2-1 Set Theory

Definitions

- A *set* is a collection of objects called *elements*. Example

$$A = \{\text{car}, \text{apple}, \text{pencil}\}$$

- The *cardinality* of the set A is the number of elements in the set. The cardinality is often denoted $|A|$. For the previous example, $|A| = 3$.
- The notation that represents "ζ is a member of A" is $\zeta \in A$.
- The notation that represents "ζ is not a member of A" is $\zeta \notin A$.
- The *empty set* is the set that contains no elements. The empty set is denoted $\varnothing$.

$$|\varnothing| = 0.$$

- A *subset* B of a set A is anther set whose elements are also elements of A. This is denoted $B \subset A$ or $B \subseteq A$. The formal mathematical definition is

$$B \subseteq A \text{ means } \zeta \in B \Rightarrow \zeta \in A.$$

  If $|A| = n$, then there are 2^n subsets of A.

- In probability theory, all sets are subsets of a set $\mathcal{S}$ called a *space*. The *space* is the set of all possible experimental outcomes.
- For any set A,

$$\varnothing \subseteq A \subseteq \mathcal{S}$$

Set Properties and Operations

- *Transitivity*: if $C \subseteq B$ and $B \subseteq A$ then $C \subseteq A$.
- *Equality*: $A = B$ if and only if (iff) $B \subseteq A$ and $A \subseteq B$.
- The *union* of two sets A and B is a set whose elements are all elements of A or of B or of both.
$$A \cup B$$
  - *The union operation is commutative*: $A \cup B = B \cup A$.
  - *The union operation is associative*: $(A \cup B) \cup C = A \cup (B \cup C)$.
- The *intersection* of two sets A and B is the set comprising all the elements that are common to the sets A and B.
$$A \cap B$$
  - *The intersection operation is commutative*: $A \cap B = B \cap A$.
  - *The intersection operation is associative*: $(A \cap B) \cap C = A \cap (B \cap C)$.
  - *Intersection distributes over union*: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
- Two sets are *mutually exclusive* or *disjoint* if they have no common elements.
$$A \cap B = \varnothing.$$
Several sets $A_1, A_2, \ldots$ are mutually exclusive if
$$A_i \cap A_j = \varnothing \quad \text{for all } i \text{ and for every } j \neq i.$$
- A *partition* of the set $\mathcal{S}$ is a set of mutually exclusive subsets $A_1, \ldots, A_n$ whose union equals $\mathcal{S}$.
$$A_1 \cup \cdots \cup A_n = \mathcal{S} \qquad A_i \cap A_j = \varnothing \text{ for } i \neq j.$$
$$\bigcup_{i=1}^{n} A_i = \mathcal{S} \qquad A_i \cap A_j = \varnothing \text{ for } i \neq j.$$
- The *complement* of a set A, denoted $\overline{A}$ is the set comprising all elemets of $\mathcal{S}$ that are not in A.
- *De Morgan's Law*:
$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$
$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

Some Results

$$B \subseteq A \Rightarrow A \cup B = A \qquad \begin{aligned} A \cup A &= A \\ A \cup \varnothing &= A \\ \mathcal{S} \cup A &= \mathcal{S} \end{aligned}$$

$$B \subseteq A \Rightarrow A \cap B = B \qquad \begin{aligned} \varnothing \cap A &= \varnothing \\ A \cap \mathcal{S} &= A \end{aligned}$$

$$\left.\begin{aligned} A \cup \overline{A} &= \mathcal{S} \\ A \cap \overline{A} &= \varnothing \end{aligned}\right\} \text{ A and $\overline{A}$ form a \textit{partition} of $\mathcal{S}$}$$

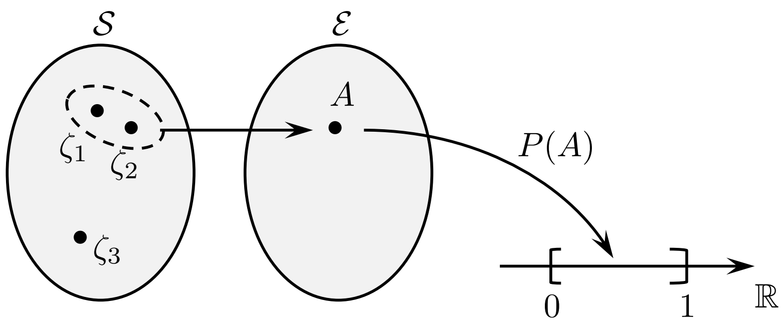
$$\overline{\overline{A}} = A$$

$$\overline{\mathcal{S}} = \varnothing$$

$$\overline{\varnothing} = \mathcal{S}$$

2-2 Probability Space

- Probability theory (in this course) is conceptualized as an experiment with possible *outcomes* $\zeta_1, \zeta_2, \ldots, \zeta_n$ where $\zeta_i \in \mathcal{S}$.
- An *event* is a subset of $\mathcal{S}$.
- The experimental outcome is *uncertain*.
- *Probability* is the measure of uncertainty.



$$
\begin{aligned}
\mathcal{S} &= \text{the set of experimental outcomes} \\
\mathcal{E} &= \text{the set of all events} \\
&= \text{the set of all subsets of } \mathcal{S} \\
P(A) &= \text{a probability measure of the event } A \\
&= P(\cdot) \text{ is a map } \mathcal{E} \to [0,1] \text{ in } \mathbb{R}
\end{aligned}
$$

Axioms of Probability for the event $A \subseteq \mathcal{S}$

1. $P(A) \geq 0$
2. $P(\mathcal{S}) = 1$
3. $A \cap B = \varnothing \quad \Rightarrow \quad P(A \cup B) = P(A) + P(B)$.

Prove these: $P(\varnothing) = 0$

$$P(\overline{A}) = 1 - P(A)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$B \subseteq A \Rightarrow P(A) \geq P(B)$$

2-3 Conditional Probability

Definition
The *conditional probability* of an event A assuming another event M is

$$P(A|M) = \frac{P(A \cap M)}{P(M)}$$

for $P(M) > 0$.

Properties

- $M \subseteq A \Rightarrow P(A|M) = 1$.
- $A \subseteq M \Rightarrow P(A|M) = \dfrac{P(A)}{P(M)}$.
- Conditional probabilities are probabilities: $P(A|M)$ satisfies the Axioms of Probability:
    1. $P(A|M) \geq 0$
    2. $P(\mathcal{S}|M) = 1$
    3. $A \cap B = \varnothing \quad \Rightarrow \quad P(A \cup B|M) = P(A|M) + P(B|M)$

Bayes Stuff for events A and B.

- Bayes Rule: $P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$
- Total Probability Theorem: Let $A_1, \ldots, A_n$ be a partition of $\mathcal{S}$. Then

$$P(B) = P(B|A_1)P(A_1) + \cdots + P(B|A_n)P(A_n).$$

- Bayes Theorem for the partition $A_1, \ldots, A_n$ of $\mathcal{S}$.

$$\begin{aligned} P(A_i|B) &= \frac{P(B|A_i)P(A_i)}{P(B)} \\ &= \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1) + \cdots + P(B|A_n)P(A_n)} \end{aligned}$$

Thomas Bayes (c. 1701 – 7 April 1761)
Presbyterian minister, statistician, and philosopher

Latinisms

Bayes' Rule $P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$

- $P(A|B)$ is called the *a posteriori* probability.
  - *a posteriori* is a Latin phrase that means "from the later."
  - *a posteriori knowledge* describes the knowledge obtained through experience or observation.
  - *a posteriori reasoning* describes the reasoning based on known facts or past events.
  - *a posteriori reasoning* draws general conclusions from specific (or particular) circumstances—the reasoning is from particular to general. An example is how court cases are argued: inferring the interior disposition of malice from the act of murder.
  - *a posteriori probability* quantifies the uncertainty about the event A given the fact observation that event B has occurred.
- $P(A)$ is called the *a priori* probability.
  - *a priori* is a Latin phrase that means "from the earlier."
  - *a priori knowledge* is independent from any experience and observation. It derives from pure reason.
  - *a priori reasoning* draws particular conclusions from general principles—the reasoning is from from general to specific. A familiar example is mathematical proofs where is a result is deduced from a principle.
  - *a priori probability* quantifies the uncertainty about the event A before any observations are made.
- $P(B|A)$ is called the *likelihood*.
- $P(B)$ does not have a name. The Total Probability Theorem in the context of Bayes' Theorem teaches that $P(B)$ may be expressed in terms of the *likelihood*.

Definition
Two events A and B are called *independent* if

$$P(A \cap B) = P(A)P(B).$$

Properties

- A and B are independent $\Rightarrow$ A and $\overline{B}$ are independent.
- A and B are independent $\Rightarrow$ $\overline{A}$ and $\overline{B}$ are independent.
- A and B are independent $\Rightarrow$ $P(A|B) = P(A)$.

Definition
Three events A_1, A_2, and A_3 are called (*mutually*) *independent* if

$$P(A_i \cap A_j) = P(A_i)P(A_j), \quad i \neq j$$

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3).$$

Properties

- $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2 \cap A_3)$
- A_1, A_2, A_3 are independent $\Rightarrow$ $P(A_1 \cap A_2 \cap \overline{A}_3) = P(A_1)P(A_2)P(\overline{A}_3)$.

Generalization
The independence of n events can be defined inductively. The events $A_1, \ldots, A_n$ are independent if any $k < n$ of them are independent and

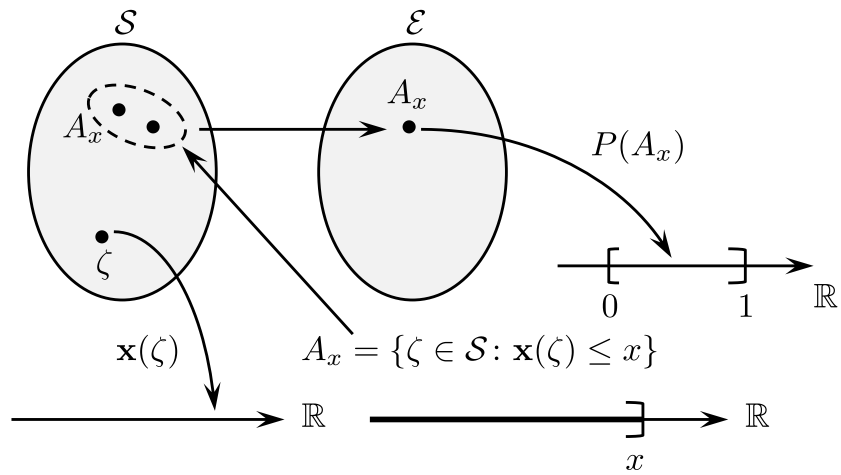
$$P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdots P(A_n)$$

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i).$$

Union Bound

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i).$$

4-1 The Concept of a Random Variable



Definition

- A *random variable* is a *map* from $\mathcal{S}$ to $\mathbb{R}$.
- For every outcome $\zeta \in \mathcal{S}$, $\mathbf{x}(\zeta)$ is a real number.
- When it is understood $\mathbf{x}(\zeta)$ is a random variable, it is customary to drop the notational dependence on ζ and write "$\mathbf{x}$ is a random variable."
- The *event* generated by the random variable $\mathbf{x}(\zeta)$ is

$$\{\zeta \in \mathcal{S} \colon \mathbf{x}(\zeta) \le x\}.$$

- The short-hand notation $\mathbf{x} \le x$ is usually used:

$$\mathbf{x} \le x \quad \text{means} \quad \{\zeta \in \mathcal{S} \colon \mathbf{x}(\zeta) \le x\}$$

"Thus $\{\mathbf{x} \le x\}$ is not a set of numbers but a *set of experimental outcomes*."

4-2 Distribution and Density Functions

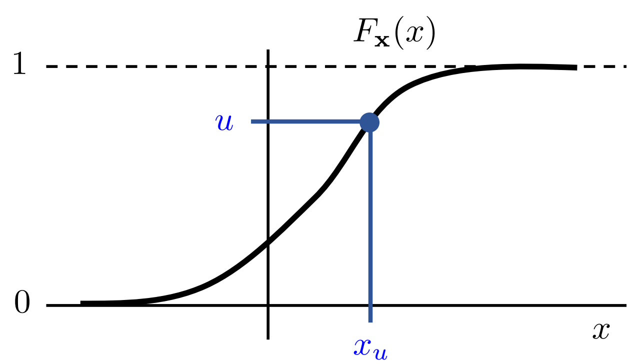
Definitions

- The elements of the set $\mathcal{S}$ contained in the event $\{\zeta\colon \mathbf{x}(\zeta) \le x\}$ change as the number x takes various values.
- The probability $P(\{\zeta\colon \mathbf{x}(\zeta) \le x\}) = P(\mathbf{x} \le x)$ is, therefore, a number that depends on x.
- This number is denoted $F_{\mathbf{x}}(x)$ and is called the (*cumulative*) *distribution function*.
- The distribution function $F_{\mathbf{x}}(x)$ is defined for every $-\infty < x < \infty$.

Percentiles

The u percentile of a rnaom variable $\mathbf{x}$ is the smallest number x_u such that

$$u = P(\mathbf{x} \le x_u) = F_{\mathbf{x}}(x_u).$$



Properties

1. End points.

$$F_{\mathbf{x}}(+\infty) = 1 \quad F_{\mathbf{x}}(-\infty) = 0.$$

2. $F_{\mathbf{x}}(x)$ is a nondecreasing function of x.

   if $x_1 < x_2$ then $F_{\mathbf{x}}(x_1) \leq F_{\mathbf{x}}(x_2)$

3. If $F_{\mathbf{x}}(x_0) = 0$ then $F_{\mathbf{x}}(x) = 0$ for every $x \leq x_0$.

4. $P(\mathbf{x} > x) = 1 - F_{\mathbf{x}}(x)$.

5. $F_{\mathbf{x}}(x)$ is continuous from the right:

$$\lim_{\epsilon \to 0} F_{\mathbf{x}}(x + \epsilon) = F_{\mathbf{x}}(x).$$

6. $P(x_1 < \mathbf{x} \leq x_2) = F_{\mathbf{x}}(x_2) - F_{\mathbf{x}}(x_1)$

7. $P(\mathbf{x} = x) = F_{\mathbf{x}}(x) - \lim_{\epsilon \to 0} F_{\mathbf{x}}(x - \epsilon)$.

8. $P(x_1 \leq \mathbf{x} \leq x_2) = F_{\mathbf{x}}(x_2) - \lim_{\epsilon \to 0} F_{\mathbf{x}}(x_1 - \epsilon)$.

$F_{\mathbf{x}}(x)$ is always continuous from the right.

$F_{\mathbf{x}}(x)$ is does not have to be continuous from the left.

If $F_{\mathbf{x}}(x)$ is not continuous, then the left and right limits are different:

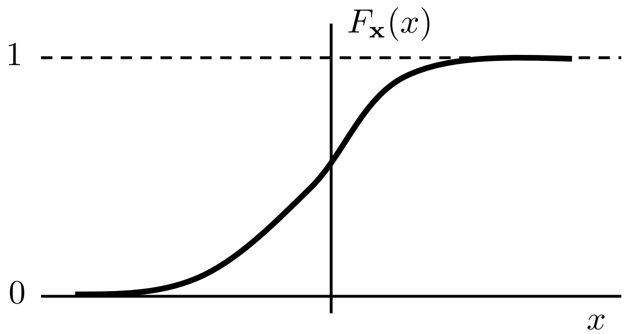
$$\lim_{\epsilon \to 0} F_{\mathbf{x}}(x + \epsilon) - \lim_{\epsilon \to 0} F_{\mathbf{x}}(x - \epsilon) > 0$$

The difference is > 0 because of Property 2.

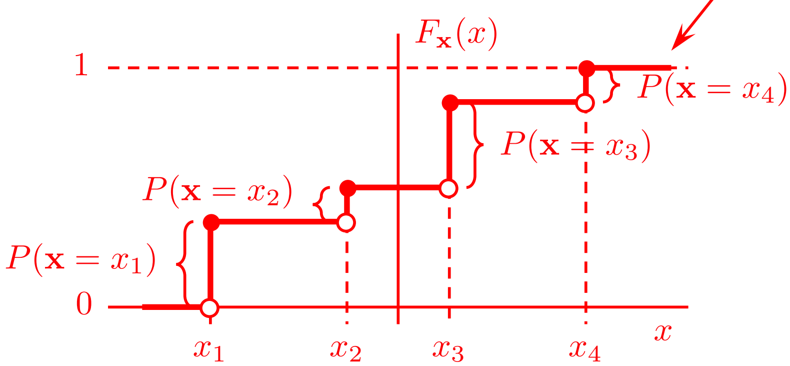
From Property 7, the difference is $P(\mathbf{x} = x)$

The only discontinuities of $F_{\mathbf{x}}(x)$ are of the jump type.

continuous distribution function



discontinuous distribution function



$F_{\mathbf{x}}(x)$ is continuous: $P(\mathbf{x} = x) = 0$

$F_{\mathbf{x}}(x)$ is discontinuous at x_1, x_2, x_3, x_4:

$$P(\mathbf{x} = x) = \begin{cases} 0 & x \notin \{x_1, x_2, x_3, x_4\} \\ > 0 & x \in \{x_1, x_2, x_3, x_4\} \end{cases}$$

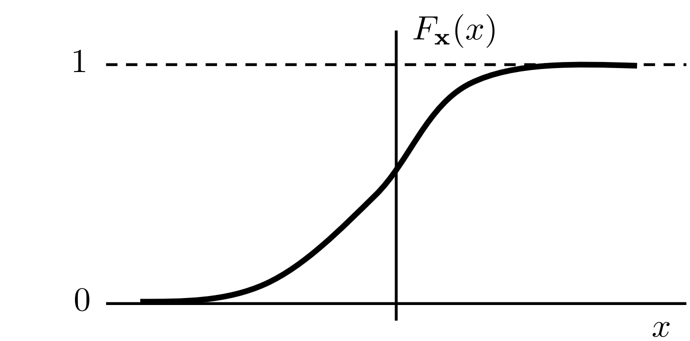
Probability Density Function (pdf)

Definition $f_{\mathbf{x}}(x) = \dfrac{d}{dx} F_{\mathbf{x}}(x)$

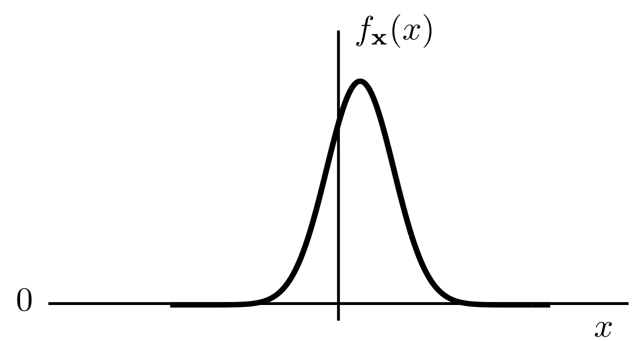
Properties

1. $f_{\mathbf{x}}(x) > 0$ for all x.
2. $F_{\mathbf{x}}(x) = \displaystyle\int_{-\infty}^{x} f_{\mathbf{x}}(u)du.$
3. $\displaystyle\int_{-\infty}^{\infty} f_{\mathbf{x}}(u)du = 1.$
4. $P(x_1 < \mathbf{x} \leq x_2) = \displaystyle\int_{x_1}^{x_2} f_{\mathbf{x}}(u)du.$

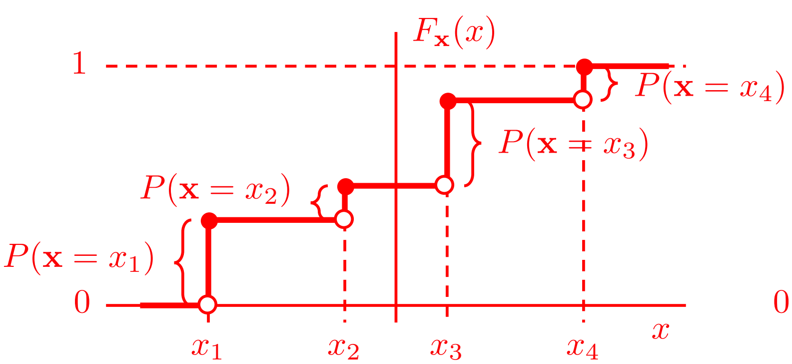
continuous distribution function



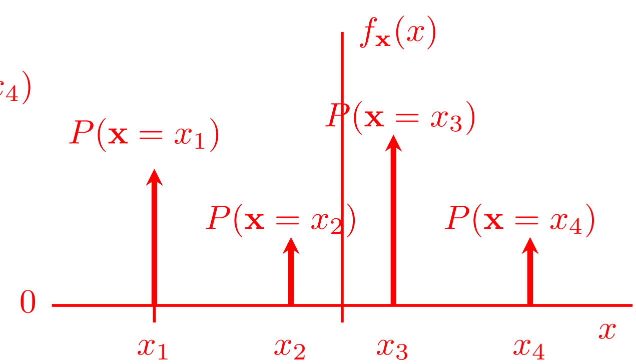
“continuous random variable”



discontinuous distribution function



“discrete random variable”



Probability Density Function: $f_{\mathbf{x}}(x) = \displaystyle\sum_{i=1}^{4} P(\mathbf{x} = x_i)\delta(x - x_i)$

Probability Mass Function: $P(\mathbf{x} = x_1),\ P(\mathbf{x} = x_2),\ P(\mathbf{x} = x_3),\ P(\mathbf{x} = x_4)$

4-4 Conditional Distributions

Definition

- The *conditional distribution* $F_{\mathbf{x}|M}(x|M)$ of a random variable $\mathbf{x}$ assuming the event M is defined as the conditional probability of the event $\mathbf{x} \le x$:

$$F_{\mathbf{x}|M}(x|M) = P(\mathbf{x} \le x|M) = \frac{P(\mathbf{x} \le x, M)}{P(M)}$$

  where the notation $(\mathbf{x} \le x, M)$ means

$$\{\zeta \in \mathcal{S} \colon \mathbf{x}(\zeta) \le x\} \cap \{\zeta \in M\}.$$

- To find $F_{\mathbf{x}|M}(x|M)$ one must, in general, know the underlying experiment.
- However, if the event M is an event that can be expressed in terms of the raon variable $\mathbf{x}$, then, for the determination of $F_{\mathbf{x}|M}(x|M)$, knowledge of $F_{\mathbf{x}}(x)$ is sufficient.

  1. $M = \{\zeta \in \mathcal{S} \colon \mathbf{x}(\zeta) \le a\}$: $F_{\mathbf{x}|\mathbf{x}\le a}(x|\mathbf{x} \le a) = \begin{cases} 1 & x \ge a \\ \dfrac{F_{\mathbf{x}}(x)}{F_{\mathbf{x}}(a)} & x < a \end{cases}$
  2. $M = \{\zeta \in \mathcal{S} \colon b < \mathbf{x}(\zeta) \le a\}$: $F_{\mathbf{x}|b<\mathbf{x}\le a}(x|b < \mathbf{x} \le a) = \begin{cases} 1 & x \ge a \\ \dfrac{F_{\mathbf{x}}(x) - F_{\mathbf{x}}(b)}{F_{\mathbf{x}}(a) - F_{\mathbf{x}}(b)} & b \le x < a \\ 0 & x < b \end{cases}$

Properties

- The *conditional distribution* $F_{\mathbf{x}|M}(x|M)$ is a distribution function.
- $F_{\mathbf{x}|M}(x|M)$ possesses all 8 properties of a distribution.

Definition

The *conditional density* $f_{\mathbf{x}|M}(x|M)$ of a random variable $\mathbf{x}$ assuming the event M is the derivative of the conditional distribution:

$$f_{\mathbf{x}|M}(x|M) = \frac{d}{dx} F_{\mathbf{x}|M}(x|M) = \lim_{\Delta x \to 0} \frac{P(x < \mathbf{x} \le x + \Delta x|M)}{\Delta x}$$

Total Probability and Bayes' Theorem

- Bayes' Rule for a random variable $\mathbf{x}$:

$$f_{\mathbf{x}|A}(x|A) = \frac{P(A|\mathbf{x}=x)}{P(A)} f_{\mathbf{x}}(x)$$
$$P(A|\mathbf{x}=x) = \frac{f_{\mathbf{x}|A}(x|A)P(A)}{f_{\mathbf{x}}(x)}.$$

- Total Probability Theorem for a random variable $\mathbf{x}$:

$$P(A) = \int_{-\infty}^{\infty} P(A|\mathbf{x}=x) f_{\mathbf{x}}(x) dx.$$

- Bayes' Theorem for a random variable $\mathbf{x}$:

$$f_{\mathbf{x}|A}(x|A) = \frac{P(A|\mathbf{x}=x) f_{\mathbf{x}}(x)}{\displaystyle\int_{-\infty}^{\infty} P(A|\mathbf{x}=x) f_{\mathbf{x}}(x) dx}$$

5-1 The Random Variable $g(x)$

Definition

- $g(x)$ is a real-valued function of the real-valued variable x.
- $\mathbf{x} = \mathbf{x}(\zeta)$ is a random variable.
- $g(\mathbf{x})$ is a new random variable.
  - $\mathbf{x}(\zeta)$ is a map $\mathcal{S} \to \mathbb{R}$.
  - For each $\zeta \in \mathcal{S}$, $\mathbf{x}(\zeta)$ is a real number.
  - $g(\mathbf{x}(\zeta))$ is another real number.
  - $g(\mathbf{x}(\zeta))$ is a composite map $\mathcal{S} \to \mathbb{R}$.
  - This composite map is called $\mathbf{y}(\zeta)$: the random variable $\mathbf{y}$.
- $\mathbf{y} = g(\mathbf{x})$ is a random variable defined by the event $\{\zeta\colon \mathbf{y}(\zeta) \leq y\}$ and its probability (the distribution function)

$$F_{\mathbf{y}}(y) = P(\{\zeta\colon \mathbf{y}(\zeta) \leq y\}) = P(\mathbf{y} \leq y).$$

- The event $\{\zeta\colon \mathbf{y}(\zeta) \leq y\}$ may also be written $\{\zeta \in \mathcal{S}\colon g(\mathbf{x}(\zeta)) \leq y\}$ so that the distribution function may also be written

$$F_{\mathbf{y}}(y) = P(\{\zeta \in \mathcal{S}\colon g(\mathbf{x}(\zeta)) \leq y\}) = P(g(\mathbf{x}) \leq y).$$

- Let $D_y = \{x \in \mathbb{R}\colon g(x) \leq y\}$. Then

$$\{\zeta \in \mathcal{S}\colon g(\mathbf{x}(\zeta)) \leq y\} = \{\zeta \in \mathcal{S}\colon \mathbf{x}(\zeta) \in D_y\}$$

  The probability of this event is

$$\begin{aligned} F_{\mathbf{y}}(y) &= P(\{\zeta \in \mathcal{S}\colon g(\mathbf{x}(\zeta)) \leq y\}) \\ &= P(\{\zeta \in \mathcal{S}\colon \mathbf{x}(\zeta) \in D_y\}) \\ &= \int_{D_y} f_{\mathbf{x}}(x)dx \end{aligned}$$

- Probability density function of $\mathbf{y}$

$$f_{\mathbf{y}}(y) = \frac{d}{dy} F_{\mathbf{y}}(y) = \frac{d}{dy} \int_{D_y} f_{\mathbf{x}}(x)dx$$

5-2 The Distribution of $g(x)$

Examples

1. $\mathbf{y} = a\mathbf{x} + b$

2. $\mathbf{y} = \mathbf{x}^2$

3. $\mathbf{y} = e^{\mathbf{x}}$

4. $\mathbf{y} = \begin{cases} 1 & \mathbf{x} > 0 \\ -1 & \mathbf{x} \leq 0 \end{cases}$

5. $\mathbf{y} = \begin{cases} \sqrt{\mathbf{x}} & \mathbf{x} > 0 \\ 0 & \mathbf{x} \leq 0 \end{cases}$

6. $\mathbf{y} = \begin{cases} 1 & \mathbf{x} > 1 \\ \mathbf{x} & -1 \leq \mathbf{x} < 1 \\ -1 & \mathbf{x} \leq -1 \end{cases}$

5-3 Mean and Variance

Definitions

- The *expected value* or *mean* of the random variable $\mathbf{x}$ is

$$E\{\mathbf{x}\} = \begin{cases} \displaystyle\int_{-\infty}^{\infty} x f_{\mathbf{x}}(x)dx & \text{continuous RV} \\ \displaystyle\sum_i x_i P(\mathbf{x} = x_i) & \text{discrete RV} \end{cases}$$

- The *conditional mean* of the random variable $\mathbf{x}$ assuming an event M is

$$E\{\mathbf{x}|M\} = \begin{cases} \displaystyle\int_{-\infty}^{\infty} x f_{\mathbf{x}|M}(x|M)dx & \text{continuous RV} \\ \displaystyle\sum_i x_i P(\mathbf{x} = x_i|M) & \text{discrete RV} \end{cases}$$

- The *mean of* $g(\mathbf{x})$ can be computed two ways.

  1. Write $\mathbf{y} = g(\mathbf{x})$ and find the pdf $f_{\mathbf{y}}(y)$ or pmf $P(\mathbf{y} = y_i)$. The mean is

$$E\{g(\mathbf{x})\} = E\{\mathbf{y}\} = \begin{cases} \displaystyle\int_{-\infty}^{\infty} y f_{\mathbf{y}}(y)dy & \text{continuous RV} \\ \displaystyle\sum_i y_i P(\mathbf{y} = y_i) & \text{discrete RV} \end{cases}$$

  2. Apply the *Law of the Unconscious Statistician*

$$E\{g(\mathbf{x})\} = \begin{cases} \displaystyle\int_{-\infty}^{\infty} g(x) f_{\mathbf{x}}(x)dx & \text{continuous RV} \\ \displaystyle\sum_i g(x_i) P(\mathbf{x} = x_i) & \text{discrete RV} \end{cases}$$

- The *variance* of the random variable $\mathbf{x}$ with mean $\mu_{\mathbf{x}} = E\{\mathbf{x}\}$ is

$$E\{(\mathbf{x} - \mu_{\mathbf{x}})^2\} = \begin{cases} \displaystyle\int_{-\infty}^{\infty} (x - \mu_{\mathbf{x}})^2 f_{\mathbf{x}}(x)dx & \text{continuous RV} \\ \displaystyle\sum_i (x_i - \mu_{\mathbf{x}})^2 P(\mathbf{x} = x_i) & \text{discrete RV} \end{cases}$$

5-4 Moments

Definitions

- The n-th *moment* of the random variable $\mathbf{x}$ is

$$m_n = E\{\mathbf{x}^n\} = \begin{cases} \displaystyle\int_{-\infty}^{\infty} x^n f_{\mathbf{x}}(x)dx & \text{continuous RV} \\ \displaystyle\sum_i x_i^n P(\mathbf{x} = x_i) & \text{discrete RV} \end{cases}$$

- The n-th *central moment* of the random variable $\mathbf{x}$ is

$$\mu_n = E\{(\mathbf{x} - m_1)^n\} = \begin{cases} \displaystyle\int_{-\infty}^{\infty} (x - m_1)^n f_{\mathbf{x}}(x)dx & \text{continuous RV} \\ \displaystyle\sum_i (x_i - m_1)^n P(\mathbf{x} = x_i) & \text{discrete RV} \end{cases}$$

- The n-th *absolute moment* of the random variable $\mathbf{x}$ is $E\{|\mathbf{x}|^n\}$.
- The n-th *absolute central moment* of the random variable $\mathbf{x}$ is $E\{|\mathbf{x} - m_1|^n\}$.
- The n-th *generalized moment* of the random variable $\mathbf{x}$ is $E\{(\mathbf{x} - a)^n\}$.
- The n-th *generalized absolute moment* of the random variable $\mathbf{x}$ is $E\{|\mathbf{x} - a|^n\}$.

5-5 Characteristic Functions

Definitions

- The *characteristic function* of the random variable $\mathbf{x}$ is

$$\Phi_{\mathbf{x}}(\omega) = \begin{cases} \displaystyle\int_{-\infty}^{\infty} f_{\mathbf{x}}(x) e^{j\omega x} dx & \text{continuous RV} \\ \displaystyle\sum_i P(\mathbf{x} = x_i) e^{j\omega x_i} & \text{discrete RV} \end{cases}$$

- The inversion formula are

$$f_{\mathbf{x}}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_{\mathbf{x}}(\omega) e^{-j\omega x} d\omega \qquad \text{continuous RV}$$

$$P(\mathbf{x} = x_i) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{\mathbf{x}}(\omega) e^{-j\omega x_i} d\omega \qquad \text{discrete RV}$$

Properties

- From the Law of the Unconscious Statistician, the characteristic function may be defined as an expectation:

$$\Phi_{\mathbf{x}}(\omega) = E\{e^{j\omega \mathbf{x}}\}.$$

- Characteristic functions are used to analyze sums of random variables. Sums of random variables are examined in Chapter 6.

6-1 Bivariate Distributions

Definitions

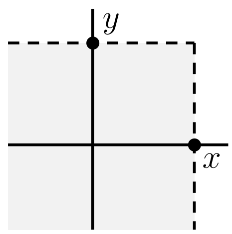
- The *joint* (*bivariate*) *distribution* of two random variables $\mathbf{x}$ and $\mathbf{y}$ is

$$F_{\mathbf{xy}}(x,y) = P(\{\zeta \in \mathcal{S}\colon \mathbf{x}(\zeta) \le x\} \cap \{\zeta \in \mathcal{S}\colon \mathbf{y}(\zeta) \le y\})$$

- Shorthand notation: $F_{\mathbf{xy}}(x,y) = P(\mathbf{x} \le x, \mathbf{y} \le y)$

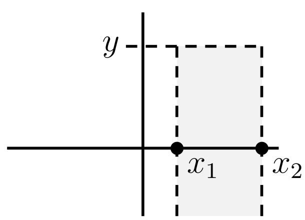
Properties

0 $F\mathbf{xy}(x,y)$ is the probability $\mathbf{x}$ and $\mathbf{y}$ are in an open rectangle.

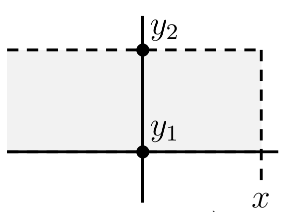


1 $F\mathbf{xy}(-\infty,y) = 0$, $F\mathbf{xy}(x,-\infty) = 0$, $F\mathbf{xy}(\infty,\infty) = 1$.

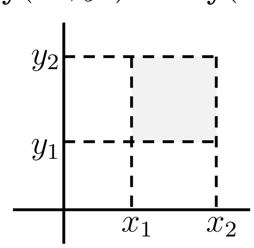
2a $P(x_1 < \mathbf{x} \le x_2, y_1 < \mathbf{y} \le y_2) = F_{\mathbf{xy}}(x_2,y) - F_{\mathbf{xy}}(x_1,y)$



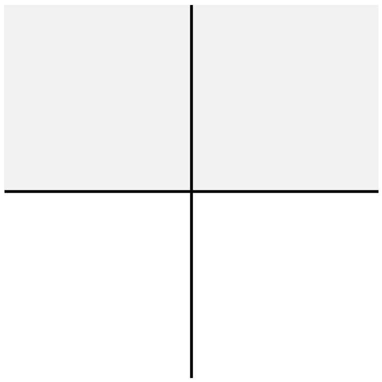
2b $P(\mathbf{x} \le x, y_1 < \mathbf{y} \le y_2) = F_{\mathbf{xy}}(x,y_2) - F_{\mathbf{xy}}(x,y_1)$

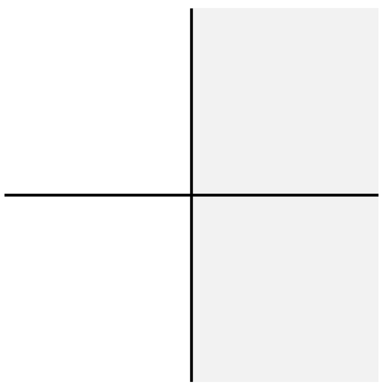


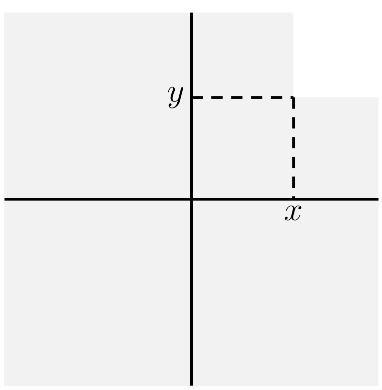
3 $P(x_1 < \mathbf{x} \le x_2, y_1 < \mathbf{y} \le y_2)$
$= F_{\mathbf{xy}}(x_2,y_2) - F_{\mathbf{xy}}(x_1,y_2) - F_{\mathbf{xy}}(x_2,y_1) + F_{\mathbf{xy}}(x_1,y_1)$.



Can you work out these?







Definitions

- The *joint density* of $\mathbf{x}$ and $\mathbf{y}$ is defined by the function

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

- If $F_{\mathbf{xy}}(x,y)$ has step discontinuities, then the pdf $f_{\mathbf{xy}}(x,y)$ contains impulses (Dirac deltas). Alternatively, the joint *probability mass function* can be used:

$$P(\mathbf{x} = x_i, \mathbf{y} = y_k)$$

Properties

1. $F_{\mathbf{xy}}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{\mathbf{xy}}(u,v)\, dv\, du$

2. Joint Statistics

$$P((\mathbf{x},\mathbf{y}) \in D) = \iint_D f_{\mathbf{xy}}(x,y)\, dx\, dy$$

3. Marginal Statistics

   (a) Marginal distribution and density/pmf of $\mathbf{x}$

$$\begin{aligned} F_{\mathbf{x}}(x) &= F_{\mathbf{xy}}(x,\infty) \\ f_{\mathbf{x}}(x) &= \int_{-\infty}^{\infty} f_{\mathbf{xy}}(x,y)\, dy && \text{jointly continuous RVs} \\ P(\mathbf{x} = x_i) &= \sum_k P(\mathbf{x} = x_i, \mathbf{y} = y_k) && \text{jointly discrete RVs} \end{aligned}$$

   (b) Marginal distribution and density/pmf of $\mathbf{y}$

$$\begin{aligned} F_{\mathbf{y}}(y) &= F_{\mathbf{xy}}(\infty,y) \\ f_{\mathbf{y}}(y) &= \int_{-\infty}^{\infty} f_{\mathbf{xy}}(x,y)\, dx && \text{jointly continuous RVs} \\ P(\mathbf{y} = y_k) &= \sum_i P(\mathbf{x} = x_i, \mathbf{y} = y_k) && \text{jointly discrete RVs} \end{aligned}$$

Definition

Two random variables $\mathbf{x}$ and $\mathbf{y}$ are called (*statistically*) *independent* if the events $\{\zeta \in \mathcal{S} \colon \mathbf{x}(\zeta) \in A\}$ and $\{\zeta \in \mathcal{S} \colon \mathbf{y}(\zeta) \in B\}$ are independent, that is, if (using shorthand notation)

$$P(\mathbf{x} \in A, \mathbf{y} \in B) = P(\mathbf{x} \in A)\, P(\mathbf{y} \in B)$$

Properties

1. If $A = \{\zeta \in \mathcal{S} \colon \mathbf{x}(\zeta) \leq x\}$ and $B = \{\zeta \in \mathcal{S} \colon \mathbf{y}(\zeta) \leq y\}$ then independence means
$$F_{\mathbf{xy}}(x, y) = F_{\mathbf{x}}(x) F_{\mathbf{y}}(y)$$
2. From property 1 for jointly continuous RVs
$$f_{\mathbf{xy}}(x, y) = f_{\mathbf{x}}(x) f_{\mathbf{y}}(y)$$
3. From property 1 (or the definition) for jointly discrete RVs
$$P(\mathbf{x} = x_i, \mathbf{y} = y_k) = P(\mathbf{x} = x_i)\, P(\mathbf{y} = y_k)$$
4. Theorem 6-1: If the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent then the random variables
$$\mathbf{z} = g(\mathbf{x}) \qquad \mathbf{w} = h(\mathbf{y})$$
are also independent.
5. Theorem 6-2: Let the random variable $\mathbf{x}$ be defined by experiment $\mathcal{S}_1$ and the random variable $\mathbf{y}$ by experiment $\mathcal{S}_2$. If the experiments $\mathcal{S}_1$ and $\mathcal{S}_2$ are independent, then the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent.

Definition

Joint Normality: Two random variables $\mathbf{x}$ and $\mathbf{y}$ are called *jointly normal* if the joint density is given by [(6-23)–(6-24)]

$$f_{\mathbf{xy}}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \exp\left\{ -\frac{1}{2(1-r^2)} \left(\frac{(x-\mu_1)^2}{\sigma_1^2} - 2r\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right) \right\}$$

for $|r| < 1$.

Properties

1. The form (6-23)–(6-24) is horrific: one cannot tell what is going on. MDR prefers the form

   $$f_{\mathbf{xy}}(x,y) = \frac{1}{2\pi\sqrt{\det(C)}} \exp\left\{ -\frac{1}{2} \begin{bmatrix} (x-\mu_1) & (y-\mu_2) \end{bmatrix} C^{-1} \begin{bmatrix} x-\mu_1 \\ y-\mu_2 \end{bmatrix} \right\}$$

   where

   $$C = \begin{bmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

   is the *covariance matrix.*

2. The *marginal density* of $\mathbf{x}$ is

   $$f_{\mathbf{x}}(x) = \int_{-\infty}^{\infty} f_{\mathbf{xy}}(x,y)\,dy = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left\{ -\frac{(x-\mu_1)^2}{2\sigma_1^2} \right\}$$

3. The *marginal density* of $\mathbf{y}$ is

   $$f_{\mathbf{y}}(y) = \int_{-\infty}^{\infty} f_{\mathbf{xy}}(x,y)\,dxy = \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left\{ -\frac{(y-\mu_2)^2}{2\sigma_2^2} \right\}$$

4. if $r = 0$ in (6-23)–(6-24) then $f_{\mathbf{xy}}(x,y) = f_{\mathbf{x}}(x)\, f_{\mathbf{y}}(y)$

Definition

We say that the joint density of two random variables $\mathbf{x}$ and $\mathbf{y}$ is *circularly symmetric* (or *symmetrical*) if it depends only on its distance from the origin, that is if

$$f_{\mathbf{xy}}(x,y) = g(r) \qquad r = \sqrt{x^2+y^2}.$$

Note: this r is not the same r used in (6-23)–(6-24)!

6-2 One Function of Two Random Variables

Let $\mathbf{x}$ and $\mathbf{y}$ be two random variables with joint pdf $f_{\mathbf{xy}}(x,y)$ and let $g : \mathbb{R}^2 \to \mathbb{R}$ be a function denoted $z = g(x,y)$. The meaning of $\mathbf{z} = g(\mathbf{x},\mathbf{y})$.

- $\zeta \in \mathcal{S}$ is an outcome.
- $\mathbf{x}(\zeta)$ is a real number.
- $\mathbf{y}(\zeta)$ is a real number.
- $g(\mathbf{x}(\zeta),\mathbf{y}(\zeta))$ is a real number.
- $g(\mathbf{x}(\zeta),\mathbf{y}(\zeta))$ is a composite map $\mathcal{S} \to \mathbb{R}$.
- Call this composite map $\mathbf{z}(\zeta)$.
- $\mathbf{z}(\zeta) = g(\mathbf{x}(\zeta),\mathbf{y}(\zeta))$ is a random variable.
- The events of $\mathbf{z}(\zeta)$ are described by the set

$$\{\zeta \in \mathcal{S} \colon \mathbf{z}(\zeta) \le z\} = \{\zeta \in \mathcal{S} \colon g(\mathbf{x}(\zeta),\mathbf{y}(\zeta)) \le z\}$$

- Let $D_z = \{(x,y) \in \mathbb{R}^2 \colon g(x,y) \le z\}$. Then

$$\{\zeta \in \mathcal{S} \colon \mathbf{z}(\zeta) \le z\} = \{\zeta \in \mathcal{S} \colon g(\mathbf{x}(\zeta),\mathbf{y}(\zeta)) \le z\} = \Big\{\zeta \in \mathcal{S} \colon \Big(\mathbf{x}(\zeta),\mathbf{y}(\zeta)\Big) \in D_z\Big\}$$

- Cumulative distribution function of $\mathbf{z}$:

$$\begin{aligned} F_{\mathbf{z}}(z) &= P(\{\zeta \in \mathcal{S} \colon \mathbf{z}(\zeta) \le z\}) \\ &= P(\{\zeta \in \mathcal{S} \colon g(\mathbf{x}(\zeta),\mathbf{y}(\zeta)) \le z\}) \\ &= P\left(\Big\{\zeta \in \mathcal{S} \colon \Big(\mathbf{x}(\zeta),\mathbf{y}(\zeta)\Big) \in D_z\Big\}\right) \\ &= \iint_{D_z} f_{\mathbf{xy}}(x,y)\,dx\,dy \end{aligned}$$

- Probability density function of $\mathbf{z}$:

$$f_{\mathbf{z}}(z) = \frac{d}{dz}F_{\mathbf{z}}(z) = \frac{d}{dz}\iint_{D_z} f_{\mathbf{xy}}(x,y)\,dx\,dy$$

6-4 Joint Moments

Given two random variables $\mathbf{x}$ and $\mathbf{y}$ and a function $g(x, y)$ $(g\colon \mathbb{R}^2 \to \mathbb{R})$, we form the random variable $\mathbf{z} = g(\mathbf{x}, \mathbf{y})$.

- The expected value of $\mathbf{z}$ is

$$E\{\mathbf{z}\} = \begin{cases} \displaystyle\int_{-\infty}^{\infty} z f_{\mathbf{z}}(z)\, dz & \text{continuous RV} \\ \displaystyle\sum_{\ell} z_\ell P(\mathbf{z} = z_\ell) & \text{discrete RV} \end{cases}$$

- From the Law of the Unconscious Statistician

$$E\{\mathbf{z}\} = \begin{cases} \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f_{\mathbf{xy}}(x, y)\, dx\, dy & \text{jointly continuous RVs} \\ \displaystyle\sum_{i}\sum_{k} g(x_i, y_k) P(\mathbf{x} = x_i, \mathbf{y} = y_k) & \text{jointly discrete RVs} \end{cases}$$

- Linearity

$$E\left\{\sum_{k=1}^{n} a_k g_k(\mathbf{x}, \mathbf{y})\right\} = \sum_{k=1}^{n} a_k E\left\{g_k(\mathbf{x}, \mathbf{y})\right\}$$

  Consequences

$$E\{\mathbf{x} + \mathbf{y}\} = E\{\mathbf{x}\} + E\{\mathbf{y}\}$$

$$E\{\mathbf{xy}\} \neq E\{\mathbf{x}\}E\{\mathbf{y}\} \quad \text{in general}$$

Definitions

$\mathbf{x}$ and $\mathbf{y}$ are two random variables with

$$E\{\mathbf{x}\} = \mu_{\mathbf{x}} \qquad E\{(\mathbf{x}-\mu_{\mathbf{x}})^2\} = \sigma_{\mathbf{x}}^2$$
$$E\{\mathbf{y}\} = \mu_{\mathbf{y}} \qquad E\{(\mathbf{y}-\mu_{\mathbf{y}})^2\} = \sigma_{\mathbf{y}}^2$$

- The *covariance* $C_{\mathbf{xy}}$ is

$$C_{\mathbf{xy}} = E\{(\mathbf{x}-\mu_{\mathbf{x}})(\mathbf{y}-\mu_{\mathbf{y}})\}$$

- The *correlation coefficient* is

$$\rho_{\mathbf{xy}} = \frac{C_{\mathbf{xy}}}{\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}} \qquad -1 \le \rho_{\mathbf{xy}} \le 1$$

- The *correlation* is

$$R_{\mathbf{xy}} = E\{\mathbf{xy}\}$$

- $\mathbf{x}$ and $\mathbf{y}$ are *uncorrelated* means

$$C_{\mathbf{xy}} = 0 \qquad \rho_{\mathbf{xy}} = 0 \qquad E\{\mathbf{xy}\} = E\{\mathbf{x}\}E\{\mathbf{y}\}$$

- $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal* means

$$R_{\mathbf{xy}} = 0$$

- Theorem 6-5: If two random variables $\mathbf{x}$ and $\mathbf{y}$ are independent $[f_{\mathbf{xy}}(x,y) = f_{\mathbf{x}}(x)f_{\mathbf{y}}(y)]$ then they are uncorrelated.

- Variance of the sum $\mathbf{z} = \mathbf{x} + \mathbf{y}$:

$$\sigma_{\mathbf{z}}^2 = \sigma_{\mathbf{x}}^2 + 2\rho_{\mathbf{xy}}\sigma_{\mathbf{x}}\sigma_{\mathbf{y}} + \sigma_{\mathbf{y}}^2$$

- *Joint moments*

$$m_{kr} = E\{\mathbf{x}^k\mathbf{y}^r\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^k y^r f_{\mathbf{xy}}(x,y)\,dx\,dy$$

- *Joint central moments*:

$$\mu_{kr} = E\left\{(\mathbf{x}-m_{10})^k\,(\mathbf{y}-m_{01})^r\right\}$$

Comments

$\mathbf{x}$ and $\mathbf{y}$ are two random variables with

$$E\{\mathbf{x}\} = \mu_{\mathbf{x}} \qquad E\{(\mathbf{x}-\mu_{\mathbf{x}})^2\} = \sigma_{\mathbf{x}}^2$$
$$E\{\mathbf{y}\} = \mu_{\mathbf{y}} \qquad E\{(\mathbf{y}-\mu_{\mathbf{y}})^2\} = \sigma_{\mathbf{y}}^2$$

- Comment on covariance

$$C_{\mathbf{xy}} = R_{\mathbf{xy}} - \mu_{\mathbf{x}}\mu_{\mathbf{y}}$$

- Comment on jointly normal random variables $\mathbf{x}$ and $\mathbf{y}$.

  MDR's preferred form is

$$f_{\mathbf{xy}}(x,y) = \frac{1}{2\pi\sqrt{\det(C)}} \exp\left\{-\frac{1}{2}\begin{bmatrix}(x-\mu_{\mathbf{x}}) & (y-\mu_{\mathbf{y}})\end{bmatrix} C^{-1}\begin{bmatrix}x-\mu_{\mathbf{x}}\\ y-\mu_{\mathbf{y}}\end{bmatrix}\right\}$$

  where

$$C = \begin{bmatrix}\sigma_{\mathbf{x}}^2 & r\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}\\ r\sigma_{\mathbf{x}}\sigma_{\mathbf{y}} & \sigma_{\mathbf{y}}^2\end{bmatrix}$$

  The parameter r here is defined by

$$r\sigma_{\mathbf{x}}\sigma_{\mathbf{y}} = E\left\{(\mathbf{x}-\mu_{\mathbf{x}})(\mathbf{y}-\mu_{\mathbf{y}})\right\} = C_{\mathbf{xy}}.$$

  That is, $r = \rho_{\mathbf{xy}}$.

- Comment on Theorem 6-5

  If two random variables are uncorrelated they are not necessarily independent. However, for normal random variables uncorrelatedness is equivalent to independence.

- Comment on moments

  For the determination of the joint statistics of $\mathbf{x}$ and $\mathbf{y}$ knowledge of their joint density is required. However, in many applications, only the first- and second-moments are used. These moments are determined in terms of the five parameters

$$\mu_{\mathbf{x}} \quad \mu_{\mathbf{y}} \quad \sigma_{\mathbf{x}}^2 \quad \sigma_{\mathbf{y}}^2 \quad \rho_{\mathbf{xy}}$$

  If $\mathbf{x}$ and $\mathbf{y}$ are jointly normal, then these parameters determine uniquely $f_{\mathbf{xy}}(x,y)$.

6-5 Joint Characteristic Functions

Definition

The *joint characteristic function* of the random variables $\mathbf{x}$ and $\mathbf{y}$ is

$$\Phi_{\mathbf{xy}}(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\mathbf{xy}}(x, y) e^{j(\omega_1 x + \omega_2 y)} dx\, dy$$

Properties

1. The *inversion formula* is

$$f_{\mathbf{xy}}(x, y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi_{\mathbf{xy}}(\omega_1, \omega_2) e^{-j(\omega_1 x + \omega_2 y)} d\omega_1\, d\omega_2$$

2. The *marginal characteristic functions* are

$$\Phi_{\mathbf{x}}(\omega) = \Phi_{\mathbf{xy}}(\omega, 0) \qquad \Phi_{\mathbf{y}}(\omega) = \Phi_{\mathbf{xy}}(0, \omega)$$

3. If $\mathbf{z} = a\mathbf{x} + b\mathbf{y}$ then

$$\Phi_{\mathbf{z}}(\omega) = E\left\{e^{j(a\mathbf{x} + b\mathbf{y})}\right\} = \Phi_{\mathbf{xy}}(a\omega, b\omega).$$

4. Independence. If the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent

$$E\left\{e^{j(\omega_1 \mathbf{x} + \omega_2 \mathbf{y})}\right\} = E\left\{e^{j\omega_1 \mathbf{x}}\right\} E\left\{e^{j\omega_2 \mathbf{y}}\right\}$$

$$\Rightarrow \quad \Phi_{\mathbf{xy}}(\omega_1, \omega_2) = \Phi_{\mathbf{x}}(\omega_1)\Phi_{\mathbf{y}}(\omega_2)$$

5. Convolution. If the random variables $\mathbf{x}$ and $\mathbf{y}$ are independent and $\mathbf{z} = \mathbf{x} + \mathbf{y}$, then

$$E\left\{e^{j\omega \mathbf{z}}\right\} = E\left\{e^{j\omega(\mathbf{x} + \mathbf{y})}\right\} = E\left\{e^{j\omega \mathbf{x}}\right\} E\left\{e^{j\omega \mathbf{y}}\right\}$$

$$\Rightarrow \Phi_{\mathbf{z}}(\omega) = \Phi_{\mathbf{x}}(\omega)\Phi_{\mathbf{y}}(\omega)$$

   Convolution of pdfs $\Leftrightarrow$ multiplication of characteristic functions.

A Random Result

If $\mathbf{x}$ and $\mathbf{y}$ are jointly normal with zero mean, then

$$E\left\{\mathbf{x}^2\mathbf{y}^2\right\} = E\left\{\mathbf{x}^2\right\} E\left\{\mathbf{y}^2\right\} + 2E^2\left\{\mathbf{xy}\right\}$$

6-6 Conditional Distributions

Definition

The *conditional distribution function* of the random variable and $\mathbf{y}$ given the event M is

$$F_{\mathbf{y}|M}(y|M) = P(\{\zeta \in \mathcal{S}\colon \mathbf{y}(\zeta) \leq y\}|\{\zeta \in \mathcal{S}\colon M \text{ occurs}\})$$

$$= \frac{P(\{\zeta \in \mathcal{S}\colon \mathbf{y}(\zeta) \leq y\} \cap \{\zeta \in \mathcal{S}\colon M \text{ occurs}\})}{P(\{\zeta \in \mathcal{S}\colon M \text{ occurs}\})}$$

or, using the short-hand notation,

$$= \frac{P(\mathbf{y} \leq y, M)}{P(M)}$$

Properties

1. $M = \mathbf{x}(\zeta) \leq x$:

$$F_{\mathbf{y}|\mathbf{x}\leq x}(y|\mathbf{x} = x) = \frac{F_{\mathbf{xy}}(x,y)}{F_{\mathbf{x}}(x)}$$

$$f_{\mathbf{y}|\mathbf{x}\leq x}(y|\mathbf{x} \leq x) = \frac{1}{F_{\mathbf{x}}(x)} \int_{-\infty}^{x} f_{\mathbf{xy}}(u,y)\,du$$

2. $M = x_1 \leq \mathbf{x}(\zeta) \leq x_2$:

$$F_{\mathbf{y}|x_1<\mathbf{x}\leq x_2}(y|x_1 < \mathbf{x} \leq x_2) = \frac{F_{\mathbf{xy}}(x_2,y) - F_{\mathbf{xy}}(x_1,y)}{F_{\mathbf{x}}(x_2) - F_{\mathbf{x}}(x_1)}$$

$$f_{\mathbf{y}|x_1<\mathbf{x}\leq x_2}(y|x_1 < \mathbf{x} \leq x_2) = \frac{1}{F_{\mathbf{x}}(x_2) - F_{\mathbf{x}}(x_1)} \int_{x_1}^{x_2} f_{\mathbf{xy}}(u,y)\,du$$

3. Same as 2 with $x_1 = x$ and $x_2 = x + \Delta x$. As $\Delta x \to 0$:

$$f_{\mathbf{y}|\mathbf{x}=x}(y|\mathbf{x} = x) = \frac{f_{\mathbf{xy}}(x,y)}{f_{\mathbf{x}}(x)}$$

4. The other way round

$$f_{\mathbf{x}|\mathbf{y}=y}(\mathbf{x}|\mathbf{y} = y) = \frac{f_{\mathbf{xy}}(x,y)}{f_{\mathbf{y}}(y)}$$

Bayes' Theorem and Total Probability

- Bayes's Rule

$$f_{\mathbf{x}|\mathbf{y}}(x|y) = \frac{f_{\mathbf{y}|\mathbf{x}}(y|x) f_{\mathbf{x}}(x)}{f_{\mathbf{y}}(y)} \qquad f_{\mathbf{y}|\mathbf{x}}(y|x) = \frac{f_{\mathbf{x}|\mathbf{y}}(x|y) f_{\mathbf{y}}(y)}{f_{\mathbf{x}}(x)}$$

- Total Probability Theorem

$$f_{\mathbf{y}}(y) = \int_{-\infty}^{\infty} f_{\mathbf{y}|\mathbf{x}}(y|x) f_{\mathbf{x}}(x)\, dx \qquad f_{\mathbf{x}}(x) = \int_{-\infty}^{\infty} f_{\mathbf{x}|\mathbf{y}}(x|y) f_{\mathbf{y}}(y)\, dy$$

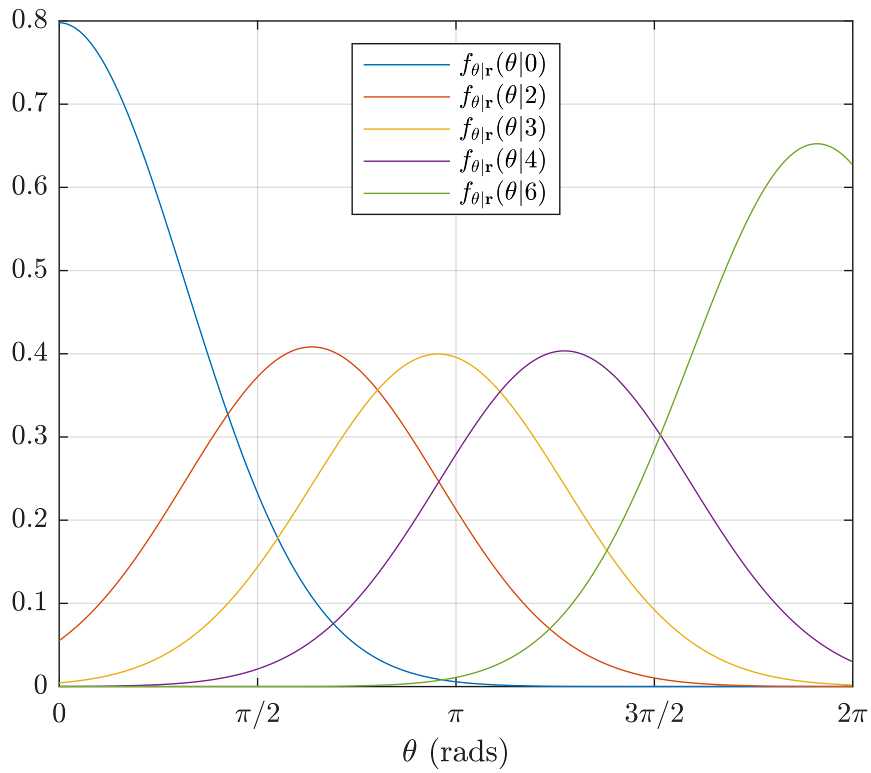
- Bayes' Theorem

$$f_{\mathbf{x}|\mathbf{y}}(x|y) = \frac{f_{\mathbf{y}|\mathbf{x}}(y|x) f_{\mathbf{x}}(x)}{\displaystyle\int_{-\infty}^{\infty} f_{\mathbf{y}|\mathbf{x}}(y|x) f_{\mathbf{x}}(x)\, dx} \qquad f_{\mathbf{y}|\mathbf{x}}(y|x) = \frac{f_{\mathbf{x}|\mathbf{y}}(x|y) f_{\mathbf{y}}(y)}{\int_{-\infty}^{\infty} f_{\mathbf{x}|\mathbf{y}}(x|y) f_{\mathbf{y}}(y)\, dy}$$

	Event A with Event B	Event A with RV $\mathbf{x}$	RV $\mathbf{x}$ with RV $\mathbf{y}$
Bayes Rule	$P(A\|B) = \frac{P(B\|A)P(A)}{P(B)}$	$P(A\|\mathbf{x}=x) = \frac{f_{\mathbf{x}\|A}(x\|A)P(A)}{f_{\mathbf{x}}(x)}$	$f_{\mathbf{x}\|\mathbf{y}}(x\|y) = \frac{f_{\mathbf{y}\|\mathbf{x}}(y\|x)f_{\mathbf{x}}(x)}{f_{\mathbf{y}}(y)}$
	$P(B\|A) = \frac{P(A\|B)P(B)}{P(A)}$	$f_{\mathbf{x}\|A}(x\|A) = \frac{P(A\|\mathbf{x}=x)f_{\mathbf{x}}(x)}{P(A)}$	$f_{\mathbf{y}\|\mathbf{x}}(y\|x) = \frac{f_{\mathbf{x}\|\mathbf{y}}(x\|y)f_{\mathbf{y}}(y)}{f_{\mathbf{x}}(x)}$
TPT	$P(A) = \sum_{i=1}^{n_B} P(A\|B_i)P(B_i)$	$P(A) = \int_{-\infty}^{\infty} P(A\|\mathbf{x}=x)f_{\mathbf{x}}(x)dx$	$f_{\mathbf{y}}(y) = \int_{-\infty}^{\infty} f_{\mathbf{y}\|\mathbf{x}}(y\|x)f_{\mathbf{x}}(x)dx$
	$P(B) = \sum_{i=1}^{n_A} P(B\|A_i)P(A_i)$	$f_{\mathbf{x}}(x) = \sum_{i=1}^{n_A} f_{\mathbf{x}\|A_i}(x\|A_i)P(A_i)$	$f_{\mathbf{x}}(x) = \int_{-\infty}^{\infty} f_{\mathbf{x}\|\mathbf{y}}(x\|y)f_{\mathbf{y}}(y)dy$
Bayes Tho- erem	$P(A\|B) = \frac{P(B\|A)P(A)}{\sum_{i=1}^{n_A} P(B\|A_i)P(A_i)}$	$P(A\|\mathbf{x}=x) = \frac{f_{\mathbf{x}\|A}(x\|A)P(A)}{\sum_{i=1}^{n_A} f_{\mathbf{x}\|A_i}(x\|A_i)P(A_i)}$	$f_{\mathbf{x}\|\mathbf{y}}(x\|y) = \frac{f_{\mathbf{y}\|\mathbf{x}}(y\|x)f_{\mathbf{x}}(x)}{\int_{-\infty}^{\infty} f_{\mathbf{y}\|\mathbf{x}}(y\|x)f_{\mathbf{x}}(x)dx}$
	$P(B\|A) = \frac{P(A\|B)P(B)}{\sum_{i=1}^{n_B} P(A\|B_i)P(B_i)}$	$f_{\mathbf{x}\|A}(x\|A) = \frac{P(A\|\mathbf{x}=x)f_{\mathbf{x}}(x)}{\int_{-\infty}^{\infty} P(A\|\mathbf{x}=x)f_{\mathbf{x}}(x)dx}$	$f_{\mathbf{y}\|\mathbf{x}}(y\|x) = \frac{f_{\mathbf{x}\|\mathbf{y}}(x\|y)f_{\mathbf{y}}(y)}{\int_{-\infty}^{\infty} f_{\mathbf{x}\|\mathbf{y}}(x\|y)f_{\mathbf{y}}(y)dy}$

	Event A with Event B	Event A with RV $\mathbf{x}$	RV $\mathbf{x}$ with RV $\mathbf{y}$
Bayes Rule	$P(A\|B) = \dfrac{P(B\|A)P(A)}{P(B)}$	$P(A\|\mathbf{x}=x) = \dfrac{P(\mathbf{x}=x\|A)P(A)}{P(\mathbf{x}=x)}$	$P(\mathbf{x}=x\|\mathbf{y}=y) = \dfrac{P(\mathbf{y}=y\|\mathbf{x}=x)P(\mathbf{x}=x)}{P(\mathbf{y}=y)}$
	$P(B\|A) = \dfrac{P(A\|B)P(B)}{P(A)}$	$P(\mathbf{x}=x\|A) = \dfrac{P(A\|\mathbf{x}=x)P(\mathbf{x}=x)}{P(A)}$	$P(\mathbf{y}=y\|\mathbf{x}=x) = \dfrac{P(\mathbf{x}=x\|\mathbf{y}=y)P(\mathbf{y}=y)}{P(\mathbf{x}=x)}$
TPT	$P(A) = \sum_{i=1}^{n_B} P(A\|B_i)P(B_i)$	$P(A) = \sum_{i=1}^{n_\mathbf{x}} P(A\|\mathbf{x}=x_i)P(\mathbf{x}=x_i)$	$P(\mathbf{y}=y) = \sum_{i=1}^{n_\mathbf{x}} P(\mathbf{y}=y\|\mathbf{x}=x_i)P(\mathbf{x}=x_i)$
	$P(B) = \sum_{i=1}^{n_A} P(B\|A_i)P(A_i)$	$P(\mathbf{x}=x) = \sum_{i=1}^{n_A} P(\mathbf{x}=x_i)\|A_i)P(A_i)$	$P(\mathbf{x}=x) = \sum_{i=1}^{n_\mathbf{y}} P(\mathbf{x}=x\|\mathbf{y}=y_i)P(\mathbf{y}=y_i)$
Bayes Tho-erem	$P(A\|B) = \dfrac{P(B\|A)P(A)}{\sum_{i=1}^{n_A} P(B\|A_i)P(A_i)}$	$P(A\|\mathbf{x}=x) = \dfrac{P(\mathbf{x}=x\|A)P(A)}{\sum_{i=1}^{n_\mathbf{y}} P(\mathbf{x}=x\|\mathbf{y}=y_i)P(\mathbf{y}=y_i)}$	$P(\mathbf{x}=x\|\mathbf{y}=y) = \dfrac{P(\mathbf{y}=y\|\mathbf{x}=x)P(\mathbf{x}=x)}{\sum_{i=1}^{n_\mathbf{x}} P(\mathbf{y}=y\|\mathbf{x}=x_i)P(\mathbf{x}=x_i)}$
	$P(B\|A) = \dfrac{P(A\|B)P(B)}{\sum_{i=1}^{n_B} P(A\|B_i)P(B_i)}$	$P(\mathbf{x}=x\|A) = \dfrac{P(A\|\mathbf{x}=x)P(\mathbf{x}=x)}{\sum_{i=1}^{n_\mathbf{x}} P(A\|\mathbf{x}=x_i)P(\mathbf{x}=x_i)}$	$P(\mathbf{y}=y\|\mathbf{x}=x) = \dfrac{P(\mathbf{x}=x\|\mathbf{y}=y)P(\mathbf{y}=y)}{\sum_{i=1}^{n_\mathbf{y}} P(\mathbf{x}=x\|\mathbf{y}=y_i)P(\mathbf{y}=y_i)}$

Plot for Example 6-42



6-7 Conditional Expected Values

Definition

Let $\mathbf{y}$ be a random variable and let $g(y)$ be a function $g : \mathbb{R} \to \mathbb{R}$. The *conditional expectation* is

$$E\{g(\mathbf{y})|M\} = \int_{-\infty}^{\infty} g(y) f_{\mathbf{y}|M}(y|M)\, dy$$

Properties

1. The *conditional mean* is

$$\mu_{\mathbf{y}|x} = E\{\mathbf{y}|x\} = \int_{-\infty}^{\infty} y f_{\mathbf{y}|x}(y|x)\, dy$$

2. The *conditional variance* is

$$\sigma^2_{\mathbf{y}|x} = E\{(\mathbf{y} - \mu_{\mathbf{y}|x})^2\} = \int_{-\infty}^{\infty} (y - \mu_{\mathbf{y}|x})^2 f_{\mathbf{y}|x}(y|x)\, dy$$

3. The conditional mean $E\{\mathbf{y}|x\}$ is a function of x.
   (a) This function of x is called a *regression line* even if the function is not a line.
   (b) Replacing x in this function by the random variable $\mathbf{x}$ creates the random variable $E\{\mathbf{y}|\mathbf{x}\}$.
   (c) The expected value of the random variable $E\{\mathbf{y}|\mathbf{x}\}$ is

$$E\{\, E\{\mathbf{y}|\mathbf{x}\}\, \} = \int_{-\infty}^{\infty} E\{\mathbf{y}|x\} f_{\mathbf{x}}(x)\, dx = E\{\mathbf{y}\}$$

4. Generalizations:
   (a) $E\{g(\mathbf{x}, \mathbf{y})|x\}$ is a function of x.

$$E\{g(\mathbf{x}, \mathbf{y})|x\} = \int_{-\infty}^{\infty} g(x, y) f_{\mathbf{y}|x}(y|x)\, dy$$

   (b) Replacing x by the random variable $\mathbf{x}$ makes $E\{g(\mathbf{x}, \mathbf{y})|\mathbf{x}\}$ a random variable.

$$E\{E\{g(\mathbf{x}, \mathbf{y})|\mathbf{x}\}\} = E\{g(\mathbf{x}, \mathbf{y})\}$$

   (c) $g(x, y) = g_1(x) g_2(y)$:

$$E\{g_1(\mathbf{x}) g_2(\mathbf{y})|x\} = E\{g_1(x) g_2(\mathbf{y})|x\} = g_1(x) E\{g_2(\mathbf{y})|x\}$$
$$E\{g_1(\mathbf{x}) g_2(\mathbf{y})\} = E\{E\{g_1(\mathbf{x}) g_2(\mathbf{y})|\mathbf{x}\}\} = E\{g_1(\mathbf{x}) E\{g_2(\mathbf{y})|\mathbf{x}\}\}$$

7-1 General Concepts

Definitions

- A *random vector* is the vector

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix}$$

  whose components are $\mathbf{x}_i$ are random variables.

- The *distribution* of $\mathbf{X}$ is the *joint distribution* of the elements of the vector. For

$$X = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$$

  The distribution of $\mathbf{X}$ is

$$F_{\mathbf{X}}(X) = P(\mathbf{x}_1 \le x_1, \ldots, \mathbf{x}_n \le x_n)$$

- If the random variables in $\mathbf{X}$ are jointly continuous then the joint density is

$$f_{\mathbf{X}}(X) = \frac{\partial^n F_{\mathbf{X}}(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n}$$

- If the random variables in $\mathbf{X}$ are jointly discrete then the probability mass function is

$$P(\mathbf{X} = X) = P(\mathbf{x}_1 = x_1, \ldots, \mathbf{x}_n = x_n)$$

- The random variables $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are called (mutually) *independent* if the events

$$\{\zeta \in \mathcal{S} \colon \mathbf{x}_1 \le x_1\}, \ldots, \{\zeta \in \mathcal{S} \colon \mathbf{x}_n \le x_n\}$$

  are independent.

Notes

- $F_{\mathbf{X}}(X) \colon \mathbb{R}^n \to [0,1] \in \mathbb{R}$
- $f_{\mathbf{X}}(X) \colon \mathbb{R}^n \to \mathbb{R}$.

Definitions

- A *random vector* is the vector

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix}$$

  whose components are $\mathbf{x}_i$ are random variables.

- The *distribution* of $\mathbf{X}$ is the *joint distribution* of the elements of the vector. For

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

  The distribution of $\mathbf{X}$ is

$$F_{\mathbf{X}}(X) = P(\mathbf{x}_1 \le x_1, \ldots, \mathbf{x}_n \le x_n)$$

- If the random variables in $\mathbf{X}$ are jointly continuous then the joint density is

$$f_{\mathbf{X}}(X) = \frac{\partial^n F_{\mathbf{X}}(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n}$$

- If the random variables in $\mathbf{X}$ are jointly discrete then the probability mass function is

$$P(\mathbf{X} = X) = P(\mathbf{x}_1 = x_1, \ldots, \mathbf{x}_n = x_n)$$

- The random variables $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are called (mutually) *independent* if the events

$$\{\zeta \in \mathcal{S} \colon \mathbf{x}_1 \le x_1\}, \ldots, \{\zeta \in \mathcal{S} \colon \mathbf{x}_n \le x_n\}$$

  are independent.

Notes

- $F_{\mathbf{X}}(X) \colon \mathbb{R}^n \to [0,1] \in \mathbb{R}$
- $f_{\mathbf{X}}(X) \colon \mathbb{R}^n \to \mathbb{R}$.

Properties

1. $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are independent:

$$F_{\mathbf{X}}(X) = F_{\mathbf{x}_1}(x_1) \cdots F_{\mathbf{x}_n}(x_n)$$
$$f_{\mathbf{X}}(X) = f_{\mathbf{x}_1}(x_1) \cdots f_{\mathbf{x}_n}(x_n)$$

2. $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are independent and identically distributed (IID):

$$F_{\mathbf{X}}(X) = F_{\mathbf{x}}(x_1) \cdots F_{\mathbf{x}}(x_n)$$
$$f_{\mathbf{X}}(X) = f_{\mathbf{x}}(x_1) \cdots f_{\mathbf{x}}(x_n)$$

where $F_{\mathbf{x}}(\cdot)$ is the common CDF and $f_{\mathbf{x}}(\cdot)$ is the common PDF.

3. Marginal Distributions

$$F_{\mathbf{x}_1}(x_1) = F_{\mathbf{X}}(x_, \infty, \ldots, \infty)$$

4. Marginal PDFs for jointly continuous RVs

$$f_{\mathbf{x}_1}(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, x_2, \ldots, x_n)\, dx_2 \cdots dx_n$$

5. Marginal PMFs for jointly discrete RVs

$$P(\mathbf{x}_1 = x_1) = \sum_{i_2} \cdots \sum_{i_n} P(\mathbf{x}_1 = x_1, \mathbf{x}_2 = x_{i_2}, \ldots, \mathbf{x}_n = x_{i_n})$$

6. Expectation (continuous RVs)

$$\begin{aligned} E\{g(\mathbf{X})\} &= E\{g(\mathbf{x}_1, \ldots, \mathbf{x}_n)\} \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n) f_{\mathbf{X}}(x_1, \ldots, x_n)\, dx_1 \cdots dx_n \end{aligned}$$

7. Expectation (discrete RVs)

$$\begin{aligned} E\{g(\mathbf{X})\} &= E\{g(\mathbf{x}_1, \ldots, \mathbf{x}_n)\} \\ &= \sum_{i_1} \cdots \sum_{i_n} g(x_{i_1}, \ldots, x_{i_n}) P(\mathbf{x}_1 = x_{i_1}, \ldots, \mathbf{x}_n = x_{i_n}) \end{aligned}$$

Vector/Matrix Definitions

Vector Definitions

- Vector, conjugate, transpose, Hermitian (conjugate transpose)

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

- Conjugate

$$X^* = \begin{bmatrix} x_1^* \\ \vdots \\ x_n^* \end{bmatrix}$$

- Transpoose

$$X^t = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$$

- Hermitian (conjugate-transpose)

$$X^H = \begin{bmatrix} x_1^* & \cdots & x_n^* \end{bmatrix}$$

Matrix Definitions

- Matrix

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{bmatrix}$$

- Conjugate

$$X^* = \begin{bmatrix} x_{11}^* & \cdots & x_{1n}^* \\ \vdots & & \vdots \\ x_{n1}^* & \cdots & x_{nn}^* \end{bmatrix}$$

- Transpose

$$X^t = \begin{bmatrix} x_{11} & \cdots & x_{n1} \\ \vdots & & \vdots \\ x_{1n} & \cdots & x_{nn} \end{bmatrix}$$

- Hermitian (conjugate-transpose)

$$X^H = \begin{bmatrix} x_{11}^* & \cdots & x_{n1}^* \\ \vdots & & \vdots \\ x_{1n}^* & \cdots & x_{nn}^* \end{bmatrix}$$

Vector Operations for $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$

- Inner product

$$X^H Y = \begin{bmatrix} x_1^* & \cdots & x_n^* \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} x_i^* y_i$$

  inner product: $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$

- Outer product

$$XY^H = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1^* & \cdots & y_n^* \end{bmatrix} = \begin{bmatrix} x_1 y_1^* & \cdots & x_1 y_n^* \\ \vdots & & \vdots \\ x_n y_1^* & \cdots & x_n y_n^* \end{bmatrix}$$

  outer product $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}^{n\times n}$

Matrix Operations

- Determinant: $\det(X)\colon \mathbb{C}^{n\times n} \to \mathbb{C}$
- Inverse: X^{-1} is the *inverse* of the square matrix X means

$$X^{-1}X = XX^{-1} = I$$

- Symmetric
  - the real matrix X is *symmetric* means $X = X^t$.
  - the complex-valued matrix X is *conjugate symmetric* or *Hermitian* means $X = X^H$.
- Unitary: the complex-valued matrix X is *unitary* means

$$XX^H = X^H X = I$$

- Eigen-decomposition $X = Q\Lambda Q^{-1}$ where

the n eigenvalues of X

$$Q = \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & & \vdots \\ v_{n1} & \cdots & v_{nn} \end{bmatrix} \quad \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

the eigenvector corresponding to λ_n

the eigenvector corresponding to λ_1

Statistical Vectors and Matrices for $\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix}$

- Mean Vector

$$\mu_{\mathbf{X}} = E\{\mathbf{X}\} = \begin{bmatrix} E\{\mathbf{x}_1\} \\ \vdots \\ E\{\mathbf{x}_n\} \end{bmatrix} = \begin{bmatrix} \mu_{\mathbf{x}_1} \\ \vdots \\ \mu_{\mathbf{x}_n} \end{bmatrix}$$

- Covariance Matrix

$$C_{\mathbf{XX}} = E\{(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{X} - \mu_{\mathbf{X}})^t\}$$

$$= E\left\{ \begin{bmatrix} \mathbf{x}_1 - \mu_{\mathbf{x}_1} \\ \vdots \\ \mathbf{x}_n - \mu_{\mathbf{x}_n} \end{bmatrix} \begin{bmatrix} (\mathbf{x}_1 - \mu_{\mathbf{x}_1}) & \cdots & (\mathbf{x}_n - \mu_{\mathbf{x}_n}) \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} (\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{x}_1 - \mu_{\mathbf{x}_1}) & \cdots & (\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{x}_n - \mu_{\mathbf{x}_n}) \\ \vdots & & \vdots \\ (\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{x}_1 - \mu_{\mathbf{x}_1}) & \cdots & (\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{x}_{\mathbf{x}_n} - \mu_{\mathbf{x}_n}) \end{bmatrix} \right\}$$

$$= \begin{bmatrix} E\{(\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{x}_1 - \mu_{\mathbf{x}_1})\} & \cdots & E\{(\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{x}_n - \mu_{\mathbf{x}_n})\} \\ \vdots & & \vdots \\ E\{(\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{x}_1 - \mu_{\mathbf{x}_1})\} & \cdots & E\{(\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{x}_n - \mu_{\mathbf{x}_n})\} \end{bmatrix}$$

$$= \begin{bmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & & \vdots \\ C_{n1} & \cdots & C_{nn} \end{bmatrix}$$

- Correlation Matrix

$$R_{\mathbf{XX}} = E\{\mathbf{X}\mathbf{X}^t\} = E\left\{ \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} \mathbf{x}_1\mathbf{x}_1 & \cdots & \mathbf{x}_1\mathbf{x}_n \\ \vdots & & \vdots \\ \mathbf{x}_n\mathbf{x}_1 & \cdots & \mathbf{x}_n\mathbf{x}_n \end{bmatrix} \right\} = \begin{bmatrix} E\{\mathbf{x}_1\mathbf{x}_1\} & \cdots & E\{\mathbf{x}_1\mathbf{x}_1\} \\ \vdots & & \vdots \\ E\{\mathbf{x}_n\mathbf{x}_1\} & \cdots & E\{\mathbf{x}_n\mathbf{x}_n\} \end{bmatrix}$$

$$= \begin{bmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & & \vdots \\ R_{n1} & \cdots & R_{nn} \end{bmatrix}$$

Statistical Matrices for

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_n \end{bmatrix}$$

- Cross-Covariance Matrix

$$C_{\mathbf{XY}} = E\{(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{Y} - \mu_{\mathbf{Y}})^t\}$$

$$= E\left\{ \begin{bmatrix} \mathbf{x}_1 - \mu_{\mathbf{x}_1} \\ \vdots \\ \mathbf{x}_n - \mu_{\mathbf{x}_n} \end{bmatrix} \begin{bmatrix} (\mathbf{y}_1 - \mu_{\mathbf{y}_1}) & \cdots & (\mathbf{y}_n - \mu_{\mathbf{y}_n}) \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} (\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{y}_1 - \mu_{\mathbf{y}_1}) & \cdots & (\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{y}_n - \mu_{\mathbf{y}_n}) \\ \vdots & & \vdots \\ (\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{y}_1 - \mu_{\mathbf{y}_1}) & \cdots & (\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{y}_n - \mu_{\mathbf{y}_n}) \end{bmatrix} \right\}$$

$$= \begin{bmatrix} E\{(\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{y}_1 - \mu_{\mathbf{y}_1})\} & \cdots & E\{(\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{y}_n - \mu_{\mathbf{y}_n})\} \\ \vdots & & \vdots \\ E\{(\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{y}_1 - \mu_{\mathbf{y}_1})\} & \cdots & E\{(\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{y}_n - \mu_{\mathbf{y}_n})\} \end{bmatrix}$$

$$= \begin{bmatrix} C_{\mathbf{x}_1\mathbf{y}_1} & \cdots & C_{\mathbf{x}_1\mathbf{y}_n} \\ \vdots & & \vdots \\ C_{\mathbf{x}_n\mathbf{y}_1} & \cdots & C_{\mathbf{x}_n\mathbf{y}_n} \end{bmatrix}$$

- Cross-Correlation Matrix

$$R_{\mathbf{XY}} = E\{\mathbf{X}\mathbf{Y}^t\} = E\left\{ \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 & \cdots & \mathbf{y}_n \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} \mathbf{x}_1\mathbf{y}_1 & \cdots & \mathbf{x}_1\mathbf{y}_n \\ \vdots & & \vdots \\ \mathbf{x}_n\mathbf{y}_1 & \cdots & \mathbf{x}_n\mathbf{y}_n \end{bmatrix} \right\} = \begin{bmatrix} E\{\mathbf{x}_1\mathbf{y}_1\} & \cdots & E\{\mathbf{x}_1\mathbf{y}_1\} \\ \vdots & & \vdots \\ E\{\mathbf{x}_n\mathbf{y}_1\} & \cdots & E\{\mathbf{x}_n\mathbf{y}\mathbf{x}_n\} \end{bmatrix}$$

$$= \begin{bmatrix} R_{\mathbf{x}_1\mathbf{y}_1} & \cdots & R_{\mathbf{x}_1\mathbf{y}_n} \\ \vdots & & \vdots \\ R_{\mathbf{x}_n\mathbf{y}_1} & \cdots & R_{\mathbf{x}_n\mathbf{y}_n} \end{bmatrix}$$

Normal Random Vector: $\mathbf{X} \sim N(\mu_{\mathbf{X}}, C_{\mathbf{XX}})$ means

$$f_{\mathbf{X}}(X) = \frac{1}{\sqrt{(2\pi)^n \det(C_{\mathbf{XX}})}} \exp\left\{-\frac{1}{2}(X - \mu_{\mathbf{X}})^T C_{\mathbf{XX}}^{-1} (X - \mu_{\mathbf{X}})\right\}$$

where

$$\mu_{\mathbf{X}} = E\{\mathbf{X}\}$$

$$C_{\mathbf{XX}} = E\left\{(\mathbf{x} - \mu_{\mathbf{X}})(\mathbf{x} - \mu_{\mathbf{X}})^t\right\}$$

Because $X \in \mathbb{R}^n$, $f_{\mathbf{X}}(X)\colon \mathbb{R}^n \to \mathbb{R}$.

If $\mathbf{X} \sim N(\mu_{\mathbf{X}}, C_{\mathbf{XX}})$, then

$$\mathbf{y} = a_1 \mathbf{x}_1 + \cdots + a_n \mathbf{x}_n = \underbrace{\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix}}_{\mathbf{X}}$$

is also normal.

This generalizes: If $\mathbf{X} \sim N(\mu_{\mathbf{X}}, C_{\mathbf{XX}})$, then

$$\underbrace{\mathbf{Y}}_{k\times 1} = \underbrace{A}_{k\times n} \underbrace{\mathbf{X}}_{n\times 1}$$

is normal: $\mathbf{Y} \sim N(\mu_{\mathbf{Y}}, C_{\mathbf{YY}})$

7-2 Conditional Densities, Characteristic Functions, and Normality

Definitions

- The *joint conditional density* of the random variables $\mathbf{x}_n, \ldots, \mathbf{x}_{k+1}$ assuming $\mathbf{x}_k, \ldots, \mathbf{x}_1$ is

$$f(x_n, \ldots, x_{k+1}|x_k, \ldots x_1) = \frac{f(x_1, \ldots, x_n)}{f(x_1, \ldots, x_k)}$$

- The *joint conditional distribution* of the random variables $\mathbf{x}_n, \ldots, \mathbf{x}_{k+1}$ assuming $\mathbf{x}_k, \ldots, \mathbf{x}_1$ is

$$\begin{aligned} &F(x_n, \ldots, x_{k+1}|x_k, \ldots x_1) \\ &\quad = \int_{u_n=-\infty}^{x_n} \cdots \int_{u_{k+1}=-\infty}^{x_{k+1}} f(u_n, \ldots, u_{k+1}|x_k, \ldots x_1)\, du_{k+1} \cdots du_n \end{aligned}$$

Properties

1. Chain Rule

$$f(x_1, \ldots, x_n) = f(x_n|x_{n-1}, \ldots, x_1) \cdots f(x_2|x_1) f(x_1)$$

2. Removing variables on the left or on the right of the conditional line:

   (a) Some example equations

$$\begin{aligned} f(x_1|x_3) &= \int_{-\infty}^{\infty} f(x_1, x_2|x_3)\, dx_2 \\ f(x_1|x_4) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1|x_2, x_3, x_4) f(x_2, x_3|x_4)\, dx_2\, dx_3 \end{aligned}$$

   (b) To remove any number of variables on the left of the conditional line, integrate with respect to them.

   (c) To remove any number of variables to the right of the conditional line, multiply by their conditional density with respect to the remaining variables on the right (of the conditional line) and integrate the product.

   (d) The Chapman-Kolmogorov equation is a special case:

$$f(x_1|x_3) = \int_{-\infty}^{\infty} f(x_1|x_2, x_3) f(x_2|x_3)\, dx_2.$$

The definitions and properties hold for discrete random variables if the pdfs are replaced by pmfs and the integrals by summations.

Definitions

- The conditional mean of $\mathbf{y}_1, \ldots, \mathbf{y}_k = g(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ given the event M is

$$E\{g(\mathbf{x}_1, \ldots, \mathbf{x}_n)|M\} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n) f(x_1, \ldots x_n|M)\, dx_1 \cdots dx_n$$

Properties

1. Special case

$$E\{\mathbf{x}_1|x_2, \ldots, x_n\} = \int_{-\infty}^{\infty} x_1 f(x_1|x_2, \ldots, x_n)\, dx_1$$

   $E\{\mathbf{x}_1|x_2, \ldots, x_n\}$ is a function of $x_2, \ldots, x_n$.

2. $E\{\mathbf{x}_1|\mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is a *random variable*. The expected value of $E\{\mathbf{x}_1|\mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is

$$E\{E\{\mathbf{x}_1|\mathbf{x}_2, \ldots, \mathbf{x}_n\}\} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} E\{\mathbf{x}_1|x_2, \ldots, x_n\} f(x_2, \ldots, x_n)\, dx_2 \cdots dx_n$$

$$= E\{\mathbf{x}_1\}$$

3. $E\{\mathbf{x}_1|x_2, x_3, x_4\}$ is a function of x_2, x_3, x_4. $E\{\mathbf{x}_1|x_2, x_3, \mathbf{x}_4\}$ is a random variable. The conditional expected value of $E\{\mathbf{x}_1|x_2, x_3, \mathbf{x}_4\}$ is

$$E\{E\{\mathbf{x}_1|x_1, x_2, \mathbf{x}_4\}\} = \int_{-\infty}^{\infty} E\{\mathbf{x}_1|x_2, x_3, x_4\} f(x_4|x_2, x_3)\, dx_4$$

$$= E\{\mathbf{x}_1|x_2, x_3\}$$

4. Generalization: To remove any number of variables on the right of the conditional expected value line, multiply by their conditional density with respect to the remaining variables on the right (of the conditional expected value line) and integrate.

The definitions and properties hold for discrete random variables if the pdfs are replaced by pmfs and the integrals by summations.

Law of Large Numbers

- $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are an IID sequence with $E\{\mathbf{x}_i\} = \mu$.
- Sample mean $\overline{\mathbf{x}} = \dfrac{\mathbf{x}_1 + \cdots + \mathbf{x}_n}{n}$
- Law of Large Numbers: $\dfrac{\mathbf{x}_1 + \cdots + \mathbf{x}_n}{n} \underset{n\to\infty}{\longrightarrow} \mu$

Central Limit Theorem

- $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are a sequence of *independent* random variables with

$$E\{\mathbf{x}_i\} = \mu_i \qquad E\{(\mathbf{x}_i - \mu_i)^2\} = \sigma_i^2$$

- The sum $\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_n$

$$E\{\mathbf{x}\} = \mu = \mu_1 + \cdots + \mu_n$$
$$E\{(\mathbf{x} - \mu)^2\} = \sigma^2 = \sigma_1^2 + \cdots + \sigma_n^2$$

- Central Limit Theorem: under certain general conditions

$$\mathbf{z} = \frac{\mathbf{x} - \mu}{\sigma} \qquad F_{\mathbf{z}}(z) \underset{n\to\infty}{\longrightarrow} \text{standard normal CDF}$$

- If the random variables $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are jointly continuous then, under the same general conditions

$$\mathbf{z} = \frac{\mathbf{x} - \mu}{\sigma} \qquad f_{\mathbf{z}}(z) \underset{n\to\infty}{\longrightarrow} \text{standard normal PDF}$$

Complex-Valued Random Variables

Statistical Vectors and Matrices for the complex-valued random vector

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_n \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 + j\mathbf{y}_1 \\ \vdots \\ \mathbf{x}_n + j\mathbf{y}_n \end{bmatrix}$$

- Mean Vector

$$\mu_{\mathbf{Z}} = E\{\mathbf{Z}\} = \begin{bmatrix} E\{\mathbf{z}_1\} \\ \vdots \\ E\{\mathbf{z}_n\} \end{bmatrix} = \begin{bmatrix} \mu_{\mathbf{z}_1} \\ \vdots \\ \mu_{\mathbf{z}_n} \end{bmatrix}$$

- Covariance Matrix

$$C_{\mathbf{ZZ}} = E\{(\mathbf{Z} - \mu_{\mathbf{Z}})(\mathbf{Z} - \mu_{\mathbf{Z}})^H\}$$

$$= E\left\{ \begin{bmatrix} \mathbf{z}_1 - \mu_{\mathbf{z}_1} \\ \vdots \\ \mathbf{z}_n - \mu_{\mathbf{z}_n} \end{bmatrix} \begin{bmatrix} (\mathbf{z}_1^* - \mu_{\mathbf{z}_1}^*) & \cdots & (\mathbf{z}_n^* - \mu_{\mathbf{z}_n}^*) \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} (\mathbf{z}_1 - \mu_{\mathbf{z}_1})(\mathbf{z}_1^* - \mu_{\mathbf{z}_1}^*) & \cdots & (\mathbf{z}_1 - \mu_{\mathbf{z}_1})(\mathbf{z}_n^* - \mu_{\mathbf{z}_n}^*) \\ \vdots & & \vdots \\ (\mathbf{z}_n - \mu_{\mathbf{z}_n})(\mathbf{z}_1^* - \mu_{\mathbf{z}_1}^*) & \cdots & (\mathbf{z}_n - \mu_{\mathbf{z}_n})(\mathbf{z}_n^* - \mu_{\mathbf{z}_n}^*) \end{bmatrix} \right\}$$

$$= \begin{bmatrix} E\{(\mathbf{z}_1 - \mu_{\mathbf{z}_1})(\mathbf{z}_1^* - \mu_{\mathbf{z}_1}^*)\} & \cdots & E\{(\mathbf{z}_1 - \mu_{\mathbf{z}_1})(\mathbf{z}_n^* - \mu_{\mathbf{z}_n})^*\} \\ \vdots & & \vdots \\ E\{(\mathbf{z}_n - \mu_{\mathbf{z}_n})(\mathbf{z}_1^* - \mu_{\mathbf{z}_1})^*\} & \cdots & E\{(\mathbf{z}_n - \mu_{\mathbf{z}_n})(\mathbf{z}_n^* - \mu_{\mathbf{z}_n}^*)\} \end{bmatrix}$$

$$= \begin{bmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & & \vdots \\ C_{n1} & \cdots & C_{nn} \end{bmatrix}$$

- Correlation Matrix

$$R_{\mathbf{ZZ}} = E\{\mathbf{Z}\mathbf{Z}^H\} = E\left\{ \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_n \end{bmatrix} \begin{bmatrix} \mathbf{z}_1^* & \cdots & \mathbf{z}_n^* \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} \mathbf{z}_1\mathbf{z}_1^* & \cdots & \mathbf{z}_1\mathbf{z}_n^* \\ \vdots & & \vdots \\ \mathbf{z}_n\mathbf{z}_1^* & \cdots & \mathbf{z}_n\mathbf{z}_n^* \end{bmatrix} \right\} = \begin{bmatrix} E\{\mathbf{z}_1\mathbf{z}_1^*\} & \cdots & E\{\mathbf{z}_1\mathbf{z}_1^*\} \\ \vdots & & \vdots \\ E\{\mathbf{z}_n\mathbf{z}_1^*\} & \cdots & E\{\mathbf{z}_n\mathbf{z}_n^*\} \end{bmatrix}$$

$$= \begin{bmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & & \vdots \\ R_{n1} & \cdots & R_{nn} \end{bmatrix}$$

Statistical Matrices for the complex-valued random vectors

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_n \end{bmatrix} \quad \mathbf{W} = \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_n \end{bmatrix}$$

- Cross-Covariance Matrix

$$C_{\mathbf{ZW}} = E\{(\mathbf{Z} - \mu_{\mathbf{Z}})(\mathbf{W} - \mu_{\mathbf{W}})^H\}$$

$$= E\left\{ \begin{bmatrix} \mathbf{z}_1 - \mu_{\mathbf{z}_1} \\ \vdots \\ \mathbf{z}_n - \mu_{\mathbf{z}_n} \end{bmatrix} \begin{bmatrix} (\mathbf{w}_1^* - \mu_{\mathbf{w}_1}^*) & \cdots & (\mathbf{w}_n^* - \mu_{\mathbf{w}_n}^*) \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} (\mathbf{z}_1 - \mu_{\mathbf{z}_1})(\mathbf{w}_1^* - \mu_{\mathbf{w}_1}^*) & \cdots & (\mathbf{z}_1 - \mu_{\mathbf{z}_1})(\mathbf{w}_n^* - \mu_{\mathbf{w}_n}^*) \\ \vdots & & \vdots \\ (\mathbf{z}_n - \mu_{\mathbf{z}_n})(\mathbf{w}_1^* - \mu_{\mathbf{w}_1}^*) & \cdots & (\mathbf{z}_n - \mu_{\mathbf{z}_n})(\mathbf{w}_n^* - \mu_{\mathbf{w}_n}^*) \end{bmatrix} \right\}$$

$$= \begin{bmatrix} E\{(\mathbf{z}_1 - \mu_{\mathbf{z}_1})(\mathbf{w}_1^* - \mu_{\mathbf{w}_1}^*)\} & \cdots & E\{(\mathbf{z}_1 - \mu_{\mathbf{z}_1})(\mathbf{w}_n^* - \mu_{\mathbf{w}_n})^*\} \\ \vdots & & \vdots \\ E\{(\mathbf{z}_n - \mu_{\mathbf{z}_n})(\mathbf{w}_1^* - \mu_{\mathbf{w}_1})^*\} & \cdots & E\{(\mathbf{z}_n - \mu_{\mathbf{z}_n})(\mathbf{w}_n^* - \mu_{\mathbf{w}_n}^*)\} \end{bmatrix}$$

$$= \begin{bmatrix} C_{\mathbf{z}_1\mathbf{w}_1} & \cdots & C_{\mathbf{z}_1\mathbf{w}_n} \\ \vdots & & \vdots \\ C_{\mathbf{z}_n\mathbf{w}_1} & \cdots & C_{\mathbf{z}_n\mathbf{w}_n} \end{bmatrix}$$

- Correlation Matrix

$$R_{\mathbf{ZW}} = E\{\mathbf{Z}\mathbf{W}^H\} = E\left\{ \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_n \end{bmatrix} \begin{bmatrix} \mathbf{w}_1^* & \cdots & \mathbf{w}_n^* \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} \mathbf{z}_1\mathbf{w}_1^* & \cdots & \mathbf{z}_1\mathbf{w}_n^* \\ \vdots & & \vdots \\ \mathbf{z}_n\mathbf{w}_1^* & \cdots & \mathbf{z}_n\mathbf{w}_n^* \end{bmatrix} \right\} = \begin{bmatrix} E\{\mathbf{z}_1\mathbf{w}_1^*\} & \cdots & E\{\mathbf{z}_1\mathbf{w}_1^*\} \\ \vdots & & \vdots \\ E\{\mathbf{z}_n\mathbf{w}_1^*\} & \cdots & E\{\mathbf{z}_n\mathbf{w}_n^*\} \end{bmatrix}$$

$$= \begin{bmatrix} R_{\mathbf{z}_1\mathbf{w}_1} & \cdots & R_{\mathbf{z}_1\mathbf{w}_n} \\ \vdots & & \vdots \\ R_{\mathbf{z}_n\mathbf{w}_1} & \cdots & R_{\mathbf{z}_n\mathbf{w}_n} \end{bmatrix}$$

Complex-Valued Multivariate Normal 1:

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_n \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 + j\mathbf{y}_1 \\ \vdots \\ \mathbf{x}_n + j\mathbf{y}_n \end{bmatrix} \quad \Rightarrow \quad \mathbf{V} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_n \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ b \\ c \end{matrix}}\right\} \mathbf{X} \\ \left.\vphantom{\begin{matrix} a \\ b \\ c \end{matrix}}\right\} \mathbf{Y} \end{matrix}$$

- PDF of $\mathbf{Z}$ is the joint PDF of $\mathbf{V}$:

$$f_{\mathbf{V}}(V) = \frac{1}{(2\pi)^n \sqrt{\det(C_{\mathbf{VV}})}} \exp\left\{ -\frac{1}{2}(V - \mu_{\mathbf{V}})^t C_{\mathbf{VV}}^{-1}(V - \mu_{\mathbf{V}}) \right\}$$

The mean vector and covariance matrix of V have special forms

$$\mu_{\mathbf{V}} = E\{\mathbf{V}\} = E\left\{ \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \right\} = \begin{bmatrix} \mu_{\mathbf{X}} \\ \mu_{\mathbf{Y}} \end{bmatrix}$$

$$\begin{aligned} C_{\mathbf{VV}} &= E\left\{ (\mathbf{V} - \mu_{\mathbf{V}})(\mathbf{V} - \mu_{\mathbf{V}})^T \right\} \\ &= E\left\{ \begin{bmatrix} \mathbf{X} - \mu_{\mathbf{X}} \\ \mathbf{Y} - \mu_{\mathbf{Y}} \end{bmatrix} \begin{bmatrix} (\mathbf{X} - \mu_{\mathbf{X}})^t & (\mathbf{Y} - \mu_{\mathbf{Y}})^t \end{bmatrix} \right\} \\ &= E\left\{ \begin{bmatrix} (\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{X} - \mu_{\mathbf{X}})^t & (\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{Y} - \mu_{\mathbf{Y}})^t \\ & \\ (\mathbf{Y} - \mu_{\mathbf{Y}})(\mathbf{X} - \mu_{\mathbf{X}})^t & (\mathbf{Y} - \mu_{\mathbf{Y}})(\mathbf{Y} - \mu_{\mathbf{Y}})^t \end{bmatrix} \right\} \\ &= \begin{bmatrix} C_{\mathbf{XX}} & C_{\mathbf{XY}} \\ C_{\mathbf{YX}} & C_{\mathbf{YY}} \end{bmatrix} \end{aligned}$$

Complex-Valued Multivariate Normal 2:

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_n \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 + j\mathbf{y}_1 \\ \vdots \\ \mathbf{x}_n + j\mathbf{y}_n \end{bmatrix}$$

- PDF of $\mathbf{Z}$ in terms of Z:

$$f_{\mathbf{Z}}(Z) = \frac{1}{\pi^n \sqrt{\det(C_{\mathbf{ZZ}})\det(\Gamma)}} \times$$
$$\exp\left\{ -\frac{1}{2} \begin{bmatrix} (Z-\mu_{\mathbf{Z}})^H & (Z-\mu_{\mathbf{Z}})^t \end{bmatrix} \begin{bmatrix} C_{\mathbf{ZZ}} & P_{\mathbf{ZZ}} \\ P_{\mathbf{ZZ}}^H & C_{\mathbf{ZZ}}^* \end{bmatrix}^{-1} \begin{bmatrix} (Z-\mu_{\mathbf{Z}}) \\ (Z-\mu_{\mathbf{Z}})^* \end{bmatrix} \right\}$$

where

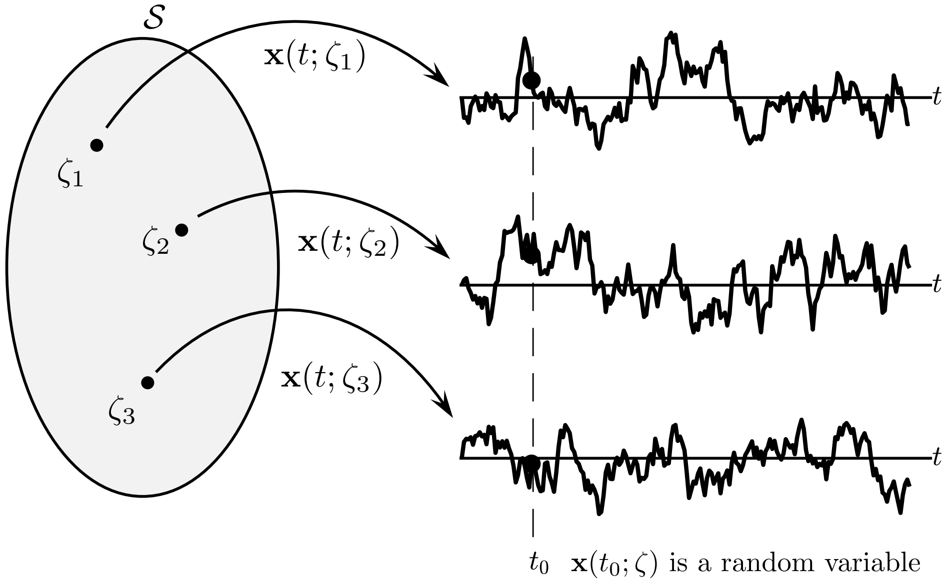
$$\begin{aligned} \mu_{\mathbf{Z}} &= E\{\mathbf{Z}\} && \text{mean vector} \\ C_{\mathbf{ZZ}} &= E\left\{(\mathbf{Z}-\mu_{\mathbf{Z}})(\mathbf{Z}-\mu_{\mathbf{Z}})^H\right\} && \text{covariance matrix} \\ P_{\mathbf{ZZ}} &= E\left\{(\mathbf{Z}-\mu_{\mathbf{Z}})(\mathbf{Z}-\mu_{\mathbf{Z}})^t\right\} && \text{pseudo-covariance matrix} \\ \Gamma &= C_{\mathbf{ZZ}}^* - P_{\mathbf{ZZ}}^H \mathbf{C}_{ZZ}^{-1} P_{\mathbf{ZZ}} \end{aligned}$$

- A *Proper* complex-valued multivariate normal random vector is one for which the pseudo-covariance matrix is the all-zeros matrix.

- The pdf for a *proper* complex-valued multivariate normal random vector is

$$f_{\mathbf{Z}}(Z) = \frac{1}{\pi^n \sqrt{\det(C_{\mathbf{ZZ}})}} \exp\left\{ -(\mathbf{Z}-\mu_{\mathbf{Z}})^H C_{\mathbf{ZZ}}^{-1} (\mathbf{Z}-\mu_{\mathbf{Z}}) \right\}$$

9-1 Definitions



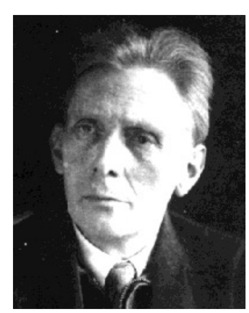
Definition

- A *stochastic process* (also called *random process*) $\mathbf{x}(t;\zeta)$ is a rule for assigning to every $\zeta \in \mathcal{S}$ a function of time.
  - A stochastic process is a *family of time functions* depending on the the parameter $\zeta \in \mathcal{S}$.
  - A stochastic process is a *function* of t and ζ.
- The functions of time that comprise the stochastic process may be either *continuous time* functions or *discrete time* functions.

Interpretations

1. If t and ζ are variables, the result is a family (or an *ensemble*) of waveforms $\mathbf{x}(t,\zeta)$.
2. If t is a variable and ζ is fixed, the result is a single function of time (or a *sample* of the stochastic process).
3. If t is fixed and ζ is variable, the result is a *random variable.*
4. If t and ζ are fixed, the result is a *number.*

στόχος, στόχου, ὁ: target, guess, conjecture



Alexsandr Yakovlevich Khinchin
(Алекса́ндр Я́ковлевич Хи́нчин)
1894-1959

Statistics of Stochastic Processes

Definition

- The *n-th order distribution* of the real-valued process $\mathbf{x}(t)$ is the joint distribution of the real-valued random variables $\mathbf{x}(t_1), \ldots, \mathbf{x}(t_n)$:

$$F_{\mathbf{x}}(x_1, \ldots, x_n; t_1, \ldots, t_n) = P(\mathbf{x}(t_1) \leq x_1, \ldots, \mathbf{x}(t_n) \leq x_n)$$

  - If the random variables are jointly continuous, then the joint cdf is a continuous function.
  - If the random variables are jointly discrete, the the joint cdf is an n-dimensional stair-step function.
  - Do not confuse time and random variable type: a continuous-*time* random process may be described by either continuous or discrete *random variables* at a fixed time instant.

- The *n-th order density function* of the real-valued process $\mathbf{x}(t)$ is joint density of the real-valued random variables $\mathbf{x}(t_1), \ldots, \mathbf{x}(t_n)$:

$$f_{\mathbf{x}}(x_1, \ldots, x_n; t_1, \ldots, t_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F_{\mathbf{x}}(x_1, \ldots, x_n; t_1, \ldots, t_n)$$

  - If the random variables are jointly continuous, then the joint pdf is smooth.
  - If the random variables are jointly discrete, then the joint pdf contains impulses (in the form of Dirac delta functions).
  - Alternatively, for jointly discrete random variables, the joint pmf may be used.
  - Do not confuse time and random variable type: a continuous-*time* random process may be described by either a continuous or discrete *random variable* at a fixed time instant.

Special cases (real-valued random processes)

- First-order density:
  1. The *first-order* distribution/density is the special case $n = 1$:
  $$F_{\mathbf{x}}(x;t) = P(\mathbf{x}(t) \leq x)$$
  $$f_{\mathbf{x}}(x;t) = \frac{\partial F_{\mathbf{x}}(x;t)}{\partial x}$$
  2. The *mean* of the random process $\mathbf{x}(t)$ is the mean of the random variable $\mathbf{x}(t)$ for fixed t and is computed from the first-order pdf
  $$\mu_{\mathbf{x}}(t) = E\{\mathbf{x}(t)\} = \int_{-\infty}^{\infty} x f_{\mathbf{x}}(x;t) dx$$
- Second-order density
  1. The *second-order* distribution/density is the special case $n = 2$:
  $$F_{\mathbf{x}}(x_1, x_2; t_1, t_2) = P(\mathbf{x}(t_1) \leq x_1, \mathbf{x}(t_2) \leq x_2)$$
  $$f_{\mathbf{x}}(x_1, x_2; t_1, t_2) = \frac{\partial^2 F_{\mathbf{x}}(x_1, x_2; t_1, t_2)}{\partial x_1 \partial x_2}$$
  2. The *autocorrelation function* is the expected value of the product $\mathbf{x}(t_1)\mathbf{x}(t_2)$ and is computed from the second order density:
  $$R_{\mathbf{xx}}(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{\mathbf{x}}(x_1, x_2; t_2, t_2)\, dx_1\, dx_2$$
  3. The *average power* of the random process $\mathbf{x}(t)$ is the value of $R_{\mathbf{xx}}(t_1, t_2)$ along the diagonal $t = t_1 = t_2$:
  $$\text{average power} = E\{\mathbf{x}^2(t)\} = R_{\mathbf{xx}}(t, t)$$
  4. The *autcovariance* of the random process $\mathbf{x}(t)$ is the covariance of the random variables $\mathbf{x}(t_1)$ and $\mathbf{x}(t_2)$ and is computed from the second order density
  $$\begin{aligned} C_{\mathbf{xx}}(t_1, t_2) &= E\{(\mathbf{x}(t_1) - \mu_{\mathbf{x}}(t_1))(\mathbf{x}(t_2) - \mu_{\mathbf{x}}(t_2))\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - \mu_{\mathbf{x}}(t_1))(x_2 - \mu_{\mathbf{x}}(t_2)) f_{\mathbf{x}}(x_1, x_2; t_1, t_2)\, dx_1\, dx_2 \end{aligned}$$
  5. The *variance* of the random process $\mathbf{x}(t)$ is the value of $C_{\mathbf{xx}}(t_1, t_2)$ along the diagonal $t = t_1 = t_2$:
  $$\text{variance} = E\{(\mathbf{x}(t) - \mu_{\mathbf{x}}(t))^2\} = C_{\mathbf{xx}}(t, t)$$
  6. The *correlation coefficient* is
  $$r_{\mathbf{xx}}(t_1, t_2) = \frac{C_{\mathbf{xx}}(t_1, t_2)}{\sqrt{C_{\mathbf{xx}}(t_1, t_1) C_{\mathbf{xx}}(t_2, t_2)}}$$

More Definitions (real-valued random processes)

- A *white* random process $\mathbf{x}(t)$ means

$$C_{\mathbf{xx}}(t_1, t_2) = q(t_1)\delta(t_1 - t_2)$$

  It is *almost always* assumed that a white random process has zero mean:

$$\mu_{\mathbf{x}}(t) = 0$$

- A *normal random process* $\mathbf{x}(t)$ means the random variables $\mathbf{x}(t_1), \ldots, \mathbf{x}(t_n)$ are jointly normal for any n and any $t_1, \ldots, t_n$.

Two real-valued random processes

- Two real-valued random processes $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are described by the joint distribution and density of the random variables

$$\mathbf{x}(t_1), \ldots, \mathbf{x}(t_n), \mathbf{y}(t_1'), \ldots, \mathbf{y}(t_m')$$

$$F_{\mathbf{xy}}(x_1, \ldots, x_n, y_1, \ldots, y_m; t_1, \ldots, t_n, t_1', \ldots, t_m') = P(\mathbf{x}(t_1) \leq x_1, \ldots, \mathbf{x}(t_n) \leq x_n, \mathbf{y}(t_1') \leq y_1, \ldots, \mathbf{y}(t_m') \leq y_m)$$

$$f_{\mathbf{xy}}(x_1, \ldots, x_n, y_1, \ldots, y_m; t_1, \ldots, t_n, t_1', \ldots, t_m') = \frac{\partial^{n+m} F_{\mathbf{xy}}(x_1, \ldots, x_n, y_1, \ldots, y_m; t_1, \ldots, t_n, t_1', \ldots, t_m')}{\partial x_1 \cdots \partial x_n \partial y_1 \cdots \partial y_m}$$

- The *cross-correlation function* is

$$R_{\mathbf{xy}}(t_1, t_2) = E\{\mathbf{x}(t_1)\mathbf{y}(t_2)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{\mathbf{xy}}(x, y; t_1, t_2)\, dx\, dy$$

- The *cross-covariance function* is

$$\begin{aligned} C_{\mathbf{xy}}(t_1, t_2) &= E\{(\mathbf{x}(t_1) - \mu_{\mathbf{x}}(t_1))(\mathbf{y}(t_2) - \mu_{\mathbf{y}}(t_2))\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_{\mathbf{x}}(t_1))(y - \mu_{\mathbf{y}}(t_2)) f_{\mathbf{xy}}(x, y; t_1, t_2)\, dx\, dy \end{aligned}$$

- Two processes $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are *uncorrelated* if

$$C_{\mathbf{xy}}(t_1, t_2) = 0 \quad \text{for every } t_1 \text{ and } t_2$$

Comments on Complex-Valued Random Processes

- A *complex-valued* random process $\mathbf{z}(t,\zeta)$ maps each $\zeta \in \mathcal{S}$ to a complex-valued waveform.

  1. If t and ζ are variables, the result is an *ensemble* of complex-valued waveforms $\mathbf{z}(t,\zeta)$.
  2. If t is variable and ζ is fixed, the result is a single complex-valued function of time: a *sample* of the random process.
  3. If t is fixed and ζ is variable, the result is a complex-valued *random variable.*
  4. If t and ζ are fixed, the result is a *complex number.*

- The n-th order distribution and density of the complex-valued process $\mathbf{z}(t)$

  – Write

$$\begin{array}{ccc} \mathbf{z}(t_1) = \mathbf{x}(t_1) + j\mathbf{y}(t_1) & \qquad & z_1 = x_1 + jy_1 \\ \vdots & & \vdots \\ \mathbf{z}(t_n) = \mathbf{x}(t_n) + j\mathbf{y}(t_n) & & z_n = x_n + jy_n \end{array}$$

  – The n-th order distribution is the joint distribution of the complex-valued random variables $\mathbf{z}(t_1), \ldots, \mathbf{z}(t_n)$

$$\begin{aligned} F_{\mathbf{z}}(z_1,\ldots,z_n;t_1,\ldots,t_n) &= P(\mathbf{x}(t_1) \le x_1,\ldots,\mathbf{x}(t_n) \le x_n, \mathbf{y}(t_1) \le y_1,\ldots,\mathbf{y}(t_n) \le y_n) \\ &= F_{\mathbf{xy}}(x_1,\ldots,x_n,y_1\ldots,y_n;t_1,\ldots,t_n) \end{aligned}$$

  – The n-th order density of $\mathbf{z}(t)$ is expressed in terms of the (real-valued) real and imaginary components of $\mathbf{z}(t)$

$$\begin{aligned} f_{\mathbf{z}}(z_1,\ldots,z_n;t_1,\ldots,t_n) &= f_{\mathbf{xy}}(x_1,\ldots,x_n,y_1,\ldots,y_n;t_1,\ldots,t_n) \\ &= \frac{\partial^{2n}}{\partial x_1 \cdots \partial x_n \partial y_1 \cdots \partial y_n} F_{\mathbf{xy}}(x_1,\ldots,x_n,y_1\ldots,y_n;t_1,\ldots,t_n) \end{aligned}$$

- First two moments

  – mean

$$\mu_{\mathbf{z}}(t) = \int_{-\infty}^{\infty} z f_{\mathbf{z}}(z;t)dz = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+jy) f_{\mathbf{xy}}(x,y;t)\,dx\,dy = \mu_{\mathbf{x}}(t) + j\mu_{\mathbf{y}}(t)$$

  – Autocorrelation $R_{\mathbf{zz}}(t_1,t_2) = E\{\mathbf{z}(t_1)\mathbf{z}^*(t_2)\}$

$$R_{\mathbf{zz}}(t_1,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1+jy_1)(x_2-jy_2) f_{\mathbf{xy}}(x_1,x_2,y_1,y_2;t_1,t_2)\,dx_1\,dx_2\,dy_1\,dy_2$$

  – Autocovariance: $C_{\mathbf{zz}}(t_1,t_2) = R_{\mathbf{zz}}(t_1,t_2) - \mu_{\mathbf{z}}(t_1)\mu_{\mathbf{z}}^*(t_2)$

Stationary Processes

Definitions

A stochastic process $\mathbf{x}(t)$ is called *strict-sense stationary* (abbreviated SSS) if its statistical properties are invariant to to a shift of the time origin.

$\Rightarrow$ $\mathbf{x}(t)$ and $\mathbf{x}(t+c)$ have the same statistics.

$\Rightarrow$ $f_{\mathbf{x}}(x_1, \ldots, x_n; t_1, \ldots, t_n) = f_{\mathbf{x}}(x_1, \ldots, x_n; t_1 + c, \ldots, t_n + c)$ for any c and for all n.

Properties

1. First-order density:
   (a) $f_{\mathbf{x}}(x;t) = f_{\mathbf{x}}(x;t+c) \Rightarrow f_{\mathbf{x}}(x;t) = f_{\mathbf{x}}(x)$
   (b) $\mu_{\mathbf{x}}(t) = \int_{-\infty}^{\infty} x f_{\mathbf{x}}(x;t)dx = \int_{-\infty}^{\infty} x f_{\mathbf{x}}(x)dx = \mu_{\mathbf{x}}$
2. Second-order density:
   (a)
   $$\begin{aligned} f_{\mathbf{x}}(x_1, x_2; t_1, t_2) &= f_{\mathbf{x}}(x_1, x_2; t_1 + c, t_2 + c) \\ \Rightarrow f_{\mathbf{x}}(x_1, x_2; t_1, t_2) &= f_{\mathbf{x}}(x_1, x_2; t_1 - t_2, 0) \\ &= f_{\mathbf{x}}(x_1, x_2; \tau, 0), \quad \tau = t_1 - t_2 \end{aligned}$$
   "Thus the joint density of the random variables $\mathbf{x}(t+\tau)$ and $\mathbf{x}(t)$ is independent of [i.e., not a function of] t and it equals $f(x_1, x_2; \tau)$."
   (b)
   $$\begin{aligned} R_{\mathbf{xx}}(t_1, t_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{\mathbf{x}}(x_1, x_2; t_1 t_2)\, dx_1\, dx_2 \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{\mathbf{x}}(x_1, x_2; \tau, 0)\, dx_1\, dx_2 \\ &= R_{\mathbf{xx}}(\tau, 0) \end{aligned}$$
   (c) It is customary to express the autocorrelation function for a WSS random process by
   $$R_{\mathbf{xx}}(\tau) = E\{\mathbf{x}(t+\tau)\mathbf{x}(t)\} = E\{\mathbf{x}(t)\mathbf{x}(t-\tau)\}$$
   (d) $C_{\mathbf{xx}}(t_1, t_2) = C_{\mathbf{xx}}(\tau) = R_{\mathbf{xx}}(\tau) - \mu_{\mathbf{x}}^2$

Consequences

1. average power of SSS process $= R_{\mathbf{xx}}(0)$
2. variance of SSS process $= C_{\mathbf{xx}}(0)$
3. correlation coefficient of SSS process: $r_{\mathbf{xx}}(\tau) = \dfrac{C_{\mathbf{xx}}(\tau)}{C_{\mathbf{xx}}(0)}$

Definitions

A stochastic process $\mathbf{x}(t)$ is called *wide-sense stationary* (abbreviated WSS) if

$$f_{\mathbf{x}}(x;t) = f_{\mathbf{x}}(x;t+c)$$
$$f_{\mathbf{x}}(x_1,x_2;t_1,t_2) = f_{\mathbf{x}}(x_1,x_2;t_1+c,t_2+c)$$

Properties

1. $\mu_{\mathbf{x}}(t) = \mu_{\mathbf{x}}$
2. $R_{\mathbf{xx}}(t_1,t_2) = R_{\mathbf{xx}}(\tau)$

A stochastic process $\mathbf{x}(t)$ is *WSS white noise* means $C_{\mathbf{xx}}(\tau) = q\delta(\tau)$.

Comments on Complex-Valued WSS Random Processes

- The complex-valued WSS process $\mathbf{z}(t) = \mathbf{x}(t) + j\mathbf{y}(t)$ is described in terms of the joint statistics of the two real-valued processes $\mathbf{x}(t)$ and $\mathbf{y}(t)$.
- The first-order density property

$$f_{\mathbf{z}}(z;t) = f_{\mathbf{z}}(z;t+c) \Rightarrow f_{\mathbf{z}}(z;t) = f_{\mathbf{z}}(z)$$

becomes

$$f_{\mathbf{xy}}(x,y;t) = f_{\mathbf{xy}}(x,y;t+c) \Rightarrow f_{\mathbf{xy}}(x,y;t) = f_{\mathbf{xy}}(x,y)$$

- The complex-valued mean is a constant:

$$\mu_{\mathbf{z}}(t) = \mu_{\mathbf{z}} \Rightarrow \mu_{\mathbf{x}}(t) + j\mu_{\mathbf{y}}(t) = \mu_{\mathbf{x}} + j\mu_{\mathbf{y}}$$

- The second-order density property

$$\begin{aligned} f_{\mathbf{z}}(z_1,z_2;t_1,t_2) &= f_{\mathbf{x}}(z_1,z_2;t_1+c,t_2+c) \\ &\Rightarrow f_{\mathbf{z}}(z_1,z_2;t_1,t_2) = f_{\mathbf{x}}(z_1,z_2;t_1-t_2,0) \end{aligned}$$

becomes

$$\begin{aligned} f_{\mathbf{xy}}(x_1,x_2,y_1,y_2;t_1,t_2) &= f_{\mathbf{xy}}(x_1,x_2,y_1,y_2;t_1+c,t_2+c) \\ &\Rightarrow f_{\mathbf{xy}}(x_1,x_2,y_1,y_2;t_1,t_2) = f_{\mathbf{xy}}(x_1,x_2,y_1,y_2;t_1-t_2,0) \end{aligned}$$

- Autocorrelation function is

$$R_{\mathbf{zz}}(\tau) = E\{\mathbf{z}(t+\tau)z^*(t)\} = E\{\mathbf{z}(t)\mathbf{z}^*(t-\tau)\}$$

Properties of the auto- and cross-correlation functions

General Random Processes

1. $R_{\mathbf{xx}}(t_1, t_2) = E\{\mathbf{x}(t_1)\mathbf{x}^*(t_2)\}$
2. $R_{\mathbf{xx}}(t_2, t_1) = R^*_{\mathbf{xx}}(t_1, t_2)$
3. $R_{\mathbf{xx}}(t, t) \geq 0$
4. $R_{\mathbf{xy}}(t_1, t_2) = E\{\mathbf{x}(t_1)\mathbf{y}^*(t_2)\}$
5. $R_{\mathbf{yx}}(t_2, t_1) = R^*_{\mathbf{xy}}(t_1, t_2)$

WSS Random Processes

1. $R_{\mathbf{xx}}(\tau) = E\{\mathbf{x}(t+\tau)\mathbf{x}^*(t)\}$
2. $R_{\mathbf{xx}}(-\tau) = R^*_{\mathbf{xx}}(\tau)$
3. $R_{\mathbf{xx}}(0) \geq 0$
4. $R_{\mathbf{xy}}(\tau) = E\{\mathbf{x}(t+\tau)\mathbf{y}^*(t)\}$
5. $R_{\mathbf{yx}}(-\tau) = R^*_{\mathbf{xy}}(\tau)$
6. $R_{\mathbf{xx}}(\tau) \leq R_{\mathbf{xx}}(0)$

From Property 6 for WSS random processes

$R_{\mathbf{xx}}(\tau_1) = R_{\mathbf{xx}}(0)$ for some $\tau_1 \neq 0 \quad \Rightarrow \quad R_{\mathbf{xx}}(\tau + \tau_1) = R_{\mathbf{xx}}(\tau)$ for all τ

$\Rightarrow \quad R_{\mathbf{xx}}(\tau)$ is *periodic* with period τ_1

$R_{\mathbf{xx}}(\tau_1) = R_{\mathbf{xx}}(\tau_2) = R_{\mathbf{xx}}(0)$ for τ_1, τ_2 noncommensurate $\quad \Rightarrow \quad R_{\mathbf{xx}}(\tau) = \text{constant}.$

9-2 Systems With Stochastic Inputs

Definitions

- A stochastic process $\mathbf{x}(t, \zeta)$ is a map from $\mathcal{S}$ to a real-valued function of time.
  - For each $\zeta \in \mathcal{S}$, $\mathbf{x}(t, \zeta)$ is a *signal*.
- System: input is a signal $x(t)$. Output is another signal $y(t)$.
  - If the input to a system is a random process, then the input/relationship applies on a sample-by-sample basis.
  - For $\zeta_i \in \mathcal{S}$,

$$\mathbf{x}(t, \zeta_i) = \text{the input signal}$$
$$\mathbf{y}(t, \zeta_i) = \text{the output signal}$$

$$\mathbf{x}(t, \zeta_i) \longrightarrow \boxed{\text{system}} \longrightarrow \mathbf{y}(t, \zeta_i)$$

- Kinds of Systems
  - The system is *deterministic* if it operates only the variable t, treating ζ as a parameter.
  - The system is called *stochastic* if it operates on both variales t and ζ.
  - If the system is specified in terms of physical elements or by an equation,
    * the system is deterministic if the elements or coefficients of the defining equations are deterministic
    * the system is stochastic of the elements or coefficients of the defining equations are random
  - Memoryless Systems: a system is called *memoryless* if its input/output relationship is given by $\mathbf{y}(t) = g(\mathbf{x}(t))$.
  - LTI Systems: a linear time-invariant system is described in the recorded lectures. The input/output relationship is given by the *convolution* of the input signal with the *impulse response* $h(t)$ of the system:

$$\mathbf{y}(t, \zeta) = \int_{-\infty}^{\infty} \mathbf{x}(u, \zeta) h(t-u)\, du = \int_{-\infty}^{\infty} \mathbf{x}(t-u, \zeta) h(u)\, du$$

  - In this class, all LTI systems will be described by linear constant-coefficient differential equations *with all zero initial conditions*

LTI System with WSS Input

$$\mathbf{x}(t) \longrightarrow \boxed{h(t)} \longrightarrow \mathbf{y}(t)$$

$$\mu_{\mathbf{y}} = \mu_{\mathbf{x}} \int_{-\infty}^{\infty} h(u)du$$

$$R_{\mathbf{xy}}(\tau) = \int_{-\infty}^{\infty} R_{\mathbf{xx}}(\tau + u)h^*(u)du$$

$$R_{\mathbf{yy}}(\tau) = \int_{-\infty}^{\infty} R_{\mathbf{xy}}(\tau - u)h(u)du$$

Special Case: $\mathbf{x}(t)$ is WSS white random process:

$$\mu_{\mathbf{x}} = 0 \qquad R_{\mathbf{xx}}(\tau) = q\delta(\tau)$$

$$\mu_{\mathbf{y}} = 0$$

$$R_{\mathbf{xy}}(\tau) = qh^*(-\tau)$$

$$R_{\mathbf{yy}}(\tau) = q \underbrace{\int_{-\infty}^{\infty} h^*(u - \tau)h(u)du}_{\rho(\tau)}$$

9-3 The Power Spectrum

Definitions

- The *power spectrum* (or *spectral density*) of a WSS random process $\mathbf{x}(t)$, real or complex, is the Fourier transform of its autocorrelation function $R_{\mathbf{xx}}(\tau) = E\{\mathbf{x}(t+\tau)\mathbf{x}(t)\}$:

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

- From the Fourier inversion formula

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

- The *cross power spectrum* of two random processes $\mathbf{x}(t)$ and $\mathbf{y}(t)$ is the Fourier transform of their cross correlation $R_{\mathbf{xy}}(\tau) = E\{\mathbf{x}(t+\tau)\mathbf{y}^*(t)\}$:

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

- From the Fourier inversion formula

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

Properties

1. $S_{\mathbf{xx}}(\omega)$ is a real-valued function of ω.
2. If $\mathbf{x}(t)$ is real, then $S_{\mathbf{xx}}(\omega)$ is real and even.
3. $S_{\mathbf{xx}}(\omega) \geq 0$ for all ω.
4. $S_{\mathbf{xy}}(\omega)$ is, in general, complex valued, even when both processes $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are real-valued processes.
5. $S_{\mathbf{xy}}(\omega) = S_{\mathbf{yx}}^*(\omega)$

TABLE 9-1

$$R(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)e^{j\omega\tau}\,d\omega \quad \leftrightarrow \quad S(\omega) = \int_{-\infty}^{\infty} R(\tau)e^{-j\omega\tau}\,d\tau$$

$$\delta(\tau) \leftrightarrow 1 \qquad\qquad 1 \leftrightarrow 2\pi\delta(\omega)$$

$$e^{j\beta\tau} \leftrightarrow 2\pi\delta(\omega-\beta) \qquad \cos(\beta\tau) \leftrightarrow \pi\delta(\omega-\beta) + \pi\delta(\omega+\beta)$$

$$e^{-\alpha|\tau|} \leftrightarrow \frac{2\alpha}{\alpha^2+\omega^2} \qquad\qquad e^{-\alpha\tau^2} \leftrightarrow \sqrt{\frac{\pi}{\alpha}}e^{-\omega^2/4\alpha}$$

$$e^{-\alpha|\tau|}\cos(\beta\tau) \leftrightarrow \frac{\alpha}{\alpha^2+(\omega-\beta)^2} + \frac{\alpha}{\alpha^2+(\omega+\beta)^2}$$

$$2e^{-\alpha\tau^2}\cos(\beta\tau) \leftrightarrow \sqrt{\frac{\pi}{\alpha}}\left[e^{-(\omega-\beta)^2/4\alpha} + e^{-(\omega+\beta)^2/4\alpha}\right]$$

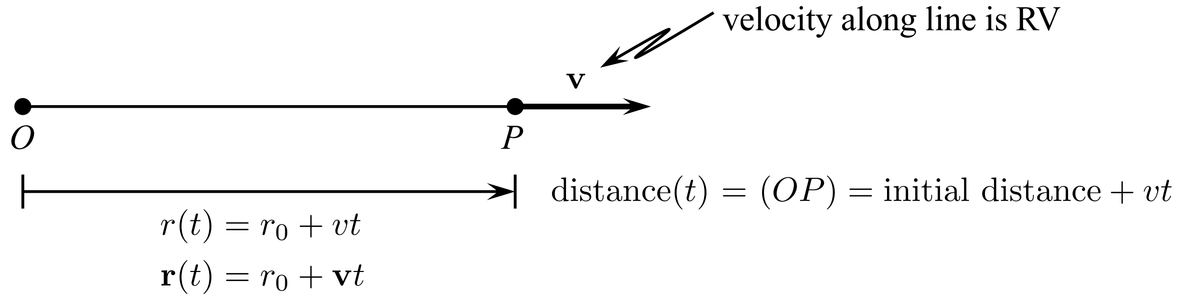
$$\begin{cases} 1 - \dfrac{|\tau|}{T} & |\tau| < T \\ 0 & |\tau| > T \end{cases} \leftrightarrow \frac{4\sin^2(\omega T/2)}{T\omega^2}$$

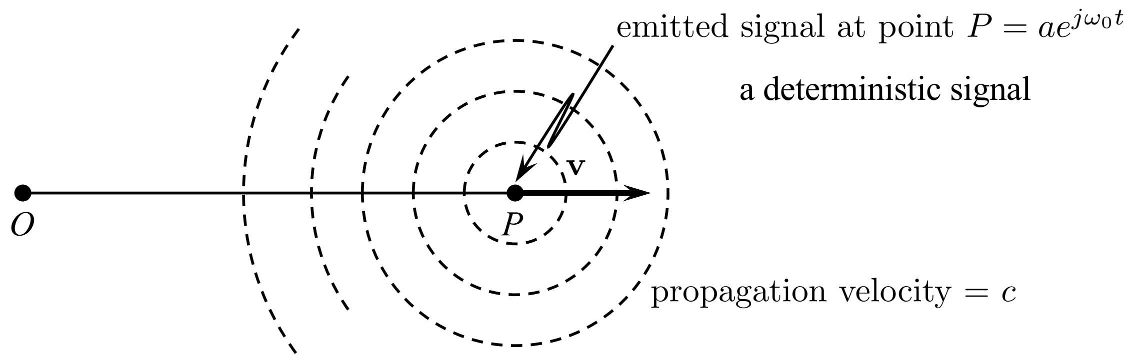
$$\frac{\sin(\sigma\tau)}{\pi\tau} \leftrightarrow \begin{cases} 1 & |\omega| < \sigma \\ 0 & |\omega| > \sigma \end{cases}$$

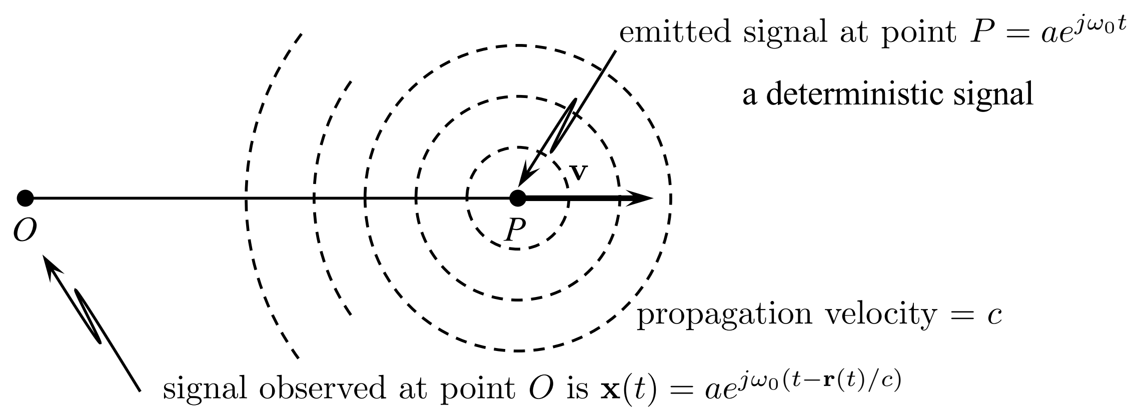
$$\left.\begin{aligned} e^{-\alpha|\tau|}\sin(\beta|\tau|) &\leftrightarrow \frac{\omega+\beta}{\alpha^2+(\omega+\beta)^2} - \frac{\omega-\beta}{\alpha^2+(\omega-\beta)^2} \\ |\tau|e^{-\alpha|\tau|} &\leftrightarrow 2\frac{\alpha^2-\omega^2}{(\alpha^2+\omega^2)^2} \end{aligned}\right\}$$

These are valid Fourier transform pairs, but the left-hand sides by themselves are not valid auto-correlation functions. (See Property 6.)

Example 9-24: Doppler Effect







Observed signal is a random process because $\mathbf{r}(t)$ is a random process.
$\mathbf{r}(t)$ is a random process because $\mathbf{v}$ is a random variable.

If $\mathbf{x}(t)$ is WSS, we can ask, "What is the power spectral density of $\mathbf{x}(t)$?"

LTI System with WSS Input

$$\mathbf{x}(t) \longrightarrow \boxed{\; h(t) \;} \longrightarrow \mathbf{y}(t)$$

$$\mu_{\mathbf{y}} = \mu_{\mathbf{x}} \int_{-\infty}^{\infty} h(u) du \qquad \qquad \mu_{\mathbf{y}} = \mu_{\mathbf{x}} H(0)$$

$$R_{\mathbf{xy}}(\tau) = \int_{-\infty}^{\infty} R_{\mathbf{xx}}(\tau + u) h^*(u) du \qquad \qquad S_{\mathbf{xy}}(\omega) = S_{\mathbf{xx}}(\omega) H^*(\omega)$$

$$R_{\mathbf{yy}}(\tau) = \int_{-\infty}^{\infty} R_{\mathbf{xy}}(\tau - u) h(u) du \qquad \qquad S_{\mathbf{yy}}(\omega) = S_{\mathbf{xy}}(\omega) H(\omega) = S_{\mathbf{xx}}(\omega) |H(\omega)|^2$$

Example 9-27

$$\mathbf{x}(t) \longrightarrow \boxed{h(t)} \longrightarrow \mathbf{y}(t)$$

(a) $\quad \mathbf{y}'(t) + c\mathbf{y}(t) = \mathbf{x}(t) \quad \text{all } t \quad \mu_{\mathbf{x}} = 0 \quad R_{\mathbf{xx}}(\tau) = q\delta(\tau)$

$$H(s) = \frac{1}{s+c}$$

$$H(\omega) = \frac{1}{c + j\omega}$$

$$|H(\omega)|^2 = \frac{1}{c^2 + \omega^2}$$

$$S_{\mathbf{yy}}(\omega) = S_{\mathbf{xx}}(\omega)|H(\omega)|^2$$

$$= \frac{q}{c^2 + \omega^2}$$

$$R_{\mathbf{yy}}(\tau) = \frac{q}{2c}e^{-c|\tau|}$$

$$H(s) = \frac{1}{s+c}$$

$$h(t) = e^{-ct}U(t)$$

$$R_{\mathbf{xy}}(\tau) = \begin{cases} qe^{c\tau} & \tau < 0 \\ 0 & \tau > 0 \end{cases}$$

$$R_{\mathbf{yy}}(\tau) = \begin{cases} \displaystyle\int_0^\infty qe^{c(\tau-u)}e^{-cu}du & \tau < 0 \\ \displaystyle\int_\tau^\infty qe^{c(\tau-u)}e^{-cu}du & \tau > 0 \end{cases}$$

$$= \begin{cases} \dfrac{q}{2c}e^{c\tau} & \tau < 0 \\ \dfrac{q}{2c}e^{-c\tau} & \tau > 0 \end{cases}$$

$$= \frac{q}{2c}e^{-c|\tau|}$$

$$\rho(\tau) = \begin{cases} \displaystyle\int_0^\infty e^{-c(u-\tau)}e^{-cu}du & \tau < 0 \\ \displaystyle\int_\tau^\infty e^{-c(u-\tau)}e^{-cu}du & \tau > 0 \end{cases} = \begin{cases} \dfrac{1}{2c}e^{c\tau} & \tau < 0 \\ \dfrac{1}{2c}e^{-c\tau} & \tau > 0 \end{cases} = \frac{1}{2c}e^{-c|\tau|}$$

$$R_{\mathbf{yy}}(\tau) = q\rho(\tau) = \frac{q}{2c}e^{-c|\tau|}$$

$$E\{\mathbf{y}^2(t)\} = R_{\mathbf{yy}}(0) = \frac{q}{2c}$$

Example 9-27

$$\mathbf{x}(t) \longrightarrow \boxed{\;h(t)\;} \longrightarrow \mathbf{y}(t)$$

(b) $\mathbf{y}''(t) + b\mathbf{y}'(t) + c\mathbf{y}(t) = \mathbf{x}(t)$ all t $\mu_{\mathbf{x}} = 0$ $R_{\mathbf{xx}}(\tau) = q\delta(\tau)$

$$H(s) = \frac{1}{s^2 + bs + c}$$

$$H(\omega) = \frac{1}{c - \omega^2 + jb\omega}$$

$$|H(\omega)|^2 = \frac{1}{(c - \omega^2)^2 + b^2\omega^2}$$

$$S_{\mathbf{yy}}(\omega) = S_{\mathbf{xx}}(\omega)|H(\omega)|^2 = \frac{q}{(c - \omega^2)^2 + b^2\omega^2}$$

$\underline{b^2 < 4c}$

$$R_{\mathbf{yy}}(\tau) = \frac{q}{2bc} e^{-\alpha|\tau|} \left(\cos(\beta\tau) + \frac{\alpha}{\beta} \sin(\beta|\tau|) \right) \qquad \alpha = \frac{b}{2} \qquad \alpha^2 + \beta^2 = c$$

$\underline{b^2 = 4c}$

$$R_{\mathbf{yy}}(\tau) = \frac{q}{2bc} e^{-\alpha|\tau|} \left(1 + \alpha|\tau| \right) \qquad \alpha = \frac{b}{2}$$

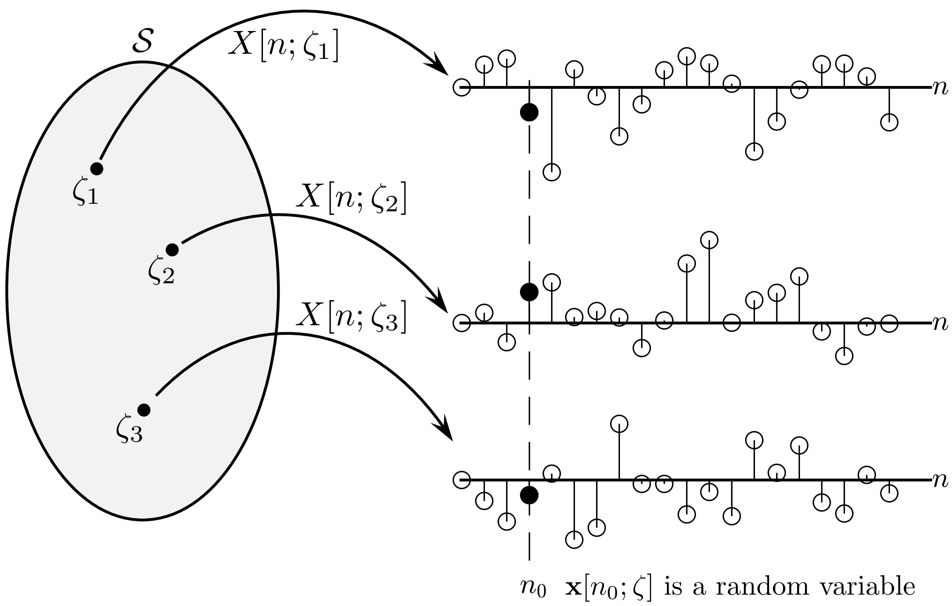
$\underline{b^2 > 4c}$

$$R_{\mathbf{yy}}(\tau) = \frac{q}{4\gamma bc} \left[(\alpha + \gamma) e^{-(\alpha - \gamma)|\tau|} - (\alpha - \gamma) e^{-(\alpha + \gamma)|\tau|} \right]$$

$$\alpha = \frac{b}{2} \qquad \alpha^2 - \gamma^2 = c$$

In all cases, $E\{\mathbf{y}^2(t)\} = \dfrac{q}{2bc}$.

9-4 Discrete-Time Random Processes



Definition

A discrete-time *stochastic process* (also called *random process*) $\mathbf{x}[n;\zeta]$ is a rule for assigning to every $\zeta \in \mathcal{S}$ a discrete-time sequence.

Interpretations

1. If n and ζ are variables, the result is a family (or an *ensemble*) of sequences $\mathbf{x}[n,\zeta]$.
2. If n is a variable and ζ is fixed, the result is a single discrete-time sequence (or a *sample* of the stochastic process).
3. If n is fixed and ζ is variable, the result is a *random variable.*
4. If n and ζ are fixed, the result is a *number.*

Statistics of Discrete-Time Stochastic Processes

Definition

- The *k-th order distribution* of the real-valued process $\mathbf{x}[n]$ is the joint distribution of the real-valued random variables $\mathbf{x}[n_1], \ldots, \mathbf{x}[n_k]$:

$$F_{\mathbf{x}}(x_1, \ldots, x_k; n_1, \ldots, n_k) = P(\mathbf{x}[n_1] \leq x_1, \ldots, \mathbf{x}[n_k] \leq x_k)$$

  - If the random variables are jointly continuous, then the joint cdf is a continuous function.
  - If the random variables are jointly discrete, the the joint cdf is a k-dimensional stair-step function.
  - Do not confuse time and random variable type: a discete-*time* random process may be described by either continuous or discrete *random variables* at a fixed time index.

- The *k-th order density function* of the real-valued process $\mathbf{x}[n]$ is joint density of the real-valued random variables $\mathbf{x}[n_1], \ldots, \mathbf{x}[n_k]$:

$$f_{\mathbf{x}}(x_1, \ldots, x_k; n_1, \ldots, n_k) = \frac{\partial^k}{\partial x_1 \cdots \partial x_k} F_{\mathbf{x}}(x_1, \ldots, x_n; n_1, \ldots, n_k)$$

  - If the random variables are jointly continuous, then the joint pdf is smooth.
  - If the random variables are jointly discrete, then the joint pdf contains impulses (in the form of Dirac delta functions).
  - Alternatively, for jointly discrete random variables, the joint pmf may be used.
  - Do not confuse time and random variable type: a discrete-*time* random process may be described by either a continuous or discrete *random variable* at a fixed time index.

Special cases (real-valued random processes)

- First-order density:

  1. The *first-order* distribution/density is the special case $k = 1$:

$$F_{\mathbf{x}}(x; n) = P(\mathbf{x}[n] \leq x)$$
$$f_{\mathbf{x}}(x; n) = \frac{\partial F_{\mathbf{x}}(x; n)}{\partial x}$$

  2. The *mean* of the random process $\mathbf{x}[n]$ is the mean of the random variable $\mathbf{x}[n]$ for fixed n and is computed from the first-order pdf

$$\mu_{\mathbf{x}}[n] = E\{\mathbf{x}[n]\} = \int_{-\infty}^{\infty} x f_{\mathbf{x}}(x; n) dx$$

- Second-order density

  1. The *second-order* distribution/density is the special case $k = 2$:

$$F_{\mathbf{x}}(x_1, x_2; n_1, n_2) = P(\mathbf{x}[n_1] \leq x_1, \mathbf{x}[n_2] \leq x_2)$$
$$f_{\mathbf{x}}(x_1, x_2; n_1, n_2) = \frac{\partial^2 F_{\mathbf{x}}(x_1, x_2; n_1, n_2)}{\partial x_1 \partial x_2}$$

  2. The *autocorrelation function* is the expected value of the product $\mathbf{x}[n_1]\mathbf{x}[n_2]$ and is computed from the second order density:

$$R_{\mathbf{xx}}[n_1, n_2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{\mathbf{x}}(x_1, x_2; n_1, n_2)\, dx_1\, dx_2$$

  3. The *average power* of the random process $\mathbf{x}[n]$ is the value of $R_{\mathbf{xx}}[n_1, n_2]$ along the diagonal $n = n_1 = n_2$:

$$\text{average power} = E\{\mathbf{x}^2[n]\} = R_{\mathbf{xx}}[n, n]$$

  4. The *autcovariance* of the random process $\mathbf{x}[n]$ is the covariance of the random variables $\mathbf{x}[n_1]$ and $\mathbf{x}[n_2]$ and is computed from the second order density

$$\begin{aligned} C_{\mathbf{xx}}[n_1, n_2] &= E\{(\mathbf{x}[n_1] - \mu_{\mathbf{x}}[n_1])(\mathbf{x}[n_2] - \mu_{\mathbf{x}}[n_2])\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - \mu_{\mathbf{x}}[n_1])(x_2 - \mu_{\mathbf{x}}[n_2]) f_{\mathbf{x}}(x_1, x_2; n_1, n_2)\, dx_1\, dx_2 \end{aligned}$$

  5. The *variance* of the random process $\mathbf{x}[n]$ is the value of $C_{\mathbf{xx}}[n_1, n_2]$ along the diagonal $n = n_1 = n_2$:

$$\text{variance} = E\{(\mathbf{x}[n] - \mu_{\mathbf{x}}[n])^2\} = C_{\mathbf{xx}}(n, n)$$

  6. The *correlation coefficient* is

$$r_{\mathbf{xx}}[n_1, n_2] = \frac{C_{\mathbf{xx}}[n_1, n_2]}{\sqrt{C_{\mathbf{xx}}[n_1, n_1] C_{\mathbf{xx}}[n_2, n_2]}}$$

More Definitions (real-valued random processes)

- A *white* random process $\mathbf{x}[n]$ means

$$C_{\mathbf{xx}}[n_1, n_2] = q[n_1]\delta[n_1 - n_2]$$

It is *almost always* assumed that a white random process has zero mean:

$$\mu_{\mathbf{x}}[n] = 0$$

- A *normal random process* $\mathbf{x}[n]$ means the random variables $\mathbf{x}[n_1], \ldots, \mathbf{x}[n_k]$ are jointly normal for any k and any $n_1, \ldots, n_k$.

Two real-valued random processes

- Two real-valued random processes $\mathbf{x}[n]$ and $\mathbf{y}[n]$ are described by the joint distribution and density of the random variables

$$\mathbf{x}[n_1], \ldots, \mathbf{x}[n_k], \mathbf{y}[n'_1], \ldots, \mathbf{y}[n'_m]$$

$$F_{\mathbf{xy}}(x_1, \ldots, x_k, y_1, \ldots, y_m; n_1, \ldots, n_k, n'_1, \ldots, n'_m) = \\ P(\mathbf{x}[n_1] \le x_1, \ldots, \mathbf{x}[n_k] \le x_k, \mathbf{y}[n'_1] \le y_1, \ldots, \mathbf{y}[n'_m] \le y_m)$$

$$f_{\mathbf{xy}}(x_1, \ldots, x_k, y_1, \ldots, y_m; n_1, \ldots, n_k, n'_1, \ldots, n'_m) = \\ \frac{\partial^{k+m} F_{\mathbf{xy}}(x_1, \ldots, x_k, y_1, \ldots, y_m; n_1, \ldots, n_k, n'_1, \ldots, n'_m)}{\partial x_1 \cdots \partial x_k \partial y_1 \cdots \partial y_m}$$

- The *cross-correlation function* is

$$R_{\mathbf{xy}}[n_1, n_2] = E\{\mathbf{x}[n_1]\mathbf{y}[n_2]\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{\mathbf{xy}}(x, y; n_1, n_2)\, dx\, dy$$

- The *cross-covariance function* is

$$\begin{aligned} C_{\mathbf{xy}}[n_1, n_2] &= E\{(\mathbf{x}[n_1] - \mu_{\mathbf{x}}[n_1])(\mathbf{y}[n_2] - \mu_{\mathbf{y}}[n_2])\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_{\mathbf{x}}[n_1])(y - \mu_{\mathbf{y}}[n_2]) f_{\mathbf{xy}}(x, y; n_1, n_2)\, dx\, dy \end{aligned}$$

- Two processes $\mathbf{x}[n]$ and $\mathbf{y}[n]$ are *uncorrelated* if

$$C_{\mathbf{xy}}[n_1, n_2] = 0 \quad \text{for every } n_1 \text{ and } n_2$$

Comments on Complex-Valued Random Processes

- A *complex-valued* random process $\mathbf{z}[n,\zeta]$ maps each $\zeta \in \mathcal{S}$ to a complex-valued discrete-time sequence.

  1. If n and ζ are variable, the result is an *ensemble* of complex-valued discrete-time sequences $\mathbf{z}[n,\zeta]$.
  2. If n is variable and ζ is fixed, the result is a single complex-valued discrete-time sequence: a *sample* of the random process.
  3. If n is fixed and ζ is variable, the result is a complex-valued *random variable.*
  4. If n and ζ are fixed, the result is a *complex number.*

- The k-th order distribution and density of the complex-valued process $\mathbf{z}[n]$

  – Write

$$\begin{aligned} \mathbf{z}[n_1] &= \mathbf{x}[n_1] + j\mathbf{y}[n_1] & \qquad z_1 &= x_1 + jy_1 \\ &\vdots & &\vdots \\ \mathbf{z}[n_k] &= \mathbf{x}[n_k] + j\mathbf{y}[n_k] & \qquad z_k &= x_k + jy_k \end{aligned}$$

  – The k-th order distribution is the joint distribution of the complex-valued random variables $\mathbf{z}[n_1], \ldots, \mathbf{z}[n_k]$

$$\begin{aligned} F_{\mathbf{z}}(z_1,\ldots,z_k;n_1,\ldots,n_k) &= P(\mathbf{x}[n_1] \le x_1,\ldots,\mathbf{x}[n_k] \le x_k, \mathbf{y}[n_1] \le y_1,\ldots,\mathbf{y}[n_k] \le y_k) \\ &= F_{\mathbf{xy}}(x_1,\ldots,x_k,y_1\ldots,y_k;n_1,\ldots,n_k) \end{aligned}$$

  – The k-th order density of $\mathbf{z}[n]$ is expressed in terms of the (real-valued) real and imaginary components of $\mathbf{z}[n]$

$$\begin{aligned} f_{\mathbf{z}}(z_1,\ldots,z_k;n_1,\ldots,n_k) &= f_{\mathbf{xy}}(x_1,\ldots,x_k,y_1,\ldots,y_k;n_1,\ldots,n_k) \\ &= \frac{\partial^{2k}}{\partial x_1\cdots\partial x_k \partial y_1 \cdots \partial y_k} F_{\mathbf{xy}}(x_1,\ldots,x_k,y_1\ldots,y_k;n_1,\ldots,n_k) \end{aligned}$$

- First two moments

  – mean

$$\mu_{\mathbf{z}}[n] = \int_{-\infty}^{\infty} z f_{\mathbf{z}}(z;n)dz = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+jy) f_{\mathbf{xy}}(x,y;n)\,dx\,dy = \mu_{\mathbf{x}}[n] + j\mu_{\mathbf{y}}[n]$$

  – Autocorrelation $R_{\mathbf{zz}}[n_1,n_2] = E\{\mathbf{z}[n_1]\mathbf{z}^*[n_2]\}$

$$R_{\mathbf{zz}}[n_1,n_2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1+jy_1)(x_2-jy_2) f_{\mathbf{xy}}(x_1,x_2,y_1,y_2;n_1,n_2)\,dx_1\,dx_2\,dy_1\,dy_2$$

  – Autocovariance: $C_{\mathbf{zz}}[n_1,n_2] = R_{\mathbf{zz}}[n_1,n_2] - \mu_{\mathbf{z}}[n_1]\mu_{\mathbf{z}}^*[n_2]$

Stationary Processes

Definitions

A stochastic process $\mathbf{x}[n]$ is called *strict-sense stationary* (abbreviated SSS) if its statistical properties are invariant to to a shift of the time origin.

$\Rightarrow$ $\mathbf{x}[n]$ and $\mathbf{x}[n+c]$ have the same statistics.

$\Rightarrow$ $f_{\mathbf{x}}(x_1,\ldots,x_k;n_1,\ldots,n_k) = f_{\mathbf{x}}(x_1,\ldots,x_k;n_1{+}c,\ldots,n_k{+}c)$ for any integer c and for all k.

Properties

1. First-order density:

   (a) $f_{\mathbf{x}}(x;n) = f_{\mathbf{x}}(x;n+c) \Rightarrow f_{\mathbf{x}}(x;n) = f_{\mathbf{x}}(x)$

   (b) $\mu_{\mathbf{x}}[n] = \displaystyle\int_{-\infty}^{\infty} x f_{\mathbf{x}}(x;n)dx = \int_{-\infty}^{\infty} x f_{\mathbf{x}}(x)dx = \mu_{\mathbf{x}}$

2. Second-order density:

   (a)
$$\begin{aligned} f_{\mathbf{x}}(x_1,x_2;n_1,n_2) &= f_{\mathbf{x}}(x_1,x_2;n_1+c,n_2+c) \\ \Rightarrow f_{\mathbf{x}}(x_1,x_2;n_1,n_2) &= f_{\mathbf{x}}(x_1,x_2;n_1-n_2,0) \\ &= f_{\mathbf{x}}(x_1,x_2;m,0), \quad m = n_1 - n_2 \end{aligned}$$

   To paraphrase: "Thus the joint density of the random variables $\mathbf{x}[n+m]$ and $\mathbf{x}[n]$ is independent of [i.e., not a function of] n and it equals $f_{\mathbf{x}}(x_1,x_2;m)$."

   (b)
$$\begin{aligned} R_{\mathbf{xx}}[n_1,n_2] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2 f_{\mathbf{x}}(x_1,x_2;n_1n_2)\,dx_1\,dx_2 \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2 f_{\mathbf{x}}(x_1,x_2;m,0)\,dx_1\,dx_2 \\ &= R_{\mathbf{xx}}[m,0] \end{aligned}$$

   (c) It is customary to express the autocorrelation function for a SSS random process by
$$R_{\mathbf{xx}}[m] = E\{\mathbf{x}[n+m]\mathbf{x}[n]\} = E\{\mathbf{x}[n]\mathbf{x}[n-m]\}$$

   (d) $C_{\mathbf{xx}}[n_1,n_2] = C_{\mathbf{xx}}[m] = R_{\mathbf{xx}}[m] - \mu_{\mathbf{x}}^2$

Consequences

1. average power of SSS process $= R_{\mathbf{xx}}[0]$
2. variance of SSS process $= C_{\mathbf{xx}}[0]$
3. correlation coefficient of SSS process: $r_{\mathbf{xx}}[m] = \dfrac{C_{\mathbf{xx}}[m]}{C_{\mathbf{xx}}[0]}$

Definitions

A stochastic process $\mathbf{x}[n]$ is called *wide-sense stationary* (abbreviated WSS) if

$$f_{\mathbf{x}}(x;n) = f_{\mathbf{x}}(x;n+c)$$
$$f_{\mathbf{x}}(x_1,x_2;n_1,n_2) = f_{\mathbf{x}}(x_1,x_2;n_1+c,n_2+c)$$

Properties

1. $\mu_{\mathbf{x}}[n] = \mu_{\mathbf{x}}$
2. $R_{\mathbf{xx}}[n_1,n_2] = R_{\mathbf{xx}}[m]$

A stochastic process $\mathbf{x}[n]$ is *WSS white noise* means $C_{\mathbf{xx}}[m] = q\delta[m]$.

Comments on Complex-Valued WSS Random Processes

- The complex-valued WSS process $\mathbf{z}[n] = \mathbf{x}[n] + j\mathbf{y}[n]$ is described in terms of the joint statistics of the two real-valued processes $\mathbf{x}[n]$ and $\mathbf{y}[n]$.
- The first-order density property

$$f_{\mathbf{z}}(z;n) = f_{\mathbf{z}}(z;n+c) \Rightarrow f_{\mathbf{z}}(z;n) = f_{\mathbf{z}}(z)$$

  becomes

$$f_{\mathbf{xy}}(x,y;n) = f_{\mathbf{xy}}(x,y;n+c) \Rightarrow f_{\mathbf{xy}}(x,y;n) = f_{\mathbf{xy}}(x,y)$$

- The complex-valued mean is a constant:

$$\mu_{\mathbf{z}}[n] = \mu_{\mathbf{z}} \Rightarrow \mu_{\mathbf{x}}[n] + j\mu_{\mathbf{y}}[n] = \mu_{\mathbf{x}} + j\mu_{\mathbf{y}}$$

- The second-order density property

$$\begin{aligned} f_{\mathbf{z}}(z_1,z_2;n_1,n_2) &= f_{\mathbf{x}}(z_1,z_2;n_1+c,n_2+c) \\ &\Rightarrow f_{\mathbf{z}}(z_1,z_2;n_1,n_2) = f_{\mathbf{x}}(z_1,z_2;n_1-n_2,0) \end{aligned}$$

  becomes

$$\begin{aligned} f_{\mathbf{xy}}(x_1,x_2,y_1,y_2;n_1,n_2) &= f_{\mathbf{xy}}(x_1,x_2,y_1,y_2;n_1+c,n_2+c) \\ &\Rightarrow f_{\mathbf{xy}}(x_1,x_2,y_1,y_2;n_1,n_2) = f_{\mathbf{xy}}(x_1,x_2,y_1,y_2;n_1-n_2,0) \end{aligned}$$

- Autocorrelation function is

$$R_{\mathbf{zz}}[m] = E\{\mathbf{z}[n+m]z^*[n]\} = E\{\mathbf{z}[n]\mathbf{z}^*[n-m]\}$$

Properties of the auto- and cross-correlation functions

General Random Processes

1. $R_{\mathbf{xx}}[n_1, n_2] = E\{\mathbf{x}[n_1]\mathbf{x}^*[n_2]\}$
2. $R_{\mathbf{xx}}[n_2, n_1] = R^*_{\mathbf{xx}}[n_1, n_2]$
3. $R_{\mathbf{xx}}[n, n] \geq 0$
4. $R_{\mathbf{xy}}[n_1, n_2] = E\{\mathbf{x}[n_1]\mathbf{y}^*[n_2]\}$
5. $R_{\mathbf{yx}}[n_2, n_1] = R^*_{\mathbf{xy}}[n_1, n_2]$

WSS Random Processes

1. $R_{\mathbf{xx}}[m] = E\{\mathbf{x}[n+m]\mathbf{x}^*[n]\}$
2. $R_{\mathbf{xx}}[-m] = R^*_{\mathbf{xx}}[m]$
3. $R_{\mathbf{xx}}[0] \geq 0$
4. $R_{\mathbf{xy}}[m] = E\{\mathbf{x}[n+m]\mathbf{y}^*[n]\}$
5. $R_{\mathbf{yx}}[-m] = R^*_{\mathbf{xy}}[m]$
6. $R_{\mathbf{xx}}[m] \leq R_{\mathbf{xx}}[0]$

From Property 6 for WSS random processes

$R_{\mathbf{xx}}[m_1] = R_{\mathbf{xx}}[0]$ for some $m_1 \neq 0 \quad \Rightarrow \quad R_{\mathbf{xx}}[m+m_1] = R_{\mathbf{xx}}[m]$ for all m

$\Rightarrow \quad R_{\mathbf{xx}}[m]$ is *periodic* with period m_1

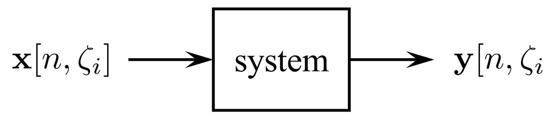
$R_{\mathbf{xx}}[1] = R_{\mathbf{xx}}[0] \quad \Rightarrow \quad R_{\mathbf{xx}}[m] = R_{\mathbf{xx}}[0]$ for all m.

Systems With Stochastic Inputs

Definitions

- A stochastic process $\mathbf{x}[n, \zeta]$ is a map from $\mathcal{S}$ to a real-valued discrete-time sequence.
  - For each $\zeta \in \mathcal{S}$, $\mathbf{x}[n, \zeta]$ is a *discrete-time signal.*
- Discrete-time system: input is a discrete-time signal $x[n]$. Output is another discrete-time signal $y[n]$.
  - If the input to a discrete-time system is a random process, then the input/relationship applies on a sample-by-sample basis.
  - For $\zeta_i \in \mathcal{S}$,

    $$\begin{aligned} \mathbf{x}[n, \zeta_i] &= \text{the input discrete-time signal} \\ \mathbf{y}[n, \zeta_i] &= \text{the output discrete-time signal} \end{aligned}$$

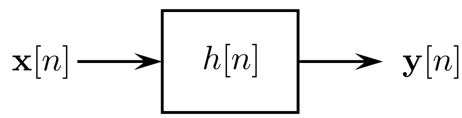


- Kinds of Systems
  - The system is *deterministic* if it operates only the variable n, treating ζ as a parameter.
  - The system is called *stochastic* if it operates on both variables n and ζ.
  - If the system is specified in terms of physical elements or by an equation,
    * the system is deterministic if the elements or coefficients of the defining equations are deterministic
    * the system is stochastic of the elements or coefficients of the defining equations are random
  - Memoryless Systems: a system is called *memoryless* if its input/output relationship is given by $\mathbf{y}[n] = g(\mathbf{x}[n])$.
  - LTI Systems: a linear time-invariant system is described in the recorded lectures. The input/output relationship is given by the *discrete-time convolution* of the input signal with the *impulse response* $h[n]$ of the system:

    $$\mathbf{y}[n, \zeta] = \sum_{k=-\infty}^{\infty} \mathbf{x}[k, \zeta] h[n-k] = \sum_{k=-\infty}^{\infty} \mathbf{x}[n-k, \zeta] h[k]$$

  - In this class, all LTI systems will be described by linear constant-coefficient difference equations *with all zero initial conditions*

LTI System with WSS Input



$$\mu_{\mathbf{y}} = \mu_{\mathbf{x}} \sum_{n=-\infty}^{\infty} h[n]$$

$$R_{\mathbf{xy}}[m] = \sum_{k=-\infty}^{\infty} R_{\mathbf{xx}}[m+k]h^*[k]$$

$$R_{\mathbf{yy}}[m] = \sum_{k=-\infty}^{\infty} R_{\mathbf{xy}}[m-k]h[k]$$

Special Case: $\mathbf{x}[n]$ is WSS white random process:

$$\mu_{\mathbf{x}} = 0 \qquad R_{\mathbf{xx}}[m] = q\delta[m]$$

$$\mu_{\mathbf{y}} = 0$$

$$R_{\mathbf{xy}}[m] = qh^*[-m]$$

$$R_{\mathbf{yy}}[m] = q \underbrace{\sum_{k=-\infty}^{\infty} h^*[k-m]h[k]}_{\rho[m]}$$

The Power Spectrum

Definitions

- The *power spectrum* (or *spectral density*) of a WSS discrete-time random process $\mathbf{x}[n]$, real or complex, is the DTFT of its autocorrelation function $R_{\mathbf{xx}}[m] = E\{\mathbf{x}[n+m]\mathbf{x}[n]\}$:

$$\mathsf{S}_{\mathbf{xx}}(e^{j\omega}) = \sum_{m=-\infty}^{\infty} R_{\mathbf{xx}}[m]e^{-j\omega m}$$

- From the DTFT inversion formula

$$R_{\mathbf{xx}}[m] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathsf{S}_{\mathbf{xx}}(e^{j\omega})e^{j\omega m}\,d\omega$$

- The *cross power spectrum* of two random processes $\mathbf{x}[n]$ and $\mathbf{y}[n]$ is the Fourier transform of their cross correlation $R_{\mathbf{xy}}[m] = E\{\mathbf{x}[n+m]\mathbf{y}^*[n]\}$:

$$\mathsf{S}_{\mathbf{xy}}(e^{j\omega}) = \sum_{m=-\infty}^{\infty} R_{\mathbf{xy}}[m]e^{-j\omega m}$$

- From the DTFT inversion formula

$$R_{\mathbf{xy}}[m] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathsf{S}_{\mathbf{xy}}(e^{j\omega})e^{j\omega m}\,d\omega$$

Properties

1. $\mathsf{S}_{\mathbf{xx}}(e^{j\omega})$ is a real-valued function of ω.
2. If $\mathbf{x}[n]$ is real, then $\mathsf{S}_{\mathbf{xx}}(e^{j\omega})$ is real and even.
3. $\mathsf{S}_{\mathbf{xx}}(e^{j\omega}) \geq 0$ for all ω.
4. $\mathsf{S}_{\mathbf{xy}}(e^{j\omega})$ is, in general, complex valued, even when both processes $\mathbf{x}[n]$ and $\mathbf{y}[n]$ are real-valued processes.
5. $\mathsf{S}_{\mathbf{xy}}(e^{j\omega}) = \mathsf{S}^*_{\mathbf{yx}}(e^{j\omega})$

The DTFT Version of TABLE 9-1

$R[m] = \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} \mathsf{S}(e^{j\omega})e^{j\omega k}\,d\omega$	$\leftrightarrow$	$\mathsf{S}(e^{j\omega}) = \displaystyle\sum_{m=-\infty}^{\infty} R[m]e^{-j\omega m}$
$\delta[m]$	$\leftrightarrow$	1
1	$\leftrightarrow$	$2\pi \displaystyle\sum_{\ell=-\infty}^{\infty} \delta(\omega - 2\pi\ell)$
$e^{j\beta m}$	$\leftrightarrow$	$2\pi \displaystyle\sum_{\ell=-\infty}^{\infty} \delta(\omega - \beta - 2\pi\ell)$
$\cos(\beta m)$	$\leftrightarrow$	$\pi \displaystyle\sum_{\ell=-\infty}^{\infty} \delta(\omega - \beta - 2\pi\ell) + \delta(\omega + \beta - 2\pi\ell)$
$\rho^{\lvert m\rvert}$	$\leftrightarrow$	$\dfrac{\rho e^{j\omega}}{1-\rho e^{j\omega}} + \dfrac{1}{1-\rho e^{-j\omega}}$ $= \dfrac{1}{1-\rho e^{j\omega}} + \dfrac{\rho e^{-j\omega}}{1-\rho e^{-j\omega}}$
$\rho^{\lvert m\rvert}\cos(\beta m)$	$\leftrightarrow$	$\dfrac{1}{2}\left[\dfrac{\rho e^{j(\omega-\beta)}}{1-\rho e^{j(\omega-\beta)}} + \dfrac{1}{1-\rho e^{-j(\omega-\beta)}} + \dfrac{\rho e^{j(\omega+\beta)}}{1-\rho e^{j(\omega+\beta)}} + \dfrac{1}{1-\rho e^{-j(\omega+\beta)}}\right]$ $= \dfrac{1}{2}\left[\dfrac{1}{1-\rho e^{j(\omega-\beta)}} + \dfrac{\rho e^{-j(\omega-\beta)}}{1-\rho e^{-j(\omega-\beta)}} + \dfrac{1}{1-\rho e^{j(\omega+\beta)}} + \dfrac{\rho e^{-j(\omega+\beta)}}{1-\rho e^{-j(\omega+\beta)}}\right]$
$\begin{cases} 1 & -M \le m \le M \\ 0 & \text{otherwise} \end{cases}$	$\leftrightarrow$	$\dfrac{\sin\left(\omega\left(\dfrac{2M+1}{2}\right)\right)}{\sin\left(\dfrac{\omega}{2}\right)}$

$$
\begin{aligned}
\rho^{|m|}\sin(\beta|m|) \leftrightarrow \frac{1}{j2}\Bigg[&\frac{1}{1-\rho e^{-j(\omega-\beta)}} - \frac{\rho e^{j(\omega-\beta)}}{1-\rho e^{j(\omega-\beta)}} + \\
&+ \frac{\rho e^{j(\omega+\beta)}}{1-\rho e^{j(\omega+\beta)}} - \frac{1}{1-\rho e^{-j(\omega+\beta)}}\Bigg] \\
= \frac{1}{j2}\Bigg[&\frac{\rho e^{-j(\omega-\beta)}}{1-\rho e^{-j(\omega-\beta)}} - \frac{1}{1-\rho e^{j(\omega-\beta)}} + \\
&+ \frac{1}{1-\rho e^{j(\omega+\beta)}} - \frac{\rho e^{-j(\omega-\beta)}}{1-\rho e^{-j(\omega+\beta)}}\Bigg]
\end{aligned}
$$

$$
|m|\rho^{|m|} \leftrightarrow \frac{\rho e^{j\omega}}{(1-\rho e^{j\omega})^2} + \frac{\rho e^{-j\omega}}{(1-\rho e^{-j\omega})^2}
$$

These are valid DTFT pairs, but the left-hand sides by themselves are not valid auto-correlation functions. (See Property 6.)

LTI System with WSS Input

$$\mathbf{x}[n] \longrightarrow \boxed{h[n]} \longrightarrow \mathbf{y}[n]$$

$$\mu_{\mathbf{y}} = \mu_{\mathbf{x}} \sum_{n-\infty}^{\infty} h[n] \qquad\qquad \mu_{\mathbf{y}} = \mu_{\mathbf{x}} \mathbf{H}(1)$$

$$R_{\mathbf{xy}}[m] = \sum_{n=-\infty}^{\infty} R_{\mathbf{xx}}[n+m]h^*[m] \qquad\qquad \mathbf{S}_{\mathbf{xy}}(e^{j\omega}) = \mathbf{S}_{\mathbf{xx}}(e^{j\omega})\mathbf{H}^*(e^{j\omega})$$

$$R_{\mathbf{yy}}[m] = \sum_{n-\infty}^{\infty} R_{\mathbf{xy}}[m-n]h[n] \qquad\qquad \mathbf{S}_{\mathbf{yy}}(e^{j\omega}) = \mathbf{S}_{\mathbf{xy}}(e^{j\omega})\mathbf{H}(e^{j\omega}) = \mathbf{S}_{\mathbf{xx}}(e^{j\omega})|\mathbf{H}(e^{j\omega})|^2$$

Example 9-27 Revisited

$$\mathbf{x}[n] \longrightarrow \boxed{h[n]} \longrightarrow \mathbf{y}[n]$$

$$\text{(a)} \quad \mathbf{y}[n] - b\mathbf{y}[n-1] = \mathbf{x}[n] \quad \text{all } n \quad \mu_{\mathbf{x}} = 0 \quad R_{\mathbf{xx}}[m] = q\delta[m]$$

$$\mathbf{H}(z) = \frac{1}{1 - bz^{-1}}$$

$$\mathbf{H}(e^{j\omega}) = \frac{1}{1 - be^{-j\omega}}$$

$$|\mathbf{H}(e^{j\omega})|^2 = \frac{1}{(1 - be^{-j\omega})(1 - be^{j\omega})}$$

$$= \frac{1}{1-b^2}\left[\frac{be^{-j\omega}}{1 - be^{-j\omega}} + \frac{1}{1 - be^{j\omega}}\right]$$

$$\mathbf{S}_{\mathbf{yy}}(e^{j\omega}) = \mathbf{S}_{\mathbf{xx}}(e^{j\omega})|\mathbf{H}(e^{j\omega})|^2$$

$$= \frac{q}{1-b^2}\left[\frac{be^{-j\omega}}{1 - be^{-j\omega}} + \frac{1}{1 - be^{j\omega}}\right]$$

$$R_{\mathbf{yy}}[m] = \frac{q}{1-b^2} b^{|m|}$$

$$\mathbf{H}(z) = \frac{1}{1 - bz^{-1}}$$

$$h[n] = b^n\, U[n]$$

$$R_{\mathbf{xy}}[m] = \begin{cases} qb^m & m \le 0 \\ 0 & m > 0 \end{cases}$$

$$R_{\mathbf{yy}}[m] = \begin{cases} \dfrac{q}{1-b^2} b^{-m} & m < 0 \\ \dfrac{q}{1-b^2} b^{m} & m \ge 0 \end{cases}$$

$$= \frac{q}{1-b^2} b^{|m|}$$

$$\rho[m] = \begin{cases} \displaystyle\sum_{n=0}^{\infty} b^{n-m} b^n & m < 0 \\ \displaystyle\sum_{n=m}^{\infty} b^{n-m} b^n & m \ge 0 \end{cases} = \begin{cases} \dfrac{b^{-m}}{1-b^2} & m < 0 \\ \dfrac{b^{m}}{1-b^2} & m \ge 0 \end{cases} = \frac{b^{|m|}}{1-b^2}$$

$$R_{\mathbf{yy}}[m] = q\rho[m] = \frac{q}{1-b^2} b^{|m|}$$

$$E\{\mathbf{y}^2[n]\} = R_{\mathbf{yy}}[0] = \frac{q}{1-b^2}$$

Example 9-27 Revisited

$$\mathbf{x}[n] \longrightarrow \boxed{h[n]} \longrightarrow \mathbf{y}[n]$$

(b) $\quad \mathbf{y}[n] - b\mathbf{y}[n-1] + c\mathbf{y}[n-2] = \mathbf{x}[n] \quad \text{all } n \quad \mu_{\mathbf{x}} = 0 \quad R_{\mathbf{xx}}[m] = q\delta[m]$

$$\mathbf{H}(z) = \frac{1}{1 - bz^{-1} + cz^{-2}}$$

$$\mathbf{H}(e^{j\omega}) = \frac{1}{1 - be^{-j\omega} + ce^{-j2\omega}}$$

$$|\mathbf{H}(e^{j\omega})|^2 = \frac{1}{(1 - be^{-j\omega} + ce^{-j2\omega})(1 - be^{j\omega} + ce^{j2\omega})}$$

$$\mathbf{S}_{\mathbf{yy}}(e^{j\omega}) = \mathbf{S}_{\mathbf{xx}}(e^{j\omega})|\mathbf{H}(e^{j\omega})|^2 = \frac{q}{(1 - be^{-j\omega} + ce^{-j2\omega})(1 - be^{j\omega} + ce^{j2\omega})}$$

$\underline{b^2 < 4c}$

$$R_{\mathbf{yy}}[m] = \frac{q}{(1+c)^2 - b^2}\,\rho^{|m|}\left[\frac{1+c}{1-c}\cos(\beta m) + \frac{b}{2\gamma}\sin(\beta|m|)\right]$$

$$\alpha = \frac{b}{2}, \quad \gamma^2 + \alpha^2 = c, \quad \rho^2 = \alpha^2 + \gamma^2 = c, \quad \beta = \operatorname{atan}\left(\frac{\gamma}{\alpha}\right)$$

$\underline{b^2 = 4c}$

$$R_{\mathbf{yy}}[m] = \frac{q}{(1-c)^2}\left[|m| + \frac{1+c}{1-c}\right]\alpha^{|m|} \qquad \alpha = \frac{b}{2}$$

$\underline{b^2 > 4c}$

$$R_{\mathbf{yy}}[m] = \frac{q}{2\gamma(1-c)}\left[\frac{\alpha+\gamma}{1-(\alpha+\gamma)^2}(\alpha+\gamma)^{|m|} - \frac{\alpha-\gamma}{1-(\alpha-\gamma)^2}(\alpha-\gamma)^{|m|}\right]$$

$$\alpha = \frac{b}{2}, \qquad \alpha^2 - \gamma^2 = c$$

In all cases, $E\{\mathbf{y}^2[n]\} = \dfrac{q(1+c)}{(1-c)[(1+c)^2 - b^2]}$.

Complex Random Variables

Chapters 2- 7

Page 75: A *complex* random variable $\mathbf{z}$ is a sum

$$\mathbf{z} = \mathbf{x} + j\mathbf{y}$$

where $\mathbf{x}$ and $\mathbf{y}$ are real random variables.

Page 77: **Note** A complex random variable $\mathbf{z} = \mathbf{x} + j\mathbf{y}$ has no distribution function because the inequality $\mathbf{x} + j\mathbf{y} \leq x + jy$ has no meaning.* The statistical properties of $\mathbf{z}$ are specified in terms of the joint distribution of the random variables $\mathbf{x}$ and $\mathbf{y}$ (see Chap. 6).

Note to students: *This statement has no meaning for two reasons. First, following the definition for real random variables, the distribution function of the complex variable $\mathbf{z}$ must be

$$F_{\mathbf{z}}(z) = \mathrm{P}\{ \mathbf{z} \leq z \}.$$

Second, anyone from the math department will tell you that complex numbers are not *ordered*. Ordered means that from the natural definition of complex numbers one cannot say one complex number is greater than another. Sure, we can conceive of all sorts of possible measures such as "greater means greater in magnitude" or $z = x + jy$ is greater than $w = u + jv$ is both $x > u$ and $y > v$." But in both cases, we are imposing a definition after the fact.

Page 143: ***Complex random variables:*** If $\mathbf{z} = \mathbf{x} + j\mathbf{y}$ is a complex random variable, then its expected value is by definition

$$E\{\mathbf{z}\} = E\{\mathbf{x}\} + jE\{\mathbf{y}\}$$

Note to students: The variance of a complex random variable is not defined here (in Chap. 5) because joint distributions have not yet been introduced: the variance of a complex random variable is defined in terms of the covariance – so we have to wait.

Section 6-1, Pages 169– 179:

Here, joint distributions and densities are introduced. This is the natural place to define the distribution and density functions of the complex random variable $\mathbf{z} = \mathbf{x} + j\mathbf{y}$. But I cannot find the definitions. So here they are

$$z = x + jy$$

$$\downarrow$$

Distribution function of $\mathbf{z} = \mathbf{x} + j\mathbf{y}$: $F_{\mathbf{z}}(z) = F_{\mathbf{xy}}(x,y) = P\{\mathbf{x} \le x,\ \mathbf{y} \le y\}$

Density function of $\mathbf{z} = \mathbf{x} + j\mathbf{y}$: $f_{\mathbf{z}}(z) = f_{\mathbf{xy}}(x,y) = \dfrac{\partial}{\partial x}\dfrac{\partial}{\partial y} F_{\mathbf{xy}}(x,y)$

Example 6-15 examines the (complex to complex) transformation $\mathbf{x} + j\mathbf{y} \to \mathbf{r}e^{j\boldsymbol{\theta}}$

1. From the definition of a complex random variable, $\mathbf{x} + j\mathbf{y}$ is described by the joint density function $f_{\mathbf{xy}}(x,y)$.
2. The complex random variable $\mathbf{r}e^{j\boldsymbol{\theta}}$ can also be described a joint density function. Here the joint density function is $f_{\mathbf{r\theta}}(r,\theta)$.
3. The transformation $\mathbf{x} + j\mathbf{y} \to \mathbf{r}e^{j\boldsymbol{\theta}}$ can be thought of as the 2×2 transformation

$$\mathbf{r} = \sqrt{\mathbf{x}^2 + \mathbf{y}^2}$$
$$\boldsymbol{\theta} = \tan^{-1}\left(\frac{\mathbf{y}}{\mathbf{x}}\right)$$

This transformation is examined in Example 6-22 (pp. 202-203).

For the special case where $\mathbf{x}$ and $\mathbf{y}$ are independent normal random variables with zero mean and common variance σ^2, the joint density function of $\mathbf{r}$ and $\boldsymbol{\theta}$ is

$$f_{\mathbf{r}\boldsymbol{\theta}}(r,\theta) = \frac{r}{2\pi\sigma^2}\exp\left\{-\frac{r^2}{2\sigma^2}\right\}$$

This joint density function describes the complex random variable $\mathbf{r}e^{j\boldsymbol{\theta}}$.

Page 249: **CORRELATION AND COVARIANCE MATRICES**. The covariance C_{ij} of two real random variables $\mathbf{x}_i$ and $\mathbf{x}_j$ is defined as in (6-163). For complex random variables

$$C_{ij} = E\left\{(\mathbf{x}_i - \eta_i)\left(\mathbf{x}_j^* - \eta_j^*\right)\right\} = E\left\{\mathbf{x}_i \mathbf{x}_j^*\right\} - \eta_i \eta_j^*$$

by definition. The variance of $\mathbf{x}_i$ is given by

$$\sigma_i^2 = E\left\{|\mathbf{x}_i - \eta_i|^2\right\} = E\{|\mathbf{x}_i|^2\} - |E\{\mathbf{x}_i\}|^2$$

COVARIANCE. The covariance C or C_{xy} of two [real] random variables $\mathbf{x}$ and $\mathbf{y}$ is by definition the number

$$C_{xy} = E\{(\mathbf{x} - \eta_x)(\mathbf{y} - \eta_y)\} \qquad \text{(6-123)}$$

where $E\{\mathbf{x}\} = \eta_x$ and $E\{\mathbf{y}\} = \eta_y$. Expanding the product in (6-123) and using (6-121) we obtain

$$C_{xy} = E\{\mathbf{xy}\} - E\{\mathbf{x}\}E\{\mathbf{y}\}$$

Notes to class:

1. The definition for the covariance of two complex random variables (at the top of this page) is first given in Chapter 7, Sequences of Random Variables. Clearly, the definition could have been given earlier. I do not know why it was not.
2. A sequence of random variables is simply a list of random variables, say $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. That is $\mathbf{x}_i$ and $\mathbf{x}_j$ from the list are two different random variables for $i \neq j$. One could call these two different random variables $\mathbf{x}$ and $\mathbf{y}$. Then the definition would have been the more familiar
$$C_{xy} = E\left\{(\mathbf{x} - \eta_x)(\mathbf{y} - \eta_y)^*\right\} = E\{\mathbf{xy}^*\} - E\{\mathbf{x}\}E\{\mathbf{y}^*\}$$
3. The term $E\{\mathbf{xy}^*\}$ or $E\{\mathbf{x}_i\mathbf{x}_j^*\}$ above is the *correlation*. I am not sure why the definition is not made explicit.

The Empty Set: Much Ado About Nothing

ECEn 670: Stochastic Processes

The great thing about mathematics is that it allows one to be upset over nothing.

–Anonymous

Definitions

On page 15, the text reads [1]

> The *empty* or *null* set is by definition the set that contains no elements. This set will be denoted by $\{\emptyset\}$.

There are several issues with this two-sentence paragraph.

1. The terms *empty set* and *null set* used to be used interchangeably. The term *null set* now has a technical definition in measure theory [2] that sets it apart from *empty set*. So, for this class, the term *empty set* is used.

2. The definition "the empty set is the set that contains no elements" works for us.

3. Notation: "This set will be denoted by $\{\emptyset\}$" is not consistent with the widely adopted conventions in mathematical set theory. In fact, I would go as far as saying the notation is incorrect. To be consistent with the conventions in set theory, the sentence should read

   > This set will be denoted by $\emptyset$.

   That is, the symbol $\emptyset$ denotes the set with no elements. Curiously, this is the notation used in the lecture slides provided by U. Pillai on the publisher's website [3]. To put the nail in the coffin, we have from Malcolm Graham [4]

   > Note that the empty set is not designated by $\{\varnothing\}$; this notation would represent a set containing one element $\varnothing$, rather than the set with no elements. Similarly, the set $\{0\}$ contains the element zero and hence is not the empty set.

The third issue warrants a little more discussion. The cardinality of the set A, denoted $|A|$ is defined as the number of elements in the set. Some examples are

$$\begin{aligned} A &= \{a\} & |A| &= 1 \\ A &= \{a, b, c\} & |A| &= 3 \\ A &= \{\{a, b\}, c\} & |A| &= 2. \end{aligned}$$

Because $\emptyset$ contains no elements $|\emptyset| = 0$. But the notation $\{\emptyset\}$ means the set containing one element: the empty set. So we have $|\{\emptyset\}| = 1$.

Notation

As of this writing, the notation for the empty set most commonly encountered is either "$\{\ \}$" or "$\varnothing$". With reference to the end of the previous section, the set containing the empty set is $\{\{\}\}$ or $\{\varnothing\}$, respectively.

The symbol $\varnothing$ for the empty set was introduced by Bourbaki in [5, pg. 4]:

> certaines propriétés ... ne sont vraies pour *aucun* élément de E ... la partie qu'elles définissent est appelée la *partie vide* de E, et designée par la notation $\varnothing$."

On the choice of the symbol $\varnothing$, André Weil (1906–1998), a student of Bourbaki, explains in [6, pg 114] (translation by Jennifer Gage via [7])

> Wisely, we had decided to publish an installment establishing the system of notation for set theory, rather than wait for the detailed treatment that was to follow: it was high time to fix these notations once and for all, and indeed the ones we proposed, which introduced a number of modifications to the notations previously in use, met with general approval. Much later, my own part in these discussions earned me the respect of my daughter Nicolette, when she learned the symbol $\varnothing$ for the empty set at school and I told her that I had been personally responsible for its adoption. The symbol came from the Norwegian alphabet, with which I alone among the Bourbaki group was familiar.

The text uses the symbol $\emptyset$ to represent the empty set. Comments:

1. The symbol $\emptyset$ looks like a zero with a slash through it. In LaTeX, the symbol is created using $\emptyset$.

2. Most authors (including your instructor) believe \emptyset produces the ugliest empty set symbol in the solar system. Evidently, the keepers of LaTeX also believe this and have provided the \varnothing command that produces the much better looking $\varnothing$. (\varnothing needs the amssymb package.)

3. In Unicode, the empty set symbol $\varnothing$ (∅) occupies code point U+2205. But many fonts in use today do not include this character and render it as a small rectangle.

4. The Greek letters ϕ (lowercase) and Φ (uppercase) are not equivalent to the symbol for the empty set, and should not be used. Never. Ever.

$$\phi \neq \varnothing \qquad \Phi \neq \varnothing.$$

The Ohio State University

$\varnothing$ is *the* empty set, not *an* empty set. This is because the empty set is unique. Two sets are equal if they both contain the same elements. Consequently, there can only be one set with no elements. In other words, if there were more than one empty set, they would be equal.

Von Neumann's Definition of Ordinals

Ordinals are an extension to the natural numbers (non-negative integers). In the simplest terms, natural numbers count things (they answer the question "how many?") and ordinal numbers tell the position of something (its "order") in a list. For a finite number of things to be counted or ordered, there is little difference between natural numbers and the ordinal numbers. Where things get interesting to mathematicians is the case where an infinite number of things need to be ordered. The trick is to identify the definition of a number that satisfies all the required mathematical properties in the limit.

Von Neumann defined ordinals in terms of the cardinality of sets. If α is the set representing an ordinal, then the next ordinal is defined by the cardinality of the sucessor set $S_\alpha = \{\alpha \cup \{\alpha\}\}$.

The starting point is the empty set! The first few Von Neumann ordinals are

$$\begin{aligned}
&0 \leftarrow \varnothing \\
&1 \leftarrow 0 \cup \{0\} = \varnothing \cup \{\varnothing\} = \{\varnothing\} \\
&2 \leftarrow 1 \cup \{1\} = \{\varnothing\} \cup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\} \\
&3 \leftarrow 2 \cup \{2\} = \{\varnothing, \{\varnothing\}\} \cup \{\{\varnothing, \{\varnothing\}\}\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} \\
&4 \leftarrow 3 \cup \{3\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}
\end{aligned}$$

References

[1] A. Papoulis and U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. Boston: McGraw-Hill, 2002.

[2] Wikipedia, "Measure (mathematics)," 2017. [Online]. Available: https://en.wikipedia.org/wiki/Measure_(mathematics)

[3] McGraw-Hill Companies, "Probability, random variables, and stochastic processes: Instructor resources," 2001. [Online]. Available: http://www.mhhe.com/engcs/electrical/papoulis/ippt.mhtml

[4] M. Graham, *Modern Elementary Mathematics*, 4th ed. San Diego, CA: Harcourt College Publishers, 1984.

[5] N. Bourbaki, *Éléments de mathématique Fasc.1: Les structures fondamentales de l'analyse; Liv.1: Theorie de ensembles.* Paris: Hermann & C[ie] Éditeurs, 1939.

[6] A. Weil, *The Apprenticeship of a Mathematician*. Basel-Boston-Berlin: Birkhaeuser Verlag, 1992.

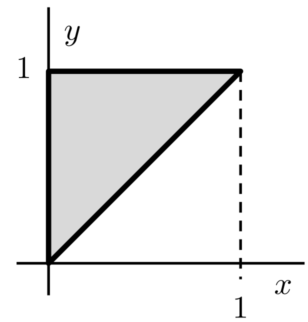
[7] J. Miller, "Earliest uses of symbols of set theory and logic," 2017. [Online]. Available: http://jeff560.tripod.com/set.html

Example 6-40: The Extended Version

ECEn 670: Stochastic Processes

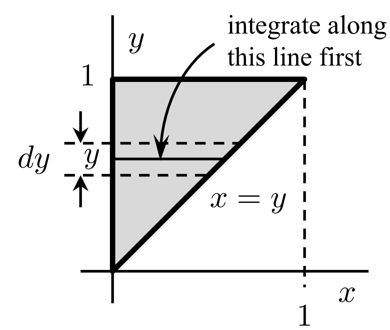
The joint density of the random variables $\mathbf{x}$ and $\mathbf{y}$ is

$$f_{\mathbf{xy}}(x,y) = \begin{cases} 2 & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$



The region of support for the joint density $f_{\mathbf{xy}}(x,y)$ is shown by the gray area in the figure to the right. The joint density function $f_{\mathbf{xy}}(x,y)$ is the constant 2 over the gray region of support and zero everywhere else in the (x,y) plane. To see that this is a valid joint density function, the double integral must be one. The double integral can be formulated two ways: $dx\,dy$ or $dy\,dx$.

The "$dx\,dy$" order is developed with the aid of the figure to the right.



1. Pick a value of y. For $0 < y < 1$, the value of y defines a horizontal line that passes through the region of support of $f_{\mathbf{xy}}(x,y)$.

2. The *area* under the slice defined by the horizontal line $y = y$ is

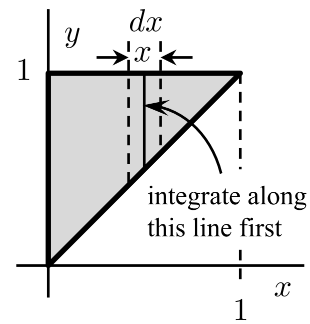
$$\text{area of the "slice" } y = y = \int_{x=0}^{y} 2\,dx$$

3. Multiplying the area under the slice $y = y$ by the incremental width dy creates the incremental *volume* for the slice.

4. The *volume* of the joint density $f_{\mathbf{xy}}(x,y)$ is obtained by summing the incremental volumes. In the limit $dy \to 0$, the sum becomes the integral:

$$\text{volume} = \int_{y=0}^{1} \int_{x=0}^{y} 2\,dx\,dy = \int_{y=0}^{1} 2y\,dy = 1.$$

The "$dy\,dx$" order is developed with the aid of the figure to the right.



1. Pick a value of x. For $0 < x < 1$, the value of x defines a vertical line that passes through the region of support of $f_{\mathbf{xy}}(x, y)$.

2. The *area* under the slice defined by the vertical line $x = x$ is

$$\text{area of the ``slice'' } x = x = \int_{y=x}^{1} 2\,dy$$

3. Multiplying the area under the slice $y = y$ by the incremental width dx creates the incremental *volume* for the slice.

4. The *volume* of the joint density $f_{\mathbf{xy}}(x, y)$ is obtained by summing the incremental volumes. In the limit $dx \to 0$, the sum becomes the integral:

$$\text{volume} = \int_{x=0}^{1} \int_{y=x}^{1} 2\,dy\,dx = \int_{x=0}^{1} 2(1-x)\,dx = 1.$$

The marginal density $f_{\mathbf{x}}(x)$: The marginal density $f_{\mathbf{x}}(x)$ is obtained from the joint density by integrating with respect to the unwanted variable y:

$$f_{\mathbf{x}}(x) = \int_{-\infty}^{\infty} f_{\mathbf{xy}}(x, y)\,dy.$$

For a given x, the integral with respect to y is along the line $x = x$:

$$\begin{aligned} f_{\mathbf{x}}(x) &= \begin{cases} \displaystyle\int_{y=x}^{1} 2\,dy & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 2(1-x) & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

The marginal density function is plotted in Figure 1.

The marginal density $f_{\mathbf{y}}(y)$: The marginal density $f_{\mathbf{y}}(y)$ is obtained from the joint density by

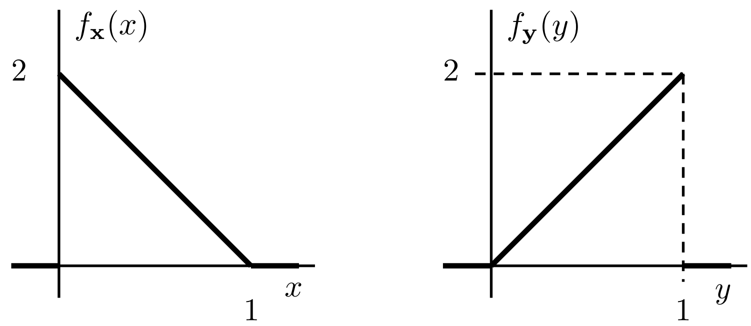


Figure 1: The marginal density functions.

integrating with respect to the unwanted variable x:

$$f_{\mathbf{y}}(y) = \int_{-\infty}^{\infty} f_{\mathbf{xy}}(x, y)\, dx.$$

For a given y, the integral with respect to x is along the line $y = y$:

$$\begin{aligned} f_{\mathbf{y}}(y) &= \begin{cases} \displaystyle\int_{x=0}^{x} 2\, dx & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 2y & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

The marginal density function is plotted in Figure 1.

Marginal statistics:

$$\begin{aligned} \mu_{\mathbf{x}} &= \int_{-\infty}^{\infty} x f_{\mathbf{x}}(x)\, dx = \int_0^1 x 2(1-x)\, dx = \frac{1}{3} \\ \sigma_{\mathbf{x}}^2 &= \int_{-\infty}^{\infty} (x - \mu_{\mathbf{x}})^2 f_{\mathbf{x}}(x)\, dx = \int_0^1 \left(x - \frac{1}{3}\right)^2 2(1-x)\, dx = \frac{1}{18} \\ \mu_{\mathbf{y}} &= \int_{-\infty}^{\infty} y f_{\mathbf{y}}(y)\, dy = \int_0^1 y 2y\, dy = \frac{2}{3} \\ \sigma_{\mathbf{y}}^2 &= \int_{\infty}^{\infty} (y - \mu_{\mathbf{y}})^2 f_{\mathbf{y}}(y)\, dy = \int_0^1 \left(y - \frac{2}{3}\right)^2 2y\, dy = \frac{1}{18}. \end{aligned}$$

Joint statistics:

$$\begin{aligned}
C_{\mathbf{xy}} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_{\mathbf{x}})(y-\mu_{\mathbf{y}})f_{\mathbf{xy}}(x,y)\,dx\,dy \\
&= \int_{y=0}^{1}\int_{x=0}^{y}\left(x-\frac{1}{3}\right)\left(y-\frac{2}{3}\right)2\,dx\,dy \\
&= 2\int_{y=0}^{1}\left(y-\frac{2}{3}\right)\left(\frac{1}{2}y^2-\frac{1}{3}y\right)\,dy \\
&= \frac{1}{36}
\end{aligned}$$

$$\rho_{\mathbf{xy}} = \frac{C_{\mathbf{xy}}}{\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}} = \frac{\frac{1}{36}}{\sqrt{\frac{1}{18}\frac{1}{18}}} = \frac{1}{2}$$

$$\begin{aligned}
R_{\mathbf{xy}} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{\mathbf{xy}}(x,y)\,dx\,dy \\
&= \int_{y=0}^{1}\int_{x=0}^{y} xy2\,dx\,dy \\
&= \int_{y=0}^{1} y\,y^2\,dy \\
&= \frac{1}{4}.
\end{aligned}$$

Note that the covariance may also be computed using

$$C_{\mathbf{xy}} = R_{\mathbf{xy}} - \mu_{\mathbf{x}}\mu_{\mathbf{y}} = \frac{1}{4} - \frac{1}{3}\times\frac{2}{3} = \frac{1}{36}.$$

Conditional densities and expectations:

$$f_{\mathbf{x}|\mathbf{y}}(x|y) = \frac{f_{\mathbf{xy}}(x,y)}{f_{\mathbf{y}}(y)} = \begin{cases} \dfrac{2}{2y} = \dfrac{1}{y} & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_{\mathbf{y}|\mathbf{x}}(y|x) = \frac{f_{\mathbf{xy}}(x,y)}{f_{\mathbf{x}}(x)} = \begin{cases} \dfrac{2}{2(1-x)} = \dfrac{1}{1-x} & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

The behavior of $f_{\mathbf{x}|\mathbf{y}}(x|y)$ is illustrated by examining $f_{\mathbf{x}|\mathbf{y}}(x|y)$ at a few trial values of y:

$$f_{\mathbf{x}|\mathbf{y}}(x|\mathbf{y} = 0.2) = \begin{cases} 5 & 0 < x < 0.2 \\ 0 & \text{otherwise} \end{cases}$$

$$f_{\mathbf{x}|\mathbf{y}}(x|\mathbf{y} = 0.5) = \begin{cases} 2 & 0 < x < 0.5 \\ 0 & \text{otherwise} \end{cases}$$

$$f_{\mathbf{x}|\mathbf{y}}(x|\mathbf{y} = 0.8) = \begin{cases} 1.25 & 0 < x < 0.8 \\ 0 & \text{otherwise} \end{cases}$$

These cases are plotted in Figure 2 (a).

The behavior of $f_{\mathbf{y}|\mathbf{x}}(y|x)$ is illustrated by examining $f_{\mathbf{y}|\mathbf{x}}(y|x)$ at a few trial values of x:

$$f_{\mathbf{y}|\mathbf{x}}(y|\mathbf{x} = 0.2) = \begin{cases} 1.25 & 0.2 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_{\mathbf{y}|\mathbf{x}}(y|\mathbf{x} = 0.5) = \begin{cases} 2 & 0.5 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_{\mathbf{y}|\mathbf{x}}(y|\mathbf{x} = 0.8) = \begin{cases} 5 & 0.5 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

These cases are plotted in Figure 2 (b).

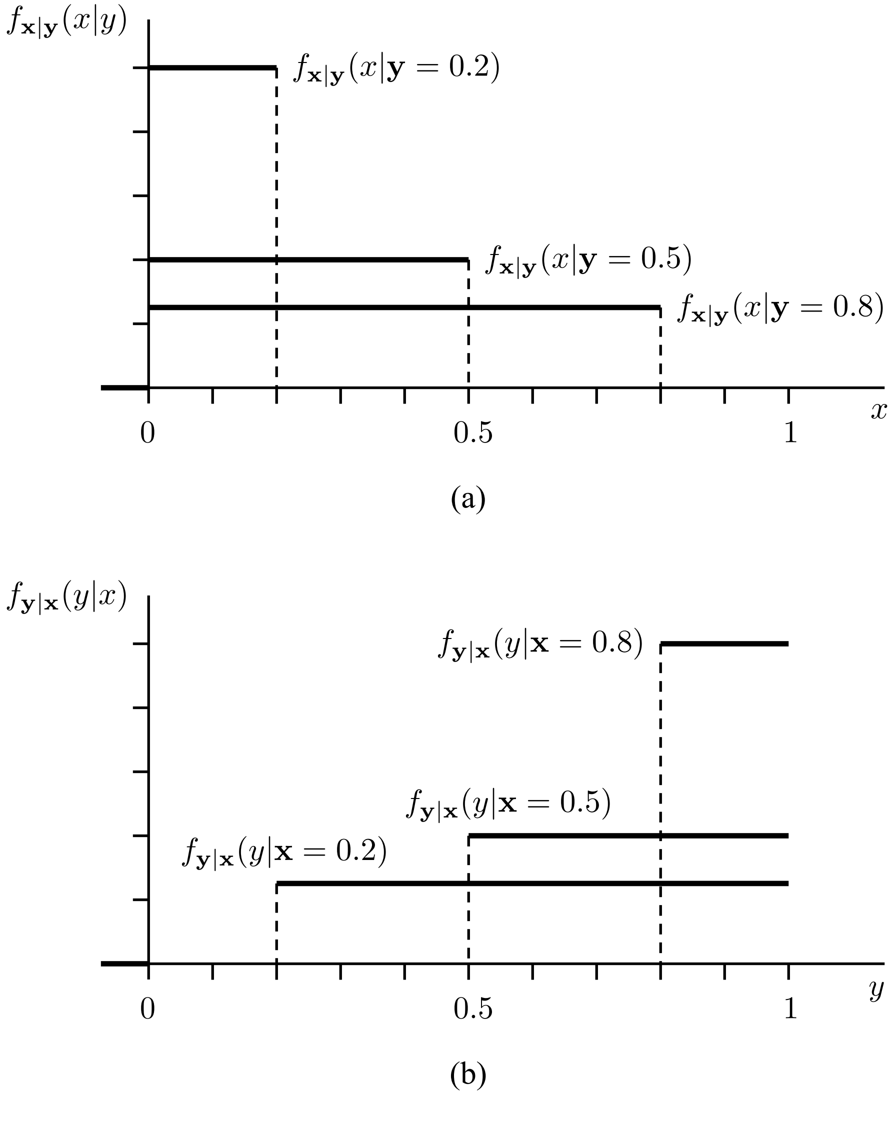


Figure 2: The conditional density functions.

The conditional expectation $E\{\mathbf{x}|\mathbf{y}=y\}$ is computed as follows:

$$\begin{aligned}
\mu_{\mathbf{x}|y} &= E\{\mathbf{x}|\mathbf{y}=y\} \\
&= \int_{-\infty}^{\infty} x f_{\mathbf{x}|\mathbf{y}}(x|\mathbf{y}=y)\,dx \\
&= \begin{cases} \displaystyle\int_{x=0}^{y} x\frac{1}{y}\,dx & 0<y<1 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \dfrac{y}{2} & 0<y<1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

$E\{\mathbf{x}|\mathbf{y}=y\}$ is a function of y. (This is true, in general.) A plot of $E\{\mathbf{x}|\mathbf{y}=y\}$ vs. y is shown in Figure 3 (a). $E\{\mathbf{x}|\mathbf{y}=y\}$ as a function of y is called a "regression line" even though in general it is not a "line."

Replacing the y in the conditional expectation with $\mathbf{y}$ produces the random variable

$$g(\mathbf{y}) = \begin{cases} \dfrac{\mathbf{y}}{2} & 0<\mathbf{y}<1 \\ 0 & \text{otherwise.} \end{cases}$$

The expected value of this random variable is

$$E\{g(\mathbf{y})\} = \int_{-\infty}^{\infty} g(y) f_{\mathbf{y}}(y)dy = \int_0^1 \frac{y}{2}\,2y\,dy = \frac{1}{3} = \mu_{\mathbf{x}}.$$

This demonstrates the property

$$E\{\,E\{\,\mathbf{x}|\mathbf{y}\,\}\,\} = E\{\mathbf{x}\}.$$

The conditional expectation $E\{\mathbf{y}|\mathbf{x}=x\}$ is computed as follows:

$$\begin{aligned}
\mu_{\mathbf{y}|x} &= E\{\mathbf{y}|\mathbf{x}=x\} \\
&= \int_{-\infty}^{\infty} y f_{\mathbf{y}|\mathbf{x}}(y|\mathbf{x}=x)\,dy \\
&= \begin{cases} \displaystyle\int_{y=x}^{1} y\frac{1}{1-x}\,dy & 0<x<1 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \dfrac{x+1}{2} & 0<x<1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

$E\{\mathbf{y}|\mathbf{x}=x\}$ is a function of x. A plot of $E\{\mathbf{y}|\mathbf{x}=x\}$ vs. x is shown in Figure 3 (b). $E\{\mathbf{y}|\mathbf{x}=x\}$ as a function of x is called a "regression line" even though in general it is not a "line."

Replacing x in the conditional expectation with $\mathbf{x}$ produces the random variable

$$g(\mathbf{x}) = \begin{cases} \displaystyle\int_{y=x}^{1} y\frac{1}{1-\mathbf{x}}\,dy & 0<\mathbf{x}<1 \\ 0 & \text{otherwise} \end{cases}$$

The expected value of this random variable is

$$E\{g(\mathbf{x})\} = \int_{-\infty}^{\infty} g(x) f_{\mathbf{x}}(x)\,dx = \int_{0}^{1} \frac{x+1}{2}\,2(1-x)\,dx = \frac{2}{3} = \mu_{\mathbf{y}}.$$

This demonstrates the property

$$E\{\,E\{\,\mathbf{y}|\mathbf{x}\,\}\,\} = E\{\mathbf{y}\}.$$

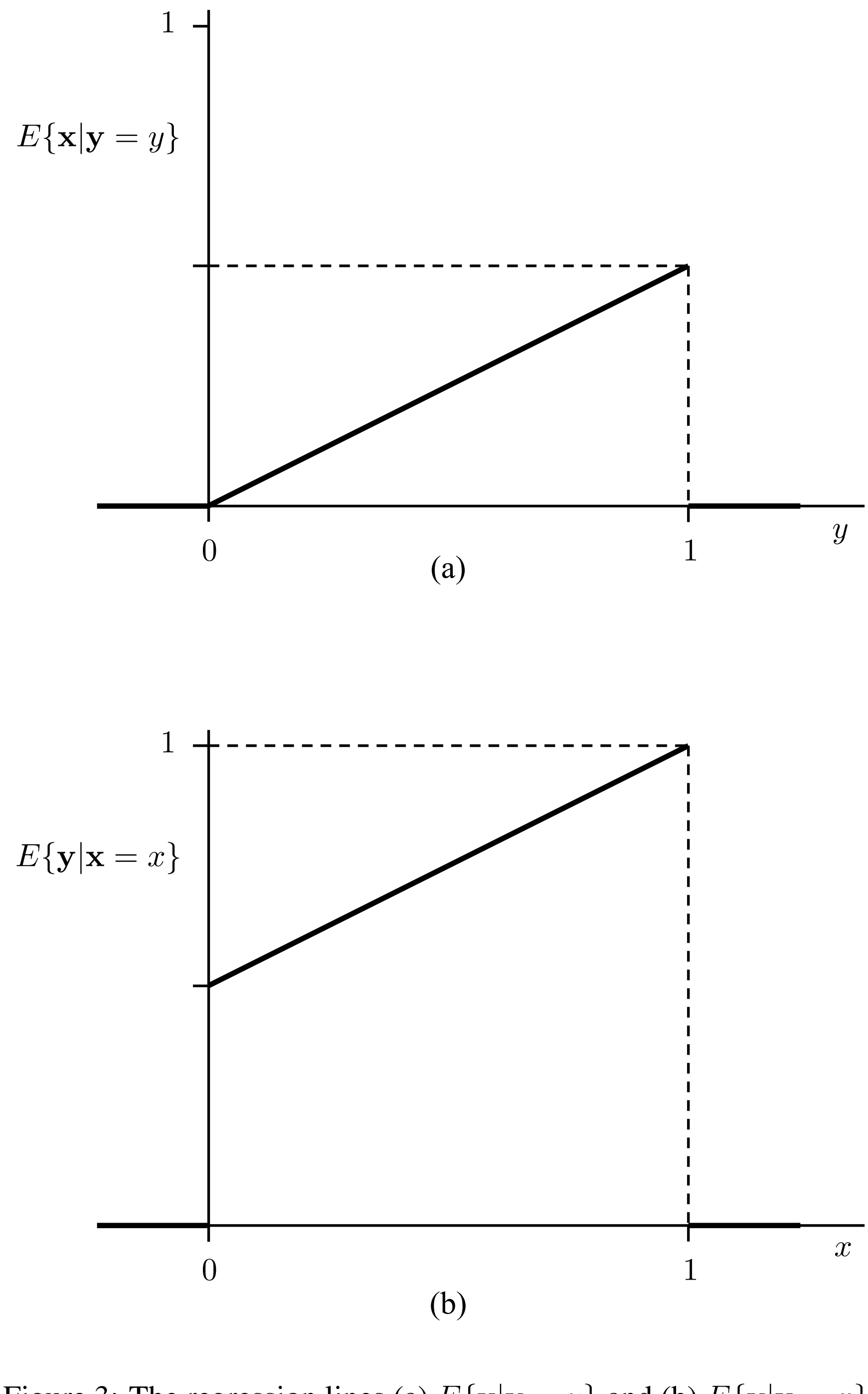
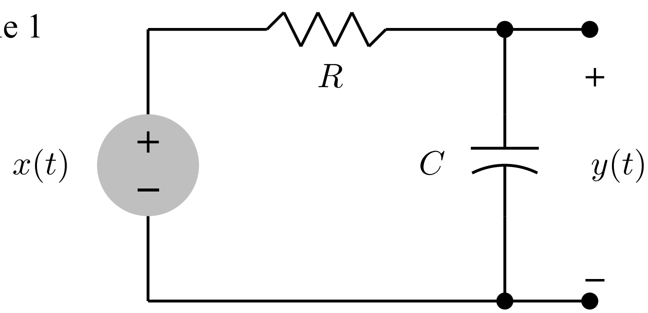


Figure 3: The regression lines (a) $E\{\mathbf{x}|\mathbf{y} = y\}$ and (b) $E\{\mathbf{y}|\mathbf{x} = x\}$.

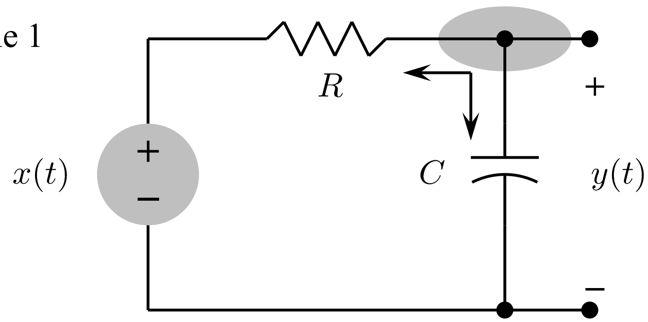
A Review of Continuous Time Linear Systems

- Description of Linear Systems
- Frequency Domain Analysis

Example 1

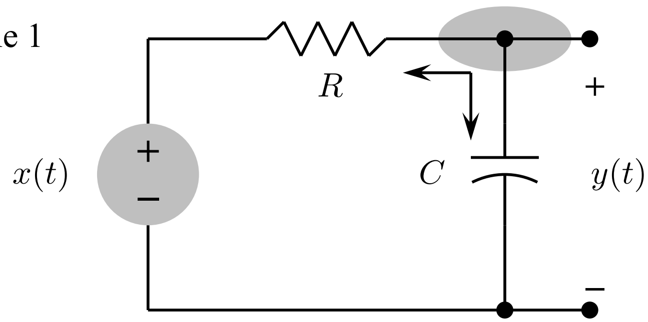


Example 1



sum of the currents leaving the node $= 0$

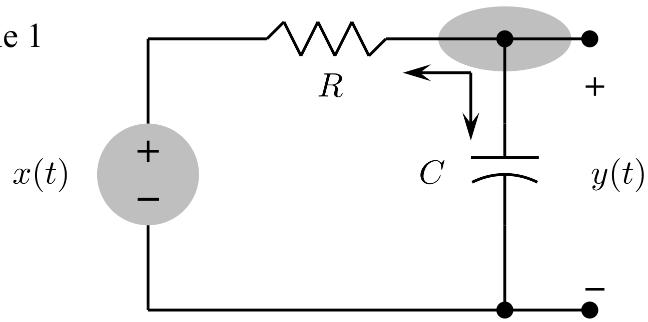
Example 1



sum of the currents leaving the node $= 0$

$$\frac{y(t)-x(t)}{R} + C\frac{d}{dt}y(t) = 0$$

Example 1

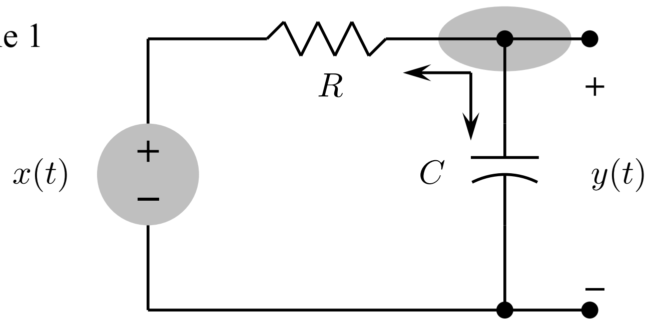


sum of the currents leaving the node $= 0$

$$\frac{y(t) - x(t)}{R} + C\frac{d}{dt}y(t) = 0$$

$$Cy'(t) + \frac{1}{R}y(t) = \frac{1}{R}x(t)$$

Example 1



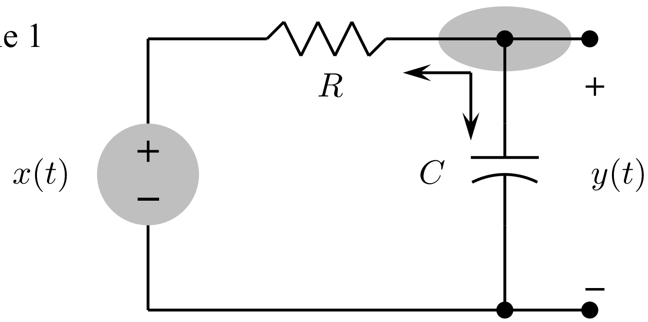
sum of the currents leaving the node $= 0$

$$\frac{y(t) - x(t)}{R} + C\frac{d}{dt}y(t) = 0$$

$$Cy'(t) + \frac{1}{R}y(t) = \frac{1}{R}x(t)$$

$$y'(t) + \frac{1}{RC}y(t) = \frac{1}{RC}x(t)$$

Example 1



sum of the currents leaving the node $= 0$

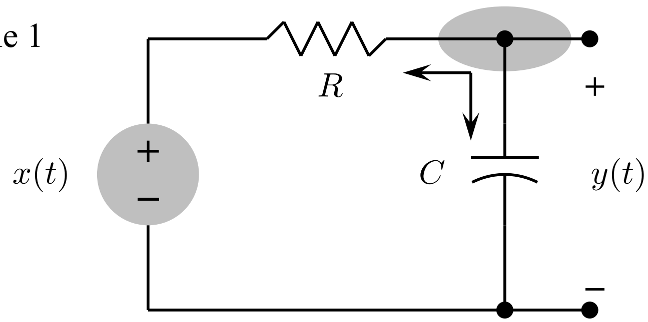
$$\frac{y(t)-x(t)}{R} + C\frac{d}{dt}y(t) = 0$$

$$Cy'(t) + \frac{1}{R}y(t) = \frac{1}{R}x(t)$$

$$y'(t) + \frac{1}{RC}y(t) = \frac{1}{RC}x(t)$$

Linear **C**onstant **C**oefficient **D**ifferential **E**quation
LCCDE

Example 1



sum of the currents leaving the node $= 0$

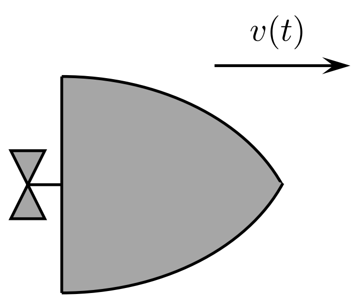
$$\frac{y(t)-x(t)}{R} + C\frac{d}{dt}y(t) = 0$$

$$Cy'(t) + \frac{1}{R}y(t) = \frac{1}{R}x(t)$$

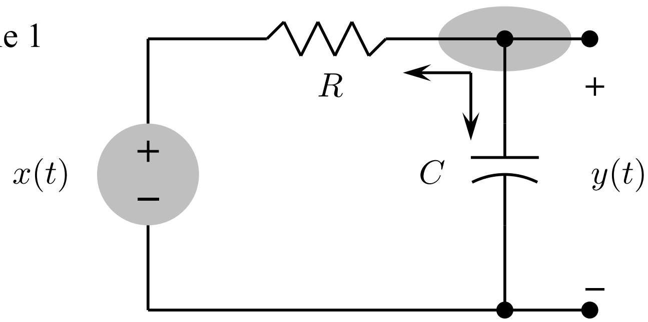
$$y'(t) + \frac{1}{RC}y(t) = \frac{1}{RC}x(t)$$

Linear **C**onstant **C**oefficient **D**ifferential **E**quation
LCCDE

Example 2



Example 1



sum of the currents leaving the node = 0

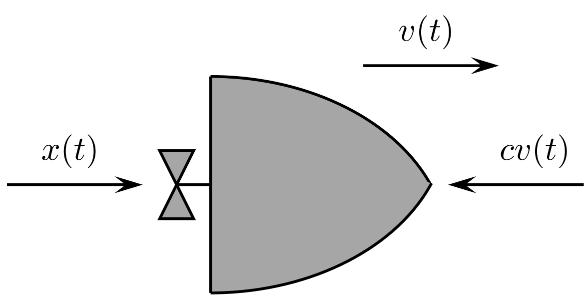
$$\frac{y(t) - x(t)}{R} + C\frac{d}{dt}y(t) = 0$$

$$Cy'(t) + \frac{1}{R}y(t) = \frac{1}{R}x(t)$$

$$y'(t) + \frac{1}{RC}y(t) = \frac{1}{RC}x(t)$$

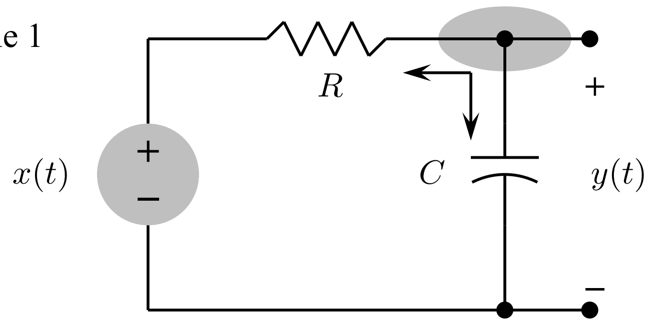
Linear **C**onstant **C**oefficient **D**ifferential **E**quation
LCCDE

Example 2



force = mass × acceleration

Example 1



sum of the currents leaving the node $= 0$

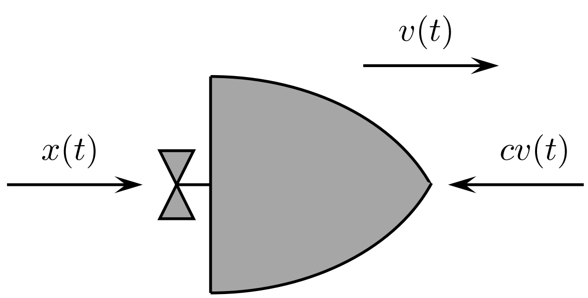
$$\frac{y(t)-x(t)}{R} + C\frac{d}{dt}y(t) = 0$$

$$Cy'(t) + \frac{1}{R}y(t) = \frac{1}{R}x(t)$$

$$y'(t) + \frac{1}{RC}y(t) = \frac{1}{RC}x(t)$$

Linear **C**onstant **C**oefficient **D**ifferential **E**quation
LCCDE

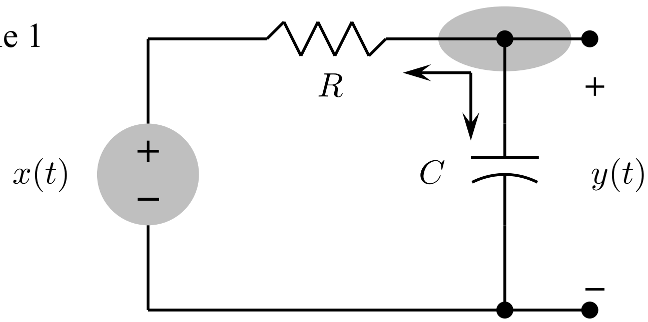
Example 2



force $=$ mass $\times$ acceleration

$$x(t) - cv(t) = m\frac{d}{dt}v(t)$$

Example 1



sum of the currents leaving the node $= 0$

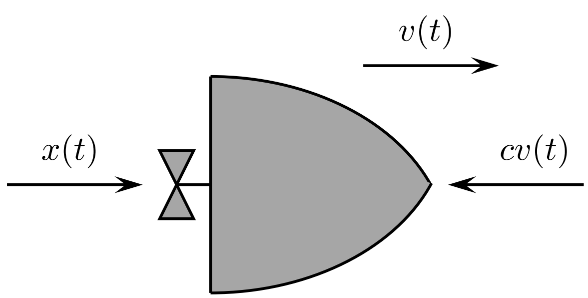
$$\frac{y(t) - x(t)}{R} + C\frac{d}{dt}y(t) = 0$$

$$Cy'(t) + \frac{1}{R}y(t) = \frac{1}{R}x(t)$$

$$y'(t) + \frac{1}{RC}y(t) = \frac{1}{RC}x(t)$$

Linear **C**onstant **C**oefficient **D**ifferential **E**quation
LCCDE

Example 2

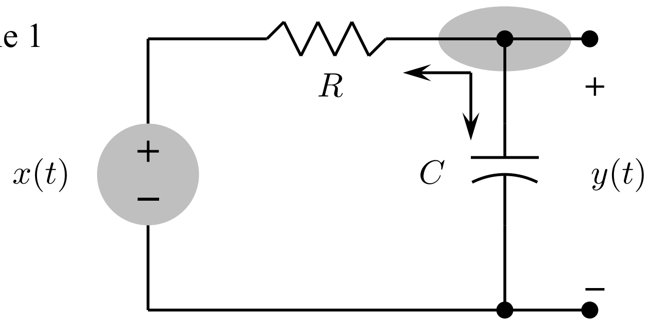


force $=$ mass $\times$ acceleration

$$x(t) - cv(t) = m\frac{d}{dt}v(t)$$

$$mv'(t) + cv(t) = x(t)$$

Example 1



sum of the currents leaving the node $= 0$

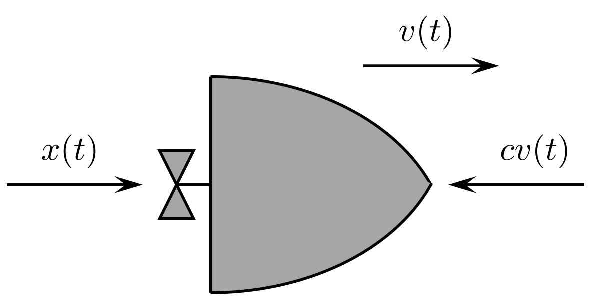
$$\frac{y(t) - x(t)}{R} + C\frac{d}{dt}y(t) = 0$$

$$Cy'(t) + \frac{1}{R}y(t) = \frac{1}{R}x(t)$$

$$y'(t) + \frac{1}{RC}y(t) = \frac{1}{RC}x(t)$$

Linear **C**onstant **C**oefficient **D**ifferential **E**quation
LCCDE

Example 2



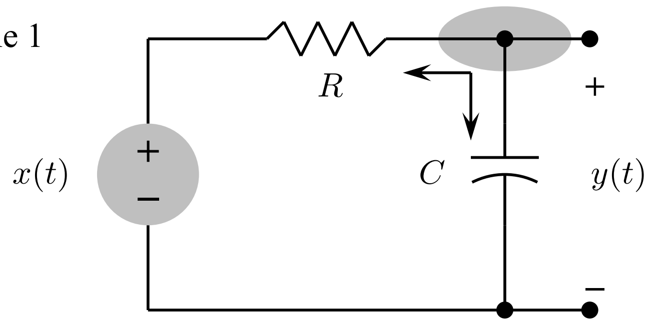
force $=$ mass $\times$ acceleration

$$x(t) - cv(t) = m\frac{d}{dt}v(t)$$

$$mv'(t) + cv(t) = x(t)$$

$$v'(t) + \frac{c}{m}v(t) = \frac{1}{m}x(t)$$

Example 1



sum of the currents leaving the node $= 0$

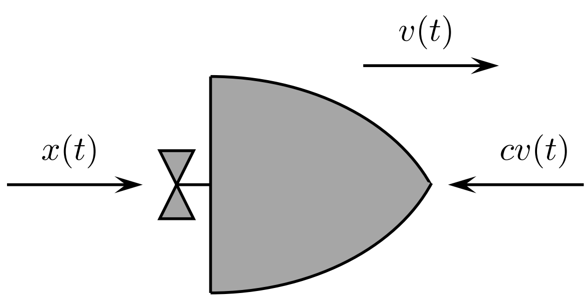
$$\frac{y(t) - x(t)}{R} + C\frac{d}{dt}y(t) = 0$$

$$Cy'(t) + \frac{1}{R}y(t) = \frac{1}{R}x(t)$$

$$y'(t) + \frac{1}{RC}y(t) = \frac{1}{RC}x(t)$$

Linear **C**onstant **C**oefficient **D**ifferential **E**quation
LCCDE

Example 2



force $=$ mass $\times$ acceleration

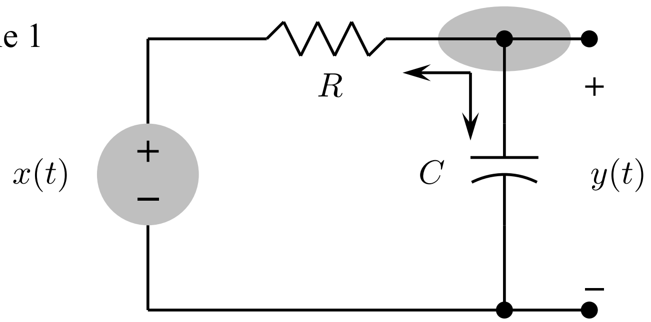
$$x(t) - cv(t) = m\frac{d}{dt}v(t)$$

$$mv'(t) + cv(t) = x(t)$$

$$v'(t) + \frac{c}{m}v(t) = \frac{1}{m}x(t)$$

$$y'(t) + \frac{c}{m}y(t) = \frac{1}{m}x(t)$$

Example 1



sum of the currents leaving the node $= 0$

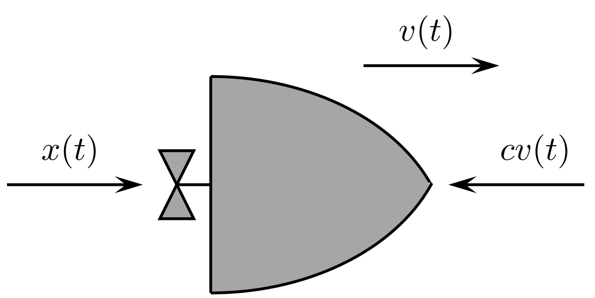
$$\frac{y(t)-x(t)}{R} + C\frac{d}{dt}y(t) = 0$$

$$Cy'(t) + \frac{1}{R}y(t) = \frac{1}{R}x(t)$$

$$y'(t) + \frac{1}{RC}y(t) = \frac{1}{RC}x(t)$$

Linear **C**onstant **C**oefficient **D**ifferential **E**quation
LCCDE

Example 2



force $=$ mass $\times$ acceleration

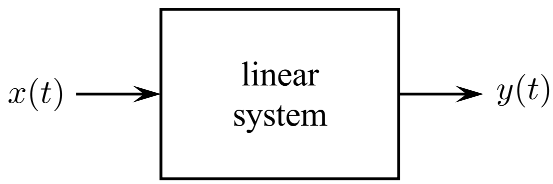
$$x(t) - cv(t) = m\frac{d}{dt}v(t)$$

$$mv'(t) + cv(t) = x(t)$$

$$v'(t) + \frac{c}{m}v(t) = \frac{1}{m}x(t)$$

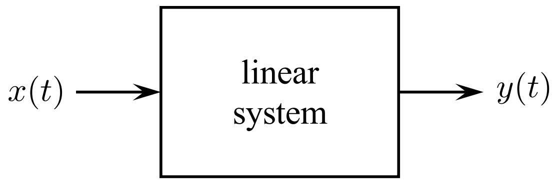
$$y'(t) + \frac{c}{m}y(t) = \frac{1}{m}x(t)$$

Linear **C**onstant **C**oefficient **D**ifferential **E**quation
LCCDE



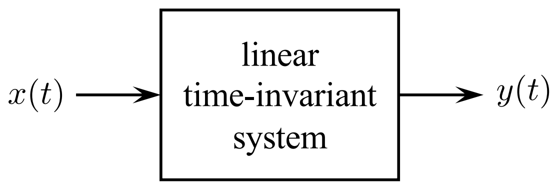
The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

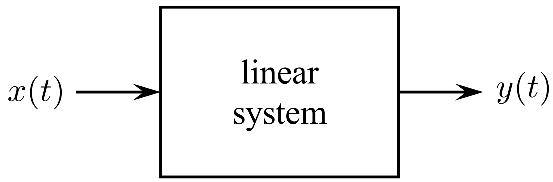


The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

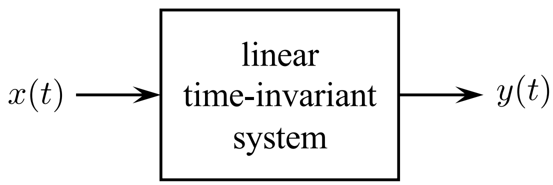


Systems defined by an LCCDE are also time-invariant

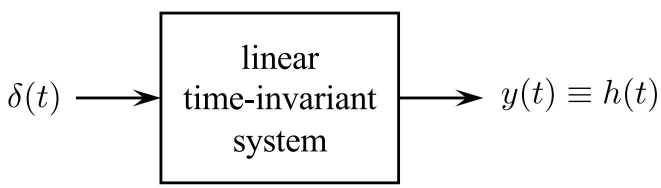


The linear systems of interest to us have an input/output relationship defined by LCCDE

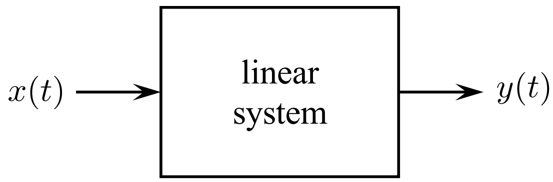
e.g., $y'(t) + ay(t) = x(t)$



Systems defined by an LCCDE are also time-invariant

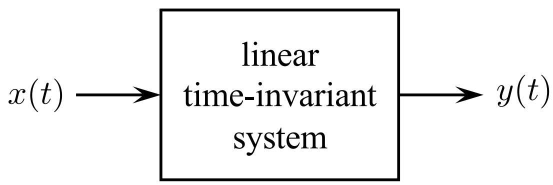


The input/output relationship of a linear time-invariant system may also be described using the impulse response and convolution.

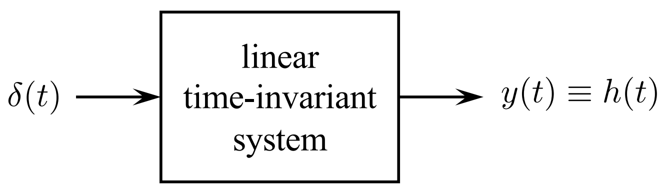


The linear systems of interest to us have an input/output relationship defined by LCCDE

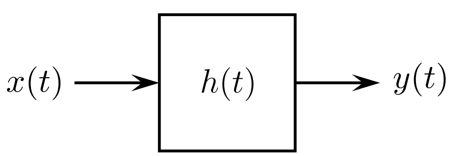
e.g., $y'(t) + ay(t) = x(t)$



Systems defined by an LCCDE are also time-invariant

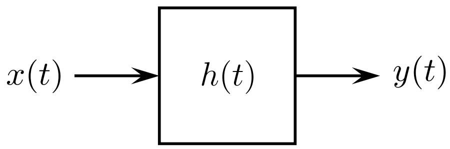


The input/output relationship of a linear time-invariant system may also be described using the impulse response and convolution.



$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

Convolution Example



$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

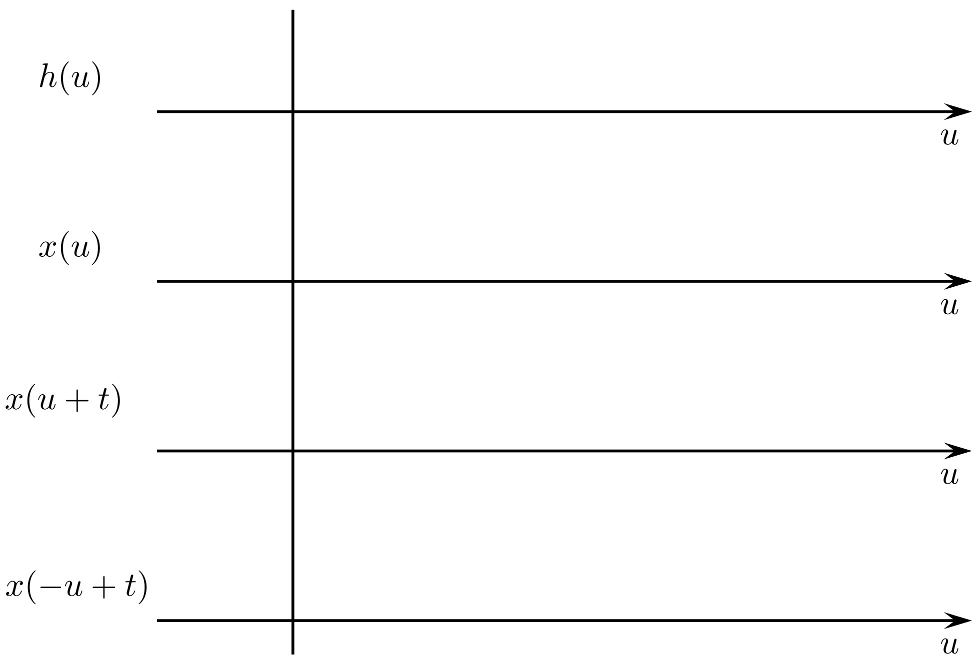
$$x(t) = U(t) \qquad h(t) = \frac{1}{a}e^{-at}U(t)$$

Convolution Example

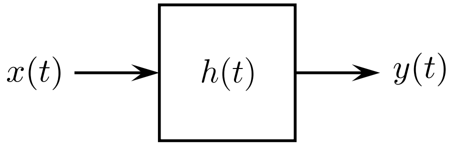
$$x(t) \longrightarrow \boxed{h(t)} \longrightarrow y(t)$$

$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

$$x(t) = U(t) \qquad h(t) = \frac{1}{a}e^{-at}U(t)$$

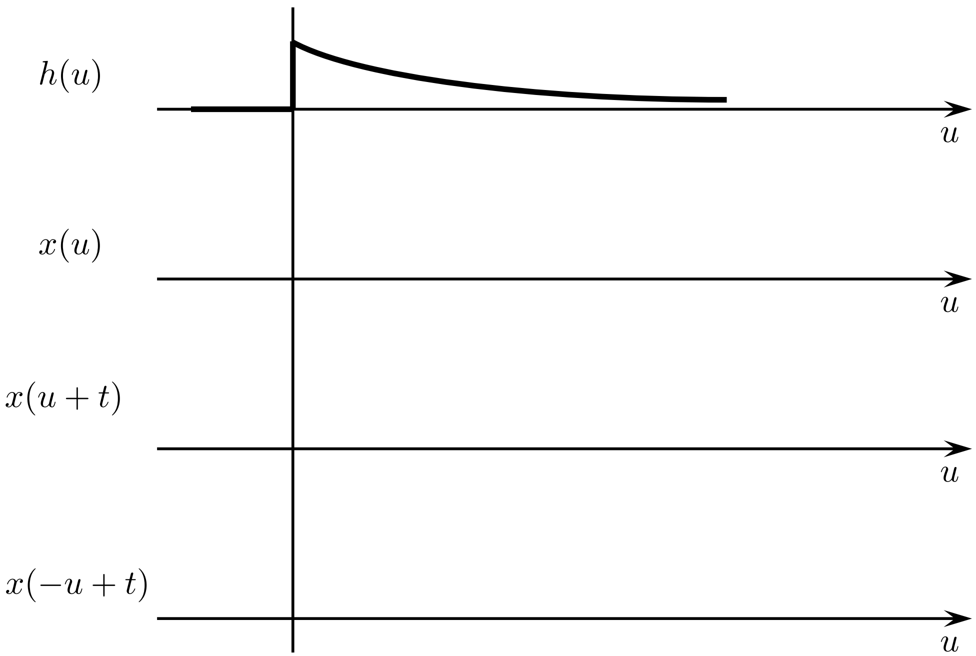


Convolution Example



$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

$x(t) = U(t) \qquad h(t) = \frac{1}{a}e^{-at}U(t)$

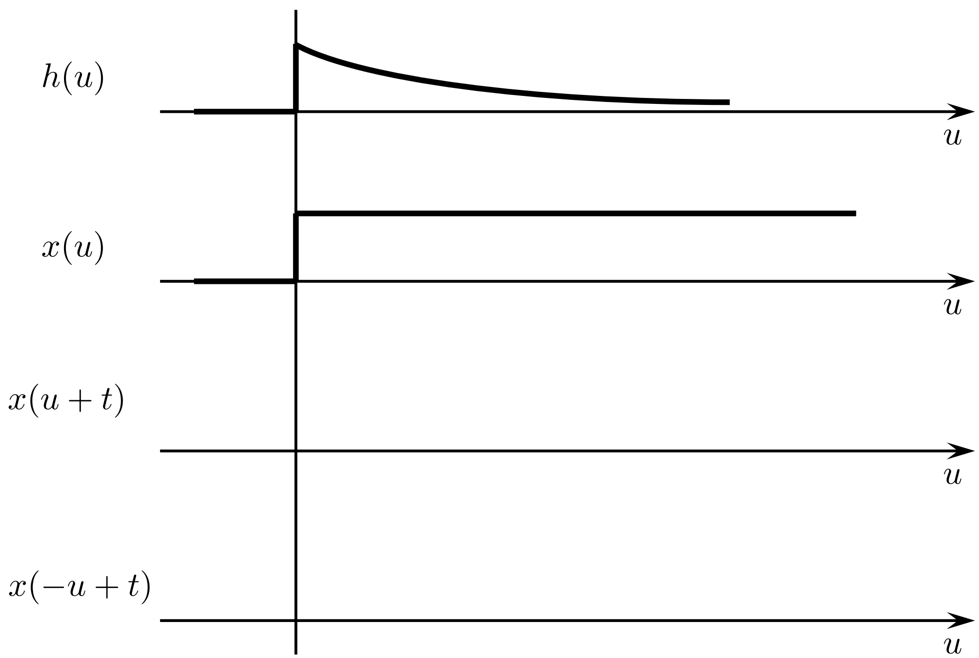


Convolution Example

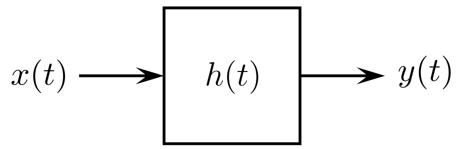
$$x(t) \longrightarrow \boxed{h(t)} \longrightarrow y(t)$$

$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

$$x(t) = U(t) \qquad h(t) = \frac{1}{a}e^{-at}U(t)$$

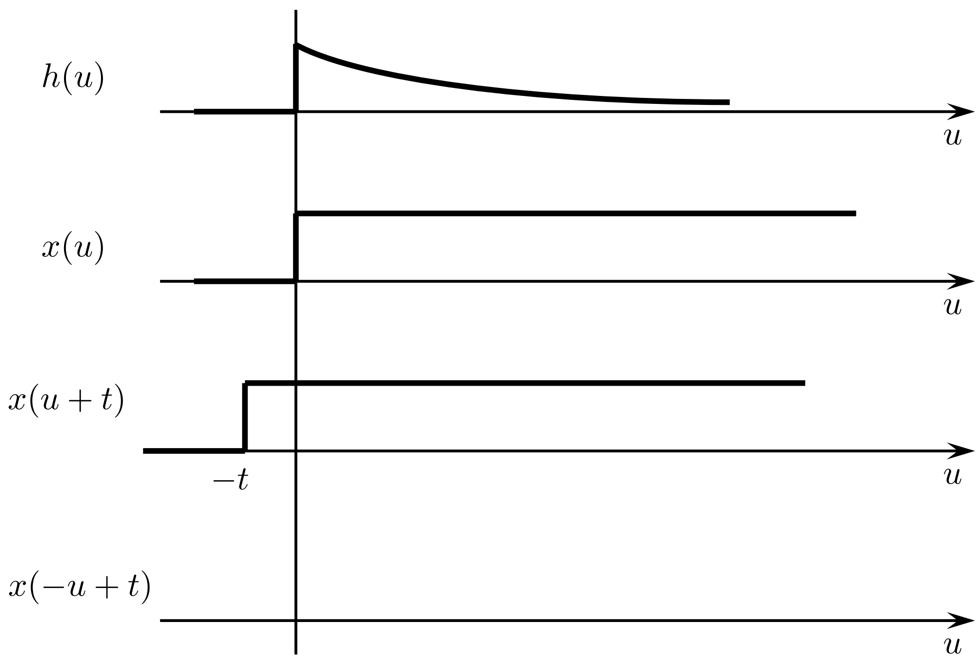


Convolution Example

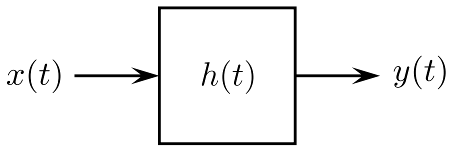


$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

$x(t) = U(t) \qquad h(t) = \frac{1}{a}e^{-at}U(t)$

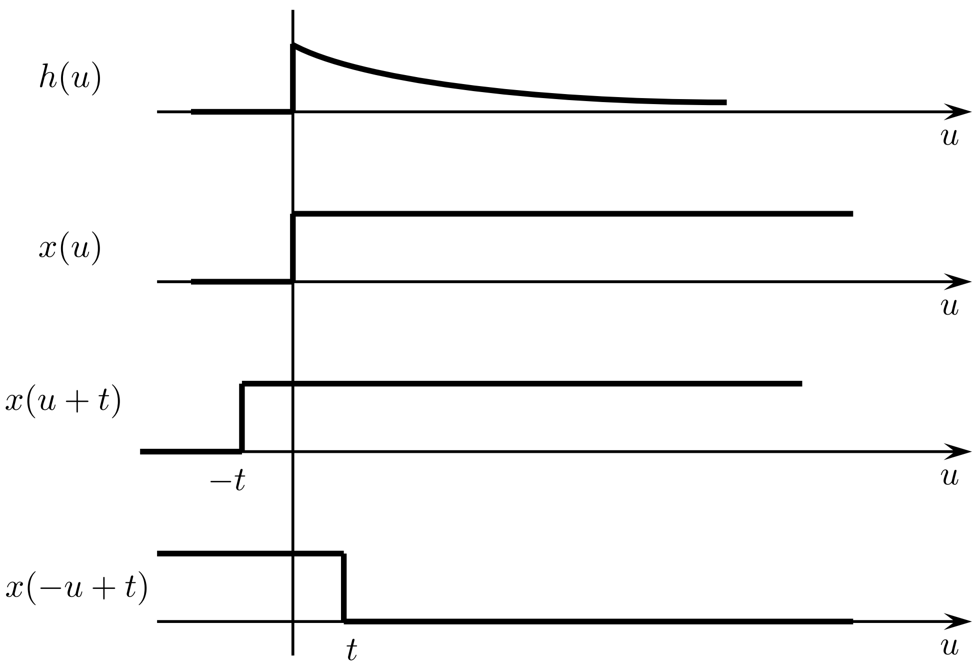


Convolution Example

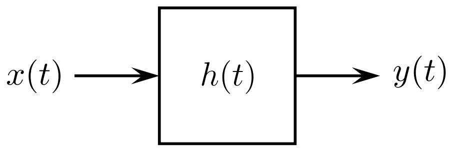


$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

$$x(t) = U(t) \qquad h(t) = \frac{1}{a}e^{-at}U(t)$$

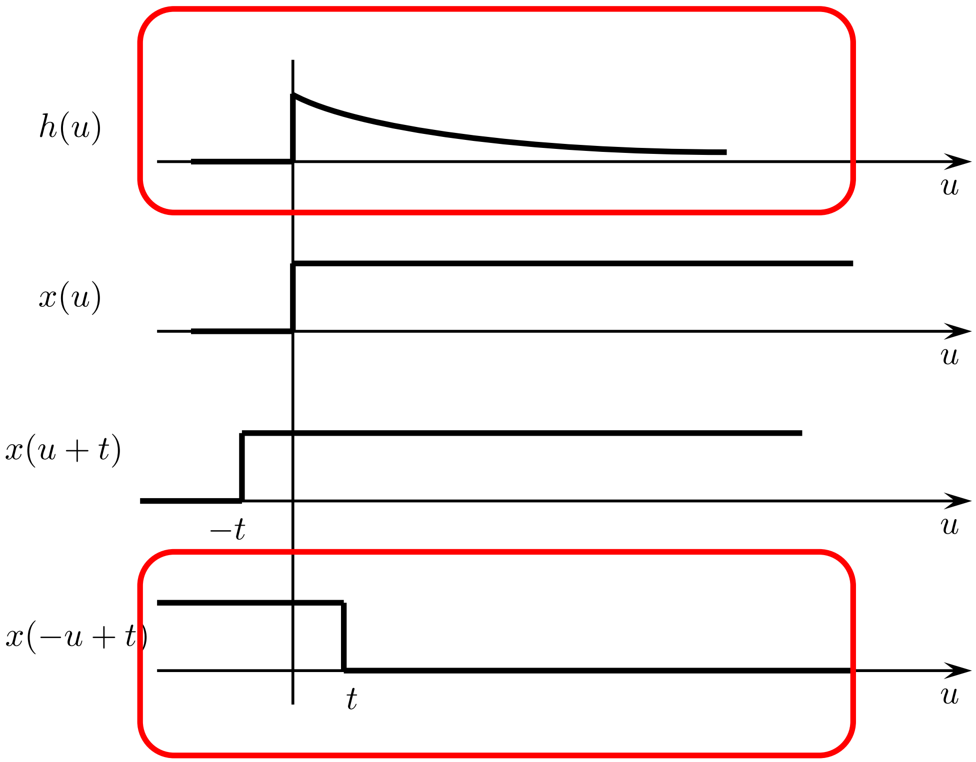


Convolution Example



$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

$$x(t) = U(t) \qquad h(t) = \frac{1}{a}e^{-at}U(t)$$

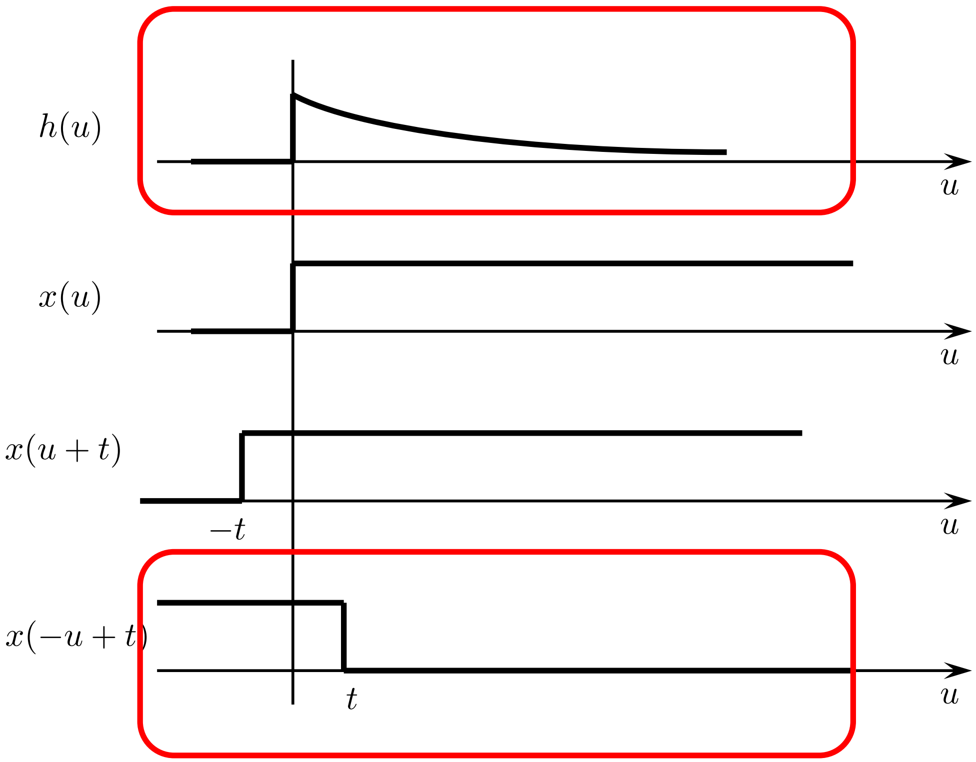


Convolution Example

$$x(t) \longrightarrow \boxed{h(t)} \longrightarrow y(t)$$

$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

$$x(t) = U(t) \qquad h(t) = \frac{1}{a}e^{-at}U(t)$$



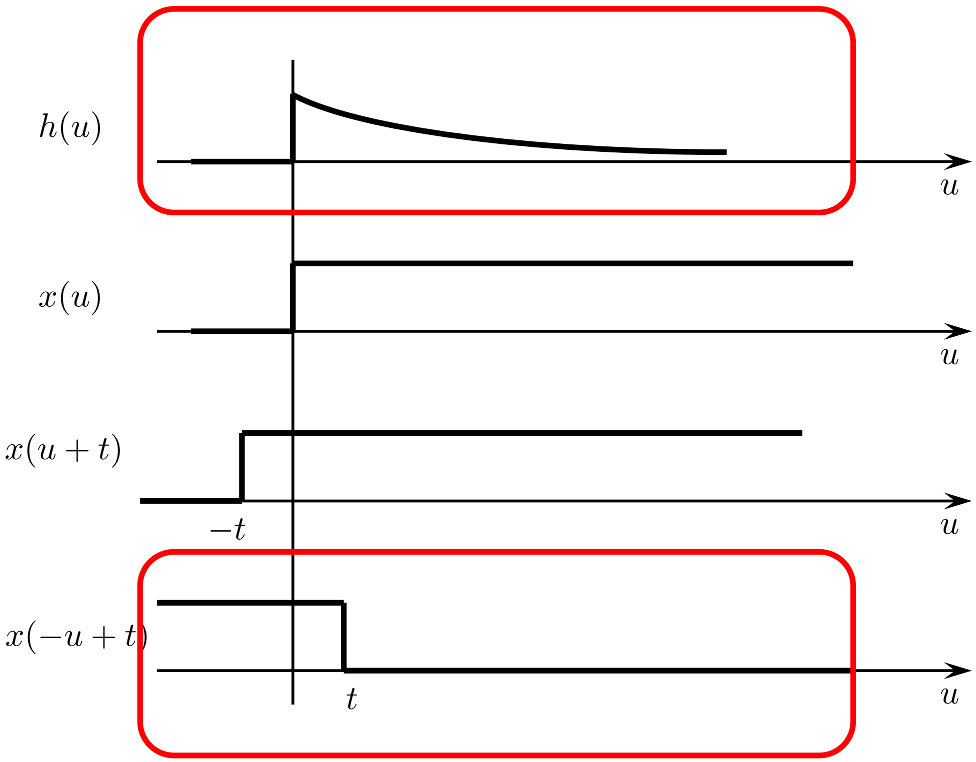
$$t < 0, \qquad y(t) = 0$$

Convolution Example

$$x(t) \longrightarrow \boxed{h(t)} \longrightarrow y(t)$$

$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

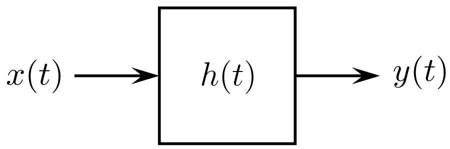
$$x(t) = U(t) \qquad h(t) = \frac{1}{a}e^{-at}U(t)$$



$$t < 0, \qquad y(t) = 0$$

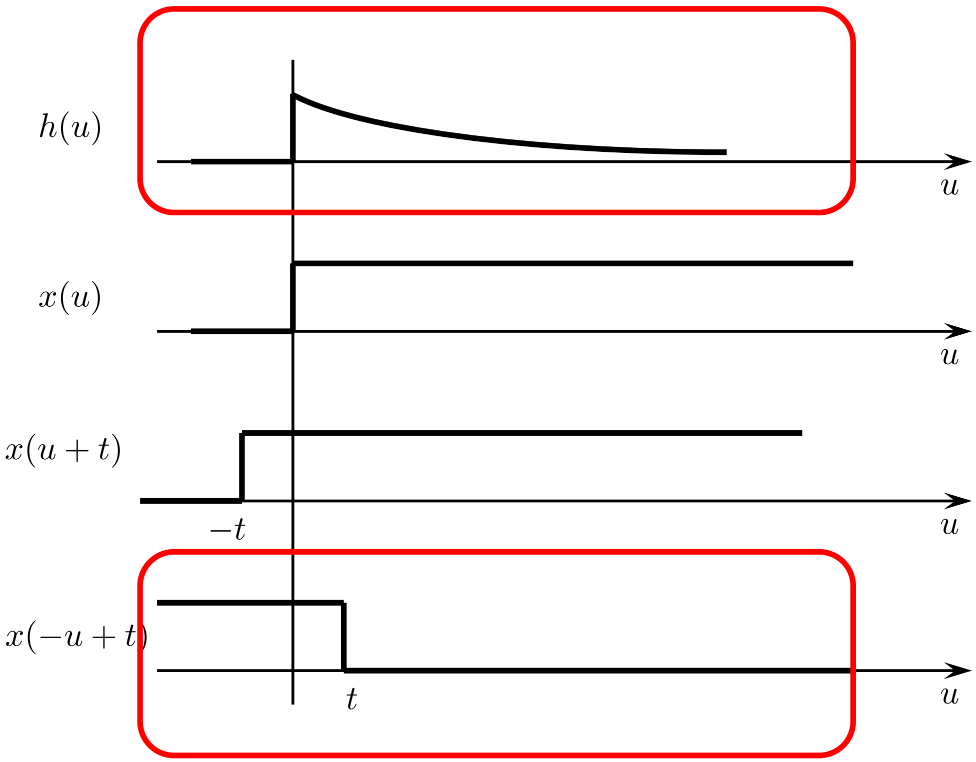
$$t \geq 0, \qquad y(t) = \int_0^t \frac{1}{a}e^{-au}du = 1 - e^{-at}$$

Convolution Example



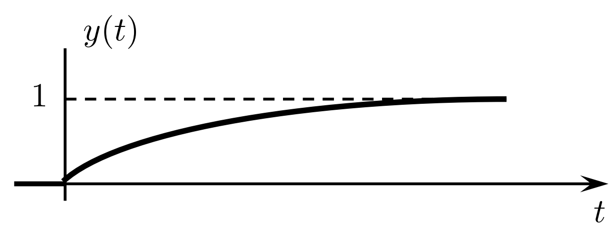
$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

$$x(t) = U(t) \qquad h(t) = \frac{1}{a}e^{-at}U(t)$$



$$t < 0, \qquad y(t) = 0$$

$$t \geq 0, \qquad y(t) = \int_0^t \frac{1}{a}e^{-au}du = 1 - e^{-at}$$



Frequency Domain Analysis

Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t) e^{-st} dt$$

$$x(t) = \frac{1}{j2\pi} \oint X(s) e^{st} ds$$

Frequency Domain Analysis

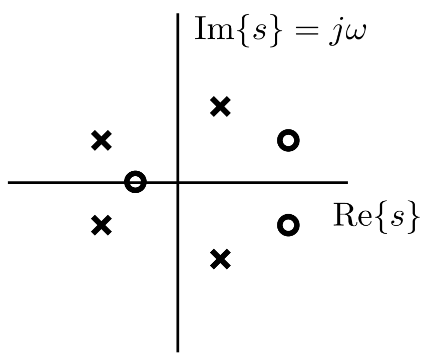
Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

$$x(t) = \frac{1}{j2\pi} \oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.



Frequency Domain Analysis

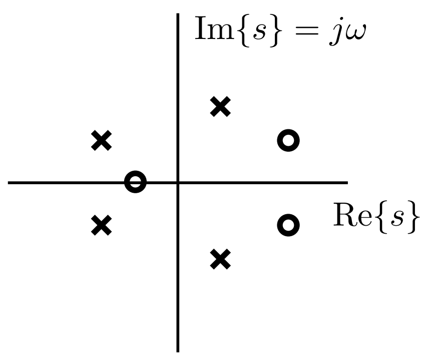
Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

$$x(t) = \frac{1}{j2\pi}\oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:
$$X(s) = \frac{B(s)}{A(s)}$$



Frequency Domain Analysis

Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

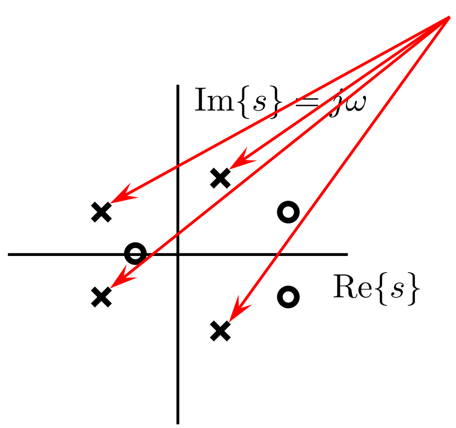
$$x(t) = \frac{1}{j2\pi}\oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:

$$X(s) = \frac{B(s)}{A(s)}$$

poles: values of s where $X(s) \to \pm\infty$
roots of $A(s)$



Frequency Domain Analysis

Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

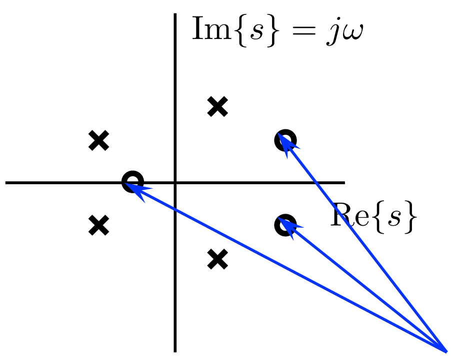
$$x(t) = \frac{1}{j2\pi} \oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:

$$X(s) = \frac{B(s)}{A(s)}$$

poles: values of s where $X(s) \to \pm\infty$
roots of $A(s)$



zeros: values of s where $X(s) = 0$
roots of $B(s)$

Frequency Domain Analysis

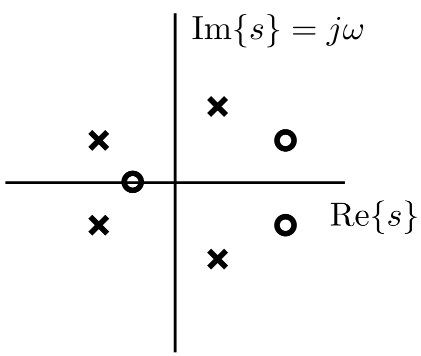
Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

$$x(t) = \frac{1}{j2\pi} \oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:
$$X(s) = \frac{B(s)}{A(s)}$$



The Laplace transform integral

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

only converges for certain values of s. These values are called *the region of convergence* (ROC).

- The ROC cannot contain any poles.
- The ROC for the Laplace transform of a stable (bounded) time-domain signal contains the $\mathrm{Im}\{s\} = j\omega$ axis.

Frequency Domain Analysis

Laplace Transform

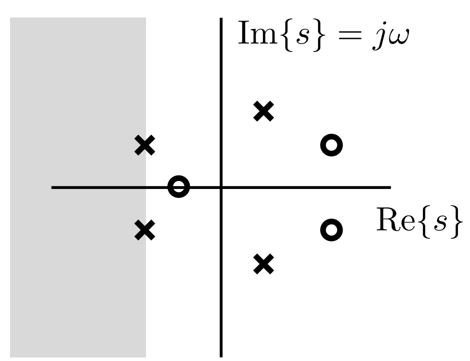
$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

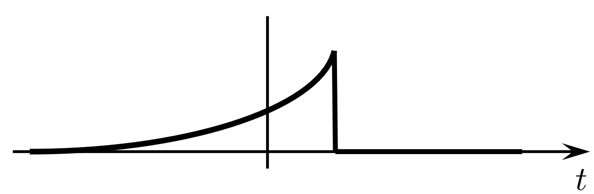
$$x(t) = \frac{1}{j2\pi}\oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:
$$X(s) = \frac{B(s)}{A(s)}$$

ROC for left-sided signals





Frequency Domain Analysis

Laplace Transform

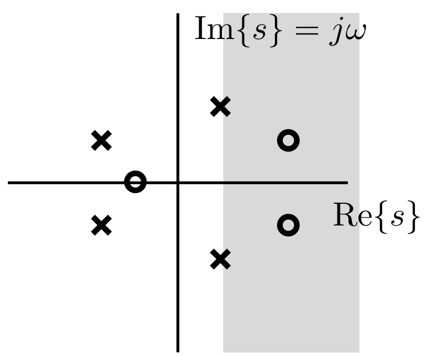
$$X(s) = \int_{-\infty}^{\infty} x(t) e^{-st} dt$$

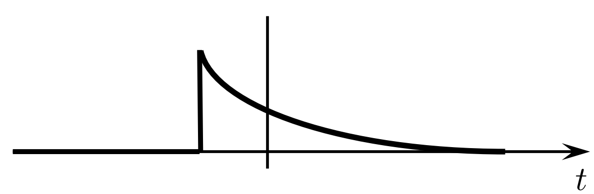
$$x(t) = \frac{1}{j2\pi} \oint X(s) e^{st} ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:
$$X(s) = \frac{B(s)}{A(s)}$$

ROC for right-sided signals





Frequency Domain Analysis

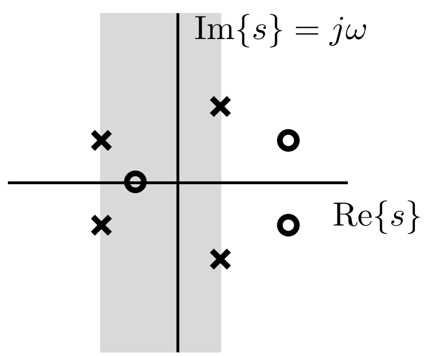
Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

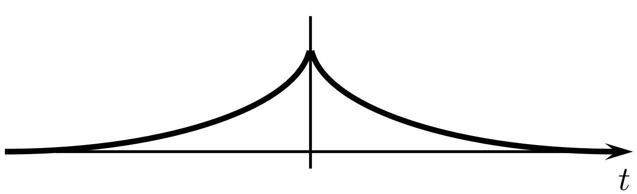
$$x(t) = \frac{1}{j2\pi} \oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:
$$X(s) = \frac{B(s)}{A(s)}$$



ROC for two-sided signals



Frequency Domain Analysis

Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

$$x(t) = \frac{1}{j2\pi} \oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:
$$X(s) = \frac{B(s)}{A(s)}$$
- Because s is a complex variable, s as a *variable of integration* defines a *contour integral* in the complex plane.

Frequency Domain Analysis

Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

$$x(t) = \frac{1}{j2\pi}\oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:
$$X(s) = \frac{B(s)}{A(s)}$$
- Because s is a complex variable, s as a *variable of integration* defines a *contour integral* in the complex plane.
  - Contour integral wizards compute the inverse Laplace transform from the definition.
  - Everyone else uses partial fraction expansion to construct tables, then uses the tables.

Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

$$x(t) = \frac{1}{j2\pi} \oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:
$$X(s) = \frac{B(s)}{A(s)}$$
- Because s is a complex variable, s as a *variable of integration* defines a *contour integral* in the complex plane.
  - Contour integral wizards compute the inverse Laplace transform from the definition.
  - Everyone else uses partial fraction expansion to construct tables, then uses the tables.

$x(t)$	$X(s)$	ROC
$\delta(t)$	1	all s
$U(t)$	$\frac{1}{s}$	Real $\{s\} > 0$
$\frac{t^{n-1}}{(n-1)!}U(t)$	$\frac{1}{s^n}$	Real $\{s\} > 0$
$e^{-at}U(t)$	$\frac{1}{s+a}$	Real $\{s\} > -a$
$\frac{t^{n-1}}{(n-1)!}e^{-at}U(t)$	$\frac{1}{(s+a)^n}$	Real $\{s\} > 0$
$\delta(t-T)$	e^{-sT}	all s
$\cos(\omega_0 t)\,U(t)$	$\frac{s}{s^2+\omega_0^2}$	Real $\{s\} > 0$
$\sin(\omega_0 t)\,U(t)$	$\frac{\omega_0}{s^2+\omega_0^2}$	Real $\{s\} > 0$
$e^{-at}\cos(\omega_0 t)\,U(t)$	$\frac{s+a}{(s+a)^2+\omega_0^2}$	Real $\{s\} > -a$
$e^{-at}\sin(\omega_0 t)\,U(t)$	$\frac{\omega_0}{(s+a)^2+\omega_0^2}$	Real $\{s\} > -a$
$\frac{d^n\delta(t)}{dt}$	s^n	all s
$\underbrace{U(t) * \cdots * U(t)}_{n \text{ times}}$	$\frac{1}{s^n}$	Real $\{s\} > 0$

Frequency Domain Analysis

Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

$$x(t) = \frac{1}{j2\pi}\oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and
  valued function of the comp
- LCCDEs always produce a
  in s:

$$X(s) = \frac{B}{A}$$

- Because s is a complex vari
  of integration defines a *co*
  complex plane.
  - Contour integral wiza
    verse Laplace transforn
  - Everyone else uses pa
    sion to construct table
    bles.

Property	Signal	Laplace Transform	ROC
	$x(t)$	$X(s)$	R_x
	$y(t)$	$Y(s)$	R_y
Linearity	$ax(t) + by(t)$	$aX(s) + bY(s)$	at least $R_x \cap R_y$
Time Shifting	$x(t - t_0)$	$e^{-st_0}X(s)$	R_x
Time Scaling	$x(at)$	$\frac{1}{\lvert a\rvert}X\left(\frac{s}{a}\right)$	all s for which s/a is in R_x
Conjugation	$x^*(t)$	$X^*(s^*)$	R_x
Convolution	$x(t) * y(t)$	$X(s)Y(s)$	at least $R_x \cap R_y$
Differentiation	$\frac{d}{dt}x(t)$	$sX(s) - x(0^-)$	at least R_x
Integration	$\int_{-\infty}^{t} x(\tau)d\tau$	$\frac{1}{s}X(s)$	at least $R_x \cap \{\text{Real}\{s\} > 0\}$
Initial Value Theorem[a]		$\lim_{t\to 0^+} x(t) = \lim_{s\to\infty} sX(s)$	
Final Value Theorem[b]		$\lim_{t\to\infty} x(t) = \lim_{s\to 0} sX(s)$	

[a] The Initial Value Theorem is valid for signals $x(t)$ that satisfy the following conditions: $x(t) = 0$ for $t < 0$ and $x(t)$ contains no impulses or higher-order singularities at $t = 0$.

[b] The Final Value Theorem is valid for signals $x(t)$ that satisfy the following conditions: $x(t) = 0$ for $t < 0$ and $x(t)$ has a finite limit as $t \to \infty$.

Frequency Domain Analysis

Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

$$x(t) = \frac{1}{j2\pi}\oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:
$$X(s) = \frac{B(s)}{A(s)}$$
- Because s is a complex variable, s as a *variable of integration* defines a *contour integral* in the complex plane.
  - Contour integral wizards compute the inverse Laplace transform from the definition.
  - Everyone else uses partial fraction expansion to construct tables, then uses the tables.

Fourier Transform

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

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

Frequency Domain Analysis

Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

$$x(t) = \frac{1}{j2\pi} \oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:
$$X(s) = \frac{B(s)}{A(s)}$$
- Because s is a complex variable, s as a *variable of integration* defines a *contour integral* in the complex plane.
  - Contour integral wizards compute the inverse Laplace transform from the definition.
  - Everyone else uses partial fraction expansion to construct tables, then uses the tables.

Fourier Transform

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

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

Comments:

- w is a real variable and $X(\omega)$ is a complex-valued function of the real variable ω.

Frequency Domain Analysis

Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

$$x(t) = \frac{1}{j2\pi} \oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:
$$X(s) = \frac{B(s)}{A(s)}$$
- Because s is a complex variable, s as a *variable of integration* defines a *contour integral* in the complex plane.
  - Contour integral wizards compute the inverse Laplace transform from the definition.
  - Everyone else uses partial fraction expansion to construct tables, then uses the tables.

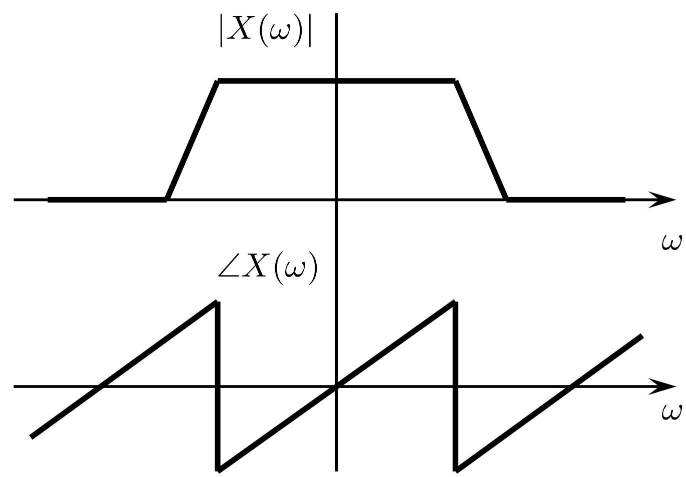
Fourier Transform

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

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

Comments:

- w is a real variable and $X(\omega)$ is a complex-valued function of the real variable ω.



Frequency Domain Analysis

Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

$$x(t) = \frac{1}{j2\pi}\oint X(s)e^{st}ds$$

Comments:

- s is a complex variable and $X(s)$ is a complex-valued function of the complex variable s.
- LCCDEs always produce a ratio of polynomials in s:
$$X(s) = \frac{B(s)}{A(s)}$$
- Because s is a complex variable, s as a *variable of integration* defines a *contour integral* in the complex plane.
  - Contour integral wizards compute the inverse Laplace transform from the definition.
  - Everyone else uses partial fraction expansion to construct tables, then uses the tables.

Fourier Transform

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

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

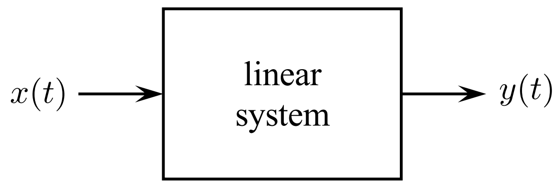
Comments:

- w is a real variable and $X(\omega)$ is a complex-valued function of the real variable ω.
- Because ω is a real variable, the integral that defines the inverse transform is the familiar integral introduced in your first and second calculus courses.
  - Integral wizards compute the inverse Fourier transform from the definition.
  - Everyone else uses uses tables.

$x(t)$	$X(\omega)$	$X(f)$
$\delta(t)$	1	1
$\delta(t-t_0)$	$e^{-j\omega t_0}$	$e^{-j2\pi f t_0}$
$u(t)$	$\frac{1}{j\omega}+\pi\delta(\omega)$	$\frac{1}{j2\pi f}+\frac{1}{2}\delta(f)$
$\frac{t^{n-1}}{(n-1)!}e^{-at}u(t)$	$\frac{1}{(a+j\omega)^n}$	$\frac{1}{(a+j2\pi f)^n}$
$\sum_{k=-\infty}^{\infty} a_k e^{jk\omega_0 t}$ $(\omega_0=2\pi f_0)$	$2\pi\sum_{k=-\infty}^{\infty} a_k\delta(\omega-k\omega_0)$	$\sum_{k=-\infty}^{\infty} a_k\delta(f-kf_0)$
$\sum_{k=-\infty}^{\infty}\delta(t-kT)$	$\frac{2\pi}{T}\sum_{n=-\infty}^{\infty}\delta\left(\omega-\frac{2\pi n}{T}\right)$	$\frac{1}{T}\sum_{n=-\infty}^{\infty}\delta\left(f-\frac{n}{T}\right)$
1	$2\pi\delta(\omega)$	$\delta(f)$
$\begin{cases} 1 & \lvert t\rvert<T \\ 0 & \lvert t\rvert>T \end{cases}$	$2T\frac{\sin(\omega T)}{\omega T}$	$2T\frac{\sin(2\pi f T)}{2\pi f T}$
$2B\frac{\sin(2\pi B t)}{2\pi B t}$	$\begin{cases} 1 & -2\pi B\le\omega\le 2\pi B \\ 0 & \text{otherwise} \end{cases}$	$\begin{cases} 1 & -B\le f\le B \\ 0 & \text{otherwise} \end{cases}$
$e^{j\omega_0 t}$ $(\omega_0=2\pi f_0)$	$2\pi\delta(\omega-\omega_0)$	$\delta(f-f_0)$
$\cos(\omega_0 t)$ $(\omega_0=2\pi f_0)$	$\pi\delta(\omega-\omega_0)+\pi\delta(\omega+\omega_0)$	$\frac{1}{2}\delta(f-f_0)+\frac{1}{2}\delta(f+f_0)$
$\sin(\omega_0 t)$ $(\omega_0=2\pi f_0)$	$\frac{\pi}{j}\delta(\omega-\omega_0)-\frac{\pi}{j}\delta(\omega+\omega_0)$	$\frac{1}{j2}\delta(f-f_0)-\frac{1}{j2}\delta(f+f_0)$
$\exp\{-a\lvert t\rvert\}$	$\frac{2a}{a^2+\omega^2}$	$\frac{2a}{a^2+(2\pi f)^2}$
$\exp\{-\pi t^2\}$	$\exp\left\{-\pi\left(\frac{\omega}{2\pi}\right)^2\right\}$	$\exp\left\{-\pi f^2\right\}$

Property	Signal	Fourier Transform in ω	Fourier Transform in f
	$x(t)$	$X(\omega)$	$X(f)$
	$y(t)$	$Y(\omega)$	$Y(f)$
Linearity	$ax(t) + by(t)$	$aX(\omega) + bY(\omega)$	$aX(f) + bY(f)$
Time Shifting	$x(t - t_0)$	$e^{-j\omega t_0} X(\omega)$	$e^{-j2\pi f t_0} X(f)$
Time Scaling	$x(at)$	$\frac{1}{\lvert a\rvert} X\left(\frac{\omega}{a}\right)$	$\frac{1}{\lvert a\rvert} X\left(\frac{f}{a}\right)$
Conjugation	$x^*(t)$	$X^*(-\omega)$	$X^*(-f)$
Convolution	$x(t) * y(t)$	$X(\omega)Y(\omega)$	$X(f)Y(f)$
Differentiation	$\frac{d}{dt}x(t)$	$j\omega X(\omega)$	$j2\pi f X(f)$
Integration	$\int_{-\infty}^{t} x(\tau)d\tau$	$\frac{1}{j\omega}X(\omega) + \pi X(0)\delta(\omega)$	$\frac{1}{j2\pi f}X(f) + \frac{1}{2}X(0)\delta(f)$
Frequency Shifting ($\omega_0 = 2\pi f_0$)	$e^{j\omega_0 t}x(t)$	$X(\omega - \omega_0)$	$X(f - f_0)$
Multiplication	$x(t)y(t)$	$\frac{1}{2\pi}X(\omega) * Y(\omega)$	$X(f) * Y(f)$
Parseval's Theorem	$\int_{-\infty}^{\infty} \lvert x(t)\rvert^2 dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \lvert X(\omega)\rvert^2 d\omega = \int_{-\infty}^{\infty} \lvert X(f)\rvert^2 df$		

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$y'(t) + ay(t) = x(t)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

$$x(t) \longrightarrow \boxed{\begin{array}{c}\text{linear}\\ \text{system}\end{array}} \longrightarrow y(t)$$

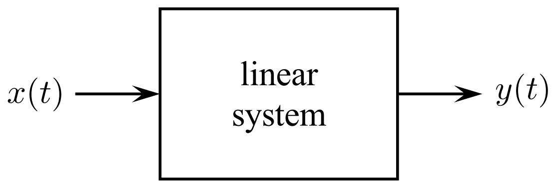
The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$y'(t) + ay(t) = x(t)$$

$$sY(s) - y(0^-) + aY(s) = X(s)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



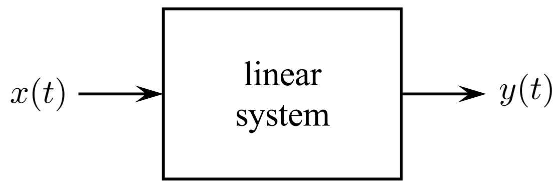
The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$y'(t) + ay(t) = x(t)$$

$$sY(s) - \cancelto{0}{y(0^-)} + aY(s) = X(s)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

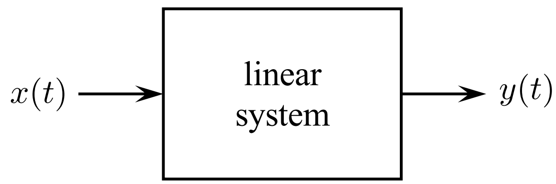
e.g., $y'(t) + ay(t) = x(t)$

$$y'(t) + ay(t) = x(t)$$

$$sY(s) - \cancelto{0}{y(0^-)} + aY(s) = X(s)$$

$$sY(s) + aY(s) = X(s)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

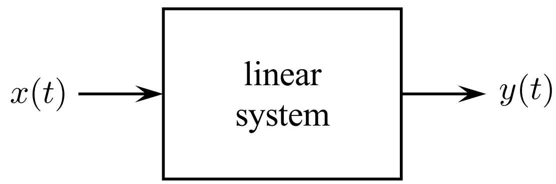
$$y'(t) + ay(t) = x(t)$$

$$sY(s) - \cancelto{0}{y(0^-)} + aY(s) = X(s)$$

$$sY(s) + aY(s) = X(s)$$

$$(s+a)Y(s) = X(s)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$y'(t) + ay(t) = x(t)$$

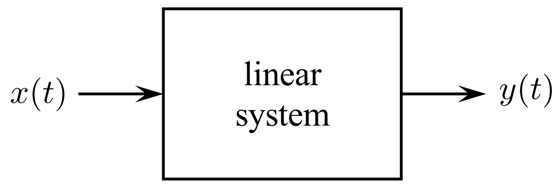
$$sY(s) - \cancelto{0}{y(0^-)} + aY(s) = X(s)$$

$$sY(s) + aY(s) = X(s)$$

$$(s + a)Y(s) = X(s)$$

$$Y(s) = \frac{1}{s+a} X(s)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$y'(t) + ay(t) = x(t)$$

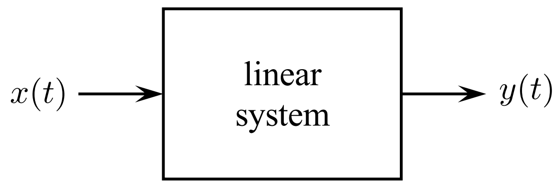
$$sY(s) - \cancelto{0}{y(0^-)} + aY(s) = X(s)$$

$$sY(s) + aY(s) = X(s)$$

$$(s+a)Y(s) = X(s)$$

$$Y(s) = \underbrace{\frac{1}{s+a}}_{\text{Transfer function } H(s)} X(s)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$y'(t) + ay(t) = x(t)$$

$$sY(s) - \cancelto{0}{y(0^-)} + aY(s) = X(s)$$

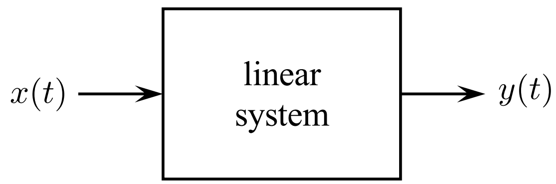
$$sY(s) + aY(s) = X(s)$$

$$(s+a)Y(s) = X(s)$$

$$Y(s) = \underbrace{\frac{1}{s+a}}_{\text{Transfer function } H(s)} X(s)$$

$$Y(s) = H(s)X(s) \quad \Rightarrow \quad y(t) = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$y'(t) + ay(t) = x(t)$$

$$sY(s) - \cancelto{0}{y(0^-)} + aY(s) = X(s)$$

$$sY(s) + aY(s) = X(s)$$

$$(s+a)Y(s) = X(s)$$

$$Y(s) = \frac{1}{s+a}X(s)$$

Transfer function $H(s)$

$$Y(s) = \frac{1}{s+a}X(s) \quad \Rightarrow \quad y(t) = \int_0^\infty e^{-au}x(t-u)du$$

$$Y(s) = H(s)X(s) \quad \Rightarrow \quad y(t) = \int_{-\infty}^\infty h(u)x(t-u)du$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

$$x(t) \longrightarrow \boxed{\text{linear system}} \longrightarrow y(t)$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

$$\text{e.g., } y'(t) + ay(t) = x(t)$$

$$y'(t) + ay(t) = x(t)$$

$$sY(s) - \cancelto{0}{y(0^-)} + aY(s) = X(s)$$

$$sY(s) + aY(s) = X(s)$$

$$(s+a)Y(s) = X(s)$$

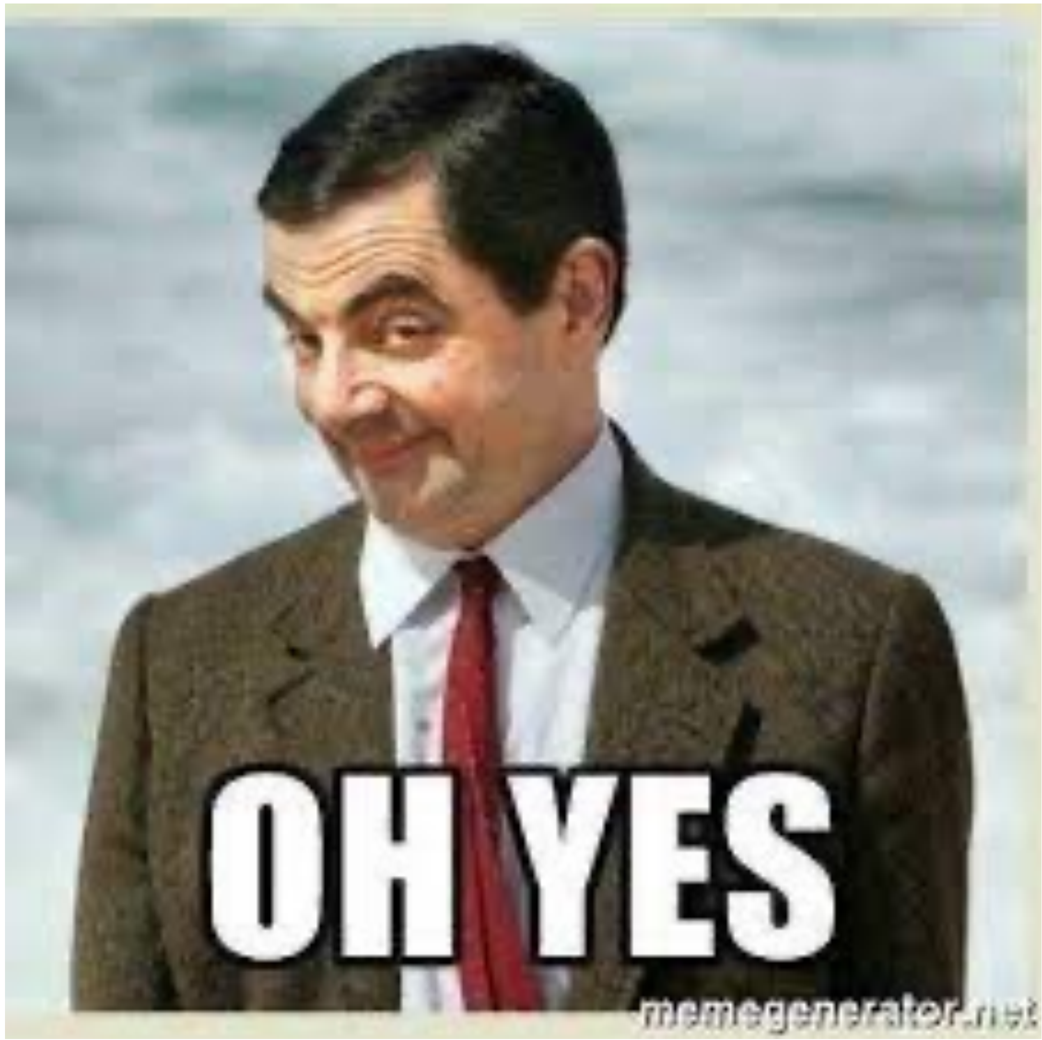
$$Y(s) = \frac{1}{s+a} X(s)$$

Transfer function $H(s)$

$$Y(s) = \frac{1}{s+a} X(s) \quad \Rightarrow \quad y(t) = \int_0^\infty e^{-au} x(t-u)\,du$$

$$Y(s) = H(s)X(s) \quad \Rightarrow \quad y(t) = \int_{-\infty}^\infty h(u) x(t-u)\,du$$

Is this really the impulse response?


OH YES
memegenerator.net

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

$$x(t) \longrightarrow \boxed{\text{linear system}} \longrightarrow y(t)$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$y'(t) + ay(t) = x(t)$$

$$sY(s) - \cancelto{0}{y(0^-)} + aY(s) = X(s)$$

$$sY(s) + aY(s) = X(s)$$

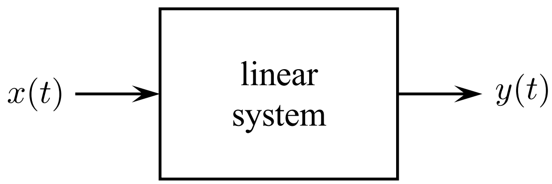
$$(s+a)Y(s) = X(s)$$

$$Y(s) = \frac{1}{s+a}X(s)$$

$$Y(s) = H(s)X(s)$$

$$x(t) = \delta(t) \quad \Rightarrow \quad X(s) = 1$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$y'(t) + ay(t) = x(t)$$

$$sY(s) - \cancelto{0}{y(0^-)} + aY(s) = X(s)$$

$$sY(s) + aY(s) = X(s)$$

$$(s+a)Y(s) = X(s)$$

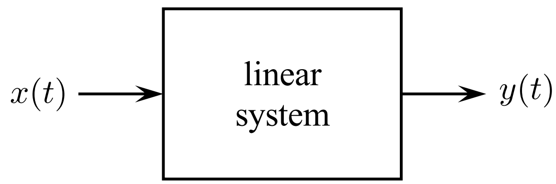
$$Y(s) = \frac{1}{s+a}X(s)$$

$$Y(s) = H(s)X(s)$$

$$x(t) = \delta(t) \quad \Rightarrow \quad X(s) = 1$$

$$Y(s) = H(s)X(s) \quad \Rightarrow \quad Y(s) = H(s)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$y'(t) + ay(t) = x(t)$$

$$sY(s) - \cancelto{0}{y(0^-)} + aY(s) = X(s)$$

$$sY(s) + aY(s) = X(s)$$

$$(s+a)Y(s) = X(s)$$

$$Y(s) = \frac{1}{s+a} X(s)$$

$$Y(s) = H(s)X(s)$$

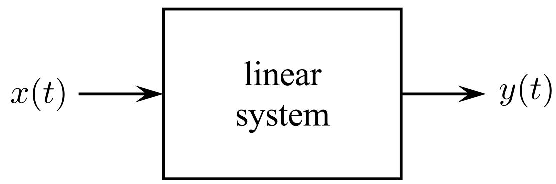
$$x(t) = \delta(t) \quad \Rightarrow \quad X(s) = 1$$

$$Y(s) = H(s)X(s) \quad \Rightarrow \quad Y(s) = H(s)$$

$$Y(s) = H(s) \quad \Rightarrow \quad y(t) = h(t)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

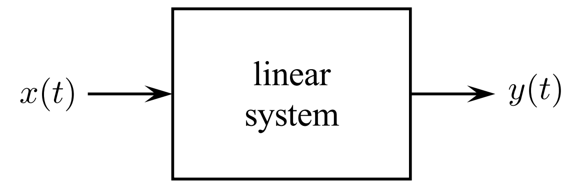


$$y'(t) + ay(t) = x(t)$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

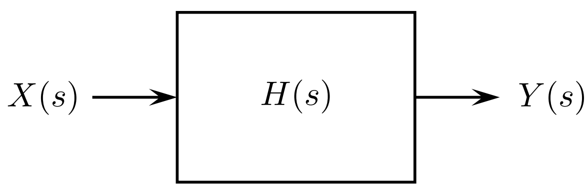
Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



$$y'(t) + ay(t) = x(t)$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$



Laplace domain transfer function

$$Y(s) = H(s)X(s)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

$$x(t) \longrightarrow \boxed{\begin{array}{c}\text{linear}\\ \text{system}\end{array}} \longrightarrow y(t)$$

$$y'(t) + ay(t) = x(t)$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$X(s) \longrightarrow \boxed{H(s)} \longrightarrow Y(s)$$

Laplace domain transfer function

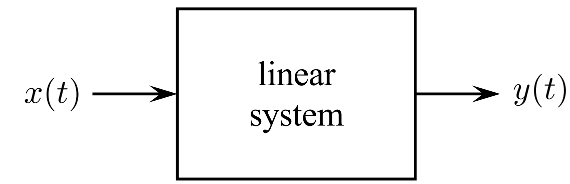
$$Y(s) = H(s)X(s)$$

$$x(t) \longrightarrow \boxed{h(t)} \longrightarrow y(t)$$

Time domain convolution

$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

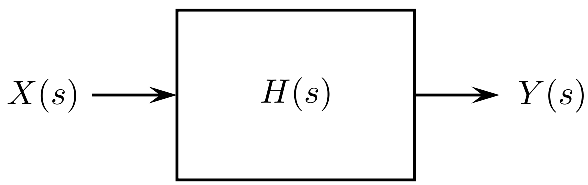
Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



$y'(t) + ay(t) = x(t)$

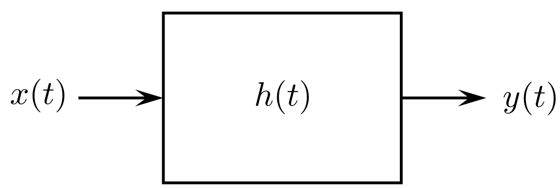
The linear systems of interest to us have an input/output relationship defined by LCCDE

$$\text{e.g., } y'(t) + ay(t) = x(t)$$



Laplace domain transfer function

$$Y(s) = H(s)X(s)$$



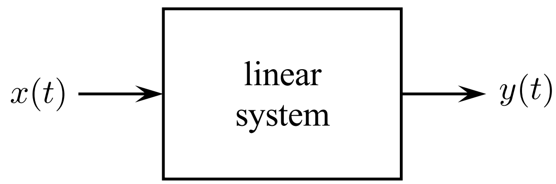
Time domain convolution

$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

To find the impulse response of an LTI system described by an LCCDE

1. Solve the LCCDE using the Laplace transform with all-zeros initial conditions.
2. Write the Laplace domain solution as $Y(s) = H(s)X(s)$.
3. Identify $H(s)$ in the solution.
4. The impulse response $h(t)$ is the inverse Laplace transform of $H(s)$.

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + Fourier Transform

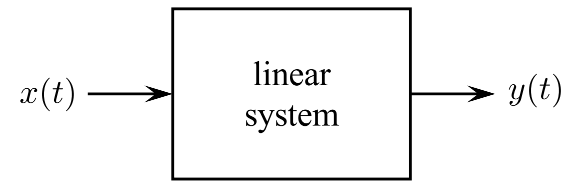


$$y'(t) + ay(t) = x(t)$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

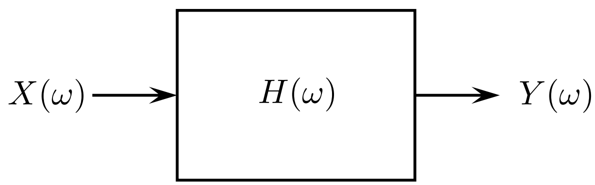
Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + Fourier Transform



$$y'(t) + ay(t) = x(t)$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

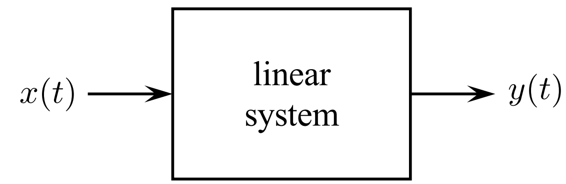
e.g., $y'(t) + ay(t) = x(t)$



Fourier domain frequency response

$$Y(\omega) = H(\omega)X(\omega)$$

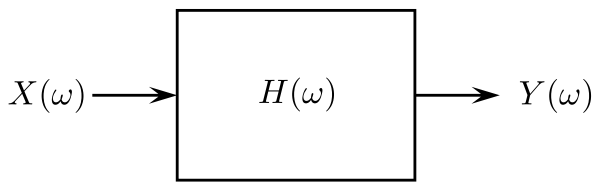
Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + Fourier Transform



$y'(t) + ay(t) = x(t)$

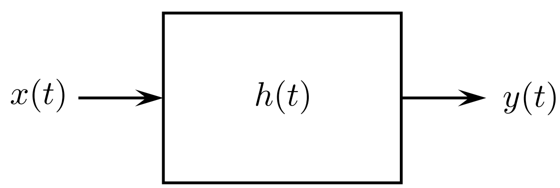
The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$



Fourier domain frequency response

$$Y(\omega) = H(\omega)X(\omega)$$



Time domain convolution

$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + Fourier Transform

$$x(t) \longrightarrow \boxed{\begin{array}{c}\text{linear}\\ \text{system}\end{array}} \longrightarrow y(t)$$

$$y'(t) + ay(t) = x(t)$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$X(\omega) \longrightarrow \boxed{H(\omega)} \longrightarrow Y(\omega)$$

Fourier domain frequency response

$$Y(\omega) = H(\omega)X(\omega)$$

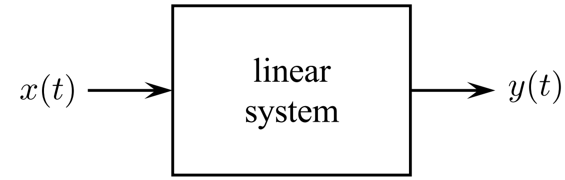
$$x(t) \longrightarrow \boxed{h(t)} \longrightarrow y(t)$$

Time domain convolution

$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

- One can solve an LCCDE using the Fourier transform with all-zeros initial conditions. But it is more common to use the Laplace transform.

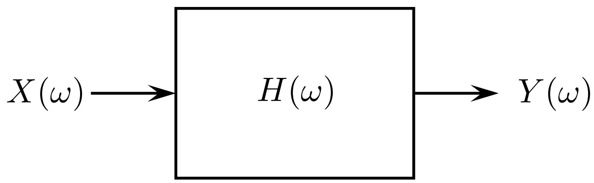
Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + Fourier Transform



$y'(t) + ay(t) = x(t)$

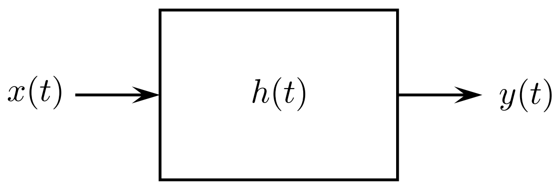
The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$



Fourier domain frequency response

$$Y(\omega) = H(\omega)X(\omega)$$



Time domain convolution

$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

- One can solve an LCCDE using the Fourier transform with all-zeros initial conditions. But it is more common to use the Laplace transform.
- After identifying the Laplace domain transfer function $H(s)$, the Fourier domain frequency response is

$$H(\omega) = H(s)\Big|_{s=j\omega}$$

This works only as long as the ROC of $H(s)$ contains the $s = j\omega$ axis.

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + Fourier Transform

$$x(t) \longrightarrow \boxed{\text{linear system}} \longrightarrow y(t)$$

$$y'(t) + ay(t) = x(t)$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y'(t) + ay(t) = x(t)$

$$X(\omega) \longrightarrow \boxed{H(\omega)} \longrightarrow Y(\omega)$$

Fourier domain frequency response

$$Y(\omega) = H(\omega)X(\omega)$$

$$x(t) \longrightarrow \boxed{h(t)} \longrightarrow y(t)$$

Time domain convolution

$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)x(t-u)du$$

- One can solve an LCCDE using the Fourier transform with all-zeros initial conditions. But it is more common to use the Laplace transform.
- After identifying the Laplace domain transfer function $H(s)$, the Fourier domain frequency response is

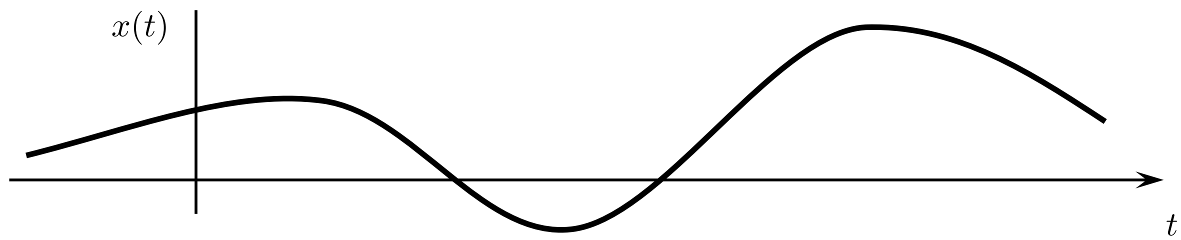
$$H(\omega) = H(s)\Big|_{s=j\omega}$$

  This works only as long as the ROC of $H(s)$ contains the $s = j\omega$ axis.
- For stable systems, this is always true.

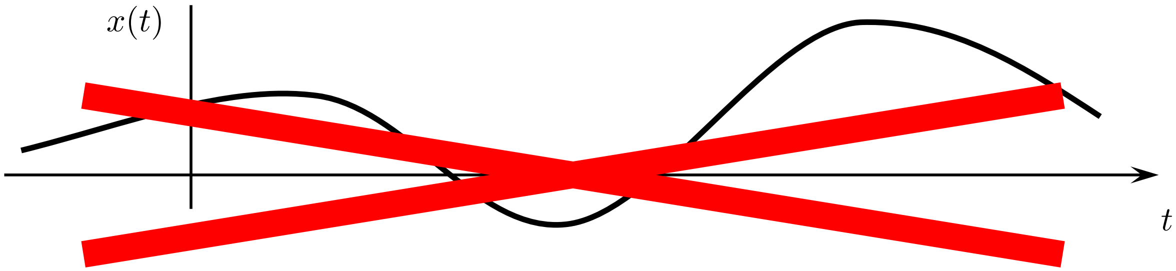
A Review of Discrete Time Linear Systems

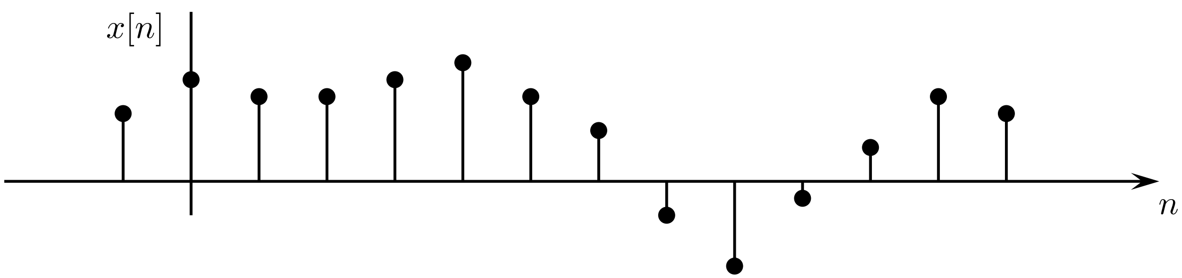
- Description of Discrete Time Signals
- A Description of Discrete Time Systems
- Frequency Domain Analysis

Discrete Time Signals

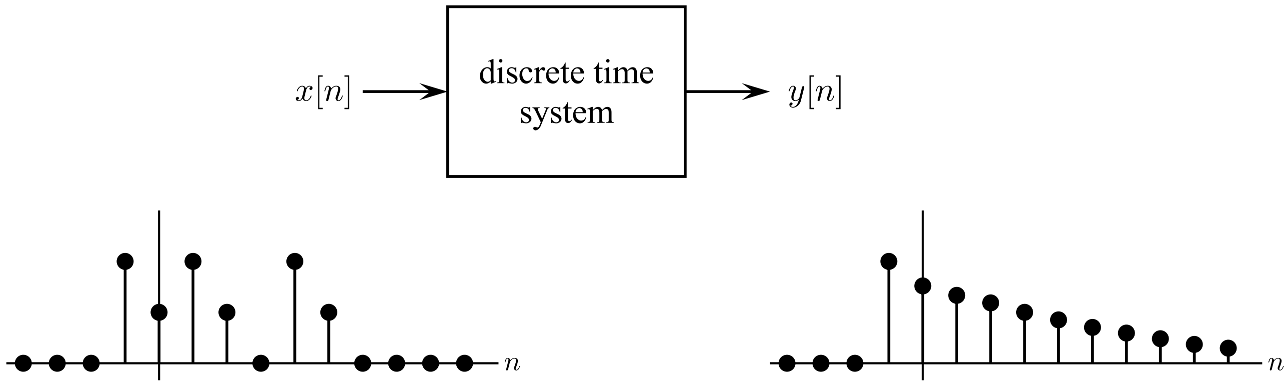


Discrete Time Signals



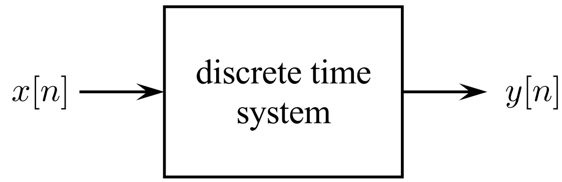


Discrete Time System



A discrete time system produces an output discrete time signal from an input discrete time signal.

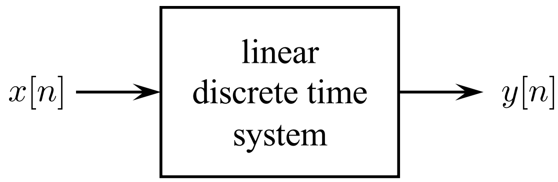
Linear Discrete Time System



$$y[n] + a_1 y[n-1] = b_0 x[n] + b_1 x[n-1]$$

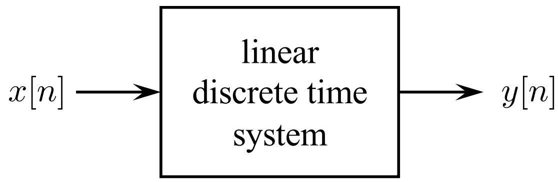
Linear **C**onstant **C**oefficient **D**ifference **E**quation
LCCDE

Recursions in the form of an LCCDE are the most popular for linear discrete time systems because only multiplications by fixed constants and additions are required.



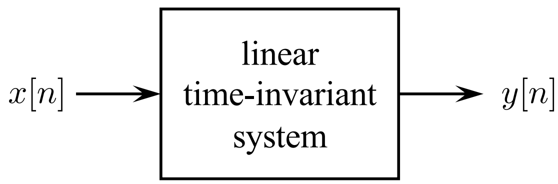
The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

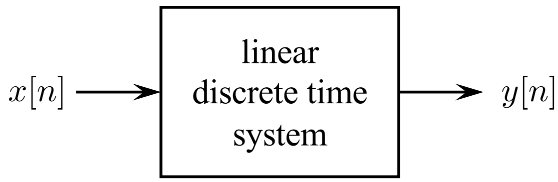


The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

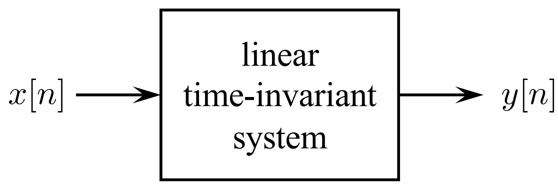


Systems defined by an LCCDE are also time-invariant

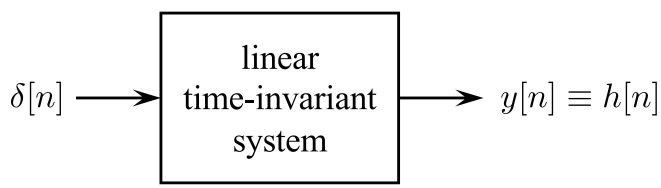


The linear systems of interest to us have an input/output relationship defined by LCCDE

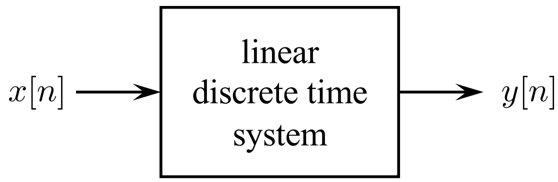
$$\text{e.g., } y[n] - ay[n-1] = x[n]$$



Systems defined by an LCCDE are also time-invariant

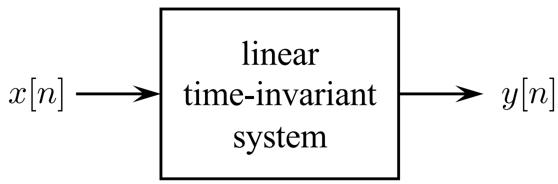


The input/output relationship of a linear time-invariant system may also be described using the impulse response and convolution.

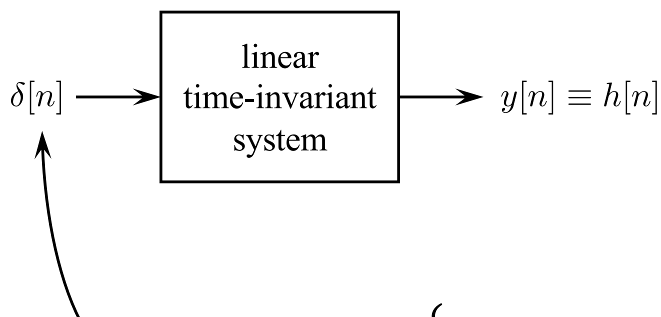


The linear systems of interest to us have an input/output relationship defined by LCCDE

$$\text{e.g., } y[n] - ay[n-1] = x[n]$$

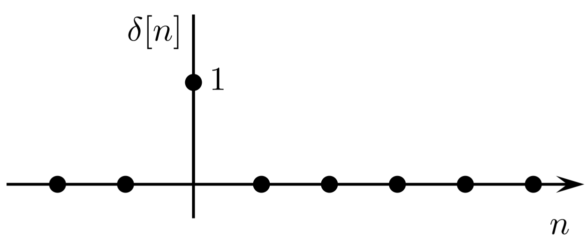


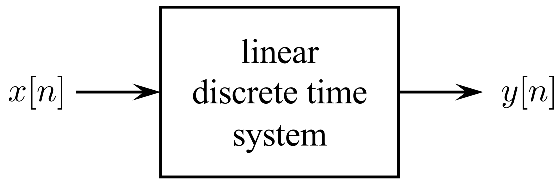
Systems defined by an LCCDE are also time-invariant



The input/output relationship of a linear time-invariant system may also be described using the impulse response and convolution.

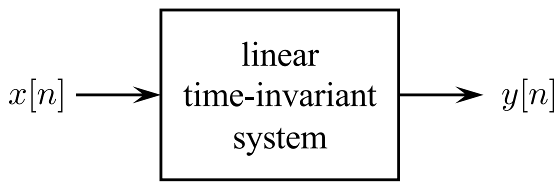
$$\text{Kronecker “delta”: } \delta[n] = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases}$$



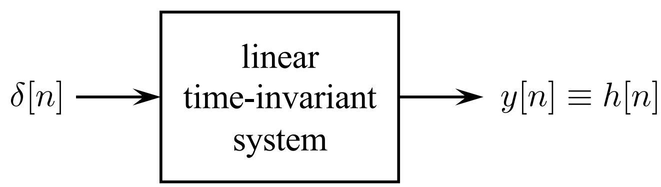


The linear systems of interest to us have an input/output relationship defined by LCCDE

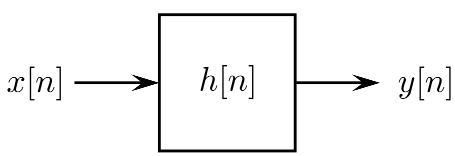
$$\text{e.g., } y[n] - ay[n-1] = x[n]$$



Systems defined by an LCCDE are also time-invariant



The input/output relationship of a linear time-invariant system may also be described using the impulse response and convolution.



$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

Convolution Example

$$x[n] \longrightarrow \boxed{h[n]} \longrightarrow y[n]$$

$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

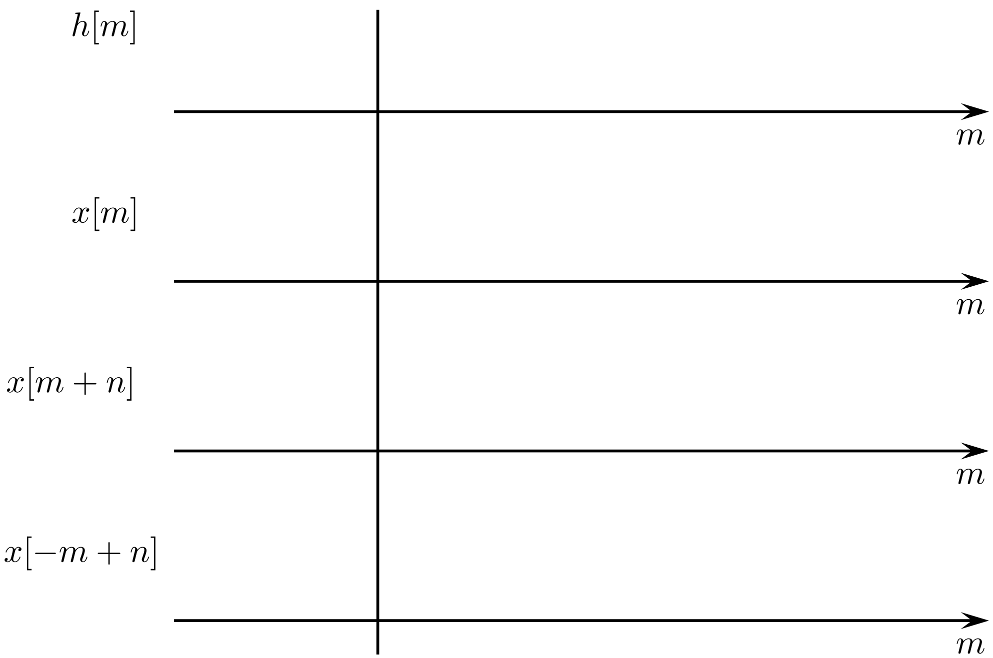
$x[n] = U[n] \qquad h[n] = a^n U[n]$

Convolution Example

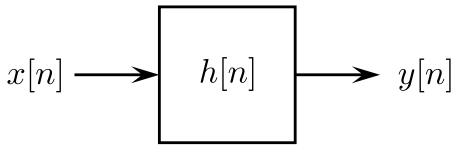
$$x[n] \longrightarrow \boxed{h[n]} \longrightarrow y[n]$$

$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

$x[n] = U[n]$ $\quad$ $h[n] = a^n U[n]$

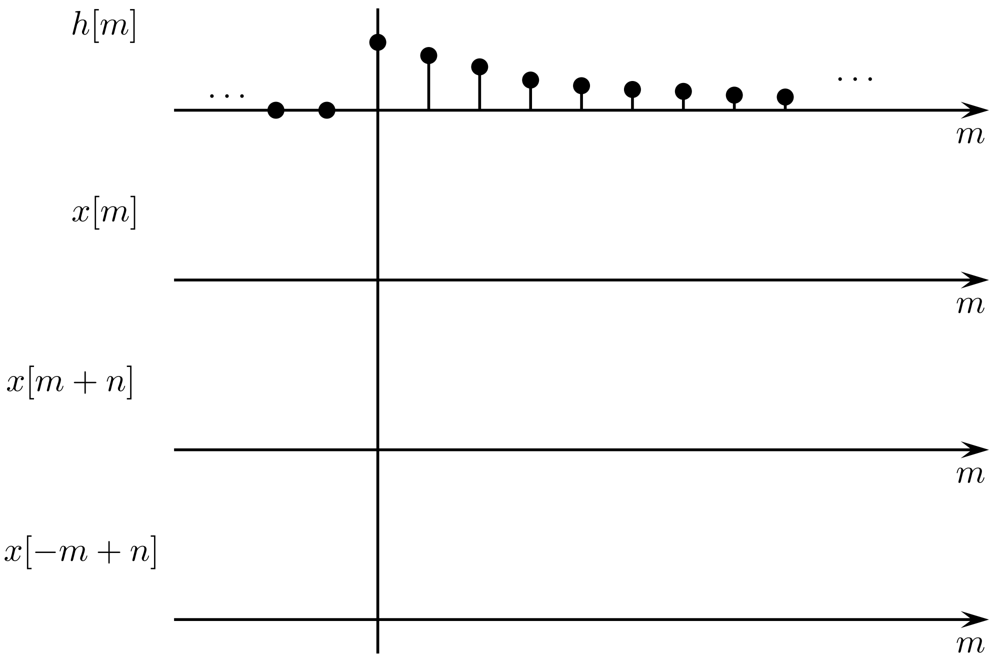


Convolution Example



$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

$x[n] = U[n] \qquad h[n] = a^n U[n]$

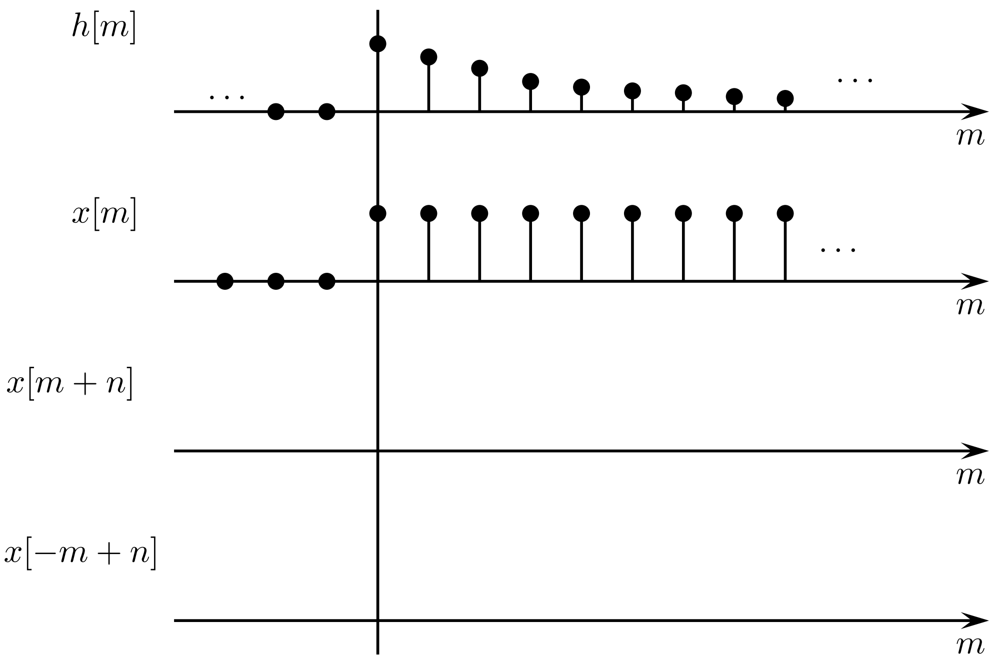


Convolution Example

$$x[n] \longrightarrow \boxed{h[n]} \longrightarrow y[n]$$

$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

$x[n] = U[n] \qquad h[n] = a^n U[n]$

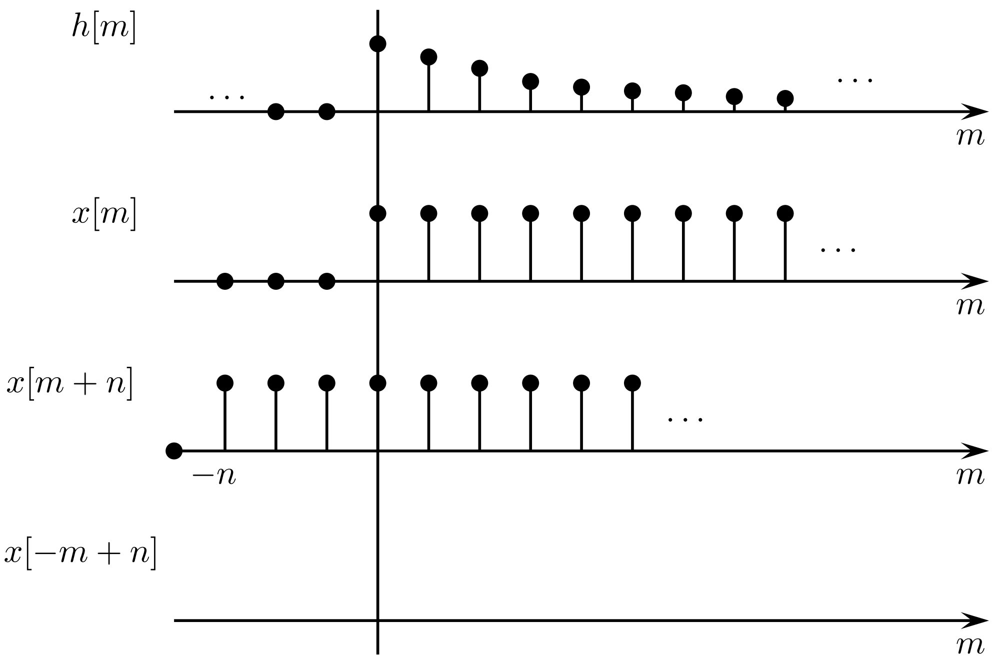


Convolution Example

$$x[n] \longrightarrow \boxed{h[n]} \longrightarrow y[n]$$

$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

$x[n] = U[n] \qquad h[n] = a^n U[n]$

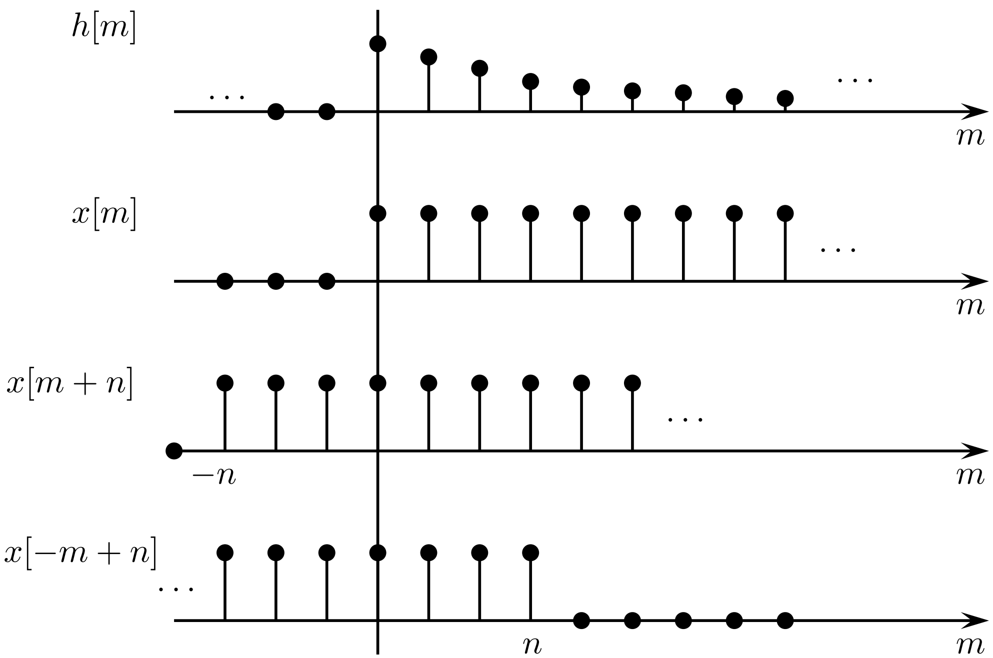


Convolution Example

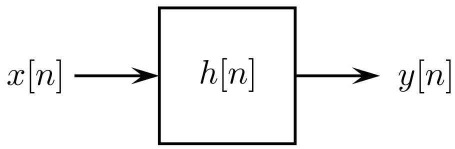
$$x[n] \longrightarrow \boxed{h[n]} \longrightarrow y[n]$$

$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

$x[n] = U[n] \qquad h[n] = a^n U[n]$

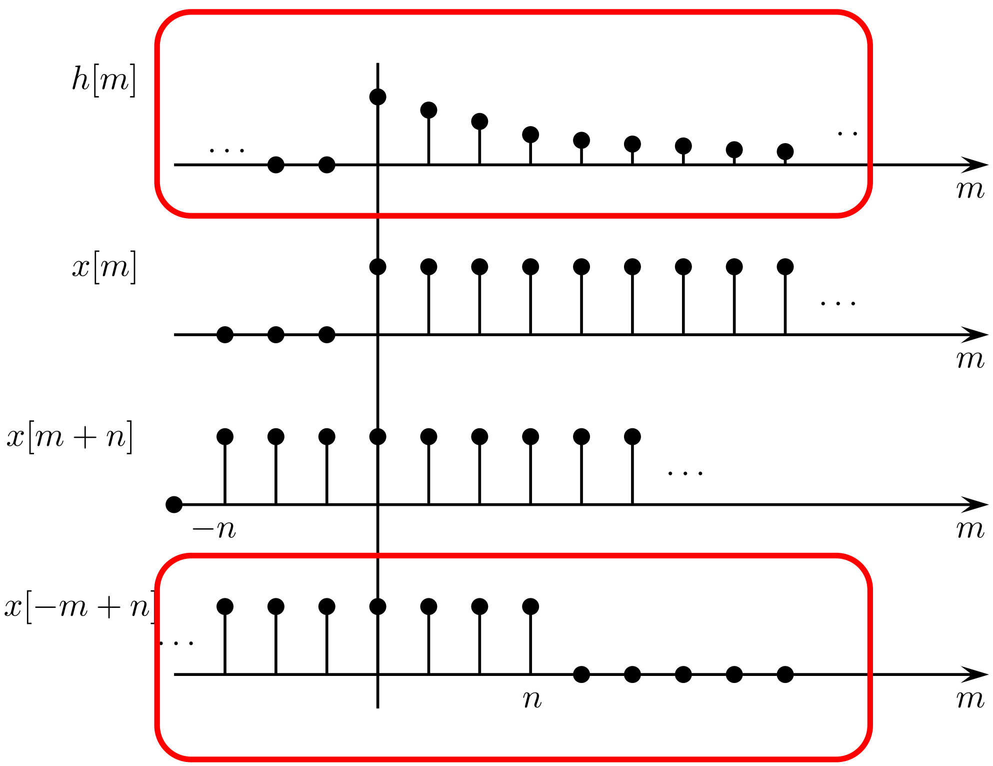


Convolution Example



$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

$x[n] = U[n] \qquad h[n] = a^n U[n]$

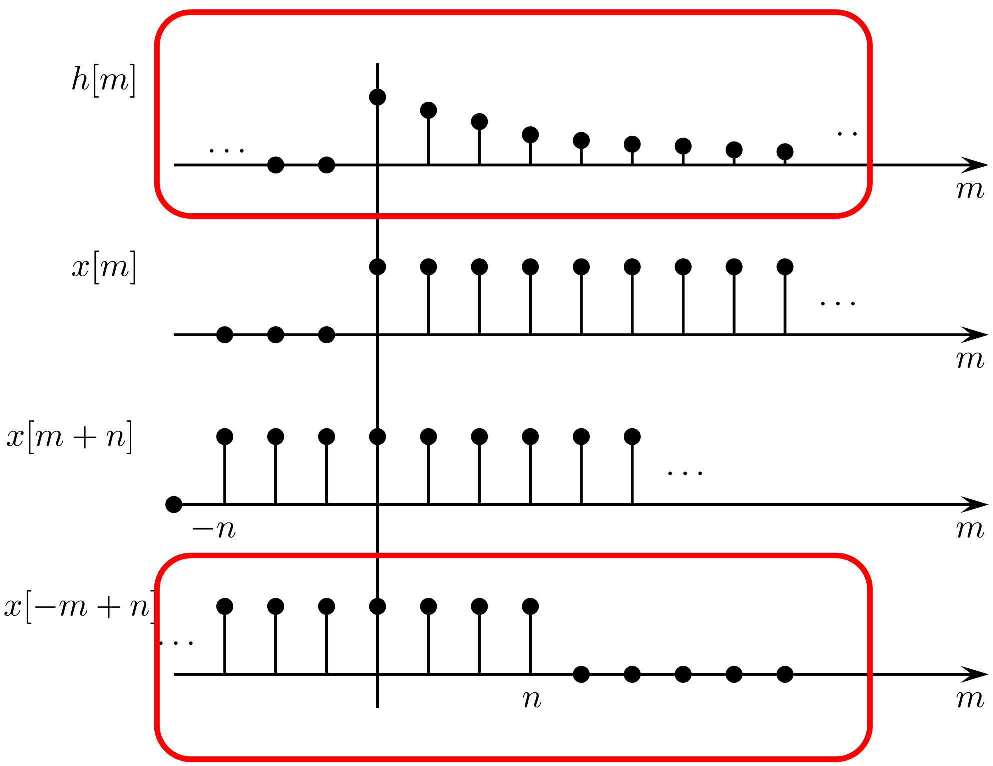


Convolution Example

$$x[n] \longrightarrow \boxed{h[n]} \longrightarrow y[n]$$

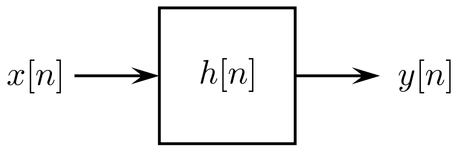
$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

$x[n] = U[n] \qquad h[n] = a^n U[n]$



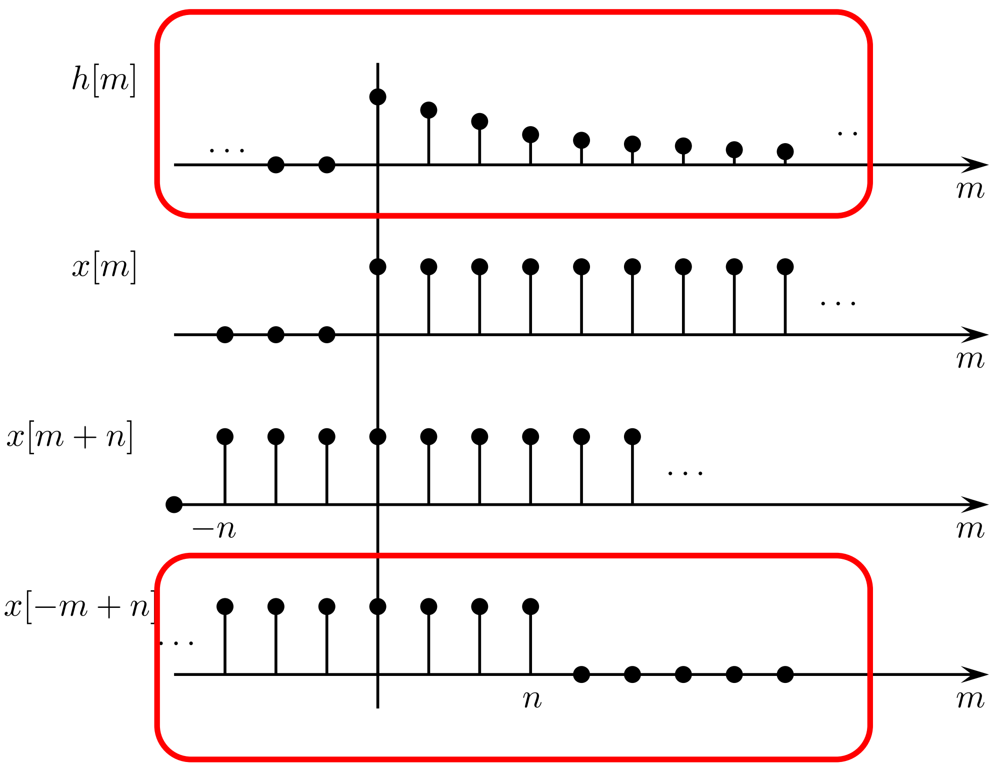
$n < 0, \qquad y[n] = 0$

Convolution Example



$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

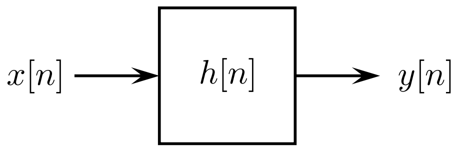
$x[n] = U[n] \qquad h[n] = a^n U[n]$



$n < 0, \qquad y[n] = 0$

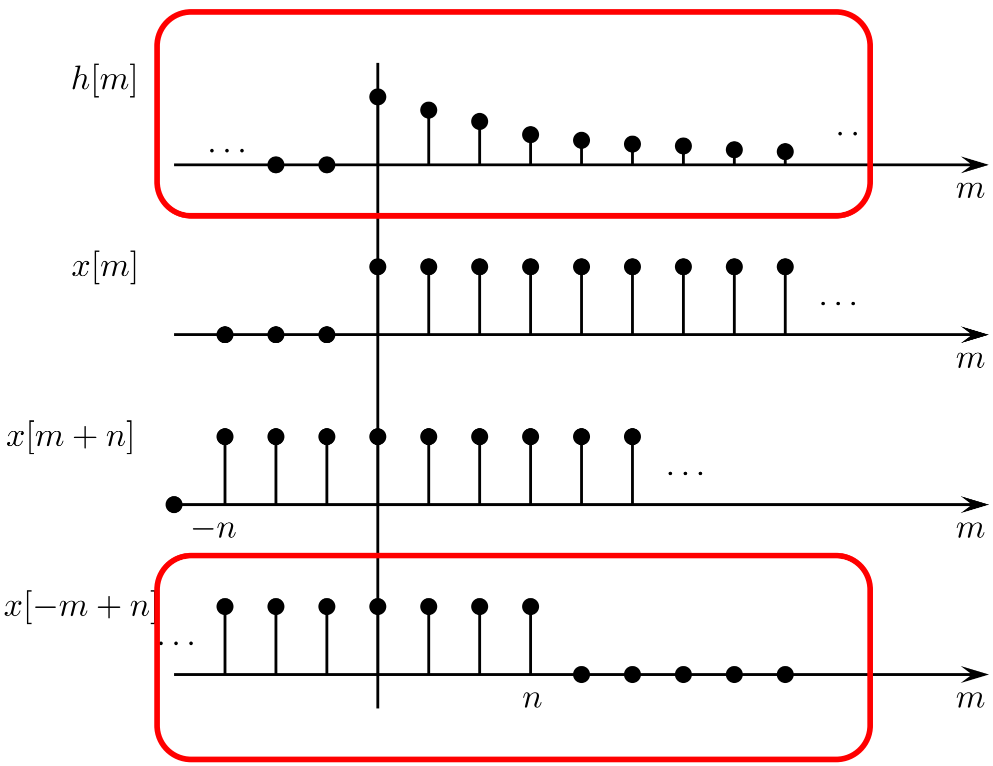
$$n \geq 0, \qquad y[n] = \sum_{m=0}^{n} a^m = \frac{1 - a^{n+1}}{1 - a}$$

Convolution Example



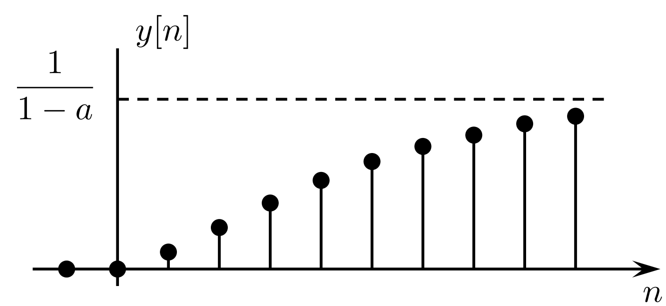
$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

$x[n] = U[n] \qquad h[n] = a^n U[n]$



$$n < 0, \qquad y[n] = 0$$

$$n \geq 0, \qquad y[n] = \sum_{m=0}^{n} a^m = \frac{1-a^{n+1}}{1-a}$$



Convolution Example

$$x[n] \longrightarrow \boxed{h[n]} \longrightarrow y[n]$$

$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

$x[n] = U[n] \qquad h[n] = a^n U[n]$

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$$

$$\sum_{n=0}^{N-1} r^n = \frac{1-r^N}{1-r}$$

$$\sum_{n=1}^{\infty} r^n = \frac{r}{1-r}$$

$$\sum_{n=0}^{\infty} r^{n-1} = \frac{1}{r(1-r)}$$

$$\sum_{n=0}^{\infty} nr^{n-1} = \frac{1}{(1-r)^2}$$

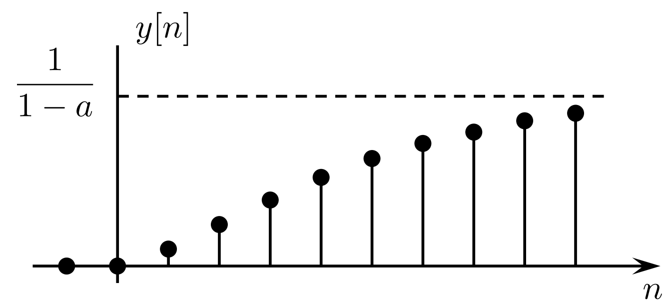
$$\sum_{n=0}^{\infty} nr^n = \frac{r}{(1-r)^2}$$

geometric sums and their variations are helpful in this discrete time business

$|r| < 1$

$$n < 0, \qquad y[n] = 0$$

$$n \geq 0, \qquad y[n] = \sum_{m=0}^{n} a^m = \frac{1-a^{n+1}}{1-a}$$



Frequency Domain Analysis

z Transform

$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Frequency Domain Analysis

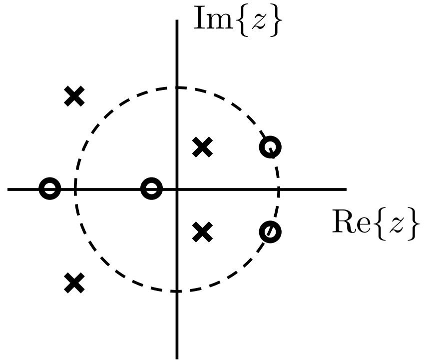
z Transform

$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.



Frequency Domain Analysis

z Transform

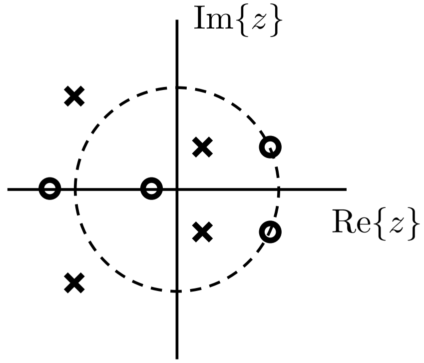
$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$



Frequency Domain Analysis

z Transform

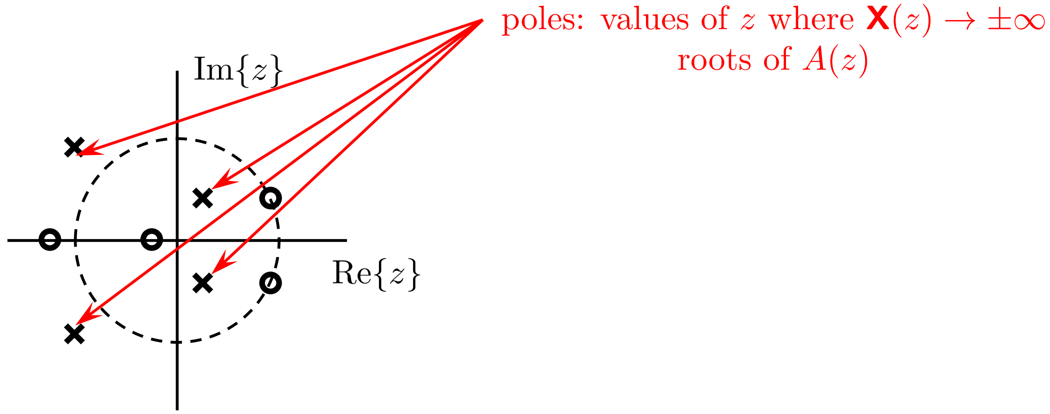
$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$



Frequency Domain Analysis

z Transform

$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

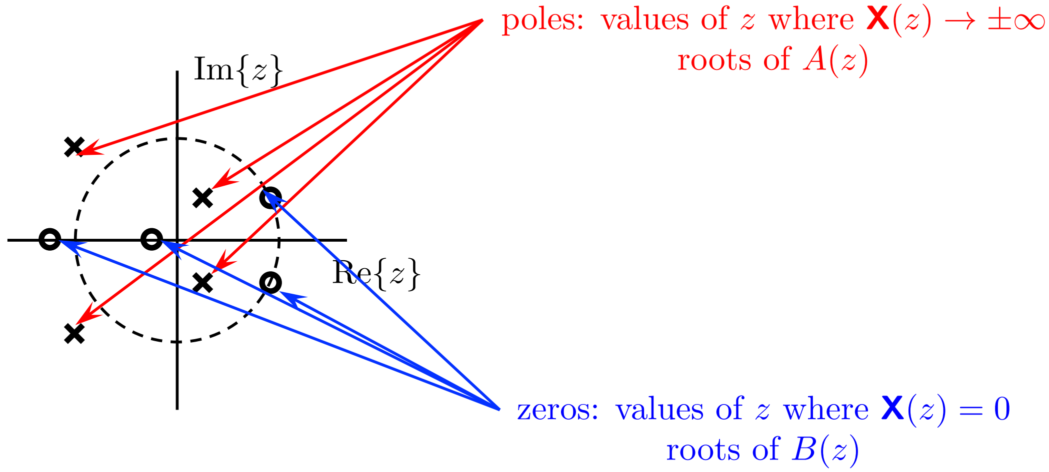
$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.

- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$



z Transform

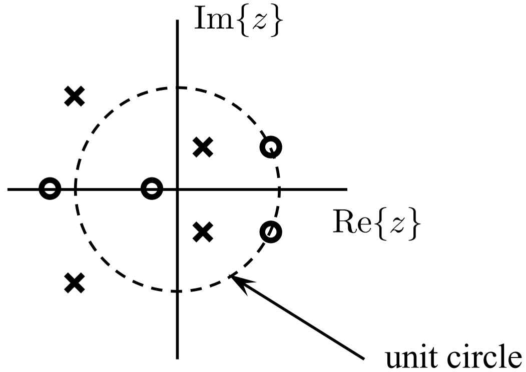
$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$



The z transform sum

$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

only converges for certain values of z. These values are called *the region of convergence* (ROC).

- The ROC cannot contain any poles.
- The ROC for the z transform of a stable (bounded) time-domain signal contains the countour $e^{j\omega}$. This contour is called the *unit circle*.

Frequency Domain Analysis

z Transform

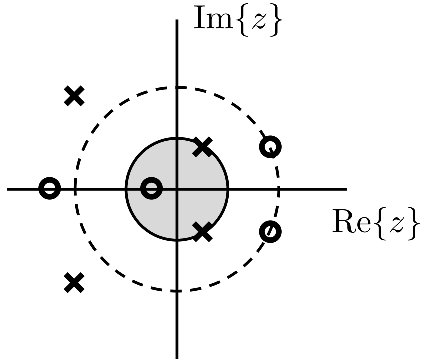
$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

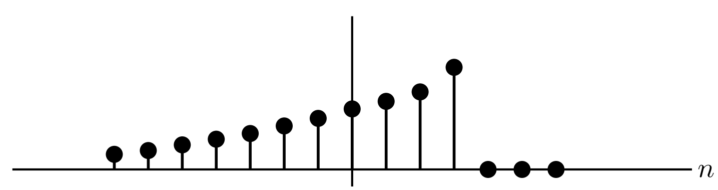
Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$



ROC for left-sided signals



Frequency Domain Analysis

z Transform

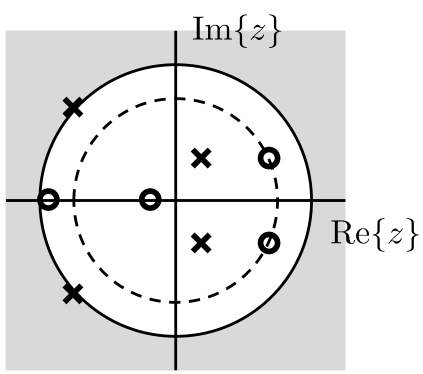
$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

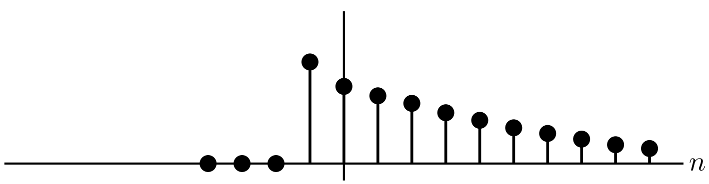
Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$



ROC for right-sided signals



Frequency Domain Analysis

z Transform

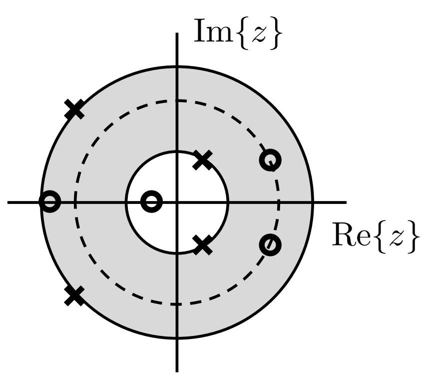
$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

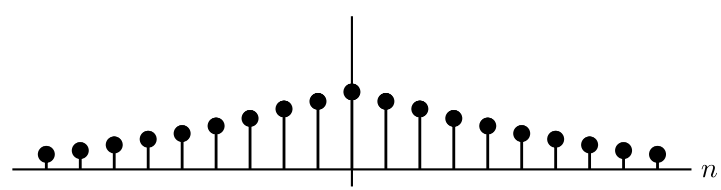
Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$



ROC for two-sided signals



z Transform

$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$

- Because z is a complex variable, z as a *variable of integration* defines a *contour integral* in the complex plane.

z Transform

$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$

- Because z is a complex variable, z as a *variable of integration* defines a *contour integral* in the complex plane.
  - Contour integral wizards compute the inverse z transform from the definition.
  - Everyone else uses partial fraction expansion to construct tables, then uses the tables.

z Transform

$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$

- Because z is a complex variable, z as a *variable of integration* defines a *contour integral* in the complex plane.
  - Contour integral wizards compute the inverse z transform from the definition.
  - Everyone else uses partial fraction expansion to construct tables, then uses the tables.

Signal	Transform	ROC
$\delta[n]$	1	all z
$U[n]$	$\frac{1}{1 - z^{-1}}$	$\lvert z\rvert > 1$
$a^n U[n]$	$\frac{1}{1 - az^{-1}}$	$\lvert z\rvert > \lvert a\rvert$
$na^n U[n]$	$\frac{az^{-1}}{(1 - az^{-1})^2}$	$\lvert z\rvert > \lvert a\rvert$
$\delta[n-m]$	z^{-m}	all z[a]
$\cos(\omega_0 n) U[n]$	$\frac{1 - \cos(\omega_0) z^{-1}}{1 - [2\cos(\omega_0)] z^{-1} + z^{-2}}$	$\lvert z\rvert > 1$
$\sin(\omega_0 n) U[n]$	$\frac{\sin(\omega_0) z^{-1}}{1 - [2\cos(\omega_0)] z^{-1} + z^{-2}}$	$\lvert z\rvert > 1$
$r^n \cos(\omega_0 n) U[n]$	$\frac{1 - r\cos(\omega_0) z^{-1}}{1 - [2r\cos(\omega_0)] z^{-1} + z^{-2}}$	$\lvert z\rvert > r$
$r^n \sin(\omega_0 n) U[n]$	$\frac{r\sin(\omega_0) z^{-1}}{1 - [2r\cos(\omega_0)] z^{-1} + z^{-2}}$	$\lvert z\rvert > r$

[a] except 0 if $m > 0$ or ∞ if $m > 0$

Property	Signal	z-transform	ROC
	$x[n]$	$\mathbf{X}(z)$	R_x
	$y[n]$	$\mathbf{Y}(z)$	R_y
Linearity	$ax[n] + by[n]$	$a\mathbf{X}(z) + b\mathbf{Y}(z)$	at least $R_x \cap R_y$
Time Shifting	$x[n - n_0]$	$z^{-n_0}\mathbf{X}(z)$	R_x
Time Scaling (Upsampling)	$\begin{cases} x(n/K) & n \text{ is a multiple of K} \\ 0 & \text{otherwise} \end{cases}$	$\mathbf{X}(z^K)$	$R_x^{1/K}$
Conjugation	$x^*(n)$	$\mathbf{X}^*(z^*)$	R_x
Convolution	$x[n] * y[n]$	$\mathbf{X}(z)\mathbf{Y}(z)$	at least $R_x \cap R_y$
First Difference	$x[n] - x[n-1]$	$(1 - z^{-1})\mathbf{X}(z)$	at least $R_x \cap \left\{\lvert z\rvert > 0\right\}$
Accumulation	$\sum_{k=-\infty}^{n} x[k]$	$\frac{1}{1 - z^{-1}}\mathbf{X}(z)$	at least $R_x \cap \left\{\lvert z\rvert > 1\right\}$
Initial Value Theorem	$x[0]$	$\lim_{z\to\infty} \mathbf{X}(z)$	
Final Value Theorem	$\lim_{N\to\infty} x[n]$	$\lim_{z\to 1}(1 - z^{-1})\mathbf{X}(z)$	poles of $(1 - z^{-1})\mathbf{X}(z)$ inside unit circle.

Frequency Domain Analysis

z Transform

$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$

- Because z is a complex variable, z as a *variable of integration* defines a *contour integral* in the complex plane.
  - Contour integral wizards compute the inverse z transform from the definition.
  - Everyone else uses partial fraction expansion to construct tables, then uses the tables.

Discrete Time Fourier Transform (DTFT)

$$\mathbf{X}(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}$$

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

Frequency Domain Analysis

z Transform

$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$

- Because z is a complex variable, z as a *variable of integration* defines a *contour integral* in the complex plane.
  - Contour integral wizards compute the inverse z transform from the definition.
  - Everyone else uses partial fraction expansion to construct tables, then uses the tables.

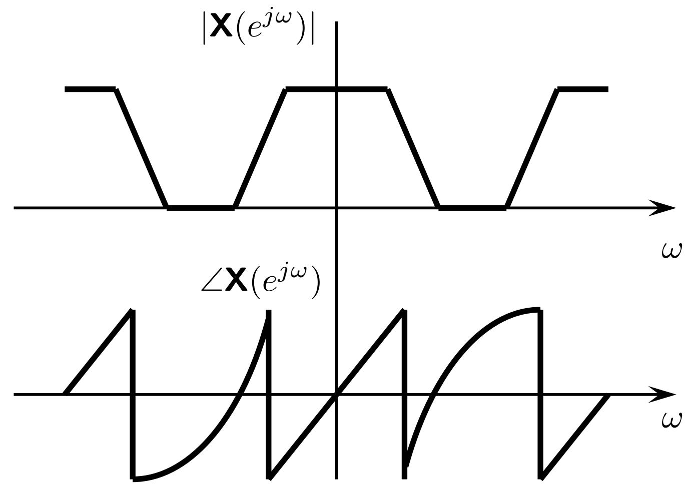
Discrete Time Fourier Transform (DTFT)

$$\mathbf{X}(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}$$

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

Comments:

- w is a real variable and $\mathbf{X}(e^{j\omega})$ is a complex-valued function of the real variable ω.



Frequency Domain Analysis

z Transform

$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$

- Because z is a complex variable, z as a *variable of integration* defines a *contour integral* in the complex plane.
  - Contour integral wizards compute the inverse z transform from the definition.
  - Everyone else uses partial fraction expansion to construct tables, then uses the tables.

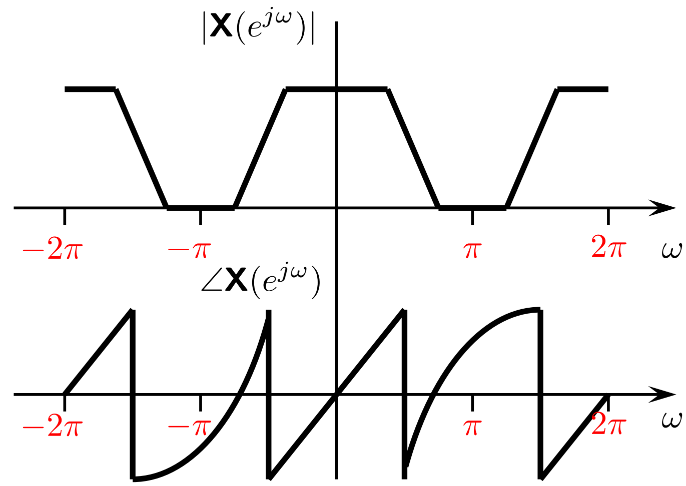
Discrete Time Fourier Transform (DTFT)

$$\mathbf{X}(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}$$

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

Comments:

- w is a real variable and $\mathbf{X}(e^{j\omega})$ is a complex-valued function of the real variable ω.
- $\mathbf{X}(e^{j\omega})$ is *periodic* in ω with period 2π.



z Transform

$$\mathbf{X}(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

$$x[n] = \frac{1}{j2\pi} \oint \mathbf{X}(z) z^{n-1} dz$$

Comments:

- z is a complex variable and $\mathbf{X}(z)$ is a complex-valued function of the complex variable z.
- LCCDEs always produce a ratio of polynomials in z^{-1}:

$$\mathbf{X}(z) = \frac{B(z)}{A(z)} = \frac{b_q z^{-q} + \cdots + b_1 z^{-1} + b_0}{a_p z^{-p} + \cdots + a_1 z^{-1} + a_0}$$

- Because z is a complex variable, z as a *variable of integration* defines a *contour integral* in the complex plane.
  - Contour integral wizards compute the inverse z transform from the definition.
  - Everyone else uses partial fraction expansion to construct tables, then uses the tables.

Discrete Time Fourier Transform (DTFT)

$$\mathbf{X}(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}$$

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

Comments:

- w is a real variable and $\mathbf{X}(e^{j\omega})$ is a complex-valued function of the real variable ω.
- $\mathbf{X}(e^{j\omega})$ is *periodic* in ω with period 2π.
- Because ω is a real variable, the integral that defines the inverse transform is the familiar integral introduced in your first and second calculus courses.
  - Integral wizards compute the inverse DTFT transform from the definition.
  - Everyone else uses uses tables.

Signal	DTFT
$x[n]$	$\mathbf{X}\left(e^{j\omega}\right)$
$\delta[n]$	1
$\delta[n-n_0]$	$e^{-j\omega n_0}$
$U[n]$	$\dfrac{1}{1-e^{-j\omega}} + \pi \sum\limits_{\ell=-\infty}^{\infty} \delta\left(\omega - 2\pi\ell\right)$
$a^n U[n], \lvert a\rvert < 1$	$\dfrac{1}{1-ae^{-j\omega}}$
$\sum\limits_{k=k_0}^{k_0+N-1} c_k e^{jk(2\pi/N)n}$	$2\pi \sum\limits_{k=-\infty}^{\infty} c_k \delta\left(\omega - \dfrac{2\pi k}{N}\right)$
$\sum\limits_{k=-\infty}^{\infty} \delta[n-kN]$	$\dfrac{2\pi}{N} \sum\limits_{\ell=-\infty}^{\infty} \delta\left(\omega - \dfrac{2\pi\ell}{N}\right)$
1	$2\pi \sum\limits_{\ell=-\infty}^{\infty} \delta\left(\omega - 2\pi\ell\right)$
$\begin{cases} 1 & \lvert n\rvert \le N \\ 0 & \lvert n\rvert > N \end{cases}$	$\dfrac{\sin\left(\omega\left[N+\frac{1}{2}\right]\right)}{\sin\left(\frac{\omega}{2}\right)}$
$e^{j\omega_0 n}$	$2\pi \sum\limits_{\ell=-\infty}^{\infty} \delta\left(\omega - \omega_0 - 2\pi\ell\right)$
$\cos\left(\omega_0 n\right)$	$\pi \sum\limits_{\ell=-\infty}^{\infty} \Big\{\delta(\omega-\omega_0-2\pi\ell) + \delta(\omega+\omega_0-2\pi\ell)\Big\}$
$\sin\left(\omega_0 n\right)$	$\dfrac{\pi}{j} \sum\limits_{\ell=-\infty}^{\infty} \Big\{\delta(\omega-\omega_0-2\pi\ell) - \delta(\omega+\omega_0-2\pi\ell)\Big\}$

Property	Signal	DTFT
	$x[n]$	$\mathbf{X}\left(e^{j\omega}\right)$
	$y[n]$	$\mathbf{Y}\left(e^{j\omega}\right)$
Linearity	$ax[n]+by[n]$	$a\mathbf{X}\left(e^{j\omega}\right)+b\mathbf{Y}\left(e^{j\omega}\right)$
Time Shifting	$x[n-n_0]$	$e^{-j\omega n_0}\mathbf{X}\left(e^{j\omega}\right)$
Time Scaling (Upsampling)	$\begin{cases} x[n/K] & n \text{ is a multiple of K} \\ 0 & \text{otherwise} \end{cases}$	$\mathbf{X}\left(e^{jK\omega}\right)$
Conjugation	$x^*[n]$	$\mathbf{X}^*\left(e^{-j\omega}\right)$
Convolution	$x[n]*y[n]$	$\mathbf{X}\left(e^{j\omega}\right)\mathbf{Y}\left(e^{j\omega}\right)$
Multiplication	$x[n]y[n]$	$\frac{1}{2\pi}\mathbf{X}\left(e^{j\omega}\right)*\mathbf{Y}\left(e^{j\omega}\right)$
First Difference	$x[n]-x[n-1]$	$\left(1-e^{-j\omega}\right)\mathbf{X}\left(e^{j\omega}\right)$
Accumulation	$\sum_{k=-\infty}^{n} x[k]$	$\frac{1}{1-e^{-j\omega}}\mathbf{X}\left(e^{j\omega}\right)$
Parseval's Relation	$\sum_{n=-\infty}^{\infty}\lvert x[n]\rvert^2=\frac{1}{2\pi}\int_{2\pi}\left\lvert\mathbf{X}\left(e^{j\omega}\right)\right\rvert^2 d\omega$	

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

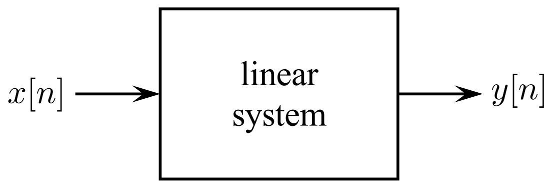
$$x[n] \longrightarrow \boxed{\begin{array}{c}\text{linear}\\ \text{system}\end{array}} \longrightarrow y[n]$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

$$y[n] - ay[n-1] = x[n]$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



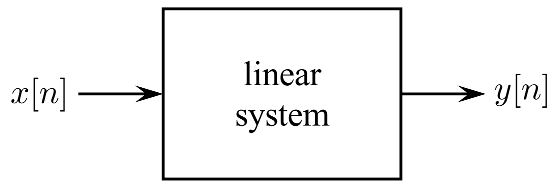
The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

$$y[n] - ay[n-1] = x[n]$$

$$\mathbf{Y}(z) - a\left(z^{-1}\mathbf{Y}(z) + y[-1]\right) = \mathbf{X}(z)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



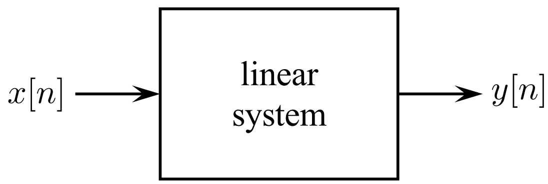
The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

$$y[n] - ay[n-1] = x[n]$$

$$\mathbf{Y}(z) - a\left(z^{-1}\mathbf{Y}(z) + \cancelto{0}{y[-1]}\right) = \mathbf{X}(z)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

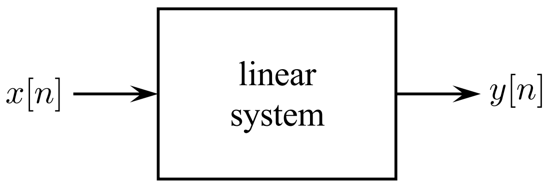
e.g., $y[n] - ay[n-1] = x[n]$

$$y[n] - ay[n-1] = x[n]$$

$$\mathbf{Y}(z) - a\left(z^{-1}\mathbf{Y}(z) + \cancelto{0}{y[-1]}\right) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) - az^{-1}\mathbf{Y}(z) = \mathbf{X}(z)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

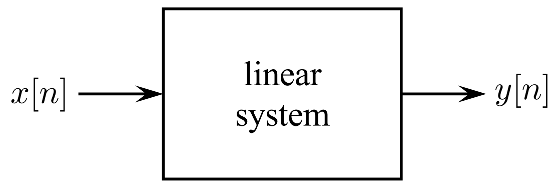
$$y[n] - ay[n-1] = x[n]$$

$$\mathbf{Y}(z) - a\left(z^{-1}\mathbf{Y}(z) + \cancelto{0}{y[-1]}\right) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) - az^{-1}\mathbf{Y}(z) = \mathbf{X}(z)$$

$$(1 - az^{-1})\mathbf{Y}(z) = \mathbf{X}(z)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

$$y[n] - ay[n-1] = x[n]$$

$$\mathbf{Y}(z) - a\left(z^{-1}\mathbf{Y}(z) + \cancelto{0}{y[-1]}\right) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) - az^{-1}\mathbf{Y}(z) = \mathbf{X}(z)$$

$$(1 - az^{-1})\mathbf{Y}(z) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) = \frac{1}{1 - az^{-1}}\mathbf{X}(z)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

$$x[n] \longrightarrow \boxed{\text{linear system}} \longrightarrow y[n]$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

$$\text{e.g., } y[n] - ay[n-1] = x[n]$$

$$y[n] - ay[n-1] = x[n]$$

$$\mathbf{Y}(z) - a\left(z^{-1}\mathbf{Y}(z) + \cancelto{0}{y[-1]}\right) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) - az^{-1}\mathbf{Y}(z) = \mathbf{X}(z)$$

$$(1 - az^{-1})\mathbf{Y}(z) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) = \underbrace{\frac{1}{1 - az^{-1}}}_{\text{Transfer function } \mathbf{H}(z)}\mathbf{X}(z)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

$$x[n] \longrightarrow \boxed{\begin{array}{c}\text{linear}\\ \text{system}\end{array}} \longrightarrow y[n]$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

$$\text{e.g., } y[n] - ay[n-1] = x[n]$$

$$y[n] - ay[n-1] = x[n]$$

$$\mathbf{Y}(z) - a\left(z^{-1}\mathbf{Y}(z) + \cancelto{0}{y[-1]}\right) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) - az^{-1}\mathbf{Y}(z) = \mathbf{X}(z)$$

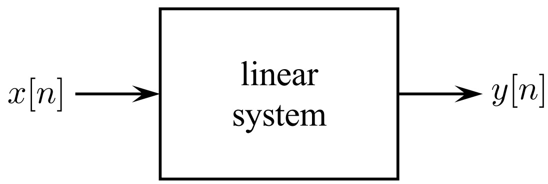
$$(1 - az^{-1})\mathbf{Y}(z) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) = \frac{1}{1 - az^{-1}}\mathbf{X}(z)$$

Transfer function $\mathbf{H}(z)$

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{X}(z) \quad \Rightarrow \quad y[n] = \sum_{m=\infty}^{\infty} h[m]x[n-m]$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

$$\text{e.g., } y[n] - ay[n-1] = x[n]$$

$$y[n] - ay[n-1] = x[n]$$

$$\mathbf{Y}(z) - a\left(z^{-1}\mathbf{Y}(z) + \cancelto{0}{y[-1]}\right) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) - az^{-1}\mathbf{Y}(z) = \mathbf{X}(z)$$

$$(1 - az^{-1})\mathbf{Y}(z) = \mathbf{X}(z)$$

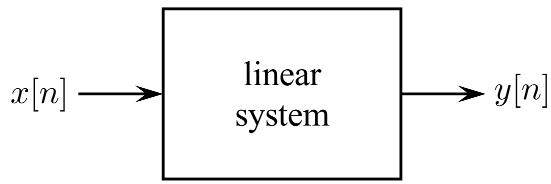
$$\mathbf{Y}(z) = \frac{1}{1-az^{-1}}\mathbf{X}(z)$$

Transfer function $\mathbf{H}(z)$

$$\mathbf{Y}(z) = \frac{1}{1-az^{-1}}\mathbf{X}(z) \quad \Rightarrow \quad y[n] = \sum_{m=0}^{\infty} a^m x[n-m]$$

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{X}(z) \quad \Rightarrow \quad y[n] = \sum_{m=\infty}^{\infty} h[m]x[n-m]$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

$$y[n] - ay[n-1] = x[n]$$

$$\mathbf{Y}(z) - a\left(z^{-1}\mathbf{Y}(z) + \cancelto{0}{y[-1]}\right) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) - az^{-1}\mathbf{Y}(z) = \mathbf{X}(z)$$

$$(1 - az^{-1})\mathbf{Y}(z) = \mathbf{X}(z)$$

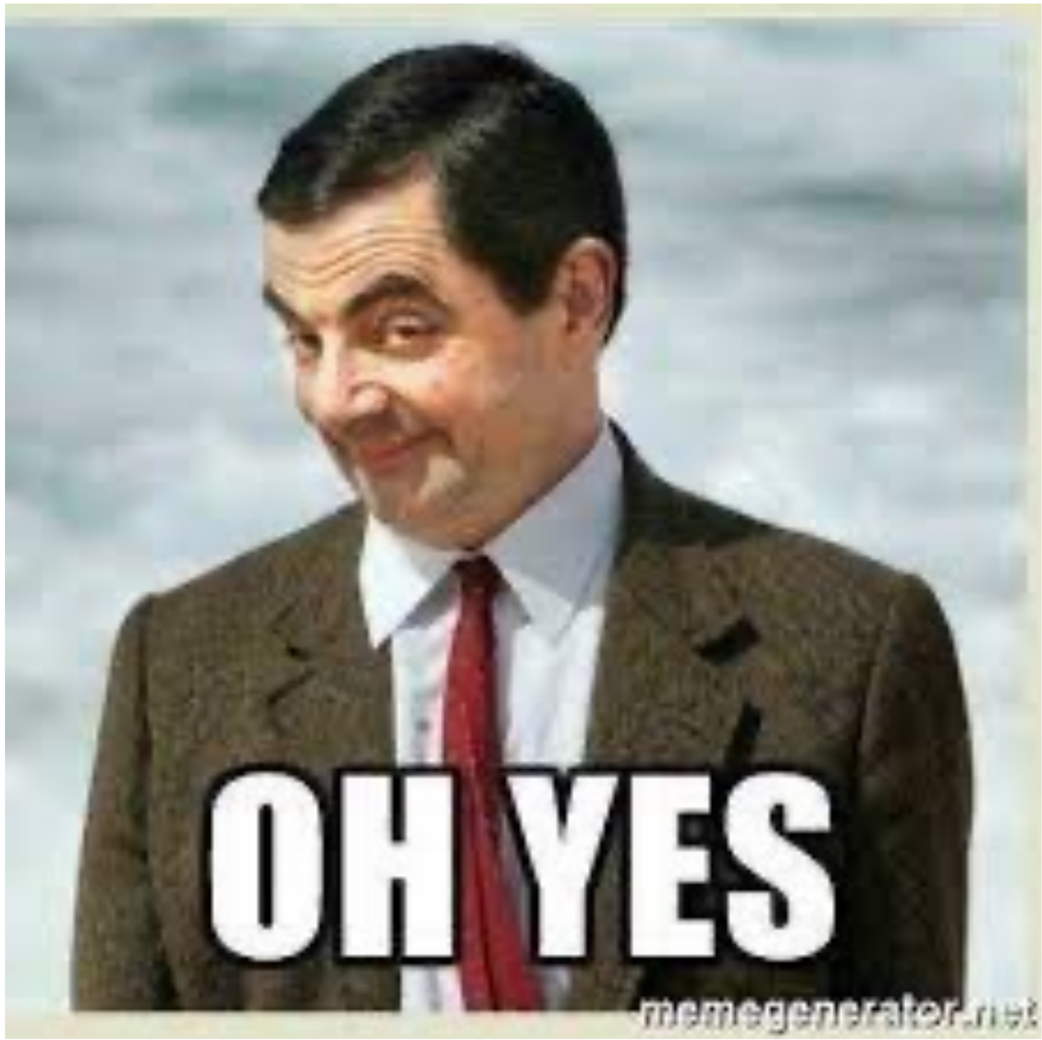
$$\mathbf{Y}(z) = \frac{1}{1 - az^{-1}}\mathbf{X}(z)$$

Transfer function $\mathbf{H}(z)$

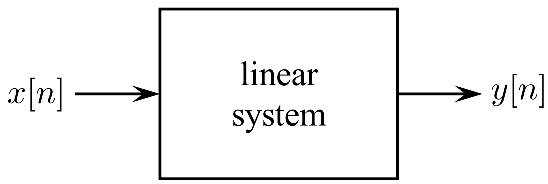
$$\mathbf{Y}(z) = \frac{1}{1 - az^{-1}}\mathbf{X}(z) \quad \Rightarrow \quad y[n] = \sum_{m=0}^{\infty} a^m x[n-m]$$

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{X}(z) \quad \Rightarrow \quad y[n] = \sum_{m=\infty}^{\infty} h[m]x[n-m]$$

Is this really the impulse response?


OH YES
memegenerator.net

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

$$y[n] - ay[n-1] = x[n]$$

$$\mathbf{Y}(z) - a\left(z^{-1}\mathbf{Y}(z) + \cancelto{0}{y[-1]}\right) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) - az^{-1}\mathbf{Y}(z) = \mathbf{X}(z)$$

$$(1 - az^{-1})\mathbf{Y}(z) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) = \frac{1}{1 - az^{-1}}\mathbf{X}(z)$$

$$x[n] = \delta[n] \quad \Rightarrow \quad \mathbf{X}(z) = 1$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

$$x[n] \longrightarrow \boxed{\begin{array}{c}\text{linear}\\ \text{system}\end{array}} \longrightarrow y[n]$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

$$\text{e.g., } y[n] - ay[n-1] = x[n]$$

$$y[n] - ay[n-1] = x[n]$$

$$\mathbf{Y}(z) - a\left(z^{-1}\mathbf{Y}(z) + \cancelto{0}{y[-1]}\right) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) - az^{-1}\mathbf{Y}(z) = \mathbf{X}(z)$$

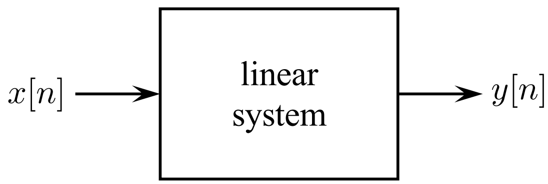
$$(1 - az^{-1})\mathbf{Y}(z) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) = \frac{1}{1 - az^{-1}}\mathbf{X}(z)$$

$$x[n] = \delta[n] \quad \Rightarrow \quad \mathbf{X}(z) = 1$$

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{X}(z) \quad \Rightarrow \quad \mathbf{Y}(z) = \mathbf{H}(z)$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

$$y[n] - ay[n-1] = x[n]$$

$$\mathbf{Y}(z) - a\left(z^{-1}\mathbf{Y}(z) + \cancelto{0}{y[-1]}\right) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) - az^{-1}\mathbf{Y}(z) = \mathbf{X}(z)$$

$$(1 - az^{-1})\mathbf{Y}(z) = \mathbf{X}(z)$$

$$\mathbf{Y}(z) = \frac{1}{1 - az^{-1}}\mathbf{X}(z)$$

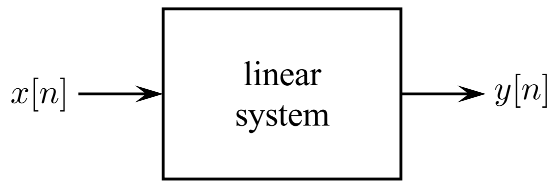
$$x[n] = \delta[n] \quad \Rightarrow \quad \mathbf{X}(z) = 1$$

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{X}(z) \quad \Rightarrow \quad \mathbf{Y}(z) = \mathbf{H}(z)$$

$$\mathbf{Y}(z) = \mathbf{H}(z) \quad \Rightarrow \quad y[n] = h[n]$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

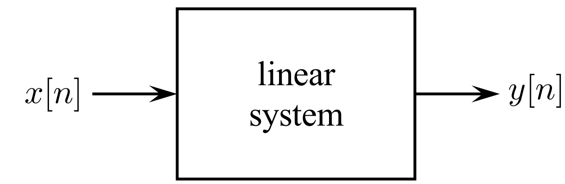


$$y[n] - ay[n-1] = x[n]$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

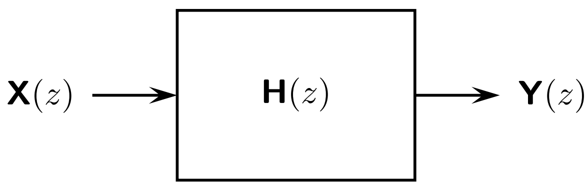
Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



$$y[n] - ay[n-1] = x[n]$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

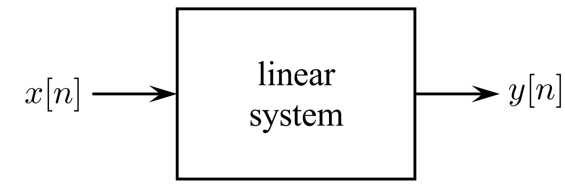
e.g., $y[n] - ay[n-1] = x[n]$



z domain transfer function

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{X}(z)$$

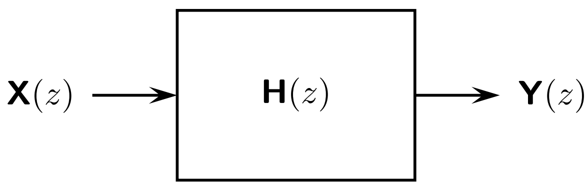
Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain



$$y[n] - ay[n-1] = x[n]$$

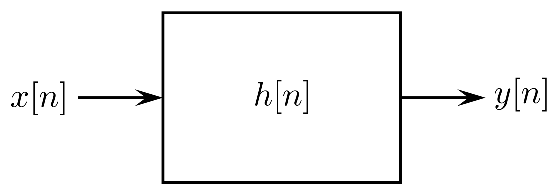
The linear systems of interest to us have an input/output relationship defined by LCCDE

$$\text{e.g., } y[n] - ay[n-1] = x[n]$$



z domain transfer function

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{X}(z)$$



Time domain convolution

$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain

$$x[n] \longrightarrow \boxed{\text{linear system}} \longrightarrow y[n]$$

$$y[n] - ay[n-1] = x[n]$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

$$\text{e.g., } y[n] - ay[n-1] = x[n]$$

$$\mathbf{X}(z) \longrightarrow \boxed{\mathbf{H}(z)} \longrightarrow \mathbf{Y}(z)$$

z domain transfer function

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{X}(z)$$

$$x[n] \longrightarrow \boxed{h[n]} \longrightarrow y[n]$$

Time domain convolution

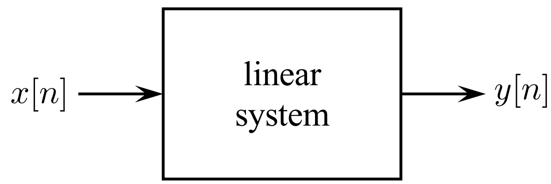
$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

To find the impulse response of an LTI system described by an LCCDE

1. Solve the LCCDE using the z transform with all-zeros initial conditions.
2. Write the z domain solution as $\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{X}(z)$.
3. Identify $\mathbf{H}(z)$ in the solution.
4. The impulse response $h[n]$ is the inverse Laplace transform of $\mathbf{H}(z)$.

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + DTFT

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + DTFT

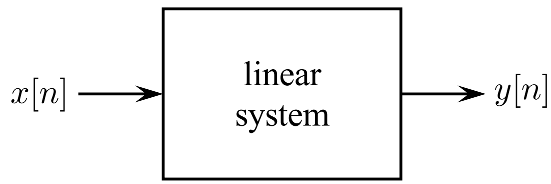


$$y[n] - ay[n-1] = x[n]$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

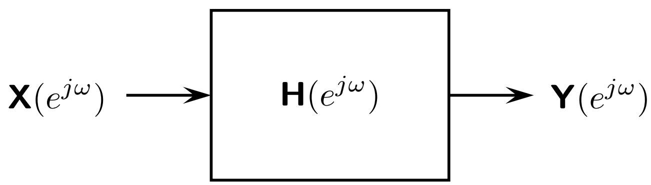
Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + DTFT



$$y[n] - ay[n-1] = x[n]$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

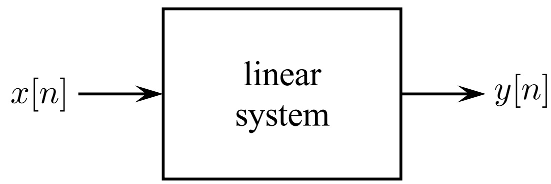
e.g., $y[n] - ay[n-1] = x[n]$



DTFT domain frequency response

$$\mathbf{Y}(e^{j\omega}) = \mathbf{H}(e^{j\omega})\mathbf{X}(e^{j\omega})$$

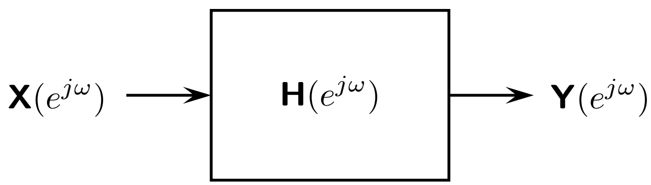
Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + DTFT



$$y[n] - ay[n-1] = x[n]$$

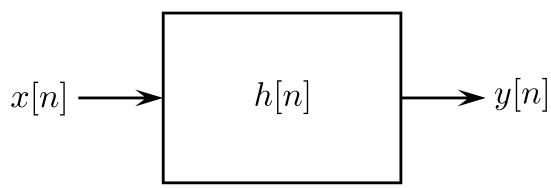
The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$



DTFT domain frequency response

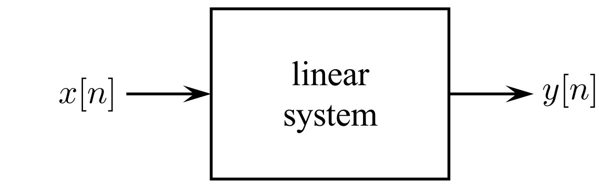
$$\mathbf{Y}(e^{j\omega}) = \mathbf{H}(e^{j\omega})\mathbf{X}(e^{j\omega})$$



Time domain convolution

$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

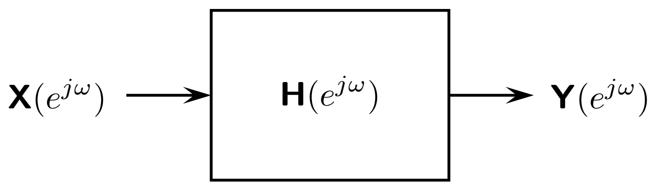
Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + DTFT



$$y[n] - ay[n-1] = x[n]$$

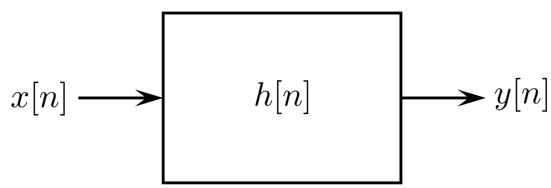
The linear systems of interest to us have an input/output relationship defined by LCCDE

$$\text{e.g., } y[n] - ay[n-1] = x[n]$$



DTFT domain frequency response

$$\mathbf{Y}(e^{j\omega}) = \mathbf{H}(e^{j\omega})\mathbf{X}(e^{j\omega})$$



Time domain convolution

$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

- One can solve an LCCDE using the DTFT with all-zeros initial conditions. But it is more common to use the z transform.

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + DTFT

$$x[n] \longrightarrow \boxed{\begin{array}{c}\text{linear}\\ \text{system}\end{array}} \longrightarrow y[n]$$

$$y[n] - ay[n-1] = x[n]$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

$$\mathbf{X}(e^{j\omega}) \longrightarrow \boxed{\mathbf{H}(e^{j\omega})} \longrightarrow \mathbf{Y}(e^{j\omega})$$

DTFT domain frequency response

$$\mathbf{Y}(e^{j\omega}) = \mathbf{H}(e^{j\omega})\mathbf{X}(e^{j\omega})$$

$$x[n] \longrightarrow \boxed{h[n]} \longrightarrow y[n]$$

Time domain convolution

$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

- One can solve an LCCDE using the DTFT with all-zeros initial conditions. But it is more common to use the z transform.
- After identifying the z domain transfer function $\mathbf{H}(z)$, the DTFT domain frequency response is

$$\mathbf{H}(e^{j\omega}) = \mathbf{H}(z)\Big|_{z=e^{j\omega}}$$

This works only as long as the ROC of $\mathbf{H}(z)$ contains the unit circle ($z = e^{j\omega}$).

Relationships: LCCDE, Impulse Response, Convolution, Frequency Domain + DTFT

$$x[n] \longrightarrow \boxed{\text{linear system}} \longrightarrow y[n]$$

$$y[n] - ay[n-1] = x[n]$$

The linear systems of interest to us have an input/output relationship defined by LCCDE

e.g., $y[n] - ay[n-1] = x[n]$

$$\mathbf{X}(e^{j\omega}) \longrightarrow \boxed{\mathbf{H}(e^{j\omega})} \longrightarrow \mathbf{Y}(e^{j\omega})$$

DTFT domain frequency response

$$\mathbf{Y}(e^{j\omega}) = \mathbf{H}(e^{j\omega})\mathbf{X}(e^{j\omega})$$

$$x[n] \longrightarrow \boxed{h[n]} \longrightarrow y[n]$$

Time domain convolution

$$y[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

- One can solve an LCCDE using the DTFT with all-zeros initial conditions. But it is more common to use the z transform.
- After identifying the z domain transfer function $\mathbf{H}(z)$, the DTFT domain frequency response is
$$\mathbf{H}(e^{j\omega}) = \mathbf{H}(z)\Big|_{z=e^{j\omega}}$$
This works only as long as the ROC of $\mathbf{H}(z)$ contains the unit circle ($z = e^{j\omega}$).
- For stable systems, this is always true.

$$S_{NN}(f) = \frac{hf}{e^{hf/(k_B \mathcal{T})} - 1}$$

$$\begin{aligned} h \quad \text{(Planck's constant)} &= 6.6261 \times 10^{-34} \quad \text{Js} \\ k_B \quad \text{(Boltzmann's constant)} &= 1.3807 \times 10^{-23} \quad \text{J/K} \\ \mathcal{T} &= \text{temperature in Kelvin} \\ f &= \text{frequency in Hz (cycles/s)} \end{aligned}$$