Lecture notes on Monte Carlo and Molecular Dynamics Simulations [*]

Mauro Rovere

Department of Physics “E. Amaldi”
Università degli Studi Roma Tre
Via della Vasca Navale 84 - Roma - Italy
e-mail: rovere@fis.uniroma3.it

1 Introduction

In these lecture notes we intend to give a short introduction to the methods of computer simulation developed to study the physics of liquids [1, 2]. With the rapid evolution of the computational technology computer simulation has become, from a simple support to theoretical approaches and experiments, a third methodology able to set up a complete phenomenology of model systems. In statistical mechanics very few models can be exactly solved. For many years this has been a limitation for the theoretical studies of liquids, since the liquid state is too far from both the ideal gas and the harmonic crystal [3]. To study liquids and different complex systems, like amorphous solid, biological matter etc.. it is necessary to propose realistic models that are too complex to be solved analytically without approximations. Numerical simulations can be used to perform virtual experiments on different fluid models making possible to test their predictive abilities. Moreover it is possible to reveal details on the behaviour of the systems that are not easy to observe in experiments. In this way computer simulation has given an essential contribution to the progress of the physics of liquids. In the same period of time the experimentalists mainly with the development of neutron scattering techniques have improved the accuracy of our knowledge of the structure and particle dynamics of the liquid state.

---

[*] School of Neutron Scattering “F. P. Ricci”, Santa Margherita di Pula, 22 Sep.- 3 Oct. 2008

Of course one cannot forget that the systems studied in computer simulation must be validated against experiments. In this respect neutron scattering technique represent one of the best experimental methods to compare with computer simulation. Neutron scattering on fluids is able to give detailed information on the microscopic structure and the dynamical behaviour of liquids more than other techniques. We refer in particular to systems containing hydrogens. From one side the results of neutron scattering experiments can be better interpreted with the support of simulation, on the other side if experiments give a detailed phenomenology on the systems they can help to refine the model to use in simulations. So the mutual interaction between neutron scattering techniques and computer simulation methods plays a key role in the understanding the properties of molecular substances.

We consider here two techniques of computer simulation. The first one to be invented was the Monte Carlo (MC) method, developed during the second world war at Los Alamos [4]. Sampling with random numbers were implemented long time before, we recall the famous examples of the naturalist Buffon in the eighteenth century and the Italian mathematician Lazerini in 1901. Fermi used random numbers to make predictions on nuclear fission processes. Metropolis, an American physicist of Greek origin, collaborated with von Neumann, Ulam and Fermi to improve the use of random sampling for applications in statistical mechanics. The name of the method was proposed by an uncle of Ulam, who was an experienced player of Casino games. Metropolis after the war was able to build a computer called MANIAC completely devoted to perform Monte Carlo simulations. The first simulation on an hard disk fluid has been published in 1953 [5].

Some years later in 1957 Alder and Wainwright [6] invented a method, called Molecular Dynamics (MD), based on the numerical solution of the Newton equations for a large number of interacting particles. At variance with Monte Carlo which is a stochastic method, Molecular Dynamics is deterministic and it is the only method able to study the dynamical behaviour of the systems, as we will see later. The simulation of Alder and Wainwright was done on a hard disk system. The first simulation on more a realistic fluid is due to Rahman [7]. He studied with MD a Lennard-Jones liquid (see below for the definition). Later in 1971 Rahman and Stillinger made the first MD simulation of liquid water [8], already simulated with MC by Barker and Watts [9]. For a complete account of the history of the computer simulation methods we refer to the already cited textbooks [1, 2].

2 Model the system of atoms for computer simulation

2.1 Born-Oppenheimer approximation

We are interested in systems composed by a large number of atoms. If the system contains N nuclei of charge Z with coordinates $R = (R_1, R_2, ..., R_N)$ and M electrons with coordinates $r = (r_1, r_2, ..., r_M)$ such that $M = ZN$ its Hamiltonian can be written as

$$H = K_N(R) + K_e(r) + V_{eN}(r, R) + V_{NN}(R) + V_{ee}(r) \tag{1}$$

where $K_{\{...\}}$ are the kinetic energy operators and the $V_{\{...\}}$ are the interaction potential between the different particles. By assuming the Born-Oppenheimer approximation the total wave function can be written as the product

$$\Psi(r, R) \simeq \chi(R) \cdot \phi(r, \{R\}) \tag{2}$$

where now the coordinates R appear as fixed parameters in the electronic problem. The approximation is based on the adiabatic decoupling of the slow motion of the nuclei with respect to the fast dynamics of the electrons. The problem becomes separated in two Schrödinger equations, for the electrons

$$[K_e(r) + V_{eN}(r, \{R\}) + V_{ee}(r)]\, \phi(r, \{R\}) = E_{el}(\{R\})\, \phi(r, \{R\}) \tag{3}$$

and for the nuclei

$$[K_N(R) + V_{NN}(R) + E_{el}(R)]\, \chi(R) = E\chi(R) \tag{4}$$

It is usual to consider Eq.(3) only for the valence electrons while the role of the core electrons is to screen the nucleus. In the hamiltonian the coulombic potential V_{eN} is replaced with an effective interaction between the valence electrons and a rigid ion (method of the pseudopotential [10]). In this way the ion dynamics is determined by the time dependent Schrödinger equation

$$\left\{ K_I(R) + \sum_l \sum_{k>l} \frac{Z_l Z_k e^2}{R_{kl}} + E_{el}(R) \right\} \chi(R, t) = i\hbar \frac{\partial \chi}{\partial t} \tag{5}$$

We can define the potential in which the ions move as

$$U^{BO}(R) = \sum_l \sum_{k>l} \frac{Z_l Z_k e^2}{R_{kl}} + E_{el}(R) \tag{6}$$

2.2 Classical approximation

Apart some anomalous case, like helium, all the systems in the fluid phases can be treated in the classical approximation [3]. In the range of temperatures and densities of fluids the thermal De Broglie wavelength

$$\Lambda = \sqrt{\frac{h}{2m\pi k_B T}} \tag{7}$$

results to be

$$\Lambda \ll a \tag{8}$$

where a is the mean distance between the particles related to the density ρ

$$a = \left(\frac{1}{4\pi\rho}\right)^{1/3} \tag{9}$$

2.3 Effective potentials

The potential to be used in computer simulation is generally approximated with an empirical potential. To obtain U^{BO} in fact it would require to solve the problem of the electrons for each move of the ions. The usual way to proceed in computer simulation of liquids is to assume that the electrons are able to follow the motion of the ions and remain on the BO equilibrium surface. Then the first step is to expand the BO potential in n-body terms

$$U^{BO}(R) = \sum_i \sum_{j>i} u^{(2)}(R_i, R_j) + \sum_i \sum_j \sum_k u^{(3)}(R_i, R_j, R_k) + ... \tag{10}$$

Since a two-body potential is more easy to use and less expensive from the point of view of computation the second step consists in assuming a pair potential based on empirical properties. The empirical potential contains some parameters to be fitted. A typical and well known potential is the Lennard-Jones (LJ), which is valid for describing noble gases in the fluid phases

$$u(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{11}$$

where ε and σ are parameters to determine by fitting empirical properties. We note that the 6-power law in the attractive part comes from the quantum mechanical calculations of the Van der Waals dipole-dipole attraction between atoms. We have not time to go in more details, we only recall that empirical potentials work well usually for closed shell atoms, like rare gases or simple ionic liquids. For other cases like silicon atoms different methods of simulations have been developed and we will briefly consider this point below (see beginning of Sec.3).

Empirical potentials are used also for many molecular fluids, for instance water. In water the hydrogen bond formation can be well reproduced by means of site potentials. The molecule is assumed, in the simplest approach, to be rigid with the proper geometry of H_2O. To model the intermolecular interaction the oxygen and the two hydrogens of the molecule are replaced with point fractional charges, two positive charges for the hydrogens and a negative one for the oxygen. In some model the charge of the oxygen is shifted with respect to the position of the atomic oxygen. The intermolecular interaction is composed by coulombic potentials between the point charges and a LJ potential between oxygens.

2.4 Basic methodology

The computer simulation is usually implemented by assuming a system of N particles in a cubic box of length L and volume $\Omega = L^3$. Since N is typically of the order of $100 \div 10000$, very far from the thermodynamics limit, to avoid surface effects, periodic boundary conditions (PBC) are applied in the three directions. In particular cases the PBC can be relaxed in some of the directions for simulating interfacial phenomena. In principle each particle would interact with all the particles in the box and their images in the periodic repetition of the box. Since the calculation of the forces is particularly time consuming one assumes the so called minimum image convention. A particle is considered at the center of a box of size L and interacts only with the other particles or alternatively with their images that are in this box. In this way the maximum range of the potential of each particles is $L/2$. Sometimes if the potential is short ranged it is truncated at a cut off $r_c \leq L/2$. The discontinuity of the potential at the cut off does not give problems in MC. In the case of MD instead the discontinuity can produce a spurious contribution to the forces so it is common to use a truncated and shifted potential: $u(r)$ becomes $u(r) - u(r_c)$ for $r \leq r_c$ and zero for $r > r_c$ [1, 2]. One must take into account in comparing different simulation studies that the results could depend on the choice of r_c. Long range corrections can be applied to thermodynamical quantities to take into account the truncation of the potential, as we will see below.

If the numerical procedure and the algorithms are correct the final results of computer simulation must be independent from the starting conditions. It is true however that at the beginning we have to place the particles in the simulation box and the configuration that we can produce is off equilibrium. By running the simulation the system finally would reach an equilibrated point of the phase space, if we are not far from ergodic conditions. It is convenient to start with the particles placed in a crystalline phase since they are sufficiently far apart. If two particles are very close the repulsion between them can produce a large positive potential energy and a very large kinetic energy. To dissipate such large energy contribution one would need a very long equilibration run. To avoid this problem it is better to start with a crystalline system and melt it at high temperatures.

2.5 Long range corrections

As said above it is usual in computer simulation to truncate the potential to avoid very long computation. We can rewrite the potential as

$$U = U_C + U_{LR} \tag{12}$$

where

$$U_C = \sum_i \sum_{j>i} u\left(r_{ij} < r_c\right) \tag{13}$$

and U_{LR} is the long range term. Now we know from the theory of liquids that the potential energy can be obtained from the pair correlation function $g(r)$

$$U = \int_0^\infty 2\pi\rho r^2 u(r) g(r) dr \tag{14}$$

If we assume that $g(r) = 1$ for $r > r_c$ we can write a simple formula for the long range correction in Eq.(12)

$$U_{LR} \approx \int_{r_c}^\infty 2\pi\rho r^2 u(r) dr \tag{15}$$

The long range corrections can be calculated once in the simulation program. For simple form of the potential, like LJ, the integral in Eq.(15) can be evaluated analytically.

When the potential is long range, like a coulombic term, the simple formula (15) does not give the right estimation of the long range contributions. Neglecting the long range corrections in a problem where the particles interact with a coulombic potential could produce wrong results. One must take into account the interaction of each particle with all the particles and periodic images. In the case of a coulombic potential we have

$$U = \frac{1}{2} \sum_{\mathbf{n}} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{Z_i Z_j}{\left|r_{ij} + \mathbf{n}L\right|} \tag{16}$$

where the sum over $\mathbf{n} = (n_x, n_y, n_z)$ is the sum over all the periodic images, for $\mathbf{n} = 0$ the terms $i = j$ must be avoided. The problem becomes similar to the calculation of the potential energy of an ionic crystal and can be treated with a similar method, called the Ewald sum method. A complete discussion of the application of the Ewald sums to computer simulation can be found in the textbooks [1, 2].

3 Molecular Dynamics Methods

The first method of simulation that we consider is Molecular Dynamics (MD). The time evolution of the atoms can be calculated on the basis of the Newton equations

$$M\ddot{r_i} = -\nabla_i U(r_1, ..., r_N) \tag{17}$$

The potential to introduce in Eq.(17) is usually approximated with a two-body effective potential. In recent years methods of first principles MD have been developed, where the classical dynamics of the atoms is coupled to a fictitious dynamics of electrons, in this way it is possible to avoid the assumption of an empirical potential. The electronic problem is solved during the evolution of the atomic dynamics, *on the fly*, in the framework of density functional theory [11–13]. The description of this type of computer simulation is however behind the scopes of the present lectures.

3.1 Molecular dynamics and statistical mechanics

If we consider system composed of N particles with an Hamiltonian

$$H(p^N, r^N) = K(p^N) + U(r^N) \tag{18}$$

the positions and the momenta of the system evolve in time according to the equations

$$\dot{r_i} = \frac{\partial H}{\partial p_i} \qquad \dot{p_i} = -\frac{\partial H}{\partial r_i} \tag{19}$$

From the time evolution of (p^N, r^N) it is possible to derive the time average value of an observable $A(p^N, r^N)$ defined as

$$\bar{A} = \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau dt A\left[p^N(t), r^N(t)\right] \tag{20}$$

Statistical mechanics is based on the ergodic hypothesis that the time average (20) is equivalent to the ensemble average

$$\langle A \rangle_{ensemble} = \bar{A} \tag{21}$$

In MD the Eq.(19) are solved numerically. In the numerical procedure the time variable is discretize by introducing a finite time step Δt, in this way $t \rightarrow t_k = k\Delta t$ and the average is performed on a finite time $\tau = n\Delta t$

$$(\bar{A})_{MD} \simeq \frac{1}{n} \sum_{k=1}^{n} A\left[p^N(t_k), r^N(t_k)\right] \tag{22}$$

In the presence of conservative forces the Hamiltonian is a conserved quantity, so the total energy $E = K + U$ is a constant. From the point of view of statistical mechanics the averages realized in MD are equivalent to averages in the microcanonical ensemble, since the number of particles and the volume are fixed and the energy is constant

$$(\bar{A})_{MD} \simeq \langle A \rangle_{microcan} \tag{23}$$

In the microcanonical ensemble the temperature fluctuates. In the absence of external forces the total momentum is conserved.

3.2 Algorithms for the time evolution

In MD the time evolution of the system takes place along discrete time steps. The trajectory of a particles is not the real one but in the single time step we use a Taylor expansion

$$r(t + \Delta t) = r(t) + \left(\frac{dr}{dt}\right)_t \Delta t + \frac{1}{2}\left(\frac{d^2r}{dt^2}\right)_t \Delta t^2 + ... \tag{24}$$

where the derivatives are determined by the forces at time t. If the expansion, for instance, includes only second order terms it is equivalent to assume a motion with constant acceleration during the time step. Of course this approximation introduces deviations from the real trajectory at each time step. Finally the system can be driven toward wrong directions in the phase space. For this reason we must be very careful in implementing the algorithms to solve the Newton equations on computers. A good algorithm must satisfies at least the following conditions:

It must be simple and fast enough

It must give stable trajectories with *enough* long time step

The temporal evolution must be reversible

It must conserve energy (and momenta)

The most used class of algorithms is called *à la Verlet*, since Verlet in 1967 proposed the first algorithm of this type [14]. We can write down the approximate equations for a step forward and a step backward starting from a time t as

$$\begin{aligned} r(t+\Delta t) &= r(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 \\ r(t-\Delta t) &= r(t) - v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 \end{aligned}$$

By summing both sides we get

$$r(t+\Delta t) = -r(t-\Delta t) + 2r(t) + a(t)\Delta t^2 \tag{25}$$

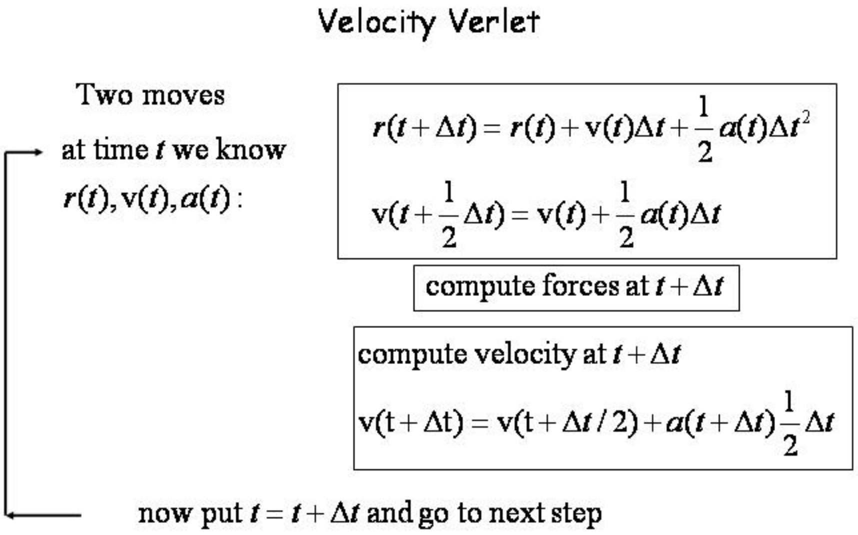
the acceleration is obtained from the potential as

$$a(t) - \frac{1}{m}\nabla U\left(r^N(t)\right) \tag{26}$$

At each time step the velocity can be calculated as

$$v(t) = \frac{r(t+\Delta t) + r(t-\Delta t}{2\Delta t} \tag{27}$$

It can be shown that this simple algorithm satisfies all the conditions written above with a time step that can be considered very long for MD, $\Delta t \simeq 10^{-15}s$ The Verlet formula has been very successful and has been used for long time, since it requires a simple routine and a small amount of memory. The only weak point is that the velocities are not calculated directly. Sometimes more precise values for the velocities are required. With the development of computers *ram* memory was not more a great problem and a modification of the pure Verlet algorithm was proposed, it is called the velocity Verlet algorithm:



There are a number of other algorithms of this class, that we have not time to describe here [1].

3.3 Temperature in the microcanonical ensemble

During the simulation the temperature can be obtained on the basis of the equipartition of the energy from the kinetic energy (in D=3)

$$\frac{3}{2} N k_B T(t) = \frac{1}{2} \sum_i m_i v_i^2 \tag{28}$$

This is the instantaneous temperature. The average temperature can be calculated from the average value of the kinetic energy

$$\frac{3}{2} N k_B T = \left\langle \frac{1}{2} \sum_i m_i v_i^2 \right\rangle \tag{29}$$

3.4 Equilibration procedure

If we want to study a system in the liquid phase it is convenient to start with the particles in a crystal, as said above, and prepare the system at high

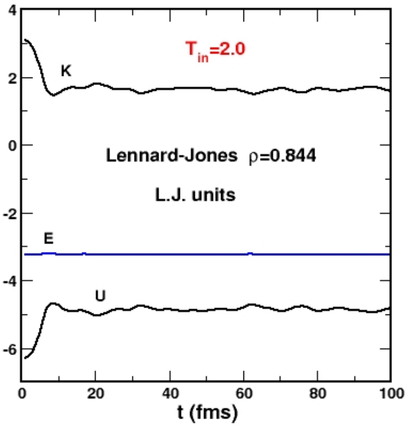


Figure 1: MD:first equilibration

temperature. In this way after few steps, 1−10 picoseconds, the crystal melts and we obtain a liquid. The initial temperature is fixed by assigning to the particles velocities that are distributed according to the Maxwell-Boltzmann distribution. Of course if one is interested in simulating a crystalline phase the starting temperature has to be below the melting.

In the initial simulation run we can check the behaviour of various quantities. In Fig.1 as an example it is shown an initial run for a system with a LJ potential. Reduced units are used where the temperature is $k_B T/\varepsilon$ and the density $\rho\sigma^3$ in terms of the LJ parameter in (11). The critical temperature of this system is estimated around $T_c \approx 1.3$.

It is important to note that the total energy remains constant while both the kinetic and the potential energies relax in few time steps and oscillate around an average value. Take in consideration that the fixed density and the final temperature we want to reach, $T = 0,72$, are in the region of the stable liquid.

To drive the system from an initial T_{in} to a lower temperature T_0 we can use a velocity rescaling procedure. Every fixed number of time steps the velocity of each particle is rescaled by a factor $f = \sqrt{T_0/T_{in}}$ so that

$$\frac{3}{2} N k_B T(t) = \frac{1}{2} \sum_i m_i (v_i/f)^2 = \frac{3}{2} N k_B T_0 \tag{30}$$

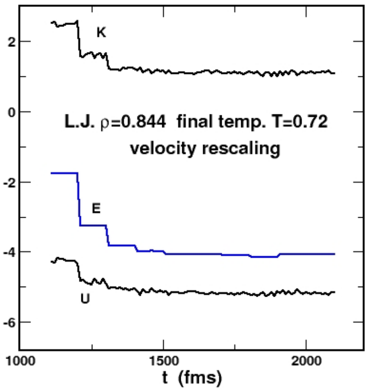


Figure 2: Energies

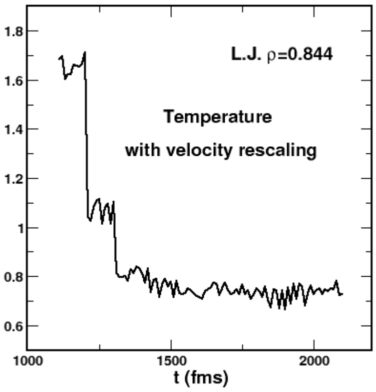


Figure 3: Temperature

In Fig.2-3 there are represented rescaling to intermediate temperatures between $T_{in} = 2.0$ and $T_0 = 0.72$. We can use intermediate temperatures to make the procedure less drastic. We see that every time the rescaling is done the energy is not conserved.

Finally, after we reached the final T, the system must be further equilibrated by switching off the velocity rescaling. We can check the behaviour of the various quantities. In particular we expect that the potential energy and the kinetic energy (the temperature) oscillate around a constant average value, see Fig.4

When the equilibration point is reached we can proceed to calculate average values in the microcanonical ensemble and/or to study the structural and dynamical properties of the liquid.

Apart for the energy also other thermodynamical quantities can be calculated. For instance the pressure can be obtained by the virial theorem

$$P\Omega = Nk_BT - \frac{1}{3}\left\langle \sum_{i=1}^{N} r_i \cdot \nabla_i U \right\rangle \tag{31}$$

Discussions about the use of this formula with PBC appear periodically in the literature, an interesting general approach to the problem can be found in a recent paper [15].

Also the radial distribution function $g(r)$ can be easily obtained from the configurations. If $n^{(2)}(r)$ is the average number of atom pairs in the range

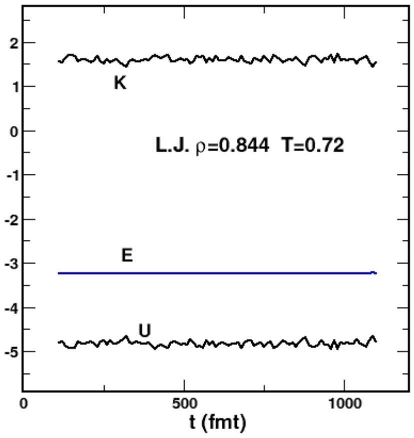


Figure 4: MD: equilibrated system

$(r, r + \delta r)$ and $n^{(id)}(r)$ is the equivalent quantity in an ideal gas

$$n^{(id)}(r) = \frac{4\pi\rho}{3}\left[(r + \delta r)^3 - r^3\right] \tag{32}$$

then

$$g(r) = n^{(2)}(r)/n^{(id)}(r) \tag{33}$$

3.5 MD in different ensembles

The microcanonical ensemble could not be the most convenient for comparing with experiments. The study of dynamical properties, like diffusion or relaxation phenomena, must be performed in this ensemble since the Hamiltonian dynamical behaviour of the system must be reproduced. It is a different story for the thermodynamical properties. Experiments are frequently performed at constant pressure and temperature in the isobaric-isothermal ensemble. On the other hand it could be sometimes more convenient to keep the temperature constant and obtain quantities averaged in the canonical ensemble. In order to perform MD in ensemble different from the microcanonical we need to use a generalized dynamics, which is not more determined by a real Hamiltonian and does not satisfy the conservation rules of the Newtonian dynamics. In practice the system is considered at contact with a reservoir. System of particles and reservoir are an isolated composed system which is microcanonical, instead the system of particles could exchange energy with the reservoir to keep the temperature of the liquid constant. To maintain the pressure constant we must relax the assumption of a constant box of simulation. The volume of the box becomes a variable which is determined by the coupling with the reservoir. In all the cases when the system evolves with a non Hamiltonian dynamics the trajectories of the particles are not realistic. There are still conservation law but they concern the total system composed by the liquid and the reservoir. To give an idea in the case a constant temperature MD, as formulated by Nosè [16], the Hamiltonian is modified by introducing a degree of freedom of the reservoir, called s

$$\tilde{H} = K + U + K_s + U_s \tag{34}$$

where to the kinetic K and the potential U energy of the system it is added a potential energy U_s.

$$U_s = (f+1)k_B T ln(s) \tag{35}$$

where f is the number of degree of freedom of the system and a kinetic energy of the thermal bath

$$K_s = \frac{1}{2}Q\dot{s} \tag{36}$$

where Q is a thermal inertia parameter.

There are a number of algorithms to perform isothermal-isobaric MD simulations. To treat problems of phase transformations in solids it is possible to implements MD where the simulation box can be deformed, this methodology was invented by Parrinello and Rahman [17].

4 Monte Carlo simulation

4.1 Monte Carlo integration and importance sampling

The Monte Carlo (MC) simulation method is based on the idea of calculating the averages of statistical mechanics with the use of numerical integration. The calculation of area or volume by using random numbers is an old idea but nowadays the modern name, introduced by Metropolis, is used for a number of numerical methods based on random numbers that are applied in different field. We consider now how MC can be used for simple integration.. Suppose we have to integrate a function $f(x)$ between a and b

$$F = \int_a^b f(x)dx \tag{37}$$

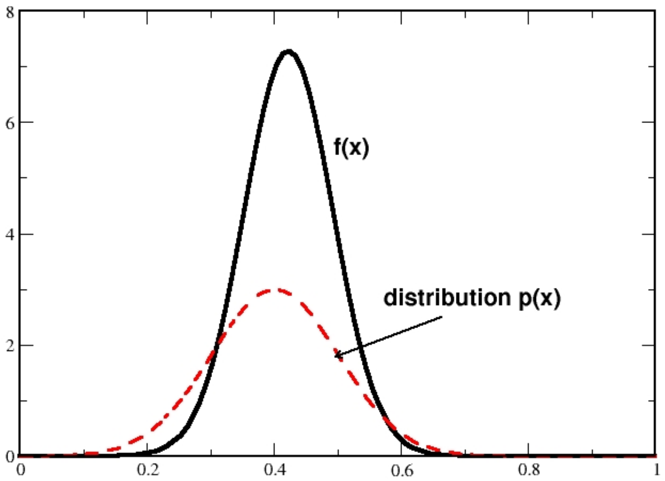


Figure 5: Importance sampling

there are a different methods to perform the numerical integration, one of them, not the more convenient in D=1, is based on the use of the random numbers that a computer can produce. These random numbers are uniformly distributed between 0 and 1. We can extract a series of random number $0 < R_i < 1$ with $i = 1, ..., n$ and calculate (37) as

$$F \approx (b-a)\frac{1}{n}\sum_{i=1}^{n} f(x_i) \qquad x_i = a + (b-a)R_i \tag{38}$$

Like any other numerical integration method also the simple MC could be more or less successful depending on the behaviour of the function $f(x)$. Suppose that the function is as in Fig.5

The important region to sample is the one where the peak is present. In this case the use of uniformly distributed random numbers could give bad results. It would useful to introduce a weighted distribution to maximize the sampling in the important region. This idea, introduced by Von Neumann, is called importance sampling. We can add in the integral a distribution $p(x)$ appropriate for our problem, for instance the Gaussian in Fig.5

$$F = \int_a^b \frac{f(x)}{p(x)}\,[p(x)dx] \tag{39}$$

where $p(x)$ is normalized

$$\int_a^b p(x)dx = 1 \tag{40}$$

Now the integral can be calculated as

$$F \approx (b-a)\frac{1}{n}\sum_{i=1}^{n}\frac{f(x_i}{p(x_i)} \tag{41}$$

where the x_i are distributed according to $p(x)$. $p(x)$ is built from the uniform distribution of the random numbers obtained from the computer. At a first sight it seems that the method can work if one knows very well the behaviour of the function $f(x)$. This could be a difficult task when we have to deal with functions in multidimensional space. Von Neumann proposed a method to guess the appropriate distribution function in more general cases. It is from these ideas of Von Neumann that Metropolis invented his algorithm.

4.2 Integrals in statistical mechanics

Consider now the Boltzmann distribution in the canonical ensemble, if $\alpha = (r_1, ..., r_N; p_1, ..., p_N)$ is a point in the phase space of our system its probability is given by

$$\rho\left[H(\alpha)\right] = \frac{e^{-H(\alpha)/k_BT}}{Q_N(T)} \tag{42}$$

where $H(\alpha) = K(p) + U(r)$ and the partition function Q_N is

$$Q_N = \sum_\alpha e^{-H(\alpha)/k_BT} \tag{43}$$

In classical statistical mechanics if we want to average quantities which do not depend on the momenta we can directly integrate (42) on the momenta. Now the points in the phase space are given by the configurations $\alpha = (r_1, ..., r_N)$ and we can use the probability

$$\rho\left[U(\alpha)\right] = \frac{e^{-U(\alpha)/k_BT}}{Z(T)} \tag{44}$$

where

$$Z = \int d\Omega e^{-U(\alpha)/k_BT} \tag{45}$$

The average value of an observable $A(\alpha)$ is obtained from

$$\langle A\rangle = \int d\alpha A(\alpha)\rho\left[U(\alpha)\right] \tag{46}$$

Is it possible to calculate the integral (46) with a numerical integration ?

4.3 Importance sampling in statistical mechanics

If we want to perform the integral (46) with the important sampling MC technique we have to introduce a distribution function

$$\langle A\rangle = \int d\alpha A(\alpha)\rho\left[U(\alpha)\right] = \int d\alpha A(\alpha)\rho\left[U(\alpha)\right]\frac{p\left(\alpha\right)}{p\left(\alpha\right)} \tag{47}$$

that can be integrated numerically as

$$\langle A\rangle \approx \frac{1}{n}\sum_{k=1}^{n} A(\alpha_k)\rho\left[U(\alpha_k)\right]\frac{1}{p\left(\alpha_k\right)} \tag{48}$$

Thought the integral (46) is multidimensional and the function cannot be plotted down we can guess on the basis of statistical mechanics that the most important region of integration has to be close to the average value $< A >$. This is also the region more sampled by the Boltzmann distribution. So it is reasonable to assume that

$$p\left(\alpha\right) = \rho\left[U(\alpha)\right] \tag{49}$$

With this choice Eq.(48) becomes a very simple formula

$$\langle A\rangle \approx \frac{1}{n}\sum_{k=1}^{n} A(\alpha_k) \tag{50}$$

Of course we can calculate the average values in the form (50) only if the sequence of configurations $\{\alpha_k\}$ is generated according to the Boltzmann distribution (49). The problem now is: how to generate the right configuration sequence ?

4.4 Markov processes

Our system can stay at each time t in one of the states $\{\alpha_1, ..., \alpha_k, ...\}$. There is a probability $p_k(t)$ that the system is in the state α_k at time t, of course this probability satisfies

$$p_k \geq 0 \qquad \sum_k p_k = 1 \tag{51}$$

We want to look at the way in which the system evolves in time. We define the conditional probability that the system could be in a state α_n at the time t_n if it was at t_{n-1} in α_{n-1}, at t_{n-2} in α_{n-2} ...at t_1 in α_1

$$W\left(\alpha_n \mid \alpha_{n-1}, ..., \alpha_1\right) \tag{52}$$

In principle the evolution is determined from what happened before. But we can restrict to Markov processes where

$$W(\alpha_n \mid \alpha_{n-1}, ..., \alpha_1) = W(\alpha_n \mid \alpha_{n-1}) \cdot W(\alpha_{n-1} \mid \alpha_{n-2}) \cdot ... W(\alpha_2 \mid \alpha_1) \quad (53)$$

In a Markov process the evolution is determined only from what happened at the preceding time step. So at each time step the system loses memory of its previous evolution.

We can define the stochastic matrix whose elements are

$$W_{ij} = W(\alpha_j \mid \alpha_i) \quad (54)$$

It satisfies the properties

$$W_{ij} \geq 0 \quad (55)$$

$$\sum_j W_{ij} = 1 \quad (56)$$

The last condition is due to the fact that the system starting from a state α_i must arrives in one of the state α_j.

With the use of the stochastic matrix we can determine how evolves in time the probability p_k defined above. It is easy to see that the probability at a time $t + \Delta t$ can be derived from the one at time t with the equation

$$p_k(t + \Delta t) = p_k(t) - \Delta t \sum_j p_k W_{kj} + \Delta t \sum_j p_j W_{jk} \quad (57)$$

In the continuous limit this becomes the so called master equation

$$\frac{dp_k}{dt} = -\sum_j p_k W_{kj} + \sum_j p_j W_{jk} \quad (58)$$

An equilibrium distribution, as the Boltzmann distribution in statistical mechanics, must satisfies

$$\frac{dp_k}{dt} = 0 \quad (59)$$

4.5 Ergodicity and detailed balance

With the stochastic matrix $\tilde{W}$ we can generate from an initial probability distribution a series of them. The probability distribution is given by the row vector $\mathbf{p} = (p_1, ..., p_k, ...)$ If we start from $\mathbf{p}^{(0)}$ we can generate

$$\mathbf{p}^{(0)} \cdot \tilde{W} = \mathbf{p}^{(1)} \quad (60)$$

This is equivalent to the equation

$$\sum_j p_j^{(0)} W_{ji} = p_i^{(1)} \tag{61}$$

then by applying again $\tilde{W}$ we can have

$$\mathbf{p}^{(1)} \cdot \tilde{W} = \mathbf{p}^{(0)} \cdot \tilde{W}^2 = \mathbf{p}^{(2)} \tag{62}$$

more generally

$$\mathbf{p}^{(0)} \cdot \tilde{W}^n = \mathbf{p}^{(n)} \tag{63}$$

Now we want to find a distribution $\mathbf{p}$ such that

$$\mathbf{p} \cdot \tilde{W} = \mathbf{p} \tag{64}$$

A distribution which satisfies (64) is an equilibrium distribution since it is invariant over transformation. The property of Eq.(64) can be obtained only if the stochastic matrix $\tilde{W}$ has one non degenerate eigenvector with an unitary eigenvalue.

We do not enter into the mathematical treatment, we only recall here that the most important condition for $\tilde{W}$ is that it must be irreducible, i. e. $\forall i, j$ must exist a finite m such that $(W^m)_{ij} > 0$. It means that all the phase space is reachable. From each point in the phase space the system can move to another point without limitations. This is an alternative formulation of the ergodic condition. Under this condition the Perron-Frobenius theorem states that the stochastic matrix has one unitary eigenvalue and the corresponding eigenvector is the limiting distribution of the Markov chain. In this way the $\mathbf{p}$ is independent of the starting point.

A sufficient, not necessary, condition for the matrix to be irreducible is that

$$p_k W_{kj} = p_j W_{jk} \tag{65}$$

From the mathematical point of view it is easy to see that from Eq.(65) we can write, by considering Eq.(56)

$$\sum_k p_k W_{kj} = \sum_k p_j W_{jk} = p_j \tag{66}$$

The condition (65) is called the microscopic reversibility or detailed balance condition. It is also easy to see that if we look at the master equation (58) we get an equilibrium probability distribution, since Eq.(59) is satisfied.

4.6 Metropolis method

We want that our probability distribution is given by

$$p_k = e^{-\beta U(\alpha_k)}/Z \tag{67}$$

where $\beta = 1/k_B T$. To satisfy the detailed balance, Eq.(65), it must be true that

$$\frac{p_k}{p_j} = \frac{W_{jk}}{W_{kj}} = e^{-\beta[U(\alpha_k)-U(\alpha_j)]} \tag{68}$$

Consider that the system goes from the state α_j to the state α_k, if we define

$$\Delta U_{jk} = U(\alpha_k) - U(\alpha_j) \tag{69}$$

from Eq.(68) we get the condition

$$\frac{W_{jk}}{W_{kj}} = e^{-\beta \Delta U_{jk}} \tag{70}$$

Now the Metropolis method is based on the following assumption

$$W_{jk} = \begin{cases} e^{-\beta \Delta U_{jk}} & \text{if} \quad \Delta U_{jk} > 0 \\ \\ 1 & \text{if} \quad \Delta U_{jk} < 0 \end{cases} \tag{71}$$

It is easy to see that the transition probability W_{jk} from j to k given by Eq.(71) satisfies Eq.(70).

- if $\Delta U_{jk} > 0$ then $\Delta U_{kj} < 0 \Rightarrow W_{jk} = e^{-\beta \Delta U_{jk}}$ and W_{kj}=1
- if $\Delta U_{jk} < 0$ then $\Delta U_{kj} > 0 \Rightarrow W_{jk} = 1$ and $W_{kj} = e^{\beta \Delta U_{jk}}$

in both cases Eq.(70) is fulfilled.

The Eq.(71) for the transition probability can be implemented on the computer with the Metropolis algorithm. The system we consider is a fluid composed of N atoms in a volume Ω at temperature T.

Metropolis Algorithm

- calculate the initial potential energy U_1
- choose randomly an atom
- move the atom to a new random position
- calculate the new pot. energy U_2 and $\Delta U = U_2 - U_1$

- calculate $F = exp(-\beta \Delta U)$
- extract a random number $0 < R < 1$
- compare F with R

  if $F > R$ then accept the new configuration 2

  else the system remains in the old configuration 1
- start again the procedure with a new atom

There is a simple physical interpretation of this algorithm that can be inferred with the help of Fig.6.

If $U_2 < U_1 \rightarrow F > 1 \Rightarrow$ the move is accepted as expected since the system goes in a state at lower energy.

If $U_2 > U_1$ the move is not rejected since the temperature is finite and there fluctuations of the energy.

> If the fluctuation is too large respect to $k_B T$ the move has an high probability of being rejected. In this case $|\Delta U| >> k_B T$ and F becomes very small so there is an high probability that $F < R$.
>
> On the contrary if $|\Delta U| \leq k_B T$ there is an high probability that $F > R$.

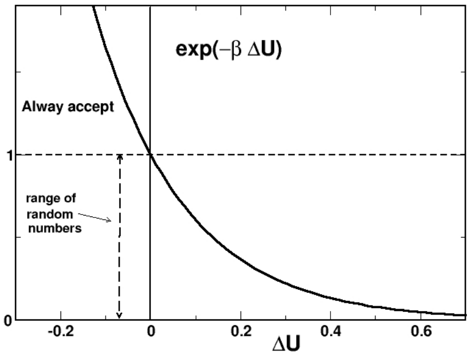


Figure 6: Function for the accepting rule in MC

4.7 Averaging on Monte Carlo steps

To study a fluid of many atoms with MC simulation we can use the procedures described in Sec. 2. As in MD we can start from a lattice at high temperature and after the melting equilibrate the system to a lower T. We move the atoms according to the Metropolis algorithm. if the number of atoms is large, as usual with the modern computers, we can try to move each atoms starting from the number 1 to the number N, the so called typewriter way of proceed. In details we can apply the algorithm in this way

- FOR $i = 1, N$
- extract R_x, R_y, R_z
- shift the atom position:

$$\begin{aligned} x' &= x + \Delta(2R_x - 1) \\ y' &= y + \Delta(2R_y - 1) \\ z' &= z + \Delta(2R_z - 1) \end{aligned} \tag{72}$$

- apply the Metropolis criterion to accept or reject the move

The parameter Δ in (73) is the maximum allowed shift. It is easy to see that if Δ is very small almost all the moves will be accepted, while if Δ is too large almost all the moves will be rejected. Δ is usually adjusted to give an acceptance ratio of 50%.

The cycle on the N atoms is called a Monte Carlo step (MCS) and it is assumed as a conventional unit for the *time* evolution of the system. It has to be noted that when the system does not go in the new configuration it is considered as it makes a transition to the old state. The thermodynamical and the structural quantities can be calculated on averaging on the MCS. We can the use, for example, Eq. 31 for the pressure and Eq. 33 for the radial distribution function. Of course the averages are done now in the canonical ensemble.

4.8 MC sampling in different ensembles

In MC simulation is easy to perform averages in ensembles different from the canonical one. For instance if we want to study the system in the NPT

(isothermal-isobaric) ensemble, we can add to the algorithm another type of move. It consists in changing the volume at constant pressure

$$\Omega' = \Omega + \Delta_\Omega (2R - 1) \tag{73}$$

The Metropolis criterion can be applied now by considering not the potential energy but the difference of the enthalpy after and before the change of volume

$$\Delta H = U\left(\Omega'\right) - U\left(\Omega\right) + p\left(\Omega' - \Omega\right) \tag{74}$$

where p is the pressure that we want to apply to the system.

The system can be also studied in the grand-canonical ensemble. In this case Ω and T are kept fixed and we have to fix also a chemical potential μ. Then we need moves able to add or remove particles from the system. This can be done by adding to the Metropolis algorithm a move of the following type

- choose randomly: remove or add a particle
- for removing: choose randomly a particle and try to remove it
- for adding: choose randomly a position in the box and try to insert the new particle

In the acceptance criterion one has to consider that the quantity to look before and after the removing/adding a particle is the potential

$$U - \mu N \tag{75}$$

We have not time to give the details of the MC algorithm in these ensemble and refer the reader to textbooks.

5 Final considerations

As said at the beginning these lecture notes are intended to give only a short introduction to the vast field of computer simulation, so we did not consider in details some important points.

We had not time for example to consider the methods used to simulate fluids composed of molecules. Molecular systems consist of particles which interact via intra and inter atomic forces. In practice one considers that the intramolecular forces are much stronger than the intermolecular ones and that the vibrational motions are very fast compared with the usual time step of simulation. So it is common to take the system as composed by rigid molecules with fixed geometry. The motion becomes a combination of the translational and rotational dynamics. In MD there are two types of algorithms that are commonly used.

In the method of quaternions the center of mass moves according to the algorithms previously described. The rotation of the atoms of the molecule around the center of mass is realized by considering the equations of motions for the Euler angles between an axis system fixed in space and one fixed with respect to the body. Instead of the Euler angles it has been found more convenient to use generalized coordinates, called quaternions, related to the Euler angles and more easy to handle numerically [1, 2].

A second method is based on the use of dynamics constraints. Each atom of the molecule moves according to the algorithms of the translational motion, but constraints are added to the hamiltonian through Lagrangian multipliers in order to preserve the geometry of the molecule [1, 2, 18].

In MC simulation rotational moves of rigid molecules must be added to the translational ones.

Like in experiments also in MC and MD the quantities evaluated are affected by errors. They are mainly due to the fact that averaging in statistical mechanics is expected to be done on independent configurations. This not rigorously true even in MC because we realize an approximate Markov chain. Moreover the averages are performed on a finite time. Depending on the type of problem and the region of the thermodynamical space we consider, the errors could be more or less relevant. Another source of errors in MC could be related to the random number generators. The routines usually found in modern computers are sometimes still not reliable for very long runs. However accurate methods for determining the errors in equilibrated averages have been developed and are implemented in normal computer simulations.

MC and MD methods can be complementary. MD is the only method to obtain a complete description of the dynamical properties. With enough computer effort it is possible to calculate the time correlation functions and the relaxation to equilibrium also in metastable states like undercooled liquid. MD can be used to study problems of non equilibrium dynamics when the system is under the effect of a perturbation and transport coefficients can be calculated.

MC is sometimes preferred in the study of phase transitions since it is more easy to implement simulations in ensemble like gran canonical where interface effects at coexistence can be avoided. It has to take into account however that the study of phase transitions, in particular of critical phenomena, would require a rigorous estimation of finite size effects, because criticality is determined by the divergence of the correlation length. In finite size systems the correlation length cannot overcome the value of the boxlength. In such cases the use of PBC is not enough to recover the lack of thermodynamical limit.

References

[1] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids.* Clarendon Press, Oxford, 1989.

[2] B. Smit and D. Frenkel. *Understanding Molecular Simulation.* Academic Press, San Diego, second edition, 2001.

[3] J. P. Hansen and I. R. McDonald. *Theory of Simple Liquids.* Academic Press, New York, third edition, 2006.

[4] N. Metropolis and S. Ulam. The monte carlo method. *J. Am. Stat. Ass.*, 44:335, 1949.

[5] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. *J. Chem. Phys.*, **21**, 1087 (1953).

[6] B. J. Alder and T. E. Wainwright. *J. Chem. Phys.*, **27**, 1208 (1957).

[7] A. Rahman. *Phys. Rev.*, **136A**, 405 (1964).

[8] A. Rahman and F. H. Stillinger. *J. Chem. Phys.*, **55**, 3336 (1971).

[9] J. A. Barker and R. O. Watts. *Chem. Phys. Lett.*, **3**, 144 (1969).

[10] G. Grosso and G. Pastori Parravicini. *Solid State Physics.* Academic Press, New York, 2000.

[11] R. Car and M. Parrinello. *Phys. Rev. Lett.*, **55**, 2471 (1985).

[12] D. Remler and P. Madden. *Mol. Phys.*, **70**, 921 (1990).

[13] G. Galli and M. Pasquarello. In M. P. Allen and D. J. Tildesley, editors, *Computer Simulation in Chemical Physics*, page 261. Kluwer, 1993. For an introduction to density functional methodologies see also: [10].

[14] L. Verlet. *Phys. Rev.*, **159**, 98 (1967).

[15] S. Morante, G. Rossi, and M. Testa. *J. Chem. Phys.*, **125**, 034101 (2006).

[16] S. Nosé. *Mol. Phys.*, **52**, 255 (1984).

[17] M. Parrinello and A. Rahman. *Phys. Rev. Lett.*, **45**, 1196 (1980).

[18] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. *J. Comput. Phys.*, **23**, 327 (1977).