t_1}(x_1)\prod_{k=2}^{n} f_{t_k|t_{k-1}}(x_k|x_{k-1}) \qquad (+)
\end{aligned}$$

where $f_{t_k|t_{k-1}}(x_k|x_{k-1})$ represents the transition density for the process and is given by

$$f_{t_k|t_{k-1}}(x_k|x_{k-1}) = f_{t_k-t_{k-1}}(x_k - x_{k-1}).$$
(independent and stationary increments/decrements)

Thus, the statistic of the Brownian motion process can be fully described by its first-order density. From (+), the Brownian motion process is obviously a Gaussian random process as well as a Wiener process.

- The autocorrelation of $X(\mu, t)$ is given by

$$\begin{aligned}
R_X(t_1,t_2) &= E\{X(\mu,t_1)X(\mu,t_2)\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2 f_{t_1,t_2}(x_1,x_2)dx_2dx_1 \\
&= \int_{-\infty}^{\infty} x_1\frac{1}{\sqrt{2\pi\alpha t_1}}\exp\{-\frac{x_1^2}{2\alpha t_1}\} \\
&\quad \cdot\int_{-\infty}^{\infty} x_2\frac{1}{\sqrt{2\pi\alpha(t_2-t_1)}}\exp\{-\frac{(x_2-x_1)^2}{2\alpha(t_2-t_1)}\}dx_2dx_1 \\
&= \int_{-\infty}^{\infty} x_1^2\frac{1}{\sqrt{2\pi\alpha t_1}}\exp\{-\frac{x_1^2}{2\alpha t_1}\}dx_1 \\
&= \alpha t_1
\end{aligned}$$

for $t_2 > t_1$. Thus,

$$R_X(t_1,t_2) = \alpha\min\{t_1,t_2\}.$$