Vibration Theory, Vol. 5

introduction to stochastic process theory

Nielsen, Søren R. K.

Publication date:
1996

Document Version
Accepted author manuscript, peer reviewed version

Citation for published version (APA):
Nielsen, S. R. K. (1996). *Vibration Theory, Vol. 5: introduction to stochastic process theory*. Dept. of Building Technology and Structural Engineering, Aalborg University. U/ No. 9603

VIBRATION THEORY, VOL. 5

Introduction to
Stochastic Process Theory

Søren R. K. Nielsen
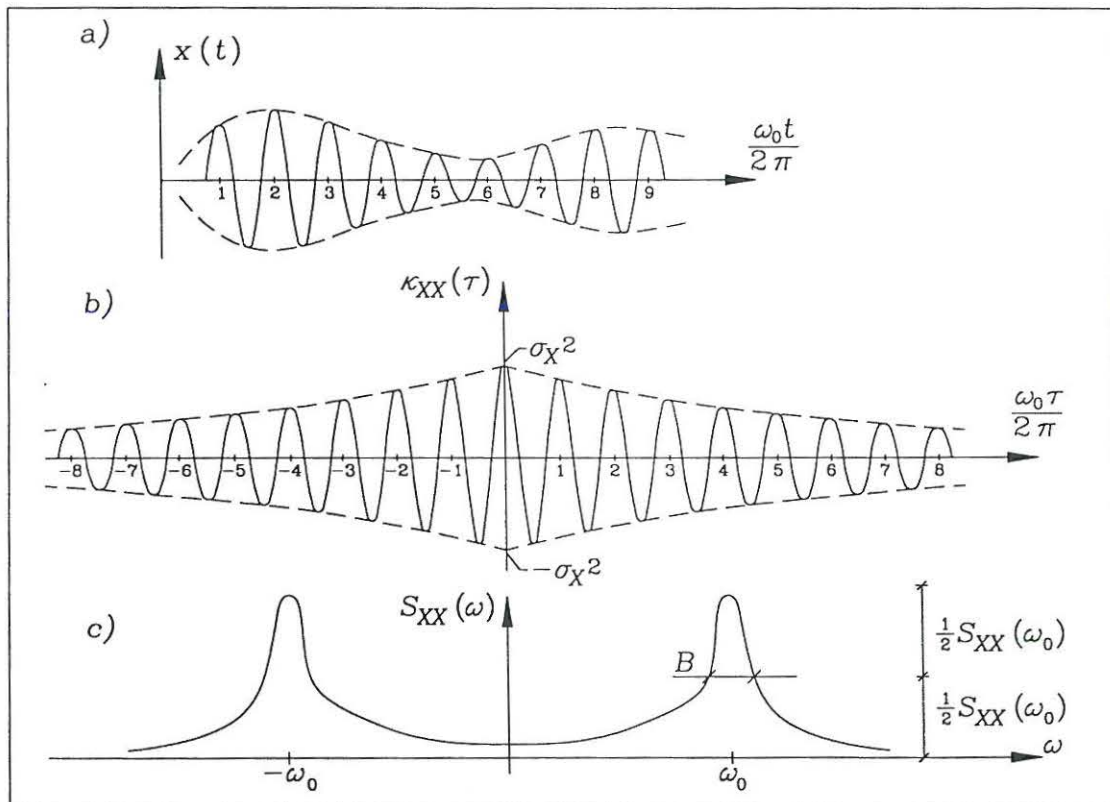
TABLE OF CONTENTS

1. REVIEW OF PROBABILITY THEORY

1.1 Stochastic variables

A probability space is a triple (Ω, σ, P), where Ω is a *sample space*, σ is a *sigma algebra* and P is a *probability measure*. A sample space Ω is a set of all probable outcomes of an experiment. The elements $\omega \in \Omega$, the sample space, are named *elementary events* or *sample points*. A set of elementary events is named an *event*, i.e. $A \subseteq \Omega$. The empty set $\emptyset$ is named the *impossible event*, and Ω itself is named the *safe event*. A sigma algebra, σ, of the sample space Ω is a nonempty set of events (i.e. a set of sets) of Ω with the property

$$1.\ A \in \sigma \ \Rightarrow \ A^c \in \sigma \tag{1.1}$$

$$2.\ A_1, ..., A_n \in \sigma \ \Rightarrow \ \cup_{i=1}^{n} A_i \in \sigma \tag{1.2}$$

Eq. (1.1) claims that if the event $A \subset \Omega$ is included in Ω, so is the complement, A^c. Eq. (1.2) says that if a finite sequence of events $A_1, ..., A_n$ are all in Ω, so is the union of the events. Actually, eq. (1.2) is claimed to be valid even if a *countable* infinite number of events is considered. Eqs. (1.1) and (1.2) are sufficient to prove that σ is closed under all conventional set operations.

Example 1.1: Intersection of events is an event

Let $A \in \sigma \wedge B \in \sigma$. From De Morgan's law it then follows

$$(A \cap B)^c = A^c \cup B^c \tag{1.3}$$

The right-hand side is an event from the axioms (1.1) and (1.2). So, the left-hand side is an event. From eq. (1.1) follows that its complement $\left((A \cap B)^c\right)^c = A \cap B$ is an event too.

A probability measure, $P : \sigma \to [0, 1]$, is a function which relates each element of σ (each event) with a number in the closed interval from 0 to 1. More precisely, the following axioms are to be fulfilled by a legal probability measure

$$1.\ \forall A \in \sigma : P(A) \in [0, 1] \tag{1.4}$$

$$2.\ P(\Omega) = 1 \tag{1.5}$$

$$3.\ \text{If } A_1, A_2, ..., A_n \text{ are mutually disjoint then } \left(\cup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) \tag{1.6}$$

Events $A, B \in \sigma$ are mutually disjoint, if their intersection $A \cap B = \emptyset$. Eq. (1.6) is requested to be valid even for countable infinite many disjoint events.

A stochastic variable $X : \Omega \to R$ is a function with the property, that

$$\forall x \in R : \{\omega \in \Omega \mid X(\omega) \le x\} \in \boldsymbol{\sigma} \tag{1.7}$$
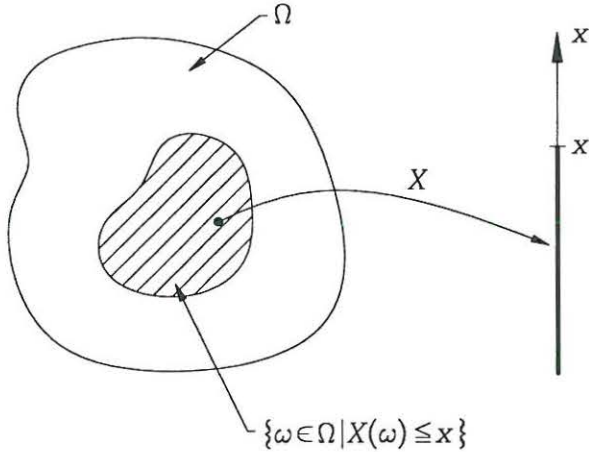

Figure 1.1: Definition of a stochastic variable.

For each $\omega \in \Omega$, the stochastic variable defines a real value $X(\omega) \in R$. For some x consider the semi-infinite interval $]-\infty, x]$. $\{\omega \in \Omega \mid X(\omega) \le x\}$ indicates the elementary events, which are mapped into the said interval. This set can always be defined, but need not be included in the sigma algebra $\boldsymbol{\sigma}$. However, if the function $X(\omega)$ qualifies as a random variable, $\{\omega \in \Omega \mid X(\omega) \le x\}$ is indeed an event (a member of $\boldsymbol{\sigma}$) no matter which semi-infinite (no matter which x) is considered. The event $\{\omega \in \Omega \mid X(\omega) \le x\}$ can then be related with a certain probability, which is denoted the *probability distribution function* $F_X(x)$ of the stochastic variable X defined as

$$F_X(x) = P\big(\{\omega \in \Omega \mid X(\omega) \le x\}\big) \equiv P(X \le x) \tag{1.8}$$

The name, probability distribution function, indicates that the probability is a function of the considered semi-infinite interval as defined by x. Set signs around events, and the argument, ω, of the random variable will usually be omitted in what follows, as shown by the abbreviated notation in the last statement of eq. (1.8). Further, stochastic variables are designated with capital letters, whereas the sample values are indicated with lower case letters.

Notice that

$$\{\omega \in \Omega \mid X(\omega) \le -\infty\} = \emptyset \tag{1.9}$$

$$\{\omega \in \Omega \mid X(\omega) \le \infty\} = \Omega \tag{1.10}$$

$$\{\omega \in \Omega \mid X(\omega) \le b\} = \{\omega \in \Omega \mid X(\omega) \le a\} \cup \{\omega \in \Omega \mid a < X(\omega) \le b\} \tag{1.11}$$

where the latter events in eq. (1.11) are mutually disjoint. From eq.(1.6) follows

$$F_X(-\infty) = 0 \tag{1.12}$$

$$F_X(+\infty) = 1 \tag{1.13}$$

$$P(a < X \le b) = F_X(b) - F_X(a) \tag{1.14}$$

The range $\Omega_X = X(\omega)$ may either be discrete (either finite or countable infinite) or continous. Correspondingly, X is named a *discrete* or a *continous stochastic variable*.
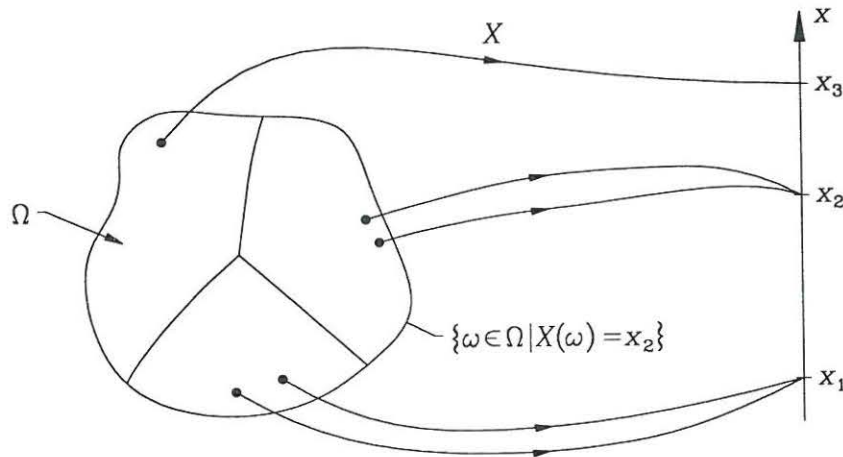
Figure 1.2: Discrete stochastic variable.

The probability function of a discrete stochastic variable is defined from

$$P_X(x_i) = P\big(\{\omega \in \Omega \mid X(\omega) = x_i\}\big) \quad , \quad i = 1, 2 \ldots \tag{1.15}$$

Below in table 1.1 some well-known probability functions for discrete stochastic variables are listed

Distribution	Probability function	Parameters	Short notation
Binomial	$P_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, 2, \ldots$	n : Number of trials p : Probability of success	$X \sim$ $B(n, p)$
Geometric	$P_X(x) = p(1-p)^{x-1}$ $x = 1, 2, \ldots$	p: Probability of success	$X \sim$ $Ge(p)$
Hyper-geometric	$P_X(x) = \dfrac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$ $x = 0, 1, 2, \ldots$	N : Total population n : Number drawn m : Favourable objects	$X \sim$ $H(N, n, m)$
Poisson	$P_X(x) = \exp(-\lambda)\frac{\lambda^n}{x!}$ $x = 0, 1, 2, \ldots$	λ: Mean number of successes	$X \sim$ $Po(\lambda)$
Pascal	$P_X(x) = \binom{n+x-1}{x} p^x (1-p)^x$ $x = 0, 1, 2, \ldots$	n : Number of successes p : Probability of success	$X \sim$ $Pa(n, s)$

Table 1.1: Probability functions for discrete stochastic variables.
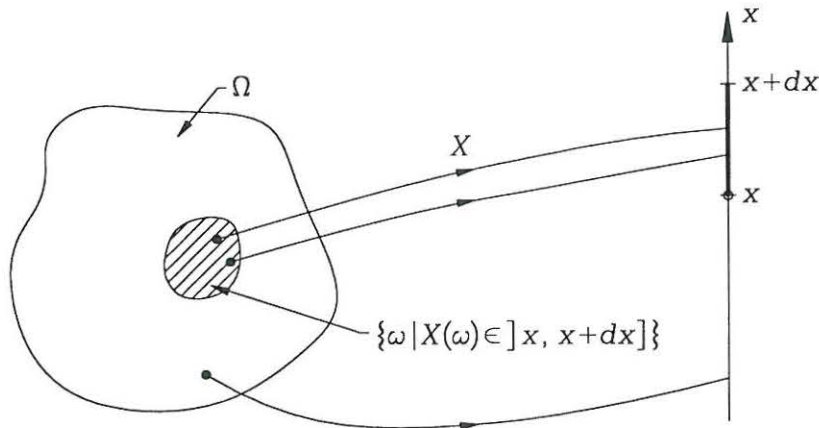
Figure 1.3: Continuous stochastic variable.

Assume that X is a continuous stochastic variable. The probability of having samples in the interval $]x, x + dx]$ follows from eq. (1.14)

$$P(X \in]x, x + dx]) = F_X(x + dx) - F_X(x) = \frac{dF_X(x)}{dx} dx \tag{1.16}$$

The *probability density function* of the stochastic variable is defined by

$$f_X(x) = \frac{dF_X(x)}{dx} \tag{1.17}$$

A continous stochastic variable is completely specified by its probability density function (pdf). Below in table 1.2 some well-known pdf's for continuous stochastic variables are listed.

Distribution	Density function	Short notation
Uniform	$f_X(x) = \begin{cases} 0 & , \quad x < a \\ \frac{1}{b-a} & , \quad a \le x \le b \\ 0 & , \quad x > b \end{cases}$	$X \sim U(a, b)$
Normal	$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) \quad , \quad x \in R$	$X \sim N(\mu, \sigma^2)$
Gamma	$f_X(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)} x^\alpha \exp(-\beta x) & , \quad x \ge 0 \end{cases}$	$X \sim Ga(\alpha, \beta)$
Exponential	$f_X(x) = \begin{cases} 0 & , \quad x < 0 \\ \beta \exp(-\beta x) & , \quad x \ge 0 \end{cases}$	$X \sim E(\beta)$
Rayleigh	$f_X(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) & , \quad x \ge 0 \end{cases}$	$X \sim R(\sigma^2)$
Weibull	$f_X(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{\alpha}{x_1}\left(\frac{x-x_0}{x_1}\right)^{\alpha-1} \exp\left(-\left(\frac{x-x_0}{x_1}\right)^\alpha\right) & , \quad x \ge 0 \end{cases}$	$X \sim W(x_0, x_1, \alpha)$

Table 1.2: Density functions for continuous stochastic variables.

In order to introduce the probability density function of a discrete stochastic variable the so-called Dirac's delta function must be defined.
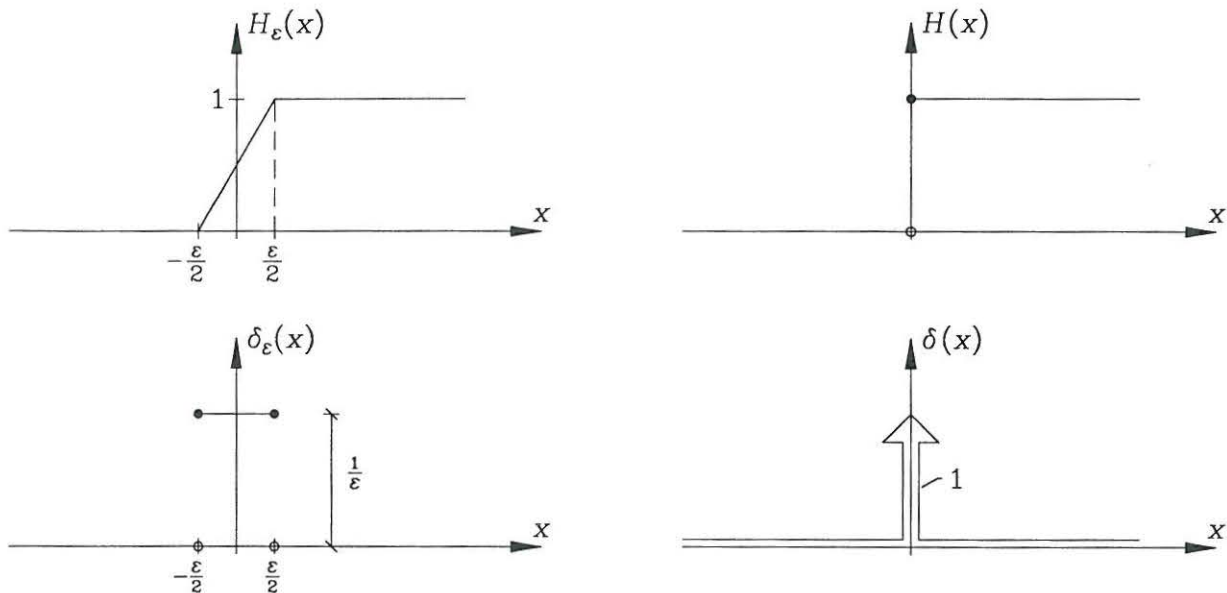
Figure 1.4: a) Ramp function. b) Derivative of ramp function.

Consider the ramp function $H_\varepsilon(x)$ shown in figure 1.4a. It's derivative $\delta_\varepsilon(x)$ as shown in figure 1.4b is characterized by being 0 outside the interval $x \in [-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}]$, where the value $\frac{1}{\varepsilon}$ is attained. The area below the graph is obviously 1 for all ε. Dirac's delta function can then be visualized as the limit of $\delta_\varepsilon(x)$ as $\varepsilon \to 0$, preserving the area 1, but with infinite value at $x = 0$. The limit of the ramp function $H_\varepsilon(x)$ as $\varepsilon \to 0$ is named the Heaviside's unit step function $H(x)$ defined as

$$H(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \tag{1.18}$$

Formally, Dirac's delta function can be thought of as the derivative of the Heaviside's unit step function. This interpretation is usual in engineering texts, but may cause winces in the face of a mathematician.

Let $f(x)$ be continuous at the point $x = x_0$. Then

$$\int_{-\infty}^{\infty} \delta(x_0 - x)f(x)dx = \lim_{\varepsilon \to 0} \int_{x_0-\varepsilon/2}^{x_0+\varepsilon/2} \delta_\varepsilon(x_0 - x)f(x)dx =$$

$$\lim_{\varepsilon \to 0} f(x_0) \int_{x_0-\varepsilon/2}^{x_0+\varepsilon/2} \frac{1}{\varepsilon}dx = f(x_0) \tag{1.19}$$

Eq. (1.19) states the fundamental property of the Dirac's delta function.

In fig. 1.5 the probability distribution function of a discrete stochastic variable is shown. It appears as a step function, which is right continuous (i.e. $F_X(x_i) = F_X(x_i+\varepsilon)$, $\varepsilon > 0$). The probability function $P_X(x_i)$ indicates the magnitude of the jumps. Formally, the probability density function of such a step function appears as a sequence of Dirac's delta-spikes with intensity $P_X(x_i)$ as shown in fig. 1.5.

6

In fig. 1.6 the corresponding probability distribution and density functions of a continuous stochastic variable are shown.

Generally, the sample space consists of discrete and continuous parts. In fig. 1.7 the probability distribution function and the probability density function of the so-called *mixed type* are shown. The probability $P_X(x_i)$ of obtaining the discrete sample x_i appears as a jump in $F_X(x)$ and as a delta-spike in $f_X(x)$. Again, the probability distribution function is right continuous at the discontinuity points.
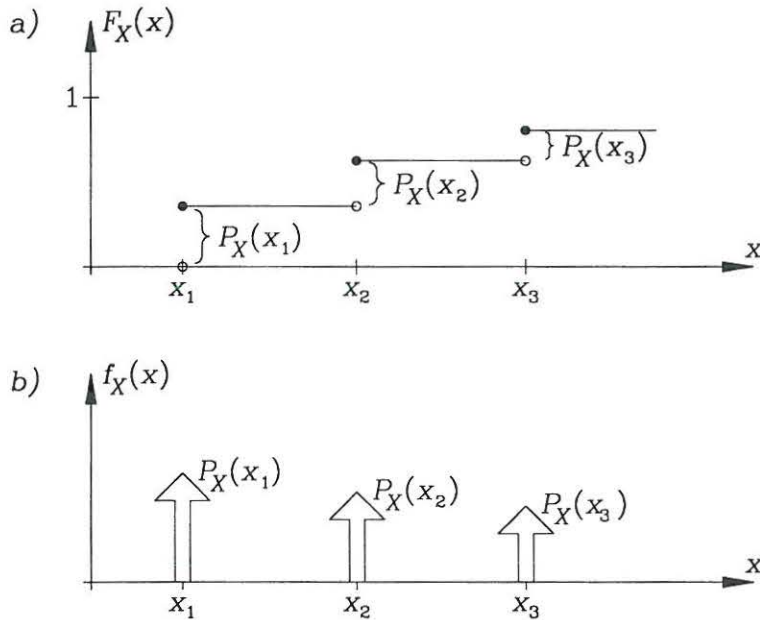
Figure 1.5: Probability distribution function and probability density function for a discrete stochastic variable.
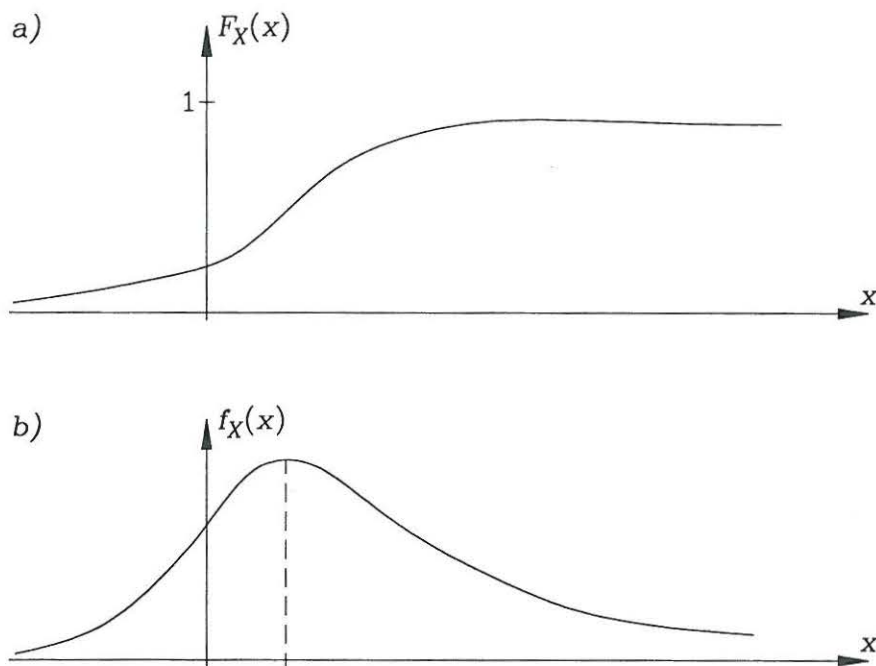
Figure 1.6: Probability distribution function and probability density function for a continuous variable.
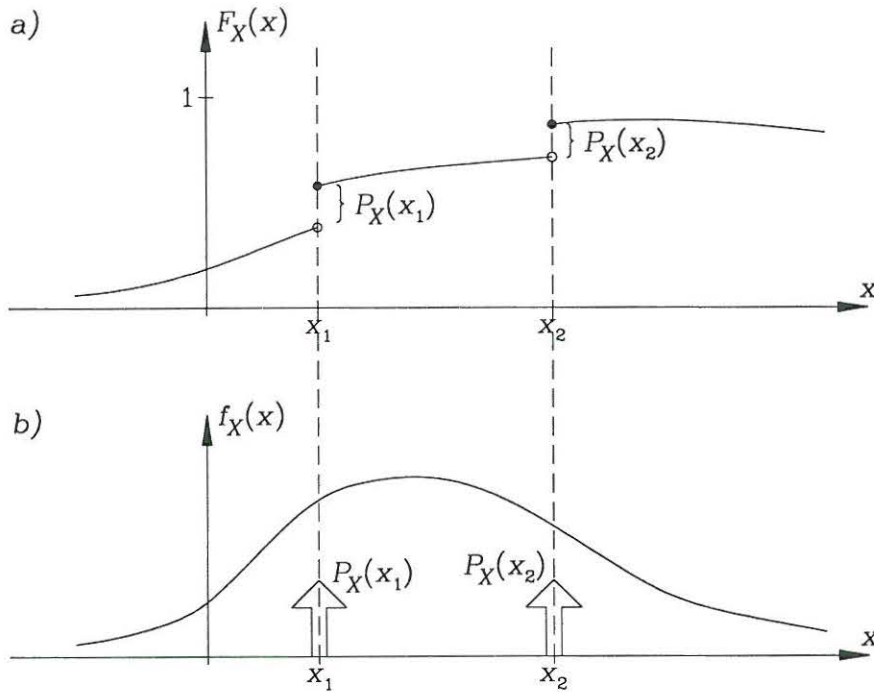
Figure 1.7: Probability distribution function and probability density function of a stochastic variable of mixed type.
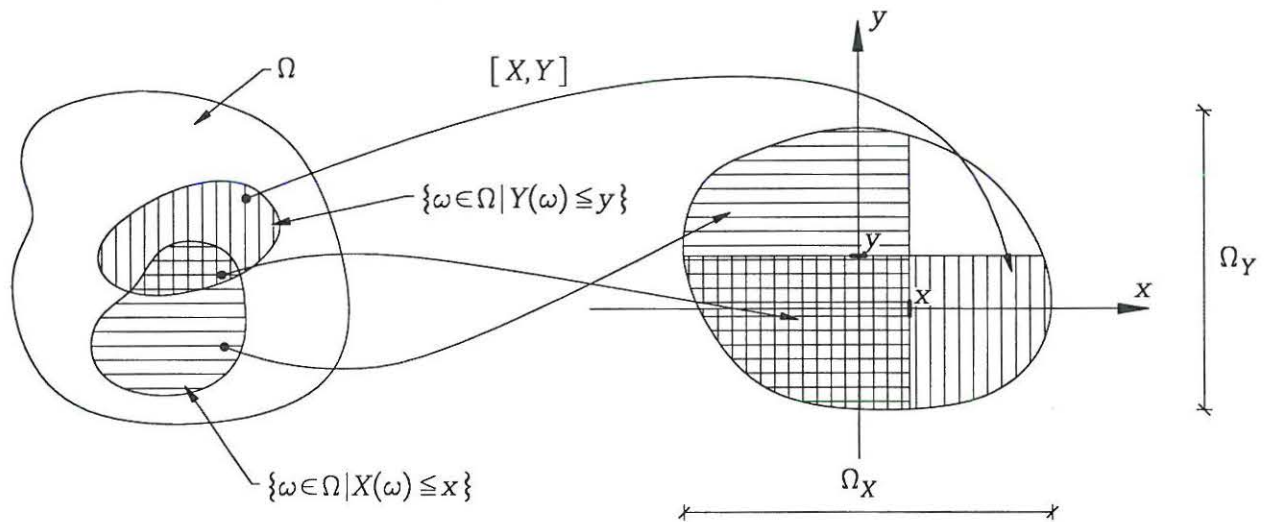
Figure 1.8: 2-dimensional stochastic variable.

Let X and Y be stochastic variables defined on the same sample space Ω. The pair is termed a *2-dimensional stochastic variable*. Consider the events $\{\omega \in \Omega \mid X(\omega) \leq x\} \in \boldsymbol{\sigma}$ and $\{\omega \in \Omega \mid Y(\omega) \leq y\} \in \boldsymbol{\sigma}$. Then, the intersection $\{\omega \in \Omega \mid X(\omega) \leq x\} \in \boldsymbol{\sigma} \cap \{\omega \in \Omega \mid Y(\omega) \leq y\} \in \boldsymbol{\sigma}$ is also an event for which a probability is defined for arbitraty x and y, see fig. 1.8.

The probability distribution, $F_{XY}(x,y)$, of the 2-dimensional stochastic variable is de-

8

fined from

$$F_{XY}(x, y) = P(\{\omega \in \Omega \mid X(\omega) \leq x\} \in \boldsymbol{\sigma} \cap \{\omega \in \Omega \mid Y(\omega) \leq y\} \in \boldsymbol{\sigma})$$

$$\equiv P(X \leq x \land Y \leq y) \tag{1.20}$$

Since $\{\omega \in \Omega \mid X(\omega) < \infty\} = \{\omega \in \Omega \mid Y(\omega) < \infty\} = \Omega$ and $\{\omega \in \Omega \mid X(\omega) < -\infty\} = \{\omega \in \Omega \mid Y(\omega) < -\infty\} = \emptyset$, it follows

$$F_{XY}(x, \infty) = P(\{X \leq x\} \cap \Omega) = P(\{X \leq x\}) = F_X(x) \tag{1.21}$$

$$F_{XY}(\infty, y) = P(\Omega \cap \{Y \leq y\}) = P(\{Y \leq y\}) = F_Y(y) \tag{1.22}$$

$$F_{XY}(\infty, \infty) = P(\Omega \cap \Omega) = P(\Omega) = 1 \tag{1.23}$$

$$F_{XY}(-\infty, y) = P(\emptyset \cap \{Y \leq y\}) = P(\emptyset) = 0 \tag{1.24}$$

$$F_{XY}(x, -\infty) = P(\{Y \leq y\} \cap \emptyset) = P(\emptyset) = 0 \tag{1.25}$$

The range Ω_{XY} of the 2-dimensional stochastic variable may either be discrete (finite or countable infinite number of points $(x_i, y_i) \in R^2$) or continuous. In the latter case Ω_{XY} forms a subset of R^2. Correspondingly, $[X, Y]$ is named a discrete or a continuous stochastic variable. $[X, Y]$ is discrete if both X and Y are discrete, continuous if both X and Y are continuous. Else, $[X, Y]$ is of the mixed type.

The probability function of a discrete stochastic variable is defined from

$$P_{XY}(x_i, y_i) = P(\{\omega \in \Omega \mid X(\omega) = x_i \land Y(\omega) = y_i\}) \tag{1.26}$$

The probability density f_{XY} of a continuous 2-dimensional stochastic variable is defined from the mixed derivative

$$f_{XY}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x, y) \tag{1.27}$$

$f_{XY}(x, y) \, dx dy$ signifies the probability of the event $\{\omega \in \Omega \mid x < X(\omega) \leq x + dx\} \cap \{\omega \in \Omega \mid y < Y(\omega) \leq y + dy\}$. Hence, $f_{XY}(x, y)$ is a non-negative function of x and y. The probability that $[X, Y]$ is in any subset of Ω_{XY} is obtained by integrating $f_{XY}(x, y)$ over that $[X, Y]$ is in any subset of Ω_{XY} i obtained by integrating $f_{XY}(x, y)$ over that subset. As an example the following probability becomes

$$P(a_1 < X \leq b_1 \land a_2 < Y \leq b_2) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f_{XY}(x, y) \, dy dx \tag{1.28}$$
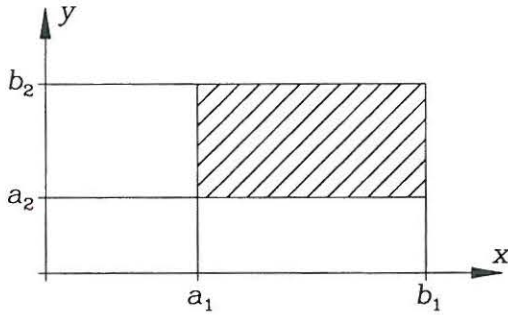
Figure 1.9: Evaluation of integral by means of probability density function.

From eqs. (1.21–1.25) and eq. (1.27) follow, see fig. 1.9.

$$F_{XY}(x,y) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f_{XY}(x,y)\,dy\,dx \qquad (1.29)$$

The integral (1.28) can be evaluated in terms of the probability distribution function as follows

$$P(a_1 < X \le b_1 \wedge a_2 < Y \le b_2) = \int\limits_{a_1}^{b_1} \int\limits_{a_2}^{b_2} f_{XY}(x,y)\,dy\,dx =$$

$$\int\limits_{-\infty}^{b_1} \int\limits_{a_2}^{b_2} f_{XY}(x,y)\,dy\,dx - \int\limits_{-\infty}^{a_1} \int\limits_{a_2}^{b_2} f_{XY}(x,y)\,dy\,dx =$$

$$\int\limits_{-\infty}^{b_1} \int\limits_{-\infty}^{b_2} f_{XY}(x,y)\,dy\,dx - \int\limits_{-\infty}^{b_1} \int\limits_{-\infty}^{a_2} f_{XY}(x,y)\,dy\,dx$$

$$-\int_{-\infty}^{a_1} \int_{-\infty}^{b_2} f_{XY}(x,y)\,dy\,dx + \int_{-\infty}^{a_1} \int_{-\infty}^{a_2} f_{XY}(x,y)\,dy\,dx =$$

$$F_{XY}(b_1,b_2) - F_{XY}(b_1,a_2) - F_{XY}(a_1,b_2) + F_{XY}(a_1,a_2) \qquad (1.30)$$

The events $A \in \boldsymbol{\sigma}$ and $B \in \boldsymbol{\sigma}$ are said to be independent if

$$P(A \cap B) = P(A)\,P(B) \qquad (1.31)$$

Now, let $A = \{\omega \in \Omega \mid X(\omega) \in]x, x+dx]\}$ and $B = \{\omega \in \Omega \mid Y(\omega) \in]y, y+dy]\}$. The stochastic variables X og Y are said to be mutually independent, if these events are independent for any x and y. Obviously, this is the case only if

$$f_{XY}(x, y) = f_X(x) f_Y(y) \tag{1.32}$$

Let $X_1, X_2, ...X_n$ be a set of stochastic variables all defined on the same sample space Ω. Then $[X_1, X_2, \ldots, X_n]$ is termed a *n-dimensional stochastic variable* or a *n*-dimensional *stochastic vector*. As a straightforward generalization of the 2-dimensional case, the distribution function of $[X_1, X_2, \ldots, X_n]$ is

$$F_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n) = P(\{X_1 \le x_1\} \cap \{X_2 \le x_2\} \cap \cdots \cap \{X_n \le x_n\})$$

$$\equiv P(X_1 \le x_1 \wedge X_2 \le x_2 \wedge \cdots \wedge X_n \le x_n) \tag{1.33}$$

Depending on whether the range $\Omega_{X_1 X_2 \ldots X_n}$ is discrete or continuous, $[X_1, X_2, \ldots, X_n]$ is said to be a discrete or a continuous n-dimensional stochastic variable. The probability function of a n-dimensional stochastic variable is defined from

$$P_{X_1 X_2 \ldots X_n}(x_{i_1}, x_{i_2}, \ldots, x_{i_n}) = P(X_1 = x_{i_1} \wedge X_2 = x_{i_2} \wedge \cdots \wedge X_n = x_{i_n}) \tag{1.34}$$

If the range is continuous the probability density function of $[X_1, X_2, \ldots, X_n]$ is defined as the nth order mixed derivative

$$f_{X_1 X_2 \ldots X_n} = \frac{\partial^n}{\partial x_1 \partial x_2 \ldots \partial x_n} F_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n) \tag{1.35}$$

Using the vector notation $\mathbf{X}^T = [X_1, X_2, \ldots, X_n]$ and $f_{\mathbf{X}}(\mathbf{x}) = f_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n)$ the probability that a continuous n-dimensional stochastic variable $\mathbf{X}$ is in an arbitrary domain $D \subseteq \Omega_{\mathbf{X}}$ can be written

$$P(\mathbf{X} \in D) = \int_D f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x} \tag{1.36}$$

The vector $[X_1, X_2, \ldots, X_n]$ is said to be a n-dimensional normally distributed stochastic variable, in short notation $\mathbf{X} \sim N(\boldsymbol{\mu}, \mathbf{C})$, if

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det(\mathbf{C}))^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \mathbf{C}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right) \tag{1.37}$$

where $\boldsymbol{\mu}$ is an n-dimensional vector and $\mathbf{C}$ is a symmetric positive definite matrix. $\det(\mathbf{C})$ is the determinant of $\mathbf{C}$ which is always positive.

Example 1.2: Two-dimensional normally distributed stochastic variable

For a two-dimensional normally distributed variable, one may write

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma 2 \end{bmatrix} \tag{1.38}$$

The matrix $\mathbf{C}$ will be positive definit, if $\rho^2 < 1$. Then eq. (1.37) attains the form

$$f_{X_1 X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{\xi_1^2 - 2\rho\xi_1\xi_2 + \xi_2^2}{2(1-\rho^2)}\right) \tag{1.39}$$

where

$$\xi_i = \frac{x_i - \mu_i}{\sigma_i}, \quad i = 1, 2 \tag{1.40}$$

1.2 Combined stochastic variables
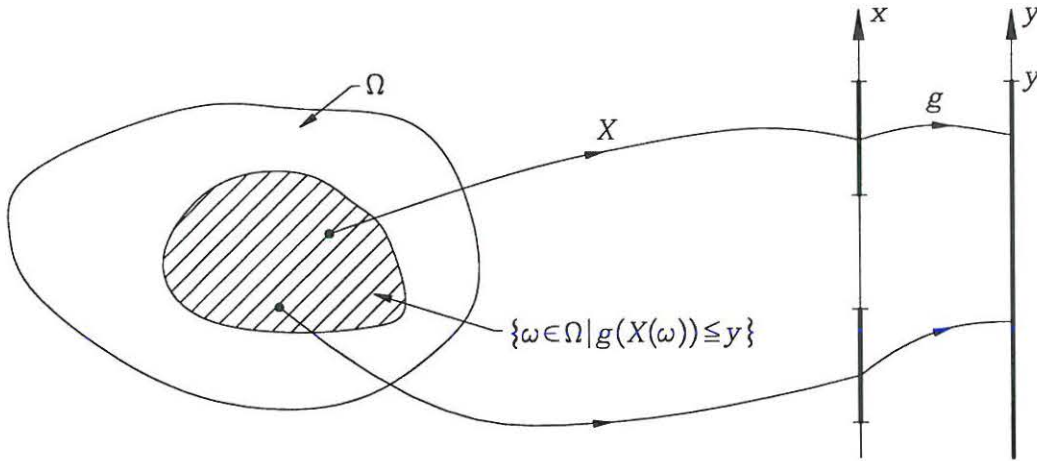
Figure 1.10: Combined stochastic variable.

Consider the real function $g : R \to R$. The combined mapping $Y : \Omega \to R, : Y(\omega) = g(X(\omega))$ can then be defined. Per definition Y is a stochastic variable, if

$$\forall y \in R : \{\omega \in \Omega \mid g(X(\omega)) \leq y\} \in \boldsymbol{\sigma} \tag{1.41}$$

Eq. (1.41) cannot be true for arbitrary real functions $g(x)$. However, the class of functions (the so-called Borel measurable functions) for which it is true is extremely large. Without any practical restrictions, eq. (1.41) can then be assumed to be fulfilled. Y is called a *combined stochastic variable*.

Assuming the distribution function $F_X(x)$ or the probability density function $f_X(x)$ of X is known, one may want to know the distribution function $F_Y(y)$ or the probability density function $f_Y(y)$ of the combined stochastic variable. In the simplest case we may assume $g(x)$ to be a differentiable monotonous increasing or decreasing function of x, and that X is a continuous stochastic variable. The probability of $\{Y \in]y, y + dy]\}$ is then equal to the probability of $\{X \in]x, x + dx]\}$, where $y = g(x)$. Hence,

$$f_Y(y)\,dy = f_X(x)\,dx \quad \Leftrightarrow$$

$$f_Y(y) = f_X(x)\frac{1}{\mid g'(x)\mid} = f_X(g^{-1}(y))\frac{1}{\mid g'(g^{-1}(y))\mid} \tag{1.42}$$

where $g'(x) = \frac{dg(x)}{dx}$ and $x = g^{-1}(y)$ signifiy the inverse function. If $g(x)$ is not monotonous, the probability distribution function $F_Y(y)$ can be evaluated from the basic definition

$$F_Y(y) = P(Y \leq y) = P(g(x) \leq y) = \int\limits_{\{x \in \Omega_X \mid g(x) \leq y\}} f_X(x)\,dx \tag{1.43}$$

where the integral is evaluated over the set $\{x \in \Omega_X \mid g(x) \leq y\}$, i.e. these points of the sample space where $g(x) \leq y$.

Example 1.3: Distribution of the square of a stochastic variable

Let $Y = g(X) = X^2$. Eq. (1.43) then provides for the probability distribution function of Y

$$\begin{aligned}
F_Y(y) &= P\left(X^2 \leq y\right) \\
&= \begin{cases} 0 & ,y < 0 \\ P\left(-\sqrt{y} \leq X \leq \sqrt{y}\right) & ,y \geq 0 \end{cases} \\
&= \begin{cases} 0 & ,y < 0 \\ F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right) & ,y \geq 0 \end{cases}
\end{aligned} \tag{1.44}$$

The probability density function of Y then becomes

$$f_Y(y) = \begin{cases} 0 & ,y \leq 0 \\ \frac{1}{2\sqrt{y}}(f_X(\sqrt{y}) + f_X(-\sqrt{y})) & ,y > 0 \end{cases} \tag{1.45}$$

Example 1.4: Inverse method
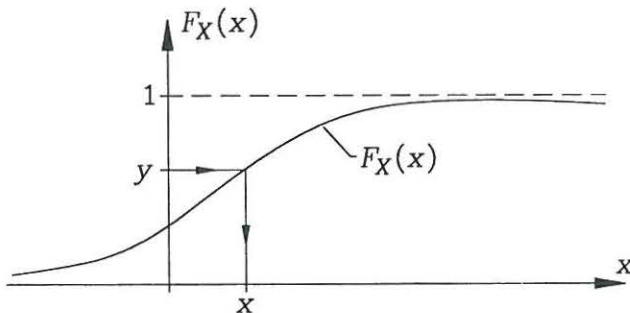
Figure 1.11: Illustration of the inverse method

Let X be a continuous stochastic variable and define $Y = F_X(X)$, i.e. $g(x)$ has been chosen as the monotonous increasing function $F_X(x)$. From eq. (1.41) it then follows for the probability distribution function of Y

$$F_Y(y) = P(Y \leq y) = P(F_X(X) \leq y)$$

$$= \begin{cases} 0 & , y \in]-\infty, 0[\\ P(X \leq F_X^{-1}(y)) & , y \in [0, 1] \\ 1 & , y \in]1, \infty[\end{cases}$$

$$= \begin{cases} 0 & , y \in]-\infty, 0[\\ F_X(F_X^{-1}(y)) = y & , y \in [0, 1] \\ 1 & , y \in]1, \infty[\end{cases} \tag{1.46}$$

The probability density function then becomes

$$f_Y(y) = \begin{cases} 0 & , y \in]-\infty, 0[\,\cup\,]1, \infty[\\ 1 & , y \in [0, 1] \end{cases} \tag{1.48}$$

Hence, $Y \sim U(0, 1)$. This observation forms the basis of the so-called *inverse method* for generating samples x of a stochastic variable on a computer. A sample of $Y \sim U(0, 1)$ is generated by a so-called pseudo random number generator, and the realizations of X are then calculated from $x = F_X^{-1}(y)$. The principle is illustrated in fig. (1.11).

Example 1.5: Log-normally distributed stochastic variable

Let $X \sim N(\mu, \sigma^2)$ and $y = g(x) = e^x$. Then $g'(x) = e^x$, $x = g^{-1}(y) = \ln y$, $g'\left(g^{-1}(y)\right) = y$. Hence,

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\ln y - \mu)^2}{2\sigma^2}\right) \frac{1}{y}, \quad y > 0 \tag{1.49}$$

A stochastic variable with the pdf given by eq. (1.49) is termed a log-normally distributed variable, $Y \sim LN\left(\mu, \sigma^2\right)$.

The principle of combined stochastic variables can be generalized to stochastic vectors. Let $\mathbf{X}(\omega)$ be a n-dimensional stochastic vector, and $\mathbf{g} : R^n \to R^m$ a m-dimensional stochastic vector function. Then a m-dimensional stochastic vector $\mathbf{Y} : \Omega \to R^m$ is defined from the combined mapping

$$\mathbf{Y}(\omega) = \mathbf{g}(\mathbf{X}(\omega)) \tag{1.50}$$

As an example consider the linear combination

$$\mathbf{Y} = \mathbf{A} + \mathbf{B}\mathbf{X} \tag{1.51}$$

$\mathbf{X}$ is an n-dimensional stochastic variable, $\mathbf{A}$ is an n-dimensional deterministic vector and $\mathbf{B}$ is an $n \times n$ dimensional non-singular deterministic matrix. The inverse transformation becomes

$$\mathbf{X} = \mathbf{B}^{-1}(\mathbf{Y} - \mathbf{A}) \tag{1.52}$$

Consider the small volume $\Delta \mathbf{x}$ placed at $\mathbf{x}$. $\Delta \mathbf{x}$ is mapped into $\Delta \mathbf{y}$ at $\mathbf{y} = \mathbf{A} + \mathbf{B}\mathbf{x}$ by the mapping (1.51). In the limit as $\Delta \mathbf{x}$ and $\Delta \mathbf{y}$ goes to zero, they are related through

$$\frac{\Delta \mathbf{y}}{\Delta \mathbf{x}} = \mid \det(\mathbf{B}) \mid \tag{1.53}$$

The probability that $\mathbf{Y} \in \Delta \mathbf{y}$ is equal to the probability that $\mathbf{X} \in \Delta \mathbf{x}$. Hence,

$$f_{\mathbf{Y}}(\mathbf{y}) \Delta \mathbf{y} = f_{\mathbf{X}}(\mathbf{x}) \Delta \mathbf{x} \Longrightarrow$$

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}\big(\mathbf{B}^{-1}(\mathbf{y} - \mathbf{A})\big) \frac{1}{\mid \det(\mathbf{B}) \mid} \tag{1.54}$$

If $\mathbf{X} \sim N(\boldsymbol{\mu}, \mathbf{C})$, it follows from eq. (1.37) that

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= \frac{\exp\left(-\frac{1}{2}\left(\mathbf{B}^{-1}(\mathbf{y} - \mathbf{A}) - \boldsymbol{\mu}\right)^T \mathbf{C}^{-1}\left(\mathbf{B}^{-1}(\mathbf{y} - \mathbf{A}) - \boldsymbol{\mu}\right)\right)}{(2\pi)^{\frac{n}{2}} (\det(\mathbf{C}))^{\frac{1}{2}} \mid \det(\mathbf{B}) \mid} \\ &= \frac{\exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{A} - \mathbf{B}\boldsymbol{\mu})^T \left(\mathbf{B}\mathbf{C}\mathbf{B}^T\right)^{-1}(\mathbf{y} - \mathbf{A} - \mathbf{B}\boldsymbol{\mu})\right)}{(2\pi)^{\frac{n}{2}} (\det(\mathbf{B}\mathbf{C}\mathbf{B}^T))^{\frac{1}{2}}} \end{aligned} \tag{1.55}$$

where the well-known results form matrix algebra $\left(\mathbf{B}^{-1}\right)^T \mathbf{C}^{-1} \mathbf{B}^{-1} = \left(\mathbf{B}\mathbf{C}\mathbf{B}^T\right)^{-1}$ and $\det(\mathbf{B}) = \det\left(\mathbf{B}^{\frac{1}{2}} \mathbf{B}^{\frac{1}{2}}\right) = \det\left(\mathbf{B}^{\frac{1}{2}}\right) \det\left(\mathbf{B}^{\frac{1}{2}}\right)$. The matrix $\mathbf{B}^{\frac{1}{2}}$ is defined as the solution of the matrix equation $\mathbf{B}^{\frac{1}{2}} \mathbf{B}^{\frac{1}{2}} = \mathbf{B}$. Eq. (1.55) shows that $\mathbf{Y} \sim N\left(\mathbf{A} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B}\mathbf{C}\mathbf{B}^T\right)$. Hence, normality is preserved under linear transformations as eq. (1.51). This result can be shown to hold, even when $\mathbf{B}$ is non-quadratic, i.e. when $\mathbf{X}$ and $\mathbf{Y}$ are of different dimensions n and m.

Example 1.6: Random phase stochastic variable

Let Y be a combined stochastic variable of $X \sim U(0, 2\pi)$ defined by the mapping

$$Y = g(X) = a\cos(b + X) \tag{1.56}$$

where $a \in R_+$ and $b \in R$. Y defined by eq. (1.56) is called a *random phase variable*. Eq. (1.56) does not define a monotonous mapping of the interval $[0, 2\pi]$ on $[-a, a]$ for any value of b. Hence, the method eq. (1.42) cannot be used.
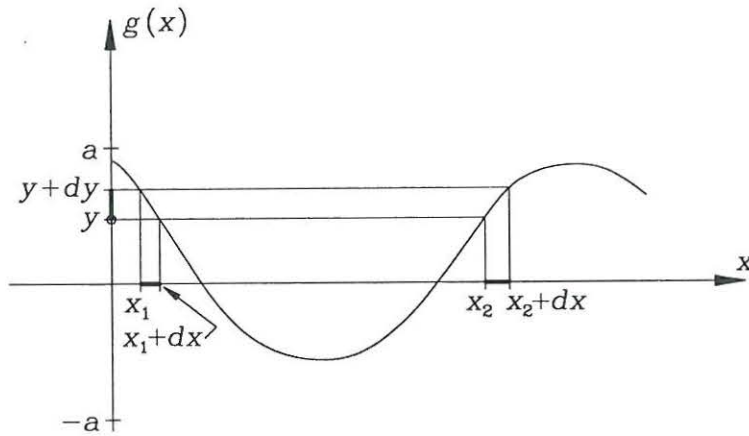
Figure 1.12: Mapping $y = g(x)$ for $\omega_0 t = \frac{\pi}{6}$.

However, as seen from figure 1.12 the event $\{Y \in [y, y + dy]\}$ occurs when $\{X \in [x_1, x_1 + dx] \cup [x_2, x_2 + dx]\}$, where x_1 and x_2 are the mapping points of the inverse mapping $g^{-1}(y)$. The length of the intervals $[x_1, x_1 + dx]$ and $[x_2, x_2 + dx]$ are both equal to dx because of the symmetry property of the mapping function. It then follows

$$f_Y(y)\,dy = f_X(x_1)\,dx + f_X(x_2)\,dx = \frac{1}{2\pi}dx + \frac{1}{2\pi}dx = \frac{1}{\pi}dx \quad \Rightarrow$$

$$f_Y(y) = \frac{1}{\pi}\left|\frac{dx}{dy}\right| = \frac{1}{\pi}\left|\frac{-1}{a\sin(b+x)}\right| = \frac{1}{\pi}\frac{1}{a\sqrt{1-\cos^2(b+x)}}$$

$$= \frac{1}{\pi}\frac{1}{\sqrt{a^2-y^2}} \tag{1.57}$$

The probability density function (1.57) is shown in figure 1.13.
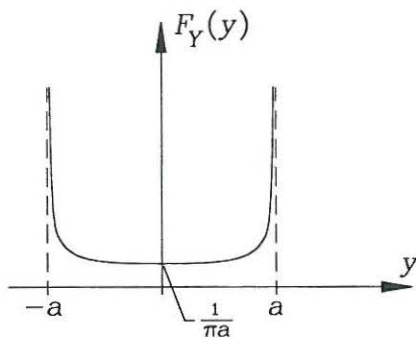
Figure 1.13: Probability density function of random phase stochastic variable.

Example 1.7: Transformations to standard normal space

Let $\mathbf{X} \sim N\left(\boldsymbol{\mu_X}, \mathbf{C_{XX}}\right)$ and select $\mathbf{A}$ and $\mathbf{B}$ in eq. (1.51) in a way that $\mathbf{Y} \sim N\left(\boldsymbol{\mu_Y}, \mathbf{C_{YY}}\right)$ fulfil

$$\boldsymbol{\mu_Y} = \mathbf{A} + \mathbf{B}\boldsymbol{\mu_X} = \mathbf{0} \tag{1.58}$$

$$\mathbf{C_{YY}} = \mathbf{B}\mathbf{C_{XX}}\mathbf{B}^T = \mathbf{I} \tag{1.59}$$

where $\mathbf{I}$ denotes the identity matrix. A solution to eq. (1.59) is

$$\mathbf{B}^T = \mathbf{C_{XX}^{-\frac{1}{2}}} = \boldsymbol{\Phi}\boldsymbol{\Lambda_{XX}^{-\frac{1}{2}}}\boldsymbol{\Phi}^{-1} \tag{1.60}$$

$$\boldsymbol{\Phi} = \left[\boldsymbol{\Phi}^{(1)}, \ldots, \boldsymbol{\Phi}^{(n)}\right] \tag{1.61}$$

$$\boldsymbol{\Lambda_{XX}^{-\frac{1}{2}}} = \begin{bmatrix} \lambda_1^{-\frac{1}{2}} & & \\ & \ddots & \\ & & \lambda_n^{-\frac{1}{2}} \end{bmatrix} \tag{1.62}$$

where $\boldsymbol{\Phi}^{(i)}$ and λ_i are the eigenvectors and eigenvalues of $\mathbf{C_{XX}}$, i.e.

$$\mathbf{C_{XX}}\boldsymbol{\Phi}^{(i)} = \lambda_i\boldsymbol{\Phi}^{(i)}, \quad i = 1, 2, \ldots, n \tag{1.63}$$

Eq. (1.63) can be assembled as

$$\mathbf{C_{XX}}\boldsymbol{\Phi} = \boldsymbol{\Phi}\boldsymbol{\Lambda_{XX}} \Rightarrow$$
$$\mathbf{C_{XX}} = \boldsymbol{\Phi}\boldsymbol{\Lambda_{XX}}\boldsymbol{\Phi}^{-1} \tag{1.64}$$

Upon insertion of eqs. (1.60) and (1.64) in eq. (1.59), the validity of eq. (1.60) is proved. with the solution obtained for $\mathbf{B}$, $\mathbf{A}$ should be selected as follows

$$\mathbf{A} = -\mathbf{B}\boldsymbol{\mu_X} \tag{1.65}$$

The indicated solution eq. (1.60) is not unique. Actually, if $\mathbf{B}$ is a solution any quantity $\mathbf{B}_1$ of the following form will do

$$\mathbf{B}_1 = \mathbf{O}\mathbf{B} \tag{1.66}$$

where $\mathbf{O}$ is an arbitrary orthonormal transformation (rotation) matrix, i.e. $\mathbf{O}$ has the property $\mathbf{O}^{-1} = \mathbf{O}^T$. Actually

$$\mathbf{B}_1\mathbf{C_{XX}}\mathbf{B}_1^T = \mathbf{O}\mathbf{B}\mathbf{C_{XX}}\mathbf{B}^T\mathbf{O}^T = \mathbf{O}\mathbf{I}\mathbf{O}^T = \mathbf{I} \tag{1.67}$$

Eq. (1.55) can then be written

$$f_{\mathbf{Y}}(\mathbf{y}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y_i^2\right) \tag{1.68}$$

Eq. (1.68) shows that all components in $\mathbf{Y}$ are mutually stochastically independent and identically distributed $N \sim (0,1)$. Hence, it has been shown that at least one linear transformation exists (actually infinitely many), which maps a normal vector $\mathbf{X} \sim \left(\boldsymbol{\mu_X}, \mathbf{C_{XX}}\right)$ into standardized independent normal vectors $\mathbf{Y} \sim N\left(\mathbf{0}, \mathbf{I}\right)$.

Finally, consider the non-linear bijective (one-to-one) mapping

$$\mathbf{Y} = \mathbf{g}\left(\mathbf{X}\right) \tag{1.69}$$

where $\mathbf{X}$ and $\mathbf{Y}$ are n-dimensional stochastic variables. Using the same arguments as leading to eqs. (1.42) and (1.54), the pdf of $\mathbf{Y}$ is evaluated as

$$f_{\mathbf{Y}} = f_{\mathbf{X}}\left(\mathbf{g}^{-1}\left(\mathbf{y}\right)\right) \frac{1}{\left| \det\left(\frac{\partial \mathbf{g}(\mathbf{g}^{-1}(\mathbf{y}))}{\partial \mathbf{x}^T}\right)\right|} \tag{1.70}$$

where $\det\left(\frac{\partial \mathbf{g}(\mathbf{g}^{-1}(\mathbf{y}))}{\partial \mathbf{x}^T}\right)$ signifies the *Jacobian* of the mapping (1.69) and $\mathbf{x} = \mathbf{g}^{-1}\left(\mathbf{y}\right)$ is the inverse mapping.

Example 1.8: Box-Muller transformation

Consider the non-linear transformation

$$\left.\begin{aligned} y_1 &= \left(-2\ln x_1\right)^{\frac{1}{2}} \cos\left(2\pi x_2\right) \\ y_2 &= \left(-2\ln x_1\right)^{\frac{1}{2}} \sin\left(2\pi x_2\right) \end{aligned}\right\} \tag{1.71}$$

Let $\mathbf{X} = [X_1, X_2]$, where $X_1, X_2 \sim U\left(0, 1\right)$ are stochastic independent. The problem is then to determine the joint probability density of $\mathbf{Y}^T = [Y_1, Y_2]$ obtained by the transformation (1.71).

With the restriction $[x_1, x_2] \in [0, 1] \times [0, 1]$ the mapping (1.71) is bijective. Hence, the joint pdf of $\mathbf{Y}^T = [Y_1, Y_2]$ can be obtained by eq. (1.70). The equations in (1.71) are squared and added. Then, it follows

$$x_1 = \exp\left(-\frac{1}{2}\left(y_1^2 + y_2^2\right)\right) \tag{1.72}$$

Further the Jacobian of the mapping (1.71) becomes

$$\det\left(\frac{\partial \mathbf{g}\overline{\mathbf{x}}}{\partial \mathbf{x}^T}\right) =$$

$$\det\left(\begin{bmatrix} -\frac{1}{x_1}(-2\ln x_1)^{-\frac{1}{2}}\cos(2\pi x_2) & -2\pi(-2\ln x_1)^{\frac{1}{2}}\sin(2\pi x_2) \\ -\frac{1}{x_1}(-2\ln x_1)^{-\frac{1}{2}}\sin(2\pi x_2) & 2\pi(-2\ln x_1)^{\frac{1}{2}}\cos(2\pi x_2) \end{bmatrix}\right) =$$

$$-\frac{2\pi}{x_1} \tag{1.73}$$

From eq. (1.72) and (1.73) it then follows

$$\left| \det\left(\frac{\partial \mathbf{g}(\mathbf{g}^{-1}(\mathbf{y}))}{\partial \mathbf{x}^T}\right)\right| = \frac{2\pi}{\exp\left(-\frac{1}{2}(y_1^2 + y_2^2)\right)} \tag{1.74}$$

$X_1 \sim U(0,1) \wedge X_2 \sim U(0,1)$ and mutually stochastic independence of X_1 and X_2 implies that

$$f_{X_1 X_2}(x_1, x_2) = \begin{cases} 1 \cdot 1 & , [x_1, x_2] \in [0,1] \times [0,1] \\ 0 & , [x_1, x_2] \notin [0,1] \times [0,1] \end{cases} \tag{1.75}$$

It then follows from eq. (1.70)

$$f_{Y_1 Y_2}(y_1, y_2) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}\left(y_1^2 + y_2^2\right)\right) = f_{Y_1}(y_1) f_{Y_2}(y_2) \tag{1.76}$$

where

$$f_{Y_i} = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y_i^2\right) \tag{1.77}$$

Hence, mutually independent stochastic variables, $X_1 \sim U(0,1)$, $X_2 \sim U(0,1)$, are transformed into mutually independent stochastic variables $Y_1 \sim N(0,1)$, $Y_2 \sim N(0,1)$ by the mapping (1.71). Eq. (1.71) is known as the *Box-Muller transformation* and may be used to generate samples $[y_1, y_2]$ of mutually independent stochastic variables $Y_1 \sim N(0,1)$ and $Y_2 \sim N(0,1)$. Initially, independent samples x_1 and x_2 are generated by a pseudo random number generator. If this sample $[x_1, x_2]$ is inserted in the right-hand side of eq. (1.71) the left-hand side will be samples of the indicated stochastic variables Y_1 and Y_2.

Next, consider the monotonous mapping $g :]0,1] \to R_+$ defined by

$$r = g(x_1) = (-2\ln x_1)^{\frac{1}{2}} \quad , \quad x_1 \in \,]0,1] \tag{1.78}$$

It follows that $x_1 = g^{-1}(r) = \exp\left(-\frac{1}{2}r^2\right)$ and $g'(x_1) = -\frac{1}{x_1}(-2\ln x_1)^{-\frac{1}{2}} \Rightarrow g'\left(g^{-1}(r)\right) = -\frac{1}{r}\exp\left(\frac{1}{2}r^2\right)$. The combined stochastic variable $R = g(X_1)$, where $X_1 \sim U(0,1)$ and $g(x_1)$ is defined by eq. (1.78), then has the probability density function of R is given by eq. (1.42)

$$f_R(r) = f_{X_1}(g^{-1}(r)) \frac{1}{|\, g'(g^{-1}(r)) \,|} = 1 \frac{1}{\frac{1}{r}\exp\left(\frac{1}{2}r^2\right)} = r\exp\left(-\frac{1}{2}r^2\right) \tag{1.79}$$

Hence, $R \sim R(1)$, i.e. Rayleigh distributed with the parameter $\sigma^2 = 1$. Consider the 2-dimensional stochastic variable $\mathbf{Y}^T = [Y_1, Y_2]$ defined by the following transformation

$$\left. \begin{array}{l} Y_1 = R\cos(2\pi X_2) \\ Y_2 = R\sin(2\pi X_2) \end{array} \right\} \tag{1.80}$$

If $R \sim R(1)$ and $X_2 \sim U(0,1)$ and R and X_2 are stochastic independent, eq. (1.80) is then equivalent to the Box-Muller transformation, and it follows that $Y_1 \sim N(0,1)$ and $Y_2 \sim N(0,1)$ and that Y_1 and Y_2 are stochastically independent. Eq. (1.80) can also be used to generate samples of mutually independent standardized normally distributed stochastic variables. In this case independent samples of R (by the inverse method) are generated along with a sample of X_2. The samples of Y_1 and Y_2 then follow upon inserting into eq. (1.80).

1.3 Expectations

The *expected value* (mean value), $E[X]$, of a discrete and a continuous stochastic variable, respectively, is defined as

$$E[X] = \sum_{i=1}^{n} x_i P_X(x_i) \tag{1.81}$$

$$E[X] = \int_{-\infty}^{\infty} x_i f_X(x) dx \tag{1.82}$$

In the following the symbox μ_X is reserved for the expected value. If the probability density function, $f_X(x)$, is visualized as a continuous distribution of mass of total magnitude 1 along the x-axis (occationally the phrase "probability mass" is heard) and $P_X(x_i)$ as concentrated mass particles at the points x_i, μ_X can be interpreted as the center of gravity of the mass distribution.

The *variance*, $Var[X]$, of X is defined as the expected value of the combined stochastic variable $(X - \mu_X)^2$. For a discrete and continuous stochastic variable it is given as

$$
\begin{aligned}
Var[X] = E\left[(X - \mu_X)^2\right] &= \sum_{i=1}^{n} (x_i - \mu_X)^2 P_X(x_i) \\
&= \sum_{i=1}^{n} x_i^2 P_X(x_i) - 2\mu_X \sum_{i=1}^{n} x_i P_X(x_i) + \mu_X^2 \sum_{i=1}^{n} P_X(x_i) \\
&= E[X^2] - \mu_X^2
\end{aligned}
\tag{1.83}
$$

$$
\begin{aligned}
Var[X] = E\left[(X - \mu_X)^2\right] &= \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) \, dx \\
&= \int_{-\infty}^{\infty} x^2 f_X(x) \, dx - 2\mu_X \int_{-\infty}^{\infty} x f_X(x) \, dx + \mu_X^2 \int_{-\infty}^{\infty} f_X(x) dx \\
&= E[X^2] - \mu_X^2
\end{aligned}
\tag{1.84}
$$

where the normalization conditions $\sum_{i=1}^{n} P_X(x_i) = 1$ and $\int_{-\infty}^{\infty} f_X(x) \, dx = 1$ have been used. The symbol σ_X^2 will be reserved for $Var[X]$. σ_X is termed the *standard deviation*. Within the mentioned mass analogy, σ_X^2 can be interpreted as the mass moment of inertia around the center of gravity. In case $\mu_X \neq 0$, the *coefficient of variation*, $V[X]$, is defined as

$$V[X] = \frac{\sigma_X}{\mu_X} \tag{1.85}$$

The abbreviation V_X will be used for $V[X]$.

Consider more generally the combined stochastic variable $Y = g(X_1, X_2, \ldots, X_n)$. The expected value of Y is computed in the following way for discrete and continuous stochastic variables

$$E[Y] = E[g(X_1, \ldots, X_n)]$$

$$= \sum_{i_1} \cdots \sum_{i_n} g(x_{i_1}, \ldots, x_{i_n}) P_{X_1 \ldots X_n}(x_{i_1}, \ldots, x_{i_n}) \tag{1.86}$$

$$E[Y] = E[g(X_1, \ldots, X_n)]$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots x_n) f_{X_1 \ldots X_n}(x_1, \ldots x_n) dx_1 \ldots dx_n \tag{1.87}$$

In case infinitely many states of the discrete n-dimensional variable $[X_1, X_2, \ldots, X_n]$ the multiple sum (1.86) may not converge. In this case the expected value of $Y = g(X_1, X_2, \ldots, X_n)$ does not exist. For the continuous n-dimensional stochastic variable the integral (1.87) may also be divergent, as the integral limits goes to infinity, so the expectation does not exist.

If $g(\mathbf{X}) = X_i$ and $g(\mathbf{X}) = (X_i - \mu_{X_i})(X_j - \mu_{X_j})$, the *expected values* and *covariances* are obtained, i.e.

$$\mu_{X_i} = E[X_i] \tag{1.88}$$

$$C_{X_i X_j} = E[(X_i - \mu_{X_i})(X_j - \mu_{X_j})] = E[X_i X_j] - \mu_{X_i} \mu_{X_j} \tag{1.89}$$

The last statement of eq. (1.89) is proved in the same way as the last statements of eqs. (1.83) and (1.84). Occasionally, we shall abbreviate the index notation to write $\mu_{X_i} = \mu_i$ and $C_{X_i X_j} = C_{ij}$. The variances are obtained from $\sigma_{X_i}^2 = C_{X_i X_i}$. The *correlation coefficients* are defined as

$$\rho_{X_i X_j} = \frac{C_{X_i X_j}}{\sigma_{X_i} \sigma_{X_j}} \tag{1.90}$$

Because $f_{X_i X_j}(x_i, x_j) \geq 0$ and $(a(X_i - \mu_{X_i}) + (X_j - \mu_{X_j}))^2 \geq 0$ it follows for all $a \in R$

$$E\left[(a(X_i - \mu_{X_i}) + (X_j - \mu_{X_j}))^2\right] \geq 0 \quad \Rightarrow$$

$$a^2 \sigma_{X_i}^2 + 2a C_{X_i X_j} + \sigma_{X_j}^2 \geq 0 \tag{1.91}$$

Since the left-hand side of the 2nd order equation (1.91) is non-negative, its discriminant must be non-positive. Using eq. (1.90) it then follows

$$4C_{X_i X_j}^2 - 4\sigma_{X_i}^2 \sigma_{X_j}^2 \leq 0 \quad \Rightarrow$$

$$\rho_{X_i X_j}^2 \leq 1 \quad \Rightarrow$$

$$-1 \leq \rho_{X_i X_j} \leq 1 \tag{1.92}$$

The variables X_i and X_j are said to be *uncorrelated*, if $\rho_{X_i X_j} = 0$. Now, let $[X_i, X_j]$ be normally distributed. From eq. (1.38) follows that the parameter ρ can be interpreted as the correlation coefficient of X_i and X_j. If X_i and X_j are uncorrelated, so $\rho = 0$, it follows from eq. (1.39) that X_i and X_j are then independent. Generally, it can be stated that normally distributed mutually uncorrelated stochastic variables are mutually stochastic independent as well, because the covariance matrix $\mathbf{C_{XX}}$ becomes diagonal, so $F_{\mathbf{X}}(\mathbf{x}) = \Pi_{i=1}^n f_{X_i}(x_i)$. It is generally *not* true that uncorrelation among stochastic variables implies mutually stochastic independence. However, if X_i and X_j are mutually independent stochastic variables, then

$$C_{X_i X_j} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - \mu_{X_i})(x_j - \mu_{X_j}) f_{X_i}(x_i) f_{X_j}(x_j) \, dx_i dx_j$$

$$= \int_{-\infty}^{\infty} (x_i - \mu_{X_i}) f_{X_i}(x_i) \, dx_i \int_{-\infty}^{\infty} (x_j - \mu_{X_j}) f_{X_j}(x_j) \, dx_j = 0 \tag{1.93}$$

Hence, the opposite statement is true. Mutually stochastic independence always implies uncorrelation. Finally, it can be shown that the mean value of the product of two uncorrelated variables is equal to the product of the mean values

$$E[X_i X_j] = C_{X_i X_j} + \mu_{X_i} \mu_{X_j} = 0 + \mu_{X_i} \mu_{X_j}$$

$$= E[X_i] E[X_j] \tag{1.94}$$

1.4 Conditional distributions

Let $[X, Y]$ be a two-dimensional stochastic variable. If $[X, Y]$ is discrete the *conditional probability function* of Y on X is defined as

$$P_{Y|X}(y_j \mid x_i) = \frac{P_{XY}(x_i, y_j)}{P_X(x_i)}, \quad P_X(x_i) > 0 \tag{1.95}$$

If $[X, Y]$ is continuous the *conditional probability density function* of Y on X is defined as

$$f_{Y|X}(y \mid x) = \frac{f_{XY}(x,y)}{f_X(x)}, \quad f_X(x) > 0 \tag{1.96}$$

$P_{Y|X}(y_j \mid x_i)$ and $f_{Y|X}(y \mid x)$ has the properties of a usual probability function and a usual probability density function, i.e. they are non-negative, and admit the normalization conditions

$$\sum_j P_{Y|X}(y_j \mid x_i) = \frac{\sum_j P_{XY}(x_i, y_j)}{P_X(x_i)} = \frac{P_X(x_i)}{P_X(x_i)} = 1 \tag{1.97}$$

$$\int_{-\infty}^{\infty} f_{Y|X}(y \mid x)\, dy = \frac{\int_{-\infty}^{\infty} f_{YX}(x,y)dy}{f_X(x)} = \frac{f_X(x)}{f_X(x)} = 1 \tag{1.98}$$

The conditional expected value of the combined stochastic variable $h(Y)$ on $X = x$ can then be defined from

$$E[h(Y) \mid X = x] = \begin{cases} \displaystyle\sum_j h(y_j)P_{Y|X}(y_j \mid x_i) \\[2mm] \displaystyle\int_{-\infty}^{\infty} h(y)f_{Y|X}(y \mid x)dy \end{cases} \tag{1.99}$$

If $h(Y) = Y$ then eq. (1.99) determines the *conditional expected value*, $\mu_{Y|X} = E[Y \mid X = x]$, and if $h(Y) = (Y - \mu_{Y|X})^2$, eq. (1.99) determines the *conditional variance*, denoted $\sigma_{Y|X}^2 = E\left[(Y - \mu_{Y|X})^2 \mid X = x\right]$. $g(x) = E[h(y) \mid x]$ is a function of x. We may then define the combined stochastic variable g. The expected value of this quantity becomes

$$E[g] = E[E[h(Y) \mid X]] = E[h(Y)] \tag{1.100}$$

Eq. (1.100) will only be proved for the continuous case. From eqs. (1.87) and (1.96) follow

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

$$= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} h(y)f_{Y|X}(y \mid x)\, dy \right) f_X(x)dx = \int_{-\infty}^{\infty} h(y)f_Y(y)dy$$

$$= E\left[h\left(Y\right)\right] \tag{1.101}$$

In the same way it can be proved that

$$E\left[E\left[Xh\left(Y\right) \mid X\right]\right] = E\left[Xh\left(Y\right)\right] \tag{1.102}$$

Example 1.9: Conditional normal probability density functions

Let $[X, Y] \sim N\left(\boldsymbol{\mu}, \mathbf{C}\right)$ where

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix} \tag{1.103}$$

$$\mathbf{C} = \begin{bmatrix} \sigma_X^2 & \sigma_X \sigma_Y \rho \\ \sigma_X \sigma_Y \rho & \sigma_Y^2 \end{bmatrix} \tag{1.104}$$

Then it follows from eq. (1.39) that the joint pdf can be written

$$f_{XY}(x, y) = \frac{1}{\sqrt{2\pi}\sigma_X \sqrt{2\pi}\sigma_Y \sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \right.$$

$$\left. \frac{1}{2}\left(\frac{y-\mu_Y - \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)}{\sigma_Y\sqrt{1-\rho^2}}\right)^2\right]$$

$$= f_X(x)f_{Y|X}\left(y \mid x\right) \tag{1.105}$$

$$f_X\left(x\right) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right) \tag{1.106}$$

$$f_{Y|X}\left(y \mid x\right) = \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\left(\frac{y-\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)}{\sigma_Y\sqrt{1-\rho^2}}\right)^2\right) \tag{1.107}$$

Eq. (1.107) shows that the conditional distribution also becomes normal with the following conditional expected value and conditional variance

$$E[Y \mid x] = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}\left(x - \mu_X\right) \tag{1.108}$$

$$Var[Y \mid x] = \sigma_Y^2\left(1 - \rho^2\right) \tag{1.109}$$

As an extension of eq. (1.107) consider the m-dimensional stochastic vector $\mathbf{X} \sim N\left(\boldsymbol{\mu_X}, \mathbf{C_{XX}}\right)$ and an n-dimensional stochastic vector $\mathbf{Y} \sim N\left(\boldsymbol{\mu_Y}, \mathbf{C_{YY}}\right)$. The covariance between $\mathbf{X}$ and $\mathbf{Y}$ is given by the matrix $\mathbf{C_{XY}}$. It can then be shown that the pdf. of $\mathbf{Y}$ on the condition $\mathbf{X} = \mathbf{x}$ is normally distributed with mean and covariance

$$\left.\begin{array}{l}\boldsymbol{\mu}_{\mathbf{Y}|\mathbf{X}} = \boldsymbol{\mu}_{\mathbf{Y}} + C_{\mathbf{XY}}^{T} C_{\mathbf{XX}}^{-1}\left(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}}\right) \\[2mm] C_{\mathbf{Y}|\mathbf{X}} = C_{\mathbf{YY}} - C_{\mathbf{XY}}^{T} C_{\mathbf{XX}}^{-1} C_{\mathbf{XY}}\end{array}\right\} \tag{1.110}$$

So, any conditional distribution of a normal vector is normal itself. Along with the preservation of normality of any marginal distribution and under linear transformations this stresses the attractive mathematical properties of the normal distribution.

If X and Y are mutually independent stochastic variables then $P_{Y|X}(y \mid x) = P_Y(y)$ and $f_{Y|X}(y \mid x) = f_Y(y)$, as seen from the following statements

$$P_{Y|X}(y \mid x) = \frac{P_{XY}(x,y)}{P_X(x)} = \frac{P_X(x)P_Y(y)}{P_X(x)} = P_Y(y) \tag{1.111}$$

$$f_{Y|X}(y \mid x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{f_X(x)f_Y(y)}{f_X(x)} = f_Y(y) \tag{1.112}$$

1.5 Convergency of a sequence of stochastic variables

Initially consider an infinite sequence of real numbers $\{x_n, \ n = 1, 2, \dots\} = \{x_1, x_2, \dots\}$. Convergency of this sequence to some limit x means that all elements x_n in the sequence deviates within an arbitrary small value ε from x after a certain number N, which of course depends on the selected value of ε. Formally this statement can be written

$$\forall \varepsilon > 0 \exists N = N(\varepsilon) : n \geq N(\varepsilon) \ \Rightarrow \ \mid x_n - x \mid < \varepsilon \tag{1.113}$$

The convergency of a sequence is often tested by the equivalent Cauchy's criterion of mutual convergence, which states that a limit x exists if and only if

$$\forall \varepsilon > 0 \exists N = N(\varepsilon) : n, m \geq N(\varepsilon) \Rightarrow \mid x_n - x_m \mid < \varepsilon \tag{1.114}$$

The following abbreviated notation is introduced

$$\lim_{n \to \infty} x_n = x \tag{1.115}$$

Next, consider an infinite sequence of random variables $\{X_n, \ n = 1, 2, \dots\}$, all defined on the same sample space Ω. For each elementary event $\omega \in \Omega$ an ordinary sequence of real numbers is generated $\{x_n, \ n = 1, 2, \dots\}$, where $x_n = X_n(\omega)$. Define $A \subseteq \Omega$ as the set of elementary events ω for which such sequences do converge. It can be shown that A is an event, i.e. the probability $P(A)$ is defined. One may then construct a stochastic variable $X : \Omega \to R$ in the following way

$$\forall \omega \in A : X(\omega) = \lim_{n \to \infty} X_n(\omega) \tag{1.116}$$

$$\forall \omega \in A^c : X(\omega) \quad \text{otherwise specified} \tag{1.117}$$

where A^c specifies the complementary event. If $P(A) = 1$, we say that the sequence of stochastic variables $\{X_n, n = 1, 2, \dots\}$ *converges with probability 1*, or *almost certain* to the stochastic variable X. The following abbreviated notation is used to indicate the limit with probability 1

$$\lim_{n \to \infty}\text{-ac } X_n = X \tag{1.118}$$

Notice, A^c need not be empty even though $P(A^c) = 0$. The notation merely states that the number of elements $\omega \in A^c$ is infinitely small compared to the number of elements in A.

The set of samples for which inequality (1.113) is fulfilled for some $n \geq N(\varepsilon)$ can be written $\{\omega \in \Omega \;||\; X_n(\omega) - X(\omega) \;|< \varepsilon\}$. Hence, the set for which the inequality is fulfilled for all $n \geq N(\varepsilon)$ can be written as the intersection of all such events, $\cap_{n=N(\varepsilon)}^{\infty} \{\omega \in \Omega \;||\; X_n(\omega) - X(\omega) \;|< \varepsilon\}$. Convergence with probability 1 implies, that the probability of this event is 1. Hence, eq. (1.118) can be written as

$$\forall \varepsilon > 0 \exists N(\varepsilon) P : \left(\cap_{n=N(\varepsilon)}^{\infty} \{\omega \in \Omega \;||\; X_n(\omega) - X(\omega) \;|< \varepsilon\}\right) = 1 \tag{1.119}$$

Eq. (1.119) can also be written

$$\forall \varepsilon > 0 : \lim_{N \to \infty} P\left(\cap_{n=N}^{\infty} \{\omega \in \Omega \;||\; X_n(\omega) - X(\omega) \;|< \varepsilon\}\right) = 1 \tag{1.20}$$

An alternative limit definition of the sequence $\{X_n, : i = 1, 2, \dots\}$ can be defined, if there exists a stochastic variable $X : \Omega \to R$, so

$$\lim_{n \to \infty} E\left[(X_n - X)^2\right] = 0 \tag{1.121}$$

Then, $\{X_n, i = 1, 2, \dots\}$ is said to be *convergent in mean square* to the limit X. Eq. (1.121) requires the convergence of the sequence of ordinary non-negative real numbers $\{E\left[(X_n - X)^2\right], n = 1, 2, \dots\}$ to the limit $x = 0$ according to eq. (1.113). Mean square convergence can also be stated in terms of an equivalent Cauchy type of criterion, saying that a mean square limit X exists to the sequence $\{X_1, X_2, \dots\}$ if and only if

$$\lim_{\substack{n \to \infty \\ m \to \infty}} E\left[(X_n - X_m)^2\right] = 0 \tag{1.122}$$

The following abbreviated notation will be used for the limit in mean square

$$\lim_{n \to \infty}\text{-ms } X_n = X \tag{1.123}$$

Example 1.10: Sequence not convergent with probability 1 and convergent in mean square

To emphasize the difference between the limit stochastic variables $\lim-ac_{n\to\infty}X_n$ and $\lim-ms_{n\to\infty}X_n$ consider the following sequence $\{X_n, n = 1, 2, ...\}$. The stochastic variables of the sequence are assumed to be mutually independent. Further, the stochastic variables are assumed to be discrete with only two states $x_1 = 0$ and $x_2 = 1$, with the probability function of X_n given by

$$P_{X_n}(x) = \begin{cases} 1 - \frac{1}{n} & , x = 0 \\ \frac{1}{n} & , x = 1 \end{cases} \tag{1.124}$$

Since the probability mass concentrate at the state $x = 0$ as $n \to \infty$, one may assume that a limit stochastic variable $X(\omega) = 0$ (a stochastic variable with the single state 0) exist in either mean square or with probability 1. To investigate this left-hand sides of eqs. (1.120) and (1.121) will be evaluated.

In case of almost certain convergence to the limit $X = 0$, consider the probability

$$P(\cap_{n=N}^{p}\{\omega \in \Omega \mid\mid X_n(\omega) - X(\omega) \mid < \varepsilon\}) =$$

$$P(\cap_{n=N}^{p}\{\omega \in \Omega \mid\mid X_n(\omega) \mid < \varepsilon\}) =$$

$$P(\cap_{n=N}^{p}\{\omega \in \Omega \mid\mid X_n(\omega) \mid = 0\}) =$$

$$P(\{\omega \in \Omega \mid X_N(\omega) = 0\}) \cdots P(\{\omega \in \Omega \mid X_{p-1}(\omega) = 0\}) P(\{\omega \in \Omega \mid X_p(\omega) = 0\}) =$$

$$\left(1 - \frac{1}{N}\right)\left(1 - \frac{1}{N+1}\right) \cdots \left(1 - \frac{1}{p-1}\right)\left(1 - \frac{1}{p}\right) =$$

$$\frac{N-1}{N}\frac{N}{N+1} \cdots \frac{p-2}{p-1}\frac{p-1}{p} =$$

$$\frac{N-1}{p} \tag{1.125}$$

In the 3rd statement it has been used that $\mid X(\omega) \mid < \varepsilon$ can only be fulfilled if $\mid X(\omega) \mid = 0$. Next, it has been used that any X_n and X_m are mutually stochastic independent, so $P(X_n = 0 \wedge X_m = 0) = P(X_n = 0) P(X_m = 0)$. From eq. (1.125) it now follows

$$P(\cap_{n=N}^{\infty}\{\omega \in \Omega \mid\mid X_n(\omega) - X(\omega) \mid < \varepsilon\}) =$$

$$\lim_{p\to\infty}(\cap_{n=N}^{p}\{\omega \in \Omega \mid\mid X_n(\omega) - X(\omega) \mid < \varepsilon\}) =$$

$$\lim_{p\to\infty}\frac{N-1}{p} = 0 \tag{1.126}$$

Hence, the left-hand side of eq. (1.120) converge to 0 as $N \to \infty$ and no limit $X = 0$ exists with probability 1 for the indicated sequence.

In case of mean square convergence to the limit $X = 0$ consider the expectation

$$E[(X_n - X)^2] = E\left[X_n^2\right] = 0^2 P_{X_n}(0) + 1^2 P_{X_n}(1) = 0^2\left(1 - \frac{1}{n}\right) + 1^2\frac{1}{n} = \frac{1}{n} \tag{1.127}$$

Since this converge to 0 as $n \to \infty$ eq. (1.121) will be fulfilled, and the limit $X = 0$ exists in mean square.

Example 1.11: Sequence convergent with probability 1 and not in mean square

Now, the previous example is considered again with a slightly changed probability function

$$P_{X_n}(x) = \begin{cases} 1 - \frac{1}{n^2} & , x = 0 \\ \frac{1}{n^2} & , x = n \end{cases} \tag{1.128}$$

The states allowed by X_n is now $X = 0$ and $X = n$ with the indicated probabilities. Corresponding to eq. (1.125) we evaluate

$$P(\cap_{n=N}^{p}\{\omega \in \Omega \mid\mid X_n(\omega) \mid= 0\}) =$$

$$P(\{\omega \in \Omega \mid X_N(\omega) = 0\}) \cdots P(\{\omega \in \Omega \mid X_{p-1}(\omega) = 0\})P(\{\omega \in \Omega \mid X_p(\omega) = 0\}) =$$

$$\left(1 - \frac{1}{N^2}\right)\left(1 - \frac{1}{(N+1)^2}\right) \cdots \left(1 - \frac{1}{(p-1)^2}\right)\left(1 - \frac{1}{p^2}\right) =$$

$$\frac{(N-1)(N+1)}{N^2} \frac{N(N+2)}{(N+1)^2} \cdots \frac{(p-2)p}{(p-1)^2} \frac{(p-1)(p+1)}{p^2} =$$

$$\frac{N-1}{N} \frac{p+1}{p} \tag{1.129}$$

From eq. (1.129) follows

$$P(\cap_{n=N}^{\infty}\{\omega \in \Omega \mid\mid X_n(\omega) - X(\omega) \mid< \varepsilon\}) = \lim_{p \to \infty} \frac{N-1}{N} \frac{p+1}{p} = \frac{N-1}{N} \Rightarrow$$

$$\lim_{N \to \infty} P(\cap_{n=N}^{\infty}\{\omega \in \Omega \mid\mid X_n(\omega) - X(\omega) \mid< \varepsilon\}) = \lim_{N \to \infty} \frac{N-1}{N} = 1 \tag{1.130}$$

Consequently, the limit with probability 1 now exists.

In case of mean square convergence to the limit $X = 0$, eq. (1.127) now becomes

$$E[(X_n - X)^2] = E[X_n^2] = 0^2 P_{X_n}(0) + n^2 P_{X_n}(n) = 0^2 \left(1 - \frac{1}{n^2}\right) + n^2 \frac{1}{n^2} = 1 \tag{1.131}$$

From eq. (1.131) now follows $\lim_{n \to \infty} E[(X_n - X)^2] = 1$, so no limit in mean square exists for the new sequence.

From the indicated examples is learned, that almost certain convergence does not imply convergence in mean square, and oppositely, that the existence of a mean square limit does not guarantee almost certain convergence.

Even if both limits exist they may still be different stochastic variables (functions of ω), even for $\omega \in A$ where the sequence of sample values converge according to eq. (1.113). However, it can be shown that if both limits exist, they can only deviate within a non-empty subset of samples $B \subset \Omega$ with the probability $P(B) = 0$.

From later applications the following 2 theorems on mean square convergency are stated.

Theorem 1.1
If $\lim-\text{ms}_{n\to\infty}X_n = X$ and $\lim-\text{ms}_{n\to\infty}Y_n = Y$ then

$$\lim_{\substack{n\to\infty \\ m\to\infty}} E[X_nY_n] = E[XY] \tag{1.132}$$

Proof:

$$E[X_nY_m - XY] =$$

$$E[(X_n - X)(Y_m - Y)] + E[(X_n - X)Y] + E[(Y_m - Y)X] \leq$$

$$E[(X_n - X)^2]^{\frac{1}{2}} E[(Y_m - Y)^2]^{\frac{1}{2}} + E[(X_n - X)^2]^{\frac{1}{2}} E[Y^2]^{\frac{1}{2}} +$$

$$E[(Y_m - Y)^2]^{\frac{1}{2}} E[X^2]^{\frac{1}{2}} \tag{1.133}$$

where use has been made of the so-called Schwarz inequality which states that for any stochastic variables X, Y

$$E[|XY|] \leq E[X^2]^{\frac{1}{2}} E[Y^2]^{\frac{1}{2}} \tag{1.134}$$

The right-hand side of eq. (1.133) converge to zero from the premises, from which eq. (1.132) follows. In order to prove the Schwarz inequality, consider the non-negative quantity $(aX + Y)^2$. The expectation of the quantity is non-negative, leading to

$$E[(aX + Y)^2] \geq 0 \Rightarrow$$

$$a^2 E[X^2] + 2aE[XY] + E[Y^2] \geq 0 \tag{1.135}$$

In analogy to the statement leading to eq. (1.192), the discriminant of the 2nd order equation on the left-hand side must be non-positive. This leads directly to the Schwarz inequality (1.134).

Theorem 1.2 Mean square convergence criterion
The sequence $\{X_n, n = 1, 2, \ldots\}$ converges in mean square to some stochastic variable X if and only if the moments $E[X_nX_m]$ exists and has the same limit no matter how the limit passings $n \to \infty$ and $m \to \infty$ are performed.

Proof:
Consider

$$E[(X_n - X_m)^2] = E[X_n^2] - 2E[X_nX_m] + E[X_m^2] \tag{1.136}$$

At the proof of the sufficiency (the "if part") of the theorem, it is assumed that $E[X_n^2]$, $E[X_nX_m]$ and $E[X_m^2]$ all converge to the same limit no matter how $n \to \infty$ and $m \to \infty$. The right-hand side of eq. (1.136) then has the limit 0 as $n, m \to \infty$. According to the Cauchy criterion, eq. (1.120) this implies the existence of a mean-square limit X.

At the proof of the necessity (the "only if" part), the existence of a mean-square limit X is assumed, and it shall be proved that $E[X_nX_m]$ converge to a limit, which is indeed $E[X^2]$ when $n \to \infty$ and

$m \to \infty$ independently. This follows by inserting X_m for Y_m and X for Y in theorem 1.1, in which case

$$\lim_{\substack{n \to \infty \\ m \to \infty}} E\left[X_n X_m\right] = E\left[X^2\right] \tag{1.137}$$

Consider the special case when all the stochastic variables Y_m, and so also the limit Y, are discrete and equal to 1, i.e. $\forall \omega \in \Omega : Y_1(\omega) = Y_2(\omega) = \cdots Y(\omega) = 1$. Then, theorem 1.1 provides

$$\lim_{n \to \infty} E\left[X_n\right] = E\left[X\right] = E\left[\lim_{n \to \infty}\text{-ms} X_n\right] \tag{1.138}$$

Eq. (1.138) shows that the operation of limit passing and expectation commute, when limit in mean square of a sequence of stochastic variables is considered. This property will occationally be used in the following.

Obviously, mean square convergence only make sense, if $\mid E\left[X_n X_m\right] \mid < \infty$ for all n, m. From the Schwarz inequality follows

$$\mid E\left[X_n X_m\right] \mid \le E\left[X_n^2\right]^{\frac{1}{2}} E\left[X_m^2\right]^{\frac{1}{2}} \tag{1.139}$$

Hence, $E\left[X_n X_m\right]$ exsits for all n, m, if

$$E\left[X_n^2\right] < \infty \, , \ n = 1, 2 \ldots \tag{1.140}$$

Eq. (1.140) must also be fulfilled by the limit in mean square, X, if this is not included in the sequence. When eq. (1.140) is fulfilled for all n, X_n is called a *2nd order stochastic variable* and $\{X_1, X_2, \ldots\}$ is called a *2nd order sequence*.

If the Schwarz inequality is applied with $Y(0) = 1$, it follows

$$\mid E\left[X\right] \mid = \mid E\left[X \cdot 1\right] \mid \le E\left[X^2\right]^{\frac{1}{2}} E\left[1^2\right]^{\frac{1}{2}} = E\left[X^2\right]^{\frac{1}{2}} \tag{1.141}$$

Even though almost certain convergency may seem more intuitively comprehensive we shall in the following mostly concentrate on convergency in mean square. This is so, because the existance of such a limit can be checked barely from the properties of the joint 2nd order moments $E\left[X_n Y_m\right]$ of the sequence according to the above theorem 1.2. The knowledge of these moments is the best information available about the joint distribution of the stochastic variables of the sequence. In contrast necessary and sufficient conditions for convergence with probability 1 is more involved.

The central limit theorem

Consider a sequence of stochastic variables $\{X_n, n = 1, 2, \ldots\}$ defined by

$$X_n = \frac{1}{n} \sum_{j=1}^{n} Y_j \tag{1.142}$$

where the stochastic variables Y_j are mutually independent and identically distributed with the mean, μ_Y, and variance, σ_Y^2. The mean value, μ_{X_n}, and variance, $\sigma_{X_n}^2$, of X_n becomes

$$\mu_{X_n} = \frac{1}{n} \sum_{j=1}^{n} E[Y_j] = \frac{1}{n} \sum_{j=1}^{n} \mu_Y = \mu_Y \tag{1.143}$$

$$\sigma_{X_n}^2 = E\left[(X_n - \mu_Y)^2\right]$$

$$= E\left[\left(\frac{1}{n} \sum_{j=1}^{n} (Y_j - \mu_Y)\right)^2\right]$$

$$= \frac{1}{n^2} \sum_{j=1}^{n} \sum_{k=1}^{n} E\left[(Y_j - \mu_Y)(Y_k - \mu_Y)\right]$$

$$= \frac{1}{n^2} \sum_{\substack{j=1 \\ }}^{n} \sum_{\substack{k=1 \\ k \neq j}}^{n} E\left[(Y_j - \mu_Y)(Y_k - \mu_Y)\right] + \frac{1}{n^2} \sum_{j=1}^{n} E\left[(Y_j - \mu_Y)^2\right] \tag{1.144}$$

Since Y_j and Y_k are stochastic independent one has

$$E\left[(Y_j - \mu_Y)(Y_k - \mu_Y)\right] = E\left[(Y_j - \mu_Y)\right] E\left[(Y_k - \mu_Y)\right] = 0, \quad j \neq k \tag{1.145}$$

The double sum in eq. (1.144) then cancel, and eq. (1.144) reduces to

$$\sigma_{X_n}^2 = \frac{1}{n^2} \sum_{j=1}^{n} \sigma_Y^2 = \frac{1}{n} \sigma_Y^2 \tag{1.146}$$

Hence,

$$\lim_{n \to \infty} E\left[(X_n - \mu_Y)^2\right] = \lim_{n \to \infty} \frac{1}{n} \sigma_Y^2 = 0 \tag{1.147}$$

This shows that the sequence $\{X_n, n = 1, 2, \ldots\}$ converge in mean square to the limit $X(\omega) = \mu_Y$ as $n \to \infty$. This result states one form of the law of large numbers. It

can even be proved that $\{X_n, n = 1, 2, \dots\}$ converge with probability 1 to the limit $X(\omega) = \mu_Y$ (Kolmogorovs law of large numbers). The proof for this is considerably more involved and is omitted.

Consider the sum of mutually independent stochastic variables Y_j

$$S_n = \sum_{j=1}^{n} Y_j \tag{1.148}$$

The mean value, μ_{S_n}, and variance, $\sigma_{S_n}^2$, follows from eq. (1.146)

$$\mu_{S_n} = \sum_{j=1}^{n} \mu_{Y_j} \tag{1.149}$$

$$\sigma_{S_n}^2 = \sum_{j=1}^{n} \sigma_{Y_j}^2 \tag{1.150}$$

Now, the significant observation is that if none of the stochastic variables Y_j in eq. (1.148) is dominating the sum S_n tends to become normally distributed with the mean value eq. (1.149) and variance eq. (1.150) as the number n of addends increases. This is obtained under fairly mild conditions even if all the stochastic variables Y_j have different probability distributions. This observation is contained in the following very important limit theorem which will be stated without proof.

Theorem 1.3: Central limit theorem
The stochastic variable S_n as defined by eq. (1.148) converge to a normally distributed random variable with the mean value (1.149) and the variance (1.150) as $n \to \infty$ independently of the probability distributions of the stochastic variables Y_j, if the following conditions are fulfilled for some $\delta > 0$

$$E[Y_j] < \infty \tag{1.151}$$

$$E\left[|Y_j - \mu_{Y_j}|^{2+\delta}\right] < \infty \tag{1.152}$$

$$\lim_{n \to \infty} \frac{\sum_{j=1}^{n} E\left[|Y_j - \mu_{Y_j}|^{2+\delta}\right]}{\sigma_{S_n}^{2+\delta}} = 0 \tag{1.153}$$

Because sums and differences of normally distributed stochastic variables become normally distributed, the theorem is trivial if $Y_j \sim N\left(\mu_{Y_j}, \sigma_{Y_j}^2\right)$.

In structural mechanics a certain load effect can often be modelled as a sum of numerous mutually independent and equally important contributions (if some components were of minor importance to the problem they should be excluded from the model!). From the central limit theorem it can then be concluded, that the said load effect is normally distributed. Aside from tractable mathematical properties, which is a legal objective

when competing models cannot be graded due to lack of information, the enormous popularity of normal distributions in science and engineering is justified mainly from the central limit theorem. There exist various generalizations of the theorem. For instance, consider the sequence of n-dimensional stochastic vectors $\{\mathbf{S}_n, \ n = 1, 2, \ldots \}$, where

$$\mathbf{S}_n = \sum_{j=1}^{n} \mathbf{Y}_j \tag{1.154}$$

$\mathbf{Y}_j$ are mutually independent n-dimensional stochastic variables with mean value $\boldsymbol{\mu}_{\mathbf{Y}_j}$ and covariance matrix $\mathbf{C}_{\mathbf{Y}_j \mathbf{Y}_j}$. Then, eq. (1.154) converges under equally mild conditions as $n \to \infty$ to a n-dimensional normal vector $\mathbf{S}_n \sim N\left(\boldsymbol{\mu}_{\mathbf{S}_n}, \mathbf{C}_{\mathbf{S}_n \mathbf{S}_n}\right)$, where

$$\boldsymbol{\mu}_{\mathbf{S}_n} = \sum_{j=1}^{n} \boldsymbol{\mu}_{\mathbf{Y}_j} \tag{1.155}$$

$$\mathbf{C}_{\mathbf{S}_n \mathbf{S}_n} = \sum_{j=1}^{n} \mathbf{C}_{\mathbf{Y}_j \mathbf{Y}_j} \tag{1.156}$$

Example 1.12: Distribution of sums of identically uniformly distributed stochastic variables

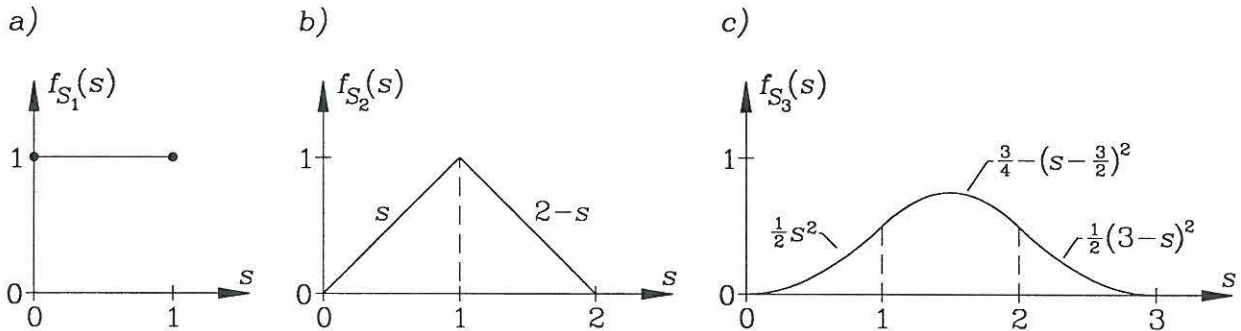Figure 1.14: Distribution of sum of uniformly distributed random varibles.

The probability distribution function of S_n as given by eq. (1.148) can be determined from the recursive formula

$$
\begin{aligned}
F_{S_n}(s) &= P\left(S_n \leq s\right) \\
&= \int_{-\infty}^{\infty} P(S_{n-1} + y \leq s \mid Y_n = y) f_{Y_n}(y) dy \\
&= \int_{-\infty}^{\infty} P(S_{n-1} + y \leq s) f_{Y_n}(y) dy
\end{aligned}
$$

$$= \int\limits_{-\infty}^{\infty} F_{S_{n-1}}(s-y) f_{Y_n}(y)\, dy \tag{1.157}$$

where it has been used that $S_{n-1} = Y_1 + Y_2 + \ldots Y_{n-1}$ is stochastic independent of Y_n Now, let $Y_j \sim U(0,1)$. Eq. (1.157) then provides the following recursive relation for the probability density function

$$F_{S_n}(s) = \int\limits_{0}^{1} F_{S_{n-1}}(s-y) \cdot 1\, dy \;\Rightarrow$$

$$f_{S_n}(s) = \int\limits_{0}^{1} f_{S_{n-1}}(s-y)\, dy, \; n = 2, 3, \ldots \tag{1.158}$$

with

$$f_{S_1}(s) = f_{Y_1}(s) = \begin{cases} 0 & , s \notin [0,1] \\ 1 & , s \in [0,1] \end{cases} \tag{1.159}$$

The result obtained from eq. (1.158) has been sketched in figure 1.14 for $n = 2, 3$. The resemblance with a normally distributed variable is allready visible for $f_{S_3}(s)$.

Example 1.13: Product of independent stochastic variables

Consider a sequence of stochastic variables $\{P_n, n = 1, 2, \ldots\}$ defined by

$$P_n = \prod_{j=1}^{n} Y_j \tag{1.60}$$

where the stochastic variables are mutually stochastic independent. All factors are assumed of equal importance to the product. The natural logarithm of P_n then becomes

$$\ln P_n = \sum_{j=1}^{n} \ln Y_j \tag{1.161}$$

If n is sufficiently large it follows from the central limit theorem that

$$\ln P_n \sim N \left(\sum_{j=1}^{n} E[\ln Y_j], \sum_{j=1}^{n} Var[\ln Y_j] \right) \tag{1.162}$$

so

$$P_n \sim LN \left(\sum_{j=1}^{n} E[\ln Y_j], \sum_{j=1}^{n} Var[\ln Y_j] \right) \tag{1.163}$$

Models made up of a product of independent contributions are also widely used in engineering.

Example 1.14: Bending moment of a simply supported beam

The horizontal simply supported beam of length l in figure 1.15 is loaded by n vertical statical loads $P_1, P_2, ..., P_n$, which are modelled as mutually independent stochastic loads with mean values μ_{P_i} and variances $\sigma^2_{P_i}$. The load P_i is acting the distance x_i from the left end. The bending moment at the distance y with sign defined in figure 1.15 can be written

$$M(y) = \sum_{i=1}^{n} h(y, x_i) P_i \qquad (1.164)$$
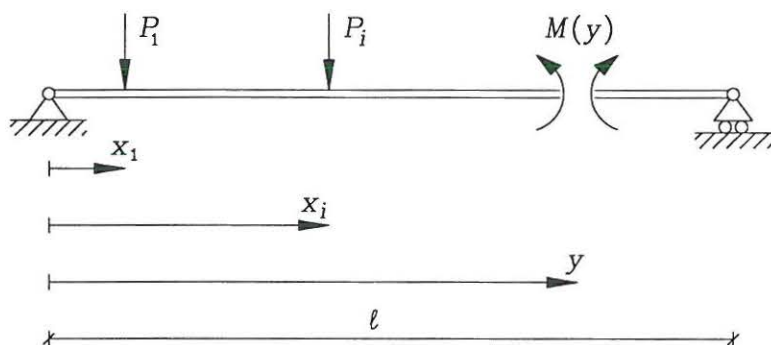
Figure 1.15: Simply supported beam loaded by n independent random loads.

The influence function, $h(y, x)$, providing the bending moment at the position y from a unit load $P = 1$ at x, is given by

$$h(y, x) = \begin{cases} (l - y)x & , x < y \\ (l - x)y & , x > y \end{cases} \qquad (1.165)$$

The expected value $\mu_{M(y)}$ and the variance $\sigma^2_{M(y)}$ of $M(y)$ become

$$\mu_{M(y)} = \sum_{i=1}^{n} h(y, x_i) \mu_{P_i} \qquad (1.166)$$

$$\sigma^2_{M(y)} = \sum_{i=1}^{n} h^2(y, x_i) \sigma^2_{P_i} \qquad (1.167)$$

If $h(y, x_i)\mu_{P_i}$ are comparable for all n loads, each of these contributes comparably to the sum in eq. (1.164). If n is large (say $n > 10$), then $M(y) \sim N\big(\mu_{M(y)}, \sigma^2_{M(y)}\big)$ according to the central limit theorem, even if none of the loads are normally distributed.

2. STOCHASTIC PROCESSES

2.1 Basic concepts

In section 1.1 the concepts of a n-dimensional real stochastic vector $\mathbf{X}^T = [X_1, X_2, ..., X_n]$ and of a stochastic sequence $\{X_1, X_2, ... X_n, ... \}$ were introduced. In both cases sets of indexed stochastic variables defined on the same sample Ω are considered. A *stochastic process* is defined as a natural generalization of these concepts as simply an indexed set $\{X(t), t \in T\}$ of stochastic variables $X(t)$ defined on the same sample Ω. The parameter t is designated the *index parameter* and T the *index set*. This can be

$T = \{1, 2, ..., n\}$: $\{X(t), t \in T\}$ is an n-dimensional stochastic variable

$T = \{1, 2, ..., n, ...\}$: $\{X(t), t \in T\}$ is a stochastic sequence,

                     or an *index discrete stochastic process*

$T \subseteq R^m$: $\{X(t), t \in T\}$ is an *index continuous stochastic process*

The stochastic variables $X(t)$ of the process may be discrete or continuous. Depending on whether the index parameter t and the states of the stochastic variables are discrete or continuous one then get the following 4 categories of stochastic processes, *index discrete state discrete processes, index discrete state continuous processer, index continuous state discrete processes, index continuous state continuous processes*. Physically, t may signify as well time as space, depending of whether a time or spatial dependent phenonomen is observed.

Example 2.1: Stochastic depth of beam
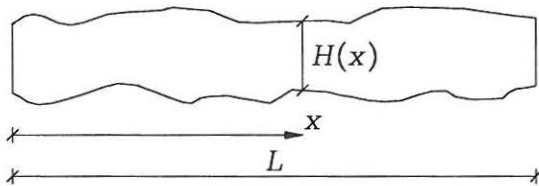
Figure 2.1: Random beam depth process.

Consider the beam of length L with rectangular cross-sections, where the depths are stochastic. At the distance x from one of the beam ends, the depth is modelled as a stochastic variable $H(x)$. The set of stochastic variables from all sections of the beam forms a stochastic process $\{H(x), x \in [0, L]\}$. The sample space Ω is formed of all beams of the length L.

In nature all processes are considered to be continuously dependent on the index parameter t (at least within the framework of classical physics). Hence, such phenonomena should be modelled as index continuous state continuous processes, as indicated in the above example 2.1. However, in practice, we do not manage to measure the depth exact at finite many sections spaced with a certain distance Δx. Further, the depths

cannot be measured below a certain minimum accuracy Δh (say 1 nm), so the observed beam depths will be discrete at the samples $n\Delta h$, $n = 1, 2, \ldots, N$, where $N\Delta h$ is the maximum depth likely to occur. Hence, any engineering application of stochastic processes inevitably involves a discretization of the index set and the sample space, so only index discrete state discrete processes can be handled. Nevertheless, we shall develop the theory also for continuous processes. These are considered the limit of the discretized process, as the number of considered index parameters and the number of discrete sample states goes to infinity.

As in the case of n-dimensional stochastic variables, the probability structure is specified by the joint distribution functions of the stochastic variables in the process. The decisive difference is that the number of stochastic variables may be infinite. With ever increasing accuracy the probability structure of the process can then be described by a sequence of joint distribution functions

$$
\left.
\begin{aligned}
& F_{X(t_1)}(x_1) \\
& F_{X(t_1)X(t_2)}(x_1, x_2) \\
& \vdots \\
& F_{X(t_1)\ldots X(t_n)}(x_1, \ldots, x_n) \\
& \vdots
\end{aligned}
\right\}
\tag{2.1}
$$

for all $t_1, t_2, \ldots t_n, \cdots \in T$. These are named the *distribution functions of the 1st order, the 2nd order*, etc. Of course the nth order distribution functions $F_{X(t_1)\ldots X(t_n)}(x_1, \ldots, x_n)$ depend on which stochastic variables are considered and may consequently be considered to be a function of both index parameters $t_1, \ldots, t_n$ besides the *state variables* $x_1, \ldots, x_n$. Obviously, the knowledge of the nth order distribution functions automatically implies that the distribution function of lower order can be obtained by marginalization, letting some of the state variables go to infinity. As an example the $(n-1)$th order distribution functions are obtained from

$$
F_{X(t_1)\ldots X(t_{n-1})}(x_1, \ldots x_{n-1}) = F_{X(t_1)\ldots X(t_{n-1})X(t_n)}(x_1, \ldots x_{n-1}, \infty)
\tag{2.2}
$$

Eq. (2.2) is known as the *consistency condition* of the sequence (2.2). Further, the nth order distribution function must fulfil the following *symmetry condition* at any order n

$$
F_{X(t_1)\ldots X(t_n)}(x_1, \ldots, x_n) = F_{X(t_{\alpha_1})\ldots X(t_{\alpha_n})}(x_{\alpha_1}, \ldots, x_{\alpha_n})
\tag{2.3}
$$

where $\alpha_1, \ldots, \alpha_n$ is an arbitraty permutation of the numbers $1, 2, \ldots, n$. Eq. (2.3) simply states

$$
P(\{X(t_1) \leq x_1\} \cap \cdots \cap \{X(t_n) \leq x_n\}) =
$$

$$P\big(\{X(t_{\alpha_1}) \le x_{\alpha_1}\} \cap \cdots \cap \{X(t_{\alpha_n}) \le x_{\alpha_n}\}\big) \tag{2.4}$$

Now the following question can be asked. If a family of finite-dimensional distribution functions within the index set T is given, under what conditions does there exist a stochastic process associated with these distribution functions. The answer to this question is contained in the so-called *Kolmogorovs extension theorem*, which states that the necessary and sufficient condition for existence of such a process is that the given family of distributions satisfies the condition of consistency and symmetry. In the case of index discrete processes (n-dimensional stochastic variables and stochastic sequences) this theorem is trivial. In the more subtle case of index continuous processes it is not trivial.

The following notation will be introduced for ease

$$F_{\{X\}}(x_1, t_1; \ldots; x_n, t_n) = F_{X(t_1)\ldots X(t_n)}(x_1, \ldots, x_n) \tag{2.5}$$

In this notation the symmetry property for the 2nd order distributions read

$$F_{\{X\}}(x_1, t_1; x_2, t_2) = F_{\{X\}}(x_2, t_2; x_1, t_1) \tag{2.6}$$

The *probability function of the nth order* for a state discrete process is defined from eq. (1.26)

$$P_{\{X\}}(x_1, t_1; \ldots; x_n, t_n) = P\big(X(t_1) = x_1 \wedge \cdots \wedge X(t_n) = x_n\big) \tag{2.7}$$

The *probability density function of the nth order* for a state continuous process is defined as indicated in eq. (1.27)

$$f_{\{X\}}(x_1, t_1; \ldots; x_n, t_n) = \frac{\partial^n}{\partial x_1 \ldots \partial x_n} F_{\{X\}}(x_1, t_1; \ldots; x_n, t_n) \tag{2.8}$$

When a certain $\omega_n \in \Omega$ is considered, the mapping points $x_n(t) = X(t, \omega_n)$ form a real function defined on T, which is termed a *realization*. If all $\omega \in \Omega$ are considered, a sequence of realizations $\{x_n(t) = X(t, \omega_n)\}$ is obtained from which the probabilistic structure may equally well be defined. In figure 2.2 is shown the realizations, when one and the same physical phenomenon is modelled by stochastic processes of the four categories.

The joint probability structure of the process may alternatively be described by a hieracy of joint stochastic moments of increasing order. In case the following expectations exist the process is defined if the following joint expected values can be defined

$$\mu_{X^{n_1} X^{n_2} \ldots X^{n_n}}(t_1, t_2, \ldots, t_n) = E\left[X^{n_1}(t_1) X^{n_2}(t_2) \cdots X^{n_n}(t_n)\right] \tag{2.9}$$

$\mu_{n_1 n_2 \dots n_n} (t_1, t_2, \dots, t_n)$ which is named the *joint moments of the order* $n_1 + n_2 + \dots + n_n$ must be provided for arbitrary order $n_1, n_2, \dots, n_n$ and arbitrary index permeations $t_1, t_2, \dots t_n$ (i.e. arbitrary stochastic variables $X(t_1), X(t_2), \dots, X(t_n)$). Further, the moments should be known no matter the number n of considered stochastic variables. Of specific interest is the following lower order joint moments.

The *mean value function* $\mu_X(t)$ specifies the expected value of the stochastic value $X(t)$

$$\mu_X(t) = E[X(t)] \tag{2.10}$$

The *auto-correlation function* $\mu_{XX}(t_1, t_2)$ specifies the joint moment of the stochastic variables $X(t_1)$ and $X(t_2)$

$$\mu_{XX}(t_1, t_2) = E[X(t_1)X(t_2)] \tag{2.11}$$

The *auto-covariance function* $\kappa_{XX}(t_1, t_2)$ specifies the covariance between the stochastic variables $X(t_1)$ and $X(t_2)$, cf. eq. (1.89)
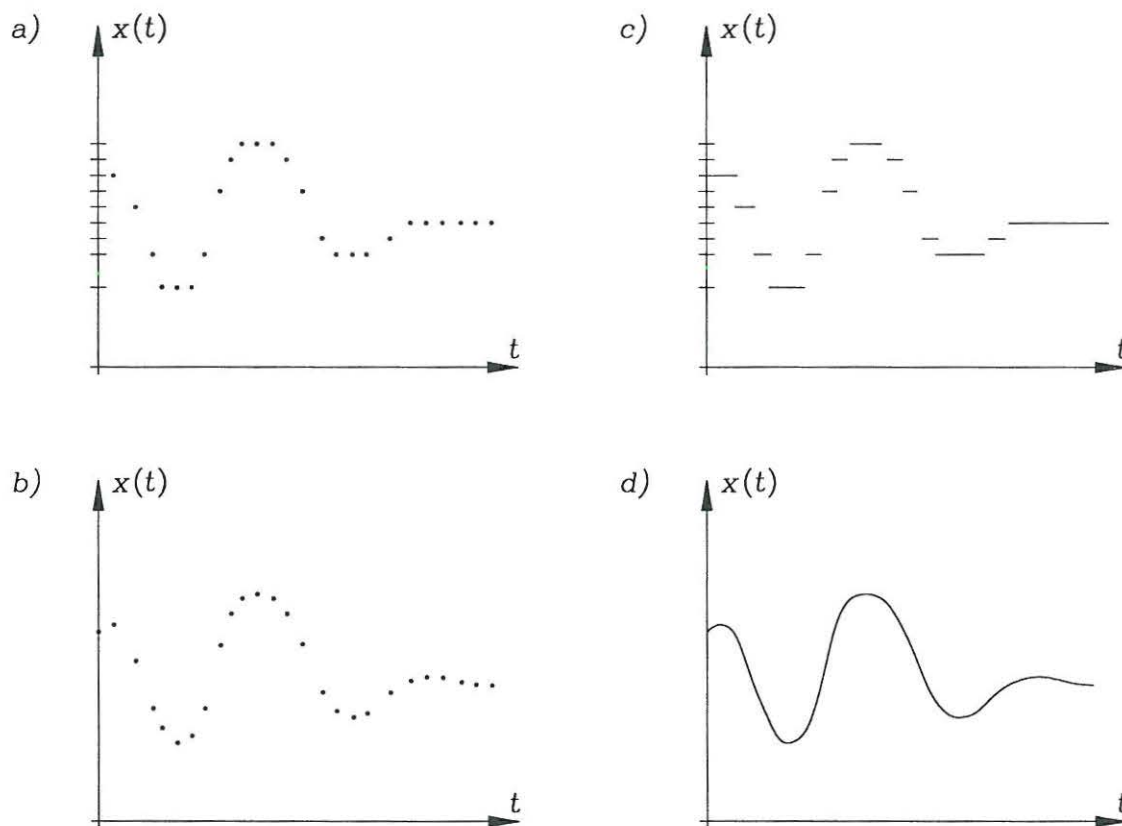

Figure 2-2: Realizations of stochastic processes. a) Index discrete state discrete process. b) Index discrete state continuous process. c) Index continuous state discrete process. d) Index continuous state continuous process.

$$\kappa_{XX}(t_1,t_2) = E\big[(X(t_1)-\mu_X(t_1))(X(t_2)-\mu_X(t_2)) = \mu_{XX}(t_1,t_2)-\mu_X(t_1)\mu_X(t_2)\big] \quad (2.12)$$

Due to the symmetry property as indicated by eq. (2.3) it follows that

$$\kappa_{XX}(t_1,t_2) = \kappa_{XX}(t_2,t_1) \tag{2.13}$$

The *variance function* $\sigma_X^2(t)$ is obtained from eq. (2.11) for $t_1 = t_2$

$$\sigma_X^2(t) = E\left[(X(t)-\mu_X(t))^2\right] \tag{2.14}$$

where $\sigma_X(t)$ is the *standard deviation function*. The *auto-correlation coefficient function* $\rho_{XX}(t_1,t_2)$ specifies the correlation coefficient of $X(t_1)$ and $X(t_2)$, cf. eq. (1.90)

$$\rho_{XX}(t_1,t_2) = \frac{\kappa_{XX}(t_1,t_2)}{\sigma_X(t_1)\sigma_X(t_2)} \tag{2.15}$$

Specification of the stochastic structure based on eq. (2.9) may render itself useless because the indicated moments cannot be estimated from the available data. Hence, there is no need to put such strong limitations on the mathematical model (the stochastic process) to ensure the existence of joint moments of arbitrary order. Normally, no more that the mean value function and the atuo-covariance function can be synthesized from the measurement. A stochastic process for which the moments $E[X^2(t)]$ exist for all $t \in T$ is called a 2nd order stochastic process. For such processes one can prove the following theorem.

Theorem 2.1: The existence of the mean value function and the auto-covariance function is garanteed for a 2nd order stochastic process

Proof:

The existence of the mean value function follows from the Schwarz inequality

$$\begin{aligned}
\mid \mu_X(t) \mid &=\mid E[X(t)] \mid=\mid E[X(t)\cdot 1] \mid \\
&\le \left(E[X^2(t)]E[1^2]\right)^{\frac{1}{2}} = E[X^2(t)]^{\frac{1}{2}} < \infty
\end{aligned} \tag{2.16}$$

In the same way the existence of the auto-correlation function is proved

$$\mid \mu_{XX}(t_1,t_2) \mid=\mid E[X(t_1)X(t_2 t)] \mid \le E[X^2(t_1)]^{\frac{1}{2}} E[X^2(t_2)]^{\frac{1}{2}} < \infty \tag{2.17}$$

The existence of the auto-covariance function then follows from eq. (2.12).

In principle any 2nd order stochastic process which has been calibrated to certain measured mean value functions and auto-covariance functions will do. However, at the modelling process further knowledge of the physics should be invoked. For instance, a certain candidate should be abandoned if it has the measured mean value function $\mu_H(x)$ and autocovariance function $\kappa_{HH}(x_1, x_2)$ of the height along the beam in example 2.1, but produces negative beam depth with a certain significant probability.

A *stochastic vector process* is a set of N-dimensional stochastic vectors $\{\mathbf{X}(t), t \in T\}$, $\mathbf{X}^T(t) = [X_1(t), \ldots, X_N(t)]$, where all stochastic processes are defined on the same sample space. A vector process can be considered a double indexed set of stochastic variables, when the elements T as well as the component indices form the index parameters. In the same manner as for scalar processes, the stochastic structure of a vector process is specified by a sequence of distribution functions

$$\left. \begin{array}{l} F_{X_{i_1}(t_1)}(x_1) \\[2mm] F_{X_{i_1}(t_1)X_{i_2}(t_2)}(x_1, x_2) \\[1mm] \vdots \\[2mm] F_{X_{i_1}(t_1)\ldots X_{i_n}(t_n)}(x_1, \ldots, x_n) \\[1mm] \vdots \end{array} \right\} \qquad (2.18)$$

for all $t_1, t_2, \ldots, t_n \in T$ and all $i_1, i_2, \ldots, i_N \in \{1, 2, \ldots, N\}$. The *distribution functions of the nth order* depend of the considered stochastic variables $X_{i_1}(t_1), \ldots, X_{i_n}(t_n)$. Besides on the state variables $x_1, \ldots, x_n$, these depend of the index parameters $t_1, \ldots, t_n$. The following notation will be applied

$$F_{\{\mathbf{X}\}}(x_1, t_1, i_1; \ldots; x_n, t_n, i_n) = F_{X_{i_1}(t_1)\ldots X_{i_n}(t_n)}(x_1, \ldots, x_n) \qquad (2.19)$$

The symmetry property of the distribution functions of the 2nd order, eq. (2.6) takes the form

$$F_{\{\mathbf{X}\}}(x_1, t_1, i_1; x_2, t_2, i_2) = F_{\{\mathbf{X}\}}(x_2, t_2, i_2; x_1, t_1, i_1) \qquad (2.20)$$

The probability function of the nth order for a state discrete vector process and the probability density function of the nth order for a state continuous vector process are designated, cf. (1.34), (1.35)

$$P_{\{\mathbf{X}\}}(x_1, t_1, i_1; \ldots; x_n, t_n, i_n) = P(X_{i_1}(t_1) = x_1 \wedge \cdots \wedge X_{i_n}(t_n) = x_n) \qquad (2.21)$$

$$f_{\{\mathbf{X}\}}(x_1, t_1, i_1; \ldots; x_n, t_n, i_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F_{\{\mathbf{X}\}}(x_1, t_1, i_1; \ldots; x_n, t_n, i_n) \qquad (2.22)$$

A stochastic vector process $\{\mathbf{X}(t), t \in T\}$ is a 2nd order stochastic process if all the coordinate processes $\{X_i(t), t \in T\}$ are 2nd order stochastic processes.

For vector processes, the prefix "cross" is usually applied to characterize the statistical moments, when stochastic variables from different coordinate processes are considered. The prefix "auto" is reserved to cases, where statistical moments for the same coordinate process $\{X_i(t), t \in T\}$ are considered.

The mean value function $\mu_{X_i}(t)$ specifies the expected value of the stochastic variable $X_i(t)$ of the ith order coordinate process

$$\mu_{X_i}(t) = E[X_i(t)] \tag{2.23}$$

The *cross-correlation function* $\mu_{X_{i_1} X_{i_2}}(t_1 t_2)$ specifies the joint moment of the stochastic variables $X_{i_1}(t_1)$ and $X_{i_2}(t_2)$ of the i_1th and i_2th coordinate process

$$\mu_{X_{i_1} X_{i_2}}(t_1 t_2) = E[X_{i_1}(t_1) X_{i_2}(t_2)] \tag{2.24}$$

The *cross-covariance function* $\kappa_{X_{i_1} X_{i_2}}(t_1, t_2)$ specifies the covariance between the stochastic variables $X_{i_1}(t_1)$ and $X_{i_2}(t_2)$, i.e.

$$\kappa_{X_{i_1} X_{i_2}} = E\left[\left(X_{i_1}(t_1) - \mu_{X_{i_1}}(t_1)\right)\left(X_{i_2}(t_2) - \mu_{X_{i_2}}(t_2)\right)\right] \tag{2.25}$$

Due to eq. (2.20) the following symmetry property applies

$$\kappa_{X_{i_1} X_{i_2}}(t_1, t_2) = \kappa_{X_{i_2} X_{i_1}}(t_2, t_1) \tag{2.26}$$

The variance function of the i_1th coordinate process $\sigma^2_{X_{i_1}}(t)$ is obtained from eq. (2.25) for $t_1 = t_2 = t$ and $i_1 = i_2$. $\sigma_{X_{i_1}}(t)$ is the standard deviation function of the i_1th coordinate process.

The *cross-correlation coefficient function* $\rho_{X_{i_1} X_{i_2}}(t_1, t_2)$ specifies the correlation coefficient of $X_{i_1}(t_1)$ and $X_{i_2}(t_2)$, cf eq. (2.15)

$$\rho_{X_{i_1} X_{i_2}}(t_1, t_2) = \frac{\kappa_{X_{i_1} X_{i_2}}(t_1, t_2)}{\sigma_{X_{i_1}}(t_1)\sigma_{X_{i_2}}(t_2)} \tag{2.27}$$

Example 2.2: Stochastic depth and width of beam

For the beam of example 2.1 the width $W(x)$ at position x is also considered as a stochastic variable. Then $\{\mathbf{X}(x), x \in [0, L]\}, \mathbf{X}^T(x) = [H(x), W(x)]$ forms a 2-dimensional vector process defined on the set Ω of beams of length L with rectangular cross-section. The symmetry property eq. (2.26) reads

$$\kappa_{HW}(x_1, x_2) = \kappa_{WH}(x_2, x_1) \tag{2.28}$$

The left-hand side as well as the right-hand side of eq. (2.28) specify the covariance between the stochastic variables $H(x_1)$ and $W(x_2)$.

The process $\{X(t), t \in T\}$ is called a *Gaussian process*, if the n-dimensional stochastic vector $\mathbf{X}^T = [X(t_1), \ldots, X(t_n)]$ is normally distributed for any order n and arbitrary index parameters $t_1, \ldots, t_n \in T$, i.e. the probability density function $f_{\{X\}}(x_1, t_1; x_2, t_2, \ldots, x_n, t_n)$ is given by eq. (1.37). The vector of expected values $\boldsymbol{\mu_X}$ and covariance matrix $\mathbf{C_{XX}}$ of $\mathbf{X}$ can be specified from the mean value functions $\mu_X(t)$ and the auto-covariance function $\kappa_{XX}(t_1, t_2)$. Hence, a Gaussian process is completely described by its mean value function and auto-covariance function.

Example 2.3: Harmonic processes

Let $R \sim R\left(\sigma^2\right)$, $\Theta \sim U\left(0, 2\pi\right)$ be mutually independent stochastic variables. Next consider the stochastic process $\{X\left(t\right), t \in R\}$ where

$$X(t) = R\cos\left(\omega_0 t + \Theta\right) = R\cos\Theta\cos\left(\omega_0 t\right) - R\sin\Theta\sin\left(\omega_0 t\right) \tag{2.29}$$

and $\omega_0 \in R_+$ is a deterministic constant. Eq. (2.29) which will be referred to as a *harmonic process*, is remarkable in the sense that it is generated from only two stochastic variables R and Θ.

The mean value function $\mu_X\left(t\right)$ and the auto-covariance function $\kappa_{XX}\left(t_1, t_2\right)$ becomes

$$\mu_X(t) = E\left[R\cos(\omega_0 t + \Theta)\right] = E[R]E\left[cos\left(\omega_0 t + \Theta\right)\right]$$

$$= \sqrt{\frac{\pi}{2}}\sigma \cdot 0 \equiv 0 \tag{2.30}$$

$$\kappa_{XX}(t_1, t_2) = E\left[R^2\cos\left(\omega_0 t_1 + \Theta\right)\cos\left(\omega_0 t_2 + \Theta\right)\right]$$

$$= E\left[R^2\right]E\left[\cos\left(\omega_0 t_1 + \Theta\right)\cos\left(\omega_0 t_2 + \Theta\right)\right]$$

$$= E\left[R^2\right]E\left[\frac{1}{2}\cos(\omega_0(t_2 - t_1)) + \frac{1}{2}\cos(\omega_0(t_2 + t_1) + 2\Theta)\right]$$

$$= \sigma^2\cos(\omega_0(t_2 - t_1)) \tag{2.31}$$

where it has been used

$$E[R] = \int\limits_0^\infty r\frac{r}{\sigma^2}\exp\left(-\frac{r^2}{2\sigma^2}\right)dr = \sqrt{\frac{\pi}{2}}\sigma \tag{2.32}$$

$$E[R^2] = \int\limits_0^\infty r^2\frac{r}{\sigma^2}\exp\left(-\frac{r^2}{2\sigma^2}\right)dr = 2\sigma^2 \tag{2.33}$$

$$E\left[\cos\left(\omega_0 t + \Theta\right)\right] = \int\limits_0^{2\pi}\cos(\omega_0 t + \theta)\frac{1}{2\pi}d\theta = 0 \tag{2.34}$$

$$E\left[\cos(\omega_0(t_1 + t_2) + 2\Theta)\right] = \int\limits_0^{2\pi}\cos\left(\omega_0(t_1 + t_2) + 2\theta\right)\frac{1}{2\pi}d\theta = 0 \tag{2.35}$$

From example 1.8 follows $Y_1 = R\cos\Theta \sim N\left(0, 1\right)$ and $Y_2 = R\sin\Theta \sim N\left(0, 1\right)$, so $X\left(t\right) = Y_1\cos\left(\omega_0 t\right) - Y_2\sin\left(\omega_0 t\right)$. Being a sum of 2 normally distributed variables it follows that $X\left(t\right)$ is normally distributed too, $X\left(t\right) \sim N\left(\mu_X\left(t\right), \sigma_X^2\left(t\right)\right)$. The mean value, $\mu_X\left(t\right) \equiv 0$ according to eq. (2.30). The variance, $\sigma_X^2(t)$, follows from eq. (2.31) for $t_1 = t_2 = t$, so $\sigma_X^2(t) = \sigma^2$. However, this does *not* imply that $\{X(t), t \in R\}$ is a Gaussian process. Actually, none of the higher order joint distributons are normal. In order to see this consider $X\left(t_1\right), X\left(t_2\right), X\left(t_3\right)$ for arbitrary t_1, t_2, t_3. Assume $X\left(t_2\right) = x_2$ and $X\left(t_3\right) = x_3$. These samples are then generated by the samples $R = r_1$, $\Theta = \theta_1$ of the basic variables of the process fulfilling

$$\left.\begin{array}{l} x_2 = r_1\cos\left(\omega_0 t_2 + \theta_1\right) \\ x_3 = r_1\cos\left(\omega_0 t_3 + \theta_1\right) \end{array}\right\} \tag{2.36}$$

The probability distribution of $X(t_1)$ on condition of $X(t_2) = x_2$ and $X(t_3) = x_3$ is then discrete at the single sample $x_1 = r_1 \cos(\omega_0 t_1 + \theta_1)$, with (r_1, θ_1) determined from eq. (2.36). If $[X(t_1), X(t_2), X(t_3)]$ had been jointly normal, the distribution of $X(t_1)$ on condition of $X(t_2)$ and $X(t_3)$ had been normal, cf. eq. (1.110).

Next consider the even more simple stochastic process $\{X(t), t \in R\}$ where

$$X(t) = a\cos(\omega_0 t + \Theta) \tag{2.37}$$

where $a, \omega_0 : \in R_+$ are deterministic constants and $\Theta \sim U(0, 2\pi)$. (2.37) will be referred to as a *random phase process*. From example 1.6 follows that the 1st order pdf of $X(t)$ is given by

$$f_{\{X\}}(x,t) = \begin{cases} 0 & ,\mid x \mid > a \\ \dfrac{1}{\pi}\dfrac{1}{\sqrt{a^2 - x^2}} & ,\mid x \mid < a \end{cases} \tag{2.38}$$

The 1st order pdf is now far from being normally distributed. The mean value function and the auto-covariance function of the process become

$$\mu_X(t) = E\left[a\cos(\omega_0 t + \Theta)\right] = a\int\limits_0^{2\pi} \cos(\omega_0 t + \theta)\frac{1}{2\pi}d\theta \equiv 0 \tag{2.39}$$

$$\kappa_{XX}(t_1, t_2) = E\left[a\cos(\omega_0 t_1 + \Theta)\, a\cos(\omega_0 t_2 + \Theta)\right]$$

$$a^2 \int\limits_0^{2\pi} \cos(\omega_0 t_1 + \theta)\cos(\omega_0 t_2 + \theta)\frac{1}{2\pi}d\theta$$

$$a^2 \int\limits_0^{2\pi} \left(\frac{1}{2}\cos(\omega_0(t_1 - t_2)) + \frac{1}{2}\cos(\omega_0(t_1 + t_2) + 2\theta)\right)\frac{1}{2\pi}d\theta$$

$$\frac{a^2}{2}\cos(\omega_0(t_1 - t_2)) \tag{2.40}$$

Example 2.4: Sums of harmonic processes

As an extension to eq. (2.29) consider the following process $\{X(t), t \in R\}$ made up of a finite sum of harmonic processes

$$X(t) = \sum_{j=1}^{N} R_j \cos(\omega_j t + \Theta_j) \tag{2.41}$$

where $\omega_j \in R_+$ are deterministic constants. The random variables $R_j \sim R\left(\sigma_j^2\right)$ and $\Theta_j \sim U(0, 2\pi)$ are mutually stochastic independent. The mean value function and auto-covariance function of eq. (2.41) becomes

$$\mu_X(t) = \sum_{j=1}^{N} E\left[R_j \cos(\omega_j t + \Theta_j)\right] = \sum_{j=1}^{N} E\left[R_j\right] E\left[\cos(\omega_j t + \Theta_j)\right] \equiv 0 \tag{2.42}$$

$$\kappa_{XX}(t_1, t_2) = E\left[X(t_1)X(t_2)\right]$$

$$= \sum_{j=1}^{N}\sum_{k=1}^{N} E\left[R_j R_k \cos\left(\omega_j t_1 + \Theta_j\right)\cos\left(\omega_k t_2 + \Theta_k\right)\right]$$

$$= \sum_{j=1}^{N} E\left[R_j^2\right] E\left[\cos\left(\omega_j t_1 + \Theta_j\right)\cos\left(\omega_j t_2 + \Theta_j\right)\right]$$

$$\sum_{\substack{j=1 \\ }}^{N}\sum_{\substack{k=1 \\ k\neq j}}^{N} E\left[R_j\right] E\left[R_k\right] E\left[\cos\left(\omega_j t_1 + \Theta_j\right)\right] E\left[\cos\left(\omega_k t_2 + \Theta_k\right)\right]$$

$$= \sum_{j=1}^{N} E\left[R_j^2\right] E\left[\frac{1}{2}\cos(\omega_j(t_2 - t_1)) + \frac{1}{2}\cos(\omega_j(t_2 + t_1) + 2\Theta_j)\right]$$

$$= \sum_{j=1}^{N} \sigma_j^2 \cos(\omega_j(t_2 - t_1)) \tag{2.43}$$

At the derivation of eq. (2.43) the mutual stochastic independency of R_j, R_k, Θ_j, Θ_k for $k \neq j$ has been used, so

$$E\left[R_j R_k \cos\left(\omega_j t_1 + \Theta_j\right)\cos\left(\omega_k t_2 + \Theta_k\right)\right] =$$

$$E\left[R_j\right] E\left[R_k\right] E\left[\cos\left(\omega_j t_1 + \Theta_j\right)\right] E\left[\cos\left(\omega_k t_2 + \Theta_k\right)\right] =$$

$$E\left[R_j\right] E\left[R_k\right] \cdot 0 \cdot 0 = 0 \tag{2.44}$$

Since, $R_j \cos\left(\omega_j + \Theta_j\right) \sim N\left(0, \sigma_j^2\right)$, $X(t)$ is made up of normally distributed variables, and then is normally distributed itself

$$X(t) \sim N(\mu_X(t), \sigma_X^2(t)) = N\left(0, \sum_{j=1}^{N} \sigma_j^2\right) \tag{2.45}$$

Still, this does not mean that $\{X(t), t \in R\}$ is a Gaussian process. However, if the number of terms, N, in the sum (2.41) is sufficiently large, and all addends contribute equally, which will be the case if the variances, σ_j^2, are all of comparable magnitude, it can be argued from the central limit theorem that the distribution of any finite dimensional subset of stochastic variables $\mathbf{X}^T = [X(t_1), \ldots, X(t_n)]$ extracted from the process will be normal as the number of terms grows to infinity. Hence, eq. (2.41) will asymptotically approach a Gaussian process.

Finally, consider the process $\{X(t), t \in R\}$ made up of a finite sum of random phase processes

$$X(t) = \sum_{j=1}^{N} a_j \cos(\omega_j t + \Theta_j) \tag{2.46}$$

where $a_j, \omega_j \in R_+$ are deterministic constants and $\Theta_j \sim U(0, 2\pi)$ are mutually stochastic independent. Using almost identical derivations as leading to eqs. (2.42) and (2.43) the mean value function and the auto-covariance function become, cf. eq. (2.40)

$$\mu_X(t) \equiv 0 \tag{2.47}$$

$$\kappa_{XX}(t_1, t_2) = \sum_{j=1}^{N} \frac{a_j^2}{2} \cos(\omega_j (t_2 - t_1)) \tag{2.48}$$

If all addends of eq. (2.46) are of equally importance, it can again be argued from the central limit theorem that the random phase process, eq. (2.46) approaches a Gaussian process as the number of terms, N, grows to infinity. The rate of convergence is generally believed to be slower than of eq. (2.41).

2.2 Continuity, differentiability and integration in mean square

First, the concepts of continuity, differentiability and integration of an ordinary function $x : T \to R$ defined on some interval $T = [a, b]$ are resumed.
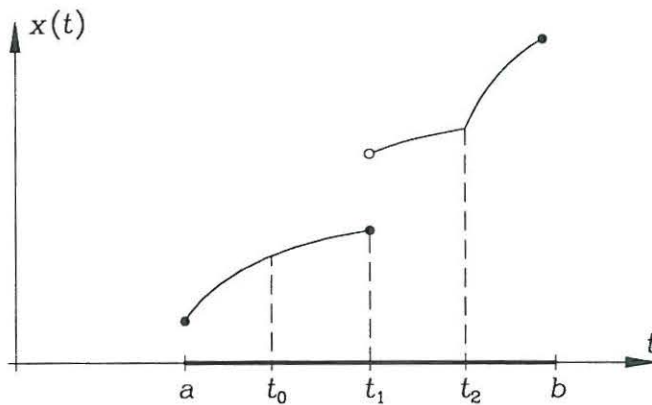

Figure 2.3: Discontinuous, non-differentiable, integrable real function

The function $x(t)$ is said to be *continuous* in some point $t_0 \in T$, if for any sequence of index values $t \in T$ the corresponding sequence of function values $\{x(t)\}$ converges to the value $x(t_0)$

$$\lim_{t \to t_0} x(t) = x(t_0) \tag{2.49}$$

The meaning of (2.49) is that the graph of $x(t)$ does not make a jump at the point t_0, so the limit of the sequence of function values $\{x(t)\}$ is the same, whether $t \to t_0$ from above or below. Obviously, the function with the graph shown in figure 2.3 is not continuous in the point t_1. The function $x(t)$ is said to be *continuous in T* if it is continuous for all $t \in T$.

The function $x(t)$ is said to be differentiable in some point t_0 if for any sequence of index values $t \in T$ the corresponding sequence of difference quotients converges to some limit $\frac{d}{dt} x(t_0)$

$$\lim_{t \to t_0} \frac{x(t) - x(t_0)}{t - t_0} = \frac{d}{dt} x(t_0) \tag{2.50}$$

The difference quotient on the left-hand side of (2.50) signifies the slope of the secant. These slopes are required to converge to one and the same limit, $\frac{d}{dt}x(t_0)$, the slope of the tangent, whether $t \to t_0$ from above or from below. If not so, the graph of $x(t)$ makes a break at the point t_0. As seen in figure 2.3, $x(t)$ is not differentiable in the point t_2. Since differentiability at a point t_0 only makes sense, if it is continuous in this point, it is not differentiable in point t_1 either. The function $x(t)$ is said to be *differentiable in* T if it is differentiable for all $t \in T$.

Next, divide the interval $[a, b]$ into n subintervals $\Delta t_i = t_i - t_{i-1}$, $t_0 = a$, $t_n = b$, and select some point $t_i^{\star} \in]t_{i-1}, t_i[$ in the interior of each of these intervals. The area, A, below the graph of $x(t)$ can then approximately be calculated by the sum

$$A_n \simeq \sum_{i=1}^{n} x(t_i^{\star})\Delta t_i, \quad \Delta t_i = t_i - t_{i-1} \tag{2.51}$$

Now, let the number of intervals $n \to \infty$, so $\max_{i=1,n} \Delta t_i \to 0$. For instance, this is obtained if $\Delta t_i = \frac{b-a}{n}$ for all intervals. If the sequence of corresponding areas $\{A_n\}$ converges to some limit A as $n \to \infty$, which is the searched area, no matter how the internal points $t_i^{\star}$ are selected within the sub-intervals, the function $x(t)$ is said to be *Riemann integrable* over the interval $[a, b]$, and (2.51) is called a *Riemann sum*. Any function, which makes a finite number of finite jumps in the interval T is Riemann integrable. So, this is also the case for the function $x(t)$ shown in figure 2.3.

Since the concepts of continuity, differentiability and integrability involves a continuous change of index parameters, it is obvious that one has to confine oneself to index continuous stochastic processes $\{X(t), t \in T\}$, where $T = [a, b]$ is some interval, if these concepts are to be extended to stochastic processes. On the other hand the stochastic processes may be as well state continuous as state discrete.

Let E be the event (i.e. the subset of elementary events $\omega \in \Omega$) for which the corresponding realizations $x(t) = X(\omega, t)$ are either continuous, differentiable or Riemann integrable in the ordinary sense as described above. Continuity and differentiability may either refer to a certain fixed index point $t_0 \in T$ or to all index values in T. If the probability $P(E)$ of the event E is equal to 1, the process $\{X(t), t \in T\}$ is said to be continuous, differentiable or Riemann integrable with *probability 1* or *almost certain*. Again, this does not mean that all the realizations $x(t) = X(\omega, t)$ are continuous, differentiable or Riemann integrable. Only, that non-continuous, non-differentiable or non-Riemann integrable realizations occur with a relative frequency within a set of realizations, which goes to zero as the number of realizations within the considered set goes to infinity.

Example 2.5 Realization of a continuous, non-differentiable stochastic process
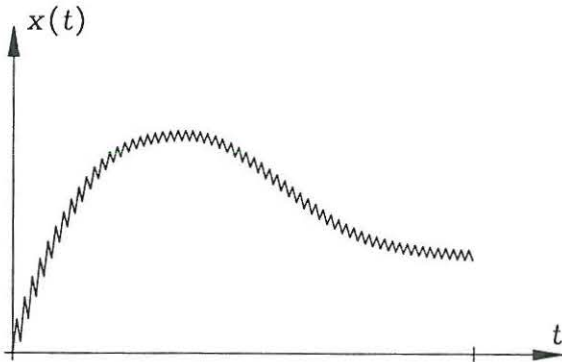

Figure 2.4: Realization of Wiener process

In figure 2.4 is shown a typical realization of a somewhat spurious process $\{X(t), t \in [0, b]\}$ with $T = [0, b]$ and where $X(0) = 0$ with probability 1. The ripples on the top of the realization are assumed to have zero wave height and zero wave length. Then, the realization is continuous. However, the tangents are everywhere vertical, so the differential coefficient changes discontinuously from $-\infty$ to $+\infty$. Obviously, a derivative does not exist. If almost all realizations of $\{X(t), t \in [0, b]\}$ have the indicated properties, $\{X(t), t \in [0, b]\}$ is continuous and non-differentiable for all $t \in [0, b]$ with probability 1. If further $\{X(t), t \in [0, b]\}$ is Gaussian the considered process is a so-called *Wiener process*, which as will later be seen, forms an important brick in the modelling of many stochastic phenomena.

Any phenomenon within the realm of structural mechanics is continuous and differentiable in time and space. Hence, these phenomena should be modelled by stochastic processes which possess these analytical properties with probability 1. Since the existence of such properties of a certain stochastic model normally requires information which goes beyond the mean value function and the auto-covariance function, and since these moments are often the only information available about the process, one has tried to define the concepts of continuity, differentiability and integrability as limits of certain sequences of stochastic variables in mean square instead of limits with probability 1. The idea is that the existence of such limits in mean square can be verified entirely from the analytical properties of the auto-correlation function $\mu_{XX}(t_1, t_2)$.

2.2.1 Continuity in mean square

A stochastic process $\{X(t), t \in [a, b]\}$ is said to be *continuous in mean square* at the index value $t_0 \in [a, b]$ if the stochastic sequence $\{X(t)\}$ converges to some limit stochastic variable $X(t_0)$ in mean square as $t \to t_0$, cf. (1.121)

$$\lim_{t \to t_0} E\left[\left(X(t) - X(t_0)\right)^2\right] = 0 \tag{2.52}$$

The process is said to be *continuous in mean square in* T if a limit $X(t_0)$ exists, fulfilling (2.52) for all $t_0 \in T$.

From the mean square convergence theorem follows that continuity in mean square is guaranteed if and only if the correlation function $\mu_{XX}(t_1, t_2)$ exists and has a limit as $t_1, t_2 \to t_0$ independently, i.e. if

$$\lim_{\substack{t_1 \to t_0 \\ t_2 \to t_0}} E[X(t_1)X(t_2)] = \lim_{\substack{t_1 \to t_0 \\ t_2 \to t_0}} \mu_{XX}(t_1, t_2) = \mu_{XX}(t_0, t_0) = E[X^2(t_0)] \qquad (2.53)$$

(2.53) states that the process $\{X(t), t \in [a, b]\}$ is continuous in mean square at $t = t_0$ if and only if the auto-correlation $\mu_{XX}(t_1, t_2)$ exists and is continuous at the diagonal $t_1 = t_2 = t_0$. Further, the process is said to be continuous in mean square in T, if this property is fulfilled for all $t_0 \in T$. Obviously, this is the case if $\{X(t), t \in [a, b]\}$ is a 2nd order process.

Example 2.6: Continuity in mean square of the Wiener process and the Poisson counting process
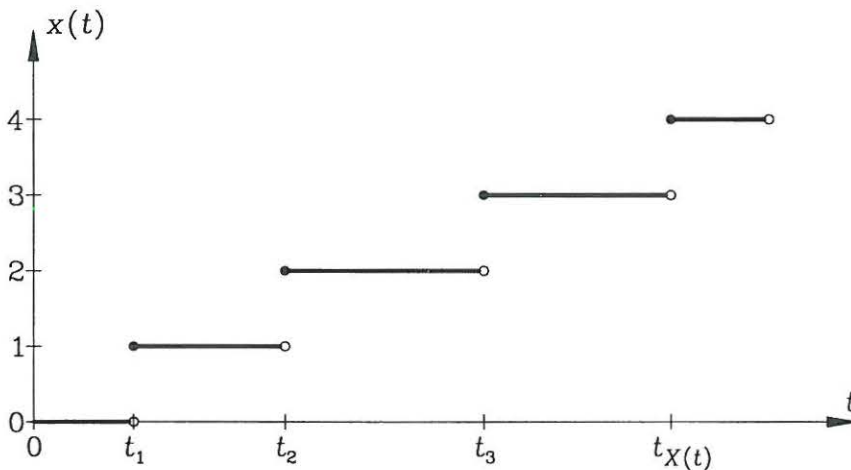

Figure 2.5: Realization of Poisson counting process.

A traffic light at a crossroad switches to red at a time, which may be selected as $t = 0$. The first car arrives to the red light at the random time $t = t_1$, the next car arrives at the time $t = t_2$, etc. As long as the traffic light has not switched back to green these cars make a queue and one may be interesting in knowing the number of cars $X(t)$, which have arrived after some elapsed time t, in order to design the length of the carriageway marking. Obviously, $X(t)$ is a stochastic variable, which changes with the elapsed time t. Then the set of such stochastic variables $\{X(t), t \in \,]0, \infty[\}$ indicating the number of waiting cars at variable instants of times forms per definition a stochastic process, which is called a *counting process*. The realizations of a counting process is piece-wise constant and jumps with the magnitude 1 each time a new event is registrated (*in casu* a new car arrives).

If further the times $t_1 \ldots t_{X(t)}$ of arrival of the cars to the crossroads are independent events, which is often assumed to be the case in little traffic, the counting process is a so-called Poisson counting process, which is another important brick in stochastic modelling. For this process it can be shown

that the auto-correlation function is given by, see section 3.3

$$\mu_{XX}(t_1, t_2) = \begin{cases} 0 & , t_1 \in]-\infty, 0[\ \vee \ t_2 \in]-\infty, 0[\\ \nu \min(t_1, t_2) + \nu^2 t_1 t_2 & , t_1, t_2 \in [0, \infty[, \ \nu \in R_+ \end{cases} \tag{2.54}$$

The corresponding correlation function of the Wiener process $\{X(t), t \in [0, \infty[\}$ with the typical realization as shown in figure 2.4 can be shown to be, see section 3.2

$$\mu_{XX}(t_1, t_2) = \begin{cases} 0 & , t_1 \in]-\infty, 0[\ \vee \ t_2 \in]-\infty, 0[\\ 2\pi S_0 \min(t_1, t_2) & , t_1, t_2 \in [0, \infty[, \ S_0 \in R_+ \end{cases} \tag{2.55}$$

Both correlation functions (2.54) and (2.55) exist and are continuous at the diagonal. Hence, as well the Wiener process as the Poisson counting process are continuous in mean square. However, the Poisson counting process is *not* continuous with probability 1 as seen from the typical realization shown in figure 2.5.

2.2.2 Differentiability in mean square

Consider the following combined stochastic variable

$$Y(t_0, h) = \frac{X(t_0 + h) - X(t_0)}{h} \tag{2.56}$$

where h may be any real number fulfilling $t_0 + h \in [a, b]$ so the stochastic variable $X(t_0 + h)$ is a member of the considered stochastic process. h may be as well positive as negative. Now, consider the limit passing $h \to 0$, in which case a sequence of stochastic variables $\{Y(t_0, h)\}$ is obtained. The stochastic process $\{X(t), t \in [a, b]\}$ is then said to be *differentiable in mean square* at $t_0 \in T$ if the stochastic sequence $\{Y(t_0, h)\}$ converges in mean square as $h \to 0$ to some limit stochastic variable, designated $\frac{d}{dt}X(t_0)$,

$$\lim_{h \to 0} E\left[\left(Y(t_0, h) - \frac{d}{dt}X(t_0)\right)^2\right] = 0 \tag{2.57}$$

The process $\{X(t), t \in [a, b]\}$ is said to be *differentiable in mean square in* T if a limit $\frac{d}{dt}X(t_0)$ exists, fulfilling (2.57) for all $t_0 \in T$. In this case, the set of limit stochastic variables $\{\frac{d}{dt}X(t), t \in [a, b]\}$ is named the *1st order derivative process* of the process $\{X(t), t \in [a, b]\}$.

From the mean square convergence theorem follows that the existence of the mean square derivative $\frac{d}{dt}X(t)$ is guaranteed if and only if $E[Y(t_0, h_1)Y(t_0, h_2)]$ exists and has the same limit no matter how the limit passings $h_1 \to 0$ and $h_2 \to 0$ are performed. Then,

$$\lim_{\substack{h_1 \to 0 \\ h_2 \to 0}} E\left[Y(t_0, h_1)Y(t_0, h_2)\right] =$$

$$\lim_{\substack{h_1 \to 0 \\ h_2 \to 0}} E\left[\frac{1}{h_1}\left(X(t_0 + h_1) - X(t_0)\right) \frac{1}{h_2}\left(X(t_0 + h_2) - X(t_0)\right)\right] =$$

$$\lim_{h_1 \to 0} \frac{1}{h_1} \left(\lim_{h_2 \to 0} \frac{1}{h_2} \left(\mu_{XX}(t_0 + h_1, t_0 + h_2) - \mu_{XX}(t_0 + h_1, t_0) \right) - \right.$$

$$\lim_{h_2 \to 0} \frac{1}{h_2} \left(\mu_{XX}(t_0, t_0 + h_2) - \mu_{XX}(t_0, t_0) \right) \right) =$$

$$\lim_{h_1 \to 0} \frac{1}{h_1} \left(\frac{\partial}{\partial t_2} \mu_{XX}(t_0 + h_1, t_0) - \frac{\partial}{\partial t_2} \mu_{XX}(t_0, t_0) \right) =$$

$$\frac{\partial^2}{\partial t_1 \partial t_2} \mu_{XX}(t_0, t_0) \tag{2.58}$$

(2.58) states that the process $\{X(t), t \in [a, b]\}$ is differentiable in mean square if and only if the 2nd mixed derivative $\frac{\partial^2}{\partial t_1 \partial t_2} \mu_{XX}(t_1, t_2)$ exists and is continuous at the diagonal $t_1 = t_2 = t_0$. The independence of the limit passing of h_1 and h_2 means that the directional derivative of the 2nd order in any direction is required to exist and be continuous at t_0. As stated in the mathematical analysis this is ensured if and only if this property is fulfilled for the mixed derivatives.

Example 2.7: 1st order derivative process of the Wiener process
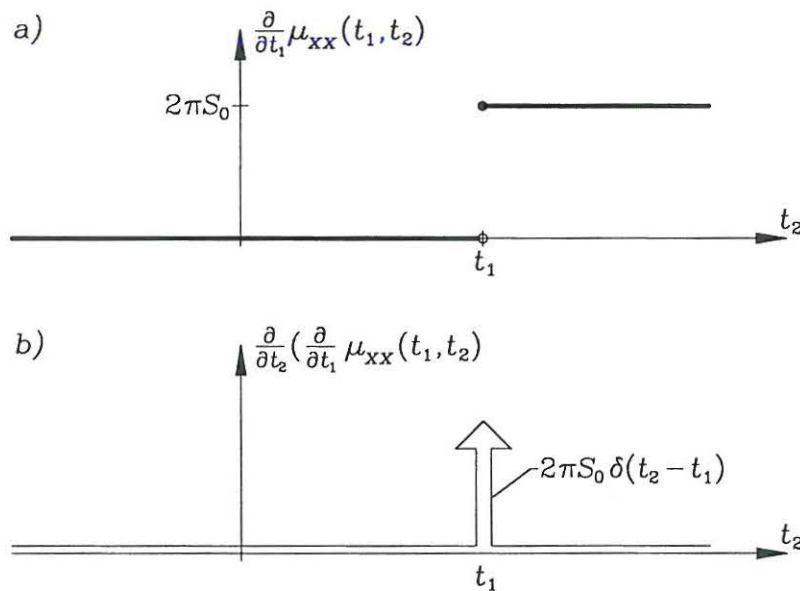

Figure 2.6: 1st and 2nd partial derivative of correlation function of Wiener process.

If almost all realizations of the Wiener process have the appearance as shown in figure 2.4 it is clear that this process is not differentiable with probability 1. Now, let us check whether it is differentiable

in the mean square. For a certain $t_2 > 0$ one has the following variation of the correlation function (2.55) as a function of t_1

$$\mu_{XX}(t_1, t_2) = \begin{cases} 0 & , t_1 < 0 \\ 2\pi S_0 t_1 & , 0 \le t_1 \le t_2 \\ 2\pi S_0 t_2 & , t_1 > t_2 \end{cases} \Rightarrow$$

$$\frac{\partial}{\partial t_1}\mu_{XX}(t_1, t_2) = \begin{cases} 0 & , t_1 < 0 \\ 2\pi S_0 & , 0 \le t_1 \le t_2 \\ 0 & , t_1 > t_2 \end{cases} = \begin{cases} 0 & , t_2 < t_1 \\ 2\pi S_0 & , t_2 \ge t_1 \end{cases} \tag{2.59}$$

In the last statement of (2.59) the intervals are specified in terms of t_2 instead of t_1. $\frac{\partial}{\partial t_1}\mu_{XX}(t_1, t_2)$ as a function of t_2 has been plotted in figure 2.6a.

The mixed derivative $\frac{\partial}{\partial t_2}\left(\frac{\partial}{\partial t_1}\mu_{XX}(t_1, t_2)\right)$ as a function of t_2 then becomes, see figure 2.6b

$$\frac{\partial}{\partial t_2}\left(\frac{\partial}{\partial t_1}\mu_{XX}(t_1, t_2)\right) = 2\pi S_0 \delta(t_2 - t_1) \tag{2.60}$$

$\delta(t)$ is Dirac's delta function, cf. (1.19). (2.60) is zero for $t_1 \ne t_2$ and goes to infinity for $t_1 = t_2$. Hence, (2.60) is far from being continuous at the main diagonal $t_1 = t_2$, and the Wiener process is not differentiable in mean square.

The 1st derivative process in mean square as a model for the derivatives of some physical phenomena is somewhat obscure, since one can not say anything about it's set of realizations. The only way to get familiar with it is to calculate it's joint moments. To shorten the notation the abbreviate $\{X'(t), t \in [a, b]\}$, $X'(t) = \frac{d}{dt}X(t)$ will be adopted for the 1st order derivative process.

The mean value function of the process becomes

$$\mu_{X'}(t) = E\left[X'(t)\right] = E\left[\frac{d}{dt}X(t)\right] = E\left[\lim_{\substack{h \to 0}}\text{-ms}\frac{X(t+h) - X(t)}{h}\right] =$$

$$\lim_{h \to 0} E\left[\frac{X(t+h) - X(t)}{h}\right] = \lim_{h \to 0} \frac{\mu_X(t+h) - \mu_X(t)}{h} = \frac{d}{dt}\mu_X(t) \tag{2.61}$$

At the derivation of (2.61) use has been made of the previously proved assertion that the operations of limit passing in mean square and expectation commute, cf. (1.138). (2.61) can equally be stated that the operations of stochastic differentiation in mean square and expectation commute $\left(E\left[\frac{d}{dt}X(t)\right] = \frac{d}{dt}E[X(t)]\right)$. Despite the suggestive appearance of this statement it should be noticed that $\frac{d}{dt}$ in front of $\frac{d}{dt}X(t)$ is a part of a symbol and not an operator. The statement should then merely be considered as a formal tool for deriving results, which can be proved to be rigorously correct in a more involved manner.

The auto-correlation function becomes

$$\mu_{X'X'}(t_1, t_2) = E\left[X'(t_1)X'(t_2)\right] =$$

$$\lim_{\substack{h_1 \to 0 \\ h_2 \to 0}} E \left[\frac{1}{h_1 h_2} \left(X(t_1 + h_1) - X(t_1) \right) \left(X(t_2 + h_2) - X(t_2) \right) \right] =$$

$$\lim_{h_1 \to 0} \frac{1}{h_1} \left(\lim_{h_2 \to 0} \frac{1}{h_2} E\left[X(t_1 + h_1) X(t_2 + h_2) - X(t_1 + h_1) X(t_2) \right] - \right.$$

$$\left. \lim_{h_2 \to 0} \frac{1}{h_2} E\left[X(t_1) X(t_2 + h_2) - X(t_1) X(t_2) \right] \right) =$$

$$\lim_{h_1 \to 0} \frac{1}{h_1} \left(\lim_{h_2 \to 0} \frac{1}{h_2} \left(\mu_{XX}(t_1 + h_1, t_2 + h_2) - \mu_{XX}(t_1 + h_1, t_2) \right) - \right.$$

$$\left. \lim_{h_2 \to 0} \frac{1}{h_2} \left(\mu_{XX}(t_1, t_2 + h_2) - \mu_{XX}(t_1, t_2) \right) \right) =$$

$$\lim_{h_1 \to 0} \frac{1}{h_1} \left(\frac{\partial}{\partial t_2} \mu_{XX}(t_1 + h_1, t_2) - \frac{\partial}{\partial t_2} \mu_{XX}(t_1, t_2) \right) =$$

$$\frac{\partial^2}{\partial t_1 \partial t_2} \mu_{XX}(t_1, t_2) \tag{2.62}$$

From the Schwarz inequality follows, cf. (1.134)

$$\left| E\left[X'(t_1) X'(t_2) \right] \right| \leq E\left[X'(t_1)^2 \right]^{\frac{1}{2}} E\left[X'(t_2)^2 \right]^{\frac{1}{2}} \tag{2.63}$$

Hence, the moment $\mu_{X'X'}(t_1, t_2)$ exists, if only the moments $E\left[X'(t_1)^2 \right] = \mu_{X'X'}(t_1, t_1)$ and $E\left[X'(t_2)^2 \right] = \mu_{X'X'}(t_2, t_2)$ exist. However, this is guaranteed, if the process is differentiable in mean square, corresponding to the limit (2.58) exists in this case. The existence of the mixed derivative (2.62) is then guaranteed for all t_1, t_2 if only this derivative exists at the main diagonal.

(2.62) can equally be given the suggestive formulation that the operations of stochastic differentiation in mean square and expectation commute, i.e.

$$E \left[\frac{d}{dt} X(t_1) \frac{d}{dt} X(t_2) \right] = \frac{\partial^2}{\partial t_1 \partial t_2} E\left[X(t_1) X(t_2) \right]$$

Combining (2.61) and (2.62) follows that the auto-covariance function of the 1st derivative process, $\kappa_{X'X'}(t_1, t_2)$ is related to the auto-covariance function of the underlying process as follows, cf. (2.12)

$$\kappa_{X'X'}(t_1, t_2) = \mu_{X'X'}(t_1, t_2) - \mu_{X'}(t_1) \mu_{X'}(t_2) =$$

$$\frac{\partial^2}{\partial t_1 \partial t_2} \mu_{XX}(t_1, t_2) - \frac{\partial}{\partial t_1} \mu_X(t_1) \frac{\partial}{\partial t_2} \mu_X(t_2) =$$

$$\frac{\partial^2}{\partial t_1 \partial t_2} (\mu_{XX}(t_1, t_2) - \mu_X(t_1)\mu_X(t_2)) = \frac{\partial^2}{\partial t_1 \partial t_2} \kappa_{XX}(t_1, t_2) \tag{2.64}$$

Generally, joint moments of the nth order are given in the same way as

$$E\left[X'(t_1) \cdots X'(t_n)\right] = \frac{\partial^n}{\partial t_1 \cdots \partial t_n} E\left[X(t_1) \cdots X(t_n)\right] \tag{2.65}$$

Hence, these joint expectations exist if the partial derivative on the right-hand side of (2.65) exists and is continuous.

Next, the *2nd derivative in mean square*, $\{X''(t), t \in [a, b]\}$, of the process $\{X(t), t \in [a, b]\}$ is defined as the 1st derivative of the process $\{X'(t), t \in [a, b]\}$. From (2.58) follows that the existence of the 2nd derivative in mean square is guaranteed if and only if

$$\frac{\partial^2}{\partial t_1 \partial t_2} \mu_{X'X'}(t_0, t_0) = \frac{\partial^4}{\partial t_1^2 \partial t_2^2} \mu_{XX}(t_0, t_0) \tag{2.66}$$

exists and is continuous at the diagonal. Hence, the existence of the 2nd order derivative requires the existence and continuity at the diagonal of the 4th order mixed derivatives of the correlation function of the underlying process.

The mean value function $\mu_{X''}(t)$ and the auto-covariance function $\kappa_{X''X''}(t_1, t_2)$ follows immediately from (2.61) and (2.64)

$$\mu_{X''}(t) = \frac{d}{dt}\mu_{X'}(t) = \frac{d^2}{dt^2}\mu_X(t) \tag{2.67}$$

$$\kappa_{X''X''}(t_1, t_2) = \frac{\partial^2}{\partial t_1 \partial t_2}\kappa_{X'X'}(t_1, t_2) = \frac{\partial^4}{\partial t_1^2 \partial t_2^2}\kappa_{XX}(t_1, t_2) \tag{2.68}$$

If t signifies time and $X(t)$ models a displacement variable, the 1st and 2nd order derivative processes are designated the *velocity process* and the *acceleration process*, respectively.

The *nth order derivative in mean square*, $\{X^{(n)}, t \in [a, b]\}$, may recursively be defined as the 1st derivative process of the $(n-1)$th order derivative process. The mean value function and the auto-covariance function becomes

$$\mu_{X^{(n)}}(t) = \frac{d^n}{dt^n}\mu_X(t) \tag{2.69}$$

$$\kappa_{X^{(n)}X^{(n)}}(t_1, t_2) = \frac{\partial^{2n}}{\partial t_1^n \partial t_2^n}\kappa_{XX}(t_1, t_2) \tag{2.70}$$

The process exists if the auto-correlation function $\mu_{X^{(n)}X^{(n)}}(t_1, t_2) = \frac{\partial^{2n}}{\partial t_1^n \partial t_2^n} \mu_{XX}(t_1, t_2)$ exists and is continuous at the main diagonal.

Generally, the nth order joint moment of the n_1th order derivative process, the n_2th order derivative process, $\ldots$, and the n_nth order derivative process is related to the nth order moment of the underlying process as follows

$$E\left[X^{(n_1)}(t_1)X^{(n_2)}(t_2)\cdots X^{(n_n)}(t_n)\right] =$$

$$\frac{\partial^{n_1+n_2+\cdots+n_n}}{\partial t_1^{n_1} \partial t_2^{n_2} \cdots \partial t_n^{n_n}} E\left[X(t_1)X(t_2)\cdots X(t_n)\right] \tag{2.71}$$

Example 2.8: 1st derivative process in mean square of a harmonic process

(2.56) for the harmonic process (2.29) can be written

$$Y(t, h) = \frac{R\cos(\omega_0(t+h) + \Theta) - R\cos(\omega_0 t + \Theta)}{h} =$$

$$R\left(-\omega_0 \sin(\omega_0 t + \Theta) - \frac{1}{2}\omega_0^2 \cos(\omega_0 t + \Theta)h + \cdots\right) \quad \Rightarrow$$

$$\lim_{h\to 0} Y(t, h) = -R\omega_0 \sin(\omega_0 t + \Theta) \tag{2.72}$$

where a Taylor expansion of $R\cos(\omega_0(t+h) + \Theta)$ to the 2nd order in h has been applied. (2.72) is valid for all $\omega \in \Omega$, i.e. for all function values $R(\omega)$ and $\Theta(\omega)$ of the stochastic variables. Hence, the 1st derivative process with probability 1 is given by (2.72).

Now,

$$\lim_{h\to 0} E\left[\left(Y(t, h) - (-R\omega_0 \sin(\omega_0 t + \Theta))\right)^2\right] =$$

$$\lim_{h\to 0} E\left[R^2\left(-\frac{1}{2}\omega_0^2 \cos(\omega_0 t + \Theta)h + \cdots\right)^2\right] =$$

$$\lim_{h\to 0} E[R^2]\left(\frac{1}{4}\omega_0^4 h^2 E[\cos^2(\omega_0 t + \Theta)] + \cdots\right) = 0 \tag{2.73}$$

Hence, $\frac{d}{dt}X(t) = -R\omega_0 \sin(\omega_0 t + \Theta)$ is also the 1st derivative process in mean square.

Example 2.9: Derivative processes in mean square of Gaussian processes

Consider a Gaussian process $\{X(t), t \in [a, b]\}$ with the 1st order derivative process $\left\{\frac{d}{dt}X(t), t \in [a, b]\right\}$. The n-dimensional stochastic variable $[Y(t_1, h_1)\ldots Y(t_n, h_n)]$ is then defined by the linear transformations, cf. (2.56)

$$\left.\begin{array}{l} Y(t_1, h_1) = \dfrac{1}{h_1}(X(t_1 + h_1) - X(t_1)) \\ \vdots \\ Y(t_n, h_n) = \dfrac{1}{h_n}(X(t_n + h_n) - X(t_n)) \end{array}\right\} \tag{2.74}$$

(2.74) defines $[Y(t_1, h_1) \ldots Y(t_n, h_n)]$ as a linear mapping of the $2n$-dimensional normally distributed stochastic vector $[X(t_1), X(t_1 + h_1) \ldots X(t_n), X(t_n + h_n)]$. Hence, $[Y(t_1, h_1) \ldots Y(t_n, h_n)]$ is normally distributed for any $h_1 \ldots h_n$. This is so also for the vector of limits in mean square as $h_1 \ldots h_n \to 0$, $\left[\frac{d}{dt} X(t_1) \ldots \frac{d}{dt} X(t_n) \right]$. This means that the 1st order derivative process of the Gaussian process $\{X(t), t \in [a, b]\}$ is also Gaussian. In this case the 1st order derivative process is completely described by it's mean value function (2.61) and it's auto-covariance function (2.64).

If it exists, it is then seen by induction that also the nth order derivative process becomes Gaussian. This property may be stated that the linear operation of mean square differentiation transforms a Gaussian stochastic process into a new Gaussian process.

2.2.3 Integration in mean square

The following Riemann integral is considered

$$y(t) = \int_a^t h(t, \tau) x(\tau) d\tau \tag{2.75}$$

Relative to (2.51) the upper limit of the integral t is now assumed to be variable. The function $h(t, \tau)$ is called a *filter function* or an *impulse response function*. Forget about the meaning of these names for the time being. Instead, consider $h(t, \tau)$ merely as weights multiplied on the integrand $x(\tau)$, which are dependent on the variable upper integration limit. The only limitation on the selection of the weights will be that the Riemann integral (2.75) exists, i.e. that the following Riemann sum converges as $n \to \infty$

$$y_n(t) = \sum_{i=1}^n h(t, \tau_i^\star) x(\tau_i^\star) \Delta \tau_i \tag{2.76}$$

for any sub-division $\tau_0 < \tau_1 < \cdots < \tau_n$, $\tau_0 = a$, $\tau_n = t$ of the interval $[a, t]$, such that $\max_{t=1 \ldots n} \Delta \tau_i \to 0$, $\Delta \tau_i = \tau_i - \tau_{i-1}$. The intermediate times in the sum, $\tau_i^\star$, may be selected anywhere within the interval $[\tau_{i-1}, \tau_i[$. $h(t, \tau)$ may be as well real as complex. In the following calculations in this section only real functions are assumed.

With the same sub-division as indicated by (2.76), consider the following combined stochastic variable $Y_n(t)$ defined as a linear combination of the stochastic variables $X(\tau_1^\star) \ldots X(\tau_n^\star)$ within the stochastic process $\{X(\tau), \tau \in [a, t]\}$

$$Y_n(t) = \sum_{i=1}^n h(t, \tau_i^\star) X(\tau_i^\star) \Delta \tau_i \tag{2.77}$$

If the same limit passing as described in relation to (2.76) is applied to (2.77) a sequence $\{Y_n(t)\}$ of stochastic variables is obtained. This sequence may or may not converge to some limit $Y(t)$, either with probability 1 or in mean square. If the sequence do converge

the following *symbol* will be applied for $Y(t)$, which is named the *weighted integral of* $X(t)$

$$Y(t) = \int\limits_a^t h(t,\tau)X(\tau)d\tau \tag{2.78}$$

If the weighted integral $Y(t)$ exists for all $t \in T$, the set of stochastic variables $\{Y(t), t \in [a,t]\}$ forms a stochastic process designated the *weighted integral process* of the process $\{X(\tau), \tau \in [a,t]\}$.

Refining ourselves to the limits of the sequence (2.77) in the mean square sense, it follows from the mean square convergence theorem that the weighted integral in mean square exists if and only if $E[Y_n(t)Y_m(t)]$ exists and has the same limit no matter how the limit passings $n \to \infty$ and $m \to \infty$ are performed. Then,

$$\lim_{\substack{n\to\infty\\m\to\infty}} E[Y_n(t)Y_m(t)] =$$

$$\lim_{\substack{n\to\infty\\m\to\infty}} E\left[\sum_{i=1}^n \sum_{j=1}^m h(t,\tau_{n,i}^\star)h(t,\tau_{m,j}^\star)X(\tau_{n,i}^\star)X(\tau_{m,j}^\star)\Delta\tau_{n,i}\Delta\tau_{m,j}\right] =$$

$$\lim_{\substack{n\to\infty\\m\to\infty}} \sum_{i=1}^n \sum_{j=1}^m h(t,\tau_{n,i}^\star)h(t,\tau_{m,j}^\star)\mu_{XX}(\tau_{n,i}^\star, \tau_{m,j}^\star)\Delta\tau_{n,i}\Delta\tau_{m,j} =$$

$$\int\limits_a^t \int\limits_a^t h(t,\tau_1)h(t,\tau_2)\mu_{XX}(\tau_1,\tau_2)d\tau_1 d\tau_2 \tag{2.79}$$

Where $\tau_{n,i}^\star$ and $\tau_{m,j}^\star$ indicate intermediate points in sub-divisions with n and m sub-intervals, and $\Delta\tau_{n,i}$ and $\Delta\tau_{m,j}$ are the corresponding interval lengths. (2.79) states that the weighted integral process of the process $\{X(t), t \in [a,b]\}$ with the weights $h(t,\tau)$ exists if the planar weighted Riemann integral of the correlation function (2.79) exists for all $t \in [a,b]$.

The mean value function of the weighted integral process becomes

$$\mu_Y(t) = E[Y(t)] = E\left[\int\limits_a^t h(t,\tau)X(\tau)d\tau\right] = E\left[\lim_{n\to\infty}\text{-ms} \sum_{i=1}^n h(t,\tau_i^\star)X(\tau_i^\star)\Delta\tau_i\right] =$$

$$\lim_{n\to\infty} E\left[\sum_{i=1}^n h(t,\tau_i^\star)X(\tau_i^\star)\Delta\tau_i\right] = \lim_{n\to\infty} \sum_{i=1}^n h(t,\tau_i^\star)E[X(\tau_i^\star)]\Delta\tau_i =$$

$$\int\limits_a^t h(t,\tau)\mu_X(\tau)d\tau \tag{2.80}$$

where commutation of the operations of expectation and limit passing in mean square has been applied. (2.80) can then be stated that the operations of expectation and stochastic integration in mean square commute, i.e.

$$E\left[\int_a^t h(t,\tau)X(\tau)d\tau\right] = \int_a^t h(t,\tau)E\left[X(\tau)\right]d\tau \tag{2.81}$$

Again, it should be noticed that $\int_a^t h(t,\tau)X(\tau)d\tau$ is a symbol for the weighted integral in mean square and not a linear operator. Hence, this statement should be considered merely as a formal tool for deriving results, for which a more rigorous derivation becomes tedious.

The auto-correlation function becomes

$$\mu_{YY}(t_1,t_2) = E\left[Y(t_1)Y(t_2)\right] =$$

$$\lim_{\substack{n\to\infty \\ m\to\infty}} E\left[\sum_{i=1}^n \sum_{j=1}^m h(t_1,\tau_{n,i}^\star)h(t_2,\tau_{m,j}^\star)X(\tau_{n,i}^\star)X(\tau_{m,j}^\star)\Delta\tau_{n,i}\Delta\tau_{m,j}\right] =$$

$$\lim_{\substack{n\to\infty \\ m\to\infty}} \sum_{i=1}^n \sum_{j=1}^m h(t_1,\tau_{n,i}^\star)h(t_2,\tau_{m,j}^\star)\mu_{XX}(\tau_{n,i}^\star,\tau_{m,j}^\star)\Delta\tau_{n,i}\Delta\tau_{m,j} =$$

$$\int_a^{t_1}\int_a^{t_2} h(t_1,\tau_1)h(t_2,\tau_2)\mu_{XX}(\tau_1,\tau_2)d\tau_1 d\tau_2 \tag{2.82}$$

Combining (2.80) and (2.82) the following expression is obtained for the auto-covariance function, cf. (2.12)

$$\kappa_{YY}(t_1,t_2) = \mu_{YY}(t_1,t_2) - \mu_Y(t_1)\mu_Y(t_2) =$$

$$\int_a^{t_1}\int_a^{t_2} h(t_1,\tau_1)h(t_2,\tau_2)\mu_{XX}(\tau_1,\tau_2)d\tau_1 d\tau_2 -$$

$$\int_a^{t_1} h(t_1,\tau_1)\mu_X(\tau_1)d\tau_1 \int_a^{t_2} h(t_2,\tau_2)\mu_X(\tau_2)d\tau_2 =$$

$$\int_a^{t_1}\int_a^{t_2} h(t_1,\tau_1)h(t_2,\tau_2)\left(\mu_{XX}(\tau_1,\tau_2)-\mu_X(\tau_1)\mu_X(\tau_2)\right)d\tau_1 d\tau_2 =$$

$$\int\limits_{a}^{t_1} \int\limits_{a}^{t_2} h(t_1, \tau_1) h(t_2, \tau_2) \kappa_{XX}(\tau_1, \tau_2) d\tau_1 d\tau_2 \tag{2.83}$$

Generally, the nth order joint moment of the weighted integral process can be proved to be

$$E\left[Y(t_1) \cdots Y(t_n)\right] =$$

$$\int\limits_{a}^{t_1} \cdots \int\limits_{a}^{t_n} h(t_1, \tau_1) \cdots h(t_n, \tau_n) E\left[X(\tau_1) \cdots X(\tau_n)\right] d\tau_1 \cdots d\tau_n \tag{2.84}$$

(2.84), which includes (2.80) and (2.82) as special cases, may suggestively be stated as the operations of expectation and stochastic integration in mean square commute.

Example 2.10: Integral process of harmonic process

For the harmonic process (2.29), the integrated process can be shown to be

$$Y(t) = \int\limits_{a}^{t} X(t) dt = \frac{R}{\omega_0} \left(\sin(\omega_0 t + \Theta) - \sin(\omega_0 a + \Theta) \right) \tag{2.85}$$

(2.85) is valid for any function values $R(\omega)$ and $\Theta(\omega)$, $\omega \in \Omega$, and then signifies the integral process with probability 1. Moreover, (2.85) can be proved to be the integral process in mean square. The proof is omitted.

Example 2.11: Weighted integral process in mean square of Gaussian process

Consider a Gaussian process $\{X(t), t \in [a, \infty[\}$. The n-dimensional stochastic variable $[Y_{n_1}(t_1) \ldots Y_{n_n}(t_n)]$ is then defined by the linear transformations, cf. (2.77)

$$\left. \begin{array}{l} Y_{n_1}(t_1) = h(t_1, \tau^{\star}_{n_1,1}) X(\tau^{\star}_{n_1,1}) \Delta \tau_{n_1,1} + \cdots + h(t_1, \tau^{\star}_{n_1,n_1}) X(\tau^{\star}_{n_1,n_1}) \Delta \tau_{n_1,n_1} \\[4pt] \quad\vdots \\[4pt] Y_{n_n}(t_n) = h(t_n, \tau^{\star}_{n_n,1}) X(\tau^{\star}_{n_n,1}) \Delta \tau_{n_n,1} + \cdots + h(t_n, \tau^{\star}_{n_n,n_n}) X(\tau^{\star}_{n_n,n_n}) \Delta \tau_{n_n,n_n} \end{array} \right\} \tag{2.86}$$

$Y_{n_i}, i = 1 \ldots n$ in (2.86) signifies the Riemann sum (2.77) using n_i intervals for the division of the interval $[a, t_i]$.

(2.86) defines $[Y_{n_1}(t_1) \ldots Y_{n_n}(t_n)]$ as a linear mapping of the $(n_1 + n_2 + \cdots + n_n)$-dimensional normally distributed stochastic vector $[X(\tau^{\star}_{n_1,1}) \ldots X(\tau^{\star}_{n_1,n_1}), X(\tau^{\star}_{n_2,1}) \ldots X(\tau^{\star}_{n_2,n_2}) \ldots X(\tau^{\star}_{n_n,1}) \ldots X(\tau^{\star}_{n_n,n_n})]$. Hence, $[Y_{n_1}(t_1) \ldots Y_{n_n}(t_n)]$ is normally distributed for any $n_1 \ldots n_n$. This is so also in the limit as $n_1 \ldots n_n \to \infty$ where the n-dimensional vector of weighted integrals in mean square, $[Y(t_1) \ldots Y(t_n)]$,

is obtained. This means that the weighted integral process of the Gaussian process $\{X(t), t \in [a, \infty[\}$ is also Gaussian. In this case the weighted integral process is completely described by it's mean value function (2.80) and it's auto-covariance function (2.83). Hence, the linear operation of mean square integration, transforms a Gaussian stochastic process into a new Gaussian process.

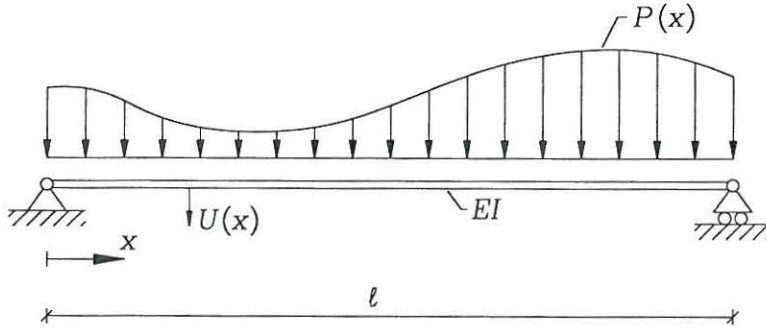## Example 2.12: Simply supported Bernoulli-Euler beam

Figure 2.7: Simply supported Bernoulli-Euler beam with random loading.

Figure 2.7 shows a plane, horizontal simply supported Bernoulli-Euler beam of the length l. Sections are identified by the coordinate x, which is measured from the left end. The bending stiffness EI is assumed to be deterministic and constant. The beam is loaded in vertical direction with a loading $P(x)$, which is considered uncertain. Hence, this loading is modelled by a stochastic process $\{P(x), x \in [0, l]\}$ with a certain mean value function $\mu_P(x)$ and auto-covariance function $\kappa_{PP}(x_1, x_2)$ signifying the deterministic variation and correlation structure of the excitation. Because the loading is uncertain the displacement $U(x)$ in the direction of the loading becomes uncertain as well, and will be modelled by a stochastic process $\{U(x), x \in [0, l]\}$. The loading and the displacement processes are related by the following stochastic differential equation and stochastic weighted integral which may be interpreted as well with probability 1 as in mean square

$$\left. \begin{aligned} EI\frac{d^4}{dx^4}U(x) &= P(x), \quad x \in]0, l[\\ U(0) &= \frac{d^2}{dx^2}U(0) = U(l) = \frac{d^2}{dx^2}U(l) = 0 \end{aligned} \right\} \tag{2.87}$$

$$U(x) = \int_0^l h(x, \tau)P(\tau)d\tau \tag{2.88}$$

$$h(x, \tau) = \begin{cases} \frac{x(l-\tau)(2\tau l - x^2 - \tau^2)}{6lEI} & , x \leq \tau \\ \frac{\tau(l-x)(2xl - x^2 - \tau^2)}{6lEI} & , x > \tau \end{cases} \tag{2.89}$$

Here, $h(t, \tau)$ signifies the influence function for the displacement, i.e. the displacement at the position x from a unit load at the position τ.

From (2.69) and (2.80) follows that the mean value function $\mu_U(x)$ of the displacement either can be obtained from the following differential equation or from the following integral

$$\left. \begin{aligned} EI\frac{d^4}{dx^4}\mu_U(x) &= \mu_P(x), \quad x \in [0, l] \\ \mu_U(0) &= \frac{d^2}{dx^2}\mu_U(0) = \mu_U(l) = \frac{d^2}{dx^2}\mu_U(l) = 0 \end{aligned} \right\} \tag{2.90}$$

$$\mu_U(x) = \int_0^l h(x, \tau)\mu_P(\tau)d\tau \tag{2.91}$$

$\mu_U(x)$ is simply the displacement of the beam if it is loaded with the mean value loading $\mu_P(x)$. The auto-covariance function of the displacement follows from (2.83)

$$\kappa_{UU}(x_1, x_2) = \int_0^l \int_0^l h(x_1, \tau_1)h(x_2, \tau_2)\kappa_{PP}(\tau_1, \tau_2)d\tau_1 d\tau_2 \tag{2.92}$$

The nth order joint moments $E[U(x_1)\cdots U(x_n)]$ can be calculated from (2.84) in case the corresponding moments, $E[P(\tau_1)\cdots P(\tau_n)]$, are available for the loading

$$E[U(x_1)\cdots U(x_n)] =$$

$$\int_0^l \cdots \int_0^l h(x_1, \tau_1)\cdots h(x_n, \tau_n)E[P(\tau_1)\cdots P(\tau_n)]d\tau_1\cdots d\tau_n \tag{2.93}$$

In civil engineering a great deal of phenomena can with acceptable accuracy be modelled by linear models, such as the Bernoulli-Euler beam in example 2.12. Since Gaussianity is preserved under such linear operations as stochastic differentiation and stochastic integration the response processes (stresses, displacements, etc.) become Gaussian too. However, even if the excitation is non-Gaussian the response processes may under certain conditions be assumed to be Gaussian. Actually, if the random variables $X(\tau_i^\star)$ in the sum (2.77) are not too strongly correlated, and all the weights $h(t, \tau_i^\star)$ are of equal magnitude, $Y_n(t)$ may be assumed to be Gaussian as $n \to \infty$ because of the central limit theorem. This is the main reason for the extensive applicability of Gaussian processes as models for response processes in structural mechanics.

2.3 Homogeneous processes

A stochastic process $\{X(t), t \in T\}$ is said to be *homogeneous in the strict sense* (or *strictly homogeneous*) if the set of finite dimensional joint distributions of the stochastic variables of the process is invariant under a linear translation $t \to t + a$ for every $a \in T$. This means

$$\left.\begin{aligned}
&F_{\{X\}}(x_1, t_1) = F_{\{X\}}(x_1, t_1 + a) \\
&F_{\{X\}}(x_1, t_1; x_2, t_2) = F_{\{X\}}(x_1, t_1 + a; x_2, t_2 + a) \\
&\quad\vdots \\
&F_{\{X\}}(x_1, t_1; \ldots; x_n, t_n) = F_{\{X\}}(x_1, t_1 + a; \ldots; x_n, t_n + a) \\
&\quad\vdots
\end{aligned}\right\} \tag{2.94}$$

(2.94) states that the joint distribution functions of the n-dimensional stochastic variables $[X(t_1)\ldots X(t_n)]$ and $[X(t_1+a)\ldots X(t_n+a)]$ are identical for any order n, any index values $t_1\ldots t_n \in T$ and any $a \in T$. (2.94) then only makes sense if T is unbounded, i.e. $T = Z^n$ for index discrete processes and $T = R^n$ for index continuous processes. Actually, if T is bounded, and t is placed on the boundary, then an $a \in T$ exists such that $t + a \notin T$ in which case (2.94) becomes meaningless.

If (2.94) only holds for $n = 1$ and $n = 2$, the process is termed *homogeneous in the weak sense* (or *weakly homogeneous*).

If especially $a = -t_1$, it is seen from (2.94) that when $\{X(t), t \in T\}$ is weakly homogeneous, the 1st order distribution functions become independent of t_1, and the 2nd order distribution functions will only depend on the index parameters t_1 and t_2 through the difference $t_2 - t_1$, i.e. $F_{\{X\}}(x, t_1) = F_{\{X\}}(x)$ and $F_{\{X\}}(x_1, t_1; x_2, t_2) = F_{\{X\}}(x_1, x_2; t_2 - t_1)$. For a strictly homogeneous process the higher order distribution functions depend on $t_1 \ldots t_n$ through the differences $t_2 - t_1, t_3 - t_1 \ldots t_n - t_1$, i.e. $F_{\{X\}}(x_1, t_1; x_2, t_2; \ldots; x_n, t_n) = F_{\{X\}}(x_1, x_2 \ldots x_n; t_2 - t_1 \ldots t_n - t_1)$. If the index parameter t denotes time, the terms *strictly* and *weakly stationary processes* are used synonymously with the designations strongly and weakly homogeneous processes.

Since $F_{\{X\}}(x)$ is independent of t for a weakly homogeneous process, the mean value function and the variance function also become independent of t, i.e. $\mu_X(t) = \mu_X$ and $\sigma_X^2(t) = \sigma_X^2$. As $F_{\{X\}}(x_1, t_1; x_2, t_2)$ is merely a function of $(t_2 - t_1)$, the auto-covariance function fulfils

$$\kappa_{XX}(t_1, t_2) = \kappa_{XX}(t_2 - t_1) \tag{2.95}$$

The symmetry property (2.13) of the auto-covariance function implies that

$$\kappa_{XX}(\tau) = \kappa_{XX}(-\tau), \quad \tau = t_2 - t_1 \tag{2.96}$$

Hence, the auto-covariance function of a weakly homogeneous process is an even function of the index difference.

The previous definitions can immediately be extended to N-dimensional stochastic vector processes $\{\mathbf{X}(t), t \in T\}$, $\mathbf{X}^T(t) = [X_1(t)\ldots X_N(t)]$. The process is strictly homogeneous if for any order n, any index parameter $(t_1, i_1)\ldots(t_n, i_n) \in T \times \{1 \ldots N\}$ and any $a \in T$

$$\left.\begin{aligned}
&F_{\{\mathbf{X}\}}(x_1, t_1, i_1) = F_{\{\mathbf{X}\}}(x_1, t_1 + a, i_1) \\
&F_{\{\mathbf{X}\}}(x_1, t_1, i_1; x_2, t_2, i_2) = F_{\{\mathbf{X}\}}(x_1, t_1 + a, i_1; x_2, t_2 + a, i_2) \\
&\quad\vdots \\
&F_{\{\mathbf{X}\}}(x_1, t_1, i_1; \ldots; x_n, t_n, i_n) = F_{\{\mathbf{X}\}}(x_1, t_1 + a, i_1; \ldots; x_n, t_n + a, i_n) \\
&\quad\vdots
\end{aligned}\right\} \tag{2.97}$$

Vector processes are weakly homogeneous, if (2.97) only holds for $n = 1$ and $n = 2$. If $\{\mathbf{X}(t), t \in T\}$ is homogeneous, it implies that all coordinate processes are homogeneous.

The opposite statement is not necessarily true. For a weakly homogeneous vector process the mean value functions become constant as a function of t, i.e. $\mu_{X_i}(t) = \mu_{X_i}$. The cross-covariance functions become a function of the index difference $t_2 - t_1$, i.e. $\kappa_{X_{i_1} X_{i_2}}(t_1, t_2) = \kappa_{X_{i_1} X_{i_2}}(t_2 - t_1)$. The symmetry property (2.13) attains the form

$$\kappa_{X_{i_1} X_{i_2}}(\tau) = \kappa_{X_{i_2} X_{i_1}}(-\tau), \quad \tau = t_2 - t_1 \tag{2.98}$$

In nature all processes will be of limited extent in space and time. Hence, no natural processes can be strictly or weakly homogeneous in the sense of the definitions (2.94) and (2.97). Application of homogeneous processes is then a matter of approximate mathematical modelling. For instance the wind velocities in a storm grow up from calm condition until a period of length say 1 hour is reached, where extreme wind velocities are reached, followed by a decaying period, where the wind velocities eventually returns to calm. It is common practice to model the central period as a weakly homogeneous process in wind engineering. This is justified if a characteristic time length scale of the structure to which the wind load model is applied is small compared to the mentioned length of 1 hour of the central part of extreme exposure.

In the remaining part of this section it is assumed that $T = R$. A necessary condition for considering a physical process as homogeneous is that the distances in the index set $\tau_0 = |t_1 - t_2|$, within which significant correlations among the stochastic variables are present, are small compared to the minimum diameter in T. For instance for the beam-depth process $\{H(x), x \in [0, L]\}$ of example 2.1 this means that the distances x_0, between which $H(x)$ and $H(x + x_0)$ are significantly correlated, fulfils $x_0 \ll L$. τ_0 will be referred to as the *correlation length* of the process. A formal quantitative definition for homogeneous processes in terms of the auto-correlation coefficient function is given by

$$\tau_0 = \int_0^\infty |\rho_{XX}(\tau)| d\tau \tag{2.99}$$

A typical variation of $\rho_{XX}(\tau)$ and the geometrical interpretation of τ_0, are shown in figure 2.8.
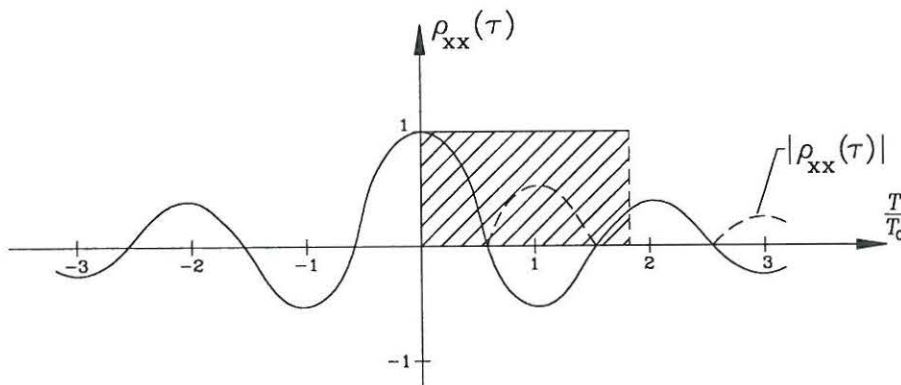

Figure 2.8: Auto-correlation coefficient function and geometrical interpretation of correlation length τ_0.

The characteristic time length of the structure exposed to the mentioned homogeneous wind loading model should then be selected as the correlation length of the displacement response process of the structure. In case the structure is sensible to dynamic loading it can be shown that $\tau_0 \simeq \frac{1}{\zeta}T_0$, where ζ is the damping ratio and T_0 is the eigenperiod of the structure (more precise the damping ratio and eigenperiod of the fundamental mode). Using $\zeta \simeq 0.01$ and $T_0 \simeq 1$ s follows $\tau_0 \simeq 100$ s. This is well below the length of the period of extreme exposure, so the application of homogeneous modelling is justified. On the other hand, in case of earthquake excitation the central part of exposure, due to the arrival of horizontal shear waves, is much shorter, say 10 s. Hence, excitations due to earthquakes must be modelled by non-homogeneous stochastic processes.

Assuming $|\kappa_{X_{i_1}X_{i_2}}(\tau)|$ to be integrable, the following *Fourier transform* pair may be defined

$$\kappa_{X_{i_1}X_{i_2}}(\tau) = \int\limits_{-\infty}^{\infty} e^{i\omega\tau} S_{X_{i_1}X_{i_2}}(\omega)d\omega \tag{2.100}$$

$$S_{X_{i_1}X_{i_2}}(\omega) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{-i\omega\tau} \kappa_{X_{i_1}X_{i_2}}(\tau)d\tau \tag{2.101}$$

(2.100) and (2.101) are termed the *Wiener-Khintchine relations*. $S_{X_{i_1}X_{i_2}}(\omega)$ is designated as the *cross-spectral density functions* (especially the *auto-spectral density function*, if $i_1 = i_2$). $\kappa_{X_{i_1}X_{i_2}}(\tau)$ is always real, whereas $S_{X_{i_1}X_{i_2}}(\omega)$ is generally complex. Since the Fourier transformation $S_{X_{i_1}X_{i_2}}(\omega)$ uniquely defines the cross-covariance function all information about the correlation structure is also contained in $S_{X_{i_1}X_{i_2}}(\omega)$. Before a physical interpretation of $S_{X_{i_1}X_{i_2}}(\omega)$ along with a formal derivation of (2.100) and (2.101) are given some useful mathematical properties of the cross-spectral density function will be proved. Substituting $u = -\tau$ into (2.101) and applying (2.96) provide

$$S_{X_{i_1}X_{i_2}}(\omega) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{i\omega u} \kappa_{X_{i_1}X_{i_2}}(-u)du = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{i\omega u} \kappa_{X_{i_2}X_{i_1}}(u)du =$$

$$\left(\frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{-i\omega u} \kappa_{X_{i_2}X_{i_1}}(u)du\right)^{\star} = S_{X_{i_2}X_{i_1}}^{\star}(\omega) \tag{2.102}$$

$S_{X_{i_2}X_{i_1}}^{\star}(\omega)$ denotes the *complex conjugate* of $S_{X_{i_2}X_{i_1}}(\omega)$. At the derivation of (2.102) it has been used that $\left(e^{i\omega u}\right)^{\star} = \left(\cos(\omega u) + i\sin(\omega u)\right)^{\star} = \cos(\omega u) - i\sin(\omega u) = e^{-i\omega u}$, so $\left(\left(e^{i\omega u}\right)^{\star}\right)^{\star} = (e^{-i\omega u})^{\star}$. Use of this relation in (2.101) provides

$$S_{X_{i_1}X_{i_2}}(\omega) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} \left(e^{i\omega\tau}\right)^{\star} \kappa_{X_{i_1}X_{i_2}}(\tau)d\tau =$$

$$\left(\frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{-i(-\omega)\tau} \kappa_{X_{i_1} X_{i_2}}(\tau) d\tau \right)^{\star} = S^{\star}_{X_{i_1} X_{i_2}}(-\omega) \qquad (2.103)$$

The corresponding results for the auto-spectral density function is obtained by setting $i_1 = i_2 = i$ in (2.102) and (2.103), i.e.

$$S_{X_i X_i}(\omega) = S^{\star}_{X_i X_i}(\omega) \qquad (2.104)$$

$$S_{X_i X_i}(\omega) = S^{\star}_{X_i X_i}(-\omega) = S_{X_i X_i}(-\omega) \qquad (2.105)$$

(2.104) and (2.105) establish the auto-spectral density as a real and symmetric function.

Since, the auto-covariance function $\kappa_{X_i X_i}(\tau)$ and the auto-spectral density function $S_{X_i X_i}(\omega)$ are symmetric functions, the Wiener-Khintchine relations for these quantities can be written as the following *cosine transform* pair

$$\kappa_{X_i X_i}(\tau) = \int\limits_{-\infty}^{\infty} \cos(\omega\tau) S_{X_i X_i}(\omega) d\omega + i \int\limits_{-\infty}^{\infty} \sin(\omega\tau) S_{X_i X_i}(\omega) d\omega =$$

$$2 \int\limits_{0}^{\infty} \cos(\omega\tau) S_{X_i X_i}(\omega) d\omega \qquad (2.106)$$

$$S_{X_i X_i}(\omega) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \cos(\omega\tau) \kappa_{X_i X_i}(\tau) d\tau - i \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \sin(\omega\tau) \kappa_{X_i X_i}(\tau) d\tau =$$

$$\frac{1}{\pi} \int\limits_{0}^{\infty} \cos(\omega\tau) \kappa_{X_i X_i}(\tau) d\tau \qquad (2.107)$$

where it has been used that $\cos(\omega\tau)$ is an even function and $\sin(\omega\tau)$ is an odd function of ω and τ. The variance function $\sigma^2_{X_i} = \kappa_{X_i X_i}(0)$ is obtained by setting $\tau = 0$ in (2.106)

$$\sigma^2_{X_i} = 2 \int\limits_{0}^{\infty} S_{X_i X_i}(\omega) d\omega \qquad (2.108)$$

Positive values of ω may be interpreted as a circular frequency, if τ represents time, whereas ω is a *wave number*, if τ represents a space variable. Although ω will be called "frequency" in the following, this distinction should be kept in mind. Then,

(2.108) shows that the auto-spectral density function can be interpreted as the density distribution of the variance $\sigma^2_{X_i}$ in the frequency space. So, $2S_{X_iX_i}(\omega)d\omega$ represents the variance contribution from harmonic components in the interval $]\omega, \omega + d\omega]$.

In order to give a further physical interpretation of the spectral density concept the theory of Fourier analysis of periodic functions will initially be reviewed. In this respect, consider a real function $x(t)$ defined in the interval $[-\frac{T}{2}, \frac{T}{2}]$. If $x(t)$ is piece-wise differentiable and the absolute value, $\mid x(t) \mid$, is integrable over the interval $[-\frac{T}{2}, \frac{T}{2}]$, $x(t)$ can be represented by the following so-called *Fourier series*

$$x(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos(\omega_n t) + b_n \sin(\omega_n t)) \tag{2.109}$$

where

$$\omega_n = n\frac{2\pi}{T}, \quad n = 1, 2, \ldots \tag{2.110}$$

$$\left.\begin{array}{l} a_n = \dfrac{2}{T}\displaystyle\int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)\cos(\omega_n t)dt, \quad n = 0, 1, 2, \ldots \\[4em] b_n = \dfrac{2}{T}\displaystyle\int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)\sin(\omega_n t)dt, \quad n = 1, 2, \ldots \end{array}\right\} \tag{2.111}$$

The right-hand side of (2.109) converges to $x(t)$ at all continuity points in $[-\frac{T}{2}, \frac{T}{2}]$. At any discontinuity point the series converges to the mean value, $\frac{1}{2}\left(x(t^-) + x(t^+)\right)$, of the limits from the left and from the right. The range of definition may be extended outside $[-\frac{T}{2}, \frac{T}{2}]$, in which case $x(t)$ determined by (2.109) becomes a periodic function with the period T. Next, use of Euler's formulas

$$\left.\begin{array}{l} \cos(\omega_n t) = \dfrac{1}{2}\left(e^{i\omega_n t} + e^{-i\omega_n t}\right) \\[2em] \sin(\omega_n t) = \dfrac{1}{2i}\left(e^{i\omega_n t} - e^{-i\omega_n t}\right) \end{array}\right\} \tag{2.112}$$

in (2.109) provides

$$x(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}\frac{1}{2}(a_n - ib_n)e^{i\omega_n t} + \sum_{n=1}^{\infty}\frac{1}{2}(a_n + ib_n)e^{-i\omega_n t} =$$

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}\frac{1}{2}(a_n - ib_n)e^{i\omega_n t} + \sum_{n=-1}^{-\infty}\frac{1}{2}(a_{-n} + ib_{-n})e^{-i\omega_{-n} t} =$$

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} \frac{1}{2}(a_n - ib_n)e^{i\omega_n t} + \sum_{n=-1}^{-\infty} \frac{1}{2}(a_n - ib_n)e^{i\omega_n t} \qquad (2.113)$$

At the derivation of (2.113) a change of summation variable from n to $-n$ is initially performed in the second sum. Next, it has been used that $\omega_{-n} = -\omega_n$, $a_n = a_{-n}$ and $b_n = -b_{-n}$, as seen from (2.110) and (2.111). Setting $b_0 = 0$, the Fourier series can then be represented in the following compact form

$$x(t) = \sum_{n=-\infty}^{\infty} e^{i\omega_n t} X(\omega_n, T)\Delta\omega \qquad (2.114)$$

where

$$\Delta\omega = \frac{2\pi}{T} \qquad (2.115)$$

$$X(\omega_n, T) = \frac{1}{2\pi}\frac{T}{2}(a_n - ib_n) = \frac{1}{2\pi}\int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)\left(\cos(\omega_n t) - i\sin(\omega_n t)\right) dt =$$

$$\frac{1}{2\pi}\int_{-\frac{T}{2}}^{\frac{T}{2}} e^{-i\omega_n t} x(t) dt \qquad (2.116)$$

$\Delta\omega$ as given by (2.115) is termed the *fundamental circular frequency*. Since, $\left(e^{i\omega_n t}\right)^{\star} = e^{-i\omega_n t}$, cf. the remarks subsequent to eq. (2.103), and because $x(t)$ is real, so $x(t)^{\star} = x(t)$, complex conjugation of the left-hand and the right-hand sides of (2.114) provides

$$x(t) = \sum_{n=-\infty}^{\infty} \left(e^{i\omega_n t} X(\omega_n, T)\right)^{\star}\Delta\omega = \sum_{n=-\infty}^{\infty} \left(e^{i\omega_n t}\right)^{\star} X^{\star}(\omega_n, T)\Delta\omega =$$

$$\sum_{n=-\infty}^{\infty} e^{-i\omega_n t} X^{\star}(\omega_n, T)\Delta\omega \qquad (2.117)$$

$x^2(t)$ can then be represented by the series

$$x^2(t) = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} X(\omega_n, T)X^{\star}(\omega_m, T)e^{i(\omega_n - \omega_m)t}\Delta\omega^2 =$$

$$\sum_{n=-\infty}^{\infty}\sum_{\substack{m=-\infty \\ m\neq n}}^{\infty} X(\omega_n, T)X^{\star}(\omega_m, T)e^{i2\pi(n-m)\frac{t}{T}}\Delta\omega^2 +$$

$$\sum_{n=-\infty}^{\infty} X(\omega_n, T) X^{\star}(\omega_n, T) \Delta\omega^2 \qquad (2.118)$$

The time average value of $x^2(t)$ over the interval $[-\frac{T}{2}, \frac{T}{2}]$ can then be calculated as follows

$$\frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt = \sum_{n=-\infty}^{\infty} X(\omega_n, T) X^{\star}(\omega_n, T) \Delta\omega^2 = \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} |X(\omega_n, T)|^2 \Delta\omega \qquad (2.119)$$

The contributions from the off-diagonal terms in the double sum (2.118) cancel, because

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} e^{i2\pi(n-m)\frac{t}{T}} dt =$$

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} \cos\left(2\pi(n-m)\frac{t}{T}\right) dt + i \int_{-\frac{T}{2}}^{\frac{T}{2}} \sin\left(2\pi(n-m)\frac{t}{T}\right) dt =$$

$$0 + i \cdot 0 = 0, \quad n \neq m \qquad (2.120)$$

Now, $X(\omega_{-n}, T) = X(-\omega_n, T) = X^{\star}(\omega_n, T)$ as follows from (2.110) and (2.116). Then, $|X(\omega_{-n}, T)|^2 = |X(\omega_n, T)|^2$. (2.119) can then be written

$$\frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt = \frac{2\pi}{T} |X(0, T)|^2 \Delta\omega + 2\frac{2\pi}{T} \sum_{n=1}^{\infty} |X(\omega_n, T)|^2 \Delta\omega = \mu_{x,T}^2 + \sigma_{x,T}^2 \quad \Rightarrow$$

$$\mu_{x,T} = \frac{2\pi}{T} X(0, T) = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) dt \qquad (2.121)$$

$$\sigma_{x,T}^2 = 2\frac{2\pi}{T} \sum_{n=1}^{\infty} |X(\omega_n, T)|^2 \Delta\omega = 2 \sum_{n=1}^{\infty} S_{xx,T}(\omega_n) \Delta\omega \qquad (2.122)$$

$$S_{xx,T}(\omega_n) = \frac{2\pi}{T} |X(\omega_n, T)|^2 \qquad (2.123)$$

$\mu_{x,T}$ signifies the time average value of $x(t)$ over the interval. Notice $X(0,T)$ is real, so $X(0,T)^\star = X(0,T)$. $\sigma_{x,T}^2$ signifies the time average value of the squared deviation from $\mu_{x,T}$, $\left(x(t) - \mu_{x,T}\right)^2$, as seen from the following derivation

$$
\frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \left(x(t) - \mu_{x,T}\right)^2 dt =
$$

$$
\frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t)dt - 2\mu_{x,T}\frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)dt + \mu_{x,T}^2\frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt =
$$

$$
\frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t)dt - \mu_{x,T}^2 = \sigma_{x,T}^2 \tag{2.124}
$$

(2.122), which is termed *Parceval's theorem*, states that contribution to $\sigma_{x,T}^2$ is made up by independent contributions $2S_{xx,T}(\omega_n)\Delta\omega$ from the circular frequencies $\omega_n = n\frac{2\pi}{T}$, $n = 1, 2, \dots$. Marc-Antoine Parceval was a French mathematician, who lived during the French Revolution. Surely, (2.122) was not invented by the father of Lohengrin. Now, (2.122) should be compared with (2.108). $\mu_{x,T}$ as given by (2.121) is an estimate of the mean value function μ_X, $\sigma_{x,T}^2$ as given by (2.124) is an estimate of the variance σ_X^2 of the homogeneous process, and $S_{xx,T}(\omega_n)$ as given by (2.123) is an estimate of the auto-spectral density function $S_{XX}(\omega_n)$ at the circular frequency ω_n. Actually, for a certain class of so-called *ergodic processes* to be considered in chapter 4, for any realization $x(t)$ of the process $\{X(t), t \in R\}$ one has

$$
\mu_X = \lim_{T\to\infty} \mu_{x,T} = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)dt \tag{2.125}
$$

$$
\sigma_X^2 = \lim_{T\to\infty} \sigma_{x,T}^2 = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \left(x(t) - \mu_X\right)^2 dt \tag{2.126}
$$

$$
S_{XX}(\omega_n) = \lim_{T\to\infty} S_{xx,T}(\omega_n) = \lim_{T\to\infty} \frac{2\pi}{T} \mid X(\omega_n,T) \mid^2 \tag{2.127}
$$

Further, in the limit $T \to \infty$ the Riemann sum (2.122) converges to the integral (2.108).

In the same way it can be shown that for any realizations $x(t)$ and $y(t)$ of the ergodic stochastic processes $\{X(t), t \in R\}$ and $\{Y(t), t \in R\}$ that the cross-spectral density function is given by

$$
S_{XY}(\omega_n) = \lim_{T\to\infty} \frac{2\pi}{T} X^\star(\omega_n,T)Y(\omega_n,T) \tag{2.128}
$$

where

$$Y(\omega_n, T) = \frac{1}{2\pi} \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{-i\omega_n t} y(t) dt \qquad (2.129)$$

Obviously, $S_{xx,T}(\omega_n) = \frac{2\pi}{T} \mid X(\omega_n, T) \mid^2 \geq 0$, which also holds in the limit as $T \to \infty$. To the extent, that the auto-spectral density function can be determined by the ergodic relation (2.127), it then follows that $S_{X_i X_i}(\omega)$ is a real symmetric non-negative function of the circular frequency ω.

The Fourier transformation as applied in (2.100) and (2.101) can formally be derived from the finite transforms (2.114) and (2.116) as $T \to \infty$

$$\left.\begin{aligned} x(t) &= \int_{-\infty}^{\infty} e^{i\omega t} X(\omega) d\omega \\[2em] X(\omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} x(t) dt \end{aligned}\right\} \qquad (2.130)$$

The Riemann sum (2.114) converges to the 1st of the integral transformations in (2.130).

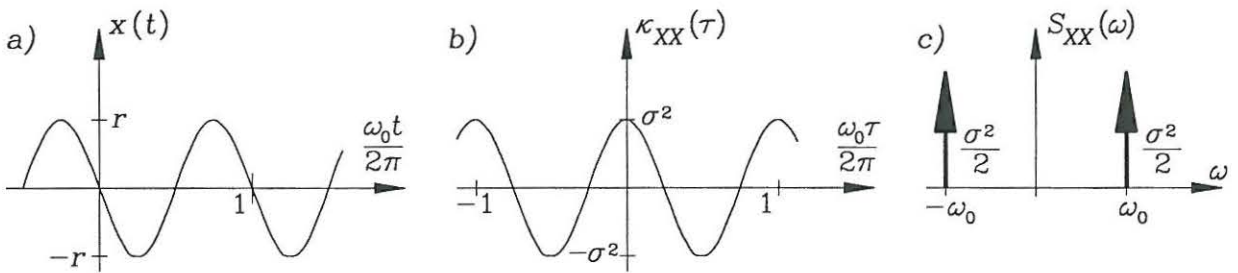## Example 2.13: Auto-spectral density function of a harmonic process

Figure 2.9: Harmonic process. a) Realization. b) Auto-covariance function. c) Auto-spectral density function.

The harmonic process considered in example 2.3 can be shown to be homogeneous in the strict sense. The auto-spectral density function of the process follows from (2.31) and (2.101)

$$S_{XX}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega\tau} \sigma^2 \cos(\omega_0\tau) d\tau = \frac{\sigma^2}{2\pi} \int_{-\infty}^{\infty} e^{-iw\tau} \frac{1}{2} \left(e^{i\omega_0\tau} + e^{-i\omega_0\tau} \right) d\tau =$$

$$\frac{\sigma^2}{4\pi} \int_{-\infty}^{\infty} \left(e^{-i(\omega-\omega_0)\tau} + e^{-i(\omega+\omega_0)\tau} \right) d\tau = \frac{\sigma^2}{2} \left(\delta(\omega - \omega_0) + \delta(\omega + \omega_0) \right) \qquad (2.131)$$

In the latter statement the following formal mutual Fourier transform of $x(\tau) = 1$ and the Dirac's delta function $X(\omega) = \delta(\omega)$ is applied

$$
\left.
\begin{aligned}
1 &= \int_{-\infty}^{\infty} e^{i\omega\tau}\delta(\omega)d\omega \\[2em]
\delta(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\tau} \cdot 1 d\tau
\end{aligned}
\right\}
\tag{2.132}
$$

This validity of the first relation of (2.132) follows from (1.19). The second relation, which is applied in (2.131) then follows from the second relation of (2.130).

A realization of the harmonic process with the amplitude $R = r$, along with the auto-covariance function, as given by (2.31), and the auto-spectral density function, as given by (2.131), are shown in figure 2.9.
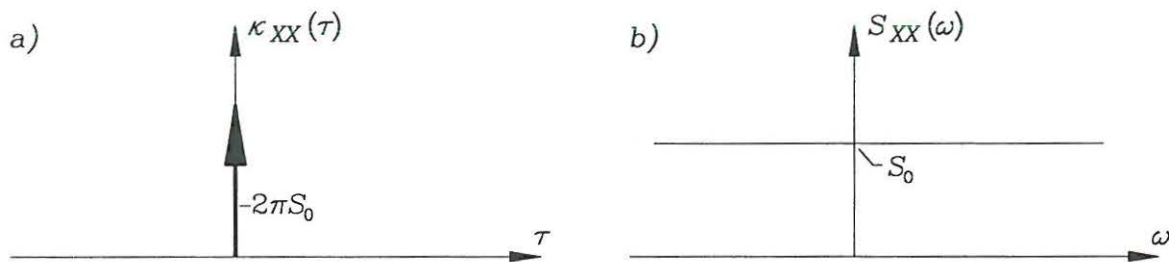
Figure 2.10: White noise. a) Auto-covariance function. b) Auto-spectral density function

A homogeneous Gaussian process $\{X(t), t \in R\}$ with mean value function $\mu_X = 0$ is termed *white noise*, if the auto-spectral density function is constant for all frequencies, i.e.

$$
S_{XX}(\omega) = S_0, \quad \omega \in R
\tag{2.133}
$$

From (2.100) and (2.132) the auto-covariance function of a white noise can be evaluated as

$$
\kappa_{XX}(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau}S_0 d\omega = 2\pi S_0 \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\tau u}du = 2\pi S_0 \delta(\tau)
\tag{2.134}
$$

where the change of integrabel variable $u = -\omega$ has been made. If especially $2\pi S_0 = 1$, the process is named *unit white noise*. The auto-covariance function and the auto-spectral density function of a white noise process, are shown in 2.10.

According to (2.134), $X(t)$ and $X(t + \tau)$ for a white noise are stochastic independent no matter how small values of $\mid \tau \mid$ are considered. The samples $x(t)$ and $x(t + \tau)$ will then not be close to each other even for small values of $\mid \tau \mid$. Hence, the realizations of white noise are discontinuous for any index value, t. Inserting (2.133) into (2.108) provides $\sigma_X^2 = \infty$, implying $X(t) \sim N(0, \infty)$. Let $a < b$. Then

$$P(a < X(t) \leq b) = \lim_{\sigma_X \to \infty} \left(\Phi \left(\frac{b}{\sigma_X} \right) - \Phi \left(\frac{a}{\sigma_X} \right) \right) = 0.5 - 0.5 = 0 \qquad (2.135)$$

$\Phi(x)$ is the distribution function of a stochastic variable $X \sim N(0,1)$. Together (2.134) and (2.135) imply that the realizations of a white noise process are discontinuous for all $t \in T$, and the samples will be beyond of any finite interval with probability 1. Surely, such a process cannot be given any rigorous mathematical definition. It is not even possible to sketch typical realizations of the process. It is common practice in engineering applications to define a white process as 1st derivative process of the Wiener process. The reason for this is of course the identity between (2.60) and (2.134). Again the lack of rigor becomes obvious, since the Wiener process has non-differentiable realizations as mentioned in example 2.5. The designation 'white noise' originates from optics, where the auto-spectral density function of sunlight is almost constant over the visible frequency range of the electromagnetic spectrum.

The harmonic process dealt with in example 2.3 and the white noise process can be considered as limiting extreme stochastic processes, where the variance is concentrated on the angular frequencies ω_0 and uniformly distributed over all frequencies, respectively. Both cases are idealizations, which are only approximately met in practice.
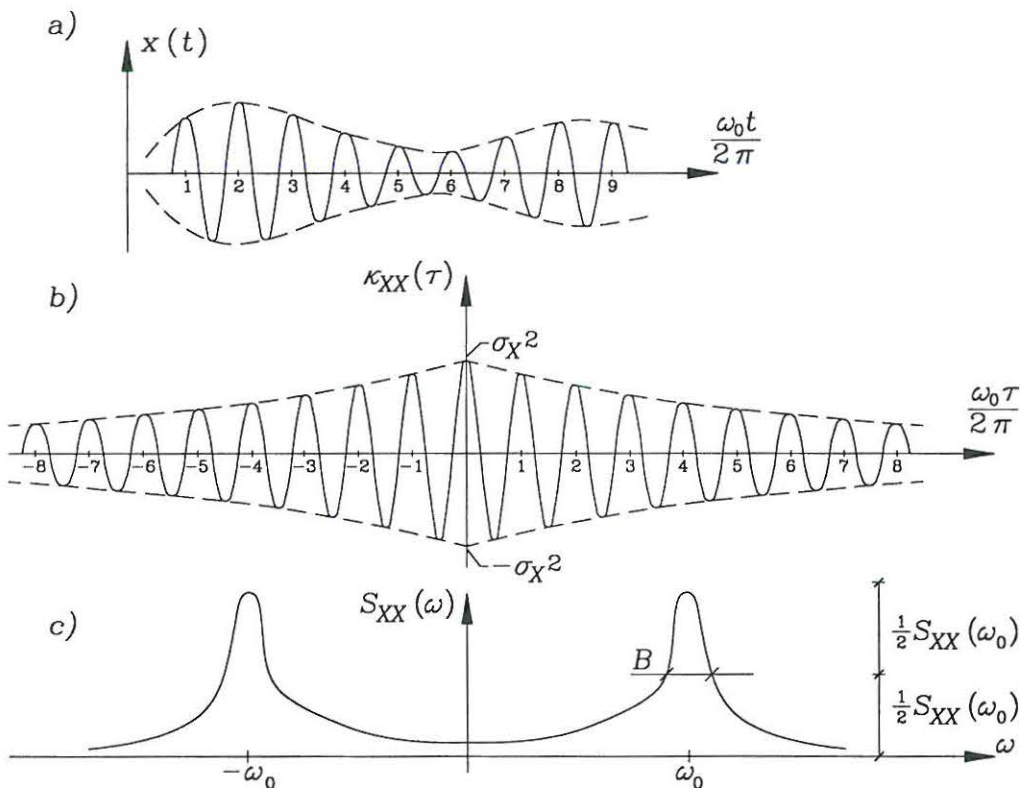

Figure 2.11: Narrow-banded process. a) Realization. b) Auto-covariance function. c) Auto-spectral density function

For physical processes with some resemblance to a harmonic process the variance will be continuously distributed over a range of frequencies, although there will be a marked concentration of variance contributions from a narrow interval close to $\pm\omega_0$. Such a process is named a *narrow-banded stochastic process*.

A realization, the auto-covariance function and the auto-spectral density function for a narrow-banded process are shown in figure 2.11. The realization is approximately harmonic with the angular frequency ω_0, although the amplitude is slowly varying in a random manner. $x(t)$ may be considered to be composed of a sum of harmonic components with slightly different angular frequencies, all in the vicinity of ω_0. The slowly varying amplitudes are due to the superposition of such components. Finally, it is observed that each crossing of the t-axis by $x(t)$ is succeeded by exactly one local maximum or local minimum.

The auto-covariance function shown in figure 2.11b appears as a cosine function with angular frequency ω_0 and with monotonously decreasing amplitudes. The correlation length τ_0 of the process indicates the values of τ, where the local maxima of $\kappa_{XX}(\tau)$ have diminished to insignificant magnitudes. As seen $\tau_0 \gg T_0 = \frac{2\pi}{\omega_0}$ for a narrow-banded process.

The auto-spectral density function of a narrow-band process is shown in figure 2.11c. As a measure of the narrow-bandedness the *half-band width B* may be introduced, defined as shown in the figure. Based on B, a non-dimensional band-width parameter ζ can be introduced as follows

$$\zeta = \frac{B}{2\omega_0} \tag{2.136}$$

Response processes in civil engineering structures sensible to dynamic excitations show narrow-banded character. In this case ω_0 is usually the fundamental circular eigenfrequency of the structure. ζ as given by (2.136) turns out to be a measure of the damping properties of the structure and is typically $\zeta \sim 0.05$.

The surface elevation at a certain site from swell waves appears to some extent as shown in figure 2.11a and can then be classified as narrow-banded.

In wind engineering the most common example of narrow-banded excitation is vortex shedding from slender flexible structures such as chimneys. In this case the vortex shedding process may be locked-in to take place at the fundamental circular eigenfrequency ω_0 of the structure and not at the Strouhal's frequency, resulting in serious resonance.

In equally vague terms, a *broad-banded process* is defined as a process where the variance contributions are uniformly distributed over a wider band of frequencies in a way that no single significantly dominating frequency can be identified in the spectrum. In this respect a broad-banded process has some resemblance to white noise. However, the crucial difference is that any physical broad-banded process will have finite variance, as well as continuous (and even differentiable) realizations.
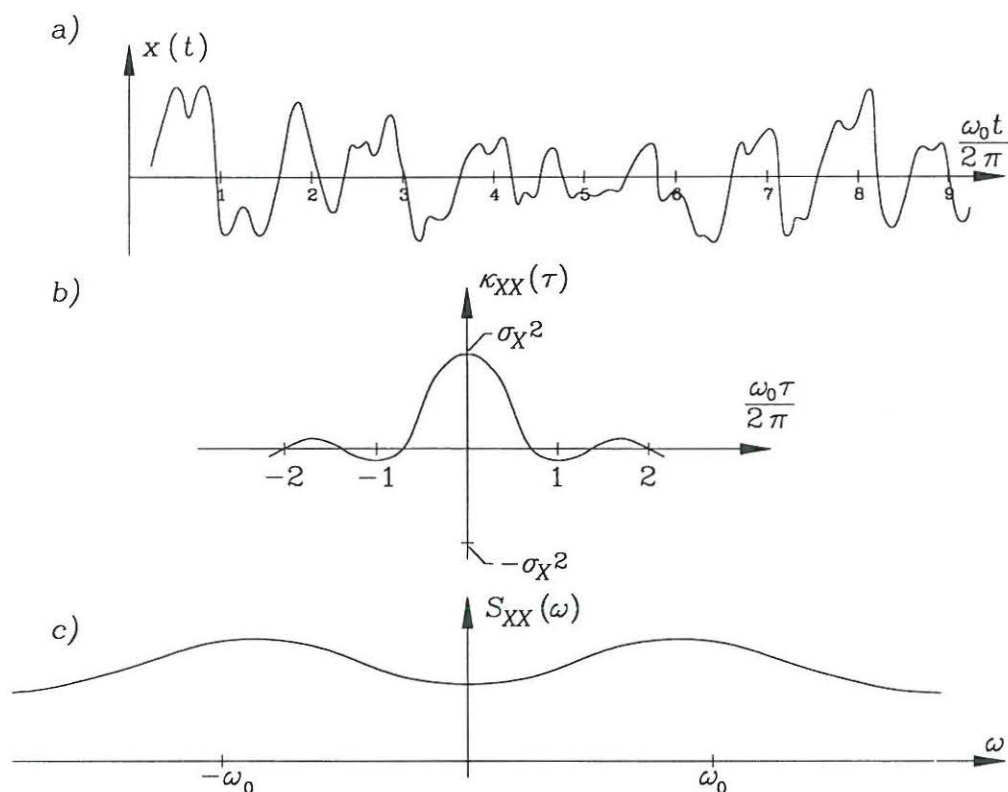
Figure 2.12: Broad-banded process. a) Realization. b) Auto-covariance function. c) Auto-spectral density function.

A realization, the auto-covariance function and the auto-spectral density function of a broad-banded process are shown in figure 2.12. The realization in figure 2.12a has a very irregular performance. Each crossing of the t-axis is quite often succeeded by more than one local maximum or local minimum contrary to the narrow-banded case.

In figures 2.12a and 2.12b the time has been normalized with respect to the *mean zero upcrossing period* $T_0 = \frac{2\pi}{\omega_0}$, which can be estimated from the realization simply by counting the number of upcrossings in a certain sample interval. As seen from figure 2.12b significant correlations are only present for time separations of magnitude T_0, i.e. the correlation length fulfils $\tau_0 \sim T_0$ for a broad-banded process. Finally, the auto-spectral density function of a broad-banded process has been shown in figure 2.12c.

Examples of broad-banded natural loading processes are the surface elevation from the so-called wind waves, i.e. waves which are still in the generation area. Gusty wind loads due to atmospheric turbulence is also of broad-banded character. Most physical phenomena, where t signifies a spatial parameter have realizations, which resemble the time series shown in figure 2.12a and consequently should be categorized as a broad-banded process in case of stochastic modelling. E.g. the yield stress along a long bar will vary in the irregular way shown in figure 2.12a around some mean value.

Negative angular frequencies cannot be interpreted physically. Due to the symmetry of $S_{XX}(\omega)$ one occasionally prefers to operate with the so-called *one-sided auto-spectral*

density function, defined as

$$S_X(\omega) = 2S_{XX}(\omega), \quad \omega \in [0, \infty[\tag{2.137}$$

For the one-sided auto-spectral density function only one index X is applied. $S_{XX}(\omega)$ will correspondingly be referred to as the *double-sided auto-spectral density function*. The Wiener-Khintchine relations (2.100), (2.101) and the result (2.108) for the variance may be expressed in terms of $S_X(\omega)$ as follows

$$\kappa_{XX}(\tau) = \int_0^\infty \cos(\omega\tau) S_X(\omega) d\omega \tag{2.138}$$

$$S_X(\omega) = \frac{2}{\pi} \int_0^\infty \cos(\omega\tau) \kappa_{XX}(\tau) d\tau \tag{2.139}$$

$$\sigma_X^2 = \int_0^\infty S_X(\omega) d\omega \tag{2.140}$$

Based on the one-sided auto-spectral density function the *spectral moments* are defined in the following way

$$\lambda_i = \int_0^\infty \omega^i S_X(\omega) d\omega, \quad i = 0, 1, 2, \ldots \tag{2.141}$$

$S_X(\omega)/\lambda_0$ has the main-characteristics of a probability density function (non-negatively and area $= 1$). The expected value of this 'pdf' is $\frac{\lambda_1}{\lambda_0}$ and the variance is $\left(\frac{\lambda_2}{\lambda_0} - \frac{\lambda_1^2}{\lambda_0^2}\right)$. Hence, the variational coefficient becomes, cf. (1.85)

$$\delta = \frac{\left(\lambda_0 \lambda_2 - \lambda_1^2\right)^{\frac{1}{2}}}{\lambda_1} = \sqrt{\frac{\lambda_0 \lambda_2}{\lambda_1^2} - 1} \tag{2.142}$$

δ was suggested as a band-width parameter by Vanmarcke.[1] For a harmonic process, $\lambda_1 = \omega_0 \lambda_0$ and $\lambda_2 = \omega_0^2 \lambda_0$, implying $\delta = 0$. Hence, small values of δ indicate narrow-bandedness.

Example 2.14: Band-limited white noise

A weakly homogeneous stochastic process is called *band-limited white noise*, if the one-sided auto-spectral density function is given as

---

[1] E. Vanmarcke: *Random fields: Analysis and Synthesis*, The MIT Press, Cambridge, Massachusetts, 1983.
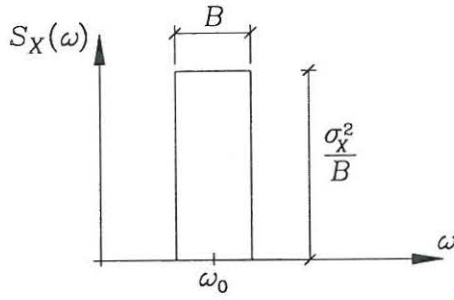
Figure 2.13: Band-limited white noise.

$$S_X(\omega) = \begin{cases} \frac{\sigma_X^2}{B} & , \omega \in \left[\omega_0 - \frac{B}{2}, \omega_0 + \frac{B}{2}\right] \\ 0 & , \omega \notin \left[\omega_0 - \frac{B}{2}, \omega_0 + \frac{B}{2}\right] \end{cases} \tag{2.143}$$

ω_0 is termed the *center frequency* and B is the *band-width* of the auto-spectral density function. These are assumed to fulfil $B \leq 2\omega_0$.

The auto-covariance function follows from (2.138)

$$\kappa_{XX}(\tau) = \int\limits_{\omega_0 - \frac{B}{2}}^{\omega_0 + \frac{B}{2}} \cos(\omega\tau) \frac{\sigma_X^2}{B} d\omega =$$

$$\frac{\sigma_X^2}{B\tau} \left[\sin\left(\left(\omega_0 + \frac{B}{2}\right)\tau\right) - \sin\left(\left(\omega_0 - \frac{B}{2}\right)\tau\right) \right] \tag{2.144}$$

The spectral moments are obtained from (2.141)

$$\lambda_i = \frac{\sigma_X^2}{B} \int\limits_{\omega_0 - \frac{B}{2}}^{\omega_0 + \frac{B}{2}} \omega^i d\omega = \frac{\sigma_X^2}{(i+1)B} \left(\left(\omega_0 + \frac{B}{2}\right)^{i+1} - \left(\omega_0 - \frac{B}{2}\right)^{i+1} \right) \tag{2.145}$$

Vanmarcke's band-width parameter is obtained by inserting (2.145) into (2.142). Eliminating B in favour of the non-dimensional band-width parameter ζ the result may be written as

$$\delta = \sqrt{\frac{1}{3}}\zeta \tag{2.146}$$

Example 2.15: Auto-spectral density function of a process made up of a sum of harmonic processes, and modelling of homogeneous Gaussian processes

In example 2.4 the auto-covariance function of the sum of harmonic processes (2.41) was evaluated and the result was given by (2.43). Using the result (2.131) for the auto-spectral density of each of the components in (2.43), the auto-spectral density function is seen to be

$$S_{XX}(\omega) = \sum_{j=1}^{N} \frac{\sigma_j^2}{2} \left(\delta(\omega - \omega_j)) + \delta(\omega + \omega_j)\right) \tag{2.147}$$
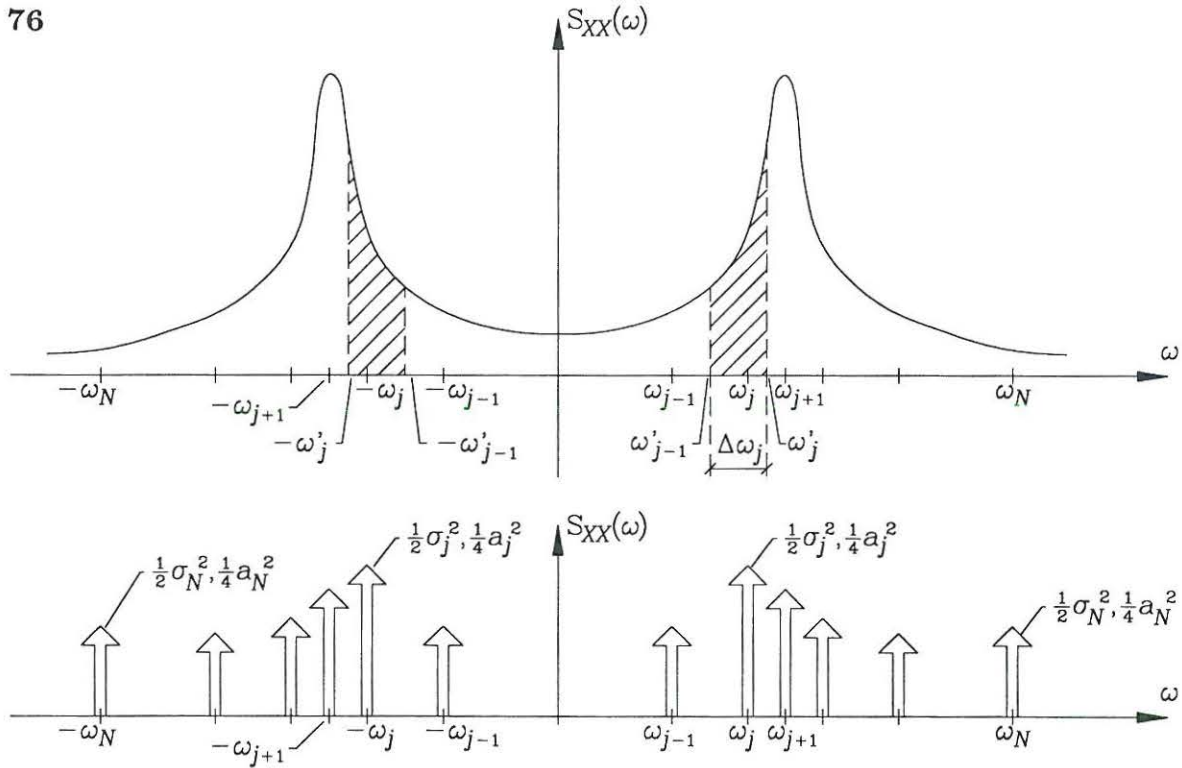
Figure 2.14: a) Discretization of zero-mean stationary Gaussian process with continuous auto-spectral density function.   b) Auto-spectral density function for the process made up of a sum of harmonic processes or a sum of random phase processes.

Similarly, for the sum of random phase processes (2.46) the auto-covariance function was evaluated as given by (2.48). In this case the auto-spectral density function can in the same way be evaluated as follows

$$S_{XX}(\omega) = \sum_{j=1}^{N} \frac{a_j^2}{4}\big(\delta(\omega - \omega_j) + \delta(\omega + \omega_j)\big) \tag{2.148}$$

(2.147) and (2.148) have been depicted in figure 2.14b.

In example 2.4 it was mentioned that as well the sum of harmonic processes as the sum of random phase processes approaches a zero-mean Gaussian process as the number of harmonic components, N, grows to infinity, if only the amplitudes σ_j and a_j are all of comparable magnitude, so no harmonic component is dominating. This observation may be used to model a zero-mean stationary Gaussian stochastic process by an equivalent continuous auto-spectral density function as shown in figure 2.14a. A number of circular frequencies $\omega_1 \ldots \omega_j \ldots \omega_N$ are selected on the frequency axis. To each of these frequencies a harmonic process $R_j \cos(\omega_j t + \Theta_j)$ is associated, which is selected so the intensity $\frac{1}{2}\sigma_j^2$ of the delta-spike at $\omega = \omega_j$ in figure 2.14b and the variance contribution of the hatched area in figure 2.14a are alike, i.e.

$$\frac{1}{2}\sigma_j^2 = \int_{\omega'_{j-1}}^{\omega'_j} S_{XX}(\omega)d\omega \quad \Rightarrow$$

$$\sigma_j^2 = 2\int_{\omega'_{j-1}}^{\omega'_j} S_{XX}(\omega)d\omega = \int_{\omega'_{j-1}}^{\omega'_j} S_X(\omega)d\omega, \quad j = 1,2\ldots N \tag{2.149}$$

where $\omega'_{j-1} = \frac{1}{2}(\omega_{j-1} + \omega_j)$ and $\omega'_j = \frac{1}{2}(\omega_j + \omega_{j+1})$ signifies the midpoints between the circular frequency points $\omega_1 \ldots \omega_N$. These points should be selected, so the variances (2.149) are all of equal magnitude, i.e. the frequency band $\Delta\omega = \omega'_j - \omega'_{j-1} = \frac{1}{2}(\omega_{j+1} - \omega_{j-1})$ should be small, where $S_X(\omega_j)$ is large, in order to insure a convergency to a Gaussian process as $N \to \infty$.

For a sum of random phase processes the corresponding replacement requires that the amplitude a_j of the jth harmonic process fulfils

$$\frac{1}{4}a_j^2 = \int\limits_{\omega'_{j-1}}^{\omega'_j} S_{XX}(\omega)d\omega \quad \Rightarrow$$

$$a_j^2 = 4 \int\limits_{\omega'_{j-1}}^{\omega'_j} S_{XX}(\omega)d\omega = 2 \int\limits_{\omega'_{j-1}}^{\omega'_j} S_X(\omega)d\omega, \quad j = 1, 2 \ldots N \tag{2.150}$$

Again the discrete circular frequencies of the replacement model, $\omega_1 \ldots \omega_N$, should be selected, so a_j, $j = 1, 2 \ldots N$ as given by (2.150) are all of equal magnitude.

2.3.1 Derivative processes of homogeneous processes

If $\{X(t), t \in R\}$ is a strictly or weakly homogeneous stochastic process, it can be shown that the 1st derivative process in mean square, $\{X'(t), t \in R\}$, becomes strictly or weakly homogeneous as well.

The mean value function and auto-covariance function in the homogeneous case follow from (2.61) and (2.64)

$$\mu_{X'}(t) = \frac{d}{dt}\mu_X \equiv 0 \tag{2.151}$$

$$\kappa_{X'X'}(t_1, t_2) = \frac{\partial^2}{\partial t_1 \partial t_2} \kappa_{XX}(t_2 - t_1) = -\frac{d^2}{d\tau^2}\kappa_{XX}(\tau), \quad \tau = t_2 - t_1 \tag{2.152}$$

In the latter statement of (2.152) $\partial t_2 = d\tau$ (with t_1 kept constant) and $\partial t_1 = -d\tau$ (with t_2 kept constant) are applied.

By induction it follows that the nth derivative $\{X^{(n)}(t), t \in R\}$ becomes homogeneous in strict or weak sense, if $\{X(t), t \in R\}$ is strictly or weakly homogeneous ($\{X''(t), t \in R\}$ becomes homogeneous, because so is $\{X'(t), t \in R\}$ and so on).

For the mth and nth derivative processes, $\{X^{(m)}(t), t \in R\}$ and $\{X^{(n)}(t), t \in R\}$, $m, n > 0$, the mean value functions becomes $\mu_{X^{(m)}}(t) \equiv 0$ and $\mu_{X^{(n)}}(t) \equiv 0$. Their cross-covariance function then becomes, cf. (2.70)

$$\kappa_{X^{(m)}X^{(n)}}(t_1, t_2) = \frac{\partial^{m+n}}{\partial t_1^m \partial t_2^n} \kappa_{XX}(t_2 - t_1) =$$

$$(-1)^m \frac{d^{m+n}}{d\tau^{m+n}} \kappa_{XX}(\tau), \quad \tau = t_2 - t_1 \tag{2.153}$$

Especially, for $m = 0$ and $n = 1$ the cross-covariance function of $\{X(t), t \in R\}$ and $\{X'(t), t \in R\}$ is obtained as

$$\kappa_{XX'}(\tau) = \frac{d}{d\tau} \kappa_{XX}(\tau), \quad \tau = t_2 - t_1 \tag{2.154}$$

Since, $\kappa_{XX}(\tau)$ is a symmetric function of τ, its derivative must be zero at $\tau = 0$. Then, $\kappa_{XX'}(0) = \frac{d}{d\tau} \kappa_{XX}(0) = 0$. This shows that for a weakly homogeneous process the stochastic variables $X(t)$ and $X'(t)$ are *uncorrelated*. Further, if $\{X(t), t \in R\}$ is Gaussian it can even be stated that $X(t)$ and $X'(t)$ are *stochastically independent*, since zero correlation implies stochastic independency in this unique case.

Using the Wiener-Khintchine relation (2.100) on the right-hand side of (2.152) and (2.153) one has

$$\kappa_{X'X'}(\tau) = -\frac{d^2}{d\tau^2} \int\limits_{-\infty}^{\infty} e^{i\omega\tau} S_{XX}(\omega) d\omega =$$

$$\int\limits_{-\infty}^{\infty} e^{i\omega\tau} \omega^2 S_{XX}(\omega) d\omega, \quad \tau = t_2 - t_1 \tag{2.155}$$

$$\kappa_{X^{(m)}X^{(n)}}(\tau) = (-1)^m \frac{d^{m+n}}{d\tau^{m+n}} \int\limits_{-\infty}^{\infty} e^{i\omega\tau} S_{XX}(\omega) d\omega =$$

$$\int\limits_{-\infty}^{\infty} e^{i\omega\tau} (-1)^m (i\omega)^{m+n} S_{XX}(\omega) d\omega, \quad \tau = t_2 - t_1 \tag{2.156}$$

Upon comparison with (2.100) the auto-spectral density function $S_{X'X'}(\omega)$ of $\{X'(t), t \in R\}$, and the cross-spectral density function $S_{X^{(m)}X^{(n)}}(\omega)$ of $\{X^{(m)}(t), t \in R\}$ and $\{X^{(n)}(t), t \in R\}$ are then seen to be given as

$$S_{X'X'}(\omega) = \omega^2 S_{XX}(\omega) \tag{2.157}$$

$$S_{X^{(m)}X^{(n)}}(\omega) = (-1)^m (i\omega)^{m+n} S_{XX}(\omega) \tag{2.158}$$

The variance function $\sigma^2_{X^{(n)}}$ of the nth derivative process is simply

$$\sigma^2_{X^{(n)}} = \int\limits_{-\infty}^{\infty} \omega^{2n} S_{XX}(\omega) d\omega = \int\limits_{0}^{\infty} \omega^{2n} S_X(\omega) d\omega = \lambda_{2n}, \quad n = 0, 1, \ldots \tag{2.159}$$

where λ_{2n} are the spectral moments defined in (2.141).

In case of homogeneous processes, the 1st derivative in mean square exists, if $\frac{d^2}{dx^2}\kappa_{XX}(\tau)$ exists and is continuous at $\tau = 0$ according to (2.58). As seen from (2.155) this is tantamount to state that $\omega^2 S_{XX}(\omega)$ is integrable over $]-\infty, \infty[$ because this is the requirement for the existence of its Fourier transform, $\kappa_{X'X'}(\tau)$, as a continuous function. Generally, $\{X^{(n)}(t), t \in R\}$ exists if $\omega^{2n} S_{XX}(\omega)$ is integrable, corresponding to the variance (2.159) is finite.

2.3.2 Integrated processes of homogeneous processes

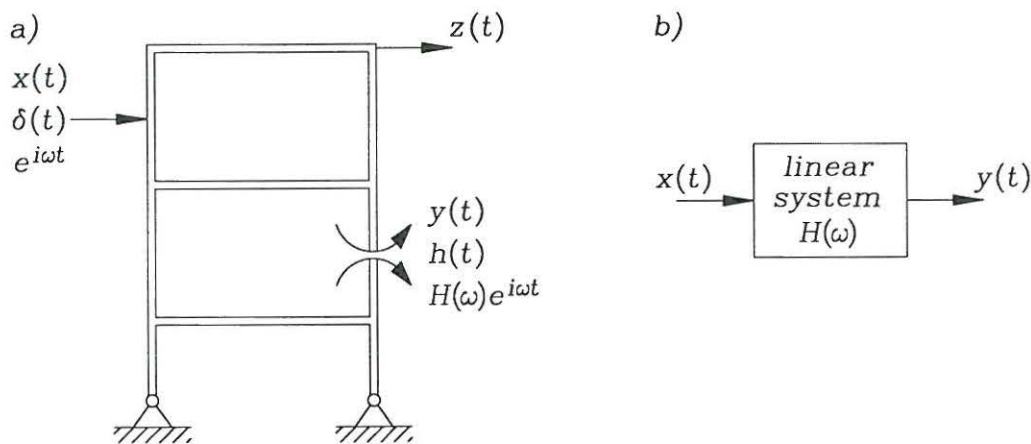### Example 2.16: Impulsive loadings of dynamic systems

Figure 2.15: a) Plane frame exposed to harmonic excitation. b) Symbolic representation of single input – single output relation.

Figure 2.15a shows a plane frame. At some point on the structure a time dependent force $x(t)$ is acting, which causes the dynamic response $y(t)$ somewhere in the structure. Physically, $y(t)$ may be a displacement, a bending moment, etc. The relationship between the input, $x(t)$, and the output, $y(t)$, may be represented symbolically as shown in figure 2.15b.

In physics the *impulse* of a time varying force $x(t)$ is defined as the time integral

$$I = \int x(t)dt \tag{2.160}$$

An *impulsive force* applied at the time τ is zero outside a short interval $[\tau, \tau + d\tau[$, where it is infinite in a way that the integral (2.160) attains a finite value. It follows from (1.19) that an impulsive force with the impulse I applied at the time τ can be written

$$x(t) = I\delta(t - \tau) \tag{2.161}$$

If $I = 1$, the impulsive force is called a *unit impulse*.

Assume the frame is at rest and a unit impulse is applied at the time $t = 0$. The response can then be written as
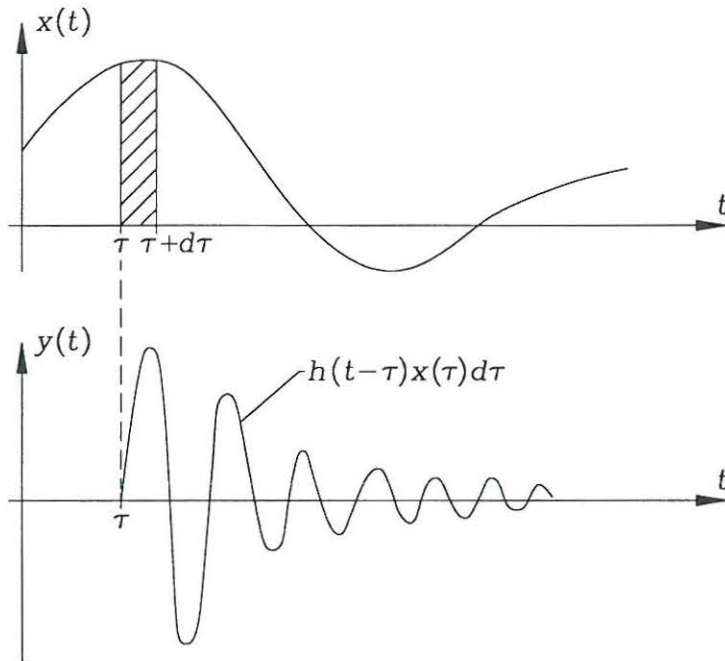
$$y(t) = h(t), \quad t > 0 \tag{2.162}$$

Figure 2.16: Effect of an impulse applied at the time τ.

If the mechanical properties (eigenfrequencies, damping properties) of the frame are not varying with time the response only depends on the elapsed time interval from the instant of time of applying the impulse. Hence, a unit impulse applied at the time $\tau < t$, will imply the response $y(t) = h(t - \tau)$ at the time t. Now, the impulse during the interval $[t, t + \Delta\tau[$ is not 1 but $x(\tau)d\tau$ according to (2.160). Hence, the response from this differential impulse becomes $y(t) = h(t - \tau)x(\tau)d\tau$, see figure 2.16. The response from all such previous differential increments can then be written as the following weighted integral

$$y(t) = \int\limits_{-\infty}^{t} h(t - \tau)x(\tau)d\tau \tag{2.163}$$

(2.163) is nothing but a systematic application of the *superposition principle* for adding load effects in linear systems. Further, future impulses applied at $\tau > t$ cannot affect the present response. This means that the impulse response function must fulfil the following *causality condition*

$$h(t) = 0, \quad t < 0 \tag{2.164}$$

(2.163) is of the weighted integral type (2.75). The reason for using the designation impulse response function for the function $h(t, \tau)$ should now be clear. In the present case $h(t, \tau) = h(t - \tau)$ is merely a function of the elapsed interval since the application of the impulse, since the structural system has been assumed to be time-invariant.

Example 2.17: Frequency response of dynamic systems

Now, assume that the structure shown in figure 2.15 is excited by a harmonic varying force

$$x(t) = X_0 \cos(\omega t) = \text{Re}(X_0 e^{i\omega t}) \tag{2.165}$$

X_0 is the amplitude and ω is the circular frequency of the excitation. The last statement of (2.165) follows from $\text{Re}(e^{ix}) = \text{Re}(\cos x + i \sin x) = \cos x$. Harmonic varying excitations of the indicated type are quite often caused by unbalanced machines operating in the building.

It is now a fact that the structure starts oscillating with the same circular frequency, ω, but with a different phase. Then $y(t)$ can be written, cf. (2.165)

$$y(t) = \text{Re}(Y_0 e^{i\omega t}) \tag{2.166}$$

$$Y_0 = H(\omega) X_0 \tag{2.167}$$

Y_0 as given by (2.167) is the amplitude of the response. The difference between (2.166) and (2.165) is that the amplitude Y_0 of the response is generally complex, whereas X_0 is real. Y_0 can then be written on the polar form

$$Y_0 = | Y_0 | e^{i\Theta} \tag{2.168}$$

(2.166) can then be written

$$y(t) = \text{Re}(| Y_0 | e^{i(\omega t + \Theta)}) = | Y_0 | \text{Re}(e^{i(\omega t + \Theta)}) = | Y_0 | \cos(\omega t + \Theta) \tag{2.169}$$

$| Y_0 |$ is the observed amplitude of the response in the experiment. The meaning of the phase Θ is that the harmonic input and output are not zero at the same time. Phase lags are caused by damping in vibratory systems.

$H(\omega)$ in (2.167) is named the *frequency response function*. Physically, $H(\omega)$ can be interpreted as the complex amplitude of the response $y(t)$, when a unit harmonic input $x(t) = \text{Re}(e^{i\omega t})$ is applied at the structure, cf. figure 2.15.

The impulse response function $h(t)$ and the frequency response function $H(\omega)$ can be characterized as functions, which describe the response of the system to standardized loadings, namely a unit impulse and a unit amplitude harmonic varying excitation. Since, both functions defines one and the same system there should be a close relationship between them. In order to derive this relationship, let $x(\tau) = e^{i\omega\tau}$ in (2.163), where the real operator has been omitted for ease. The response then becomes $y(t) = H(\omega)e^{i\omega t}$. Hence, (2.163) provides

$$H(\omega)e^{i\omega t} = \int\limits_{-\infty}^{t} h(t - \tau)e^{i\omega\tau}\,d\tau \quad \Rightarrow$$

$$H(\omega) = \int\limits_{-\infty}^{t} h(t - \tau)e^{-i\omega(t-\tau)}\,d\tau = -\int\limits_{\infty}^{0} h(u)e^{-i\omega u}\,du = \int\limits_{0}^{\infty} e^{-i\omega u} h(u)\,du \tag{2.170}$$

where the change of integration variable $u = t - \tau$ has been applied. This is possible because the integration is with respect to τ, so t is constant. The lower integration limit in (2.170) can be replaced with $-\infty$ because of the causality condition (2.164), so one has that the frequency response function, $H(\omega)$, is the Fourier transform of the impulse response function, i.e.

$$H(\omega) = \int\limits_{-\infty}^{\infty} e^{-i\omega u} h(u) du \qquad (2.171)$$

From (2.130) then follows that $h(u)$ is obtained by the inverse relation

$$h(u) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{i\omega u} H(\omega) d\omega \qquad (2.172)$$

A system, for which the impulse response function fulfils $h(t, \tau) = h(t - \tau)$ is called *time-invariant*. The reason for this designation follows from the discussion in example 2.16.

Now, assume that $\{X(t), t \in R\}$ is a strictly or weakly homogeneous process, and the system is time-invariant. Corresponding to (2.78) the integrated process is written

$$Y(t) = \int\limits_{-\infty}^{t} h(t - \tau) X(\tau) d\tau \qquad (2.173)$$

The mean value function becomes, cf. (2.80)

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

$$\mu_X \int\limits_{-\infty}^{\infty} h(u) du = \mu_Y \text{ (constant)} \qquad (2.174)$$

where the change of integration variable to $u = t - \tau$ has been applied. In (2.83) the substitutions $u_1 = t_1 - \tau_1$ and $u_2 = t_2 - \tau_2$ provides

$$\kappa_{YY}(t_1, t_2) = \int\limits_{-\infty}^{t_1} \int\limits_{-\infty}^{t_2} h(t_1 - \tau_1) h(t_2 - \tau_2) \kappa_{XX}(\tau_2 - \tau_1) d\tau_1 d\tau_2 =$$

$$\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} h(u_1) h(u_2) \kappa_{XX}(t_2 - t_1 + u_1 - u_2) du_1 du_2 = \kappa_{YY}(t_2 - t_1) \qquad (2.175)$$

The auto-spectral density function of $\{Y(t), t \in R\}$ follows from (2.101) and (2.175)

$$S_{YY}(\omega) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{-i\omega\tau} \kappa_{YY}(\tau)d\tau =$$

$$\frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{-i\omega\tau} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} h(u_1)h(u_2)\kappa_{XX}(\tau + u_1 - u_2)du_1 du_2 d\tau =$$

$$\int\limits_{-\infty}^{\infty} e^{i\omega u_1}h(u_1)du_1 \int\limits_{-\infty}^{\infty} e^{-i\omega u_2}h(u_2)du_2 \cdot$$

$$\frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{-i\omega(\tau+u_1-u_2)}\kappa_{XX}(\tau + u_1 - u_2)d\tau \tag{2.176}$$

In the innermost integral of (2.176) u_1 and u_2 are constants, and a new integration variable $u = \tau + u_1 - u_2$ can be introduced with $du = d\tau$. This integral is then seen to be $S_{XX}(\omega)$. The next integral is seen to be $H(\omega)$. Finally, the outermost integral is evaluated to be $H^\star(\omega)$. (2.176) can then be written as

$$S_{YY}(\omega) = H^\star(\omega)H(\omega)S_{XX}(\omega) = \mid H(\omega) \mid^2 S_{XX}(\omega) \tag{2.177}$$

The result for the outermost integral follows because, cf. (2.102)

$$H^\star(\omega) = \left(\int\limits_{-\infty}^{\infty} e^{-i\omega u}h(u)du \right)^\star = \int\limits_{-\infty}^{\infty} \left(e^{-i\omega u}\right)^\star h(u)du = \int\limits_{-\infty}^{\infty} e^{i\omega u}h(u)du \tag{2.178}$$

(2.177), which is one of the most important relations in spectral analysis of response processes in linear systems, has been illustrated in figure 2.17. $S_{XX}(\omega)$ is most often broad-banded in vibration problems in civil engineering. In this case $\mid H(\omega) \mid^2$ is usually narrow-banded, which is merely a statement that civil engineering structures are lightly damped. The figure then shows that the response processes in structural dynamics most often are narrow-banded, i.e. their sample curves appear as shown in figure 2.11a. ω_0 in figure 2.17 signifies the fundamental circular eigenfrequency of the structure.

For $\omega = 0$, (2.171) provides

$$H(0) = \int\limits_{-\infty}^{\infty} h(u)du = \int\limits_{-\infty}^{\infty} h(t-\tau) \cdot 1 d\tau = \frac{1}{k} \tag{2.179}$$

Hence, (2.174) can be written as

$$\mu_Y = \frac{1}{k}\mu_X \tag{2.180}$$

84

In (2.179) $H(0)$ can be interpreted as the response, when a statical load of magnitude 1 has been applied for infinitely long time. Hence, $H(0) = \frac{1}{k}$ signifies the statical influence coefficient, cf. (1.165). k is the corresponding stiffness coefficient. (2.180) can then be interpreted as the *static response* of the system to the static excitation $x(t) = \mu_X$.
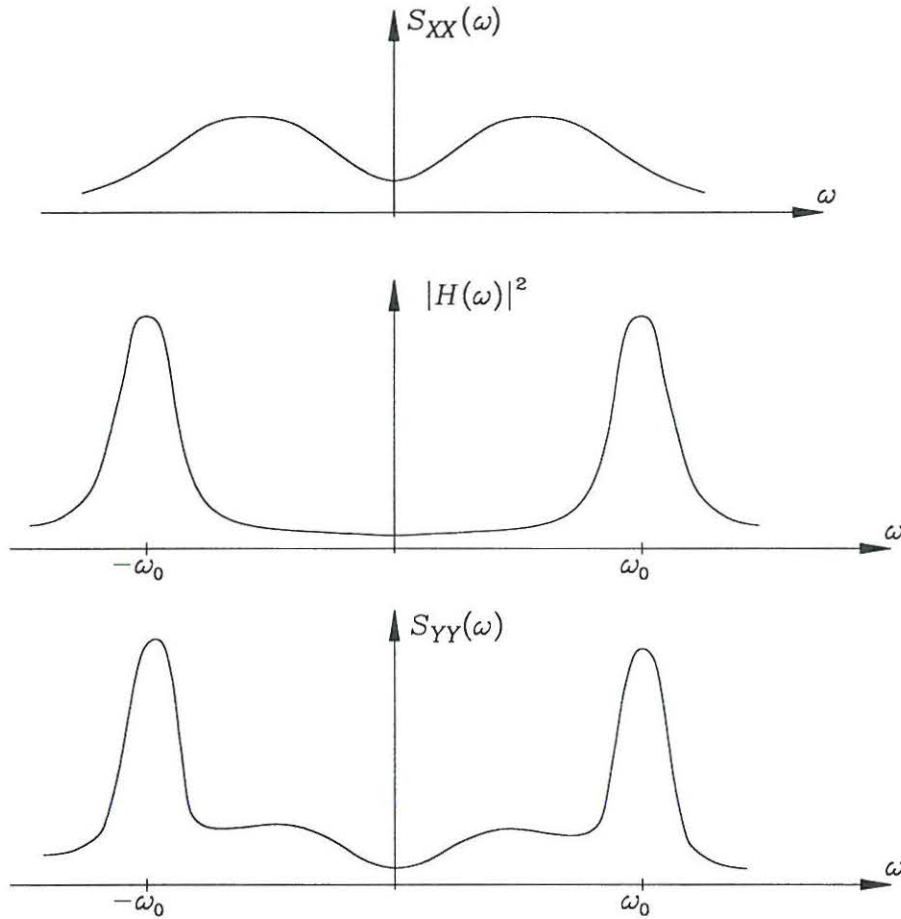

Figure 2.17: Broad-banded excitation process filtered through a narrow-banded filter.

For the plane frame in figure 2.15a one may now consider several response quantities $y(t)$ and $z(t)$ generated by the same time-dependent loading $x(t)$. For the frame shown in figure 2.15a the extra response quantity, $z(t)$, may represent the horizontal top storey displacement. Let the response $y(t)$ be determined by the frequency response function $H_{YX}(\omega)$ and the response $z(t)$ by the frequency response function $H_{ZX}(\omega)$, see figure 2.18. The corresponding impulse response functions $h_{YX}(t)$ and $h_{ZX}(t)$ are determined by (2.172). If the loading is modelled as a weakly stationary stochastic process $\{X(t), t \in R\}$ one may then ask for the cross-correlation of the response processes $\{Y(t), t \in R\}$ and $\{Z(t), t \in R\}$.

The cross-covariance function becomes, cf. (2.175)

$$\kappa_{YZ}(t_1, t_2) = \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} h_{YX}(t_1 - \tau_1) h_{ZX}(t_2 - \tau_2) \kappa_{XX}(\tau_2 - \tau_1) d\tau_1 d\tau_2 =$$

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} h_{YX}(u_1)h_{ZX}(u_2)\kappa_{XX}(t_2 - t_1 + u_1 - u_2)du_1 du_2 \qquad (2.181)$$
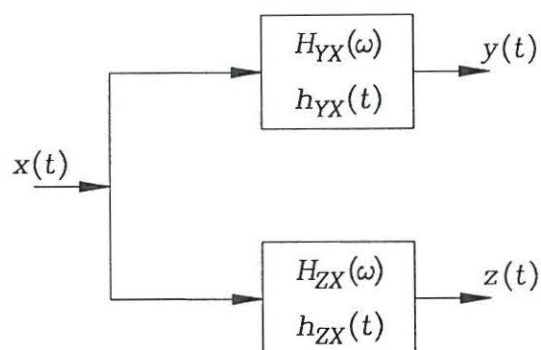

Figure 2.18: Symbolic representation of single input – multiple output system.

The cross-spectral density function of $\{Y(t), t \in R\}$ and $\{Z(t), t \in R\}$ then follows from (2.101). With a derivation simular to the one leading to (2.177) the result becomes

$$S_{YZ}(\omega) = H_{YX}^{\star}(\omega)H_{ZX}(\omega)S_{XX}(\omega) \qquad (2.182)$$

Example 2.18: Cross-spectral density of particle velocity and acceleration processes in long-crested sea states
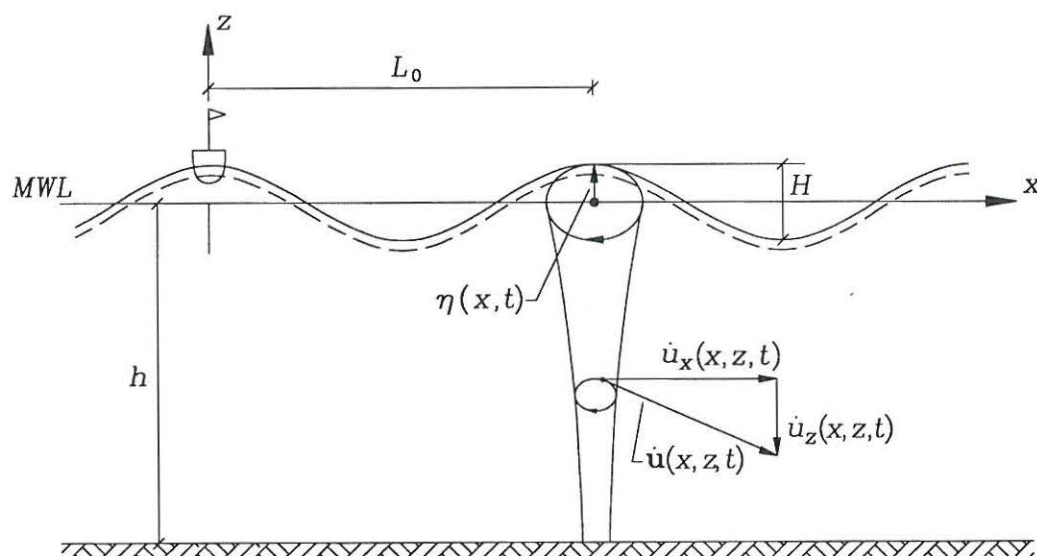

Figure 2.19: Planar linear wave.

86

A planar 2-dimensional irregular sea-state is considered. Initially, a Cartesian (x, y, z)-coordinate system is defined. The (x, y)-plane is horizontal and placed in the mean water level (MWL). The x-axis is directed along the direction of propagation and the z-axis is orientated in the upwards direction. The sea-bottom is horizontal and the water depth is h. The sea-bottom is then defined by the coordinate plane $z = -h$, see figure 2.19.

The surface elevation $\eta(x, t)$ is measured by a wave gauge at the position x_0. The surface elevation at the measure point is modelled by a zero-mean weakly stationary stochastic process $\{\eta(0, t), t \in R\}$. Using the theory of (2.123) the corresponding surface elevation auto-spectral density function $S_{\eta\eta}(\omega)$ can be estimated from the measured time series.

In offshore engineering loadings are caused by the horizontal fluid velocity $\dot{u}_x(x, z, t)$ and horizontal fluid acceleration $\ddot{u}_x(x, z, t)$. The cross-spectral density functions of these quantities will be determined, assuming that fluid velocities and fluid accelerations can be derived from the surface elevation by *linear wave theory* (Airy wave theory).
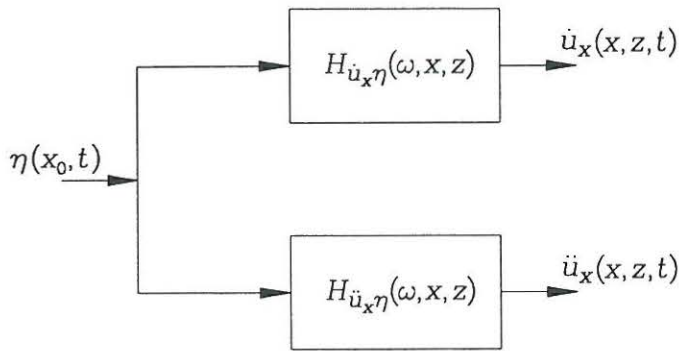

Figure 2.20: Symbolic representation of single input – multiple output system for long crested sea-states

In the present case $\{\eta(x_0, t), t \in R\}$ is the basic generating input to the system. As usual, the frequency response functions $H_{\dot{u}_x\eta}(\omega, x, z)$ and $H_{\ddot{u}_x\eta}(\omega, x, z)$ at the position (x, z) is preliminary determined considering a harmonic varying input $\eta(x_0, t)$, i.e.

$$\eta(x_0, t) = \text{Re}\Big(a \exp(i(\omega t - k(\omega)x_0 + \Theta)) \Big) \tag{2.183}$$

(2.183) can be interpreted as a regular wave-let propagating along the x-axis with the wave-height $H = 2a$, the circular frequency ω, the wave number $k(\omega)$ and the phase Θ. The wave number is related to the circular frequency ω by the *dispersion relation*[1]

$$k \tanh(kh) = \frac{\omega^2}{g} \tag{2.184}$$

where g is the *acceleration of gravity*. For mathematically defined circular frequencies $\omega < 0$, the wave number is formally defined as $k(\omega) = -k(-\omega)$, i.e. $k(\omega)$ becomes an odd function of ω. Since $\tanh(-x) = -\tanh(x)$, $k(\omega)$ also fulfils (2.184) for $\omega < 0$, if only this relation is fulfilled for positive circular frequencies. (2.183) describes the surface elevation at the position x_0. The corresponding surface elevation at the arbitrary position x becomes

$$\eta(x, t) = \text{Re}\Big(a \exp(i(\omega t - k(\omega)x + \Theta)) \Big) =$$

---

[1] R. Wiegel: *Oceanographical engineering*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1964.

$$\text{Re}\left(e^{-ik(\omega)(x-x_0)}a\exp\left(i(\omega t - k(\omega)x_0 + \Theta)\right)\right), \quad x_0 = 0 \tag{2.185}$$

The fluid velocity and the flow acceleration at the position (x,z), $z \in [-h,0]$, then follows from usual linear wave theory[2]

$$u_{\dot{x}}(x,z,t) = \text{Re}\left(H_{\dot{u}_x\eta}(\omega,x,z)a\exp\left(i(\omega t - k(\omega)x_0 + \Theta)\right)\right) \tag{2.186}$$

$$u_{\ddot{x}}(x,z,t) = \frac{\partial}{\partial t}u_{\dot{x}}(x,z,t) = \text{Re}\left(H_{\ddot{u}_x\eta}(\omega,x,z)a\exp\left(i(\omega t - k(\omega)x_0 + \Theta)\right)\right) \tag{2.187}$$

$$H_{\dot{u}_x\eta}(\omega,x,z) = \omega\,\frac{\cosh(k(\omega)(z+h))}{\sinh(k(\omega)h)}\,e^{-ik(\omega)(x-x_0)} \tag{2.188}$$

$$H_{\ddot{u}_x\eta}(\omega,x,z) = (i\omega)\omega\,\frac{\cosh(k(\omega)(z+h))}{\sinh(k(\omega)h)}\,e^{-ik(\omega)(x-x_0)} \tag{2.189}$$

The requested cross-spectral densities then follow immediately from (2.182)

$$S_{\dot{u}_x\dot{u}_x}(\omega; x_1,z_1; x_2,z_2) = H^\star_{\dot{u}_x\eta}(\omega,x_1,z_1)H_{\dot{u}_x\eta}(\omega,x_2,z_2)S_{\eta\eta}(\omega) \tag{2.190}$$

$$S_{\dot{u}_x\ddot{u}_x}(\omega; x_1,z_1; x_2,z_2) = H^\star_{\dot{u}_x\eta}(\omega,x_1,z_1)H_{\ddot{u}_x\eta}(\omega,x_2,z_2)S_{\eta\eta}(\omega) \tag{2.191}$$

$$S_{\ddot{u}_x\ddot{u}_x}(\omega; x_1,z_1; x_2,z_2) = H^\star_{\ddot{u}_x\eta}(\omega,x_1,z_1)H_{\ddot{u}_x\eta}(\omega,x_2,z_2)S_{\eta\eta}(\omega) \tag{2.192}$$

$S_{\dot{u}_x\dot{u}_x}(\omega; x_1,z_1; x_2,z_2)$ specifies the cross-spectral density function of the fluid velocity processes $\{u_{\dot{x}}(x_1,z_1,t), t \in R\}$ and $\{u_{\dot{x}}(x_2,z_2,t), t \in R\}$, etc. It should be noted that the expressions (2.190), (2.191) and (2.192) do not depend on the position x_0 of the wave gauge. As an example

$$S_{\dot{u}_x\dot{u}_x}(\omega; x_1,z_1; x_2,z_2) =$$

$$\omega\,\frac{\cosh(k(\omega)(z_1+h))}{\sinh(k(\omega)h)}\,e^{ik(\omega)(x_1-x_0)}\,\omega\,\frac{\cosh(k(\omega)(z_2+h))}{\sinh(k(\omega)h)}\,e^{-ik(\omega)(x_2-x_0)}S_{\eta\eta}(\omega) =$$

$$\omega^2\,\frac{\cosh(k(\omega)(z_1+h))\cosh(k(\omega)(z_2+h))}{\sinh^2(k(\omega)h)}\,e^{-ik(\omega)(x_2-x_1)} \tag{2.193}$$

Without any restriction the position of the wave gauge can then be chosen as the origin of the coordinate system, so $x_0 = 0$.

3. EXAMPLES OF STOCHASTIC PROCESSES

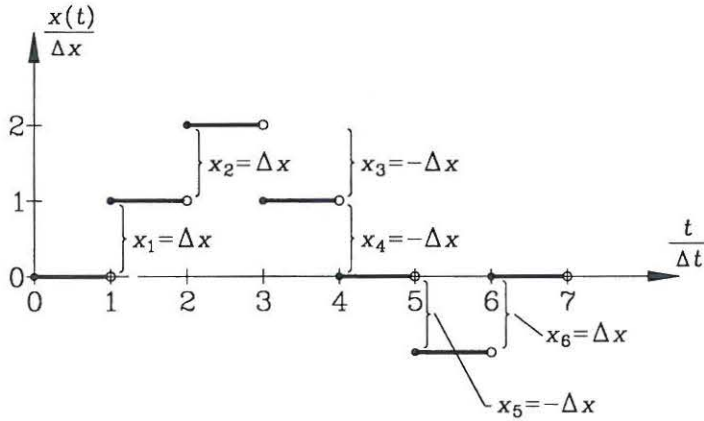## 3.1 One-dimensional random walk process

Figure 3.1: Realization of a one-dimensional random walk process.

Consider a motion starting at the origin along the x-axis. At the instants of time $t = i\Delta t$, $i = 1, 2, \ldots$ a jump of magnitude Δx is performed either in the forward or the backward direction. The jump in the forward direction is performed with the probability p and the jump in the backward direction with the probability $1 - p$. The magnitude of the jump is then described by the discrete stochastic variable X_i with the probability function $P(X_i = \Delta x) = p$ and $P(X_i = -\Delta x) = 1 - p$. In case the sequence $\{X_1, X_2, \ldots\}$ are mutually independent and identical distributed discrete stochastic variable the described process is called a one-dimensional random walk process. It is assumed that X_i is realized at the time $t = i\Delta t$, $i = 1, 2, \ldots$ The one-dimensional random walk process can formally be described as a stochastic process $\{X(t), t \in [0, \infty[\}$, where

$$X(t) = \sum_{i=1}^{\left[\frac{t}{\Delta t}\right]} X_i, \quad X(0) = 0 \tag{3.1}$$

$[x]$ signifies the integer part of x (e.g. $[2.00] = 2$, $[2.99] = 2$). Hence, for any $t \in [i\Delta t, (i+1)\Delta t[$, one has $X(t) = X_1 + \cdots X_i$. Obviously, $\{X(t), t \in R\}$ has a discontinuous staircase like appearance as shown in figure 3.1 and should then be classified as an index continuous state discrete process.

In order to calculate the mean value function and the auto-covariance function of the process the following expectations are calculated for any X_i and X_j

$$E[X_i] = \Delta x P(X_i = \Delta x) - \Delta x P(X_i = -\Delta x) =$$

$$\Delta x p - \Delta x(1 - p) = \Delta x(2p - 1) \tag{3.2}$$

$$E[X_i^2] = \Delta x^2 P(X_i = \Delta x) + (-\Delta x)^2 P(X_i = -\Delta x) =$$

$$\Delta x^2 p + \Delta x^2(1 - p) = \Delta x^2 \tag{3.3}$$

$$E[X_iX_j] = E[X_i]E[X_j] = \Delta x^2 (2p-1)^2, \quad i \neq j \tag{3.4}$$

In (3.4) it has been used that X_i and X_j are stochastic independent for $i \neq j$, so $E[X_iX_j] = E[X_i]E[X_j]$. The mean value function of the one-dimensional random walk then becomes

$$\mu_X(t) = E[X(t)] = \begin{cases} 0 & , \quad \frac{t}{\Delta t} < 1 \\ \sum_{i=1}^{[\frac{t}{\Delta t}]} E[X_1] & , \quad \frac{t}{\Delta t} \geq 1 \end{cases} =$$

$$\begin{cases} 0 & , \quad \frac{t}{\Delta t} < 1 \\ \sum_{i=1}^{[\frac{t}{\Delta t}]} \Delta x(2p-1) & , \quad \frac{t}{\Delta t} \geq 1 \end{cases} =$$

$$\begin{cases} 0 & , \quad \frac{t}{\Delta t} < 1 \\ \Delta x(2p-1)[\frac{t}{\Delta t}] & , \quad \frac{t}{\Delta t} \geq 1 \end{cases} =$$

$$\Delta x(2p-1)\left[\frac{t}{\Delta t}\right] \tag{3.5}$$

In the last statement it has been used that $\left[\frac{t}{\Delta t}\right] = 0$ for $t < \Delta t$, so the requirement $\mu_X(t) = 0$ for $\frac{t}{\Delta t} < 1$ is automatically fulfilled.

At the evaluation of the auto-correlation function of the process it is temporarily assumed that $t_1 \leq t_2$. Then, from (3.2), (3.3) and (3.4)

$$\mu_{XX}(t_1,t_2) = \begin{cases} 0 & , \quad \frac{t_1}{\Delta t} < 1 \\ \sum_{i=1}^{[\frac{t_1}{\Delta t}]} \sum_{j=1}^{[\frac{t_2}{\Delta t}]} E[X_iX_j] & , \quad \frac{t_1}{\Delta t} \geq 1 \end{cases} =$$

$$\begin{cases} 0 & , \quad \frac{t_1}{\Delta t} < 1 \\ \sum_{i=1}^{[\frac{t_1}{\Delta t}]} E[X_i^2] + \sum_{i=1}^{[\frac{t_1}{\Delta t}]} \sum_{j=1}^{[\frac{t_2}{\Delta t}]} E[X_iX_j] & , \quad \frac{t_1}{\Delta t} \geq 1 \end{cases} =$$

$$\begin{cases} 0 & , \quad \frac{t_1}{\Delta t} < 1 \\ \sum_{i=1}^{[\frac{t_1}{\Delta t}]} \Delta x^2 + \sum_{i=1}^{[\frac{t_1}{\Delta t}]} \sum_{\substack{(j=1) \\ (j \neq i)}}^{[\frac{t_2}{\Delta t}]} \Delta x^2 (2p-1)^2 & , \quad \frac{t_1}{\Delta y} \geq 1 \end{cases} =$$

$$\begin{cases} 0 & , \quad \frac{t_1}{\Delta t} < 1 \\ \sum_{i=1}^{[\frac{t_1}{\Delta t}]} \Delta x^2 + \sum_{i=1}^{[\frac{t_1}{\Delta t}]} \sum_{\substack{(j=1) \\ (j \neq i)}}^{[\frac{t_2}{\Delta t}]} \Delta x^2 (2p-1)^2 & , \quad \frac{t_1}{\Delta t} \geq 1 \end{cases} =$$

$$\begin{cases} 0 & , \quad \frac{t_1}{\Delta t} < 1 \\ \left[\frac{t_1}{\Delta t}\right] \Delta x^2 + \sum_{i=1}^{[\frac{t_1}{\Delta t}]} \left(\left[\frac{t_2}{\Delta t}\right]-1\right) \Delta x^2 (2p-1)^2 & , \quad \frac{t_1}{\Delta t} \geq 1 \end{cases} =$$

$$\begin{cases} 0 & , \quad \frac{t_1}{\Delta t} < 1 \\ \left[\frac{t_1}{\Delta t}\right] \Delta x^2 \left(1 + \left(\left[\frac{t_2}{\Delta t}\right]-1\right)(2p-1)^2\right) & , \quad \frac{t_1}{\Delta t} \geq 1 \end{cases} =$$

$$\left[\frac{t_1}{\Delta t}\right] \Delta x^2 \left(1 + \left(\left[\frac{t_2}{\Delta t}\right] - 1\right)(2p-1)^2\right) =$$

$$\left[\frac{t_1}{\Delta t}\right] \Delta x^2 4p(1-p) + \mu_X(t_1)\mu_X(t_2), \quad t_1 \leq t_2 \tag{3.6}$$

where again $\left[\frac{t_1}{\Delta t}\right] = 0$ for $t_1 < \Delta t$ has been used. Further, in the last statement of (3.6), (3.5) has been inserted.

Similarly, for $t_2 \leq t_1$ it can be proved

$$\mu_{XX}(t_1, t_2) = \left[\frac{t_2}{\Delta t}\right] \Delta x^2 4p(1-p) + \mu_X(t_1)\mu_X(t_2), \quad t_2 \leq t_1 \tag{3.7}$$

(3.6) and (3.7) can then be combined into

$$\mu_{XX}(t_1, t_2) = \min\left(\left[\frac{t_1}{\Delta t}\right], \left[\frac{t_2}{\Delta t}\right]\right) \Delta x^2 4p(1-p) + \mu_X(t_1)\mu_X(t_2) \tag{3.8}$$

The indicated auto-correlation function with the value $\mu_{XX}(t_1 t) = \left[\frac{t}{\Delta t}\right] \Delta x^2 4p(1-p) + \left(\left[\frac{t}{\Delta t}\right] \Delta x(2p-1)\right)^2$, but is not continuous on the diagonal $t_1 = t_2 = t$. Hence, the one-dimensional random walk process is not continuous in the mean-square cf. (2-53). It is of course neither continuous with the probability 1.

The auto-covariance function then follows from (2.12), (3,5) and (3.8)

$$\kappa_{XX}(t_1, t_2) = \mu_{XX}(t_1, t_2) - \mu_X(t_1)\mu_X(t_2) =$$

$$\min\left(\left[\frac{t_1}{\Delta t}\right], \left[\frac{t_2}{\Delta t}\right]\right) \Delta x^2 4p(1-p) \tag{3.9}$$

Especially, the variance function becomes

$$\sigma_X^2(t) = \left[\frac{t}{\Delta t}\right] \Delta x^2 4p(1-p) \tag{3.10}$$

Let $P_{\{X\}}(i\Delta x, j\Delta t) = P\big(X(j\Delta t) = i\Delta x\big)$ signify the probability function of the 1st order at the times $t = j\Delta t$, $j = 0, 1, 2, \ldots$ og the process. Obviously, this quantity fulfils the difference equation

$$P_{\{X\}}(i\Delta x, (j+1)\Delta t) =$$

$$P_{\{X\}}\big((i-1)\Delta x, j\Delta t\big)p + P_{\{X\}}\big((i+1)\Delta x, j\Delta t\big)(1-p) \tag{3.11}$$

(3.11) is called the *masterequation* for the one-dimensional random walk process in physics, and describes the development of $P_{\{X\}}(i\Delta x, j\Delta t)$ in space and time. In order to solve (3.11) the initial distribution $P_{\{X\}}(i\Delta x, 0) = 0, i = 0$, at the time $t = 0$ must be known. *Deterministic start* at the point $x = i_0\Delta x$ means that $P_{\{X\}}(i_0\Delta x, 0) = 1$ and $P_{\{X\}}(i\Delta x, 0) = 0$, $i \neq i_0$.

As the number of jumps, $\left[\frac{t}{\Delta t}\right]$, goes to infinity the sum on the right-hand side of (3.1) approaches a normally distributed random variable as follows from the central limit theorem, see section (1.6). From the multi-dimensional generalization of the central limit theorem it can equally be argued that any n-dimensional vector $[X(t_1)\ldots X(t_n)]$ approaches normality as $\min\left(\left[\frac{t_1}{\Delta t}\right]\ldots\left[\frac{t_n}{\Delta t}\right]\right)$ goes to infinity, i.e. when all stochastic variables in the vector are generated by infinitely many jumps. Hence, (3.1) approaches a Gaussian process as the number of jumps goes to infinity.

3.2 Wiener processes. Gaussian white noise

A Wiener process $\{W(t), t \in [0, \infty[\}$ is defined as a one-dimensional random walk in which $p = \frac{1}{2}$, and where $\Delta t \to 0$ and $\Delta x \to 0$ in a way that throughout the limit passing one has

$$\frac{\Delta x^2}{\Delta t} = 2\pi S_0 \tag{3.12}$$

where S_0 is a positive constant.

Since the interval, Δt, between the jumps goes to zero it follows that infinitely many jumps take place in a finite interval $[0, t]$. From the concluding statements of section 3.1 it then follows that a Wiener process is a Gaussian process. Since, the magnitude, Δx, of the jumps goes to zero the sample curves will be continuous. A Wiener process can then be classified as an index continuous state continuous process.

Since a Wiener process is a Gaussian process it is completely defined by its mean value function $\mu_W(t)$ and its auto-covariance function $\kappa_{WW}(t_1, t_2)$, which will be determined next.

Applying $p = \frac{1}{2}$ in (3.5) it follows

$$\mu_X(t) \equiv 0 \tag{3.13}$$

(3.12) holds throughout the limit passing as $\Delta t \to 0$. Hence, one also has for the Wiener process as well

$$\mu_W(t) \equiv 0 \tag{3.14}$$

For any $t \in [0, \infty[$ one has

$$\frac{t_i}{\Delta t} = \left[\frac{t_i}{\Delta t}\right] + \alpha_i \,, \; i = 1, 2, \ldots \tag{3.15}$$

where $\alpha_i \in [0, 1[$. For $p = \frac{1}{2}$, (3.9) can then be written

$$\kappa_{XX}(t_1, t_2) = \min\left(t_1 - \alpha_1 \Delta t, t_2 - \alpha_2 \Delta t\right) \frac{\Delta x^2}{\Delta t} =$$

$$\min\left(t_1 - \alpha_1 \Delta t, t_2 - \alpha_2 \Delta t\right) 2\pi S_0 \quad \Rightarrow$$

$$\kappa_{WW}(t_1, t_2) = \lim_{\Delta t \to 0} \kappa_{XX}(t_1, t_2) = \min(t_1, t_2) 2\pi S_0 \tag{3.16}$$

(3.14) and (3.16) prove the assertions of the Wiener process in example 2.6, cf. (2.55). That the realizations have the appearance shown in figure 2.4 is clear from the construction of the process as a limit case of a one-dimensional random walk process.

Especially, the variance function becomes

$$\sigma_W^2(t) = t \cdot 2\pi S_0 \tag{3.17}$$

Consequently, $W(t) \sim N\left(0, \sigma_W^2(t)\right)$ with the variance given by (3.17).

Formally, a Gaussian white noise process has previously in section 2.3 been defined as the derivative process of a Wiener process. From the construction of the Wiener process as the limit of a one-dimensional random walk process, it follows why this must change discontinuously from $-\infty$ to $+\infty$ infinitely often as explained in section 2.3 subsequent to (2.135)

In (3.11) $p = \frac{1}{2}$ is inserted, and the following reformulation is performed

$$\frac{P_{\{X\}}(i\Delta x, (j+1)\Delta t) - P_{\{X\}}(i\Delta x, j\Delta t)}{\Delta t} =$$

$$\frac{1}{2} \frac{\Delta x^2}{\Delta t} \frac{P_{\{X\}}((i-1)\Delta x, j\Delta t) - 2P_{\{X\}}(i\Delta x, j\Delta t) + P_{\{X\}}((i+1)\Delta x, j\Delta t)}{\Delta X^2} \tag{3.18}$$

For Δx sufficiently small one has $P_{\{X\}}(i\Delta x, j\Delta t) \simeq f_{\{W\}}(i\Delta x, j\Delta t)\Delta x$, where $f_{\{W\}}(x, t)$ signifies the probability density function of the 1st order of the Wiener process. Inserting this result into (3.18) and performing the limit operations $\Delta x \to 0 \wedge \Delta t \to 0 \wedge \frac{\Delta x^2}{\Delta t} = 2\pi S_0$ for which the one-dimensional random walk process converge to a Wiener process, (3.18) is seen to converge to the following partial differential equation, which is the master equation for the Wiener process.

$$\left.\begin{array}{l} \dfrac{\partial f_{\{W\}}(x, t)}{\partial t} = \pi S_0 \dfrac{\partial^2 f_{\{W\}}(x, t)}{\partial x^2} \quad , \quad t > 0 \\[3mm] f_{\{W\}}(x, 0) = \delta(x) \end{array}\right\} \tag{3.19}$$

where the following well-known finite difference operators have been used

$$\frac{\partial f_{\{W\}}(x, j\Delta t)}{\partial t} = \frac{f_{\{W\}}(x, (j+1)\Delta t) - f_{\{W\}}(x, j\Delta t)}{\Delta t} + O(\Delta t) \tag{3.20}$$

$$\frac{\partial^2 f_{\{W\}}(i\Delta x, t)}{\partial x^2} = \frac{f_{\{W\}}((i-1)\Delta x, t) - 2f_{\{W\}}(i\Delta x, t) + f_{\{W\}}((i+1)\Delta x, t)}{\Delta x^2} + O(\Delta x) \quad (3.21)$$

(3.19) is termed the Fokker-Planet equation of the problem, and is a diffusion equation identical to the non-stationary one-dimensional heat equation. In the present case (3.19) specifies the diffusion of the probability mass along the x-axis from the initial deterministic position at $x = 0$, corresponding to the indicated initial distribution $f_{\{W\}}/x, 0) = \delta(x)$. The solution of (3.19) is given as

$$f_{\{W\}}(x, t) = \frac{1}{\sqrt{2\pi}\sigma_W(t)} \exp\left(-\frac{1}{2}\frac{x^2}{\sigma_W^2(t)}\right) \quad (3.22)$$

which is proved by insertion into (3.19).

$\sigma_W(t)$ is given by (3.19). Hence, $W(t) \sim N\left(0, \sigma_w^2(t)\right)$ in agreement with the previous stated result.

A stochastic process $\{X(t), t \in [0, \infty[\}$ is said to have *independent increments*, if for any $0 \leq t_0 < t_1 < \cdots < t_n$ the stochastic variables $X(t_1) - X(t_0)$, $X(t_2) - X(t_1) \ldots X(t_n) - X(t_{n-1})$ are stochastic independent. The quantity

$$\Delta X(t_i) = X(t_{i+1}) - X(t_i) \quad (3.23)$$

is called the *increment*. It will now be shown that the Wiener process $\{W(t), t \in [0, \infty[\}$ has independent increments. Since the Wiener process is Gaussian, its increments, $\Delta W(t_i)$, become jointly normal distributed. To prove that they are mutually independent it then suffices to prove that they are uncorrelated. However, for any $i \neq j$ one has

$$E\left[\Delta W(t_i)\Delta W(t_j)\right] = E\left[(W(t_{i+1}) - W(t_i))(W(t_{j+1}) - W(t_j))\right] =$$

$$E\left[W(t_{i+1})W(t_{j+1})\right] - E\left[W(t_{i+1})W(t_j)\right] -$$

$$E\left[W(t_i)W(t_{j+1})\right] + E\left[W(t_i)W(t_j)\right] =$$

$$2\pi S_0 \min(t_{i+1}, t_{j+1}) - 2\pi S_0 \min(t_{i+1}, t_j) -$$

$$2\pi S_0 \min(t_i, t_{j+1}) + 2\pi S_0 \min(t_i, t_j) \quad (3.24)$$

Assume $i < j$. Then, $\min(t_{i+1}, t_{j+1}) = t_{i+1}$, $\min(t_{i+1}, t_j) = t_{i+1}$, $\min(t_i, t_{j+1}) = t_i$ and $\min(t_i, t_j) = t_i$. Hence,

$$E\left[\Delta W(t_i)\Delta W(t_j)\right] = 2\pi S_0 t_{i+1} - 2\pi S_0 t_{i+1} - 2\pi S_0 t_i + 2\pi S_0 t_i = 0 \quad (3.25)$$

Next, it is shown in the same way that $E\left[\Delta W(t_i)\Delta W(t_j)\right] = 0$ for $j > i$, so one has

$$E\left[\Delta W(t_i)\Delta W(t_j)\right] = 0 , \quad i \neq j \tag{3.26}$$

(3.26) proves the assertion that increments of a Wiener process are uncorrelated and hence stochastic independent. The property that a Wiener process has independent increments is unique among Gaussian processes. Hence, a Wiener process can be defined as a Gaussian process with index set $T = [0,\infty[$, where $P(W(0) = 0) = 1$, and which has independent increments.

Setting $i = j$ in (3-24) one obtains

$$E[\Delta W^2(t_i)] = 2\pi S_0 \left(\min(t_{i+1},t_{i+1}) - \min(t_{i+1},t_i) - \min(t_i,t_{i+1}) + \min(t_i,t_i)\right) =$$

$$2\pi S_o\left(t_{i+1} - t_i - t_i + t_i\right) = 2\pi S_0\left(t_{i+1} - t_i\right) \tag{3.27}$$

Hence, the increments $\Delta W(t_i)$ are mutually independent stochastic variables $N\left(0, 2\pi S_0 \left(t_{i+1} - t_i\right)\right)$.

3.3 Poisson processes. Compound poisson processes

The Poisson process as well as the renewal process to be considered in the next section belong to a class of processes known as counting processes, $\{N(t), t \in]0,\infty]\}$, which were preliminary introduced in example 2.6. $N(t)$ signifies the random number of arrivals during the interval $]0, t]$ up to and including the present time t of some discrete events arriving at the random instants of time $t_1, t_2 \ldots t_{N(t)}$. In example 2.6 the considered events were the arrivals of cars to a cross-road with the traffic light switched to red.

Counting processes are index continuous state discrete stochastic processes. Hence, the specification of their probabilistic structure involves the determination of the family of joint probability functions $P_{\{N\}}(n,t) = P\left(N(t) = n\right)$, $P_{\{N\}}(n_1,t_1;n_2,t_2) = P\left(N(t_1) = n_1 \wedge N(t_2) = n_2\right), \ldots$. Generally, it will be assumed that the counting process only counts the events which are arriving after the time $t = 0$, i.e. $P(N(0) = 0) = 1$.

Formally, a Poisson process $\{N(t), t \in [0,\infty[\}$ is defined as the counting process fulfilling the following conditions

1. $\{N(t), t \in]0,\infty]\}$ has independent increments.

2.   The probability of one arrival in the interval $]t, t + dt]$ is equal to $\nu(t)dt$, where $\nu(t)$ is positive.

3.   The probability of more than 1 arrival in the interval $]t, t+dt]$ is ignorable compared to $\nu(t)dt$.

If ν is constant with time the Poisson process is said to be *stationary*. $\nu(t)$ is called the *intensity* of the Poisson process, and can according to the 2nd condition be interpreted as the probability per unit time of an event.

Counting processes for which the 3rd requirement is fulfilled are called *regular counting processes*. Also the renewal counting process of the next section is regular.

In order to derive the master equation for the Poisson counting process, consider the event $A_k = \{N(t - dt) = n - k \wedge N(t) - N(t - dt) = k\}$ that $n - k$ events have arrived in $]0, t - dt]$ and k events have arrived in $]t - dt, t]$. Since, A_k is just one of several events resulting in n arrivals during $]0, t]$ the event $\{N(t) = n\}$ of exactly n arrivals in $]0, t]$ can be written as the union $A_0 \cup A_1 \cup A_2 \cup \cdots$. Further, the events $A_0, A_1, A_2, \ldots$ must be mutually disjoint, because these events specifies different number of arrivals in $]t - dt, t]$. From (1.6) then follows

$$P\big(N(t) = n\big) = P(A_0 \cup A_1 \cup A_2 \cup \cdots) = P(A_0) + P(A_1) + P(A_3) + \cdots =$$

$$P\big(N(t - dt) = n \wedge N(t) - N(t - dt) = 0\big) +$$

$$P\big(N(t - dt) = n - 1 \wedge N(t) - N(t - dt) = 1\big) +$$

$$P\big(N(t - dt) = n - 2 \wedge N(t) - N(t - dt) = 2\big) + \cdots =$$

$$P\big(N(t - dt) = n\big) P\big(N(t) - N(t - dt) = 0\big) +$$

$$P\big(N(t - dt) = n - 1\big) P\big(N(t) - N(t - dt) = 1\big) +$$

$$P\big(N(t - dt) = n - 2\big) P\big(N(t) - N(t - dt) = 2\big) + \cdots \tag{3.28}$$

In the last statement of (3.28) the 1st condition of independent increments of the Poisson process has been used. Then, $N(t - dt)$ and $N(t) - N(t - dt)$ become independent stochastic variables, so their joint probabilities split into a product, cf. (1.32). Using the 3rd condition that $P\big(N(t) - N(t - dt) > 1\big) = 0$, (3.28) provides the following differential equation for the 1st order probability function for $dt \to 0$

$$P_{\{N\}}(n, t) = P_{\{N\}}(n, t - dt)\big(1 - \nu(t - dt)dt\big) + P_{\{N\}}(n - 1, t - dt)\nu(t - dt)dt \quad \Rightarrow$$

$$\frac{d}{dt} P_{\{N\}}(n, t) = \frac{1}{dt}\big(P_{\{N\}}(n, t) - P_{\{N\}}(n, t - dt)\big) =$$

$$\big(P_{\{N\}}(n - 1, t) - P_{\{N\}}(n, t)\big)\nu(t) \tag{3.29}$$

(3.29) should be solved with the initial value $P_{\{N\}}(n, 0) = P(N(0) = n) = 0$, $n > 1$, specifying that the probability that any number of events n has arrived at the time $t = 0$ is zero, corresponding to $P_{\{N\}}(0, 0) = P\big(N(0) = 0\big) = 1$. The solution of (3.29) with the indicated initial condition can be shown to be

$$P_{\{N\}}(n, t) = \frac{1}{n!}\left(\int_0^t \nu(\tau)d\tau\right)^n \exp\left(-\int_0^t \nu(\tau)d\tau\right), \quad n = 0, 1, \ldots \tag{3.30}$$

That (3.30) fulfils (3.29) with the indicated initial values is most easily proved by insertion. As seen from table 1.1, $P_{\{N\}}(n,t)$ has a Poisson distribution with the parameter $\lambda = \int_0^t \nu(\tau)d\tau$. Especially, for a stationary Poisson process $\lambda = \nu t$ and (3.30) becomes

$$P_{\{N\}}(n,t) = \frac{1}{n!}(\nu t)^n \exp(-\nu t), \quad n = 0, 1, \dots \tag{3.31}$$

Let $X \sim \mathrm{Po}(\lambda)$. Then the expected value and 2nd order moment become

$$E[X] = \sum_{n=0}^{\infty} nP(X = n) = \sum_{n=0}^{\infty} n\frac{\lambda^n}{n!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{n=1}^{\infty}\frac{\lambda^{n-1}}{(n-1)!} =$$

$$\lambda e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!} = \lambda e^{-\lambda}e^{\lambda} = \lambda \tag{3.32}$$

$$E[X^2] = \sum_{n=0}^{\infty} n^2 P(X = n) = \sum_{n=0}^{\infty} n^2\frac{\lambda^n}{n!}e^{-\lambda} = \sum_{n=0}^{\infty}\left(\frac{n(n-1)+n}{n!}\right)\lambda^n e^{-\lambda} =$$

$$\lambda^2 e^{-\lambda}\sum_{n=2}^{\infty}\frac{\lambda^{n-2}}{(n-2)!} + \lambda e^{-\lambda}\sum_{n=1}^{\infty}\frac{\lambda^{n-1}}{(n-1)!} =$$

$$\lambda^2 e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!} + \lambda e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!} = \lambda^2 + \lambda \tag{3.33}$$

From (3.32) follows immediately that the mean value function of $\{N(t), t \in]0, \infty]\}$ in the stationary case is given by

$$\mu_N(t) = E[N(t)] = \lambda = \nu t \tag{3.34}$$

In order to evaluate the auto-correlation function $\mu_{NN}(t_1, t_2)$ it is temporarily assumed that $t_1 \le t_2$. Since $N(t_2) = N(t_1) + (N(t_2) - N(t_1))$, and $N(t_1)$ and $N(t_2) - N(t_1)$ are stochastic independent, because the Poisson process has independent increments, it follows from (3.33) and (3.34) that

$$\mu_{NN}(t_1, t_2) = E[N(t_1)N(t_2)] = E[N^2(t_1)] + E[N(t_1)(N(t_2) - N(t_1))] =$$

$$E[N^2(t_1)] + E[N(t_1)]E[N(t_2) - N(t_1)] =$$

$$E[N^2(t_1)] - (E[N(t_1)])^2 + E[N(t_1)]E[N(t_2)] =$$

$$(\nu t_1)^2 + \nu t_1 - (\nu t_1)^2 + (\nu t_1)(\nu t_2) = \nu t_1 + \nu^2 t_1 t_2, \quad t_1 \le t_2 \tag{3.35}$$

Similarly, for $t_2 \geq t_1$ it can be proved

$$\mu_{NN}(t_1, t_2) = \nu t_2 + \nu^2 t_1 t_2, \quad t_2 \leq t_1 \tag{3.36}$$

(3.35) and (3.36) can be combined into

$$\mu_{NN}(t_1, t_2) = \nu \min(t_1, t_2) + \nu^2 t_1 t_2 \tag{3.37}$$

(3.37) process the preliminary assertion (2.54). The auto-covariance function then becomes

$$\kappa_{NN}(t_1, t_2) = \mu_{NN}(t_1, t_2) - \mu_N(t_1)\mu_N(t_2) = \nu \min(t_1, t_2) \tag{3.38}$$

Especially, the variance function becomes

$$\sigma_N^2(t) = \nu t \tag{3.39}$$

$N(t + dt) - N(t)$ signifies the number of events in $]t, t + dt]$. Using the 3rd condition of the Poisson counting process the expected value of this increment becomes

$$E\left[N(t + dt) - N(t)\right] = \sum_{n=0}^{\infty} nP\left(N(t + dt) - N(t) = n\right) =$$

$$0 \cdot P\left(N(t + dt) - N(t) = 0\right) + 1 \cdot P\left(N(t + dt) - N(t) = 1\right) +$$

$$2 \cdot P\left(N(t + dt) - N(t) = 2\right) + \cdots =$$

$$1 \cdot P\left(N(t + dt) - N(t) = 1\right) = \nu(t)dt \tag{3.40}$$

Hence, $\nu(t)$ can alternatively be interpreted as the expected number of events per unit time. The identity of this interpretation and the previous one is a consequence of the regularity condition 3 of the counting process.

Especially, for a stationary Poisson process follows for any $s, t > 0$

$$P\left(N(t + s) - N(s) = k\right) = P\left(N(t) = k\right) = \frac{(\nu t)^k}{k!} e^{-\nu t} \tag{3.41}$$
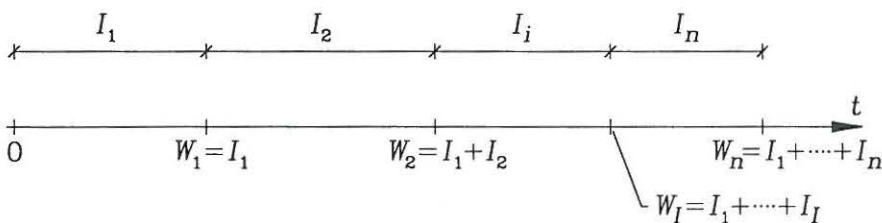

Figure 3.2: Realization of interarrival and waiting times.

Let I_1 be the time of the first event, and more generally let I_n be the time between the $(n-1)$th and the nth events. Then, the sequence of random variables $\{I_n, n = 1, 2, \dots\}$ is called the sequence of *interarrival times*. The *waiting time*, W_n, of the nth event is then defined as

$$W_n = I_1 + \cdots + I_n, \quad n = 1, 2, \dots \tag{3.42}$$

where $W_0 = 0$.

For the interarrival times the following theorem is valid.

Theorem 3.1

The interarrival times I_n, $n = 1, 2, \dots$, of a stationary Poisson counting process are independent and identical exponential distributed, $I_n \sim \mathrm{E}(\nu)$.

First, the theorem is proved for $n = 1$.

$$F_{I_1}(t) = P(I_1 \le t) = 1 - P(I_1 > t) = 1 - P(N(t) = 0) = 1 - e^{-\nu t} \quad \Rightarrow$$

$$f_{I_1}(t) = \frac{d}{dt} F_{I_1}(t) = \nu e^{-\nu t}, \quad t > 0 \tag{3.43}$$

For I_2 the conditional distribution is initially derived by assuming that the first event takes place at some time s. Then from (3.41)

$$F_{I_2|I_1}(t \mid s) = P(I_2 \le t \mid I_1 = s) = 1 - P(I_2 > t \mid I_1 = s) =$$

$$1 - P(N(t+s) - N(s) = 0 \mid I_1 = s) =$$

$$1 - P(N(t+s) - N(s) = 0) \quad \text{(due to independent increments)} \;\; =$$

$$1 - P(N(t) = 0) \quad \text{(due to the stationarity)} \;\; = 1 - e^{-\nu t} \tag{3.44}$$

Hence, the conditional distribution $F_{I_2|I_1}(t \mid s)$ is independent of I_1 and $I_2 \sim \mathrm{E}(\nu)$. Recursively it can be shown that each arrival time I_n is independent and exponentially distributed, which proves the theorem.

A *renewal process* can be defined as a counting process for which the interarrival times are independent and identical distributed, determined by some distribution function $F_I(t)$. Hence, a Poisson counting process can be characterized as a renewal process with exponential distributed interarrival times, $F_I(t) = 1 - e^{-\nu t}$.

(3.42) can be written

$$W_n = W_{n-1} + I_n \tag{3.45}$$

Since the interarrival times are independent, $W_{n-1} = I_1 + \cdots + I_{n-1}$ and I_n are stochastic independent. The distribution function $F_{W_n}(t)$ of the waiting time of the nth event then fulfils, cf. (1-1.57)

$$F_{W_n}(t) = P(W_n \le t) = P(W_{n-1} + I_n \le t) =$$

$$\int\limits_0^\infty P\big(W_{n-1} \leq t - u \mid I_n = u\big) f_{I_n}(u)\,du = \int\limits_0^\infty F_{W_{n-1}}(t-u) f_I(u)\,du =$$

$$\int\limits_0^t F_{W_{n-1}}(t-u) f_I(u)\,du, \quad n = 1, 2, \ldots \tag{3.46}$$

where it has been used that W_{n-1} and I_n are independent. In the last statement of (3.46) the upper integration limit has been reduced from ∞ to t because $F_{W_{n-1}}(u) = 0$ for $u < 0$. For $f_I(u) = \nu e^{-\nu u}$ the solution of (3.46) can be shown to be

$$F_{W_n}(t) = \sum_{i=n}^\infty \frac{(\nu t)^i}{i!} e^{-\nu t} \tag{3.47}$$

(3.47) can be proved by insertion into (3.46). However, this is more easily seen from the fact that $W_n \leq t$ if n or more events have occurred in the interval $]0, t]$. Hence,

$$P(W_n \leq t) = P\big(N(t) \geq n\big) = \sum_{i=n}^\infty P\big(N(t) = i\big) = \sum_{i=n}^\infty \frac{(\nu t)^i}{i!} e^{-\nu t} \tag{3.48}$$
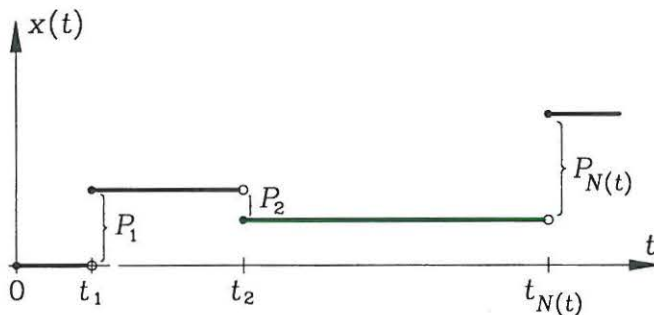

Figure 3.3: Realization of compound Poisson Process.

The jumps of the Poisson counting process at the instant of times $t_1, t_2 \ldots t_{N(t)}$ are always of magnitude 1 as shown in figure 2.5. Now, consider a more general case, where the magnitude of the jumps at the time t_i is P_i, where $P_1, P_2 \ldots P_{N(t)}$ are identical distributed and mutually independent stochastic variables with the probability density function $f_P(p)$. Moreover, $P_1, P_2 \ldots P_{N(t)}$ are stochastic independent of the random arrival times. The indicated process $\{X(t), t \in [0, \infty[\}$ is named a *compound Poisson process*. A realization of the process is shown in figure 3.3. From its construction it is clear that the compound Poisson process also has independent increments and

exponential distributed interarrival times. Obviously, the value $X(t)$ at the time t of the compound Poisson is the sum of all the previous jumps $P_1, P_2, \ldots, P_{N(t)}$. Thus

$$X(t) = \sum_{i=1}^{N(t)} P_i \tag{3.49}$$

The mean value, auto-covariance function and variance function of the compound Poisson process can be shown to be

$$\mu_X(t) = E[P] \int_0^t \nu(\tau)d\tau \tag{3.50}$$

$$\kappa_{XX}(t_1, t_2) = E[P^2] \int_0^{\min(t_1, t_2)} \nu(\tau)d\tau \tag{3.51}$$

$$\sigma_X^2 = E[P^2] \int_0^t \nu(\tau)d\tau \tag{3.52}$$

3.4 Renewal processes

The Poisson counting process was initially presented in example 2.6 as a specification of the number of cars arriving at a traffic light, which has switched to red. The Poisson counting process is characterized by the fact that the interarrival times between car arrivals are mutual independent and identically exponentially distributed (in the stationary case), which indeed is the case in little traffic. However, in heavy traffic the interarrival times are no longer exponential distributed. Instead, the cars tend to arrive with a constant (deterministic) interarrival time. In order to handle such cases, it seems quite naturally to generalize the stationary Poisson counting process to allow for other distributions of the interarrival times than the exponential distribution. If the interarrival times $I_1, I_2 \ldots I_n, \ldots$ of the counting process $\{N(t), t \in [0, \infty[\}$ are mutually independent and identically distributed stochastic variables with the distribution function $F_I(t)$, the counting process is called a *renewal process*. Correspondingly, the events registered by the counting process are called *renewals*.

The waiting time W_n until the arrival of the nth event is given by (3.42). Since, the interarrival times are still mutually independent and identically distributed, the distribution function $F_{W_n}(t)$ must still fulfil the recursive integral equation (3.46).

Analog to (3.48) the probability function $P_{\{N\}}(n, t) = P(N(t) = n)$ of the renewal counting process can be calculated as follows

$$P(W_n \leq t) = P(N(t) \geq n) \quad \Rightarrow$$

$$P_{\{N\}}(n, t) = P(N(t) = n) = P(N(t) \geq n) - P(N(t) \geq n+1) = F_{W_n}(t) - F_{W_{n+1}}(t) \tag{3.53}$$

The mean value function, $\mu_N(t)$, of the renewal process is called the *renewal function*. $\mu_N(t) = E[N(t)]$ signifies the mean number of renewals up to and including the time t. Using (3.53) one has

$$\mu_N(t) = E[N(t)] = \sum_{n=1}^{\infty} nP(N(t) = n) = \sum_{n=1}^{\infty} n\big(F_{W_n}(t) - F_{W_{n+1}}(t)\big) =$$

$$\sum_{n=1}^{\infty} \big(nF_{W_n}(t) - (n+1)F_{W_{n+1}}(t) + F_{W_{n+1}}(t)\big) =$$

$$\big(F_{W_1}(t) + 2F_{W_2}(t) + 3F_{W_3}(t) + \cdots\big) - \big(2F_{W_2}(t) + 3F_{W_3}(t) + \cdots\big) +$$

$$\big(F_{W_2}(t) + F_{W_3}(t) + \cdots\big) =$$

$$F_{W_1}(t) + F_{W_2}(t) + F_{W_3}(t) + \cdots = \sum_{n=1}^{\infty} F_{W_n}(t) \tag{3.54}$$

From (3.46) and (3.54) it then follows

$$\mu_N(t) = F_{W_1}(t) + \sum_{n=2}^{\infty} F_{W_n}(t) = F_{W_1}(t) + \sum_{n=1}^{\infty} F_{W_{n+1}}(t) =$$

$$F_I(t) + \sum_{n=1}^{\infty} \int_0^t F_{W_n}(t-u)f_I(u)du =$$

$$F_I(t) + \int_0^t \left(\sum_{n=1}^{\infty} F_{W_n}(t-u)\right) f_I(u)du =$$

$$F_I(t) + \int_0^t \mu_N(t-u)f_I(u)du \tag{3.55}$$

where it has been used that $W_1 = I_1$, so $F_{W_1}(t) = F_I(t)$. (3.55) is called the *renewal equation*, and it is an integral equation for the determination of the renewal function.

The derivative, $\nu_N(t) = \frac{d}{dt}\mu_N(t)$, of the renewal function is called the *renewal density*. Upon differentiating of (3.55) with respect to t this is seen to fulfil the integral equation

$$\nu_N(t) = f_I(t) + \int_0^t \nu_N(t-u)f_I(u)du + \mu_N(t-t)f_I(t) =$$

$$f_I(t) + \int_0^t \nu_N(t-u)f_I(u)du \tag{3.56}$$

where it has been used that $\mu_N(0) = 0$. (3.56) is also sometimes called the *renewal equation*.

The Poisson counting process is uniquely defined as a regular counting process with independent increments. Since a stationary Poisson counting process is the unique renewal process with exponential distributed interarrival times, it follows that other renewal processes cannot have independent increments. This fact makes the analysis of renewal counting processes somewhat more involved.

Example 3.1: Renewal processes with Gamma distributed interarrival times
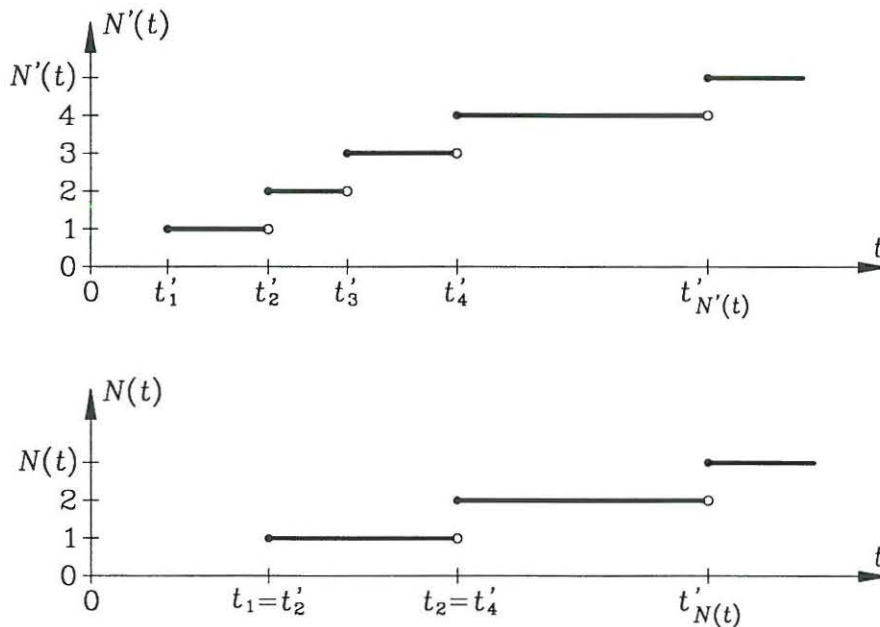

Figure 3.4: a) Realization of stationary Poisson counting process. b) Realization of renewal counting process with Gamma distributed interarrival times.

In figure 3.4a is shown a realization of a stationary Poisson counting process $\{N'(t), t \in [0,\infty[\}$ with mean arrival rate ν, i.e. the interarrival times are exponentially distributed, $I_n \sim E(\nu)$. The events, which arrive at the random times $t_1', t_2', t_3', t_4', \ldots$, may specify significant loadings on a structure. If the loadings are not likely to destroy the structure totally, it may be decided as repair strategy only to repair the structure after every second load event. The renewals of the structure then take place at the times $t_1 = t_2', t_2 = t_4', \ldots$, see figure 3.4b. The number of renewals is specified by the counting

process $\{N(t), t \in [0, \infty[\}$, which picks out every second of the load events as specified by the Poisson counting process. Of course, the probabilistic structure of this counting process is of interest in order to estimate the future repair expenses of the structure.

Obviously, the considered counting process is a renewal counting process with the interarrival time I defined by

$$I = I_1 + I_2 \tag{3.57}$$

where I_1 and I_2 are mutually independent, identically exponential distributed stochastic variables with the parameter ν, $I_1 \sim \mathrm{E}(\nu)$, $I_2 \sim \mathrm{E}(\nu)$. The distribution function of I can then be obtained from, cf. (3.46)

$$F_I(t) = \int\limits_0^t F_{I_1}(t-u) f_{I_2}(u) du = \int\limits_0^t \left(1 - e^{-\nu(t-u)}\right) \nu e^{-\nu u} du =$$

$$1 - e^{-\nu t} - \nu t e^{-\nu t} \quad \Rightarrow$$

$$f_I(t) = \frac{d}{dt} F_I(t) = \nu^2 t e^{-\nu t} \tag{3.58}$$

Since, $\Gamma(1+1) = 1 \cdot \Gamma(1) = 1$, (3.58) is seen to specify a Gamma distribution with the parameters $\alpha = 1$ and $\beta = \nu$, $I \sim \mathrm{Ga}(1, \nu)$, cf. table 1.2.

The probability of n or more renewals, $P(N(t) \geq n)$, is equal to the probability of $2n$ or more load events as determined by the Poisson counting process, i.e.

$$P(N(t) \geq n) = P\left(N'(t) \geq 2n\right) \tag{3.59}$$

Then, the probability function becomes, cf. (3.31)

$$P_{\{N\}}(n, t) = P(N(t) = n) = P(N(t) \geq n) - P(N(t) \geq n+1) =$$

$$P(N'(t) \geq 2n) - P(N'(t) \geq 2n+2) =$$

$$\sum_{i=2n}^{\infty} \frac{1}{i!}(\nu t)^i e^{-\nu t} - \sum_{i=2n+2}^{\infty} \frac{1}{i!}(\nu t)^i e^{-\nu t} = \sum_{i=2n}^{2n+1} \frac{1}{i!}(\nu t)^i e^{-\nu t} \tag{3.60}$$

Next, the renewal function can be calculated

$$\mu_N(t) = \sum_{n=1}^{\infty} n P(N(t) = n) = \sum_{n=1}^{\infty} n \left(\frac{1}{(2n)!}(\nu t)^{2n} + \frac{1}{(2n+1)!}(\nu t)^{2n+1}\right) e^{-\nu t} =$$

$$\sum_{n=1}^{\infty} \frac{1}{2} \left(\frac{2n}{(2n)!}(\nu t)^{2n} + \frac{2n+1-1}{(2n+1)!}(\nu t)^{2n+1}\right) e^{-\nu t} =$$

$$\frac{1}{2}\nu t \sum_{n=1}^{\infty} \left(\frac{(\nu t)^{2n-1}}{(2n-1)!} + \frac{(\nu t)^{2n}}{(2n)!}\right) e^{-\nu t} - \frac{1}{2} \sum_{n=1}^{\infty} \frac{(\nu t)^{2n+1}}{(2n+1)!} e^{-\nu t} \tag{3.61}$$

Now, $e^{\nu t} = \sum_{i=0}^{\infty} \frac{(\nu t)^i}{i!} = 1 + \sum_{n=1}^{\infty} \left(\frac{(\nu t)^{2n-1}}{(2n-1)!} + \frac{(\nu t)^{2n}}{(2n)!} \right)$ and $e^{-\nu t} = \sum_{i=0}^{\infty} \frac{(-\nu t)^i}{i!}$. So, $\frac{1}{2}(e^{\nu t} - e^{-\nu t}) = \sum_{n=0}^{\infty} \frac{(\nu t)^{2n+1}}{(2n+1)!} = \nu t + \sum_{n=1}^{\infty} \frac{(\nu t)^{2n+1}}{(2n+1)!}$. (3.60) can then be written

$$\mu_N(t) = \frac{1}{2}\nu t(e^{\nu t} - 1)e^{-\nu t} - \frac{1}{2} \left(\frac{1}{2}(e^{\nu t} - e^{-\nu t}) - \nu t \right) e^{-\nu t} = \frac{\nu}{2}t - \frac{1}{4} + \frac{1}{4}e^{-2\nu t} \tag{3.62}$$

As seen from (3.62), $\mu_N(t)$ approaches asymptotically to $\frac{\nu}{2}t$ as $t \to \infty$, i.e. to the mean value function of a Poisson counting process with the mean arrival rate $\frac{\nu}{2}$. However, the indicated renewal process *will not* approach to a Poisson counting process with the mean arrival rate $\frac{\nu}{2}$. Of course, this is so because all the interarrival times remain Gamma distributed, $\mathrm{Ga}(1, \nu)$, and do not converge to an exponential distribution, $\mathrm{E}(\frac{\nu}{2})$.

The indicated results (3.58), (3.60) and (3.61) can be generalized to the case, where the renewal of the structure is performed after every kth event of the Poisson counting process. Then, the interarrival time I of the renewal process is defined by

$$I = I_1 + \cdots + I_k, \quad k = 1, 2, \ldots \tag{3.63}$$

where $I_j \sim \mathrm{E}(\nu)$, $j = 1 \ldots k$. It can then be shown that in this case I becomes Gamma distributed with the parameters $\alpha = k - 1$ and $\beta = \nu$, $I \sim \mathrm{Ga}(k - 1, \nu)$, i.e.

$$f_I(t) = \frac{\nu(\nu t)^{k-1}}{(k-1)!} e^{-\nu t}, \quad t \geq 0 \tag{3.64}$$

where it has been used that $\Gamma(k) = (k-1)!$.

The probability function corresponding to (3.60) becomes

$$P_{\{N\}}(n, t) = P(N(t) = n) = \sum_{i=kn}^{kn+k-1} \frac{1}{i!}(\nu t)^i e^{-\nu t}, \quad k = 1, 2, \ldots \tag{3.65}$$

The renewal function can also be calculated analytically as follows

$$\mu_N(t) = \frac{\nu t}{k} + \frac{1}{k} \sum_{n=1}^{k-1} \frac{\gamma_n}{1 - \gamma_n} (1 - e^{-\nu t(1-\gamma_n)}) \tag{3.66}$$

where $\gamma_n = e^{i 2\pi \frac{n}{k}}$, $n = 1, 2 \ldots k - 1$.

3.5 Markov chains

A state discrete stochastic process $\{X(t), t \in T\}$, where the index parameter t signifies the time, is called a *state discrete Markov process* if for any $t_1 < t_2 < \cdots < t_n < t_{n+1}$ in T the following relation exists between the conditional probability functions

$$P_{\{X\}}(x_{n+1}, t_{n+1} | x_n, t_n; \ldots ; x_2, t_2; x_1, t_1) =$$

$$P_{\{X\}}(x_{n+1}, t_{n+1} | x_n, t_n), \quad t_n > t_{n-1} > \cdots > t_2 > t_1 \tag{3.67}$$

The finite or countable infinitely many states of the Markov process have been designated $x_1, x_2 \ldots x_i, \ldots$.

Assume, that up to and including the present time t_n the following states of the process have been observed $X(t_1) = x_1, X(t_2) = x_2 \ldots X(t_n) = x_n$, and consider a set of realizations passing through these observations. The left-hand side of (3.67) then specifies the relative number of these realizations, which realizes the value x_{n+1} at the next future time t_{n+1}. Next, consider another and larger set of realizations, which only passes through the most recent observation $X(t_n) = x_n$ but may attain arbitrarily values at the preceeding instants of observation $t_1, t_2 \ldots t_{n-1}$. The right-hand side of (3.67) then specifies the relative number of these realizations, which realizes the value x_{n+1} at the future time t_{n+1}. If $\{X(t), t \in T\}$ is a Markov process the two constructed conditional probabilities will be identical for all states x_{n+1} at any future time t_{n+1}. Hence, the probability of the Markov process $X(t)$ to take the value x_{n+1} at the time t_{n+1} under condition that its values have been observed at some preceeding instants of time, only depends on the knowledge of its most recent observed value $X(t_n) = x_n$ at the present time t_n, whereas the past observations are of no importance.

The conditional probability function on the right-hand side of (3.67) is known as the *transition probability function* of the Markov process. A state discrete Markov process is completely described by its 1st order probability function $P_{\{X\}}(x, t)$ and its transition probability function. For example the 2nd and 3rd order probability functions become

$$P_{\{X\}}(x_2, t_2; x_1, t_1) = P_{\{X\}}(x_2, t_2 | x_1, t_1) P_{\{X\}}(x_1, t_1), \quad t_2 > t_1 \tag{3.68}$$

$$P_{\{X\}}(x_3, t_3; x_2, t_2; x_1, t_1) = P_{\{X\}}(x_3, t_3 | x_2, t_2; x_1, t_1) P_{\{X\}}(x_2, t_2; x_1, t_1) =$$

$$P_{\{X\}}(x_3, t_3 | x_2, t_2) P_{\{X\}}(x_2, t_2 | x_1, t_1) P_{\{X\}}(x_1, t_1), \quad t_3 > t_2 > t_1 \tag{3.69}$$

Next, a compatibility equation is derived, which a Markov process necessarily must fulfil, if the relation (3.67) holds for arbitrarily index values $t_1 < t_2 < \cdots < t_{n-1} < t_n$. For any state discrete stochastic process (not necessarily Markovian) follows for arbitrary $t_i < t_k < t_j$

$$P_{\{X\}}(x_j, t_j; x_i, t_i) = \sum_{x_k} P_{\{X\}}(x_j, t_j; x_k, t_k; x_i, t_i) =$$

$$\sum_{x_k} P_{\{X\}}(x_j, t_j | x_k, t_k; x_i, t_i) P_{\{X\}}(x_k, t_k; x_i, t_i) \quad \Rightarrow$$

$$P_{\{X\}}(x_j, t_j | x_i, t_i) = \frac{P_{\{X\}}(x_j, t_j; x_i, t_i)}{P_{\{X\}}(x_i, t_i)} =$$

$$\sum_{x_k} P_{\{X\}}(x_j, t_j | x_k, t_k; x_i, t_i) P_{\{X\}}(x_k, t_k | x_i, t_i), \quad t_i < t_k < t_j \tag{3.70}$$

where the summation is performed over all intermediate states x_k. If now $\{X(t), t \in T\}$ is a Markov process, (3.70) reduces to

$$P_{\{X\}}(x_j, t_j | x_i, t_i) = \sum_{x_k} P_{\{X\}}(x_j, t_j | x_k, t_k) P_{\{X\}}(x_k, t_k | x_i, t_i), \quad t_i < t_k < t_j \tag{3.71}$$

(3.71), which is the requested compatibility equation, is known as the *Chapman-Kol-mogorov-Smoluchowski equation*. The relation (3.71) is interpreted in figure 3.5. The relation simply sums up the probabilities of all intermediate sample paths x_k from the initial state x_i at time t_i to the final state x_j at the time t_j.
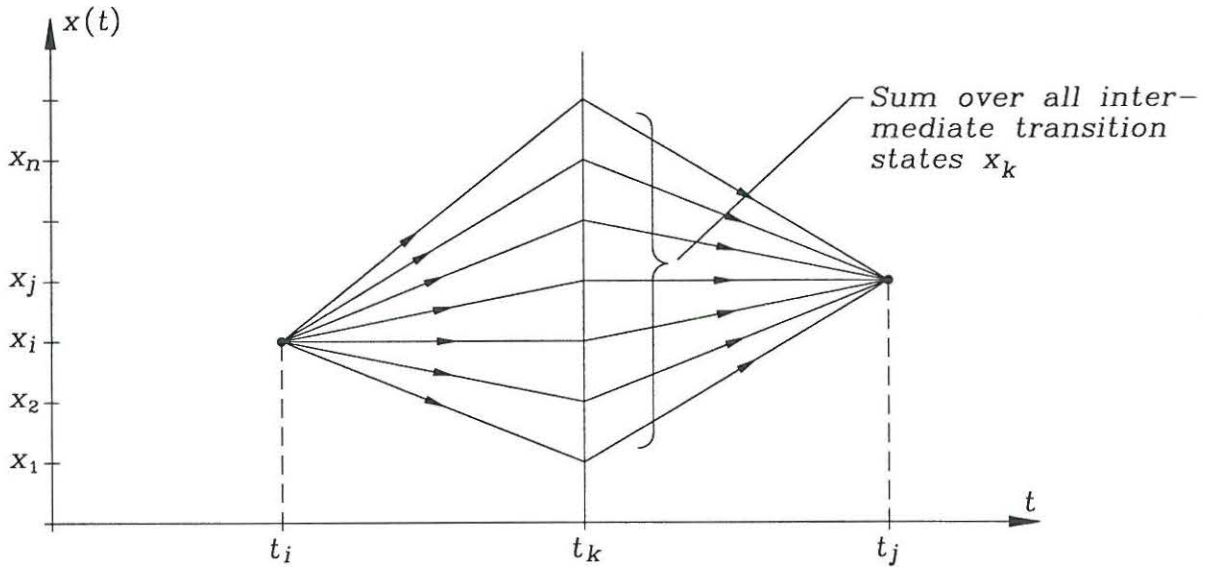

Fig. 3.5: Interpretation of the Chapman-Kolmogorov-Smoluchowski equation.

In the remaining part of this section it is assumed that the index set $T = \{t_1, t_2 \ldots t_n, \ldots\}$ of the Markov process is discrete. Then the index discrete state discrete Markov process $\{X(t), t \in \{t_1, t_2 \ldots t_n, \ldots\}\}$ is called a *Markov chain*. Quite often the instants of observation have equidistant interval Δt, so $t_k = k\Delta t$, $k = 1, 2, \ldots$, although this need not be so in the general case.

At the time t_n one may be interested in the transition probabilities $P_{ji}^{(n)}$ from the state x_i at the time t_n to the state x_j at the next observation time t_{n+1}, i.e.

$$P_{ji}^{(n)} = P_{\{X\}}(x_j, t_{n+1} | x_i, t_n) \tag{3.72}$$

These probabilities can be assembled in the so-called *transition probability matrix* $\mathbf{P}^{(n)}$, so $P_{ji}^{(n)}$ forms the element in the jth row and the ith column.

Obviously, $\sum_{x_j} P_{\{X\}}(x_j, t_{n+1} | x_i, t_n) = 1$, so the columns of $\mathbf{P}^{(n)}$ fulfil

$$\sum_{j=1}^{\infty} P_{ji}^{(n)} = 1 \tag{3.73}$$

The Markov chain is called *stationary* if the transition probability matrix $\mathbf{P}^{(n)}$ does not depend on n. Then the transition probability matrix from states at time t_1 to states

at time t_2, from states at time t_2 to states at time t_3, etc. are all identical, i.e. $\mathbf{P}^{(1)} = \mathbf{P}^{(2)} = \cdots = \mathbf{P}$. Stationary Markov chains are met, when index continuous Markov processes, which are also stationary in the strict sense, are observed at equidistant instants of time $t_k = k\Delta t$, $k = 1, 2, \ldots$. In the rest of this section the Markov chains are assumed to be stationary.

Let $\boldsymbol{\pi}^{(1)}$ be the vector of 1st order probabilities $\pi_i^{(1)} = P_{\{X\}}(x_i, t_1)$, $i = 1, 2, \ldots$ for all states at the initial time t_1. The 1st order probabilities at the next time of observation t_2 then follows from marginalization of (3.68)

$$\pi_j^{(2)} = P_{\{X\}}(x_j, t_2) = \sum_{x_i} P_{\{X\}}(x_j, t_2; x_i, t_1) =$$
$$\sum_{x_i} P_{\{X\}}(x_j, t_2 | x_i, t_1) P_{\{X\}}(x_i, t_1) = \sum_{i=1}^{\infty} P_{ji} \pi_i^{(1)} \tag{3.74}$$

In vector form (3.74) can be written

$$\boldsymbol{\pi}^{(2)} = \mathbf{P}\boldsymbol{\pi}^{(1)} \tag{3.75}$$

Now, $\boldsymbol{\pi}^{(2)}$ forms the initial value for the next transition from states at the time t_2 to the states at the following time of observation t_3. From (3.75) follows

$$\boldsymbol{\pi}^{(3)} = \mathbf{P}\boldsymbol{\pi}^{(2)} = \mathbf{PP}\boldsymbol{\pi}^{(1)} \tag{3.76}$$

Proceeding in this manner the 1st order probabilities at the time t_k after $(k-1)$ transitions is seen to be given as

$$\boldsymbol{\pi}^{(k)} = \mathbf{P}^{k-1}\boldsymbol{\pi}^{(1)} \tag{3.77}$$

where $\mathbf{P}^{k-1} = \mathbf{PP}\cdots\mathbf{P}$ ($\mathbf{P}$ multiplied by itself $k-1$ times).

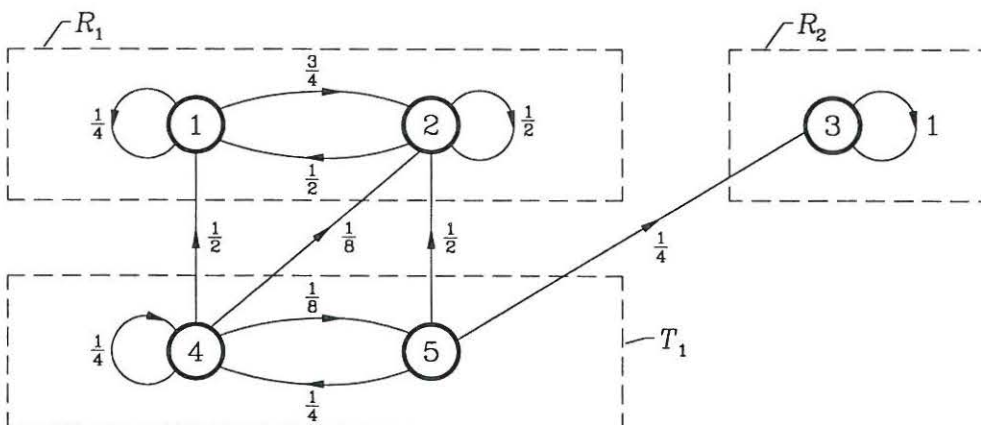**Example 3.2: 5 state Markov chain**

Fig. 3.6: State transition diagram for 5 state Markov chain.

Consider a Markov chain $\{X_1, X_2, \ldots\}$ with the 5 states $x_1, x_2 \ldots x_5$. The transition probability density matrix is given as

$$
\mathbf{P} = \begin{bmatrix}
\frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
\frac{3}{4} & \frac{1}{2} & 0 & \frac{1}{8} & \frac{1}{2} \\
0 & 0 & 1 & 0 & \frac{1}{4} \\
0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} \\
0 & 0 & 0 & \frac{1}{8} & 0
\end{bmatrix}
= \begin{bmatrix}
\mathbf{P}_{R_1 R_1} & 0 & \mathbf{P}_{R_1 T_1} \\
0 & \mathbf{P}_{R_2 R_2} & \mathbf{P}_{R_2 T_1} \\
0 & 0 & \mathbf{P}_{T_1 T_1}
\end{bmatrix}
\tag{3.78}
$$

The structure of the transition probability matrix can be illustrated by a state transition diagram as the one shown in figure 3.6.

In the state transition diagram shown in figure 3.6 hatched line boxes R_1, R_2 and T_1 have been drawn around the states $1+2$, 3 and $4+5$. The main difference between the boxes R_1 and R_2 and the box T_1 is that for the former all paths are orientated into the box, while for the box T_1 all paths are orientated out of the box. States within a box with only inwards orientated paths are called *recurrent states*, whereas states within a box with only outwards orientated paths are called *transient states*.

If the system at some time enters a recurrent state, it will continue to cycle only between the states belonging to the same box in all future. The name 'recurrent' indicates that once attained, such a state will be occupied at least one time more with probability 1 in infinitely many future transitions. On the other hand the probability is zero that the system remains within a box of transient states after infinitely many transitions. The system may move forth and back between the states of the system, as is the case with states 4 and 5 in figure 3.6. However, at each transition there is a leakage of probability mass out of the box, and the probability of staying within the box eventually drops to zero.

Recurrent states are further divided into so-called *positive recurrent states* and *null recurrent states*. The meaning of these concepts can be explained as follows. Assume a certain recurrent state x_j is occupied at a certain time. One may then define the number of transitions, N_j, until the state x_j is reoccupied again for the first time. If x_j is positive recurrent the system will in average return to x_j in finite many transitions, i.e. $E[N_j] < \infty$. If infinitely many transitions in average are needed until x_j is attained again, i.e. $E[N_j] = \infty$, x_j is classified as a null recurrent state.

Example 3.3: Average recurrent length of state 1 in figure 3.6

On condition of state 1 is occupied at a certain instant of time the probability of staying in state 1 is $\frac{1}{4}$, so $P(N_1 = 1) = \frac{1}{4}$. State 1 can be reoccupied for the first time after 2 transitions, if the system goes to state 2 in the first transition and then immediately back to state 1 in the second transition. The probability for this event is $P(N_1 = 2) = \frac{3}{4} \cdot \frac{1}{2}$. Generally, state 1 is reoccupied for the first time after $n > 2$ transitions, if the system goes from state 1 to state 2 in the first transition, stays in state 2 in the succeeding $n - 2$ transitions and, eventually, returns to state 1 in the nth transition. The probability of this event is $P(N_1 = n) = \frac{3}{4} \cdot \left(\frac{1}{2}\right)^{n-2} \cdot \frac{1}{2} = 3\left(\frac{1}{2}\right)^{n+1}$, $n > 2$. Hence, the probability function of

N_1 becomes

$$P_{N_1}(n) = P(N_1 = n) = \begin{cases} \frac{1}{4} & , n = 1 \\ 3\left(\frac{1}{2}\right)^{n+1} & , n \geq 2 \end{cases} \tag{3.79}$$

One can then calculate the expectation

$$E[N_1] = \sum_{n=1}^{\infty} nP(N_1 = n) = 1\frac{1}{4} + \sum_{n=2}^{\infty} n3\left(\frac{1}{2}\right)^{n+1} =$$

$$2.50 + 3\sum_{n=1}^{\infty} \frac{n}{2^n} - 3\sum_{n=0}^{\infty} \frac{1}{2^n} = 2.50 \tag{3.80}$$

where the results $\sum_{n=1}^{\infty} \frac{n}{2^n} = 2.0$, $\sum_{n=0}^{\infty} \frac{1}{2^n} = 2,0$ have been used. Since, $E[N_1] = 2.50 < \infty$, state 1 is positive recurrent.

It can be shown that all states of a Markov chain can be classified into sets of recurrent states $R_1, R_2, \ldots$ and sets of transient states $T_1, T_2, \ldots$. Using a suitable numbering, the transition probability matrix can then be written

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{R_1 R_1} & \mathbf{0} & \cdots & \mathbf{P}_{R_1 T_1} & \mathbf{P}_{R_1 T_2} & \cdots \\ \mathbf{0} & \mathbf{P}_{R_2 R_2} & \cdots & \mathbf{P}_{R_2 T_2} & \mathbf{P}_{R_2 T_2} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{P}_{T_1 T_1} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{P}_{T_2 T_2} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{bmatrix} \tag{3.81}$$

where $\mathbf{P}_{R_i R_i}$ is the transition probability matrix of the recurrent states within the class R_i, $\mathbf{P}_{R_1 T_j}$ is the transition probability matrix from transient states within T_j to recurrent states within R_i, and $\mathbf{P}_{T_i T_i}$ is transition probability matrix from transient states within $T_i T_i$. The transition matrix $\mathbf{P}_{T_i T_j}$ of transient states within T_j to transient states within $\mathbf{T}_i$ is $\mathbf{0}$, since non transitions into transients states are possible per definition. The division as given by (3.81) has also been emphasized in the last statement of (3.78) for the example shown in example 3.1.

A state x_j for which $P_{jj} = 1$ is called an *absorbing state*. State 3 in figure 3.6 is absorbing. If such a state is attained at some instant of time, the system will remain there in all future. Obvious, an absorbing state is also a recurrent state.

A state x_j for which $P_{jj} = 0$ is called a *reflecting state*. State 5 in figure 3.6 is reflecting. If the system arrives at such a state at a certain instant of time, it will leave the state again with probability 1 during the following transition. A reflecting state may be as well recurrent as transient.

A state x_j is *accessible* from the state x_i if at least one path can be found on the state transition diagram leading from the state x_i to the state x_j in the direction of the

arrows. As an example, state 3 in figure 3.6 is accessible from state 5 and from state 4 via state 5. On the other hand neither state 3, 4, or 5 are accessible from the states 1 and 2.

A Markov chain is called *irreducible* if all states are accessible from all other states. Obviously such a Markov chain cannot have any transient states, and only one class R_1 of recurrent states.

Finally, a Markov chain is said to be *periodic* with the period $N > 1$, if after N transitions the system returns to its initial state with the probability 1. Periodic Markov chains are only possible if all states are reflecting.

Now, the prerequisites have been made to consider the problem of determining the probability distribution, $\pi^{(\infty)}$, obtained from (3.77) after infinitely many transitions as $k \to \infty$. If such a limit exists, the Markov chain is said to have a *limiting distribution* or a *stationary distribution*. If it exists, the limiting distribution must be invariant to further transitions, i.e.

$$\pi^{(\infty)} = \mathbf{P}\pi^{(\infty)} \tag{3.82}$$

(3.71) determines the limiting distribution as the eigenvector to $\mathbf{P}$, normalized so $\sum_{j=1}^{\infty} \pi_j^{(\infty)} = 1$, and with the associated eigenvalue $\lambda = 1$. If the eigenvalue $\lambda = 1$ is simple, it means that a unique solution $\pi^{(\infty)}$ is obtained to (3.82). Mathematically, it can be proved that such a unique solution is obtained only if the Markov chain is irreducible, aperiodic and all states are positive recurrent. Since, the distribution of a periodic Markov chain repeats itself with a period $N > 1$, and not with the period $N = 1$ required for the stationary distribution, the restriction to aperiodic Markov chains is obvious. Since the initial distribution is not entering (3.82), it follows that the limiting distribution obtained after infinitely many transitions is the same in this case, no matter which initial distribution is applied. If the Markov chain is not irreducible, corresponding to multiple classes $R_1, R_2, \ldots$ of positive recurrent states are present, multiple solutions to (3.82) are obtained, one solution belonging to each recurrent class. A reduced relation similar to (3.82) is then obtained for the state variables in each recurrent class, R_i, i.e.

$$\pi_{R_i}^{(\infty)} = \mathbf{P}_{R_i R_i} \pi_{R_i}^{(\infty)} \tag{3.83}$$

The state variables not belonging to the considered recurrent class are all zero. Any linear combination of the stationary distribution for each recurrent class, suitably normalized so the sum of the probabilities is 1, is also a stationary distribution. Which of these infinite many stationary distribution is eventually obtained after infinite many transitions, now depends on the initial distribution.

Since the eigenvalue $\lambda = 1$ is known, the solution of the n-dimensional eigenvalue problems (3.82) or (3.83) can easily be obtained as follows. One arbitrary state variable is selected, and (3.82) and (3.83) are then rearranged into a system of $n - 1$ linear equations, which is next solved in terms of the selected state variable. The actual magnitude of the components then follows from the normalization condition $\sum_{j=1}^{\infty} \pi_j^{(\infty)} = 1$.
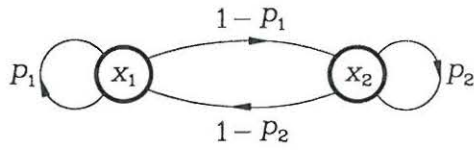
Example 3.4: Two state Markov chain

Figure 3.7: State transition diagram of two state Markov chain.

Assume, that a Markov chain can only attain the two states x_1 and x_2 with the following transition probability matrix

$$\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} p_1 & 1 - p_2 \\ 1 - p_1 & p_2 \end{bmatrix}, \quad p_1, p_2 \in [0, 1] \tag{3.84}$$

$P_{11} = p_1$ specifies the probability of staying in the state x_1, and $P_{21} = 1 - p_1$ is the probability of making a transition to the state x_2 on condition of starting in the state x_1. Similarly, p_2 specifies the probability of staying in the state x_2, and $1 - p_2$ is the probability of making a transition to x_1 on condition of starting in the state x_2. The structure of the transition probability matrix can be illustrated by the state transition diagram shown in figure 3.7. As seen the Markov chain is irreducible. If $p_1 \in]0, 1[$ or $p_2 \in]0, 1[$ it is also aperiodic and positive recurrent. Hence, a unique stationary distribution exists in this case.

By induction it can be proved that

$$\mathbf{P}^k = \frac{1}{2 - p_1 - p_2} \begin{bmatrix} 1 - p_2 & 1 - p_2 \\ 1 - p_1 & 1 - p_1 \end{bmatrix} +$$

$$\frac{(p_1 + p_2 - 1)^k}{2 - p_1 - p_2} \begin{bmatrix} 1 - p_1 & -(1 - p_2) \\ -(1 - p_1) & 1 - p_2 \end{bmatrix} \tag{3.85}$$

The steps of the induction proof is that (3.85) is first verified to hold for $k = 1$. Next, $\mathbf{P}^k$ is assumed to be given by (3.85), and finally $\mathbf{P}^{k+1} = \mathbf{P}^k \mathbf{P}$ is calculated and shown to have the same form.

Assume, $p_1 \in]0, 1[$ or $p_2 \in]0, 1[$. The limit cases $p_1 = p_2 = 0$ and $p_1 = p_2 = 1$ are then excluded and need special consideration. Then $p_1 + p_2 - 1 \in]-1, 1[\Rightarrow (p_1 + p_2 - 1)^k \to 0$ as $k \to \infty$. Hence, (3.85) converges to the limit

$$\mathbf{P}^\infty = \frac{1}{2 - p_1 - p_2} \begin{bmatrix} 1 - p_2 & 1 - p_2 \\ 1 - p_1 & 1 - p_1 \end{bmatrix} \tag{3.86}$$

Let the initial distribution $\boldsymbol{\pi}^{(1)}$ be given as

$$\boldsymbol{\pi}^{(1)} = \begin{bmatrix} \alpha \\ 1 - \alpha \end{bmatrix}, \quad \alpha \in [0, 1] \tag{3.87}$$

The stationary distribution then becomes

$$\boldsymbol{\pi}^{(\infty)} = \frac{1}{2 - p_1 - p_2} \begin{bmatrix} 1 - p_2 & 1 - p_2 \\ 1 - p_1 & 1 - p_1 \end{bmatrix} \begin{bmatrix} \alpha \\ 1 - \alpha \end{bmatrix} = \frac{1}{2 - p_1 - p_2} \begin{bmatrix} 1 - p_2 \\ 1 - p_1 \end{bmatrix} \tag{3.88}$$

As expected, the stationary distribution is independent of the initial distribution as specified by α in this case, i.e. the stationary distribution is unique.

If $p_1 = p_2 = 0$ both states are reflecting. The system switches forth and back between the two states. Obviously, this corresponds to a periodic system with the period $N = 2$, and no stationary distribution is present.

If $p_1 = p_2 = 1$ both states are absorbing. The system is no longer irreducible, but breaks down into 2 recurrent classes, one for each state and the following limiting distributions are obtained, cf. (3.83).

$$\boldsymbol{\pi}^{\infty} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{\pi}^{\infty} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{3.89}$$

Which stationary state is attained depends on which state the system starts from, i.e. the stationary distribution now depends on the initial state.

Next, let us solve the same problem based on (3.82). Without restriction let us solve the equations for π_1^{∞} in terms of π_2^{∞}

$$\begin{bmatrix} \pi_1^{\infty} \\ \pi_2^{\infty} \end{bmatrix} = \begin{bmatrix} p_1 & 1 - p_2 \\ 1 - p_1 & p_2 \end{bmatrix} \begin{bmatrix} \pi_1^{\infty} \\ \pi_2^{\infty} \end{bmatrix} \quad \Rightarrow \quad \pi_1^{\infty} = \frac{1 - p_2}{1 - p_1} \pi_2^{\infty} \tag{3.90}$$

(3.90) is obtained for both equations of (3.90), reflecting that $\lambda = 1$ is an eigenvalue. Next, the normalization condition $\pi_1^{(\infty)} + \pi_2^{(\infty)} = 1$ is applied, so

$$\pi_2^{\infty} k \left(\frac{1 - p_2}{1 - p_1} \right) + \pi_2^{\infty} = 1 \quad \Rightarrow \quad \pi_2^{\infty} = \frac{1 - p_1}{2 - p_1 - p_2} \quad , \quad \pi_1^{\infty} = \frac{1 - p_2}{2 - p_1 - p_2} \tag{3.91}$$

which is identical to (3.88).

Example 3.5: Stationary distributions of 5 state Markov chain

The 5 state Markov chain of example 3.2 is considered again. Since there are 2 classes of recurrent states R_1 and R_2, two stationary distributions are present.

For the recurrent class R_1 the stationary states are determined by (3.81) and (3.83)

$$\begin{bmatrix} \pi_1^{(\infty)} \\ \pi_2^{(\infty)} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \pi_1^{(\infty)} \\ \pi_2^{(\infty)} \end{bmatrix} = \tag{3.92}$$

Solving the equations for $\pi_2^{(\infty)}$ in terms of $\pi_1^{(\infty)}$ the solution $\pi_2^{(\infty)} = \frac{3}{2} \pi_1^{(\infty)}$ is obtained. Next, the normalization condition $\pi_1^{(\infty)} + \pi_2^{(\infty)} = 1$ provides

$$\pi_1^{(\infty)} + \frac{3}{2} \pi_1^{(\infty)} = 1 \quad \Rightarrow$$
$$\pi_1^{(\infty)} = \frac{2}{5} \quad , \quad \pi_2^{(\infty)} = \frac{3}{5} \tag{3.93}$$

Hence, the stationary distribution of the Markov chain for the recurrent class R_1 becomes

$$\boldsymbol{\pi}_1^{(\infty)} = \begin{bmatrix} \pi_1^{(\infty)} \\ \pi_2^{(\infty)} \\ \pi_3^{(\infty)} \\ \pi_4^{(\infty)} \\ \pi_5^{(\infty)} \end{bmatrix} = \begin{bmatrix} \frac{2}{5} \\ \frac{3}{5} \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{3.94}$$

For the recurrent class R_2 the stationary distribution is $\pi_3^{(\infty)} = 1$. Then, the stationary distribution of the Markov chain becomes

$$\boldsymbol{\pi}_2^{(\infty)} = \begin{bmatrix} \pi_1^{(\infty)} \\ \pi_2^{(\infty)} \\ \pi_3^{(\infty)} \\ \pi_4^{(\infty)} \\ \pi_5^{(\infty)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{3.95}$$

Any linear combination $\alpha\boldsymbol{\pi}_1^{(\infty)} + (1 - \alpha)\boldsymbol{\pi}_2^{(\infty)}$, $\alpha \in]0, 1[$ is again a stationary distribution.

4. ERGODIC PROCESSES

If a physical phenomenon is to be modelled by a stochastic process $\{X(t), t \in [0, \infty[\}$, the statistical properties of the process such as the mean value function and auto-correlation function must be determined from available measurements. Assume N number of independent time series of $x_1(t), x_2(t) \ldots x_N(t)$ are available. The mean value function and the auto-correlation function can then be estimated from the averages

$$\mu_X(t) \simeq \frac{1}{N} \sum_{n=1}^{N} x_n(t) \tag{4.1}$$

$$\mu_{XX}(t_1, t_2) \simeq \frac{1}{N} \sum_{n=1}^{N} x_n(t_1) x_n(t_2) \tag{4.2}$$

where the estimate (4.1) is applied on the right-hand side of (4.2). The accuracy of the estimates depends on the N number of available time series.
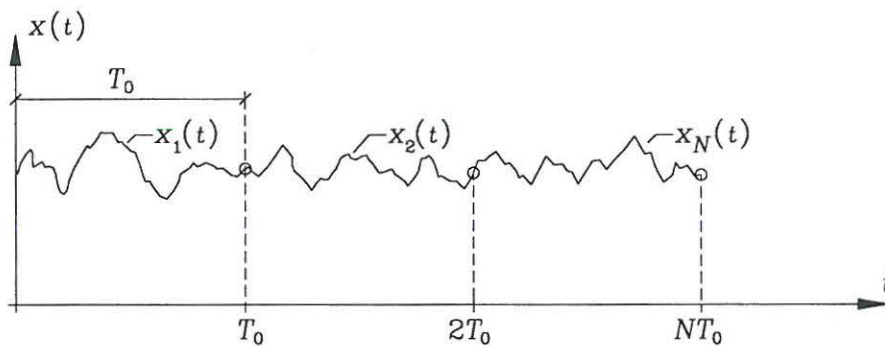

Figure 4.1: Realization of ergodic process.

Quite often only a single time series is available. However, if $\{X(t), t \in [0, \infty[\}$ can be assumed to be homogeneous from the physical nature of the problem, the probabilistic structure will be invariant under a shift T_0 in the index set. If further T_0 is sufficiently large, so that $X(t), X(t+T_0) \ldots X(t+(N-1)T_0)$ are all mutually independent stochastic variables, one may divide the available time series $x(t)$ into N sub-time series $x_1(t) = x(t), x_2(t) = x(t + T_0) \ldots x_N(t) = x(t + (N - 1)T_0)$ as shown in figure 4.1. These cannot be considered as independent realizations of a stochastic process $\{X(t), t \in [0, \infty[\}$. Actually, $x_1(T_0)$ and $x_2(0)$ are identical. Nevertheless, the stochastic variables producing the samples entering (4.1), (4.2) can be considered identically distributed and mutually independent stochastic variables. This heuristic approach suggests that the statistical properties of the process may be estimated based on such sub-time series.

Next, the interval $[0, T_0]$ is further sub-divided into M equidistant intervals of length $\Delta t = \frac{T_0}{M}$. Due to the mandatory stationarity assumptions of the process, (4.1) should provide the same value μ_X at all times $t = 0, \Delta t \ldots (M-1)\Delta t$. Applying the average of these local estimates as the final estimate for μ_X, we have

$$\mu_X \simeq \frac{1}{M} \sum_{m=1}^{M} \frac{1}{N} \sum_{n=1}^{N} x_n((m-1)\Delta t) =$$

$$\frac{1}{MN\Delta t} \sum_{m=1}^{M} \sum_{n=1}^{N} x_n((m-1)\Delta t)\Delta t \xrightarrow{\Delta t \to 0} \frac{1}{T} \int_0^T x(t)dt \tag{4.3}$$

$$T = MN\Delta t = NT_0 \tag{4.4}$$

where $x(t)$ is the original available time series. Using a similar averaging technique for local estimates of $\mu_{XX}(t, t+\tau)$ as given by (4.1) at times $t = 0, \Delta t \ldots (M-1)\Delta t$, the auto-correlation function may be estimated from

$$\mu_{XX}(\tau) \simeq \frac{1}{T-\tau} \int_0^{T-\tau} x(t)x(t+\tau)dt, \quad \tau > 0 \tag{4.5}$$

By (4.3) and (4.4) the problem of estimating the mean value function and the auto-covariance function has been reduced to the evaluation of certain time averages of a time series $x(t)$. The length T of the time series must be sufficiently large, so that an extra increase of T does not alter the estimates significantly. T_0 should be at least the correlation length of the process in order to ensure the mutual independence of the samples $x_n(t)$. N should be larger than 50 to ensure the stabilization of the averages (4.1) and (4.2). Hence, T must be at least of magnitude 50 times the correlation length of the process. This of course demands longer time series for narrow-banded than for broad-banded processes.

Even if the right-hand sides of (4.1) and (4.2) have stabilized for a given sample interval T, the results are only at any value, if limits are independent of the specific realization of $\{X(t), t \in [0, \infty[\}$ to be applied. This is not necessarily the case for an arbitrary homogeneous process. However, if the same limit values are obtained for almost all realizations $x(t)$ applied on the the right-hand side of (4.3) and (4.5) the process $\{X(t), t \in [0, \infty[\}$ is said to be *ergodic in the mean value with probability 1* and *ergodic in the auto-correlation function with probability 1*, respectively. Ergodicity of higher order moments with probability 1 can be defined in the same way. If the process is ergodic in the moments of arbitrary order it is said to be *strictly ergodic*.

Ergodicity with probability 1 is closely related to the definition of the corresponding stochastic integrals with probability 1. Similarly to the definition of stochastic integrals

in mean square ergodicity in the mean square can also be defined. In this respect consider the following stochastic integrals in mean square

$$M_{X,T} = \frac{1}{T} \int_0^T X(t)dt \tag{4.6}$$

$$M_{XX,T}(\tau) = \frac{1}{T-\tau} \int_0^{T-\tau} X(t)X(t+\tau)dt, \quad \tau > 0 \tag{4.7}$$

The stochastic process $\{X(t), t \in [0, \infty[\}$ is then said to be *ergodic in the mean value in mean square* if lim-ms $M_{X,T} = \mu_X$ as $T \to \infty$. This means, cf. (1.121)

$$\lim_{T\to\infty} E\left[(M_{X,T} - \mu_X)^2\right] = 0 \quad \Rightarrow$$

$$\lim_{T\to\infty} \left(E\left[\frac{1}{T^2} \int_0^T \int_0^T X(t_1)X(t_2)dt_1 dt_2\right] - 2\mu_X E\left[\frac{1}{T} \int_0^T X(t)dt\right] + \mu_X^2 \right) =$$

$$\lim_{T\to\infty} \frac{1}{T^2} \int_0^T \int_0^T E\left[X(t_1)X(t_2)\right] dt_1 dt_2 -$$

$$2\mu_X \lim_{T\to\infty} \frac{1}{T} \int_0^T E\left[X(t)\right] dt + \mu_X^2 = 0 \tag{4.8}$$

where (2.81) has been used for the evaluation of the expectations of the stochastic integrals. Since the process is homogeneous, $E[X(t)] = \mu_X$ and $E[X(t_1)X(t_2)] = \mu_{XX}(t_2 - t_1) = \kappa_{XX}(t_2 - t_1) + \mu_X^2$, see (2.12). lim-ms $M_{X,T} = \mu_X$ as $T \to \infty$ is then obtained if

$$\lim_{T\to\infty} \frac{1}{T^2} \int_0^T \int_0^T \left(\kappa_{XX}(t_2 - t_1) + \mu_X^2\right) dt_1 dt_2 -$$

$$2\mu_X \lim_{T\to\infty} \frac{1}{T} \int_0^T \mu_X dt + \mu_X^2 = 0 \quad \Rightarrow$$

$$\lim_{T\to\infty} \frac{1}{T^2} \int_0^T \int_0^T \kappa_{XX}(t_2 - t_1)dt_1 dt_2 = 0 \tag{4.9}$$
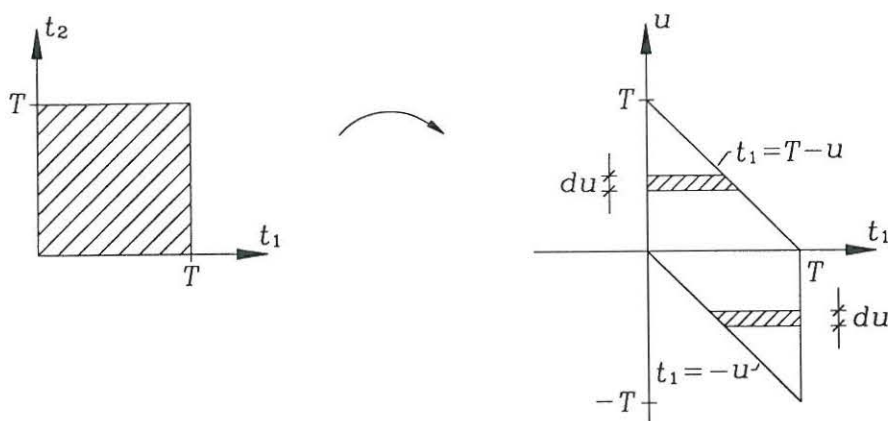
Figure 4.2: Mapping of integration areas.

In order to evaluate the planar integral in (4.9) further the change of integration variable $u = t_2 - t_1$ is performed in the outermost integral. Then, see figure 4.2

$$\int\limits_0^T \int\limits_0^T \kappa_{XX}(t_2 - t_1)dt_1 dt_2 =$$

$$\int\limits_{-T}^0 \left(\int\limits_{-u}^T \kappa_{XX}(u)dt_1 \right) du + \int\limits_0^T \left(\int\limits_0^{T-u} \kappa_{XX}(u)dt_1 \right) du =$$

$$\int\limits_{-T}^0 \kappa_{XX}(u)(T + u)du + \int\limits_0^T \kappa_{XX}(u)(T - u)du =$$

$$2 \int\limits_0^T \kappa_{XX}(u)(T + |u|)du = 2T \int\limits_0^T \kappa_{XX}(u) \left(1 + \frac{|u|}{T} \right) du \qquad (4.10)$$

where first it has been applied that $\kappa_{XX}(u)$ is a constant as a function of t_1, and hence can be set outside the innermost integration with respect to t_1. Secondly, the symmetry property (2.96) has been used. From (4.9) and (4.10) then follows that the process $\{X(t), t \in [0, \infty[\}$ is ergodic in the mean square if and only if its auto-covariance function fulfils

$$\lim_{T \to \infty} \frac{1}{T} \int\limits_0^T \kappa_{XX}(u) \left(1 + \frac{u}{T} \right) du = 0 \qquad (4.11)$$

For a certain $\tau > 0$ consider the stochastic variable $Y_\tau(t) = X(t)X(t + \tau)$. (4.7) can then be written

$$M_{XX,T}(\tau) = \frac{1}{T-\tau} \int_0^{T-\tau} Y_\tau(t)dt \qquad (4.12)$$

From the similarity between (4.6) and (4.12) it follows that the stochastic integral (4.12) converges in mean square to the limit $\mu_{XX}(\tau) = E[Y_\tau(t)] = E[X(t)X(t+\tau)]$, and the process hence is ergodic in the auto-correlation function in mean square, if and only if

$$\lim_{T \to \infty} \frac{1}{T-\tau} \int_0^{T-\tau} \kappa_{Y_\tau Y_\tau}(u) \left(1 + \frac{u}{T-\tau}\right) du = 0 \qquad (4.13)$$

where $\kappa_{Y_\tau Y_\tau}(u)$ is the auto-covariance function of the process $\{Y_\tau(t), t \in [0, \infty[\}$ defined from

$$\kappa_{Y_\tau Y_\tau}(u) = E\left[(Y_\tau(t) - \mu_{XX}(\tau))(Y_\tau(t+u) - \mu_{XX}(\tau))\right] =$$

$$E\left[(X(t)X(t+\tau) - \mu_{XX}(\tau))(X(t+u)X(t+u+\tau) - \mu_{XX}(\tau))\right] \qquad (4.14)$$

Now, it can be stated that the random variables $M_{X,T}$ and $M_{XX,T}(\tau)$ provide unbiased and consistent estimates of μ_X and $\mu_{XX}(\tau)$, when a certain realization $x(t)$ is inserted on the right-hand side of (4.6) and (4.7). That the estimates are unbiased for any T follows from taking the expectation on both sides of (4.6) and (4.7) and succeeding use of (2.81)

$$E[M_{X,T}] = E\left[\frac{1}{T} \int_0^T X(t)dt\right] = \frac{1}{T} \int_0^T E[X(t)]dt = \frac{1}{T} \int_0^T \mu_X dt = \mu_X \qquad (4.15)$$

$$E[M_{XX,T}(\tau)] = \frac{1}{T-\tau} \int_0^{T-\tau} E[X(t)X(t+\tau)]dt =$$

$$\frac{1}{T-\tau} \int_0^{T-\tau} \mu_{XX}(\tau)dt = \mu_{XX}(\tau) \qquad (4.16)$$

That the estimates are consistent, i.e. that the variances $E\left[(M_{X,T} - \mu_X)^2\right]$ and $E\left[(M_{XX,T}(\tau) - \mu_{XX}(\tau))^2\right]$ goes to zero as $T \to \infty$, is simply another statement for the ergodic in mean square statements.

It can be proved that ergodicity in statistical moments of a certain order implies ergodicity of all the moments of lower order. Hence, ergodicity in the covariance implies ergodicity in the mean value.

Without proof it is finally stated that the consistency of the estimates (2.122) and (2.123) to the limits σ_X^2 and $S_{XX}(\omega_n)$ is guaranteed if only the process $\{X(t), t \in [0, \infty[\}$ is ergodic in the auto-covariance function in mean square. The estimate (2.122) is unbiased for any T, whereas (2.123) usually is biased.
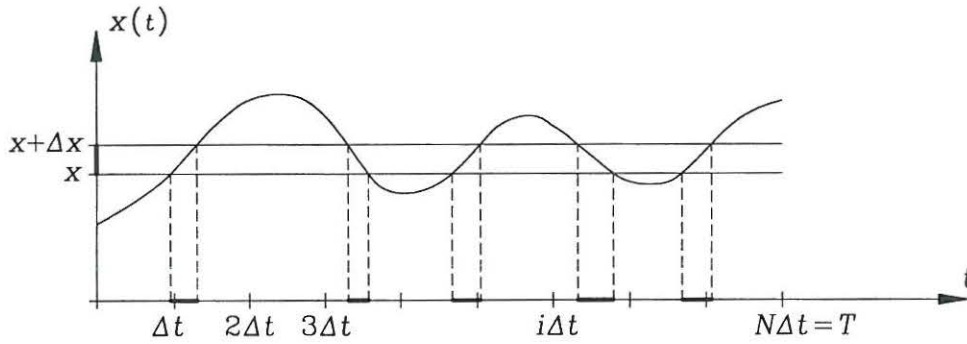

Figure 4.3: Estimation of probability density function by ergodic sampling.

Not only the statistical moments, but also the joint pdfs of the process can be estimated by ergodic sampling from an available time series. The procedure for doing this has been illustrated in figure 4.3. The sample interval is divided into N equidistant intervals of the length $\Delta t = \frac{T}{N}$. The interval $]i\Delta t, (i+1)\Delta t]$, $i = 0 \ldots N-1$, is associated with the sample $x_i = x(i\Delta t)$. The sample interval of $X(t)$ is in the same way divided into disjoint intervals. The probability of samples in the interval $]x, x+\Delta x]$ is approximately equal to $f_{\{X\}}(x)\Delta x$. This probability can then be estimated by the fraction $\frac{\Delta N(x)}{N}$, where $\Delta N(x)$ is the number of samples x_i for which $x_i \in]x, x+\Delta x]$. One then has

$$f_{\{X\}}(x) \simeq \frac{1}{\Delta x} \frac{\Delta N(x)}{N} \tag{4.17}$$

(4.17) is supposed to give the correct answer as $N \to \infty$ and $\Delta x \to 0$. Now,

$$f_{\{X\}}(x) \simeq \frac{1}{\Delta x} \frac{\Delta N(x)\Delta t}{N\Delta t} \xrightarrow{\Delta t \to 0} \frac{1}{\Delta t} \frac{1}{T} \int_0^T 1_{]x,x+\Delta x]}(x(t))dt \tag{4.18}$$

$$1_A(y) = \begin{cases} 1 & , y \in A \\ 0 & , y \notin A \end{cases} \tag{4.19}$$

Hence, $f_{\{X\}}(x)\Delta x$ is estimated as the time-average of $\frac{1}{\Delta x} 1_{]x,x+\Delta x]}(x(t))$, where the indicator function for the interval $A =]x, x + \Delta x]$ is defined by (4.19). In figure 4.3 $1_{]x,x+\Delta x]}(x(t)) = 1$ for the parts of the abscissa axis marked with a bold line. As the

probability density function is tantamount to the knowledge of all statistical moments of the considered process, it is only expected to be estimated correctly using the indicated procedure if the process is strictly ergodic.

The probability distribution function specifies the relative number of samples in $]-\infty, x]$. Hence, it can be estimated from the following time average

$$F_{\{X\}}(x) \simeq \frac{1}{T} \int_0^T 1_{]-\infty, x]}(x(t)) \, dt \qquad (4.20)$$

Example 4.1: Probability density function of harmonic process by ergodic sampling
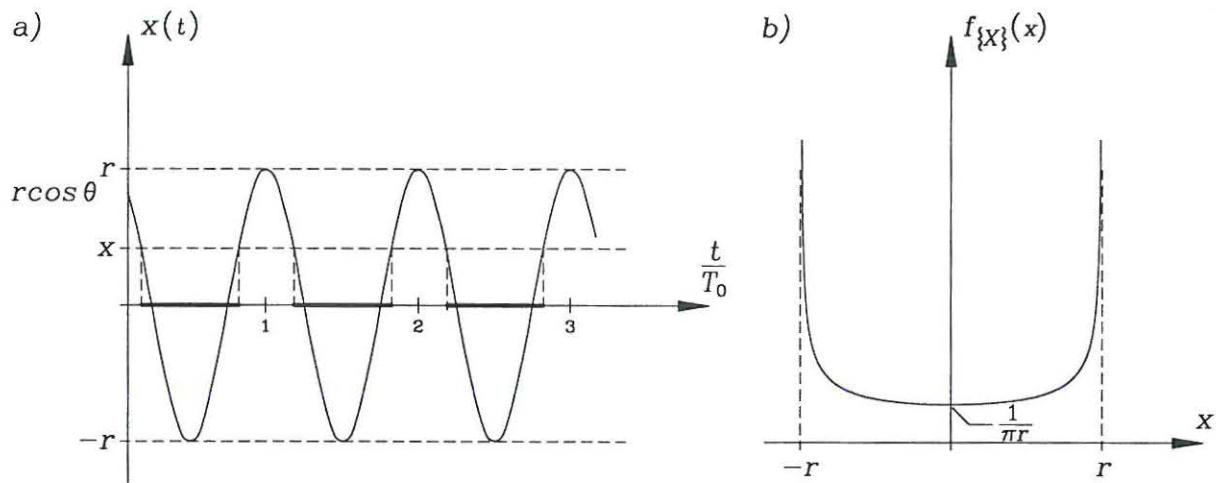

Figure 4.4: a) Realization of harmonic process. b) Probability density function obtained by ergodic sampling.

Consider the harmonic process defined by (2.29). The first order pdf is given by

$$f_{\{X\}}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \qquad (4.21)$$

Here the mean value function is taken as $\mu_X = 0$. Let r and θ denote the amplitude and the phase of the realization, respectively. The corresponding realization $x(t) = r\cos(\omega_0 t + \theta)$ has been shown in figure 4.4a. Using this time series in (4.3) and (4.5), the following results are obtained as $T \to \infty$

$$\mu_X = 0 \qquad (4.22)$$

$$\kappa_{XX}(\tau) = \frac{r^2}{2} \cos(\omega_0 \tau) \qquad (4.23)$$

(4.22) is the correct mean value function of the harmonic process, cf. (2.30), which is consequently ergodic in the mean value. In general the sample variance $\frac{r^2}{2}$ is different from σ^2. Hence, (4.23) is

different from (2.31), and the harmonic process is not ergodic in the auto-covariance function. Because of the periodicity, the distribution function (4.20) determined by the ergodic sampling approach as $T \to \infty$, is equal to the length of the bold-line marked intervals on the abscissa of figure 4.4a in proportion to $T_0 = \frac{2\pi}{\omega_0}$. Then,

$$F_{\{X\}}(x) = \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x}{r}, \quad x \in [-r, r] \tag{4.24}$$

The first order pdf is obtained by differentiation with respect to x

$$f_{\{X\}}(x) = \begin{cases} \frac{1}{\pi\sqrt{r^2 - x^2}} & , \ |x| < r \\ 0 & , \ |x| > r \end{cases} \tag{4.25}$$

(4.25), which has been shown in figure 4.4b, is completely different from (4.21). This is due to the limited ergodicity properties of the harmonic process.

Instead of the harmonic process, the random phase process (2.37) is next considered. For this process strictly ergodicity is present. The pdf obtained by ergodic sampling is still given by (4.25) with $r = a$ and is seen to be exactly identical to (2.38).

LITERATURE

[1]   Soong, T.T.: *Random Differential Equations in Science and Engineering*, Academic Press, New York and London, 1973.

[2]   Lin, Y.K.: *Probabilistic Theory of Structural Dynamics*, McGraw-Hill, New York, 1967.

[3]   Gray, R.M. and Davidson, L.E.: *Random Processes. A Mathematical Approach for Engineers*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1986.

[4]   Cramér, H. and Leadbetter, M.R.: *Stationary and Related Stochastic Processes. Sample Function Properties and Their Applications*, John Wiley and Sons, New York, 1967.

[5]   Vanmarcke, E.: *Random Fields: Analysis and Synthesis*, The MIT Press, Cambridge Ma., 1988.

[6]   Osaki, S.: *Applied Stochastic System Modelling*, Springer-Verlag, Berlin, 1992.