CHAPTER 1

STOCHASTIC VARIABLES AND STOCHASTIC PROCESSES

1.1. Probability Theory

An experiment (or a trial of some process) is performed whose outcome (results) is uncertain: it depends on chance. A collection of all possible elementary (or individual) outcomes is called the **sample space** (or **phase space**, or **range**) and is denoted by Ω. If the experiment is tossing a pair of distinguishable dice, then $\Omega = \{(i, j) \mid 1 \le i, j \le 6\}$. For the case of an experiment with a fluctuating pressure Ω is the set of all real functions $\Omega = (0, \infty)$. An observable event A is a subset of Ω; this is written in the form $A \subset \Omega$. In the dice example we could choose an even, for example, as $A = \{(i, j) \mid i + j = 4\}$. For the case of fluctuating pressures we could use the subset $A = (p_0 > 0, \infty)$.

Not every subset of Ω is observable (or interesting). An example of a non-observable event appears when a pair of dice are tossed and only their spots are counted, $\Omega = \{(i, j), 2 \le i + j \le 12\}$. Then elementary outcomes like (1, 2), (2, 1) or (3, 1), (2, 2), (1, 3) are not distinguished.

Let Γ be the set of observable events for one single experiment. Then Γ must include the certain event of Ω, and the impossible event of $\varnothing$ (the empty set). For every $A \subset \Gamma$, A^c the complement of A, satisfies $A^c \subset \Gamma$ and for every $B \subset \Gamma$ the union and intersection of events, $A \cup B$ and $A \cap B$, must pertain also to Γ. Γ is called an **algebra** of events. In many cases there are countable unions and intersections in Γ. Then it is sufficient to assume that

$$\bigcup_{n=1}^{\infty} A_n \in \Gamma, \quad \text{if } A_n \in \Gamma.$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

An algebra with this property is called a **sigma algebra**. In measure theory, the elements of Γ are called measurable sets and the pair of (Γ, Ω) is called a **measurable space**.

A finite measure $\Pr(\mathrm{A})$ defined on Γ with

$$0 \le \Pr(\mathrm{A}) \le 1, \quad \Pr(\varnothing) = 0, \quad \Pr(\Omega) = 1,$$

is called the probability and the triple $(\Gamma, \Omega, \Pr)$ is referred to as the **probability space**. The set function Pr assigns to every event A the real number $\Pr(\mathrm{A})$. The rules for this set function are along with the formula above

$$\Pr(\mathrm{A}^{\mathrm{c}}) = 1 - \Pr(\mathrm{A});$$
$$\Pr(\mathrm{A}) \le \Pr(\mathrm{B}); \quad \Pr(\mathrm{B} \backslash \mathrm{A}) = \Pr(\mathrm{B}) - \Pr(\mathrm{A}) \quad \text{for } \mathrm{A} \subset \mathrm{B} \in \Gamma.$$

The probability measure $\Pr(\Gamma)$ on Ω is thus a function $\Pr(\Gamma) \to [0, 1]$ and it is generally derived with Lebesque integrations that are defined on Borel sets.

We introduced this formal concept because it can be used as the most general way to introduce axiomatically the probability theory (see e.g. Chung, [1.1]). We will not follow this procedure but we will introduce heuristically stochastic variables and their probabilities.

Definition 1.1. (Stochastic variables)
A **random** (or **stochastic**) **variable** $\mathrm{X}(\omega), \omega \in \Omega$ is a real valued function defined on the sample space Ω. In the following we omit the parameter ω whenever no confusion is possible. □

Definition 1.2. (Probability of an event)
The **probability** of an event equals the number of elementary outcomes divided by the total number of all elementary outcomes, provided that all cases are equally likely. □

Example
For the case of a discrete sample space with a finite number of elementary outcome we have, $\Omega = \{\omega_1, \ldots, \omega_n\}$ and an event is given by $\mathrm{A} = \{\omega_1, \ldots, \omega_k\}, 1 \le k \le n$. The probability of the event A is then $\Pr(\mathrm{A}) = k/n$. ♣

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

Definition 1.3. (Probability distribution function and probability density)
In the continuous case, the **probability distribution function** (PDF) $\mathrm{F_X}(x)$ of a vectorial stochastic variable $\mathrm{X} = (\mathrm{X}_1, \ldots, \mathrm{X}_n)$ is defined by the monotonically increasing real function

$$\mathrm{F_X}(x_1, \ldots, x_n) = \Pr(\mathrm{X}_1 \le x_1, \ldots, \mathrm{X}_n \le x_n), \tag{1.1}$$

where we used the convention that the variable itself is written in upper case letters, whereas the actual values that this variable assumes are denoted by lower case letters.

The **probability density** $\mathrm{p_X}(x_1, \ldots, x_n)$ (PD) of the random variable is then defined by

$$\mathrm{F_X}(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} \mathrm{p_X}(u_1, \ldots, u_n)\mathrm{d}u_1 \cdots \mathrm{d}u_n \tag{1.2}$$

and this leads to

$$\frac{\partial^n \mathrm{F_X}}{\partial x_1 \cdots \partial x_n} = \mathrm{p_X}(x_1, \ldots, x_n). \tag{1.3}$$

Note that we can express (1.1) and (1.2) alternatively if we put

$$\Pr(x_{11} \le \mathrm{X}_1 \le x_{12}, \ldots, x_{n1} \le \mathrm{X}_n \le x_{n2})$$
$$= \int_{x_{11}}^{x_{12}} \cdots \int_{x_{n1}}^{x_{n2}} \cdots \mathrm{p_X}(x_1, \ldots, x_n)\mathrm{d}x_1 \cdots \mathrm{d}x_n. \tag{1.1a}$$

The conditions to be imposed on the PD are given by the positiveness and the normalization condition

$$\mathrm{p_X}(x_1, \ldots, x_n) \ge 0; \quad \int \cdots \int \mathrm{p_X}(x_1, \ldots, x_n)\mathrm{d}x_1 \cdots \mathrm{d}x_n = 1. \tag{1.4}$$

In the latter equation we used the convention that integrals without explicitly given limits refer to integrals extending from the lower boundary $-\infty$ to the upper boundary ∞. □

In a continuous phase space the PD may contain Dirac delta functions

$$\mathrm{p}(x) = \sum_k \mathrm{q}(k)\delta(x - k) + \hat{\mathrm{p}}(x); \quad \mathrm{q}(k) = \Pr(x = k), \tag{1.5}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

where $\mathrm{q}(k)$ represents the probability that the variable x of the discrete set equals the integer value k. We also dropped the index X in the latter formula. We can interpret it to correspond to a PD of a set of discrete states of probabilities $\mathrm{q}(k)$ that are embedded in a continuous phase space S. The normalization condition (1.4) yields now

$$\sum_k \mathrm{q}_k + \int_s \hat{\mathrm{p}}(x)\mathrm{d}x = 1.$$

Examples (discrete Bernoulli and Poisson distributions)
First we consider the **Bernoulli distribution**

$$\text{(i)}\;\; \mathrm{q}_k^{(\mathrm{B})} = \Pr(x = k) \equiv b(k, n, p) = \binom{n}{k} \mathrm{p}^k (1-\mathrm{p})^{n-k}; \quad k = 0, 1, \ldots$$

and then we introduce the **Poisson distribution**

$$\text{(ii)}\;\; \pi_k(\lambda t) = \Pr(x = k) = \frac{(\lambda t)^k \exp(-\lambda t)}{k!}; \quad k = 0, 1, \ldots.$$

In the appendix of this chapter we will give more details about the Poisson distribution. We derive there the Poisson distribution as limit of Bernoulli distribution

$$\pi_k(\lambda t) = \lim_{n\to\infty} b(k, n, p = \lambda t/n).$$ ♣

In the following we will consider in almost all cases only continuous sets.

1.2. Averages

The sample space and the PD define together completely a stochastic variable. To introduce observable quantities we consider now averages. The **expectation value** (or the **average**, or the **mean value**) of a function $\mathrm{G}(x_1, \ldots, x_n)$ of the stochastic variables $x_1, \ldots, x_n$ is defined by

$$\langle \mathrm{G}(x_1, \ldots, x_n)\rangle = \int \cdots \int \mathrm{G}(x_1, \ldots, x_n)\mathrm{p}_\mathrm{X}(x_1, \ldots, x_n)\mathrm{d}x_1 \cdots \mathrm{d}x_n. \tag{1.6}$$

In the case of a discrete variable we must replace to integral in (1.6) by a summation. We obtain then with the use of (1.5) for $\hat{\mathrm{p}}(x)$

$$\langle \mathrm{G}(x_1, \ldots, x_n)\rangle = \sum_{k_1} \cdots \sum_{k_n} \mathrm{G}(k_1, \ldots, k_n)\mathrm{q}(k_1, \ldots, k_n). \tag{1.7}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

There are two rules for the application of the averages:

(i) a and b are two deterministic constants and $\mathrm{G}(x_1, \dots x_n)$ and $\mathrm{H}(x_1, \dots, x_n)$ are two functions of the random variables $x_1, \dots, x_n$. Then we have

$$\begin{aligned}&\langle a\mathrm{G}(x_1, \dots, x_n) + b\mathrm{H}(x_1, \dots, x_n)\rangle \\ &\quad = a\langle \mathrm{G}(x_1, \dots, x_n)\rangle + b\langle \mathrm{H}(x_1, \dots, x_n)\rangle,\end{aligned} \tag{1.8a}$$

and

(ii)

$$\langle\langle \mathrm{G}(x_1, \dots, x_n)\rangle\rangle = \langle \mathrm{G}(x_1, \dots, x_n)\rangle. \tag{1.8b}$$

Now we consider two scalar random variables x and y, their joint PD is $\mathrm{p}(x, y)$. If we do not have more information (observed values) of y, we introduce the two **marginal PD**'s $\mathrm{p_X}(x)$ and $\mathrm{p_Y}(y)$ of the single variables x and y

$$\mathrm{p_X}(x) = \int_{\mathrm{S}_y} \mathrm{p}(x, y)\mathrm{d}y; \quad \mathrm{p_Y}(y) = \int_{\mathrm{S}_x} \mathrm{p}(x, y)\mathrm{d}x, \tag{1.9a}$$

where we integrate over the phase spaces S_x (S_y) of the variables $x(y)$. The normalization condition (1.4) yields

$$\int_{\mathrm{S}_x} \mathrm{p_X}(x)\mathrm{d}x = \int_{\mathrm{S}_y} \mathrm{p_Y}(y)\mathrm{d}y = 1. \tag{1.9b}$$

Definition 1.4. (Independence of variables)
We consider n random variables $x_1, \dots, x_n$, x_1 to be independent of the other variables $x_2, \dots, x_n$ if

$$\langle x_1 x_2 \cdots x_n\rangle = \langle x_1\rangle\langle x_2 \cdots x_n\rangle. \tag{1.10a}$$

We see easily that a sufficient condition to satisfy (1.10a) is

$$\mathrm{p}(x_1, \dots, x_n) = \mathrm{p}_1(x_1)\mathrm{p}_{n-1}(x_2, \dots, x_n), \tag{1.10b}$$

where $\mathrm{p}_k(\dots)$, $k < n$ denotes the marginal probability distribution of the corresponding variables. □

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

The **moments** of a PD of a scalar variable x are given by

$$\langle x^n \rangle = \int \mathrm{p}(x) x^n \mathrm{d}x; \quad n \in \mathrm{N},$$

where n denotes the order of the moment. The first order moment $\langle x \rangle$ is the average of x and we introduce the **variance** σ^2 by

$$\sigma^2 = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 \geq 0. \tag{1.11}$$

The random variable $x - \langle x \rangle$ is called the **standard deviation**.

The average of the of the Fourier transform of a PD is called the **characteristic function**

$$\begin{aligned} \mathrm{G}(k_1, \ldots, k_n) &= \langle \exp(ik_r x_r) \rangle \\ &= \int \mathrm{p}(x_1, \ldots, x_n) \exp(ik_r x_r) \mathrm{d}x_1 \cdots \mathrm{d}x_n, \end{aligned} \tag{1.12}$$

where we applied a summation convention $k_r x_r = \sum_{j=1}^{n} k_j x_j$. This function has the properties $\mathrm{G}(0, \ldots, 0)1; \mid \mathrm{G}(k_1, \ldots, k_n) \mid \leq 1$.

Example

The **Gaussian** (or **normal**) PD of a scalar variable x is given by

$$\mathrm{p}(x) = (2\pi)^{-1/2} \exp(-x^2/2); \quad -\infty < x < \infty. \tag{1.13a}$$

Hence we obtain (see also EX 1.1)

$$\langle x^{2n} \rangle = \frac{(2n)!}{2^n n!}; \quad \sigma^2 = 1; \quad \langle x^{2n+1} \rangle = 0. \tag{1.13b}$$

A stochastic variable characterized by $\mathrm{N}(m, s)$ is a normal distributed variable with the average m and the variance s. The variable x distributed with the PD (1.13a) is thus called a **normal distributed variable** with N(0, 1).

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

A Taylor expansion of the characteristic function $\mathrm{G}(k)$ of (1.13a) yields with (1.12)

$$\mathrm{G}(k) = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle x^n \rangle. \tag{1.14a}$$

We define the **cumulants** κ_m by

$$\ln \mathrm{G}(k) = \sum_{m=1}^{\infty} \frac{(ik)^m}{m!} \kappa_m. \tag{1.14b}$$

A comparison of equal powers of k gives

$$\begin{aligned} &\kappa_1 = \langle x \rangle; \quad \kappa_2 = \langle x^2 \rangle - \langle x \rangle^2 = \sigma^2; \\ &\kappa_3 = \langle x^3 \rangle - 3\langle x^2 \rangle \langle x \rangle + 2\langle x \rangle^3; \ldots. \end{aligned} \tag{1.14c}$$

♣

Definition 1.5. (Conditional probability)
We assume that A, B $\subset \Gamma$ are two random events of the set of observable events Γ. The **conditional probability** of A given B (or knowing B, or under the hypothesis of B) is defined by

$$\Pr(\mathrm{A} \mid \mathrm{B}) = \Pr(\mathrm{A} \cap \mathrm{B})/\Pr(\mathrm{B}); \quad \Pr(\mathrm{B}) > 0.$$

Thus only events that occur simultaneously in A and B contribute to the conditional probability.

Now we consider n random variables $x_1, \ldots, x_n$ with the joint PD $\mathrm{p}_n(x_1, \ldots, x_n)$. We select a subset of variables $x_1, \ldots, x_s$. and we define a conditional PD of the latter variables, knowing the remaining subset $x_{s+1}, \ldots, x_n$, in the form

$$\begin{aligned} &\mathrm{p}_{s|n-s}(x_1, \ldots, x_s \mid x_{s+1}, \ldots, x_n) \\ &\quad = \mathrm{p}_n(x_1, \ldots, x_n)/\mathrm{p}_{n-s}(x_{s+1}, \ldots, x_n). \end{aligned} \tag{1.15}$$

Equation (1.15) is called **Bayes's rule** and we use the marginal PD

$$\mathrm{p}_{n-s}(x_{s+1}, \ldots, x_n) = \int \mathrm{p}_n(x_1, \ldots, x_n) \mathrm{d}x_1 \cdots \mathrm{d}x_s, \tag{1.16}$$

where the integration is over the phase space of the variables $x_1 \cdots x_s$. Sometimes is useful to write to Bayes's rule (1.15) in the form

$$\mathrm{p}_n(x_1, \ldots, x_n) = \mathrm{p}_{n-s}(x_{s+1}, \ldots, x_n) \mathrm{p}_{s|n-s}(x_1, \ldots, x_s \mid x_{s+1}, \ldots, x_n). \tag{1.15$'$}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

We can also rearrange (1.15′) and we obtain

$$\mathrm{p}_n(x_1,\ldots,x_n) = \mathrm{p}_s(x_1,\ldots,x_s)\mathrm{p}_{n-s|s}(x_{s+1},\ldots,x_n \mid x_1,\ldots,x_s). \tag{1.15''}$$

□

Definition 1.6. (Conditional averages)
The **conditional average** of the random variable x_1, knowing $x_2,\ldots,x_n$, is defined by

$$\begin{aligned}\langle x_1 \mid x_2,\ldots,x_n\rangle &= \int x_1 \mathrm{p}_{1|n-1}(x_1 \mid x_2,\ldots,x_n)\mathrm{d}x_1\\ &= \int x_1 \mathrm{p}_n(x_1 \mid x_2,\ldots,x_n)\mathrm{d}x_1/\mathrm{p}_{n-1}(x_2,\ldots,x_n).\end{aligned} \tag{1.17}$$

Note that (1.17) is a random variable.

The rules for this average are in analogy to (1.8)

$$\langle ax_1 + bx_2 \mid y\rangle = a\langle x_1 \mid y\rangle + b\langle x_2 \mid y\rangle, \quad \langle\langle x \mid y\rangle\rangle = \langle x \mid y\rangle. \tag{1.18}$$

□

Example
We consider a scalar stochastic variable x with its PD $\mathrm{p}(x)$. An event A is given by $x \in [a, b]$. Hence we have

$$\mathrm{p}(x \mid \mathrm{A}) = 0 \ \forall x \notin [a, b],$$

and

$$\mathrm{p}(x \mid \mathrm{A}) = \mathrm{p}(x)\Big/\int_a^b \mathrm{p}(s)\mathrm{d}s; \quad x \in [a, b].$$

The conditional PD is thus given by

$$\langle x \mid \mathrm{A}\rangle = \int_a^b x\mathrm{p}(x)\mathrm{d}x\Big/\int_a^b \mathrm{p}(s)\mathrm{d}s.$$

For an exponentially distributed variable x in $[0, \infty]$ we have $\mathrm{p}(x) = \lambda\exp(-\lambda x)$. Thus we obtain for $a \geq 0$ the result

$$\langle x \mid x \geq a\rangle = \int_a^\infty x\exp(-\lambda x)\mathrm{d}x\Big/\int_a^\infty \exp(-\lambda x)\mathrm{d}x = a + 1/\lambda.$$

♣

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

1.3. Stochastic Processes, the Kolmogorov Criterion and Martingales

In many applications (e.g. in irregular phenomena like blood flow, capital investment, or motions of molecules, etc.) one encounters a family of random variables that depend on continuous or discrete parameters like the time or positions. We refer to $\{\mathrm{X}(t,\omega), t \in \mathrm{I}, \omega \in \Omega\}$, where I is set of (continuous or discrete) parameters and $\mathrm{X}(t,\omega) \in \mathrm{R}_n$, as a **stochastic process** (random process or stochastic (random) function). If I is a discrete set it is more convenient to call $\mathrm{X}(t,\omega)$ a time series and to use the phrase process only for continuous sets. If the parameter is the time t then we use $\mathrm{I} = [t_0, \mathrm{T}]$, where t_0 is an initial instant. For a fixed value of $t \in \mathrm{I}, \mathrm{X}(t,\omega)$ is a random variable and for every fixed value of $\omega \in \Omega$ (hence for every observation) $\mathrm{X}(t,\omega)$ is a real valued function. Any observation of this process is called a **sample function** (realization, trajectory, path or orbit) of the process.

We consider now a finite variate PD of a process and we define the time dependent probability density functions (PDF) in analogy to (1.1) in the form

$$
\begin{aligned}
\mathrm{F_X}(x,t) &= \Pr(\mathrm{X}(t) < x);\\
\mathrm{F_{X,Y}}(x,t;y,s) &= \Pr(\mathrm{X}(t) < x, \mathrm{Y}(s) < y);\\
&\dots\\
\mathrm{F}_{\mathrm{X}_1,\dots,\mathrm{X}_n}(x_1,t_1;\dots;x_n,t_n) &= \Pr(\mathrm{X}_1(t) < x_1, \mathrm{X}_n(t) < x_n),
\end{aligned}
\tag{1.19}
$$

where we omit the dependence of the process $\mathrm{X}(t)$ on the chance variable ω, whenever no confusion is possible. The system of PDF's satisfies two classes of conditions:

(i) Symmetry
If $\{k_1,\dots,k_n\}$ is a permutation of $1,\dots,n$ then we obtain

$$
\mathrm{F}_{\mathrm{X}_1,\dots,\mathrm{X}_n}(x_{k_1},t_{k_1};\dots;x_{k_n},t_{k_n}) = \mathrm{F}_{\mathrm{X}_1,\dots,\mathrm{X}_n}(x_1,t_1;\dots;x_n,t_n).
\tag{1.19a}
$$

(ii) Compatibility

$$
\begin{aligned}
&\mathrm{F}_{\mathrm{X}_1,\dots,\mathrm{X}_n}(x_1,t_1;\dots;x_r,t_r;\infty,t_{r+1};\dots;\infty,t_n)\\
&\quad = \mathrm{F}_{\mathrm{X}_1,\dots,\mathrm{X}_r}(x_1,t_1;\dots;x_r,t_r).
\end{aligned}
\tag{1.19b}
$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

The rules to calculate averages are still given by (1.6) where the corresponding PD is derived by (1.3) and where the PDF's of (1.19) are used

$$\mathrm{p}(x_1, t_1; \ldots; x_n, t_n) = \frac{\partial^n}{\partial x_1(t_1)\cdots\partial x_n(t_n)} \mathrm{F}_{\mathrm{X}_1,\ldots,\mathrm{X}_n}(x_1, t_1; \ldots; x_n, t_n).$$

One would expect that a stochastic process at a high rate of irregularity (expressed e.g. by high values of intensity constants, see Chapter 2) would exhibit sample functions (SF) with a high degree of irregularity like jumps ore singularities. However, **Kolmogorov's criterion** gives a condition for continuous SF:

Theorem 1.1. (Kolmogorov's criterion)
A bivariate distribution is necessary to give information about the possibility of continuous SF. If and only if (IFF)

$$\langle |\mathrm{X}_1(t_1) - \mathrm{X}_2(t_2)|^a \rangle \le c \mid t_1 - t_2 \mid^{1+b}; \quad a, b, c > 0; \quad t_1, t_2 \in [t_0, \mathrm{T}], \tag{1.20}$$

then the stochastic process $\mathrm{X}(t)$ posses almost certainly (AC, this symbol is discussed in Chapter 5) continuous SF. However, the latter are nowhere differentiable and exhibit jumps, and higher order derivatives singularities. ✾

We will use later the Kolmogorov's criterion to investigate SF of Brownian motions and of stochastic integrals.

Definition 1.7. (Stationary process)
A process $x(t)$ is **stationary** if its PD is independent of a time shift τ

$$\mathrm{p}(x_1, t_1 + \tau; \ldots; x_n, t_n + \tau) = \mathrm{p}(x_1, t_1; \ldots; x_n, t_n). \tag{1.21a}$$

Equation (1.21a) implies that all moments are also independent of the time shift

$$\begin{aligned} &\langle x(t_1 + \tau)x(t_2 + \tau)\cdots x(t_k + \tau)\rangle \\ &\quad = \langle x(t_1)x(t_2)\cdots x(t_k)\rangle; \quad \text{for } k = 1, 2 \ldots. \end{aligned} \tag{1.21b}$$

A consequence of (1.25a) is given by

$$\begin{aligned} \langle x(t)\rangle &= \langle x\rangle, \quad \text{independent of } t; \\ \langle x(t)x(t+\tau)\rangle &= \langle x(0)x(\tau)\rangle = g(\tau). \end{aligned} \tag{1.21c}$$

□

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

The **correlation matrix** is defined by

$$c_{ik} = \langle z_i(t_1) z_k(t_2) \rangle; \quad z_i(t_1) = x_i(t_1) - \langle x_i(t_1) \rangle. \tag{1.22}$$

Thus, we have

$$c_{ik} = \langle x_i(t_1) x_k(t_2) \rangle - \langle x_i(t_1) \rangle \langle x_k(t_2) \rangle. \tag{1.23}$$

The diagonal elements of this matrix are called **autocorrelation functions** (we do not employ a summation convention)

$$c_{ii} = \langle z_i(t_1) z_i(t_2) \rangle.$$

The nondiagonal elements are referred to as **cross-correlation functions**. The **correlation coefficient** (the **nondimensional correlation**) is defined by

$$\mathrm{r}_{ik} = \frac{\langle x_i(t_1) x_k(t_2) \rangle - \langle x_i(t_1) \rangle \langle x_k(t_2) \rangle}{\sqrt{\langle x_i^2(t_1) \rangle - \langle x_i(t_1) \rangle^2} \sqrt{\langle x_k^2(t_2) \rangle - \langle x_k(t_2) \rangle^2}}. \tag{1.24}$$

For stationary processes we have

$$\begin{aligned} c_{ik}(t_1, t_2) &= \langle z_i(0) z_k(t_2 - t_1) \rangle = c_{ik}(t_2 - t_1); \\ c_{ki}(t_1, t_2) &= \langle z_k(t_1) z_i(t_2) \rangle = \langle z_k(t_1 - t_2) z_i(0) \rangle = c_{ik}(t_1 - t_2). \end{aligned} \tag{1.25}$$

A stochastic function with $c_{ik} \equiv 0$ is called an **uncorrelated function** and we obtain

$$\langle x_i(t_1) x_k(t_2) \rangle = \langle x_i(t_1) \rangle \langle x_k(t_2) \rangle. \tag{1.26}$$

Note that the condition of noncorrelation (1.26) is weaker than the condition of statistical independence.

Example

We consider the process $\mathrm{X}(t) = \mathrm{U}_1 \cos t + \mathrm{U}_2 \sin t$. $\mathrm{U}_{1,2}$ are independent stochastic variables independent of the time. The moments of the latter are given by $\langle \mathrm{U}_k \rangle = 0$, $\langle \mathrm{U}_k^2 \rangle = a = \text{const}; k = 1, 2$, $\langle \mathrm{U}_1 \mathrm{U}_2 \rangle = 0$. Hence we obtain $\langle \mathrm{X} \rangle = 0; c_{\mathrm{xx}}(s, t) = a \cos(t - s)$. ♣

Remark (Statistical mechanics and stochastic differential equations)
In Chapter 2 we will see that stochastic differential equations or "stochastic mechanics" can be used to investigate a single mechanical system in the presence of stochastic influences (white or colored

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

noise). We use concepts that are similar to those developed in statistical mechanics such as probability distribution functions, moments, Markov properties, ergodicity, etc. We solve the stochastic differential equation (analytically, but in most cases numerically) and one solution represents a realization of the system. Repeating the solution process we obtain another realization and in this way we are able to calculate the moments of the system. An alternative way to calculate the moments would be to solve the Fokker–Planck equation (see: Chapter 3) and then use the corresponding solution to determine the moments. To establish the Fokker–Planck equation we will use again the coefficients of the stochastic differential equation.

Statistical Mechanics works with the use of ensemble averages. Rather than defining a single quantity (e.g. a particle) with a PD $\mathrm{p}(x)$, one introduces a fictitious set of an arbitrary large number of M quantities (e.g. particles or thermodynamic systems) and these M non-interacting quantities define the ensemble. In case of interacting particles, the ensemble is made up by M different realizations of the N particles. In general, these quantities have different characteristic values (temperature, or energy, or values of N) x, in a common range. The number of quantities having a characteristic value between x and $x + \mathrm{d}x$ defines the PD. Therefore, the PD is replaced by density function for a large number of samples. One observes a large number of quantities and averages the results. Since, by definition, the quantities do not interact one obtains in this way a physical realization of the ensemble. The averages calculated with this density function are referred to as ensemble averages and a system where ensemble averages equal time averages is called an ergodic system. In stochastic mechanics we say that a process with the property that the averages defined in accordance with (1.6) equal the time averages, represents an ergodic process.

An other stochastic process that posses SF of some regularity is called a **martingale**. This name is related to "fair games" and we give a discussion of this expression in a moment.

In everyday language, we can state that the best prediction of a martingale process $\mathrm{X}(t)$ conditional on the path of all Brownian

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

motions up to $s < t$ is given by previous value $\mathrm{X}(s)$. To make this idea precise we formulate the following theorem:

Theorem 1.2. (Adapted process)
We consider a probability space $(\Gamma, \Omega, \Pr)$ with an increasing family (of sigma algebras of Γ) of events $\Gamma_s \in \Gamma_t, 0 \leq s < t$ (see Section 1.1). A process $\mathrm{X}(s, \omega); \omega \in \Omega, s \in [0, \infty)$ is called Γ_s-adapted if it is Γ_s-measurable. An Γ_s**-adapted process** can be expanded into a (the limit) of a sequence of Brownian motions $\mathrm{B}_u(\omega)$ with $u < s$ (but not $u > s$). ✾

Example

For $n = 2, 3, \ldots; 0 \leq \lambda \leq t$ we see that the processes

(i) $\mathrm{G}_1(t, \omega) = \mathrm{B}_{t/n}(\omega), \quad \mathrm{G}_2(t, \omega) = \mathrm{B}_{t-\lambda}(\omega),$
(ii) $\mathrm{G}_3(t, \omega) = \mathrm{B}_{nt}(\omega), \quad \mathrm{G}_4(t, \omega) = \mathrm{B}_{t+\lambda}(\omega),$

are Γ_t-adapted, respectively, not adapted. ♣

Theorem 1.3. (martingale process)
A process $\mathrm{X}(t)$ is called a **martingale** IFF it is adapted and the condition

$$\langle \mathrm{X}_t \mid \Gamma_s \rangle = \mathrm{X}_s \; \forall\, 0 \leq s \leq t < \infty, \tag{1.27}$$

is almost certainly (AC) satisfied.

If we replace the equality sign in (1.27) by $\leq$ ($\geq$) we obtain a super (sub) martingale. We note that martingales have no other discontinuities than at worst finite jumps (see Arnold [1.2]). ✾

Note that (1.27) defines a stochastic process. Its expectation $\langle\langle \mathrm{X}_t \mid \Gamma_s \rangle\rangle = \langle \mathrm{X}_s \rangle; s \leq t$ is a deterministic function.

An interesting property of a martingale is expressed by

$$\Pr(\sup \mid \mathrm{X}(t) \mid \geq c) \leq \langle \mid \mathrm{X}(b) \mid^p \rangle / c^p; \quad c > 0; \quad p \geq 1, \tag{1.28}$$

where sup is the supremum of the embraced process in the interval $[a, b]$. (1.28) is a particular version of the Chebyshev inequality, that

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

will be derived in EX 1.2. We apply later the concept of martingales to Wiener processes and to stochastic integrals.

Finally we give an explanation of the phrase "martingale". A gambler is involved in a fair game and he has at the start the capital $\mathrm{X}(s)$. Then he should posses in the mean at the instant $t \geq s$ the original capital $\mathrm{X}(s)$. This is expressed in terms of the conditional mean value $\langle \mathrm{X}_t \mid \mathrm{X}_s \rangle = \mathrm{X}_s$. Etymologically, this term comes from French and means a system of betting which seeks the amount to be wagered after each win or loss.

1.4. The Gaussian Distribution and Limit Theorems

In relation (1.13) we have already introduced a special case of the Gaussian (normal distributed) PD (GD) for a scalar variable. A generalization of (1.13) is given by the $\mathrm{N}(m, \sigma^2)$ PD

$$\mathrm{p}(x) = (2\pi\sigma^2)^{-1/2} \exp[-(x-m)^2/(2\sigma^2)]; \quad \forall\, x \in [-\infty, \infty] \tag{1.29}$$

where m is the average and $\sigma^2 = \langle x^2 \rangle - m^2$ is the variance. The multivariate form of the Gaussian PD for the set of variables $x_1, \ldots, x_n$ has the form

$$\mathrm{p}(x_1, \ldots, x_n) = \mathrm{N} \exp\left(-\frac{1}{2} \mathrm{A}_{ik} x_i x_k - b_k x_k \right), \tag{1.30a}$$

where we use a summation convention. The normalization constant N is given by

$$\mathrm{N} = (2\pi)^{-n/2} [\mathrm{Det(A)}]^{1/2} \exp\left(-\frac{1}{2} \mathrm{A}^{-1}_{ik} b_i b_k \right). \tag{1.30b}$$

We define the characteristic function of (1.30) has the form

$$\mathrm{G}(k_1, \ldots, k_n) = \exp\left[-\mathrm{A}^{-1}_{uv} \left(\frac{1}{2} k_u k_v - i k_u b_v \right) \right]. \tag{1.31}$$

An expansion of (1.31) WRT powers of k yields the moments

$$\langle x_i \rangle = -\mathrm{A}^{-1}_{ik} b_k, \tag{1.32a}$$

and the covariance is given by

$$c_{ik} = \langle (x_i - \langle x_i \rangle)(x_k - \langle x_k \rangle) \rangle = \mathrm{A}^{-1}_{ik}. \tag{1.32b}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

This indicates that the GD is completely given, if the mean value and the covariance matrix are evaluated. The n variables are uncorrelated and thus are independent if A^{-1} and hence A itself are diagonal.

The higher moments of n-variate GD with zero mean are particularly easy to calculate. To show this, we recall that for zero mean we have $b_k = 0$ and we obtain the characteristic function with the use of (1.31) and (1.32) in form of

$$\begin{aligned} \mathrm{G} &= \exp\left[-\frac{1}{2}\langle x_u x_v\rangle k_u k_v\right] \\ &= 1 + \frac{1}{2}\langle x_u x_v\rangle z_u z_v + \frac{1}{8}\langle x_u x_v\rangle\langle x_p x_q\rangle z_u z_v z_p z_q + \cdots ; \\ z_r &= ik_r; \quad \forall u, v, p, q, r = 1, 2, \ldots . \end{aligned} \tag{1.33}$$

A comparison of equal powers of z in (1.33) and in a Taylor expansion of the exponent in (1.31) shows that all odd moments vanish

$$\langle x_a x_b x_c\rangle = \langle x_a x_b x_c x_d x_e\rangle = \cdots 0.$$

We also obtain with restriction to $n = 2$ (bivariate GD)

$$\begin{aligned} \langle x_k^4\rangle &= 3\langle x_k^2\rangle^2; \quad \langle x_k^3 x_p\rangle = 3\langle x_k^2\rangle\langle x_1 x_2\rangle, \quad i, p = 1, 2; \\ \langle x_1^2 x_2^2\rangle &= \langle x_1^2\rangle\langle x_2^2\rangle + 2\langle x_1 x_2\rangle^2. \end{aligned} \tag{1.34}$$

In the case of a trivariate PD we face additional terms of the type $\langle x_k x_p^2 x_r\rangle = 2\langle x_k x_p\rangle\langle x_p x_r\rangle + \langle x_k x_r\rangle\langle x_p^2\rangle$. The higher order variate and higher order moments can be calculated in analogy to the results (1.34).

We give also the explicit formula of the bivariate Gaussian (see also EX 1.3)

$$\mathrm{p}(x, y) = \mathrm{N}_2 \exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{\xi^2}{a} - \frac{2r\xi\eta}{\sqrt{ab}} + \frac{\eta^2}{b}\right]\right\}, \tag{1.35a}$$

$$\xi = x - \langle x\rangle, \quad \eta = y - \langle y\rangle,$$

with

$$\frac{1}{\mathrm{N}_2} = 2\pi\sqrt{ab(1-r^2)}; \quad \sigma_x^2 = a; \quad \sigma_y^2 = b, \tag{1.35b}$$

and where $r = r_{12}$ is defined as the cross correlation coefficient (1.24). For $\sigma_x = \sigma_y = 1$ and $\langle x\rangle = \langle y\rangle = 0$ in (1.35) we can expand the latter

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

formula and we obtain

$$\mathrm{p}(x,y) = (2\pi)^{-1}\exp\bigl[-(x^2+y^2)/2\bigr]\sum_{k=0}^{\infty}\frac{r^k}{k!}\mathrm{H}_k(x)\mathrm{H}_k(y), \qquad (1.36)$$

where $\mathrm{H}_k(x)$ is the k-th order Hermite polynomial (see Abramowitz and Stegun [1.3]). Equation (1.36) is the basis of the "Hermitian-chaos" expansion in the theory of stochastic partial differential equations.

In EX 1.3 we show that conditional probabilities of the GD (1.35a) are Gaussian themselves.

Now we consider two limit theorems. The first of them is related to GD and we introduce the second one for later use.

1.4.1. *The central limit theorem*

We consider the random variable

$$\mathrm{U} = \frac{x_1 + \cdots + x_n}{\sqrt{n}}; \quad \langle x_k\rangle = 0, \qquad (1.37)$$

where x_k are identically independent distributed (IID) (but not necessarily normal) variables with zero mean and variance $\sigma^2 = \langle x_k^2\rangle$. We find easily $\langle \mathrm{U}\rangle = 0$ and $\langle \mathrm{U}^2\rangle = \sigma^2$.

The **central limit theorem** says that U tends in the limit $n \to \infty$ to a $\mathrm{N}(0, \sigma^2)$ variable with a PD given by (1.13a). To prove this we use the independence of the variables x_k and we perform the calculation of the characteristic function of the variable U with the aid of (1.12)

$$\begin{aligned}\mathrm{G_U}(k) &= \int \mathrm{d}x_1\mathrm{p}(x_1)\cdots\int \mathrm{d}x_n\mathrm{p}(x_n)\cdots\exp\left[ik(x_1+\cdots+x_n)/\sqrt{n}\right]\\ &= \left[\mathrm{G}_x(k/\sqrt{n})\right]^n\\ &= \left[1 - \frac{k^2\sigma^2}{2n} + \mathrm{O}(n^{-3/2})\right]^n \to \exp(-k^2\sigma^2/2) \quad \text{for } n\to\infty.\end{aligned} \qquad (1.38)$$

We introduced in the second line of (1.38) the characteristic function of one of the individual random functions according to (1.14a); (1.38) is the characteristic function of a GD that corresponds indeed to

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

$N(0, \sigma^2)$. Note that this result is independent of the particular form of the individual PD's $p(x)$. It is only required that $p(x)$ has finite moments. The central limit theorem explains why the Gaussian PD plays a prominent role in probability and stochastics.

1.4.2. *The law of the iterated logarithm*

We give here only this theorem and refer the reader for its derivation to the book Chow and Teichler [1.4]. y_n is the partial sum of n IID variables

$$y_n = x_1 + \cdots + x_n; \quad \langle x_n \rangle = \beta, \quad \langle (x_n - \beta)^2 \rangle = \sigma^2. \tag{1.39}$$

The theorem of the iterated logarithm states that there exists AC an asymptotic limit

$$-\sigma \leq \lim_{n\to\infty} \frac{y_n - n\beta}{\sqrt{2n \ln[\ln(n)]}} \leq \sigma. \tag{1.40}$$

Equation (1.40) is particular valuable in case of estimates of stochastic functions and we will use it later to investigate Brownian motions. We will give a numerical verification of (1.40) in program F18.

1.5. Transformation of Stochastic Variables

We consider transformations of an n-dimensional set of stochastic variables $x_1, \ldots, x_n$ with the PD $p_{X_1 \cdots X_n}(x_1, \ldots, x_n)$. First we introduce the PD of a linear combination of random variables

$$z = \sum_{k=1}^{n} \alpha_k x_k, \tag{1.41a}$$

where the α_k are deterministic constants. The PD of the stochastic variable z is then defined by

$$P_Z(z) = \int dx_1 \cdots \int dx_n \delta\left(z - \sum_{k=1}^{n} \alpha_k x_k\right) p_{X_1 \cdots X_n}(x_1, \ldots, x_n). \tag{1.41b}$$

Now we investigate transformations of the stochastic variables $x_1, \ldots, x_n$. The new variables are defined by

$$u_k = u_k(x_1, \ldots, x_n), \quad k = 1, \ldots, n. \tag{1.42}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

The inversion of this transformation and the Jacobian are

$$x_k = \mathrm{g}_k(u_1, \ldots, u_n), \quad \mathrm{J} = \partial(x_1, \ldots, x_1)/\partial(u_1, \ldots, u_1). \tag{1.43}$$

We infer from an expansion of the probability measure (1.1a) that

$$\begin{aligned} \mathrm{dp}_{\mathrm{X}_1 \cdots \mathrm{X}_n} &= \Pr(x_1 \leq \mathrm{X}_1 \leq x_1 + \mathrm{d}x_1, \ldots, x_n \leq \mathrm{X}_n \leq x_n + \mathrm{d}x_n) \\ &= \mathrm{p}_{\mathrm{X}_1 \cdots \mathrm{X}_n}(x_1, \ldots, x_n)\mathrm{d}x_1 \cdots \mathrm{d}x_n \\ &\qquad \text{for } \mathrm{d}x_k \to 0,\ k = 1, \ldots, n. \end{aligned} \tag{1.44a}$$

Equation (1.44a) represents the elementary probability measure that the variables are located in the hyper plane

$$\prod_{k=1}^{n} [x_k, x_k + \mathrm{d}x_k].$$

The principle of **invariant elementary probability measure** states that this measure is invariant under transformations of the coordinate system. Thus, we obtain the transformation

$$\mathrm{dp}_{\mathrm{U}_1 \cdots \mathrm{U}_n} = \mathrm{dp}_{\mathrm{X}_1 \cdots \mathrm{X}_n}. \tag{1.44b}$$

This yields the transformation rule for the PD's

$$\begin{aligned} &\mathrm{p}_{\mathrm{U}_1 \cdots \mathrm{U}_n}(u_1(x_1, \ldots, x_n), \ldots, u_n(x_1, \ldots, x_n)) \\ &\quad = \mid \det(\mathrm{J}) \mid \mathrm{p}_{\mathrm{X}_1 \cdots \mathrm{X}_n}(x_1, \ldots, x_n). \end{aligned} \tag{1.45}$$

Example (The Box–Miller method)
As an application we introduce the transformations method of **Box–Miller** to generate a GD. There are two stochastic variables given in an elementary cube

$$\mathrm{p}(x_1, x_2) = \begin{pmatrix} 1 & \forall\, 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1 \\ 0 & \text{elsewhere} \end{pmatrix}. \tag{1.46}$$

Note that the bivariate PD is already normalized. Now we introduce the new variables

$$\begin{aligned} y_1 &= \sqrt{-2\ln x_1}\cos(2\pi x_2), \\ y_2 &= \sqrt{-2\ln x_1}\sin(2\pi x_2). \end{aligned} \tag{1.47}$$

The inversion of (1.47) is

$$x_1 = \exp\left[-\left(y_1^2 + y_2^2\right)/2\right]; \quad x_2 = \frac{1}{2\pi}\mathrm{arc\ tan}(y_2/y_1).$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

According to (1.45) we obtain the new bivariate PD

$$\mathrm{p}(y_1, y_2) = \mathrm{p}(x_1, x_2)\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \frac{1}{2\pi}\exp\left[-\left(y_1^2 + y_2^2\right)/2\right], \tag{1.48}$$

and this the PD of two independent N(0, 1) variables.

Until now we have only covered stochastic variables that are time-independent or stochastic processes for the case that all variables belong to the same instant. In the next section we discuss a property that is rather typical for stochastic processes. ♣

1.6. The Markov Property

A process is called a **Markov** (or **Markovian**) **process** if the conditional PD at a given time t_n depends only on the immediately prior time t_{n-1}. This means that for $t_1 < t_2 < \cdots < t_n$

$$\mathrm{p}_{1|n-1}(y_n, t_n \mid y_1, t_1; \ldots; y_{n-1}, t_{n-1}) = \mathrm{p}_{1|1}(y_n, t_n \mid y_{n-1}, t_{n-1}), \tag{1.49}$$

and the quantity $\mathrm{p}_{1|1}(y_n, t_n \mid y_{n-1}, t_{n-1})$ is referred to as **transition probability distribution** (TPD).

A Markov process is thus completely defined if we know the two functions

$$\mathrm{p}_1(y_1, t_1) \quad \text{and} \quad \mathrm{p}_2(y_2, t_2 \mid y_1, t_1) \quad \text{for } t_1 < t_2.$$

Thus, we obtain for $t_1 < t_2$ (see (1.15″) and note that we use a semicolon to separate coordinates that belong to different instants)

$$\mathrm{p}_2(y_1, t_1; y_2, t_2) = p_1(y_1, t_1)\mathrm{p}_{1|1}(y_2, t_2 \mid y_1, t_1), \tag{1.50.1}$$

and for $t_1 < t_2 < t_3$

$$\begin{aligned}&\mathrm{p}_3(y_1, t_1; y_2, t_2; y_3, t_3)\\ &\quad = \mathrm{p}_1(y_1, t_1)\mathrm{p}_{1|1}(y_2, t_2 \mid y_1, t_1)\mathrm{p}_{1|1}(y_3, t_3 \mid y_2, t_2).\end{aligned} \tag{1.50.2}$$

We integrate equation (1.50.2) over the variable y_2 and we obtain

$$\mathrm{p}_2(y_1, t_1; y_3, t_3) = p_1(y_1, t_1)\int \mathrm{p}_{1|1}(y_2, t_2 \mid y_1, t_1)\mathrm{p}_{1|1}(y_3, t_3 \mid y_2, t_2)\mathrm{d}y_2. \tag{1.51}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

Now we use

$$\mathrm{p}_{1|1}(y_3, t_3 \mid y_1, t_1) = \mathrm{p}_2(y_1, t_1; y_3, t_3)/p_1(y_1, t_1),$$

and we obtain from (1.51) the **Chapman–Kolmogorov equation**

$$\mathrm{p}_{1|1}(y_3, t_3 \mid y_1, t_1) = \int \mathrm{p}_{1|1}(y_2, t_2 \mid y_1, t_1)\mathrm{p}_{1|1}(y_3, t_3 \mid y_2, t_2)\mathrm{d}y_2. \tag{1.52}$$

It is easy to verify that a particular solution of (1.52) is given by

$$\mathrm{p}_{1|1}(y_2, t_2 \mid y_1, t_1) = [2\pi(t_2 - t_1)]^{-1/2} \exp\{-(y_2 - y_1)^2/[2(t_2 - t_1)]\}. \tag{1.53}$$

We give in EX 1.4 hints how to verify (1.53).

We can also integrate the identity (1.50.1) over y_1 and we obtain

$$\mathrm{p}_1(y_2, t_2) = \int \mathrm{p}_1(y_1, t_1)\mathrm{p}_{1|1}(y_2, t_2 \mid y_1, t_1)\mathrm{d}y_1. \tag{1.54}$$

The latter relation is an integral equation for the function $\mathrm{p}_1(y_2, t_2)$. EX 1.5 gives hints to show that the solution to (1.54) is the Gaussian PD

$$\mathrm{p}_1(y, t) = (2\pi t)^{-1/2} \exp[-y^2/(2t)]; \quad \lim_{t \to 0_+} \mathrm{p}_1(y, t) = \delta(y). \tag{1.55}$$

In Chapter 3 we use the Chapman–Kolmogorov equation (1.52) to derive the master equation that is in turn applied to deduce the Fokker–Planck equation.

1.6.1. *Stationary Markov processes*

Stationary Markovian processes are defined by a PD and transition probabilities that depend only on the time differences. The most important example is the **Ornstein–Uhlenbeck-process** that we will treat in Section 2.1.3 and 3.4. There we will prove the formulas for its PD

$$\mathrm{p}_1(y) = (2\pi)^{-1/2} \exp(-y^2/2), \tag{1.56.1}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

and the transition probability

$$p_{1|1}(y_2, t_2 \mid y_1, t_1) = [2\pi(1-u^2)]^{-1/2} \exp\left\{-\frac{(y_2 - uy_1)^2}{[2(1-u^2)]}\right\};$$
$$u = \exp(-\tau); \quad p_{1|1}(y_2, t_1 \mid y_1, t_1) = \delta(y_2 - y_1). \tag{1.56.2}$$

The Ornstein–Uhlenbeck-process is thus stationary, Gaussian and Markovian. A theorem from Doob [1.5] states that this is apart from trivial process, where all variables are independent — the only process that satisfies all the three properties listed above. We continue to consider stationary Markov processes in Section 3.1.

1.7. The Brownian Motion

Brown discovered in year 1828 that pollen submerged in fluids show under collisions with fluid molecules, a completely irregular movement. This process is labeled with $y := \mathrm{B}_t(\omega)$, where the subscript is the time. It is also called a **Wiener** (white noise) **process** and labeled with the symbol W_t (WP) that is identical to the Brownian motion: $\mathrm{W}_t \equiv \mathrm{B}_t$. The WP is a Gaussian [it has the PD (1.55)] and a Markov process.

Note also that the PD of the Wiener process (WP) — given by (1.55) — satisfies a parabolic partial differential equation (called **Fokker–Planck equation**, see Section 3.2)

$$\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial^2 p}{\partial x^2}. \tag{1.57}$$

We calculate the characteristic function $\mathrm{G}(u)$ and we obtain according to (1.12)

$$\mathrm{G}(u) = \langle \exp(iu\mathrm{W}_t)\rangle = \exp(-u^2 t/2), \tag{1.58a}$$

and we obtain the moments in accordance with (1.13b)

$$\langle \mathrm{W}_t^{2k}\rangle = \frac{(2k)!}{2^k k!}t^k; \quad \langle \mathrm{W}_t^{2k+1}\rangle = 0; \quad k \in \mathrm{N}_0. \tag{1.58b}$$

We use the Markovian properties now to prove the independence of Brownian increments. The latter are defined

$$y_1, y_2 - y_1, \ldots, y_n - y_{n-1} \quad \text{with } y_k := \mathrm{W}_{t_k}; \quad t_1 \le \cdots \le t_n. \tag{1.59}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

We calculate explicitly the joint distribution given by (1.50) and we obtain with the use (1.53) and (1.55)

$$\begin{aligned}\mathrm{p}_2(y_1,t_1;y_2,t_2) = [(2\pi)^2 t_1(t_2-t_1)]^{-1/2}\exp\{-y_1^2/(2t_1)\\ -(y_2-y_1)^2/[2(t_2-t_1)]\},\end{aligned} \tag{1.60}$$

and

$$\begin{aligned}&\mathrm{p}_3(y_1,t_1;y_2,t_2;y_3,t_3)\\ &\quad= [(2\pi)^3 t_1(t_2-t_1)(t_3-t_2)]^{-1/2}\exp\{-y_1^2/(2t_1)\\ &\qquad -(y_2-y_1)^2/[2(t_2-t_1)]-(y_3-y_2)^2/[2(t_3-t_2)]\},\\ &\mathrm{p}_4(y_1,t_1;y_2,t_2;y_3,t_3;y_4,t_4)\\ &\quad= [2\pi(t_4-t_3)]^{-1/2}\mathrm{p}_3(y_1,t_1;y_2,t_2;y_3,t_3)\\ &\qquad\times\exp\{-(y_4-y_3)^2/[2(t_4-t_3)]\}.\end{aligned} \tag{1.61}$$

We see that the joint PD's of the variables $y_1, y_2-y_1, y_3-y_2, y_4-y_3$ are given in (1.60) and (1.61) in a factorized form and this implies the independence of these variables. To prove the independence of the remaining variables $y_4-y_3,\ldots,y_n-y_{n-1}$ we would only have to continue the process of constructing joint PD's with the aid of (1.49).

In EX 1.6 we prove the following property

$$\langle y_1(t_1)y_2(t_2)\rangle = \min(t_1,t_2) \equiv t_1\wedge t_2. \tag{1.62}$$

Equation (1.62) also demonstrates that the Brownian motion is not a stationary process, since the autocorrelation does not depend on the time difference $\tau = t_2-t_1$ but it depends on $t_2\wedge t_1$.

To apply Kolmogorov's criterion (1.20) we choose $a=2$ and we obtain with (1.58b) and (1.62) $\langle [y_1(t_1)-y_2(t_2)]^2\rangle = |t_2-t_1|$. Thus we can conclude with the choice $b=c=1$ that the SF of the WP are ac continuous functions. The two graphs Figures 1(a) and 1(b) are added in this section to indicate the continuous SF.

We apply also the law of iterated logarithm to the WP. To this end we consider the independent increments y_k-y_{k-1} where we $t_k = k\Delta t$ with a finite time increment Δt. This yields for the partial sum in (1.39)

$$\sum_{k=1}^{n}(y_k-y_{k-1}) = y_n = y_{n\Delta t};\quad \alpha=\langle y_k\rangle = 0;\quad \langle (y_k-y_{k-1})^2\rangle = \Delta t.$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

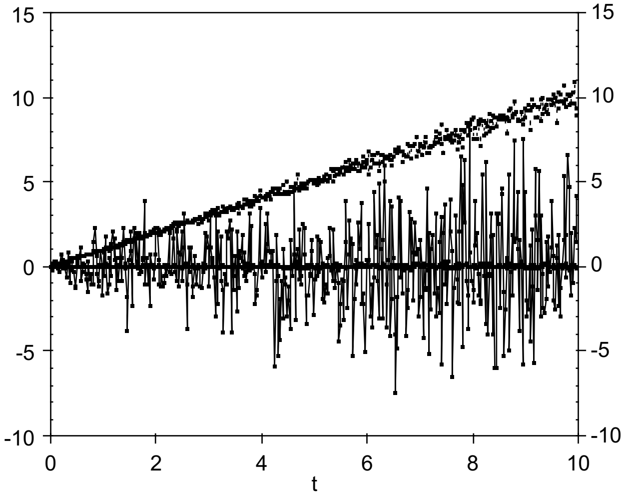

Fig. 1(a). The Brownian motion B_t versus the time axis. Included is a graph of the numerically determined temporal evolution of the mean value and the variance.

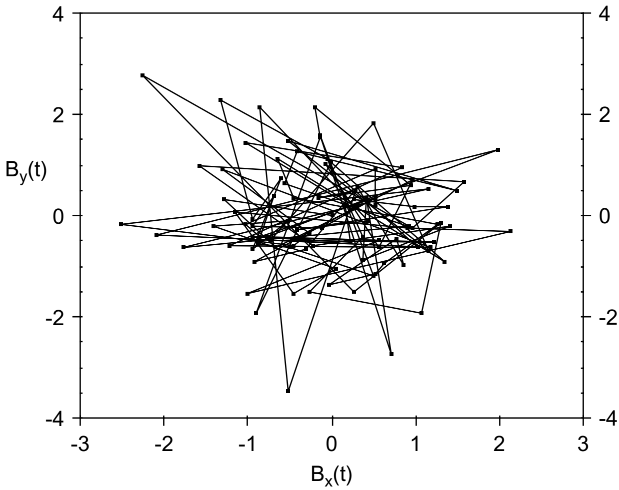

Fig. 1(b). The planar Brownian motion with $x = B_t^1$ and $y = B_t^2 \cdot B_t^k$, $k = 1, 2$ are independent Brownian motions.

We substitute the results of the last line into (1.40) and we obtain

$$-\sqrt{\Delta t} \leq \lim_{n\to\infty} \frac{W_{n\Delta t}}{\sqrt{2n\ln(\ln(n))}} \leq \sqrt{\Delta t}.$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

The assignment of $t := n\Delta t$ into the last line and the approximation $\ln(t/\Delta t) \to \ln(t)$ for $t \to \infty$ gives the desired result for the AC asymptotic behavior of the WP

$$-1 \le \lim_{n\to\infty} \frac{\mathrm{W}_t}{\sqrt{2t\ln(\ln(t))}} \le 1. \tag{1.63}$$

We will verify (1.63) in Chapter 5 numerically.

There are various equivalent definitions of a Wiener process. We use the following:

Definition 1.8. (Wiener process)
A WP has an initial value of $\mathrm{W}_0 = 0$ and its increments $\mathrm{W}_t - \mathrm{W}_s$, $t > s$ satisfies three conditions. They are

(i) independent and (ii) stationary (the PD dependence on $t - s$) and (iii) $\mathrm{N}[0, t - s]$ distributed.

As a consequence of these three conditions WP exhibits continuous sample functions with probability 1. □

There are also WP's that do not start at zero. There is also a generalization of the WP with discontinuous SF. We will return to this point at the end of Section 1.7.

Now we show that a WP is a martingale

$$\langle \mathrm{B}_s \mid \mathrm{B}_u \rangle = \mathrm{B}_u; \quad s > u. \tag{1.64}$$

We prove (1.64) with the application of the Markovian property (1.53). We use (1.17) write

$$\begin{aligned}\langle \mathrm{B}_s \mid \mathrm{B}_u \rangle = \langle y_2, s \mid y_1, u \rangle &= \int y_2 \mathrm{p}_{1|1}(y_2, s \mid y_1, u)\mathrm{d}y_2 \\ &= \frac{1}{\sqrt{2\pi(s-u)}} \int y_2 \exp\{-(y_2 - y_1)^2/[2(s-u)]\}\mathrm{d}y_2 \\ &= y_1 \equiv \mathrm{B}_u.\end{aligned}$$

This concludes the proof of (1.64).

A WP has also the following properties. The translated quantity $\hat{\mathrm{W}}_t$ and the scaled quantity $\tilde{\mathrm{W}}_t$ defined by

$$t, a > 0 : \hat{\mathrm{W}}_t = \mathrm{W}_{t+a} - \mathrm{W}_a \quad \text{and} \quad \tilde{\mathrm{W}}_t = \frac{1}{a}\mathrm{W}_{a^2 t}, \tag{1.65}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

are also a Brownian motion. To prove (1.65) we note first that the averages of both variables are zero $\langle \hat{\mathrm{W}}_t \rangle = \langle \tilde{\mathrm{W}}_t \rangle = 0$. Now we have to show that both variables satisfy also the condition for the auto correlation. We prove this only for the variable $\tilde{\mathrm{W}}_t$ and leave the second part for the EX 1.7. Thus, we put

$$\langle \tilde{\mathrm{W}}_t \tilde{\mathrm{W}}_s \rangle = \langle \mathrm{B}_{a^2 t} \mathrm{B}_{a^2 s} \rangle / a^2 = \frac{(a^2 t) \wedge (a^2 s)}{a^2} = t \wedge s.$$

So far, we considered exclusively scalar WP's. In the study of partial differential equations we need to introduce a set of n independent WP's. Thus, we generalize the WP to the case of an independent WP's that define a vector of a stochastic processes

$$x_1(t_1), \ldots, x_n(t_n); \quad t_k \geq 0. \tag{1.66}$$

The corresponding PD is then

$$\begin{aligned} \mathrm{p}(x_1, \ldots, x_n) &= \mathrm{p}_{x_1}(x_1) \ldots \mathrm{p}_{x_n}(x_n) \\ &= (2\pi)^{-n/2} \prod_{k=1}^{n} t_k^{-1/2} \exp\left[-x_k^2/(2t_k)\right]. \end{aligned} \tag{1.67}$$

We have assumed independent stochastic variables (like the orthogonal basic vectors in the case of deterministic variables) and this independence is expressed by the factorized multivariate PD (1.67).

We define an n-dimensional WP (or a **Wiener sheet** (WS)) by

$$\mathrm{M}_t^{(n)} = \prod_{k=1}^{n} x_k(t_k); \quad t = (t_1, \ldots, t_n). \tag{1.68}$$

Now we find how we can generalize the Definition 1.8 to the case of n stochastic processes. First, we prove easily that the variable (1.68) has a zero mean

$$\left\langle \mathrm{M}_t^{(n)} \right\rangle = 0. \tag{1.69}$$

Thus, it remains to calculate the autocorrelation (1.62). We use the independence of the set of variables $x_k(t_k), k = 1, \ldots, n$ and we obtain with the use of the bivariate PD (1.61) with $y_1 = x_k(t_k)$;

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

$y_2 = x_k(s_k)$ and factorize the result for the independent variables. Hence, we obtain

$$\left\langle \mathrm{M}_t^{(n)} \mathrm{M}_s^{(n)} \right\rangle = \prod_{k=1}^{n} \langle x_k(t_k) x_k(s_k) \rangle; \;\; t = (t_1, \ldots, t_n); \;\; s = (s_1, \ldots, s_n).$$

The evaluation of the last line yields with (1.62)

$$\left\langle \mathrm{M}_t^{(n)} \mathrm{M}_s^{(n)} \right\rangle = \prod_{k=1}^{n} t_k \wedge s_k. \tag{1.70}$$

The relations (1.69) and (1.70) show now that process (1.68) is an n-WP.

In analogy to deterministic variables we can now construct with stochastic variables curves, surfaces and hyper surfaces. Thus, a curve in 2-dimensional WS and surfaces on 3-dimensional WS are given by

$$\mathrm{C}_t = \mathrm{M}_{t,f(t)}^{(2)}; \quad \mathrm{S}_{t_1,t_2} = \mathrm{M}_{t_1,t_2,\mathrm{g}(t_1,t_2)}^{(3)}.$$

We give here only two interesting examples.

Example 1
Here we put

$$\mathrm{K}_t = \mathrm{M}_{a,b}^{(2)}; \quad a = \exp(t), \quad b = \exp(-t); \quad -\infty < x < \infty.$$

This defines a stochastic hyperbola with zero mean and with the autocorrelation

$$\begin{aligned} \langle \mathrm{K}_t \mathrm{K}_s \rangle &= \langle x_1(e^t) x_1(e^s) \rangle \langle x_2(e^{-t}) x_2(e^{-s}) \rangle \\ &= (e^t \wedge e^s)(e^{-t} \wedge e^{-s}) = \exp(-|t-s|). \end{aligned} \tag{1.71}$$

The property (1.71) shows this process is not only a WS but also a stationary Ornstein–Uhlenbeck process (see Section 1.5.1). ♣

Example 2
Here we define the process

$$\mathrm{K}_t = \exp[-(1+c)t]\mathrm{M}_{a,b}^{(2)}; \quad a = \exp(2t), \quad b = \exp(2ct); \quad c > 0. \tag{1.72}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

Again we see, that stochastic variable defined (1.72) has zero mean and the calculation of its autocorrelation yields

$$\begin{aligned}\langle \mathrm{K}_t \mathrm{K}_s \rangle &= \exp[-(1+c)(t+s)]\langle x_1(e^{2t})x_1(e^{2s})\rangle\langle x_2(e^{2ct})x_2(e^{2cs})\rangle \\ &= \exp[-(1+c)(t+s)](e^{2t} \wedge e^{2s})(e^{2ct} \wedge e^{2cs}) \\ &= \exp[-(1+c)|t-s|]. \end{aligned} \tag{1.73}$$

The latter equation means that the process (1.72) is again an Ornstein–Uhlenbeck process. Note also that because of $c > 0$ there is no possibility to use (1.73) to reproduce the result of the previous example. ♣

Just as in the case of one parameter, there exist for WS's also scaling and translation. Thus, the stochastic variables

$$\begin{aligned} \mathrm{H}_{u,v} &= \frac{1}{ab}\mathrm{M}^{(2)}_{a^2u,b^2v}; \\ \mathrm{L}_{u,v} &= \mathrm{M}^{(2)}_{u+a,v+b} - \mathrm{M}^{(2)}_{u+a,b} - \mathrm{M}^{(2)}_{a,v+b} - \mathrm{M}^{(2)}_{a,b}, \end{aligned} \tag{1.74}$$

are also WS's. The proof of (1.74) is left for EX 1.8.

We give in Figures 1(a) and 1(b) two graphs of the Brownian motion.

At the end of this section we wish to mention that the WP is a subclass of a **Levy process** $\mathrm{L}(t)$. The latter complies with the first two conditions of the Definition 1.8. However, it does not possess normal distributed increments. A particular feature of normal distributed process x is the vanishing of the skewness $\langle x^3\rangle/\langle x^2\rangle^{3/2}$. However, many statistical phenomena (like hydrodynamic turbulence, the market values of stocks, etc.) show remarkable values of the skewness. This means that a GD (with only two parameter) is not flexible enough to describe such phenomena and it must be replace by a PD that contains a sufficient number of parameters. An appropriate choice is the normal inverted Gaussian distribution (NIGD) (see Section 4.4). The NIGD distribution does not satisfy the Kolmogorov criterion. This means that the sample functions of the Levy process $\mathrm{L}(t)$ is equipped with SF that jump up and down at arbitrary instances t. To get more information about the Levy process we refer the reader to the work of Ikeda & Watanabe [1.6] and of Rydberg

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

[1.7]. In Section 4.4 we will give a short description of the application of the NIGD in economics theories.

1.8. Stochastic Integrals

We need stochastic integrals (SI) when we attempt to solve a stochastic differential equation (SDE). Hence we introduce a simple first order ordinary SDE

$$\frac{\mathrm{dX}}{\mathrm{d}t} = a(\mathrm{X}(t), t) + b(\mathrm{X}(t), t)\xi_t; \quad \mathrm{X}, a, b, t \in \mathrm{R}. \tag{1.75}$$

We use in (1.75) the deterministic functions a and b. The symbol ξ_t indicates the only stochastic term in this equation. We assume

$$\langle \xi_t \rangle = 0; \quad \langle \xi_t \xi_s \rangle = \delta(t-s). \tag{1.76}$$

The spectrum of the autocorrelation in (1.76) is constant (see Section 2.2) and in view of this ξ_t is referred as white noise and any term proportional to ξ_t is called a noisy term. These assumptions are based on a great variety of physical phenomena that are met in many experimental situations.

Now we replace (1.75) by a discretization and we put

$$\Delta t_k = t_{k+1} - t_k > 0; \quad \mathrm{X}_k = \mathrm{X}(t_k);$$
$$\Delta \mathrm{X}_k = \mathrm{X}_{k+1} - \mathrm{X}_k; \quad k = 0, 1, \ldots.$$

The substitution into (1.75) yields

$$\Delta \mathrm{X}_k = a(\mathrm{X}_k, t_k)\Delta t_k + b(\mathrm{X}_k, t_k)\Delta \mathrm{B}_k;$$
$$\Delta \mathrm{B}_k = \mathrm{B}_{k+1} - \mathrm{B}_k; \quad k = 1, 2, \ldots$$

where we used

$$\xi(t_k) = \frac{\Delta \mathrm{B}_k}{\Delta t_k}; \quad \xi(t) = \lim_{\Delta t_k \to 0} \frac{\mathrm{dB}(t)}{\mathrm{d}t}. \tag{1.77}$$

A precise derivation of (1.77) is given in Section 2.2. Thus we can write (1.75) in terms of

$$\mathrm{X}_n = \mathrm{X}_0 + \sum_{s=0}^{n-1} \Big[a(\mathrm{X}_s, t_s)\Delta t_s + b(\mathrm{X}_s, t_s)\Delta \mathrm{B}_s\Big]; \quad \mathrm{X}_0 = \mathrm{X}(t_0). \tag{1.78}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

What happens in the limit $\Delta t_k \to 0$? If there is a "reasonable" limit of the last term in (1.78) we obtain as solution of the SDE (1.75)

$$\mathrm{X}(t) = \mathrm{X}(0) + \int_0^t a(\mathrm{X}(s), s)\mathrm{d}s + \text{``}\int_0^t b(\mathrm{X}(s), s)\mathrm{dB}_s\text{''}. \tag{1.79}$$

The first integral in (1.79) is a conventional integral of Riemann's type and we put the stochastic (noisy) integral into inverted commas. The irregularity of the noise does not allow to calculate the stochastic integral in terms of a Riemann integral. This is caused by the fact that the paths of the WP are nowhere differentiable. Thus we find that a SI depends crucially on the decomposition of the integration interval.

We assumed in (1.75) to (1.79) that $b(\mathrm{X}, t)$ is a deterministic function. We generalize the problem of the calculation of a SI and we consider a stochastic function

$$\mathrm{I} = \int_0^{\mathrm{T}} \mathrm{f}(w, s)\mathrm{dB}_s. \tag{1.80}$$

We recall that Riemann integrals of the type ($\mathrm{g}(s)$ is a differentiable function)

$$\int_0^{\mathrm{T}} \mathrm{f}(s)\mathrm{dg}(s) = \int_0^{\mathrm{T}} \mathrm{f}(s)\mathrm{g}'(s)\mathrm{d}s,$$

are discretized in the following manner

$$\int_0^{\mathrm{T}} \mathrm{f}(s)\mathrm{dg}(s) = \lim_{n\to\infty} \sum_{k=0}^{n-1} \mathrm{f}(s_k)[\mathrm{g}(s_{k+1}) - \mathrm{g}(s_k)].$$

Thus, it is plausible to introduce a discretization of (1.80) that takes the form

$$\mathrm{I} = \sum_k \mathrm{f}(s_k, \omega)(\mathrm{B}_{k+1} - \mathrm{B}_k). \tag{1.81}$$

In Equation (1.81) we used s_k as time-argument for the integrand f. This is the value of s that corresponds to the left endpoint of the discretization interval and we say that this decomposition does not

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

look into the future. We call this type of integral an **Ito integral** and write

$$\mathrm{I_I} = \int_0^{\mathrm{T}} \mathrm{f}(s,\omega)\mathrm{dB}_s. \tag{1.82}$$

An other possible choice is to use the midpoint of the interval and with this we obtain the **Stratonovich integral**

$$\mathrm{I_S} = \int_0^{\mathrm{T}} \mathrm{f}(s,\omega)\mathrm{o}\,\mathrm{dB}_s = \sum_k \mathrm{f}(\tilde{s}_k,\omega)(\mathrm{B}_{k+1}-\mathrm{B}_k); \quad \tilde{s}_k = \frac{1}{2}(t_{k+1}+t_k). \tag{1.83}$$

Note that the symbol "o" between integrand and the stochastic differential is used to indicate Stratonovich integrals.

There are, of course, an uncountable infinity of other decompositions of the integration interval that yield to different definitions of a SI. It is, however, convenient to take advantage only of the Ito and the Stratonovich integral. We will discuss their properties and find out which type of integrals seems to be more appropriate for the use in the analysis of stochastic differential equations.

Properties of the Ito integral

(a) We have for deterministic constants $a < b < c, \alpha, \beta \in \mathrm{R}$.

$$\int_a^c [\alpha \mathrm{f}_1(s,\omega) + \beta \mathrm{f}_2(s,\omega)]\mathrm{dB}_s = \alpha \mathrm{I}_1 + \beta \mathrm{I}_2; \quad \mathrm{I}_k = \int_a^c \mathrm{f}_k(s,\omega)\mathrm{dB}_s. \tag{1.84}$$

Note that (1.84) remains also valid for Stratonovich integrals. The proof of (1.84) is trivial.

In the following we give non-trivial properties that apply, however, exclusively to Ito integrals. Now we need a definition:

Definition 1.9. (non-anticipative or adapted functions)
The function $\mathrm{f}(t, \mathrm{B}_s)$ is said to be **non-anticipative** (or **adapted**, see also Theorem 1.2) if it depends only on a stochastic variable of the past: B_s appears only for arguments $s \le t$. Examples for a non-anticipative functions are

$$\mathrm{f}(s,\omega) = \int_0^s \mathrm{g}(u)\mathrm{dB}_u; \quad \mathrm{f}(s,\omega) = \mathrm{B}_s. \qquad \square$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

Now we list further properties of the Ito integrals that include non anticipative functions $\mathrm{f}(s, \mathrm{B}_s)$ and $\mathrm{g}(s, \mathrm{B}_s)$.

(b)
$$\mathrm{M}_1 \equiv \left\langle \int_0^a \mathrm{f}(s, \mathrm{B}_s) \mathrm{dB}_s \right\rangle = 0. \tag{1.85}$$

Proof.
We use (1.81) and obtain

$$\mathrm{M}_1 = \left\langle \sum_k \mathrm{f}(s_k, \mathrm{B}_k)(\mathrm{B}_{k+1} - \mathrm{B}_k) \right\rangle.$$

But we know that B_k is independent of $\mathrm{B}_{k+1} - \mathrm{B}_k$. The function $\mathrm{f}(s_k, \mathrm{B}_k)$ is thus also independent of $\mathrm{B}_{k+1} - \mathrm{B}_k$. Hence we obtain

$$\mathrm{M}_1 = \sum_k \langle \mathrm{f}(s_k, \mathrm{B}_k) \rangle \langle \mathrm{B}_{k+1} - \mathrm{B}_k \rangle = 0.$$

This concludes the proof of (1.85).

(c) Here we study the average of a product of integrals and we show that

$$\mathrm{M}_2 \equiv \left\langle \int_0^t \mathrm{f}(s, \mathrm{B}_s) \mathrm{dB}_s \int_0^t \mathrm{g}(u, \mathrm{B}_u) \mathrm{dB}_u \right\rangle = \int_0^t \langle \mathrm{f}(s, \mathrm{B}_s) \mathrm{g}(s, \mathrm{B}_s) \rangle \mathrm{d}s. \tag{1.86}$$

Proof.

$$\mathrm{M}_2 = \sum_{m,n} \langle \mathrm{f}(s_m, \mathrm{B}_m)(\mathrm{B}_{m+1} - \mathrm{B}_m) \mathrm{g}(s_n, \mathrm{B}_n)(\mathrm{B}_{n+1} - \mathrm{B}_n) \rangle.$$

We have to distinguish three subclasses: (i) $n > m$, (ii) $n < m$ and (iii) $n = m$.

Taking into account the independence of the increments of WP's we see that only case (iii) contributes non-trivially to M_2. This yields

$$\mathrm{M}_2 = \sum_n \langle \mathrm{f}(s_n, \mathrm{B}_n) \mathrm{g}(s_n, \mathrm{B}_n)(\mathrm{B}_{n+1} - \mathrm{B}_n)^2 \rangle.$$

But we know that $\mathrm{f}(s_n, \mathrm{B}_n)\mathrm{g}(s_n, \mathrm{B}_n)$ is again a function that is independent of $(\mathrm{B}_{n+1} - \mathrm{B}_n)^2$. We use (1.62) and obtain

$$\langle (\mathrm{B}_{n+1} - \mathrm{B}_n)^2 \rangle = \langle \mathrm{B}_{n+1}^2 - 2\mathrm{B}_{n+1}\mathrm{B}_n + \mathrm{B}_n^2 \rangle = t_{n+1} - t_n = \Delta t_n,$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

and thus we get

$$\begin{aligned}\mathrm{M}_2 &= \sum_n \langle \mathrm{f}(s_n, \mathrm{B}_n)\mathrm{g}(s_n, \mathrm{B}_n)\rangle\langle(\mathrm{B}_{n+1} - \mathrm{B}_n)^2\rangle \\ &= \sum_{n=1}^{\infty} \langle \mathrm{f}(s_n, \mathrm{B}_n)\mathrm{g}(s_n, \mathrm{B}_n)\rangle \Delta t_n.\end{aligned}$$

The last relation tends for $\Delta t_n \to 0$ to (1.86).

(d) A generalization of the property (c) is given by

$$\mathrm{M}_3 \equiv \left\langle \int_0^a \mathrm{f}(s, \mathrm{B}_s)\mathrm{dB}_s \int_0^b \mathrm{g}(u, \mathrm{B}_u)\mathrm{dB}_u \right\rangle = \int_0^{a\wedge b} \langle \mathrm{f}(s, \mathrm{B}_s)\mathrm{g}(s, \mathrm{B}_s)\rangle \mathrm{d}s. \tag{1.87}$$

To prove (1.87) we must distinguish to subclasses (i) $b = a + c > a$ and (ii) $a = b + c > b; c > 0$. We consider only case (i), the proof for case (ii) is done by analogy. We derive from (1.86) and (1.87).

$$\begin{aligned}\mathrm{M}_3 &= \mathrm{M}_2 + \left\langle \int_0^a \mathrm{f}(s, \mathrm{B}_s)\mathrm{dB}_s \int_a^b \mathrm{g}(u, \mathrm{B}_u)\mathrm{dB}_u \right\rangle \\ &= \mathrm{M}_2 + \sum_n \sum_{m>n} \langle \mathrm{f}(s_n, \mathrm{B}_n)\mathrm{g}(s_m, \mathrm{B}_m)\Delta\mathrm{B}_n\Delta\mathrm{B}_m\rangle.\end{aligned}$$

But we see that $\mathrm{f}(s_n, \mathrm{B}_n)$ and $\Delta\mathrm{B}_n$ are independent of $\mathrm{f}(s_m, \mathrm{B}_m)$ and $\Delta\mathrm{B}_m$. Hence, we obtain

$$\mathrm{M}_3 = \mathrm{M}_2 + \sum_n \langle \mathrm{f}(s_n, \mathrm{B}_n)\Delta\mathrm{B}_n\rangle \sum_{m>n} \langle \mathrm{g}(s_m, \mathrm{B}_m)\Delta\mathrm{B}_m\rangle = \mathrm{M}_2,$$

where we use (1.85). This conclude the proof of (1.87) for case (i).

Now we calculate an **example**

$$\mathrm{I}(t) = \int_0^t \mathrm{B}_s\mathrm{dB}_s. \tag{1.88a}$$

First of all we obtain with the use of (1.85) and (1.86) the moments of the stochastic variable (1.88a)

$$\begin{gathered}\langle \mathrm{I}(t)\rangle = 0; \quad \langle \mathrm{I}(t)\mathrm{I}(t+\tau)\rangle = \int_0^\gamma \langle \mathrm{B}_s^2\rangle \mathrm{d}s = \int_0^\gamma s\mathrm{d}s = \gamma^2/2; \\ \gamma = t \wedge (t+\tau).\end{gathered} \tag{1.88b}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

We calculate the integral with an Ito decomposition

$$\mathrm{I} = \sum_k \mathrm{B}_k(\mathrm{B}_{k+1} - \mathrm{B}_k).$$

But we have

$$\begin{aligned}\Delta \mathrm{B}_k^2 &\equiv (\mathrm{B}_{k+1}^2 - \mathrm{B}_k^2) = (\mathrm{B}_{k+1} - \mathrm{B}_k)^2 + 2\mathrm{B}_k(\mathrm{B}_{k+1} - \mathrm{B}_k) \\ &= (\Delta \mathrm{B}_k)^2 + 2\mathrm{B}_k(\mathrm{B}_{k+1} - \mathrm{B}_k).\end{aligned}$$

Hence we obtain

$$\mathrm{I}(t) = \frac{1}{2}\sum_k \left[\Delta\left(\mathrm{B}_k^2\right) - \left(\Delta \mathrm{B}_k\right)^2\right].$$

We calculate now the two sums in (1.90) separately. Thus we obtain the first place

$$\begin{aligned}\mathrm{I}_1(t) &= \sum_k \Delta\left(\mathrm{B}_k^2\right) = \left(\mathrm{B}_1^2 - \mathrm{B}_0^2\right) + \left(\mathrm{B}_2^2 - \mathrm{B}_1^2\right) + \cdots + \left(\mathrm{B}_\mathrm{N}^2 - \mathrm{B}_{\mathrm{N}-1}^2\right) \\ &= \mathrm{B}_\mathrm{N}^2 \to \mathrm{B}_t^2,\end{aligned}$$

where we used $\mathrm{B}_0 = 0$. The second integral and its average are given by

$$\begin{aligned}\mathrm{I}_2(t) &= \sum_k \left(\Delta \mathrm{B}_k\right)^2 = \sum_k \left(\mathrm{B}_{k+1}^2 - 2\mathrm{B}_{k+1}\mathrm{B}_k + \mathrm{B}_k^2\right); \\ \langle \mathrm{I}_2(t)\rangle &= \sum_k \Delta t_k = t.\end{aligned}$$

The relation $\langle \mathrm{I}_2(t)\rangle = t$ gives not only the average but also the integral $\mathrm{I}_2(t)$ itself. However, the direct calculation of $\mathrm{I}_2(t)$ is impractical and we refer the reader to the book of Øksendahl [1.8], where the corresponding algebra is performed. We use instead an indirect proof and show that the quantity z (the standard deviation of $\mathrm{I}_2(t)$) is a deterministic function with the value zero. Thus, we put $z = \mathrm{I}_2(t) - t$. The mean value is clearly $\langle z\rangle = 0$ and we obtain

$$\langle z^2\rangle = \langle \mathrm{I}_2^2(t) - 2t\mathrm{I}_2(t) + t^2\rangle = \langle \mathrm{I}_2^2(t)\rangle - t^2.$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

But we have

$$\langle \mathrm{I}_2^2(t)\rangle = \sum_k \sum_m \langle (\Delta \mathrm{B}_k)^2 (\Delta \mathrm{B}_m)^2\rangle. \tag{1.88c}$$

The independence of the increments of the WP's yields

$$\langle (\Delta \mathrm{B}_k)^2 (\Delta \mathrm{B}_m)^2\rangle = \langle (\Delta \mathrm{B}_k)^2\rangle \langle (\Delta \mathrm{B}_m)^2\rangle + \delta_{km}\langle \Delta \mathrm{B}_k^4\rangle,$$

hence we obtain with the use of the results of EX 1.6

$$\langle \mathrm{I}_2^2(t)\rangle = \left(\sum_k \langle (\Delta \mathrm{B}_k)^2\rangle\right)^2 + \sum_k \langle (\mathrm{B}_{k+1} + \mathrm{B}_k)^4\rangle = t^2 + \sum_k (t_{k+1} - t_k)^2.$$

However, we have

$$\sum_k (t_{k+1} - t_k)^2 = \sum_k (\Delta t)^2 = t\Delta t \to 0 \quad \text{for } \Delta t \to 0,$$

and this indicates that $\langle z^2\rangle = 0$.

This procedure can be pursued to higher orders and we obtain the result that all moments of z are zero and thus we obtain $\mathrm{I}_2(t) = t$.

Thus, we obtain finally

$$\mathrm{I}(t) = \int_0^t \mathrm{B}_s \mathrm{dB}_s = \frac{1}{2}\big(\mathrm{B}_t^2 - t\big). \tag{1.89}$$

There is a generalization of the previous results with respect to higher order moments. We consider here moments of a stochastic integral with a deterministic integrand

$$\mathrm{J}_k(t) = \int_0^t \mathrm{f}_k(s)\mathrm{dB}_s; \quad k \in \mathbf{N}. \tag{1.90}$$

These integrals are a special case of the ones in (1.82) and we know from (1.85) that the mean value of (1.90) is zero. The covariance of (1.90) is given by (see (1.86))

$$\langle \mathrm{J}_k(t)\mathrm{J}_m(t)\rangle = \int_0^t \mathrm{f}_k(s)\mathrm{f}_m(s)\mathrm{d}s.$$

But we can obtain formally the same result if we put

$$\langle \mathrm{dB}_s \mathrm{dB}_u\rangle = \delta(s-u)\mathrm{d}s\mathrm{d}u. \tag{1.91}$$

A formal justification of (1.91) is given in Chapter 2 in connection with formula (2.41). Here we show that (1.91) leads to a result that

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

is identical to the consequences of (1.86)

$$\langle \mathrm{J}_k(t)\mathrm{J}_m(t)\rangle = \int_0^t \mathrm{f}_k(s)\int_0^t \mathrm{f}_m(u)\langle \mathrm{dB}_s\mathrm{dB}_u\rangle$$
$$= \int_0^t \mathrm{f}_k(s)\int_0^t \mathrm{f}_m(u)\delta(s-u)\mathrm{d}s\mathrm{d}u = \int_0^t \mathrm{f}_k(s)\mathrm{f}_m(s)\mathrm{d}s.$$

We also know that B_t and hence dB_t are Gaussian and Markovian. This means that all odd moments of the integral (1.90) must vanish

$$\langle \mathrm{J}_k(t)\mathrm{J}_m(t)\mathrm{J}_r(t)\rangle = \cdots = 0. \tag{1.92a}$$

To calculate higher order moments we use the properties of the multivariate GD and we put for the 4th order moment of the differential

$$\langle \mathrm{dB}_p\mathrm{dB}_q\mathrm{dB}_u\mathrm{dB}_v\rangle = \langle \mathrm{dB}_p\mathrm{dB}_q\rangle\langle \mathrm{dB}_u\mathrm{dB}_v\rangle + \langle \mathrm{dB}_p\mathrm{dB}_u\rangle\langle \mathrm{dB}_q\mathrm{dB}_v\rangle$$
$$+ \langle \mathrm{dB}_p\mathrm{dB}_v\rangle\langle \mathrm{dB}_q\mathrm{dB}_u\rangle = [\delta(p-q)\delta(u-v)$$
$$+ \delta(p-u)\delta(q-v) + \delta(p-v)\delta(q-u)]\mathrm{d}p\mathrm{d}q\mathrm{d}u\mathrm{d}v.$$

Note that the 4th order moment of the differential of WP's has a form similar to an isotropic 4th order tensor. Hence, we obtain

$$\langle \mathrm{J}_j(t)\mathrm{J}_m(t)\mathrm{J}_r(t)\mathrm{J}_s(t)\rangle = \int_0^t \mathrm{f}_j(\alpha)\mathrm{f}_m(\alpha)\mathrm{d}\alpha \int_0^t \mathrm{f}_r(\beta)\mathrm{f}_s(\beta)\mathrm{d}\beta$$
$$+ \int_0^t \mathrm{f}_j(\alpha)\mathrm{f}_r(\alpha)\mathrm{d}\alpha \int_0^t \mathrm{f}_m(\beta)\mathrm{f}_s(\beta)\mathrm{d}\beta$$
$$+ \int_0^t \mathrm{f}_j(\alpha)\mathrm{f}_s(\alpha)\mathrm{d}\alpha \int_0^t \mathrm{f}_m(\beta)\mathrm{f}_r(\beta)\mathrm{d}\beta.$$

This leads in a special case to

$$\langle \mathrm{J}_k^4(t)\rangle = 3\langle \mathrm{J}_k^2(t)\rangle^2. \tag{1.92b}$$

Again, this procedure can be carried out also for higher order moments and we obtain

$$\left\langle \mathrm{J}_k^{2\mu+1}(t)\right\rangle = 0; \quad \left\langle \mathrm{J}_k^{2\mu}(t)\right\rangle = 1.3.\ldots(2\mu-1)\left\langle \mathrm{J}_k^2(t)\right\rangle^\mu; \quad \mu \in \mathbf{N}. \tag{1.92c}$$

Equation (1.92) signifies that the stochastic Ito-integral (1.90) with the deterministic integrand $\mathrm{f}_k(s)$ is $\mathrm{N}[0, \int_0^t \mathrm{f}_k^2(s)\mathrm{d}s]$ distributed. However, one can also show that the Ito-integral with the non-anticipative

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

integrand

$$\mathrm{K}(t) = \int_0^t \mathrm{g}(s, \mathrm{B}_s)\mathrm{dB}_s, \tag{1.93a}$$

is, in analogy to the stochastic integral with the deterministic integrand,

$$\mathrm{N}[0, \tau(t)]; \quad \tau(t) = \int_0^t \langle \mathrm{g}^2(u, \mathrm{B}_u)\rangle \mathrm{d}u, \tag{1.93b}$$

distributed (see Arnold [1.2]). The variable $\tau(t)$ is referred to as intrinsic time of the stochastic integral (1.93a). We use this variable to show with Kolmogorov's Theorem (1.20) that (1.93a) posses continuous SF. The Ito integral

$$x_k = \int_0^{t_k} \mathrm{g}(u, \mathrm{B}_u)\mathrm{dB}_u, \quad t_1 = t > t_2 = s,$$

with

$$\langle x_k\rangle = 0; \quad \left\langle x_k^2\right\rangle = \tau(t_k) = \tau_k; \quad k = 1, 2, \quad \langle x_1 x_2\rangle = \tau_2,$$

has according to (1.35a) the joint PD

$$\mathrm{p}_2(x_1, x_2) = [(2\pi)^2\tau_1(\tau_1 - \tau_2)]^{-1/2} \exp\left[-\frac{x_1}{2\tau_1} - \frac{(x_1 - x_2)^2}{2(\tau_1 - \tau_2)}\right].$$

Yet, the latter line is identical with the bivariate PD of the Wiener process (1.60) if we replace in the latter equation the t_k by τ_k. Hence, we obtain from Kolmogorov's criterion $\langle[x_1(\tau_1) - x_{21}(\tau_2)]^2\rangle = \mid \tau_1 - \tau_2 \mid$ and this guarantees the continuity of the SF of the Ito-integral (1.93a). A further important feature of Ito integrals is their martingale property. We verify this now for the case of the integral (1.89). To achieve this, we generalize the martingale formula (1.64) for the case of arbitrary functions of the Brownian motions

$$\langle \mathrm{f}(y_2, s) \mid \mathrm{f}(y_1, t)\rangle = \int \mathrm{f}(y_2, s)\mathrm{p}_{1|1}(y_2, s \mid y_1, t)\mathrm{d}y_2 = \mathrm{f}(y_1, t);$$
$$y_k = \mathrm{B}_{t_k}; \ \forall\ s > t, \tag{1.94}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

where $p_{1|1}$ is given by (1.53). To verify now the martingale property of the integral (1.89) we specify (1.94) to

$$\langle \mathrm{I}(y_2, s) \mid \mathrm{I}(y_1, t)\rangle = \frac{1}{\sqrt{\beta\pi}} \int \left(y_2^2 - s\right) \exp[-(y_2 - y_1)^2/\beta]\mathrm{d}y_2.$$

The application of the standard substitution (see EX 1.1) yields

$$\begin{aligned} \langle \mathrm{I}(y_2, s) \mid \mathrm{I}(y_1, t)\rangle &= \frac{1}{\sqrt{\beta\pi}} \int \left(y_1^2 - s + 2y_1 z\sqrt{\beta} + \beta z^2\right) \exp(-z^2)\mathrm{d}z \\ &= \frac{1}{2}\left(y_1^2 - s + \beta/2\right) = \mathrm{I}(y_1, t). \end{aligned} \tag{1.95}$$

This concludes the proof that the Ito integral (1.89) is a martingale. The general proof that all Ito integrals are martingales is given by Øksendahl [1.8]. However, we will encounter the martingale property for a particular class of Ito integrals in the next section.

To conclude this example we add here also the Stratonovich version of the integral (1.89). This yields (the subscript s indicates a Stratonovich integral)

$$\begin{aligned} \mathrm{I}_s(t) &= \int_0^t \mathrm{B}_s \mathrm{o}\, \mathrm{dB}_s = \frac{1}{2}\sum_k (\mathrm{B}_{k+1} + \mathrm{B}_k)(\mathrm{B}_{k+1} - \mathrm{B}_k) \\ &= \frac{1}{2}\sum_k^{\infty} \Delta\mathrm{B}_k^2 \to \frac{\mathrm{B}_t^2}{2}. \end{aligned} \tag{1.96}$$

The result (1.96) is the “classical” value of the integral whereas the Ito integral gives a non classical result. Note also the significant differences between the Ito and Stratonovich integrals. Even the moments do not coincide since we infer from (1.96)

$$\langle \mathrm{I}_s(t)\rangle = \frac{t}{2} \quad \text{and} \quad \langle \mathrm{I}_s(t)\mathrm{I}_s(u)\rangle = \frac{1}{4}[tu + 2(t \wedge u)^2].$$

It is now easy to show that the Stratonovich integral I_s is not a martingale. We obtain this result if we drop the term s in second line of (1.95)

$$\langle \mathrm{I}_s(y_2, s) \mid \mathrm{I}_s(y_1, t)\rangle = \frac{1}{2}\left(y_1^2 + \beta/2\right) \neq \mathrm{I}_s(y_1, t). \qquad ♣$$

Hence, we may summarize the properties of the Ito and Stratonovich integrals. The Stratonovich concept uses all the transformation rules of classical integration theory and thus leads in many

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

applications to an easy way of performing the integration. Deviating from the Ito integral, the Stratonovich integral does, however, not posses the effective rules to calculated averages such as (1.85) to (1.87) and they do not have the martingale property. In the following we will consider both integration concepts and their application in solution of SDE.

We have calculated so far only one stochastic integral and we continue in the next section with helpful rules perform the stochastic integration.

1.9. The Ito Formula

We begin with the differential of a function $\Phi(\mathrm{B}_t, t)$. Its Ito differential takes the form

$$\mathrm{d}\Phi(\mathrm{B}_t, t) = \Phi_t \mathrm{d}t + \Phi_{\mathrm{B}_t} \mathrm{dB}_t + \frac{1}{2}\Phi_{\mathrm{B}_t\mathrm{B}_t}(\mathrm{dB}_t)^2. \tag{1.97.1}$$

Formula (1.97.1) contains the non classical term that is proportional to the second derivative WRT B_t. We must supplement (1.97.1) by a further non classical relation

$$(\mathrm{dB}_t)^2 = \mathrm{d}t. \tag{1.97.2}$$

Thus, we infer from (1.97.1,2) the final form of this differential

$$\mathrm{d}\Phi(\mathrm{B}_t, t) = \left(\Phi_t + \frac{1}{2}\Phi_{\mathrm{B}_t\mathrm{B}_t}\right)\mathrm{d}t + \Phi_{\mathrm{B}_t}\mathrm{dB}_t. \tag{1.98}$$

Next we derive the Ito differential of the function $\mathrm{Y} = \mathrm{g}(x, t)$ where x is the solution of the SDE

$$\mathrm{d}x = a(x, t)\mathrm{d}t + b(x, t)\mathrm{dB}_t. \tag{1.99.1}$$

In analogy to (1.97.1) we include a non classical term and put

$$\mathrm{dY} = \mathrm{g}_t \mathrm{d}t + \mathrm{g}_x \mathrm{d}x + \frac{1}{2}\mathrm{g}_{xx}(\mathrm{d}x)^2,$$

We substitute $\mathrm{d}x$ from (1.99.1) into the last line and apply the non classical formula

$$(\mathrm{d}x)^2 = (a\mathrm{d}t + b\mathrm{dB}_t)^2 = b^2\mathrm{d}t; \quad (\mathrm{d}t)^2 = \mathrm{d}t\mathrm{dB}_t = 0; \quad (\mathrm{dB}_t)^2 = \mathrm{d}t, \tag{1.99.2}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

and this yields

$$\mathrm{dY} = \left(\mathrm{g}_t + a\mathrm{g}_x + \frac{b^2}{2}\mathrm{g}_{xx}\right)\mathrm{d}t + b\mathrm{g}_x\mathrm{dB}_t. \tag{1.99.3}$$

The latter equation is called the Ito formula for the total differential of function $\mathrm{Y} = \mathrm{g}(x,t)$ given the SDE (1.99.1). (1.99.3) contains the non classical term $b^2\mathrm{g}_{xx}/2$ and it differs thus from the classical (or Stratonovich) total differential

$$\mathrm{dY}_c = (\mathrm{g}_t + a\mathrm{g}_x)\mathrm{d}t + b\mathrm{g}_x\mathrm{dB}_t. \tag{1.100}$$

Note that both the Ito and the Stratonovich differentials coincide if $\mathrm{g}(x,t)$ is a first order polynomial of the variable x.

We postpone a sketch of the proof of (1.99) for a moment and give an example of the application of this formula. We use (1.99.1) in the form

$$\mathrm{d}x = \mathrm{dB}_t, \quad \text{or} \quad x = \mathrm{B}_t \quad \text{with } a = 0, \quad b = 1, \tag{1.101a}$$

and we consider the function

$$\mathrm{Y} = \mathrm{g}(x) = x^2/2; \quad \mathrm{g}_t = 0; \quad \mathrm{g}_x = x; \quad \mathrm{g}_{xx} = 1. \tag{1.101b}$$

Thus we obtain from (1.99.3) and (1.101b)

$$\mathrm{dY} = \mathrm{d}(x^2/2) = \mathrm{d}t/2 + \mathrm{B}_t\mathrm{dB}_t,$$

and the integration of this total differential yields

$$\int_0^t \mathrm{d}(x^2/2) = \int_0^t \mathrm{d}(\mathrm{B}_s^2/2) = \mathrm{B}_t^2/2 = t/2 + \int_0^t \mathrm{B}_s\mathrm{dB}_s$$

and the last line reproduces (1.89). ♣

We give now a sketch of the proof of the Ito formula (1.99) and we follow in part considerations of Schuss [1.9]. It is instructive to perform this in detail and we do it in four consecutive steps labeled with S_1 to S_4.

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

$\underline{\mathrm{S}_1}$

We begin with the consideration of the stochastic function $x(t)$ given by

$$x(v) - x(u) = \int_u^v a(x(s), s)\mathrm{d}s + \int_u^v b(x(s), s)\mathrm{dB}_s, \tag{1.102}$$

where a and b are two differentiable functions. Thus, we obtain the differential of $x(t)$ if we put in (1.102) $v = u + \mathrm{d}t$ and let $\mathrm{d}t \to 0$

$$\mathrm{d}x(u) = a(x(u), u)\mathrm{d}u + b(x(u), u)\mathrm{dB}_u. \tag{1.103}$$

Before we pass to the next step we consider two important examples

Example 1. (integration by parts)
Here we consider a deterministic function f and a stochastic function Y and we put

$$\mathrm{Y}(\mathrm{B}_t, t) = \mathrm{g}(\mathrm{B}_t, t) = \mathrm{f}(t)\mathrm{B}_t. \tag{1.104a}$$

The total differential is in both (Ito and Stratonovich) cases (see (1.98) with $\Phi_{\mathrm{B}_t\mathrm{B}_t} = 0$) given by the exact formula

$$\mathrm{dY} = \mathrm{d}[\mathrm{f}(t)\mathrm{B}_t] = \mathrm{f}(t)\mathrm{dB}_t + \mathrm{f}'(t)\mathrm{B}_t\mathrm{d}t. \tag{1.104b}$$

The integration of this differential yields

$$\mathrm{f}(t)\mathrm{B}_t = \int_0^t \mathrm{f}'(s)\mathrm{B}_s\mathrm{d}s + \int_0^t \mathrm{f}(s)\mathrm{dB}_s. \tag{1.105a}$$

Subtracting the last line for $t = u$ from the same relation for $t = v$ yields

$$\mathrm{f}(v)\mathrm{B}_v - \mathrm{f}(u)\mathrm{B}_u = \int_u^v f'(s)\mathrm{B}_s\mathrm{d}s + \int_u^v \mathrm{f}(s)\mathrm{dB}s. \tag{1.105b}$$

♣

Example 2. (Martingale property)
We consider a particular class of Ito integrals

$$\mathrm{I}(t) = \int_0^t \mathrm{f}(u)\mathrm{dB}_u, \tag{1.106}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

and show that $\mathrm{I}(t)$ is a martingale. First we realize that the integral $\mathrm{I}(t)$ is a particular case of the class (1.93a) with $\mathrm{g}(u, B_u) = \mathrm{f}(u)$. Hence we know that the variable (1.106) is normal distributed and posses the intrinsic time given by (1.93b). Its transition probability $\mathrm{p}_{1|1}$ is defined by (1.53) with $t_j = \tau(t_j)$; $y_j = \mathrm{I}(t_j)$; $j = 1, 2$. This concludes the proof that the integral (1.106) obeys a martingale property like (1.27) or (1.64). ♣

S₂

Here we consider the product of two stochastic functions subjected to two SDE with constant coefficients

$$\mathrm{d}x_k(t) = a_k \mathrm{d}t + b_k \mathrm{dB}_t; \quad a_k, b_k = \text{const}; \;\; k = 1, 2, \tag{1.107}$$

with the solutions

$$x_k(t) = a_k t + b_k \mathrm{B}_t; \quad x_k(0) = 0. \tag{1.108}$$

The task to evaluate $\mathrm{d}(x_1 x_2)$ is outlined in EX 1.9 and we obtain with the aid of (1.89)

$$\mathrm{d}(x_1 x_2) = x_2 \mathrm{d}x_1 + x_1 \mathrm{d}x_2 + b_1 b_2 \mathrm{d}t. \tag{1.109}$$

The term proportional to $b_1 b_2$ in (1.109) is non classical and it is a mere consequence of the non classical term in (1.89).

The relation (1.109) was derived for constant coefficients in (1.107). One may derive (1.109) under the assumption of step-function for the functions a and b in (1.106) and with that one can approximate differentiable functions (see Schuss [1.9]).

We consider now two examples

Example 1

We take put $x_1 = \mathrm{B}_t, x_2 = \mathrm{B}_t^2$. Thus, we obtain with an application of (1.101b) and (1.109)

$$\mathrm{dB}_t^3 = \mathrm{B}_t \mathrm{dB}_t^2 + \mathrm{B}_t^2 \mathrm{dB}_t + 2\mathrm{B}_t \mathrm{d}t = 3(\mathrm{B}_t \mathrm{d}t + \mathrm{B}_t^2 \mathrm{dB}_t).$$

The use of the induction rule yields the generalization

$$\mathrm{dB}_t^k = k\mathrm{B}_t^{k-1} \mathrm{dB}_t + \frac{k(k-1)}{2} \mathrm{B}_t^{k-2} \mathrm{d}t \tag{1.110}$$

♣

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

Example 2

Here we consider polynomials of the Brownian motion

$$\mathrm{P}_n(\mathrm{B}_t) = c_0 + c_1\mathrm{B}_t + \cdots + c_n\mathrm{B}_t^n; \quad c_k = \text{const.} \tag{1.111}$$

The application of (1.110) to (1.111) leads to

$$\mathrm{dP}_n(\mathrm{B}_t) = \mathrm{P}'_n(\mathrm{B}_t)\mathrm{dB}_t + \frac{1}{2}\mathrm{P}''_n(\mathrm{B}_t)\mathrm{d}t; \quad {}' = \partial/\partial\mathrm{B}_t. \tag{1.112}$$

The relation (1.112) is also valid for all functions that can be expanded in form of polynomials. ♣

$\underline{\mathrm{S}_3}$

Here we consider the product

$$\Phi(\mathrm{B}_t, t) = \varphi(\mathrm{B}_t)\mathrm{g}(t), \tag{1.113}$$

where g is a deterministic function. The use of (1.109) yields

$$\begin{aligned}
\mathrm{d}\Phi(\mathrm{B}_t, t) &= \mathrm{g}(t)\mathrm{d}\varphi(\mathrm{B}_t) + \varphi(\mathrm{B}_t)\mathrm{g}'(t)\mathrm{d}t \\
&= \left(\varphi'\mathrm{d}\,\mathrm{B}_t + \frac{1}{2}\varphi''\mathrm{d}t\right)\mathrm{g} + \varphi\mathrm{g}'(t)\mathrm{d}t \qquad (1.114) \\
&= \left(\varphi\mathrm{g}' + \frac{1}{2}\varphi''\mathrm{g}\right)\mathrm{d}t + \mathrm{g}\varphi'\mathrm{dB}_t.
\end{aligned}$$

But we also have

$$\left(\frac{\partial}{\partial t} + \frac{1}{2}\frac{\partial^2}{\partial\mathrm{B}_t^2}\right)\Phi = \mathrm{g}'\varphi + \frac{1}{2}\mathrm{g}\varphi''. \tag{1.115}$$

Thus, we obtain

$$\mathrm{d}\Phi = \left(\frac{\partial}{\partial t} + \frac{1}{2}\frac{\partial^2}{\partial\mathrm{B}_t^2}\right)\Phi\mathrm{d}t + \frac{\partial\Phi}{\partial\mathrm{B}_t}\mathrm{dB}_t. \tag{1.116}$$

Equation (1.116) applies, in the first place, only to the function (1.113). However, the use of the expansion

$$\Phi(\mathrm{B}_t, t) = \sum_{k=1}^{\infty}\varphi_k(\mathrm{B}_t)\mathrm{g}_k(t), \tag{1.117}$$

shows that (1.116) is valid for arbitrary functions and this proves (1.98).

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

$\underline{S_4}$

In this last step we do not apply the separation (1.113) or (1.117) but we use a differentiable function of the variables (x, t), where x satisfies a SDE of the type (1.107)

$$\begin{gathered}\Phi(\mathrm{B}_t, t) \equiv \mathrm{g}(x, t) = \mathrm{g}(at + b\mathrm{B}_t, t); \quad x = a\mathrm{d}t + b\mathrm{dB}_t; \quad a, b = \mathrm{const.} \\ \Phi_t = a\mathrm{g}_x + \mathrm{g}_t; \quad \Phi_{\mathrm{B}_t} = b\mathrm{g}_x; \quad \Phi_{\mathrm{B}_t\mathrm{B}_t} = b^2\mathrm{g}_{xx}.\end{gathered} \tag{1.118}$$

Thus we obtain with (1.116)

$$\mathrm{dg} = \left(\frac{\partial \mathrm{g}}{\partial t} + a\frac{\partial \mathrm{g}}{\partial x} + \frac{b^2}{2}\frac{\partial^2 \mathrm{g}}{\partial x^2}\right)\mathrm{d}t + b\frac{\partial \mathrm{g}}{\partial x}\,\mathrm{dB}_t. \tag{1.119}$$

The relation (1.119) represents the Ito formula (1.99.3) (for constant coefficients a and b). As before, we can generalize the proof and (1.119) is valid for arbitrary coefficients $a(x, t)$ and $b(x, t)$.

We generalize now the Ito formula for the case of a **multivariate process**. First we consider K functions of the type

$$y_k = y_k\big(\mathrm{B}_t^1, \ldots, \mathrm{B}_t^{\mathrm{M}}, t\big); \quad k = 1, 2, \ldots, \mathrm{K},$$

where $\mathrm{B}_t^1, \ldots, \mathrm{B}_t^{\mathrm{M}}$ are M independent Brownian motions. We take advantage of the summation convention and obtain the generalization of (1.97.1)

$$\begin{gathered}\mathrm{d}y_k\big(\mathrm{B}_t^1, \ldots, \mathrm{B}_t^{\mathrm{M}}, t\big) = \frac{\partial y_k}{\partial t}\mathrm{d}t + \frac{\partial y_k}{\partial \mathrm{B}_t^r}\mathrm{dB}_t^r + \frac{1}{2}\frac{\partial^2 y_k}{\partial \mathrm{B}_t^r \mathrm{B}_t^s}\mathrm{dB}_t^r\mathrm{dB}_t^s; \\ k = 1, \ldots, \mathrm{K}; \quad r, s = 1, \ldots, \mathrm{M}.\end{gathered} \tag{1.120}$$

We generalize (1.97.2) and put

$$\mathrm{dB}_t^r\mathrm{dB}_t^s = \delta_{rs}\mathrm{d}t, \tag{1.121}$$

and we obtain (see (1.98))

$$\mathrm{d}y_k\big(\mathrm{B}_t^1, \ldots, \mathrm{B}_t^{\mathrm{M}}, t\big) = \left(\frac{\partial y_k}{\partial t} + \frac{1}{2}\frac{\partial^2 y_k}{\partial \mathrm{B}_t^r \mathrm{B}_t^r}\right)\mathrm{d}t + \frac{\partial y_k}{\partial \mathrm{B}_t^r}\mathrm{dB}_t^r. \tag{1.122}$$

Now we consider a set of n SDEs

$$\begin{gathered}\mathrm{dX}_k = a_k(\mathrm{X}_1, \ldots, \mathrm{X}_n, t)\mathrm{d}t + b_{kr}(\mathrm{X}_1, \ldots, \mathrm{X}_n, t)\mathrm{dB}_t^r; \\ k = 1, 2, \ldots, n; \quad r = 1, 2, \ldots, \mathrm{R}.\end{gathered} \tag{1.123}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

We wish to calculate the differential of the function

$$\mathrm{Z}_k = \mathrm{Z}_k(\mathrm{X}_1, \ldots, \mathrm{X}_n, t); \quad k = 1, \ldots, \mathrm{K}. \tag{1.124}$$

The differential reads

$$\begin{aligned}
\mathrm{dZ}_k &= \frac{\partial \mathrm{Z}_k}{\partial t}\,\mathrm{d}t + \frac{\partial \mathrm{Z}_k}{\partial \mathrm{X}_m}\mathrm{dX}_m + \frac{1}{2}\frac{\partial^2 \mathrm{Z}_k}{\partial \mathrm{X}_m \partial \mathrm{X}_u}\mathrm{dX}_m \mathrm{dX}_u \\
&= \frac{\partial \mathrm{Z}_k}{\partial t}\mathrm{d}t + \frac{\partial \mathrm{Z}_k}{\partial \mathrm{X}_m}\big(a_m \mathrm{d}t + b_{mr}\mathrm{dB}_t^r\big) \\
&\quad + \frac{1}{2}\frac{\partial^2 \mathrm{Z}_k}{\partial \mathrm{X}_m \partial \mathrm{X}_u}\big(a_m \mathrm{d}t + b_{mr}\mathrm{dB}_t^r\big)\big(a_u \mathrm{d}t + b_{us}\mathrm{dB}_t^s\big); \\
&\quad m, n = 1, 2, \ldots, n; \;\; r = 1, 2, \ldots, \mathrm{R}.
\end{aligned} \tag{1.125}$$

The n-dimensional generalization of the rule (1.99.2) is given by

$$\mathrm{d}\big(\mathrm{B}_t^r \mathrm{dB}_t^u\big) = \delta_{ru}\mathrm{d}t; \quad (\mathrm{d}t)^2 = \mathrm{dB}_t^r\,\mathrm{d}t = 0. \tag{1.126}$$

Thus, we obtain the differential of the vector valued function (1.124)

$$\begin{aligned}
\mathrm{dZ}_k &= \left(\frac{\partial \mathrm{Z}_k}{\partial t} + a_m \frac{\partial \mathrm{Z}_k}{\partial \mathrm{X}_m} + \frac{1}{2} b_{mr} b_{ur} \frac{\partial^2 \mathrm{Z}_k}{\partial \mathrm{X}_m \partial \mathrm{X}_u}\right)\mathrm{d}t \\
&\qquad + b_{mr}\frac{\partial \mathrm{Z}_k}{\partial \mathrm{X}_m}\,\mathrm{dB}_t^r.
\end{aligned} \tag{1.127}$$

Now we conclude this section with two examples.

Example 1

A stochastic process is given by

$$\mathrm{Y}_1 = \mathrm{B}_t^1 + \mathrm{B}_t^2 + \mathrm{B}_t^3; \quad \mathrm{Y}_2 = \big(\mathrm{B}_t^2\big)^2 - \mathrm{B}_t^1 \mathrm{B}_t^3.$$

We obtain for the SDE in the form (1.120) corresponding to the last line

$$\begin{aligned}
\mathrm{dY}_1 &= \mathrm{dB}_t^1 + \mathrm{dB}_t^2 + \mathrm{dB}_t^3; \\
\mathrm{dY}_2 &= \mathrm{d}t + 2\mathrm{B}_t^2 \mathrm{dB}_t^2 - \big(\mathrm{B}_t^3 \mathrm{dB}_t^1 + \mathrm{B}_t^1 \mathrm{dB}_t^3\big).
\end{aligned} \qquad \clubsuit$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

Example 2
Here we study a single stochastic process under the influence of two independent Brownian motions

$$\mathrm{d}x = a(x,t)\mathrm{d}t + \beta(x,t)\mathrm{dB}_t^1 + \gamma(x,t)\mathrm{dB}_t^2. \tag{1.128}$$

The differential of the function $\mathrm{Y} = \mathrm{g}(x,t)$ has the form

$$\mathrm{dY} = \left[\mathrm{g}_t + a\mathrm{g}_x + \frac{1}{2}(\beta^2+\gamma^2)\frac{\partial^2 \mathrm{g}}{\partial x^2}\right]\mathrm{d}t + \mathrm{g}_x\left(\beta\mathrm{dB}_t^1 + \gamma\mathrm{dB}_t^2\right).$$

We consider now the special case

$$\mathrm{g} = \ln x; \quad a = rx; \quad \beta = ux; \quad \gamma = \sigma x; \quad r, u, \sigma = \text{const},$$

and we obtain

$$\mathrm{d}(\ln x) = [r - (u^2+\sigma^2)/2]\mathrm{d}t + \left(u\mathrm{dB}_t^1 + \sigma\mathrm{dB}_t^2\right). \tag{1.129}$$

We will use (1.129) in Section 2.1 of the next chapter. ♣

We introduced in this chapter some elements of the probability theory and added the basic ideas about SDE. For readers who wish to get more deeply involved in the abstract theory of probability and in particular with the measure theory we suggest they consider the following books: Chung & Aitsahia [1.10], Ross [1.11], Mallivan [1.12], Pitman [1.13] and Shiryaev [1.14].

Appendix: Poisson Processes

In many applications appears there a random set of countable points driven by some stochastic system. Typical examples are arrival times of customers (at the desk of an office, at the gate of an airport, etc.), the birth process of an organism, the number of competing building projects for a state budget. The randomness in such phenomena is conveniently described by Poisson distributed variables.

First we verify that the Poisson distribution is the limit of the Bernoulli distribution. We substitute for the argument p in the Bernoulli distribution in Section 1.1 the value $p = \alpha/n$ and this

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

yields

$$b(k,n,\alpha/n) = \binom{n}{k}\left(\frac{\alpha}{n}\right)^k\left(1-\frac{\alpha}{n}\right)^{n-k};$$
$$b(0,n,\alpha/n) = \left(1-\frac{\alpha}{n}\right)^n \to \exp(-\alpha) \quad \text{for } n \to \infty. \tag{A.1}$$

Now we put

$$\frac{b(k+1,n,\alpha/n)}{b(k,n,\alpha/n)} = \frac{\alpha}{k+1}\frac{n-k}{n}\left(1-\frac{\alpha}{n}\right)^{-1} \to \frac{\alpha}{k+1},$$

and this yields

$$b(1,n,\alpha/n) \to \alpha\exp(-\alpha); \quad b(2,n,\alpha/n) \to \frac{\alpha^2}{2}\exp(-\alpha); \dots$$
$$b(k,n,\alpha/n) \to \frac{\alpha^n}{n!}\exp(-\alpha) = \pi_k(\alpha).$$

Definition. (Homogeneous Poisson process (HPP))
A random point process $\mathrm{N}(t)$, $t \geq 0$ on the real axis is a **HPP** with a constant intensity λ if it satisfies the three conditions

(a) $\mathrm{N}(0) = 0$.
(b) The random increments $\mathrm{N}(t_k) - \mathrm{N}(t_{k-1})$; $k = 1, 2, \dots$ are for any sequence of times $0 \leq t_0 < t_1 < \cdots < t_n < \cdots$ mutually independent.
(c) The random increments defined in condition (b) are Poisson distributed of the form

$$\Pr([\mathrm{N}(t_{r+1}) - \mathrm{N}(t_r)] = k) = \frac{(\lambda\tau_r)^k\exp(-\lambda\tau_r)}{k!};$$
$$\tau_r = t_{r+1} - t_r, \;\; k = 0, 1, \dots; \;\; r = 1, 2, \dots. \tag{A.2}$$

□

To analyze the sample paths we consider the increment $\Delta\mathrm{N}(t) = \mathrm{N}(t+\Delta t) - \mathrm{N}(t)$. Its probability has, for small values of Δt, the form

$$\Pr(\Delta\mathrm{N}(t) = k) = \frac{(\lambda\Delta t)^k}{k!}\exp(-\lambda\Delta t) \to \begin{bmatrix} 1-\lambda\Delta t \text{ for } k=0 \\ \lambda\Delta t \quad \text{for } k=1 \\ \mathrm{O}(\Delta t^2) \text{ for } k \geq 2 \end{bmatrix}. \tag{A.3}$$

Equation (A.3) means that for $\Delta t \to 0$ the probability that $\mathrm{N}(t+\Delta t)$ is most likely the one of $\mathrm{N}(t)$ $(\Pr([\mathrm{N}(t+\Delta t) - \mathrm{N}(t)] = 0) \approx 1)$.

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

However, the part of (A.3) with $\Pr([\mathrm{N}(t+\Delta t)-\mathrm{N}(t)]=1)\approx\lambda\Delta t$ indicates that there is small chance for a jump with the height unity. The probability of jumps with higher heights $k=2,3,\ldots$ corresponding to the third part of (A.3) is subdominantly small and such jumps do not appear.

We calculate of the moments of the HPP in two alternative ways. (i) We use (1.5) with (A.2) to obtain

$$\begin{aligned}\langle x^m\rangle &= \int \mathrm{p}(x)x^m\mathrm{d}x = \sum_{k=0}^{\infty} k^m \Pr(x=k)\\ &= \exp(-\alpha)\sum_{k=0}^{\infty} k^m\alpha^k/k!; \quad \alpha=\lambda t,\end{aligned} \tag{A.4}$$

or we apply (ii) the concept of the generating function defined by

$$\begin{aligned}&\mathrm{g}(z)\equiv\sum_{k=0}^{\infty} z^k \Pr(x=k), \quad \text{with } \mathrm{g}'(1)=\langle x\rangle,\\ &\mathrm{g}''(1)=\langle x^2\rangle-\langle x\rangle,\ldots;\;\; \mathrm{g}'=\frac{\mathrm{dg}}{\mathrm{d}z}.\end{aligned} \tag{A.5}$$

This leads in the case of an HPP to

$$\begin{aligned}\mathrm{g}(z) &= \sum_{k=0}^{\infty} z^k\alpha^k\exp(-\alpha)/k! = \exp(-\alpha)\sum_{k=0}^{\infty}(z\alpha)^k/k!\\ &= \exp[\alpha(z-1)].\end{aligned} \tag{A.6}$$

In either case we obtain

$$\langle \mathrm{N}(t)\rangle=\lambda t, \quad \langle \mathrm{N}^2(t)\rangle=(\lambda t)^2+\lambda t. \tag{A.7}$$

We calculate now the PD of the sum x_1+x_2 of two independent HPP's. By definition this yields

$$\begin{aligned}\Pr([x_1+x_2]=k) &= \Pr\left(\sum_{j=0}^{k}[x_1=j, x_2=k-j]\right)\\ &= \sum_{j=0}^{k}\Pr(x_1=j, x_2=k-j)\\ &= \sum_{j=0}^{k}\Pr(x_1=j)\Pr(x_2=k-j)\end{aligned}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

$$= \sum_{j=0}^{k} \exp[-(\theta_1)]\frac{\theta_1^j}{j!}\exp[-(\theta_2)]\frac{\theta_2^{k-j}}{(k-j)!}$$

$$= \exp[-(\theta_1+\theta_2)]\sum_{j=0}^{k}\theta_1^j\theta_2^{k-j}\binom{k}{j}\Big/ k!$$

$$= \exp[-(\theta_1+\theta_2)](\theta_1+\theta_2)^k/k!. \tag{A.8}$$

If the two variables are IID ($\theta = \theta_1 = \theta_2$) (A.8) reduces to

$$\Pr([x_1+x_2]=k) = \exp(-2\theta)(2\theta)^k/k!. \tag{A.9}$$

Poisson HPP's play important roles in Markov process (see Brémaud [1.16]). In many applications these Markov chains are iterations driven by "white noise" modeled by HPP's. Such iterations arise in the study of the stability of continuous periodic phenomena, in the biology and economics, etc. We consider the form of iterations

$$x(t+s) = \mathrm{F}(x(s), \mathrm{Z}(t+s)); \quad s,t \in \mathrm{N}_0 \tag{A.10}$$

where t, s are discrete variables and $x(t)$ is a discrete random variable driven by the white noise $\mathrm{Z}(t+s)$. An important particular case is $\mathrm{Z}(t+s) := \mathrm{N}(t+s)$ with a PD

$$\Pr(\mathrm{N}(t+s)=k) = \exp(u)u^k/k!; \quad u = \theta(t+s).$$

The **transition probability** is the matrix governing the transition from state i to state k.

Examples

(i) Random walk
This is an iteration of a discrete random variable $x(t)$

$$x(t) = x(t-1) + \mathrm{N}(t); \quad x(0) = x_0 \in \mathrm{N}. \tag{A.11}$$

$\mathrm{N}(t)$ is HPP with $\Pr([\mathrm{N}(t)=k]) = \exp(-\lambda t)(\lambda t)^k/k!$. Hence, we obtain the transition probability

$$\begin{aligned}\mathrm{p}_{ji} &= \Pr(x(t)=j, x(t-1)=i) = \Pr([i+\mathrm{N}(j)]=j)\\ &= \Pr(\mathrm{N}(j)=j-1).\end{aligned}$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

(ii) Flip-Flop processes
The iteration takes here the form

$$x(t) = (-1)^{\mathrm{N}(t)}. \tag{A.12}$$

The transition matrix takes the form

$$\mathrm{p}_{-1,1} = \Pr(x(t+s) = 1 \mid x(s) = -1) = \Pr(\mathrm{N}(t) = 2k+1) = \alpha;$$
$$\mathrm{p}_{1,1} = \Pr(x(t+s) = 1 \mid x(s) = 1 = \Pr(\mathrm{N}(t) = 2k) = \beta,$$

with

$$\alpha = \sum_{k=0}^{\infty} \exp(-\lambda t)(\lambda t)^{2k+1}/(2k+1)! = \exp(-\lambda t)\sinh(\lambda t);$$
$$\beta = \sum_{k=0}^{\infty} \exp(-\lambda t)(\lambda t)^{2k}/(2k)! = \exp(-\lambda t)\cosh(\lambda t). \qquad \clubsuit$$

Another important application of HPP is given by a 1D approach to turbulence elaborated by Kerstein [1.17] and [1.18]. This model is based on the turbulence advection by a random map. A triplet map is applied to a shear flow velocity profile. An individual event is represented by a mapping that results in a new velocity profile. As a statistical hypothesis the author assumes that the temporal rate of the event is governed by a Poisson process and the parameter of the map can be sampled from a given PD. Although this model was applied to 1D turbulence, its results go beyond this limit and the model has a remarkable power of prediction experimental data.

Exercises

EX 1.1. Calculate the mean value $\mathrm{M}_n(s,t) = \langle(\mathrm{B}_t - \mathrm{B}_s)^n\rangle$, $n \in \mathbf{N}$. Hint: Use (1.60) and the standard substitution $y_2 = y_1 + z\sqrt{2(t_2 - t_1)}$, where z is a new variable. Show that this yields

$$\mathrm{M}_n = \frac{[2(t_2 - t_1)]^{n/2}}{\sqrt{\pi}} \int \exp(-v^2)\mathrm{d}v \int \exp(-z^2)z^n \mathrm{d}z.$$

The gamma function is defined by (see Ryshik & Gradstein [1.15])

$$\int \exp(-z^2)z^n \mathrm{d}v = \begin{cases} \Gamma((n+1)/2)\,\forall\, n = 2k; \\ 0\,\forall\, n = 2k+1; \quad k \in \mathbf{N}, \end{cases}$$

$$\Gamma(1/2) = \sqrt{\pi}, \quad \Gamma(n+1) = n\Gamma(n).$$

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

Verify the result

$$\mathrm{M}_{2n} = \pi^{-1/2}[2(t_2 - t_1)]^n \Gamma((2n+1)/2).$$

EX 1.2. We consider a 1D random variable X with the mean μ and the variance σ^2. Show that the latter can be written in the form ($\mathrm{f_X}(x)$ is the PD $\mu = \langle x \rangle; \varepsilon > 0$)

$$\sigma^2 \geq \left(\int_{\mu+\varepsilon}^{\infty} + \int_{-\infty}^{\mu-\varepsilon} \right) \mathrm{f_X}(x)(x-\mu)^2 \mathrm{d}x; \quad \varepsilon.$$

For $x \leq \mu - \varepsilon$ and $x \geq \mu + \varepsilon \Rightarrow (x-\mu)^2 \geq \varepsilon^2$, this yields

$$\sigma^2 \geq \varepsilon^2 \left[1 - \int_{\mu-\varepsilon}^{\mu+\varepsilon} \mathrm{f_X}(x) \mathrm{d}x \right] = \varepsilon^2 \,\mathrm{Pr}(|\mathrm{X} - \mu| \geq \varepsilon),$$

and this gives the **Chebyshev inequality** its final form

$$\mathrm{Pr}\{|\mathrm{X} - \mu| \geq \varepsilon\} \leq \sigma^2/\varepsilon^2.$$

The inequality governing martingales (1.28) is obtained with considerations similar to the derivation of the Chebyshev inequality.

EX 1.3.

(a) Show that we can factorize the bivariate GD (1.35a) with zero mean and equal variance ($\langle x \rangle = \langle y \rangle = 0; \sigma^2 = a = b$) in the form

$$\mathrm{p}(x,y) = \gamma^{-1/2}\mathrm{p}(x)\mathrm{p}((y - rx)/\sqrt{\gamma}); \quad \gamma = (1 - r^2),$$

where $\mathrm{p}(x)$ is the univariate GD (1.29).

(b) Calculate the conditional distribution (see 1.17) of the bivariate GD (1.35a). Hint: (c is the covariance matrix)

$$\mathrm{p}_{1|1}(x \mid y) = \sqrt{c_{yy}/(2\pi \mathrm{D})}\, \exp\left[-c_{yy}(x - yc_{xy}/c_{yy})/(2\mathrm{D})\right];$$

$$c = \begin{pmatrix} \langle x^2 \rangle & \langle xy \rangle \\ \langle xy \rangle & \langle y^2 \rangle \end{pmatrix}; \quad \mathrm{D} = \det(c).$$

Verify that the latter line corresponds to a $\mathrm{N}[yc_{xy}/c_{yy}, c_{xx} - c_{xy}^2/c_{yy}]$ distribution.

EX 1.4. Prove that (1.53) is a solution of the Chapman–Kolmogorov equation (1.52)

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

Hint: The integrand in (1.52) is given by

$$T = p_{1|1}(y_2, t_2 \mid y_1, t_1) p_{1|1}(y_3, t_3 \mid y_2, t_2).$$

Use the substitution

$$u = t_2 - t_1 > 0, \quad v = t_3 - t_2 > 0; \quad t_3 - t_1 = v + u > 0,$$

introduce (1.53) into (1.52) and put

$$\begin{aligned} &T = (4\pi^2 uv)^{-1/2} \exp(-\lambda); \\ &\lambda = (y_3 - y_2)^2/(2v) + (y_2 - y_1)^2/(2u) = a_2 y_2^2 + a_1 y_2^1 + a_0, \end{aligned}$$

with $a_k = a_k(y_1, y_3), k = 1, 2, 3$. Use the standard substitution (see EX 1.1) to obtain

$$\int T dy_2 = (4\pi uv)^{-1/2} \exp[-F(y_3, y_2)] \int \exp(-K) dy_2;$$

$$F = \frac{4a_0 a_2 - a_1^2}{4a_2}; \quad K = a_2 \left(y_2 + \frac{a_1}{2a_2} \right)^2,$$

and compare the result of the integration with the right hand side of (1.52).

EX 1.5. Verify that the solution of (1.54) is given by (1.55). Prove also its initial condition.

Hint: To verify the initial condition use the integral

$$g = (2\pi t)^{-1/2} \lim_{t \to 0_+} \int_{-\infty}^{\infty} \exp[-y^2/(2t)] H(y) dy,$$

where $H(y)$ is a continuous function. Use the standard substitution in its form $y = \sqrt{2t} z$.

To verify the solution (1.55) use the same substitution as in EX 1.4.

EX 1.6. Calculate the average $\langle y_1^m(t_1) y_2^n(t_2) \rangle$; $y_k = B_{t_k}$, $k = 1, 2$; $n, m \in N$ with the use of the Markovian bivariate PD (1.60).

Hint: Use standard substitution of the type given in EX 1.2.

EX 1.7. Verify that the variable $\hat{B}_t$ defined in (1.65) has the autocorrelation $\langle \hat{B}_t \hat{B}_s \rangle = t \wedge s$. To perform this task we calculate for a

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

fixed value of $a > 0$

$$\langle \hat{\mathrm{B}}_t \hat{\mathrm{B}}_s \rangle = \langle \mathrm{B}_{t+a}\mathrm{B}_{s+a} \rangle - \langle \mathrm{B}_{t+a}\mathrm{B}_a \rangle - \langle \mathrm{B}_{s+a}\mathrm{B}_a \rangle + \langle \mathrm{B}_a^2 \rangle$$
$$= s \wedge t + a - a - a + a = s \wedge t.$$

EX 1.8. Prove that the scaled and translated WS's defined in (1.74) are WS's.

Hint: To cover the scaled WS's, put

$$\mathrm{H}_{u,v} = \frac{1}{ab}\mathrm{M}^{(2)}_{a^2u,b^2v} = \frac{1}{ab}x_1(a^2u)x_2(b^2v).$$

Because of $\langle x_1(\alpha)x_2(\beta) \rangle = 0$ we have $\langle \mathrm{H}_{u,v} \rangle = 0$. Its autocorrelation is given by

$$\langle \mathrm{H}_{u,v}\mathrm{H}_{p,q} \rangle = \frac{1}{(ab)^2}\langle x_1(a^2u)x_1(a^2p) \rangle \langle x_2(b^2v)x_2(b^2q) \rangle$$
$$= \frac{1}{(ab)^2}(a^2u) \wedge (a^2p)(b^2v) \wedge (b^2q) = (u \wedge p)(v \wedge q).$$

For the case of the translated quantity use the consideration of EX 1.7.

EX 1.9. Verify the differential (1.109) of two linear stochastic functions.

Hint: According to (1.89) we have $\mathrm{dB}_t^2 = 2\mathrm{B}_t\mathrm{dB}_t + \mathrm{d}t \ldots$

EX 1.10. Show that the "inverted" stochastic variables

$$\mathrm{Z}_t = t\mathrm{B}_{1/t}; \quad \mathrm{H}_{s,t} = st\mathrm{M}^{(2)}_{1/s,1/t},$$

are also a WP (Z_t) and a WS ($\mathrm{H}_{s,t}$).

EX 1.11. Use the bivariate PD (1.60) for a Markov process, to calculate the two-variable characteristic function of a Brownian motion. Verify the result

$$\mathrm{G}(u,v) \equiv \langle \exp[i(u\mathrm{B}_1 + v\mathrm{B}_2)] \rangle$$
$$= \exp\left\{-\left[\frac{1}{2}(u^2t_1 + v^2t_2) + 2uv(t_1 \wedge t_2)\right]\right\}, \quad \mathrm{B}_k = \mathrm{B}_{t_k}$$

and compare its 1D limit with (1.58a).

EX 1.12. Calculate the probability P of a particle to stay in the interior of the circle $\mathrm{D} = \{(x,y) \in \mathrm{R}_2 \mid x^2 + y^2 \leq \mathrm{R}\}$.

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

Hint: Assume that components of the vector (x, y) are statistically independent use the bivariate GD (1.35) with zero mean to calculate

$$\mathrm{P}[\mathrm{B}_t \in \mathrm{D}] = \iint_{\mathrm{D}} \mathrm{p}(x, y)\mathrm{d}x\mathrm{d}y.$$

EX 1.13. Consider the Brownian motion on the perimeter of ellipses and hyperbolas

(i) ellipses

$$x(t) = \cos(\mathrm{B}_t), \quad y(t) = \sin(\mathrm{B}_t),$$

(ii) hyperbolas

$$x(t) = \cosh(\mathrm{B}_t), \quad y(t) = \sinh(\mathrm{B}_t).$$

Use the Ito formula to obtain the corresponding SDE and calculate $\langle x(t)\rangle$ and $\langle y(t)\rangle$.

EX 1.14. Given the variables

$$\mathrm{Z}_1 = \left(\mathrm{B}_t^1 - \mathrm{B}_t^2\right)^4 + \left(\mathrm{B}_t^1\right)^5; \quad \mathrm{Z}_2 = \left(\mathrm{B}_t^1 - \mathrm{B}_t^2\right)^3 + \left(\mathrm{B}_t^1\right)^6,$$

where B_t^1 and B_t^2 are independent WP's. Find the SDE's governing dZ_1 and dZ_2.

EX 1.15. The random function

$$\mathrm{R}(t) = [\left(\mathrm{B}_t^1\right)^2 + \cdots + \left(\mathrm{B}_t^n\right)^2]^{1/2},$$

is considered as the distance of an n-dimensional vector of independent WP's from the origin. Verify that its differential has the form

$$\mathrm{dR}(t) = \sum_{k=1}^{n} \mathrm{B}_t^k \mathrm{dB}_t^k/\mathrm{R} + \frac{n-1}{2\mathrm{R}}\mathrm{d}t.$$

EX 1.16. Consider the stochastic function

$$x(t) = \exp(a\mathrm{B}_t - a^2 t/2); \quad a = \text{const}.$$

(a) Show that

$$x(t) = x(t-s)x(s).$$

Hint: Use (1.65).

(b) Show that $x(t)$ is a martingale.

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html

EX 1.17. The **Wiener–Lévy Theorem** is given by

$$\mathrm{B}_t = \sum_{k=1}^{\infty} \lambda_k \int_0^t \psi_k(z)\mathrm{d}z, \tag{E.1}$$

where λ_k is a set of IID $\mathrm{N}(0,1)$ variables and $\psi_k; k = 1, 2, \ldots$ is a set of orthonormal functions in [0, 1]

$$\int_0^1 \psi_k(z)\psi_m(z)\mathrm{d}z = \delta_{km}.$$

Show that (E.1) defines a WP.

Hint: The autocorrelation is given by $\langle \mathrm{B}_t \mathrm{B}_s \rangle = t \wedge s$. Show that

$$\frac{\partial}{\partial t}\langle \mathrm{B}_t \mathrm{B}_s \rangle = \frac{\partial}{\partial t}(t \wedge s) = \sum_{k=1}^{\infty} \psi_k(t) \int_0^s \psi_k(z)\mathrm{d}z.$$

Multiply the last line by $\psi_m(t)$ and integrate the resulting equation from zero to unity.

EX 1.18. A bivariate PD of two variables x, y is given by $\mathrm{p}(x, y)$.

(a) Calculate the PD of the "new" variable z and its average for (i) $z = x \pm y$; (ii) $z = xy$.

Hint: Use (1.41b).

(b) Find the PD $\mathrm{f}_{\mathrm{UV}}(u, v)$ for the "new" variables $u = x + y; v = x - y$.

EX 1.19. The Ito representation of a given stochastic processes $\mathrm{F}(t, \omega)$ has the form

$$\mathrm{F}(t, \omega) = \langle \mathrm{F}(t, \omega) \rangle + \int_0^t \mathrm{f}(s, \omega)\mathrm{dB}_s,$$

where $\mathrm{f}(s, \omega)$ is an other stochastic process. Find $\mathrm{f}(s, \omega)$ for the particular cases

(i) $\mathrm{F}(t, \omega) = \mathrm{const}$; (ii) $\mathrm{F}(t, \omega) = \mathrm{B}_t^n$; $n = 1, 2, 3$; (iii) $\mathrm{F}(t, \omega) = \exp(\mathrm{B}_t)$.

EX 1.20. Calculate the probability of n identically independent HPP's [see (A.8)].

STOCHASTIC DIFFERENTIAL EQUATIONS IN SCIENCE AND ENGINEERING (With CD-ROM)
© World Scientific Publishing Co. Pte. Ltd.
http://www.worldscibooks.com/physics/5806.html