

Chapter 5

Monte Carlo Method

The Monte Carlo (MC) method is associated with a whole class of numerical methods, where the using of the random number generator (RNG) plays the central role. The Monte Carlo method together with the Molecular Dynamics method is the main method of computer simulation, especially for investigation of systems in the thermal equilibrium state.

In fact, we already met with the MC method in the previous Chapter 4 when we discussed the numerical solution of the Langevin equation (more exactly, the method of solution of the Langevin equation is more logically to call as the hybrid of MD and MC methods). Recall that in the solution of the Langevin equation the RNG has been used to imitate the fluctuating force acting on the subsystem under investigation from the thermostat. This example may lead to an impression that the MC method is an imitation one, i.e. that it gives some approximate description of a real system. But this impression is wrong. All approximations are connected with the model under consideration only, while the MC method gives the mathematically rigorous solution of the model with any preliminary stated accuracy.

In a general case all MC methods reduce to the same scheme. Namely, let we need to calculate the value of a variable a . To do this, we have to invent a random variable ξ such that the mathematical expectation of ξ will coincide with the value of a ,

$$a = \langle \xi \rangle \equiv M\{\xi\}, \quad (5.1)$$

where

$$M\{\xi\} = \frac{1}{N} \sum_{n=1}^N \xi_n \text{ at } N \rightarrow \infty. \quad (5.2)$$

According to the probability theory, for a finite N the computational error of the MC method will always be proportional to $N^{-1/2}$. Consequently, the length of the sequence $\{\xi_n\}$ must be very long as a rule.

Naturally the following questions arise:

1. In what a way should ξ be chosen, or, more specifically, what should the corresponding probability distribution $P(\xi)$ be equal to?
2. How can the corresponding Markov chain, i.e. the sequence of random numbers with the given distribution $P(\xi)$, be constructed?

Actually the MC method has found a wide application with appearing fast electronic computers only. Metropolis and Ulam [121] were the first who have done this in 1949.

5.1 Generation of random numbers

5.1.1 Generation of the standard random numbers

Evidently that for exploring the MC method, first of all we are in need of a generator of random numbers. Here let us consider the so called *standard random number generator* (SRNG), and then we will show how we can get random numbers with any given probability distribution $P(\xi)$ from the standard random numbers.

The *standard random numbers* (SRNs) are the random numbers γ uniformly distributed within the unit interval $(0,1)$. The sequence of the SRNs $\{\gamma_n\}$ must satisfy the following four requirements:

(1) Its probability distribution must be equal one,

$$P(\gamma) = 1. \quad (5.3)$$

(2) Its mathematical expectation should be the following:

$$M\{\gamma\} \equiv \langle \gamma \rangle = \frac{1}{2}. \quad (5.4)$$

(3) The dispersion has to be equal to

$$D\{\gamma\} \equiv \langle (\gamma - \langle \gamma \rangle)^2 \rangle = \frac{1}{12}. \quad (5.5)$$

Recall that according to the definition, the dispersion for a function $A(\xi)$ of a random variable ξ , which is defined in an interval (a, b) with a probability distribution $P(\xi)$, is equal to

$$D\{A\} \equiv \langle (A - \langle A \rangle)^2 \rangle = \int_a^b d\xi P(\xi) [A(\xi) - M\{A\}]^2, \quad (5.6)$$

where

$$M\{A\} \equiv \langle A \rangle = \int_a^b d\xi P(\xi) A(\xi). \quad (5.7)$$

(4) There must be no correlation between different elements of the sequence $\{\gamma_n\}$, i.e., it should be fulfilled the relationship

$$\left\langle \left(\gamma_i - \frac{1}{2} \right) \left(\gamma_k - \frac{1}{2} \right) \right\rangle = \left\langle \gamma_i - \frac{1}{2} \right\rangle \left\langle \gamma_k - \frac{1}{2} \right\rangle = 0 \quad (5.8)$$

for any $i \neq k$.

In what a way can we get random numbers with a computer? The decisive factor here is the fact that any application of the MC method to physical problems requires a very long sequence of random numbers. Consequently, such methods of obtaining the random numbers (RNs) as, for example, their getting from a preliminary prepared table, or the setting of the RNs with an external noise generator, occur to be ineffective (a large table cannot be stored, and for the noise generator it is too difficult to control the quality of the obtained RNs).

The only realistic method turns to be the calculation of the RNs with the recurrence first-order formula

$$\gamma_{n+1} = f(\gamma_n). \quad (5.9)$$

Clearly that in this way we can get only pseudo-random numbers but not the true random numbers. The sequence $\{\gamma_n\}$ calculated with Eq. (5.9), always has a finite period simply in consequence of that the set of different numbers available with a computer, is finite. Therefore, if some number in the sequence $\{\gamma_n\}$ is been met in the second time, then according to Eq. (5.9) all next numbers will be repeated too. Thus, the problem is to find such a function $f(\gamma)$ that the period of the obtained sequence $\{\gamma_n\}$ will be as long as possible.

In order to choose an appropriate function $f(\gamma)$ in Eq. (5.9), let us combine the consequent random numbers $\gamma_1, \gamma_2, \dots$ in pairs $(\gamma_1, \gamma_2), (\gamma_3, \gamma_4), \dots$, and let us consider these pairs as coordinates of points in a unit square (see Fig. 5.1). Clearly that these points must uniformly fill in the whole square, i.e., the graph of the function $y = f(x)$ should fill in the square as dense as possible. On the other hand, the function $f(x)$ has to be maximally simple (for example, linear), because we need a very large amount of RNs. To satisfy these two requirements, we may take the function

$$f(x) = gx \pmod{1} \quad (5.10)$$

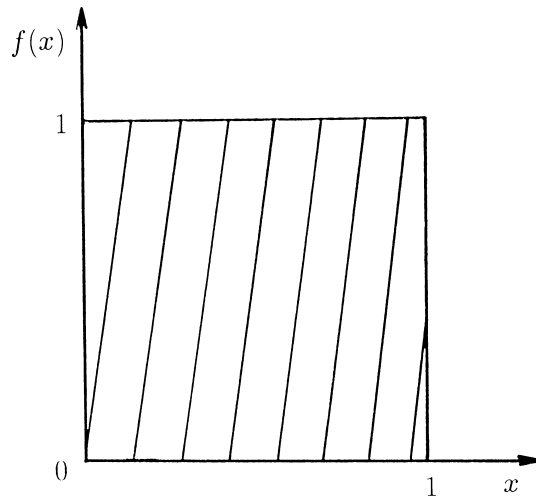


Figure 5.1: The function $f(x) = gx \pmod{1}$ for the random number generator.

with a sufficiently large g ($g \gg 1$) as is shown in Fig. 5.1. Recall that we met already Eq. (5.10) in Chapter 2 devoted to the stochastic theory. Namely, Eq. (5.10) for $g = 2$ is nothing else that the logistic map (2.9) for the $r = 4$ case, when its strange attractor takes the whole interval $(0,1)$. In fact, instead of Eq. (5.10) it is more convenient to use the following recurrence formula for integer numbers m_n (D. Lemer):

$$m_{n+1} = gm_n \pmod{M}. \quad (5.11)$$

Here \pmod{M} is the operation of taking of the integer remainder from the division by M , g and M have to be relatively simple, and the inequality $g < \sqrt{M}$ should be satisfied. The needed standard random numbers γ are then obtained with the equation

$$\gamma_n = m_n/M. \quad (5.12)$$

The described method is known as the *congruential method*, or the *residual method*. Note that the period of the sequence $\{\gamma_n\}$ obtained in this way, cannot exceed M , and that the sequence itself depends on the choice of the initial value m_0 . Usually for the 32-bit computers one takes $M = 2^{31} - 1 = 2147483647$ (the so-called Mersenn's number) and $g = 7^5 = 16807$ (another choice is $g = 48271$ or $g = 69621$ [185]), for the 36-bit computers, $M = 2^{36}$ and $g = 5^{15}$, for the 40-bit computers, $M = 2^{40}$ and $g = 5^{17}$, and for the 49-bit computers, $M = 2^{48}$ and $g = 11^{13}$. In the case of $g \ll M$, however, the algorithm (5.11) is not too good — if one meets with a small γ_n value, the values of the following random numbers will be small too. Because the multiplication and especially the division are comparative slow computer operations, it is better to write a subroutine for the random number generator using computer instructions. Concrete programs for RNGs as well as the bibliography on this question may be found, e.g., in Heerman's book [1] (see also [9, 185, 186]).

The main requirements to the SRNG in the MC method are the following:

(a) The random number generator should work with high speed, and the period of the resulting sequence must be sufficiently long. For this, it is desirable to use the 32-bit computer (or, better, the 64-bit computer), and the double precision would be used.

(b) The each RNG must be tested before its exploration, i.e., the validity of the conditions (5.3), (5.4), (5.5), and (5.8) must be verified. This point is especially actual in using the RNG developed for a computer of one type, on the computer of an another type. At this, it is desirable to test not only the whole sequence $\{\gamma_n\}$, but also its separate sections (consisting, for example, of 2000 sequent elements), as well as to test, whether the period of the random sequence is sufficiently long for the problem under investigation. The detailed description of different tests for checking the quality of RNs may be found, e.g., in the book of Sobol' [122]. However, it must be noted that any test cannot guarantee the getting of "good" random numbers, although it can save from the using of wittingly "bad" sequences at least.

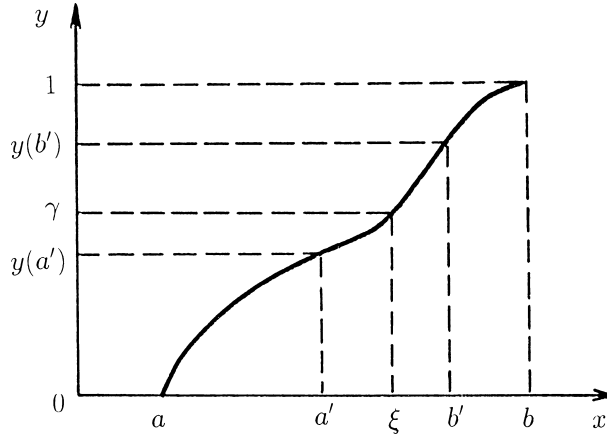


Figure 5.2: The transformation of the standard random numbers γ into the random numbers ξ with a given probability distribution $P(\xi)$.

(c) To have a more reliable results, it is desirable to carry out the MC simulation with different SRNGs. For example, some runs may be carried out not only with the “main” SRNG, but also with an another SRNG, and then the results have to be compared.

5.1.2 Generation of random numbers with a given distribution

Often in the MC simulation we need a sequence of random numbers $\{\xi_n\}$ defined in an interval (a, b) with a given probability distribution $P(\xi)$. Let us prove that the values of ξ can be found from the equation

$$\int_a^\xi P(x) dx = \gamma, \quad (5.13)$$

i.e., to get a next random number ξ , we have to take the next standard random number γ and then to solve the equation (5.13).

To prove, let us consider the function

$$y(x) = \int_a^x P(x) dx. \quad (5.14)$$

According to the definition of the probability distribution, we have

$$P(x) > 0 \text{ and } \int_a^b P(x) dx = 1. \quad (5.15)$$

Thus, it follows that $y(a) = 0$, $y(b) = 1$, and $y'(x) = P(x) > 0$, i.e., the function $y(x)$ monotonically increases from 0 to 1 as shown in Fig. 5.2. Consequently, any line $y = \gamma$, where $0 < \gamma < 1$, intersects the graph of the function $y = f(x)$ in one and only one point (the abscissa of which is just taken as ξ), and Eq. (5.13) always has a single solution only.

Now let take an arbitrary interval (a', b') inserted into the interval (a, b) . Points from the interval $a' < x < b'$ are in one-to-one correspondence with the ordinates of the function $y(x)$, and these ordinates satisfy the inequality $y(a') < y < y(b')$. Therefore, if ξ belongs to the interval $a' < x < b'$, the corresponding γ is to belong to the interval $y(a') < y < y(b')$ and vice versa (see Fig. 5.2). Thus,

$$\text{Prob}\{a' < \xi < b'\} = \text{Prob}\{y(a') < \gamma < y(b')\}. \quad (5.16)$$

But because γ is uniformly distributed within the interval $(0, 1)$, we have

$$\text{Prob}\{y(a') < \gamma < y(b')\} = y(b') - y(a') = \int_{a'}^{b'} P(x) dx. \quad (5.17)$$

From Eqs. (5.16) and (5.17) it follows that

$$\text{Prob} \{a' < \xi < b'\} = \int_{a'}^{b'} P(x) dx, \quad (5.18)$$

and Eq. (5.18) just means that the random value ξ which is the root of Eq. (5.13), has the probability distribution $P(x)$.

Let us consider several specific examples of obtaining of RNs with a given $P(x)$.

(1) The random number ξ is uniformly distributed within an interval (a, b) , if its distribution function is constant in this interval,

$$P(x) = \frac{1}{b-a}, \quad a < x < b. \quad (5.19)$$

In this case the integral (5.13) is trivial, and it follows that

$$\xi = a + (b-a)\gamma. \quad (5.20)$$

(2) Clearly that the three of random numbers $(\gamma_1, \gamma_2, \gamma_3)$ correspond to the coordinates of points which uniformly fill in the unit cube. On the other hand, if we are in need of random points uniformly filled in the sphere of a radius R , we have to pass to spherical coordinates r, φ and θ (radius, azimuthal and polar angles respectively). Then each random point within the sphere is defined by three of random points (r, φ, θ) , which are solutions of the equations

$$\frac{1}{\frac{4}{3}R^3} \int_0^r dr r^2 = \gamma_1, \quad (5.21)$$

$$\frac{1}{2\pi} \int_0^\varphi d\varphi = \gamma_2, \quad (5.22)$$

$$\frac{1}{2} \int_0^\theta d\theta \sin \theta = 1 - \gamma_3. \quad (5.23)$$

From Eqs. (5.21) to (5.23) it follows that

$$r = R\gamma_1^{1/3}, \quad (5.24)$$

$$\varphi = 2\pi\gamma_2, \quad (5.25)$$

$$\theta = \arccos(2\gamma_3 - 1). \quad (5.26)$$

(3) One of important problems where the MC method is used, is the modeling of a random free path for a particle moving in a scattered or dissipating medium. In this case we are in need of random numbers ξ defined in the infinite interval $(0, \infty)$ with the probability distribution

$$P(x) = \frac{1}{\lambda} \exp\left(-\frac{x}{\lambda}\right), \quad (5.27)$$

where $\lambda = M\{\xi\}$ is the mean free path. The solution of Eq. (5.13) in this case yields the relationship $\xi = -\lambda \ln(1 - \gamma)$, or, that is the same,

$$\xi = -\lambda \ln \gamma, \quad (5.28)$$

because γ and $(1 - \gamma)$ are distributed in the interval $(0,1)$ in the same way. The generalization of the distribution (5.27) is the so called *gamma-distribution*, which is characterized by the distribution function

$$P(x) = \frac{1}{(n-1)!} x^{n-1} e^{-x}, \quad (5.29)$$

where $n \geq 1$ is an integer. In this case the random numbers can be obtained with the formula [122]

$$\xi = -\ln(\gamma_1 \gamma_2 \dots \gamma_n). \quad (5.30)$$

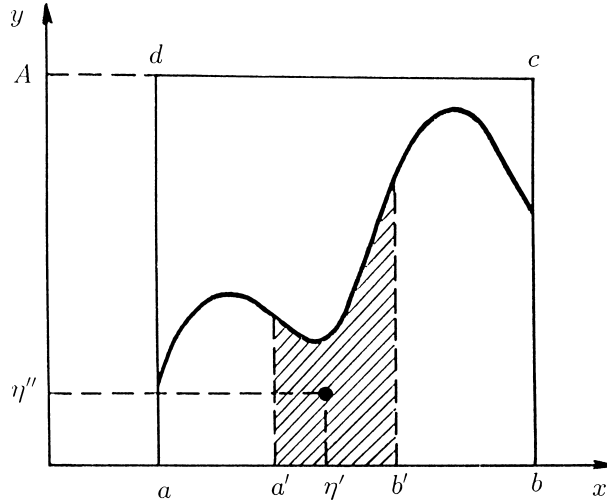


Figure 5.3: Illustration of Neumann's method.

(4) Now let us consider a situation when the integral of $P(x)$ cannot be expressed through elementary functions, or when the function $P(x)$ was determined in an experiment, so that it exists in a table form only. If the random value ξ is defined in a finite interval (a, b) and its distribution function is bounded, $P(x) \leq A$, the raffling of ξ may be done by Neumann's method in the following way:

- (i) choose two sequent values γ' and γ'' from the standard random number sequence $\{\gamma_n\}$, and construct the random point $\Gamma(\eta', \eta'')$ with the coordinates $\eta' = a + \gamma'(b - a)$ and $\eta'' = \gamma'' A$;
- (ii) if the point Γ occurs to be below the curve $y = P(x)$, we put $\xi = \eta'$; otherwise, when the point Γ lies above the curve $y = P(x)$, the pair (η', η'') is to be thrown out, and we are taking the next pair (γ', γ'') .

Indeed, the random points Γ are uniformly distributed within the rectangular “ $abcd$ ” which has the area $A(b - a)$ (see Fig. 5.3). The probability that the point Γ will occur to be below the curve $y = P(x)$ and will not be thrown out, is equal to the ratio of areas

$$\frac{\int_a^b P(x) dx}{A(b - a)} = \frac{1}{A(b - a)}. \quad (5.31)$$

But the probability that the point Γ will occur to be below the curve $y = P(x)$ in the interval $a' < x < b'$, is equal to the following ratio of areas:

$$\frac{\int_{a'}^{b'} P(x) dx}{A(b - a)}. \quad (5.32)$$

Consequently, among all values of ξ which have been picked out, a part of those getting into the interval (a', b') , is equal to the ratio

$$\frac{\int_{a'}^{b'} P(x) dx}{A(b - a)} \bigg/ \frac{1}{A(b - a)} = \int_{a'}^{b'} P(x) dx, \quad (5.33)$$

q.e.d. Note that the sequence $\{\xi_n\}$ obtained in this way, is approximately four times shorter than the original sequence $\{\gamma_n\}$.

(5) Very often we need the so called *normal random numbers* ξ which are defined in the interval $(-\infty, +\infty)$ with the probability distribution

$$P(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad (5.34)$$

so that $M\{\xi\} = 0$ and $D\{\xi\} = 1$. To produce them, we may use the central limit theorem of the probability theory which states that if $\{\xi_n^{(1)}\}$, $\{\xi_n^{(2)}\}$, \dots are independent sequences of random numbers with the same

distribution function, then the sum of N such random numbers, $\xi_n = \sum_{i=1}^N \xi_n^{(i)}$, is to produce at $N \rightarrow \infty$ the sequence $\{\xi_n\}$ of normal random numbers, and

$$M\{\xi\} = NM\{\xi^{(i)}\} \text{ and } D\{\xi\} = ND\{\xi^{(i)}\}. \quad (5.35)$$

Therefore, the normal random numbers can be obtained from the standard random numbers with the relationship

$$\xi \approx \sqrt{\frac{3}{N}} \sum_{n=1}^N (2\gamma_n - 1), \quad (5.36)$$

where $N \gg 1$. For practical use one usually put $N = 12$, so that Eq. (5.36) takes the form

$$\xi \approx \left(\sum_{n=1}^{12} \gamma_n \right) - 6. \quad (5.37)$$

In this case the obtained sequence $\{\xi_n\}$ is in twelve times shorter than the original sequence $\{\gamma_n\}$. Recall that namely such random numbers have been used in solution of the Langevin equation.

From the normal random numbers ξ with the help of a simple renormalization

$$\tilde{\xi} = a + \sigma\xi \quad (5.38)$$

we can get the *Gauss random numbers* $\tilde{\xi}$ distributed with the probability function

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right), \quad (5.39)$$

for which $M\{\tilde{\xi}\} = a$ and $D\{\tilde{\xi}\} = \sigma^2$.

(6) However, the described above method, Eq. (5.37), is not the best one — to get one normal random number, we have to use twelve standard random numbers. Instead [194], one may use the Box-Muller transformation [195], which allows us to transform uniformly distributed random variables to a new set of random variables with a Gaussian distribution. Starting with two independent standard random numbers $\gamma_{1,2}$ and making the transformation $\xi_1 = \sqrt{-2\ln\gamma_1} \cos(2\pi\gamma_2)$, $\xi_2 = \sqrt{-2\ln\gamma_1} \sin(2\pi\gamma_2)$, we obtain two new independent random numbers $\xi_{1,2}$ which have a Gaussian distribution with zero mean and a standard deviation of one. For practical use, the polar form of the Box-Muller transformation is both faster and more robust numerically [196]. The algorithm of it is the following [194]:

```
float x1, x2, w, y1, y2;
do {
    x1 = 2.0 * ranf() - 1.0;
    x2 = 2.0 * ranf() - 1.0;
    w = x1 * x1 + x2 * x2;
} while ( w >= 1.0 );
w = sqrt( ( -2.0 * ln( w ) ) / w );
y1 = x1 * w;
y2 = x2 * w;
```

where `ranf()` is the routine to obtain a standard random number.

(7) Finally, sometimes we need discrete random numbers. For example, such a situation emerges in step-by-step simulation of some process, when at a given step with a probability p (at $\xi = 0$) the process will go in one way, and with a probability $1-p$ (at $\xi = 1$), in the another way. The discrete random numbers ξ_n can be obtained from the SRNs γ_n by the rule

$$\xi_n = \begin{cases} 0 & \text{if } \gamma_n \leq p, \\ 1 & \text{if } \gamma_n > p. \end{cases} \quad (5.40)$$

A more detailed discussion of the questions mentioned in this section may be found, e.g., in Ref. [122].

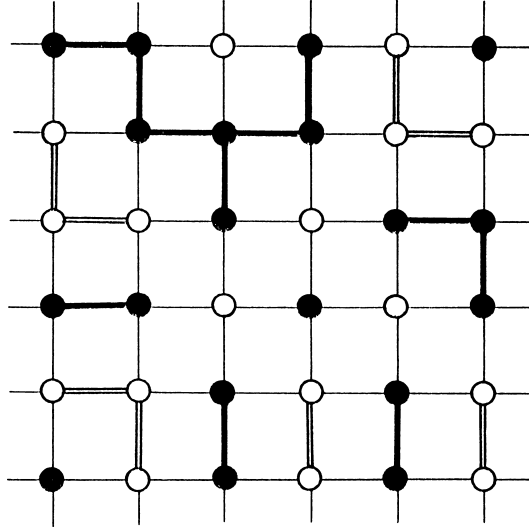


Figure 5.4: Square lattice for the nodal problem of the percolation theory. The sites united into “black” clusters are connected by thick lines, and those united into “white” clusters, by double lines.

5.2 Some applications of the MC method

5.2.1 Percolation theory

Let us consider a square lattice and suppose that each its site is either occupied by an atom with a probability p , or it is free with the probability $(1 - p)$ as is shown in Fig. 5.4. Let us assume that two occupied sites are *coupled* to one another, if they are either nearest neighbors (NN), or if they can be connected by a chain of the NN occupied sites. Let us call by the (“black”) *cluster* the set of occupied sites in which any two sites are coupled (similarly one can define the “white” cluster consisting of free sites, see Fig. 5.4). Clearly that at $p \rightarrow 0$ the average cluster size tends to zero, while in the limit $p \rightarrow 1$ the cluster will occupy the whole lattice. Naturally it arises the question, at what value $p = p_c^s$ does the infinite cluster emerge in the first time? If we suppose that the “bonds” are conducting, and we solder electrodes to the opposite lattice sides, then a current between the electrodes will pass when $p > p_c^s$ only. The value p_c has thus been called by the *percolation threshold*.

The problem of the same class emerges when we are considering a metallic mesh where the each section (i.e., a piece connecting the nearest neighboring sites) has a resistance r . Let us again solder the electrodes to the opposite mesh sides, and measure the system conductivity σ . If now we begin to break the sections with the probability $1 - p$, than the conductivity σ will decrease, and at $p \leq p_c^b$ the conductivity will fall down to zero.

The former among the described above problems is known as the *node problem*, and the latter, as the *bond problem*. According to Hammereli’s theorem, for the same lattice the inequality $p_c^b \leq p_c^s$ is fulfilled. Intuitively it is clear, because in order to blockade one site, we have to break all the bonds which come together to this site.

Note that the bond problem can be transformed into the node problem but for an another lattice called the *covered lattice*. The covered lattice is constructed in the following way: put the site of the covered lattice into the middle point of the each bond of the original lattice, and then connect two sites of the covered lattice to each other, if and only if the corresponding bonds of the original lattice come to the same site. The resulting covered lattice may occur to be more complicated than the original one. For example, the covered lattice for the square lattice cannot be placed on a plane.

The described problems are called the *percolation problems*. Analytical results in the percolation theory are not numerous, and the most natural way of their solution is the MC method. For this, we have to

Table 5.1: Percolation thresholds for lattices of different kinds (analytical results are denoted by asterisk; $p_0 = 2 \sin(\pi/18) \approx 0.347296$) (from [124] and [188])

spatial dimensionality	lattice	p_c^b (bond problem)	p_c^s (node problem)
2	triangular	p_0^*	0.5*
	square	0.5	0.592746
	honeycomb	$(1 - p_0)^*$	0.70
3	simple cubic	0.24881	0.3117 ± 0.0003
	b.c.c.	0.18	0.25
	f.c.c.	0.12	0.20
	diamond	0.39	0.43

construct an array of the necessary dimensionality, and to numerate the array elements analogously to the sites of the lattice. To each array element, the value 0 or 1 (e.g., the free or occupied site respectively) is given with the help of the generator of discrete random numbers. The obtained array is then analyzed to find the clusters (the corresponding algorithms may be found, e.g., in Appendix to the Müller-Krumbhaar paper [123]).

Apart from the calculation of the percolation threshold p_c , it is interesting to study the average cluster size as a function of $p_c - p$ at $p < p_c$, $\langle N \rangle = f_N(p_c - p)$, as well as to investigate the dependence of the system conductivity on $p - p_c$ at $p > p_c$, $\sigma = f_\sigma(p - p_c)$. Numerical results show that $f_N(x)$ and $f_\sigma(x)$ are power functions, so that the percolation processes belong to the class of critical phenomena [187].

In addition to the square lattice, one may study the triangular or hexagonal lattices as well as three-dimensional lattices of different symmetry. Some results for the percolation thresholds obtained by the MC method, are summarized in Table 5.2.1. Besides, the problem may be generalized by accounting of more removed bonds, or by filling the lattice by atoms of different kinds, *etc.*

Percolation models have found a wide application in description of different physical phenomena, such as **(1)** jump conductivity of an amorphous semiconductor, **(2)** destroying of the spontaneous magnetization of a ferromagnetic, when it is doped by nonmagnetic impurities, **(3)** formation of a polymer gel, **(4)** penetration of gas molecules through a pore filter, *etc.* A more detailed discussion of percolation problems may be found, e.g., in Refs. [124, 125, 188].

5.2.2 Penetration of neutrons through an absorbing plate

Now let us consider a standard example from the neutron physics, namely the problem of shielding of radiation. For the sake of concreteness let take the simplest variant of the problem, the penetration of neutrons through a plane.

Let the neutron beam is forwarded normal to a uniform infinite plane of a width h . When a neutron collides with an atom in the plane, it may either be absorbed with a probability p_c , or may be scattered with the probability $p_s = 1 - p_c$. For the sake of simplicity we assume that the neutron energy remains unchanged at the scattering, and that any direction of the neutron after the scattering is equiprobable (the latter is sometimes true for substances consisting of heavy atoms). Figure 5.5 demonstrates possible scenarios of the neutron motion: (a) the neutron passes through the plane, (b) the neutron is absorbed in the plane, and (c) the neutron is reflected from the plane. Let suppose also that between the scattering events the neutron moves freely, and let its motion is characterized by a mean path length λ_0 , so that the random path lengths λ are distributed with the probability density $P(\lambda) = \lambda_0^{-1} \exp(-\lambda/\lambda_0)$.

Let we have to calculate the probability P^+ that the neutron passes through the plane, the probability P^- of the neutron reflection, and the probability P^c of the neutron absorption in the plane. To describe the corresponding computation algorithm, first we have to introduce the following variables. Let N denote the number of incoming neutrons, N^- denote the number of neutrons reflected from the plane, N^+ denote the number of neutrons passed through the plane, N^c denote the number of neutrons absorbed in the plane, and x

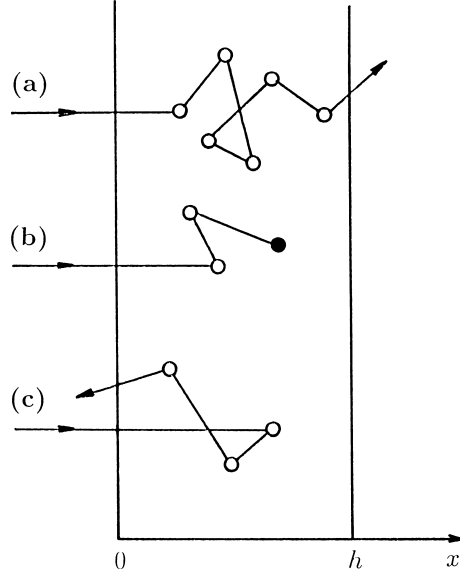


Figure 5.5: Penetration of neutrons through a plane: (a) passing through, (b) absorption in, and (c) reflection from the plane.

denote the horizontal coordinate of the neutron. At the beginning we have to put $N = N^+ = N^- = N^c = 0$. The further step-by-step operation of the program is the following.

Step 1. Suppose that the neutron comes to the plane from the left-hand side (see Fig. 5.5) so that we put $x := 0$ and

$$N := N + 1 \quad (5.41)$$

(here the symbol “:=” designates the assign operator. Recall that in the standard FORTRAN the symbol “=” is used as the assign operator). Put also $\mu := 1$ for the variable μ which will be needed at next steps.

Step 2. With the help of the standard random number generator let us “raffle” the free path length λ :

$$\lambda = -\lambda_0 \ln \gamma. \quad (5.42)$$

Step 3. Calculate the new neutron coordinate after its free run for the distance λ :

$$x := x + \lambda\mu. \quad (5.43)$$

Step 4. Test the condition that the neutron has passed through the plane,

$$x \geq h. \quad (5.44)$$

If this condition is fulfilled, the calculation of the trajectory of the given neutron is finished, and the counter of the passed neutrons is increased by one,

$$N^+ := N^+ + 1. \quad (5.45)$$

In the opposite case we have to test the reflection condition

$$x \leq 0. \quad (5.46)$$

If this condition is fulfilled, the calculation of the trajectory is finished too, and the counter of the reflected neutrons is increased by one,

$$N^- := N^- + 1. \quad (5.47)$$

Last, if both conditions (5.44) and (5.46) are not fulfilled, i.e. if $0 < x < h$, then, consequently, the neutron must undergo a scattering in the plane, and we have to raffle its “fate” at the scattering. For this, let us take the next standard random number γ and test the absorption condition

$$\gamma < p_c. \quad (5.48)$$

If the latter inequality is fulfilled, the calculation of the trajectory is finished again, and the counter of the absorbed neutrons is increased by one,

$$N^c := N^c + 1. \quad (5.49)$$

For the all three cases when the calculation of the given trajectory has been finished, we have to check whether the number of incoming neutrons N reaches the preliminary designed number N_{\max} , and either to stop the computation, or to return to the step 1.

Step 5. If the neutron was not gone out from the plane as well as it was not absorbed, we assume that the neutron undergoes the scattering at the point with the coordinate x . Then we have to raffle a new direction of its velocity. Because we have assumed in our model that any direction of the scattering is equiprobable, the cosine of the scattering angle, μ , must be uniformly distributed in the interval $(-1, 1)$. From this it follows that the value of μ should be raffled by the formula

$$\mu = 2\gamma - 1. \quad (5.50)$$

Then we have to return to the step 2 (it must be noted that each value γ must be used not more than one times, so that for the calculation of one section of the trajectory we are in need of three different SRNs).

When all N_{\max} trajectories have been calculated, the necessary probabilities are defined approximately as

$$P^+ \approx N^+/N, \quad P^- \approx N^-/N, \quad \text{and} \quad P^c \approx N^c/N. \quad (5.51)$$

The computation method described above is quite natural but, unfortunately, it is imperfect. In particular, for a “thick” plane it is too difficult to obtain a reliable value for P^+ , because a majority of neutrons is absorbed in the plane, and the resulting value of N^+ occurs to be too small. In this case we may use other variants of the MC method, for example, the so-called *balance method*. Let us suppose that at the initial time moment a “pack” consisting of a large number w of neutrons, penetrates into the plane, so that the operator (5.41) is to be replaced by $N := N + w$, and then the “pack” moves as a whole along the same trajectory. Consequently, when the “pack” leaves the plane, we have to use the operators $N^\pm := N^\pm + w$ instead the operators (5.45) or (5.47) respectively. When the “pack” collides with a plane’s atom, the number of neutrons absorbed from the “pack” is equal wp_c , while the number of scattered neutrons is equal wp_s in average. Therefore, instead of raffling the absorption process we may simply add the value wp_c to the counter of absorbing neutrons, $N^c := N^c + wp_c$, and then we may continue to follow for the motion of the scattered “pack” which is assumed to be scattered as a whole into the same direction becoming “thinner” in the p_s^{-1} times, i.e. $w := wp_s$. The rest of the algorithm described above remains the same.

The gain in accuracy in the balance method is achieved owing to the fact that now all the trajectories are going out from the plane, no trajectories stop in the plane (a rigorous prove may be found in the Sobol’ book [122]). The value w is called usually as the neutron “weight”, and instead of the “pack” consisting of w neutrons, one says about a single neutron with the weight w . The initial weight w_0 at the step 1 is usually taken as one. This is not in a contradiction with the reasons of a “large pack”, because it can be seen that all w values obtained at the calculation of a given trajectory, contain the initial value w_0 as the common multiplier.

Note that for walls of a complicated shape the MC method is in fact the unique method of calculating the probabilities P^\pm .

5.2.3 Calculation of definite integrals

Let us now consider the problem of approximate calculation of definite integrals. Let we have a function $y = g(x)$ defined in the interval (a, b) , and let this function is bounded so that we can choose the number A such that

$$A > y_{max} \equiv \max g(x). \quad (5.52)$$

Let we have to calculate the integral

$$I = \int_a^b dx g(x). \quad (5.53)$$

Because the computation of definite integrals is equivalent to the calculation of areas, we may proceed in the following way. Consider the rectangular $abcd$ (Fig. 5.3) and fill it uniformly by points with random coordinates. Let the total number of points is M , and the number of points under the curve $y = g(x)$, is M_I . Then the area under the curve may be approximately calculated as

$$I \approx (M_I/M)(b-a)A. \quad (5.54)$$

However, the described computation method is too primitive and leads to large errors, especially in the cases when the function $g(x)$ has, for example, a sharp maximum within the interval (a, b) . For a more accurate solution of the problem under investigation the so-called *preferable choice method* is used usually.

Let us take a random variable ξ defined in the interval (a, b) with a probability distribution $P(\xi)$, and with the help of this variable let us calculate a new random variable η according to the formula

$$\eta = g(\xi)/P(\xi). \quad (5.55)$$

Then one can see that

$$M\{\eta\} = \int_a^b dx P(x) [g(x)/P(x)] = I, \quad (5.56)$$

and this gives the method of computation of I :

$$I \approx \frac{1}{N} \sum_{n=1}^N \eta_n = \frac{1}{N} \sum_{n=1}^N \frac{g(\xi_n)}{P(\xi_n)}. \quad (5.57)$$

Intuitively it is clear that the lower is the dispersion of the random variable η , the higher will be the accuracy of the described method. Indeed, with the help of the central limit theorem of the probability theory it can be proved (e.g., see [122]) that with a very high probability (equal 0.997) the error of the approximation (5.57) does not exceed the value $3\sqrt{D\{\eta\}/N}$. On the other hand, we can lower $D\{\eta\}$ with the help of a suitable choice of the primary random variable ξ . For this, let us calculate the dispersion of the random variable η :

$$D\{\eta\} = M\{\eta^2\} - I^2 = \int_a^b dx [g^2(x)/P(x)] - I^2. \quad (5.58)$$

Let prove that this expression achieves its minimum when $P(x)$ is proportional to $|g(x)|$. To find the minimum of the dispersion $D\{\eta\}$ given by Eq. (5.58) for any possible choice of $P(x)$, let us use the well known inequality

$$\left[\int_a^b dx |u(x)v(x)| \right]^2 \leq \left[\int_a^b dx u^2(x) \right] \left[\int_a^b dx v^2(x) \right]. \quad (5.59)$$

Putting $u(x) = g(x)/\sqrt{P(x)}$ and $v(x) = \sqrt{P(x)}$, from the inequality (5.59) we get

$$\left[\int_a^b dx |g(x)| \right]^2 \leq \left[\int_a^b dx \frac{g^2(x)}{P(x)} \right] \left[\int_a^b dx P(x) \right] = \int_a^b dx \frac{g^2(x)}{P(x)}. \quad (5.60)$$

Equations (5.58) and (5.60) mean that

$$D\{\eta\} \geq \left[\int_a^b dx |g(x)| \right]^2 - I^2. \quad (5.61)$$

It leaves to prove that the lowest boundary is achieved when $P(x)$ is proportional to $|g(x)|$. Let

$$P_g(x) = \frac{|g(x)|}{\int_a^b dx |g(x)|}. \quad (5.62)$$

It is not difficult to calculate that for the probability density (5.62) we have

$$\int_a^b dx \left[g^2(x)/P_g(x) \right] = \left[\int_a^b dx |g(x)| \right]^2, \quad (5.63)$$

and the dispersion $D\{\eta\}$ is indeed equal to the right-hand side of Eq. (5.61).

Note that it is practically impossible to find the “best” density (5.62), because for this we have to calculate the integral $\int_a^b |g(x)| dx$. The calculation of the latter integral is the problem equivalent in its complexity to the original one. Besides, we should not choose too complicated functions $P(x)$, because the raffling of the ξ values becomes a very complicated procedure. But the reasons, described above, can and must be used in choosing of $P(x)$.

For a practical sake, one-dimensional integrals (5.53) are not been calculated by the MC method, there exist more accurate quadrature formulas for these purposes. But when we are meeting with multi-dimensional integrals, the situation changes: quadrature formulas become extremely complicated, while the MC method may be applied almost without changes. Already for integrals over the three-dimensional space, an accuracy of the MC method becomes comparable with the accuracy of quadrature formulas, and in cases of integrals over a space of dimensionality more than three, the MC method becomes the unique numerical method of computation of definite integrals. In particular, the MC method may be used in calculation of path integrals which are widely used in statistical physics [126], quantum physics [127] and quantum field theory [128].

5.3 The MC method in statistical physics

5.3.1 Idea of the method

When the system is in the thermodynamically equilibrium state, then, according to Boltzmann’s hypothesis, its dynamics is stochastic, i.e. it undergoes to statistical laws of the probability theory. Consequently, it is natural to expect that the MC method should be the most suitable one in order to study the equilibrium state. Indeed, in the previous Chapter 4 we have used already the MC method in numerical solution of the Langevin equation. Using the stationary solution of this equation for calculation of time averages, we can study thermodynamical characteristics of the canonical ensemble. Note that a value of the friction coefficient η in this case may be chosen arbitrary (usually one takes $\eta \sim 1$), because the equilibrium characteristics do not depend on η .

On the other hand, instead of the averages over time, we may calculate the averages over the phase space,

$$\langle A \rangle = \int dq A(q) P_{\text{eq}}(q). \quad (5.64)$$

Here $A(q)$ is a measurable physical quantity, q is the coordinate in the $2\nu N_p$ -dimensional phase space of the system under investigation ($\nu = 1, 2$ or 3 is the spacial dimensionality and N_p is the number of particles), and $P_{\text{eq}}(q)$ is the equilibrium Maxwell-Boltzmann distribution,

$$P_{\text{eq}}(q) = Z^{-1} \exp[-\beta E(q)], \quad (5.65)$$

where $\beta = (k_B T)^{-1}$, $E(q)$ is the total system energy, and Z is the partition function,

$$Z = \int dq \exp[-\beta E(q)], \quad (5.66)$$

which ensures the normalization of the probability distribution (5.65).

Usually the measurable $A(q)$ depends only on the coordinates of particles, and the total energy $E(q)$ may be presented as a sum of kinetic and potential energies. In this case the integrals over momenta in Eqs. (5.64) and (5.66) can be taken analytically and, moreover, they are mutually cancelled in the expression for $\langle A \rangle$. Therefore in what follows we will let $E(q)$ in Eqs. (5.65) and (5.66) denote the potential energy $V(q)$ only, and the integration in Eqs. (5.64) and (5.66) is to be carried out only over the spacial coordinates of particles.

The integrals (5.64) and (5.66) are the νN_p -dimensional integrals over the configurational space of the system under investigation. They can be found numerically by the unique method, namely by the MC method. As has been described in the previous Sec. 5.2.3, for this we have to construct a random sequence $\{\xi_n\}$ of points in the νN_p -dimensional configurational space such that the probability distribution density of this sequence, $P(q)$, is to be close as possible to the under-integral functions in Eqs. (5.64) and (5.66). Then we can calculate $\langle A \rangle$ with the formula

$$\langle A \rangle = \lim_{N \rightarrow \infty} \frac{N^{-1} \sum_{n=1}^N P^{-1}(\xi_n) A(\xi_n) \exp[-\beta V(\xi_n)]}{N^{-1} \sum_{n=1}^N P^{-1}(\xi_n) \exp[-\beta V(\xi_n)]}. \quad (5.67)$$

Taking into account that the functions $A(q)$ are smooth usually, the choice $P(q) = P_{\text{eq}}(q)$ has to be the optimal one. In this case Eq. (5.67) reduces to the arithmetic average,

$$\langle A \rangle = \frac{1}{N} \sum_{n=1}^N A(\xi_n), \quad N \rightarrow \infty. \quad (5.68)$$

Thus, the problem under investigation reduces to a search of the sequence of random points in the phase space, ξ_1, ξ_2, \dots , distributed with the probability density $P_{\text{eq}}(q)$ which has to coincide with the Boltzmann distribution. Because each point in the phase space corresponds to a given state of the system, the random points ξ_n predominantly fill those part of the phase space, where the probability to find the system is high. That is to say that the microstates ξ_n are to be sampled according to their importance. The method of construction of the random sequence $\{\xi_n\}$ with the distribution density $P_{\text{eq}}(q)$, based on the procedure of random walk of a point in the configurational space of the system, was firstly proposed by Metropolis *et al.* [129]. It consists in the following.

Let us turn to a discrete time $t_n = n \Delta t$ (further we may put $\Delta t \rightarrow 0$) and introduce some distribution density $P(q, t_n)$ which corresponds to the probability that the system is in the point q of the configuration space at the time t_n . Introduce also the value $W(q \rightarrow q')$ which describes the conditional probability of the transition of the system from the point q to the point q' per unit of time. Then, we can write the following evolution equation for the function $P(q, t_n)$:

$$\begin{aligned} \Delta P(q, t_n) &\equiv P(q, t_n) - P(q, t_{n-1}) \\ &= \sum_{q'} W(q' \rightarrow q) P(q', t_{n-1}) \Delta t - \left(\sum_{q'} W(q \rightarrow q') \Delta t \right) P(q, t_{n-1}), \end{aligned} \quad (5.69)$$

which is known as the *master equation*. The first term in the right-hand side of Eq. (5.69) rises $P(q, t_n)$ owing to transitions of the system to the state q from other points q' of the phase space. The rate of the increasing is equal to the probability of the transition for time Δt , $W(q' \rightarrow q) \Delta t$, multiplied by the probability that the system was in the state q' , $P(q', t_{n-1})$. Analogously, the second term of Eq. (5.69) describes the transitions from the state q to other points q' of the phase space. In the limit $\Delta t \rightarrow 0$ Eq. (5.69) takes the form

$$\frac{\partial P(q, t)}{\partial t} = \sum_{q'} [W(q' \rightarrow q) P(q', t) - W(q \rightarrow q') P(q, t)]. \quad (5.70)$$

It can be proved that this equation has a stationary solution $P_0(q)$, and that any initial probability distribution $P(q, t = 0)$ approaches to the $P_0(q)$ as $t \rightarrow \infty$, provided the following two conditions are satisfied (the proof may be found, e.g., in [130]): **(1)** $W \geq 0$, and **(2)** the *ergodicity condition*, or the *connectivity condition*, according to which it must exist a possibility of the transition from any one point of the phase space to any other point for a finite number of steps.

Then, for the stationary distribution, Eq. (5.70) takes the form

$$\sum_{q'} [W(q' \rightarrow q) P_0(q') - W(q \rightarrow q') P_0(q)] = 0. \quad (5.71)$$

The simplest way to satisfy this equation is to put zero each term of the sum (5.71), i.e.

$$W(q' \rightarrow q) P_0(q') - W(q \rightarrow q') P_0(q) = 0. \quad (5.72)$$

Equation (5.72) is called the *detailed balance condition*.

In order that $P_0(q)$ is to coincide with the necessary function $P_{\text{eq}}(q)$ (i.e., with Boltzmann's distribution (5.65)), we have to choose the transition probabilities $W(q \rightarrow q')$ which satisfy the condition

$$\frac{W(q \rightarrow q')}{W(q' \rightarrow q)} = \frac{P_{\text{eq}}(q')}{P_{\text{eq}}(q)} = \exp\left(-\frac{\Delta V}{k_B T}\right), \quad (5.73)$$

where $\Delta V = V(q') - V(q)$. Then the necessary random sequence of states of the system, ξ_1, \dots, ξ_N , distributed with the probability density $P_{\text{eq}}(q)$, may be obtained as the stationary solution of the master equation (5.69), i.e. as the sequence of states obtained with the help of this equation when the system has achieved the equilibrium state.

The rates $W(\dots)$ in Eq. (5.73) may be chosen in different ways. The simplest and the most convenient is the choice

$$W = \begin{cases} \exp(-\Delta V/k_B T) & \text{if } \Delta V > 0, \\ 1 & \text{if } \Delta V < 0. \end{cases} \quad (5.74)$$

Thus, the transition to a state with a lower potential energy occurs with the probability one, while the transition to a state with a higher potential energy, with the probability $\exp(-\Delta V/k_B T)$ which is always lower than one.

5.3.2 Calculation scheme

The Metropolis procedure may be realized with the help of the following algorithm.

First let us consider the canonical ensemble consisting of N_p particles imbedded into some fixed volume at a constant temperature T . Let choose an initial configuration ξ_0 , and initiate the counter of states n , $n := 0$.

Step A1. Put $n := n + 1$. If $n > N$, where N is the preliminary stated length of the sequence ξ_n , the calculation is finished.

Step A2*. With the help of the standard random number generator, select the j -th particle by the formula $j := \text{int}(\gamma N_p)$.

Step A3*. In a random way, choose the vector $\Delta \vec{r}$ which gives the direction and the length of the displacement, and displace the j -th particle on $\Delta \vec{r}$ (the direction may be raffled by Eqs. (5.25) and (5.26), and the length, with the help of Gauss random numbers). In the result we get a trial configuration ξ^* .

Step A4. Calculate the change of the potential energy, ΔV , caused by the displacement of the randomly selected particle j on the random vector $\Delta \vec{r}$, $\Delta V = V(\vec{r}_j + \Delta \vec{r}) - V(\vec{r}_j)$.

Step A5. If in the result of this displacement the potential energy decreases, $\Delta V < 0$, the trial configuration is accepted as the next element of the sequence, $\xi_n := \xi^*$.

Step A6. In the opposite case, when the potential energy increases, $\Delta V > 0$, take a next standard random number γ , which has to be compared with $\exp(-\Delta V/k_B T)$. In this case it may be two variants: **(1)** if $\gamma \leq \exp(-\Delta V/k_B T)$, we put $\xi_n := \xi^*$ again; **(2)** but if $\gamma > \exp(-\Delta V/k_B T)$, then the trial configuration ξ^* should be throw out as an “unsuccessful” one, and the original configuration is duplicated, $\xi_n := \xi_{n-1}$.

Step A7. Return to the step **A1**.

The Metropolis procedure can be generalized for the case of grand canonical ensemble, where the chemical potential μ is fixed, while the number of particles may vary. Recall that in this case the averages are determined by the expression

$$\langle A \rangle = \sum_{N_p=0}^{\infty} \int dp dq A(q, N_p) P_{\text{eq}}(p, q, N_p), \quad (5.75)$$

where $P_{\text{eq}}(p, q, N_p)$ is defined by the formula

$$P_{\text{eq}}(p, q, N_p) = \frac{1}{\Xi N_p!} \exp \{ -\beta [E(p, q, N_p) - \mu N_p] \}, \quad (5.76)$$

and the grand partition function is equal to

$$\Xi = \sum_{N_p=0}^{\infty} \frac{1}{N_p!} \int dp dq \exp \{ -\beta [E(p, q, N_p) - \mu N_p] \}. \quad (5.77)$$

Repeating the arguments which have led to the expression (5.73), one can obtain that with accounting of the integration over momenta, the transition probabilities now must satisfy the condition

$$\frac{W(\{q, N_p\} \rightarrow \{q', N'_p\})}{W(\{q', N'_p\} \rightarrow \{q, N_p\})} = \frac{N_p!}{N'_p!} a^{N'_p - N_p} \exp(-\beta \Delta V), \quad (5.78)$$

where

$$\Delta V = V(q', N'_p) - V(q, N_p), \quad (5.79)$$

$$a = (mk_B T / 2\pi \hbar^2)^{\nu/2} \exp(\beta \mu), \quad (5.80)$$

$\nu=1, 2$ or 3 is the system dimensionality, and \hbar is the Planck constant.

The algorithm is modified in the following way.

Step B1. Coincides with the step **A1**.

Step B2. Select in a random way one among the following three operations:

- (a) to leave the number of particles unchanged (if the SRN γ occurs to be within the interval $0 < \gamma < 1/3$),
- (b) to increase N_p by one (if $1/3 \leq \gamma < 2/3$), or
- (c) to decrease N_p by one (if $2/3 \leq \gamma < 1$).

Step B3a. In the case of the variant **B2(a)**, we have to carry out the steps **A2** to **A6**, and then to return to the step **B1** instead of the **A1**.

Step B3b. In the case of the variant **B2(b)**, we have to choose in a random way the coordinates for a new particle, and to add it to the system, $N_p := N_p + 1$, obtaining in such a way a trial configuration ξ^* .

Step B3c. In the case of the variant **B2(c)**, we have to select in a random way the number $j = \text{int}(\gamma N_p)$, and to take out the j -th particle from the system, $N_p := N_p - 1$, thus obtaining the configuration ξ^* .

Step B4. Calculate the change of the potential energy for the transition from the old configuration ξ_{n-1} to the trial configuration ξ^* , $\Delta V = V(\xi^*) - V(\xi_{n-1})$.

Step B5. Calculate the value $b = [a/(N_p + 1)] \exp(-\beta \Delta V)$ in the case of creation of a particle (the variant (b)), or the value $b = (N_p/a) \exp(-\beta \Delta V)$ for the case of the destruction of the particle (the variant (c)). If $b \geq 1$, the trial configuration is accepted, $\xi_n := \xi^*$.

Step B6. But if $b < 1$, we call the next standard random number γ and compare it with b . If $\gamma \leq b$, then we again put $\xi_n := \xi^*$. But if $\gamma > b$, the configuration ξ^* is to be throw out, and we put $\xi_n := \xi_{n-1}$.

Step B7. Return to the step **B1**.

Note that depending on the model under investigation, a concrete realization of the steps **A2*** and **A3*** marked by asterisk, may differ from those proposed above. Note also that in principle with the help of the MC method we may consider the microcanonical ensemble as well, if we introduce a “demon” which causes artificially the random point (the configuration of the system) to walk in the phase space closely to the surface of constant energy. Besides, the MC method may be modified in order to consider the *NPT* ensemble too. The corresponding algorithms are described, e.g., in [131] and in Heerman’s book [1].

With the help of the algorithms described above, the sequence of configurations (or points in the phase space) $\xi_0, \xi_1, \xi_2, \dots, \xi_n, \dots, \xi_N$ is constructed. This sequence creates the Markov chain, which approaches to the stationary distribution $P_{\text{eq}}(q)$ for large n . Because for the calculation of equilibrium characteristics of the system we must use a part of the Markov chain $\{\xi_n\}$, which corresponds to the stationary distribution, an initial part of the sequence should be thrown out. When the sequence $\{\xi_n\}$ is been found, any average can be calculated using Eq. (5.68). A length of the sequence should be such that $N/N_p \sim 10^3 \div 10^6$; for $N \sim 10^4 N_p$ an error in calculation of averages is about 1% [132]. Because the raffling of each configuration is in need of a several standard random numbers, the period of the sequence $\{\gamma_n\}$ produced by the SRNG, must be very large, it has to exceed the value $(10^4 \div 10^7)N_p$. The MC method allows us to simulate systems consisting of an essentially larger number of particles (larger than $500^3 \approx 10^8$) than the MD method.

The Markov chain constructed with the algorithm described above, is not too satisfactory, because the consequent configurations differ one from another by the position of one particle as maximum, while the other particles remain at the same positions. Therefore, a real difference of consequent configurations is not too large, and such Markov chain is not very suitable for calculation of thermal characteristics. In order to improve this deficiency, one may construct a new Markov chain $\{\tilde{\xi}_n\}$, selecting from the original chain $\{\xi_n\}$ the each m -th configuration, i.e.

$$\tilde{\xi}_1 = \xi_{i_0}, \tilde{\xi}_2 = \xi_{i_0+m}, \tilde{\xi}_3 = \xi_{i_0+2m}, \dots, \quad (5.81)$$

where the number i_0 corresponds to the configuration when the chain has achieved the equilibrium state. It is easy to see that if we take $m > N_p$, then the sequent states in the chain $\{\tilde{\xi}_n\}$ with a high probability will differ one from another by displacements of all particles.

The another deficiency of the described above method is connected with the fact that at each step we are trying to shift a single particle only, i.e., a collective motion of particles is completely ignored. In a result, it exists a probability that the system may be locked in a metastable state, and the escape from this state may be possible with a simultaneous jump of few particles (this is especially dangerous at an “unsuccessful” choice of the initial configuration ξ_0). To improve this deficiency, we may complicate the algorithm, for example, to include the possibility of simultaneous displacements of two particles at the steps **A2–A3**. A more powerful method based on Darwing’s laws, will be described below in Sec. 5.6.

Unfortunately, with the help of the Metropolis procedure we can calculate only derivatives of the free energy with respect to thermodynamic variables (for example, the specific heat $c_V = (\partial F / \partial T)_V$, or the magnetization $M = (\partial F / \partial h)_T$, where h is the magnetic field). But this method does not allow to calculate directly the partition function Z and, consequently, the free energy $F = -k_B T \ln Z$ and the entropy $S = (E - F) / k_B T$. To improve the situation, one may try to find the entropy by another methods, for example, using its definition

$$S = k_B \ln \mathcal{N}, \quad (5.82)$$

where \mathcal{N} is the number of ways to place the system in the phase space [133]. The another method is to calculate the value $A(a) = \partial F / \partial a$ for all values of the variable a from 0 to a_0 , and then to find F with the help of numerical integration,

$$F(a_0) = \int_0^{a_0} da A(a) + F(0). \quad (5.83)$$

Here, however, we must know the value $F(0)$.

A more detailed description of applications of the MC method in physics may be found, e.g., in the books edited by Binder [132, 134, 135].

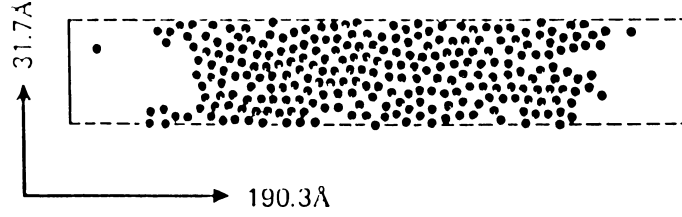


Figure 5.6: Snapshot pictures of the x - y atomic positions of the 256 Lennard-Jones argon atoms for the 10^7 -th configuration in the Monte Carlo simulation for the two-dimensional liquid-vapor interface. The horizontal broken lines of the computation box depict periodic boundary conditions and the vertical lines depict hard walls (after [88]).

5.4 Modeling of phase transitions

The investigation of phase transitions is, perhaps, the main application of the MC method. Because it is described in a number of review articles and monographs (e.g., see books [132, 134, 135, 136]), the further description will be rather brief.

For the MC studies of different states of the system (such as gas, liquid and crystalline states) and phase transitions (PT) between these states, the model is choosing with the same principles as in the MD method described above in Sec. 3.1. Namely, we have to choose the dimensionality of the system, then to take the law of interaction between the particles, and to put the boundary conditions. Because the number of particles in the MC method may essentially exceed the number of particles accepted in the MD method, the MC simulation may be successfully employed for three-dimensional systems as well. Usually in this case the periodic boundary conditions are imposed. Otherwise, if we are using the free or fixed boundary conditions, it emerges the problem how to distinguish the body and surface contributions. For this, we may make several computations with different numbers of particles N_p . Indeed, for the homogeneous (one-phase) state of the system, the dependence of some intensive characteristic A versus N_p may be presented approximately in the form

$$A(N_p) \approx a_v N_p + a_s N_p^{2/3}. \quad (5.84)$$

Then, constructing the graph of the dependence $A(N_p)/N_p$ versus $N_p^{-1/3}$, we can find the body contribution a_v from the point of intersection of the dependence with the ordinate axis, and the surface contribution a_s from the tangent of the inclination angle. Note, however, that the procedure of separation of the contributions is essentially complicated near the point of the second-order phase transition, where extensions of fluctuationally arising domains start to exceed the system size, and the value a_v begins to depend on N_p .

5.4.1 Continuum models

For studying of continuum models (i.e. the models where the particles can move continuously in space), the Metropolis algorithm described in the previous section, is explored usually. The average displacement $\langle \Delta \vec{r} \rangle$ for the shift of a particle per one MC step, is selected during the program training in such a way that to obtain the maximally fast approach to the equilibrium state. In this case it is useful to compare the results of the MC modeling with those of the MD simulation of the same model.

Summary of the results on the MC modeling of homogeneous systems as well as the phase transitions in these systems, may be found in many works, in particular, in the paper of Levesque *et al.* (see Chapter 2 in the book [132]); see also the bibliography in Abraham's survey [88]. Note only that when the interparticle interaction is modeled by the hard-core potential (3.2), the system exhibits only a single first-order phase transition, the gas-liquid PT. But already a crude inclusion of attraction with the help of the step potential (3.3), leads to the existence of both phase transitions, the gas-liquid and liquid-solid PTs.

As an example, let us consider the Abraham results [88], where the system of $N_p = 256$ atoms interacting by the 12–6 Lennard-Jones potential (3.5) were studied. The system was imbedded into the rectangular box (see Fig. 5.6), two boundaries of which (parallel to the OY axis) were the elastically reflected walls, while on the boundaries parallel to the OX axis, the periodic boundary conditions were imposed. Such a geometry has allowed to get a planar (more exactly, a linear for the two-dimensional system) boundary of separation of two phases, the gas and liquid ones (if we impose the same boundary conditions in both directions, the separation boundary should be a circle). It was obtained the sequence of MC configurations of a length $N \sim 10^5 N_p$. One of the configurations is shown in Fig. 5.6. One can see a sharp boundary of separation of two phases. However, the surface of separation of the phases occurs to be “rough” (unlike from a usually smooth separation surface in three-dimensional systems); a width of the transition layer is about $(3 \div 4)r_0$ (r_0 is the parameter of the LJ potential; for **Ar** atoms $r_0 = 3.822 \text{ \AA}$).

5.4.2 Lattice models

Perhaps, the most successful applications of the MC method are connected with lattice models. Recall that these models assume that particles may occupy only sites of some (one-, two-, or three-dimensional) spatial lattice characterized by a given symmetry. At the same time, each particle may carry several (usually discrete) internal degrees of freedom, for example, a spin in the case of magnetics or a kind of the atom in modeling of binary alloys. Consequently, with each site j of the lattice we have to associate one or several variables which take, for example, one among two values, $n_j = 0$ (the free site) or $n_j = 1$ (the site occupied by an atom).

The potential energy in the lattice model consists of one-particle and many-particle terms. The one-particle term describes the particle’s energy in the given lattice site, and it may include a dependence on internal degrees of freedom. If the lattice has a complex elementary cell, different sites in the cell may be characterized by different one-particle energies. When disordered systems are modeled, the one-particle energy may take random values.

The interparticle interaction is described with the help of introduction of the parameters v_{ij} , which are equal the energy of interaction of the particles occupied the sites i and j (this energy may also include a dependence on internal degrees of freedom). Emphasize that if the occupation number n_j is allowed to take only two values 0 and 1, then the lattice model supposes a short-range interparticle repulsion with a potential of the hard-core type from the very beginning. In the simplest case one assumes that $v_{ij} \neq 0$ only for the particles which occupy the nearest neighboring sites. Often, however, we have to take into account the interaction of more removed neighbors, as well as the anisotropy of interaction. Additionally to the pairwise interaction v_{ij} , sometimes it is necessary to include a “trio” interaction v_{ijk} .

The term corresponding to the kinetic energy of the system, is absent in the lattice models. As the result, wave processes (phonons, spin waves, *etc.*) are lost when the lattice models are used.

In time evolution of the system, the particles may “jump” to nearest neighboring or more removed free sites of the lattice. When the system is in contact with some reservoir of particles, the destruction of some particles and the creation of new particles (at the free lattice sites) may take place. Besides, the particles may change the value of their internal degrees of freedom. It is assumed, however, that such acts are carried out “instantly” in time, and that they are characterized by some probabilities (of the jump, destruction, creation, *etc.*) per unit of time.

We have already met with applications of the lattice models in Sec. 5.2.1, where we discussed the percolation problems. In that case it has been assumed $v_{ij} = 0$, and in the result of simulation we have obtained the geometry of connected clusters.

An interesting problem is the investigation of random walks of particles in the lattice. If we “mark” one of the particles, for example, with the help of introduction of some (fictitious) internal degree of freedom, then in this way we can calculate the self-diffusion coefficient D_s .

Two-dimensional (2D) variant of the lattice model is widely used in modeling of adsorption systems.

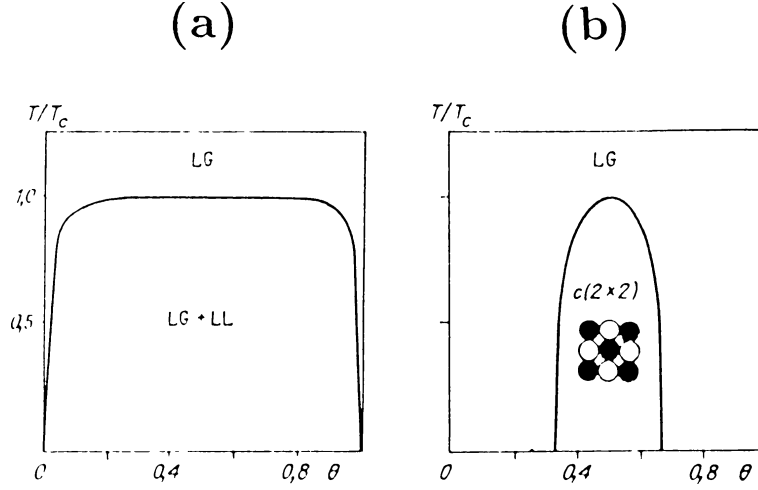


Figure 5.7: Phase diagram for the lattice-gas model with the square lattice and interaction of nearest neighbors only: (a) attraction, and (b) repulsion of atoms. LG is the lattice-gas phase, LL is the lattice-liquid phase, and $c(2 \times 2)$ is the ordered phase. $T_c \approx 0.567|v_{ij}|$ (after [139]).

The system Hamiltonian in this case has a form

$$H = \sum_j (\varepsilon_{(j)} - \mu) n_j + \frac{1}{2} \sum_{i \neq j} v_{ij} n_i n_j + \frac{1}{6} \sum_{i \neq j \neq k} v_{ijk} n_i n_j n_k + \dots \quad (5.85)$$

Here the occupation number n_j takes the values 0 or 1, $\varepsilon_{(j)}$ is the one-particle energy, and μ is the chemical potential of the particle's reservoir, the latter is to be taken into account only in simulation of the adlayer which is in contact with a surrounding gas phase.

When the adlayer is isolated from the gas, the number of particles $N_p = \sum_j \langle n_j \rangle$ is to be fixed, and for the elementary MC step we must take the jump of an atom to the nearest neighboring (or a more removed) free site, $(n_i = 0, n_j = 1) \rightleftharpoons (n_i = 1, n_j = 0)$ (the so-called *Kawasaki dynamics* [137]). On the other hand, when we are studying the adlayer which is in contact with a gas phase, the elementary MC steps must include adsorption and desorption processes, $(n_j = 0) \rightleftharpoons (n_j = 1)$ (the so-called *Glauber dynamics* [138]).

From the sequence of configurations $\{\xi_n\}$ we can calculate the correlation function,

$$G(i - j) = \langle (n_i - \langle n_i \rangle)(n_j - \langle n_j \rangle) \rangle, \quad (5.86)$$

so that we can determine the phase state of the adlayer and construct the phase diagram of the system. Unlike from the solely two-dimensional system (i.e. the film adsorbed on “smooth” substrate without a potential relief), the 2D system subjected to the external periodic substrate potential (recall that the lattice model corresponds to the extreme case of infinitely high amplitude of the substrate relief) may be in a “true” crystalline phase, i.e. in the phase with nondecaying (oscillating) correlation function (5.86) at a temperature $T \neq 0$. For description of ordered phases one introduces the so-called *order parameter* which is a tensor function in a general case (e.g., see [76] or [87]).

As the simplest example, let us consider the square lattice with interaction of nearest neighbors only. If the atoms attract each other, $v_{ij} < 0$, the system exhibits the first-order PT. The phase diagram of the system is shown in Fig. 5.7a, where the full curve shows the dependence of the critical temperature T_{crit} on the “coverage” $\theta = N_p/N_L$ (N_p is the number of adatoms and N_L is the total number of the lattice sites). At $T > T_{\text{crit}}(\theta)$ the adlayer is in the lattice-gas (LG) phase, while at $T < T_{\text{crit}}(\theta)$ there coexist simultaneously two phases, the LG phase and the lattice-liquid (LL) phase. Note that the most direct proof, that a phase transition studied with the help of the computer experiment corresponds to the first-order PT, is, perhaps, the observation of a hysteresis at slow increasing and decreasing of the temperature during the simulation.

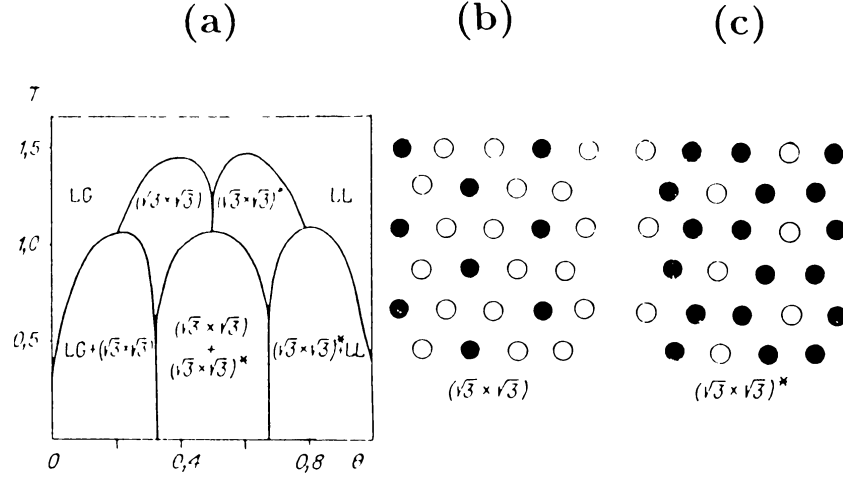


Figure 5.8: (a) Phase diagram for the lattice-gas model on the triangular lattice for repulsive interaction of nearest and next-nearest neighbors, (b) the structure $(\sqrt{3} \times \sqrt{3})$, and (c) the structure $(\sqrt{3} \times \sqrt{3})^*$ (black and white circles correspond to occupied and free lattice sites respectively) (after [143]).

It has to be noted also that the determination of the PT point itself is a rather difficult problem because of finiteness of the system under simulation (e.g., see [136, 140, 141] for more details).

Otherwise, if the atoms repel each other, $v_{ij} > 0$, the system exhibits the second-order PT of the “order–disorder” type. The corresponding phase diagram calculated by Binder and Landau [139], is presented in Fig. 5.7b. In this case at $T < T_{\text{crit}}(\theta)$ the whole adlayer is in the ordered state with the $c(2 \times 2)$ structure (see Fig. 5.7b). Emphasize that the ground state of the $c(2 \times 2)$ structure is doubly degenerated: from one ground-state configuration we may obtain the another (equivalent) ground-state configuration by the shift of all atoms by one lattice constant. At $T = 0$ one of these states is realized in the adlayer, but at $T \neq 0$ the domains corresponding to the another ground state configuration, are created fluctuationally. With T increasing, the number of such domains rises, and at $T = T_{\text{crit}}(\theta)$ the domains are completely mixed so that the ordered phase becomes totally destroyed. Recall that close to the point of the second-order PT, the correlation radius tends to infinity, and the phase transition takes place simultaneously in the whole system. Evidently that owing to the finiteness of the system under simulation, a region of the phase diagram near the PT points, is “smeared”, and this may lead to essential errors in determination of $T_{\text{crit}}(\theta)$.

The phase diagrams described above, correspond to the case of the canonical ensemble, where the number of adatoms N_p is fixed. On the other hand, when the adlayer is in contact with a gas, first we have to find the average number of adatoms, $\langle N_p \rangle$, which corresponds to a given value of the chemical potential of the gas (or to the gas pressure), and then to use the phase diagrams presented above.

Recall that the described model has an analytical solution in a single point only, namely at $\theta = \frac{1}{2}$ (the famous Onsager solution [142] of the two-dimensional Ising model for zero magnetic field).

As a more complicated example, Fig. 5.8 demonstrates the phase diagram calculated by Landau [143], for the triangular lattice with accounting the repulsive interactions of nearest and next-nearest neighbors. An interesting peculiarity of this phase diagram is a degeneration (the mixture of ordered phases) near the $\theta = \frac{1}{2}$ point.

When the model takes into account pairwise interatomic interactions only, the Hamiltonian (5.85) is invariant with respect to the substitution “occupied site” \rightleftharpoons “free site” [so that $\theta \rightleftharpoons (1 - \theta)$], and the phase diagram is symmetric with respect to the $\theta = \frac{1}{2}$ line. However, the accounting of trio interactions ($v_{ijk} \neq 0$) destroys this symmetry.

The comparison of the phase diagram calculated by the MC method with that measured experimentally, helps to extract an important information on the interaction between adatoms. Note that the employing of analytical approaches for the construction of phase diagrams, is very complicated owing to the fact that

fluctuation effects in 2D systems are too strong and, consequently, approximate theories of the mean-field type are inapplicable. Note also that the description of a submonolayer adsorbed film with the help of the lattice model is justified, if the following two conditions are satisfied simultaneously. First, it should be fulfilled the inequality

$$k_B T \ll \varepsilon_a, \quad (5.87)$$

where ε_a is the activation energy for the transition of an adatom to a neighboring lattice site. When the inequality (5.87) is satisfied, the adatoms are situated in the lattice sites (at the minima of the substrate potential relief) for an almost the whole time, and from time to time only they are jumping to another sites. Second, it should be satisfied the condition

$$|v_{ij}| \ll \varepsilon_a, \quad (5.88)$$

because in the opposite case we can not neglect by displacements of adatoms from the relief minima caused by interatomic interactions. These two conditions are satisfied, for example, for gas atoms or molecules (such as **O**, **H**, **CO**, **NO**, *etc.*) chemically adsorbed on a metal substrate.

In investigation of multilayer adsorbed films, as well as in studying of clean crystal surfaces, it is often used an another variant of the lattice model, the so-called *solid on solid* (SOS) model introduced by Kossel. This model assumes that each site of the two-dimensional lattice is associated with the discrete variable h_j ($0 \leq h_j < \infty$ for the simulation of adlayers, and $-\infty < h_j < +\infty$ for the simulation of clean surfaces) which has a meaning of a height of the column of adatoms growing over a given lattice site. The one-particle contribution into the system potential energy is assumed to be proportional to the column height h_j , $\sum_j \varepsilon(h_j) h_j$, and for accounting of interatomic interaction parallel the surface it is assumed that the interaction energy of two nearest neighboring columns of heights h_i and h_j is determined by the difference of their heights. Thus, the Hamiltonian of the SOS model has the form

$$H = \sum_j (\varepsilon(h_j) - \mu) h_j + \frac{1}{2} \phi \sum_{\langle i \neq j \rangle} |h_i - h_j|^m, \quad (5.89)$$

where ϕ is the parameter of interaction of the nearest neighboring columns (in the case of their attraction we have to put $\phi > 0$), and the exponent m is usually taken to be $m = 1$. The sum in the second term of the right-hand side of Eq. (5.89) is carried out over the nearest neighboring sites only. Notice that when the interval of h_j variation is restricted by the values 0 and 1 only, the Hamiltonian (5.89) reduces to the Hamiltonian (5.85) considered above. The special feature of the SOS model is that it forbids the existence of “overhanging” structures, i.e. there must be no bubbles in the adlayer. As the elementary MC acts in the SOS model we have to take the adsorption-desorption acts, $(h_j) \rightleftharpoons (h_j \pm 1)$, as well as the acts of diffusion along the surface, $(h_i, h_j) \rightleftharpoons (h_i - 1, h_j + 1)$.

The SOS model may be used for studying of the roughing of the surface with increasing of temperature. Examples of the MC configurations obtained in this case, are presented in Fig. 5.9 taken from [132]. More often, however, the SOS model is used to study the crystal growth (see Sec. 5.5).

The most known among the lattice models is the Ising model, where all sites are assumed to be occupied by particles having an internal degree of freedom s_j , called usually by spin, and s_j is assumed to take, for example, two values $s_j = \pm 1$ (along or opposite the external magnetic field h). The Hamiltonian of the Ising model has the form

$$H = -h \sum_j s_j + \frac{1}{2} \sum_{i \neq j} J_{ij} s_i s_j, \quad (5.90)$$

where J_{ij} is called the “exchange integral”. For modeling of ferromagnetics we have to put $J_{ij} < 0$, and for the case of antiferromagnetics, $J_{ij} > 0$. Recall that the two-dimensional version of the Ising model for the $h = 0$ case with accounting of interaction of nearest neighbors only, has the analytical solution [142].

It is easy to see that the substitution $s_j = 1 - 2n_j$ transforms the Ising model into the lattice-gas model (5.85). As the elementary step in the Ising model, a spin in a randomly chosen site j is upset usually, $(s_j = +1) \rightleftharpoons (s_j = -1)$ (in the lattice-gas model this corresponds to the case of an adlayer in contact with a gas).

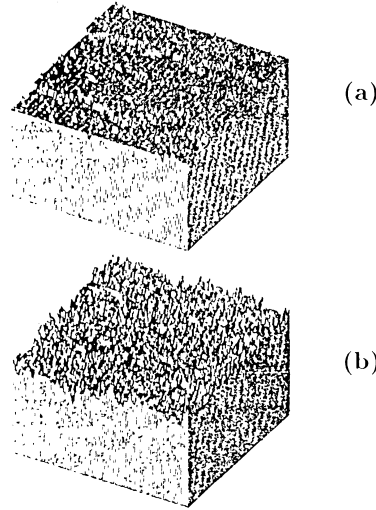


Figure 5.9: Simulation of crystal roughening in the framework of the Kossel (solid-on-solid) model. (a) At low temperature the surface is smooth, and (b) at high temperature the surface is rough (from [132]).

Note one more that in the Ising model at $h = 0$ there exists an effective violation of ergodicity so that at low temperatures (lower than the critical temperature T_{crit}) the ground state of the system is doubly degenerated. For example, in the case of ferromagnetics the total magnetization $M = \sum_j \langle s_j \rangle$ may be either positive or negative. Because of finiteness of the system investigated in computer simulation, there always exists a nonzero probability of the transition from one ground state to the another ground state. The probability of this transition is determined by an activation barrier, the height of which is proportional to $N_L^{(\nu-1)/\nu}$ [144]. Therefore, when we are calculating the total magnetization, we have to accumulate for averaging the absolute values of M .

In conclusion, let us recall also the Geisenberg model, where with each lattice site it is associated a continuous variable, for example, a vector of unit length \vec{s}_j , which may be directed arbitrary. In this case the model Hamiltonian takes the form

$$H = -\vec{h} \sum_j \vec{s}_j + \frac{1}{2} J \sum_{\langle i \neq j \rangle} \vec{s}_i \vec{s}_j, \quad (5.91)$$

where \vec{h} is the vector corresponding to the external magnetic field.

The modeling of magnetic systems with Ising- or Geisenberg-type Hamiltonians is a very wide topic, which lies beyond the scope of the present course. For a more detailed discussion of this problem we refer a reader to special sources (e.g., see Landau's article in the book [139]).

5.5 Relaxation dynamics

The Metropolis technique described above in Sec. 5.3, was intended to investigate the thermodynamically equilibrium state of a system. However, the approach to the equilibrium state lies in the foundation of this technique from the very beginning. Consequently, the master equation (5.69) may be interpreted as the equation which describes “time evolution” of the system. However, the dynamics obtained in this way, is not true, but the *Markovian dynamics*, i.e. the dynamics “without a memory of the past”. Such a dynamics is called the *relaxation dynamics*. In more detail, the Metropolis procedure uses the detailed balance condition (5.72) (known also as the *microcanonical reversibility condition*), which is usually not satisfied in real dynamics. For the same reasons, inertional effects fall out the description, and in the result, wave processes are lost in the relaxation dynamics.

As the temporal coordinate in the relaxation dynamics, the Monte Carlo step per particle is used which may be connected with the real time approximately only. Recall that for the correct description of the equilibrium state, the transition probabilities per unit time, W_{ij} , must satisfy the unique condition, the detailed balance condition (5.72) which fixes only the ratio of the forward and backward transitions $\xi_{n-1} \leftrightarrow \xi_n$. On the other hand, a rate of dynamical processes in the system is determined not only by the ratio W_{ij}/W_{ji} , but by the values of individual probabilities W_{ij} as well. According to the Kramers theory (Sec. 4.5), the probability W_{ij} itself depends exponentially on the activation energy for the given transition, which in turn depends on both configurations ξ_{n-1} and ξ_n (for example, on the positions and states of the neighboring atoms interacting with the “jumping” atom). The incorporation of these reasons into consideration significantly increases the number of model parameters. However, in order to connect a given MC step with the real time, we have to take into account the reasons described above and to modify the Metropolis algorithm in an appropriate way.

Nevertheless, the relaxation dynamics is very interesting, because it exhibits many characteristic features of true dynamics of the system. Besides, the relaxation dynamics may be carried out for an essentially longer time interval as well as for systems consisting of a much larger number of particles than the MD simulation of the real dynamics.

The main idea of using of the Metropolis procedure in studying the system dynamics, is to take as the initial configuration a state ξ_0 which is realistic from the physical viewpoint. For example, in investigation of the submonolayer adsorbed film, we may take for ξ_0 the configuration where the atoms occupy the left-half part of the lattice only. Then the sequence of configurations $\{\xi_n\}$ obtained with the Metropolis algorithm, will describe the system evolution to the equilibrium state, i.e. the “smoothing” of the initial step in atomic concentration to the state where the atoms are uniformly distributed over the surface [145]. Fixing the MC configurations through the equal time intervals, calculating the concentration of adatoms as a function of the spacial coordinate, and processing the simulation data with the help of the Motana-Boltzmann technique, we can find the chemical diffusion coefficient D_c as a function of the concentration of adatoms analogously as it is done in laboratory experiments on surface diffusion [92]. Besides, we may analyze in detail the evolution of the diffusion front which usually is characterized by a fractal structure [189].

As the next example, let us note the using of the MC simulation in investigation of dynamics of phase transitions, for example, the crystallization of a liquid when it is cooling below the critical temperature T_{crit} of the crystal-liquid phase transition. In this case the Metropolis procedure is carried out for a temperature T lower than T_{crit} , while as the initial configuration, we have to choose a configuration corresponding to the thermal equilibrium state of the system at $T = T_{\text{ini}} > T_{\text{crit}}$ (such a configuration may be obtained from the preliminary performed MC simulation for the $T = T_{\text{ini}}$ case or, more crude, simply by taking the random atomic coordinates, the latter corresponds to the $T_{\text{ini}} = \infty$ case). Then, generating the sequence of the MC configurations $\{\xi_n\}$, we may follow in detail how, owing to thermal fluctuations, small “drops” or “clusters” of the new phase emerge on the background of the old (unstable at the simulation temperature) phase, and how these drops either growth to macroscopic sizes (if they exceed a “critical” size) and coagulate, or disappear [146]. The results of such a modeling may be used to test different theoretical predictions. Note, however, that if the model uses a cutting radius r^* for the interatomic potential (as it was done in the particle-particle MD method), the simulation results for the dynamics of the first-order phase transition occur to be very sensitive to a choice of the value r^* [99]. Besides of such “homogeneous” nucleation, a large interest has the simulation of the “heterogeneous” nucleation, i.e. the nucleation at various irregularities of the system such as impurities, surfaces, interfaces, *etc.*

When the procedure described above, is carried out for a temperature $T \ll T_{\text{crit}}$, we may simulate a process of phase separation at fast cooling of, e.g., a liquid alloy. For example, Fig. 5.10 demonstrates snapshot pictures for the two-dimensional lattice model of a binary A-B alloy [147], where a process of crystalline growth is seen clearly. Such computer experiments are very important for the metallurgy.

The simulation technique for continuous phase transitions is the same. For example, Salid and Binder [148] investigated the submonolayer film with the coverage $\theta = 1/2$, adsorbed on the centered rectangular (rhomboidal) lattice. This model describes an adsorption on the (110) face of the b.c.c. crystal, for example,

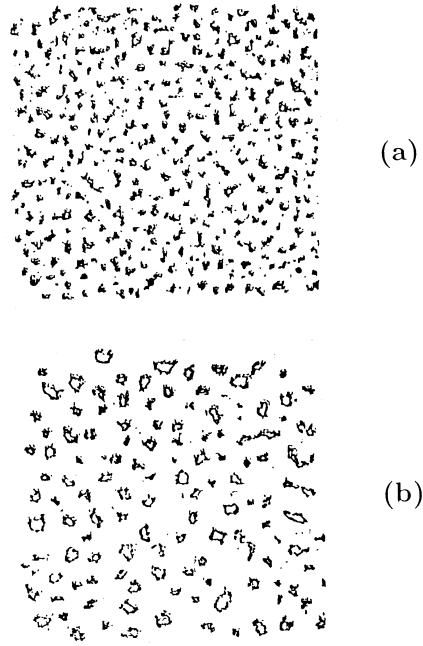


Figure 5.10: Snapshot pictures for the A-B alloy modeled by the 200x200 lattice-gas system at $\langle n_B \rangle = 0.2$ for (a) $t = 500$ and (b) $t = 8925$ Monte Carlo steps per atom after quench to the temperature $T = 0.6T_{\text{crit}}$. Black points correspond to the B-atoms, while the A-atoms are not shown (after Rao *et al.* [147]).

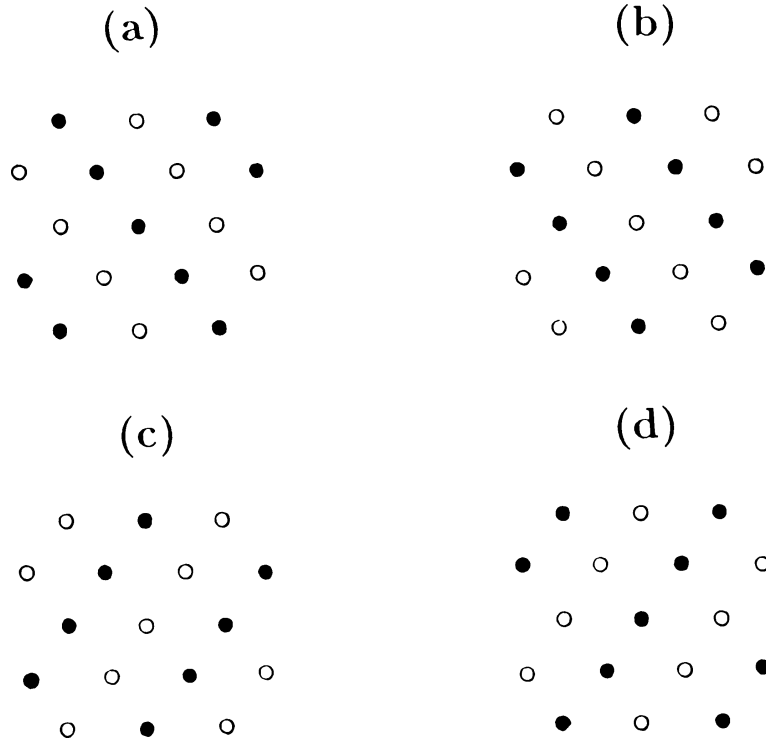


Figure 5.11: Four different realizations of the $p(2 \times 1)$ structure on the square lattice.

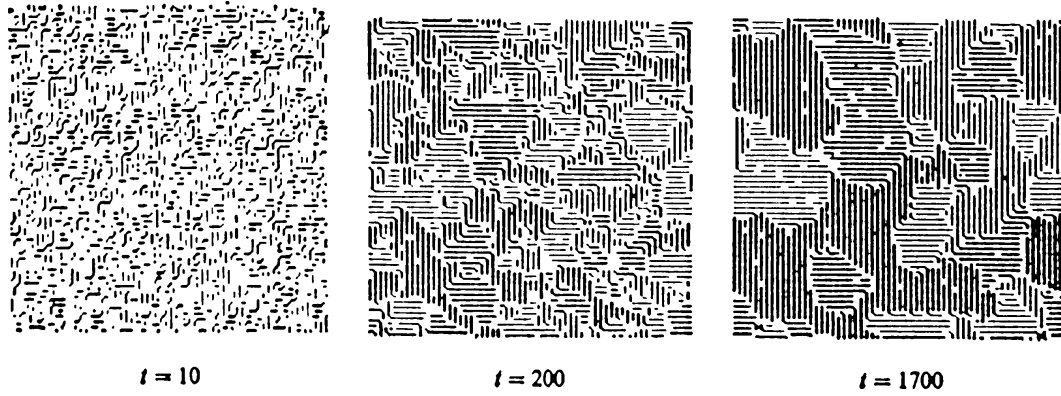


Figure 5.12: Snapshot simulation pictures for the time evolution of the quenched lattice gas model. The symmetry of the model corresponds to the $p(2 \times 1)$ **O-W**(110) adsorption system. Differently shaded regions correspond to the $p(2 \times 1)$ domains with different realizations of the $p(2 \times 1)$ structure as is shown in Fig. 5.11. The denoted “time” t corresponds to the number of the Monte Carlo step per atom (after [148]).

the **O-W**(110) adsystem. The $T = 0$ ground state of this system corresponds to the $p(2 \times 1)$ structure, and this ground state is four-fold degenerated as shown in Fig. 5.11. The results of the MC simulation at quench of the system to a temperature $T < T_{\text{crit}}$ shown in Fig. 5.12, clearly demonstrate the dynamics of fluctuationally-driven creation and further evolution of domains of the $p(2 \times 1)$ structure.

It is important to note that the relaxation dynamics is very sensitive to the concrete realization of elementary MC steps. For example, when we investigate an adlayer with the help of the two-dimensional lattice-gas model with Kawasaki dynamics, the equilibrium state does not depend on the distance of the atomic jump in a single MC act. But the system dynamics obtained with the jumps to the nearest neighboring sites only, differs essentially from the dynamics which assumes additionally the possibility of “long jumps” to more removed lattice sites.

The relaxation dynamics is widely used to study the heterogeneous crystal growth. For example, the growth of the first layer on some substrate may be investigated with the help of the two-dimensional lattice model with the Hamiltonian (5.85), while the growth of three-dimensional crystals, with the help of the SOS model (5.89). In both cases we have to take into account the contact of the system with the particle reservoir (gas over the crystal surface). As the initial configuration it is natural to take an empty lattice, and as the elementary MC steps, the adsorption-desorption processes as well as the diffusion jumps. Recall that with the SOS model we can simulate the growth of the perfect crystal only, because the creation of bubbles (vacancies) is forbidden in this model. Such a model is not too realistic at temperatures close to the melting temperature. But for simulation of the crystal growth from a beam or a solution at strongly nonequilibrium conditions, the SOS model is quite satisfactory.

It is important to note that if in the SOS model we take for the chemical potential μ of the particle’s reservoir (gas) a value, which is not equal to the value of the chemical potential μ_s characteristic for the equilibrium state of an ideal crystal at the given temperature, then the growth process at $\mu > \mu_s$ (or the evaporation of the crystal if $\mu < \mu_s$) will continue for an infinitely long time, because the model reservoir is “inexhaustible”. In this case the sequence of the MC configurations $\{\xi_n\}$ approaches some stationary (but nonequilibrium) state as $n \rightarrow \infty$. Because new atoms income to the surface at random sites, the shape of the growing crystal is rough always (the so-called *kinetic rough*). A degree of this rough is high, when the rate of adsorption-desorption processes significantly exceeds the rate of diffusion, because the diffusion processes try to “smooth” surface defects.

From experiments it is known that the heterogeneous crystal growth may be carried out by one among three different mechanisms: first, by the *Frank-van der Merwe* (layer-by-layer) mechanism, second, by the *Vollmer-Weber* mechanism (when separate islands of three-dimensional crystals start to grow from the

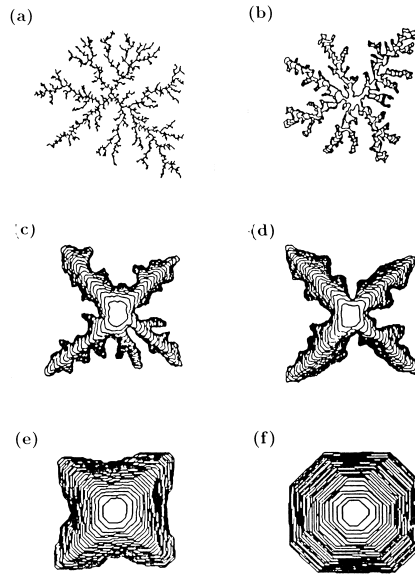


Figure 5.13: Diffusion limited growth of two-dimensional crystals for different diffusion times n_τ . (a) $n_\tau = 0$, 3600 atoms, (b) $n_\tau = 10$, 11000 atoms, (c) $n_\tau = 50$, 19000 atoms, (d) $n_\tau = 100$, 23000 atoms, (e) $n_\tau = 1000$, 23000 atoms, and (f) $n_\tau = 10000$, 27000 atoms aggregated into the crystal (after [152]).

very beginning), and third, by the *Stranski-Krastanov* mechanism (when the first layer grows by the first mechanism, while the second and consequent layers, by the second mechanism). All these scenarios may be studied with the MC simulation, if we take into account in the Hamiltonian (5.89) of the SOS model, that the one-particle energies $\varepsilon_{(h)}$ and the parameters ϕ for the interatomic interactions, may take one values in the first layer, and another ones in the second and consequent layers [149].

Finally, choosing an appropriate initial configuration ξ_0 , we may model the crystal growth over the screw dislocation, or on the stepped (vicinal) surface, *etc.* (e.g., see Müller-Krumbhaar’s article in the book [132]).

In the simulation process we may support artificially an unstable state, for example, by injecting constantly new particles into the system as it was described above. As an example, let us consider also the so-called *diffusion limited crystal growth* (e.g., see [150, 151]), and, for a change, consider a case when the surface of the growing crystal is not flat.

Above (in Sec. 3.4) in the discussion of the simulation of the crystal growth by the MD method, we have noted that the energy of incoming atoms may essentially exceed the average thermal energy of the substrate atoms. In this case the incoming atom gives the surplus of energy into the substrate for some “thermalization time” τ . During the time τ the atom has an abnormally high mobility, and it may do several jumps on the surface even at zero substrate temperature. Now we demonstrate how this process may be modeled with the MC method [152].

Let us take the 2D lattice model and place in its center a “nucleus” (a single immobile atom), to which the incoming atoms can “stick”. Then, let us inject atoms one by one into the lattice (for example, from the left-hand side of the lattice), and allow for the atoms to walk randomly over the lattice until they meet with the nucleus. Then, let us suppose that after the meeting with the nucleus, the atom can carry out on the nucleus’ surface a limited number of random walks only, e.g., n_τ walks, and after that the atom must be irreversibly aggregated on the visiting site having the maximum number of nearest neighbors. If $n_\tau \gg 1$, then with a high probability during the diffusion over the surface of the growing crystal, the atom will find a site with the lowest potential energy (i.e., with the maximum number of neighbors), so that the atom will join itself to a “true” site of the growing crystal. In a result, the growing crystal is the “ideal” isometric

octagonal crystal bounded by equally developed (10) and (11) faces as is shown in Fig. 5.13f. Otherwise, if $n_\tau = 0$, the incoming atom “sticks” to a random place on the surface, and in the result the random fractal pattern with the fractal dimension $d = 1.70$ is growing (see Fig. 5.13a). Finally, at intermediate values of the allowed number of random walks n_τ , the so-called *dendritic growth* is observed (for example, see the $n_\tau = 100$ case in Fig. 5.13d) [190]. In the latter case, it is growing a clear anisotropic pattern with protrusions, which are then transformed into well defined branches, and latter produce new protrusions, *etc.*

5.6 Accelerating relaxation to equilibrium

As was shown in the previous Sec. 5.5, the Metropolis procedure imitates successfully the system evolution to the equilibrium state. In real life, however, a single “attempt” of the transition to a new configuration takes a time $\sim 10^{-13}$ second, and all particles carry out these attempts simultaneously. Consequently, for one second a real system makes $\sim 10^{13}N_p$ attempts. On the other hand, in the computer experiment the particles try to pass to a new site one by one, and the number of these attempts per one second does not exceed $10^5 \div 10^8$. Therefore, “physical” times of approach to equilibrium in laboratory and computer experiments differ in $(10^5 \div 10^8)N_p$ times! Thus, if the time of approach to equilibrium in a laboratory experiment is, for example, about few seconds, in the computer experiment we should wait for the achieving of equilibrium for few years as minimum.

One may indicate four classes of systems characterized by slow relaxation to equilibrium.

(a) All models where some conservation laws must be satisfied, exhibit the so-called *hydrodynamical slowing* (e.g., see [153]).

(b) All systems near the point of the second-order phase transition, exhibit the so-called *critical slowing*.

(c) The so-called *metastable slowing* is observed at first-order phase transitions, because for formation of a new phase, first islands (drops) of the new phase with critical sizes should be created, and this process typically include overcoming of an activation barrier.

(d) Some systems (such as glasses, overcooling alloys, *etc.*) have a number of metastable states characterized by an energy which is only slightly above the energy of the ground state. During the cooling, these systems with a very high probability are locked in one of the metastable states, and then the system may get out from the metastable state with simultaneous motion of many particles only.

Evidently that in all these cases, the Metropolis algorithm is in need of a modification. An original method of acceleration of the relaxation to equilibrium in the MC simulation, based on Darwin’s evolution laws, was proposed by Shumway and Sethna [154]. The idea of the method consists in using of “adaptable” Monte Carlo steps, which are modified during the system evolution in such a way that to achieve the most fast approach to the equilibrium.

Recall briefly the standard Metropolis procedure. At each step of construction of the sequence $\{\xi_n\}$ of the MC configurations, we have constructed from the old configuration ξ_{n-1} the trial configuration ξ^* , and then we have been decided, either to take ξ^* as the new configuration ξ_n , or to throw it out. At this step, we have selected randomly only one variant among a set of variants of the ξ^* construction (for example, we have chosen randomly the particle’s number, the length and the direction of the displacement). Let the set of different variants of construction of ξ^* from ξ_{n-1} constitutes some set (“population”) \mathcal{K} consisting of a finite number K of elements. Some of these variants (elements of the set \mathcal{K}) approach the system to the equilibrium, others are wittingly unacceptable. The idea of the Shumway and Sethna method is to select successful variants and throw out the unsuccessful ones and, moreover, just during the simulation to modify purposefully the procedure of the ξ^* construction with the aim to lower the number of unsuccessful attempts.

For this, let associate some “health” coefficient h_k , where $k = 1, 2, \dots, K$, with each “member” of the “population” \mathcal{K} (i.e., with each variant of the ξ^* construction). At the beginning of the construction of the sequence $\{\xi_n\}$, we put $h_k := 1$ for all k . Then, during the simulation we will change the coefficients h_k , increasing them for the “successful” variants of the ξ^* construction, and decreasing for the “unsuccessful” members of the population \mathcal{K} . Namely, let at the n -th step of construction of the sequence $\{\xi_n\}$, we have

chosen with the help of the random number generator one of the members of the population \mathcal{K} , for example, with the number k_0 ($1 \leq k_0 \leq K$), as the variant of the ξ^* construction. Then, after the construction and testing the trial configuration ξ^* , we change the health coefficient for the k_0 -th element in the following way: if the transition $\xi_{n-1} \rightarrow \xi^*$ lowers the system potential energy, $\Delta V < 0$, we put

$$h_{k_0} := \alpha h_{k_0} + \beta |\Delta V|, \quad (5.92)$$

where α and β are some positive numerical coefficients, and $\alpha < 1$. Otherwise, if the attempt of the transition leads to increasing of the potential energy, $\Delta V > 0$, we simply put

$$h_{k_0} := \alpha h_{k_0}. \quad (5.93)$$

Clearly that if some of the variants of the ξ^* construction will lead to increasing of the system energy for the each time, its health coefficient will constantly decrease.

Now let us introduce some “critical health level” h_{crit} ($0 < h_{\text{crit}} < 1$). In the case when the value h_k for some variant k will become lower than the critical level, $h_k < h_{\text{crit}}$, let throw out the k -th variant from the set \mathcal{K} of the ξ^* construction, and instead of it let us introduce some new variant of the ξ^* construction. For this, let choose two new elements k and k' from the set \mathcal{K} in a random way, but taking into account their health coefficients (i.e., an element with a higher h_k value should be chosen with a higher probability). Then, using the k and k' variants, let us combine a new variant of the ξ^* construction, and include this variant into the set \mathcal{K} instead of the removing one. The concrete procedure of construction of the new element depends on the model under investigation. For example, if the variants k' and k'' describe the displacement of one of the system particles on distances $\Delta r'$ and $\Delta r''$ correspondingly, then as the new variant we may take the displacement of the particle on the distance $\Delta r' + \Delta r''$ or $\Delta r' - \Delta r''$. Or, if the procedure of the ξ^* construction includes a simultaneous shift of few particles, and the variants k' and k'' correspond to shifts of n' and n'' particles respectively, then the new variant may correspond to the simultaneous shift of $n' + n''$ particles, *etc.* Clearly that the health coefficient of the new variant must be put to one.

During the carrying out of the procedure described above, the “unsuccessful” elements of the population \mathcal{K} will be throw out (“died off”), while the “successful” elements will “survive”, their health coefficients will increase and, moreover, they will be “reproduced”, creating new elements (“offsprings”) instead of the “died off”. However, this procedure has a nonzero probability that at some step the population \mathcal{K} may loss of diversity and collapse. To avoid such a scenario, from time to time we may introduce “mutations” into the population \mathcal{K} , for example, multiplying a randomly chosen element by some random number. A rate of producing of new “mutations” has to be not too slow (because this will not save the population against the collapse) as well as not too high (because in the latter case the useful role of the health coefficients will be lost). An optimum may correspond to the rate of “mutation” producing, which is of the same order as the rate of creation of “offsprings”.

The concrete values for the coefficients α , β and h_{crit} may be chosen at the training of the computer program. Shumway and Sethna [154] applied the described method to study the Frenkel-Kontorova model (see below Sec. 6.4.2), where with temperature decreasing the system is locked in a glass-like state. Using the method described above, they achieved the acceleration of the approach to equilibrium (comparing with the standard Metropolis algorithm) in $10^4 \div 10^6$ times. During the simulation, the maximal distance of atomic displacements per one MC step were risen from $0.1a_0$ at the beginning to $0.6a_0$ at the end of simulation, and the number of simultaneously shifted atoms were increased from one to few dozens. Note, however, that the sequence $\{\xi_n\}$ calculated by the described above method, is not Markovian.

5.7 Comparison of the MD and MK methods

In conclusion let us compare the applicability of the two main methods of computer simulation to statistical physics problems.

Molecular dynamics method

1. One may check, whether the system under investigation is integrable or non-integrable, and in the latter case one can test, does the “developed chaos” exists in the system, i.e., may the thermal equilibrium be achieved.
2. Temporal averages are calculated.
3. The method may be used for investigation of the far from equilibrium systems, as well as for studying non-stationary processes.
4. The method gives the true dynamics of the system, including the description of wave processes.
5. The method is usually realized with a complex computational algorithm, needs a long processor time on a powerful computer, and can simulate a system consisting of a comparative small number of particles.

From this comparison one can see that each of the methods has its own dignities as well as deficiencies. Therefore, the choice of a suitable method first of all depends on the problem under investigation.

Monte Carlo method

1. The existence of a contact with a thermal bath, which guaranties that the thermal equilibrium will always be achieved, is postulated.
2. Configurational averages (the averages over the ensemble) are calculated.
3. The applicability of the method is strictly verified for the thermodynamically equilibrium state of the system only.
4. The method gives the relaxation dynamics of the system, which only approximately imitates the approach to the equilibrium. All vibrational processes are lost.
5. The method may be realized with a comparative simple and universal algorithm, allows to model systems of a large number of particles during an essentially longer time interval.