Chapter 16
Some Basics of Stochastic Processes

16.1 Introduction

This chapter is devoted to an introduction to some basic concepts of stochastic processes. This introduction serves two purposes: First of all, it allows for a more systematic treatment of non-deterministic methods in Computational Physics which is certainly necessary if we really aim at an understanding of these methods. The second reason can be found in the elementary importance of stochastics in modern theoretical physics and chemistry in general. Hence, many of the concepts elaborated within this chapter will be of profound importance in subsequent chapters. For instance, we present a discussion of diffusion theory in Chap. 17 as a motivating example.

The reader not familiar with the basics of probability theory [1–4] is highly encouraged to at least consult Appendix E before proceeding. In particular, we are going to apply the notation introduced in Appendix E throughout this chapter.

The basics of stochastic processes will be discussed within five sections including this introduction. In Sect. 16.2 we introduce basic definitions associated with stochastic processes in general. Here we discuss concepts which will serve as a basis for an understanding of the methods presented within the subsequent sections. Section 16.3 deals with a special class of stochastic processes, the so called MARKOV processes. As we shall see, these processes are of fundamental importance for statistical physics and for computational methods. Moreover, in Sect. 16.4 we consider so called MARKOV-chains which are discrete MARKOV processes defined on a discrete time span. This will serve as the basis of a very important method in computational physics, the so called MARKOV-Chain Monte Carlo technique. We already encountered a simple example of this method in Sect. 14.3 and in Chap. 15, the METROPOLIS algorithm. Finally, in Sect. 16.5 continuous-time MARKOV-chains will be discussed, in particular, discrete MARKOV processes on a continuous time span. These processes are very important, for instance, in diffusion theory as will be demonstrated in Chap. 17.

B.A. Stickler, E. Schachinger, *Basic Concepts in Computational Physics*,
DOI 10.1007/978-3-319-27265-8_16

A discussion of detailed balance will also be included in the section on MARKOV processes, Sect. 16.3, although detailed balance follows from physical arguments. Detailed balance has already been introduced in our discussion of the METROPOLIS algorithm, Sect. 14.3.

16.2 Stochastic Processes

The following discussion is primarily restricted to one-dimensional processes.

A stochastic process is a time dependent process depending on randomness [5–7]. From a mathematical point of view, a stochastic process $Y_X(t)$ is a random variable Y which is a function of another random variable X and time $t \geq 0$:

$$Y_X(t) = f(X, t) \; . \tag{16.1}$$

Here we apply the notation of Appendix E and denote random variables by capital letters, such as X, and their realization by lower case characters, such as x. Consequently, the realization of a stochastic process is described by

$$Y_x(t) = f(x, t) \; . \tag{16.2}$$

The set of all possible realizations of $Y_X(t)$ spans the *state space* of the stochastic process. We note that it is in principle not necessary to define t as the time in a classical sense. It suffices to denote $t \in T$, where T is a totally ordered set such as, for instance, $T = \mathbb{N}$ the natural numbers. The set T is referred to as the *time span*. We distinguish four different scenarios:

- discrete state space, discrete time span,
- continuous state space, discrete time span,
- discrete state space, continuous time span,
- continuous state space, continuous time span.

Stochastic processes on a continuous time span are referred to as continuous-time stochastic processes.

Suppose the random variable X follows the pdf $p_X(x)$. It is then an easy task to calculate averages $\langle Y(t) \rangle$ of the stochastic process $Y_X(t)$ via

$$\langle Y(t) \rangle = \int \mathrm{d}x \, Y_x(t) p_X(x) \; . \tag{16.3}$$

This concept is easily extended to multiple times $t_1, t_2, \ldots, t_n$ by

$$\langle Y(t_1) Y(t_2) \cdots Y(t_n) \rangle = \int \mathrm{d}x \, Y_x(t_1) Y_x(t_2) \cdots Y_x(t_n) p_X(x) \; , \tag{16.4}$$

which defines the *moments* of the stochastic process [1, 5, 8]. Similar to the concept of the correlation coefficient (see Appendix, Sect. E.10) we define the so called *auto-correlation function* $\kappa(t_1, t_2)$:

$$\begin{aligned}\kappa(t_1, t_2) &= \frac{\langle [Y(t_1) - \langle Y(t_1)\rangle]\,[Y(t_2) - \langle Y(t_2)\rangle]\rangle}{\sqrt{\left\langle [Y(t_1) - \langle Y(t_1)\rangle]^2 \right\rangle \left\langle [Y(t_2) - \langle Y(t_2)\rangle]^2 \right\rangle}} \\ &= \frac{\langle Y(t_1)Y(t_2)\rangle - \langle Y(t_1)\rangle\,\langle Y(t_2)\rangle}{\sqrt{\mathrm{var}[Y(t_1)]\mathrm{var}[Y(t_2)]}} \\ &= \frac{\gamma[Y(t_1), Y(t_2)]}{\sqrt{\mathrm{var}[Y(t_1)]\mathrm{var}[Y(t_2)]}}\ . \end{aligned} \tag{16.5}$$

The function $\gamma[Y(t_1), Y(t_2)]$ is referred to as the *auto-covariance function* and is defined as

$$\gamma[Y(t_1), Y(t_2)] = \mathrm{cov}[Y(t_1), Y(t_2)]\ . \tag{16.6}$$

We proceed by defining the pdf of a stochastic process $Y_X(t)$. The pdf $p_1(y, t)$, which describes the probability that the stochastic process $Y_X(t)$ takes on its representation y at time t, is given by (see Sect. 14.2)

$$p_1(y, t) = \int \mathrm{d}x\, \delta[y - Y_x(t)] p_X(x)\ . \tag{16.7}$$

We define, in analogy, the pdf $p_n(y_1, t_1, y_2, t_2, \ldots, y_n, t_n)$ which describes the probability that the stochastic process takes on the realization y_1 at time t_1, y_2 at time t_2, ..., and y_n at time t_n for arbitrary n:

$$\begin{aligned} p_n(y_1, t_1, y_2, t_2, \ldots, y_n, t_n) = \int \mathrm{d}x\, \delta[y_1 - Y_x(t_1)]\delta[y_2 - Y_x(t_2)] \cdots \\ \times \delta[y_n - Y_x(t_n)] p_X(x)\ . \end{aligned} \tag{16.8}$$

This is referred to as the *hierarchy of pdfs*. We note the following important properties of the pdf $p_n(y_1, t_1, y_2, t_2, \ldots, y_n, t_n)$ [8]:

- $p_n(y_1, t_1, y_2, t_2, \ldots, y_n, t_n) \geq 0\ ,$ (16.9)
- $p_n(\ldots, y_k, t_k \ldots, y_\ell, t_\ell, \ldots) = p_n(\ldots, y_\ell, t_\ell \ldots, y_k, t_k, \ldots)\ ,$ (16.10)
- $\int \mathrm{d}y_n\, p_n(y_1, t_1, \ldots, y_n, t_n) = p_{n-1}(y_1, t_1, \ldots, y_{n-1}, t_{n-1})\ ,$ (16.11)
- $\int \mathrm{d}y p_1(y, t) = 1\ .$ (16.12)

The moments defined in Eq. (16.4) can also be expressed with the help of the pdfs p_n by

$$\langle Y(t_1)Y(t_2)\cdots Y(t_n)\rangle = \int \mathrm{d}y_1 \ldots \mathrm{d}y_n y_1 \cdots y_n p_n(y_1, t_1, \ldots, y_n, t_n) \; . \tag{16.13}$$

Conditional pdfs $p_{\ell|k}$ can also be introduced. They describe the probability that we have y_{k+1} at $t_{k+1}, \ldots, y_{k+\ell}$ at $t_{k+\ell}$ if there existed y_1 at $t_1, \ldots, y_k$ at t_k via

$$p_{\ell|k}(y_{k+1}, t_{k+1}, \ldots, y_{k+\ell}, t_{k+\ell}|y_1, t_1, \ldots, y_k, t_k) = \frac{p_{k+\ell}(y_1, t_1, \ldots, y_{k+\ell}, t_{k+\ell})}{p_k(y_1 t_1, \ldots, y_k, t_k)} \; . \tag{16.14}$$

It follows that

$$\int \mathrm{d}y_2 \, p_{1|1}(y_2, t_2|y_1, t_1) = 1 \; . \tag{16.15}$$

Let us give some further definitions [8]:

- A stochastic process is referred to as a *stationary process* if the moments defined in Eq. (16.4) are invariant under a time-shift Δt:

$$\langle Y(t_1)Y(t_2)\cdots Y(t_n)\rangle = \langle Y(t_1 + \Delta t)Y(t_2 + \Delta t)\cdots Y(t_n + \Delta t)\rangle \; . \tag{16.16}$$

  In particular, one has $\langle Y(t)\rangle = \text{const}$ and the auto-covariance depends only on the time difference $|t_1 - t_2|$:

$$\gamma(t_1, t_2) = \mathrm{cov}[Y(t_1), Y(t_2)] = \mathrm{cov}[Y(0), Y(|t_1 - t_2|)] \equiv \gamma(t_1 - t_2) \; . \tag{16.17}$$

  It is understood that $\gamma(t) = \gamma(-t)$. Moreover, we have

$$p_n(y_1, t_1 + \Delta t, \ldots, y_n, t_n + \Delta t) = p_n(y_1, t_1, \ldots, y_n, t_n) \; , \tag{16.18}$$

  and in particular, $p_1(y, t) = p_1(y)$.
- A *time-homogeneous* process is a stochastic process whose conditional pdfs are stationary

$$p_{1|1}(y_2, t_2|y_1, t_2 - \tau) = p_{1|1}(y_2, s_2|y_1, s_2 - \tau) \; , \tag{16.19}$$

  for all t_2, τ, s_2. The pdf $p_{1|1}$ is referred to as *transition probability*.
- A *process of stationary increments* is a stochastic process $Y_X(t)$ for which the difference $Y_X(t_2) - Y_X(t_1)$ is stationary for all $t_2 - t_1$, with $t_2 > t_1 \geq 0$. This means, in particular, that the pdf of this process depends only on the time difference $t_2 - t_1$. The quantities $Y_X(t_2) - Y_X(t_1)$ are referred to as *increments*.

- A *process of independent increments* is a stochastic process $Y_X(t)$ for which the differences

$$Y_X(t_2) - Y_X(t_1), Y_X(t_3) - Y_X(t_2), \ldots, Y_X(t_n) - Y_X(t_{n-1}) \ ,$$

  are independent for all $t_n > t_{n-1} > \ldots > t_2 > t_1$.
- A LÉVY *process* is a continuous-time stochastic process with stationary independent increments which starts with $Y_X(0) = 0$.
- A *Gaussian process* is a stochastic process $Y_X(t)$ for which all finite linear combinations of $Y_X(t)$, $t \in T$ follow a normal distribution (see Appendix, Sect. E.7). We shall come back to this kind of process in Chap. 17.
- A WIENER *process* is a continuous-time stochastic process with independent increments which starts with $Y_X(0) = 0$ and for which the increments $Y_X(t_2) - Y_X(t_1)$ follow a normal distribution with mean 0 and variance $t_2 - t_1$. The WIENER process is a special case of a LÉVY process. One of the main applications of the WIENER process is to study *Brownian motion* or diffusion. This process will be discussed in more detail in Sect. 16.3 and in Chap. 17.
- The *random walk* is the discrete analogy to the WIENER process [9–11]. This means in particular that if the step size of the random walk goes to zero, the WIENER process is reestablished. This point will be elucidated in Chap. 17.

After stating the most important definitions, we proceed to the next section in which the attention is on a special class of stochastic processes, the so called MARKOV processes.

16.3 MARKOV Processes

A MARKOV process is a stochastic process $Y_X(t)$ for which the conditional pdf $p_{1|n-1}$ satisfies for arbitrary n and $t_1 < t_2 < \cdots < t_n$ the relation

$$p_{1|n-1}(y_n, t_n | y_1, t_1, \ldots, y_{n-1}, t_{n-1}) = p_{1|1}(y_n, t_n | y_{n-1}, t_{n-1}) \ . \tag{16.20}$$

Hence, a MARKOV process is a process in which any state y_n, t_n is uniquely defined by its precursor state y_{n-1}, t_{n-1} and is independent of the entire rest of the past [12]. MARKOV processes are of particular importance in natural sciences because of their rather simple structure. This will become clear throughout the rest of this book.

We note in passing that a process with independent increments is *always* a MARKOV process because

$$Y_X(t_{n+1}) = Y_X(t_n) + [Y_X(t_{n+1}) - Y_X(t_n)] \ , \tag{16.21}$$

is satisfied. Since the increment $Y_X(t_{n+1}) - Y_X(t_n)$ is independent of all previous increments which gave rise to $Y_X(t_n)$ *by definition*, $Y_X(t_{n+1})$ depends only on $Y_X(t_n)$, which is exactly the MARKOV property (16.20).

The quantity $p_{1|1}(y_n, t_n|y_{n-1}, t_{n-1})$ which appears in Eq. (16.20) is referred to as *transition probability*. Given the transition probability $p_{1|1}$ together with the pdf p_1, one can construct the whole hierarchy of pdfs (16.8) of the MARKOV process by calculating successively [8]:

$$\begin{aligned} p_2(y_1, t_1, y_2, t_2) &= p_{1|1}(y_2, t_2|y_1, t_1)p_1(y_1, t_1) \ , \\ p_3(y_1, t_1, y_2, t_2, y_3, t_3) &= p_{1|2}(y_3, t_3|y_1, t_1, y_2, t_2)p_2(y_1, t_1, y_2, t_2) \ , \\ &= p_{1|1}(y_3, t_3|y_2, t_2)p_{1|1}(y_2, t_2|y_1, t_1)p_1(y_1, t_1) \ , \\ &\vdots \quad \vdots \end{aligned} \tag{16.22}$$

Here we employed definition (16.14) and in the second step of the second equation we employed for $p_{1|2}$ the MARKOV property (16.20).

The fact that the whole hierarchy of pdfs can be constructed by repeating the steps illustrated in Eq. (16.22) reveals the rather simple structure of MARKOV processes. However, Eq. (16.22) contains another useful information. We regard the pdf p_3 of (16.22) for three successive times $t_1 < t_2 < t_3$. First we integrate the left-hand side with respect to y_2 which yields with the help of property (16.10):

$$\int \mathrm{d}y_2 \, p_3(y_1, t_1, y_2, t_2, y_3, t_3) = p_2(y_1, t_1, y_3, t_3) \ . \tag{16.23}$$

Hence, we have

$$p_2(y_1, t_1, y_3, t_3) = p_1(y_1, t_1) \int \mathrm{d}y_2 \, p_{1|1}(y_3, t_3|y_2, t_2)p_{1|1}(y_2, t_2|y_1, t_1) \ , \tag{16.24}$$

or after dividing both sides by $p_1(y_1, t_1)$ and by keeping in mind Eq. (16.14) we arrive at:

$$p_{1|1}(y_3, t_3|y_1, t_1) = \int \mathrm{d}y_2 \, p_{1|1}(y_3, t_3|y_2, t_2)p_{1|1}(y_2, t_2|y_1, t_1) \ . \tag{16.25}$$

This equation is known as the CHAPMAN-KOLMOGOROV equation [8, 13]. The interpretation of this equation is straight-forward: the transition probability from (y_1, t_1) to (y_3, t_3) is equivalent to the transition probability from (y_1, t_1) to (y_2, t_2) multiplied by the transition probability from (y_2, t_2) to (y_3, t_3) when summed over all intermediate positions y_2. This is illustrated in Fig. 16.1.

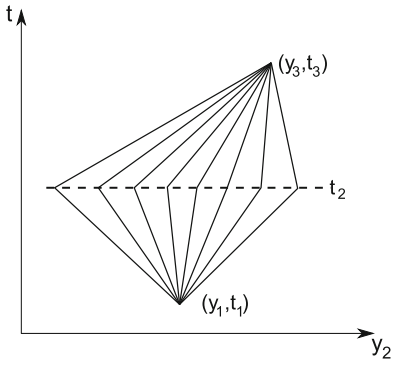

Fig. 16.1 Illustration of the CHAPMANN-KOLMOGOROV equation

We state a very important theorem: Any two non-negative functions $p_{1|1}$ and p_1 uniquely define a MARKOV process if the CHAPMAN-KOLMOGOROV equation (16.25) is obeyed and if

$$p_1(y_2, t_2) = \int \mathrm{d}y_1 p_{1|1}(y_2, t_2|y_1, t_1) p_1(y_1, t_1) \,, \tag{16.26}$$

which follows immediately from the first equation in Eqs. (16.22) in combination with property (16.10).

As a first example we consider one of the most important MARKOV processes in physics, the WIENER process [10]. Its importance stems from its application to the description of Brownian motion, the random motion of dust particles on a fluid surface. (In Chap. 17 we take a closer look at diffusion phenomena.) The transition probability of the WIENER process is of the form[1]

$$p_{1|1}(y_2, t_2|y_1, t_1) = \frac{1}{\sqrt{2\pi(t_2 - t_1)}} \exp\left[-\frac{(y_2 - y_1)^2}{2(t_2 - t_1)}\right] . \tag{16.27}$$

The initial condition is given by

$$p_1(y_1, t_1 = 0) = \delta(y_1) \,. \tag{16.28}$$

A straight-forward calculation proves that (16.27) indeed obeys the CHAPMAN-KOLMOGOROV equation (16.25). Moreover, we deduce from Eqs. (16.28) together with (16.26) that

$$p_1(y, t) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{y^2}{2t}\right) . \tag{16.29}$$

[1]This form is equivalent to the above definition of the WIENER process, in particular to the requirement of normally distributed increments with variance $t_2 - t_1$.

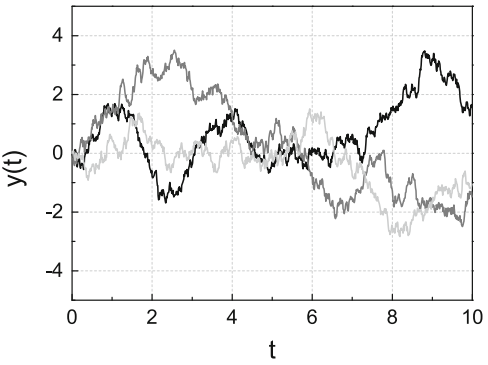

Fig. 16.2 Three possible realizations of the WIENER process

The WIENER process is easily realized on the computer. We regard the one-dimensional case and start at the origin $Y_X(0) = 0$. As per definition the increments $Y_X(t + \mathrm{d}t) - Y_X(t)$ follow a normal distribution $\mathcal{N}(\mathrm{d}y|0, \mathrm{d}t)$ of mean zero and variance $\mathrm{d}t$. Hence, we start with $y_0 = 0$ at time $t_0 = 0$, sample the displacement $\mathrm{d}y$ within a time-step $\mathrm{d}t$ from $\mathcal{N}(\mathrm{d}y|0, \mathrm{d}t)$ and calculate the next position at time $t_1 = t_0 + \mathrm{d}t$ which is given by: $y_1 = y_0 + \mathrm{d}y$.[2] This process is repeated until a certain time limit has been reached. Figure 16.2 presents the result of three such calculations which have been started using different seeds.

Let us mention a second very important MARKOV process, the POISSON process. The POISSON process is particularly interesting for problems involving *waiting times*, such as the decay of some radioactive nucleus. However, we shall also come across the POISSON process within the context of diffusion in Chap. 17. The transition probability of the POISSON process is defined as

$$p_{1|1}(n_2, t_2|n_1, t_1) = \frac{(t_2 - t_1)^{n_2 - n_1}}{(n_2 - n_1)!} \exp\left[-(t_2 - t_1)\right] \ . \tag{16.30}$$

Here it is understood that $n_1, n_2 \in \mathbb{N}$ and $n_2 > n_1$. Hence, the POISSON process counts the number of occurrences n_2 of a certain event until the time t_2 is reached under the premise that n_1 events have already occurred at time t_1. The POISSON process is initialized by the pdf

$$p_1(n_1, t_1 = 0) = \delta_{n_1 0} \ , \tag{16.31}$$

here δ_{ij} is the KRONECKER-δ. Hence we have according to Eq. (16.26)

$$p_1(n, t) = \sum_{n_1} p_{1|1}(n, t|n_1, t_1 = 0) p_1(n_1, t_1 = 0) = \frac{t^n}{n!} \exp(-t) \ , \tag{16.32}$$

[2] Alternatively, we may sample $\mathrm{d}y$ from a normal distribution with variance 1 and multiply it by $\sqrt{\mathrm{d}t}$. This follows from a simple transformation of variables.

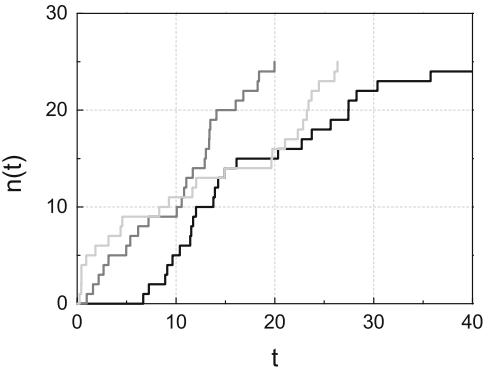

Fig. 16.3 Three possible realizations of a POISSON process

which is a POISSON distribution (see Appendix, Sect. E.4 and, for instance, Ref. [14]). Let us briefly consider the time between two events. Suppose we had n_1 events at time t_1. Then, we calculate the probability that at time $t_2 = t_1 + \tau$ we still counted $n_2 = n_1$ events, thus, nothing happened. We have

$$p_{1|1}(n_1, t_1 + \tau | n_1, t_1) = \exp(-\tau) \tag{16.33}$$

and the waiting times are independent and follow an exponential distribution. We may simulate the POISSON process by starting at $t_1 = 0$ with $n_1 = 0$ and by increasing $n_2, n_3, \ldots$ by one, i.e. $n_{i+1} = n_i + 1$ after successive waiting times $\tau_1, \tau_2, \ldots$ which we sample from the exponential distribution (see Sect. 13.2) until a final count N has been reached. The result of such a procedure is illustrated in Fig. 16.3 where $n(t)$ vs t has been plotted for three runs started by different seeds.

Finally, we remark that for a time-homogeneous MARKOV process the transition probability $p_{1|1}(y_2, t_2|y_1, t_1)$ depends by definition on the time difference $t_2 - t_1 \equiv \tau$ rather than explicitly on the two times t_1 and t_2 and is usually denoted by $T_\tau(y_2, y_1)$.

We turn now our attention to another very important general concept of MARKOV processes, the *master equation* [8]. This equation is in fact the differential form of the CHAPMAN-KOLMOGOROV equation. We regard the CHAPMAN-KOLMOGOROV equation (16.25) for three successive times $t_1 < t_2 < t_3 = t_2 + \tau$ where τ is assumed to be small, i.e. $\tau \ll 1$. We expand the conditional pdf $p_{1|1}$ in a TAYLOR series with respect to τ:

$$p_{1|1}(y_3, t_2 + \tau | y_2, t_2) = p_{1|1}(y_3, t_2 | y_2, t_2) + \tau \frac{\partial}{\partial \tau} p_{1|1}(y_3, t_2 + \tau | y_2, t_2)\Big|_{\tau=0} + \mathcal{O}(\tau^2) \,. \tag{16.34}$$

In order to transform this equation into a more transparent form, we introduce the transition rate $w(y_3|y_2, t_2)$ from y_2 to y_3, with $y_2 \neq y_3$:

$$w(y_3|y_2, t_2) = \frac{\partial}{\partial \tau} p_{1|1}(y_3, t_2 + \tau | y_2, t_2)\Big|_{\tau=0} \,. \tag{16.35}$$

In addition, we note that the first term on the right-hand side of Eq. (16.34) has to be of the form:

$$p_{1|1}(y_3, t_2|y_2, t_2) = \delta(y_3 - y_2) \ . \tag{16.36}$$

On the other hand, we defined the transition rate (16.35) only for elements $y_2 \neq y_3$ and, thus, we denote the remaining part (i.e. $y_2 = y_3$) by $a(y_2, t_2)$. All this allows us to rewrite Eq. (16.34) as

$$p_{1|1}(y_3, t_2 + \tau|y_2, t_2) = [1 + a(y_2, t_2)]\delta(y_3 - y_2) + \tau w(y_3|y_2, t_2) \ , \tag{16.37}$$

where we neglected terms of order $\mathcal{O}(\tau^2)$. The pdf $p_{1|1}(y_3, t_2 + \tau|y_2, t_2)$ is subject to the normalization (16.15) and this provides us with the required condition to determine $a(y_2, t_2)$:

$$a(y_2, t_2) = -\tau \int \mathrm{d}y_3 w(y_3|y_2, t_2) \ . \tag{16.38}$$

Hence, the term $1 + a(y_2, t_2)$ describes the probability that no event occurs within the time interval $[t_2, t_2 + \tau]$. The expansion (16.37) is inserted into the CHAPMAN-KOLMOGOROV equation (16.25) with the result:

$$\begin{aligned} \frac{p_{1|1}(y_3, t_2 + \tau|y_1, t_1) - p_{1|1}(y_3, t_2|y_1, t_1)}{\tau} = \int \mathrm{d}y_2 \big[& w(y_3|y_2, t_2) p_{1|1}(y_2, t_2|y_1, t_1) \\ & - w(y_2|y_3, t_2) p_{1|1}(y_3, t_2|y_1, t_1) \big] \ . \end{aligned} \tag{16.39}$$

Finally, we arrive, in the limit $\tau \to 0$, at the master equation:

$$\begin{aligned} \frac{\partial}{\partial t} p_{1|1}(y, t|y', t') = \int \mathrm{d}y'' \big[& w(y|y'', t) p_{1|1}(y'', t|y', t') \\ & - w(y''|y, t) p_{1|1}(y, t|y', t') \big] \ . \end{aligned} \tag{16.40}$$

We multiply this equation by $p_1(y', t')$ and integrate over y'. This results in the master equation for the pdf $p_1(y, t)$

$$\frac{\partial}{\partial t} p_1(y, t) = \int \mathrm{d}y' \left[w(y|y', t) p_1(y', t) - w(y'|y, t) p_1(y, t) \right] \ , \tag{16.41}$$

where we made use of the property (16.26).

Let us briefly discuss this result: In its derivation we assumed the state space to be continuous. However, if a master equation for a discrete state space is required the integral is to be replaced by a sum over the discrete states. On the other hand, the physical interpretation of such an equation is straight-forward: The time evolution

of the quantity $p_1(y,t)$ is governed by the sum over all transitions into state y (first term) minus all transitions out of state y. We remark that master equations occur commonly in physical applications; for instance, the collision integral of the BOLTZMANN equation is of a similar form. The transitions rates $w(y|y',t)$ can be determined explicitly in many applications in physics.[3]

Furthermore, if the system is in a stationary state then it is described by a stationary distribution $p_1(y,t) = p_1(y)$ and we obtain from (16.41) the relation

$$\int \mathrm{d}y' w(y|y',t)p_1(y') = \int \mathrm{d}y' w(y'|y,t)p_1(y) \;, \tag{16.42}$$

which is referred to as *global balance*. The much stronger condition

$$w(y|y',t)p_1(y') = w(y'|y,t)p_1(y) \;, \tag{16.43}$$

is referred to as *detailed balance* and will be discussed next.

The task now is to prove that the equilibrium distribution function $p_e(X)$ of a classical physical system will, under certain restrictions, indeed fulfill detailed balance. (This proof was given by N.G. VAN KAMPEN [5].) The next steps of the proof become more transparent if a vector $x = (q_k, p_k)^T \in \mathbb{R}^{6N}$ is introduced which represents the phase space trajectory of the N particles constituting the system under investigation. This trajectory is determined by HAMILTON's equations of motion [16–18]:

$$\dot{q}_k = \frac{\partial}{\partial p_k} H(x) \;, \qquad \dot{p_k} = -\frac{\partial}{\partial q_k} H(x) \;. \tag{16.44}$$

Furthermore, $Y_X(t)$ denotes a stochastic process which describes some observable of the physical system. We require that $Y_X(t)$ is invariant under time reversal. We also assume the equilibrium distribution function $p_e(x)$ to be invariant under time reversal, which in most cases is equivalent to the requirement that the HAMILTON function $H(x)$ is invariant under time reversal. The operation of time reversal will be indicated by bared variables:

$$\bar{t} = -t \;, \quad \bar{x} = (q_k, -p_k)^T \;. \tag{16.45}$$

[3] As an example we quote FERMI's golden rule [15], where the transition rate $w_{nn'}$ from unperturbed states n to n' is of the form

$$w_{nn'} = \frac{2\pi}{\hbar} |H'_{nn'}| \rho(E_n) \;,$$

where $H'_{nn'}$ are the matrix elements of the perturbation Hamiltonian H' and $\rho(E_n)$ denotes the density of states of the unperturbed system.

Hence, the above assumptions result in

$$\overline{Y_x(t)} = Y_{\bar{x}}(\bar{t}) = Y_{\bar{x}}(-t) = Y_x(t) \ , \tag{16.46}$$

and

$$\overline{p_e(x)} = p_e(\bar{x}) = p_e(x) \ . \tag{16.47}$$

In particular, we deduce from Eq. (16.46) that

$$Y_{\bar{x}}(0) = Y_x(0) \ , \tag{16.48}$$

and

$$Y_{\bar{x}}(t) = Y_x(-t) \ . \tag{16.49}$$

We calculate the pdf p_2 from

$$p_2(y_1, 0, y_2, t) = \int \mathrm{d}x \delta[y_1 - Y_x(0)] \delta[y_2 - Y_x(t)] p_e(x) \ . \tag{16.50}$$

However, since we integrate over the whole phase space we recognize that the volume is invariant under a change $\mathrm{d}x \to \mathrm{d}\bar{x}$. Thus, we can change the variable of integration from x to $\bar{x}$ and this results in:

$$\begin{aligned} p_2(y_1, 0, y_2, t) &= \int \mathrm{d}\bar{x} \delta[y_1 - Y_{\bar{x}}(0)] \delta[y_2 - Y_{\bar{x}}(t)] p_e(\bar{x}) \\ &= \int \mathrm{d}x \delta[y_1 - Y_x(0)] \delta[y_2 - Y_x(-t)] p_e(x) \\ &= p_2(y_2, -t, y_1, 0) \\ &= p_2(y_2, 0, y_1, t) \ . \end{aligned} \tag{16.51}$$

We obtain immediately:

$$p_{1|1}(y_2, t|y_1, 0) p_e(y_1) = p_{1|1}(y_1, t|y_2, 0) p_e(y_2) \ . \tag{16.52}$$

Differentiation of this equation with respect to t together with definition (16.35) yields for small values of t

$$w(y_2|y_1) p_e(y_1) = w(y_1|y_2) p_e(y_2) \ , \tag{16.53}$$

which is the condition of detailed balance, Eq. (16.43), for stationary distributions.

It should be noted at this point that detailed balance in physical systems is strongly connected to the entropy growth (the H-theorem by BOLTZMANN [19]).

Here, detailed balance was based on the condition that the stochastic process $Y_X(t)$ was invariant under time reversal and that the equilibrium distribution $p_e(x)$ had the same property. This has in most cases the consequence that the HAMILTON function is also invariant under time reversal transformations. However, a detailed discussion of these properties is far beyond the scope of this book.

We continue our presentation with so called MARKOV-chains which are a special class of MARKOV processes. MARKOV-chains will prove to be very important for the understanding of MARKOV-chain Monte Carlo techniques, such as the METROPOLIS algorithm.

16.4 MARKOV-Chains

A MARKOV-chain is a time-homogeneous MARKOV process defined on a discrete time span and in a discrete state space [20–22]. Hence, we express the time instances by integers $T = \mathbb{N}$, $t_n = n$ where $n \in \mathbb{N}$ and possible outcomes are indexed by integers $Y_X(t_n) \in \{m\}$ where $m \in \mathbb{N}$. As a first consequence of the discreteness of the state space we replace all pdfs $p(\cdot)$ by probabilities $\Pr(\cdot)$. Hence the MARKOV property (16.20) reads

$$\Pr(Y_{n+1} = y | Y_n = y_n, \ldots, Y_1 = y_1) = \Pr(Y_{n+1} = y | Y_n = y_n) \ , \tag{16.54}$$

where we applied the notation $Y_n \equiv Y_X(t_n)$ and $y_n \in \{m\}$ is one particular realization out of the discrete state space. Since we assume the transition probabilities to be independent of the actual time, we can define a *transition matrix* $P = \{p_{ij}\}$ via

$$p_{ij} = \Pr(Y_{n+1} = j | Y_n = i) \ . \tag{16.55}$$

Consequently, we write

$$\Pr(Y_n = i_n, Y_{n-1} = i_{n-1}, \ldots, Y_0 = i_0) = \Pr(Y_0 = i_0) p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} \ . \tag{16.56}$$

We note that the transition matrix is a *stochastic matrix*, a matrix with only non-negative elements such that the sum of each row is equal to one. Furthermore, one can prove that the product of two stochastic matrices results, again, in a stochastic matrix.

We define the *state vector at time n*, $\pi^{(n)} = \{\pi_i^{(n)}\}$ as

$$\pi_i^{(n)} = \Pr(Y_n = i) \ . \tag{16.57}$$

From the marginalization rule (Appendix, Sect. E.6) follows for the particular case $n = 1$

$$\Pr(Y_1 = i) = \sum_k \Pr(Y_0 = k) \Pr(Y_1 = i | Y_0 = k) \,, \tag{16.58}$$

or with the help of the definitions (16.55) and (16.57):

$$\pi_i^{(1)} = \sum_k p_{ki} \pi_k^{(0)} \,, \quad \forall i. \tag{16.59}$$

Hence, we get for $n = 1$

$$\pi^{(1)} = \pi^{(0)} P \,, \tag{16.60}$$

and for $n = 2$

$$\pi^{(2)} = \pi^{(1)} P = \pi^{(0)} P^2 \,. \tag{16.61}$$

Obviously,

$$\pi^{(n)} = \pi^{(0)} P^n \,, \tag{16.62}$$

follows for arbitrary n. Hence the probability matrix for an n step transition $P^{(n)}$ is given by $P^{(n)} = P^n$. We immediately deduce that the CHAPMAN-KOLMOGOROV equation for MARKOV-chains is fulfilled since

$$P^{(n)} P^{(m)} = P^n P^m = P^{n+m} = P^{(n+m)} \,, \tag{16.63}$$

for two integers n and m.

Let us cite some further definitions in order to classify MARKOV-chains [5, 8, 20–22]:

- The notation $i \to j$ means *state i leads to state j* and is true whenever there is a path of length n, $i_0 = i, i_1, \ldots, i_n = j$ such that all $p_{i_k i_{k+1}} > 0$ for $k = 0, 1, \ldots, n-1$. This is equivalent to $(P^n)_{ij} > 0$.
- The notation $i \leftrightarrow j$ means *state i communicates with state j*. This relation is true whenever $i \to j$ and $j \to i$.
- A class of states is given if (i) all states within one class communicate with each other and (ii) two states of different classes never communicate with each other. These classes are referred to as the *irreducible* classes of the MARKOV-chain.
- An *irreducible* MARKOV-*chain* is a MARKOV-chain in which the whole state space forms an irreducible class, i.e. *all* states communicate with each other.
- A *closed set of states* is a set of states which never leads to states which are outside of this set.

- An *absorbing state* is a state which does not lead to any other states: It forms itself a closed set. An absorbing state can be reached from the outside but there is no escape from it.
- A state is referred to as *transient* if the probability of returning to the state is less than one.
- A state is referred to as *recurrent* if the probability of returning to the state is equal to one.
- Furthermore, we call a state *positive recurrent* if the expectation value of the first return time is less than infinity and *null recurrent* if it is infinity. We may formulate this in a more mathematical language: The time of first return to state i is defined via

$$T_i = \inf\left(n \geq 1 : X_n = i | X_0 = i\right) \ . \tag{16.64}$$

The probability that we return to state i for the first time after n steps is defined as

$$f_{ii}^n = \Pr(T_i = n) \ . \tag{16.65}$$

Hence, a state is referred to as recurrent if

$$F_i = \sum_n f_{ii}^n = 1 \ , \tag{16.66}$$

positive recurrent if

$$\langle T_i \rangle = \sum_n n f_{ii}^n < \infty \ , \tag{16.67}$$

and null recurrent if

$$\langle T_i \rangle = \sum_n n f_{ii}^n = \infty \ . \tag{16.68}$$

We note that we also have $\langle T_i \rangle = \infty$ if state i is transient. Furthermore, one can show that a state is only recurrent if

$$\sum_n p_{ii}^n = \infty \ . \tag{16.69}$$

- A state is referred to as *periodic* if the return time of the state can only be a multiple of some integer $d > 1$.
- A state is referred to as *aperiodic* if $d = 1$.
- We call a state *ergodic* if it is positive recurrent and aperiodic.
- A MARKOV-chain is called *ergodic* if all its states are *ergodic*.

We give some useful theorems in the context of the above definitions: First of all, it can be proved that if a MARKOV-chain is irreducible it follows that either *all* states are transient, or *all* states are null recurrent, or *all* states are positive recurrent.

Furthermore, a theorem by KOLMOGOROV states that if a MARKOV-chain is irreducible and aperiodic then the limit

$$\pi_j = \lim_{n\to\infty} \pi_j^{(n)} = \frac{1}{\langle T_j \rangle} , \tag{16.70}$$

exists. It follows from the above discussion that if all states j are transient or null recurrent we have

$$\pi_j = 0 , \tag{16.71}$$

and if all states j are positive recurrent, we have

$$\pi_j \neq 0 . \tag{16.72}$$

In this case the state vector $\pi = \{\pi_j\}$ is referred to as the *stationary distribution* or *equilibrium distribution*. We note that in this context the term equilibrium does not mean that nothing changes, but that the system *forgets* its own past. In particular, as soon as the system reaches the stationary distribution, it is independent of the initial state $\pi^{(0)}$.

We concentrate now on equilibrium distributions. It follows from Eq. (16.62) that π satisfies:

$$\pi = \pi P . \tag{16.73}$$

Thus, π is the *left-eigenvector* to the transition probability matrix P with eigenvalue 1. We note that Eq. (16.73) states a homogeneous eigenvalue problem: The solution is only determined up to a constant multiplicator (see Sect. 8.3). However, it is clear that the vector π satisfies

$$\sum_j \pi_j = 1 . \tag{16.74}$$

One can prove that the unique solution of the eigenvalue problem (16.73) together with the normalization condition (16.74) for n states can be written as

$$\pi = e \cdot (P - E - I)^{-1} , \tag{16.75}$$

where e is an n-element row vector containing only ones, E is a $n \times n$ matrix containing only ones and I is the $n \times n$ identity.

Let us briefly elaborate on this point: if it is possible to construct a MARKOV-chain which possesses a unique stationary distribution, we know that it will

definitely reach this distribution independent of the choice of initial conditions. The existence as well as the form of the stationary distribution is clearly determined by the transition probabilities p_{ij}. A sufficient condition for a unique stationary distribution to exist is the requirement of reversibility. A MARKOV-chain is referred to as *reversible* if

$$p_{ij}\pi_i = p_{ji}\pi_j \ , \quad \forall i,j, \tag{16.76}$$

i.e. if the transition probabilities ensure *detailed balance* for the stationary distribution π.

Now we are in a position to understand better why detailed balance was such an important concept of the METROPOLIS algorithm discussed in Sect. 14.3: Invoking the detailed balance condition ensures that for all possible initial states the MARKOV-chain converges toward the equilibrium distribution for which detailed balance is fulfilled. Of course, the convergence time will highly depend on the choice of the initial state as well as on the choice of the transition matrix. Hence, we can generate random numbers with the help of such a MARKOV-chain and after a *thermalization* period these numbers will follow the required pdf. Methods based on this concept are commonly referred to as MARKOV-*chain Monte Carlo sampling methods* [23–26].

We give a brief example, the *spread of a rumor*. Let Z_1 and Z_2 be two distinct versions of a report. If a person receives report Z_1 it will pass this report on as Z_1 with probability $(1-p)$ or as Z_2 with probability p. An alternative is that the person receives Z_2 and passes it on as Z_2 with probability $(1-q)$ or modifies it to Z_1 with probability q. We summarize

- $\Pr(Z_1 \to Z_1) = (1-p) = p_{11}$,
- $\Pr(Z_1 \to Z_2) = p = p_{12}$,
- $\Pr(Z_2 \to Z_1) = q = p_{21}$,
- $\Pr(Z_2 \to Z_2) = (1-q) = p_{22}$.

The transition matrix is of the form

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} . \tag{16.77}$$

We note that the two states communicate with each other $Z_1 \leftrightarrow Z_2$, hence the MARKOV-chain is irreducible. Furthermore, since the process can reach either state Z_1 or Z_2 within a single time step, it is clearly aperiodic. Let us briefly investigate the probabilities of first recurrence f_{ii}^n after n steps. Due to the theorem by KOLMOGOROV it is sufficient to investigate the state Z_1 since the MARKOV-chain is irreducible and it follows that also Z_2 has the same recurrence properties. We note the following possible paths for a first return to state Z_1:

- $1:\quad \Pr(Z_1 \to Z_1) = (1-p) = f_{11}^1$,
- $2:\quad \Pr(Z_1 \to Z_2 \to Z_1) = pq = f_{11}^2$,

- $3:\quad \Pr(Z_1 \to Z_2 \to Z_2 \to Z_1) = p(1-q)q = f_{11}^3$,
- $n:\quad \Pr(Z_1 \to Z_2 \to \cdots \to Z_2 \to Z_1) = p(1-q)^{n-2}q = f_{11}^n$.

The probability of returning to Z_1 is, see Eq. (16.66),

$$\begin{aligned} F_1 &= \sum_{n=1}^{\infty} f_{11}^n \\ &= (1-p) + pq \sum_{n=0}^{\infty} (1-q)^n \\ &= (1-p) + pq \frac{1}{1-(1-q)} \\ &= 1 \ , \end{aligned} \tag{16.78}$$

where we employed that $0 < (1-q) < 1$ as well as the convergence of the geometric series. Hence state Z_1 is recurrent and, therefore, also state Z_2. We calculate the expectation value of the first return time $\langle T_1 \rangle$:

$$\begin{aligned} \langle T_1 \rangle &= \sum_{n=1}^{\infty} n f_{11}^n \\ &= (1-p) + pq \sum_{n=0}^{\infty} (n+2)(1-q)^n \\ &= (1-p) + 2pq \underbrace{\sum_{n=0}^{\infty} (1-q)^n}_{=\frac{1}{q}} + pq \sum_{n=0}^{\infty} n(1-q)^n \\ &= 1 + p - pq(1-q) \frac{\mathrm{d}}{\mathrm{d}q} \underbrace{\sum_{n=0}^{\infty} (1-q)^n}_{=\frac{1}{q}} \\ &= 1 + p + \frac{p}{q}(1-q) \\ &= \frac{p+q}{q} \ . \end{aligned} \tag{16.79}$$

Hence, the states Z_1 and Z_2 are positive recurrent as long as $p \neq 0$ and $q \neq 0$. This means that an equilibrium distribution exists and it can be obtained from Eq. (16.70).

We have

$$\pi_1 = \frac{1}{\langle T_1 \rangle} = \frac{q}{p+q} \,. \tag{16.80}$$

Due to the normalization condition (16.74) we obtain

$$\pi_2 = 1 - \pi_1 = \frac{p}{p+q} \,, \tag{16.81}$$

and, therefore,

$$\langle T_2 \rangle = \frac{p+q}{p} \,. \tag{16.82}$$

Since all states are positive recurrent and aperiodic, the above MARKOV-chain is ergodic. Finally, we remark that this example also fulfills detailed balance since

$$\pi_1 p_{12} = \frac{qp}{p+q} = \pi_2 p_{21} \,. \tag{16.83}$$

Let us briefly interpret this example: Suppose the original, true version Z_1 of a report is '*Mr. X is going to resign*' while Z_2 is just the opposite: '*Mr. X is* not *going to resign*'. The property of irreducibility of the MARKOV-chain reflects the fact that there is no version of the report which cannot be reached or modified. Moreover, we just demonstrated that the process is positive recurrent: Even if the probability p that Z_1 was modified to Z_2 is very small and the probability q that Z_2 was modified to Z_1 is very high, the report will infinitely often return to version Z_2 with probability one. This means that the public will be told infinitely often that Mr. X is *not* going to resign with probability one. The equilibrium probabilities π_1 and π_2 display the asymptotic probability of versions one and two of the report, respectively. However, as has already been emphasized, this does not mean that the report cannot be modified in equilibrium, it simply displays the fact that the probabilities reached a steady state. Finally, we note an interesting effect in passing: Suppose that the probabilities that any of the two versions is modified is very small but equal, i.e. $p = q \ll 1$. Then the equilibrium distribution is

$$\pi_1 = \pi_2 = \frac{1}{2} \,, \tag{16.84}$$

and the public will believe Z_1 and Z_2 with the same probability after some time independently of the initial version and also independent of the actual decision of Mr. X. Detailed balance expresses the property that the probability of receiving Z_2 and passing it on as Z_1 is the same as the probability of receiving Z_2 and passing it on as Z_1.

We close this section with a final remark: It is an easy task to generalize the ideas of MARKOV-chains to continuous state spaces since we already introduced the required tools in Sect. 16.3. Let $\pi(x)$ denote the stationary distribution density and $p(x|y)$ the accompanying transition rate pdf. Then relation (16.73) transforms into

$$\pi(x) = \int \mathrm{d}y \pi(y) p(x|y) \,, \tag{16.85}$$

together with

$$\int \mathrm{d}x \pi(x) = 1 \,, \tag{16.86}$$

the usual normalization of pdfs. In this case, the condition of detailed balance is given by

$$\pi(x) p(y|x) = \pi(y) p(x|y) \,, \tag{16.87}$$

which is equivalent to Eq. (16.52).

16.5 Continuous-Time MARKOV-Chains

A generalization of the results of the previous sections to a continuous time span is straight-forward. We define the continuous-time MARKOV-chain as a time-homogeneous MARKOV process on a discrete state space but with a continuous time span, $t \geq 0$. Thus

$$\Pr[X(t+s) = n | X(s) = m] = p_{nm}(t) \,, \tag{16.88}$$

is *independent* of $s \geq 0$. In this case the transition matrix $P(t) = \{p_{ij}(t)\}$ is an explicit function of time t. Its elements $p_{nm}(t)$ have the following four properties:

(a) All matrix elements $p_{nm}(t)$ of the transition matrix P are positive:

$$p_{nm}(t) \geq 0 \,, \qquad \forall t > 0 \,. \tag{16.89}$$

(b) The usual normalization of the rows of the transition matrix P is valid:

$$\sum_m p_{nm}(t) = 1 \,, \qquad \forall n \text{ and } t > 0 \,. \tag{16.90}$$

(c) As for every MARKOV process, the transition matrix of the continuous time MARKOV-chain obeys the CHAPMAN-KOLMOGOROV equation:

$$\sum_k p_{nk}(t)p_{km}(t') = p_{nm}(t+t') \,, \tag{16.91}$$

which can alternatively be expressed as

$$P(t+t') = P(t)P(t') \,. \tag{16.92}$$

(d) We assume $p_{nm}(t)$ to be a continuous function of t and that:

$$\lim_{t\to 0} p_{nm}(t) = \begin{cases} 1 & \text{for } n = m \,, \\ 0 & \text{for } n \neq m \,. \end{cases} \tag{16.93}$$

It follows from this equation that the matrix elements $p_{nm}(t)$ can be written as

$$p_{nm}(t) = \begin{cases} 1 + q_{nn}\, t + \mathscr{O}(t^2) & \text{for } n = m \,, \\ q_{nm}\, t + \mathscr{O}(t^2) & \text{for } n \neq m \,, \end{cases} \tag{16.94}$$

where we introduced with $\{q_{nm}\} = Q$ the *transition rate matrix*. The transition rate matrix Q obeys:

(a) All off-diagonal elements q_{nm}, $n \neq m$, are non-negative since

$$q_{nm} = \lim_{t\to 0} \frac{p_{nm}(t)}{t} \geq 0 \qquad \text{for } n \neq m \,. \tag{16.95}$$

(b) All diagonal elements q_{nn} are non-positive since

$$q_{nn} = -\lim_{t\to 0} \frac{1 - p_{nn}(t)}{t} \leq 0 \,. \tag{16.96}$$

(c) Differentiating Eq. (16.90) with respect to t yields that the sum over all elements in a row is equal to zero. Therefore, we conclude:

$$q_{nn} = -\sum_{\substack{m \\ n\neq m}} q_{nm} \,. \tag{16.97}$$

Moreover, differentiating the CHAPMAN-KOLMOGOROV equation with respect to t or t' gives the KOLMOGOROV forward – or KOLMOGOROV backward equations

$$\dot{P}(t) = P(t)Q \qquad \text{and} \qquad \dot{P}(t) = QP(t) \,, \tag{16.98}$$

respectively. We obtain $P(t) = \exp(Qt)$ where the exponential function of a matrix has to be interpreted as

$$\exp(Qt) = \sum_{k=0}^{\infty} \frac{t^k}{k!} Q^k \,, \tag{16.99}$$

where $Q^0 = I$ is the identity matrix.

We define s as the time of the first jump of our process for the particular case $X(0) = n$

$$s = \min\left[t | X(t) \neq X(0)\right] \,. \tag{16.100}$$

It can be shown that $P_n(s > t)$, the probability that the jump occurs at *some* time $s > t$, is given by

$$P_n(s > t) = \exp(q_{nn} t) \,, \tag{16.101}$$

where we note that $q_{nn} \leq 0$. Moreover,

$$P_n[X(s) = m] = -\frac{q_{nm}}{q_{nn}} \,, \tag{16.102}$$

and the process starts again at time s and in state m. This means that in a continuous-time MARKOV-chain the waiting times between two consecutive jumps are *exponentially distributed*. One of the simplest examples of a continuous time MARKOV-chain is the POISSON process, discussed in Sect. 16.3.

Summary

This chapter introduced the concept of stochastic processes $Y_X(t)$ as 'time' dependent processes depending on randomness. Y was a random variable which depended on another random variable X and t, the time. All realizations of $Y_X(t)$ spanned the state space. Each stochastic process was coupled to a pdf which described the probability that the process took on the realization y at time t. In the course of this introduction a series of general properties which classify such processes were defined. This was followed by the discussion of a particular class of stochastic processes, the MARKOV processes. They had the remarkable property that a future realization of the process solely depended on its current realization and not on the history how this current realization had been reached (MARKOV property). A huge class of processes in physics and related sciences is Markovian in nature. The next refinement in our discussion was the introduction of MARKOV-chains. These were processes defined on a discrete time span and in a discrete state space.

This allowed to replace the pdfs by probabilities. Again, various specific properties of MARKOV-chains opened the possibility of a distinctive classification. A very important observation was that under certain conditions a MARKOV-chain reached a stationary or equilibrium distribution and that it definitely arrived at this distribution independent of the choice of initial conditions. Moreover, detailed balance was obeyed by this equilibrium condition. This observation was the backbone of MARKOV-chain Monte Carlo sampling methods, in particular of the METROPOLIS algorithm. Finally, continuous-time MARKOV-chains were discussed.

Problems

1. Write a program to simulate the WIENER process in one dimension. Follow the method explained in Sect. 16.2 and perform the following analysis:
   a. Illustrate graphically some typical sample paths.
   b. Calculate the mean $\langle x(t)\rangle$ and the variance $\mathrm{var}[x(t)]$ by restarting the process several times with different seeds and plot the result.
   c. Measure the position x of the particle at a particular time t for several runs (with different seeds) and illustrate the result $p(x,t)$ graphically.
2. Realize numerically a POISSON process according to the instructions given in Sect. 16.2. Again, plot some typical sample paths. Moreover, calculate the mean waiting time $\langle\tau\rangle$ as well as the variance $\mathrm{var}(\tau)$ numerically as well as analytically.

References

1. Chow, Y.S., Teicher, H.: Probability Theory, 3rd edn. Springer Texts in Statistics. Springer, Berlin/Heidelberg (1997)
2. Kienke, A.: Probability Theory. Universitext. Springer, Berlin/Heidelberg (2008)
3. Stroock, D.W.: Probability Theory. Cambridge University Press, Cambridge (2011)
4. von der Linden, W., Dose, V., von Toussaint, U.: Bayesian Probability Theory. Cambridge University Press, Cambridge (2014)
5. van Kampen, N.G.: Stochastic Processes in Physics and Chemistry. Elsevier, Amsterdam (2008)
6. Gardiner, C.: Stochastic Methods. Springer Series in Synergetics. Springer, Berlin/Heidelberg (2009)
7. Chiasson, J.: Introduction to Probability Theory and Stochastic Processes. Wiley, New York (2013)
8. Breuer, H.P., Petruccione, F.: Open Quantum Systems, chap. 1. Clarendon Press, Oxford (2010)
9. Montroll, E.W., Schlesinger, M.F.: The wonderful world of random walks. In: Lebowitz, J.L., Montroll, E.W. (eds.) Non-equilibrium Phenomena II. Studies in Statistical Mechanics, vol. 11. North-Holland, Amsterdam (1984)
10. Rudnik, J., Gaspari, G.: Elements of the Random Walk. Cambridge University Press, Cambridge (2010)

11. Ibe, O.C.: Elements of Random Walk and Diffusion Processes. Wiley, New York (2013)
12. Marcus, M.B., Rosen, J.: Markov Processes, Gaussian Processes, and Local Times. Cambridge University Press, Cambridge (2011)
13. Papoulis, A., Pillai, S.: Probability, Random Variables and Stochastic Processes. McGraw Hill, New York (2001)
14. Knuth, D.: The Art of Computer Programming, vol. II, 3rd edn. Addison Wesley, Menlo Park (1998)
15. Ballentine, L.E.: Quantum Mechanics. World Scientific, Hackensack (1998)
16. Arnol'd, V.I.: Mathematical Methods of Classical Mechanics, 2nd edn. Graduate Texts in Mathematics, vol. 60. Springer, Berlin/Heidelberg (1989)
17. Scheck, F.: Mechanics, 5th edn. Springer, Berlin/Heidelberg (2010)
18. Goldstein, H., Poole, C., Safko, J.: Classical Mechanics, 3rd edn. Addison-Wesley, Menlo Park (2013)
19. Pathria, R.K., Beale, P.D.: Statistical Mechanics, 3rd edn. Academic, San Diego (2011)
20. Norris, J.R.: Markov Chains. Cambridge Series in Statistical and Probabilistic. Cambridge University Press, Cambridge (1998)
21. Modica, G., Poggiolini, L.: A First Course in Probability and Markov Chains. Wiley, New York (2012)
22. Graham, C.: Markov Chains: Analytic and Monte Carlo Computations. Wiley, New York (2014)
23. Fishman, G.S.: Monte Carlo: Concepts, Algorithms and Applications. Springer Series in Operations Research. Springer, Berlin/Heidelberg (1996)
24. Doucet, A., de Freitas, N., Gordon, N. (eds.): Sequential Monte Carlo Methods in Practice. Information Science and Statistics. Springer, Berlin/Heidelberg (2001)
25. Kalos, M.H., Whitlock, P.A.: Monte Carlo Methods, 2nd edn. Wiley, New York (2008)
26. Kroese, D.P., Taimre, T., Botev, Z.I.: Handbook of Monte Carlo Methods. Wiley, New York (2011)