Chapter 14
A Brief Introduction to Monte-Carlo Methods

14.1 Introduction

This chapter presents a brief introduction to Monte-Carlo methods in general, and to Monte-Carlo integration as well as to the METROPOLIS-HASTINGS algorithm in particular. A detailed discussion of the fundamental concepts involved is postponed to Chap. 16. The introduction given here is not supposed to be self-contained and methods will be introduced without reference to their background.

The notion of Monte-Carlo methods, Monte-Carlo algorithms or Monte-Carlo techniques is not well defined. In particular, the term *Monte-Carlo* summarizes a wide field of methods which are based on the sampling of random numbers [1–3]. In general, the advantage of Monte-Carlo algorithms lies in their computational strength. In many cases it is simply not feasible to employ deterministic methods due to their very high computational cost. However, in many cases the use of methods based on random sampling is also motivated by the nature of the processes to be described. We mentioned in the previous chapter as a typical example the radioactive decay of some nucleus. This process is believed to be purely stochastic in nature.

The development of Monte-Carlo techniques was initialized in the 1940s by J. VON NEUMANN, S. M. ULAM and N. METROPOLIS who coined the term *Monte-Carlo methods*. One of the earliest illustrations of the principle of Monte-Carlo techniques in general, and of Monte-Carlo integration in particular is the Monte-Carlo approximation of π. The discussion which follows now includes the essential ideas of Monte-Carlo integration.

We regard the unit square characterized by the corner points $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$. The area A_s of this square is one. We insert a quarter-circle of radius $r = 1$ which, consequently, possesses the area $A_c = \pi/4$. Suppose, we are throwing darts on this unit square in such a way that the impact points are uniformly distributed;

B.A. Stickler, E. Schachinger, *Basic Concepts in Computational Physics*,
DOI 10.1007/978-3-319-27265-8_14

then the probability P that a certain dart becomes stuck within the interior of the quarter-circle is given by

$$P = \frac{A_c}{A_s} = A_c = \frac{\pi}{4} = 0.785398\ldots\ . \tag{14.1}$$

From a probabilistic point of view, we have after N throws of which n hit the interior of the quarter-circle the probability:

$$P = \lim_{N\to\infty} \frac{n}{N}\ . \tag{14.2}$$

The strategy is clear: we draw uniformly distributed random numbers x_i, y_i from the interval $[0, 1]$. These are the intersection points of the darts. Repeating this *experiment* several times and counting the number of hits n within the quarter-circle allows us to approximate π via

$$P = \frac{\pi}{4} \approx \frac{n}{N}\ . \tag{14.3}$$

The resulting approximation of π will be strongly influenced by the number of experiments N as well as by the performance of the random number generator used. Table 14.1 lists computed approximations of π for different numbers of experiments N as they were obtained with the help of a linear congruential generator. Linear congruential generators have been introduced and discussed in Sect. 12.2. The parameters used to initialize the generators are given in the caption of the table. Furthermore, Fig. 14.1 illustrates the result after $N = 10^3$ experiments for both generators.

Table 14.1 Approximate values $\pi_a^{(i)}$ obtained with the method discussed in the text. The linear congruential generators are initialized by the following parameters: generator (1): $a = 7^5$, $c = 0$, $m = 2^{31} - 1$, and $x_0 = 281$ (PARK-MILLER) and generator (2): $a = 7^5$, $c = 0$, $m = 2^{11}$, and $x_0 = 281$. We also give the absolute errors $|\pi_a^{(i)} - \pi|$

N	$\pi_a^{(1)}$	$\lvert\pi_a^{(1)} - \pi\rvert$	$\pi_a^{(2)}$	$\lvert\pi_a^{(2)} - \pi\rvert$
10	2.8000	0.34159	2.8000	0.34159
10^2	2.9200	0.22159	3.1600	0.01841
10^3	3.1600	0.01841	3.1840	0.04241
10^4	3.1304	0.01119	3.1868	0.04521
10^5	3.1358	0.00579	3.1875	0.04589
10^6	3.1393	0.00229	3.1875	0.04599
10^7	3.1413	0.00028	3.1875	0.04599

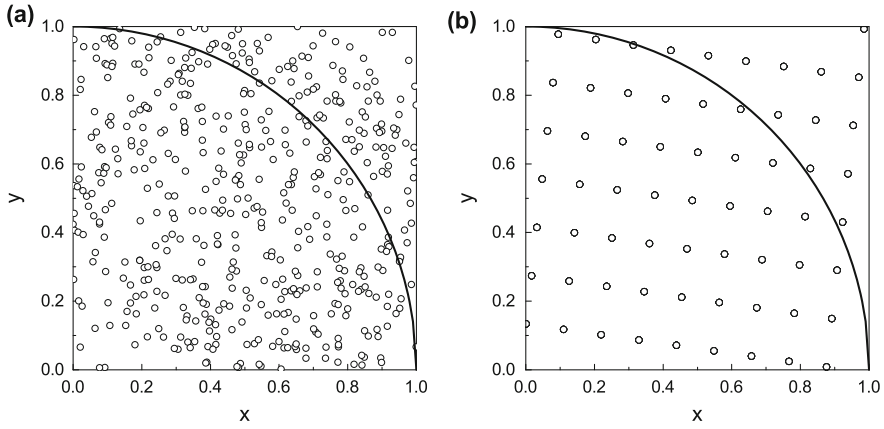

Fig. 14.1 $N = 10^3$ uniformly distributed random numbers within the unit-square. Frame **(a)** gives the results for generator (1) while frame **(b)** is for generator (2). The number of elements within the quarter-circle indicated by the *solid line* determines the value of $\pi_a^{(i)}$. The inferior result of the approximation obtained with generator (2) [frame **(b)**] originates in correlations between the x and y coordinates

14.2 Monte-Carlo Integration

We generalize the ideas formulated above and consider a function $f(x) \geq 0$ for $x \in [a, b] \subset \mathbb{R}$ where the area of interest is

$$A = \int_a^b \mathrm{d}x f(x) \ . \tag{14.4}$$

We denote

$$\xi = \max_{x \in [a,b]} f(x) \ , \tag{14.5}$$

and obtain using the above example

$$A = A_s \lim_{N \to \infty} \frac{n}{N} \ , \tag{14.6}$$

where n is the number of random points under the curve indicated schematically in Fig. 14.2. The area A_s is given by

$$A_s = (b - a)\xi \ , \tag{14.7}$$

and the random numbers $r_i = (x_i, y_i)$ are uniformly distributed within the intervals $x_i \in [a, b]$ and $y_i \in [0, \xi]$. This method is referred to as *hit and miss integration* [4].

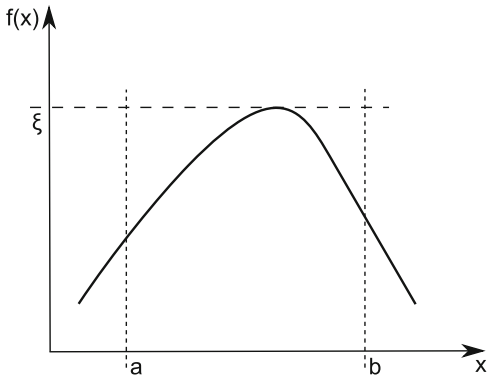

Fig. 14.2 Schematic illustration of the Monte-Carlo integration technique

Another way to perform a Monte-Carlo integration is the so called *mean-value integration*. It is essentially based on the mean value theorem of calculus which we already employed in our discussion of quadrature in Chap. 3. We restate it here for the sake of a more transparent presentation: The mean-value theorem states that if $f(x)$ is a continuous function for $x \in [a, b]$ then there exists a $z \in (a, b)$ such that

$$\int_a^b \mathrm{d}x f(x) = f(z)(b-a) \; . \tag{14.8}$$

The function value $f(z) \equiv \langle f \rangle$ is referred to as the *expectation value* or *mean value* of $f(x)$. We know from probability theory [5–7] that the expectation value can be approximated by the arithmetic mean $\overline{f}$

$$\frac{1}{b-a}\int_a^b \mathrm{d}x' f(x') \simeq \overline{f} \pm \sqrt{\frac{\overline{f^2} - \overline{f}^2}{N}} \; , \tag{14.9}$$

with the error given by the standard error, Eq. (E.14). The arithmetic mean $\overline{f}$, on the other hand, is given by

$$\overline{f} = \frac{1}{N}\sum_{i=1}^{N} f(x_i) \; , \tag{14.10}$$

and consequently

$$\overline{f^2} = \frac{1}{N}\sum_{i=1}^{N} f^2(x_i) \; . \tag{14.11}$$

Note that here the variables x_i are assumed to be uniformly distributed random numbers within the interval $[a, b]$. (This result will immediately be discussed in

more detail.) However, first of all we note from the law of large numbers, Eq. (E.25), that this approach is exact in the limit $N \to \infty$:

$$\frac{1}{b-a}\int_a^b \mathrm{d}x' f(x') = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} f(x_i)\ . \tag{14.12}$$

Let us now consider the more general case which, in the end, will guide us to a very prominent formulation of Monte-Carlo integration. We want to estimate the expectation value

$$\langle f\rangle = \int \mathrm{d}x f(x)p(x)\ , \tag{14.13}$$

where $x \in \mathbb{R}^d$ and $p(x)$ is a pdf. A typical example is the calculation of the thermal expectation value in statistical physics where the pdf $p(x)$ is given by the normalized BOLTZMANN distribution

$$p(x) = \frac{1}{Z}\exp\left[-\frac{E(x)}{k_\mathrm{B}T}\right]\ . \tag{14.14}$$

Here $E(x)$ denotes the energy as a function of the parameter $x \in \mathbb{R}^d$, k_B stands for BOLTZMANN's constant, T is the temperature, and the normalization factor Z is referred to as the canonical partition function [8–11].

Equation (14.13) may be rewritten as

$$\langle f\rangle = \int \mathrm{d}x f(x)p(x) = \int \mathrm{d}f\, f\, q(f)\ , \tag{14.15}$$

where we introduced the probability density $q(f)$ of f via

$$q(f) = \int \mathrm{d}x\, \delta\left[f - f(x)\right] p(x)\ , \tag{14.16}$$

with $\delta(\cdot)$ DIRAC's δ-distribution. Let us briefly explain how we arrived at this definition. Let the cdf $P(x)$ be defined by[1]

$$P(x) = \Pr(X \le x) = \int_{-\infty}^{x} \mathrm{d}x\, p(x)\ . \tag{14.17}$$

[1] Please note that according to the conventions established in Appendix E capital letters denote random variables.

We define in analogy the cdf $Q(f)$:

$$Q(f) = \Pr(F \leq f) = \Pr[f(X) \leq f] \,. \tag{14.18}$$

Note that we distinguish between the function $f(X)$ of the random variable X [which follows the pdf $p(X)$] and the particular function value $f \in \mathbb{R}$. Furthermore, the probability $\Pr[f(X) \leq f]$ can be rewritten as

$$\Pr[f(X) \leq f] = \sum_n \Pr(a_n \leq X \leq b_n) \,, \tag{14.19}$$

where the values $a_n < b_n$ are the ordered intersection points $a_1 < b_1 < a_2 < b_2 < \ldots < a_N < b_N$ chosen in such a way that

$$f(a_n) = f(b_n) = f \,, \quad \text{and} \quad f[x \in (a_n, b_n)] < f \,. \tag{14.20}$$

It is a matter of the particular form of $f(x)$ whether or not the boundary points have to be included. Equation (14.19) can be rewritten:

$$\Pr(a_n \leq X \leq b_n) = P(b_n) - P(a_n) = \int_{a_n}^{b_n} \mathrm{d}x\, p(x) \,. \tag{14.21}$$

The pdf $q(f)$ is related to the cdf $Q(f)$ via

$$q(f) = \frac{\mathrm{d}}{\mathrm{d}f} Q(f) \,, \tag{14.22}$$

and we obtain

$$\begin{aligned} q(f) &= \sum_n \frac{\mathrm{d}}{\mathrm{d}f} \Pr(a_n \leq X \leq b_n) \\ &= \sum_n \frac{\mathrm{d}}{\mathrm{d}f} \int_{a_n}^{b_n} \mathrm{d}x\, p(x) \\ &= \sum_n \left[\frac{\mathrm{d}b_n}{\mathrm{d}f} p(b_n) - \frac{\mathrm{d}a_n}{\mathrm{d}f} p(a_n) \right] \\ &= \sum_n \left[\left(\frac{\mathrm{d}f(x)}{\mathrm{d}x} \right)^{-1} p(x) \bigg|_{x=b_n} - \left(\frac{\mathrm{d}f(x)}{\mathrm{d}x} \right)^{-1} p(x) \bigg|_{x=a_n} \right] . \end{aligned} \tag{14.23}$$

However, we know from Eq. (14.20) that:

$$\left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=b_n} \overset{!}{>} 0 \qquad \text{and} \qquad \left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=a_n} \overset{!}{<} 0 \,. \tag{14.24}$$

We introduce the intersection points x_k where $x_1 < x_2 < \ldots < x_K$ and $K = 2N$ (if the boundary points are not included) for which $f(x_1) = f(x_2) = \ldots = f(x_K) = f$. Hence, Eq. (14.23) may be rewritten as

$$\begin{aligned} q(f) &= \sum_k \left|\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|^{-1}_{x=x_k} p(x_k) \\ &= \sum_k \frac{p(x_k)}{|f'(x_k)|} \,. \end{aligned} \tag{14.25}$$

We want to improve this result and remember that the DIRAC δ-distribution of an arbitrary function $g(y)$ can be expressed as [12]

$$\delta[g(y)] = \sum_i \frac{\delta(y - y_i)}{|g'(y_i)|} \,, \tag{14.26}$$

where the y_i are the zeros of $g(y)$, i.e. $g(y_i) = 0$. Hence, we arrive at the final form of Eq. (14.23)[2]:

$$q(f) = \int \mathrm{d}x \delta\left[f - f(x)\right] p(x) \,. \tag{14.27}$$

We note, furthermore, that:

$$\int \mathrm{d}f \, q(f) = \int \mathrm{d}f \int \mathrm{d}x \delta\left[f - f(x)\right] p(x) = \int \mathrm{d}x \, p(x) = 1 \,. \tag{14.28}$$

[2] We give an example. Suppose $f(x) = \exp(x)$. Then we deduce that

$$\delta\left[f - \exp(x)\right] = \frac{\delta(x - \ln f)}{f} \,,$$

and, consequently,

$$q(f) = \frac{p(\ln f)}{f} \,.$$

A second example was given in Chap. 13 where we derived the pdf (13.5).

As a result, the variance of f, var (f), can be expressed as

$$\text{var}\,(f) = \int \mathrm{d}x\,[f(x) - \langle f\rangle]^2\, p(x) = \int \mathrm{d}f\,[f - \langle f\rangle]^2\, q(f)\ . \tag{14.29}$$

Let us define in a next step the arithmetic mean of $f(X)$

$$\mathscr{F} = \frac{1}{N}\sum_{i=1}^{N} f(x_i) = \frac{1}{N}\sum_{i=1}^{N} f_i\ , \tag{14.30}$$

calculated with the help of N random numbers. Hence, we have

$$\langle \mathscr{F}\rangle = \langle f\rangle\ , \tag{14.31}$$

and

$$\text{var}\,(\mathscr{F}) = \frac{\text{var}\,(f)}{N}\ , \tag{14.32}$$

according to Appendix E. It follows from the central limit theorem, Appendix Sect. E.8, that for large values of N, the pdf of $\mathscr{F}$, $p(\mathscr{F})$ converges to a normal distribution (E.43) with $\langle \mathscr{F}\rangle$ and var $(\mathscr{F})$:

$$p(\mathscr{F}) \approx \mathscr{N}\left(\mathscr{F}\,\middle|\,\langle f\rangle\,, \frac{\text{var}\,(f)}{N}\right)\ . \tag{14.33}$$

Based on this property $\langle f\rangle$ can be estimated from:

$$\langle f\rangle = \mathscr{F} \pm \sqrt{\frac{\text{var}\,(f)}{N}} = \frac{1}{N}\sum_{i=1}^{N} f(x_i) \pm \sqrt{\frac{\text{var}\,(f)}{N}}\ . \tag{14.34}$$

Here, the random numbers x_i are sampled from the pdf $p(x)$. This method is the most prominent formulation of Monte-Carlo integration.

We shall briefly discuss some properties of this method. We deduce from Eq. (14.34) that the error scales like $N^{-\frac{1}{2}}$. In contrast to the integration methods we discussed in Chap. 3, N is no longer the number of grid-points but the number of random numbers sampled.[3] In principle, the error scaling is worse than in the case of classical integrators. For instance, in the case of the central rectangular rule (Sect. 3.2) we had an error scaling of N^{-2} when summed over the whole interval. However, we obtained this result for the one-dimensional case, in higher dimensions we will certainly need much more grid-points. On the other hand, in Eq. (14.34) N

[3] Nevertheless, there is certainly some conceptual similarity between grid-points and random numbers within this context.

corresponds to the number of d-dimensional random numbers x. Hence, Monte-Carlo integration can be of advantage whenever one has to deal with complicated, high dimensional integrals. In contrast, restricted to one dimension it is in most cases not an improvement of the methods discussed already.

Monte-Carlo integration can also be of advantage whenever the integrand $f(x)$ is not well behaved. In such a case a very fine grid would be required to compute a reasonable estimate of the true value of the integral. Monte-Carlo integration offers a very convenient alternative due to its conceptual simplicity [13].

It is certainly a drawback of Monte-Carlo integration in its formulation (14.34), that the error is also proportional to $\sqrt{\mathrm{var}\,(f)}$ which is a yet unknown quantity. One has to approximate it with an adequate estimator, for instance with the help of the sampling variance. Moreover, if the variance $\mathrm{var}\,(f)$ diverges, the central limit theorem does not hold and the procedure (14.34) is no longer justified and will fail for sure.

Closely related to the problem of how to determine $\mathrm{var}\,(f)$, is the question of how many random numbers should be drawn. In most cases an iterative approach is the most promising strategy. In a first step N random numbers are drawn and the integral is computed using Eq. (14.34). Then another set of N random numbers is sampled and Eq. (14.34) is reevaluated now using all $2N$ random numbers. If the change in the resulting estimate of the integral is less than some given tolerance ϵ, the loop is terminated otherwise another set of N random numbers is added.

We mention that this form of Monte-Carlo integration can be improved particularly by sampling only from points which dominantly contribute to the integral. This method is referred to as *importance sampling* [13–16] and will be discussed in more detail later on.

14.3 The METROPOLIS Algorithm: An Introduction

The METROPOLIS algorithm is a more sophisticated method to produce random numbers from given distributions. In fact, the METROPOLIS algorithm is a special form of the rejection method (Sect. 13.3). This section introduces the algorithm on a very basic level which will, in the end, allow a first glance at an interesting model out of statistical physics, namely the ISING model. It will be discussed in Chap. 15 and a more detailed discussion of the METROPOLIS algorithm will be postponed to Sect. 16.4.

The METROPOLIS algorithm is particularly useful to treat problems in statistical physics where thermodynamic expectation values of some observable O are of interest [8–11]. They are defined as

$$\langle O \rangle = \int \mathrm{d}x O(x) q(x) \,, \tag{14.35}$$

where x is a set of parameters and $q(x)$ is the BOLTZMANN distribution (14.14). The set of parameters x could be, for instance, the position- and momentum-space coordinates of N different particles. In most cases x is a high dimensional object which makes classical numerical integration (Chap. 3) cumbersome. Instead Monte-Carlo integration is employed and the integral (14.35) is approximated with the help of Eq. (14.34) by

$$\langle O \rangle \approx \frac{1}{N} \sum_{i=1}^{N} O(x_i) \pm \sqrt{\frac{\mathrm{var}\,(O)}{N}} \,, \tag{14.36}$$

where the uncorrelated random numbers x_i, $i = 1, 2, \ldots, N$ are sampled from the pdf, Eq. (14.14). We recognize immediately the problem: we need to know the exact functional form of $q(x)$ if we want to apply either the inverse transformation method or the rejection method discussed in Chap. 13. However, the partition function Z itself is determined by an integral which can be approximated using Eq. (14.36). We set

$$q(x) = \frac{p(x)}{Z} \,, \tag{14.37}$$

and

$$Z = \int \mathrm{d}x\, p(x) \tag{14.38}$$

follows from the normalization of $q(x)$. The METROPOLIS algorithm was designed to avoid precisely this problem. We concentrate on a pdf which is of the form (14.37), but $q(x)$ must not necessarily be described by a normalized BOLTZMANN distribution, Eq. (14.14). Thus, $p(x)$ is arbitrary but it ensures that

$$\int \mathrm{d}x\, q(x) = 1 \Longleftrightarrow \int \mathrm{d}x\, p(x) = Z \,, \tag{14.39}$$

and $q(x) \geq 0$ for all x. In other words, $q(x)$ is a pdf. Suppose we already have a sequence $x_0, x_1, \ldots, x_n = \{x_n\}$ of parameters which indeed follows the pdf $q(x)$.[4] We now add to the last element of this sequence x_n a small perturbation δ and set

$$x_t = x_n + \delta \,. \tag{14.40}$$

Note that the perturbation δ is of the same dimension as the vector x. Similar to the rejection method we seek for a criterion which helps us to decide whether or not the test value x_t can be accepted as the next element of the sequence $\{x_n\}$.

[4] The question of how one can obtain such a sequence will be discussed in Sect. 16.3.

The METROPOLIS method proposes an acceptance probability of the form

$$\Pr(A|x_t, x_n) = \begin{cases} 1 & \text{if} \quad \dfrac{q(x_t)}{q(x_n)} \geq 1 \ , \\ \dfrac{q(x_t)}{q(x_n)} & \text{otherwise.} \end{cases} \tag{14.41}$$

Hence, if $\Pr(A|x_t, x_n) = 1$, we set $x_{n+1} = x_t$, and if $\Pr(A|x_t, x_n) < 1$, we draw a random number $r \in [0, 1]$ and accept x_t if $r \leq \Pr(A|x_t, x_n)$ and reject x_t otherwise. We note that in this formulation the knowledge of the normalization factor Z is no longer required since it follows from Eq. (14.37) that

$$\frac{q(x_t)}{q(x_n)} = \frac{p(x_t)}{p(x_n)} \ . \tag{14.42}$$

Consequently we rewrite Eq. (14.41) as

$$\Pr(A|x_t, x_n) = \min\left(\frac{p(x_t)}{p(x_n)}, 1\right) = p(x_t|x_n) \ , \tag{14.43}$$

where we introduced in the last step a more compact notation.

A discussion of the underlying concepts and why the choice (14.41) indeed samples random numbers according to the pdf $q(x)$ requires some basic knowledge of stochastics in general and of MARKOV-chains in particular. This is the reason why we postponed this discussion to Chap. 16. Nevertheless, there is a particular property, referred to as *detailed balance* which requires our attention because it is crucial for the METROPOLIS algorithm: Let $p(x_t|x_n)$ denote the pdf for the probability that a random number x_t is generated from the random number x_n as defined in Eq. (14.43). Then the condition of detailed balance is defined as

$$p(x_t|x_n)q(x_n) = p(x_n|x_t)q(x_t) \ . \tag{14.44}$$

In words: The probability $p(x_t|x_n)$ that a random number x_t is generated from a random number x_n times the probability $q(x_n)$ that the random number x_n occurred at all is equal to the probability $p(x_n|x_t)$ that the random number x_n is generated from x_t times the probability $q(x_t)$ that x_t occurred. Detailed balance is motivated by physics and is a condition of thermodynamic equilibrium.

Let us briefly demonstrate that the METROPOLIS algorithm (14.43) satisfies detailed balance: We distinguish three different cases: (i) Suppose that $p(x_t|x_n) = p(x_n|x_t) = 1$. From Eq. (14.43) we note that this is only possible if $p(x_t) = p(x_n)$ and therefore $q(x_t) = q(x_n)$ which is already Eq. (14.44) for this particular case. (ii)

We assume that $p(x_t|x_n) = 1$ but $p(x_n|x_t) \neq 1$. It follows from Eq. (14.43) that

$$\begin{aligned} p(x_n|x_t)q(x_t) &= \frac{p(x_n)}{p(x_t)}q(x_t) \\ &= q(x_n) \ . \end{aligned} \tag{14.45}$$

This corresponds to Eq. (14.44) for $p(x_t|x_n) = 1$. Note that we made use of definition (14.37) in order to achieve this result. (iii) Finally, we find for $p(x_n|x_t) = 1$ and $p(x_t|x_n) \neq 1$ that

$$\begin{aligned} p(x_t|x_n)q(x_n) &= \frac{p(x_t)}{p(x_n)}q(x_n) \\ &= q(x_t) \ , \end{aligned} \tag{14.46}$$

which, again, is Eq. (14.44). Hence, the METROPOLIS algorithm (14.43) indeed obeys detailed balance.

So far the question of how to choose the initialization point x_0 of the sequence stayed unanswered. This is clearly not a trivial problem and it is strongly related to one of the major disadvantages of the METROPOLIS algorithm, namely that subsequent random numbers (x_n, x_{n+1}) are strongly correlated. One of the most pragmatic approaches is to choose a starting point x_0 at random out of the parameter space and then discard it together with the first few members of the sequence. This approach is strongly motivated by a clear physical picture: The sequence of random numbers resembles the evolution of the physical system from an arbitrary initial point x_0 toward equilibrium which manifests itself in the condition of detailed balance. Hence, the approach of discarding the first few members of the sequence is referred to as *thermalization*.

The integral of interest, Eq. (14.35) is then approximated with the help of Eq. (14.36), where the random numbers $x_k, x_{k+1}, \ldots, x_{k+N}$ are used, if the thermalization required k steps. There is a remedy which helps to reduce correlations between subsequent random numbers within the sequence which is based on a similar strategy. In particular, the modified sequence

$$x_k, x_{k+\ell}, x_{k+2\ell}, \ldots \ , \tag{14.47}$$

generated by discarding ℓ intermediate random numbers will reduce correlations between the members of this final sequence of random numbers.

Summary

This chapter set the stage for an important numerical tool in Computational Physics: the Monte-Carlo techniques. It started with the conceptual transparent task of how to calculate π using a sequence of uniformly distributed random numbers of the

range [0, 1]. This established the so-called hit and miss technique. It moved on to a discussion of Monte-Carlo integration in a more formal way and discussed in detail the error involved by this type of integration as opposed to the error experienced by deterministic methods. The conclusion was, that Monte-Carlo integration was certainly preferable whenever estimates of high dimensional integrals were required and it also had advantages when the integrand was heavily structured. The second part of this chapter dealt with the METROPOLIS algorithm which allowed to generate a sequence of random numbers from some pdf $p(x)$. It was conceptually similar to the rejection method discussed earlier. The mathematical background which is more involved was not discussed within this first contact with the METROPOLIS algorithm. Instead, the emphasis was to demonstrate that this algorithm obeyed detailed balance a property purely based on physics as a condition of thermodynamic equilibrium. It was, furthermore, pointed out that the random numbers generated by this algorithm were highly correlated and some strategies to remedy this problem were discussed.

References

1. Gentle, J.E.: Random Number Generation and Monte Carlo Methods. Statistics and Computing. Springer, Berlin/Heidelberg (2003)
2. Kalos, M.H., Whitlock, P.A.: Monte Carlo Methods, 2nd edn. Wiley, New York (2008)
3. Kroese, D.P., Taimre, T., Botev, Z.I.: Handbook of Monte Carlo Methods. Wiley, New York (2011)
4. Fishman, G.S.: Monte Carlo: Concepts, Algorithms and Applications. Springer Series in Operations Research. Springer, Berlin/Heidelberg (1996)
5. Chow, Y.S., Teicher, H.: Probability Theory. Springer Texts in Statistics, 3rd edn. Springer, Berlin/Heidelberg (1997)
6. Kienke, A.: Probability Theory. Universitext. Springer, Berlin/Heidelberg (2008)
7. Stroock, D.W.: Probability Theory. Cambridge University Press, Cambridge (2011)
8. Schwabl, F.: Statistical Mechanics. Advanced Texts in Physics. Springer, Berlin/Heidelberg (2006)
9. Halley, J.W.: Statistical Mechanics. Cambridge University Press, Cambridge (2006)
10. Pathria, R.K., Beale, P.D.: Statistical Mechanics, 3rd edn. Academic, San Diego (2011)
11. Hardy, R.J., Binek, C.: Thermodynamics and Statistical Mechanics: An Integrated Approach. Wiley, New York (2014)
12. Ballentine, L.E.: Quantum Mechanics. World Scientific, Hackensack (1998)
13. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes in C++, 2nd edn. Cambridge University Press, Cambridge (2002)
14. Doucet, A., de Freitas, N., Gordon, N. (eds.): Sequential Monte Carlo Methods in Practice. Information Science and Statistics. Springer, Berlin/Heidelberg (2001)
15. Ripley, B.D.: Stochastic Simulation. Wiley, New York (2006)
16. Landau, D.P., Binder, K.: A Guide to Monte Carlo Simulations in Statistical Physics, 3rd edn. Cambridge University Press, Cambridge (2009)