Analysis of mathematical methods and principles of molecular dynamics and monte carlo method

Ruiqing Zhang[1]

[1]Jinan Xinhang Experimental Foreign Language School, Jinan, Shandong, China, 250103

rosier214@163.com

Abstract. In this era, molecular dynamics and Monte Carlo methods have become the primary simulation methods. With the emergence of computer simulation research methods, it has been possible to solve the sizeable computational volume and some other problems in the simulation process. At present, molecular dynamics simulations have taken an essential place in the substantial computational system and have a remarkable ability to solve multi-body problems. Therefore, researchers widely use it in many fields such as physics, chemistry, and materials science. Meanwhile, the Monte Carlo method is also a very effective statistical simulation method. This method can far surpass ordinary integration in efficiency with guaranteed computational accuracy. Furthermore, the derived kinetic Monte Carlo method can simulate and study dynamics problems. It can be seen that both simulation methods play an vital role in various disciplines. Therefore, it is very significant to understand the mathematical principles behind them and to know their advantages and disadvantages.

Keywords: molecular dynamics simulation, monte carlo simulation, pair potential and many-body potential, ensemble, Van der Waals force (VDW), kinetic monte carlo simulation.

1. Introduction

One of the challenges throughout the history of physics and chemistry has been to discuss a complete system with multiple molecules interacting. The advent of molecular dynamics simulations has provided a solution to this problem. The application of molecular dynamics can be traced back to 1957 when Alder [1] first applied the method to the problem of solid-liquid phase transitions in hard-sphere systems. Theoretical architectures alone may be able to support only two molecules in a system where they are mutually present. In such a complicated system, to predict the motion of all atoms, it is necessary to apply a considerable number of differential and integral mathematical equations to its analysis. The emergence of molecular dynamics has successfully overcome the problem of solving many-body systems. So with such a large amount of computation, scientists chose to use computers to simulate the motion between molecules. With the rapid development in the field of computers, a crucial numerical prediction method—Computer simulation, is widely used. Its excellent accuracy and independence from external factors, such as ambient temperature, and air pressure, have brought many researchers into the arms of computer simulation.

Computer simulation is applied chiefly to study and dissect a system of atoms containing vast orders of magnitude to predict the next and previous states of the particles. For example, it is possible to determine the coordinates of an atom at the next moment and its acceleration. The current mainstream computational methods contain Monte Carlo simulation methods, first principles computational methods, and molecular dynamics simulations. At the same time, the emergence of molecular dynamics simulations has solved microscopic problems that were previously poorly solved in various scientific disciplines and are commonly used to calculate and estimate numerical values. This paper focuses on the basic principles of molecular dynamics simulation methods and Monte Carlo methods, and analyzes the mathematical techniques they contain. By organizing and summarizing these two methods, this paper will provide some references for future research on simulation methods and provide a theoretical basis for optimizing mathematical methods.

2. Molecular dynamics simulation

Molecular dynamics simulation is a technique to study the motion of molecules on a microscopic scale. Molecular dynamics simulations involve solving the equations of motion for each molecule and calculating its position and velocity at each step. A typical molecular dynamics simulation consists of several steps representing one or more time steps. In this approach, the complete account is taken of the fact that the interactions between molecules, atoms, and functional groups in the system affect the dynamic evolution of the whole system. We can model how the system evolves by following all possible paths the system may take in these times. Finally, the trajectories of the particles are obtained after continuous iterative solving. The main advantage of molecular dynamics simulation is that it allows us to see what happens when we change specific parameters (such as pressure) in the simulation. We can then compare these results with experiments performed on natural systems under similar conditions to gain insight into how these parameters affect the behavior of our model system. However, the shortcoming of molecular dynamics is that the small order of magnitude usually limits the full-time scale it can simulate, typically limited to us or less.

The main simulation flow diagram of molecular dynamics is shown in Figure 1, containing total initial conditions, boundary conditions, potential functions, system synthesis, and the equations of motion.

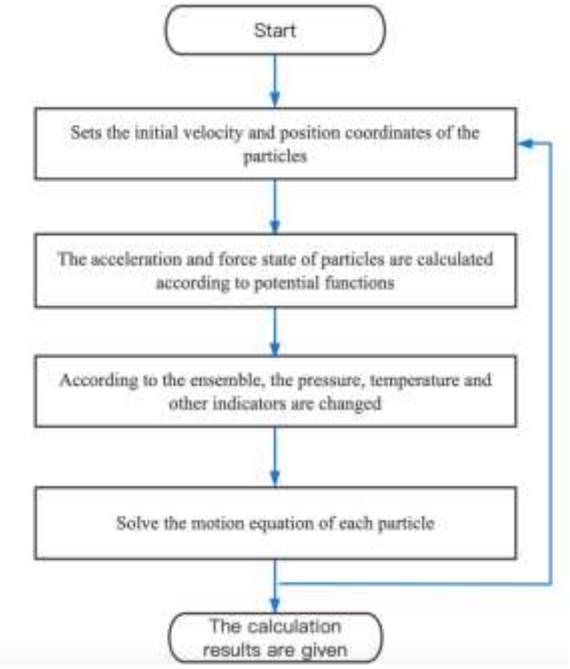

Figure 1. The flowchart of Molecular dynamics simulation.

In terms of principles, molecular dynamics is mainly based on Newton's second law equations of motion to calculate numerical data such as coordinates and accelerations of different atoms in the system. In the system, each atom receives forces from different directions due to interaction potentials between the atoms. Therefore, based on the initial values of the indicators of the object and the interaction potential between the atoms, the forces on each atom in the potential field can be derived, which can be solved by substituting Newton's second law based on the computer's preset parameters of the system environment.

There are two main components of the molecular potential energy, which are the internal potential energy int of the molecule and the Van der Waals force [2] (VDW) between molecules:

$$\mathrm{U} = \mathrm{U_{int}} + \mathrm{U_{VDW}} \tag{1}$$

where the VDW can be considered as the integral of the VDW between each atom in the system:

$$\mathrm{U_{VDW}} = \mathrm{u_1} + \mathrm{u_2} + \mathrm{u_3} + \ldots + \mathrm{u_{10}} + \ldots = \sum_{\mathrm{a=1}}^{\mathrm{n-1}} \sum_{+1}^{\mathrm{n}} \mathrm{u_{ab}}(\mathrm{r_{ab}}) \tag{2}$$

The force on atom a at moment t can be deduced from Newton's second law as follows:

$$F_a(t) = m_a \frac{\partial^2 r(t)_a}{\partial^2 t_a} \tag{3}$$

In the above equation, the coordinates of the particle a are represented as r_a. Since there is no need to consider electron motion in the molecular dynamics system, the motion of so atoms in the system will follow Newton's second law:

$$F_a(t) = -\frac{\partial U}{\partial r_{ab}} \tag{4}$$

In the above equation, m_a expresses the mass of particle a and r_{ab} represents the distance between the interaction of atom a and atom b. According to the equations of mechanics and the relationship between energy, it can be found that the force F_a on atom a at moment t as the first order derivative of the potential energy function U with respect to the coordinates r_a.

The aggregate force on atom a at moment t is $F_a(t)$, the mass of atom a is m_a, and the position vector of atom a is r_a. After a series of calculations using last two formulas above, it can be predicted that the coordinates and velocity of the atom at a fixed moment, and we can in turn deduce the trajectory of an object of any size, which can also be called the trajectory of the atom.

2.1. Inter-atomic interaction potential

An important thing in running molecular dynamics is the potential function. As the name implies, the function used to describe the interactions between atoms in a system is called the potential function. In 1903, a scientist named G. Mie [3] first proposed that the interaction potential between particles is composed of repulsive and attractive interactions between atoms. Since the potential function can represent the laws of motion of atoms, material properties can often be determined as well. And for most elements, the availability of the potential can become difficult because of the need to consider experimental data such as specific mechanical properties parameters and to consider systems consisting of multiple elements combined. Therefore, the importance of the correct selection of the potential function cannot be overstated and usually affects the accuracy and precision of the data results. So far, the study of potential functions has not unified a uniform expression. In the following, we will introduce several common types of potential functions.

2.2. Pair potential and many-body potential

A pair potential means the potential of two atoms relative to each other, so it simply represents the interaction between two particles, independent of the other atoms in the system. Some common pair potentials are Lennard-Jones potential, Morse potential, Born-Mayer potential and Johnson potential, etc. For practical studies, scientists often do not know to consider the interaction between two atoms but rather the many-body interactions in the whole multi-particle system. That is why the many-body potential emerged. Some of the commonly used many-body potentials are Embedded Atom Method potential, Modified Embedded Atom Method potential, Finnis-Sinclair potential, etc. Throughout simulation history, both pair and multi-body potentials have played a prominent role in molecular dynamics simulations.

2.2.1. Lennard-Jones (L-J) potential. The Lennard-Jones potential is probably the most widespread two-body potential function. The expression for the L-J potential [4] was invented by J. E. Lennard-Jones as early as 1924. This potential is often used in the study of the field of noble gases, where it is used as a potential function to describe the interaction between molecules of noble gases. However, the L-J potential may at the same time express only weak inter-atomic forces, such as van der Waals forces, which represent an attraction between two molecules of tiny order of magnitude. The Lennard-Jones (L-J) potential function can be expressed analytically as:

$$\mathrm{U_{ab}(r)} = 4\varepsilon_{\mathrm{ab}}\left[\left(\frac{\sigma_{\mathrm{ab}}}{\mathrm{r}}\right)^{12} - \left(\frac{\sigma_{\mathrm{ab}}}{\mathrm{r}}\right)^{6}\right] \tag{5}$$

In the upper formula, $\mathrm{U_{ab}(r)}$ expresses the magnitude of the potential energy between atom a and atom b, r indicates the distance between the atomic coordinates, and $\varepsilon_{\mathrm{ab}}$ denotes the parameter of the atomic size.

2.2.2. Morse potential. The Morse potential function emerged to solve the problem of vibrations in diatomic molecular systems. In 1924, P. M. Morse proposed the Morse potential function to solve the related quantum mechanical problem of the vibration spectrum of diatomic molecules [5]. The expression of the Morse potential function is as follows:

$$U_{ab}(r) = A\left[e^{-2\alpha(r_{ab}-r_0)} - e^{-\alpha(r_{ab}-r_0)}\right] \tag{6}$$

In the upper equation, A, α and $r_0$are the parameters in formula. The potential function parameters obtained mainly by studying data such as modulus of elasticity and binding energy in the material.

2.2.3. Embedded Atom Method (EAM) Potential. The EAM potential is a potential commonly used to describe metallic systems.The EAM potential treats each atom as an impurity in a whole composed of all other atoms.The whole is an embedded atom potential, which is divided into two main parts: a pair potential for the interaction between two atoms, and a second mosaic potential representing the particle in the context of other local electron clouds[6].The expression for the total potential energy in the EAM is:

$$U_{total} = \sum_{a} F_a(\rho) + \frac{1}{2}\sum_{\mathrm{b}\neq\mathrm{a}} \emptyset_{\mathrm{ab}}(\mathrm{r_{ab}})$$

$$\rho = \textstyle\sum_{\mathrm{b}\neq\mathrm{a}} \mathrm{f_a}(\mathrm{r_{ab}}) \tag{7}$$

In this case, U_{total} is the total potential energy, a and b respectively represent the different atoms, and $F_a(\rho)$ is the embedding energy expression. The formula for the second term is the pair potential term without an explicit expression. The second expression demonstrates the integration of the electron cloud density of the other atoms in the system at atom a as ρ. Where, $\mathrm{r_{ab}}$expresses the coordinate distance between atom a and atom b and $\mathrm{f_a}(\mathrm{r_{ab}})$ represents the background charge density.

2.3. Ensemble

The ensemble is a concept in statistical physics that represents the set of all micro-states under a system. There may be a vast number of micro-states contained in a macro-state, but their overall trend can present the same macro-state, and the ensemble is featured with numerous and independent characteristics. Typically, the ensemble is used in conjunction with the actual environment. The most regular ensembles used in molecular dynamics simulations are the canonical ensemble (NVT), the constant-pressure, constant-temperature ensemble (NPT), and the constant-pressure, constant-enthalpy ensemble (NPH).

2.3.1. Canonical ensemble (NVT). In the canonical ensemble (NVT), the number of atoms in the system, the temperature, and the system's volume are always constant, and the total momentum is zero. This also means that the energy in the system will become energy entangled with the energy of the external environment, and there will be energy transfer between them. To keep the system at a constant temperature, the system should be in thermal equilibrium with the external environment. The temperature within the system can be controlled by adjusting the speed of atoms.

2.3.2. Micro-canonical ensemble (NVE). The micro-canonical ensemble (NVE) is a comprehensive statistical system. In the NVE, the total number of particles and their volume are kept constant. The energy within the complex does not pass between the energy of the external environment, so its energy remains constant. During the simulation, the relationship between the velocity of the atoms in the system and the temperature is given by:

$$\frac{1}{2}mv_a^2 = \frac{3}{2}k_bT \tag{8}$$

In the above expressions, m represents the mass of atom a, v_a denotes the velocity of atom a, T expresses the thermodynamic temperature of the atom, and k_b represents the Boltzmann constant of the atom.
Thus, it can be seen that the kinetic energy in a system with particle number N as:

$$\overline{E_k} = \frac{3}{2}Nk_bT \tag{9}$$

Generally, when manipulating the system, we are required to give the atoms in the system the appropriate atomic coordinates and their velocities, so that the final energy value of the system fits the energy value needed for the simulation.

3. Monte Carlo Simulation

The Monte Carlo model method is a very effective statistical simulation method. The rapid advancement of computer technology has made it possible to use this method to generate numerous random instances of the object under study by enumeration as a way to analyze the laws followed by the object of research and to solve some complex problems.

In essence, the Monte Carlo method uses a computer to simulate random numbers distributed in different locations to identify the patterns of events that occur according to their specific probability distribution. Take the application of Monte Carlo methods to molecular dynamics simulations as an example, and this method can infer the position and velocity of the next moment of a particle that satisfies the physical process of Newtonian mechanics. Its events are calculated based on probabilities. For instance, if object A is in a coordinate system, the probability of its displacement to the right is 0.3, the probability of its displacement to the left is 0.4, the probability of its displacement upward is 0.2, and the probability of its displacement downward is 0.1. The total probability is 1. A process like this, where it is not certain in which direction object A will go at the next moment, is a manifestation of the principle of the Monte Carlo method. Because although the theoretical probability of the object going in the direction of the probability is relatively large, everything is uncertain, with a certain degree of randomness.

Nowadays, researchers often use computers to generate random numbers about different distributions of study objects, and get a large number of data results by simulating a large number of random examples. Afterward, the large amount of data obtained is used to make the final simulation results infinitely close to the expected value [7].

Moreover, the Monte Carlo method also has unique advantages, such as low computational complexity in multi-parameter events and the ability to generate data with randomness [8].

3.1. Kinetic Monte Carlo Simulation

The kinetic Monte Carlo method is an advanced Monte Carlo method. For the Monte Carlo method, it cannot simulate the situation when there are more particles in the system because its sampling sample is relatively small and not enough to support to derive the actual procedure. So scientists improved the Monte Carlo method and invented the Metropolis Monte Carlo method. This method improves on the sampling process of the original Monte Carlo method by making its sampling into a trade-off based on the probability value of the sampled results. This Metropolis sampling method is also the basis for the annealing algorithm [9]. In the Metropolis Monte Carlo method first generates random states and looks at the energy difference between the newly developed state and the initial state in the system to determine the basis for the trade-off. That is, the ability to jump to another state depends on the energy difference between the two beginning and end states.

$$P_{a\to b} = 1 \quad E_a \geq E_b$$

$$P_{a\to b} = \exp\left(-\frac{E_b - E_a}{kT}\right) \quad E_a \geq E_b \tag{10}$$

The two formulas above represent the tendency of the system to move from a high energy state to a low energy state and the system to move from a lower energy state to a higher energy state, respectively.

As with the annealing algorithm, the Metropolis Monte Carlo method will have a certain probability of rejecting the current optimal solution at each step when determining whether a state migration is going to occur. Since the optimal solution at this point may be a locally optimal solution rather than a globally optimal solution, the existence of this probability allows the method to have the probability of going beyond the local to find the global optimal solution. For the study of state changes in molecular dynamics, it is necessary to consider the transition states between states. The kinetic Monte Carlo method successfully introduces the amount of time into the model, taking into account the rate of migration between states and the state energy barriers, allowing this method to be applied to molecular dynamics.

For a single atomic migration, this can be expressed by the following expression：

$$P_{jmp} = v_0 \exp\left(-\frac{E_a}{k_b T}\right) \tag{11}$$

Where P_{jmp} represents the probability of an inter-atomic leap, v_0 is the frequency at which the atom tries to jump, E_a expresses the amount of energy required for the jump, k_b is the Boltzmann constant, and T is the absolute temperature. After calculating the probability of migration and the corresponding activation energy at each step, random numbers are generated at random to simulate the choice of events, i.e., representing the occurrence or non-occurrence of a transition:

$$\Delta t = -\frac{\ln(\text{random})}{P_{total}} \tag{12}$$

In the above equation, Δt denotes the time required for the leap, P_{total} denotes the sum of the probabilities of all possible events of the leap in the current system, and random is the random number generated to generate the probability of a certain rejection of the optimal solution in the range [0,1]. If this random number is constant, then the total probability of the migration is high, and the time required is shorter.

3.2. Process of Kinetic Monte Carlo Simulation

(1) Given the initial configuration;

(2) Calculate the total probability of migration in the studied system;

(3) Randomly generate random numbers and select the migration events to be executed and calculate the specific time required;

(4) Calculate the atomic coordinates and other physical quantities in the new system;

(5) Determine whether the number of simulation steps or the total simulation time potential has reached the desired result. If the simulation is successful, the simulation ends, and the structure is

output. If not, repeat steps 2, 3, 4, and 5 from step 2 until the desired result is reached and output.

4. Conclusion

This paper discusses two mainstream simulation methods, molecular dynamics simulations, and Monte Carlo simulations. With regards to molecular dynamics, the primary simulation process and its principles and mathematical methods are analyzed, exposing how it predicts atomic motion trajectories. Then, the inter-atomic interaction potential and the ensemble are described in detail, and its mathematical formulas are analyzed. Through these, it is easy to see that molecular dynamics simulations are a no-brainer for solving many-body problems. However, the total time scale that can often be simulated by molecular dynamics simulations is usually limited by the small order of magnitude. In Part II, this paper introduces the application of Monte Carlo simulation methods and their derivatives to kinetic Monte Carlo simulations. Although the Monte Carlo method is efficient and stochastic, it is not suitable to be applied to systems with larger systems. In contrast, the improved kinetic Monte Carlo method can be applied to dynamics. Upon detailed analysis of the mathematical theory behind it, it is found that the kinetic Monte Carlo method possesses the probability of being able to jump out of the local optimal solution and then potentially discover the global optimal solution.

References

[1] ALDER B. J., WAINWRIGHT T. E. Phase transition for a hard sphere system [J]. J.Chem.Phys., 1957, 27(5): 1208-1209.
[2] TA Halgren. Journal of the American Chemical Society [J]. ACS Publications, 1992.
[3] MIE G. Zur kinetischen Theorie der einatomigen Korper [J]. Ann. Phys., 1903, 316(8): 657-697.
[4] JONES J. E. On the determination of molecular fields. II. From the equation of state of a gas [J]. Proceedings of the Royal Society of London., 1924, 106(738): 463-477.
[5] MORSE P. M. Diatomic molecules according to the wave mechanics. II. Vibrational Levels [J]. Physical review, 1929, 34(1): 57.
[6] Daw M. S., Baskes M. Embedded-atom method: Derivation and application impurities, surfaces, and other defects in metals [J]. Physical Review B, 1984, 29, 6443-6453.
[7] Raychaudhuri S. Introduction to monte carlo simulation [C]. 2008 Winter simulation conference. IEEE, 2008: 91-100.
[8] Kroese D. P., Brereton T., Taimre T., et al. Why the monte carlo method is so important today [J]. Wiley Interdisciplinary Reviews: Computational Statistics, 2014, 6(6): 386-392.
[9] Aarts E., Korst J., Simulated Annealing and Boltzmann Machines [M]. 1989.
[10] Van Laarhoven, P.J., Aarts E. H., Simulated Annealing. In simulated Annealing: Theory and Application, Springer, 1987, pp.7-15.